Chapter 2 : Seafloor-Consistent Pegleg Multiple Attenuation

2.1 Introduction

The first observations of field data movies made by my colleagues - Ottolini, Sword, and Thorson (Sword, 1981) - revealed coherent motion of anomalous packets of reflected energy. This motion always stands out on any seismic data - land or marine - even over "pancake" geology. In marine "off-end" shooting geometry, the motion always runs from near to far offsets as the movie frames run from earlier to later shot gathers. Furthermore, the packet velocities are correlated with the speed of the boat. These observations can be explained by assigning a laterally varying reflectivity to the geology.

In the previous chapter we saw that there was much to be gained by accounting for the different water column times at shot and geophone locations. In this chapter we take the (empirically justified) view that the texture of the seafloor is also a first order effect. We therefore use a model which allows each part of the seafloor to have its own spectral signature. This allows us to statistically exploit the inherent consistency in the pegleg multiple response from shot to shot. For this reason we term the approach a "seafloor-consistent multiple suppression".

One of the proven techniques of modelling surface effects in reflection seismology is the "surface-consistent statics model" (Taner et al., 1974; Wiggins et al., 1976). Historically this has been used to smooth out trace-to -trace static time shifts, permitting a more coherent stack. Recently, Taner and Coburn (1980) introduced the closely related idea of a surface-consistent frequency response model to handle the statics problem. They modelled each seismic trace, $D(\omega)$, as a separable product of source, geophone, midpoint, and offset responses in the frequency domain. In particular, they assumed that

$$D(\omega) \approx S(\omega)\, G(\omega)\, Y(\omega)\, H(\omega) \qquad (2.1.1)$$

Modelling the seismic trace in this way allows for removal of near-surface absorption/transmission anomalies and facilitates conventional surface-consistent statics corrections.

One might well ask if a similar approach to the multiple suppression problem can be taken. In what sense can one talk about a "seafloor-consistent multiple suppression"?

2.2 The Decomposition Model

We adopt a convolutional model very similar to (2.1.1). The only refinement is the addition of an "average response" factor, A(ω). This last term depends mainly on the average shot waveform - something we will *not* attempt to deconvolve here. Our goal is to attenuate the water column reverberations (leaving as much primary as possible intact!). The model, therefore, is:

$$D(s,g,\omega) \approx S(s,\omega)\, G(g,\omega)\, Y(y,\omega)\, H(h,\omega)\, A(\omega) \qquad (2.2.1)$$

Intuitively, we expect S to contain shot ghost responses, water reverberation effects characteristic of shot location and residual shot waveforms. Receiver ghosts as well as water reverberations under the hydrophone locations are embedded in G. Any anomalous response that is characteristic of midpoint is handled by the Y factor. H is included to model variation of the seafloor reflection coefficient with offset.

Taking logarithms on both sides of equation (2.2.1) gives:

$$\ln D_{sg} \approx \ln S_s + \ln G_g + \ln Y_{\frac{s+g}{2}} + \ln H_{\frac{s-g}{2}} + \ln A \qquad (2.2.2)$$

where s denotes shot index and g, receiver index. Equation (2.2.2) holds independently for each frequency and is a complex-algebraic equation. The real part contains log-amplitude information and the imaginary part defines the phase.

Ideally, we would like to get imaginary as well as real solutions to equation (2.2.2). This was attempted but frustrated by the issue of phase unwrapping. The "phase unwrapping" problem (Tribolet, 1979) is encountered whenever complex logarithm processing of seismic data is attempted. There are ambiguities of 2π in a discretely sampled phase spectrum when the slope of the phase curve becomes large. This is commonly the case in any band where the signal to noise ratio is poor, such as the very low frequency band.

For deep water - deep enough to allow discrimination between the bubble pulse and the seafloor reverberation response - it turns out that a considerable amount can be done with only a knowledge of the amplitude responses in equation (2.2.2). There are a number of a-priori constraints we can place on the phase. In particular, the shot and receiver reverberation effects are known to be causal and have time-limited inverses. As we shall see in Section 2.7, Wiener-Levinson spectral whitening techniques seem to correctly handle phase estimation.

2.3 Reduction of Free Parameter Count

The standard predictive method of multiple suppression involves estimating a Weiner-Levinson (W-L) inverse filter for each seismic trace. The W-L filter is a stationary model fit to the data. If there are N_s shotpoints, N_h offsets and N_t adjustable W-L filter points for each trace, the total number of free parameters describing the multiples is $N_s N_h N_t$.

In all model fitting it is desirable to describe the phenomenon of interest with as few free parameters as possible. The separable model of equation (2.1.1) gives a very parsimonious description of the seismic dataset. It reduces the number of independent spatial parameters from a total of $N_s N_h N_t$ (as in standard trace-by-trace predictive deconvolution) to $(N_s+N_g+N_y+N_h)N_t$. In most cases this amounts to a very significant reduction. If there are fewer free parameters involved in the multiple suppression then we are less likely to attenuate the primaries while suppressing multiples. This improves the post-stack primary to multiple ratio.

The problem of estimating least square solutions for the logarithms of S,G,Y, and H in equation (2.2.2) is completely analogous to the classic residual statics estimation problem. The problem is overdetermined (since there are $N_s N_h$ equations for only $N_s+N_g+N_y+N_h$ unknowns) but is known to be underconstrained in the long wavelength components of the solution. This second aspect of the problem will become apparent when we examine real data solutions.

The complexity of this problem is increased by a factor of **nf** over conventional residual statics, where **nf** is the number of frequency planes over which we wish to do the decomposition. The method of solution follows.

2.4 Least Square Model Decomposition

After extracting the average magnitude, $\ln A(\omega)$, from the s-g plane, our problem is to find the min $E(S,G,Y,H)$, where:

$$E = \sum_s \sum_g (D_{sg} - S_s - G_g - Y_{\frac{s+g}{2}} - H_{\frac{s-g}{2}})^2 \qquad (2.4.1)$$

In equation (2.4.1), D,S,G,Y, and H are redefined to be the log-magnitudes of their values in equation (2.2.2). The s,g subscripts range over the appropriate regions of the s,g plane. The minimum coincides with the vanishing of the partial derivatives of E with respect to S_k, G_l, Y_p, and H_q. Setting, for example,

$$\frac{\partial E}{\partial S_k} = 0$$

yields

$$\sum_g [D_{kg} - S_k - G_g - Y_{\frac{k+g}{2}} - H_{\frac{k-g}{2}}] = 0 \qquad (2.4.2)$$

or

$$S_k = \frac{1}{N_g} \sum_g (D_{kg} - G_g - Y_{\frac{k+g}{2}} - H_{\frac{k-g}{2}}) \qquad (2.4.3a)$$

Similarly,

$$G_l = \frac{1}{N_s} \sum_s (D_{sl} - S_s - Y_{\frac{s+l}{2}} - H_{\frac{s-l}{2}}) \qquad (2.4.3b)$$

Making the index transformation

$$m = \frac{s+g}{2} \quad ; \quad n = \frac{s-g}{2}$$

we can rewrite equation (2.4.1) as:

$$E = \sum_m \sum_n (D_{m+n,m-n} - S_{m+n} - G_{m-n} - Y_m - H_n)^2 \qquad (2.4.4)$$

Zeroing $\partial E / \partial Y_p$ and $\partial E / \partial H_q$ in (2.4.4) gives:

$$Y_p = \frac{1}{N_h} \sum_n (D_{p+n,p-n} - S_{p+n} - G_{p-n} - H_n) \qquad (2.4.3c)$$

$$H_q = \frac{1}{N_y} \sum_m (D_{m+q,m-q} - S_{m+q} - G_{m-q} - Y_m) \qquad (2.4.3d)$$

Equations (2.4.3a-d) tell us what we may have intuitively expected; the L_2 norm solutions to the decomposition problem can be obtained by averaging model residuals over the direction orthogonal to the response component of interest. It is worth noting in passing that L_1 solutions can be obtained by replacing the averaging operators in equations (2.4.3) with median operators.

The solution strategy involves cycling through equations (2.4.3) for all desired values of k,l,p and q until convergence at the spatial wavelengths of interest is obtained. (This turns out to be equivalent to solving the implied system of equations by "Gauss-Seidel" iteration). The outermost loop of the algorithm is over frequency.

An alternative approach to this problem involves working in the midpoint wavenumber domain rather than the space domain (Appendix 2.A). This is similar to Marcoux's (1981) approach to the statics problem. In the statics case, the problem decouples in the

wavenumber domain, allowing an explicit solution. Although our problem does not fully decouple with this approach, it does make the computations better organized.

2.5 Decomposition of Flemish Cap Data

To test the concepts presented to this point in Chapter 2, we return to the Flemish Cap dataset - partially displayed in Figures 1.2 and 1.3. The data is both 48 trace and 48 fold with 25 metre shot spacing and 50 metre group spacing. The near offset is 273 metres - the equivalent of 11 shot points.

A 2 to 4 second window of two successive common shot gathers is displayed in Figure 2.5.1. It is more difficult to pick out the split pegleg multiples on gathers than on common offset sections. The arrow just above 3 seconds points to the near offset node of a split pegleg event. The steeply dipping events on the far offsets are water bottom multiples.

The algorithm of Section 2.4 was run on the inner 24 offsets of this line. The outer 24 offsets were excluded since they contained water bottom multiples which violate the vertical reverberation path assumption of the model of equation (2.2.1). Figures 2.5.2 show the logarithms of the amplitude responses of the various model components. They are plotted as functions of frequency (fraction of Nyquist) and in-line coordinate.

The average response, A, is much smoother and of higher amplitude than the residual responses. The offset response shows a trend of diminishing amplitude towards larger offset. It also has some colour but this differs markedly from the colour of the other three responses, S,G and Y. The main feature of the S and G plots is a trend towards higher "quefrencies" from left to right. This is due to the increased water depth or reverberation time at the right end of the section (see Figure 1.2). The S and G plots have much more lateral continuity and energy than the Y plot. They appear to be more sensitive to the seafloor than the midpoint response. This observation is reinforced by the fact that the best visual correlation between the S and G plots occurs when their two right edges are offset by 11 shot points (exactly the distance of the near offset). The useful bandwidth of the data appears to run from 5% to 50% of the Nyquist frequency. There is no horizontally coherent pattern in any of the residual model responses outside this passband.

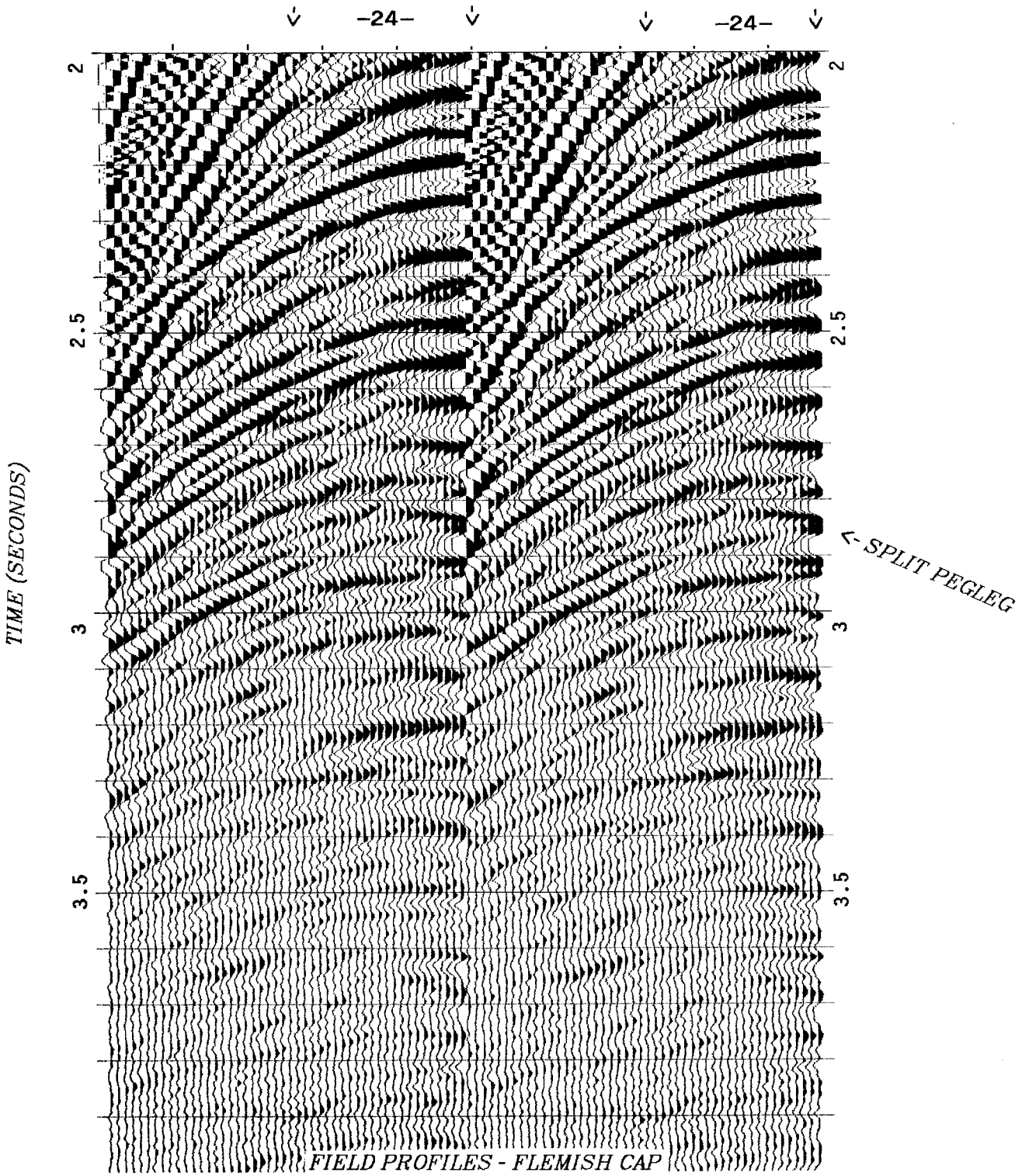


FIG. (2.5.1). Two common shot gathers from the Flemish Cap line shown in Figures 1.2 and 1.3. The arrow just above 3 seconds points to a split pegleg multiple. Only the 24 nearest offsets on each profile were used in the decomposition.

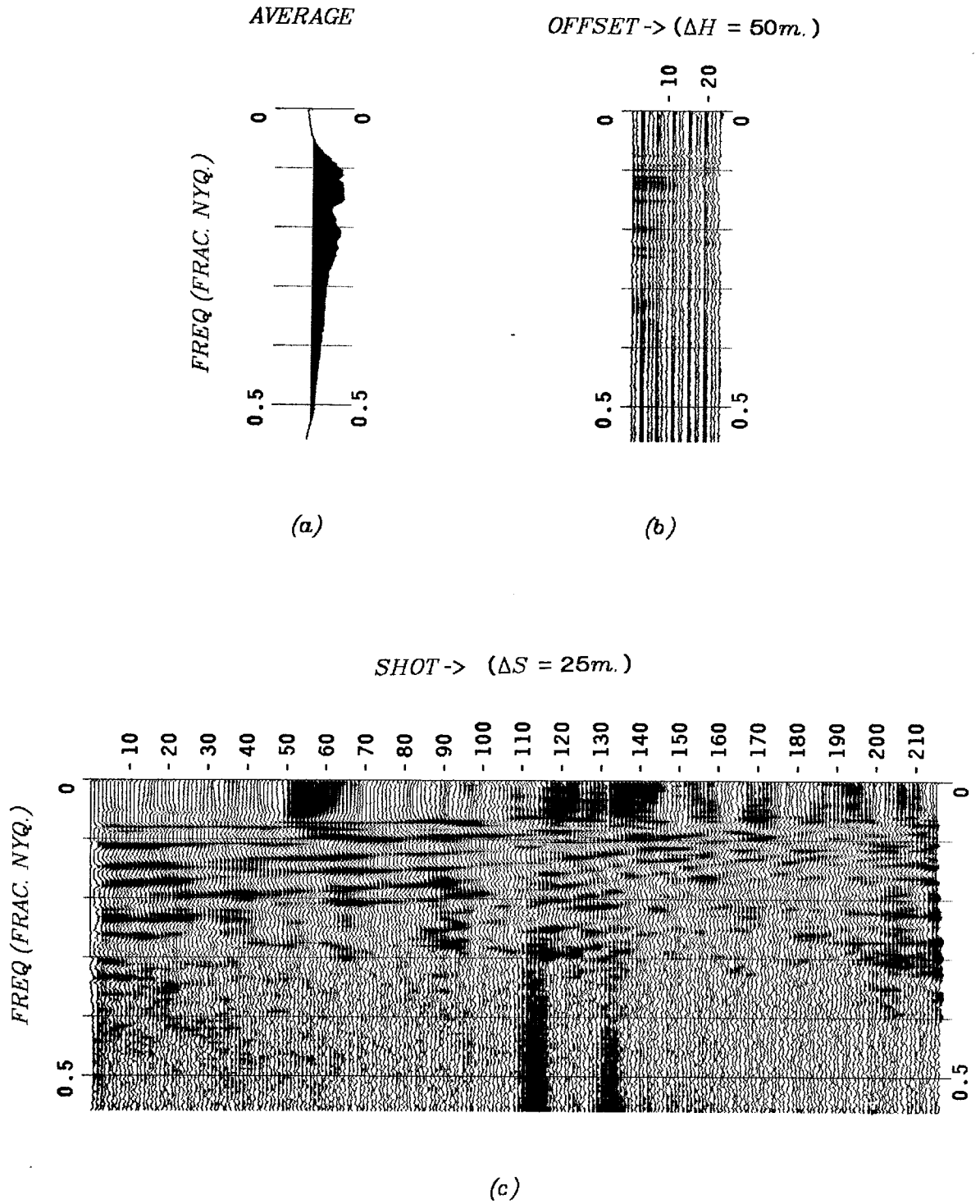
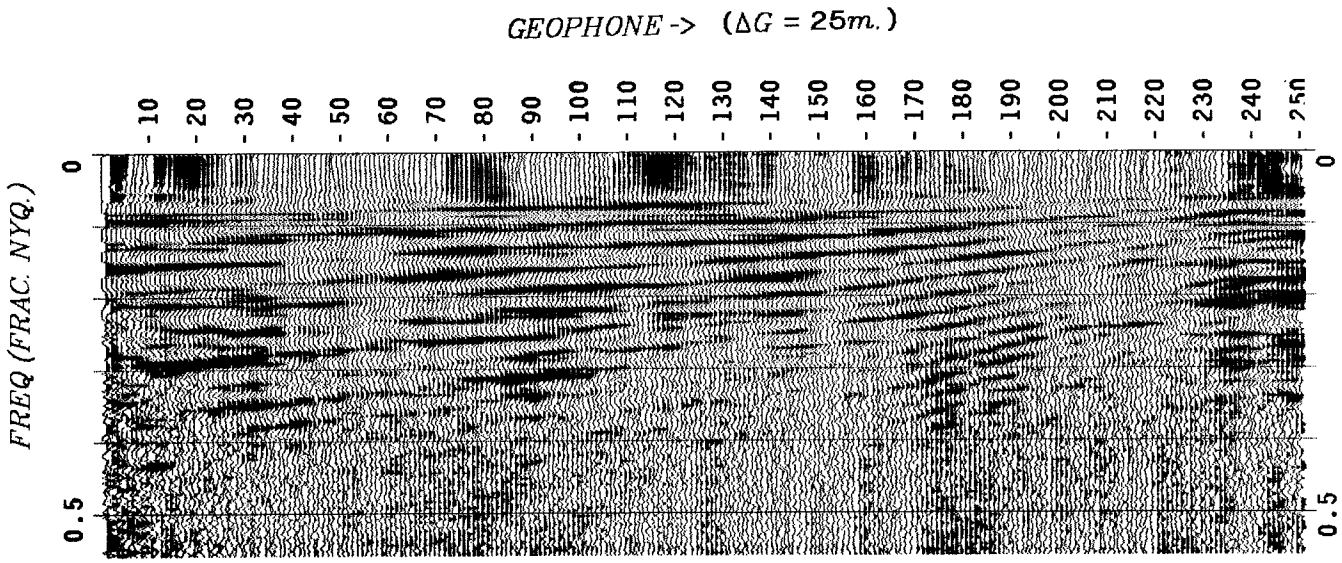
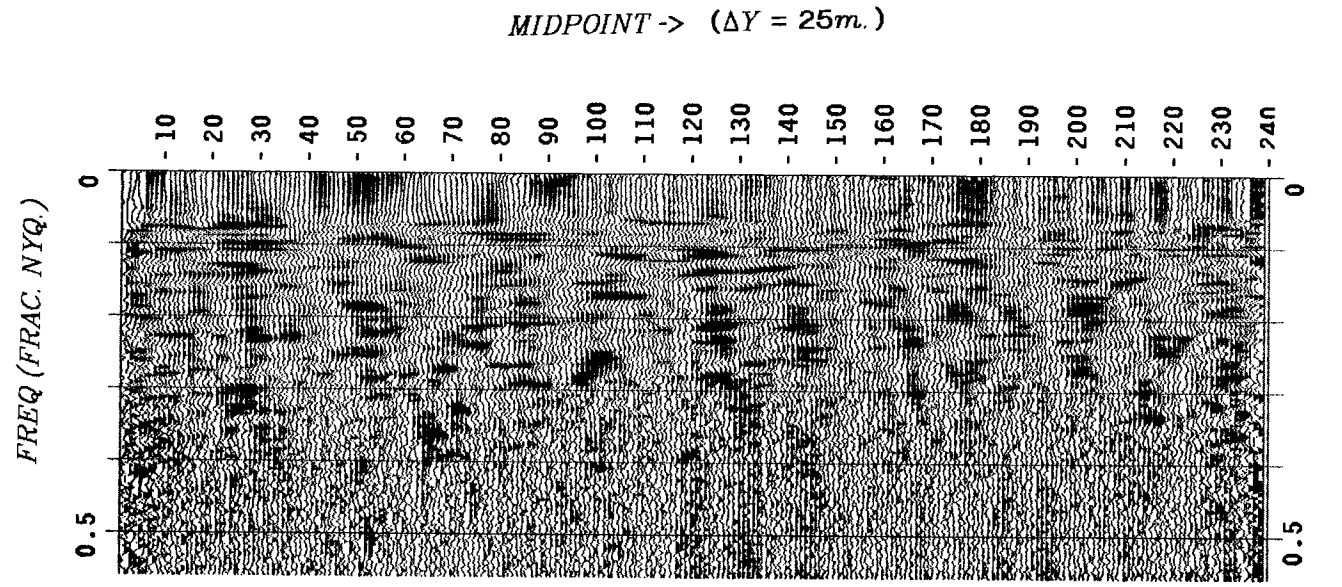


FIG. (2.5.2a-c). Logarithm plots of (a) average, (b) offset, and (c) shot responses are displayed clockwise from top left. Plots (b) and (c) are displayed after 4 iterations of equations 2.4.3. Maximum value of (a) is about 6 times the clip level of (b) and (c).



(d)



(e)

FIG. (2.5.2d-e). Logarithm plots of (d) geophone, and (e) midpoint amplitude responses are displayed top to bottom. Clip level is same as in plots (b) and (c) of previous page.

2.6 Convergence Properties

Since the model decomposition problem as posed in Section 2.4 is underconstrained in the very long wavelength solutions, such solutions are only conditionally stable. Although low wavelength instability was never observed to be a problem in practice, the potential for it is always there. It is possible, for example, for a large, positive D.C. solution in S to be cancelled by an equally large but negative D.C. component in G without changing the value of E in equation (2.4.1). A program implementing equations 2.4.3 did cause arithmetic overflow when allowed to iterate a large number of times overnight.

The algorithm can be "bulletproofed" by redefining the error functional to be

$$E' = E + \lambda \left(\sum_s S_s^2 + \sum_g G_g^2 + \sum_p Y_p^2 + \sum_q H_q^2\right) \qquad (2.6.1)$$

with $\lambda > 0$. This change adds a term of λ to each of the denominators of equations (2.4.3).

The plots of Figure (2.5.2) were made after four iterations of equations (2.4.3) in the order "d,a,b,c". In order to get a handle on the errors involved in this solution we iterated four times again from the same initial data but in the order "b,a,c,d". The difference between the two final S amplitude responses is displayed in Figure 2.6.1 at the same clip level as the plots of Figure 2.5.2. This plot shows that the main errors are confined to long wavelengths The corresponding difference plots for the other model components (not displayed) support this contention. We can therefore be confident that the decomposition is unique at wavelengths less than a cable length or, in this case, about fifty traces.

Figure 2.6.2 gives some idea of the speed of convergence of equations (2.4.3). In this figure the log-amplitude versus frequency of the traces of a particular common shot gather (S.P. #100) is displayed after odd-numbered iterations of the algorithm of Section 2.4. The displayed quantity is

$$D_{kg} - G_g - Y_{\frac{k+g}{2}} - H_{\frac{k-g}{2}} \qquad (2.6.2)$$

for one fixed shotpoint, k. The display changes with each iteration as G,Y and H are updated. Convergence is obtained in about four iterations. The current estimate of the S response - obtained by averaging over the g coordinate - is displayed to the right of each gather.

An obvious difference between the first and third iterations is the power level on the near (leftmost) offsets. The slanted events, which stack out in the sum trace, represent the water column reverberations at the geophone locations. They diverge with increasing offset, since for a fixed shot point the water depth increases with offset.

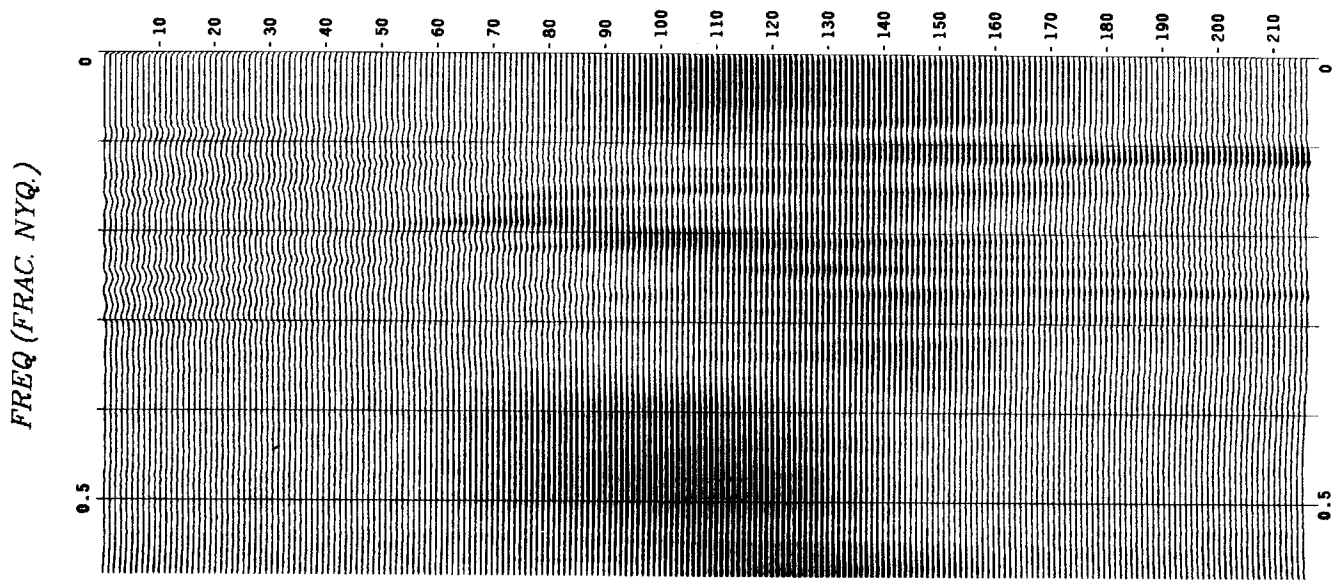


FIG. (2.6.1). Estimate of the error in shot response. Most of the error energy is in the long spatial wavelength components, i.e., wavelengths greater than a cable length.

2.7 Suppressing the Multiples

The plots of Section 2.5 represent the anomalous spectra - characteristic of shot, geophone, and midpoint locations. Our goal is to remove anomalous colour and power associated with seafloor location. The algorithm used for this is shown in Figure 2.7.1.

We start with a database of log-amplitude spectra, S_s and G_g, obtained from the seismic data by equations 2.4.3. For each trace, d_{sg}, we form the sum $S_s + G_g$, multiply by 2, and exponentiate to obtain a model power spectrum. It is essential to pay careful attention to scale factors in the definition of S and G up to the exponentiation stage since these have a nonlinear effect in further processing. Factors of 2 and $\log_e 10$ are most likely to cause trouble. Having obtained a power spectrum, we do an inverse Fourier transform to obtain an autocorrelation. We then use a standard predictive deconvolution (Weiner-Levinson) algorithm to obtain a causal inverse filter, a_{sg}, to the anomalous power spectrum. The multiple-suppressed trace is the convolution of a_{sg} with d_{sg}.

An important issue in the design of a_{sg} is the selection of prediction distance and filter length. For the Flemish Cap data the prediction distance in the W-L procedure was chosen to be slightly less than the minimum seafloor time. This was done so as to leave the phase of

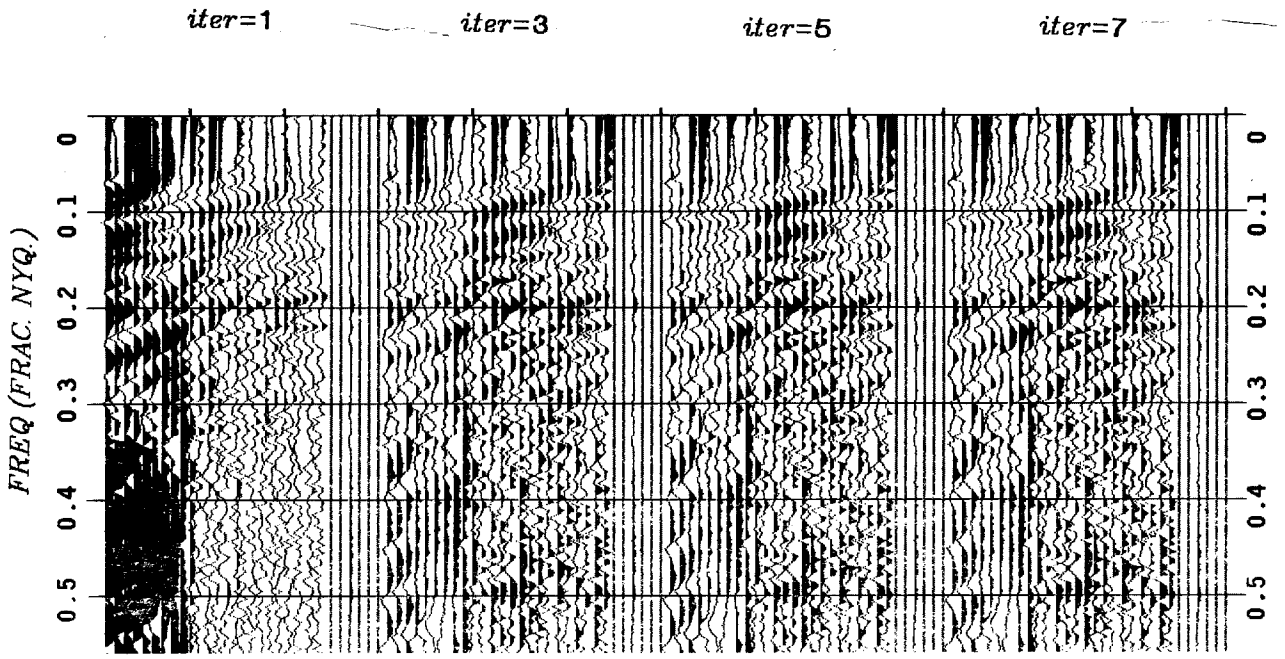


FIG. (2.6.2). Log-amplitude plots of a common shot gather after correction for geophone, midpoint, and offset effects. The quantity in expression (2.6.2) is displayed after odd-numbered iterations of equations (2.4.3). These traces are summed to give the S estimate for this particular shotpoint.

the bubble pulse unchanged. The filter length was chosen just long enough to include the maximum seafloor time across the section. The choice of prediction distance may become critical in shallow water - a case yet to be tested with real data. This point is further discussed in Section 9 of this chapter.

The next three figures demonstrate the results of applying the above algorithm to the Flemish Cap data of Chapter 1. Figure 2.7.2 is a 2400% brute CDP stack of this data. Figure 2.7.3 shows the result of the processing sequence of Figure 2.7.1. The primary near 3 seconds now stands out very clearly and the multiples are virtually eliminated. A weaker primary - some 200 msec. below the strongest primary is also now visible. The bubble pulse remains in the data, since it is contained in the average response which is not backed out.

For comparison, Figure 2.7.4 shows a stack of the data after standard predictive deconvolution processing. A gap of 200 msec. was used with an operator of 120 msec. Although this conventional processing was successful in removing the second order pegleg from the final result, it failed to do a good job on the first order pegleg.

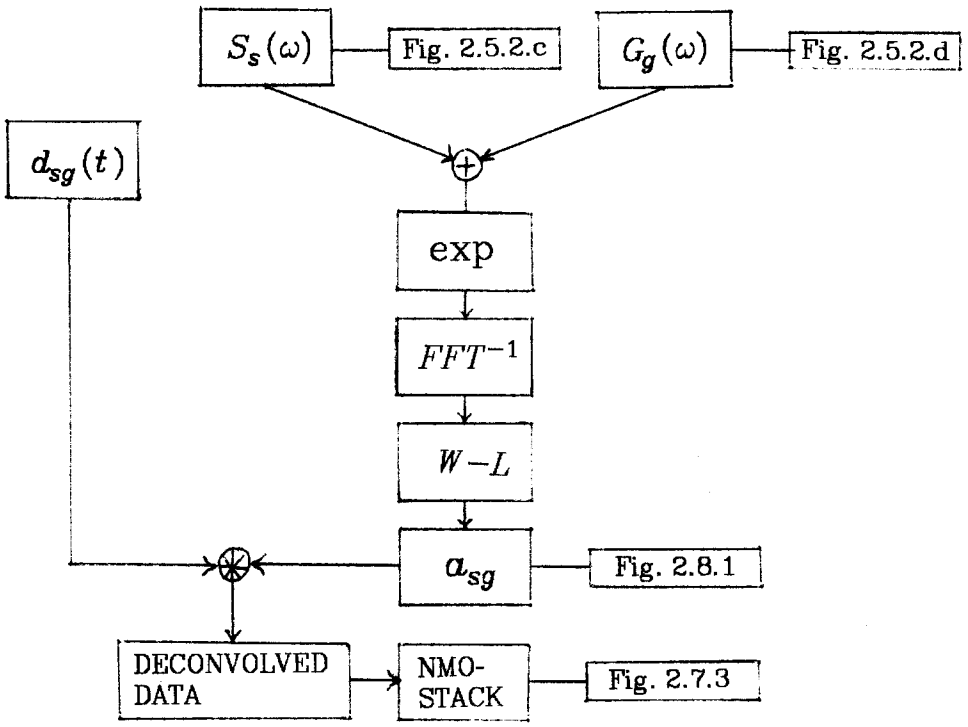


FIG. (2.7.1). Procedure for processing the seismic trace, d_{sg}, recorded at s'th shotpoint and g'th geophone location.

Figures 2.7.2 through 2.7.4 show a definite continuum of quality from a worst case situation (2.7.2), to a better one (2.7.4), to a best result (2.7.3).

2.8 Movies of Inverse Operators

A better understanding of the methods of this chapter can be obtained from movies of the inverse operator, α_{sg} (Fig. 2.8.1). In each frame of Figure 2.8.1 the vertical axis is time and the horizontal axis is shotpoint location. Each frame is displayed at a constant but successively larger offset and may be described by the formula

$$Frame_{g-s}(s,t) = Levinson\ [\ FFT^{-1}(e^{2[S_s(\omega)+G_g(\omega)]})] \qquad (2.8.1)$$

which summarizes the right half of Figure 2.7.1. It is difficult to appreciate all of the

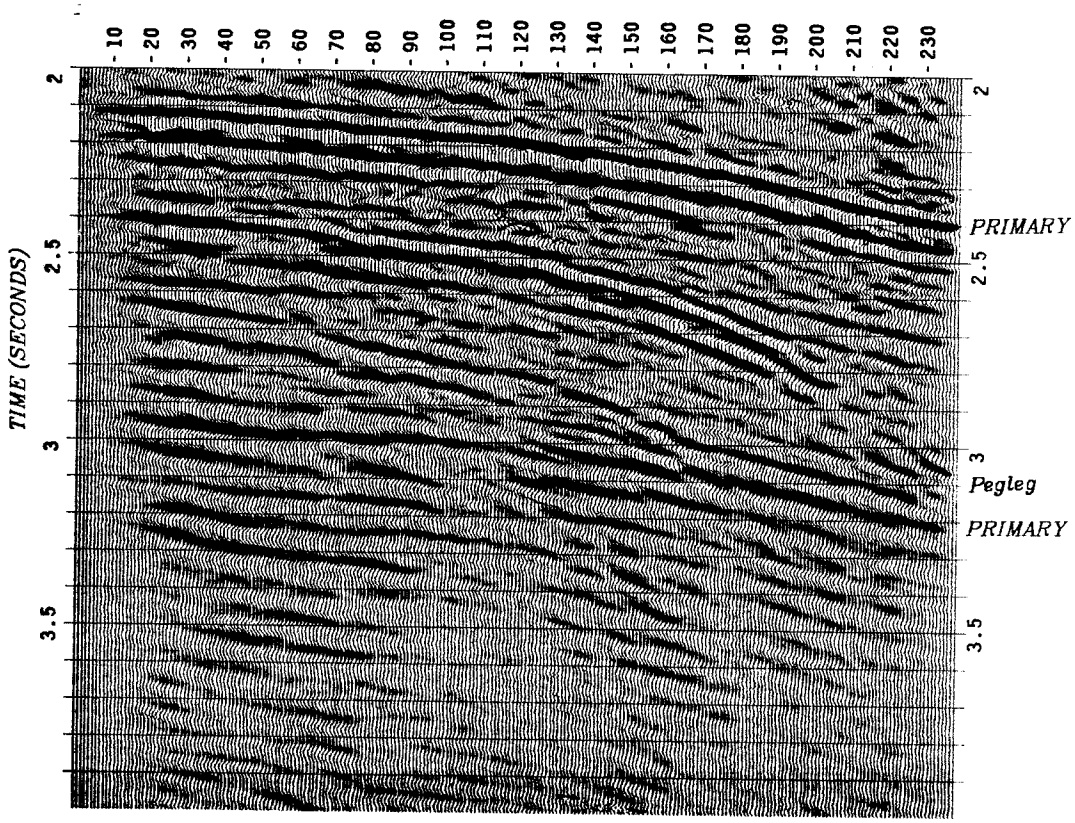


FIG. (2.7.2). Flemish Cap 2400% brute CDP stack.

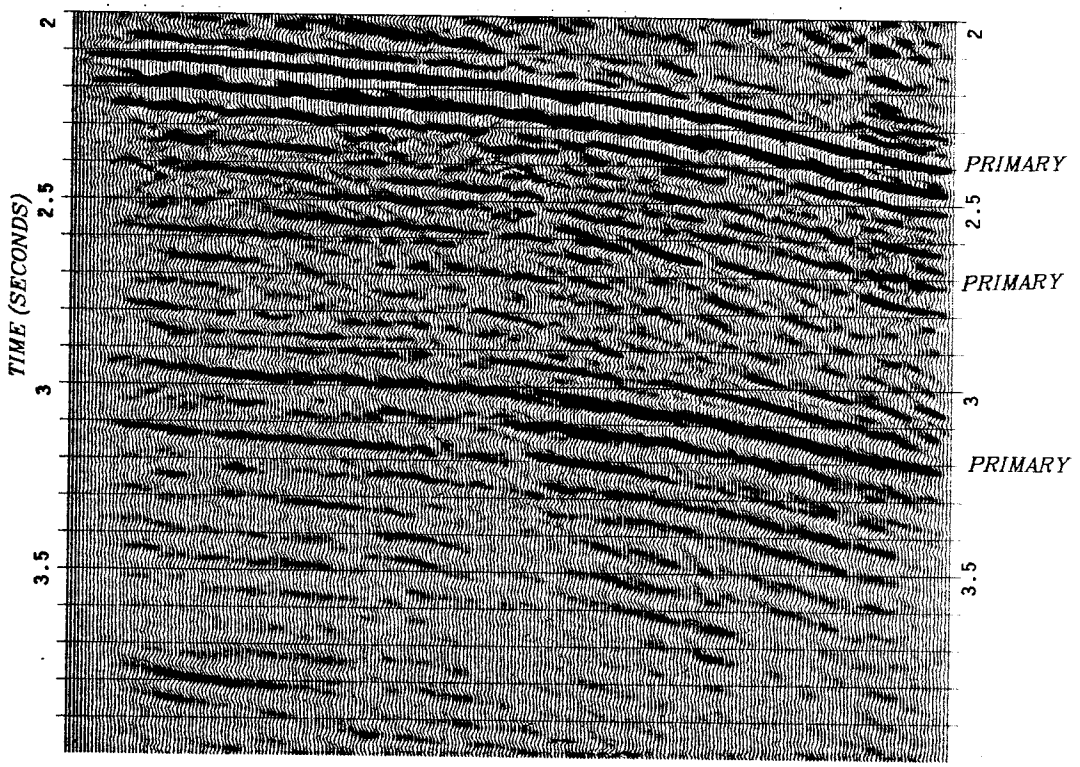


FIG. (2.7.3). Flemish Cap - CDP stack of processed data.

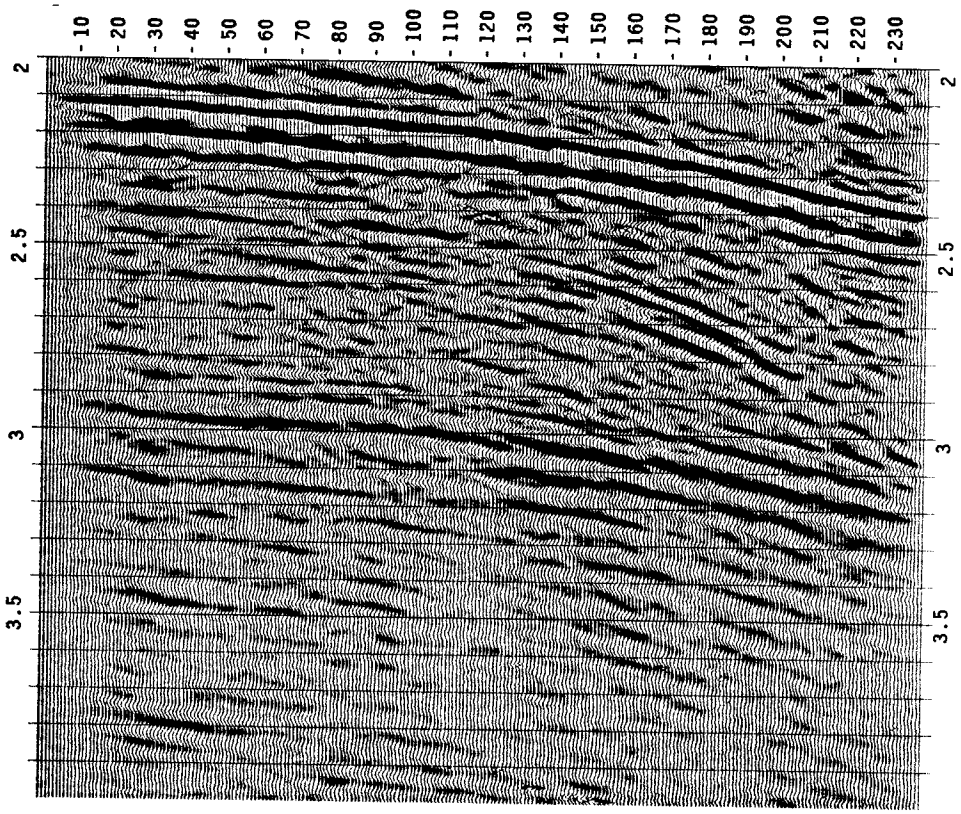


FIG. (2.7.4). "Standard Multiple Suppression Processing". Predictive decon was run before stack with a 200 msec. gap and 120 msec. filter. The second order pegleg multiple was attenuated but the first order pegleg remains quite strong.

subtleties in the movie from a series of still frames; nevertheless, the main phenomena can still be seen.

In viewing the movie, the eye can pick out two distinct branches in the total image. One branch is motionless and is associated with the shotpoint reverberations. (The shots are fixed in this frame of reference). The other branch appears to slide over top of the background and is associated with the geophone reverberations. The overall moving image looks like the pincers of a crab opening and closing as offset increases and decreases. The moving part of the pincer is the upper part since the geophone locations move into shallower water as offset increases.

It is not significant that there is motion in this movie since the underlying mathematical construction calls for two images moving with respect to each other. Even if random

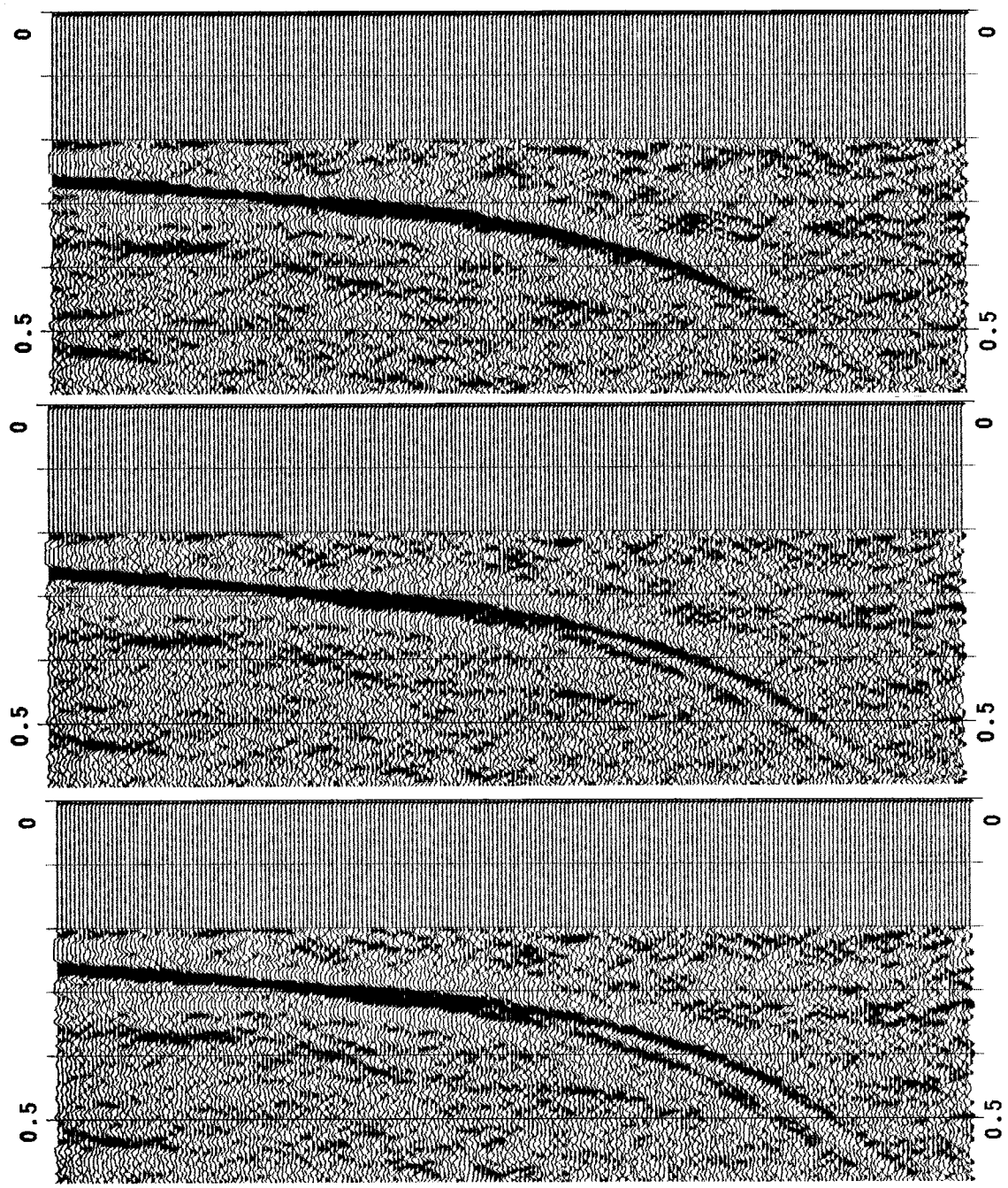


FIG. (2.8.1). Frames from a movie of inverse operators for the Flemish Cap data. Horizontal axis is shotpoint location. Vertical axis is time from 0 to 0.5 secs. Each frame is plotted at a constant but successively larger offset.

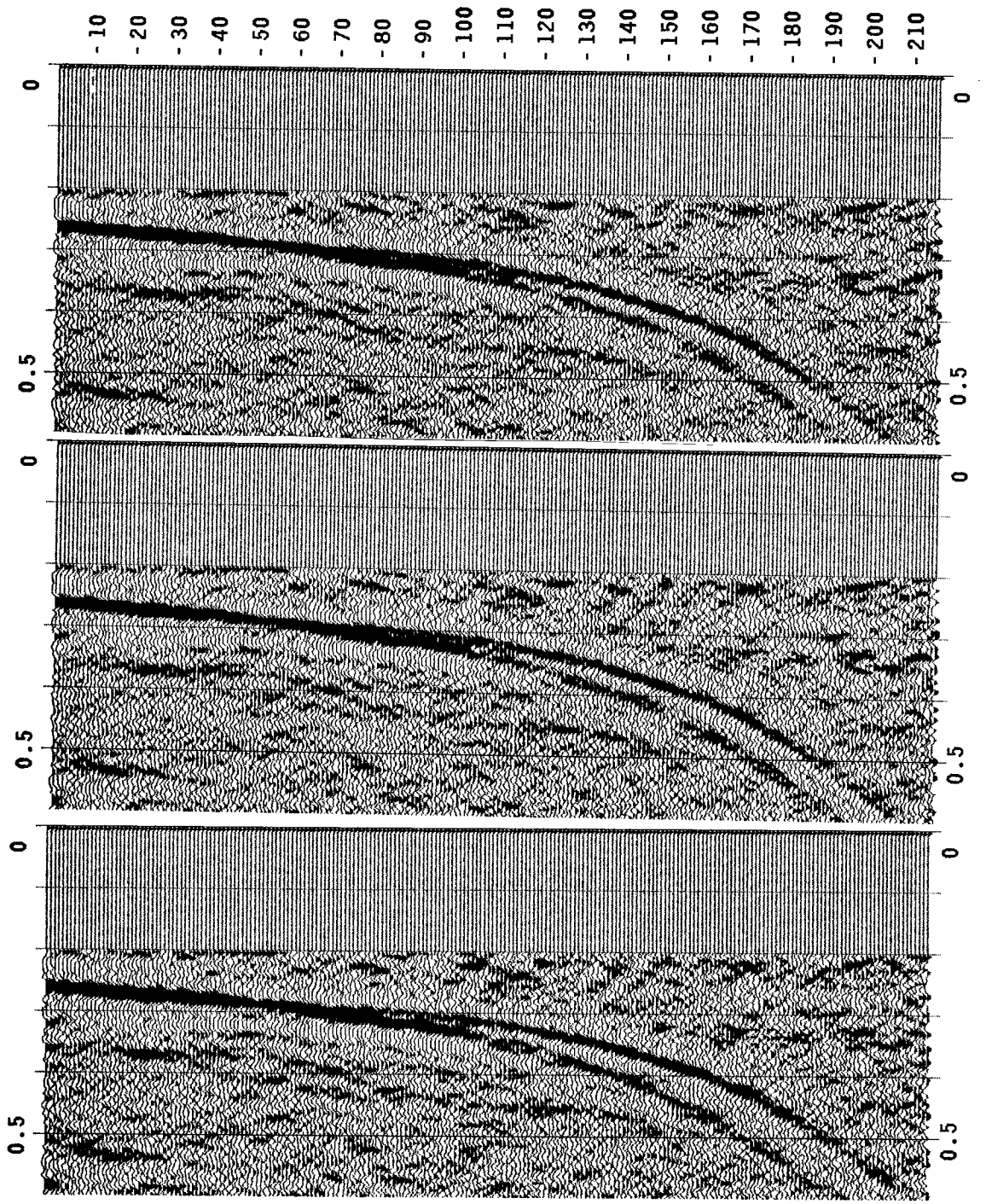


FIG. (2.8.1-continued). As offset increases, the difference between shot and geophone reverberation times increases.

numbers were substituted for S_s and G_g this movie would still give the effect of motion. What *is* significant, however, is the spatial coherence of the two seafloor branches and the fact that they coincide at zero offset. These results are so convincing that we no longer need to restrict ourselves to the best available data.

2.9 The Shallow Water Problem

The Flemish Cap data is an example of a deep water multiple problem. By deep water, we mean that the decay time of the bubble pulse, τ_b, is less than the minimum water column reverberation time, τ_s. (i.e. $\tau_b < \tau_s$). A great deal of seismic data is recorded over shallow bottom areas such as the Gulf Coast, where $\tau_b > \tau_s$. Here, also, the multiples are usually more subtle than those of the Flemish Cap since the bottom is not as hard. Is this method of any use in these areas? We believe that the answer is yes - but only if the bubble pulse is constant from trace to trace.

In general, the anomalous shot response, S, will be composed of two factors: an anomalous bubble response, S_b, and an anomalous water column reverberation response, S_w, at the shot location. In the same spirit, we can model the anomalous geophone response , G, as a product of G_b and G_w. If the survey is carefully controlled, then both S_b and G_b should be small with respect to A. The remaining factors, S_w and G_w, are both minimum phase on physical grounds. We are therefore justified in backing them out with Wiener-Levinson techniques. If this is done, we will be left with a dataset whose colour cannot be attributed to any seafloor-consistent response. This should aid further deconvolution by leaving an apparent bubble which is more consistent from trace to trace.

Conclusions

This chapter has presented a method for statistically suppressing the water column reverberations associated with each shot and geophone location under the assumption of near-vertical incidence propagation in the water. Water bottom multiples which do not obey this assumption must be eliminated by pre-processing. Due to the method's similarity to the surface-consistent statics problem we call it a "seafloor-consistent pegleg attenuation". Since the number of free model parameters is greatly reduced in comparison with trace-by-trace predictive deconvolution, primary events are not severely attacked. The method is illustrated by application to a deep water section with extremely encouraging results. The technique may also be useful with shallow marine data, although further development is needed to confirm this.

REFERENCES

Marcoux, M.O. (1981), On the Resolution of Statics, Structure, and Residual Normal Moveout, Geophysics, v.46, No.7, pp. 984-993.

Sword, C. (1981), SEP Goes to The Movies, SEP report no. 28, pp.11-20.

Taner, M.T, and Coburn, K.W. (1980), Surface Consistent Estimation of Source and Receiver Response Functions, paper G-43, 1980 SEG meeting, Houston.

Taner, M.T., Koehler, F., and Alhilali, K.A. (1974), Estimation and Correction of Near - Surface Time Anomalies, Geophysics, v.39, pp 441-463.

Tribolet, J.M. (1979), Seismic Applications of Homomorphic Signal Processing: Prentice-Hall (Signal Processing Series).

Wiggins, R.A., Larner, K.L., and Wisecup, R.D. (1976), Residual Statics Analysis as a General Linear Inverse Problem, Geophysics, v.41, no.5, pp 922-938.

Appendix 2.A : Model Decomposition in Midpoint Wavenumber Space

The advantage of working in midpoint wavenumber (k_y) space is that the S,G and Y solution components decouple for each k_y. Using the same symbols as in Section 2.4 we write the residual model error (equation 2.4.4) as

$$E = \sum_h \sum_y [D(y,h) - S(y+h) - G(y-h) - Y(y) - H(h)]^2 \qquad (2.A.1)$$

Fourier transforming over the midpoint axis and using Parseval's theorem gives

$$E = \sum_h \sum_{k_y} || \tilde{D}(k_y,h) - e^{ik_y h}\tilde{S}(k_y) - e^{-ik_y h}\tilde{G}(k_y) - \tilde{Y}(k_y) - H(h) ||^2 \qquad (2.A.2)$$

Now, for $each\ k_y$ we set

$$\frac{\partial E(k_y)}{\partial \tilde{S}^*(k_y)} = 0 \qquad (2.A.3)$$

giving

$$\sum_h e^{-ik_y h}(\tilde{D} - e^{ik_y h}\tilde{S} - e^{-ik_y h}\tilde{G} - \tilde{Y} - H(h)) = 0 \qquad (2.A.4)$$

or

$$\tilde{S} = \frac{1}{N_h}\sum_h [(\tilde{D} - \tilde{Y} - H(h))e^{-ik_y h} - \tilde{G}e^{-2ik_y h}] \qquad (2.A.5)$$

Similarly,

$$\tilde{G} = \frac{1}{N_h}\sum_h [(\tilde{D} - \tilde{Y} - H(h))e^{ik_y h} - \tilde{S}e^{2ik_y h}] \qquad (2.A.6)$$

$$\tilde{Y} = \frac{1}{N_h}\sum_{h}(\tilde{D} - e^{ik_y h}\tilde{S} - e^{-ik_y h}\tilde{G} - H(h)) \qquad (2.A.7)$$

and

$$H(h_l) = \frac{1}{N_y}\sum_{k_y} \text{Re}\,[\tilde{D}(k_y,h_l) - e^{ik_y h}\tilde{S} - e^{-ik_y h}\tilde{G} - \tilde{Y}]\ ;\ 1 \le l \le N_h \qquad (2.A.8)$$

Note that the unknowns $\tilde{S},\tilde{G},\tilde{Y}$ are now (complex) scalars instead of vectors, as they were in Section 2.4.