VIII. NEWTON-TYPE ALGORITHM

In the preceding chapter an algorithm was developed to maximize $U(X,\alpha_1,2)$ where $\alpha_1 > 2$. We believe the algorithm is optimum for extracting signals in noise which is on the uniform side of the Gaussian. These algorithms concentrate on pushing up events above the noise level. Tests on synthetics indicate it does not try to suppress the smaller amplitudes. Maximizing a norm ratio like $U(X,2,1)$ seems to balance the effort between pushing up large events and suppressing noise. Several modifications to the Wiggins-type algorithm to accomplish this were tried without success.

In this chapter Newton's method is used to derive an algorithm which is valid for arbitrary α_1 and α_2. It is compared to the Wiggins-type algorithm and steepest descent methods. An implementation in the time domain is developed and several results presented.

A. Derivation

Newton's method (32) for iteratively extremizing the multidimensional nonlinear function $U(X,\alpha_1,\alpha_2)$ has the form

$$X^{k+1} = X^k - [\nabla^2 U(X^k,\alpha_1,\alpha_2)]^{-1} \nabla U(X^k,\alpha_1,\alpha_2) \qquad (8\text{-}1)$$

where

X^k is the k-th estimate of the reflectivity series,

∇U is the gradient vector,

$$[\nabla U]_i = \frac{\partial U(X^k,\alpha_1,\alpha_2)}{\partial x_i} \quad ,$$

and

$\nabla^2 U$ is the Hessian matrix

$$[\nabla^2 U]_{ij} = \frac{\partial^2 U(X^k, \alpha_1, \alpha_2)}{\partial x_i \, \partial x_j} \quad .$$

For now only a single channel is considered. Later the channels will be averaged over to produce a multichannel estimate of the inverse filter.

The gradient vector is derived from Equation (6-7) as

$$[\nabla U]_i = \left(\frac{|x_i|^{\alpha_1 - 1}}{\mu_1} - \frac{|x_i|^{\alpha_2 - 1}}{\mu_2} \right) \operatorname{sgn}(x_i) \tag{8-2}$$

where

$$\mu_1 = \frac{1}{n} \sum_{i=1}^{n} |x_i|^{\alpha_1} \tag{8-3}$$

and

$$\mu_2 = \frac{1}{n} \sum_{i=1}^{n} |x_i|^{\alpha_2} \quad . \tag{8-4}$$

If the current estimate of the reflectivity series has distribution α_3, μ_1 and μ_2 can be expressed in terms of the scale β as

$$\mu_1 = \frac{\Gamma\left(\frac{\alpha_1 + 1}{\alpha_3}\right)}{\Gamma\left(\frac{1}{\alpha_3}\right)} \beta^{\alpha_1} \tag{8-5}$$

$$\mu_2 = \frac{\Gamma\left(\frac{\alpha_2 + 1}{\alpha_3}\right)}{\Gamma\left(\frac{1}{\alpha_3}\right)} \beta^{\alpha_2} \quad . \tag{8-6}$$

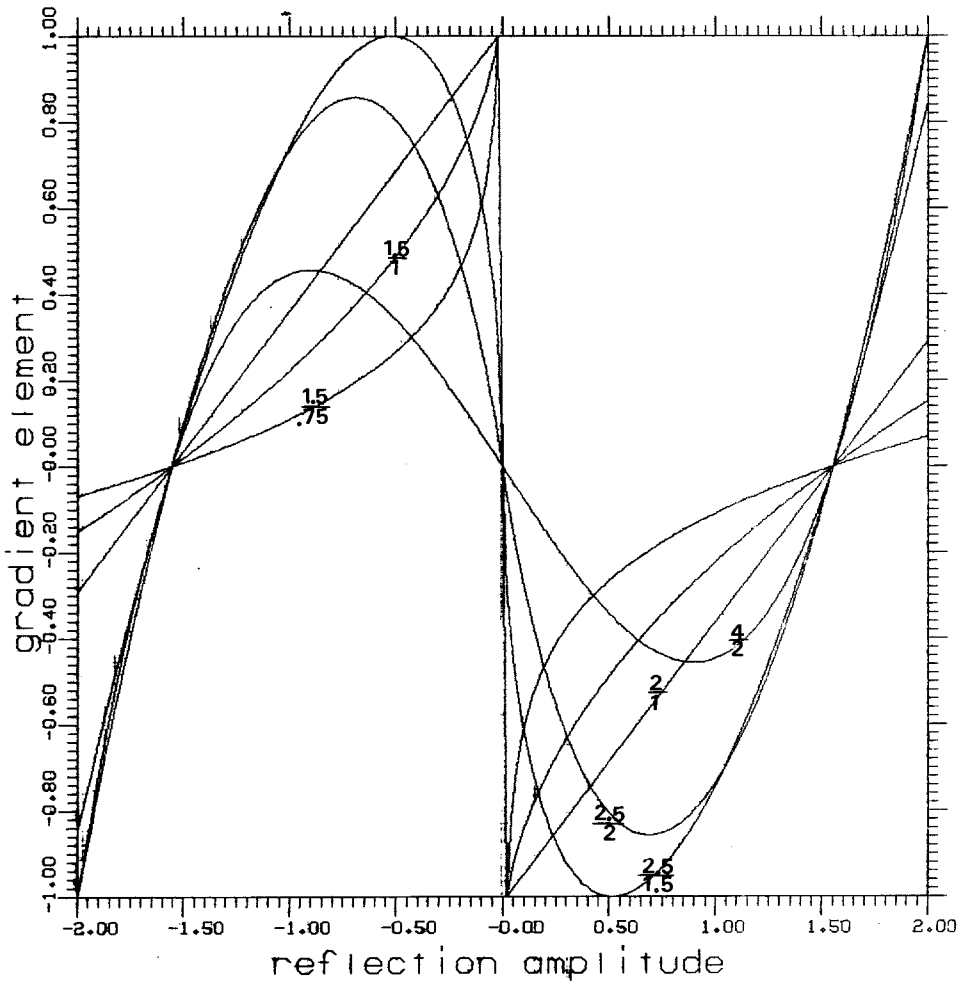
Using these results in Equation (8-2) gives the gradient

$$[\nabla U]_i = \left[\frac{|x_i|^{\alpha_1 - 1}}{\beta^{\alpha_1} \Gamma\left(\frac{\alpha_1+1}{\alpha_3}\right)} - \frac{|x_i|^{\alpha_2 - 1}}{\beta^{\alpha_2} \Gamma\left(\frac{\alpha_2+1}{\alpha_3}\right)} \right] \Gamma\left(\frac{1}{\alpha_3}\right) \operatorname{sgn}(x_i) \tag{8-7}$$

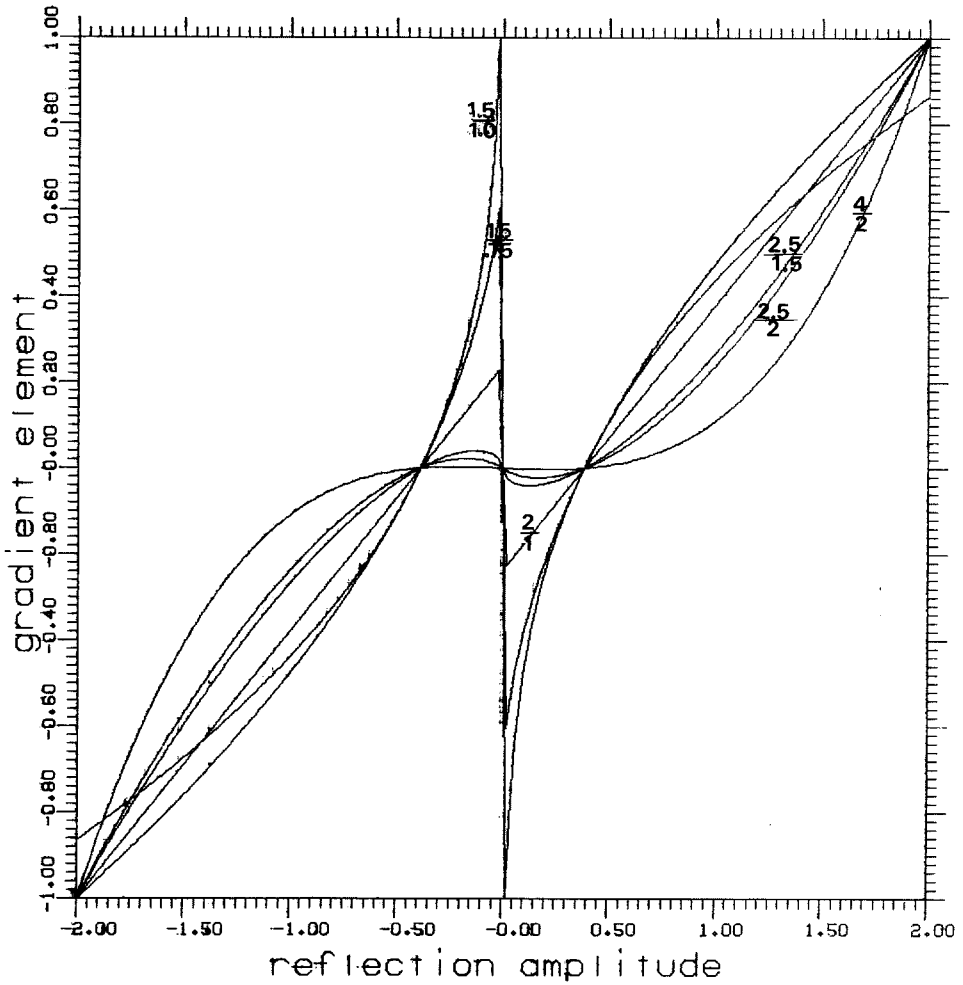
This can be expressed in terms of the variance by substituting Equation (5-7) in (8-7), yielding

$$[\nabla U]_i = \left\{ \frac{|x_i|^{\alpha_1 - 1}}{\left[\frac{\sigma^2 \Gamma\left(\frac{1}{\alpha_3}\right)}{\Gamma\left(\frac{3}{\alpha_3}\right)} \right]^{\alpha_1/2}} - \frac{|x_i|^{\alpha_2 - 1}}{\left[\frac{\sigma^2 \Gamma\left(\frac{1}{\alpha_3}\right)}{\Gamma\left(\frac{3}{\alpha_3}\right)} \right]^{\alpha_2/2}} \right\} \operatorname{sgn}(x_i) \tag{8-8}$$

Two plots of this function are displayed in Figure 8-1 for various combinations of α_1 and α_2 with x varying. Each plot is for a different current distribution α_3.



$\alpha_3 = 2.5$



$\alpha_3 = .75$

FIGURE 8-1. Plots of gradient element size versus reflection amplitudes for different norm ratios and different current distributions. The reflection amplitude scale is in standard deviations. Each curve represents a different norm ratio and is labeled α_1/α_2. The gradient element size varied greatly and was scaled to (-1,1) for plotting. Ratios with $\alpha_2 > 1$ are characterized by pushing up the large amplitude; when $\alpha_2 < 1$ the gradients concentrate on suppressing the small amplitudes. Changing α_3 only affects the zero crossing, not the shape of the curves.

The diagonal elements of the Hessian matrix are

$$[\nabla^2 U]_{ii} = \frac{(\alpha_1-1)\ |x_i|^{\alpha_1-2}}{\mu_1} - \frac{(\alpha_2-1)\ |x_i|^{\alpha_2-2}}{\mu_2}$$

$$- \frac{\alpha_1}{n}\left(\frac{|x_i|^{\alpha_1-1}}{\mu_1}\right)^2 + \frac{\alpha_2}{n}\left(\frac{|x_i|^{\alpha_2-1}}{\mu_2}\right)^2$$

$$+ 2\delta(x_i)\left(\frac{|x_i|^{\alpha_1-1}}{\mu_1} - \frac{|x_i|^{\alpha_2-1}}{\mu_2}\right) \qquad (8\text{-}9)$$

and the off-diagonal elements are

$$[\nabla^2 U]_{ij} = \left(\frac{\alpha_1}{n}\ \frac{|x_i x_j|^{\alpha_1-1}}{\mu_1^2} - \frac{\alpha_2}{n}\ \frac{|x_i x_j|^{\alpha_1-1}}{\mu_2^2}\right) \operatorname{sgn}(x_i)\ \operatorname{sgn}(x_j) \qquad (8\text{-}10)$$

Inversion of the Hessian matrix is not computationally practical because of its large size. Since the x's are random variables the Hessian can be written as the sum of a diagonal matrix and a zero mean random matrix whose norm is believed to be zero for increasing n. The diagonal matrix is found by taking the expected value of the Hessian. This approximately yields a constant matrix written

$$E([\nabla^2 U]) = \bar{c}\ I \qquad (8\text{-}11)$$

where

$$\bar{c} = \frac{\alpha_1-1}{\mu_1} E(|x|^{\alpha_1-2}) - \frac{\alpha_2-1}{\mu_2} E(|x|^{\alpha_2-2}) \qquad (8\text{-}12)$$

and I is the identity matrix.

The last term of Equation (8-9) has a zero expectation because the probability of a random variable being exactly zero is zero so

$$E[\delta(x_i)] = 0 \ . \tag{8-13}$$

The off-diagonal elements go to zero because they are signed and also scaled by $1/n$. The third and fourth terms of Equation (8-9) go to zero because they are inversely proportional to $n\beta^2$. This can be shown by using Equation (5-7) when taking the expectation of Equation (8-9).

The constant given by Equation (8-12) can be simplified using the expression for the absolute moments given by Equation (5-6). Assume the current estimate of the reflectivity series is described by the shape parameter α_3, then

$$\overline{c} = \frac{\alpha_3}{\beta^2}\left[\frac{\Gamma\left(\frac{\alpha_1+\alpha_3-1}{\alpha_3}\right)}{\Gamma\left(\frac{\alpha_1+1}{\alpha_3}\right)} - \frac{\Gamma\left(\frac{\alpha_2+\alpha_3-1}{\alpha_3}\right)}{\Gamma\left(\frac{\alpha_2+1}{\alpha_3}\right)}\right] \tag{8-14}$$

In deriving Equation (8-14), the integral

$$(\alpha-1)\int_0^\infty x^{\alpha-2}\, e^{-x}\, dx = (\alpha-1)\ \Gamma(\alpha-1) \tag{8-15}$$

had to be evaluated for $\alpha > 0$. The integral

$$\int_0^\infty x^{\alpha-2}\, e^{-x}\, dx \tag{8-16}$$

is defined only for $\alpha > 1$. Because

$$(\alpha-1)\ \Gamma(\alpha-1) = \Gamma(\alpha) \tag{8-17}$$

the pole at $\alpha = 1$ is simple and removable. Equation (8-14) is, therefore, valid for positive α_1, α_2, and α_3.

Using the maximum likelihood estimator for the scale, Equation (5-14) gives an estimate of the constant as

$$\bar{c} = \alpha_3 \left(\frac{\alpha_3 \sum_{i=1}^{n} |x_i|^{\alpha_3}}{n} \right)^{-2/\alpha_3} \left[\frac{\Gamma\left(\frac{\alpha_1+\alpha_3-1}{\alpha_3}\right)}{\Gamma\left(\frac{\alpha_1+1}{\alpha_3}\right)} - \frac{\Gamma\left(\frac{\alpha_2+\alpha_3-1}{\alpha_3}\right)}{\Gamma\left(\frac{\alpha_2+1}{\alpha_3}\right)} \right] \tag{8-18}$$

which is the estimate used in the current algorithm. A less efficient estimate for seismic data written in terms of the variance is

$$\bar{c} = \frac{\alpha_3 \Gamma\left(\frac{3}{\alpha_3}\right)}{\sigma^2 \Gamma\left(\frac{1}{\alpha_3}\right)} \left[\frac{\Gamma\left(\frac{\alpha_1+\alpha_3-1}{\alpha_3}\right)}{\Gamma\left(\frac{\alpha_1+1}{\alpha_3}\right)} - \frac{\Gamma\left(\frac{\alpha_2+\alpha_3-1}{\alpha_3}\right)}{\Gamma\left(\frac{\alpha_2+1}{\alpha_3}\right)} \right] \tag{8-19}$$

This estimate was derived using Equation (5-7) to relate the variance and scale. Figure 8-2 displays the constant for several combinations of α_1 and α_2 versus α_3.

When using Newton's method to maximize a nonlinear function, the Hessian matrix must be negative-definite for a maximum to be obtained. This requires the constant $\bar{c}$ to be negative. Figure 8-2 shows it to be negative so long as the estimated reflectivity series is described by a shape parameter less than two.

With seismic data the parameter is sometimes estimated as slightly greater than two. This usually happens when the signal to noise ratio is small. If this occurs in the multichannel case the seismogram can be left out of the estimation. In the single channel case the shape parameter

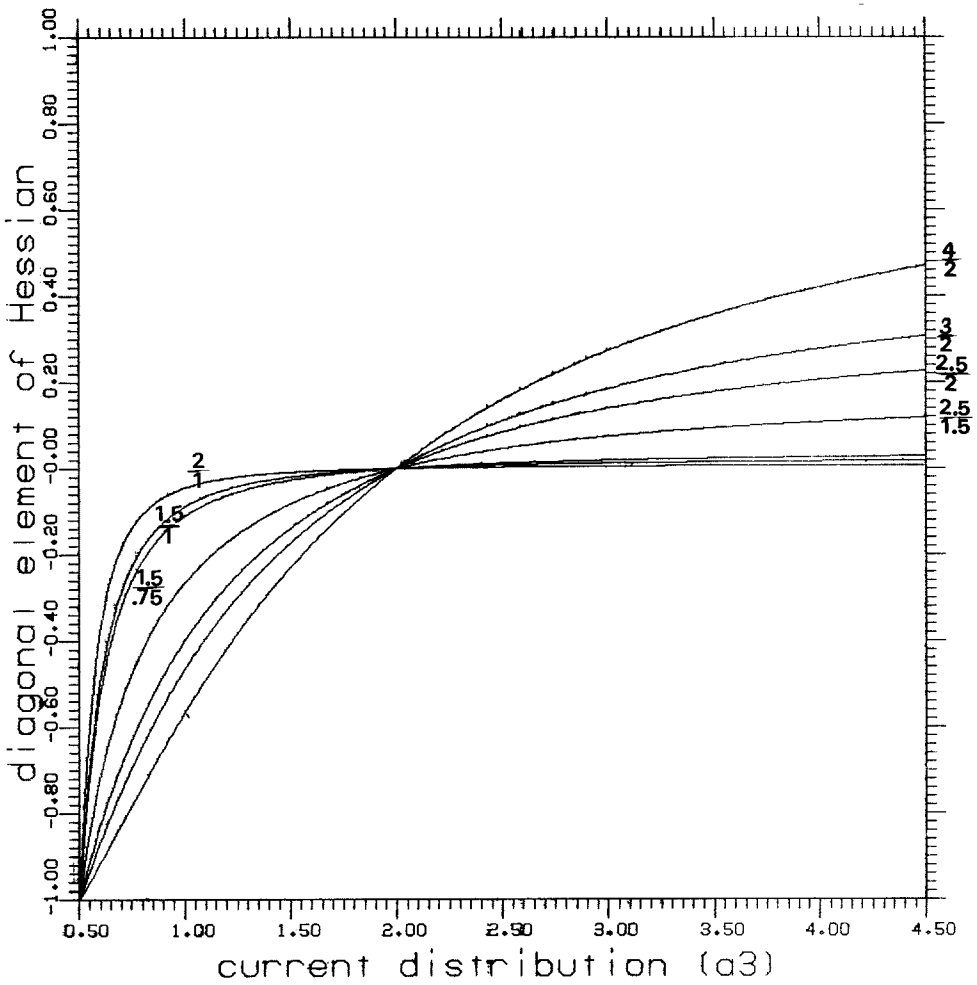


FIGURE 8-2. The expected value of the diagonal element of the Hessian is plotted against different current distributions. Each curve is labeled by the norm ratio α_1/α_2. The amplitudes were scaled to the interval (-1,1). The variance of the current distribution is held constant.

must be arbitrarily set below two.

Define perturbations to the currently estimated reflectivity series as

$$dX^{k+1} = X^{k+1} - X^{k} \quad . \tag{8-20}$$

From Equation (8-1) the perturbations are expressible in terms of the variance as

$$dX^{k+1} = [\nabla^2 U]^{-1} \nabla U \tag{8-21}$$

where $[\nabla^2 U]^{-1}$ is defined by Equation (8-19) and ∇U by Equation (8-8). This function is displayed in Figure 8-3.

The Hessian constant, $\bar{c}$, is a random variable and thus may be assigned a variance. Let c represent the first two terms of Equation (8-9). Assuming all other terms of the Hessian go to zero in expectation, the variance, $\sigma_{\bar{c}}^2$, is

$$\sigma_{\bar{c}}^2 = E[c^2] - \bar{c}^2 \quad . \tag{8-22}$$

Figure 8-4 is a plot of $\bar{c}/\sigma_{\bar{c}}$ as a function of α_3 for different norm ratios. It indicates that the variance decreases relative to $\bar{c}$ as the current distribution gets spikier. No results are presented for the 2/1 case because the variance is undefined for α_1 or $\alpha_2 \leq 1.5$.

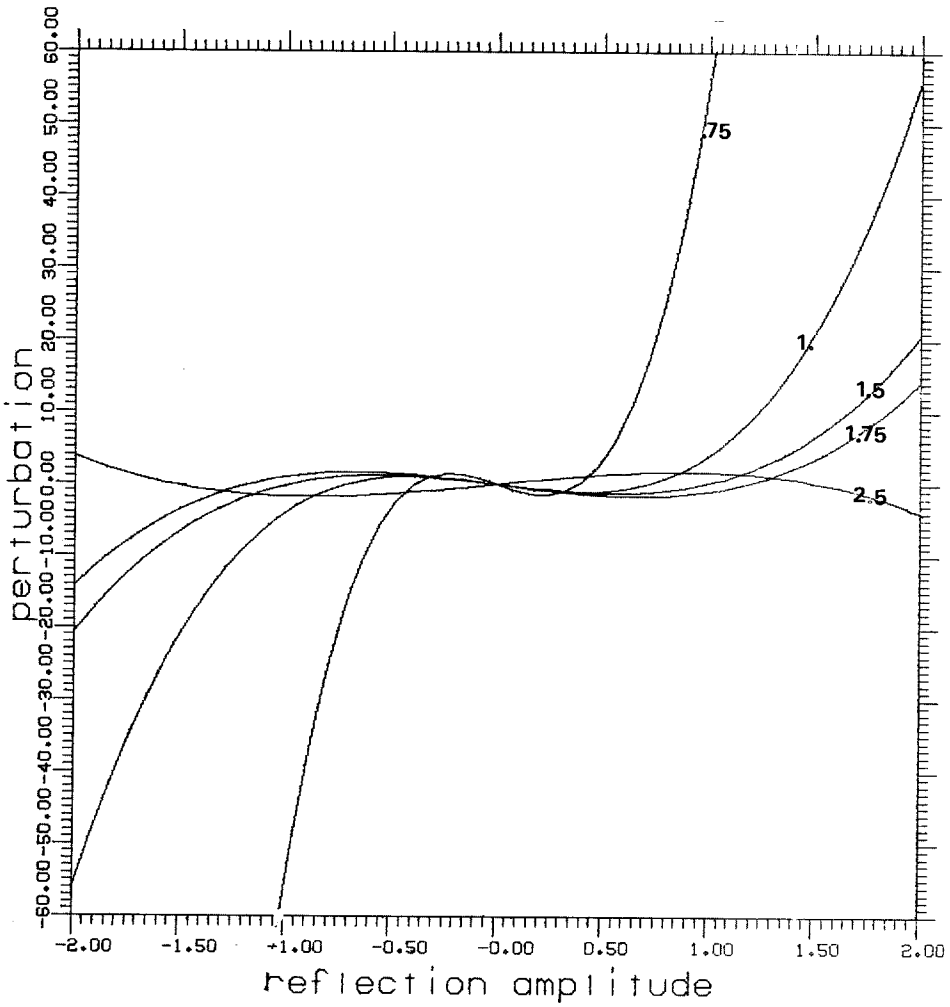


FIGURE 8-3. Perturbations to the reflectivity series relative to estimated reflection amplitudes. Both axes are in units of standard deviations. Each curve represents the perturbations for different current distributions. The norm ratio was $\alpha_1 = 4$ and $\alpha_2 = 2$.

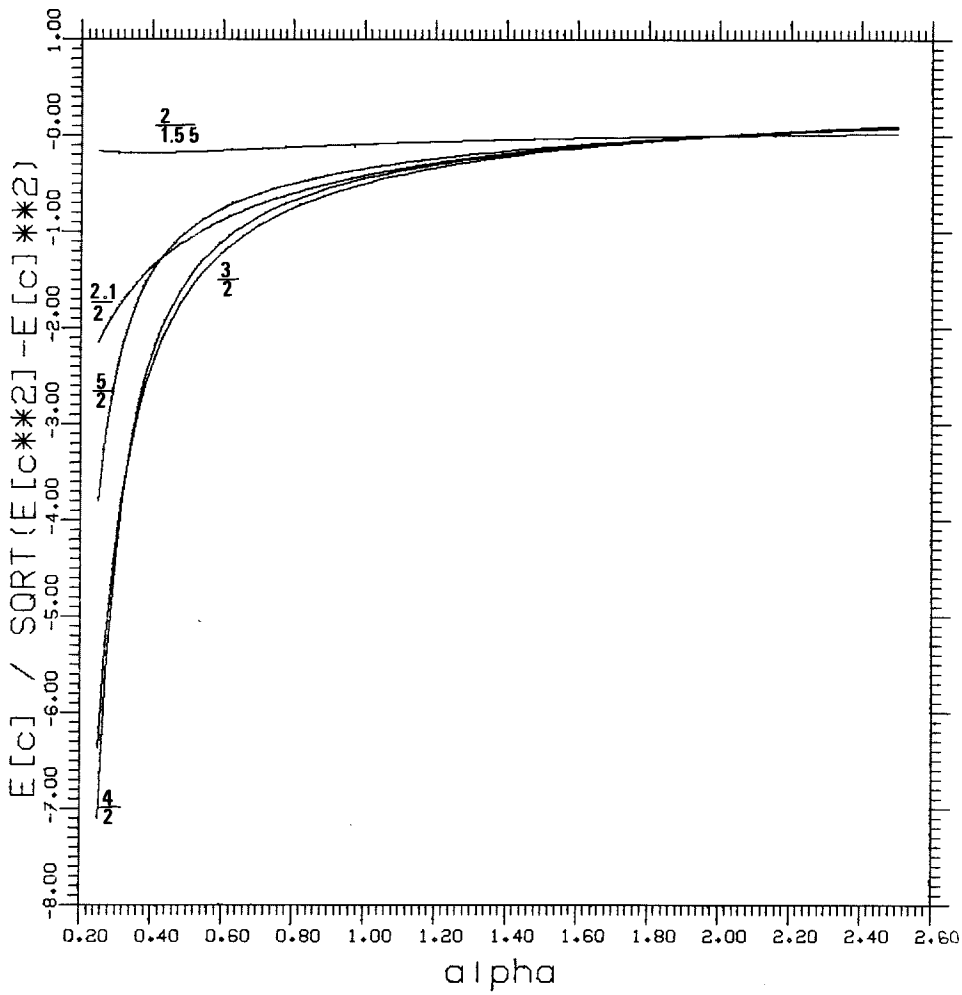
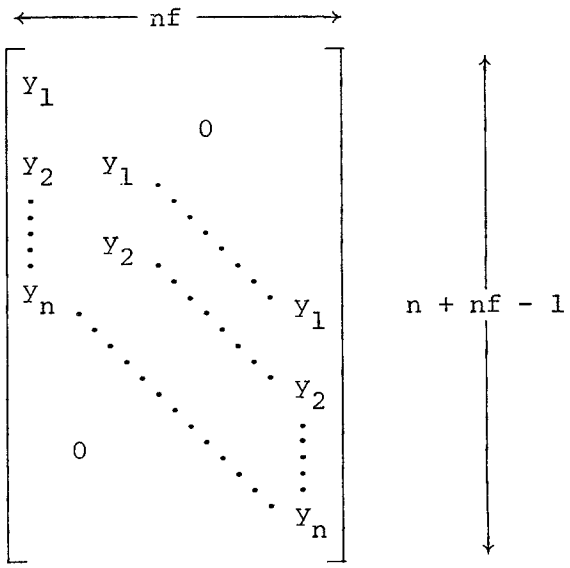


FIGURE 8-4. Each curve describes the distance to move in the direction of the gradient divided by its standard deviation. They are labeled by the norm ratio α_1/α_2 and plotted versus the shape parameter α_3. For the 4/2 curve, the value of -2 indicates the distance is $\bar{c} \pm \bar{c}/2$.

C. Time Domain Implementation

So far the Newton-type method has been expressed in terms of the reflectivity series. Now it is related to the inverse filter. Let [Y] represent the convolution matrix



where nf is the inverse filter length.

The k-th estimate of the reflectivity series, x^k, is the convolution of the input and k-th estimate of the inverse filter written (small letters now denote vectors),

$$[Y]f^k = x^k \quad . \tag{8-23}$$

Substituting this into Equation (8-1), Newton's method, gives

$$[Y]df^{k+1} = -[\nabla^2 U(x^k,\alpha_1,\alpha_2)]^{-1} \nabla U(x^k,\alpha_1,\alpha_2) \tag{8-24}$$

where

$$df^{k+1} = f^{k+1} - f^k \tag{8-25}$$

is an unknown perturbation to the current filter estimate.

Because the length of the filter is less than that of the reflectivity series the system is overdetermined. The least squares estimate of the filter perturbations is found by solving the system

$$[\bar{Y}\ Y]\ df^{k+1} = -[\bar{Y}]\ [\nabla^2 U(X^k,\alpha_1,\alpha_2)]^{-1}\ \nabla U(X^k,\alpha_1,\alpha_2) \tag{8-26}$$

where $[\bar{Y}]$ is the transpose of [Y]. Since the expectation of the Hessian matrix is a constant diagonal matrix the system can be rewritten

$$c[\bar{Y}\ Y]\ df^{k+1} = -[\bar{Y}]\ \nabla U(X^k,\alpha_1,\alpha_2) \tag{8-27}$$

where c is the constant given by Equation (8-18). Moving the constant to the LHS is essential to the multichannel algorithm as its inverse blows up when the current distribution is near Gaussian. This is illustrated by Figure 8-5.

For the multichannel case the filter is assumed common to all channels. Summing Equation (8-27) over channels yields the system

$$\left[\sum_{j=1}^{m} c_j\ \bar{Y}_j\ Y_j\right] df^{k+1} = -\sum_{j=1}^{m} [\bar{Y}_j]\ \nabla U(X_j{}^k,\alpha_1,\alpha_2) \tag{8-28}$$

which can be written

$$R\ df^{k+1} = -dg^{k+1} \tag{8-29}$$

where

R is an nf x nf Toeplitz matrix. It is a weighted sum over channels of the input autocorrelation matrices,

df^{k+1} is a vector of length nf representing unknown perturbations to the k-th filter,

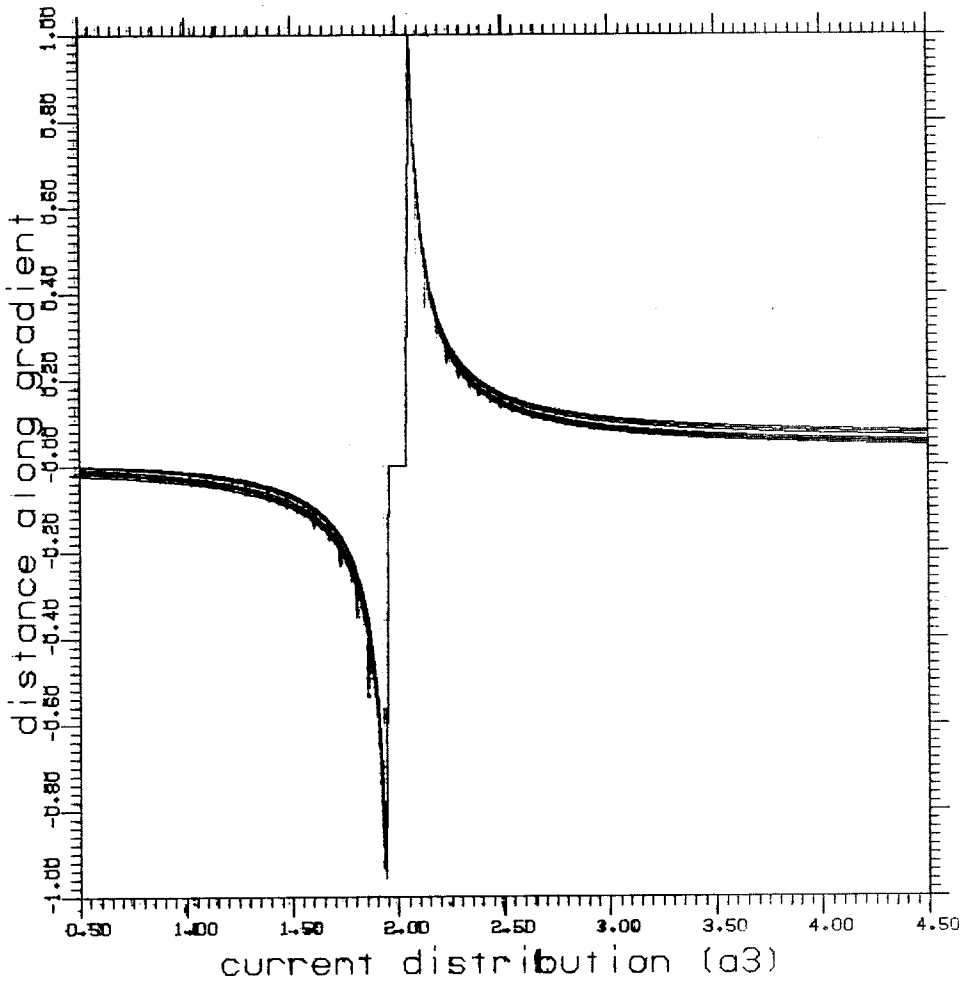


FIGURE 8-5. The inverse of the Hessian constant is similar to a distance to move in the direction of the gradient. When the current distribution goes to Gaussian the distance goes to infinity. Distances for distributions near Gaussian are not plotted.

and

dg^{k+1} is a vector of length nf. It is a weighted sum over channels of the crosscorrelations between the inputs and perturbations to the k-th estimated reflectivity series.

The inverse of the matrix R is efficiently computed using the *Levinson recursion* (33). To stabilize the inversion a *ridge regression technique* (34) is easily implemented. The solution is written

$$df^{k+1} = -[R + \varepsilon I]^{-1} dg^{k+1} \tag{8-30}$$

where ε is the constant used for ridge regression. It is dependent on the signal to noise ratio and in the current algorithm it defaults to .001. It might be useful to estimate a different constant for each channel based on the currently estimated distribution of its reflectivity series. This has not yet been tried.

After the system [Equation (8-29)] is solved the filter perturbations are added to the previously estimated filter. Convolving this filter with the inputs gives a new estimate of the reflectivity series. If the initial estimate of the reflectivity series is noisy, nearly Gaussian, or far from the "true" solution, the algorithm can become unstable.

A partial remedy is to scale the filter perturbations so the resulting estimate of the reflectivity series maximizes some easily computed objective function. Let

$$f^{k+1} = f^{k} + \lambda df^{k+1} \tag{8-31}$$

where λ is called the step size or scale. The k+1-th estimate of the reflectivity series is

$$x^{k+1} = [Y]f^{k+1}$$
$$= [Y]f^{k} + \lambda[Y]\ df^{k+1}$$
$$= x^{k} + \lambda\ dx^{k+1} \tag{8-32}$$

where dx^{k+1} is the perturbation to the reflectivity series.

Generalized kurtosis,

$$K(1,x^{k} + \lambda\, dx^{k+1}) = \frac{\frac{1}{n}\sum_{i=1}^{n} |x_i^{k} + \lambda\, dx_i^{k+1}|^2}{\left(\frac{1}{n}\sum_{i=1}^{n} |x_i^{k} + \lambda\, dx_i^{k+1}|\right)^2} \tag{8-33}$$

is an easily computed objective function. The optimum step size is obtained by varying λ until kurtosis is maximized.

The *Fibonacci method* (32) is a simple search procedure for determining the step size. It requires the objective function to be unimodal within a given search interval, $[\ell_1, r_1]$. The Fibonacci numbers are recursively defined by

$$F_0 = 0,\ F_1 = 1,\ F_i = F_{i-1} + F_{i-2};\ i = 2,3,... \tag{8-34}$$

The interval reduction is accomplished by comparing the values of $K(1;X^{k} + \lambda_i^{1}\ dx^{k+1})$ and $K(1,X^{k} + \lambda_i^{\dagger}\ dx^{k+1})$ where

$$\lambda_i^{1} = \ell_i + \frac{F_{N-i}}{F_{N+2-i}}\ (r_i - \ell_i) \tag{8-35}$$

and

$$\lambda_i^{\dagger} = \ell_i + \frac{F_{N+1-i}}{F_{N+2-i}}\ (r_i - \ell_i)\ . \tag{8-36}$$

If the function's value at λ_i^{1} is greater than or equal to that at $\lambda_i^{\dagger}$, the right side of the interval is modified by setting $r_i = \lambda_i^{\dagger}$, otherwise

the left side is increased by setting $\ell_i = \lambda_i^1$. The accuracy of the method is determined by the number of function evaluations, N. For example, eleven evaluations will bracket the optimum step size within one percent of the initial search interval.

Figure 8-6 shows the estimated distribution (inversely proportional to kurtosis) of the reflectivity series versus step size for successive iterations. It is clearly unimodal for a positive step size. That the Newton-type method predicts the step size fairly accurately is evidenced by the fact that the minimum at all iterations is approximately one.

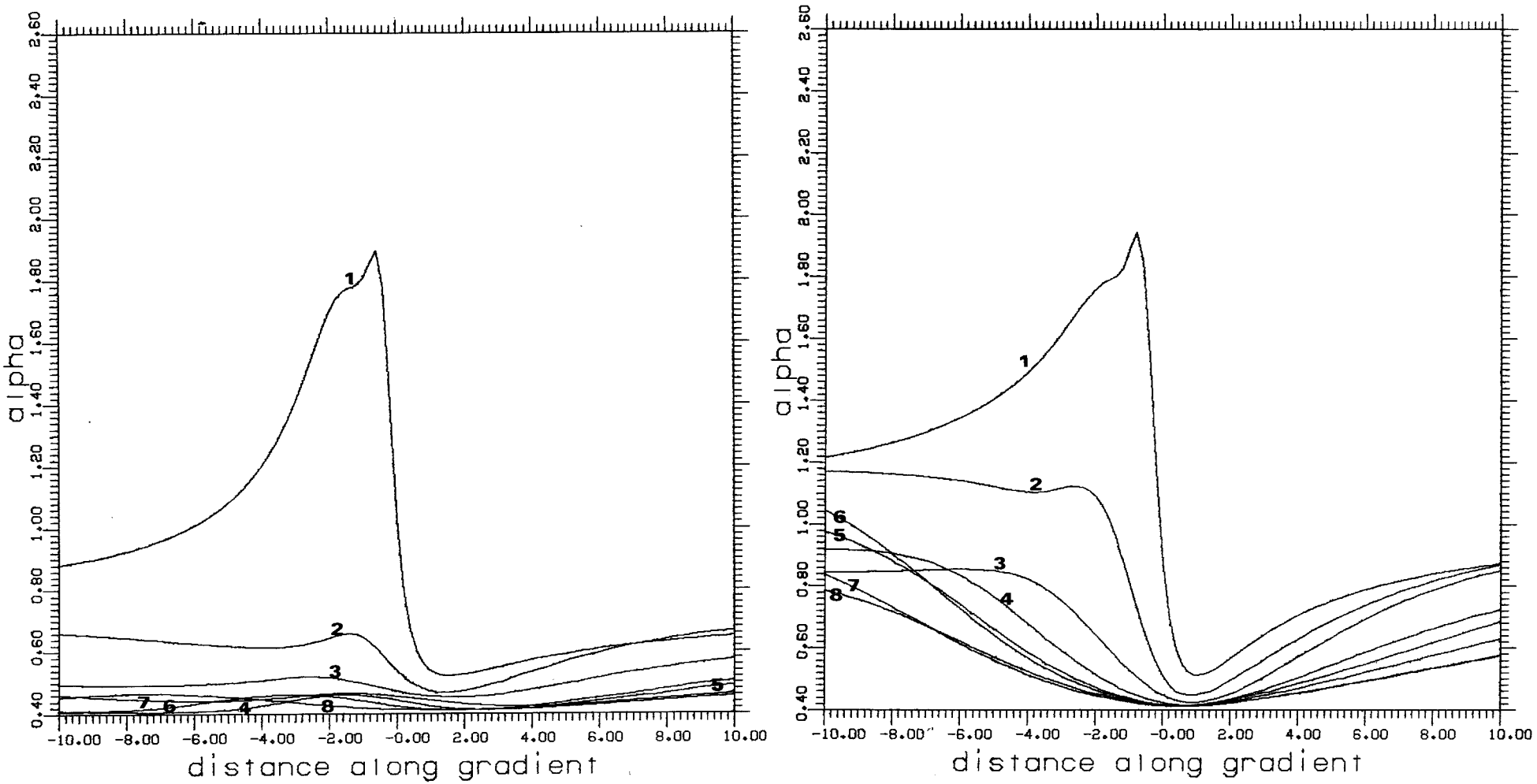


FIGURE 8-6. The distribution of the estimated reflectivity series for varying step size is displayed. A step size of 1 is estimated by the Newton-type method. Each curve represents the distribution at different iterations. On the right the norm ratio was $\alpha_1 = 2.5$, $\alpha_2 = 1.5$. On the left the ratio was $\alpha_1 = 4$, $\alpha_2 = 2$. The near offset trace of Figure 8-7 was being deconvolved.

C. Relation Between Wiggins and Newton-type Methods

The Wiggins-type method iteratively solves

$$\left[\sum_{j=1}^{m} \frac{\bar{Y}_j Y_j}{\frac{1}{n}\sum_{i=1}^{n} (x_{ij}^k)^2}\right] f^{k+1} = \sum_{j=1}^{m} \frac{\bar{Y}_j Z(X_j^k)}{\frac{1}{n}\sum_{i=1}^{n} |x_{ij}^k|^{\alpha_1}} \tag{8-37}$$

for an estimate of the inverse filter. $Z(X_j^k)$ represents a vector whose elements are the magnitudes of the current reflectivity series to the (α_1-1)-th power where their signs are retained. Since

$$f^{k+1} = f^k + df^{k+1} \tag{8-38}$$

and

$$X_j^k = [Y_j]\, f_j^k \,, \tag{8-39}$$

Equation (8-37) can be rewritten in terms of filter perturbations as

$$\left[\sum_{j=1}^{m} \frac{\bar{Y}_j Y_j}{\frac{1}{n}\sum_{i=1}^{n} (x_{ij}^k)^2}\right] df^{k+1} = \sum_{j=1}^{m} \bar{Y}_j \left[\frac{Z(X_j^k)}{\frac{1}{n}\sum_{i=1}^{n} |x_{ij}^k|^{\alpha_1}} - \frac{X_j^k}{\frac{1}{n}\sum_{i=1}^{n} (x_{ij}^k)^2}\right] . \tag{8-40}$$

This is similar to the Newton-type algorithm which maximizes $U(X;\alpha_1,2)$ by solving

$$\left[\sum_{j=1}^{m} c_j \bar{Y}_j Y_j\right] df^{k+1} = -\sum_{j=1}^{m} \bar{Y}_j \left[\frac{Z(X_j^k)}{\frac{1}{n}\sum_{i=1}^{n} |x_{ij}^k|^{\alpha_1}} - \frac{X_j^k}{\frac{1}{n}\sum_{i=1}^{n} (x_{ij}^k)^2}\right] \tag{8-41}$$

where c_j is a constant determined by the estimated distribution of X_j^k.

The two methods are identical except for a scaling of the autocorrelations, $c(\alpha_1,\alpha_3)$, because from Equation (8-19), with $\alpha_2 = 2$,

$$c_j = \frac{-c(\alpha_1,\alpha_3)}{\frac{1}{n}\sum (x_{ij}{}^k)^2} \quad . \tag{8-42}$$

D. An Alternate Algorithm

The Newton-type method can become unstable if the current estimated distributions vary greatly from channel to channel. In practice it is generally stable and converges quickly. The gradient vector defined by Equation (8-2) is orthogonal to the current estimate of the reflectivity series. Because of this the method is equivalent to steepest descent (32) where the Hessian constant represents the distance, determined by each channel's distribution, to move in the direction of the gradient.

An alternate method is to equally weight each channel's contribution to the sums in Equation (8-28). This has been accomplished by dividing the autocorrelations by their zero lag,

$$e_j = \sum_{i=1}^{n} y_{ij}{}^2 \tag{8-43}$$

and scaling the crosscorrelations by

$$d_j = \frac{1}{\sqrt{\left(\sum_{i=1}^{n} dx_{ij}{}^2\right)\left(\sum_{i=1}^{n} y_{ij}{}^2\right)}} \quad . \tag{8-44}$$

The system of equations represented by Equation (8-29) is solved and the step size then determined to maximize the kurtosis. Computationally this

modification is attractive because the input autocorrelations are only computed once. This advantage is balanced by the cost of a linear search to determine the optimum step size.

E. Application to Field-Recorded Data

The marine seismic gather presented in the last chapter is used here to show how the Newton-type method performs. Figure 8-7 displays the 48-trace gather used as input to the algorithm. The recorded gather was time-shifted to approximately flatten the reflection from the seafloor. The shifted gather was divided into eight groups of six traces. Each group had a different exponential gain applied to approximately correct for the effects of spherical divergence.

Figure 8-8 is the deconvolution achieved by maximizing U(X,2,1) over the six near-offset traces. The estimated inverse filter was applied to the entire gather. The deconvolution is noticeably simpler in appearance, especially the near-offset traces. The Newton-type method with no limiting of step size was used. It converged in five iterations.

Figure 8-9 is the deconvolution obtained by maximizing U(X,2,1) where a different inverse filter was estimated for each group. The Newton method with no limit on the step size became unstable for several of the groups. This is believed to be due to high noise levels, end effects, and the channels not being independent.

Many methods were tried to improve the robustness of the algorithm. Figure 8-10 is a result obtained after eleven iterations where

$$U[X, 2\hat{\alpha}, \max(\hat{\alpha}, 1)] \tag{8-45}$$

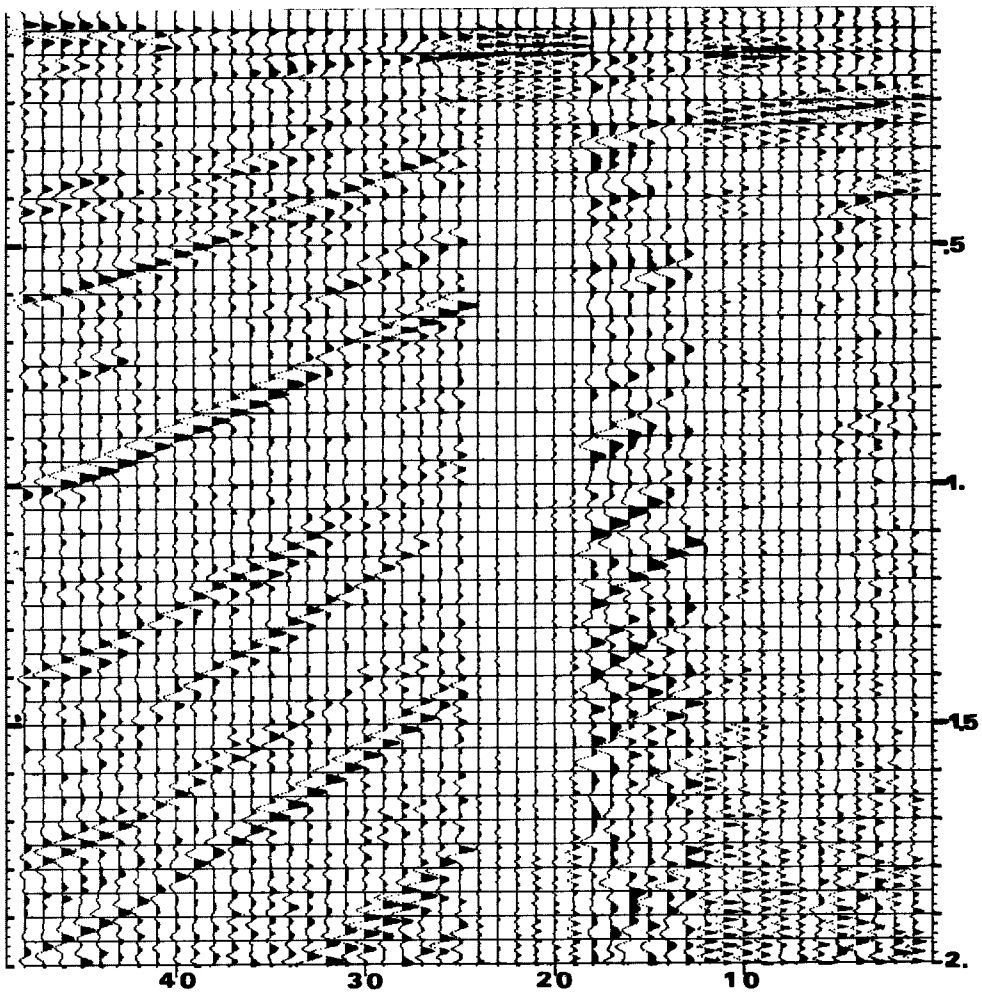


FIGURE 8-9. Results obtained using U(X;2,1) to estimate an inverse filter for each six-trace subgroup of the gather.

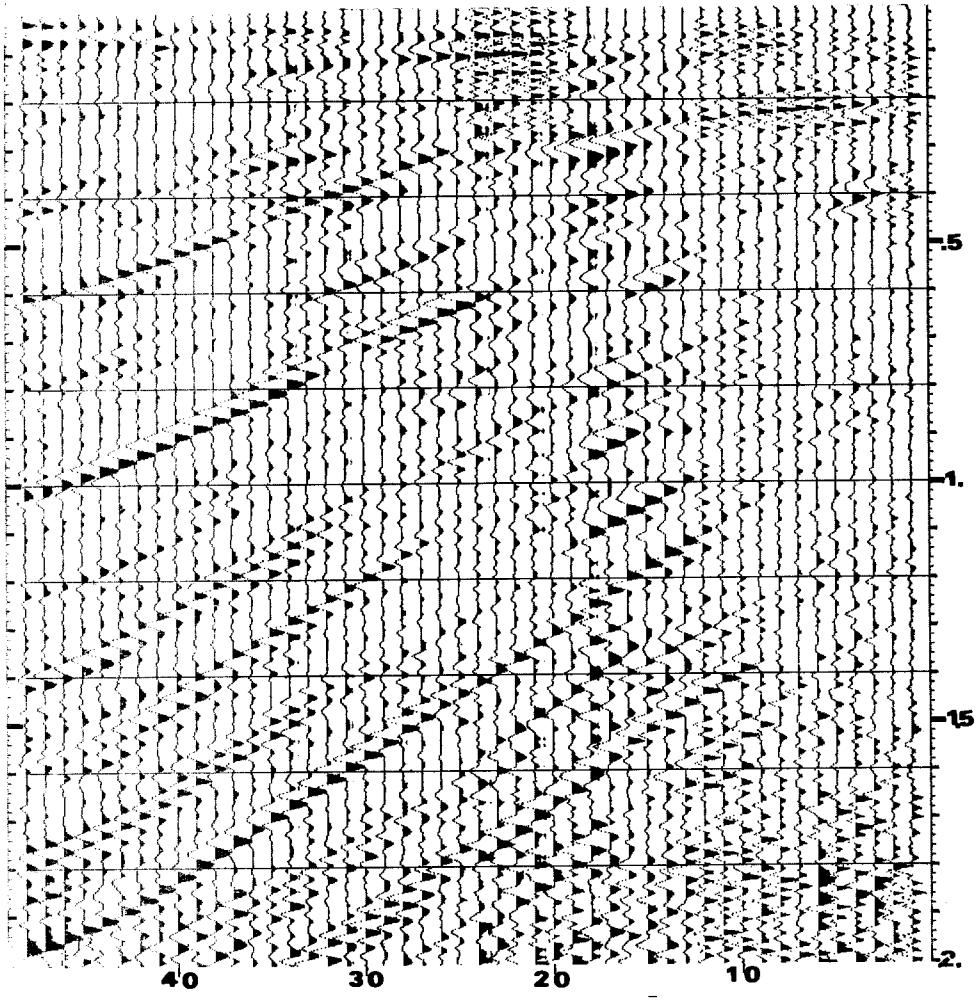


FIGURE 8-10. Results obtained using U[X,2$\hat{\alpha}$,max($\hat{\alpha}$,1)] to estimate a filter for each six-trace group.

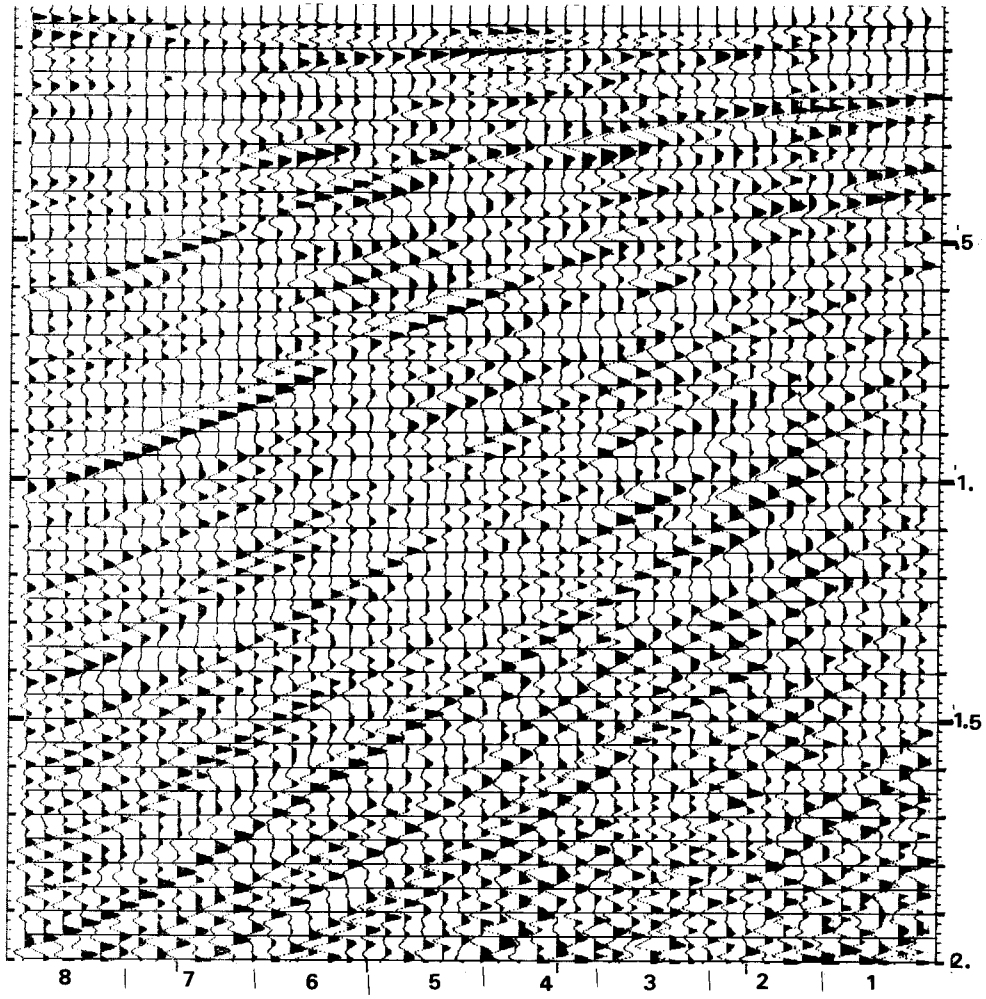


FIGURE 8-7. Inputs to deconvolution algorithm. Time shifting and an exponential gain correction were applied.

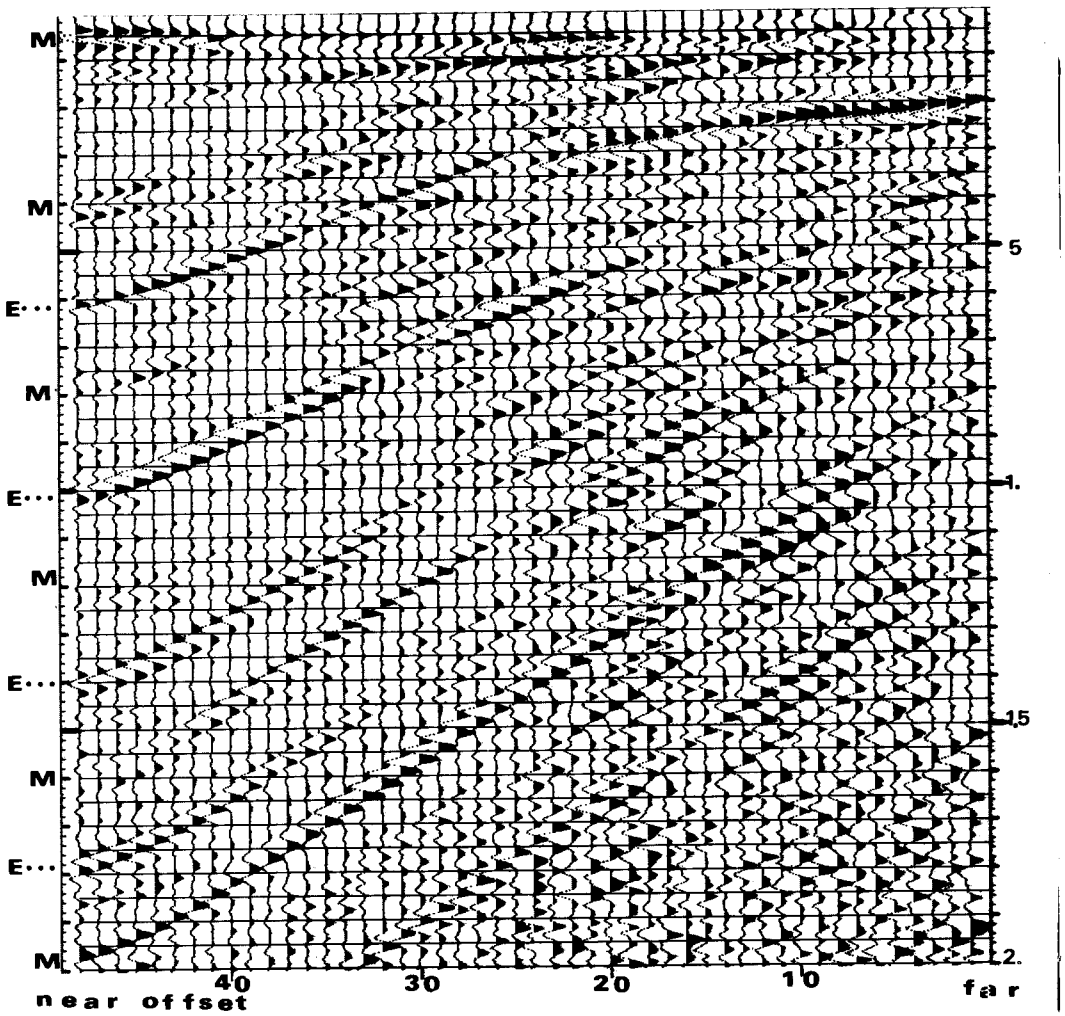


FIGURE 8-8. Results obtained using U(X,2,1) to estimate a filter from traces 42-48. The filter was applied to the entire gather. The seafloor and its multiples are marked on the left by M. A strong primary and its peglegs are marked by E.

was maximized using a limited step. In this method the distribution of each channel's estimated reflectivity series, $\hat{\alpha}$, was used to determine the norm ratio. The limiting of step size was determined from only one trace of the gather because of core limitations. The results are considerably better than those of Figure 8-9. The noise in the lower right-hand corner is believed due to end effects.

Figure 8-11 is a near-offset section extracted from the same line as the shot gather displayed in Figure 8-7. Prominent bubble oscillations mask primary events in several areas. The ungained section was deconvolved using the Newton-type method to maximize U(X,2,1). The initial filter was 138 points long. A spike was placed at the 25th point. Figure 8-12 is the result obtained after three iterations. The bubble oscillations were effectively suppressed. Precursors before the seafloor reflection, a problem caused by the Wiggins-type algorithm, are absent. This is attributed to the sensitivity of U(X,2,1) to very small reflection amplitudes. U(X,4,2), being sensitive only to the larger amplitudes, yielded noticeable precursors when applied to this dataset.

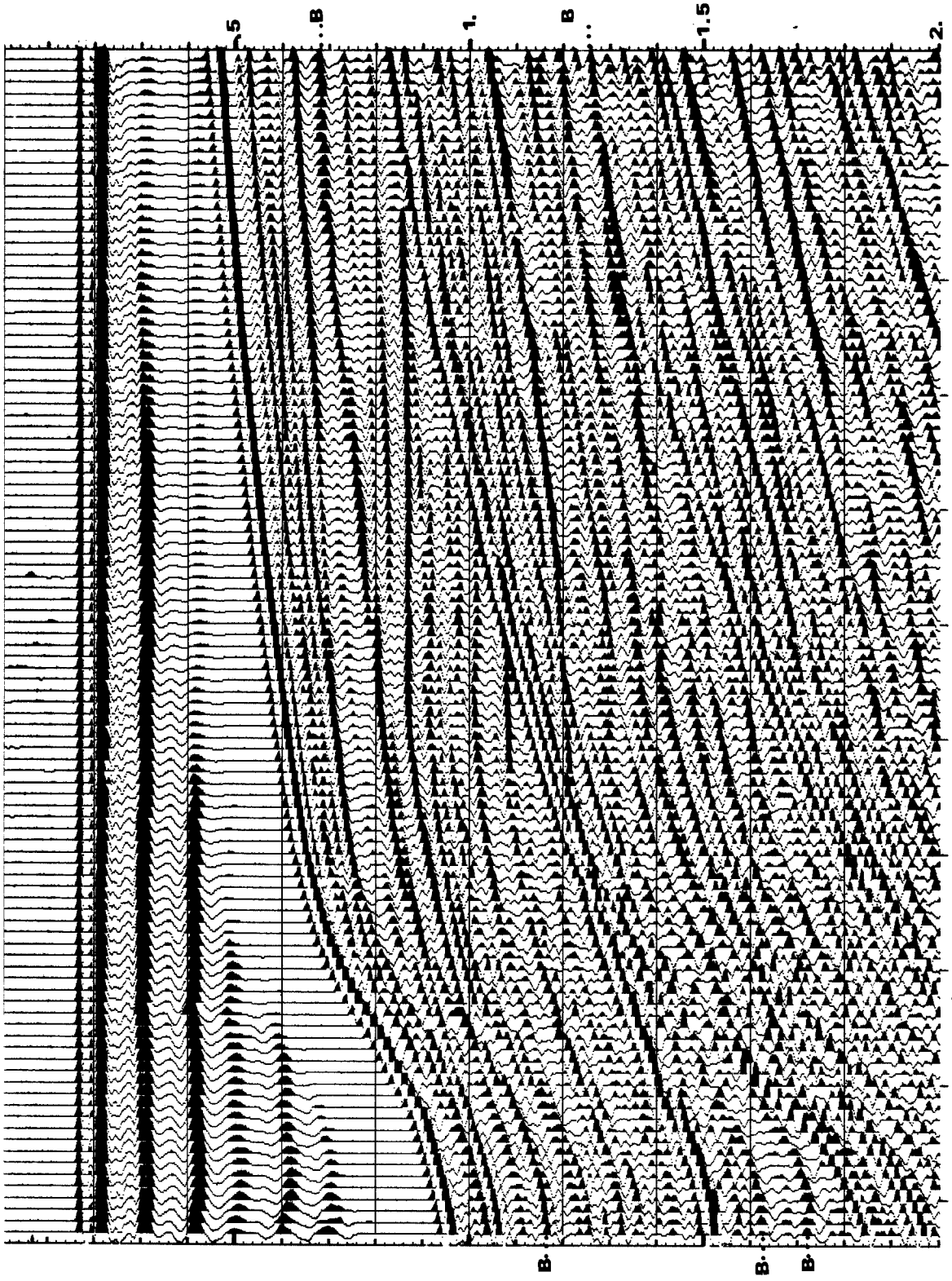


FIGURE 8-11. Several areas of this near-offset section display the airgun bubble masking primary reflections. These areas are indicated by B's. The traces were AGC'ed before plotting.

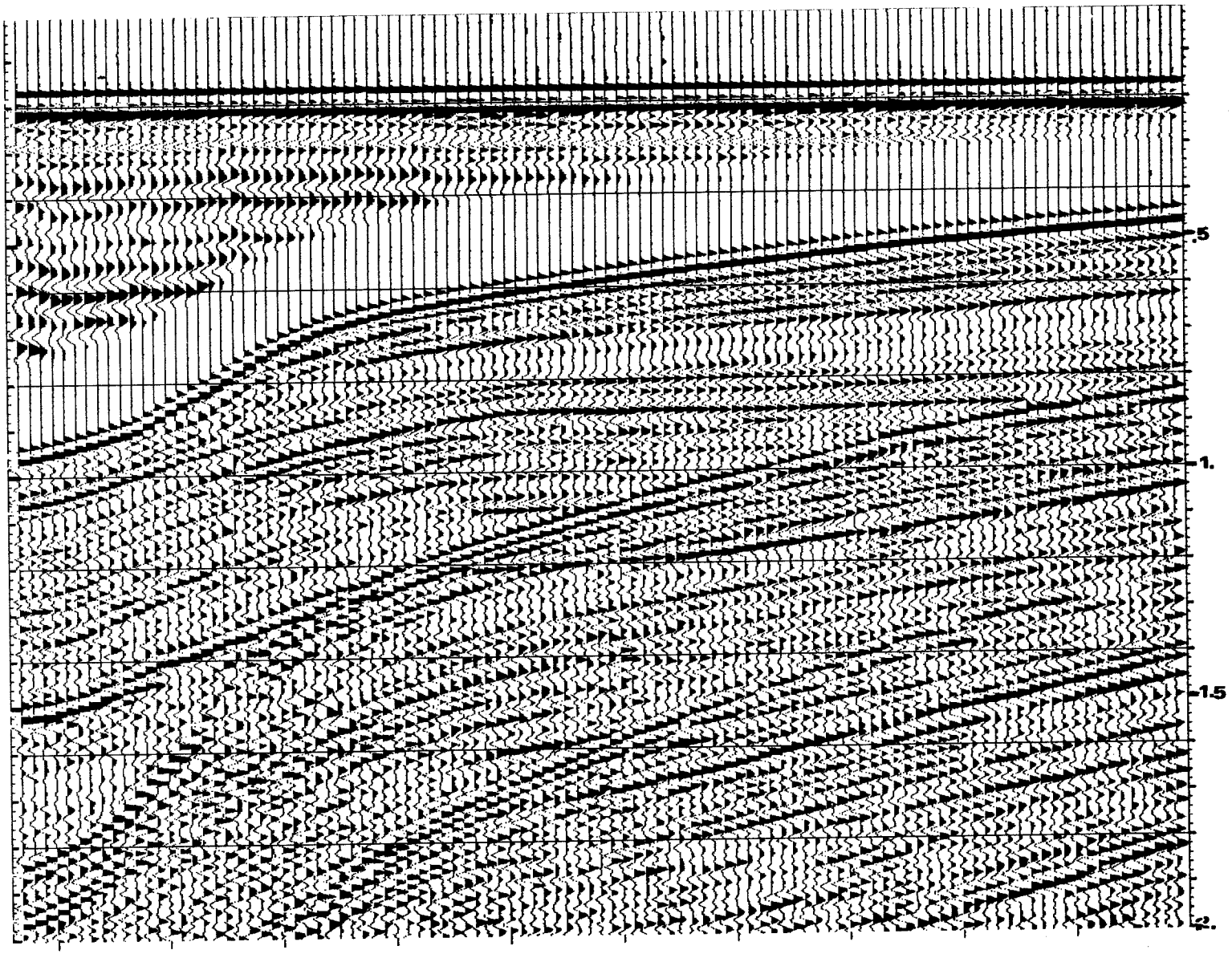


FIGURE 8-12. Deconvolved near-offset section obtained by maximizing U(X,2,1). The bubbles oscillating labeled on Figure 8-11 are suppressed. What is left of the direct arrival indicates the airgun signature varies slightly from shot to shot.