Chapter 3

Application

Conventional data processing does not involve inversion of huge matrices, and missing data are often assumed to be zero. Nevertheless, it often works very well. In this chapter we first see what conventional processing is in terms of the formulation of chapter 2 (multichannel transposition). We then apply multichannel inversion to both the design of three-dimensional surveys and to data processing.

§ 3.1 Multichannel transposition

If there is no spatial aliasing, then $N = 1$ in equation (2.6.2), and we have

$$\begin{pmatrix} \mathbf{d}_1(k) \\ \mathbf{d}_2(k) \\ \mathbf{d}_3(k) \\ \mathbf{d}_4(k) \\ \mathbf{d}_5(k) \end{pmatrix} = \begin{pmatrix} \mathbf{D}^+{}_1(k) \\ \mathbf{D}^+{}_2(k) \\ \mathbf{D}^+{}_3(k) \\ \mathbf{D}^+{}_4(k) \\ \mathbf{D}^+{}_5(k) \end{pmatrix} \Big(\mathbf{m}(k) \Big) \ . \tag{3.1.1}$$

In conventional processing, DMO is applied to each common-offset section ($\mathbf{D}_j \, \mathbf{d}_j$), and then the results are stacked ($\sum_j$):

$$\begin{aligned} \text{stack} &= \sum_{j=1}^{J} \mathbf{D}_j \, \mathbf{d}_j \\ &= \sum_{j=1}^{J} \mathbf{D}_j \, \mathbf{D}^+{}_j \ \mathbf{m} \\ &\approx J \ \mathbf{m} \ . \end{aligned} \tag{3.1.2}$$

The matrix $\sum \mathbf{D}_j \, \mathbf{D}^+{}_j$ is shown in Figure 3.1.1. It is practically the identity matrix times the CDP-fold, J.

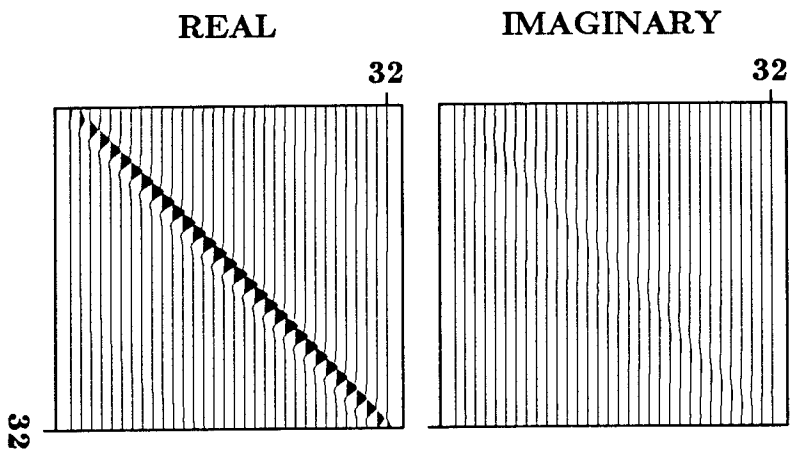


FIG. 3.1.1. The matrix $\sum \mathbf{D}_j \mathbf{D}^+{}_j$.

The processing in (3.1.2) is the application of the transpose of the matrix in (3.1.1),

$$\begin{pmatrix} \mathbf{D}^+{}_1(k) \\ \mathbf{D}^+{}_2(k) \\ \mathbf{D}^+{}_3(k) \\ \mathbf{D}^+{}_4(k) \\ \mathbf{D}^+{}_5(k) \end{pmatrix}^* = \begin{pmatrix} \mathbf{D}_1(k) & \mathbf{D}_2(k) & \mathbf{D}_3(k) & \mathbf{D}_4(k) & \mathbf{D}_5(k) \end{pmatrix}, \qquad (3.1.3)$$

because for each block,

$$\mathbf{D}^+{}_j{}^*(k) = \mathbf{D}_j(k) \; . \qquad (3.1.4)$$

The transpose operator is often called the inverse — for example, diffraction is not the inverse, but the transpose of migration (Claerbout, 1985). Nercessian et al (1985) found that migration is the transpose of the operator they were inverting.

We have seen in equation (3.1.2) and in Figure 3.1.1 that transpose processing (DMO stacking) is adequate when no data are missing. Now let us see if this approach is appropriate when the data are spatially aliased.

First we write equation (2.6.2) as

$$\mathbf{d} = \mathbf{G}\mathbf{m} \; . \qquad (3.1.5)$$

Then we define

$$\hat{\mathbf{m}} \equiv \mathbf{G}^* \mathbf{d} \; . \qquad (3.1.6)$$

To see what $\hat{\mathbf{m}}$ is, consider an example of $N=3$ and $J=4$:

$$\mathbf{G} = \begin{pmatrix} \mathbf{D}^+{}_1(k-\kappa) & \mathbf{D}^+{}_1(k) & \mathbf{D}^+{}_1(k+\kappa) \\ \mathbf{D}^+{}_2(k-\kappa) & \mathbf{D}^+{}_2(k) & \mathbf{D}^+{}_2(k+\kappa) \\ \mathbf{D}^+{}_3(k-\kappa) & \mathbf{D}^+{}_3(k) & \mathbf{D}^+{}_3(k+\kappa) \\ \mathbf{D}^+{}_4(k-\kappa) & \mathbf{D}^+{}_4(k) & \mathbf{D}^+{}_4(k+\kappa) \end{pmatrix} . \tag{3.1.7}$$

The transpose of $\mathbf{G}$ is

$$\mathbf{G}^* = \begin{pmatrix} \mathbf{D}_1(k-\kappa) & \mathbf{D}_2(k-\kappa) & \mathbf{D}_3(k-\kappa) & \mathbf{D}_4(k-\kappa) \\ \mathbf{D}_1(k) & \mathbf{D}_2(k) & \mathbf{D}_3(k) & \mathbf{D}_4(k) \\ \mathbf{D}_1(k+\kappa) & \mathbf{D}_2(k+\kappa) & \mathbf{D}_3(k+\kappa) & \mathbf{D}_4(k+\kappa) \end{pmatrix} , \tag{3.1.8}$$

hence,

$$\begin{pmatrix} \hat{\mathbf{m}}(k-\kappa) \\ \hat{\mathbf{m}}(k) \\ \hat{\mathbf{m}}(k+\kappa) \end{pmatrix} = \begin{pmatrix} \mathbf{D}_1(k-\kappa) & \mathbf{D}_2(k-\kappa) & \mathbf{D}_3(k-\kappa) & \mathbf{D}_4(k-\kappa) \\ \mathbf{D}_1(k) & \mathbf{D}_2(k) & \mathbf{D}_3(k) & \mathbf{D}_4(k) \\ \mathbf{D}_1(k+\kappa) & \mathbf{D}_2(k+\kappa) & \mathbf{D}_3(k+\kappa) & \mathbf{D}_4(k+\kappa) \end{pmatrix} \begin{pmatrix} \mathbf{d}_1(k) \\ \mathbf{d}_2(k) \\ \mathbf{d}_3(k) \\ \mathbf{d}_4(k) \end{pmatrix} . \tag{3.1.9}$$

The data, $\mathbf{d}$, are given for the frequencies $|k| < \kappa/2$, (because for $|k| > \kappa/2$ the spectrum replicates). $\hat{\mathbf{m}}$ is defined in equation (3.1.9) for a wider range of frequency: $|k| < 3/2\kappa$:

$$\hat{\mathbf{m}}(k) = \begin{cases} \sum_{j=1}^{J} \mathbf{D}_j(k)\mathbf{d}_j(k+\kappa) & \text{for } -3\kappa/2 < k < -\kappa/2 \\ \sum_{j=1}^{J} \mathbf{D}_j(k)\mathbf{d}_j(k) & \text{for } -\kappa/2 < k < \kappa/2 \\ \sum_{j=1}^{J} \mathbf{D}_j(k)\mathbf{d}_j(k-\kappa) & \text{for } \kappa/2 < k < 3\kappa/2 \end{cases} \tag{3.1.10}$$

Equation (3.1.10) describes DMO (application of $\mathbf{D}_j$) and stacking ($\sum_j$), using the groundless assumption

$$\mathbf{d}(k-\kappa) = \mathbf{d}(k) = \mathbf{d}(k+\kappa) . \tag{3.1.11}$$

The replication of the Fourier transform of the data in equation (3.1.11) is equivalent to adding zero traces between data traces in the space domain. $\mathbf{G}^*$ is therefore DMO stacking using zero traces in place of missing traces.

It is easy to see that $\mathbf{G}^*$ is not the inverse of $\mathbf{G}$ by looking at the matrix $\mathbf{G}^*\mathbf{G}$ (Figure 3.1.2). $\mathbf{G}^*\mathbf{G}(k)$ is not the identity matrix: it includes off-diagonal elements that are as large as the main diagonal.

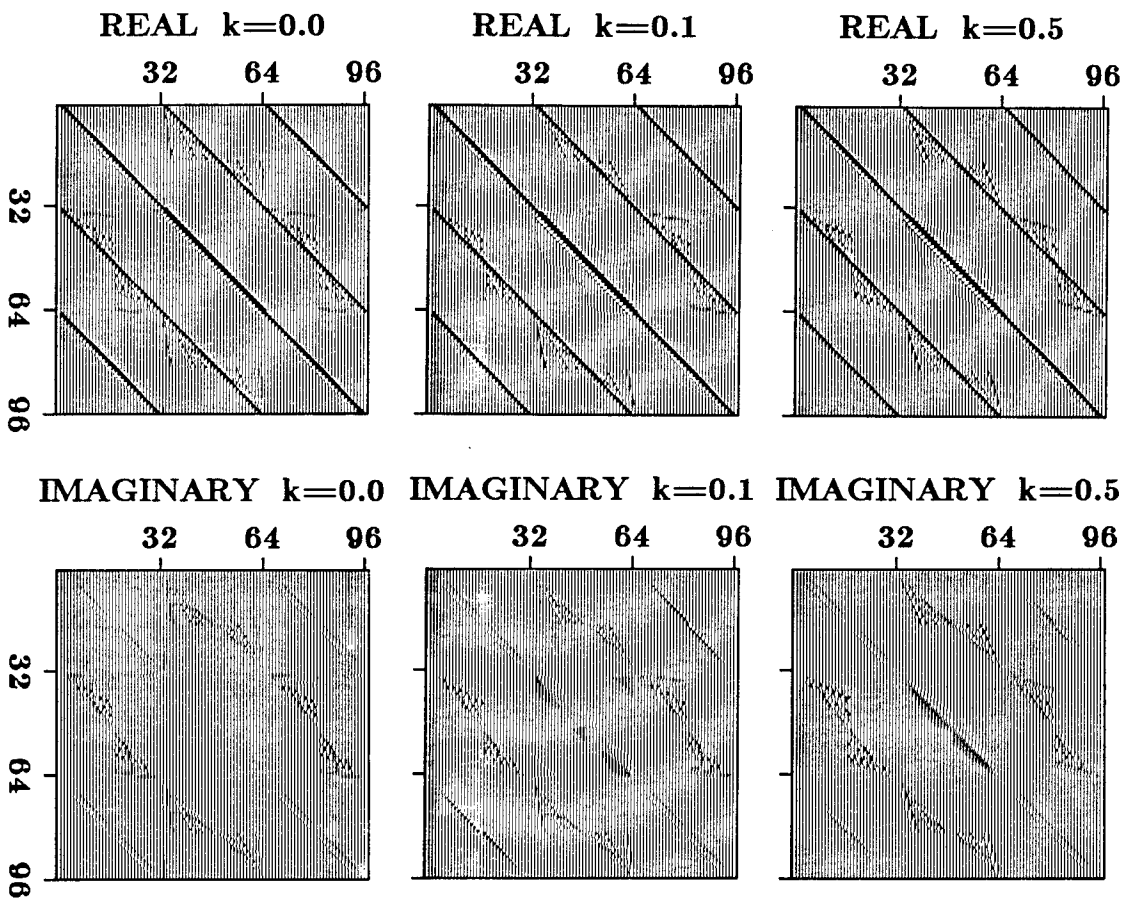


FIG. 3.1.2. $\mathbf{G}^*\mathbf{G}$ for various k's. They are approximately sparse but they are not the identity matrix.

§ 3.2 Missing data = zero data ?

Zero data are of course not the same as missing data, but when an inversion problem is "solved" by transposition, it is of no consequence to assume that the two are equivalent. For example consider an inversion $\mathbf{d} = \mathbf{Gm}$:

$$\begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \\ g_{41} & g_{42} & g_{43} \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} . \tag{3.2.1}$$

If d_2 is missing we are left with

$$\begin{pmatrix} d_1 \\ d_3 \\ d_4 \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{31} & g_{32} & g_{33} \\ g_{41} & g_{42} & g_{43} \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} , \tag{3.2.2}$$

which is not the same as,

$$\begin{pmatrix} d_1 \\ 0 \\ d_3 \\ d_4 \end{pmatrix} = \begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \\ g_{41} & g_{42} & g_{43} \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} . \tag{3.2.3}$$

However, application of the transpose of the matrix in equation (3.2.2)

$$\begin{pmatrix} \hat{m}_1 \\ \hat{m}_2 \\ \hat{m}_3 \end{pmatrix} = \begin{pmatrix} g_{11} & g_{31} & g_{41} \\ g_{12} & g_{32} & g_{42} \\ g_{13} & g_{33} & g_{43} \end{pmatrix} \begin{pmatrix} d_1 \\ d_3 \\ d_4 \end{pmatrix} , \tag{3.2.4}$$

would give the same result as would application of the transpose of the matrix in equation (3.2.3)

$$\begin{pmatrix} \hat{m}_1 \\ \hat{m}_2 \\ \hat{m}_3 \end{pmatrix} = \begin{pmatrix} g_{11} & g_{21} & g_{31} & g_{41} \\ g_{12} & g_{22} & g_{32} & g_{42} \\ g_{13} & g_{23} & g_{33} & g_{43} \end{pmatrix} \begin{pmatrix} d_1 \\ 0 \\ d_3 \\ d_4 \end{pmatrix} . \tag{3.2.5}$$

We can use the same operator $\mathbf{G}^*$, whether data are missing or not if we set all missing data to zero.

§ 3.3 Generalized inversion

The inversion $\mathbf{d} = \mathbf{Gm}$ can be solved by minimizing the norm of the error vector

$$\mathbf{e} = \mathbf{d} - \mathbf{Gm} \ . \tag{3.3.1}$$

Least squares would minimize the L_2 norm of $\mathbf{e}$,

$$E = \mathbf{e}^* \mathbf{e} = \left(\mathbf{d}^* - \mathbf{m}^* \mathbf{G}^*\right)\left(\mathbf{d} - \mathbf{Gm}\right) \ . \tag{3.3.2}$$

The minimum is where

$$0 = \frac{1}{2}\nabla_{\mathbf{m}} E = -\mathbf{G}^* \mathbf{d} + \mathbf{G}^* \mathbf{Gm} \ , \tag{3.3.3}$$

so

$$\mathbf{m} = \left(\mathbf{G}^* \mathbf{G}\right)^{-1} \mathbf{G}^* \mathbf{d} \ . \tag{3.3.4}$$

If $\mathbf{G}^* \mathbf{G}$ is singular, we can use an a priori model $\mathbf{m}_0$, and a positive-definite matrix $\mathbf{C}$, to minimize

$$E = \mathbf{e}^* \mathbf{e} + \left(\mathbf{m}-\mathbf{m}_0\right)^* \mathbf{C}\left(\mathbf{m}-\mathbf{m}_0\right) \ . \tag{3.3.5}$$

Minimizing E with respect to $\mathbf{m}$ we get

$$0 = \frac{1}{2}\nabla_{\mathbf{m}} E = -\mathbf{G}^* \mathbf{d} + \mathbf{G}^* \mathbf{Gm} + \mathbf{C}\left(\mathbf{m}-\mathbf{m}_0\right) \ , \tag{3.3.6}$$

so

$$\mathbf{m} = \left(\mathbf{G}^* \mathbf{G} + \mathbf{C}\right)^{-1}\left(\mathbf{G}^* \mathbf{d} + \mathbf{Cm}_0\right) \ . \tag{3.3.7}$$

The matrix $\mathbf{G}^* \mathbf{G}+\mathbf{C}$ cannot be singular because $\mathbf{C}$ is positive definite and $\mathbf{G}^* \mathbf{G}$ is nonnegative definite. The use of the matrix $\mathbf{C}$ is called damping; it is important when $\mathbf{G}^* \mathbf{G}$ is singular or ill conditioned.

Equation (3.3.7) can be applied in two steps:

(1) Prestack partial migration and stacking:

$$\hat{\mathbf{m}} = \mathbf{G}^* \mathbf{d} \ . \tag{3.3.8}$$

$\mathbf{G}^*$ is DMO with zero traces in place of missing traces in the input to DMO.

(2) Poststack multichannel inversion: solving

$$\hat{\mathbf{m}}+\mathbf{Cm}_0 = \left(\mathbf{G}^* \mathbf{G}+\mathbf{C}\right)\mathbf{m} \ . \tag{3.3.9}$$

The approximate sparsity of $\mathbf{G}^* \mathbf{G}$, as shown in Figure 3.1.2, may be useful in a poststack multichannel inversion.

§ 3.4 Synthetic data example

A synthetic zero-offset section with reflections from a 90 ° dipping reflector and from a flat reflector is shown in Figures 3.4.1a and 3.4.2a. The midpoint interval is 25 m. The synthetic data (Figure 3.4.1b) are four CMP gathers, 200 m apart. Each gather has five traces: zero offset to 400 m half-offset.

The result of the application of $\mathbf{G}^*$ is shown in Figures 3.4.1c and 3.4.2b. The interpolation is done with $N=8$, reducing the midpoint interval from 200 to 25 m. Poststack correction $\left(\mathbf{G}^* \mathbf{G}+\mathbf{C}\right)^{-1}$, with the section in 3.4.1c as input, generated the section shown in Figures 3.4.1d and 3.4.2c. The improvement resulting from the multichannel inversion is substantial, in particular for the flat reflector. The dipping reflector is relatively well interpolated by DMO alone. The V patterns were explained by Rocca and Ronen (1984).

§ 3.5 The design of three-dimensional surveys

The operator $\mathbf{G}$ is determined by the recording geometry. We have seen that if there is no aliasing then $\mathbf{G}$ is unitary. Achieving no aliasing in three dimensions may be impractical, but we can at least achieve a nonsingular $\mathbf{G}$.

Assume that a (marine) 3-D survey is made by performing many 2-D surveys in parallel. The line spacing Δy, is too big, so there is a spatial-aliasing problem in the cross-line direction. The in-line midpoint interval Δx, is small enough to avoid in-line aliasing: the summation over n_x in equation (2.14.2) reduces to

$$\mathbf{d_h}(k_x, k_y) = \sum_{n_y} \mathbf{D_h^+}(k_x, k_y - n_y \kappa_y)\mathbf{m}(k_x, k_y - n_y \kappa_y) \quad , \tag{3.5.1}$$

with $n_x = 0$ inside A in equation (2.13.3):

$$A^2 = 1 + \left(\frac{h_x k_x + h_y k_y - h_y n_y \kappa_y}{\omega t} \right)^2 . \tag{3.5.2}$$

If in each of the 2-D lines in our 3-D survey all shots and receivers are on the same line (x axis), then the offset vectors are always in the x direction, $h_y = 0$ in equation (3.5.2), and A becomes

$$A^2 = 1 + \left(\frac{h_x k_x}{\omega t} \right)^2 . \tag{3.5.3}$$

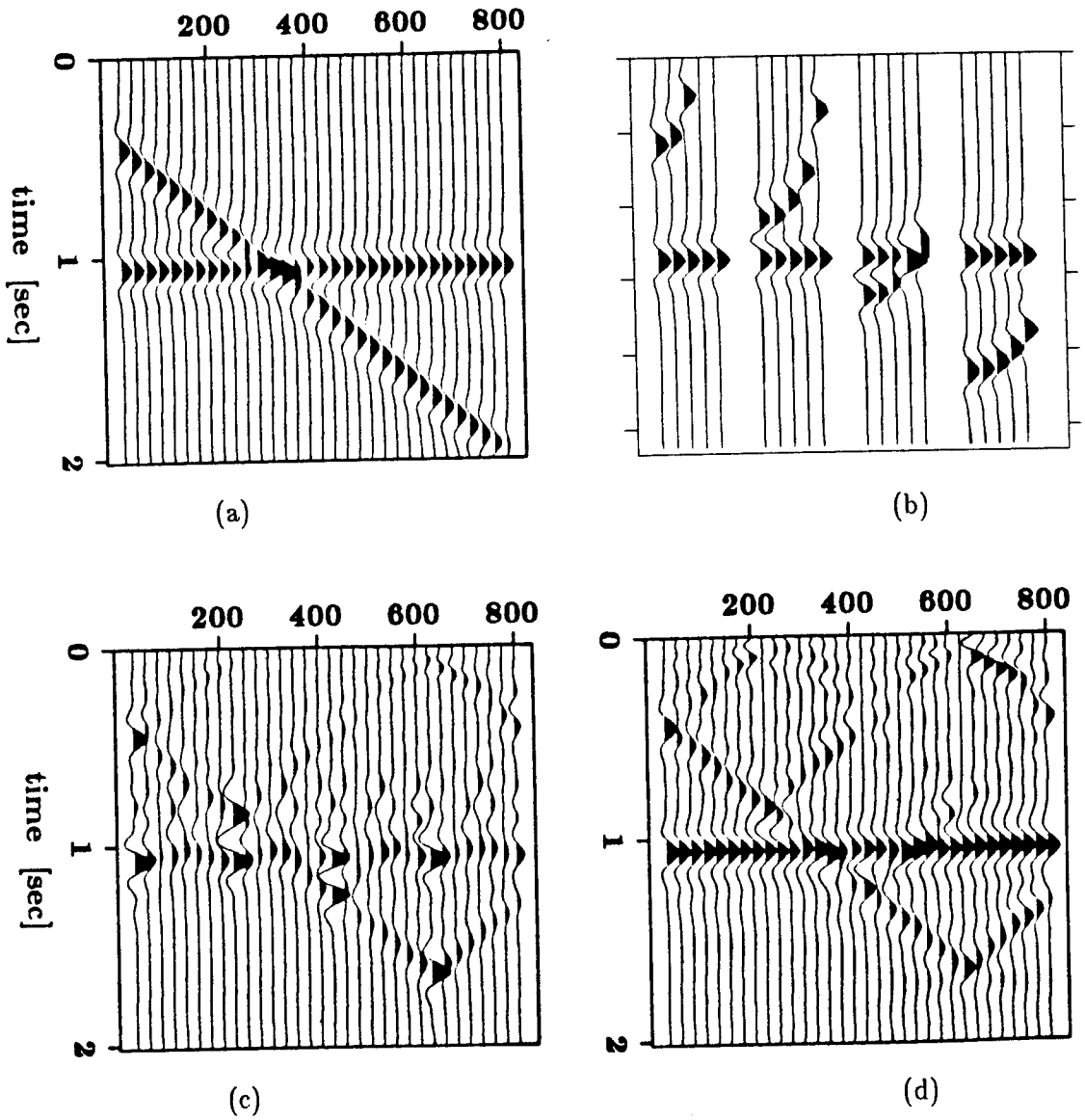


FIG. 3.4.1. Synthetic example. (a) The true model: a well-sampled zero-offset section. (b) The data: four gathers after NMO (the dipping reflector is over-corrected by the NMO), 200 m midpoint interval. (c) Transpose processing (the result of DMO stacking after addition of zero traces). (d) The result of multichannel inversion.

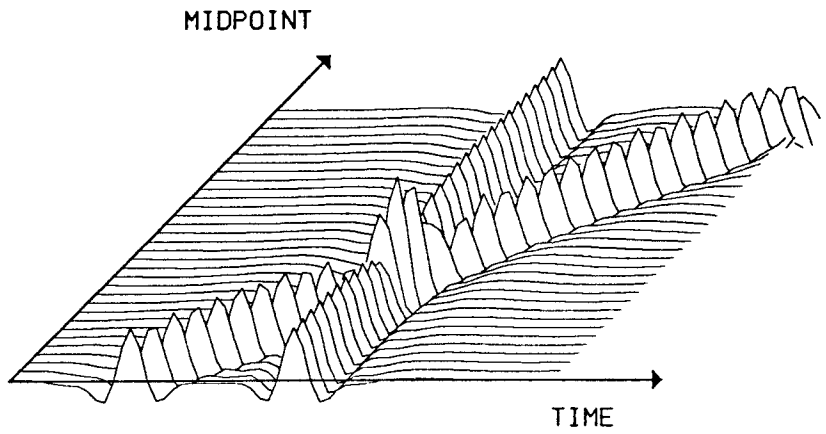
A and consequently $\mathbf{D_h}^+$ are independent of n_y:

$$\mathbf{D_h}^+(k_x, k_y - n_y \kappa_y) = \mathbf{D_h}^+(k_x, k_y) , \tag{3.5.4}$$

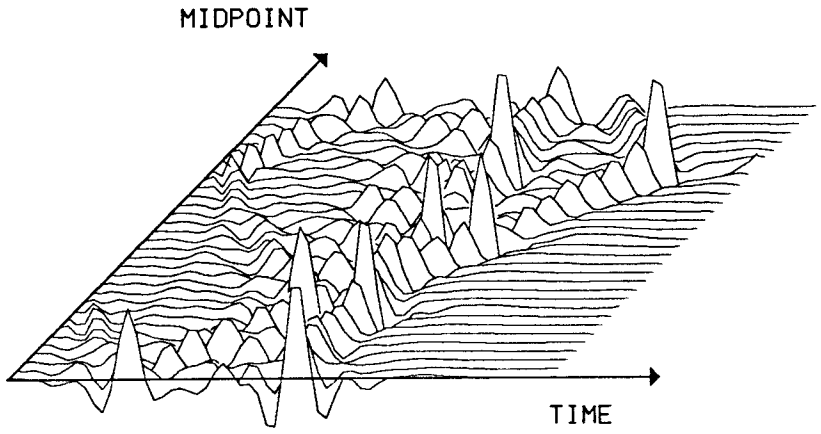
they can be moved out of the summation in equation (3.5.1):

$$\mathbf{d_h}(k_x, k_y) = \mathbf{D_h}^+(k_x, k_y) \sum_{n_y} \mathbf{m}(k_x, k_y - n_y \kappa_y) \quad . \tag{3.5.5}$$

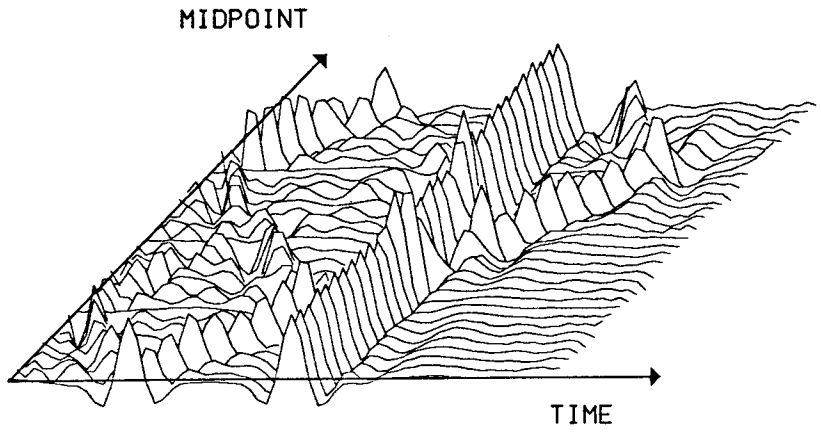
FIG. 3.4.2. Synthetic example (hidden line plots).



(a) The model: a well-sampled, zero-offset section.



(b) The result of transposition $\mathbf{G}^* \mathbf{d}$.



(c) The result of inversion $\left(\mathbf{G}^* \mathbf{G}+\mathbf{C}\right)^{-1} \mathbf{G}^* \mathbf{d}$.

In a matrix form,

$$\begin{pmatrix} \mathbf{d}_{\mathbf{h}_1}(k_x,k_y) \\ \cdot \\ \mathbf{d}_{\mathbf{h}_j}(k_x,k_y) \\ \cdot \\ \cdot \end{pmatrix} = \begin{pmatrix} \mathbf{D}_{\mathbf{h}_1}(k_x,k_y) & \mathbf{D}_{\mathbf{h}_1}(k_x,k_y) & \mathbf{D}_{\mathbf{h}_1}(k_x,k_y) \\ \cdot & \cdot & \cdot \\ \mathbf{D}_{\mathbf{h}_j}(k_x,k_y) & \mathbf{D}_{\mathbf{h}_j}(k_x,k_y) & \mathbf{D}_{\mathbf{h}_j}(k_x,k_y) \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} \mathbf{m}(k_x,k_y-\kappa_y) \\ \mathbf{m}(k_x,k_y) \\ \mathbf{m}(k_x,k_y+\kappa_y) \end{pmatrix} \quad (3.5.6)$$

The block columns of the matrix are all the same, and the system (3.5.6) is ill conditioned.

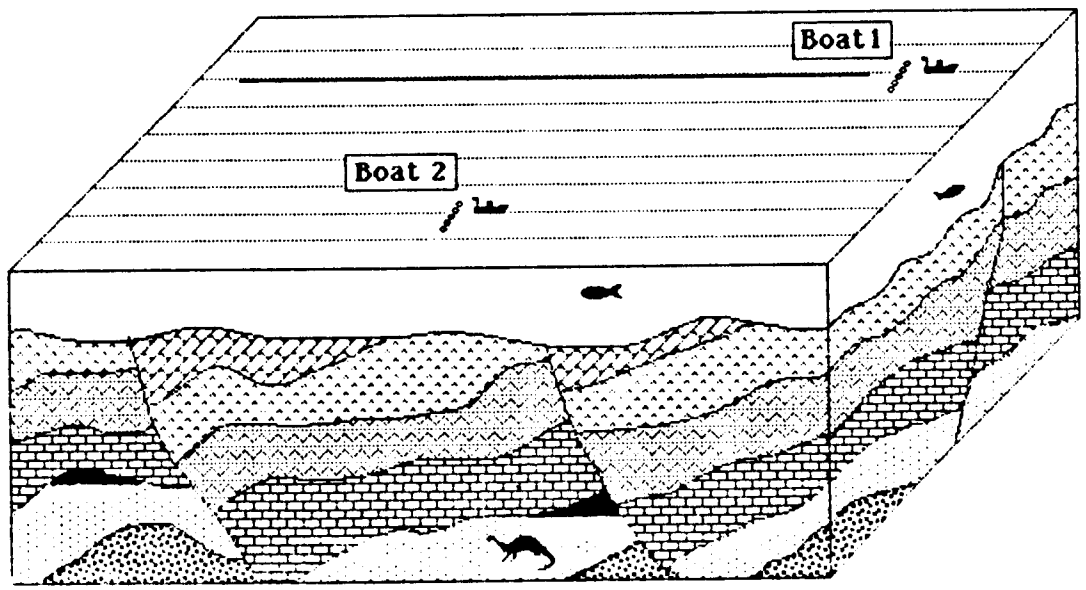


FIG. 3.5.1. Two-boat recording geometry: one cable, alternating shots; the line spacing is one fifth the distance between the boats.

Cross-line offset can be obtained by the recording geometry shown in Figure 3.5.1: two boats in parallel shooting alternatively; only one boat records. Boat 1 produces in-line offsets: h_y =0, and h_x varies from the near to the far half-offset. Boat 2 produces the cross-line offsets: h_y =500 m (if the distance between the boats is 1000 m), and h_x varies along the cable. The cross-line offset is some odd number times the line spacing (five in Figure 3.5.1), so the cross-line midpoints fall between recording lines.

The matrix relation for this experiment is

$$\begin{pmatrix} \mathbf{d}_{\mathbf{h}_1}(\mathbf{k}) \\ \cdot \\ \cdot \\ \mathbf{d}_{\mathbf{h}_{J/2}}(\mathbf{k}) \\ \hline \mathbf{d}_{\mathbf{h}_{J/2+1}}(\mathbf{k}) \\ \cdot \\ \cdot \\ \mathbf{d}_{\mathbf{h}_J}(\mathbf{k}) \end{pmatrix} = \begin{pmatrix} \mathbf{D}_{\mathbf{h}_1}(k_x, k_y) & \mathbf{D}_{\mathbf{h}_1}(k_x, k_y) & \mathbf{D}_{\mathbf{h}_1}(k_x, k_y) \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \mathbf{D}_{\mathbf{h}_{J/2}}(k_x, k_y) & \mathbf{D}_{\mathbf{h}_{J/2}}(k_x, k_y) & \mathbf{D}_{\mathbf{h}_{J/2}}(k_x, k_y) \\ \hline \mathbf{D}_{\mathbf{h}_{J/2+1}}(k_x, k_y - \kappa_y) & \mathbf{D}_{\mathbf{h}_{J/2+1}}(k_x, k_y) & \mathbf{D}_{\mathbf{h}_{J/2+1}}(k_x, k_y + \kappa_y) \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \mathbf{D}_{\mathbf{h}_J}(k_x, k_y - \kappa_y) & \mathbf{D}_{\mathbf{h}_J}(k_x, k_y) & \mathbf{D}_{\mathbf{h}_J}(k_x, k_y + \kappa_y) \end{pmatrix} \begin{pmatrix} \mathbf{m}(k_x, k_y - \kappa_y) \\ \mathbf{m}(k_x, k_y) \\ \mathbf{m}(k_x, k_y + \kappa_y) \end{pmatrix} . \tag{3.5.7}$$

Each block $\mathbf{D}_{\mathbf{h}_j}(k_x, k_y - n_y \kappa_y)$ depends on its n_y: n_y is 1 in the left column, 0 in the center, and -1 in the right column. The dependency is through A:

$$A^2 = 1 + \left(\frac{h_x k_x + h_y k_y - h_y n_y \kappa_y}{\omega t} \right)^2 . \tag{3.5.8}$$

The columns are independent because they have different n_y's, and $h_y \neq 0$ in some of their elements. The upper $J/2$ block rows in the matrix of equation (3.5.7) come from the in-line data, shot and recorded from boat 1; $h_y = 0$ in all of them, but h_x varies from row to row. The lower $J/2$ rows come from the cross-line data, shot from boat 2 and recorded from boat 1; $h_y = 500$ in all of them, and h_x varies from row to row. Without boat 2, we would not have the lower $J/2$ equations, and the matrix would be ill conditioned because its block columns would all be the same.

Additional independent channels can be obtained if each boat uses alternating multiple sources as shown in Figure 3.5.1: each boat tows a linear array of five sources, 12.5 m apart, extending to 25 m on each side of the boat. In every shot point, only one (out of five) fires. This generates a range of cross-line offsets. If the distance between boat 2 to the middle of the cable is 1000 m, the following cross-line offsets are recorded: -25,-12.5,0,12.5 and 25 m from boat 1; 975,987.5,1000,1012.5 and 1025 m from boat 2.

§ 3.6 3-D simulated by 2-D

If $k_x = 0$ then $\mathbf{h} \cdot \mathbf{k} = h_y k_y$; as if the survey were two-dimensional. The case $k_x = 0$ in equation (3.5.7) is important for reflectors that do not dip in the in-line direction (they may dip in the cross-line direction). Since h_x is always multiplied by k_x in equation (3.5.8), the upper $J/2$ block rows of the matrix of equation (3.5.7) are all the

same, (because they can differ only in their h_x). Similarly, the lower $J/2$ block rows are all the same. The condition of the matrix for the k_x=0 case is as if we had performed a 2-D experiment using only two offsets: zero and the cross-line offset.

The cross-line aliasing of the experiment illustrated in Figure 3.5.1, when five alternating sources per boat are used, can be simulated by taking a shot interval of 200 m, and ten offsets at -25,-12.5,0,12.5,25,975,987.5,1000,1012.5 and 1025 m on each shot profile.

§ 3.7 In and out of DMO space

Multichannel inversion, as defined in equation (2.6.1), requires inversion of matrices whose sizes are in the thousands. Since the cost of matrix inversion is proportional to the cube of the size of the matrix, it seems that multichannel inversion is impossible on today's computers. Fortunately, there are practical iterative techniques that produce good results, at a reasonable cost.

Conjugate-gradient inversion is an iterative method; each iteration involves the application of the forward operator, and the application of its transpose. Application of the transpose of the matrix of equation (2.6.2) is equivalent to DMO stacking (§3.1). Application of the forward matrix is transpose DMO stacking. The forward operator is going out of DMO space; its transpose, into DMO space.

Conjugate-gradient inversion of an $N \times N$ matrix converges within N iterations. Few iterations, however, often provide the solution. For multichannel inversion, I found that about four iterations were enough. Each iteration of the conjugate-gradient program involves performing DMO-stacking ($\mathbf{G}^*$) and transpose DMO-stacking ($\mathbf{G}$), so the total cost in four iterations is up to that of eight DMOs. Actually, the cost is less than that because Fourier transforms are not repeated, data are not resorted and operators are re-used.

Application of the matrix in equation (2.6.2) is still a costly operation, proportional to nt^2. Using this matrix is equivalent to using Hale's DMO method (1984). There are other DMO methods whose cost is proportional to $nt \log nt$ (DMO with the log transform [§2.8]; Shurtleff's method, [1984]; Fowler's method, [1984]), or even to nt (Bolondi et al, 1982; Salvador and Savelli, 1982; Jakubowicz, 1984). The alternative methods involve approximations, which might be perfectly acceptable in practice, but might complicate my study of multichannel inversion. I chose to use matrix multiplication but in practice more efficient methods can be used.

§ 3.8 Selective damping in conjugate-gradient inversion

The conjugate-gradient method of inversion is a popular one, described in many textbooks (for example Golub and Van Loan, 1983; Luenberger, 1984). I used a conjugate-gradient program written by Paige and Saunders (1982).

Damping is necessary when the matrix $\mathbf{G}$ is singular, and if many conjugate-gradient iterations are allowed. To incorporate damping into the inversion

$$\mathbf{d} = \mathbf{G}\mathbf{m} \,, \tag{3.8.1}$$

we can solve the system

$$\overline{\mathbf{d}} = \overline{\mathbf{G}}\mathbf{m} \,, \tag{3.8.2}$$

instead of (3.8.1). The modified data

$$\overline{\mathbf{d}} = \begin{pmatrix} \mathbf{d} \\ \vec{\mathbf{m}}_0 \end{pmatrix} , \tag{3.8.3}$$

include a modified a priori model $\vec{\mathbf{m}}_0$ And the matrix $\mathbf{G}$ is modified to

$$\overline{\mathbf{G}} = \begin{pmatrix} \mathbf{G} \\ \Lambda \end{pmatrix} , \tag{3.8.4}$$

where the matrix Λ is square and nonsingular. The solution to equation (3.8.2) is equivalent to the generalized inversion of equation (3.3.7) if $\Lambda^2 = \mathbf{C}$ and $\Lambda\vec{\mathbf{m}}_0 = \mathbf{C}\mathbf{m}_0$.

§ 3.9 Field-data example

A small part of a marine line from the Gulf of Mexico (Figure 1.1.3) is the database for an example of the effect of multichannel inversion. The original data do not have any aliasing problem: a shot interval of 25 m and 240 channels, 12.5 m apart, provide a 6.25 m midpoint interval on the 6000% stack. Spatial aliasing was introduced by using only a small part of the data in such a way that 3-D cross-line aliasing is simulated (§3.6). The 12.5 m group interval enabled simulation of the alternating multiple shots of Figure 3.5.1. I simulated five alternate sources, 12.5 m apart, with a line spacing of 150 m. The simulated distance between the boats was 750 m. Because data were not recorded at zero and negative offsets, I used five near traces: 221 to 271 m instead of -25 to +25 (and five far traces: 971 to 1021 m). A stacking chart of the data is shown in Figure 3.9.1. The data that were actually used in the inversion are shown in Figure 3.9.2. The coverage is not uniform, as it usually is in 3-D data: nonuniform coverage is natural for the data; it is the model, not the data, which should be uniformly and adequately sampled.

The models obtained by the first six iterations are shown in Figure 3.9.3. Multichannel inversion made a substantial improvement compared to DMO with zero traces in place of missing data (first iteration), but it still needs to be improved. The spatial spectra are clearly unbalanced (Figure 3.9.4): they are too strong in integer multiples of κ, while they should decay monotonically as the spatial frequency increases.

I applied a simple frequency balance program in the (t, k) domain: if the energy in the frequency k was higher than the energy in the previous frequency then the spectrum in k was divided by its envelope (AGC in time) and multiplied by the envelope of the next frequency. I used leaky integration of the absolute values to calculate envelopes. The balanced spectra and models are shown in Figures 3.9.5 and 3.9.6.

The final result (Figure 3.9.7d) is on the one hand much better than what is achieved by DMO stacking (Figure 3.9.7c), but on the other hand not precisely the true solution: the recorded near-offset section (Figure 3.9.7a): most noticeable are wrap around artifacts due to the Fourier-transform in space, and spurious events due to the truncation of the model because of the assumption that the model is zero in the evanescent zone.

§ 3.10 Summary

The application of multichannel inversion in reflection seismology is in both the acquisition and the processing of data.

The conclusion for acquisition is that the offset vectors in three-dimensional surveys should be in various directions.

The conclusion for data processing is that conventional processing (including prestack partial or full migration) is equivalent to the application of the transpose of the matrix we need to invert. This processing is adequate only when data are not missing; in the presence of spatial aliasing inversion gives much better results than transposition.

To apply the multichannel inversion, I implemented two alternative methods: a poststack process (on synthetic data) and an iterative, conjugate-gradient scheme (on field data). In both methods, the first step, or first iteration, is the conventional processing.

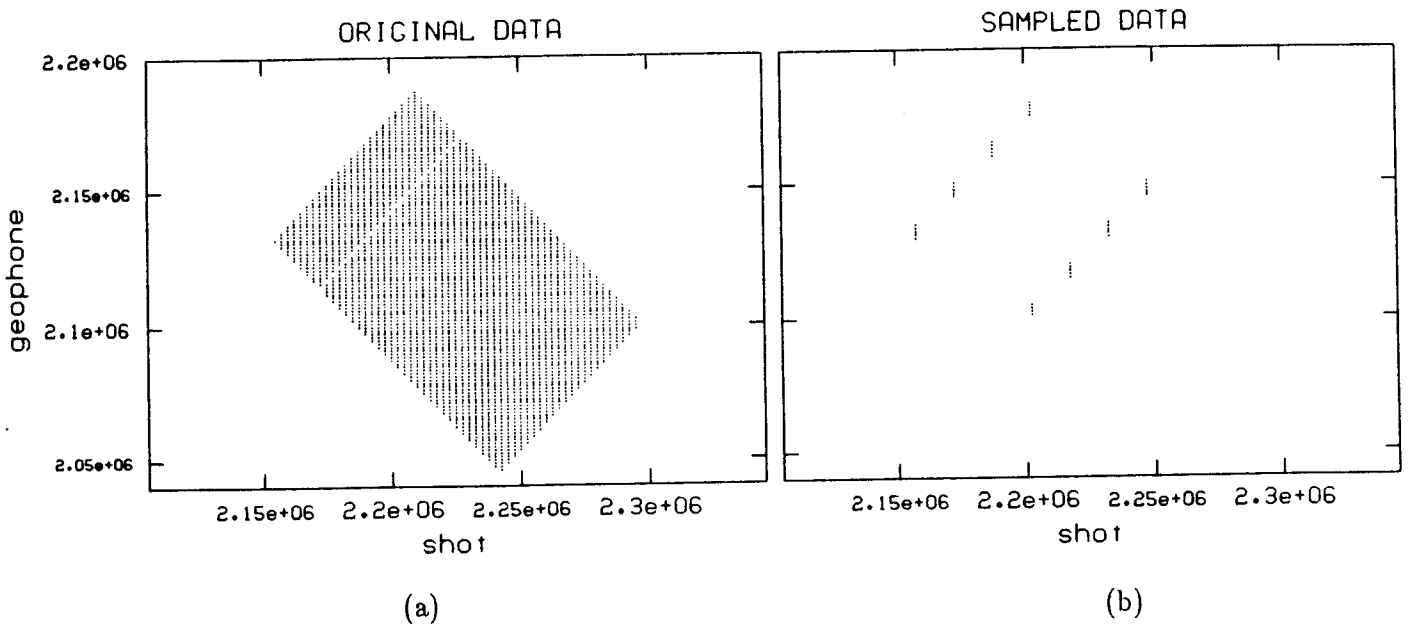


FIG. 3.9.1. (a) Stacking chart of the original data: 25 m shot interval, 12.5 m group interval (one offset is missing because of a dead receiver). (b) Stacking chart of the sampled data that simulated the cross-line direction in the experiment of Figure 3.5.1: 150 m shot interval, 10 channels per shot in two groups of 5, 750 m apart.

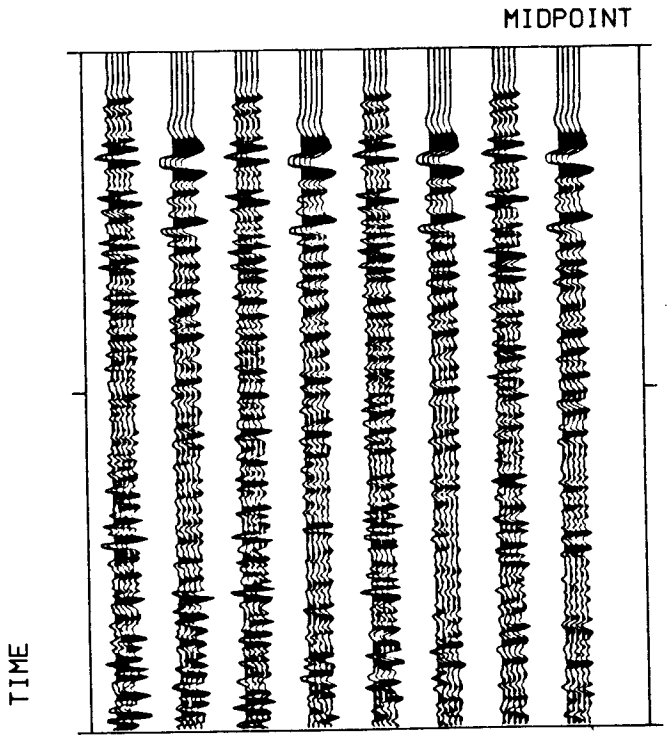


FIG. 3.9.2. The data. The coverage is not uniform; every trace was NMOed and plotted at its midpoint position.

FIG. 3.9.3. The models of the first six iterations and the true near-offset section. Convergence is achieved within four iterations.

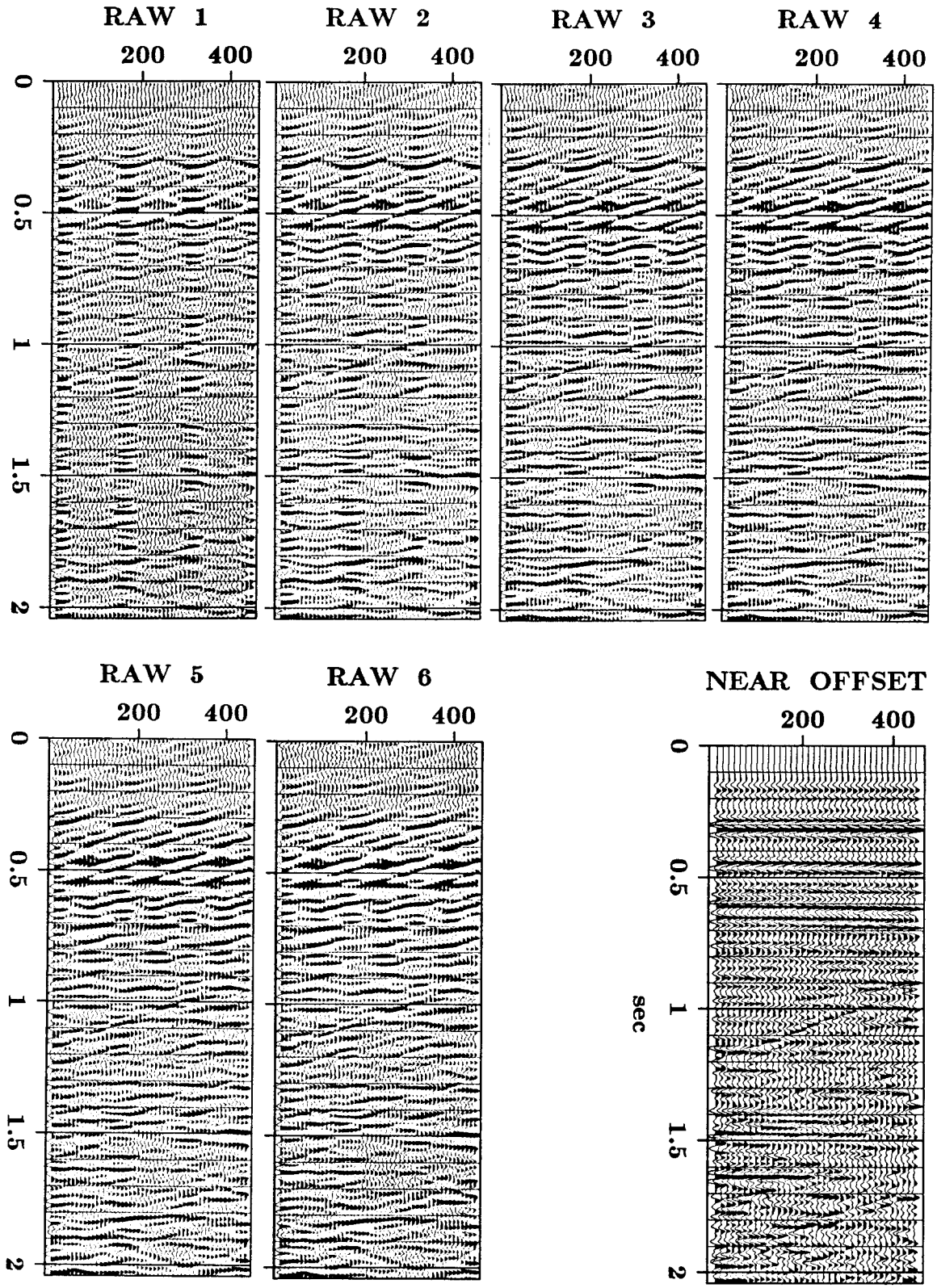


FIG. 3.9.4. Spatial spectra of the first six iterations and the true spectrum. They are clearly unbalanced.

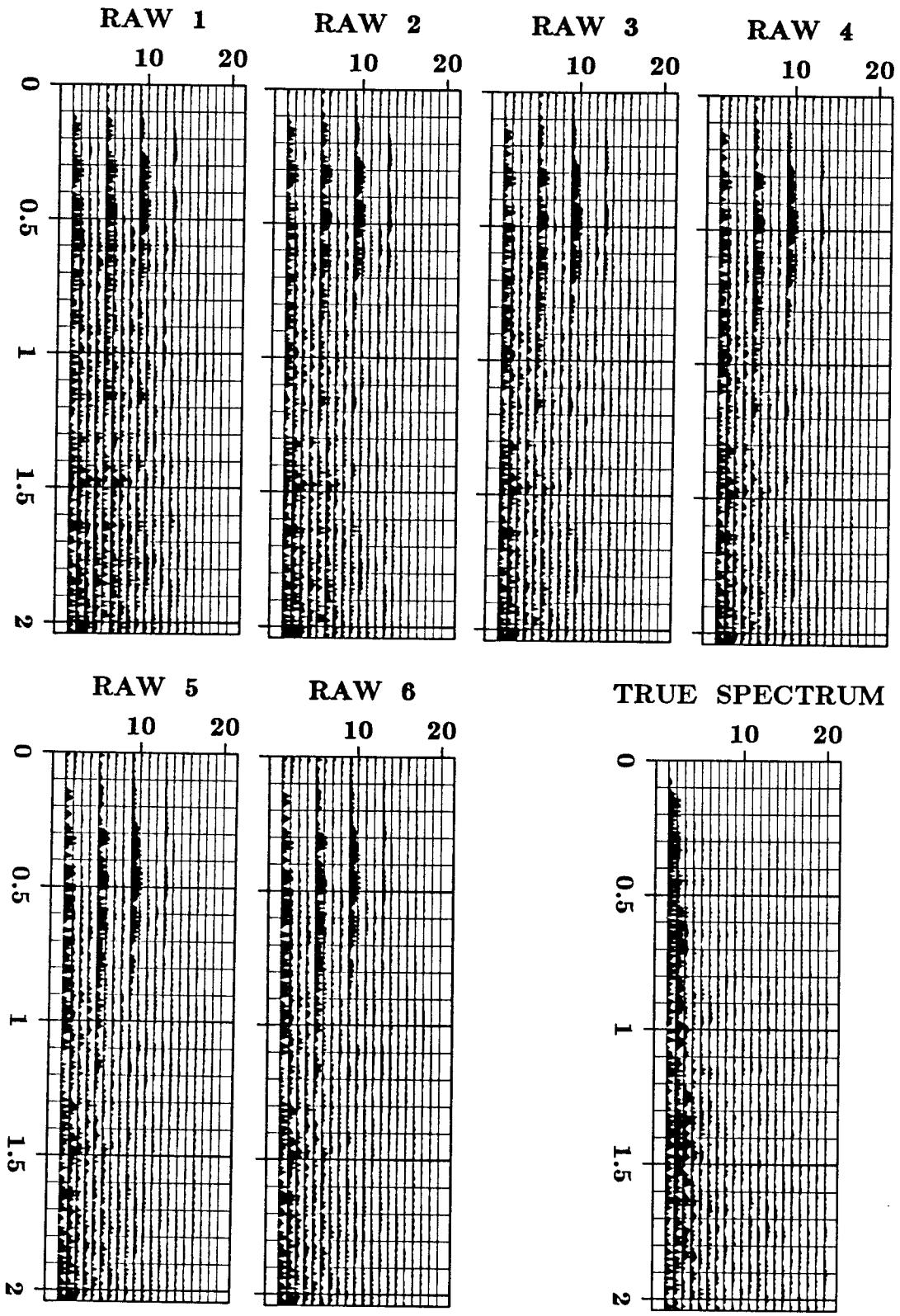


FIG. 3.9.5. Balanced spectra of the first six iterations and the true spectrum.

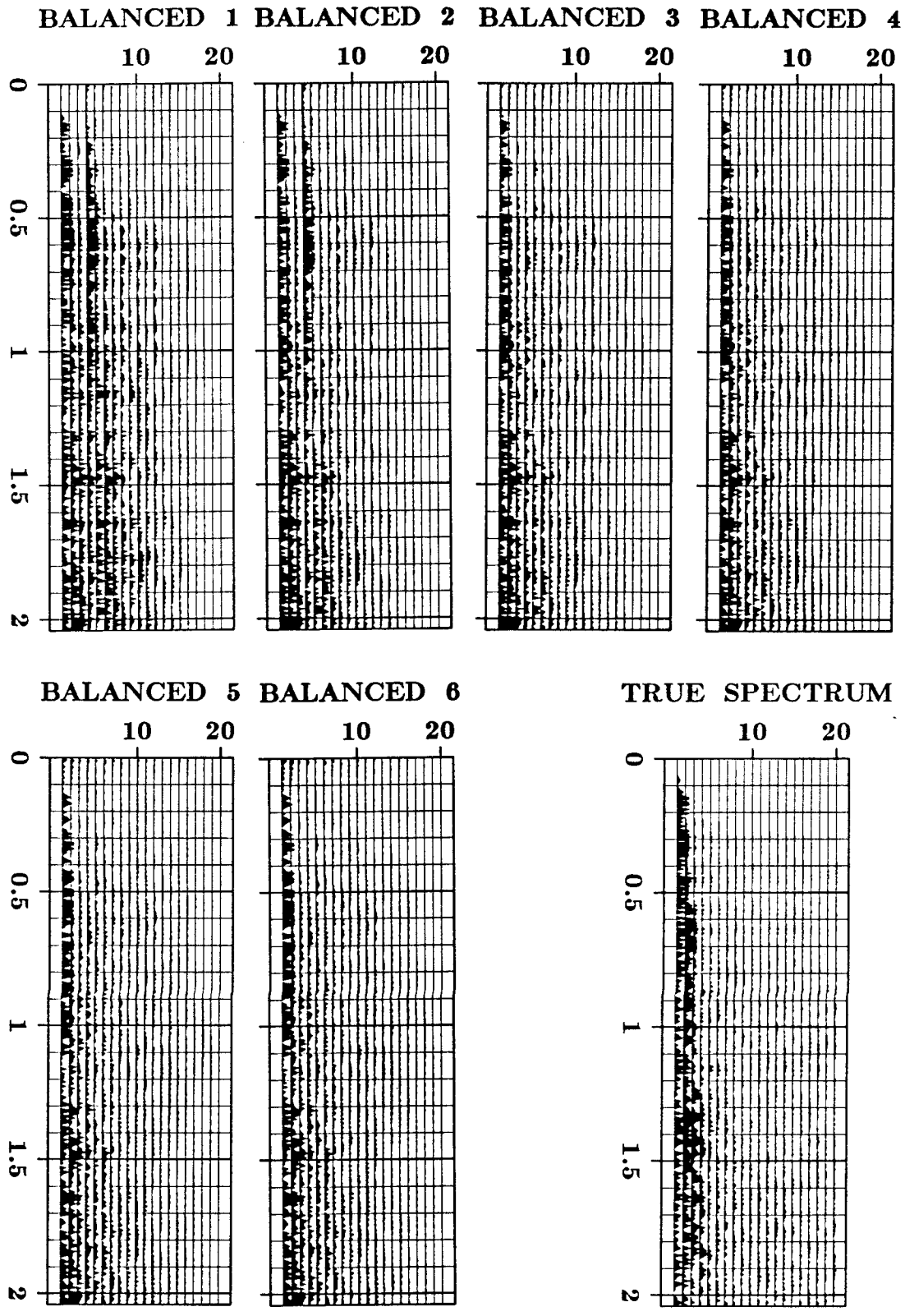
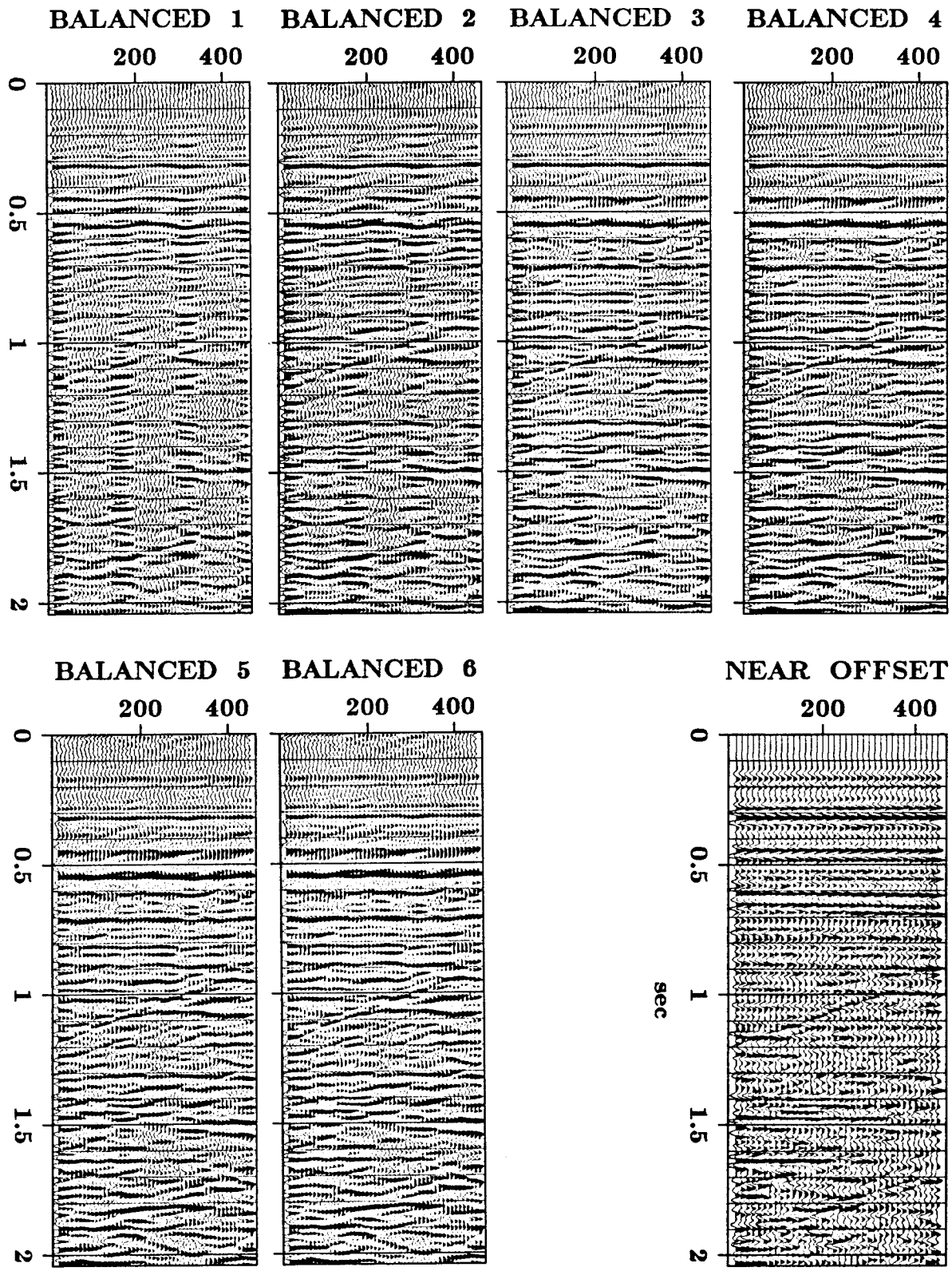


FIG. 3.9.6. The models of the first six iterations, with spectral balance and the true near-offset section.



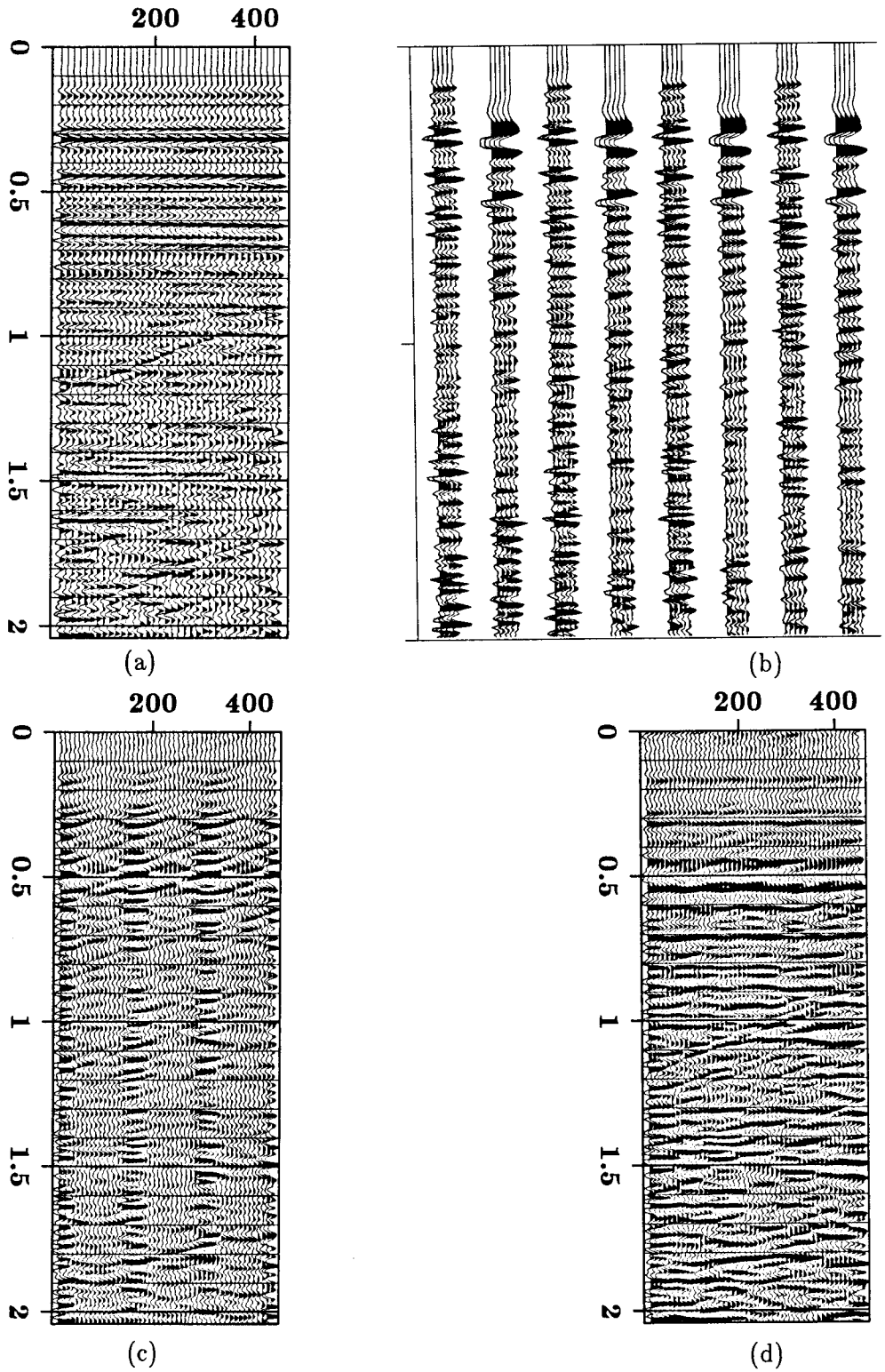


FIG. 3.9.7. Summary of the field data example. (a) Approximately the model: near-offset section with 25 m sampling interval. (b) The data: the coverage is not uniform; every trace is NMOed and plotted at its midpoint. (c) Unbalanced first iteration: the result of DMO stacking. (d) Balanced sixth iteration: flat and dipping reflectors are well interpolated.