Overcoming spatial aliasing in reflection seismology

Shuki Ronen

ABSTRACT

I propose a method of overcoming spatial aliasing in two steps:

(1) Pre-stack: Dip moveout with zero traces between data traces.

(2) Post-stack: A correction applied in the frequency domain.

The derivation of this method is based on pre-stack partial migration and on the Fourier analysis of sampling and aliasing.

INTRODUCTION

Spatial aliasing is one of the main problems of reflection seismology. Avoiding spatial aliasing by a dense recording geometry is costly, especially in three-dimensional surveys. So, before demanding more data in a dense recording, one would do better to use the redundancy in the multi-channel data he already has.

Interpolation schemes that attempt to overcome spatial aliasing usually make a priory assumptions, e.g. a limited number of dips, or a predetermined dominant dip azimuth. Predetermined assumptions are unreliable: the stacked section might have many crossing dips, the assumed dominant dip azimuth might be incorrect, etc. An intelligent interpolation program may find a model that fits the assumptions of the program's author or user, but would this be the real model?

Dip move-out (DMO) offers an opportunity for an interpolation, without unsafe assumptions: the DMO operator smears data between the shot and the receiver, as it sends the data from the CMP to the CDP. In 3-D surveys, this can be used to interpolate in the cross-line direction, using cross-line offsets. There are difficulties, however, with this simple approach; for example, the CDP coincides with the CMP for flat

reflectors, and the interpolation by DMO of flat reflectors, as shown by Rocca and Ronen (1984), is not as good as one might expect.

Interpolation methods rarely have difficulties with flat reflectors. To me, the difficulty the DMO interpolation had with flat reflectors was an indication that it was not the correct operation, and a motivation to examine the theory more carefully. I believe that in this paper I present the complete wave equation interpolator.

1. THE FORWARD PROBLEM

The object of this chapter is to formulate linear relations between seismic data and an earth model, using sampling theory and pre-stack partial migration.

1.1. Matrix form of aliasing

The function $m(x)$ has the Fourier transform,

$$m(x) \supset m(k) = \int_{-\infty}^{\infty} dx \; e^{i 2\pi k x} \; m(x) \; . \tag{1.1.1}$$

Suppose that $m(k)$ is band-limited,

$$m(k) = 0 \quad \text{for} \quad |k| \geq B \; . \tag{1.1.2}$$

The sampling theorem claims that if $B \leq 1/2$, then $m(x)$ can be fully restored from the sequence

$$d(x) = \text{III}(x) m(x) \; , \tag{1.1.3}$$

where the sampling symbol is defined by

$$\text{III}(x) = \sum_{n=-\infty}^{\infty} \delta(x-n) \; . \tag{1.1.4}$$

$\delta(x)$ is Dirac's impulse symbol. What is the relation between $d(k)$ and $m(k)$?

The sampling symbol Fourier transforms to itself,

$$\text{III}(x) \supset \text{III}(k) \; . \tag{1.1.5}$$

So, the multiplication in equation (1.1.3) transforms to the convolution,

$$d(x) \supset d(k) = \text{III}(k) \; * \; m(k) = \sum_{n=-\infty}^{\infty} m(k-n) \; . \tag{1.1.6}$$

If $B < 1/2$ then $m(k) = d(k)$ for $-1/2 < k < 1/2$. If $B > 1/2$, however, the replications, $m(k-n)$, overlap, as shown in Figure (1-1h), and $m(k) \neq d(k)$.

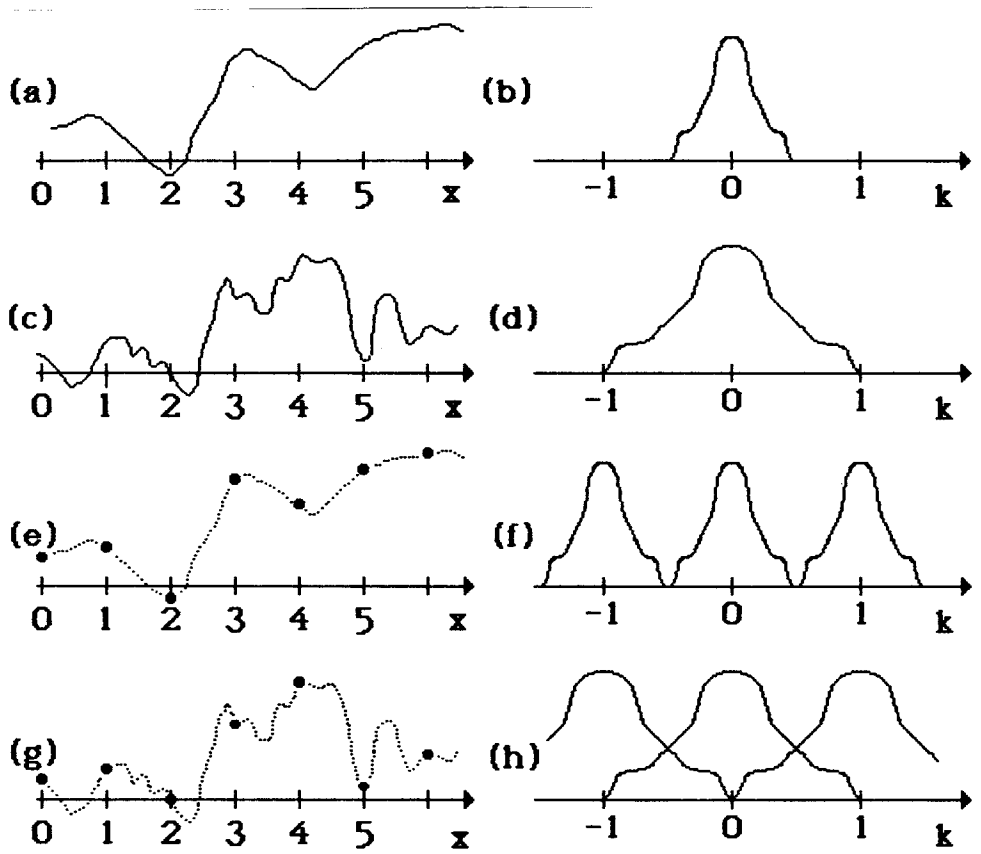


FIG. 1-1. Sampling and aliasing: (a) A function and (b) its schematic Fourier transform. (c) Another function and (d) its Fourier transform. (e) Samples from the function in (a) and (f) their corresponding Fourier transform: replication of the transform in (b); the sampling rate is adequate. (g) Samples from the function in (c) and (h) their corresponding Fourier transform: the sampling rate is not adequate: there is no way to recover the function in (c) from the sequence of samples in (g).

Define N as the smallest integer such that $N \geq 2B$. If N is odd, (Figure (1-2a)), then for $-1/2 < k < 1/2$,

$$d(k) = \sum_{n=-(N-1)/2}^{(N-1)/2} m(k-n) . \tag{1.1.7a}$$

If N is even, (Figure (1-2b)), then for $0 < k < 1$,

$$d(k) = \sum_{n=-N/2+1}^{N/2} m(k-n) . \tag{1.1.7b}$$

In any case, there are N contributing to $d(k)$ for every k. In a matrix form, for $N=3$,

$$d(k) = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} m(k-1) \\ m(k) \\ m(k+1) \end{pmatrix} , \tag{1.1.8}$$

which is an under-determined system: d gives us some information about m, but is not sufficient to determine m.

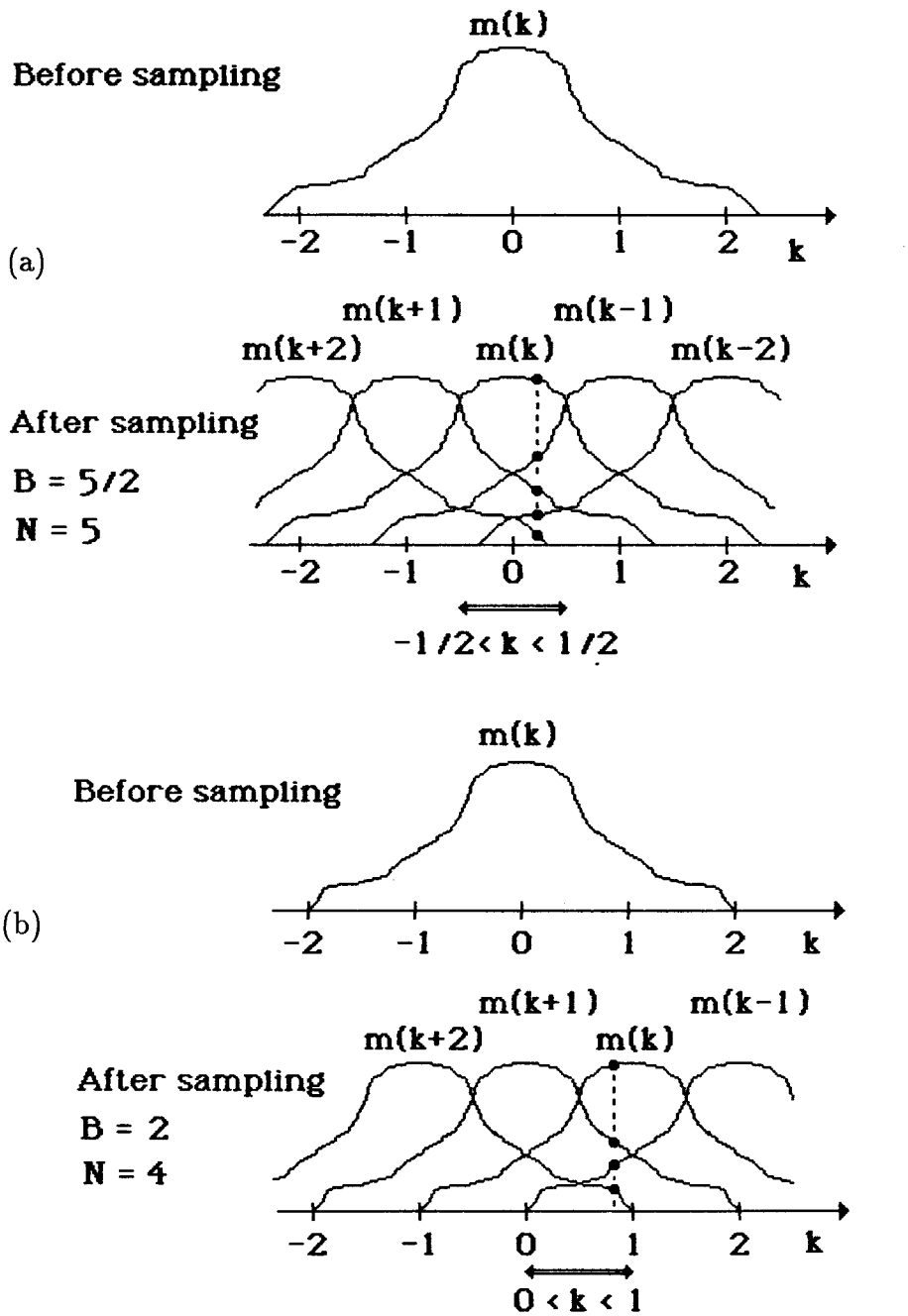


FIG. 1-2. Aliasing for odd (a) and for even (b) N.

1.2. Multi-channel scalar case

Suppose the function $m(x)$ is filtered prior to sampling by two convolution filters, h_1 and h_2, in parallel as shown in Figure (1-3). If one filter is half a unit delay operator,

$$h_1(x) = \delta(x - 1/2), \tag{1.2.1}$$

and the other is

$$h_2(x) = \delta(x), \tag{1.2.2}$$

then the data are

$$d_1(x) = \text{III}\,(x)m(x)\,, \tag{1.2.3}$$

and

$$d_2(x) = \text{III}\,(x)m(x - 1/2)\,, \tag{1.2.3}$$

and we can effectively double the sampling rate by intercalating the two sequences, d_1 and d_2; m can be reconstructed even if B is 1 and the analog to digital converters (A/D) sampling rate is 1 unit. (each channel is aliased with $N=2$). This is a trivial example of how multi-channel coverage can be used to overcome aliasing. Anther example is when the h_j are different differentiation operators (Linden, 1959). It is important that those h_j filters are different or else they will not contain more information than one channel (except the statistical improvement in signal to noise ratio).

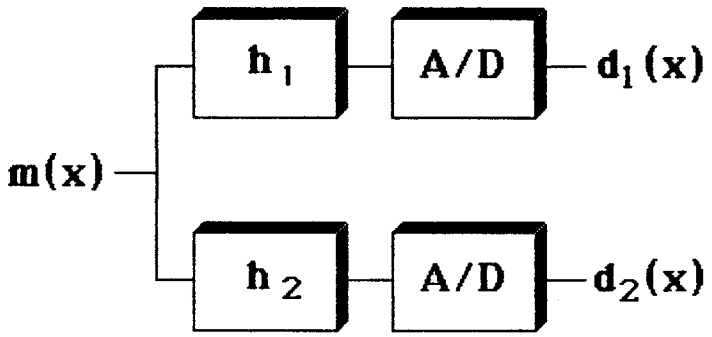


FIG. 1-3. Di-channel data, scalar case.

For two general linear-time-invariant filters h_1 and h_2, two sequences are available,

$$d_1(x) = \text{III}\,(x)\left[h_1(x) * m(x)\right] \supset d_1(k) = \sum_n h_1(k-n)m(k-n)\,, \qquad (1.2.4)$$

and

$$d_2(x) = \text{III}\,(x)\left[h_2(x) * m(x)\right] \supset d_2(k) = \sum_n h_2(k-n)m(k-n)\,. \qquad (1.2.5)$$

If $N = 2$ then, according to equation (1.1.7b),

$$d_1(k) = h_1(k)m(k) + h_1(k-1)m(k-1)\,, \qquad (1.2.6)$$

and

$$d_2(k) = h_2(k)m(k) + h_2(k-1)m(k-1)\,, \qquad (1.2.7)$$

or, in a matrix form,

$$\begin{pmatrix} d_1(k) \\ d_2(k) \end{pmatrix} = \begin{pmatrix} h_1(k) & h_1(k-1) \\ h_2(k) & h_2(k-1) \end{pmatrix} \begin{pmatrix} m(k) \\ m(k-1) \end{pmatrix} \qquad (1.2.8)$$

To find m we may be able to invert,

$$\begin{pmatrix} m(k) \\ m(k-1) \end{pmatrix} = \frac{1}{h_1(k)h_2(k-1) - h_1(k-1)h_2(k)} \begin{pmatrix} h_2(k-1) & -h_1(k-1) \\ -h_2(k) & h_1(k) \end{pmatrix} \begin{pmatrix} d_1(k) \\ d_2(k) \end{pmatrix} \qquad (1.2.9)$$

If there are J channels (Figure (1-4)) and for a general N. The system of equations is,

$$d_j(k) = \sum_n h_j(k-n)m(k-n) \qquad \text{for } j = 1,2,\cdots J\,, \qquad (1.2.10)$$

or, in a matrix form, for a 5×3 case,

$$\begin{pmatrix} d_1(k) \\ d_2(k) \\ d_3(k) \\ d_4(k) \\ d_J(k) \end{pmatrix} = \begin{pmatrix} h_1(k-1) & h_1(k) & h_1(k+1) \\ h_2(k-1) & h_2(k) & h_2(k+1) \\ h_3(k-1) & h_3(k) & h_3(k+1) \\ h_4(k-1) & h_4(k) & h_4(k+1) \\ h_J(k-1) & h_J(k) & h_J(k+1) \end{pmatrix} \begin{pmatrix} m(k-1) \\ m(k) \\ m(k+1) \end{pmatrix}\,. \qquad (1.2.11)$$

We may invert the J by N matrix to find $m(k)$. Note that the inversion extrapolates in the k direction: the data are given for $|k| < 1/2$, the model is found for $|k| < B = 3/2$, because $d(k-1) = d(k)$, but $m(k-1) \neq n(k)$ if $B > 1/2$.

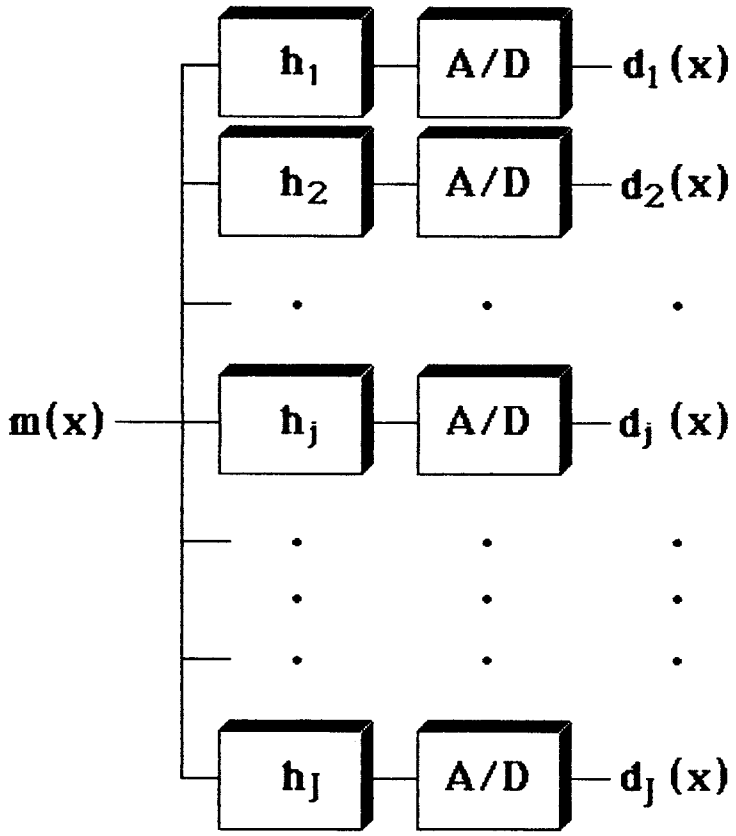


FIG. 1-4. J channels, scalar case.

1.3. Multi-channel seismic case without spatial aliasing

The goal of this section is to define the model $\mathbf{m}$ and the data $\mathbf{d}$, and to find $\mathbf{G}$, such that there is a linear relation

$$\mathbf{d} = \mathbf{G}\mathbf{m} \ . \tag{1.3.1}$$

One might choose the elasticity and density of the earth and the seismic wavelet as his model, and the field data as his data. The problem with this choice is that the relation between the data and the model is not linear: if the elastic coefficient is doubled, the data will change in a complicated way. In the context of spatial aliasing, I prefer to define the model and data in a way that the relation between them is linear.

I choose $\mathbf{d}$ to be the data after conventional pre-stack processing except DMO, (after gain deconvolution and NMO), and $\mathbf{m}$ to be the hypothetical zero- offset section we would record if we made an adequately sampled zero-offset experiment. As shown in

Figure (1-5), **d** is after NMO, and **m** is before migration. The purpose of this choice is to decouple the problem of velocity estimation from the problem of spatial aliasing: NMO and migration are strongly dependent on velocity, DMO is relatively velocity independent.

Each common-offset section, $\mathbf{d}_j$, in Figure (1-6), comes from the zero-offset section, **m**, via a $\mathbf{D}^+{}_j$ filter. It is shown in appendix A that the relation between the time-space Fourier transform of the zero-offset section $m(k,\omega)$ and the space Fourier transform of a common-offset section, $d_j(t,k)$ (at half offset h_j) is the inverse DMO

$$d_j(k,t) = \int d\omega A^{-1} e^{-i\omega A t}\, m(\omega,k)\,, \tag{1.3.2}$$

where, A is $\sqrt{1+[h_j k/\omega t]^2}$. Equation (1.3.2) is not a process I would apply, rather it is a physical process I would like to undo.

My model assumes the existence of an invertible operator which maps a nonzero-offset section to the zero-offset section. This assumption was made by Fabio Rocca and other authors who derived DMO as extrapolation in offset. Dave Hale made a less restrictive assumption that the zero-offset section is obtained from summing over <u>all</u> offsets, and not necessarily from each offset. The consistency and practical equivalency of these two assumptions are shown in appendix B.

A discrete form of in equation (1.3.2) is

$$d_j(k,t) = \sum_{\omega} A^{-1} e^{-i\omega A t}\, m(k,\omega)\,. \tag{1.3.3}$$

This is a matrix-vector multiplication

$$\mathbf{d}_j(k) = \mathbf{D}^+{}_j(k)\mathbf{m}(k)\,, \tag{1.3.4}$$

where the vectors are the common-offset section

$$\mathbf{d}_j(k) = \begin{pmatrix} d_j(k,1) \\ \cdot \\ d_j(k,t) \\ \cdot \\ d_j(k,nt) \end{pmatrix}, \tag{1.3.5}$$

and the zero-offset section

$$\mathbf{m}(k) = \begin{pmatrix} m(k,1) \\ \cdot \\ m(k,\omega) \\ \cdot \\ m(k,n\omega) \end{pmatrix}, \tag{1.3.6}$$

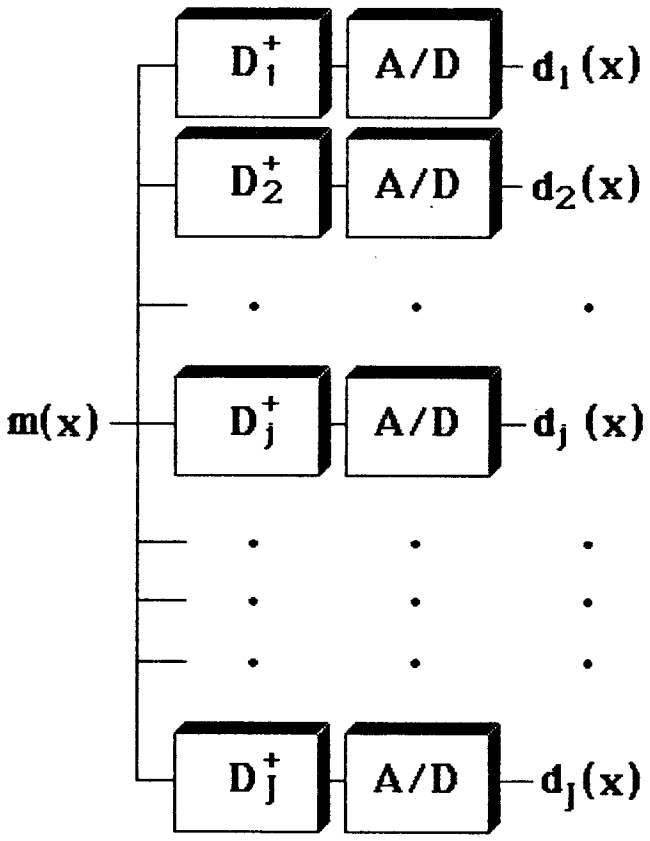


FIG. 1-5. A model for multi-channel seismic data.

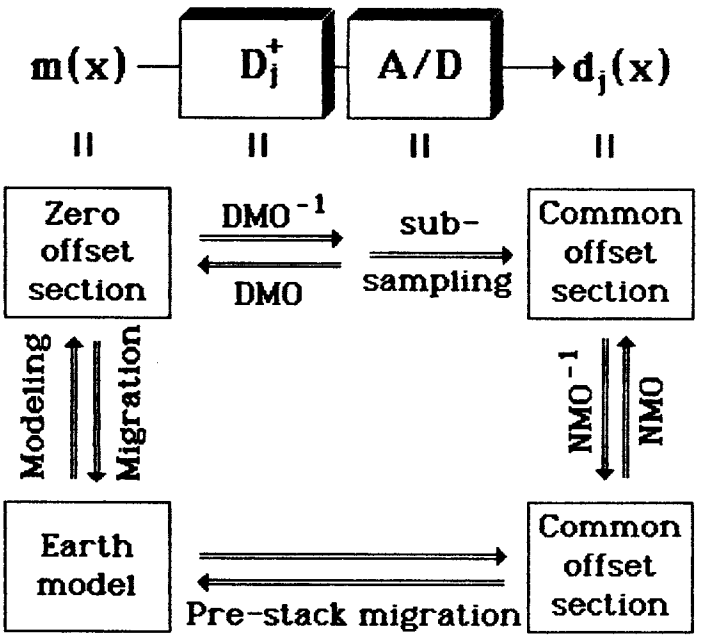


FIG. 1-6. **m** is the zero-offset section, each **d** sub j is a constant offset section after NMO. The $\mathbf{D}^+$ operators are inverse DMO.

and the matrix $\mathbf{D}^+$ is the inverse DMO operator: an $nt \times n\omega$ matrix, whose value in the t-th row and the ω-th column is

$$\left[\mathbf{D}^+{}_j(k)\right]_{\omega,t} = A^{-1}(h_j\, k/\omega t)\, e^{-i\omega A(h_j\, k/\omega t)t} \quad . \tag{1.3.7}$$

$\mathbf{D}$, the inverse of $\mathbf{D}^+$ is the DMO. (also known as pre-stack partial migration and offset continuation). The subject of DMO is enjoying increasing popularity since DEVILISH (Judson et al, 1978) and the works of Yilmaz (1980) and of Deregowski and Rocca (1982). Finite-differencing methods were developed first: DEVILISH, Yilmaz's method, Bolondi et. al. (1982), Salvador and Savelly, (1982). Then came the exact operators by slant-stacks/radial-trace, (Ottolini, 1982) and by Fourier-transform (Hale, 1983). Jakubowicz (1984) proposed a method in the dip domain, and Fowler (1984) is developing DMO in velocity space. In practice, many use integral summation methods.

1.4. Spatially aliased multi-channel seismic data

It's time to merge sampling theory and migration.

According to equations (1.1.7) and (1.3.4), a sampled common-offset section has the form,

$$\mathbf{d}_j(k) = \sum_n \mathbf{D}^+{}_j(k-n)\mathbf{m}(k-n) \quad . \tag{1.4.1}$$

For all the offsets, the full system is

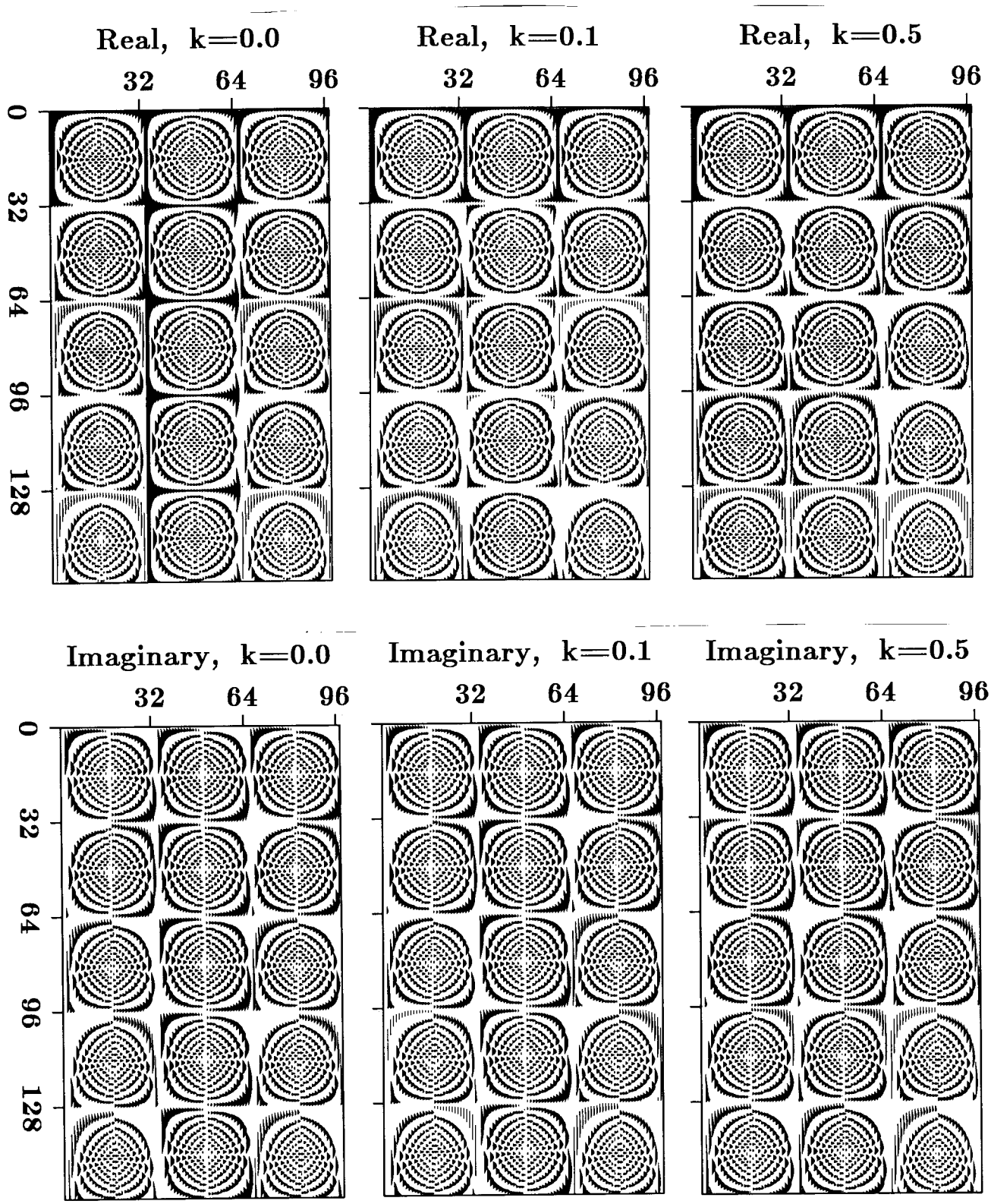
$$\begin{pmatrix} \mathbf{d}_1(k) \\ \mathbf{d}_2(k) \\ \mathbf{d}_3(k) \\ \mathbf{d}_4(k) \\ \mathbf{d}_J(k) \end{pmatrix} = \begin{pmatrix} \mathbf{D}^+{}_1(k-1) & \mathbf{D}^+{}_1(k) & \mathbf{D}^+{}_1(k+1) \\ \mathbf{D}^+{}_2(k-1) & \mathbf{D}^+{}_2(k) & \mathbf{D}^+{}_2(k+1) \\ \mathbf{D}^+{}_3(k-1) & \mathbf{D}^+{}_3(k) & \mathbf{D}^+{}_3(k+1) \\ \mathbf{D}^+{}_4(k-1) & \mathbf{D}^+{}_4(k) & \mathbf{D}^+{}_4(k+1) \\ \mathbf{D}^+{}_J(k-1) & \mathbf{D}^+{}_J(k) & \mathbf{D}^+{}_J(k+1) \end{pmatrix} \begin{pmatrix} \mathbf{m}(k-1) \\ \mathbf{m}(k) \\ \mathbf{m}(k+1) \end{pmatrix} . \tag{1.4.2}$$

The only difference between equation (1.4.2) and the scalar case of equation (1.2.11) is that here are vectors and matrices, $\mathbf{d}_j$, $\mathbf{m}$, and $\mathbf{D}^+{}_j$, while there were scalars, d_j, m, and h_j. The matrix of equation (1.4.2) is shown in Figure (1-7).

Equation (1.4.2) is a system of $J \times nt$ equations, (one equation for every time sample of every common-offset section), with $N \times n\omega$ unknowns, (N $m(k-n)$'s, each $n\omega$ long). The matrix has a J by N blocks of $\mathbf{D}^+$ sub-matrices, each $\mathbf{D}^+$ is $nt \times n\omega$ big. We want to solve this system for all $0 < k < 1/2$.

Solving equation (1.4.2) not only continues the data in the offset direction as DMO should do, it also continues the data in the wavenumber , k, direction from the low

FIG. 1-7. $\mathbf{G}(k)$ for various k's. $J=5$ and $N=3$. Each block is an inverse DMO operator $\mathbf{D}^{+}{}_{j}(k-n)$ of nt by $n\omega$ (here $nt=n\omega=32$).



frequencies $0 < k < 1/2$ on which the aliased data $\mathbf{d}_j$ are given, to the full range $0 < k < B$ required to describe the zero-offset section $\mathbf{m}$.

2. SEPARATE-CHANNELS INVERSION

Conventional processing does not involve inversion of any huge system of equations like (1.4.2). The object of this chapter is to see what conventional processing is, in terms of the formulation of the last section.

2.1. Conventional processing without DMO

When DMO is not performed, the NMO-stack is

$$\hat{\mathbf{m}}(k) = \sum_{j=1}^{J} \mathbf{d}_j(k) \tag{2.1.1}$$

$$= \sum_{j=1}^{J} \sum_{n} \mathbf{D}^+{}_j(k-n)\mathbf{m}(k-n) \ ,$$

according to equation (1.4.1).

Aliasing is not overcome, moreover, even without aliasing ($N=1$) the stack, $\hat{\mathbf{m}}$, is not the zero-offset section, $\mathbf{m}$,

$$\hat{\mathbf{m}}(k) = \left\{ \sum_{j=1}^{J} \mathbf{D}^+{}_j(k) \right\} \mathbf{m}(k) \ , \tag{2.1.2}$$

because of the filter $\sum_j \mathbf{D}^+{}_j$.

2.2. Conventional processing with DMO - no aliasing case

When DMO is done, it is applied separately to every common-offset section as shown in Figure (2-1). I will next show that DMO is the required operator if the data are not aliased.

If there is no spatial aliasing, then $N = 1$ in equation (1.4.2), and we have,

$$\begin{pmatrix} \mathbf{d}_1(k) \\ \mathbf{d}_2(k) \\ \mathbf{d}_3(k) \\ \mathbf{d}_4(k) \\ \mathbf{d}_J(k) \end{pmatrix} = \begin{pmatrix} \mathbf{D}^+{}_1(k) \\ \mathbf{D}^+{}_2(k) \\ \mathbf{D}^+{}_3(k) \\ \mathbf{D}^+{}_4(k) \\ \mathbf{D}^+{}_J(k) \end{pmatrix} \Big(\mathbf{m}(k) \Big) \ . \tag{2.2.1}$$

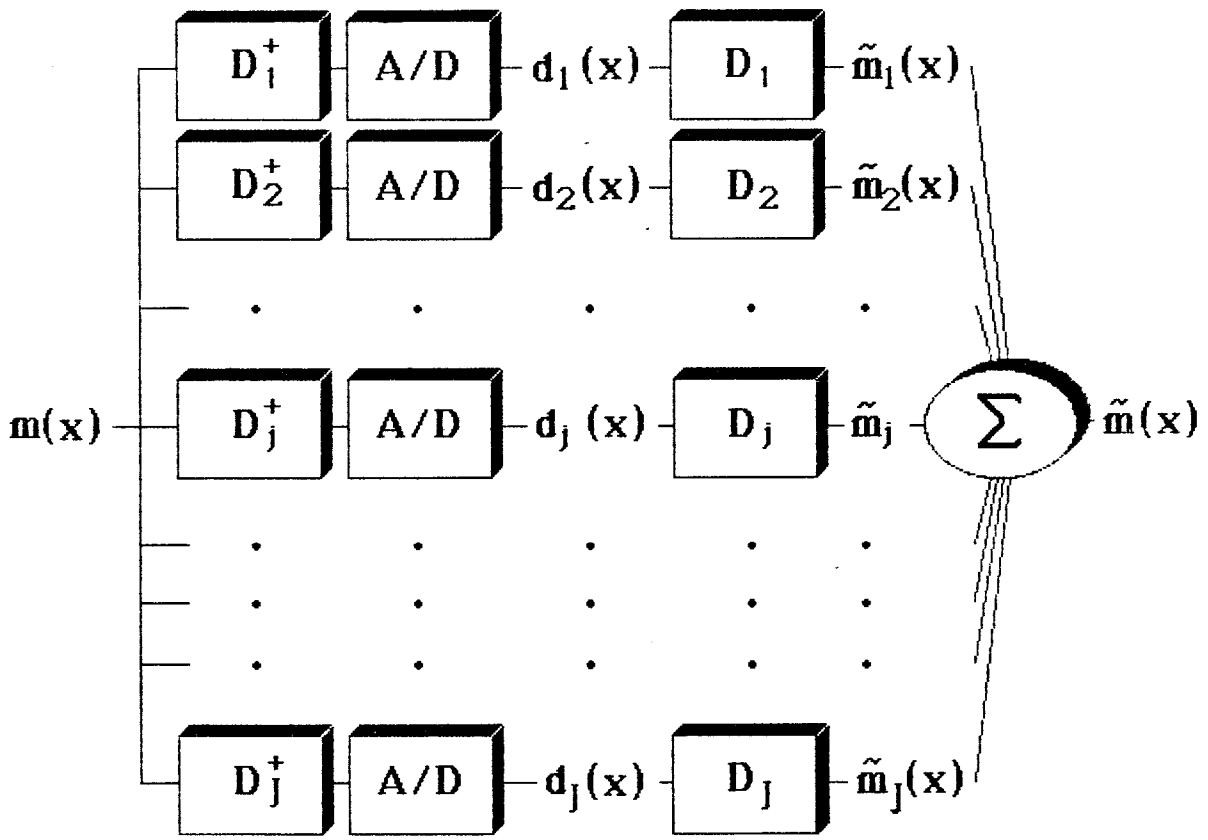


FIG. 2-1. Separate inversion of J channels.

$\mathbf{m}(k)$ is a $n\omega$ long vector, each $\mathbf{d}$' is a nt long vector, and the $\mathbf{D}^+$'s are matrices. Applying DMO is equivalent to multiplying equation (2.2.1) by the block diagonal matrix,

$$\begin{pmatrix} \mathbf{D}_1 & & & \\ & \mathbf{D}_2 & & 0 \\ & & \cdot & \\ & 0 & & \\ & & & \mathbf{D}_J \end{pmatrix}, \tag{2.2.2}$$

obtaining,

$$\begin{pmatrix} \mathbf{D}_1\mathbf{d}_1 \\ \mathbf{D}_2\mathbf{d}_2 \\ \cdot \\ \cdot \\ \mathbf{D}_J\mathbf{d}_J \end{pmatrix} = \begin{pmatrix} \mathbf{D}_1\mathbf{D}^+{}_1 \\ \mathbf{D}_2\mathbf{D}^+{}_2 \\ \cdot \\ \cdot \\ \mathbf{D}_J\mathbf{D}^+{}_J\mathbf{I} \end{pmatrix} \Big(\mathbf{m}(k)\Big) = \begin{pmatrix} \mathbf{I} \\ \mathbf{I} \\ \cdot \\ \cdot \\ \mathbf{I} \end{pmatrix} \Big(\mathbf{m}(k)\Big) . \tag{2.2.3}$$

The left-hand side is the data after NMO and DMO, the unknown is the zero-offset section. The identity matrices result from applying DMO after inverse DMO. This is an overdetermined system: each of the J processed common-offset sections, $\mathbf{D}_j \mathbf{d}_j$, should be equal to the same zero-offset section, $\mathbf{m}$. Solving equation (2.2.3) by least squares (minimizing the L_2 norm of the error) is equivalent to conventional stacking,

$$\mathbf{m} = \frac{1}{J} \sum_{j=1}^{J} \mathbf{D}^{+}{}_j \mathbf{d}_j \ . \tag{2.2.4}$$

2.3. Conventional processing with DMO - aliased case

Consider first the scalar di-channel example of Figure (1-3). Separate inversion as shown in Figure (2-2), inverts every channel separately,

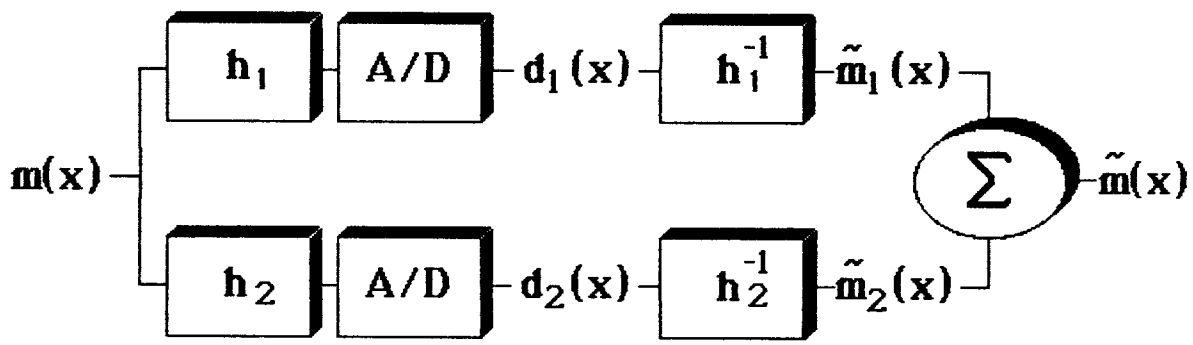


FIG. 2-2. Separate inversion of two channels.

$$\tilde{m}_1(k) = h_1^{-1}(k) d_1(k) \ ,$$

and,

$$\tilde{m}_2(k) = h_2^{-1}(k) d_2(k) \ , \tag{2.3.1}$$

then averages the results,

$$\tilde{m}(k) = \frac{1}{2} \begin{pmatrix} \mathrm{I} & \mathrm{I} \end{pmatrix} \begin{pmatrix} \tilde{m}_1(k) \\ \tilde{m}_2(k) \end{pmatrix}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} h_1^{-1}(k) & 0 \\ 0 & h_2^{-1}(k) \end{pmatrix} \begin{pmatrix} h_1(k) & h_1(k-1) \\ h_2(k) & h_2(k-1) \end{pmatrix} \begin{pmatrix} m(k) \\ m(k-1) \end{pmatrix}$$

$$= m(k) + \frac{1}{2}\left(\frac{h_1(k-1)}{h_1(k)} + \frac{h_2(k-1)}{h_2(k)} \right) m(k-1) \, . \tag{2.3.2}$$

The separate channel inversion of equation (2.3.2) is different from the inversion in equation (1.2.9). The difference between $\tilde{m}(k)$ and $m(k)$ is a term proportional to $m(k-1)$. This is aliasing noise: it vanishes when there is no aliasing. Its magnitude depends on the data-independent function

$$R(k) = \frac{1}{2}\left(\frac{h_1(k-1)}{h_1(k)} + \frac{h_2(k-1)}{h_2(k)} \right) . \tag{2.3.3}$$

In the seismic case, each common-offset section section after DMO is

$$\mathbf{D}_j(k)\mathbf{d}_j(k) = \mathbf{D}_j(k)\sum_n \mathbf{D}^+{}_j(k-n)\mathbf{m}(k-n) \tag{2.3.4}$$

$$= \sum_n \left\{ \mathbf{D}_j(k)\mathbf{D}^+{}_j(k-n) \right\} \mathbf{m}(k-n) \, .$$

stacking after DMO gives

$$\sum_j \mathbf{D}_j(k)\mathbf{d}_j(k) = \sum_j \sum_n \left\{ \mathbf{D}_j(k)\mathbf{D}^+{}_j(k-n) \right\} \mathbf{m}(k-n) \tag{2.3.5}$$

$$= \sum_n \left\{ \sum_j \mathbf{D}_j(k)\mathbf{D}^+{}_j(k-n) \right\} \mathbf{m}(k-n)$$

$$= \sum_n \mathbf{R}(k,n)\mathbf{m}(k-n) \, .$$

$\mathbf{R}$, the analogous of the R in equation (2.3.3), is implicitly defined. Interesting applications of the analysis of the remaining aliasing noise are weighted stacking, and use of the space derivative. (Rocca and Ronen, 1984).

2.4. The transpose operator

In many applications the transpose operator is practically or approximately, the inverse (Claerbout, this SEP report).

If equation (1.4.2) is written as

$$\mathbf{d} = \mathbf{G}\mathbf{m} \, , \tag{2.4.1}$$

then the transpose, $\mathbf{G}^*$ is a transform from the data domain, where $\mathbf{d}$ is, to the model space, where $\mathbf{m}$ and $\hat{\mathbf{m}}$ are

$$\hat{\mathbf{m}} \equiv \mathbf{G}^* \mathbf{d} \ . \tag{2.4.2}$$

Maybe it is not obvious that $\mathbf{G}^*$ is DMO after zero traces are put where missing traces are, but that's exactly what it is. To see that, consider $N=3$ and $J=4$,

$$\mathbf{G} = \begin{pmatrix} \mathbf{D}^+{}_1(k-1) & \mathbf{D}^+{}_1(k) & \mathbf{D}^+{}_1(k+1) \\ \mathbf{D}^+{}_2(k-1) & \mathbf{D}^+{}_2(k) & \mathbf{D}^+{}_2(k+1) \\ \mathbf{D}^+{}_3(k-1) & \mathbf{D}^+{}_3(k) & \mathbf{D}^+{}_3(k+1) \\ \mathbf{D}^+{}_4(k-1) & \mathbf{D}^+{}_4(k) & \mathbf{D}^+{}_4(k+1) \end{pmatrix} . \tag{2.4.3}$$

From appendix A, the transpose of each block, $\mathbf{D}^+$, is the DMO, $\mathbf{D}$

$$\mathbf{D}^+{}_j{}^*(k) = \mathbf{D}_j(k) \ , \tag{2.4.4}$$

so the transpose of the whole thing is,

$$\mathbf{G}^* = \begin{pmatrix} \mathbf{D}^+{}_1{}^*(k-1) & \mathbf{D}^+{}_2{}^*(k-1) & \mathbf{D}^+{}_3{}^*(k-1) & \mathbf{D}^+{}_4{}^*(k-1) \\ \mathbf{D}^+{}_1{}^*(k) & \mathbf{D}^+{}_2{}^*(k) & \mathbf{D}^+{}_3{}^*(k) & \mathbf{D}^+{}_4{}^*(k) \\ \mathbf{D}^+{}_1{}^*(k+1) & \mathbf{D}^+{}_2{}^*(k+1) & \mathbf{D}^+{}_3{}^*(k+1) & \mathbf{D}^+{}_4{}^*(k+1) \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{D}_1(k-1) & \mathbf{D}_2(k-1) & \mathbf{D}_3(k-1) & \mathbf{D}_4(k-1) \\ \mathbf{D}_1(k) & \mathbf{D}_2(k) & \mathbf{D}_3(k) & \mathbf{D}_4(k) \\ \mathbf{D}_1(k+1) & \mathbf{D}_2(k+1) & \mathbf{D}_3(k+1) & \mathbf{D}_4(k+1) \end{pmatrix} . \tag{2.4.5}$$

Thus,

$$\begin{pmatrix} \hat{\mathbf{m}}(k-1) \\ \hat{\mathbf{m}}(k) \\ \hat{\mathbf{m}}(k+1) \end{pmatrix} = \begin{pmatrix} \mathbf{D}_1(k-1) & \mathbf{D}_2(k-1) & \mathbf{D}_3(k-1) & \mathbf{D}_4(k-1) \\ \mathbf{D}_1(k) & \mathbf{D}_2(k) & \mathbf{D}_3(k) & \mathbf{D}_4(k) \\ \mathbf{D}_1(k+1) & \mathbf{D}_2(k+1) & \mathbf{D}_3(k+1) & \mathbf{D}_4(k+1) \end{pmatrix} \begin{pmatrix} \mathbf{d}_1(k) \\ \mathbf{d}_2(k) \\ \mathbf{d}_3(k) \\ \mathbf{d}_4(k) \end{pmatrix} \tag{2.4.6}$$

$\hat{\mathbf{m}}$ is defined in equation (2.4.6) for the frequencies $-3/2<k<3/2$. For $|k| < 1/2$,

$$\hat{\mathbf{m}}(k) = \begin{pmatrix} \mathbf{D}_1(k) & \mathbf{D}_2(k) & \mathbf{D}_3(k) & \mathbf{D}_4(k) \end{pmatrix} \begin{pmatrix} \mathbf{d}_1(k) \\ \mathbf{d}_2(k) \\ \mathbf{d}_3(k) \\ \mathbf{d}_4(k) \end{pmatrix}$$

$$= \sum_{j=1}^{4} \mathbf{D}_j(k)\mathbf{d}_j(k) \tag{2.4.7}$$

from the middle row of equation (2.4.6). Equation (2.4.7) is just DMO and stacking.

For $-3/2 < k < -1/2$, we use the upper ($n=1$) row of (2.4.6),

$$\hat{\mathbf{m}}(k) = \begin{pmatrix} \mathbf{D}_1(k) & \mathbf{D}_2(k) & \mathbf{D}_3(k) & \mathbf{D}_4(k) \end{pmatrix} \begin{pmatrix} \mathbf{d}_1(k+1) \\ \mathbf{d}_2(k+1) \\ \mathbf{d}_3(k+1) \\ \mathbf{d}_4(k+1) \end{pmatrix}$$

$$= \sum_{j=1}^{J} \mathbf{D}_j(k)\mathbf{d}_j(k+1) \; . \tag{2.4.8}$$

And similarly, for $1/2 < k < 3/2$,

$$\hat{\mathbf{m}}(k) = \sum_{j=1}^{J} \mathbf{D}_j(k)\mathbf{d}_j(k-1) \; . \tag{2.4.9}$$

This can be done by replicating the Fourier transform of the data

$$\mathbf{d}(k-1) = \mathbf{d}(k) = \mathbf{d}(k+1) \; , \tag{2.4.10}$$

which is equivalent to adding zero traces between data traces in the space domain.

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

2.5. Field data example

Application of $\mathbf{G}^*$ to field data is shown in Figures (2-3) to (2-18). NMO-stack of data from the gulf of Mexico is shown in Figure (2-3): dipping reflectors were lost in stacking due to the dip-filtering term of equation (2.1.2). With DMO (Figure (2-4)), the dipping reflectors show up but they are spatially aliased and lost in migration (Figure (2-5)), even though DMO has been performed!

Application of $\mathbf{G}^*$ with $N=4$ is shown in Figure (2-6). (actually a finite differencing equivalent of $\mathbf{G}^*$ was applied). The midpoint interval was reduced to 12.5 m and the dipping reflectors are now adequately sampled. However, there is aliasing noise, especially around the very shallow (because of the mute), and the very deep (because the DMO operator is small down-deep), reflectors. The aliasing noise was then reduced by using weighted stacking, (Figures (2-7) and (2-8)), and by use of the spatial derivative (figures (2-9) and (2-10)).

Figures (2-11) to (2-18) describe results of similar processing, but the midpoint interval was increased to 100 m, by eliminating every other CDP gather. I used $\mathbf{G}^*$ with $N=8$ to reduce the sampling interval down to 12.5 m.

When I prepared this example to the last SEG, I did not know that I was using a transpose instead of an inverse. The indication that I was not using the right operator is that flat reflectors seem not less (and perhaps more) noisy than dipping reflectors.

FIG. 2-3. NMO-stack. Dipping reflectors were lost in stacking.

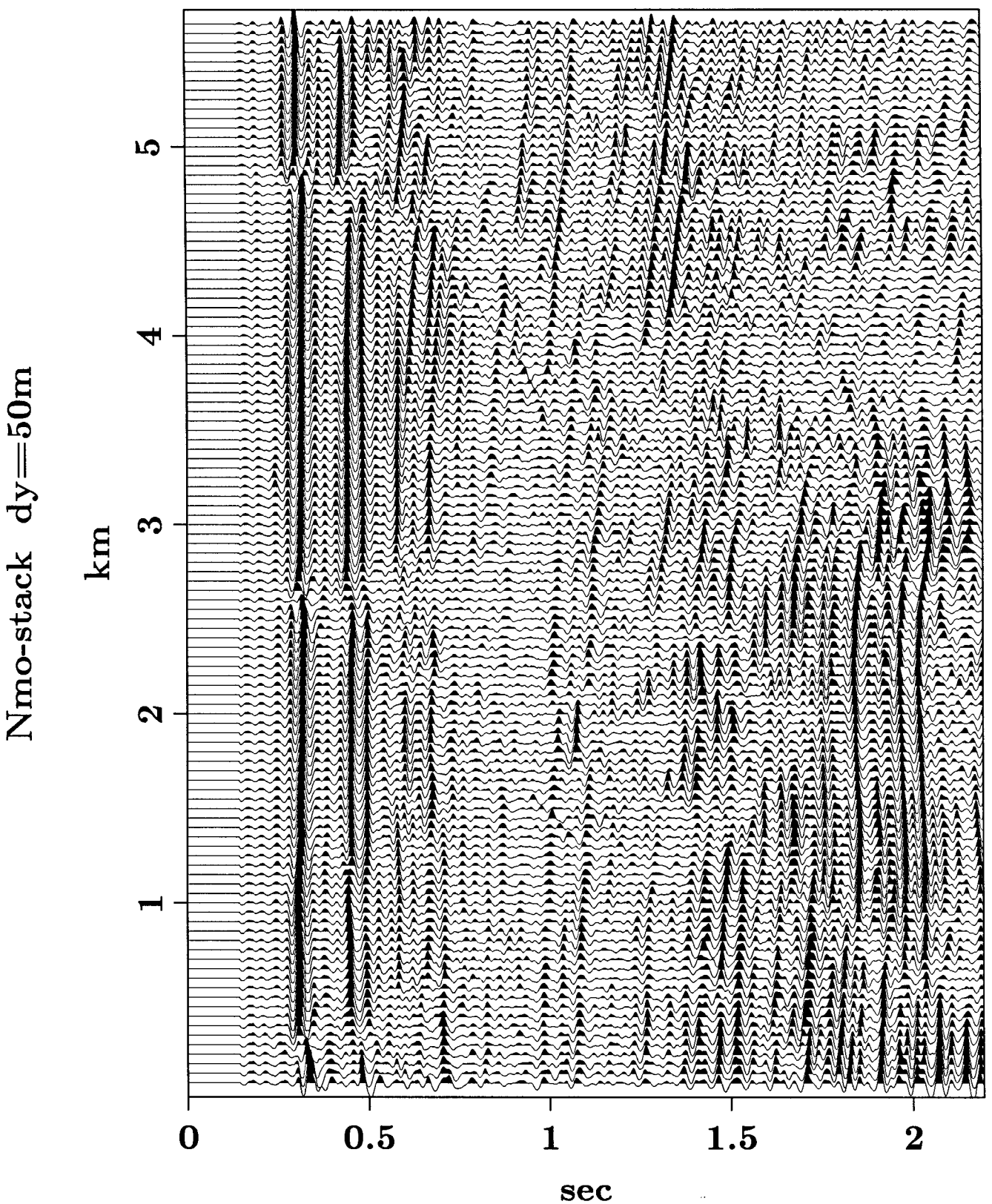


FIG. 2-4. The same data with DMO: the dipping reflectors show up (at 1 sec under 1.5 km and 4 km) but the stack is spatially aliased with midpoint interval of 50 m.

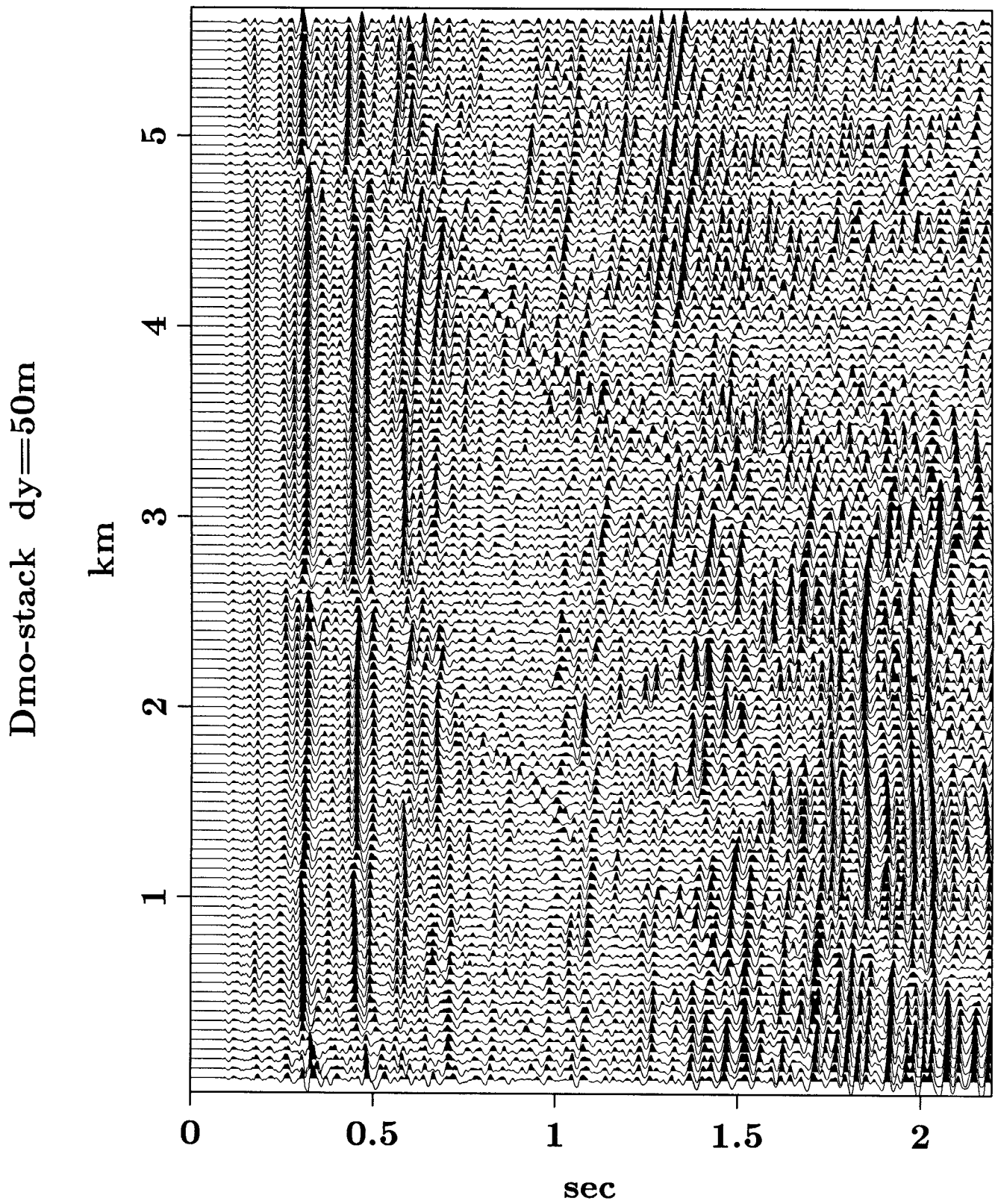


FIG. 2-5. Migration of the stack in Figure (2-4). The dipping reflectors are lost because of the spatial aliasing.

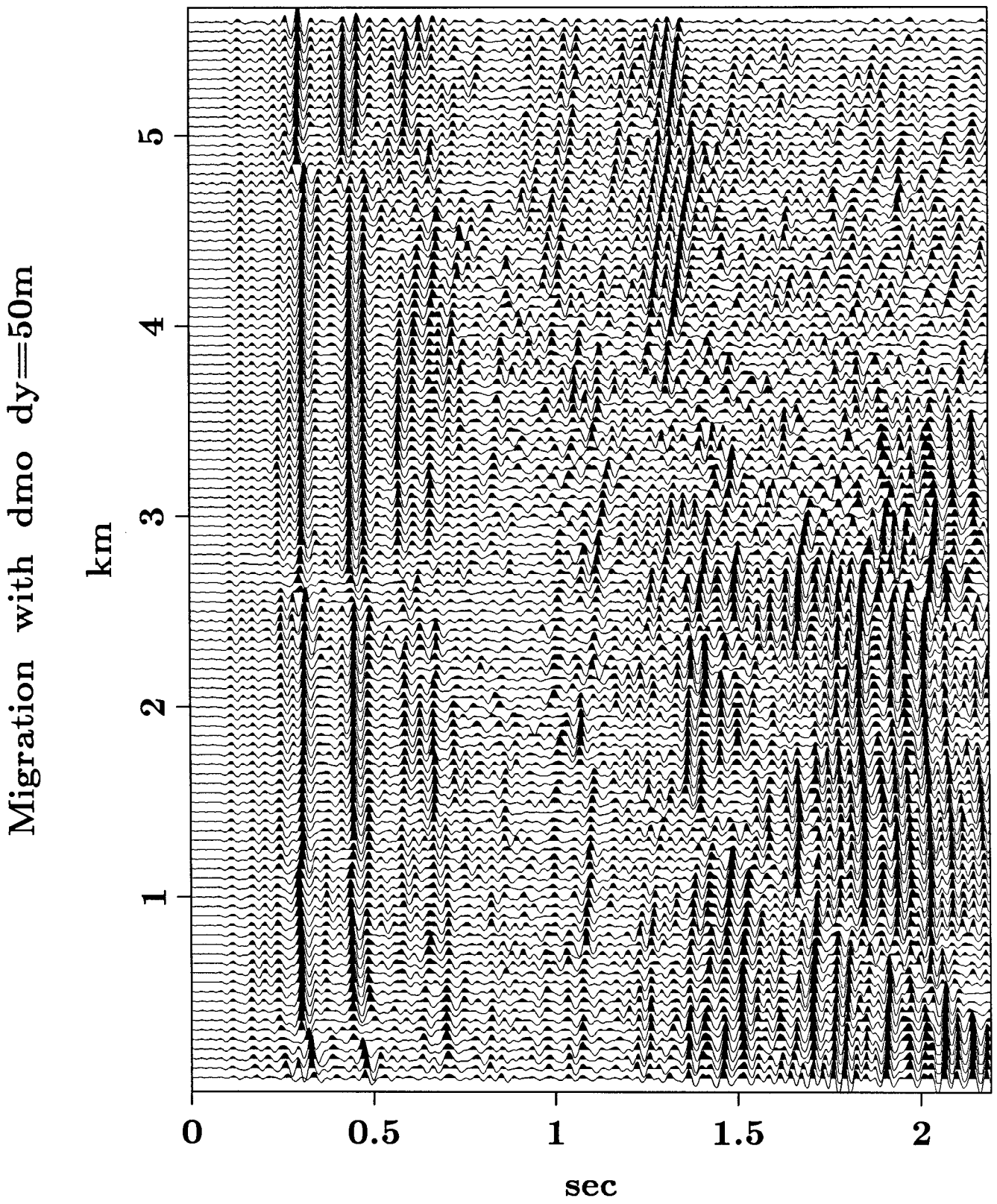


FIG. 2-6. DMO-stack of the same data. The midpoint interval was reduced to 12.5 m by application a finite differencing equivalent to $\mathbf{G}^*$ with $N=4$.

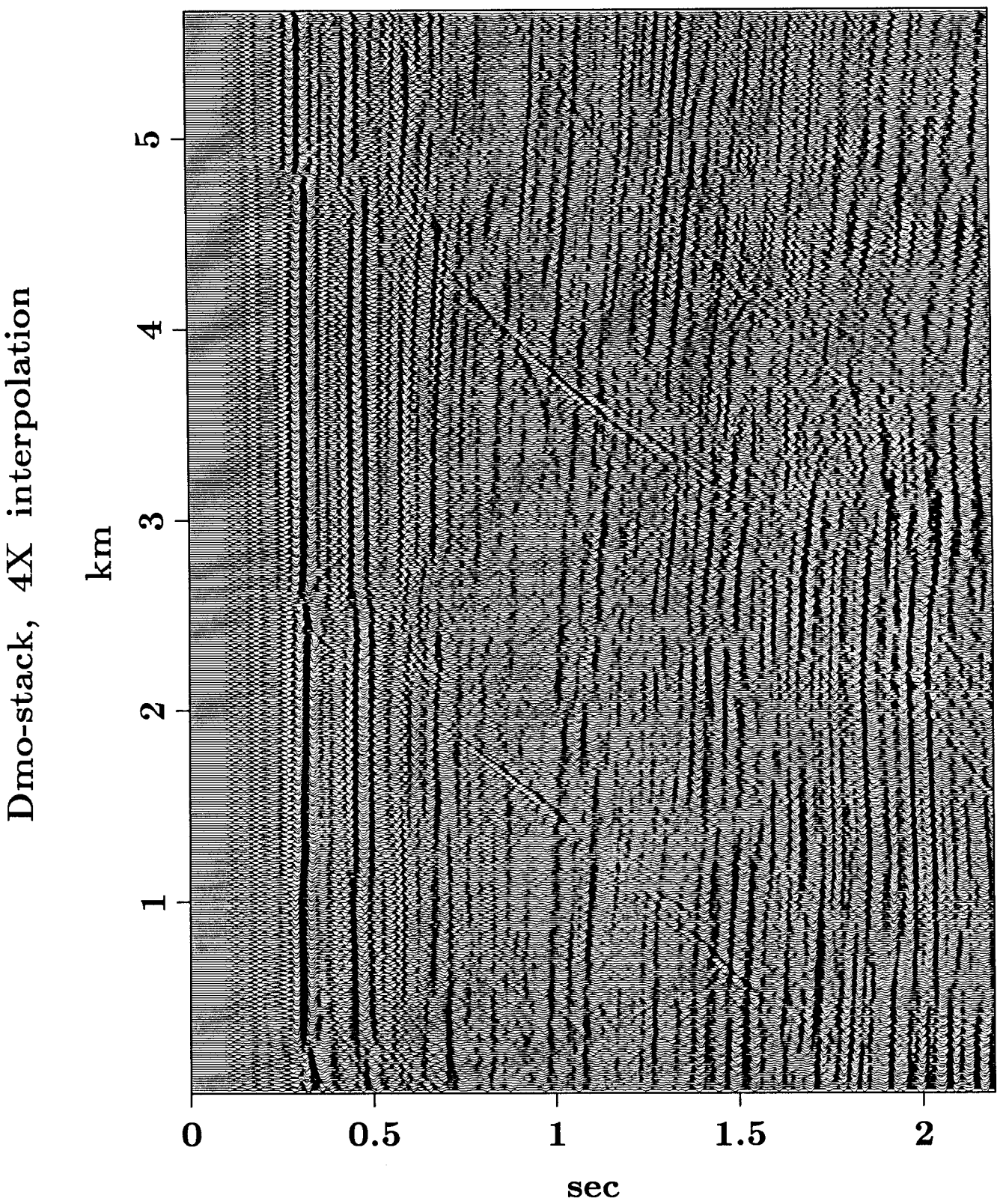


FIG. 2-7. Reduction of the aliasing noise by using weighted stack.

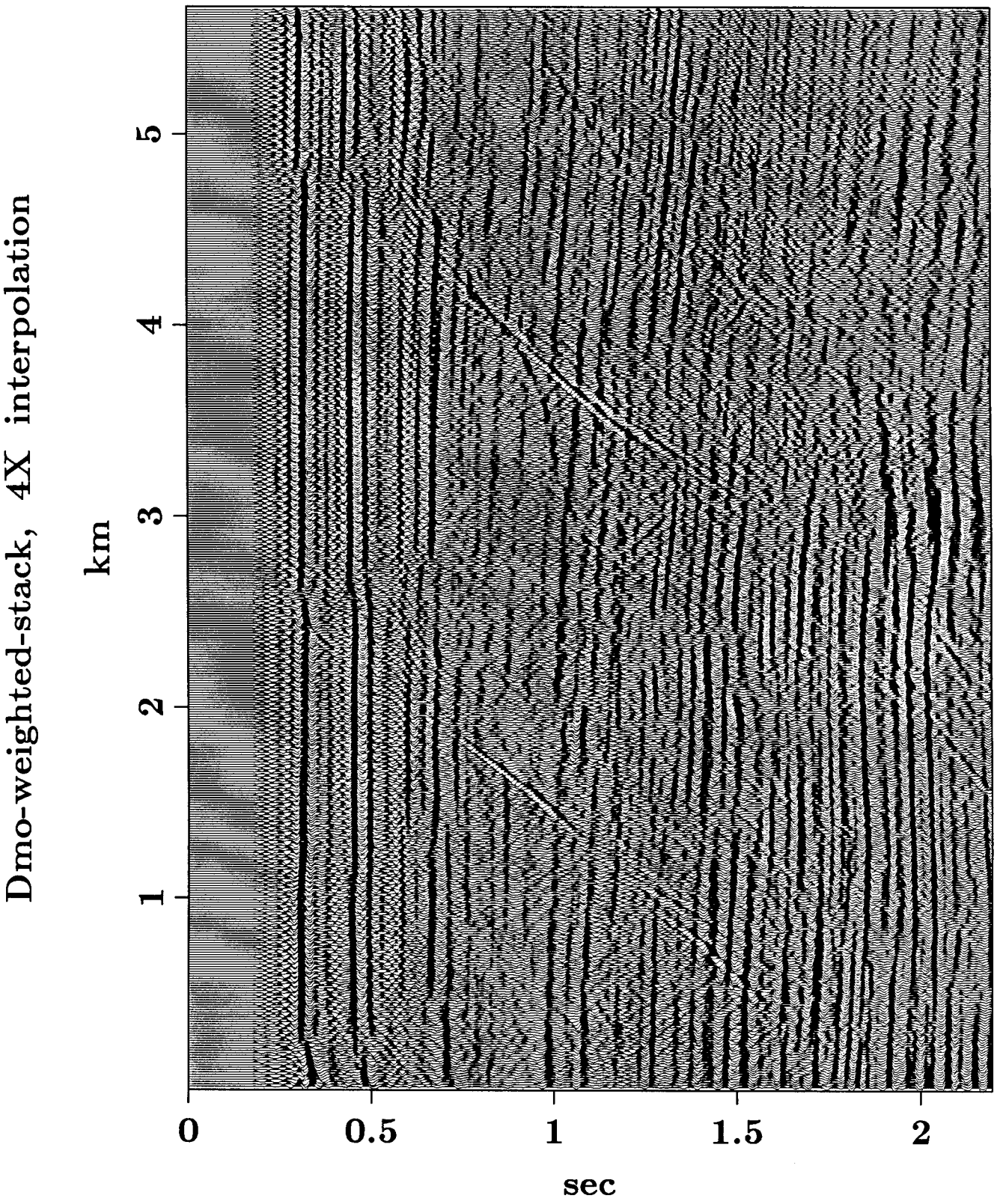


FIG. 2-8. Migration of the stack of Figure (2-7).

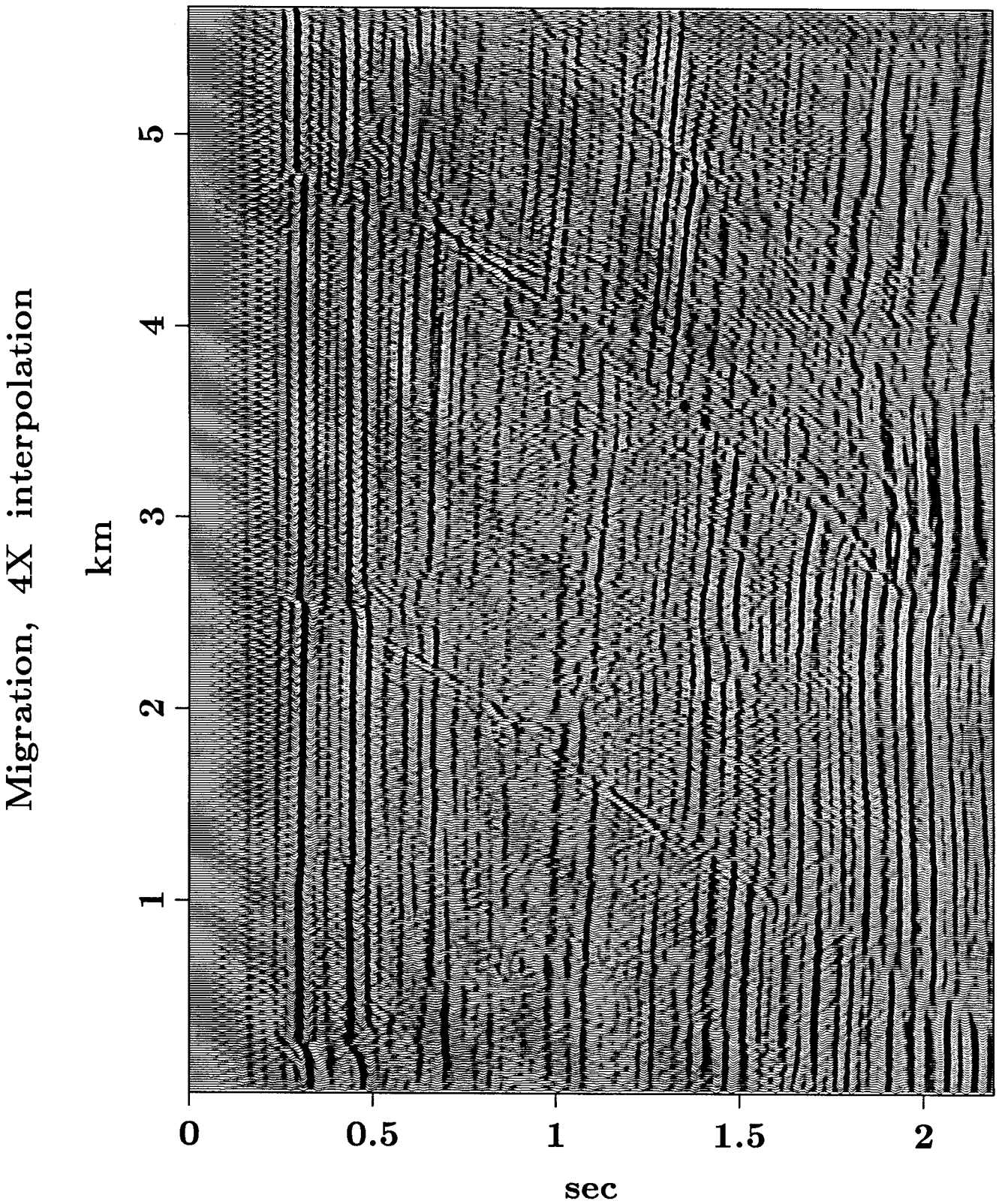


FIG. 2-9. DMO-stack using the space derivative. The aliasing noise is reduced.

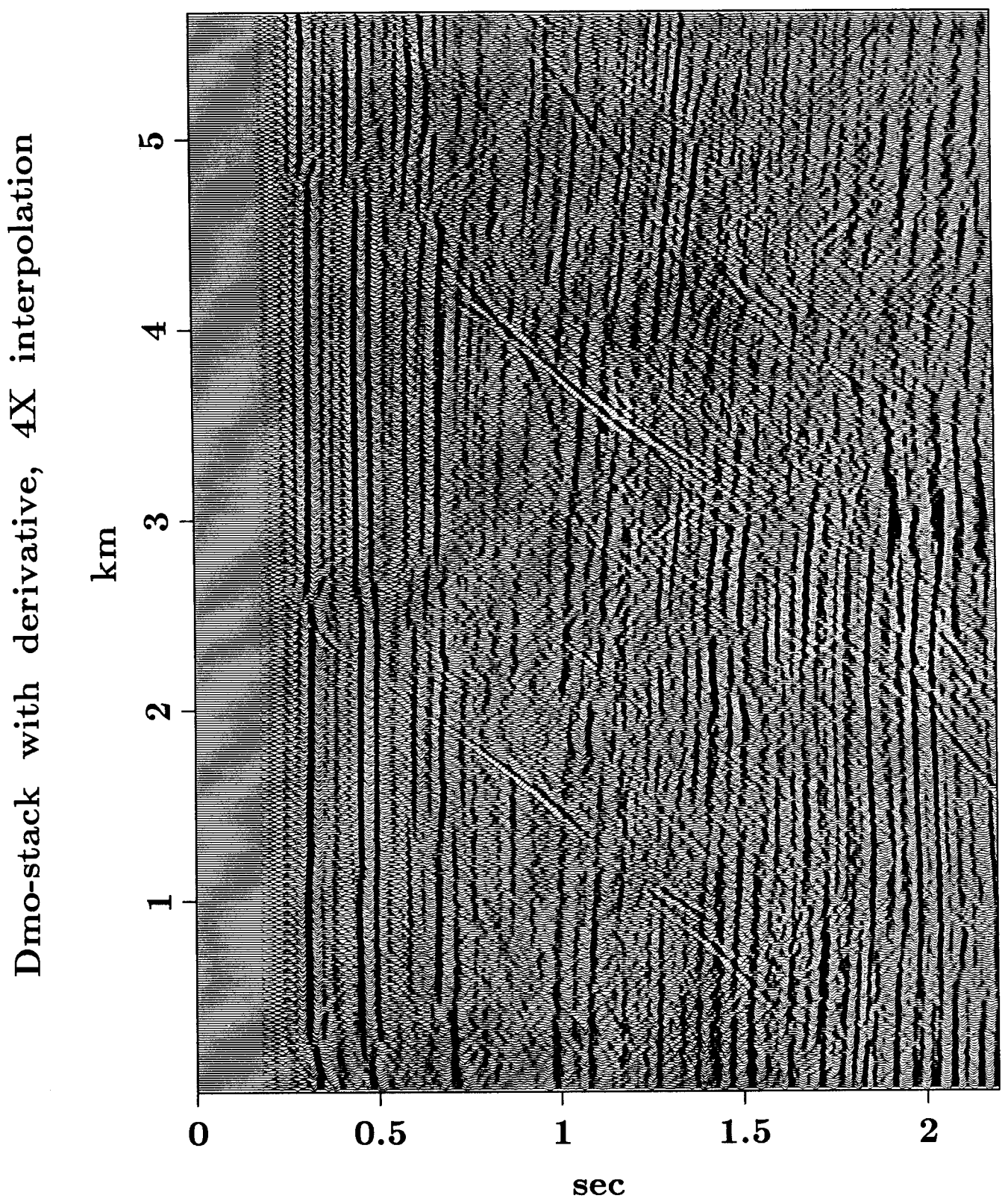


FIG. 2-10. Migration of the stack in Figure (2-9).

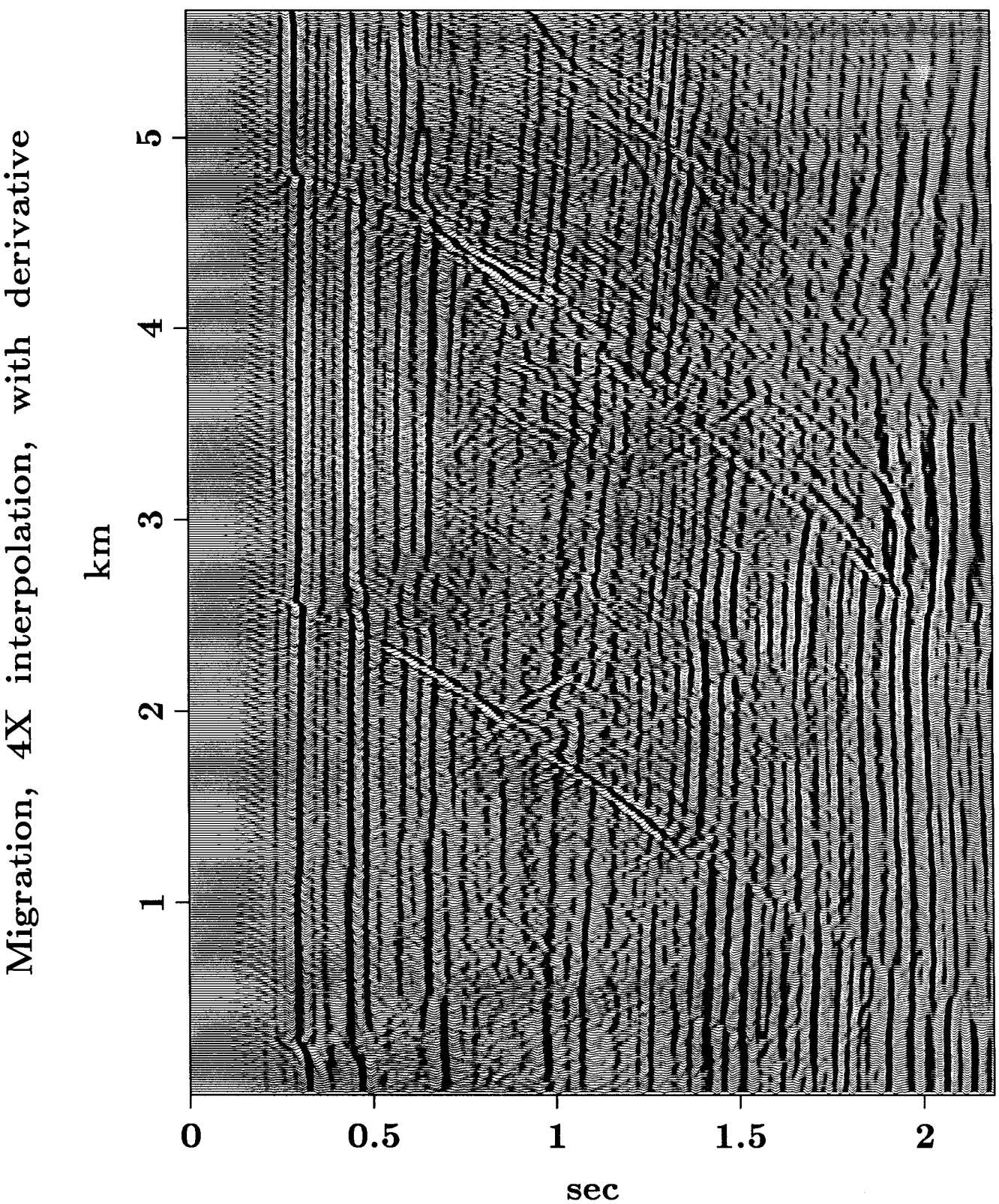


FIG. 2-11. NMO-stack of every other midpoint of the data in Figure (2-3). The midpoint interval is 100 m.

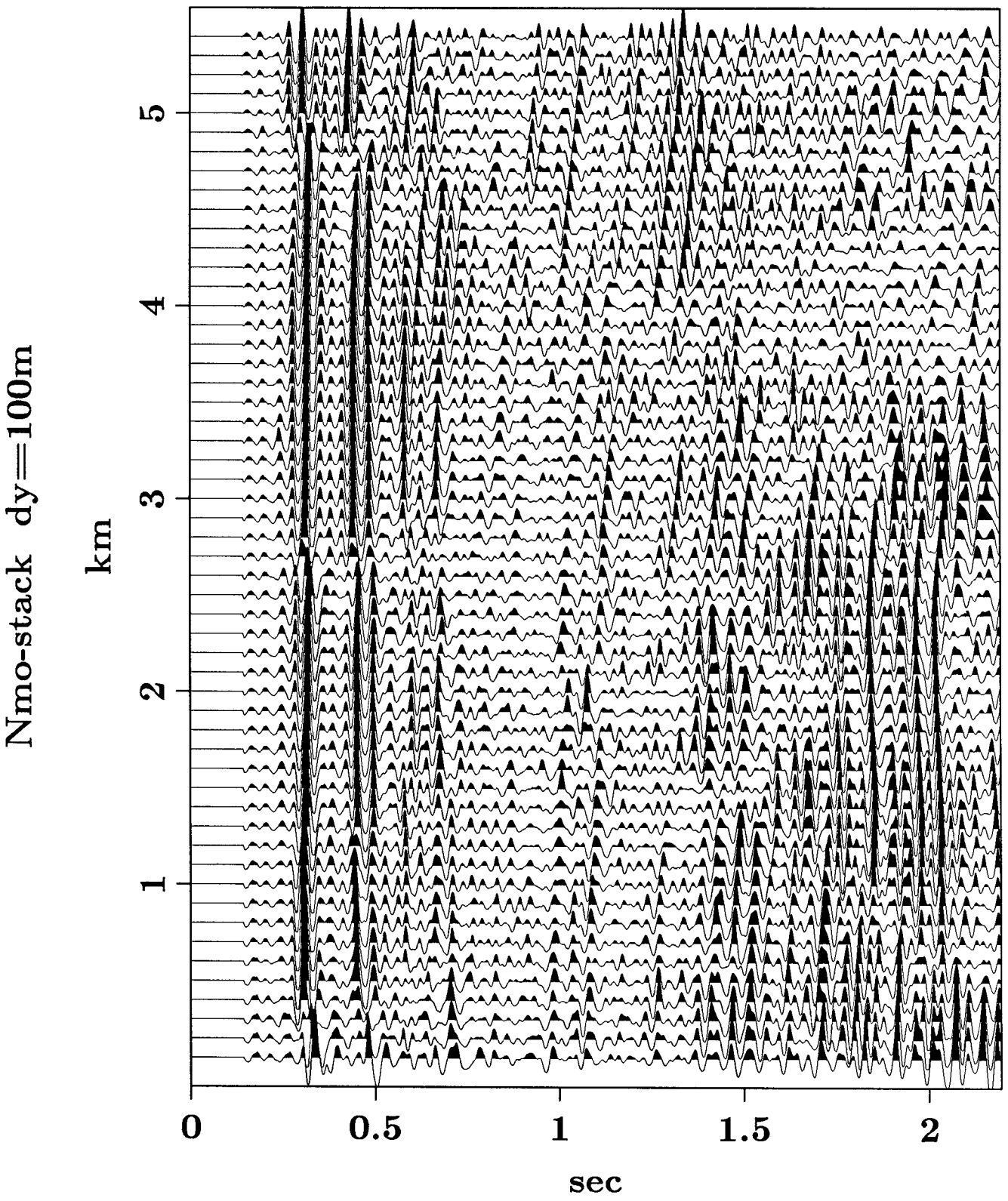


FIG. 2-12. DMO-stack of the data in Figure (2-14). Anything but the flat reflectors is badly aliased with a midpoint interval of 100 m.

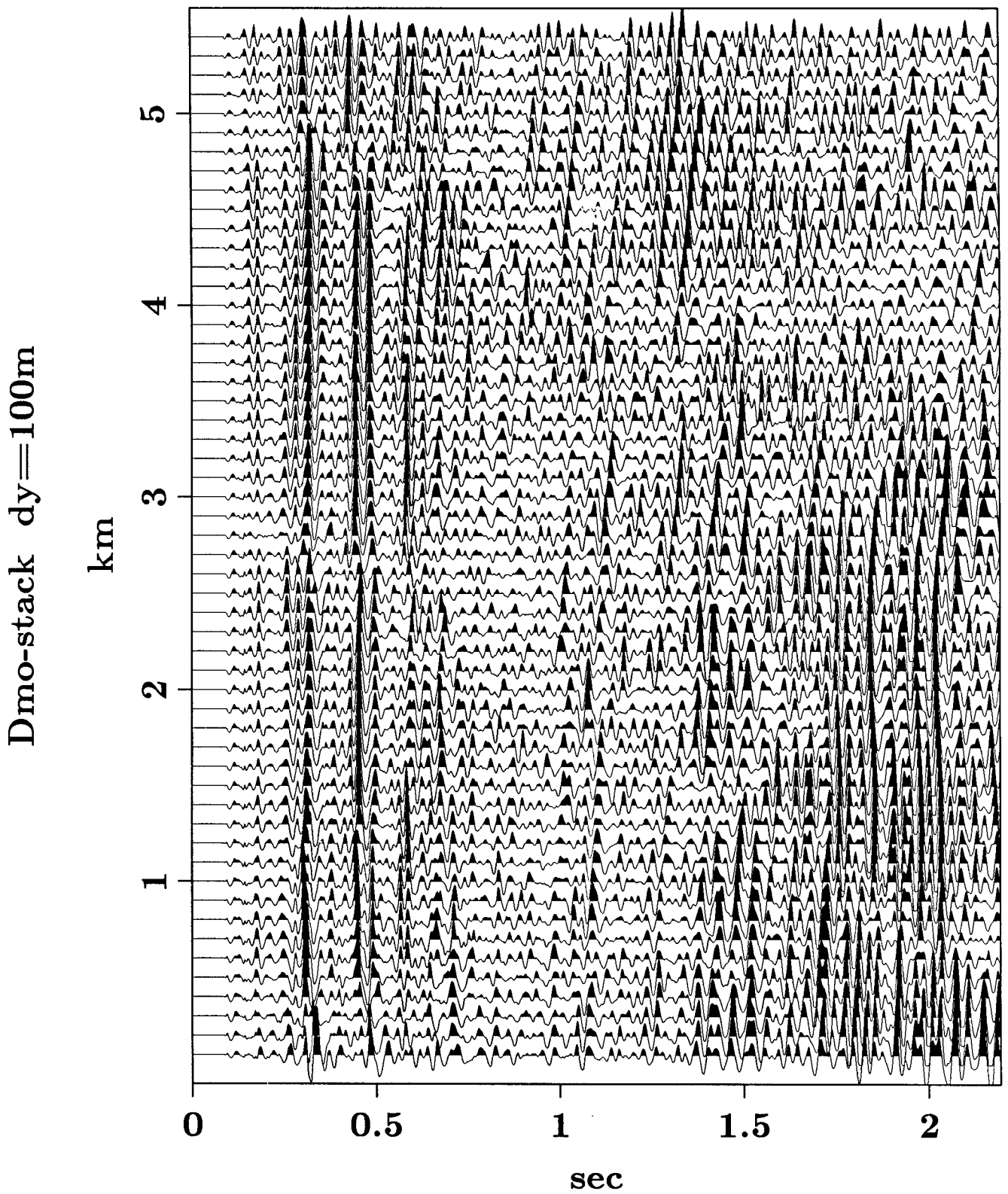


FIG. 2-13. Migration of the stack of Figure (2-12).

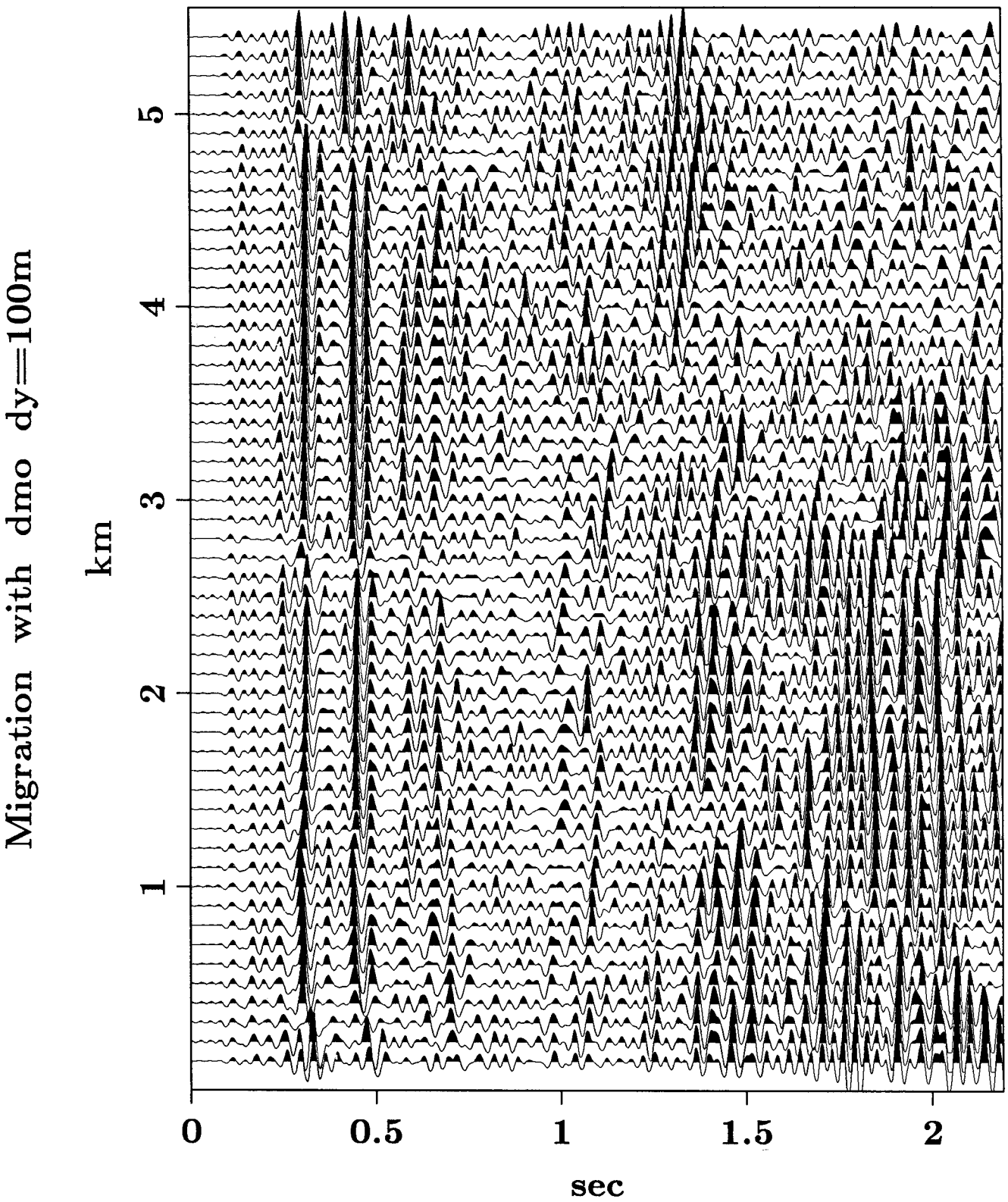


FIG. 2-14. Applying $\mathbf{G}^*$ with $N=8$ to the same data of Figure (2-11). The midpoint interval is now 12.5 m.

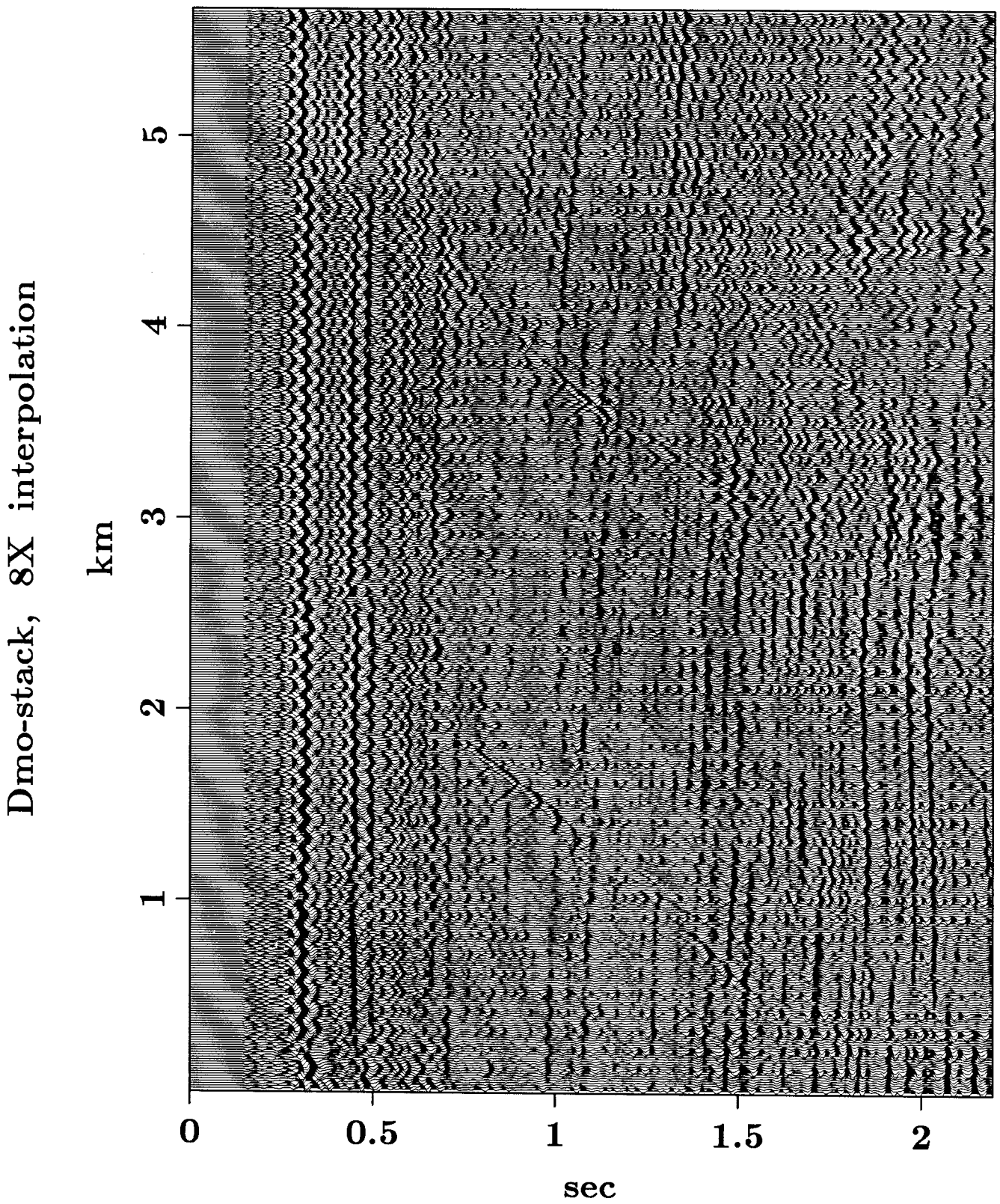


FIG. 2-15. The same as Figure (2-14) but with weighted stack.

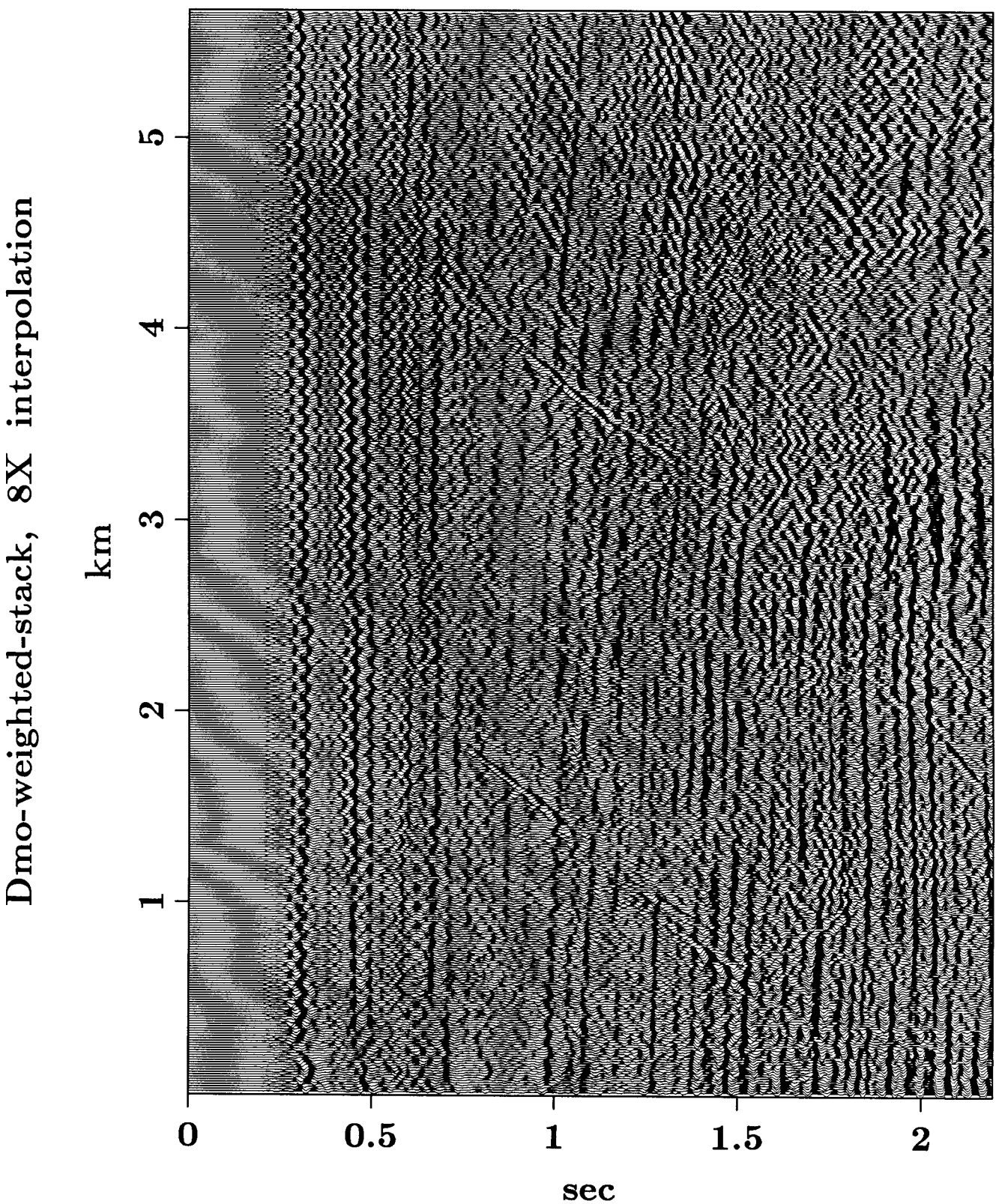


FIG. 2-16. Migration of the stack in Figure (2-15).

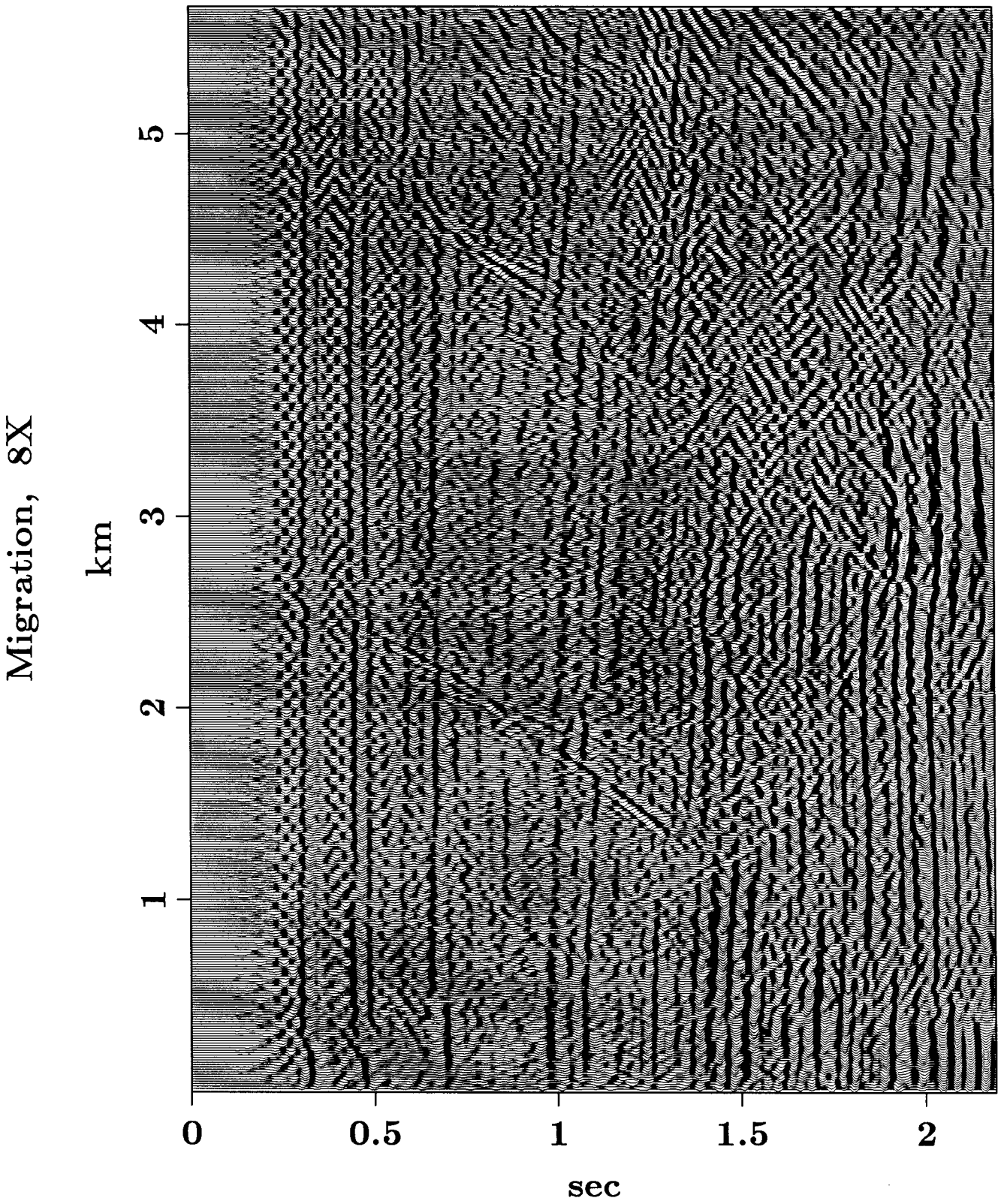


FIG. 2-17. Stack with derivative, $N=8$ in $\mathbf{G}^*$.

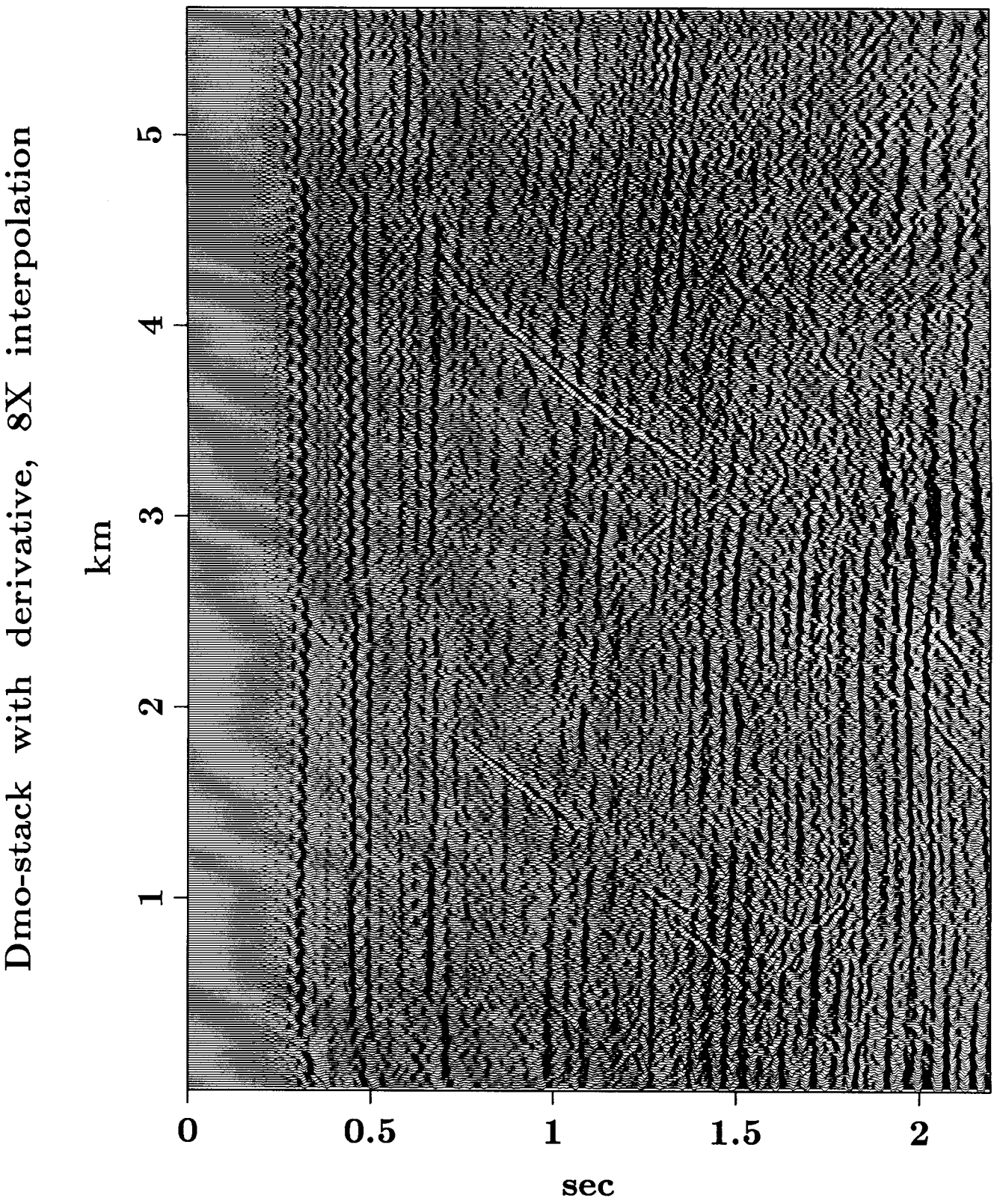
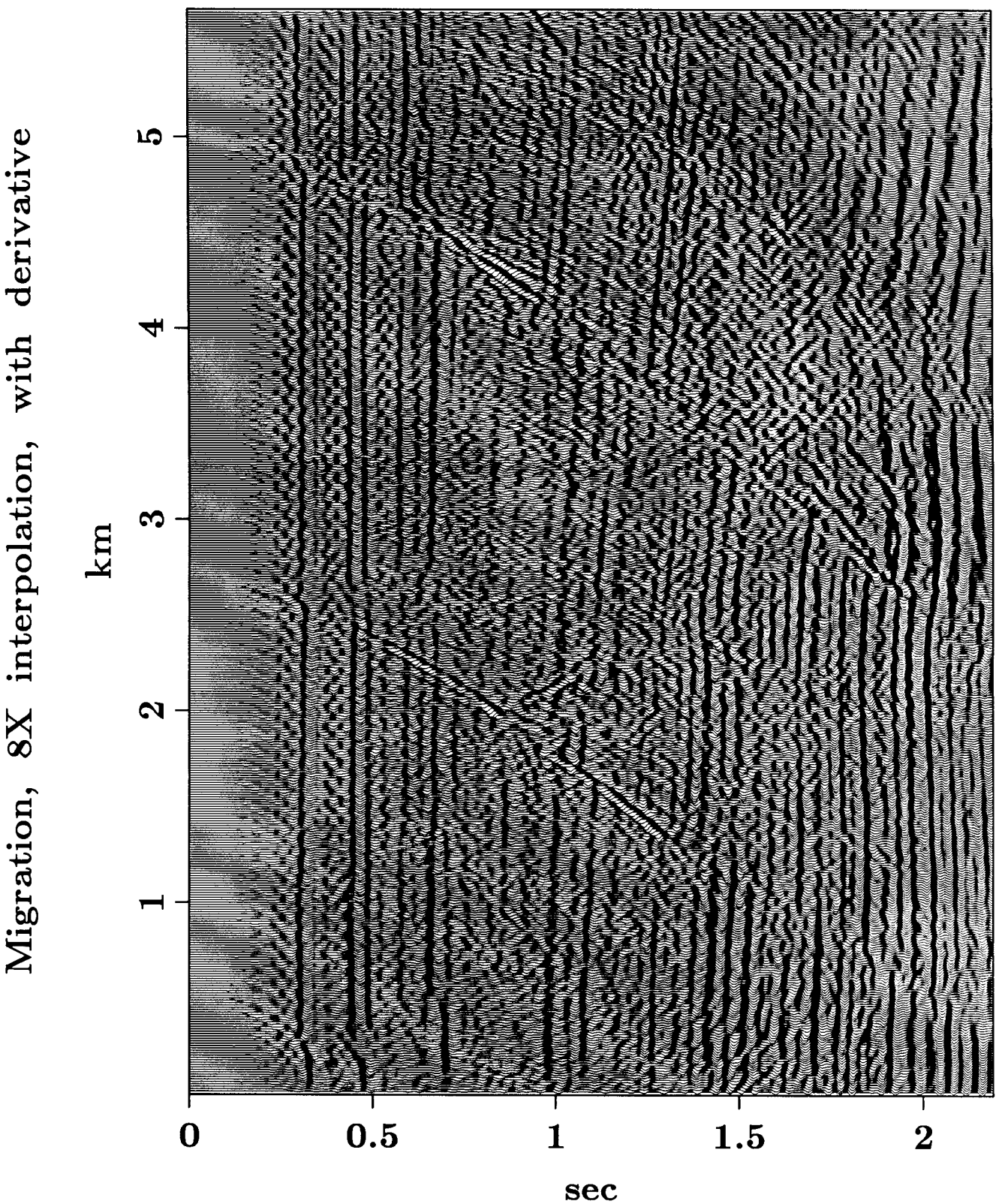


FIG. 2-18. Migration of the stack of Figure (2-17).



3. MULTI-CHANNEL INVERSION

3.1. Generalized inversion

The equation,

$$\mathbf{d} = \mathbf{G}\mathbf{m} \ , \tag{3.1.1}$$

can be solved by minimizing the error,

$$\mathbf{e} = \mathbf{d} - \mathbf{G}\mathbf{m} \ . \tag{3.1.2}$$

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

$$E = \mathbf{e}^* \mathbf{e} \ . \tag{3.1.3}$$

A more general solution to equation (3.1.1) is to minimize

$$\begin{aligned} E &= \mathbf{e}^* \mathbf{e} + \mathbf{m}^* \mathbf{C}\mathbf{m} \ , \\ &= \Big(\mathbf{d}^* - \mathbf{m}^* \mathbf{G}^*\Big)\Big(\mathbf{d} - \mathbf{G}\mathbf{m}\Big) + \mathbf{m}^* \mathbf{C}\mathbf{m} \\ &= \mathbf{d}^* \mathbf{d} - \mathbf{d}^* \mathbf{G}\mathbf{m} - \mathbf{m}^* \mathbf{G}^* \mathbf{d} + \mathbf{m}^* \mathbf{G}^* \mathbf{G}\mathbf{m} + \mathbf{m}^* \mathbf{C}\mathbf{m} \ . \end{aligned} \tag{3.1.4}$$

Where $\mathbf{C}$ is positive definite. $\mathbf{C}$ is sometimes called damping, it is important when $\mathbf{G}^* \mathbf{G}$ is singular or ill conditioned.

Minimization of E with respect to $\mathbf{m}$ gives,

$$0 = \frac{1}{2} \nabla_{\mathbf{m}} E = -\mathbf{G}^* \mathbf{d} + \mathbf{G}^* \mathbf{G}\mathbf{m} + \mathbf{C}\mathbf{m} \ . \tag{3.1.5}$$

So,

$$\mathbf{m} = \Big(\mathbf{G}^* \mathbf{G} + \mathbf{C}\Big)^{-1} \mathbf{G}^* \mathbf{d} \ . \tag{3.1.6}$$

$\mathbf{G}^* \mathbf{G}+\mathbf{C}$ cannot be singular if $\mathbf{C}$ is positive definite.

Equation (3.1.6) can be applied in two steps:

(1) Pre-stack partial migration (DMO): As I have shown, $\mathbf{G}^*$ is DMO with zero traces in place of missing traces in the input to DMO.

(2) Post-stack multi-channel inversion: Solving

$$\hat{\mathbf{m}} = [\mathbf{G}^* \mathbf{G}+\mathbf{C}]\mathbf{m} \ . \tag{3.1.7}$$

Steps (1) and (2) produce a well sampled stacked section, $\mathbf{m}$, ready for migration. To obtain the reflectors we need post-stack migration.

3.2. Multi-channel inversion

The singular values of $\mathbf{G}(k)$ are shown in Figure (3-1). For $k=0$ and for $k=0.5$ there is apparently a rank deficiency of 32, but there are constrains because $m(x,t)$ is real.

$\mathbf{G}^*\mathbf{G}$ is approximately sparse, as can be seen in Figure (3-2); each block is approximately diagonal so the inverse of each block is readily obtained. This is encouraging because a typical size of $\mathbf{G}^*\mathbf{G}$ for field data might be 4096×4096 if N=4 and $n\omega$=1024.

I am currently trying various damping matrices $\mathbf{C}$. So far I have considered diagonal $\mathbf{C}$'s of the form

$$[\mathbf{C}]_{ij} = \delta_{ij}\left(e_0 + e_1\frac{(k_i/\omega_i)^n}{1+(k_i/\omega_i)^n}\right), \tag{3.2.1}$$

where k_i is determined by the block, and ω_i is within the block. The damping of equation (3.2.1) discriminates against evanescent frequencies. A particular choice of $\mathbf{C}$ is shown in Figure (3-3).

3.3. Synthetic example

Figure (3-4a) shows the model: a zero-offset section with a 90^o dipping event and a flat event. The midpoint interval is 25 m.

The synthetic data are four CMP gathers, 200 m apart, each has 5 traces: zero-offset to 400 m half offset.

The section 3-4b is DMO-stack without any interpolation, it is hopelessly aliased and post stack sinc interpolation (Figure (3-4c)), can interpolate the flat event but is unable to interpolate the dipping event.

The results of the application of $\mathbf{G}^*$ are shown in Figure (3-4d). The number of offsets is $J=5$ and the interpolation is with $N=8$, so the system (1.4.2) is underdetermined!

Figures (3-4e) and (3-4f) show the results of applying the post-stack multi-channel inversion of equation (3.1.7) with different $\mathbf{C}$'s: (3-4e) is when $\mathbf{C}=0.1\,\mathbf{I}$, and (3-4f) is when $\mathbf{C}$ is as shown in Figure (3-3). The multichannel inversions are also shown in the hidden line plots of Figure (3-5). The improvement in performing the multichannel inversion is substantial, in particular for the flat reflector. The dipping reflector was relatively well interpolated just by the DMO in Figure (3-4d). The V patterns were explained by Rocca and Ronen (1984).

FIG. 3-1. The singular values of $\mathbf{G}(k)$ for various k's.

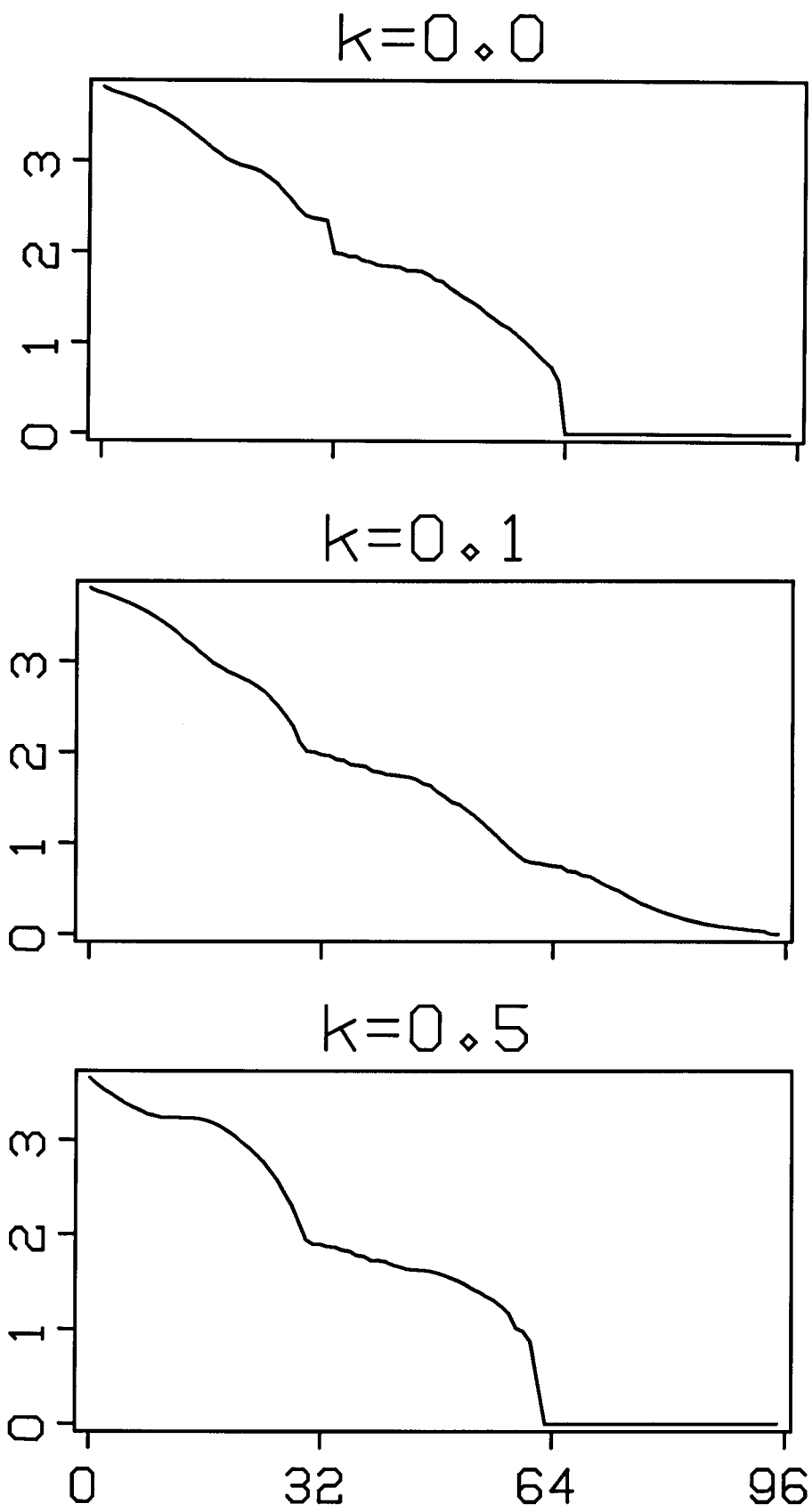
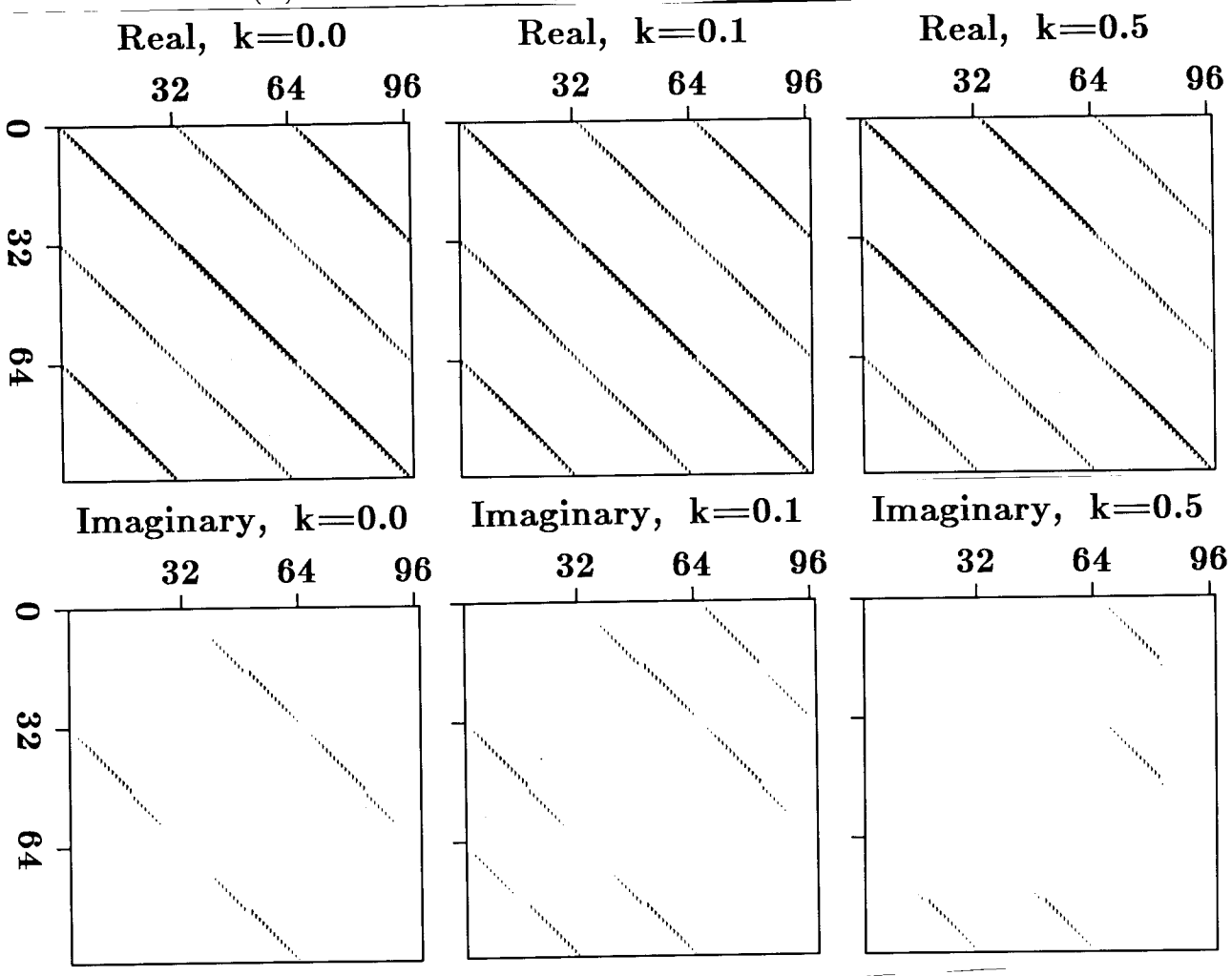


FIG. 3-2. $\mathbf{G}^* \mathbf{G}(k)$ for various k's. They are approximately sparse.



The real part of $\mathbf{G}^* \mathbf{G}(k = 0.1)$.

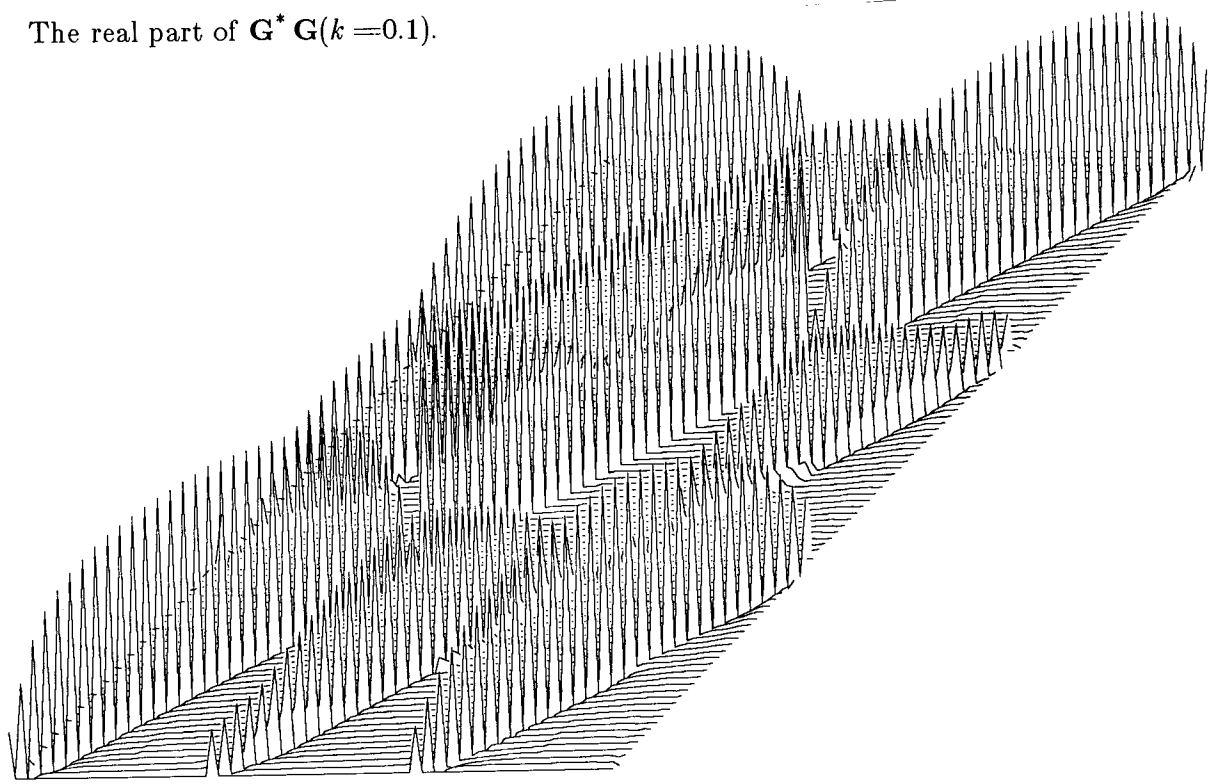


FIG. 3-3. The diagonal of $\mathbf{C}(k)$ for various k's, with e_0=0.05, e_1=1 and n=4 in equation (3.2.1).

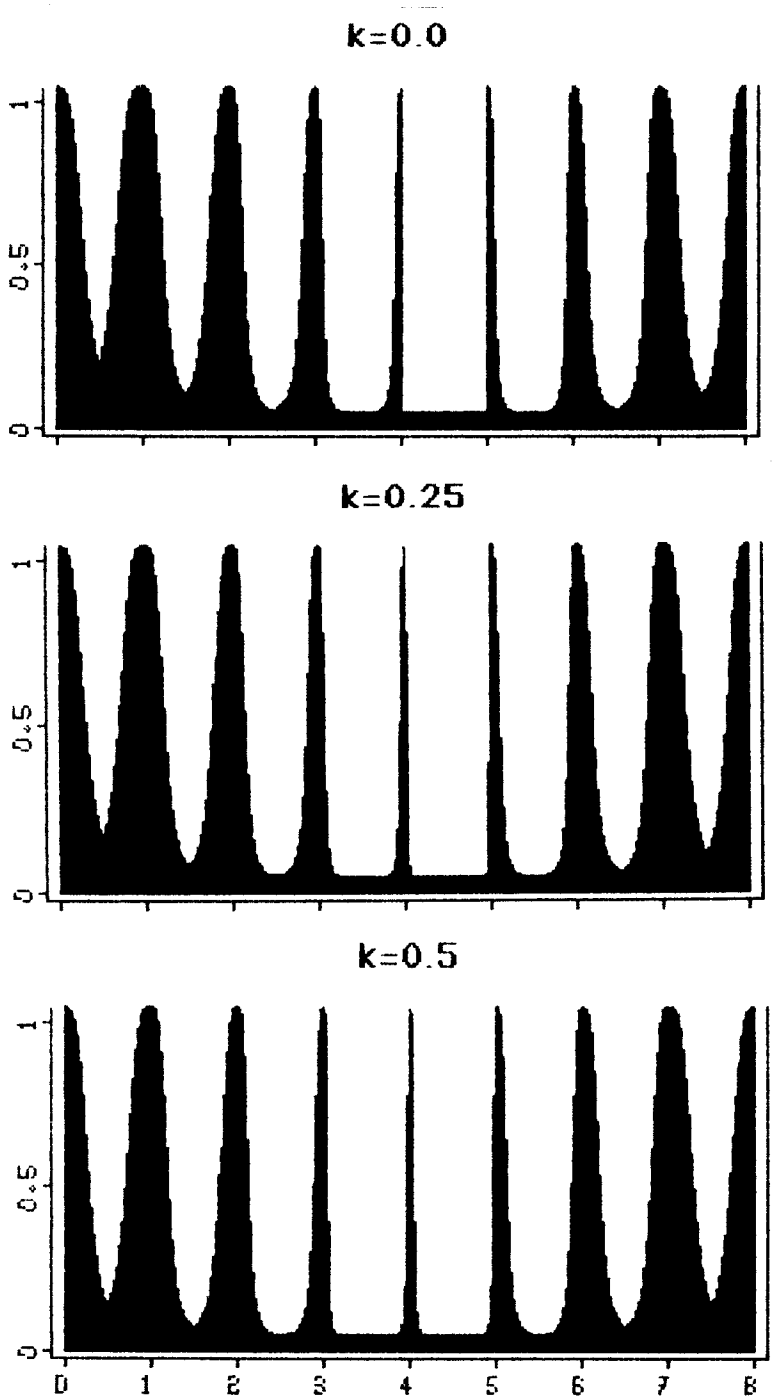


FIG. 3-4. Synthetic example for the post stack multi-channel inversion.

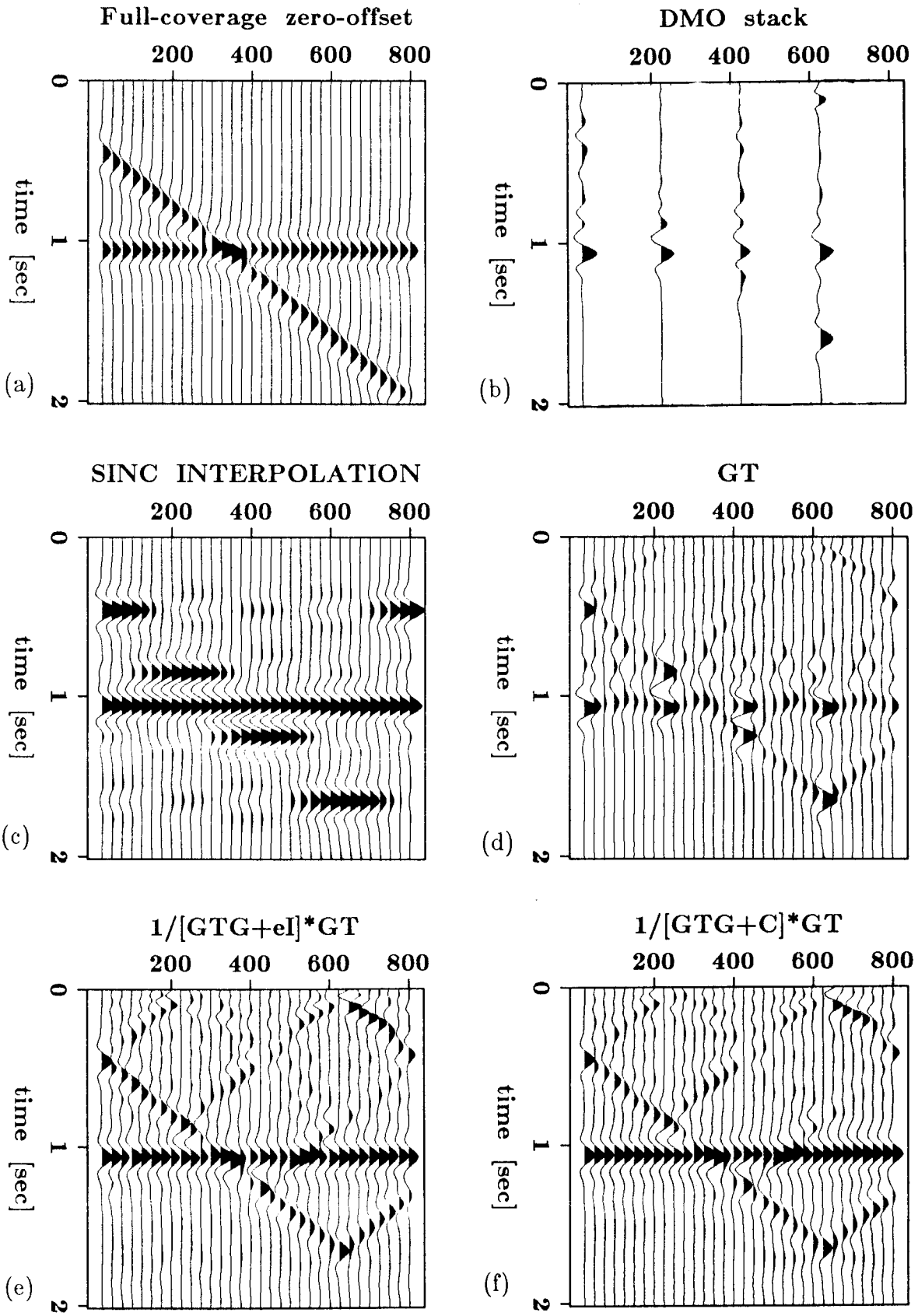
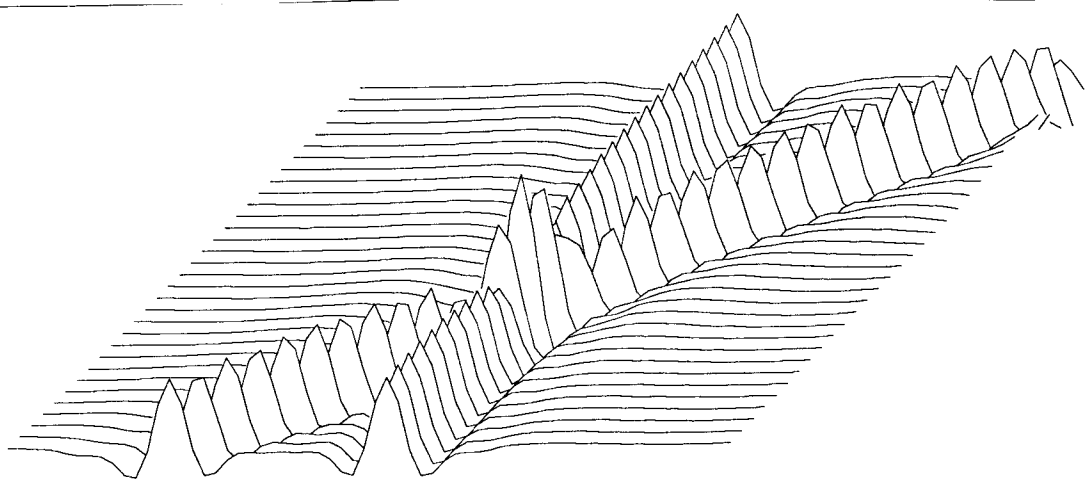
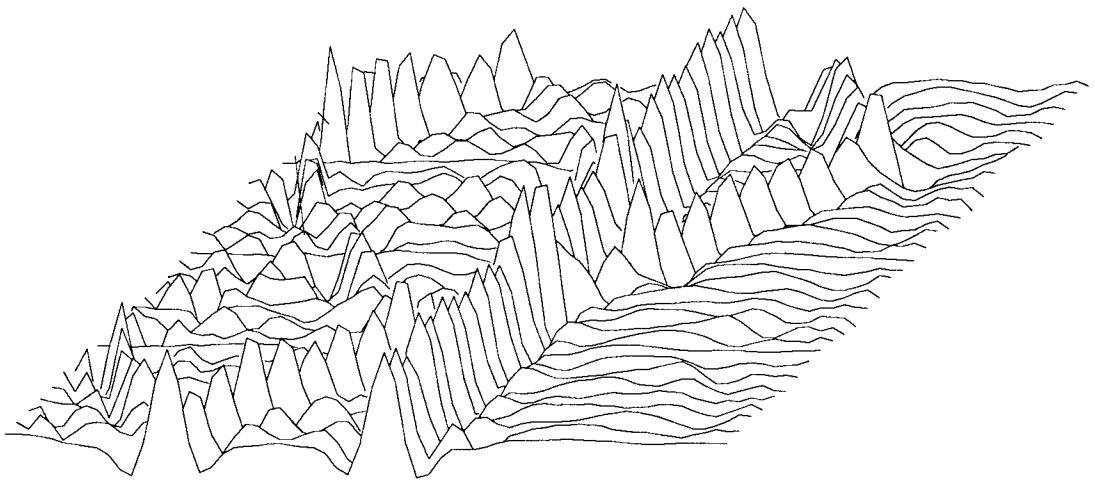


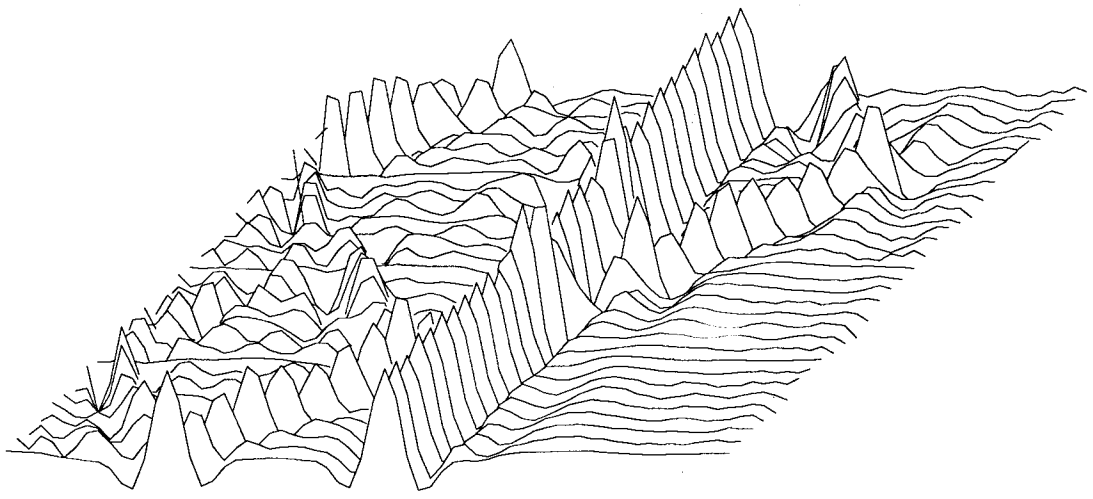
FIG. 3-5.



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



(b) Inversion with e_0=0.1; e_1=0, n =4, in equation (3.2.1).



(c) Inversion with e_0=0.05; e_1=1, n =4, in equation (3.2.1).

CONCLUSIONS

Spatial aliasing in multi-channel data can be overcome by a combination of pre-stack and post-stack processes. The pre-stack process is based on the DMO operation, it is equivalent to applying the transpose, $\mathbf{G}^*$, of a certain operator $\mathbf{G}$. The post-stack process, so far tested only on synthetic data, is the generalized inverse of $\mathbf{G}^* \mathbf{G}$. In analogy to tomography, the pre-stack step is back projection, the second is the rho filter.

The ability to recover resolution from aliased data and the limitations of this ability, should be taken into account when 3-D surveys are designed.

REFERENCES

Bolondi G., Loinger E., and Rocca F., 1982, Offset continuation of seismic sections: Geophysical Prospecting, **30**, 813-828.

Claerbout, J.F., 1985, What is the transpose operation? : SEP, **42**.

Deregowski, S.M., and Rocca, F., 1981, Geometrical optics and wave theory of constant offset sections in layered media: Geophysical Prospecting, **29**, 384-406.

Fowler, P., 1984, Velocity independent imaging: Presented at the 54th annual international SEG meeting, Atlanta.

Hale, D., 1983, Dip Move-out by Fourier transform: Ph.D. thesis, Stanford University. SEP, **36**.

Jakubowicz, H., 1984, A simple efficient method of pre-stack partial migration: Presented at the 54th annual international SEG meeting, Atlanta.

Judson, D.R., Schultz, P.S., and Sherwood, J.W.C., 1978, Equalizing the stacking velocities of dipping events via Devilish: Presented at the 48th annual international SEG meeting, San Francisco; brochure released by Digicon Geophysical Corp.

Linden, D.A., 1959, A discussion of Sampling theorems: Proc. I.R.E., **47**, 1219.

Ottolini, R., 1982, Migration of reflection seismic data in angle-midpoint coordinates: Ph.D. thesis, Stanford University. SEP, **33**.

Rocca, F., and Ronen, J., 1984, Improving resolution by dip-moveout: Presented at the 54th annual SEG meeting, Atlanta.

Salvador L. and Savelli S., 1982, Offset continuation for seismic stacking: Geophysical prospecting **30**, 829-849.

Yilmaz, O., 1979, Prestack Partial Migration: Ph.D. thesis, Stanford University. SEP, **18**.

APPENDIX A: INVERSE DMO

Abstract

The inverse DMO is given by

$$d(t,k,h) = \int d\omega\, A^{-1}\, e^{-i\omega A t}\, m(\omega,k)\ ,$$

where $d(t,k,h)$ is a common-offset section after NMO and Fourier transform over space, $m(\omega,k)$ is the zero-offset section after Fourier transform over time and space, h is half the offset between shot and receiver, and A is $\sqrt{1+[hk/\omega t]^2}$.

Pre-stack full migration

The migration of a spike, $\delta(x)\,\delta(y)\,\delta(t-t_h)$, on a common-offset section, for constant velocity, produces the ellipsoid

$$\left(\frac{x}{vt_h/2}\right)^2 + \left(\frac{y}{vt_n/2}\right)^2 + \left(\frac{z}{vt_n/2}\right)^2 = 1\ . \tag{A-1}$$

t_h is the arrival time of the spike on the common-offset section, t_n is the normal move-out (NMO) time,

$$t_n^2 = t_h^2 - (2h/v)^2\ . \tag{A-2}$$

Eliminating t_h from equation (A-1) produces

$$\frac{x^2}{1+\left(2h/vt_n\right)^2} + y^2 + z^2 = \left(vt_n/2\right)^2\ . \tag{A-3}$$

The change of variable,

$$\chi^2 = \frac{x^2}{1+\left(2h/vt_n\right)^2}\ , \tag{A-4}$$

compresses the ellipsoid (A-3) to the sphere

$$\chi^2 + y^2 + z^2 = \left(vt_n/2\right)^2\ , \tag{A-5}$$

which is the zero-offset migration of a spike at the NMO time t_n. Zero-offset migration has the dispersion relation

$$\left(\frac{v}{2}\right)^2 \left(k_\chi^2 + k_y^2 + k_z^2\right) = \omega_n^2 \tag{A-6}$$

Inserting the Fourier-transform of change of variables (A-4),

$$k_\chi^2 = k_x^2 \left[1 + \left(2h / v t_n \right)^2 \right], \tag{A-7}$$

into equation (A-6) gives the important result,

$$\left(\frac{v}{2} \right)^2 \left(k_x^2 + k_y^2 + k_z^2 \right) = \omega_n^2 - \left(\frac{h k_x}{t_n} \right)^2 . \tag{A-8}$$

Which is the dispersion relation for pre-stack full migration.

Pre-stack partial migration

Pre-stack full migration (Figure (A-1)) can be done by,

(1) NMO.

(2) Pre-stack partial migration (DMO):

Substituting,

$$\omega_0^2 = \omega_n^2 - \left(\frac{h k_x}{t_n} \right)^2 , \tag{A-9}$$

in the right hand side of equation (A-8). Equation (A-9) is time dependent: ω_n and t_n appear together, it is velocity independent (but constant velocity was assumed).

(3) Zero-offset migration:

$$(v/2)^2 \left(k_x^2 + k_y^2 + k_z^2 \right) = \omega_0^2 . \tag{A-10}$$

Inverse DMO

The relation (A-9) can be used to derive the relation between the zero-offset section m and the common-offset (after normal move-out) section d.

Starting from the Fourier transform relation

$$d(t, k_x, h) = \int d\omega_n \, e^{-i \omega_n t_n} \, d(\omega_n, k_x) , \tag{A-11}$$

changing variables ω_n to ω_0, we have

$$= \int d\omega_0 \left| \frac{d\omega_n}{d\omega_0} \right| e^{-i \omega_n [\omega_0, t_n, h, k_x] t_n} \, d(\omega_n [\omega_0, t_n, h, k_x], k_x, h)$$

$$= \int d\omega_0 \left| \frac{d\omega_n}{d\omega_0} \right| e^{-i \omega_n [\omega_0, t_n, h, k_x] t_n} \, m(\omega_0, k_x) . \tag{A-12}$$

Defining

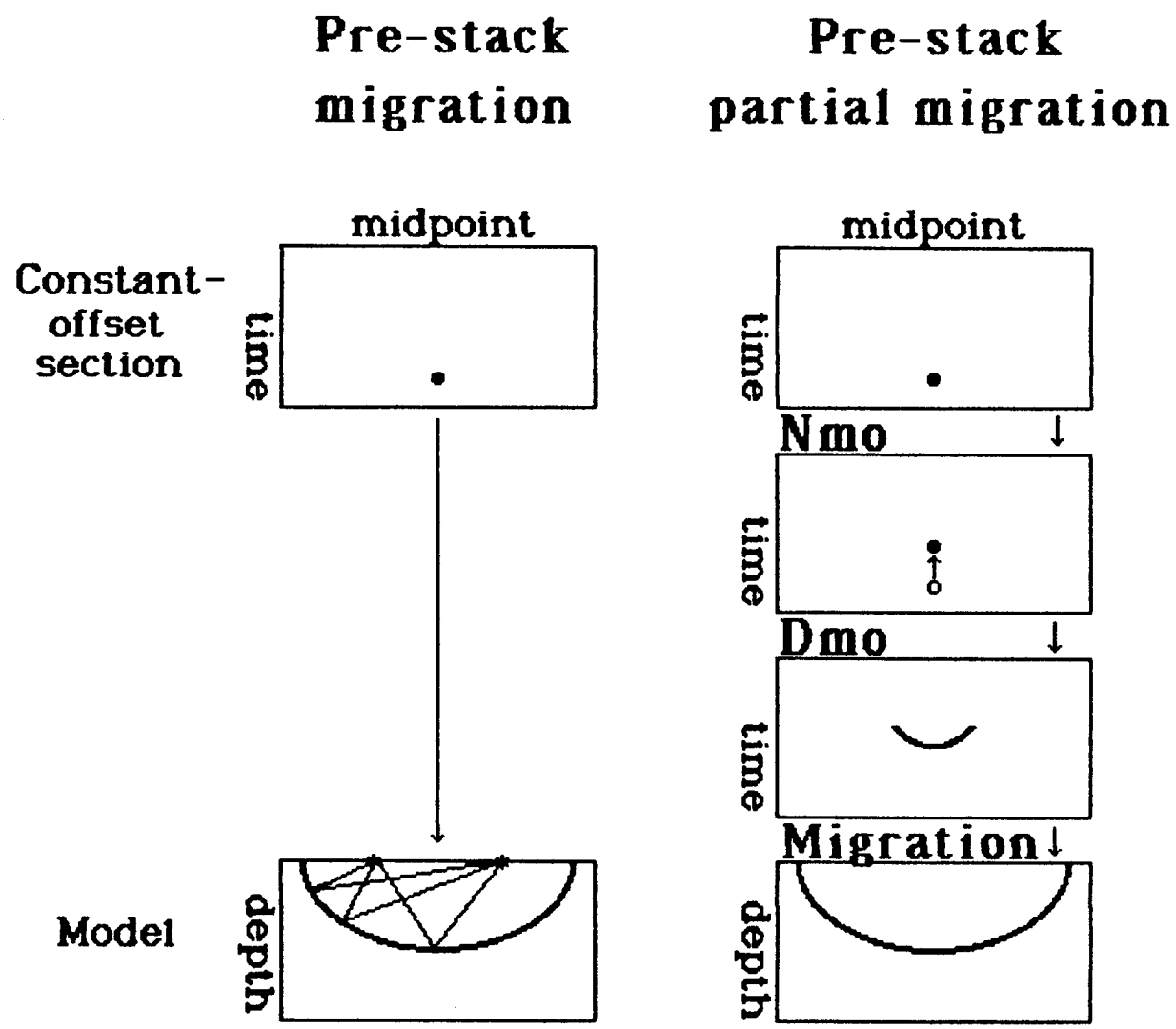


FIG. A-1. Decomposition of pre-stack full migration to NMO, DMO, and post-stack migration.

$$A = \sqrt{1 + [hk_x / \omega_0 t_n]^2}\,, \tag{A-13}$$

we have, from equation (A-9),

$$\omega_n = A\,\omega_0\,, \tag{A-14}$$

and

$$\left| \frac{d\,\omega_n}{d\,\omega_0} \right| = \left| \frac{\omega_0}{\omega_n} \right| = A^{-1}\,. \tag{A-15}$$

Substituting equations (A-14) and (A-15) into (A-12), we get

$$d\,(t_n\,,k_x\,,h\,) = \int d\,\omega_0\,A^{-1}\,e^{-i\,\omega_0 A t_n}\,m\,(\omega_0,k_x\,)\;.$$

This is the pseudo-inverse of DMO.

Impulse response

Deregowski and Rocca (1981) showed that DMO will smear an impulse at $(x=0,t=t_n\,)$ to the ellipse

$$t_0^{\,2}\,(x\,) = t_n^{\,2}\left(1-\frac{x^{\,2}}{h^{\,2}}\right) \tag{A-16}$$

The inverse-DMO will therefore smear an impulse at $(x=0,t=t_0)$, to the curve,

$$t_n^{\,2}(x\,) = \frac{t_0^{\,2}}{\left(1-\dfrac{x^{\,2}}{h^{\,2}}\right)}\;. \tag{A-17}$$

The DMO and inverse DMO impulse responses are shown in Figure (A-2). Applications of inverse DMO to Figure (A-2a) ($\mathbf{D^{+}D}\approx\mathbf{I}$), and DMO to Figure (A-2b) ($\mathbf{DD^{+}}\approx\mathbf{I}$), are shown in Figure (A-3).

Conclusion of appendix A

The DMO and its pseudo-inverse are given by,

DMO	$m\,(\omega) = \int dt\;\; A^{-1}\,e^{-i\,\omega A t}\,d\,(t\,,h\,)$
Inverse-DMO	$d\,(t\,,h\,) = \int d\,\omega\,A^{-1}\,e^{\,i\,\omega A t}\,m\,(\omega)$

The DMO is a unitary-like operator since its pseudo-inverse is also its complex conjugate.

APPENDIX B: DMO AND OFFSET CONTINUATION

Abstract

Dip move-out by offset continuation (Bolondi et. al., 1983) and this thesis assume that the output of the DMO continues every nonzero-offset section to the zero-offset section,

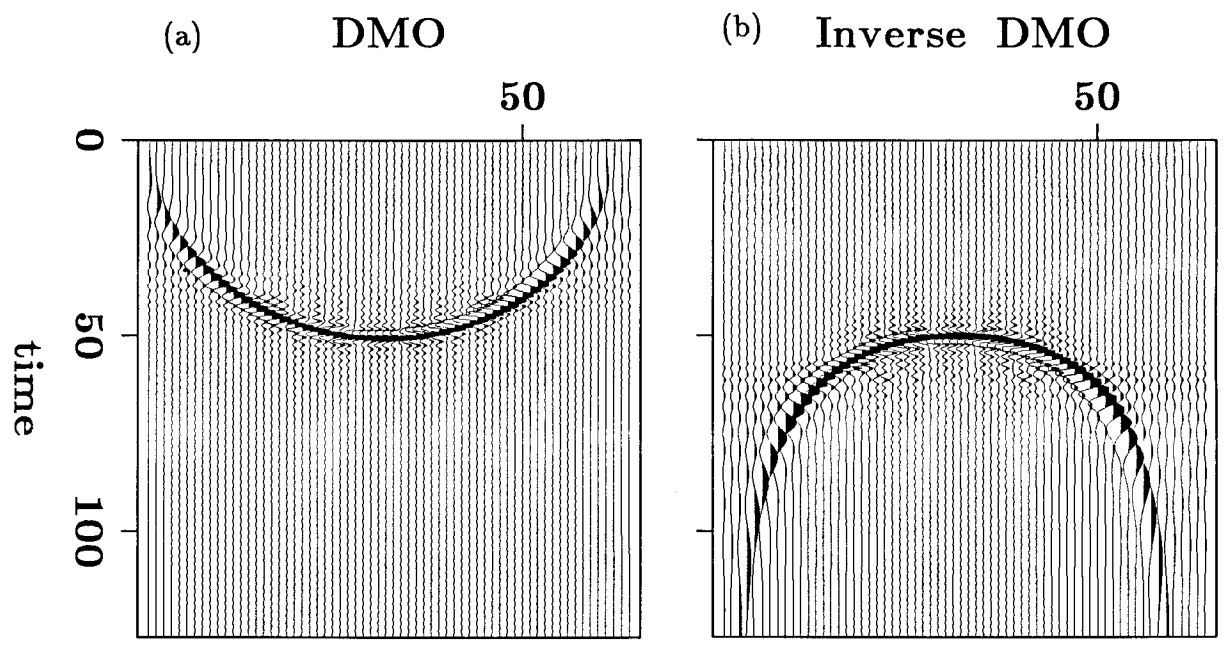


FIG. A-2. (a) DMO impulse response. (b) Inverse-DMO impulse response.

$$\mathbf{m} = \mathbf{D}_h \, \mathbf{d}_h \quad . \tag{B-1}$$

$\mathbf{d}_h$ is a common-offset section after NMO. $\mathbf{m} = \mathbf{d}_0$ is the zero-offset section, $\mathbf{D}_h$ is the DMO operator.

Hale (1984) made a less restrictive assumption that the zero-offset section is the sum over all offsets after DMO,

$$\mathbf{d}_0 = \mathbf{m} = \int dh \, \mathbf{D}_h \, \mathbf{d}_h \quad . \tag{B-2}$$

To what extent are equations (B-1) and (B-2) equivalent?

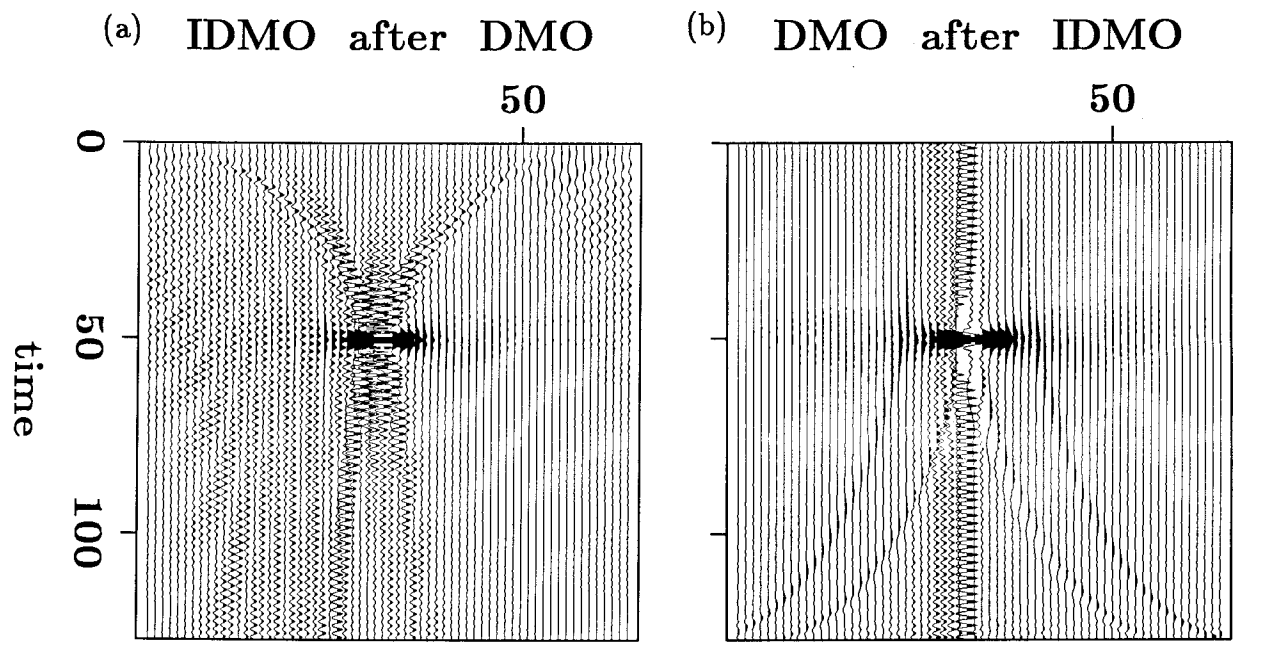
If (B-1) then (B-2)?

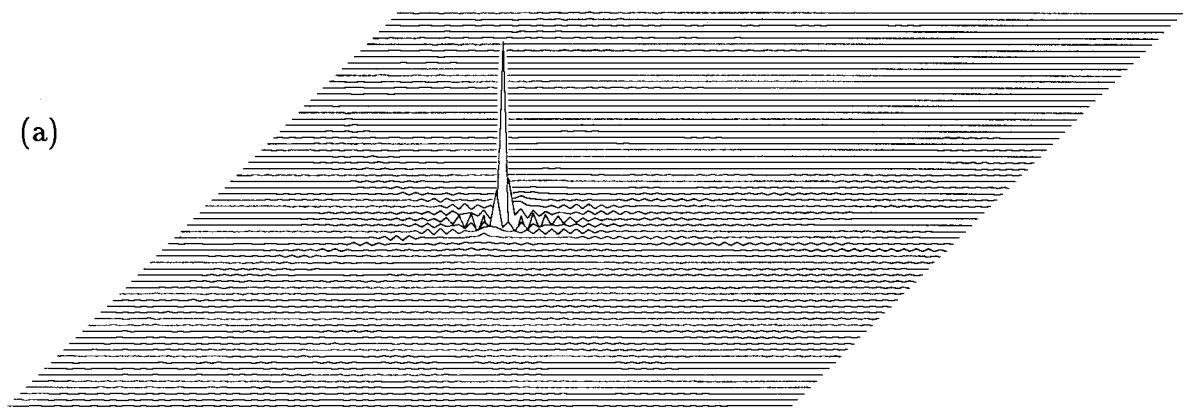
This is easy to show. If equation (B-1) is correct, then

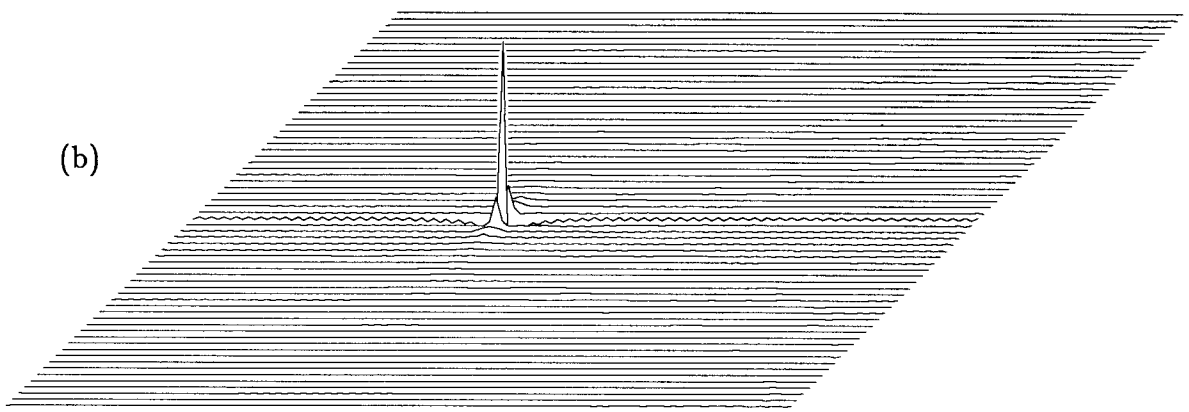
$$\int dh \, \mathbf{D}_h \, \mathbf{d}_h \; = \int dh \, \mathbf{m} = \mathbf{m} \int dh \; = \mathbf{m} \; . \tag{B-3}$$

(Stacking is assumed to be with normalized weights).

FIG. A-3. (a) Inverse-DMO after DMO of an impulse. (b) DMO after inverse-DMO of an impulse.







If (B-2) then (B-1)?

It is sufficient to show that there is an operator $\mathbf{D}_h^+$ that can produce any nonzero-offset section $\mathbf{d}_h$ from the zero-offset section $\mathbf{m}$,

$$\mathbf{d}_h = \mathbf{D}_h^+ \mathbf{m} \ . \tag{B-4}$$

If $\mathbf{D}_h^+$ exists then $\mathbf{D}_h$ of (B-1) is its pseudoinverse. The singularity of these operators means that $\mathbf{d}_h$ should be in the range of $\mathbf{D}_h^+$. (Similarly we ignore evanescent energy in migration)

Assuming (B-2), let's see if $\mathbf{D}^+$ as defined in Table A-1 can produce (B-1),

$$\begin{aligned} \mathbf{D}_h^+ \mathbf{m} &= \mathbf{D}_h^+ \int dh' \ \mathbf{D}_{h'} \ \mathbf{d}_{h'} \qquad \text{(B-5)} \\ &= \int d\omega A^{-1} e^{i\omega A t} \int dh' \int dt' \ A'^{-1} e^{-i\omega A' t'} \ d(h', k, t') \\ &= \int dh' \int dt' \left\{ \int d\omega A^{-1} A'^{-1} e^{i\omega [A t - A' t']} \right\} d(h', k, t') \end{aligned}$$

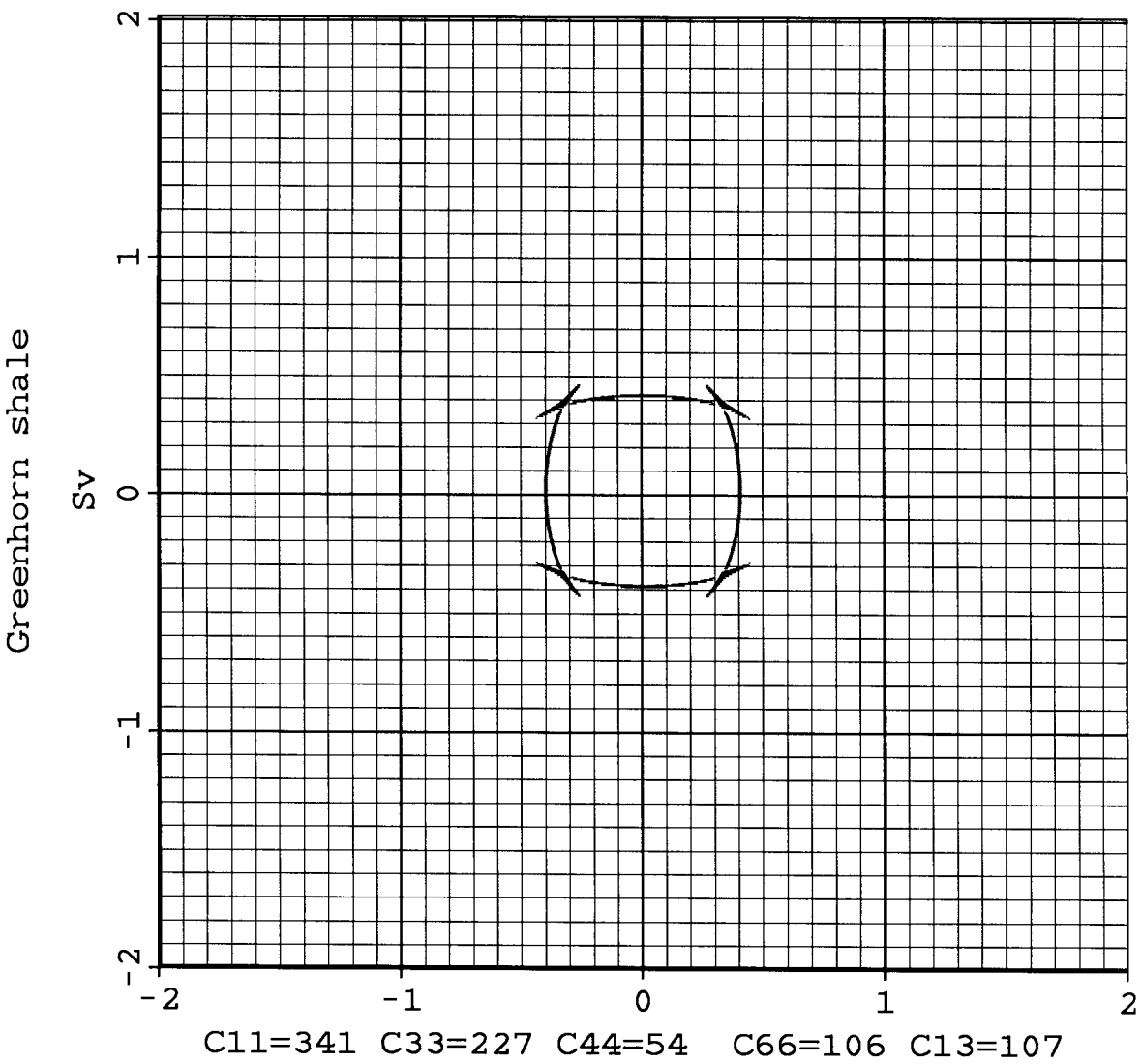
where $A' = \sqrt{1+(kh'/\omega t')^2}$. To show that $\mathbf{D}_h^+ \mathbf{m}$ is $d(h, k, t)$, I need to show that

$$\int d\omega A^{-1} A'^{-1} e^{i\omega [A t - A' t']} = \delta(h - h') \delta(t - t') \tag{B-6}$$

This integral is not trivial. However, we can see immediately that the phase of the integrand is zero for all ω when $h = h'$ and $t = t'$. The invertibility of the DMO operator was shown in Figure (A-3).

Conclusion of appendix B

The assumptions (B-1) and (B-2) are consistent with each other and they are equivalent in practice.


Greenhorn shale
Sv
2
1
0
-1
-2
-2
-1
0
1
2
C11=341 C33=227 C44=54 C66=106 C13=107