

Seismic trace interpolation in the F - X domain

S. Spitz*

ABSTRACT

Interpolation of seismic traces is an effective means of improving migration when the data set exhibits spatial aliasing. A major difficulty of standard interpolation methods is that they depend on the degree of reliability with which the various geological events can be separated. In this respect, a multichannel interpolation method is described which requires neither a priori knowledge of the directions of lateral coherence of the events, nor estimation of these directions.

The method is based on the fact that linear events present in a section made of equally spaced traces may be interpolated exactly, regardless of the original spatial interval, without any attempt to determine their true dips. The predictability of linear events in the f - x domain allows the missing traces to be expressed as the output of a linear system, the input of which consists of the recorded traces. The interpolation operator is obtained by solving a set of linear equations whose coefficients depend only on the spectrum of the spatial prediction filter defined by the recorded traces.

Synthetic examples show that this method is insensitive to random noise and that it correctly handles curvatures and lateral amplitude variations. Assessment of the method with a real data set shows that the interpolation yields an improved migrated section.

INTRODUCTION

The performance of any multichannel data processing depends heavily on the spatial sampling interval. In particular, if the frequency content of the seismic traces is adequate for the resolution of thin beds, too large a spatial sampling interval leads to aliasing which adversely affects migration, resulting in poor lateral resolution of the subsurface image. It is well known that an alternative to expensive spatial sampling is interpolation of the seismic traces. Thus,

the ability to generate unrecorded samples from undersampled data, without affecting the frequency bandwidth necessary for vertical resolution, has a direct influence on the cost of a seismic survey.

In order to interpolate seismic traces recorded with spatial aliasing, some extra information is needed to overcome the ambiguities imposed by the sampling theorem. The usual assumption is that the seismic traces of the original section (or, more precisely, the traces pertaining to a spatio-temporal processing gate) are made of a limited number of linear events. Whether this assumption is correct or not depends on the data, on the size of the spatio-temporal processing gate and, ultimately, on the ability of the chosen method to generate an output free of artifacts and of spatial aliasing.

Because knowledge of the dips allows unambiguous unraveling of aliased linear events, the standard interpolation techniques are mostly focused on the problem of finding the true directions of the coherent events in the input section (see, for instance, Bardan, 1987). These directions usually result from a local dip search performed in small spatio-temporal gates. The desired samples are generated afterwards, by interpolating the amplitudes along the direction of lateral coherence of each event. The local dip search may amount to a scan of the multichannel coherence function and may involve several dips, or a single dip (Larner et al., 1981) with obvious setbacks in the case of crossing events. However, since the spatio-temporal processing gate consists of only a few traces, in order to allow for curved events, such an automatic search technique may provide misleading directions if noise is present. Unless local dips are picked manually, as proposed by Pieprzak and McClean (1988), some a priori information, such as local dip limits, has to be supplied. It is also unclear how the model-free pattern-recognition technique proposed by King et al. (1984) will handle noisy data formed of aliased events produced by complex tectonics.

This paper takes the position that the trace interpolation should be carried out without any attempt to solve the difficult problem of separating geological reflectors. To this effect, a method to interpolate linear events without any reference to their true dips is developed in the next section.

Manuscript received by the Editor January 24, 1990; revised manuscript received November 27, 1990.

*Compagnie Generale de Geophysique, Massy, France.

©1991 Society of Exploration Geophysicists. All rights reserved.

For the sake of clarity, the multichannel properties of such events are described in the appendices. Numerical examples on synthetic and real data sets demonstrate the method in the third section.

INTERPOLATION OF LINEAR EVENTS

It is shown in this section that N equally spaced traces made of L events (each event being invariant from trace to trace except for its constant dip) may be exactly interpolated without any reference to the dips, if L is known and if it is less than N . Using the same notation as in Appendix A, each input trace g_k may be modeled in the frequency domain as follows:

$$g_k(f) = \sum_{j=1}^L a_j(f) z_j^{k-1}(f), \quad k = 1, \dots, N, \quad (1)$$

where $a_j(f)$ is the Fourier transform of the wavelet associated with the event j and where $z_j(f)$ is the j th event phase shift at the frequency f , $z_j(f) = \exp(2\pi i f p_j)$, corresponding to the time shift p_j between adjacent traces. The forward-backward one-step prediction filter components $P_j(f)$ (see Canales, 1984 and Appendix A) may be determined in the least-squares sense from the following set of equations:

$$g_k(f) = \sum_{j=1}^L P_j(f) g_{k-j}(f), \quad k = L+1, \dots, N, \quad (2a)$$

$$g_k^*(f) = \sum_{j=1}^L P_j^*(f) g_{k+j}^*(f), \quad k = 1, \dots, N-L, \quad (2b)$$

where * denotes complex conjugate.

The interpolation considered throughout this section halves the original trace interval (first order interpolation); the results can be generalized easily to higher order interpolations. If this first order interpolation is exact, each Fourier transformed trace of the output section, $g_k^i(f)$, allows an expansion in terms of the L phase shifts similar to equation (1):

$$g_k^i(f) = \sum_{j=1}^L a_j(f) z_j^{k-1}(f), \quad k = 1, \dots, 2N-1. \quad (3)$$

The forward-backward prediction filter components $P_j^i(f)$ express the traces $g_k^i(f)$ of the output section in terms of the preceding L traces and of the following L traces:

$$g_k^i(f) = \sum_{j=1}^L P_j^i(f) g_{k-j}^i(f), \quad k = L+1, \dots, 2N-1, \quad (4a)$$

$$g_k^i(f) = \sum_{j=1}^L P_j^{i*}(f) g_{k+j}^i(f), \quad k = 1, \dots, 2N-L-1. \quad (4b)$$

The expressions (4) provide a relationship between the missing traces and the known traces. This relationship may be formally expressed as follows:

$$\underline{\mathbf{A}}(\mathbf{P}'(f)) \begin{bmatrix} g_2^i(f) \\ g_4^i(f) \\ \vdots \\ g_{2N-2}^i(f) \end{bmatrix} = \underline{\mathbf{B}}(\mathbf{P}'(f)) \begin{bmatrix} g_1^i(f) \\ g_3^i(f) \\ \vdots \\ g_{2N-1}^i(f) \end{bmatrix}. \quad (5)$$

The notation $\underline{\mathbf{A}}(\mathbf{P}'(f))$, $\underline{\mathbf{B}}(\mathbf{P}'(f))$ emphasizes the fact that these two matrices depend only on the components of the prediction vector $\mathbf{P}'(f)$. If all the components of the vector \mathbf{P}' are known on the signal bandwidth, expression (5) represents a system of $2(2N-L-1)$ equations with $2N-1$ unknowns, which can be solved in the usual least squares sense. From a formal point of view, expression (5) defines the missing traces as the output of a linear system defined by the multichannel filter:

$$[\underline{\mathbf{A}}^+(\mathbf{P}'(f))\underline{\mathbf{A}}(\mathbf{P}'(f))]^{-1}\underline{\mathbf{A}}^+(\mathbf{P}'(f))\underline{\mathbf{B}}(\mathbf{P}'(f))$$

and by the known traces. Here $\underline{\mathbf{A}}^+$ stands for $\underline{\mathbf{A}}$ transpose and complex conjugate.

Obviously, the prediction vector $\mathbf{P}'(f)$ is unknown but can be determined from the input traces prediction filter $\mathbf{P}(f)$. To proceed, note that the phase shifts in the output and input sections are related by $z_j^i(f) = z_j(f/2)$, since the interpolation sought consists of dividing the original trace interval by a factor of 2. Moreover, the analytical expressions (A-5) of a given component of $\mathbf{P}'(f)$ or $\mathbf{P}(f)$ prediction filters involve only products of the corresponding phase shifts. It follows that the spectrum of $\mathbf{P}'(f)$ is determined completely:

$$P_j^i(f) = P_j(f/2), \quad j = 1, \dots, L. \quad (6)$$

The above expression defines a spectral estimate of the components of the prediction filter $\mathbf{P}'(f)$ for which the low frequency components of the prediction vector $\mathbf{P}(f)$ play a key role. It follows that the higher the interpolation order, the more the low-frequency content of the input traces drives the interpolation. If the input data are band limited, the remark does not imply that the high-frequency content of the output traces is affected, because the input and output in expression (5) are at the same frequency. However, for a given time window, the estimation of $\mathbf{P}(f/2)$ (for instance by zero padding in the time domain) may be misleading, owing to the lower limit of the data spectrum. Indeed, if f_1 denotes this lower limit, the expressions (5) and (6) involve the estimation of the \mathbf{P} vector components at the frequency $f_1/2$, which is out of the bandwidth. In order to overcome this difficulty, it is shown in Appendix B that despite possible limitation of the signal bandwidth, the L spectra $P_1(f), \dots, P_L(f)$ may be defined on the full frequency band, using a forward and backward prediction technique along the frequency axis, for the vector \mathbf{P} .

The sequence (2) [possibly (B-3)], (6), (5) defines a procedure for an exact interpolation of linear events. One advantage of the method is that no information on the dips is required, expressions (5) and (6) being an alternative to a multidip search or to a broad-band solution of the nonlinear optimization problem defined by (A-10).

NUMERICAL EXAMPLES

The main difficulty when applying the method to real data (or to synthetic data, where some kind of noise is present) is the choice of the “length” L of the prediction vector \mathbf{P} . This is very similar to the problem of determining the order of an AR process. The position taken in this paper is that a trial with L amounting to about half the number of traces pertaining to the spatio-temporal processing gate gives a fair indication of whether L should be increased/decreased, or the processing gate changed. It is obvious, however, that a better procedure would be the introduction of some objective criterion to select, on the signal bandwidth, the optimal L value for each processing gate.

A common feature of multichannel adaptive processing sequences is their relative insensitivity to the presence of random noise. This is not surprising, since these sequences take advantage of signal redundancy in order to attenuate the

noise component. In this respect, the proposed method involves two minimizations, in the least-squares sense, of the random noise energy: the first minimization occurs when the prediction vector \mathbf{P} is determined from the input traces [expressions (2)] and the second when the interpolated traces are constructed [expression (5)]. Figures 1 and 2 illustrate the ability of the algorithm to generate unrecorded samples from spatially undersampled data, in the case of aliased linear events when random noise is present.

The true advantage of a multichannel interpolation method lies in its ability to process a large number of traces in order to account for the low spatial frequencies of the data. The number of traces processed together is limited by the ability of the particular chosen algorithm to handle curved horizons and lateral amplitude variations. In this respect, to consider curved events as made piece-wise of linear events can only limit the processing to a small number of traces and therefore

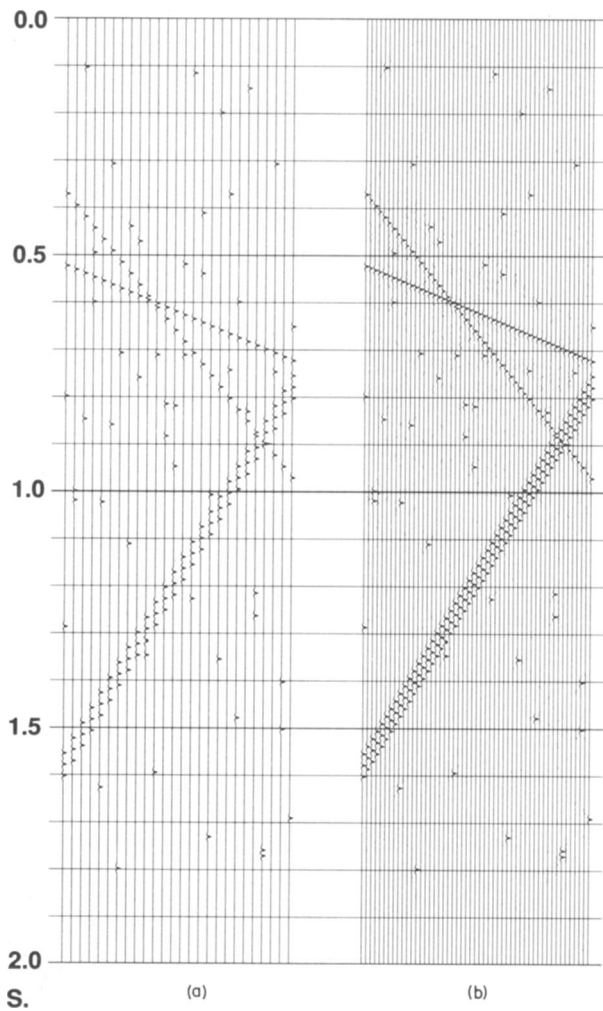


FIG. 1. (a) Input data made of three aliased linear events and sparse random noise. The noise and the signal have the same spectrum. (b) First order interpolated data. The interpolation has been carried out using a spatio-temporal processing gate which covered the entire input. The order of the prediction filter was $L = 3$. Note that the parallel events have been correctly interpolated.

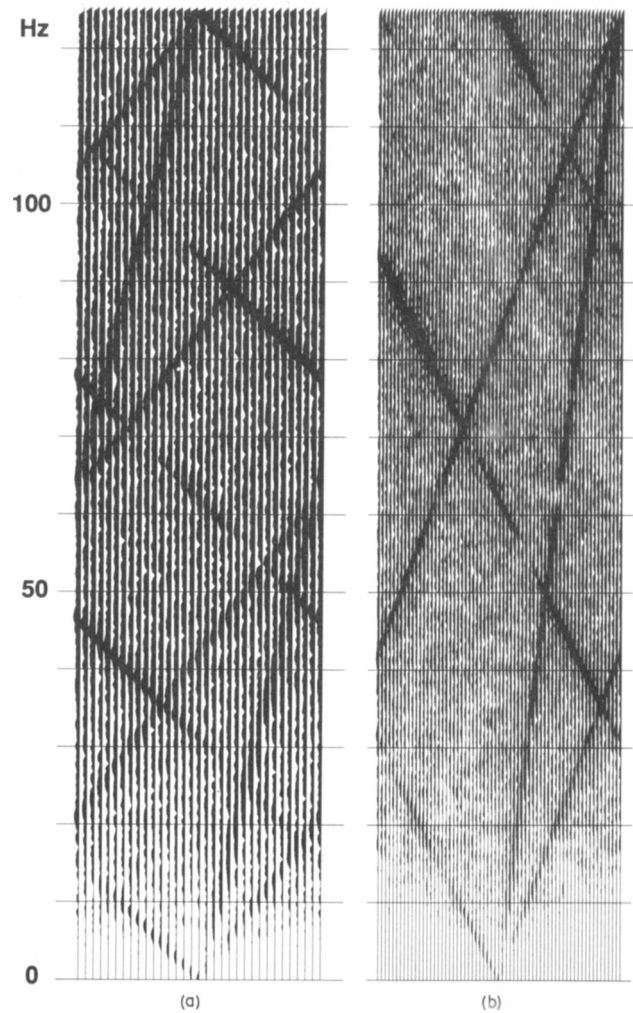


FIG. 2. Amplitude f - k spectra. The horizontal axes represent normalized wavenumbers, from -0.5 to 0.5 cycles. (a) f - k amplitude spectrum of the input shown in Figure 1(a). The unit scale of the normalized wavenumber is $1/32$ cycles. (b) f - k amplitude spectrum of the interpolated data set shown in Figure 1(b). The unit scale of the normalized wavenumber is $1/64$ cycles. Note that the event in Figure 1(a) beginning at 525 ms is no longer aliased.

to an even smaller number of dips. The underlying idea of the method presented in this paper is that it represents a suboptimal solution to the problem of linear events interpolation. Indeed, the one-step predictability in the f - x domain, which is the basis of the method, may be seen as a necessary, but not sufficient, condition for the signal to be formed of linear events. It follows from the discussion in Appendix A that L linear events on N traces are predictable in the f - x domain with filters of length L which involve lags higher than 1. Moreover, the components of a lag- n prediction filter are nonlinearly related to the components of the lag-1 prediction filter. The quality of the interpolation in Figure 3 shows that the proposed algorithm, based on lag-1 prediction, may accommodate curvatures and strong lateral amplitude variations.

The interpolation method is checked on a part of a stacked section from the North Sea (Figure 4). The processing input (Figure 5) consists of a second order decimation of the reference, and the restored section (Figure 6) is compared in Figure 7 to the original data set. Because interpolations are not carried out for their own sake, the influence of the

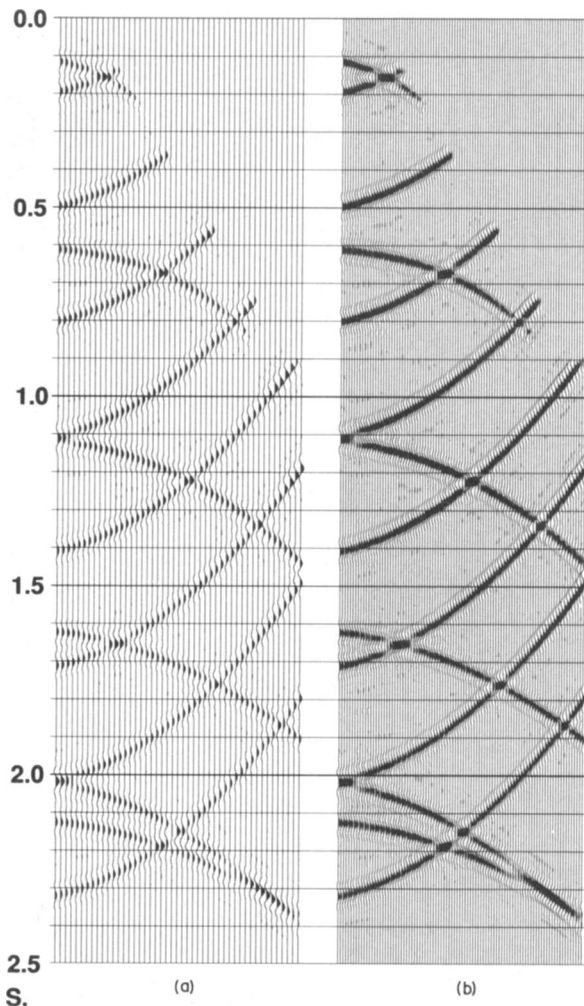


FIG. 3. (a) Input data set made of curved events displaying lateral amplitude variations. (b) First order interpolated data. The interpolation has been carried out using a spatio-temporal processing gate which covered the entire input. The order of the prediction filter was $L = 15$.

processing on migration is also examined. The original migrated section (Figure 8) and the migrated restored section (Figure 10) are compared in Figure 11. The choice of the finite-difference scheme to perform migration was motivated mainly by its more severe requirements with respect to aliasing (see Figure 9).

CONCLUSIONS

This paper has described a multichannel, model-free interpolation method. The choice of the f - x domain follows from the fact that, in the case of a section made of linear events, any possible aliasing does not affect the components of the spatial prediction filter. The method does not address the difficult problem of separating the events present in the section to be processed. This may be seen as a definite advantage over traditional seismic interpolation techniques.

Although derived under the assumption that the input section is made of linear events, the method represents only a suboptimal solution to the problem of interpolating such a section. The numerical examples presented show that some departure from this basic assumption can be tolerated. In the presence of random noise, the method handles events with moderate curvatures which display some lateral amplitude variations, without trading vertical resolution for lateral resolution.

The method is also derived under the assumption that the spatial sampling of the input section is regular. However, interpolating an irregularly sampled section violates the method's assumptions in the same way as interpolating a regularly spaced section made of curved events displaying some lateral amplitude variations. It is then expected that a first order interpolation, for instance, will calculate an acceptable trace midway between the existing traces, even if the spatial sampling is slightly irregular. At this stage it is unclear how to modify the theory in order to interpolate and, at the same time, transform an irregular grid into a regular one.

REFERENCES

- Bardan, V., 1987, Trace interpolation in seismic data processing: *Geophys. Prosp.*, **35**, 343-358.
- Canales, L. L., 1984, Random noise reduction: Presented at the 54th Ann. Internat. Mtg., Soc. Explor. Geophys., Expanded Abstracts, 525-527.
- Gulunay, N., 1986, FXDECON and complex Wiener prediction filter: Presented at the 56th Ann. Internat. Mtg., Soc. Explor. Geophys., Expanded Abstracts, 279-281.
- Kay, S. M., and Marple Jr., S. L., 1981, Spectrum analysis—A modern perspective, *Proceedings of the IEEE*, **69**, 1380-1419.
- King, G. A., Leong, T. K., and Flinchbaugh, B. E., 1984, The role of interpolation in seismic resolution: Presented at the 54th Ann. Internat. Mtg., Soc. Explor. Geophys., Expanded Abstracts, 766-767.
- Larner, K., Gibson, B., and Rothman, D., 1981, Trace interpolation and the design of seismic surveys: *Geophysics*, **46**, 407.
- Leaney, W. S. P., and Esmersoy, C., 1989, Parametric decomposition of offset VSP wave fields: Presented at the 59th Ann. Internat. Mtg., Soc. Explor. Geophys., Expanded Abstracts, 26-29.
- Mars, J., Glangeaud, F., Lacoume, J. L., Fourmann, J. M., and Spitz, S., 1987, Separation of seismic waves: Presented at the 57th Ann. Internat. Mtg., Soc. Explor. Geophys., Expanded Abstracts, 489-492.
- Pieprzak, A. W., and McClean, J. W., 1988, Trace interpolation of severely aliased events: Presented at the 58th Ann. Internat. Mtg., Soc. Explor. Geophys., Expanded Abstracts, 658-660.
- Pisarenko, V. F., 1973, The retrieval of harmonics from a covariance function: *Geophys. J. Roy. Astr. Soc.*, **33**, 347-366.

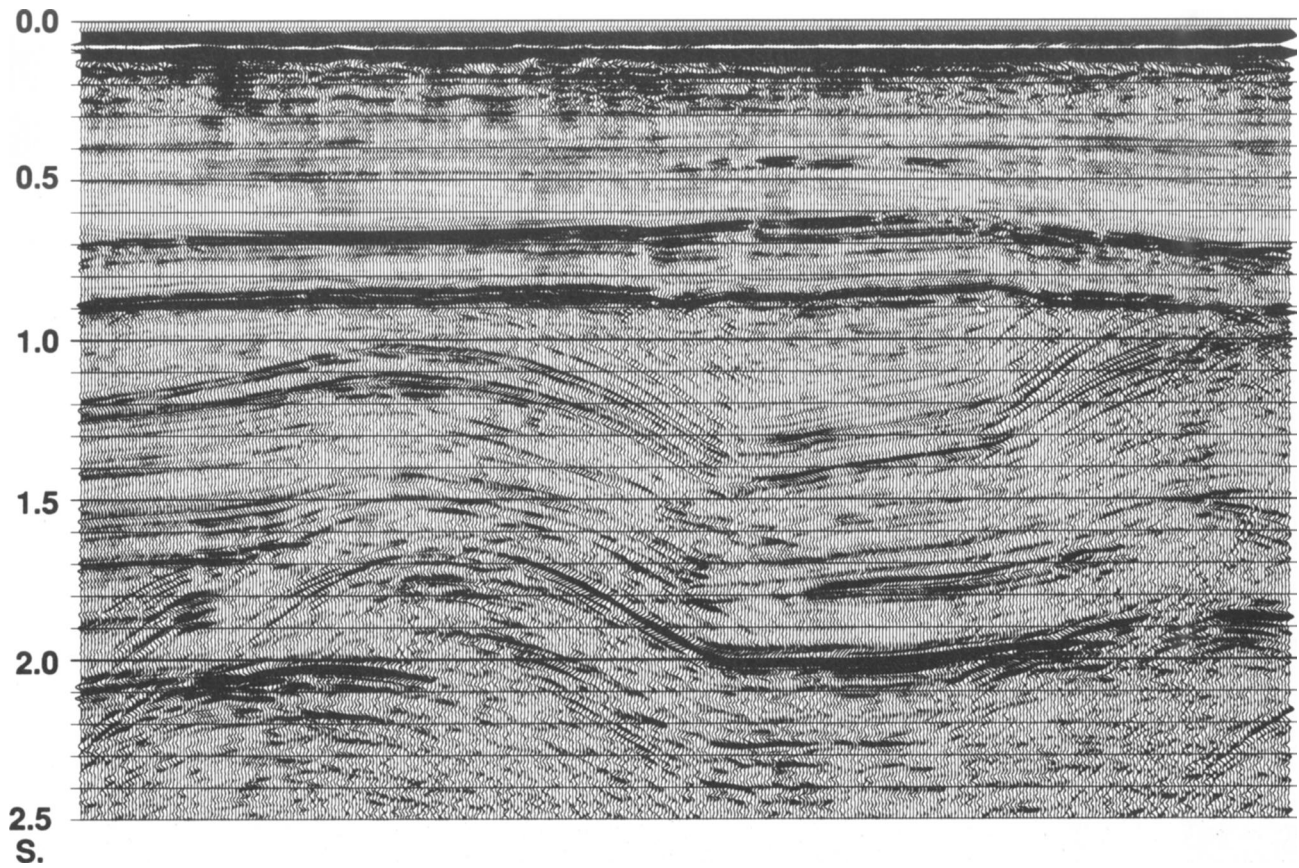


FIG. 4. Reference data set made of 352 traces. The trace interval is 25 m.

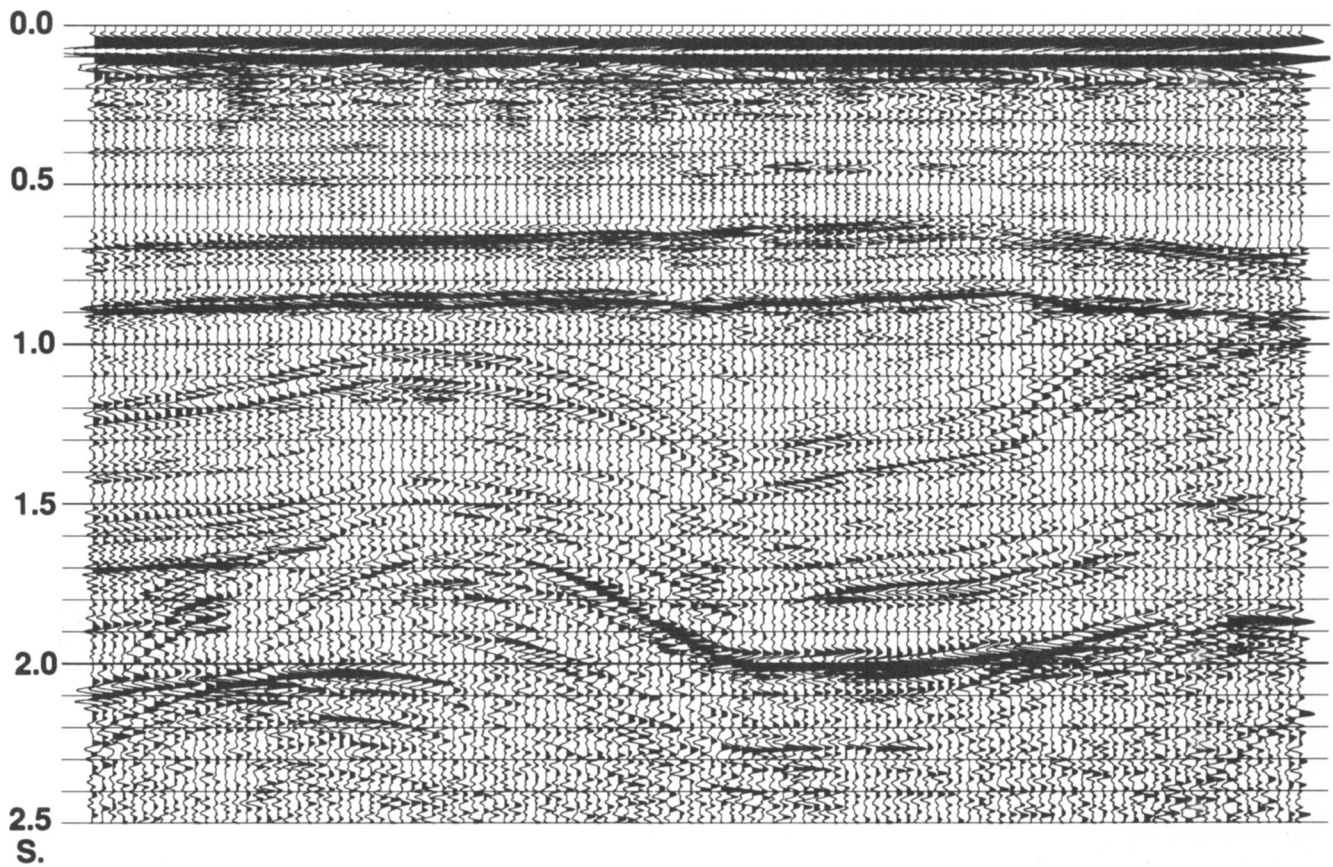


FIG. 5. Input to interpolation. The trace interval is 75 m.

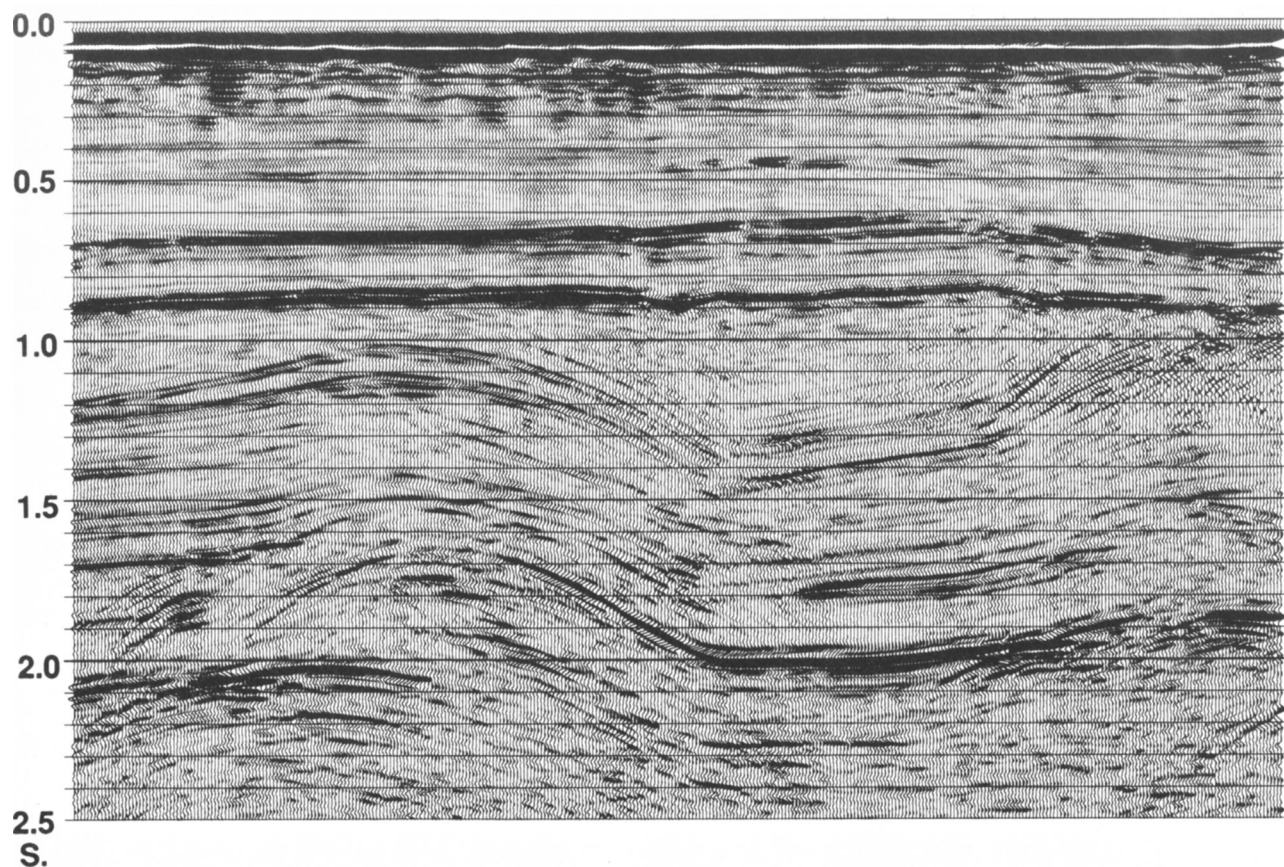


FIG. 6. Second order interpolation. The processing has been carried out using overlapping spatio-temporal gates made of 31 input traces and 256 ms. The order of the prediction filter was $L = 9$.

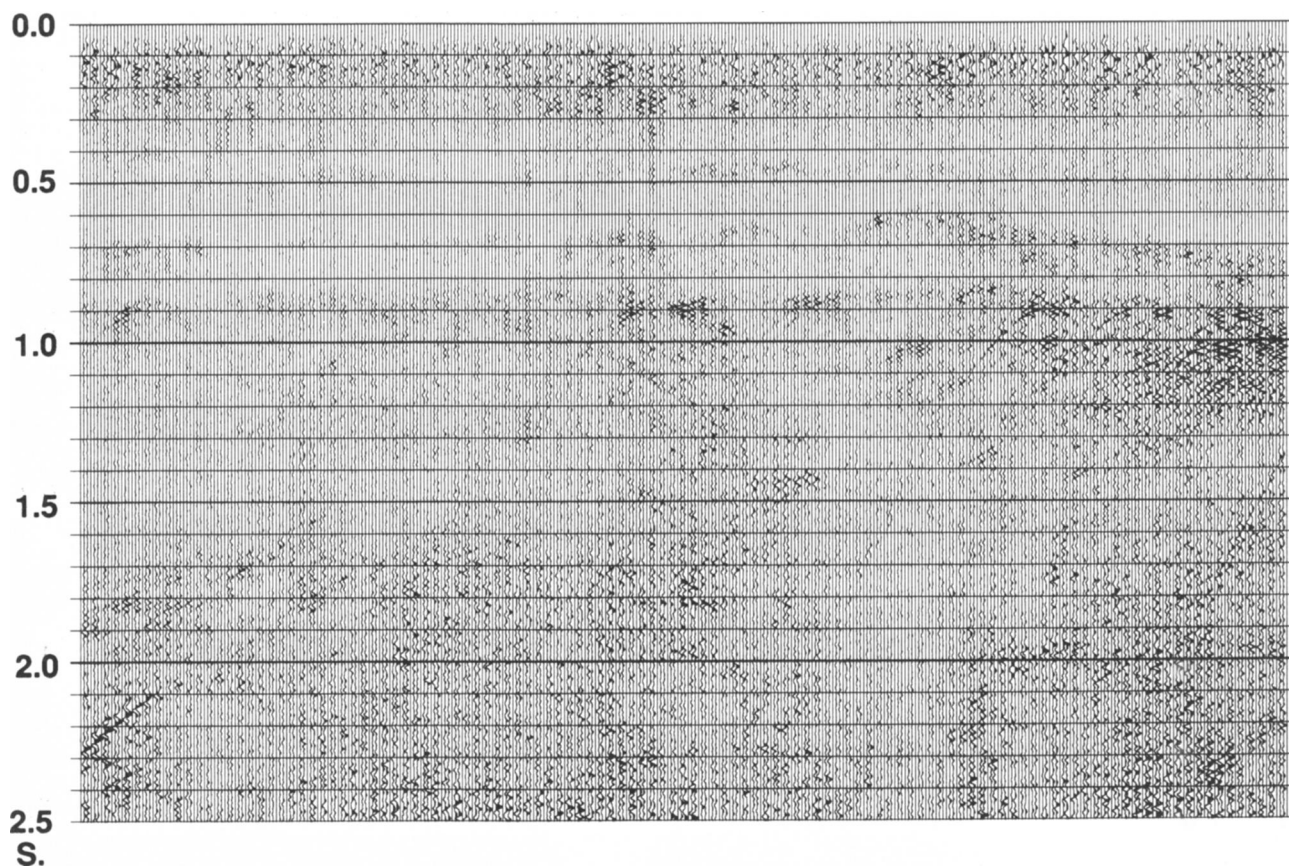


FIG. 7. Difference between the reference section (Figure 4) and the interpolated section (Figure 6). For display purposes, this section has been amplified by a factor of 1.25.

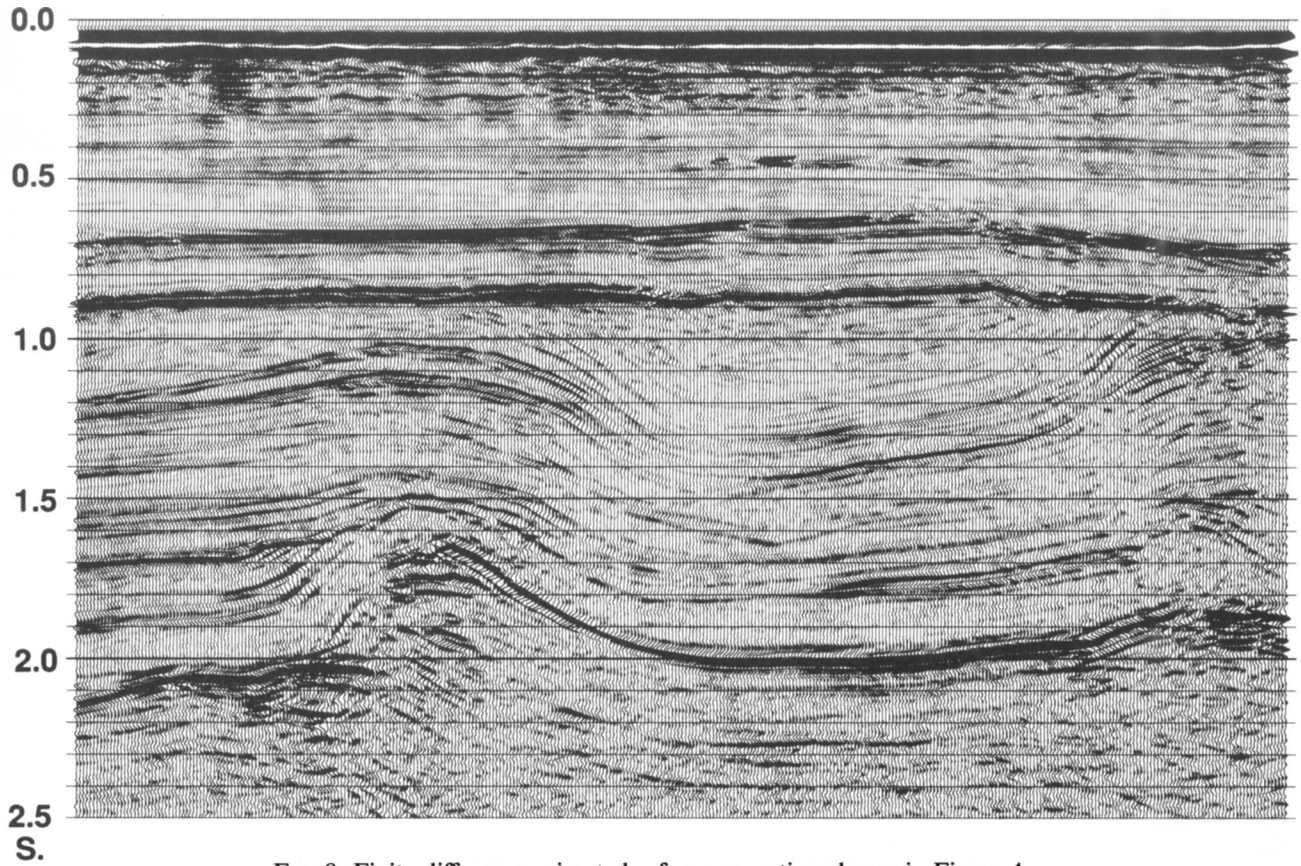


FIG. 8. Finite-difference migrated reference section shown in Figure 4.

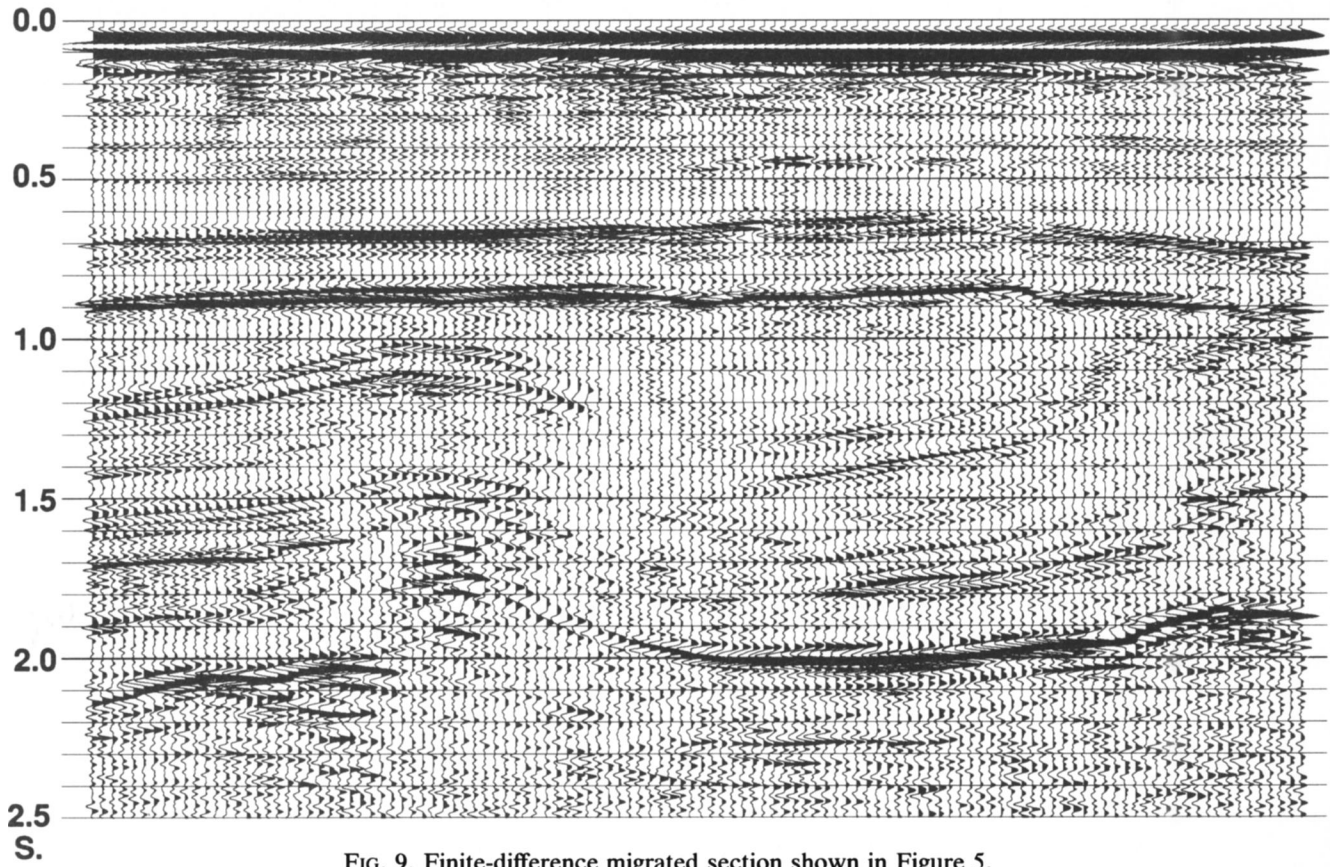


FIG. 9. Finite-difference migrated section shown in Figure 5.

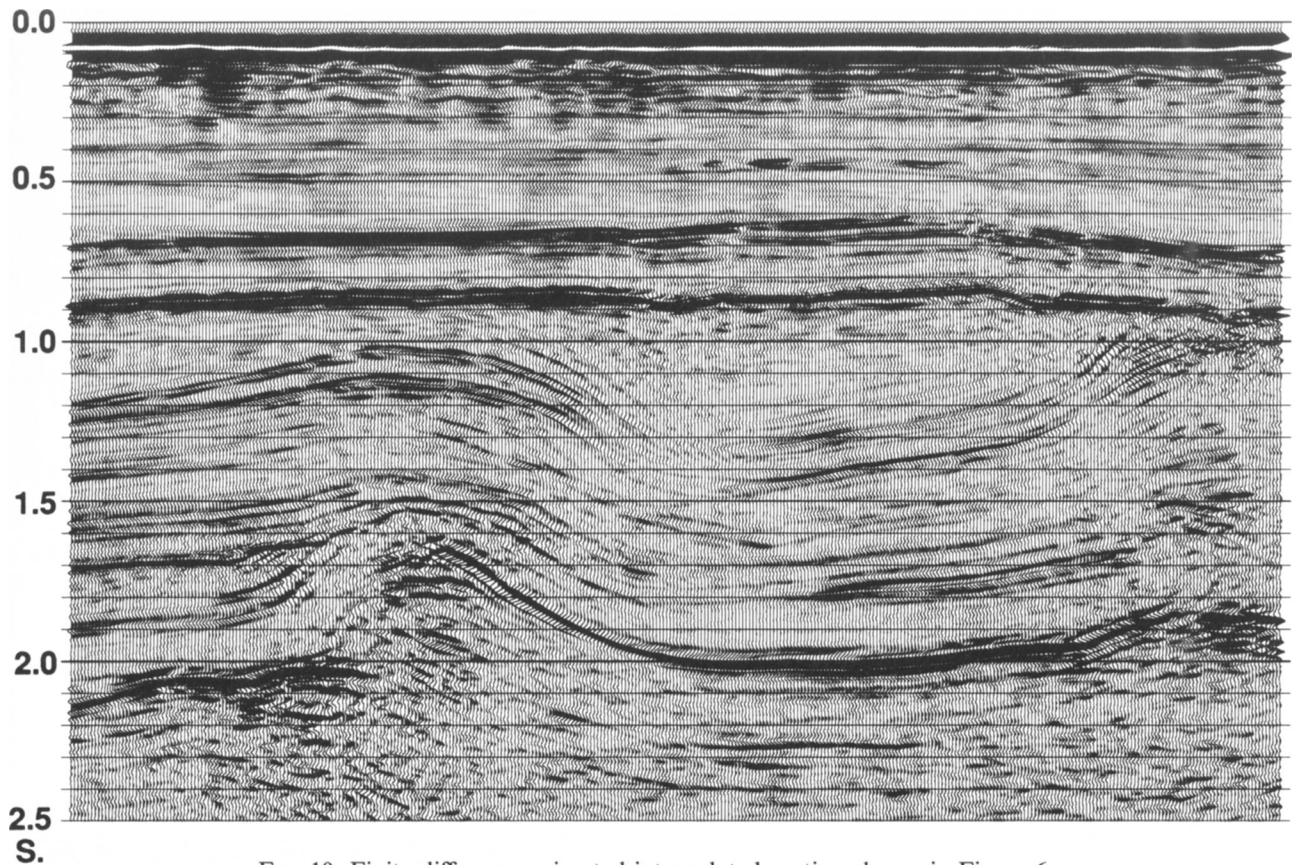


FIG. 10. Finite-difference migrated interpolated section shown in Figure 6.

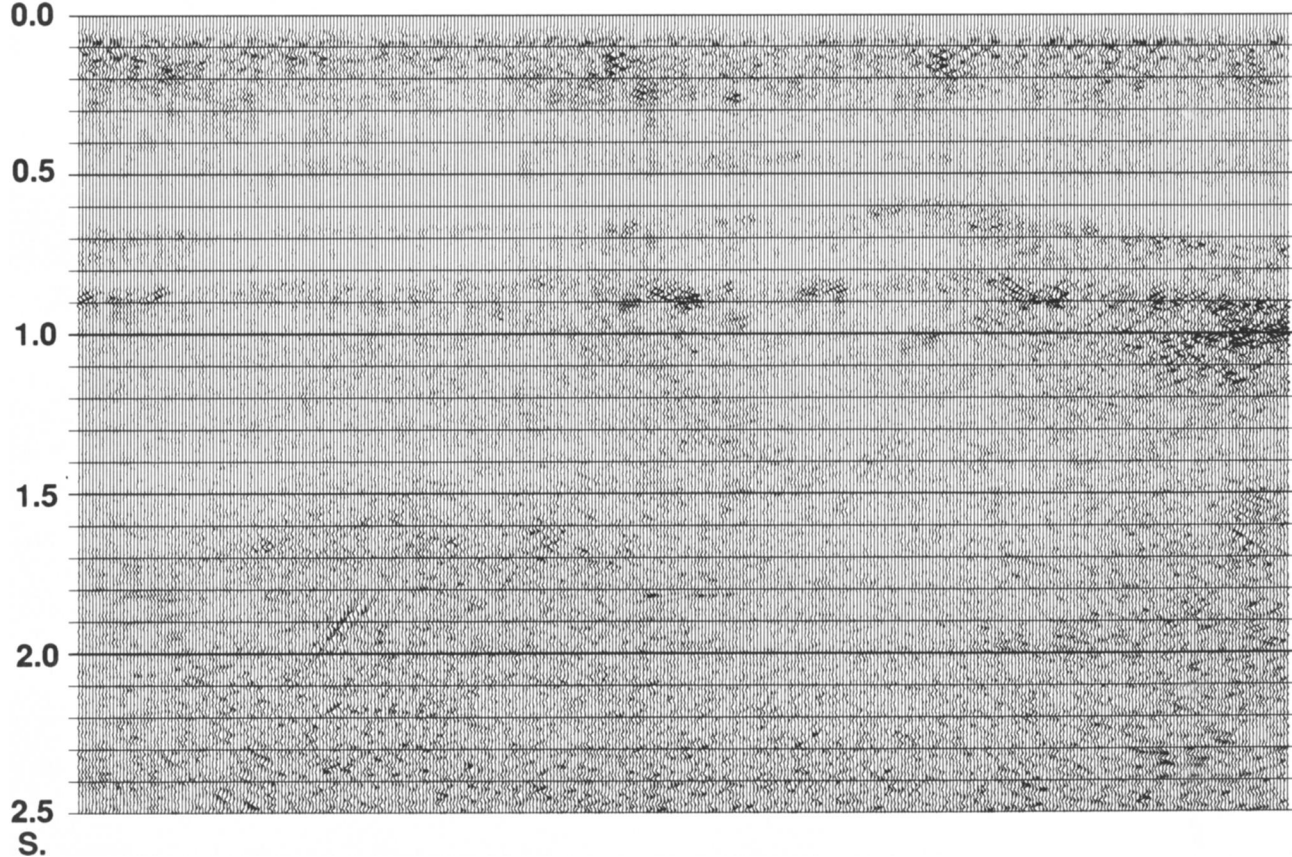


FIG. 11. Difference between the migrated reference section (Figure 8) and the migrated second order interpolated section (Figure 10). For display purposes, this section has been amplified by a factor of 1.25.

APPENDIX A

PREDICTION AND SEPARATION OF LINEAR EVENTS

Throughout this appendix, a section of N equally spaced traces is considered, whose signal content is composed of linear events. Each event is characterized by its gradient p_j , the number L of events with different gradients being less than N . In the frequency domain the N -channel datum is a vector $\mathbf{x}^T(f) = [x_1(f), \dots, x_N(f)]$, where T stands for transpose. The assumed model is $\mathbf{x}(f) = \mathbf{g}(f) + \mathbf{n}(f)$, where \mathbf{g} and \mathbf{n} denote the signal and noise vectors respectively. The noise is assumed uncorrelated on the various traces and uncorrelated with the signal. The complex z -transform along the j th event defines the transfer function $\mathbf{S}_j^T(f) = [1, z_j(f), \dots, z_j^{N-1}(f)]$, with $z_j(f) = \exp(2\pi i f p_j)$. The system of linear events which forms the signal may then be modeled as a linear combination of the L steering vectors $\mathbf{S}_j(f)$:

$$\mathbf{g}(f) = \sum_{j=1}^L a_j(f) \mathbf{S}_j(f). \quad (\text{A-1})$$

The complex coefficient $a_j(f)$ is the Fourier transform of the wavelet associated with event j . It can be very long in the time domain, as parallel arrivals are defined as a single event. No assumption is made on the $a_j(f)$ bandwidth in relation to the space interval, so that the signal may display some spatial aliasing. Defining the $N \times L$ matrix $\mathbf{S}(f) = [\mathbf{S}_1(f), \dots, \mathbf{S}_L(f)]$, the datum $\mathbf{x}(f)$ becomes, in matrix notations,

$$\mathbf{x}(f) = \mathbf{S}(f)\mathbf{a}(f) + \mathbf{n}(f). \quad (\text{A-2})$$

Since the L steering vectors are linearly independent (except for some frequencies in the broad-band, for which it may happen that some steering vectors are identical), the Van Der Monde matrix \mathbf{S} is of rank L . Any row may then be written as a linear combination of L other rows. In particular, the $(L + 1)$ th row, consisting of the steering vectors components on the $(L + 1)$ th trace, may be written as a linear combination of the first L rows as follows:

$$(z_1^L, \dots, z_L^L) = (P_L, \dots, P_1) \times \begin{bmatrix} 1 & \dots & 1 \\ z_1 & & z_L \\ \cdot & & \cdot \\ z_1^{L-1} & & z_L^{L-1} \end{bmatrix}, \quad (\text{A-3})$$

where the frequency dependence has been dropped.

The knowledge of the steering vectors determines \mathbf{P} at each frequency. Conversely, the $P_1(f), \dots, P_L(f)$ spectra determine the L phase shifts, since from (A-3) these phase shifts are the roots of the polynomial:

$$Z_L = z^L - P_1 z^{L-1} - \dots - P_L. \quad (\text{A-4})$$

The analytical expression of each component P_j in terms of the phase shifts is obtained from the expansion of the polynomial $Z_L = (z - z_1)(z - z_2) \dots (z - z_L)$:

$$\begin{aligned} P_1 &= z_1 + \dots + z_L, \\ P_2 &= -(z_1 z_2 + \dots + z_{L-1} z_L), \\ &\cdot \\ &\cdot \\ P_L &= (-)^{L+1} z_1 z_2 \dots z_L. \end{aligned} \quad (\text{A-5})$$

Each component P_j has $\binom{L}{j}$ terms, where $\binom{L}{j}$ denotes the binomial coefficient. These components are not independent, since $|P_L|^2 = 1$ and, from the above expressions:

$$P_L^* P_j = -P_{L-j}^*, \quad j = 1, \dots, L-1. \quad (\text{A-6})$$

It is instructive to see how the roots move on the unit circle as f varies. At the DC , all the roots are concentrated in $(1, 0)$. As the frequency increases, each root z_j advances clockwise or anticlockwise in steps proportional to its gradient p_j . In the process, a root may cross the real negative axis, revealing the aliasing of the corresponding event. It may also happen that, due to the aliasing, two roots become identical at some particular frequency. At such frequencies the corresponding events are indiscernable and the rank of the matrix \mathbf{S} is less than L .

Obviously, $\mathbf{P}(f)$ has to be determined from the noisy datum-vector $\mathbf{x}(f)$. Since the components of the \mathbf{P} -vector are invariant with respect to translations defined across the array of N traces, it follows from (A-3) and (A-1) that the signal component on the k th trace, $g_k(f)$, may be predicted from the signal components on the preceding traces $g_{k-L}(f), \dots, g_{k-1}(f)$. The expression (A-6) implies that this signal component may also be predicted from the signal components on the traces which follow, $g_{k+1}(f), \dots, g_{k+L}(f)$. Thus, at each frequency of the signal bandwidth, the components of the forward-backward one-step prediction error filter are the solution of the nonwindowed system of normal equations:

$$\mathbf{R} \begin{bmatrix} 1 \\ -P_1 \\ \cdot \\ -P_L \end{bmatrix} = \begin{bmatrix} e \\ 0 \\ \cdot \\ 0 \end{bmatrix}. \quad (\text{A-8})$$

In the expression above, e denotes the minimum prediction error energy. The components of the $(L + 1) \times (L + 1)$ matrix \mathbf{R} are defined by:

$$R_{ij}(f) = \sum_{k=0}^{N-L-1} (x_{k+L+2-j} x_{k+L+2-i}^* + x_{k+i} x_{k+j}^*). \quad (\text{A-9})$$

At this stage, the advantages of an economical solution of (A-8) have to be carefully weighed in the light of any possible drawbacks. If a windowed version of (A-8) is used, which leads to Levinson's recursion, the vector \mathbf{P} solution cannot correctly describe the system of linear events (Gulunay, 1986). This is due to the fact that windowing assumes null traces outside the original array of N traces, while the

invariance of vector \mathbf{P} for translations performed across this array implies that the events are present outside as well. Some drawbacks may also arise if Burg's nonwindowed algorithm is used. Indeed, it may be checked that the prediction filter components (A-5) do not verify Levinson's recursion constraint which leads to Burg's algorithm (see, for instance, Kay and Marple, 1981).

When all the components of the prediction vector are determined on the signal bandwidth, each event in the section may be separated from the others. The separation involves determining the steering vectors and estimating, in the least-squares sense, the coefficients $a_j(f)$ in (A-1). The construction of the steering vectors implies correct sorting of the roots of $Z_L(f)$ in the frequency domain. In particular, the correct identification of the roots as f varies implies that the aliasing is unraveled. However, correct sorting of the

roots may be impossible to achieve when the components of the prediction filter are to some extent corrupted by noise. An alternative consists of overall determination of the L dips on the broad band, which are solutions of the nonlinear optimization problem:

$$\sum_f |Z_L(f)|^2 \text{ minimum.} \quad (\text{A-10})$$

It is noteworthy that the problem of separating linear events may be described following other approaches. However, any approach leads to a polynomial of the type (A-4), described first by Pisarenko (1973), and to a separation which involves either sorting its roots at each frequency, or a nonlinear optimization problem on the broad band (see for instance Mars et al., 1987, and Leaney and Esmersoy, 1989).

APPENDIX B

PREDICTABILITY OF THE P -VECTOR COMPONENTS

In this appendix it is shown that the prediction vector components can be defined on the full frequency band, despite possible limitation of the signal band width. To this effect it is assumed that the coefficients $a_1(f), \dots, a_L(f)$ in expression (1) vanish outside a frequency band $f_1 \leq f \leq f_2$, defined by $f_1 = k_1 \Delta f$ and $f_2 = k_2 \Delta f$, where Δf is the frequency interval. Inside this frequency band, expression (2) defines the L components of the prediction vector \mathbf{P} . At the frequency $k \Delta f$, the phase shifts are the roots of the polynomial $Z_L(k \Delta f)$ defined in (A-4). If the prediction vector is noise-free, each root z_j moves on the unit circle in constant steps proportional to the gradient p_j , when the frequency varies. The position of the root on the circle may then be predicted, for instance from its past positions. It is then natural to seek a similar predictability of the actual components of the vector \mathbf{P} .

It is easily derived from expression (A-5) that the component P_L is predictable one step ahead:

$$P_L(k \Delta f) = P_L(\Delta f) P_L((k-1) \Delta f), \quad (\text{B-1})$$

$$k = k_1 + 1, \dots, k_2.$$

It may be checked that the component P_1 is also predictable:

$$P_1(k \Delta f) = \sum_{j=1}^L P_j(\Delta f) P_1((k-j) \Delta f), \quad (\text{B-2})$$

$$k = k_1 + L, \dots, k_2.$$

To proceed, consider the component P_2 . This component, defined in expression (A-5), is made of $M = L(L-1)/2$ terms of the form $z_i z_j$. Let $y_1 = z_1 z_2, \dots, y_M = z_{L-1} z_L$, denote these M phase shifts. Then P_2 has exactly the same form as P_1 in terms of these M new phase shifts, $P_2 = -(y_1 + \dots +$

$y_M)$. It follows from (B-2) that P_2 is predictable on the frequency axis, the length of the one-step ahead prediction filter being $L(L-1)/2$.

In general, each component of the vector \mathbf{P} is predictable one step ahead on the frequency axis:

$$P_j(k \Delta f) = \sum_{m=1}^M Q_{j,m}(\Delta f) P_j((k-m) \Delta f), \quad (\text{B-3a})$$

$$k = k_1 + M, \dots, k_2.$$

It may be shown that this forward prediction may be completed with the corresponding one-step backwards prediction:

$$P_j(k \Delta f) = \sum_{m=1}^M Q_{j,m}^*(\Delta f) P_j((k+m) \Delta f), \quad (\text{B-3b})$$

$$k = k_1, \dots, k_2 - M.$$

The length M of each prediction filter $Q_j(\Delta f), j = 1, \dots, L$, is given by the binomial coefficient $\binom{L}{j}$. The components of the $Q_j(\Delta f)$ filters are in general nonlinear expressions of the prediction filter $\mathbf{P}(\Delta f)$ components, except for $j = 1$ and $j = L$ [expressions (B-1) and (B-2) respectively]. For instance, if $L = 3$, $Q_{21}(\Delta f) = -P_2(\Delta f)$, $Q_{22}(\Delta f) = -P_1(\Delta f) P_3(\Delta f)$ and $Q_{22}(\Delta f) = (P_3(\Delta f))^2$.

The expressions (B-3) show that each prediction filter $Q_j(\Delta f)$ is the solution of a nonwindowed system of normal equations similar to expression (A-8), which involves the prediction filter component P_j on the bandwidth. The result is that any component of the filter \mathbf{P} may be defined by extrapolation on the full frequency band.