SEISMIC DATA PROCESSING

ÖZDOĞAN YILMAZ

INVESTIGATIONS
IN GEOPHYSICS

SOCIETY OF EXPLORATION GEOPHYSICICISTS

Series: Investigations in Geophysics,
Volume 2

Edwin B. Neitzel, Series Editor

SEISMIC DATA PROCESSING

ÖZDOĞAN YILMAZ

Stephen M. Doherty, Editor

Society of Exploration Geophysicists
Post Office Box 702740/ Tulsa, OK 74170-2740

ISBN 0-931830-41-9 (Series 1)
ISBN 0-931830-40-0 (Volume 2)

Library of Congress Catalog Number: 87-061718

Published 1987 by
Society of Exploration Geophysicists
P.O. Box 702740
Tulsa, OK 74170-2740

Second printing 1988, Third printing 1988,
Fourth printing 1989

Printed in the United States of America

To My Parents

CONTENTS

Chapter 1 FUNDAMENTALS

Chapter 2 DECONVOLUTION

Chapter 3 **VELOCITY ANALYSIS, STATICS CORRECTIONS, AND STACKING**

Chapter 7 **SLANT STACK AND APPLICATIONS**

Chapter 8 **SPECIAL TOPICS**

PREFACE

The seismic method plays a prominent role in the search for hydrocarbons. Seismic exploration consists of three main stages: data acquisition, processing, and interpretation. This book is intended to help the seismic analyst understand the fundamentals of the techniques used in processing seismic data. In particular, emphasis is given to the practical aspects of data analysis.

Topics in this book are treated in two phases. First, each process is described from a physical viewpoint, with less emphasis on mathematical development. In doing so, geometric means are used extensively to help the reader gain the physical insight into the different processes. Second, the geophysical parameters that affect the fidelity of the resulting output from each process are critically examined via an extensive series of synthetic and real data examples.

For the student of reflection seismology and new entrants to the seismic industry, this book tries to provide insights into the practical aspects of the application of the theory of time series and waves. For experienced seismic explorationists, this book should serve as a refresher and handy reference. However, it is not just meant for the seismic analyst. Explorationists who would like to gain a practical background in seismic data processing without any mathematical burden also should benefit from it. Nevertheless, for the more theoretically inclined, a mathematical treatise on the main subjects is provided in the appendixes.

The seismic analyst is confronted daily with the important tasks of:

1. selecting a proper sequence of processing steps appropriate for the field data under consideration,
2. selecting an appropriate set of parameters for each processing step, and
3. evaluating the resulting output from each processing step, then diagnosing any problems caused by improper parameter selection.

There is a well-established sequence for standard seismic data processing. The three principal processes—deconvolution, stacking, and migration—make up the foundation of routine processing. There also are some auxiliary processes that help improve the effectiveness of the principle processes. Questions often arise as to the kind of auxiliary processes that should be used and when they should be applied. For example, if shot records contain an abundance of source-generated coherent noise, then dip filtering may be valuable before deconvolution. Beam steering may be necessary to improve the signal-to-noise ratio while reducing the number of channels in processing by a factor of as much as 4. Residual statics corrections often are required for improving velocity estimation and stacking.

In a daily production environment, many questions arise concerning the optimal parameter selection for each process. Some of the most repeatedly asked questions are: What is a good length for the deconvolution operator? What should the prediction lag be? What should the design gate for the operator be? How should the correlation window be chosen in residual statics computations? What kind of aperture width should one select in Kirchhoff migration? What is the optimum depth step size in finite-difference migration? Many more questions could be included in this list of questions. To help answer these questions, a large number of examples using both field and synthetic data and describing a wide range of processing parameters are provided.

Since the old adage "a picture is worth a thousand words" is especially apt in a discussion of seismic data processing, figures make up the major portion of this textbook. In preparing some of the figures, Darran Lucas, Mike Cox, Gregg Godkin, Tania Bachus, Tomaso Gabrieli, Dave Nichols and Dave Hill kindly provided their assistance. Thanks are due to the oil companies and contractors for supplying data and some figures for which specific acknowledgement is made in the figure captions. I express my deep appreciation to Soraya Brombacher and her colleagues for the art work on most of the figures. I also extend my appreciation to Meg LaVergne, who put the final touch on many figures and computer-drafted the flow diagrams. Thanks also to the members of the playback group at Western's London Digital Center. Sally Humphreys, Jan Mitchell, and Vivian Millson helped key the text into the word processor.

I acknowledge the review work done on some chapters by Jon Claerbout, Sven Treitel, John Sherwood and Fred Hilterman. I also thank many colleagues at Western Geophysical Company who participated in reviewing parts of the earlier drafts. Special thanks to Diane Parker, who did an outstanding job of editing for style, proofreading, and preparing the camera-ready copy of the book. Thanks to Lynn Griffin for helping to bring the text up to the SEG standards of publication. Special thanks also are due to my editor, Steve Doherty, for his excellent and comprehensive review. I also appreciate his valuable recommendations in revising the text and the help he provided in bringing clarity to the text. I extend my deepest and wholehearted appreciation to my wife, Hulya, for her everlasting encouragement. I acknowledge the support given to me by Carl Savit, Damir Skerl, Aftab Alam and Ken Larner. And finally, thanks are due to Western Geophysical Company of America for the support provided to me in writing this textbook.

Özdoğan Yilmaz

INTRODUCTION

Seismic data processing strategies and results are strongly affected by the field acquisition parameters. The common-midpoint (CMP) recording is the most widely used seismic data acquisition technique. By providing redundancy (measured as the fold of coverage) in the seismic experiment, it improves signal quality. Figure 1 shows seismic data collected along the same traverse in 1965 with single-fold coverage, in 1967 with four-fold coverage, and in 1981 with twelve-fold coverage. Although these different vintages of data have been subjected to different treatments in processing, the fold of coverage has caused the most difference in the signal level of the final sections.

Surface conditions have a lot to do with the quality of data collected in the field. Figure 2 shows a stacked section. Part of the line between midpoints A and B is over an area covered with karstic limestone. Note the continuous reflections between 2 and 3 s outside the limestone-covered zone. These reflections abruptly disappear under the problem zone in the middle. The lack of events is not the result of a subsurface void of reflectors. Rather, it is caused by a low signal-to-noise (S/N) ratio resulting from energy scattering and absorption in the highly porous surface limestone.

Besides surface conditions, environmental and demographic restrictions can have a significant impact on field data quality. Figure 3 shows a stacked section. The part of the line between midpoints A and B is through a village. In the village, the vibroseis source was not operated with full power. Hence, not enough energy penetrated into the earth. Although surface conditions were similar along the entire line, the risk of property damage resulted in poor signal quality in the middle portion of the line.

Other factors, such as weather conditions, care taken during recording, and condition of the recording equipment, also influence data quality. So we see that seismic data are collected often in less than ideal conditions. Hence, we can only hope to suppress the noise and enhance the signal in processing to the extent allowed by the quality of the data acquisition.

In addition to field acquisition parameters, seismic data processing results also depend on the tools used for processing. We must ask the question: Are the tools used today in seismic data processing adequate? To gain insight into this question, consider the three principal processes: deconvolution, CMP stacking, and migration. Deconvolution quite often improves temporal resolution by collapsing the seismic wavelet to approximately a spike and suppressing reverberations on some field data (Figure 4). The problem with deconvolution is that the accuracy of its output may not always be self-evident unless it can be compared with well data. The main reason for this is that our model for deconvolution is nondeterministic in character.

Common-midpoint stacking is the most robust of the three principal processes. By using redundancy in CMP recording, stacking can significantly suppress uncorrelated noise, thereby increasing the S/N ratio (Figure 1). It also can attenuate a large part of the coherent noise in the data, such as guided waves and multiples. The normal moveout (NMO) correction before stacking is done using the primary velocity function. Because multiples have larger moveout than primaries, they are undercorrected and, hence, attenuated during stacking (Figure 5). The main problem with CMP stacking is that it is based on the hyperbolic moveout assumption. Although this assumption may be violated in areas with severe structural complexities, seismic data acquired in many parts of the world seem to satisfy this assumption reasonably well.

Finally, migration collapses diffractions and moves dipping events to their supposedly true subsurface locations. In other words, migration is an imaging process. Because it is based on the wave equation, migration also is a deterministic process. The migration output often is self-evident; that is, you can tell whether the output is properly migrated. When the output is not self-evident, this uncertainty often can be traced to the imprecision of the velocity information available for input to the migration program. Figure 6 shows an in-line stacked section from a land three-dimensional (3-D) survey and the same in line after 3-D migration using three different velocities:

(Text continued on page 7)

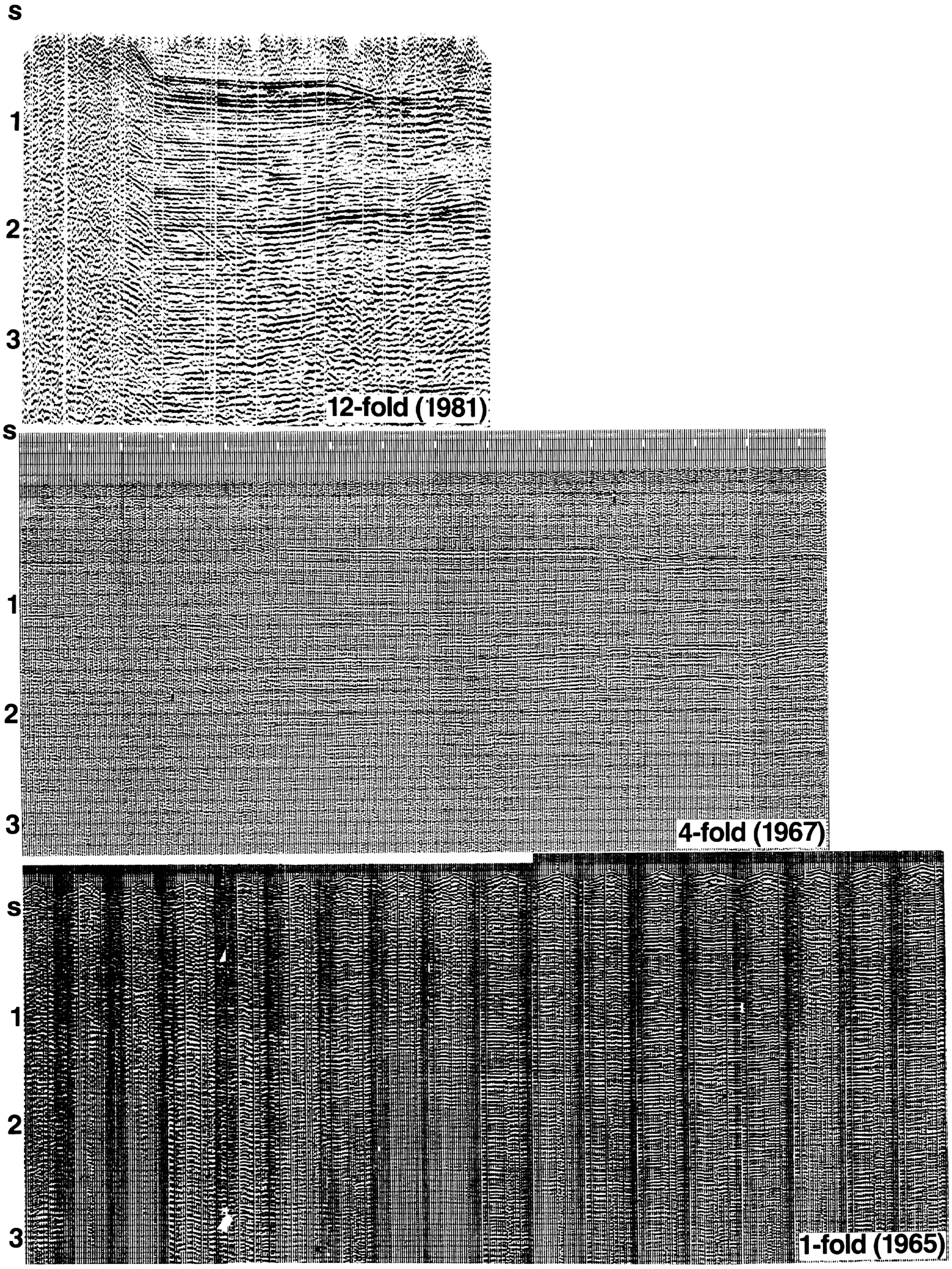


FIG. 1. A single-fold section (bottom) obtained in 1965, a four-fold stacked section obtained in 1967 (center), and a twelve-fold stacked section obtained in 1981 (top) along the same line traverse. The horizontal scales are different; the part of the single-fold section that is shown corresponds to the right half of the twelve-fold section. (Data courtesy Turkish Petroleum Corp.)

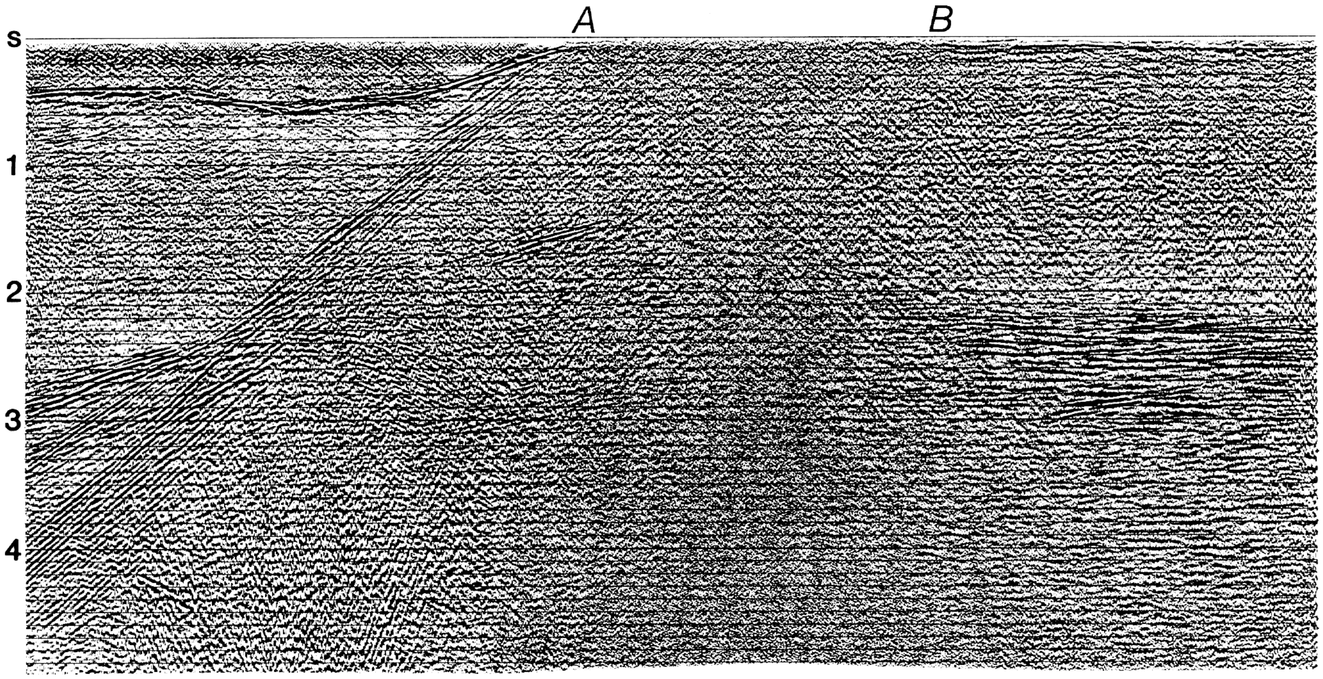


FIG. 2. The poor signal between midpoints A and B on this stacked section is caused by a karstic limestone on the surface.

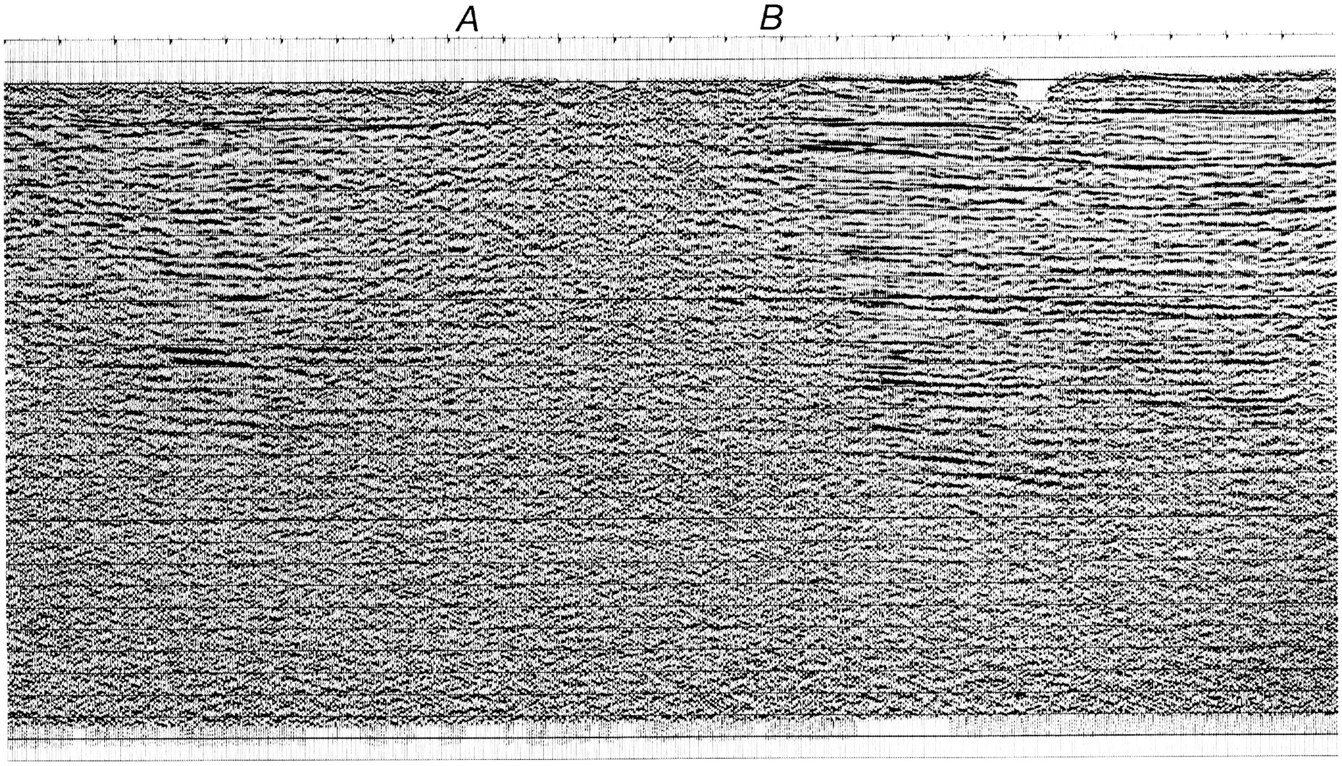


FIG. 3. A village is situated between midpoints A and B. The poor signal in that zone of the stacked section is caused by operating the vibroseis source at low power.

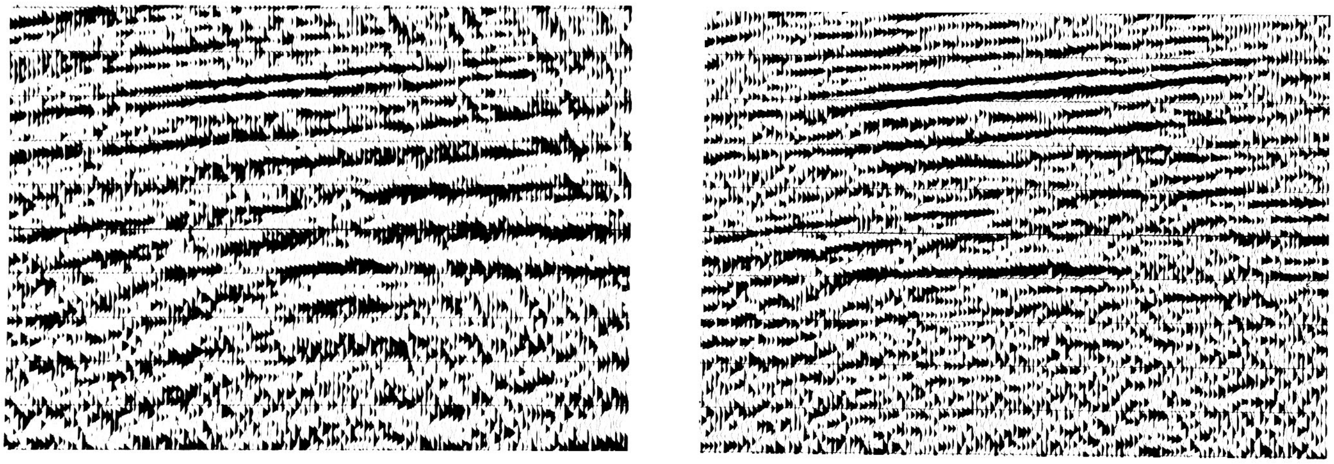


FIG. 4. A CMP stacked section without (left) and with (right) deconvolution. (Data courtesy Gulf Oil Co. Nigeria Ltd.)

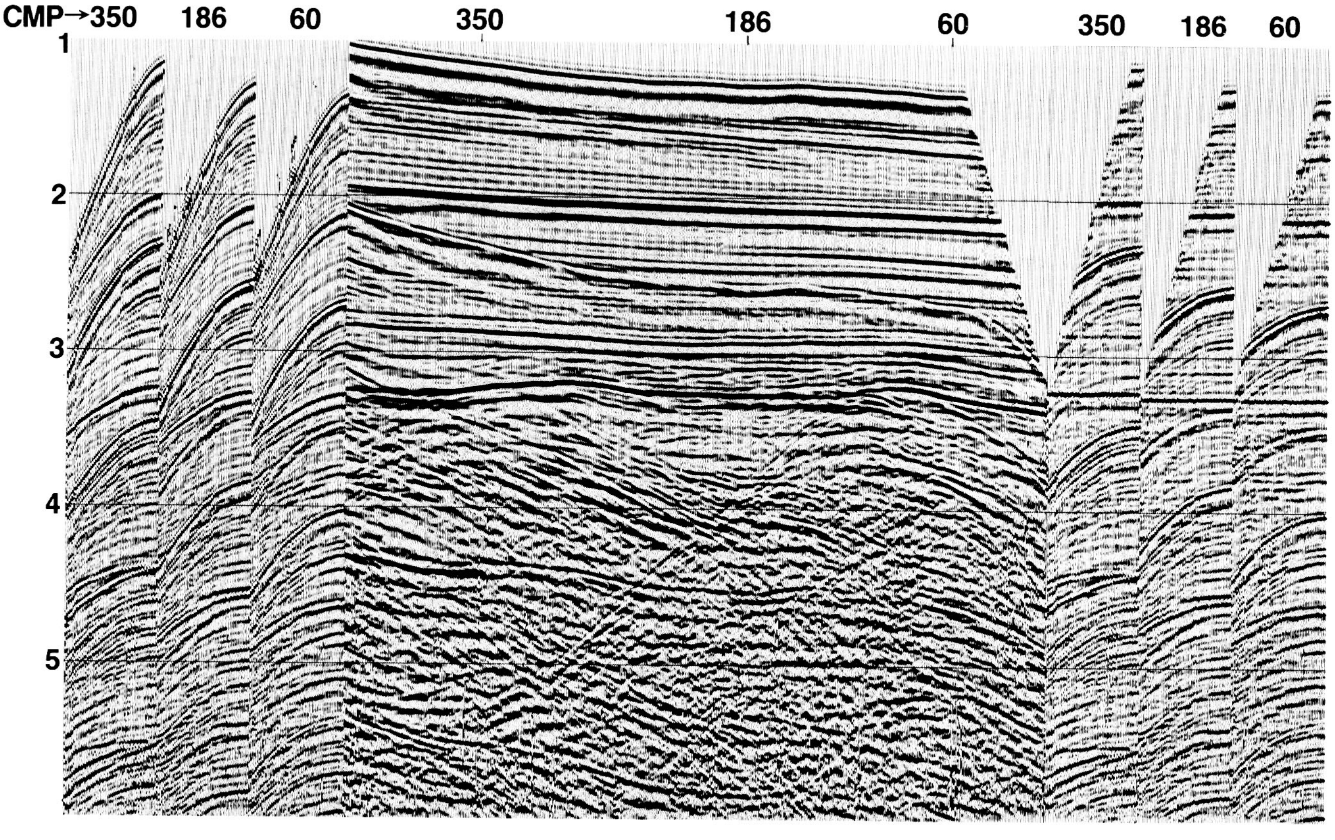


FIG. 5. Three CMP gathers before (left) and after (right) NMO correction. Note that the primaries have been flattened and the multiples undercorrected after NMO correction, hence, multiple energy has been attenuated on the stacked section (center) relative to primary energy. (Data courtesy Petro-Canada Resources.)

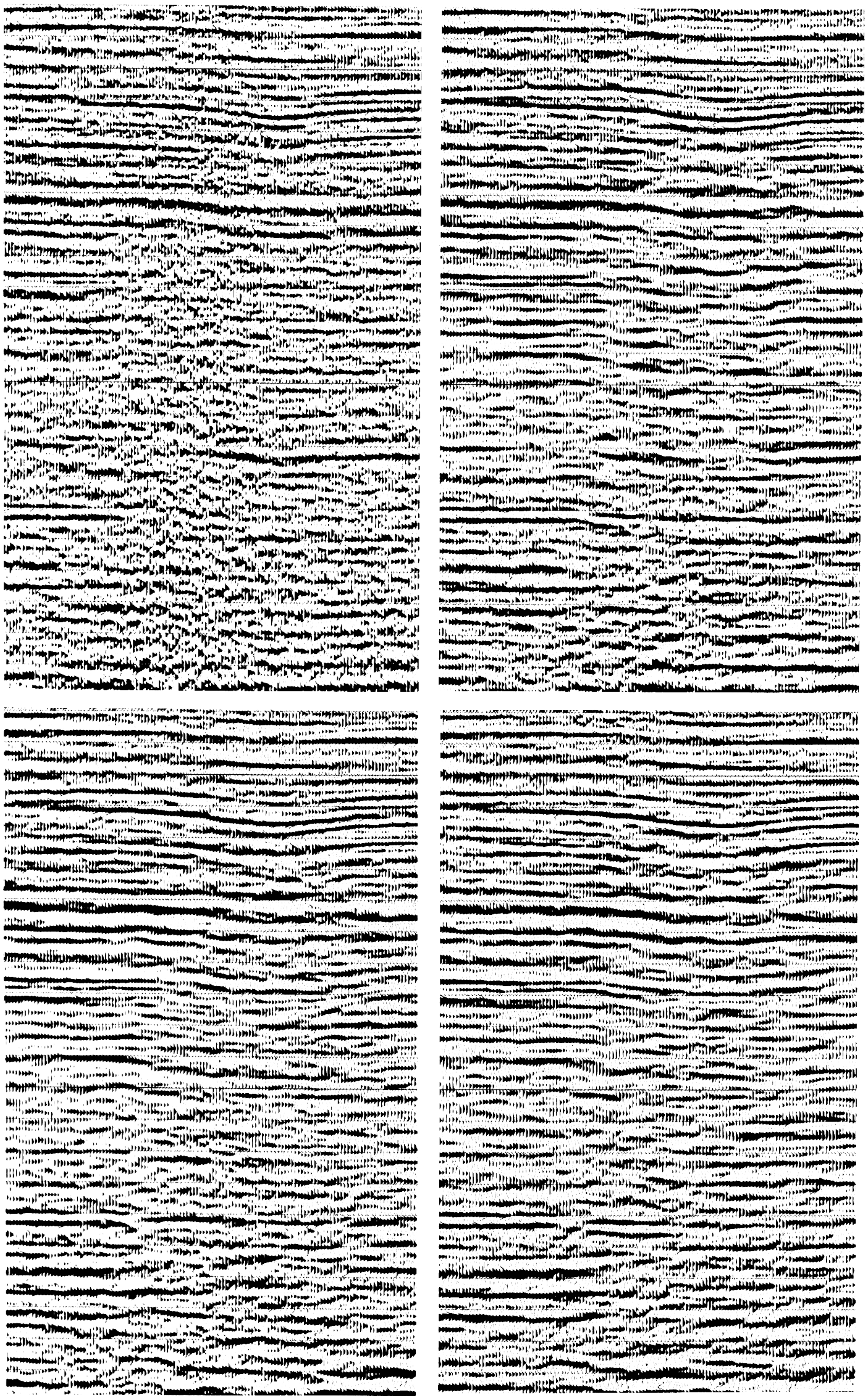


FIG. 6. An in-line stacked section (top left) and the same section after 3-D migrations using 90 percent (top right), 100 percent (bottom left), and 110 percent (bottom right) of the smoothed stacking velocities.

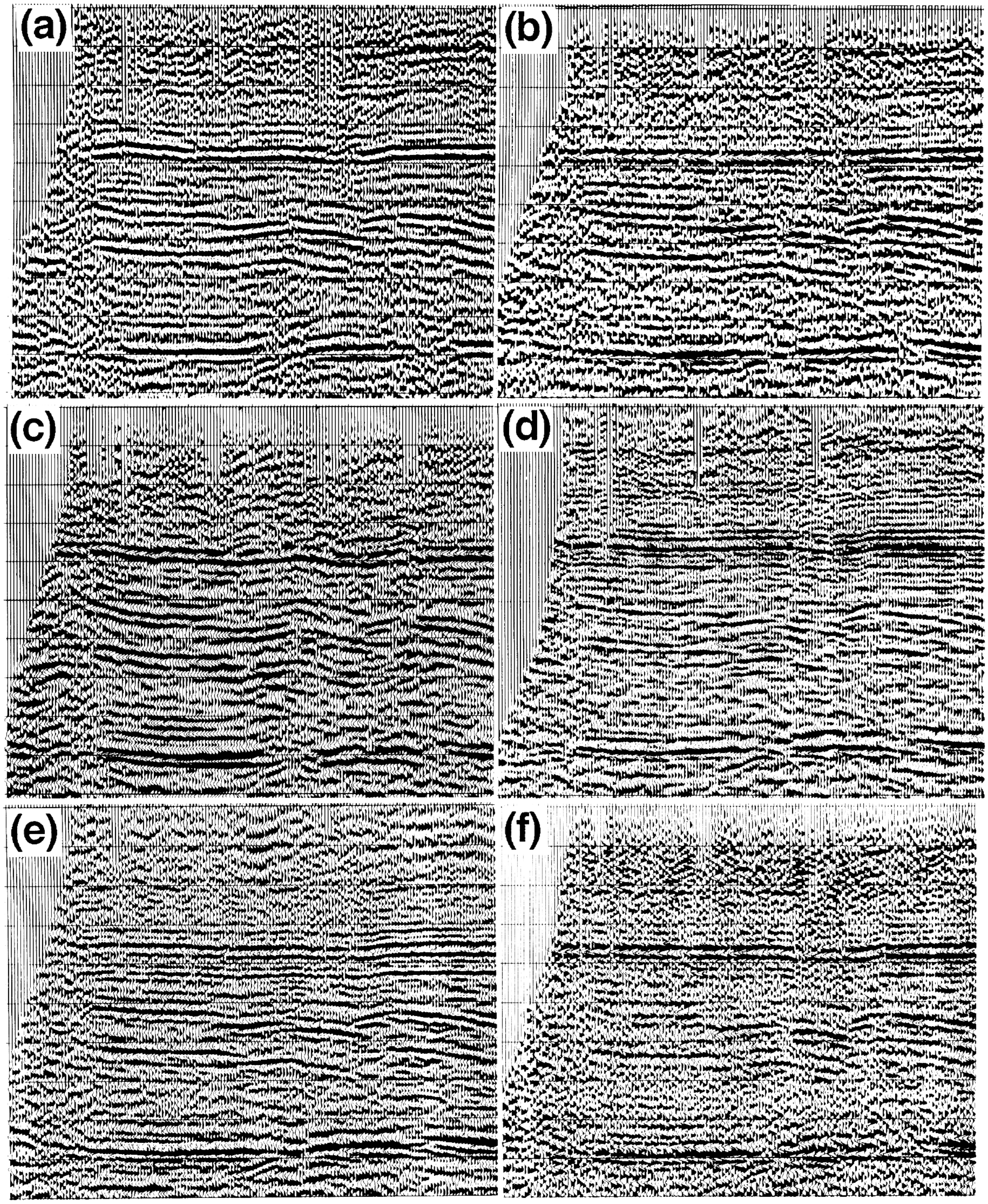


FIG. 7. A seismic line processed by six different contractors. (Data courtesy British Petroleum Development, Ltd.; Carless Exploration Ltd.; Clyde Petroleum Plc.; Goal Petroleum Plc.; Premier Consolidated Oilfields Plc.; and Tricentrol Oil Corporation Ltd.)

90, 100, and 110 percent of smoothed stacking velocities. Although dips are extremely gentle in this stacked section, the interpreter is concerned with subtle faults and small closures. There are subtle differences between the migrated sections; it is difficult to determine which part of which migrated section represents the true subsurface picture.

The above discussion on migration leads us to a weak link between seismic data and geology. *That weak link is velocity.* We can be proud of the migrated stacked section as long as the vertical axis is time. However, when the geologist asks for a section in depth, we often hedge. Again, because of the uncertainty in velocity estimation, the depth section never is entirely reliable. The weak link of velocity is a fundamental problem in seismic exploration.

There is another fundamental problem with today's seismic data processing. Even when starting with the same raw data, the result of processing done by one organization seems to be different from that done by another organization. The example shown in Figure 7 demonstrates this problem. The same data have been processed by six different contractors. Note the significant differences in frequency content, S/N ratio, and degree of structural continuity from one section to another. These differences often stem from differences in the choice of parameters and the detailed aspects of implementation of processing algorithms. For example, all the contractors applied residual statics corrections in generating the sections in Figure 7. However, the programs each contractor used to estimate residual statics most likely differed in handling the correlation window, selecting the traces used for crosscorrelation with the pilot trace, and statistically treating the correlation peaks.

Another aspect of seismic data processing is the generation of artifacts while trying to enhance signal. A good seismic program not only performs the task for which it is written but also generates minimum numerical artifacts. One of the features that makes a production program different from a research program, which is aimed at testing whether the idea works or not, is refinement of the algorithm in the production program to minimize artifacts. Processing can be hazardous if artifacts overpower the intended action of the program.

The ability of the seismic analyst often is as important as the effectiveness of the algorithms in determining the quality of the final product from data processing. There are many examples of excellent processing using mediocre software. There are also examples of the reverse. The example shown in Figure 7 implies that seismic data processing is not completely within the realm of scientific objectiveness. Data processing seems to have a flavor of interpretive character; some even consider it an art. In the following chapters, the basic tools for seismic data processing are described. It is up to the seismic analyst to make good use of these tools.

1

Fundamentals

1.1 INTRODUCTION

The Fourier transform is a fundamental ingredient of seismic data analysis. It applies to almost all stages of processing. A given time series, such as a seismic trace, can be uniquely and completely described as a sum of a number of sinusoids, each with a unique peak amplitude, frequency, and a phase-lag (relative alignment). This process is achieved by the forward Fourier transform. Conversely, the seismic trace can be synthesized given the individual frequency components. This process is achieved by the inverse Fourier transform. A brief mathematical treatise of the Fourier transform is given in Appendix A.

Seismic data processing algorithms often can be described or implemented more simply in the frequency domain than in the time domain. In Section 1.2, the one-dimensional (1-D) Fourier transform is introduced and some basic properties of time series in both time and frequency domains are discussed. Many of the processing techniques (single-or multi-channel) involve an operand (seismic trace) and an operator (filter). A simple application of Fourier analysis is in the design of zero-phase frequency filters, typically in the form of band-pass filtering.

In Section 1.3, 40 common-shot gathers that were recorded in different parts of the world with different types of sources and recording instruments are examined (Yilmaz and Cumro, 1983). Various types of seismic energy are introduced; namely, reflections, refractions, coherent noise such as multiples, guided waves, side-scattered energy and ground roll, and ambient random noise.

In Section 1.4, a summary of the basic data processing sequence is presented with field data examples. There are three primary stages in seismic data processing; each is aimed at improving seismic resolution. (Seismic resolution refers to the ability to separate two events that are very close together.)

1. *Deconvolution* is performed along the time axis to increase temporal resolution by compressing the basic seismic wavelet to approximately a spike and suppressing reverberating wave trains.
2. *Stacking* compresses the offset dimension, thus reducing seismic data volume to the plane of the zero-offset seismic section and increasing the signal-to-noise (S/N) ratio.
3. *Migration* commonly is performed on the stacked section (which is assumed to be the zero-offset section) to increase lateral resolution by collapsing diffractions and moving dipping events to their true subsurface positions.

Secondary processes are implemented at certain stages to condition the data and improve the performance of deconvolution, stacking, and migration. When coherent noise is dip filtered, for example, deconvolution and velocity analysis may be improved. Residual statics corrections also improve velocity analysis and, thereby, the quality of the stacked section.

Gain types are discussed in Section 1.5. Gain is a time-variant scaling in which the scaling function commonly is derived from the data. Gain often is applied to seismic data for display purposes. Another example of gain application is correction for wavefront divergence, which is done early in processing. (Wavefront divergence is the decrease in amplitude of a wavefront because of geometric spreading of seismic waves.) A further example of gain is automatic gain control (AGC), which brings up weak reflection zones in seismic data. However, an AGC-type gain can destroy signal character and must, therefore, be considered with caution.

The two-dimensional (2-D) Fourier transform (Section 1.6) is a way to decompose a seismic wave field, such as a common-shot gather, into its plane-wave components, each with a certain frequency propagating at a certain angle from the vertical. Therefore, the 2-D Fourier transform can describe processes like migration and frequency-wavenumber (*f-k*) filtering. A common application of *f-k* filtering is the rejection of coherent noise by dip filtering. This procedure is illustrated by field data examples.

1.2 THE 1-D FOURIER TRANSFORM

Consider the following experiment. Hold a spring at one end and attach a weight to the other end. Pull the weight down a certain amount, say 0.8 units of distance. Release the weight. Assume that the spring is elastic; i.e., it bounces up and down *ad infinitum*. Set the time to zero at the onset of motion. Displacement of the weight as a function of time should vary between the peak amplitudes (+0.8, −0.8). If you had a device that could trace the amplitude of the displacement as a function of time, it would produce a sinusoidal curve as shown in Figure 1-1 (frame 1). Measure the time interval between two consecutive peaks; you find that it is 0.080 s (80 ms). This time interval is called the period of the spring and it depends on the spring constant, a measure of spring stiffness. We say that the spring has completed one cycle of motion in a single period of time. Count the number of cycles within one second. This should be 12.5 cycles, which is called the frequency associated with the spring motion. One cycle per second (cps) is one hertz (Hz). Note that 1/0.080 s = 12.5 Hz; that is, frequency is the inverse of the period.

To continue, repeat the above experiment using a spring with a higher stiffness. Give the second spring a peak displacement amplitude of 0.4 units. The motion of the spring is traced as another sinusoid in Figure 1-1 (frame 2). The period and frequency of the spring are 0.040 s and 25 Hz, respectively. To keep track of these measurements, plot the peak amplitude of each spring as a function of frequency. These are the *amplitude spectra* shown in Figure 1-1.

Working with two identical springs, release spring 1 from a peak amplitude displacement of 0.8 units and set the time to zero at the onset of the motion. When spring 1 passes through the zero amplitude position, set spring 2 in motion from the same peak amplitude displacement

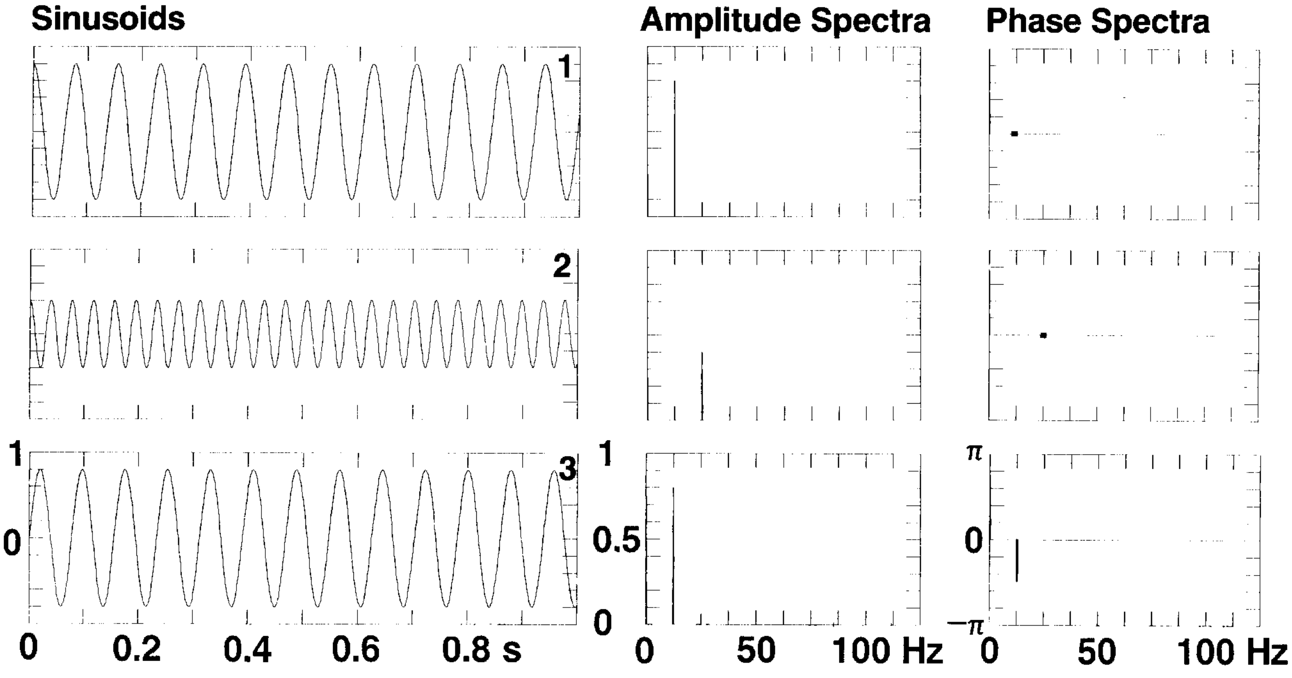


FIG. 1-1. Tracing the motion of a spring in time yields a sinusoidal curve. The peak amplitude represents the maximum displacement of the weight at the end of the spring from the unstretched position. (Positive amplitude corresponds to spring motion in the upward direction.) The time between the two consecutive peaks is the period of the sinusoid, the inverse of which is called frequency. Finally, the time delay of the onset of one spring relative to another is defined as a phase-lag. Phase spectra (the negative of phase-lag spectra) distinguish sinusoids 1 and 3.

(0.8). The motion of spring 1 is plotted in frame 1, while the motion of spring 2 is plotted in frame 3 of Figure 1-1. Because the springs were set to motion with the same peak amplitude displacement, the amplitude spectra of the two sinusoidal time functions should be identical. However, a difference is noted between the time functions in frames 1 and 3. In particular, when the sinusoid in frame 1 takes the peak amplitude value, the sinusoid in frame 3 takes the zero amplitude value. There was a time delay (20 ms) equivalent to one-quarter of a full cycle in setting spring 2 in motion relative to spring 1. This time delay is the difference between the two sinusoids shown in frames 1 and 3. A full cycle is equivalent to 360 degrees or 2π radians. Therefore, a time delay of one-quarter of a cycle is equivalent to a +90-degree *phase-lag*.

Phase is defined as the negative of phase-lag (Robinson and Treitel, 1980). Thus, a negative time shift corresponds to a positive phase value. Note that in Figure 1-1, if we apply a time shift of one-quarter of a full cycle (i.e., 20 ms) to the sinusoid in frame 3 in the negative time direction, we obtain the sinusoid in frame 1. Although their amplitude spectra are identical, these two sinusoids can be distinguished based on their *phase spectra* as seen in Figure 1-1.

The experiment is completed. What is learned? First, the motion of an elastic spring can be described by a sinusoidal time function. Second, and more important, a complete description of a sinusoidal motion is given by its frequency, peak amplitude, and phase. This experiment teaches us how to describe spring motion as a function of time and frequency.

Now imagine an ensemble of many springs, each with a sinusoidal motion with a specific frequency, peak amplitude, and phase. The sinusoidal responses of all the members are shown in Figure 1-2. Suppose the motions of the individual springs are superimposed by adding all the traces. The result is a *time-dependent* signal that is represented by the first trace in Figure 1-2 (as indicated by the asterisk). The superposition (or *synthesis*) allows us to transform the motion from frequency to time domain. This transformation is reversible; that is, the time-domain signal can be broken down (or *analyzed*) into its sinusoidal components in the frequency domain.

Mathematically, this two-way process is achieved by the Fourier transform. In practice, the standard algorithm used on digital computers is the Fast Fourier Transform (FFT). Analysis of a time-dependent signal into its frequency components is done by forward Fourier transform, while synthesis of all the frequency components to the time-domain signal is done by inverse Fourier transform.

Figure 1-3 is a display of the Fourier transform of the time-dependent signal from Figure 1-2. The amplitude and phase spectra constitute a more condensed frequency-domain representation of the sinusoids in Figure 1-2. We can clearly see the parallelism between the two types of displays. In particular, the amplitude spectrum has a large and a relatively small peak at about 20 and 40 Hz,

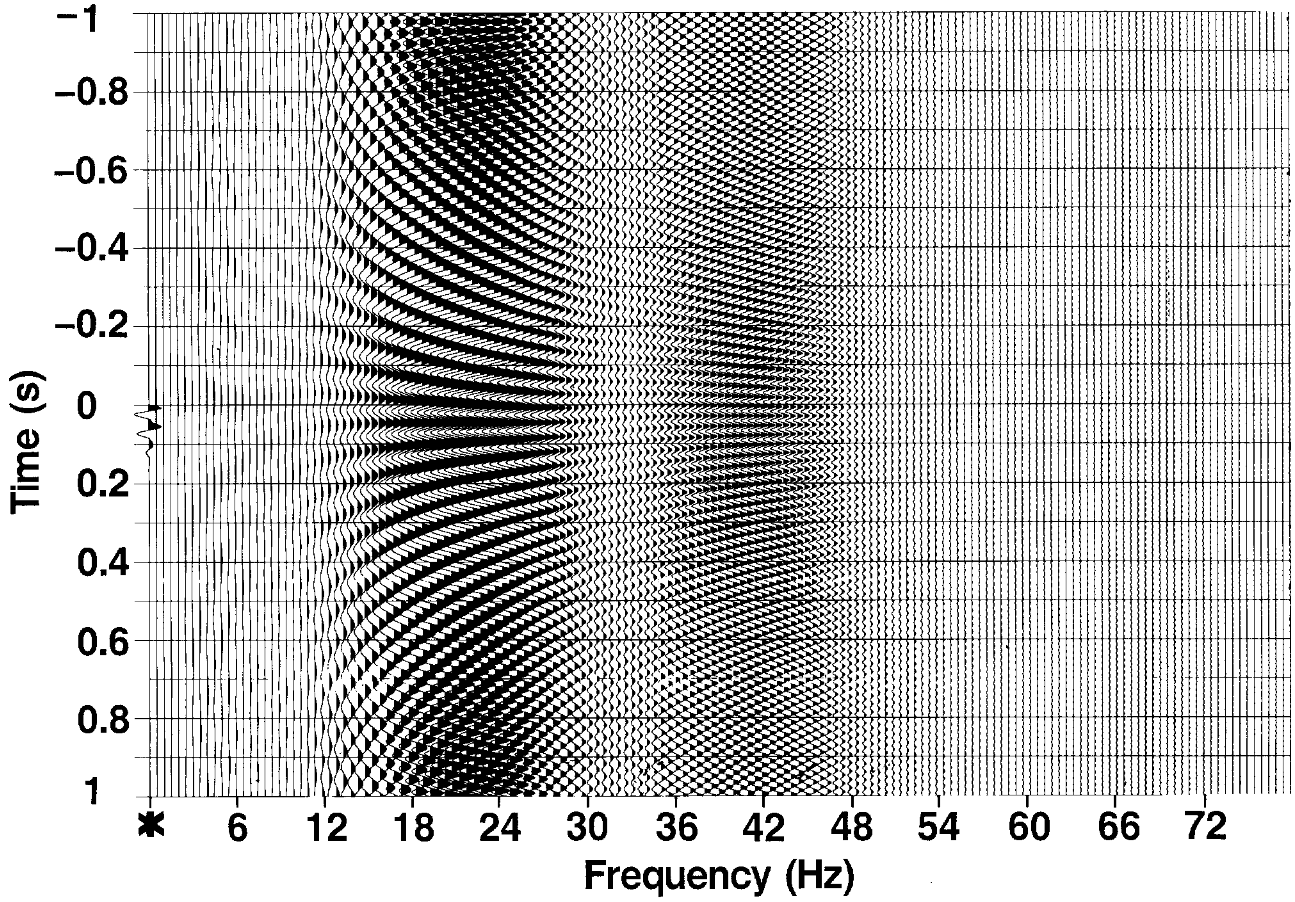


FIG. 1-2. An ensemble of sinusoidal motions with different frequency, peak amplitude, and phase-lag can be superimposed to synthesize a time-dependent waveform on the trace as indicated by the asterisk.

respectively. Darker bands corresponding to larger peak amplitudes occur in Figure 1-2 at about the same frequencies. On the other hand, zones of weak amplitudes at about 30 Hz and at the low- and high-frequency ends of the spectrum also are apparent in both types of representations. Remember that the amplitude spectrum curve represents the peak amplitudes of the individual sinusoidal components as a function of frequency.

Now examine the phase spectrum. From the spring experiment, recall that the time delay of a particular frequency component also was expressed as a phase-lag. To better trace phase-lag as a function of frequency, a part of Figure 1-2 is magnified in Figure 1-4. Follow the positive peaks that are intersected by the zero timing line. Note that the peaks fall above the zero timing line (i.e., the negative time values) at the low-frequency end of the spectrum. They then cross over to the positive side of the time axis at about 20 Hz and stay on that side over the rest of the frequency axis. The path that they followed in Figure 1-4 can be plotted as the phase spectrum of Figure 1-3. If all the peaks were aligned along the zero timing line in Figure 1-4, then the corresponding time-domain signal would have a zero-phase spectrum. In this case, all the sinusoids would reinforce each other, causing a maximum peak value at zero time (Figure 1-11).

The physical significance of the amplitude spectrum is easier to understand than that of the phase spectrum. These two spectra are discussed further in this chapter. Basic mathematical details of the Fourier transform are given in Appendix A.

1.2.1 Frequency Aliasing

A seismic signal is a continuous time function. In digital recording, the continuous (analog) seismic signal is sampled at a fixed rate in time, called the *sampling interval* (or sampling rate). Typical values of sampling intervals range between 1 and 4 ms for most reflection seismic work. High resolution studies require sampling intervals as small as 0.25 ms. Figure 1-5 shows a continuous signal in time. The discrete samples that might actually be recorded are shown by dots. A discrete time function is called a *time series*. The bottom curve in Figure 1-5 is an attempted reconstruction of the original continuous signal, which is shown as the curve on top. Note that the reconstructed signal lacks the details present in the original analog signal. These details correspond to high-frequency components that were lost by sampling. If a smaller sampling interval were chosen, then the reconstructed signal would more accurately represent the original signal. For the extreme case of a zero sampling interval, the continuous signal can be represented exactly.

Is there a measure of the restorable frequency bandwidth of the digitized data? A time series (e.g., a seismic trace) with a 2-ms sampling interval and the corresponding amplitude spectrum are shown at the top of Figure 1-6. In general, given the sampling interval Δt, the highest frequency that can be restored is $1/(2\Delta t)$. This is called the *Nyquist frequency*. For this example, $\Delta t = 2$ ms. Therefore, the Nyquist frequency is 250 Hz. The original time series was resampled to obtain a series with 4- and 8-ms sampling intervals. The corresponding Nyquist frequencies are 125 and 62.5 Hz, respectively. Figure 1-6 also shows the series (as reconstructed back to 2 ms for plotting purposes) sampled at 4 and 8 ms with their amplitude spectra. Note that the coarser the sampling interval, the smoother the series. Smoothness results from a loss of high frequencies as seen in the amplitude spectra. Frequency components between 125 and 250 Hz, which are present in the time series with the 2-ms sampling interval, seem to be absent in the series resampled to 4 ms. Likewise, frequency components between 62.5 and 250 Hz seem to be absent from the series resampled to 8 ms. Can these frequencies be recovered? No. Once a continuous signal is digitized, the highest frequency that can be restored is the Nyquist frequency.

We may think that when the time series sampled at 4 or 8 ms is interpolated back to a 2-ms sampling interval, those high frequencies should return. As stated earlier, the time series in Figure 1-6 with 4- and 8-ms sampling intervals actually were reconstructed by interpolation back to 2 ms to get the same number of samples as the original series for plotting with the same scale. Interpolation does not recover the frequencies lost by sampling; it only generates extra samples.

The implication for sampling the continuous signal in the field is an important one. If the earth signal had frequencies, say up to 150 Hz, then the 4-ms sampling interval would cause loss of the band between 125 and 150 Hz.

Consider the sinusoid in Figure 1-7. This signal is resampled as before to 4 and 8 ms. The amplitude spectra indicate that all three have the same frequency, 25 Hz. Nothing happened to the signal after resampling it to a coarser sampling interval. Now examine the higher frequency sinusoid (75 Hz) in Figure 1-8. It appears the same at both 2- and 4-ms sampling. However, resampling to 8 ms changed the signal and made it appear to be a lower-frequency sinusoid. The resampled signal has a frequency of 50 Hz as seen in the amplitude spectrum. The Nyquist frequency for an 8-ms sampling interval is 62.5 Hz. The true signal frequency is 75 Hz. As a result of resampling, the signal was lost, but reappeared as a signal with lower frequency (50 Hz). We say that it *folded back* onto the spectrum after resampling. Finally, a 150-Hz sinusoid resampled to 4 and 8 ms is shown in Figure 1-9. This time, the 4-ms sampling made the signal appear as a 100-Hz signal, while the 8-ms sampling made it appear as a 25-Hz signal.

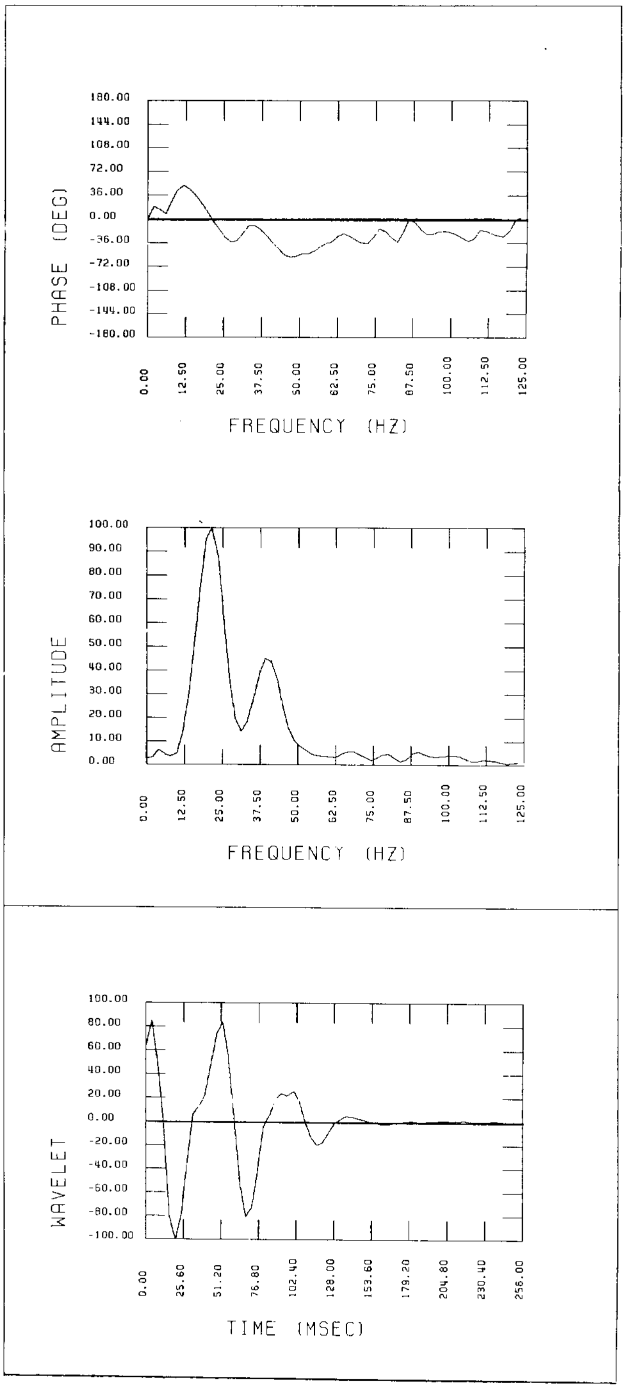


FIG. 1-3. The information from Figure 1-2 can be condensed into amplitude and phase spectra. Each point along the amplitude spectrum curve corresponds to the peak amplitude of the sinusoid at that frequency plotted as a trace in Figure 1-2. Note the equivalence of the two peaks in the amplitude spectrum with the two high-amplitude zones in Figure 1-2. Each point along the phase spectrum corresponds to the time delay of a peak or trough along the sinusoid at that frequency with respect to the timing line at t = 0. Note the equivalence of the phase curve with the trend of a positive peak from trace to trace in Figure 1-4.

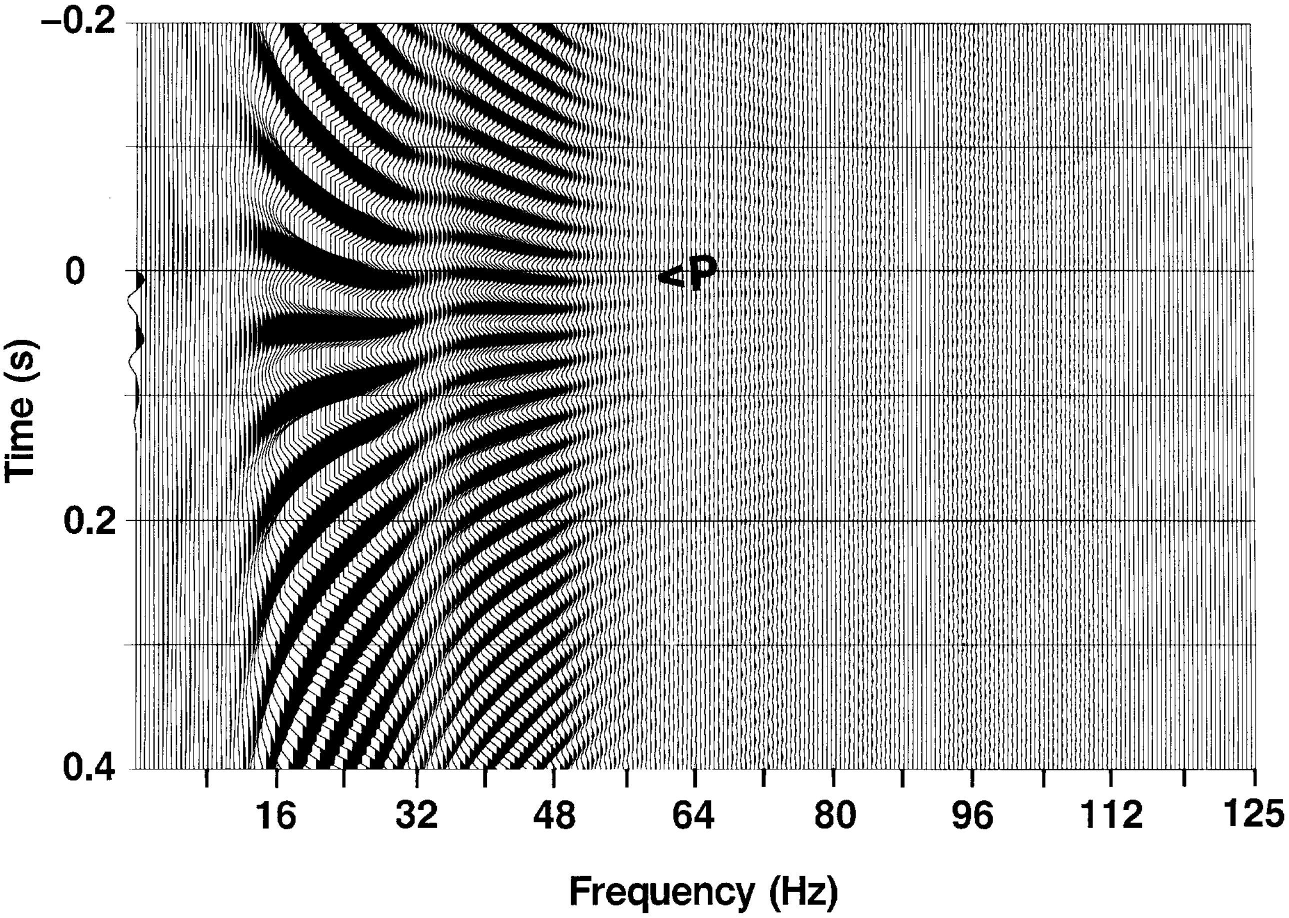


FIG. 1-4. A portion of Figure 1-2 is enlarged to better delineate the trend of the phase curve from trace to trace; i.e., from one frequency component to another. Compare the trend indicated by the positive peak P with the phase spectrum in Figure 1-3.

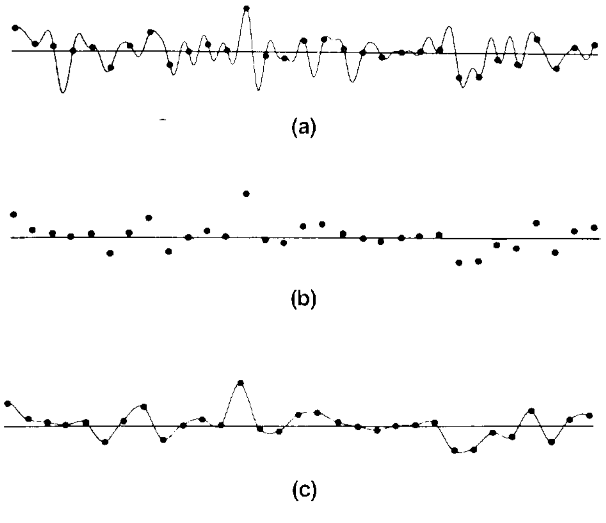


FIG. 1-5. When digitized, a continuous analog signal loses frequencies above the Nyquist frequency: (a) continuous analog signal, (b) digitized signal, (c) reconstructed analog signal. (Adapted from Rothman, 1981.)

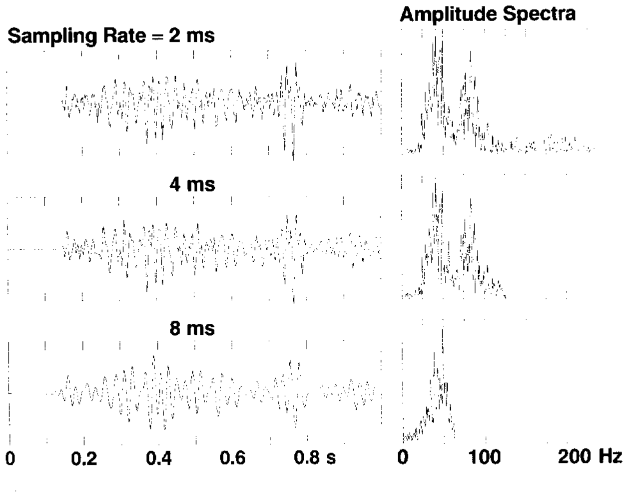


FIG. 1-6. A time series sampled at 2 ms has a Nyquist frequency of 250 Hz. Resampling to 4 and 8 ms confines the frequency band to 125 and 62.5 Hz, respectively. Note the loss of high frequencies at larger sampling intervals.

By using a single-frequency sinusoid, we see that frequencies above the Nyquist really are not lost after sampling, but reappear at frequencies below the Nyquist. Now consider the superposition of two sinusoids with 12.5- and 75-Hz frequencies as shown in Figure 1-10. Digitization of this signal at 2- and 4-ms sampling intervals does not alter the original signal, since its frequency components are below the Nyquist frequencies associated with 2- and 4-ms sampling intervals (250 and 125 Hz, respectively). However, when the signal is digitized at a coarser sampling interval, such as 8 ms, the amplitude spectrum changes. The 12.5-Hz component is not affected, because 8-ms sampling still is sufficient to sample this low-frequency component. On the other hand, the 75-Hz component is seen as a lower-frequency component (50 Hz). Once again, note that those frequencies in the original signal above the Nyquist frequency corresponding to the chosen sampling interval are folded back in the amplitude spectrum of the digitized version of the signal. This analysis can be extended to many sinusoids of different frequencies. In particular, the discrete time series derived from too coarse sampling (undersampling) of a continuous signal actually contains contributions from high-frequency components of that continuous sig-

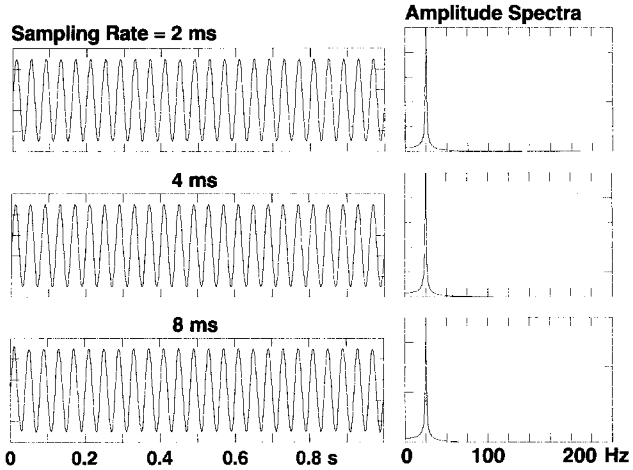


FIG. 1-7. A 25-Hz sinusoid sampled at 2 ms remains unchanged when resampled at 4 or 8 ms.

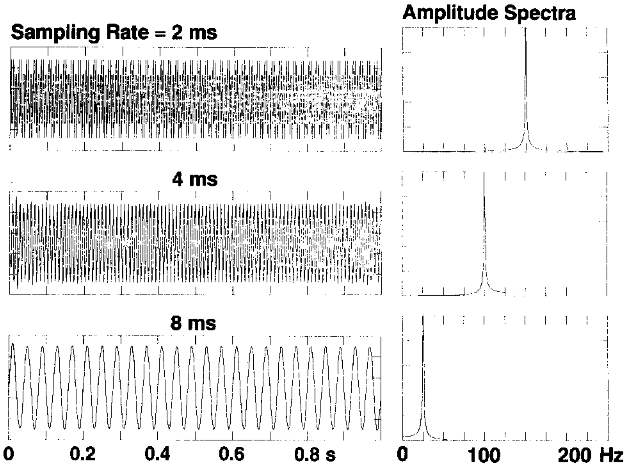


FIG. 1-9. A 150-Hz sinusoid sampled at 2 ms appears as a 100-Hz sinusoid when resampled at 4 ms and as a 25-Hz sinusoid when resampled at 8 ms. (Amplitude modulation in the 2-ms sinusoid is due to limitations in plotting very high frequency signals.)

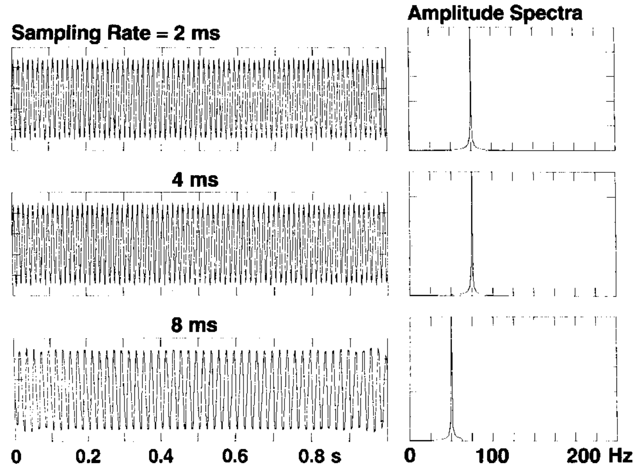


FIG. 1-8. A 75-Hz sinusoid sampled at 2 ms remains unchanged when resampled at 4 ms, but appears as a 50-Hz sinusoid when resampled at 8 ms. Hence, the latter is the alias of the original sinusoid.

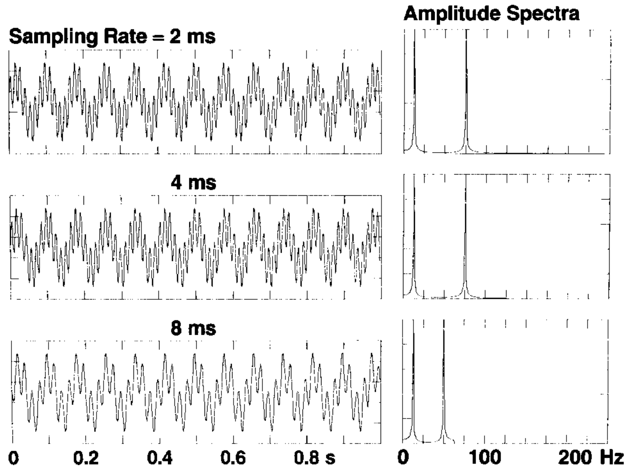


FIG. 1-10. A time series synthesized from two sinusoids at 12.5 and 75 Hz at 2-ms sampling rate remains unchanged when resampled at 4 ms. However, at 8 ms, its high-frequency component shifts from 75 to 50 Hz, while its low-frequency component remains the same.

nal. Those high frequencies fold back onto the spectrum of the discrete time series and appear as lower frequencies. The phenomenon that is caused by undersampling the continuous signal is termed *aliasing*.

To compute the alias frequency f_a, use the relation $f_a = |2mf_N - f_s|$, where f_N is the folding frequency, f_s is the signal frequency, and m is an integer such that $f_a < f_N$. For example, suppose that f_s = 65 Hz, f_N = 62.5 Hz which corresponds to 8 ms sampling rate. The alias frequency then is $f_a = |2 \times 62.5 - 65| = 60$ Hz.

In conclusion, undersampling has two effects: (a) band limiting the spectrum of the continuous signal, with the maximum frequency being the Nyquist, and (b) contamination of the digital signal spectrum by high frequencies beyond the Nyquist, which may have been present in the continuous signal. Nothing can be done about the first problem. The second problem is of practical importance. To keep the recoverable frequency band between zero and the Nyquist frequency free from aliased frequencies, a high-cut antialiasing filter is applied in the field before analog-to-digital (A/D) conversion of seismic signals. This filter eliminates those frequency components that would have been aliased during sampling. Typically, the high-cut antialiasing filter has a cutoff frequency that is either three-quarters or half of the Nyquist frequency. This filter rolls off steeply so that frequencies above the Nyquist are highly attenuated.

1.2.2 Phase Considerations

In Section 1.2, a time-dependent signal was synthesized from its frequency components. Consider a signal with a zero-phase spectrum. Figure 1-11 shows sinusoids with frequencies ranging from approximately 1 to 32 Hz. All of these sinusoids have zero-phase lag; thus, the peak amplitudes align at $t = 0$. The time-domain signal on the trace identified by an asterisk in Figure 1-11 is derived by summing all these sinusoids. This summation is an inverse Fourier transform. Such a time-domain signal is called a *wavelet*. A wavelet usually is considered a transient signal, that is, a signal with a finite duration. It has a start time and an end time, and its energy is confined between these two time positions. The wavelet that was just constructed is symmetric around $t = 0$ and has a (positive) peak amplitude at $t = 0$. Such a wavelet is called zero phase. In fact, the wavelet was synthesized using the zero-phase sinusoids of equal peak amplitude.

A zero-phase wavelet is symmetric with respect to zero time and peaks at zero time. Figure 1-12 shows the results of applying a linear phase shift to the sinusoids in Figure 1-11. Linear phase shift is defined as *phase = constant · frequency*. The wavelet (identified by an asterisk) has shifted in time by −0.2 s, but its shape has not changed. *Thus, a linear phase shift is equivalent to a*

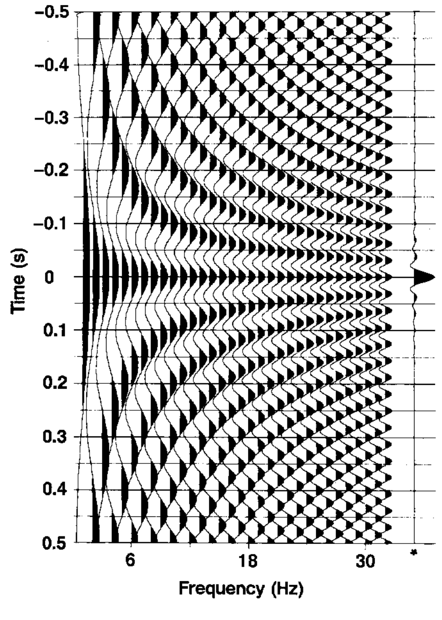


FIG. 1-11. Summation of a discrete number of sinusoids with no phase-lag, but with the same peak amplitude, yields a band-limited symmetric wavelet represented by the trace on the right (denoted by an asterisk). This is a zero-phase wavelet.

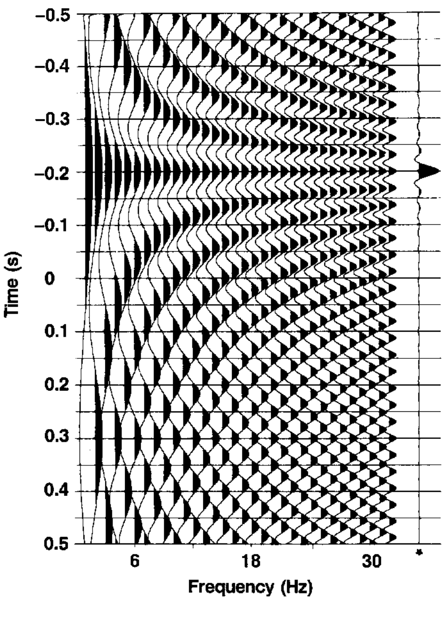


FIG. 1-12. The same sinusoidal components as in Figure 1-11, but with a −0.2 s constant time delay. When summed, they yield a band-limited symmetric wavelet that is represented by the trace on the right (denoted by an asterisk). This wavelet is the same as that shown in Figure 1-11, except that it is shifted in time by −0.2 s. This time shift is related to the linear phase spectrum associated with the summed frequency components.

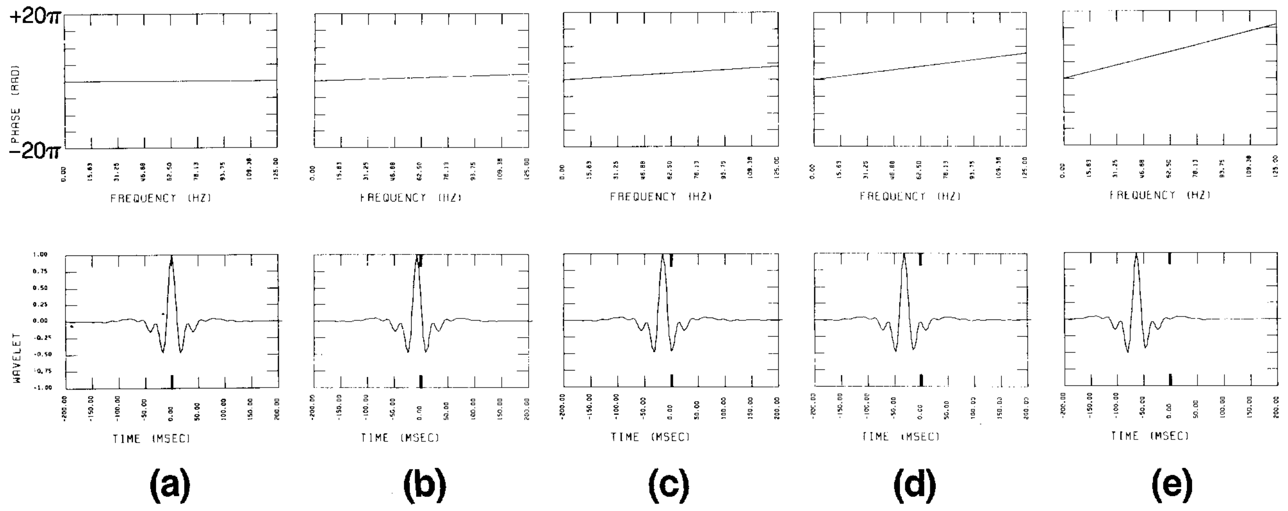


FIG. 1-13. Starting with a zero-phase wavelet (a), linear phase shifts are applied to shift the wavelet in time without changing its shape. The slope of the linear phase function is related to the time shift.

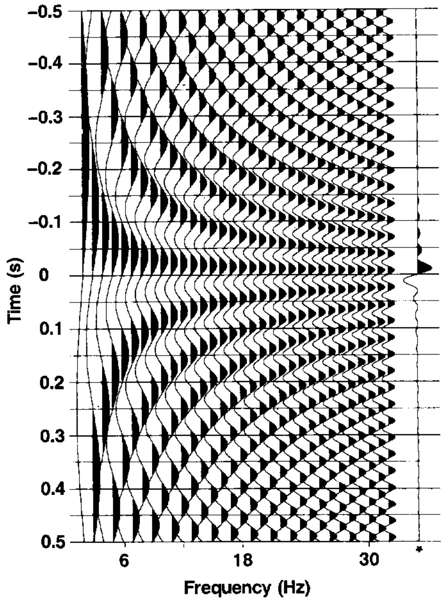


FIG. 1-14. The same sinusoidal components as in Figure 1-11 but with a constant 90-degree phase shift applied to each. The zero crossings are aligned at t = 0. Summation of these sinusoids yields an antisymmetric wavelet that is represented by the trace on the right (denoted by an asterisk).

constant time shift. The slope of the line describing the phase spectrum is proportional to the time shift.

The wavelet can be shifted by any amount of time simply by changing the slope of the line that describes the phase spectrum (Figure 1-13). Starting with the zero-phase wavelet, Figure 1-13 shows the effect of increasing amounts of linear phase shift on a zero-phase wavelet. Although not shown, by changing the sign of the slope in the phase spectrum, the wavelet can be shifted in the opposite time direction.

If a 90-degree phase shift is applied to each of the sinusoids in Figure 1-11, as shown in Figure 1-14, then the zero crossings are aligned at $t = 0$. The result of this summation yields the wavelet shown on the trace identified by an asterisk. An antisymmetric wavelet is produced. Note that the two wavelets in Figures 1-11 and 1-14 have the same amplitude spectrum because they have the same frequency content. The difference lies in their phase spectra. The wavelet in Figure 1-11 has zero-phase spectrum, while that in Figure 1-14 has a constant-phase spectrum (+90 degrees). Therefore, the difference in wavelet shape is due to the difference in their phase spectra.

Figure 1-15 shows the effect of various amounts of constant phase shift on a zero-phase wavelet. The 90-degree phase shift converts the zero-phase wavelet to an antisymmetric wavelet. The 180-degree phase shift changes the polarity of the zero-phase wavelet. The 270-degree phase shift changes the polarity of the zero-phase wavelet, while converting it to an antisymmetric wavelet. Finally, the 360-degree phase shift retains the shape of the original wavelet. *A constant phase shift changes the shape of a wavelet. In particular, a 90-degree phase shift converts a symmetric wavelet to an antisymmetric wavelet, while a 180-degree phase shift changes its polarity.*

So far, two basic phase spectra have been examined; namely, linear and constant phase shifts. The combined

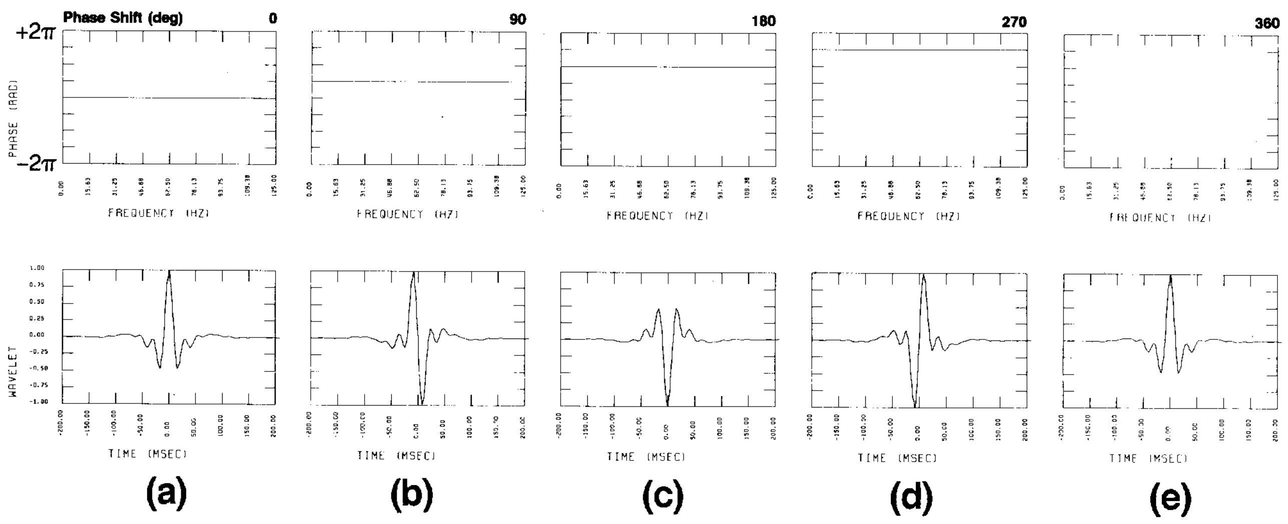


FIG. 1-15. Starting with the zero-phase wavelet (a), its shape is changed by applying constant phase shifts. A 90-degree phase shift converts the zero-phase wavelet to an antisymmetric wavelet (b), while a 180-degree phase shift reverses its polarity (c). A 270-degree phase shift reverses the polarity, while making the wavelet antisymmetric (d). Finally, a 360-degree phase shift does not modify the wavelet (e).

effect of the two now is considered. The phase spectrum is a function defined as $a + b \cdot frequency$, where a is the constant phase shift and b is the slope of the linear phase shift. Figure 1-16 shows the results of applying a 90-degree constant phase plus a linear phase component to the sinusoids in Figure 1-11. The zero-phase wavelet with the same amplitude spectrum as that in Figure 1-11 was shifted in time by −0.2 s because of the linear phase shift, and converted to an antisymmetric form because of the constant 90-degree phase shift.

Other variations in phase spectrum are shown in Figure 1-17. The zero-phase wavelet (Figure 1-17a) can be modified to different shapes simply by changing the phase spectrum. It can be modified so much that it may not resemble the original wavelet shape as illustrated by the last example (Figure 1-17d). *By keeping the amplitude spectrum unchanged, the wavelet shape can be changed by modifying the phase spectrum.*

1.2.3 Time-Domain Operations

Consider a reflectivity sequence represented by the time series (1, 0, $\frac{1}{2}$). Also consider an impulsive source that causes an explosion at $t = 0$ with an amplitude of 1. The response of the reflectivity sequence to an impulse is called the *impulse response*. This physical process can be described as follows:

Time of Onset	Reflectivity Sequence	Source	Response
0	1 0 $\frac{1}{2}$	1	1 0 $\frac{1}{2}$

One unit time later, suppose that the impulsive source generates an implosion with an amplitude of $-\frac{1}{2}$. This

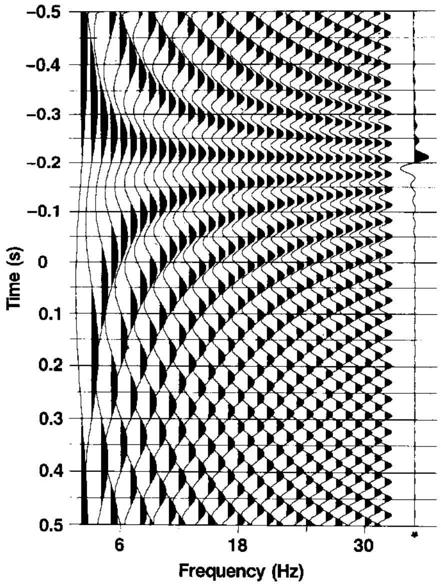


FIG. 1-16. A linear (Figure 1-12) combined with a constant phase shift (Figure 1-14) results in a time-shifted antisymmetric wavelet. The wavelet is represented by the trace on the right (denoted by an asterisk).

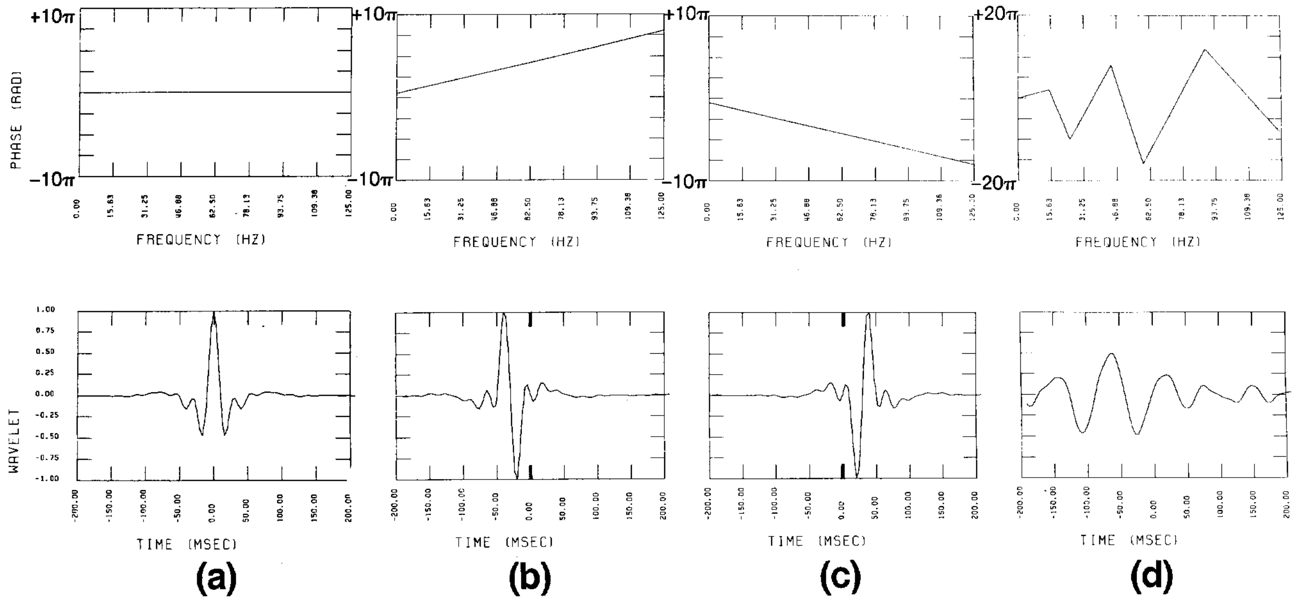


FIG. 1-17. The shape of a zero-phase wavelet (a) can be modified by introducing a nonzero-phase spectrum of any form as in (b), (c), (d).

process is described as:

Time of Onset	Reflectivity Sequence	Source	Response
1	$1 \ 0 \ \frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2} \ 0 \ -\frac{1}{4}$

Note that the response is the reflectivity sequence scaled by impulse strength. Since a general source function is considered to be a sequence of explosive and implosive impulses, the individual impulse responses are added to obtain the combined response. This process is called linear superposition and is described in Table 1-1.

In Table 1-1, the asterisk denotes *convolution*. The response of the reflectivity sequence $(1, 0, \frac{1}{2})$ to the source wavelet $(1, -\frac{1}{2})$ was obtained by convolving the two series. This is done computationally as shown in Table 1-2.

A fixed array is set up from the reflectivity sequence. The source wavelet is reversed (folded) and moved (lagged) one sample at a time. At each lag, the elements that align are multiplied and the resulting products are summed. The mechanics of convolution are described in Table 1-3. The number of elements of output array $c(k)$ is given by $m + n - 1$, where m and n are the lengths of the operand array $a(i)$ and the operator array $b(j)$, respectively.

When the roles of the arrays in Table 1-2 are interchanged, the output array in Table 1-4 results. Note that the output response is identical to that in Table 1-2. Hence, it does not matter which array is fixed and which is moved as the convolution is performed.

Seismic processing often requires measurement of the similarity or time alignment of two traces. *Correlation* is another time-domain operation that is used to make such measurements. Consider the following two wavelets:

Wavelet 1: $(2, 1, -1, 0, 0)$

Wavelet 2: $(0, 0, 2, 1, -1)$

Although these wavelets are identical in shape, wavelet 2 is shifted by two samples with respect to wavelet 1. The time lag at which they are most similar can be determined. To do this, perform the operation on wavelet 1 as described in Table 1-3 without reversing wavelet 2 (omit Step 0). This is *crosscorrelation* and the result is shown in Table 1-5. Crosscorrelation measures how much two time series resemble each other. Crosscorrelation of a time series with itself is known as *autocorrelation*.

From Table 1-5, note that maximum correlation occurs at lag -2. This suggests that if wavelet 2 were shifted two samples back in time, then these two wavelets would have maximum similarity.

Table 1-6 shows the crosscorrelation values that result when the arrays are interchanged. This time the maximum correlation occurs at lag 2. Thus, if wavelet 1 were shifted two samples forward in time, then these two wavelets would have maximum similarity. Also note that, unlike convolution, crosscorrelation is not commutative; i.e., the output depends on which array is fixed and which is moving (compare the results listed in Tables 1-5 and 1-6).

Table 1-7 shows the autocorrelation lags of wavelet 1. Note that maximum correlation occurs at zero lag, an important property of autocorrelation. Moreover, the autocorrelation function is symmetric. This is a property of real time series. Therefore, only one side of the autocorrelation needs to be computed.

It is heuristically shown in Section 1.2.4 that convolution in the time domain is equivalent to multiplication in the frequency domain (Bracewell, 1965). Since correlation is

Table 1-1. Linear superposition.

Time of Onset	Reflectivity Sequence			Source		Response			
0	1	0	$\frac{1}{2}$	1		1	0	$\frac{1}{2}$	
1	1	0	$\frac{1}{2}$		$-\frac{1}{2}$		$-\frac{1}{2}$	0	$-\frac{1}{4}$
Superposition:				1	$-\frac{1}{2}$	1	$-\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{4}$

Expressed differently:

$$(1, 0, \tfrac{1}{2}) * (1, -\tfrac{1}{2}) = (1, -\tfrac{1}{2}, \tfrac{1}{2}, -\tfrac{1}{4})$$

Table 1-2. Convolution of source wavelet (1, $-\frac{1}{2}$) with the reflectivity sequence (1, 0, $\frac{1}{2}$).

	Reflectivity Sequence				Output Response
	1	0	$\frac{1}{2}$		
$-\frac{1}{2}$	1				1
	$-\frac{1}{2}$	1			$-\frac{1}{2}$
		$-\frac{1}{2}$	1		$\frac{1}{2}$
			$-\frac{1}{2}$	1	$-\frac{1}{4}$

Table 1-3. Mechanics of the convolutional process.

Fixed Array:

a(1) *a*(2) *a*(3) *a*(4) *a*(5) *a*(6) *a*(7) *a*(8)

Moving Array:

b(1) *b*(2) *b*(3)

Given two arrays, *a*(*i*) and *b*(*j*):

Step 0 = Reverse moving array *b*(*j*).
Step 1 = Multiply in the vertical direction.
Step 2 = Add the products and write as output point *c*(*k*).
Step 3 = Shift array *b*(*j*) one sample to the right and repeat Steps 1 and 2.

Table 1-4. Convolution of reflectivity sequence (1, 0, $\frac{1}{2}$) with source wavelet (1, $-\frac{1}{2}$).

Source Wavelet						Output Response
		1	$-\frac{1}{2}$			
$\frac{1}{2}$	0	1				1
	$\frac{1}{2}$	0	1			$-\frac{1}{2}$
		$\frac{1}{2}$	0	1		$\frac{1}{2}$
			$\frac{1}{2}$	0	1	$-\frac{1}{4}$

Table 1-5. Crosscorrelation of wavelet 1 with wavelet 2.

													Output	Lag
				2	1	−1	0	0						
0	0	2	1	−1									−2	−4
	0	0	2	1	−1								1	−3
		0	0	2	1	−1							6	−2
			0	0	2	1	−1						1	−1
				0	0	2	1	−1					−2	0
					0	0	2	1	−1				0	1
						0	0	2	1	−1			0	2
							0	0	2	1	−1		0	3
								0	0	2	1	−1	0	4

Table 1-6. Crosscorrelation of wavelet 2 with wavelet 1.

													Output	Lag
				0	0	2	1	−1						
2	1	−1	0	0									0	−4
	2	1	−1	0	0								0	−3
		2	1	−1	0	0							0	−2
			2	1	−1	0	0						0	−1
				2	1	−1	0	0					−2	0
					2	1	−1	0	0				1	1
						2	1	−1	0	0			6	2
							2	1	−1	0	0		1	3
								2	1	−1	0	0	−2	4

Table 1-7. Autocorrelation of wavelet 1.

													Output	Lag
				2	1	−1	0	0						
2	1	−1	0	0									0	−4
	2	1	−1	0	0								0	−3
		2	1	−1	0	0							−2	−2
			2	1	−1	0	0						1	−1
				2	1	−1	0	0					6	0
					2	1	−1	0	0				1	1
						2	1	−1	0	0			−2	2
							2	1	−1	0	0		0	3
								2	1	−1	0	0	0	4

a convolution without reversing the moving array (see Table 1-3), a similar frequency-domain operation also applies to correlation. Figure 1-18 is a summary of frequency-domain descriptions of convolution and correlation.

From Figure 1-18, note that both convolution and correlation produce an output with a spectral bandwidth that is common to both of the input series. The immediate example is the band-pass filtering process. Also note that phases are additive in case of convolution and subtractive in case of correlation (Bracewell, 1965). For autocorrelation, this implies that the output series is zero phase. This fact already was verified by the example in Table 1-7 where it was shown that the autocorrelation is symmetric with respect to zero lag.

As a measure of similarity, crosscorrelation is used widely at various stages of data processing. For instance, traces in a common-midpoint (CMP) gather are crosscorrelated with a pilot trace to compute residual statics shifts (see Section 3.4). Again, the fundamental basis for computing velocity spectra is crosscorrelation. The building blocks of the Wiener filter (Section 2.6) are crosscorrelation of the desired output waveform with the input wavelet and autocorrelation of the input wavelet.

One other important process is the vibroseis correlation. This involves crosscorrelation of a frequency-modulated source (sweep) signal with the recorded vibroseis trace. The sweep is a frequency-modulated vibroseis source input to the ground. The convolutional model for vibroseis data is described in Section 2.7.7. Figure 1-19 shows a vibroseis sweep signal, a recorded common-source gather, and the correlated gather. The sweep length is 10 s with a frequency band of 6 to 60 Hz. The 15-s uncorrelated vibroseis record yields a 5-s correlated record. Note that the early part of the uncorrelated record contains low-frequency energy with increasingly higher frequencies at late times. This is because an *upsweep* (frequency increasing with time) signal was used in this data example.

1.2.4 Frequency Filtering

What happens to a wavelet when its amplitude spectrum is changed while its zero-phase character is preserved? To begin, consider the wavelet (summed trace 1) resulting from superposition of two very low-frequency components in Figure 1-20. Then add increasingly higher frequency components to the Fourier synthesis (summed traces 2 through 5). Note that the wavelet in the time domain is compressed as the frequency bandwidth (the range of frequencies summed) is increased. Ultimately, if all the frequencies in the inverse Fourier transformation are included, then the resulting wavelet becomes a spike, as seen in Figure 1-21 (summed trace 6). Therefore, a spike is characterized as the in-phase synthesis of all frequencies from zero to the Nyquist. For all frequencies, the amplitude spectrum of a spike is unity, while its phase spectrum is zero.

Figure 1-22 shows five zero-phase wavelets, synthesized as shown in Figure 1-20. Note that all of them have band-limited amplitude spectra. A zero-phase band-limited wavelet can be used to *filter* a seismic trace. The output trace contains only those frequencies that make up the wavelet used in filtering. The time-domain representation of the wavelet is the *filter operator*. The individual time samples of this operator are the *filter coefficients*. The process described here is zero-phase frequency filtering, since it does not modify the phase spectrum of the input trace, but merely band limits its amplitude spectrum.

Frequency-domain filtering involves multiplying the amplitude spectrum of the input seismic trace by that of the filter operator. The procedure is described in Figure 1-23. On the other hand, the filtering process in the time domain involves convolving the filter operator with the input time series. Figure 1-24 is a summary of the filter design and its time-domain application. The frequency- and time-domain formulations of the filtering process (Figures 1-23 and 1-24) are based on the following important concept in time series analysis: *Convolution in the time domain is equivalent to multiplication in the frequency domain. Similarly, convolution in the frequency domain is equivalent to multiplication in the time domain* (Bracewell, 1965).

Frequency filtering can be in the form of band-pass, band-reject, high-pass (low-cut), or low-pass (high-cut) filters. All of these filters are based on the same principle: construct a zero-phase wavelet with an amplitude spectrum that meets one of the four specifications. Band-pass filtering is used most because a seismic trace typically contains some low-frequency noise, such as ground roll, and some high-frequency ambient noise. The usable seismic reflection energy usually is confined to a bandwidth of approximately 10 to 70 Hz, with a dominant frequency around 30 Hz.

Band-pass filtering is performed at various stages in data processing. If necessary, it can be performed before deconvolution to suppress remaining ground roll energy and high-frequency ambient noise that otherwise would contaminate signal autocorrelation. Narrow band-pass filtering may be necessary before crosscorrelating traces in a CMP gather with a pilot trace for use in estimating residual statics shifts (Section 3.4). Band-pass filtering also can be performed before computing crosscorrelations during construction of the velocity spectrum for improved velocity picking (Section 3.3). Finally, it is a standard practice to apply a time-variant band-pass filter to stacked data (this section).

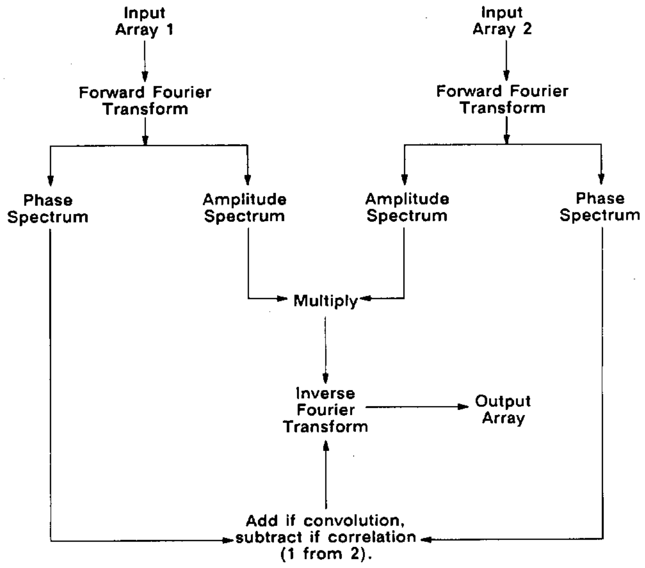


FIG. 1-18. The frequency-domain description of convolution and correlation.

Practical Aspects of Filter Design

Application of a filter in the frequency or time domain (Figures 1-23 and 1-24) yields basically identical results. In practice, the time-domain approach is favored, since convolution involving a short array, such as a filter operator, is more economical than doing Fourier transforms.

From Figure 1-22, the fundamental property of frequency filters can be stated as: *The broader the bandwidth, the more compressed the filter operator; thus, fewer filter coefficients are required.* This property also follows from the fundamental concept that the effective time span of a time series is inversely proportional to its effective spectral bandwidth (Bracewell, 1965).

In designing a band-pass filter, the goal is to pass a certain bandwidth with little or no modification, and to largely suppress the remaining part of the spectrum as much as practical. Initially, it appears that this goal can be met by defining the desired amplitude spectrum for the filter operator as follows:

$$A(f) = \begin{cases} 1, & f_1 < f < f_2 \\ 0, & \text{elsewhere,} \end{cases}$$

where f_1 and f_2 are the cutoff frequencies. This is known

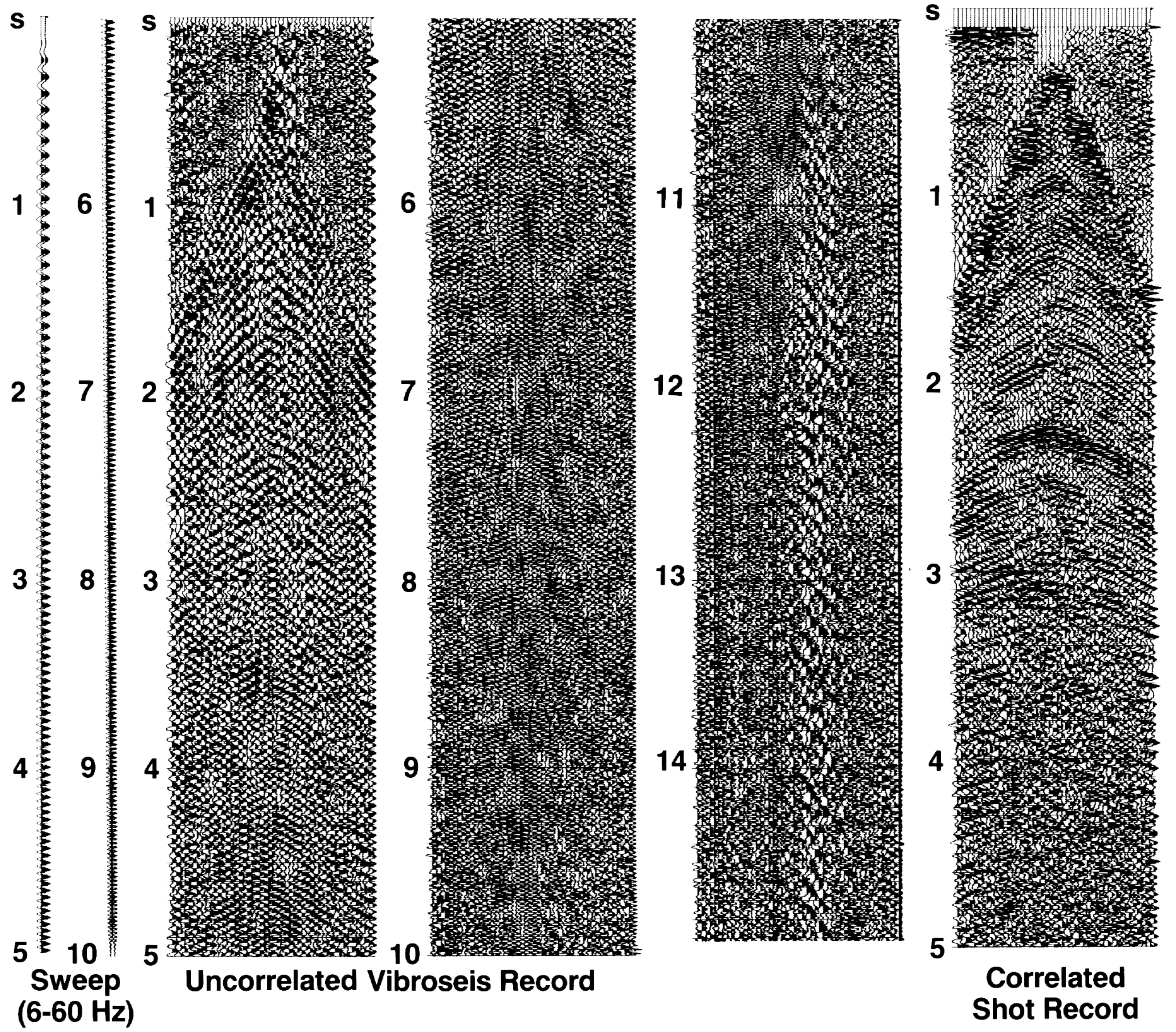


FIG. 1-19. Vibroseis correlation: the sweep signal is correlated with the recorded vibroseis record to get correlated field data. A 10-s sweep and 15-s recorded data yield a 5-s correlated record.

as the boxcar amplitude spectrum. To analyze the properties of such a filter, perform the following sequence of operations:

Define a Boxcar
Amplitude Spectrum
and Zero-Phase Spectrum
↓
Inverse FFT
↓
Filter Operator
↓
Truncate
↓
Forward FFT
↓
Compute Actual
Amplitude Spectrum

Figure 1-25a shows the results of this sequence of operations. The operator is on top and the actual and desired (boxcar) amplitude spectra are superimposed on the bottom. The actual spectrum has a ringy character. This is known as the Gibbs phenomenon (Bracewell, 1965), and results from representing a boxcar with a finite number of Fourier coefficients. From a practical standpoint, the ringing is undesirable, since some of the frequencies in the passband are amplified, while others are attenuated. Additionally, some of the frequencies in the reject zones on both sides of the boxcar are passed. What is the remedy? Instead of defining the desired passband as a boxcar, assign slopes on both sides (as shown in Figure 1-25b) and thus define the passband as a trapezoid. Note that the actual and desired amplitude spectra are now closer in agreement and the operator is

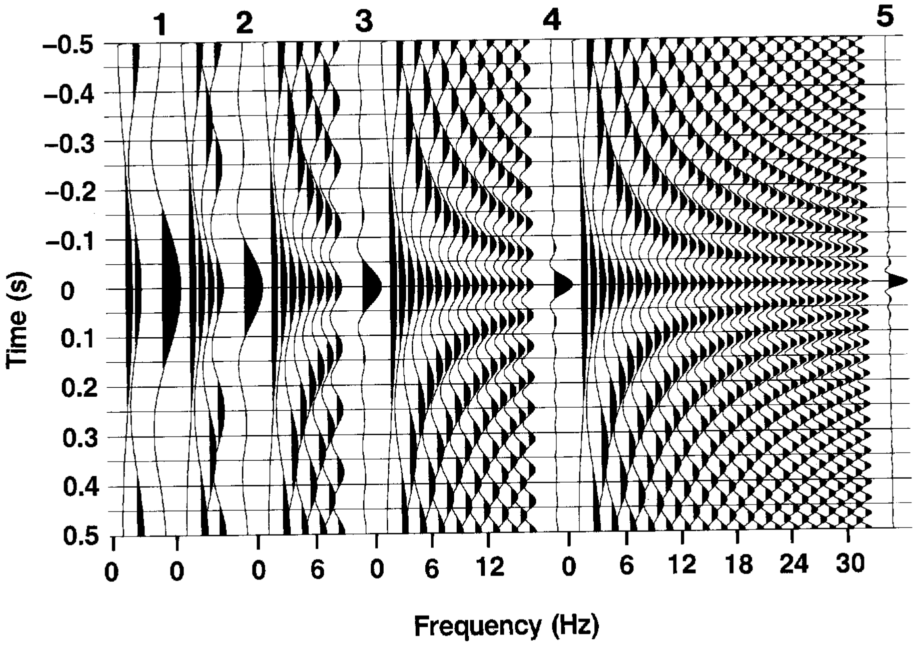


FIG. 1-20. The summation of zero-phase sinusoids with identical peak amplitudes. Trades resulting from each summation are numbered from 1 to 5. As the frequency bandwidth is increased, the synthesized zero-phase wavelet is increasingly compressed.

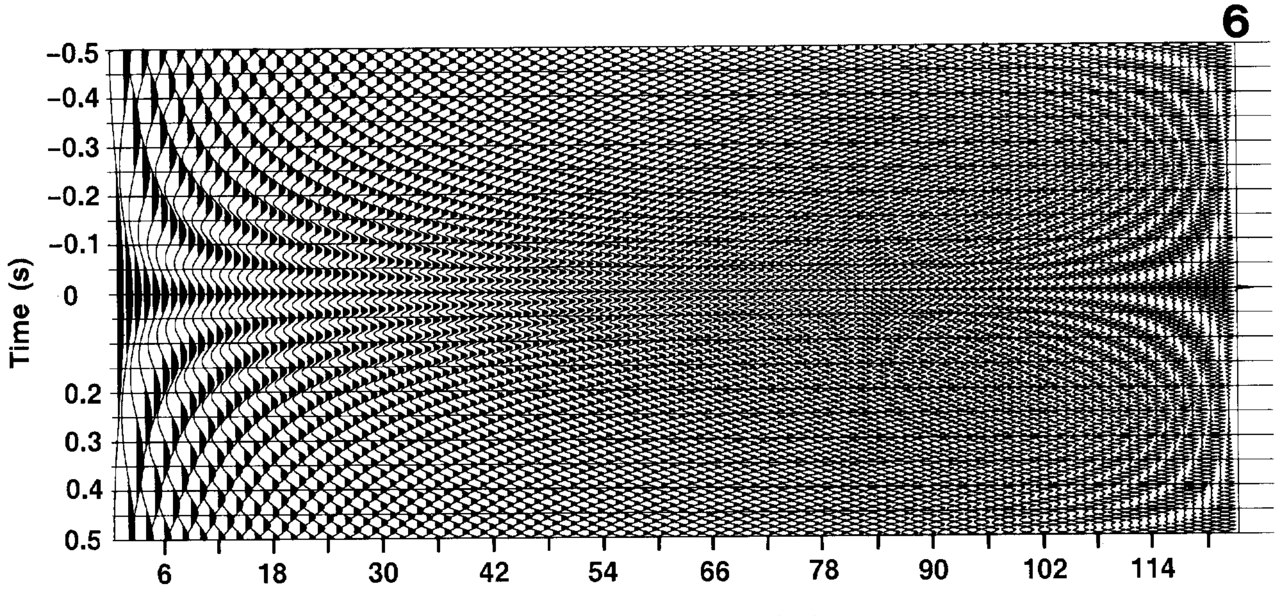


FIG. 1-21. The output wavelet becomes a spike (summed trace 6) when the summation includes sinusoids at all frequencies up to the Nyquist frequency. Compare this with output traces 1 through 5 in Figure 1-20. What causes the peculiar pattern of amplitude modulation above approximately 84 Hz?

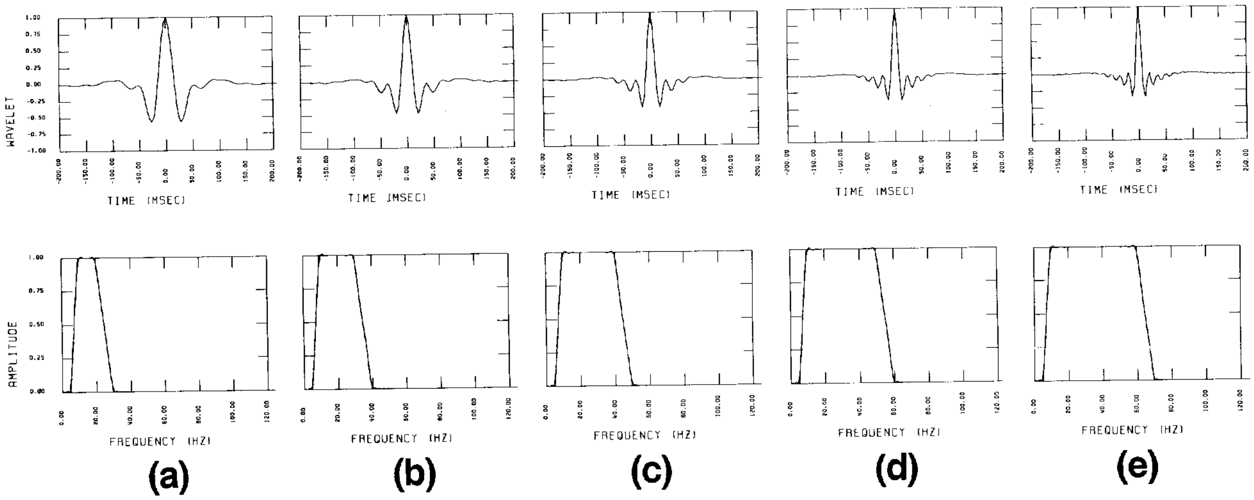


FIG. 1-22. A series of zero-phase wavelets (top row) and their respective amplitude spectra (bottom row). As bandwidth is increased, the wavelet is more compressed in time.

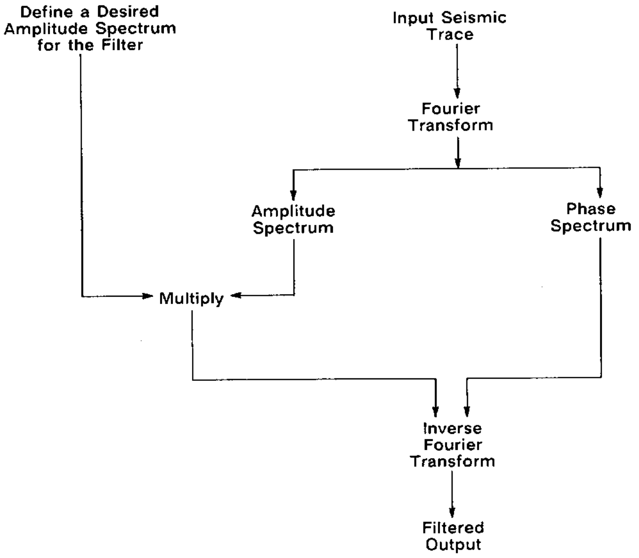


FIG. 1-23. Design and application of a zero-phase filter in the frequency domain.

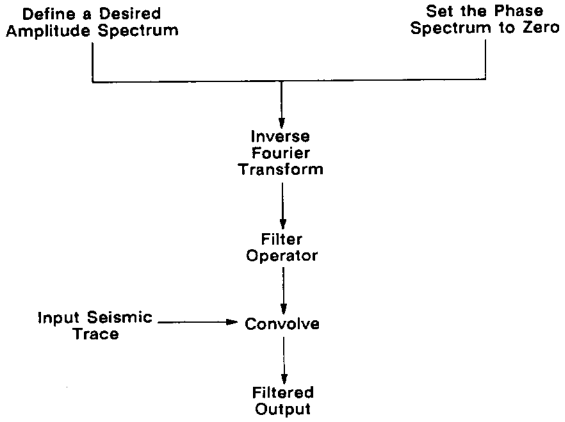


FIG. 1-24. Design of a zero-phase frequency filter and its application in the time domain.

more compact (it has fewer nonzero coefficients). However, in achieving a more compact operator, the shape of the desired spectrum has been compromised and the passband is broader than intended. The trapezoid slopes must be sufficiently gentle to achieve a satisfactory result as in Figure 1-25c, where the actual and desired spectra are approximately equal and the operator is compact. This is most desirable in practice, since it is best to work with operators that are as short as possible. It is recommended that a gentler slope be assigned on the high-frequency side relative to the low-frequency side of the passband. Finally, while defining the passband as a trapezoid, smoothing also must be applied at the corner frequencies (A, B, C, and D, as indicated in Figure 1-25c). This must be done because the Fourier transform exists only for continuous functions (Bracewell, 1965).

How short can the operator be? Figure 1-26 shows a sequence of increasingly longer operators. Excessive truncation causes a large deviation from the desired amplitude spectrum even though reasonable slopes were provided to the passband. This is noted in Figure 1-26, where solid bars indicate operator length. Extension of the operator length brings the desired and actual spectra closer. However, there is a certain length beyond which nearly zero coefficients are added to the operator. Note that the frequency bandwidth is inversely proportional to the effective length of the operator (Figure 1-22). This criterion is used to established the operator length.

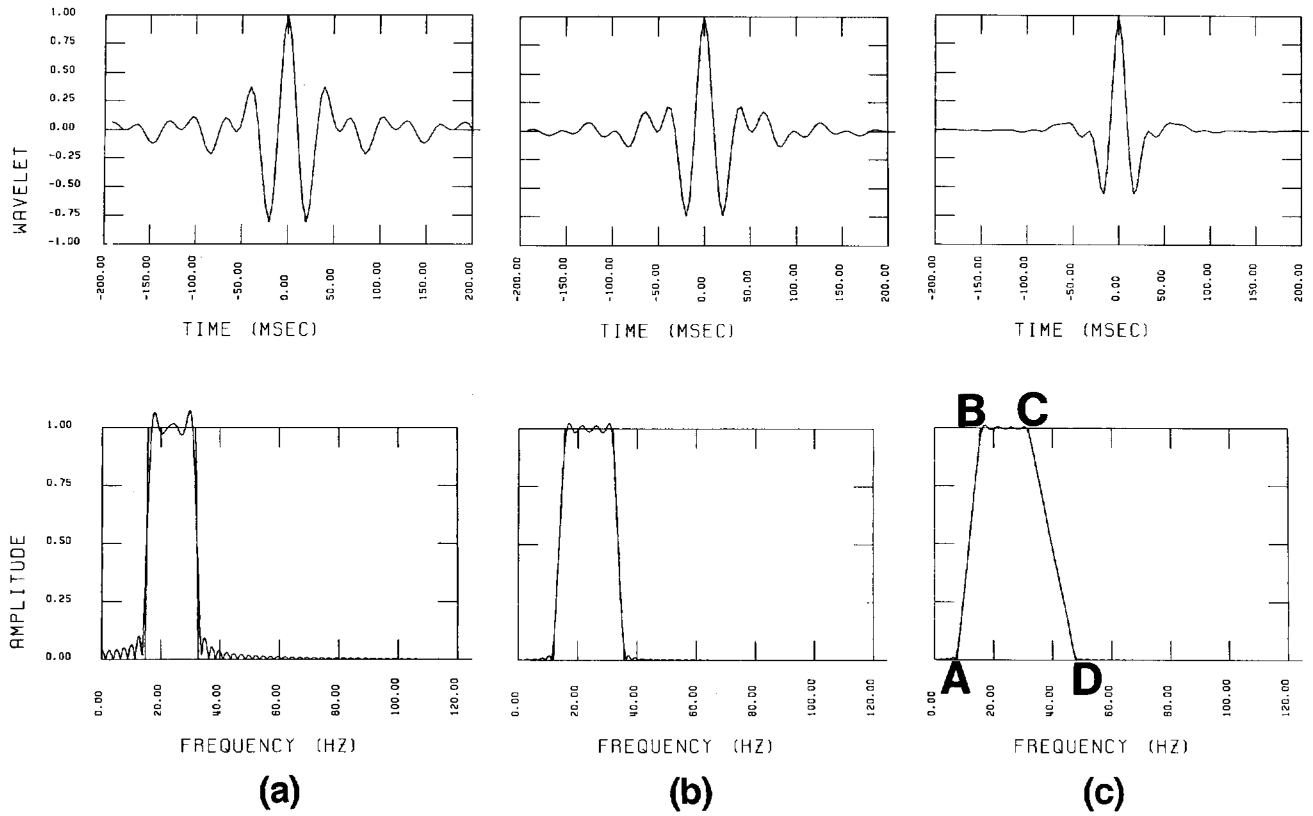


FIG. 1-25. Three zero-phase wavelets (top row) and their respective amplitude spectra (bottom row). (a) The steeply defined slopes of the passband cause ripples in the wavelet and the actual amplitude spectrum. (b) A moderate and (c) gentle slope help eliminate the ripples. Refer to the text for a discussion of points A, B, C, and D.

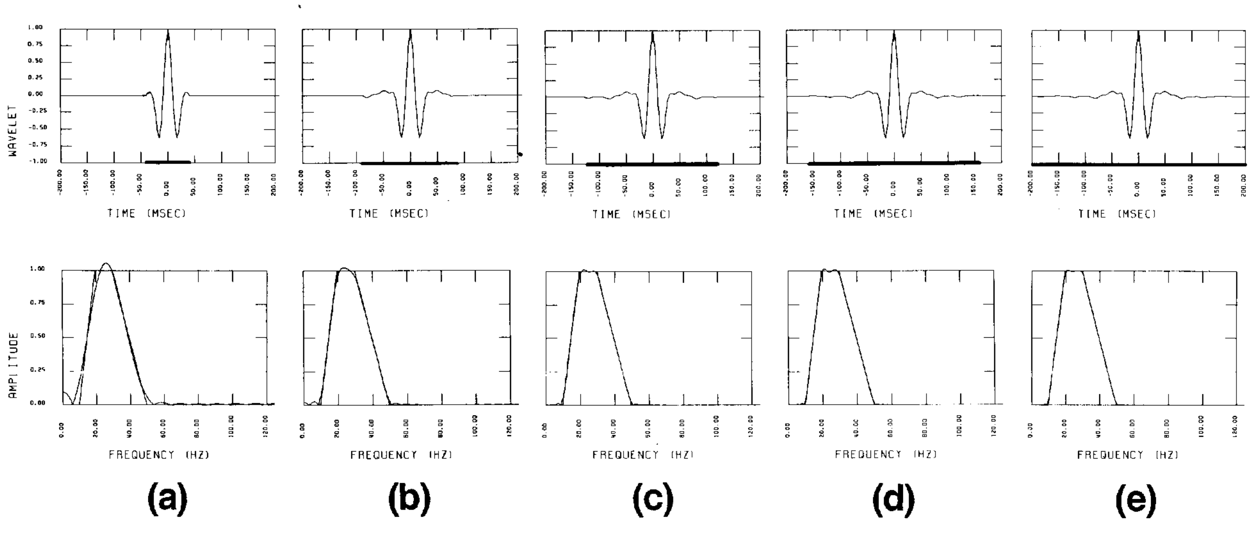


FIG. 1-26. Solid bars indicate the live length (nonzero coefficients) of the band-pass filter operator. Severe truncation (a) causes significant departure of the actual amplitude spectrum from the desired (trapezoid) amplitude spectrum, which is the same in all five cases.

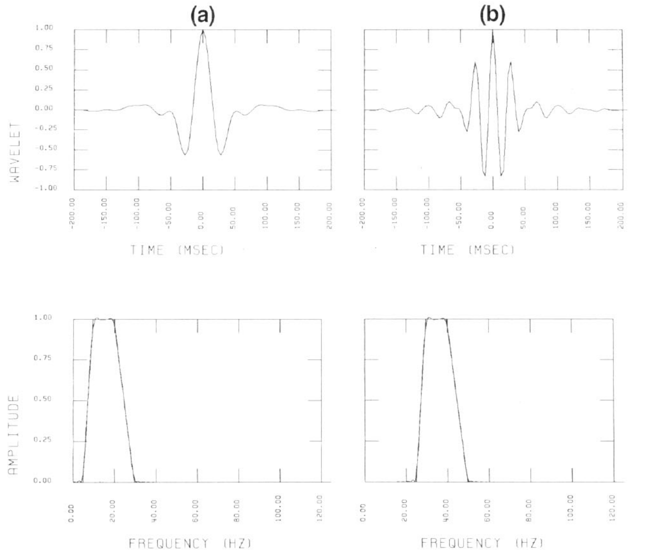


FIG. 1-27. Two wavelets (top row) with the same bandwidth (bottom row); the passband of the left wavelet is centered at 15 Hz, while that of the right wavelet is centered at 35 Hz. Both wavelets have ripples, although one is low and the other is high frequency in character. Just having low or high frequencies does not suffice; both are needed to increase temporal resolution.

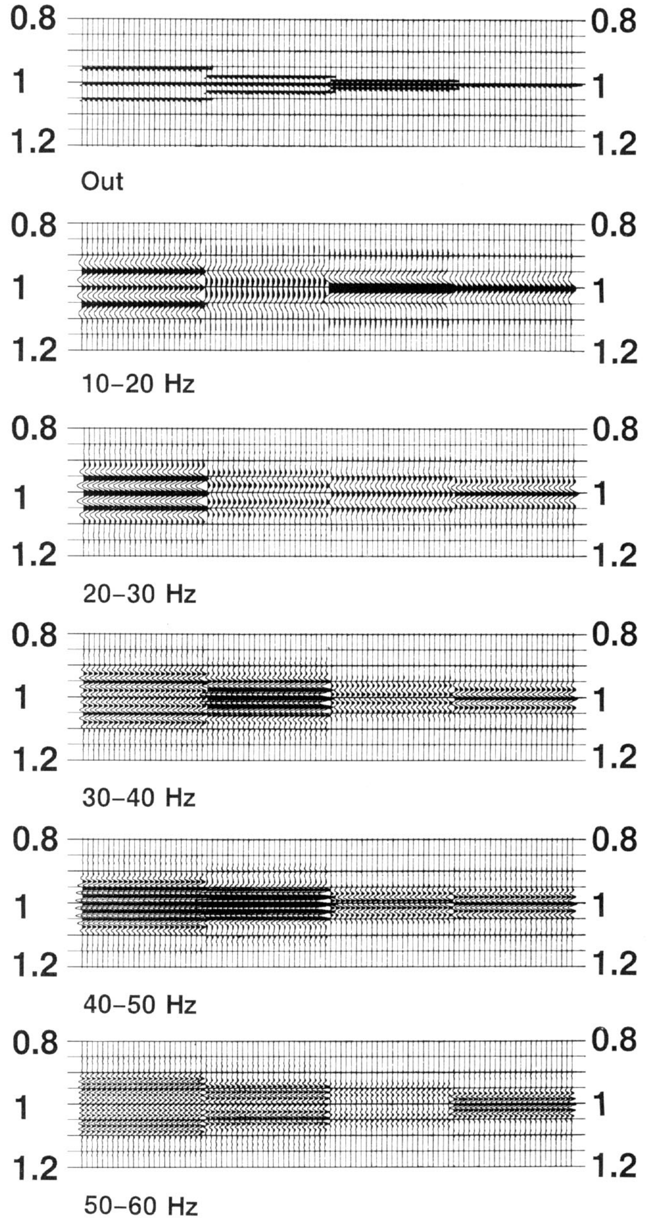


FIG. 1-28. The top section is a reflectivity model that consists of, from left to right, three reflectors with 48-ms separation, three reflectors with 24-ms separation, three reflectors with 12-ms separation, and a single reflector at 1 s. Band-limited responses (the same bandwidth, 10 Hz, centered at different frequencies) do not provide good resolution.

Bandwidth and Vertical Resolution

Frequency filtering is intimately tied to vertical (temporal) resolution of seismic data. Consider the filter operators in Figure 1-27. Both have the same effective bandwidth (the difference between the high-cut and low-cut frequencies); therefore, the envelopes of the two operators are identical. The greater ringyness of the second operator (Figure 1-27b) results from its lower bandwidth ratio (the ratio of the high-cut to the low-cut frequency).

There is a common misunderstanding that only high frequencies are needed to increase temporal resolution. This is not true. The top frame in Figure 1-28 shows a single reflector and three sets of closely situated reflectors with 48-, 24-, and 12-ms time separations. A series of narrow band-pass filters is applied to these data as shown in the lower frames. The reflectors with the 48-ms separation are resolved reasonably well by using the 10 to 20-Hz bandwidth. However, the more closely situated reflectors cannot be resolved with this filter. For the 20 to 30-Hz bandwidth, again, the 48-ms reflectors are reasonably separated. Nevertheless, none of the narrow band-pass filters provides good resolution. Just having low or high frequencies does not improve temporal resolution. Both low and high frequencies are needed to increase temporal resolution. This is demonstrated further in Figure 1-29. Note that closely situated reflectors can be resolved with increasingly broader bandwidth. The 10 to 30-Hz bandwidth is sufficient to resolve the reflectors with 48-ms separation. The 10 to 50-Hz bandwidth is sufficient to resolve the reflectors with 24-ms separation. Finally, the 10 to 100-Hz bandwidth is needed to resolve the reflectors that are separated by 12 ms. There is a close relationship between the amount of separation and the desired bandwidth. This is discussed in Section 8.3.

Time-Variant Filtering

The seismic spectrum, especially the high-frequency end, is subject to absorption along the propagation path because of the intrinsic attenuation of the earth (Section 1.5). Consider the portion of a stacked section and its narrow band-pass filtered panels in Figure 1-30. Signal is present from top to bottom within the 10 to 20-, 20 to 30-, and 30 to 40-Hz bands. Noise is noted below 3.2 s in the 40 to 50-Hz band. This noise quickly builds up to shallower times at higher frequency bands. For example, the 50 to 60-Hz band has useful signal down to 2.6 s, while the 60 to 70- and 70 to 80-Hz bands have useful signal down to only 1.8 s. Higher frequency bands of useful signal are confined to the shallow part of the section. Thus, temporal resolution is greatly reduced in the deeper portions of the section. From a practical standpoint, the time-variant character of the signal bandwidth requires an application of frequency filters in a time-varying manner. By doing this, the ambient noise, which begins to dominate the signal at late times, is excluded and a cleaner section is obtained. From Figure 1-30, the following time-variant filter (TVF) parameters are selected:

Time, ms	Filter Band, Hz
0	10-70
1800	10-70
2600	10-60
3200	10-50
4000	10-40

In practice, the filters are blended across adjacent time windows to establish a smooth transition of the passband regions. For some data, the bandwidth may be kept quite large from top to bottom. The stacked section in Figure 1-31 can tolerate wide-band filtering from early to late times. A second band-pass series of filter scans, which is shown in Figure 1-32, allows an assessment of the right choice of the bandwidth for a given time gate. Here, we start with a narrow band-pass filter at the low-frequency end of the spectrum and gradually broaden the passband by including higher frequencies.

Time-variant filters typically are applied on stacked data. A uniform bandwidth must be established when filtering two sets of data that may have different vintages, source types, or noise levels. This is especially significant when trying to tie two lines and follow a reflector across them. The interpreter uses the frequency character of a marker horizon as a reference in the tracking procedure. Therefore, two intersecting lines should be filtered so that the reflection character is consistent from one to the other, thus simplifying the interpretation.

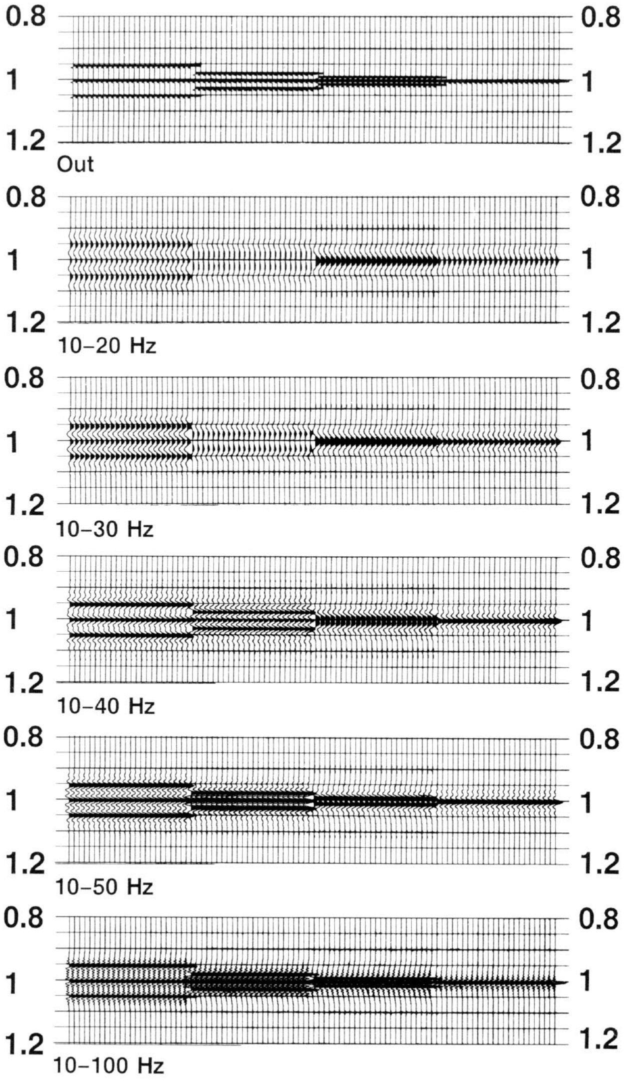


FIG. 1-29. The top section is the same reflectivity model as in Figure 1-28. Reflectors with large separation (48 ms) are resolvable with a bandwidth as low as 10 to 30 Hz; however, reflectors with smaller separation (24 and 12 ms) require increasingly larger bandwidths for resolution.

1.3 WORLDWIDE ASSORTMENT OF COMMON-SHOT GATHERS

Forty common-shot gathers, both land and marine, from North and South America, Europe, the Middle East, North Africa, and the Far East, are presented in Figure 1-33. Source types are vibroseis, Geoflex,* dynamite, air gun, Maxipulse,** Aquapulse,** and Aquaseis.* The recording parameters, including the number of traces, number of samples per trace, sampling interval, trace interval, and inner offset, are indicated in Table 1-8. Study the field records to learn how to recognize different types of waves. For display purposes, an instantaneous type of gain (AGC) (Section 1.5) was applied to all 40 records. These records will be referred to by their record numbers in the following discussions.

The main goal in processing reflection seismic data is to enhance genuine reflection signal by suppressing unwanted energy in the form of coherent and random ambient noise. In the following paragraphs, shot gathers are examined to point out the different types of seismic energy.

Record 1 is a correlated vibroseis dataset. (For vibroseis correlation, refer to Section 1.2.3.) A number of reflec-

---

*Registered trademark of Imperial Chemical Industries.

**Registered trademark of Western Geophysical Company of America.

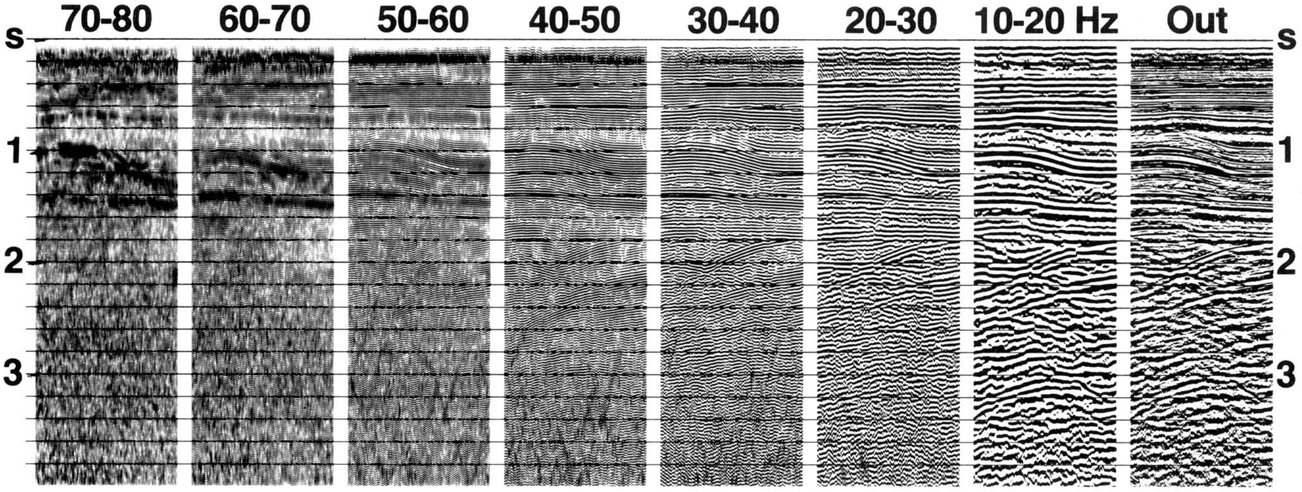


FIG. 1-30. The right-most panel is a portion of a CMP stack without filtering. The remaining panels show the same data with different narrow-band filters applied.

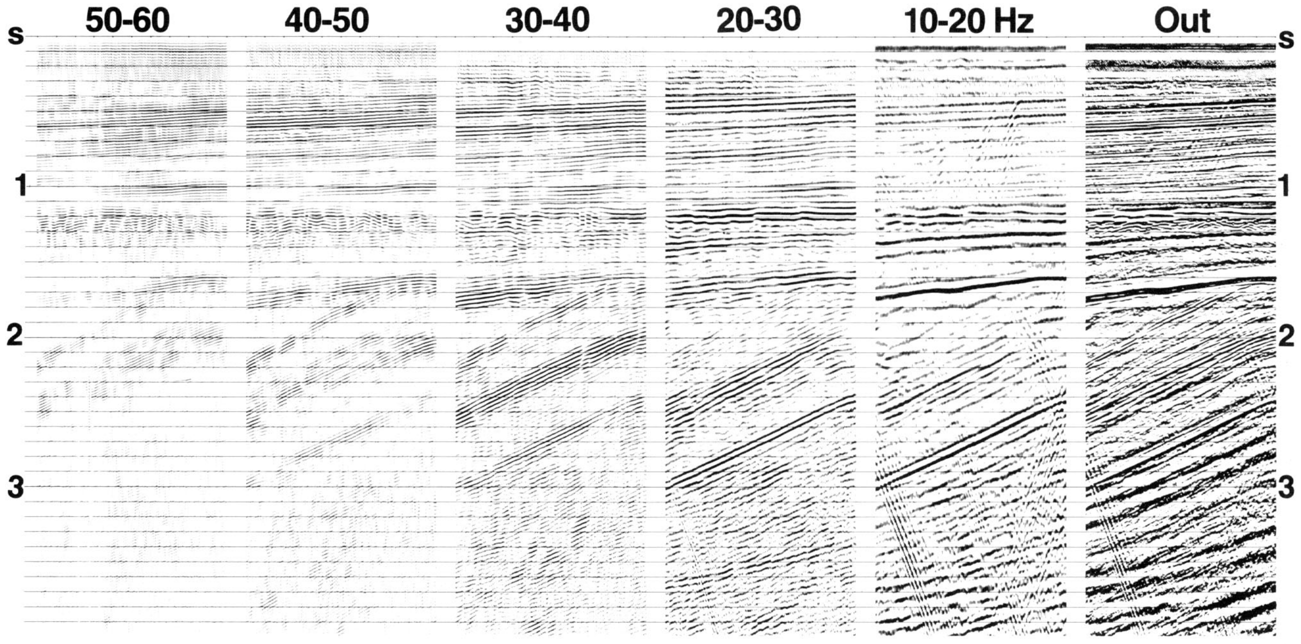


FIG. 1-31. The right-most panel is a portion of a CMP stack without filtering. The remaining panels show the same data with different narrow-band filters applied.

tions are present in this record with reasonably good S/N ratio. A genuine reflection is recognized on common-shot gathers by its hyperbolic nature. Reflections behave the same way on common-midpoint gathers. A flat horizon with no dip yields a symmetric hyperbola on both common-shot and common-midpoint gathers recorded using split-spread geometry. (In split-spread geometry, the source is located somewhere in the middle of the receiver cable, usually at the center.) A dipping horizon yields a skewed hyperbola on a common-shot gather, while still yielding a symmetric hyperbola on a common-midpoint gather. Reciprocity of sources and receivers provides this symmetry. From the reflection hyperbolas in Record 1, note that the subsurface is made up of nearly horizontally flat layers. Any irregularity in the shape of the moveout hyperbola can be attributed to near-surface effects and lateral variations in velocity.

Record 2 is an asymmetric shot gather. Note the reflection energy between 1 and 2 s, with rather irregular moveout. Record 3, which was obtained by using dynamite, contains a series of reflections with nearly perfect hyperbolic moveout, especially between 1 and 3 s. This

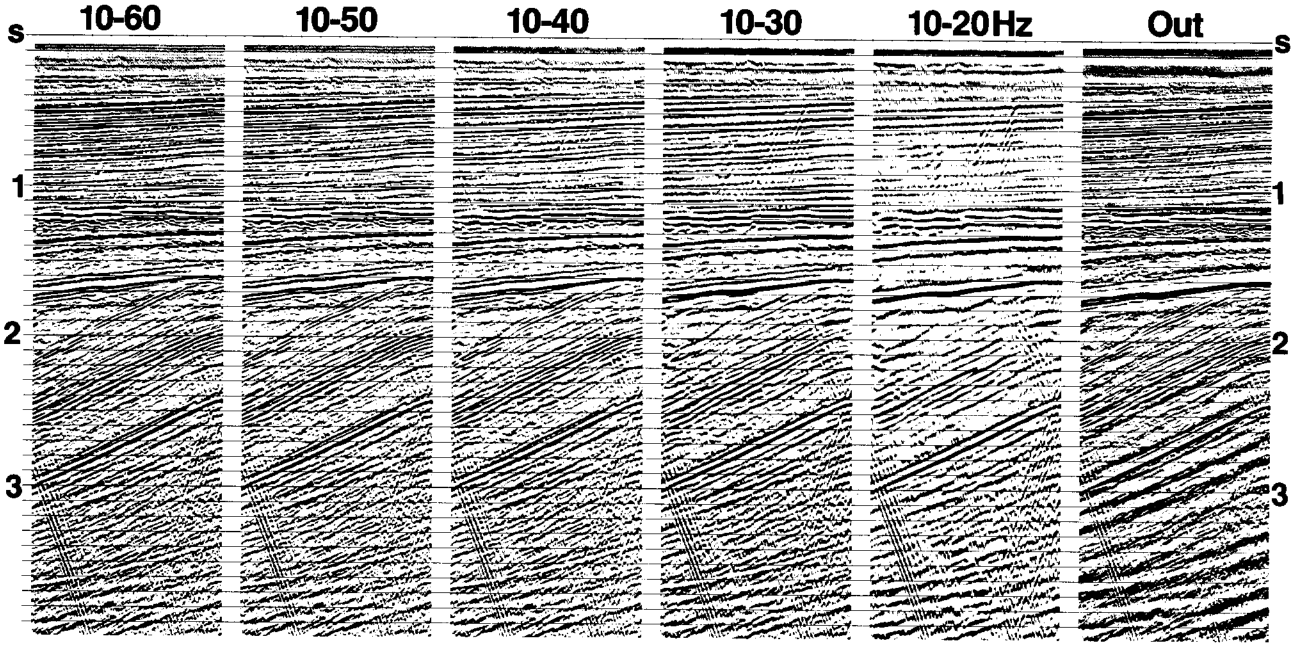


FIG. 1-32. The right-most panel is a portion of a CMP stack without filtering. The remaining panels show the same data with different band-pass filters (with increasingly wider passbands) applied.

record is from the days of analog recording. It is not uncommon to digitize old analog data and process it with modern techniques.

Record 4 contains events with complex moveout between 2.5 and 3.5 s. Events A, B, and C have skewed hyperbolic moveouts, which, in this case, suggests that they are dipping up toward the left. Also note the traveltime distortions along the moveouts due to (most likely) near-surface irregularities.

Record 5 has some ground roll energy, which is characterized by its low-frequency, high-amplitude appearance, particularly on short-offset traces. This kind of energy typically is suppressed in the field by using a proper receiver array. Record 6 contains weak and strong, nearly flat reflectors (A and B). The break in the reflection hyperbola (C_1C_2) suggests the presence of a fault (C_1 on the upthrown side and C_2 on the downthrown side). Again, note the ground roll energy with its dispersed low-frequency character on inside traces (event D).

Record 7 contains three interesting events. Event A is a skewed hyperbola, which suggests that it is dipping up toward the left, while event B is nearly symmetric, which suggests a flat dip. Finally, event C shows a discontinuity F along its moveout curve, indicating the presence of a fault.

Record 8 shows a record with excellent signal quality. This dynamite record has a number of reflections and associated interbed reverberations. Note the progressive decrease in the S/N ratio at late times. This is true for almost all seismic data. Event A has large moveout because it is shallow, while event B has small moveout because it is deep. (Linear energy C is referred to in Exercise 1.21.)

Record 9, which is a correlated vibroseis dataset, has a series of reflections and ground roll. Unlike data from impulsive sources such as dynamite, first breaks in vibroseis data may not be distinguishable (compare, for example, Records 8 and 9). This is because the correlated vibroseis record contains some of the side lobes of the sweep signal autocorrelation. Note the increase in random noise in the later part of the record below 3 s.

Record 10 contains two strong shallow reflectors, A and B, in addition to ground roll energy C. Also, a bundle of energy with extremely large moveout is noted between 2.5 and 5 s (D_1-D_2). This coherent noise may be attributed to side-scattered energy, which is caused by inhomogeneities in the subsurface (particularly at the water bottom) that behave as point sources.

Record 11 contains four prominent reflections. This record is from Alaska, where the thickness of the permafrost layer can be irregular. Such near-surface irregularities can have lateral dimensions that range from less than a group interval to wavelengths that are several times a cable length. As seen on the right flank of the hyperbolas (events A, B, C, and D), these irregularities cause substantial time shifts in reflection arrivals. Such distortions in moveout could be dynamic (time-dependent) or static (time-independent). They should be corrected before stacking. Except for these distortions, all events seem to have symmetric, hyperbolic moveouts that indicate nearly horizontally layered substrata.

Record 12 is a field record with a low S/N ratio. A complex subsurface structure is implied between 2 and 3.5 s.

From Record 13, note the high-frequency hyperbolic energy S that is associated with a side scatterer, possibly

Table 1-8. Parameter index of a worldwide assortment of common-shot gathers.

Record Number	Area	Number of Samples per Trace	Number of Traces	Sampling Interval, ms	Trace Interval, F or M	Inner Offset, F or M	Source
1	South Texas	1275	48	4	330 F	990 F	Vibroseis*
2	West Texas	1025	120	4	100 F	400 T	Vibroseis
3**	Louisiana	1500	24	4	340 F	340 F	Dynamite
4	Turkey	1275	48	4	100 M	250 M	Vibroseis
5	South America	3000	48	2	100 M	200 M	Dynamite
6	Far East	1250	48	4	100 M	150 M	Dynamite
7	South America	2600	48	2	100 M	300 M	Vibroseis
8	Central America	1300	96	4	50 M	100 M	Dynamite
9	Alaska	1000	96	4	220 F	990 F	Vibroseis
10	North Africa	1325	120	4	25 M	300 M	Vibroseis
11	Alaska	1000	96	4	220 F	990 F	Vibroseis
12	Mississippi	1275	48	4	330 F	990 F	Vibroseis
13	Offshore Offshore	2025	48	4	220 F	875 F	Air gun
14	Offshore Texas	1525	48	4	220 F	690 F	Aquapulse
15	Offshore Canada	2500	48	2	25 M	360 M	Air gun
16	South America	1275	48	4	25 M	233 M	Air gun
17	South America	2000	48	4	50 M	250 M	Air gun
18	Offshore Louisiana	1500	120	4	82 F	716 F	Air gun
19	Turkey	1250	216	4	10 M	50 M	Dynamite
20	South Aleutians	2025	120	4	82 F	921 F	Air gun
21	Denver Basin	1550	48	2	220 F	220 F	Vibroseis
22	Williston Basin	1550	48	2	110 F	110 F	Vibroseis
23	San Juaquin	1550	48	2	220 F	220 F	Vibroseis
24	Arctic	3000	48	2	220 F	220 F	Aquaflex
25	Alberta	2000	96	2	50 M	50 M	Dynamite
26	Alberta	1500	48	2	67 M	67 M	Dynamite
27	Canada	1791	92	4 (1-28) (29-92)	50 M 25 M	200 M	Air gun
28	Canada	2500	48	2	25 M	300 M	Air gun
29	Offshore Spain	2000	48	4	50 M	250 M	Maxipulse
30	Offshore Crete	2125	96	4	25 M	230 M	Air gun
31	North Sea	1550	96	4	25 M	228 M	Air gun
32	North Sea	1550	96	4	25 M	178 M	Air gun
33	North Sea	1625	96	4	25 M	200 M	Air gun
34	Celtic Sea	1500	60	4	50 M	253 M	Air gun
35	Denmark	2500	52	2	100 M	100 M	Dynamite
36	Middle East	1024	48	4	50 M	250 M	Vibroseis
37	Turkey	1000	48	4	75 M	187 M	Vibroseis
38	North Africa	2500	60	2	100 M	100 M	Vibroseis
39	Middle East	2500	60	2	50 M	100 M	Geoflex
40	West Africa	2600	96	2	30 M	120 M	Dynamite

* All vibroseis records have been correlated.
** Analog recording.

at the water bottom.

Record 14 has three identifiable reflections: A, B, and C. Reverberations and multiples also make up a significant portion of the data.

Record 15 is a marine record. The hard water bottom causes refraction arrival A. This shot gather primarily contains guided waves, which are manifested as linear trends such as B, C, and D. The genuine reflection E has little moveout.

Guided waves are trapped within a water layer and travel in the horizontal direction. They are dispersive; i.e., each frequency component travels at a different speed, which is called *horizontal phase velocity*. Their behavior is variable, primarily dictated by water bottom conditions and the thickness of the water layer. They are an important source of coherent noise and are mainly confined to the supercritical region of propagation, where no transmission occurs into the substratum. The nature of guided waves is analyzed in Section 7.3.

Wave packet A in Record 16 is made up entirely of guided waves. Direct arrivals B carry the highest frequency components, while lower frequencies C arrive earlier. Moderate frequencies D make up the later portion of the dispersive wave packet. This record has a reflection E and long-period multiples M1 through M4. The reflection and its multiples also have an accompanying reverberating wave train that is nearly 300 ms long.

Record 17 is longer than the common length (4 to 6 s) used in seismic data acquisition. There is no apparent signal after 4 s; nevertheless, very weak signals can sometimes be uncovered by stacking.

Record 18 has some events worth mention. Dispersive waves A, which include the head wave and direct arrivals make up the early portion of the record. Some reflections, B, C, and D are followed by short-period reverberations. In the deeper part of the record, note the events with extremely large moveout E, which is unusual for deep data. These events represent the side-scattered coherent noise.

Record 19 is a walk-away noise test. It is actually a composite of six shot records. The receiver cable was held constant while the shots were moved away without

(Text continued on page 40)

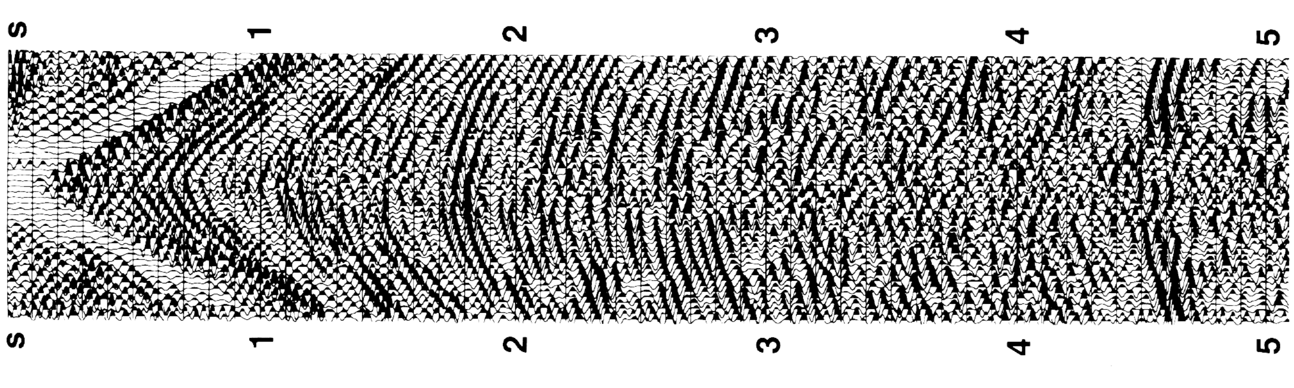
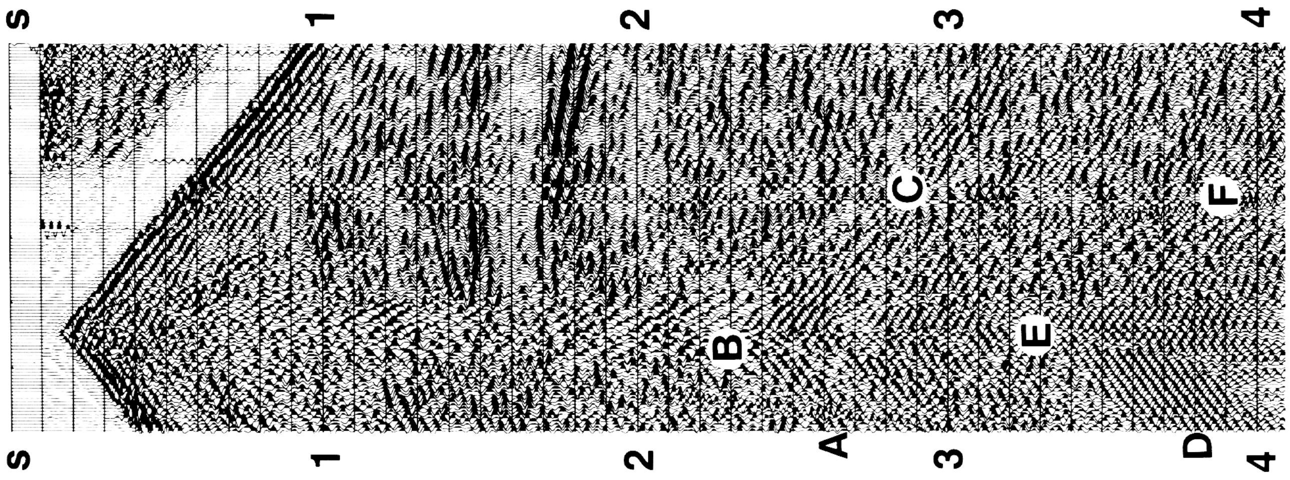
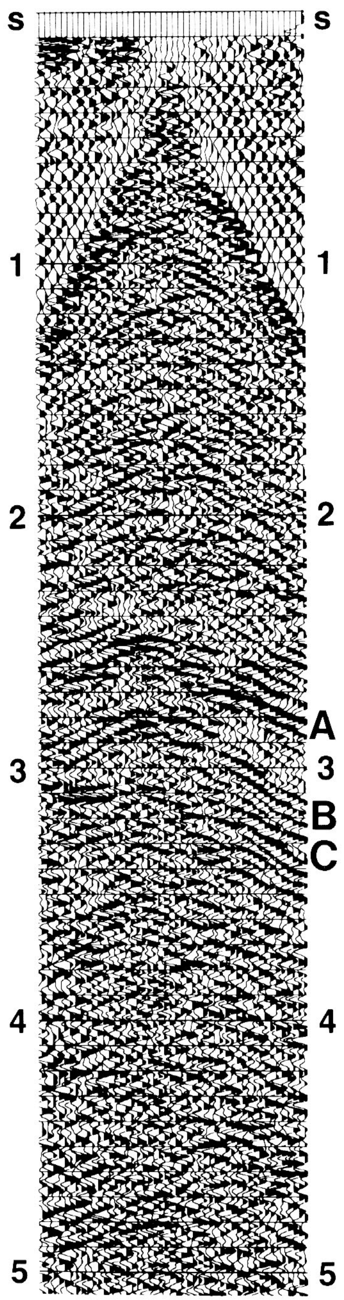


FIG. 1-33. Record 1. For this and the following records, refer to Table 1-8 for recording parameters.

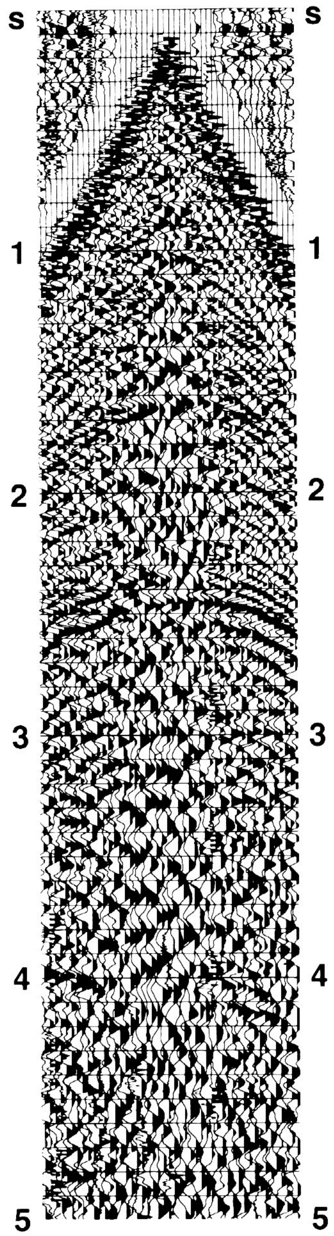


Record 2. Identify events ABC and DEF.

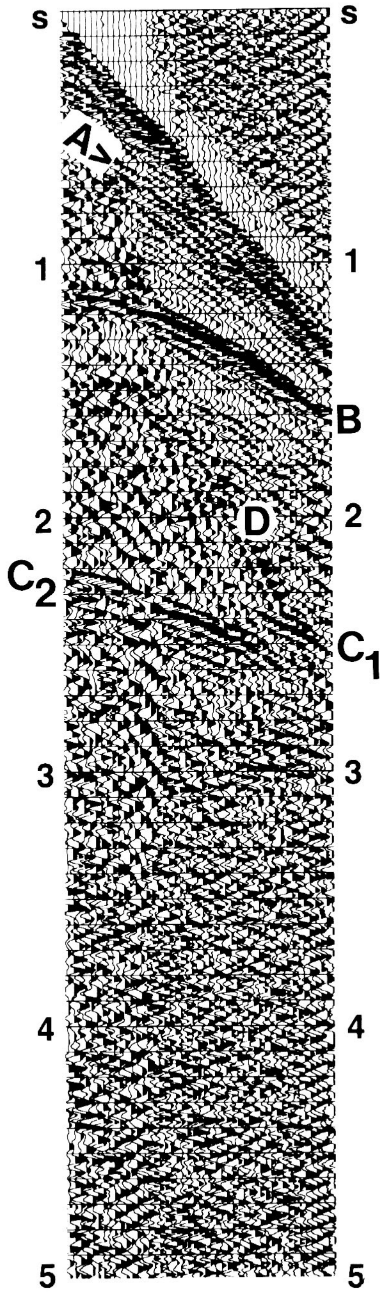
Record 3.



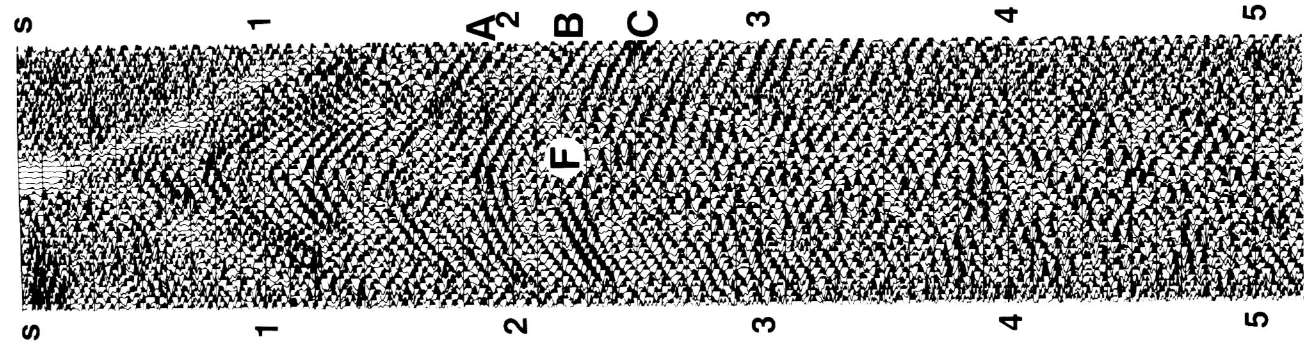
Record 4. What is the cause of the energy preceding the first breaks?



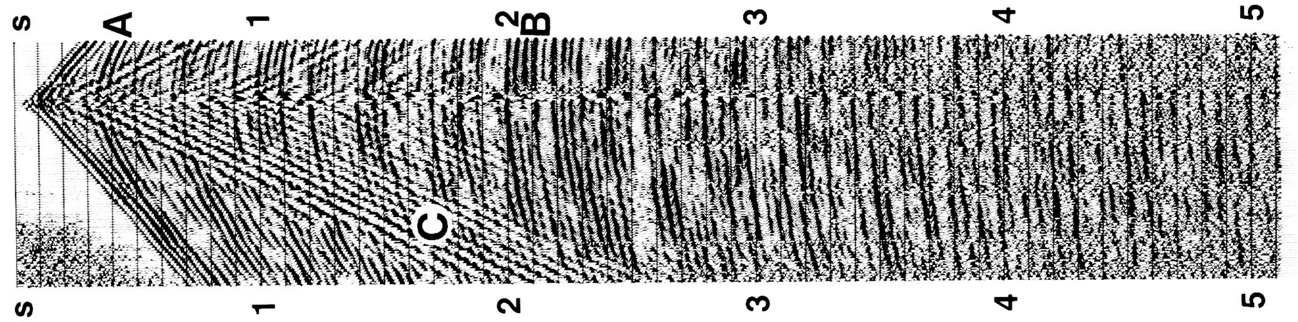
Record 5.



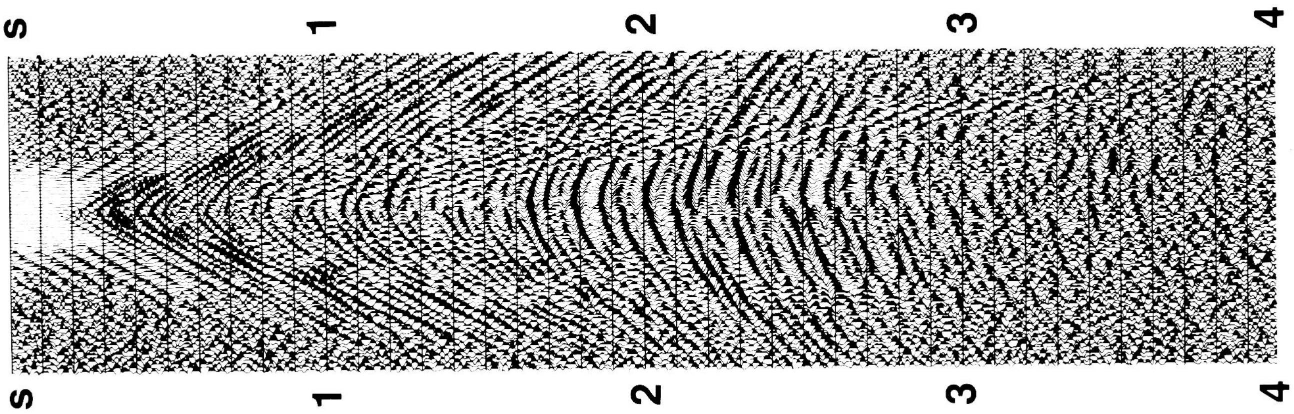
Record 6.



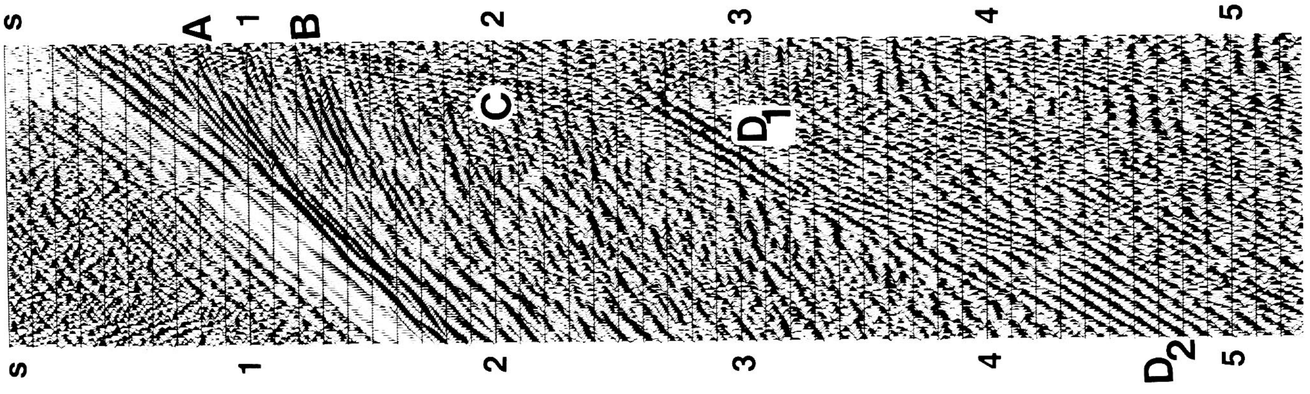
Record 7.



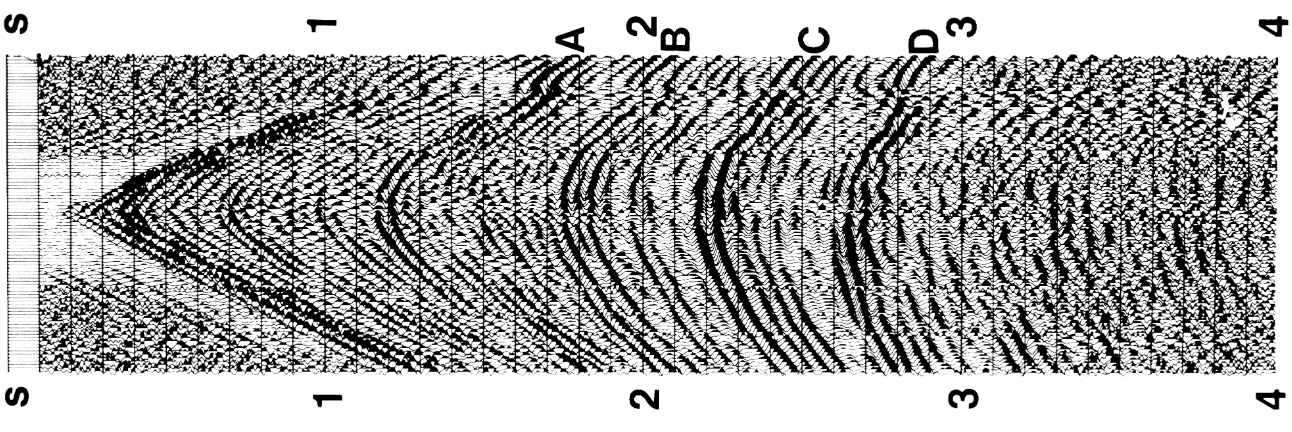
Record 8.



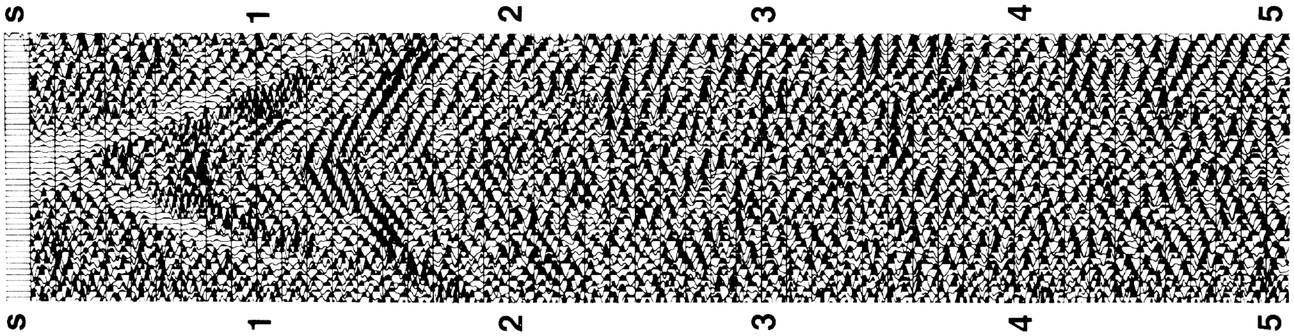
Record 9.



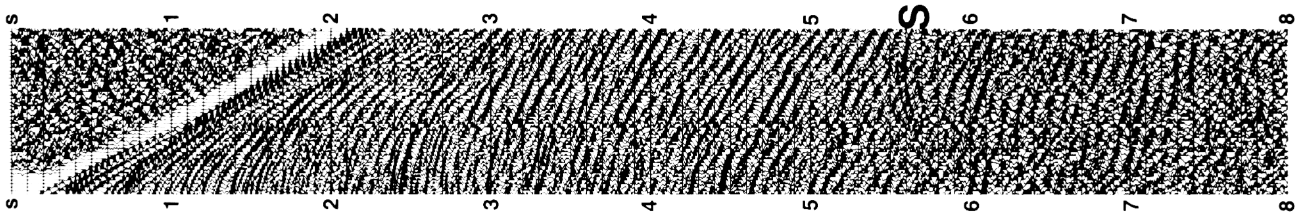
Record 10.



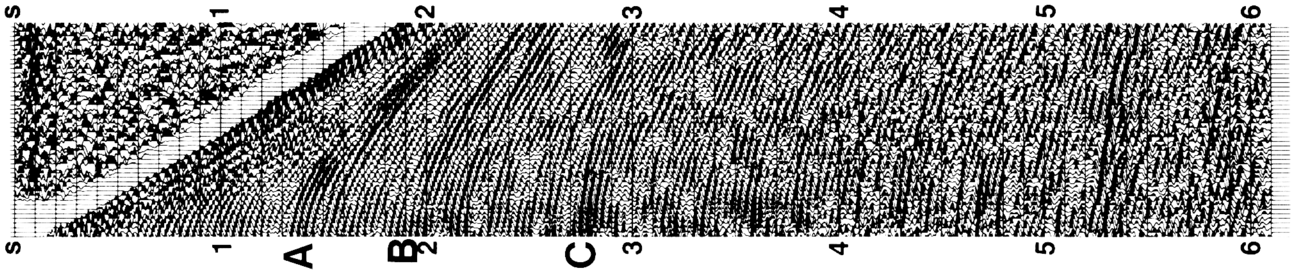
Record 11.



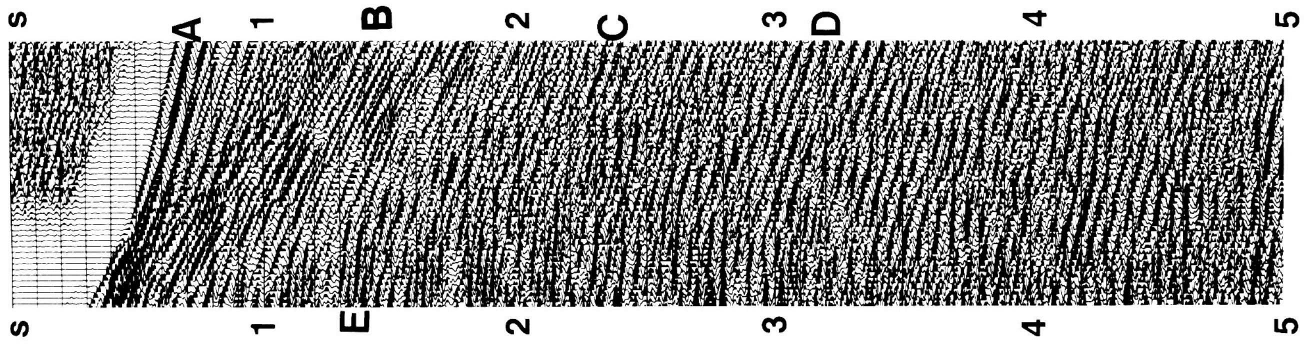
Record 12.



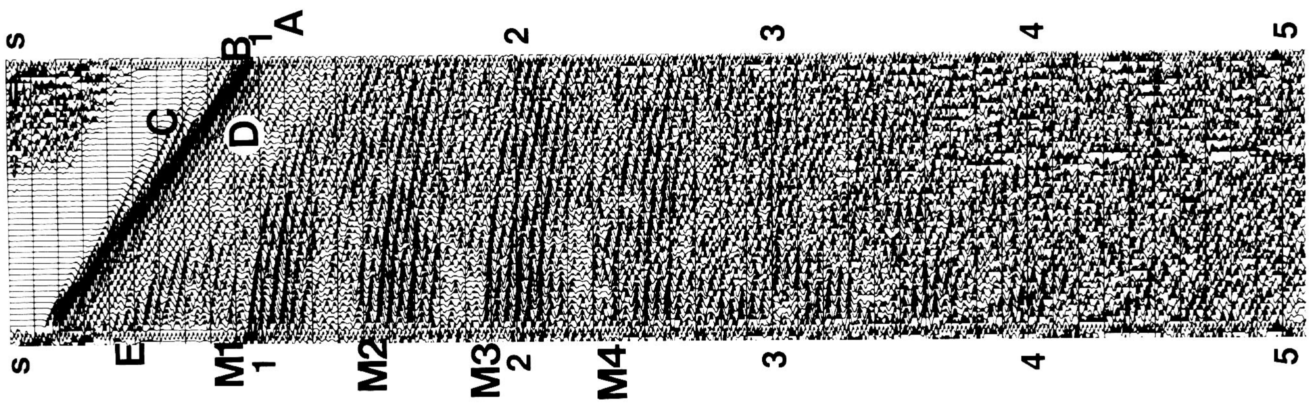
Record 13.



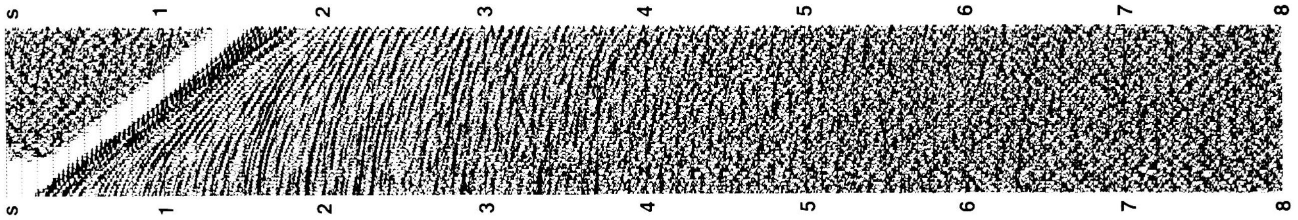
Record 14.



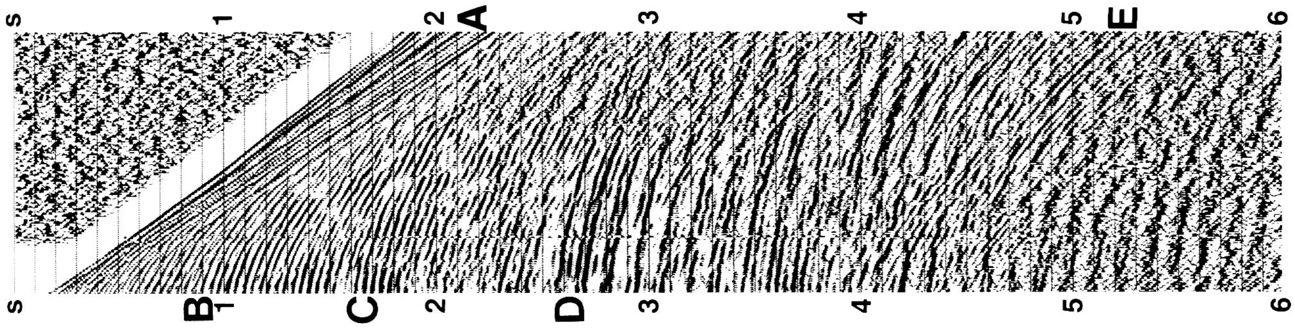
Record 15.



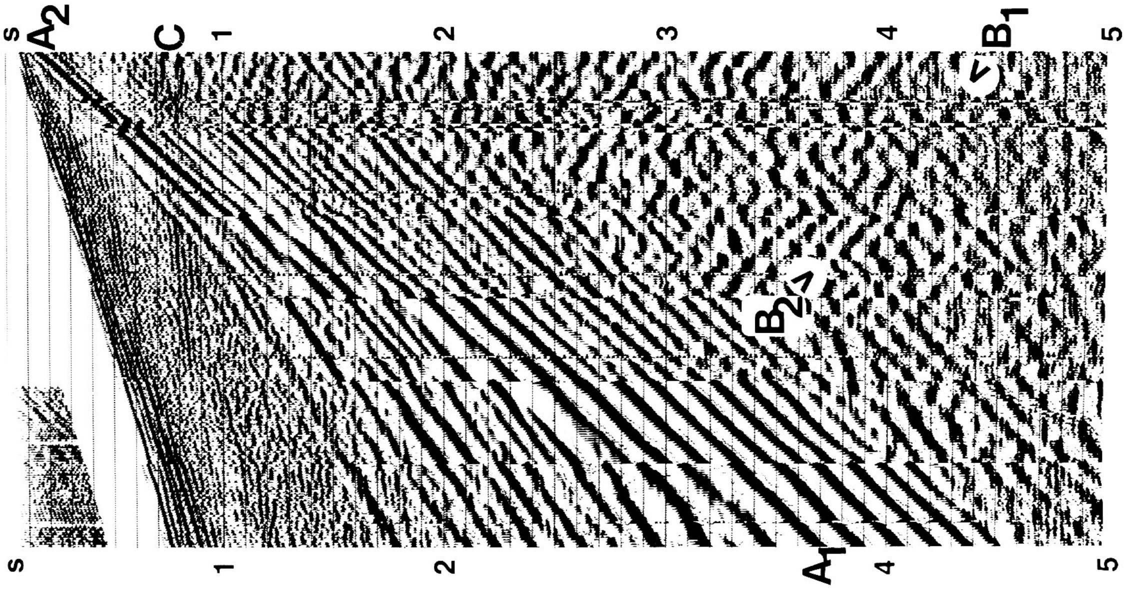
Record 16.



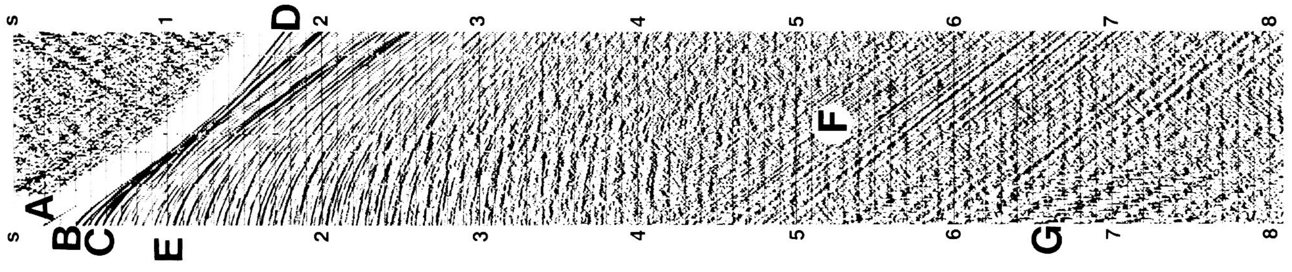
Record 17.



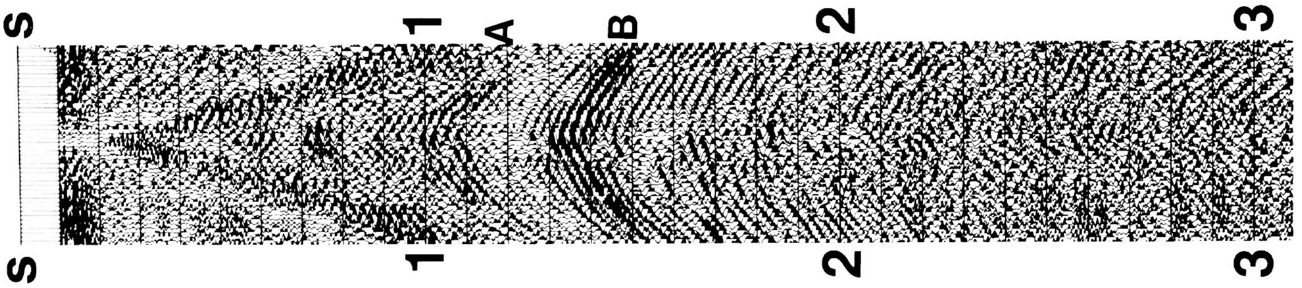
Record 18. Identify the direct wave on this record.



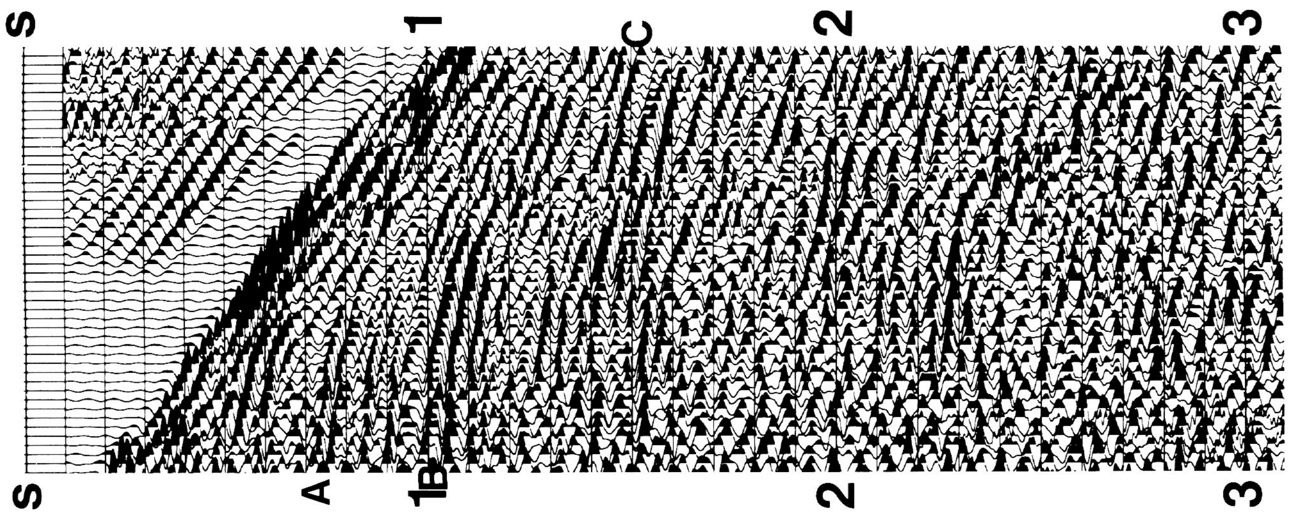
Record 19.



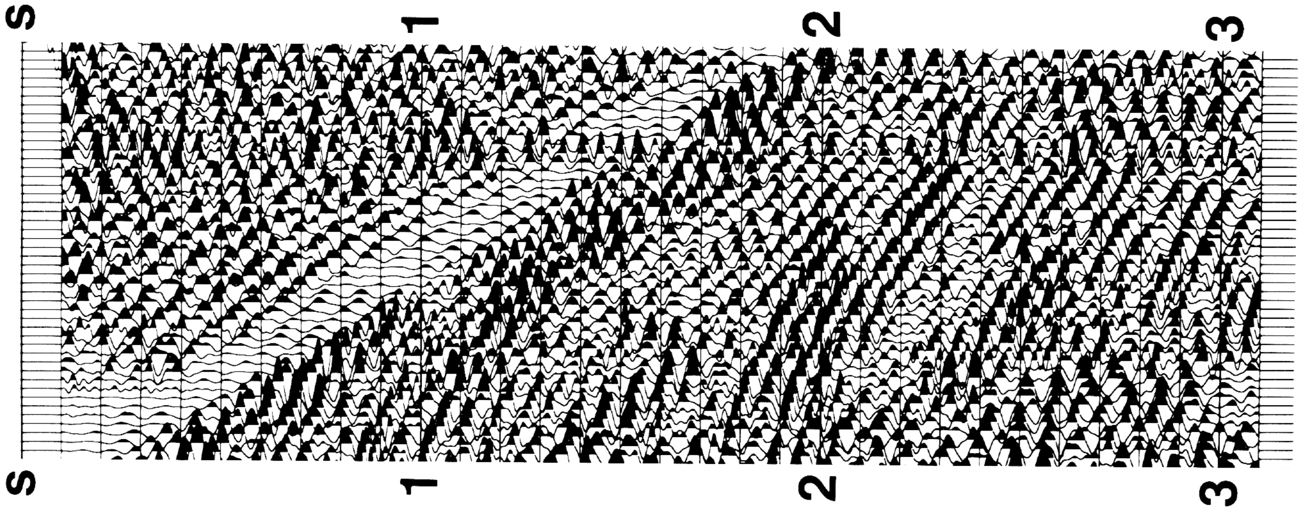
Record 20.



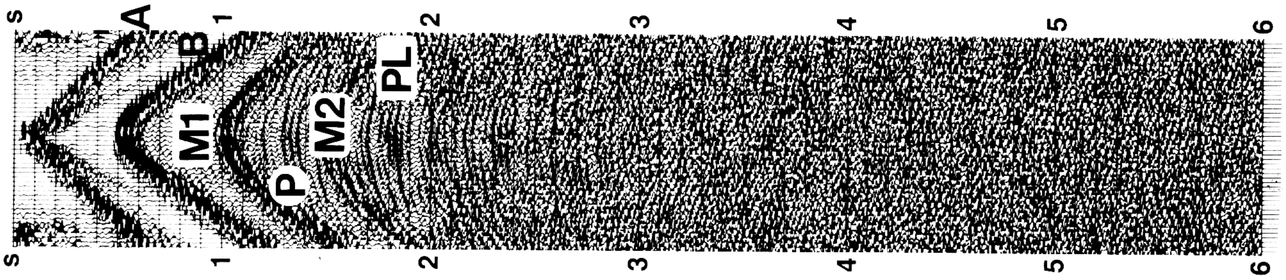
Record 21.



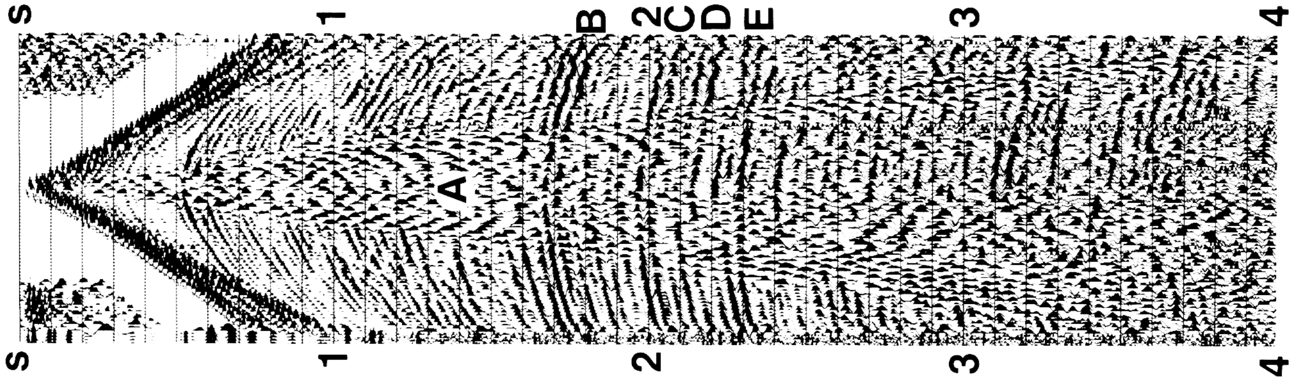
Record 22.



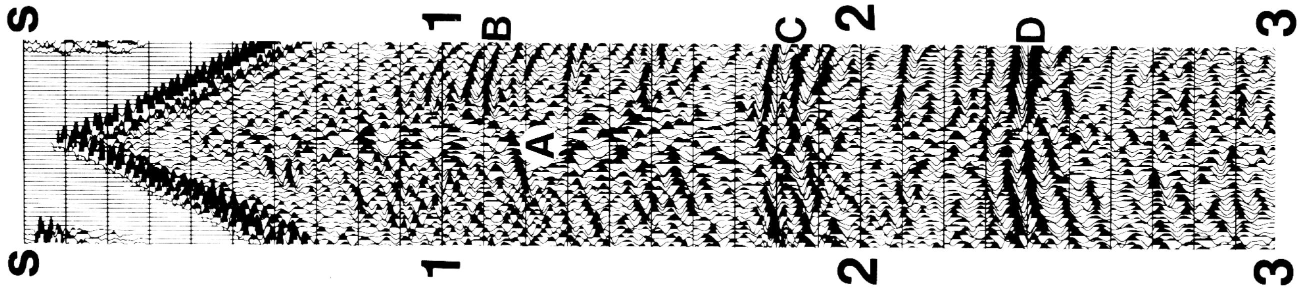
Record 23.



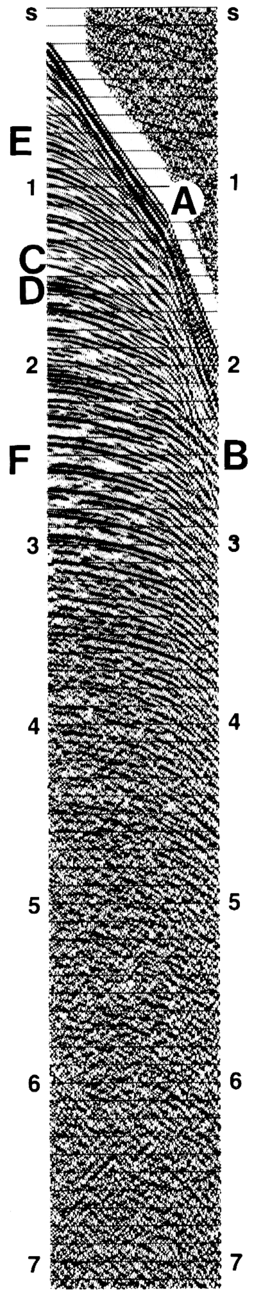
Record 24.



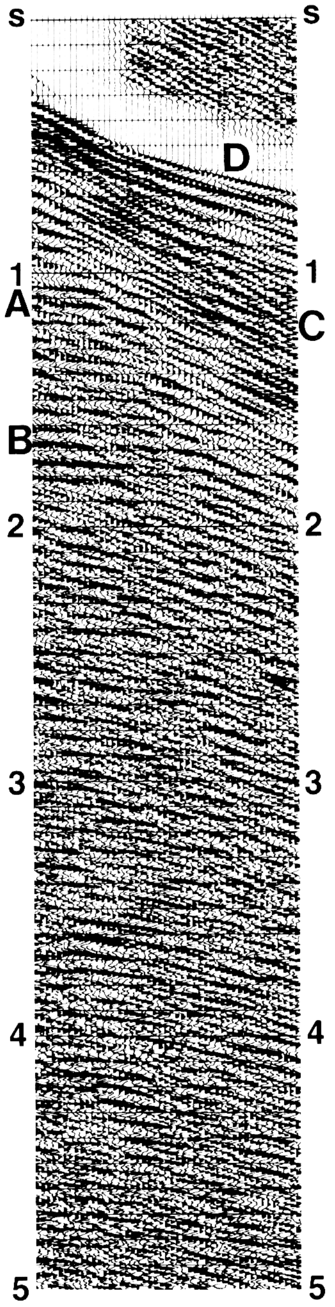
Record 25.



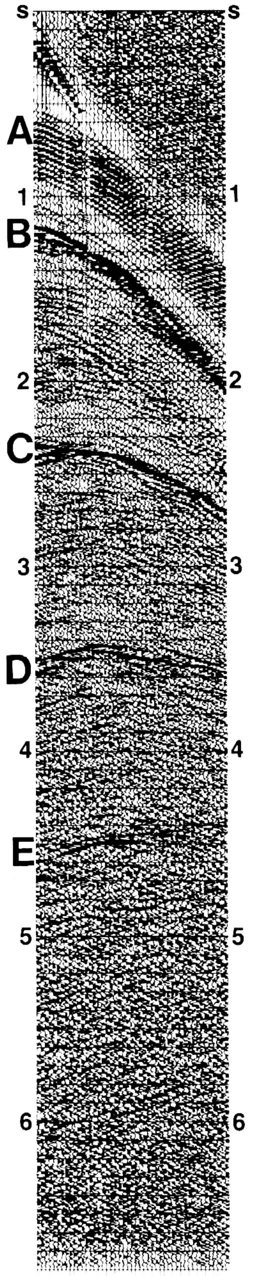
Record 26.



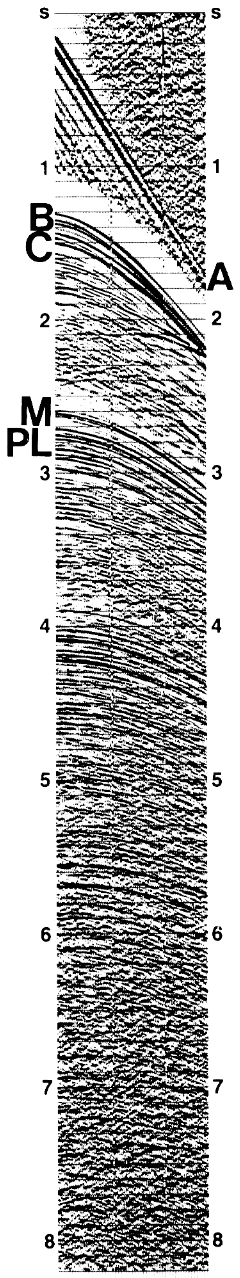
Record 27. Identify the direct wave on this record.



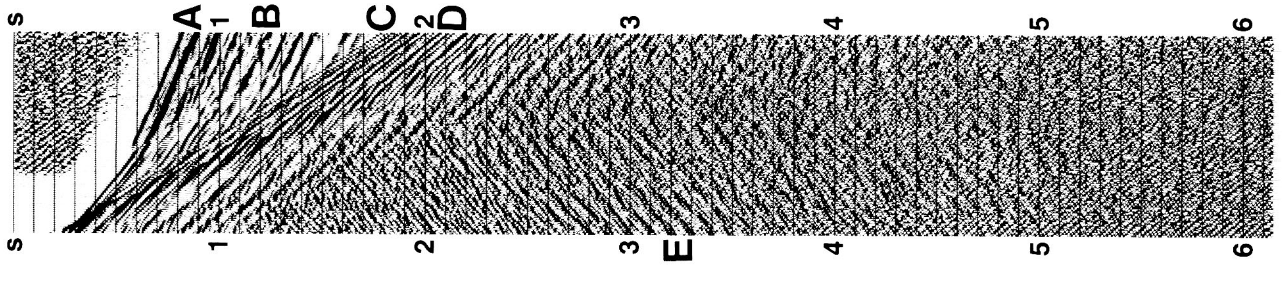
Record 28.



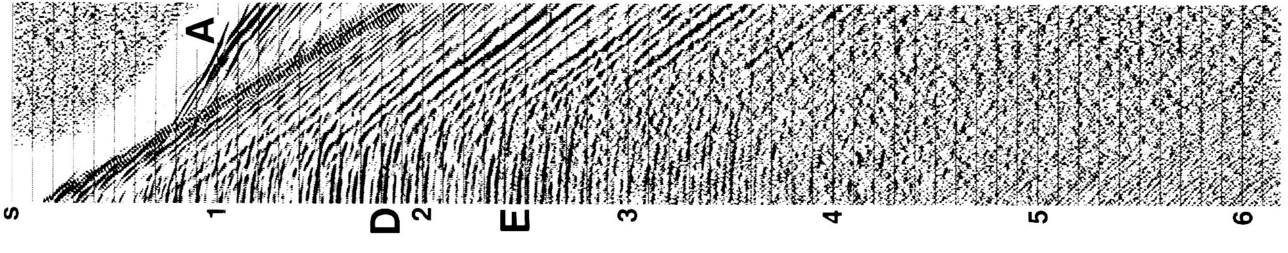
Record 29.



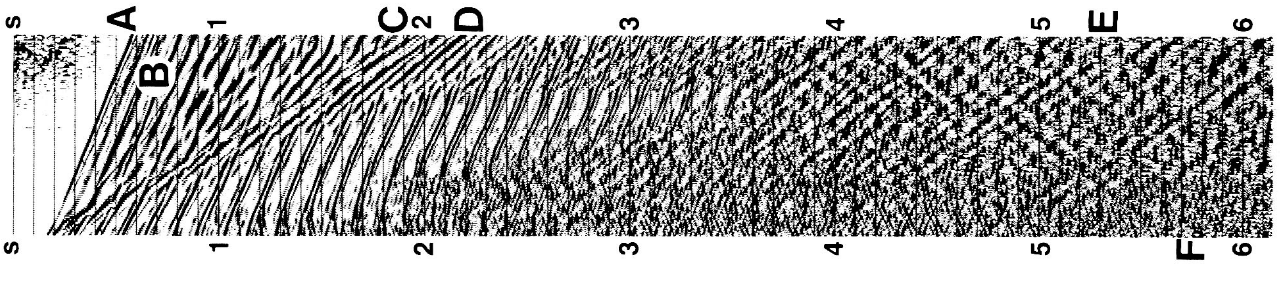
Record 30.



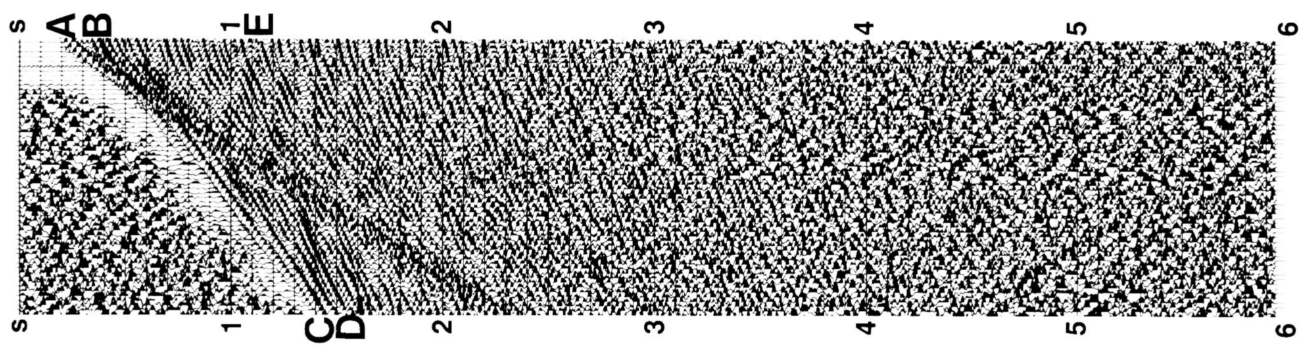
Record 31.



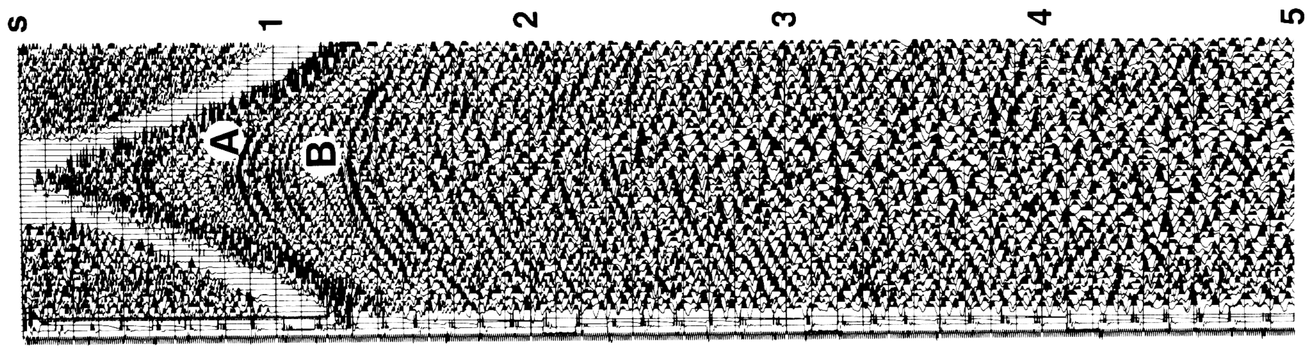
Record 32.



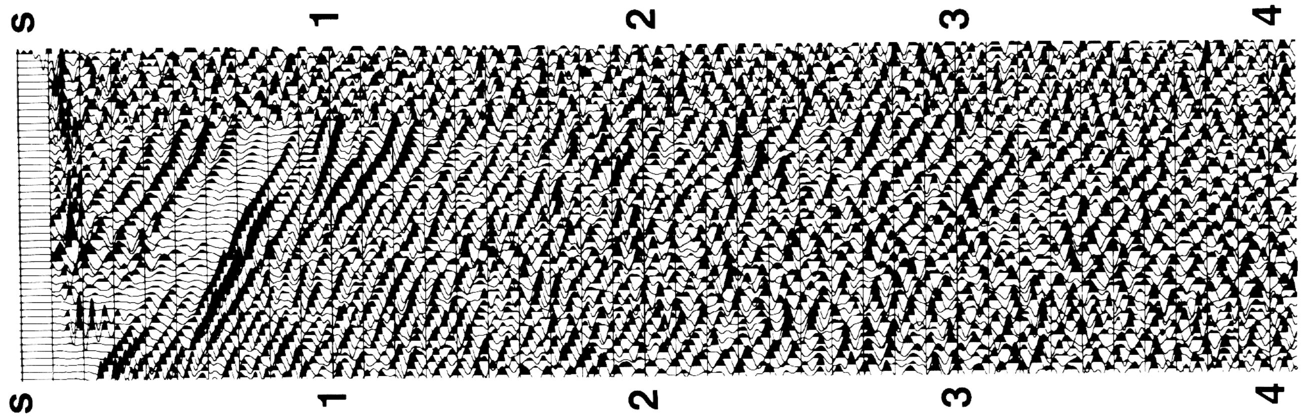
Record 33.



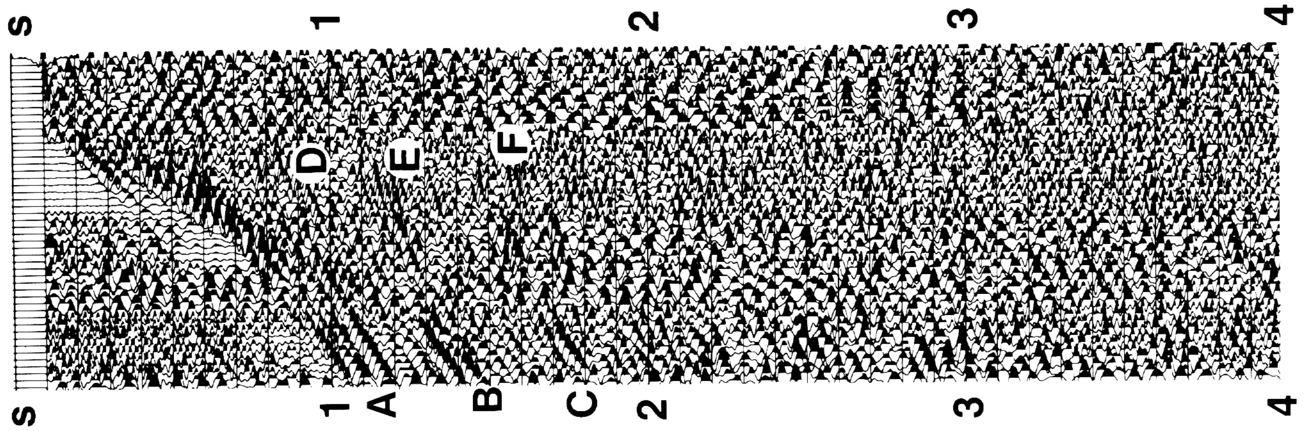
Record 34.



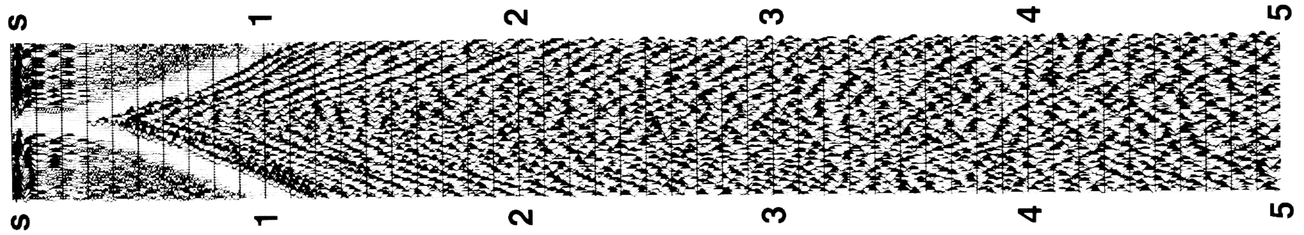
Record 35.



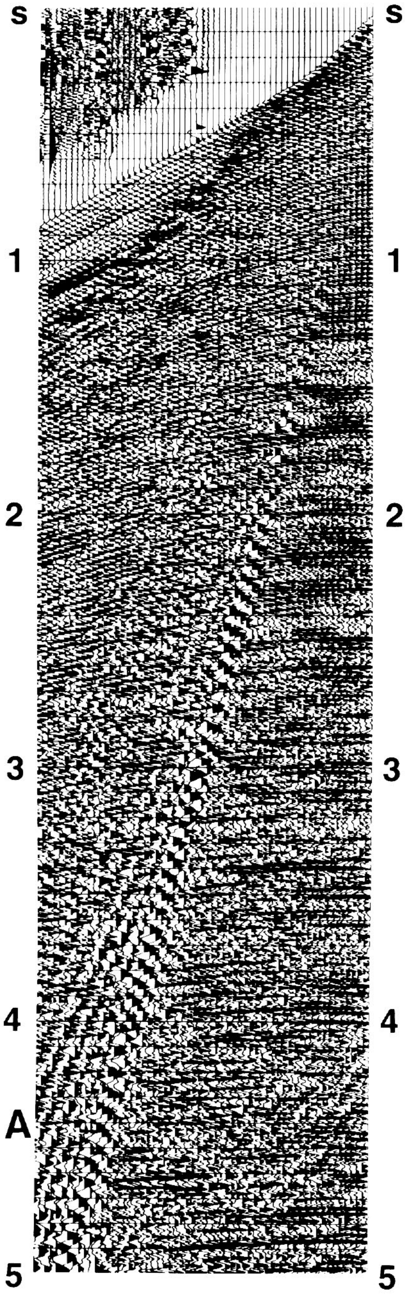
Record 36.



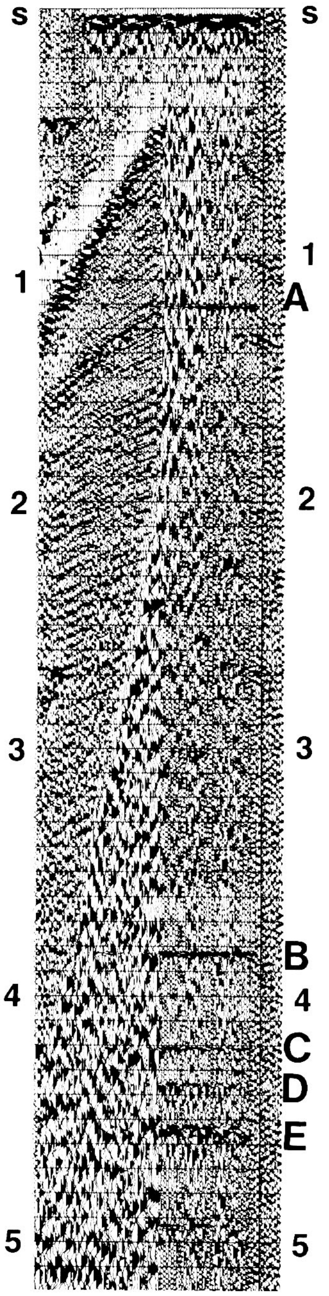
Record 37.



Record 38.



Record 39.



Record 40.

overlap. The receiver group interval is 10 m. The receivers in each group were bundled together without array forming. This allowed all signal and noise wavelengths to be recorded. Horizontal wavelength is determined by measuring the dominant frequency (the reciprocal of the time between successive peaks or troughs) and the horizontal phase velocity (the reciprocal of stepout $\Delta t/\Delta x$ of the unwanted ground roll. The horizontal wavelength then is used to design the receiver array length that is needed to suppress this energy (see Exercise 1.2). Wave packet A_1A_2 (between 1.7 and 4.6 on the far-left trace) is an excellent example of ground roll. The linear coherent energy with opposite dip B_1B_2 is the backscattered component. Reflection C is being disrupted by ground roll energy.

Ground roll is different from guided waves, although both are dispersive. Ground roll is one type of Rayleigh wave that arises because of the coupling of compressional waves (*P*) and the vertical component of shear waves (*SV*) that propagate along the free surface (Grant and West, 1965). On the other hand, guided waves are one kind of compressional wave that travels within a layer just as sound waves travel in an organ pipe.

Record 20 (marine) shows a variety of wave types. Direct arrivals A are significantly suppressed by receiver arrays in the field. We can see the water bottom reflection B on short-offset traces. Note a shallow reflector C and associated refraction arrival D. At 1 s, another reflector E is seen. Much of the energy between 1 and 3 s most likely is multiples associated with B, C, and E. Linear noise (possibly cable noise) F and the lower-frequency propeller noise G appear in the deeper portion of the record after 4 s.

Record 21 (vibroseis) has a weak A and a strong B shallow reflector. Below 2 s, ambient noise dominates the record. Record 22 is another vibroseis record. Note reflection arrivals A, B, and C. Although reasonably hyperbolic, there are some fluctuations in traveltimes that may be attributed to near-surface complexity (as inferred by the first breaks). Record 23 shows similar characteristics.

Record 24 is a marine shot gather that was acquired with an Aquaseis source. Direct arrivals A, the water-bottom reflection B, and first-order multiples M1, M2, are easily recognized. A primary reflection P and its peg-leg multiple PL also are distinguished.

Record 25 (land) has very good signal quality. In addition to the several primary reflections, note ground roll energy A. It is predominantly low frequency and travels with low group velocity (the speed with which the energy in a wave packet travels). Also note the near-surface effects that cause traveltime distortions along the right flanks of reflections B, C, D, and E. Record 26 (obtained by using dynamite) does not have a well-developed ground roll energy; however, it is still recognizable from its low-frequency character A. Traveltime paths that correspond to reflections (for example, B, C, and D) have been disrupted by ground roll and possibly distorted by irregularities in the near surface.

Record 27 (marine) is interesting. Note the change in the cable geometry A (see Table 1-8). There is a well-developed dispersive wave packet B that spans between 1.9 and 2.9 s at the far-offset trace. This includes the head wave and direct arrivals. Also note the distinct moveout difference between events C and D. Event C, with a larger moveout, belongs to the short-period multiple wave train that is associated with the water-bottom reflection. Event D, with a small moveout, is a primary reflection with its own peg-leg series F.

The air-gun data in Record 28 contain high-velocity reflections with little moveout. Note the predominant guided wave packet C that spans between 0.7 and 1.9 s at the far-offset trace. It is due to the strong water-bottom refractor D.

Record 29 is puzzling (see Exercise 1.4). The skewness of the reflection hyperbolas (B, C, D, and E) increases in depth.

Record 30 is a deep-water shot record. The direct arrival A, water-bottom reflection B, and shallow reflection C, can be identified easily. First-order water-bottom multiples M and peg-leg multiples PL, which are associated with the shallow reflector C, also are prominent in this record.

Record 31 is a shot record primarily made up of guided waves. The following elements are identified: A is refraction arrival, B is its multiples, C is direct arrival, D is the dispersive medium-frequency components of guided waves between 1.8 and 3 s at the far-offset trace, and E is backscattered energy.

Record 32 is another marine record that contains strong guided wave energy. A refractor A, direct arrival B, and dispersed wave packet C span between 1.3 and 4 s at the far-offset trace. There also is subcritical reflection energy, most of which is reverberation D, E.

The events on Records 33, 34, and 35 are referred to in Exercises 1-7, 1-8, and 1-9. The leftmost four traces in Record 35 are associated with the recording channels that are used to record auxiliary information.

Record 36 seems to have no events with hyperbolic moveout. Record 37 has a few reflections, AD, BE, CF; however, they are buried in strong ambient noise. Record 38 contains virtually no reflections. A strong dispersive wave (ground roll) makes up the early part of the record, while the remaining part contains primarily random noise. Processing cannot yield signal from field data without signal. At best, it suppresses whatever noise is in the field data and brings up the reflection energy that is buried in the noise. Seismic data must not be acquired with the attitude, "Don't worry, processing will bring out the signal."

Record 39 is a Geoflex record containing strong ground roll energy A. Additionally, the record from top to bottom contains a strong high-frequency reverberating wave train and short-period multiples that are associated with the water bottom and, perhaps, a few shallow reflectors. Record 40 has a small usable segment; i.e., the longer offset traces on the left side between 1 and 4 s. The remaining part of the record contains strong random noise and transient noise, A, B, C, D, and E, which are attributed to electronic instrument noise (possibly resulting from weather conditions).

To summarize, field records contain (a) reflections, (b) coherent noise, and (c) random ambient noise. One important aspect of data processing is to uncover genuine reflections by suppressing noise of various types.

Reflections are recognized by their hyperbolic traveltimes. If the reflecting interface is horizontal, then the apex of the reflection hyperbola is situated at zero offset. On the other hand, if it is a dipping interface, then the reflection hyperbola is skewed in the updip direction.

There are several wave types under the coherent noise category:

1. Ground roll is recognized by low frequency, strong amplitude, and low group velocity. It is the vertical component of dispersive surface waves. In the field, receiver arrays are used to eliminate ground roll. Ground roll can have strong backscattered components due to lateral inhomogeneities in the near-surface layer.
2. Guided waves are persistent, especially in shallow marine records in areas with hard water bottom. The water layer makes a strong velocity contrast with the substratum, which causes most of the energy to be trapped within and guided laterally through the water layer. The dispersive nature of these waves makes them easily recognized on shot records. Guided waves also make up the early arrivals. The stronger the velocity contrast between

the water layer and the substratum, the smaller the critical angle; thus, more guided-wave energy is trapped in the supercritical region. When there is a strong velocity contrast, refraction energy propagates in the form of a head wave. Guided waves also are found in land records. These waves are largely attenuated by CMP stacking. Because of their prominently linear moveout, in principle they also can be suppressed by dip filtering techniques. One such filtering technique is based on 2-D Fourier transformation of the shot record. This is discussed in Section 1.6.2. Another approach is based on slant stacking, which is described in Section 7.4.

3. Side-scattered noise commonly occurs at the water bottom, where there is no flat, smooth topography. Irregularities of varying size act as point scatterers, which cause diffraction arrivals with table-top trajectories. They can be on or off the vertical plane of the recording cable. These arrivals typically exhibit a large range of moveouts, depending on the spatial position of the scatterers in the subsurface.
4. Cable noise is linear and low in amplitude and frequency. It primarily appears on shot records as late arrivals.
5. The air wave with a 300 m/s velocity can be a serious problem when shooting with surface charges such as Geoflex, Poulter, or land air gun. Perhaps the only effective way to remove air waves is to zero out the data on shot gathers along a narrow corridor containing this energy (notch muting). It often is impossible to recover any data arriving after the air wave on Poulter data.
6. Power lines also cause noisy traces in the form of a monofrequency wave. A monofrequency wave may be 50 or 60 Hz, depending on where the field survey was conducted. Notch filters often are used in the field to suppress such energy.
7. Multiples are secondary reflections with interbed or intrabed raypaths. Guided waves include supercritical multiple energy. Multiples are attacked by methods, which are based on moveout discrimination, and prediction theory, which uses the periodic behavior of multiples. The most effective moveout-based suppression technique often is CMP stack with inside-trace mute (Section 8.2). Prediction theory should be particularly effective, at least in theory, in the slant-stack domain (Section 7.5).

Random noise has various sources. A poorly planted geophone, wind motion, transient movements in the vicinity of the recording cable, wave motion in the water that causes the cable to vibrate, and electrical noise from the recording instrument – all can cause ambient noise. The net result of scattered noise from many scatterers in the subsurface also contributes to random noise (Larner et al., 1983).

In Section 1.5, it is noted that energy propagating within the earth is subject to decay in amplitude because of wavefront divergence and frequency-dependent absorption from the intrinsic attenuation of rocks. Signal strength therefore decreases in time, while random noise persists and eventually dominates. Unfortunately, gain corrections to restore signal strength at later times boost random noise in the process. Fortunately, CMP stack suppresses a significant part of the uncorrelated random noise (Section 1.4.3).

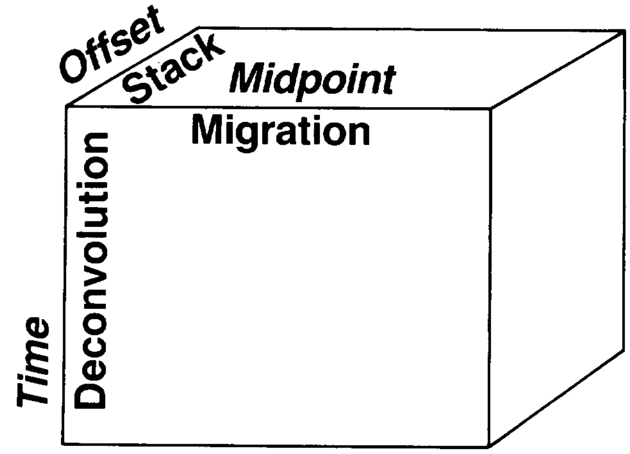


FIG. 1-34. Seismic data volume represented in processing coordinates, midpoint-offset-time. Deconvolution acts on the data along the time axis and increases temporal resolution. Stacking compresses the data volume in the offset direction and yields the plane of stacked section (the front face of the prism). Migration then moves dipping events to their true subsurface positions and collapses diffractions, thus increasing lateral resolution.

1.4 BASIC DATA PROCESSING SEQUENCE

Since the introduction of digital recording, a routine sequence in seismic data processing has evolved. This basic sequence now is described to gain an overall understanding of each step.

There are three primary stages in processing seismic data. In their usual order of application, they are:

1. Deconvolution
2. Stacking
3. Migration.

Figure 1-34 represents the seismic data volume in processing coordinates; namely, midpoint-offset-time. Deconvolution acts along the time axis. It removes the basic seismic wavelet (the source time function modified by various effects of the earth and recording system) from the recorded seismic trace and thereby increases temporal resolution. Deconvolution achieves this goal by compressing the wavelet. Stacking also is a process of compression. In particular, the data volume in Figure 1-34 is reduced to a plane of midpoint-time at zero offset (the frontal face of the box) first by applying normal moveout correction to traces from each CMP gather

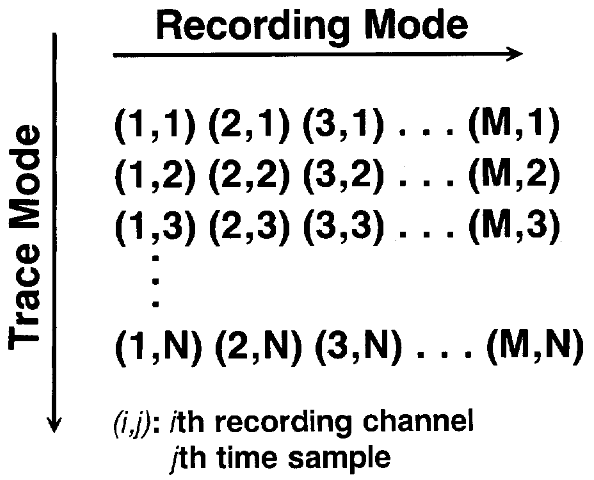


FIG. 1-35. Seismic data are recorded in rows of samples; samples at the same time at consecutive channels. Demultiplexing involves sorting the data into columns of samples—all the time samples in one channel followed by those in the next channels.

(Section 3.2), then by summing them along the offset axis. The result is a *stacked section*. (The terms *stacked section*, *CMP stack*, and *stack* often are used synonymously.) The CMP stack is an approximation to zero-offset section. Finally, migration commonly is applied to stacked data. Migration is a process that collapses diffractions and maps dipping events on a stacked section to their true subsurface locations. In this respect, migration is a spatial deconvolution process that improves spatial resolution.

All other processing techniques may be considered secondary in that they help improve the effectiveness of the primary processes. For example, dip filtering may need to be applied before deconvolution to remove coherent noise so that the autocorrelation estimate is based on reflection energy that is free from such noise. Wide band-pass filtering also may be needed to remove very low- and high-frequency noise. Before deconvolution, correction for geometric spreading is necessary to compensate for loss of amplitude due to wavefront divergence. Velocity analysis, which is an essential step for stacking, can be improved by multiple suppression and residual statics corrections.

Keep in mind that the success of a process depends not only on the proper choice of parameters pertinent to that particular process, but also on the effectiveness of the previous processing stages. As outlined above, conventional processing is based on certain assumptions. Many of the secondary processes are designed to make data compatible with the assumptions of the three primary processes. Deconvolution assumes a stationary, vertically incident, minimum-phase source wavelet and white reflectivity series that is free of noise. Stacking assumes hyperbolic moveout, while migration is based on a zero-offset (primaries only) wave field assumption. A pessimist could claim that none of these assumptions is valid. He may be right; however, when applied, these techniques do provide results that are close to the true subsurface picture. This is because these three processes are robust and their performance is not very sensitive to the underlying assumptions in their theoretical development.

The following sequence describes basic seismic data processing.

1.4.1 Preprocessing

Field data are recorded in a multiplexed mode using a certain type of format. The data first are demultiplexed as described in Figure 1-35. Mathematically, demultiplexing is seen as transposing a big matrix so that the columns of the resulting matrix can be read as seismic traces recorded at different offsets with a common shotpoint (Figure 1-36). At this stage, the data are converted to a convenient format that is used throughout processing. This format is determined by the type of processing system and the individual company.

Preprocessing also involves trace editing. Noisy traces, traces with transient glitches (see Figure 1-33, Record 40), or monofrequency signals (see Figure 1-33, Record 3) are deleted; polarity reversals (see Figure 1-33, Record 2) are corrected.

A gain recovery function is applied on the data to correct for the amplitude effects of wavefront (spherical) divergence (Figure 1-37). This amounts to applying a geometric spreading function, which depends on traveltime, and an average primary velocity function, which is associated with primary reflections in a particular survey area (Section 1.5). Additionally, an exponential gain function may be used to compensate for attenuation losses. As an option, it may be desirable to filter the data with a wide band-pass filter before deconvolution.

Finally, field geometry is incorporated with the seismic data. This may precede any gain correction that is offset-dependent. Based on surveying information, coordinates of shot and receiver locations for all traces are stored on trace headers. Changes in shot and receiver locations are properly handled based on the information available in the observer's log. Many types of processing problems arise from incorrectly setting up the field geometry. Regardless of how meticulously processing parameters are chosen, the quality of a stacked section can be degraded severely because of incorrect field geometry. For land data, elevation statics are applied at this stage to reduce traveltimes to a common datum level. This level may be flat or vary (float) along the line.

1.4.2 Deconvolution

The step that follows preprocessing is deconvolution. Typically, prestack deconvolution is aimed at improving temporal resolution by compressing the effective source wavelet contained in the seismic trace to a spike (spiking deconvolution). Predictive deconvolution (Sections 2.6

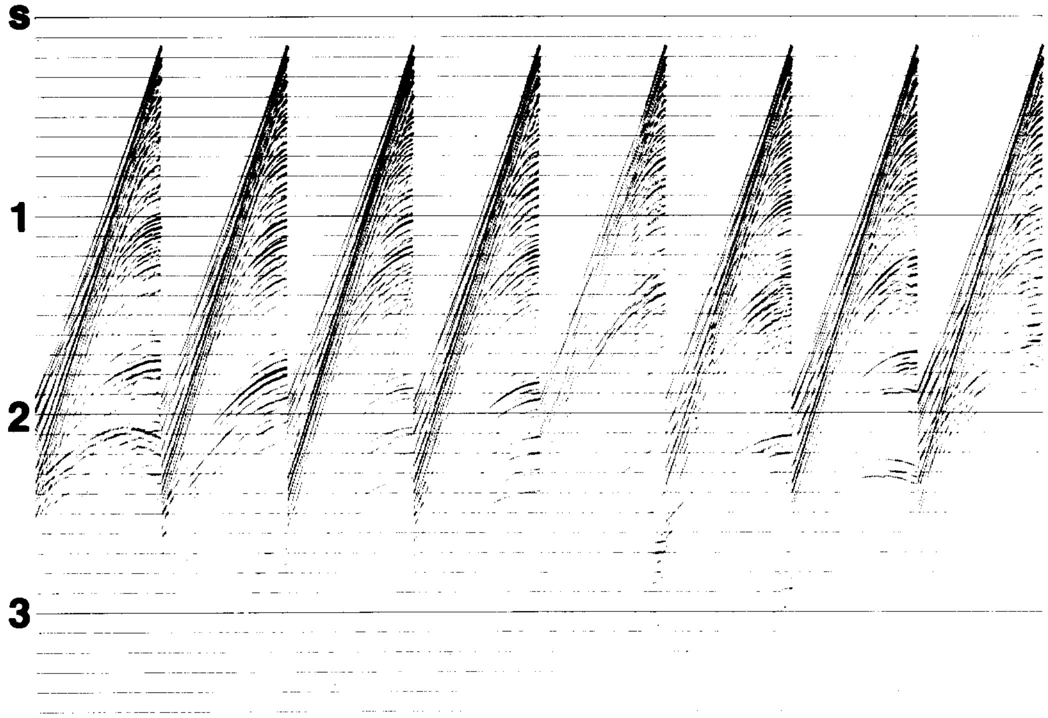


FIG. 1-36. Common-shot gathers just after demultiplexing. These are from an offshore survey. Note the strong amplitudes at the early part and the relatively weaker energy at the deeper part of the records. Such decay in amplitude primarily is due to wavefront divergence.

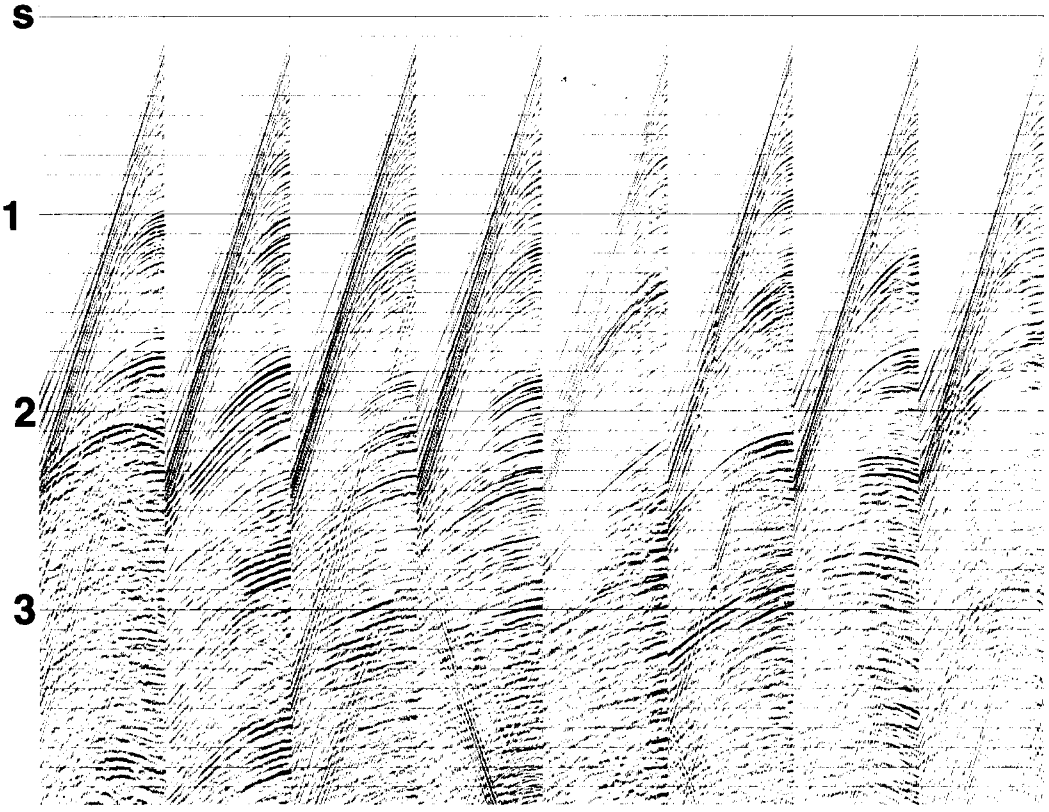


FIG. 1-37. Common-shot gathers of Figure 1-36 after correcting for the amplitude effects of wavefront divergence. Note the restored amplitudes at late times on the records.

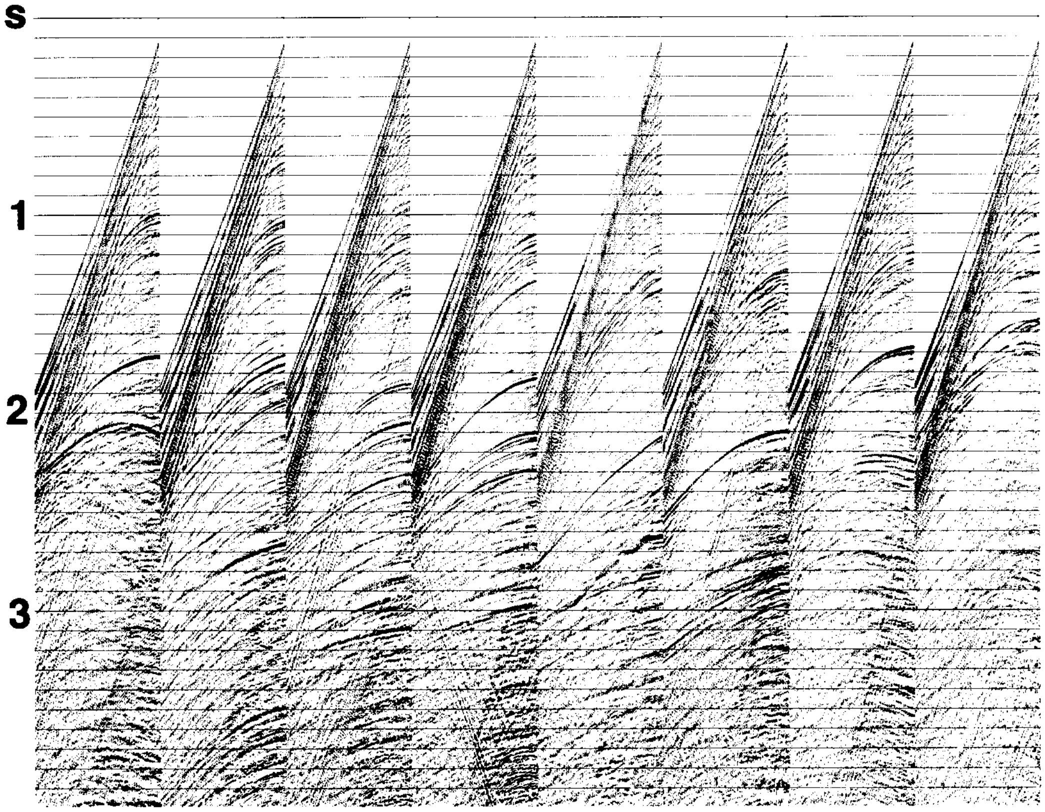


FIG. 1-38. Common-shot gathers of Figure 1-37 after spiking deconvolution. By examining some of the individual reflections and comparing them with Figure 1-37, note how the wavelet and reverberatory energy that trails behind each reflection is significantly suppressed by deconvolution.

and 2.7) with a prediction lag (commonly termed gap) that is equal to the first or second zero crossing of the autocorrelation function also is commonly used. Deconvolution techniques used in conventional processing are based on optimum Wiener filtering (Section 2.6).

Figure 1-38 shows the common-shot gathers after spiking deconvolution. Wavelet compression is most obvious when compared with the gathers in Figure 1-37. Since spiking deconvolution broadens the spectrum of seismic data, traces contain much more high-frequency energy after deconvolution. Because both high-frequency noise and signal are boosted, the data often need filtering with a wide band-pass filter after deconvolution. In addition, some kind of trace balancing is used after deconvolution to bring the data to a common root-mean-squared (rms) level, as shown in Figure 1-39. (Balancing is described in Section 1.5.)

1.4.3 CMP Sorting

After the initial signal processing (described above), the data are transformed from shot-receiver to midpoint-offset coordinates. This is CMP sorting, which requires field geometry information. Note that the term *common depth point* (CDP) often is used instead of common midpoint.

Seismic data acquisition with multifold coverage is done in shot-receiver (s,g) coordinates. Figure 1-40 is a schematic depiction of the recording geometry. On the other hand, seismic data processing conventionally is done in midpoint-offset (y,h) coordinates. The required coordinate transformation is achieved by sorting the data into CMP gathers. Each individual trace is assigned to the midpoint between the shot and receiver locations associated with that trace. Those traces with the same midpoint location are grouped together, making up a CMP gather. Figure 1-41 depicts the geometry of a CMP gather. Note that CDP gather is equivalent to a CMP gather only when reflectors are horizontal and velocities do not vary horizontally. However, when there are dipping reflectors in the subsurface, these two gathers are not equivalent and only the term CMP gather should be used. Selected CMP gathers obtained from sorting the deconvolved shot gathers (Figure 1-39) are shown in Figure 1-42.

Figure 1-43 shows the superposition of shot-receiver (s,g) and midpoint-offset (y,h) coordinates. The (y,h) coordinates have been rotated 45 degrees relative to the (s,g) coordinates. The dotted area represents the coverage used in recording the seismic profile along the midpoint axis, Oy. Each dot represents a seismic trace with the time axis perpendicular to the plane of paper. The

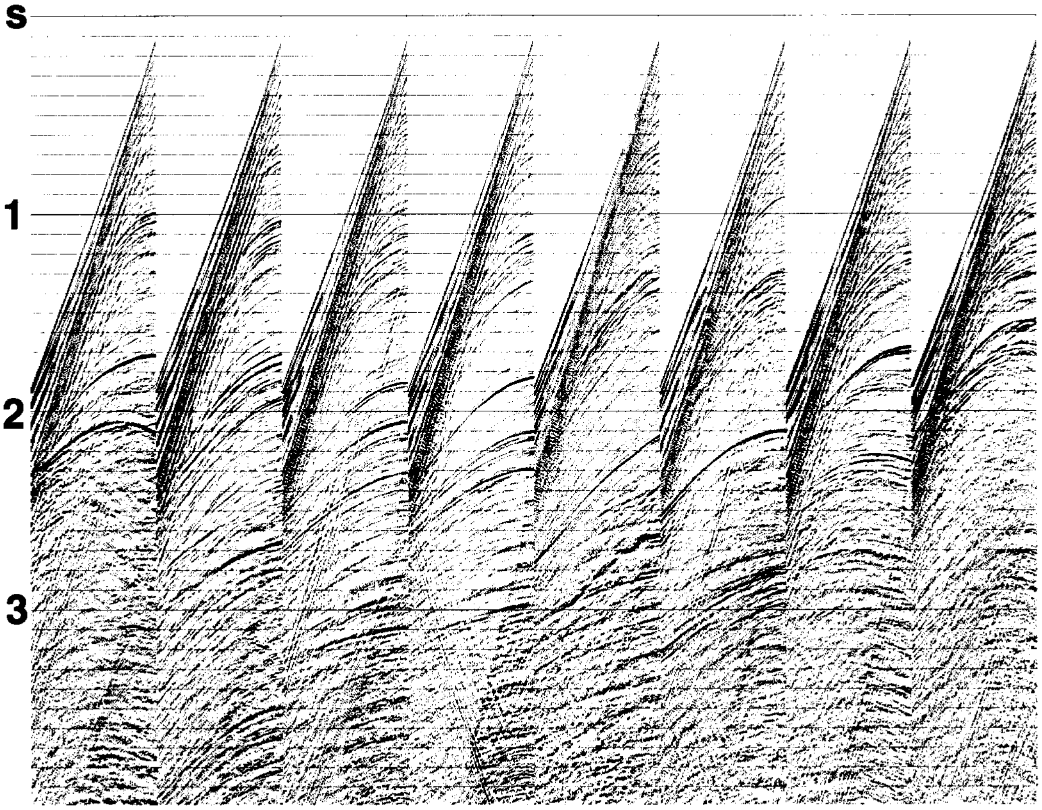


FIG. 1-39. Common-shot gathers of Figure 1-38 after trace balancing. Balancing is a time-invariant scaling of amplitudes to a common rms level for all traces.

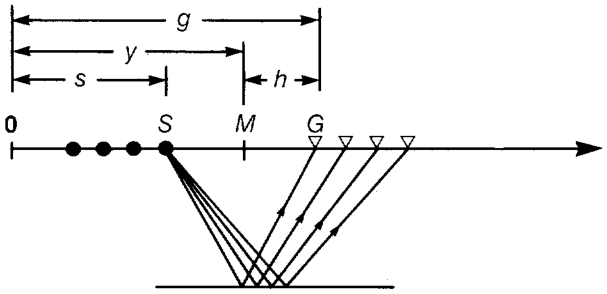


FIG. 1-40. Seismic data acquisition is done in shot-receiver (s,g) coordinates. The raypaths shown are associated with a planar horizontal reflector from a shotpoint S to several receiver locations G. The processing coordinates, midpoint-(half) offset, (y,h) are defined in terms of (s,g): $y = (g + s)/2$, $h = (g - s)/2$. The shot axis here points opposite the profiling direction, which is to the left.

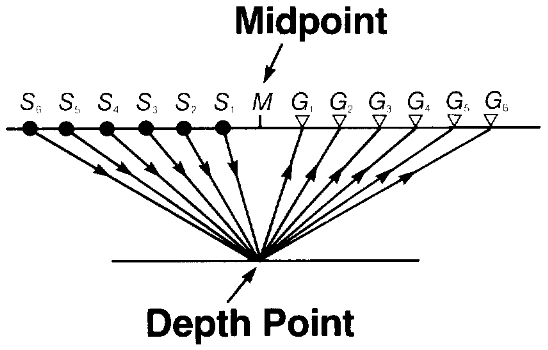


FIG. 1-41. Seismic data processing is done in midpoint-offset (y,h) coordinates. The raypaths shown are associated with a single CMP gather. A CMP gather is identical to a CDP gather if the depth point were on a horizontally flat reflector and if the medium above were horizontally layered.

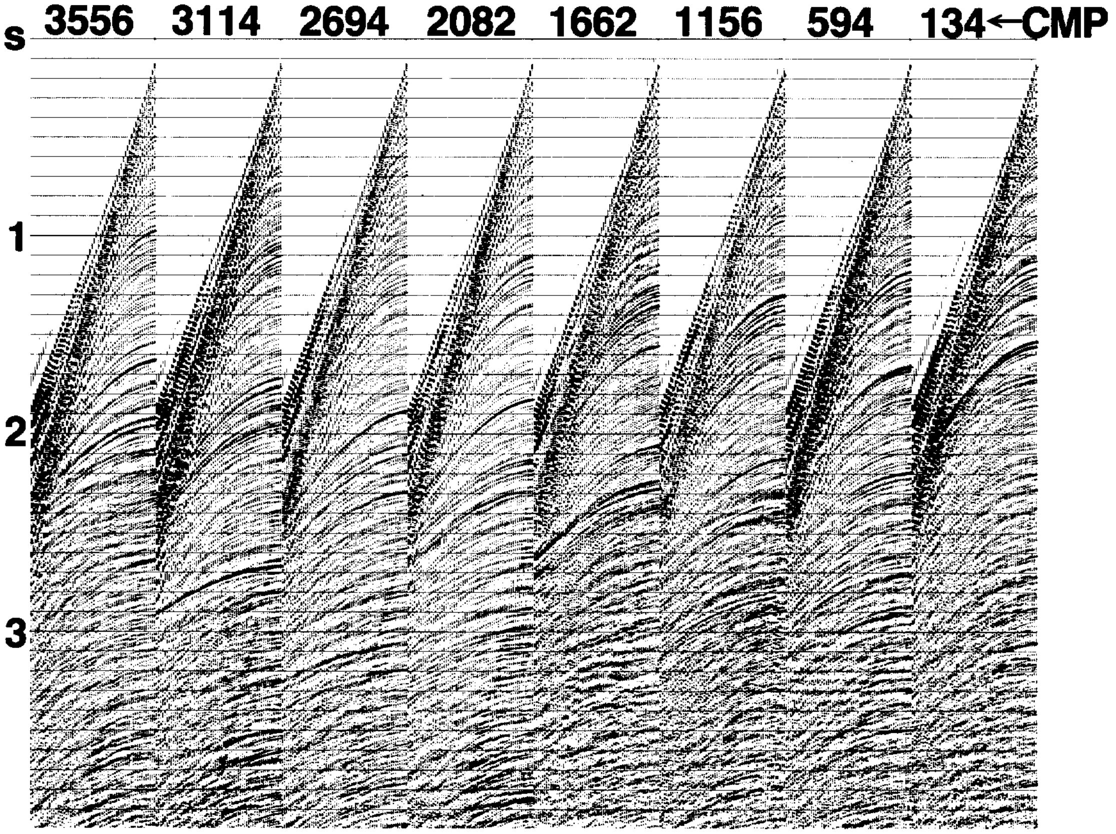


FIG. 1-42. Selected CMP gathers corresponding to the same data as in Figure 1-39.

following gather types are identified in Figure 1-43:

1. Common-Shot Gather (shot record, field record)
2. Common-Receiver Gather
3. Common-Midpoint Gather (CMP gather, CDP gather)
4. Common-Offset Section (constant-offset section)
5. CMP Stacked Section (zero-offset section).

The recording cable length is *FG* and the line length is *AD*. The number of dots along the CMP gathers (cross-section 3) is equal to the CMP fold. The fold tapers off at the ends of the profile (segments *AB* and *CD*). Full-fold coverage along the line is at midpoints over segment *BC*. The diagram in Figure 1-43 is known as a *stacking chart* and is useful when setting up the geometry of a line for preprocessing. If there is a missing shot or a bad receiver, the affected midpoints are identified easily (see Exercise 1.12).

The CMP recording technique, which was patented in the 1950s and published later (Mayne, 1962), uses redundant recording to improve the S/N ratio. To achieve redundancy, multiple sources per trace n_s, multiple receivers per trace n_g, and multiple offset coverage of the same subsurface point n_f, are used in the field. Given the total number of elements in the recording system, $N = n_s \cdot n_g \cdot n_f \cdot$, the signal amplitude-to-rms noise ratio theoretically is improved by a factor of $\sqrt{N}$. This improvement factor is based on the assumptions that the reflection signal on traces of a CMP gather is identical and the random noise is mutually uncorrelated from trace to trace (Sengbush, 1983). Because these assumptions do not strictly hold in practice, the S/N ratio improvement gained by stacking is somewhat less than $\sqrt{N}$. Common-midpoint stacking also attenuates coherent noise such as multiples (Mayne, 1962), guided waves, and ground roll. This is because reflected signal and coherent noise usually have different stacking velocities.

1.4.4 Velocity Analysis

In addition to providing an improved S/N ratio, multifold coverage with nonzero-offset recording yields velocity information about the subsurface (Chapter 3). Velocity analysis is performed on selected CMP gathers or groups of gathers. The output from one type of velocity analysis is a table of numbers as a function of velocity versus two-way zero-offset time (velocity spectrum). These numbers represent some measure of signal coherency along the hyperbolic trajectories governed by velocity, offset, and traveltime. Figure 1-44 shows the velocity spectra at the CMP locations indicated in Figure 1-42. Velocity-time pairs are selected from these spectra based on maximum coherency peaks. These velocity functions then are spatially interpolated between the analysis points across the entire profile as shown in Figure 1-45, to supply a velocity function for each CMP gather along the profile.

In areas with complex structure, velocity spectra often fail to provide sufficient accuracy in velocity picks. When this is the case, the data are stacked with a range of constant velocities, and the constant-velocity stacks themselves are used in picking velocities (Section 3.3).

1.4.5 NMO Correction and Stacking

The velocity field (Figure 1-45) is used in normal moveout (NMO) correction of CMP gathers (Section 3.2). Figure 1-46 shows the CMP gathers in Figure 1-42 after moveout correction. Note that events are virtually flattened across the offset range; i.e., the offset effect has been removed from traveltimes. However, as a result of moveout correction, traces are stretched in a time-varying manner, causing their frequency content to shift toward the low end of the spectrum. Frequency distortion increases at shallow times and large offsets. To prevent the degradation of especially shallow events, the distorted zone is deleted (muted) before stacking (Figure 1-47). Finally, a CMP stack is obtained (Figure 1-48) by summing over the offset. The stack is the frontal face of the data prism shown in Figure 1-34.

1.4.6 Residual Statics Corrections

One extra step is needed before stacking for most land and some shallow-water data. From the NMO-corrected gathers in Figure 1-49a, note that the events in CMP 216 are not as flat as they are in the other gathers. The moveout in CMP gathers does not always conform to a perfect hyperbolic trajectory. This often is because of near-surface velocity irregularities that cause a static or dynamic distortion problem. Lateral velocity variations due to complex overburden can cause moveouts that could be negative; i.e., a reflection event arrives on long-offset traces before it arrives on short-offset traces. Close examination of the velocity spectra indicates that some are easier to pick (Figure 1-50a) than others (Figure 1-51a). The velocity spectrum that corresponds to CMP 297 has sharp coherency peaks that are associated with a distinctive velocity trend. However, the velocity spectrum that corresponds to CMP 188 is noisy and difficult to interpret.

To improve stacking quality, residual statics corrections (Section 3.4) are performed on the moveout-corrected CMP gathers. This is done in a surface-consistent manner; that is, time shifts are dependent only on shot and receiver locations, not on the raypaths from shots to receivers. The estimated residual corrections are applied to the original CMP gathers with no NMO correction.

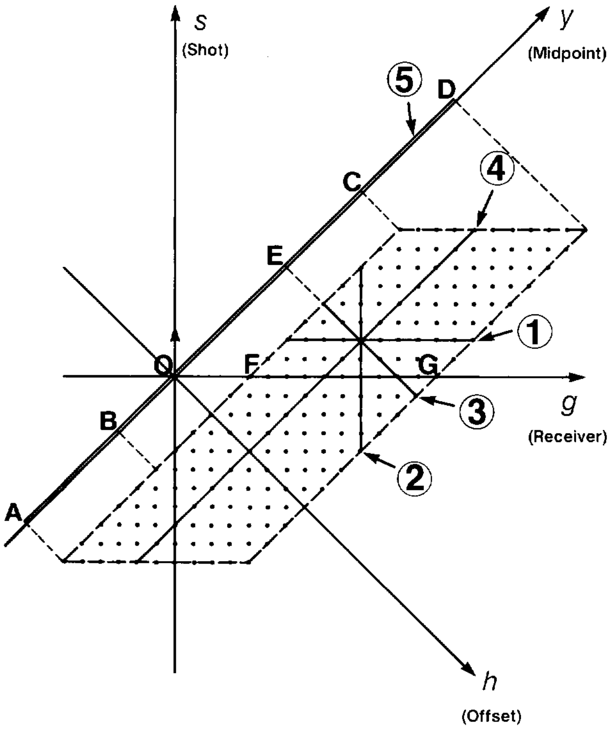


FIG. 1-43. A hypothetical stacking chart (modified from Claerbout, 1976). Each dot represents a single trace with the time axis perpendicular to the plane of the page. Shot-geophone (s,g), and midpoint-offset (y,h) coordinates are superimposed with the (y,h) plane rotated 45 degrees with respect to the (s,g) plane. Here, (1) is a common-shot gather, (2) is a common-receiver gather, (3) is a CMP gather, (4) is a common-offset section, and (5) is a CMP stacked section. The remaining notation is defined in the text.

Velocity analyses then are often repeated to improve the velocity picks (Figures 1-50b and 1-51b). With this improved velocity field, the CMP gathers are NMO-corrected (Figure 1-49b). Finally, the gathers are stacked as shown in Figure 1-52b. For comparison, the stack without the residual statics corrections is shown in Figure 1-52a. Reflection continuity over the problem zone (between midpoints 53-245) has been improved.

1.4.7 Poststack Processing

Predictive deconvolution (Sections 2.6 and 2.7) is sometimes effective in suppressing reverberations or short-period multiples and in further whitening the spectrum. Time-variant band-pass filtering (Section 1.2.4) is used to suppress noisy frequency bands. Finally, some type of gain (Section 1.5) is applied to bring up weak reflections (compare Figure 1-48 with Figure 1-53). For true amplitude preservation, time-variant scaling of stack amplitudes is avoided; instead, a relative amplitude compensa-

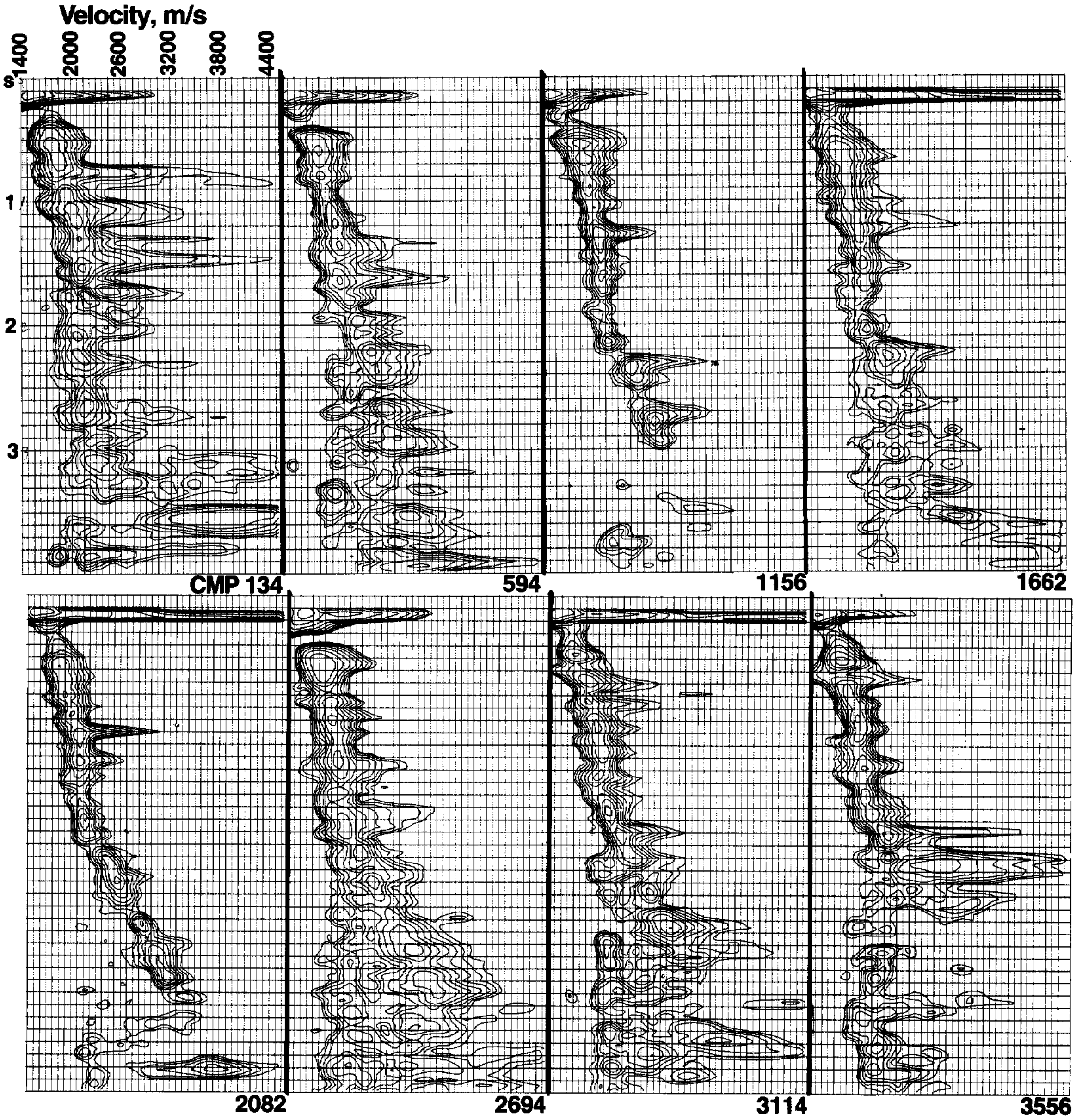


FIG. 1-44. Velocity spectra derived from the CMP gathers in Figure 1-42. Note the general trend common to all velocity functions and the progressive loss of velocity resolution in the trends at late times.

tion function that is constant from trace to trace is applied (Section 1.5). This is a slowly time-varying gain function that amplifies weak late reflections without destroying the amplitude relationships from trace to trace that may be due to subsurface reflectivity.

1.4.8 Migration

Dipping events then are moved to their true subsurface positions and diffractions are collapsed by migrating the stacked section using the medium velocity (Figure 1-54).

The conventional processing sequence is summarized in Figure 1-55. Each of the processes described above is considered in detail in subsequent chapters.

1.5 GAIN APPLICATIONS

Gain is a time-variant scaling in which the scaling function is based on a desired criterion. Often, gain is applied to seismic data for display. Another gain application is for spherical spreading correction. A field record represents

(Text continued on page 56)

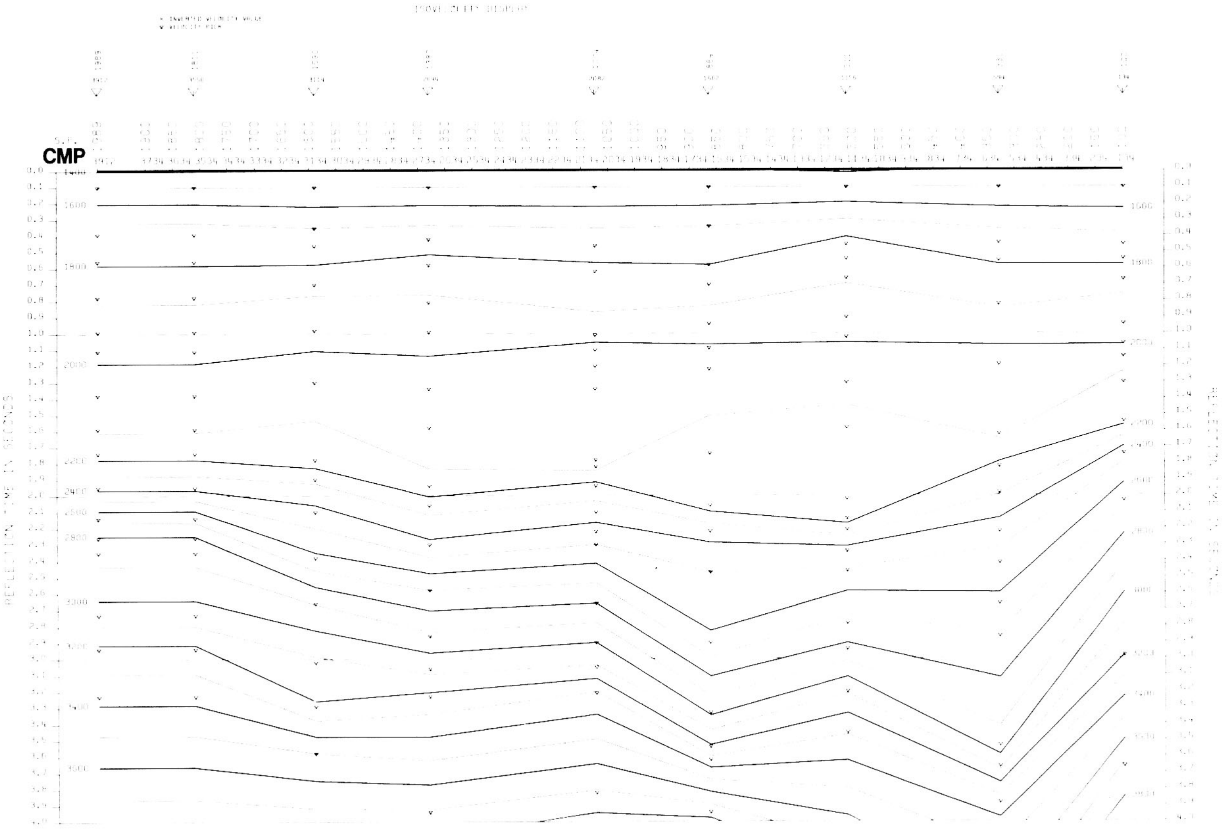


FIG. 1-45. Stacking velocity field over the length of the seismic line under consideration. This isovelocity contour map was derived using the velocity picks from the spectra in Figure 1-44. When compared with Figure 1-48, this figure reflects the structural trend in the subsurface. Triangular symbols on top indicate the locations of the velocity analyses in Figure 1-44.

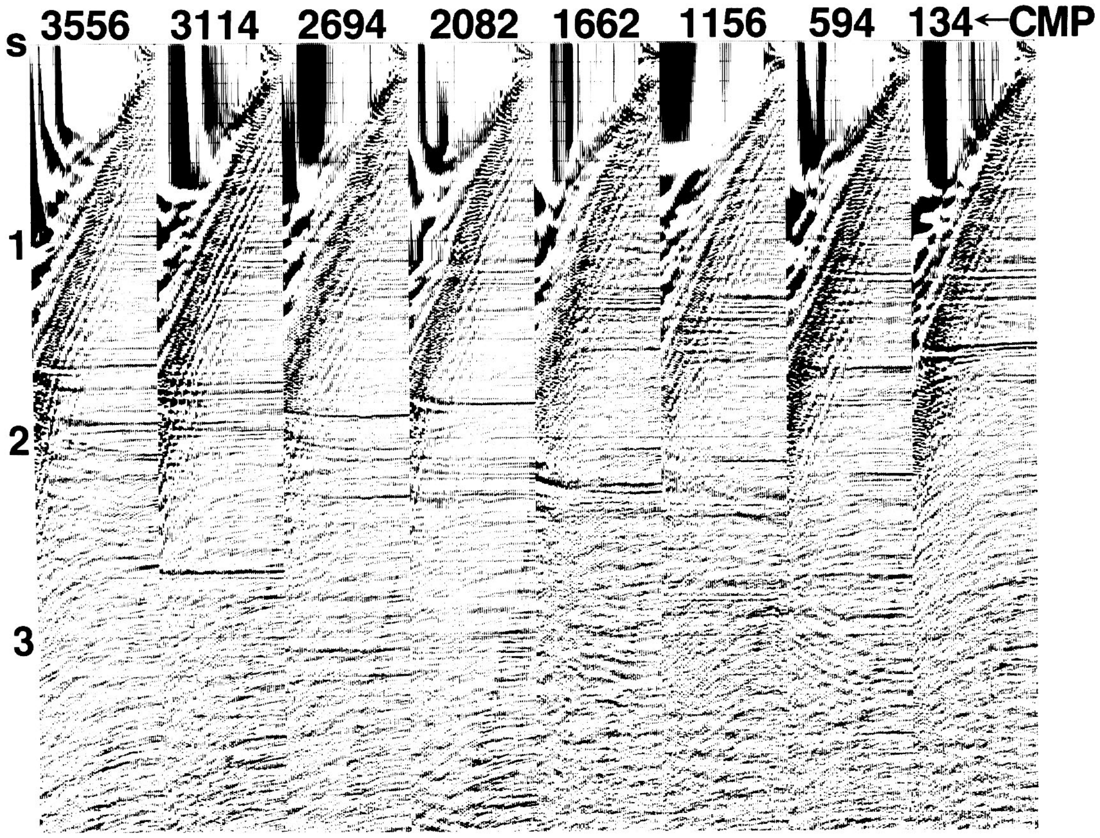


FIG. 1-46. The CMP gathers of Figure 1-42 after NMO correction using the velocity picks derived from the spectra in Figure 1-44. Note the stretching effect at the shallow part of the gathers, particularly at the far offsets. By properly selecting the velocities, the primaries are flattened. (Compare the individual events in this figure and in Figure 1-42.)

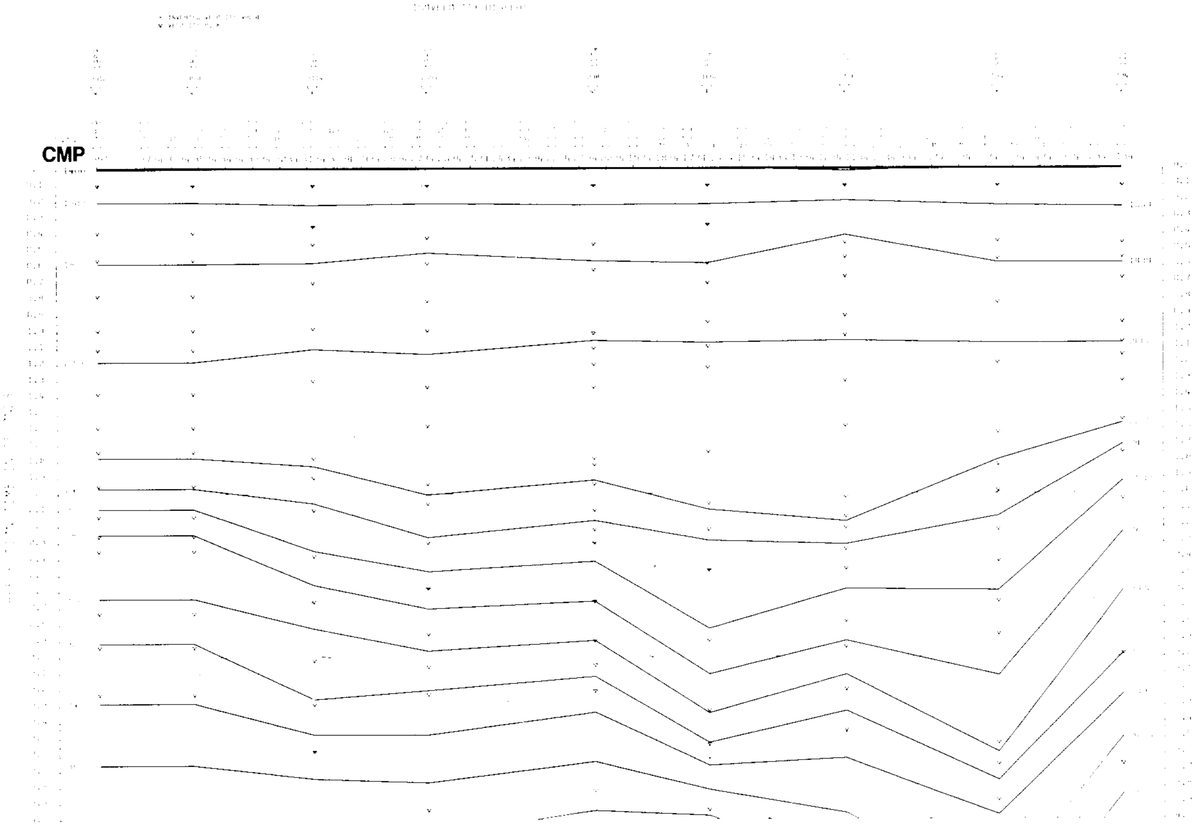


FIG. 1-45. Stacking velocity field over the length of the seismic line under consideration. This isovelocity contour map was derived using the velocity picks from the spectra in Figure 1-44. When compared with Figure 1-48, this figure reflects the structural trend in the subsurface. Triangular symbols on top indicate the locations of the velocity analyses in Figure 1-44.

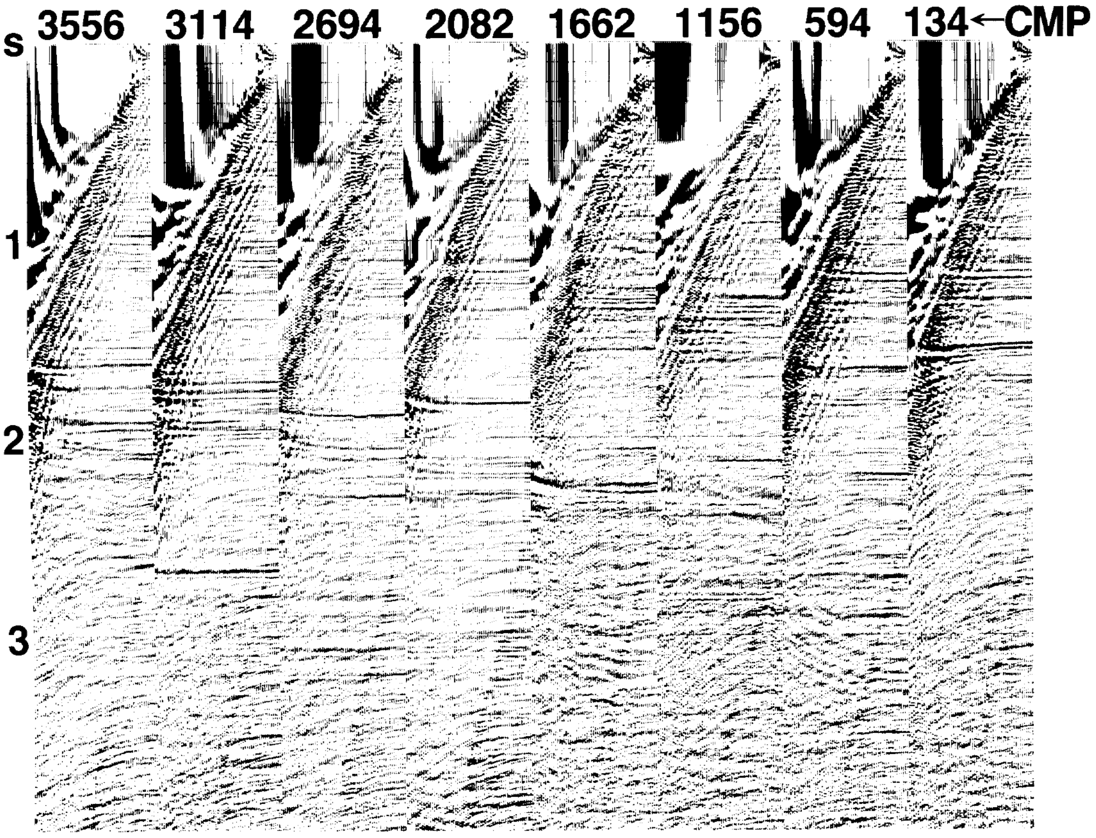


FIG. 1-46. The CMP gathers of Figure 1-42 after NMO correction using the velocity picks derived from the spectra in Figure 1-44. Note the stretching effect at the shallow part of the gathers, particularly at the far offsets. By properly selecting the velocities, the primaries are flattened. (Compare the individual events in this figure and in Figure 1-42.)

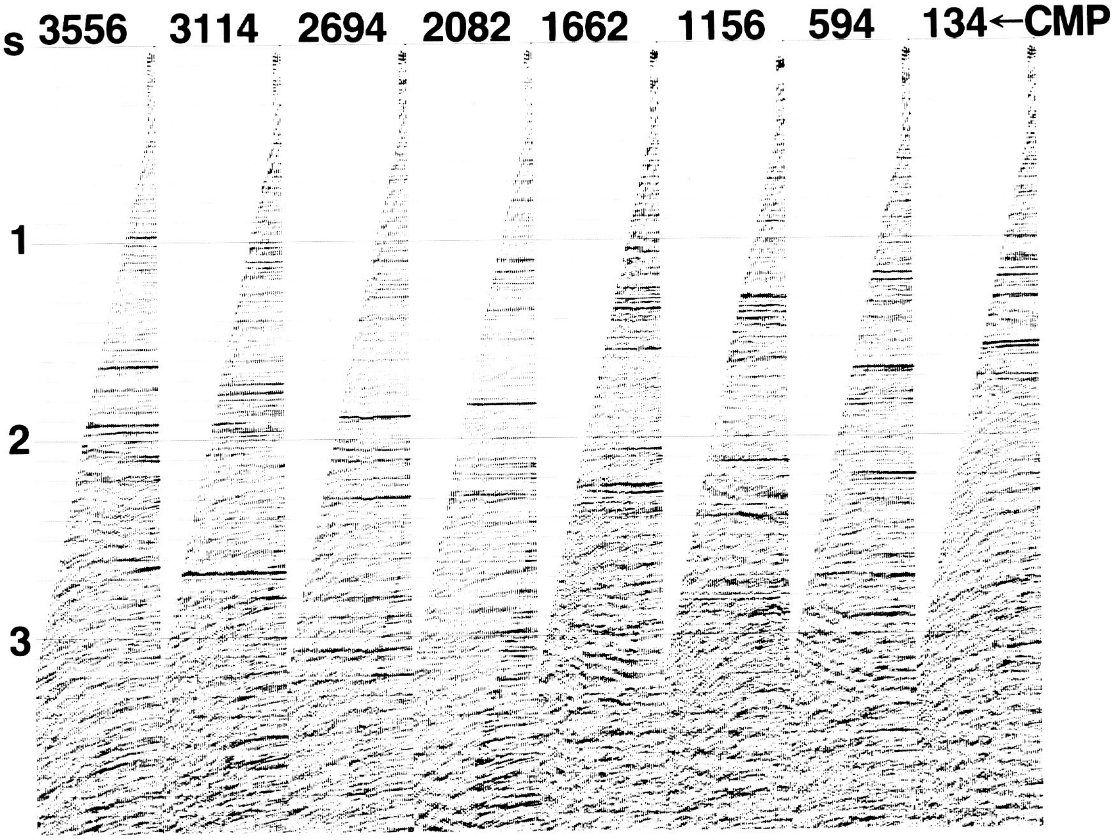


FIG. 1-47. The CMP gathers of Figure 1-46 after muting the stretched zones. By muting, the degrading effect of the stretched signal (very low frequency) on stacking quality is eliminated.

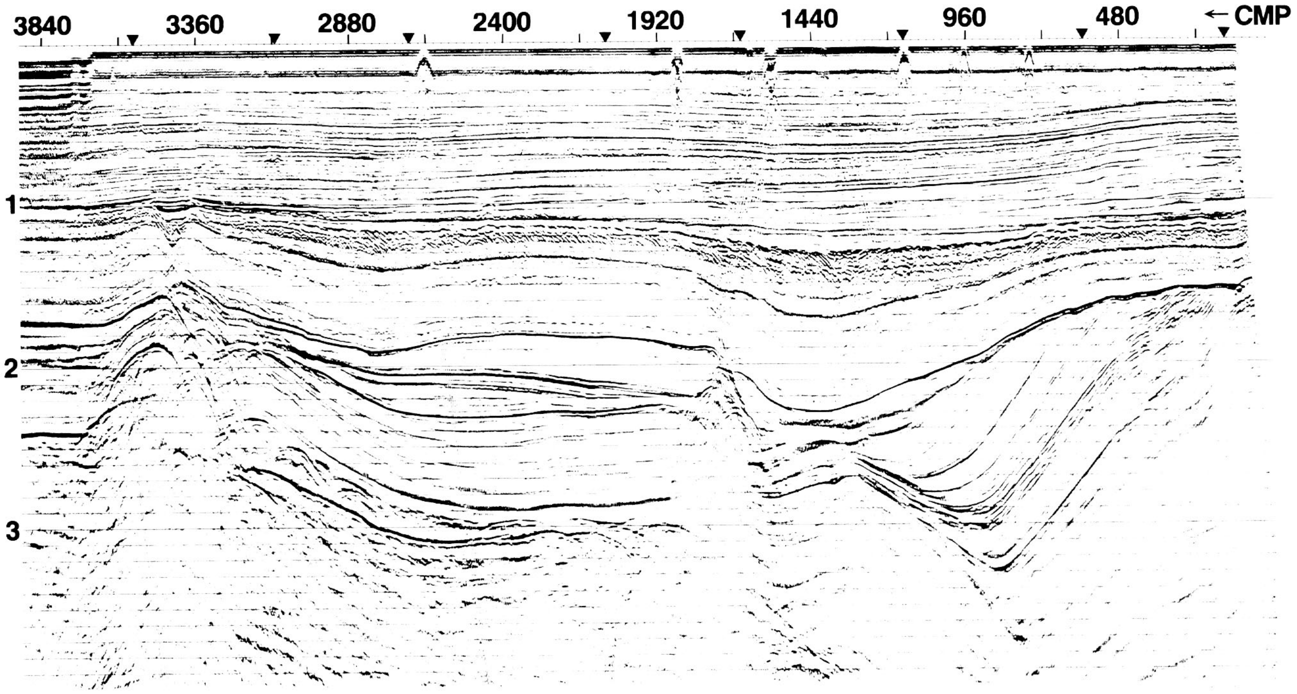


FIG. 1-48. The CMP stack associated with the gathers in Figure 1-47. The triangles refer to the locations of the velocity analyses in Figure 1-44.

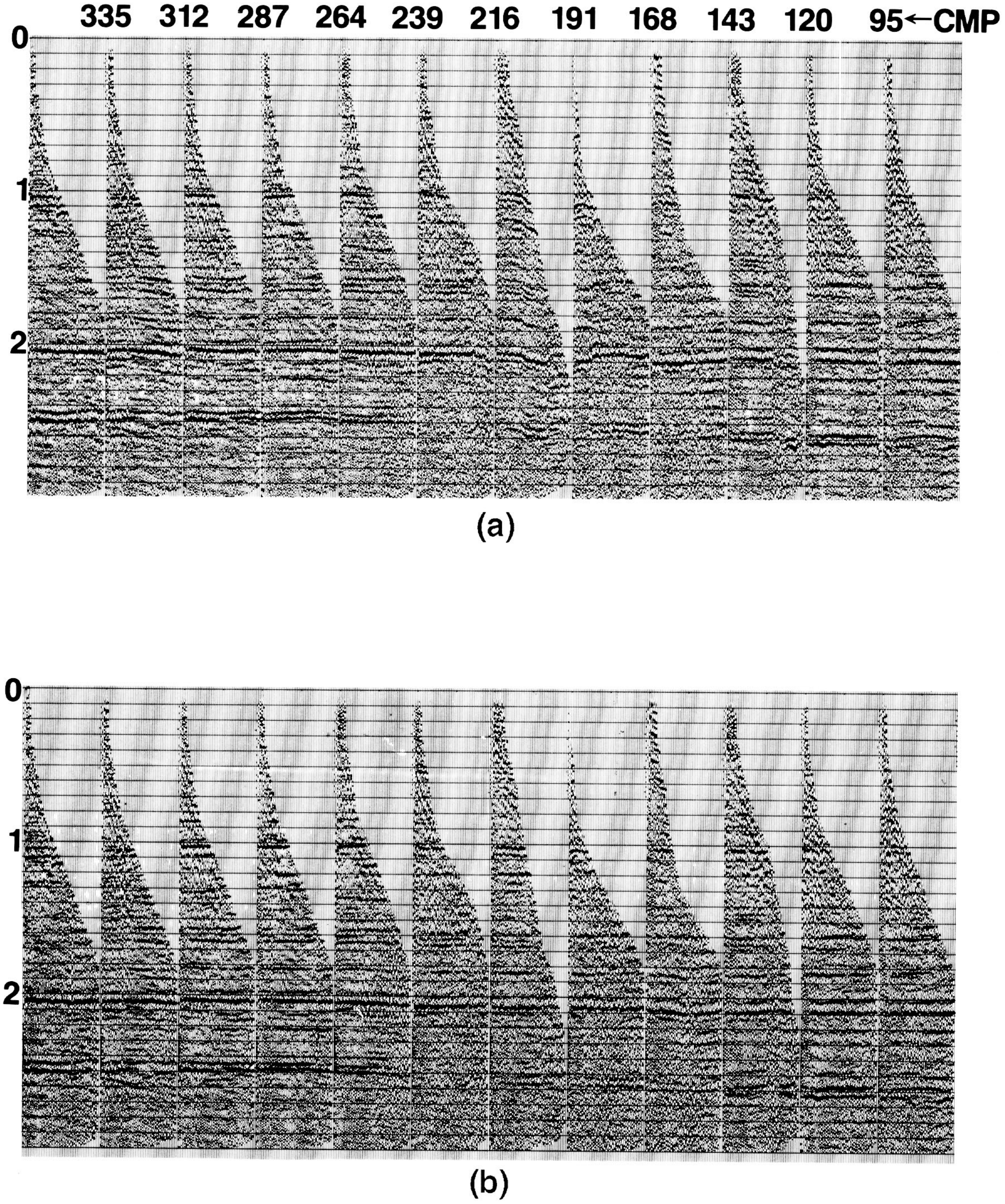


FIG. 1-49. The NMO-corrected CMP gathers from a land seismic line (a) before and (b) after residual statics corrections. Note that the distorted events (CMP 191, 216) have nearly been flattened (b).

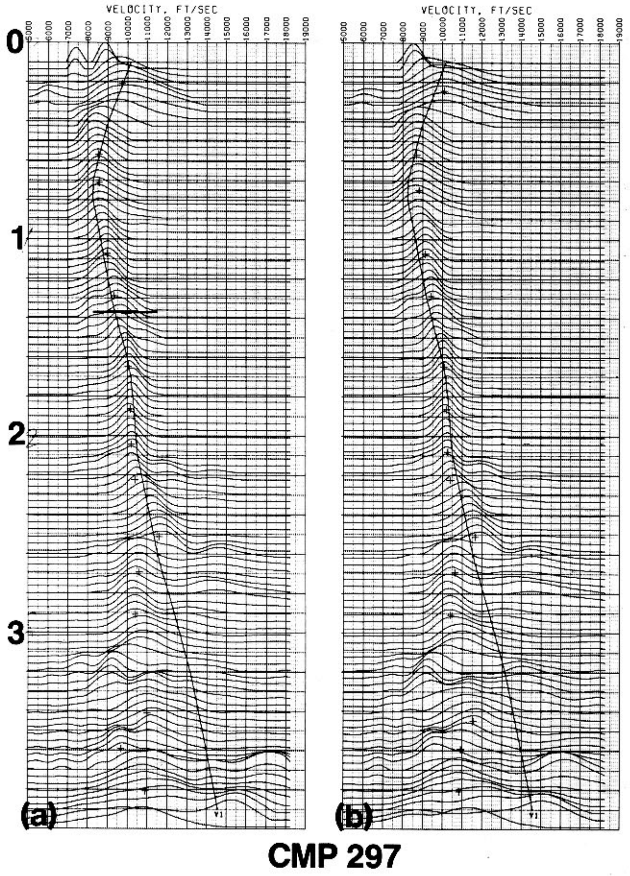


FIG. 1-50. Velocity spectra derived from the same data as in Figure 1-49, (a) before and (b) after residual statics corrections. Note that no significant difference exists between spectra derived from CMP gather 297 with and without the application of residual statics corrections. Reflection times in this gather did not have significant residual static shifts.

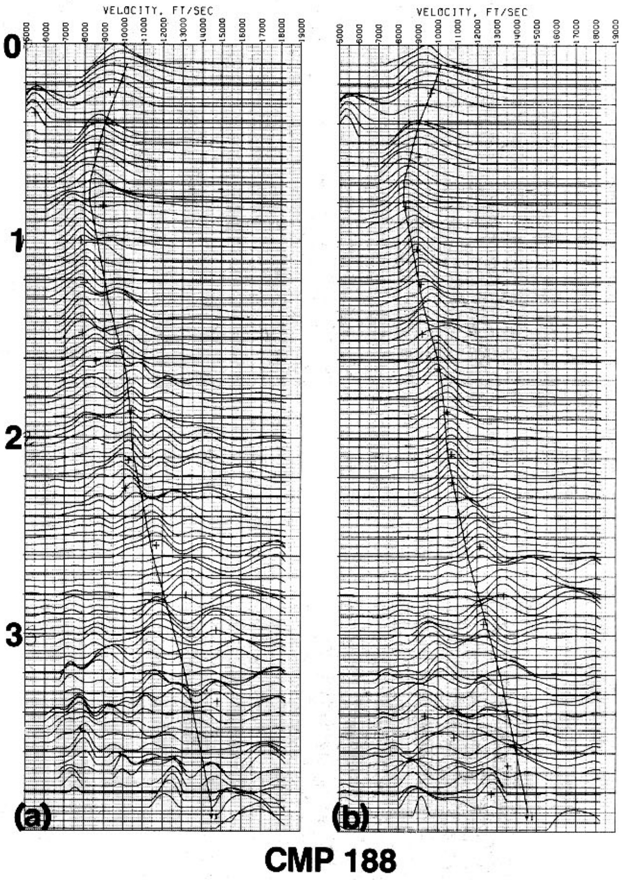
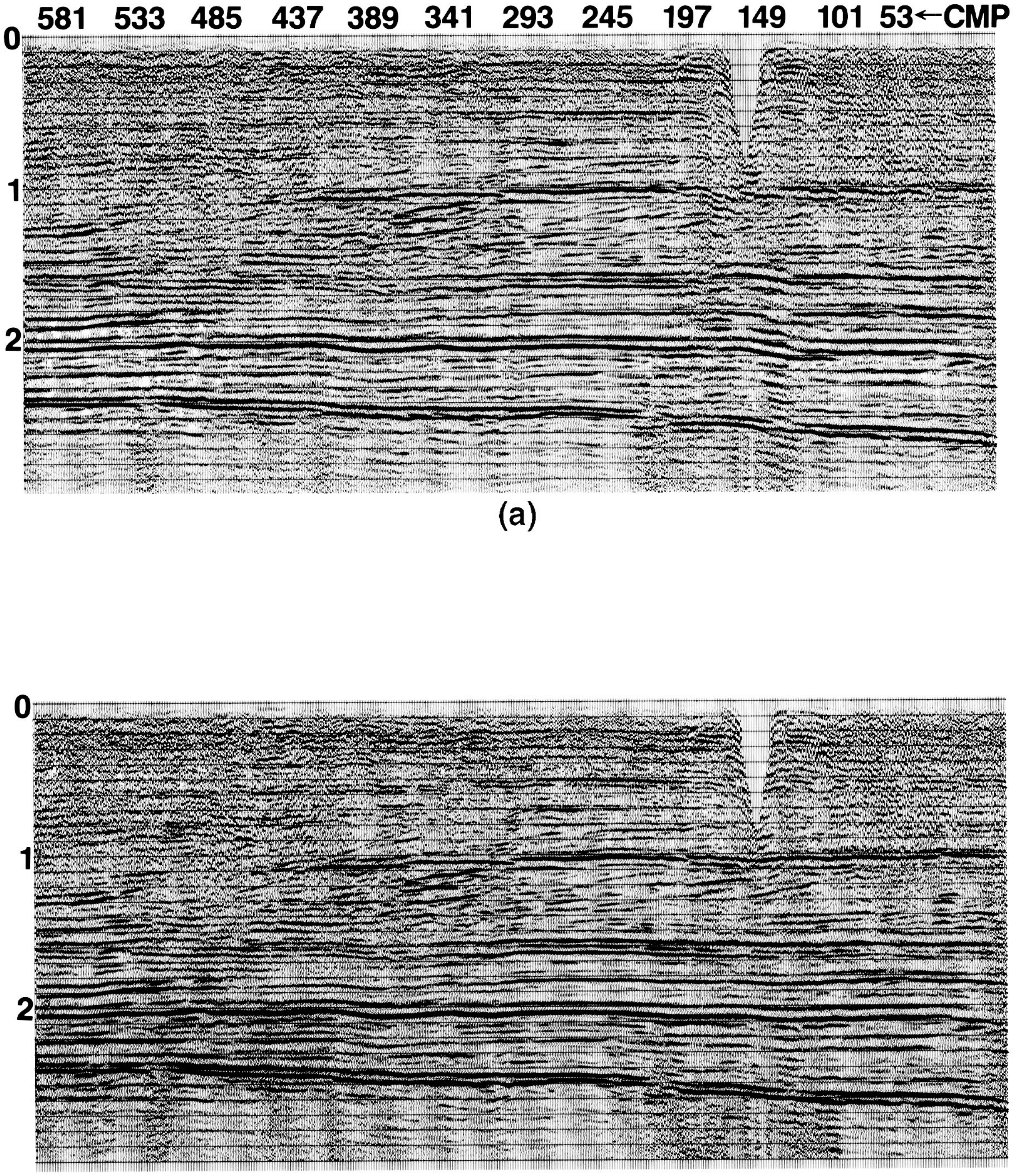


FIG. 1-51. Velocity spectra derived from the same data as in Figure 1-49, (a) before and (b) after residual statics corrections. Note the improvement after correcting down to 2.6 s.



(a)

(b)

FIG. 1-52. The CMP stacks derived from the gathers in Figure 1-49. The stack (a) without residual statics corrections shows false structure and poor coherence in the vicinity of CMP 149-197. Both are eliminated by correcting for residual statics (b).

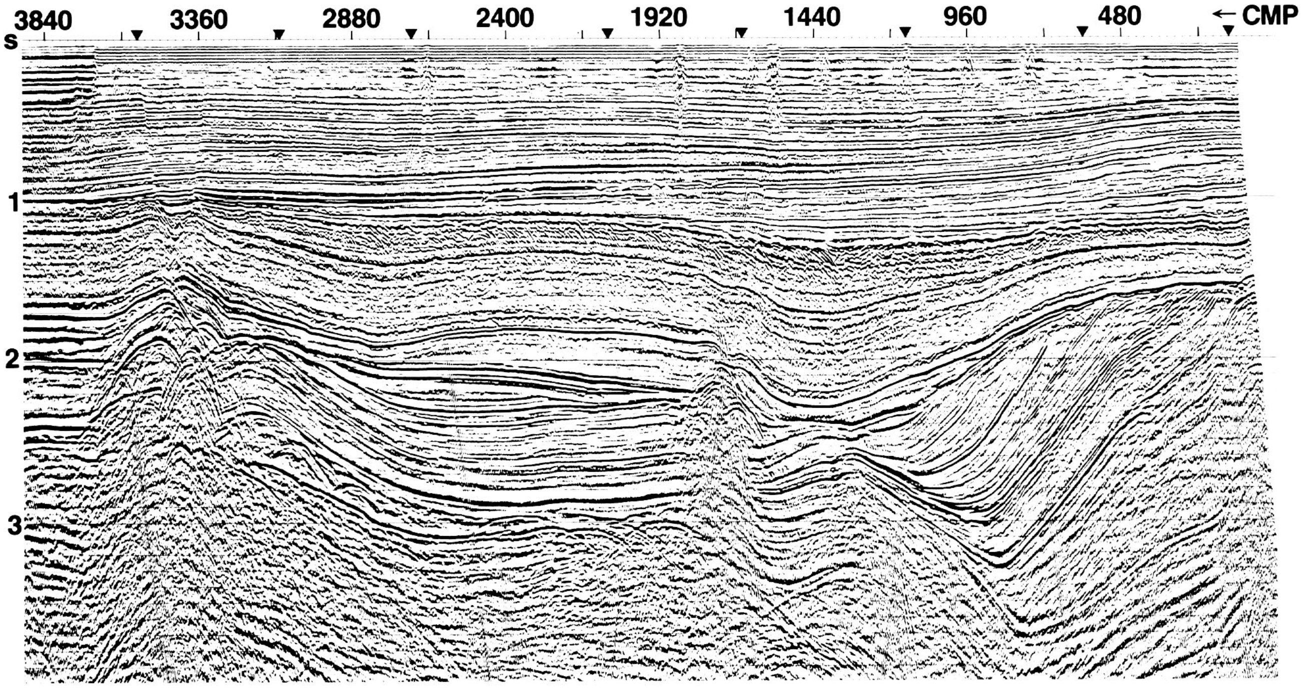


FIG. 1-53. The CMP stack associated with the gathers in Figure 1-47. This stack is the same as in Figure 1-48 with a time-variant filter (TVF) followed by an rms gain application.

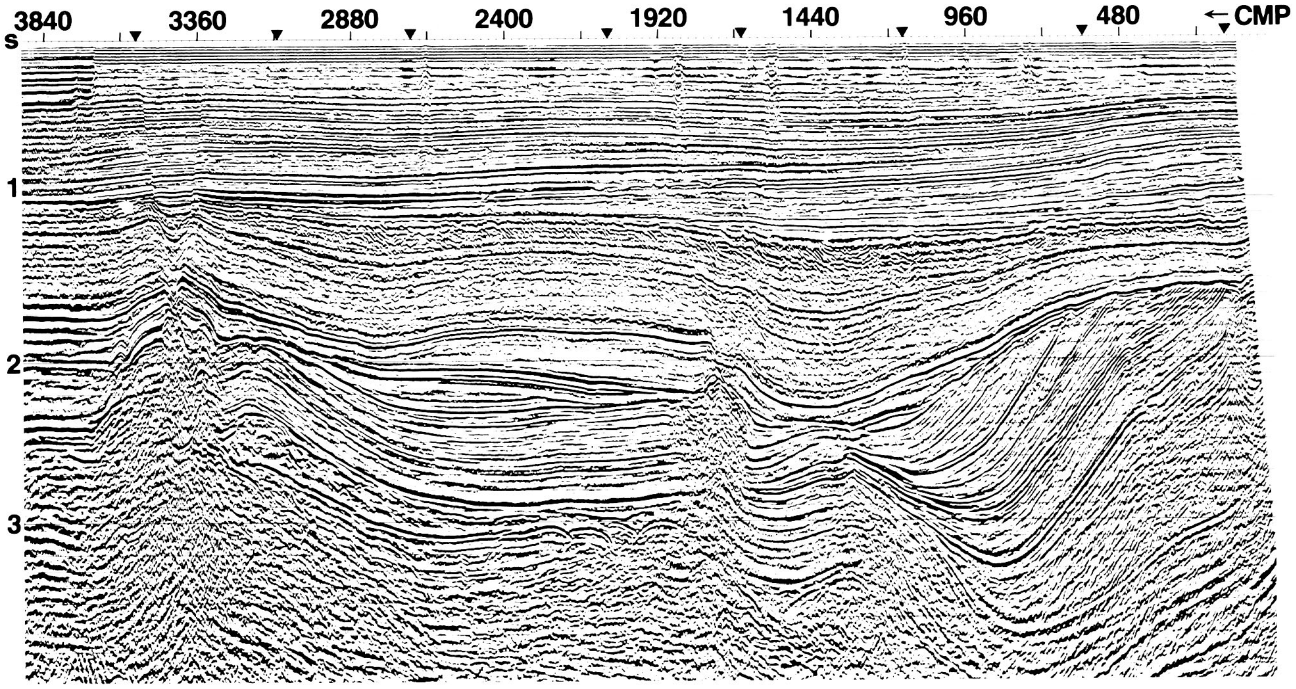


FIG. 1-54. Migrated CMP stack. The input to migration is the filtered stacked section. The migration output was gained as in Figure 1-53.

a wave field that is generated by a single shot. Conceptually, a single shot is thought of as a point source that generates a spherical wave field. The earth has two effects on a propagating wave field. In a homogeneous medium, energy density decays proportionately to $(1/r^2)$, where r is the radius of the wavefront. Wave amplitude is proportional to the square root of energy density; it decays as $(1/r)$. In practice, velocity usually increases with depth, which causes further divergence of the wavefront and a more rapid decay in amplitudes with distance. Second, the frequency content of the initial source signal changes in a time-variant manner as it propagates. In particular, high frequencies are more rapidly absorbed than low frequencies. This is because of the intrinsic attenuation in rocks.

Attenuation mechanisms still are the subject of extensive research. However, one plausible mechanism for attenuation is related to pore fluids. As the wavefront passes through rocks, the fluids that are present in the pores are disturbed. This disturbance is greater in partially saturated rocks than fully saturated rocks. Pore fluids consume part of the energy of the propagating wave field, which causes a frequency-dependent decay.

From Figure 1-56, note the wavefront divergence and frequency absorption on the field record. The first panel represents field data without any gain recovery function applied. Note the gradual decay in amplitude at later times. This record was filtered with a series of 10-Hz-wide band-pass filters. The signal of the 10 to 20-Hz panel exists down to about 6 s. On the 20 to 30-Hz panel, however, signal is visible only down to about 4 s. Moving to the higher frequency panels, note that the signal level mainly is confined to increasingly shallower times. Now apply the geometric spreading correction to the original field record in the leftmost panel of Figure 1-56. The result is shown in the leftmost panel of Figure 1-57. The amplitude level has been restored at late traveltimes. Filter panels of this record also are shown in Figure 1-57. When the filter panels in Figures 1-56 and 1-57 are compared with the same passband, we see that the geometric spreading correction brought up some of the signal level at late times. However, note that the geometric spreading correction did not restore the amplitudes of the high frequencies as much as it restored the low frequencies, since the high frequencies were subject to stronger attenuation. The effect of attenuation now must be removed by modifying the amplitude spectrum of the signal, thereby making it broader. Deconvolution is one process that is used to achieve this goal. Time-variant spectral whitening is another. Both processes are described in Chapter 2.

As discussed earlier in this section, wave amplitudes decay as $(1/r)$, where r is the radius of the spherical wavefront. This is true for a homogeneous medium without attenuation. For a layered earth, amplitude decay can be described approximately by $1/[v^2(t) \cdot t]$ (Newman, 1973). Here, t is the two-way traveltime and $v(t)$ is the rms velocity (Section 3.2) of the primary reflections (those reflected only once) averaged over a survey area. Therefore, the gain function for geometric spreading compensation is defined by $g(t) = [v(t)/v(0)]^2 [t/t(0)]$, where $v(0)$ is the velocity value at specified time t(0). A more rigorous offset-dependent and time-dependent description of the geometric spreading correction function also can be used.

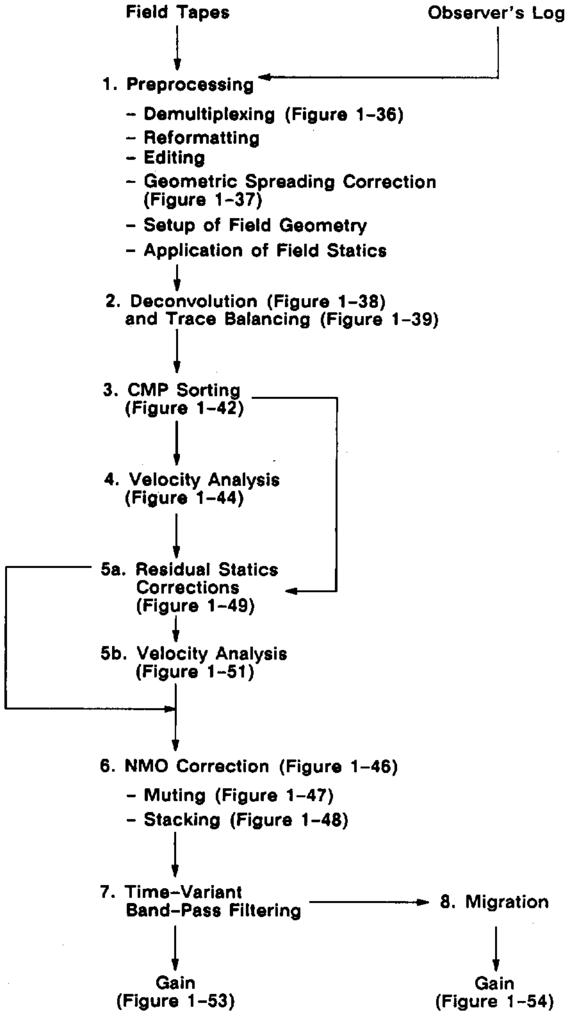


FIG. 1-55. A conventional processing flowchart.

Signal-level decay is evident in the field records in Figure 1-58. Note the weak appearance of reflections, particularly below 1 s. This does not mean that there are no strong reflections below this time. Because of the amplitude decay resulting from wavefront divergence, no signal is seen at late times. As stated previously, this earth effect must be removed to bring up any signal that may be present in later parts of the record.

The same shot records after geometric spreading correction are shown in Figure 1-59. While reflections have been brought up in strength, noise components in the data also have been boosted. This is one undesirable aspect of any type of gain application. Besides ambient noise, coherent noise in the data may be boosted as shown in Figure 1-60. By using the primary velocity function in correcting for geometric spreading, the amplitudes of the dispersive coherent noise and multiples have

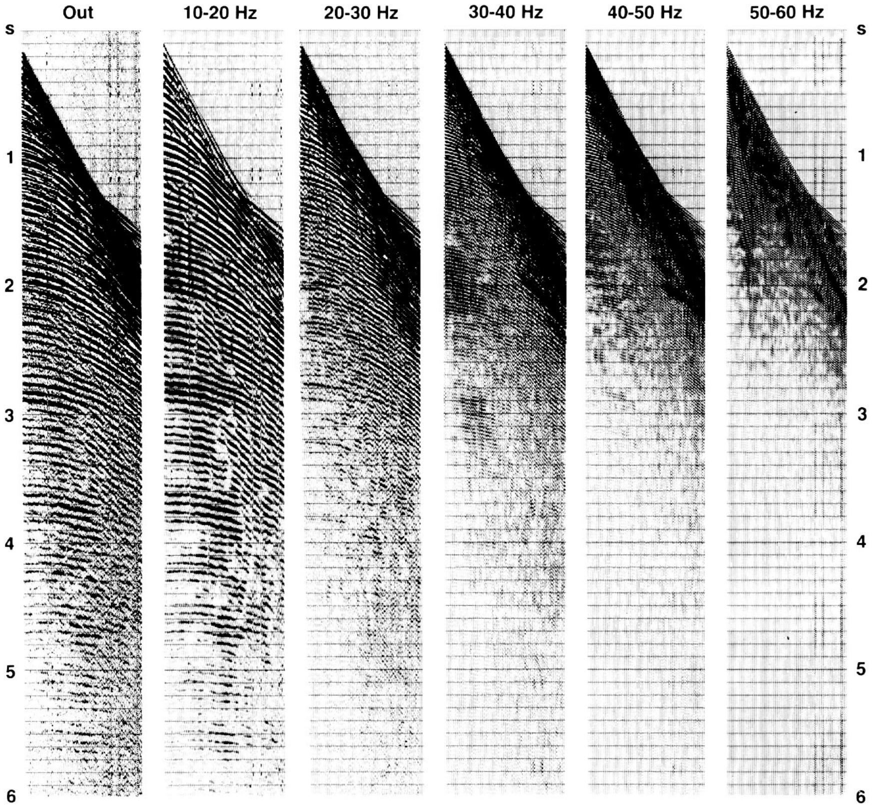


FIG. 1-56. A raw field record with no geometric spreading correction (leftmost panel) and its band-pass filtered versions. Note that the larger reflection amplitudes are confined to shallower times at increasingly higher frequency bands.

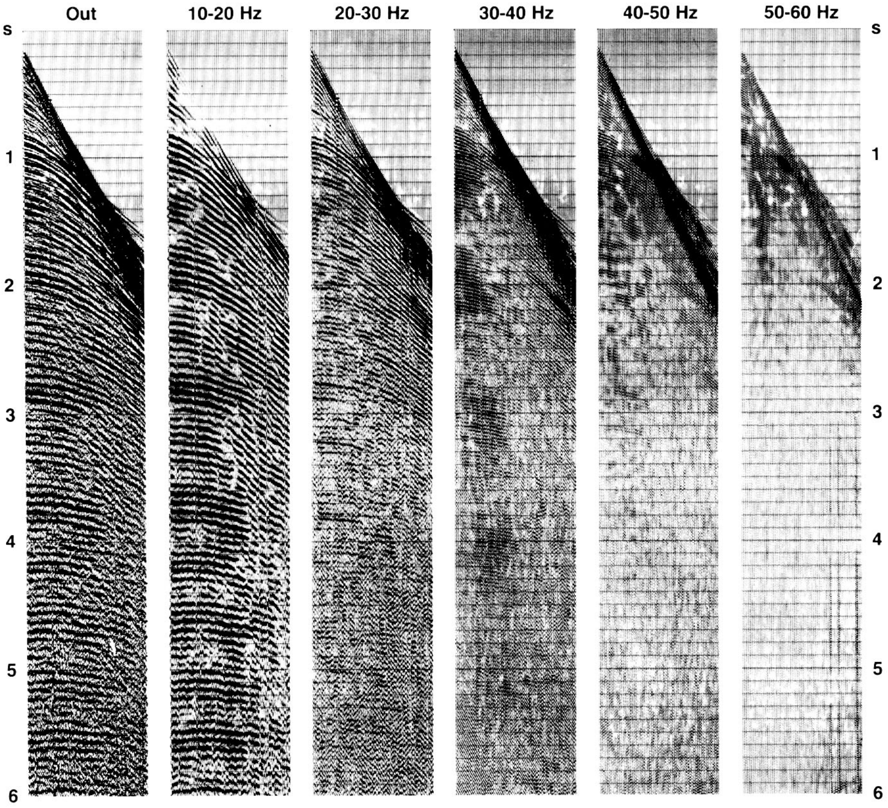


FIG. 1-57. The same field record as in Figure 1-56 (leftmost panel) after correcting for geometric spreading. Amplitudes are restored but frequency absorption remains.

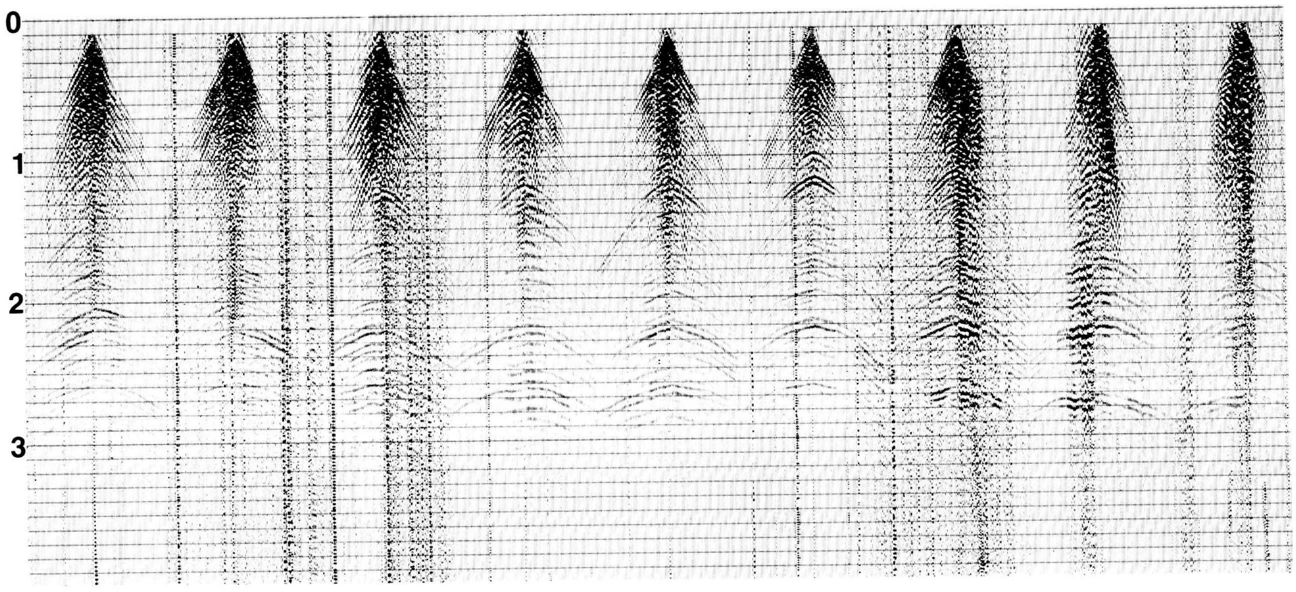


FIG. 1-58. Raw field records from a land survey. Note the rapid decay in amplitudes at late times.

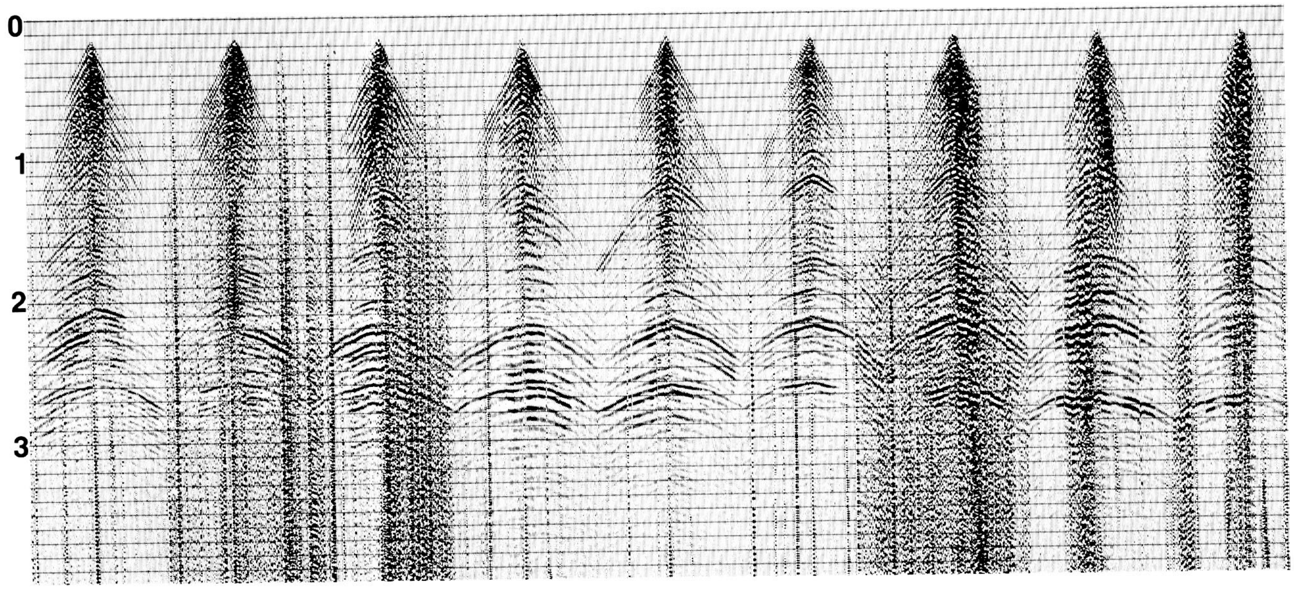


FIG. 1-59. The same field records as in Figure 1-58 after correcting for geometric spreading. The amplitudes have been restored at late times. Unfortunately, ambient noise also has been strengthened.

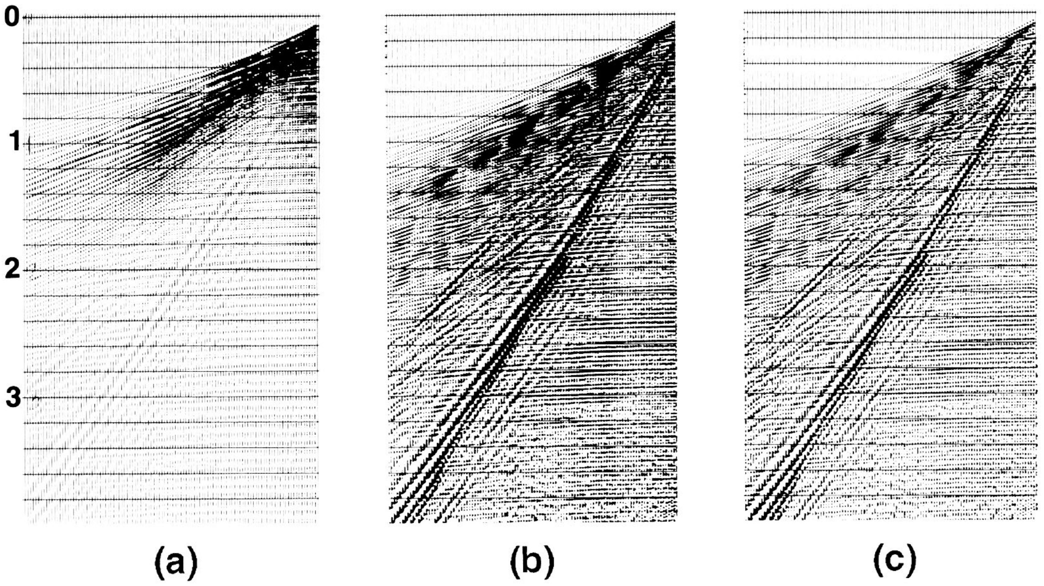


FIG. 1-60. (a) A raw field record from a marine survey. Before correcting for geometric spreading, refraction and guided wave energy dominate the record. (b) After the geometric spreading correction, while reflection amplitudes have been restored, multiples and coherent noise also have been boosted. (c) To a degree, trace balancing can bring down the level of this undesired energy.

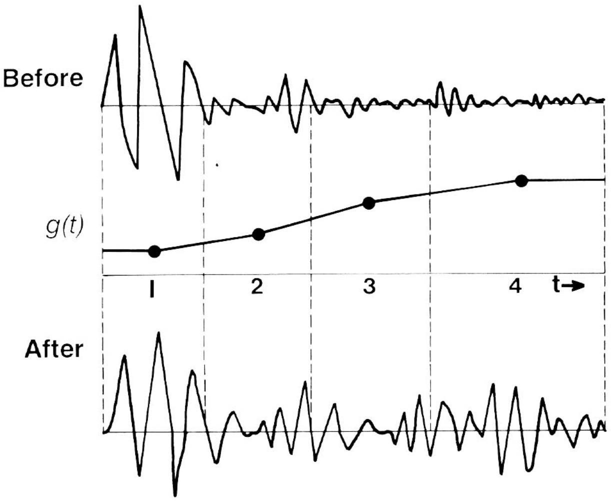


FIG. 1-61. Gain is a time-variant scaling based on a function, *g(t)*. Based on some criteria, this function is defined at the time samples (shown by solid circles) that are usually at the center of specified time gates along the trace as indicated by 1, 2, 3, and 4. Gain application simply involves multiplying *g(t)* by the input trace amplitudes.

been overcorrected. Another example of overcorrected multiples is shown in Figure 1-57. (Compare the leftmost panel with its equivalent in Figure 1-56.)

Various types of gain criteria are used in practice. Based on a desired criterion, a gain function, *g(t)*, is derived from the data and multiplied with trace amplitudes at each time sample. This is illustrated in Figure 1-61. The gain function is specified or estimated at the time samples indicated by the dots and interpolated between these samples. Three common types of gain are described in the following paragraphs.

1.5.1 Programmed Gain Control

Programmed gain control (PGC) is the simplest type of gain. Referring to part of a stacked section in Figure 1-62, a gain function can be defined by interpolating between some scalar values specified at particular time samples. Larger scalar values naturally would be assigned at late times. In panels 1 through 4 in Figure 1-62, the applied PGC factors are indicated by the circled numbers at corresponding time values. When moving from panel 1 to panel 4, the amplitudes are increasingly magnified. Rather than picking the scalars in a qualitative manner, the envelope of the ungained trace can be computed and smoothed. The envelope, which is the curve drawn by smoothly connecting the adjacent peaks (or troughs) along the trace, is a reliable attribute that describes amplitude decay rate. The PGC function then is the inverse of the trace envelope. A single PGC function is

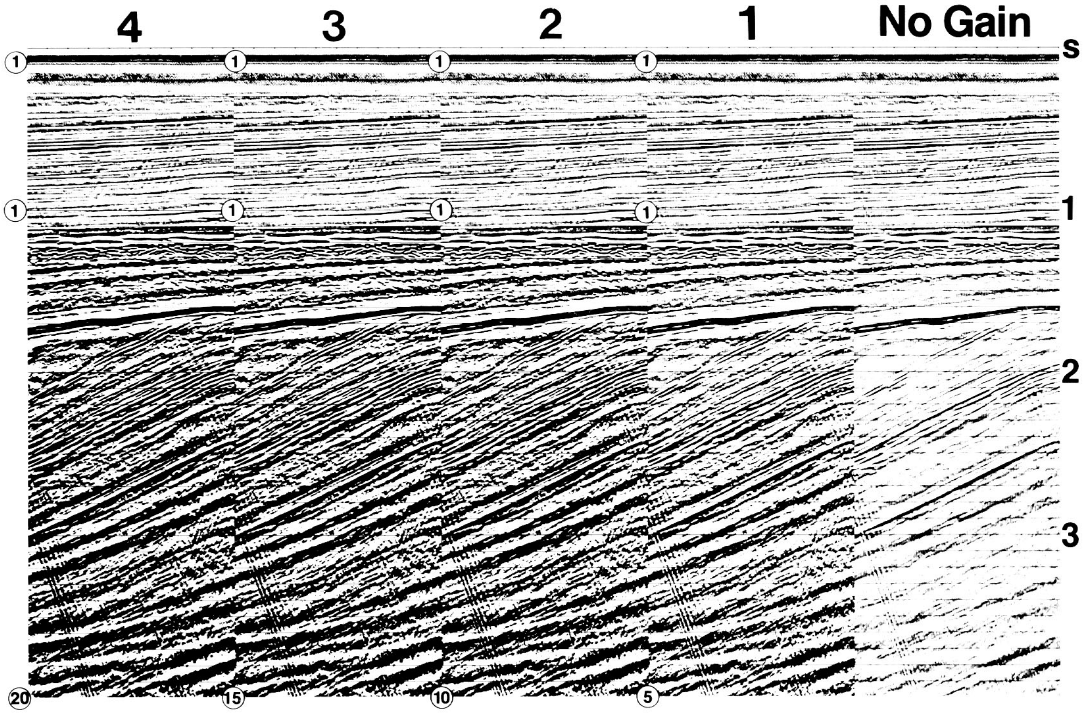


FIG. 1-62. A portion of a CMP stack before and after application of four different PGC functions. The scale factors used in constructing the gain functions are indicated by the circled numbers at the times of application.

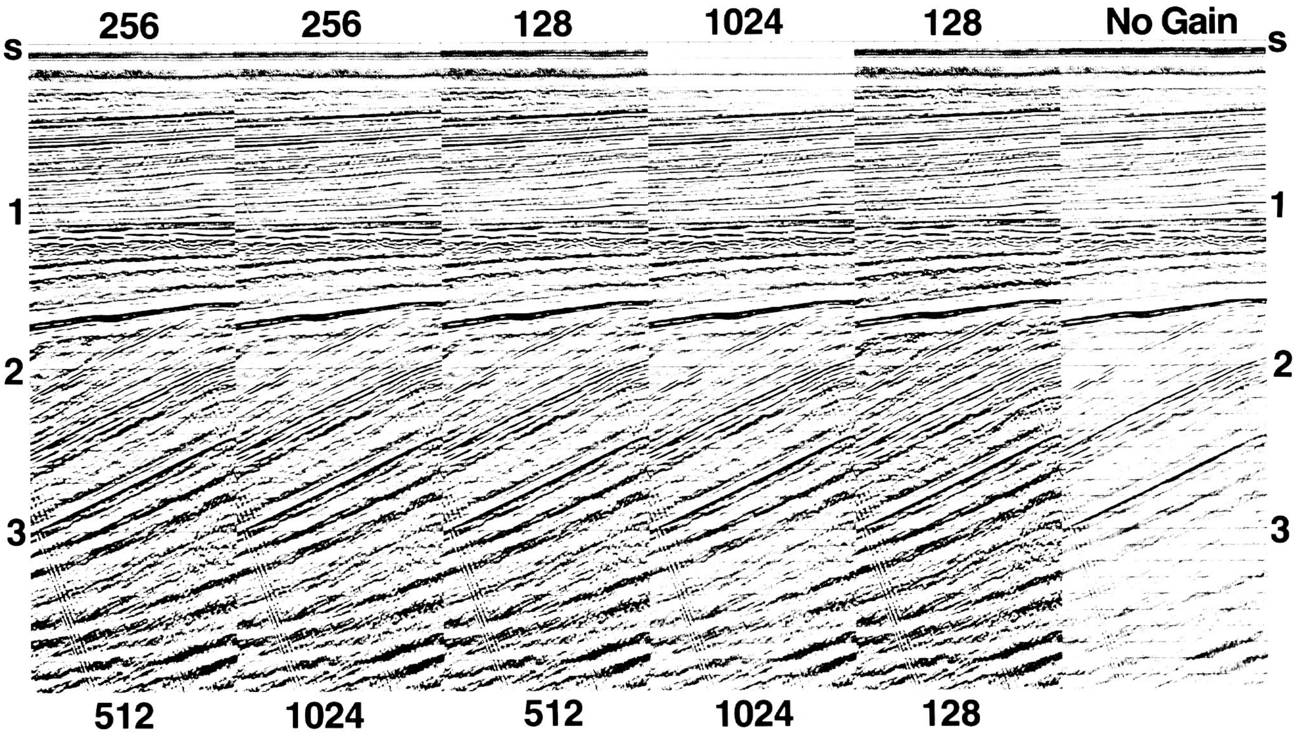


FIG. 1-63. A portion of a CMP stack before and after application of five different rms AGC functions. Numbers on the top and bottom indicate the starting and ending gain window sizes in milliseconds.

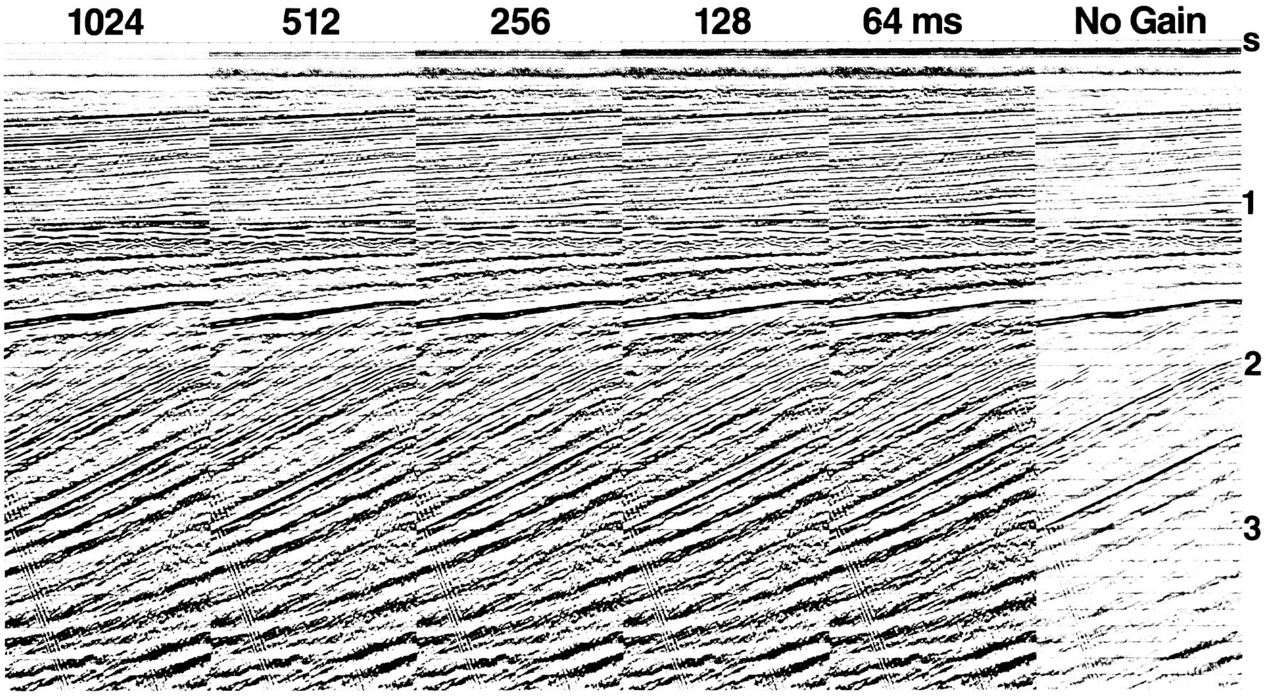


FIG. 1-64. A portion of a CMP stack before and after application of five different instantaneous AGC functions. The numbers on top indicate gain window sizes in milliseconds.

applied to all traces in a gather or stacked section to preserve the relative amplitude variations in the lateral direction.

1.5.2 RMS Amplitude AGC

The rms amplitude AGC gain function is based on rms amplitude within a specified time gate on an input trace. This gain function is computed as follows. The input trace is subdivided into fixed time gates. First, the amplitude of each sample in a gate is squared. Second, the mean of these values is computed and its square root is taken. This is the rms amplitude over that gate. The ratio of a desired rms amplitude (say 2000) to the actual rms value is assigned as the value of the gain function at the center of the gate. Typically, we start out with a certain gate length, say 256 ms, at the shallow part of the trace. Gate length can be kept either constant or it can be increased systematically down the trace. At each gate center, the value of the gain function is computed as described above. Function $g(t)$ then is interpolated between the gate centers. Note that the specified time gates are stationary, i.e., they do not slide down the trace.

Figure 1-63 shows the ungained data and a series of rms-gained sections. The minimum and maximum gate lengths are indicated at the top and bottom of each panel, respectively. When the gate used in the computation is kept small, say 128/128 ms (minimum/maximum gate lengths), then strong reflections become less distinct.

1.5.3 Instantaneous AGC

Instantaneous AGC is one of the most common gain types used. This gain function is computed as follows. First, the mean absolute value of trace amplitudes is computed within a specified time gate. Second, the ratio of the desired rms level to this mean value is assigned as the value of the gain function. Unlike the rms amplitude AGC, this value is assigned to any desired time sample of the gain function within the time gate, say the nth sample of the trace, rather than to the sample at the center of the gate. The next step is to move the time gate one sample down the trace and compute the value of the gain function for the $(n + 1)$th time sample, and so on. No interpolation is therefore needed to define this gain function. Figure 1-64 shows the ungained data and a series of instantaneous AGC-gained sections. Gate lengths are indicated on top of each panel. Very small time gates can cause a significant loss of signal character by boosting zones that contain small amplitudes. This occurs with the 64-ms AGC output. In processing, this is called a fast AGC. In the other extreme, if a large time gate is selected, then the effectiveness of the AGC process is lessened. In practice, 256- to 1024-ms AGC time gates are commonly chosen.

All of the above gain operations modify the trace amplitudes by function $g(t)$ in a time-varying manner. In true amplitude processing, it is necessary to display the data without applying a time-varying data-dependent gain function. However, some amplitude scaling is always necessary for display, since plotters require input data amplitudes to fall in a specific range. Trace balancing (trace equalization) schemes are used for this type of scaling. The balance factor is defined as the ratio of the desired rms to the rms amplitude that is computed from a specified time window. A separate balance factor is computed for and applied to each trace, individually. Alternatively, a single balance factor based on a selected trace within a group of traces can be applied to the entire group. This is called relative trace balancing. Note that trace balancing amounts to scaling the trace by using a single factor that is time-invariant (equivalent to a single-window rms AGC). Figure 1-60 shows rms trace balancing of field data following the geometric spreading correction. Trace balancing commonly is applied immediately after deconvolution (Figure 1-39), and on final stacks using large gates.

In summary, gain is applied to data for various reasons. Geometric spreading correction is applied to compensate for wavefront divergence early in processing, before deconvolution. Also before deconvolution, an exponential gain may be applied to compensate for attenuation losses. The AGC-type gain functions are applied to seismic data to bring up weak signals. These gain functions are time-variant. Third, trace balancing really is not a gain, rather, it is balancing each trace in a group of traces so that they all have the same desired rms amplitude level. Gain must be used with care, since it can destroy signal character. For example, a fast AGC makes strong reflections indistinguishable from weak reflections.

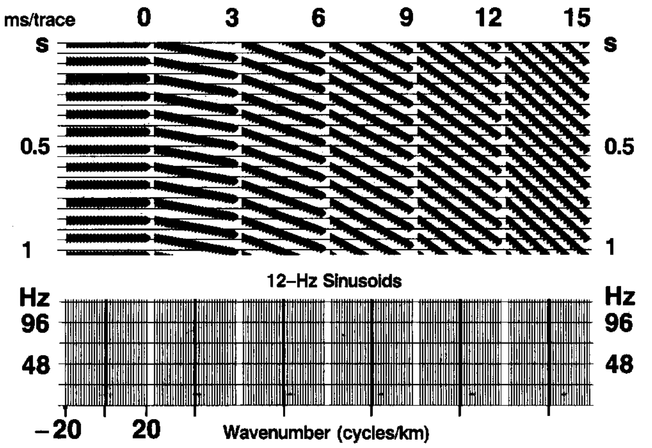


FIG. 1-65. Top row: Six gathers, each containing 12-Hz monofrequency events with different dips ranging from 0 to 15 ms/trace. Trace spacing is 25 m. Bottom row: Their respective amplitude spectra. The small dark spots on the spectra represent the mapping of events on the gathers. The solid vertical lines are the frequency axis. If the positive dips are defined as downdip from left to right, then all events map onto the positive quadrant in the frequency-wavenumber *(f-k)* plane. This is the first in a series of six figures that describes mapping of monofrequency signals in the *(f,k)* domain (Figures 1-65 through 1-70).

1.6 THE 2-D FOURIER TRANSFORM

Multichannel processing operations can be loosely defined as those that must operate on several data traces simultaneously. Multichannel processes can be useful in discriminating against noise and enhancing signal on the basis of a criterion that can be distinguished from trace to trace, such as dip or moveout. The 2-D Fourier transform is a basis for both analysis and implementation of multichannel processes.

Consider the six zero-offset sections in Figure 1-65. The trace spacing is 25 m with 24 traces per section. All have monochromatic events with 12-Hz frequency, but with dips that vary from 0 to 15 ms/trace. From the 1-D Fourier transform discussion (Section 1.2), we know about frequency, particularly *temporal frequency*, or the number of cycles per unit time. This is the Fourier dual for the time variable. However, a seismic wave field is not only a function of time, but also a function of a space variable (offset or midpoint axis). The Fourier dual for the space variable is defined as *spatial frequency*, which is the number of cycles per unit distance, or *wavenumber*. Just as the temporal frequency of a given sinusoid is determined by counting the number of peaks within a unit time, say 1 s, the wavenumber of a dipping event is determined by counting the number of peaks within a unit distance, say 1 km, along the horizontal direction. Just as the temporal Nyquist frequency is defined as [1/(2 · sampling interval)], the Nyquist wavenumber is defined as [1/(2 · trace interval)]. For all of the sections in Figures 1-65 through 1-70, the Nyquist wavenumber is 20 cycles/km, since the trace interval is 25 m.

To compute the wavenumber that is associated with the section corresponding to the 15 ms/trace dip in Figure 1-65, follow a peak or trough across the section. First compute the total time dip across the section:

(23 traces/section) · (15 ms/trace) = (345 ms/section).

Then convert this to cycles by dividing by the (temporal) period:

(345 ms/section)/[(1000 ms/s)/(12 cycles/s)] =
(4.14 cycles/section).

The spatial extent of the section is 575 m; therefore, the 15 ms/trace wavenumber associated with this dip and the 12-Hz frequency is

(4.14 cycles/section)/(0.575 km/section) = 7.2 cycles/km.

To continue this discussion, we will map these sections to the plane of temporal frequency versus spatial wave-

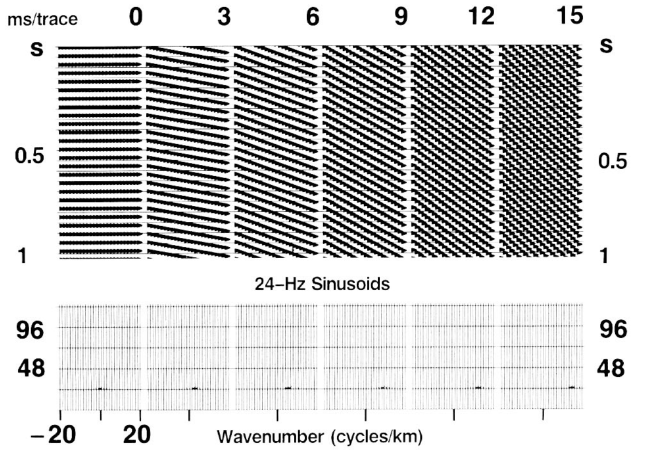


FIG. 1-66. The same as Figure 1-65, except using 24-Hz monofrequency events.

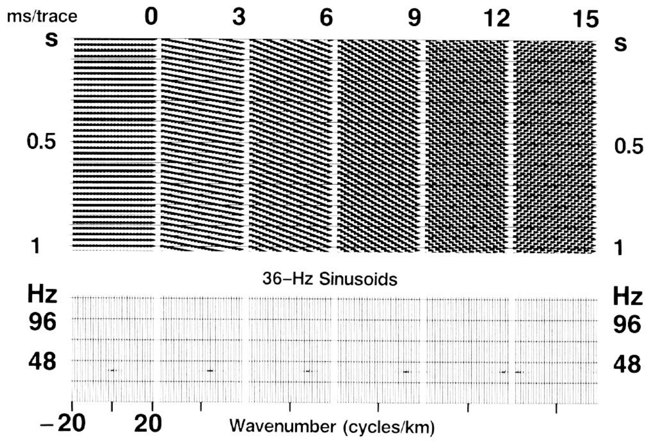


FIG. 1-67. The same as Figure 1-65, except using 36-Hz monofrequency events.

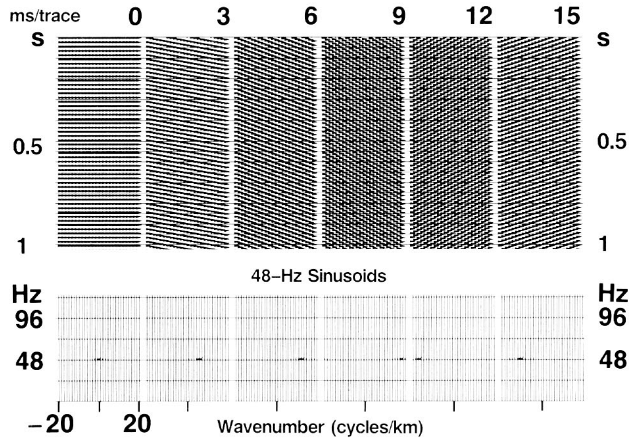


FIG. 1-68. The same as Figure 1-65, except using 48-Hz monofrequency events.

number, then look at two quadrants of this plane. The following convention will be used: Events with downdip to the right are assigned positive dip, while events with updip to the right are assigned negative dip. Additionally, positive dips map onto the right quadrant, which corresponds to positive wavenumbers, while negative dips map onto the left quadrant, which corresponds to negative wavenumbers.

The plane of frequency-wavenumber (the f-k plane) appears at the bottom of each section in Figure 1-65. The section with zero-dip events maps onto a single point on the frequency axis at 12 Hz. Zero dip is equivalent to zero wavenumber. The magnitude of the spike corresponds to the peak amplitude of the sinusoids that make up the traces in the section. Therefore, the f-k plane actually represents the 2-D amplitude spectrum of the (t,x) section. These data have been transformed from the time-space domain to the frequency-wavenumber domain. This process is mathematically described by the 2-D Fourier transform.

There is a practical relationship between the four variables: space-time, (t,x), and their Fourier duals, (f,k). Measure the inverse of the stepout in the 15 ms/trace section in Figure 1-65 by following a peak, trough, or zero crossing from trace to trace. Stepout is defined as the slope, dt/dx. In this case, the inverse of the stepout is:

$$dx/dt = 575 \text{ m}/0.345 \text{ s} = 1.67 \text{ km/s}.$$

Now, compute the ratio:

$$f/k = (12 \text{ cycles/s})/(7.2 \text{ cycles/km}) = 1.67 \text{ km/s}.$$

From this, the inverse of slope dt/dx measured in the (t,x) space along a constant phase is equal to the ratio of the frequency to the wavenumber associated with the event. Therefore, while retaining fixed stepout dt/dx, doubling the frequency means doubling the wavenumber.

Note that all sections in Figure 1-65 have the same frequency component. However, from 0 to 15 ms/trace, the number of peaks increases horizontally across each section. That is, for a given frequency, higher dips are assigned to higher wavenumbers, as seen on the f-k plots.

From Figures 1-65 through 1-70, consider the same dip components, but at different frequencies. Map each individual section to the f-k plane. Nothing unusual happens until the section with 15 ms/trace dip at 36 Hz is reached in Figure 1-67. Here there is no positive dip. In fact, as a whole, the section displays a checkerboard character. It is difficult to determine whether the dip is positive or negative.

At 48 Hz (Figure 1-68), the correct dip direction is seen in the first four sections. However, the fifth section, which corresponds to the 12 ms/trace positive dip, shows a negative dip. Therefore, it is mapped onto the negative quadrant, which is the wrong quadrant for this section. This dip component (12 ms/trace) at this frequency (48 Hz) is *spatially aliased*. In fact, any dip greater than 12 ms/trace is spatially aliased at this frequency.

In the next set of sections in Figure 1-69, spatial aliasing occurs at 60 Hz for a 9 ms/trace dip. Spatial aliasing not only causes mapping to the wrong quadrant, but also

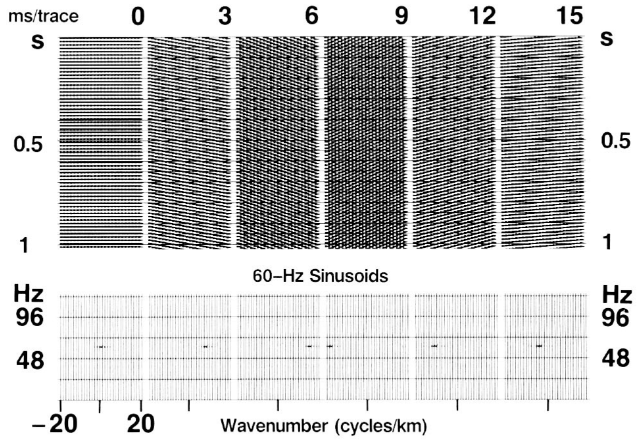


FIG. 1-69. The same as Figure 1-65, except using 60-Hz monofrequency events.

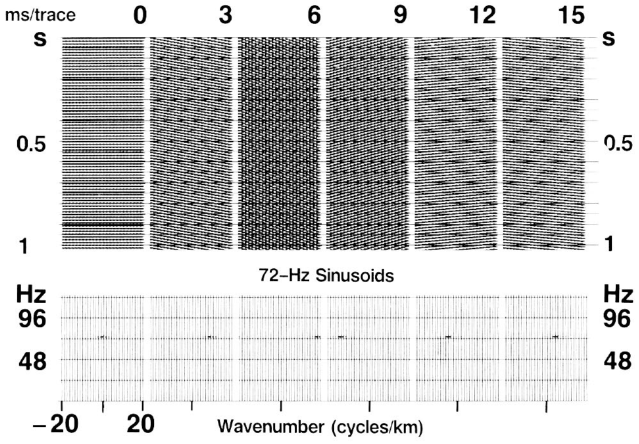


FIG. 1-70. The same as Figure 1-65, except using 72-Hz monofrequency events.

causes mapping with the wrong dip. One obvious example of this is mapping a 15 ms/trace dip at 60 Hz (Figure 1-69). Finally, at 72 Hz (Figure 1-70), the 6 ms/trace dip component is on the verge of spatial aliasing. Moreover, the 15 ms/trace dip component is spatially aliased twice; it folds back to the positive-dip quadrant and appears at a lower dip.

This same analysis can be used for the negative-dip components. From Figures 1-65 through 1-70, note that each section as a whole was mapped onto a single point in the frequency-wavenumber domain. Each section has an associated unique frequency and wavenumber assigned to it. These zero-offset sections can be considered representations of plane waves that propagate at a unique angle from the vertical and carry a monochromatic signal. The wavefront is defined as the line of constant phase, while the direction of propagation is perpendicular to the wavefront. Since a seismic wave field is a superposition of many dips and frequencies, it is equivalent to the synthesis of many plane-wave components. In this respect, the physical meaning of the 2-D Fourier transform is important, for it is a way to decompose a wave field into its plane-wave components.

A recorded wave field is a composite of many dip and frequency components, such as those shown in Figures 1-65 through 1-70. Suppose that sections with the same dip, but with different frequencies, are superimposed. The composite sections are shown in Figure 1-71 with the composite amplitude spectra below each section. For a given dip, all frequency components map onto the *f-k* plane along a straight line that passes through the origin. The higher the dip, the closer the radial line in the f-k domain is to the wavenumber axis. The zero-dip components map along the frequency axis. From the 9, 12, and 15 ms/trace dips, note that the spatially aliased frequencies are located along the linear segments that wrap around to the opposite quadrant in the amplitude spectrum. The steeper the dip, the lower the frequency at which spatial aliasing occurs.

So far, a discrete number of frequencies was considered. For a continuum of frequency components associated with a single dip, we anticipate that they would map along a straight, continuous line in the *f-k* domain. This is shown in Figure 1-72. The dipping event in Figure 1-73 is spatially aliased beginning at approximately 21 Hz.

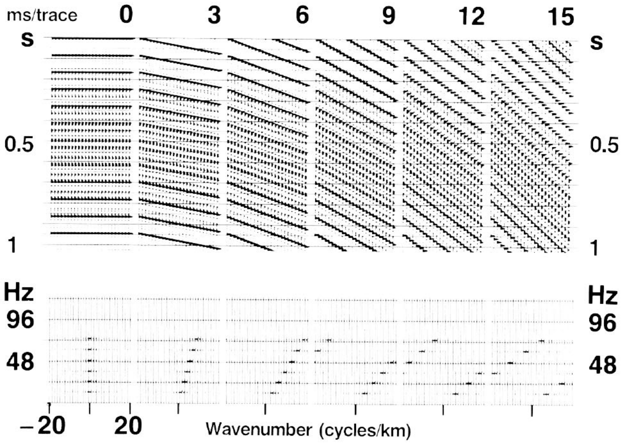


FIG. 1-71. Top row: Six gathers, each formed by summing gathers of the like dips in Figures 1-65 through 1-70. The trace spacing is 25 m. Bottom row: Respective amplitude spectra.

Examination of the monochromatic single-dip sections in Figures 1-65 through 1-70 shows that each section maps onto a single point in the *(f,k)* domain. An extension of this observation is made in Figure 1-74. Events with the same dip in the *(t,x)* space, regardless of their location, map onto a single radial line in the *(f,k)* space. When events are spatially aliased, the radial line wraps around at the Nyquist wavenumber (Figure 1-75). These concepts have important practical implications, for they lead to *f-k* dip filtering (Section 1.6.2). Events with different dips that may interfere in the *(t,x)* domain can be isolated in the *(f,k)* domain.

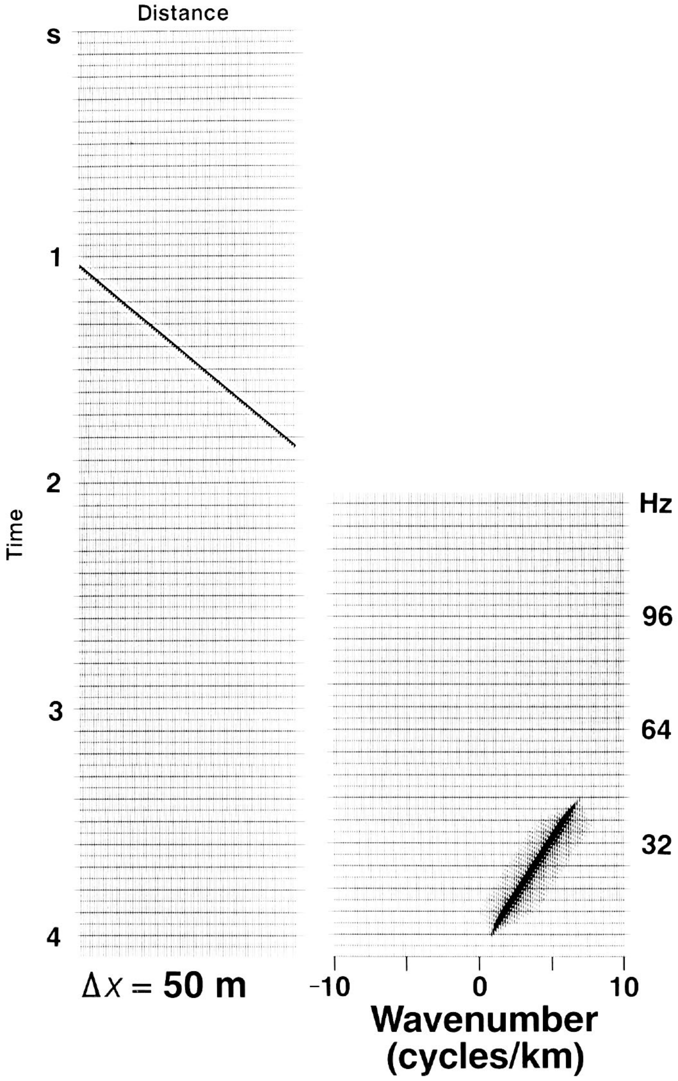


FIG. 1-72. A single, isolated dipping event and its 2-D amplitude spectrum. No frequency is spatially aliased. What is the dip of the event in milliseconds per trace?

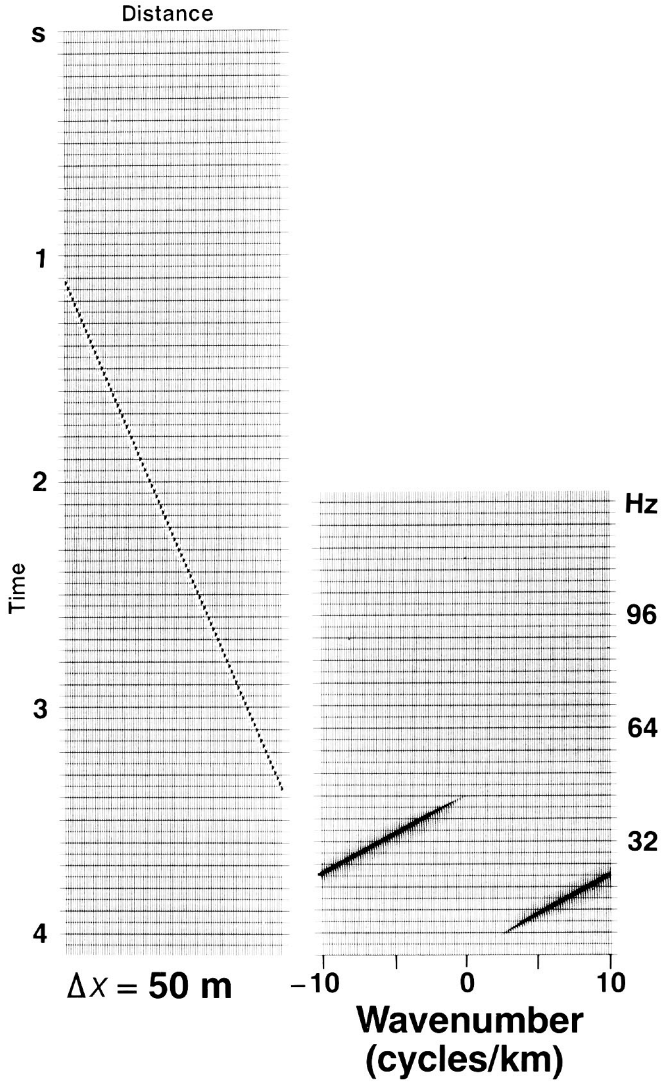


FIG. 1-73. A single, isolated dipping event and its 2-D amplitude spectrum. Frequencies beyond 21 Hz are spatially aliased.

The numerical computation of the 2-D Fourier transform involves two 1-D Fourier transforms. Figure 1-76 shows the steps that are involved. A brief mathematical formulation of the 2-D Fourier transform is given in Appendix A.

1.6.1 Spatial Aliasing

While studying the *(f,k)* plane, spatial aliasing was discussed. Actually, spatial aliasing has serious effects on the performance of multichannel processes such as *f-k* filtering (Section 1.6.2) and migration (Section 4.3.5). Because of spatial aliasing, these processes can perceive events with steep dips at high frequencies as different from what they actually are (such as the 15 ms/trace dips in Figures 1-68 through 1-70). Therefore, they are not treated properly. For example, migration moves the spatially aliased frequency components in the wrong direction and generates a dispersive noise that degrades the quality of the migrated section. This problem is discussed in Section 4.3.

How is spatial aliasing avoided? Compare the models in Figures 1-72 and 1-73. Both have the same frequency content, 6 to 42 Hz. The data in Figure 1-73 are spatially aliased because the dipping event is steeper than in Figure 1-72. Some ways to avoid spatial aliasing follow:

1. Apply time shifts so that the steep events appear to have lower dips. However, this could change the dips that were low to higher dips, making them spatially aliased. Nevertheless, this often is a feasible solution for certain situations.
2. If a low-pass filter were applied to the traces in

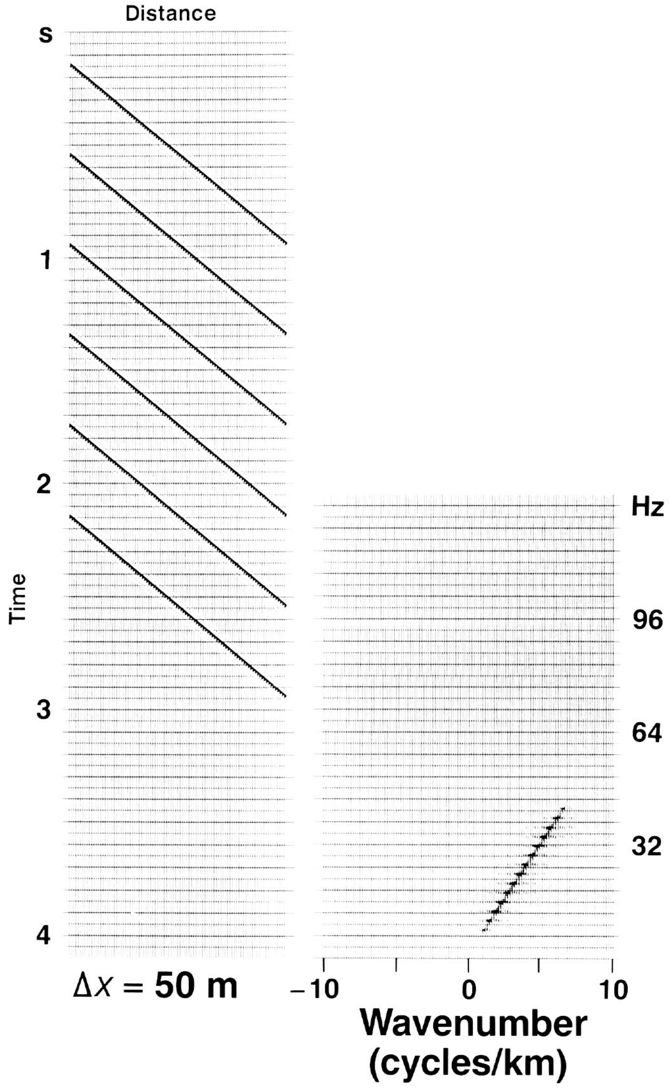


FIG. 1-74. Six events with identical dip map onto the same radial line in the (f,k) domain. No frequencies are aliased. The dip of these events is the same as that of the single event in Figure 1-72. What distinguishes this spectrum from that in Figure 1-72?

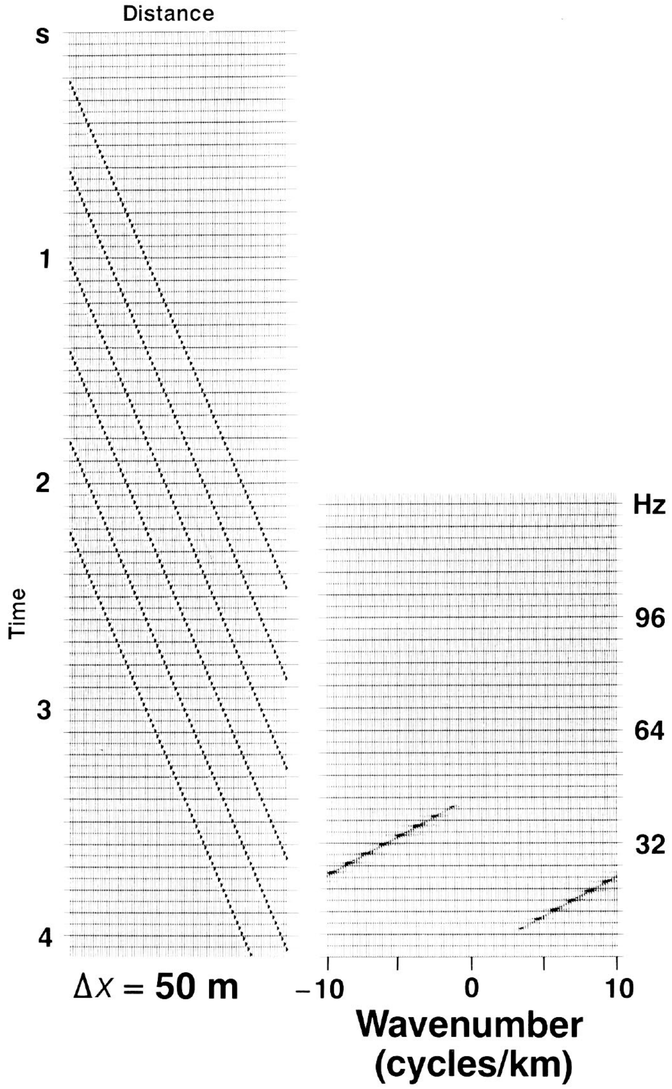


FIG. 1-75. Six events with identical dip map onto the same radial line in the (f,k) domain. Frequencies beyond 21 Hz are aliased. The dip of these events is the same as that of the single event in Figure 1-73.

Figure 1-73 so that the frequencies up to 21 Hz were retained, then the segment that is wrapped around to the negative quadrant of the amplitude spectrum is removed. Although spatial aliasing is eliminated, a significant part of the recorded frequency band is lost. This approach is not desirable.

3. Figure 1-77 shows a single dipping event recorded with three different trace spacings. The amplitude spectra suggest a third approach to solving the spatial aliasing problem. Note that the coarser the trace spacing, the more frequencies are spatially aliased. The same frequency bandwidth is kept in all three cases. The 12.5-m trace spacing provides a frequency band with no spatial aliasing. For a 25-m trace spacing, frequencies beyond 36 Hz are spatially aliased; while for a 50-m trace spacing, frequencies beyond 18 Hz are spatially aliased. For this

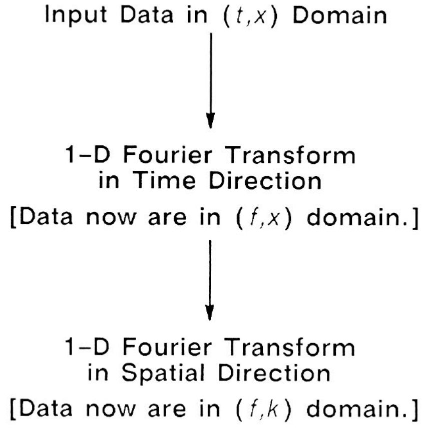


FIG. 1-76. Computation of 2-D Fourier transform.

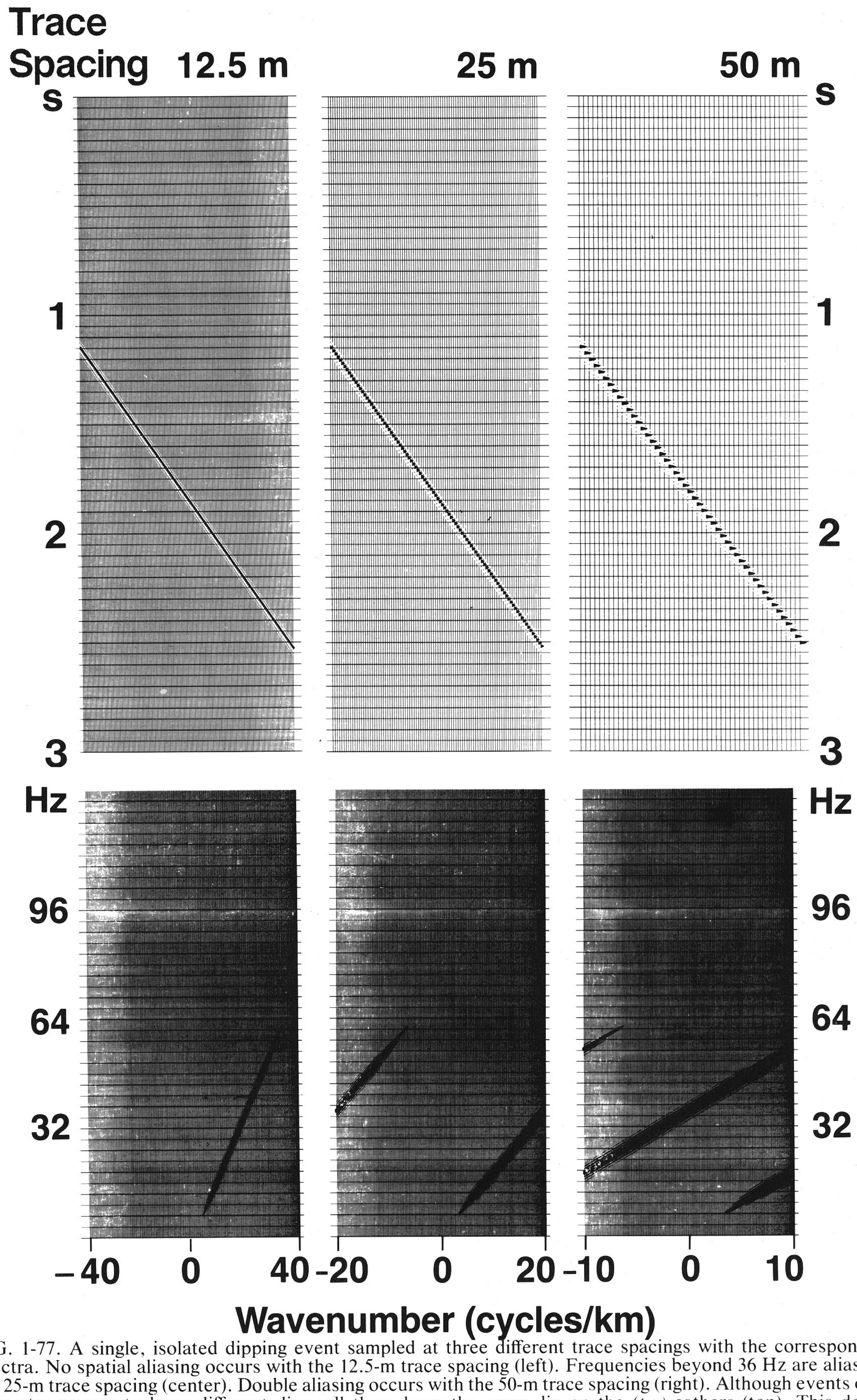


FIG. 1-77. A single, isolated dipping event sampled at three different trace spacings with the corresponding *f-k* spectra. No spatial aliasing occurs with the 12.5-m trace spacing (left). Frequencies beyond 36 Hz are aliased with the 25-m trace spacing (center). Double aliasing occurs with the 50-m trace spacing (right). Although events on the *f-k* spectra appear to have different dips, all three have the same dip on the (t,x) gathers (top). This deceptive character is because of the different horizontal scales used in displaying the *f-k* spectra.

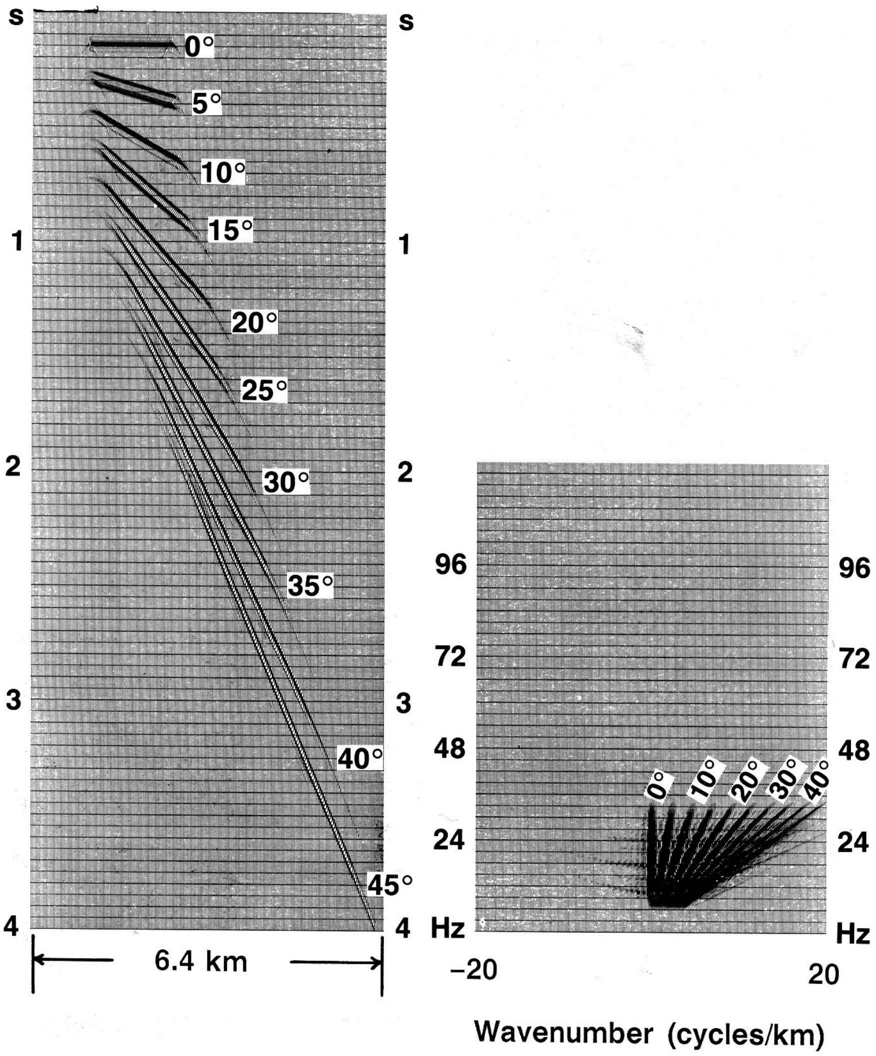


FIG. 1-78. A zero offset section (256 traces with 25-m trace spacing) containing 10 dipping events and its 2-D amplitude spectrum. No frequencies are aliased.

latter case, spatial aliasing is so severe that the aliased frequencies wrap around the wavenumber axis twice. From this, we see that spatial aliasing can be avoided by selecting a sufficiently small trace spacing. This approach requires either a data-dependent interpolation scheme to generate extra traces or modification of the field recording geometry. If the latter approach were taken, more shots and/or more recording channels are needed.

At this point, only the synthesis of a single dipping event from a discrete number of frequency components has been considered. This analysis now extends to a range of dips. Figure 1-78 shows a section with dips that vary from 0 to 45 degrees and the corresponding 2-D amplitude spectrum. These same dips, but with higher frequency content, also are seen in Figure 1-79. Events with 0-, 5-, 10-, and 15-degree dips are not spatially aliased. The 20-degree dip is aliased at nearly 72 Hz, the 30-degree dip at nearly 48 Hz, and the 45-degree dip at nearly 36 Hz. Again, the steeper the dip, the lower the frequency at which spatial aliasing occurs.

Given a dip value, how is the maximum unaliased frequency determined? Consider the 20-degree dipping event in Figure 1-79. First, measure the dip in milliseconds per trace. There are 256 traces in the (t,x) model with 25-m trace spacings. The 20-degree dip is equivalent to 7 ms/trace. Frequency components with periods less than twice the dip are spatially aliased. Thus, given the dip in milliseconds per trace, the threshold frequency at which spatial aliasing begins is 500 per dip. In the present case, the threshold frequency is $500/7 = 72$ Hz. This is verified by examining the amplitude spectrum in Figure 1-79. Derivation of the threshold frequency formula is considered in Section 4.3.5.

Figure 1-80 shows three field records and their 2-D amplitude spectra, known as *f-k* spectra. By now, it is easy to recognize and relate various events on the shot gathers to those on the *f-k* spectra. Event A is the high-

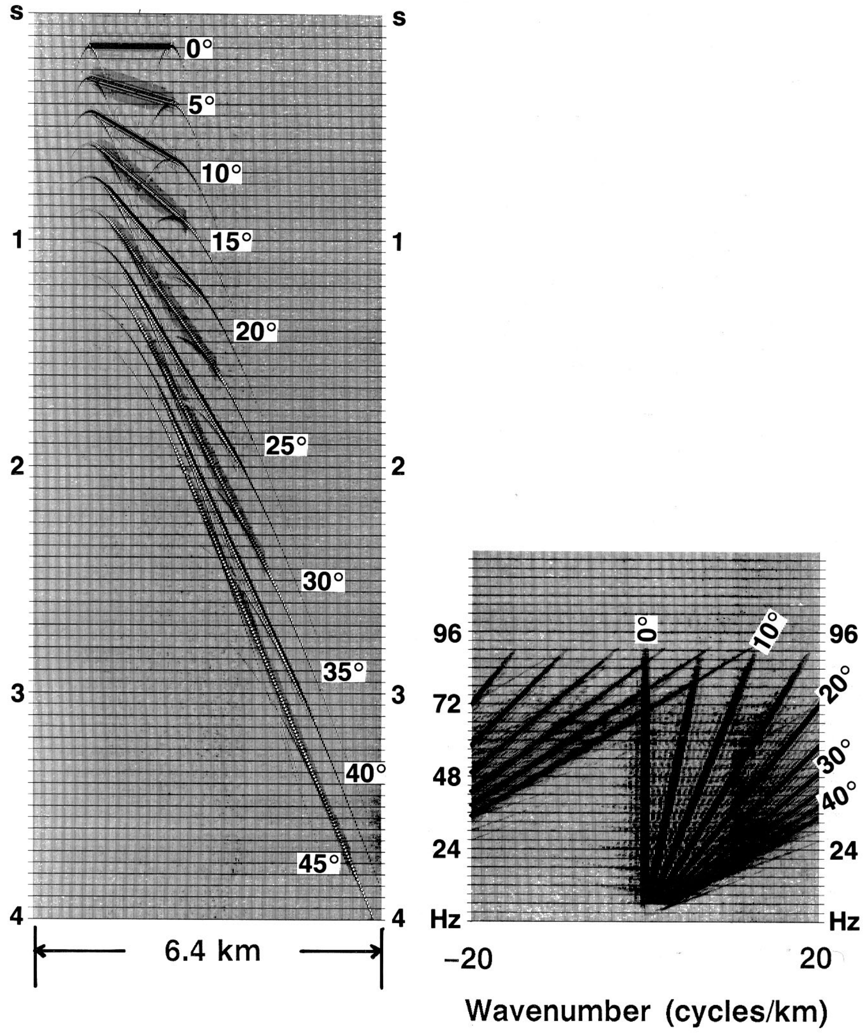


FIG. 1-79. A zero-offset section (256 traces with 25-m trace spacing) containing 10 dipping events and its 2-D amplitude spectrum. Steeper dips are aliased at increasingly lower frequencies.

amplitude dispersive coherent noise with very low group velocity. When the spatial extent of these waves broadens, bandwidth in the wavenumber direction becomes smaller. Conversely, when the spatial extent becomes smaller, the event, such as G, spants a wider wavenumber bandwidth on the *f-k* plot (compare events G and F). Events B and C are parts of the guided wave packet. Event C contains aliased energy above 42 Hz (indicated by D on the *f-k* plot). Primaries and associated multiples are mapped into region E between the frequency axis and event C.

1.6.2 *F-K* Dip Filtering

Events that dip in the *(t,x)* plane can be separated in the *(f,k)* plane by their dips. This allows us to eliminate certain types of unwanted energy from the data. In particular, coherent linear noise (in the form of ground roll), guided waves, and side-scattered energy commonly obscure the genuine reflections that may be present in recorded data. These types of noise usually are isolated from the reflection energy in the *(f,k)* space. From the field record in Figure 1-81a, note how ground roll energy can dominate the data. This is a walk-away noise test that is composed of six shot records. Ground roll is a type of dispersive waveform that propagates along the surface and is low-frequency, large-amplitude in character. Typically, ground roll is suppressed in the field by the receiver array.

Figure 1-81b is the 2-D amplitude spectrum of the composite shot gather in Figure 1-81a. Here, various types of energy are well isolated from one another. Ground roll A, its back-scattered component B, and guided waves C, are identifiable. Reflections D are situated around the frequency axis. As shown in Figure 1-81c, a fan is imposed on this spectrum within which the undesired energy is

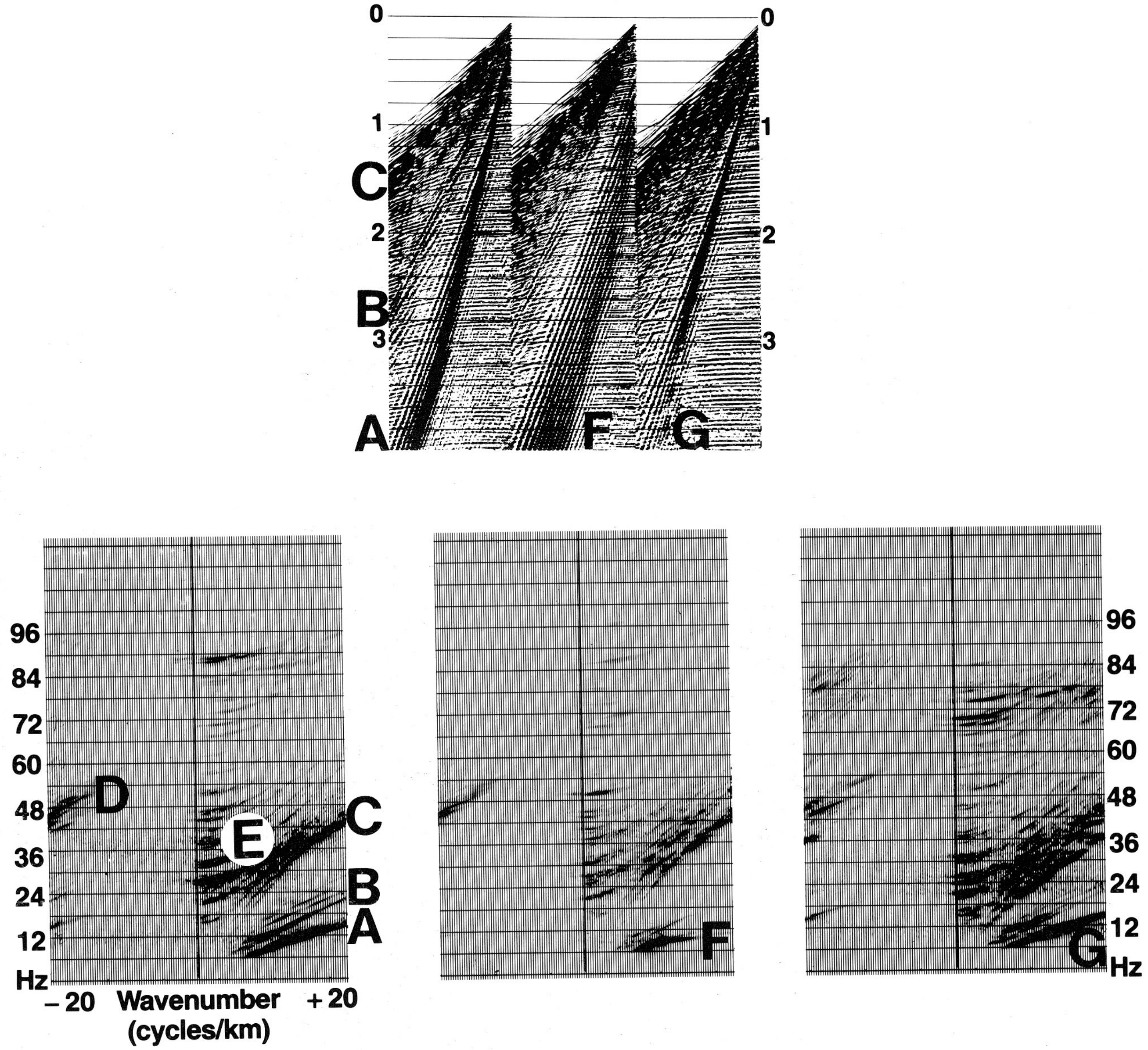
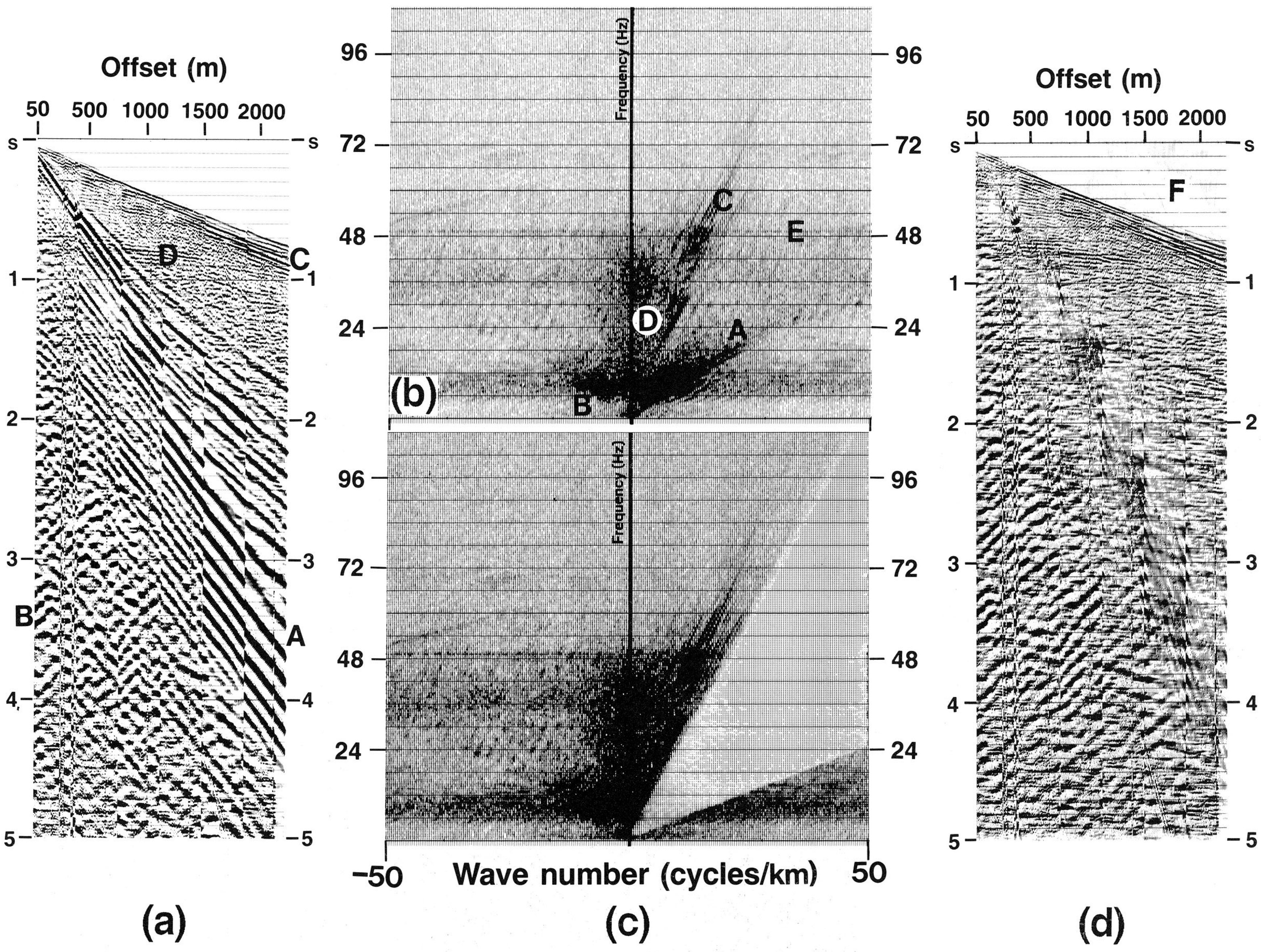


FIG. 1-80. Three common-shot gathers (top) and their *f-k* spectra (bottom). (The symbols are discussed in the text.) Dip convention: An event maps onto a positive-dip quadrant on the *f-k* spectrum if it dips down moving from near to far offsets. (Data courtesy Deminex Petroleum.)

zeroed out. This is followed by inverse mapping back to the *(t,x)* space. The resulting composite shot gather in Figure 1-81d is virtually free of ground roll energy, except for the back-scattered component. Zeroing out a fan in the *(f,k)* space is one implementation of the process known as *f-k* dip filtering. Figure 1-82 shows the steps involved in this process.

Practical issues associated with the 2-D Fourier transform and the choice of a fan reject zone are outlined below:

1. Conventional implementations of the Fourier transform itself produce wraparound noise. This is apparent in Figure 1-81d, location F. To avoid this problem, the data must be extended beyond the ranges of the spatial and temporal axes by padding with zeroes. The size of the input gather typically is increased by a factor of 4, which is equivalent to doubling the length in *t* and *x*. This increases the cost, but removes the wraparound effects.
2. The fan width must not be too narrow. This follows from previous observations of the 1-D Fourier analysis of frequency filters. If the bandwidth of the reject zone were narrow, then the *(t,x)* response of the dip filter would have a large array of nonzero elements. Fortunately, coherent noise with large stepouts, such as ground roll, often is isolated in *(f,k)* space from the zone that includes the reflection signal. This is demonstrated by the example in



(a) (c) (d)

FIG. 1-81. (a) Composite field record obtained from a walk-away noise test. Trace spacing = 10 m, A = ground roll, B = a backscattered, component of A, C = dispersive guided waves, D = primary reflection. (b) The *f-k* spectrum of this field record. Dip convention is the same as in Figure 1-80. Event E is referred to in Exercise 1-17. (c) The *f-k* spectrum of the field record after rejecting ground roll energy A. Compare this with the *f-k* spectrum (b) of the original record. For display purposes, each is normalized with respect to its own maximum. (d) Dip-filtered field record whose *f-k* spectrum is shown in (c). Compare this record with the original in (a). (Data courtesy Turkish Petroleum Corp.)

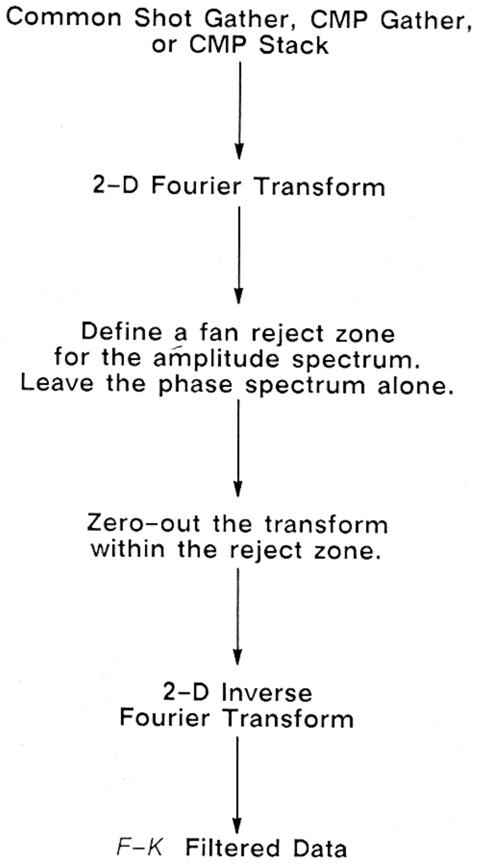


FIG. 1-82. The *f-k* dip filtering.

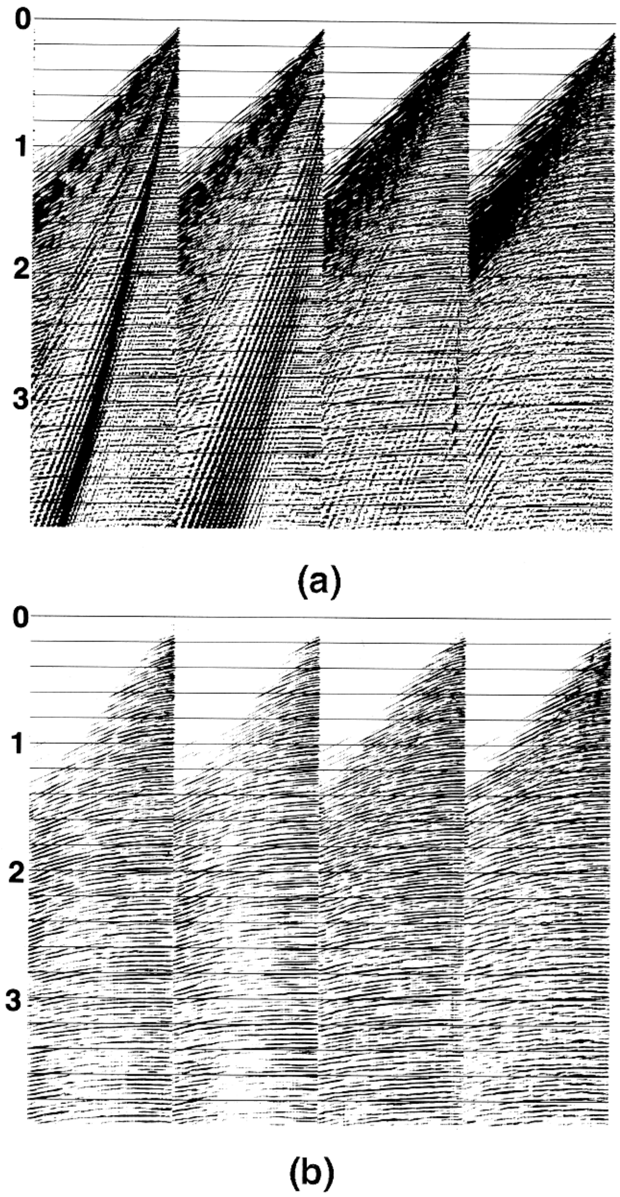


FIG. 1-83. Four shallow marine records (a) before and (b) after *f-k* dip filtering to remove coherent linear noise. The coherent noise seen in these records is primarily of guided wave type. (Data courtesy Deminex Petroleum.)

Figure 1-81b. In such cases, ground roll energy A is suppressed without damaging the reflection signal by using a large fan (Figure 1-81c).

3. As for the 1-D frequency filters, the amplitude spectrum of the *f-k* filter must not have sharp boundaries. There must be a smooth transition from the reject zone to the pass zone. This is accomplished by tapering the fan edges, which is analogous to using slopes in frequency filtering. The amount of tapering must be large enough to be effective. On the other hand, it must not be so wide that it suppresses signal in the pass zone. Extra precaution is taken at low frequencies. As the fan thins to a zero width at the origin of the *(f,k)* plane, as in a wedge, the actual reject zone may invade the low-frequency components of the pass zone. This invasion occurs because the fan cannot get too narrow. It may be desirable to stop the reject zone just short of the low frequencies. This effectively excludes the low frequencies from the *f-k* dip filtering action.
4. Spatial aliasing often causes poor *f-k* filter performance. The fan reject zone must be extended to the spatially aliased frequency components. A practical approach to this problem is to apply time shifts to the data before *f-k* filtering so that the unwanted signal appears at lower dips, thus eliminating the spatial aliasing effects. The time shifts then are removed after *f-k* filtering. Unfortunately, this may not always work, since events not spatially aliased before may be spatially aliased after linear moveout.

The reject/pass zones in the *f-k* spectra do not need to be constrained to fan shape. Figure 1-83 shows four common-shot gathers, while Figure 1-84 shows their *f-k* spectra before and after *f-k* filtering. Quadrants left of the spectra mostly contain spatially aliased data. By keeping the fan shape in the right quadrant, while zeroing out

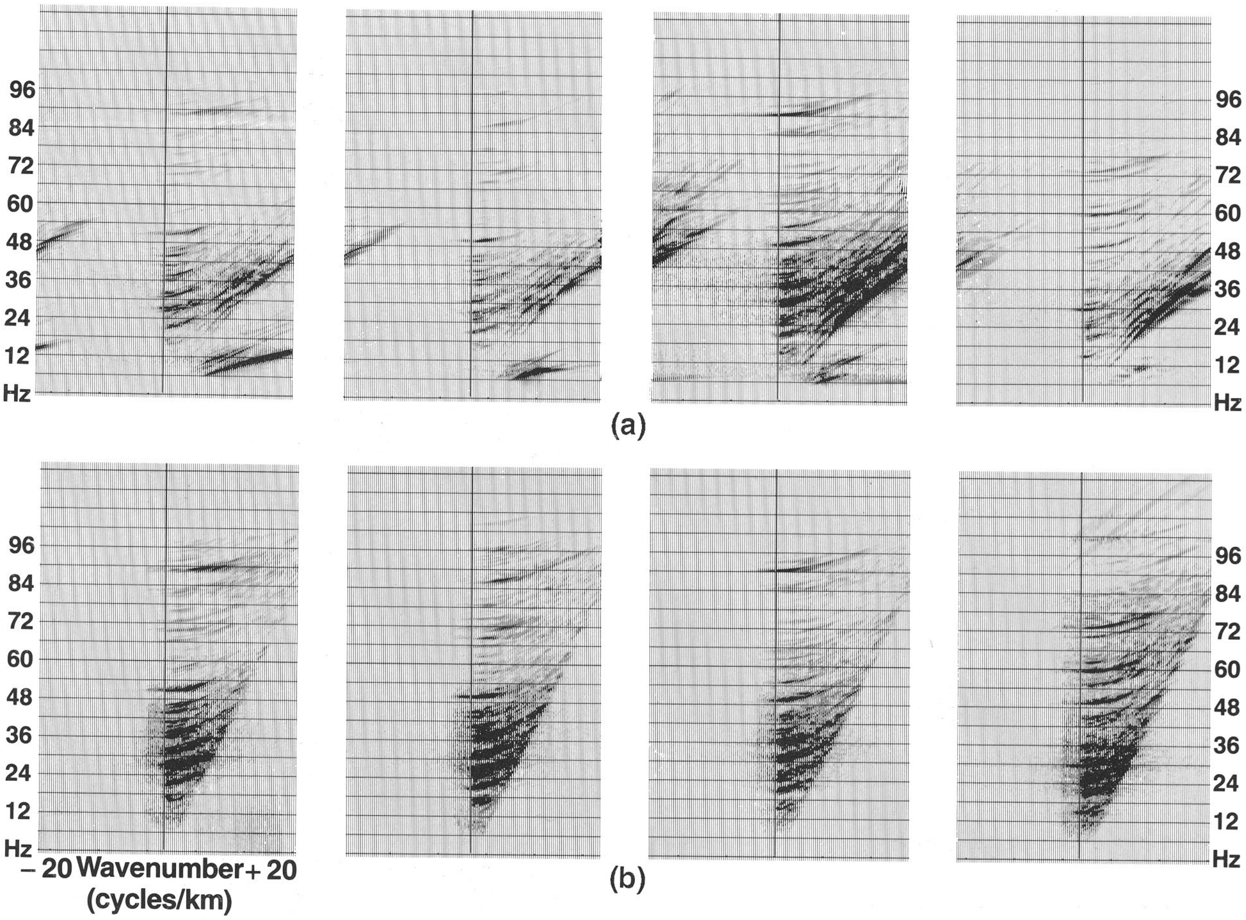


FIG. 1-84. The *f-k* spectra of the shot gathers in Figure 1-83 (a) before and (b) after *f-k* dip filtering. The dip convention is the same as for Figure 1-80.

most of the left quadrant, the coherent noise trains, including the aliased energy, are eliminated.

To what type of gather is *f-k* dip filtering applied? Because aliasing is a serious concern in this process, apply *f-k* filtering on the shot gathers rather than on the CMP gathers, since CMP gathers can have twice the trace spacing of the shot gathers. Two neighboring CMP gathers can be interleaved before applying *f-k* dip filtering, then split afterward, thus alleviating aliasing. Coherent noise as seen on the shot gathers in Figure 1-83 is suppressed in the shot-gather domain. When sorted into CMP gathers, no remnant of this noise is left in the data (Figure 1-85).

Figure 1-86a shows the stacked section derived from the CMP gathers in Figure 1-85a. Note that CMP stacking has suppressed most of the coherent noise. The *f-k* dip filtering after stack can be effective in suppressing any remaining coherent noise (Figure 1-86c). This result is comparable to the prestack *f-k* filtering result (Figure 1-86b). Coherent linear noise on stacked data also can be suppressed by poststack migration processes that incorporate dip filtering.

Two types of coherent linear noise that deserve special attention are guided waves and side-scattered energy. Guided waves manifest themselves as linear noise on both common-shot (Figure 1-83a) and CMP gathers (Figure 1-85a), but are attenuated largely by stacking (Figure 1-86a). Side-scattered energy manifests itself as linear noise on common-shot gathers (Figure 1-87a), is not apparent on CMP gathers (Figure 1-87b), but reappears as linear noise on stacked sections (Figure 1-88a) (Larner et al., 1983). The side-scattered energy has a large moveout range depending on the position of the scatterer. This type of energy stacks at high velocities along the linear flanks of its traveltime curve. We then anticipate that the linear noise seen on a stacked section, particularly at later times, is mostly scattered energy along the flanks of its traveltime curve, stacked together with any high-velocity primary energy. If *f-k* filtering were not applied to common-shot gathers containing side-scattered energy, then a stacked section with coherent linear noise could result (Figure 1-88a). Figure 1-88b shows that this noise can be suppressed by *f-k* dip filtering the stacked section. Could this noise suppression be done better? If an *f-k* filter were applied on shot records, then the stack in Figure 1-88c results. When compared to the result in Figure 1-88b, the result in Figure 1-88c seems to offer better resolution. The prestack *f-k* filtered data also

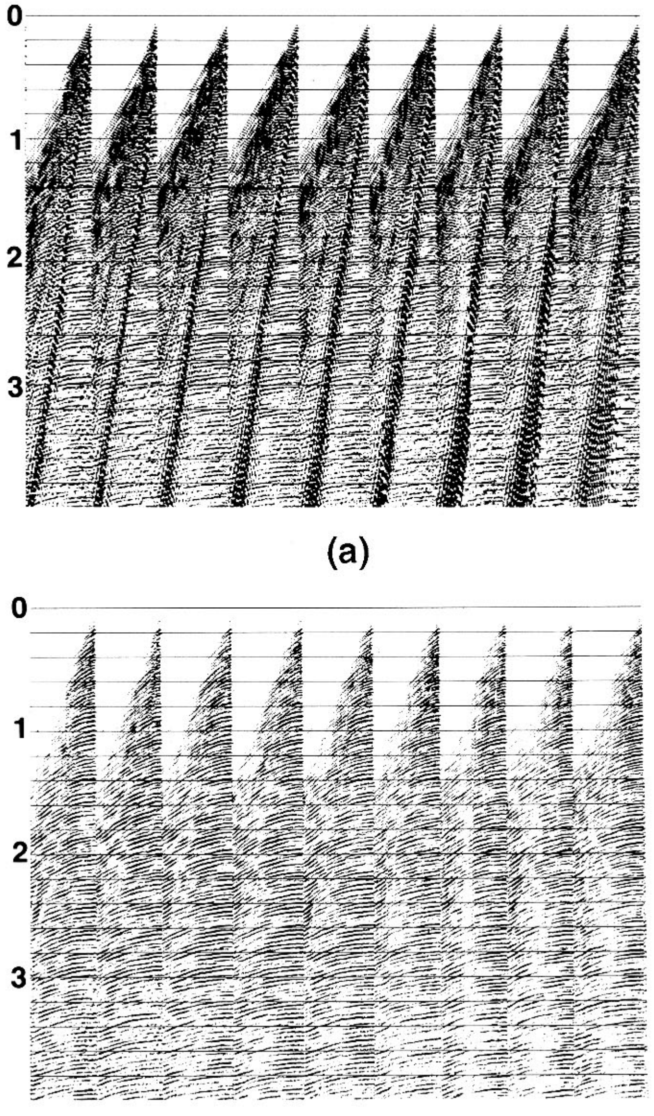


FIG. 1-85. The CMP gathers from a shallow marine survey (a) before and (b) after *f-k* dip filtering to remove coherent linear noise. (These are the same data as in Figure 1-83.) (Data courtesy Deminex Petroleum).

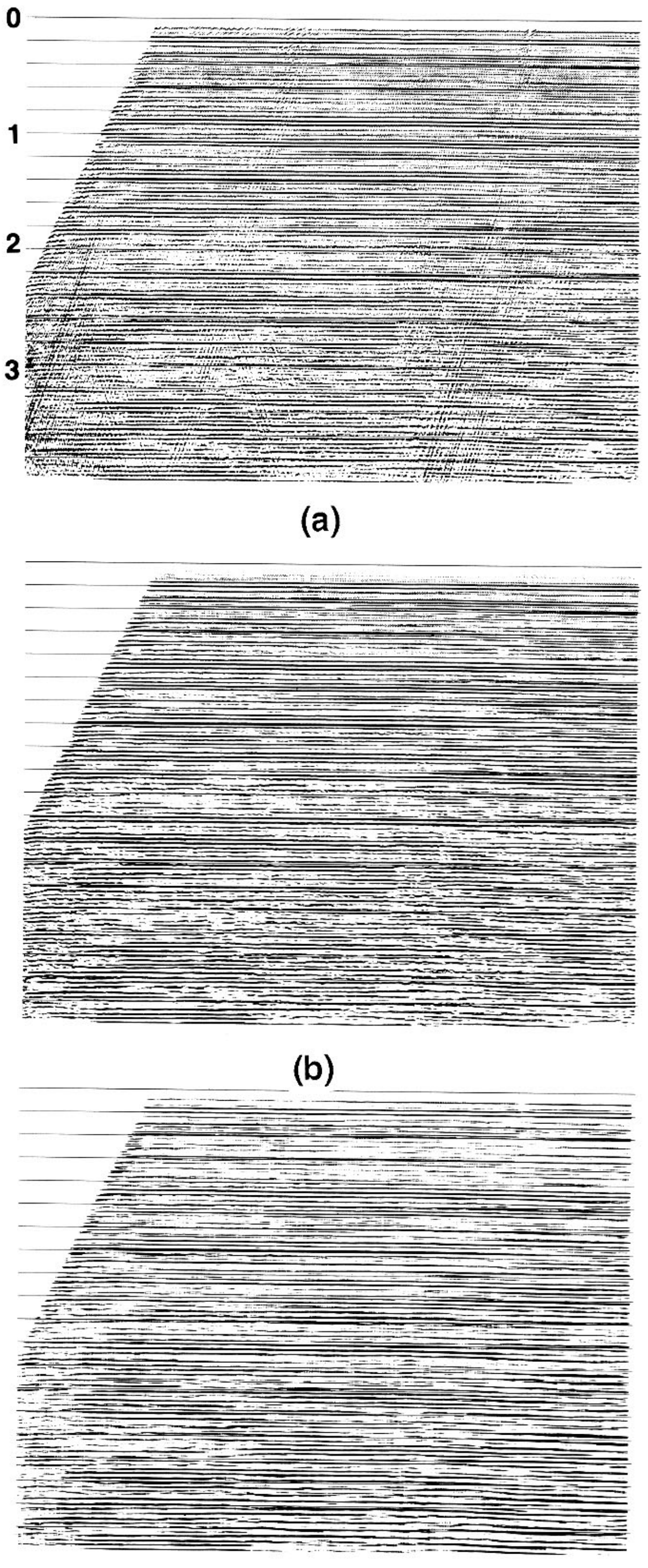
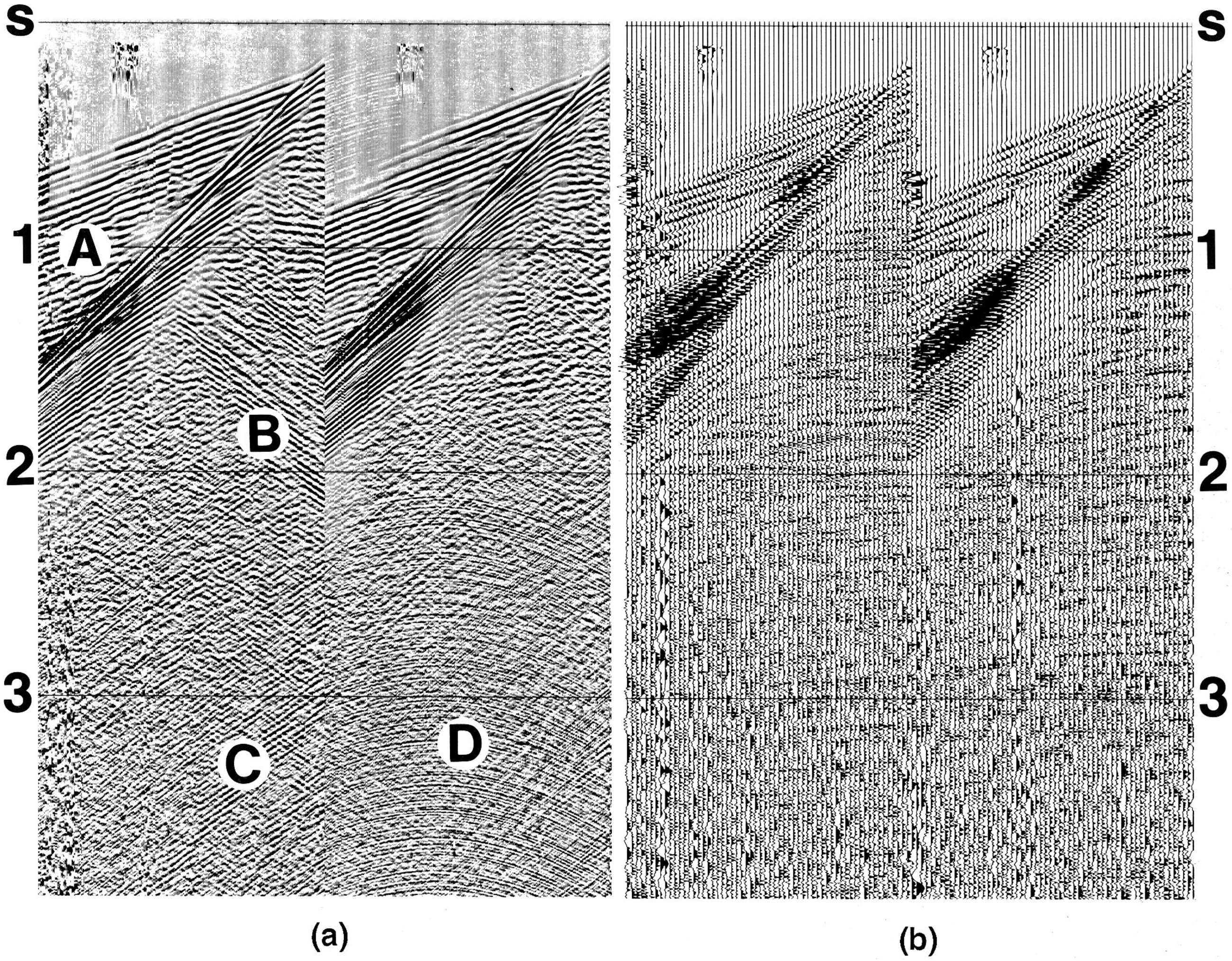


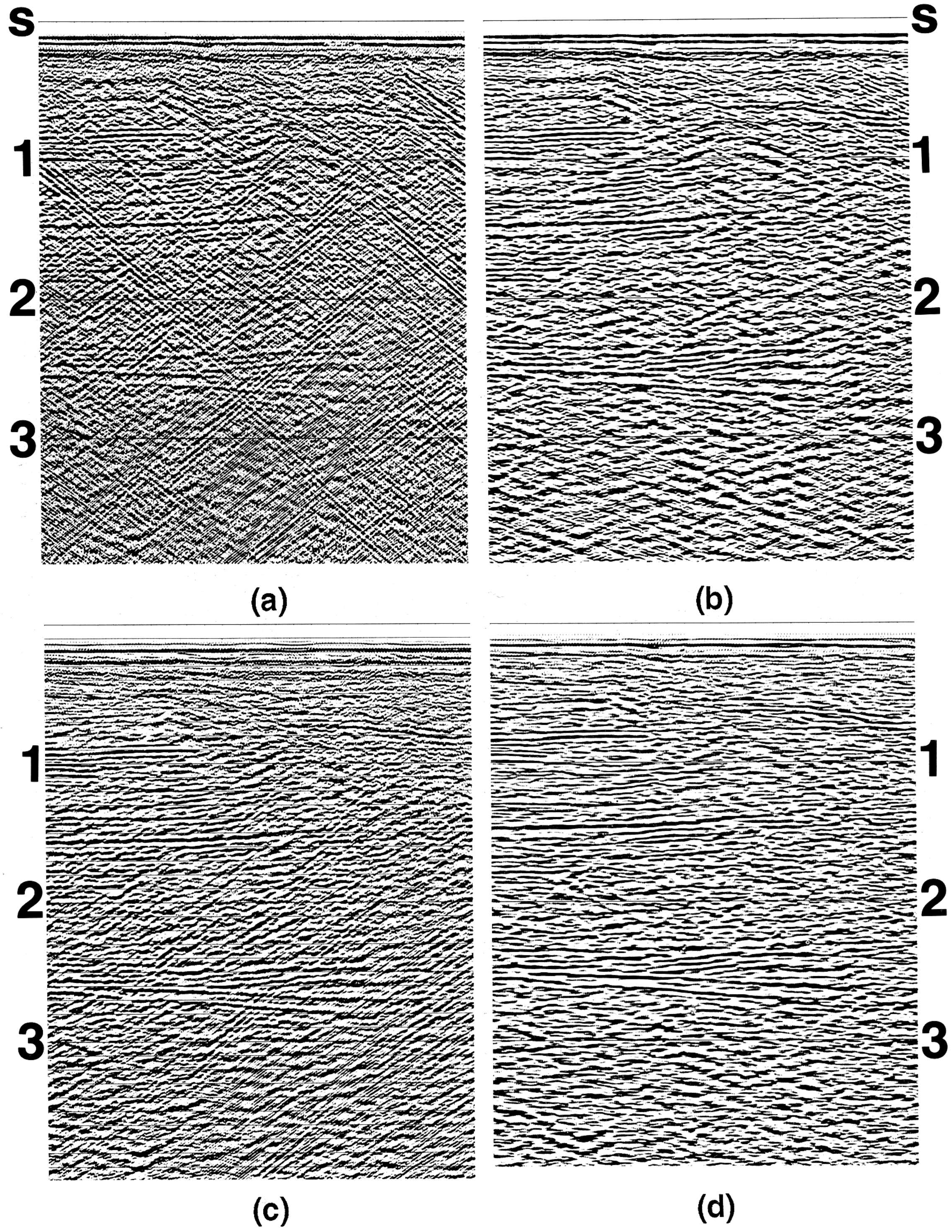
FIG. 1-86. (a) CMP stack with some coherent linear noise, *f-k* dip filtering (b) before and (c) after stack. Selected CMP gathers are shown in Figure 1-85. (Data courtesy Deminex Petroleum).



(a)

(b)

FIG. 1-87. (a) Two common-shot gathers and (b) two CMP gathers from the same line. CMP stack is shown in Figure 1-88a. Note the presence of coherent linear noise: A = guided wave packet (down to 2 s at far offset); B, C, = side-scattered noise dipping both from right to left and vice versa. What is the source of the energy D with hyperbolic traveltimes? (Data courtesy Taylor Woodrow Energy Ltd.)



(a) (b)

(c) (d)

FIG. 1-88. (a) CMP stack contaminated by coherent noise. Associated common-shot and CMP gathers are shown in Figure 1-87. (b) The same CMP stack *f-k* filtered after stack. (c) CMP stack *f-k* filtered before stack. (d) CMP stack *f-k* filtered twice before stack; first, in common-shot domain, second, in common-receiver domain. (Data courtesy Taylor Woodrow Energy Ltd.)

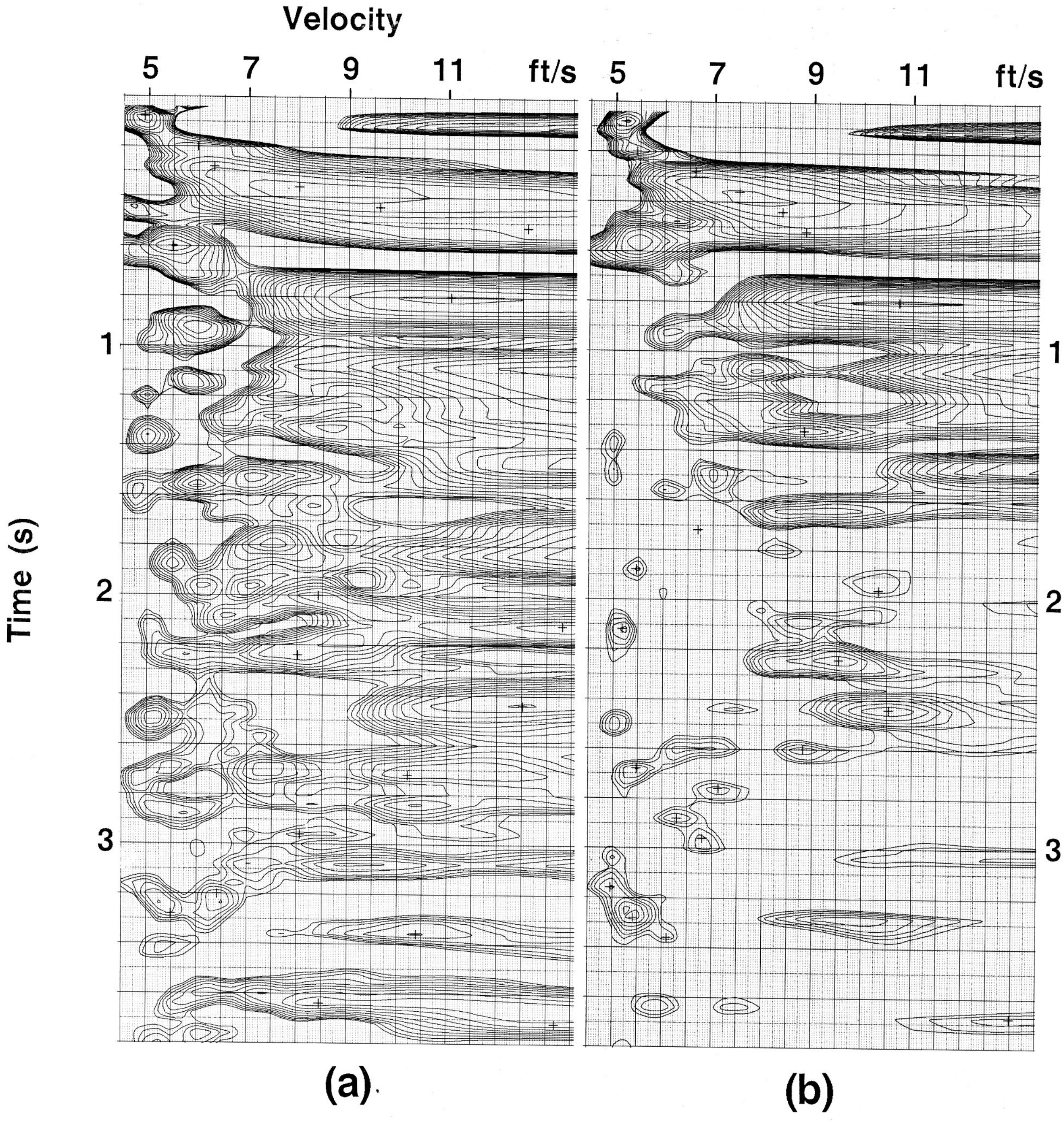


FIG. 1-89. Velocity spectrum associated with the data in Figure 1-87b (a) without and (b) with *f-k* dip filtering.

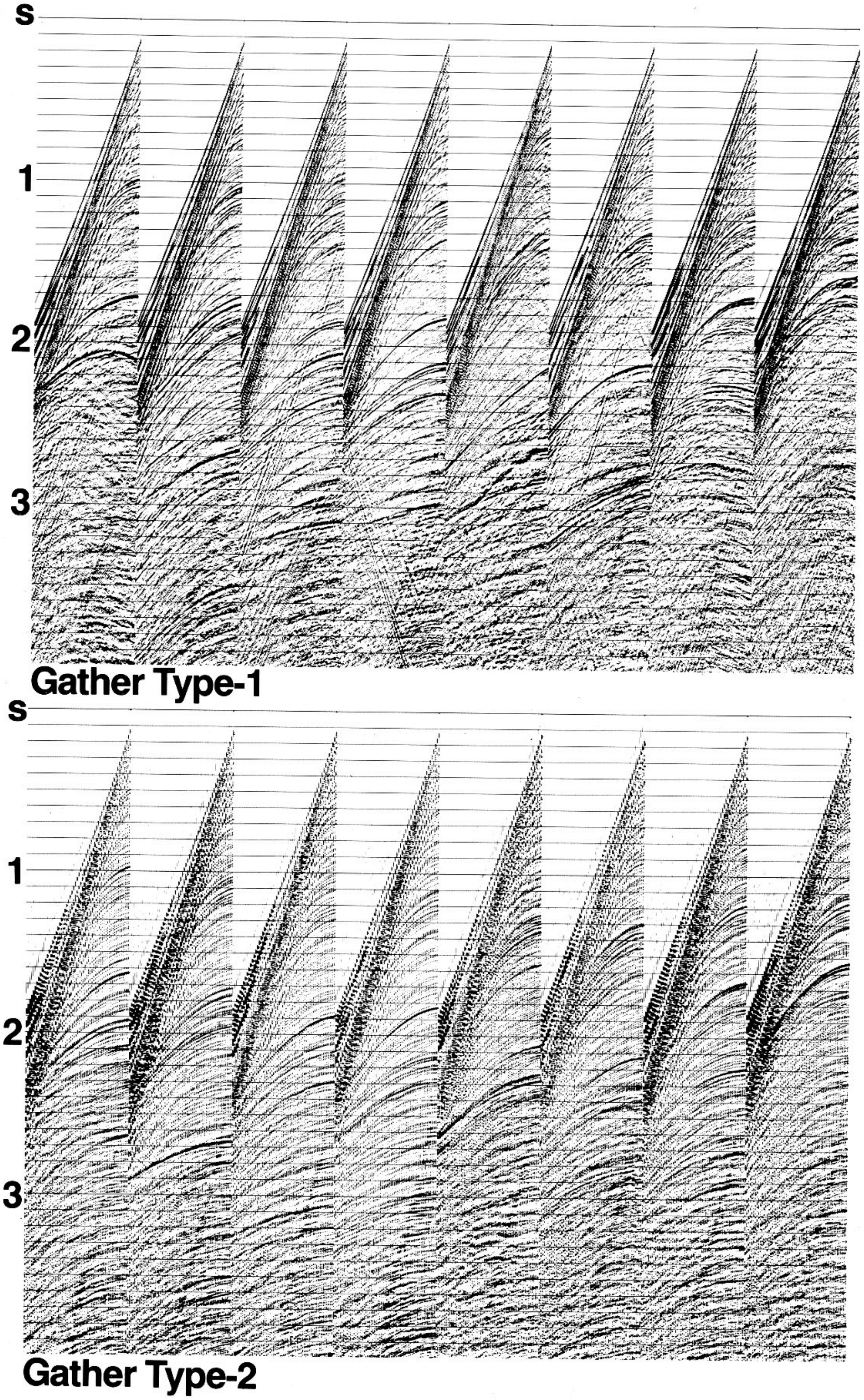


FIG. 1-90. A field dataset displayed in two different domains, common-shot and common-midpoint (see Exercise 1.13).

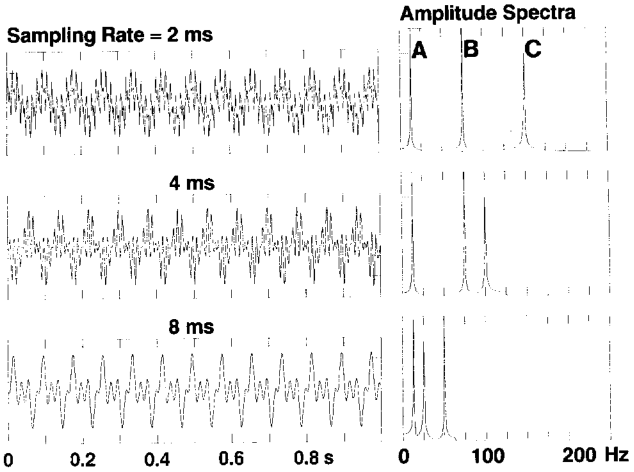


FIG. 1-91. A signal with three frequency components, A, B, and C, sampled at three different rates, 2, 4, and 8 ms. Frequency aliasing occurs at coarser sampling intervals (see Exercise 1-19).

yield an improved velocity analysis (Figure 1-89).

Practical experience with *f-k* filtering proves that an even better stack may result when both common-shot and common-receiver gathers are *f-k* filtered. The resulting stack is shown in Figure 1-88d. Compare this with Figures 1-88b and 1-88c. If *f-k* filtering were done on only the shot gathers, then a good portion of coherent noise may remain in the data. On the other hand, enhanced smearing of the data because of the second pass of *f-k* filtering can degrade reflector definition.

In summary, 2-D Fourier transformation is a way to decompose a wave field into its plane-wave components. Each plane wave carries a monochromatic signal that propagates a certain angle from the vertical. Events with the same dip in the *(t,x)* plane, regardless of location, are mapped onto a single line in the radial direction on the *(f-k)* plane. This is the basis for *f-k* dip filtering. It amounts to defining and applying a reject fan in the transform domain, then inverse transforming the data back to the *(t,x)* domain.

REFERENCES

Bracewell, R. N., 1965, The Fourier transform and its applications: McGraw-Hill Book Co.

Claerbout, J. F., 1976, Fundamentals of geophysical data processing: McGraw-Hill Book Co.

Grant, F. S. and West, G. F., 1965, Interpretation theory in applied geophysics: McGraw-Hill Book Co.

Larner, K. L., Chambers, R., Yang, M., Lynn, W., and Wai, W., 1983, Coherent noise in marine seismic data: Geophysics, **48**, 854-886.

Mayne, W. H., 1962, Common-reflection-point horizontal data stacking techniques: Geophysics, **27**, 927-938.

Newman, P., 1973, Divergence effects in a layered earth: Geophysics, **38**, 481-488.

Robinson, E. A. and Treitel, S., 1980, Geophysical signal analysis: Prentice-Hall, Inc.

Rothman, D., Larner, K. L., and Chambers, R., 1981, Trace interpolation and design of 3-D surveys: Presented at the Ann. Eur. Assoc. Explor. Geophys. Mtg.

Sengbush, R.L., 1983, Seismic exploration methods: Internat. Human Res. Dev. Corp.

Yilmaz, O. and Cumro, D., 1983, Worldwide assortment of field seismic records: Tech. Rep., Western Geophysical Company.

EXERCISES

Exercise 1-1. Using the hyperbolic traveltime equation, compute the average velocity down to reflector A in Record 8 (Figure 1-33). Assume a constant velocity between A and the surface. Table 1-8 (Record 8) contains all the information needed for your computations.

Exercise 1-2. Refer to the walk-away noise test record in Record 19 (Figure 1-33). Measure the phase velocity *(dx/dt)* of the ground roll energy at location A_1. Also measure dominant frequency at the same location. Then, estimate the dominant wavelength (velocity/dominant frequency) of the ground roll. The receiver array length needed to suppress this energy in the field should be equal to or greater than the longest noise wavelength. Table 1-8 (Record 19) contains all the information needed for your computations.

Exercise 1-3. Measure the group velocity *(xt)* of ground roll energy A in Record 25 (Figure 1-33). The required information is in Table 1-8 (Record 25).

Exercise 1-4. What is event A in Record 29 (Figure 1-33)? Are events C, D, and E multiples of B?

Exercise 1-5. Refer to Record 30 (Figure 1-33). Compute the water velocity using both the direct arrivals A and the water-bottom reflection B. Use one-way time for A and two-way time for B.

Exercise 1-6. Examine the traveltimes for water-bottom reflection B and its first multiple M in Record 30 (Figure 1-33). In particular, measure the time difference between the two, BM, at both the near and far traces. Are they equal? If not, account for the difference. (The answer to this exercise is the subject of Section 7.5.)

Exercise 1-7. Identify events A, B, C, D, E, and F in Record 33 (Figure 1-33).

Exercise 1-8. Identify events A, B, C, D, and E in Record 34 (Figure 1-33).

Exercise 1-9. Make inferences on the dips of events A and B in Record 35 (Figure 1-33).

Exercise 1-10. Consider the three primary stages in conventional processing: deconvolution, stacking, and migration. Place these three processes in order of importance and give reasons why.

Exercise 1-11. Compute fold n_f, midpoint spacing Δy, and CMP trace spacing Δf, for each of the following recording geometries:

Number of Channels	Shot Spacing, m	Receiver Spacing, m
Group 1		
6	25	25
6	50	25
6	75	25
6	100	25
6	200	25
Group 2		
6	25	25
6	25	50
6	25	75
6	25	100
6	25	200
Group 3		
4	25	25
6	25	25
8	25	25
10	25	25
12	25	25

Use a stacking chart as in Figure 1-43. Plot shot spacing Δs as a function of n_f, Δy, and Δf for group 1. Plot receiver spacing Δg as a function of n_f, Δy, and Δf for group 2. Finally, plot the number of channels n_c as a function of n_f, Δy, and Δf for group 3. Show that (a) fold can be increased by increasing the number of recording channels or by decreasing the shot spacing; (b) the midpoint interval can be made finer by decreasing the receiver spacing. The fold of coverage is given by $n_f = (n_g \Delta g)/(2 \Delta s)$, where n_g is the number of receiver channels. This relationship holds for any $(\Delta g/\Delta s)$ ratio for which the fold has an integer value less than or equal to $n_g/2$.

Exercise 1-12. Suppose that the shot associated with gather 1 in Figure 1-43 is missing. Identify the midpoints that are affected by this missing shot; i.e., the midpoints with a lower fold of coverage. Suppose the receiver associated with gather 2 in Figure 1-43 is missing. Identify the midpoints that are affected by it.

Exercise 1-13. From Figure 1-90, identify the common-shot and common-midpoint gathers.

Exercise 1-14. Under what circumstance would you prefer to assign the computed value for the instantaneous AGC gain function value to a time sample that precedes the center of the moving time gate?

Exercise 1-15. Consider an analog 200-Hz sinusoidal signal. If it were sampled at an 8-ms interval, what is its alias frequency?

Exercise 1-16. From the first arrivals in Record 27 (Figure 1-33), note the change in cable geometry at location A. What is the ratio of the receiver group intervals used to the left and right of A.

Exercise 1-17. Identify event E in Figure 1-81b.

Exercise 1-18. Gain application involves multiplying the gain function with the seismic trace (Figure 1-61). This process is equivalent to convolving the Fourier transform of the gain function with that of the input trace. Describe the effect of the gain application in the frequency domain.

Exercise 1-19. Refer to Figure 1-91. Which peaks on the amplitude spectra of the time series sampled at 4 and 8 ms correspond to peaks A, B, and C from the 2-ms sampled series?

Exercise 1-20. Consider the recording geometry in Figure 1-40. Sketch the traveltime curves on a common-shot gather associated with point scatterers (a) beneath the cable, (b) behind, and (c) in front of the cable. Assume all scatterers are on the plane of recording.

Exercise 1-21. What is event C in Record 8 (Figure 1-33)?

2

Deconvolution

2.1 INTRODUCTION

Deconvolution is a process that improves the temporal resolution of seismic data by compressing the basic seismic wavelet. Deconvolution normally is applied before stack; however, it also is common to apply deconvolution to stacked data. Figure 2-1 shows a stacked section with and without prestack deconvolution. Deconvolution has yielded a section with much higher temporal resolution. The ringy character of the stack without deconvolution limits resolution considerably.

The effect of deconvolution also is seen on data before stack. Figure 2-2 shows some CMP gathers from a marine line before and after deconvolution. Again, note that the prominent reflections stand out more distinctly on the deconvolved gathers. Deconvolution has removed a considerable amount of ringyness, while it has compressed the waveform at each of the prominent reflections. The stacked sections associated with these CMP gathers are shown in Figure 2-3. The improvement observed on the deconvolved CMP gathers also are noted on the corresponding stacked section.

Some NMO-corrected CMP gathers from a land line are shown in Figure 2-4. Note that on the gathers in Figure 2-4b and on the stacked section in Figure 2-5b, deconvolution has removed much of the reverberating energy and compressed the wavelet.

Deconvolution sometimes does more than just wavelet compression; it can remove a significant part of the multiple energy from the section. Note that the stacked section in Figure 2-6b, which was obtained from deconvolved CMP gathers, has been enhanced between 2 and 4 s. Unfortunately, deconvolution does not always work this well on multiples (Figure 2-7).

To understand deconvolution, first we need to examine the building blocks of a recorded seismic trace. The earth is composed of layers of rocks with different lithology and physical properties. Seismically, rock layers are defined by the densities and velocities with which seismic waves propagate through them. The product of density and velocity is called *seismic impedance*. The impedance contrast between adjacent rock layers causes the reflections that are recorded along a surface profile. Thus, the recorded seismogram can be modeled as a convolution of the earth's impulse response with the seismic wavelet. This wavelet has many components, including source signature, recording filter, surface reflections, and geophone response. The earth's impulse response is what would be recorded if the wavelet were just a spike. The impulse response comprises primary reflections (reflectivity series) and all possible multiples. Ideally, deconvolution should compress the wavelet components and eliminate multiples, leaving only the earth's reflectivity in the seismic trace.

After a general discussion on the convolutional model in Section 2.2, the concept of an inverse filter is introduced in Section 2.3. Basically, the inverse filter, when convolved with the seismic wavelet, converts it to a spike. When applied to a seismogram, the inverse filter should yield the earth's impulse response. The least-squares inverse filter is discussed in Section 2.4. The fundamental assumption underlying the deconvolution process (with the usual case of unknown source wavelet) is that of minimum phase. This issue is dealt with in Section 2.5. The optimum Wiener filter, which has a wide range of applications, is discussed in Section 2.6.

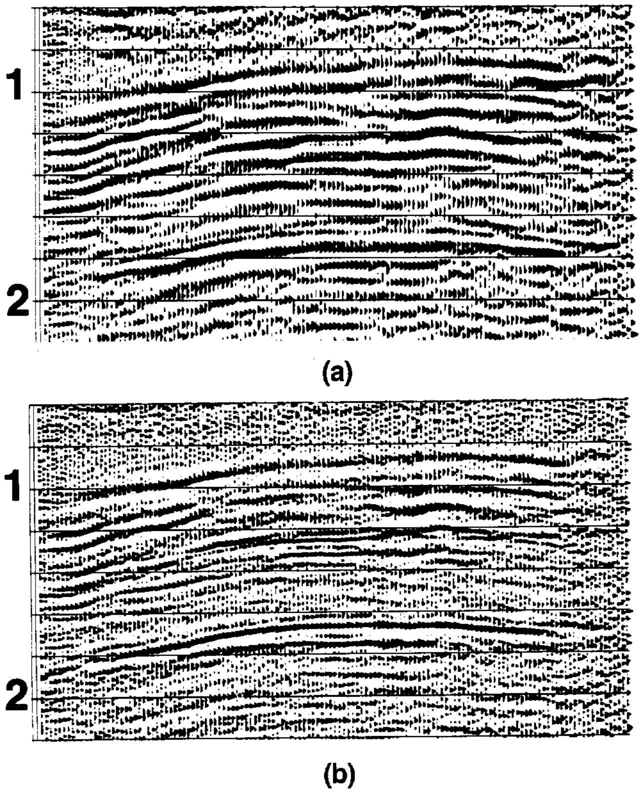


FIG. 2-1. Interpreters prefer the crisp, finely detailed appearance of the deconvolved section (b) as opposed to the blurred, ringy appearance of the section without deconvolution (a). (Data courtesy Phillips Petroleum.)

The Wiener filter converts the seismic wavelet into any desired shape. For example, much like the inverse filter, a Wiener filter can be designed to convert the seismic wavelet into a spike. However, the Wiener filter differs from the inverse filter in that it is optimal in the least-squares sense. Also, the resolution (spikiness) of the output can be controlled by designing a Wiener *prediction error filter*. Converting the seismic wavelet into a spike is like asking for a perfect resolution. In practice, because of noise in the seismogram and assumptions made about the seismic wavelet and the recorded seismogram, spiking deconvolution is not always desirable. Finally, the prediction error filter can be used to remove periodic components, i.e., multiples, from the seismogram.

The mathematical treatment of deconvolution is found in Appendix B. However, several numerical examples, which provide the theoretical groundwork from a heuristic viewpoint, are given in the text. Much of the early theoretical work on deconvolution came from the MIT Geophysical Analysis Group, which was formed in the mid-1950s.

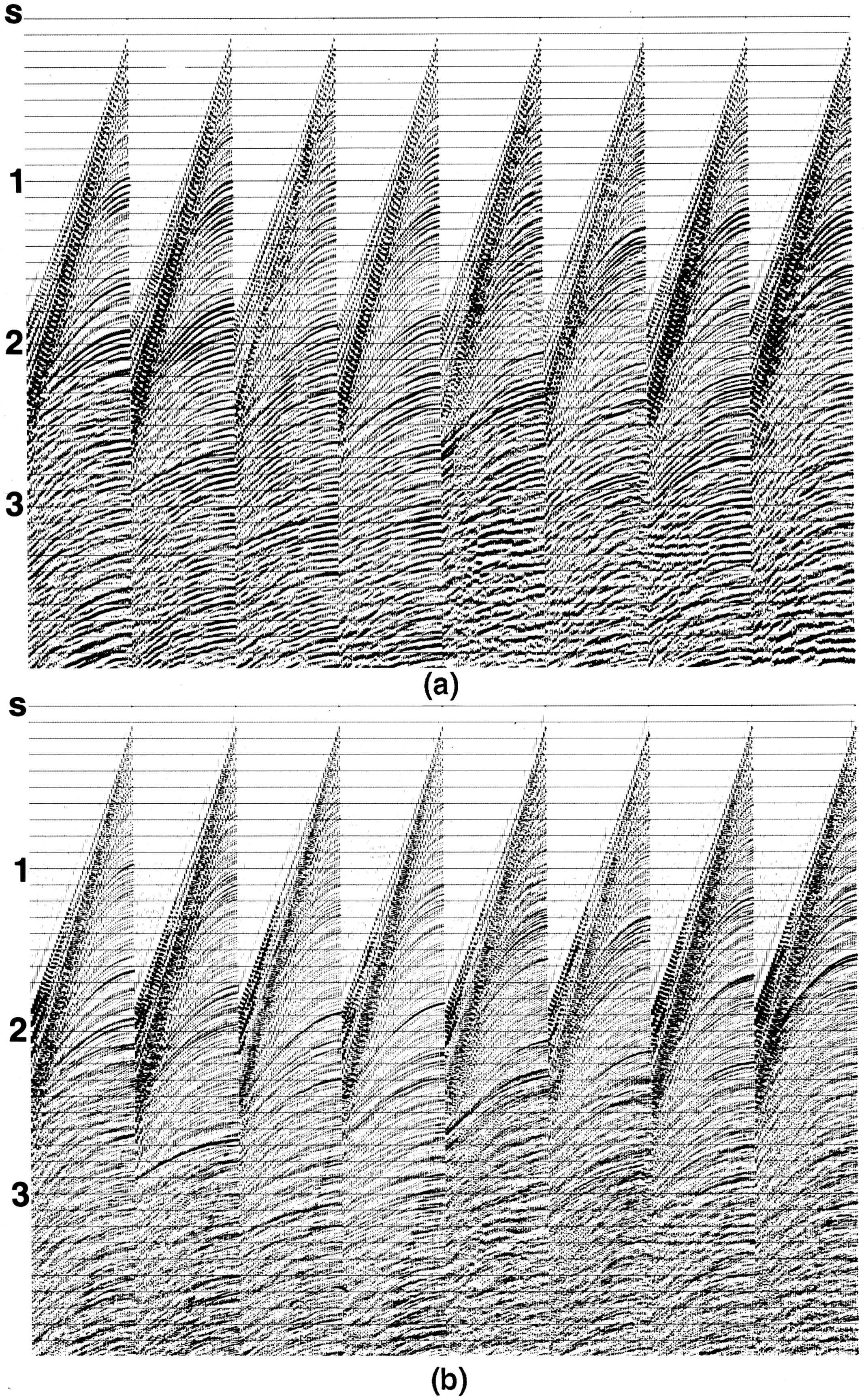


FIG. 2-2. Note the prominent reflections on the deconvolved gathers (b). The reverberating wavetrains would make it difficult to distinguish prominent reflections on the undeconvolved gathers (a).

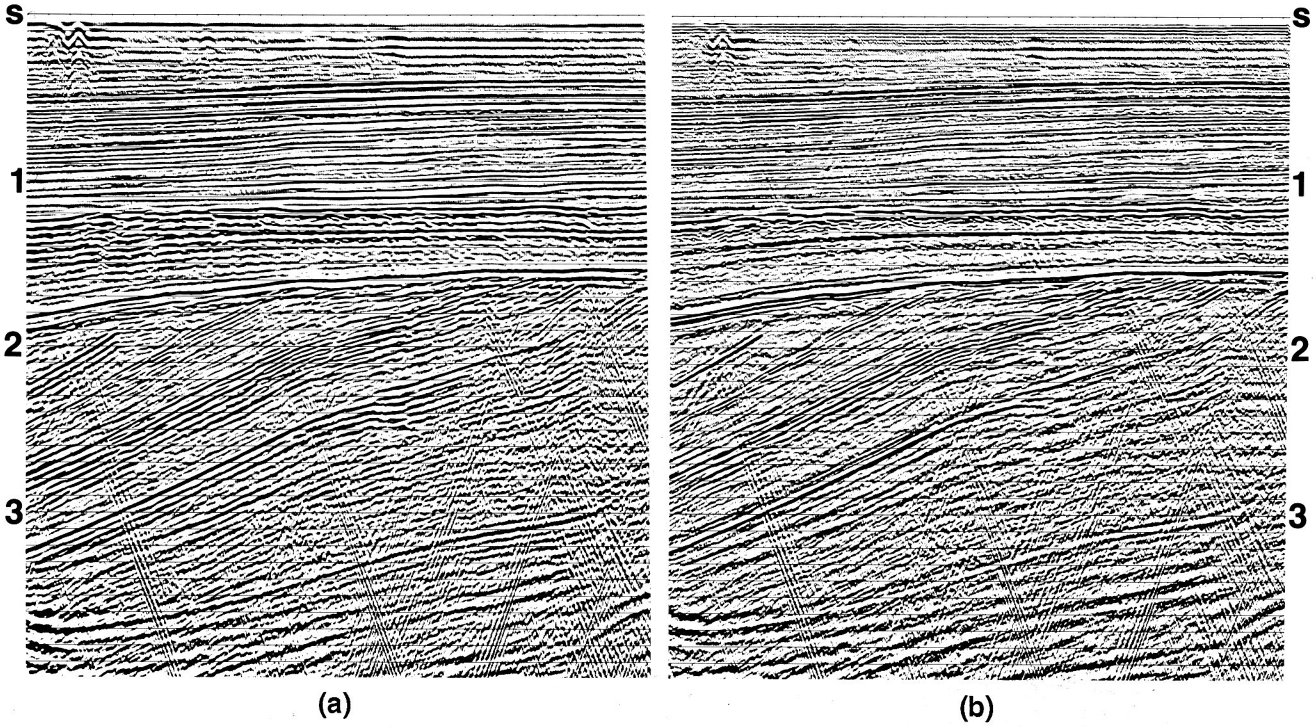


FIG. 2-3. With which stacked section would you prefer to work? (a) The section obtained from the undeconvolved gathers of Figure 2-2a, or (b) the section obtained from the deconvolved gathers of Figure 2-2b.

2.2 THE CONVOLUTIONAL MODEL

A sonic log segment is shown in Figure 2-8a. The sonic log is a plot of interval velocity as a function of depth. Here, velocities were measured between the 1000- to 5400-ft depth interval at 2-ft intervals. The velocity function was extrapolated to the surface by a linear ramp. The sonic log exhibits a strong low-frequency component with a distinct blocky character representing gross velocity variations. Actually, it is this low-frequency component that normally is estimated through velocity analysis of CMP gathers (Section 3.3).

In many sonic logs, the low-frequency component is an expression of the general increase of velocity with depth due to compaction. In some sonic logs, however, the low-frequency component exhibits a blocky character (Figure 2-8a), which is due to large-scale lithologic variations. Based on this blocky character, we may define layers of constant interval velocity (Table 2-1), each of which can be associated with a geologic formation (Table 2-2).

The sonic log also has a high-frequency component superimposed on the low-frequency component. These rapid fluctuations can be attributed to changes in rock properties that are local in nature. For example, the limestone layer can have interbeddings of shale and sand. Porosity changes also can affect interval velocities within a rock layer. Note that well-log measurements have a limited accuracy; therefore, some of the high-frequency variations, particularly those associated with a first arrival that is strong enough to trigger one receiver but not the other in the log tool (cycle skips), are not due to changes in lithology.

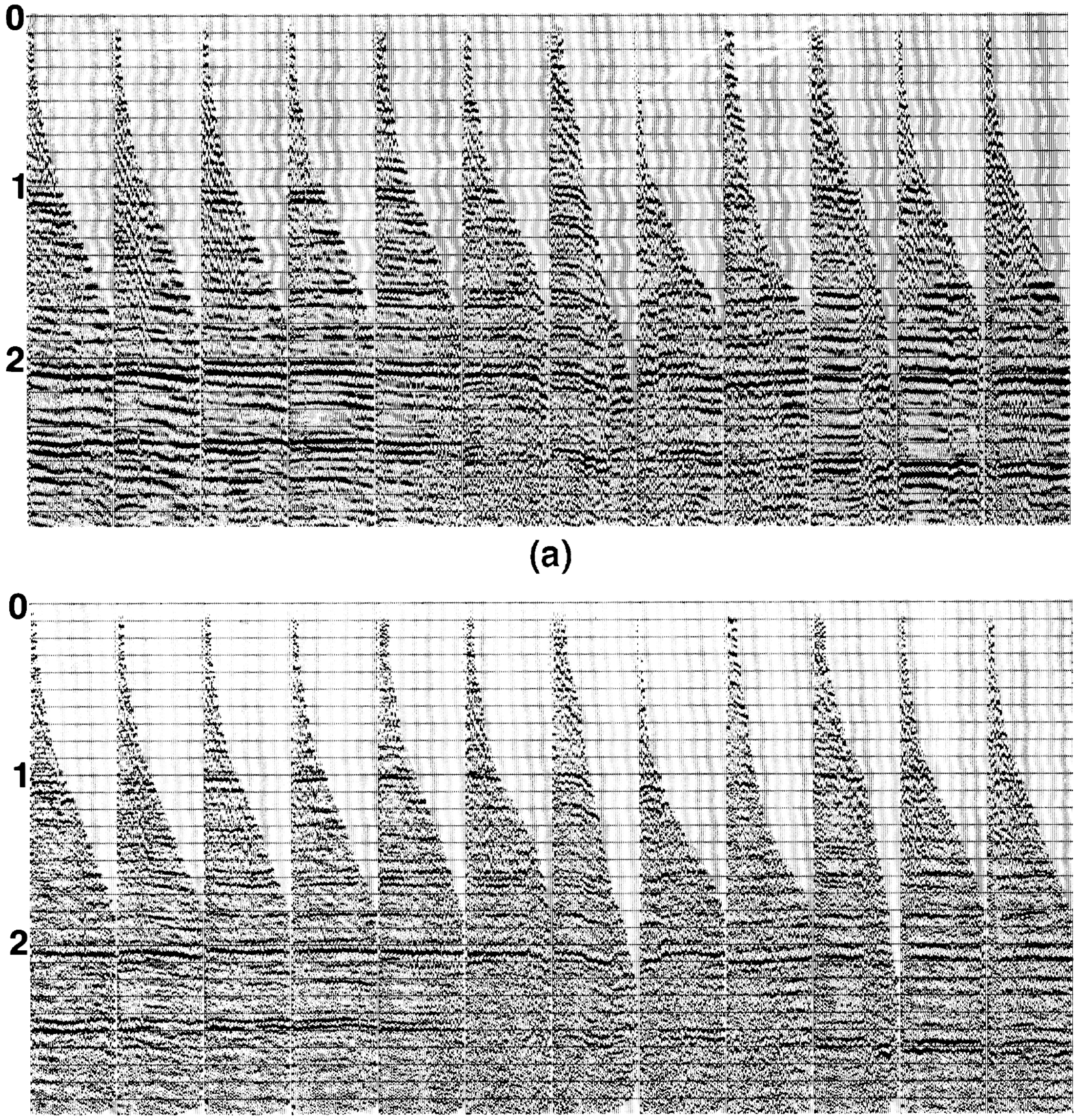
Well-log measurements of velocity and density provide a link between seismic data and the geology of the substrata. We now explain the relationship between log measurements and the recorded seismic trace. Seismic impedance is defined as the product of density and velocity. From well-log measurements, we find that the vertical density gradient is often much smaller than the vertical velocity gradient. Therefore, we often assume that the impedance contrast between rock layers is essentially due to the velocity contrast only. This assumption is not always the case. The reason we can get away with it is that the density gradient usually has the same sign as the velocity gradient. Hence, the impedance function derived from the velocity function only should be correct within a scale factor.

The first set of assumptions that is used to build the forward model for the seismic trace follows:

Assumption 1a. The earth is made up of horizontal layers of constant velocity.

Assumption 1b. The source generates a compressional plane wave that impinges on layer boundaries at normal incidence. Under such circumstances, no shear waves are generated.

Assumption 1a is violated in both structurally complex areas and in areas with gross lateral facies changes.



(a)

(b)

FIG. 2-4. Some NMO-corrected gathers associated with the stacked sections in Figure 2-5, (a) before, (b) after deconvolution. Deconvolution has removed the ringy character from the data.

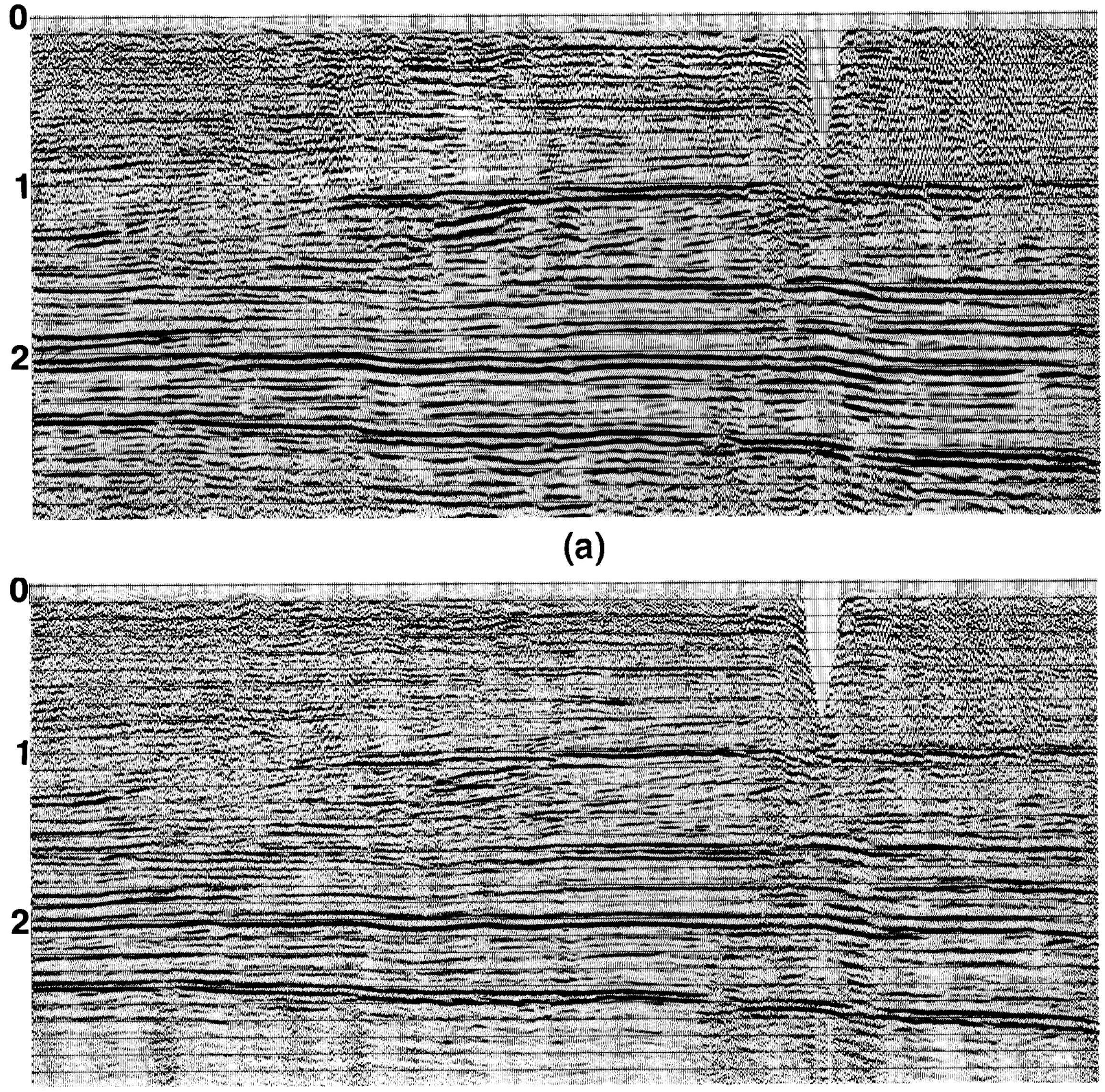


FIG. 2-5. Deconvolution helps distinguish prominent reflections with ease (b). However, on a section without deconvolution (a), reflections are buried in reverberating energy. Selected CMP gathers for both sections are shown in Figure 2-4.

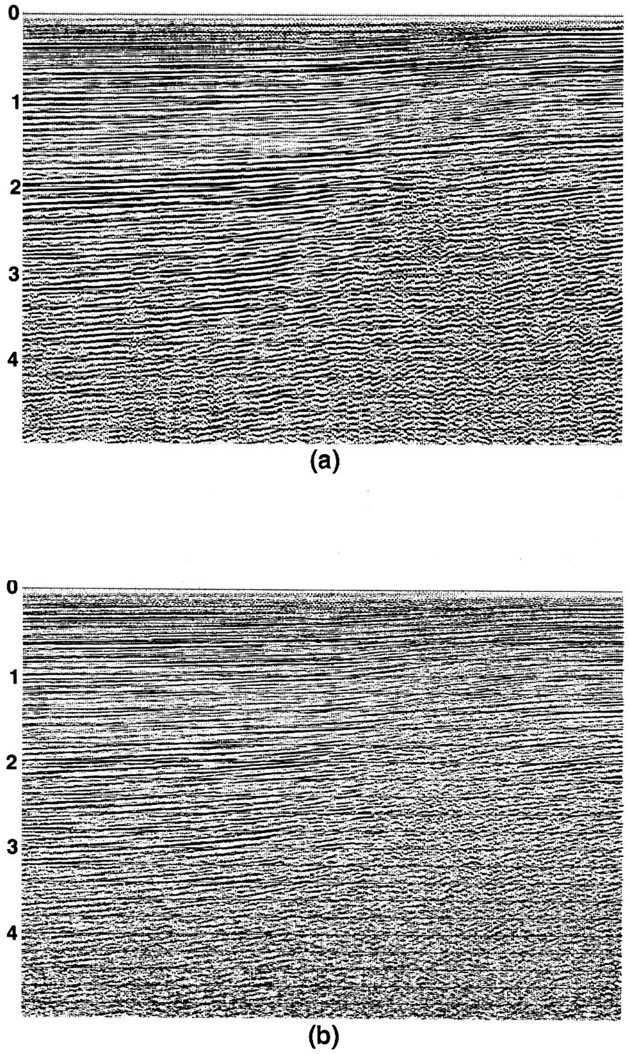


FIG. 2-6. CMP stacks with (a) no deconvolution before stack, (b) spiking deconvolution before stack. Deconvolution can remove a significant amount of multiple energy from seismic data. (Data courtesy Elf Aquitane and partners.)

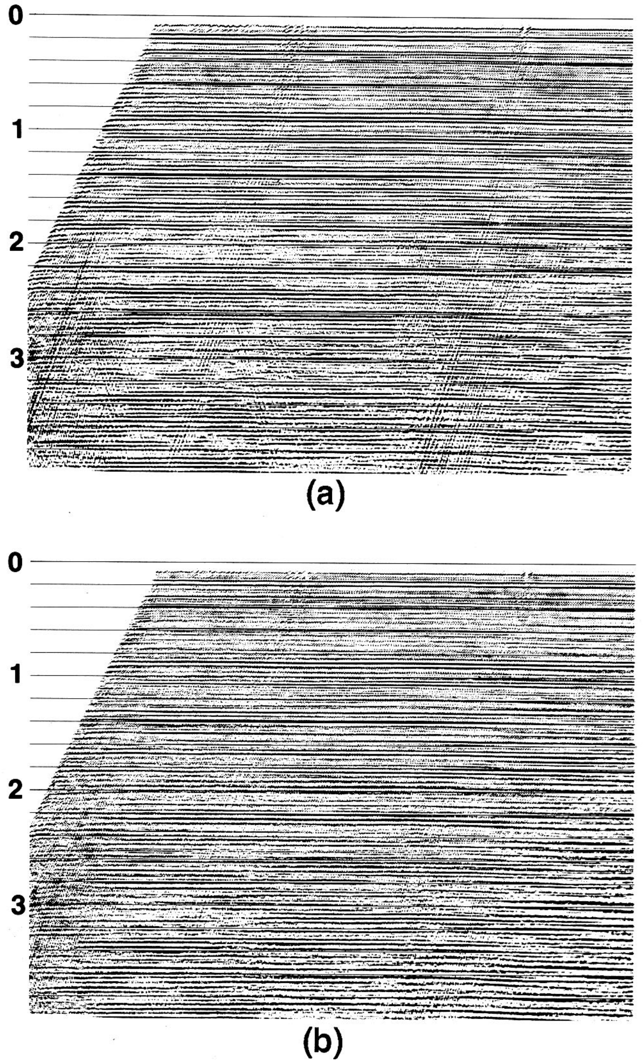


FIG. 2-7. CMP stacks with (a) no deconvolution before stack, (b) spiking deconvolution before stack. The CMP stack without deconvolution contains multiples of all types. Has deconvolution improved the section? (Data courtesy Deminex Petroleum.)

Table 2-1. The interval velocity trend obtained from the sonic log in Figure 2-8a.

Layer Number	Interval Velocity, ft/s	Depth Range, ft
1	21000	1000–2000
2	19000*	2000–2250
3	18750	2250–2500
4	12650	2500–3775
5	19650	3775–5400

* The velocity in this layer gradually decreases.

Table 2-2. Stratigraphic identification associated with layering described in Table 2-1.

Layer Number	Lithologic Unit
1	Limestone
2	Shaly limestone with gradual increase in shale content
3	Shaly limestone
4	Sandstone
5	Dolomite

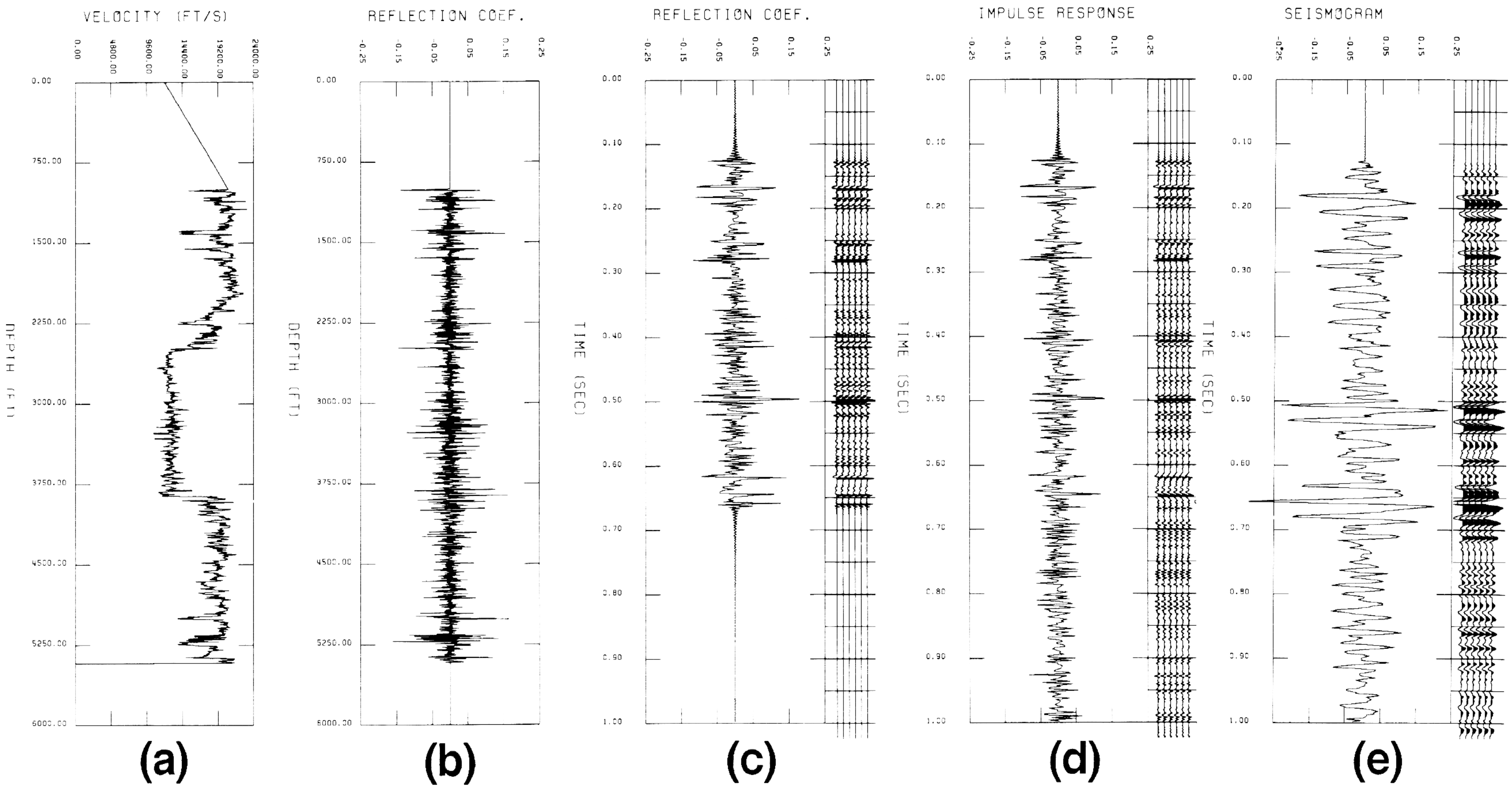


FIG. 2-8. (a) A segment of a measured sonic log, (b) the reflection coefficient series derived from (a), (c) the series in (b) after converting the depth axis to two-way time axis, (d) the impulse response that includes the primaries (c) and multiples, (e) the synthetic seismogram derived from (d) convolved with the source wavelet in Figure 2-11. One-dimensional seismic modeling means getting *e* from *a*. Deconvolution yields *d* and *c*, while 1-D inversion means getting *a* from *d*. Identify the event on *a* and *b* that corresponds to the big spike at 0.5 s in *c*. Impulse response *d* is a composite of the primaries *c* and all types of multiples.

Assumption 1b implies that zero offset data must be used. However, zero offset never is recorded. On the other hand, if the layer boundaries were deep in relation to cable length, we assume that the angle of incidence at a given boundary is small and ignore the angle dependence of reflection coefficients. Based on assumptions 1a and 1b, the reflection coefficient (for pressure or stress) c, which is associated with the boundary between, say, layers 1 and 2, is defined as $c = (I_2 - I_1)/(I_2 + I_1)$, where I is the seismic impedance associated with each layer given by the product of density ρ and velocity v.

If we also assume that density is invariant with depth or that it does not vary as much as velocity, then $c \approx (v_2 - v_1)/(v_2 + v_1)$. For vertical incidence, the reflection coefficient is the ratio of the reflected wave amplitude to the incident wave amplitude. Moreover, from its definition, the reflection coefficient is seen as the ratio of the change in acoustic impedance to twice the average acoustic impedance. If v_2 were greater than v_1, then the reflection coefficient would be positive. If v_2 were less than v_1, then the reflection coefficient would be negative. Reflection coefficient series $c(z)$, where z is the depth variable, is derived from sonic log $v(z)$ and is shown in Figure 2-8b. The position of each spike gives the depth of the layer boundary, while the magnitude of each spike corresponds to the fraction of a unit-amplitude downward-traveling incident plane wave that would be reflected from the layer boundary.

To convert the reflection coefficient series $c(z)$ (Figure 2-8b), derived from the sonic log into a time series $c(t)$, select a sampling interval, say 2 ms. Then use the velocity information in the log (Figure 2-8a) to convert the depth axis to a two-way vertical time axis. The result of this conversion is shown in Figure 2-8c, both as a conventional wiggle trace and as a variable area and wiggle trace (the same trace repeated six times to highlight strong reflections). The reflection coefficient series $c(t)$ represents the reflectivity of a series of fictitious layer boundaries that are separated by an equal time interval; namely, the *sampling rate*. This earth model, which has layers of equal traveltime (Goupillaud, 1961), is used since seismic waves are recorded in time. The major events in this reflectivity series are from the boundary between layers 2 and 3 located around 0.3 s, and the boundary between layers 4 and 5 located near 0.5 s.

So far, the reflection coefficient series (Figure 2-8c) that was constructed is composed only of *primary* reflections (energy that was reflected only once). To get a complete 1-D response of the horizontally layered earth model (assumption 1a), multiple reflections of all types (surface, intrabed and interbed multiples) must be included. If the source were unit-amplitude spike, then the recorded zero-offset seismogram would be the impulse response of the earth, which includes primary and multiple reflections. Here, the Kunetz method (Claerbout, 1976) is used to obtain such an impulse response. The impulse response derived from the reflection coefficient series in Figure 2-8c is shown in Figure 2-8d with the variable area and wiggle display.

The characteristic pressure wave created by an impulsive source, such as dynamite or air gun, is called the signature of the source. All signatures can be described as band-limited wavelets of finite duration; for example, the measured signature of an Aquapulse source in Figure 2-9. As this waveform travels into the earth, two things happen. First, its overall amplitude decays because of wavefront divergence. Second, frequencies are attenuated because of the absorption effects of rocks (see Section 1.5). The progressive change of the source wavelet in time and depth also is shown in Figure 2-9. At any given time, the wavelet is not the same as it was at the onset of source excitation. This time-dependent change in waveform is called *nonstationarity*. A compensation for nonstationarity usually is applied before deconvolution. Wavefront divergence is removed by applying a spherical spreading function (Section 1.5). Frequency attenuation is compensated for by the processing techniques discussed in Section 2.8. As a result, the simple convolutional model discussed here does not incorporate nonstationarity. This leads to the following assumption:

> Assumption 2. The source waveform does not change as it travels in the subsurface; i.e., it is stationary.

A convolutional model for the recorded seismogram now can be proposed. Suppose a vertically propagating downgoing plane wave with source signature (Figure 2-10a) travels in depth and encounters a layer boundary at 0.2-s two-way time. The reflection coefficient associated with the boundary is represented by the spike in Figure 2-10b. As a result of reflection, the source wavelet replicates itself such that it is scaled by the reflection coefficient. If we have a number of layer boundaries represented by the individual spikes in Figures 2-10d through 2-10f, then the wavelet replicates itself at those boundaries in the same

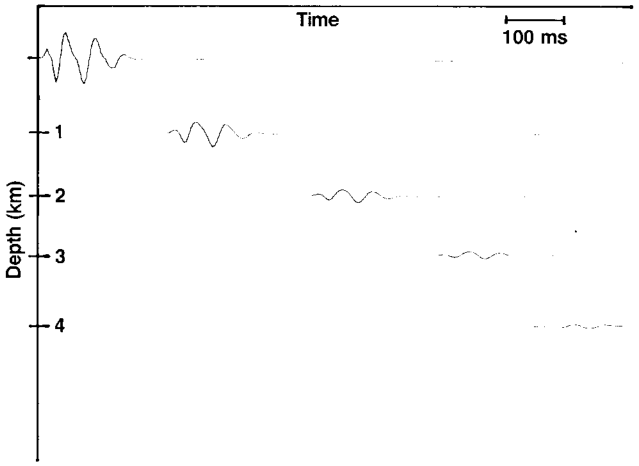


FIG. 2-9. A seismic source wavelet after onset takes the form shown at top left. As the wavelet travels into the earth, the amplitude level drops (geometric spreading) and a loss of high frequencies occurs (frequency absorption).

manner. If the reflection coefficient is negative, then the wavelet replicates itself with its polarity reversed, as in Figure 2-10e.

Now consider the ensemble of the reflection coefficients in Figure 2-10g. The response of this sparse spike series to the basic wavelet is a superposition of the individual impulse responses. This linear process is called the *principle of superposition*. It is achieved computationally by convolving the basic wavelet with the reflectivity series (Figure 2-10g). The convolutional process already was demonstrated by the numerical example in Section 1.2.3. The response of the sparse spike series to the basic wavelet in Figure 2-10g has some important characteristics. Note that for events at 0.2 and 0.35 s, we identify two layer boundaries. However, to identify the three closely spaced reflecting boundaries from the composite response (at around 0.6 s), the source waveform must be removed to obtain the sparse spike series. This removal process is just the opposite of the convolutional process used to obtain the response of the reflectivity series to the basic wavelet. The reverse process appropriately is called *deconvolution*.

The principle of superposition now is applied to the impulse response derived from the sonic log in Figure 2-8d. Convolution of a source signature with the impulse response yields the synthetic seismogram illustrated in

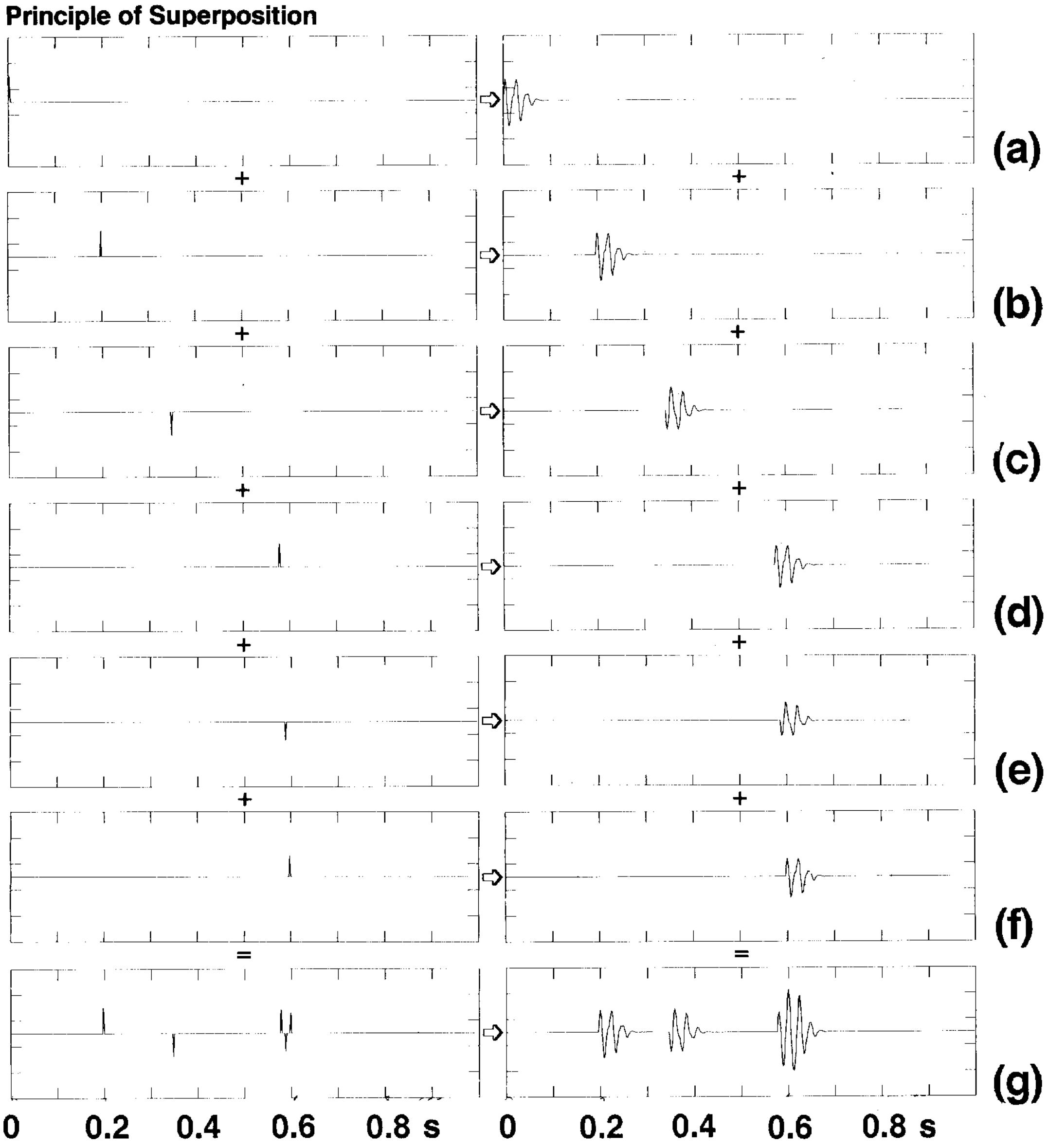


FIG. 2-10. A wavelet *a* traveling in the earth repeats itself when it encounters a reflector along its path (*b*, *c*, *d*, *e*, *f*). The left column represents the reflection coefficients, while the right column represents the response to the wavelet. Amplitudes of the response are scaled by the reflection coefficient. The resulting seismogram (bottom right) represents the composite response of the earth's reflectivity (bottom left) to the wavelet (top right).

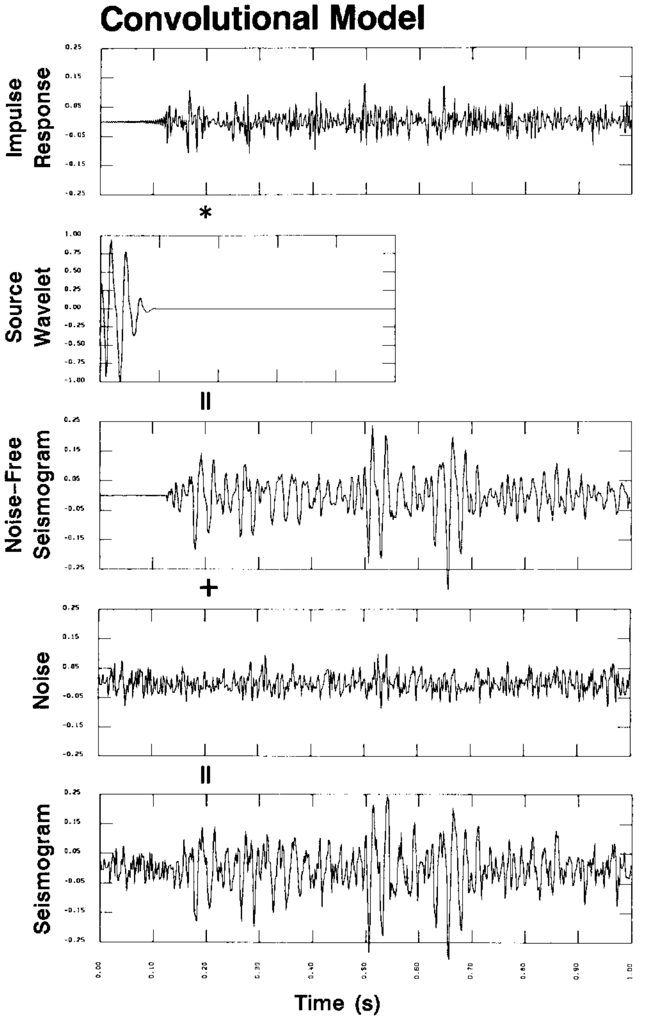


FIG. 2-11. The top frame is the same as in Figure 2-8d. The asterisk denotes convolution. The recorded seismogram (bottom frame) is the sum of the noise-free seismogram and the noise trace. This figure is equivalent to equation (2.1).

Figure 2-11. The synthetic seismogram also is shown in Figure 2-8e. This 1-D zero-offset seismogram is free of random ambient noise. For a more realistic representation of a recorded seismogram, noise is added (Figure 2-11). The *convolutional model* of the recorded seismogram now is complete. The building blocks of the convolutional model are described by

$$x(t) = w(t) * e(t) + n(t), \tag{2.1}$$

where

$x(t)$ = recorded seismogram
$w(t)$ = basic seismic wavelet
$e(t)$ = earth's impulse response
$n(t)$ = random ambient noise
$*$ = denotes convolution.

Equation (2.1) is the mathematical formulation of the convolutional model described in Figure 2-11. Deconvolution tries to recover the reflectivity series (strictly speaking, the impulse response) from the recorded seismogram.

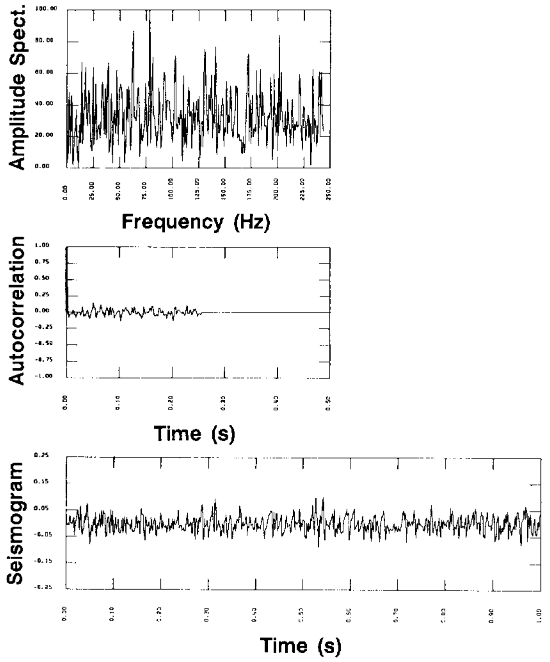


FIG. 2-12. A pure random series with infinite length has a flat amplitude spectrum and an autocorrelogram that is zero at all lags except the zero lag. What distinguishes a random signal from a spike (1,0,0,. . .)?

Alternatives to the convolutional model given by equation (2.1) do exist. One alternative that perhaps rigorously characterizes what goes on in the earth, when compared with the above formulation, is described in Appendix B.6. Nevertheless, the above formulation is the most accepted model for the 1-D seismogram.

The random noise present in the recorded seismogram has several sources. External sources are wind motion, environmental noise, or a geophone loosely coupled to the ground. Internal noise can arise from the recording instruments. A pure noise seismogram and its characteristics are shown in Figure 2-12. A pure random noise series has a white spectrum, that is, it contains all the frequencies. In turn, this implies that the autocorrelation function is a spike at zero lag and zero at all other lags. From Figure 2-12, note that these characteristic requirements are reasonably satisfied.

Now examine the equation for the convolutional model. All that normally is known in equation (2.1) is $x(t)$, the recorded seismogram. The earth's impulse response $e(t)$ must be estimated everywhere except at the location of wells with good sonic logs. The source waveform $w(t)$ also normally is unknown. In certain cases, however, the source waveform is partly known; for example, the signature of an air-gun array can be measured. However, what is measured is only the waveform at the very onset of excitation of the source array, and not the wavelet that is recorded at the receiver. Finally, there is no a priori

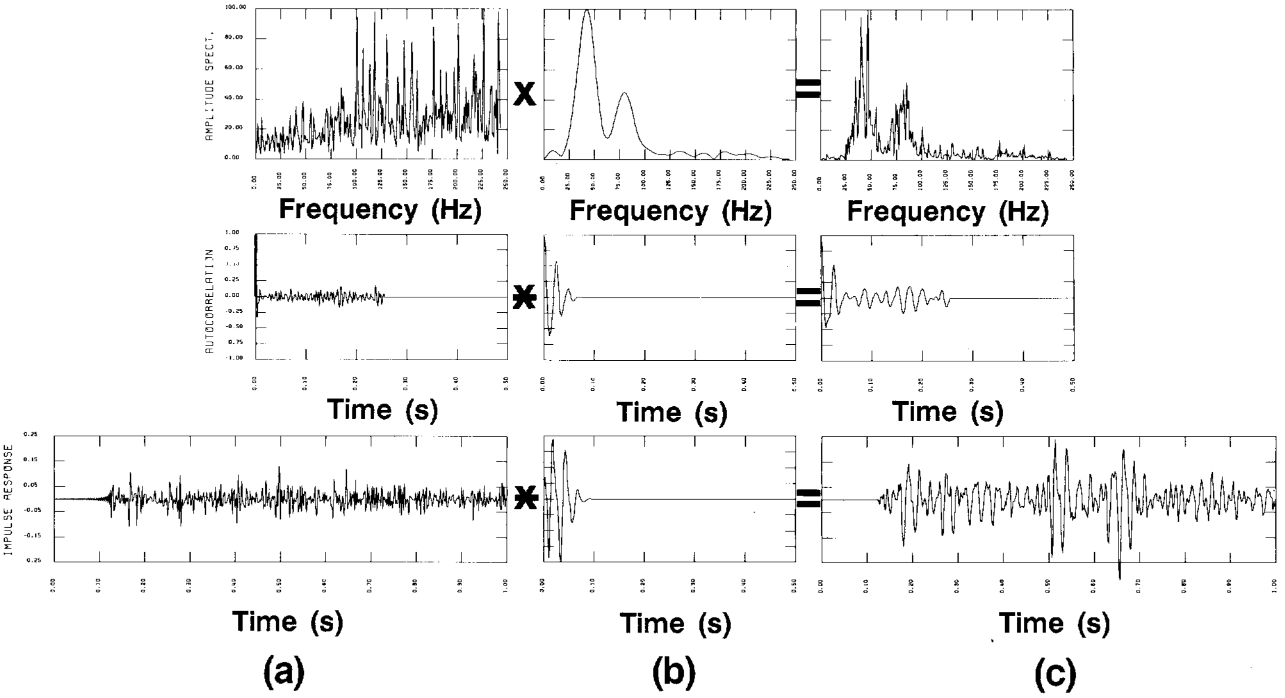


FIG. 2-13. Convolution of the earth's impulse response (a) with the wavelet (b) (equation 2.2) yields the seismogram (c) (bottom row). This process also is convolutional in terms of their autocorrelograms (middle row) and multiplicative in terms of their amplitude spectra (top row). Assumption 5 (white reflectivity) is based on the similarity between autocorrelograms and amplitude spectra of the impulse response and wavelet.

knowledge of the ambient noise $n(t)$. We now have three unknowns, one known, and one single equation [equation (2.1)]. Can this problem be solved? Pessimists would say no. However, in practice, deconvolution is applied to seismic data as an integral part of conventional processing. Furthermore, over the years, deconvolution has become an effective method to increase temporal resolution. To solve for the unknown $e(t)$, further assumptions must be made.

Assumption 3. The noise component $n(t)$ is zero.
Assumption 4. The source waveform is known.

Under these assumptions, we have one equation, equation (2.1), and one unknown. In reality, however, neither of the above two assumptions normally is valid. Therefore, the convolutional model is examined further in the next section, this time in the frequency domain, to relax assumption 4. If the source waveform were known (such as the recorded source signature), then the solution to the deconvolution problem is *deterministic*. In Section 2.3, one such method of solving for $e(t)$ is considered. If the source waveform were unknown (the usual case), then the solution to the deconvolution problem is *statistical*. The Wiener prediction theory (Section 2.6) provides one method of statistical deconvolution.

2.2.1 The Convolutional Model in the Frequency Domain

The convolutional model for the noise-free seismogram (assumption 3) is represented by the following equation:

$$x(t) = w(t) * e(t). \tag{2.2}$$

Convolution in the time domain is equivalent to multiplication in the frequency domain (Section 1.2.3). In particular, the amplitude spectrum of the seismogram equals the product of the amplitude spectra of the seismic wavelet and the earth's impulse response [equation (B.5)]. Figure 2-13 shows the amplitude spectra (top row) of the impulse response $e(t)$, the seismic wavelet $w(t)$, and the seismogram $x(t)$. The impulse response is the same as that shown in Figure 2-8d. The similarity in the overall shape between the amplitude spectrum of the wavelet and that of the seismogram is apparent. In fact, a smoothed version of the amplitude spectrum of the seismogram is nearly indistinguishable from the amplitude spectrum of the wavelet. It generally is thought that the rapid fluctuations observed in the amplitude spectrum of a seismogram are a manifestation of the earth's impulse response; while the basic shape primarily is associated with the source wavelet.

Mathematically, the similarity between the amplitude spectra of the seismogram and the wavelet suggests that the amplitude spectrum of the earth's impulse response must be nearly flat (see Appendix B.1). By examining the amplitude spectrum of the impulse response in Figure 2-13, we see that it spans virtually the entire spectral bandwidth. As seen in Figure 2-12, a time series that represents a random process has a flat (white) spectrum over the entire spectral bandwidth. From close examination of the impulse response amplitude spectrum in

Figure 2-13, we see that it is not entirely flat; the high-frequency components have a tendency to gradually strengthen. Thus, reflectivity is not entirely a random process. In fact, this has been observed in the spectral properties of reflectivity functions derived from a worldwide selection of sonic logs (Walden and Hosken, 1984).

We now study the autocorrelation functions (middle row, Figure 2-13) of the impulse response, seismic wavelet, and synthetic seismogram. Note that the autocorrelation functions of the basic wavelet and seismogram also are similar. This similarity is confined to lags for which the autocorrelation of the wavelet is nonzero. Mathematically, the similarity between the autocorrelograms of the wavelet and the seismogram suggests that the impulse response has an autocorrelation function that is small at all lags except the zero lag (see Appendix B.1). The autocorrelation function of the random series in Figure 2-12 also has similar characteristics. However, there is one subtle difference. When compared, Figures 2-12 and 2-13 show that autocorrelation of the impulse response has a significantly large negative lag value following the zero lag. This is not the case for autocorrelation of random noise. The positive peak (zero lag) followed by the smaller negative peak in the autocorrelogram of the impulse response arises from the spectral behavior discussed above. In particular, the positive peak and the adjacent, smaller negative peak of the autocorrelogram together nearly act as a fractional derivative operator, which has a ramp effect in the amplitude spectrum (see Appendix A).

The above observations made on the amplitude spectra and autocorrelation functions (Figure 2-13) imply that reflectivity is not entirely a random process. Nonetheless, the following assumption almost always is made about reflectivity to eliminate assumption 4.

Assumption 5. Reflectivity is a random process. This implies that the seismogram has the characteristics of the seismic wavelet in that their autocorrelations and amplitude spectra are similar.

This assumption is the key in implementing the predictive deconvolution. It allows the autocorrelation of the seismogram, which is known, to be substituted for the autocorrelation of the seismic wavelet, which is unknown. In Section 2.6, we see that as a result of assumption 5, an inverse filter can be estimated directly from the autocorrelation of the seismogram. For this type of deconvolution, Assumption 4, which is almost never met in reality, is not required.

2.3 INVERSE FILTERING

If a filter operator *a(t)* were defined such that convolution of *a(t)* with the known seismogram *x(t)* yields an estimate of the earth's impulse response *e(t)*, then

$$e(t) = a(t) * x(t). \tag{2.3}$$

By substituting equation (2.3) into equation (2.2), we get

$$x(t) = w(t) * a(t) * x(t). \tag{2.4}$$

When *x(t)* is eliminated from both sides of the equation, the following expression results:

$$\delta(t) = w(t) * a(t), \tag{2.5}$$

where $\delta(t)$ represents the Kronecker delta function

$$\delta(t) = \begin{cases} 1, & t = 0 \\ 0, & \text{elsewhere.} \end{cases} \tag{2.6}$$

By solving equation (2.5) for the filter operator *a(t)*, we obtain

$$a(t) = \delta(t) * w'(t), \tag{2.7}$$

where *w'(t)* is the inverse of the seismic wavelet *w(t)*, which is assumed to be known for the moment. Therefore, the filter operator needed to compute the earth's impulse response from the recorded seismogram turns out to be the mathematical inverse of the seismic wavelet. Equation (2.5) implies that the inverse filter converts the basic wavelet to a spike at $t = 0$. Likewise, the inverse converts the seismogram to a series of spikes that defines the earth's impulse response. Therefore, inverse filtering is a method of deconvolution, provided the source waveform is known (deterministic deconvolution). The steps in inverse filtering are summarized in Figure 2-14.

How do we compute the inverse of the seismic wavelet? This is accomplished mathematically by using the *z*-transform. For example, let the basic wavelet be a two-

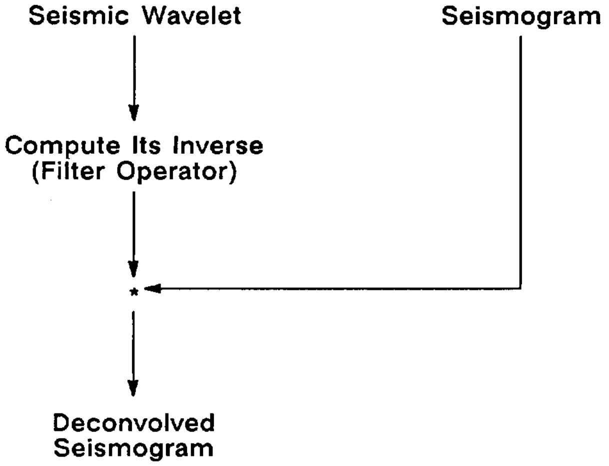


FIG. 2-14. A flowchart for inverse filtering.

point time series given by $(1, -\frac{1}{2})$. The z-transform of this wavelet is defined by the following polynomial:

$$W(z) = 1 - (\tfrac{1}{2})z. \tag{2.8}$$

The power of variable z is the number of unit time delays associated with each sample in the series. The first term has zero delay, so z is raised to zero power. The second term has unit delay, so z is raised to first power. Hence, the z-transform of a time series is a polynomial in z, whose coefficients are the values of the time samples.

A natural relationship exists between the z-transform and the Fourier transform (Appendix A). The z-variable is defined as

$$z = e^{-i\omega\Delta t},$$

where ω is angular frequency and Δt is sampling interval.

The inverse of the wavelet $w'(t)$ is obtained by polynomial division of the z-transform [equation (2.8)]:

$$W'(z) = 1/[1 - (\tfrac{1}{2})z] = 1 + (\tfrac{1}{2})z + (\tfrac{1}{4})z^2 + \cdots \tag{2.9}$$

The inverse time series is the coefficients of $W'(z)$; i.e., $[1,(\frac{1}{2}), (\frac{1}{4}), \ldots]$. This is also the filter operator $a(t)$. Note that it has an infinite number of coefficients, although they decay rapidly. As in any filtering process, in practice the operator is truncated.

First consider a two-point operator $[1,(\frac{1}{2})]$. Convolution of this operator with the wavelet as shown in Table 2-3 yields the result $[1,0, (-\frac{1}{4})]$. The ideal result is a zero-delay spike (1,0,0). Although not ideal, the actual result is spikier than the input wavelet, $[1,(-\frac{1}{2})]$.

Can the result be improved by including one more coefficient in the inverse filter as shown in Table 2-4? The output from the three-point filter is $[1,0,0,(-\frac{1}{8})]$. This is a more accurate representation of the desired output (1,0,0,0) than that achieved with the output from the two-point filter. Note that there is less energy leaking into the nonzero lags of the output from the three-point filter. Therefore, it is spikier. As more terms are included in the inverse filter, the output is closer to being a spike at zero lag. Since the number of points allowed in the operator length is limited, in practice the result never is a perfect spike.

Table 2-3. Convolution of the truncated inverse filter $[1, (\frac{1}{2})]$ with the input wavelet $[1,(-\frac{1}{2})]$.

	1	$-\frac{1}{2}$		Output
$\frac{1}{2}$	1			1
	$\frac{1}{2}$	1		0
		$\frac{1}{2}$	1	$-\frac{1}{4}$

Table 2-4. Convolution of the truncated inverse filter $[1,(\frac{1}{2}),(\frac{1}{4})]$ with the input wavelet $[1,(-\frac{1}{2})]$.

		1	$-\frac{1}{2}$			Output
$\frac{1}{4}$	$\frac{1}{2}$	1				1
	$\frac{1}{4}$	$\frac{1}{2}$	1			0
		$\frac{1}{4}$	$\frac{1}{2}$	1		0
			$\frac{1}{4}$	$\frac{1}{2}$	1	$-\frac{1}{8}$

The inverse of the input wavelet $[1,(-\frac{1}{2})]$ has coefficients that rapidly decay to zero. What about the inverse of the input wavelet $[(-\frac{1}{2}),1]$? Here, polynomial division gives the divergent series $(-2,-4,-8, \ldots)$. Truncate this series and convolve the two-point operator with the input wavelet $[(-\frac{1}{2}),1]$ as shown in Table 2-5.

The actual output is $(1,0,-4)$, while the desired output is (1,0,0). Not only is the result far from the desired output, but also it is less spiky than the input wavelet $[(-\frac{1}{2}),1]$. The reason for this poor result is that the inverse filter coefficients increase in time rather than decay. When truncated, the larger coefficients actually are excluded from the computation. If we kept the coefficient -8 in the above example, then the actual output would be $(1,0,0,-8)$, which is also a bad approximation to the desired output (1,0,0,0).

2.4 LEAST-SQUARES INVERSE FILTERING

When the input wavelet is well-behaved, such as $(1,-\frac{1}{2})$ as opposed to $(-\frac{1}{2},1)$, then the inverse filtering described in Section 2.3 yields a good approximation to a spike output. Can we do even better than that? Formulate the following problem: Given the input wavelet $[1,(-\frac{1}{2})]$, find a two-term filter (a,b) such that the error between the actual output and the desired output (1,0,0) is minimum in the least-squares sense. Compute the actual output by convolving the filter (a,b) with the input wavelet $[1,(-\frac{1}{2})]$. Table 2-6 shows convolution of the filter (a,b) with the input wavelet $[1,(-\frac{1}{2})]$.

The cumulative energy of the error L is defined as the sum of the squares of the differences between the coefficients of the actual and desired outputs:

$$L = (a - 1)^2 + (b - a/2)^2 + (-b/2)^2. \tag{2.10}$$

The task is to find coefficients (a,b) so that L takes its minimum value. This requires variation of L with respect to the coefficients (a,b) to vanish (Appendix B.5). By simplifying equation (2.10), taking the partial derivatives of quantity L with respect to a and b, and setting the results to zero, we get

$$(\tfrac{5}{2})a - b = 2, \tag{2.11a}$$

and

$$(\tfrac{5}{2})b - a = 0. \tag{2.11b}$$

We have two equations and two unknowns; namely, the filter coefficients (a,b). The so-called normal set of equations [equations (2.11a) and (2.11b)] can be put into the

Table 2-5. Convolution of the truncated inverse filter $[(-2,-4)]$ with input wavelet $[(-\frac{1}{2}),1]$.

	$-\frac{1}{2}$	1		Output
-4	-2			1
	-4	-2		0
		-4	-2	-4

Table 2-6. Convolution of filter (a,b) with input wavelet $[1,(-\frac{1}{2})]$.

	1	$-\frac{1}{2}$		Actual Output	Desired Output
b	a			a	1
	b	a		$b - a/2$	0
		b	a	$-b/2$	0

following convenient matrix form:

$$\begin{bmatrix} \frac{5}{2} & -1 \\ -1 & \frac{5}{2} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}. \tag{2.12}$$

By solving for the filter coefficients, we obtain $(a,b) = [(\frac{20}{21}), (\frac{8}{21})]$. This filter is applied to the input wavelet as shown in Table 2-7.

To quantify the spikiness of this result and compare it with the result from the inverse filter in Table 2-3, compute the energy of the errors made in both (Table 2-8). The least-squares filter yields less error when trying to convert the input wavelet $[1,(-\frac{1}{2})]$ to a spike at zero lag (1,0,0).

We now examine the performance of the least-squares filter with the input wavelet $[(-\frac{1}{2}),1]$. Note that the inverse filter produced unstable results for this wavelet (Table 2-5). We want to find a two-term filter (a,b) that, when convolved with the input wavelet $[(-\frac{1}{2}),1]$, yields an estimate of the desired spike output (1,0,0). As before, the least-squares error between the actual output and the desired output should be minimal (Table 2-9).

The cumulative energy of the error is given by:

$$L = (-a/2 - 1)^2 + (-b/2 + a)^2 + b^2. \tag{2.13}$$

By simplifying, taking the partial derivatives of quantity L with respect to a and b, and setting the results to zero, we get:

$$(\tfrac{5}{2})a - b = -1, \tag{2.14a}$$

and

$$(\tfrac{5}{2})b - a = 0. \tag{2.14b}$$

In matrix form:

$$\begin{bmatrix} \frac{5}{2} & -1 \\ -1 & \frac{5}{2} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}. \tag{2.15}$$

Table 2-7. Convolution of input wavelet $[1,(-\frac{1}{2})]$ with filter coefficients $[(\frac{20}{21}), (\frac{8}{21})]$

	1	$-\frac{1}{2}$		Output
$\frac{8}{21}$	$\frac{20}{21}$			$\frac{20}{21}$
	$\frac{8}{21}$	$\frac{20}{21}$		$-\frac{2}{21}$
		$\frac{8}{21}$	$\frac{20}{21}$	$-\frac{4}{21}$

Table 2-8. Errors made in two-term inverse and least-squares filtering.

Input: $[1,(-\frac{1}{2})]$
Desired Output: (1,0,0)

	Actual Output			Energy of Error
Inverse Filter	1	0	$-\frac{1}{4}$	$\frac{1}{16}$
Least-Squares Filter	$\frac{20}{21}$	$-\frac{2}{21}$	$-\frac{4}{21}$	$\frac{1}{21}$

By solving for the filter coefficients, $(a,b) = [(-\frac{10}{21}), (-\frac{4}{21})]$ results. Table 2-10 shows this filter applied to the input wavelet.

Table 2-11 shows the results from the inverse filter and least-squares filter quantified. The error made by the least- squares filter is, again, much less than the error made by the truncated inverse filter. However, both filters yield larger errors for input wavelet $[(-\frac{1}{2}),1]$ (Table 2-11) as compared to errors for wavelet $[1,(-\frac{1}{2})]$ (Table 2-8). The reason for this is discussed in Section 2.5.

2.5 MINIMUM PHASE

Two input wavelets, wavelet 1: $[1,(-\frac{1}{2})]$ and wavelet 2: $[(-\frac{1}{2}),1]$, were used for numerical analyses of the inverse filter and least-squares inverse filter in the preceding two sections. The results indicated that the error in converting wavelet 1 into a zero-lag spike is less than the error in converting wavelet 2 (see Tables 2-8 and 2-11). Wavelet 1 is closer to being a zero-delay spike (1,0,0) than wavelet 2. On the other hand, wavelet 2 is closer to being a delayed spike (0,1,0) than wavelet 1. We conclude that the error is reduced if the desired output closely resembles the energy distribution in the input series. Wavelet 1 has more energy at the onset, while wavelet 2 has more energy concentrated at the end.

Figure 2-15 shows three wavelets with the same amplitude spectrum, but with different phase-lag spectra. As a result, their shapes differ. The wavelet on top has more energy concentrated at the onset, the wavelet in the middle has its energy concentrated at the center, and the wavelet at the bottom has most of its energy concentrated at the end. From Section 1.2.2, we know that the shape of a wavelet can be altered by changing the phase spectrum without modifying the amplitude spectrum.

Table 2-9. Convolution of filter (a,b) with input wavelet $[(-\frac{1}{2}),1]$.

	$-\frac{1}{2}$	1		Actual Output	Desired Output
b	a			$-\frac{a}{2}$	1
	b	a		$-\frac{b}{2} + a$	0
		b	a	b	0

Table 2-10. Convolution of input wavelet $[(-\frac{1}{2}), 1]$ with filter coefficients $[(-\frac{10}{21}), (-\frac{4}{21})]$.

	$-\frac{1}{2}$	1	Output
$-\frac{4}{21}$	$-\frac{10}{21}$		$\frac{5}{21}$
	$-\frac{4}{21}$	$-\frac{10}{21}$	$-\frac{8}{21}$
	$-\frac{4}{21}$	$-\frac{10}{21}$	$-\frac{4}{21}$

Table 2-11. Errors made in the two-term inverse and least-squares filtering.

Input: $[(-\frac{1}{2}),1]$
Desired Output: (1,0,0)

	Actual Output			Energy of Error
Inverse Filter	1	0	−4	16
Least-Squares Filter	$\frac{5}{21}$	$-\frac{8}{21}$	$-\frac{4}{21}$	$\frac{336}{441}$

Table 2-12. Cumulative energy.

	Cumulative Energy at Time Sample		
Wavelet	0	1	2
A	16	16	17
B	4	13	17
C	4	13	17
D	1	1	17

We say that a wavelet is *minimum phase* if its energy is maximally concentrated at its onset. Similarly, a wavelet is *maximum phase* if its energy is maximally concentrated at its end. Finally, in all in-between situations, the wavelet is *mixed phase*. Note that a *wavelet* is defined as a transient waveform with a finite duration; i.e., it is *realizable*. A minimum-phase wavelet is one-sided in that it is zero before $t = 0$. A wavelet that is zero for $t<0$ is called *causal*. These definitions are consistent with intuition in that physical systems respond to an excitation only after that excitation. Their response also is of finite duration. In summary, a minimum-phase wavelet is *realizable* and *causal*. These observations are quantified by considering the following four, three-point wavelets (Robinson, 1966):

Wavelet A: (4,0,−1)
Wavelet B: (2,3,−2)
Wavelet C: (−2,3,2)
Wavelet D: (−1,0,4)

Now compute the cumulative energy of each wavelet at any one time. Cumulative energy is computed by adding squared amplitudes as shown in Table 2-12. These values are plotted in Figure 2-16. Note that all four wavelets have the same amount of total energy; i.e., 17 units. However, the rate at which the energy builds up is significantly different for each wavelet. For example, with wavelet A, the energy builds up rapidly close to its total value at the very first time lag. The energy for wavelets B and C builds up relatively slowly. Finally, the energy accumulates at the slowest rate for wavelet D. From Figure 2-16, note that the energy curves for wavelets A and D form the upper and lower boundaries. Wavelet A has the *least energy delay*, while wavelet D has the *largest energy delay*. Given a fixed amplitude spectrum as in Figure 2-17, the wavelet with the least energy delay is called *minimum delay*, while the wavelet with the largest energy delay is called *maximum delay*. This is the basis for Robinson's energy delay theorem: *A minimum-phase wavelet has the least energy delay.*

Time delay is equivalent to a phase-lag. Figure 2-18 shows the phase spectra of the four wavelets (Robinson, 1966). Note that wavelet A has the least phase change across the frequency axis; we say it is minimum phase. Wavelet D has the largest phase change; we say it is maximum phase. Finally, wavelets B and C have phase changes between the two extremes; hence, they are mixed phase. Since all four wavelets have the same amplitude spectrum (Figure 2-17) and the same power spectrum, they should have the same autocorrelation. This is verified as shown in Table 2-13, where only one side of the autocorrelation is tabulated, since a real time series has a symmetric autocorrelation (Section 1.2.3).

Note that zero lag of the autocorrelation is equal to the total energy contained in each wavelet; i.e., 17 units. This is true for any wavelet; that is, the area under the power spectrum is equal to the zero-lag value of the autocorrelation function. This is rigorously stated by Parseval's theorem (Appendix A).

The process by which the seismic wavelet is compressed into a zero-lag spike is called *spiking deconvolution*. In Sections 2.3 and 2.4, filters that achieve this goal were studied. Their performance depends not only on filter length (compare the results of the two-and three-point filters in Section 2.3), but also on whether the input wavelet is minimum phase. The spiking deconvolution operator is strictly the inverse of the wavelet. If the wavelet were minimum phase, then we would get a stable inverse, which also is minimum phase. The term *stable* means that the filter coefficients make a convergent series. Specifically, the coefficients decrease in time (and vanish at $t = \infty$); therefore, the filter has finite energy. This is the case for the wavelet $[1,(-\frac{1}{2})]$ with an inverse $[1,(\frac{1}{2}),(\frac{1}{4}), \ldots]$. The inverse is a stable spiking deconvolution filter. On the other hand, if the wavelet were maximum phase, then it does not have a stable inverse.

Table 2-13. Autocorrelation lags of Wavelets A, B, C, and D.

					Output
Wavelet A					
4	0	−1			
4	0	−1			17
	4	0	−1		0
		4	0	−1	−4
Wavelet B					
2	3	−2			
2	3	−2			17
	2	3	−2		0
	2	3	−2		−4
Wavelet C					
−2	3	2			
−2	3	2			17
	−2	3	2		0
		−2	3	2	−4
Wavelet D					
−1	0	4			
−1	0	4			17
	−1	0	4		0
		−1	0	4	−4

Table 2-14. Autocorrelation lags of input wavelet [1, $(-\frac{1}{2})$].

			Output
1	$-\frac{1}{2}$		
1	$-\frac{1}{2}$		$\frac{5}{4}$
	1	$-\frac{1}{2}$	$-\frac{1}{2}$

Table 2-15. Crosscorrelation lags of desired output (1,0,0) with input wavelet [1,$(-\frac{1}{2})$].

			Output
1	0	0	
1	$-\frac{1}{2}$		1
	1	$-\frac{1}{2}$	0

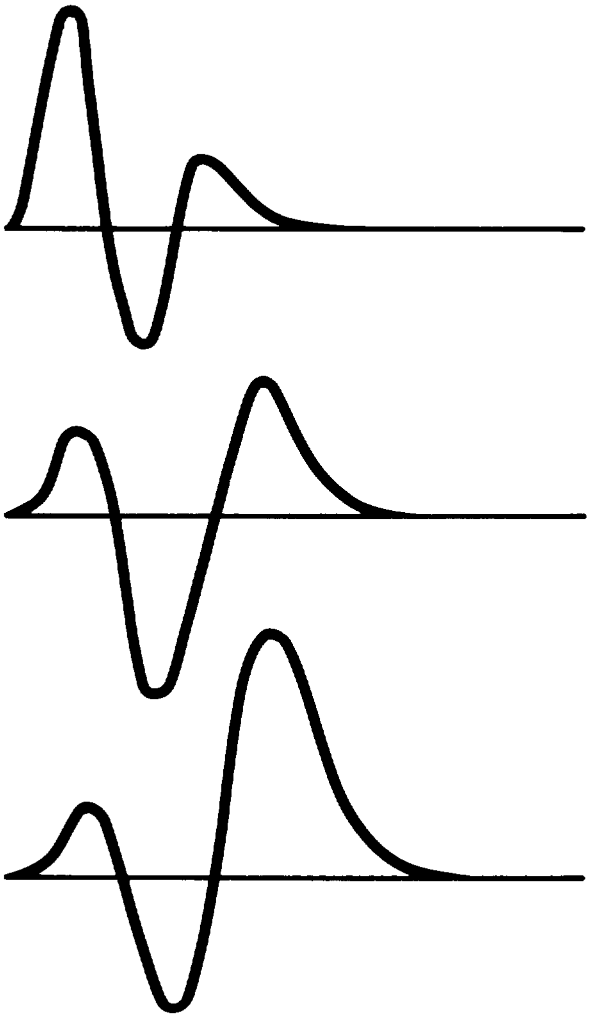


FIG. 2-15. A wavelet has a finite duration. If its energy is maximally front-loaded, then it is minimum-phase (top). If its energy is concentrated mostly in the middle, then it is mixed-phase (middle). Finally, if its energy is maximally end-loaded, then the wavelet is maximum-phase. A quantitative analysis of this phase concept is provided in Figure 2-16.

This is the case for the wavelet [$(-\frac{1}{2})$,1], whose inverse is given by the divergent series [−2,−4,−8, . . .]. Finally, a mixed-phase wavelet does not have a stable inverse. This discussion leads us to assumption 6.

> Assumption 6. The seismic wavelet is minimum phase. Therefore, it has a minimum-phase inverse.

This assumption is required by spiking deconvolution. In Section 2.6, we see that a wavelet can be converted into a delayed spike even if it is not minimum phase.

2.6 OPTIMUM WIENER FILTERS

Return to the desired output that was considered when studying inverse and least-squares filters; i.e., the zero-delay spike (1,0,0). The autocorrelation of the input wavelet [1,$(-\frac{1}{2})$] is shown in Table 2-14.

Within a scale factor of 2, this output is the same as the first column of the 2 × 2 matrix on the left side of equation (2.12), which was the equation that we solved to obtain the least-squares inverse filter. Equation (2.12) can be rewritten as

$$2\begin{bmatrix} \frac{5}{4} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{5}{4} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}. \tag{2.16}$$

Divide both sides by 2 to obtain:

$$\begin{bmatrix} \frac{5}{4} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{5}{4} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{2.17}$$

Now compute the crosscorrelation of the desired output (1,0,0) with the input wavelet [1,$(-\frac{1}{2})$] (Table 2-15).

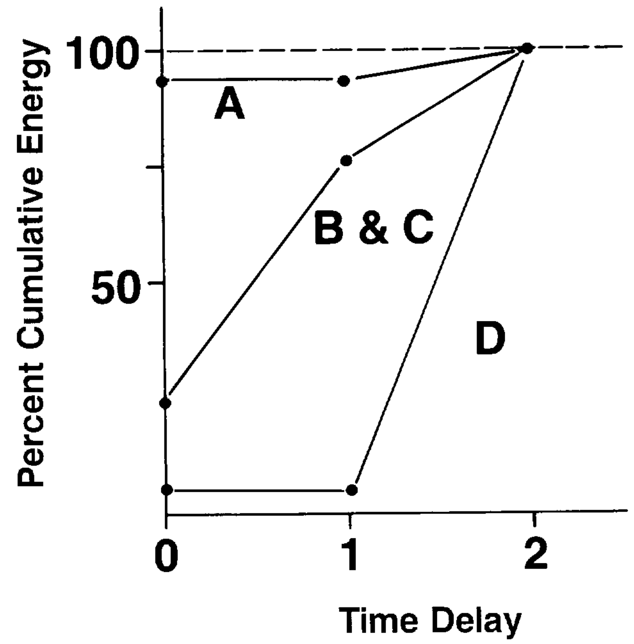


FIG. 2-16. A quantitative analysis of the minimum-and maximum-phase concept. The fastest rate of energy buildup in time occurs when the wavelet is minimum-phase *A*. The slowest rate occurs when the wavelet is maximum-phase D.

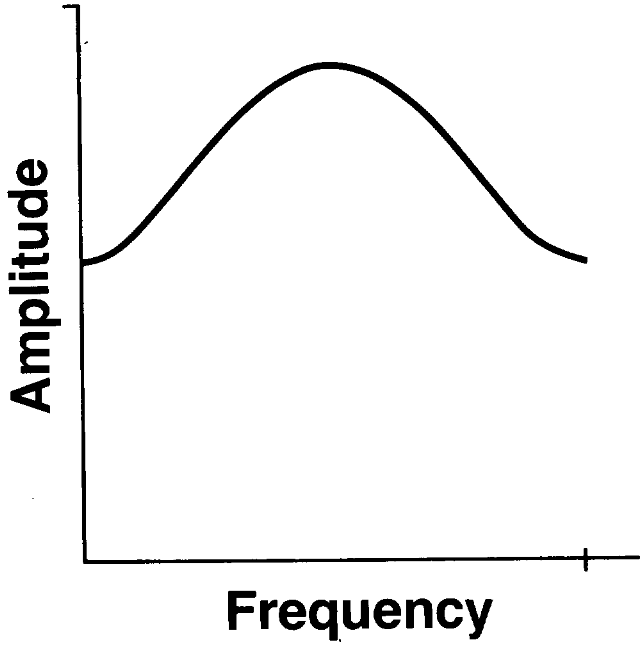


FIG. 2-17. All wavelets in Figure 2-16 (*A*, *B*, *C*, and *D*) have the same amplitude spectrum as shown above. (Adapted from Robinson, 1966.)

The output is the same as the column matrix on the right side of equation (2.17). In general, the elements of the matrix on the left side are the lags of the autocorrelation of the input wavelet, while the elements of the column matrix on the right side are the lags of the crosscorrelation of the desired output with the input wavelet.

Now perform similar operations for wavelet $[(-\frac{1}{2}),1]$. The autocorrelation of this wavelet is given in Table 2-16. Note that autocorrelation of wavelet $[(-\frac{1}{2}),1]$ is identical to that of wavelet $[1,(-\frac{1}{2})]$ (Table 2-14). As discussed in Section 2.5, this is an important property of a group of wavelets with the same amplitude spectrum. By rewriting the matrix equation, equation (2.15),

$$2\begin{bmatrix} \frac{5}{4} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{5}{4} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}. \tag{2.18}$$

Divide both sides by 2 to obtain:

$$\begin{bmatrix} \frac{5}{4} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{5}{4} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ 0 \end{bmatrix}. \tag{2.19}$$

The crosscorrelation of the desired output (1,0,0) with input wavelet $[(-\frac{1}{2}),1]$ is given in Table 2-17. Note that this result is the same as the right side of equation (2.19). Moreover, the elements of the matrix on the left side are the autocorrelation lags of the input wavelet. Matrix equations (2.17) and (2.19) were used to derive the least-squares inverse filters (Section 2.4). These filters then were applied to the input wavelets to compress them to zero-lag spike. The left side matrices in equations (2.17)

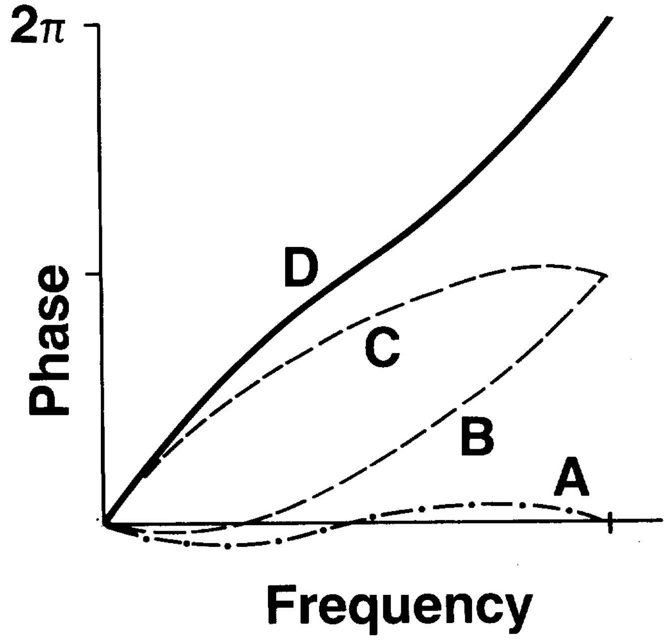


FIG. 2-18. Phase-lag spectra of the wavelets defined in Figure 2-16. They have the common amplitude spectrum of Figure 2-17. (Adapted from Robinson, 1966.)

Table 2-16. Autocorrelation lags of input wavelet $[(-\frac{1}{2}),1]$.

			Output
$-\frac{1}{2}$	1		Output
$-\frac{1}{2}$	1		$\frac{5}{4}$
	$-\frac{1}{2}$	1	$-\frac{1}{2}$

Table 2-17. Crosscorrelation of desired output (1,0,0) with input wavelet $[1,(-\frac{1}{2})]$.

			Output
1	0	0	Output
$-\frac{1}{2}$	1		$-\frac{1}{2}$
	$-\frac{1}{2}$	1	0

and (2.19) are made up of the autocorrelation lags of the input wavelets. Additionally, the column matrices on the right side are made up of lags of the crosscorrelation of the desired output, i.e., a zero-lag spike, with the input wavelets. These observations were generalized by Wiener to derive filters that convert the input to any desired output. The general form of the matrix equation for a filter of length n is (Robinson and Treitel, 1980):

$$\begin{bmatrix} r_0 & r_1 & r_2 & \cdots & r_{n-1} \\ r_1 & r_0 & r_1 & \cdots & r_{n-2} \\ r_2 & r_1 & r_0 & \cdots & r_{n-3} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ r_{n-1} & r_{n-2} & r_{n-3} & \cdots & r_0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ \cdot \\ \cdot \\ \cdot \\ g_{n-1} \end{bmatrix}. \tag{2.20}$$

Here r_i, a_i, and g_i, $i = 0,1,2,3, \ldots, n-1$ are the autocorrelation of the input wavelet, the desired filter coefficients, and the crosscorrelation of the desired output with the input wavelet, respectively.

The optimum Wiener filter a_i is optimum in that the least-squares error between the actual and desired outputs is minimum. When the desired output is the zero-lag spike $(1,0,0, \ldots, 0)$, then the Wiener filter is identical to the least-squares inverse filter. In other words, the latter really is a special case of the former. The Wiener filter applies to a large class of problems in which any desired output can be considered, not just the zero-lag spike.

Five choices for the desired output are:

Type 1: Zero-lag spike
Type 2: Spike at arbitrary lag
Type 3: Time-advanced form of input series
Type 4: Zero-phase wavelet
Type 5: Any desired arbitrary shape.

These desired output forms will be discussed in the following sections.

The general form of the normal equations [equation (2.20)] was arrived at through numerical examples for the special case where the desired output was a zero-lag spike. Appendix B.5 contains a concise mathematical treatment of the optimum Wiener filters. Figure 2-19 outlines the design and application of a Wiener filter.

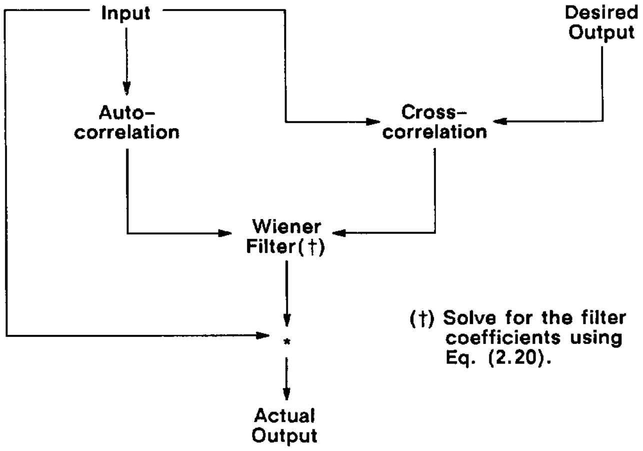


FIG. 2-19. A flowchart for Wiener filter design and application.

Determination of the Wiener filter coefficients requires solution of the so-called normal equations, equations (2.20). Although the 2 × 2 matrix was solved by hand in our numerical examples, the solution is costly for long filter operators. From equation (2.20), note that the autocorrelation matrix is symmetric. This special matrix, called the Toeplitz matrix, can be solved by Levinson recursion, a computationally efficient scheme. To do this, compute a two-point filter, derive from it a three-point filter, and so on, until the n-point filter is derived (Claerbout, 1976). In practice, filtering algorithms based on the optimum Wiener filter theory are known as Wiener-Levinson algorithms.

2.6.1 Spiking Deconvolution

The process with type 1 desired output (zero-lag spike) is called spiking deconvolution. Crosscorrelation of the desired spike $(1,0,0, \ldots, 0)$ with input wavelet $(x_0,x_1, x_2, \ldots, x_{n-1})$ yields the series $(x_0,0,0, \ldots, 0)$. The generalized form of the normal equations, equation (2.20), takes the special form:

$$\begin{bmatrix} r_0 & r_1 & r_2 & \cdots & r_{n-1} \\ r_1 & r_0 & r_1 & \cdots & r_{n-2} \\ r_2 & r_1 & r_0 & \cdots & r_{n-3} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ r_{n-1} & r_{n-2} & r_{n-3} & \cdots & r_0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}. \tag{2.21}$$

Equation (2.21) was scaled by $(1/x_0)$. The least-squares inverse filter, which was discussed in Section 2.4, has the same form as the matrix equation, equation (2.21). Therefore, spiking deconvolution is mathematically identical to

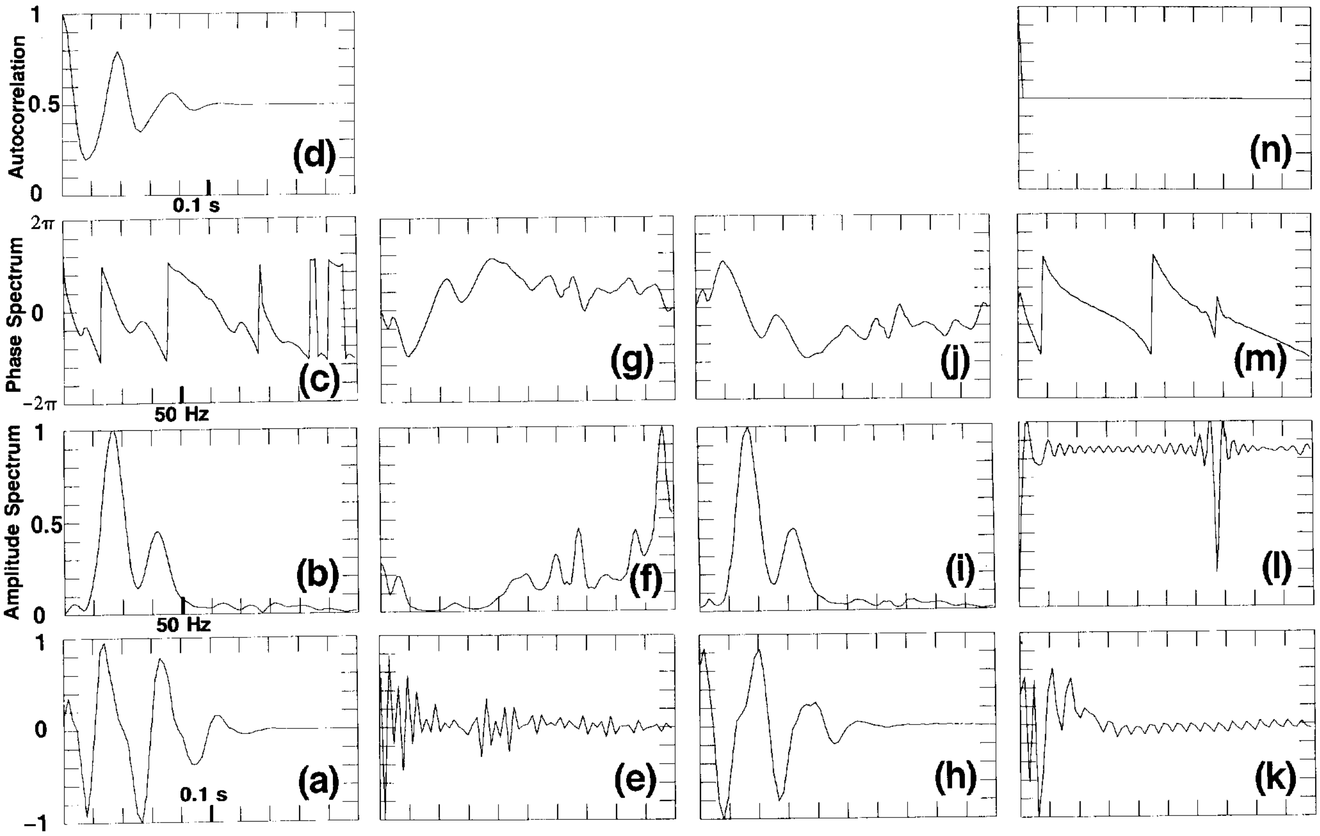


FIG. 2-20. Starting with wavelet (a), autocorrelogram (d) is computed to derive spiking deconvolution operator (e). This operator and its inverse (h) are minimum-phase. The inverse of the deconvolution operator has the same amplitude spectrum as that of the original wavelet. [Compare (i) to (b).] Its phase spectrum is simply the sign-reversed phase spectrum of the spiking deconvolution operator. [Compare (j) to (g).] If operator *e* were convolved with original wavelet *a*, then output *k* results. Although the output spectrum nearly is flat (l), it is far from being a spike at $t = 0$ (n). This desired output (n) is obtained if the input is minimum-phase wavelet (h), rather than mixed-phase wavelet (a).

least-squares inverse filtering. A distinction is made in practice between the two types of filtering. The autocorrelation matrix on the left side of equation (2.21) is computed from the input seismogram (assumption 5) in case of spiking deconvolution (statistical deconvolution), whereas it is directly computed from the known source wavelet in case of least-squares inverse filtering (deterministic deconvolution).

Figure 2-20 is a summary of spiking deconvolution based on the Wiener-Levinson algorithm. Frame a is the input mixed-phase wavelet. Its amplitude spectrum (frame b) indicates that the wavelet has most of its energy confined to a 10- to 50-Hz range. The autocorrelation function (frame d) is used in equation (2.21) to compute the spiking deconvolution operator (frame e). The amplitude spectrum of the operator (frame f) is approximately the inverse of the amplitude spectrum of the input wavelet (frame b). (The approximation improves as operator length increases.) This should be expected, since the goal of spiking deconvolution is to flatten the output spectrum. Application of this operator to the input wavelet gives the result shown in frame k.

Ideally, we would like to get a zero-lag spike, as shown in frame n. What went wrong? Assumption 6 was violated by the mixed-phase input wavelet (frame a). Frame h shows the inverse of the deconvolution operator. This is the minimum-phase equivalent of the input mixed-phase wavelet in frame a. Both wavelets have the same amplitude spectrum (frames b and i), but their phase spectra are significantly different (frames c and j). Since spiking deconvolution is equivalent to least-squares inverse filtering, the minimum-phase equivalent is merely the inverse of the deconvolution operator. Therefore, the amplitude spectrum of the operator is the inverse of the amplitude spectrum of the minimum-phase equivalent (frames f and i), and the phase spectrum of the operator is the negative of the phase spectrum of the minimum-phase wavelet, as seen in frames g and j. One way to extract the seismic wavelet, provided it is minimum phase, is to compute the spiking deconvolution operator and find its inverse.

In conclusion, if the input wavelet is not minimum phase, then spiking deconvolution cannot convert it to a perfect zero-lag spike (frame k). Although the amplitude spectrum is virtually flat (frame l), the phase spectrum of the output is not minimum phase (frame m). Finally, note

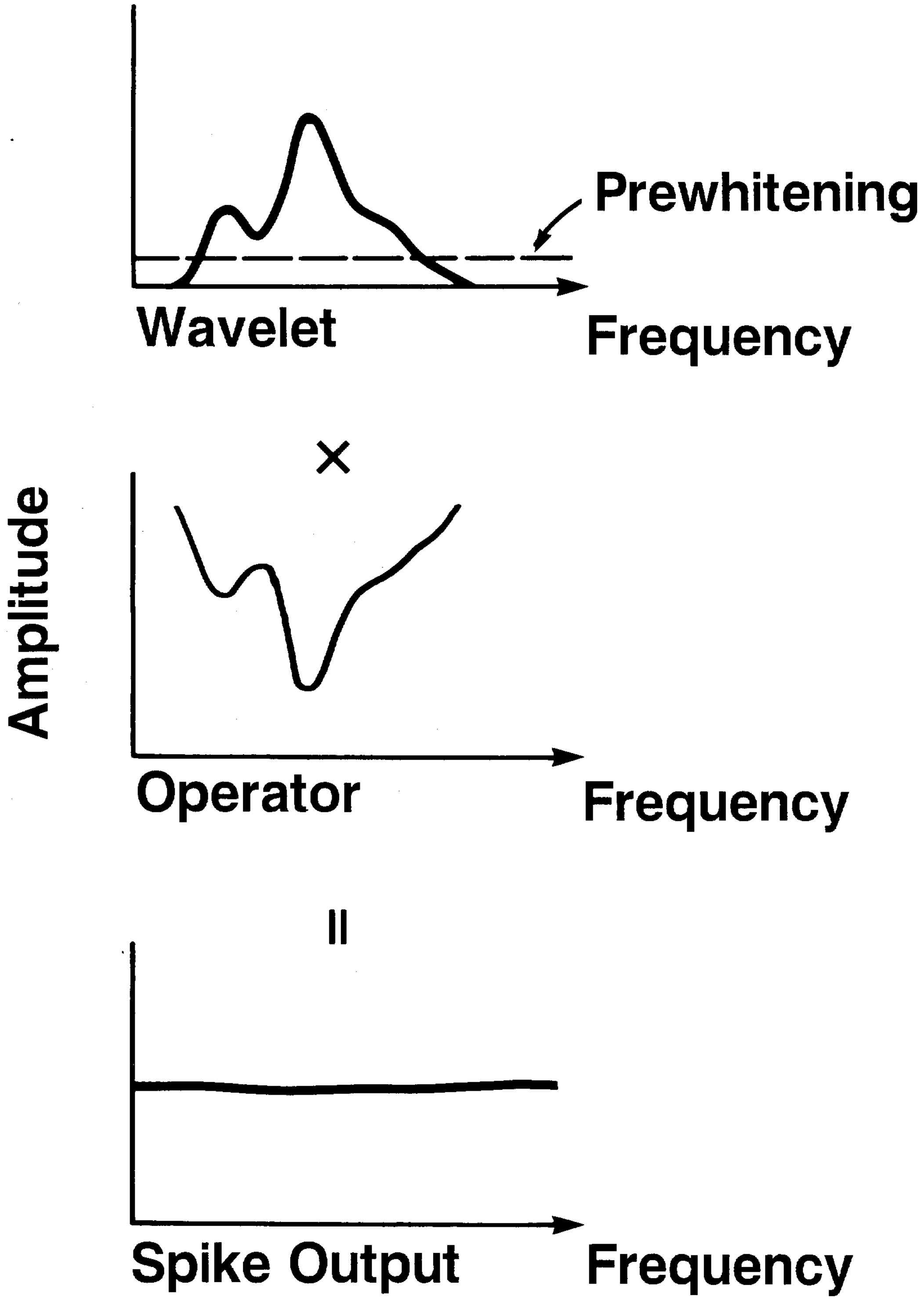


FIG. 2-21. Prewhitening amounts to adding a bias to the amplitude spectrum of the seismogram to be deconvolved. This prevents dividing by zero since the amplitude spectrum of the inverse filter (middle) is the inverse of that of the seismogram (top).

that the spiking deconvolution operator is the inverse of the minimum-phase equivalent of the input wavelet. This wavelet may or may not be minimum phase.

2.6.2 Prewhitening

From the preceding section, we know that the amplitude spectrum of the spiking deconvolution operator is (approximately) the inverse of the amplitude spectrum of the input wavelet. This is illustrated in Figure 2-21. What if we had zeroes in the amplitude spectrum of the input wavelet? To study this, apply a minimum-phase band-pass filter (see Exercise 2.10) with a wide passband (3-108 Hz) to the minimum-phase wavelet of Figure 2-20 (frame h). Deconvolution of the filtered wavelet does not produce a perfect spike, instead, a spike accompanied by a high-frequency pre-and post-cursor results (Figure 2-22). This poor result occurs because the deconvolution operator tries to boost the absent frequencies, as seen from the amplitude spectrum of the output. Can this problem occur in a recorded seismogram? Situations in which the input amplitude spectrum has zeroes rarely occur. There is always noise in the seismogram and it is additive in both the time and frequency domains. Moreover, numerical noise, which also is additive in the frequency domain, is generated during processing. However, to ensure numerical stability, an artificial level of *white noise* is introduced before deconvolution. This is called *prewhitening* and is illustrated in Figure 2-21. Prewhitening is achieved by adding a constant to the zero lag of the autocorrelation function. If the percent prewhitening is given by a number, $0 \leq \varepsilon < 1$, then the following normal equations [from equation (2.21)] must be solved.

$$\begin{bmatrix} (1+\varepsilon)r_0 & r_1 & r_2 & \cdots & r_{n-1} \\ r_1 & (1+\varepsilon)r_0 & r_1 & \cdots & r_{n-2} \\ r_2 & r_1 & (1+\varepsilon)r_0 & \cdots & r_{n-3} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ r_{n-1} & r_{n-2} & r_{n-3} & \cdots & (1+\varepsilon)r_0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}. \tag{2.22}$$

Adding a constant εr_0 to the zero lag of the autocorrelation function is the same as adding white noise to the spectrum, with its total energy equal to that constant. The effect of the prewhitening level on performance of the deconvolution is discussed in Section 2.7.3.

2.6.3 Wiener Shaping Filters

Spiking deconvolution had trouble compressing wavelet $[(-\frac{1}{2}),1]$ to a zero-lag spike (1,0,0) (Table 2-11). In terms of energy distribution, this input wavelet is more similar to a delayed spike, such as (0,1,0), than it is to a zero-lag spike, (1,0,0). Perhaps a filter that converts wavelet $[(-\frac{1}{2}),1]$ to a delayed spike would yield less error than the filter that shapes it to a zero-lag spike. Follow the flowchart in Figure 2-19 to design and apply a filter to convert $[(-\frac{1}{2}),1]$ to the delayed spike (0,1,0). First, compute the crosscorrelation (Table 2-18).

From Table 2-16, we know the autocorrelation of the input wavelet. By substituting the results from Tables 2-16 and 2-18 into the matrix equation, equation (2.20), we get

$$\begin{bmatrix} \frac{5}{4} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{5}{4} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ -\frac{1}{2} \end{bmatrix}. \tag{2.23}$$

By solving for the filter coefficients, $(a,b) = [(\frac{16}{21}),(-\frac{2}{21})]$ results. This filter is applied to the input wavelet as shown in Table 2-19. The energy of the least-squares error between the actual and desired outputs is $(\frac{84}{441})$. From Table 2-11, note that the error for a zero-lag spike was $(\frac{336}{441})$. This shows that there is less error when converting wavelet $[(-\frac{1}{2}),1]$ to the delayed spike (0,1,0) than to zero-lag spike (1,0,0). Test the spike with even more delay (0,0,1) and compute the error.

In general, for any given input wavelet, a series of desired outputs can be defined as delayed spikes. The least-squares errors then can be plotted as a function of delay. The delay (lag) that corresponds to the least error is chosen to define the desired delayed spike output. The actual output from the Wiener filter using this optimum delayed spike should be the most compact possible result.

The process that has a type 5 desired output (any desired arbitrary shape) is called wavelet shaping. The filter that does this is called a Wiener shaping filter. In fact, type 2 (delayed spike) and type 4 (zero-phase wavelet) desired outputs are special cases of the more general wavelet shaping.

Figure 2-23 shows a series of wavelet shapings that use delayed spikes as desired outputs. The input is a mixed-phase wavelet. Filter length was held constant in all eight cases. Note that the zero-delay spike case (spiking deconvolution) does not always yield the best result (Figure 2-23a). A delay in the neighborhood of 60 ms (Figure 2-23e) seems to yield an output that is closest to being a perfect spike. Typically, the process is not very sensitive to the amount of delay once it is close to the optimum delay. If the input wavelet were minimum-phase, then the optimum delay of the desired output spike generally is zero. On the other hand, if the input wavelet were mixed-phase, as illustrated in Figure 2-23, then the optimum delay is nonzero. Finally, if the input wavelet were maximum-phase, then the optimum delay is the length of that wavelet (Robinson and Treitel, 1980).

Can we not delay the desired spike output (Figure 2-23) and obtain a better result than we got from spiking deconvolution? This goal is achieved by applying a constant-time shift (60 ms in Figure 2-23) to a delayed spike result. Better yet, the same result can be obtained by shifting the shaping filter operator as much as the delay in the spike and applying it to the input wavelet. Such a filter operator is two sided (noncausal), since it has coefficients for negative and positive time values. The one-sided filter defined along the positive time axis has an *anticipation component*, while the filter defined along the negative time axis has a *memory component* (Robinson and Treitel, 1980). The two-sided filter has an anticipation component and a memory component. Figure 2-24 shows a series of shaping filterings with two-sided Wiener filters for various spike delay values.

Figure 2-25 shows examples of wavelet shaping. The input wavelet (trace b) is the same mixed-phase wavelet as in Figure 2-24 (top left frame). This wavelet is shaped into zero-phase wavelets with three different bandwidths (traces c, d, and e). The industry jargon for this process is *dephasing*. Figure 2-25 shows another wavelet shaping in which the input wavelet was converted to its minimum-phase equivalent (trace f). This conversion is often ap-

Table 2-18. Crosscorrelation of desired output (0,1,0) with input $[(-\frac{1}{2}),1]$.

0	1	0	Output
$-\frac{1}{2}$	1		1
	$-\frac{1}{2}$	1	$-\frac{1}{2}$

Table 2-19. Convolution of input wavelet $[(-\frac{1}{2}),1]$ with filter coefficients $[(\frac{16}{21}),(-\frac{2}{21})]$.

	$-\frac{1}{2}$	1		Output
$-\frac{2}{21}$	$\frac{16}{21}$			$-\frac{8}{21}$
	$-\frac{2}{21}$	$\frac{16}{21}$		$\frac{17}{21}$
		$-\frac{2}{21}$	$\frac{16}{21}$	$-\frac{2}{21}$

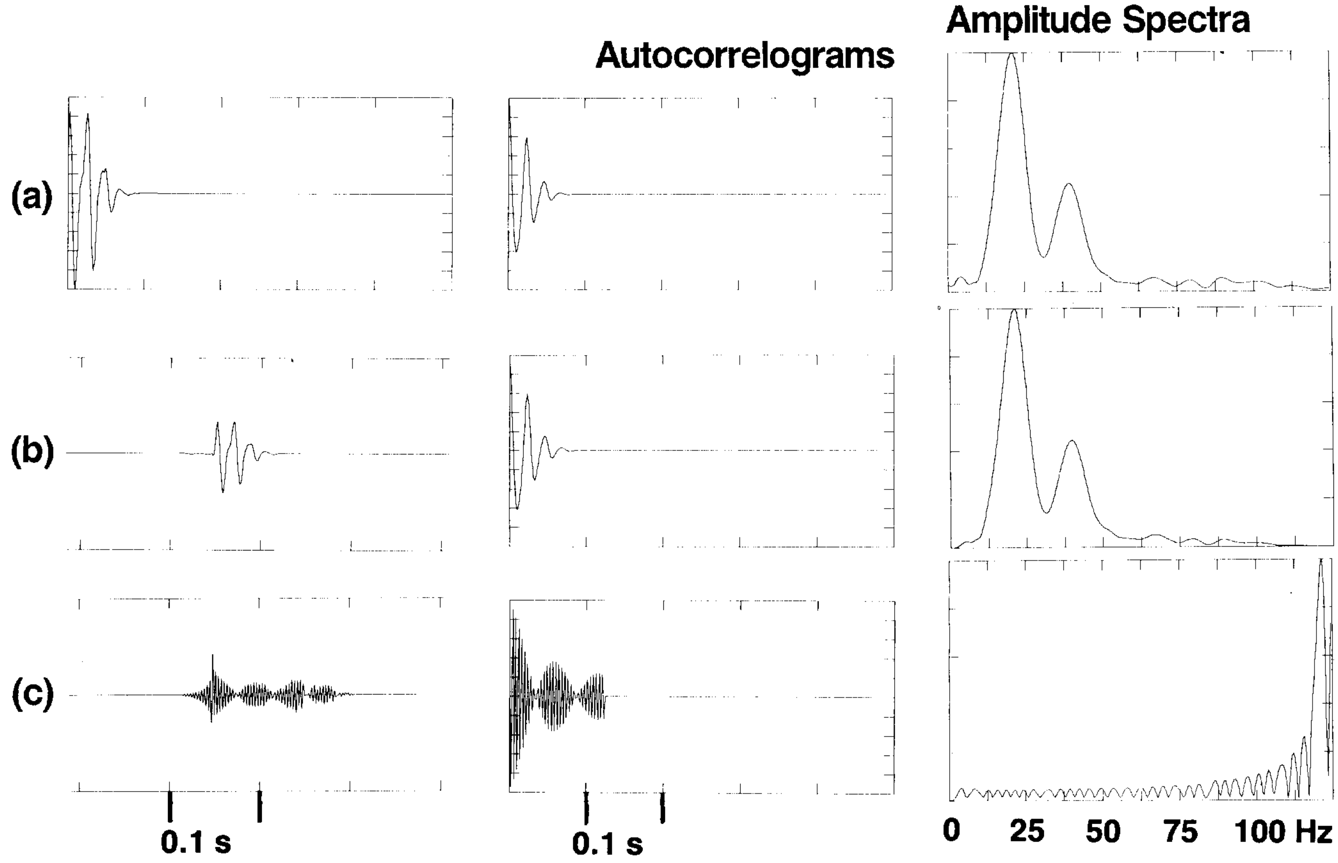


FIG. 2-22. (a) Minimum-phase wavelet, (b) band-pass filtering, (c) deconvolved. The amplitude spectrum of the band-pass filtered wavelet has zeroes above 108 Hz (middle row); therefore, the inverse filter derived from it yields unstable results (bottom row). The time delays on the wavelets in the left frames of the middle and bottom rows are for display purposes only.

plied to recorded air-gun signatures. Figure 2-26 shows examples of a recorded air-gun signature that was shaped into its minimum-phase equivalent and into a spike. When the input is the recorded signature, then the wavelet shapings in Figure 2-26 are called *signature processing*. Wavelet shaping requires knowledge of the input wavelet to compute the column on the right side of equation (2.20). If it is unknown, which is the case in reality, then it can be estimated statistically from the data. A minimum-phase wavelet, which then is shaped to a zero-phase wavelet, is derived from this estimate.

Wavelet processing is a term that is used with flexibility. The most common meaning refers to estimating (somehow) the basic wavelet embedded in the seismogram, designing a shaping filter to convert the estimated wavelet to a desired form, usually a broad-band zero-phase wavelet (Figure 2-25), and finally, applying the shaping filter to the seismogram. Another type of wavelet processing involves wavelet shaping in which the desired output is the zero-phase wavelet with the same amplitude spectrum as that of the input wavelet (Figure 2-27). Note that this type of wavelet processing does not try to flatten the spectrum, but only tries to correct for the phase of the input wavelet, which sometimes is assumed to be minimum-phase.

2.6.4 Predictive Deconvolution

The type 3 desired output, a time-advanced form of the input series, suggests a prediction process. Given the input $x(t)$, we want to predict its value at some future time $(t + \alpha)$, where α is prediction lag. Wiener showed that the filter used to estimate $x(t + \alpha)$ can be computed by using a special form of the matrix equation, equation (2.20) (Robinson and Treitel, 1980). Since the desired output $x(t + \alpha)$ is the time-advanced version of the input $x(t)$, we need to specialize the right side of equation (2.20) for the prediction problem.

Consider a five-point input time series x_i, where $i = 0,1,2,3,4$ and set $\alpha = 2$. The autocorrelation of x_i is computed in Table 2-20. The crosscorrelation between the desired output $x(t+2)$ and the input $x(t)$ is shown in Table 2-21.

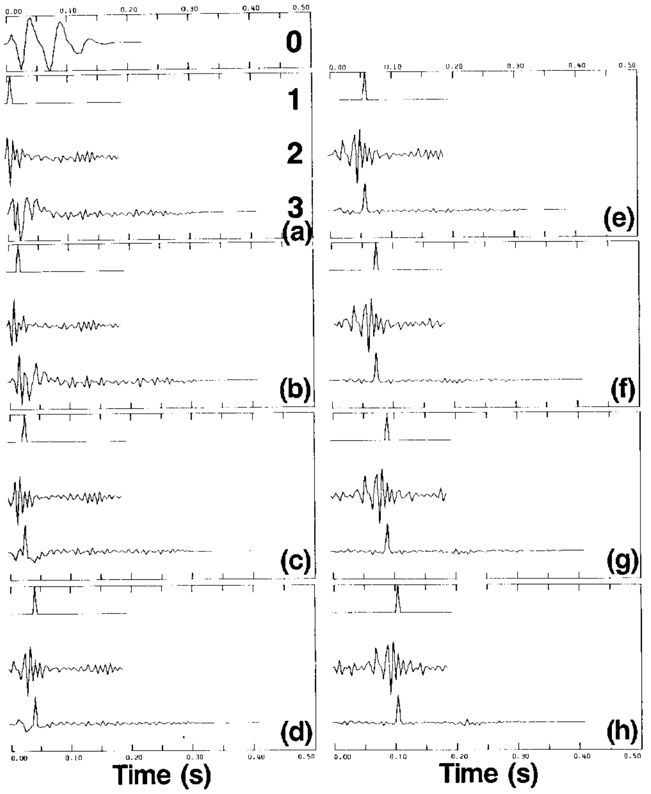


FIG. 2-23. Shaping filtering. (0) Input wavelet, (1) desired output, (2) shaping filter operator, (3) actual output. Here, the purpose is to convert the mixed-phased wavelet (0) to a series of delayed spikes (*a* through *h*) by using a one-sided operator (anticipation component only). The best result is with a 60-ms delay (e).

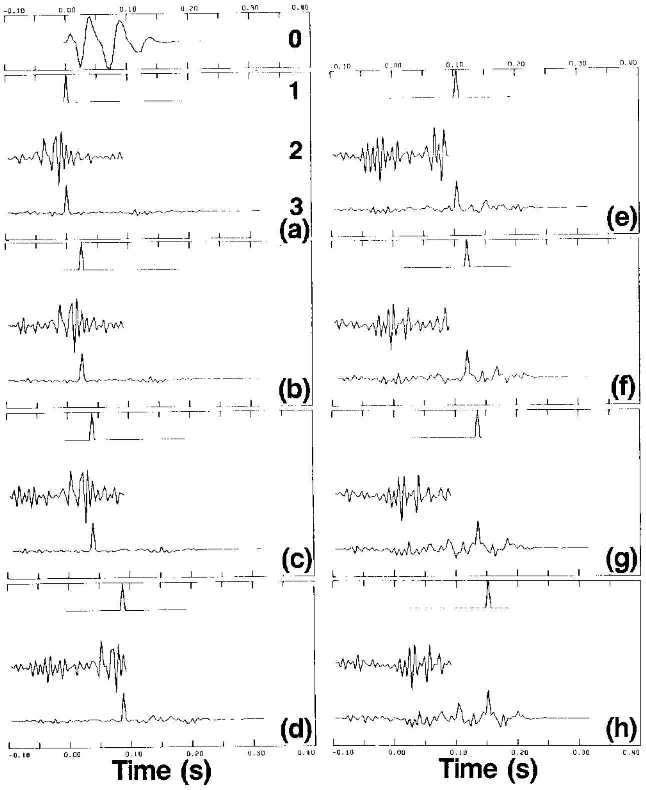


FIG. 2-24. Shaping filtering. (0) Input wavelet, (1) desired output, (2) shaping filter operator, (3) actual output. Here, the purpose is to convert the mixed-phase wavelet (0) to a series of delayed spikes (*a* through *h*) using a two-sided operator (with memory and anticipation components). The best result is obtained with a zero-delay spike using a two-sided filter (a).

Compare the results in Tables 2-20 and 2-21. Note that $g_i = r_{i+\alpha}$, $\alpha = 2$, and $i = 0,1,2,3,4$. Equation (2.20) is rewritten as follows:

$$\begin{bmatrix} r_0 & r_1 & r_2 & r_3 & r_4 \\ r_1 & r_0 & r_1 & r_2 & r_3 \\ r_2 & r_1 & r_0 & r_1 & r_2 \\ r_3 & r_2 & r_1 & r_0 & r_1 \\ r_4 & r_3 & r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} r_2 \\ r_3 \\ r_4 \\ r_5 \\ r_6 \end{bmatrix} . \tag{2.24}$$

The prediction filter coefficients a_i, where $i = 0,1,2,3,4$, can be computed from equation (2.24). Actual output now can be computed (Table 2-22).

Since we are trying to predict the time-advanced form of input, the actual output is an estimate of the series $x_{i+\alpha}$, where $\alpha = 2$. The prediction error series is given in Table 2-23.

The results in Table 2-23 suggest that the error series can be obtained more directly by convolving the input series with a filter with coefficients $(1,0,-a_0, -a_1, -a_2, -a_3, -a_4)$ (Table 2-24).

The results for e_2, e_3, e_4, e_5, and e_6 are identical (Tables 2-23 and 2-24). Since the series $(a_0, a_1, a_2, a_3, a_4)$ is called the *prediction filter*, it is natural to call the series $(1,0,-a_0, -a_1, -a_2, -a_3, -a_4)$ the *prediction error filter*. When applied on the input series, this filter yields the

Table 2-20. Autocorrelation lags of the input series $[x_0, x_1, x_2, x_3, x_4]$.

$r_0 = x_0^2 + x_1^2 + x_2^2 + x_3^2 + x_4^2$
$r_1 = x_0x_1 + x_1x_2 + x_2x_3 + x_3x_4$
$r_2 = x_0x_2 + x_1x_3 + x_2x_4$
$r_3 = x_0x_3 + x_1x_4$
$r_4 = x_0x_4$
$r_5 = 0$
$r_6 = 0$

Table 2-21. Crosscorrelation between desired output $x(t + 2)$ and input $x(t)$.

$g_0 = x_0x_2 + x_1x_3 + x_2x_4$
$g_1 = x_0x_3 + x_1x_4$
$g_2 = x_0x_4$
$g_3 = 0$
$g_4 = 0$

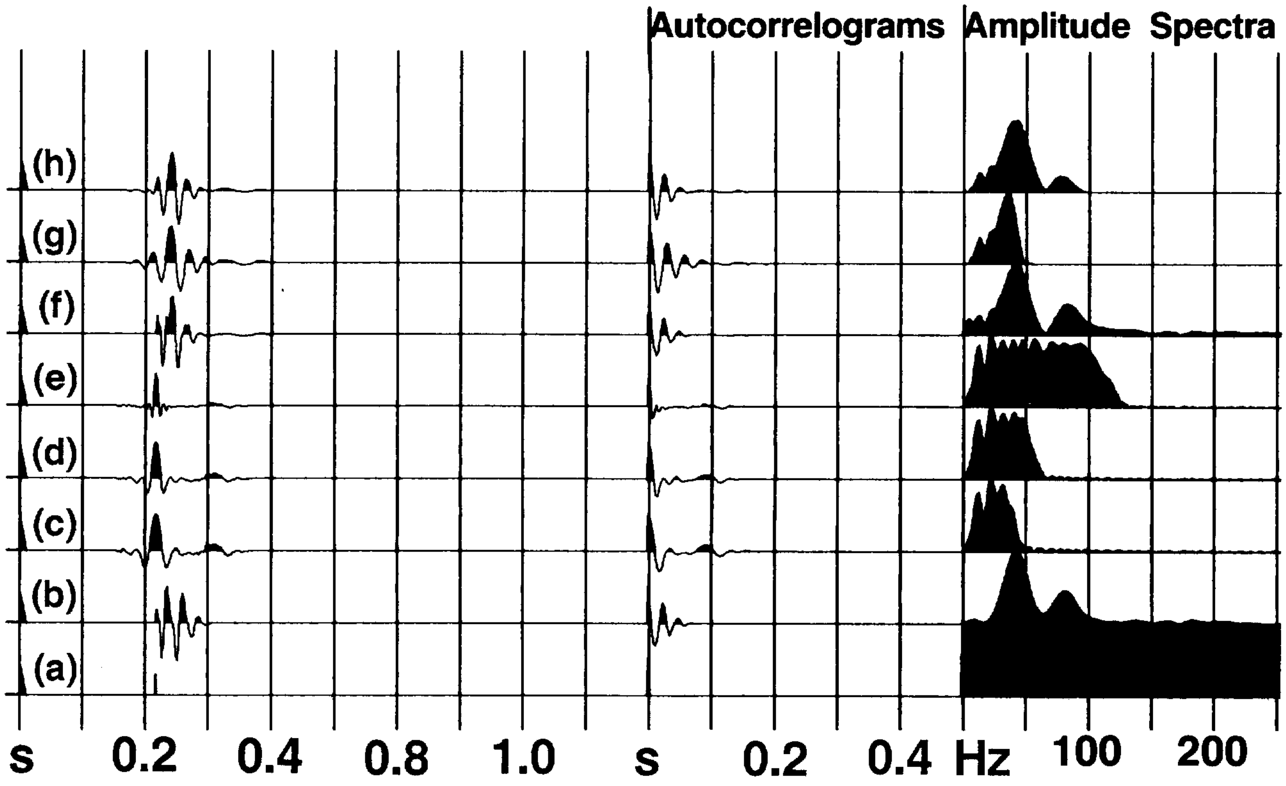


FIG. 2-25. Shaping filtering with various desired outputs. (a) Impulse response, (b) input seismogram. Here *c*, *d*, and *e* show three possible desired outputs that are band-limited zero-phase wavelets, while *f* shows a desired output that is the minimum-phase equivalent of the input wavelet *b*. Finally, *g* and *h* are desired outputs that are band-pass filtered versions of *f*.

error series in the prediction process.

Why place so much emphasis on the error series? Consider the prediction process as it relates to a seismic trace. A time series can be predicted at a future time $t + \alpha$, where α is the prediction lag (or prediction distance). A seismic trace often has a predictable component (multiples) with a periodic rate of occurrence. According to assumption 5, anything else, such as genuine reflections, is unpredictable.

Some may claim that reflections are predictable as well; this may be the case if deposition is cyclic. However, this type of deposition is not often encountered. While the prediction filter yields the predictable component (the multiples) of a seismic trace, the remaining unpredictable part, the error series, is essentially the reflection series.

Equation (2.24) can be generalized for the case of an n-long prediction filter and an α-long prediction lag.

$$\begin{bmatrix} r_0 & r_1 & r_2 & \cdots & r_{n-1} \\ r_1 & r_0 & r_1 & \cdots & r_{n-2} \\ r_2 & r_1 & r_0 & \cdots & r_{n-3} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ r_{n-1} & r_{n-2} & r_{n-3} & \cdots & r_0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} r_\alpha \\ r_{\alpha+1} \\ r_{\alpha+2} \\ \cdot \\ \cdot \\ \cdot \\ r_{\alpha+n-1} \end{bmatrix}. \tag{2.25}$$

Table 2-22. Convolution of prediction filter $a(t)$ with input series to compute actual output $y(t)$.

$$\begin{aligned} y_0 &= a_0x_0 \\ y_1 &= a_1x_0 + a_0x_1 \\ y_2 &= a_2x_0 + a_1x_1 + a_0x_2 \\ y_3 &= a_3x_0 + a_2x_1 + a_1x_2 + a_0x_3 \\ y_4 &= a_4x_0 + a_3x_1 + a_2x_2 + a_1x_3 + a_0x_4 \end{aligned}$$

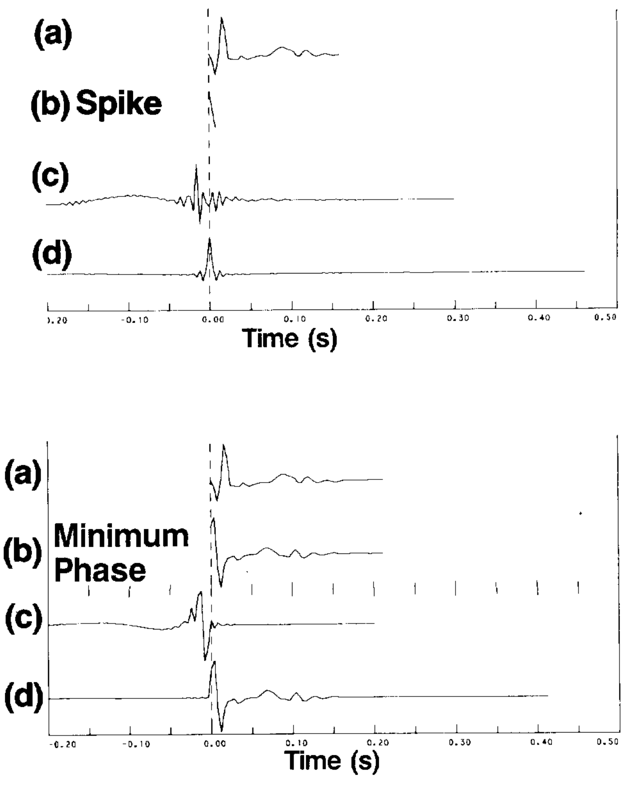


FIG. 2-26. Signature processing: (a) Recorded signature, (b) desired output, (c) shaping operator, (d) shaped signature. The desired output is a zero-delay spike (top) and the minimum-phase equivalent of the recorded signature (bottom).

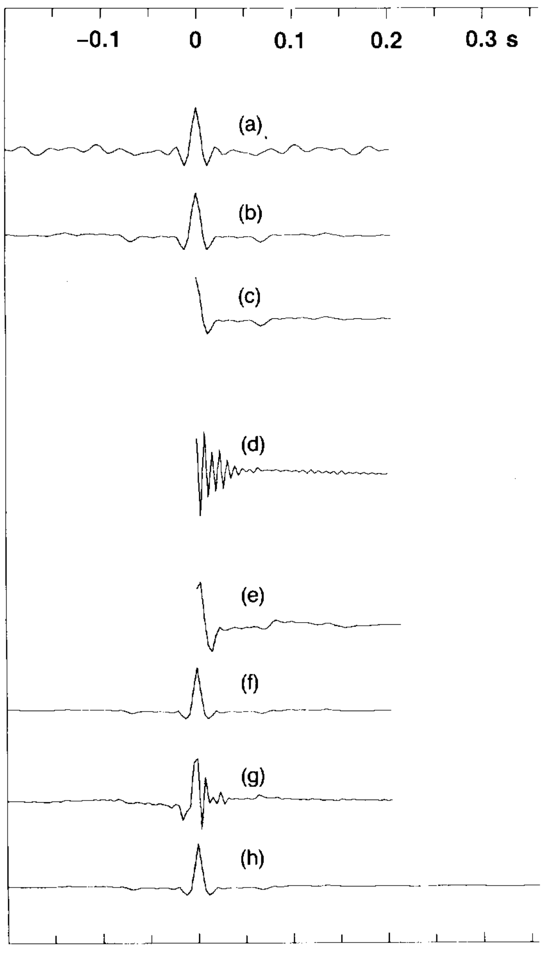


FIG. 2-27. Wavelet processing. An autocorrelogram a, estimated from the seismic trace, is used after smoothing (b) to compute the spiking deconvolution operator d. (Here c is just a one-sided version of b.) The inverse of this operator is the minimum-phase wavelet e, which is sometimes assumed to be the basic wavelet contained in the original seismic trace. It is easy to compute its zero-phase equivalent f and design a shaping filter g that converts the minimum-phase wavelet e to the zero-phase wavelet f. The actual output is h, which should be compared with f. Note that the zero-phase equivalent f has the same amplitude spectrum as the minimum-phase wavelet e.

Table 2-23. The error series $e_{i+2} = x_{i+2} - y_i$.

$e_2 = x_2 - a_0x_0$

$e_3 = x_3 - a_1x_0 - a_0x_1$

$e_4 = x_4 - a_2x_0 - a_1x_1 - a_0x_2$

$e_5 = 0 - a_3x_0 - a_2x_1 - a_1x_2 - a_0x_3$

$e_6 = 0 - a_4x_0 - a_3x_1 - a_2x_2 - a_1x_3 - a_0x_4$

Table 2-24. Convolution of filter coefficients $[1, 0, -a_i]$, $i = 0,1,2,3,4$, with input series x_i, $i = 0,1,2,3,4$.

$e_0 = x_0$

$e_1 = x_1$

$e_2 = x_2 - a_0x_0$

$e_3 = x_3 - a_1x_0 - a_0x_1$

$e_4 = x_4 - a_2x_0 - a_1x_1 - a_0x_2$

$e_5 = 0 - a_3x_0 - a_2x_1 - a_1x_2 - a_0x_3$

$e_6 = 0 - a_4x_0 - a_3x_1 - a_2x_2 - a_1x_3 - a_0x_4$

Note that design of the prediction filters requires only autocorrelation of the input series. There are two approaches to predictive deconvolution. The prediction filter may be designed using equation (2.25) and applied on input series as described in Figure 2-28.

Alternatively, the prediction error filter can be designed and convolved with the input series as described in Figure 2-29.

Now consider the special case of unit prediction lag, $\alpha = 1$. Equation (2.25) takes the following form:

$$\begin{bmatrix} r_0 & r_1 & r_2 & r_3 & r_4 \\ r_1 & r_0 & r_1 & r_2 & r_3 \\ r_2 & r_1 & r_0 & r_1 & r_2 \\ r_3 & r_2 & r_1 & r_0 & r_1 \\ r_4 & r_3 & r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \end{bmatrix}. \qquad (2.26)$$

By augmenting the right side to the left side,

$$\left[\begin{array}{c|ccccc} -r_1 & r_0 & r_1 & r_2 & r_3 & r_4 \\ -r_2 & r_1 & r_0 & r_1 & r_2 & r_3 \\ -r_3 & r_2 & r_1 & r_0 & r_1 & r_2 \\ -r_4 & r_3 & r_2 & r_1 & r_0 & r_1 \\ -r_5 & r_4 & r_3 & r_2 & r_1 & r_0 \end{array}\right] \begin{bmatrix} 1 \\ a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \qquad (2.27)$$

Add one row and move the negative sign to the column matrix that represents the filter coefficients to get:

$$\left[\begin{array}{c|ccccc} r_0 & r_1 & r_2 & r_3 & r_4 & r_5 \\ \hline r_1 & r_0 & r_1 & r_2 & r_3 & r_4 \\ r_2 & r_1 & r_0 & r_1 & r_2 & r_3 \\ r_3 & r_2 & r_1 & r_0 & r_1 & r_2 \\ r_4 & r_3 & r_2 & r_1 & r_0 & r_1 \\ r_5 & r_4 & r_3 & r_2 & r_1 & r_0 \end{array}\right] \begin{bmatrix} 1 \\ -a_0 \\ -a_1 \\ -a_2 \\ -a_3 \\ -a_4 \end{bmatrix} = \begin{bmatrix} L \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad (2.28)$$

where $L = r_0 - r_1a_0 - r_2a_1 - r_3a_2 - r_4a_3 - r_5a_4$. Note that there are six unknowns, $\{a_0, a_1, a_2, a_3, a_4, L\}$, and six equations. Solution of these equations yields the unit-delay prediction error filter, $\{1, -a_0, -a_1, -a_2, -a_3, -a_4\}$.

We can rewrite equation (2.28) as follows:

$$\begin{bmatrix} r_0 & r_1 & r_2 & r_3 & r_4 & r_5 \\ r_1 & r_0 & r_1 & r_2 & r_3 & r_4 \\ r_2 & r_1 & r_0 & r_1 & r_2 & r_3 \\ r_3 & r_2 & r_1 & r_0 & r_1 & r_2 \\ r_4 & r_3 & r_2 & r_1 & r_0 & r_1 \\ r_5 & r_4 & r_3 & r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix} = \begin{bmatrix} L \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \qquad (2.29)$$

where $b_0 = 1$, $b_i = -a_{i-1}$ and $i = 1,2,3,4,5$. This equation has a familiar structure. In fact, except for the scale factor L, it has the same form as equation (2.21), which yields the coefficients for the least-squares zero-delay inverse filter. Note that this inverse filter is the same as the prediction error filter with unit prediction lag, except for a scale factor. Therefore, spiking deconvolution actually is a special case of predictive deconvolution with unit prediction lag.

We now know that predictive deconvolution is a general process that encompasses spiking deconvolution. In general, the following statement can be made: Given an input wavelet of length $(n + \alpha)$, the prediction error filter contracts it to an α-long wavelet, where α is the prediction lag (Peacock and Treitel, 1969). When $\alpha = 1$, the procedure is called *spiking deconvolution*.

Figure 2-30 interrelates the various filters discussed in this chapter and indicates the kind of process they imply. From Figure 2-30, note that Wiener filters can be used to solve a wide range of problems. In particular, predictive deconvolution is an integral part of seismic data processing that is aimed at compressing the seismic wavelet, thereby increasing temporal resolution. In the limit, it can be used to spike the seismic wavelet and obtain an estimate for reflectivity.

Review the following assumptions that underlie predictive deconvolution:

Assumption 1a.	The earth is made up of horizontal layers of constant velocity.
Assumption 1b.	The source generates a compressional plane wave that impinges on layer boundaries at normal incidence. Under such circumstances, no shear waves are generated.
Assumption 2.	The source waveform does not change as it travels in the subsurface; i.e., it is stationary.

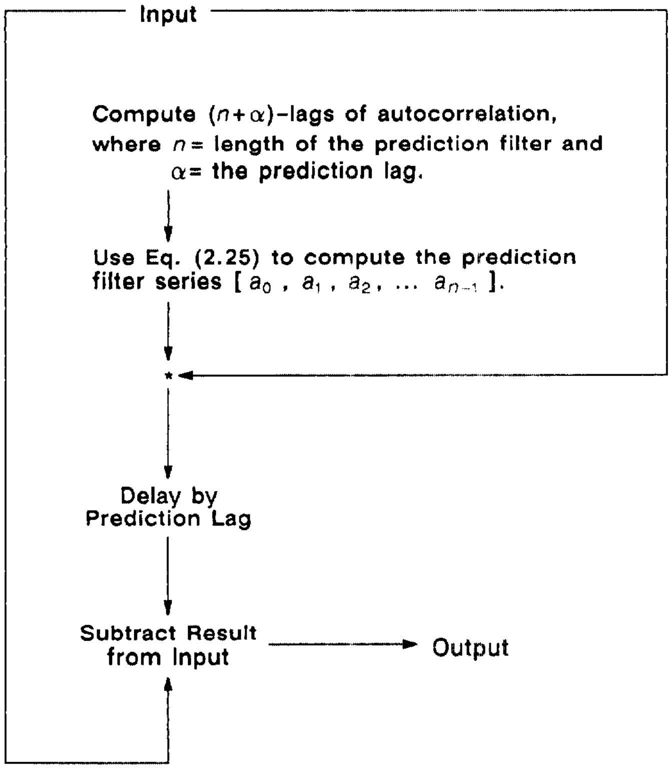


FIG. 2-28. A flowchart for predictive deconvolution using prediction filters.

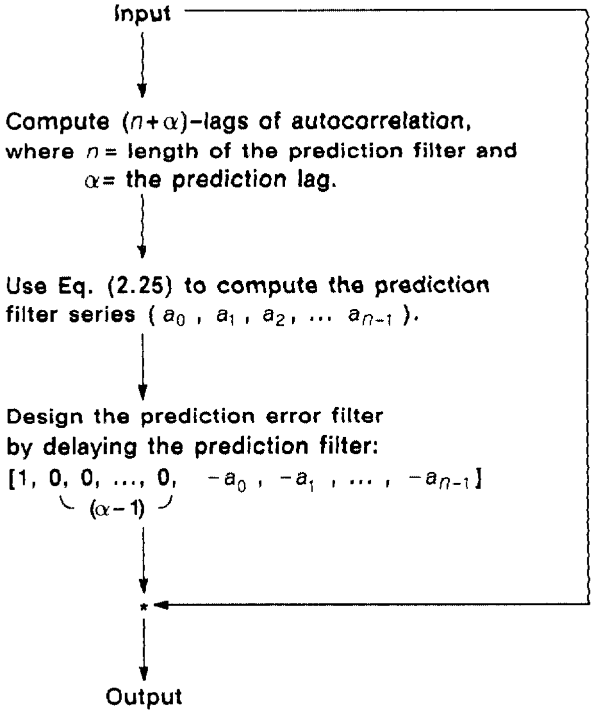


FIG. 2-29. A flowchart for predictive deconvolution using prediction error filters.

Assumption 3. The noise component $n(t)$ is zero.
Assumption 4. Reflectivity is a random process. This implies that the seismogram has seismic wavelet characteristics since their autocorrelations and amplitude spectra are similar.
Assumption 5. The seismic wavelet is minimum phase. Therefore, it has a minimum-phase inverse.

Assumption 1a, 1b, and 2 are the basis for the convolutional model of the recorded seismogram (Section 2.2). In practice, deconvolution often yields good results in areas where these three assumptions are not strictly valid. Assumption 2 can be relaxed in practice by considering a time-variant deconvolution. In this technique, a seismogram is divided into a number of time gates, typically three or more. Deconvolution operators then are designed from each gate and convolved with data within that gate. Not a good deal can be done about assumption 3. However, noise can be minimized in the recording process. Deconvolution operators can be designed using time gates and frequency bands with low noise level. Poststack deconvolution can be used in an effort to take advantage of the noise reduction inherent in the stacking process.

If the source wavelet were minimum-phase and known, then a perfect result could be obtained from deconvolution in the noise-free case (Figures 2-31 and 2-32, trace c). If assumption 4 were violated and if the source waveform were not known, then you would have problems (trace d). The quality of the output from the spiking deconvolution is degraded further when the source wavelet is not minimum-phase (Figures 2-33 and 2-34); that is, assumption 5 is violated. Finally, if there were noise in the data, that is, when assumption 3 is violated, then the result of the deconvolution would be unacceptable (Figure 2-35).

2.7 PREDICTIVE DECONVOLUTION IN PRACTICE

Figures 2-31 through 2-35 test our confidence in the usefulness of predictive deconvolution. In reality, deconvolution has been applied to billions of seismic traces; most of the time it has yielded satisfactory results. Figures 2-31 through 2-35 emphasize the critical assumptions that underlie predictive deconvolution. When deconvolution does not work on some data, the most probable reason is that one or more of the above assumptions has been violated. In this section, a series of numerical experiments will be performed to examine the validity of these assumptions. The purpose of these experiments is to gain a basic understanding of deconvolution from a practical point of view.

(Text continued on page 114)

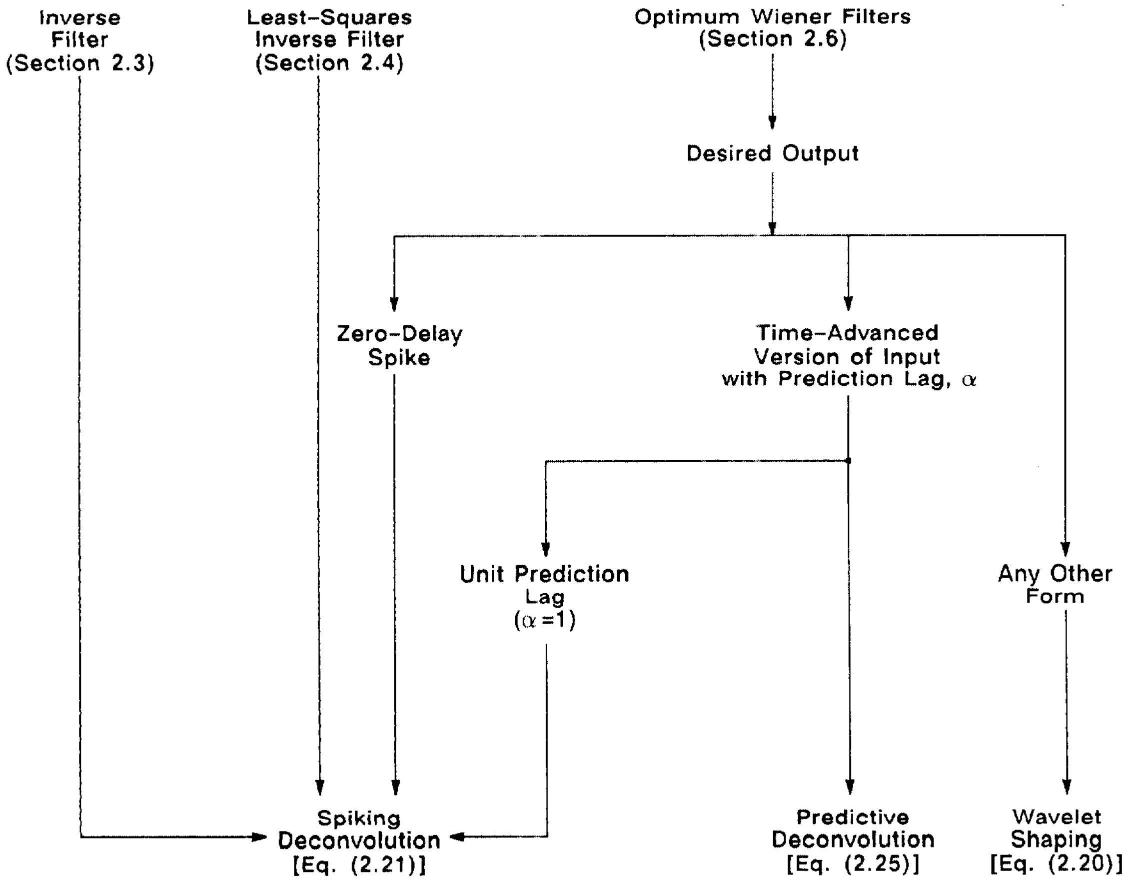


FIG. 2-30. A flowchart for interrelations between various deconvolution filters.

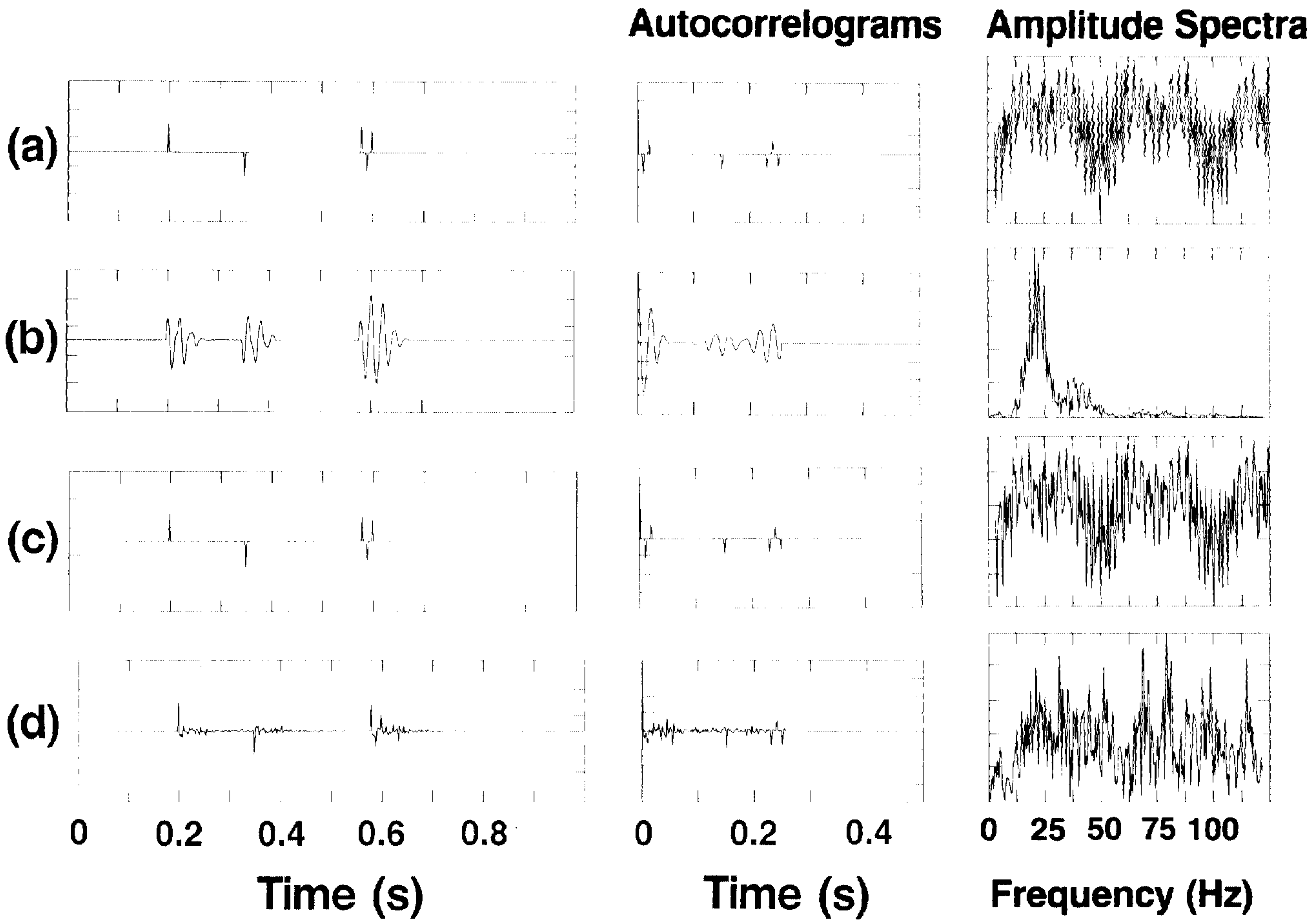


FIG. 2-31. (a) Impulse response, (b) seismogram, (c) spiking deconvolution using known, minimum-phase wavelet. (d) deconvolution assuming an unknown, minimum-phase source wavelet. Impulse response (a) is a sparse-spike series. For an unknown source wavelet (in violation of assumption 4), spiking deconvolution yields a less than perfect result. [Compare (c) and (d).]

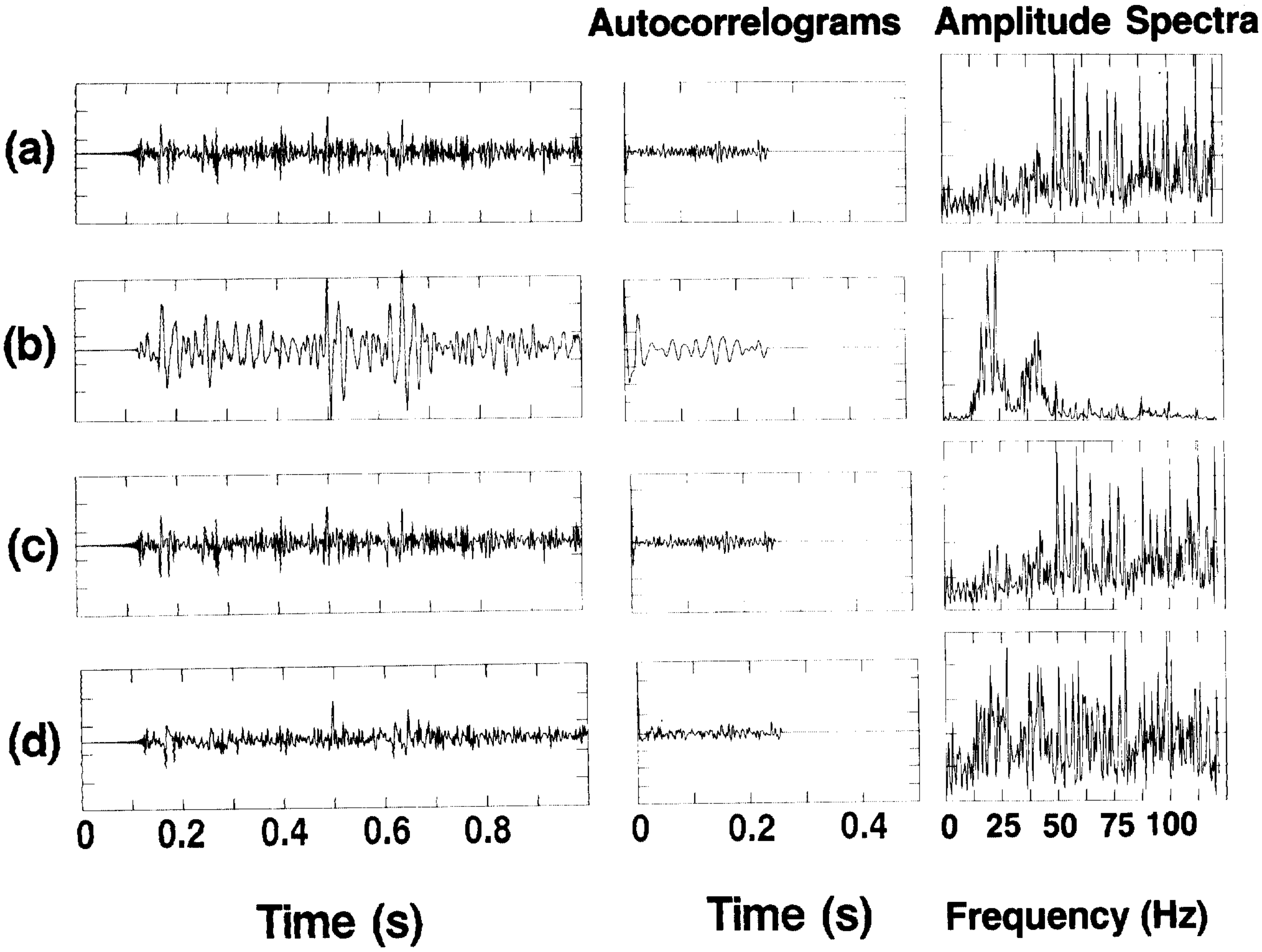


FIG. 2-32. (a) Impulse response, (b) seismogram, (c) spiking deconvolution using known, minimum-phase source wavelet, (d) deconvolution assuming an unknown, minimum-phase source wavelet. Impulse response (a) is based on a sonic log (Figure 2-8a). For the unknown source wavelet (in violation of assumption 4), spiking deconvolution yields a less than perfect result. [Compare (c) and (d).]

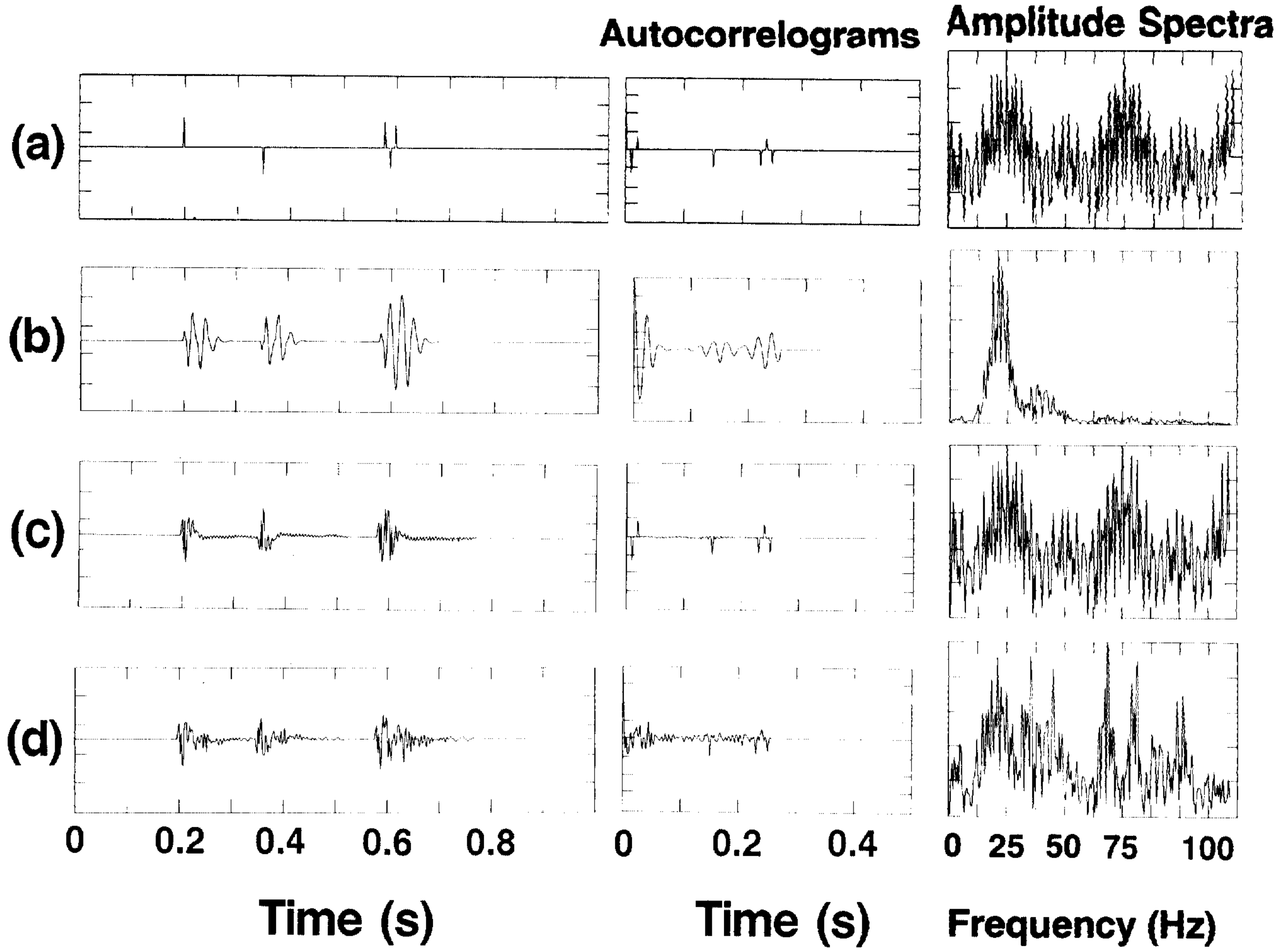


FIG. 2-33. (a) Impulse response, (b) seismogram, (c) deconvolution using a known, mixed-phase source wavelet, (d) deconvolution assuming an unknown, mixed-phase source wavelet. Impulse response (a) is a sparse-spike series. For a mixed-phase source wavelet (in violation of assumption 5), spiking deconvolution yields a degraded output (d), even when the wavelet is known (c).

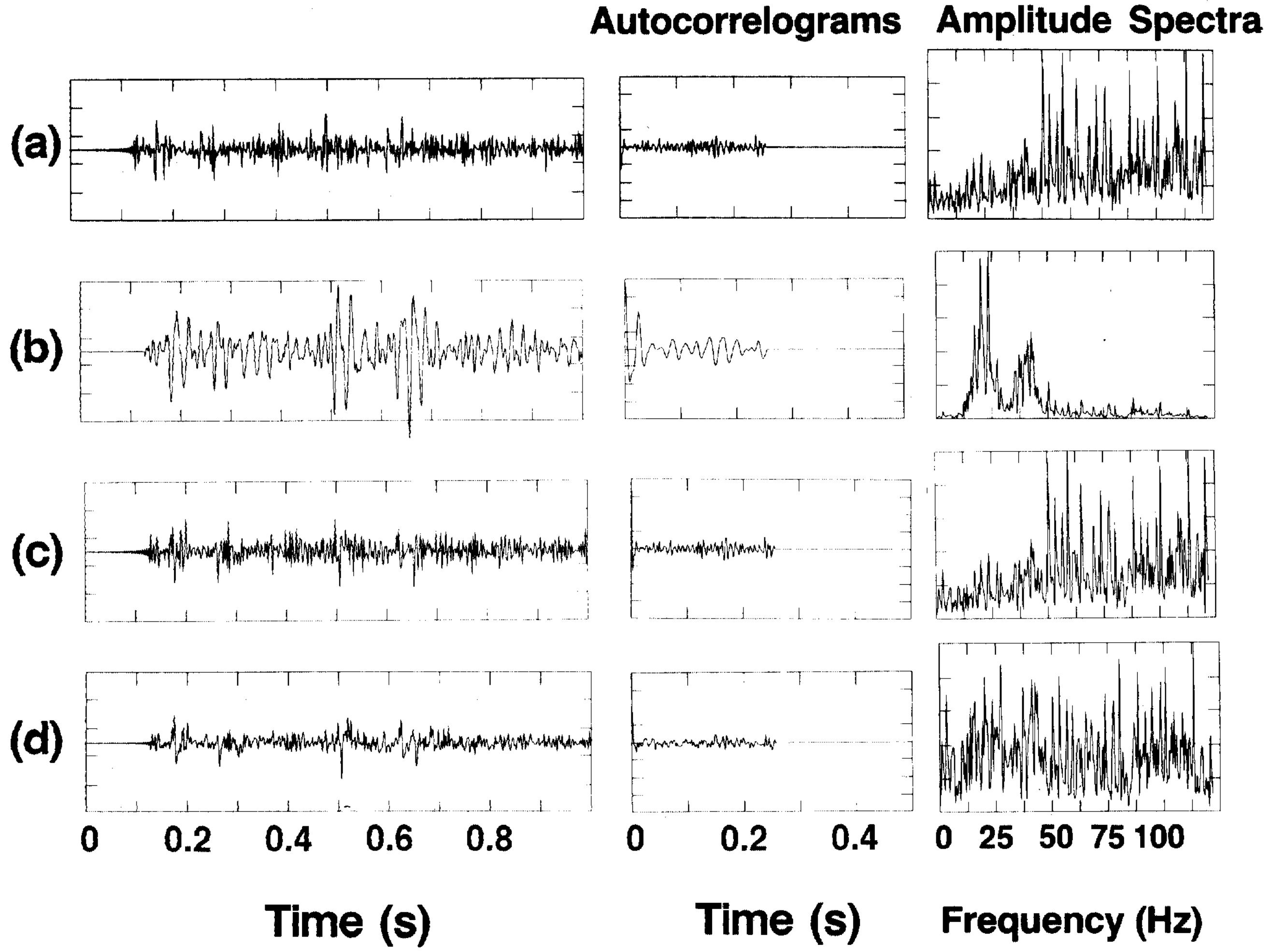


FIG. 2-34. (a) Impulse response, (b) seismogram, (c) deconvolution using a known, mixed-phase source wavelet, (d) deconvolution assuming an unknown, mixed-phase source. Impulse response (a) is based on the sonic log of Figure 2-8a. For the mixed-phase source wavelet (in violation of assumption 5), spiking deconvolution yields a degraded output (d) even when the wavelet is known (c).

2.7.1 Operator Length

We start with a single, isolated minimum-phase wavelet as shown in Figure 2-36 (trace b). Assumptions 1 through 5 are satisfied for this wavelet. The ideal result of spiking deconvolution is a zero-lag spike, as indicated by trace a. In this and the following numerical analyses, we refer to the autocorrelogram and amplitude spectrum (plotted with linear scale) of the output from each deconvolution test to better evaluate the results. In Figure 2-36 and the following figures, n, α, and ϵ refer to operator length of the prediction filter, prediction lag, and percent prewhitening, respectively. (The length of the prediction error filter then is $n + \alpha$.) In Figure 2-36, prediction lag is unity and equal to the 2-ms sampling rate; prewhitening is zero percent. Operator length varies as indicated in the figure. Short operators yield spikes with small-amplitude and relatively high-frequency tails. The 128-ms-long operator gives an almost perfect spike output. Longer operators whiten the spectrum further, bringing it closer to the spectrum of the impulse response.

The action of spiking deconvolution on the seismogram derived by convolving the minimum-phase wavelet with a sparse-spike series is similar (Figure 2-37) to the case of the single isolated wavelet (Figure 2-36). Recall that spiking deconvolution basically is inverse filtering where the operator is the least-squares inverse of the seismic wavelet. Therefore, an increasingly better result should be obtained when more and more coefficients are included in the inverse filter.

Now consider the real situation of an unknown source wavelet. Based on assumption 4, autocorrelation of the input seismogram rather than that of the seismic wavelet is used to design the deconvolution operator. The result of using the trace rather than the wavelet autocorrelation is shown in Figure 2-38. Deconvolution recovers the gross aspects of the spike series, trace a. However, note that the deconvolved traces have spurious small-amplitude spikes trailing each of the real spikes. We see that

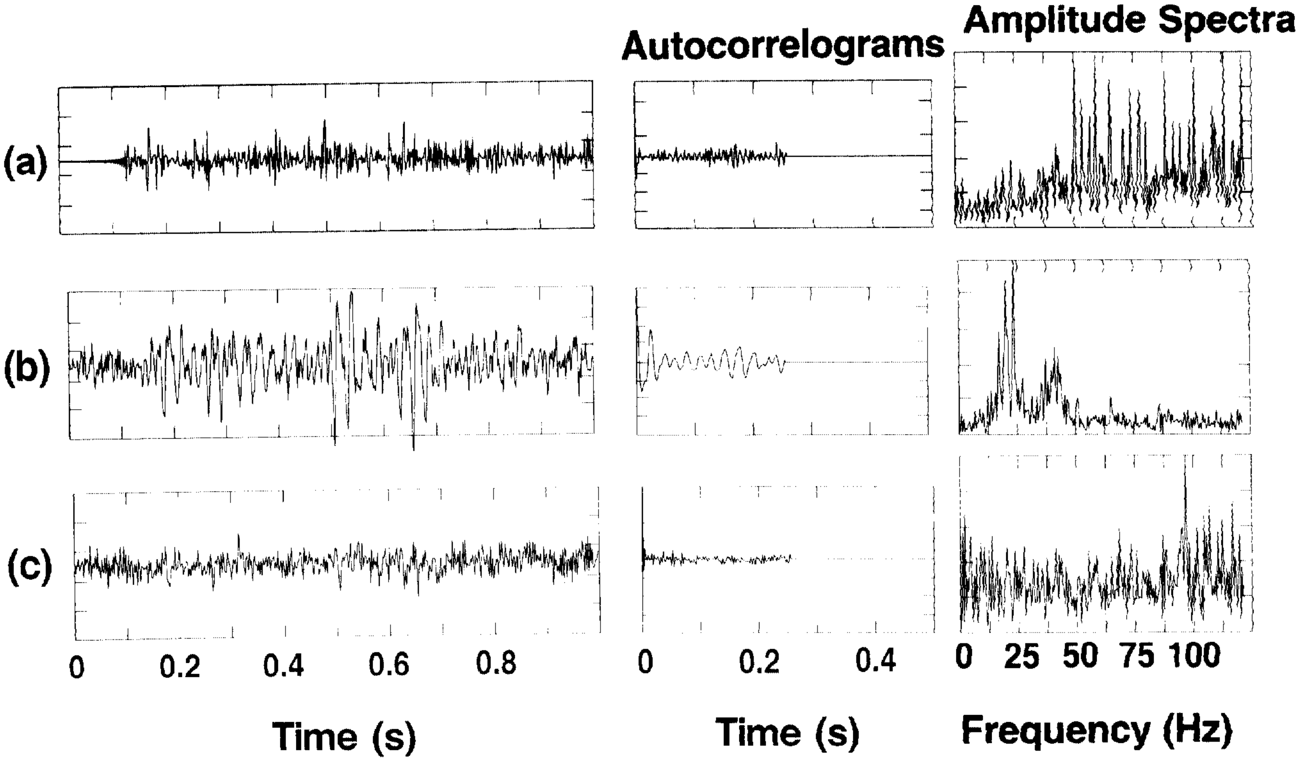


FIG. 2-35. (a) Impulse response, (b) seismogram with noise, (c) deconvolution assuming an unknown, mixed-phase source wavelet. Impulse response (a) is based on the sonic log of Figure 2-8a. In the presence of random noise (in violation of assumption 3), spiking deconvolution can produce a result with no relation to the earth's reflectivity [compare (a) to (c).]

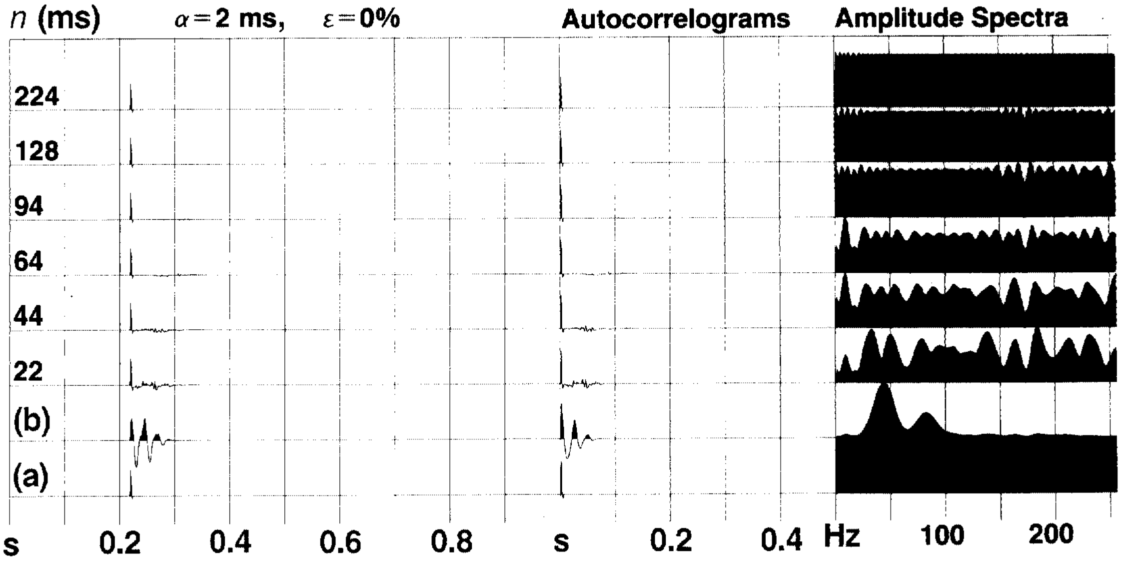


FIG. 2-36. Test of operator length for a single, isolated input wavelet, where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Impulse response. (b) seismogram with minimum-phase source wavelet.

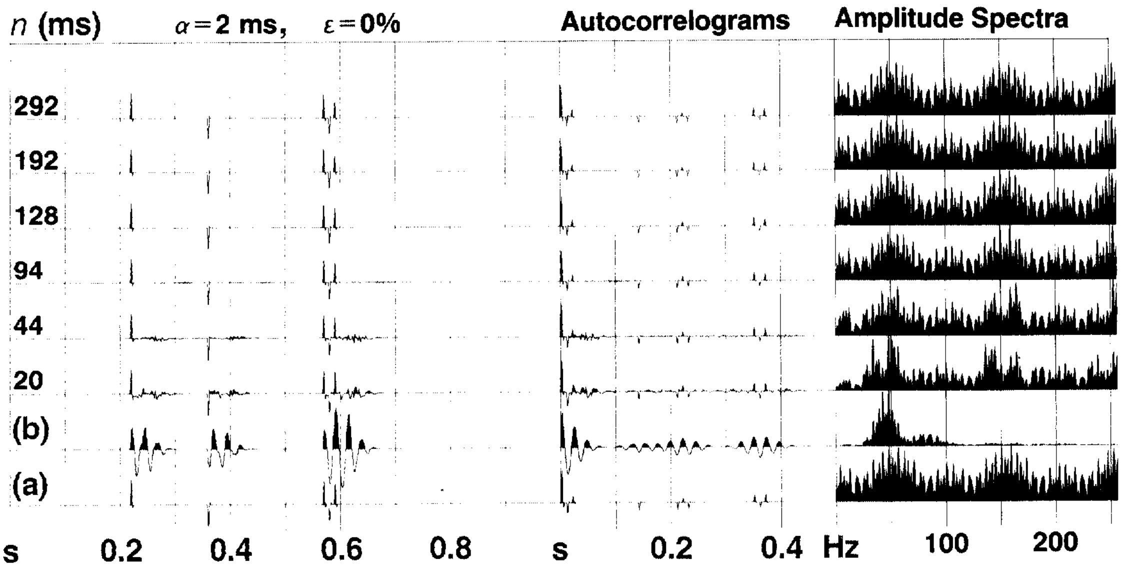


FIG. 2-37. Test of operator length where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Impulse response, (b) seismogram with known, minimum-phase source wavelet.

increasing operator length does not indefinitely improve the results; on the contrary, more and more spurious spikes are introduced. Very short operators produce the same type of noise spikes as in Figures 2-36 and 2-37. Examine the series of deconvolution tests in Figure 2-38 and note that the 94-ms operator does the best job. Compare the autocorrelogram of trace b in Figure 2-38 with that of trace b in Figure 2-36. Note that only the first 100 ms represent the autocorrelation of the source wavelet. This explains why the 94-ms operator worked best; that is, the autocorrelation lags (Figure 2-38, trace b) beyond 94 ms do not represent the seismic wavelet.

Consider the seismogram in Figure 2-39, where the wavelet is assumed to be unknown. Deconvolution has restored the spikes that correspond to major reflections in the impulse response, trace b, with some success. The 64-ms operator is a good choice.

The mixed-phase wavelet in Figure 2-40 shows what can happen when assumption 5 is violated. The wavelet in Figure 2-36 is the minimum-phase equivalent of the mixed-phase wavelet in Figure 2-40. Both wavelets have the same autocorrelograms and amplitude spectra. Hence, the deconvolution operators for both wavelets are identical. Because the minimum-phase assumption was violated, deconvolution does not convert the mixed-phase wavelet to a perfect spike. Instead, the deconvolved output is a complicated high-frequency wavelet. Also note that the dominant peak in the output is negative, while the impulse response has a positive spike. This difference in sign can happen when a mixed-phase wavelet is deconvolved. Increasing the operator length further whitens the spectrum; however, the 128-ms operator yields a result that cannot be improved further by longer operators.

The seismogram obtained from the mixed-phase wavelet and the sparse-spike series (used in the preceding figures) is shown in Figure 2-41. The 94-ms operator gives the best result. This also is the case in Figure 2-42, where both assumptions 4 and 5 are violated. The situation with the seismogram in Figure 2-43 is not so good. The spikes that correspond to major reflections in the impulse response were restored; however, there are some timing errors and polarity reversals. (Compare these results with those in Figure 2-39 for the events between 0.2 and 0.3 s and 0.6 and 0.7 s.) The 64-ms operator gives an output that cannot be improved by longer operators.

What kind of operator length should be used for spiking deconvolution? To select an operator length, ideally we want to use the autocorrelation of the unknown seismic wavelet. Fortunately, the autocorrelation of the input seismogram has the characteristics of the wavelet autocorrelation (assumption 4). Therefore, it seems appropriate that we should use the part of the autocorrelation obtained from the input seismogram that should most resemble the autocorrelation of the unknown seismic wavelet. That part is the first transient zone in the autocorrelation, as seen by comparing the autocorrelations of trace b in Figure 2-36 and trace c in Figure 2-39. The autocorrelations of trace b in Figure 2-40 and trace c in Figure 2-43 suggest the same principle.

2.7.2 Prediction Lag

So far, we have learned that predictive deconvolution has two uses: (a) spiking deconvolution, the case of unit prediction lag, and (b) predicting the input seismogram at a future time defined by the prediction lag. Case (b) is used to predict and suppress multiples. The effect of the prediction lag parameter now is examined from an interpretive point of view. Consider the single, isolated minimum-phase wavelet in Figure 2-44. Here, operator length

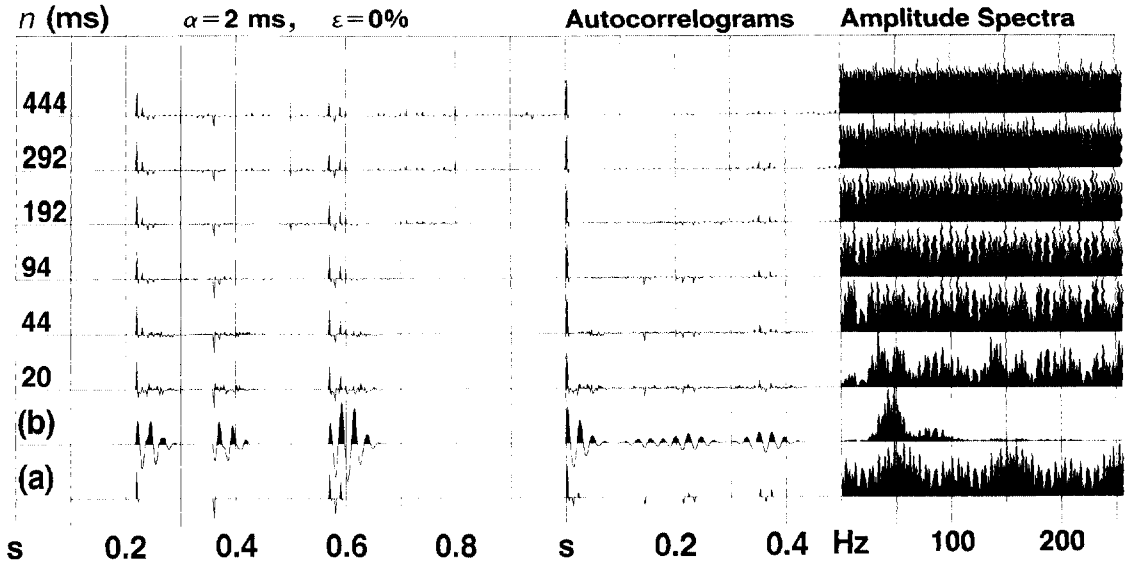


FIG. 2-38. Test of operator length where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Impulse response, (b) seismogram with unknown, minimum-phase source wavelet.

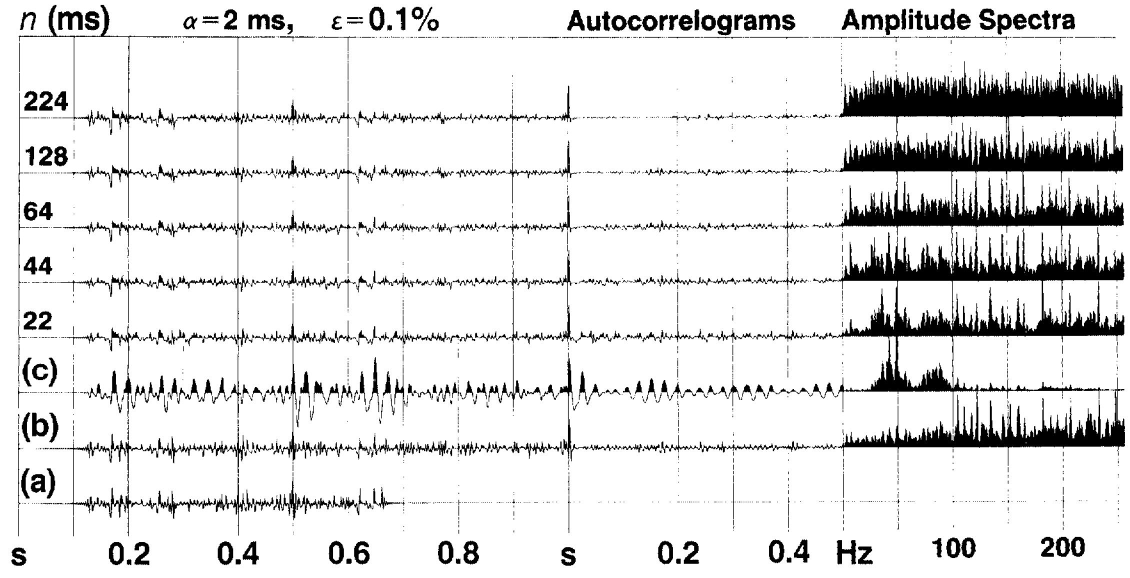


FIG. 2-39. Test of operator length where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Reflectivity, (b) impulse response, (c) seismogram with unknown, minimum-phase source wavelet.

and percent prewhitening are kept constant, while prediction lag is varied. When prediction lag is equal to the sampling rate, then the result is equivalent to spiking deconvolution. Predictive deconvolution using a prediction lag greater than unity yields a wavelet of finite duration instead of a spike. Given an input wavelet of $\alpha+n$ samples, predictive deconvolution using prediction filter with length n and prediction lag α converts this wavelet into another wavelet that is α samples long. The first α lags of the autocorrelation are preserved, while the next n lags are zeroed out. Additionally, the amplitude spectrum of the output increasingly resembles that of the input wavelet as prediction lag is increased (Figure 2-44). At a 94-ms prediction lag, predictive deconvolution does nothing to the input wavelet because almost all the lags of its autocorrelation have been left untouched. This experiment has an important practical implication: Under the ideal, noise-free conditions, resolution on the output from predictive deconvolution can be controlled by adjusting the prediction lag. Unit prediction lag implies the highest resolution, while a larger prediction lag implies less than full resolution. However, in reality, these assessments are dictated by the signal-to-noise ratio. The deconvolved output using a unit prediction lag contains

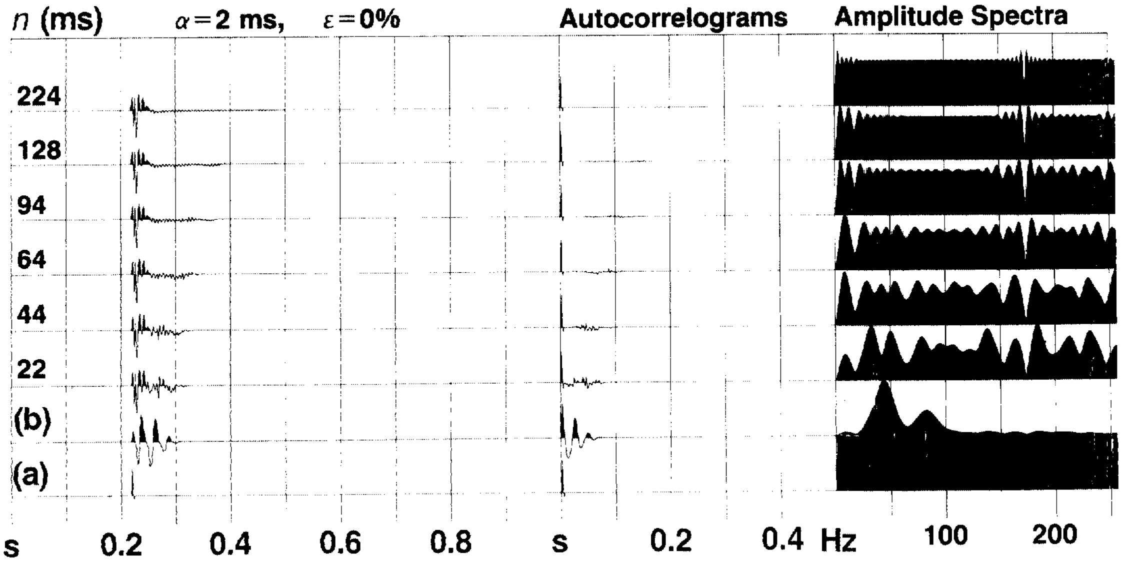


FIG. 2-40. Test of operator length for a single, isolated input wavelet where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Impulse response, (b) seismogram with mixed-phase source wavelet.

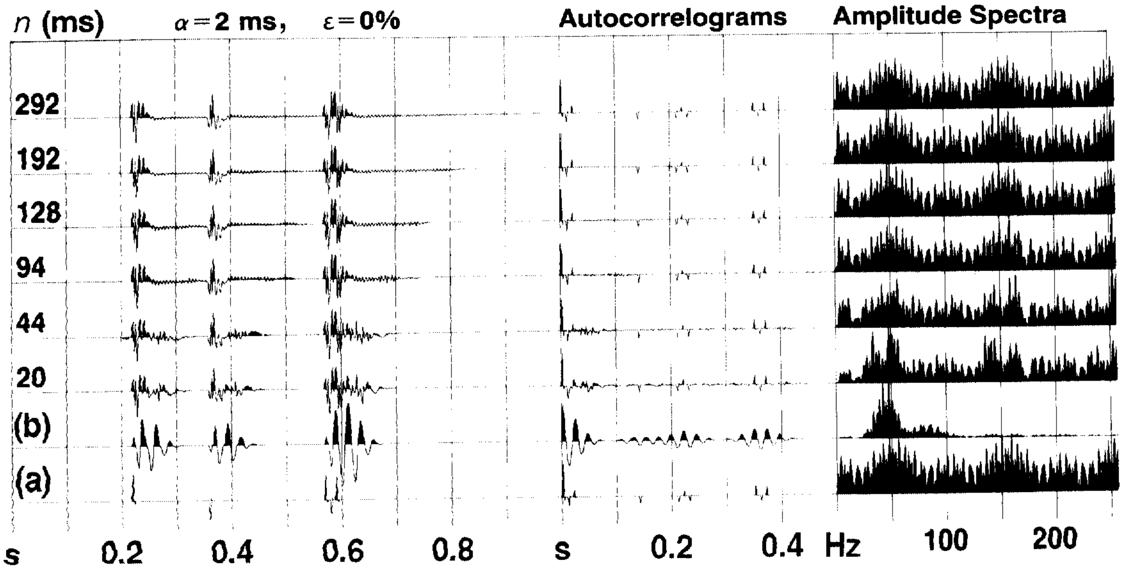


FIG. 2-41. Test of operator length where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Impulse response, (b) seismogram with known, mixed-phase source wavelet.

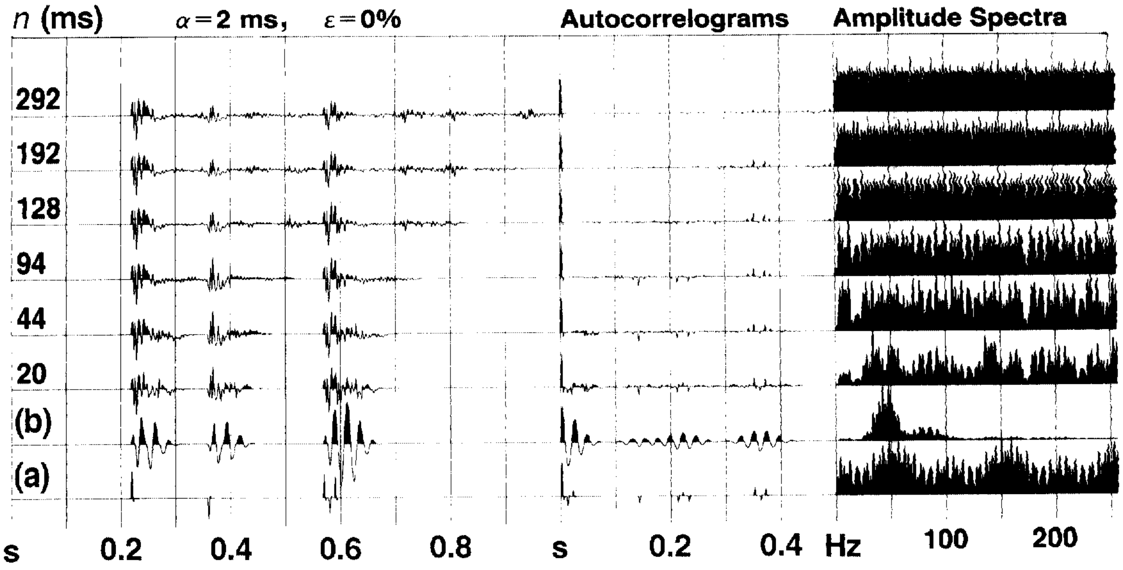


FIG. 2-42. Test of operator length where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Impulse response, (b) seimogram with unknown, mixed-phase source wavelet.

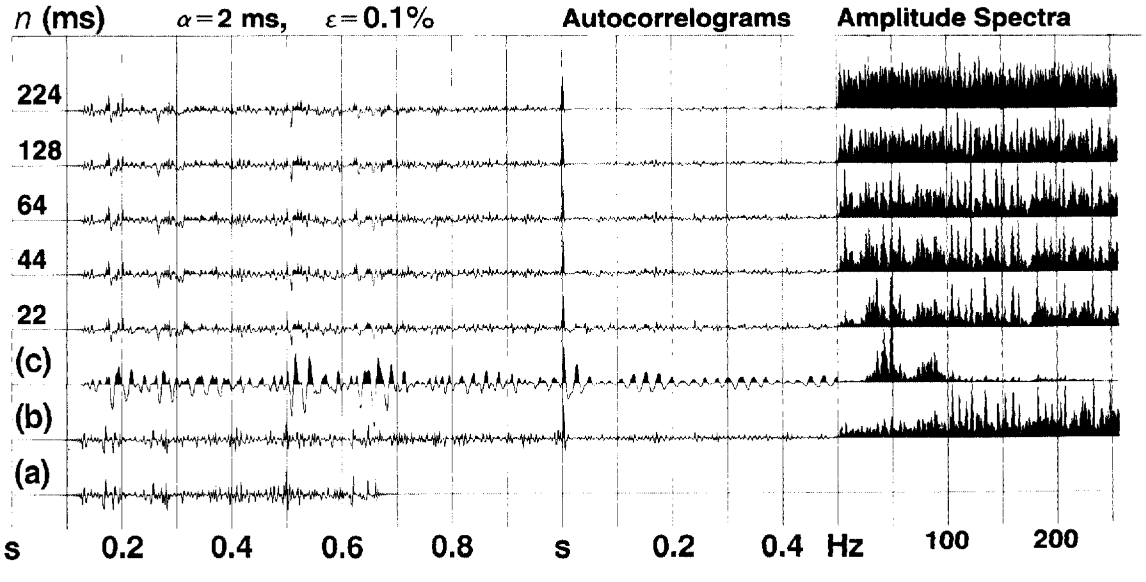


FIG. 2-43. Test of operator length where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Reflectivity, (b) impulse response, (c) seismogram with unknown, mixed-phase source wavelet.

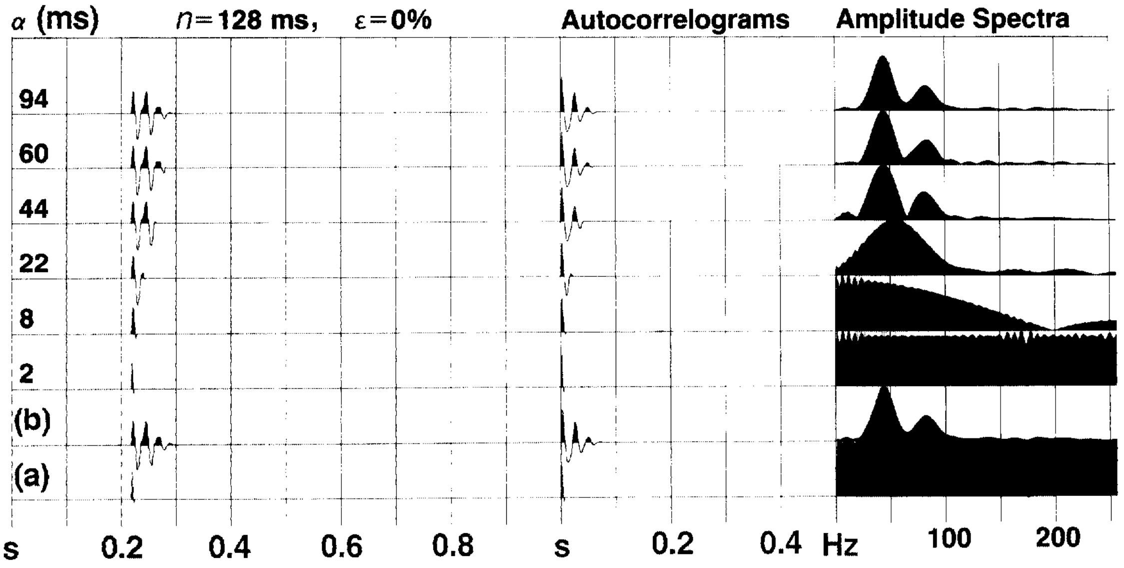


FIG. 2-44. Test of prediction lag for a single, isolated input wavelet where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Impulse response, (b) seismogram with minimum-phase source wavelet.

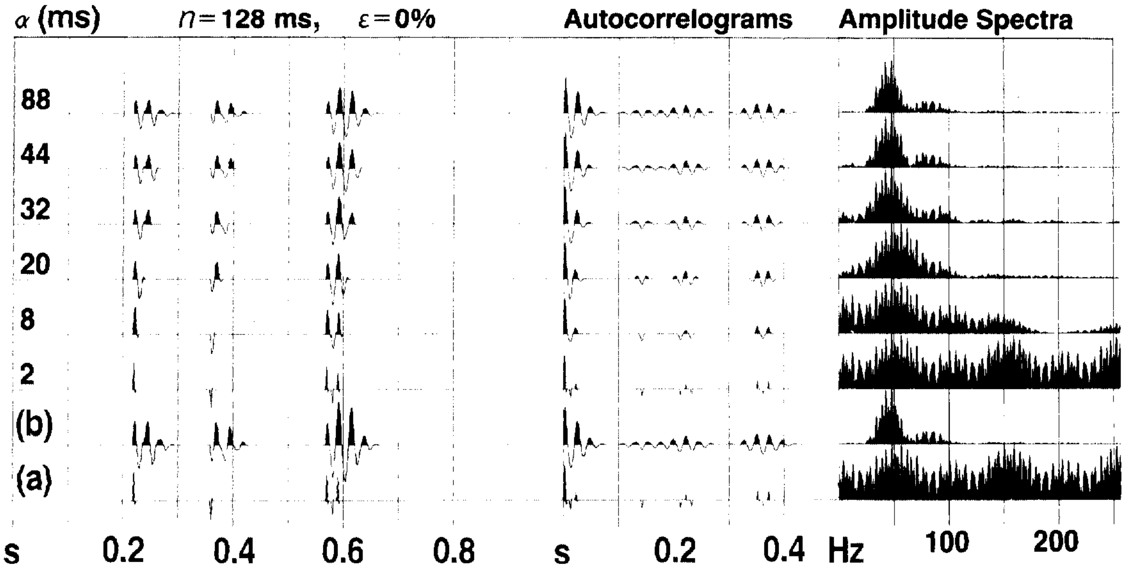


FIG. 2-45. Test of prediction lag where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Impulse response, (b) seismogram with known, minimum-phase source wavelet.

high frequencies; nevertheless, resolution may be degraded if the high-frequency energy is mostly noise, not signal.

In Figure 2-44, prediction lags of 8 and 22 ms correspond to the first and second zero crossings on autocorrelation of the input wavelet, respectively. The first zero crossing produces a spike with some width, while the second zero crossing lag produces a wavelet with a positive and negative lobe.

The relationship between prediction lag and whitening also holds for the sparse-spike series in Figure 2-45 and when the input wavelet is unknown (Figure 2-46). The effect of prediction lag on the output from predictive deconvolution of a synthetic seismogram, which was obtained from the sonic log (Figure 2-8a), is demonstrated in Figures 2-47 and 2-48. As the prediction lag increases, the output spectrum becomes increasingly less broadband. Predictive deconvolution of seismograms constructed from the mixed-phase wavelet again demonstrates that output resolution can be controlled by adjusting prediction lag (Figures 2-49 through 2-53).

If prediction lag is increased, then the output amplitude spectrum becomes increasingly band-limited. The output also can be band-limited by applying a band-pass filter on

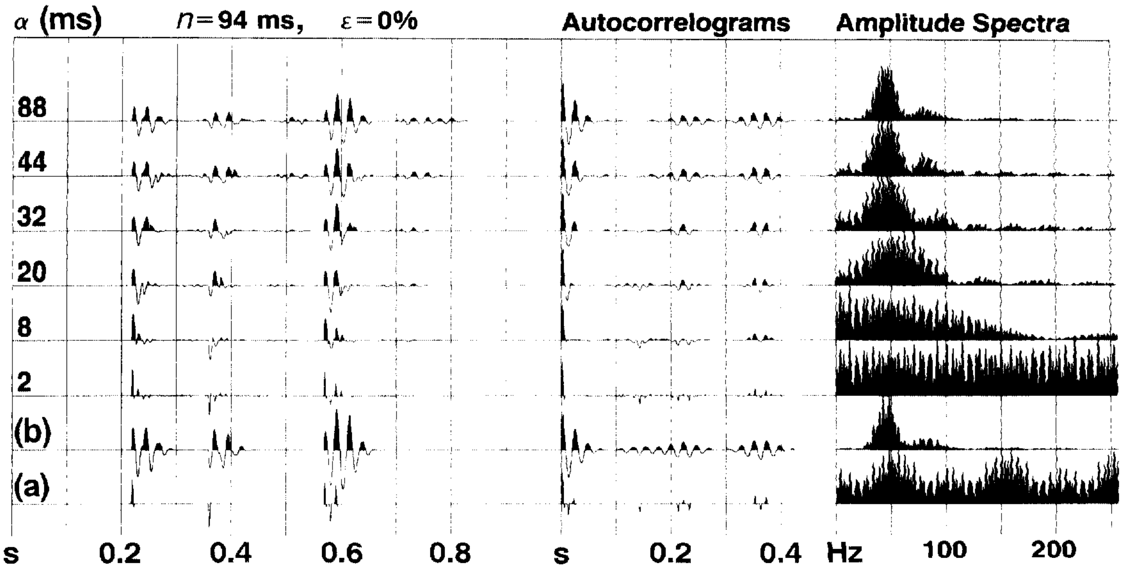


FIG. 2-46. Test of prediction lag where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Impulse response, (b) seismogram with unknown, minimum-phase source wavelet.

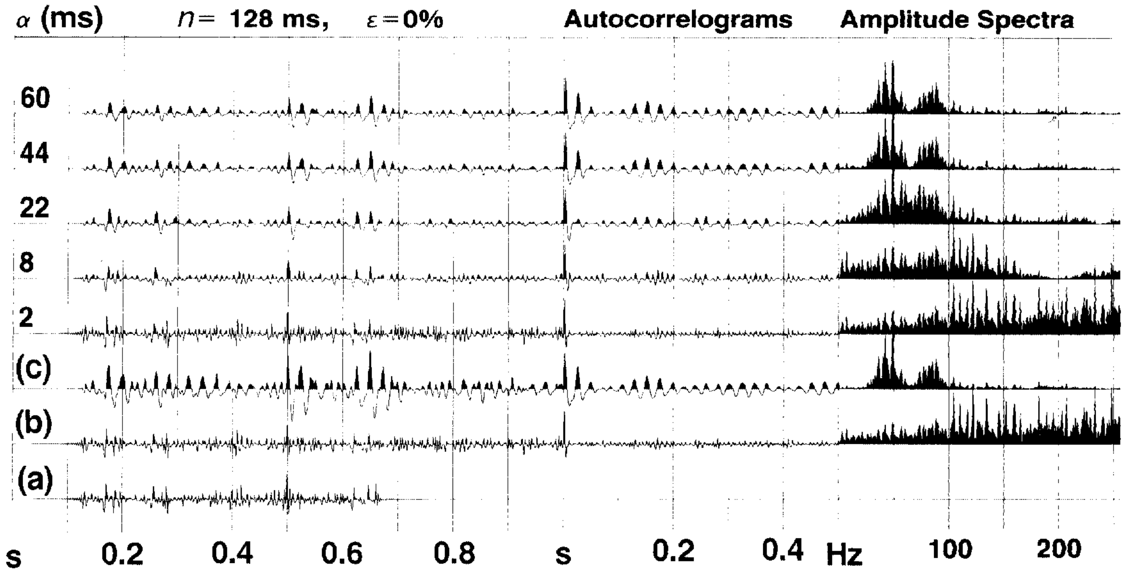


FIG. 2-47. Test of prediction lag where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Reflectivity, (b) impulse response, (c) seismogram with known minimum-phase source wavelet.

the spiking deconvolution output. Are these two ways of band-limiting equivalent? Refer to the results from both the minimum-and mixed-phase wavelets in Figures 2-44 and 2-49, respectively. Note that the output of the 22-ms prediction lag has an amplitude spectrum that is band-limited to approximately 0 to 100 Hz. However, the spectral shape within this bandwidth is not a boxcar, but rather similar to that of the input wavelet. The boxcar shape would be the case if a band-pass filter (0 to 100 Hz) were applied on the output of the spiking deconvolution (2-ms prediction lag). Hence, spiking deconvolution followed by band-pass filtering is not equivalent to predictive deconvolution with a prediction lag greater than unity.

In conclusion, if prediction lag is increased, the output from predictive deconvolution becomes less spiky. This effect can be used to our advantage, since it allows the bandwidth of deconvolved output to be controlled by adjusting prediction lag. The application of spiking deconvolution to field data is not always desirable, since it boosts high-frequency noise in the data. The most prominent effect of the nonunity prediction lag is suppression of the high-frequency end of the spectrum and preservation of the overall spectral shape of the input data. This

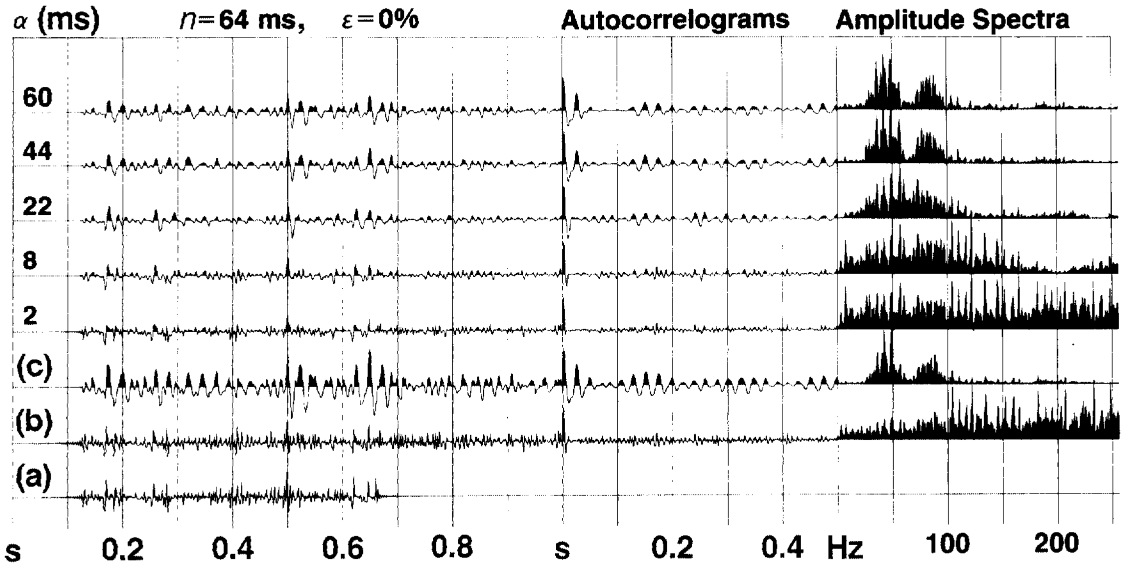


FIG. 2-48. Test of prediction lag where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Reflectivity, (b) impulse response, (c) seismogram with unknown, minimum-phase source wavelet.

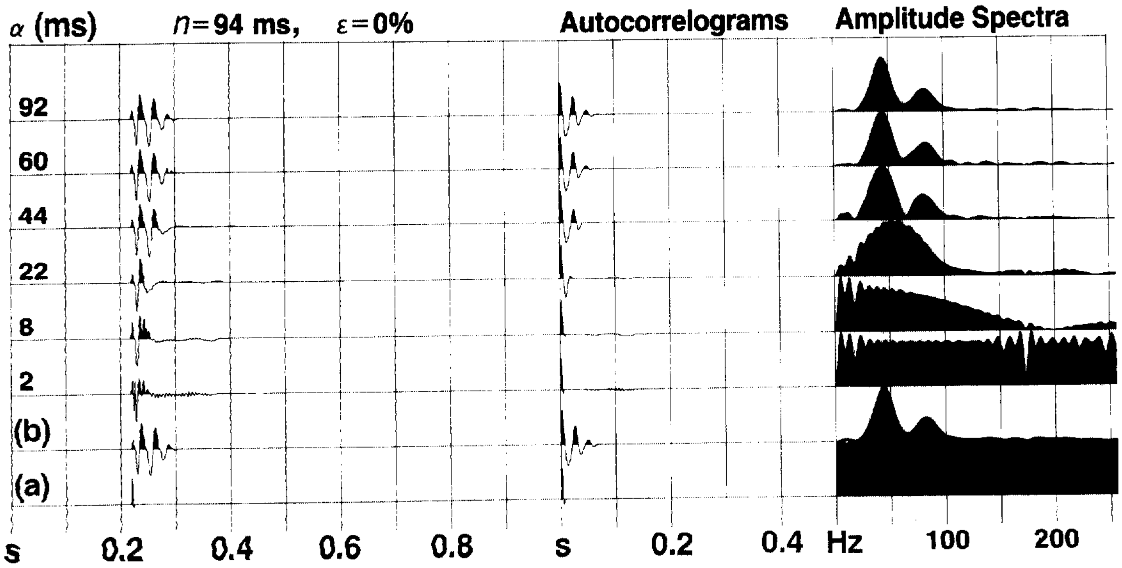


FIG. 2-49. Test of prediction lag for a single, isolated input wavelet where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Impulse response, (b) seismogram with mixed-phase source wavelet.

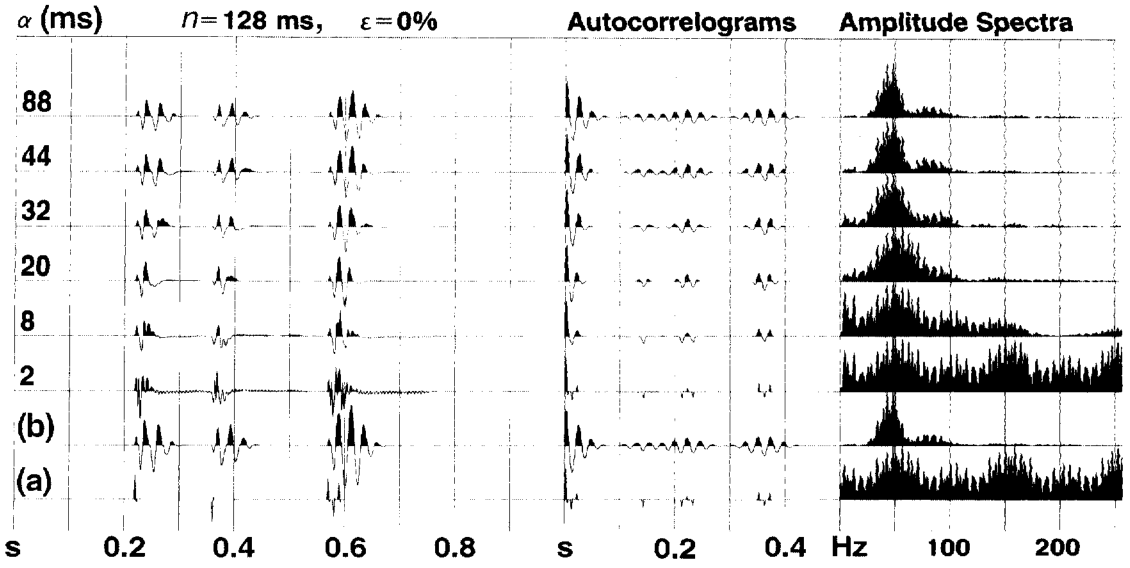


FIG. 2-50. Test of prediction lag where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Impulse response, (b) seismogram with known, mixed-phase source wavelet.

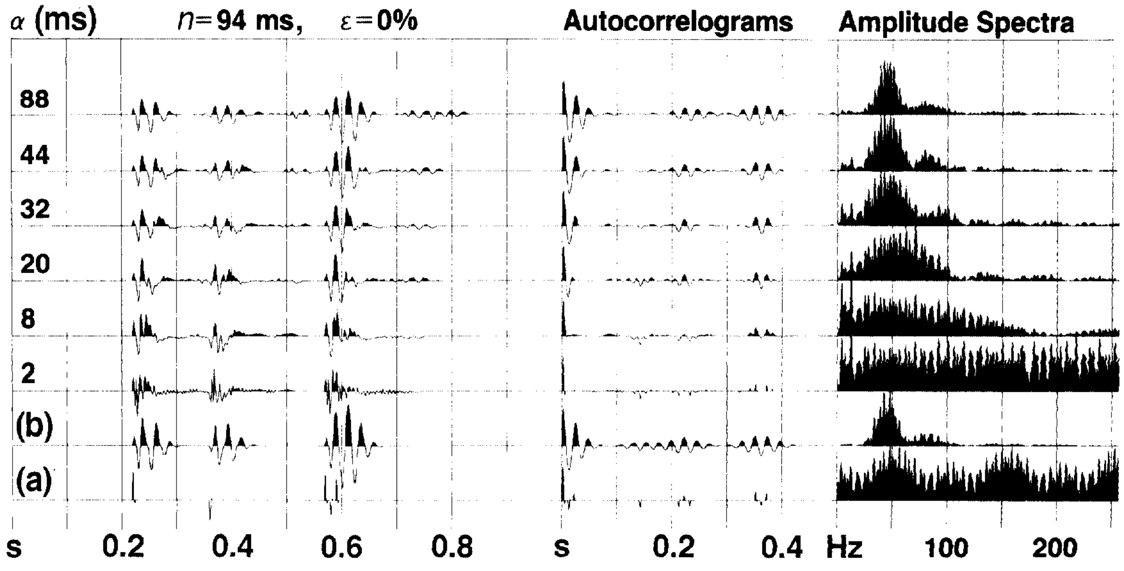


FIG. 2-51. Test of prediction lag where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Impulse response, (b) seismogram with unknown, mixed-phase source wavelet.

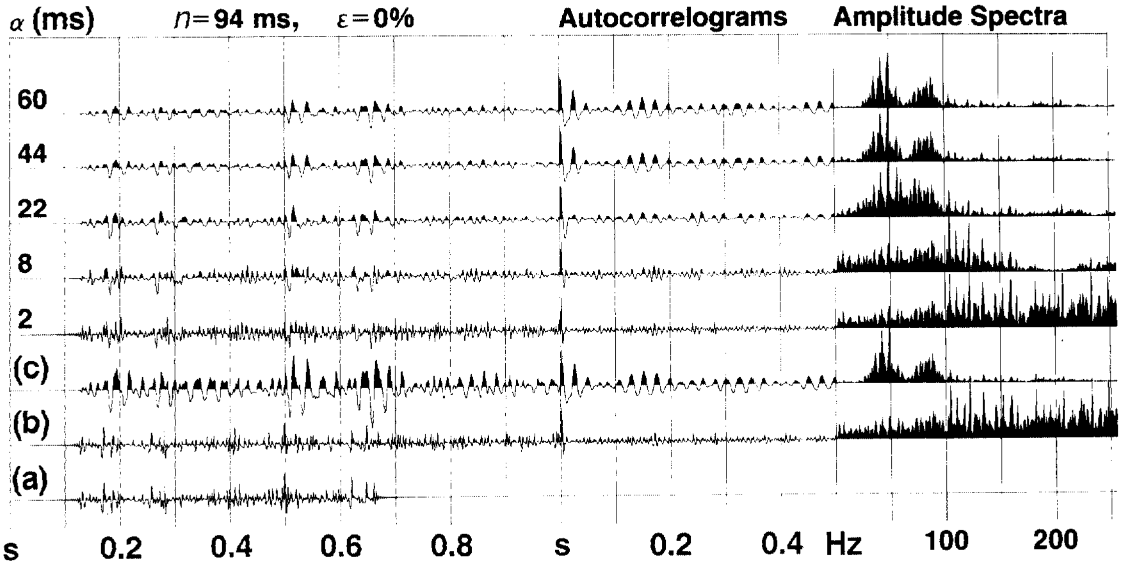


FIG. 2-52. Test of prediction lag where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Reflectivity, (b) impulse response, (c) seismogram with known, mixed-phase source wavelet.

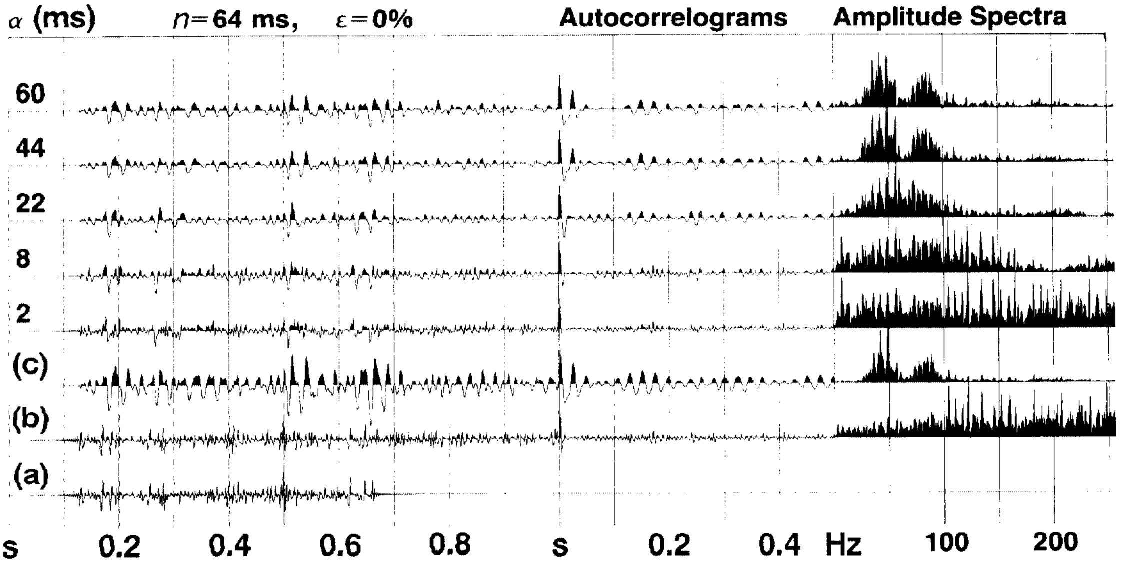


FIG. 2-53. Test of prediction lag where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Reflectivity, (b) impulse response, (c) seismogram with unknown, mixed-phase source wavelet.

effect is seen in Figures 2-48 and 2-53, which correspond to the minimum-and mixed-phase seismic wavelets. If prediction lag is increased further, then the low-frequency end of the spectrum is affected as well, making the output more band-limited.

2.7.3 Percent Prewhitening

The reasons for prewhitening were discussed in Section 2.6.2. Consider the single, isolated minimum-phase wavelet in Figure 2-54. Keep the operator length and prediction lag constant and vary the percent prewhitening. Note that the effect of varying prewhitening is similar to that of varying the prediction lag; that is, the spectrum increasingly becomes less broadband as the percent prewhitening is increased. Compare Figure 2-44 with Figure 2-54. Note that prewhitening narrows the spectrum without changing much of the flatness character, while larger prediction lag narrows the spectrum and alters its shape, making it look more like the spectrum of the input seismic wavelet. These characteristics also can

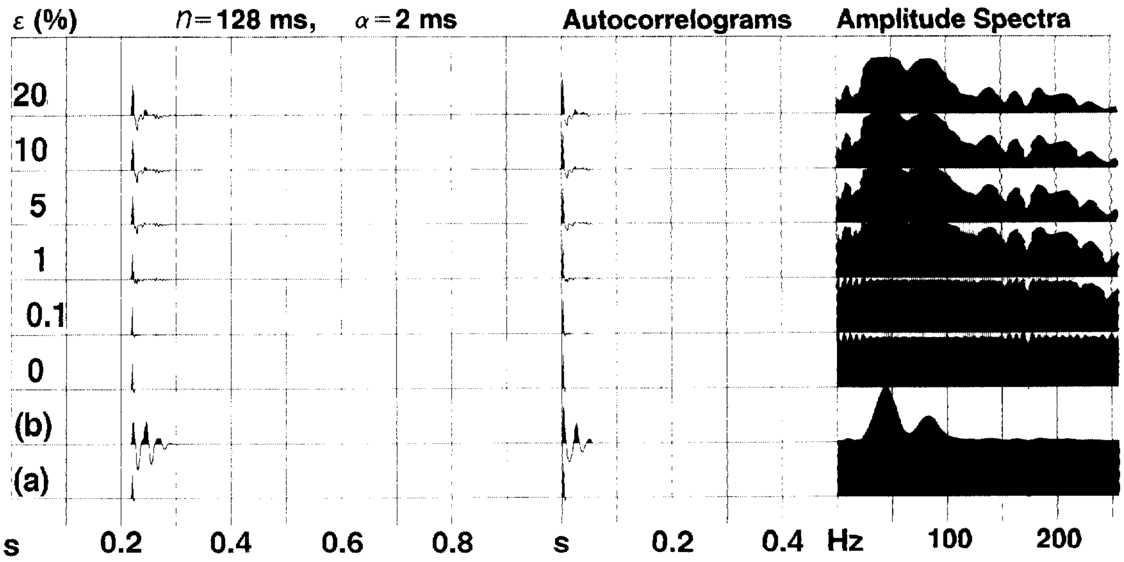


FIG. 2-54. Test of percent prewhitening for a single, isolated input wavelet where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Impulse response, (b) seismogram with minimum-phase source wavelet.

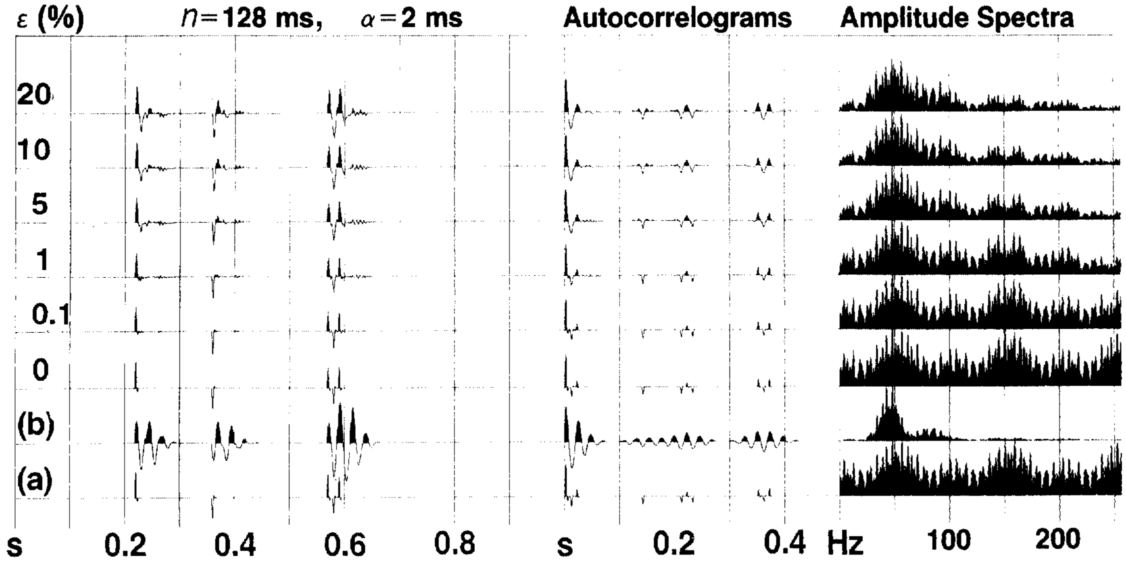


FIG. 2-55. Test of percent prewhitening where n = operator length, α = prediciton lag, and ε = percent prewhitening. (a) Impulse response, (b) seismogram with known, minimum-phase source wavelet.

be inferred from the shapes of the output wavelets. Prewhitening preserves the spiky character of the output, although it adds a low-amplitude, high-frequency tail (Figure 2-54). On the other hand, increasing prediction lag produces a wavelet with a duration equal to the prediction lag (Figure 2-44).

The effect of prewhitening on the sparse-spike train seismogram with a known and unknown minimum-phase wavelet is shown in Figures 2-55 and 2-56, respectively. The effect of prewhitening on deconvolution of the synthetic seismogram obtained from the sonic log (Figure 2-8a) is shown in Figures 2-57 and 2-58 for known and unknown minimum-phase wavelets. Prewhitening tests using the mixed-phase wavelet are shown in Figure 2-59. Finally, the combined effects of a prediction lag that is greater than unity and prewhitening for the single, isolated wavelet are shown in Figure 2-60. These figures demonstrate that prewhitening narrows the output spectrum, making it band-limited. In particular, the tests in Figures 2-54 and 2-59 using the single, isolated minimum- and mixed-phase wavelets suggest that spiking deconvolution with some prewhitening is somewhat equivalent to

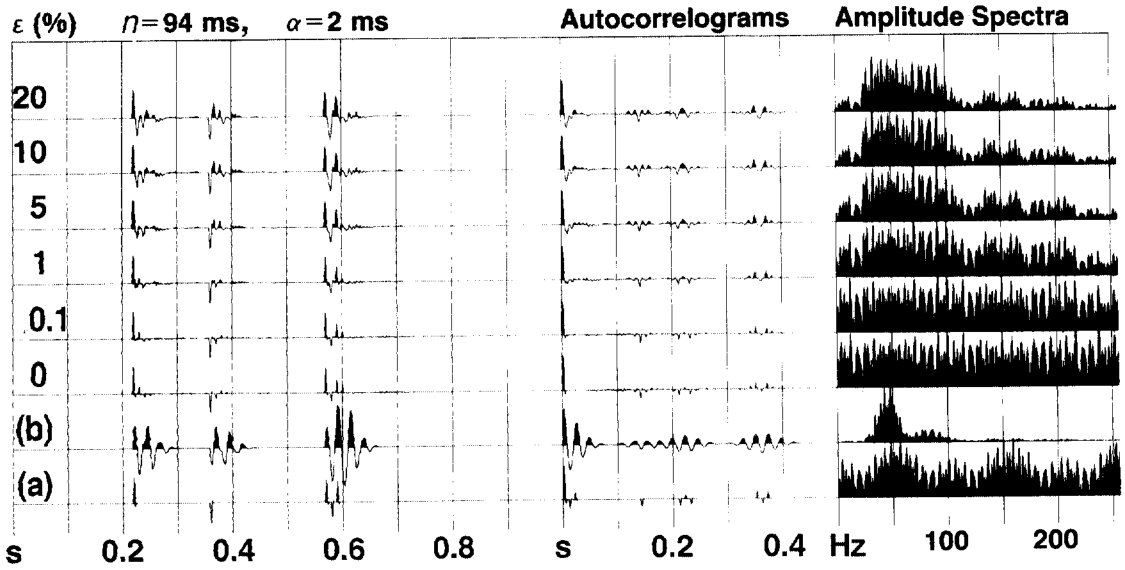


FIG. 2-56. Test of percent prewhitening where *n* = operator length, α = prediction lag, and ε = percent prewhitening. (a) Impulse response, (b) seismogram with unknown, minimum-phase source wavelet.

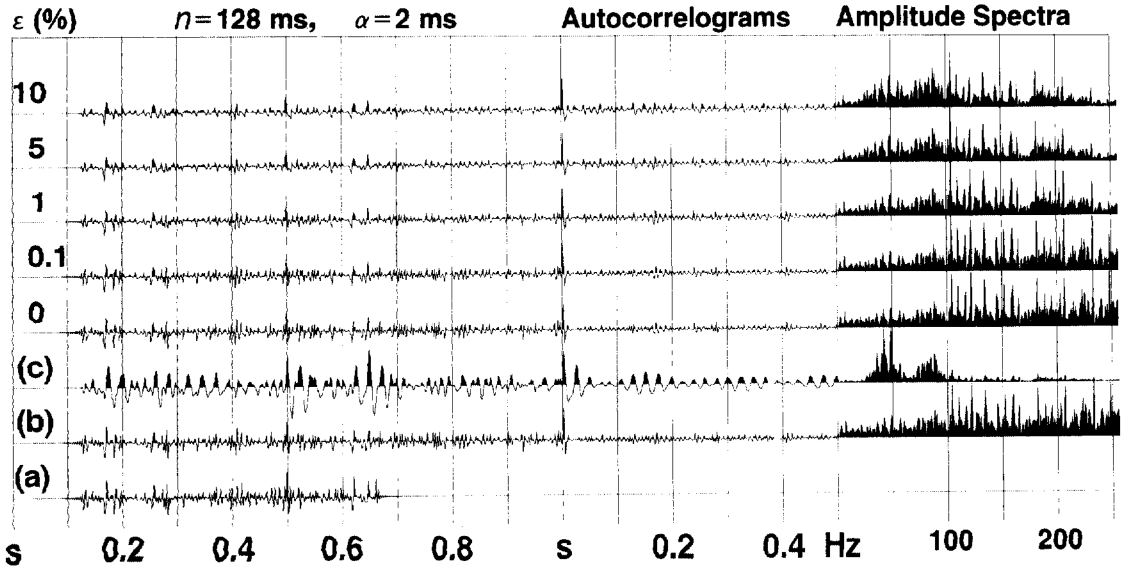


FIG. 2-57. Test of percent prewhitening where *n* = operator length, α = prediction lag, and ε = percent prewhitening. (a) Reflectivity, (b) impulse response, (c) seismogram with known, minimum-phase source wavelet.

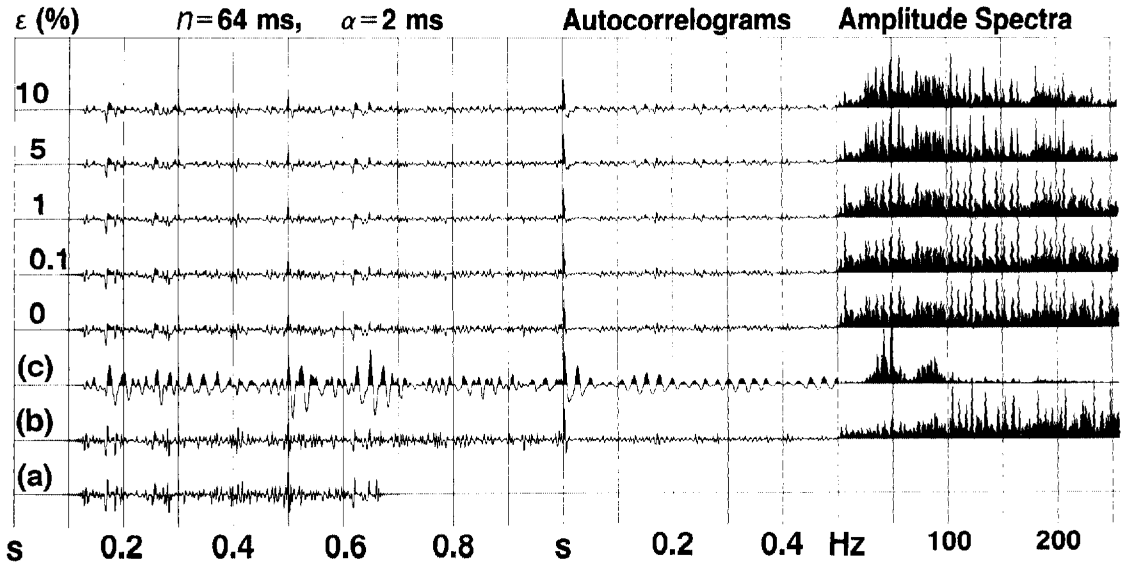


FIG. 2-58. Test of percent prewhitening where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Reflectivity, (b) impulse response, (c) seismogram with unknown, minimum-phase source wavelet.

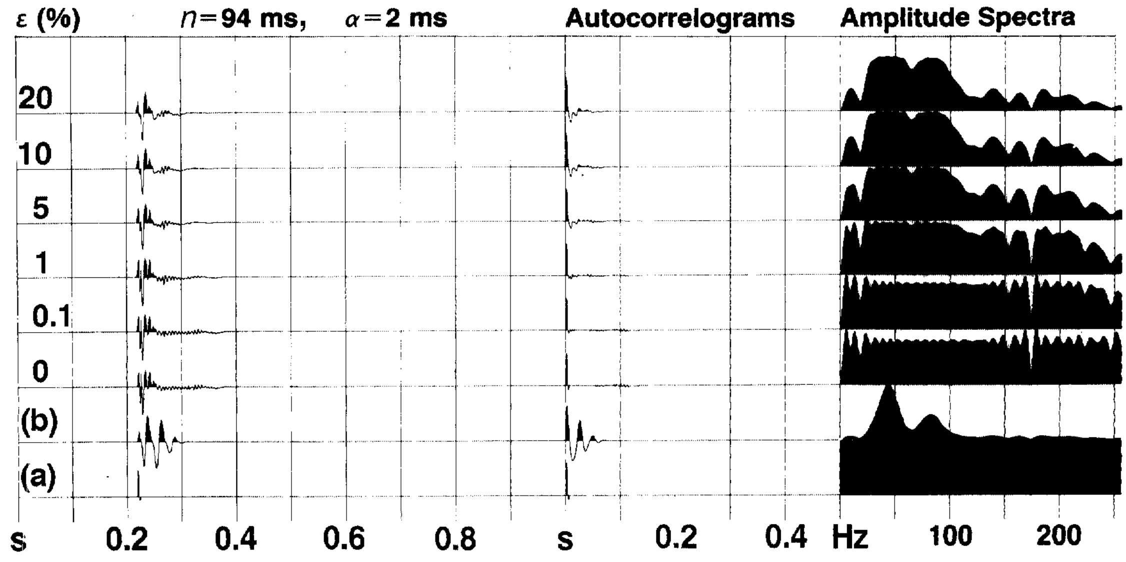


FIG. 2-59. Test of percent prewhitening for a single, isolated input wavelet where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Impulse response, (b) seismogram with mixed-phase source wavelet.

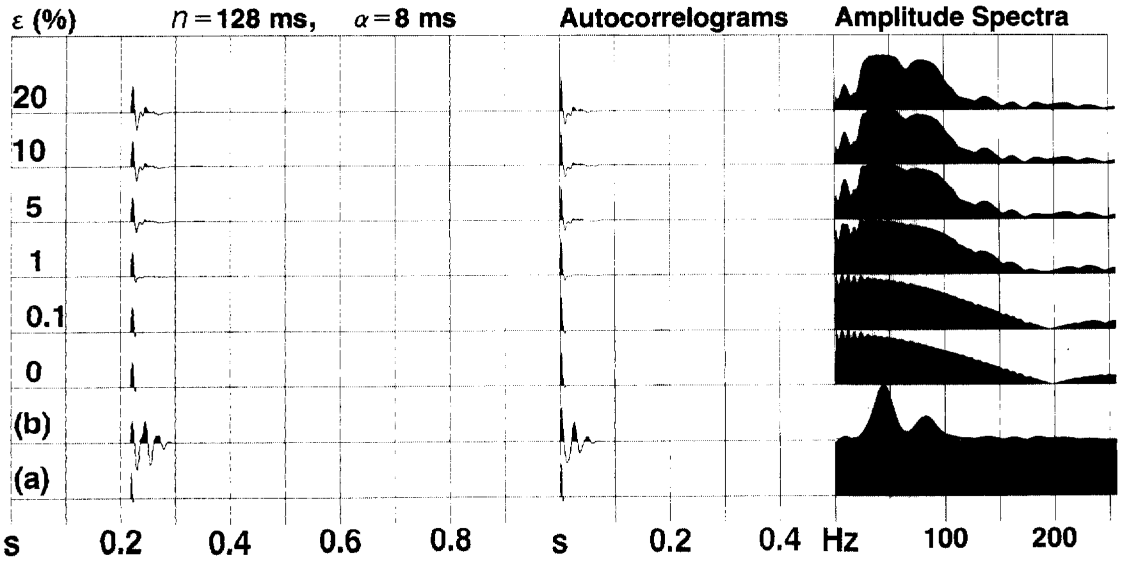


FIG. 2-60. Test of percent prewhitening for a single, isolated input wavelet where n = operator length, α = prediction lag, and ε = percent prewhitening. (a) Impulse response, (b) seismogram with minimum-phase source wavelet.

spiking deconvolution without prewhitening followed by post-deconvolution broad band-pass filtering. However, this is not exactly true, for prewhitening still leaves some relatively suppressed energy at the high-frequency end of the spectrum. From Figure 2-60, we infer that predictive deconvolution with a prediction lag greater than unity and with some prewhitening yields a result somewhat equivalent to a spiking deconvolution followed by band-pass filtering.

In conclusion, we can say that prewhitening yields a band-limited output. However, the effect is less controllable when compared to varying the prediction lag. By varying prediction lag, we have some idea of the output bandwidth, since it is related to prediction lag. The smaller the prediction lag, the broader the output bandwidth. Prewhitening is used only to ensure that numerical instability in solving the Toeplitz matrix is avoided. In practice, typically 0.1 to 1 percent prewhitening is standard.

2.7.4 Effect of Random Noise on Deconvolution

We assume that the noise component in the recorded seismogram is zero (assumption 3). The autocorrelation of ideal random noise is zero at all lags except the zero lag (Figure 2-12). Therefore, the effect of random noise on deconvolution operators should be somewhat similar to the effect of prewhitening. Both effects modify the diagonal of the autocorrelation matrix, making it more dominant (equations (2.22)). However, the noise component also slightly modifies the nonzero lags of the autocorrelation. Compare the autocorrelograms of traces b in Figures 2-54 and 2-61. In Figure 2-54, an isolated minimum-phase wavelet was considered, while in Figure 2-61, random noise was added to the same wavelet. The output wavelet shape from spiking deconvolution of the noisy wavelet using a 128-ms operator is similar to the output from spiking deconvolution of the wavelet without noise, using the same operator length but with, say, 20 percent prewhitening. This result has practical importance: Prewhitening is equivalent to adding perfect random noise to the system. Since random noise usually is in the system, only a minute amount, say 0.1 percent, of the white noise needs to be added to the seismogram for numerical stability.

The effect of random noise on the performance of deconvolution is examined further in Figures 2-62 and 2-63. These results should be compared with their noiseless counterparts in Figures 2-39 and 2-43, respectively. Observe that the noise component has a harmful effect on deconvolution. For example, when comparing Figures 2-39 and 2-62, note that the deconvolution result from the noisy seismogram has spurious spikes (e.g., between 0.5 and 0.6 s), which could be interpreted as genuine reflections. Noisy field data, which yield better stack when not treated by deconvolution, have been noted. Only by testing can we determine whether deconvolution performs satisfactorily on data with a severe noise problem.

2.7.5 Multiple Suppression

We have learned that a prediction filter predicts periodic events, like multiples, in the seismogram. The prediction error filter yields the unpredictable component of the seismogram; i.e., the reflectivity series. For example, consider the simple case of water-bottom multiples. If the reflection coefficient of the water bottom is c_w and if

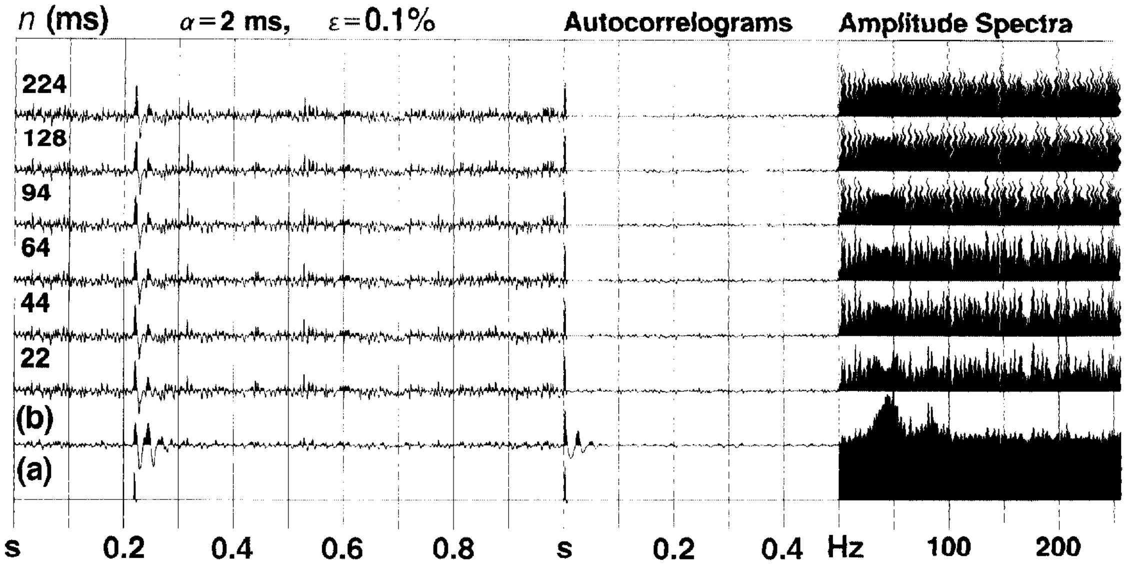


FIG. 2-61. Effect of random noise on deconvolution performance. Input seismogram (b) associated with reflectivity (a) contains a single, isolated wavelet (at around 0.2 s) buried in random noise. Here n = operator length, α = prediction lag, and ε = percent prewhitening.

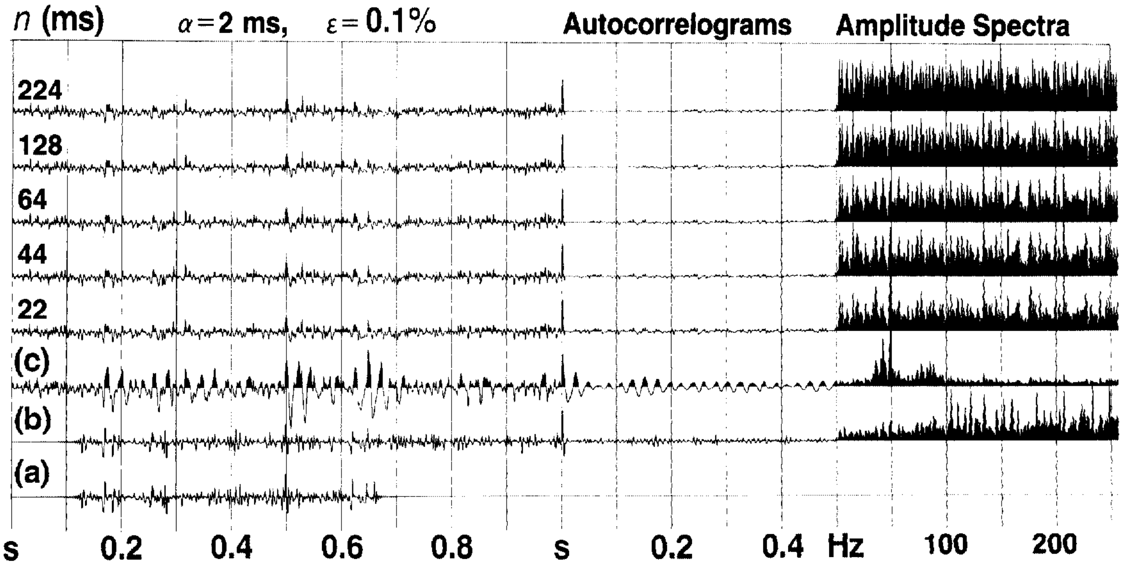


FIG. 2-62. Effect of random noise on deconvolution performance. (a) Reflectivity, (b) impulse response. Input seismogram (c) with unknown, minimum-phase source wavelet is noise contaminated. Here n = operator length, α = prediction lag, and ε = percent prewhitening.

water depth is equivalent to a two-way time t_w, then the time series is:

$$(1, 0, \ldots, 0, -c_w, 0, \ldots, 0, c_w^2, 0, \ldots, 0, -c_w^3, \cdots)$$

as shown in Figure 2-64, traces b and c. The separation between the spikes is t_w in trace b. Note that the periodicity in the time series (trace b or c) manifests itself in the amplitude spectrum as periodic peaks (or notches). The greater the spike separation in time, the closer the peaks (or notches) in the amplitude spectrum.

The noise-free convolutional model for the seismogram that contains the water-bottom multiples can be written as:

$$x(t) = w(t) * m(t) * e(t), \tag{2.30}$$

where $m(t)$ represents the water-layer reverberation spike train (Figure 2-64, trace b), and $e(t)$ now represents the earth's impulse response excluding multiples associated with the water bottom. Predictive deconvolution can suppress this periodic component in the seismogram as

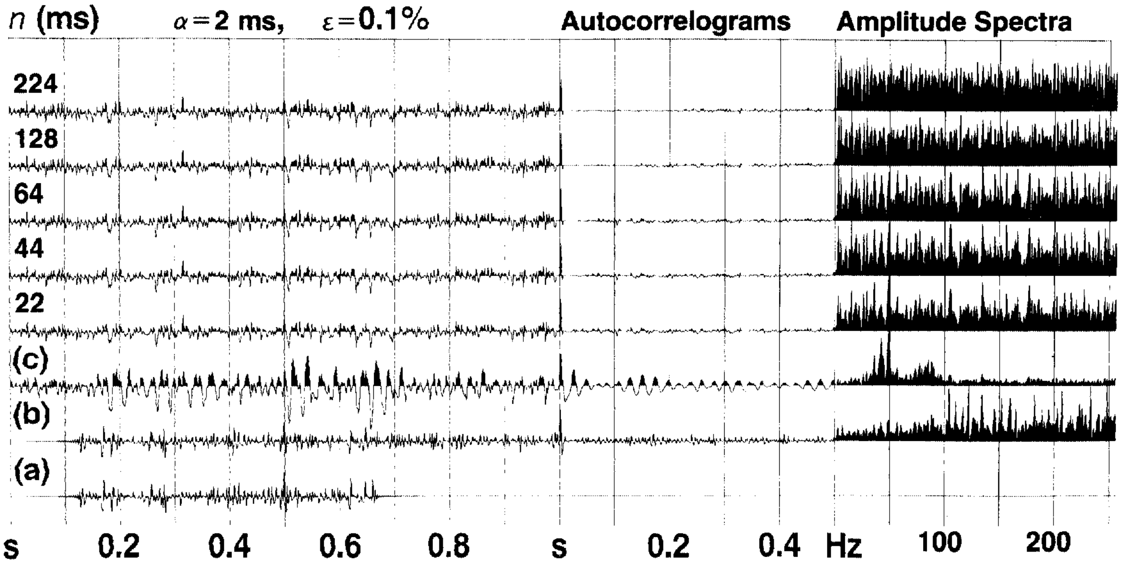


FIG. 2-63. Effect of random noise on deconvolution performance. (a) Reflectivity, (b) impulse response. Input seismogram (c) with unknown, mixed-phase source wavelet is noise contaminated. Here n = operator length, α = prediction lag, and ε = percent prewhitening.

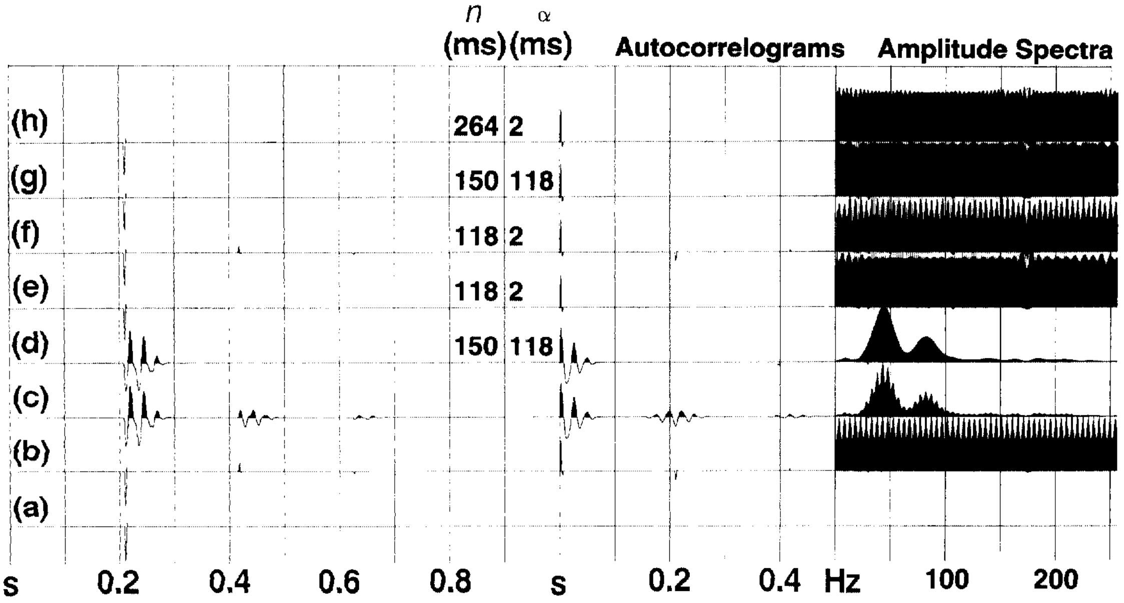


FIG. 2-64. (a) Reflectivity, (b) impulse response, (c) seismogram. Two-step deconvolution aimed at suppressing the multiple wave train, then spiking the remaining primary wavelet (d) to (e). The process can be performed in the reverse order (f) to (g). Here, n = operator length and α = prediction lag.

demonstrated by trace d. Note the two distinct goals for predictive deconvolution: (a) spiking the seismic wavelet $w(t)$, and (b) predicting and suppressing multiples $m(t)$. The first goal is achieved using an operator with unit prediction lag, while the second is achieved using an operator with a prediction lag greater than unity. The autocorrelation of the input trace can be used to determine the appropriate prediction lag for multiple suppression. Periodicity due to multiples is evident in the autocorrelogram of trace c in Figure 2-64, as an isolated series of energy packets in the neighborhood of 0.2 and 0.4 s. Prediction lag should be chosen to bypass the first part of the autocorrelogram that represents the seismic wavelet. Operator length should be chosen to include the first isolated energy packet in the autocorrelogram. After applying predictive deconvolution, we are left with only the water-bottom primary reflection. Isolated bursts in the autocorrelogram have been suppressed, while periodic peaks in the amplitude spectrum have been eliminated (trace d). If desired, the basic wavelet can be compressed

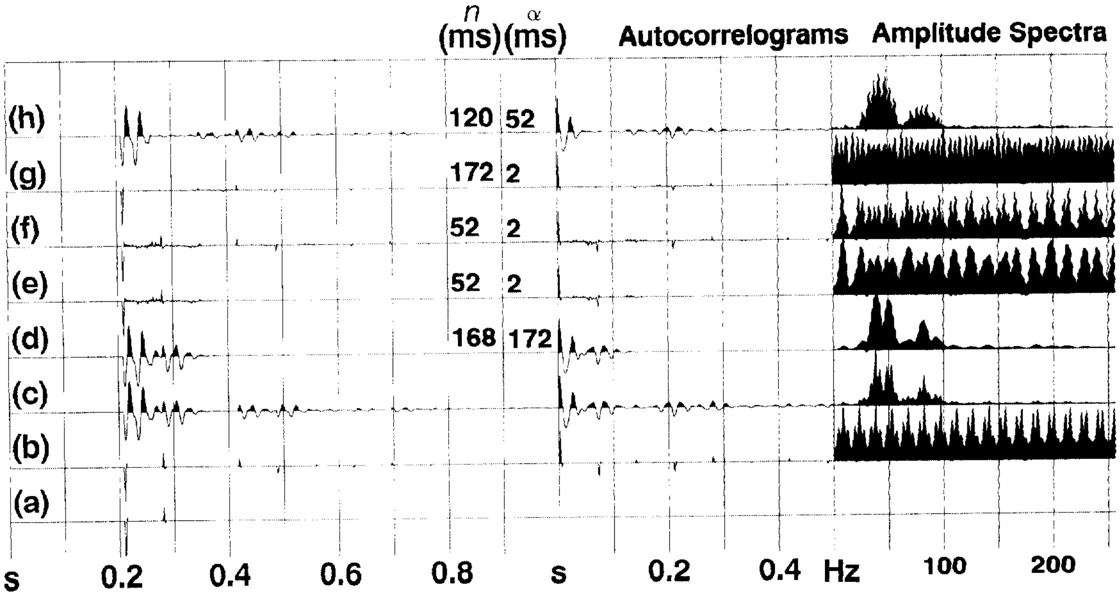


FIG. 2-65. Predictive deconvolution for multiple suppression. (a) Reflectivity, (b) impulse response, (c) seismogram. Two-step deconvolution: Predictive deconvolution (d) followed by spiking deconvolution (e). Traces (f), (g), and (h) result when single-step deconvolution is applied to the input trace (c), using the operator lengths *n* and prediction lags α, as indicated.

into a spike (trace e) by applying spiking deconvolution to the output of the predictive deconvolution (trace d). The sequence can be interchanged by first applying spiking deconvolution (trace f) followed by predictive deconvolution (trace g).

By using a sufficiently long spiking deconvolution operator, two goals are achieved in one step as seen in trace h. However, this approach can be dangerous if genuine reflections are unintentionally suppressed. This is the case in Figure 2-65. Here, the water-bottom reflection is followed by a deeper event at about 0.28 s (trace a). The impulse response contains water-bottom multiples and the peg-leg multiples that are due to the deeper reflector (trace b). The amplitude spectrum has peaks that come in pairs, indicating the presence of two different periodic components in the seismogram. Careful choice of predictive deconvolution parameters yields a trace with only the wavelets associated with the water bottom and deep reflector (trace d). This is followed by a spiking deconvolution that yields two spikes representative of the water bottom and deep primary (trace e). Spiking deconvolution alone produces the reflection coefficient series and the spikes that represent the multiples (trace f). If a longer spiking deconvolution operator is used, then the primary reflection easily can be eliminated (trace g). If a predictive deconvolution operator is used with an improper parameter choice, then again the primary reflection easily can be eliminated (trace h).

How can we ensure that no primaries are destroyed by deconvolution? Examine the autocorrelogram of trace c in Figure 2-65. The first 50 ms represent the seismic wavelet. This is followed by a burst between 50 to 170 ms that represents correlation of the water bottom and primary. The isolated burst between 170 to 340 ms represents the actual multiple series (both the peg-legs and water-bottom multiples). The prediction lag must be chosen to bypass the first part of the autocorrelogram, which represents the seismic wavelet and possible correlation between the primaries. Operator length must be chosen to include the first isolated burst, in this case between 170 to 340 ms.

It is only with vertical incidence and zero-offset recording that periodicity of the multiples is preserved. Therefore, predictive deconvolution aimed at multiple suppression may not be entirely effective when applied to nonzero-offset data, such as common-shot or common-midpoint data. Predictive deconvolution sometimes is applied to CMP stacked data in an effort to suppress multiples. The performance of such an approach can be unsatisfactory, because the amplitude relationships between multiples often are grossly altered by the stacking process, primarily because of velocity differences between primaries and multiples. Also, geometric spreading compensation by using primary velocity function adversely affects the amplitudes of multiples on nonzero-offset data. There is one domain in which the periodicity and amplitudes of multiples are preserved. This is the slant stack domain. In Section 7.5, the application of predictive deconvolution to slant stack data for multiple suppression is discussed.

2.7.6 Field Data Examples

The deconvolution parameters now are examined using field data examples. Figure 2-66 shows a CMP gather that contains five prominent reflections at around 1.1, 1.35, 1.85, 2.15, and 3.05 s. The gather also contains strong reverberations associated with these reflections. In Figures 2-66 through 2-69 and Figures 2-75 and 2-76, the input CMP gather was 6 s long; only the first 4 s of the results are displayed. The examination of the deconvolution parameters will begin with an analysis of the time gate to estimate the autocorrelation function. A first gate selected may be the entire length (6 s) of the record as seen in panel a. The solid lines on the CMP gathers refer to the gate start and end times. The autocorrelogram of the record is shown at the bottom of each panel. A second choice might be to exclude the deeper part of the record where ambient noise dominates. The start of the gate is chosen as the first arrival path (panel b). A third choice may be to exclude not only the deeper portion, but also the early part of the record that contains energy corresponding to the guided waves (panel c). These waves travel within the water layer and are not part of the signal reflected from the substrata.

By comparing the autocorrelograms from these different windows, note that the third choice best represents the reverberatory character of the data (panel c) over most of the offsets. All of the traces in the autocorrelogram within approximately the first 150 ms have a common appearance. This early portion of the autocorrelogram characterizes the basic seismic wavelet contained in the data. In general, the autocorrelation window should include the part of the record that contains useful reflection signal and should exclude coherent or incoherent noise. An autocorrelation function contaminated by noise is undesirable since the deconvolution process is most effective on noise-free data (assumption 3).

Another aspect of the autocorrelation window is length. Figure 2-66, panel d, shows the autocorrelogram estimated from a narrow window. The autocorrelogram estimated from the narrower part of the time gate (the right side of the record) in some data cases may lack the characteristics of the reverberations and basic seismic wavelet. In general, any autocorrelation function is biased; that is, the first lag value is computed from, say, n nonzero samples, the second lag value is computed from n − 1 nonzero samples, and so on. If n is not large enough, then there can be an undesirable biasing effect. How large should the data window be to avoid such biasing? If the largest autocorrelation lag used in designing the deconvolution operator were m, an accepted rule of thumb is that the number of data samples should be no less than $8m$.

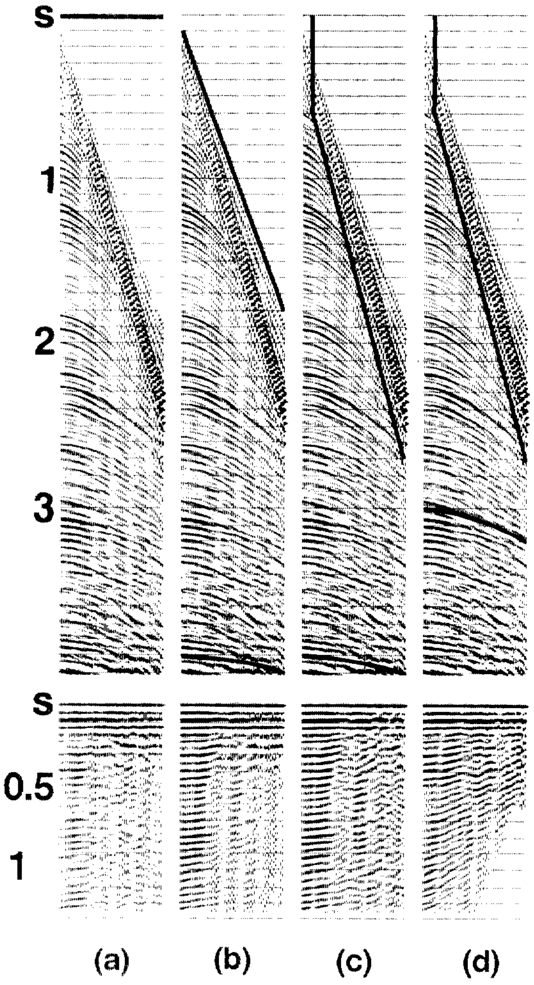


FIG. 2-66. An autocorrelation window test used to design deconvolution operators. The solid bars indicate the window boundaries. The entire 6-s length was included in a. The autocorrelograms are displayed beneath the records.

Now that the autocorrelation window is determined, we examine operator length. In Figure 2-67, prediction lag (4 ms, the same as the sampling rate) and percent prewhitening (0.1 percent) are fixed. The autocorrelograms (at the bottom of each gather) are displayed for diagnostic purposes. From the analyses of the single spike, sparse spike, and reflectivity models, the short (40-ms) operator leaves some residual energy that corresponds to the basic wavelet and reverberating wave train in the record. For a spiking deconvolution with a 160-ms-long operator, no remnant of the energy is associated with the basic wavelet and reverberations. Any operator longer than 160 ms does not change the result significantly. From the output of the 160-ms operator, note that the prominent reflections (at 1.1, 1.35, 1.85, and 2.15 s at the near offset) have been uncovered, the seismic wavelet has been compressed, and the reverberations have been significantly suppressed.

The effect of prediction lag now is examined. In Figure 2-68, the 160-ms operator length and 0.1 percent prewhitening are fixed, while prediction lag is varied. If prediction lag were increased, then the deconvolution process would be increasingly less effective in broadening the spectrum, and the autocorrelograms contain increasingly more energy at nonzero lags. In the extreme, the deconvolution process is ineffective for a 128-ms prediction lag. In practice, common values for the prediction lag are unity (spiking deconvolution) or the first or second zero crossing of the autocorrelation function (predictive deconvolution).

Finally, the percent prewhitening is varied, while the 4-

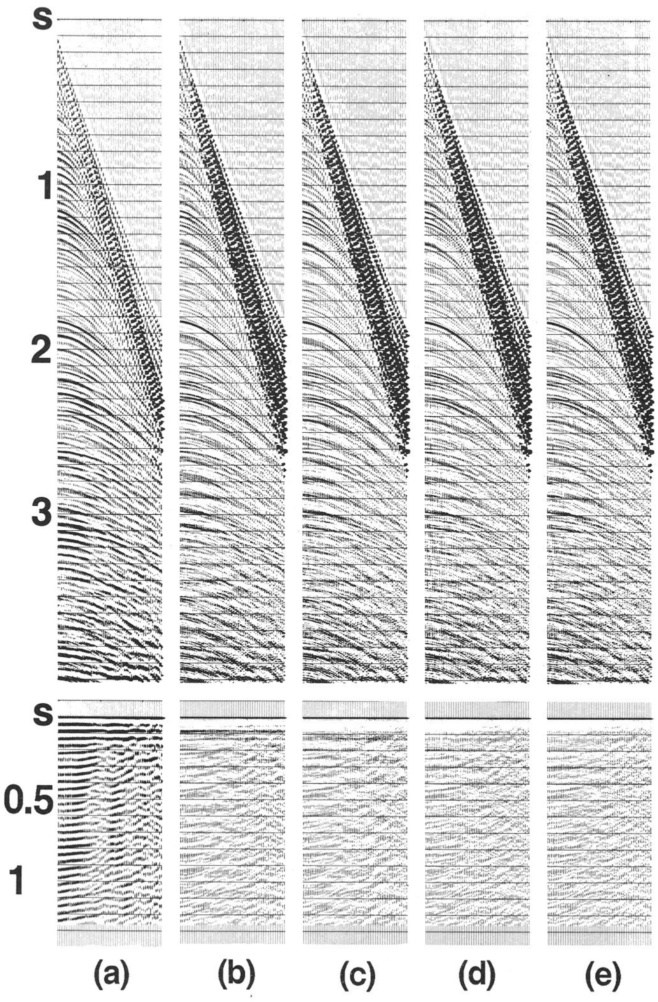


FIG. 2-67. Test of operator length. The corresponding autocorrelogram is beneath each record. The window used in autocorrelation estimation is shown in Figure 2-66c. (a) Input gather. Deconvolution using prediction lag = 4 ms (spiking deconvolution), 0.1 percent prewhitening, and prediction filter operator lengths (b) 40 ms, (c) 80 ms, (d) 160 ms, (e) 240 ms.

ms prediction lag and 160-ms operator length are fixed. These tests are shown in Figure 2-69. By increasing the percent prewhitening, the deconvolution process becomes less effective. The high end of the spectrum is not flattened as much as the rest of the spectrum (Figure 2-54). Note that the autocorrelograms contain increasingly more energy at nonzero lags with increasing percent prewhitening. In practice, it is not advisable to assign a large percent of prewhitening. Typically, a value between 0.1 and 1 percent is sufficient to ensure stability in designing the deconvolution operator.

Nonstationarity was discussed in Sections 1.5 and 2.2. The time-variant character of the seismic wavelet (Figure 2-9) often requires a multiwindow deconvolution. Figure 2-70 is a field record that was deconvolved by using three time gates. The autocorrelograms from gates 1, 2, and 3 are shown in Figure 2-71. Note the difference in character of the reverberatory energy from one gate to another. The shallow gate (1) has more high-frequency signal than the middle gate (2); while the middle gate has more high-frequency signal than the deeper gate (3). For best results, we must design different deconvolution operators from different parts of the record and apply them to the corresponding time gates. Three windows usually are sufficient to handle the nonstationary character of the seismic signal. Another example of single-and multiwindow deconvolution is shown in Figures 2-72 and 2-73. Here, autocorrelograms from different gates do not show significant variations. Therefore, it probably does not make any difference whether a single or multigate deconvolution is used. (Compare the outputs from both tests.) In these figures, the record is shown after deconvolution followed by a wide band-pass filter application. Since the amplitude spectrum of the input data is flattened as a result of spiking deconvolution, both the high-frequency ambient noise as much as the high-frequency components of the signal are boosted. Therefore, the output of spiking deconvolution often is filtered with a wide band-pass operator.

The data sometimes must be preconditioned for deconvolution. If the data were too noisy, then a wide band-pass filter could be necessary before deconvolution. If there is significant coherent noise in the data, dip filtering (as discussed in Sections 1.6.2 and 7.4) can be applied before deconvolution so that coherent noise is not included in the autocorrelation estimate.

Poststack deconvolution often is considered for several reasons. First, a residual wavelet almost always is present on the stacked section. This is because none of the underlying assumptions for deconvolution is completely met in real data; therefore, deconvolution never can completely compress the basic wavelet contained in prestack data to a spike. Second, since a CMP stack is an approximation to the zero-offset section, predictive deconvolution aimed at removing multiples may be a viable process after stack. Figure 2-74 is an example of poststack deconvolution. After deconvolution, the wavelet is compressed further and the marker horizons are better characterized.

Finally, Figure 2-75 shows an example of signature deconvolution. In marine seismic exploration, the far-field signature of the source array can be recorded. The idea is to apply a deterministic deconvolution to remove the source signature, then to apply predictive deconvolution. The convolutional model is given by

$$x(t) = s(t) * w(t) * e(t), \tag{2.31}$$

where $s(t)$ is the source signature recorded in the far-field just before it travels down into the earth, which has an impulse response $e(t)$. Since $s(t)$ is recorded, an inverse filter can be deterministically designed, as discussed in Section 2.4, then applied to the recorded seismogram to remove $s(t)$ from equation (2.31). The unknown wavelet $w(t)$ includes the propagating effects in the earth and the response of the recording system. This remaining wavelet then is removed by the statistical method of spiking

(text continued on page 139)

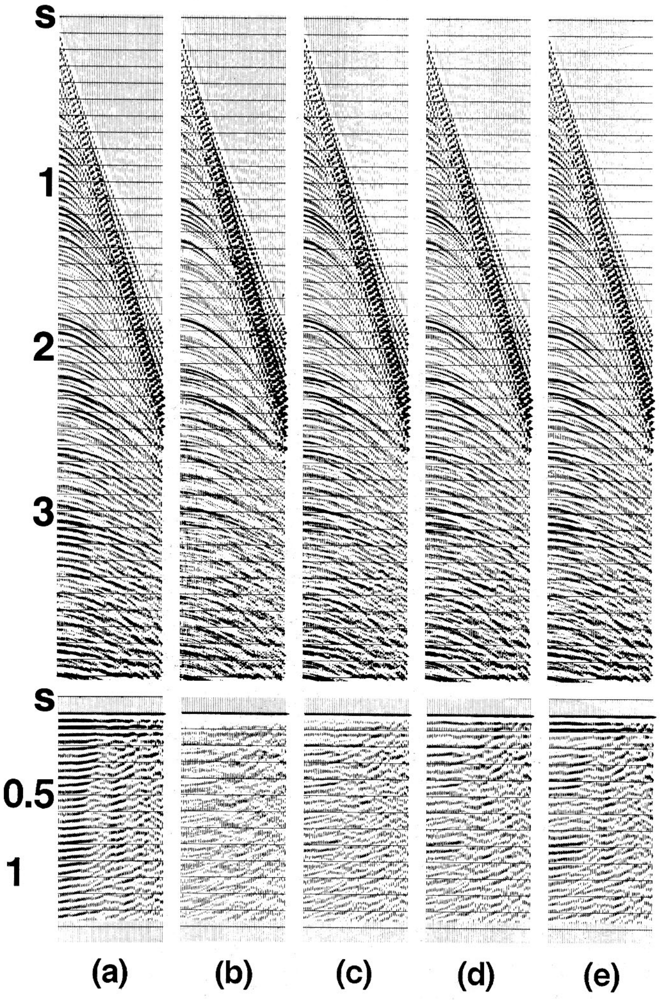


FIG. 2-68. Test of prediction lag. The corresponding autocorrelogram is beneath each record. The window used in autocorrelation estimation is shown in Figure 2-66c. (a) Input gather. Deconvolution using prediction filter operator length = 160 ms, 0.1 percent prewhitening, and prediction lags (b) 12 ms, (c) 32 ms, (d) 64 ms, (e) 128 ms.

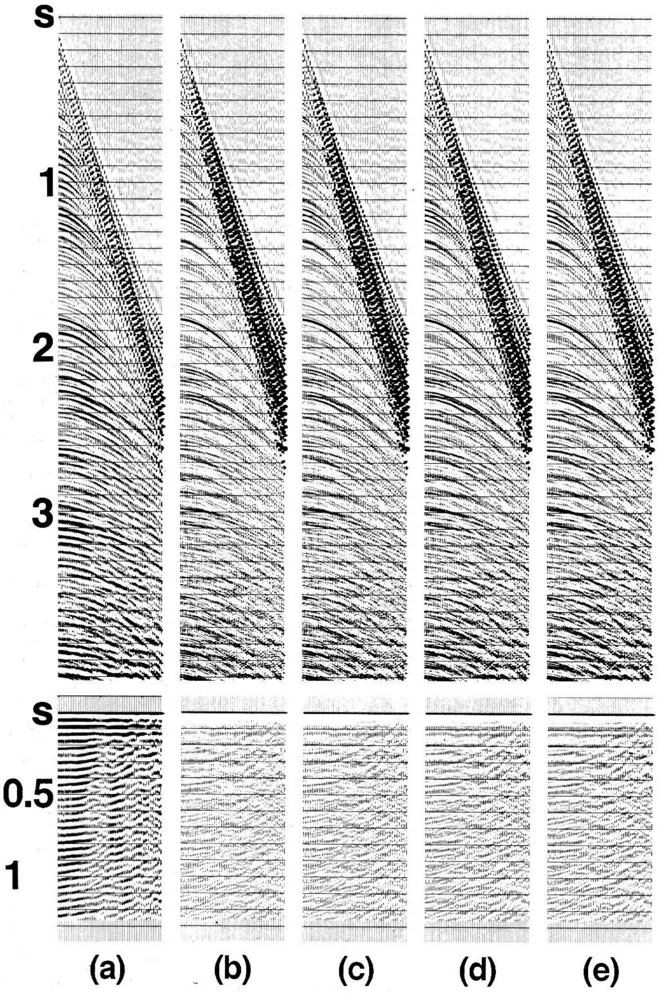


FIG. 2-69. Test of percent prewhitening. The corresponding autocorrelogram is beneath each record. The window used in autocorrelation estimation is shown in Figure 2-66c. (a) Input gather. Deconvolution using prediction filter operator length = 160 ms, prediction lag = 4 ms (spiking deconvolution), and percent prewhitening (b) 1, (c) 4, (d) 16, (e) 32.

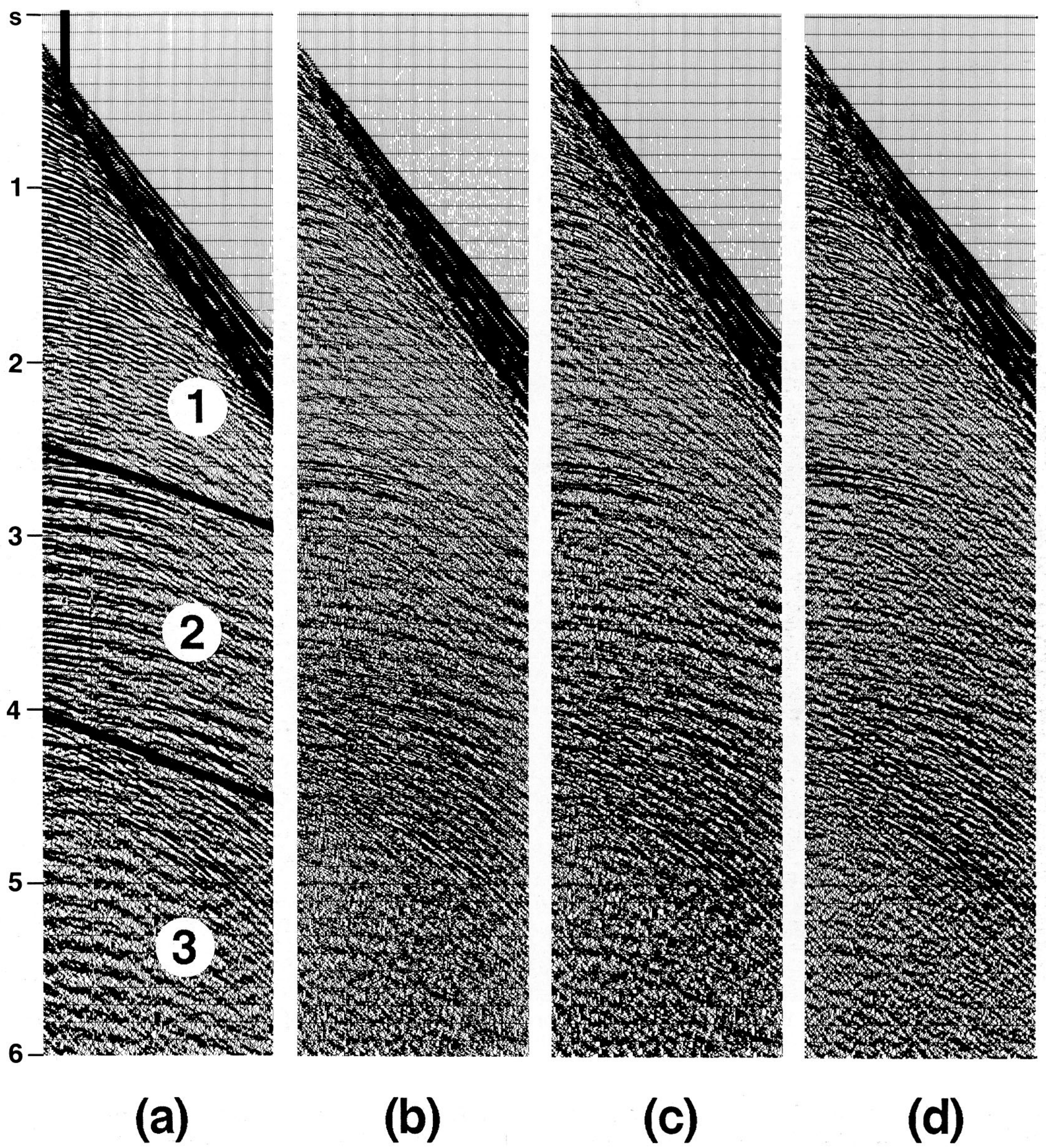


FIG. 2-70. Three-window deconvolution. The solid bars indicate window boundaries. With data from each window, a deconvolution operator is designed and applied to the data in that window. The operators are blended across the window boundaries. (a) Input gather. Deconvolution using operator length = 160 ms and prediction lags (b) 4 ms, (c) 12 ms, (d) 32 ms.

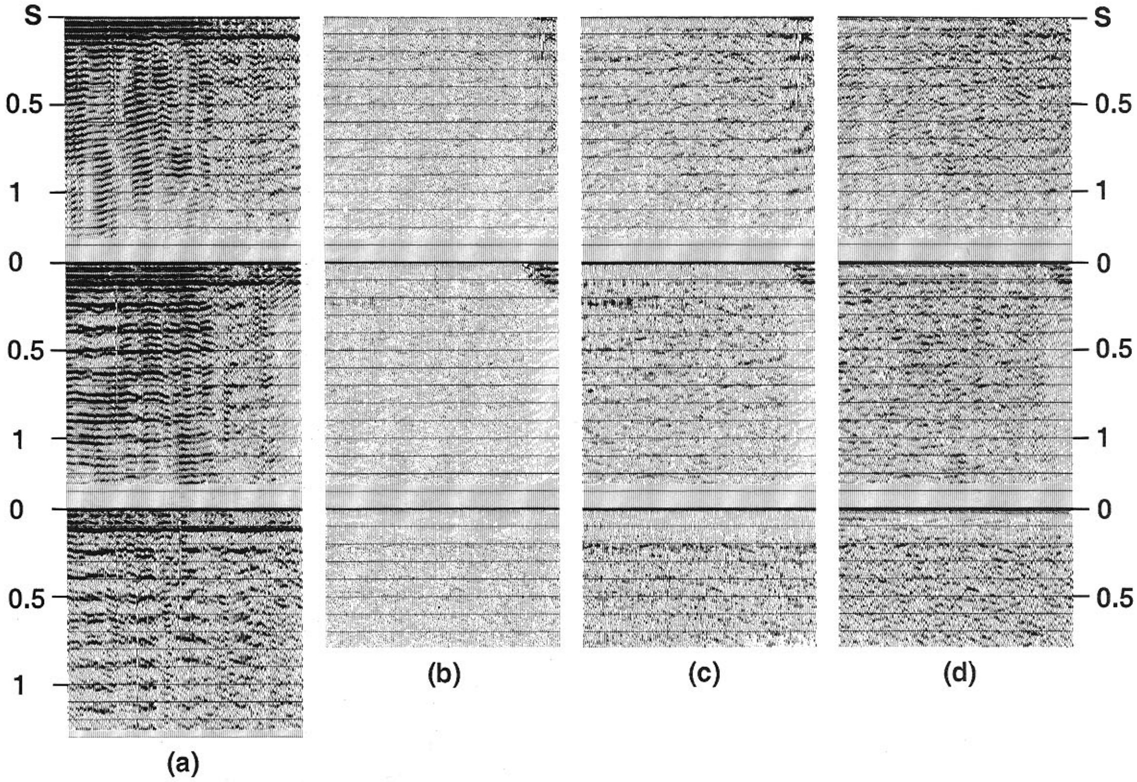


FIG. 2-71. Autocorrelograms associated with the records shown in Figure 2-70. Note the difference in character between windows 1, 2, and 3 (from top to bottom in the leftmost panel). Windows are indicated in Figure 2-70a.

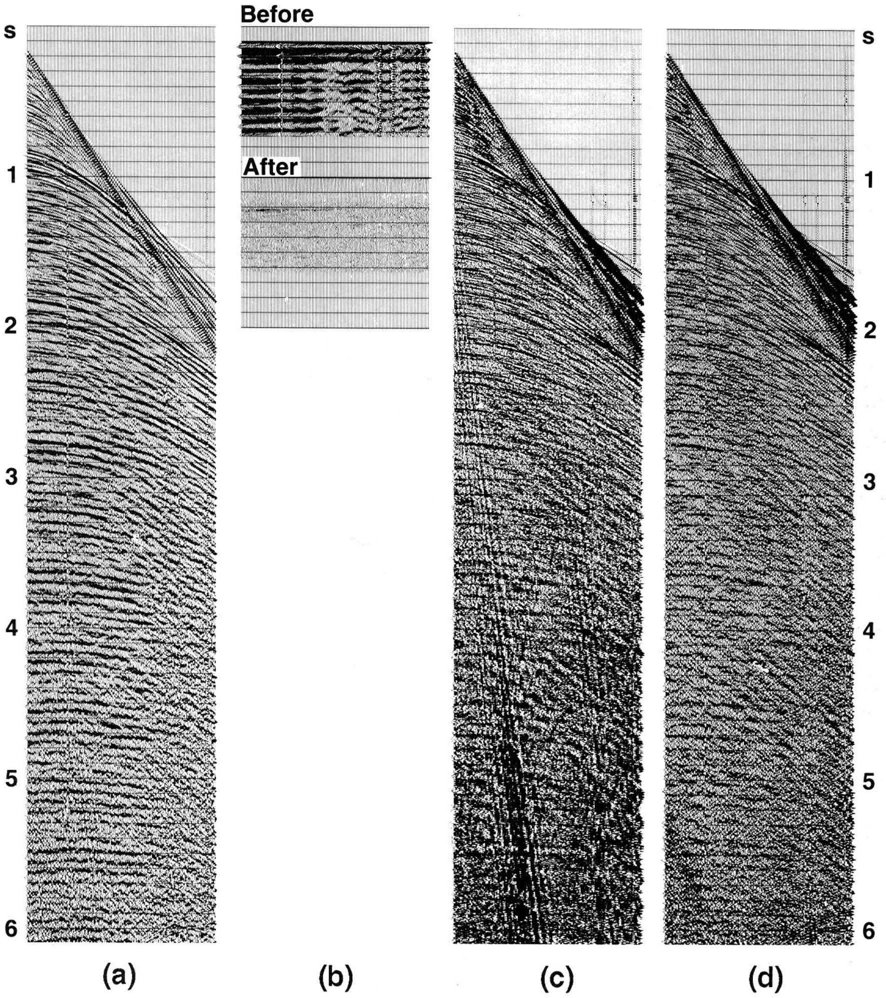


FIG. 2-72. Spiking deconvolution (c) on a shot record (a) followed by band-pass filtering (d). (b) Autocorrelograms before and after deconvolution.

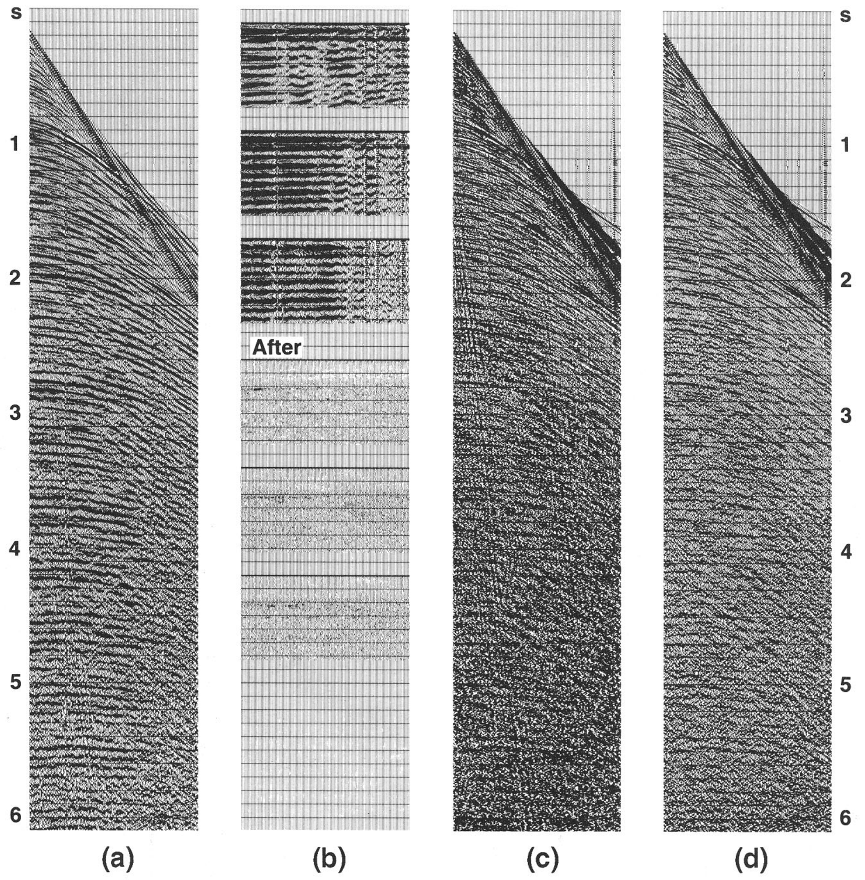


FIG. 2-73. Three-window deconvolution on the same shot record as in Figure 2-72. In this case, there is no significant difference between the characters of the autocorrelograms estimated from three windows. (a) Input gather, (b) auto correlograms before and after spiking deconvolution, (c) three-window spiking deconvolution on (a), (d) band-pass filtering on (c).

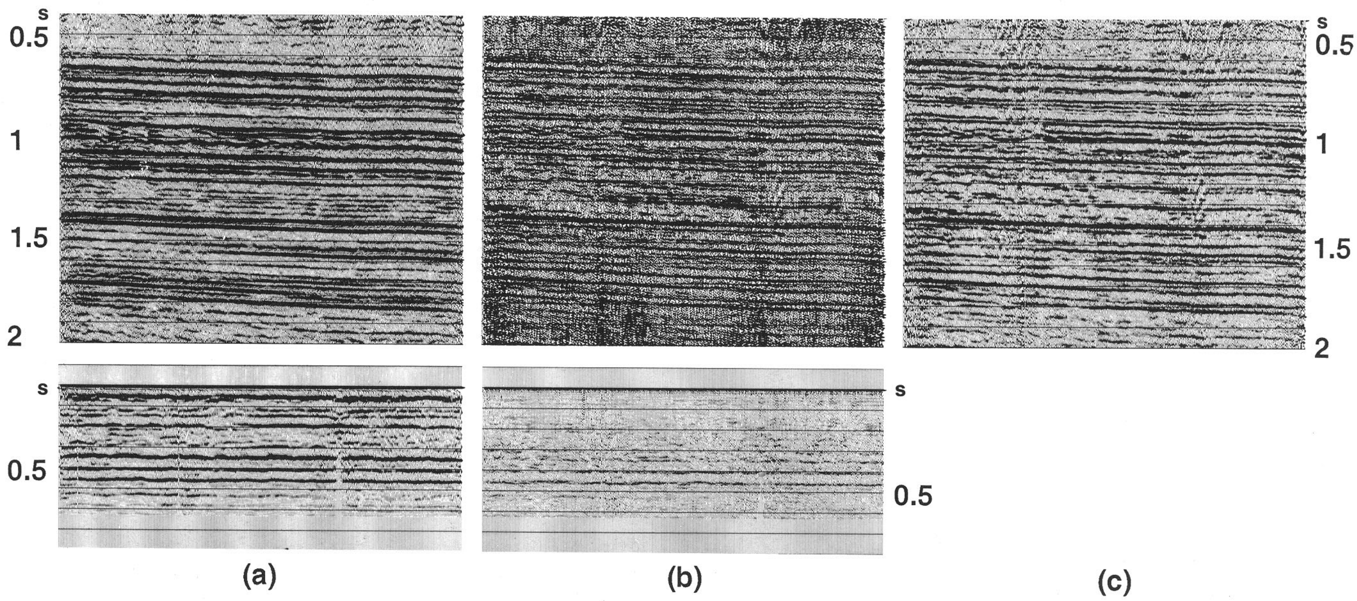


FIG. 2-74. Deconvolution after stack. The bottom frames are the autocorrelograms of the sections above. (a) CMP stack with no deconvolution, (b) spiking deconvolution after stack, followed by band-pass filtering (c).

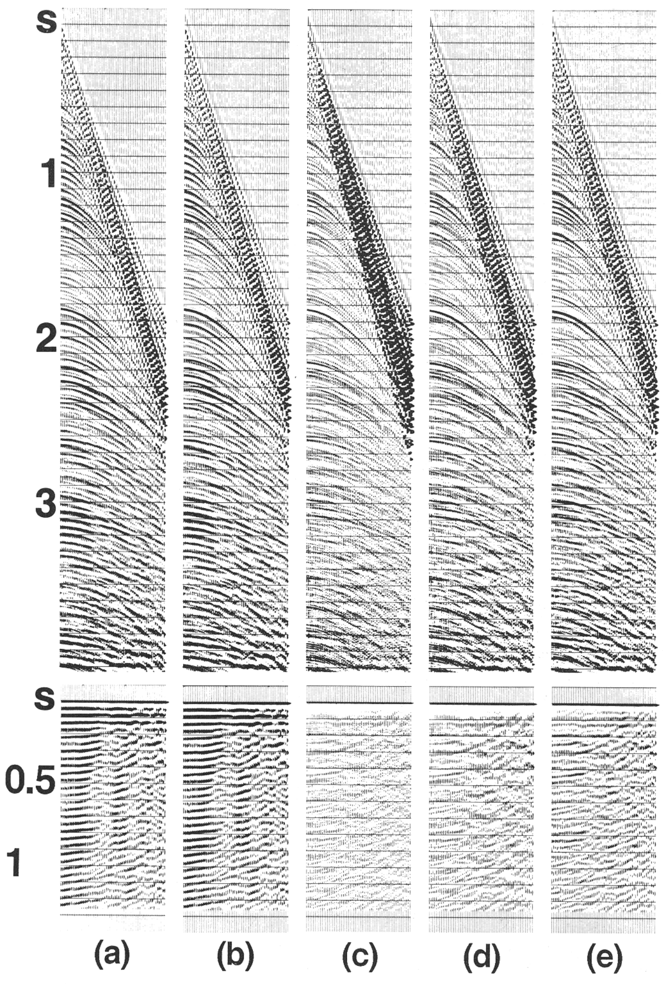


FIG. 2-75. Signature processing. A shaping filter is designed to convert the recording signature [*s(t)* in equation (2.31)] to its minimum-phase equivalent and apply it to the input record (a). Note that the output (b) has the same bandwidth as the input (a). The output (b) then has been processed by predictive deconvolution using operator length = 160 ms. and prediction lags (c) 4 ms (spiking deconvolution). (d) 12 ms. (e) 32 ms.

deconvolution as discussed in Section 2.6. If equation (2.31) were compared to equation (2.2), we would see that the old *w(t)* of equation (2.2) would be split into two parts; i.e., the source signature *s(t)*, which is the known component, and the new *w(t)*, which is the unknown component.

There are two ways to handle *s(t)*. One way is to convert it to its minimum-phase equivalent followed by predictive deconvolution (Figure 2-75). Another way is to convert *s(t)* into a spike followed by predictive deconvolution (Figure 2-76). The results shown in Figures 2-75 and 2-76 (center panels) should be compared with Figure 2-67d. Since the source was not minimum-phase in this case, Figure 2-75c should be better than Figure 2-67d. Is it?

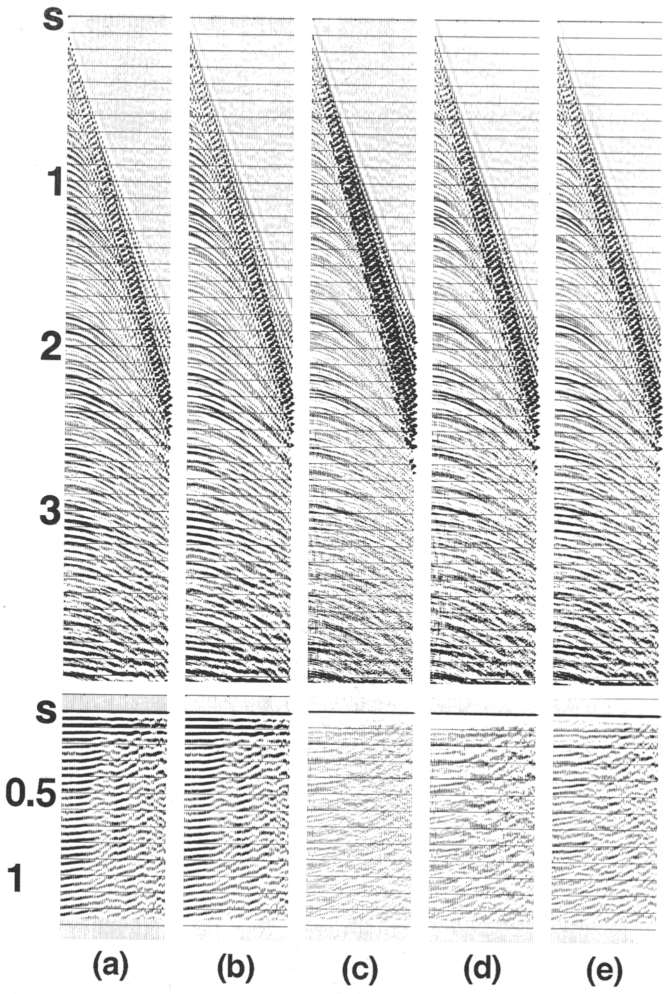


FIG. 2-76. Signature processing. A shaping filter is designed to convert the recorded signature [*s(t)* in equation (2.31)] to a spike and apply it to the input record (a). The output (b) then has been processed by predictive deconvolution using operator length = 160 ms and prediction lags (c) 4 ms (spiking deconvolution), (d) 12 ms. (e) 32 ms. Note that the output from signature processing (b) does not appear to have white spectrum; the *w(t)* component [equation (2.31)] still needs to be removed.

Results of signature processing depend on the accuracy of the recorded signature. Figure 2-77 shows a CMP stack with and without signature processing. Note the difference between the CMP stack with no deconvolution (Figure 2-77a) and with spiking deconvolution before stack (Figure 2-77b). The same data were signature processed first by converting the measured signature to its minimum-phase equivalent followed by spiking deconvolution (Figure 2-77c). When compared with Figure 2-77b, it can be argued that the reflection continuity in some places has improved. Note that just converting the signature to a spike is not sufficient (Figure 2-77d); the *w(t)* component still remains to be removed [equation (2.31)].

2.7.7 Vibroseis Deconvolution

The vibroseis source is a long-duration sweep signal in the form of a frequency-modulated sinusoid that is tapered on both ends. Just as a convolutional model was proposed for the marine seismogram given by equation (2.31), a similar convolutional model can be proposed for the vibroseis seismogram:

$$x(t) = s(t) * w(t) * e(t), \tag{2.32}$$

where *x(t)* is the recorded seismogram, *s(t)* is the sweep signal, *w(t)* is the seismic wavelet [with the same meaning as in equation (2.31)], and *e(t)* is the earth's impulse response. Convolutions in equation (2.32) become multiplications in the frequency domain:

$$X(\omega) = S(\omega)W(\omega)E(\omega). \tag{2.33}$$

In terms of amplitude A (ω) and phase ϕ (ω) spectra,

$$A_x(\omega) = A_s(\omega)A_w(\omega)A_e(\omega) \tag{2.34a}$$

and

$$\phi_x(\omega) = \phi_s(\omega) + \phi_w(\omega) + \phi_e(\omega). \tag{2.34b}$$

Crosscorrelation of the recorded seismogram *x(t)* with the sweep signal *s(t)* is equivalent to multiplying equation (2.34a) by $A_s(\omega)$ and subtracting $\phi_s(\omega)$ from equation (2.34b). The correlated vibroseis seismogram *x′(t)* there-

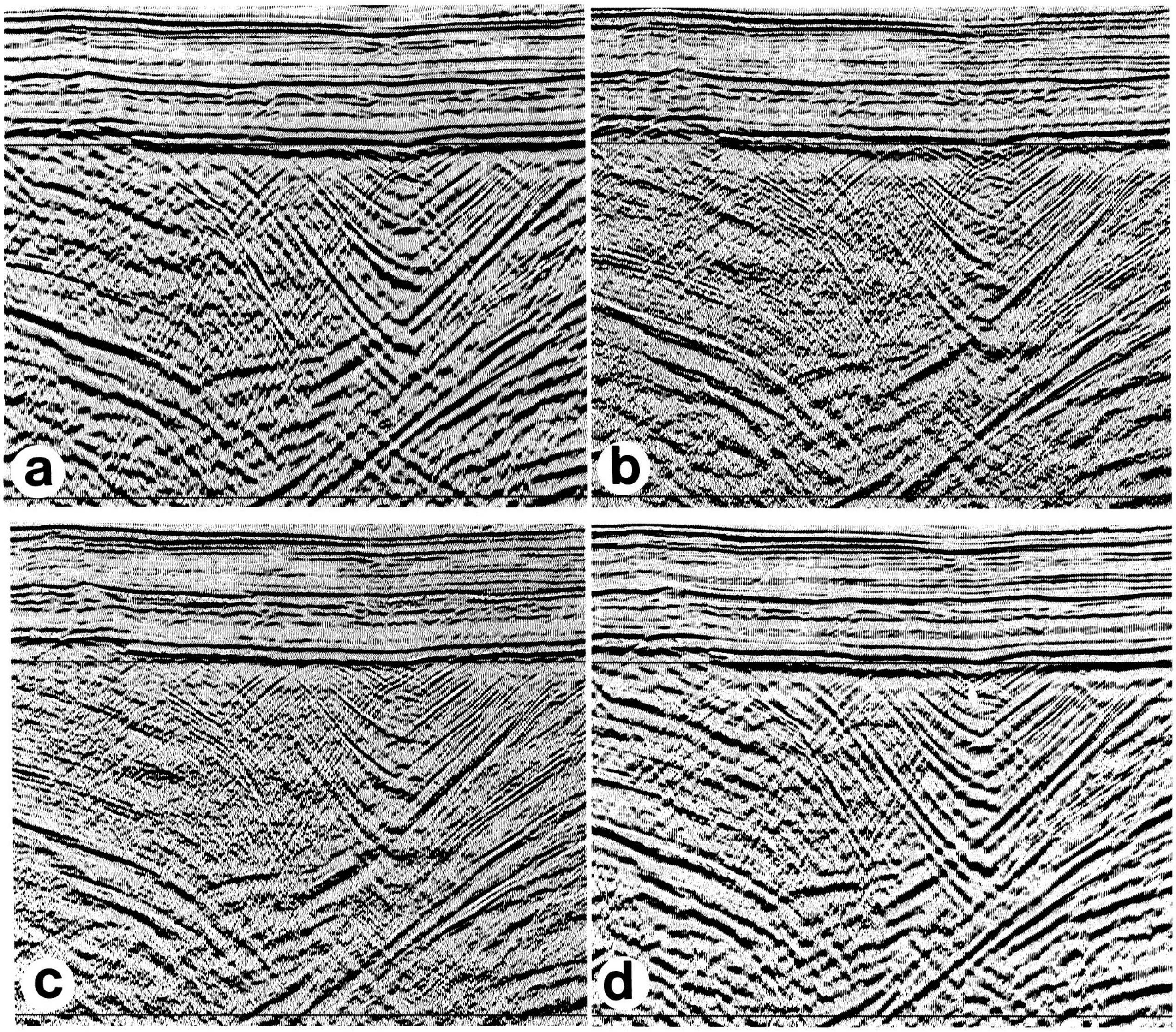


FIG 2-77. A CMP stack (a) with no deconvolution, (b) with spiking deconvolution before stack, (c) with signature processing (minimum-phase conversion of the measured signature) followed by spiking deconvolution, (d) with signature processing only (conversion of the signature to a spike).

fore would have the following amplitude and phase spectra:

$$A'(\omega) = A_s^2(\omega) A_w(\omega) A_e(\omega) \qquad (2.35a)$$

and

$$\phi'(\omega) = \phi_w(\omega) + \phi_e(\omega). \qquad (2.35b)$$

The inverse Fourier transform of $A_s^2(\omega)$ yields the autocorrelation of the sweep signal, which is called the Klauder wavelet *k(t)*. Returning to the time domain, equations (2.35) yield

$$x'(t) = k(t) * w(t) * e(t). \qquad (2.36)$$

Since it is an autocorrelation, the Klauder wavelet is zero-phase. Convolution of *k(t)* with the assumingly minimum-phase wavelet *w(t)* yields a mixed-phase wavelet. Because spiking deconvolution is based on the minimum-phase assumption, it cannot properly recover *e(t)* from vibroseis data.

One approach to deconvolution of vibroseis data is to apply a zero-phase inverse filter to remove *k(t)*, followed by a minimum-phase deconvolution to remove *w(t)*. The amplitude spectrum of the inverse filter is defined as $1/A_s^2(\omega)$. In practice, problems arise because of zeroes in the spectrum that are caused by the band-limited nature of the Klauder wavelet. Inversion of an amplitude spectrum, which has zeroes, yields an unstable operator. To circumvent this problem, a small percent of white noise, say 0.1 percent, usually is added before inverting the Klauder wavelet spectrum.

Another approach is to design a filter that converts the Klauder wavelet to its minimum-phase equivalent (Ristow and Jurczyk, 1975). A way to get the minimum-phase spectrum from a given amplitude spectrum is described in Appendix B.4 and is included in the discussion on frequency-domain deconvolution. If the Klauder wavelet were converted to its minimum-phase equivalent, then equation (2.35b) would take the form:

$$\phi'(\omega) = \phi_k(\omega) + \phi_w(\omega) + \phi_e(\omega). \qquad (2.37)$$

If we assume that *w(t)* is minimum-phase and if we make *k(t)* minimum-phase, then the result of their convolution also is minimum-phase. Spiking deconvolution now is applicable since the minimum-phase assumption is satisfied. There is a 90-degree phase difference in some vibrator systems between the control sweep signal and the baseplate response. As an option, we may want to subtract out this phase difference. Figure 2-78 shows the recommended sequence of operations for vibroseis processing.

Figure 2-79 shows how the flowchart in Figure 2-78 is used with a synthetic reflectivity series. By including the step to convert the Klauder wavelet into its minimum-phase equivalent before spiking deconvolution, a closer representation of the impulse response is produced (compare steps k and l with m).

Although sound in theory, the above scheme may have problems in practice. Fundamental issues, such as whether the convolutional model given in equation (2.32) really represents what goes on in the earth, are not resolved. Vibroseis data often are deconvolved as dynamite data, without converting the Klauder wavelet to its minimum-phase equivalent. An example of deconvolution of a correlated vibroseis record is shown in Figure 2-80. In this example, spiking deconvolution by itself is not effective in improving temporal resolution.

Despite the fact that the basic minimum-phase assumption is violated for vibroseis data as compared to explosive data, spiking deconvolution without conversion of the Klauder wavelet to its minimum-phase equivalent seems to work for most field data. Figures 2-81 and 2-82 are examples of vibroseis data before and after spiking deconvolution, respectively. Prominent reflections after deconvolution were enhanced and reverberations were significantly suppressed. Nonetheless, the problem of tying vibrator lines to lines recorded with other sources, say dynamite, is more difficult if the vibrator data have not been phase-corrected. Field systems now exist to do minimum-phase vibroseis correlation in the field.

2.8 THE PROBLEM OF NONSTATIONARITY

Figure 2-83 shows a CMP gather and its filtered versions before deconvolution. The filter scans show that there is signal between 10 to 40 Hz. Note that higher frequencies are confined to the shallower part of the gather. The same field record after spiking deconvolution is shown in Figure 2-84. Filter scans of the deconvolved record also are shown in this figure. A comparison of the records before and after deconvolution demonstrates the effects of the process; in particular, compression of the wavelet and broadening of the spectrum. The input signal level above 40 Hz is relatively weaker than that below 40 Hz. However, deconvolution has attempted to reduce the differences between the signal levels within different frequency bands by flattening the spectrum. The flattening was more effective in the shallow part of the record than in the deeper part. We know that the effective source wavelet is not stationary. Attenuation of the higher frequencies increases as the wavelet travels into the subsurface. Although a multiwindow deconvolution process was used in this case, spectral flattening was not achieved over the entire length of the record because of the large degree of nonstationarity of the data.

Nonstationarity is primarily the result of the effects of wavefront divergence and frequency attenuation. The usual approach to reduce nonstationarity is to apply processes designed to compensate for the above effects before deconvolution. Wavefront divergence is corrected by applying a geometric spreading function (Section 1.5). As yet, a method to compensate for attenuation has not been discussed. Attenuation is measured by a quantity called the quality factor *Q*. An infinite *Q* means that there is no attenuation. This factor can change in depth and in the lateral direction. If we had an analytic form for an

(Text continued on page 147)

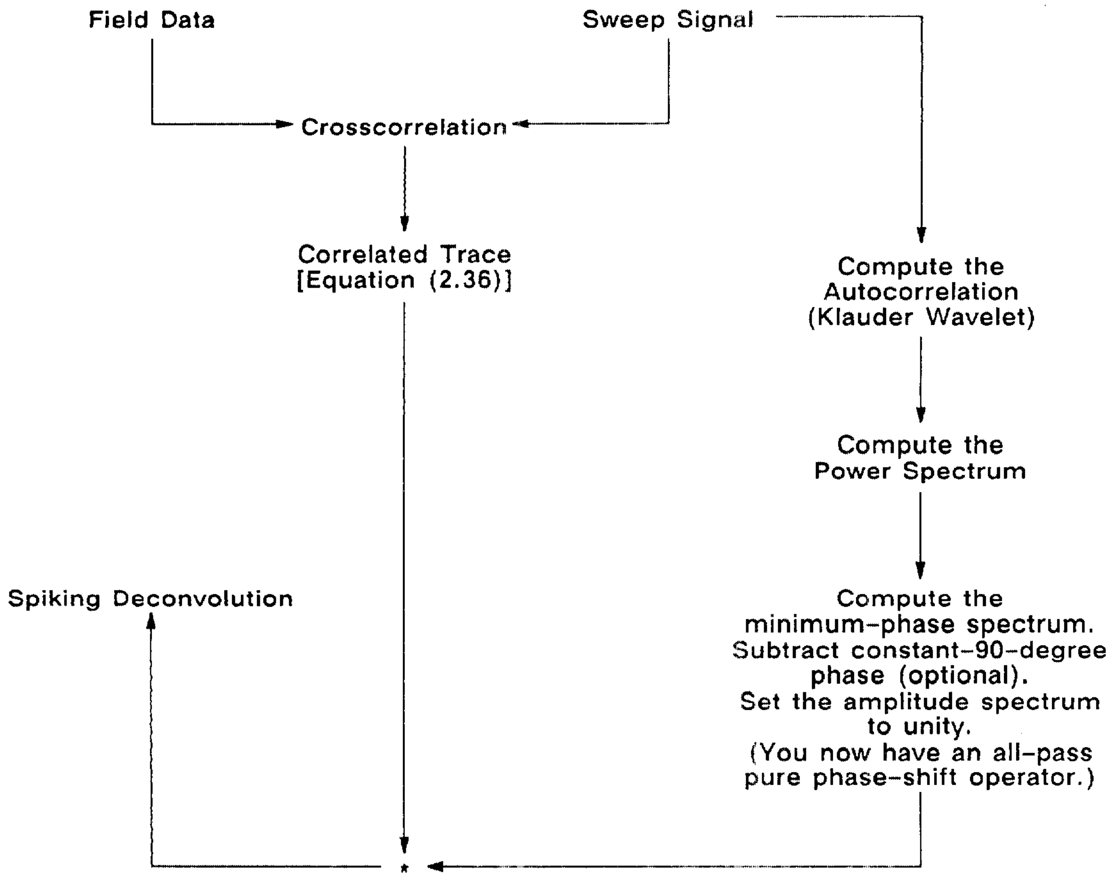


FIG. 2-78. A flowchart for vibroseis deconvolution.

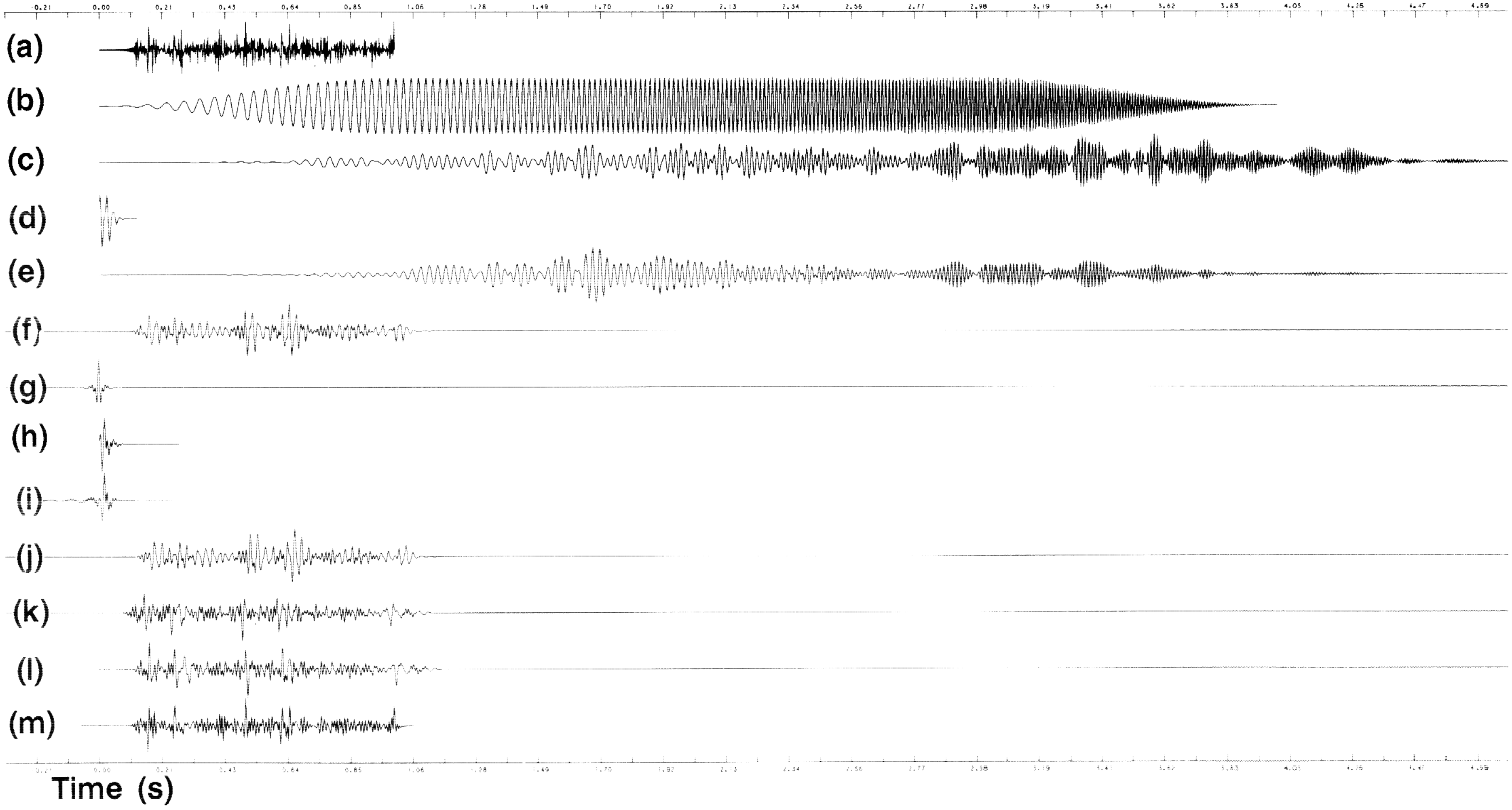


FIG. 2-79. An illustration of vibroseis deconvolution as defined in Figure 2-78. (a) Impulse response spike series; (b) vibroseis sweep signal (10 to 120-Hz bandwidth); (c) convolution of (a) with (b); (d) minimum-phase source wavelet; (e) convolution of (d) with (c); (f) crosscorrelation of the sweep signal (b) with (e); (g) Klauder wavelet; (h) minimum-phase equivalent of (g); (i) the operator that converts (g) to (h); (j) the operator (i) applied to (f); (k) spiking deconvolution of (f); (l) spiking deconvolution of (j); (m) 10 to 120-Hz zero- phase band-pass filtered version of (a).

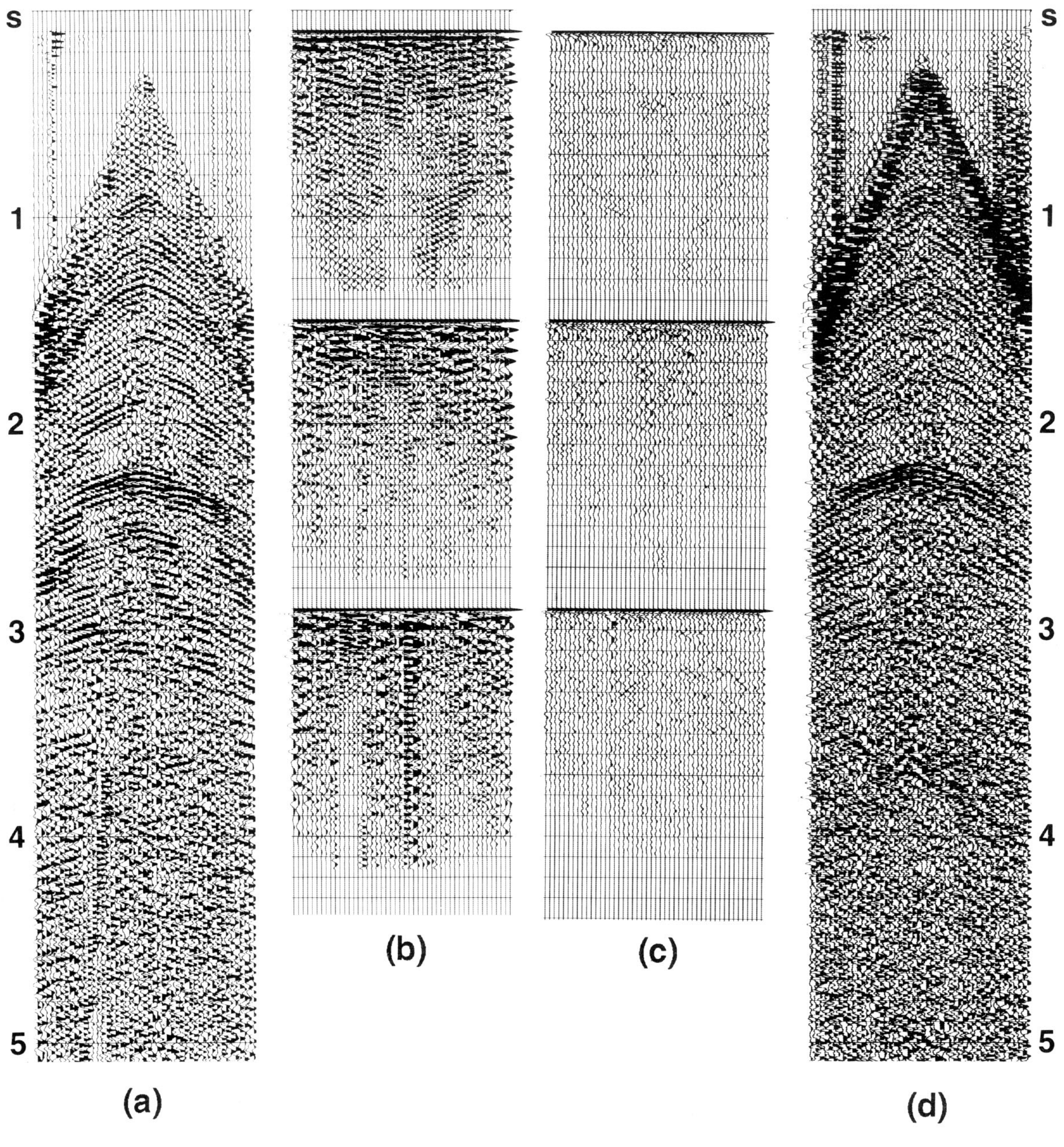


FIG. 2-80. Deconvolution of a vibroseis record. Three windows were used. (a) Correlated vibroseis record and its autocorrelograms (b), (d) spiking deconvolution output and its autocorrelograms (c).

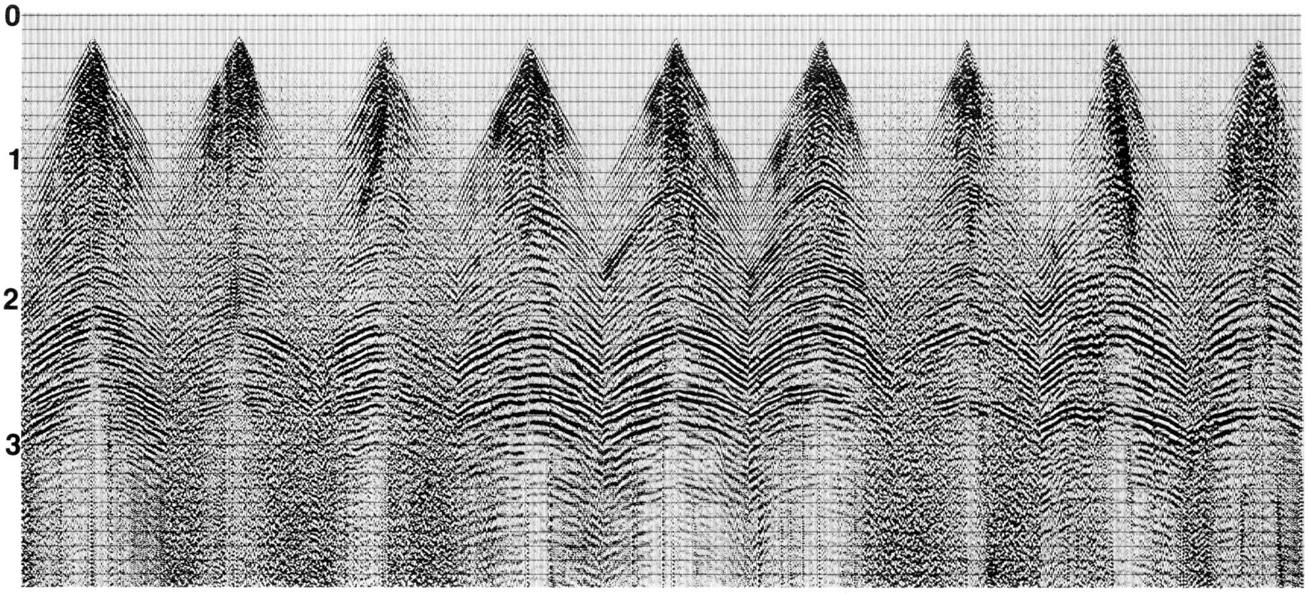


FIG. 2-81. Common-shot gathers with no deconvolution (vibroseis source). Geometric spreading correction and trace balancing have been applied.

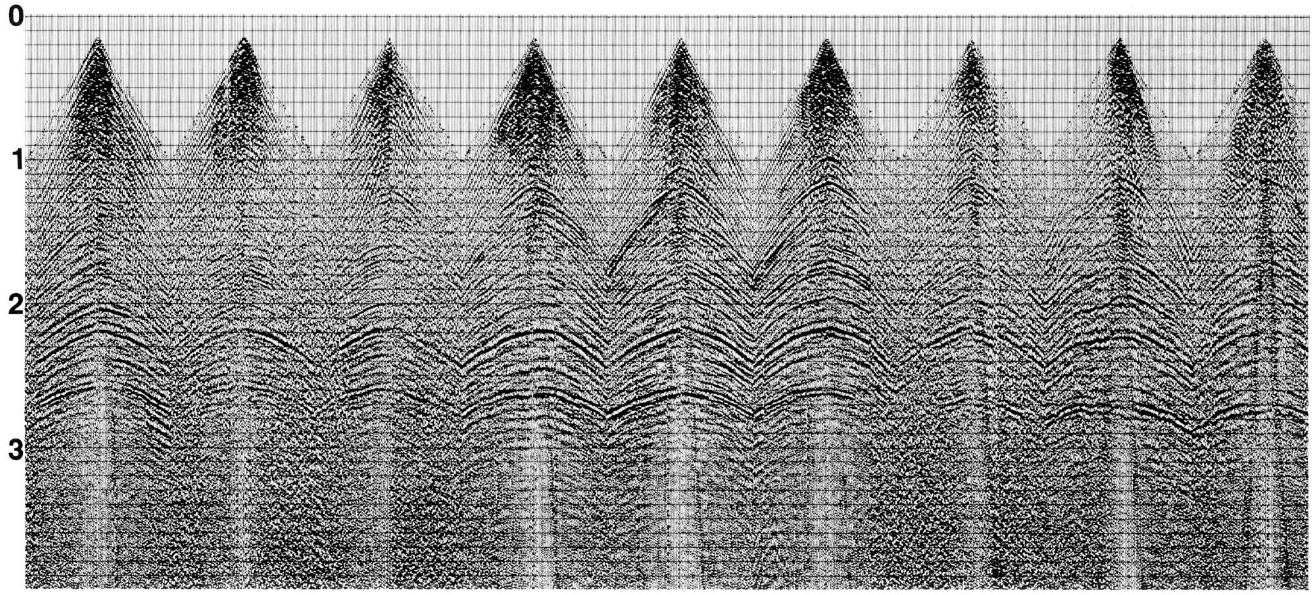


FIG. 2-82. Spiking deconvolution applied to vibroseis data from Figure 2-81.

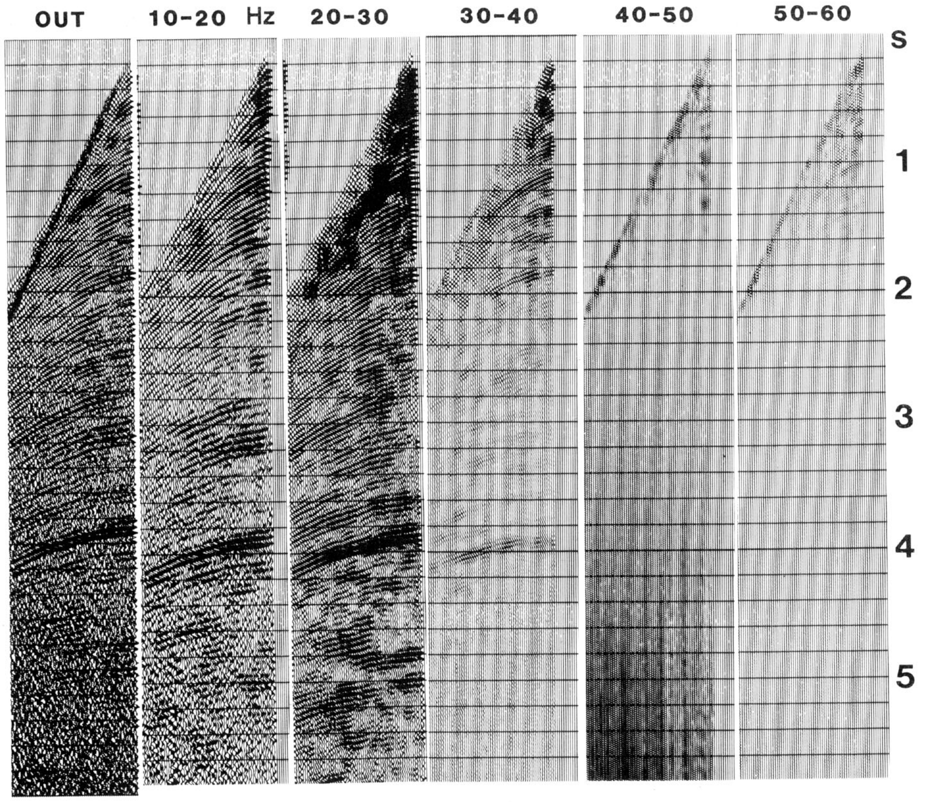


FIG. 2-83. A field record (leftmost panel) with its band-pass filtered versions.

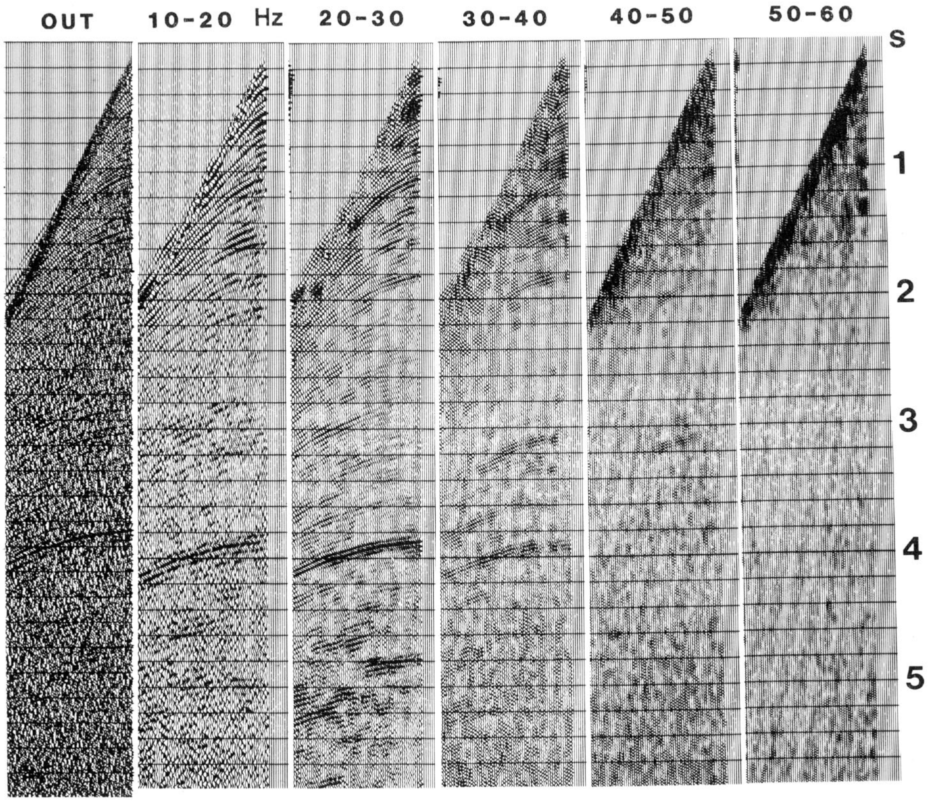


FIG. 2-84. Spiking deconvolution applied to the field record in Figure 2-83 (leftmost panel), followed by application of a series of band-pass filters.

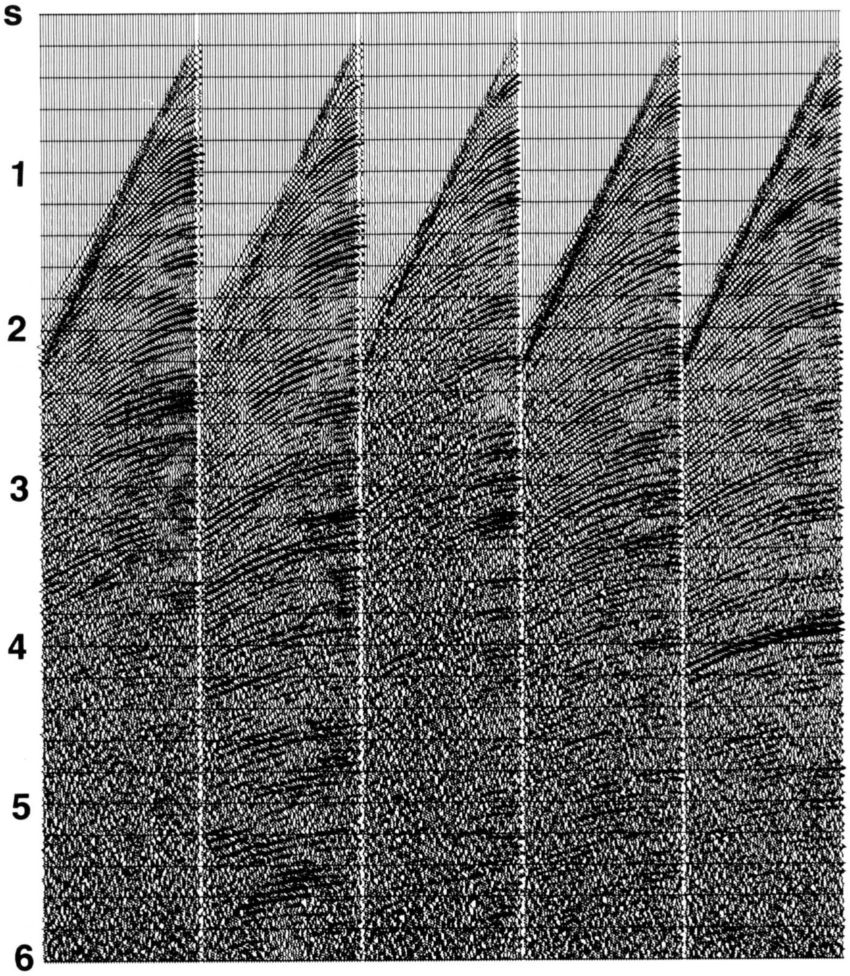


FIG. 2-87. CMP gathers with no deconvolution.

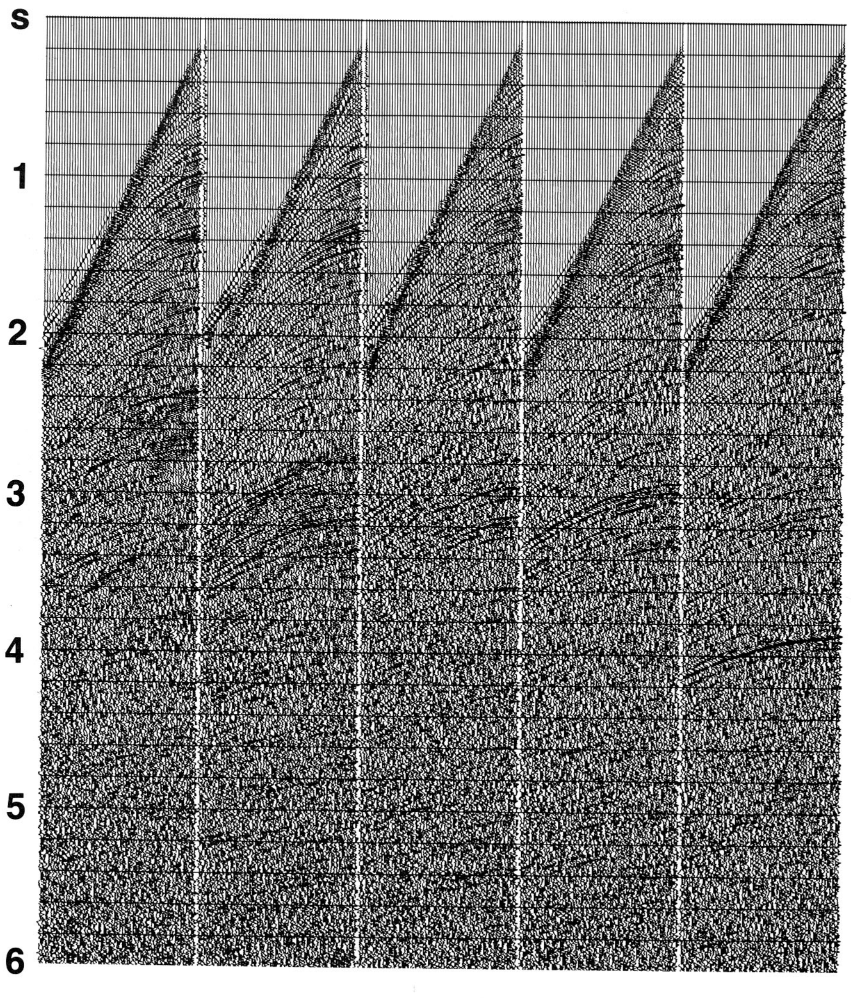


FIG. 2-88. Spiking deconvolution applied to the CMP gathers in Figure 2-87.

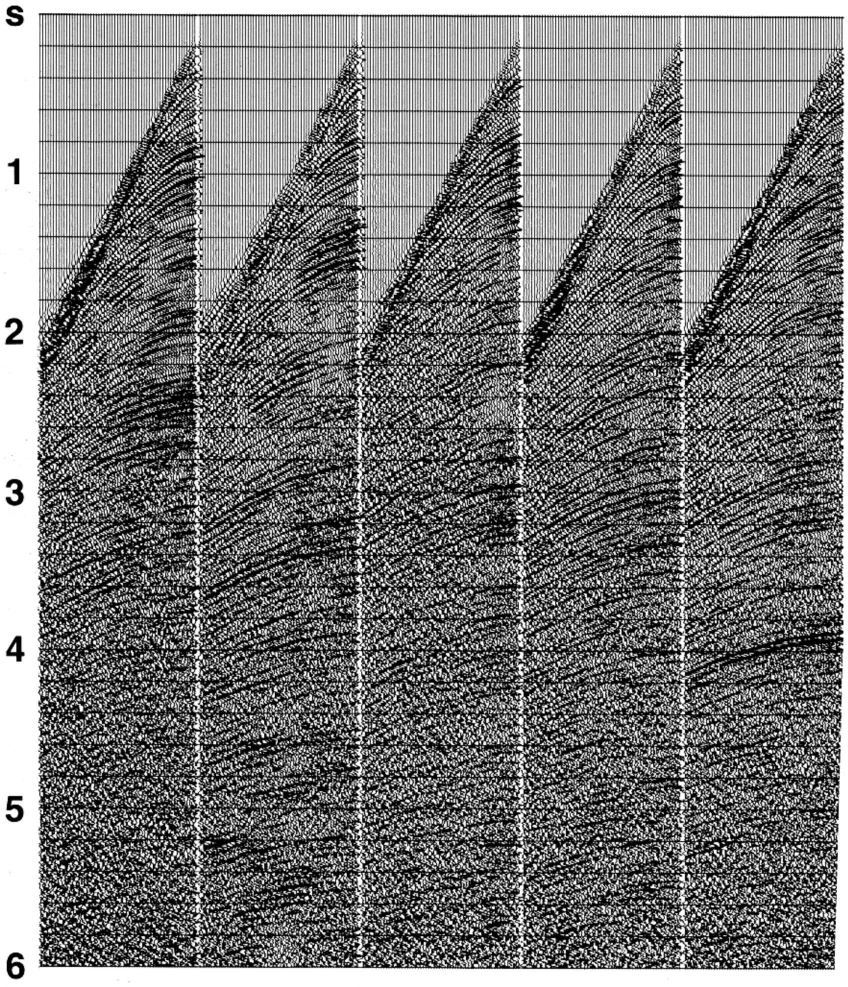


FIG. 2-89. Time-variant spectral whitening (TVSW) applied to the CMP gathers in Figure 2-87. Compare this with Figures 2-88, 2-91, and 2-92.

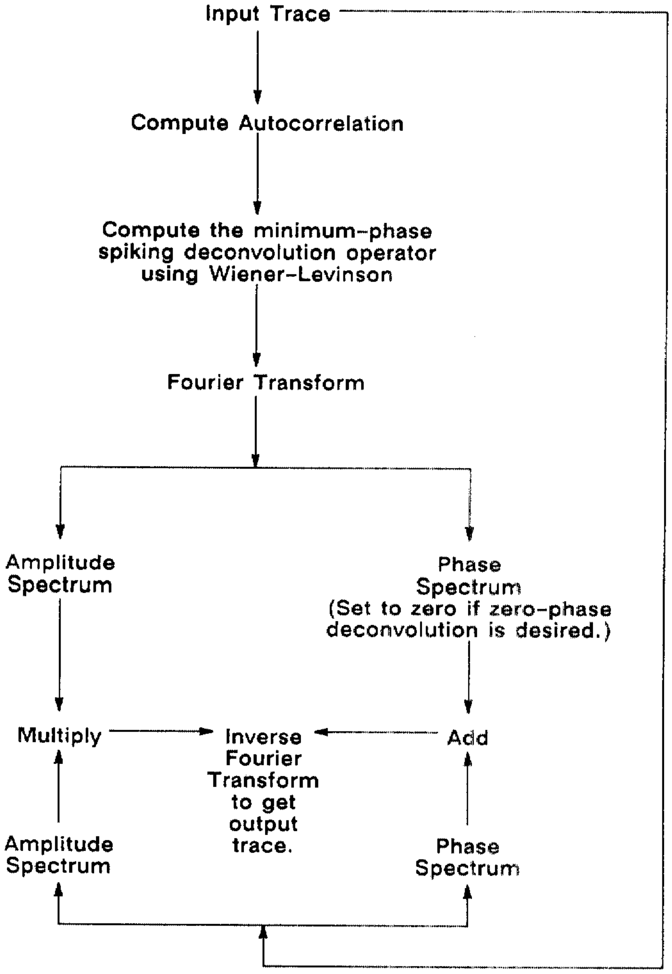


FIG. 2-90. A flowchart for a frequency-domain deconvolution.

Although the domains of operations may differ, both minimum-phase frequency-domain and Wiener-Levinson deconvolution techniques should yield equivalent results. Differences between the results shown in Figures 2-88 and 2-91 mainly are due to their computational aspects.

The zero-phase frequency-domain deconvolution aimed at achieving TVSW requires partitioning the input seismogram into small time gates, as well as designing and applying the process described in Figure 2-90 to each gate, individually. Figure 2-92 shows the field records after zero-phase frequency-domain deconvolution. The output is comparable to the TVSW output shown in Figure 2-89.

Throughout the development of deconvolution, several alternatives have been proposed to better solve the deconvolution problem. Still, predictive deconvolution is used more than the other methods, although the minimum-phase and white reflectivity sequence assumptions have been key issues of concern. The following formal processing sequence for deconvolution theoretically should yield optimum results:

1. Apply the geometric spreading compensation function v^2t. This removes the amplitude loss due to wavefront divergence.
2. Apply an exponential gain or minimum-phase inverse Q filter (Hale, 1982). This compensates for frequency attenuation.
3. Optionally apply signature deconvolution to marine data. For vibroseis data, apply the filter that converts the Klauder wavelet to its minimum-phase equivalent.
4. Apply surface-consistent deconvolution (Appendix B.6). This takes care of the lateral variations effect on the wavelet because of inhomogeneities in the vicinity of sources and receivers. In Step 2, the vertical variations effect on the wavelet is handled.
5. Apply TVSW after stack and before TVF. This provides further flattening of the spectrum within the signal bandwidth without affecting phase.
6. Optionally apply predictive deconvolution after stack to remove short-period reverberations.

Basically, the idea is to do as much deterministic deconvolution as possible. Inverse Q, signature deconvolution, and the filter that converts the Klauder wavelet to its minimum-phase equivalent are deterministic operators. Any remaining issues then are handled by statistical means. However, for most data cases, just doing the geometric spreading correction followed by predictive deconvolution is adequate. This is all that is required in most data cases. The scheme outlined above has not been in use for long; its usefulness or its practicality still is uncertain. The inverse Q filter still is a research topic because of the issues concerning its phase response.

REFERENCES

Backus, M. M., 1959, Water reverberations: Their nature and elimination: Geophysics, **24**, 233-261.

Claerbout, J. F., 1976, Fundamentals of geophysical data processing: McGraw-Hill Book Co.

Gibson, B. and Larner, K. L., 1982, Comparison of spectral flattening techniques: unpublished technical document, Western Geophysical Company.

Goupillaud, P., 1961, An approach to inverse filtering of near-surface layer effects from seismic records: Geophysics, **26**, 754-760.

Hale, I. D., 1982, Q-adaptive deconvolution: Stanford Exploration Project, Report Number 30, 133-158.

Kjartannson, E., 1979, Constant Q—wave propagation and attenuation: J. Geophys. Res., **84**, 4737-4748.

Peacock, K. L. and Treitel, S., 1969, Predictive deconvolution—theory and practice: Geophysics, **34**, 155-169.

Ristow, D. and Jurczyk, D., 1975, Vibroseis deconvolution: Geophys. Prosp., **23**, 363-379.

Robinson, E. A., 1967, Multichannel time series analysis with digital computer programs: Holden-Day.

Robinson, E. A., 1966, Multichannel z-transforms and minimum-delay: Geophysics, **31**, 482-500.

Robinson, E. A. and Treitel, S., 1980, Geophysical signal analysis: Prentice-Hall, Inc.

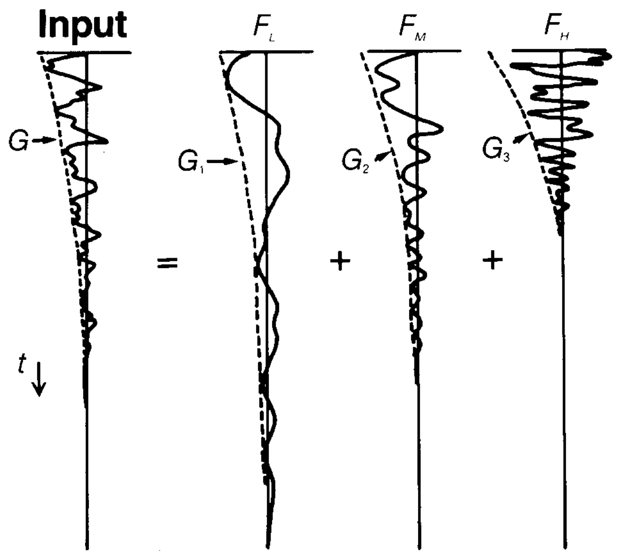


FIG. 2-85. A schematic illustration of rate of decay of frequencies in a seismic trace (Gibson and Larner, 1982).

attenuation function, then it would be easy to compensate for its effect. Several models for Q have been proposed. The constant-Q model is quite plausible and the easiest with which to deal (Kjartannson, 1979). However, the big problem of estimating Q from seismic data still remains. This problem recently has been given a lot of attention and the inverse Q-filter has crept into the processing flow.

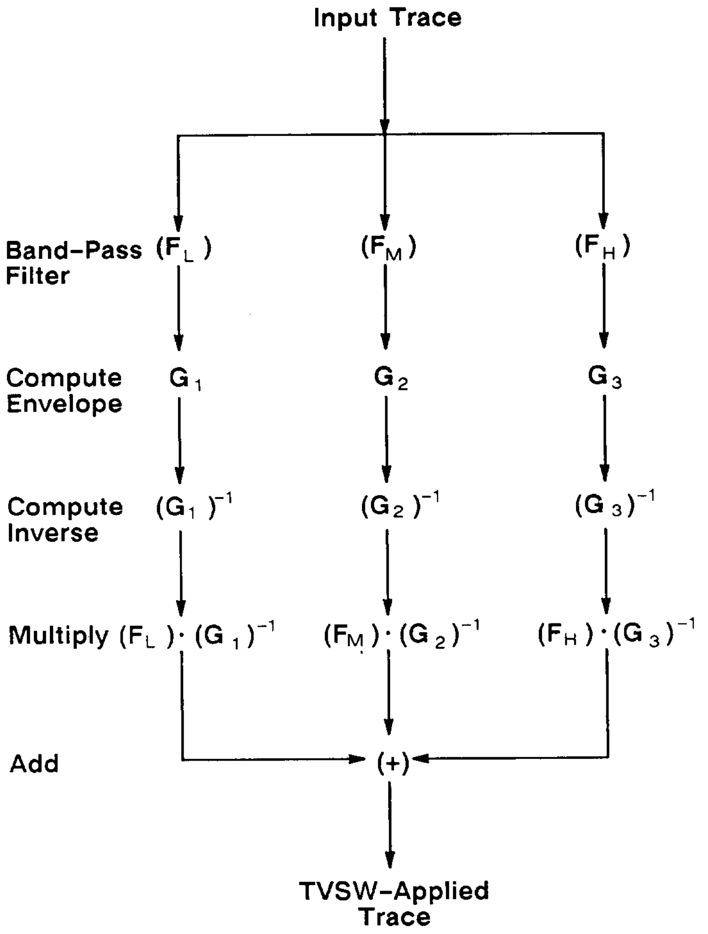


FIG. 2-86. A flowchart for time-variant spectral whitening.

2.8.1 Time-Variant Spectral Whitening

Frequency attenuation and a way to compensate for it are illustrated in Figure 2-85. Let us assume that we have an input seismogram with amplitudes decaying in time, as depicted. Now apply a series of narrow band-pass filters to this trace. Examine the field record in Figure 2-83 and note that the low-frequency component of trace F_L has a lower decay rate than the moderate-frequency component F_M. Likewise, the moderate-frequency component has a lower decay rate than the high-frequency component of the signal F_H. A series of gain functions, such as G_1, G_2, G_3, can be computed to describe the decay rates for each frequency band. This is done by computing the envelope of the band-pass filtered traces. The inverses of these gain functions then are applied to each frequency band and the results are summed. This is the time-variant spectral whitening (TVSW) process. The number of filter bands, the width of each band, and the overall bandwidth of application of TVSW are parameters that can be prescribed for a particular application. Figure 2-86 shows the steps involved in the TVSW process, which can be thought of as a process that both corrects for attenuation effects and partially deconvolves the effective source wavelet.

Figures 2-87 and 2-88 show some field records before and after spiking deconvolution, respectively. Note that this process not only has compressed the wavelet, but also has tried to suppress any reverberations in the data. On the other hand, TVSW mainly has compressed the wavelet without changing much of the ringy character of the data (Figure 2-89). Also note that little was done explicitly to the phase. Therefore, the action of TVSW may be close to a zero-phase deconvolution, although there is no rigorous theoretical proof of this. In practice, one of the main differences between TVSW and conventional deconvolution is that the former seems to be able to do a better job of flattening the amplitude spectrum. This difference can be significant for broad-band data with large dynamic range.

Spectral flattening can be achieved by an alternate approach in the frequency domain. As discussed in Appendix B.4, minimum-phase spiking deconvolution can be formulated in the frequency domain. Alternatively, we can flatten the amplitude spectrum without touching the phase. This is called zero-phase frequency-domain deconvolution. When performed over multiple time gates down the trace, it is essentially equivalent to TVSW. If we only want to flatten the spectrum, then the approach shown in Figure 2-90 can be taken.

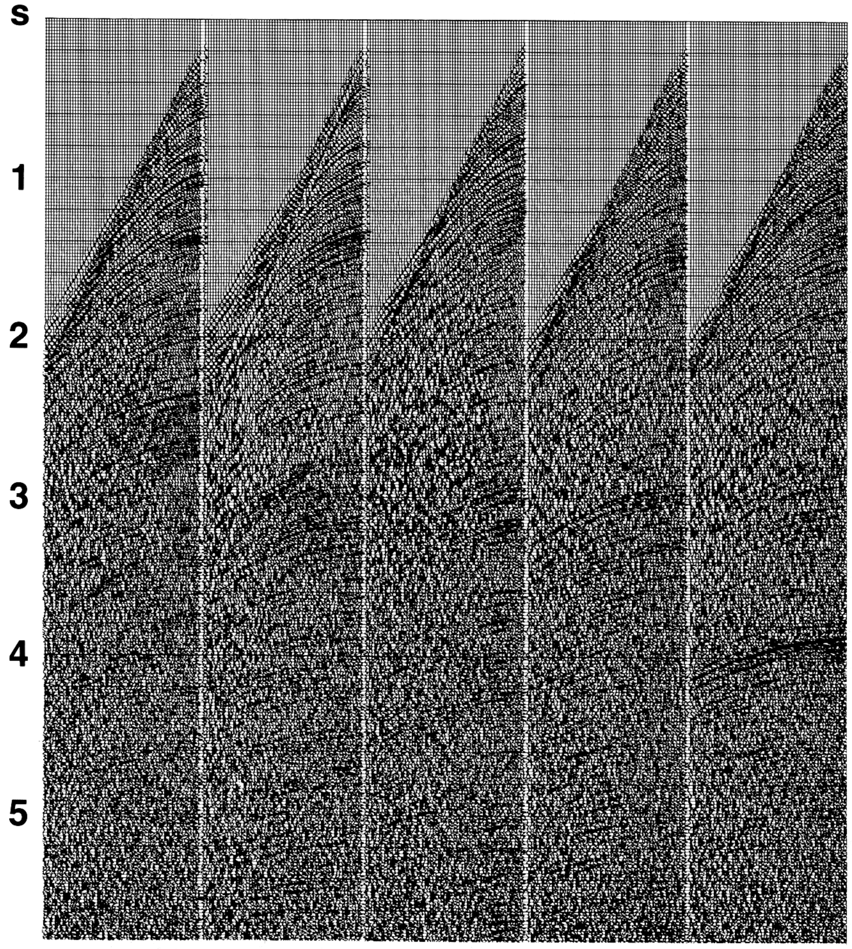


FIG. 2-91. Minimum-phase frequency-domain deconvolution applied to the CMP gathers in Figure 2-87. Compare this with Figures 2-88, 2-89, and 2-92.

Taner, M. T. and Coburn, K., 1981, Surface-consistent deconvolution: Presented at the 51st Ann. Int. Soc. Explor. Geophys. Mtg.

Treitel, S. and Robinson, E. A., 1966, The design of high resolution filters: Inst. Electr. Electron. Eng., **GE-4**, *1*.

Walden, A. T. and Hosken, J. W. J., 1984, An investigation of the spectral properties of primary reflections coefficients: Presented at the 46th Ann. Eur. Assoc. Explor. Geophys. Mtg.

Yilmaz, O., 1974, The problems of resolution and reverberations in reflection seismology: J. Geophys. Soc. Turkey, 5, 2.

EXERCISES

Exercise 2-1. Write the z-transform of wavelet $[1,0,(-\frac{1}{4})]$. Design a three-term inverse filter and apply it to the original. Hint: The z-transform of the wavelet can be written as a product of two doublets, $[1,(-\frac{1}{2})z]$ and $[1,(\frac{1}{2})z]$.

Exercise 2-2. Consider the following set of wavelets:

Wavelet A: $(3,-2,1)$
Wavelet B: $(1,-2,3)$

a. Plot the percent cumulative energy as a function of time delay. Use Robinson's energy delay theorem to determine the minimum-and maximum-phase wavelet.

b. Set up matrix equation (2.21) for each wavelet, compute the spiking deconvolution operators, then apply them.

c. Let the desired output be (0,0,1,0). Set up matrix equation (2.20) for each wavelet, compute the shaping filters, and apply them. Find that the least error occurs for wavelet A with the zero-delay spike and for wavelet D with the delayed spike.

Exercise 2-3. Consider wavelet A in Exercise 2-2. Set up matrix equation (2.22), where $\epsilon = 0.001, 0.01, 0.1$. Note that $\epsilon = 0$ already is assigned in Exercise 2-2. As the percent prewhitening increases, the spikiness of the deconvolution output decreases.

Exercise 2-4. Consider a multiple series associated with a water bottom with a reflection coefficient c_w and two way time t_w. Design an inverse filter to suppress the multiples. [This is called the Backus filter (Backus, 1959)].

Exercise 2-5. Consider the following earth model:

Two-Way Time,
ms
0 ____ Surface
500 ____ Water Bottom
750 ____ Deep Reflector

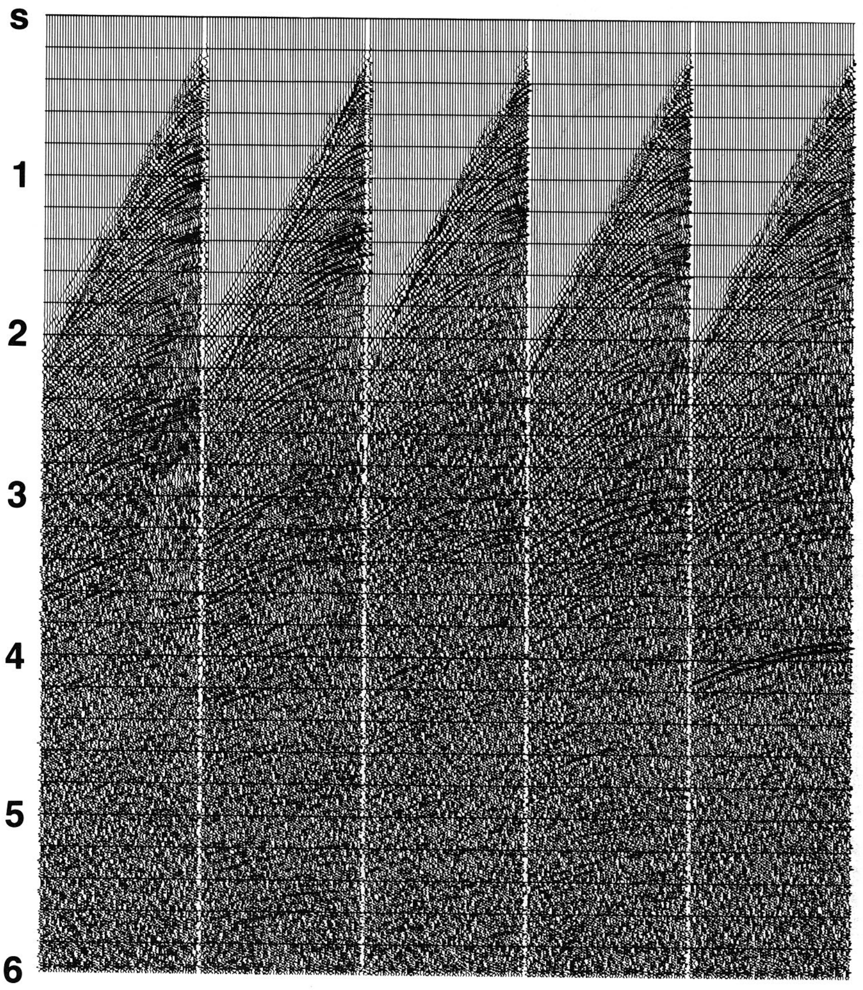


FIG. 2-92. Zero-phase frequency-domain deconvolution applied to the CMP gathers in Figure 2-87. Compare this with Figures 2-88, 2-89, and 2-91.

What prediction lag and operator length should you choose to suppress:

a. water-bottom multiples, and
b. peg-leg multiples?

Exercise 2-6. Refer to Figure 2-86. Consider the following three bandwidths with TVSW application:

FL: 10 to 30 Hz
FH: 30 to 50 Hz
FM: 50 to 70 Hz

What kind of slopes should you assign to each bandwidth so that the output trace has an amplitude spectrum that is unity over the 10 to 70-Hz bandwidth?

Exercise 2-7. If the signal character down the trace changes rapidly (strong nonstationarity), should you consider narrow or broad bandwidths for the filters used in TVSW?

Exercise 2-8. Consider a minimum-phase wavelet and the following two processes applied to it:

a. Spiking deconvolution followed by 10- to 50-Hz zero-phase band-pass filtering. b. Shaping filter to convert the minimum-phase wavelet to a 10- to 50-Hz zero-phase wavelet. What is the difference between the two outputs?

Exercise 2-9. How would you design a minimum-phase band-pass filter operator?

Exercise 2-10. Consider (a) convolving a minimum-phase wavelet with a zero-phase wavelet, (b) adding two minimum-phase wavelets. Are the resulting wavelets minimum-phase?

Exercise 2-11. Consider the sinusoid shown in Figure 1-1 (frame 1) as input to spiking deconvolution. What is the output?

3

Velocity analysis, statics corrections, and stacking

3.1 INTRODUCTION

A sonic log represents direct measurement of the velocity with which seismic waves travel in the earth as a function of depth. Seismic data, on the other hand, provide an indirect measurement of velocity. Based on these two types of information, the exploration seismologist derives a large number of different types of velocity, such as interval, apparent, average, root-mean-square (rms), instantaneous, phase, group, normal moveout (NMO), stacking, and migration velocities. However, the velocity that can reliably be derived from seismic data is the velocity that yields the best stack. Assuming a layered media, stacking velocity is related to normal moveout velocity. This, in turn, is related to rms velocity [equation (3.4)], from which the average and interval velocities are derived.

Interval velocity is the average velocity in an interval between two reflectors. Several factors influence interval velocity within a rock unit with a certain lithologic composition:

1. Pore shape
2. Pore pressure
3. Pore fluid saturation
4. Confining pressure
5. Temperature.

These factors have been investigated extensively under laboratory conditions. Figure 3-1 shows laboratory measurements of velocity as a function of the confining pressure in a Bedford limestone sample with pores in the form of microcracks. The experiment was conducted using enclosed samples to control the pore fluid pressure independent of the confining pressure. Both compressional (*P*) and shear (*S*) wave velocities increase with increasing confining pressure. More specifically, velocity generally increases rapidly with confining pressure at small confining pressures, then gradually levels off at high confining pressures (Figure 3-1). The reason for this is that as the confining pressure increases, pores close. However, at a high confining pressure, not much deformable pore space is left. Therefore, any further increase in the confining pressure will not cause a significant increase in velocity. From Figure 3-1, note that *P*-wave velocity is greater than *S*-wave velocity, regardless of confining pressure. This is true for any rock type. Finally, from Figure 3-1, we see the effect of fluid saturation in pores. The saturated rock sample has a higher *P*-wave velocity than the dry sample at low confining pressure (why is that?). At high confining pressures, *P*-wave velocity in the dry sample approaches the magnitude of the *P*-wave velocity in the saturated sample. Note also that the *P*-wave velocity in the saturated sample does not change as rapidly as in the dry sample. This is because the fluid is almost as incompressible as the rock. Whether the pores are filled with fluid or not has little effect on *S*-wave velocity, since fluids cannot support shear-wave propagation.

We now examine velocity as a function of confining pressure for an enclosed sample of Berea sandstone with rounded pores (Figure 3-2). Again, note the increase in velocity with increasing confining pressure. The important difference between this sample and the one in Figure 3-1 is the range of magnitude of the velocity. The rock with microcracks has a higher velocity than the rock with rounded pores at any given confining pressure. The reason for this is that it is easier to close the pores formed as microcracks than it is to close those that are round.

The most prominent factor influencing velocity in a rock of given lithology and porosity probably is confining pressure. This type of pressure arises from the overburden and increases with depth. It is generally true that velocity increases with depth. However, because of factors such as pore pressure, there may be inversion in the velocity within a layer. Figure 3-3 shows the variation of velocity with depth for various types of lithology. Tertiary clastics, which usually are less indurated than other rocks, occupy the low-velocity end of the graph. They generally start out with a velocity that ranges from 1.5 to 2.5 km/s at or near the surface, then gradually increase to from 4.5 to 5.5 km/s at depths greater than 5 km. Carbonates with high porosity occupy the central portion of the graph, starting at about 3 km/s and increasing to nearly 6 km/s. On the other hand, low-porosity carbonates have a smaller range of variation in velocity. If there is not much pore space to close, then the confining pressure cannot cause much of an increase in velocity.

This chapter discusses ways to estimate velocities from seismic data. Velocity estimation requires the data recorded at nonzero offsets provided by common-midpoint (CMP) recording. With estimated velocities, we can correct for nonzero offset and compress the recorded data volume (in midpoint-offset-time coordinates) to a stacked section (Figure 1-34).

For a single constant-velocity horizontal layer, the traveltime curve as a function of offset is a hyperbola (Section 3.2). The time difference between traveltime at a given offset and at zero offset is called *normal moveout* (NMO). The velocity required to correct for normal moveout is called the *normal moveout velocity*. For a single horizontal reflector, the NMO velocity is equal to the velocity of the medium above the reflector. For a dipping reflector, the NMO velocity is equal to the medium velocity divided by the cosine of the dip angle. When the dipping reflector is viewed in three dimensions, then the azimuth angle (between the dip direction and the profiling direction) becomes an additional factor. Traveltime as a function of offset from a series of plane horizontal isovelocity layers is approximated by a hyperbola. This approximation is better at small offsets than large offsets. For short offsets, the NMO velocity for a horizontally layered earth is equal to the rms velocity down to the layer boundary under consideration. In a medium composed of layers with arbitrary dips, the traveltime equation gets complicated. However, in practice, as long as dips are gentle and spread is small (less than reflector depth), the hyperbolic assumption still can be made. For layer boundaries with arbitrary shapes, the hyperbolic assumption breaks down.

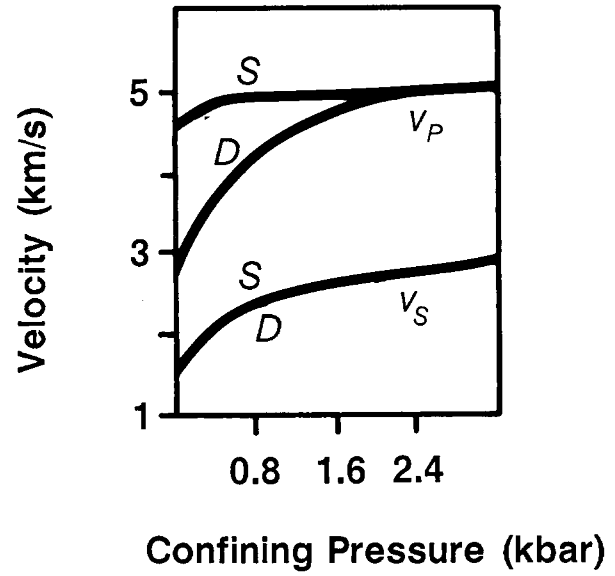


FIG. 3-1. Change of P- and S-wave velocities as a function of confining pressure observed in dry and water-saturated Bedford limestone samples with pores in the form of microcracks. Fluid volume has been kept constant during measurements. Here, S = saturated, D = dry, v_P = P-wave velocity, and v_S = S-wave velocity. (Adapted from Nur, 1981.) 1981.)

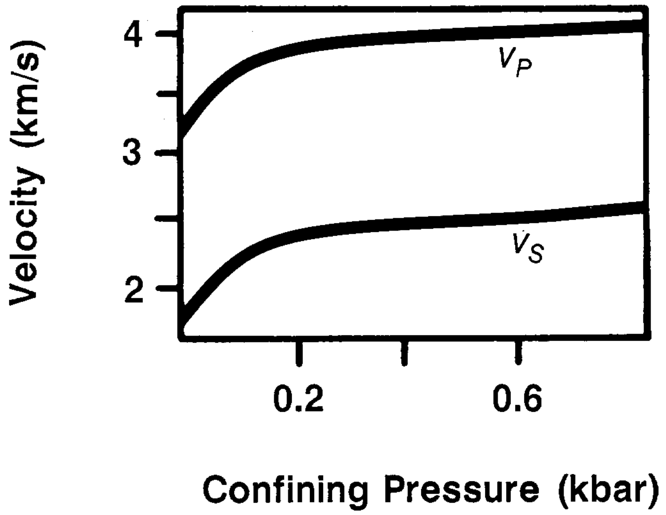


FIG. 3-2. Change of P- and S-wave velocities as a function of confining pressure observed in Berea sandstone samples with rounded pores. Fluid volume has been kept constant during measurements. Here, v_P = P-wave velocity and v_s = S-wave velocity. (Adapted from Nur, 1981.)

There is a difference between the NMO and stacking velocities that often is ignored in practice. The NMO velocity is based on the small-spread hyperbolic traveltime (Taner and Koehler, 1969; Al-Chalabi, 1973), while stacking velocity is based on the hyperbola that best fits data over the entire spread length. Nevertheless, stacking velocity and NMO correction velocity generally are considered equivalent.

Conventional velocity analysis is based on the hyperbolic assumption. Various approaches to velocity analysis are discussed in Section 3.3. The hyperbolic traveltime equation is linear in the (t^2 − x^2) plane. Zero-offset time and stacking velocity for a given reflector can be estimated from the line that best fits the traveltime picks plotted on the (t^2) versus (x^2) plane. Another way to estimate the NMO velocity is to apply different NMO corrections to a CMP gather using a range of constant velocity values, then display them side by side. The velocity that best flattens each event as a function of offset is picked as its NMO velocity. Alternatively, a small portion of a line can be stacked with a range of constant velocity values. These constant-velocity stacks then can be plotted in the form of a panel, called CVS panel. Stacking velocities that yield the desired stack then can be picked from the CVS panel.

Another commonly used velocity analysis technique is based on computing the *velocity spectrum*. The idea is to display some measure of signal coherency on a graph of velocity versus two-way zero-offset time. The underlying principle is to compute the signal coherency on the CMP gather in small time gates that follow a hyperbolic trajectory. Stacking velocities are interpreted from velocity spectra by choosing the velocity function that produces the highest coherency at times with significant event amplitudes.

Occasionally the stacking velocity variation needs to be determined in detail along a particular reflector. Horizon-oriented velocity analysis provides the stacking velocity variation in the lateral direction along a particular horizon of interest. Various practical aspects of this method also are discussed based on a real data example.

Reflection traveltimes are not always hyperbolic in horizontally layered media. One reason that traveltime often deviates from a perfect hyperbola is the presence of static time shifts caused by near-surface velocity variations. Statics can severely distort the reflection hyperbola when there are large surface elevation changes or when the weathering layer varies horizontally. Residual statics variations usually remain in the data even after initial corrections for estimated weathering layer variations and elevation changes (i.e., field statics) (Section 3.6). Therefore, corrections for these residual statics normally must be estimated and applied to CMP gathers before stacking. Estimation is done after a preliminary NMO correction using either a regional velocity function or information from a series of preliminary velocity analyses along the line. Following the residual statics corrections, velocity analyses usually are repeated to upgrade the velocity picks for stacking. Various aspects of residual statics corrections are discussed in Sections 3.4 and 3.5.

As a final note, the velocities required by stacking and migration are not necessarily the same. In fact, for data collected parallel to the dip direction of a single dipping reflector, stacking velocity is the velocity of the medium

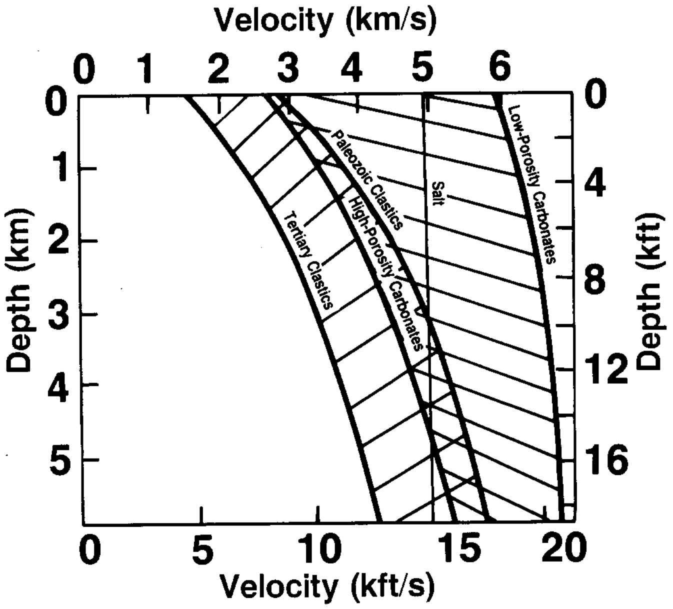


FIG. 3-3. Velocity range for rocks of different lithologic compositions at different depths of burial. (Adapted from Sheriff, 1976; courtesy American Association of Petroleum Geologists.)

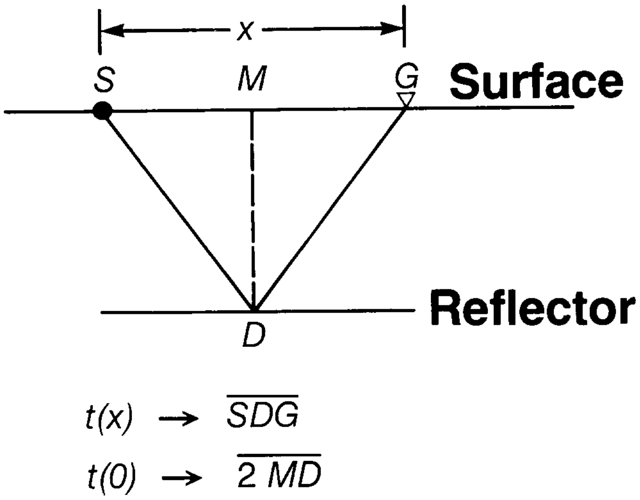


FIG. 3-4. The NMO geometry for a single horizontal reflector [refer to equation (3.1)].

above the reflector, divided by the cosine of the dip angle, while the migration velocity is the velocity of the medium itself. In other words, stacking velocity is sensitive to dip angle, while migration velocity is not. In Section 4.5, a wave-theoretical method for determining migration velocities is introduced.

3.2 NORMAL MOVEOUT

Figure 3-4 shows the simple case of a single horizontal layer. At a given midpoint location M, the traveltime along the raypath from shot position S to depth point D, then back to receiver position G is $t(x)$. Using the Pythagorean theorem, the traveltime equation as a function of offset is

$$t^2(x) = t^2(0) + x^2/v^2, \tag{3.1}$$

where x is the distance (offset) between the source and receiver positions, v is the velocity of the medium above the reflecting interface, and $t(0)$ is twice the traveltime along the vertical path MD. Note that vertical projection of depth point D to the surface, along the normal to the reflector, coincides with midpoint M. This is the case only when the reflector is horizontal. Equation (3.1) describes a hyperbola in the plane of two-way time versus offset. Figure 3-5 is an example of traces in a common midpoint (CMP) gather. The figure also represents a common depth point (CDP) gather, since all the raypaths associated with each source-receiver pair reflect from the same subsurface depth point D. The offset range in Figure 3-5 is 0 to 3150 m, with a 50-m trace separation. The medium velocity above the reflector is 2264 m/s. All of the traces in this CMP gather contain a reflection from the same depth point. The difference between the two-way time at a given offset $t(x)$ and the two-way zero-offset time $t(0)$ is called NMO. From equation (3.1), we see that velocity can be computed when offset x and two-way times $t(x)$ and $t(0)$ are known.

Once the NMO velocity is estimated, the traveltimes can be corrected to remove the influence of offset as shown in Figure 3-6. Traces in the NMO-corrected gather then are summed to obtain a stack trace at the particular CMP location.

The numerical procedure involved in hyperbolic moveout correction is illustrated in Figure 3-7. The idea is to find the amplitude value at A' on the NMO-corrected gather from the amplitude value at A on the original CMP gather. Given quantities $t(0)$, x, and v_{NMO}, compute $t(x)$ from equation (3.1). Assume that this is 1003 ms. If the sampling interval were 4 ms, then this time is equivalent to the 250.25 sample index. Therefore, the amplitude value must be computed at this time using the amplitudes at the neighboring integer sample values.

The NMO correction is given by the difference between $t(x)$ and $t(0)$:

$$\begin{aligned}\Delta t_{\mathrm{NMO}} &= t(x) - t(0)\\ &= t(0)\left\{\left[1 + \left(\frac{x}{v_{\mathrm{NMO}}\, t(0)}\right)^2\right]^{1/2} - 1\right\}.\end{aligned} \tag{3.2}$$

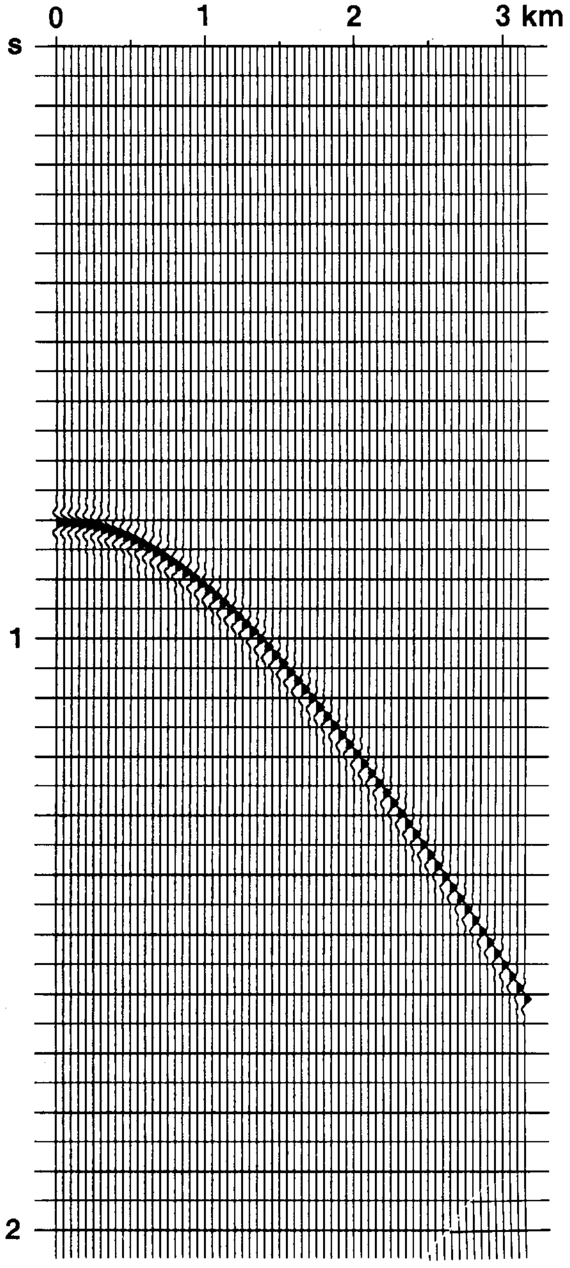


FIG. 3-5. Synthetic CMP gather associated with the geometry in Figure 3-4. Traveltime curve for a flat reflector is a hyperbola with its apex at zero-offset trace.

Table 3-1. NMO correction as a function of offset x and two-way zero-offset time for a given velocity function.

		Δt_{NMO}, s	
$t(0)$, s	v_{NMO}, m/s	$x = 1000$ m	$x = 2000$ m
0.25	2000	0.309	0.780
0.5	2500	0.140	0.443
1	3000	0.054	0.201
2	3500	0.020	0.080
4	4000	0.008	0.031

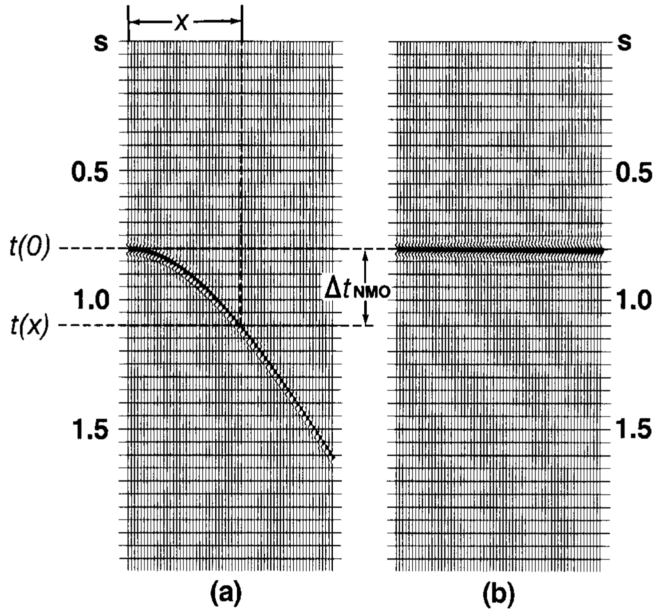


FIG. 3-6. NMO correction [equation (3.2)] involves mapping nonzero-offset traveltime $t(x)$ onto zero-offset traveltime $t(0)$. (a) Before and (b) after NMO correction.

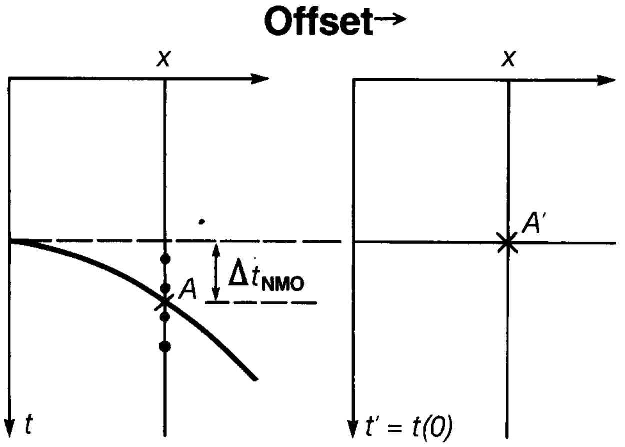


FIG. 3-7. NMO correction using a computer. For given integer value $t(0)$, velocity, and offset, compute $t(x)$ using equation (3.1). The amplitude at $t(x)$ value, denoted by A, does not necessarily fall onto an input integer sample location. By using two samples on each side of $t(x)$ (denoted by solid dots), we can interpolate between the four amplitude values to compute the amplitude value at $t(x)$. This amplitude value then is mapped onto output integer sample $t(0)$ (denoted by A') at the corresponding offset.

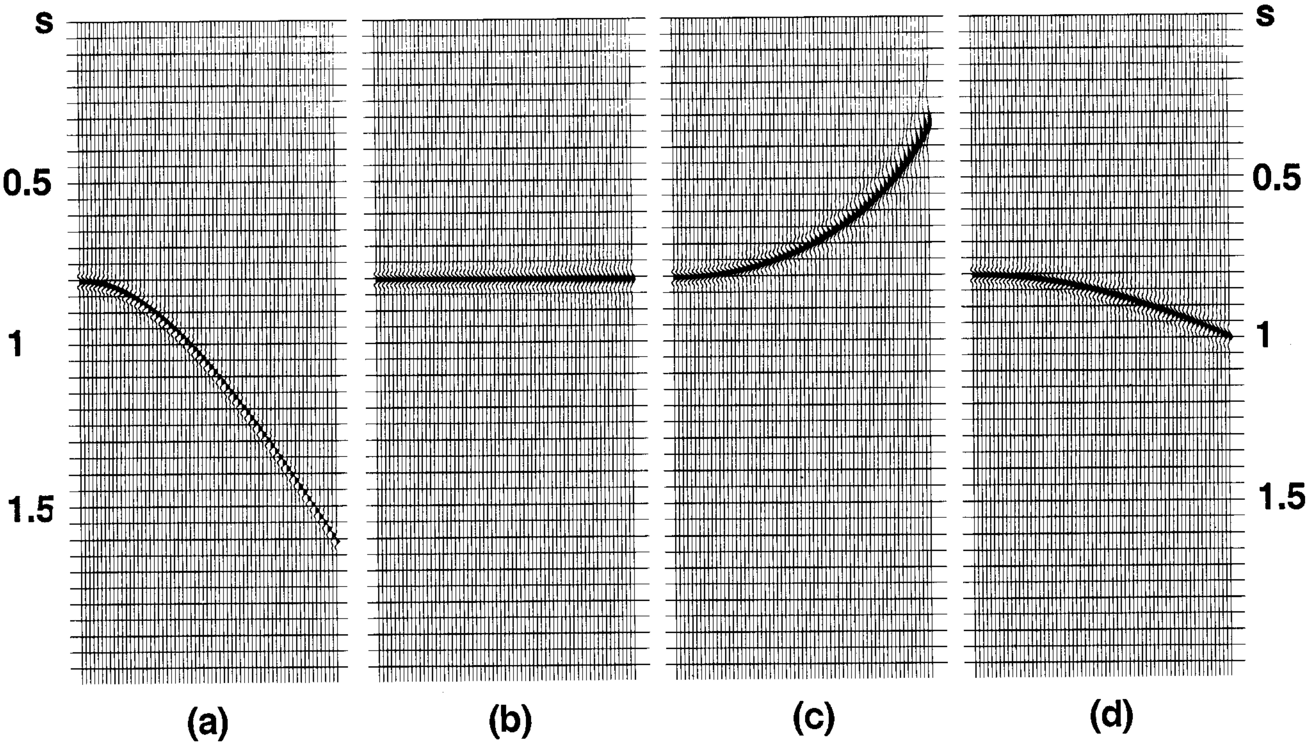


FIG. 3-8. (a) CMP gather containing a single event with a moveout velocity of 2264 m/s; (b) NMO-corrected gather using the appropriate moveout velocity; (c) overcorrection because too low a velocity (2000 m/s) was used in equation (3.2); (d) undercorrection because too high a velocity (2500 m/s) was used in equation (3.2).

Table 3-1 shows the moveout corrections for two different offset values using a realistic velocity function that increases with reflector depth. From Table 3-1, note that NMO increases with offset and decreases with depth. The NMO also is smaller for larger velocity values.

For a flat reflector with an overlying homogeneous medium, the reflection hyperbola can be corrected for offset if the correct medium velocity is used in the NMO equation. From Figure 3-8, if a velocity higher than the actual medium velocity (2264 m/s) is used, then the hyperbola is not completely flattened. This is called undercorrection. On the other hand, if a lower velocity is used, then overcorrection results. Figure 3-8 illustrates the basis of conventional velocity analysis. NMO correction is applied to the input CMP gather using a number of trial constant velocity values in equation (3.2). The velocity that best flattens the reflection hyperbola is the velocity that best corrects for NMO before stacking the traces in the gather. Furthermore, for a simple case of a single horizontal reflector, this velocity also is equal to the velocity of the medium above the reflector.

3.2.1 NMO in a Horizontally Stratified Earth

We now consider a medium composed of horizontal isovelocity layers (Figure 3-9). Each layer has a certain thickness that can be defined in terms of two-way zero-offset time. The layers have interval velocities (v_1, v_2, . . .,v_N), where N is the number of layers. Consider the raypath from source S to depth point D, back to receiver R, associated with offset x at midpoint location M. Taner and Koehler (1969) derived the traveltime equation for this path as:

$$t^2(x) = C_0 + C_1x^2 + C_2x^4 + C_3x^6 + \cdots, \qquad (3.3)$$

where $C_0 = t^2(0)$, $C_1 = 1/v_{\text{rms}}^2$, and C_2, C_3, . . . are complicated functions that depend on layer thicknesses and interval velocities. The rms velocity v_{rms} down to the reflector on which depth point D is situated is defined as

$$v_{\text{rms}}^2 = \frac{1}{t(0)} \sum_{i=1}^{N} v_i^2 \Delta t_i(0), \qquad (3.4)$$

where Δt_i is the vertical two-way time through the ith layer and $t(0) = \sum_{k=1}^{i} \Delta t_k$. By making the *small-spread approximation* (offset small compared to depth), the series in equation (3.3) can be truncated as follows:

$$t^2(x) = t^2(0) + x^2/v_{\text{rms}}^2. \qquad (3.5)$$

When equations (3.1) and (3.5) are compared, we see that the velocity required for NMO correction for a horizontally stratified medium is equal to the rms velocity, provided the small-spread approximation is made.

How much error is caused by dropping the higher order terms in equation (3.3)? Figure 3-10a shows a CMP gather based on the velocity model in Figure 3-11. Traveltimes to all four reflectors were computed by the raypath integral equations (Grant and West, 1965) that exactly describe wave propagation in a horizontally layered earth with a given interval velocity model. We now replace the layers above the second shallow event at t(0) = 0.8 s with a single layer with velocity equal to the rms velocity down to this reflector; i.e., 2264 m/s. The resulting traveltime curve, computed using equation (3.5), is shown in Figure 3-10b. This procedure is repeated for the deeper events at $t(0)$ = 1.2 and 1.6 s in Figures 3-10c and d. Note that the traveltime curves in Figures 3-10b, c, and d are perfect hyperbolas. How different are the traveltime curves in Figure 3-10a from these hyperbolas? After careful examination, note that the traveltimes are slightly different for the shallow events at $t(0)$ = 0.8 and 1.2 s only at large offsets, particularly beyond 3 km. By dropping the higher order terms, we approximate the reflection times in a horizontally layered earth with a *small-spread hyperbola*.

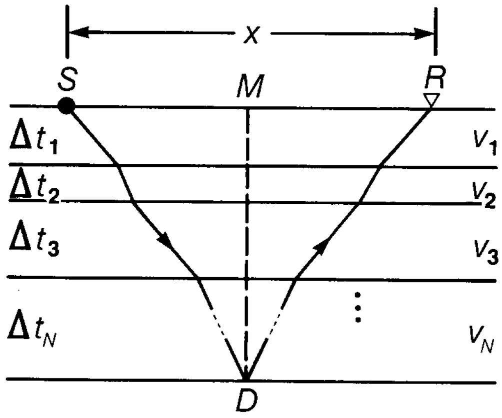


FIG. 3-9. NMO for a horizontally layered earth model [geometry for equation (3.3)].

3.2.2 NMO Stretching

Figure 3-12b shows the CMP gather in Figure 3-10a after NMO correction. The rms velocity function shown in Figure 3-11 was used in equation (3.5) for this correction. As a result of the NMO correction, a frequency distortion occurs, particularly for shallow events and at large offsets. This is called NMO stretching and is illustrated in Figure 3-13. The waveform with a dominant period T is stretched so that its period T', after NMO correction, is greater than T. Stretching is a frequency distortion in which events are shifted to lower frequencies. Stretching is quantified as

$$\Delta f/f = \Delta t_{\mathrm{NMO}}/t(0) \tag{3.6}$$

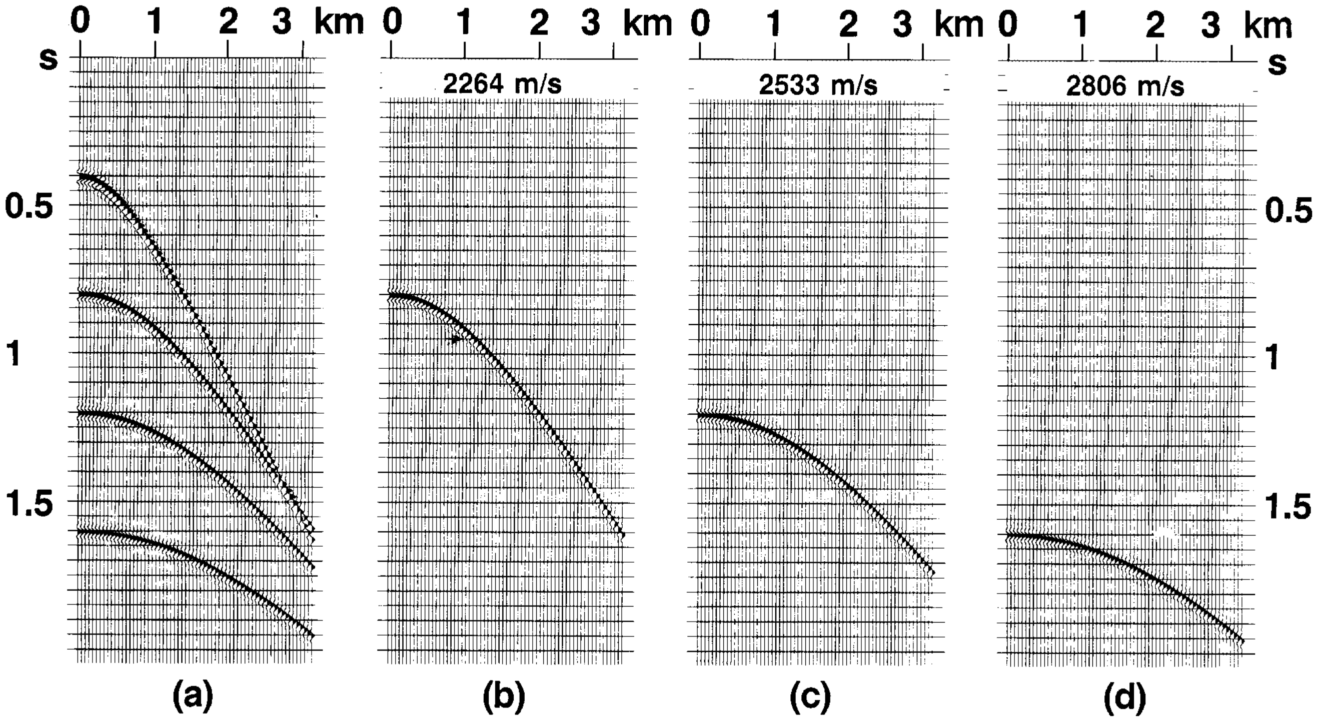


FIG. 3-10. (a) A synthetic CMP gather derived from the velocity function depicted in Figure 3-11; (b), (c), and (d) are CMP gathers derived from the rms velocities (indicated at the top of each gather) associated with the second, third, and fourth reflectors from the top. The traveltimes in (a) were derived using the raypath integral equations for horizontally layered earth.

Table 3-2. NMO stretching.

$t(0)$, s	v_{nmo}, m/s	% $\Delta f/f$ for	
		x = 1000 m	x = 2000 m
0.25	2000	123	312
0.5	2500	28	89
1	3000	5	20
2	3500	1	4
4	4000	0.2	0.8

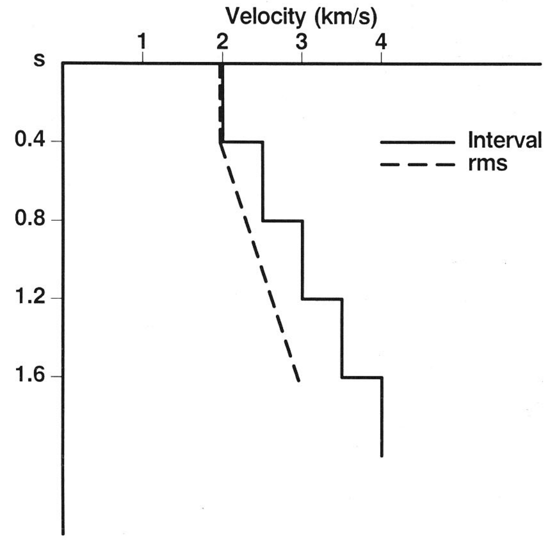
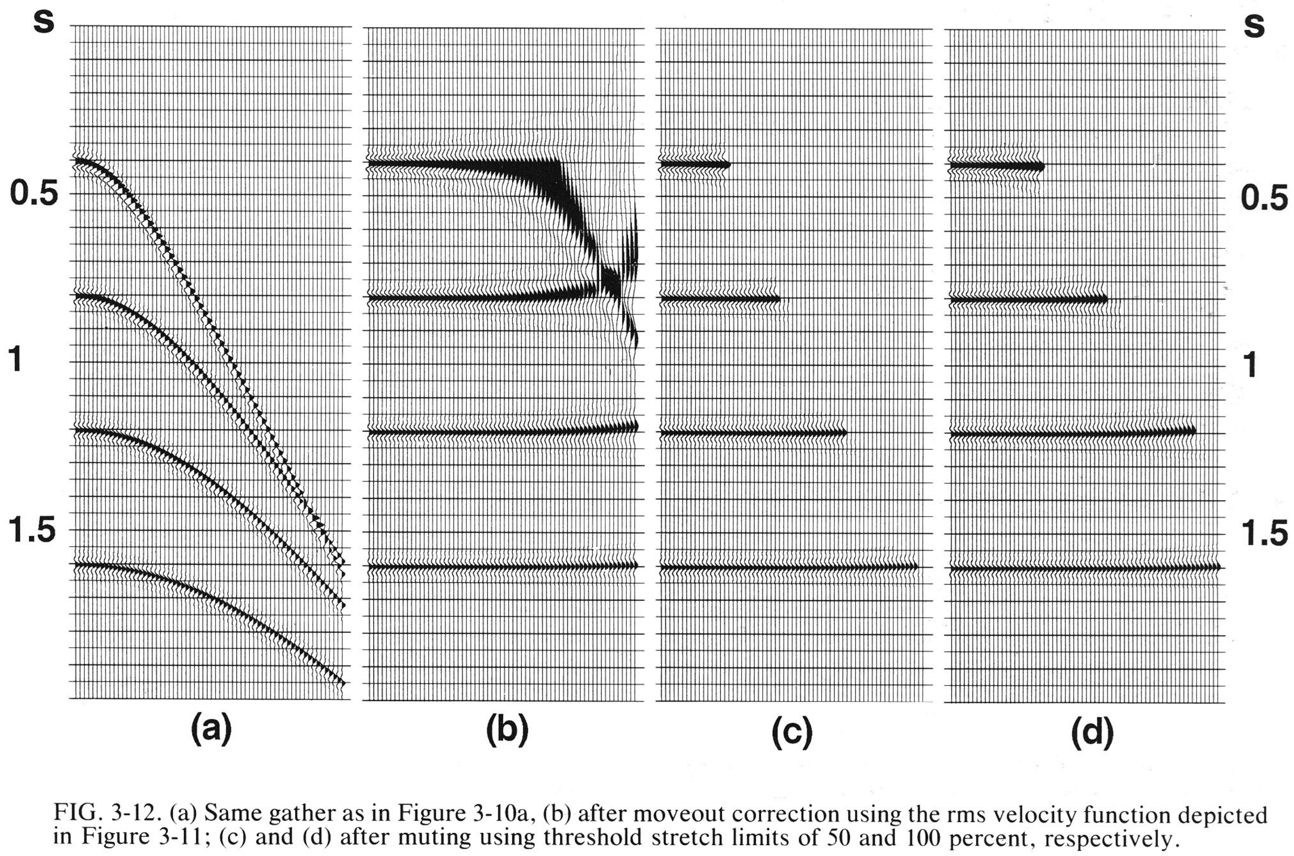


FIG. 3-11. A hypothetical velocity function used in generating the synthetic CMP gather in Figure 3-10a.

where f is the dominant frequency, Δf is change in frequency, and Δt_{NMO} is given by equation (3.2). Derivation of equation (3.6) is left for Exercise 3.14.

Table 3-2 lists the percent frequency changes due to the NMO stretching associated with the velocity function in Table 3-1. Note that stretching mainly is confined to large offsets and shallow times. For example, a waveform with a 30-Hz dominant frequency at 2000-m offset and $t(0)$ = 0.25 s shifts to nearly 10 Hz after NMO correction.

Because of the stretched waveform at large offsets, stacking the NMO-corrected CMP gather (Figure 3-12b) will severely damage the shallow events. This problem can be solved by muting the stretched zones in the gather. Automatic muting is done by using the quantitative definition of stretching given in equation (3.6). Figures 3-12c and d show two versions of the CMP gather after NMO correction and muting; one version has a stretch limit of 50 percent, while the other has a stretch limit of 100 percent. The 50 percent stretch limit does not show significant frequency distortion. However, the stretch limit can be extended to 100 percent because we want to include as much of the CMP gather in the stack as possible without degradation. A trade-off exists between the signal-to-noise (S/N) ratio and mute. In particular, if the S/N ratio is good, then it may be preferable to mute more than stretch mute requirements to preserve signal bandwidth. On the other hand, if the S/N ratio is poor, it may be necessary to accept a large amount of stretch to get any events on the stack. A real data example is provided in Figure 3-14. Here, the stretched zone is seen



(a) (b) (c) (d)

FIG. 3-12. (a) Same gather as in Figure 3-10a, (b) after moveout correction using the rms velocity function depicted in Figure 3-11; (c) and (d) after muting using threshold stretch limits of 50 and 100 percent, respectively.

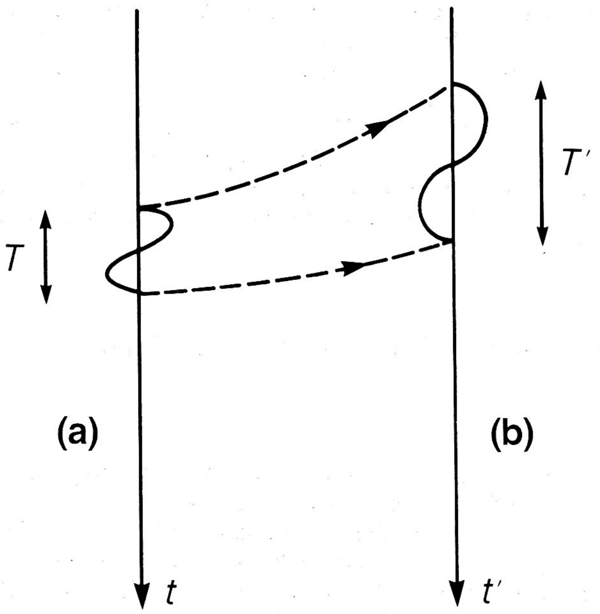


FIG. 3-13. A signal (a) with a period of T is stretched to a signal (b) with a period of $T' > T$ after NMO correction.

as the low-frequency zone at the shallow part of the CMP gathers without mute applied.

Another method for optimum selection of the mute zone is to progressively stack the data. Figure 3-15a is an NMO-corrected CMP gather without mute applied. Figure 3-15b shows the stack traces derived from the CMP gather (Figure 3-15a). The rightmost trace is the same as the rightmost trace in the input CMP gather. The second trace from the right is the sum of the two near-offset traces, and so on, progressively increasing stacking fold. The leftmost trace is the full-fold stack of the input CMP gather. By following the waveform along a certain event and observing where changes occur, the mute zone is derived as shown in Figure 3-15b. Figure 3-15c is the result of a poor mute choice based on the picks from the input gather (Figure 3-15a). Excessive muting can be hazardous. Large offsets often are needed to effectively suppress multiple reflections. A similar procedure can be followed to determine an inside mute. This time, the stacking fold is progressively increased in the near-offset direction.

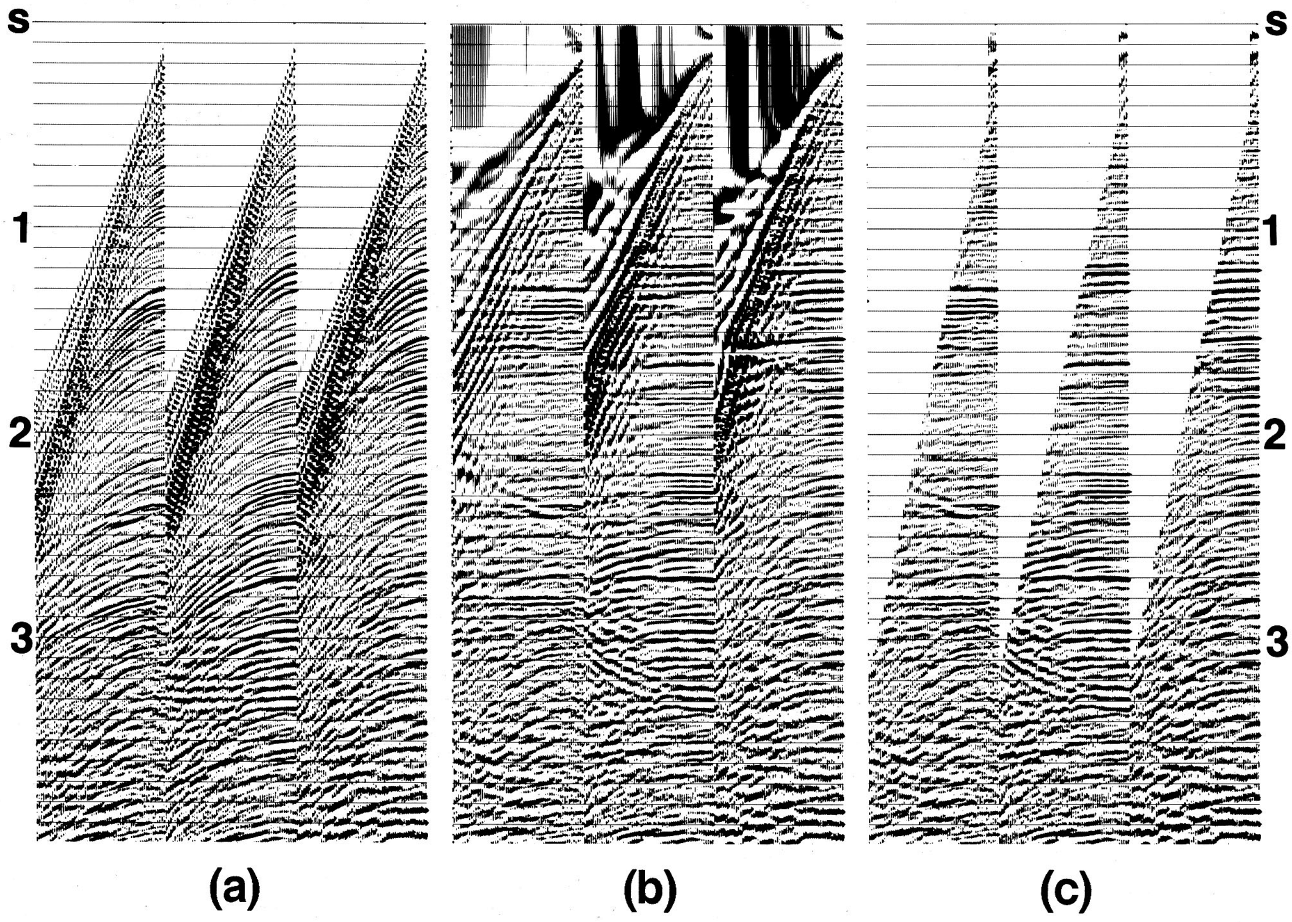


FIG. 3-14. NMO correction and muting of a stretched zone on field data; (a) CMP gathers, (b) NMO correction, (c) mute.

3.2.3 NMO for a Dipping Layer

Figure 3-16 shows a single dipping layer. We want to compute the traveltime from source location S to the reflector at depth point D, then back to receiver location G. For the dipping layer, midpoint M no longer is a vertical projection of the depth point to the surface. The terms *CDP gather* and *CMP gather* are equivalent only when the earth is horizontally stratified. When there is subsurface dip or lateral velocity variation, the two gathers are different. Midpoint M remains common to all of the source-receiver pairs within the gather, regardless of dip. Depth point D, however, is different for each source-receiver pair in a CMP gather recorded over a dipping reflector. Levin (1971), using the geometry of Figure 3-16, derived the following 2-D traveltime equation for a layer dipping at angle ϕ:

$$t^2(x) = t^2(0) + x^2 \cos^2 \phi / v^2. \tag{3.7}$$

Again, this is the equation of a hyperbola. However, the NMO velocity now is given by the medium velocity divided by the cosine of the dip angle:

$$v_{\mathrm{NMO}} = v/\cos \phi. \tag{3.8}$$

Proper stacking of a dipping event requires a velocity that is greater than the velocity of the medium above the reflector. Levin extended his work to a dipping plane interface in a three-dimensional (3-D) subsurface geometry as shown in Figure 3-17. In this case, the NMO velocity depends not only on the dip of the interface, but also on the source-receiver azimuth:

$$v_{\mathrm{NMO}} = v/(1 - \sin^2 \phi \cos^2 \theta)^{1/2}. \tag{3.9}$$

The azimuth, θ, is the angle between the structural dip direction and the direction of the actual profile (Figure 3-17).

An apparent dip angle is defined by

$$\sin \phi' = \sin \phi \cos \theta. \tag{3.10}$$

Using this definition, the NMO velocity given by equation (3.8) is rewritten as

$$v_{\mathrm{NMO}} = v/\cos \phi'. \tag{3.11}$$

This equation is identical to equation (3.8), which is for a dipping layer in a two-dimensional (2-D) subsurface geometry, except that equation (3.11) refers to apparent dip, while equation (3.8) refers to true dip angle.

Levin (1971) plotted the ratio of v_{NMO}/v [equation (3.9)] as a funciton of dip and azimuth. His results are reproduced in Figure 3-18. The horizontal axis is the azimuth. When the line is a dip line, the azimuth is zero; when it is a strike line, the azimuth is 90 degrees. The ratio v_{NMO}/v is significant when the line is shot at or near the structural dip direction. Levin (1971) also plotted v_{NMO}/v along the dip line as a function of small structural dip angles. This result is reproduced in Figure 3-19. When the dip does not exceed 15 degrees, then the ratio v_{NMO}/v is close to unity. There is a 4 percent difference between v_{NMO} and v for a 15-degree dip.

In conclusion, the NMO velocity for a dipping layer, whether 2-D or 3-D, depends on dip angle. A horizontal layer with a high velocity can yield the same moveout as a dipping layer with a low velocity. This is illustrated in Figure 3-20.

3.2.4 NMO for Several Layers with Arbitrary Dips

Figure 3-21 shows a 2-D subsurface geometry that is composed of a number of layers, each with arbitrary dip. We want to compute the traveltime from source location S to depth point D, then back to receiver location G, which is associated with midpoint M. Note that the CMP ray from midpoint M hits the dipping interface at normal incidence at D', which is not the same as D. The zero-offset time is the two-way time along the raypath from M to D'. Hubral and Krey (1980) derived the expression for traveltime $t(x)$ along SDG as

$$t^2(x) = t^2(0) + \frac{x^2}{v_{\mathrm{NMO}}^2} + \text{higher order terms}, \tag{3.12}$$

where the NMO velocity is

$$v_{\mathrm{NMO}}^2 = \frac{1}{t(0) \cos^2 \beta_0} \sum_{i=1}^{N} v_i^2 \, \Delta t_i(0) \prod_{k=1}^{i-1} \left(\frac{\cos^2 \alpha_k}{\cos^2 \beta_k} \right). \tag{3.13}$$

The angles are defined in Figure 3-21. For a single dipping layer, equation (3.13) reduces to equation (3.8). Moreover, for a horizontally stratified earth, equation (3.13) reduces to equation (3.4). As long as the dips are gentle and the spread is small, the traveltime equation is approximately represented by a hyperbola [equation (3.5)] and the velocity required for NMO correction is approximately the rms velocity function [equation (3.4)].

Table 3-3 summarizes the NMO velocity obtained from various earth models.

After making the small-spread and small-dip approximations, moveout is hyperbolic for all cases and is given by

$$t^2(x) = t^2(0) + x^2/v_{\mathrm{NMO}}^2. \tag{3.14}$$

The hyperbolic moveout velocity should be distinguished from the stacking velocity that optimally allows stacking of traces in a CMP gather. The hyperbolic form is used to define the best stacking path:

$$t_{\mathrm{st}}^2(x) = t_{\mathrm{st}}^2(0) + x^2/v_{\mathrm{st}}^2, \tag{3.15}$$

where v_{st} is the velocity that allows the best fit of the traveltime curve $t_{\mathrm{st}}(x)$ on a CMP gather to a hyperbola within the spread length. This hyperbola is not necessarily the small-spread hyperbola implied by equation (3.14).

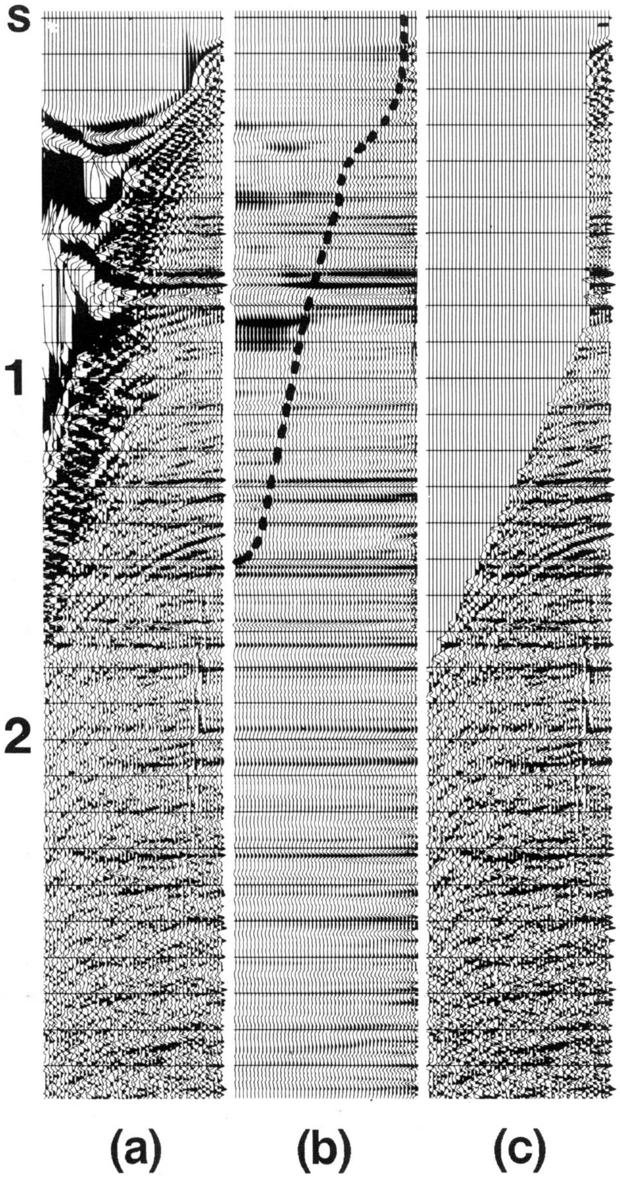


FIG. 3-15. Optimum mute selection. Starting with the NMO-corrected CMP gather in panel (a), a substack gather is obtained, (b). The rightmost trace in this gather is the same as that in the original gather, (a). The second trace from the right is the stack of the two near traces of the original gather. Finally, the leftmost trace is the full-fold stack obtained from the original gather. The area above the dotted line in (b) is the mute zone. Panel (c) is the result of a poor mute choice based on picks from the original gather (a). Compare (b) and (c). Refer to (a) to note a preference between the mutes in (b) and (c).

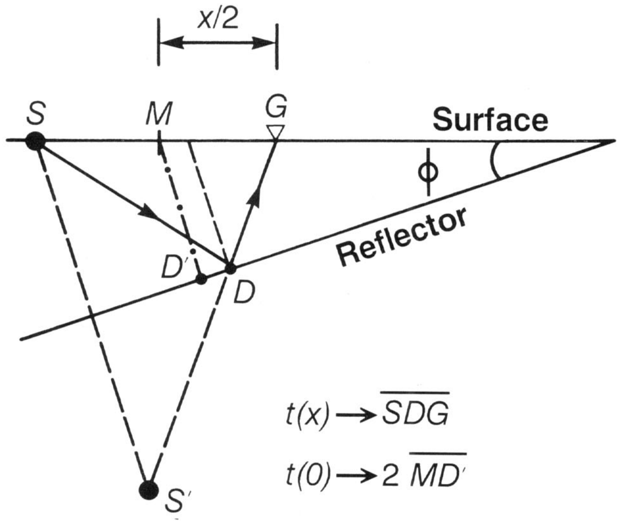


FIG. 3-16. Geometry for NMO of a single dipping reflector [refer to equation (3.7)].

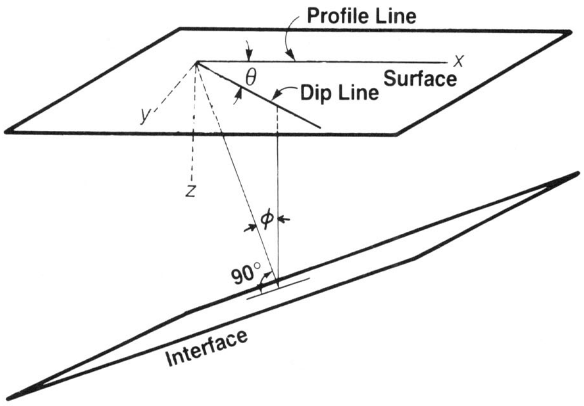


FIG. 3-17. Geometry for a dipping planar interface used in deriving the 3-D moveout velocity [refer to equation (3.9)], where ϕ = dip angle, θ = azimuth angle. (Adapted from Levin, 1971.)

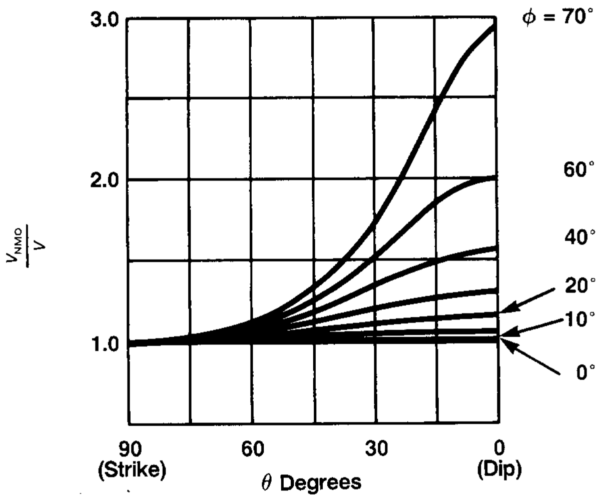


FIG. 3-18. Graphic representation of the 3-D moveout equation derived from the geometry of Figure 3-17, where ϕ = dip angle and θ = azimuth angle. Moveout velocity is identical to medium velocity if the shooting direction is in the strike-line direction (θ = 90 degrees). The largest difference between the moveout velocity and the medium velocity occurs along the dip line direction (θ = 0 degree) at large dip angles. (Adapted from Levin, 1971.)

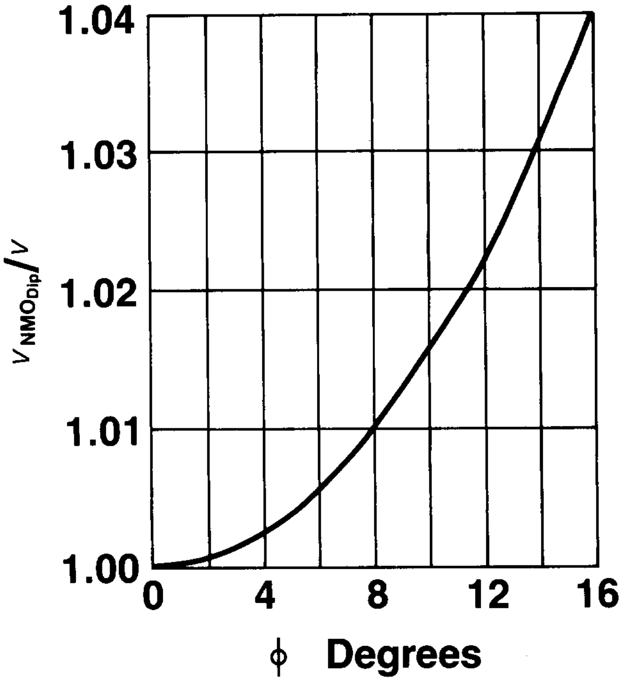


FIG. 3-19. Graphic representation of the 3-D moveout equation derived from Figure 3-17 for zero azimuth angle (θ = 0 degree, dip line) and small dip angles (ϕ = 0 to 15 degrees). (Adapted from Levin, 1971.)

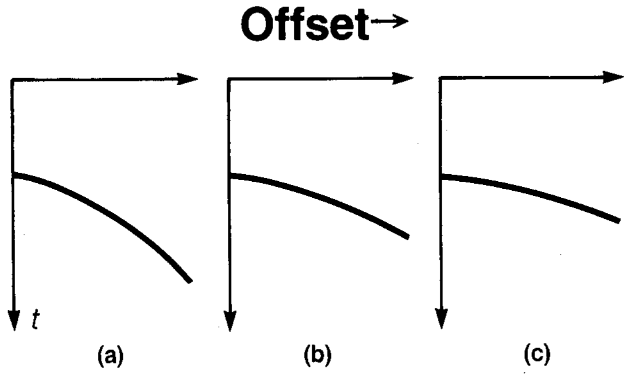


FIG. 3-20. Moveout for low-velocity event (a) is larger than for high-velocity event (b). Moveout for low-velocity dipping event (c) may not be distinguishable from high-velocity horizontal event (b). These observations are direct consequences of equation (3.7).

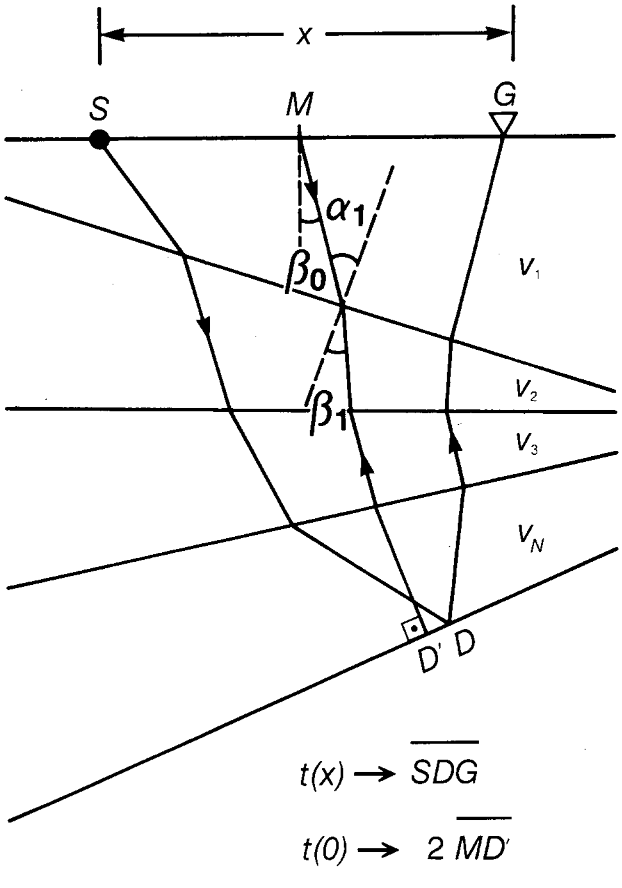


FIG. 3-21. Geometry for the moveout for a dipping interface in an earth model with layers of arbitrary dips on top [refer to equation (3.12)]. (Adapted from Hubral and Krey, 1980.)

Table 3-3. NMO velocity for various earth models.

Model	NMO Velocity
Single Horizontal Layer	Velocity of the medium above the reflecting interface.
Horizontally Stratified Earth	The rms velocity function, provided spread is small.
Single Dipping Layer	Medium velocity divided by cosine of the dip angle.
Multilayered Earth with Arbitrary Dips	The rms velocity function, provided the spread is small and the dips are gentle.

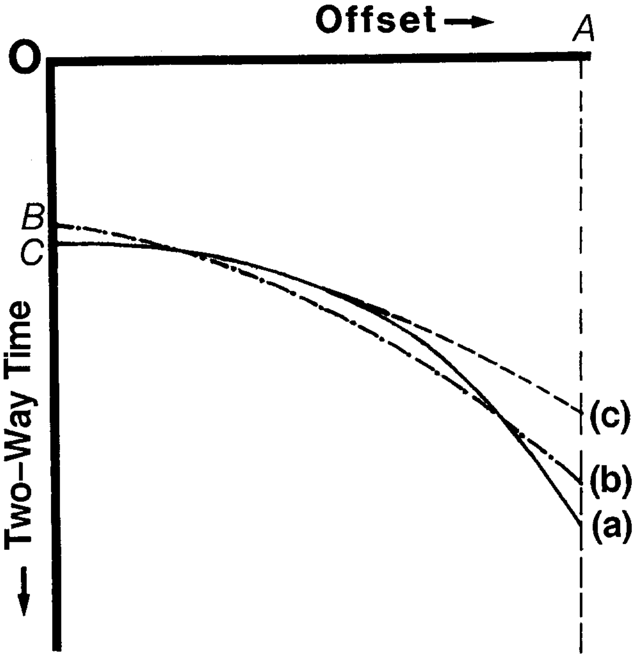


FIG. 3-22. The equation for moveout velocity is derived by assuming a small-spread hyperbola [equation (3.14)]. On the other hand, stacking velocity is derived from the best-fit hyperbola over the entire spread length [equation (3.15)]. Here, (a) is the actual traveltime, (b) is best-fit hyperbola over the offset range *OA*, and (c) is small-spread hyperbola. (Adapted from Hubral and Krey, 1980.)

Table 3-4. Estimated stacking and actual rms velocities for the synthetic model in Figure 3-23.

$t(0)$, s	Estimated Stacking Velocities $t^2 - x^2$, m/s	Actual rms Velocities, m/s
0.4	2000	2000
0.8	2264	2264
1.2	2519	2533
1.6	2828	2806

The difference becomes significant beyond a certain maximum offset value. This is illustrated in Figure 3-22. Referring to Figure 3-22, note that the observed two-way zero-offset time $OC = t(0)$ in equation (3.14) can be different from the two-way zero-offset time $OB = t_{\rm st}(0)$ associated with the best-fit hyperbola [equation (3.15)]. This occurs, for example, if some heterogeneity exists in the velocity layers above a reflector under consideration. The difference between the stacking velocity $v_{\rm st}$ and NMO velocity $v_{\rm NMO}$ is called *spread-length bias* (Al-Chalabi, 1973; Hubral and Krey, 1980). From equations (3.14) and (3.15), the smaller the spread length, the closer the optimum stacking hyperbola to the small-spread hyperbola, hence the smaller the difference between $v_{\rm st}$ and $v_{\rm NMO}$.

3.3 VELOCITY ANALYSIS

Normal moveout is the basis for determining velocities from seismic data. Computed velocities can in turn be used to correct for NMO so that reflections are aligned in the traces of a CMP gather before stacking. From equation (3.15), we can develop a practical way to determine stacking velocity from a CMP gather. Equation (3.15) describes a line on the $t^2(x)$ versus x^2 plane. The slope of the line is $1/v_{\rm st}^2$ and the intercept value $x = 0$ is $t(0)$. The synthetic gather in Figure 3-23 was derived from the velocity model in Figure 3-11. The rightmost frame of Figure 3-23 shows the picked traveltimes of four events at a number of offsets plotted on the (t^2, x^2) plane. To find the stacking velocity for a given event, the points corresponding to that event have been connected by a straight line. The inverse of the slope of the line is the square of the stacking velocity. (In practice, least-squares fitting can be used to define the line slopes.) A comparison between the computed stacking velocities and the actual rms velocities is made in Table 3-4.

The $t^2 - x^2$ velocity analysis is a reliable way to estimate stacking velocities. The accuracy of the method depends on the S/N ratio, which affects the quality of picking. In Figure 3-23, results are compared with the velocity spectrum (center frame) approach, which is discussed later in the section. A real data example is shown in Figure 3-24. Velocities estimated from the t^2-x^2 analysis are shown by triangles on the velocity spectrum. Note that agreement between the t^2-x^2 approach and the picks from the velocity spectrum are satisfactory.

Claerbout (1978) proposed a way to determine interval velocities manually from CMP gathers. The basic idea is illustrated in Figure 3-25. First, measure the slope along a

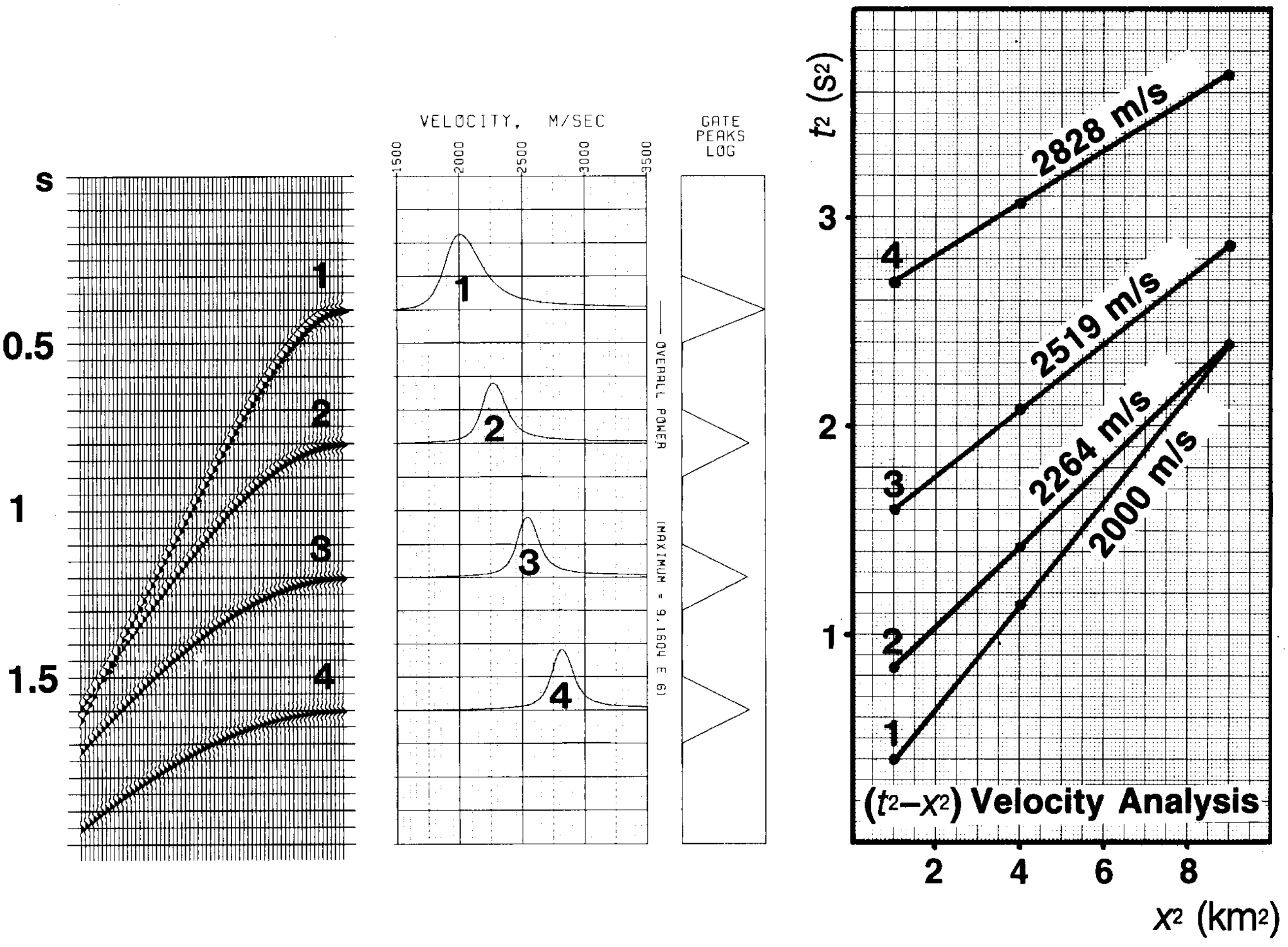


FIG. 3-23. The ($t^2 - x^2$) velocity analysis applied to the synthetic gather derived from the velocity function depicted in Figure 3-11. The center panel is the velocity spectrum based on equation (3.19).

slanted path that is tangential to both the top and bottom reflections of the interval of interest (slope 1). Then, connect the two tangential points and measure the slope of this line (slope 2). The interval velocity then is equal to the square root of the product of the two slope values. The accuracy of this method primarily depends on the S/N ratio.

The method of constant velocity scans of a CMP gather is an alternative technique for velocity analysis. Figure 3-26 is a CMP gather that has been repeatedly NMO corrected using a range of constant velocity values between 5000 and 13,600 ft/s. The NMO-corrected gathers are displayed beside each other in the form of a panel. Follow the NMO for event A. Note that this event is overcorrected at small velocities, and undercorrected at high velocities. The event is flat on the NMO-corrected gather that corresponds to the 8300 ft/s velocity; thus, 8300 ft/s is the stacking velocity for event A. Event B is flat on the NMO-corrected gather that corresponds to the 8900 ft/s velocity. By proceeding in this way, we can build up a velocity function that is appropriate for the NMO correction of this gather.

The most important reason to obtain a reliable velocity function is to get the best quality stack of signal. Therefore, stacking velocities often are estimated from data stacked with a range of constant velocities on the basis of stacked event amplitude and continuity. Figure 3-27 illustrates this approach. Here, a portion of a line containing 24 CMP gathers (typical values range from 24 to 48 CMP gathers, but may include the entire line) has been NMO-corrected and stacked with a range of constant velocity values. The resulting 24-trace CMP stacks then were displayed as a panel, where stacking velocity increases from right to left. Stacking velocities are picked directly from the constant-velocity stack (CVS) panel by choosing the velocity that yields the best stack response at a selected event time. Note the coherent noise that is present on the data. The deep event at 3.6 s seems to stack at a wide range of velocity values. This demonstrates the decrease in resolution of velocity estimates that occurs for deeper events. The problem stems from the fact that moveout generally decreases in depth.

The constant velocities used in the CVS method described above should be chosen with care. There are two issues to consider besides the expected range of actual velocities in the subsurface: (1) The range of velocities needed to stack the data and (2) the spacing between trial stacking velocities. In choosing a range, consideration should be given to the fact that dipping events and useful out-of-plane reflections may have anomalously high stacking velocities. In choosing the spacing of constant velocities, keep in mind that it is moveout, not velocity, that is the basis for velocity estimation. Thus, it is better to scan in increments of equal Δt_{NMO} than equal v_{NMO}. This prevents oversampling of the high-velocity events and undersampling of the low-velocity events. A good way to choose $\Delta(\Delta t_{\mathrm{NMO}})$ is to pick it so that the moveout difference between adjacent trial velocities at the maximum offset to be stacked is approximately $\frac{1}{3}$ of the dominant period of the data (Doherty, 1986, Personal Communication). Shallow data have short maximum offsets because of muting, while deep data have large dominant periods. Thus, the number of trial stacking velocities needed to adequately sample the data can be reduced by a fair amount.

The CVS method is especially useful in areas with complex structure (see Exercise 3.7). In such areas, it allows the interpreter to directly choose the stack with the best possible event continuity. (Often the stacking velocities themselves are of minimal importance.) Constant velocity stacks often contain many CMP traces and sometimes consist of an entire line.

The velocity spectrum method is described in the next section. Unlike the CVS method, it is based on the crosscorrelation of the traces in a CMP gather, and not on lateral continuity of the stacked events. Because of this,

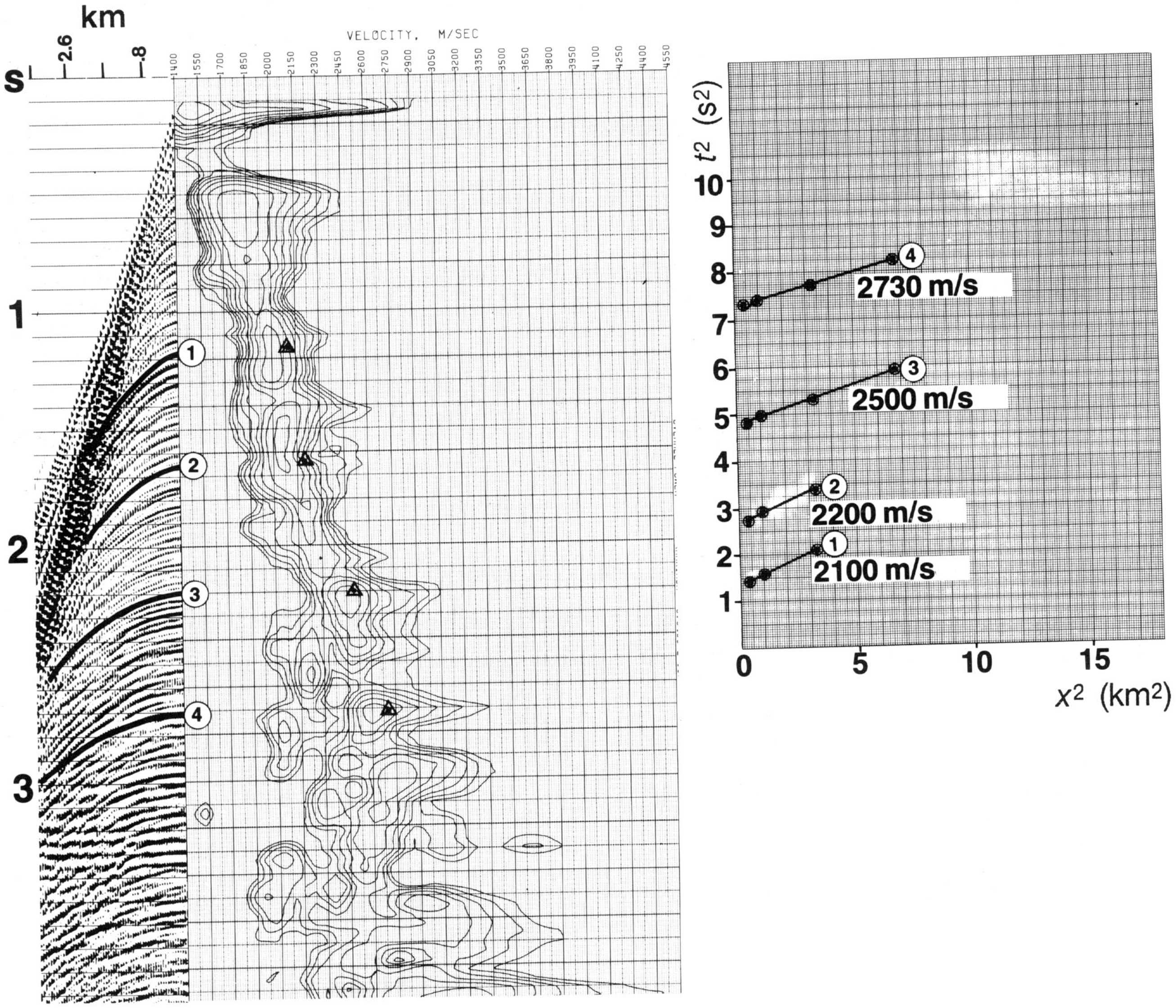


FIG. 3-24. The $(t^2 - x^2)$ velocity analysis applied to a CMP gather. The triangles on the velocity spectrum [center panel, based on equation (3.19)] represent velocity values derived from the slopes of the lines shown on the graph at the right.

when compared to the CVS method, it is more suitable for data with a multiple reflections problem and somewhat less suitable for a complex structure problem.

3.3.1 The Velocity Spectrum

The input CMP gather in Figure 3-28a contains a single reflection hyperbola from a flat interface. The medium velocity above the reflector is 3000 m/s. Suppose that this gather repeatedly is NMO-corrected and stacked using a range of constant velocity values from 2000 to 4300 m/s. Figure 3-28b displays the resultant stack traces for each velocity side by side on a plane of velocity versus two-way zero-offset time. This is called the velocity spectrum (Taner and Koehler, 1969). We have transformed the data from the offset versus two-way time domain (Figure 3-28a) to the stacking velocity versus two-way zero-offset time domain (Figure 3-28b).

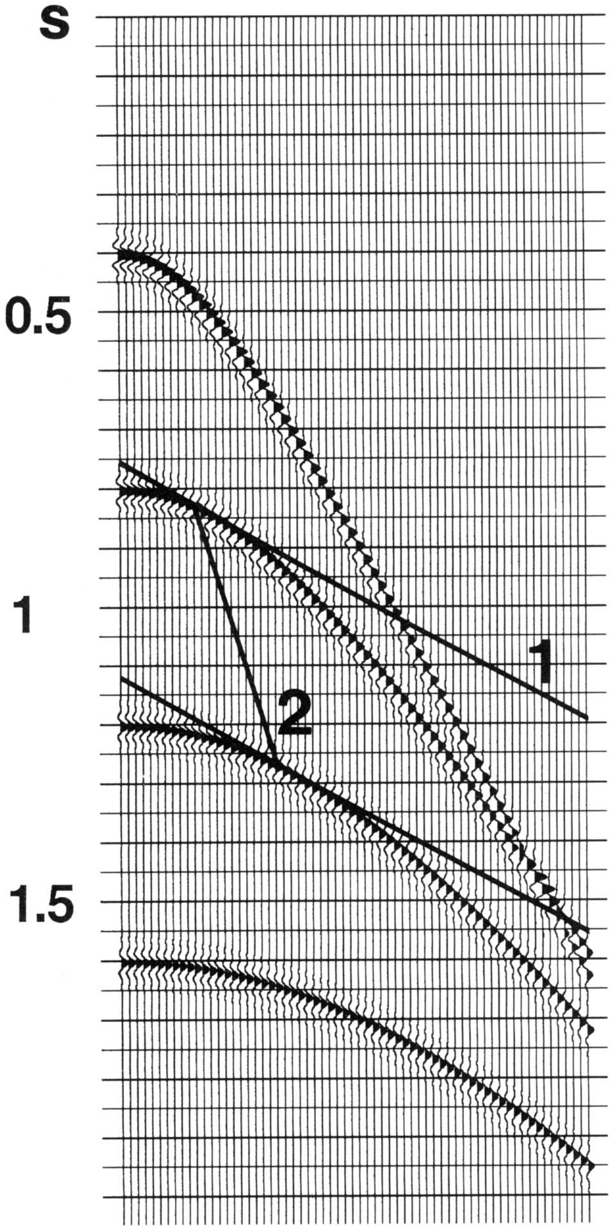


FIG. 3-25. The interval velocity between two reflectors is equal to the square-root of the products of the slope values measured as shown above. This is the same gather as in Figure 3-10a. Trace spacing is 50 m, slope 1 = 3150/0.43, slope 2 = 550/0.44, and, thus, the interval velocity between 0.8 and 1.2 s is 3026 m/s.

The highest stacked amplitude occurs with a velocity of 3000 m/s. This is the velocity that should be used to stack the event in the input CMP gather. The low-amplitude horizontal streak on the velocity spectrum is due to the contribution of small offsets, while the large-amplitude region on the spectrum is due to the contribution of the full range of offsets (Sherwood and Poe, 1972). Hence, we need long offsets for good resolution on the velocity spectrum.

A CMP gather associated with a layered earth model is shown in Figure 3-29a. Based on the stacked amplitudes, the following picks for stacking velocity function are made from the velocity spectrum (Figure 3-29b): 2700, 2800, and 3000 m/s. These picks correspond to the shallow, middle, and deep events, respectively. The velocity spectrum not only can provide the stacking velocity function, but also allows one to distinguish between primary and multiple reflections (Section 8.2).

The quantity displayed on the velocity spectra in Figures 3-28b and 3-29b is the stacked amplitude. When the S/N ratio of the input data is poor, then the stacked amplitude may not be the best display quantity. The aim in velocity analysis is to obtain picks that correspond to the best *coherency* of the signal along a hyperbolic trajectory over the entire spread length of the CMP gather. Neidell and Taner (1971) summarized various types of coherency measures that can be used as attributes in computing velocity spectra.

Consider the CMP gather with a single reflection in Figure 3-30. *Stacked amplitude* is defined as:

$$s_t = \sum_{i=1}^{M} f_{i,\,t(i)}, \tag{3.16}$$

where $f_{i,t(i)}$ is the amplitude value on the ith trace at two-way time *t(i)*. Here, *M* is the number of traces in the CMP gather. Two-way time *t(i)* lies along the trial stacking hyperbola:

$$t(i) = [t^2(0) + x_i^2/v_{\mathrm{st}}^2]^{1/2}. \tag{3.17}$$

Normalized stacked amplitude is defined as

$$NS = \frac{|s_t|}{\sum_i |f_{i,\,t(i)}|}. \tag{3.18}$$

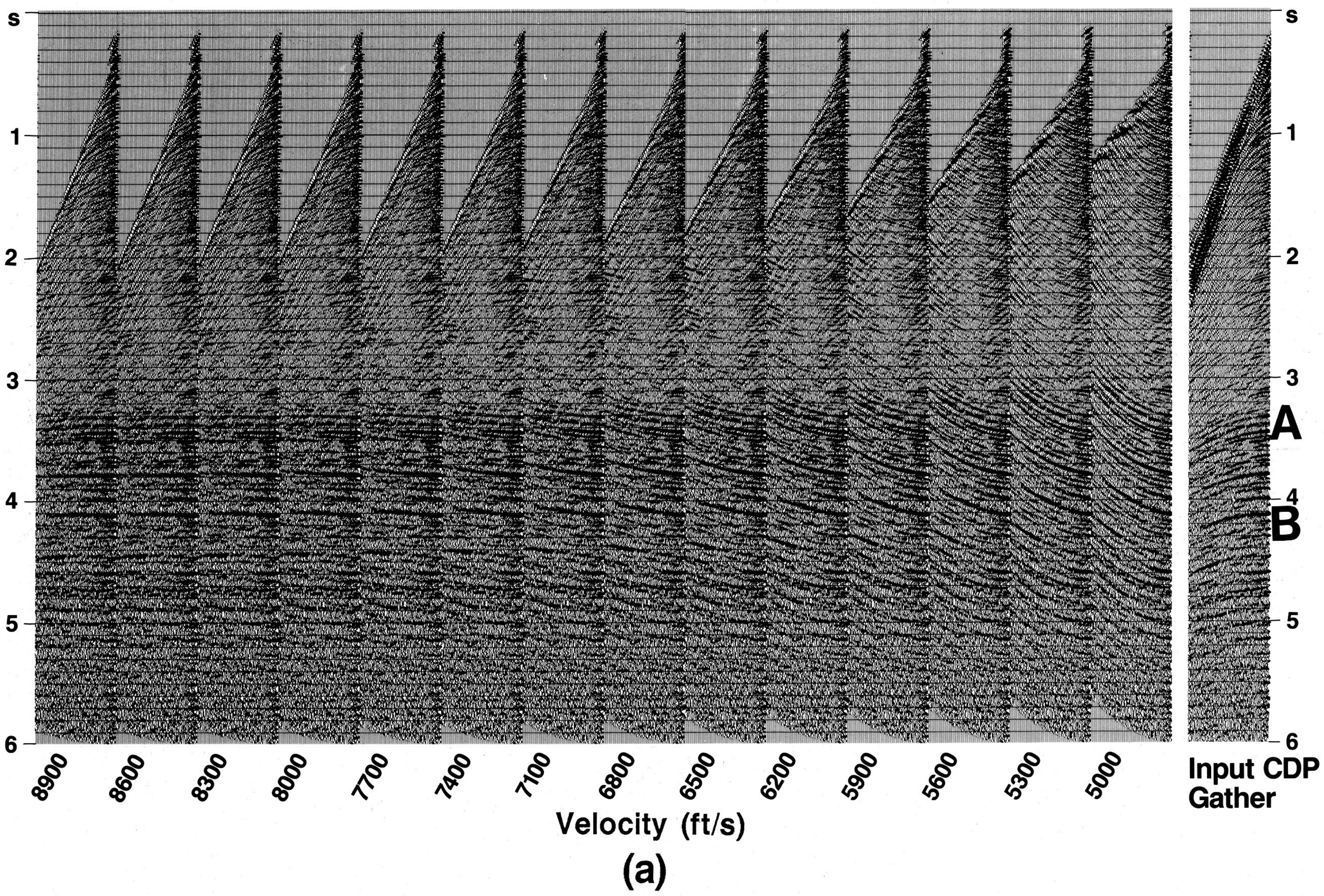


FIG. 3-26a. Constant-velocity moveout corrections applied to a CMP gather (5000 to 8900 ft/s).

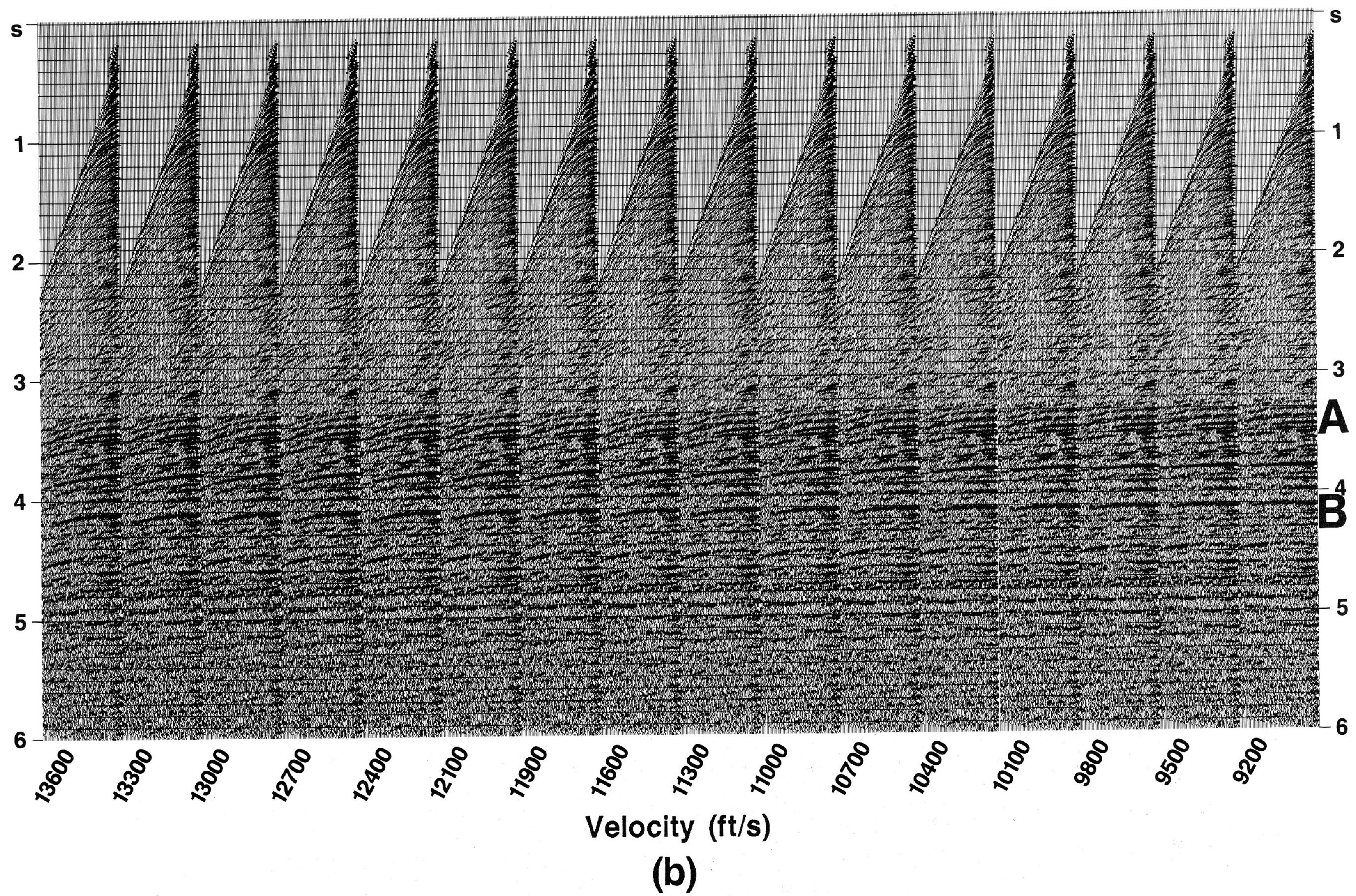


FIG. 3-26b. Constant-velocity moveout corrections applied to the CMP gather in Figure 3-26b (9200 to 13,600 ft/s).

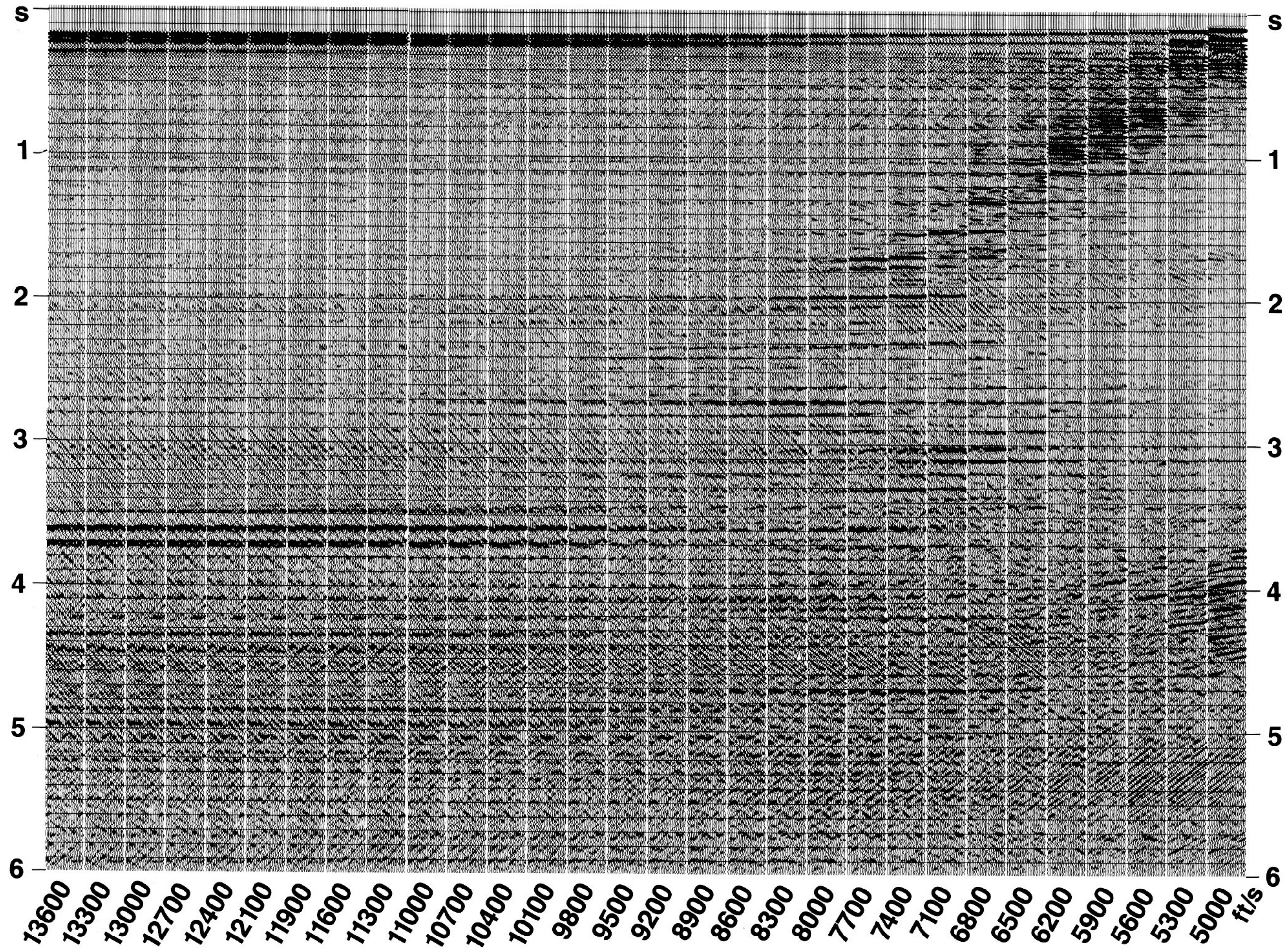


FIG. 3-27. Constant-velocity stacks of 24 CMP gathers (5000 to 13,600 ft/s).

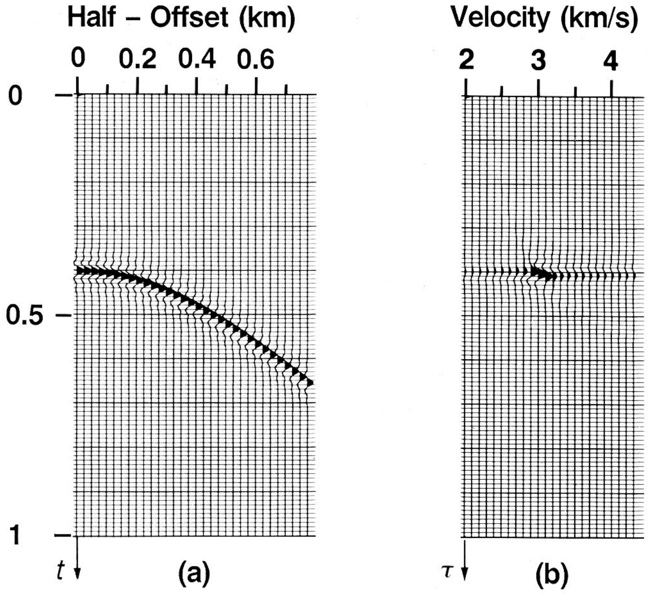


FIG. 3-28. Mapping the offset axis to the velocity axis. Each trace in the $[v, \tau = t(0)]$ gather (b) is a stack of the traces in the CMP gather (a) using a constant velocity NMO correction.

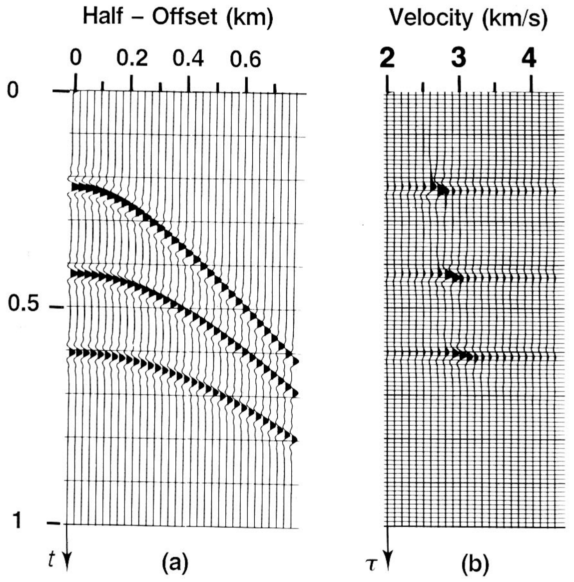


FIG. 3-29. Mapping the offset axis to the velocity axis. Each trace in the $[v, \tau = t(0)]$ gather (b) is a stack of the traces in the CMP gather (a) using a constant velocity NMO correction.

The range of NS is $0 \leq NS \leq 1$. Equation (3.18) implies that coherency as defined here is the normalized stacked amplitude.

Another quantity that is used in velocity spectrum calculations is the *unnormalized crosscorrelation* sum within a time gate that follows the path corresponding to the trial stacking hyperbola across the CMP gather. The expression for the unnormalized crosscorrelation sum is given by:

$$CC = \frac{1}{2} \sum_{t} \left\{ \left[\sum_{i=1}^{M} f_{i,\,t(i)} \right]^2 - \sum_{i=1}^{M} f_{i,\,t(i)}^2 \right\}$$

$$= \frac{1}{2} \sum_{t} \left[s_t^2 - \sum_{i} f_{i,\,t(i)}^2 \right], \qquad (3.19)$$

where CC can be interpreted as half the difference between the output energy of the stack and the input energy. A *normalized* form of CC is another attribute that often is used in velocity spectrum calculations and is given by:

$$NCC = \frac{2}{M(M-1)} \sum_{t} \sum_{k=1}^{M-1} \sum_{i=1}^{M-k} \frac{f_{i,\,t(i)} f_{i+k,\,t(i+k)}}{\left[\sum_{t} f_{i,\,t(i)}^2 \sum_{t} f_{i+k,\,t(i+k)}^2 \right]^{1/2}}. \qquad (3.20)$$

The *energy-normalized crosscorrelation* sum is given by:

$$ECC = \frac{2}{(M-1)} \frac{CC}{\sum_{t} \sum_{i=1}^{M} f_{i,\,t(i)}^2}. \qquad (3.21)$$

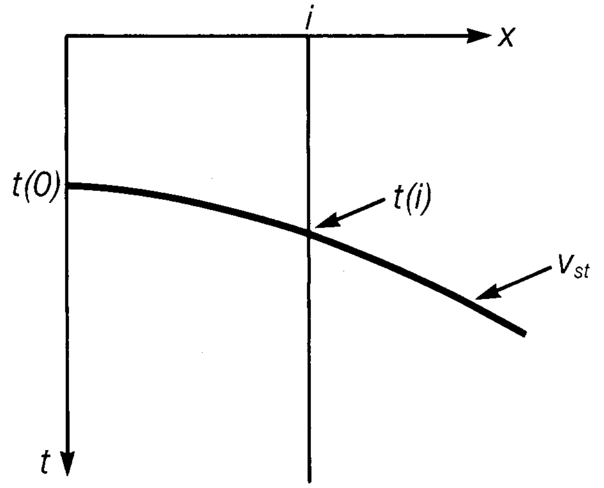


FIG. 3-30. Stacked amplitude. Amplitudes $f_{i,t(i)}$ along the best-fit hyperbola [equation (3.17)] defined by optimum stacking velocity v_{ST} are summed to get the stacked amplitude s_t [equation (3.16)].

The range of ECC is $-[1/(M\text{-}1)] < ECC \leq 1$. Finally, *semblance*, which is the normalized output-to-input energy ratio, is given by:

$$NE = \frac{1}{M} \frac{\sum_{t} s_t^2}{\sum_{t} \sum_{i=1}^{M} f_{i,\,t(i)}^2}. \qquad (3.22)$$

The following expression shows the relation of NE to ECC:

$$ECC = \frac{1}{M-1} (M \cdot NE - 1). \qquad (3.23)$$

The range of NE is $0 \leq NE \leq 1$.

Table 3-5 shows the values of the attributes defined by equation (3.16) and equations (3.18) through (3.22) for the special case of a two-fold CMP gather where the second trace is a scaled version of the first as indicated by equation (3.24):

$$\left. \begin{aligned} f_{1,t} &= f_t \\ f_{2,t} &\doteq af_t \end{aligned} \right\} \qquad (3.24)$$

Several conclusions can be made from the results in Table 3-5. Note that stacked amplitude is sensitive to trace polarity. The unnormalized crosscorrelation offers better standout of strong reflections on the velocity spectrum, while the normalized or energy-normalized crosscorrelation brings out weak reflections on the velocity spectrum. As equation (3.23) implies, semblance is no more than a biased version of the energy-normalized crosscorrelation sum.

The velocity spectrum normally is not displayed as shown in Figures 3-28b or 3-29b. Instead, two popular types of displays are used to pick velocities in the form of a *gated row plot* or a *contour plot* as shown in Figure 3-31. Many prefer the contour plot, although the gated row plot also is used extensively. Another quantity that helps picking is the maxima of the coherency values from each time gate displayed as a function of time next to the velocity spectrum, as shown in Figure 3-31. Unless otherwise indicated, the unnormalized correlation was used to construct the velocity spectrum of the synthetic CMP gather that is used in subsequent discussions.

3.3.2 Factors Affecting Velocity Estimates

Velocity estimation from seismic data is limited in accuracy and resolution for the following reasons:

1. Spread length
2. Stacking fold
3. S/N ratio
4. Muting
5. Time gate length
6. Velocity sampling
7. Choice of coherency measure
8. True departures from hyperbolic moveout
9. Bandwidth of data.

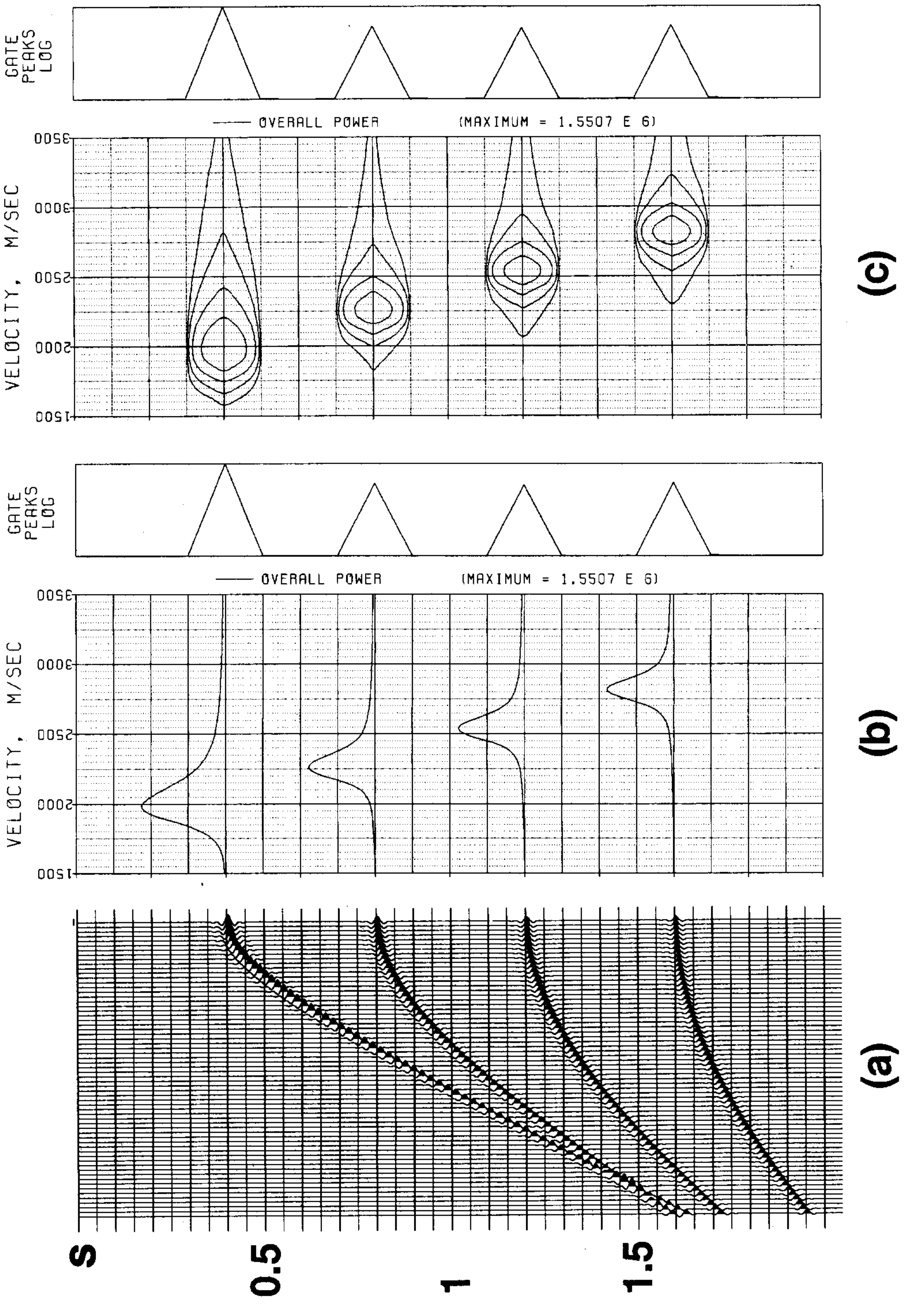


FIG. 3-31. Two ways of displaying velocity spectrum derived from the CMP gather (a), (b) gated raw plot, and (c) contour plot.

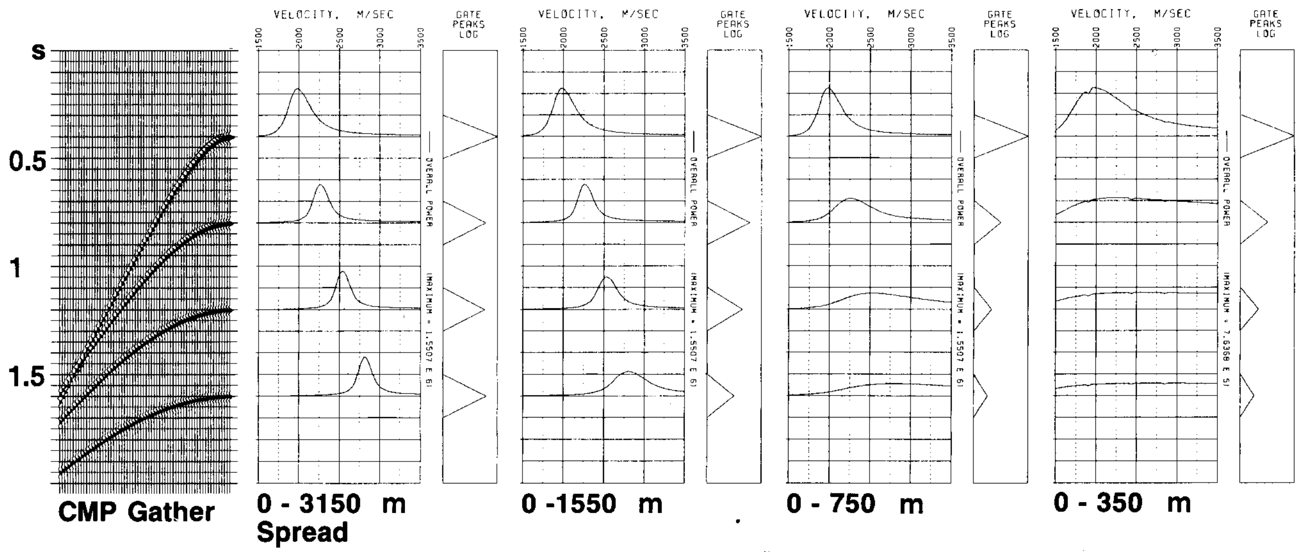


FIG. 3-32. Effect of spread length on velocity resolution. Lack of long offsets causes loss of resolution, especially at later times.

Table 3-5. Various measures of coherency as applied to the two-fold case given by equation (3.24).

Attribute	$a = 0.5$	$a = -0.5$
Stacked Amplitude [equation (3.16)]	$1.5\,f(t)$	$0.5\,f(t)$
Coherency, *NS* [equation (3.18)]	1	0.333
Unnormalized Crosscorrelation Sum, *CC* [equation (3.19)]	$0.5\sum_t f^2(t)$	$-0.5\sum_t f^2(t)$
Normalized Crosscorrelation Sum, *NCC* [equation (3.20)]	1	1
Energy-Normalized Crosscorrelation Sum, *ECC* [equation (3.21)]	0.8	−0.8
Semblance, *NE* [equation (3.22)]	0.9	0.1

Figure 3-32 is a synthetic CMP gather with velocity spectra generated by using gradually decreasing spread lengths. Lack of large-offset information means lack of the significant moveout required for velocity discrimination. Note the loss in sharpness of the peaks in the velocity spectra computed from the small-spread portion of the CMP gather. Resolution decreases first in the deeper part of the spectrum where there is little moveout (Table 3-1). Figure 3-33 shows velocity spectra computed from a real data set using spread lengths as indicated. The broadened peaks caused by the use of smaller spreads indicate loss of resolution in the velocity spectrum. This problem may be compounded by poor S/N ratio or residual statics shifts. An example of residual statics effect is shown in Figure 3-34. As spread length is made smaller, velocity becomes indistinct.

What if only the far offsets are included when computing the velocity spectrum? Although far-offset data are needed to better resolve the velocity picks, there is a stretching problem in the far-offset region. Therefore, a velocity spectrum computed on the basis of only the far-offset region of a CMP gather suffers from the effects of muting at shallow times. This problem is demonstrated in Figure 3-35, where spread is increasingly confined to the far-offset region of the input CMP gather. Note the loss of coherency peaks from the shallow events (due to muting) and the further degradation of the coherency peaks corresponding to deeper events. The lesson is clear: adequate resolution in the velocity spectrum can only be obtained with a sufficiently large spread that spans both near and far offsets. This is analogous to the lesson learned in Section 1.2.4 on temporal resolution, which requires both low and high frequencies.

Stacking fold plays a significant part in the degree of resolution achieved from velocity spectra. In contemporary seismic data acquisition, it is common to record data with 120, 240, or more channels. For computational savings, high-fold data sometimes are reduced to a low-fold equivalent gather by partial stacking. The idea is to stack a number of traces in a CMP gather from adjacent offsets to produce a CMP gather with lower fold. For example, a reduction of fold from 64 to 16 amounts to producing one output trace for each set of four adjacent input traces. Partial stacking involves differential NMO application to each group of adjacent traces using a reasonable, previously estimated velocity function so that primaries are aligned before stacking. The CMP gather in Figure 3-36 was partially stacked down to 32-, 16-, and 8-fold gathers. Corresponding velocity spectra also are shown in Figure 3-36. No harm was done by reducing the fold to 32. Even the 16-fold data seem to produce accurate picks. However, use of lower fold significantly shifts the peaks in the spectrum. Reducing fold by partial stacking merely to save computation must not be done at the expense of accuracy.

Noise in seismic data has a direct effect on the quality of a velocity spectrum. Add band-limited random noise to the

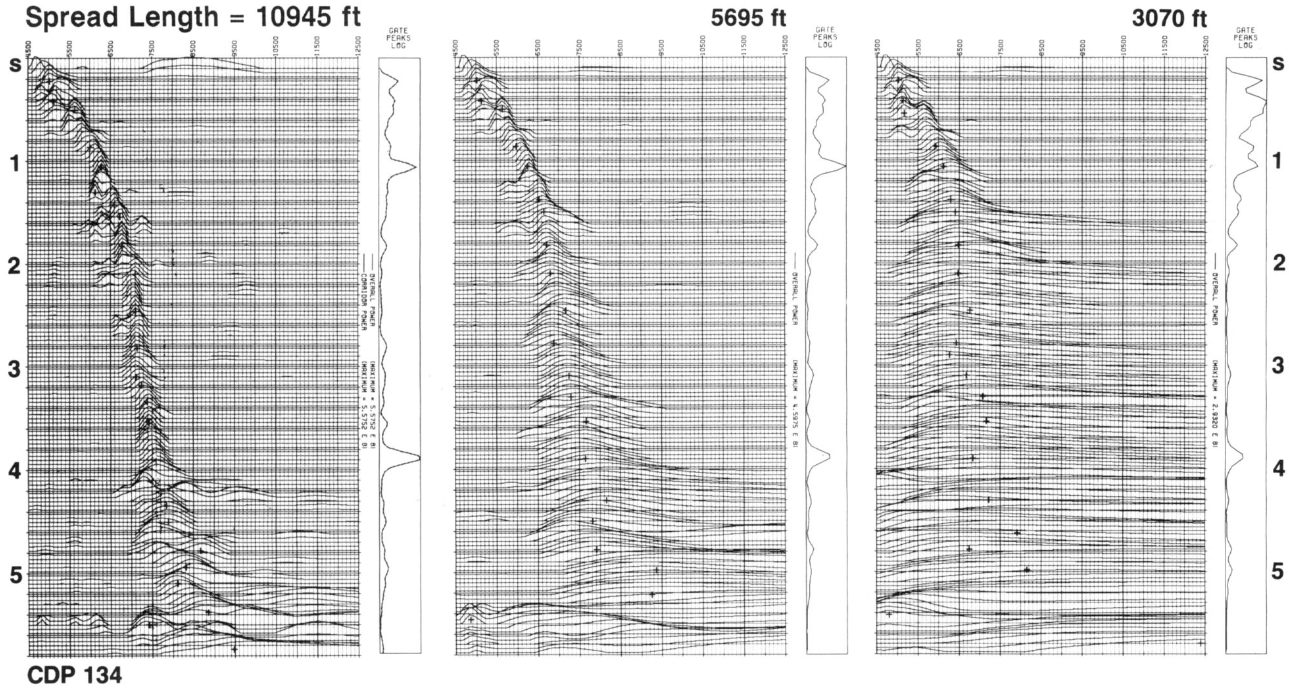


FIG. 3-33. Missing long-offset traces cause loss of resolution on the velocity spectra, especially at later times.

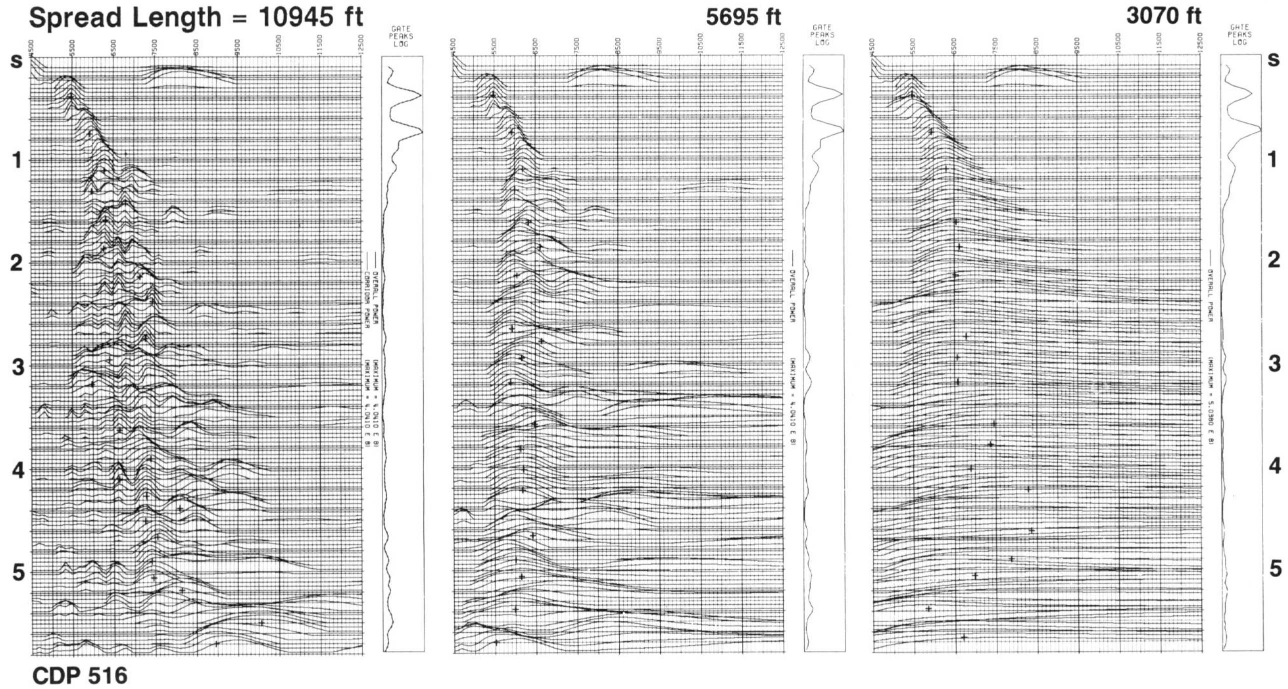


FIG. 3-34. Velocity spectra can be severely distorted, particularly in areas with statics problems. When coupled with a lack of long-offset traces, the velocity function that is derived can be misleading (center and right).

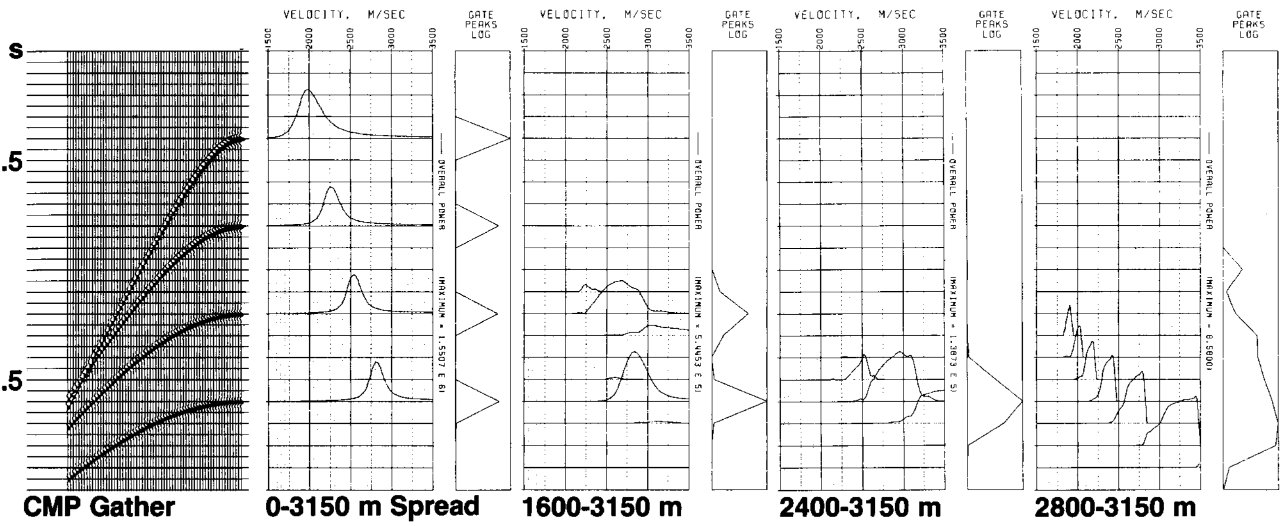


FIG. 3-35. Lack of near-offset traces can degrade the velocity spectrum. Note the loss of information at shallow times and poor picks at later times.

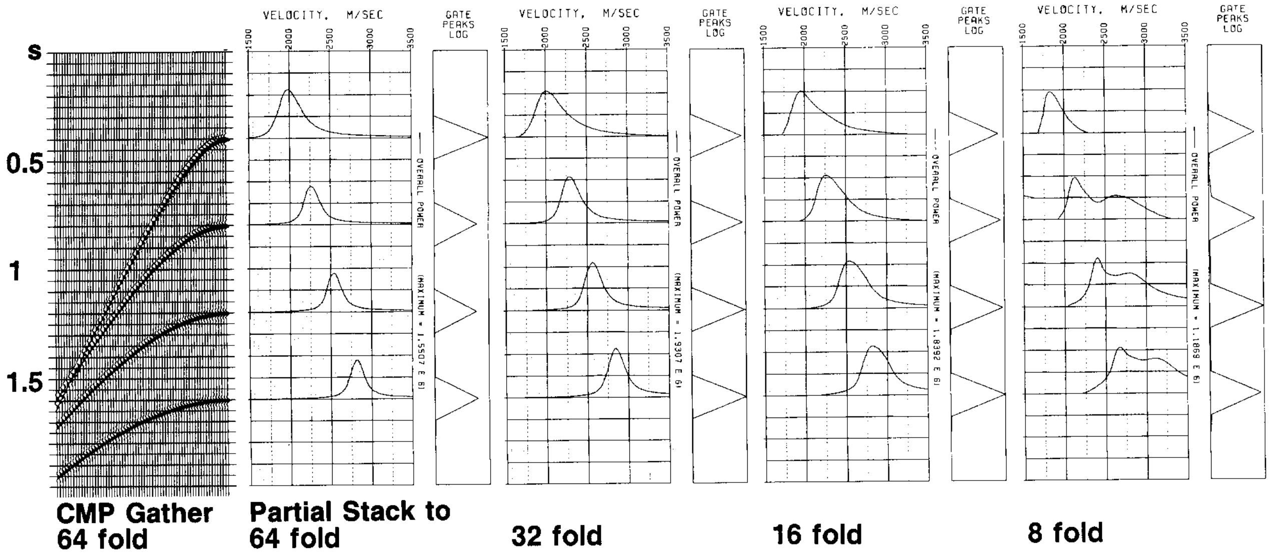


FIG. 3-36. Partial stacking can save money. However, do not use partial stacking if it could degrade the velocity spectrum. (In this example, 8-fold partial stacking is too much.)

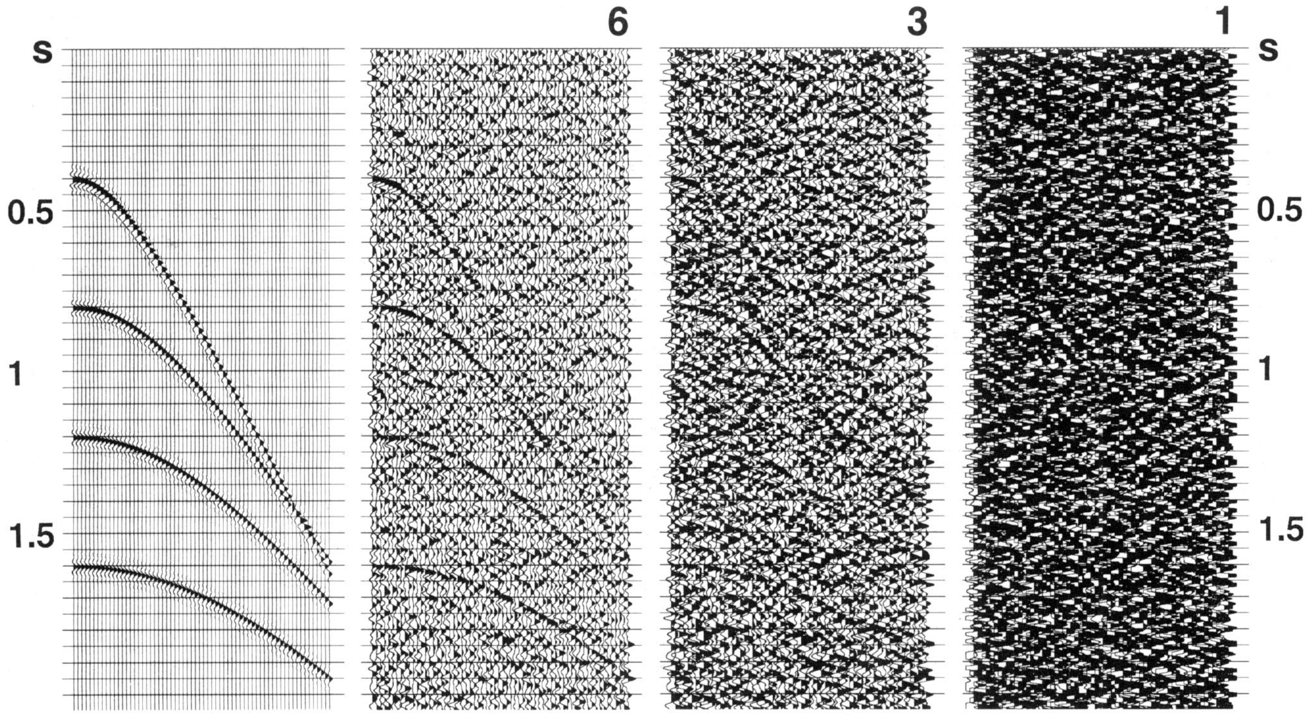


FIG. 3-37. The synthetic CMP gather derived from the velocity function depicted in Figure 3-11 and the same gather with noise added at various strengths. The numbers on top represent the ratio of peak signal amplitude to peak noise amplitude.

CMP gather at increasingly higher levels of amplitude (Figure 3-37). The corresponding velocity spectra are shown in gated row plot form in Figure 3-38 and, for comparison, in contour form in Figure 3-39. The velocity spectrum distinguishes signal along hyperbolic paths even with high levels of random noise. (Refer to the velocity spectrum for SNR = 3 in Figure 3-38.) This is because of the power of crosscorrelation in measuring coherency. The accuracy of the velocity spectrum is limited when the S/N ratio is poor. Refer to SNR = 1 in Figures 3-38 or 3-39. The event at 0.8 s still can be picked, but the others are difficult to distinguish.

As a result of moveout correction, the waveform along a reflection hyperbola is stretched (Section 3.2.2). Stretching is more severe in the shallow part of the moveout-corrected gather, especially at large offsets. The stretched zone must be muted to prevent degradation of the stacked amplitudes associated with shallow events. However, muting reduces fold in the stacking process for shallow data (Figure 3-14c). It also has an adverse effect on the velocity spectrum, for it causes weakening of the peak amplitude that falls within the mute zone, as demonstrated in Figure 3-40. These peaks must be corrected for the weakening effect of the muting process. This is done by multiplying stacked amplitudes by a scale factor equal to the ratio of the actual multiplicity to the number of live traces in the mute zone.

The velocity spectrum is computed along hyperbolic search paths for a range of constant velocity values, or constant Δt_{NMO}. The hyperbolic path spans a certain two-way time gate at zero-offset. If the gate length is chosen too small, then the computational cost increases. If it is chosen too coarsely, then the spectrum suffers especially from lack of temporal resolution. Figure 3-41 shows velocity spectra computed with four different gate lengths. For comparison, the same spectra are shown in Figure 3-42 as contour plots. In practice, gate length is chosen between one-half and one times the dominant period of the signal, typically 20 to 40 ms. Since the dominant period can be time-variant (small in early and large in late times), gate length can be specified accordingly.

The velocity range used in the analysis must be chosen carefully; it should span the velocities that correspond to those of primary reflections present on the CMP gather. The velocity increment must not be too coarse, for it can degrade the resolution, especially for high-velocity events.

Several options are considered in constructing the velocity spectrum. Partial stacking is one option that already was discussed. Subsampling (decimating in time) the data before velocity analysis is another. Band-pass filtering and automatic gain control (AGC) sometimes can improve the crosscorrelation process, especially when the input gather has poor S/N ratio. Another way to improve the quality of a velocity spectrum is to use several neighboring CMP gathers in the analysis. Figure 3-43 shows six neighboring CMP gathers. By using the first CMP gather in the group, we get the velocity spectrum in Figure 3-44a. There are two ways to analyze these

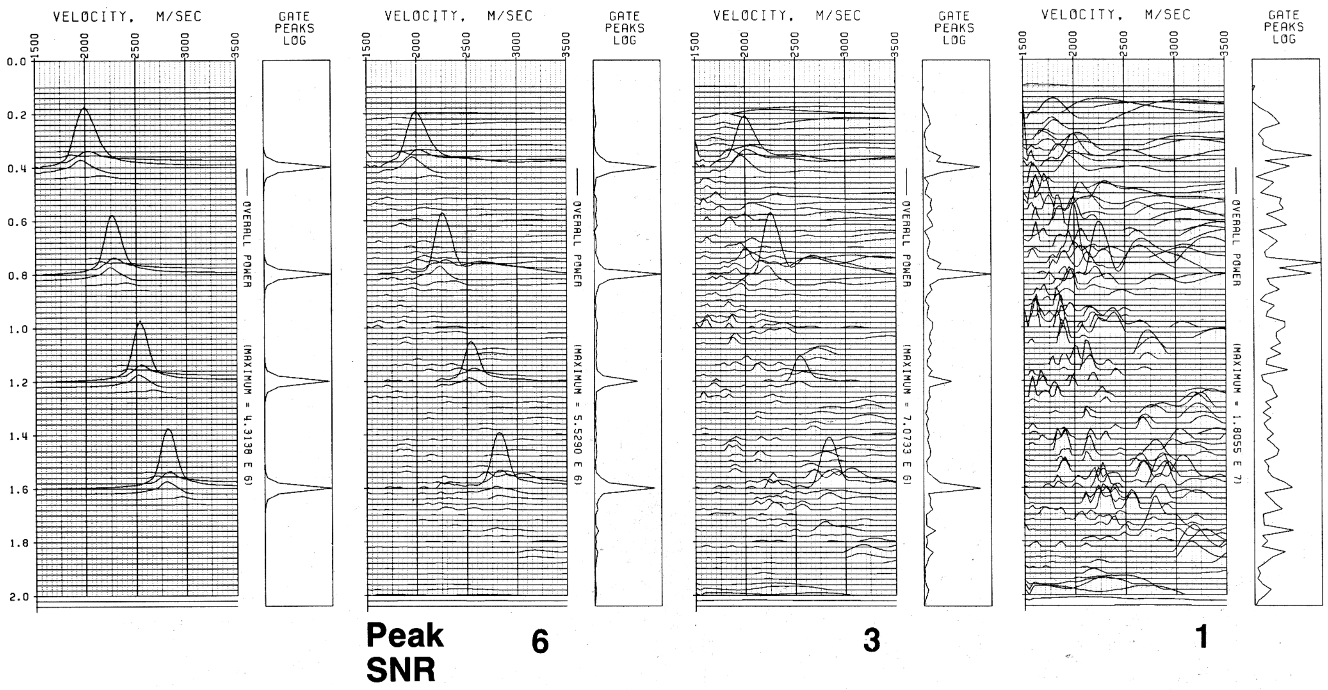


FIG. 3-38. Velocity spectra derived from the CMP gathers in Figure 3-37. The display mode is gated row.

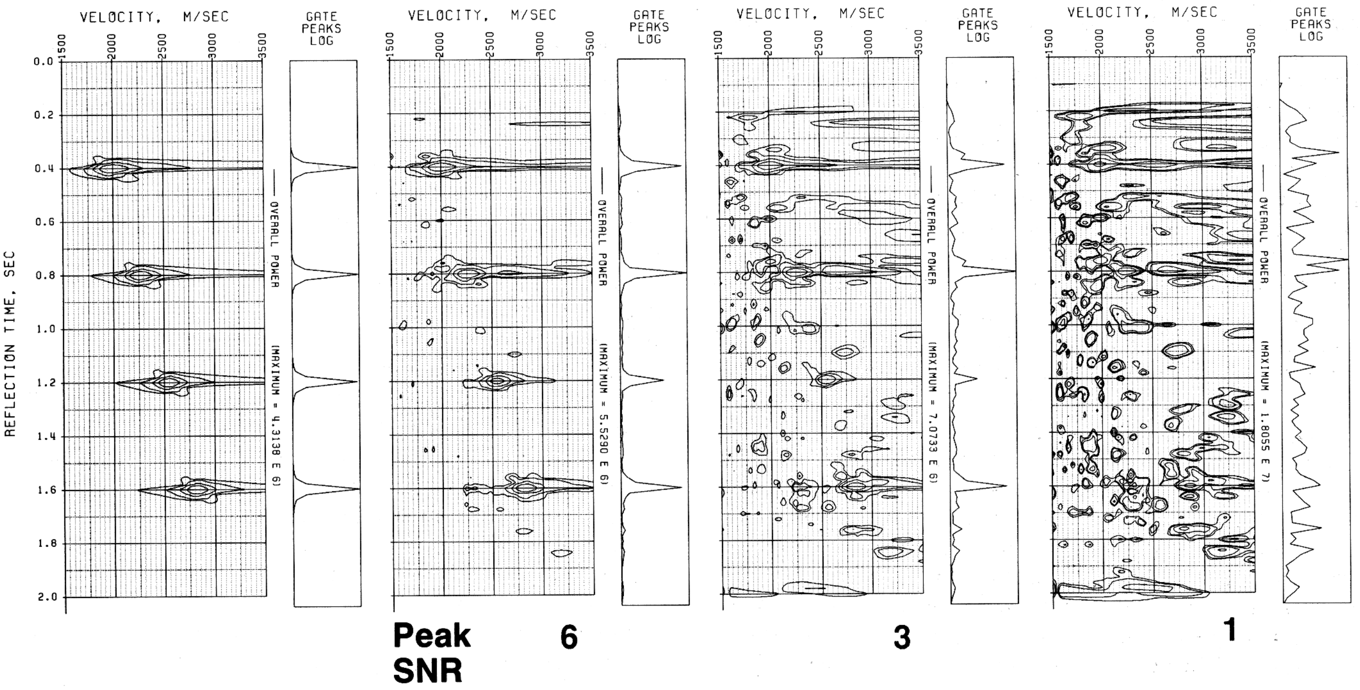


FIG. 3-39. Velocity spectra derived from the CMP gathers in Figure 3-37. The display mode is contour. See Figure 3-38 for the gated row plots of the same spectra for comparison.

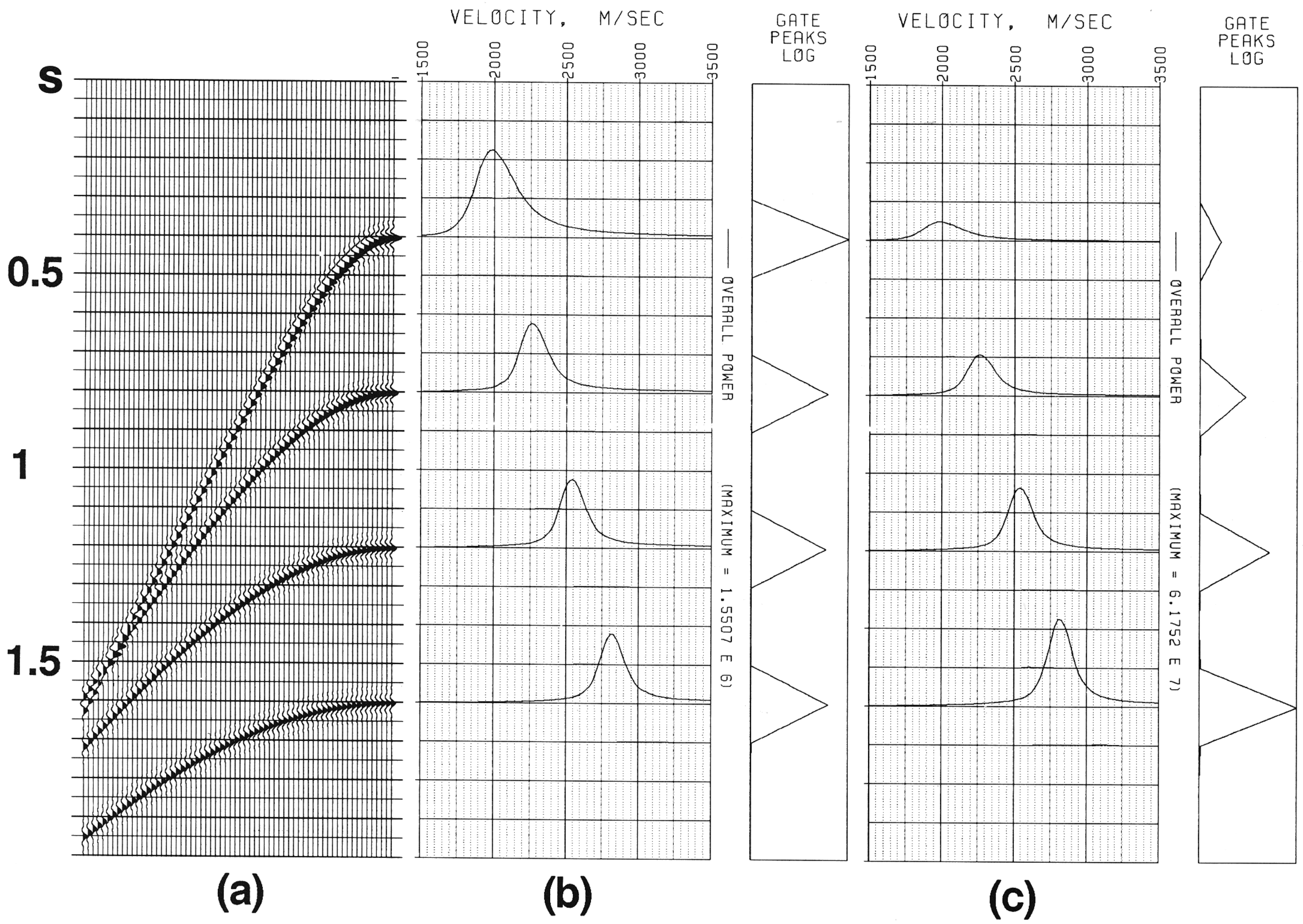


FIG. 3-40. Muting effect on correlation values: (a) CMP gather, (b) mute compensated, (c) no compensation.

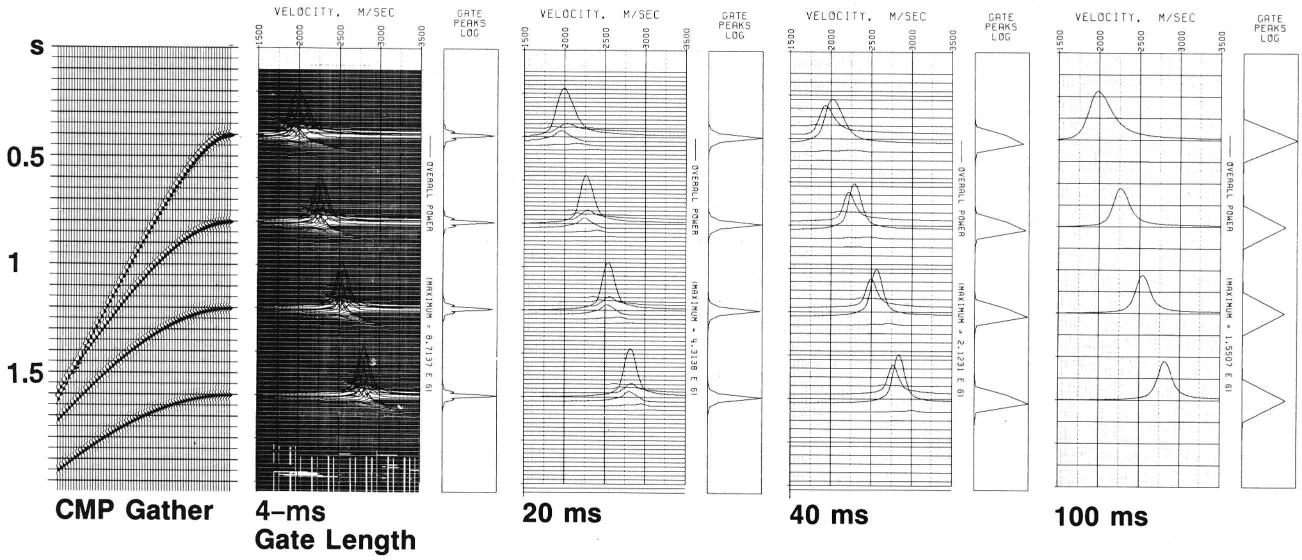


FIG. 3-41. Too small a correlation gate length increases the computational cost, while too large a value can lower resolution.

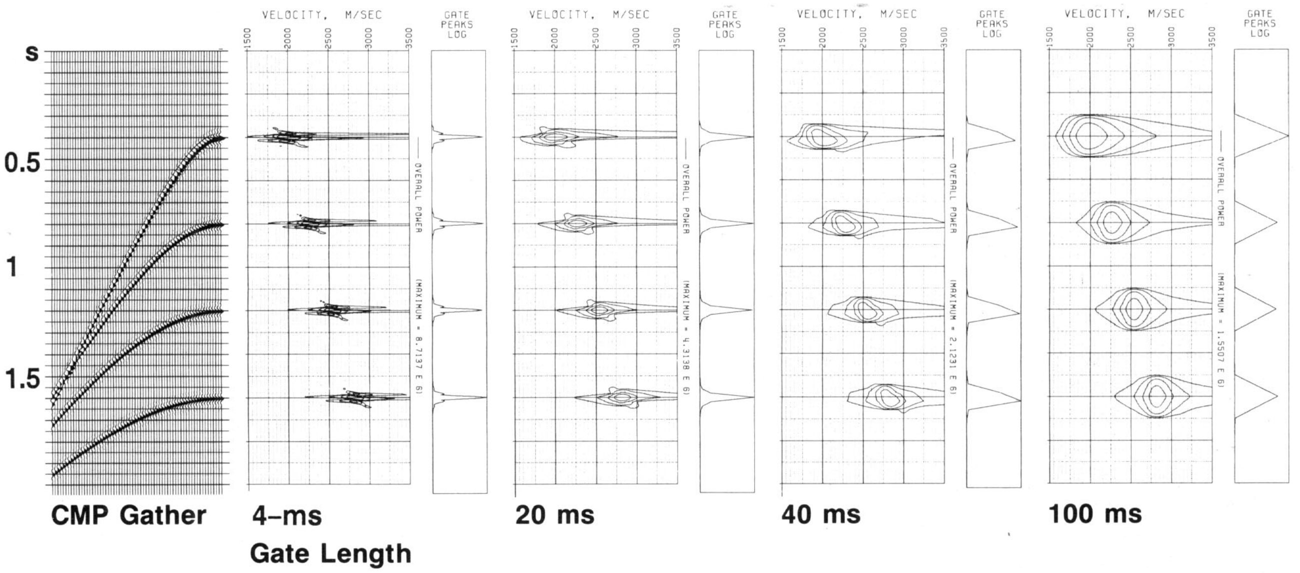


FIG. 3-42. The same velocity spectra as in Figure 3-41 displayed using the contour display mode for comparison.

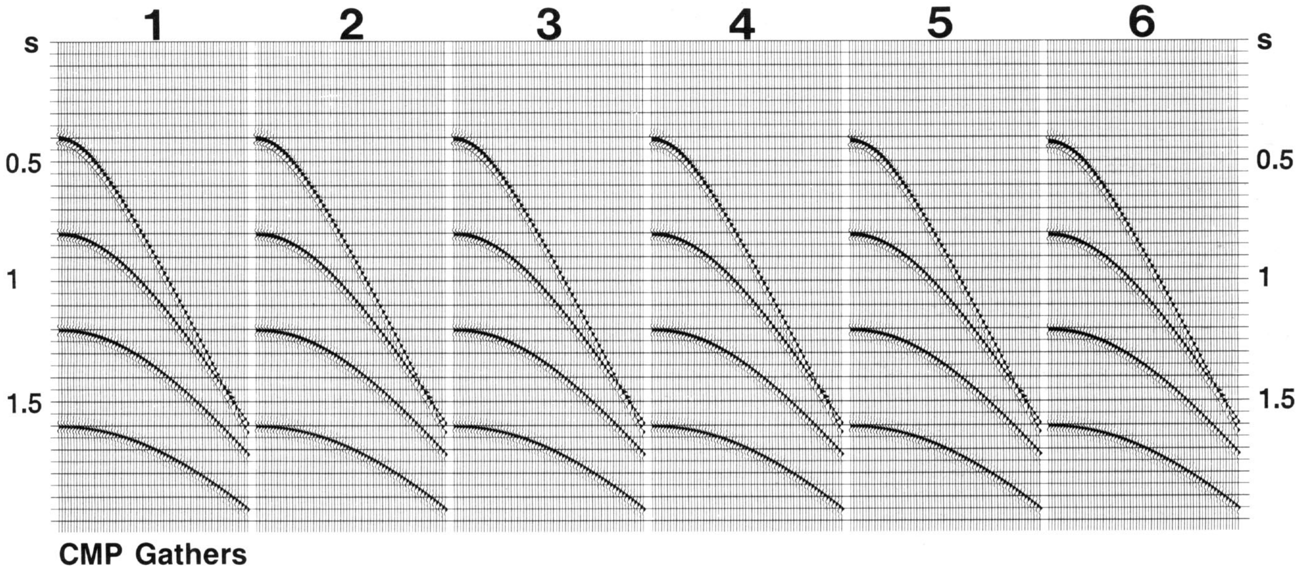


FIG. 3-43. The CMP gathers associated with six neighboring midpoint locations. The reflectors have a gentle downdip from left to right.

gathers. One way is to sum the gathers and compute the velocity spectrum from the sum. This is shown in Figure 3-44b. Another way is to compute the velocity spectra from each individual gather and sum the spectra as shown in Figure 3-44c. Clearly, the former is less costly than the latter. In practice, the number of CMP gathers that may be used must be chosen so that there is negligible dip across the gathers under consideration. If the structural dip is significant, then the number of CMP gathers included in the velocity analysis must be kept small or dip effects must be accommodated. Note that the peak corresponding to the shallow event in Figure 3-44b is smaller than its counterpart in Figure 3-44c. Look closely at the CMP gathers in Figure 3-43 and note the slight difference in traveltimes from gather to gather, especially for the shallow event. Summing these gathers distorts the hyperbolic path and causes degradation in the velocity spectrum.

When the input gather has a significant noise level, some smoothing may be done on the velocity spectrum matrix. This is done by averaging over velocity or time gates, or by some combination of the two. Another way to suppress small-amplitude correlation peaks that may be related to the ambient noise level in the data is to apply some percentage of bias to the correlation values. Biasing refers to subtracting a constant value from the correlation values over the entire velocity spectrum. Various combinations of averaging and biasing of correlation values also are used in practice. Finally, for computational efficiency, correlation values may be computed within a specified velocity corridor (Figure 3-45). The corridor must be chosen so that it spans the velocity variations vertically and laterally in the survey area.

Experience in a survey area helps when picking appropriate stacking velocities for primary reflections from velocity spectra. Acceptable velocity errors vary depending on use of the estimated velocities (Table 3-6).

3.3.3 Horizon Velocity Analysis

One way to estimate velocities with the accuracy required for detailed structural or stratigraphic studies is to analyze the particular horizon of interest continuously. Such detailed velocity estimation is called Horizon Velocity Analysis (HVA). Horizon velocity analysis is an efficient way to get velocity information at every CMP location along selected key horizons, as opposed to the conventional velocity analysis that provides velocity information at every time gate at selected CMP locations.

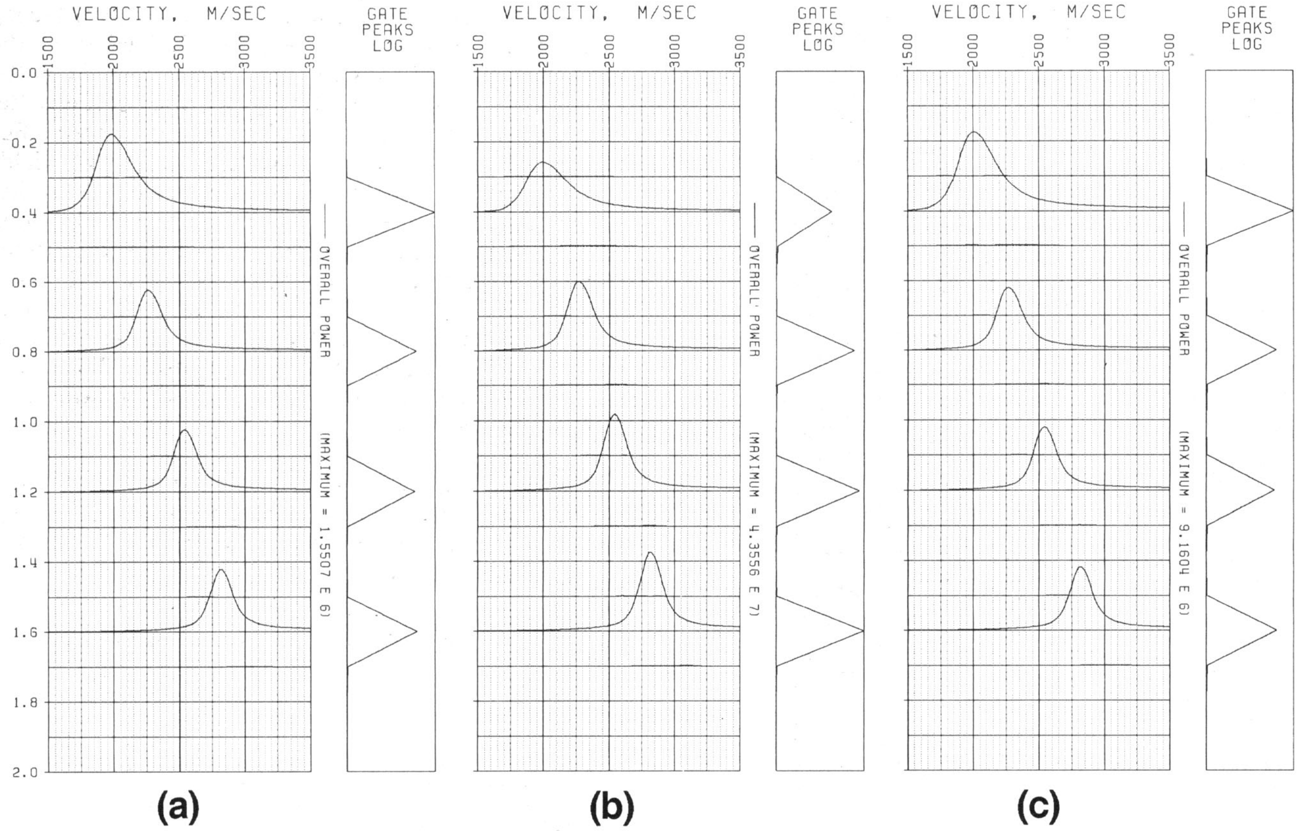


Figure 3-44. (a) Velocity spectrum derived from a single CMP gather (CMP 1) in Figure 3-43, (b) from the sum of six CMP gathers, (c) from the sum of six individual velocity spectra. Panel (b) is relatively cheaper than panel (c).

The underlying principle is the same as that of the velocity spectrum. The output coherency values derived from hyperbolic time gates are displayed as a function of velocity and CMP position. Correlation values are computed from a gate that includes the horizon of interest. Horizon times are digitized and input to the horizon velocity analysis. Figures 3-46 and 3-47 are a stacked section and an HVA over five horizons. Similar types of computational details, such as smoothing and biasing, are considered as for velocity spectrum.

Whenever there are structural discontinuities on a stacked section, HVA is carried out on segments of the horizon that are separated by faults. Horizon velocity analysis can improve a stacked section in preparation for poststack depth migration (Section 5.2). This is somewhat surprising, since HVA still is based on hyperbolic moveout, while data that require depth migration often have complex moveout. Nevertheless, in practice, HVA provides the detailed lateral velocity variations along a marker horizon, which may be missed by conventional velocity analyses that are sparsely spaced along the line. Consider horizon A in Figure 3-48, which is below the salt dome S. The salt dome behaves as a complex overburden, causing the raypaths that are associated with the underlying reflectors to bend. Note the rapid lateral changes in velocity (Exercise 3.16) and improvement in the CMP stack after using the HVA picks.

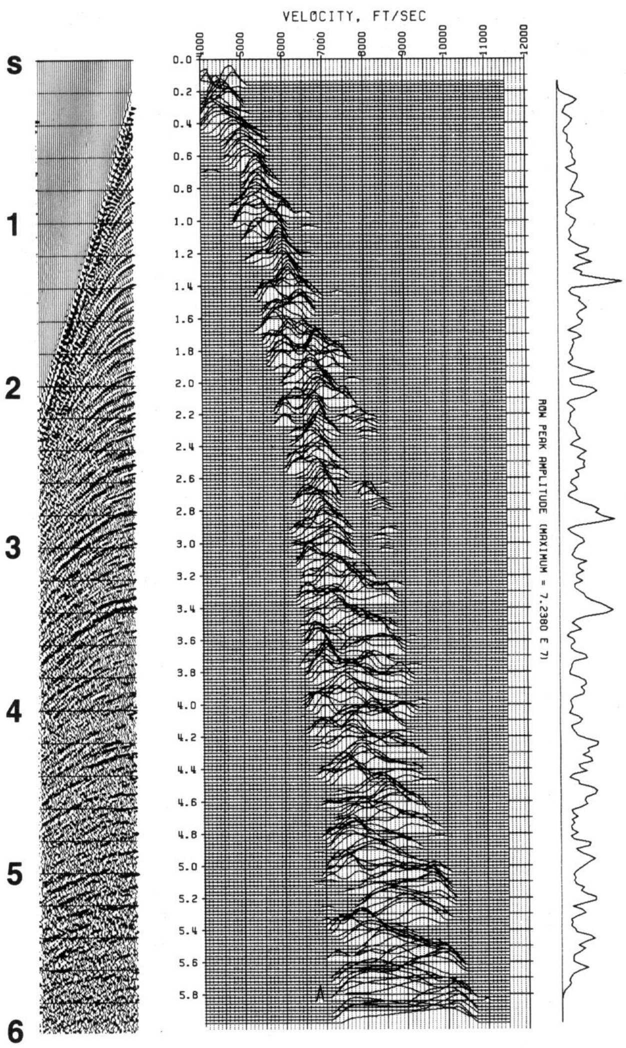


FIG. 3-45. To save computation, velocity spectrum can be estimated within a velocity corridor dictated by the dominant velocity trend.

3.4 RESIDUAL STATICS CORRECTIONS

Reflection times often are affected by irregularities in the near-surface. This is best demonstrated by the real data example in Figure 3-49. While the shot gathers on the left contain reflections with reasonably well-behaving hyperbolic moveout, those on the right have reflections that significantly depart from regular hyperbolic moveout. Although such distortions can be caused by a structural complexity deeper in the subsurface, more often they result from near-surface irregularities. Field statics corrections remove a significant part of these traveltime distortions from the data. Nevertheless, these corrections usually do not account for rapid changes in elevation, the base of weathering, and weathering velocity.

Figures 3-50a and 3-51a are selected CMP gathers (with field statics corrections) that were NMO corrected using a set of preliminary velocity picks derived from the velocity analyses in Figure 3-52. Deviations from the hyperbolic trends on the CMP gathers significantly degrade the quality of some of the velocity spectra. For instance, velocity analysis at CMP location 188 yields relatively poorer quality picks than those from other velocity analyses. The CMP gathers in the neighborhood of CMP location 188 have more traveltime distortions compared to some other CMP gathers (Figure 3-50a). The resulting stacked section could be misleading in that residual statics may cause dim spots along the reflection horizons as well as false structures (Figure 3-53a), particularly between midpoints 101 to 245. False structures also are apparent on the rms AGC gained stack (Figure 3-54a) in which dim spots may not be apparent.

Obviously a more correct picture of the subsurface should be attained from data corrected for rapidly varying near-surface effects. After making these *residual*

Table 3-6. Acceptable velocity errors (after Schneider, 1971).

Use of Velocity	Acceptable Error, % rms	Acceptable Error, % Interval
NMO corrections for conventional stack.	2-10	—
Structural anomaly detection: 100-ft anomaly at 10,000-ft depth.	0.5	—
Gross lithologic identification: 1000-ft interval at 10,000-ft depth.	0.7	10
Stratigraphic detailing: 400-ft interval at 10,000-ft depth.	0.1	3

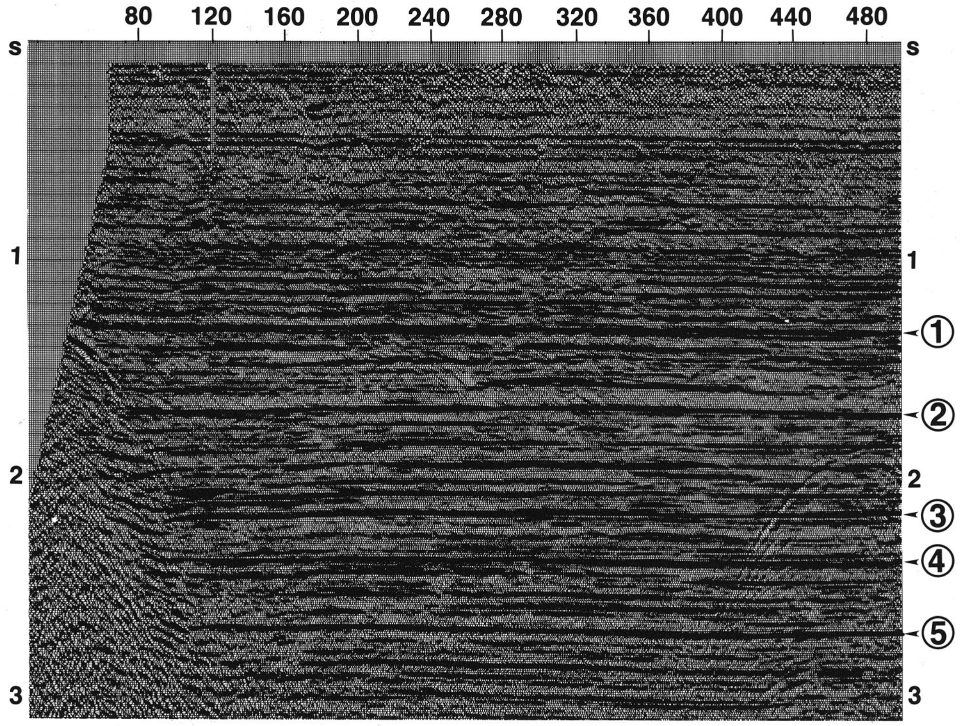


FIG. 3-46. A stacked section with five marker horizons as indicated.

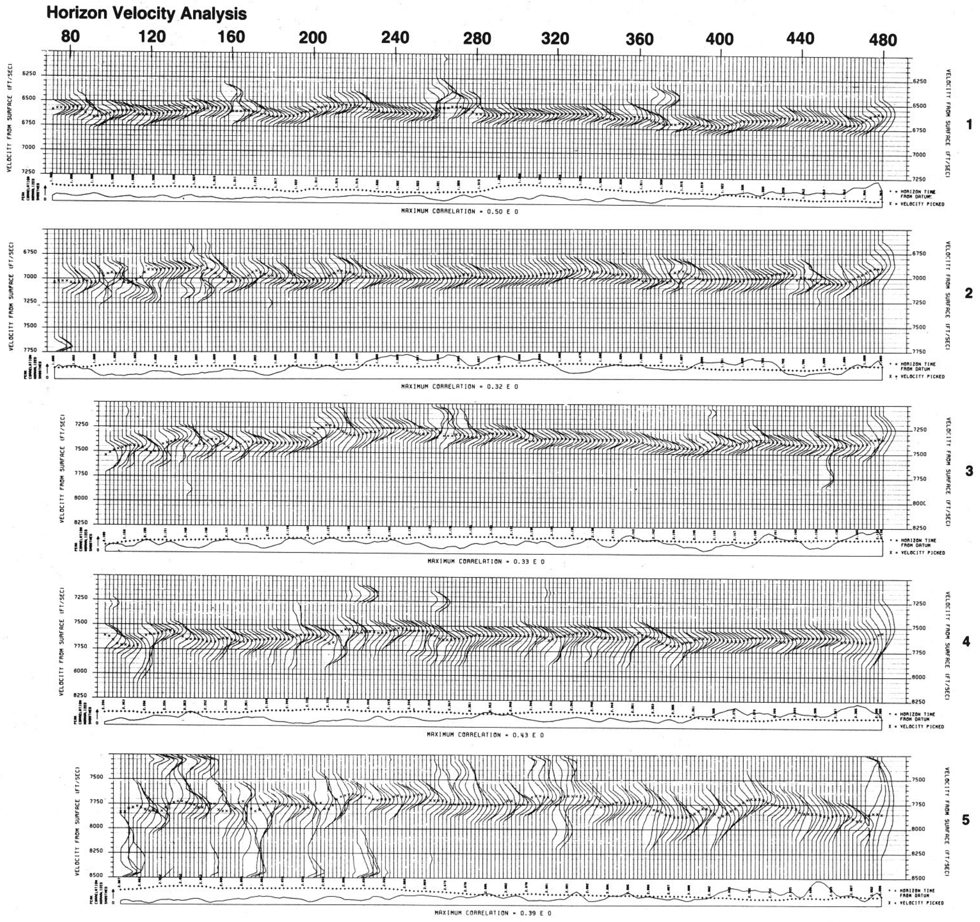


FIG. 3-47. Horizon velocity analyses along five marker horizons indicated in Figure 3-46. The vertical and horizontal axes in each panel are stacking velocity and CMP axis, respectively.

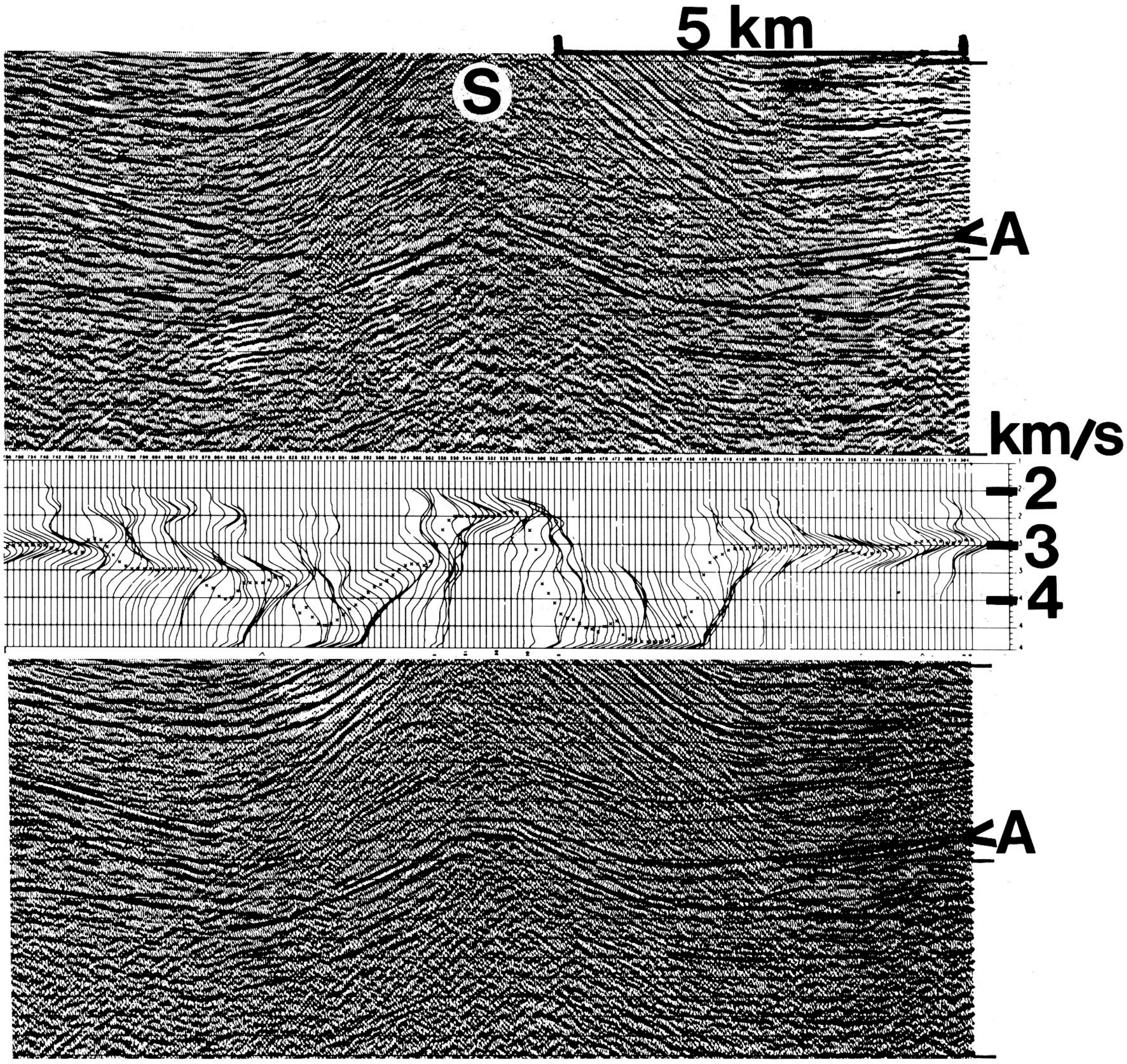


FIG. 3-48. A CMP stacked section obtained by sparsely spaced conventional velocity analysis (top) and by HVA (bottom). The HVA for horizon A below the salt dome (*S*) is shown in the center.

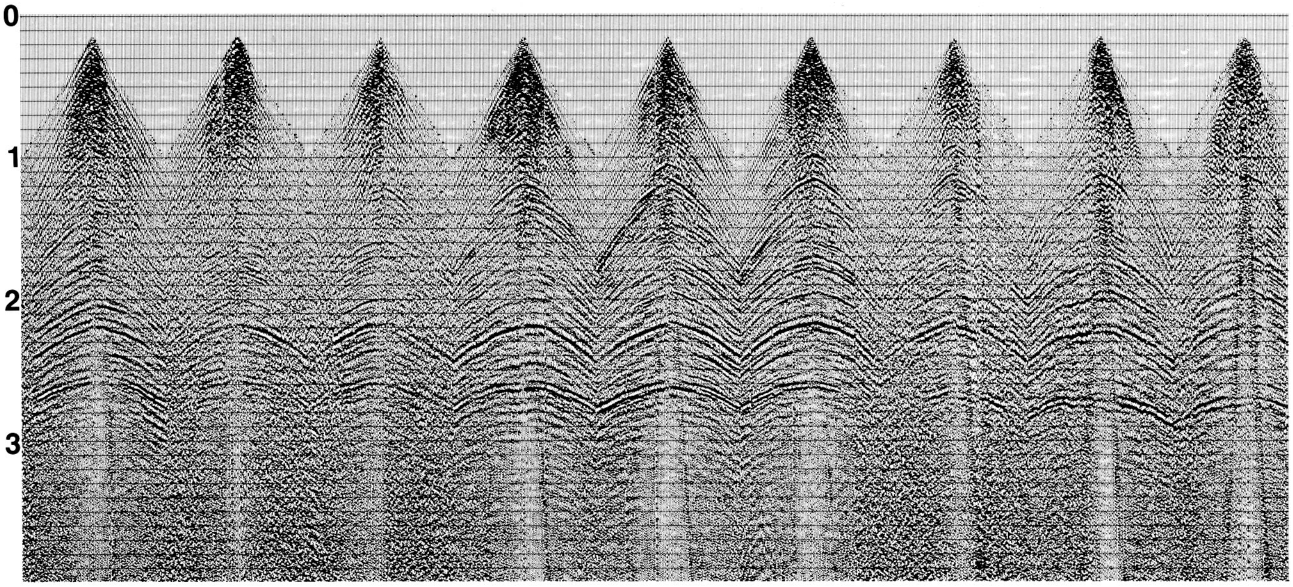


FIG. 3-49. Common-shot gathers from a land profile. Note the departures from hyperbolic traveltimes on the gathers at the right.

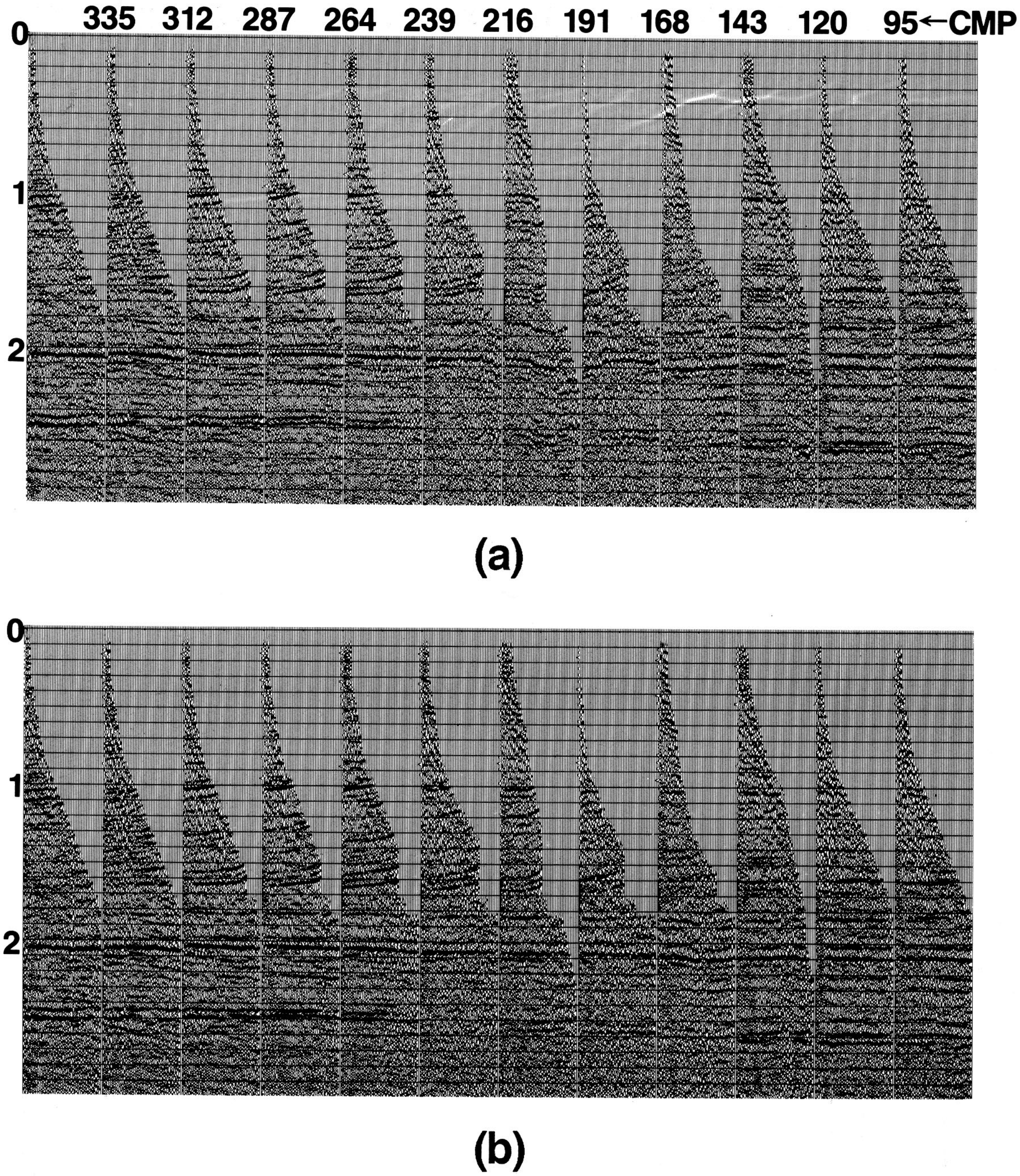


FIG. 3-50. CMP gathers from a land profile: (a) before residual statics corrections, (b) after residual statics corrections. (Shot gathers are shown in Figure 3-49.) NMO correction was applied using preliminary velocity picks derived from the spectra in Figure 3-52. The CMP stacks are shown in Figure 3-53.

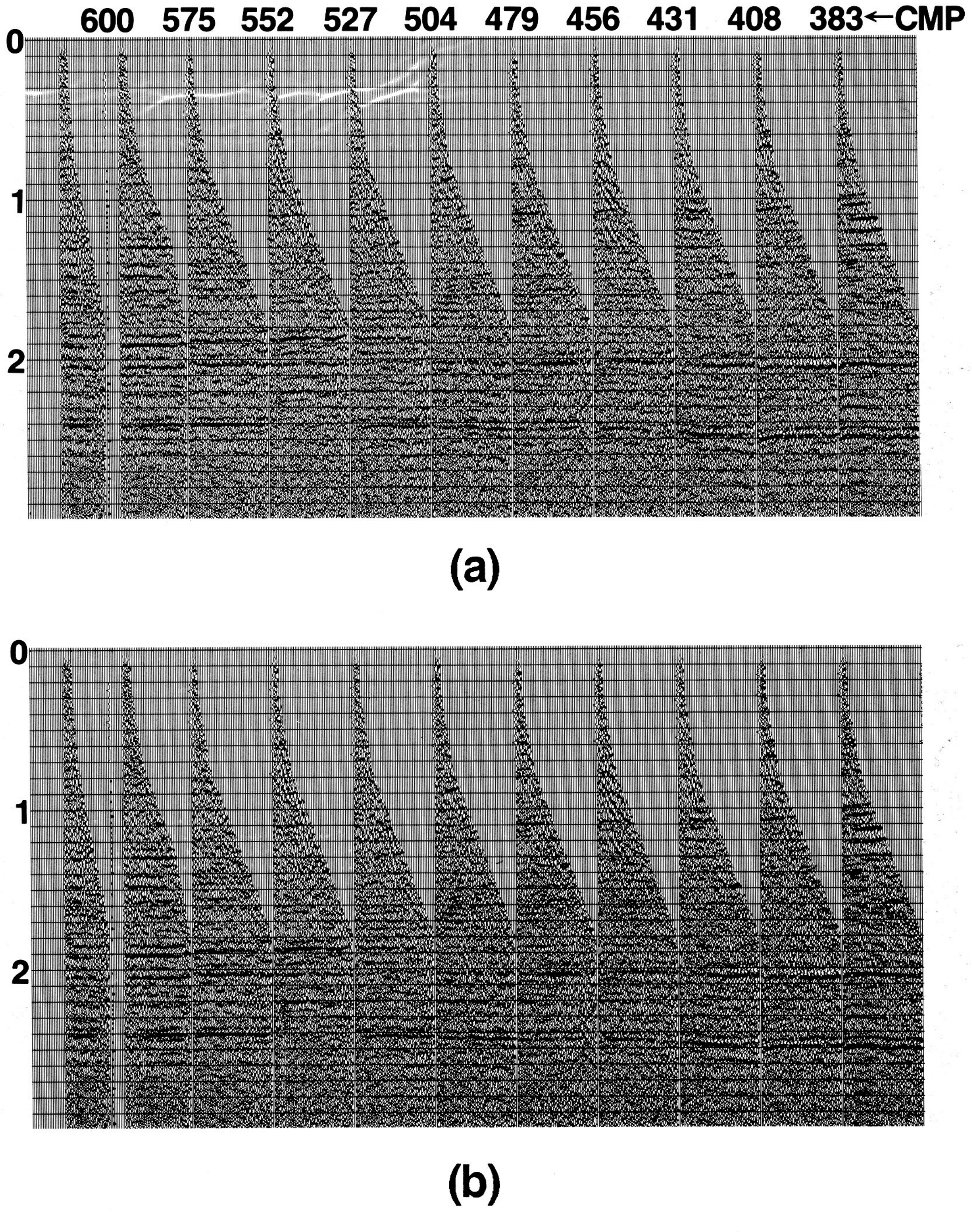


FIG. 3-51. CMP gathers from the same land line as in Figure 3-50 (a) before and (b) after residual statics corrections. This part of the line does not have as severe a statics problem as that shown in Figure 3-50. The NMO correction was applied using preliminary velocity picks derived from the spectra in Figure 3-52.

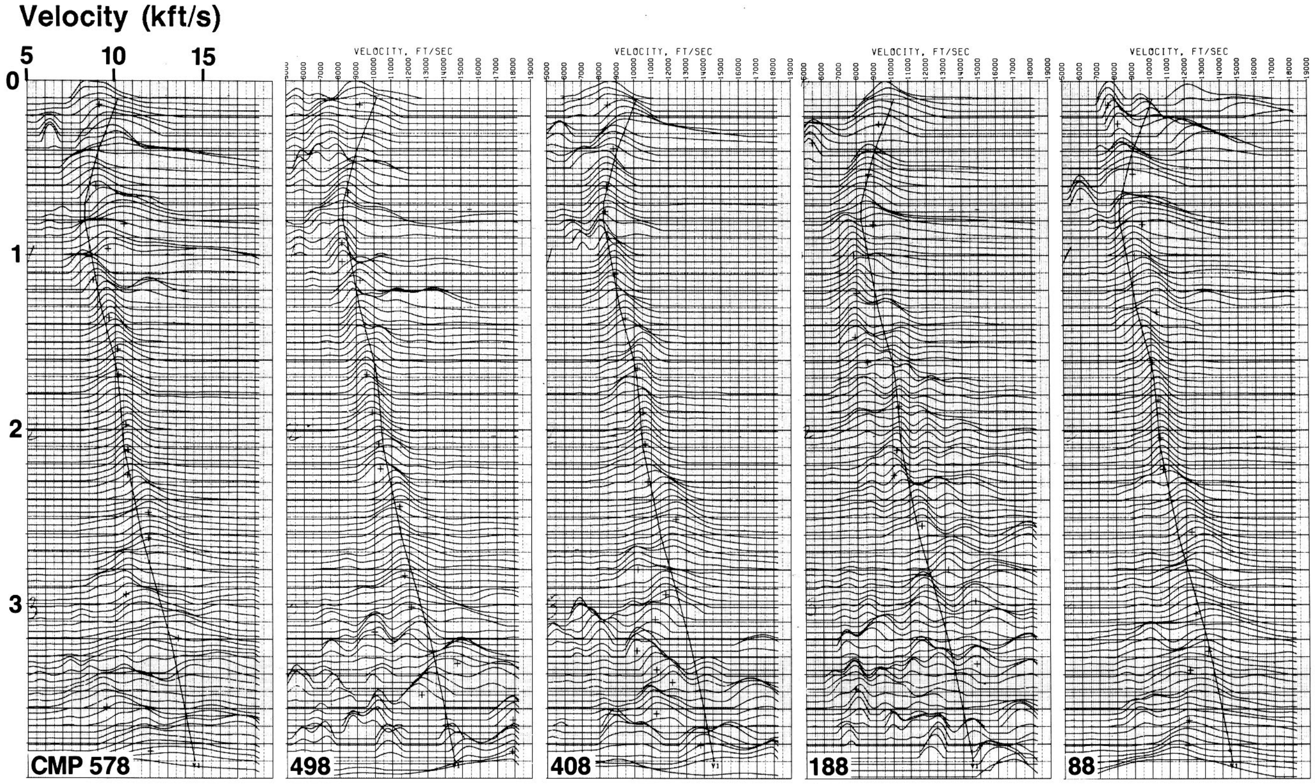


FIG. 3-52. Velocity analyses (before residual statics corrections) along the land line shown in Figure 3-53. Figures 3-50 and 3-51 show selected CMP gathers.

statics corrections, the CMP gathers with traveltime deviations show better alignment of reflections (Figure 3-50b), while those that did not require such corrections are unchanged (Figure 3-51b). After the residual statics corrections, the ungained (Figure 3-53b) and gained stacked sections (Figure 3-54b) show improvement of continuity of reflections as well as significant elimination of false structures (refer to the segment between midpoints 101 to 245).

Following the residual statics corrections, velocity analyses usually are repeated to update the velocity picks (Figure 3-55). Comparison of Figures 3-52 and 3-55 shows that residual statics corrections have improved the velocity analysis. The same CMP gathers after NMO corrections using the updated velocity picks are shown in Figures 3-56a and 3-57a, while the same gathers after residual statics corrections are shown in Figures 3-56b and 3-57b. Comparison of the CMP gathers before and after residual statics corrections shows significant elimination of time deviations. The resulting stacked sections using the revised velocity estimates are shown in Figure 3-58, while the gained stacks are shown in Figure 3-59.

Residual statics corrections usually are discussed in terms of applications to land data. However, in certain cases, residual statics corrections have produced dramatic improvement in marine data. Areas with irregular water-bottom topography in shallow water (less than 25 m), and areas with rapidly varying velocity in the sediments near the water bottom are places where statics corrections have been successful.

Figure 3-60 shows a commonly used flowchart for residual statics corrections and velocity analysis aimed at producing an optimum stacked section. In practice, this flowchart usually is augmented with additional quality control steps. It often is necessary to examine CMP gathers and velocity analyses after residual statics corrections. Diagnostic tools allow determination of the magnitude of these corrections. For example, common-shot and common-receiver gathers indicate relative static shifts from one receiver location to another and from one shot location to another (Figures 3-61 and 3-62, respectively). Also, common-shotpoint and common-receiver-point stacks can be used with common-receiver and common-shot gathers, respectively. A common-shot-point stack should indicate the range of magnitude of shot static shifts (Figure 3-63); a common-receiver-point stacked section should indicate the range of magnitude of receiver static shifts (Figure 3-64) along the line. These displays enable determination of an optimum maximum allowable shift to consider while estimating (picking)

(Text continued on page 196)

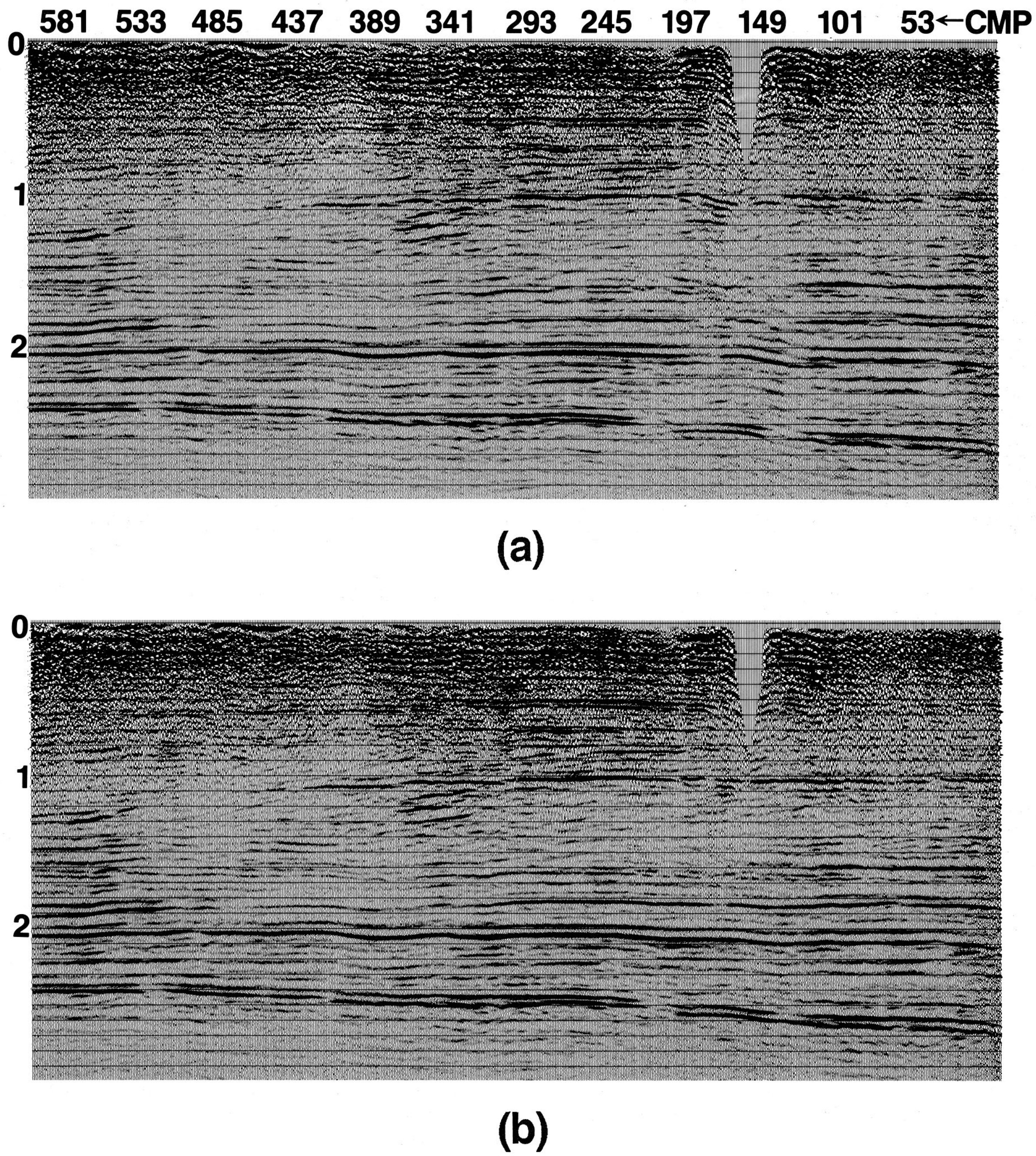


FIG. 3-53. CMP stacks derived from gathers in Figures 3-50 and 3-51. Stack (a) before and (b) after residual statics corrections. NMO correction was applied using preliminary velocity picks derived from the spectra in Figure 3-52.

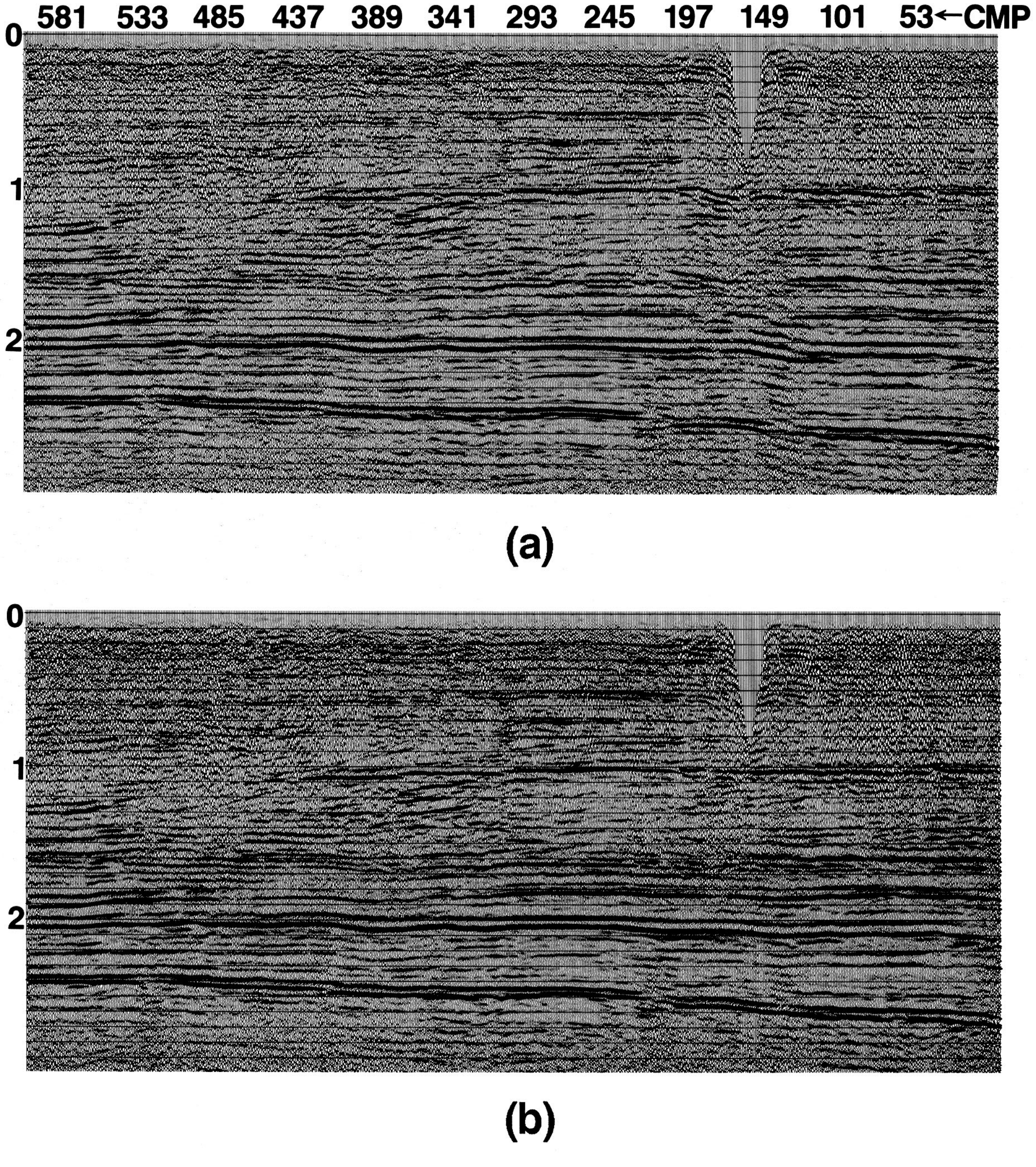


FIG. 3-54. The same CMP stacks as in Figure 3-53 with rms gain applied. Gained stacks (a) before and (b) after residual statics corrections.

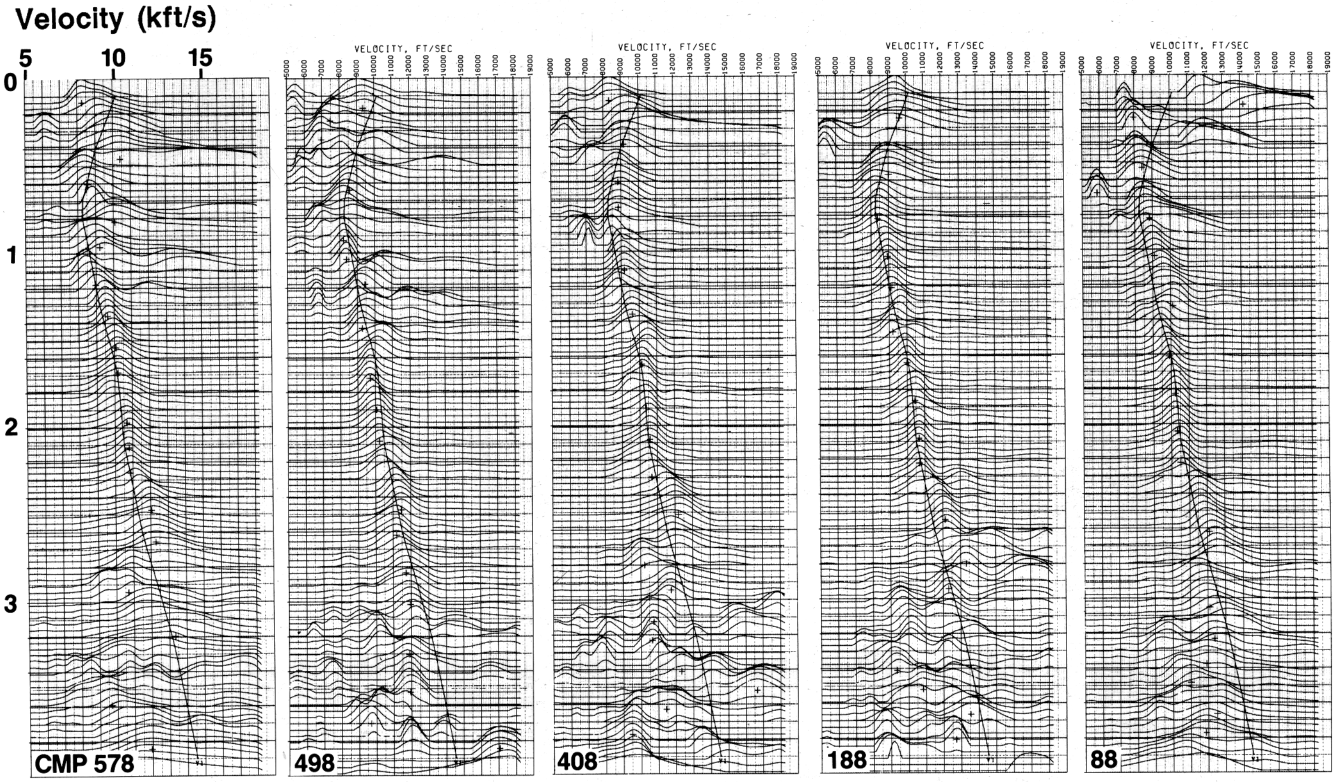


FIG. 3-55. Velocity analyses (after residual statics corrections) along the line shown in Figure 3-53. Figures 3-56 and 3-57 show CMP gathers.

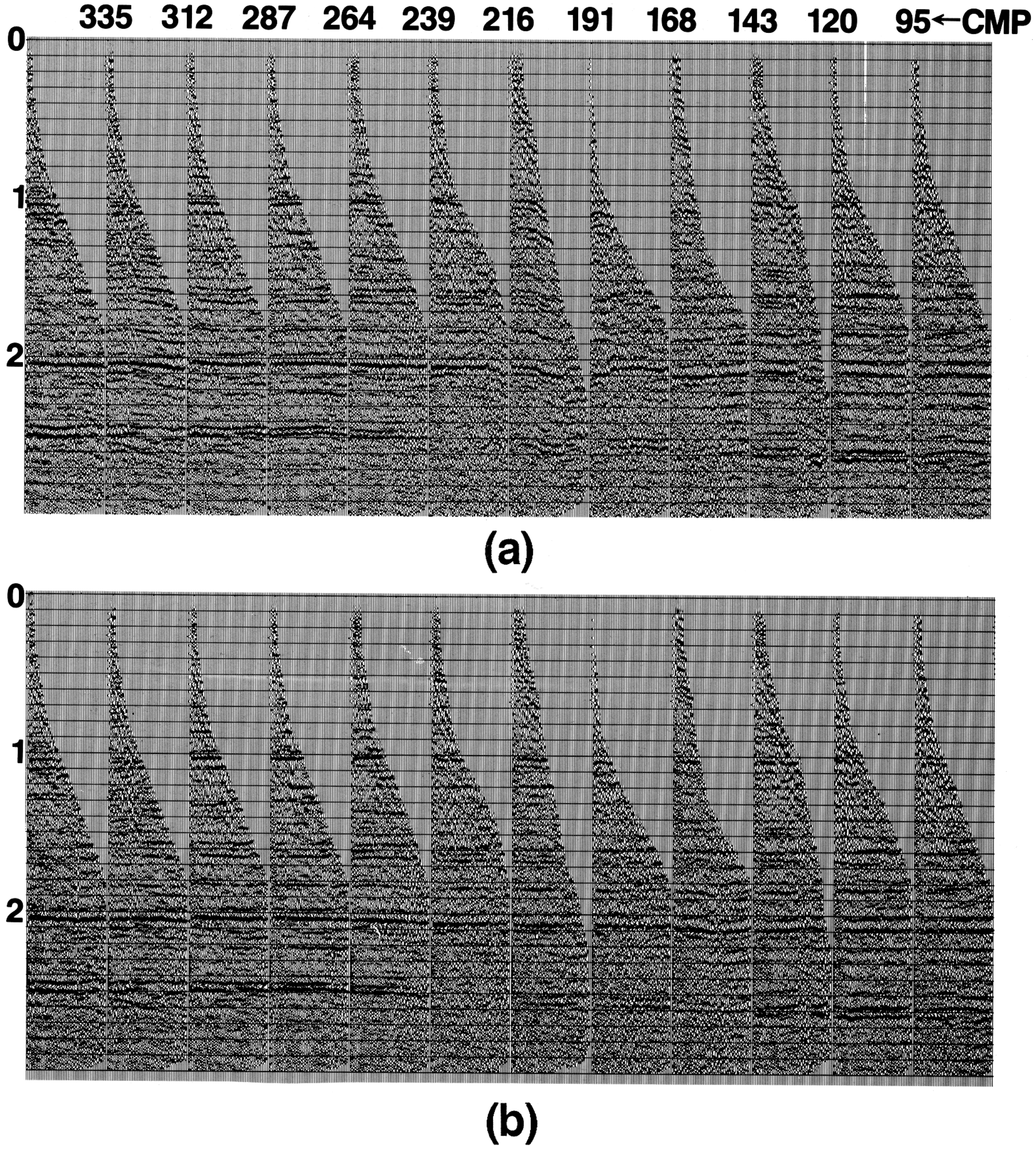


FIG. 3-56. CMP gathers from the land line shown in Figure 3-58. CMP gathers (a) without and (b) with residual statics corrections. (Figure 3-49 shows the shot gathers from the same line.) NMO correction was applied using final velocity picks derived from the spectra in Figure 3-55.

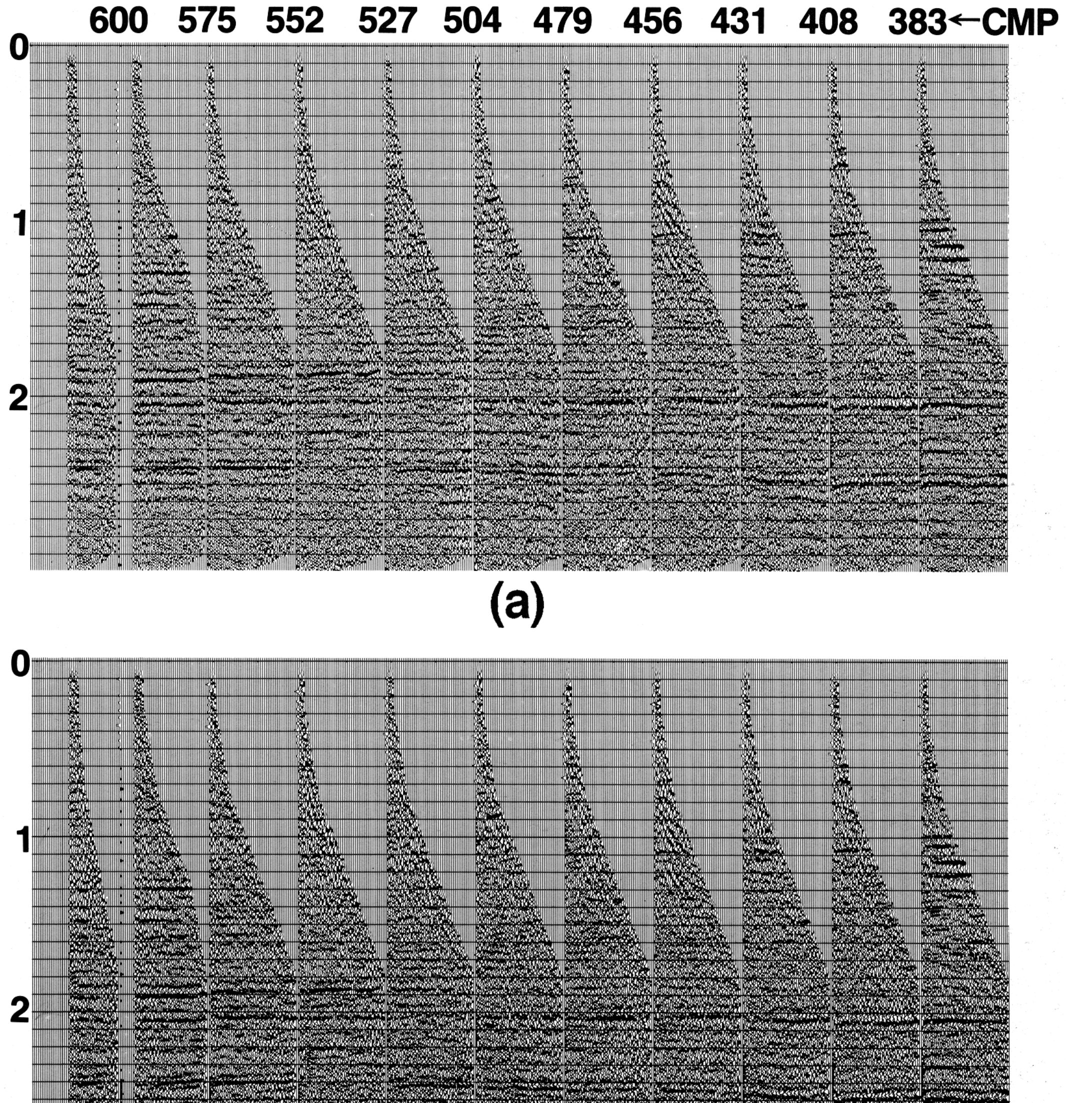


FIG. 3-57. CMP gathers from the same land line as in Figure 3-56. CMP gathers (a) without and (b) with residual statics corrections. This part of the line does not have as severe a statics problem as the part in Figure 3-56. NMO correction was applied using final velocity picks derived from the spectra in Figure 3-55.

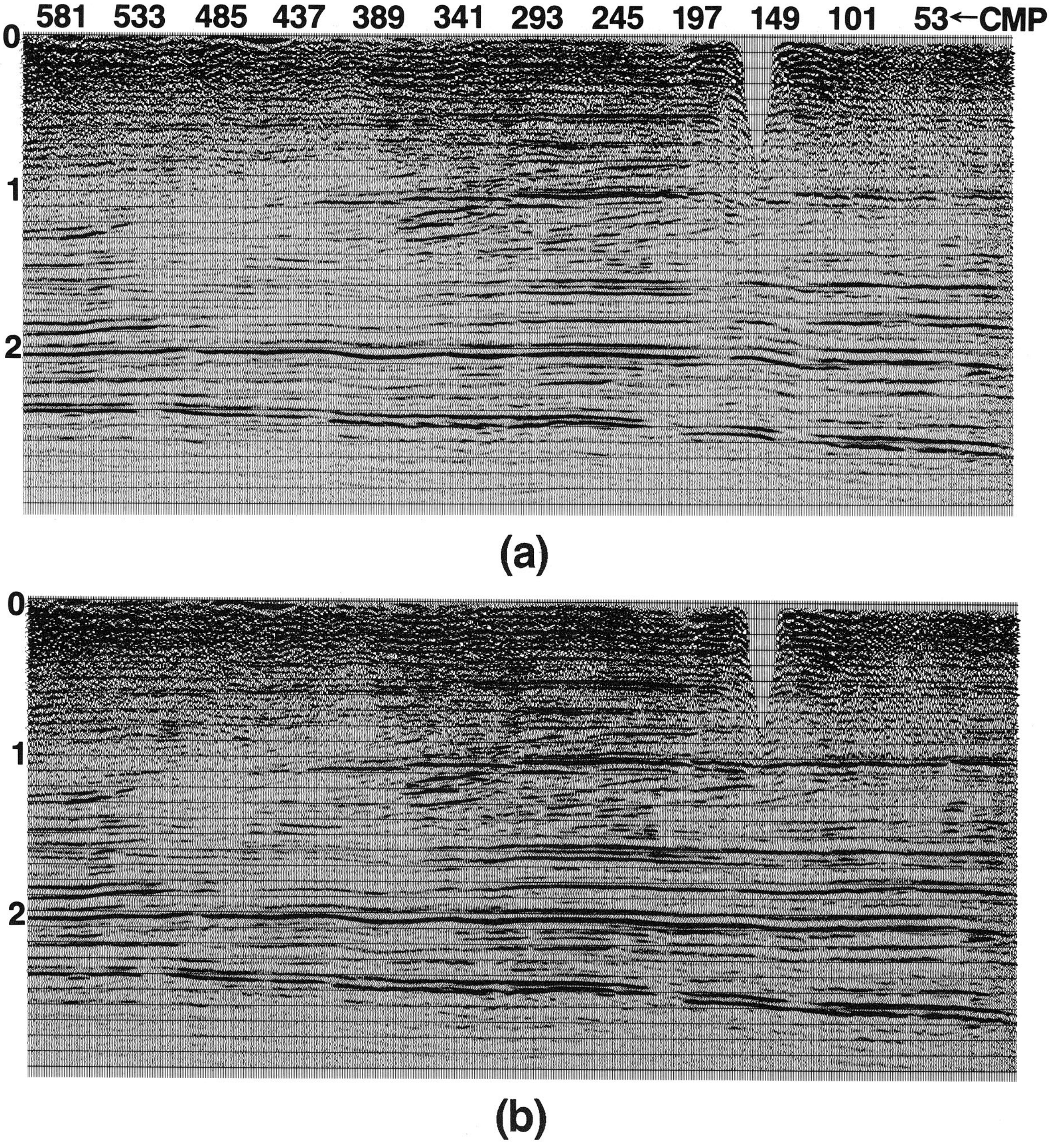


FIG. 3-58. CMP stacks derived from the gathers in Figures 3-56 and 3-57. Stack (a) before and (b) after residual statics corrections and using revised velocity estimates from Figure 3-55.

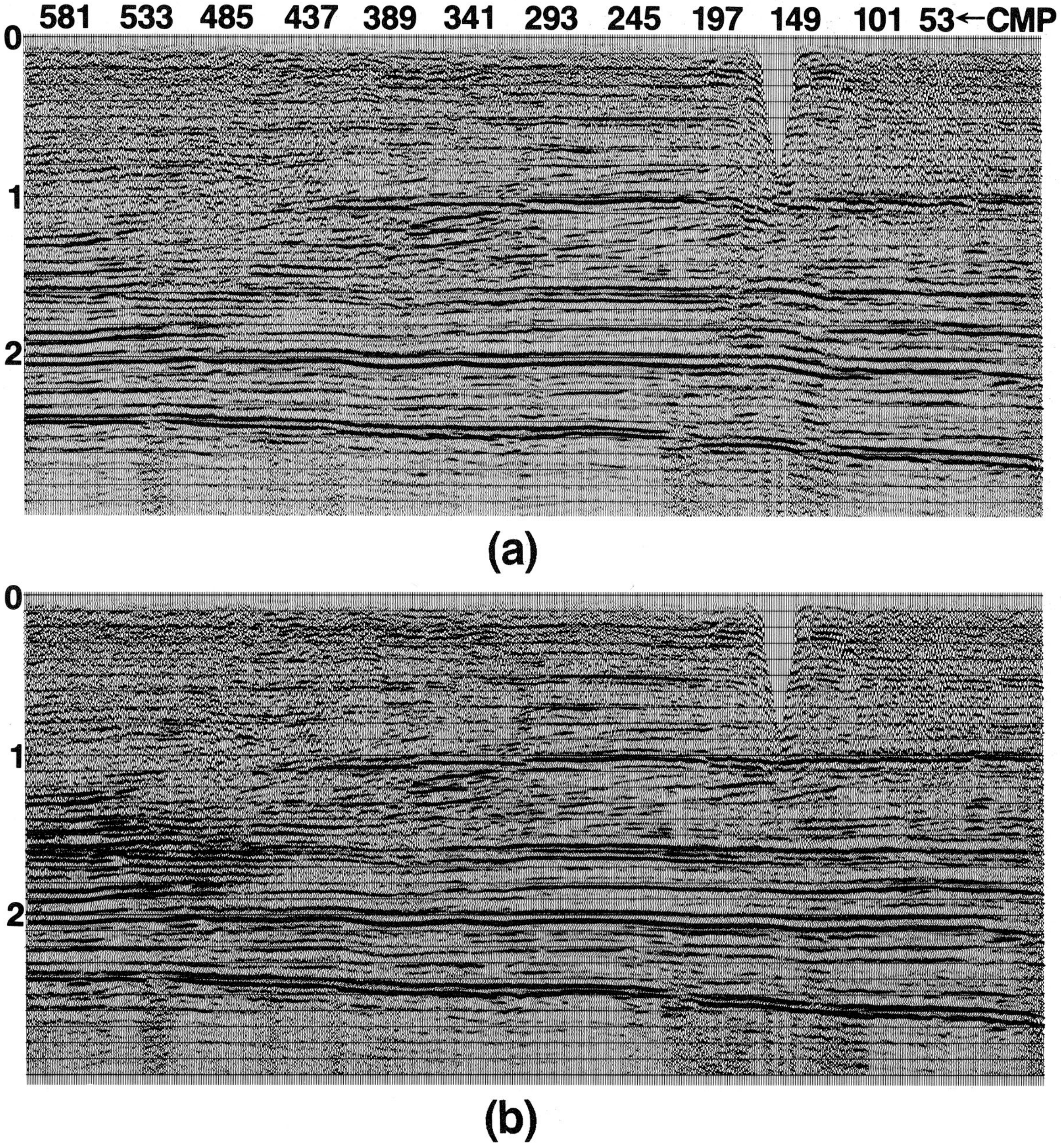


FIG. 3-59. The same CMP stacks as in Figure 3-58 with rms gain applied. Gained stacks (a) before and (b) after residual statics corrections and using revised velocity estimates from Figure 3-55.

residual statics from the data. From the example in Figures 3-63 and 3-64, the receiver component of static shifts is greater than the shot component.

3.4.1 Surface-Consistent Residual Statics Corrections

Static deviations from a perfect hyperbolic traveltime trajectory are illustrated in Figure 3-65. After NMO correction, the misalignment of the waveform across the CMP gather will yield a poor stack trace. We want to estimate the time shifts from the time of perfect alignment, then correct for them by using an automatic procedure. To do this, a model is needed for the moveout-corrected traveltime from a source location to a depth point on a reflecting horizon, then back to a receiver location. Figure 3-66 shows the geometry and notation that will be used in defining this model. The key assumption in the commonly used traveltime model discussed here is that residual statics are surface consistent (Hileman et al., 1968; Taner et al., 1974). This means that static shifts are time delays that solely depend on source or receiver locations at the surface, not on raypaths in the subsurface. This assumption is valid if all raypaths, regardless of source-receiver offset, are vertical in the near surface layering. Since the weathered layer usually has quite a low velocity and refraction at its base, which tends to make travel paths vertical, the surface-consistent assumption usually is a good one. The assumption may not be good for a high-velocity permafrost layer that causes rays to bend away from vertical.

The traveltime t_{ijh} that corresponds to the jth source station, the ith receiver station, and the kth [$k = (\mathrm{i} + j)/2$] midpoint along the hth horizon can be approximately modeled as

$$t_{ijh} = s_j + r_i + G_{kh} + M_{kh} x_{ij}^2, \tag{3.25}$$

where s_j is the residual static time shift associated with the jth source station, r_i is the residual static time shift associated with the ith receiver station, and G_{kh} is the difference in two-way time at a reference CMP location (typically CMP location 1) and the traveltime at the kth CMP location along the hth horizon. This term refers to structural variations along the horizon and is called the structural term. The term $M_{\mathrm{kh}}\, x_{ij}^2$ is the residual moveout that accounts for the imperfect moveout correction of the hth horizon. The coefficient M has the dimensions of (time/distance^2).

To be specific in analysis of the system of equations implied by equation (3.25), assume NS shot locations, NR receiver locations, and NG CMP locations. Then define the fold as NF. The purpose is to decompose the observed traveltimes estimated (picked) from the data t'_{ijh} to their individual components as defined on the right side of equation (3.25). The number of time picks (or individual equations) is equal to NG · NF. The number of unknowns is $NS + NR + NG + NG$. Typically, $NG \cdot NF > NS + NR + NG + NG$; there are more equations than unknowns. This is a least-squares problem in which we must minimize the sum of the least-squares error energy between the observed picks t'_{ijh} and the modeled times t_{ijh}:

$$E = \sum_{ijh}(t_{ijh} - t'_{ijh})^2. \tag{3.26}$$

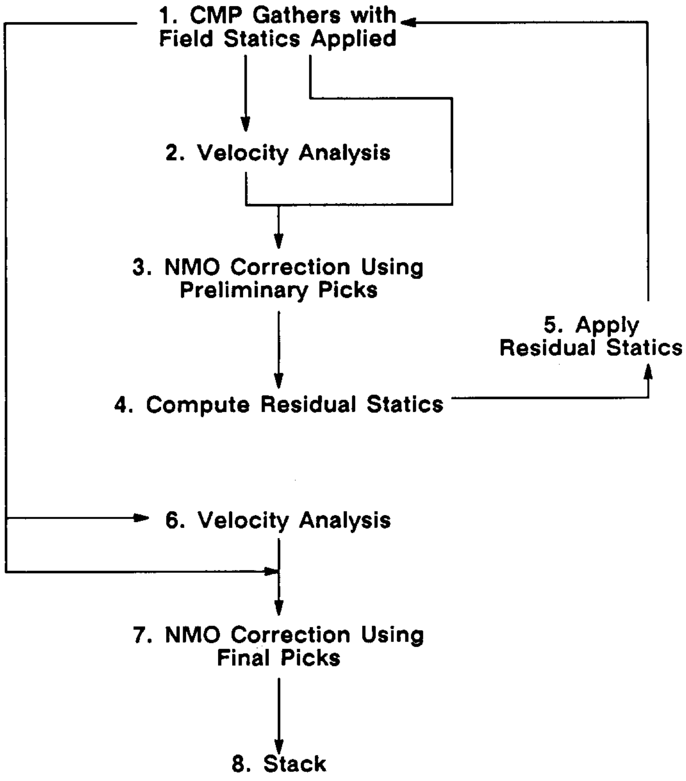


FIG. 3-60. Processing flowchart with residual statics corrections.

Residual statics corrections involve three phases:

1. Picking time values t'_{ijh}.
2. Decomposition of t'_{ijh} into its components: source and receiver statics, structural and residual moveout terms [equation (3.25)].
3. Application of derived source and receiver terms s_j and r_i, respectively, to traveltimes on the pre-NMO-corrected CMP gathers.

The picking phase relates to estimating traveltimes t'_{ijh} from the data. Several picking schemes are in use in the industry. The one described here often is known as a pilot trace scheme. Starting with the CMP gathers that are NMO corrected using a preliminary velocity function(s), trace amplitudes are scaled to a common rms amplitude in the time gates to be used for picking. For clarity, consider a time gate over the kth midpoint gather as illustrated in Figure 3-65. (It is preferable to start with a

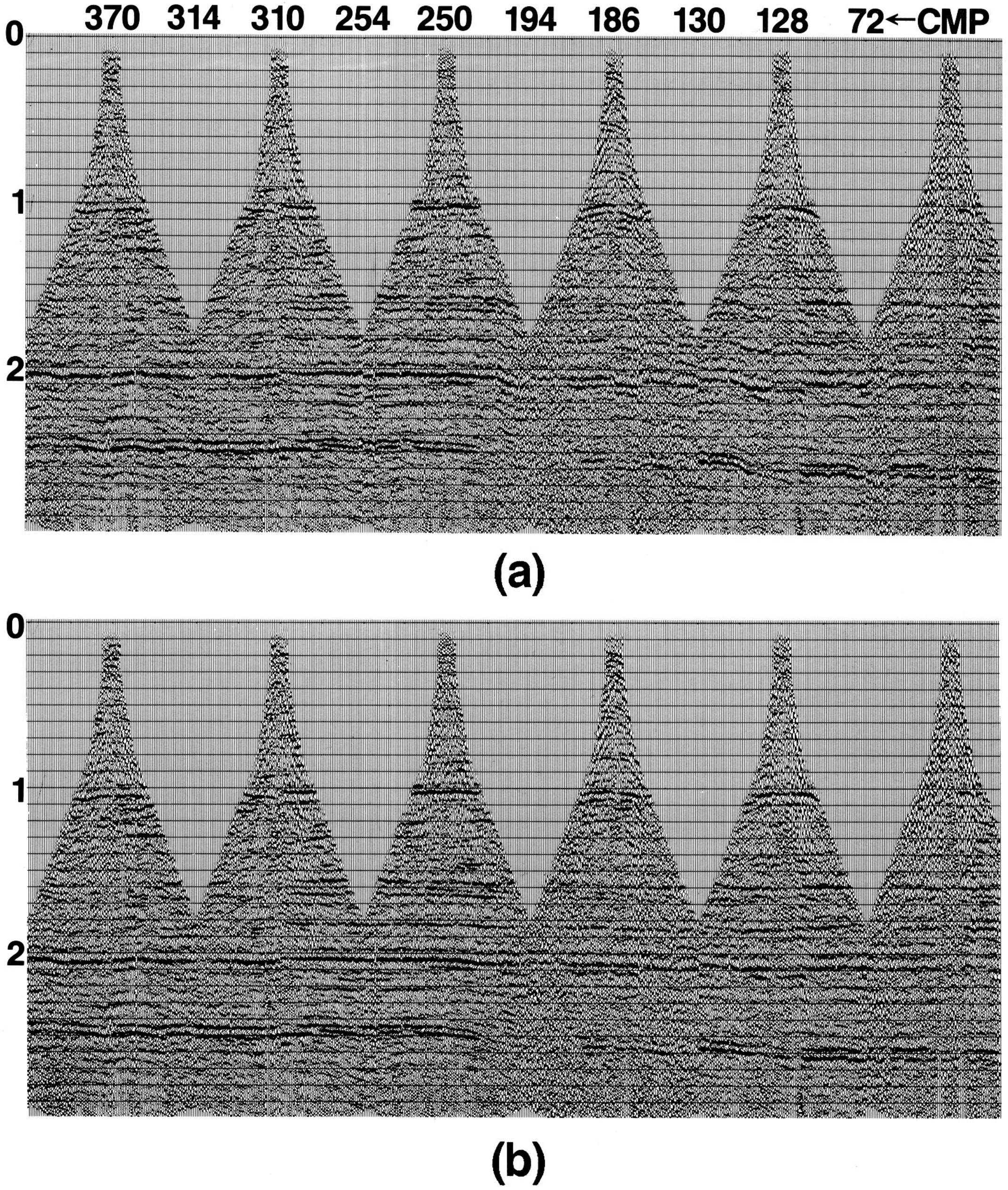


FIG. 3-61. Moveout-corrected common-shot gathers from the same land line as in Figure 3-49. Common-shot gathers (a) before and (b) after residual statics corrections.

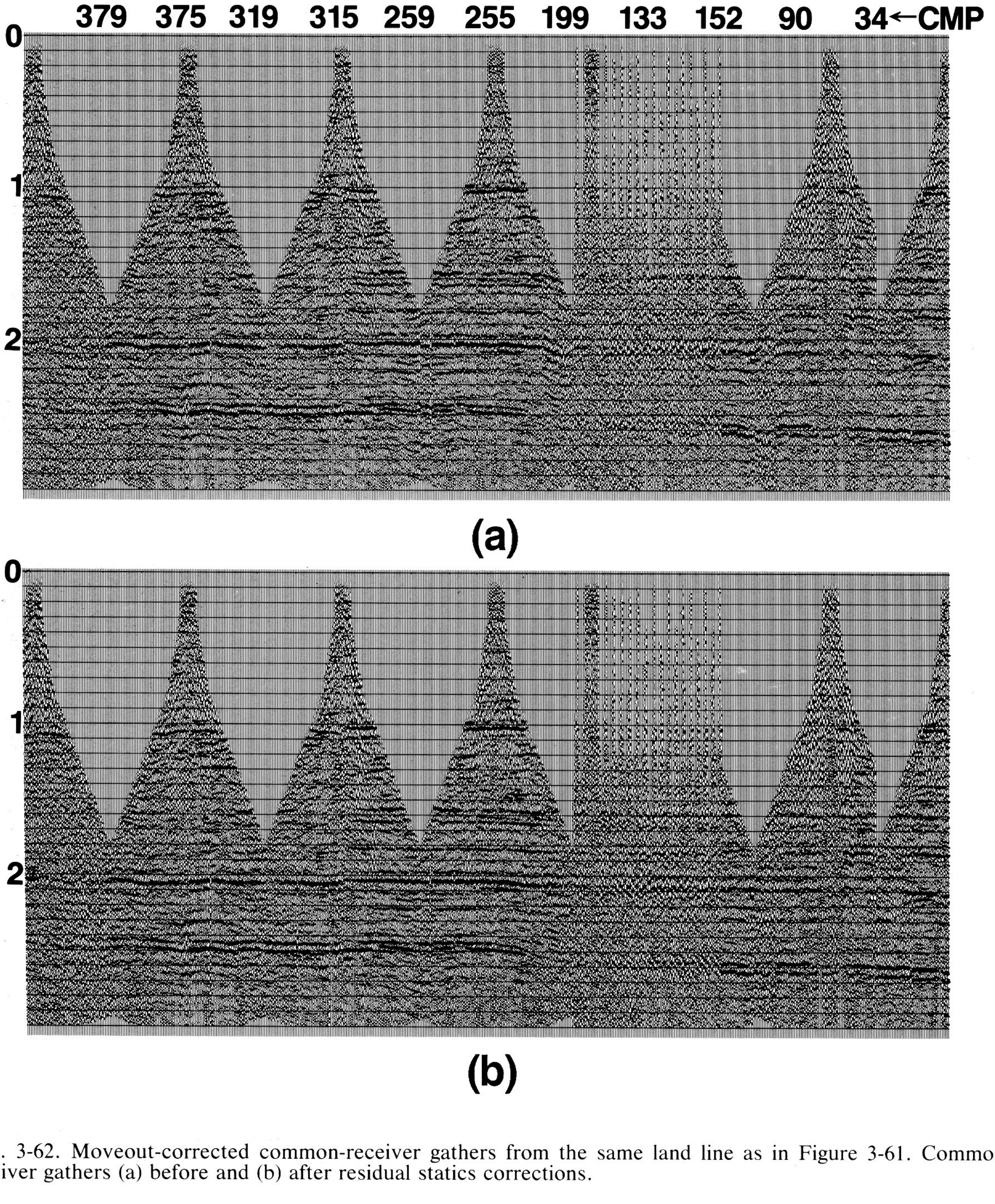


FIG. 3-62. Moveout-corrected common-receiver gathers from the same land line as in Figure 3-61. Common-receiver gathers (a) before and (b) after residual statics corrections.

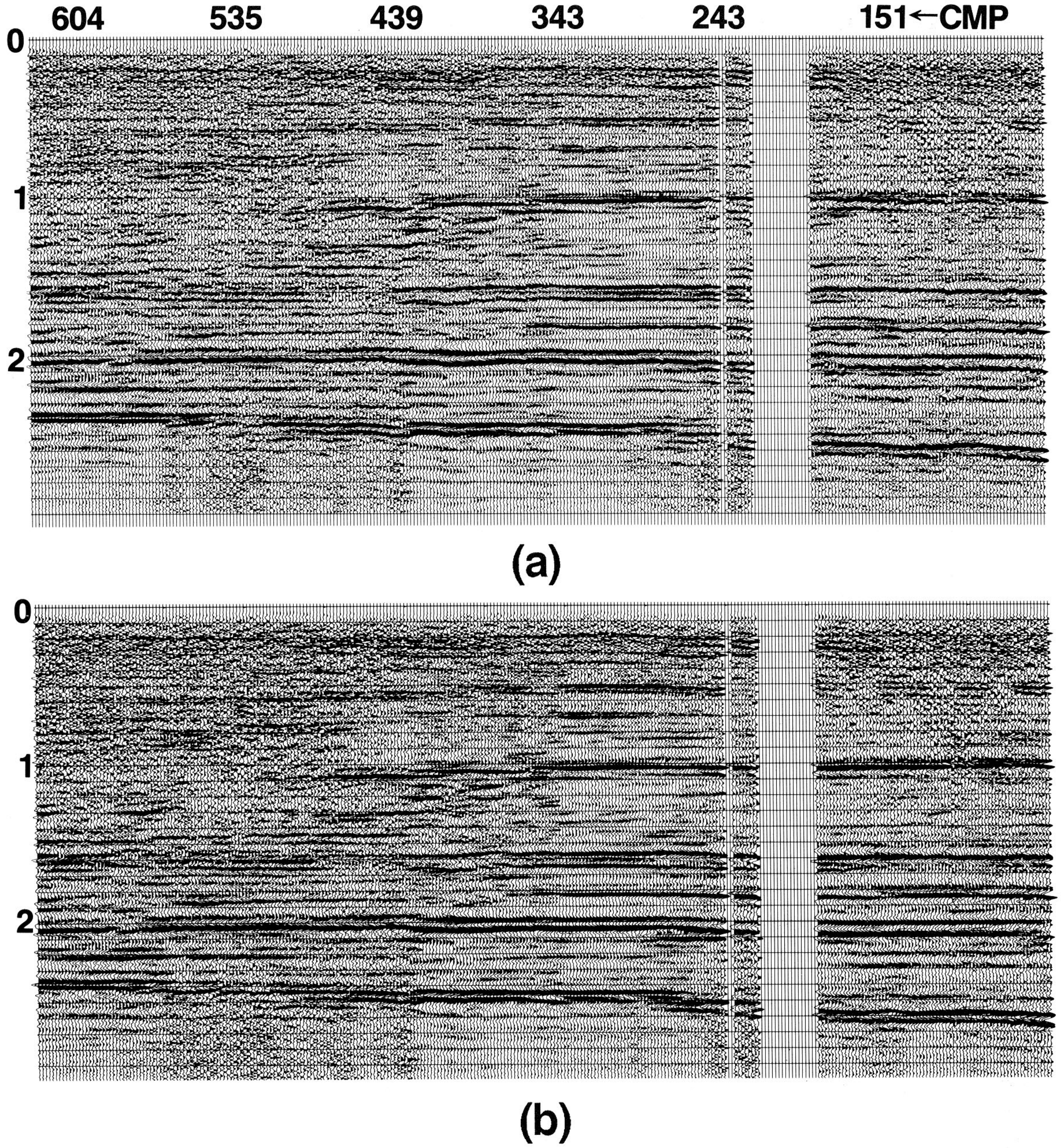


FIG. 3-63. A diagnostic display for residual statics corrections. Common-shotpoint (CSP) stack (a) before and (b) after residual statics corrections. Note the missing shots between CMP 151 and 243. The CSP stack can be used to estimate the magnitude and spatial variation of shot statics along a line.

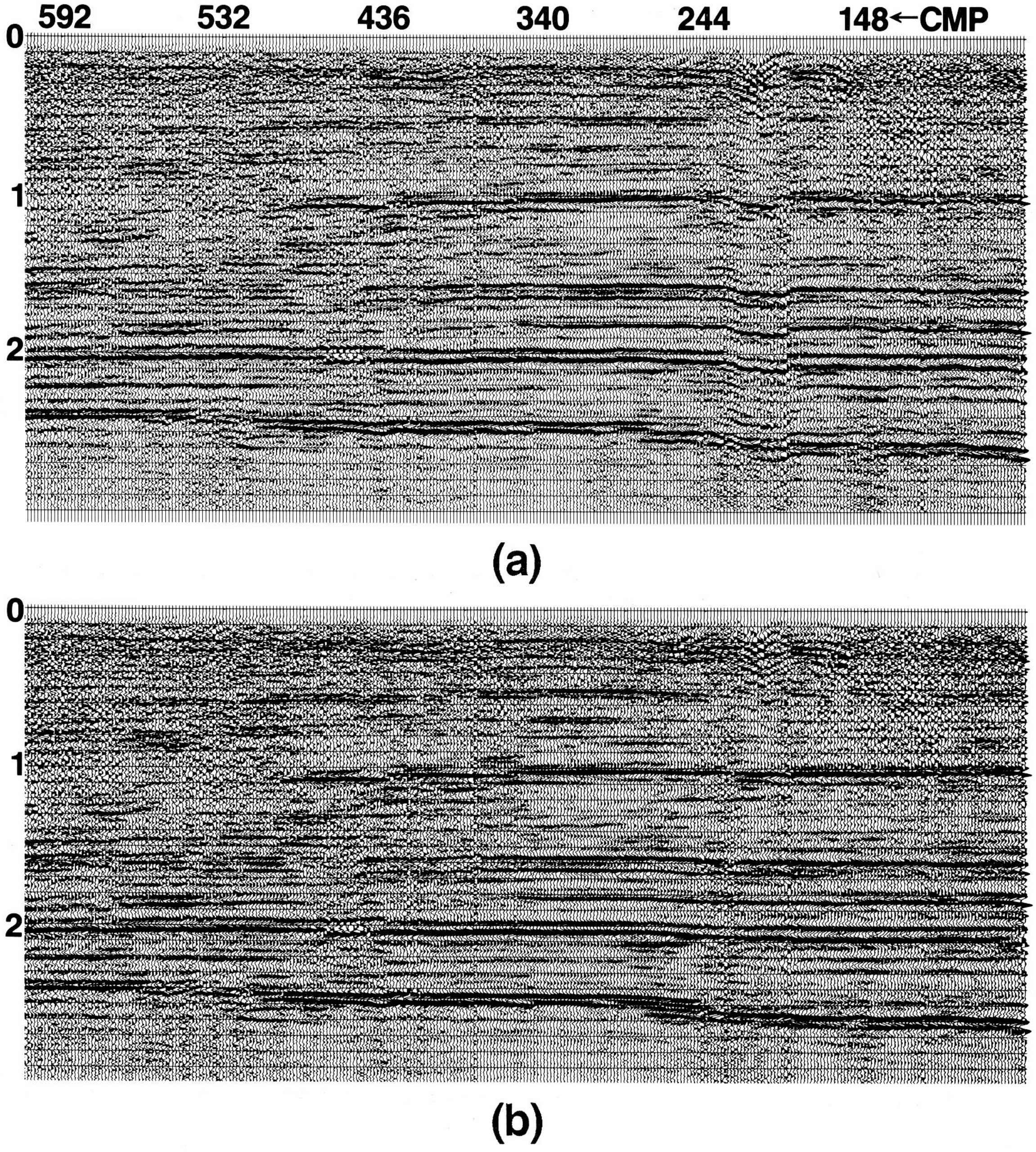


FIG. 3-64. A diagnostic display for residual statics corrections. Common-receiver-point (CRP) stack (a) before and (b) after residual statics corrections. The CRP stack can be used to estimate the magnitude and spatial variation of receiver statics along a line.

gather that has a good S/N ratio.) A stack trace then is constructed within the time gate h. Each individual trace in the gather is crosscorrelated with the stack trace. Time shifts t'_{ijh}, which correspond to maximum crosscorrelations, are picked. A preliminary pilot trace is constructed by stacking the time-shifted traces in the gather. This preliminary pilot trace is, in turn, crosscorrelated with the original traces in the gather and new values for t'_{ijh} are computed. A final pilot trace is constructed again by stacking the original traces shifted by the new values for t'_{ijh}. This final pilot trace is crosscorrelated with the traces of the next gather to construct the preliminary pilot trace for that gather. The process is performed this way on all CMP gathers moving to the left and right from the starting point. The picked time deviations t'_{ijh} are passed on to the next phase, which involves decomposing these time picks into components as defined by equation (3.25).

Several practical issues are involved in the picking phase. Band-pass filtering often helps to estimate a reliable time shift that corresponds to the peak crosscorrelation value. Selection of the time window used for crosscorrelation is another important factor. If needed, the window should

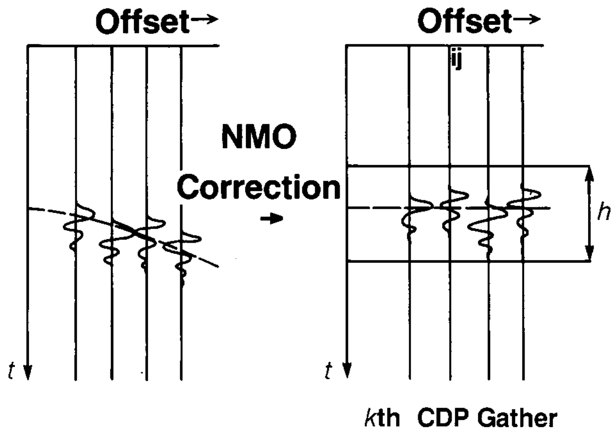


FIG. 3-65. Picking traveltime deviations from NMO-corrected gathers.

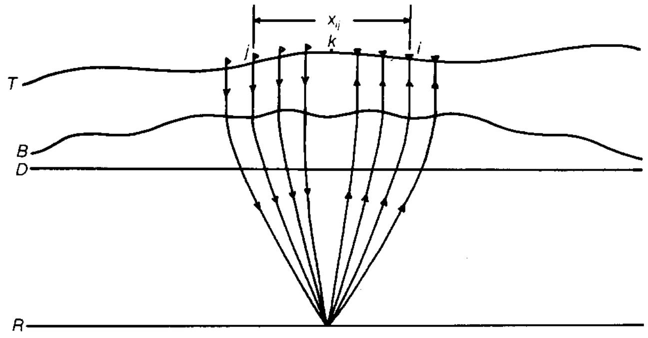


FIG. 3-66. Surface-consistent statics model [refer to equation (3.25)]. Here, T = surface topography, B = base of weathering layer, D = datum to which static corrections are made, R = deep reflector, j = shot station index, i = receiver station index, k = midpoint location index, x_{ij} = offset between the shot and receiver stations.

be allowed to change laterally so that it follows the marker horizon(s). A maximum threshold for the correlation shift can be imposed to prevent unrealistically large time deviations from being passed on to the second phase (decomposition). Any time deviations greater than a specified maximum allowable shift can be set to that shift, or rejected altogether. Alternatively, the rejected shift can be replaced by a secondary correlation peak value. Finally, the input CMP gathers must be NMO-corrected by using a regional velocity function or velocities derived from a preliminary velocity analysis. Detailed examination of these parameters is given in Section 3.5.

The next step in residual statics corrections involves least-squares decomposition of the time picks t'_{ijh}. For a basic understanding of this step, consider the following special problem. Suppose there are four observations t'_i, measured at receiver locations x_i, where i = 1,2,3,4. We want to fit the observed data into a straight line $t(x) = a + bx$, which is best in a least-squares error sense. Start with the following set of equations:

$$\left.\begin{aligned} t_1 &\approx a + b\,x_1 \\ t_2 &\approx a + b\,x_2 \\ t_3 &\approx a + b\,x_3 \\ t_4 &\approx a + b\,x_4 \end{aligned}\right\} . \tag{3.27}$$

There are two unknowns, a and b, and four equations. This problem is similar to decomposing the time picks into various components as described by equation (3.25). We define the error series e_i, such that

$$\left.\begin{aligned} -t_1 + a + b\,x_1 &= e_1 \\ -t_2 + a + b\,x_2 &= e_2 \\ -t_3 + a + b\,x_3 &= e_3 \\ -t_4 + a + b\,x_4 &= e_4 \end{aligned}\right\} . \tag{3.28}$$

We want to minimize the energy of the cumulative errors as defined by

$$E = \sum_i e_i^2 . \tag{3.29}$$

The energy for the ith error is computed by squaring both sides of equation (3.28):

$$e_i^2 = (-t_i + a + bx_i)^2 . \tag{3.30}$$

By substituting into equation (3.29) and summing over i = 1,2,3,4, we get

$$E = \sum_i t_i^2 + 4a^2 + b^2 \sum_i x_i^2 - 2a \sum_i t_i - 2b \sum_i x_i t_i + 2ab \sum_i x_i . \tag{3.31}$$

To find the best line, we want to determine unknowns a and b so that error energy E is minimal. The sum E of the squared errors will attain a minimum if a and b are chosen so that

$$\partial E/\partial a = 8a - 2\sum_i t_i + 2b \sum_i x_i = 0 \tag{3.32a}$$

and

$$\partial E/\partial b = 2b \sum_i x_i^2 - 2 \sum_i x_i t_i + 2a \sum_i x_i = 0. \tag{3.32b}$$

We now have two equations with two unknowns that can be put into a matrix form:

$$\begin{bmatrix} 4 & \sum_i x_i \\ \sum_i x_i & \sum_i x_i^2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_i t_i \\ \sum_i x_i t_i \end{bmatrix} , \tag{3.33}$$

that can be solved for a and b. The minimum energy between the estimated model and the actual data values then can be computed by solving for a and b and substituting into equation (3.31).

Consider the time values at various x locations specified in Table 3-7.

Table 3-7. Time picks t_i' at various x_i locations.

i	x_i	t_i'
1	1	2.4
2	2	2.9
3	3	3.6
4	4	4.1

Using the values in Table 3-7, the elements of the coefficient matrix on the left side and the column matrix on the right side of equation (3.33) are:

$$\begin{bmatrix} 4 & 10 \\ 10 & 30 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 13 \\ 35.4 \end{bmatrix}. \tag{3.34}$$

The solution for equation (3.34) is $a = 1.8$ and $b = 0.58$. Note that we are not constrained by the number of observations when setting up the least-squares problem. This problem involves more equations than unknowns.

The least-squares approach has a wide range of applications in applied geophysics. We also formulate the inverse filtering used in deconvolution based on a least-squares procedure (Appendix B.5). The Wiener filter itself is based on a least-squares estimation of a future sample point in a given time series. Here, we discussed another application of the least-squares method, which involved more equations than unknowns. In Chapter 8, when discussing 2-D surface data processing, we encounter the problem of fitting irregularly spaced observations into a uniform grid. This involves least-squares fitting to a local plane (Appendix F).

We now return to the problem of residual statics estimation and refer to equation (3.26). Substitute for t_{ijh} from equation (3.25) and minimize the error energy E by requiring

$$\partial E/\partial s_j = \partial E/\partial r_i = \partial E/\partial G_k = \partial E/\partial M_k = 0, \tag{3.35}$$

which yields $(NS + NR + NG + NG)$ equations, and that many unknowns.

These equations can be solved for the residual statics associated with NS source locations, NR receiver locations, NG structural terms, and NG residual moveout terms. In seismic data processing, the number of these terms can be large. It is easy to solve the problem with the two parameters given by equation (3.34). However, when dealing with a large number of linear equations, the solution must be done accurately and efficiently. A number of solution techniques are in use. Wiggins et al. (1976) used the Gauss-Seidel iterative procedure to solve for equations (3.35). This procedure is best described by returning to the earlier line fitting example and solving for a and b in equation (3.34). When written as normal equations,

$$4a + 10b = 13 \tag{3.36a}$$

and

$$10a + 30b = 35.4. \tag{3.36b}$$

These equations are rearranged as

$$a = 3.25 - 2.500b \tag{3.37a}$$

and

$$b = 1.18 - 0.333a. \tag{3.37b}$$

Since the Gauss-Seidel technique is iterative, starting values are needed. To start the iteration, set $a = b = 0$. Substituting $b = 0$ in equation (3.37a), we obtain $a = 3.25$. Putting this updated solution into equation (3.37b) gives $b = 0.0977$. At the end of the first iteration, $a = 3.25$ and $b = 0.0977$. For the second iteration, put $b = 0.0977$ into equation (3.37a) to get $a = 3.0075$. Put $a = 3.0075$ into equation (3.37b) to get $b = 0.1785$. This iterative procedure is continued in Table 3-8.

The solution from the iterative procedure slowly converges toward the actual values, $a = 1.8$, $b = 0.58$. Convergence is not always guaranteed. Nevertheless, convergence can be attained with the Gauss-Seidel method, provided the unknowns are ordered properly; i.e., iteration starts with the correct unknown. This problem is addressed in Exercises 3.18 and 3.19. The advantage of the Gauss-Seidel method is its ability to solve the large number of simultaneous equations rapidly.

The question of when to terminate the iterations remains. The rate at which the solution changes can be examined after each iteration. The computation is stopped when this rate falls below a specified threshold value.

The starting values for solving the normal equations, equations (3.35), can be chosen as $s_j = r_i = G_{kh} = M_{kh} = 0$, as in the simple numerical example. To some extent, the best order of iteration depends on the exact nature of the statics problem at hand. One order of iteration that commonly is in use follows: Compute the structural term G, the residual moveout term M, the static shift associated with source location s, then the static shift associated with receiver location r. The procedure cycles back to G in the next iteration and continues on until convergence is satisfactory. The order in which the individual terms are computed theoretically can be interchanged. However, the above order forces long wavelength variations of the time picks to be concentrated mainly in the structural term. This leads to a lesser number of iterations (typically 2 or 3) for wavelength components of the statics that are less than half the largest data offset. A large number of iterations are required to handle static variations that are

Table 3-8. Gauss-Seidel iteration for solving equations (3.34) for *a* and *b*. (Values were rounded off for tabulation.)

Actual Values: $a = 1.8$, $b = 0.58$.

Iteration	a	b
1	3.25	0.0977
5	2.4918	0.3502
10	2.0712	0.4902
15	1.9031	05462
20	1.8358	0.5686
25	1.8090	0.5779
29	1.7997	0.5807

greater than the maximum offset.

After computing the individual static shifts associated with each source and receiver location, these shifts are passed on to the next phase, where the shifts are applied to the pre-NMO-corrected gather traces. In areas with extremely poor S/N ratio or complicated near-surface variations, multiple passes of residual statics corrections may be necessary. In other words, output from the first pass may be NMO-corrected, new picks estimated, decomposed, and applied, and so on.

The question remains as to whether equations (3.25) are independent of each other. This is important, for it tells whether the solution is unique or not. For simplicity, consider the case of zero structure and zero residual moveout. In that case, equations (3.25) take the form:

$$t_{ij} = s_j + r_i. \tag{3.38}$$

Consider only four unknowns: s_j, where $j = 1,2$, and r_i, where $i = 1,2$, which yield four equations:

$$\left.\begin{aligned} t_{11} &= s_1 + r_1 \\ t_{12} &= s_2 + r_1 \\ t_{21} &= s_1 + r_2 \\ t_{22} &= s_2 + r_2 \end{aligned}\right\}. \tag{3.39}$$

However, from a close examination of equations (3.39), we note that

$$t_{11} + t_{22} = t_{12} + t_{21}.$$

Therefore, one of the equations in equations (3.39) is redundant, leaving three independent equations for four unknowns. This simple exercise shows that the statics solution obtained by solving equations (3.35) suffers from an uncertainty. In particular, the solution obtained by, say, the Gauss-Seidel iteration, is nonunique. It is but one of many possible solutions. In fact, the provided solution may not even be physically reasonable. Because of the problem of fewer *independent* equations than unknowns, Gulunay (1985) suggests that a constraint be imposed on the statics problem. A plausible constraint is that the difference between shot and receiver statics be minimal. It can be argued that this may not be valid, as is the case with dynamite data in which shots and receivers do not occupy the same physical locations. Other possible constraints can be in the form of restrictions on the spatial variation of structure, moveout, or statics terms themselves; all are used in various practical implementations.

3.5 RESIDUAL STATICS CORRECTIONS IN PRACTICE

Residual statics corrections consist of three stages: picking, decomposing, and applying residual static shifts. The picking phase determines the effectiveness of residual statics corrections. How reliable are the t'_{ijh} values that are derived from crosscorrelations? This depends on the choice of crosscorrelation window and signal quality. We now examine various parameters involved in the picking phase.

3.5.1 Maximum Allowable Shift

Consider the synthetic model in Figure 3-67. This data set was created by using the field geometry of a real seismic line. The CMP traces were derived from the first trace of the first CMP location from that real line. This trace first was zeroed out within selected time gates, then treated with the shot and receiver static shifts in Figure 3-67 in a surface-consistent manner and with a structure term that depended on midpoint location only (subsurface-consistent). The shot and receiver static shifts were varied from +32 to −32 ms. Finally, the synthetic traces were blended with a band-limited random noise, whose strength varied spatially. (The noise level was set to zero at both ends of the profile and to maximum at the center.) The stacked section constructed from data before treatment with synthetic shot and receiver static shifts is shown in Figure 3-68. Once residual statics corrections are made, the stacked section in Figure 3-67 should resemble that in Figure 3-68. Note how degrading the effect of the synthetic shot and receiver static shifts is on the continuity of reflections in Figure 3-67. The sole effect of random noise is shown in Figure 3-68.

Consider three different tests of residual statics corrections: one with a small maximum allowable shift (24 ms in Figure 3-69), one with a moderately sized shift (80 ms in Figure 3-70), and one with a fairly large shift (192 ms in Figure 3-71). All three tests had the same input CMP gathers. All three were run using the same set of parameters except for maximum allowable shift. The maximum value of combined shot and receiver static shifts for any given trace implied by the model in Figure 3-67 is ±64 ms. When the maximum allowable shift is insufficient (a value of less than 64 ms), then the derived static shifts (Figure 3-69) are significantly smaller than the actual shifts (Figure 3-67). Thus, stack quality, although significantly improved when compared to that in Figure 3-67, is far from the quality of the section shown in Figure 3-68.

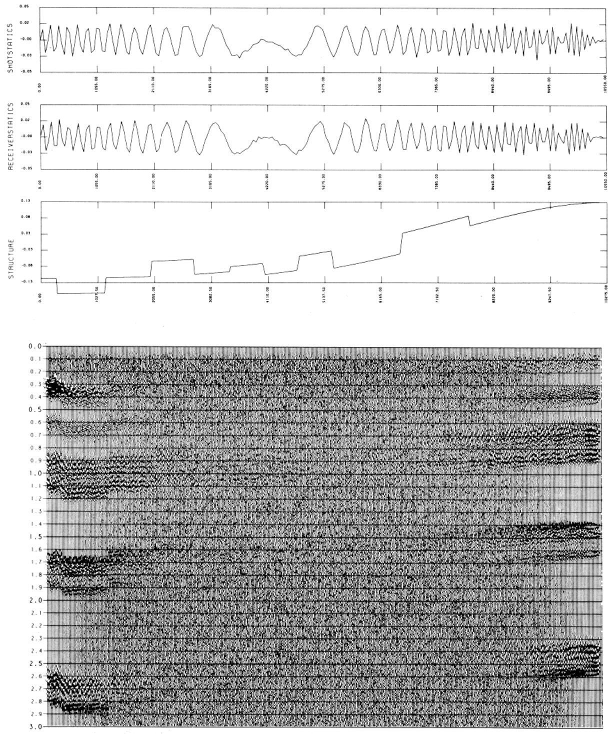


FIG. 3-67. A synthetic data set. Shot and receiver statics were applied on moveout-corrected gathers in a surface-consistent manner, while the structure term was applied in a subsurface-consistent manner. These terms are plotted above the stacked section. Random noise was added to prestack data with a spatially varying S/N ratio.

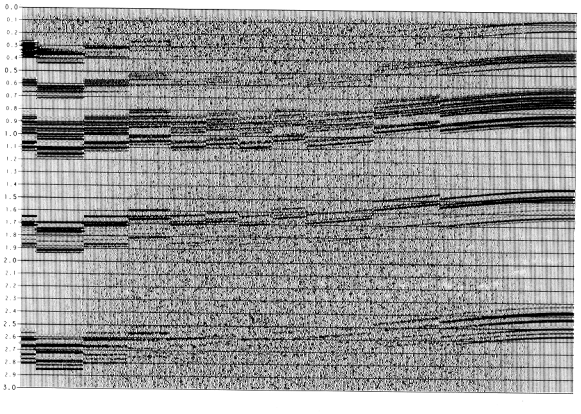


FIG. 3-68. The difference between Model 1 (shown here) and Model 2 (in Figure 3-67) is that Model 1 has not had synthetic shot or receiver statics applied.

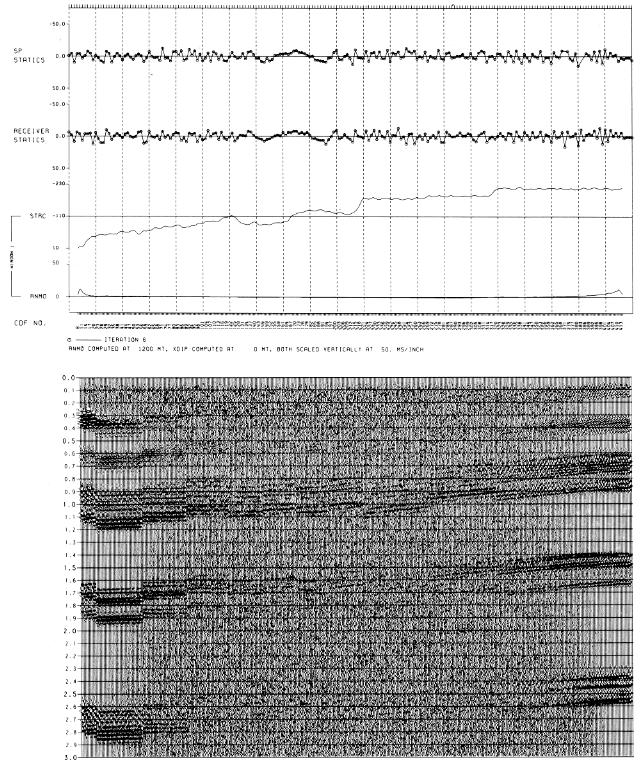


FIG. 3-69. Model 2 stack after residual statics corrections. The derived shot, receiver, and structure terms are plotted at top. Compare these estimates with the actual values in Figure 3-67. The maximum allowable shift is 24 ms.

When the maximum allowable shift is sufficient, then the derived static shifts (Figure 3-70) are like the actual shifts imposed on the input model shown on the graph in Figure 3-67. Also, stacking quality (Figure 3-70) improved so that it now is comparable to the no-static model (Figure 3-68). Reasonable results (Figure 3-71) also were obtained for the case allowing excessively large maximum allowable shift, up to 192 ms. However, this result does not imply that we can be liberal on the upper bound of maximum shift in real data situations. In the presence of short-period multiple or reverberation energy, or data with narrow bandwidth or high noise level, crosscorrelation can yield a multiple number of peaks and cause uncertainty in the estimated time shifts (cycle skipping). In this case, large maximum allowable shift could cause anomalously large time shifts to be picked. Based on the tests shown in Figures 3-68 through 3-71, the maximum allowable shift used in the picking phase should be

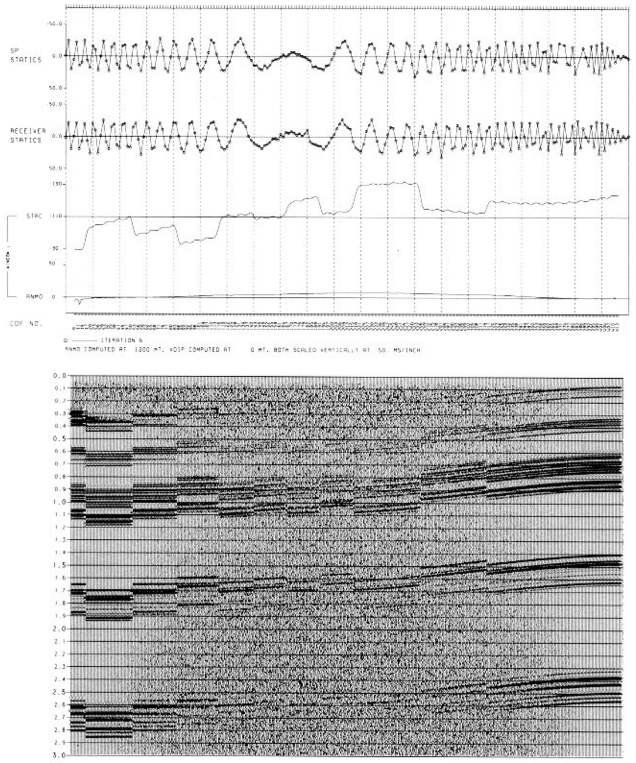


FIG. 3-70. Model 2 stack after residual statics corrections. The derived shot, receiver, and structure terms are plotted at top. Compare these estimates with the actual values in Figure 3-67. The maximum allowable shift is 80 ms.

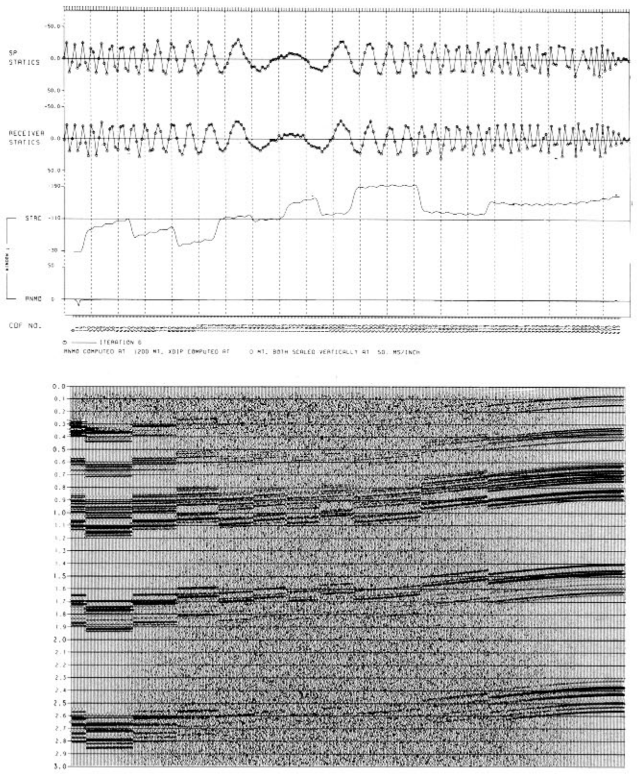


FIG. 3-71. Model 2 stack after residual statics corrections. The derived shot, receiver, and structure terms are plotted at top. Compare these estimates with the actual values in Figure 3-67. The maximum allowable shift is 192 ms.

greater than all possible combined shot and receiver static shifts at any given location along the profile. On the other hand, cycle skips, especially in poor S/N ratio conditions, also are more likely to occur if the maximum allowable shift is greater than the dominant period of the data.

We may argue that the result of cascading a number of small-shift residual statics solutions is as good as a single-step large-shift solution. This approach might have the effectiveness of the large-shift solution, while avoiding the possibility of cycle skipping. Unfortunately, cascading small-shift solutions does not work. Starting with the CMP gathers associated with the stack in Figure 3-67, we get the CMP gathers corrected for shot and receiver statics based on a 24-ms shift (first pass). The stack is shown in Figure 3-69. Using these gathers, a new statics solution was derived and applied to the data (second pass). This process was repeated for the third and fourth times. The result of this last iteration (Figure 3-72) does not have the quality of the solution derived with the 80-ms shift (Figure 3-70).

Now consider maximum allowable shift on the field data shown in Figures 3-49 through 3-64. Figure 3-73 shows CMP gathers from the problem zone of the profile in Figure 3-53. Insufficient maximum allowable shift does not completely correct for all static shifts. (Refer to the panels for 24- and 40-ms shifts in Figure 3-73.) Excessively large shifts (such as 120- or 160-ms shifts) do not seem to harm this particular data set. While common-shotpoint (CSP) stacks (Figure 3-74) indicate small shot-static shifts, common-receiver-point (CRP) stacks (Figure 3-75) indicate a zone of significant receiver-static shifts. Again, small maximum allowable shifts have not corrected completely for these statics anomalies. The ultimate judgment is made by examining the stack response and plots of the estimated statics themselves. From Figure 3-76 (ungained stack responses), it is clear that the maximum allowable shift must be adequate to accommodate the combined shot and receiver statics present in the data at any location along the profile.

3.5.2 Correlation Window

From the same field data example (Figure 3-53), the results of residual statics corrections (for the right half of the stacked section) are examined by using different correlation windows while keeping all other parameters constant. The maximum allowable shift was 80 ms in these tests. From Figure 3-77, note that a correlation window confined to the mute zone (400 to 1200 ms) is not desirable. It does not provide sufficient statistics because of the low fold of coverage and the shortness of data available for crosscorrelation with the pilot traces. With high-fold data, the mute zone problem is handled to a degree by limiting the correlation to small offsets. In this particular part of the profile, a large window including both the mute zone and deep data (800 to 2300 ms), a deep window (1700 to 2300 ms), a deep large window (1400 to 2800 ms), or a deep narrow window (1500 to 1700 ms) made no difference. This is probably because of the good S/N ratio in this part of the profile. These observations are verified by the diagnostics in Figures 3-78 and 3-79. Ungained stack responses are shown in Figure 3-80. In particular, note the relatively poor stack response using a window confined to the mute zone (400 to 1200 ms).

Choice of the correlation window is more critical in conditions of poor S/N ratio. Refer to the same diagnostics for the left half of the stacked section in Figure 3-53. These diagnostics are shown in Figures 3-81 through 3-83. Again, a correlation window confined to the mute zone not only provides an inadequate solution, as in the previous case (Figure 3-77), but also can be devastating, as shown in Figure 3-81. In this case, the CMP gathers with no corrections have better signal quality. It now is apparent that a narrow window with weak signal, even if it is outside the mute zone (such as the 1500- to 1700-ms window), may not provide sufficient statistics. The CSP stack (Figure 3-82) and the CRP stack (Figure 3-83) show the undesirable aspects of choosing a window within the mute zone or of choosing a window that is too narrow. The ungained stacked sections (Figure 3-84) clearly demonstrate the adverse effects of an improper choice of correlation window. Moreover, note the poor quality pilot traces for each CMP gather in Figure 3-85 (to the left of midpoint 377). The shot and receiver static solutions shown in the graphs above the pilot traces are totally unreliable. These test results suggest choosing a correlation window that (a) contains as much signal as possible to improve correlation values, and (b) is large enough and is outside the mute zone whenever possible.

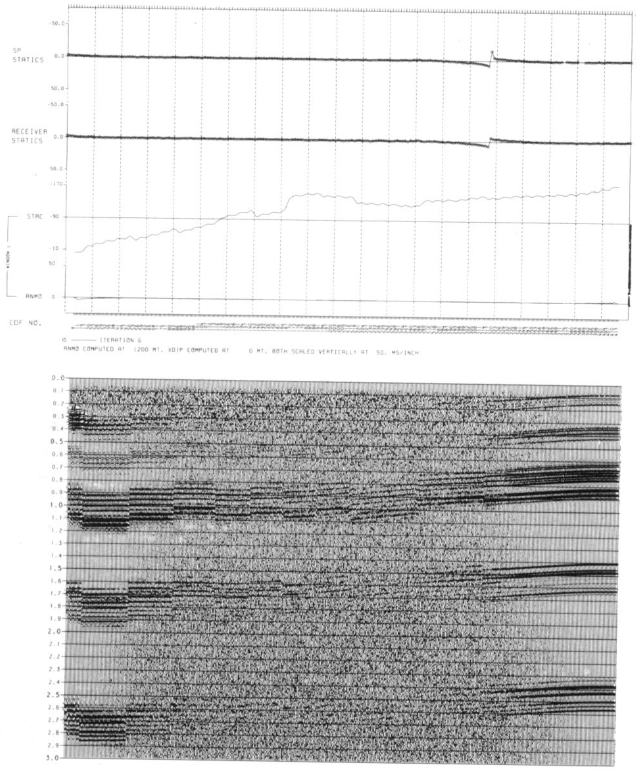


FIG. 3-72. Model 2 stack after residual statics corrections. The derived shot, receiver, and structure terms are plotted at top. This is the output from the four iterative passes of the statics estimation and application. First-pass results are in Figure 3-69. The maximum allowable shift is 24 ms on each pass.

3.5.3 Other Considerations

The synthetic data set in Figure 3-86 is the same as that in Figure 3-67, with the addition of residual moveout shifts. Residual moveout was introduced into the data by inverse NMO correcting the CMP gathers associated with the stack in Figure 3-67 using a v_1 (t), then by NMO correcting using a v_2 $(t) \neq v_1$ (t). The solution (Figure 3-87) implies some traveltime distortion at the edges of the stacked section due to low fold of coverage. (Compare this with the solution in Figure 3-70). Otherwise, stack response seems to be satisfactory. As long as the residual moveout variations are not large within the correlation window, the computed residual statics solution should be adequate. Use of more than one small correlation window during the picking phase may help minimize the effect of time-dependent residual moveout.

In some areas, the S/N ratio is so poor that a second pass of residual statics corrections must be done. The idea is that the first pass of residual statics corrections improves the signal to such a degree that a second pass should remove the residuals remaining from the first pass. For

(text continued on page 220)

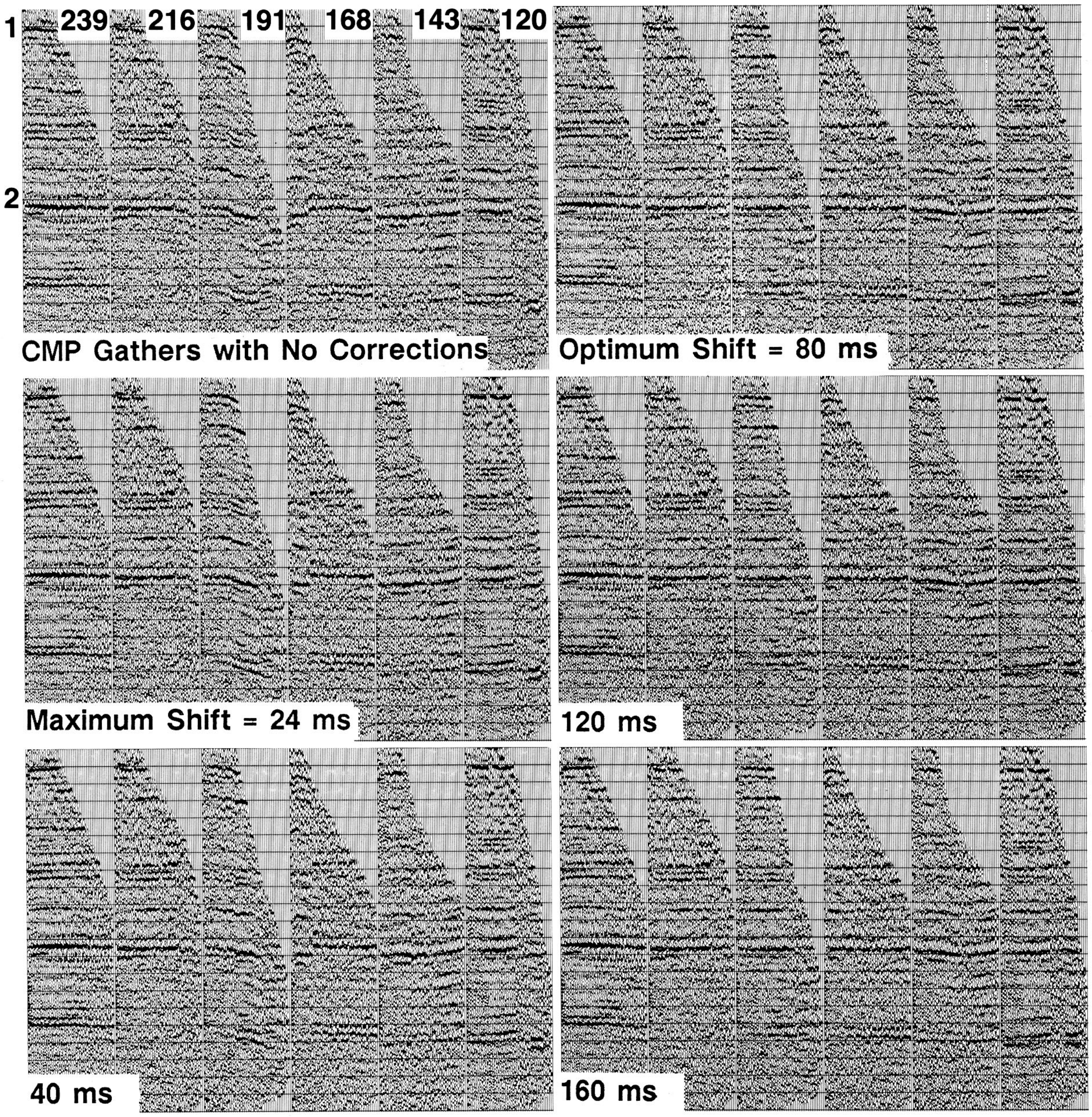


FIG. 3-73. Test of maximum allowable shift. CMP gathers after residual statics corrections using five different maximum allowable shifts. Figure 3-76 shows the CMP stacks.

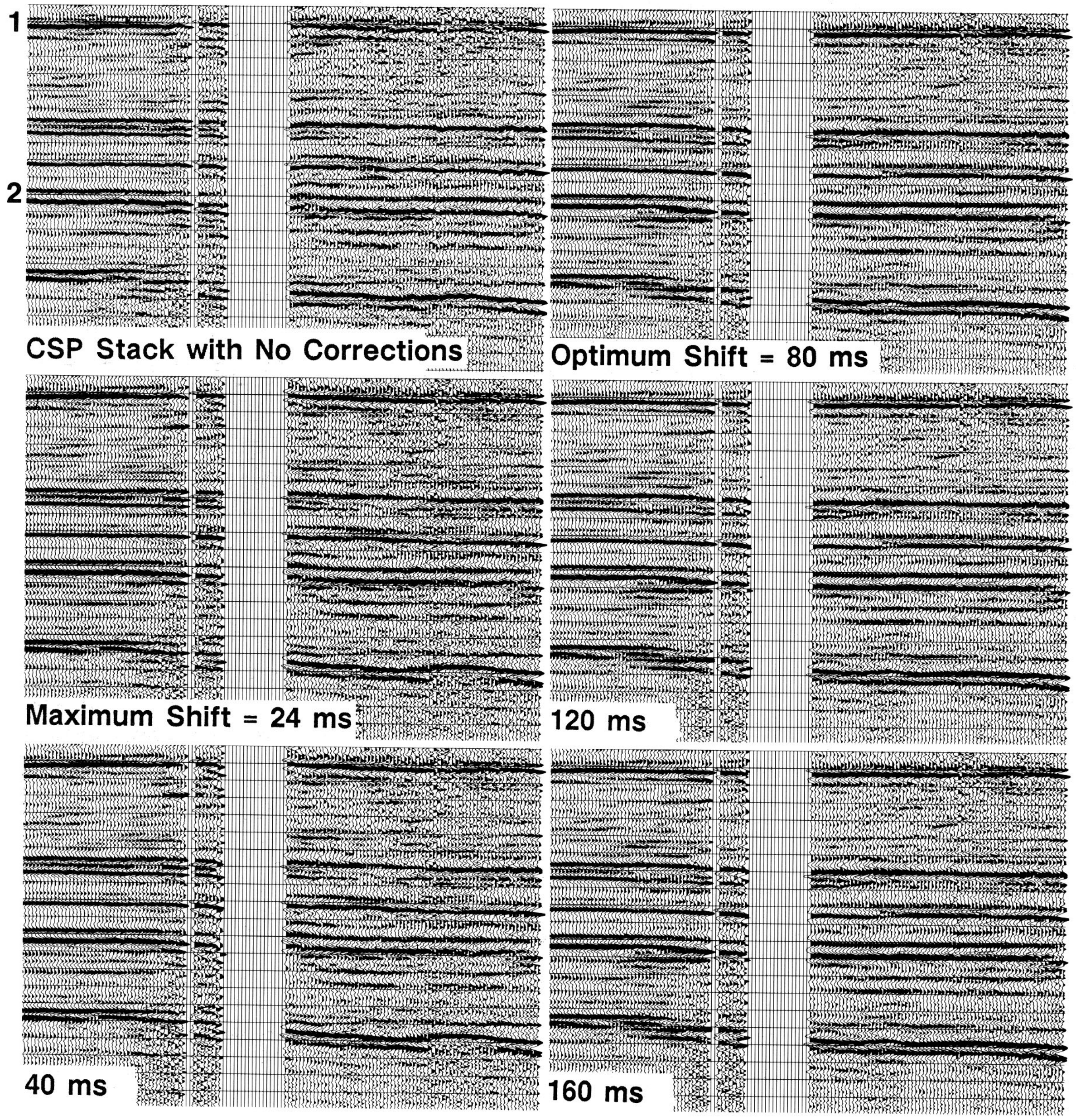


FIG. 3-74. Diagnostics for maximum allowable shift tests (Figure 3-73) showing CSP stacked data.

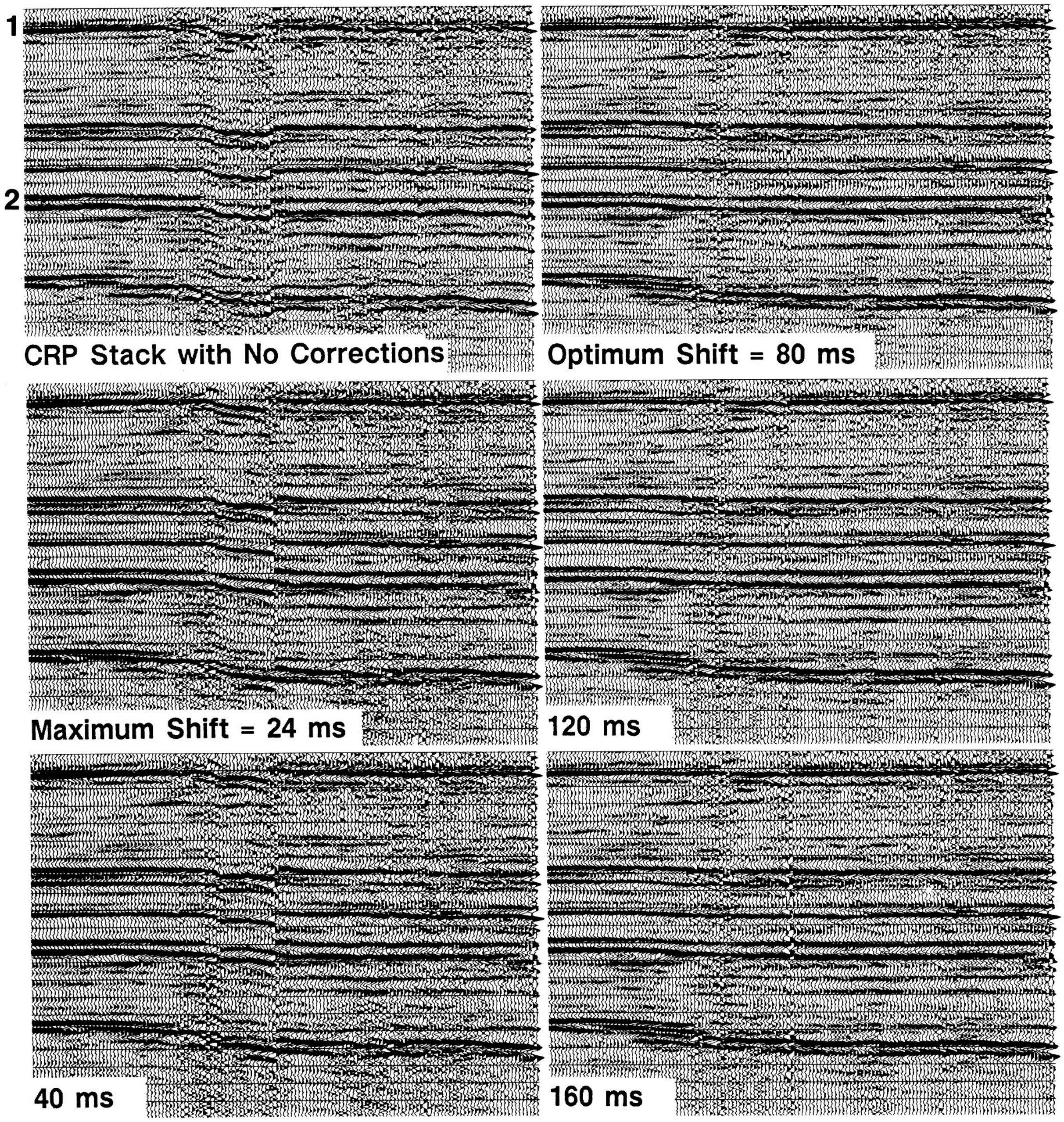


FIG. 3-75. Diagnostics for maximum allowable shift tests (Figure 3-73) showing CRP stacked data.

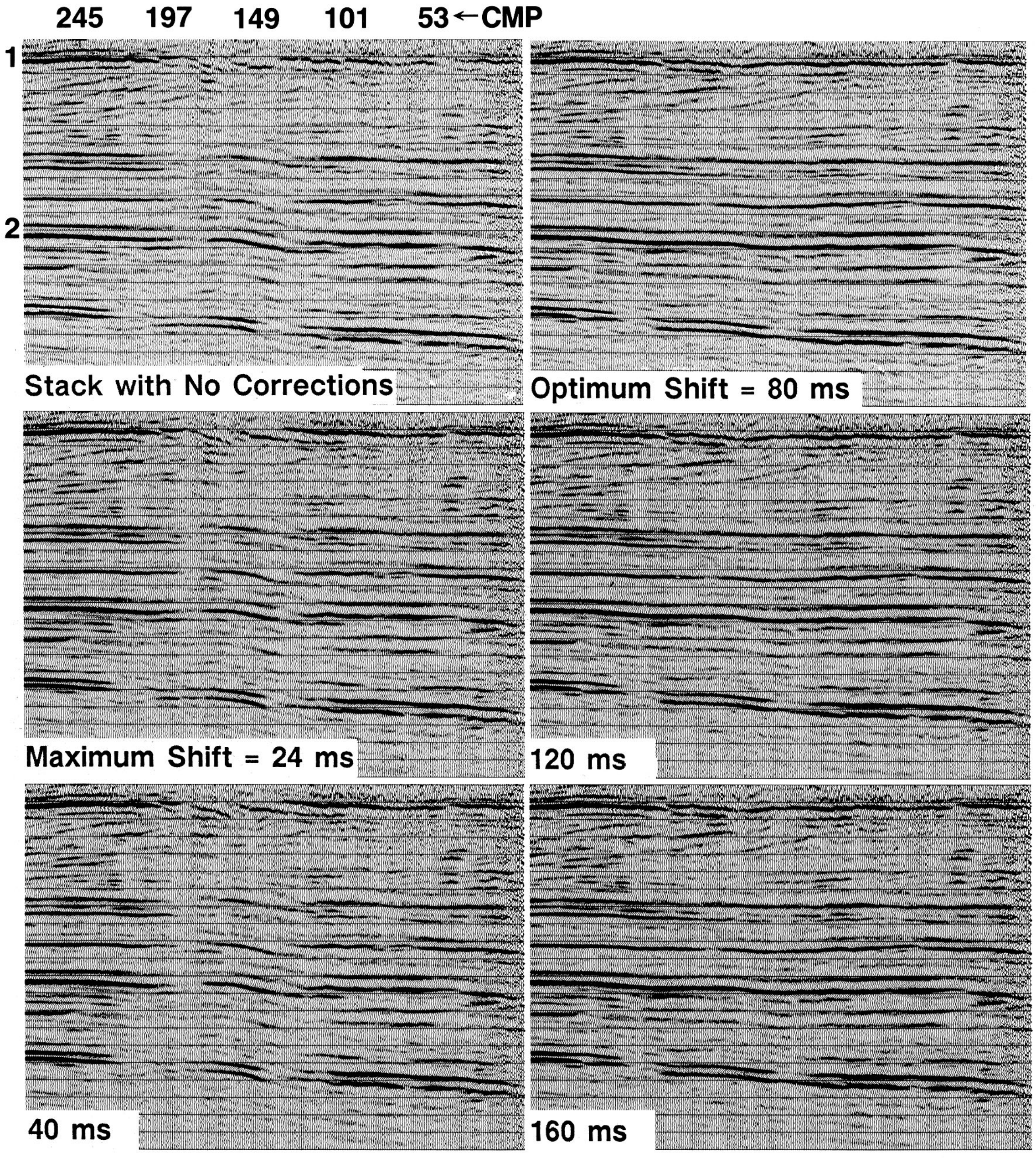


FIG. 3-76. Test of maximum allowable shift: CMP stacks after residual statics corrections using five different maximum allowable shifts. Figure 3-73 shows CMP gathers.

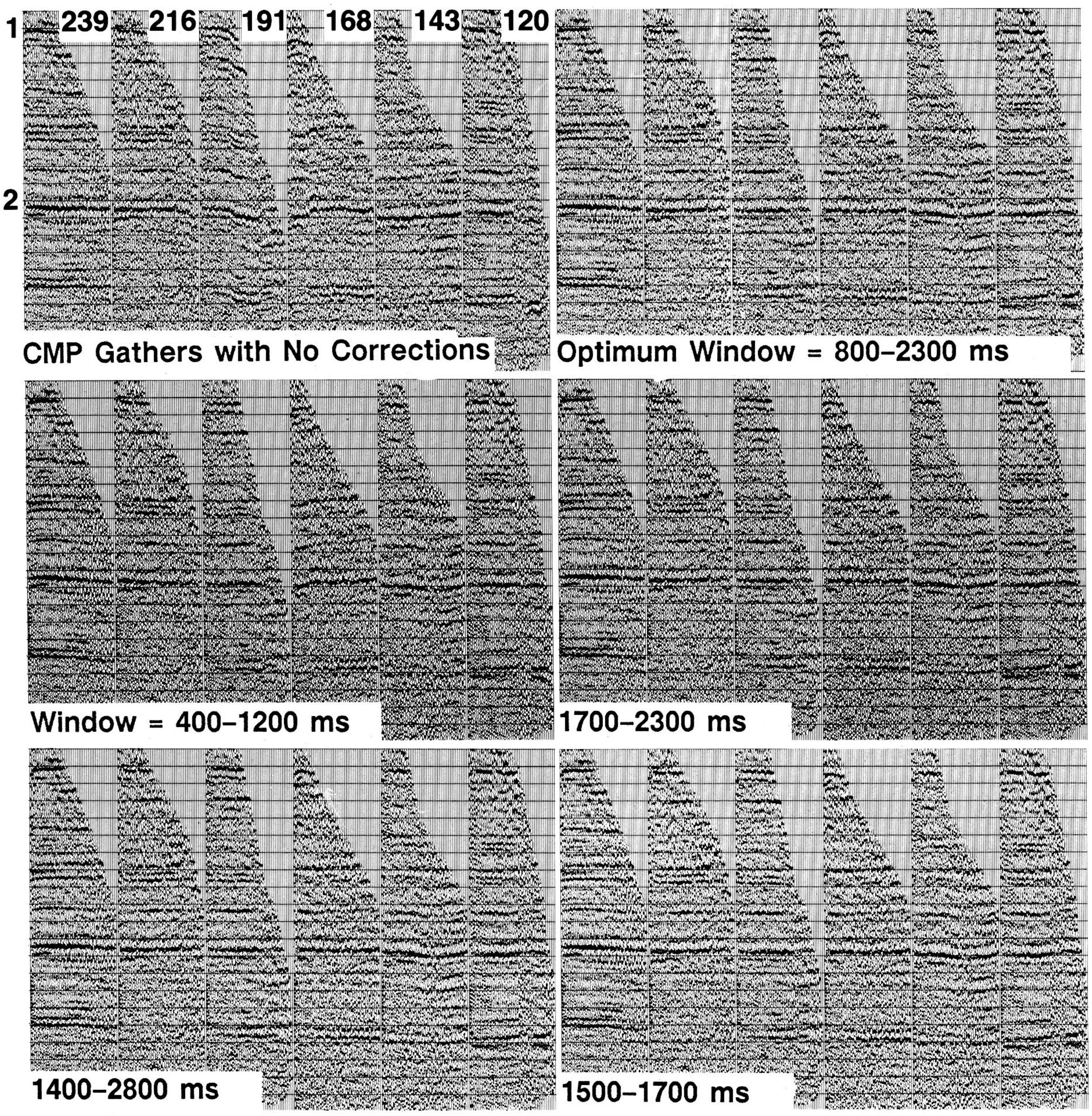


FIG. 3-77. Test of correlation window: CMP gathers after residual statics corrections using five different correlation windows. Figure 3-80 shows the CMP stacks.

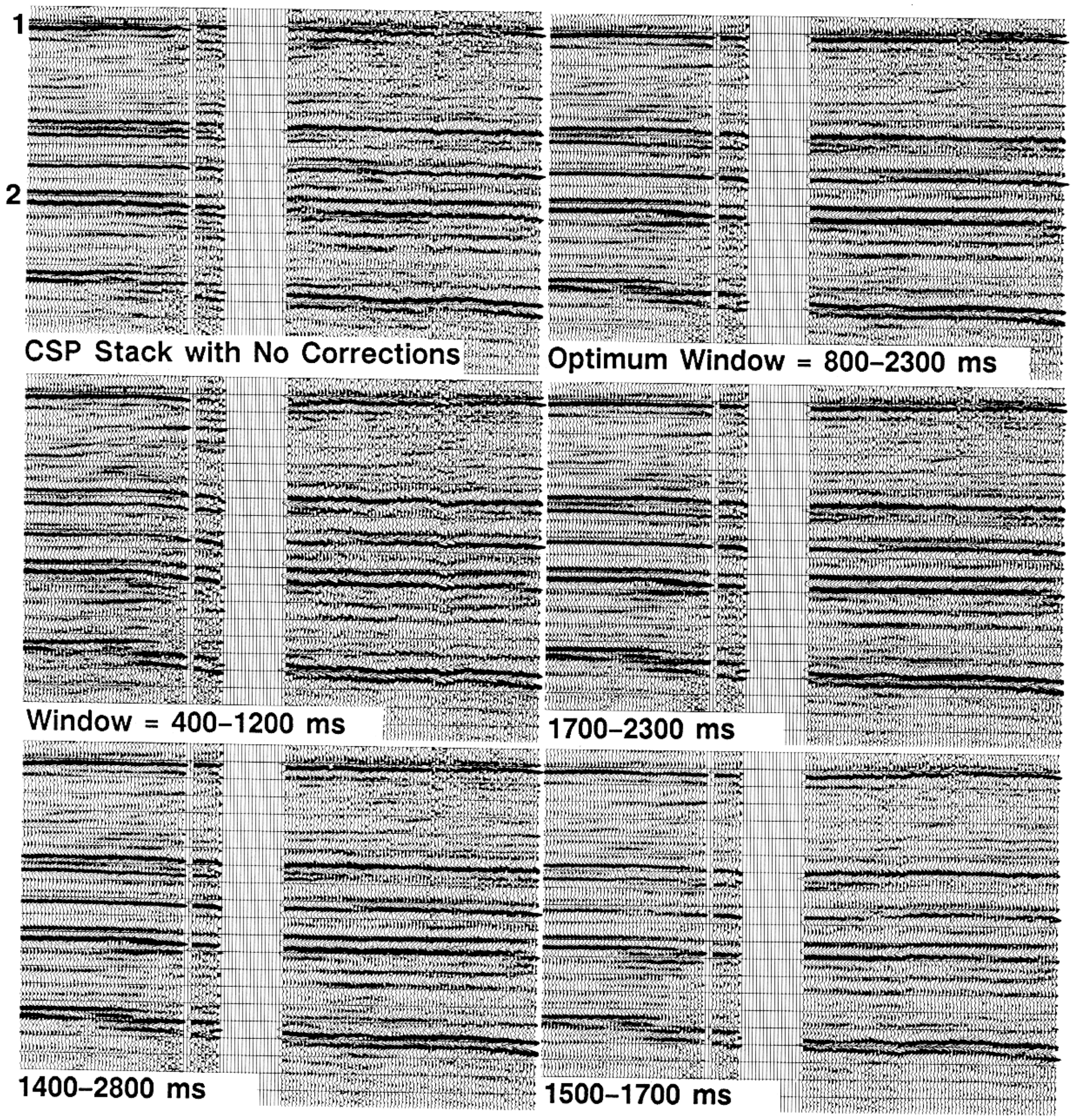


FIG. 3-78. Diagnostics for correlation window tests (Figure 3-77) showing CSP stacked data.

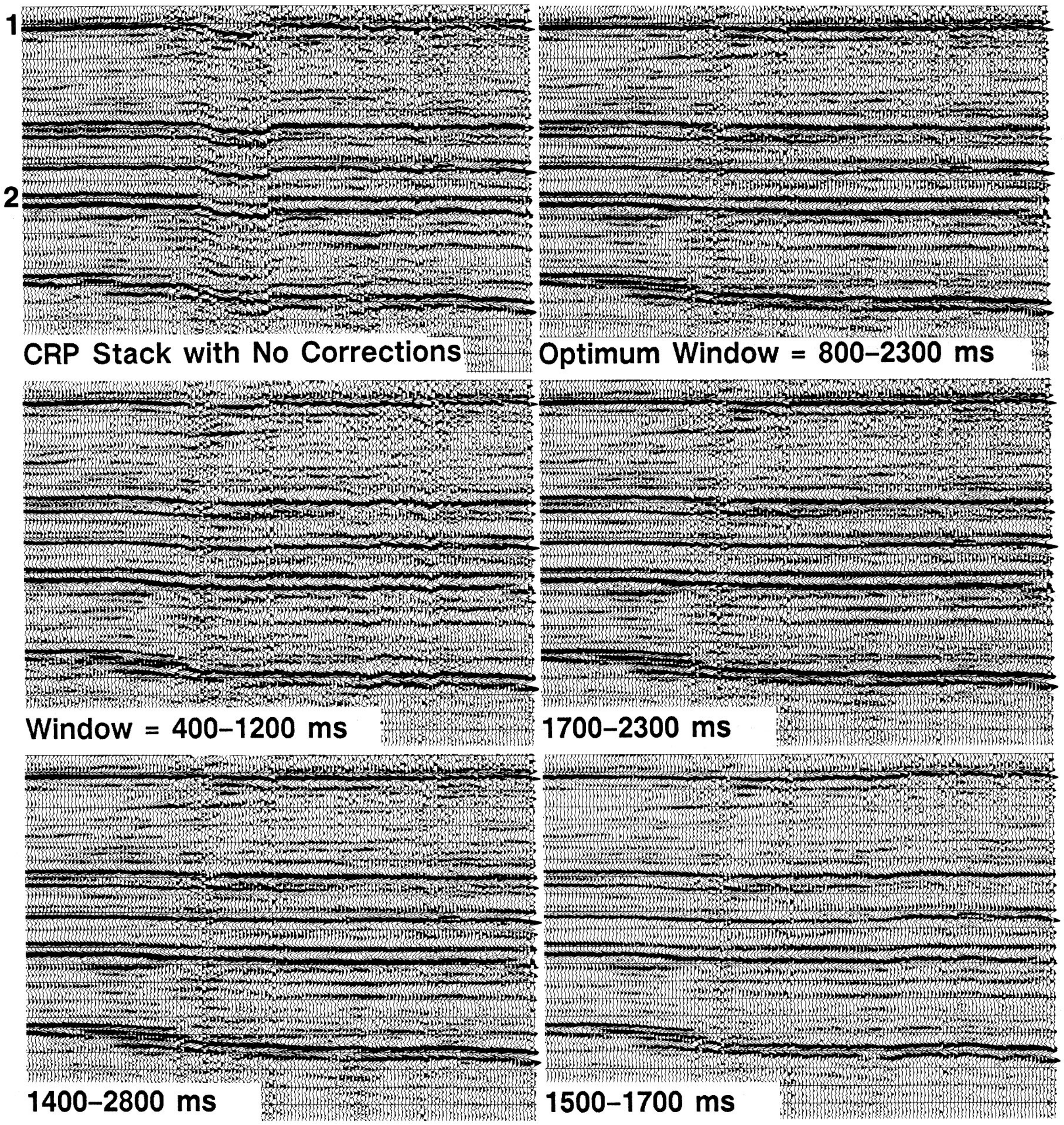


FIG. 3-79. Diagnostics for correlation window tests (Figure 3-77) showing CRP stacked data.

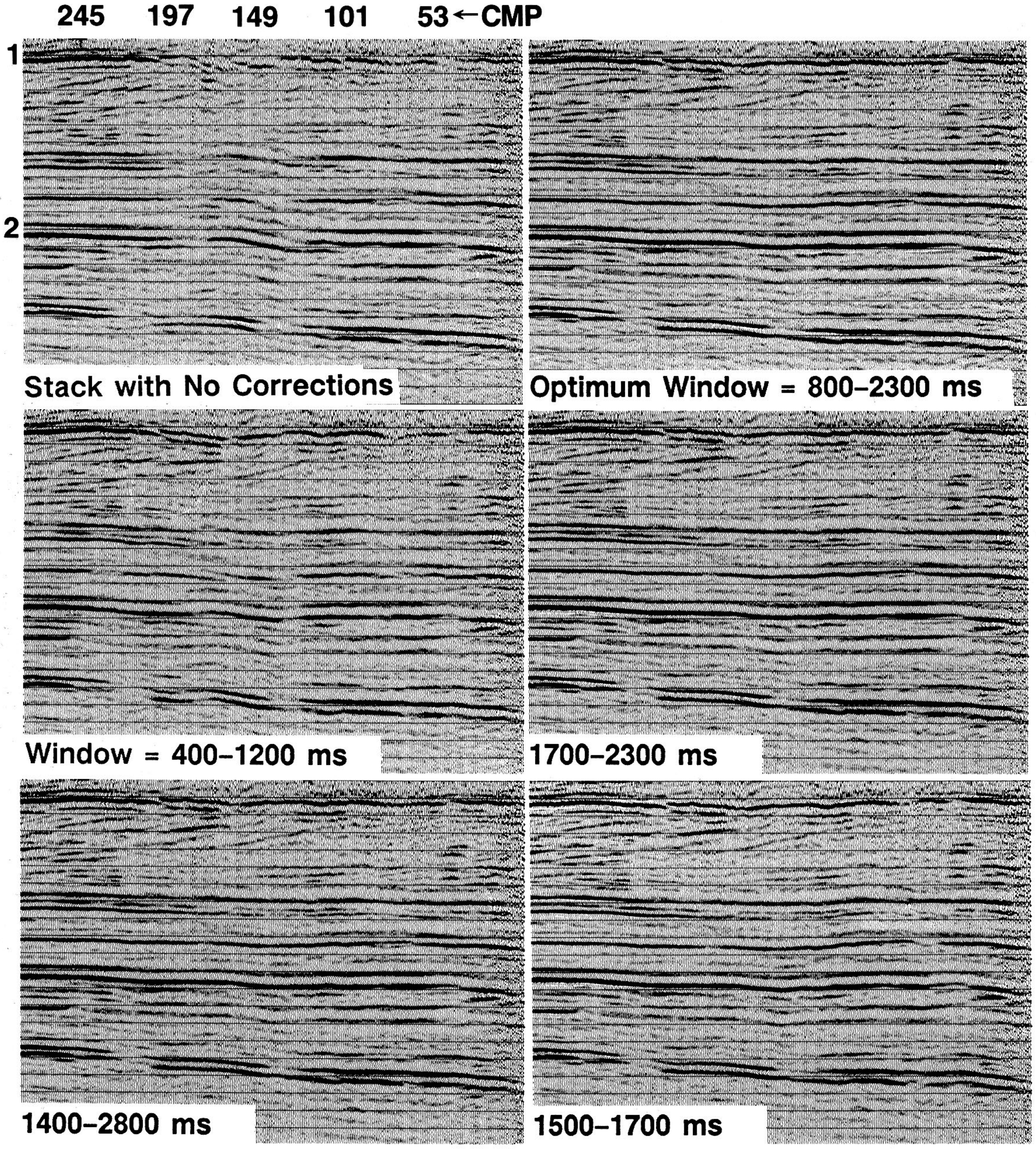


FIG. 3-80. Test of correlation window: CMP stacks after residual statics corrections using five different correlation windows. CMP gathers are shown in Figure 3-77.

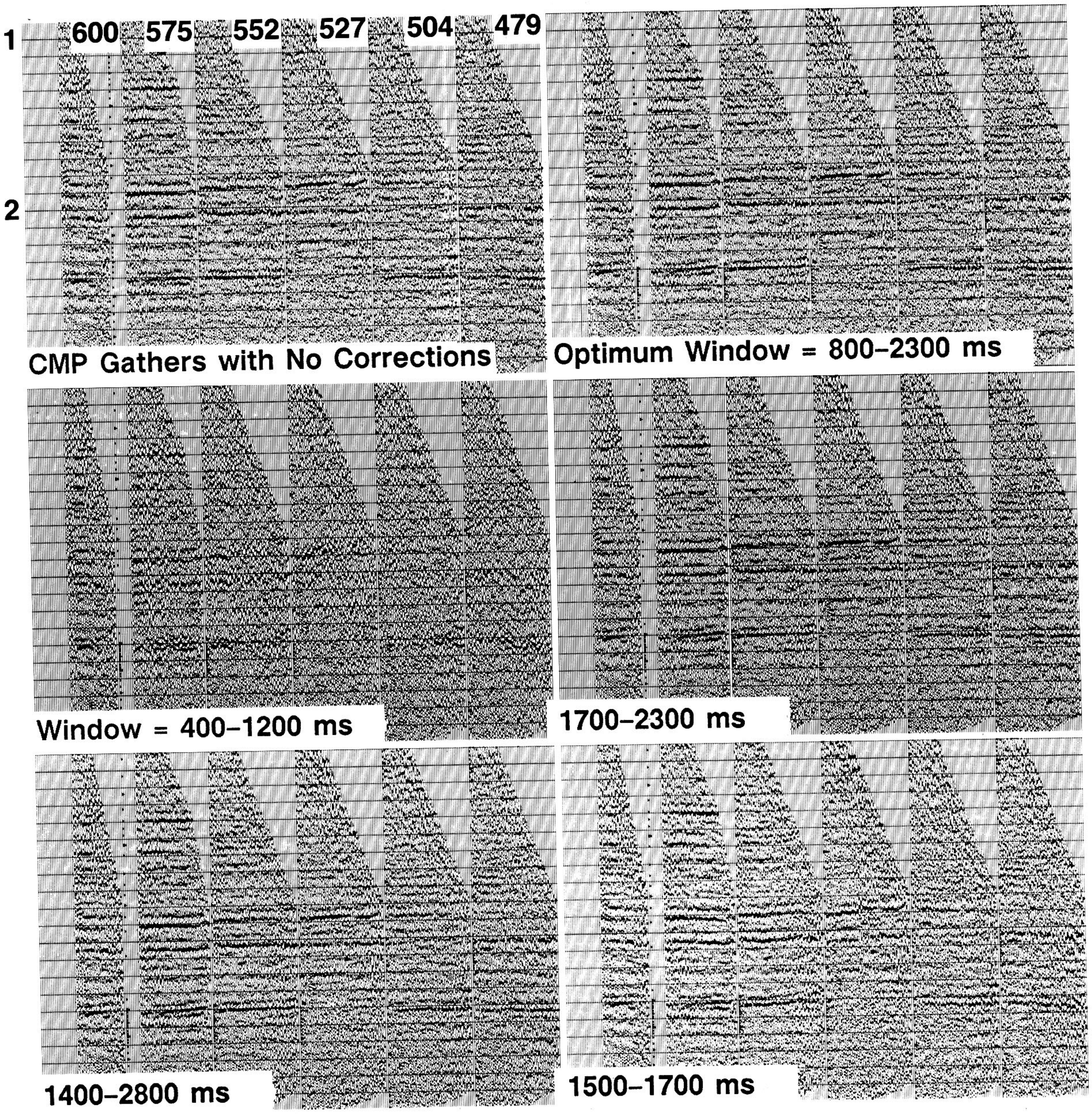


FIG. 3-81. Test of correlation window: CMP gathers after residual statics corrections using five different correlation windows. CMP stacks are shown in Figure 3-84.

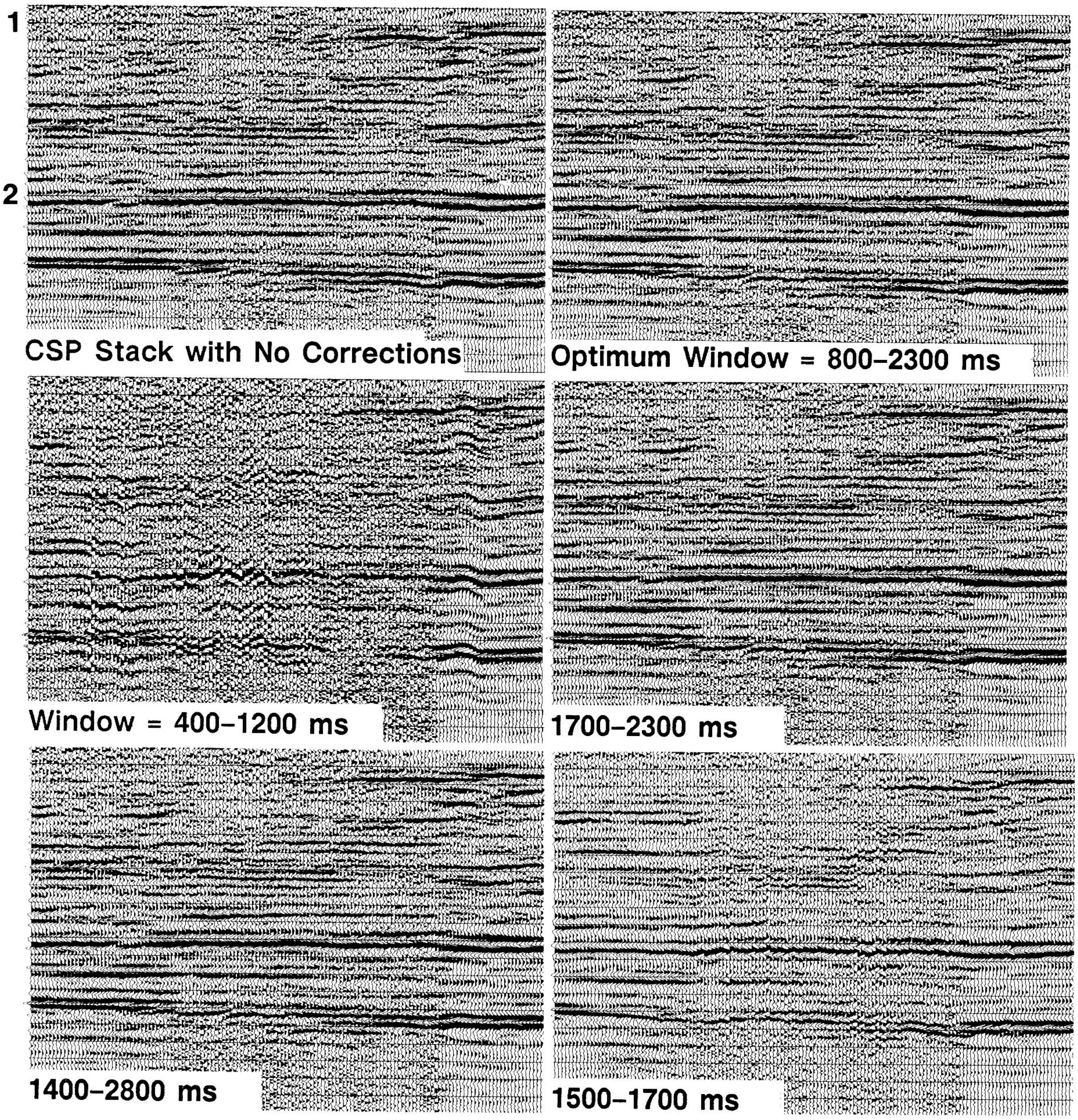


FIG. 3-82. Diagnostics for correlation window tests (Figure 3-81) showing CSP stacked data.

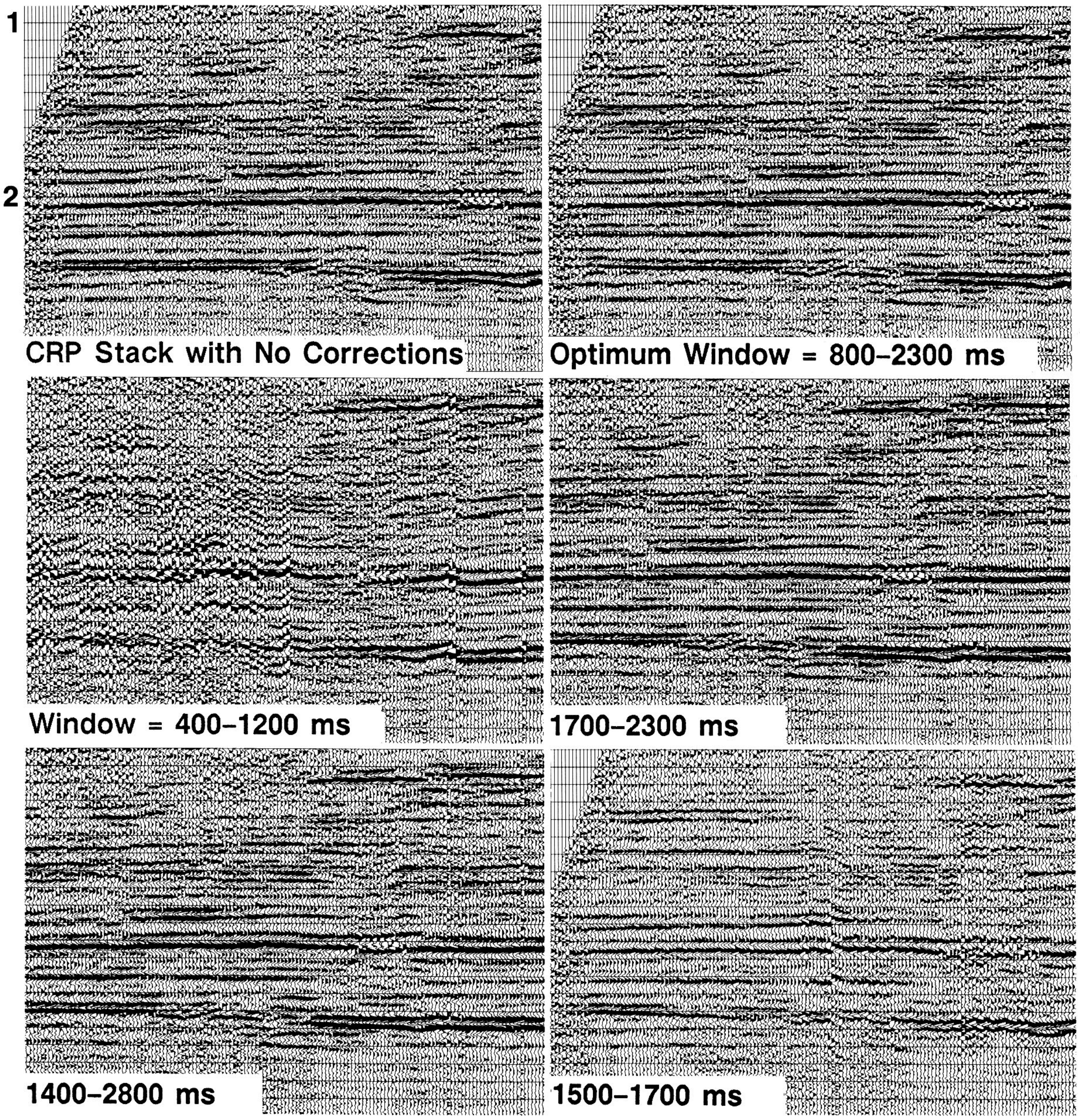
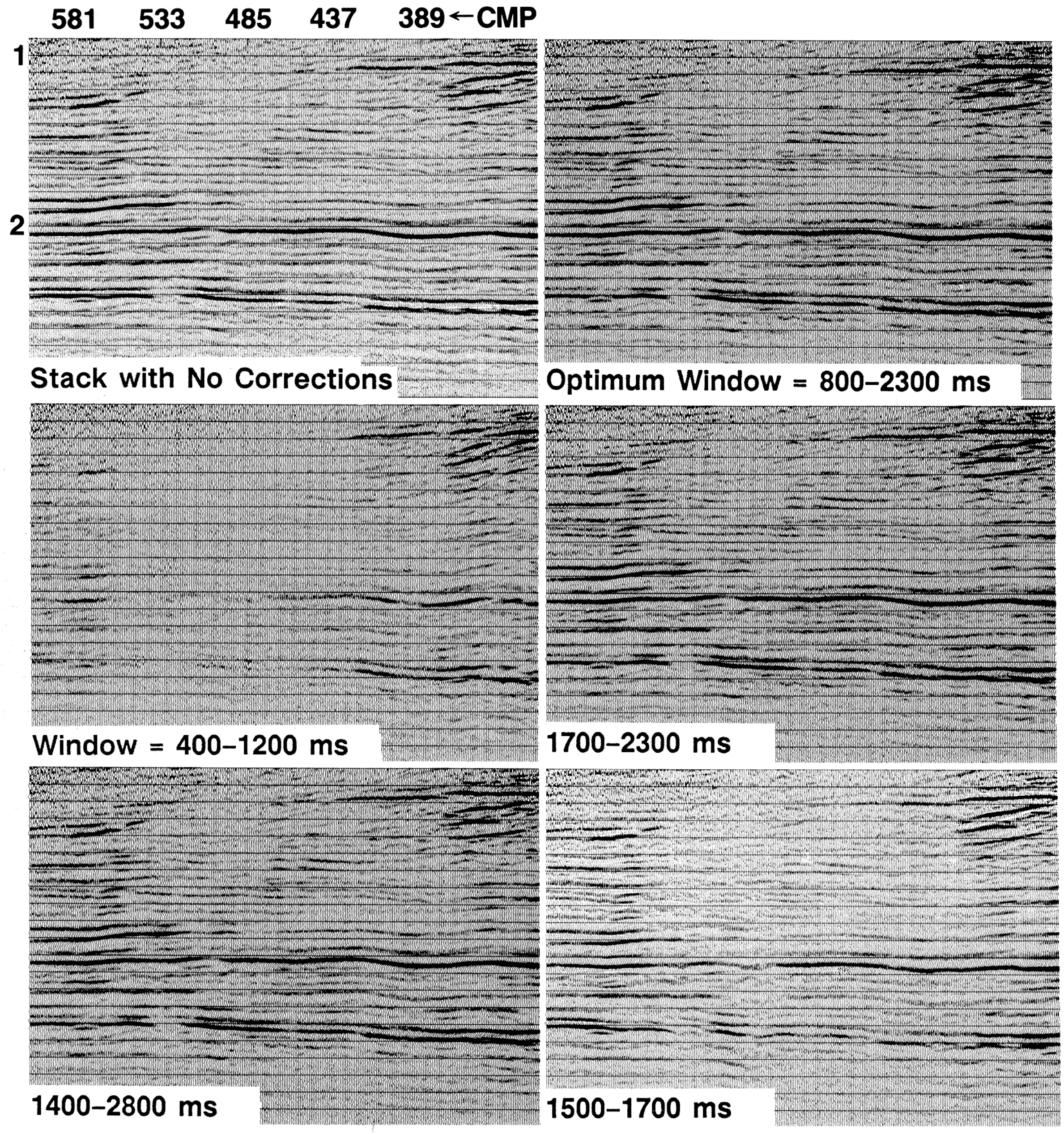


FIG. 3-83. Diagnostics for correlation window tests (Figure 3-81) showing CRP stacked data.



Stack with No Corrections

Optimum Window = 800–2300 ms

Window = 400–1200 ms

1700–2300 ms

1400–2800 ms

1500–1700 ms

FIG. 3-84 Test of correlation window: CMP stacks after residual statics corrections using five different correlation windows. CMP gathers are shown in Figure 3-81.

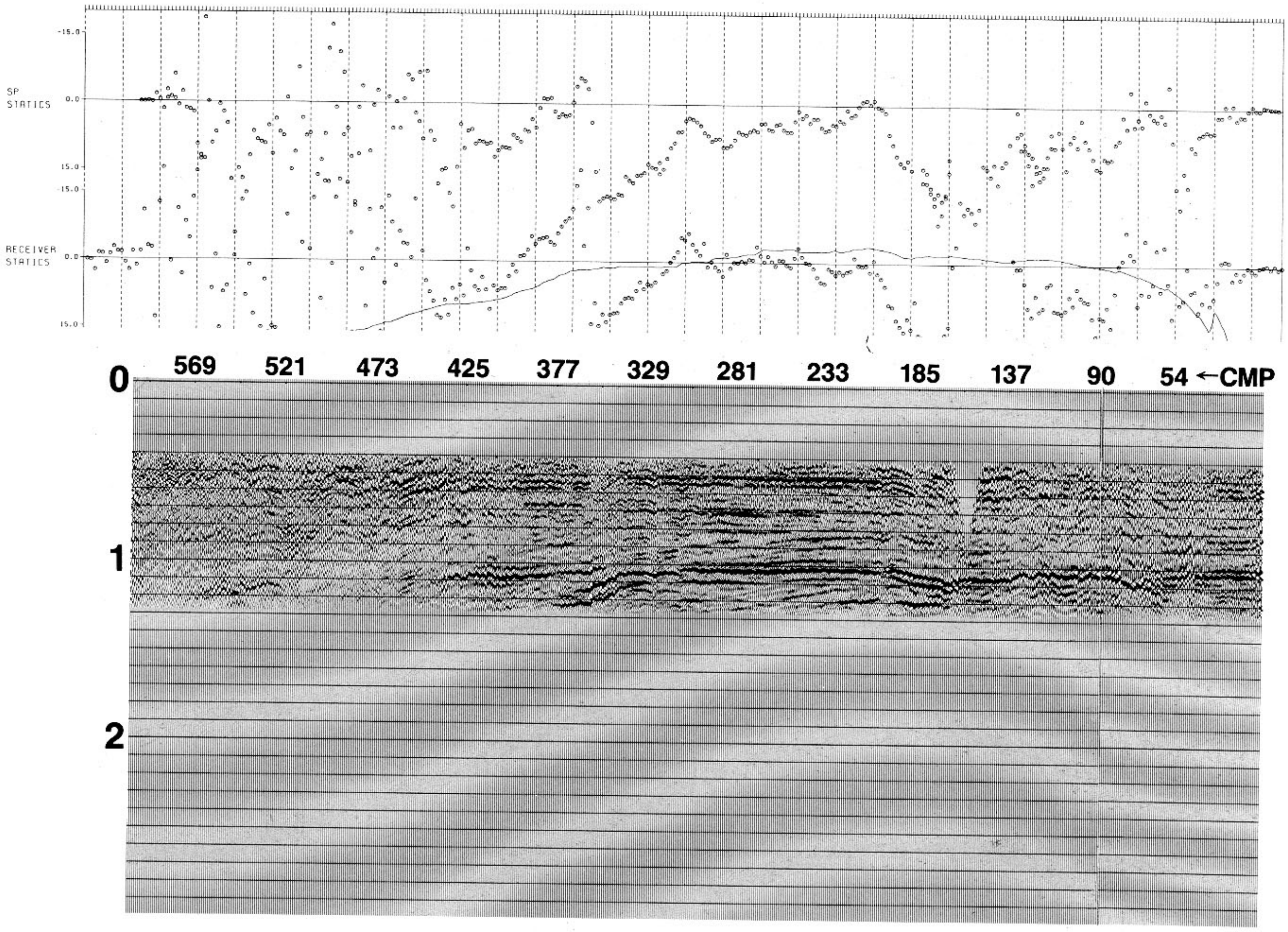


FIG. 3-85. Pilot traces (bottom) and shot and receiver statics solutions (top) for the 80-ms shift and the 400- to 1200-ms window. CMP stack is shown in Figure 3-84.

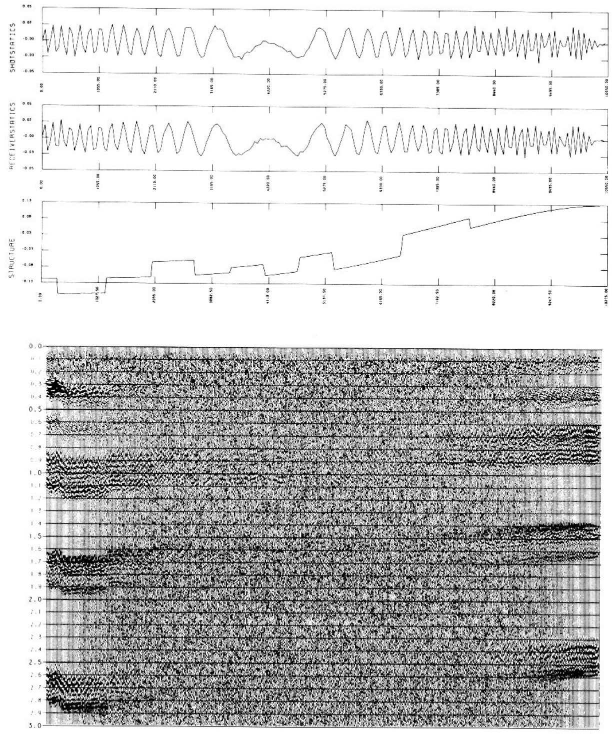


FIG. 3-86. A synthetic data set. In addition to the surface-consistent shot and receiver statics and the subsurface-consistent structure term, residual moveout shifts were introduced to the CMP gathers used in Model 2 of Figure 3-67.

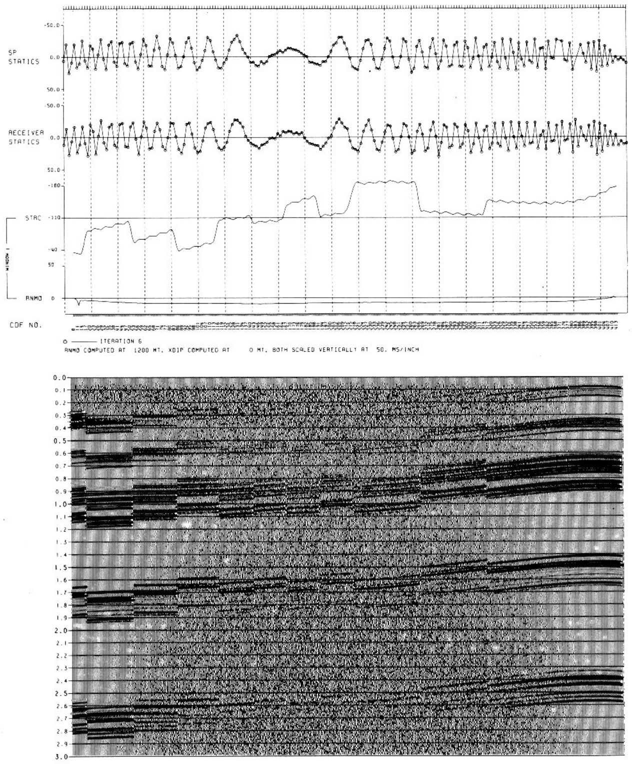


FIG. 3-87. Residual statics corrections applied to Model 3 data shown in Figure 3-86.

the second pass, the input are CMP gathers that already were corrected for residual statics. Velocity estimates must be revised between passes. Figures 3-88 and 3-89 show two different segments (A and B) of a section before and after residual statics corrections that were done in two passes. Diagnostic plots of the shot and receiver statics shown in Figures 3-90 and 3-91 indicate that the first pass has taken out a significant part of the static shifts. On the other hand, the second pass was most effective in segment B, where the S/N ratio is relatively poorer. Repeated estimation and application of the residual statics and velocity estimation is common in some processing systems. Use of a large number of trace correlations and multiple correlation peaks in a statics program tends to minimize the number of passes required.

3.6 REFRACTION STATICS

An important question in estimating shot and receiver statics is accuracy of the results as a function of wavelengths of static anomalies. Figure 3-92 is a synthetic data set that is identical to that in Figure 3-67, except for additional long wavelength shot and receiver static components. (Compare the graphic displays in Figures 3-67 and 3-92.) From the solution in Figure 3-93, note that the long wavelength components of the statics were badly underestimated. A significant difference between the stacked sections, in terms of horizon times, is apparent in Figures 3-70 and 3-93. The surface-consistent solution discussed in Section 3.4 cannot solve for the long wavelength static components. These usually are modeled as part of the structure term.

This problem of long wavelength statics variations also is seen in the field data in Figure 3-94. Residual statics corrections (based on reflections as described in Section 3.4) yield a much improved stack response (Figure 3-95). The short wavelength static shifts (less than a spread length) cause traveltime distortions in CMP gathers, and thus degrade stacking quality. However, merely improving the stack response by correcting for short-wavelength statics (Figure 3-95) may not be sufficient. In particular, note the structural features between midpoints *A* and *B*. These probably are caused by long wavelength statics variations. The long wavelength statics problem can be detected on the stacked section in Figure 3-95 by tracking

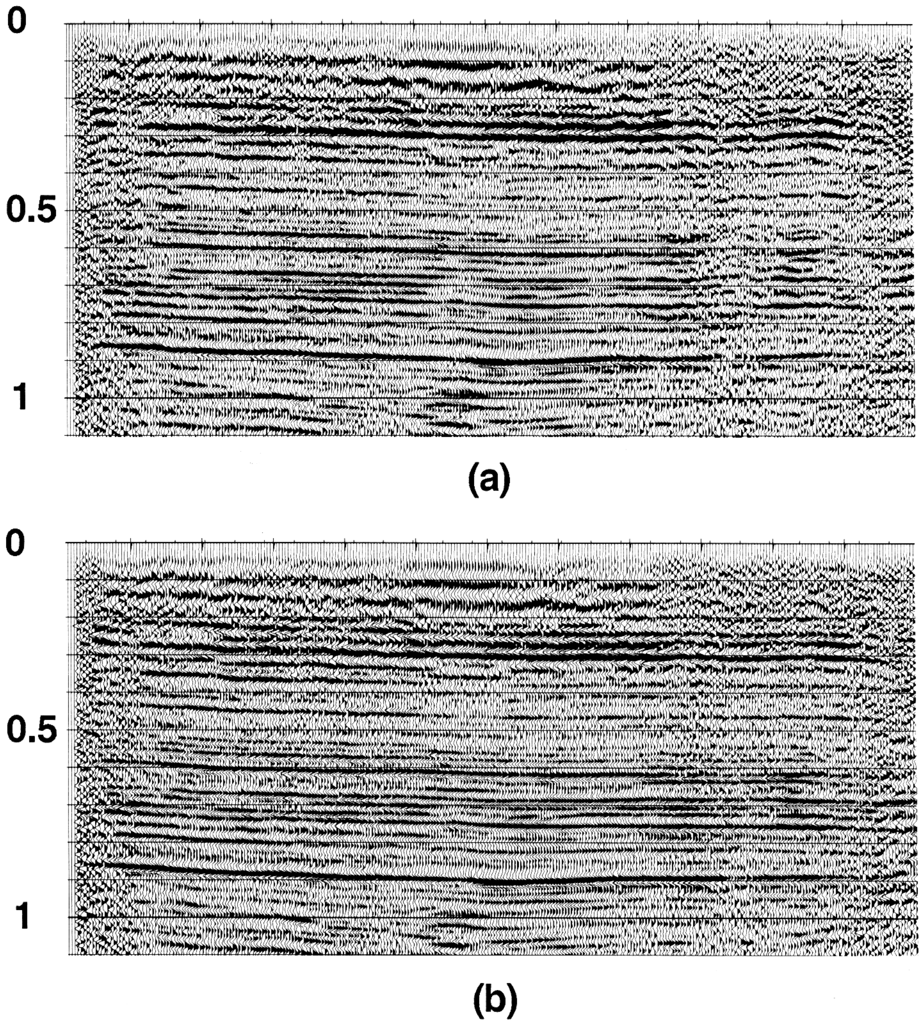


FIG. 3-88. First portion of a land line illustrating the improvement on velocities and CMP stacking as a result of residual statics corrections. Stack *A* (a) before residual statics corrections and using preliminary velocity picks and (b) after two passes of residual statics corrections and using final velocity picks.

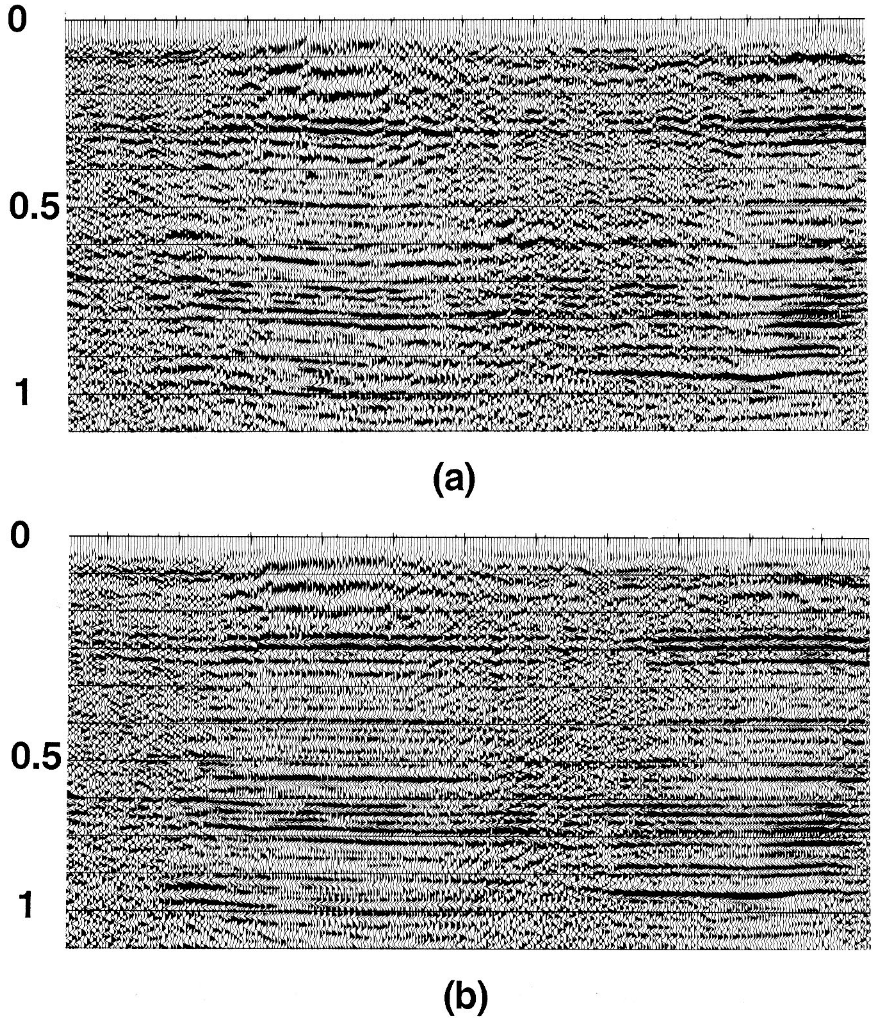


FIG. 3-89. Second portion of the land profile shown in Figure 3-88 illustrating the improvement on velocities and CMP stacking as a result of residual statics corrections. Stack *B* (a) before residual statics corrections and using preliminary velocity picks and (b) after two passes of residual statics corrections and using final velocity picks.

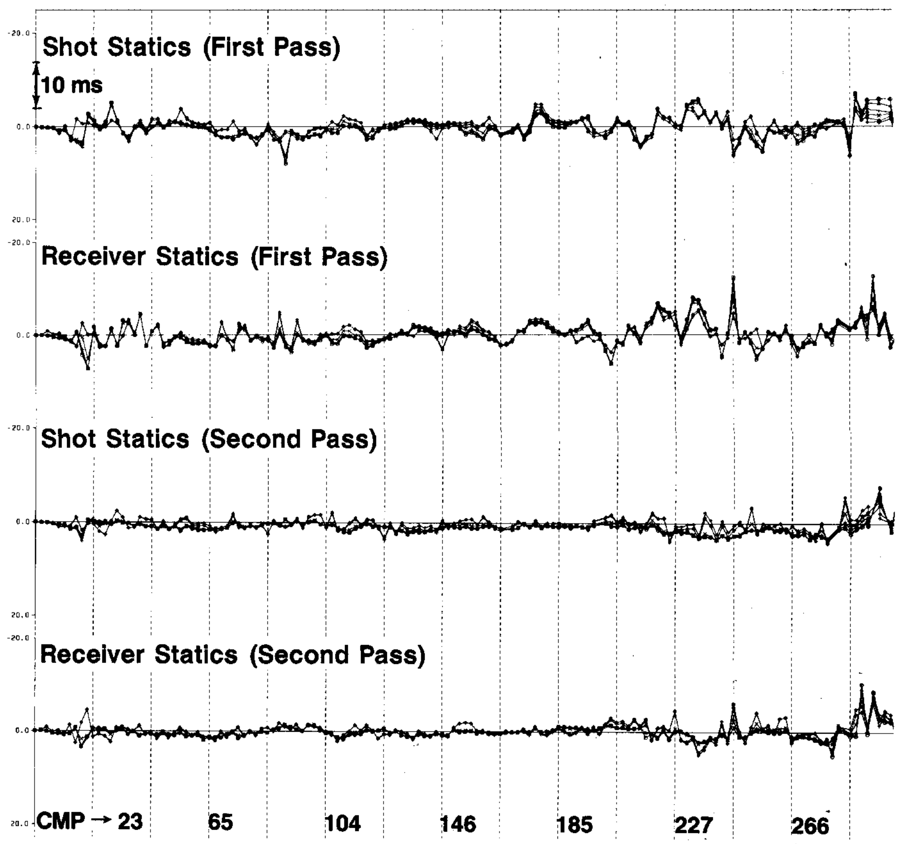


FIG. 3-90. Diagnostics for segment *A* from the residual statics corrections applied on the first portion of the land profile in Figure 3-88.

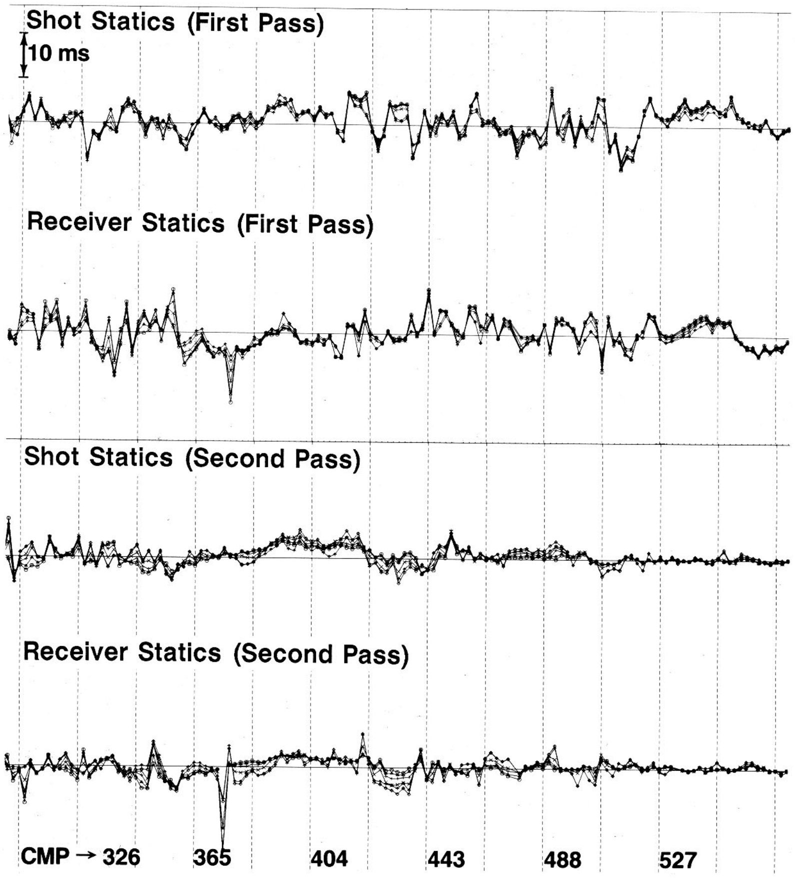


FIG. 3-91. Diagnostics for segment *B* from the residual statics corrections applied on the second portion of the land profile in Figure 3-89.

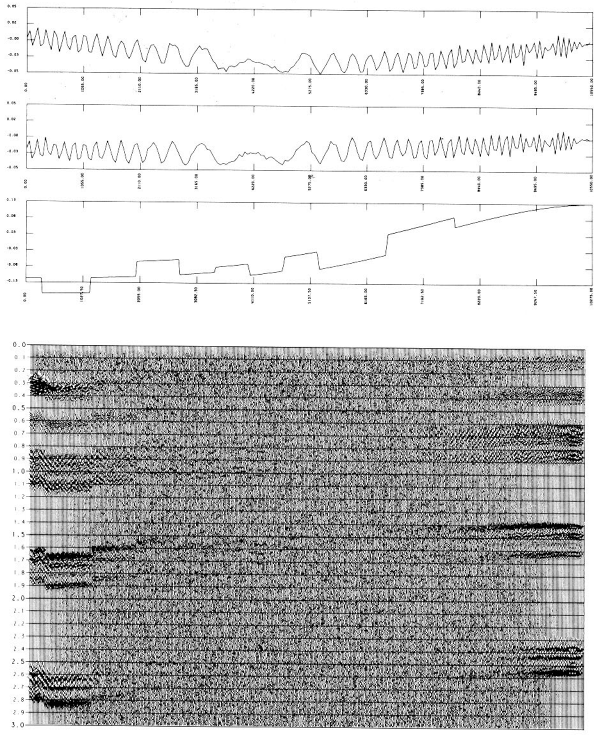


FIG. 3-92. A synthetic data contaminated with long-period statics.

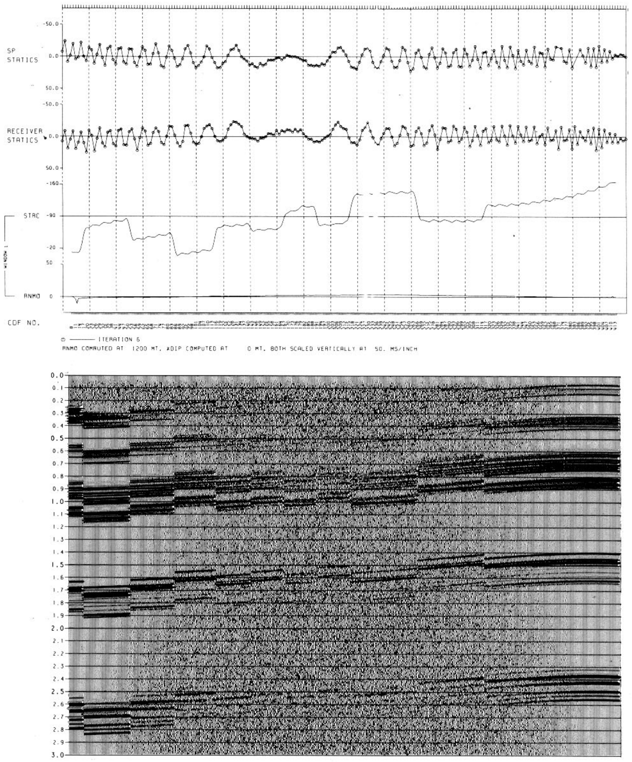


FIG. 3-93. Residual statics corrections applied to model 4 data in Figure 3-92.

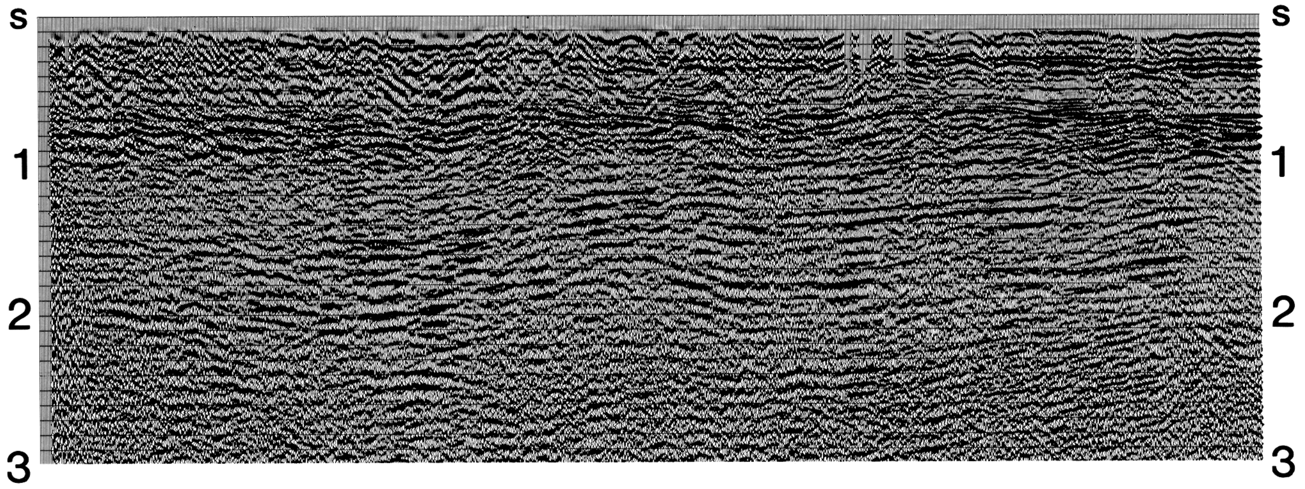


FIG. 3-94. A CMP stack obtained by applying only field statics to correct for elevation changes. (Data courtesy Nederlandse Aardolie Maatschappij B.V.)

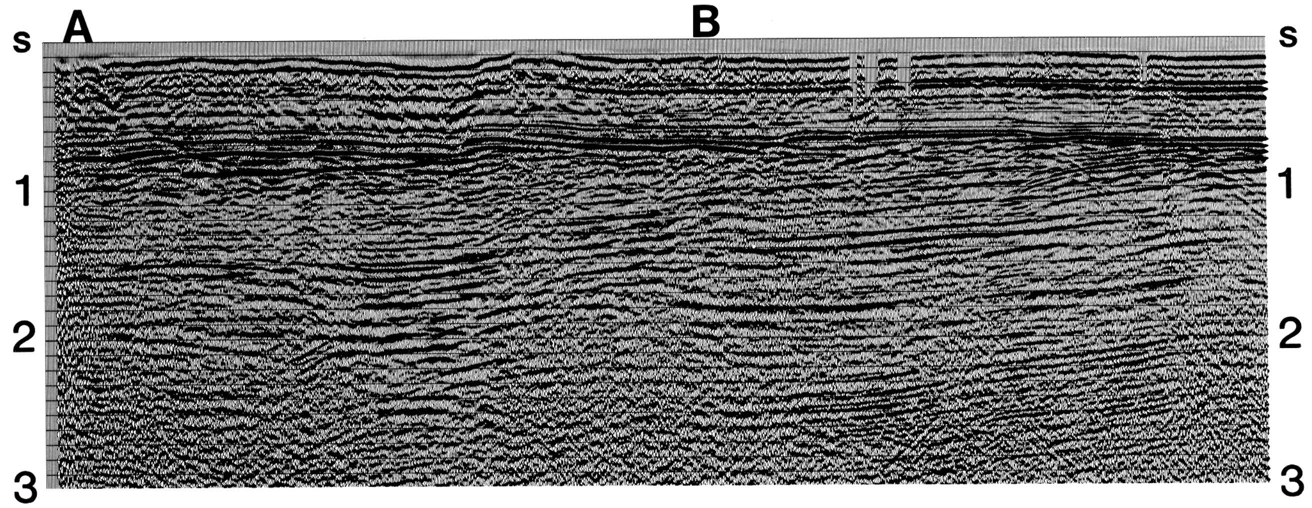


FIG. 3-95. Same data as in Figure 3-94 after reflection-based residual statics corrections (Section 3.4). (Data courtesy Nederlandse Aardolie Maatschappij B.V.)

the shallowest horizon. This suggests that the field statics applied to the data were not adequate.

Residual statics corrections are needed because field statics and datum corrections almost never totally compensate for the effects of near-surface velocity variations. This is because the near-surface velocity variations are not known and therefore exact corrections cannot be made. The method of surface-consistent statics estimation based on reflections works well in accommodating short-wavelength variations, but performs poorly in handling the long-wavelength variations. The main reason for this is that the input to reflection-based statics algorithms is arrival time differences between traces, not absolute times. Refraction-based statics methods are based on absolute first-break arrival times and, in theory, are able to estimate the long-period statics components.

3.6.1 Field Statics Corrections

It is now appropriate to review various methods of field statics corrections (Figure 3-96). If shots (denoted by S) are located below the weathering layer, then the total static correction for application to the trace associated with midpoint M is $t_D = t_S + t_R$, where t_S and t_R are the shot and receiver static corrections, respectively, down to a specified datum D. From the geometry of Figure 3-96, the field statics correction can be computed by:

$$t_D = \frac{2E_D - (E_S - D_S) - (E_R - D_R)}{v_b} - t_{UH}, \quad (3.40)$$

where E_D is the datum elevation and E_S and R_R are the surface elevations at the shot and receiver stations, respectively, D_S is the depth of the shot hole beneath the shot station, and D_R is the depth of the shot hole near the receiver station, t_{UH} is the uphole time measured at the receiver location (the time associated with the distance D_R in Figure 3-96). Finally, v_b is the bedrock (subweathering) velocity that may be derived from a deep uphole survey (to a point well below the weathering layer) conducted in the area. An uphole survey involves placing shots down the hole at various depth levels, then recording the arrivals at the surface near the hole. Alternatively, shots and receivers can be reciprocated if there is a caving problem down the hole. The hole must be deep enough to reach below the weathering layer. This provides a plot of time versus depth from which the bedrock velocity is obtained.

In land surveys, shots are not always placed in the bedrock for economic reasons, especially in areas with a thick weathering layer. Also, impulsive sources are not always used. Instead, surface sources such as vibroseis often are used. When surface sources or sources in shallow holes are used, the refracted arrivals can, at least in theory, be used to compute the static correction t_D down to a specified datum.

Consider the raypath geometry of Figure 3-97 with first-break time plot. For simplicity, consider a flat surface T and base of weathering B. From the refraction theory (Dobrin, 1960; Grant and West, 1965), note that the inverse of the slope of the line associated with the refracted wave arrivals is equal to the bedrock velocity v_b. Also note that the inverse of the slope of the line associated with the direct wave arrivals is equal to the velocity of the weathering layer v_w. By plotting the first breaks, these velocities and intercept time t_i, the time at $x = 0$, are estimated. From this, it is easy to show that depth to the bedrock z_W is given by

$$z_w = \frac{v_b v_w t_i}{2(v_b^2 - v_w^2)^{1/2}}. \tag{3.41}$$

Proof of this formula is left to Exercise 3.21. We assume that $v_b > v_w$. After computing z_w, the total static correction to the specified datum level can be derived:

$$t_D = -\frac{2z_w}{v_w} + \frac{2(E_D - E_S + z_w)}{v_b}, \tag{3.42}$$

where E_S is the surface elevation. If there is a difference between the elevations of shot and receiver stations, then an additional elevation correction using the bedrock velocity is required. Moreover, if the shots are located in boreholes, then the measured uphole time also must be incorporated into equation (3.42). The estimated static correction [equation (3.42)] is an average value over a distance that can range from the critical distance to the spread length, depending on the number of traces used in estimating the bedrock velocity. Nevertheless, more than one shotpoint is within a spread length. Therefore, an adequate definition of the near-surface model can be achieved and datum corrections can be computed for the entire profile.

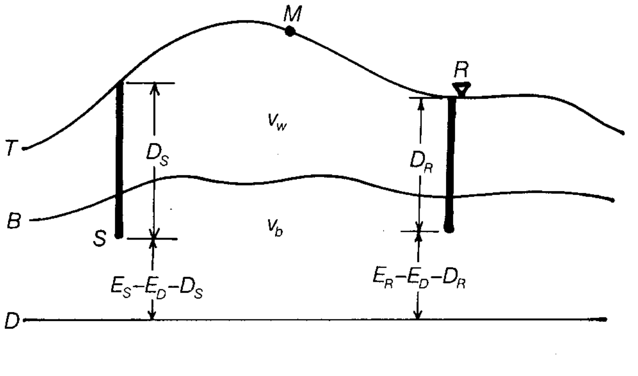


FIG. 3-96. A near-surface model for static corrections when shots are situated below the weathering layer. Here, S = shot, E_S = elevation at the shot station on the ground, R = receiver, E_R = elevation at the receiver station on the ground, T = surface topography, B = base of weathering, D = datum, E_D = datum elevation, v_w = weathering velocity, and v_b = bedrock velocity.

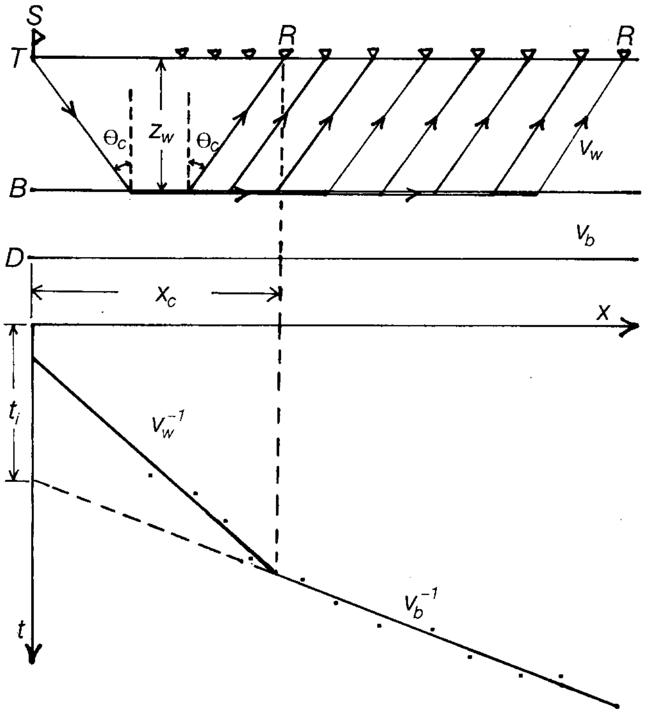


FIG. 3-97. Geometry for refracted arrivals. Here, T = ground level, B = base of weathering, D = datum level, θ_c = critical angle, and x_c = crossover distance.

3.6.2 The Plus-Minus Method

It often is difficult to use first breaks to estimate the intercept time and velocities for the weathering layer and bedrock. This is primarily because the base of weathering typically is an undulating surface, which makes traveltime plots difficult to interpret. Traveltime plots also are affected by severe elevation changes. Also, a typical field cable layout does not provide a sufficient number of channels inside the crossover distance x_c (Figure 3-97) for a reliable estimate of the weathering velocity or thickness. In most cases, v_w cannot be measured and a reasonable value is assumed for it.

Hagedoorn (1959) formulated a method to indirectly estimate intercept time and bedrock velocity; i.e., velocity below the refractor. The method still requires picking the first breaks. However, it does not require interpreting the traveltime plot (Figure 3-97). (Interpretation means drawing the linear segments for the direct and refracted waves.) Figure 3-98a shows three raypaths associated with shot-receiver pairs AD, DG, and AG. The basis of Hagedoorn's method involves computing two time values, the plus and minus times:

$$t_+ = t_{ABCD} + t_{DEFG} - t_{ABFG} \tag{3.43a}$$

and

$$t_- = t_{ABCD} - t_{DEFG} + t_{ABFG}. \tag{3.43b}$$

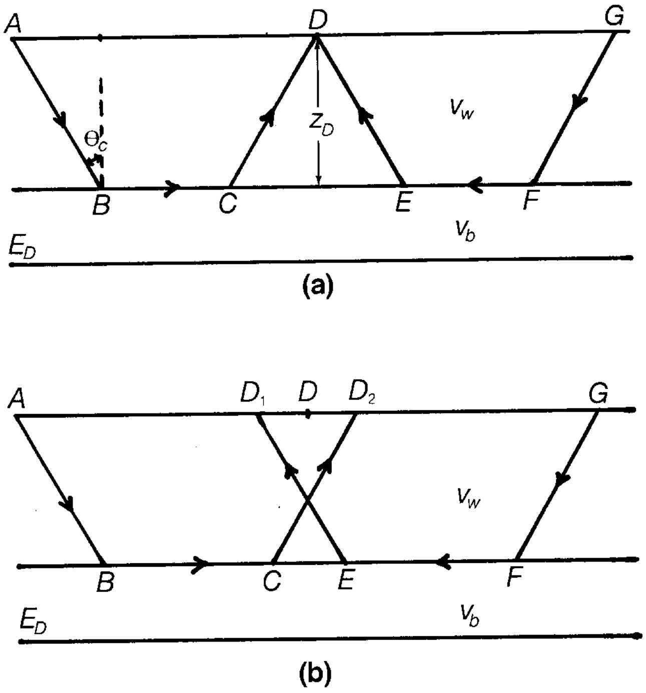


FIG. 3-98. (a) Geometry for the plus-minus method. (b) Geometry for the GRM (refer to Figure 3-96 for symbol definitions).

The times given on the right side of these equations are the measured (picked) values from the first breaks for the three raypaths shown in Figure 3-98a. From the raypath configuration (Figure 3-98a), note that plus time is indeed the intercept time and that minus time is related to bedrock velocity (Exercise 3.22). Thus, Hagedoorn's *plus-minus method* involves:

1. Picking the first breaks.
2. Computing the plus-minus times [equations (3.43a) and (3.43b)].
3. Deriving from these the intercept time and bedrock velocity.
4. Assuming a value for weathering velocity.
5. Computing the depth z_D to bedrock below station D (Figure 3-98a) from equation (3.41).
6. Computing the static shift at that station:

$$t_D = -\frac{z_D}{v_w} + \frac{(E_D - E_S + z_D)}{v_b},$$

where E_S and E_D are the surface and datum elevations at station D (Figure 3-98a).

Again, uphole and elevation corrections are needed before making the plus-minus static corrections.

In practice, raypaths that are suitable for first-break picking and are coincident at station D are not always found. Palmer (1981) generalized Hagedoorn's method of use of raypaths (Figure 3-98b). Palmer's technique, the generalized reciprocal method (GRM), takes into account offset separation D_1 D_2 when computing the plus time:

$$t_+ = t_{ABCD_2} + t_{D_1EFG} - t_{ABFG} - D_1D_2/v_b. \tag{3.44}$$

Note that more than one combination of raypaths associated with different separations of D_1 D_2 can be used to measure (pick) the traveltimes on the right sides of equations (3.43a), (3.43b), and (3.44). Consequently, there may be more than one estimate of the plus-minus times at a given station D. By carefully editing the first breaks, these estimates can be refined and reduced to a single estimate for each station. As a final note, first-break picking can be done automatically, interactively, manually, or as a combination thereof. To reliably pick, first apply linear moveout (LMO) to the data. Once picking is done, the LMO correction is reversed. Note that effectiveness of both reflection-and refraction-based methods of statics corrections depends on the reliability of the picking process. Apart from S/N ratio, indistinct first breaks (such as in vibroseis) sometimes can make picking consistent first breaks difficult.

Refer to the field data example in Figure 3-94. This stack was obtained by correcting for only elevation statics. Figure 3-99 is the same profile after correcting for statics using the GRM. Compare Figures 3-94 and 3-99 and note the elimination of long wavelength statics, which cause false structural features. The plus-minus or GRM both try to correct for all wavelengths of statics caused by variations in the near-surface model; i.e., undulations along the base of the weathering layer. Any remaining (residual) short wavelength statics should be corrected by using

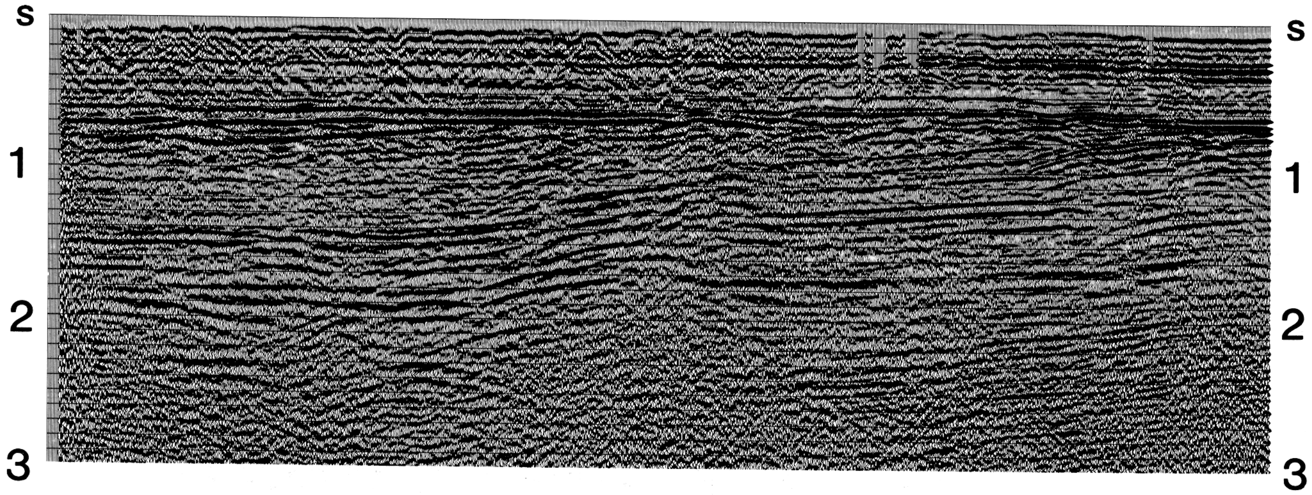


FIG. 3-99. Same data as in Figure 3-94 with static corrections using the GRM. (Data courtesy Nederlandse Aardolie Maatschappij B.V.)

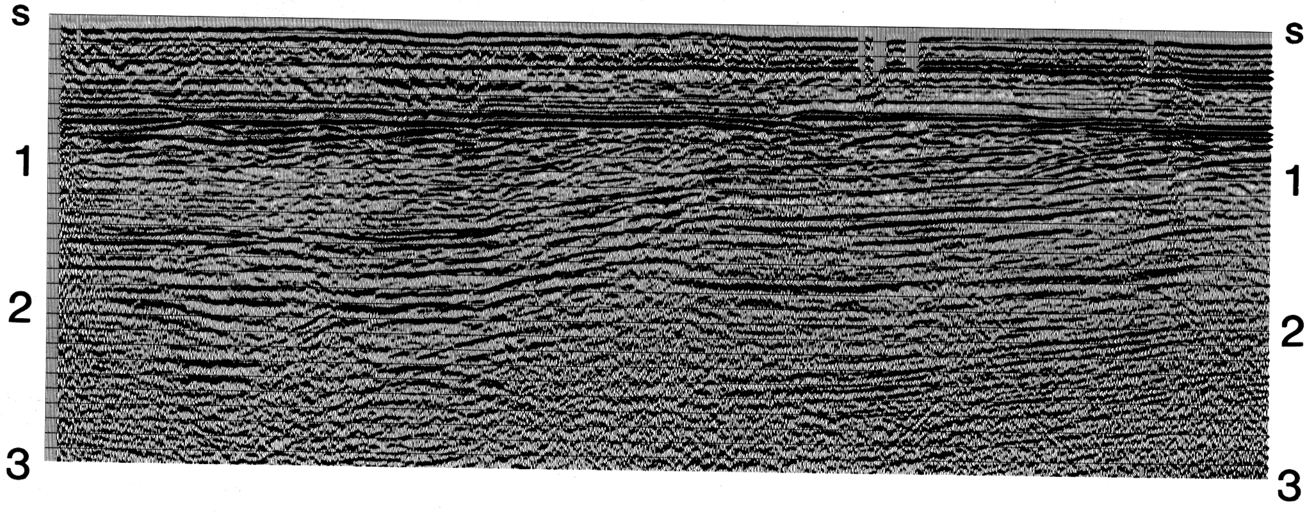


FIG. 3-100. Same data as in Figure 3-99 after (reflection-based) residual statics corrections. Compare this figure with Figure 3-95. (Data courtesy Nederlandse Aardolie Maatschappij B.V.)

a reflection-based method, such as that described in Section 3.4. The combined solution yields the section in Figure 3-100. This section should be compared with Figure 3-95.

To derive the near-surface model, these methods use the observed traveltimes from refracted arrivals that are assumed to be associated with the base of weathering. A problem arises when a near-surface model with more than one layer needs to be defined. This is the case in areas covered with glacial tilts and sand dunes. Several specialized techniques have been devised for these problems. Hampson and Russell (1984) applied generalized linear inversion (GLI) to the statics problem. This is an iterative, model-based approach that provides flexibility in defining a near-surface model consisting of arbitrarily parameterized multilayers. The process begins by computing the refracted arrival times from an assumed initial near-surface model. These computed traveltimes then are compared with the actual first-break picks (observed traveltimes). The procedure tries to minimize the difference between the computed and observed traveltimes by iteratively modifying model parameters for the near surface (such as velocities and thicknesses).

3.6.3 The Least-Squares Method

If a problem can be described by a model defined with a linear system of equations similar to traveltime equation (3.25), then a least-squares decomposition can be used to derive a solution. A brief description is given of the application of this method to the decomposition of ob-

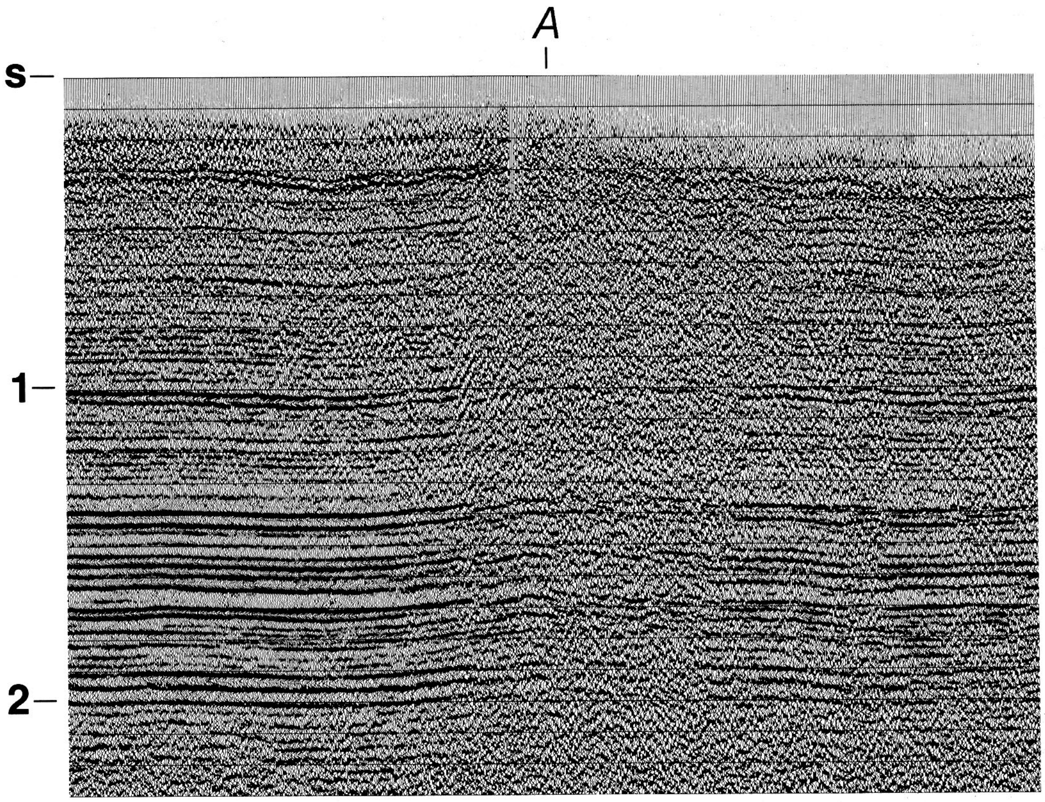


FIG. 3-101. A CMP stack from a land survey with field statics applied.

served traveltimes from refracted arrivals into shot and receiver statics components.

Consider the raypath geometry in Figure 3-98a from the jth shot location (A) to the ith receiver location (D). The traveltime t_{ij} for the refracted arrival can be expressed as a linear combination of three terms (Farrell and Euwema, 1984):

$$t_{ij} = s_j + r_i + x_{ij}/v_b, \tag{3.45}$$

where $x_{ij} = AD$ and

$$s_j = \frac{z_S(v_b^2 - v_w^2)^{1/2}}{v_b v_w} \tag{3.46a}$$

and

$$r_i = \frac{z_R(v_b^2 - v_w^2)^{1/2}}{v_b v_w} \tag{3.46b}$$

for the general case of different depths to bedrock below the shot (z_S) and receiver (z_R) stations. The unknown quantities are s_j, the portion of the intercept time associated with the jth source location, and r_i, the portion of the intercept time associated with the ith receiver location. Since the refracted arrivals typically are corrected for linear moveout (LMO) using an estimated refractor velocity, the third term in equation (3.45) must be interpreted as residual LMO.

Because equation (3.45) does not contain a structure term, any long wavelength static anomaly is partitioned between the other terms. This is not the case for the reflection-based residual statics model that was based on equation (3.25). Thus, following the field statics corrections (to account for elevation changes), consider the two-stage residual statics corrections:

1. Refraction-based statics corrections to remove long wavelength anomalies.
2. Reflection-based residual statics corrections to remove any remaining short wavelength static shifts.

Note that the linear traveltime model in equation (3.45) suffers from uncertainty in the value for the weathering layer velocity, as was the case for the plus-minus method and GRM.

Figure 3-101 is a stacked section with only the field statics applied. The bulge at midpoint location A probably is caused by a long wavelength statics anomaly. Start with CMP gathers (Figure 3-102a) and apply LMO to them (Figure 3-102b). Assuming that the first breaks correspond to a near-surface refractor, we use the estimated velocity from the first breaks (usually from a portion of the cable) to do this LMO. The stacked section of the shallow part of the data after LMO is shown in

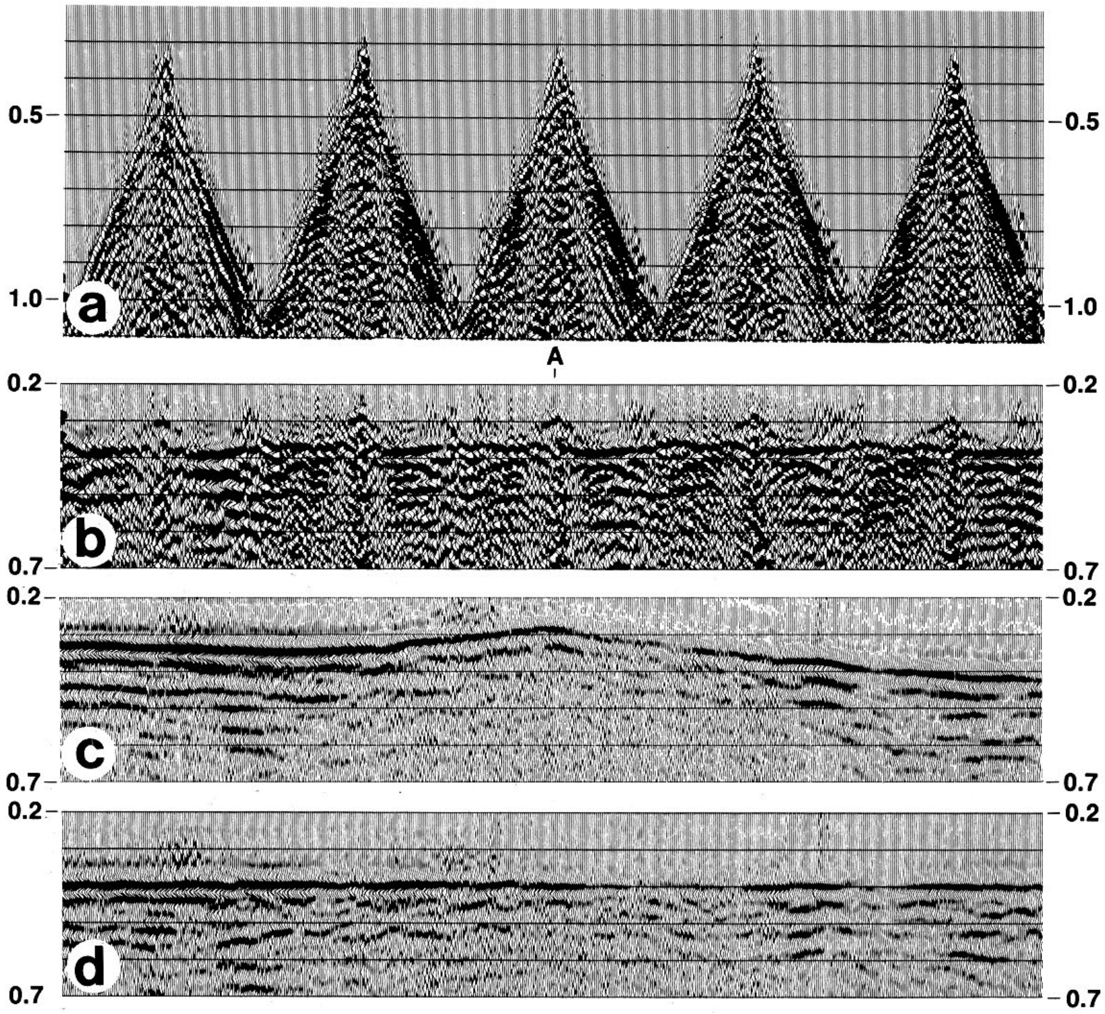


FIG. 3-102. (a) Selected CMP gathers from the profile in Figure 3-101; (b) CMP gathers after LMO correction; (c) stack of the LMO-corrected gathers as shown in (b); (d) stack of the LMO-corrected gathers after long-period statics were removed; compare with (c).

Figure 3-102c. This section is the equivalent of the pilot trace section that is associated with the reflection-based statics corrections. (An example of this is shown in Figure 3-85.) Traveltime deviations are estimated from the LMO-corrected gathers (Figure 3-102b) and are decomposed into shot and receiver intercept time components based on equation (3.45). These intercept times are used to compute shot and receiver static shifts, which are then applied to the CMP gathers shown in Figure 3-102a. The section, which corresponds to that in Figure 3-102c, after the refraction-based statics corrections clearly indicates removal of the significant long wavelength statics anomaly centered at midpoint location A (Figure 3-102d). The CMP stacked section after the refraction-based statics corrections shown in Figure 3-103 no longer contains the false structure (compare with Figure 3-101). This long wavelength anomaly cannot be removed by the reflection-based statics corrections (Figure 3-104). However, the latter removed any short wavelength statics components that were present in the data. By cascading the two corrections, we get the improved section in Figure 3-105.

REFERENCES

Al-Chalabi, M., 1973, Series approximations in velocity and traveltime computations: Geophys. Prosp., **21**, 783-795.

Claerbout, J. F., 1978, How to derive interval velocities using a pencil and straight edge: Stanford Exploration Project Report No. 14.

Dobrin, M. B., 1960, Introduction to geophysical prospecting: McGraw-Hill Book Co.

Doherty, S. M., 1986, Personal communication.

Farrell, R. C. and Euwema, R. N., 1984, Refraction Statics: Proc. Inst. Electr. Electron. Eng., **72**, 1316-1329.

Grant, F. S. and West, G. F., 1965, Interpretation theory in

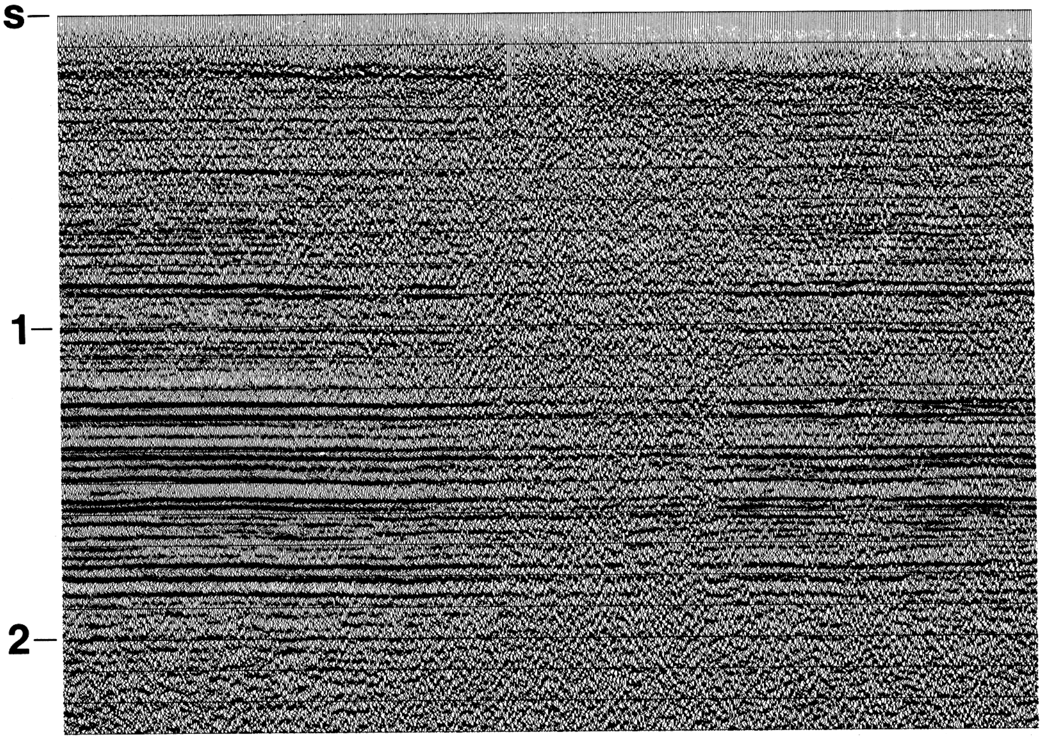


FIG. 3-103. CMP stack after long-period statics were removed using refraction arrivals; compare with Figure 3-101.

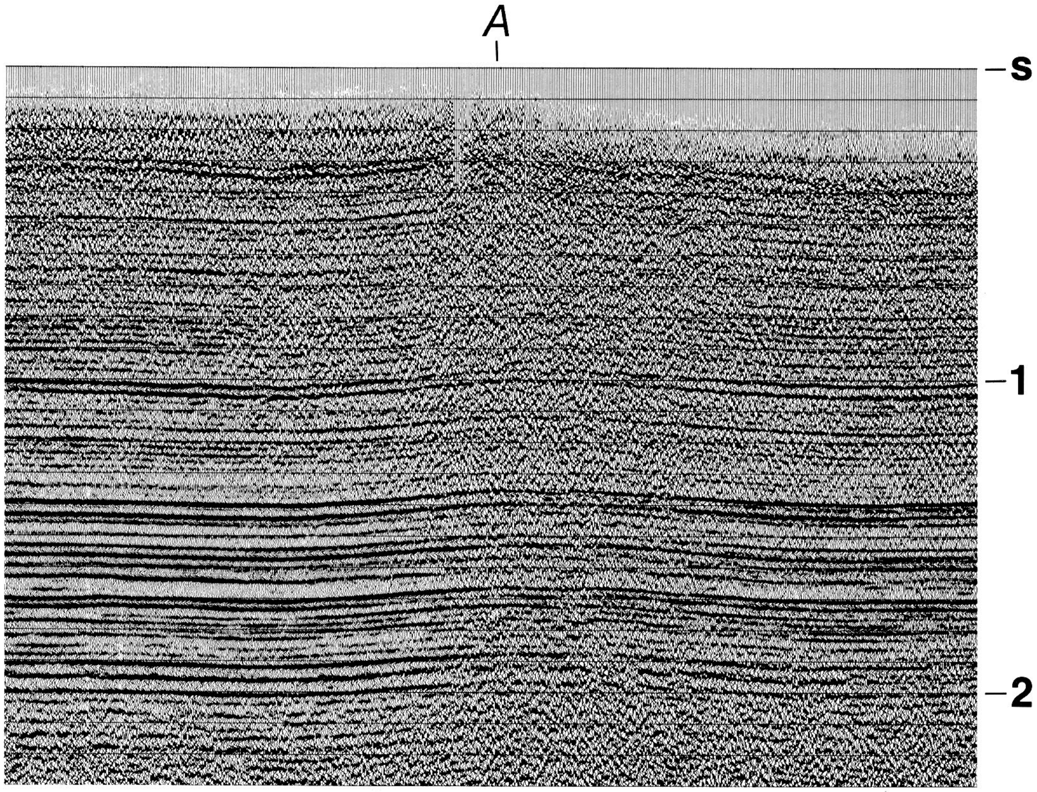


FIG. 3-104. CMP stack after residual (reflection-based) statics corrections; compare with Figure 3-101.

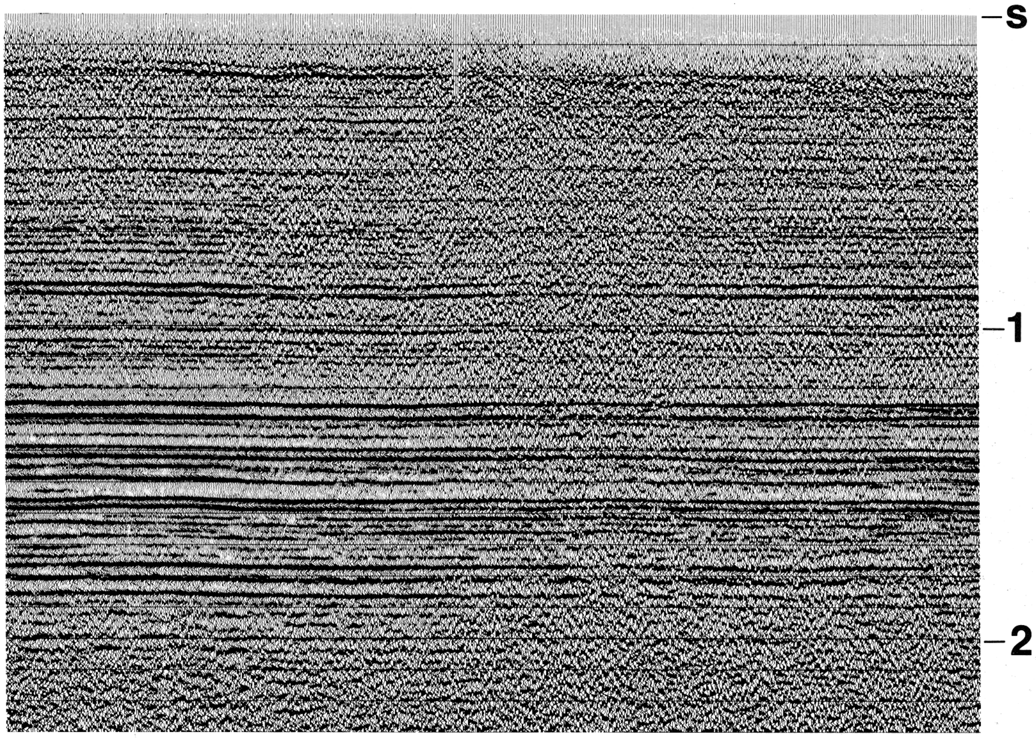


FIG. 3-105. CMP stack after refraction-and reflection-based statics corrections applied; compare with Figures 3-101, 3-103, and 3-104.

applied geophysics: McGraw-Hill Book Co.

Gulunay, N., 1985, A new method for the surface-consistent decomposition of statics using diminishing residual matrices (DRM): Presented at the 55th Ann. Int. Soc. Explor. Geophys. Mtg.

Hagedoorn, J. G., 1959, The plus-minus method of interpreting seismic refraction sections: Geophys. Prosp., **7**, 158-182.

Hampson, D. and Russell, B., 1984, First-break interpretation using generalized inversion: J. Can. Soc. Explor. Geophys., **20**, 40-54.

Hileman, J. A., Embree, P., and Pfleuger, J. C., 1968, Automated static corrections: Geophys. Prosp., **16**, 326-358.

Hubral, P. and Krey, T., 1980, Interval velocities from seismic reflection time measurements: Soc. Explor. Geophys. Monograph.

Levin, F. K., 1971, Apparent velocity from dipping interface reflections: Geophysics, **36**, 510-516.

Neidell, N. S. and Taner, M. T., 1971, Semblance and other coherency measures for multichannel data: Geophysics, **34**, 482-497.

Nur, A., 1981, Physical properties of rocks: Soc. Explor. Geophys. Short Course Notes.

Palmer, D., 1981, The generalized reciprocal method of refraction seismic interpretation: Geophysics, **46**, 1508-1518.

Schneider, W. A., 1971, Developments in seismic data processing and analysis (1968-1970): Geophysics, **36**, 1043-1073.

Schneider, W. A. and Kuo, S., 1985, Refraction modeling for statics corrections: Presented at the 55th Ann. Int. Soc. Explor. Geophys. Mtg.

Sheriff, R. E., 1976, Inferring stratigraphy from seismic data: Am. Assn. Petr. Geol. Bulletin, **60**, 528-542.

Sherwood, J. W. C. and Poe, P. H., 1972, Constant velocity stack and seismic wavelet processing: Geophysics, **37**, 769-787.

Taner, M. T. and Koehler, F., 1969, Velocity spectra—digital computer derivation and applications of velocity functions: Geophysics, **39**, 859-881.

Taner, M. T., Koehler, F., and Alhilali, K. A., 1974, Estimation and correction of near-surface time anomalies: Geophysics, **41**, 441-463.

Wiggins, R. A., Larner, K. L., and Wisecup, R. D., 1976, Residual statics analysis as a general linear inverse problem: Geophysics, **41**, 922-938.

Yilmaz, O. and Chambers, R., 1984, Migration velocity analysis by wave field extrapolation: Geophysics, **49**, 1664-1674.

EXERCISES

Exercise 3-1. Why does salt have anomalously high velocity (4.5 to 5.5 km/s)?

Exercise 3-2. Measure the traveltimes corresponding to offset values of 1 and 3 km in Figure 3-5. Then compute the velocity above the reflector and verify that it is 2264 m/s. Note that the zero-offset trace is not recorded; therefore, $t(0)$ normally is not a known quantity.

Exercise 3-3. Refer to Figure 3-25 and compute the interval velocities between events at $t(0) = 0.4$ and 0.8 s and at $t(0) = 1.2$ and 1.6 s using Claerbout's straightedge method. Verify your results by comparing them with the interval velocity function in Figure 3-11.

Exercise 3-4. Refer to Figure 3-8. The CMP gather in Figure 3-8a contains a hyperbola. Following NMO corrections and using 2000 (Figure 3-86) and 2500 m/s (Figure 3-8c) velocities, are the curves hyperbolic?

Exercise 3-5. Make velocity picks from the constant velocity scan panel in Figure 3-26.

Exercise 3-6. Make velocity picks from the CVS panel in Figure 3-27. Identify some multiples.

Exercise 3-7. Make velocity picks from the CVS panel in Figure 3-106.

Exercise 3-8. Derive Levin's relation [equation (3.7)] for the NMO velocity for a single dipping reflector from the geometry of Figure 3-16.

Exercise 3-9. Consider two intersecting lines. Would you expect that velocity analyses at the intersection point would yield the same velocity function?

Exercise 3-10. Which is correct: velocity analysis from datum or velocity analysis from surface?

Exercise 3-11. Why does equation (3.3) contain only terms with even powers?

Exercise 3-12. Derive an expression for interval velocity as a function of rms velocity using equation (3.4).

Exercise 3-13. Derive Levin's relation for the 3-D case from the geometry of Figure 3-17.

Exercise 3-14. Derive the relationship for NMO stretching described by equation (3.6).

Exercise 3-15. Fill the missing elements in on the following table. Average velocity v_{avg}, which relates vertical traveltime to depth in a horizontally layered medium, is defined as:

$$v_{\text{avg}} = \left(\sum_{i=1}^{N} v_i \Delta t_i\right) \Big/ \left(\sum_{i=1}^{N} \Delta t_i\right),$$

where $\Delta t_i = \Delta z_i / v_i$, Δz_i = layer thickness, v_i = interval velocity. The rms velocity is given by equation (3.4).

Layer Thickness, m	Interval Velocity, m/s	RMS Velocity, m/s	Average Velocity, m/s
200	2000		
		*	*
300	3000		
		*	*
400	4000		
		*	*
350	3500		
		*	*
500	5000		

Exercise 3-16. Explain why the velocity for horizon A in Figure 3-48 behaves as shown in the HVA display below the salt dome, S. Note that salt is the high-velocity layer.

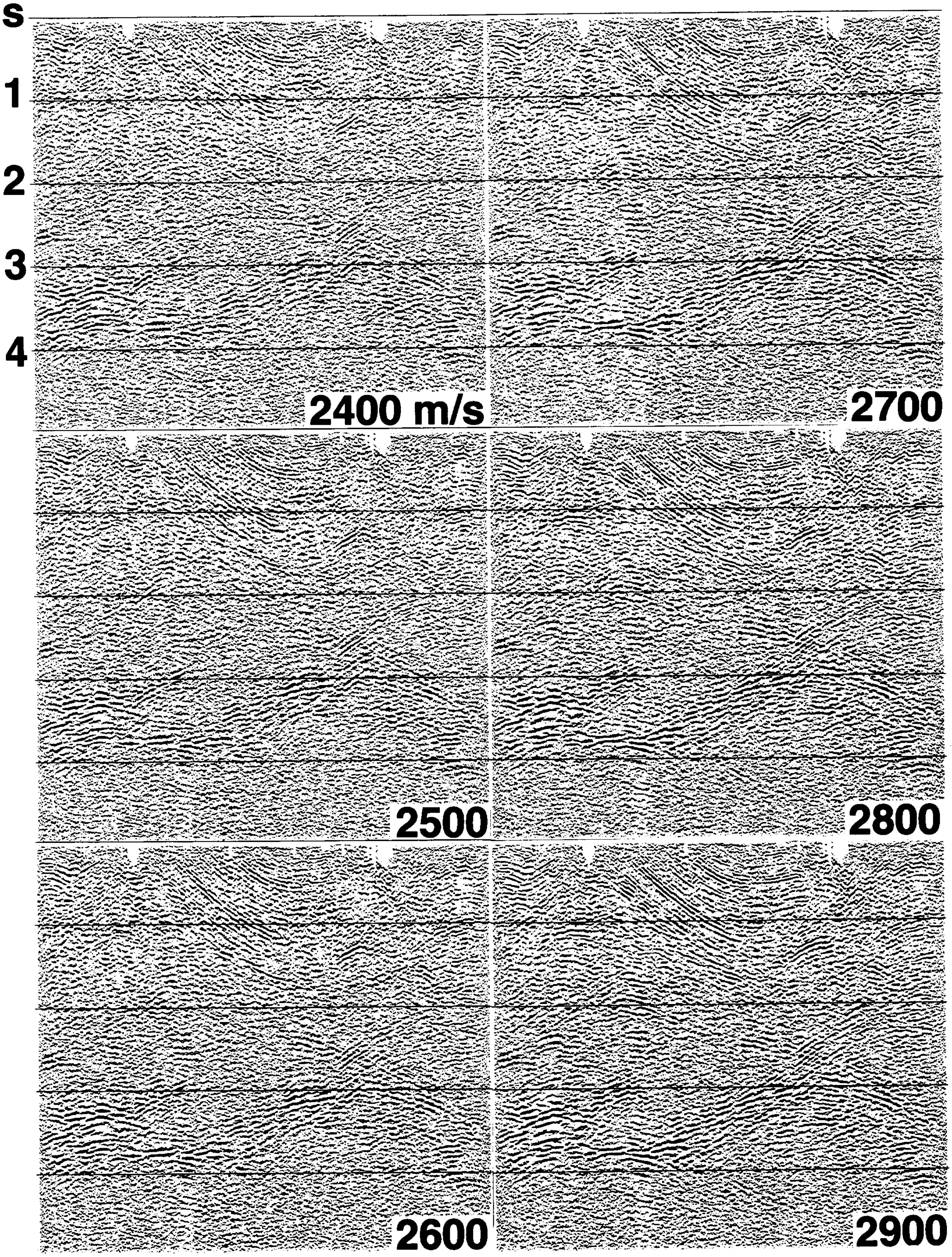


FIG. 3-106. A CVS panel for a line from a complex structure area (see Exercise 3.7).

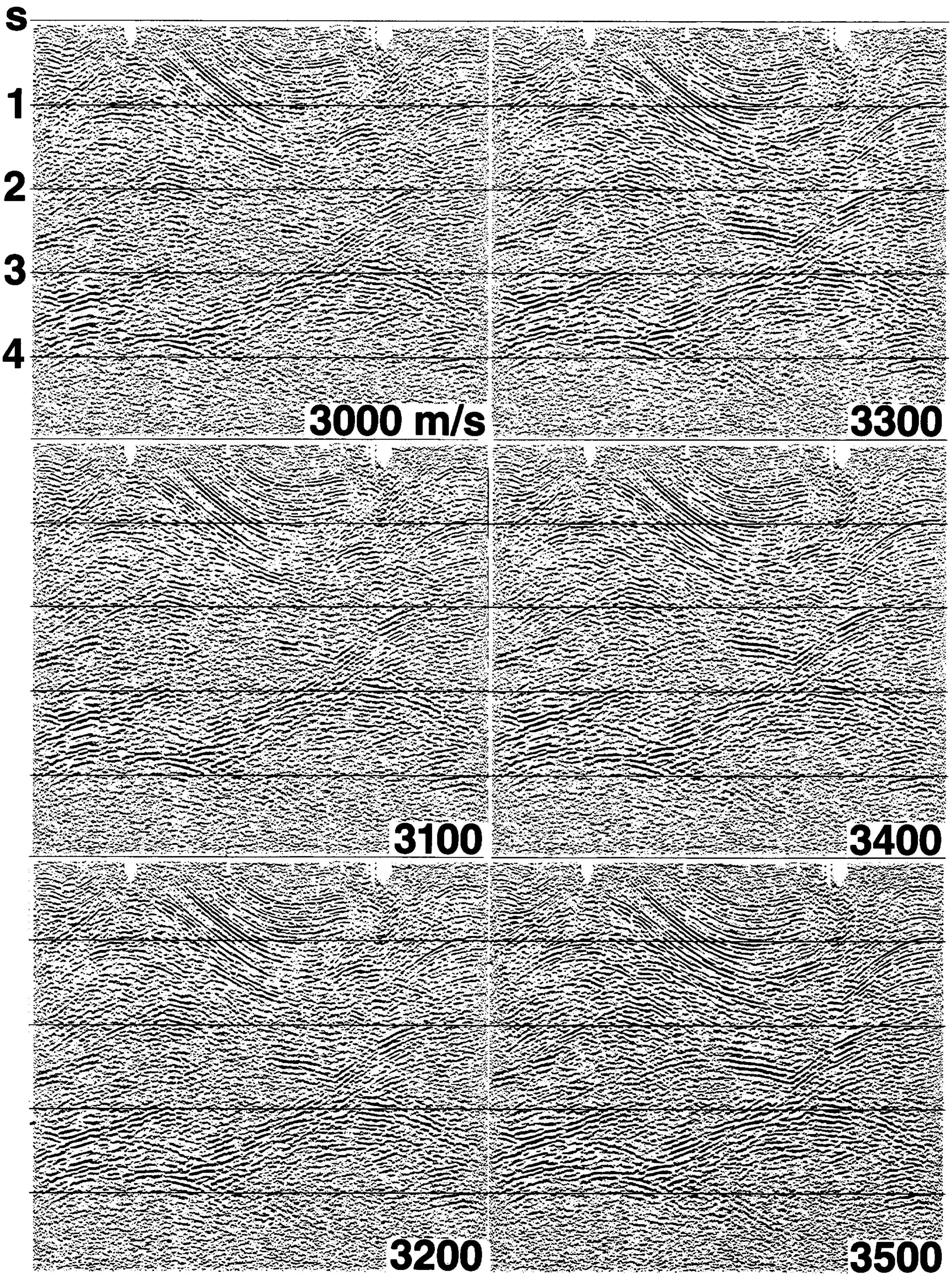


FIG. 3-106 continued

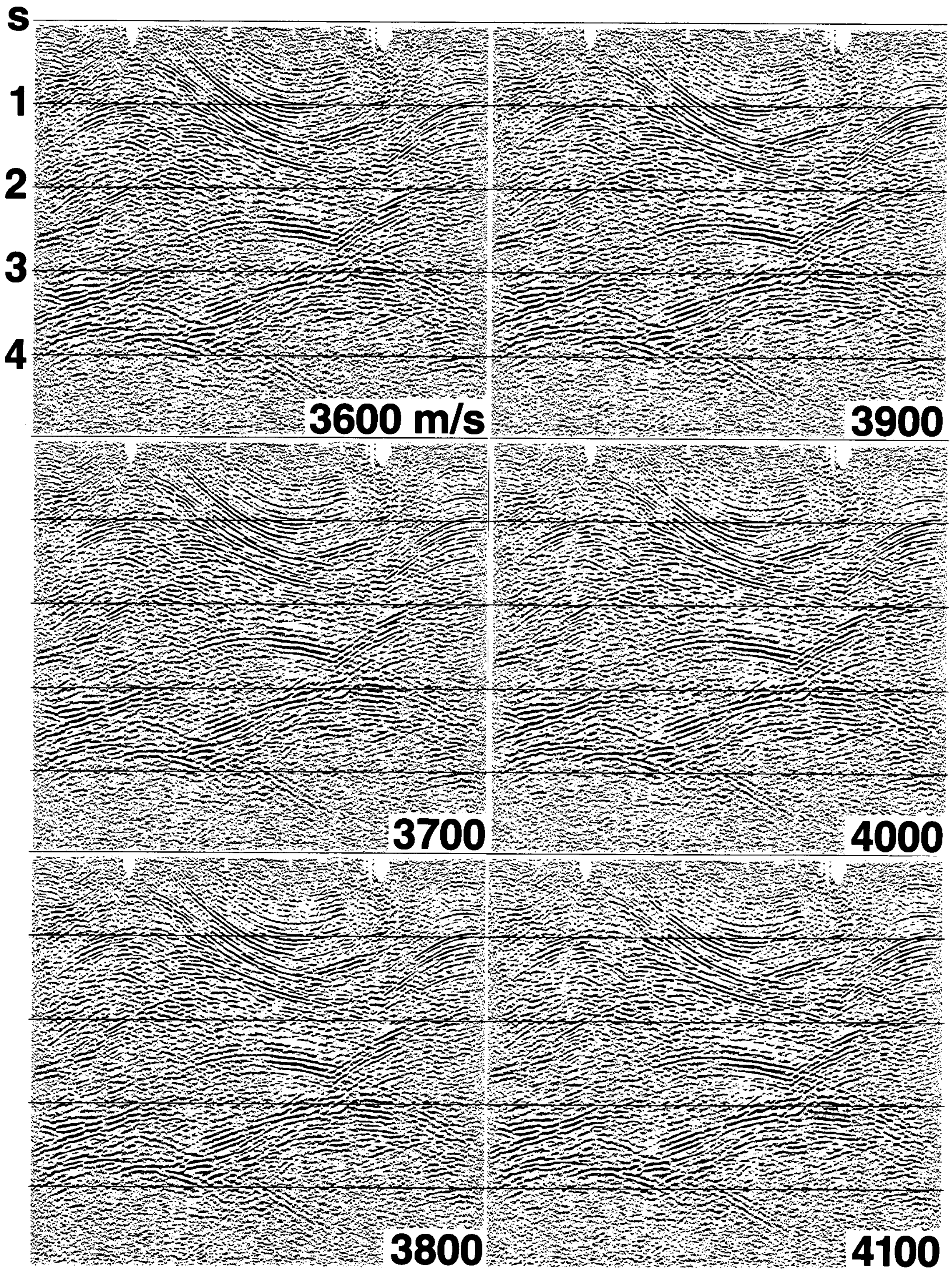


FIG. 3-106 continued

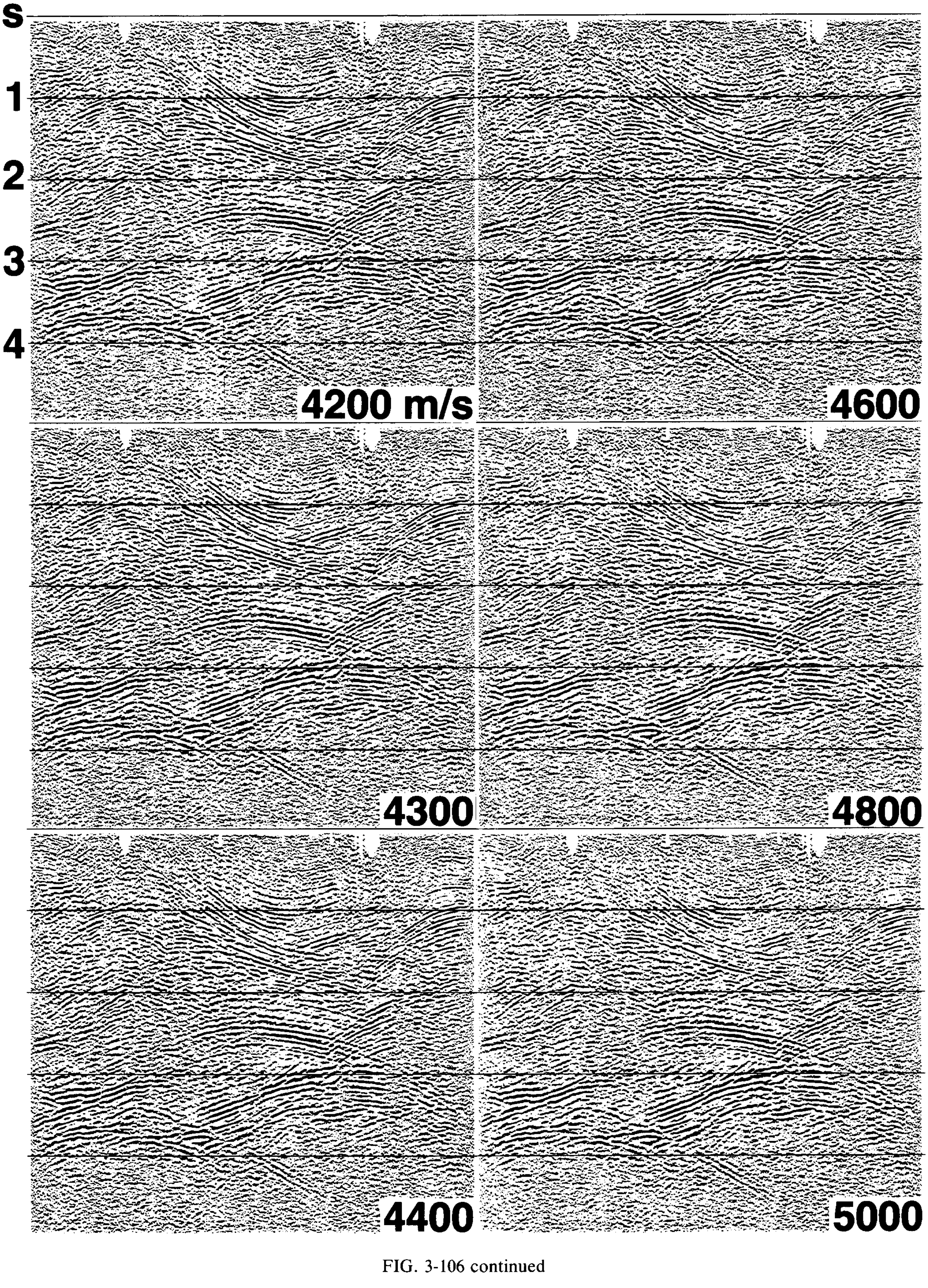


FIG. 3-106 continued

Exercise 3-17. Suppose you want to fit a set of observed traveltimes to a parabola of the form $t(x) = a + bx + cx^2$. The tabulated input values are given below.

i	x_i	Observed t'_i
1	0	0.4
2	1	1.1
3	2	3.5
4	3	7.9
5	4	14.4

You have five equations and three unknowns.
Set up the least-squares problem and solve for a, b, and c.

Exercise 3-18. Solve the system:

$$x_1 - 2x_2 = 1$$
$$x_1 + 4x_2 = 4$$

by the Gauss-Seidel iterative method. Verify the results by solving these equations by the method of substitution to obtain the correct solution: $x_1 = 2$ and $x_2 = 0.5$.

Exercise 3-19. Solve the system:

$$x_1 + 4x_2 = 4$$
$$x_1 - 2x_2 = 1$$

by the Gauss-Seidel iterative method. Note that this is the same problem as in Exercise 3.18, except that the order of the equations is reversed. The solution should be the same. You will find that the solution cannot be obtained because the process of iteration will not converge. This demonstrates the importance of ordering the equations when solving by the Gauss-Seidel method.

Exercise 3-20. Write equations (3.38) and (3.39) for i = 1,2,3 and j = 1,2,3. You will find that there are six unknowns, but five independent equations.

Exercise 3-21. Using the geometry of refracted arrivals depicted in Figure 3-97, derive the expression for depth to bedrock [equation (3.41)].

Exercise 3-22. Using the geometry of Figure 3-98a, show that the plus time defined by equation (3.43a) is the intercept time and minus time is related to the reciprocal of the bedrock velocity.

Exercise 3-23. Can you use refraction-based statics techniques in permafrost areas?

Exercise 3-24. Which of the following adversely affects the quality of velocity spectrum: long wavelength or short wavelength (less than a cable length) statics anomalies?

Exercise 3-25. In what way does inside muting of a CMP gather affect the velocity spectrum?

4

Migration

4.1 INTRODUCTION

Migration moves dipping reflectors into their true subsurface positions and collapses diffractions, thereby delineating detailed subsurface features such as fault planes. In this respect, migration can be viewed as a form of spatial deconvolution that increases spatial resolution. Figure 4-1 shows a stacked section before and after migration. This stacked section contains a salt dome with steep flanks. Figure 4-1 also shows a sketch of two prominent features: one is the diffraction hyperbola *D*, which originates at the tip of the salt dome; and the reflection *B*, off the flank of the salt dome. After migration, note that the diffraction has collapsed to its apex *P* and the dipping event has moved to a subsurface location *A*, which is at or near the salt dome flank.

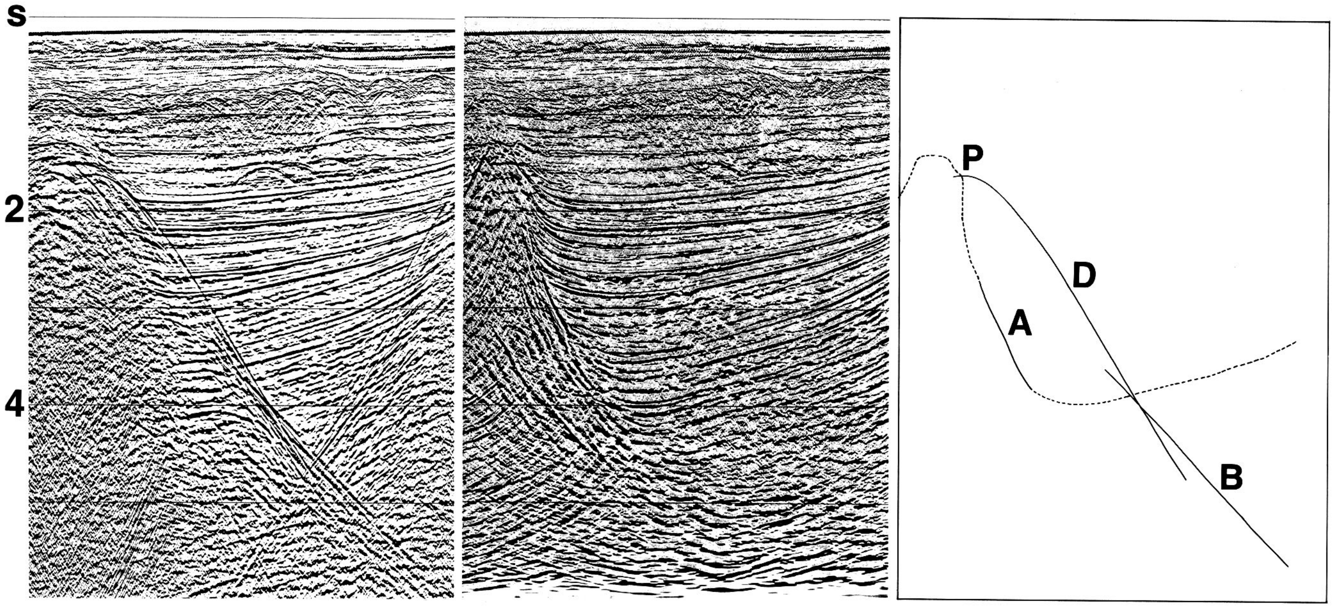
Figure 4-2 is an example with a different type of structural feature. The stack contains a zone of horizontal reflectors down to 1 s. After migration, these events are virtually unchanged. Note the prominent unconformity that represents an ancient erosional surface just below 1 s. On the stacked section, the unconformity appears complex, while on the migrated section, it becomes interpretable. The *bowties* on the stacked section are untied and turned into synclines on the migrated section. The deeper event in the neighborhood of 3 s is the multiple associated with the unconformity above. When treated as a primary and migrated with the primary velocity, it is migrated excessively (overmigrated).

The unmigrated section in Figure 4-3 contains an abundance of *diffractions* that is associated with intense faulting. The clarity of the migrated section definitely is preferable. An interpreter would find it easier to position the faults and thereby derive a reliable time structure map using the migrated section.

There do not have to be obviously complex structures on a stacked section to warrant migration. Figure 4-4 shows a stacked section that primarily is made up of parallel reflectors dipping to the left. Occasional disruptions of reflection continuity, which give rise to diffractions, also are seen. These disruptions are due to growth faults that are subtle on the stacked section, but become apparent on the migrated section. From these four examples (Figures 4-1 through 4-4), note that migration does not displace horizontal events, rather, it moves dipping events in the updip direction and collapses diffractions, thus enabling us to delineate faults.

The goal of migration is to make the stacked section appear similar to the geologic cross section along the seismic line. Ideally, we want to get a depth section from the stacked section. However, the migrated section commonly is displayed in time. One reason for this is that velocity estimation based on seismic and other data always is limited in accuracy. Therefore, depth conversion is not completely accurate. Another reason is that interpreters prefer to evaluate the validity of migrated sections by comparing them to the unmigrated data. Therefore, it is preferable to have both sections displayed in time. The migration process that produces a migrated time section is called *time migration*. Time migration, the main theme of Chapter 4, is appropriate as long as lateral velocity variations are mild to moderate.

When the lateral velocity gradients are significant, time



(a) (b) (c)

FIG. 4-1. (a) CMP stack, (b) migration, (c) sketch of a prominent diffraction *D* and a dipping event before (*B*) and after (*A*) migration. Migration moves the dipping event *B* to its assumed true subsurface position *A* and collapses the diffraction *D* to its apex *P*. The dotted line indicates the boundary of the salt dome.

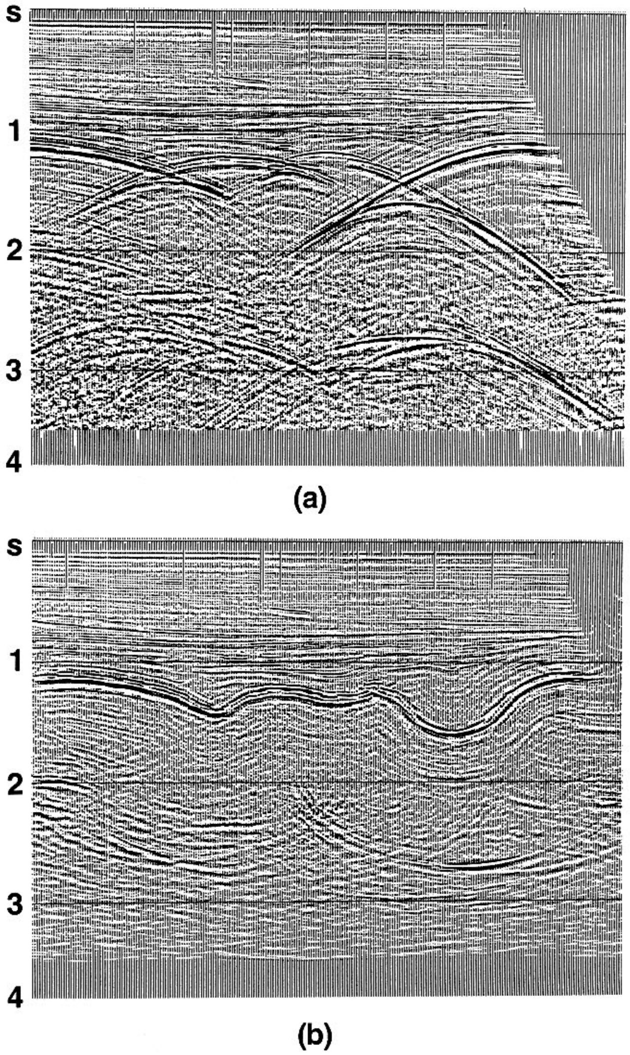


FIG. 4-2. Migration (b) unties the bowties on the stacked section (a) and turns them into synclines. (Taner and Koehler, 1977; figure courtesy Seiscom-Delta, Inc.)

migration does not produce the true subsurface picture. Instead, we need to use *depth migration* (Section 5.2), the output of which is a depth section. Consider the data from an area with intense salt tectonics in Figure 4-5. The top and base of the salt are indicated by *T* and *B*, respectively. The section has been time migrated properly above the top of the salt. However, note the apparent crisscrossing (overmigration) of the reflection associated with the base of the salt. The overmigration is the result of inadequate treatment by time migration of the effects of severe raypath bendings caused by the strong velocity contrast between the salt layer and the overlying rocks.

The tectonic settings associated with salt diapirism, overthrusting, or irregular water-bottom topography usually are three dimensional in character. A stacked section really is the response of a 3-D subsurface on a 2-D plane of profile. Therefore, 2-D data analyses are not completely valid for such 3-D data. Figure 4-6a is a cross-line stacked section from a land 3-D survey. Figure 4-6b is a 2-D migration of this section, while Figure 4-6c is the same section after 3-D migration of the entire survey. In particular, note the significant difference in the imaging of the top of salt *T* and base of salt *B*. In 2-D migration, we assume that the stacked section does not contain any energy that comes from outside the plane of recording (sideswipe). Imaging the subsurface in the third dimension is discussed in Section 6.5.

When a stacked section is migrated, we use the migration theory applicable to data recorded with a coincident source and receiver (zero-offset). To develop a conceptual framework for discussing migration of zero-offset data, we now examine two types of recording schemes. A zero-offset section is recorded by moving a single source and a single receiver along the line with no separation between them. The recorded energy follows raypaths that are normal incidence to reflecting interfaces. This recording geometry is not realizable in reality. Now consider an alternative geometry that will produce the same seismic section. Imagine exploding sources that are located along the reflecting interfaces (Loewenthal et al., 1976). Also, consider one receiver located on the surface at each CMP location along the line. The sources explode in unison and send out waves that propagate upward. The waves are recorded by the receivers at the surface. The seismic section that results from this so-called *exploding reflectors model* is largely equivalent to the zero-offset section, with one important distinction. The zero-offset section is recorded as two-way traveltime (from source to reflection point to receiver), while the exploding reflectors model is recorded as one-way traveltime (from the reflection point at which the source is located to the receiver). To make the sections compatible, we can imagine that the velocity of propagation is half the true medium velocity for the exploding reflectors model. The equivalence between the zero-offset section and the exploding reflectors model is not quite exact, particularly when strong lateral velocity variations exist along the line (Kjartansson and Rocca, 1979).

These concepts now are applied to the velocity-depth model in Figure 4-7. Visualize source-receiver pairs placed along the earth's surface at every tenth midpoint. In this case, a zero-offset section is modeled. At midpoint 130, five different arrivals are associated with rays that are normal incidence to the first interface. Alternatively, imagine receivers placed along the earth's surface at every tenth midpoint and sources placed along the interface where the rays emerge at the right angle to the interface (equivalent to the normal-incidence rays of the zero-offset section). In the latter case, the velocities indicated in Figure 4-7 must be halved to match the time axis with that associated with the zero-offset section. The interface can be sampled more densely by placing receivers and sources at closer spacing (Figure 4-8). The next deeper interface can be modeled (that is, the traveltime curve can be generated) by placing sources along this interface and leaving the receivers where they were on the surface (Figure 4-9). Finally, the same experiment can be repeated for the third interface (Figure 4-10). To

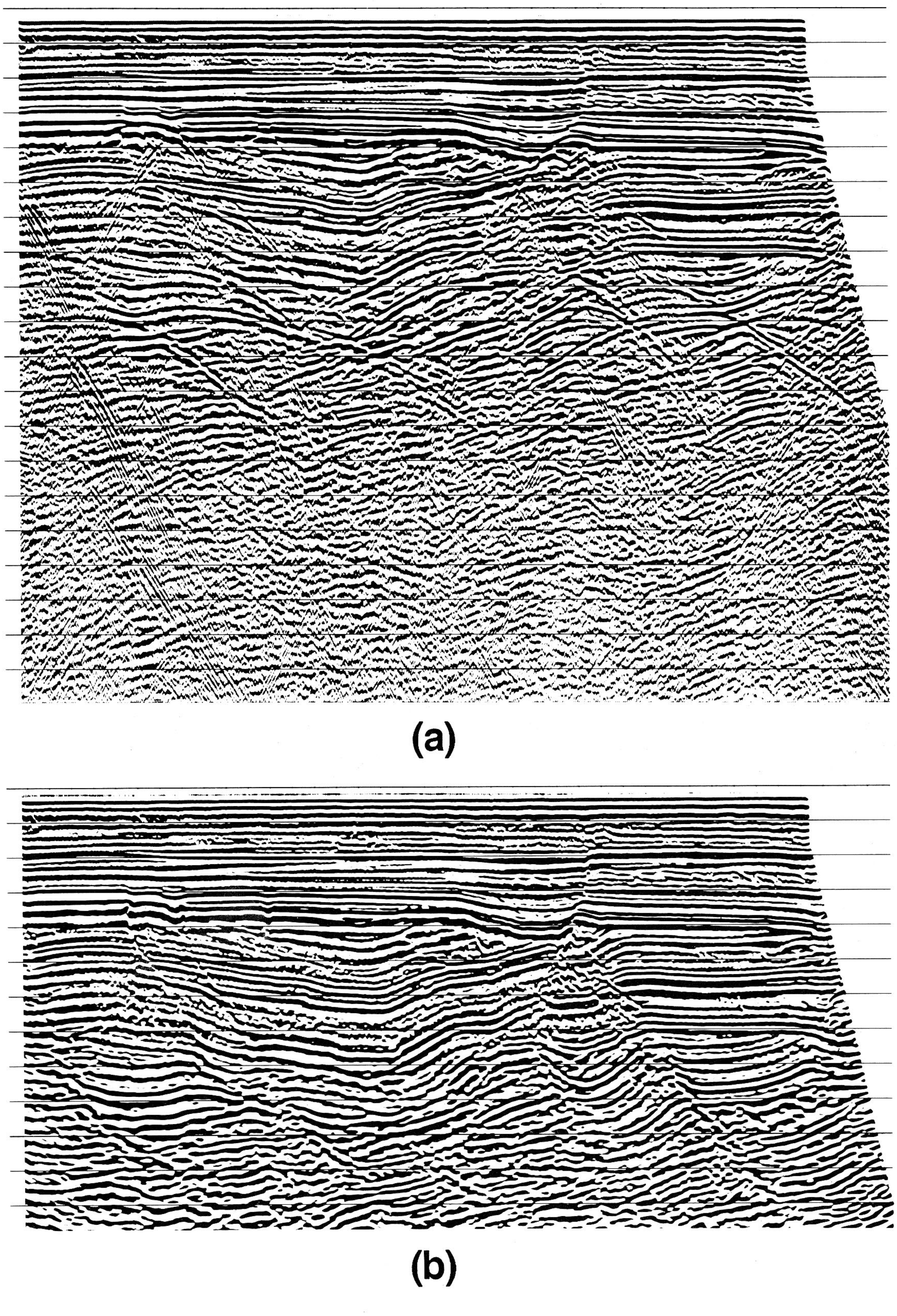


FIG. 4-3. (a) CMP stack, (b) migration. Migration collapses diffractions and delineates fault planes.

derive the composite response from the velocity-depth model in Figure 4-11a, individual responses from each interface are superimposed (those in Figures 4-8, 4-9, and 4-10). The result is shown in Figure 4-11b. We can imagine that sources were placed at all three interfaces and turned on simultaneously. Such an experiment would cause the rays emerging from the three interfaces to be recorded at receivers placed on the surface, along the line. A more familiar display of the time section equivalent to this forward model (Figure 4-11b) is shown in Figure 4-12a. The shallow complex interface (horizon 2 in Figure 4-11a) caused the complicated response of the two simple interfaces (horizons 3 and 4) in this zero-offset time section.

How valid is the assumption that a stacked section is equivalent to a zero-offset section? The conventional CMP recording geometry provides the wave field at nonzero offsets. During processing, we collapse the offset axis by stacking the data onto the midpoint-time plane at zero offset. In doing this, we normally assume hyperbolic moveout. Figure 4-13 shows selected CMP gathers modeled from the velocity-depth profile in Figure 4-11a. Because of the presence of strong lateral velocity variations, the hyperbolic assumption may not be appropriate for some reflections on some CMP gathers (Figure 4-13a); however, it may be valid for others (Figure 4-13b). We obtain a stacked section (Figure 4-12b) that resembles the zero-offset section (Figure 4-12a) to the extent that the hyperbolic moveout assumption is valid. The assumption that a conventional stacked section is equivalent to a zero-offset section also is violated to varying degrees in the presence of strong multiples and conflicting dips with different stacking velocities (Section 4.4). Although migration of unstacked data is discussed in Section 4.4, we mainly focus on migration after stack in this chapter. Table 4-1 is an overview of different migration procedures applied to different types of seismic data (2-D, 3-D, stacked, and unstacked). Table 4-1 also summarizes the circumstances necessary for invoking these migration types.

The one-way (in depth) scalar wave equation is the basis for common migration algorithms. These algorithms do not explicitly model multiple reflections, converted waves, surface waves, or noise. Any such energy present in data input to migration is treated as primary reflections. The mathematical basis of the wave field extrapolation and migration is described in Appendix C.

In this chapter, we start with a discussion of migration principles (Section 4.2). Several migration techniques then are discussed in their historical order of development as outlined below. The first migration technique developed was the semicircle superposition method that was used before the age of computers. Then came the diffraction-summation technique, which is based on summing the seismic amplitudes along a diffraction hyperbola whose curvature is governed by the medium velocity. The Kirchhoff summation technique introduced later (Schneider, 1978) basically is the same as the diffraction summation technique, with added amplitude and phase corrections applied to the data before summation. These corrections make the summation consistent with the wave equation in that they account for spherical spreading (Section 1.5), the obliquity factor (angle-dependency of amplitudes), and the phase shift inherent to Huygens' secondary sources (Section 4.2.1).

Another migration technique (Claerbout and Doherty, 1972) is based on the idea that a stacked section can be modeled as an upcoming zero-offset wave field generated by exploding reflectors. Using that model, migration can be conceptualized as consisting of wave field extrapolation (in the form of downward continuation) followed by imaging. To understand imaging, consider the shape of a wave field at observation time $t = 0$ generated by an exploding reflector. Since no time has elapsed and, thus, no propagation has occurred, the wavefront shape must be the same as the reflector shape that generated the wavefront. The fact that the wavefront shape at $t = 0$ corresponds to the reflector shape is called the *imaging principle*. To define the reflector geometry from a wave

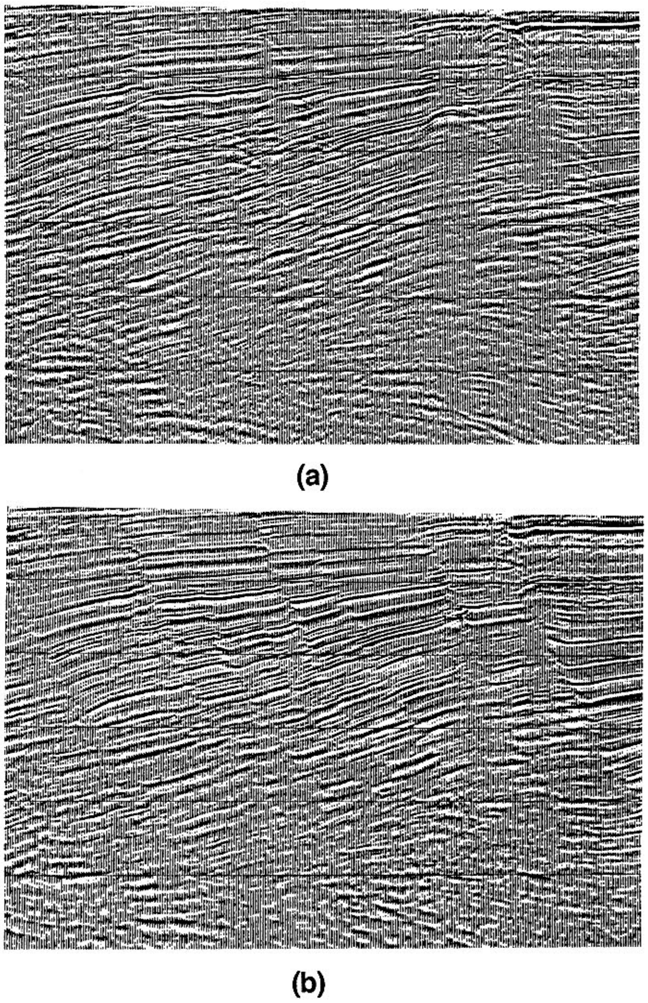


FIG. 4-4. (a) CMP stack, (b) migration. Migration collapses subtle diffractions associated with the growth faults and makes detailed structural interpretation easier. (Taner and Koehler, 1977; figure courtesy Seiscom-Delta, Inc.)

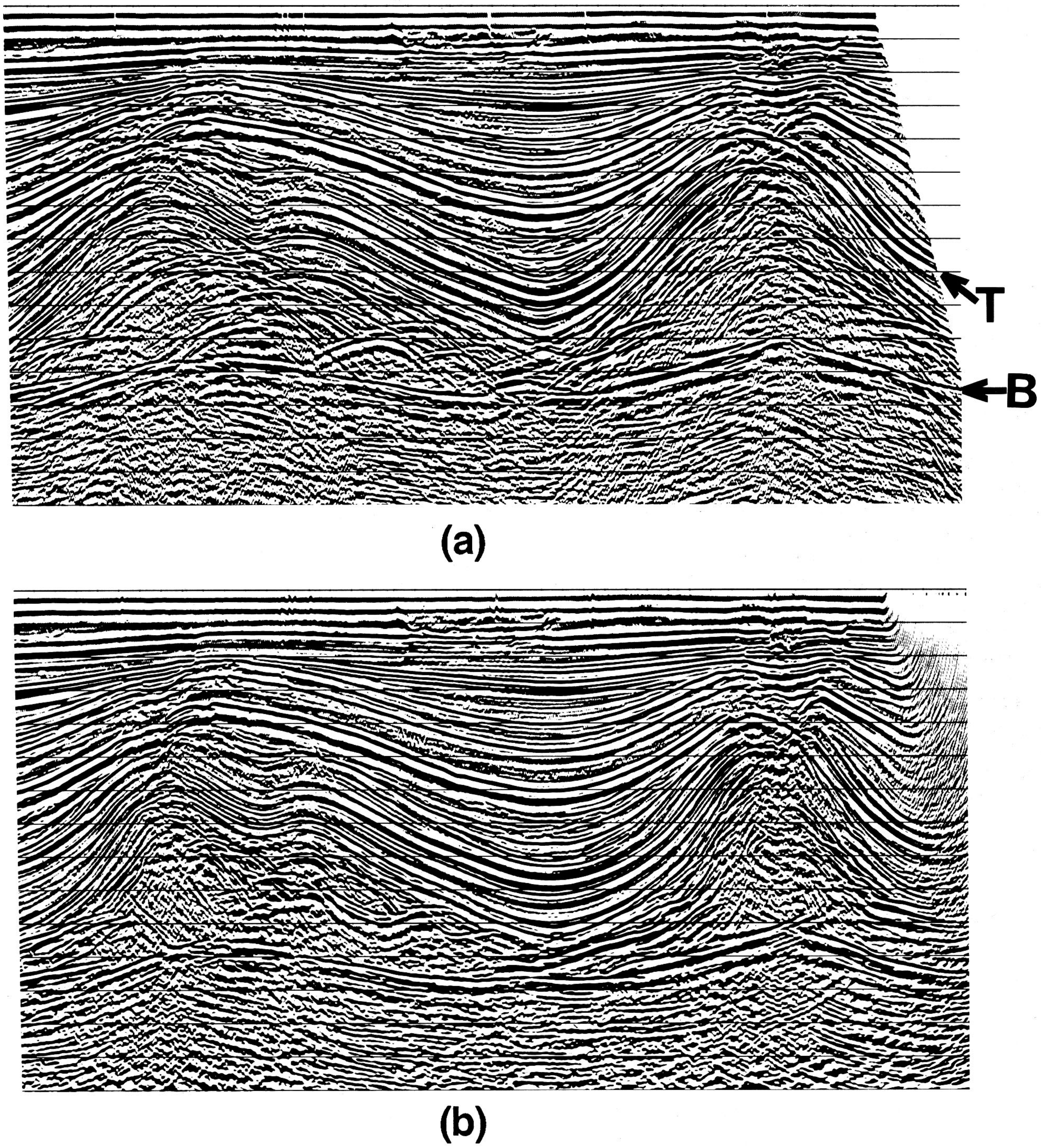


FIG. 4-5. (a) CMP stack and (b) its migration. Time migration treats the top of salt *T* properly, while it fails to image the salt base *B* accurately. Depth migration must be done to handle (*B*) (Section 5.2).

Table 4-1. Migration types.

Type	Discussion
Stack	The section the interpreter always wants.
Depth Conversion Along Vertical Raypaths	Strictly valid only for velocity that varies with depth, without structural dip.
Time Migration	Needed when the stacked section contains diffractions or structural dip. Valid for vertically varying velocity. Acceptable for mild lateral velocity variations.
Depth Migration	Needed when the stacked section contains structural dip and large lateral velocity gradients.
Prestack Partial Migration (PSPM)	Poststack migration is acceptable when the stacked section is equivalent to a zero-offset section. This is not the case for conflicting dips with different stacking velocities or large lateral velocity gradients. PSPM [Dip Moveout (DMO)] provides a better stack that can be migrated after stack. However, PSPM only solves the problem of conflicting dips with different stacking velocities.
Full-Time Migration Before Stack	The output is a migrated section. No intermediate unmigrated stacked section is produced. This often is not what the interpreter wants; he must have an unmigrated stacked section and its migrated form. Nevertheless, this is the rigorous solution to the problem of conflicting dips. PSPM is a simplification of this process.
Depth Migration Before Stack	Needed when there are extremely strong lateral velocity gradients that cannot be treated properly by stacking.
3-D Time Migration After Stack	Needed when the stack contains dipping events that are out of the profile plane (crossdips). After stack, this is the most common type of 3-D migration.
3-D Depth Migration After Stack	Needed when the problem of strong lateral velocity variation involves 3-D subsurface structural complexity.
3-D Time Migration Before Stack	Needed when PSPM fails and when the stack contains crossdips.
3-D Depth Migration Before Stack	What everyone would like to have if computer time were abundant and if the 3-D subsurface velocity model were known accurately.

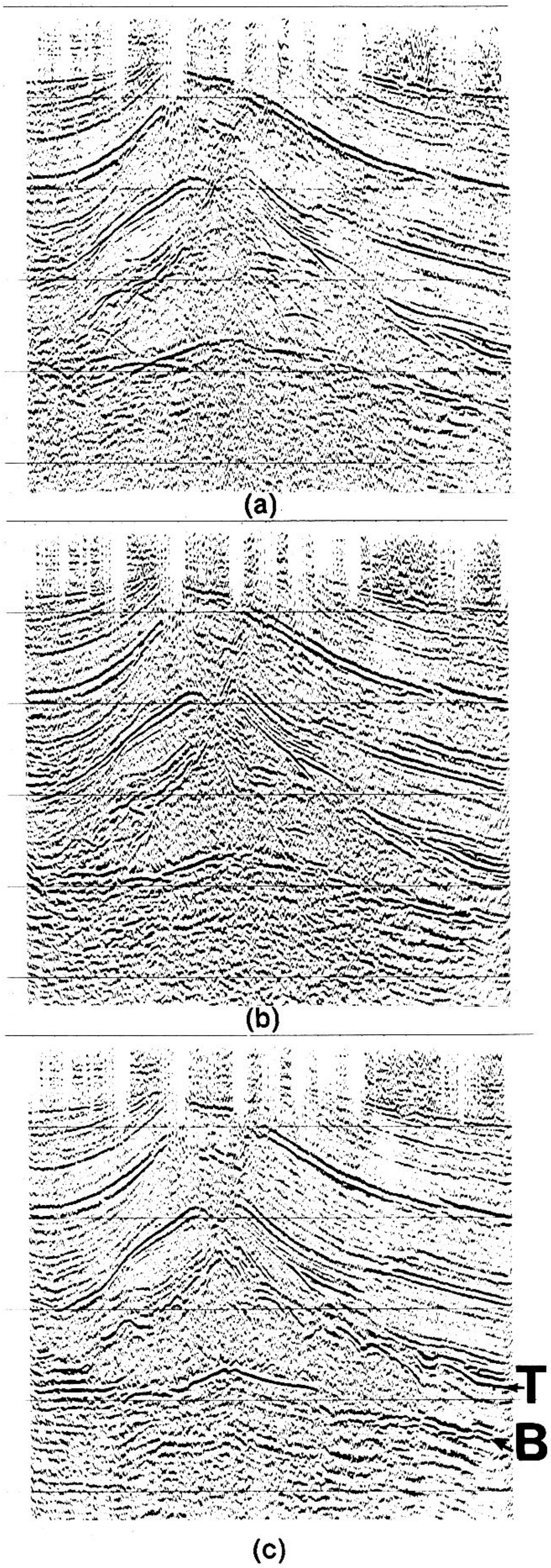


FIG. 4-6. A 2-D CMP stack (a) truly represents a 3-D wave ► field cross section. Thus, it can contain energy from outside the 2-D profile plane. A 2-D migration (b) is inadequate when this kind of energy is present on the 2-D CMP stacked section. (c) Clear imaging of the salt structure requires both 3-D data collection and 3-D migration (Chapter 6). (Data courtesy Nederlandse Aardolie Maatschappij B.V.)

field recorded at the surface, we only need to extrapolate the wave field back in depth, then monitor the energy arriving at $t = 0$. The reflector shape at any particular extrapolation depth directly corresponds to the wavefront shape at $t = 0$.

Downward continuation of wave fields can be implemented conveniently using finite-difference solutions to the scalar wave equation. Migration methods based on such implementations are called finite-difference migration (Section 4.2.2). Claerbout (1985) has provided a comprehensive theoretical foundation of finite-difference migration and its practical aspects.

After the developments on Kirchhoff summation and finite-difference migrations, Stolt (1978) introduced migration by Fourier transform (Section 4.2.3). This method involves a coordinate transformation from frequency (the transform variable associated with the input time axis) to vertical wavenumber axis (the transform variable associated with the output depth axis), while keeping the horizontal wavenumber unchanged. The Stolt method is based on a constant-velocity assumption; however, Stolt modified his method to handle the types of velocity variations for which time migration is acceptable. A recent book by Stolt and Benson (1986) combines theory with practice in the field of migration. Another frequency-wavenumber migration is the phase-shift method (Gazdag, 1978) (Section 4.2.3). This method is based on the idea that downward continuation amounts to a phase shift in the frequency-wavenumber domain. The imaging principle is invoked by summing over the frequency components of the extrapolated wave field at each depth step.

After the discussion on basic principles, many practical aspects of the Kirchhoff summation, finite difference, and frequency-wavenumber (Stolt and phase-shift) migration algorithms are considered in Section 4.3. Regardless of the migration algorithm used, the interpretability of a migrated section is dependent on the quality of the stacked section (how closely it approximates a zero-offset section), the signal-to-noise (S/N) ratio, and the velocities used in migration. Migration requires the true medium velocity. If we use a velocity model that is significantly different from the medium velocity, then the migrated section can be misleading. A procedure for estimating migration velocities is presented in Section 4.5.

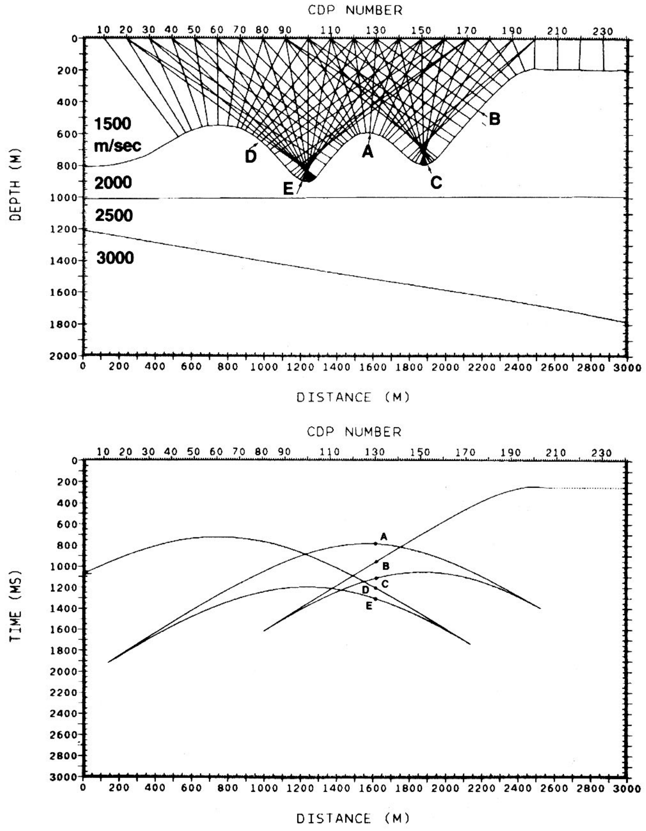


FIG. 4-7. A velocity-depth model (top) and the zero-offset response (bottom) obtained from normal-incidence rays. Note the five arrivals (*A*, *B*, *C*, *D*, and *E*) at CMP 130, all from the water bottom.

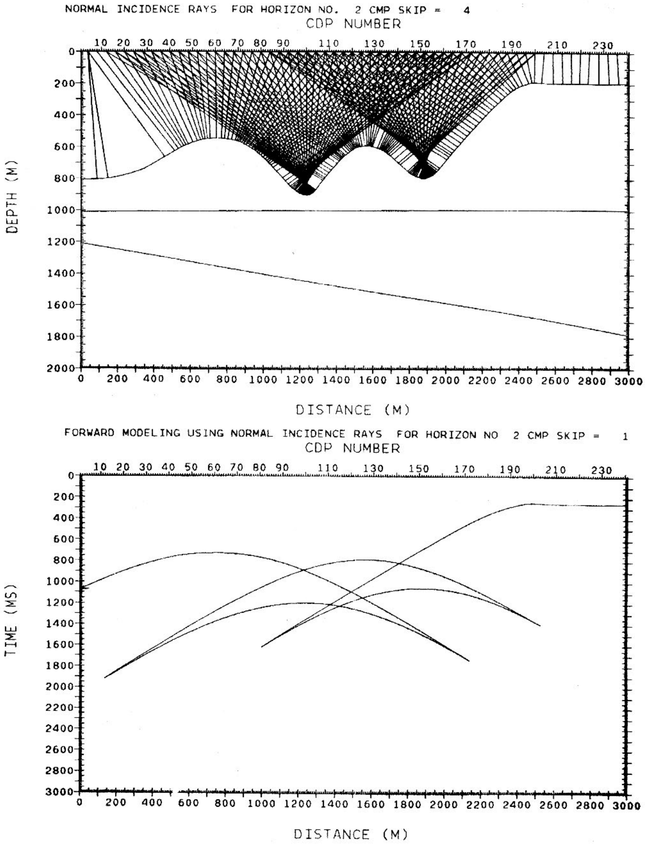


FIG. 4-8. Exploding-reflector modeling of the water bottom. The time section is equivalent to a zero-offset section; i.e., the vertical axis is in two-way time.

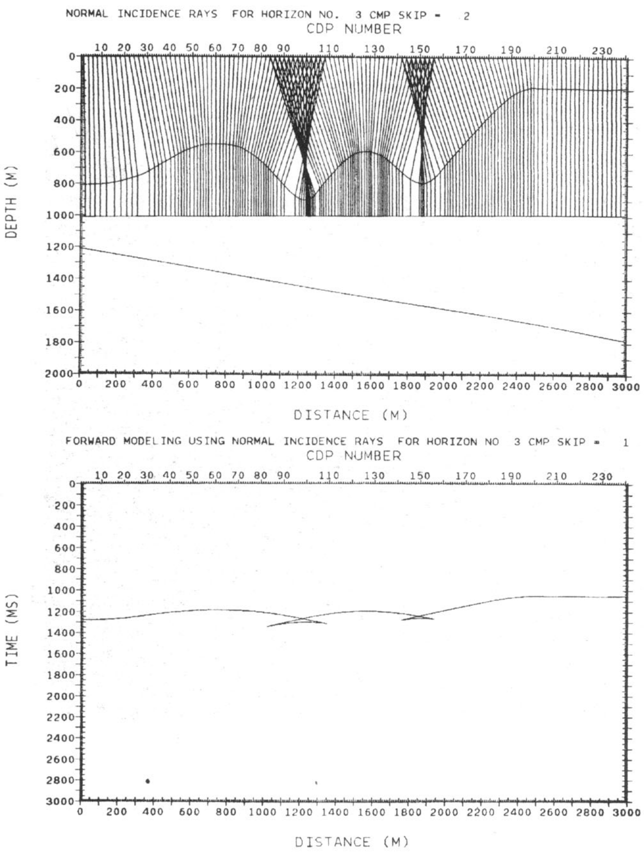


FIG. 4-9. Exploding-reflector modeling of the flat reflector. The time section is equivalent to a zero-offset section; i.e., the vertical axis is in two-way time.

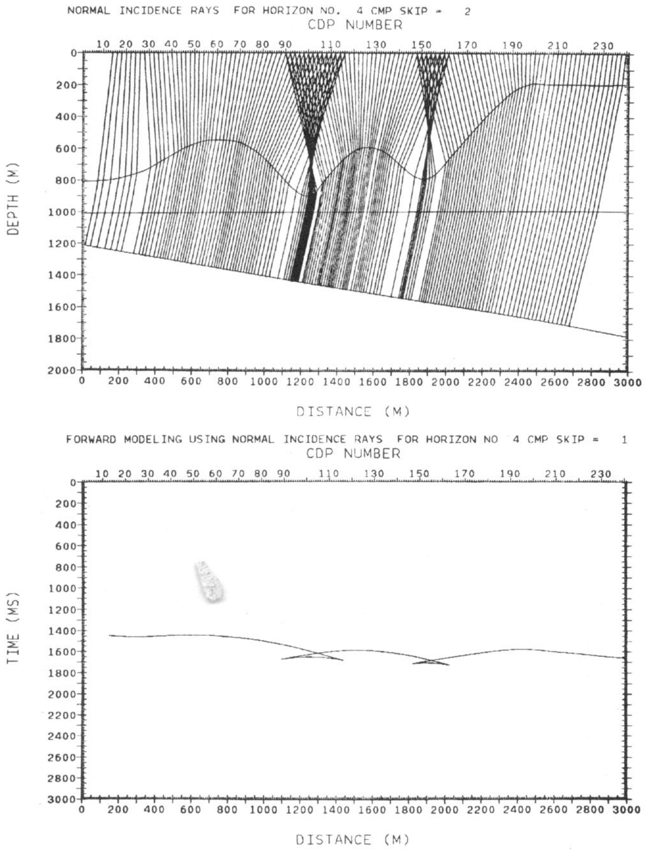


FIG. 4-10. Exploding-reflector modeling of the dipping reflector. The time section is equivalent to a zero-offset section; i.e., the vertical axis is in two-way time.

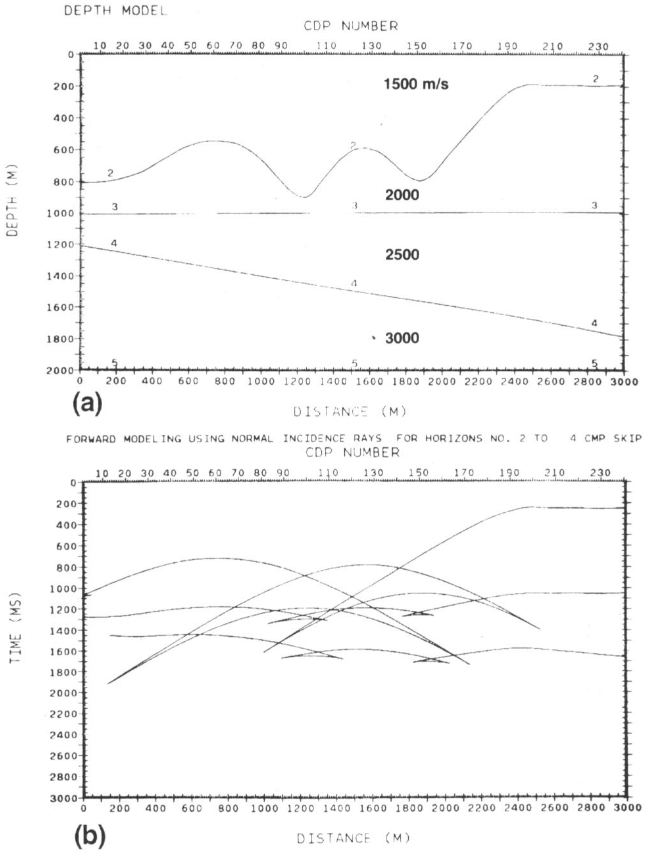


FIG. 4-11. Superposition of the normal-incidence responses from the water bottom, flat, and dipping reflectors shown in Figures 4-8, 4-9, and 4-10. Seismic modeling goes from the depth model (top) to the traveltime profile (bottom). The inverse process is seismic imaging (migration).

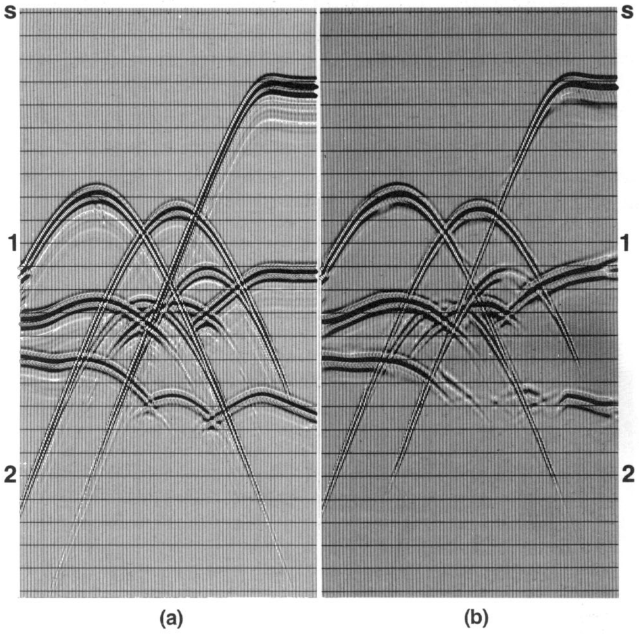


FIG. 4-12. (a) The zero-offset section equivalent to the traveltime profile in Figure 4-11b; (b) the CMP stack generated from the CMP gathers as in Figure 4-13. (Modeling by Deregowski and Barley; courtesy British Petroleum.)

4.2 MIGRATION PRINCIPLES

Consider the dipping reflector *CD* of the geologic (depth) section in Figure 4-14a. We want to obtain a zero-offset section along the profile *Ox*. As we move the source-receiver pair *(s,g)* along *Ox*, the first normal-incidence arrival from the dipping reflector is recorded at location *A*. In this discussion, we assume a constant-velocity medium $v = 1$ so that time and depth coordinates become interchangeable. The reflection arrival at location *A* is indicated by point *C′* on the zero-offset time section in Figure 4-14b. As we move from location *A* to the right, normal-incidence arrivals are recorded from the dipping reflector *CD*. The last arrival is recorded at location *B*, which is indicated by point *D′* in Figure 4-14b. (In this experiment, diffractions off the edges of reflector *CD* are excluded to simplify the discussion.) Compare the geologic section in Figure 4-14a, which is in depth, with the zero-offset seismic section in Figure 4-14b, which is in time. The true subsurface position of reflector *CD* is superimposed onto the time section for comparison. Clearly the true geologic position of reflector *CD* is not the same as the reflection event position *C′D′*.

From this simple geometric construction, note that the reflection on the time section *C′D′* must be *migrated* to its true subsurface position *CD*. The following observations can be made from this geometric description of migration (Figure 4-14):

1. The dip angle of the reflector in the geologic section is greater than in the time section; thus, migration steepens reflectors.
2. The length of the reflector, as seen in the geologic section, is shorter than in the time section; thus, migration shortens reflectors.
3. Migration moves reflectors in the updip direction.

The example in Figure 4-1 demonstrates the above observations. In particular, the dipping event *(B)* moved in the updip direction, became shorter, and steepened after migration *(A)*.

As mentioned in the previous section, conventional migration output is displayed in time, as is the input stacked section. To distinguish the two time axes, we will denote the time axis on the stacked section as *t* and the time axis on the migrated section as τ. Conversion of the time axis

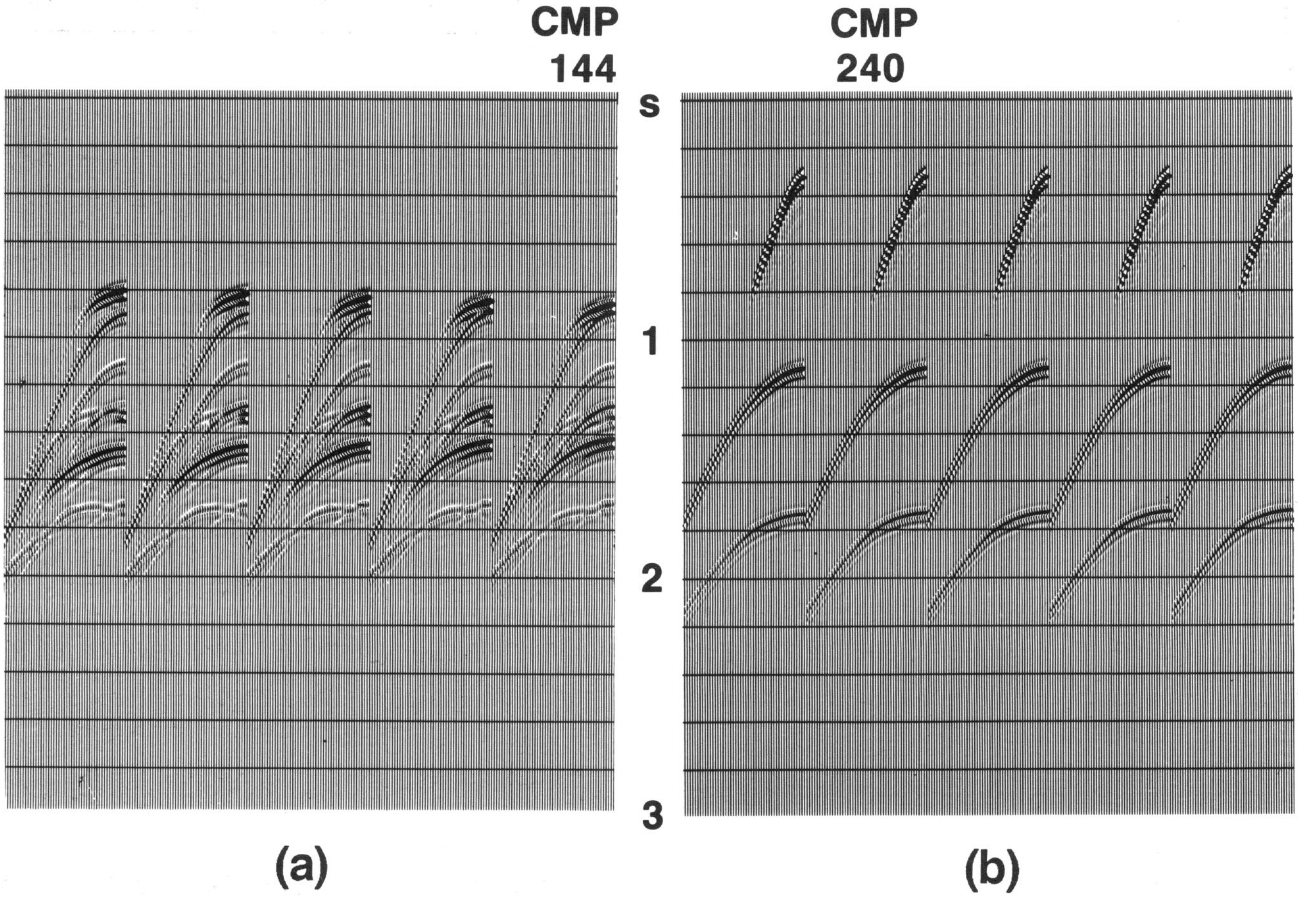


FIG. 4-13. Selected CMP gathers modeled from the velocity-depth profile in Figure 4-11a. (a) Gathers from the complex part of the depth model, (b) gathers from the simpler part of the depth model. CMP locations are indicated in Figure 4-11a. (Modeling by Deregowski and Barley; courtesy British Petroleum.)

to the depth axis is done by using the relation $z = v\tau/2$. We will examine the horizontal and vertical displacements as seen on the migrated time section. The amount of horizontal and vertical displacement that takes place in migrating dipping reflector $C'D'$ to its true subsurface position CD can be quantified. From Figure 4-15, consider a reflector segment AB. Assume that AB migrates to $A'B'$ and that point C on AB migrates to point C' on $A'B'$. The horizontal and vertical (time) displacements d_x and d_t and the dip angle after migration $\bar{\theta}_t$ (all measured on the migrated time section) can be expressed in terms of medium velocity v, traveltime t, and apparent dip of the reflector as seen on the unmigrated time section θ_t. Chun and Jacewitz (1981) derived the following formulas:

$$d_x = (v^2 t \tan \theta_t)/4 \tag{4.1}$$

$$d_t = t\{1 - [1 - (v^2 \tan^2 \theta_t)/4]^{1/2}\} \tag{4.2}$$

$$\tan \bar{\theta}_t = \tan \theta_t/[1 - (v^2 \tan^2 \theta_t)/4]^{1/2}, \tag{4.3}$$

where $\tan \theta_t = \Delta t/\Delta x$, as measured on the unmigrated time section.

To gain a better understanding of these expressions, we consider a numerical example. For a realistic velocity function that increases with depth, consider five reflecting segments at various depths. For simplicity, assume that quantity $\Delta t/\Delta x$ is the same for all (10 ms per 25-m trace spacing). From the expressions in equations (4.1), (4.2), and (4.3), compute the horizontal and vertical displacements d_x and d_t and the dips (in ms/trace) after migration. The results are summarized in Table 4-2.

Table 4-2. Horizontal and vertical displacements of points on dipping reflectors at various depths and changes in dip angle as measured on a time section as a result of migration.

t (s)	v (m/s)	d_x (m)	d_t (s)	θ_t (ms/trace)	$\bar{\theta}_t$ (ms/trace)
1	2500	625	0.134	10	11.5
2	3000	1800	0.400	10	12.5
3	3500	3675	0.858	10	14.0
4	4000	6400	1.600	10	16.7
5	4500	10125	2.820	10	23.0

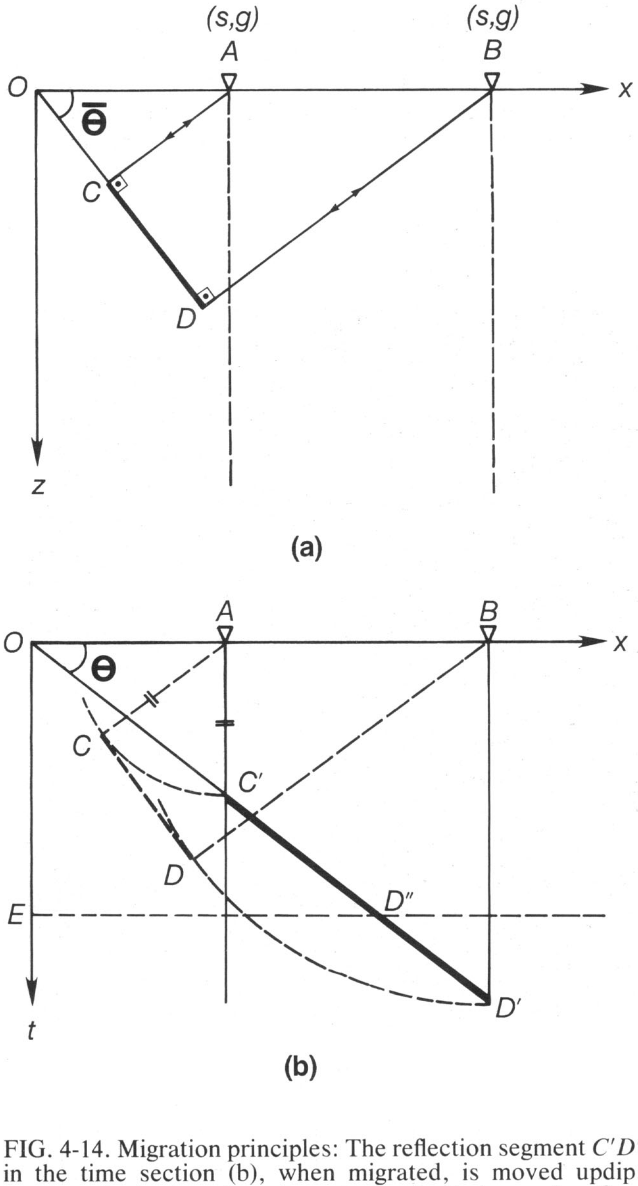


FIG. 4-14. Migration principles: The reflection segment $C'D'$ in the time section (b), when migrated, is moved updip, steepened, shortened, and mapped onto its true subsurface location CD (a). (Adapted from Chun and Jacewitz, 1981.)

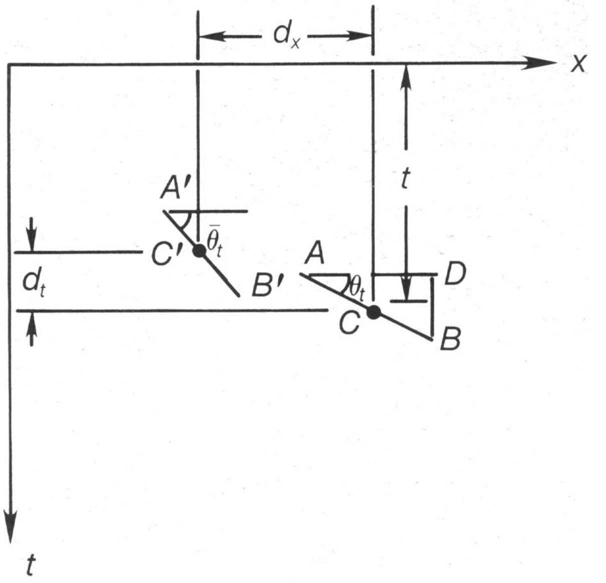


FIG. 4-15. Quantitative analysis of the migration process. Point C on dipping reflector AB is moved to C' after migration. The amount of horizontal displacement d_x, vertical displacement d_t, and dip angle after migration $\bar{\theta}_t$ is calculated from equations (4.1), 4.2), and (4.3). Here, $\Delta x = AD$, $\Delta t = BD$.

Note that the dips after migration are greater than the dip before migration. The deeper the event, the more migration takes place; e.g., at 4 s, the horizontal displacement is more than 6 km and the vertical displacement is 1.6 s. From equation (4.1), note that the horizontal displacement d_x increases with the time of event t. Moreover, d_x is a function of the velocity squared. If there is a 20 percent error in the velocity used in migration, then the event is misplaced by an error of 44 percent. Vertical displacement d_t also increases with time [equation (4.2)].

In Figure 4-14, assume that the zero-offset section was recorded only between surface locations A and B. As a result of migration, we see that dipping event $C'D'$ migrates out of the recorded section. Therefore, the data on a stacked section are not necessarily confined to the subsurface below the seismic line. The converse is even more important; the structure below the seismic line may not be recorded on the seismic section. In areas with a structural dip, line length must be chosen by considering the horizontal displacements of dipping events from the structures causing the events. This is an important consideration, especially in 3-D seismic work. The areal surface coverage of a survey usually is larger than the areal subsurface coverage of interest. Additionally, the recording time must be long enough. For example, if only OE seconds were recorded (Figure 4-14), then the recorded segment $C'D''$ would yield only part of the complete image CD. An excellent example of recording deeper in time and with longer line length for steeper dips is shown in Figure 4-1. Proper imaging of the salt dome boundary required that data be recorded for more than 6 s.

The migration concepts described above are demonstrated further by the dipping events model in Figure 4-16. The edge diffractions are included here. The dipping reflectors on the zero-offset section are steepened, shortened, and moved in the updip direction as a result of migration. Note that the steeper the dip, the more the event moves during migration.

So far, only linear reflectors were considered. We now must consider a more realistic geologic situation that involves curved reflecting interfaces (Figure 4-17). Here we have three synclines and a small anticlinal feature. The synclines appear as bowties on the zero-offset section. By using the principles deduced from the geometry of Figure 4-14, note that as a result of migration, segment A of the bowtie moves in the updip direction to the left. Similarly, segment B moves to the right, while flat-topped segment C does not move much at all. Consequently, after migration the flanks of bowties associated with synclines are opened up. On the other hand, the small anticline seems to be broader on the zero-offset section than it is on the migrated section. Again note that segment D moves updip to the right, while segment E moves updip to the left as a result of migration. Thus, synclines broaden and anticlines compress as a result of migration. Migration velocities also affect the apparent size of the structure; higher velocities mean more migration and, hence, smaller anticlinal structure.

Why does a syncline look like a bowtie on the stacked section? The answer is in Figure 4-18, where a symmetric syncline is modeled. Given the subsurface picture in Figure 4-18a, the normal-incidence rays can be computed to derive the zero-offset section in Figure 4-18b. Only five CMP locations are shown for clarity. At locations 2 and 4, there are two distinct arrivals, while at location 3, there are three distinct arrivals. By filling in the intermediate raypaths, the bowtie character of the syncline can be constructed on the time section. Complete the procedure to trace the traveltime curve in Figure 4-18b.

Two field data examples containing synclinal and anticlinal structures are shown in Figures 4-19 and 4-20. In Figure 4-19, note that the synclinal feature broadens and the anticlinal feature narrows as a result of migration. In Figure 4-20, the bowties associated with two small synclinal basins A and B grow larger in depth. After migration, the bowties are untied and the synclines are delineated.

4.2.1 Kirchhoff migration

Claerbout (1985) uses the harbor example in Figure 4-21 to describe the physical principles of migration. Assume that a storm barrier exists at some distance z_3 from the beach and that there is a gap in the barrier. Imagine a calm afternoon breeze that comes from the ocean as a plane incident wave. The wavefront is parallel to the storm barrier. As we walk along the beach line, we see a wavefront different from a plane wave. The gap on the storm barrier has acted as a secondary source and generated the semicircular wavefront that is propagating toward the beach. If we did not know about the storm barrier and the gap, we might want to lay out a receiver cable along the beach to record the approaching waves. This experiment is illustrated in Figure 4-22 with the recorded time section. Physicists call the gap on the barrier a *point aperture*. It is somewhat similar to a *point source*, since both generate circular wavefronts. However, the amplitudes on the wavefront that propagate outward from a point source are isotropic, while those from a point aperture are angle-dependent. The point aperture on the barrier acts as a *Huygens' secondary source*.

From the beach experiment, we find that Huygens' secondary source responds to a plane incident wave and generates a semicircular wavefront in the (x,z) plane. The response in the (x,t) plane is the diffraction hyperbola shown in Figure 4-23.

Imagine that the subsurface consists of points along each reflecting horizon that behave much as the gap on the storm barrier. From Figure 4-24, these points act as Huygens' secondary sources and produce hyperbolic traveltime curves. Moreover, as the sources (the points on the reflecting interface) get closer to each other, superposition of the hyperbolas produces the response of the actual reflecting interface (Figure 4-25). In terms of the harbor example, this is like assuming that the barrier is wiped out by a storm so that the primary incident plane wave reaches the beach without modification. The diffraction hyperbolas, which are due to sharp discontinuities at both ends of the reflector in Figure 4-25, remain.

(Text continued on page 258)

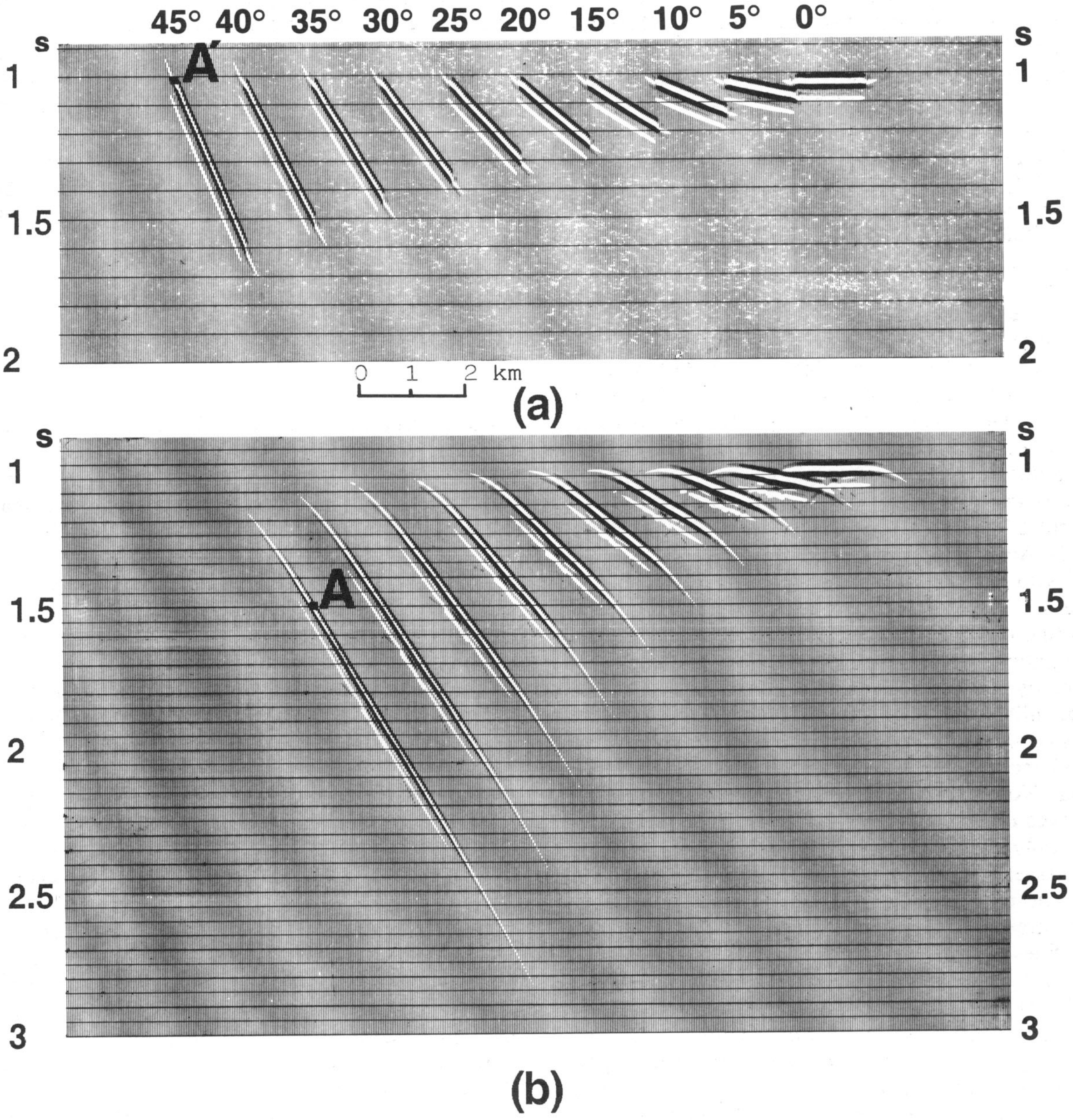


FIG. 4-16. (a) Dipping reflectors on a subsurface model (constant medium velocity, 3000 m/s) and (b) on a zero-offset section (trace spacing is 25 m). Migration of the zero-offset section yields the subsurface model. Points *A* and *A′* are referenced in Exercise 4.2.

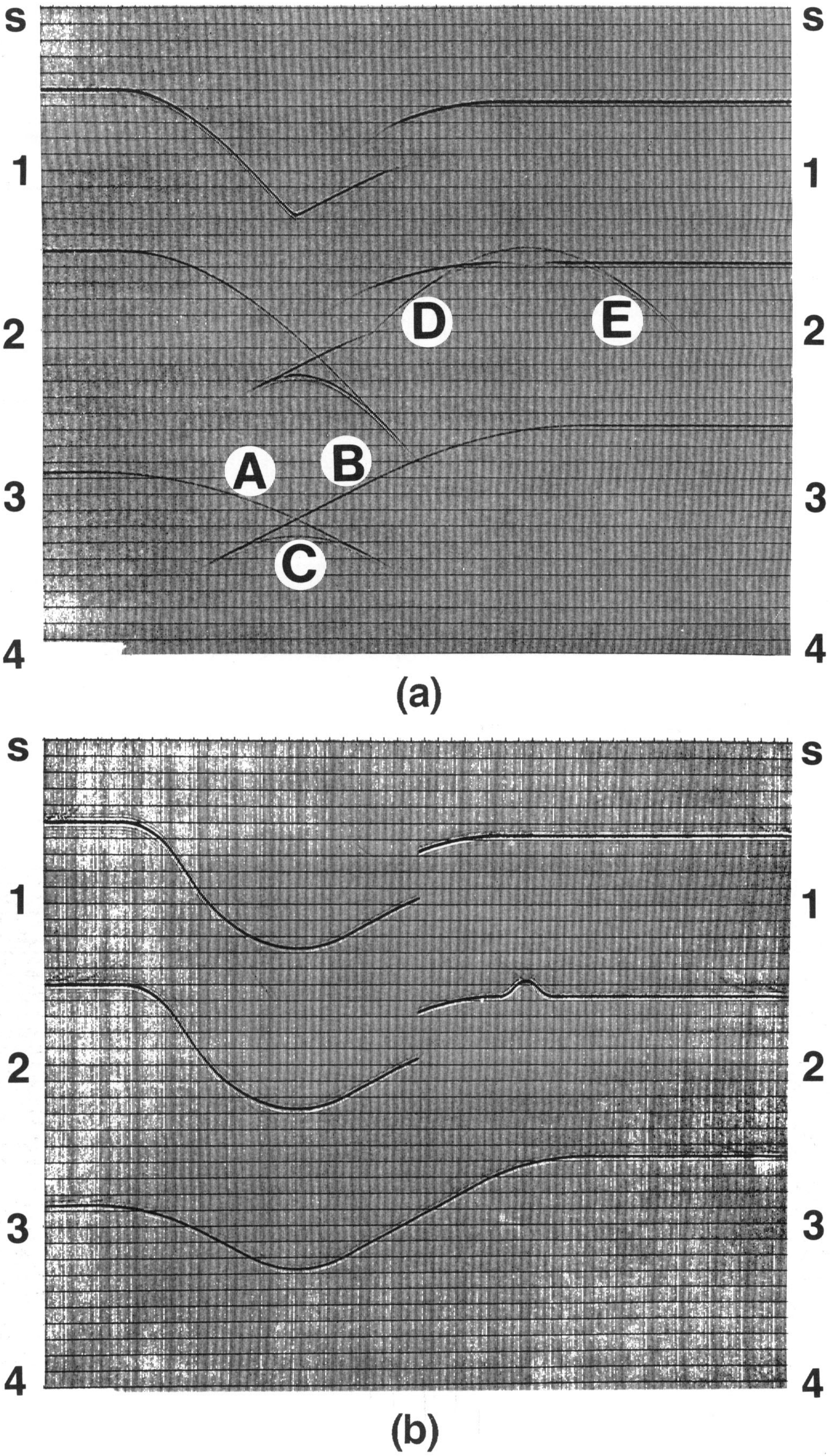


FIG. 4-17. Curved reflecting interfaces (synclines and anticlines) (a) before and (b) after migration. See text for details. (Modeling courtesy Union Oil Company.)

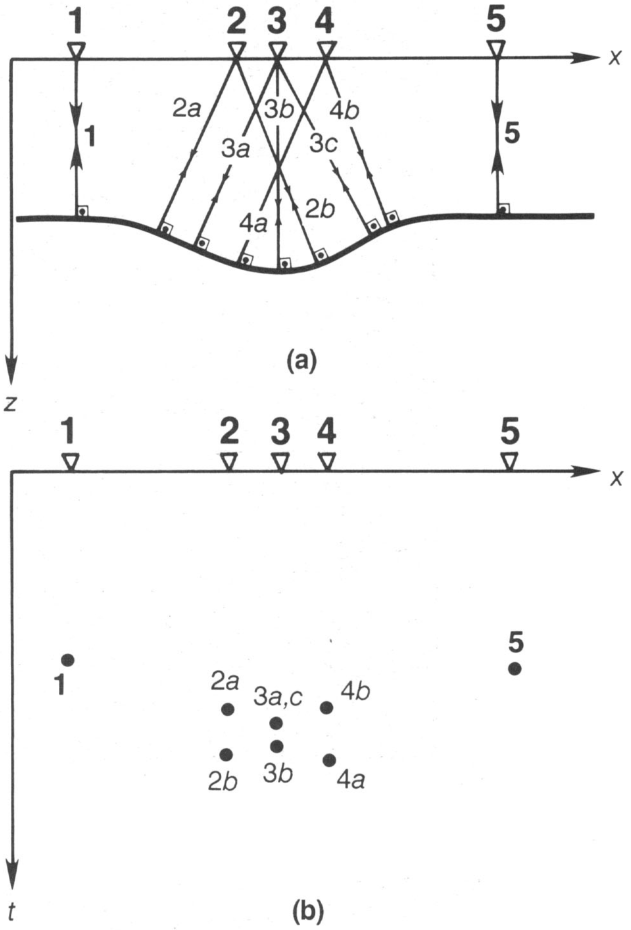


FIG. 4-18. (a) A depth model consisting of a synclinal reflector; (b) the corresponding traveltime section. Trace the bowtie in the time section.

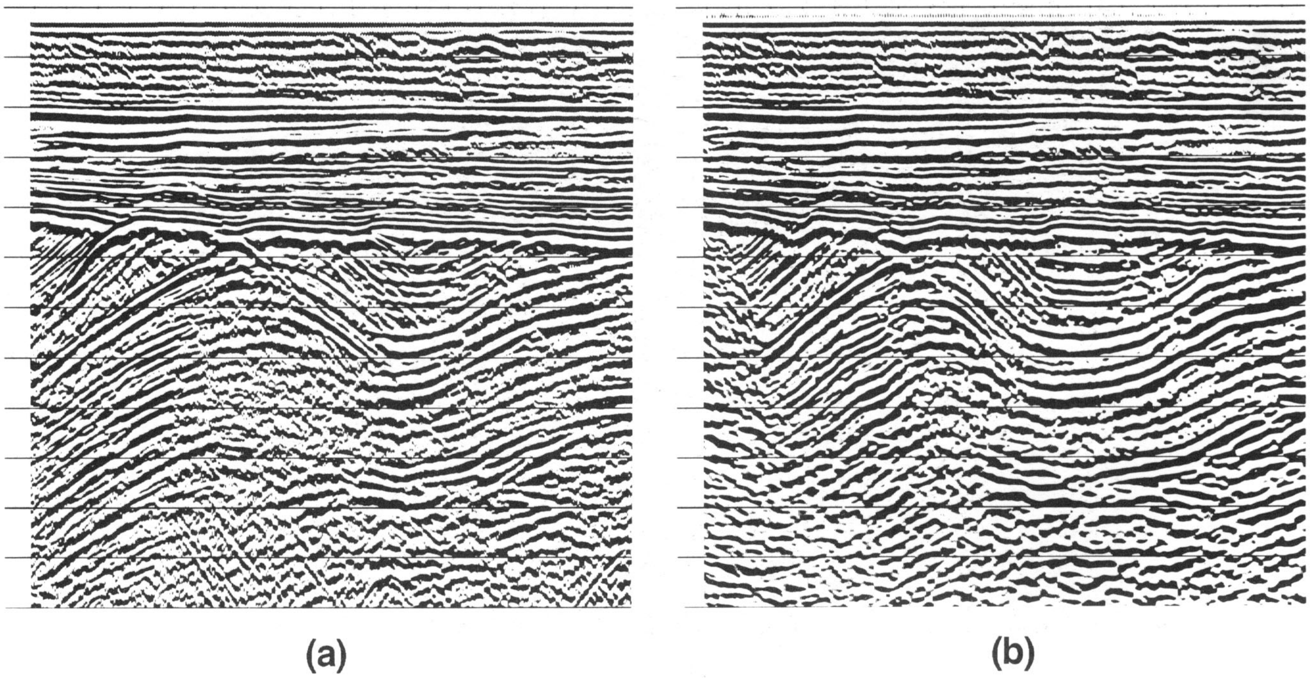


FIG. 4-19. (a) CMP stack, (b) migration. Anticlines appear bigger, while synclines appear smaller than their actual sizes on the unmigrated section (a).

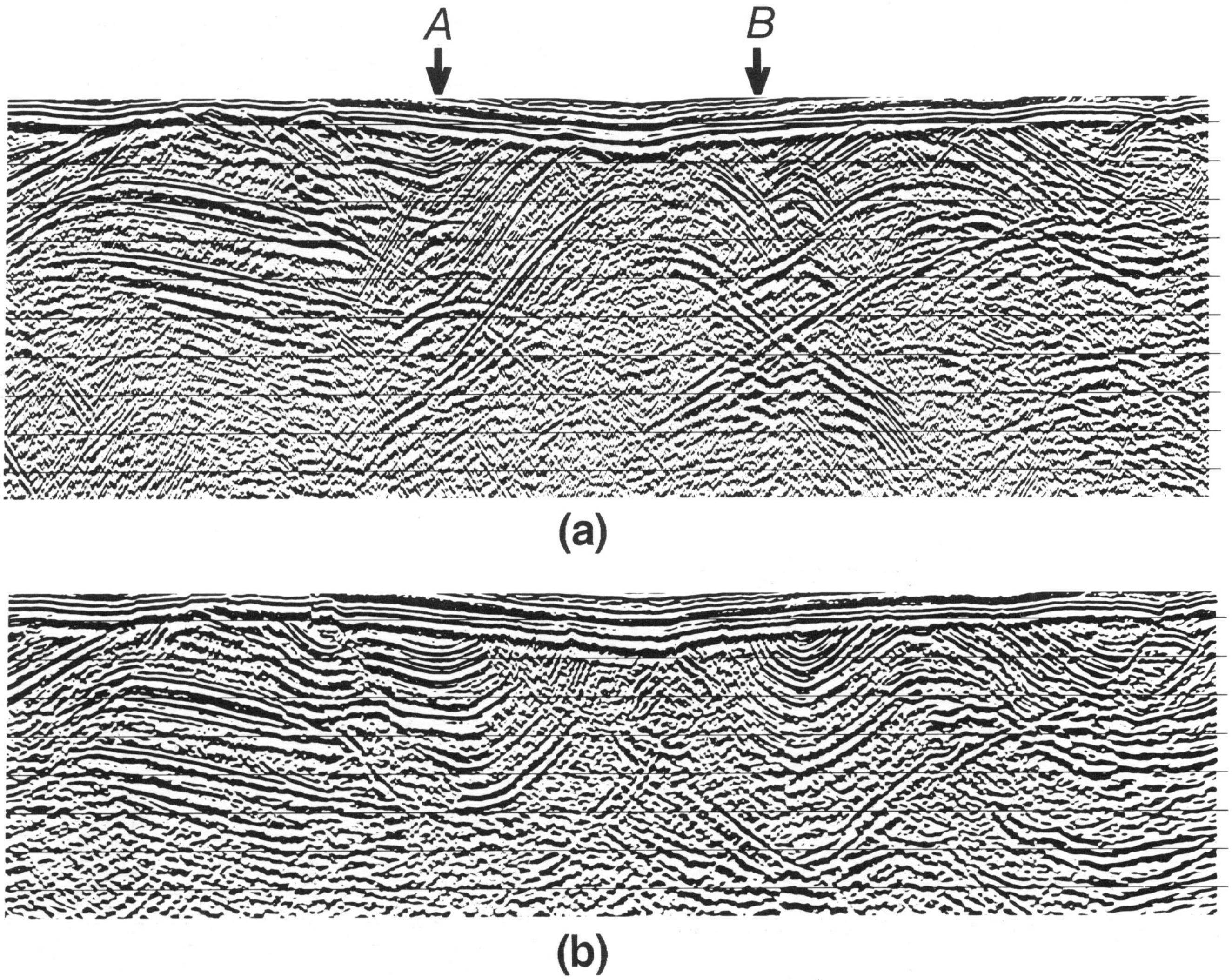


FIG. 4-20. (a) CMP stack, (b) migration. Migration unties the bowties and turns them into synclines below *A* and *B*.

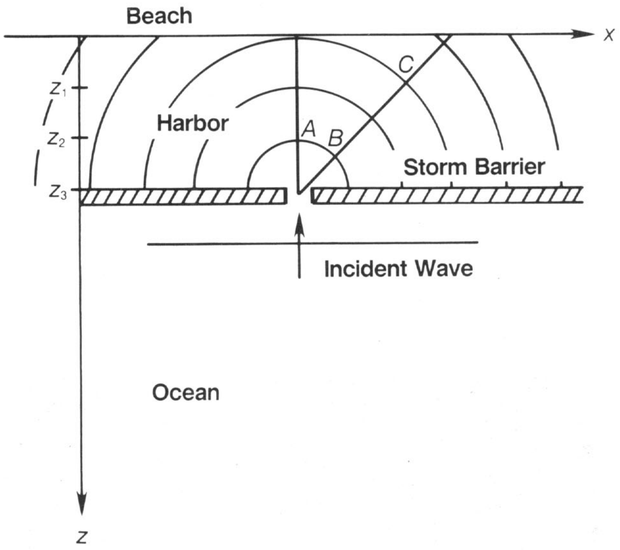


FIG. 4-21. The gap in the barrier acts as Huygens' secondary source, causing the circular wavefronts that approach the beach line. (Adapted from Claerbout, 1985.)

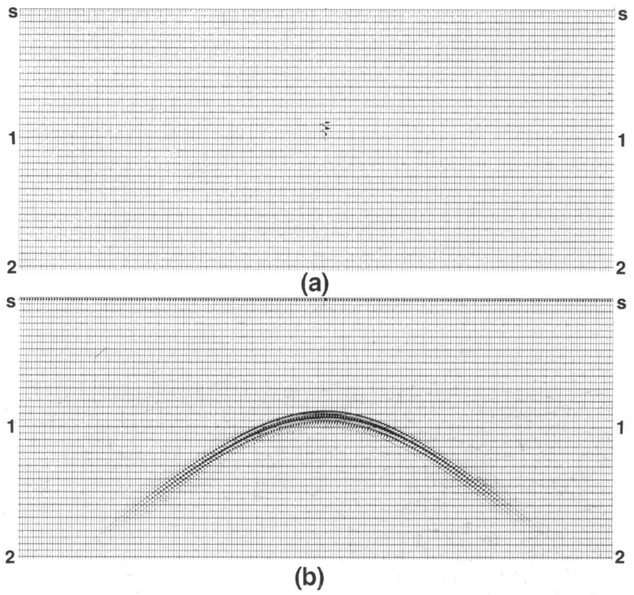


FIG. 4-23. A point that represents a Huygens' secondary source in the depth section (a) maps onto a diffraction hyperbola on the zero-offset time section (b). The vertical axis in this section is two-way time, while the vertical axis in the time section in Figure 4-22 is one-way time.

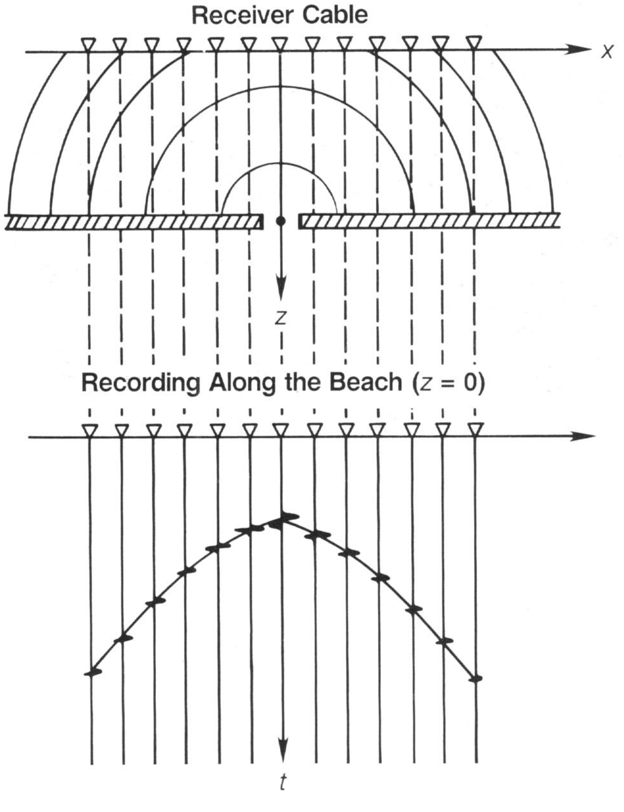


FIG. 4-22. Waves recorded along the beach generated by Huygens' secondary source (the gap in the barrier in Figure 4-21) have a hyperbolic traveltime trajectory.

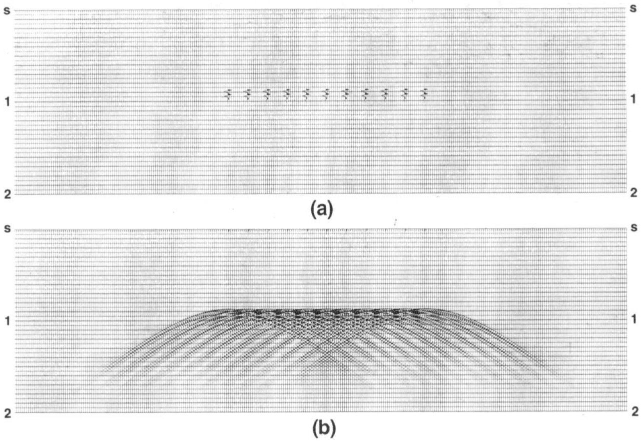


FIG. 4-24. Superposition of the zero-offset response (b) of a discrete number of Huygens' secondary sources in the depth section (a).

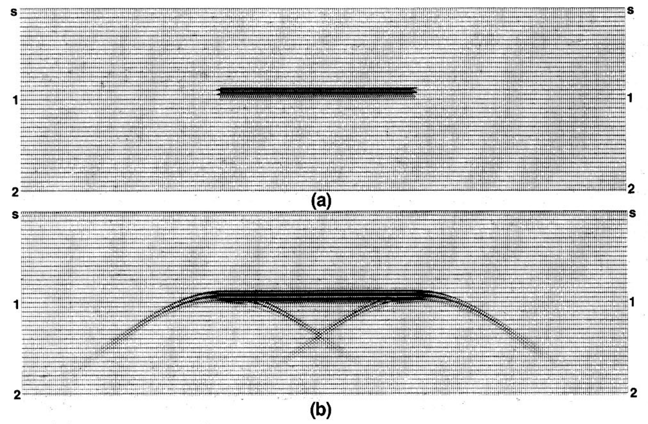


FIG. 4-25. Superposition of the zero-offset response (b) of a continuum of Huygens' secondary sources in the depth section (a).

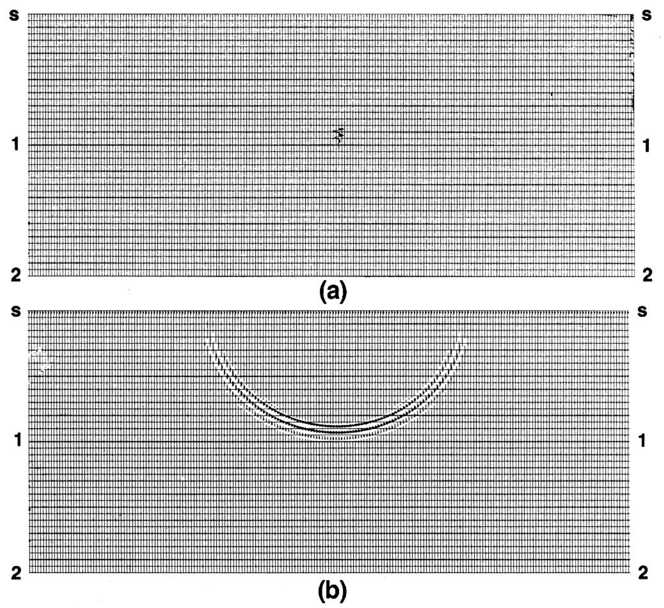


FIG. 4-26. Principles of migration based on semicircle superposition. (a) Zero-offset section (trace interval, 25 m; constant velocity, 2500 m/s), (b) migration. A point in time section (a) maps onto a semicircle in depth section (b).

These hyperbolas are equivalent to diffractions seen at fault boundaries on stacked sections.

In summary, we find that reflectors in the subsurface can be visualized as being made up of many points that act as Huygens' secondary sources. We also find that the zero-offset section consists of superposition of the many hyperbolic traveltime responses. Moreover, when there are discontinuities (faults) along the reflector, diffraction hyperbolas often stand out.

Huygens' secondary source signature is a semicircle in the (x,z) plane and a hyperbola in the (x,t) plane. This characterization of point sources in the subsurface leads to two practical migration schemes. Figure 4-26a shows a zero-offset section that consists of a single arrival at a single trace. This event migrates to a semicircle (Figure 4-26b). From Figure 4-26, note that the zero-offset section recorded over a constant-velocity earth model consisting of a semicircular reflecting interface contains a single blip of energy at a single trace as in Figure 4-26a. Since this recorded section consists of an impulse, the migrated section in Figure 4-26b can be called the *migration impulse response*. An alternate scheme for migration results from the observation that a zero-offset section consisting of a single diffraction hyperbola migrates to a single point (Figure 4-27b).

The first method of migration is based on the superposition of semicircles, while the second method is based on summation of amplitudes along hyperbolic paths. The first method was used before the age of digital computers. The second method, which is known as the *diffraction summation method*, was the first computer implementation of migration. The migration scheme based on the semicircle superposition consists of mapping the amplitude at a sample in input (x,t) space of the unmigrated time section onto a semicircle in output (x,z) space. The migrated section is formed as a result of superposition of the many semicircles. The migration scheme based on diffraction summation consists of searching the input data in (x,t) space for energy that would have resulted if a diffracting source (Huygens' secondary source) were located at a particular point in the output (x,z) space. This search is carried out by summing the amplitudes in (x,t) space along the diffraction curve that corresponds to Huygens' secondary source at each point in the (x,z) space. The result of this summation then is mapped onto the corresponding point in the (x,z) space.

Diffraction summation is a straightforward summation of amplitudes along the hyperbolic trajectory whose curvature is governed by the velocity function. The equation for this trajectory can be derived from the geometry of Figure 4-27. Assuming a horizontally layered velocity model, the velocity function used to compute the traveltime trajectory is the rms velocity at the apex of the hyperbola at time $t(0)$ (see Section 3.2.1). From the triangle COA in the upper frame,

$$t^2(x) = t^2(0) + 4x^2/v_{\mathrm{rms}}^2 . \tag{4.4}$$

Having computed the input time $t(x)$, the amplitude at input location B is placed on the output section at location A, corresponding to the output time $\tau = t(0)$ at

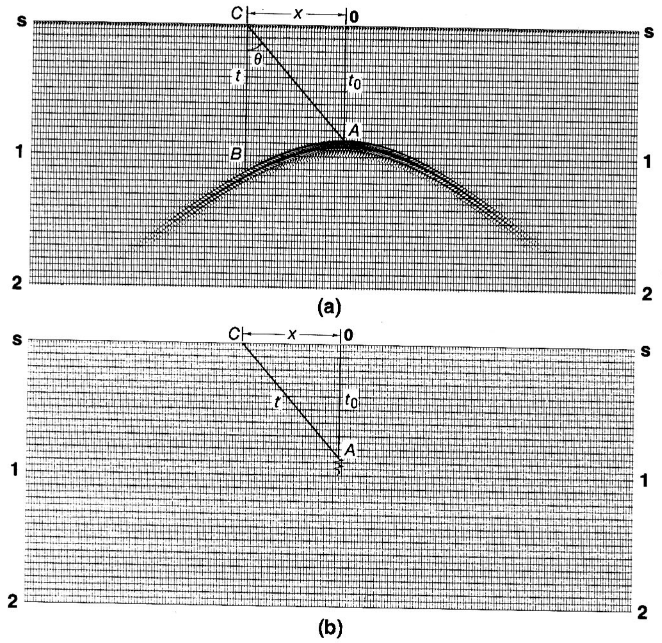


FIG. 4-27. Principles of migration based on diffraction summation. (a) Zero-offset section (trace interval, 25 m; constant velocity, 2500 m/s), (b) migration. The amplitude at *B* along the flank is mapped onto apex *A* by hyperbolic traveltime equation (4.4).

the apex of the hyperbola. From Section 3.2.1, reflection traveltimes in a layered earth approximate small-spread hyperbolas. This may seem to impose a serious restriction on the aperture width (i.e., the lateral extent of the diffraction hyperbola) in the summation process. However, the small-spread approximation is valid even at large distances from the apex and the errors associated with it are insignificant at late times. In practice, this approximation almost never is an issue.

Now consider several factors associated with the amplitude and phase behavior of the waveform along the diffraction hyperbola. From Figure 4-21, given the alternative of standing at location *A* or *B*, we intuitively think that it is safer to stand at location *B*. This is because the wave amplitude at location *A*, which is on the *z*-axis, is stronger than the wave amplitude at location *B*, which is at an oblique angle from the *z*-axis. This is one difference between a point source with uniform amplitude response at all angles and the point aperture that produces a wavefront with angle-dependent amplitudes. This angle dependence of amplitudes should be considered before summation. To perform this *obliquity factor* correction, the amplitude at location *B* in Figure 4-27 is scaled by the cosine of the angle between *BC* and *CA* before it is placed at output location *A*.

Another factor is the spherical divergence of wave amplitudes. Again, from Figure 4-21, given the alternative of standing at location *B* or *C*, we prefer to stand at location *C*. The reason for this is that the wave amplitude along the wavefront at location *C*, which is farther from the point aperture source, is weaker than the wave amplitude at location *B*. Wave energy decays as $(1/r^2)$, where r is the distance from the source to the wavefront, while amplitudes decay as $(1/r)$. Thus, amplitudes must be scaled by factor $(1/r)$ before summation for wave propagation in three dimensions.

Finally, there is a third factor that involves the inherent property of Huygens' secondary source waveform. This factor is difficult to explain from a physical viewpoint. Nevertheless, it is obvious from Figure 4-25 that Huygens' secondary sources must respond as a wavelet along the hyperbolic paths with a certain phase and frequency characteristic. Otherwise, there would be no amplitude cancellation when they are close to one another. The waveform that results from the summation must be restored in both phase and amplitude.

In summary, we must consider the following three factors before diffraction summation:

1. The obliquity factor or the directivity factor, which describes the angle dependence of amplitudes and is given by the cosine of the angle between the direction of propagation and the vertical axis z.
2. The spherical spreading factor, which is proportional to $(1/vr)^{1/2}$ for 2-D wave propagation, and $(1/vr)$ for 3-D wave propagation.
3. The wavelet shaping factor, which is designed with a 45-degree constant phase spectrum and an amplitude spectrum proportional to the square root of frequency for 2-D migration. For 3-D migration, the phase shift is 90 degrees and the amplitude is proportional to frequency.

The diffraction summation method of migration, which incorporates these three factors, is called the Kirchhoff migration. To perform this method, multiply the input data by the obliquity and spherical spreading factors. Then apply the filter with the above specifications and sum along the hyperbolic path that is defined by equation (4.4). Place the result on the migrated section at time $\tau = t(0)$ corresponding to the apex of the hyperbola. In practice, the order of the filter application (specified by factor 3) and summation can be interchanged without sacrificing accuracy because the summation is a linear process and the filter is independent of time and space. The velocity used in equation (4.4) is taken as the rms velocity, which can be allowed to vary laterally. However, lateral variation in velocity distorts the hyperbolic nature of the diffraction pattern and somehow must be considered. The value for the rms velocity typically is that of the output time sample; that is, the apex of the hyperbola.

What was determined from a physical point of view in the foregoing discussion can be rigorously described by the integral solution to the scalar wave equation. Berkhout (1980), Schneider (1978), and Berryhill (1979) are excellent references for the mathematical treatment of the Kirchhoff migration method. The integral solution of the scalar wave equation gives the output wave field $P_{\text{out}}\ (x,z,t)$ at a subsurface location (x,z) from the zero-offset wave field $P_{\text{in}}\ (x_{\text{in}},z = 0,t)$,

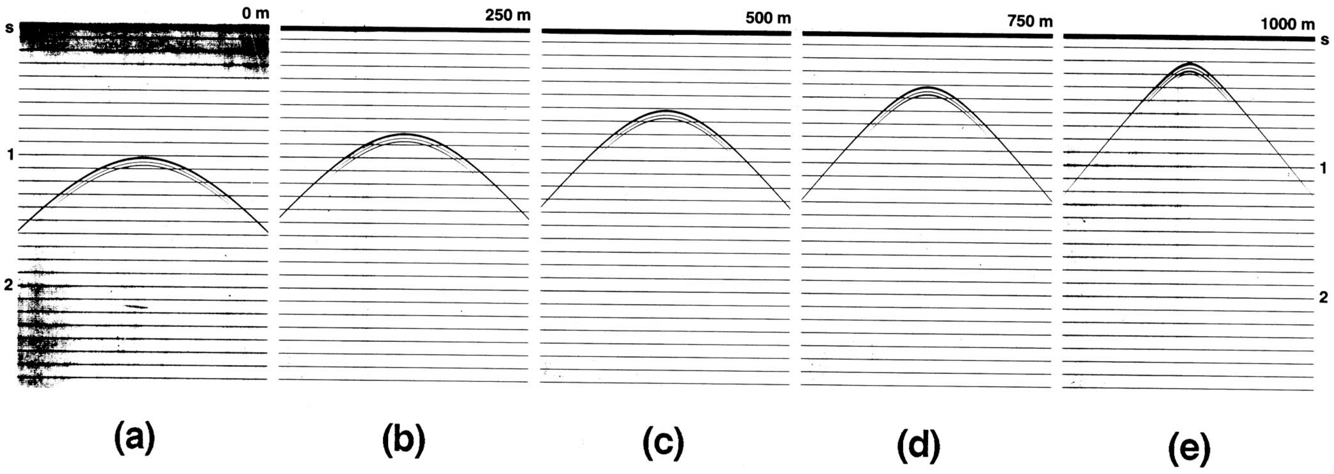


FIG. 4-28. Moving the receiver cable in the harbor experiment (Figure 4-21) from the beach into the water at discrete intervals parallel to the beach line. Numbers on top indicate the distance of the receiver cable from the beach line.

which is measured at the surface ($z = 0$). The integral solution used in seismic migration has two terms:

$$P_{\text{out}}(x, z, t) = \frac{1}{2\pi}\int dx\left[\frac{\cos\theta}{r^2}P_{\text{in}}(x_{\text{in}}, z = 0, t - r/v) + \frac{\cos\theta}{vr}\frac{\partial}{\partial t}P_{\text{in}}(x_{\text{in}}, z = 0, t - r/v)\right], \qquad (4.5)$$

where v is the rms velocity at the output point (x,z) and $r = [(x_{\text{in}} - x)^2 + z^2]$, which is the distance between the input (x_{in}, $z = 0$) and the output (x,z) points. Equation (4.5) can be used to compute the wave field at any depth z. To obtain the migrated section at an output time τ, equation (4.5) must be evaluated at $z = v\tau/2$ and the imaging principle must be invoked by mapping amplitudes of the resulting wave field at $t = 0$ onto the migrated section at output time τ. The complete migrated section is obtained by performing the integration [equation (4.5)] and setting $t = 0$ for each output location. Note that the integration is over the surface of measurement; for two dimensions, this is along the seismic line. The first term, the near-field term, is proportional to $(1/r^2)$, therefore, its contribution is negligible as compared to the second term, which is proportional to $(1/r)$. The first term usually is dropped in practical implementation of the integral. It is the second term, the far-field term, that is used in migration. The time derivative of the measured wave field yields the 90-degree phase shift and adjustment of the amplitude spectrum by the ramp function of frequency (see Appendix A, Table A-1). For 2-D migration, the half-derivative of the wave field is used. This is equivalent to the 45-degree phase shift and adjustment of the amplitude spectrum by a function defined as the square root of frequency. Finally, the far-field term is proportional to the cosine of the angle of propagation (the directivity term or the obliquity factor) and is inversley proportional to $vr = v^2t$ (the spherical spreading term) in three dimensions. In two dimensions, the spherical spreading term is $(vr)^{1/2}$.

4.2.2 Finite-Difference Migration

To describe the physical basis of finite-difference migration, recall the harbor example of Figure 4-21. Instead of taking the section recorded along the beach, which contains the diffraction hyperbola, then collapsing it to get the migrated section in Figure 4-27, consider the following alternative procedure. Again, start with the wave field recorded along the beach (Figure 4-28a). Assume that the barrier is 1250 m from the beach. Now move the recording cable into the water, 250 m from the beach. Start recording at the instant the plane wave hits the barrier. The recorded section is shown in Figure 4-28b. Move the cable 500 m from the beach and record the section in Figure 4-28c, followed by a recording 750 m from the beach (Figure 4-28d). Finally, 1000 m from the beach, record the section shown in Figure 4-28e. Note that each recording yields a hyperbola in which the apex moves closer to zero time. The actual extent of the recording cable is denoted by the solid line on top of each frame. Had we recorded at the barrier (1250 m from the beach), the apex of the hyperbola would be positioned at $t = 0$.

In Kirchhoff migration, the diffraction hyperbola is collapsed by summing the amplitudes, then placing them at the apex. The alternative approach implied by the result of the experiment shown in Figure 4-28 is to use the hyperbola recorded a distance away from the beach to construct the hyperbola that would be recorded at another distance closer to the source of the diffraction hyperbola. The process is stopped when the hyperbola collapses to its apex. In the harbor experiment, this collapse

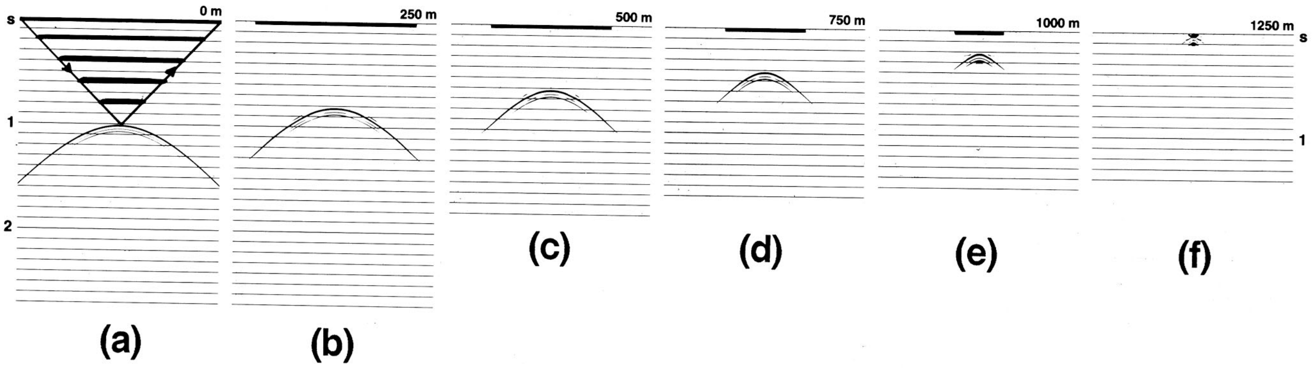


FIG. 4-29. Computer simulation of the experiment illustrated in Figure 4-28. Here, we downward continue the receivers at discrete depth intervals. The numbers on top indicate the distance of the receiver cable from the surface, $z = 0$.

occurs when the receiver cable coincides with the barrier, or, equivalently, when $t = 0$. This is called the *imaging principle*.

The harbor experiment described above can be simulated in the computer. Pretend that moving the receiver cable from the beach into the water closer to the barrier is like moving the receiver cable from the surface down into the earth closer to the reflectors. Think of the gap on the barrier as equivalent to a point (diffractor) on a reflecting interface causing the diffraction hyperbola (Figure 4-29a). Start with the wave field recorded at the surface and downward continue the receivers in depth at finite intervals. Downward continuation of the surface wave field can be considered equivalent to lowering the receivers into the earth. The computer-simulated wave fields at these different depths are shown in Figure 4-29. By applying the imaging principle at each depth, the entire wave field is imaged. The final output from this process is the migrated section. The last section [panel (f)] at 1250 m has only one arrival at $t = 0$. The recording cable is on the storm barrier and the arrival from the gap occurs at $t = 0$. As the cable moved into the ocean and recorded closer to the barrier, the recorded diffraction hyperbola arrived earlier, became shorter, and more compressed. It collapsed to a point when the receivers coincided with the storm barrier over which the source point forms a gap.

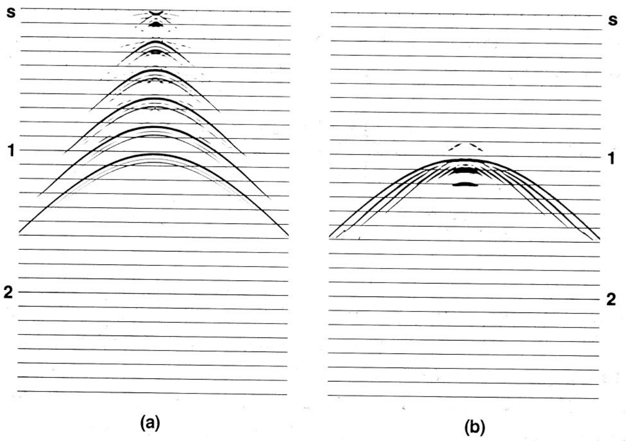


FIG. 4-30. (a) Superposition of the time sections in Figure 4-29; (b) removing the translational effect by retardation to place the energy at the apex of the hyperbola initially obtained along the beach line.

There is one important difference between the physical experiment in Figure 4-28 and the computer-simulated downward continuation experiment in Figure 4-29. While the receiver cable is the same at each step in Figure 4-28, the effective cable length gets shorter and shorter toward the source (the gap in the barrier) in Figure 4-29. This is because we started by recording the wave field at the surface (Figure 4-28a) with a finite cable length. The recorded information is confined to within the two raypaths depicted on the section in Figure 4-29a. As the cable moves closer to the source, the effective receiver cable containing the information is confined to smaller and smaller lengths. Although receivers are lowered vertically, energy moves down along raypaths it originally took on the way up. To relate these recordings at different depths (Figure 4-29), we superimpose them as shown in Figure 4-30a. Moreover, the recordings can be shifted so that the apexes of the hyperbolas coincide and are positioned at a time that is equivalent to the distance from the surface to the diffractor as shown in Figure 4-30b. This is called *time retardation*.

Reconsider the results from the computer simulation of the harbor experiment in Figure 4-29. Suppose we stopped recording at a depth of 1000 m before the barrier. The original hyperbola in Figure 4-29a was partially collapsed at this depth (Figure 4-29e). Therefore, downward continuing to a depth short of the true depth of the source causes undermigration. Diffractions (and dipping events) also are undermigrated if incorrectly low velocities are used for migration. By putting these two cases together, we know that insufficient downward continuation and too low velocities yield an undermigrated section. Assume that the recording continues and passes beyond barrier position z_3 (Figure 4-21). We infer that the focused energy on the section at this depth (Figure 4-29f) would propagate through the focal point and turn into hyperbolas that are the mirror images of those in Figures 4-29a through e. We have downward continued more than necessary. This yields overmigration, which also is caused by incorrectly high velocities. From these observations, note that downward continuing to a wrong depth is like downward continuing with the wrong velocity (Doherty and Claerbout, 1974).

Another important issue to consider is how often the extrapolated wave field should be computed. When going from one frame to another in Figure 4-29, what should the depth step size be? This is discussed in detail later in Section 4.3.2.

A migration method that uses the principle of downward continuation is called *finite-difference migration*. This migration technique is based on the differential solution to the scalar wave equation. A simple numerical example illustrates the finite-difference method of solving differential equations (Claerbout, 1985). Assume that you have \$100 today. Given an annual inflation rate of 10 percent, for the same buying power next year, you need \$110. A computer algorithm can determine the face value of the present \$100 in future years:

Operator	Data	Time
−1.1	100	│
1.0	x	│
	·	│
	·	│
	·	↓

Given the present value, 100, find future values in the data column.

The following equation solves for unknown x:

$$(1)(x) + (-1.1)(100) = 0. \tag{4.6}$$

We used a two-point operator and aligned it with the data column as indicated above. The present value, 100, is extrapolated in this way to future years:

By using equation (4.6), we have

−1.1	100	
1.0	x	$(1.0)x + (-1.1)100 = 0,\ x = 110.$
	100	
−1.1	110	
1.0	x	$(1.0)x + (-1.1)110 = 0,\ x = 121.$
	100	
	110	
−1.1	121	
1.0	x	$(1.0)x + (-1.1)121 = 0,\ x = 133.1$, and so on.

By moving the operator down in the time direction, we extrapolate the data column into the future. Equation (4.6) is generalized as

$$(1)P(t + 1) + (-1.1)P(t) = 0, \tag{4.7}$$

which is rewritten in the following form:

$$P(t + 1) - P(t) = (0.1)P(t), \tag{4.8}$$

where t is the time variable and P is the quantity being extrapolated. Instead of defining the time interval as one unit, we can define it as an arbitrary increment of time Δt. Assume that the inflation rate is a. Equation (4.8) then takes the more general form

$$P(t + \Delta t) - P(t) = aP(t). \tag{4.9}$$

Alternatively, we could use the average of the present and future values on the right side of this equation:

$$P(t + \Delta t) - P(t) = \frac{a}{2}\,[P(t + \Delta t) + P(t)]. \tag{4.10}$$

Equations (4.9) and (4.10) now can be put into the form of equation (4.6):

$$P(t + \Delta t) + (-1 - a)P(t) = 0, \tag{4.11}$$

and

$$(1 - a/2)P(t + \Delta t) + (-1 - a/2)P(t) = 0. \tag{4.12}$$

By using either equation (4.11) or equation (4.12), we compute the future values of $P(t)$ from a given initial value:

$-1-a$	$P(t)$	or	$-1-a/2$	$P(t)$
1	$P(t + \Delta t)$		$1-a/2$	$P(t + \Delta t)$
Explicit Operator			Implicit Operator	

The operator in which the coefficient of the future value $P(t + \Delta t)$ is unity is called the *explicit operator*. Stability (the problem of wave amplitudes growing from one extrapolation step to another) of the finite-difference solution is an issue with this type of operator. Implicit schemes produce stable results because of averaging on the right side of equation (4.10), the so-called Crank-Nicolson scheme. For the differential equations used in finite-difference migration algorithms (e.g., the parabolic equation described in Appendix C.2), scalar a becomes a matrix coefficient. Implicit schemes require inversion of this matrix. However, no inversion is needed with explicit schemes, since future values can be written explicitly in terms of only past values.

Equation (4.11) is rewritten by redefining scalar a to $a \cdot \Delta t$:

$$\frac{P(t + \Delta t) - P(t)}{\Delta t} = aP(t). \tag{4.13}$$

The left side of equation (4.13) is the discrete representation of the continuous derivative of P with respect to time *dP*/*dt*. Therefore, equation (4.13) is the *finite-difference equation* that corresponds to the following *differential equation:*

$$\frac{dP}{dt} = aP(t). \tag{4.14}$$

We have derived the differential equation that describes the inflation of money [equation (4.14)]. Now consider the analysis in reverse order. We start with the differential equation, equation (4.14), and write the corresponding difference equation, equation (4.13), which is the equation that is solved in the computer. This equation is written in either the explicit [equation (4.11)] or implicit [equation (4.12)] form to extrapolate the present value of P to the future.

This example illustrates how finite-difference schemes can solve differential equations in the computer. The scalar wave equation can be treated in a similar, but more complicated manner. Complications arise because it is a partial differential equation that contains the second derivatives of the wave field with respect to depth, time, and spatial axes. Setting up the computer algorithm is more involved and is not discussed here. Claerbout (1976, 1985) provides details of various aspects of the finite-difference migration methods.

Appendix C.2 describes a derivation of the parabolic (15-degree) approximation to the one-way scalar wave equation [equation (C.39)]. This approximation is the basis for the most commonly used finite-difference algorithm. Equation (C.39) is rewritten below:

$$\frac{\partial^2 Q}{\partial \tau \partial t} = \frac{v^2}{8}\frac{\partial^2 Q}{\partial y^2}, \tag{4.15}$$

where Q is the retarded wave field, t is the input time, τ is the output time, and y is the midpoint coordinate.

Equation (4.15) is the basis for the 15-degree time migration. It is derived from the dispersion relation (see Appendix C.2), assuming that velocity varies vertically. However, in practice, the velocity function in equation (4.15) can be varied laterally, provided it is smooth.

Equation (4.15) only accounts for collapsing diffraction energy to the apex of the traveltime curve, thus it is called the diffraction term. When lateral velocity variations are significant, the diffraction curve is somewhat like a skewed hyperbola with its apex shifted laterally away from the diffraction source. This lateral shift is accounted for by the thin-lens term in equation (C.35). The problem of lateral velocity variations is discussed in great detail in Section 5.1. If the lateral velocity variations are significant, then the thin-lens term is not negligible. The differential equation that corresponds to the thin-lens part of equation (C.35) is obtained by inverse Fourier transformation:

$$\frac{\partial Q}{\partial z} = 2\left[\frac{1}{\bar{v}(z)} - \frac{1}{v(y,z)}\right]\frac{\partial Q}{\partial t}. \tag{4.16}$$

Migration algorithms that implement both the diffraction and thin-lens terms in equation (C.35) generally are two-step schemes that alternately solve these two terms. In other words, to propagate one depth step, first apply the diffraction term on wave field Q. The thin-lens term then is applied to the output from the diffraction calculation. A migration method that includes the effects of the thin-lens term is called *depth migration*, since the output section is in depth. Depth migration (Section 5.2) is warranted if, and only if, there are strong lateral variations in velocity; in this case, the coefficient of the thin-lens term cannot be negligible.

As discussed in detail in Section 4.3.2, in practice the 15-degree finite-difference migration can handle dips up to 35 degrees with sufficient accuracy. However, there are steep-dip approximations to the one-way scalar wave equation. The 45-degree equation and its extension to steeper dips [equation (C.50)] are discussed in Appendix C.3, while the practical aspects of steep-dip migration are covered in Section 4.3.4. Steep-dip finite-difference algorithms may be more conveniently implemented in the frequency-space domain than in the time-space domain.

Boundary and initial conditions are needed to solve the differential equations. The initial condition for migration is the recorded wave field at the surface $z = 0$. Also, in migration we assume that the wave field is zero after a maximum observation time, typically the end time of the recorded trace. Then there are the side boundaries, beyond which assumptions must be made about the form of the wave field.

In the (x,z,t) coordinates, the seismic section is represented by the (x,t) plane, while the migrated section (earth) is represented by the (x,z) plane. Finite-difference migration, as discussed here, extrapolates the (x,t) plane in finite increments of z and outputs the wave field at $t = 0$ at each step (Figure 4-31). Another migration method, known as *reverse time migration* (Baysal et al., 1983), extrapolates an initially zero (x,z) plane backward in time, bringing in the seismic data $P(x,z = 0,t)$ as a boundary condition (at $z = 0$) at each time step. At time $t = 0$, this (x,z) plane contains the migration result $P(x,z,t = 0)$ (Figure 4-32).

Note that both the Kirchhoff and finite-difference approaches are based on the same equation, namely, the scalar wave equation. They only differ in terms of solving this equation. Kirchhoff migration is based on the integral solution, while finite-difference migration is based on the differential solution. Regardless of the method, all migration techniques incorporate the imaging principle. One important advantage of finite-difference techniques over the other migration methods is their ability to better handle lateral velocity variations.

4.2.3 Frequency-Wavenumber Migration

Stolt (1978) introduced the Fourier transform methods in migration. Since then, additional papers have appeared on the subject, some on theory, and some on the practical

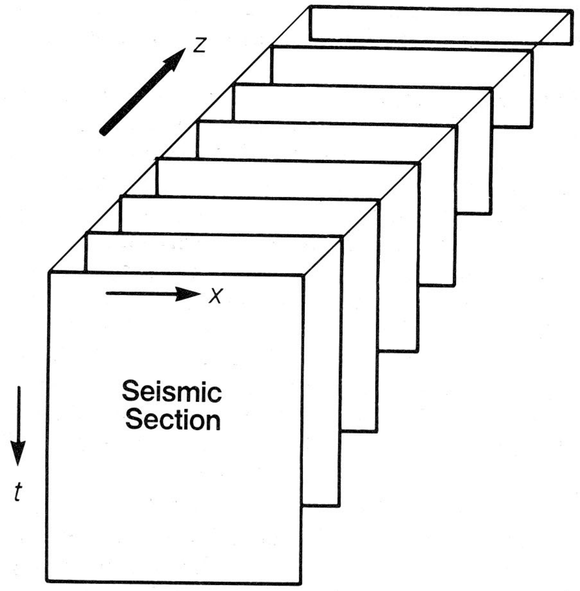


FIG. 4-31. The seismic section represented by the (x, t) plane at the surface, $z = 0$, can be downward continued to obtain the time sections at discrete depth levels. The direction of extrapolation is indicated by the thick arrow. The migrated section is represented by the (x,z) plane at $t = 0$ (imaging principle).

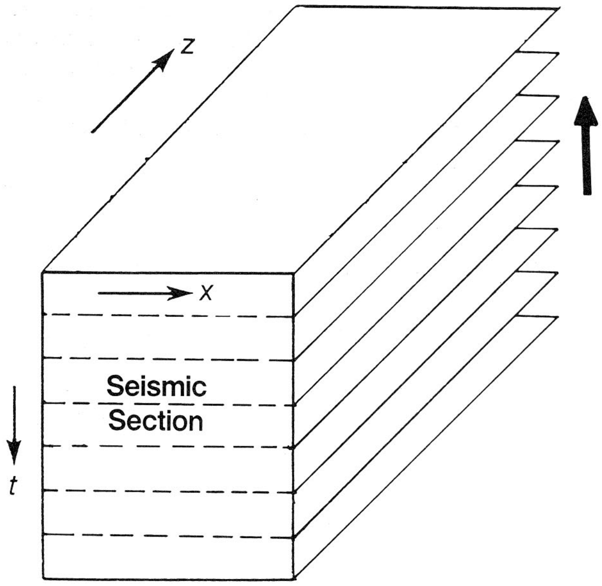


FIG. 4-32. Reverse time migration: Start with an all-zero (x,z) plane at the bottom of the data cube and extrapolate backward in time toward $t = 0$ to compute snapshots of the (x,z) plane at different times. These snapshots of the subsurface are indicated by the horizontal planes; the direction of extrapolation is indicated by the thick arrow. At each time level, include the boundary value (x slice at $z = 0$, indicated by the dotted lines) into the (x,z) plane from the seismic section. The migrated section is the (x,z) plane at $t = 0$ (the top horizontal plane).

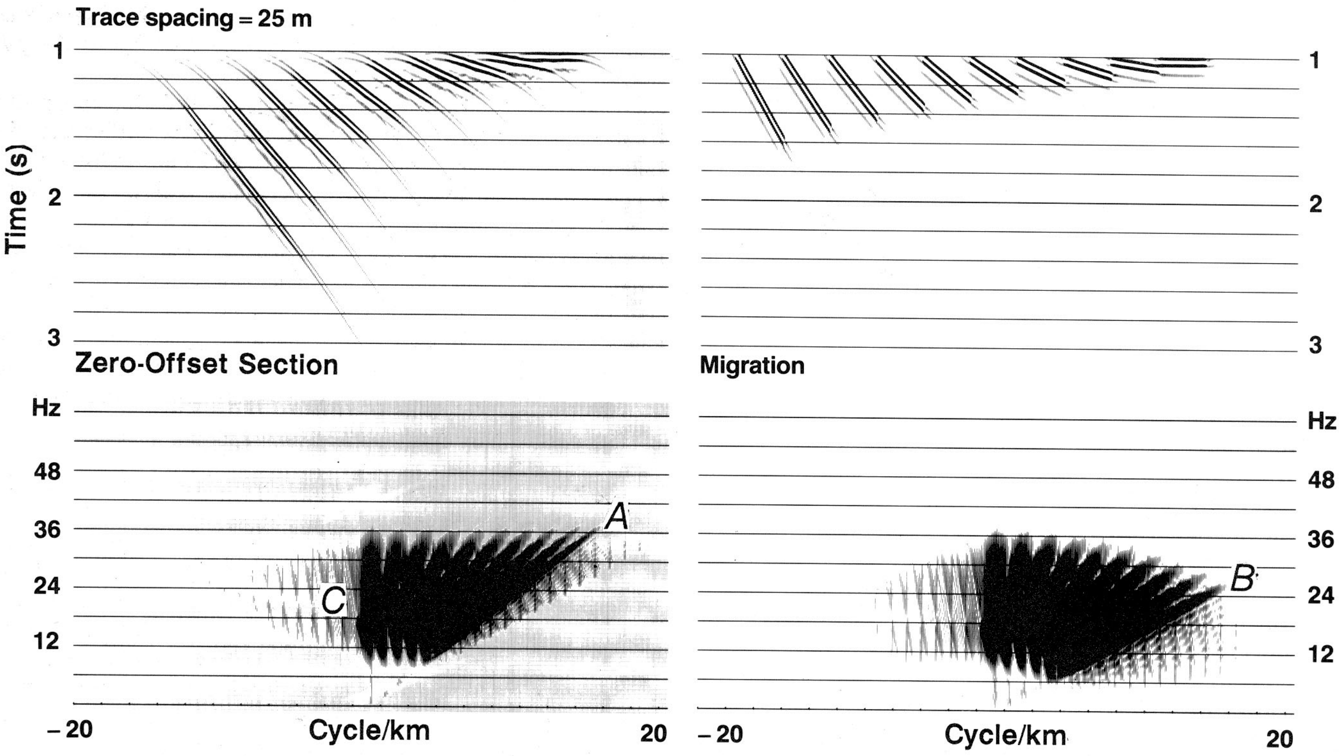


FIG. 4-33. Migration of dipping reflectors in the (t,x) (top) and (f,k) (bottom) domains. Energy A, which is associated with the steepest dip, maps onto B after migration. Exercise 4.14 refers to C. Medium velocity is constant, 3500 m/s.

aspects of the frequency-wavenumber migration. Gazdag (1978) published his work on the phase-shift method, which led to a further understanding of wave field extrapolation in the transform domain. Frequency-wavenumber *(f-k)* migration is not as easily explained as the Kirchhoff or finite-difference migration from a physical point of view. Chun and Jacewitz (1981) provide good insight into the principles of *f-k* migration.

In Section 1.6, we saw that dipping events in (t,x) space map into the *f-k* domain as radial lines. The steeper the dip, the closer the radial line is to the wavenumber axis. Figure 4-33 is an *f-k* representation of dipping events before and after migration. Note that migration rotates the radial lines in the 2-D amplitude spectrum outward and away from the zero wavenumber axis. This rotation is akin to the opening of a fan. The steepest event represented by radial line *A* maps onto radial line *B* after migration.

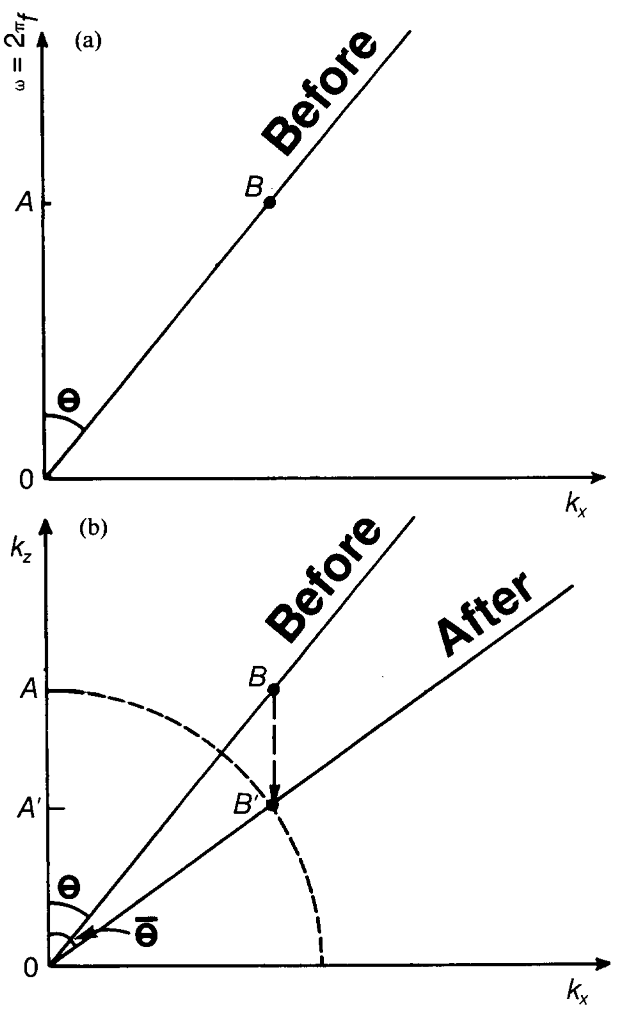


FIG. 4-34. Migration in the (f,k) domain. (a) A dipping reflector is represented by a radial line *OB* in the (f,k) plane. (b) After migration, the radial line *OB* maps onto another radial line *OB'*, while *B* maps onto *B'*. The horizontal wavenumber is invariant under migration. For comparison, the *f-k* response of the dipping event before migration (a) has been superimposed on that after migration (b). (Adapted from Chun and Jacewitz, 1981.)

Migration of a dipping event in the *f-k* domain is shown in Figure 4-34. The vertical axis is the frequency axis ω for the premigrated event, and it is the vertical wavenumber k_z associated with the depth axis for the postmigrated event. Note that this figure is the *f-k* equivalent of Figure 4-14. In both figures, we assume velocity equal to 1. *F-K* migration maps lines of constant frequency *AB* in the (ω,k_x) plane to circles *AB'* in the (k_z,k_x) plane. Therefore, migration maps point *B* vertically onto point *B'*. Note that in this process, the horizontal wavenumber k_x does not change as a result of mapping. When this mapping is completed, the dipping event *OB* is mapped along *OB'* after migration; thus, the dip angle after migration $(\bar{\theta})$ is greater than before migration (θ). For comparison, these two radial lines are shown on the same plane (k_z,k_x).

We now examine the diffraction hyperbola and its collapse to the apex after migration in the *f-k* domain. Figure 4-35 depicts a diffraction hyperbola in the (t,x) and (f,k) domains. We imagine that the hyperbola is made up of a series of dipping segments, such as *A,B,C,D,* and *E*. The zero-dip segment (*A*) is at the apex, while the steepest dip segment (*E*) is along the asymptotes. In (f,k) space, the zero-dip segment A maps along the frequency axis, while the dipping segments *B, C,* and *D* map along radial lines, increasingly away from the frequency axis. Finally, the asymptotic tail *E* maps along the radial line that represents the boundary between the propagation and the evanescent region. (The evanescent region corresponds to the energy that is located at or greater than 90 degrees from the vertical.) The opposite side of the hyperbola maps onto the second quadrant (negative k_x) in the *f-k* plane.

The analysis illustrated in Figure 4-35 is based on the representation of the hyperbola as a series of discrete dip segments. In the continuous case, a diffraction hyperbola is represented by a fan in the frequency domain. Migration opens this fan up as shown in Figure 4-36.

A curious fact emerges from the *f-k* plot of the migrated section in Figure 4-36. We expect migration to collapse the diffraction hyperbola to a point at the apex. The spectrum of this migrated section really should be more like that in Figure 4-37, that is, a rectangle. Why the difference between this spectrum and the spectrum after migration in Figure 4-36?

If you start with a point and model it, you get the diffraction hyperbola in Figure 4-37. However, in reality we deal with a diffraction hyperbola as shown in Figure 4-36. The hyperbolas do not look different in the (t,x) domain, but note the difference in their *f-k* spectra. The *f-k* spectrum of the real-life diffraction, which is always subjected to band-pass filtering (Figure 4-36), is missing the energy above the roughly 42-Hz line that is present in the *f-k* spectrum of the modeled diffraction curve (Figure 4-37). These missing high frequencies cause the difference between the spectra after migration.

Mathematical development of the frequency-wavenumber migration techniques is left to Appendix C.4. Figure 4-38 is a flowchart of the phase-shift method; its

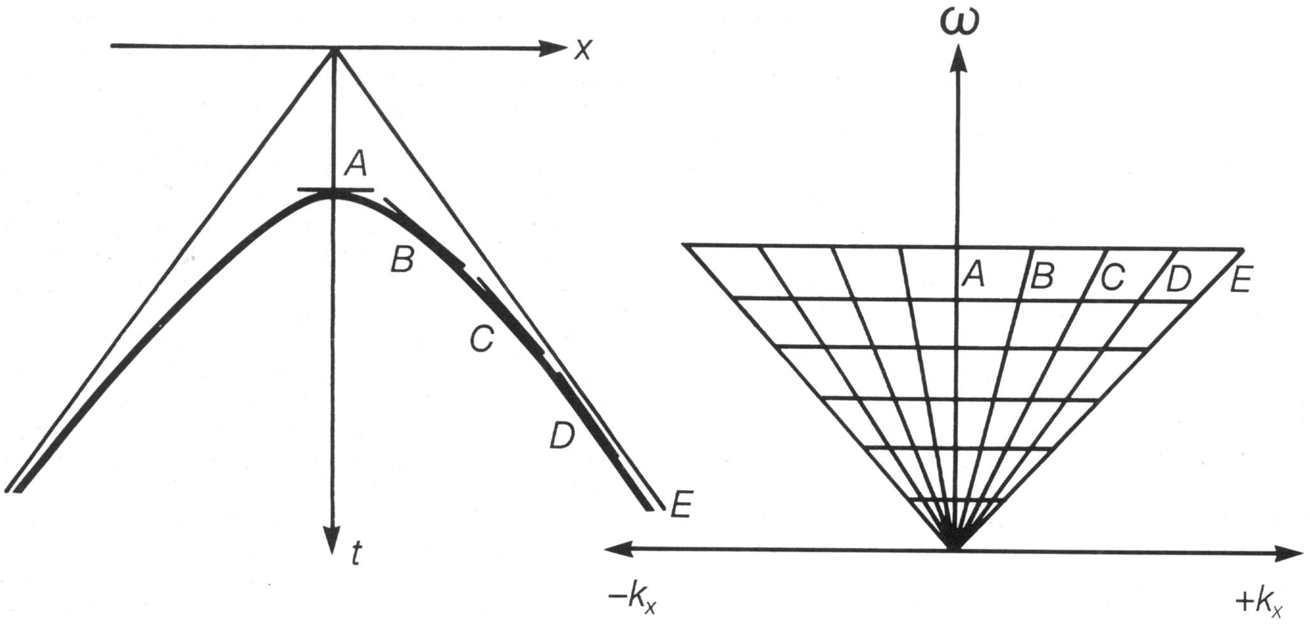


FIG. 4-35. A hyperbola maps onto a triangular area in the (*f*,*k*) plane. (See text for details.)

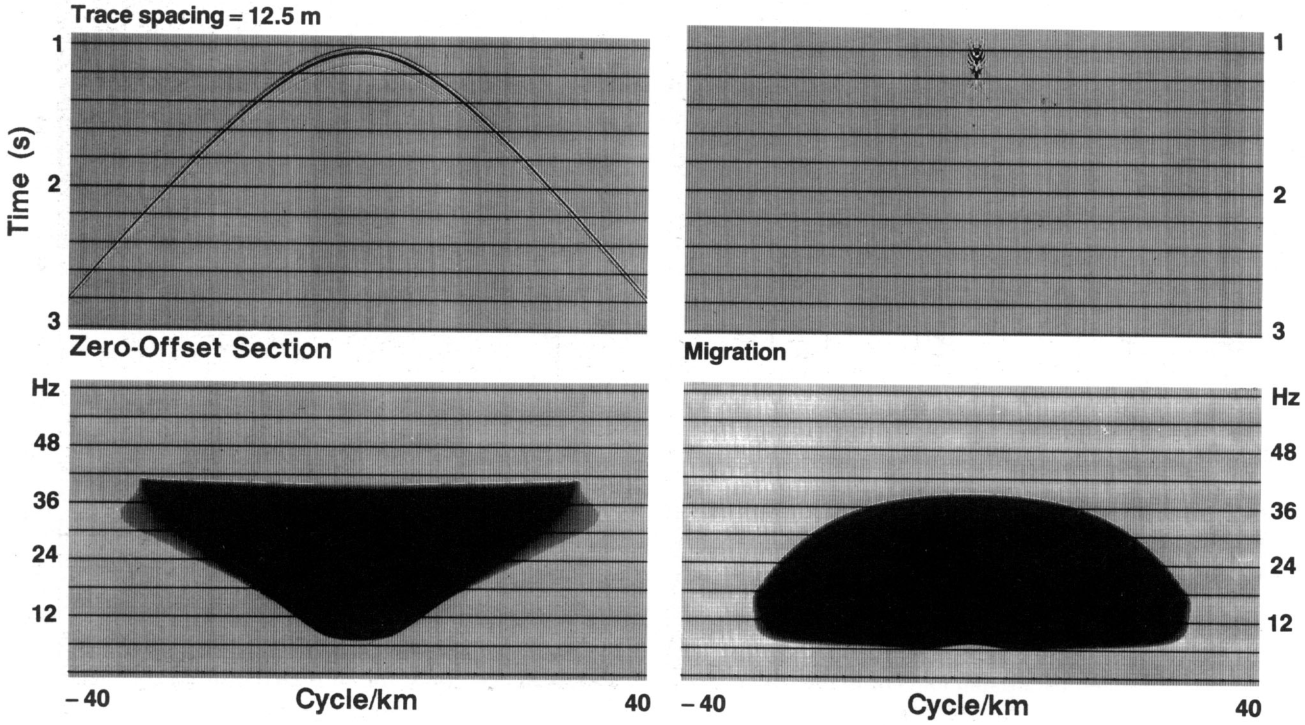


FIG. 4-36. A diffraction hyperbola and its migration in the (*t*,*x*) (top) and (*f*,*k*) (bottom) domains.

foundation is embedded in equation (C.7). Note that downward continuation involves a pure phase-shifting operation $\exp(-ik_z z)$. At each z step, a new extrapolation operator with the velocity defined for that z value is computed. As for any other migration, we need to invoke the imaging principle ($t = 0$) at each extrapolation step to obtain the migrated section. The imaging condition $t = 0$ is met by summing over all frequency components of the extrapolated wave field at each depth step. This is easily shown from the integral representing the inverse Fourier transform of the extrapolated wave field [refer to equation (C.52)]. The procedure of downward continuation and imaging is repeated until the entire wave field is migrated.

The phase-shift method (Gazdag, 1978) can only handle vertically varying velocities. Gazdag and Squazzero (1984) extended the method to handle lateral velocity variations. To achieve this, first the input wave field is extrapolated by the phase-shift method using a multiple number of laterally constant velocity functions and a series of reference wave fields are created. The imaged wave field then is computed by interpolation from the reference wave fields. This migration method is known as *phase-shift-plus-interpolation*.

If the medium velocity is constant, then migration can be expressed as a direct mapping (Stolt, 1978) from temporal frequency ω to vertical wavenumber, k_z (see Figure 4-34). Figure 4-39 is a flowchart of the Stolt algorithm; the mathematical details are left to Appendix C.4.

Stolt's algorithm for constant velocity is efficient because of the direct mapping given by equation (C.53). However, it involves interpolation in the *f-k* space to compute the corresponding ω for given k_x and k_z values. The question is whether this method has any practical value, since we never have a constant-velocity earth.

Stolt extended his method to handle velocity variations (Appendix C.4). For the variable-velocity case, Stolt's extension consists of (1) modifying the input wave field to make it appear as if it were the response of a constant-velocity earth, (2) applying the constant-velocity algorithm outlined in Figure 4-39, and (3) reversing the original modification of the input wave field. This modification is essentially a type of stretching of the time axis (see Appendix C.4) to make the reflection times approximately equivalent to those recorded for a constant-velocity earth. The nature of stretching is described by the so-called *stretch factor* W. The constant-velocity case is equivalent to $W = 1$. A practical use of the constant-

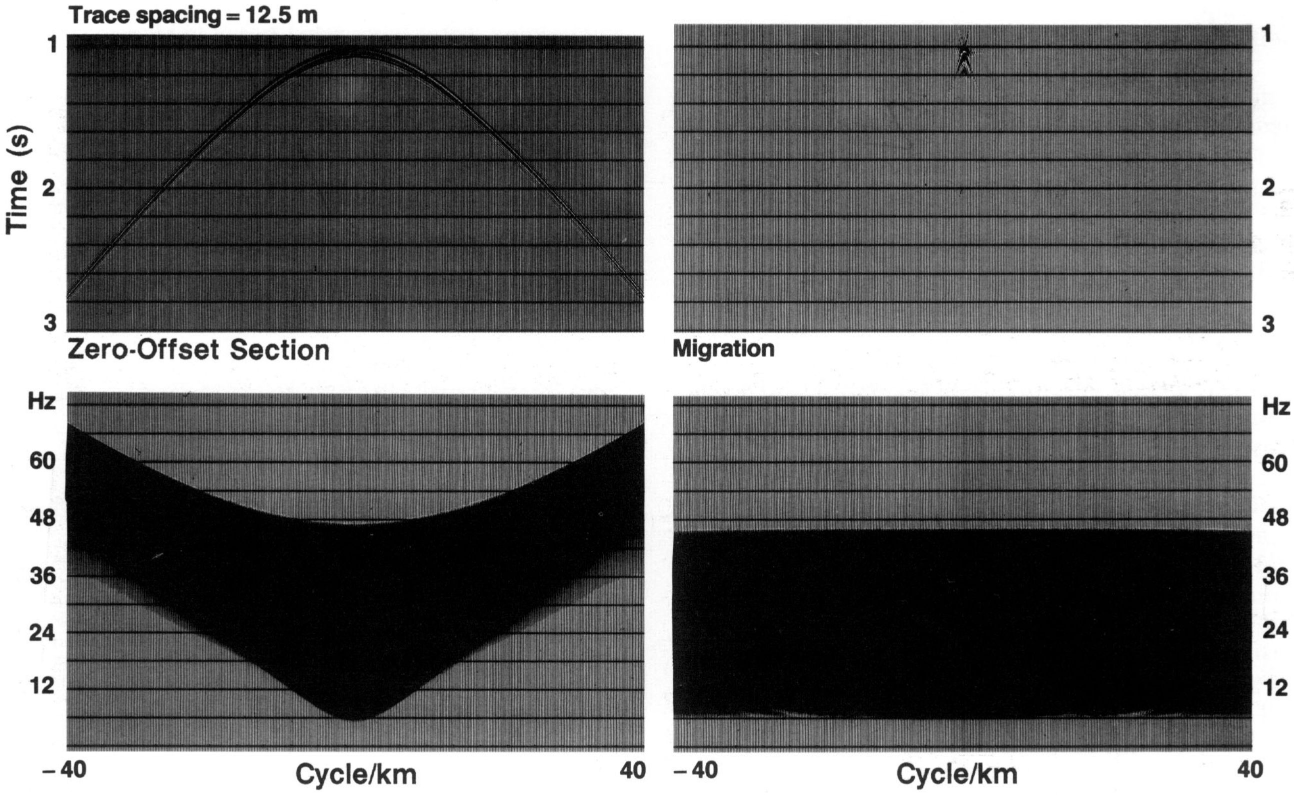


FIG. 4-37. Migration of a diffraction hyperbola in the (t,x) (top) and (f,k) (bottom) domains. This hyperbola is not dealt with in reality. The band-pass filtered hyperbola shown in Figure 4-36 is dealt with instead. Note the difference in the amplitude spectra of the two hyperbolas.

velocity algorithm is made in residual migration (Section 4.3.3 and Appendix C.5).

Note that the phase-shift and Stolt migration outputs normally are displayed in two-way vertical zero-offset time $\tau = 2z/v$, as are the outputs from the finite-difference and Kirchhoff migrations. In practice, mapping in the *f-k* domain really is from (ω, k_x) to (ω_τ, k_x) rather than to (k_z, k_x), where ω_τ is the Fourier dual of τ, and is simply k_z scaled by $v/2$. Specifically, equations (C.52) and (C.55) can be expressed again for the phase-shift and Stolt methods, respectively, in terms of τ, rather than z, by using equation (C.38a) in place of equation (C.53).

One important concept must be pointed out from the above analysis. From equation (C.38a), we see that for a constant k_x, $\omega > \omega_\tau$; thus, migration shifts the bandwidth to lower frequencies. This is analogous to the conclusion derived in relation to the NMO correction, since the latter also causes data stretching to lower frequencies (Section 3.2.2).

4.3 MIGRATION IN PRACTICE

In this section, the parameters that affect performance of Kirchhoff summation, finite-difference, and the *f-k* migration methods are discussed. In Kirchhoff migration, the important parameters are the aperture width used in summation and the maximum dip to migrate. In finite-difference and phase-shift migrations, the depth step size needs to be selected properly. The stretch factor is important in Stolt migration. The responses of these methods to velocity errors also are examined. All practical aspects are discussed using a synthetic model that consists of a number of dipping events and another model that consists of a diffraction hyperbola. Real data examples also are used to evaluate the choice of optimum parameters.

Throughout this section, migration results of different algorithms using various parameters are compared with a

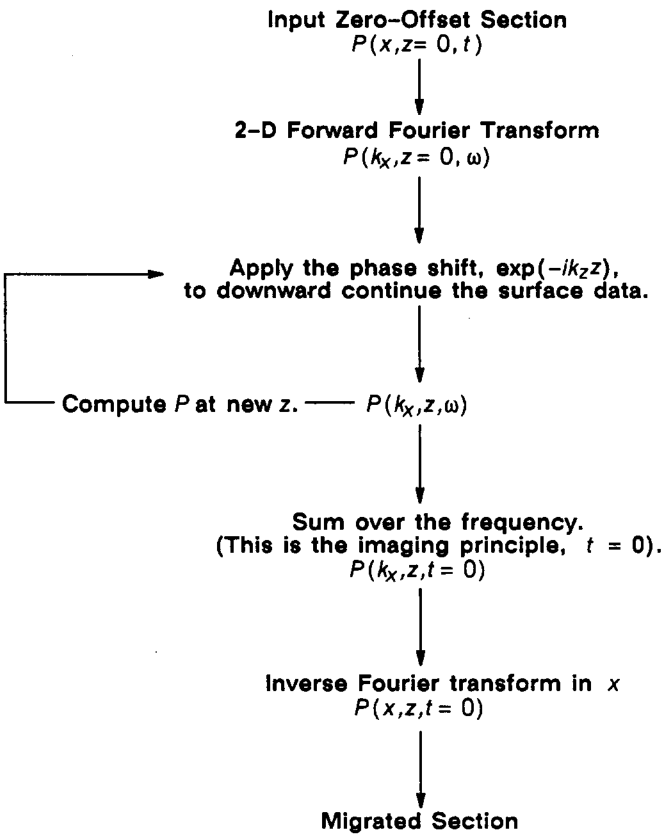


FIG. 4-38. Flowchart for Gazdag's phase-shift method of migration.

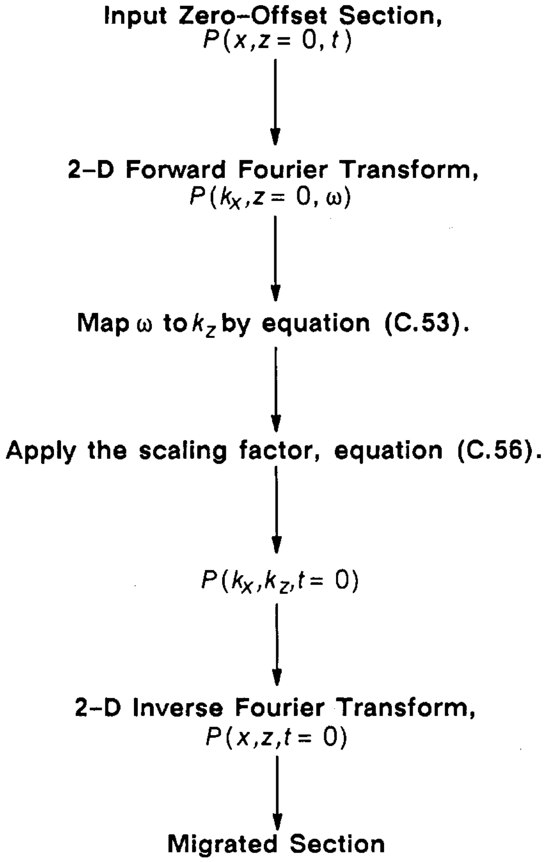


FIG. 4-39. Flowchart for constant-velocity Stolt migration.

desired migration. In all cases, this desired migration was obtained using the phase-shift method with appropriate parameters and velocities. This does not imply that the phase-shift method always provides a desirable output; it only means that the data examples in this section were chosen so that the phase-shift algorithm is appropriate. The choice of the phase-shift method was a compromise; it handles dips of up to 90 degrees and velocities that can only vary vertically.

4.3.1 Kirchhoff Migration in Practice

Before a migration algorithm is used on field data, its impulse response must be tested. A band-limited impulse response is generated by using an input that contains an isolated wavelet on one trace only. The ideal migration process should produce a semicircle. Kirchhoff migration produces the section shown in Figure 4-40d. The impulse response indicates that Kirchhoff migration can accurately handle dips up to 90 degrees. The dip on a migration impulse response is measured as the angle θ between the vertical and a specified radial direction. Note that migration can be limited to smaller dips (Figure 4-40).

From the previous section, we know that Kirchhoff migration involves a summation of amplitudes along diffraction hyperbolas. Given the rms velocity at a particular time sample of a particular input trace, a diffraction hyperbolic path is overlaid on the input section with its apex at that time sample. In theory, a diffraction hyperbola extends to infinite time and distance. In practice, we have to deal with a truncated hyperbolic summation path. The spatial extent that the actual summation path spans, called the *migration aperture*, is measured in terms of the number of traces the hyperbolic path spans.

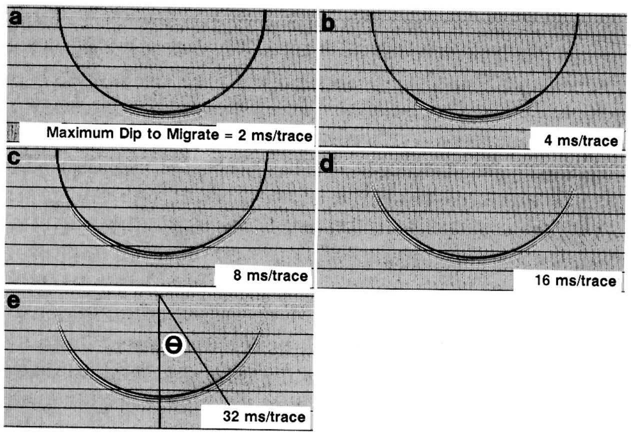


FIG. 4-40. Migration can be confined to a range of dips present on a seismic section. The impulse response for the dip-limited migration operator is a truncated semicircle. Dip angle Θ is measured from the vertical axis.

The curvature of the diffraction hyperbola is governed by the velocity function. Figure 4-41a shows a number of low-velocity diffraction hyperbolas, while Figure 4-41b shows a number of high-velocity hyperbolas. A low-velocity hyperbola has a narrower aperture when compared to a high-velocity hyperbola. This agrees with our intuition and previous discussions; that is, high velocity means more migration. In practice, we deal with a velocity function that at least varies with depth. The diffraction hyperbolas can have different curvatures depending on the velocity value at a given time sample (Figure 4-41c). Because of the vertical variation in velocity, aperture width generally is time variant. For the usual case in which velocity increases with depth, migration of the shallow part of the section requires a narrow aperture, while migration of the deep portion requires a wider aperture (Figure 4-41c). This implies that, given the same dip, deep events migrate farther than shallow events.

Aperture Width

An important parameter in practical implementation of Kirchhoff migration is aperture width. Figure 4-42 shows a zero-offset diffraction hyperbola (8 ms/trace dip along the asymptotes) and migrations using four different aperture widths. The smaller the aperture, the less capable the migration is in collapsing the diffraction hyperbola. In this case, use of an aperture width that is equal to the width of the input section (half aperture, 256 traces) yields the best result.

Figure 4-43 shows a synthetic zero-offset section that consists of a number of dipping events ranging from 0 to 45 degrees in increments of 5 degrees. Aperture width is closely related to the horizontal displacement that takes place in migration [equation (4.1)]. The number of traces an event migrates is NX = horizontal displacement/trace interval. Therefore, the aperture width that is required is $2NX + 1$. Figure 4-43 also shows Kirchhoff migrations of the dipping events using four different aperture widths. Small-aperture migration (half aperture, 35 traces) eliminates steeply dipping events on the output section. Increasing the aperture width allows proper migration of the steeply dipping events. From this we see that using too small an aperture width causes a dip filtering action during migration because a small aperture excludes the steeper flanks of the diffraction hyperbola from the summation. For any given time, the optimal value for the aperture width is defined by twice the maximum horizontal displacement in migration for the steepest dip of interest in the input section. In this case, the horizontal displacement associated with the 45-degree dipping event is computed by substituting the values for $v = 3500$ m/s, $\Delta t/\Delta x = 12$ ms/trace, and $t = 2$ s in equation (4.1). The value for the horizontal displacement is 118 traces, giving an aperture width of 237 traces. We typically consider somewhat larger values to allow for velocity errors.

A good way to determine aperture width is to generate diffraction hyperbolas as shown in Figure 4-41c using the

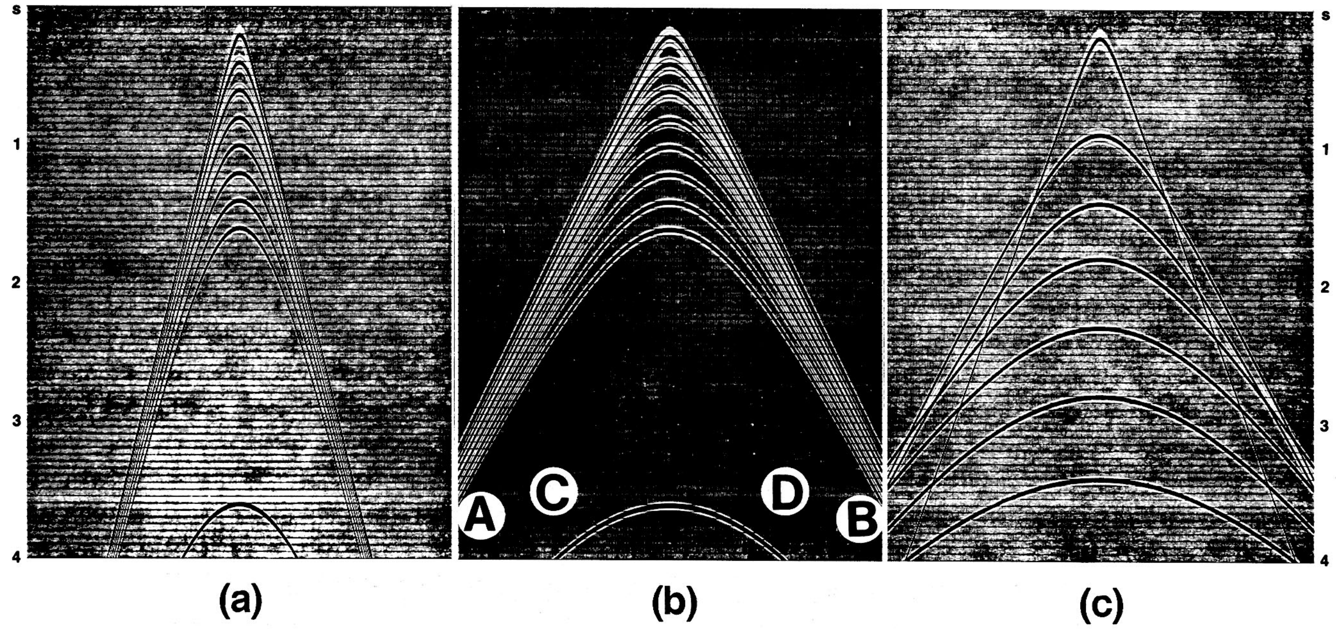


FIG. 4-41. Summation paths for Kirchhoff migration in a medium with (a) low velocity (2000 m/s), (b) high velocity (4000 m/s), and (c) vertically varying velocities. Migration aperture is small for low velocities and large for high velocities. (See text for details.)

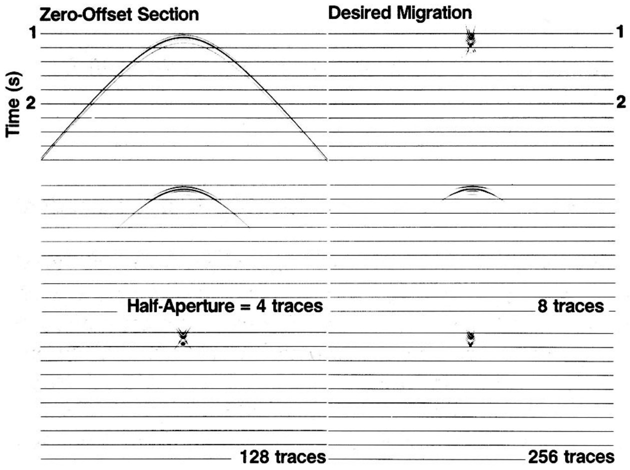


FIG. 4-42. Tests for aperture width in Kirchhoff migration. Insufficient aperture width causes incomplete collapse of diffraction hyperbola.

regionally averaged *v(z)*-type velocity. Clearly, the larger the aperture width, the more traces are used in the summation. For the dipping events in Figure 4-43, the optimal value of the half-aperture width is 150 traces; increasing the width to 300 traces resulted in no further improvement.

A test of aperture width on the CMP stacked data example is shown in Figure 4-44. The small-aperture migration (half-aperture 40 traces) causes smearing in the deeper part of the section. This smearing effect destroys the dipping events and produces spurious horizontally dominant events. Smearing gradually disappears with increasing aperture. It is apparent that the optimum half aperture is 160 traces.

Figure 4-45 is the deeper portion of a stacked section with migrations using different aperture sizes. The smearing effect is much more noticeable at small aperture. The main difference between the stacked sections in Figures 4-44 and 4-45 is that the latter, being deeper in time, contains a large amount of noise. This smearing phenomenon was not noticed in the synthetic model in Figure 4-43 because it did not contain any noise. We now see that choice of aperture width is more critical than we originally thought. In particular, a small aperture changes the noise characteristics of the section.

Why do we see horizontally dominant smearing with small-aperture migration? To answer this, we must do a simple experiment with a section that contains only random noise (Figure 4-46a). The velocity function used in migration increases in time. We see two interesting phenomena on the migrated sections using three different apertures. First, in all three cases there is more smearing of noise in the deeper part of the data, where the velocities generally are higher than in the shallower part. Second, there is relatively more smearing in the small-aperture migration compared with others at a given time in the section. Moreover, this smearing is characterized by horizontally dominant spurious events, especially in

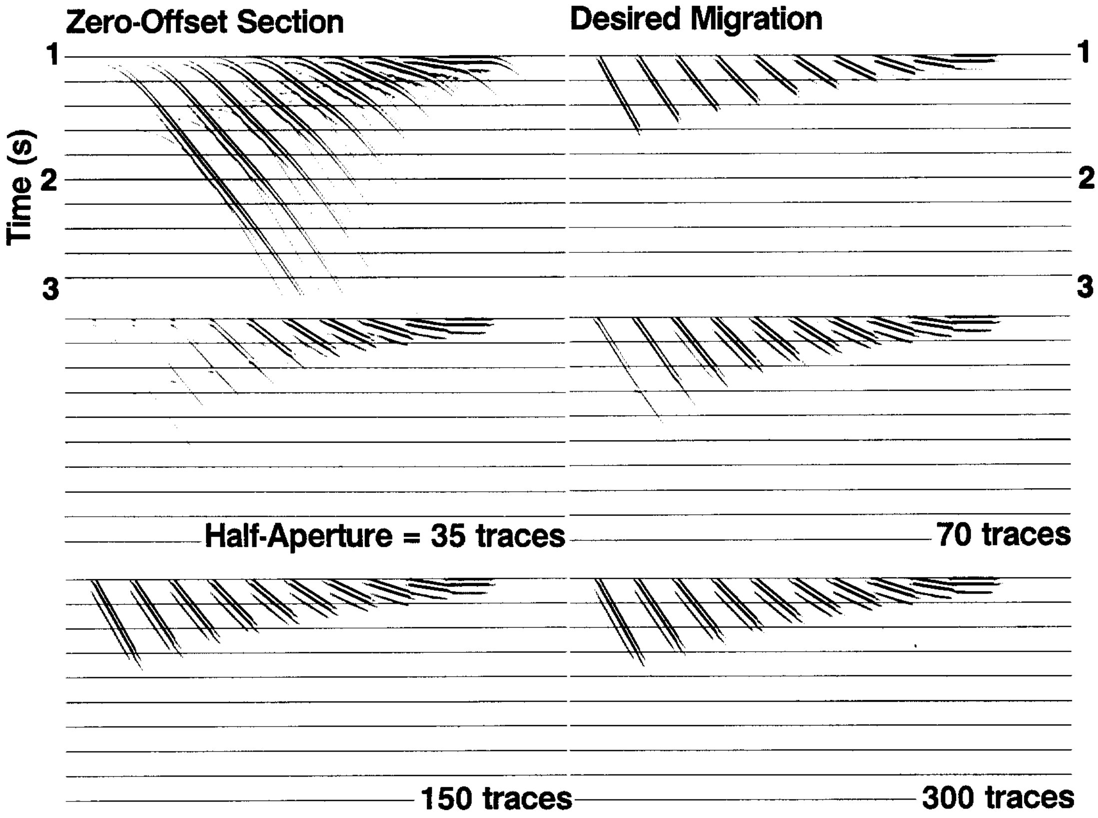


FIG. 4-43. Tests for aperture width in Kirchhoff migration. Insufficient aperture width causes removal of steeply dipping reflectors.

the deeper part of the section. Note that even with a large aperture, some smearing still is present in the deepest part of the section in Figure 4-46d. As indicated in Figure 4-41b, because summation stops at the bottom of the section, the effective aperture at late times CD is much smaller than that used in other parts of the section AB. Remember that summation using very small aperture includes only the apex portion of the diffraction hyperbola, where dips are nearly flat. Therefore, the small aperture with a dip filtering action passes flat or nearly flat events; i.e., those horizontal wavenumber components that are zero or nearly zero.

In conclusion, the following assessments are made concerning the choice of aperture width. Excessively small aperture width causes suppression of steeply dipping events and rapidly varying amplitude changes. Excessively small aperture width organizes random noise, especially in the deeper part of the section, as horizontally dominant spurious events. Excessively large aperture means more computer time. More importantly, large apertures can degrade the migration quality in poor signal-to-noise ratio conditions. Suppose deeper data in the section are noisy. Use of large aperture will cause this noise to creep into the good shallow data. Aperture width always is a compromise with noise. Sometimes it is better to use a smaller aperture than theoretically is required to avoid the adverse effect of noise on the migrated signal. Noise considerations may even require a time-dependent aperture width.

It is recommended that the aperture width be kept constant in migrating all lines from a particular survey so that an overall uniformity in amplitude characteristics on the migrated sections is maintained. In practice, a regional velocity function and the steepest dip in the survey area are used to compute the optimal aperture width that can be used over the entire set of data from the area [equation (4.1)].

Maximum Dip to Migrate

During migration, we can specify the maximum dip we want migrated in the section. This may be useful when we need to suppress the steeply dipping coherent noise. Figure 4-47 shows migrations of the dipping events with four different maximum allowable dips. For a 4 ms/trace dip limit, events with dips greater than this value are suppressed. Similarly, for an 8 ms/trace dip value, events with dips greater than this value are suppressed. When the dip value is 12 ms/trace, no suppression occurs, since all events on the input section have dips less than this value. Limiting the dip parameter is a way to reduce computational cost, since dip parameter is related to aperture width [equation (4.1)], which determines the

(Text continued on page 276)

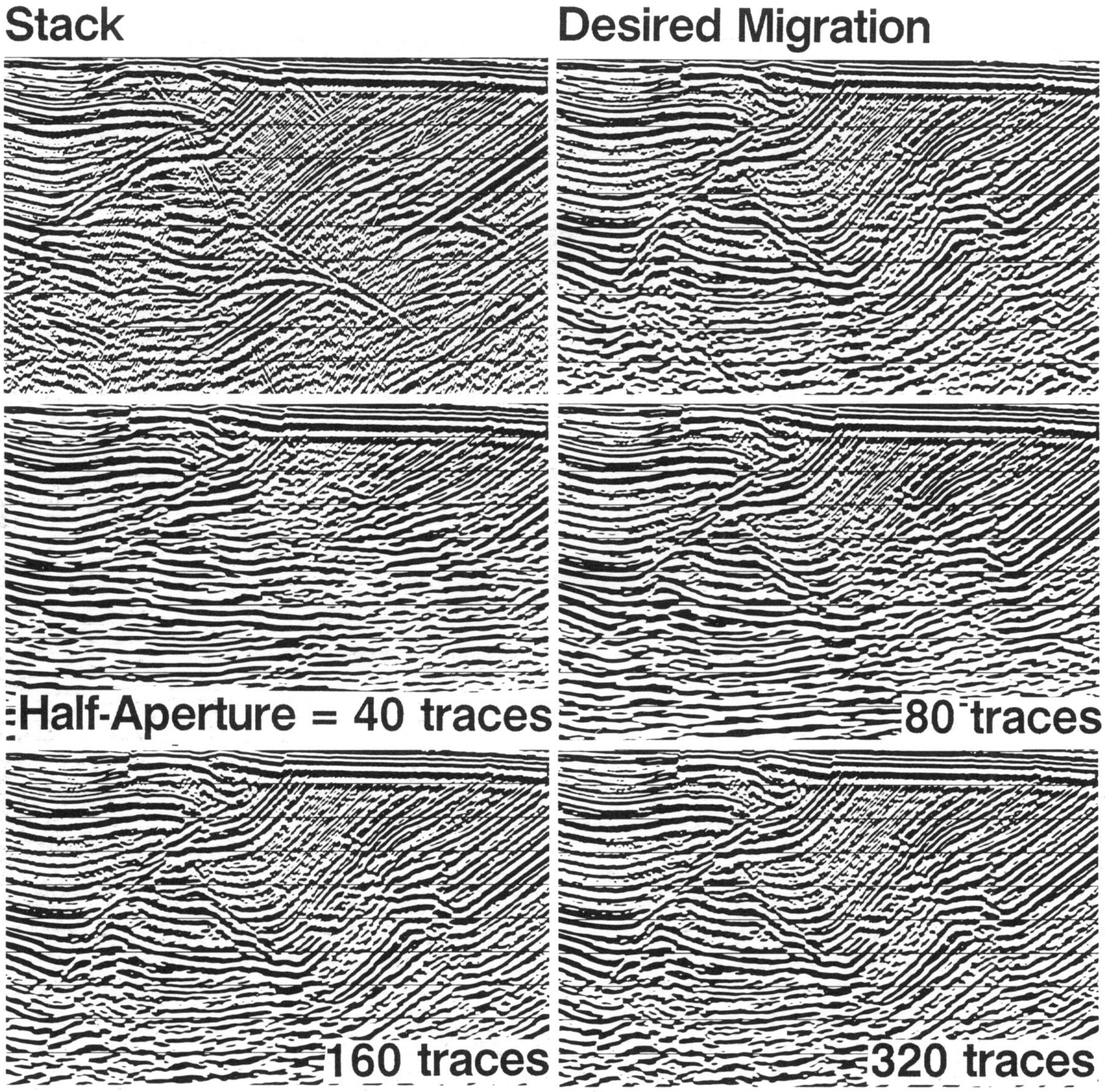


FIG. 4-44. Tests for aperture width in Kirchhoff migration. Insufficient aperture width causes removal of steeply dipping events.

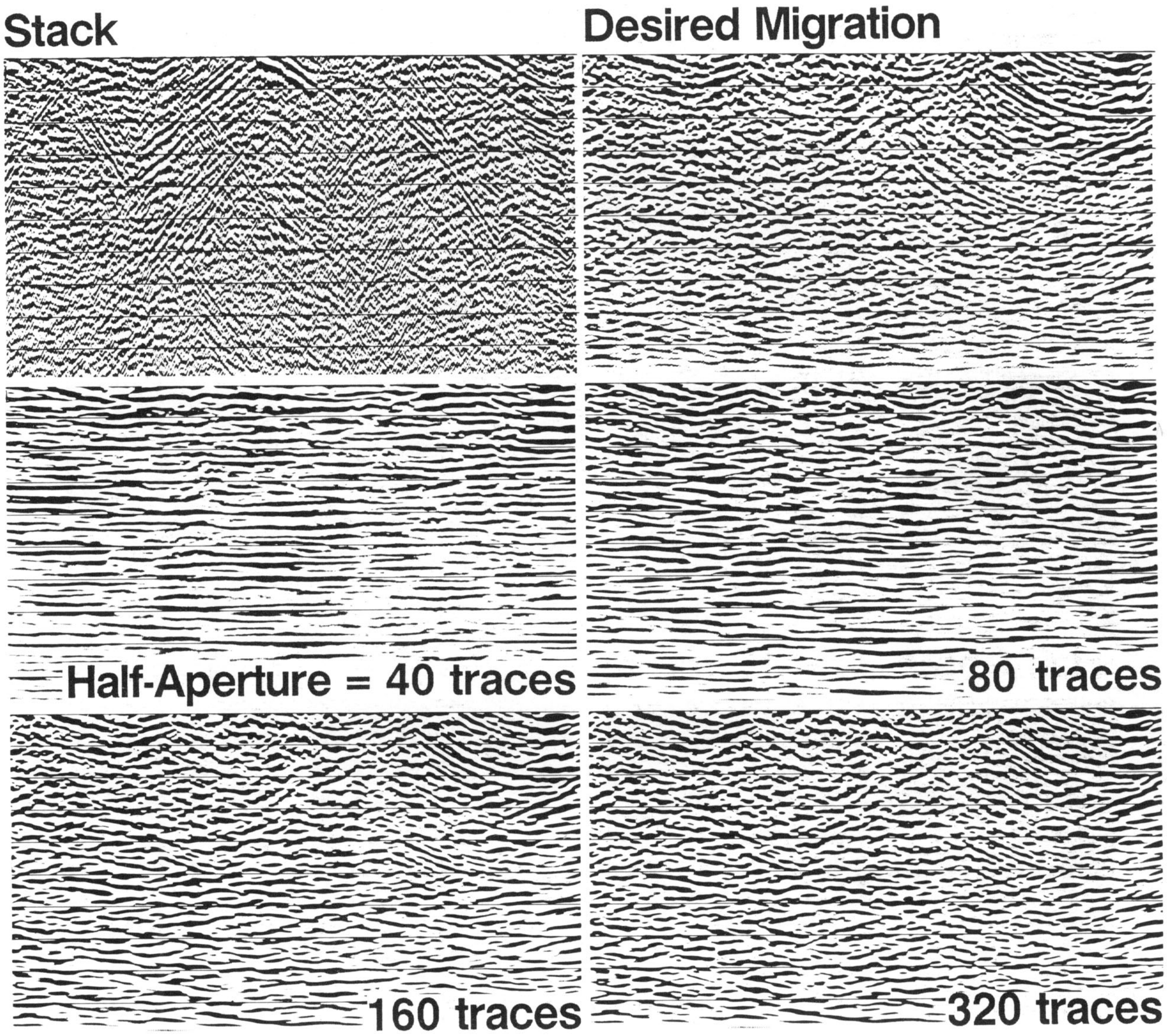


FIG. 4-45. Tests for aperture width in Kirchhoff migration. Insufficient aperture width causes spurious horizontal events in deep, noisy part of a stacked section.

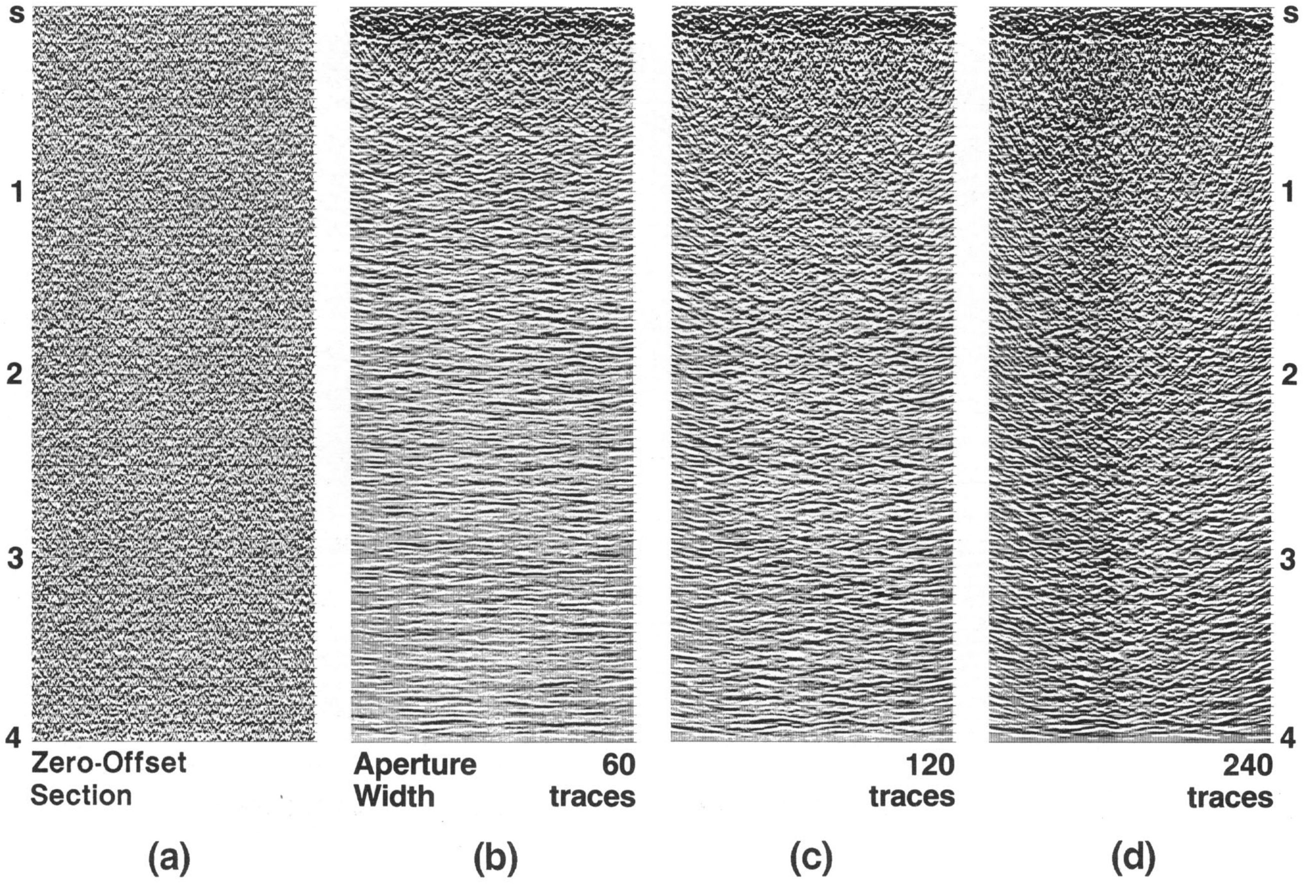


FIG. 4-46. Tests for aperture width in Kirchhoff migration. Input to migration is a pure-noise section (a). Note the spurious horizontal events in the deeper part of the section after migration using small aperture (60 traces); these gradually disappear at increasingly large apertures.

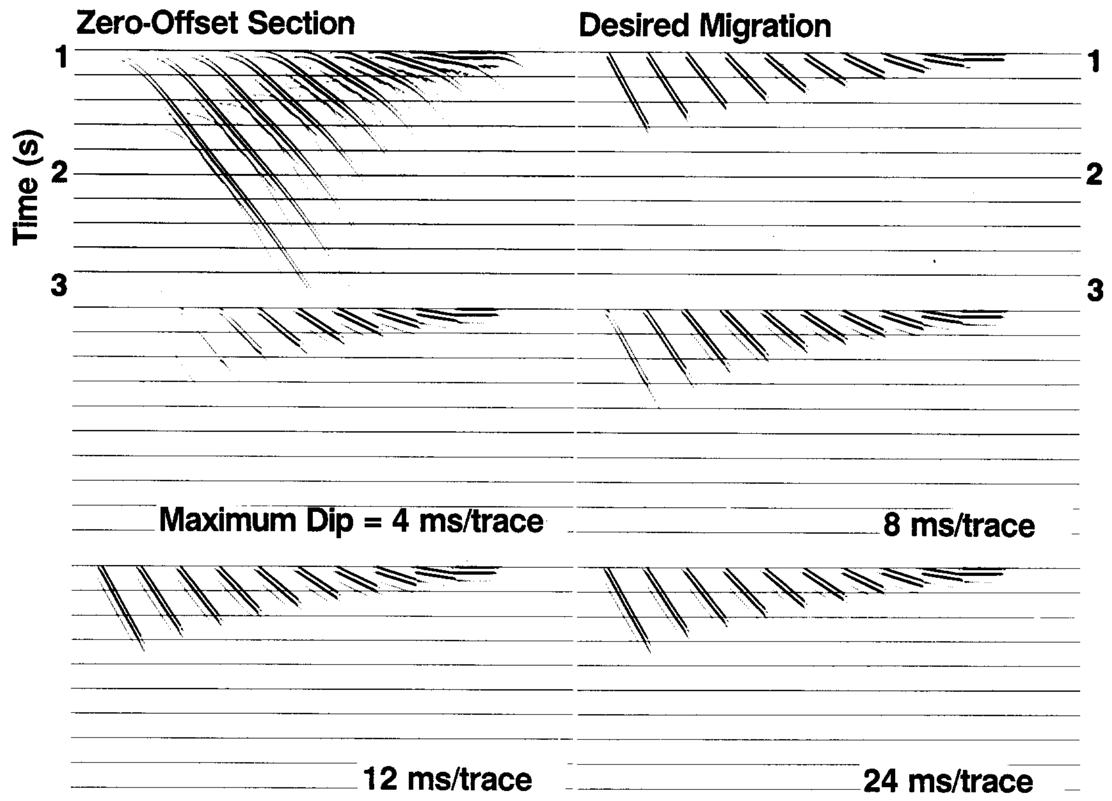


FIG. 4-47. Tests for maximum dip to migrate in Kirchhoff migration. Any event with a dip greater than the maximum specified dip value is filtered out during migration.

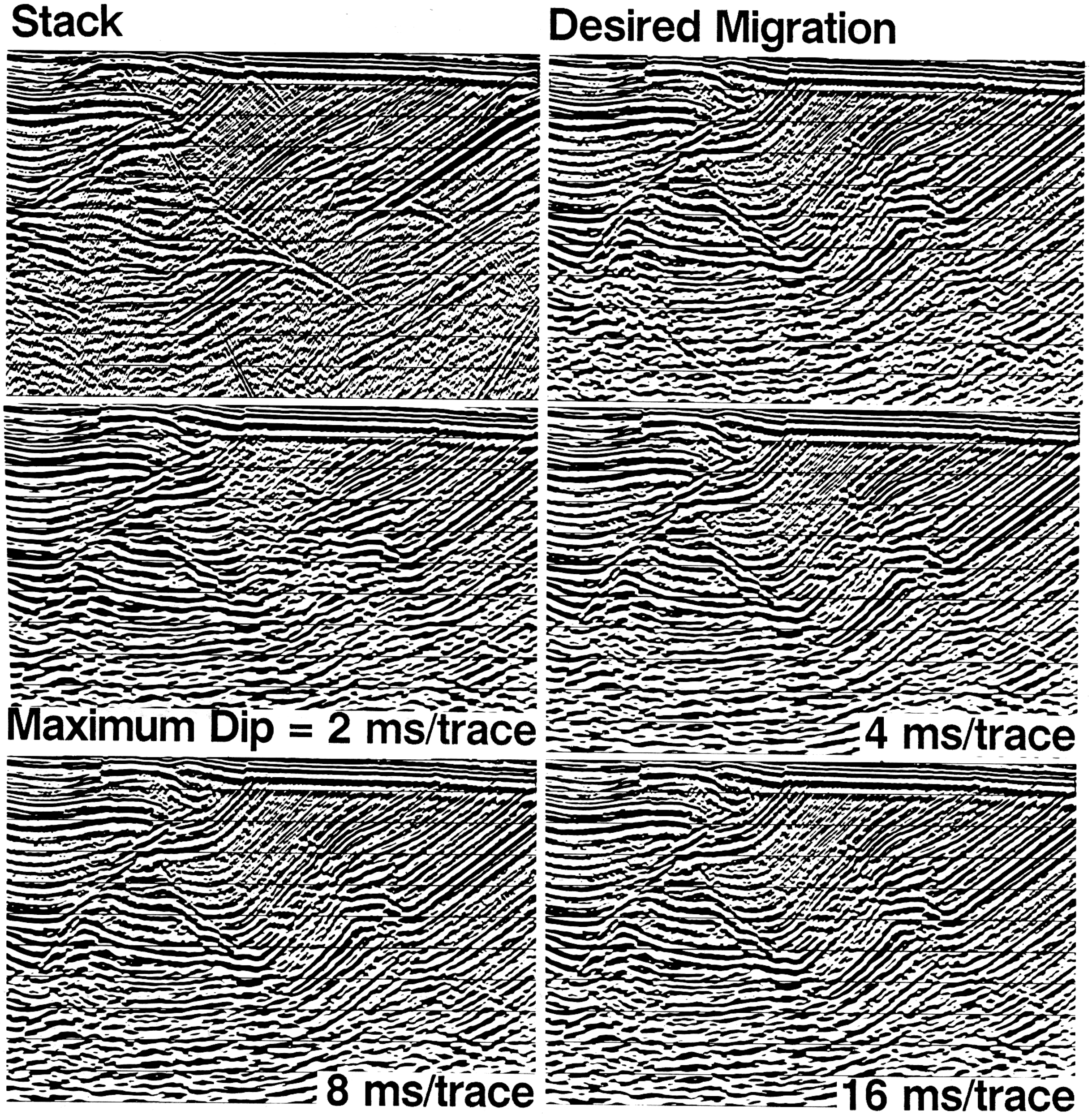


FIG. 4-48. Tests for maximum dip to migrate in Kirchhoff migration. A low value for maximum dip to migrate can be hazardous. All dips of interest must be preserved during migration.

cost. From Figure 4-40, note that the Kirchhoff migration impulse response can be limited to various maximum dips. The smaller the maximum allowable dip, the smaller the aperture. This combination of maximum aperture width and maximum dip limit determines the actual effective aperture width used in migration. In particular, diffraction hyperbolas along which summation is done are truncated beyond the specified maximum dip limit.

A field data example of testing the maximum dip parameter is shown in Figure 4-48. Some steep dips are lost on the section that corresponds to the 2 ms/trace maximum allowable dip. The 8 ms/trace dip appears to be optimum. The maximum dip parameter must be chosen carefully so that the steep dips of interest in the input section are preserved. Finally, dip value can be changed spatially and in time; however, practical implementation can be cumbersome.

Velocity Errors

We now examine the response of Kirchhoff migration to velocity errors. Figure 4-49 shows the diffraction hyperbola and migrations using the 2500 m/s medium velocity

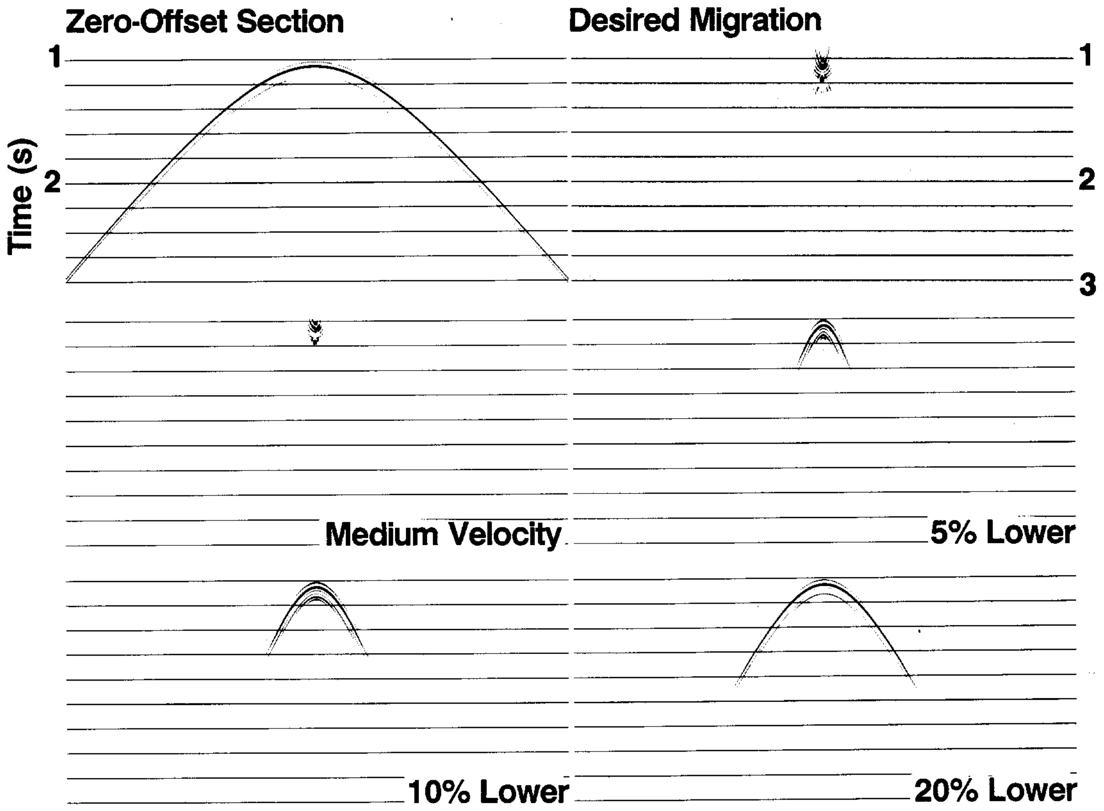


FIG. 4-49. Tests for velocity errors in Kirchhoff migration. Velocities lower than the actual medium velocity cause undermigration, that is, an incomplete collapse of diffractions.

and 5, 10, and 20 percent lower velocities. With increasingly lower velocities, the diffraction hyperbola is collapsed less and less; it is undermigrated. On the other hand, with increasingly higher velocities, it is overmigrated (Figure 4-50). Similarly, the under-and over-migration effects due to the use of erroneously low or high velocities on the dipping events model are seen in Figures 4-51 and 4-52, respectively. The medium velocity for the dipping events model is 3500 m/s. For comparison, the correct position of the event with the steepest dip *AB* is superimposed on the results of migrations with different velocities. Note that the amount of undermigration and overmigration because of velocity errors is more for steeper dips. (Compare these with the desired migration.) Remember that the steeper the dip, the more sensitive migration is to velocity errors as quantitatively concluded earlier [equations (4.1) and (4.2)].

Tests of the velocity error effect on field data are shown in Figures 4-53 and 4-54. From the migrated sections in Figure 4-53, note that the bowtie becomes increasingly less resolved at lower velocities, which indicates undermigration. The stacked section in Figure 4-54 contains what seems to be strong diffractions, but really are tight synclines and anticlines. (Refer to the result of the desired migration.) In Figure 4-54, the events cross each other on the output section that is migrated with 20 percent higher velocities than those believed reasonable, which indicates overmigration.

4.3.2 Finite-Difference Migration in Practice

From the numerical example in Section 4.2.2, finite-difference migration is implemented using either an explicit or implicit scheme [equations (4.11) and (4.12), respectively]. In this section, the implicit finite-difference migration algorithm, which is based on the parabolic approximation to the scalar wave equation (Appendix C.2) is considered. This approximation theoretically limits the algorithm to handle dips up to 15 degrees. However, in practice, it can handle dips up to 35 degrees with sufficient accuracy.

First, as we did for the Kirchhoff migration, we examine the impulse response of the implicit scheme (Figure 4-55). The desired migration of a single isolated wavelet on a single trace in the zero-offset section is a semicircle; this is the impulse response of a migration algorithm with no dip limitation (up to 90 degrees). The impulse response of the 15-degree equation is, in theory, an ellipse (Claerbout, 1985) as seen in Figure 4-55. The nature of the dispersive noise pattern inside the ellipses is discussed in the next section on depth step size. Isolated noise spikes in field data can introduce such noise patterns on migrated sections.

The parts of the responses above the small circles correspond to the evanescent energy, while the parts below

(Text continued on page 283)

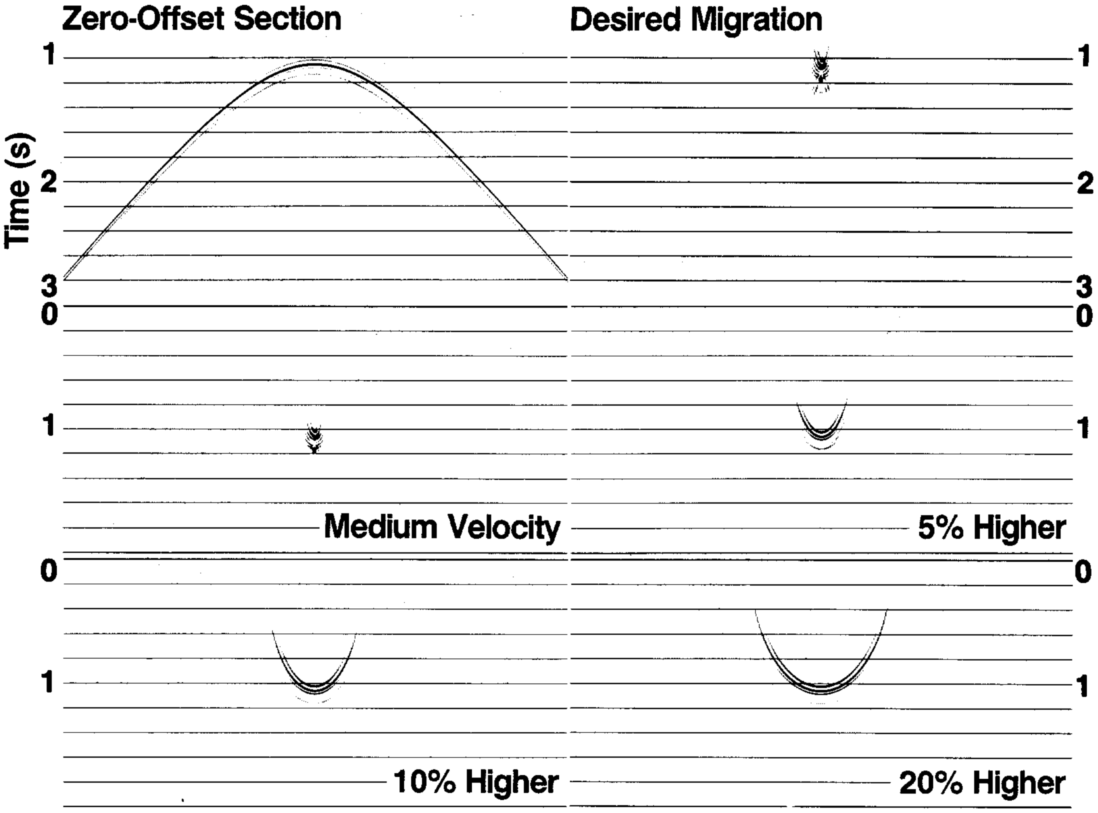


FIG. 4-50. Tests for velocity errors in Kirchhoff migration. Velocities higher than the actual medium velocity cause overmigration of a diffraction hyperbola.

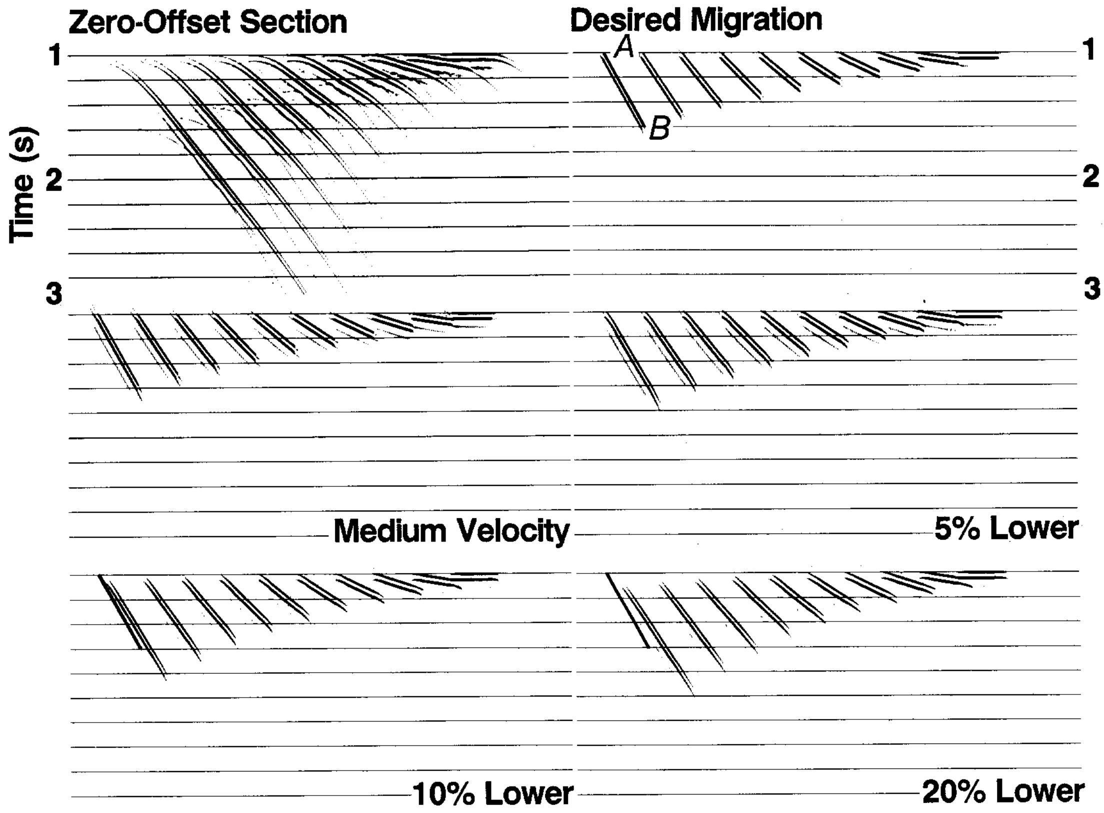


FIG. 4-51. Tests for velocity errors in Kirchhoff migration. Velocities lower than the actual medium velocity cause undermigration of dipping events.

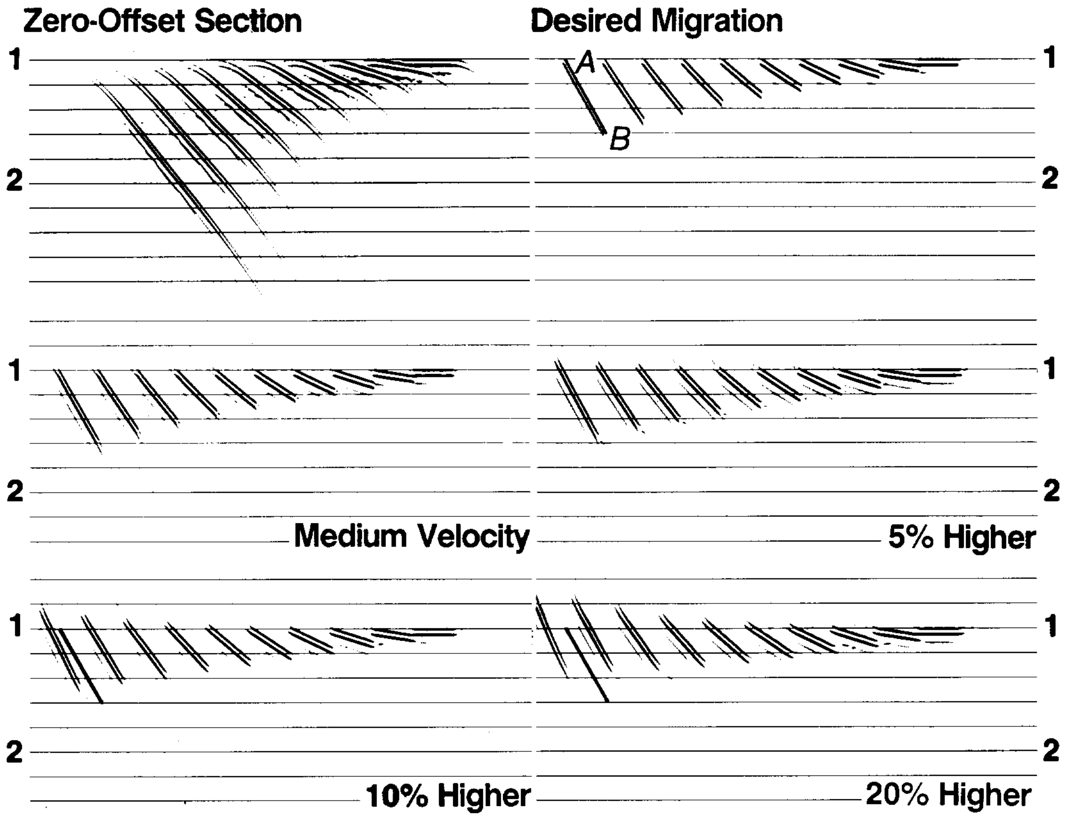
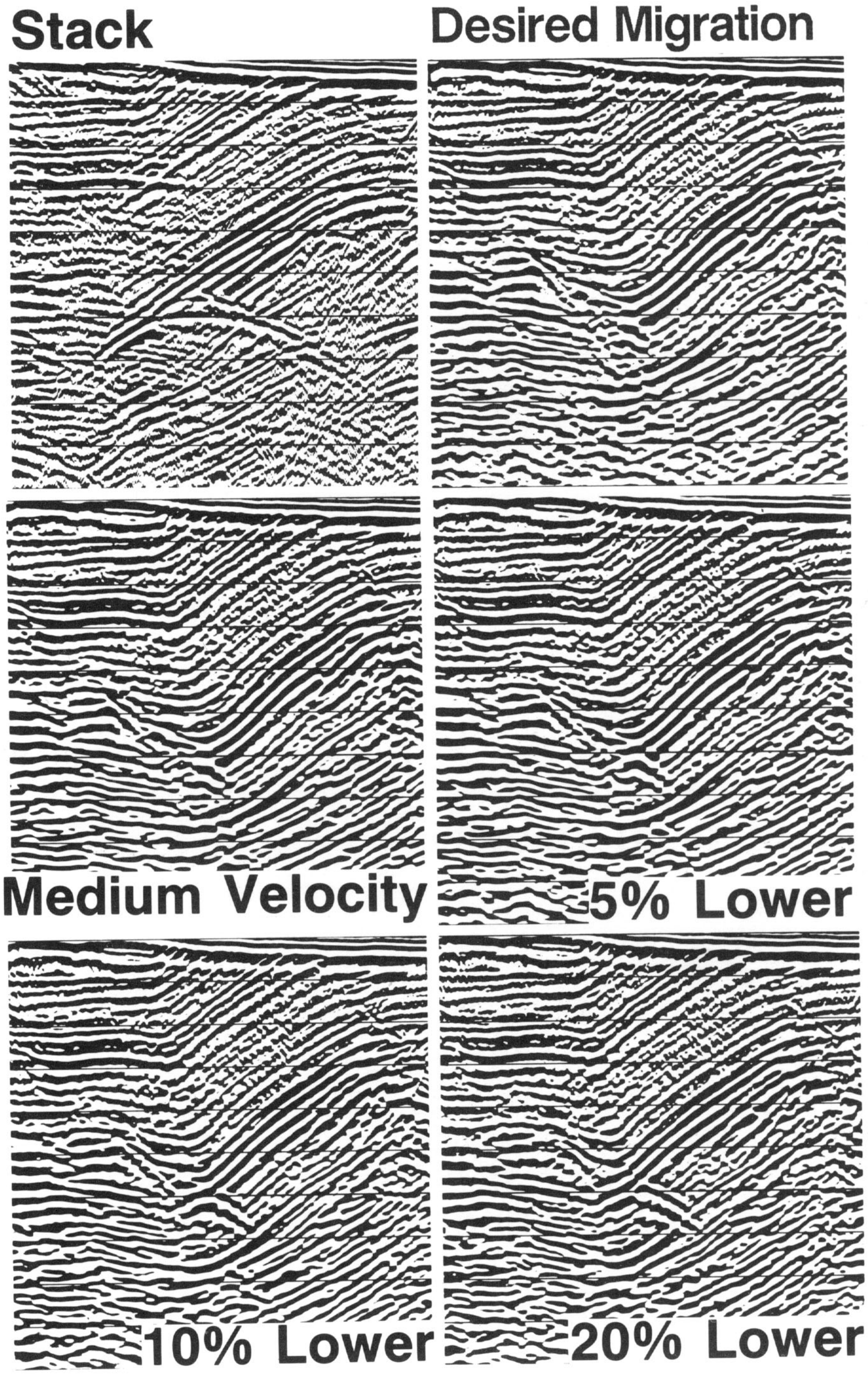


FIG. 4-52. Tests for velocity errors in Kirchhoff migration. Velocities higher than the actual medium velocity cause overmigration of dipping events. The top part of the steepest dip in the section migrated with 20 percent higher velocity has migrated out of the section.



Stack

Desired Migration

Medium Velocity

5% Lower

10% Lower

20% Lower

FIG. 4-53. Tests for velocity errors in Kirchhoff migration. Undermigration manifested as inadequate handling of the bowtie is caused by the use of velocities lower than velocities believed to be optimum (indicated as medium velocity).

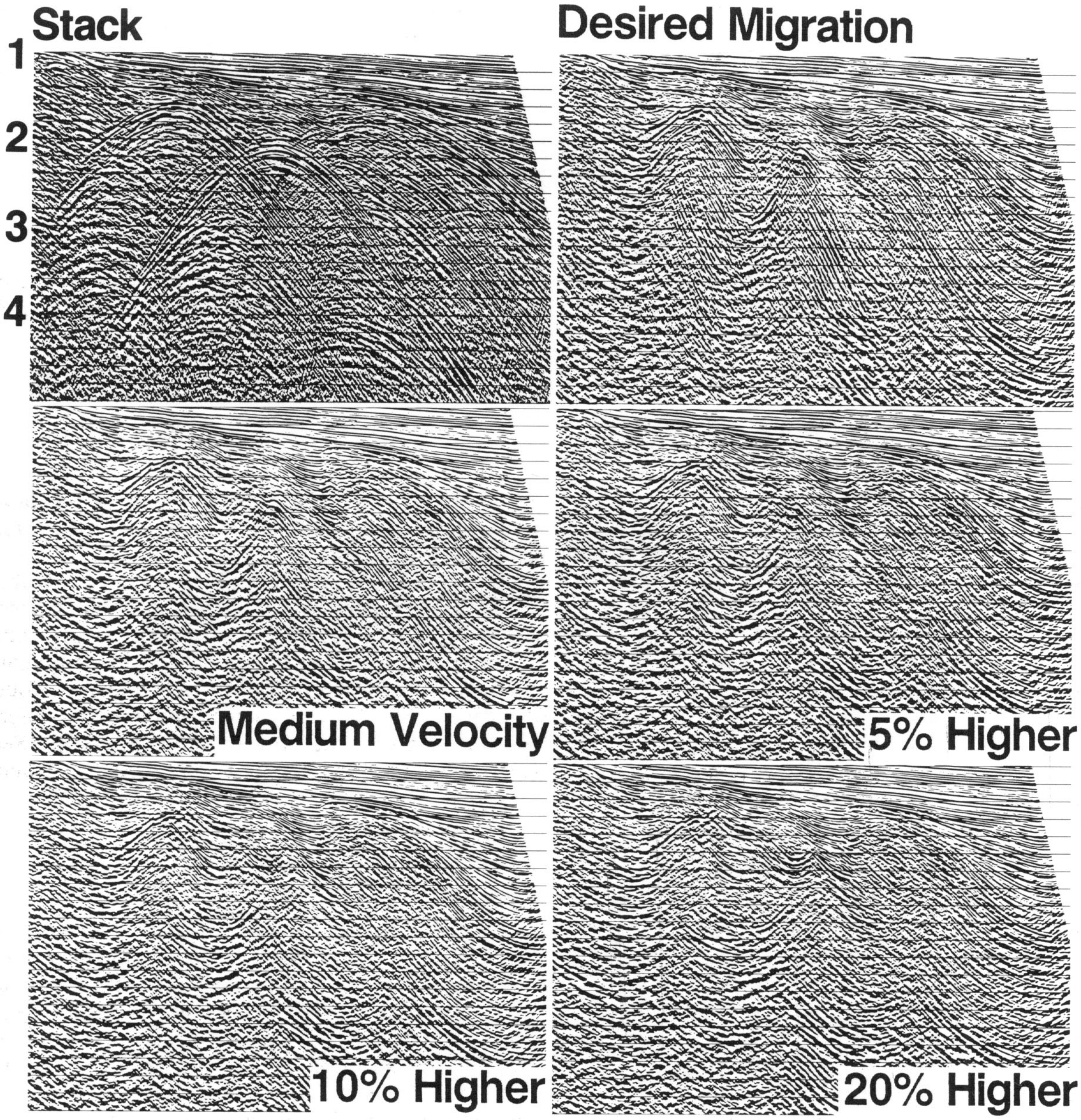


FIG. 4-54. Tests for velocity errors in Kirchhoff migration. Overmigration manifested as crossing events is caused by the use of velocities higher than velocities believed to be optimum (indicated as medium velocity). (Data courtesy Meridian Oil Inc.)

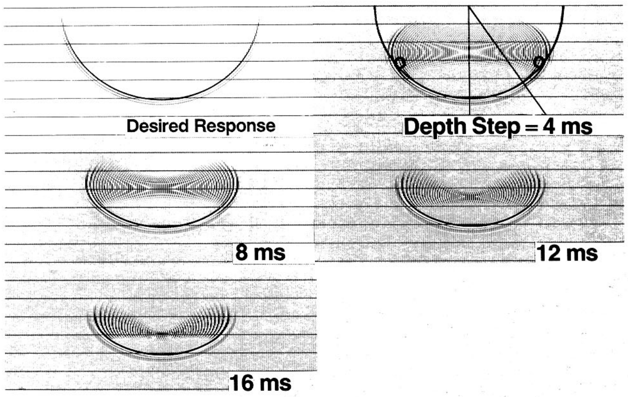


FIG. 4-55. The impulse response (depth step = 4 ms) of the 15-degree equation is an ellipse. Small circles define the boundary between the propagation zone (below the circles) and the evanescent zone (above the circles). The latter seems to be suppressed at larger depth steps. For comparison, the desired response has been superimposed on the impulse response (4-ms frame) of the 15-degree equation.

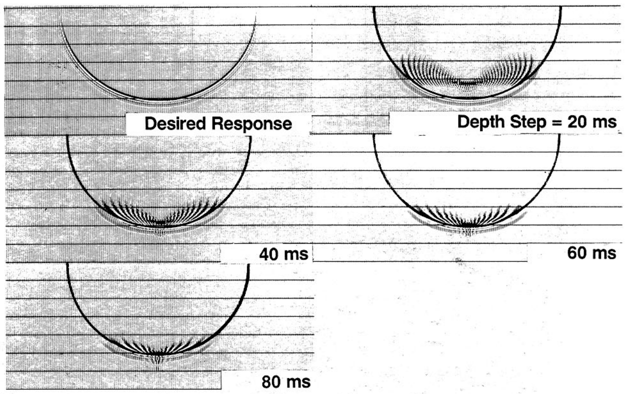


FIG. 4-56. Effect of 20- to 80-ms depth step size on the impulse response of the 15-degree equation. For comparison, the desired response has been superimposed on the impulse responses of the 15-degree equation.

the circles correspond to the propagating energy (Claerbout, 1985). The parts below the circles are the useful part of the response. The evanescent energy travels horizontally and is characterized by imaginary wavenumbers in the z-direction. Imaginary z-wavenumbers occur when the quantity in the square root in equation (C.28) becomes negative. For the exact wave equation, the evanescent energy normally decays rapidly in depth and, thus, is not expected to be present in recorded wave fields. However, the impulse response of the 15-degree finite-difference algorithm suggests propagation in the region of evanescence. This is not desirable; the parts of the elliptical wavefront above the small circles should be removed. Use of a depth step size that is greater than the input sampling rate (Figure 4-55) tends to suppress the response in the evanescent region. Of course, if excessively large depth steps are used, the wavefront is truncated further in the propagating zone (Figure 4-56).

The impulse response with a 4-ms depth step (Figure 4-55) is used to estimate the maximum dip that the implicit finite-difference algorithm can handle without serious amplitude distortions or phase errors. This is done by superimposing the desired semicircular response and measuring the angle between the indicated lines. For the implicit case, this angle is about 35 degrees. Hence, the 15-degree equation can be used to migrate dips up to approximately 35 degrees with sufficient accuracy. This is primarily because errors associated with finite-difference approximations used in particular implementations of the 15-degree equation usually are adjusted to cancel some of the theoretical error associated with the (continuous) differential equation.

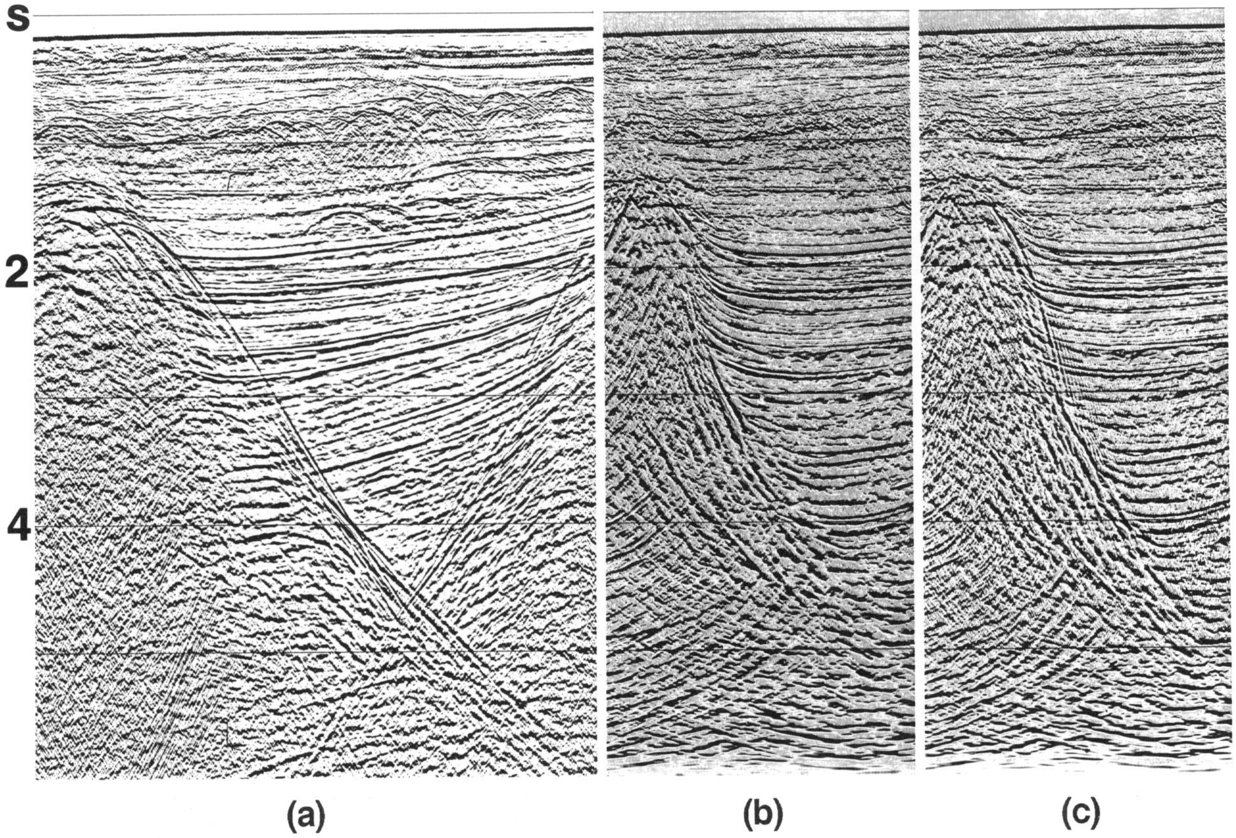
The dip-limited nature of the parabolic equation causes undermigration of the steep flanks of diffractions and steeply dipping events. This is demonstrated in Figure 4-57. The two prominent features, diffraction D and dipping event B, are located as shown in Figure 4-58 before and after migration. Nevertheless, for many field data applications, finite-difference migration based on the parabolic equation is adequate.

A steep-dip implicit finite-difference migration technique is presented in Section 4.3.4. This algorithm is based on the 45-degree approximation to the scalar wave equation and is implemented in the frequency-space domain. A mathematical discussion of the basic theory is found in Appendix C.3. Use of this steep-dip finite-difference scheme is made not only in time migration (Section 4.3.4), but also in depth migration (Section 5.2) and 3-D migration (Section 6.5).

Depth Step Size

Downward continuation of a wave field recorded at the surface of the earth is a fundamental ingredient of seismic migration. The process takes place in the computer at discrete depth intervals (Section 4.2.2). Depth step size governs performance of finite-difference migration. Inappropriate specification of this parameter can cause artifacts in the migrated section. Figure 4-59 shows the constant-velocity diffraction hyperbola and the implicit finite-difference migrations using four different depth steps. Large depth steps cause undermigration as well as kinks along the flank of the diffraction curve (especially apparent in 60- and 80-ms cases). At smaller depth steps, such as the 20- and 40-ms cases, more energy collapses to the apex. Smaller depth steps do not significantly improve focusing of the diffraction hyperbola (Figure 4-60).

Undermigration of the diffraction energy along the steep flanks of the hyperbola is due to the parabolic approximation to the scalar wave equation. The dispersive noise that accompanies the undermigrated energy is an effect of approximating differential operators with difference operators. The accuracy of this approximation decreases at



(a) (b) (c)

FIG. 4-57. (a) CMP stack, (b) desired migration by phase-shift method, (c) 15-degree finite-difference migration. The finite-difference migration based on the parabolic equation has the inherent property of undermigrating steep flanks of the diffractions and steeply dipping events. Here, depth step size = 20 ms. See Figure 4-58 for a sketch of the migration results.

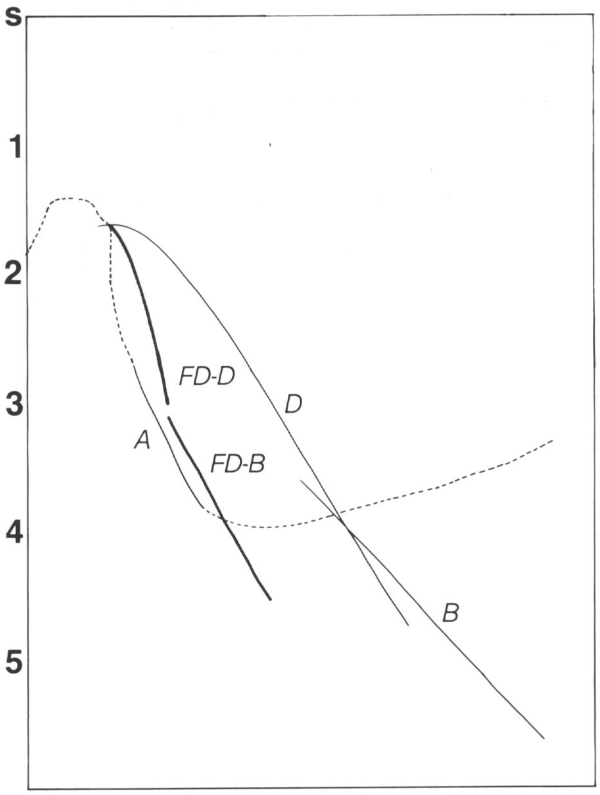


FIG. 4-58. A sketch of the diffraction *D* and steeply dipping event before (*B*) and after (*A*) desired migration from the sections in Figure 4-57. Here, *FD-B* = finite-difference migration of the dipping event and *FD-D* = finite-difference migration of the diffraction.

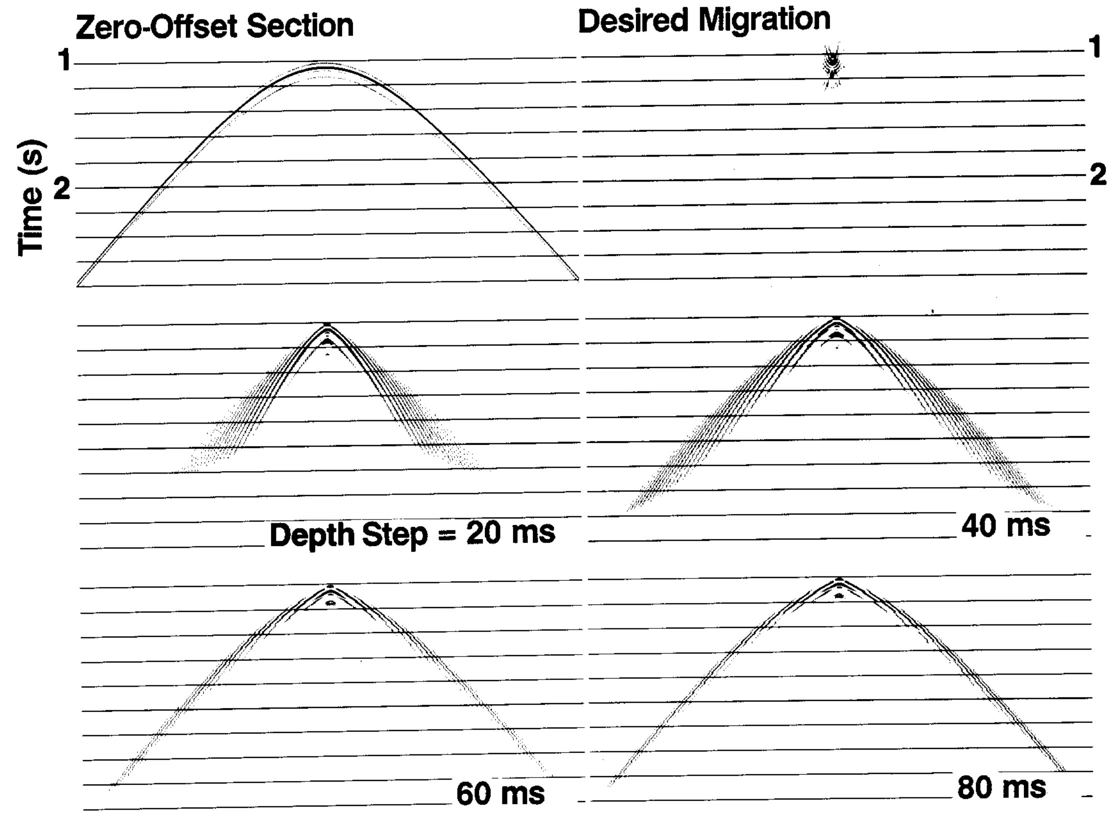


FIG. 4-59. Effect of 20- to 80-ms depth step size on the 15-degree equation's ability to collapse the diffraction hyperbola.

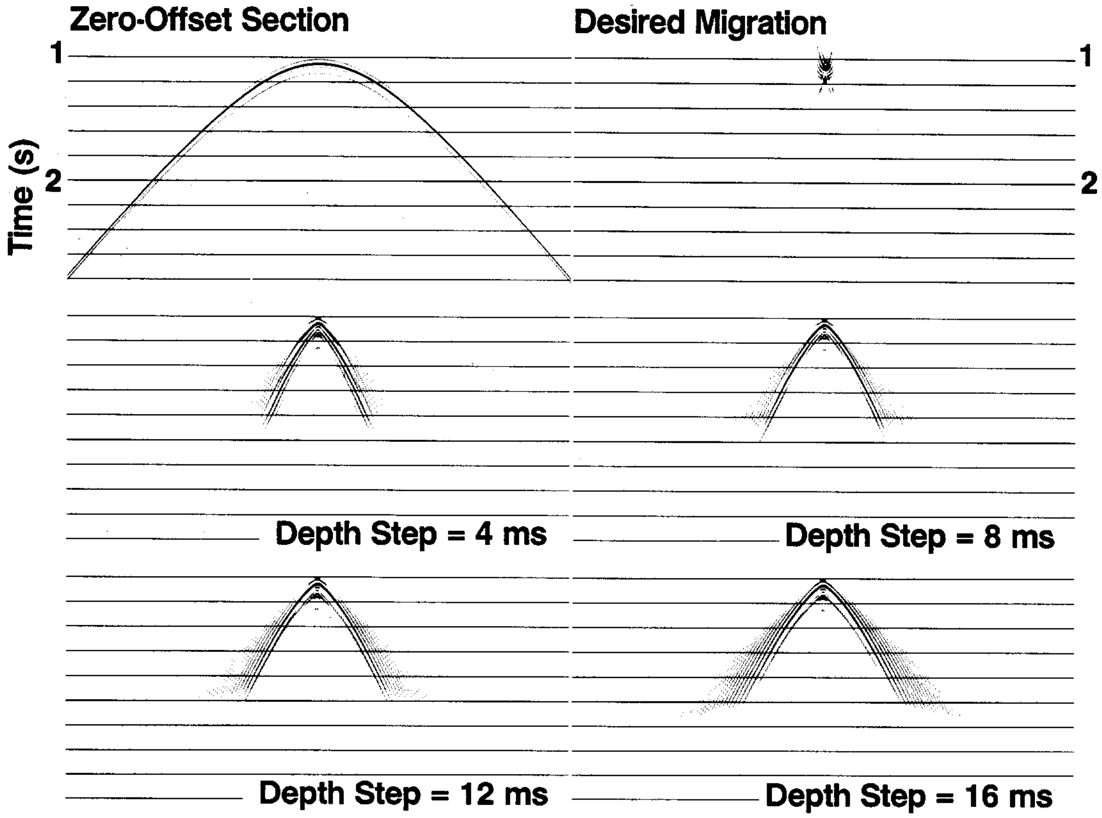


FIG. 4-60. Effect of 4- to 16-ms depth step size on the 15-degree equation's ability to collapse the diffraction hyperbola.

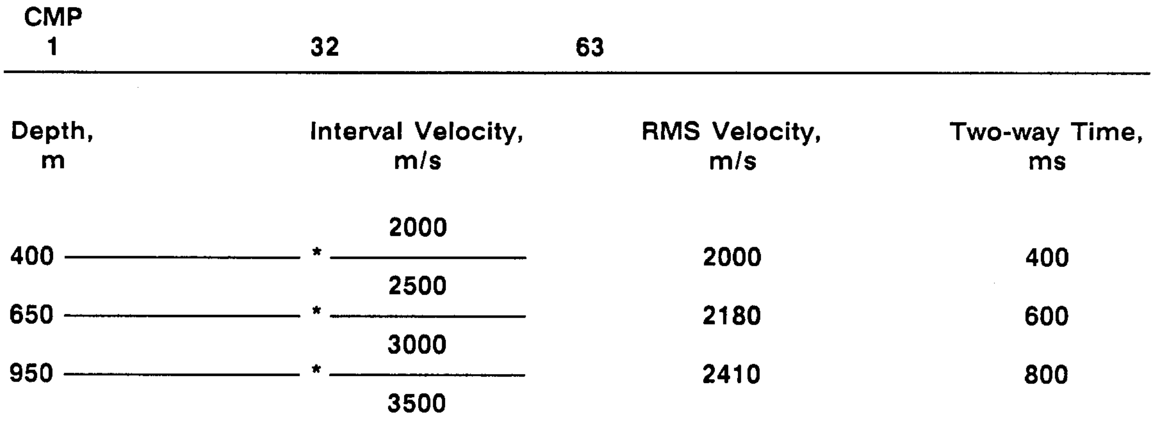


FIG. 4-61. Depth model for three point scatterers buried in a vertically varying velocity medium. The asterisks indicate the positions of the point scatterers.

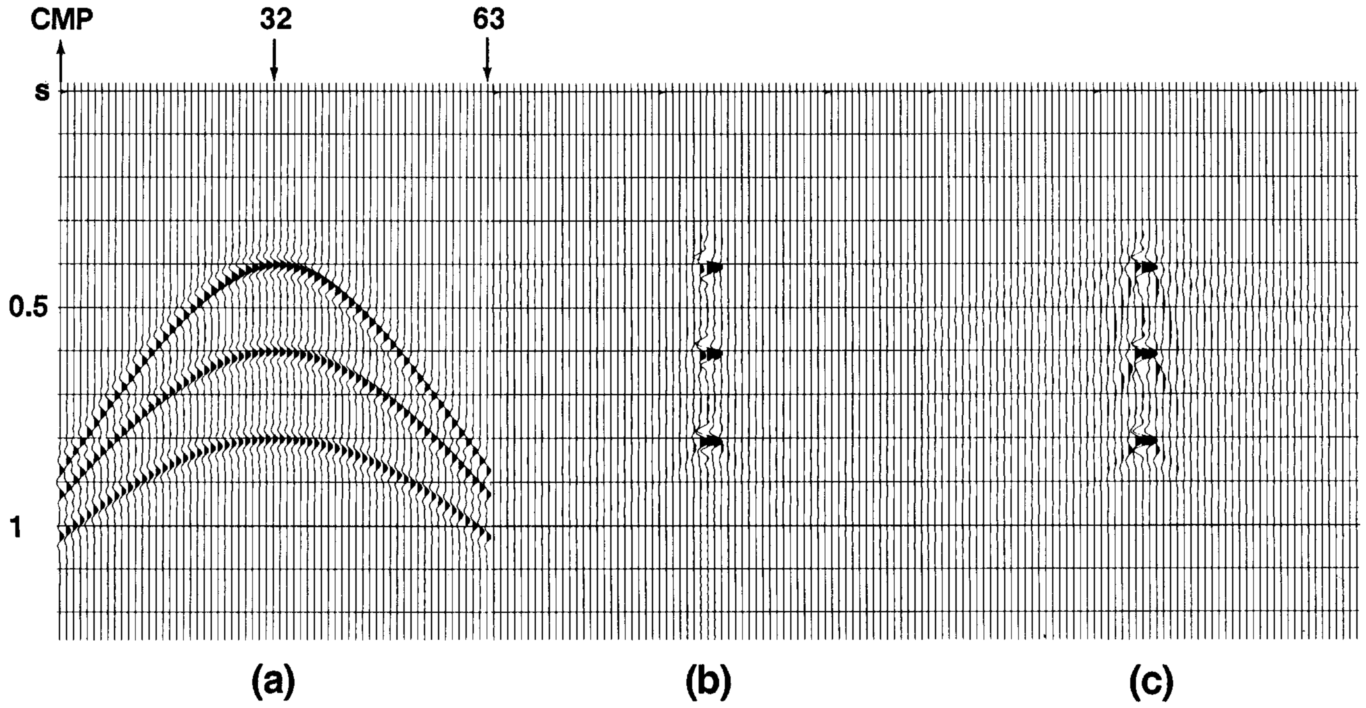


FIG. 4-62. Migration of diffractions in a horizontally layered medium. The velocity-depth model is shown in Figure 4-61. (a) Zero-offset section; (b) desired migration by phase shift method; (c) 15-degree finite-difference migration.

large frequencies and wavenumbers (Claerbout, 1985). Thus, the dispersive noise becomes less with smaller trace spacing and sampling in depth and time. For example, the difference operator of equation (4.13) becomes an increasingly better approximation to the differential operator of equation (4.14) as Δt is made smaller. To strongly emphasize the presence of the dispersive noise, migrated sections have been displayed with the same display gain level as the input section. The dispersion normally is much less pronounced on field data.

Diffractions in a layered medium respond to the parabolic approximation in a similar manner. Consider the velocity model in Figure 4-61. The zero-offset section corresponding to this velocity model is shown in Figure 4-62a. Note that the undermigration effect is more pronounced for the two shallow scatterers when compared to the deepest scatterer (Figure 4-62c). The zero-offset responses for the shallow scatterers have steeper flanks than the deepest scatterer; thus, the 15-degree equation has difficulty migrating these steep flanks.

Figure 4-63 shows the dipping events model and implicit finite-difference migration results using four different depth step sizes. For comparison, the event with the steepest dip *AB* is superimposed on the results. We can make the following inferences:

1. Increasing depth step size causes more and more undermigration at increasingly steep dips.
2. The waveform along reflectors is dispersed at steep dips and large depth steps.
3. Kinks occur along reflectors at discrete intervals that correspond to the depth step size. Kinks are more pronounced at increasingly steeper dips.

The first inference is due to the parabolic approximation, while the second is due to differencing approximations. The third is due to gradual undermigration toward the bottom of each depth step. The kinks are good for diagnostics; their presence indicates that the depth step size that is used is too coarse for the dips present in the data. In that case, smaller depth step size should be used; then the kinks disappear altogether (Figure 4-64). Nevertheless, kinks that characterize undermigration can be eliminated by a local adjustment of migration velocities or interpolation between the wave fields at the adjacent depth steps. (Kinks were suppressed in the field data examples in this section.)

It is apparent from Figure 4-63 that migration with 20-ms depth step has the least dispersion with the least undermigration; i.e., optimum accuracy in event positioning. Further decreasing depth step size does not significantly improve migration. For example, Figure 4-64 shows that migration with 4-ms depth step (equal to the sampling interval) has some dispersion along the steeply dipping reflectors as postcursors rather than precursors, which are seen in large depth steps. This behavior suggests that taking smaller depth steps does not necessarily mean a better quality migration free of the artifacts that occur with the finite-difference method.

Figures 4-65 and 4-66 show the migrations of the stacked section in Figure 4-57a using five different depth steps.

(Text continued on page 292)

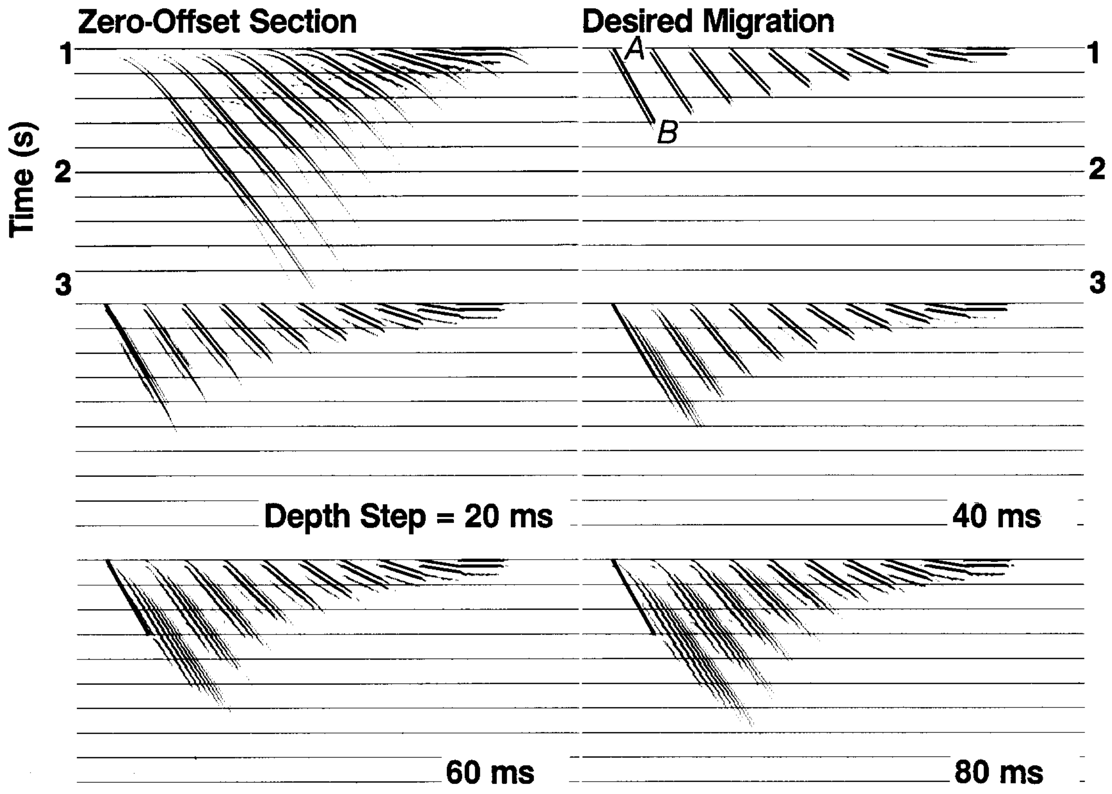


FIG. 4-63. Effect of 20- to 80-ms depth step size on the 15-degree equation's ability to migrate dipping reflectors.

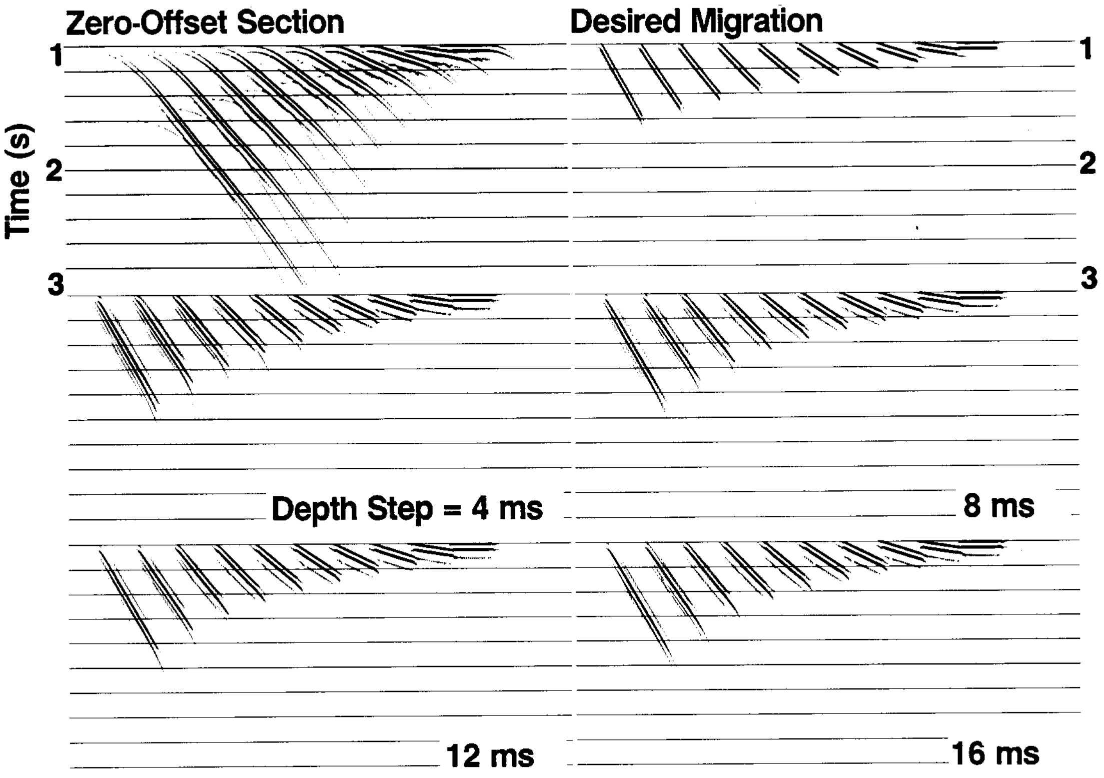


FIG. 4-64. Effect of 4- to 16-ms depth step size on the 15-degree equation's ability to migrate dipping reflectors.

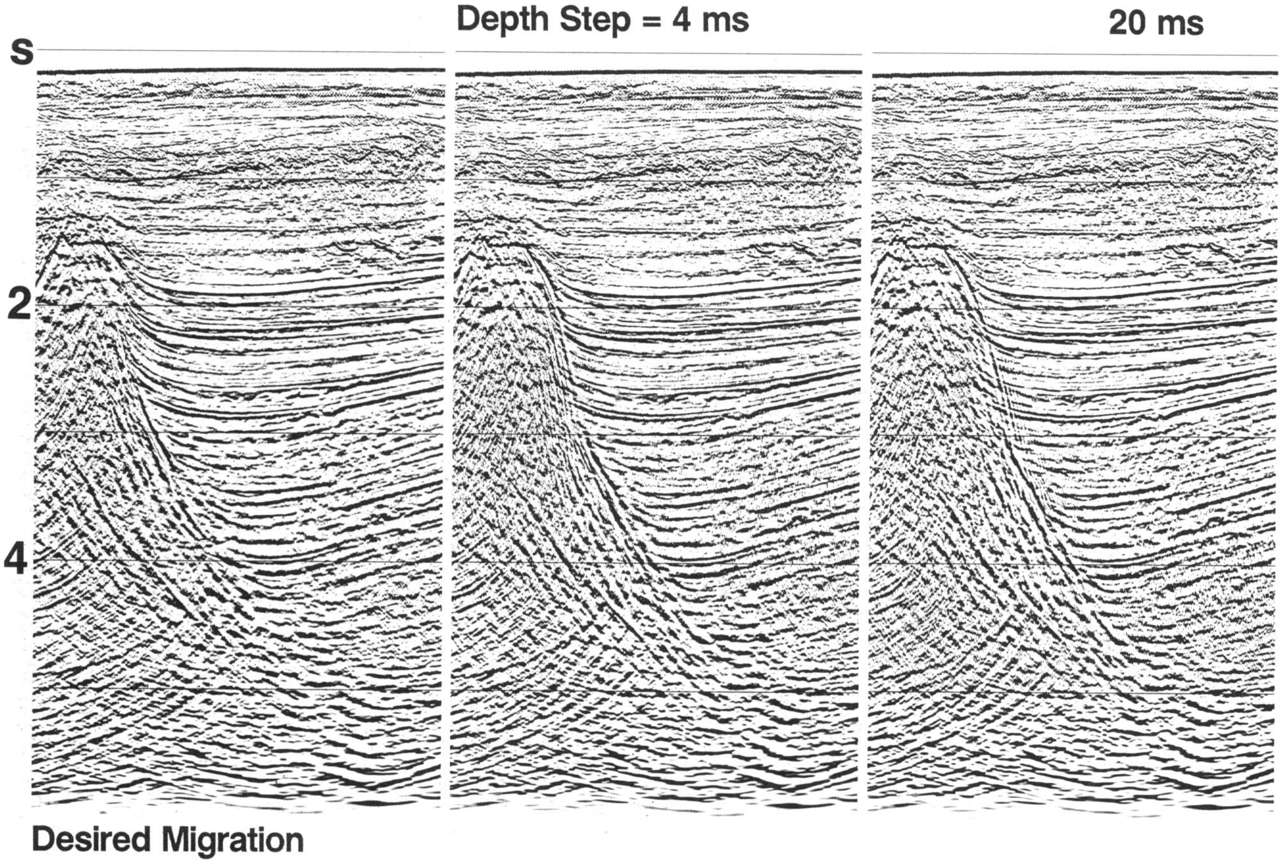


FIG. 4-65. Finite-difference migration, depth step tests. The larger the depth step, the more undermigration occurs. The input stack is shown in Figure 4-57a.

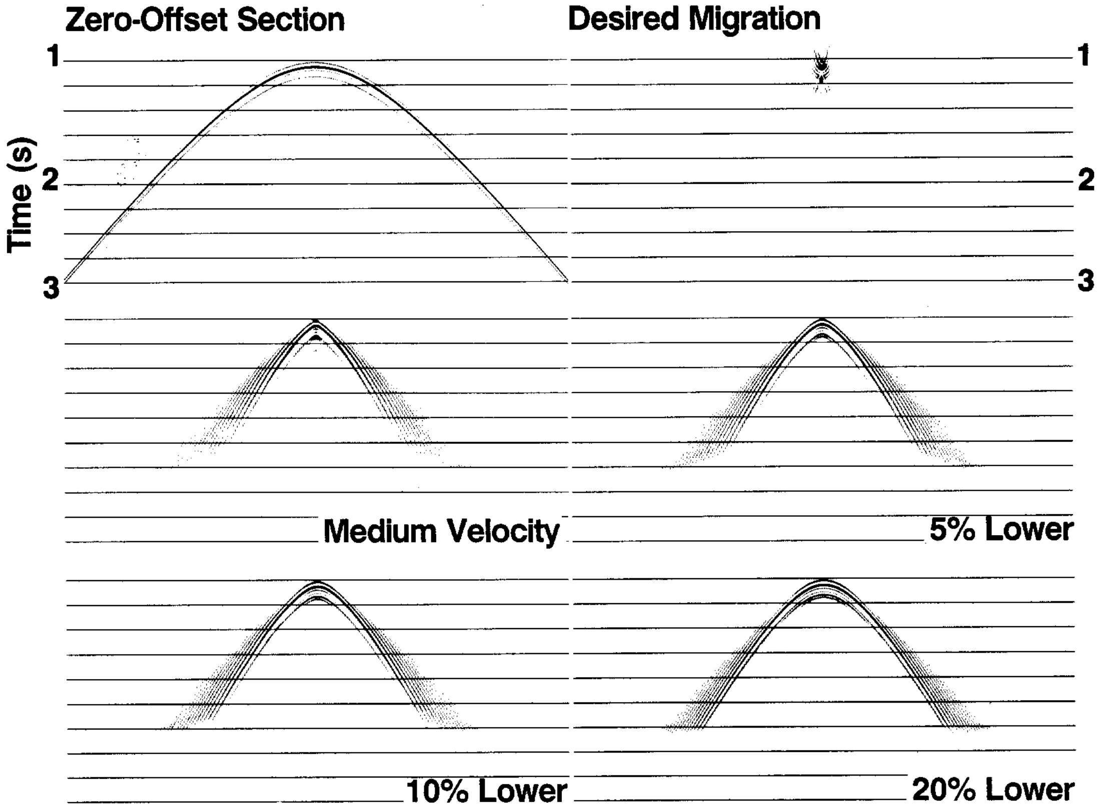


FIG. 4-67. Undermigration caused by the use of velocities lower than medium velocities reinforced with the 15-degree approximation. Compare these results with those obtained from the Kirchhoff summation (Figure 4-49). Depth step size = 20 ms.

First, note that even with the optimum 20-ms depth step, the finite-difference approach yields an undermigrated section when compared to the desired migration. Second, as depth step size is increased, more undermigration occurs. Dispersion along the diffraction hyperbola is apparent at larger depth steps and is very similar to that observed in the synthetic model (Figure 4-59). The reflection off the flank of the salt dome also is increasingly undermigrated at larger depth steps.

Unfortunately, we do not have complete freedom in selecting the depth step size used in the finite-difference migration algorithms. For economic reasons, we want to take as large a depth step as possible. However, the implicit finite-difference schemes limit us to a depth step size range that will minimize the problems discussed above. The explicit schemes require relatively smaller depth step sizes for stability. Selection of an optimal depth step size that minimizes the undermigration and dispersive noise depends on a complicated interaction between the velocity function used in migration, sampling intervals in time and space, frequency content of the data, and dips present in the subsurface (Yilmaz, 1983).

Specifically, a depth step size may be optimal for a particular set of these parameters, but the same depth step may not be optimal for a different combination. Furthermore, given the differential equation (e.g., the parabolic equation), implementational details of the differencing scheme can have a visible impact on the fidelity of migration output. Diet and Lailly (1984) discuss an optimization procedure to minimize the dispersive noise caused by the differencing approximations. The main point to remember is that migration of steep dips generally requires a small depth step size. Practical considerations suggest that depth step size should be between one-half and one-full dominant period of the seismic data to be migrated (i.e., from 20 to 40 ms), depending on the steepness of the dips in the data.

Velocity Errors

We now examine the response of the implicit algorithm to velocity errors. Figure 4-67 shows the diffraction hyperbola and its migrations using the 2500 m/s medium velocity and 5, 10, and 20 percent lower velocities. Because of the 15-degree approximation, the diffraction hyperbola already is undermigrated with the medium velocity. At increasingly lower velocities, the undermigration effect becomes more serious. However, unlike Kirchhoff migration (Figure 4-50), the overmigration effect with a finite-difference algorithm is less pronounced

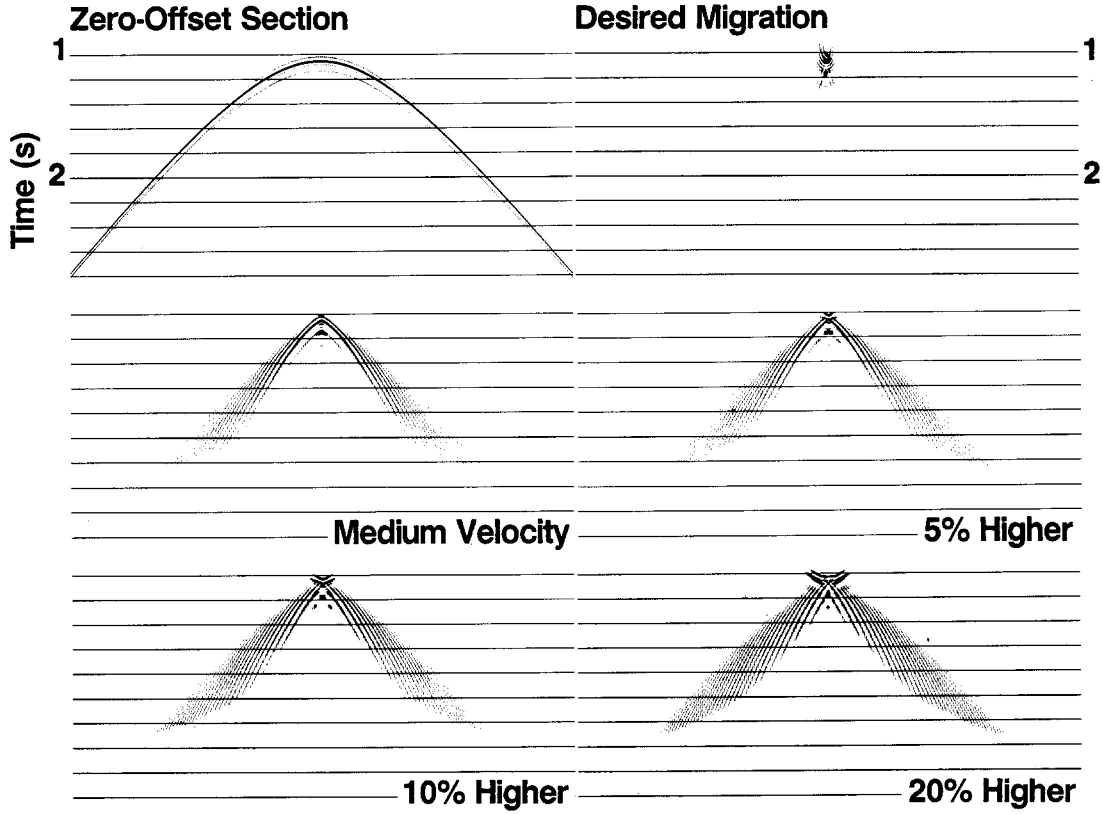


FIG. 4-68. Overmigration caused by the use of velocities higher than medium velocities in 15-degree finite-difference migration. Compare these results with those obtained from the Kirchhoff summation (Figure 4-50). Depth step size = 20 ms.

(Figure 4-68). Nevertheless, increasing dispersive noise is seen with higher velocities since higher velocities mean more migration and, thus, more dispersion. The inherently undermigrating character of the 15-degree finite-difference scheme seems to be counteracted by the normally overmigrating effect of higher velocities. Should we use velocities with greater than medium velocity when migrating with the finite-difference method? Appendix C.6 contains a brief theoretical discussion on the effect of velocity errors in migration with the parabolic equation.

Figure 4-69 is the dipping events model with migrations using the 3500 m/s medium velocity and velocities that are 5, 10, and 20 percent lower. For comparison, the correct position of the event with the steepest dip *AB* is superimposed on the migrated sections. Because of the parabolic approximation, the finite-difference algorithm inherently causes undermigration, even with the correct velocity.

As in any other migration method, velocity errors cause events to be mispositioned at increasingly steeper dips. Comparison of the finite-difference results (Figure 4-70) with those of the Kirchhoff summation (Figure 4-52) reveals that velocity errors cause more mispositioning of steep dips with the Kirchhoff summation than with the finite-difference method. The finite-difference method is rather robust in that it can tolerate large velocity errors. Therefore, use of a steep-dip algorithm, such as the Kirchhoff summation and f-k methods, requires a more accurate definition of velocities than a dip-limited algorithm, such as the 15-degree finite-difference method.

Again, migration with higher velocities causes a slight increase of dispersive noise, which accompanies the steepest dips in the section (Figure 4-70, the 20 percent higher velocity case). This is accompanied by a real overmigration effect of steep dips. The undermigrating nature of the algorithm seems to be compensated for by increasing the velocities by perhaps as much as 10 percent (also see Appendix C.6).

Velocity error tests on field data are shown in Figures 4-71 and 4-72. Figure 4-73 is a sketch of the under-and overmigration effects. When using velocities greater than medium velocities, note that overmigration is not as pronounced with the finite-difference migration (based on the parabolic equation) as it is with a 90-degree algorithm, such as the Kirchhoff or phase-shift method. (Compare Figure 4-72 with Figure 4-88.) On the other hand, when using velocities lower than medium velocities, the undermigration effect is more pronounced with the finite-difference migration based on the 15-degree approximation in comparison with the 90-degree algorithms, such as the Kirchhoff or phase-shift method. (Compare Figure 4-71 with Figure 4-87.) A way to compensate for the

(Text continued on page 298)

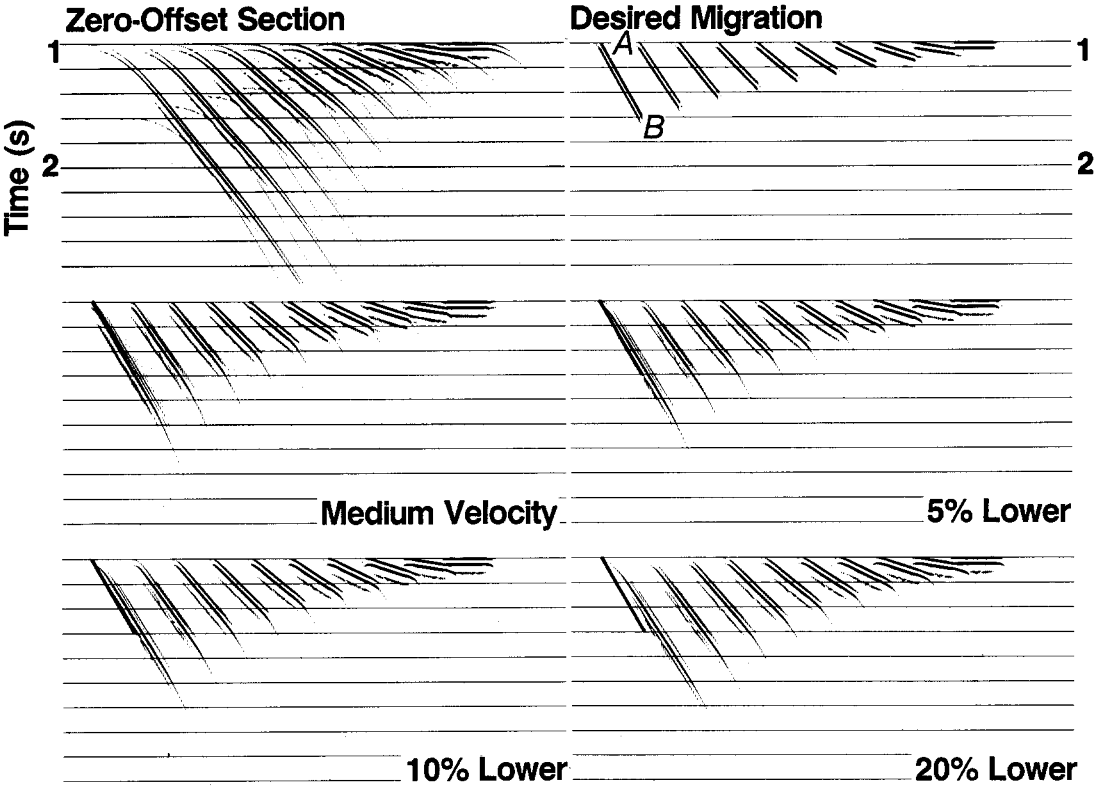


FIG. 4-69. Undermigration caused by velocities lower than medium velocities reinforced with the 15-degree approximation. Compare these results with those obtained from the Kirchhoff summation (Figure 4-51). Depth step size = 20 ms.

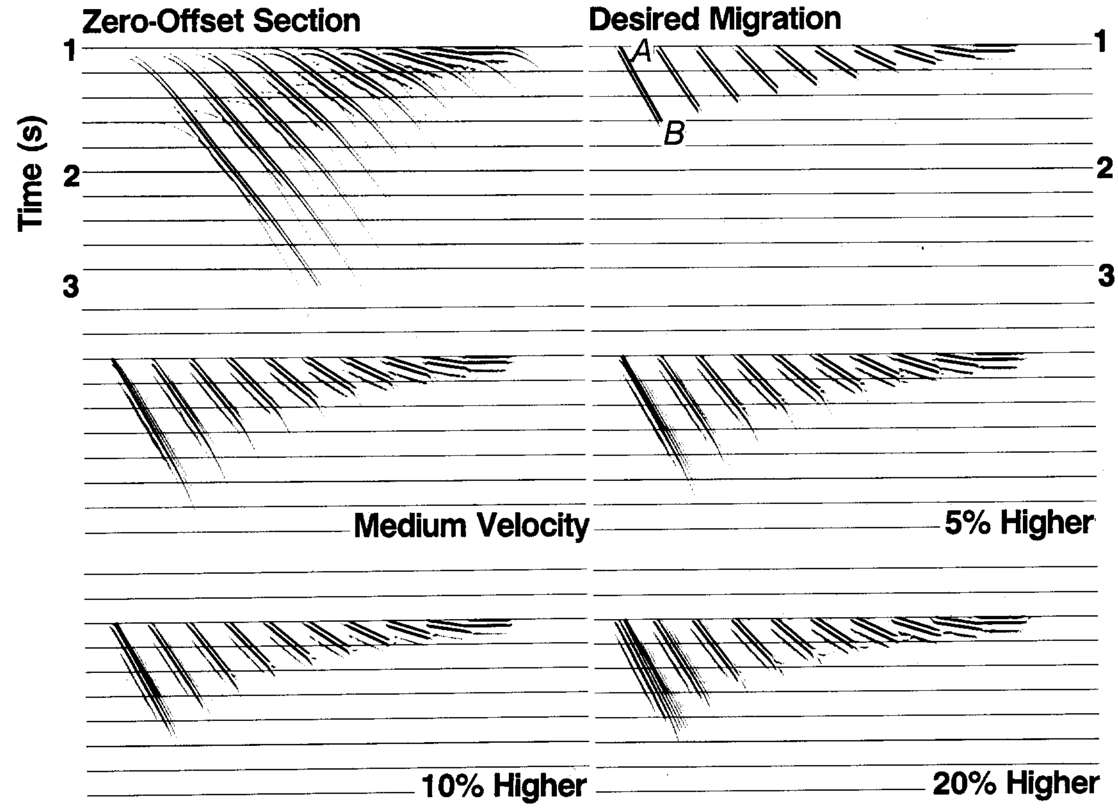


FIG. 4-70. Tests of velocity error in 15-degree finite-difference migration. Compare this with the desired migration and with the results obtained from the Kirchhoff summation (Figure 4-52). Depth step size = 20 ms.

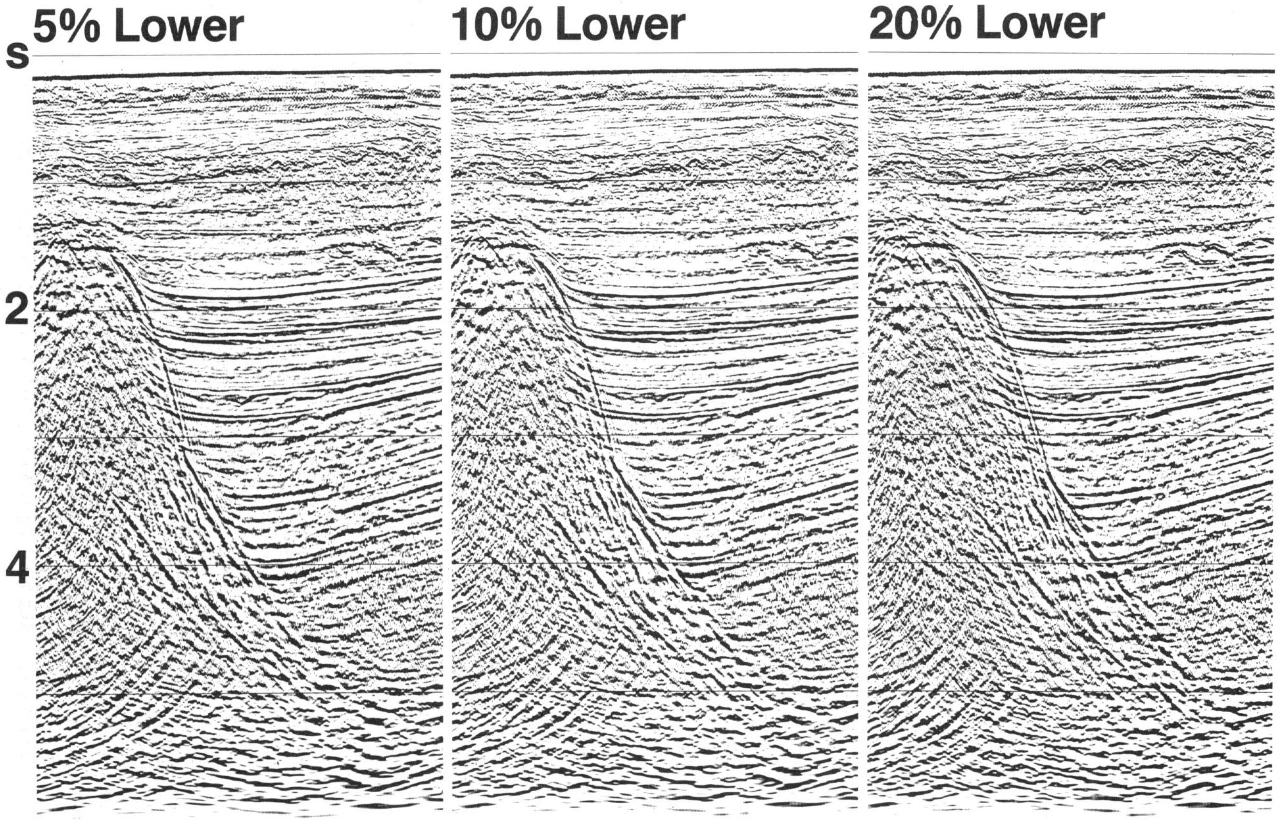


FIG. 4-71. Tests of velocity error in 15-degree finite-difference migration. The input stack and the desired migration are shown in Figure 4-57. See Figure 4-73 for a sketch of the migration results.

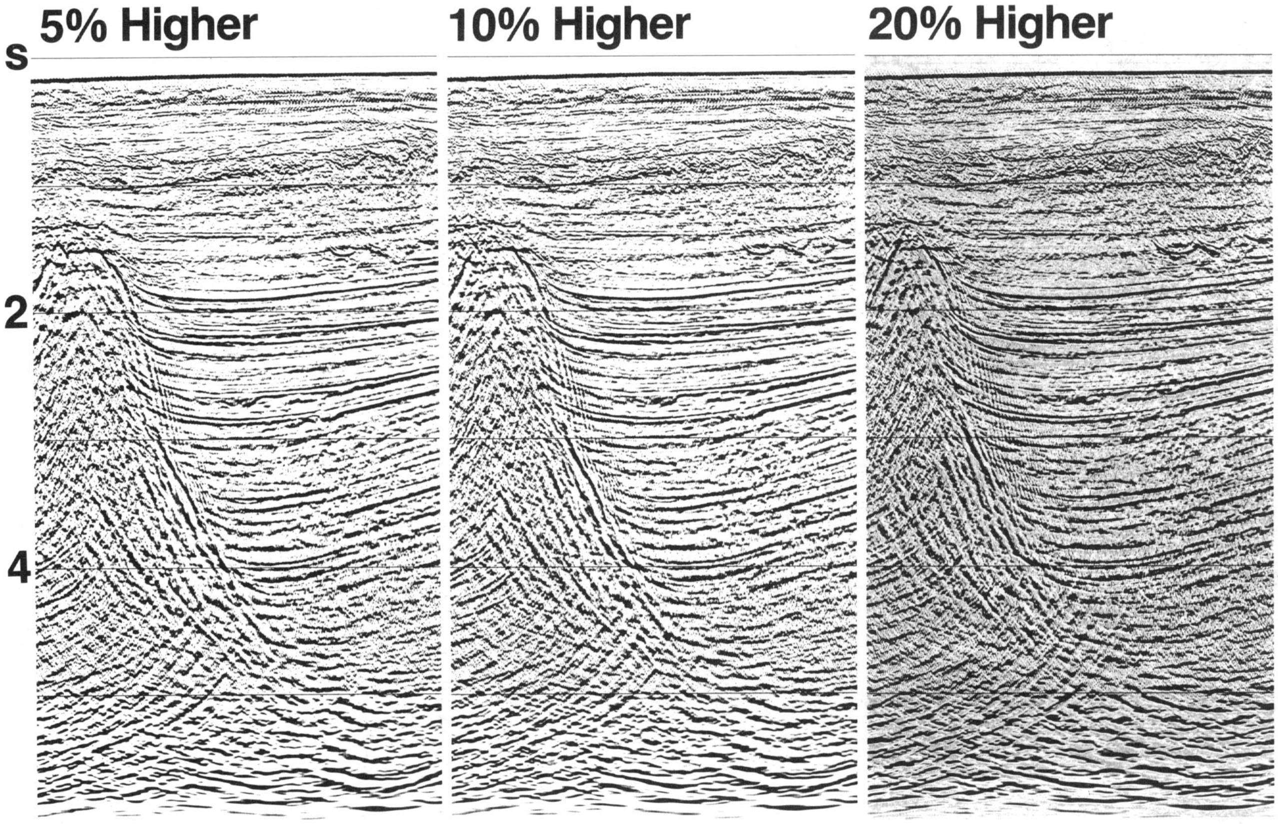


FIG. 4-72. Tests of velocity error in 15-degree finite-difference migration. The input stack and the desired migration are shown in Figure 4-57. See Figure 4-73 for a sketch of the migration results.

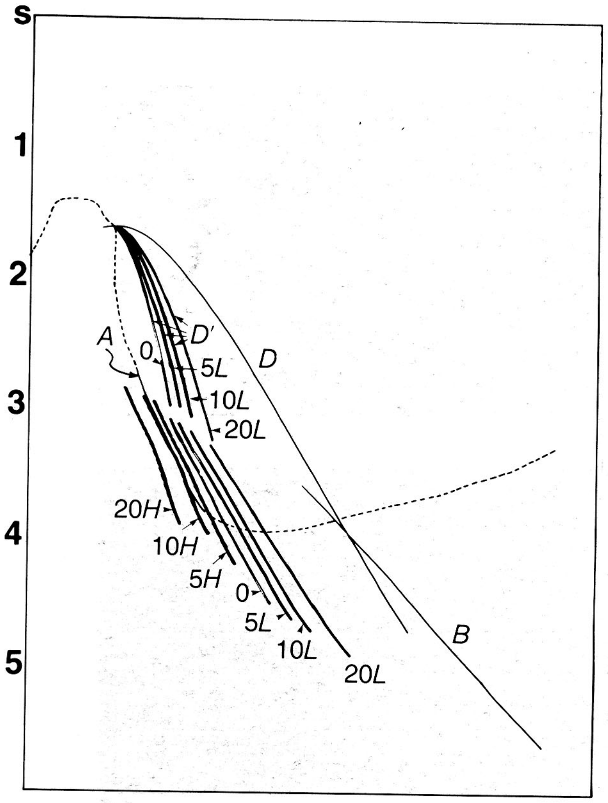


FIG. 4-73. The undermigration and overmigration effects from Figures 4-71 and 4-72. Here, B = dipping event before and A = dipping event after desired migration, D = diffraction before and D' = diffraction after finite-difference migrations, with L = percent lower and H = percent higher velocity.

undermigration is to modify migration velocities so that they increase with dip (Section C.6). For example, the best match between the desired migration and finite-difference results for the dipping event in Figure 4-73 occurred when 10 percent higher velocities were used in the finite-difference migration. In practice, it is common to use a percentage of the laterally smoothed stacking velocities in finite-difference migration without correcting them for a specific dip.

4.3.3 Frequency-Wavenumber Migration in Practice

Two different methods of migration are implemented in the frequency-wavenumber domain. The Stolt method is exact for a constant-velocity medium, while the phase-shift method is exact for a medium with vertical velocity variations. The Stolt method can be extended to the case of a medium with arbitrary velocity (arbitrary to the extent that time migration is valid).

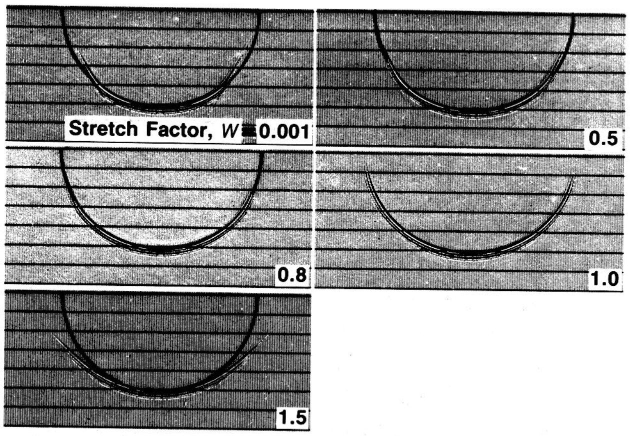


FIG. 4-74. By varying the stretch factor W, the impulse response of the exact 90-degree migration operator (semicircle) is modified. For comparison, the desired response has been superimposed on the Stolt migration impulse responses.

Stolt Stretch Factor, W

The practical aspects of the Stolt method are considered first. As discussed in Section 4.2.3, the generalized Stolt method of migration involves converting the time section to an approximately constant-velocity section, which then is migrated by the constant-velocity Stolt algorithm. This conversion is essentially stretching in the vertical (time) direction. Once the section is migrated in the stretched domain, it is converted back to the original time domain. The generalized Stolt method must be distinguished from the constant-velocity algorithm. The constant-velocity algorithm is accurate for dips up to 90 degrees for a constant-velocity medium. The generalized method approximately accounts for velocity variations by stretching the section.

Stretching is defined by the stretch factor W. In his original paper, Stolt (1978) discusses implementation of the W factor. Although W is a complicated function of velocity and stretch coordinate variables, it often is set to a scalar (Appendix C.4). The theoretical range of W is between 0 and 2. To get an insight into the effects of the stretch factor, refer to the impulse responses in Figure 4-74, where a single, isolated wavelet on a single trace is migrated using various stretch factors. Here, $W = 1$ corresponds to the exact constant-velocity Stolt algorithm. So, setting $W < 1$ compresses the impulse response inward along its steep flanks, while setting $W > 1$ opens it up. Thus, the value of W partially controls the aperture of the generalized Stolt algorithm. The farther W is from 1, the more limited the aperture becomes. A value of $W < 1$ implies undermigration at steeper dips, while a

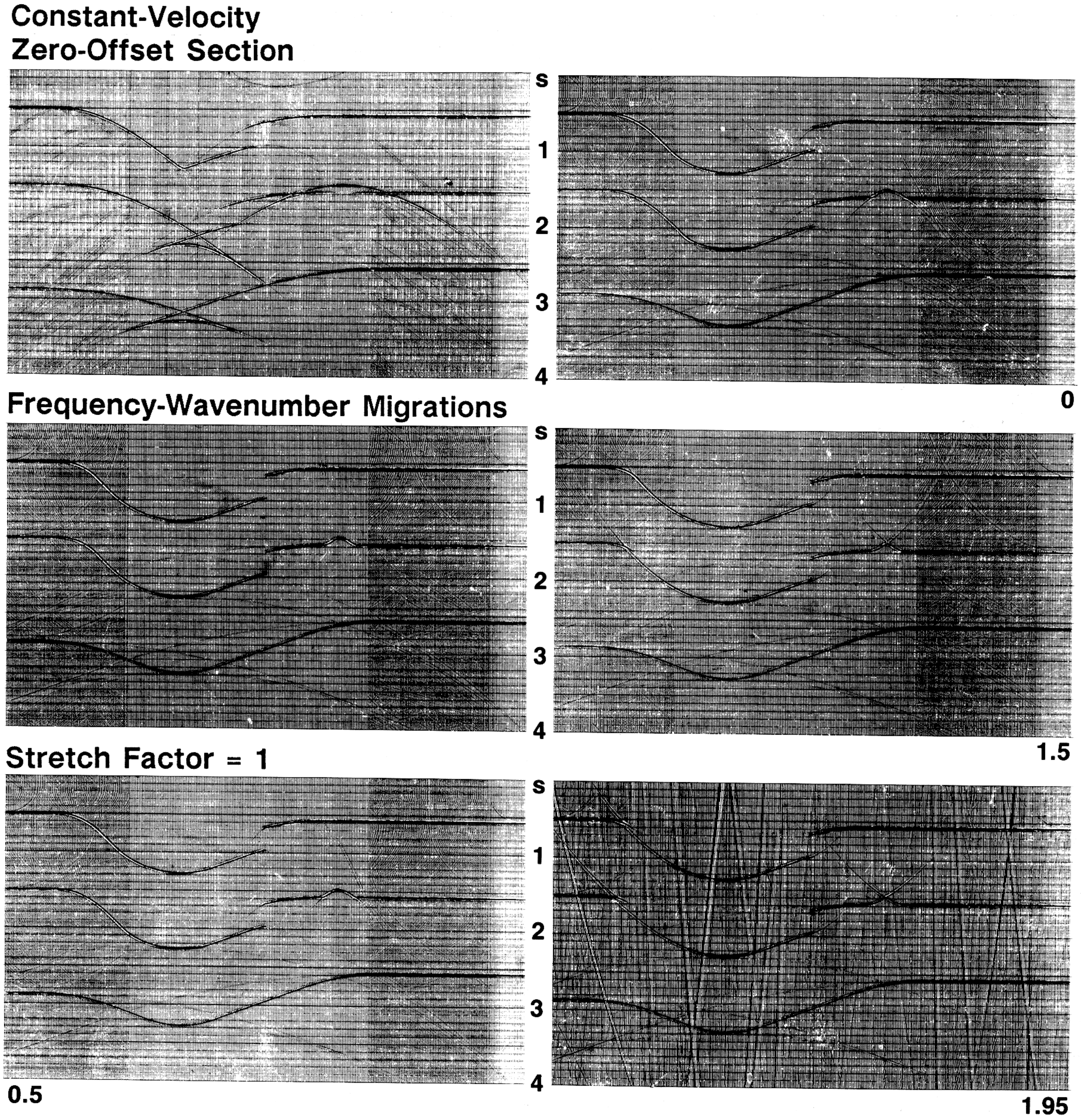


FIG. 4-75. Effect of the stretch factor on migration output. Here, $W < 1$ causes undermigration, $W > 1$ causes overmigration. (Modeling courtesy Union Oil Company.)

value of $W > 1$ implies overmigration at steeper dips, if the medium velocity is constant. Although not strictly implied by the impulse responses in Figure 4-74, when using a stretch factor different from 1, the Stolt algorithm tries to mimic a wavefront in a variable velocity medium (Stolt, 1978), while compromising on the ability to migrate steeper dips. Experience has proven that the Stolt migration with stretch produces acceptable results provided velocity variations are within limits of time migration.

Consider the subsurface model and the migration results in Figure 4-75. Stretch factor $W = 1$ produces the best migrated section because the zero-offset section was derived using a constant-velocity value. For $0 < W < 1$, the algorithm produces an undermigrated section, while for $1 > W > 2$, it produces an overmigrated section. These observations are in agreement with the impulse responses in Figure 4-74. The near-vertical streaks in the section with $W = 1.95$ are because of wraparound effects.

The generalized Stolt algorithm produces the best result when $W = 1$, provided we have constant velocity. Since this is never the case, we should examine the algorithm for a vertically varying velocity medium. Figure 4-76 shows the impulse responses for different values of W. Velocity varies linearly from $t = 0$ to $t = 4$ s between 2000 and 4000 m/s. For different W values, the portions of the

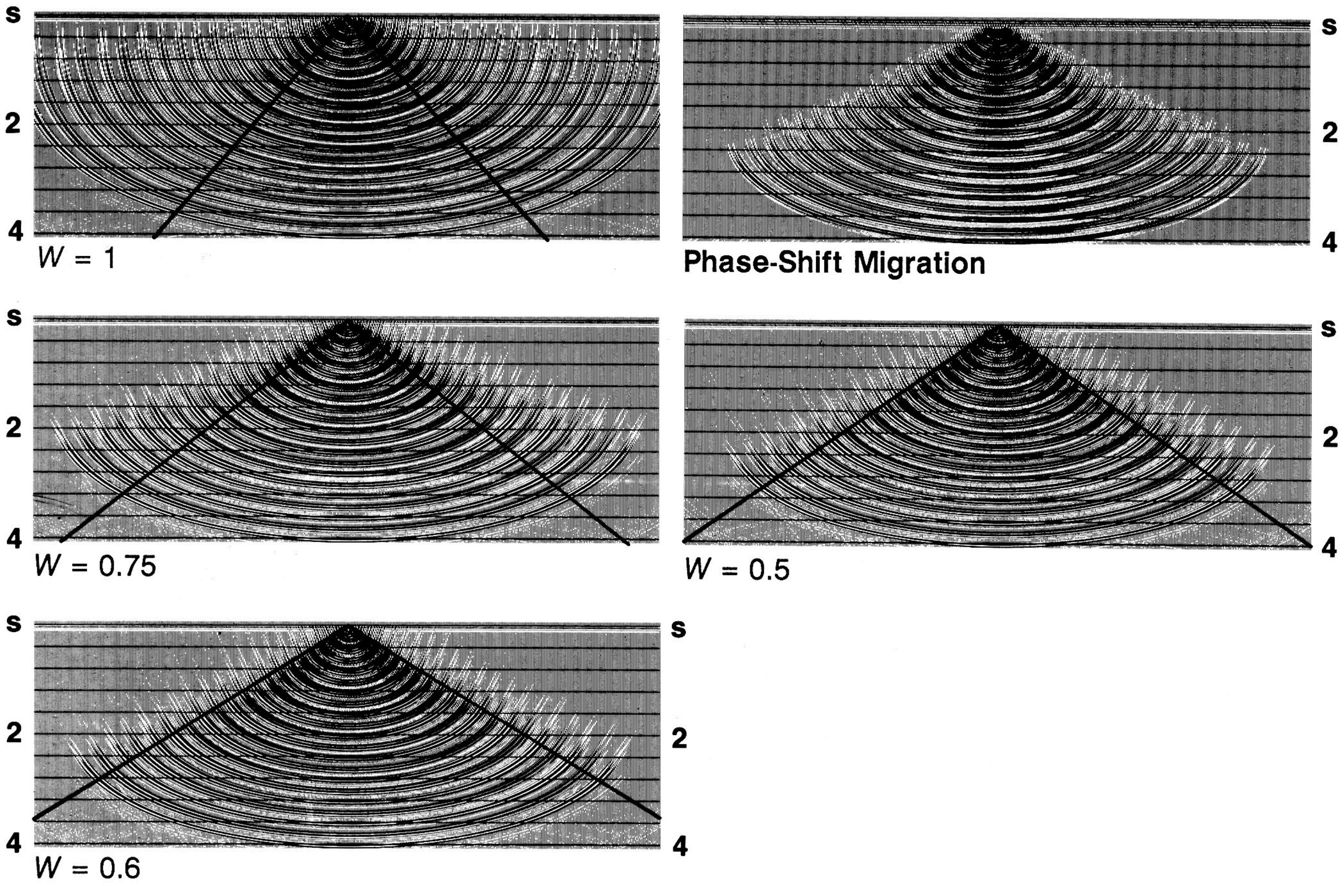


FIG. 4-76. Effect of the stretch factor when migrating with a vertically varying velocity function.

wavefronts that best match the (exact) phase-shift migration result are between the solid lines. For a vertically varying velocity medium, $W = 1$ no longer is the desired factor. In Figure 4-76, the Stolt method is accurate over the widest range of dip angles when $W = 0.6$. In general, the larger the velocity gradient, the farther the optimum W is from 1. Strictly speaking, the optimum value for W is even different at different times. In practice, wavefront plots, like those in Figure 4-76, can be generated using both the phase-shift and Stolt methods for a vertically varying regional velocity function. The W that yields the best fit at the largest angular aperture then is used to migrate the data with the Stolt method.

Figures 4-77 and 4-78 show migrations of the CMP stack in Figure 4-57a using different values of W. Migration velocities are varied only in the vertical direction. Figure 4-79 is a sketch of the migration results for the diffraction *(D)* and steeply dipping event *(B)* off the flank of the salt dome. The best match between the desired migration and the Stolt method (with stretch) is for $W = 0.5$.

Residual Migration

Although the constant-velocity Stolt algorithm may seem to be a purely academic exercise, it has a useful application. Consider the zero-offset section that consists of three point scatterers in a layered earth model (Figure 4-80a). The model parameters are provided in Figure 4-61. Finite-difference migration has difficulty collapsing these diffractions (Figure 4-80b). How can we help the 15-degree finite-difference method perform better in the presence of steep dips? The answer is to lower the apparent dip perceived by migration (Rothman et al., 1985). First we migrate the zero-offset section with the constant-velocity Stolt algorithm using the lowest value, 2000 m/s, in the $v(z)$ function. The result is shown in Figure 4-80d. Then we take this section and migrate it again (Figure 4-80e) using the appropriate *residual* velocity (Appendix C.5) and the 15-degree finite-difference method. When compared with the single-step finite-difference migration (Figure 4-80b), note the superior performance of the residual migration. Also compare this with the desired migration that was obtained by the phase-shift method (Figure 4-80c). Appendix C.5 contains a mathematical discussion on residual migration.

A field data example is shown in Figure 4-81 with a sketch of the migration results in Figure 4-82. The single-step finite-difference result has the typical undermigrated character (Figure 4-58). The 1500 m/s constant-velocity Stolt migration followed by the finite-difference migration seems to produce an output that is reasonably close to the

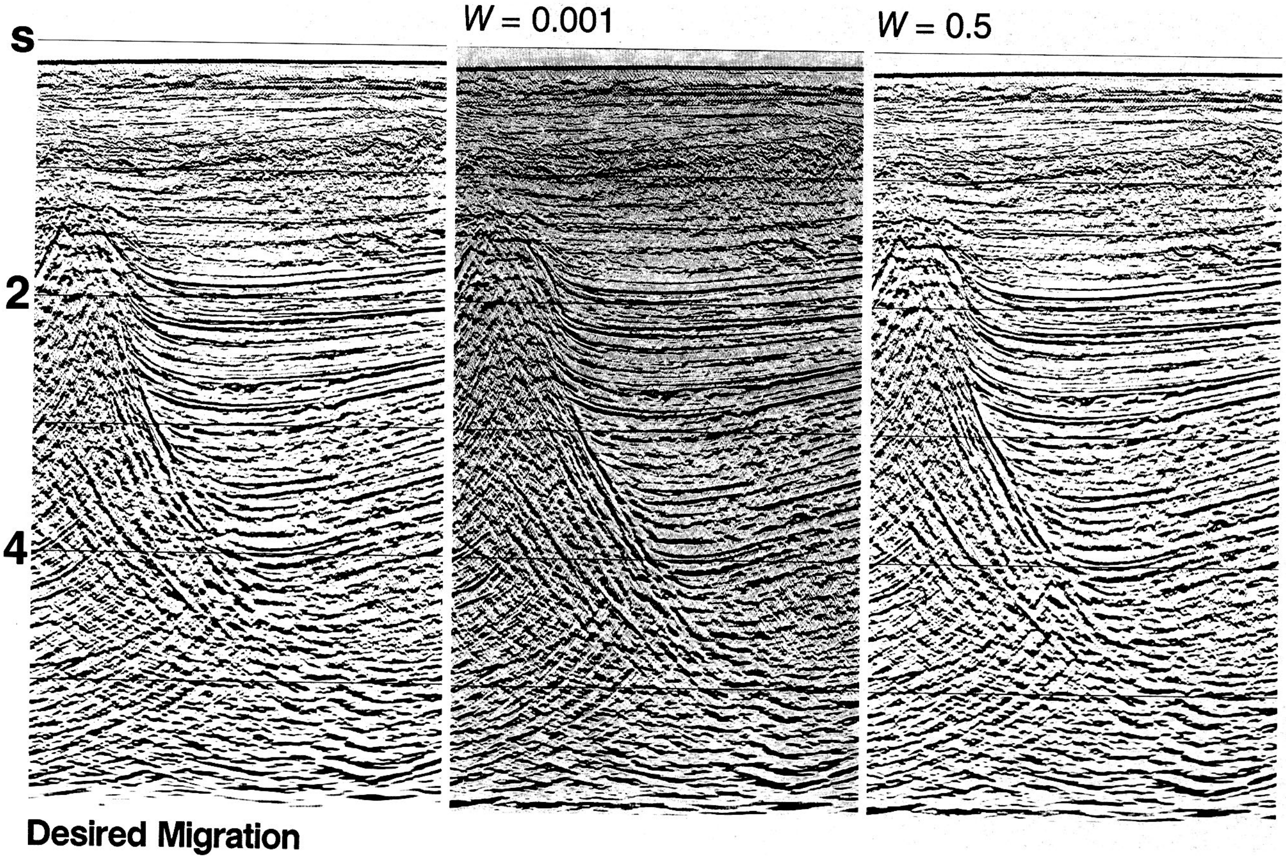


FIG. 4-77. Effect of the stretch factor on migration. The input stack is shown in Figure 4-57a.

desired migration. This desired result was obtained by the phase-shift method, which is more expensive than a combined application of the constant-velocity Stolt followed by the 15-degree finite-difference algorithms.

A limitation of residual migration is that an adequate migration is not always achieved. This is the case in Figure 4-81, since after residual migration, the dipping event still is slightly undermigrated (see the sketch in Figure 4-82). Undermigration occurs because the apparent dip perceived by the second-pass migration still may be too large to be handled accurately. The second-pass migration typically is performed using a dip-limited finite-difference algorithm. From equation (C.9b), note that the lower the velocity used in migration, the smaller the dip that is perceived by migration. Thus, if the residual velocity function given in equation (C.63) is not too different from the original velocity function (because of a large vertical gradient), then residual migration may not be adequate.

Maximum Dip to Migrate

We now consider Gazdag's phase-shift method, which is another type of *f-k* migration (Section 4.2.3). This method is suitable for practical use since it allows vertical variations in velocity and is accurate for up to 90-degree dips. Figure 4-83 shows the impulse response of the phase-shift algorithm. As with the Kirchhoff summation method, migration with the phase-shift method can be limited to smaller dips by truncating the semicircular wavefront. This dip filtering capability is useful in rejecting coherent noise from the stacked section while doing the migration. On the other hand, if the action of migration is constrained to small dip values, then the steeply dipping reflectors unintentionally can be filtered out. Edge effects also are pronounced when a very narrow range of dips is passed. Note the linear streaks on the impulse response with a dip limit of 2 ms/trace (Figure 4-83). On the field data example shown in Figure 4-84, severe dip filtering action of the 2 ms/trace maximum dip has caused smearing and eliminated virtually all of the signal contained in the section.

Depth Step Size

Figure 4-85 is the dipping-events model migrated with the phase-shift method using different depth step sizes. Since the phase-shift method is based on the dispersion relation [equation (C.28)] of the exact one-way wave equation, we

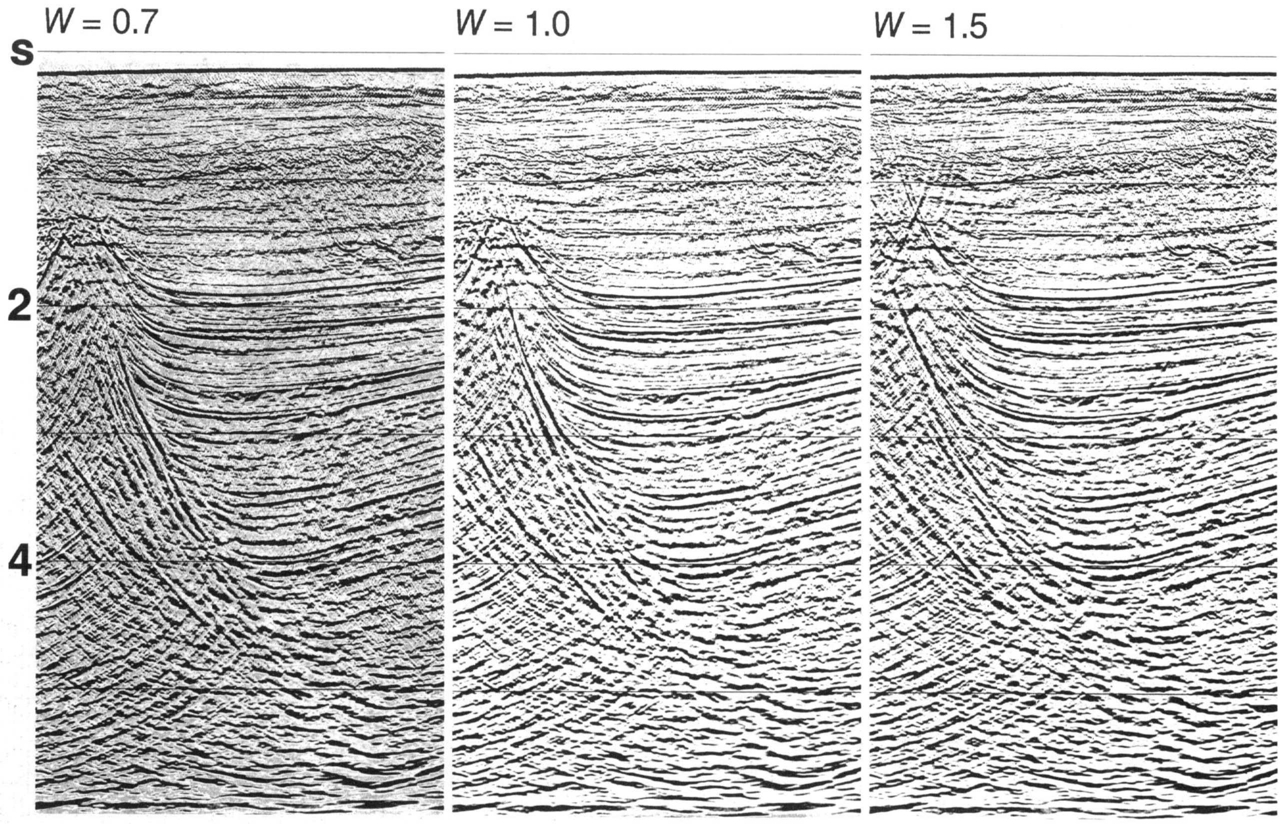


FIG. 4-78. Effect of the stretch factor on migration. Note overmigration with large values of *W*. The input stack is shown in Figure 4-57a. See Figure 4-77 for the desired migration.

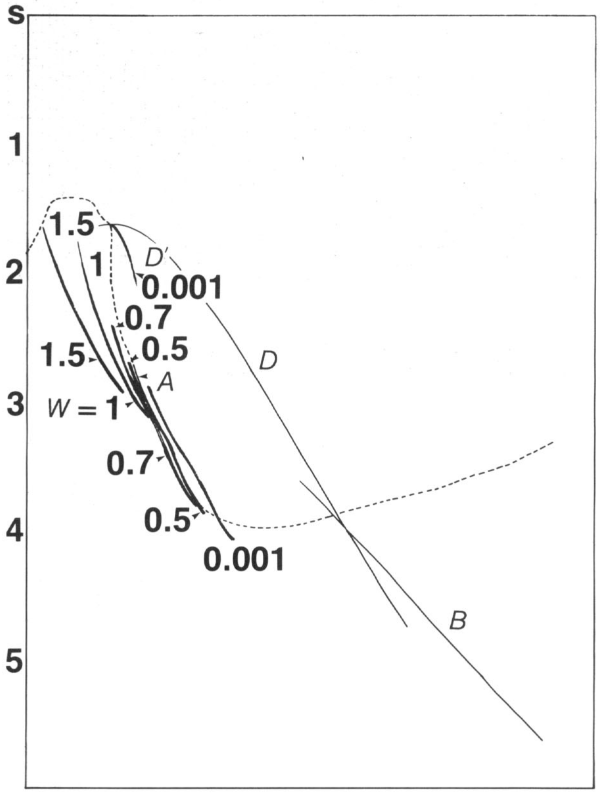


FIG. 4-79. The combined results of the migrations from Figures 4-77 and 4-78. Here, *B* = dipping event before and *A* = dipping event after desired migration, *D* = diffraction before and *D′* = diffraction after Stolt migration with stretch. Numbers represent different stretch factors *W*.

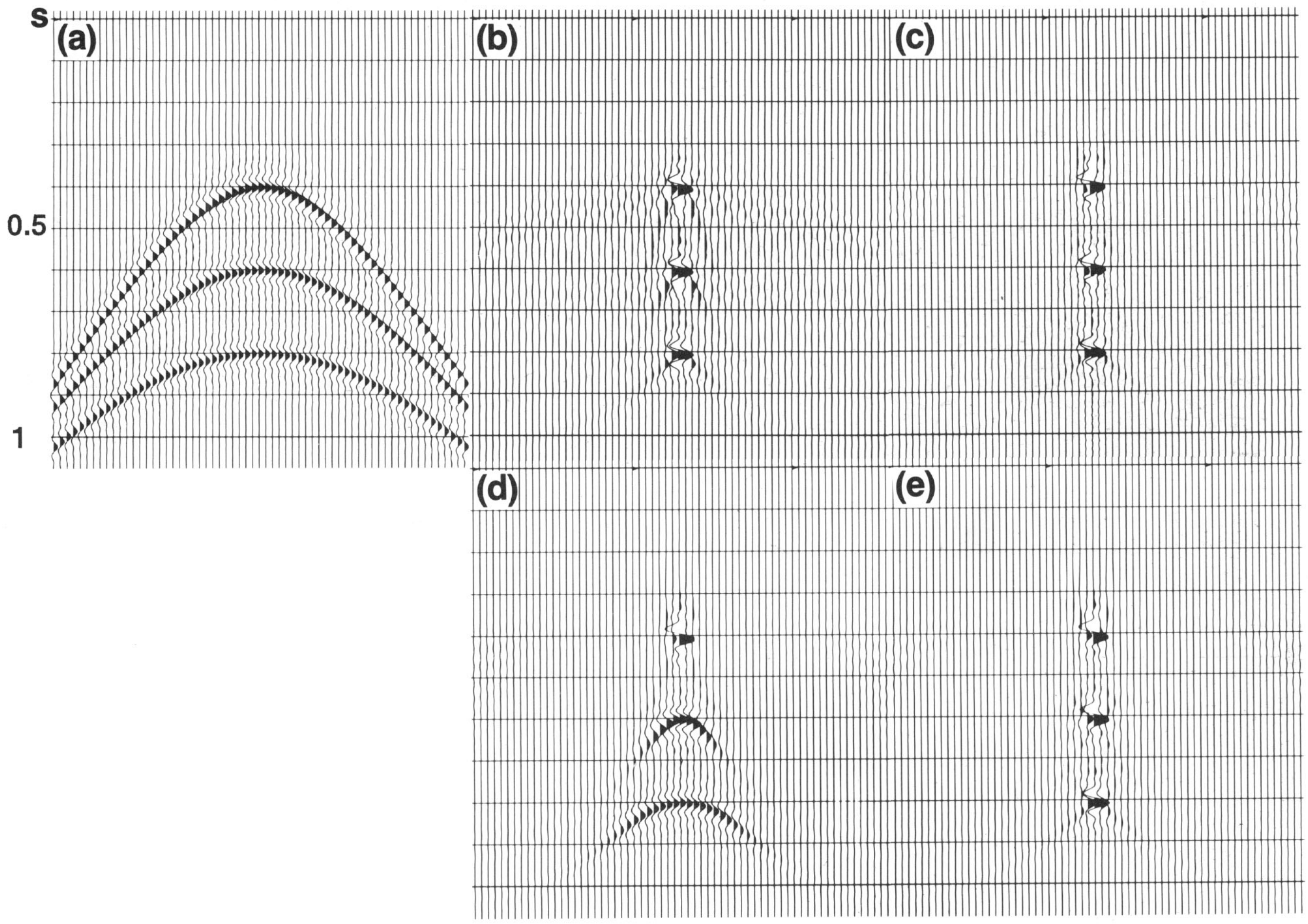


FIG. 4-80. Principle of residual migration. (a) Zero-offset section, $v(z)$ model, (b) 15-degree finite-difference migration, (c) desired migration using phase-shift method, (d) constant-velocity Stolt migration, where v = 2000 m/s, (e) panel (d) followed by 15-degree finite difference migration.

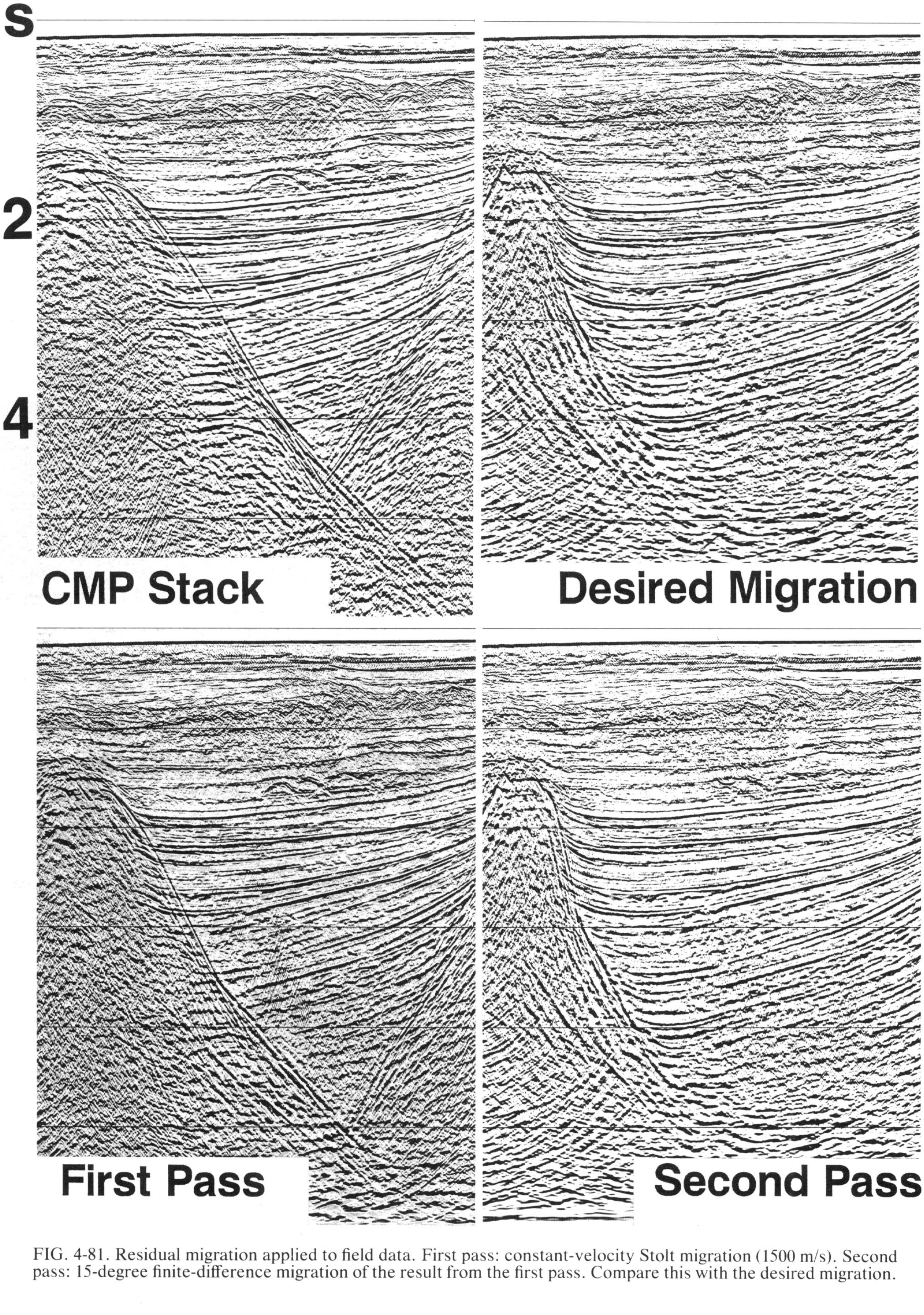


FIG. 4-81. Residual migration applied to field data. First pass: constant-velocity Stolt migration (1500 m/s). Second pass: 15-degree finite-difference migration of the result from the first pass. Compare this with the desired migration.

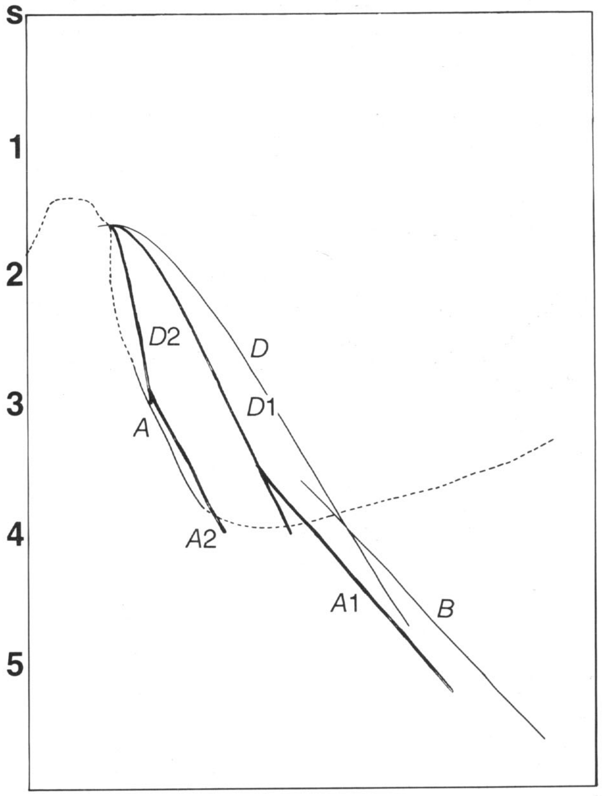


FIG. 4-82. The combined migration results from Figure 4-81. Here, B = dipping event before and A = dipping event after desired migration; D = diffraction; $A1$, $D1$ = after first pass (Stolt); $A2$, $D2$ = after second pass (15-degree finite difference).

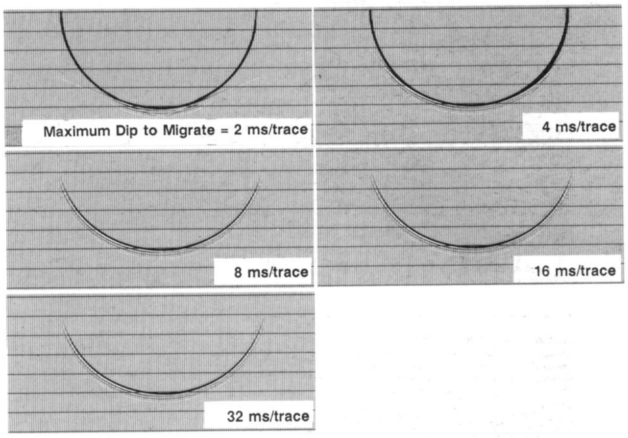


FIG. 4-83. The impulse response of the *f-k* migration operator becomes truncated semicircular when a maximum dip limit is imposed. For comparison, the desired response has been superimposed on the *f-k* responses.

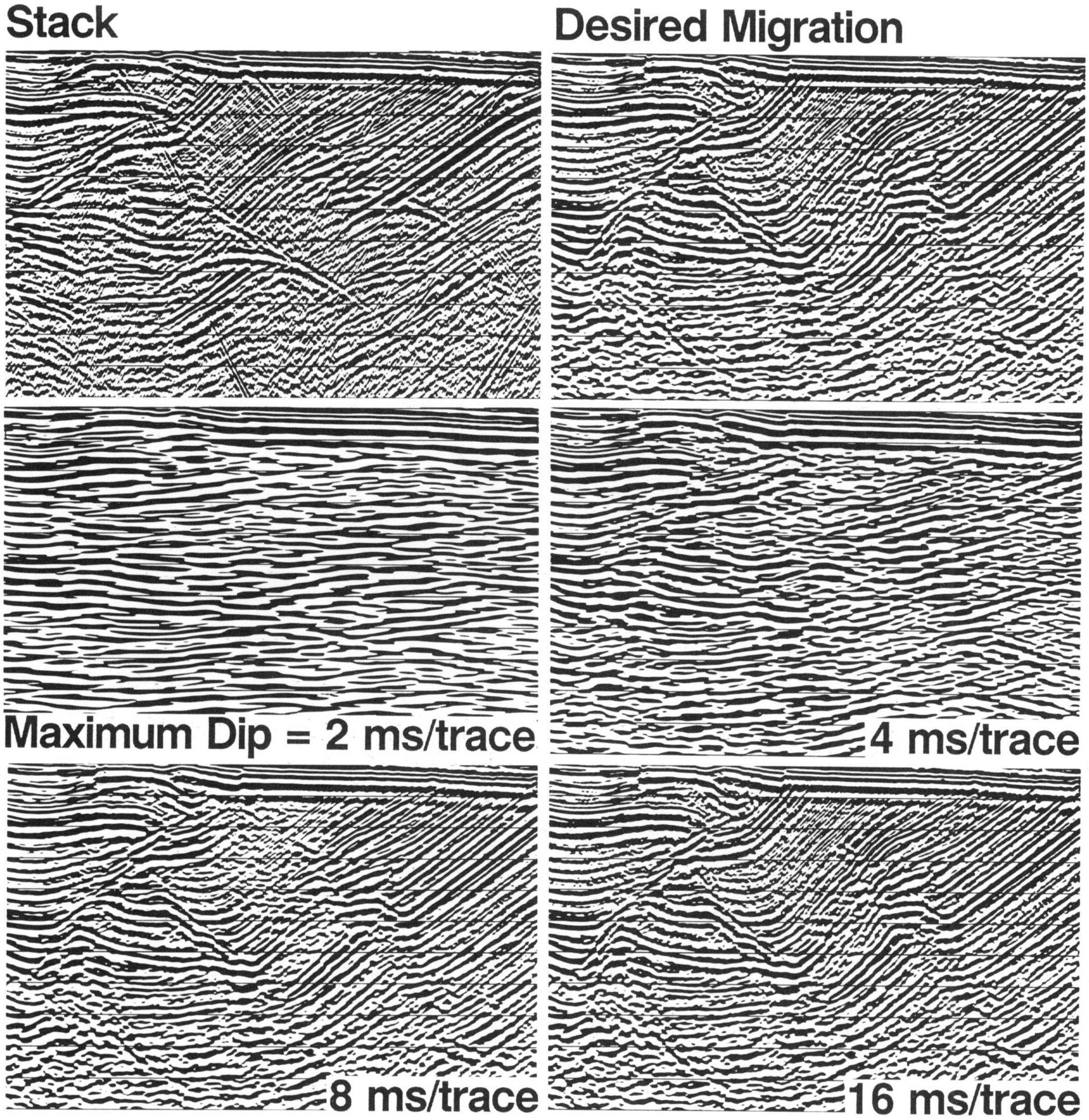


FIG. 4-84. Tests for maximum dip to migrate in phase-shift migration. An excessively small dip limit can be hazardous in migration.

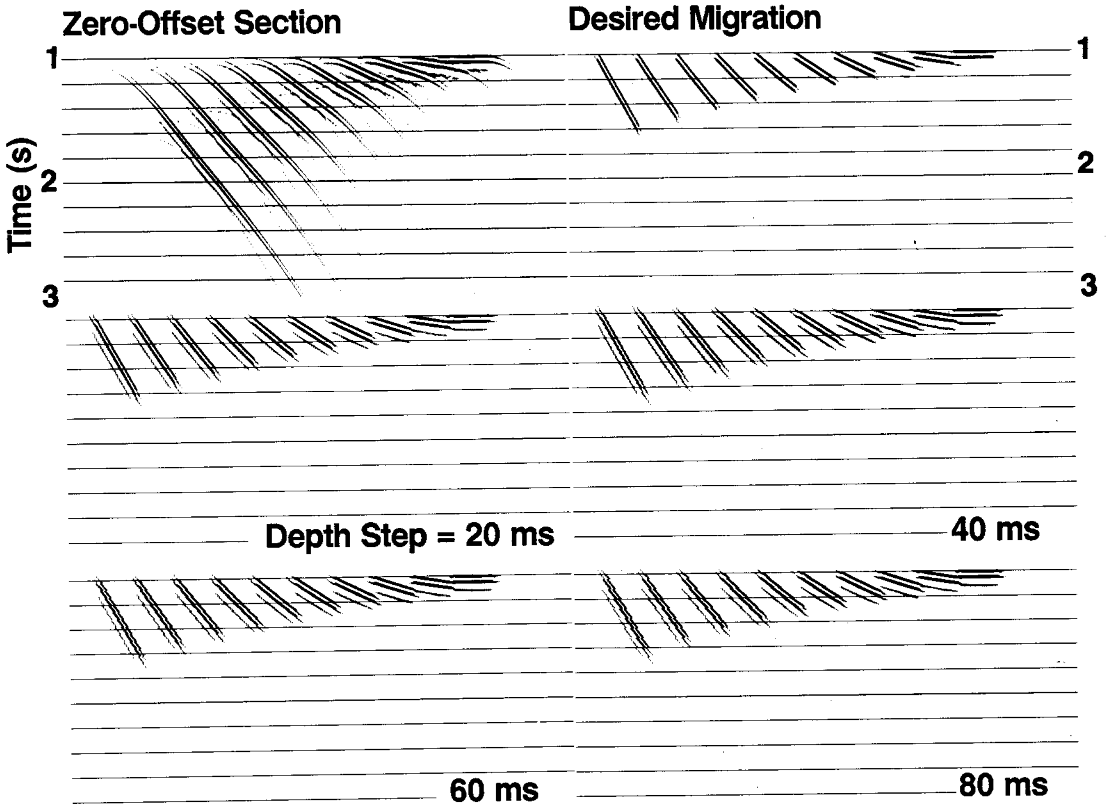


FIG. 4-85. The effect of a 20- to 80-ms depth step size on the phase-shift migration output of dipping reflectors.

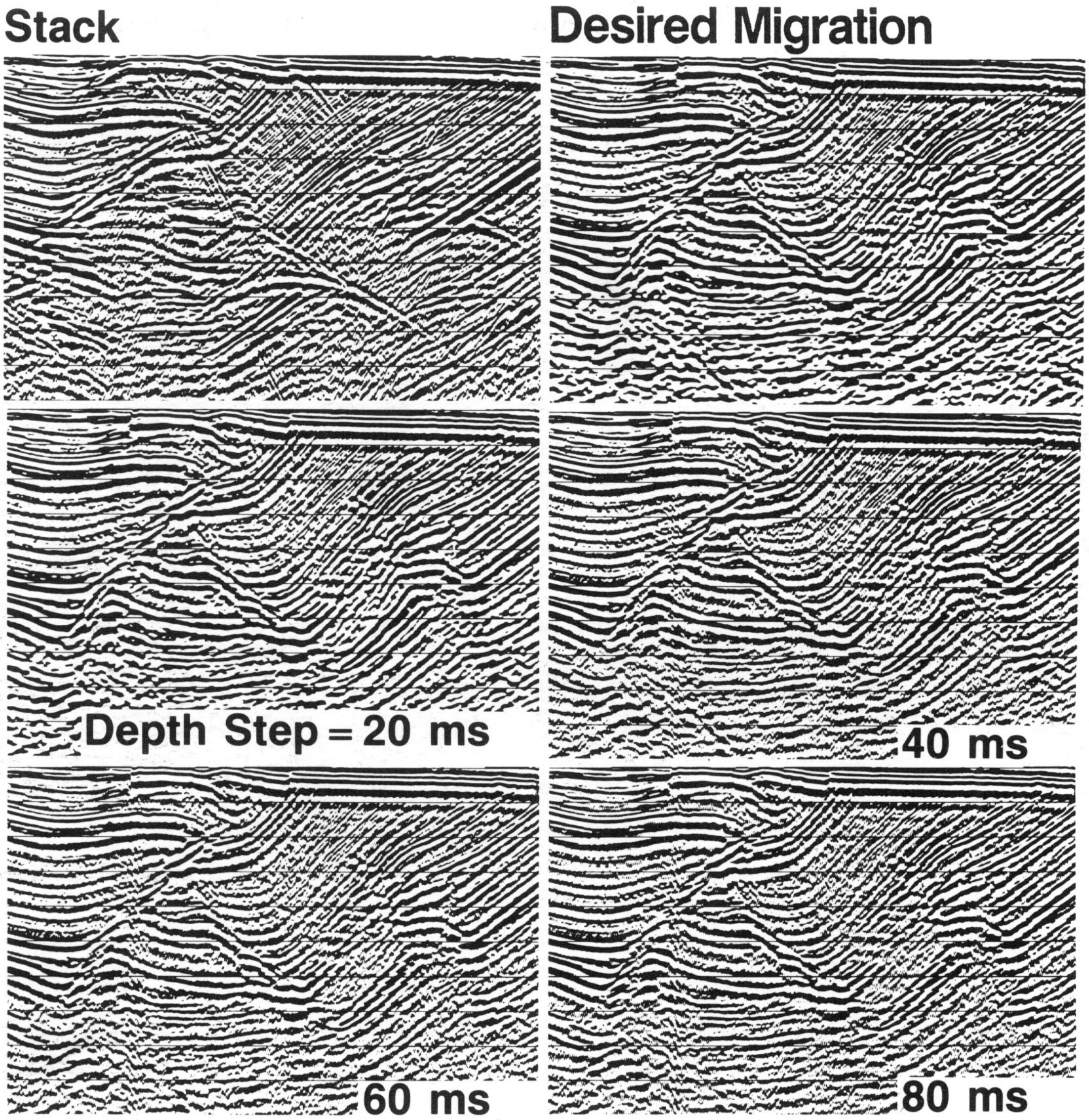


FIG. 4-86. The effect of a 20- to 80-ms depth step size on the phase-shift migration output of field data.

do not expect undermigration. However, we see discontinuities along the reflectors at intervals equal to the depth step size, which is similar to the finite-difference results. As with the finite-difference algorithms, the problem occurs along the steeper dips first; therefore, the steep dips require smaller depth steps (Figure 4-85). Because of the band-limited nature of seismic data, very small depth steps are not needed. Note that migration with a 20-ms depth step produces a section without spurious kinks along the reflectors; this is comparable to the desired migration results. Although not shown, the smaller depth steps yield results that are virtually identical to the 20-ms result.

Depth step size tests on field data are shown in Figure 4-86. Unlike the finite-difference results, all the phase-shift migrations with different depth step sizes produce equally adequate results. The only problem with large depth steps is the kinks along the steep dips. In principle, as long as there is no aliasing in the z-direction, the depth step kinks can be eliminated by a local interpolation process. In practice, the depth step size used in migration with the phase-shift method typically is taken between the half-and full-dominant period of the wave field; e.g., 20 to 40 ms, depending on steepness of the dips present in the section.

Velocity Errors

The phase-shift method responds to velocity errors much as the Kirchhoff summation method. All the synthetic model examples for the Kirchhoff migration can be considered as tests for the phase-shift method (Figures 4-49 through 4-52, inclusively). Figures 4-87 and 4-88 are

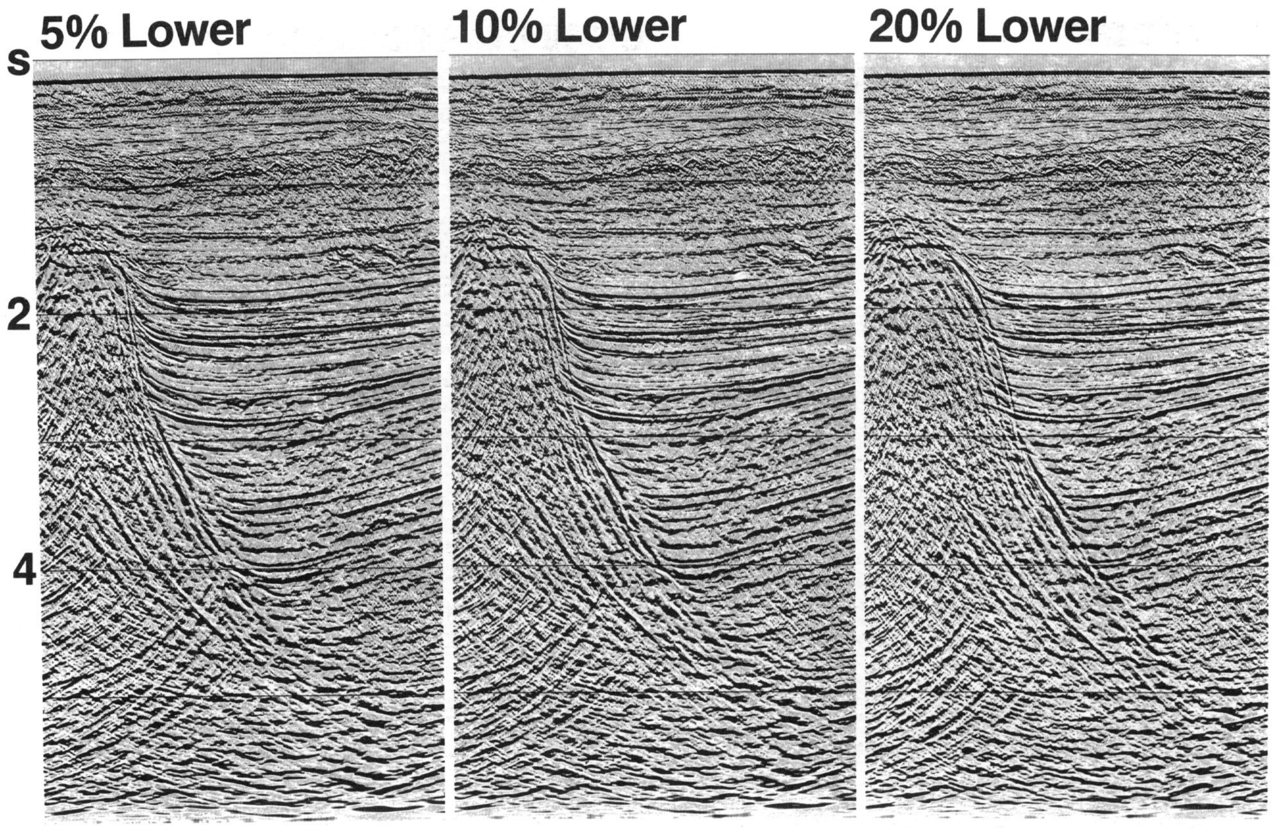


FIG. 4-87. Tests for velocity errors in phase-shift migration; undermigration caused by the use of too low velocities. The input stack and the desired migration are shown in Figure 4-57.

field data examples using too low and too high velocities, respectively. Figure 4-89 shows a sketch of the combined results of these migrations. Clearly, velocities that are too low cause undermigration of dipping events and incomplete collapse of diffractions. On the other hand, velocities that are too high cause overmigration of dipping events. Comparison of the compiled results in Figure 4-89 with those in Figure 4-73 indicates that a steep-dip algorithm, such as the *f-k* method, is more sensitive to velocity errors than a small-dip algorithm, such as the finite-difference method based on the parabolic equation. For example, note the larger departure of the dipping event from its correct position *A* when the velocity is 20 percent higher than the proper migration velocities in the *f-k* result (Figure 4-89) as compared to the finite-difference result (Figure 4-73).

4.3.4 Frequency-Space Migration

In the preceding sections, practical aspects of the finite-difference, Kirchhoff summation, and frequency-wavenumber migration methods were examined. All three have some advantages and some disadvantages. We observed that the Kirchhoff summation method can handle dips up to 90 degrees, but is limited in handling lateral velocity variations. Finite-difference methods based on the parabolic approximation to the scalar wave equation can handle dips reliably only up to 35 degrees. On the other hand, these methods have no problem with lateral velocity variations. The frequency-domain techniques have no dip limitation, but are restricted in handling velocity variations.

In this section, we discuss another finite-difference technique with the advantages of handling steep dips and all types of velocity variations. A mathematical discussion is provided in Appendix C.3. The method is implemented in the hybrid domain of frequency-space (ω,x); thus, it is commonly referred to as the omega-x or f-x method. The 15-degree equation is derived from the Taylor expansion of the dispersion relation [equation (C.29)]. The omega-x method is based on the continuous fractions expansion [equation (C.42)], which allows wider angle approximations. Kjartansson (1979) implemented the so-called 45-degree equation for both modeling and migration. This equation can be upgraded to be accurate for dips up to 65 degrees by simply tuning some coefficients (Appendix C.3). Higher-order operators can be obtained by successive application of a number of operators like the 45-degree operator (Ma, 1981) with a different set of coefficients (Lee and Suh, 1985) (Figure 4-90).

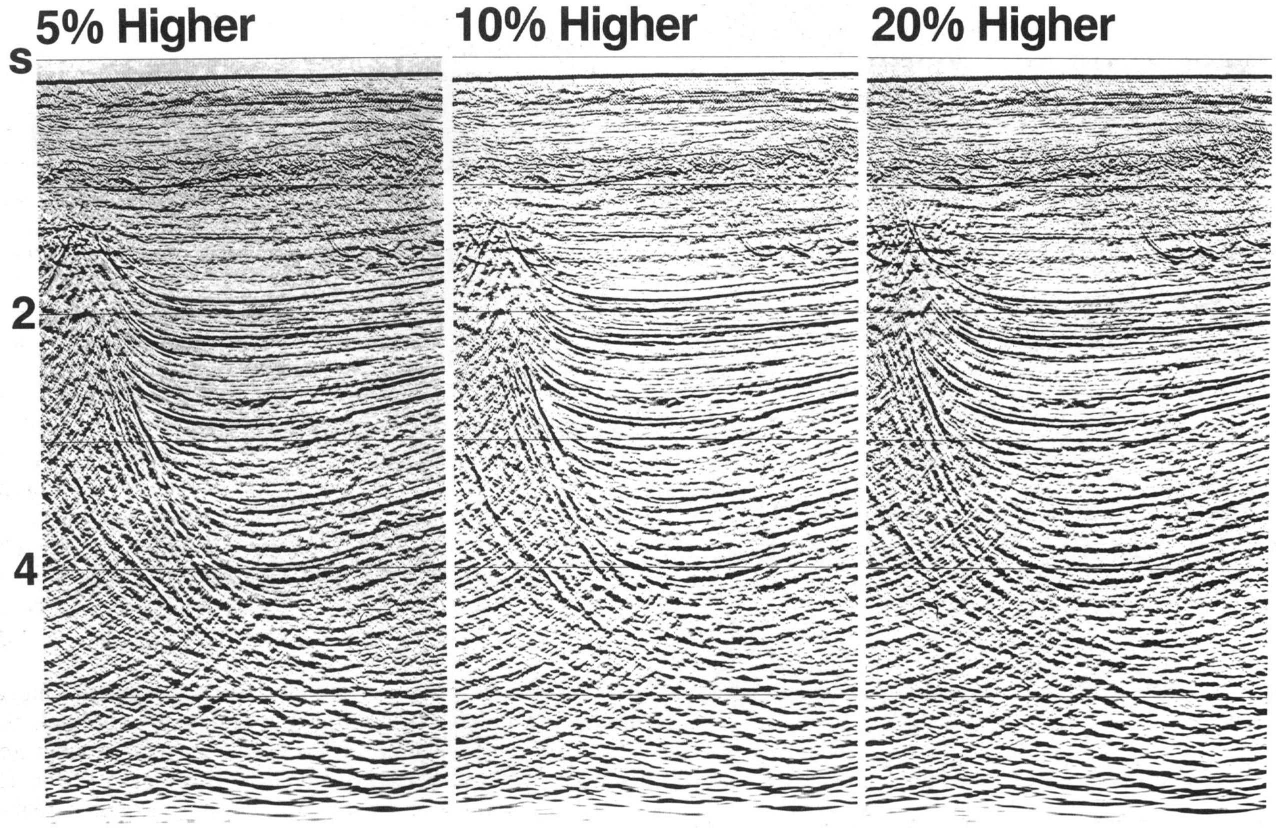
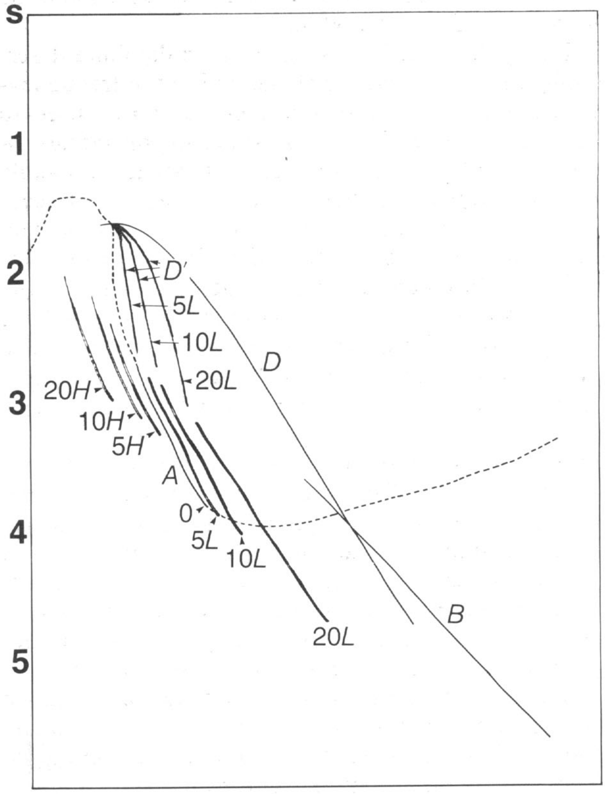


FIG. 4-88. Tests for velocity errors in phase shift migration; overmigration caused by the use of too high velocities. The input stack and the desired migration are shown in Figure 4-57.



◄FIG. 4-89. The combined results of migrations from Figures 4-87 and 4-88. Here, *B* = dipping event before and *A* = dipping event after desired migration; *D* = diffraction, and *L* = percent lower and *H* = percent higher velocity.

As discussed in Section 4.3.2, the most important parameter in finite-difference methods is the depth step size used in downward continuing the wave field recorded at the surface. We want to choose an optimum depth step size that is large (for economic reasons), yet yields a tolerable error in positioning the events after migration.

A 15-degree equation behaves somewhat differently than a 45-degree equation insofar as the depth step size. Reexamine the results of the tests done on the dipping events and diffraction models in Section 4.3.2 using the 15-degree finite-difference algorithm. Repeating the same tests using the omega-*x* 65-degree algorithm, we obtain the results in Figures 4-91 through 4-96.

At large depth steps, the algorithm undermigrates the diffraction hyperbola (Figure 4-91) as in the 15-degree case (Figure 4-59). At small depth steps (4 to 16 ms), the algorithm causes overmigration of the diffraction hyperbola (Figure 4-92), unlike the case in the 15-degree equation (Figure 4-60).

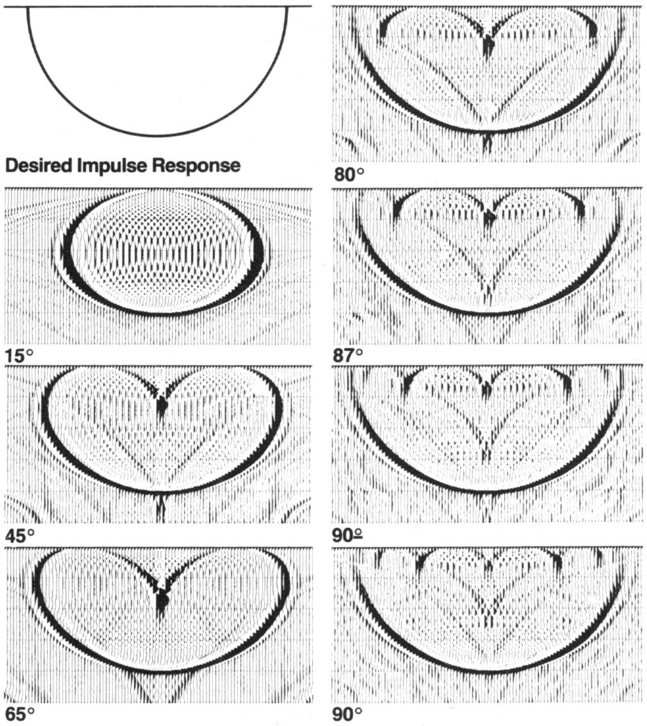


FIG. 4-90. Impulse responses of the frequency-space migration operators with various degrees of approximations to the one-way scalar wave equation. (See Appendix C.3 for the theoretical basis of these responses.)

At large depth steps, steeper dips are undermigrated (Figure 4-93), as in the 15-degree case (Figure 4-63). For comparison, the event with the steepest dip *AB* was superimposed on the results of migrations using different depth steps. At small depth steps (4 to 16 ms), the steeper dips seem to be overmigrated (Figure 4-94), unlike the case in the 15-degree equation (Figure 4-64). Also note the familiar dispersive noise (Figure 4-91 and 4-93) seen in all finite-difference migration results (regardless of the dip limitation). Comparison with the 15-degree counterparts (Figures 4-59 and 4-63) suggests that performance of the 45-degree equation in terms of the dispersive noise and undermigration is not much better than the 15-degree equation. This is particularly true for steep dips migrated with large depth steps. Theoretically, the 45-degree differential equation is more accurate than the 15-degree differential equation. However, once discretized, the difference in performance between these two equations can be less [Diet and Lailly (1984)]. A good finite-difference migration program uses differencing schemes that maintain the dip accuracy implied by the corresponding differential equation.

It is apparent that a higher order approximation, such as the 65-degree equation, provides a smaller range of choice for optimum depth size as compared to the 15-degree equation. In particular, note that the optimum depth step size is 20 ms for the models shown here; any departure from this value can cause either undermigration or overmigration. Nevertheless, there is no doubt that the 65-degree algorithm can migrate steeper dips and collapse diffractions more accurately than the 15-degree equation (compare Figure 4-59 with Figure 4-91, and Figure 4-63 with Figure 4-93, 20-ms depth step).

A field data example is shown in Figure 4-95. Here, the data were migrated using three different approximations in frequency-space domain: 15, 45, and 65 degrees. As we go to the higher-degree approximation, the collapse of the diffraction becomes complete and the steeply dipping event is migrated more accurately. Compare these results with the desired migration in Figure 4-57b. Also note the similar results obtained from the 15-degree *t-x* algorithm (Figure 4-57c) and the 15-degree omega-*x* algorithm (Figure 4-95a).

Now consider the test for velocity error using the omega-*x* algorithm. When velocities lower than medium velocity are used, the diffraction hyperbola gets undermigrated (Figure 4-96), but not as much as in the case of the 15-degree equation (Figure 4-67). When velocities higher than medium velocity are used, the diffraction hyperbola gets overmigrated (Figure 4-97), more so than with the 15-degree equation (Figure 4-68). Again we see that steep-dip algorithms are more sensitive to velocity errors. Similar conclusions are made for the dipping events model from Figures 4-98 and 4-99, which should be compared with Figures 4-69 and 4-70.

We briefly discussed a 65-degree steep-dip finite-difference migration technique implemented in the frequency-space domain. In practice, this method can be used to migrate dips up to 80 degrees. Another important advantage of the method is in its exceptional ability to handle velocity variations, whether vertical or lateral. Its accuracy for the lateral velocity problem results from the fact that the time shift associated with the thin-lens term [equation (4.16)] can be implemented exactly in the frequency domain. For these reasons, the algorithm is most appropriate for depth migration required to image targets beneath complex structures (Section 5.2), and 3-D imaging of the subsurface (Section 6.5 and Appendix C.8).

The omega-*x* migration also has the important operational advantage that each frequency can be processed separately. This property can significantly reduce computer memory requirements and, thus, decrease input/output operations for large data sets such as 3-D surveys. Also, in omega-*x* migration, some accuracy features can be conveniently implemented. For example, wave extrapolation can be limited to a specified signal bandwidth. Each frequency component can be downward continued using an optimum depth step size that yields a minimum acceptable phase error, leading to a minimum amount of dispersive noise on the migrated section.

(Text continued on page 320)

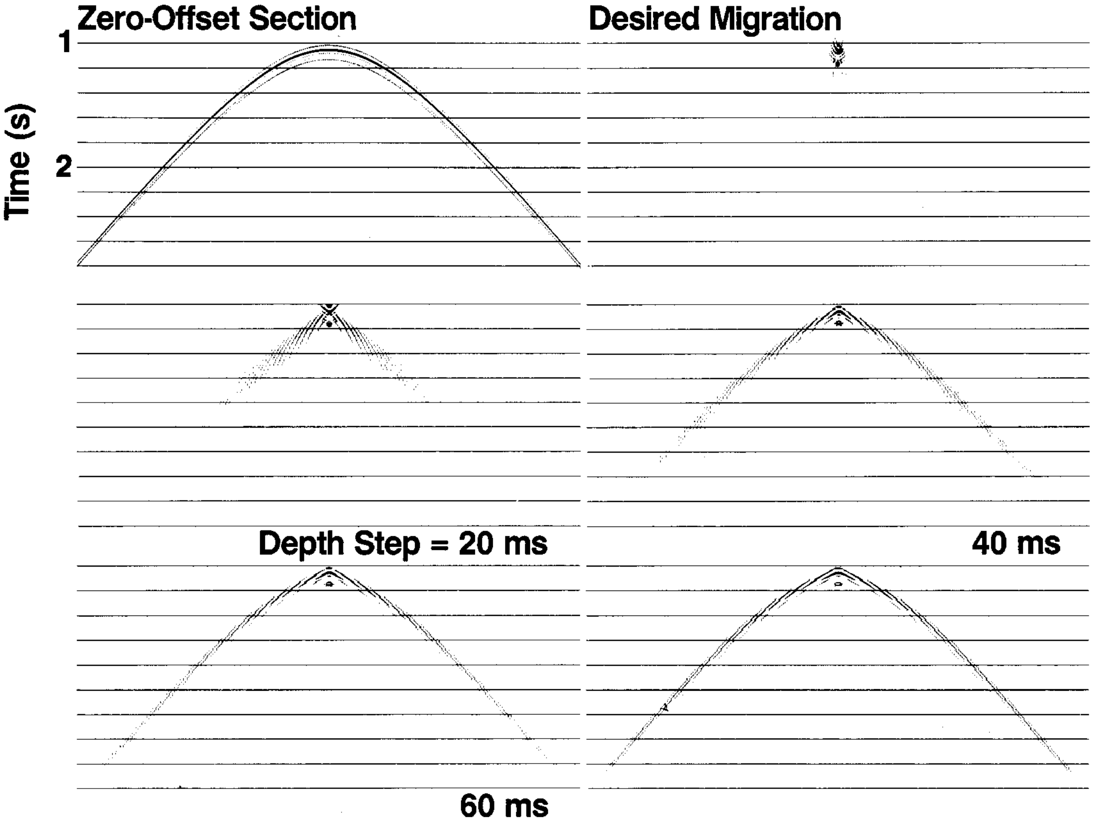


FIG. 4-91. Effect of 20- to 80-ms depth step size on the ability of the 65-degree equation in collapsing the diffraction hyperbola. Note the undermigration at large depth steps.

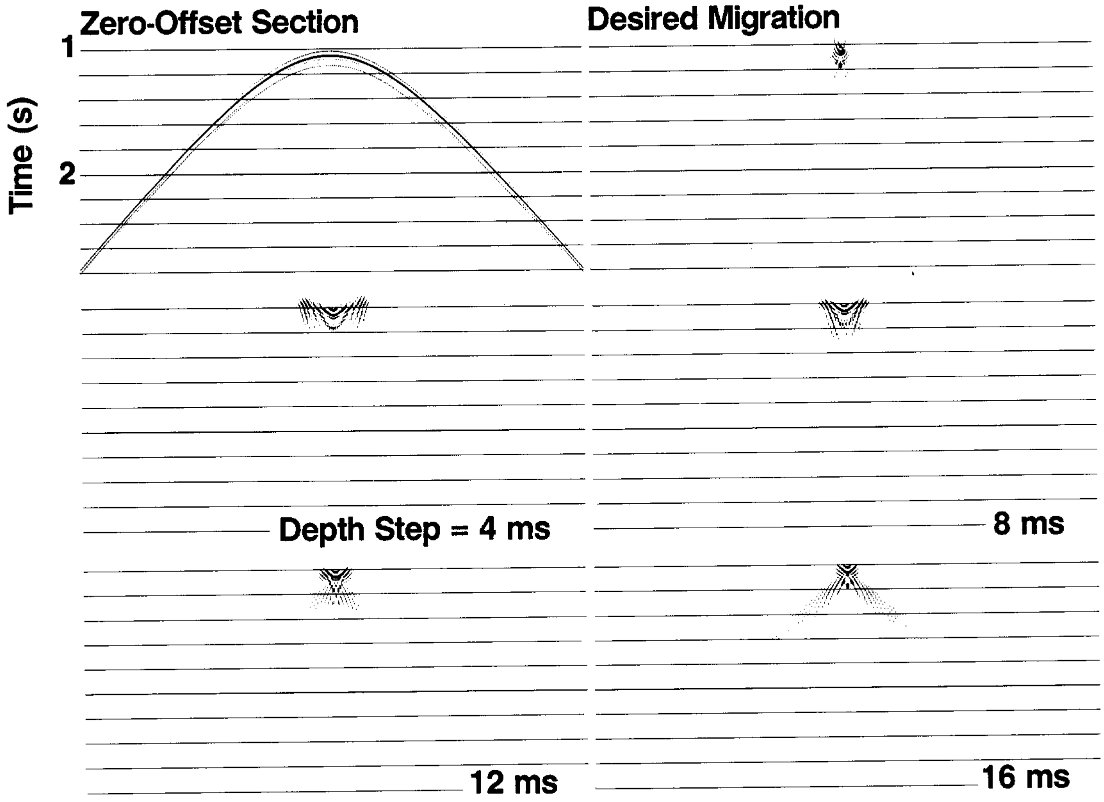


FIG. 4-92. Effect of 4- to 16-ms depth step size on the ability of the 65-degree equation in collapsing the diffraction hyperbola.

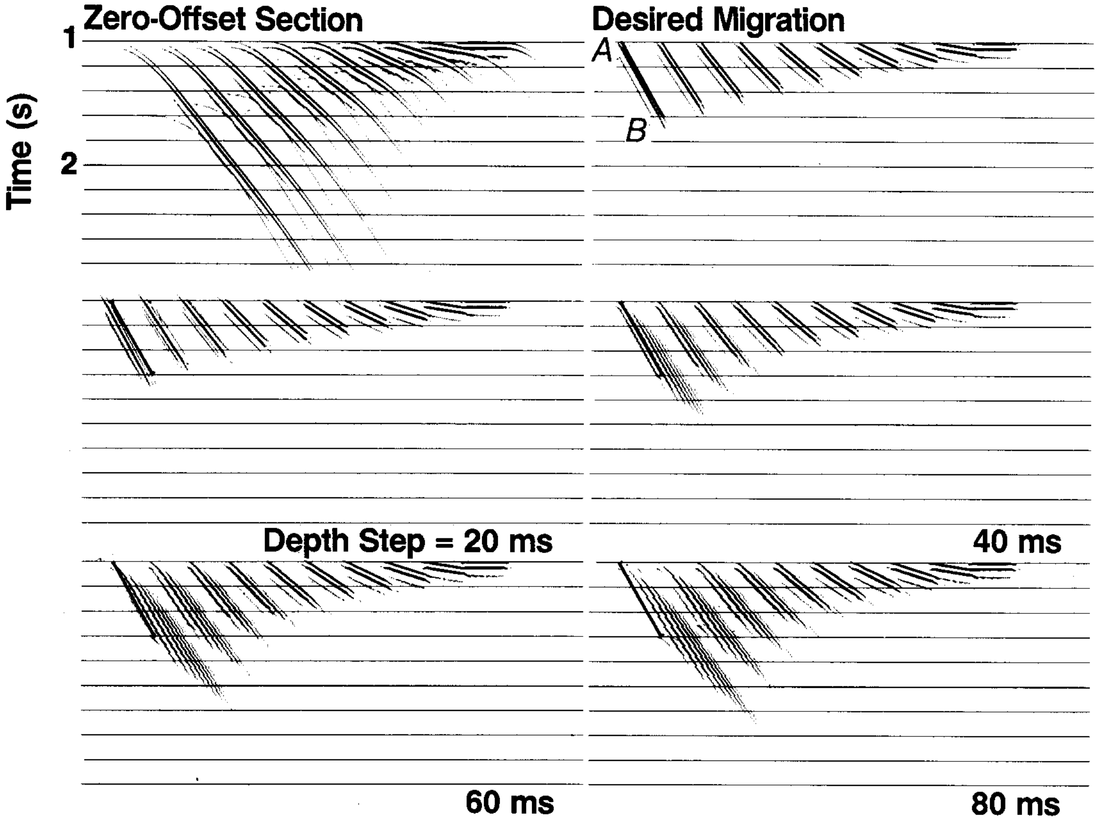


FIG. 4-93. Effect of 20- to 80-ms depth step size on the ability of the 65-degree equation in migrating dipping events. Note the undermigration at large depth steps.

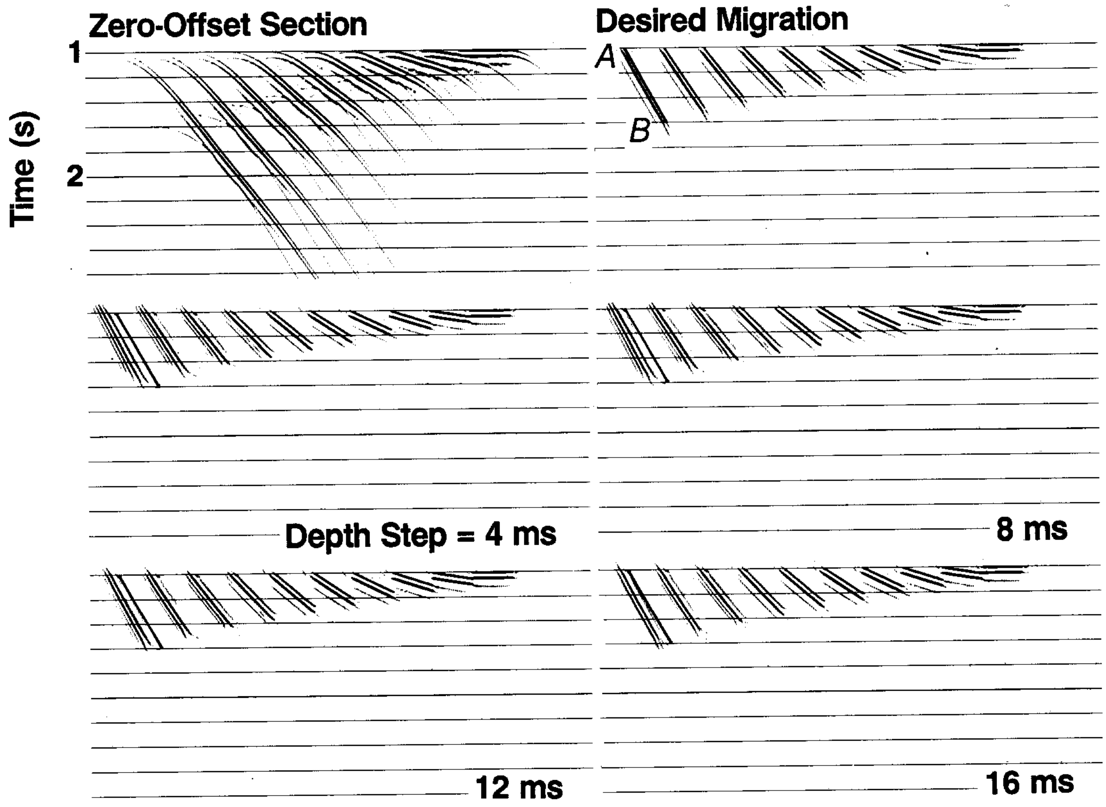


FIG. 4-94. Effect of 4- to 16-ms depth step size on the ability of the 65-degree equation in migrating dipping events. Note the overmigration at small depth steps.

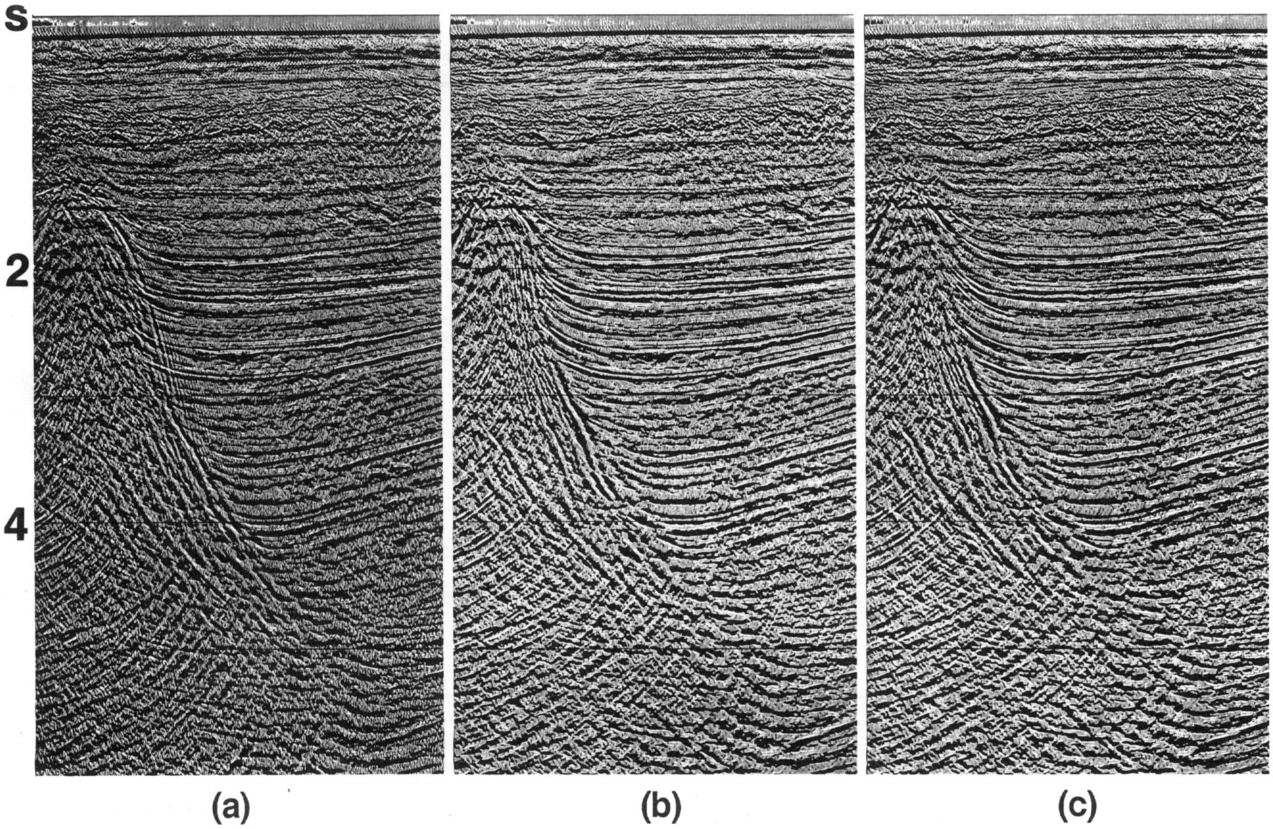


FIG. 4-95. Frequency-space migration of the CMP stack in Figure 4-57a. (The desired migration is shown in Figure 4-57b.) Note the increasingly better imaging obtained using higher angle approximations: (a) 15 degrees, (b) 45 degrees, (c) 65 degrees.

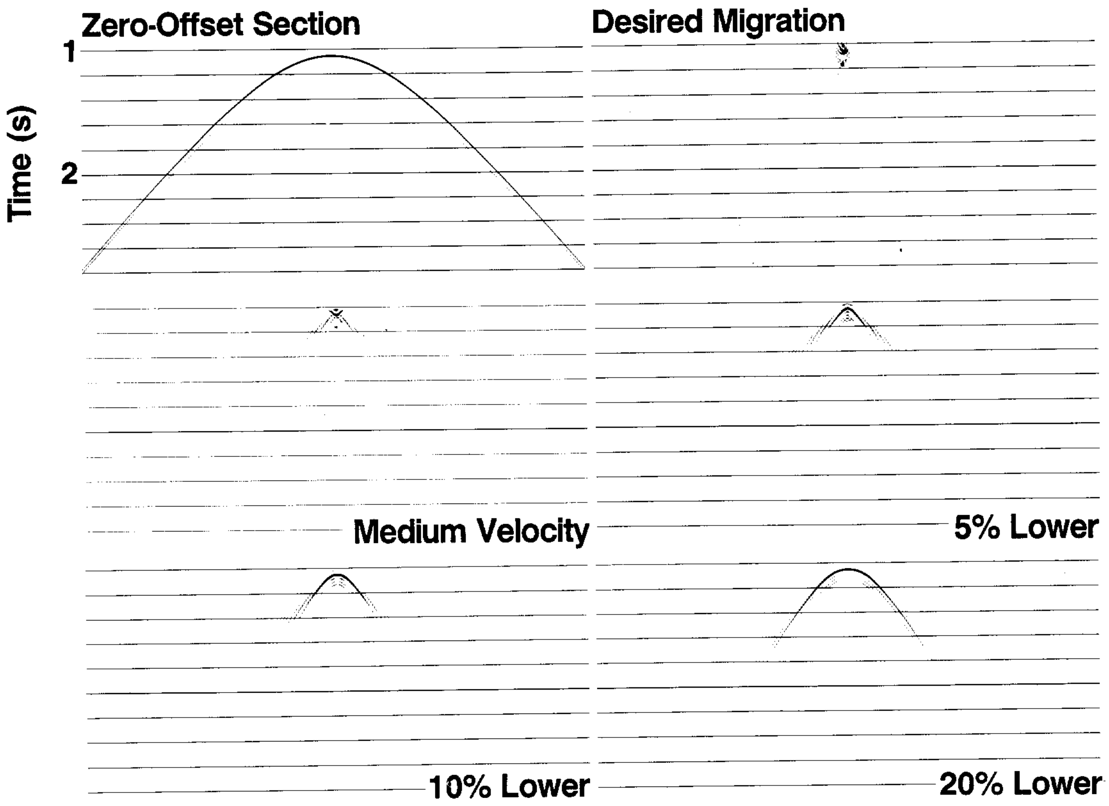


FIG. 4-96. Undermigration of the diffraction hyperbola with the 65-degree equation using velocities lower than the medium velocity.

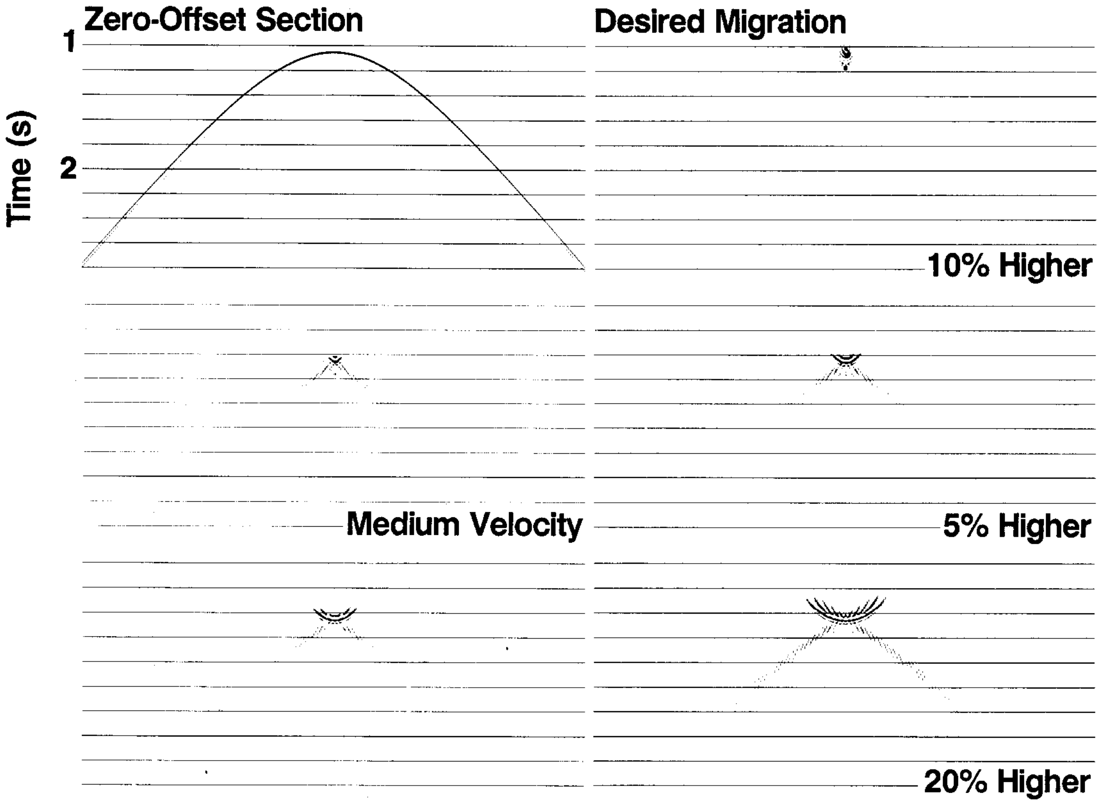


FIG. 4-97. Overmigration of the diffraction hyperbola with the 65-degree equation using velocities higher than the medium velocity.

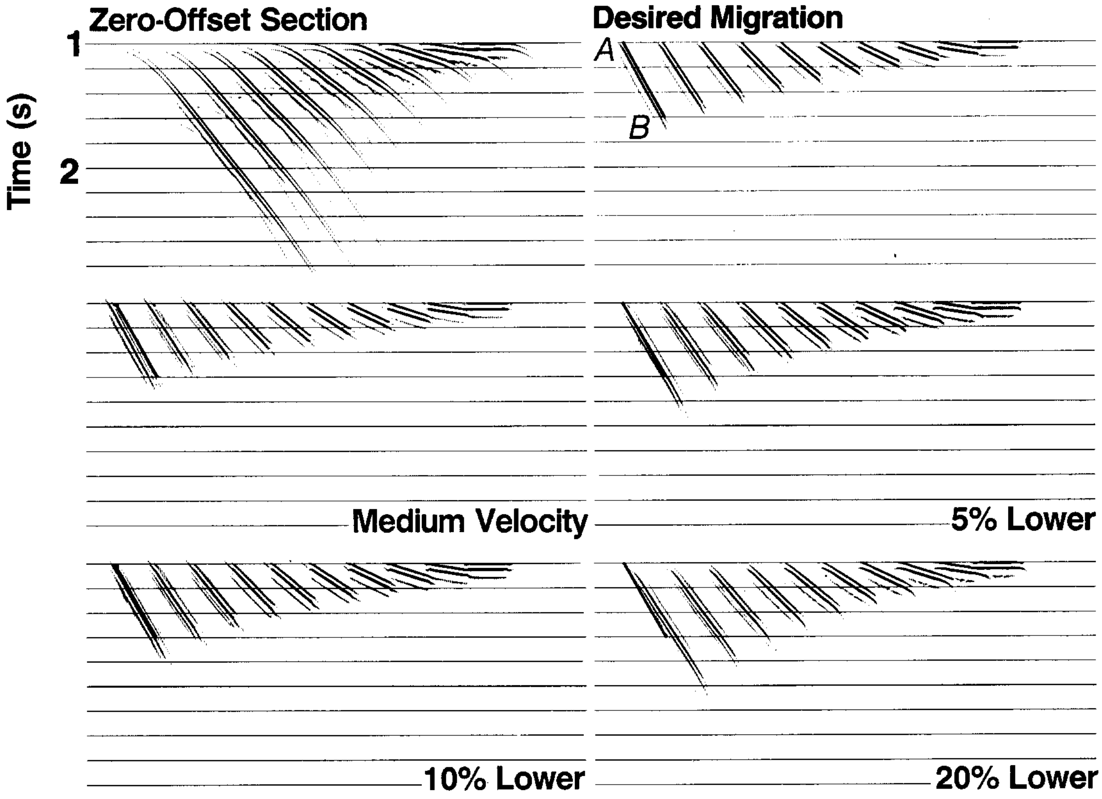


FIG. 4-98. Undermigration of dipping events with the 65-degree equation using velocities lower than the medium velocity.

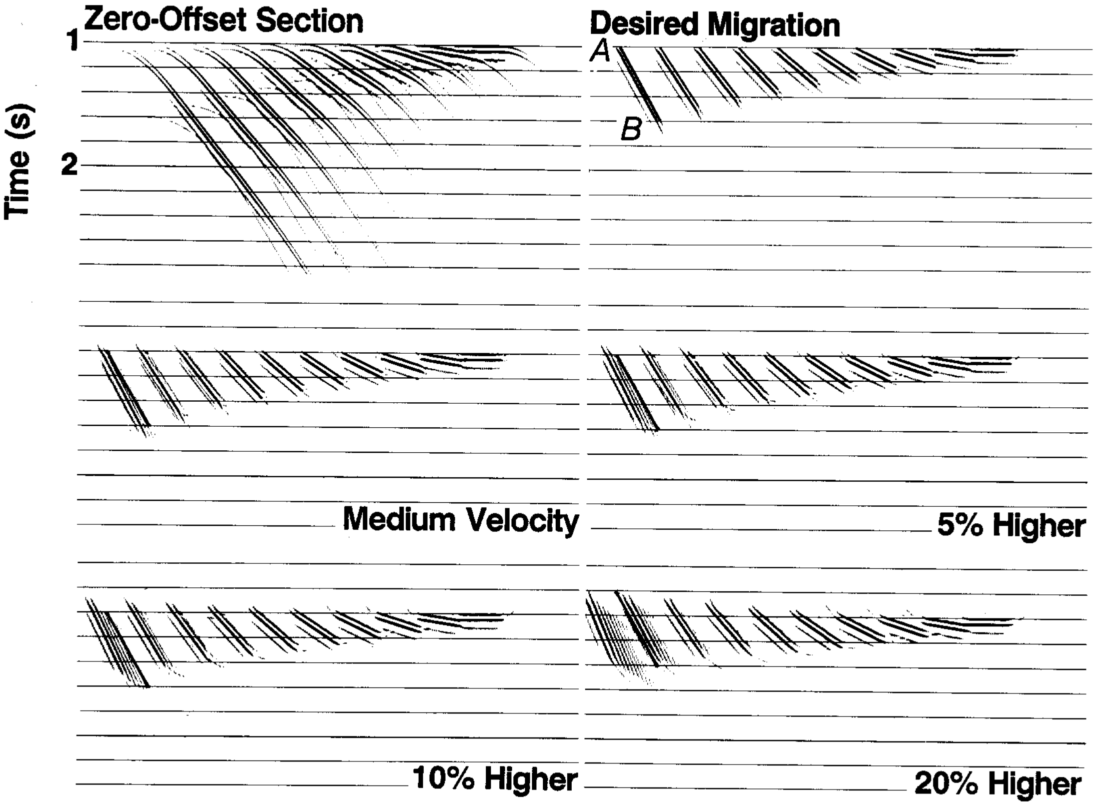


FIG. 4-99. Overmigration of dipping events with the 65-degree equation using velocities higher than the medium velocity.

4.3.5 Migration and Spatial Aliasing

The concept of spatial aliasing was presented in Section 1.6.1. The effect of spatial aliasing on the migration process is examined here. Consider the following experiment with the dipping events model. Start with the original zero-offset section and migrate it. Results and the *f-k* spectra are shown in Figure 4-33. Throw away every other trace from the zero-offset section and migrate the resampled section. Trace spacing now is 50 m. From the *f-k* spectrum of the resampled section (Figure 4-100), note that the three steepest dips are aliased. The steepest dip (45 degrees) becomes aliased at 24 Hz. The aliased energy is in the left quadrant of the *f-k* plot; hence, migration perceives this energy as if it is dipping opposite to the dominant dip direction seen on the input section. Since migration moves energy in the updip direction, in this case the aliased energy moves to the right, while the unaliased energy moves to the left. The noise seen on the migrated section in Figure 4-100 is the aliased energy after migration. It is dispersed. Since each frequency component of the aliased energy is perceived to have a different dip by the migration schemes, the displacement of the energy after migration is frequency dependent. Also note the slight loss of temporal resolution along the steep dips. The part of the frequency band associated with the aliased energy has been dispersed away during migration.

The spatial aliasing problem gets more serious as the input model is resampled to a 100-m trace spacing (Figure 4-101). This time, not only are much lower frequencies (down to 12 Hz at the steepest dip) aliased, but also the aliased energy is aliased. (Note the wraparound to the right quadrant.) On the migrated section, noise is introduced by this complex aliased energy. Migration has completely failed to image the steep dips. Almost all of the energy associated with these dips has dispersed away from the true position of these events. This happened because sections were migrated at increasingly larger trace spacings.

Figure 4-102 shows migrations of the diffraction hyperbola at different trace spacings. Splitting of the aliased and unaliased energy as a result of migration is well illustrated by this diffraction model. The unaliased energy moves in the updip direction and collapses to the apex of the diffraction hyperbola, while the aliased energy moves away from the flanks, causing the migration noise seen on these figures. The general effect of spatial aliasing is the same for all migrations: Kirchhoff summation, finite-difference, and frequency-wavenumber. Regardless of the method, migration suffers from spatial aliasing.

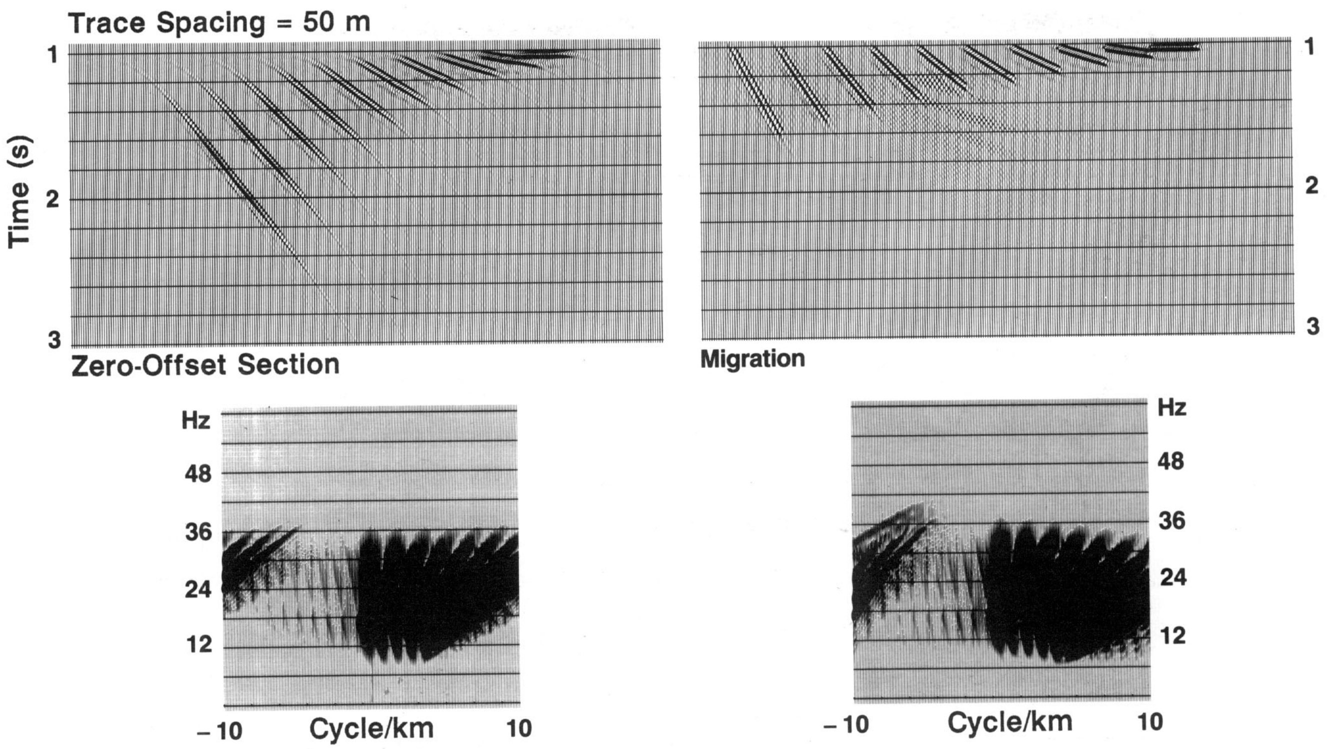


FIG. 4-100. Migration of dipping reflectors in the (t,x) (top) and $(f\text{-}k)$ (bottom) domains. Compare this with Figures 4-33 and 4-101.

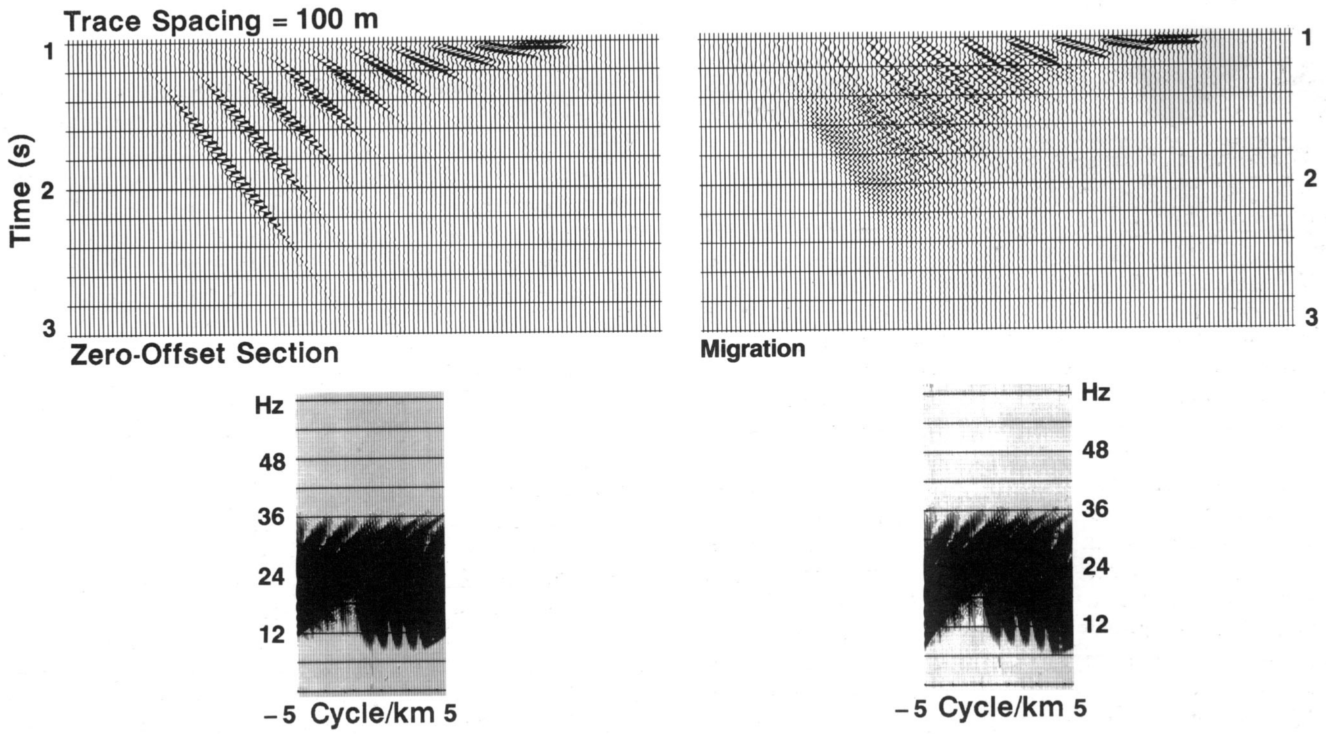


FIG. 4-101. Migration of dipping reflectors in the (t,x) (top) and (f,k) (bottom) domains. Compare this with Figures 4-33 and 4-100.

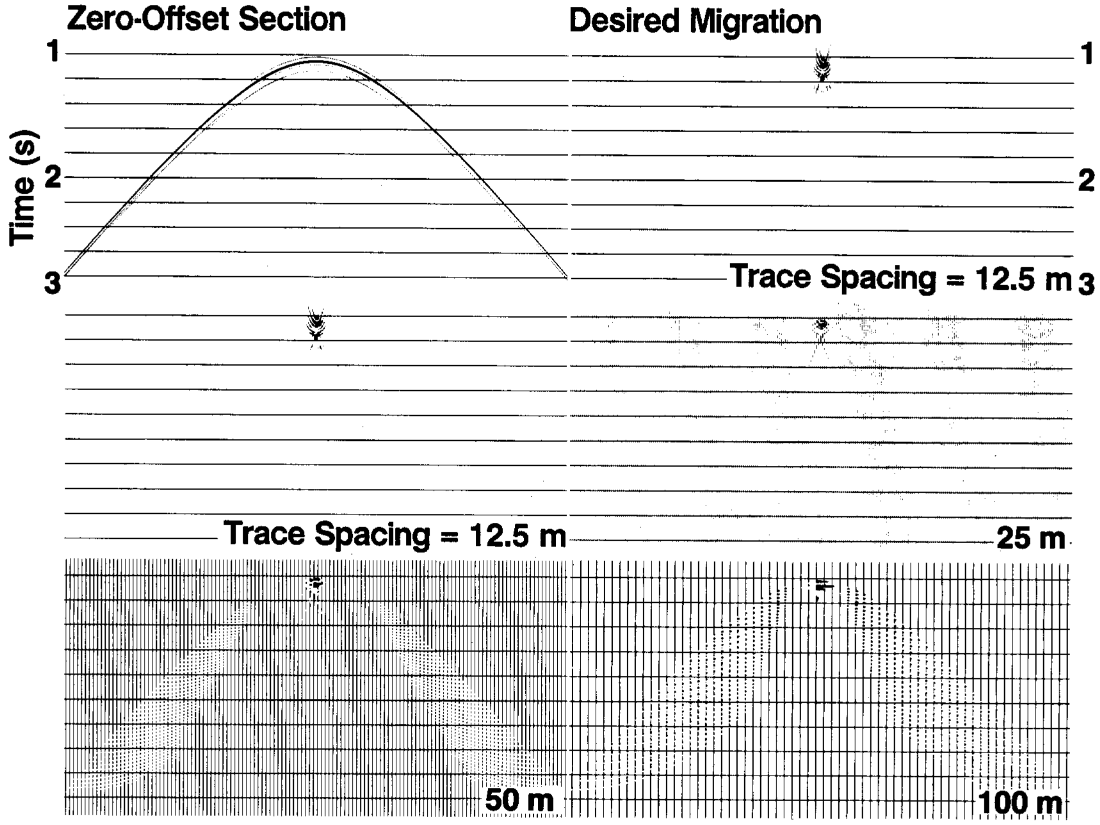


FIG. 4-102. Migration of a diffraction hyperbola using different trace spacings.

Aside from the spatial aliasing noise, dispersive noise also is seen on data migrated with a finite-difference algorithm (Section 4.3.2). This dispersive noise is an indication of data close to being aliased (Claerbout, 1985) and is demonstrated by the dipping events model in Figure 4-103. With a 50-m trace spacing, note the familiar precursors along the steeply dipping events and compare them with the migrated section in Figure 4-100. In Figure 4-103, for a 50-m trace spacing, the spatial aliasing noise *B* is distinguished from the dispersive noise *A*. Both types of noise are largely suppressed in case of data with a smaller trace spacing (e.g., 25 m). The 15-degree finite-difference method behaves similarly for the diffraction hyperbola example (Figure 4-104). Note the undermigration that is due to the 15-degree dip limitation, the dispersive noise *A*, which is due to the finite-difference approximations, and the spatially aliased energy *B*, which split away from the unaliased part that collapsed to the apex.

The effect of spatial aliasing on migration of field data is demonstrated in Figures 4-105 and 4-106. We see the original stacked section and its resampled versions at coarser trace spacings. From the migrations of these four stacked sections with coarser trace spacings, note the loss of spatial resolution (Figure 4-106).

What is the remedy for spatial aliasing noise in migration? Arrange the sequence of Figures 4-33, 4-100, and 4-101 from coarser to finer trace spacing. Note that the deleterious effect of spatial aliasing in migration disappears as we go to finer trace spacings. To avoid spatial aliasing, we must record with sufficiently fine CMP trace interval. Since we have a finite number of receiver groups and a finite cable length, we must select a trace interval that is neither so small it is wasteful, nor so large it causes spatial aliasing.

An optimum CMP trace interval can be computed as follows. Consider a dipping reflector with a dip angle of θ (Figure 4-107). Also consider a normal-incidence plane wave with a dominant period T recorded at the surface with a trace separation Δx. (This is the zero-offset case where Δx is the CMP trace interval.) From the geometry in Figure 4-107, we derive the following expression for the maximum threshold frequency that is not aliased for a given dip, velocity, and CMP trace interval:

$$f_{\text{threshold}} = v/(4\Delta x \sin \theta), \tag{4.17}$$

where v = medium velocity, Δx = trace interval, and θ = structural dip.

Table 4-3 shows the evaluation of equation (4.17) for a

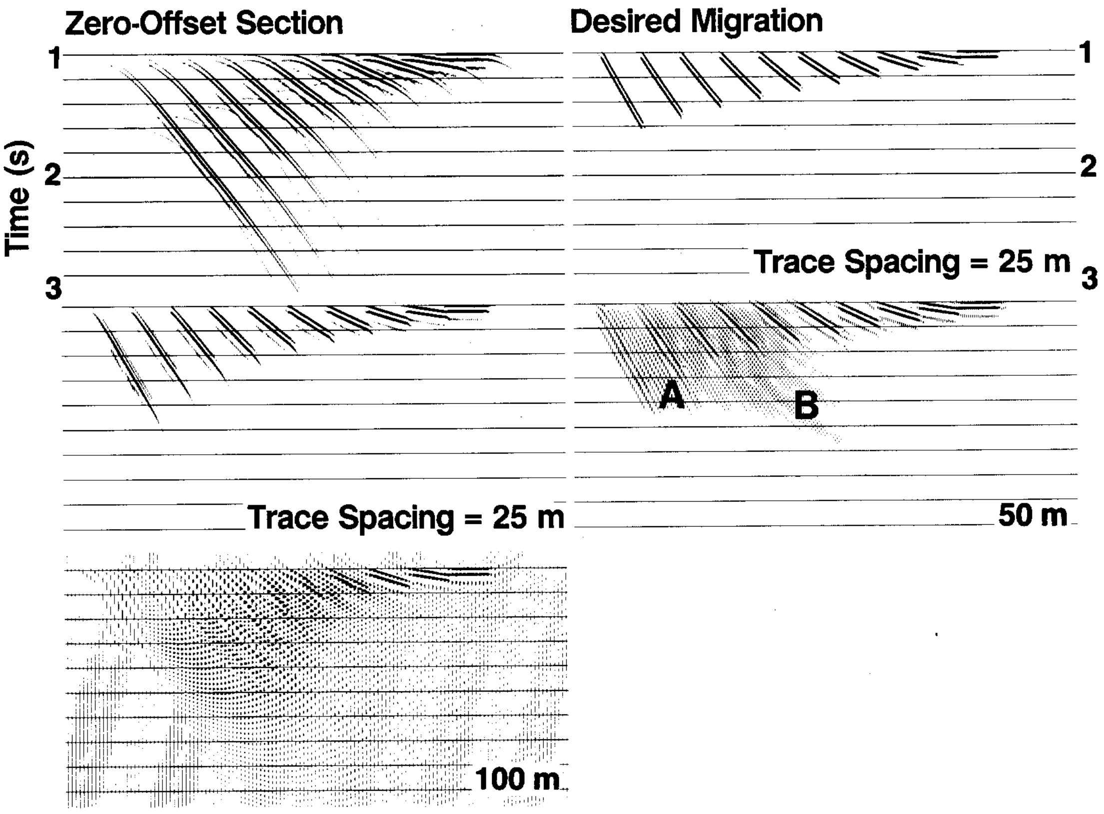


FIG. 4-103. Finite-difference (15-degree) migration of dipping events using different trace spacings. Note the presence of dispersive noise (*A*) due to finite-difference approximations in addition to the spatial aliasing noise (*B*). Compare the migration result from the 50-m trace spacing with that in Figure 4-100.

particular numerical example. Equation (4.17) also can be expressed in terms of receiver group interval 2 Δx. Suppose the maximum dip is 30 degrees. If the sampling interval were 4 ms, then the Nyquist frequency is 125 Hz. After an antialias filter, the frequency band extends up to 90 Hz, provided the high-cut filter is at three-quarters of the Nyquist. For a bandwidth without spatial aliasing, we have to select a 12.5-m CMP trace interval.

If we do not have control over the specification of the field parameters (group interval and sampling rate), there are two ways to minimize spatial aliasing. The first approach would be to filter out the aliased frequencies. This is undesirable, since it severely limits vertical and lateral resolutions (Section 8.3). The second approach would be to do trace interpolation before migration. Section 6.5.4 discusses the interpolation of aliased data. In Section 1.6.1, we found that the smaller the trace interval, the higher the Nyquist in the spatial wavenumber direction (Figure 1-77) and, thus, the less the likelihood of aliasing high-frequency data. A schematic illustration of this phenomenon is shown in Figure 4-108. Start with the spectral bandwidth that spans *COA* in the spatial wavenumber axis, where *A* is the location of the Nyquist wavenumber, and *ON* in the temporal frequency axis, where *N* is the location of the Nyquist frequency. Dip components 1 and 2 are aliased beyond frequency values *AT* and *AS*, respectively. Extend the wavenumber axis to *DOB* by making the trace interval half the original. Event 1 no longer is spatially aliased within the frequency bandwidth *ON*. Event 2 still is aliased beyond the frequency value *BV*. However, at this point and beyond, there may be no significant energy, so further extension of the wavenumber axis may not be necessary. Another important point is that if the temporal frequency band only extended up to *OG* to start with, extension of the wavenumber axis to *DOB* also would result in Event 2

Table 4-3. Frequency threshold for spatial aliasing. Velocity is 3000 m/s.

	Threshold Frequency (Hz) for CMP Trace Interval (m)			
Dip (deg) θ	12.5	25	37.5	50
10	346	173	115	86
20	175	88	58	44
30	120	60	40	30
40	93	47	31	23

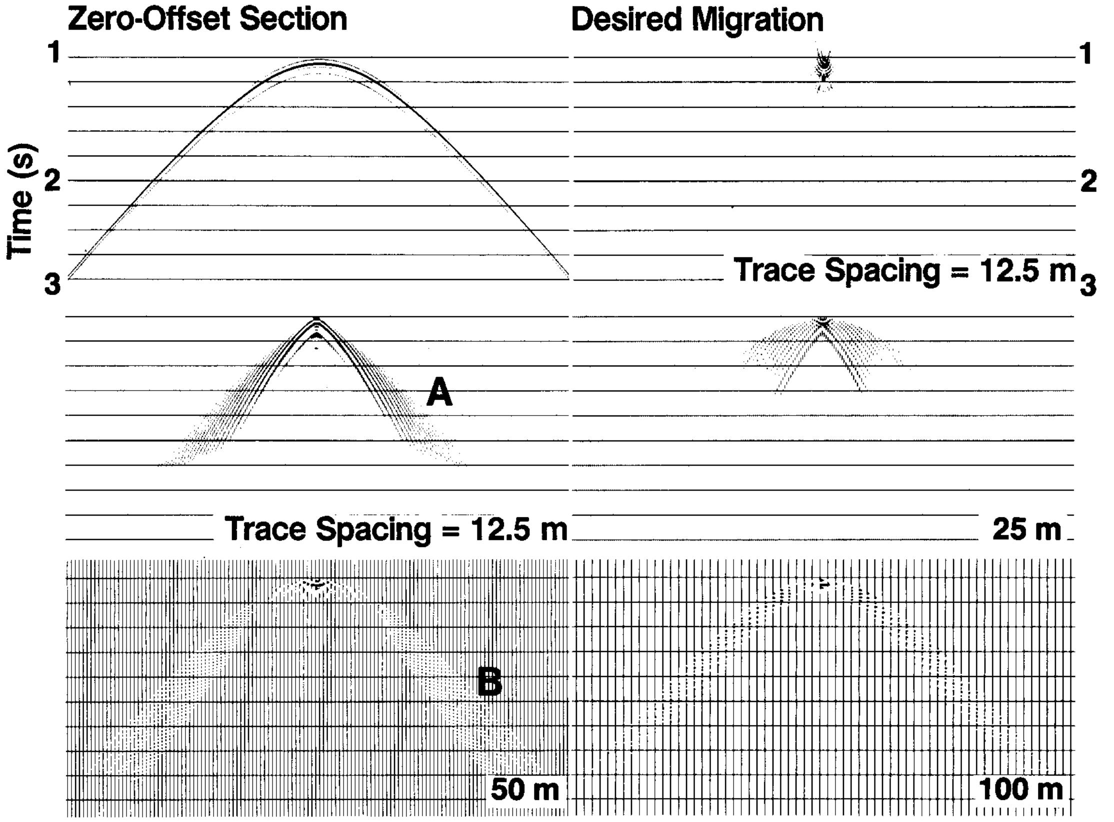


FIG. 4-104. Finite-difference (15-degree) migration of a diffraction hyperbola using different trace spacings. Note the dispersive noise (*A*) and spatial aliasing noise (*B*).

being unaliased. Thus, the amount of trace interpolation that is required also depends on temporal bandwidth as well as on structural dip.

Trace interpolation often is necessary when dealing with 3-D data and old data recorded with a large group interval. In a typical 3-D survey, the in-line trace interval may be as little as 12.5 m, while the trace interval in the cross-line direction, i.e., the line spacing, may be as much as 100 m. Therefore, interpolation is required before migration in the cross-line direction. You do not necessarily interpolate down to the in-line trace spacing; instead, depending on the maximum structural dip and velocity in the area, the optimum trace spacing for interpolation in the cross-line direction can be computed using equation (4.17). Section 6.5.4 provides more information on trace interpolation in relation to 3-D migration.

Besides the spatially aliased data problem, there is also the problem of aliased migration operators. In particular, for a low-velocity hyperbola, Kirchhoff summation may require more than one sample per trace. A good example is migration of the sea-bottom reflection using Kirchhoff summation, which results in some energy in the form of precursors above the migrated sea-bottom reflection, when only one point per trace is included in the summation.

4.3.6 Migration and Ambient Noise

In this section, the effect of migration on ambient noise in CMP stacked data is studied. Figure 4-109 shows a random noise section and the results of migrations using the three methods discussed in this chapter. Velocity linearly increases from 2000 m/s at the top to 4000 m/s at the bottom of the section. Note the smearing of amplitudes, especially at the bottom of the migrated sections. The amplitude and frequency characteristics of the input section in Figure 4-109 are virtually unchanged in the interior portions of the migrated sections. However, note the smearing of amplitudes at the bottom and side boundaries of these sections.

Ambient noise commonly dominates the deep portion of a stacked section where velocity is high. Therefore, noise organization due to migration generally is more severe in the deeper part of a stacked section where the S/N ratio is lower and velocities are higher. A field data example is shown in Figure 4-110. In addition to smearing effects, the migrated section of Figure 4-110 also has *smiles*, which are due to sparsely distributed bursts of amplitude in the input section. (Keep in mind that a single spike on the time section migrates to a semicircle on the

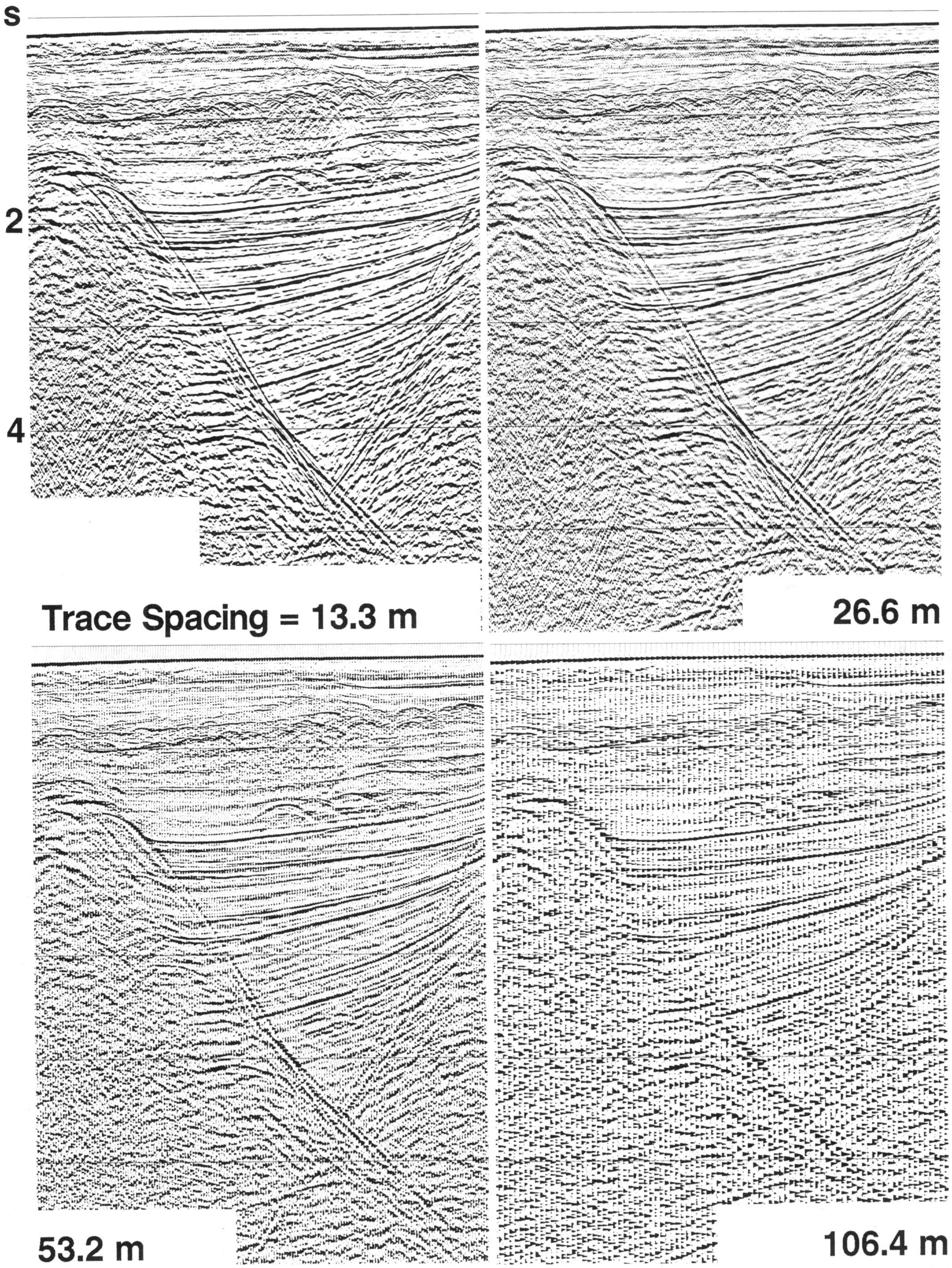


FIG. 4-105. The same CMP stack data set with different trace spacings.

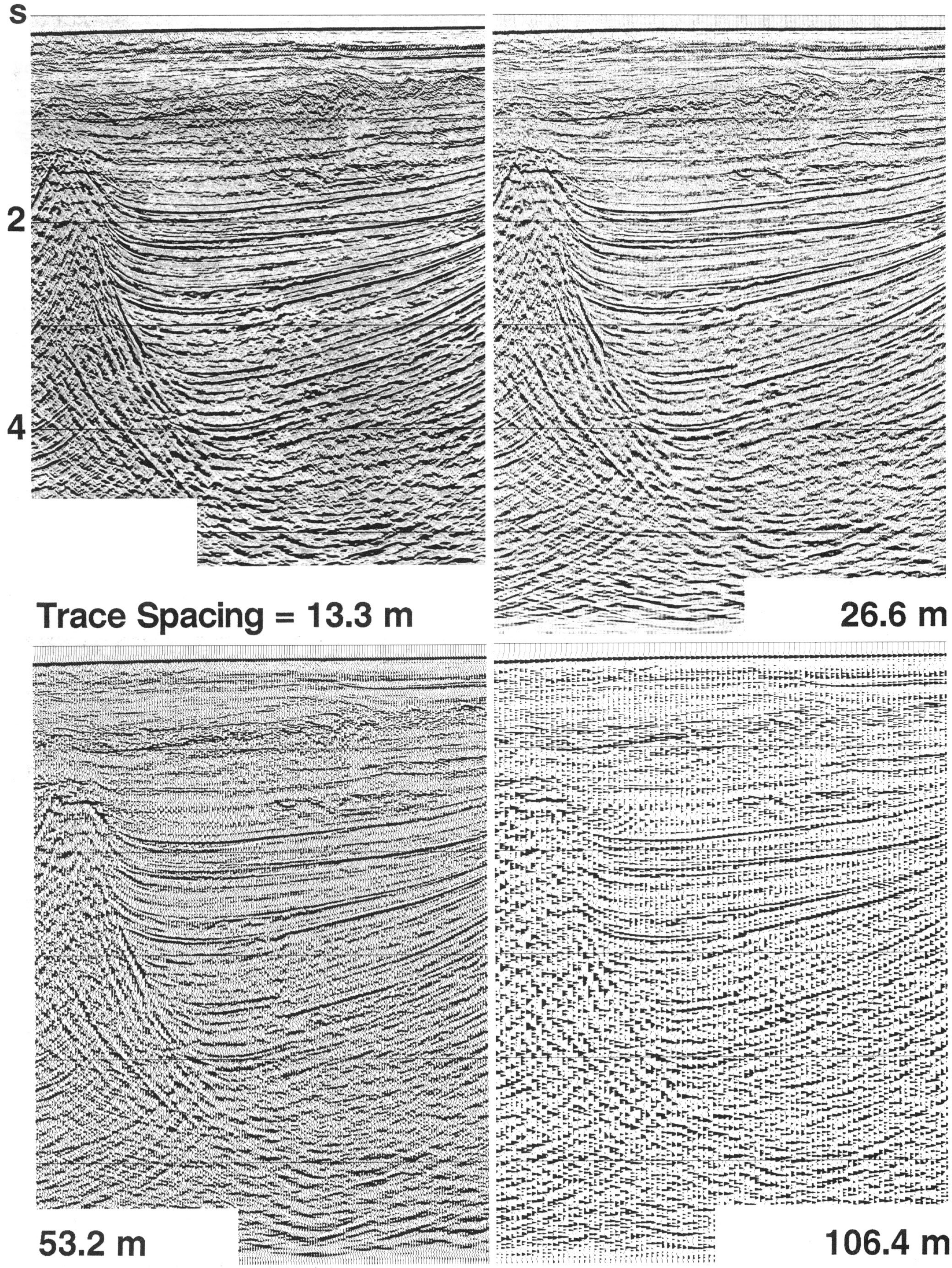


FIG. 4-106. Phase-shift migrations of the CMP stacked sections in Figure 4-105.

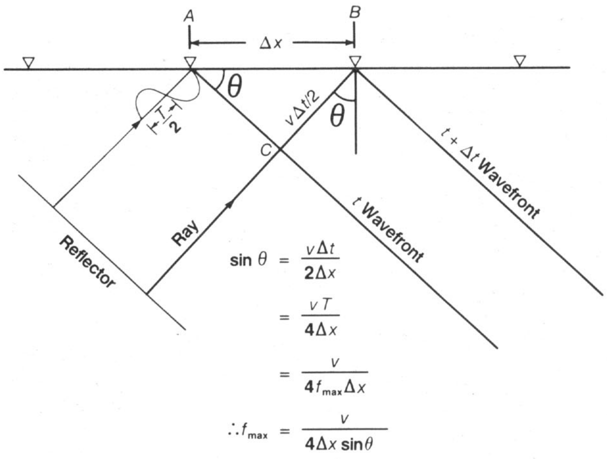


FIG. 4-107. Derivation of the threshold frequency for spatial aliasing [equations (4.17)]. Spatial aliasing occurs when the time difference between the arrivals at receivers *A* and *B* is one-half period (*T*/2) apart.

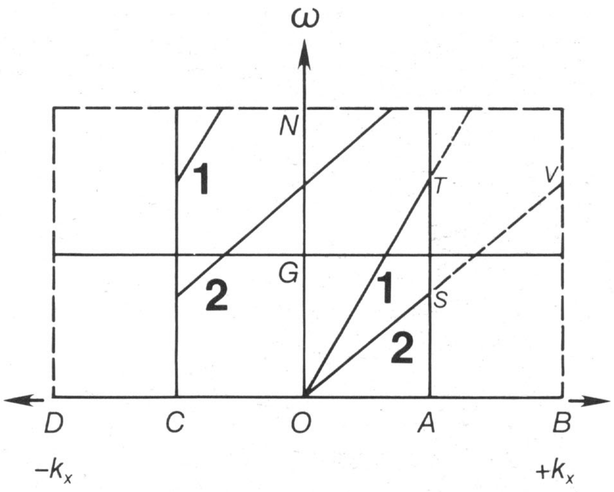


FIG. 4-108. Two dipping events in the (*f*,*k*) domain.

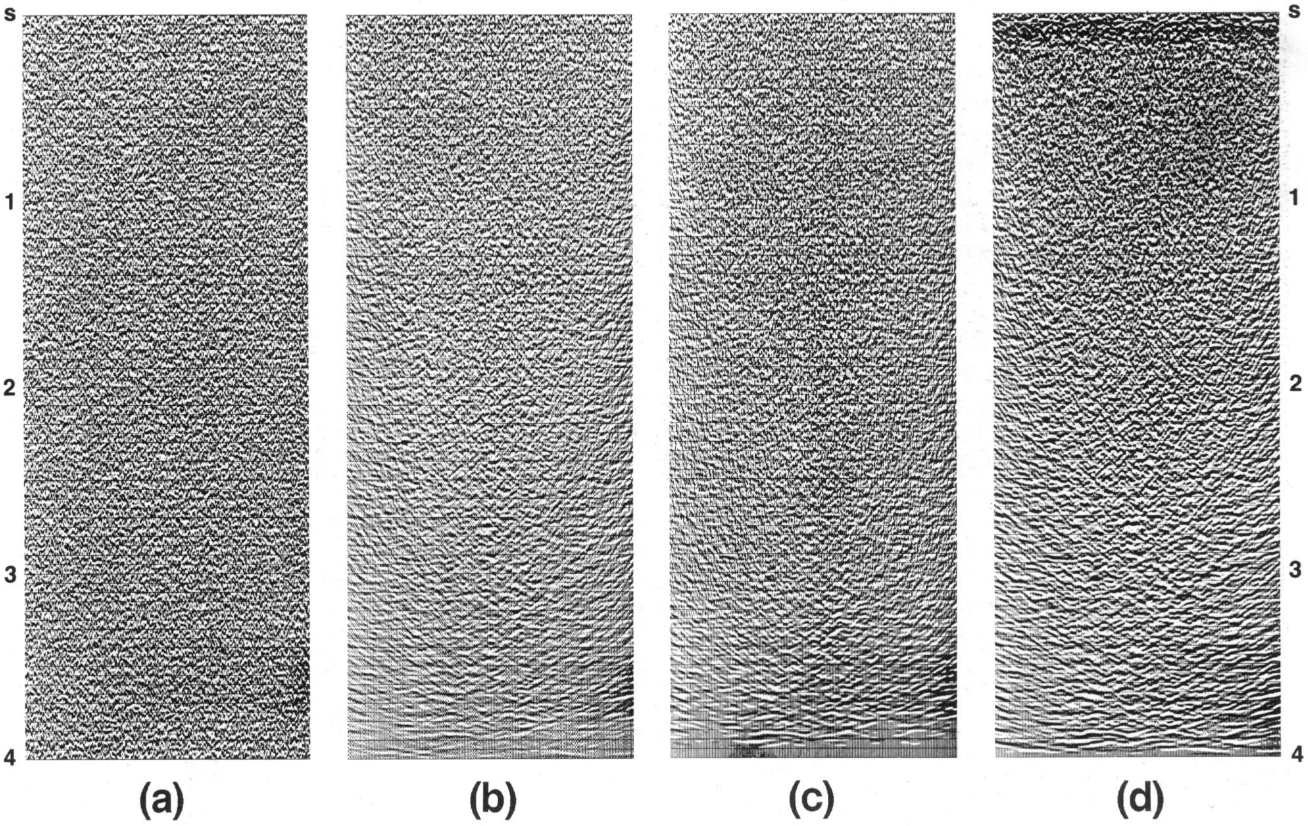


FIG. 4-109. Response of the three migration methods to random noise. (a) Zero-offset section for the $v(z)$ model; (b) *f-k*, (c) finite-difference, (d) Kirchhoff migrations.

depth section.)

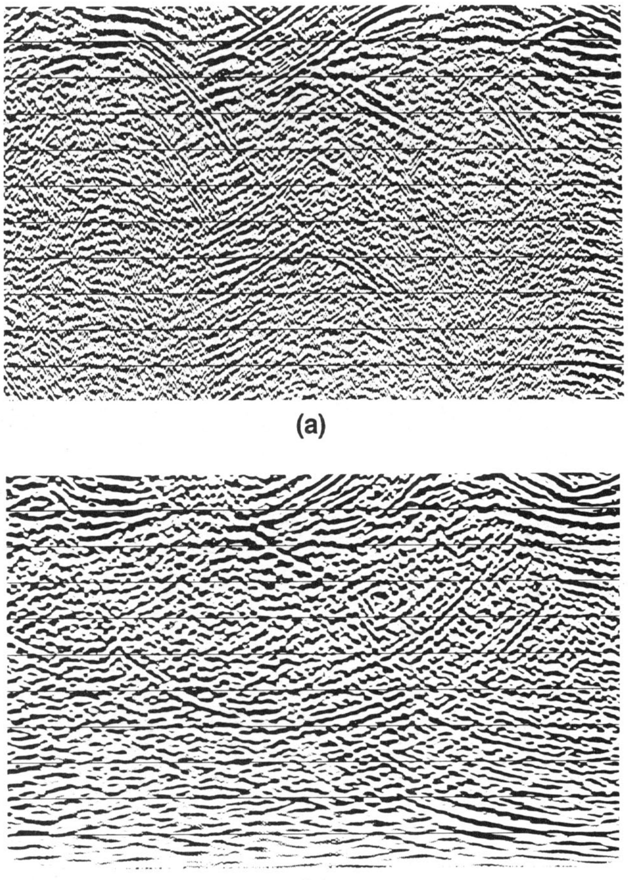
We have already seen the adverse effect of an improper choice of aperture width in Kirchhoff summation (Figure 4-46). A narrow aperture can introduce strong smearing as spurious, nearly horizontal events. A similar effect occurs for all types of migration algorithms if the maximum dip to migrate is severely restricted (Figure 4-84).

4.3.7 Migration and Profile Length

Because of economic and environmental restrictions, a seismic line may have to be recorded in the field with a shorter length than desired. To see the effect of line length on migration, we will examine the migrations of the decreasing lengths of the same CMP stack (Figure 4-111). Migration of the smaller portions, *BD* and *CD*, results in an increasingly smeared section, particularly in the deeper parts. We conclude that short seismic lines really are not suited for migration. If the section is too short, two effects occur. First, there is not enough space in the section for dipping events to move during migration. Second, side boundary effects contaminate a very large percentage of the final migrated section. The real solution to the problem is to record long lines. With a general idea of structural dip in an area, the field geophysicist must consider the additional spatial extent that is required by migration. (Refer to the discussion on Figure 4-14.) This is especially important in 3-D surveys in which the surface areal coverage must be extended beyond the subsurface areal coverage so that steep dips and structural discontinuities can be recorded and imaged properly (Section 6.3.1). The problem with 3-D is that cost increases as the square of the survey dimension, so that temptation to record too small a survey is great.

Regardless of line lengths, there are additional problems associated with the side boundaries of the stacked section input to migration. All migration algorithms implicitly make some assumption about the nature of data outside the side boundaries of the input stacked section. The simple assumptions, zero amplitude or zero gradient at the side boundaries of the section, cause data that should migrate past the edge to be reflected back into the section. To prevent this, traces of zero amplitude often are appended to the edges of the section. This allows the dipping events to move freely into the zero-amplitude region during migration. If the events that would migrate off the input section are not needed, then they often are suppressed by using absorbing side boundary conditions (Clayton and Engquist, 1980).

Figure 4-112 shows a section with significant smearing caused by side boundary effects. The wavefront character that dominates the left boundary of the migrated section down to the bottom of the mute zone can be explained using the principle of semicircle superposition for migration. Consider a dipping event *A* that extends down to the edge of the section as in the sketch in Figure 4-112. After migration *B*, note the remainder of the semicircular wavefront *C* on the left side. This wavefront did not cancel out during superposition because no data were available beyond the left boundary of the section. Another source of edge effects is the presence of amplitude bursts at or near the edge of the stacked section associated with a low S/N ratio due to low fold. The edge effects on the left boundary below the mute zone in the migrated section (Figure 4-112) probably stem from the improper amplitude level on the CMP stacked section. This amplitude imbalance is due to changes in fold at the end of the line and subsequent normalization problems.


(a)

(b)

FIG. 4-110. Migration (b) of the noisy part (deep in time) of a stacked section (a).

4.4 MIGRATION BEFORE STACK

We have learned several methods of migration applicable to stacked data. We assume that a stacked section is an appropriate representation of a zero-offset section. All the methods discussed so far work on zero-offset data.

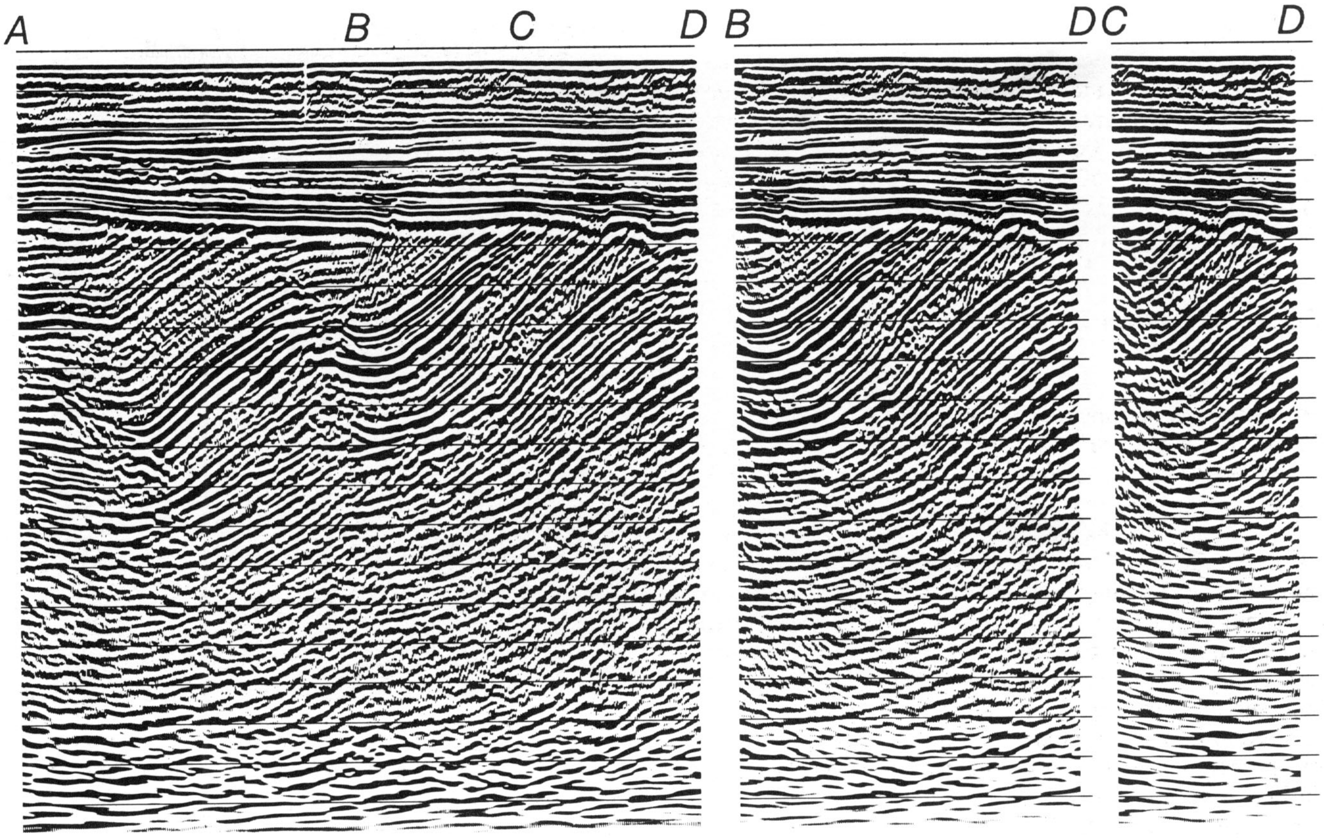
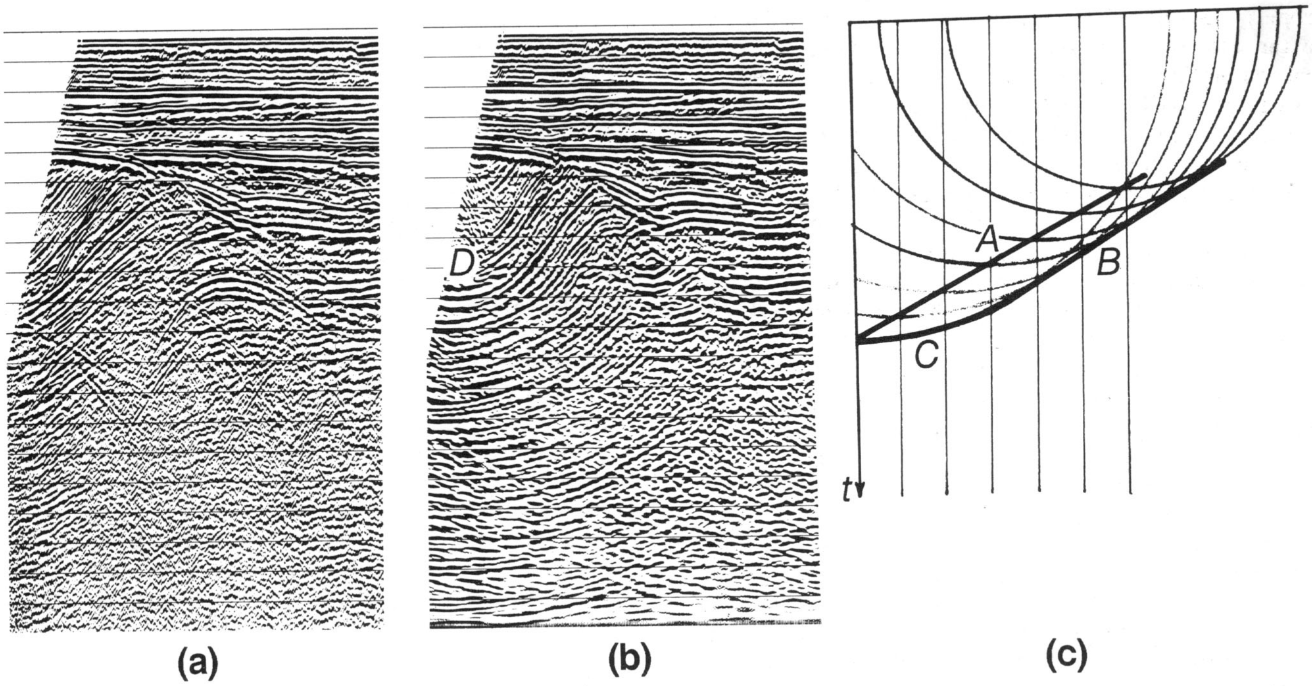


FIG. 4-111. Short profiles are not desirable for migration.



(a) (b) (c)

FIG. 4-112. Edge effects in migration. (a) CMP stack, (b) migration. Is there a syncline with its trough at location D? The sketch (c) illustrates the edge effect problem using the semicircle superposition method of migration.

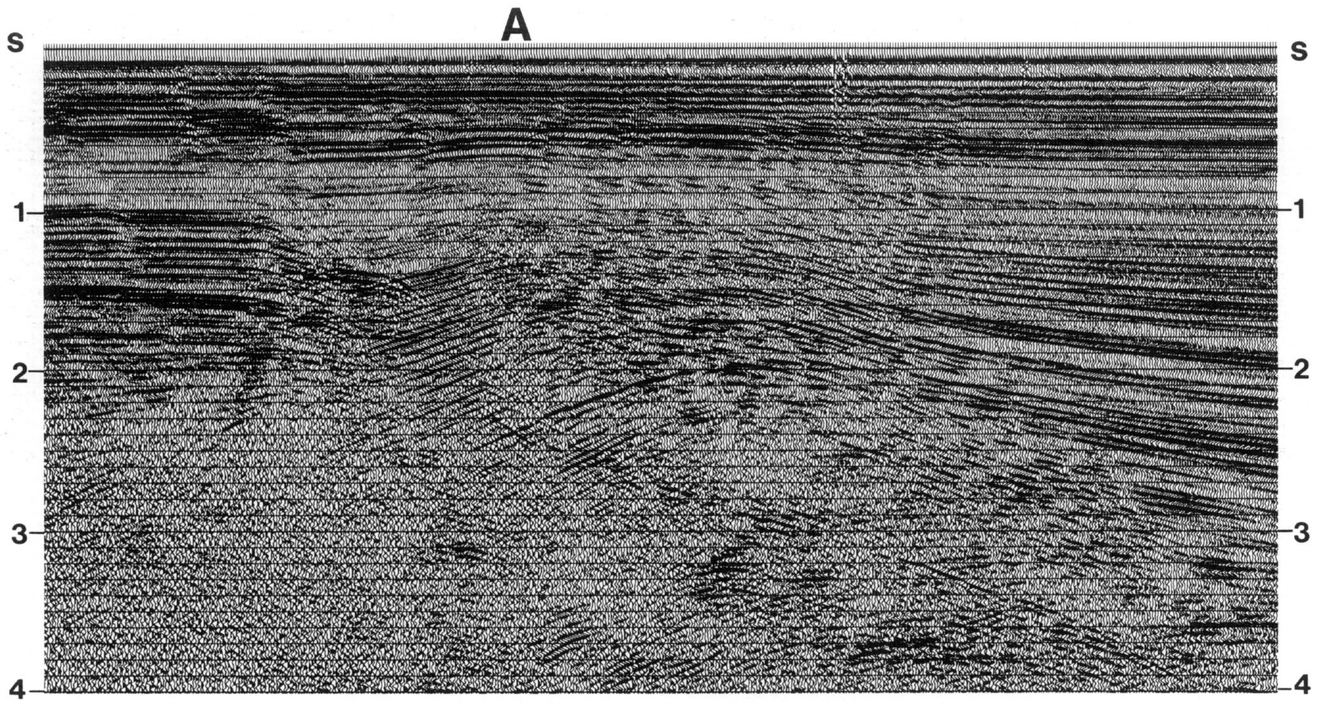


FIG. 4-113. A CMP stack containing conflicting dips along a major fault. (From a Western Geophysical brochure.)

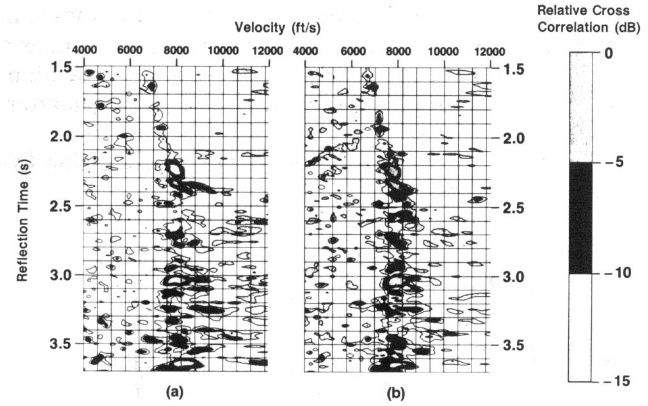


FIG. 4-114. Velocity analysis at location *A* as indicated in Figure 4-113; (a) before and (b) after prestack partial migration. (From a Western Geophysical brochure.)

We now examine the validity of the zero-offset assumption.

Consider the stacked section in Figure 4-113. This section contains a major fault plane that dips down from left to right, and many auxiliary faults that intersect this main fault. These fault patterns commonly occur in areas in which subsidence has taken place. We see reflections associated with the sedimentary strata, as well as faint indications of plane reflections. Along major fault zones, such as those in Figure 4-113, there is often a clear case of conflicting dips. From the velocity analysis in Figure 4-114a (at location A in Figure 4-113), two distinct correlation peaks, which occur at about 2.35 s associated with two conflicting dips. We will have a problem when trying to pick a velocity function from this velocity spectrum. We normally would pick along the predominant velocity trend. This leads to rejection of the pick associated with the fault-plane reflection. As a result, the amplitude of that event on the stacked section will be lower than one would like.

Similar problems exist for diffractions associated with faults and salt dome boundaries. From Figure 4-115, diffraction hyperbolas require different stacking velocities along the flanks. At the apex of the hyperbola and its vicinity, the optimum stacking velocity is the medium velocity (Figure 4-115b). Along the steep flank, the optimum stacking velocity is significantly greater than the medium velocity (Figure 4-115c). Ideally, we want a stack that best fits a zero-offset section with well-defined diffraction hyperbolas (Figure 4-115a).

An obvious example of conflicting dips associated with a salt dome boundary is shown in Figure 4-116. While picking velocities in favor of the flat reflections, the steep flank of the diffraction hyperbola *D*, which is from the tip (*P*) of the salt dome, is not stacked with sufficient strength (Figure 4-116a). This inadequate diffraction definition then causes the migration after stack to respond poorly (Figure 4-116c). The stacked section with a better definition of the diffraction energy (Figure 4-116b) yields a somewhat improved migration (Figure 4-116d). The interpreter may make a different interpretation from these migrated sections over the salt dome flank, in particular, between points *P* and *A*. The steep-dip event

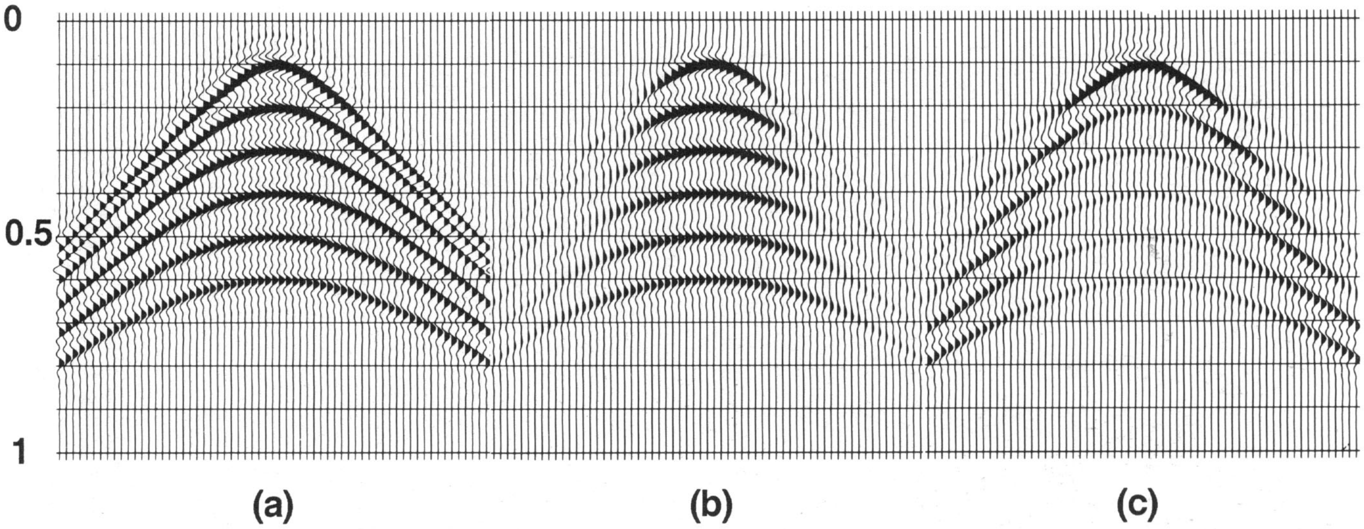


FIG. 4-115. Stack response of six point scatterers buried in a constant-velocity earth model (3000 m/s) as depicted in Figure 4-119. (a) Zero-offset section, (b) stack with NMO velocity of 3000 m/s, (c) stack with NMO velocity of 3600 m/s.

just below and to the right of point *A* (Figure 4-116d) has been migrated from deeper times into the section and is not related to the diffraction *D*.

When reflections occur at the same time with different stacking velocities, stacking quality degrades (Figure 4-116a). This is no surprise; in Section 3.2, we noted that stacking velocities are dip dependent from Levin's equation. Therefore, when a flat event is intersected by a dipping event, we can only choose a stacking velocity in favor of one of these events, not both. This is not the case for a zero-offset section, for it contains all events, regardless of dip (Figure 4-115a). Thus, we learn that in the presence of conflicting dips, stack no longer is equivalent to a zero-offset section.

Since all the migration theory discussed so far is based on the zero-offset assumption, we expect that migration after stack is not valid for conflicting dips with different stacking velocities. Why not consider migration before stack? A theory for imaging nonzero-offset data based on the double-square-root equation [equation (C.21)] is provided in Appendix C.1. Migration before stack can be done by Kirchhoff summation based on the nonzero-offset traveltime equation for a point scatterer, which can be derived from the DSR equation (Clayton, 1978). Instead of summing along the zero-offset diffraction hyperbolas (Figure 4-115a), amplitudes are summed along the nonzero-offset diffraction traveltime trajectories (Figure 4-120a). As with the zero-offset case, the velocity field dictates the curvature of these summation paths. Each common-offset section is imaged separately in this way; the results then are superimposed to produce the migrated section. The migration-before-stack section that corresponds to the field data in Figure 4-113 is shown in Figure 4-117. Clearly migration before stack produces a superior section; all dips now are present in the section. Figure 4-117 should be compared to the migration-after-stack section in Figure 4-118.

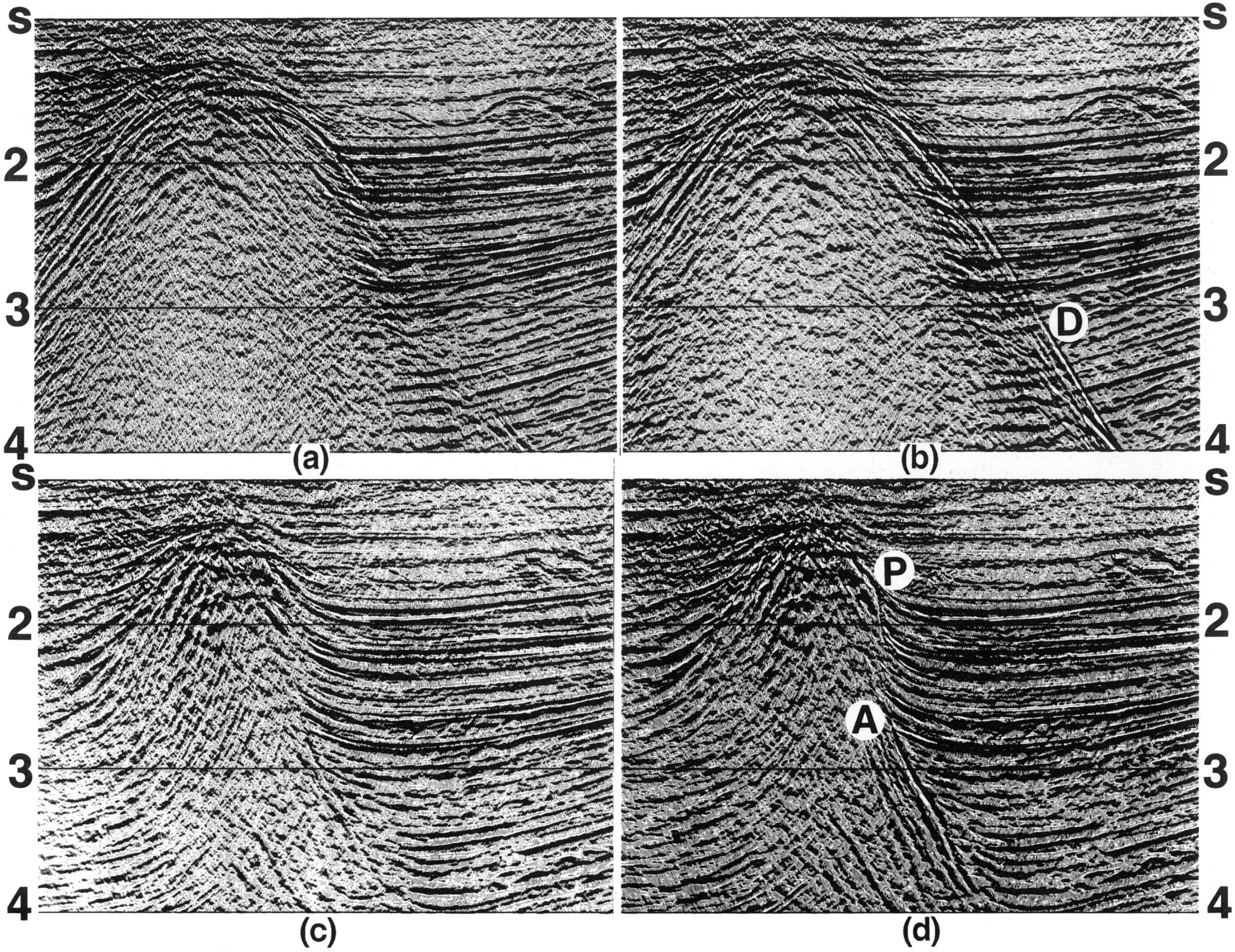
A simplified flowchart of conventional processing follows:

Field Data
↓
Zero-Dip NMO
↓
Stack
↓
Zero-Offset Migration
(Migration after Stack)

Instead of the conventional processing flowchart, we may consider DSR processing for handling conflicting dips:

Field Data
↓
v(x,z) → DSR Operator
↓
Migration Before Stack

Although we can solve the conflicting dips problem by migration before stack, other problems are associated with this approach. First, it is expensive. For 48-fold data, 48 different migrations would have to be done and the imaged sections would have to be superimposed.


S
2
3
4
(a)
(b)
(c)
(d)
D
P
A

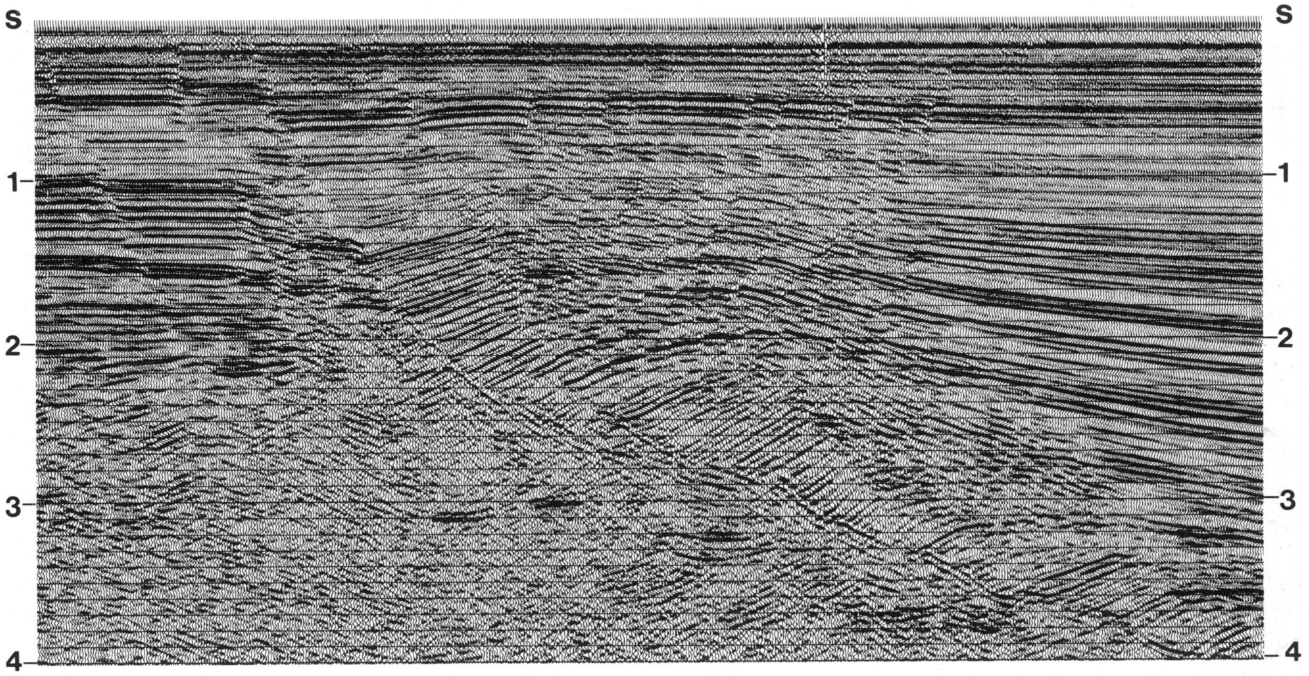


FIG. 4-117. Prestack migration (Kirchhoff summation) of the data in Figure 4-113. (From a Western Geophysical brochure.)

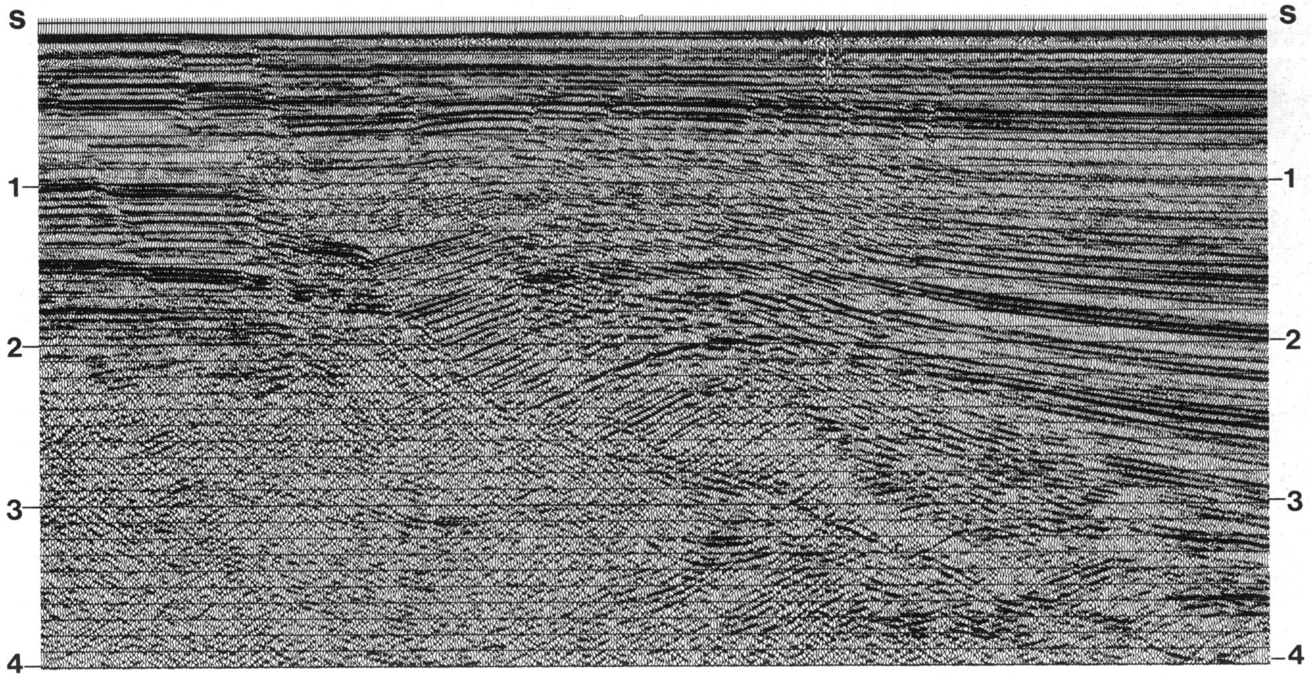


FIG. 4.118. Poststack migration (Kirchhoff summation) of the section in Figure 4-113. (From a Western Geophysical brochure.)

◂ FIG. 4-116. (a) CMP stack with a salt dome, (b) stack derived from dip moveout corrected data, (c) phase-shift migration of (a), (d) phase-shift migration of (b). (See text for details.)

Certainly partial stacking before migration can be considered to reduce cost. However, this can be done only to a degree, without compromising the quality of the output section. Second, as for any other migration method, migration before stack also requires knowledge of velocities and thus is sensitive to the effects of velocity errors. These effects are most severe at steep dips, precisely where migration before stack should be most useful. Finally, there is also the problem of spatially aliased common-offset sections (especially with land data).

Migration before stack provides exactly what its name implies: a migrated section. You cannot produce an intermediate product that is an unmigrated stacked section. The interpreter prefers to have the unmigrated stacked section in addition to the migrated stacked section partly because limited accuracy in velocity estimation makes the migrated section (on its own) an unreliable interpretation tool.

4.4.1 Prestack Partial Migration (Dip Moveout)

Can we improve on conventional processing to get a *better stack*, that is, an unmigrated section with all dips preserved? The following flowchart is an alternative to the convential processing flowchart:

Field Data
↓
Zero-Dip NMO
↓
Prestack Partial Migration
(PSPM)
↓
Better Stack
↓
Zero-Offset Migration
(Migration After Stack)

What does PSPM stand for? Since our problem with a conventional stack is conflicting dips, we must reconsider the NMO equation (Levin, 1971) for a single dipping reflector:

$$t^2(x) = t^2(0) + \frac{x^2 \cos^2 \theta}{v^2}, \qquad (4.18)$$

where θ is reflector dip, v is the velocity of the medium above the reflector, and x is the source-receiver separation (Figure 3-16). If the moveout term is divided into two parts, then

$$t^2(x) = t^2(0) + \frac{x^2}{v^2} - \frac{x^2 \sin^2 \theta}{v^2}. \qquad (4.19)$$

The first part of the moveout is associated with the zero-dip normal moveout (NMO), while the second part is associated with dip moveout (DMO), since it is the moveout related to reflector dip. The above equation implies that we can first apply NMO correction using medium velocity, then apply DMO correction. Later, we show that the DMO process can be achieved by prestack partial migration (PSPM). Unlike the NMO term, which is implemented in CMP domain, the DMO term needs to be implemented in a domain in which dips can be recognized, such as the common-offset domain.

From equation (4.19), we can easily assess the properties of the DMO term. First, it has no effect on zero-offset data ($x = 0$), regardless of dip. Second, the steeper the dip, the larger the correction. Third, the lower the velocity, the larger this correction. This also implies that the shallower the event, the more significant this term, since lower velocities generally are found in shallow parts of the seismic data. Now that the properties of this term have been discussed, we can implement it.

The conflicting dips problem has been investigated extensively. Doherty (1975) first introduced the wave extrapolation equations for nonzero-offset data. Sherwood et al. (1978) devised a method for mapping nonzero-offset data to zero offset in the presence of conflicting dips with different stacking velocities. Based on a wave-theoretical formulation to account for the difference between DSR processing and conventional processing, Yilmaz and Claerbout (1980) suggested a PSPM technique for solving the conflicting dips problem. This theory had one important drawback. Although valid for the layered earth velocity model, it was based on a small-offset approximation. Deregowski and Rocca (1981) recast the theory for PSPM in a form similar to Kirchhoff migration. They called their correction term for dipping events *dip moveout* (DMO). The PSPM and DMO terms now are used synonymously. Ottolini (1982) developed the PSPM equations in Snell-midpoint coordinates, the domain of constant-ray-parameter sections (Appendix D). This approach is theoretically accurate for layered medium as well as for all offsets and dips. However, the cost of creating constant-ray parameter sections may prohibit its implementation. This method was followed by another unique approach that involves offset continuation (Rocca et al., 1982); that is, mapping a far-offset section to a near-offset section, thereby collapsing all offsets to zero offset. Hale (1983) formulated a DMO method in the *f-k* domain. This method is exact for constant velocity, can handle all dips and offsets, and is accurate as long as the vertical velocity gradient is moderate. Fowler (1984) developed a mapping technique to correct dip-dependency of stacking velocities in the (v, k_y, ω) domain. Finally, French et al. (1984) developed a partial migration technique that tries to account for variations in source-receiver azimuths, a form particularly suitable for 3-D applications. All of these techniques are confined to vertical velocity variations for which time migration is appropriate. However, PSPM cannot solve stack imperfections caused by lateral velocity variations. (The problem of lateral velocity variations is discussed in Chapter 5.)

The Hale method is computationally intensive. It is like doing NMO correction in the frequency domain. In practice, often some form of Kirchhoff implementation of PSPM (such as Deregowski and Rocca, 1981) currently is used. This approach also is immediately applicable to 3-D seismic data. Nevertheless, because of its exactness for constant velocity, the DMO process is described qualitatively using the Hale method.

We rewrite equation (4.19) in the following form:

$$t^2(x) = t_n^2(x) - x^2p^2, \qquad (4.20)$$

where $t_n^2(x) = t^2(0) + x^2/v^2$ and $p = \sin\theta/v$, the ray parameter. In the *f-k* domain, $p = k_x/2\omega_0$, where ω_0 is the transform variable associated with the two-way zero-offset time $t(0)$. Once the NMO-corrected common-offset data are Fourier transformed in midpoint direction, dip and velocity are explicitly removed from the DMO term x^2p^2. Thus, in the *f-k* domain, the DMO correction process requires neither dip specification nor velocity information.

Figure 4-119, a depth model we will use to illustrate the DMO process and related practical issues, consists of six point scatterers buried beneath the center midpoint (CMP 32). The selected common offset sections associated with this subsurface model are shown in Figure 4-120a. The offset range is 50 to 1550 m in 50-m increments. The well-known nonhyperbolic table-top trajectories are apparent at large offset.

The CMP gathers from the model of Figure 4-119 are shown in Figure 4-120b. Only selected gathers that span the right side of the center midpoint are displayed, since the common-offset sections are symmetric with respect to the center midpoint (CMP 32). Note that the traveltimes at the center midpoint are perfectly hyperbolic, while the traveltimes at CMP gathers away from the center are increasingly nonhyperbolic. The following sequence describes DMO processing for these data.

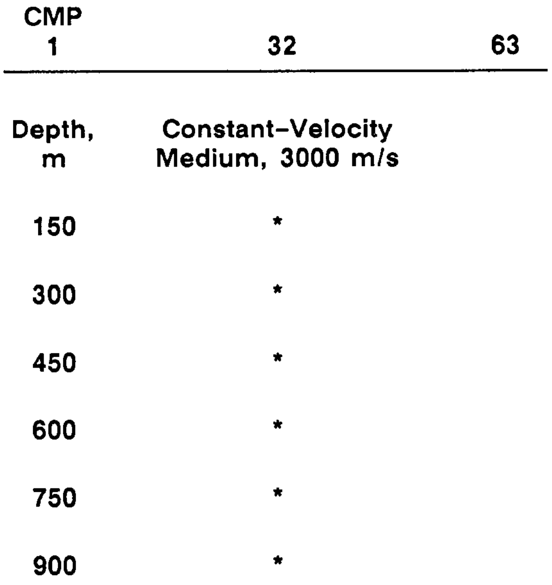
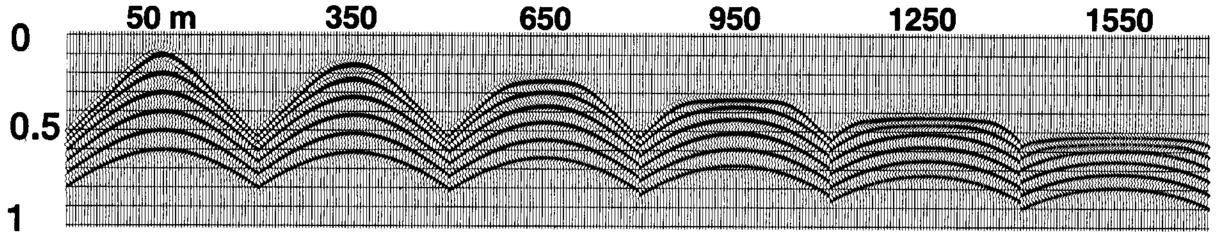
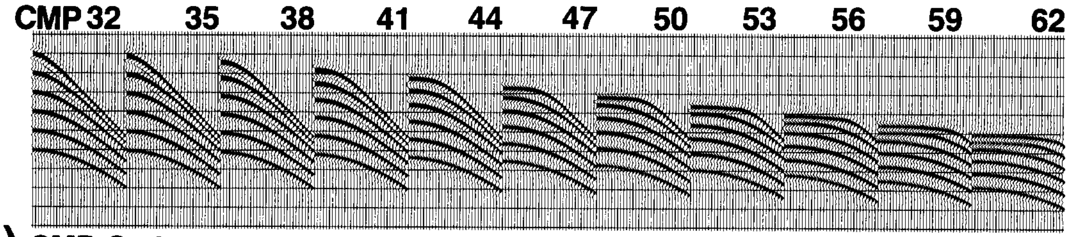


FIG. 4-119. Depth model of six point scatterers buried in a constant-velocity medium. The asterisks indicate the positions of the point scatterers.

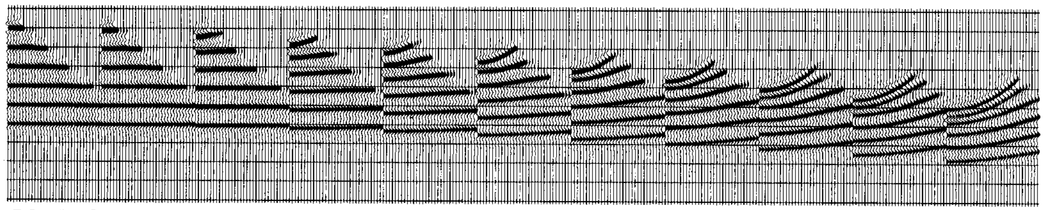
1. Figure 4-120c shows the NMO-corrected gathers, with stretch muting applied. The medium velocity (3000 m/s) was used for NMO correction, an essential requirement for subsequent DMO correction [equation (4.19)]. As a result, the events at and in the vicinity of the center midpoint (CMP 32) are flat after NMO correction, while the events at midpoints away from the center midpoint are increasingly overcorrected.
2. The stacked section derived from these gathers (Figure 4-120c) is shown in Figure 4-121b. Because medium velocity was used for NMO correction, the stack response is best for zero dip. Note the poor stack response along the steeply dipping flanks. The desired section is the zero-offset section in Figure 4-121a.
3. We sort the NMO-corrected gathers (Figure 4-120c) into common offset sections for DMO processing. These are shown in Figure 4-120d.
4. Each common offset section is individually corrected for DMO (Figure 4-120e). Note the following effects of DMO:
   a. DMO is a partial migration process. The flanks of the nonhyperbolic trajectories have been moved updip just enough to make them look like zero-offset trajectories, which are hyperbolic. As a result, each common-offset section after NMO and DMO corrections is approximately equivalent to the zero-offset section (Figure 4-121a).
   b. This partial migration is subtly different from conventional migration in one respect. Unlike conventional migration, the DMO action becomes greater at increasingly shallow depths.
   c. The DMO action also is greater at increasingly large offsets. In fact, it does nothing to the zero-offset section.
   d. Finally, as with conventional migration, the steeper the event, the greater partial migration takes place, flat events remaining unaltered.
5. Following the DMO correction, the data are sorted back into CMP gathers (Figure 4-120f). Compare this to the CMP gathers without DMO correction (Figure 4-120c). The DMO correction left the zero-dip events unchanged (at and in the vicinity of CMP 32), while it substantially corrected steeply dipping events on the CMP gathers away from the center midpoint (CMP 32). The events on the CMP gathers now are flattened. Also, since DMO correction is a migration-like process, it causes the energy to move from a CMP gather to neighboring gathers in the updip direction. Energy depletion at the CMP gathers farther from the center midpoint occurred because there was no other CMP gather to contribute energy beyond CMP 63.



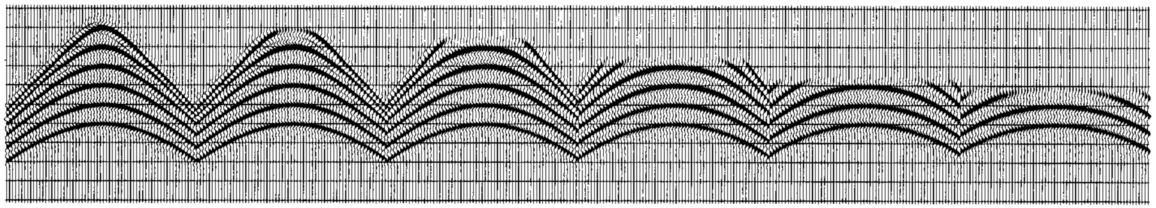
(a) Common Offset Sections Constant Velocity = 3000 m/s



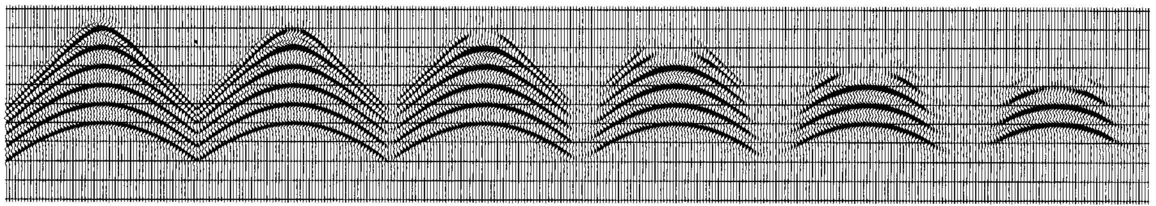
(b) CMP Gathers



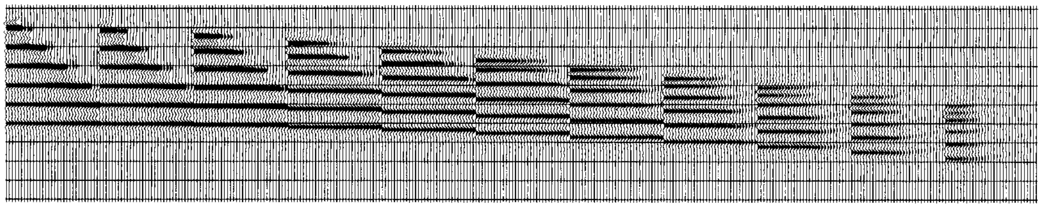
(c) CMP Gathers NMO



(d) Common Offset Sections NMO



(e) Common Offset Sections NMO + DMO



(f) CMP Gathers NMO + DMO

FIG. 4-120. Intermediate results from DMO processing the nonzero-offset synthetic data derived from the depth model in Figure 4-119.

6. Stacking the NMO- and DMO-corrected gathers (Figure 4-120f) yields a section (Figure 4-121c) that more closely represents the zero-offset section (Figure 4-121a) than the stacked section without the DMO correction (Figure 4-121b). Note the enhanced stack response along the steeply dipping flanks in Figure 4-121c. (The sections all have the same display gain.)
7. Migrations of the sections in Figure 4-121 are shown in Figure 4-122. Note the poor quality of focusing without DMO (Figure 4-122b). On the other hand, the stack with DMO correction yields a migration that is comparable in quality to that of the zero-offset section. Based on these model studies, we conclude that NMO + DMO + stack + time migration after stack is approximately equivalent to full-time migration before stack.

The previous discussion was based on a constant velocity assumption. To be of practical use, DMO must be applicable to data with velocity gradients. Figure 4-61 shows the depth model we will use to investigate DMO in the case of vertical velocity variations. The depth model consists of three point scatterers buried beneath the center midpoint (CMP 32) in a medium with a horizontally layered velocity structure.

Selected common offset sections and CMP gathers associated with this subsurface model are shown in Figures 4-123a and 4-123b. The same processing sequence is followed as was used for the constant-velocity model (Figure 4-120). The NMO correction (Figure 4-123c) before DMO correction is done using the rms velocity function indicated in Figure 4-61. The stacked sections with and without DMO are shown in Figure 4-124 with the zero-offset section based on the velocity model in

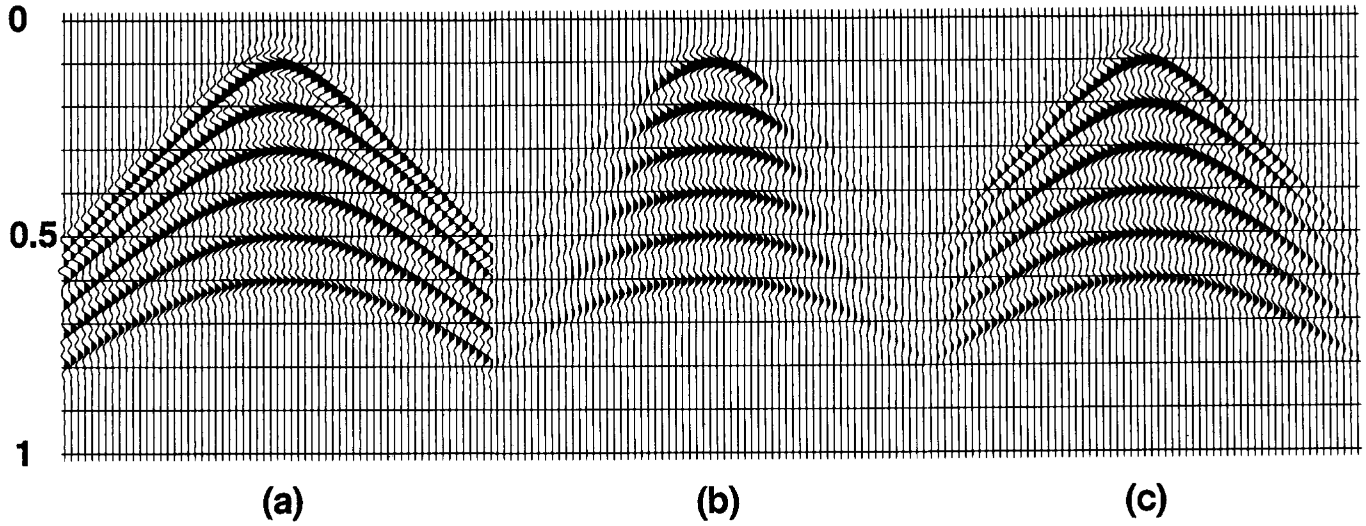


FIG. 4-121. (a) Zero-offset section associated with the depth model in Figure 4-119, (b) stack derived from the CMP gathers in Figure 4-120c, (c) DMO stack derived from the CMP gathers in Figure 4-120f.

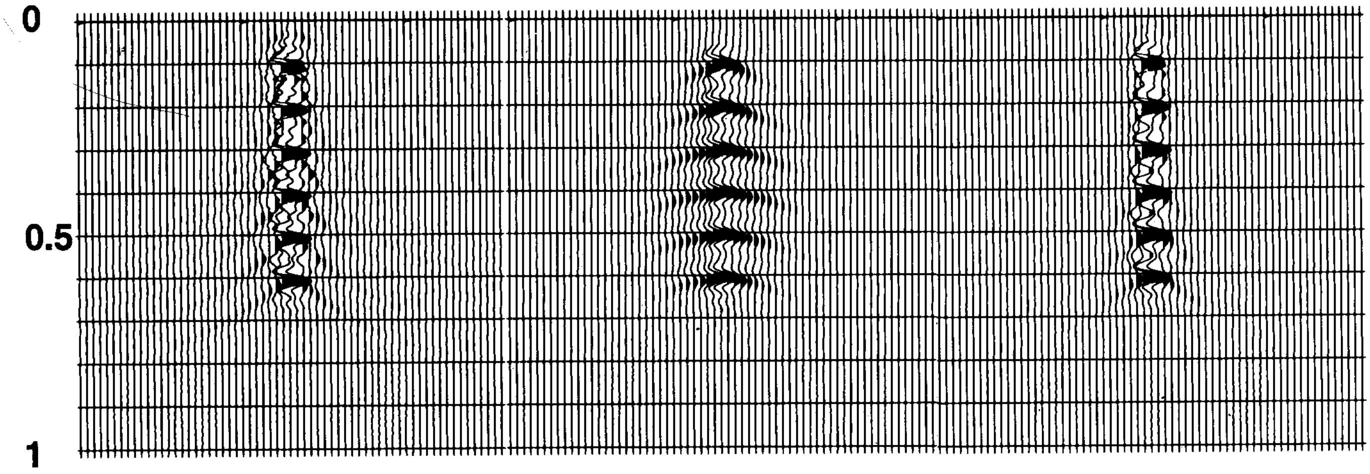
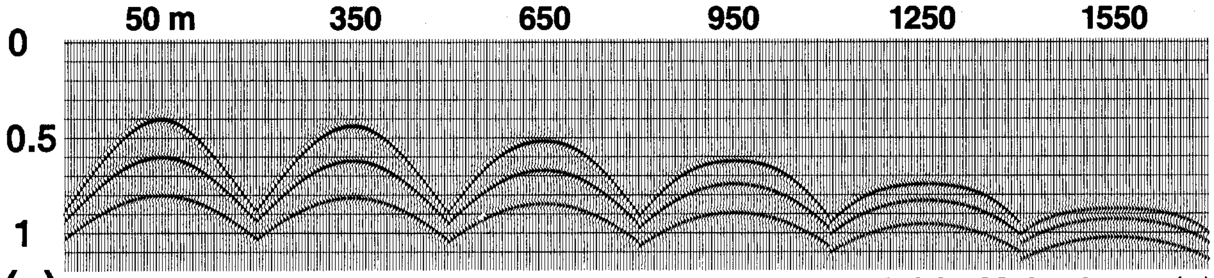
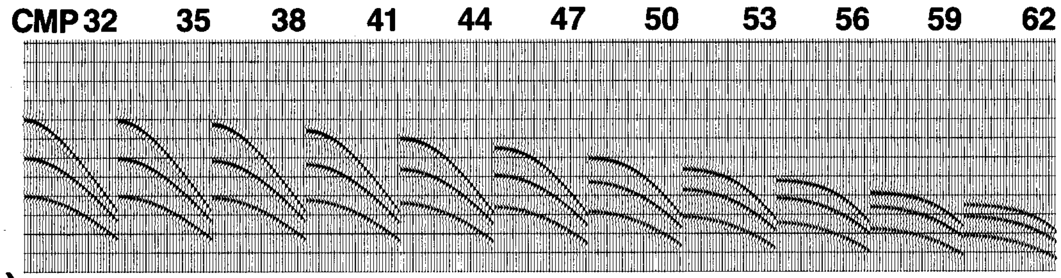


FIG. 4-122. Migrations of the sections in Figure 4-121.

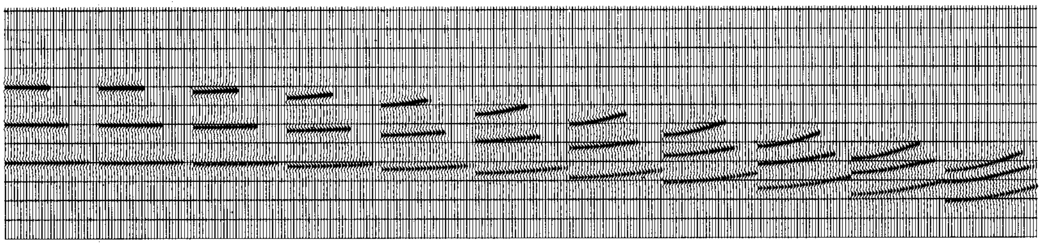


(a) Common Offset Sections

Variable Velocity, $v(z)$

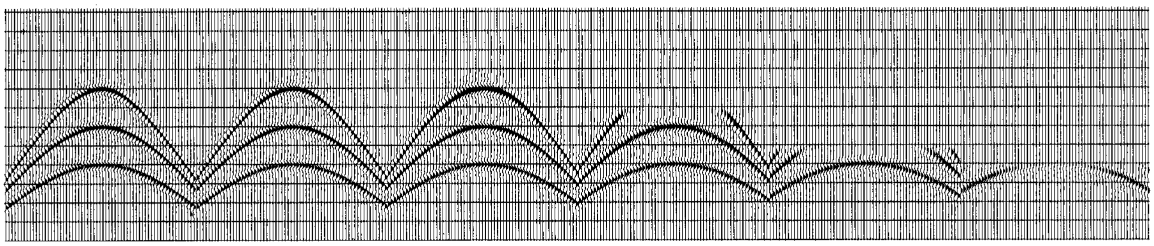


(b) CMP Gathers



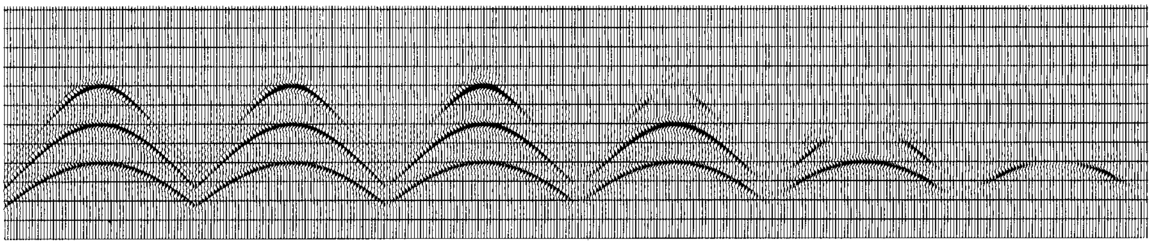
(c) CMP Gathers

NMO



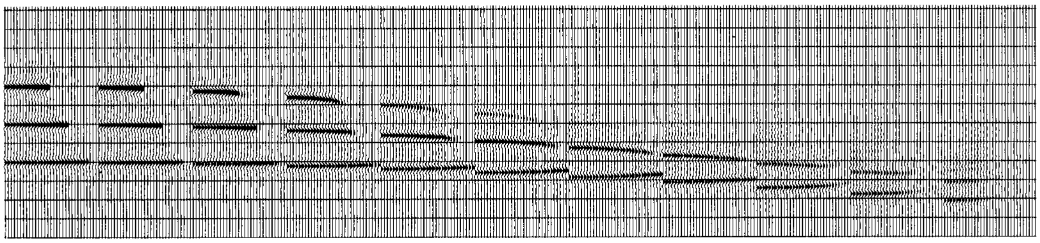
(d) Common Offset Sections

NMO



(e) Common Offset Sections

NMO + DMO



(f) CMP Gathers

NMO + DMO

◄FIG. 4-123. Intermediate results from DMO processing the nonzero-offset synthetic data derived from the depth model in Figure 4-61.

Figure 4-61. Despite the imperfect alignment of events after NMO and DMO corrections on the CMP gathers (Figure 4-123f), DMO has improved the stacking quality (compare Figures 4-124b and 4-124c). The misalignment at far offsets seen in Figure 4-123f can be attributed to the constant-velocity limitation of the Hale method. Hale (1983) showed that, provided the vertical velocity gradient is moderate, the constant-velocity DMO correction is adequate.

What happens if the NMO correction were applied with the wrong velocity? The DMO process requires an input that is NMO corrected using the medium velocity [equation (4.19)]. Thus, for field data, we try to pick a $v(z)$ function from the flattest part of the section for NMO correcting the data. The optimum stacking velocities are not used because they depend on dip. However, it is the stacking velocities that are picked from conventional velocity analyses. There is always the possibility that an accurate dip-independent velocity function will not be determined for NMO correcting the input data before DMO correction. The constant-velocity model in Figure 4-119 is used to examine this problem.

Assume that the velocity used for NMO correction is 20 percent higher than the velocity that should be used; that is, the medium velocity. Start with the CMP gathers in Figure 4-120b and apply the NMO correction using the incorrect velocity (3600 m/s). The results are shown in Figure 4-125a. Note the undercorrection at some gathers due to the high velocity used. By following the sequence described earlier, we get the results in Figure 4-125. Note that events no longer are aligned after the first NMO and DMO corrections (Figure 4-125d). Therefore, the stack obtained from these gathers is not expected to be any better than the conventional stack derived from the gathers in Figure 4-125a. The stacked sections are shown in Figure 4-126.

Perhaps the CMP stack can be improved by repicking the velocities after DMO correction. To test this idea, consider the following procedure. First, apply inverse NMO correction (Figure 4-125e) to the gathers with the velocity function that was used in the first NMO correction step (Figure 4-125a). Then, assuming we pick the correct velocity function, use it on the second NMO correction (Figure 4-125f). Significant improvement is seen when these gathers are stacked (Figure 4-127c). For a fair comparison, refer to the conventional stack with the 3000 m/s repicked velocity in Figure 4-127b. Although not shown, similar conclusions are reached from tests using velocities that are too low for NMO correction before DMO processing. From these experiments, we arrive at the flowchart in Figure 4-128 for DMO processing.

We now examine dipping events. Figure 4-129 shows the zero-offset section that consists of events with dips from 0 to 45 degrees with a 5-degree increment. Medium velocity is constant (3500 m/s). From the migrated section (Figure 4-130a), the dipping events lap onto the flat

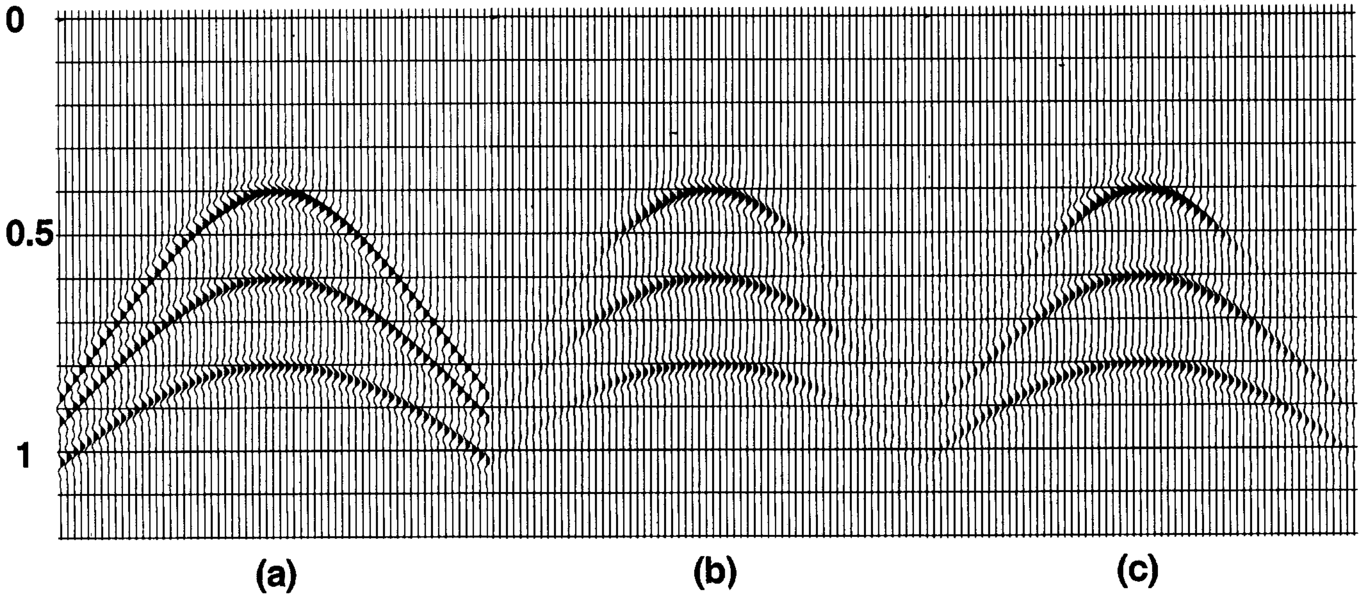
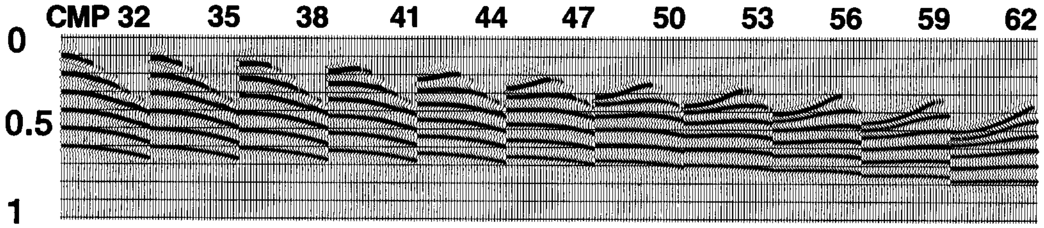
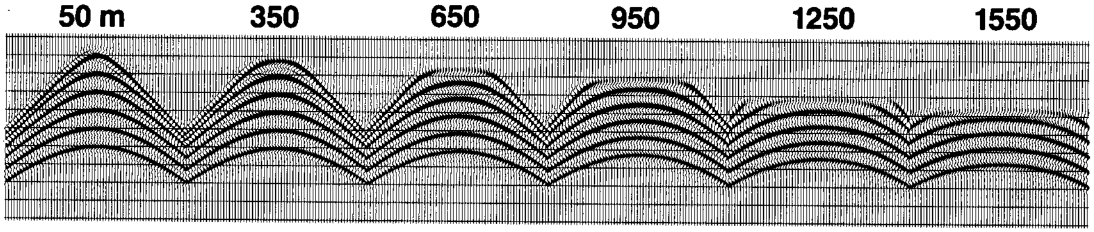


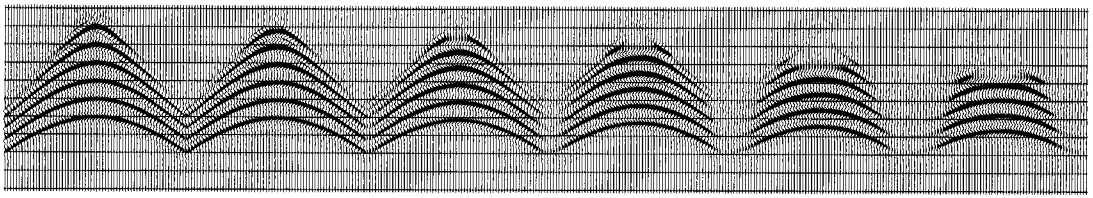
FIG. 4-124. (a) Zero-offset section associated with the depth model in Figure 4-61, (b) stack derived from the CMP gathers in Figure 4-123c, (c) DMO stack derived from the CMP gathers in Figure 4-123f.



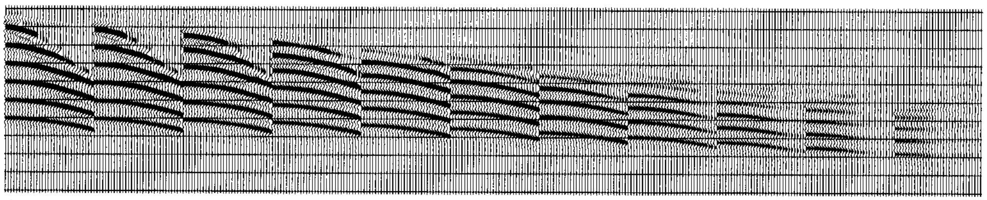
(a) CMP Gathers NMO$_1$ (3600 m/s)



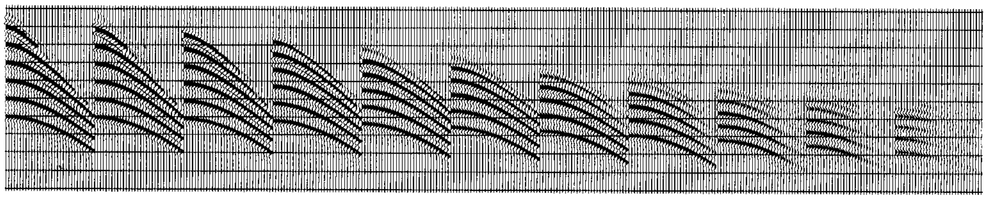
(b) Common Offset Sections NMO$_1$



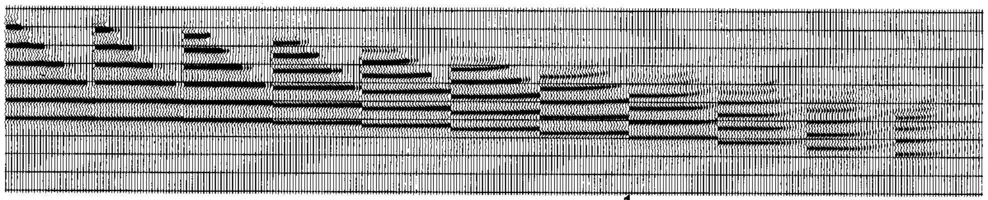
(c) Common Offset Sections NMO$_1$ + DMO



(d) CMP Gathers NMO$_1$ + DMO



(e) CMP Gathers NMO$_1$ + DMO + NMO$_1^{-1}$



(f) CMP Gathers NMO$_1$ + DMO + NMO$_1^{-1}$ + NMO$_2$ (3000 m/s)

FIG. 4-125. Intermediate results from DMO processing the nonzero-offset synthetic data derived from the depth model in Figure 4-119.

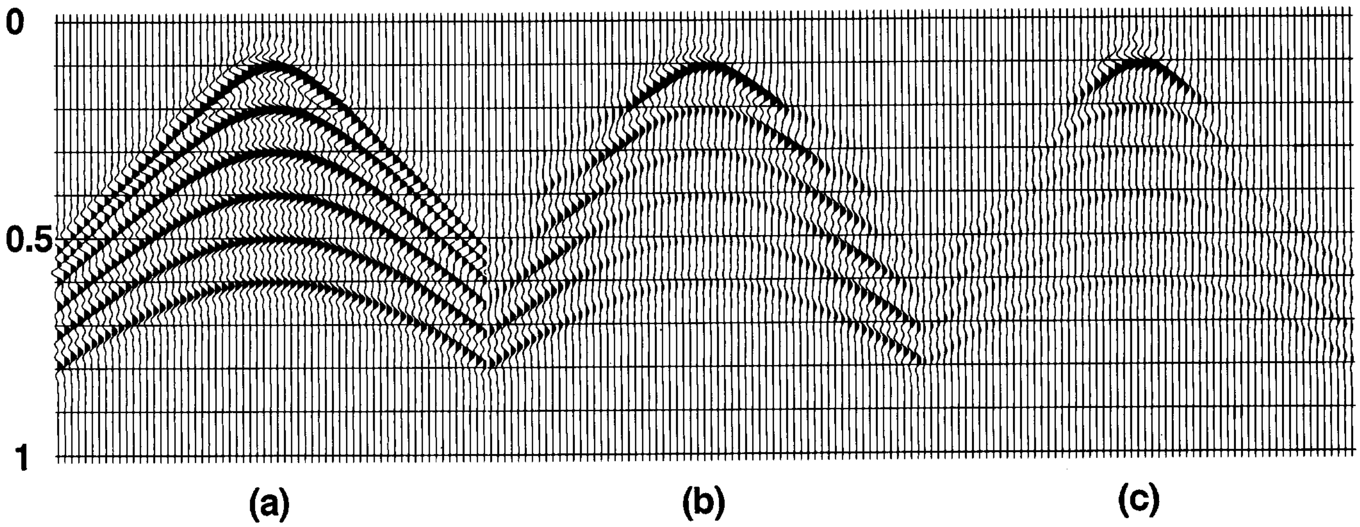


FIG. 4-126. (a) Zero-offset section associated with the depth model in Figure 4-119, (b) stack derived from the CMP gathers in Figure 4-125a, (c) DMO stack derived from the CMP gathers in Figure 4-125d.

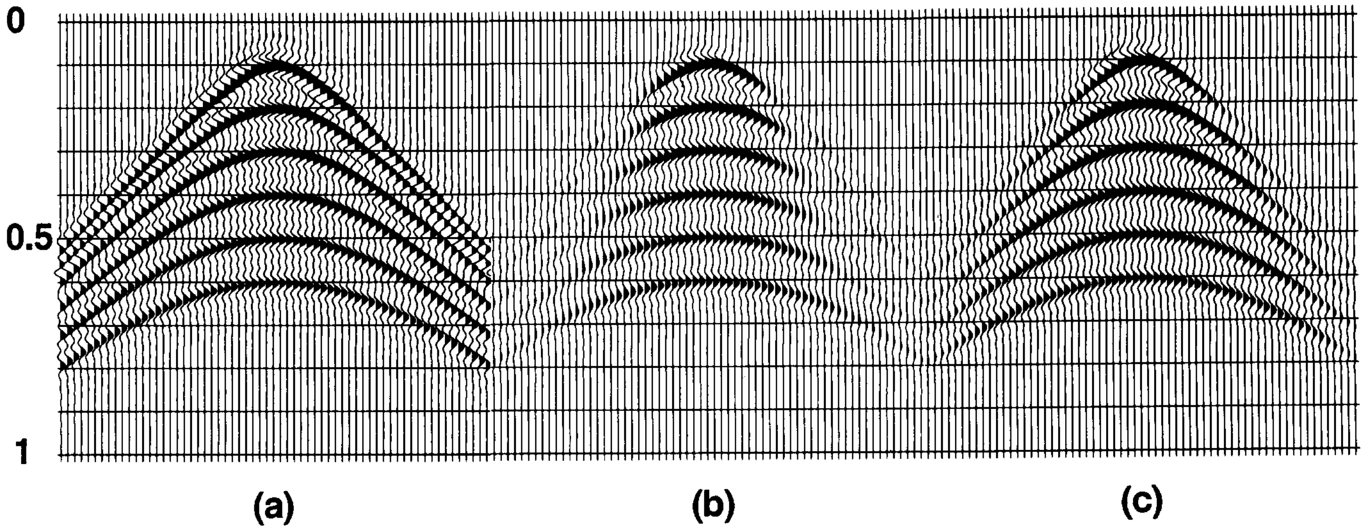


FIG. 4-127. (a) Zero-offset section associated with the depth model in Figure 4-119, (b) stack derived from the CMP gathers in Figure 4-120c, (c) DMO stack derived from the CMP gathers in Figure 4-125f.

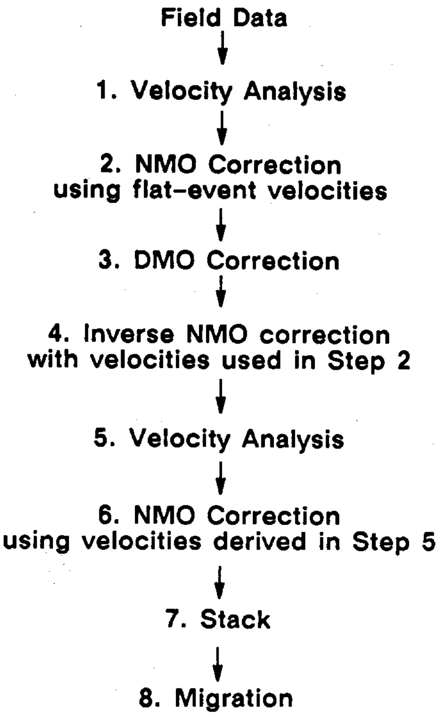


FIG. 4-128. DMO processing flowchart.

event, thus simulating pinchouts. Several velocity analyses were performed along the line; an example is shown in Figure 4-131a. Note the dip-dependent correlation peaks. By using the optimum stacking velocities picked from the densely spaced velocity analyses, we apply NMO correction to the CMP gathers, then stack them (Figure 4-129b). Aside from the conflicting dips at location *A*, stack response is close to the zero-offset section (Figure 4-129a). The DMO processing requires NMO correction using medium velocity. Stack response using the medium velocity (Figure 4-129c) clearly degrades at increasing dips. By applying DMO correction to the NMO-corrected gathers, we get the improved stacked section in Figure 4-129d. The DMO stack is closest to the zero-offset section (Figure 4-129a).

DMO also yields dip-corrected velocity functions that can be used in subsequent migration after stack. Refer to the velocity analysis in Figure 4-131b and note that all events have correlation peaks at 3500 m/s, which is the medium velocity for this model data set. Migrations of the sections in Figure 4-129 are shown in Figure 4-130. From the stacked sections (Figure 4-129) and their migrations (Figure 4-130), note that the improvement from DMO sometimes is marginal on CMP stacks.

Now return to the field data example in Figure 4-113. Following DMO correction, the duality in velocity picks at 2.35 s was eliminated; that is, the velocities were

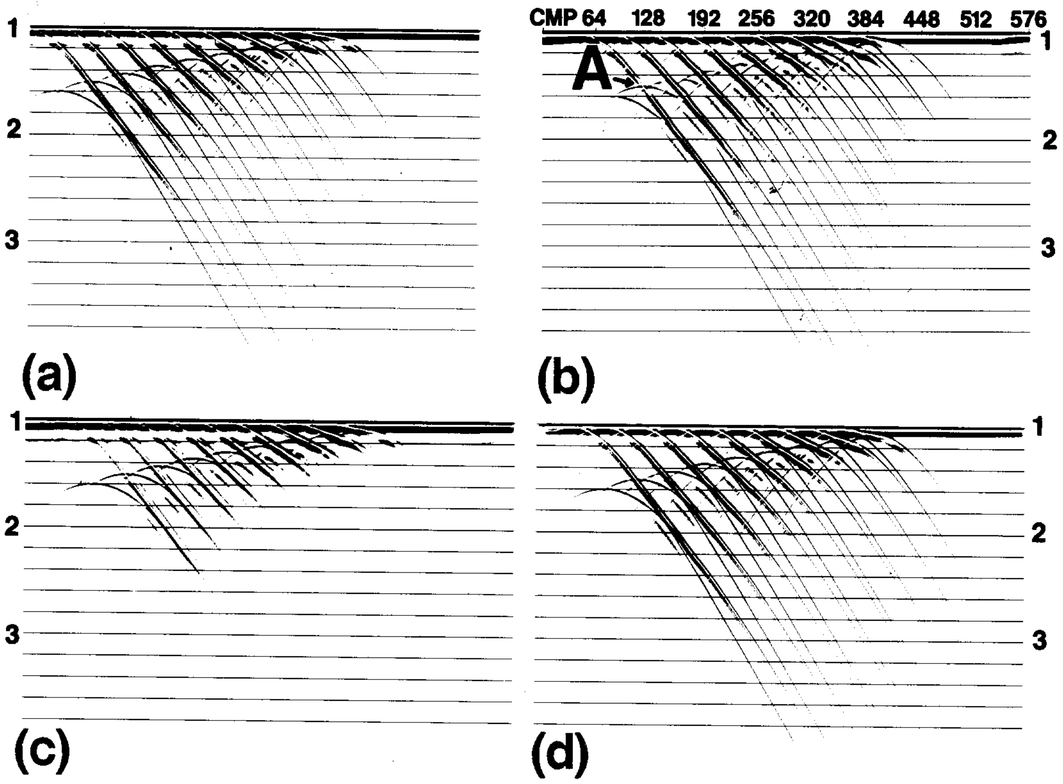


FIG. 4-129. DMO processing of dipping events. (a) Zero-offset section, medium velocity = 3500 m/s, (b) stack using optimum velocity picks (Figure 4-131a), (c) stack using medium velocity, (d) DMO stack.

corrected for dip, as shown in Figure 4-114b. The stacked section with PSPM processing is shown in Figure 4-132. It is the migration of this section (Figure 4-133) that should be compared with full migration before stack (Figure 4-117). With DMO processing, not only was a better stack obtained (compare Figures 4-113 and 4-132), but also velocities were dip-corrected (Figure 4-114b), resulting in a better migration after stack (compare Figures 4-118 and 4-133). The imaging quality is similar to full migration before stack (Figure 4-117).

From Levin's equation for the 3-D dipping plane interface, equation (3.9), note that stacking velocity not only is dip dependent, but also is azimuth dependent. The azimuth is measured by the angle between the source-receiver direction and the direction of dip line. Therefore, 3-D implementation of DMO should correct the stacking velocities for dip and azimuth (Jakubowicz et al., 1984).

The DMO process is not always without problems. Multiples can be enhanced by DMO (see Exercise 4.5). The constant-velocity assumption can even suppress

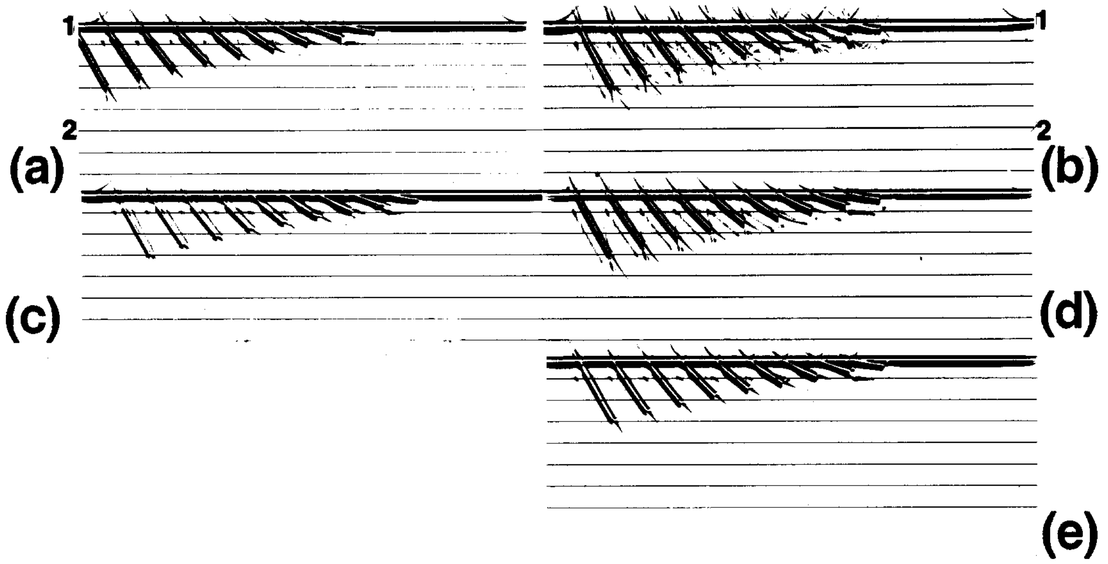


FIG. 4-130. Migrations (a) of the section in Figure 4-129a, (b) the section in Figure 129b, (c) the section in Figure 129c, (d) the section in Figure 129d, and (e) migration before stack of the dipping events model.

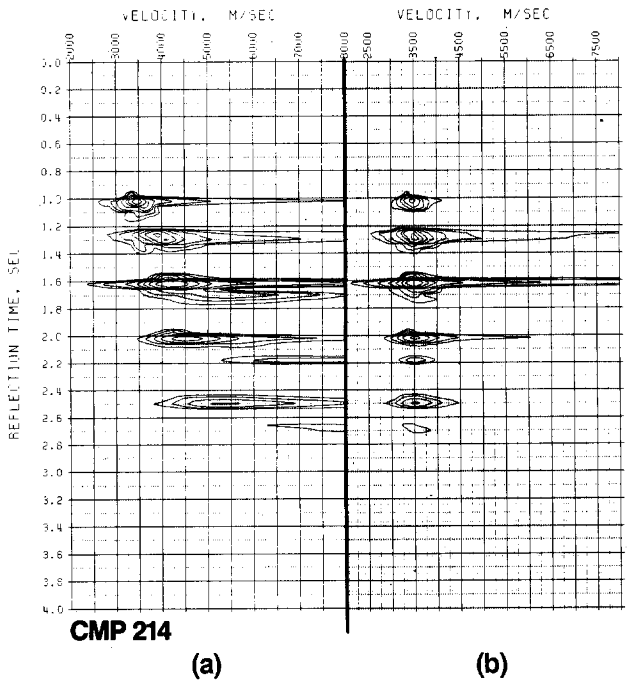


FIG. 4-131. Velocity analysis (a) before and (b) after DMO. The optimum stack is shown in Figure 4-129b.

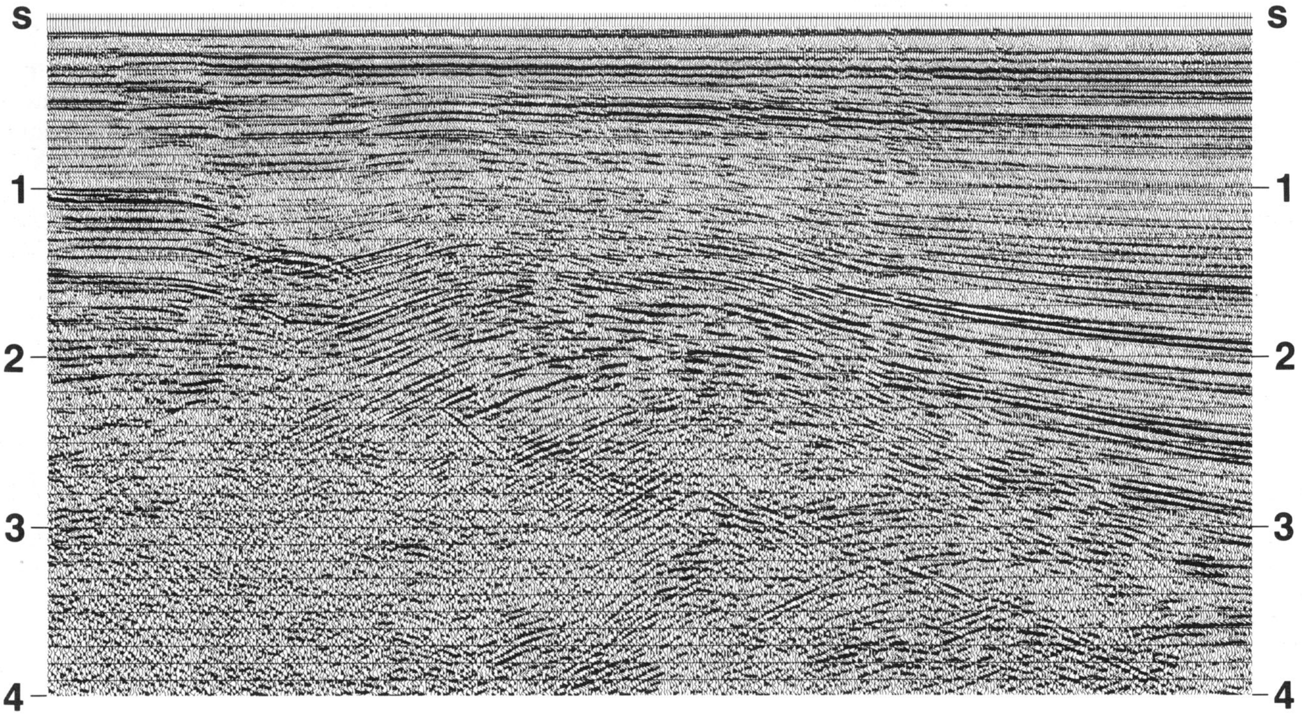


FIG. 4-132. CMP stack with PSPM; compare this section with Figure 4-113. (From a Western Geophysical brochure.)

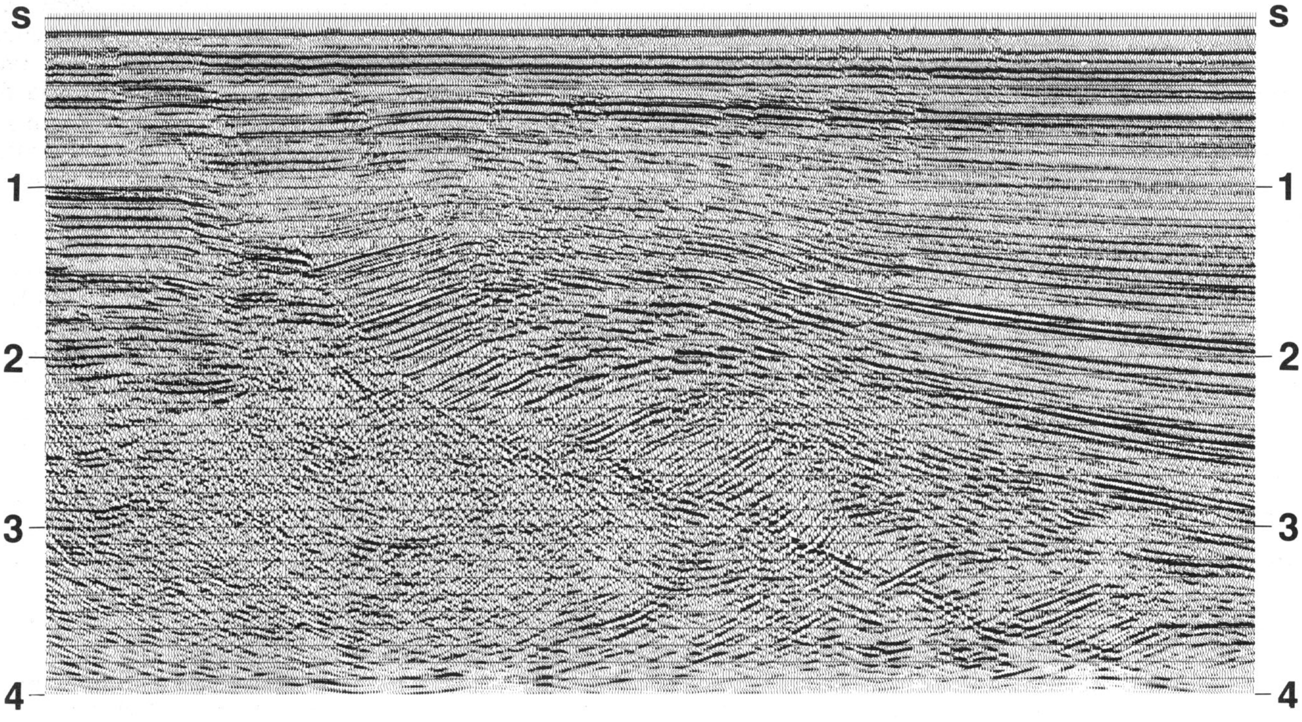


FIG. 4-133. Migration of the section in Figure 4-132. Compare this section with Figures 4-117 and 4-118. (From a Western Geophysical brochure.)

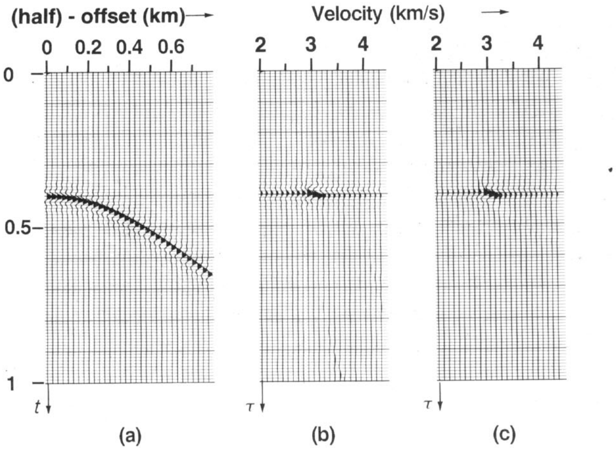


FIG. 4-134. (a) A CMP gather containing a single reflector in a constant-velocity medium, (b) velocity spectrum derived by migrating the CMP gather using a number of constant velocities and displaying the zero-offset trace from each migration, (c) velocity spectrum derived by NMO correcting and stacking the CMP gather using the same range of constant velocities as in (b). Each trace in (b) is a zero-offset trace, while each trace in (c) is a CMP stacked trace (Yilmaz and Chambers, 1984).

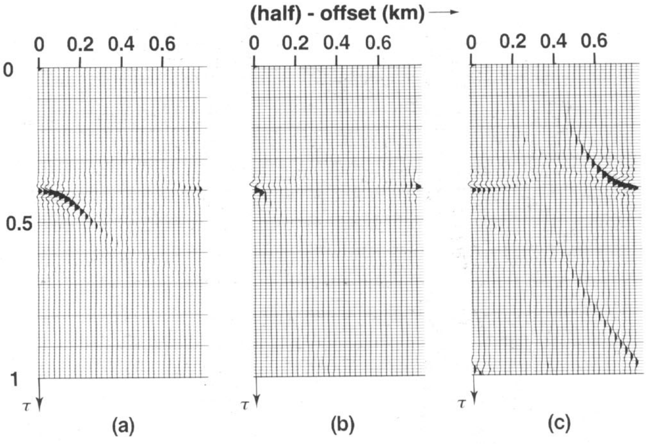


FIG. 4-135. Migrations of the CMP gather in Figure 4-134a using (a) lower than medium, (b) medium, and (c) higher than medium velocity. (Note the wraparound effect in (c) that is due to the phase-shift method used in this analysis.) (Yilmaz and Chambers, 1984).

dipping events. This is the case with a shallow dipping event and a deep flat event (Black et al., 1985). If velocity increases with depth (the usual case), then these two events can arrive about the same time and have similar moveouts. Following DMO application, nothing happens to the flat event, while the dipping event shifts to a lower velocity value. The dipping event drifts away from the velocity function for flat events and thus is suppressed during stacking. The constant-velocity assumption (Hale, 1983) is not made when formulating PSPM as Yilmaz and Claerbout (1980) and Sherwood et al. (1978) did. On the other hand, these methods have other limitations, such as dip or offset. Note that the rigorous solution to the conflicting dips problem is full time migration before stack (compare Figures 4-117 and 4-133). The PSPM (DMO) process can provide the opportunity to revise velocities for this rigorous solution.

4.5 MIGRATION VELOCITY ANALYSIS

Velocity estimation, CMP stacking, and migration generally are considered independent processes. However, they all have a common theoretical base: the scalar wave equation. Solution of this equation allows downward extrapolation of a seismic wave field recorded at the earth's surface. In turn, downward extrapolation provides a basis for CMP stacking and migration (Clayton, 1978; Yilmaz and Claerbout, 1980). Because the processes of CMP stacking and migration require velocity information, they also can be used to obtain a velocity estimate (Taner and Koehler, 1969; Gardner et al., 1974).

For example, consider the problem of conventional velocity estimation for stacking. Figure 4-134a shows a CMP gather over a single horizontal reflector. Select a constant velocity, apply NMO correction, and stack the traces in the gather. Next, place this stack trace onto the plane of velocity versus two-way zero-offset vertical time (v,τ) as shown in Figure 4-134c. By stacking the gather with different constant velocity values, the entire (v,τ) plane is filled with stacked amplitudes. (In this section, the variable τ is specifically used to define the time coordinate for moveout-corrected stacked data and the time equivalent of the wave field continuation depth.)

We now consider the migration process. For a horizontally stratified earth, as in Figure 4-134a, we cannot distinguish between a CMP gather and a common-shot gather (CSG). Moreover, since a CSG is a true wave field created by a single shot and recorded by many receivers, it seems reasonable that the CMP gather in Figure 4-134a can be migrated by treating the reflection hyperbola as if it were a diffraction hyperbola. Assuming there is no velocity information, we migrate with various trial velocity values and evaluate the results. Figure 4-135 shows three different attempts at migrating the CMP gather of Figure 4-134a. In one attempt (Figure 4-135a), too low a velocity was used; hence, the event was undermigrated. In another attempt, too high a velocity was used and the event was overmigrated (Figure 4-135c). When the velocity in the migration equals the medium velocity, we expect the diffraction hyperbola to collapse to its apex, which is at the zero-offset trace (Figure 4-135b).

What is the implication of this experiment for velocity estimation? Since the correct velocity produces a well-compressed event at the apex of the hyperbola, this velocity can be estimated by evaluating the quality of focusing at zero offset. To evaluate focusing, we pick out the zero-offset traces from the various attempts at migration with different velocities and place them side by side. This produces a display of velocity versus two-way zero-offset time as shown in Figure 4-134b.

When Figures 4-134b and 4-134c are compared, an almost identical character is noted. The resolution of velocity information obtained in the two approaches is equally degraded by data limitations such as maximum source-to-receiver offset and the absence of short-offset traces.

No distinct difference exists between migration and stacking velocity when the subsurface medium is horizontally layered as in Figure 4-134. However, for dipping reflectors, the two types of velocity differ. Stacking velocity is sensitive to the dip of the reflecting interface (Levin, 1971), while, in theory, migration velocity is independent of dip (Hubral and Krey, 1980). Therefore, for seismic migration, we must use a velocity field that is corrected for dips present in the data. As a result, any procedure that obtains velocities suitable for migration must use data from a number of neighboring CMP gathers.

The migration velocity analysis technique described in this section is based on wave field extrapolation. Velocity estimation is carried out from unstacked seismic data in midpoint-offset coordinates. This technique enables us to incorporate dip information that is essential for migration velocity estimation.

The idea of a velocity analysis that is based on differential solutions of the scalar wave equation first was introduced by Doherty and Claerbout (1974). They used the 15-degree finite-difference migration algorithm and worked with single CMP gathers. Gonzalez-Serrano and Claerbout (1979) later extended the wave equation velocity analysis to slant-midpoint coordinates and worked with linearly moveout-corrected CMP gathers. The method discussed here operates in the Fourier transform domain using the exact form of the double-square-root (DSR) operator (Yilmaz and Claerbout, 1980).

Mathematical details of the method presented here are found in Appendix C.7. Using equation (C.77), we map the downward-continued zero-offset amplitudes $P(y,h = 0,\tau,t)$ from each (t,τ) plane (the *image plane*) onto the corresponding $(v,\tau = t)$ plane, where τ is the two-way vertical zero-offset time. This mapping involves an approximation valid for small offsets and small dips. The velocity information is contained as amplitudes in the (y,τ,v) volume. In practice, we prefer to display the envelope function of the mapped amplitude since we are interested only in total signal energy and not phase. Here, the envelope is given by the running sum of squared amplitudes over a small time gate; e.g., 20 ms. Figure 4-136 summarizes the main computational steps involved in this migration velocity analysis based on wave field extrapolation.

To demonstrate the procedure outlined in Figure 4-136, we now perform some experiments using the DSR operator. Figure 4-137 shows two common-offset sections over a number of point scatterers buried in a constant-velocity earth, where $v = 3000$ m/s. Using a constant velocity for extrapolation, $v_e = 3000$ m/s, the image planes were produced for each midpoint. Two such planes corresponding to Midpoints 1 and 5 in Figure 4-137 are shown in Figure 4-138. The (v, τ) planes (Figure 4-139) then were generated from the image planes by the mapping procedure described in Appendix C.7. Peak amplitudes for all events occur at the correct medium velocity (3000 m/s). We expect the diffraction hyperbolas to migrate to the apexes beneath Midpoint 1, where the point scatterers are located. Note that in Figure 4-138, almost all the energy is in the image plane corresponding to Midpoint 1; just five midpoints away, at Midpoint 5, the migrated energy is very low.

How do we interpret the image planes? If we used the true medium velocity in downward extrapolation, then, according to the imaging principle, we would see all the events along the diagonal $\tau = t$, the *image line*, on the image plane. This happens in Figure 4-138, because a 3000 m/s extrapolation velocity is used, which is just the velocity used in generating the model in Figure 4-137.

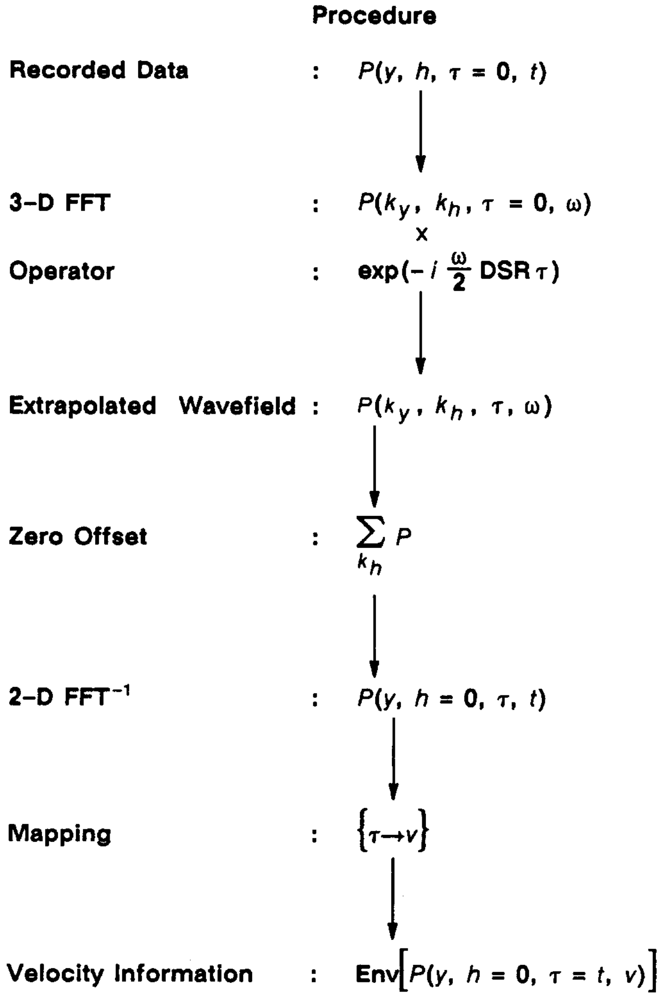


FIG. 4-136. Computational steps involved in the migration velocity analysis discussed in Section 4.5 (Yilmaz and Chambers, 1984). (See Appendix C.7 for the mathematical details.)

Any displacement of peak energy from the image line means that the velocity value used for downward extrapolation differs from that of the event. This displacement also is the basis for mapping from the image plane to the (v, τ) plane by equation (C.77).

This mapping is investigated further with the model in Figure 4-140, in which velocity increases with depth. In Figure 4-141b, note that the top and middle events fall to the left of the image line suggesting that the velocity used in extrapolation (v_e = 3000 m/s) is greater than the velocities associated with these events. The bottom event falls on the image line, implying that its associated velocity is nearly the same as the extrapolation velocity. These observations are confirmed in the corresponding (v, τ) planes in Figure 4-142. While true stacking velocity values for the three events are 2700, 2850, and 3000 m/s, the velocities interpreted from Figure 4-142b are about 2500, 2800, and 3000 m/s. Thus, the migration-based velocity estimate for the shallow event is in error by

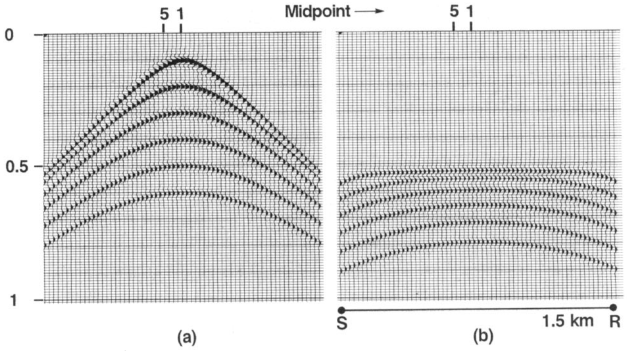


FIG. 4-137. Common-offset data derived from a constant-velocity earth model consisting of six point scatterers beneath midpoint 1, where (a) is zero-offset and (b) is far offset (Yilmaz and Chambers, 1984).

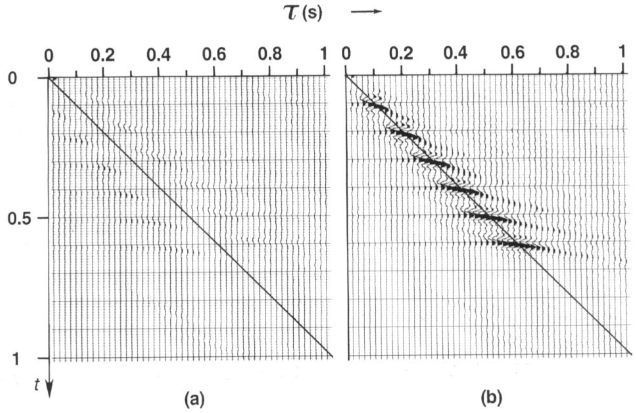


FIG. 4-138. Image planes corresponding to midpoints 1 and 5 as indicated in Figure 4-137, where (a) is CMP 5 and (b) is CMP 1 (Yilmaz and Chambers, 1984).

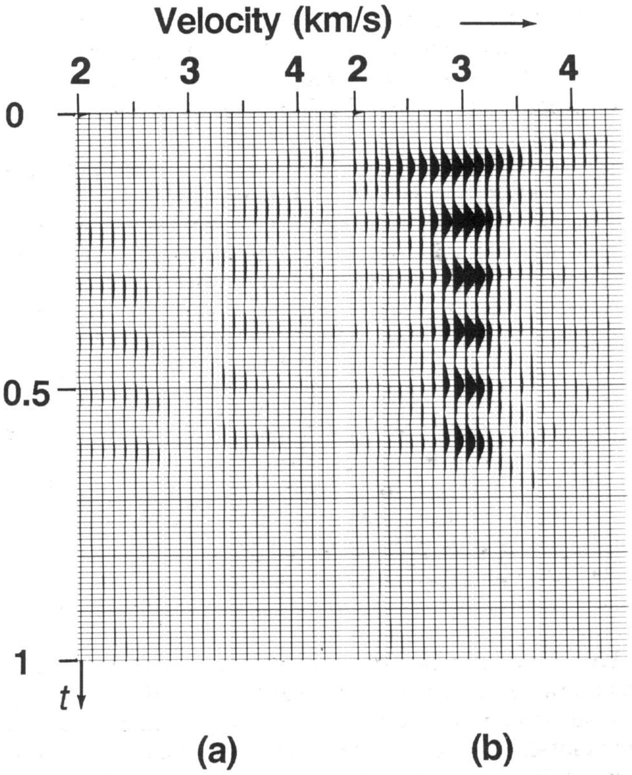


FIG. 4-139. The (v, τ) planes corresponding to midpoints 1 and 5 derived from the image planes in Figure 4-138, where (a) is CMP 5 and (b) is CMP 1 (Yilmaz and Chambers, 1984).

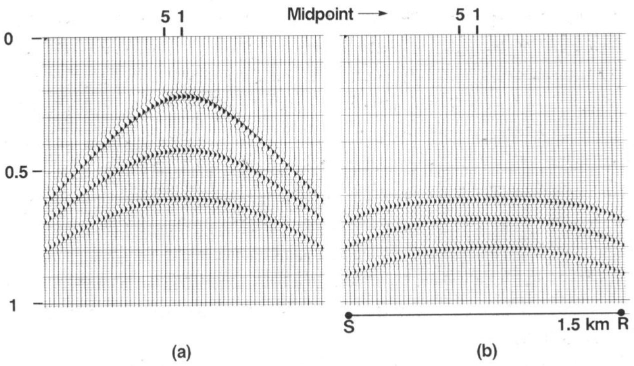


FIG. 4-140. Common-offset data based on a horizontally layered earth model containing three point scatterers located beneath midpoint 1 on the boundaries between constant-velocity layers, where (a) is zero-offset and (b) is far offset (Yilmaz and Chambers, 1984).

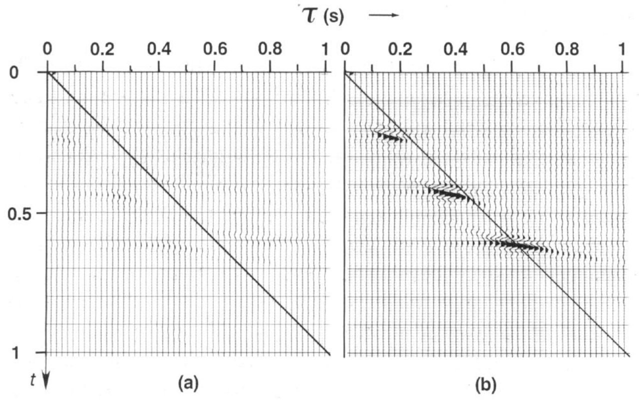


FIG. 4-141. Image planes corresponding to midpoints 1 and 5 as indicated in Figure 4-140, where (a) is CMP 5 and (b) is CMP 1 (Yilmaz and Chambers, 1984).

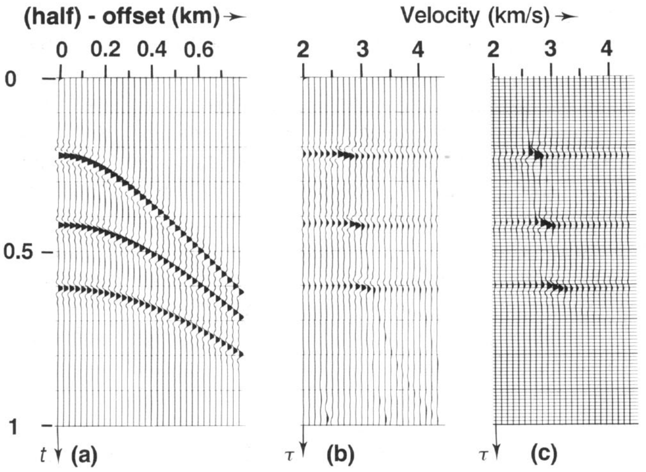


FIG. 4-143. (a) CMP gather at location 1 as indicated in Figure 4-140; (b) and (c) are velocity spectra derived from this gather by the methods of Figures 4-134b and 4-134c, respectively (Yilmaz and Chambers, 1984).

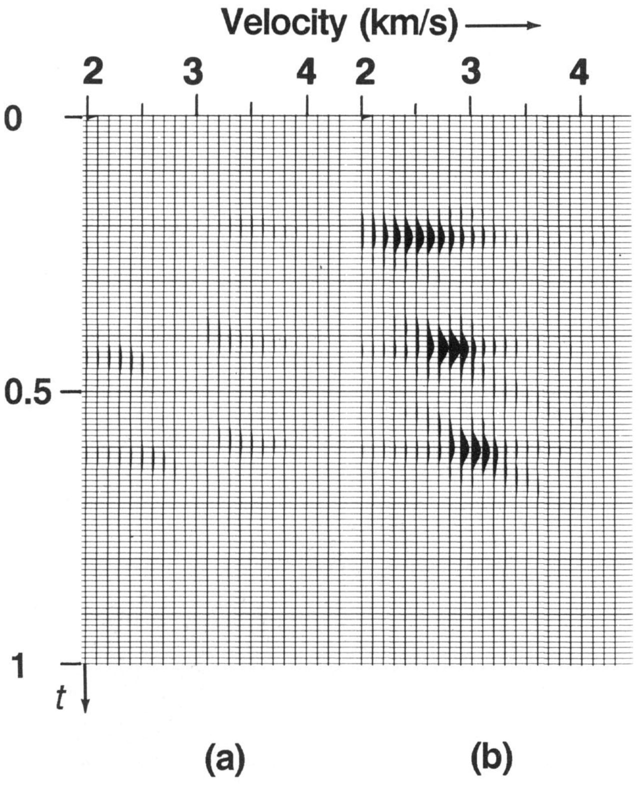


FIG. 4-142. The (v, τ) planes corresponding to midpoints 1 and 5 derived from the image planes in Figure 4-141, where (a) is CMP 5 and (b) is CMP 1 (Yilmaz and Chambers, 1984).

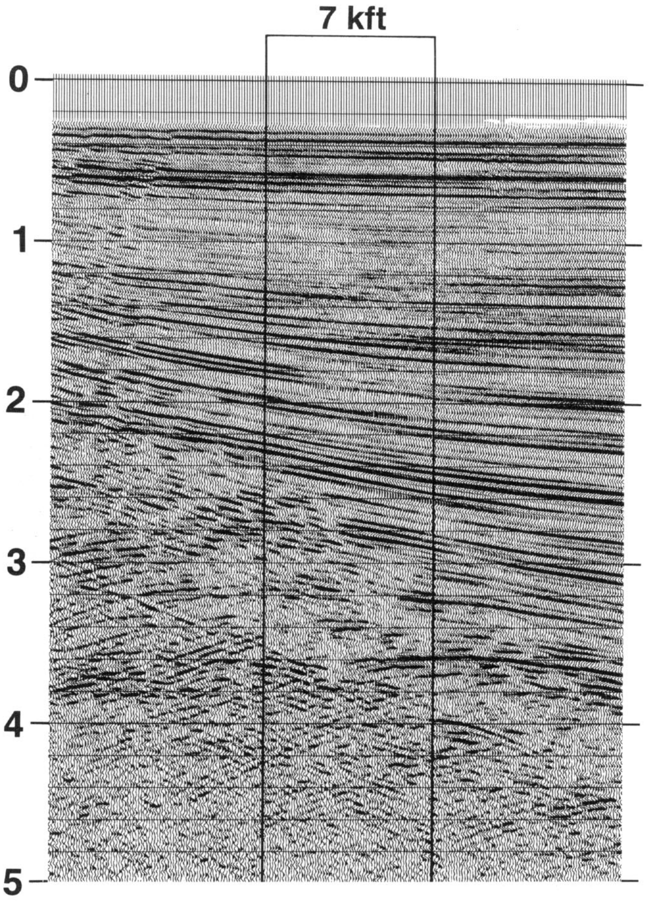


FIG. 4-144. A CMP stacked section. The center portion was used in the migration velocity analysis in Figures 4-145, 4-146, and 4-147 (Yilmaz and Chambers, 1984).

approximately 8 percent.

To determine the reason for the velocity error, we will consider a migration-based velocity analysis of our synthetic data example that does not involve the approximate mapping step. Figure 4-143a is a CMP gather from Midpoint 1 in the zero-dip region of the depth variable velocity model associated with the constant-offset sections in Figure 4-140. The migration velocity analysis on this gather (Figure 4-143b) was done by extrapolating the surface wave field $P(k_h, \omega, \tau = 0)$ repeatedly with different constant velocities in steps of $\Delta\tau = \Delta t$ (the sampling rate). The zero-offset trace from each attempt was collected after this effort, abandoning the rest of the migrated CMP gather.

Interpretation of the velocity analysis in Figure 4-143b reveals correct stacking velocities for the three events, including the shallowest. Clearly the error observed in Figure 4-142 is attributable to the mapping [equation (C.77)]. Note that the error does not occur because of depth variability of velocity, but instead, because the single extrapolation velocity used differed from the medium velocity.

The conventional velocity analysis for Midpoint 1 of this model is shown in Figure 4-143c for comparison. Note the familiar NMO stretching that is apparent in the shallow event. In other respects, both the results (Figures 4-143b and 4-143c) are comparable.

Field Data Example

Figure 4-144 is a CMP stack from offshore Texas. A 7000-ft portion (64 midpoints each with 48 offset traces) of the profile was used for migration velocity analysis. For computation efficiency, the data were windowed into 1024-ms time gates with 50 percent overlap. The image planes for one particular midpoint are shown in Figure 4-145. Different extrapolation velocities picked from a specified regional velocity function are used in each time gate. The velocity scan used in mapping then is carried out within a corridor around this function. Because different extrapolation velocities are used in successive segments, a given event appears at different values of τ in adjacent time segments. The step size in τ used in migration also varies from one segment to another, increasing from 24 ms for the shallow time gate to 80 ms for the deepest time gate.

The resulting velocity analysis for the central midpoint is shown in Figure 4-146b. The corresponding stacking velocity analysis based on the single CMP gather for that midpoint is shown in Figure 4-146c. The most obvious difference between the two results is the lack of shallow information in the migration-based (v, τ) plane. This shortcoming is attributed to spatial aliasing and lack of long-offset data in the shallow time gate. The problem can be partly eliminated by increasing the length of the time

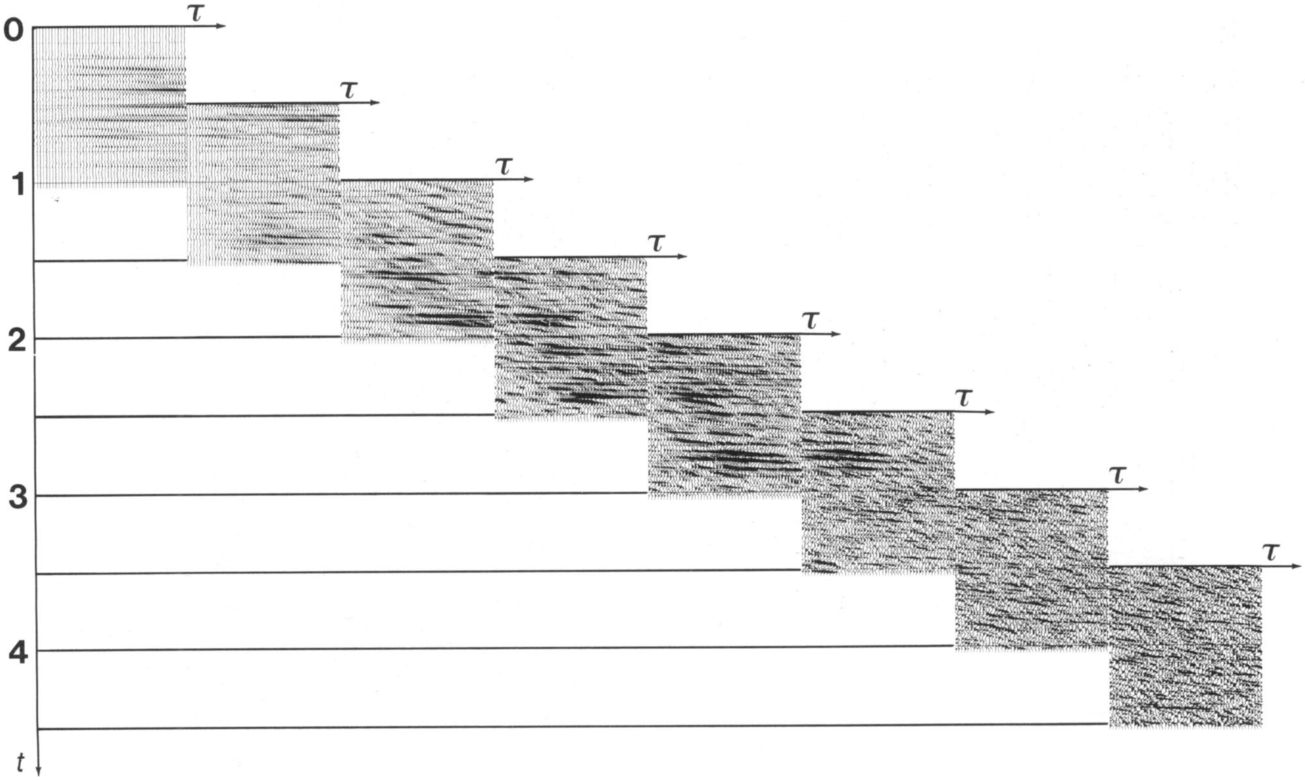


FIG. 4-145. Image planes associated with the center midpoint from the zone of interest in Figure 4-144 (Yilmaz and Chambers, 1984).

gate used in the velocity analysis. With the shortcut time-windowing approach described above, the shallowest time segment did not include the large-offset data necessary for velocity resolution.

In conventional practice, to improve the quality of velocity picks, velocity analyses from a number of neighboring CMP gathers often are summed. Figure 4-147c shows the result of stacking velocity analysis for data from the six adjacent CMP gathers indicated in Figure 4-147a. For the migration-based method, the (v, τ) planes corresponding to these gathers were summed. The result is shown in Figure 4-147b. Note the improvement in quality when compared to the results from the single CMP gathers shown in Figures 4-146b and 4-146c. Because the events have dip, the derived migration velocities are lower (by up to 4.5 percent) than the velocities derived from the stacking velocity analysis.

The velocity analysis described in this section does not handle lateral variations in velocity. It is based on a Fourier-transform domain formulation with only vertically varying velocity used in extrapolation. This method may be particularly efficient for the dip-corrected velocity estimate needed for time migration.

REFERENCES

Baysal, E., Kosloff, D.D., and Sherwood, J.W.C., 1983, Reverse time migration: Geophysics, **48**, 1514-1524.

Berkhout, A.J., 1980, Seismic migration—Imaging of acoustic energy by wave field extrapolation: Elsevier Science Publ. Co., Inc.

Berryhill, J.R., 1979, Wave-equation datuming: Geophysics, **44**, 1329-1333.

Black, J.L., McMahon, I.T., Meinardus, H., and Henderson, I., 1985, Applications of prestack migration and dip moveout: Paper presented at the 55th Ann. Int. Soc. Explor. Geophys. Mtg.

Chun, J.H. and Jacewitz, C., 1981, Fundamentals of frequency-domain migration: Geophysics, **46**, 717-732.

Claerbout, J.F., 1976, Fundamentals of geophysical data processing: McGraw-Hill Book Co.

Claerbout, J.F., 1985, Imaging the earth's interior: Blackwell Scientific Publications.

Claerbout, J.F. and Doherty, S.M., 1972, Downward continuation of moveout-corrected seismograms: Geophysics, **37**, 741-768.

Clayton, R., 1978, Common midpoint migration: Stanford

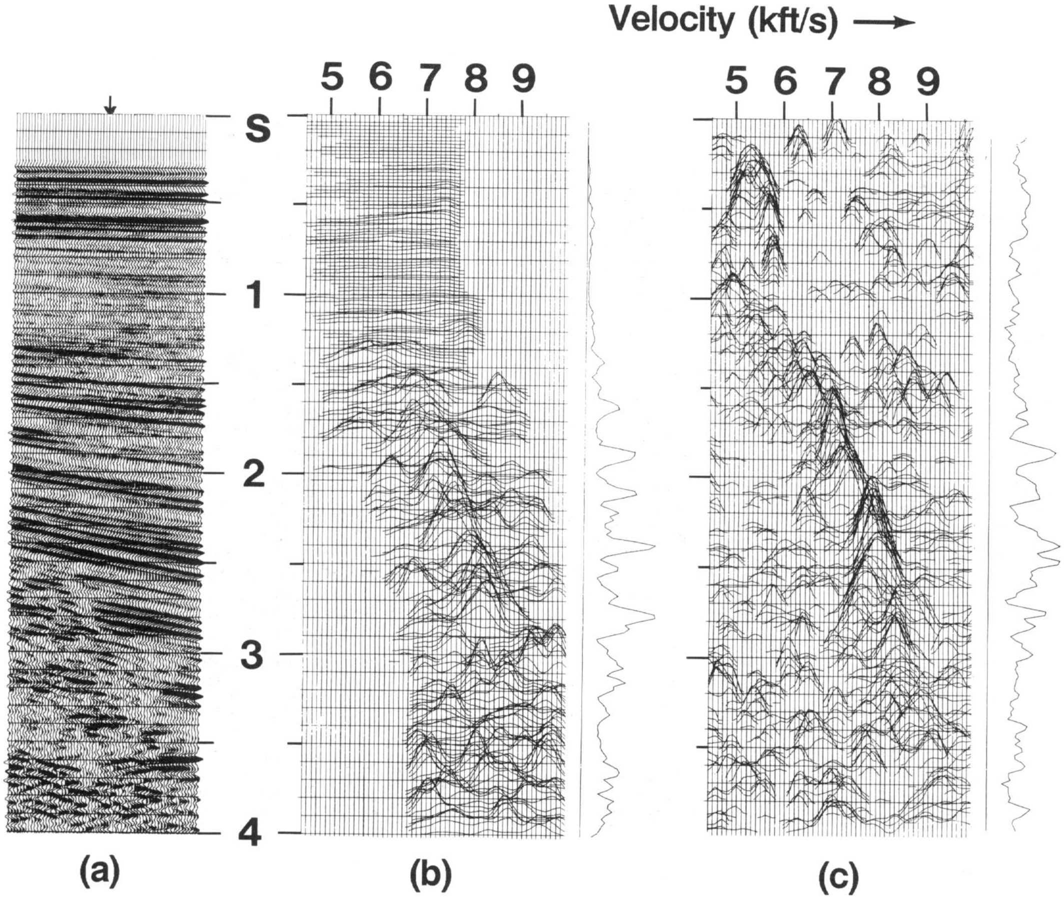


FIG. 4-146. (a) A portion of the CMP stacked section in Figure 4-144 that is under study, (b) the velocity spectrum based on the procedure in Figure 4-136 using the image planes in Figure 4-145, (c) the conventional velocity analysis as discussed in Section 3.1 (Yilmaz and Chambers, 1984).

Exploration Project Report No. 14, Stanford University.

Clayton, R. and Engquist, B., 1980, Absorbing side boundary conditions for wave-equation migration: Geophysics, **45**, 895-904.

Deregowski, S. and Rocca, F., 1981, Geometrical optics and wave theory for constant-offset sections in layered media: Geophys. Prosp., **29**, 374-387.

Diet, J.P. and Lailly, P., 1984, Choice of scheme and parameters for optimal finite-difference migration in 2-D: Paper presented at the 54th Ann. Int. Soc. Explor. Geophys. Mtg.

Doherty, S.M., 1975, Structure independent seismic velocity estimation: Ph.D. thesis, Stanford University.

Doherty, S.M. and Claerbout, J.F., 1974, Velocity analysis based on the wave equation: Stanford Exploration Project Report No. 1, Stanford University.

Fowler, P., 1984, Velocity-independent imaging of seismic reflectors: Presented at the 54th Ann. Int. Soc. Explor. Geophys. Mtg.

French, W.S., Perkins, W.T., and Zoll, R.M., 1984, Partial migration via true CRP stacking: Presented at the 54th Ann. Int. Soc. Explor. Geophys. Mtg.

Gardner, G.H.F., French, W.S., and Matzuk, T., 1974, Elements of migration and velocity analysis: Geophysics, **39**, 811-825.

Gazdag, J., 1978, Wave-equation migration by phase shift: Geophysics, **43**, 1342-1351.

Gazdag, J. and Squazzero, P., 1984, Migration of seismic data by phase shift plus interpolation: Geophysics, **49**, 124-131.

Gonzalez-Serrano, A. and Claerbout, J.F., 1979, Wave-equation velocity analysis: Stanford Exploration Project Report No. 16, Stanford University.

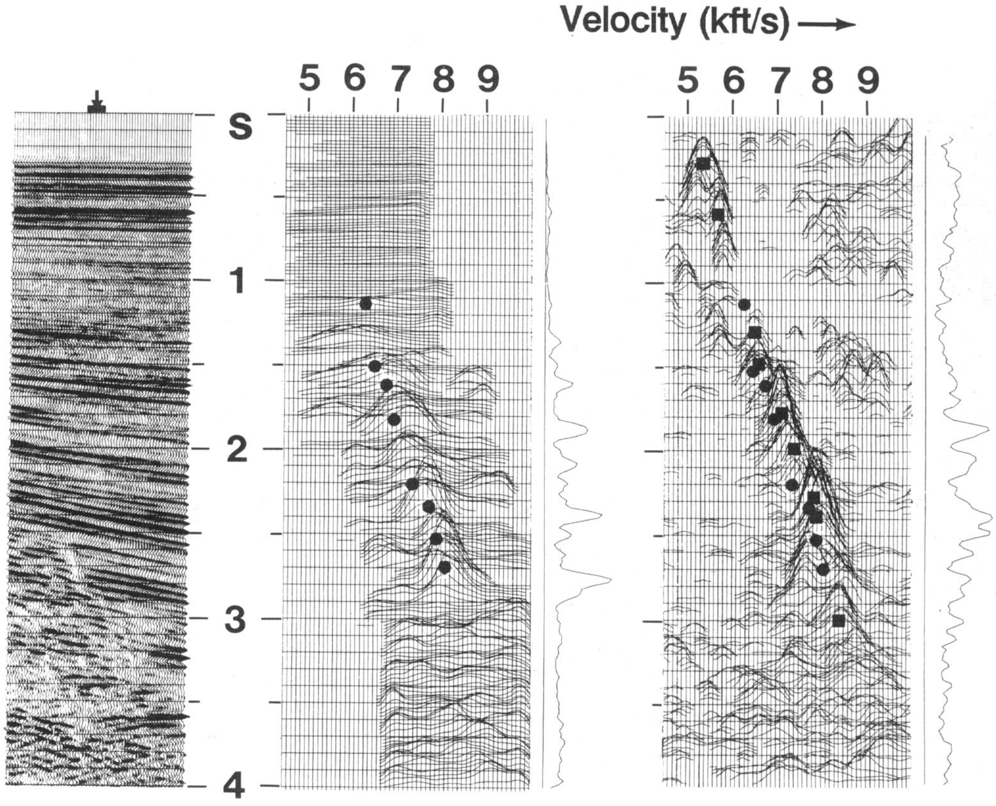
Hale, I.D., 1983, Dip moveout by Fourier transform: Ph.D. thesis, Stanford University.

Hubral, P. and Krey, T., 1980, Interval velocities from seismic reflection time measurements: Soc. Expl. Geophys. Monograph.

Jakubowicz, H., Beasley, C., and Chambers, R., 1984, A comprehensive solution to problems in processing of 3-D data: Paper presented at the 54th Ann. Int. Soc. Explor. Geophys. Mtg.

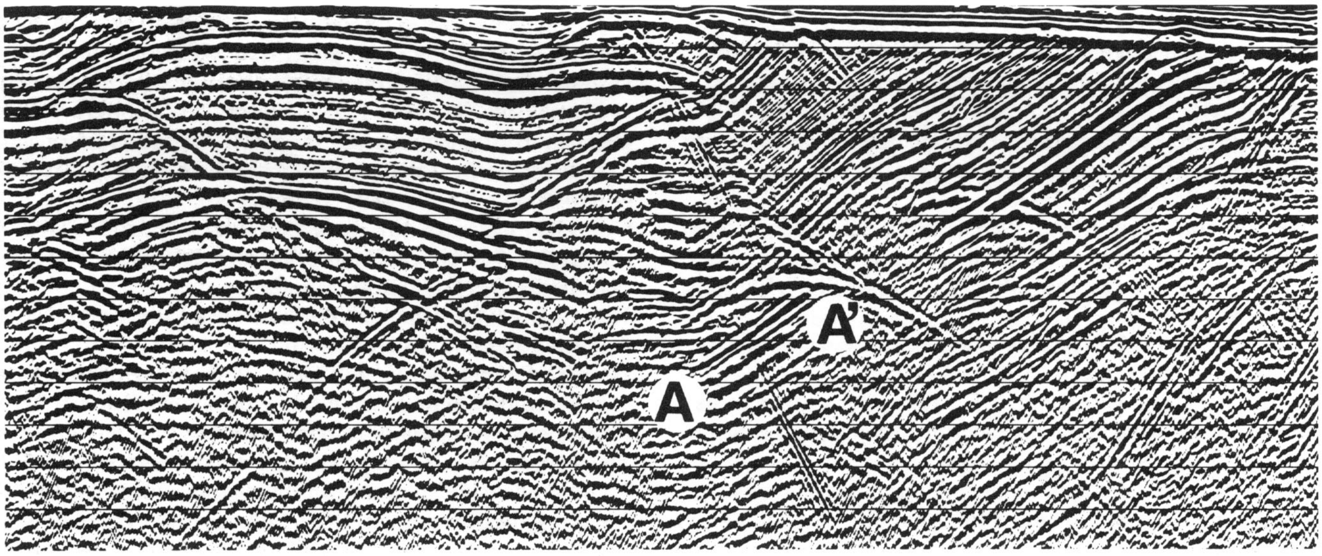
Kjartansson, E., 1979, Modeling and migration by the monochromatic 45-degree equation: Stanford Exploration Project Report No. 15, Stanford University.

Kjartansson, E. and Rocca, F., 1979, The exploding reflec-

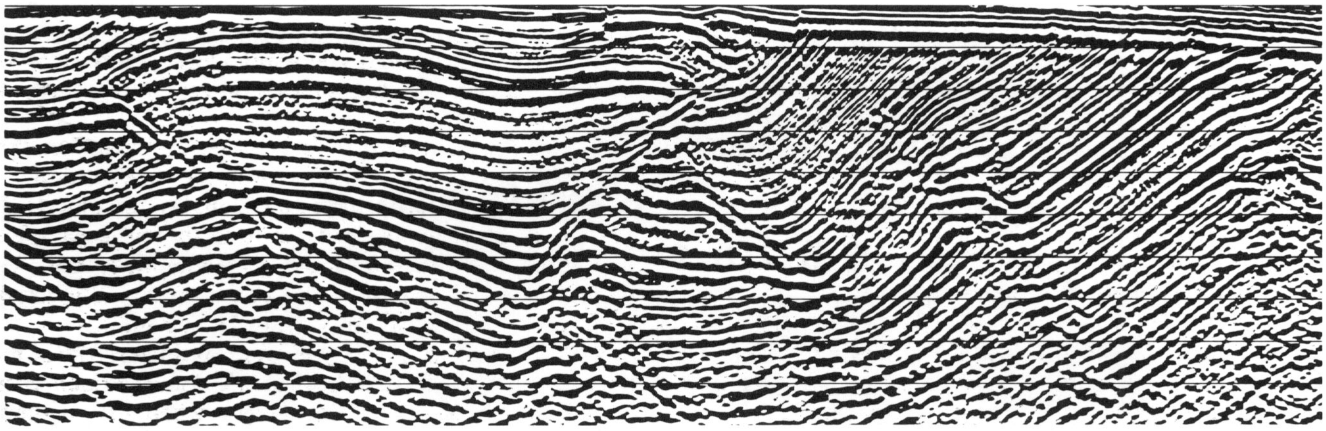


(a) (b) (c)

FIG. 4-147. (a) A portion of the CMP stacked section in Figure 4-144 that is under study; (b) the velocity spectrum based on the procedure in Figure 4-136, followed by averaging over six midpoints; (c) the conventional velocity analysis as discussed in Section 3.1 followed by averaging over six midpoints (Yilmaz and Chambers, 1984).



Stack



Migration

FIG. 4-148. Locate dipping event *AA′* on the migrated section (Exercise 4.8).

tor model and laterally variable media: Stanford Exploration Project Report No. 16, Stanford University.

Lee, M.W. and Suh, S.H., 1985, Optimization of one-way wave equations: Geophysics, **50**, 1634-1637.

Levin, F.K., 1971, Apparent velocity from dipping interface reflections: Geophysics, **36**, 510-516.

Loewenthal, D., Lu, L., Roberson, R., and Sherwood, J.W.C., 1976, The wave equation applied to migration: Geophys. Prosp., **24**, 380-399.

Ma, Z., 1981, Finite-difference migration with higher-order approximation: Presented at the 1981 Joint Mtg. Chinese Geophys. Soc. and Soc. Explor. Geophys.

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

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

Rothman, D., Levin, S., and Rocca, F., 1985, Residual migration: Applications and limitations: Geophysics, **50**, 110-126.

Schneider, W., 1978, Integral formulation for migration in two and three dimensions: Geophysics, **43**, 49-76.

Sherwood, J.W.C., Schultz, P.S., and Judson, D.R., 1978, Equalizing the stacking velocities of dipping events via Devilish: Presented at the 48th Ann. Int. Soc. Explor. Geophys. Mtg.

Stolt, R.H., 1978, Migration by Fourier transform: Geophysics, **43**, 23-48.

Stolt, R.H. and Benson, A.K., 1986, Seismic migration: Theory and practice, Geophysical Press.

Taner, M.T. and Koehler, F., 1969, Velocity spectra—digital computer derivation and applications of velocity functions: Geophysics, **32**, 859-881.

Taner, M.T. and Koehler, F., 1977, Wave-equation migration: Technical brochure, Seiscom-Delta, Inc.

Yilmaz, O., 1983, Choice of optimal depth step size in finite difference migration: Technical document, Western Geophysical Company.

Yilmaz, O. and Chambers, R., 1984, Migration velocity analysis by wave field extrapolation: Geophysics, **49**, 1664-1674.

Yilmaz, O. and Claerbout, J.F., 1980, Prestack partial migration: Geophysics, **45**, 1753-1777.

EXERCISES

Exercise 4.1. Consider the special case of a 90-degree dipping reflector in Figure 4-14a. Sketch the corresponding zero-offset time section.

Exercise 4.2. Measure the apparent dip ($\Delta t/\Delta x$) on the steepest reflector on the zero-offset section in Figure 4-16b at 1.5 s. Using the relations given by equations (4.1), (4.2), and (4.3), compute the horizontal and vertical displacements due to migration. Next compute the angle after migration. Verify the results by comparing them with the migrated section in Figure 4-16a. Point *A* should migrate to point *A'*.

Exercise 4.3. Consider Huygens' secondary sources along a dipping reflector. Sketch the zero-offset section by superimposing the individual responses from these sources. Remember, to do zero-offset modeling, you must map a point in the (x,z) plane along a hyperbola in the (x,t) plane with its apex as the input point.

Exercise 4.4. Are the traveltime trajectories on the overmigrated sections in Figure 4-50 hyperbolic?

Exercise 4.5. Examine the following cases in terms of the effect of DMO on multiples. Check the table where appropriate.

The effect of DMO is:

	To Suppress Multiples	To Enhance Multiples	No Difference
In the presence of:			
Flat Primary Flat Multiple			
Dipping Primary Flat Multiple			
Flat Primary Dipping Multiple			
Dipping Primary Dipping Multiple			

Exercise 4.6. Which is more sensitive to velocity errors, stack or migration?

Exercise 4.7. For which case is spatial aliasing a more serious problem, the low-velocity or high-velocity medium?

Exercise 4.8. Locate the dipping event *AA'* on the migrated section in Figure 4-148.

Exercise 4.9. A point in the (x,t) plane is mapped onto a semicircle in the (x,z) plane. Where does it map onto the (x,τ) plane, where $\tau = 2z/v$?

Exercise 4.10. Derive equations (C.38a) and (C.38b) in Appendix C.2.

Exercise 4.11. Refer to Figure 4-26. It suggests that if the subsurface consisted of a semicircular reflector (b), then the zero-offset response would be as in (a). What would the subsurface be like if you obtained (a) using a source-receiver pair with a finite separation between them? (Hint: see Figure C-2.)

Exercise 4.12. Suppose you specified the wrong trace spacing in your migration job. What effect does it have, overmigration or undermigration? Assume that you supplied the wrong sampling rate (in time). What effect does it have on migration output?

Exercise 4.13. Should you apply DMO before or after multiple suppression that is based on velocity discrimination?

Exercise 4.14. Refer to Figure 4-33. To what does the energy denoted by C correspond on the time section?

Exercise 4.15. Derive the expressions for horizontal and vertical displacements d_x and d_z) and the dip angle after migration ($\bar{\theta}_z$) as seen on the migrated depth section. To do this refer to Figure 4.14 and substitute $t = 2z/v$, $\tan\theta_t = (2\tan\theta_z)/v$ in equation (4.1), (4.2), and (4.3).

Exercise 4.16. Suppose you want to do zero-offset recording of the steep flank of a salt dome. Which case would require a longer line length when the medium velocity along the raypath is (a) constant or (b) vertically increasing?

Exercise 4.17. How would Figure 4-59 look if you used a 15-degree phase-shift migration algorithm?

5

Imaging Beneath Complex Structures

5.1 INTRODUCTION

Although the migration methods discussed in Chapter 4 are based on a layered media assumption, simple modifications of the basic algorithms make them accurate for situations with mild lateral velocity variations. For example, rms velocities can be varied laterally in Kirchhoff migration. In the finite-difference method, as long as lateral velocity variations are mild, the thin-lens term can be dropped (see Appendix C.2), and the velocity function used in the diffraction term can be varied laterally. In the *f-k* method (Stolt migration), lateral velocity variations are accommodated by varying the stretch factor between 0 and 1. Even when velocity varies, the output from these three methods is still a time section, thus, the term *time migration*.

The situation is different when strong lateral velocity variations are encountered. In that case, simple algorithmic modifications no longer provide adequate accuracy, and depth migration (Judson et al., 1980), rather than time migration, must be done. While both migration types use a diffraction term for collapsing energy along a diffraction hyperbola to its apex, only depth migration algorithms implement the additional thin-lens term that explicitly accounts for lateral velocity variations. Unlike time migration, the output from depth migration is a depth section. For geologically meaningful output, the velocity model must be more accurate for a depth migration than for a time migration. In the next section, the distinction between time and depth migrations is demonstrated using velocity-depth models with varying degrees of complexity.

Strong lateral velocity variations often are associated with a complex overburden structure. An example is the imbricate structures in folded belts involving both Paleozoic and younger rocks. Strong lateral velocity variations also are associated with salt diapirism. The imaging of target horizons beneath the salt layer is complicated by raypath distortions from the complex overburden. Another type of geologic environment with strong lateral velocity variations is an irregular water bottom. Strong lateral velocity variations also are found in regions with strong lateral facies changes. For example, a lithologic change from dolomite to limestone to evaporite to clastics may be associated with a large change in velocity in the horizontal direction.

Complex structures often are three dimensional. In this chapter, we will assume that the seismic line is along the dip direction and that the recorded wave field is two dimensional. The validity of this assumption will be examined in Section 6.5.

Time or depth migration after stack can produce a plausible geologic image of the subsurface, provided conventional CMP stack closely represents a zero-offset section. This is not the case in situations with either (a) conflicting events with different stacking velocities or (b) strong lateral velocity variations. For the first problem, although events with conflicting dips have hyperbolic moveout on CMP gathers, the events may not be optimally stacked with one velocity. The rigorous solution is to do time migration before stack. For the second problem, complex, nonhyperbolic moveout usually is associated with the reflections beneath a complex overburden with lateral velocity variations. The rigorous solution in this case is to do depth migration before stack.

The first problem was discussed in Section 4.4.1, where we saw that NMO followed by prestack partial migration (DMO), CMP stack, and time migration after stack largely is equivalent to doing a time migration before stack. In Section 5.3.2, we find that the second problem can be addressed by prestack layer replacement followed by NMO, CMP stack, and time migration after stack (with depth conversion) since this sequence basically is equivalent to doing a full depth migration before stack. In both cases, the approaches are based on the philosophy of revising velocity estimates and obtaining an improved unmigrated stacked section. Applications of these alternative approaches do have limitations. In certain cases, prestack partial migration (DMO) can enhance multiples (Section 4.4.1). Prestack layer replacement may not be practical when the complex overburden involves more than one velocity layer.

Besides DMO and prestack layer replacement, another way to improve the unmigrated stacked section is to do a horizon velocity analysis (HVA) along key horizons (Section 3.3.3). Sometimes limiting the offset range in stacking to small offsets also can produce an improved stack. A reason for this is that raypaths, through a complex structure, can differ greatly over the large range of offsets within a CMP gather. The small offsets may have raypath similarities that would yield a better stack. A disadvantage of this approach is that the signal-to-noise ratio may be compromised by stacking the near offsets only. These alternative approaches are used whenever possible to obtain an improved stack. Imaging then is completed by doing time migration after stack. If these attempts to improve the CMP stack fail, time or depth migration should be tried before stack, depending on the problem.

5.2 DEPTH MIGRATION

Lateral velocity variations often are associated with steep dips. Hence, a depth migration algorithm should handle steep dips well. The 65-degree omega-*x* migration algorithm (Section 4.3.4) is used in all migrations (time and depth) in this section. The omega-*x* scheme is particularly suitable for depth migration because the application of the thin-lens term amounts to a complex multiplication in the frequency domain (Appendix C.3).

The problem of lateral velocity variation will be studied using a point diffractor buried in a media with five different types of velocity-depth models. The first velocity model is shown in Figure 5-1. The corresponding zero-offset section consists of a perfect diffraction hyperbola. Therefore, only the diffraction term is needed in imaging the scatterer (see Appendix C.2). The surface projection of the point scatterer (the source), which is at CMP 240 and indicated by the vertical arrow, is aligned with the apex of the diffraction hyperbola. After time migration, the diffraction hyperbola is collapsed to its apex which, in this case, is coincident with the CMP location of the point diffractor.

Consider what happens when the point diffractor is lowered into the second layer as shown in Figure 5-2. Raypaths from the scatterer to the surface are bent at the interface between the first and second layers according to Snell's law of refraction. The zero-offset section, which also is shown in Figure 5-2, is approximately hyperbolic. From Section 3.2, we note that in a horizontally layered earth model, traveltimes are governed by the hyperbolic moveout equation. However, moveout is hyperbolic only within the small-spread approximation. The velocity associated with this approximate hyperbola is the vertical rms velocity down to the diffractor. Suppose the velocity-depth model of Figure 5-2 is replaced with that of Figure 5-3, where the first layer's velocity now corresponds to the rms velocity of the diffractor of Figure 5-2 (1790 m/s). The zero-offset section associated with this new model is a perfect hyperbola (Figure 5-3). The traveltimes in the zero-offset section that were derived from the original model (Figure 5-2) differ only negligibly from the hyperbolic traveltimes of this section at the far flanks. Moreover, the apex of this approximate hyperbola coincides with the surface projection of the diffractor (indicated by an arrow). Therefore, only time migration is needed to image a diffractor that is buried in a horizontally layered earth model. This time migration can be performed either by Kirchhoff summation, which uses the rms velocity, or by finite-difference or *f-k* methods. The latter methods honor the horizontally layered velocity model and the associated raypath bendings at the interfaces.

Suppose the point diffractor is placed in the third layer as shown in Figure 5-4. We now no longer have a hyperbolic diffraction response. Moreover, the response is skewed so that the apex A of the traveltime trajectory does not coincide with the lateral position B of the diffractor. As expected, time migration partially focuses the energy toward its apex A, which is left of the actual lateral position of diffractor B.

To properly focus the energy and place it laterally to its true subsurface position B, depth migration must be done as shown in Figure 5-4. The depth-migrated image is aligned with true subsurface position B. Lateral positioning is accomplished by the thin-lens term. The amount of lateral shift is the lateral distance AB between the apex of the skewed diffraction curve A and the actual location of the scatterer B. This shift depends on the amount of ray bending that occurs at the interfaces above the image point.

From Figure 5-4, note that the apex of skewed diffraction curve A coincides with the surface position of the ray that emerges vertically. This special ray, the *image ray*, was first recognized by Hubral (1977). The image ray associated with the point diffractor in Figure 5-4 is roughly at midpoint 200. The diffractor itself is located beneath midpoint 240. Therefore, the lateral shift is equivalent to 40 midpoints.

There is no lateral shift for the horizontally layered earth model (Figure 5-2), since there is no lateral velocity variation. The image ray emerges at the surface location coincident with that of the diffractor. For a mild lateral velocity variation, as in Figure 5-5, the lateral shift is less than 10 midpoints. For some objectives, this small lateral shift and the complete focusing may not be critical; therefore, time migration may be as good as depth migration. In those cases, the coefficient of the thin-lens term is negligibly small, which is why we often can get away with time migration in areas of mild to moderate lateral velocity variations.

The imaging problem is complicated when the overburden is as complex as that in Figure 5-6. Here, because of the bowties, a distorted diffraction curve indicates a false structure. The resulting complexity can give rise to more than one image ray. In this case, three image rays emerge at around midpoints 160, 250, and 370. Time migration fails here and imaging of this scatterer can only be achieved by depth migration.

Some examples of lateral velocity variations were studied in Figures 5-1 through 5-6. The image ray behavior and the quality of focusing determines whether time or depth migration should be done. If the starting and end points of the image ray have the same CMP location (Figure 5-2), only time migration is needed. However, if the image ray deviates several CMP locations (Figure 5-4), then depth migration may be required depending on the nature of the exploration objective. A small amount of deviation (Figure 5-5) usually implies a well-focused time migration result, and, hence, a good picture of the geometric form of the subsurface. It is the reflector form that is most critical to interpreters since the lateral shifts defined by the image rays can be applied to the data (Hubral, 1977) or to the structure map after time migration. Large image ray deviations imply grossly incorrect focusing and, thus, imply depth migration rather than time migration. Finally, if more than one image ray is associated with a subsurface point (Figure 5-6), then depth migration is imperative.

These observations on the point diffractor models now are extended to a velocity-depth model that involves the reflecting interfaces in Figure 5-6. The image rays associated with this model are shown in Figure 5-7. There is no deviation from the vertical along the image rays down to horizon 2. Therefore, depth migration is not needed to image this horizon. On the other hand, the image rays significantly deviate from the vertical as they travel down to horizons 3 and 4. For example, the image ray starting at CMP location 140 reaches horizon 4 beneath approximately CMP location 180; a lateral shift of 40 midpoints. Proper imaging of these two horizons is achievable only by depth migration.

Since the image ray plot is a powerful diagnostic tool in determining the type of migration (time or depth) given the velocity-depth model, we can speculate that image rays might be useful for converting time-migrated sections to depth sections. From Figure 5-4, note that time migration collapses the energy to apex A of the diffraction curve that coincides with the image ray location at the surface. The time migration output can be converted

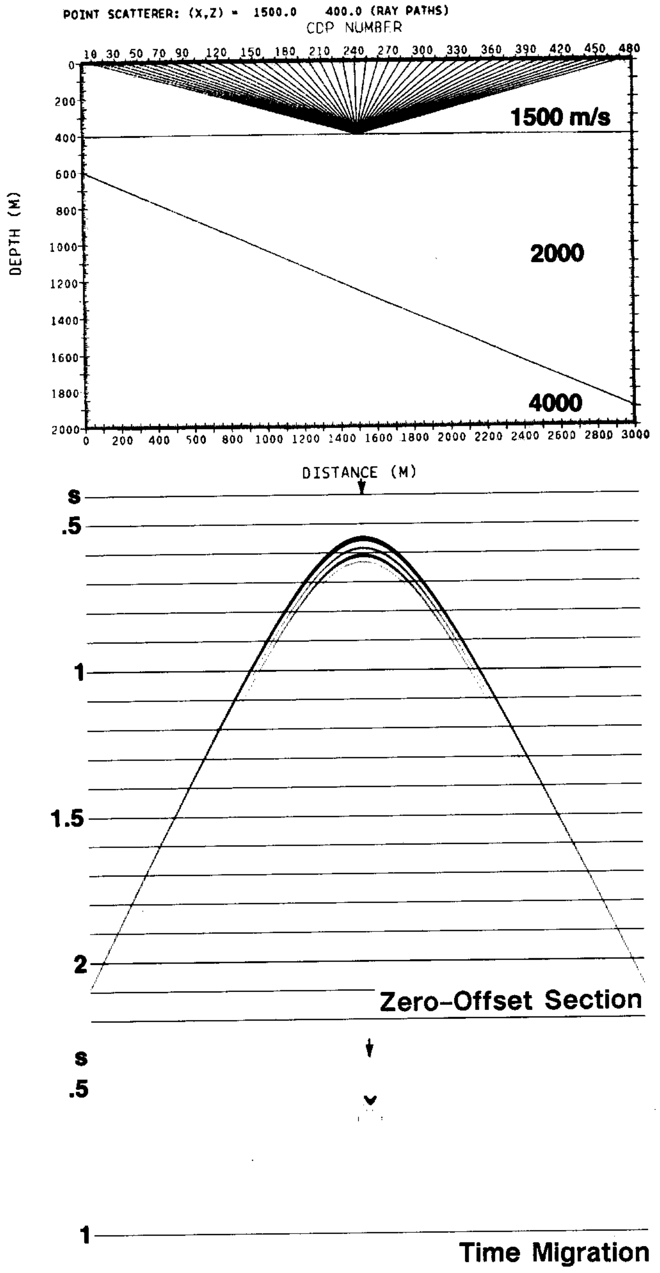


FIG. 5-1. The response of a point diffractor buried in a constant-velocity medium (top frame) is a hyperbola (middle frame). Time migration (bottom frame) is adequate in imaging this diffractor. The arrows indicate the surface projection of the true lateral position of the diffractor.

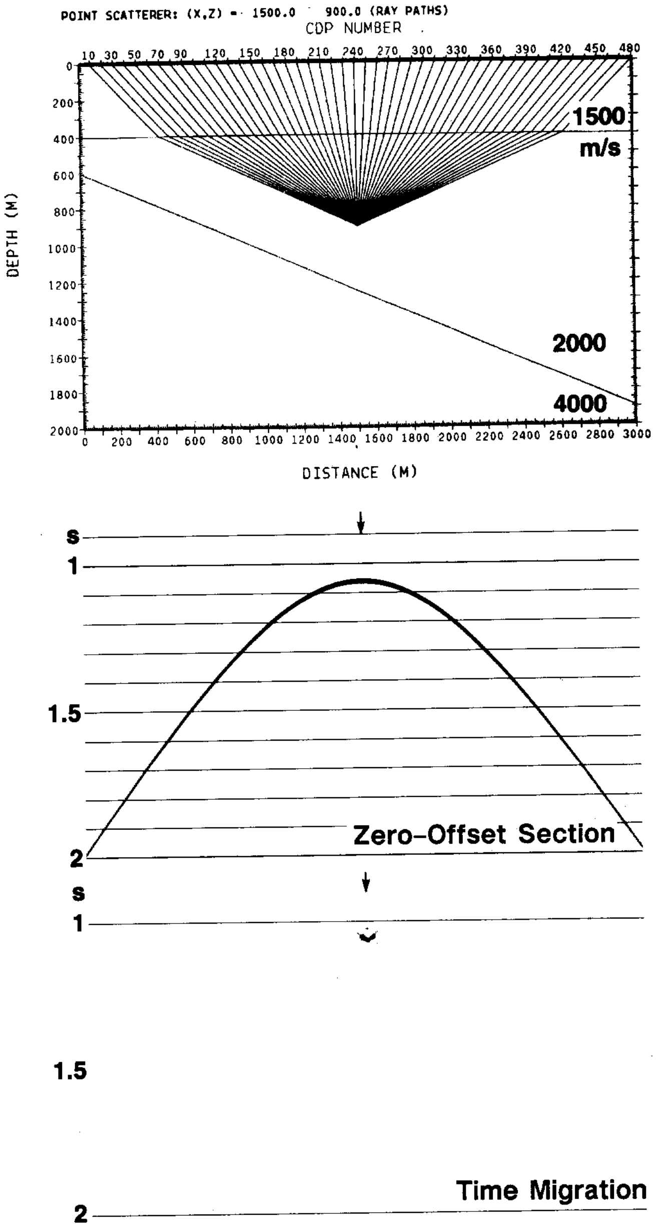


FIG. 5-2. The response of a point diffractor buried in a layered medium (top frame) is approximately a hyperbola (middle frame). Time migration (bottom frame) still is adequate for imaging the diffractor. The arrows indicate the surface projection of the true lateral position of the diffractor.

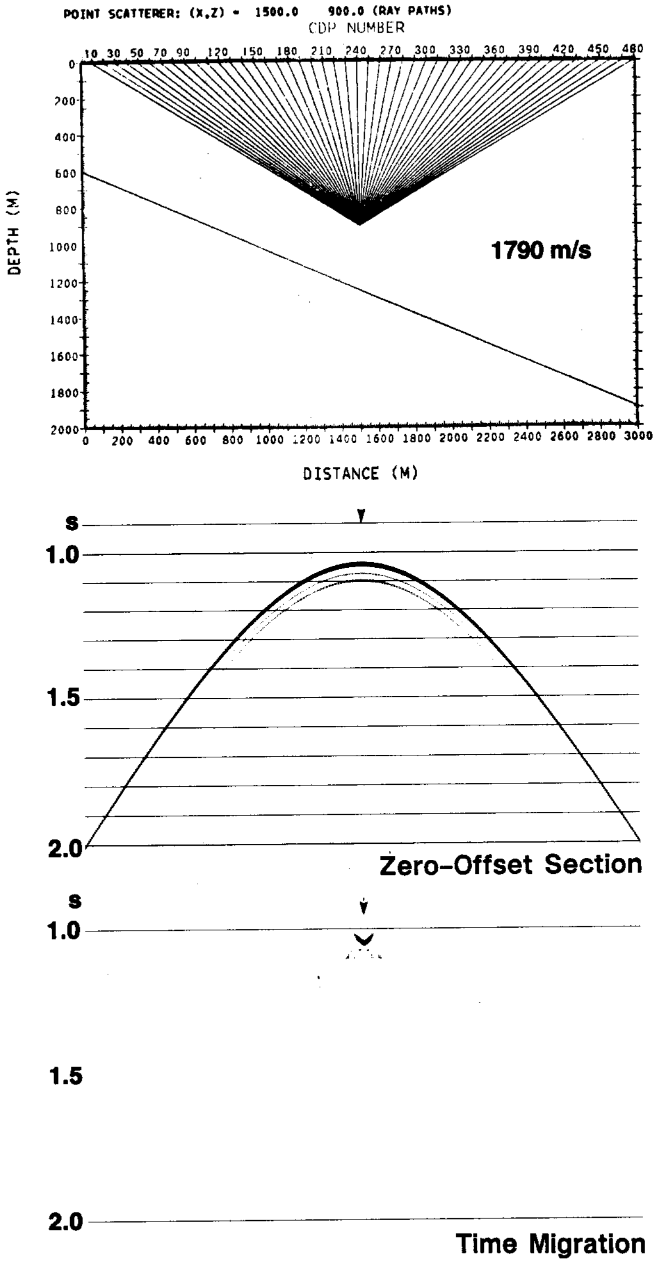


FIG. 5-3. The response of a point diffractor buried in a layered medium (Figure 5-2) is nearly equivalent to the response of another point diffractor buried in a constant-velocity medium (top frame), where the constant velocity is the rms value of the layered velocity function at the depth of the original diffractor in Figure 5-2. Traveltimes associated with the zero-offset sections in Figures 5-2 and 5-3 depart from one another only at the very far flanks of the diffraction curve. Time migration (bottom frame) is adequate for imaging the scatterer.

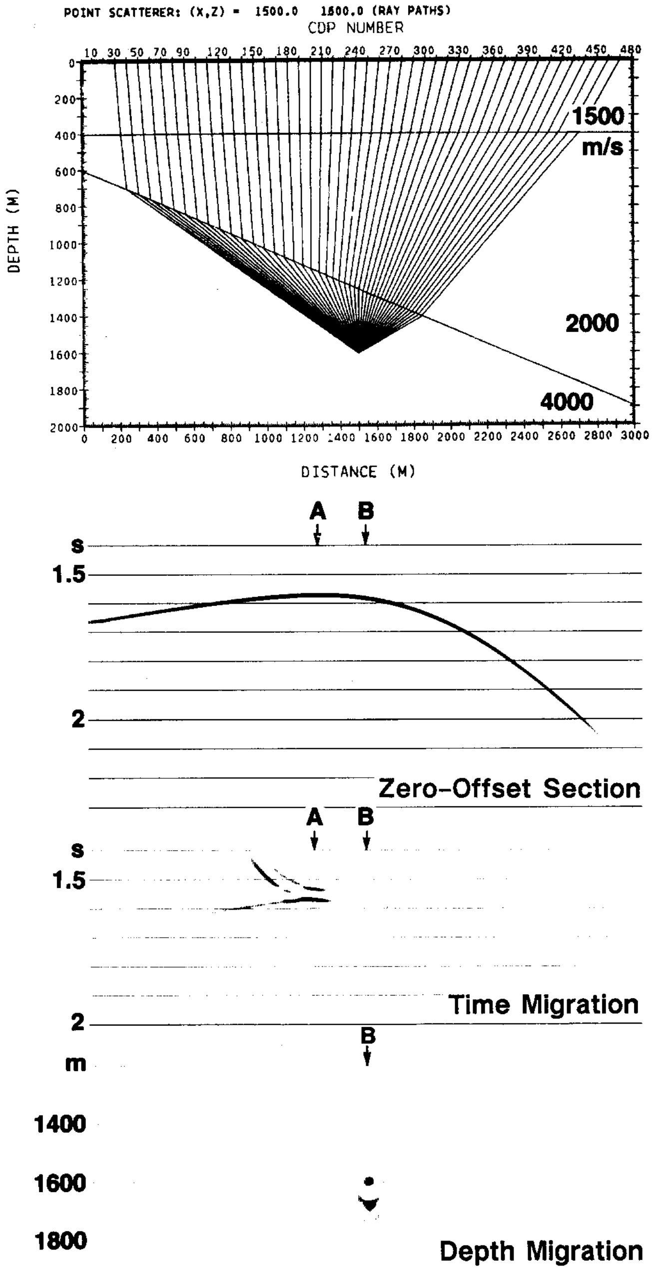


FIG. 5-4. The response of a point diffractor buried in a medium with strong lateral velocity variation (top frame) is a skewed hyperbola with its apex shifted to the left of the true position. Time migration no longer is a valid process. Depth migration is needed.

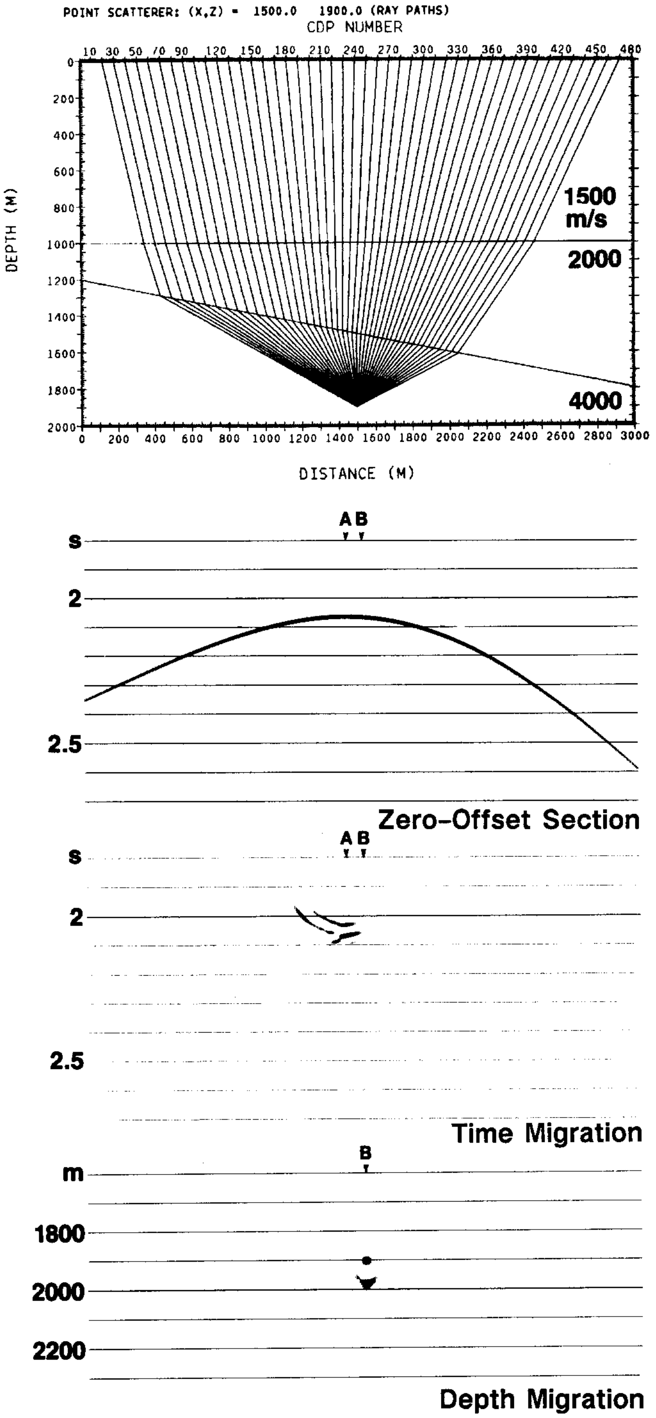


FIG. 5-5. The response of a point diffractor buried in a medium with mild lateral velocity variation (top frame) is a slightly skewed hyperbola with its apex shifted to the left of the true position. Time migration still can be an acceptable way to image the diffractor.

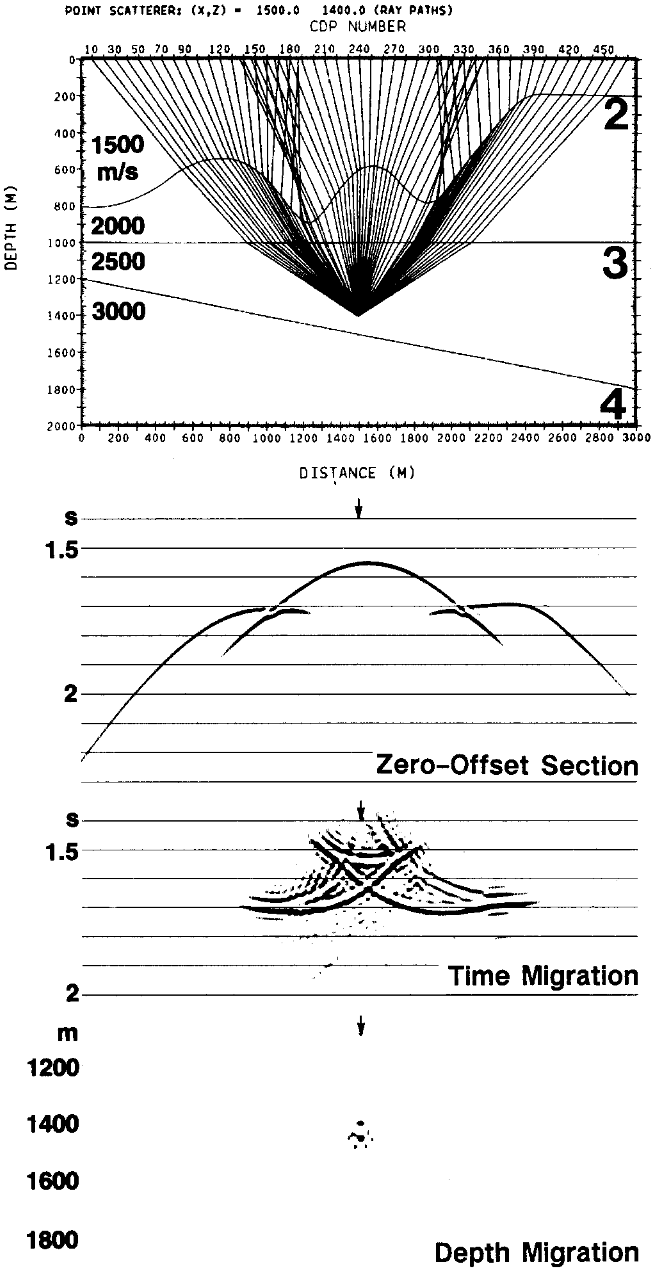


FIG. 5-6. The response of a point diffractor buried in a medium with severe lateral velocity variation (top frame) is a distorted traveltime curve that implies false structural features. Time migration no longer is acceptable; depth migration must be done.

to depth along the image rays, rather than along the vertical rays (Hubral, 1977).

Mapping along the image rays performs some of the action that is associated with the thin-lens term. Remember that at each downward continuation step, the action of the thin-lens term is vertical time shift that depends on spatial velocity variation. Since the thin-lens and diffraction terms are applied in an alternate manner as the wave field is downward continued in depth, the effects of these two terms are strongly coupled when the lateral velocity variation is severe (as in Figure 5-6). When velocity variation is moderate to strong, these two terms often can be fully separated and applied consecutively without significant error (Larner et al., 1981). Full separation implies that correction for the effects of the thin-lens term can be done either before or after time migration. If correction is done after time migration, image ray mapping should be used. If correction is done before time migration, mapping using vertical time shifts usually is applied. In practice, correction before time migration often performs better, since it tends to provide a better focused migration result.

Lateral velocity variations can be mild to moderate (as in Figure 5-5), strong (Figure 5-4), or severe (Figure 5-6). Four ways to obtain a *depth section* and the conditions that are necessary for them to be valid are shown in Figure 5-8.

Time migration is valid strictly for the vertically varying velocity model. Because it directly implements the thin-lens term, depth migration requires a detailed velocity model in which all the lateral velocity variations are considered. How do we derive a detailed velocity model? If the detailed velocity model were known exactly, we also would know the subsurface geologic model, and, hence, would have no need to do the migration. In this respect, depth migration can be viewed as a means of testing an initial geologic hypothesis. Consequently, most good depth migrations are the result of an iterative process similar to the one shown in Figure 5-9. Start with a velocity model based on the best stack and other available information, such as well data. The CMP stack is time-migrated using this initial velocity model. The output is depth-converted along the vertical rays and a new velocity-depth model is produced. The CMP stack then is depth-migrated using this initial velocity-depth model and the output is interpreted. If the interpretation does not conform to the initial velocity-depth model, then the model is modified and the depth migration is done again. Iteration continues until the model input to depth migration matches the depth-migrated section (Figure 5-9). (In practice, to reduce cost, some of the initial iterations often are done with ray tracing methods using dip bars picked from the time sections.) This iterative approach should yield a result that converges to the geologic model. However, there is no guarantee that this is a unique solution. A detailed demonstration of such an iterative depth migration is provided in Section 5.2.2.

Besides velocity-depth model building, the zero-offset assumption for a stack is another problem. Complex, nonhyperbolic moveout on CMP data associated with severe lateral velocity variations requires depth migra-

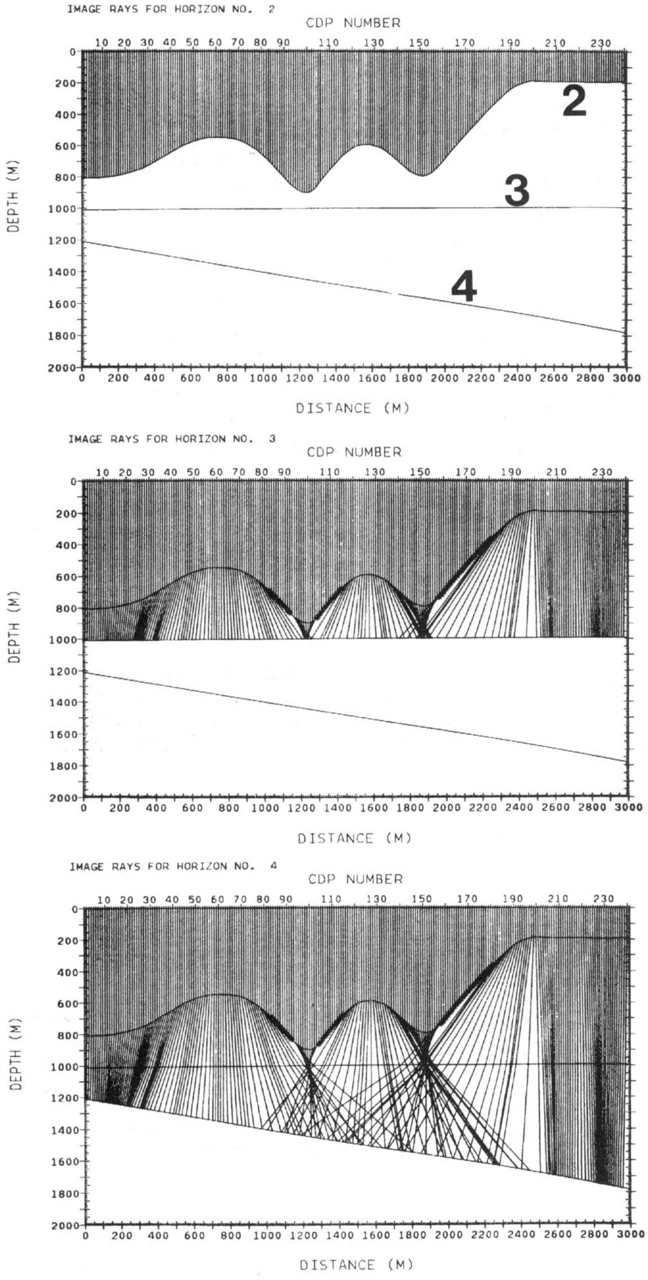


FIG. 5-7. Image rays through the velocity-depth model in Figure 5-6. Note the deviation from the vertical of the image rays as they travel through the complex structure.

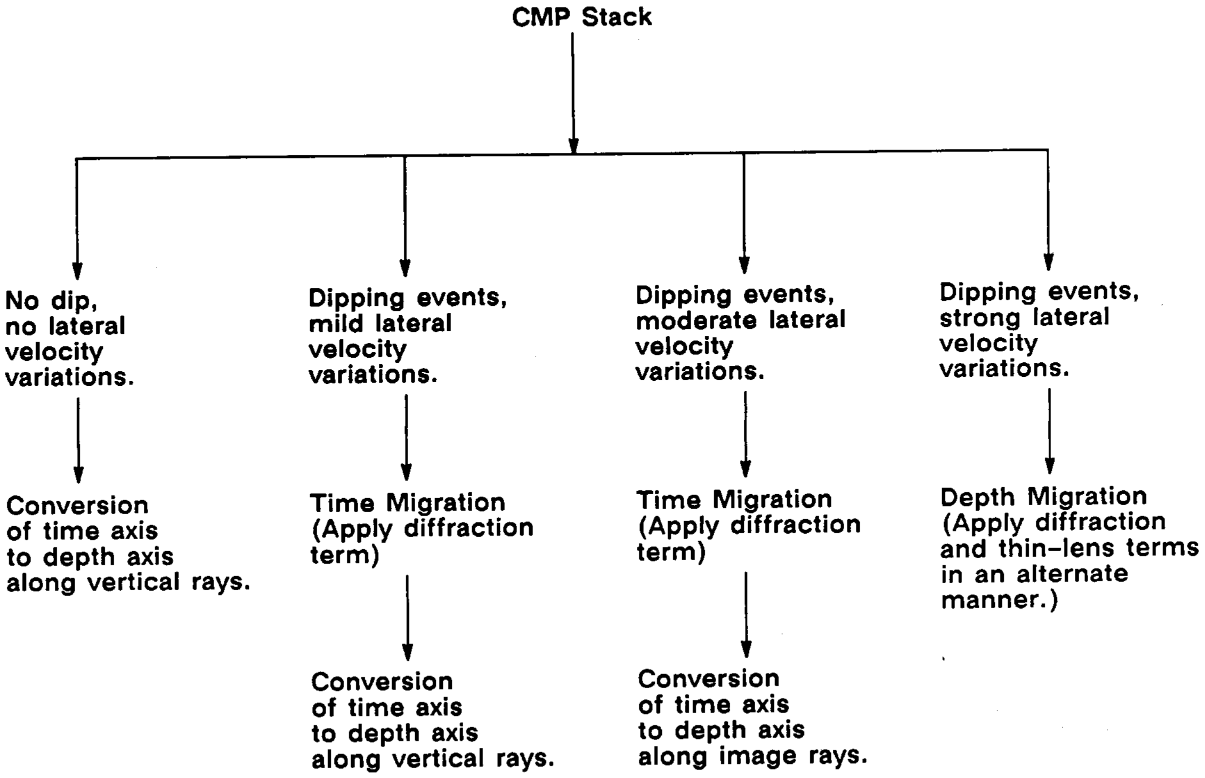


FIG. 5-8. Depth conversion schemes for different degrees of lateral velocity variations. An alternative to the depth migration approach (rightmost column) is to do vertical conversion to a nonlaterally variable velocity model followed by time migration and vertical stretch to depth. In theory, this is incorrect. In practice, it often is acceptable because the difference between the results from this approach and depth migration falls within the limits of the uncertainty in velocity.

tion before stack. However, the stack can be used to derive a plausible velocity-depth model, which then can be used to image data before stack.

Three types of commonly encountered complex overburden structures are considered in the following sections: irregular water bottoms, salt diapirs, and imbricate structures associated with overthrust tectonics.

5.2.1 Irregular Water Bottom

Consider the irregular water-bottom topography in Figure 4-11a. Forward modeling using normal-incidence rays yields the traveltime curves in Figure 4-11b. Note that simple linear events in the depth model seem to be complicated on the time section because of a complex overburden; i.e., the irregular water bottom causes complicated raypath distortions. The zero-offset section obtained from the depth model is shown in Figure 5-10a. Clearly time migration (Figure 5-10b) is not the way to

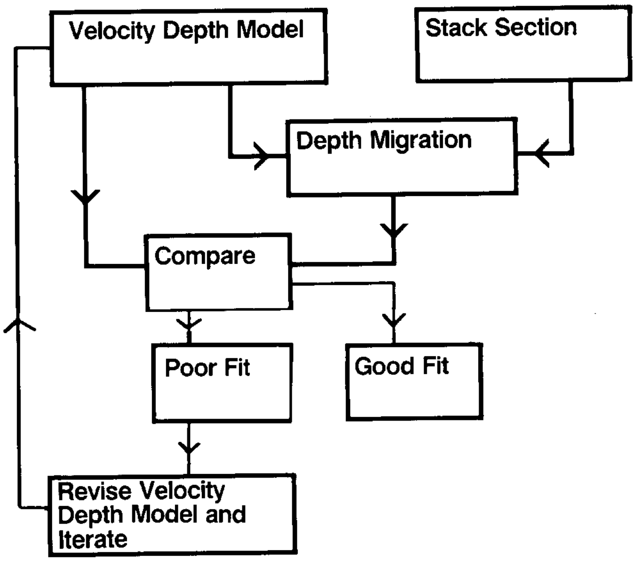


FIG. 5-9. An iterative depth migration procedure.

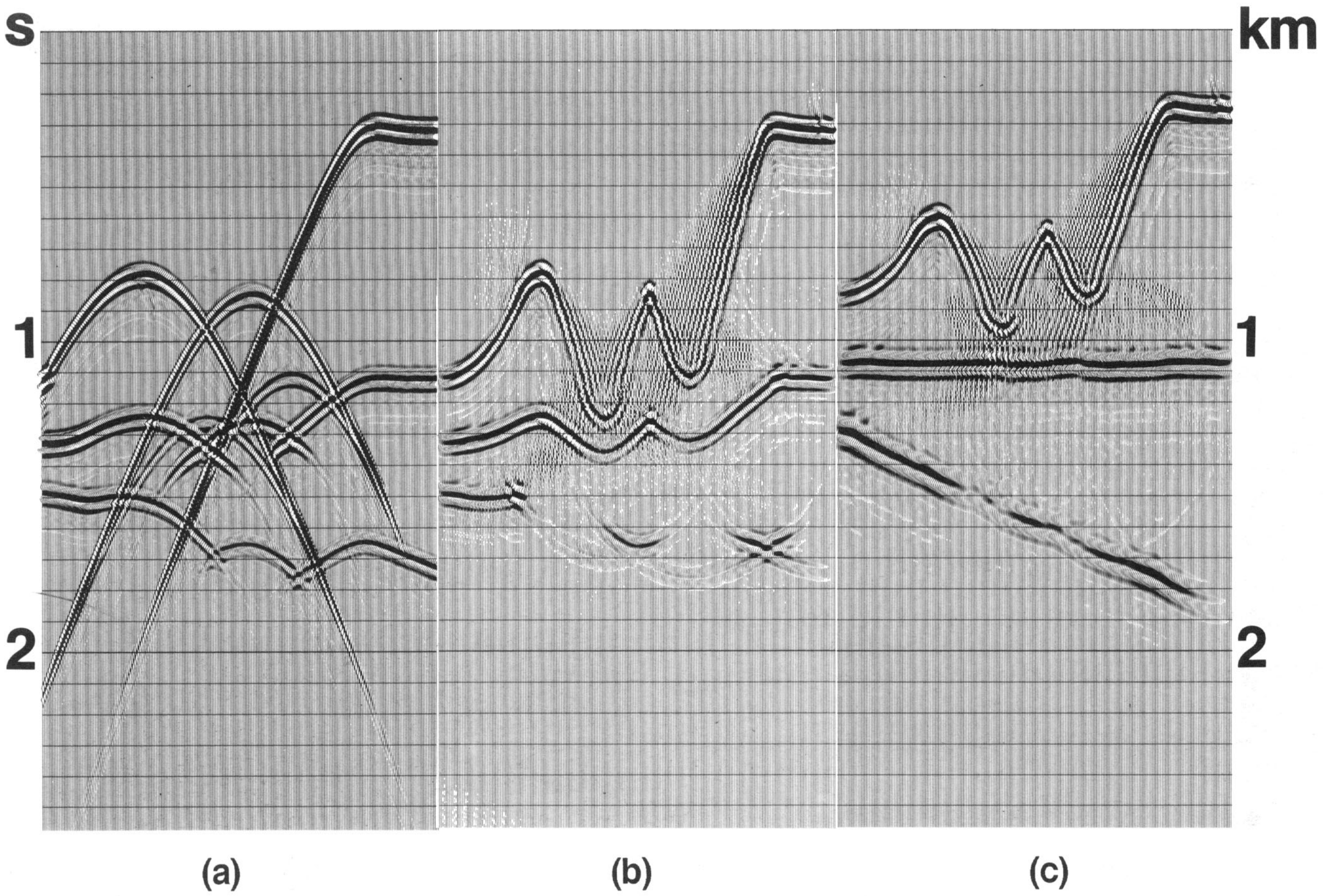


FIG. 5-10. (a) Zero-offset section derived from the depth model in Figure 4-11a. (b) Although time migration imaged the water bottom correctly, it yielded false structures that are associated with the two reflectors below. (c) Depth migration gives the true subsurface picture when the velocity-depth model is correct.

construct the correct image beneath such a complex overburden. Instead, depth migration is needed (Figure 5-10c). Time migration images the water bottom correctly; however, it treats the interferences along horizons 3 and 4 (as annotated in Figure 5-6) as bowties and converts them to synclines as shown in Figure 5-10b. Depth migration yields a more realistic subsurface image, provided the velocity-depth model is accurate. The velocity-depth model used to derive Figure 5-10c is the true velocity-depth model from Figure 4-11a.

The raypath diagrams in Figure 5-7 show why depth migration is needed to image events below an irregular water bottom. The image rays associated with the irregular water bottom (top frame) itself do not depart from the vertical, and, thus, depth migration is not required to image this horizon. On the other hand, strong lateral velocity variations across the water bottom interface cause the image rays associated with the horizons below to depart from the vertical (middle and bottom frames). Thus, depth migration is required to image these two horizons.

As shown for the zero-offset section (Figure 5-10), time migration of the stacked section (Figure 5-11a) is not the solution to complex overburden problem (Figure 5-11b). Figure 5-11c shows the results of depth migration on the CMP stack for the same velocity-depth model used for Figure 5-10. The results in Figure 5-11 are inferior because the CMP stack does not accurately represent a zero-offset section. This occurs because the complex overburden causes raypath distortions that cause anomalous, nonhyperbolic moveout patterns in prestack data, as seen in the CMP gathers of Figure 4-13a, which are associated with the stacked section in Figure 5-11a. Comparisons of depth migrations from the zero-offset and stacked sections (Figures 5-10c and 5-11c, respectively), demonstrate that this complex overburden problem really should be handled before stack. For accurate imaging by poststack depth migration, both an accurate velocity-depth model and a stacked section that closely represents a zero-offset section are needed. When an accurate stack is not available, a full prestack depth migration may be required.

Depth migration before stack involves alternate downward continuation of common-shot and common-receiver gathers at discrete depth steps (Schultz and Sherwood, 1980; Stolt, 1978). This scheme requires sorting the data

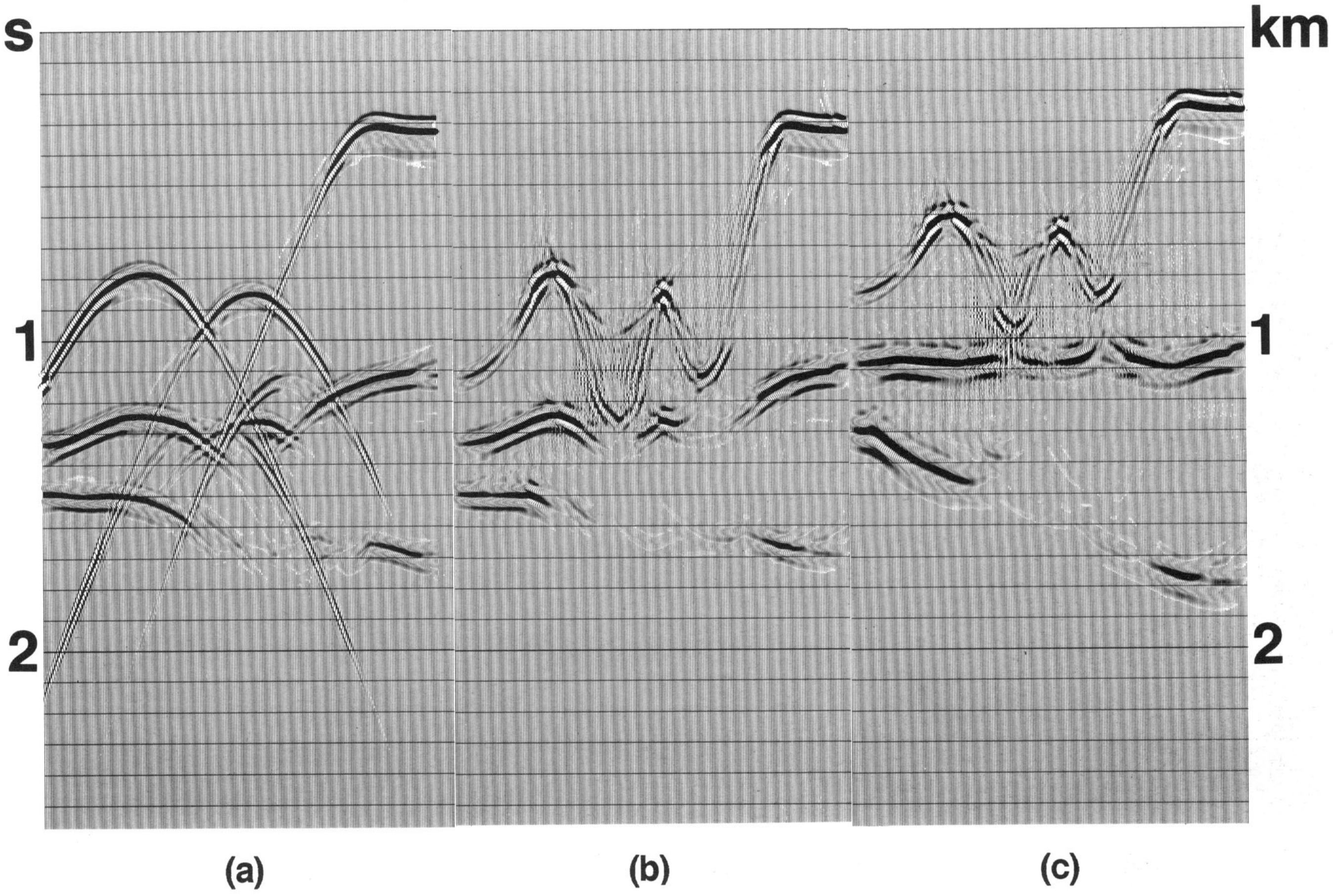


FIG. 5-11. (a) Stacked section derived from the depth model in Figure 4-11a. Time migration (b) yields false structures, while depth migration (c) does not provide the true subsurface picture, even if the correct velocity-depth model is used.

back and forth between shot and receiver gathers at each depth step. One way to cut cost is to stop the process just below the complex overburden. Another way is to bypass prestack depth migration completely and use other processes such as hand static correction and stacking programs based on trace correlation or HVA (Section 3.3.3) to improve stacking. The improved stacked section often can be accepted as a reasonable representation of the zero-offset section.

With this in mind, refer to the field data example in Figure 5-12a. Note the false structures along the unconformity T due to the irregular water-bottom topography. These are especially noticed below midpoint locations A, B, and C. Sagging of the reflection that is associated with the unconformity at these locations is attributed to the low-velocity overburden, here, the water layer. In addition to the complex overburden, this section also contains strong diffracted multiples. Once again, time migration does not provide the solution for such a complex overburden as in Figure 5-12b. Distortions along the unconformity still exist. A velocity-depth model was derived from the seismic data for use in depth migration of the CMP stacked section. Based on the initial depth migration results, the velocity-depth model was revised as shown in Figure 5-12c. Depth migration of the CMP stack based on the revised velocity-depth model is shown in Figure 5-12d. Note how the distortions of the irregular water bottom are corrected. (Compare this with Figure 5-12b.) There still is some disruption of the reflection from the unconformity beneath midpoint location C. However, the revised velocity-depth model suggests that what was input as a velocity model is similar to the structure on the depth migration after this second iteration. Further improvement is likely to be minimal.

5.2.2 Salt Diapir

A typical model for a salt diapir structure is shown in Figure 5-13a. Such structures commonly are encountered in offshore West Africa, the Gulf of Mexico, the Red Sea, and the North Sea. In these cases, the salt layer has significantly high velocity, which causes a large velocity contrast with the surrounding layers. The zero-offset

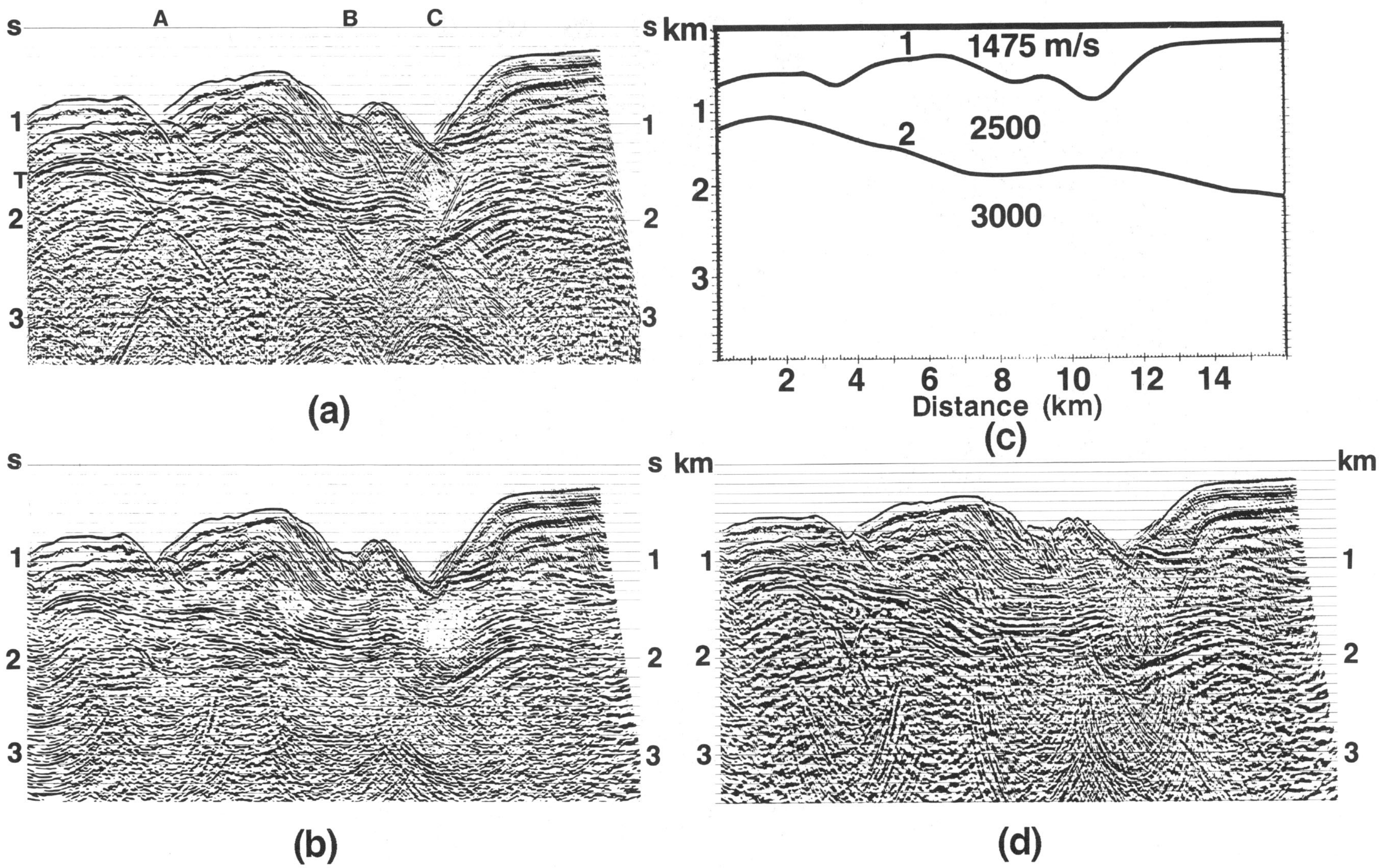


FIG. 5-12. (a) Conventional CMP stack from an area with irregular water-bottom topography. (b) Time migration of the conventional CMP stack. (c) Velocity-depth model used in depth migration of the CMP stack. horizon 1 represents the water bottom; horizon 2 represents the major unconformity denoted by *T* in (a). (d) Depth migration of the conventional CMP stack (a) using the velocity-depth model (c). (Data courtesy Hispanoil.)

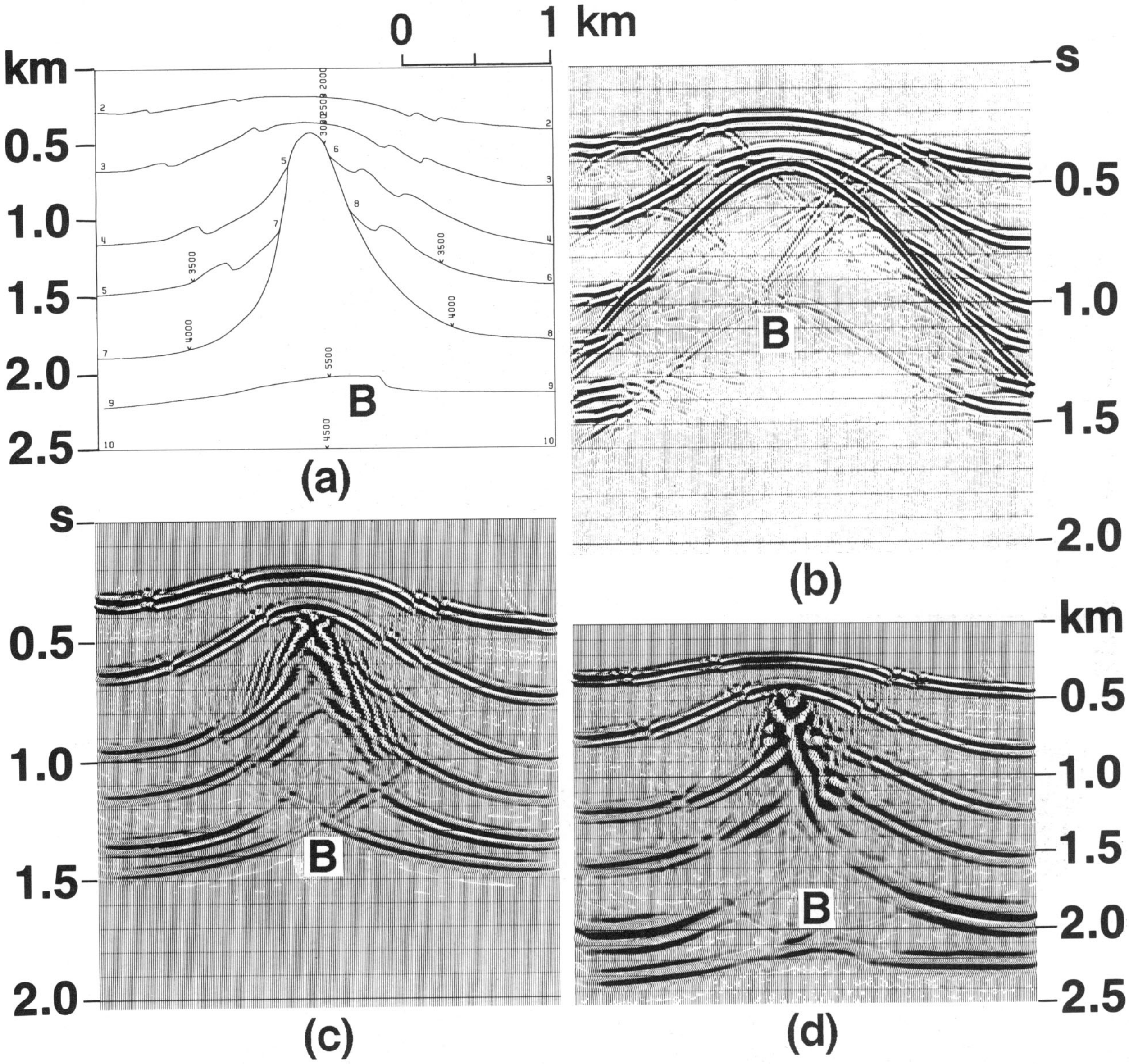


FIG. 5-13. (a) Velocity-depth model for a salt diapiric structure. (b) Zero-offset section derived from the velocity-depth model in (a). Event B is the reflection from the salt base. (c) Time migration of the zero-offset section. (d) Depth migration of the zero-offset section using the true velocity-depth model in (a). (Modeling by Deregowski and Barley, courtesy British Petroleum.)

section derived from this depth model is shown in Figure 5-13b. The event associated with the flanks of the salt dome is mostly diffraction off the tip of the salt dome; the profile length is too short to model the reflections along the flanks. The base of the salt layer B, which is pulled up because of the high-velocity overburden (the diapiric structure), is of interest. Time migration is acceptable if the target is above the salt layer (such as entrapments along the flanks), since there are no severe raypath distortions (Figure 5-13c). However, if the target horizon is beneath the salt diapir, then there is a complex overburden problem. From Figure 5-13c, note that the base of the salt layer B is not imaged properly. On the other hand, depth migration delineates the structure beneath the salt layer, including the fault present at the salt base, quite accurately (Figure 5-13d).

The velocity-depth model used in the depth migration in Figure 5-13d is the true model (Figure 5-13a) that was used to create the zero-offset section that was the input to the depth migration. In practice, we do not know the exact depth model. How, then, do we create a plausible depth model? We must use all of the available information: velocities estimated from the seismic data, well data, and the geologist's model of the subsurface. Time migration can help build part of the depth model in which there is no overburden problem. An iterative depth migration procedure, such as that described in Figure 5-9, then can be used to derive a final velocity-depth model.

Figure 5-14 shows a field data example of a salt diapir. Note how reflections below the salt dome at about 2 s are pulled up in the middle of the section. This CMP stacked section was obtained using stacking velocities derived from HVA. The HVA improved the quality of reflections, especially beneath the salt dome (Figure 3-48). Time migration yields an inadequate image of the base of the salt layer B (Figure 5-14).

Now consider an iterative depth migration to better image the base of the salt. Follow this procedure:

1. Start with your best quality CMP stacked section (Figure 5-14a) and do a time migration (Figure 5-14b).
2. Convert the time-migrated section to a depth section along *vertical rays* (Figure 5-14c). This conversion should be reasonably valid for all parts of the CMP stack that are not adversely affected by the presence of an overburden. In this example, except for location A and its vicinity, depth conversion output is considered a fairly reliable initial velocity-depth model. Image rays cannot be used for depth conversion because the velocity-depth model is needed for them and the velocity model is not yet known.
3. Digitize the depth-converted section along major key horizons. The result is shown in Figure 5-15 (top left frame).
4. Using this initial velocity-depth model, depth migrate the CMP stacked section shown in Figure 5-14a. The depth-migrated section is shown in Figure 5-15 (top right).
5. Lay the initial velocity-depth model over the depth-migrated section and modify the model in places where the two disagree.
6. Using this revised velocity-depth model, repeat the depth migration process.
7. Repeat Steps 5 and 6 until the velocity-depth model matches the depth migration output. The match is excellent in this example after three iterations.

We still do not know whether the last output from depth migration reflects the true subsurface picture. The solution is not unique; no method can provide a unique solution because the velocity values entered into the velocity-depth model can be perturbed in many ways and still can match the outputs from depth migration. However, using external information such as well data, the number of plausible solutions can be narrowed down considerably.

5.2.3 Imbricate Structures in Overthrust Belts

A typical model for imbricate structures in overthrust belts is shown in Figure 5-16a. Target zones can be on top of culminations (location A), or on the foreland structure (location B). Figure 5-16b shows a zero-offset section derived from this model. Time migration suffices in delineating targets on the imbricate structures (location A in Figure 5-16a), as shown in Figure 5-16c. However, only depth migration (Figure 5-16d) accurately images the targets beneath the culminated structure (location B in Figure 5-16a).

Figure 5-17a is a data example of overthrust structure, similar to the structure in the synthetic model in Figure 5-16a. Time migration (Figure 5-17b) is adequate for imaging any target in the imbricate structure, but would not yield a correct picture for a target below the imbricate structure.

Another example of an overthrust structure is shown in Figure 5-18a. The data were recorded in 1972 and were time migrated (Figure 5-18b) using the diffraction summation technique. Note that the migration result is reliable within the imbricate structures above 2 s, but events below have been overmigrated. No depth migration was done on these data.

Data quality from overthrust regions can be poor. Such geologic provinces usually have rugged surface topography, which can result in difficult field conditions and cause severe statics problems. Culminated structures also cause energy to scatter. It usually is only after drilling a few wells that one can reasonably interpret and derive a plausible velocity-depth model in these provinces. Figure 5-19 shows a time-migrated overthrust section. Several of the oil fields on top of the culminated structural highs were discovered by the gravity method. Because of an allocthonous overburden, these structural highs are not distinctive on the seismic section. An integrated geophysical effort is essential for exploring

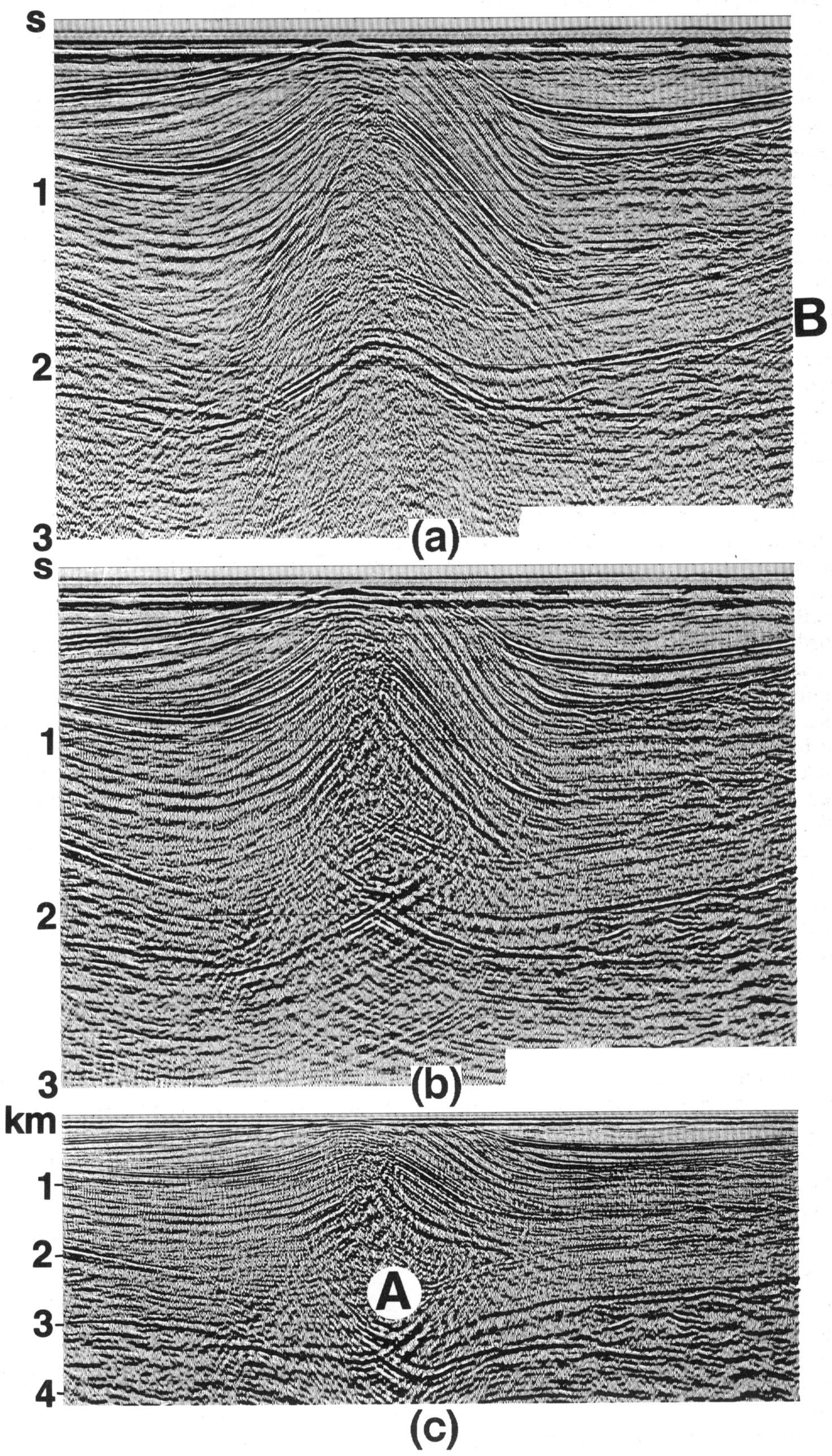


FIG. 5-14. (a) A CMP stacked section containing reflections from a salt dome. (b) Time migration of the CMP stack. Note the similarity between the overmigrated base of the salt layer in this section and in the time-migrated model section in Figure 5-13c. (c) Vertical time-to-depth conversion of the time-migrated section.

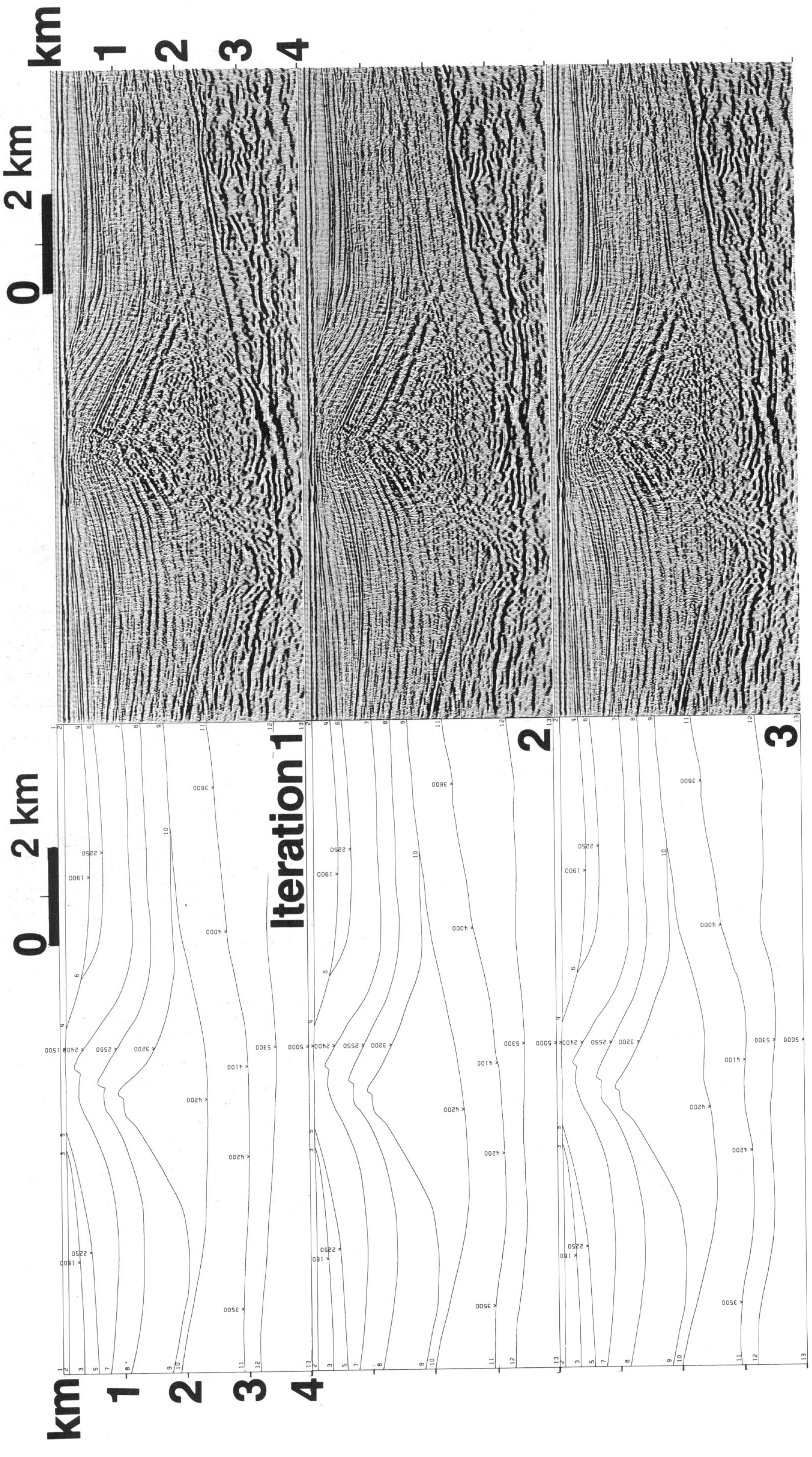


FIG. 5-15. Iterative depth migration of the CMP stack in Figure 5-14a. See text for details.

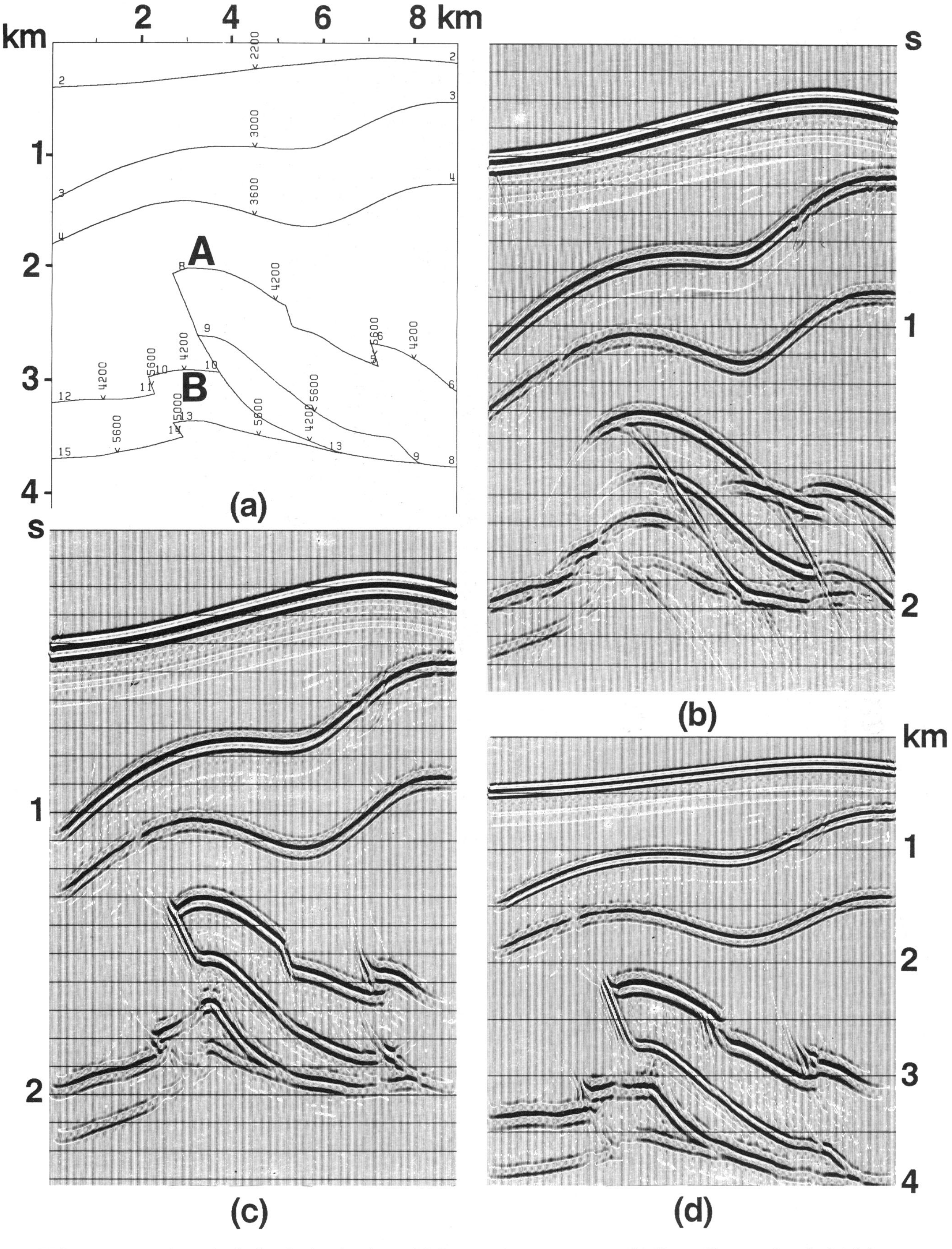


FIG. 5-16. (a) A hypothetical velocity-depth model for overthrust structure. (b) Zero-offset section derived from this velocity-depth model. (c) Time migration of the zero-offset section. (d) Depth migration of the zero-offset section in (b) using the velocity-depth model in (a).

natural resources in difficult areas such as overthrust belts.

Complex geologic structures often are three dimensional. We should not expect 2-D time or depth migration to produce an entirely accurate image of the subsurface in geologic provinces such as overthrust belts or salt tectonics. Figure 5-20 shows a stacked section and its migration from an overthrust region. In particular, many of the events below 3 s are believed to be sideswipe energy (that is, energy that comes from outside the plane of recording). See Section 6.5 for 3-D migration.

5.3 LAYER REPLACEMENT

Consider the depth model in Figure 5-21a. In Section 5.2, we noted that raypath bending occurs because of the complexity of the overburden and results in distortions and disruptions of the underlying target reflections. Without the velocity contrast (Figure 5-21b), the raypaths would not bend and there would be no need for depth migration. Figure 5-21 suggests that replacing the overburden velocity with the substratum velocity (layer replacement) can be a viable alternative to using depth migration for removing the undesirable effects of complex overburden on the substrata.

A technique for layer replacement (Yilmaz and Lucas, 1986) based on wave-equation datuming (Berryhill, 1979, 1984) is presented. Berryhill's datuming technique involves extrapolating a known wave field at a specified datum of arbitrary shape to another datum, also of arbitrary shape. Wave extrapolation is performed using the Kirchhoff integral solution to the scalar wave equation. It incorporates both the near-field and far-field terms [equation (4.5)]. The velocity used in extrapolation is that of the medium confined between the input datum and the output datum.

Figure 5-22a shows a simple case of datuming. Here, a zero-offset section is computed over three point scatterers buried beneath midpoint location A in a medium with the layered velocity structure as denoted between Figures 5-22a and 5-22b. The point scatterers are situated at the layer interfaces at 800, 1300, and 1900 m depths. The traveltime curve from the shallow point scatterer is a hyperbola. The curves of the deeper ones are nearly hyperbolic (Section 3.2.1). Given this wave field at the surface ($z = 0$), downward continue it using the 2000 m/s velocity of the first layer and compute the wave field at the first interface, $z = 800$ m. The result is shown in Figure 5-22b. As expected, the hyperbola associated with the shallow point scatterer largely collapses to its apex since this scatterer is located at the first interface. Because the receivers now are closer to the other two deeper scatterers, events associated with them also are compressed. Figure 5-22b shows the zero-offset section that would be recorded if the receivers were placed along the first interface. The energy from the shallow point scatterer on this section now arrives at $t = 0$ because the datum for this section is the interface at which the scatterer is located.

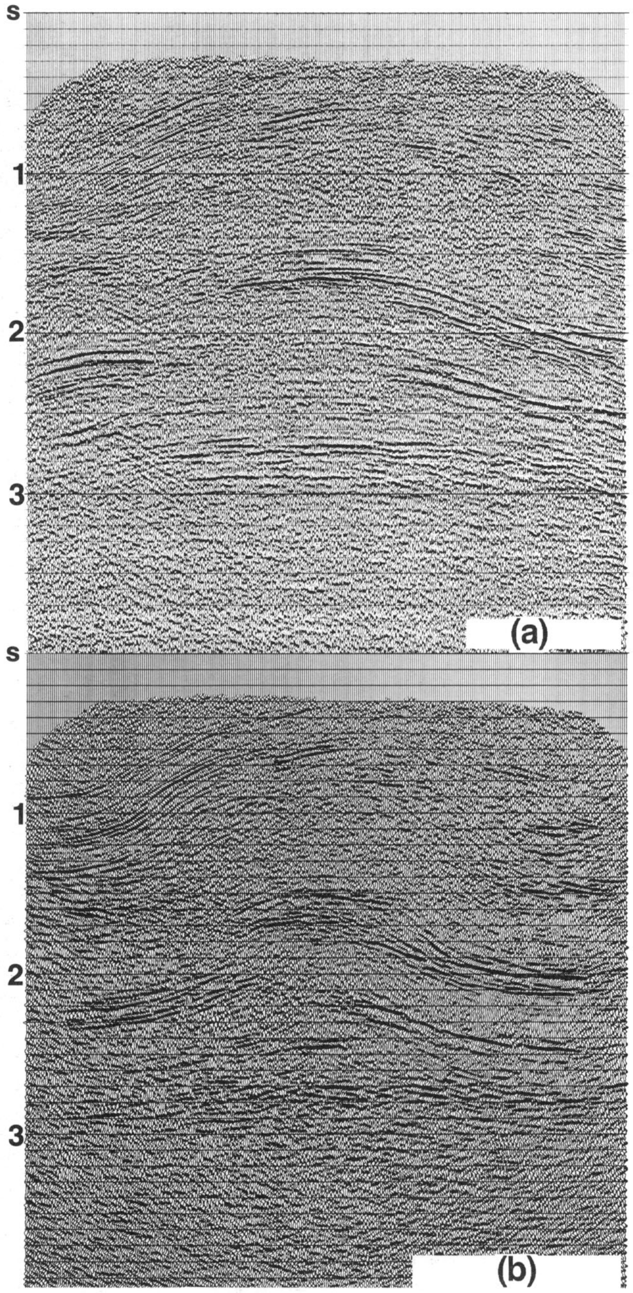


FIG. 5-17. (a) A field data example involving overthrust tectonics and (b) time migration of this stacked section.

The wave field at the first interface ($z = 800$ m), which is shown in Figure 5-22b, now is extrapolated back up to the surface ($z = 0$) using the velocity of the second layer (2500 m/s). Figure 5-22c shows the result, which is compared to the zero-offset section in Figure 5-22d. The latter was derived independently from the velocity model denoted between Figures 5-22c and 5-22d. The first layer with 2000 m/s velocity was *replaced* with the second layer with 2500 m/s velocity.

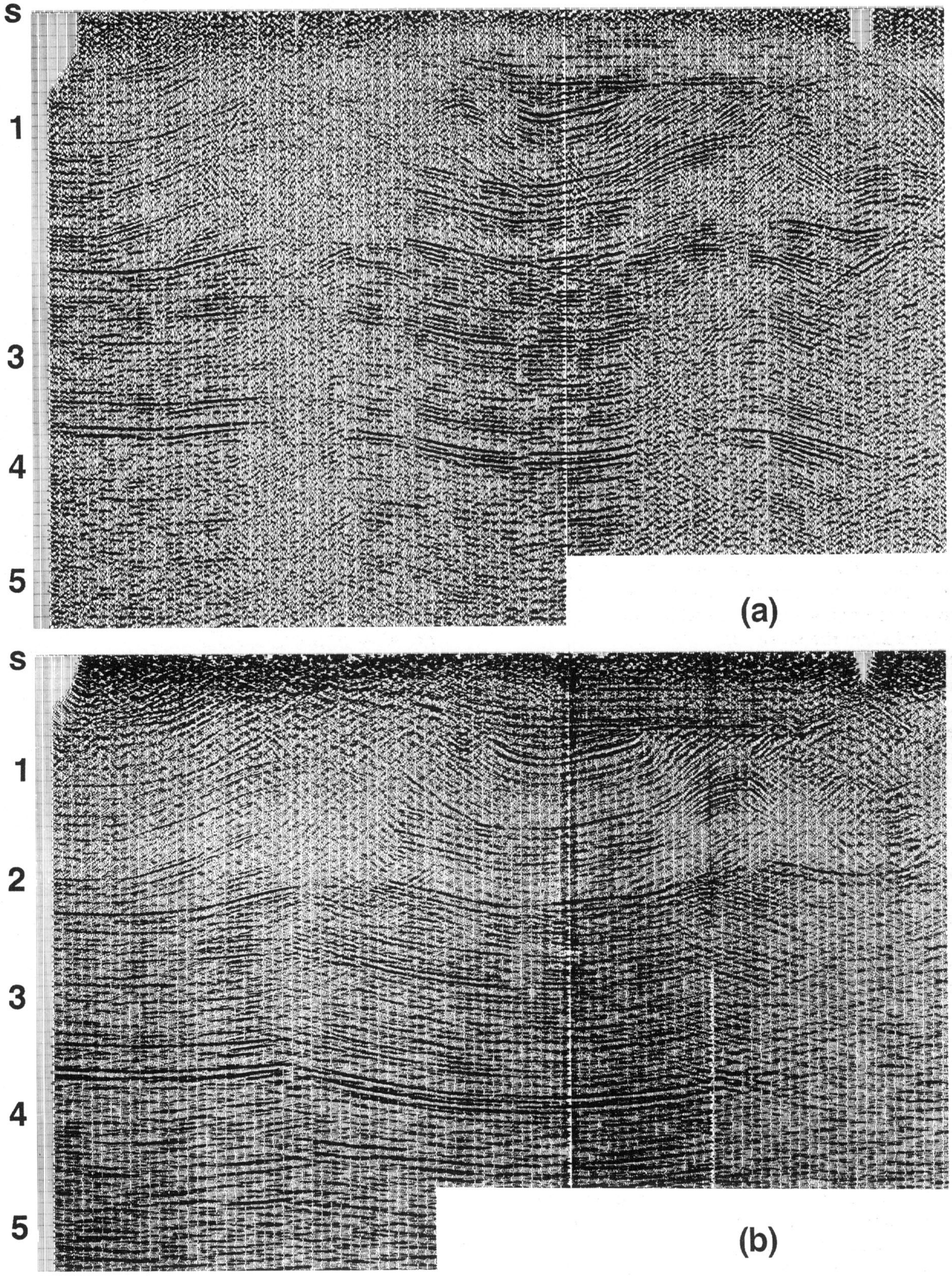


FIG. 5-18. (a) A CMP stack from an area with overthrust tectonics and (b) time migration. (Data courtesy Amoco Production Company.)

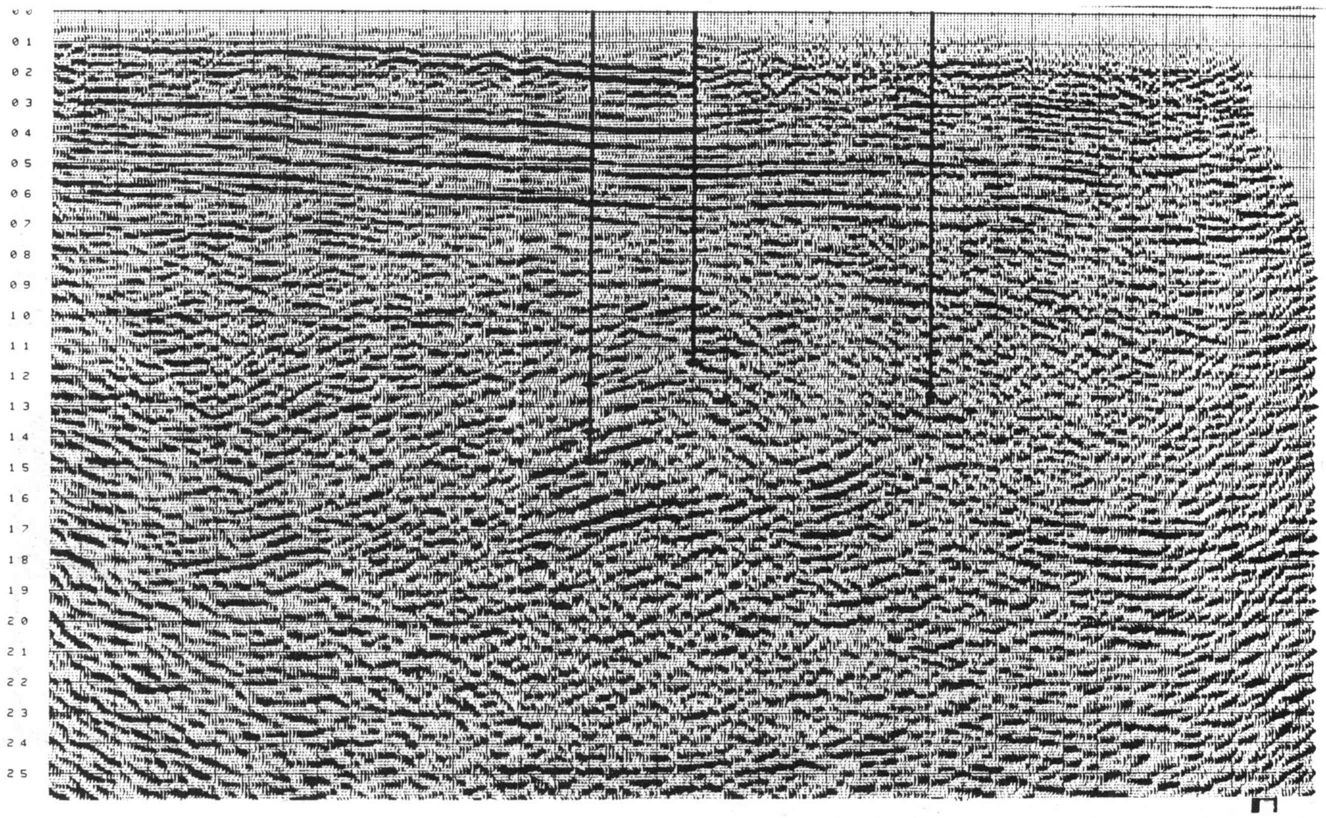


FIG. 5-19. A migrated stacked section from an overthrust belt. The culmination of the imbricate structures was delineated and verified only after drilling a number of wells (indicated by the vertical bars). (Courtesy Turkish Petroleum Corporation.)

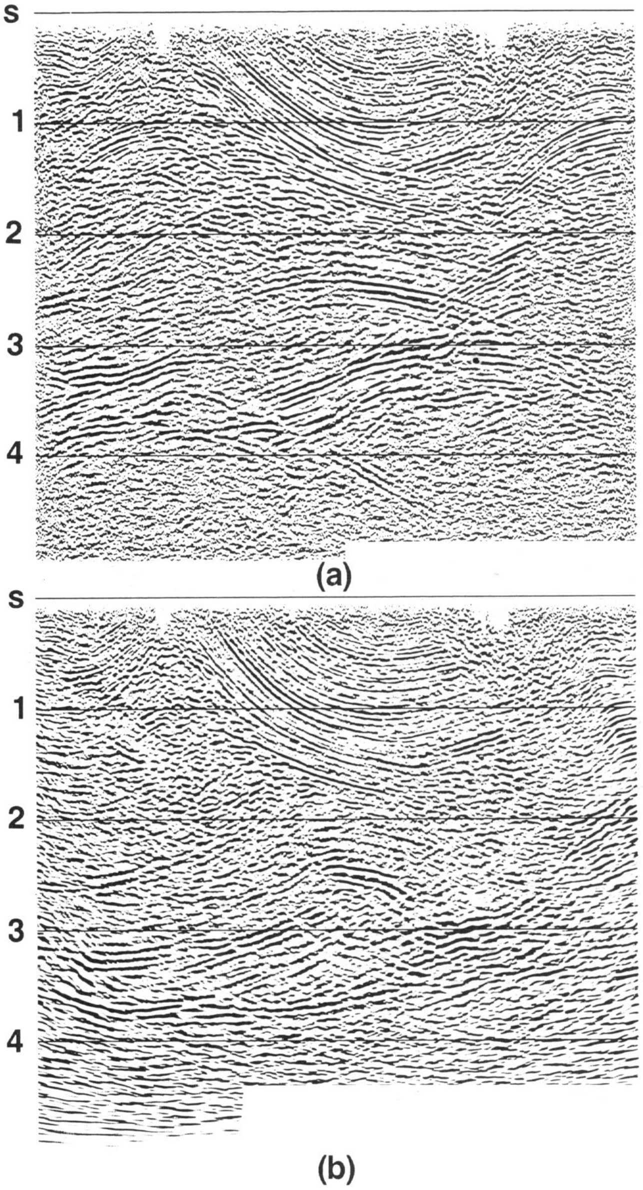


FIG. 5-20. (a) A CMP stacked section from an overthrust region and (b) its time migration. (Data courtesy Occidental Oil of Pakistan Ltd. and Pakistan Oil Fields Ltd.)

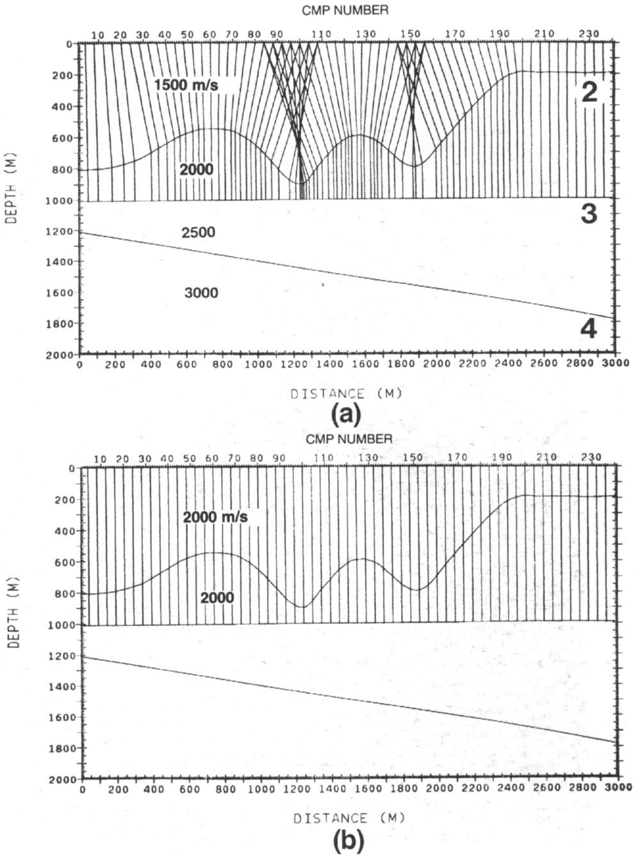


FIG. 5-21. (a) Velocity contrast between the overburden and substratum causes raypath bending at the interface between the two. (b) Replacing the overburden velocity with the velocity of the substratum eliminates raypath bending.

In principle, datuming can be performed by any extrapolation method, such as phase-shift, finite-difference, or Kirchhoff summation. Nevertheless, the Kirchhoff summation is more convenient in handling datum surfaces with arbitrary shapes.

It is important to distinguish between datuming and migration. Datuming produces an unmigrated time section at a specified datum *z(x)*, which can be arbitrary in shape. Migration involves computing the wave field at all depths from the wave field at the surface. In this respect, datuming is an ingredient of migration, when migration is done as a downward continuation process. In addition to downward continuation, migration requires invoking the imaging principle $t = 0$.

Wave-equation datuming has several practical applications: horizon flattening, layer replacement, and forward modeling. These are performed in either prestack or poststack mode. The main difference between the two implementations is that the velocity must be halved when doing poststack datuming to conform to the exploding reflectors model (Section 4.1). Horizon flattening involves downward continuing from one marker horizon to another. When this is done successively for all marker horizons in a section, the technique is useful in reconstructing the past structural history in a given survey area (Taner et al., 1982).

Seismic modeling involves a series of upward continuations through a specified set of velocity interfaces starting at the bottom of the model and ending at the top. An example is provided in Section 8.4.

5.3.1 Poststack Layer Replacement

Now consider another practical application of wave-equation datuming to 2-D surface seismic data: removing the degrading effect of an irregular water bottom topography on the continuity and geometry of reflections below. This problem is particularly severe in areas with a strong velocity contrast between the water layer and the substratum. Despite the usual 3-D nature of the problem, the 2-D interpretation of the target reflections often can be

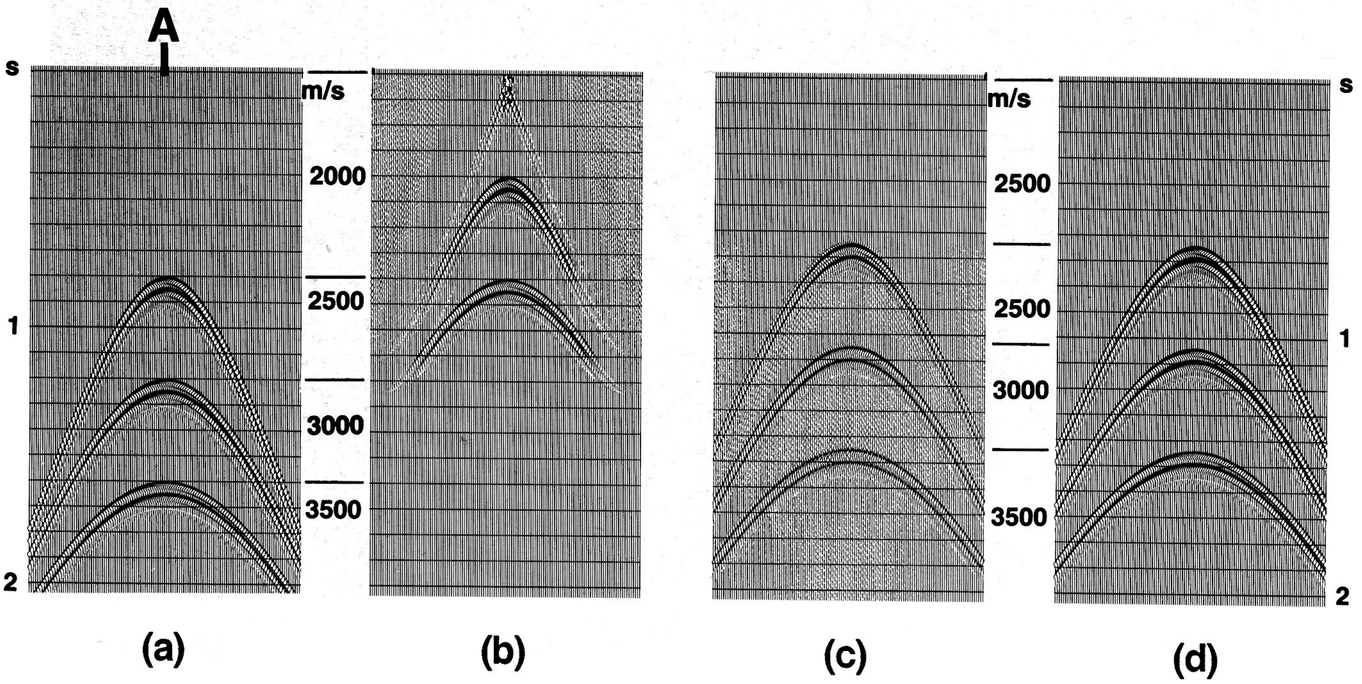


FIG. 5-22. Layer replacement by Kirchhoff summation. (a) Input zero-offset section (at datum level $z = 0$) based on a depth model consisting of three point scatterers buried at 800, 1300, 1900 m depths. Interval velocities are indicated on the right side of the section. (b) Step 1: Downward continuation of the wave field from $z = 0$ to datum level $z = 800$ m using a velocity of 2000 m/s. (c) Step 2: Upward continuation of (b) back to $z = 0$ using a velocity of 2500 m/s. (d) The zero-offset section derived independently using interval velocities that are indicated on the left side of the section. This section should be compared to (c).

improved by replacing the velocity of the water layer with the velocity of the substratum.

The first step in poststack layer replacement involves downward continuing the wave field at the surface (Figure 5-23a) to the water bottom (horizon 2 in Figure 5-21a) using the water velocity in extrapolation. The intermediate result is shown in Figure 5-23b. Note that in this horizon-flattened section, the water-bottom reflection is at $t = 0$, which means that all receivers are situated on the irregular water bottom.

If we specified the overburden velocity or the water-bottom topography incorrectly, then the water-bottom reflection would not be at $t = 0$. In this respect, the intermediate section becomes a useful diagnostic tool before moving to the next step.

The second step in poststack layer replacement involves upward continuation of the intermediate wave field (Figure 5-23b) back to the surface using the velocity of the substratum (2000 m/s). Figure 5-23c is the zero-offset section after layer replacement. A zero-offset section was created from the same model as Figure 5-23a, except that the first layer velocity was 2000 m/s. It is shown in Figure 5-23d for comparison. Comparison of the results from layer replacement (Figure 5-23c) and depth migration (Figure 5-10c) shows that both processes have removed the effects of the complex overburden. However, note that layer replacement only requires accurate representation of the overburden, while depth migration requires accurate representation of the entire velocity-depth model. Also note that the output from depth migration is a migrated depth section (Figure 5-10c), while the output from layer replacement is an unmigrated time section (Figure 5-23c). After eliminating the complex overburden effect, this section only requires time migration.

5.3.2 Prestack Layer Replacement

As with depth migration, layer replacement after stack can remove the effect of complex overburden, provided the input section accurately represents a zero-offset section. However, the complex overburden causes raypath distortions that generate anomalous, nonhyperbolic

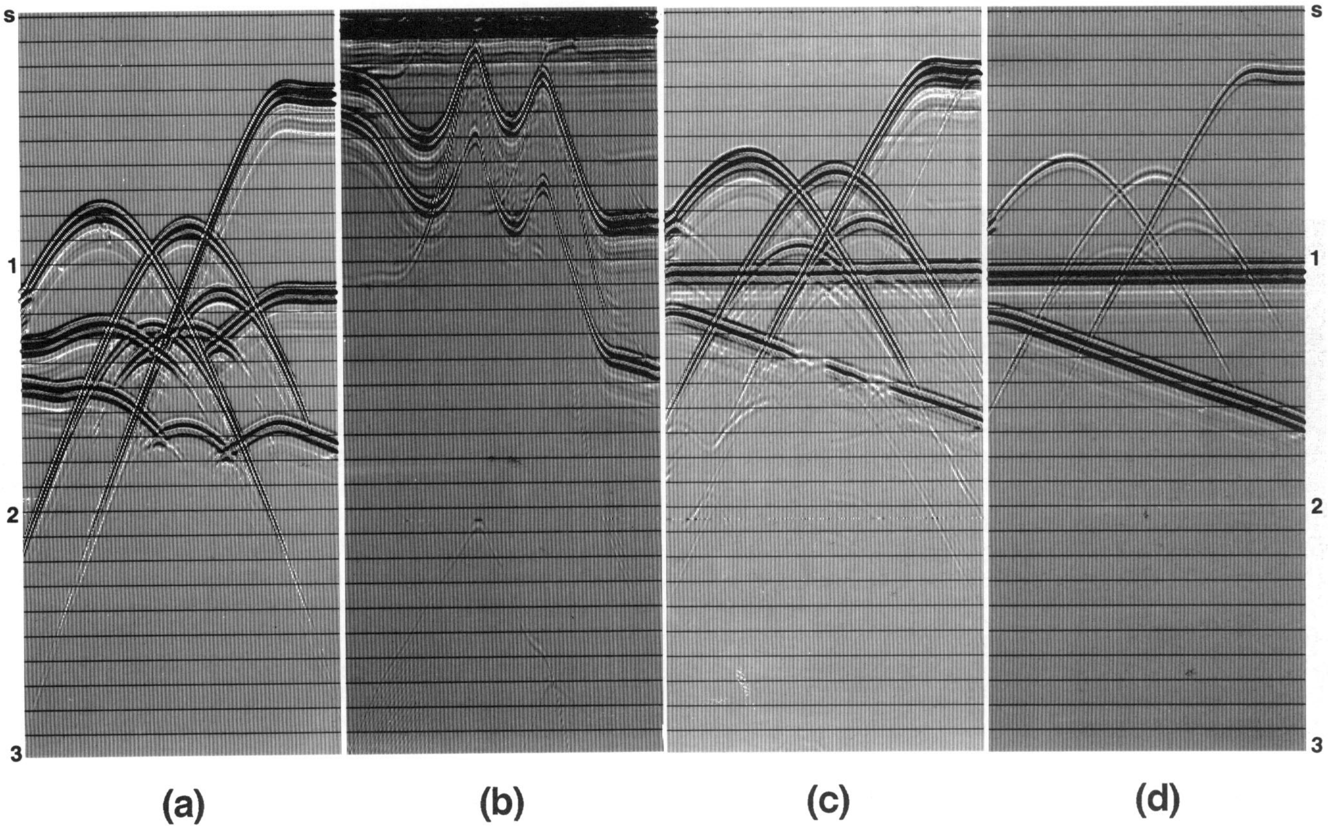
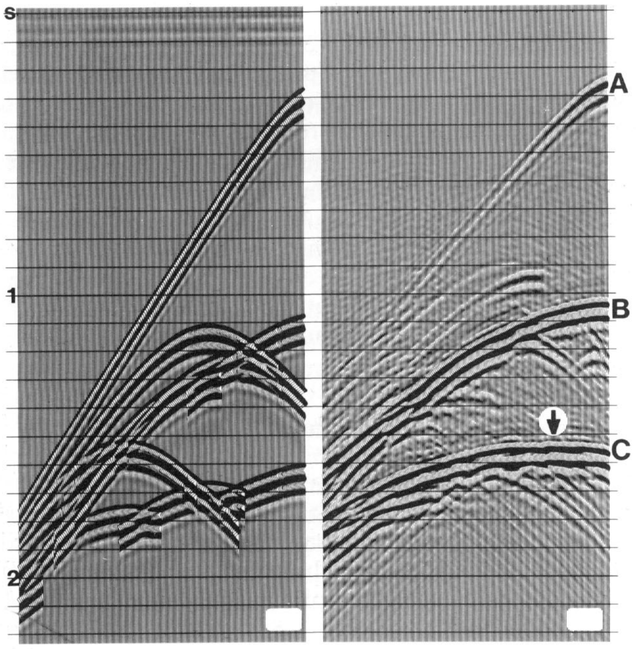


FIG. 5-23. Poststack layer replacement involves two steps. Step 1: The zero-offset section (a) is extrapolated down to the water bottom (horizon 2 in Figure 5-21a) using half the water velocity. Section (b) is obtained when the receivers are placed along the water bottom. Step 2: This intermediate wave field is extrapolated back up to the surface using the velocity of the stratum below the water bottom. The resulting zero-offset section (c) can be compared against the zero-offset section derived independently (d) using the same velocity-depth model as in Figure 5-21a, except the overburden velocity is the same as that of the substratum (2000 m/s).



(a) (b)

FIG. 5-24. (a) A synthetic common-shot gather from the complicated part of the velocity-depth model in Figure 5-6 before (a) and after (b) layer replacement. Limitations in modeling with ray tracing cause abrupt terminations along the moveout curves and amplitude glitches in (a). In turn, these caused the spurious diffractions in (b) during the upward continuation steps.

moveout patterns in prestack data. Depth migration does not produce an accurate image of the subsurface from the stacked section, even when the velocity-depth model is known (Schultz and Sherwood, 1980). Similarly, poststack layer replacement does not remove the effect of complex overburden, even if its geometry is known accurately, because the input stacked section differs from the zero-offset section. Nevertheless, based on simple, ray-trace modeling, we can determine whether poststack layer replacement can delineate the underlying structure. If the effort yields unsatisfactory results, then prestack layer replacement should be performed.

Complex moveout is evident in the common-shot and CMP gathers in Figures 5-24a and 5-25a, respectively. These were synthesized from the depth model in Figure 5-6 by using a modeling program based on ray tracing that neither includes diffractions nor properly models amplitudes. The offset range is 50 to 2387.5 m with 12.5-m receiver spacing. A total of 437 shot gathers was generated, each with 192 traces. The coverage is uniform along the line and is 96-fold.

Starting with the common-shot gathers, prestack layer replacement involves (a) downward continuation of all receivers to the output datum using the overburden velocity, (b) sorting the data to common receiver gathers, (c) downward continuation of shots to the same output datum using the overburden velocity, (d) upward continuation of shots back to the surface using the velocity of the substratum, (e) sorting the data back to common shot gathers, and (f) upward continuation of receivers back to the surface using the substratum velocity.

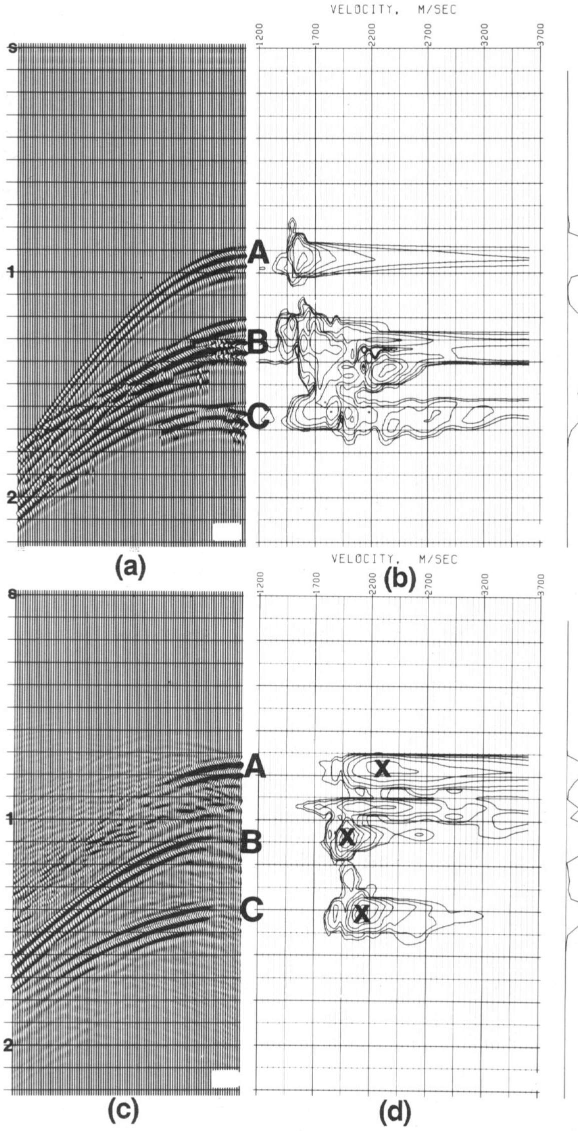


FIG. 5-25. A CMP gather and velocity spectrum from the complicated part of the velocity-depth model in Figure 5-6 before [(a) and (b)] and after [(c) and (d)] layer replacement.

This series of operations eliminates the traveltime distortions associated with the water bottom, as seen in the common-shot and CMP gathers (Figures 5-24b and 5-25c, respectively). Despite undesirable effects due to limitations of modeling with ray tracing, the complexity of the reflections associated with the two events (horizons 3 and 4 in Figure 5-6) was reduced by layer replacement. Once the complex overburden effect is removed, these events have the hyperbolic moveout seen in Figure 5-24b. Events A, B, and C are associated with horizons 2, 3, and 4, respectively, as shown in Figure 5-6. Note that horizon 3 is flat. Therefore, the apex of the hyberbola is at the near-offset trace (event B). On the other hand, horizon 4 dips to the right. Therefore, the apex of the hyberbola shifts updip (as indicated by the arrow in Figure 5-24b).

Note the improvement in the velocity estimates after the layer replacement for reflections beneath the complex overburden (compare Figures 5-25b and 5-25d). The velocity analysis before layer replacement yields a good pick for the water-bottom reflection A, while picks B and C, which are associated with the deeper layers (horizons 3 and 4 in Figure 5-6), are not distinct. Similarly, events B and C also are indistinct on the maximum correlation curve plotted on the right side of the velocity spectrum. After the layer replacement (Figure 5-25d), note that the picks (denoted by x) associated with the three events are distinct in the velocity spectrum and the correlation curve. The dipping water-bottom reflection A now has considerably higher moveout velocity (2300 m/s in Figure 5-25d), as compared to the original (1600 m/s in Figure 5-25b). The velocity pick for the flat event B from Figure 5-25d is 2000 m/s, which is the velocity of the medium above this reflector after layer replacement.

Just as it is true for any wave extrapolation-based process, the wave-theoretical layer replacement technique described here suffers from spatial aliasing. With modern data acquisition, we can record with many channels at small shot and group intervals. Therefore, spatial aliasing should not be a big problem with more recent data. With this provision, prestack layer replacement followed by poststack time migration is a practical alternative to depth migration before stack in areas where simple geology is overlain by an overburden with a strong lateral velocity variation.

The mathematical details of wave field extrapolation based on the Kirchhoff integral are given in Berryhill's original paper (Berryhill, 1979). For poststack layer replacement, we assume that the stack is a zero-offset section, which is equivalent to the exploding reflectors wave field, provided the velocity in the medium is halved (Section 4.1). We assume that sources already are situated along the reflectors in the subsurface. Thus, we only need to move the receivers from one datum to another during downward and upward continuation steps. For prestack layer replacement, each common-shot or common-receiver gather is extrapolated independently. In particular, the wave field at a point on the output datum is computed using all the traces in the input gather. Although not an issue in the present data example, during downward continuation steps of shots and receivers, the output gather should be computed beyond the extent of the input gather to prevent possible loss of steeply dipping events. For prestack layer replacement, the velocity used in the extrapolation is that of the medium between the input datum and the output datum.

Now consider the field data example in Figure 5-12a. Layer replacement will be done on these data before and after stack to remove the effect of the irregular water bottom. First, we must define the geometry of the overburden, in this case, the water-bottom topography. Since the overburden velocity is constant (1475 m/s), we migrate the CMP stack section using the constant-velocity Stolt method, digitize the water bottom, and convert it to depth as shown in Figure 5-26a.

Consider poststack layer replacement. First we downward continue the CMP stack (Figure 5-12a) by using the water velocity from the surface to the water bottom to get the horizon-flattened section in Figure 5-26b. (As usual, we assume that the CMP stack is a zero-offset wave field.) The water-bottom reflection is approximately flat and is situated at $t = 0$, which indicates that the velocity-depth model in Figure 5-26a is fairly accurate. Although incorrect in shape, reflections are quite continuous on this section, indicating that we achieve a focusing effect by downward extrapolation. The bottom of the section reflects the mirror image of the water-bottom topography.

The next step is to take this wave field back to the surface, this time using the velocity of the substratum (2500 m/s). Luckily, the velocity derived from the seismic data is fairly constant across the section for the substratum. The section after upward continuation is the result of layer replacement and is shown in Figure 5-26c. After eliminating the effect of the complex overburden, only time migration is needed to image this section (Figure 5-26d). The section with time migration now can be compared against the poststack depth migration in Figure 5-12d. Despite the different approaches, these two results infer similar structural interpretation. Although not shown, the exact comparison should be made between depth migration and depth-converted time migration.

The results of prestack layer replacement now are considered. Starting with the common-shot gathers and using the depth model in Figure 5-26a, we perform downward and upward continuations on common-shot and common-receiver gathers. The intermediate steps for selected common-shot gathers are shown sequentially in Figure 5-27. Note the arrival time of the water-bottom reflection at $t = 0$ on the gathers in step 2 after downward continuing both the shots and receivers to the water bottom.

A selected set of CMP gathers before and after layer replacement, sorted from the shot gathers in steps 0 and 4, respectively, is shown in Figure 5-28. Because the data contain strong diffracted multiples, it is difficult to evaluate the layer replacement results. As in any other one-way extrapolation technique, wave-equation datuming treats multiples as primaries. From the velocity analysis in Figure 5-29, we are encouraged by the picking we can do on the velocity spectrum after layer replacement. However, the stack has the final say and is shown in Figure 5-30a. Again, once time-migrated, it should provide an accurate image of the substratum, free from the effects of complex overburden (compare Figure 5-30b with Figures 5-12d and 5-26d).

The interpretation of the unconformity *T* from the results of prestack layer replacement (Figures 5-30a and b) closely agrees with the proposed depth model in Figure 5-12c. The unconformity is continuous beneath midpoint C, where there is a structural high. Below and to the left of midpoint A, the unconformity extends down to the right with some tensional faults into the ancient continental shelf below midpoint B, then finally extends to the continental slope of that age to the right of the structural high below midpoint C. The present-day situation is just

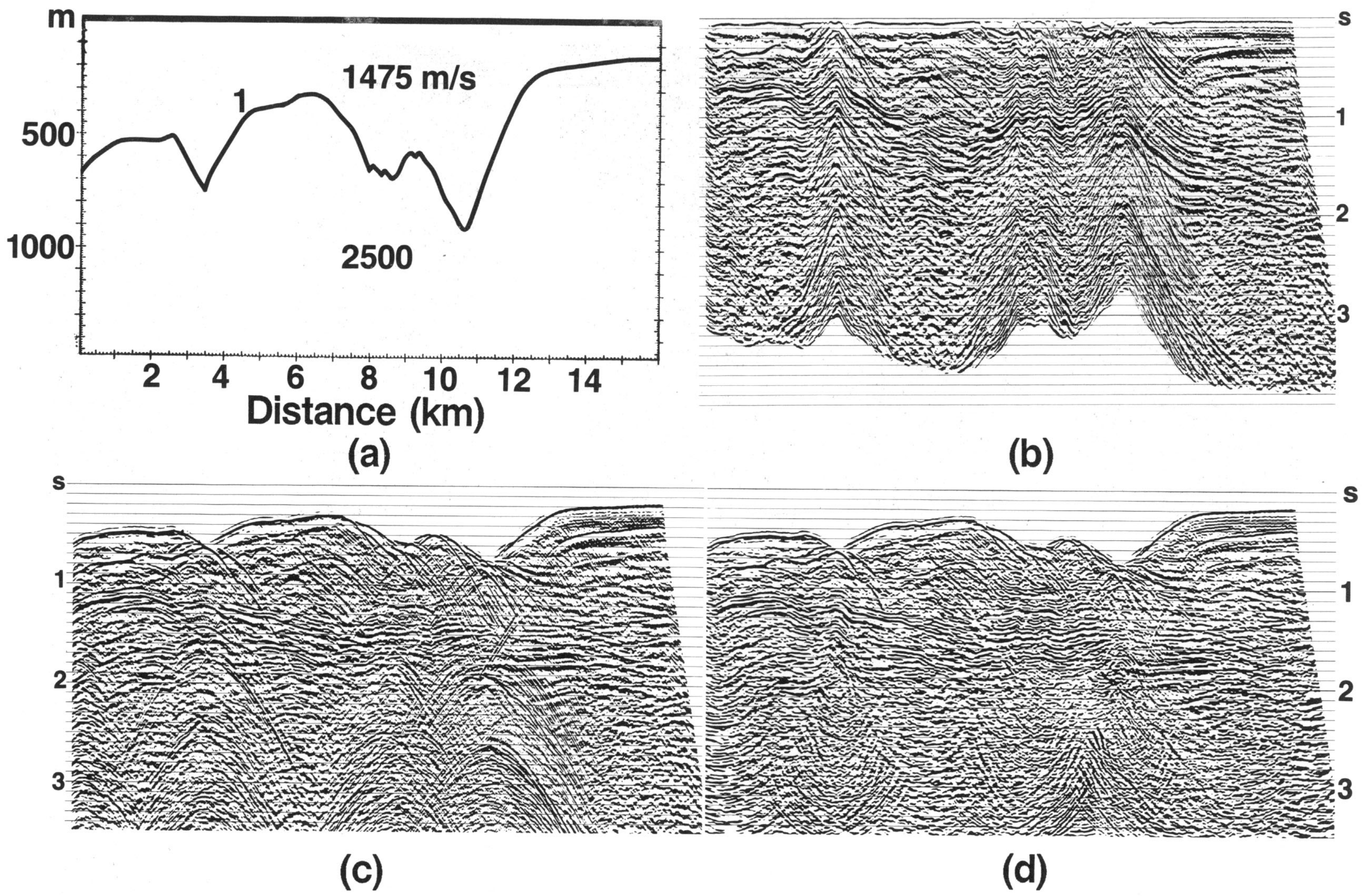


FIG. 5-26. (a) Velocity-depth model derived from the constant-velocity (1475 m/s) Stolt migration of the CMP stack in Figure 5-12a. (b) The stacked section in Figure 5-12a after extrapolating from the surface to the water bottom [horizon 1 in (a)] using water velocity. The water-bottom reflection is at $t = 0$. This is the first step in poststack layer replacement. (c) Upward continuation of the wave field in (b) from the water bottom to the surface using the water-bottom velocity. This is the second step in poststack layer replacement. (d) Time migration of the section in (c). The replaced water layer has a velocity of 2500 m/s. (Data courtesy Hispanoil.)

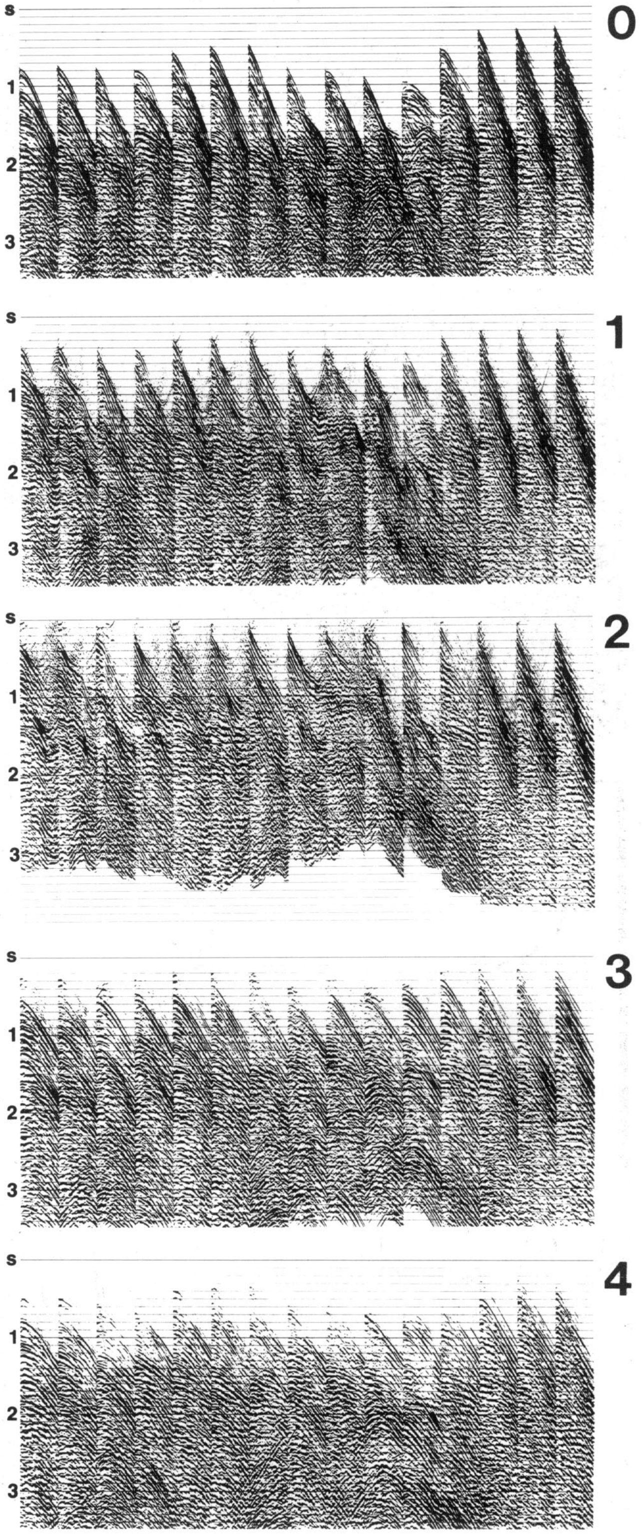


FIG. 5-27. Steps involved in prestack layer replacement. These shot gathers are from the field data example in Figure 5-12a. The near-surface model is shown in Figure 5-26a. (0) Common-shot gathers at the surface, (1) downward continuation of receivers to the water bottom with v = 1475 m/s, (2) downward continuation of shots to the water bottom with v = 1475 m/s, (3) upward continuation of shots to the surface with v = 2500 m/s, (4) upward continuation of receivers to the surface with v = 2500 m/s. (Data courtesy Hispanoil.)

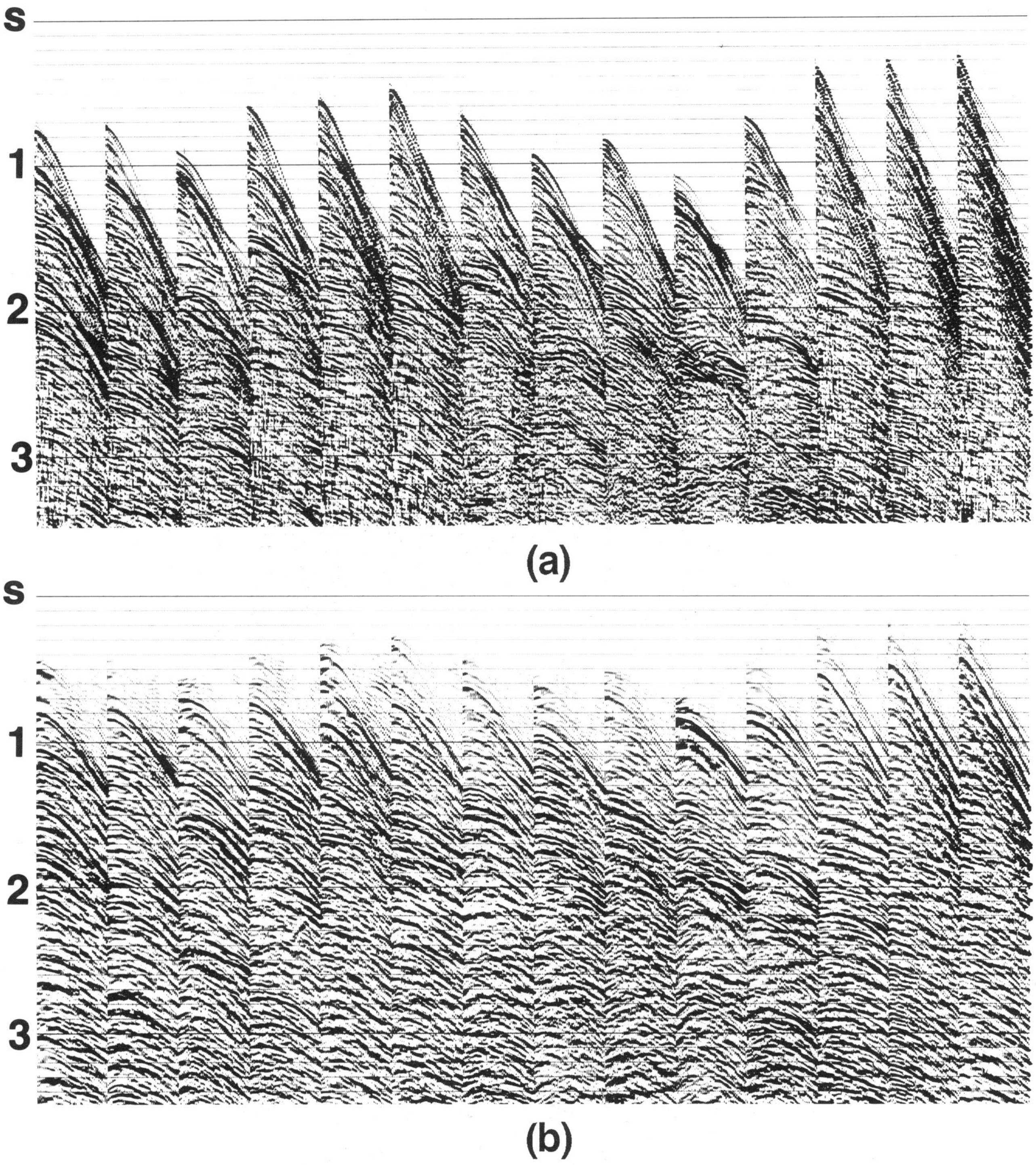


FIG. 5-28. CMP gathers associated with common-shot gathers from steps 0 and 4 in Figure 5-27 (a) before and (b) after layer replacement. (Data courtesy Hispanoil.)

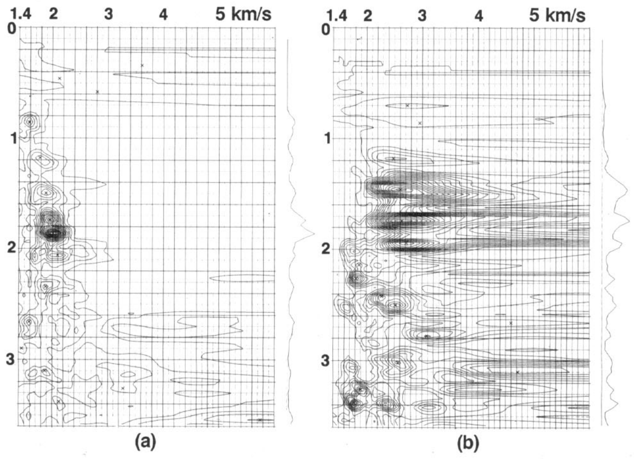


FIG. 5-29. Velocity analysis in the vicinity of midpoint C in Figure 5-12a (a) before and (b) after layer replacement.

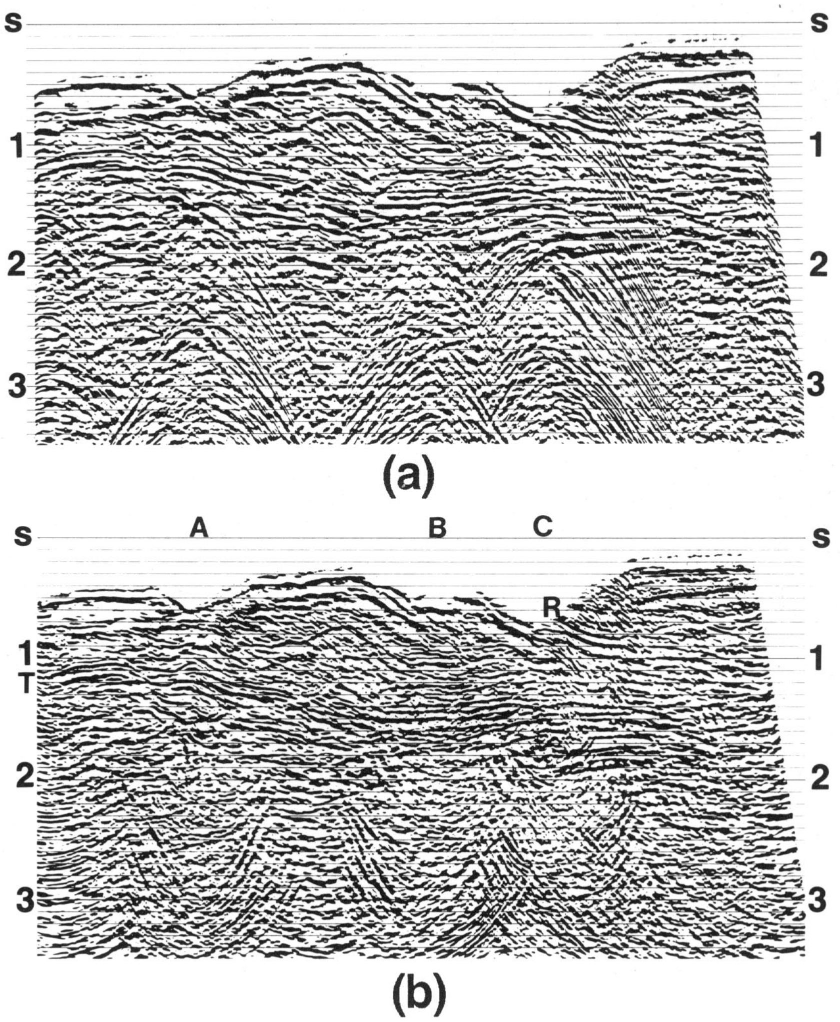


FIG. 5-30. (a) CMP stack derived from the gathers (Figure 5-28b) after layer replacement. Compare this with the conventional stack and poststack layer replacement result in Figures 5-12a and 5-26c, respectively. (b) Time migration of the stacked section in (a). Compare this with time migration of the poststack layer replacement result and depth migration after stack in Figures 5-26d and 5-12d, respectively. (Data courtesy Hispanoil.)

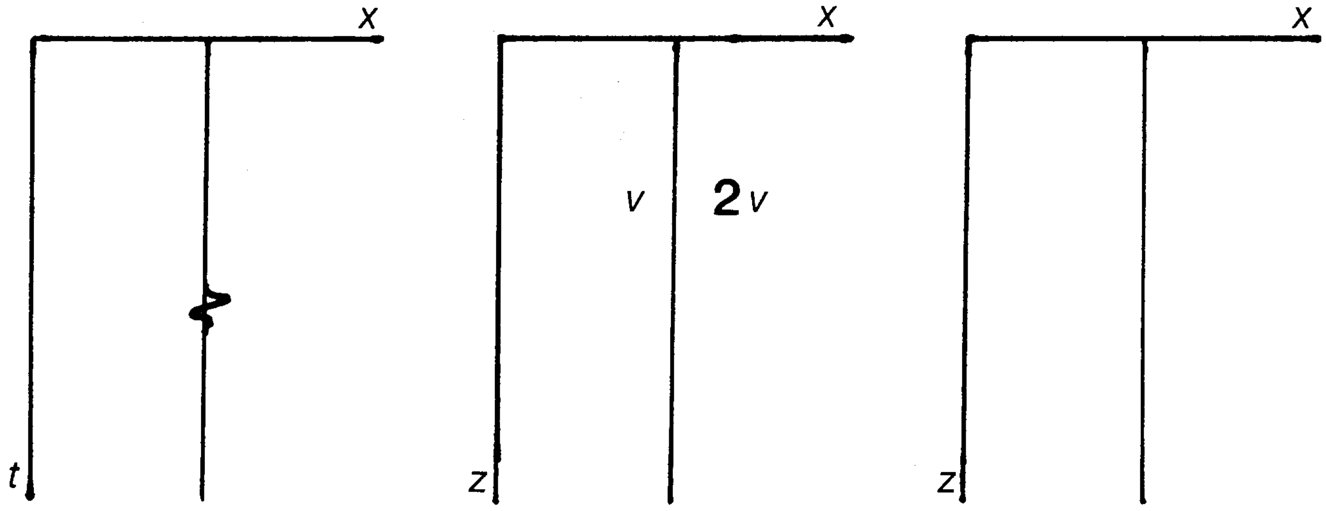


FIG. 5-31. (a) A zero-offset section containing all-zero traces except the one shown, which has a single isolated wavelet; (b) the velocity-depth model for depth migration; (c) the (x,z) plane (depth-migrated section) see Exercise 5.1).

the opposite, with the continental shelf on the right. Note that shallow reflector R now is conformable with the rest of the shallow section above the unconformity.

For the irregular water-bottom case, it was shown that prestack layer replacement followed by NMO, stack, and time migration after stack (with depth conversion) basically is equivalent to full depth migration before stack. Full depth migration before stack is costly, while prestack layer replacement provides an unmigrated stacked section and an opportunity to upgrade velocity analyses once the effect of the complex moveout is eliminated.

The situation may be more complicated when dealing with a complex overburden problem that involves more than one interface. In principle, datuming can be applied layer by layer and the effect of the overburden can be removed. This could be costly; a depth migration might just as well be done. Most seismic data conform to hyperbolic moveout. Nonhyperbolic moveout often can be handled by conventional CMP stack. If a good quality stack results, then imaging after stack is adequate. The case is rare where moveout is so complex that there is no choice but to resort to the prestack datuming or migration to solve the problem.

REFERENCES

Berryhill, J.R., 1979, Wave-equation datuming: Geophysics, **44**, 1329-1344.

Berryhill, J.R., 1984, Wave-equation datuming before stack: Presented at the 54th Ann. Int. Soc. Explor. Geophys. Mtg.

Hubral, P., 1977, Time migration—some ray-theoretical aspects: Geophys. Prosp., **25**, 738-745.

Judson, D.R., Lin, J., Schultz, P.S., and Sherwood, J.W.C., 1980, Depth migration after stack: Geophysics, **45**, 376-393.

Larner, K.L., Hatton, L., and Gibson, B., 1981, Depth migration of imaged time sections: Geophysics, **46**, 734-750.

Schultz, P.S. and Sherwood, J.W.C., 1980, Depth migration before stack: Geophysics, **45**, 361-375.

Stolt, R.H., 1978, Migration by Fourier transform: Geophysics, **43**, 23-48.

Taner, M.T., Cook, E.E., Neidell, N.S., 1982, Paleo-seismic and color acoustic impedance sections: applications of downward continuation in structural and stratigraphic context: Presented at the 52nd Ann. Int. Soc. Explor. Geophys. Mtg.

Yilmaz, O. and Lucas, D., 1986, Prestack layer replacement: Geophysics, **51**, 1355-1369.

EXERCISES

Exercise 5.1. Consider the zero-offset (x,t) section depicted in Figure 5-31. Using the velocity-depth model in the same figure, sketch the depth-migrated (x,z) section.

Exercise 5.2. Correlate the horizons on the velocity-depth model in Figure 5-15 (bottom left) with the section in Figure 5-15 (bottom right).

6

3-D Seismic Exploration

6.1 INTRODUCTION

Subsurface geological features of interest in hydrocarbon exploration are three dimensional in nature. Examples are salt diapirs, overthrust and folded belts, major unconformities, reefs, and deltaic sands. A 2-D seismic section is a cross section of a 3-D seismic response. Despite the fact that a 2-D section contains signal from all directions, including out-of-plane of the profile, 2-D migration normally assumes that all the signal comes from the plane of the profile itself. Although out-of-plane reflections (sideswipes) often are recognizable by the experienced seismic interpreter, the out-of-plane signal often causes 2-D migrated sections to mistie. These misties are due to inadequate imaging of the subsurface resulting from the use of 2-D rather than 3-D migration. On the other hand, 3-D migration of 3-D data provides an adequate and detailed 3-D image of the subsurface, leading to a more reliable interpretation.

Special attention must be given to 3-D survey design and acquisition. A typical marine 3-D survey is carried out by shooting closely spaced parallel lines (line shooting). A typical land or shallow water 3-D survey is done by laying out a number of receiver lines parallel to each other and placing the shotpoints in the perpendicular direction (swath shooting).

In marine 3-D surveys, the shooting direction (boat track) is called the *in-line direction*; for land 3-D surveys, the receiver cable is along the in-line direction. The direction that is perpendicular to the in-line direction in a 3-D survey is called the cross-line direction. In contrast to 2-D surveys in which line spacing can be as much as 1 km, the line spacing in 3-D surveys can be 50 m or less. This dense coverage requires an accurate knowledge of shot and receiver locations.

The size of the survey area is dictated by the areal extent of the subsurface target zone and the aperture size required for adequate imaging of that target zone. This imaging requirement means that the areal extent of a 3-D survey almost always is larger than the areal extent of the objective. A few hundred thousand to a few million traces normally are collected during a 3-D survey. Because 3-D surveys are costly, most are aimed at detailed delineation of already discovered oil and gas fields.

The basic principles of 2-D seismic data processing still apply to 3-D processing. In 2-D seismic data processing, traces are collected as common-midpoint (CMP) gathers. In 3-D data processing, traces are collected as common-cell gathers (bins). These gathers are used in velocity analysis and common-cell stacks are generated. A common-cell gather coincides with a CMP gather for swath shooting. Typical cell sizes are 25 × 25 m for land surveys and 12.5 × 37.5 m for marine surveys.

Conventional 3-D recording geometries often complicate the process of stacking the data in a common-cell gather. Cable feathering in marine 3-D surveys can result in traveltime deviations from a single hyperbolic moveout within a common-cell gather. For land 3-D surveys, azimuth-dependent moveout (Section 3.2.3) within a common-cell gather is an issue.

After stacking, the 3-D data volume is often (but not always) migrated in two stages. First, a 2-D migration is applied along the in-line direction. Then the data are sorted and a second pass of 2-D migration is applied along the cross-line direction. Before the second pass of migration, the data sometimes need to be trace-interpolated along the cross-line direction to avoid spatial aliasing.

The 3-D data volume then is available to the interpreter as vertical sections in both the in-line and cross-line directions and as horizontal sections (time slices). The time slices allow the interpreter to generate contour maps for marker horizons. The interactive environment provides an effective and efficient means for interpretation of the sheer volume of 3-D migrated seismic data. Fault correlations, horizon tracking, horizon flattening, and some image processing techniques can be adapted to the interactive environment to help improve interpretation.

6.2 WHY 3-D?

Consider the earth model in Figure 6-1, which consists of a dipping plane interface in a homogeneous medium. Examine line A along the dip direction. If the survey consisted of lines parallel to the dip direction, then a 2-D assumption about the subsurface would be valid. No out-of-plane signal would be recorded and 2-D migration of these dip lines would be correct, as shown in Figure 6-2a. After 2-D migration, point D beneath surface point X is migrated updip to its true subsurface position D′. Now consider line B, which is in the strike direction in Figure 6-1. (Line B ties line A at surface position X.) The reflection from the subsurface point D′ in Figure 6-2a is recorded on both lines A and B at their intersection point X. The event on line B is an out-of-plane reflection, while the event on line A is not. However, the reflection on the strike line from the dipping interface shows no dip (Figure 6-2b). Since migration does not alter the position of flat events, the migrated section for the strike line is identical to the corresponding unmigrated section. If we tied dip line A with strike line B after migrating both sections, a mistie would occur (Figure 6-2b). (However, the unmigrated sections will tie.)

The subsurface generally dips in many directions in areas where structural traps are being explored. Thus, it is not possible to identify the in-line direction as either a dip or strike direction. This is the case for line C in Figure 6-1. The apparent dip perceived on this line is less than the true dip of the plane interface that is perceived along the dip line. Examine the positioning after migration of point D beneath intersection point X for all three lines in the plan view in Figure 6-3. Point D is moved to the true subsurface position D′ along dip line A. The same point does not move after migration along strike line B. It moves to D″ along line C. A second pass of migration is suggested in the direction perpendicular to line C to move the already migrated energy from D″ to its true subsurface position D′.

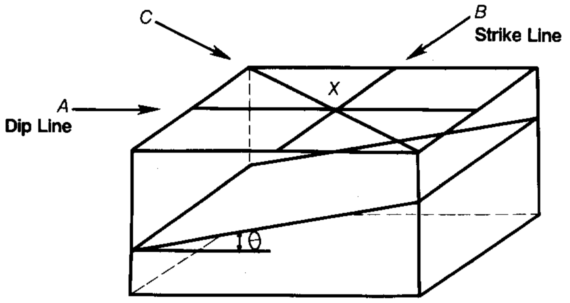


FIG. 6-1. A subsurface model consisting of a single dipping plane interface in a homogeneous medium. Line A is in the dip direction, line B is in the strike direction, and line C is in an arbitrary direction. After migration, data recorded at the intersection point X on these three lines migrate to positions corresponding to different subsurface locations. Schematic illustrations of the migrations are shown in Figures 6-2 and 6-3. (Adapted from Workman, 1984.)

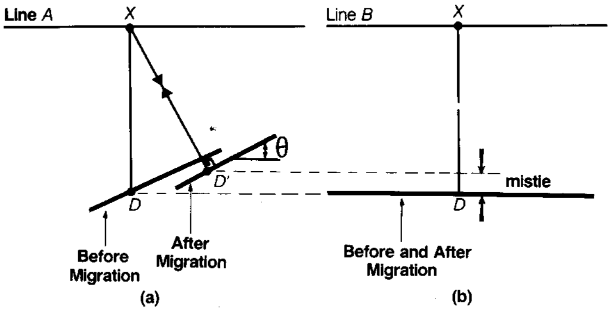


FIG. 6-2. (a) Migration along the dip line and (b) along the strike line over the depth model in Figure 6-1. Point D after migration is moved updip to D' along dip line A. Point D does not move after migration along strike line B. This causes the mistie indicated between the two migrated sections.

Although principles of 3-D migration are discussed in Section 6.5, we need to assess the interpretational differences between 2-D and 3-D migrations. Figure 6-4 shows an in-line (left column) and a cross-line (right column) stacked section from a land 3-D survey and their 2-D and 3-D migrations. Note that 3-D migration has yielded a better definition of the top (T) of the salt dome and better delineation of the faults along the base (B) of the salt dome. There is no doubt that the interpretation based on 2-D imaging is significantly different from interpretation based on 3-D imaging.

Figure 6-5 is another example of the significant improvement available from the interpretation of a 3-D migrated section. Note that the two salt domes and the syncline between are delineated better after 3-D migration. Three-dimensional migration often produces surprisingly different sections from 2-D migrated sections. The example in Figure 6-6 shows a no-reflection zone on the 2-D migrated section, while the same zone contains a series of continuous reflections on the 3-D migrated section that are easily correlated with reflections outside that zone.

As noted earlier, 2-D migration can introduce misties between 2-D lines in the presence of dipping events. Two-dimensional migration cannot adequately image the subsurface, while 3-D migration eliminates these misties by completing the imaging process. This is demonstrated in Figure 6-7 in which the correlation of an in line and a cross line can be examined at their intersection (indicated by the vertical bar). The slight mistie problem that is evident on the 2-D migrated sections (especially apparent between 1.3 and 2 s) is eliminated on the 3-D migrated sections.

From the field data examples, we see that 3-D migration provides complete imaging of the 3-D subsurface geology. In contrast, 2-D migration can yield inadequate results. The difference between 2-D seismic and 3-D seismic is the way in which migration is performed.

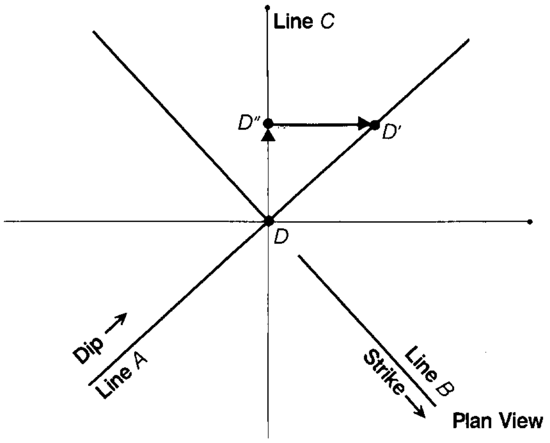


FIG. 6-3. Plan view of the migrations at the intersection point along the three lines indicated in Figure 6-1. Point D moves to D', its true subsurface position, along dip line A. Point D does not move on strike line B. The same point moves to D'' along line C, which is in an arbitrary direction. Complete imaging is achieved by migrating the data again along the direction perpendicular to C to move the energy from D'' to D'.

Dense coverage on top of the target zone, say a 25-m in-line trace spacing and a 25-m cross-line trace spacing, will not necessarily provide adequate subsurface imaging unless migration is performed in a 3-D sense.

Figure 6-8 shows an increasingly modified and improved interpretation made from seismic data that were obtained from detailed 2-D surveys in an area between 1964 and 1970. The reconnaissance survey in the first year of exploration (1964) consisted of only a few lines. A preliminary time structure map based on this initial survey inferred a structural closure with a northwesterly trend. In the second year of exploration (1965), more lines were shot in the same directions as before and the structural closure was verified somewhat. Lines were

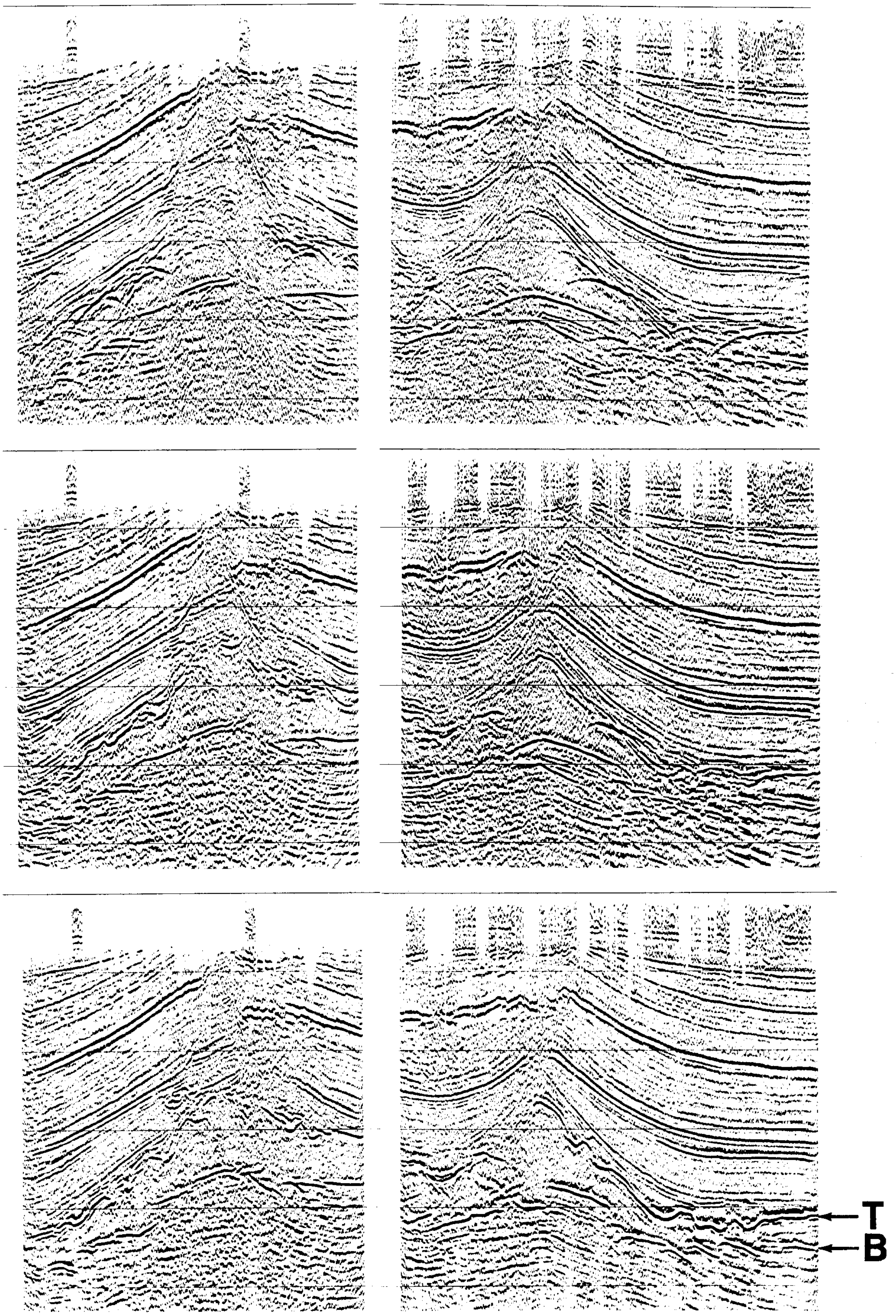


FIG. 6-4. An in-line (top left) and a cross-line (top right) stacked section from a land 3-D survey. Also shown are the 2-D (center row) and 3-D (bottom row) migrations of the two stacked sections. (Data courtesy Nederlandse Aardolie Maatschappij B.V.)

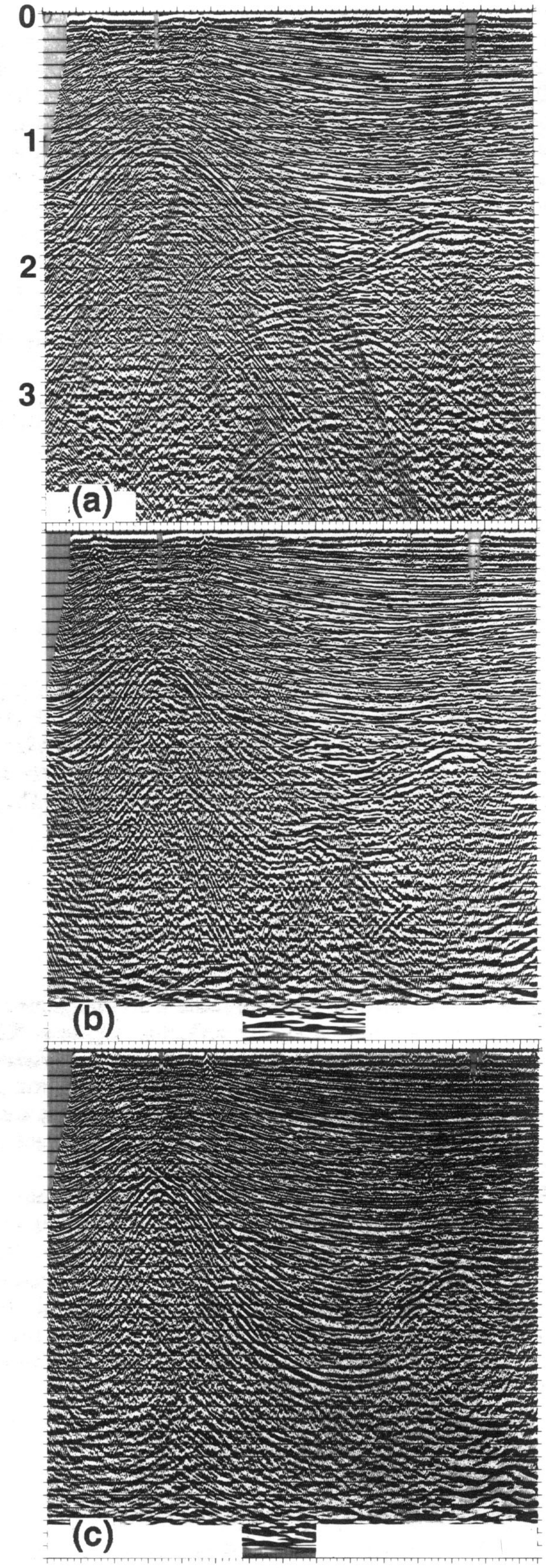


FIG. 6-5. (a) An in-line stacked section from a marine 3-D survey, (b) 2-D migration, (c) 3-D migration. (Data courtesy Amoco Europe and West Africa, Inc.)

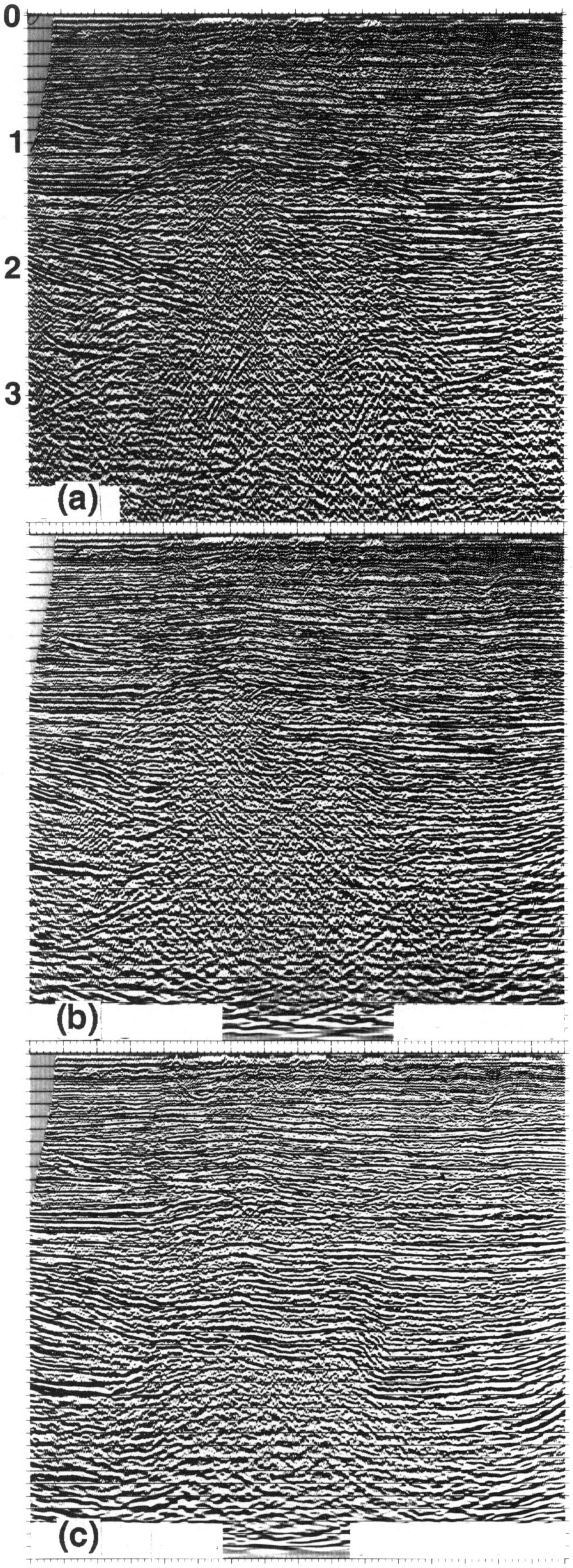
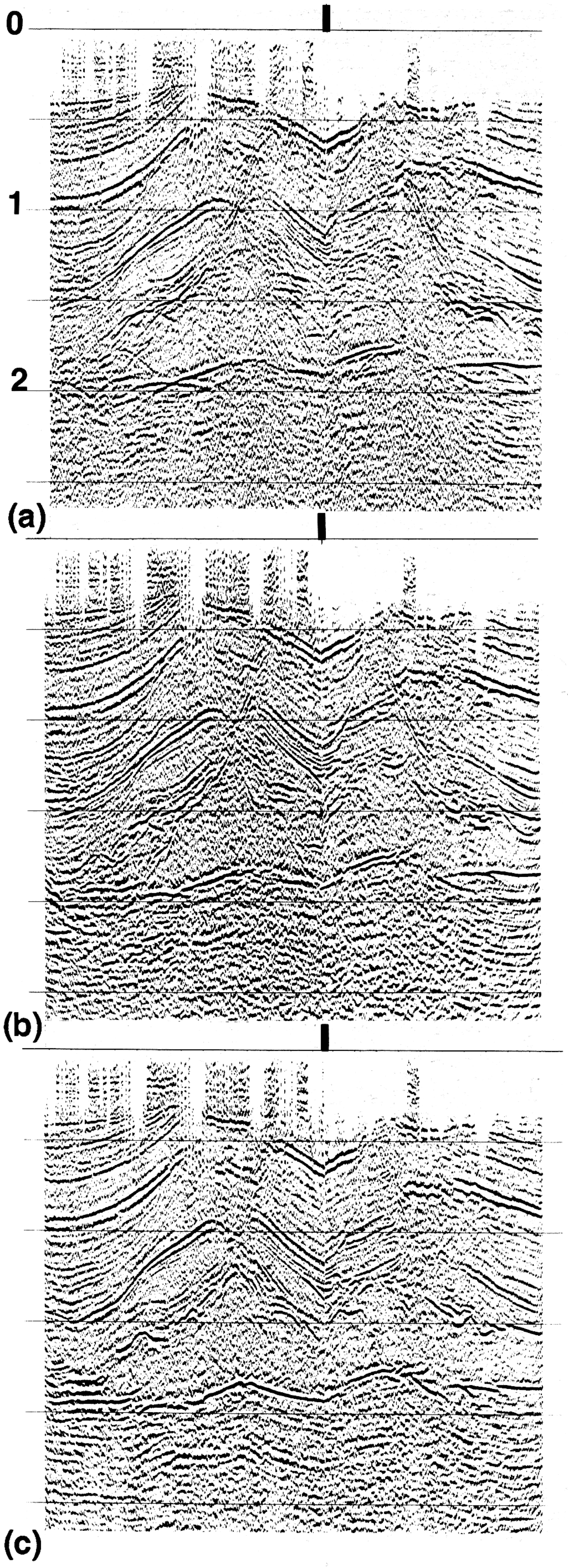


FIG. 6-6. (a) Another in-line stacked section from the same marine 3-D survey as in Figure 6-5, (b) 2-D migration, (c) 3-D migration. (Data courtesy Amoco Europe and West Africa, Inc.)



◄FIG. 6-7. Two intersecting perpendicular lines (intersection point indicated by the vertical bar). These data are from the same 3-D survey as in Figure 6-4. (a) Stacked sections (an in-line section on the left and a cross-line section on the right), (b) 2-D migrations, (c) 3-D migrations. (Data courtesy Nederlandse Aardolie Maatschappij B.V.)

placed in the northeasterly direction in the third year (1966) to delineate the structure. This change in survey direction made it possible to map some faults in the area. In the following years (1967, 1968, and 1970), more lines were shot in the northeast-northwest and southwest-southeast directions. These additional surveys allowed denser coverage of the subsurface and led to more detailed, improved interpretational results. As we examine the location map of 1970, which includes all the lines surveyed in the area between 1964 and 1970, we must ask whether it would have been better to conduct a 3-D survey and image the subsurface in a 3-D sense. Three-dimensional surveying can lead to a high degree of accuracy and reliability in interpretation. Since the 3-D data volume provides detailed constraints on interpretational decisions, drilling programs based on 3-D usually have high success rates.

6.3 3-D SURVEY DESIGN AND ACQUISITION

The ultimate goal of a 3-D survey is to obtain a 3-D migrated wave field. The fidelity of this migration depends on stack quality and accuracy in velocity estimation. However, two other factors control the fidelity of migration: migration aperture and spatial sampling. They also can dictate the design of the field survey.

6.3.1 Migration Aperture

Figure 4-14a shows a depth profile that represents a subsurface model with dipping reflector segment CD buried in a homogeneous medium. Zero-offset modeling using normal-incidence rays yields the time section in Figure 4-14b. Although not shown in the figure, the time section also should include diffractions off the edges of the reflecting segment.

Migration moves event C′D′ on the time section to its true subsurface position CD, which is overlaid on the time section for comparison. The horizontal extent of the zone of interest is OA. If the line length were confined to OA during recording, then the time section would be blank. On the other hand, if the recording were confined to line segment AB, then event C′D′ would be absent from the migrated section. Although the target is confined to line segment OA, the time section must be recorded over a longer segment OB. The line length also must be long enough to include a significant part of the diffractions that may be present in the data. Additionally, the recording time must be long enough to include diffraction tails and all of the dipping events of interest. The spatial (in the horizontal direction) and temporal (in the vertical direction) displacement of a point on a dipping event resulting from migration depends on medium velocity, depth, and dip of the event [Figure 4-15 and equations (4.1), (4.2), and (4.3)]. Thus, line length and line position on the surface must be chosen carefully based on the migration aperture needed to adequately image the target zone of interest.

These considerations apply just as well to 3-D surveys. Figure 6-9 is depth contour map of a fictitious structural high. The subsurface extent of the objective portion of structure is indicated by the smaller rectangle. Using the principles discussed above and in Section 4.2 (based on Figure 4-14), the actual survey size needed to define the objective area is outlined by the larger rectangle.

Note that the survey area does not have to be extended equally in all directions. The northern flank of the structure is the steepest part. Therefore, the survey area must be extended most in that direction. Extensions in other directions are determined accordingly. Another consideration in extending the survey area is the required additional length in profile to achieve full-fold coverage over the already extended survey area. A typical subsurface anomaly with a lateral extent of, say, 3 × 3 km may require a 3-D survey over an area as large as 9 × 9 km.

6.3.2 Spatial Sampling

Spatial aliasing was discussed in detail in Section 1.6.1 and in relation to migration in Section 4.3.5. The spatial aliasing problem is caused by coarse spatial sampling of the wave field to be migrated; i.e., the stacked section. The spatial sampling of final stacked data (without trace interpolation) is defined by the recording parameters. Therefore, receiver spacing, cross-line spacing, and the

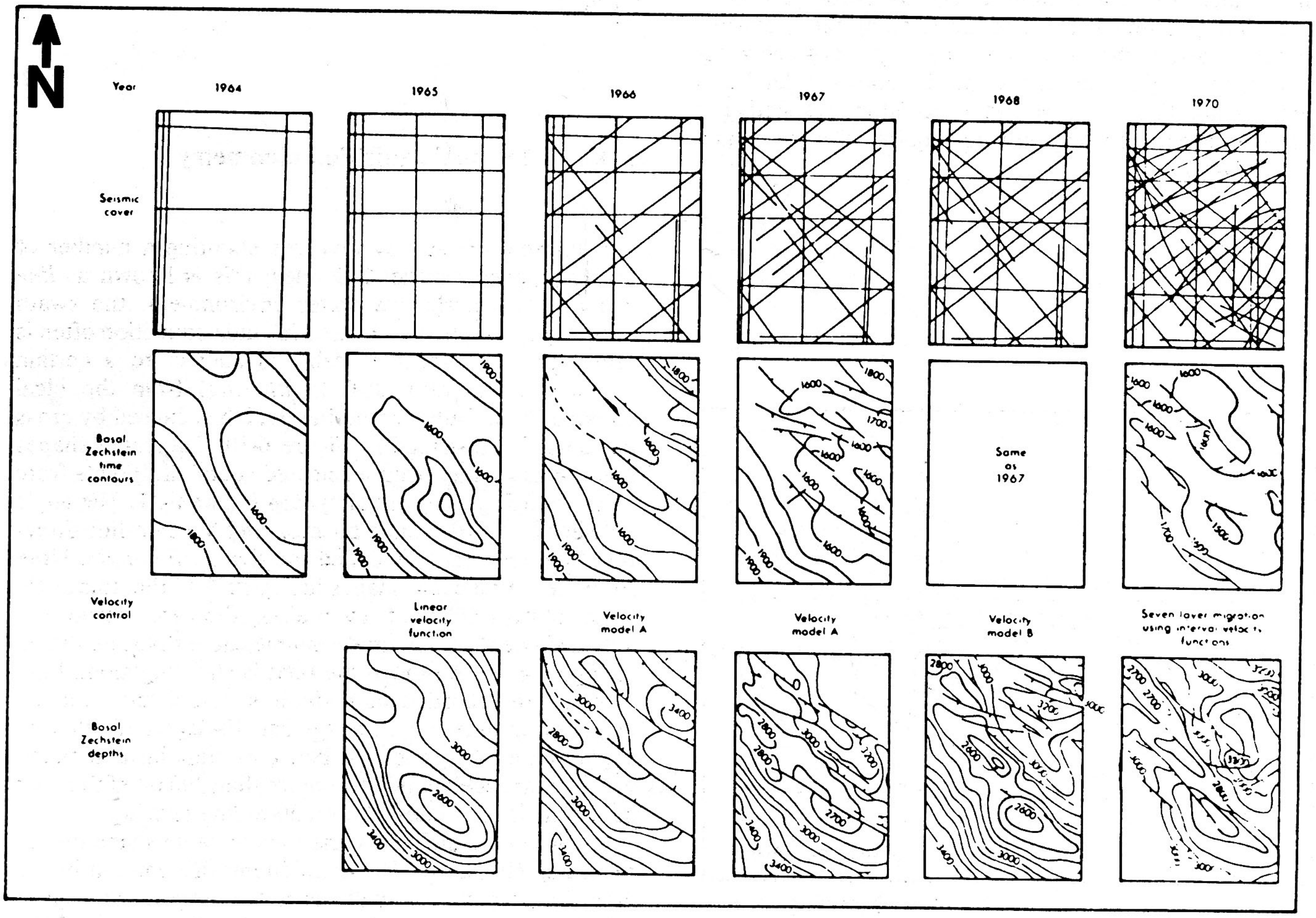


FIG. 6-8. Progress of seismic exploration over a structure from 1964 to 1970. Top row: Base maps of seismic lines. Middle row: Time structure maps. Bottom row: Depth structure maps. As more data were acquired, the interpretation was improved and modified accordingly (Homabrook, 1974).

cross-line direction used in the field must be chosen carefully.

From Figure 4-107, note that a relationship exists between the trace spacing on a stacked section, dip, and the frequency at which spatial aliasing begins to occur. Imagine normal-incidence rays recorded at two receivers, A and B. In the constant-velocity case, the angle between the surface and the wavefront is the true dip of the reflector from which these rays emerged. There is a time delay equivalent to travel path CB between the receivers at A and B. If this time delay is half the period of a given frequency component of the signal arriving at the receivers, then that frequency is at the threshold of being aliased. From the relationship indicated by Figure 4-107, note that the maximum frequency that is not aliased gets smaller at increasingly steeper dips, lower velocities, and coarser trace spacing. From this relationship, an optimum trace spacing can be derived for the in-line and cross-line directions, provided a regional velocity field and the subsurface dips are known. Typical trace spacings in the in-line and cross-line directions in marine 3-D surveys are 12.5 to 25 m and 37.5 to 75 m, respectively. Even if the trace spacing in the cross-line direction is as small as possible, it usually is greater than that in the in-line direction, for economic reasons. Because of this, trace interpolation may be required along the cross-line direction before the migration stage. Typical trace spacings in the in-line and cross-line directions in land 3-D surveys are 12.5 to 25 m and 25 to 50 m, respectively. Trace interpolation may not be required for some land 3-D surveys.

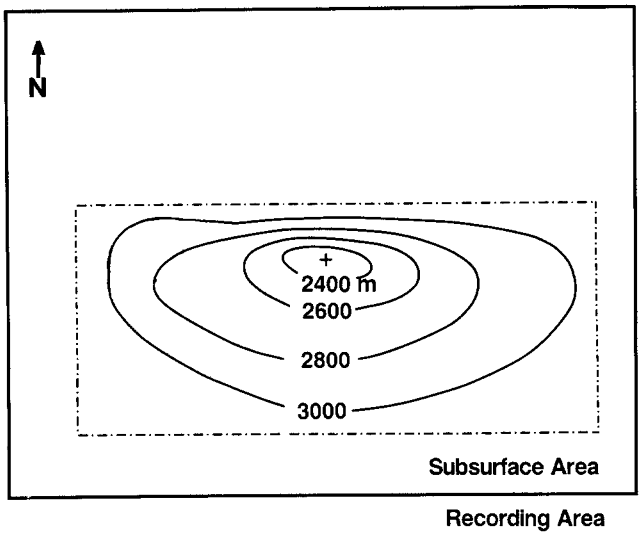


FIG. 6-9. To account for migration aperture, the size of a 3-D survey over a subsurface structure normally is greater than that of the structure itself.

6.3.3 Other Considerations

Almost all of the field operational aspects of 2-D acquisition are applicable to 3-D surveys. For example, selection of navigation and recording equipment depends on field conditions. The operating environment also must be considered. In the marine environment, water depth, tides, currents, sea conditions, fishing and shipping activity, and obstacles such as drilling platforms, wrecks, reefs, and fish traps must be considered. On land, environmental restrictions, accessibility, topography, cultivation, and demographic restrictions are factors that can affect survey design and acquisition. Because of these restrictions, careful planning and some adjustment of nominal shooting geometry often are required to achieve acceptable fold and offset distribution. Dual-boat operations can achieve more uniform coverage and shorten field time. Accurate surveying is a necessity for 3-D surveys, since data are collected with such dense spatial sampling; statics resulting from line-to-line surveying errors can seriously degrade 3-D migrations. In fact, some claim that positioning error, rather than economics, is the limiting factor for line spacing in marine 3-D surveys.

6.3.4 Marine Acquisition Geometry

A marine 3-D survey involves shooting a number of closely spaced parallel 2-D lines; this is known as *line shooting*. (In a shallow water environment, the swath shooting technique that is used in land acquisition often is preferred.) The receiver cable is subject to a certain amount of sideways drift (*feathering*) from the ideal streamer line. Cable feathering, which is caused by cross currents, is illustrated in Figure 6-10. The cable shapes that are associated with a selected set of shotpoints from a marine 3-D survey are shown in Figure 6-11. The angle between the actual cable position and the shot-line direction (the boat track) is called the *feathering angle*. Note from the actual cable shapes in Figure 6-11 that this angle is not always constant, even along the cable associated with a single shot. Assume a simple cable shape as shown in Figure 6-10. Although the boat is shooting along Line 2, data are recorded at midpoints associated with the neighboring lines. For a typical 10-degree feathering angle and a 2400-m-long cable, the midpoint associated with the far receiver is offset more than 200 m off the shot line. This is four lines off at a 50-m line spacing.

Because of significant variations in cable shape during recording (Figure 6-11), the midpoint distribution in the cross-line direction usually is not as regular as Figure 6-10 implies. We must know exactly where each of the receivers is located along the cable, as well as shotpoint location. By precisely determining the shot and receiver coordinates, we can determine the midpoint locations and gather them into cells (Figure 6-10) for stacking and migration.

Navigation data collected on the survey vessel normally include boat location, source location, and cable compass readings. There are 8 to 12 digital compasses along a typical marine cable. Readings from these devices allow computation of the (x,y) coordinates of the cable compasses. Cable shape then is computed based on a curve-fitting procedure that rejects any anomalous measurements. Navigation data are analyzed during processing; quality control is carried out to derive the final shot-receiver locations.

Cable shapes and source locations from a marine 3-D survey are shown in Figure 6-12 in their entirety. Note the cable feathering that is especially apparent at the edges of the survey. Midpoint locations associated with the source-receiver locations in this figure are plotted in Figure 6-13. There are holes in coverage over the survey area. These low-fold areas must be filled in during acquisition by shooting more lines at appropriate locations. If these deficiencies are discovered in processing, it is costly to send the vessel back for further acquisition. Quality control work therefore must be carried out on board to monitor coverage.

6.3.5 Land Acquisition Geometry

Land 3-D acquisition commonly is carried out by swath shooting in which receiver cables are laid out in parallel lines (in-line direction) and shots are positioned in a perpendicular direction (cross-line direction). Figure 6-14 shows swath shooting with 6 receiver cables, each with 80 receiver groups, 50-m apart. The receiver cables are spaced apart from left to right at 100, 200, 100, 200, and 100 m distances. Shooting is perpendicular to the swath starting from the far left and moving in and away to the right of the swath. The swath shooting method yields a wide range of source-receiver azimuths, which can be a concern during velocity analysis (Section 6.4.2). (The source-receiver azimuth is the angle between a reference line, such as a receiver line or a dip line, and the line that passes through the source and receiver stations.) The main advantage of swath shooting is that it is economical. As one shot line is completed, the receiver cables are moved up along the swath a number of stations (equivalent to the shot-line spacing) and shooting is repeated. The shooting geometry in Figure 6-14 provided a 25 × 25-m common-cell gather. Once one swath is

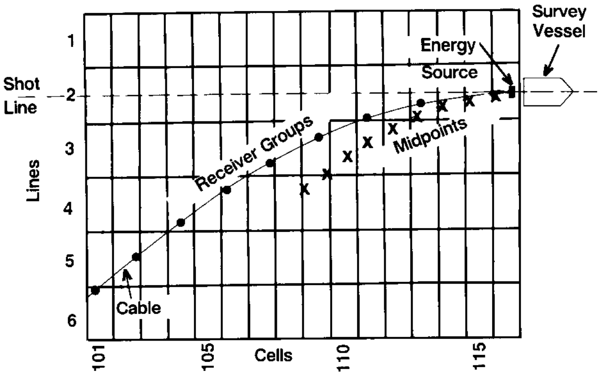


FIG. 6-10. Crosscurrents cause the marine cable to feather and drift sideways from the shot-line direction. This causes the midpoints to spread in the cross-line direction. When data are sorted into common-cell gathers, each cell contains midpoints associated with more than one source line. (Adapted from Workman, 1984.)

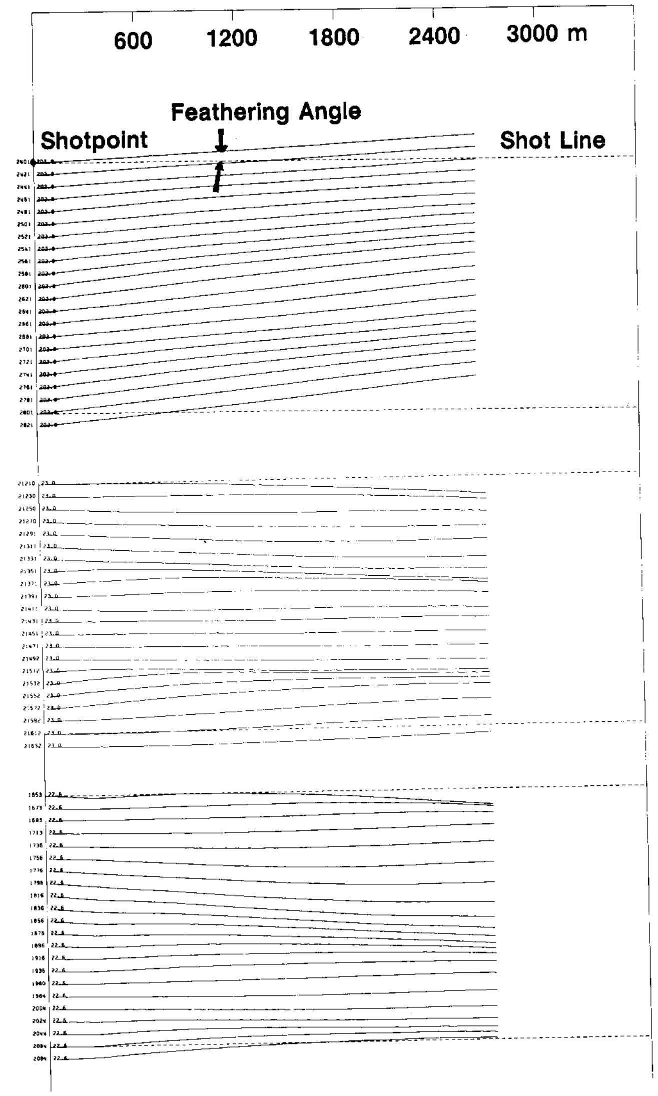


FIG. 6-11. Cable shapes for selected shots from a marine 3-D survey. Note that the cable shape varies from shot to shot, making the feathering angle nonuniform. Variations in cable shape cause variations in the distribution of midpoints within cells.

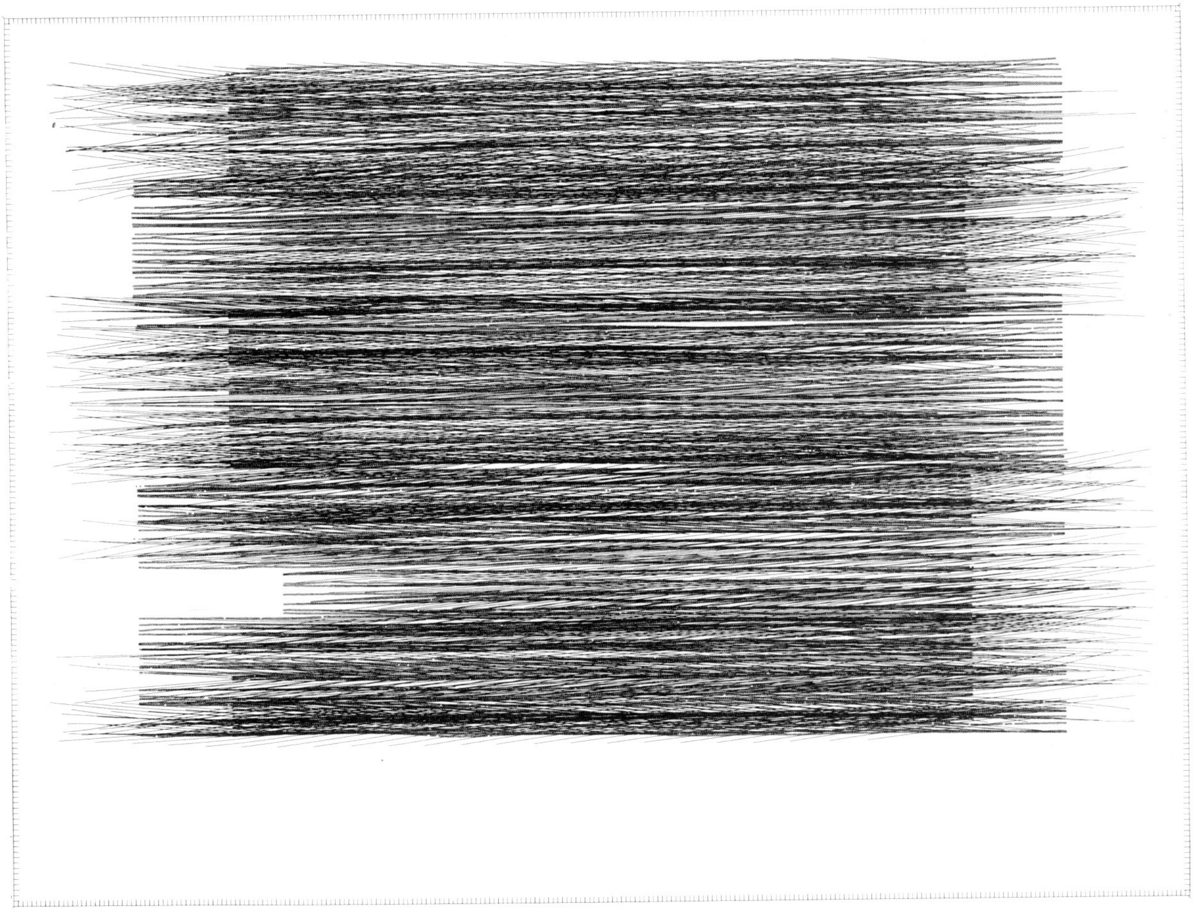


FIG. 6-12. Cable shapes and source locations for an entire marine 3-D survey.

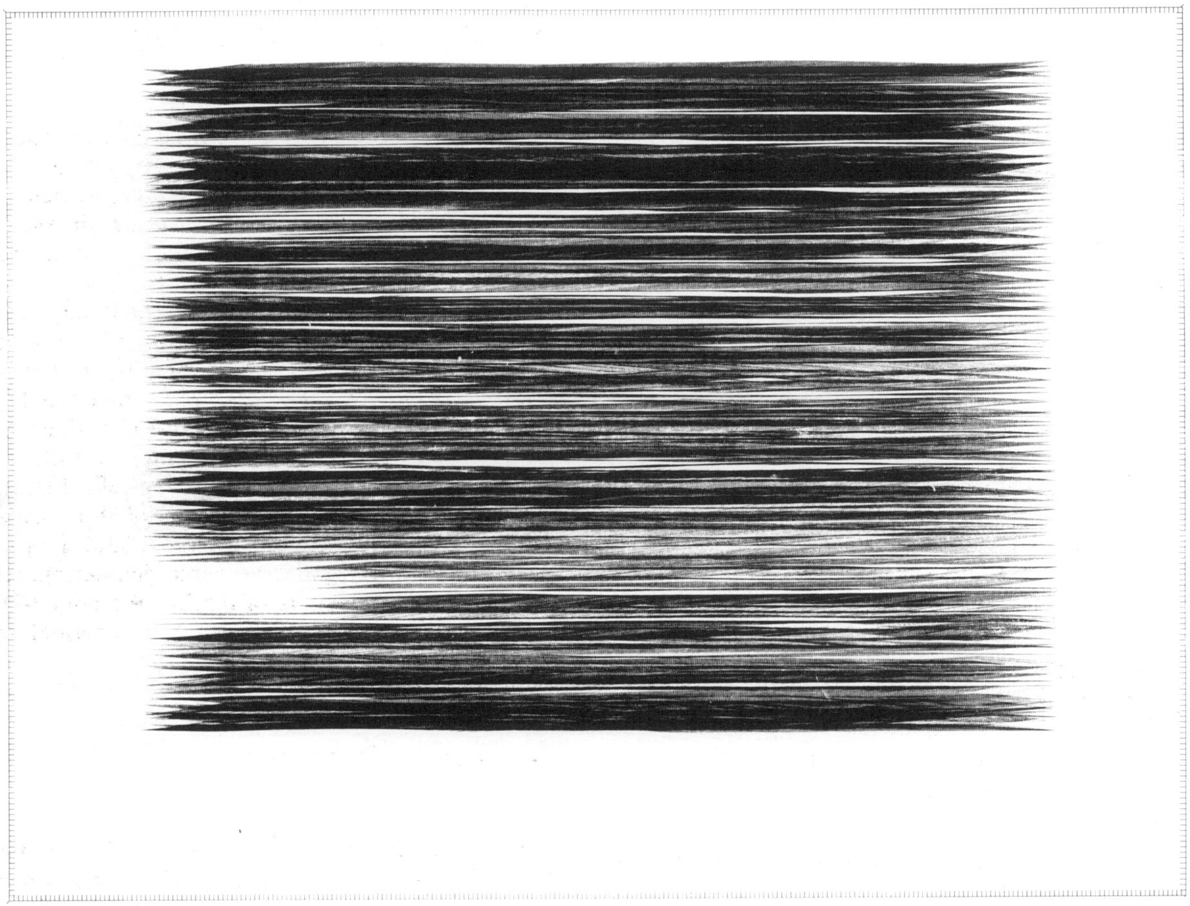


FIG. 6-13. Midpoint locations derived from the cable shapes in Figure 6-12.

completed, another swath parallel to it is recorded; this procedure is repeated over the entire survey area. Although not exercised in some 3-D surveys, it is important to leave some receiver cables on the ground to ensure proper coupling of statics between swaths. A complete survey plan, including the 12 swaths of receivers and shot locations, is shown in Figure 6-15. Because of operating conditions, uniform coverage was not achievable over the entire survey area. The average fold for the survey in Figure 6-15 is 12, increasing to 24 toward the bottom right side of the area.

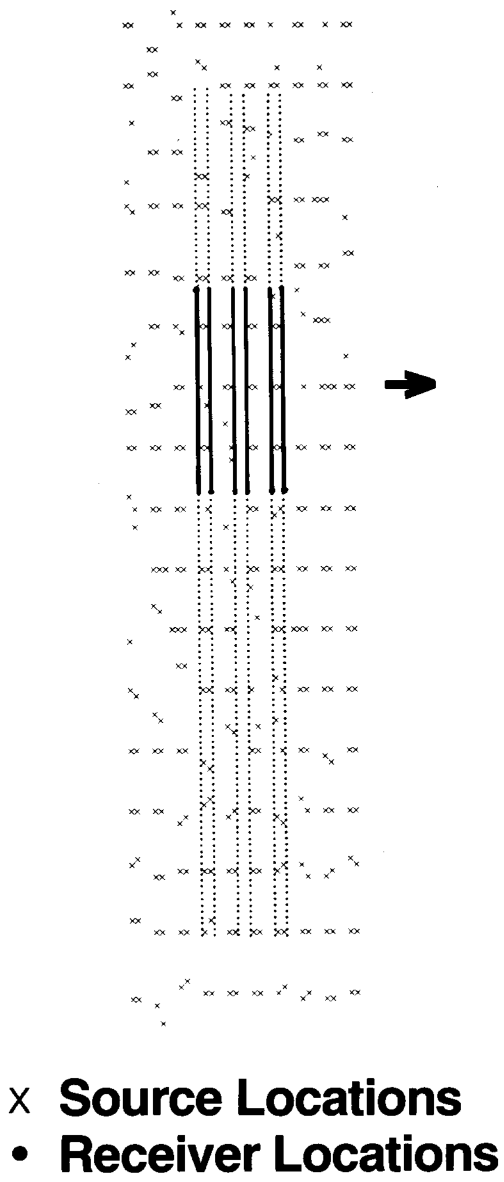


FIG. 6-14. The swath shooting geometry used in a land 3-D survey. Solid lines indicate the actual receiver cables for a particular shot line aligned with the arrow indicated. Receiver cables (80 groups each) and shots are moved up along the swath. When the end of the swath is reached, the next swath starts until all of the survey area is covered.

There are other types of shooting geometries for land 3-D surveys. Because of environmental and topographic restrictions, swath shooting is not always possible. For example, coverage under a topographic high or a lake can be achieved by shooting around it, making a complete loop with the shots and receivers located on the loop. (This type of geometry is to be avoided if at all possible, since it usually is expensive and often results in large azimuth and fold variations in common-cell gathers.)

6.4 3-D DATA PROCESSING

Almost all concepts of 2-D seismic data processing apply to 3-D data processing. Additional complications do arise in 3-D geometry quality control, statics, velocity analysis, and migration. Editing traces with high-level noise, geometric spreading correction, deconvolution and trace balancing, field statics applications (for land and shallow water data) are done in the preprocessing stage. In conventional 2-D processing, traces are collected into CMP gathers, while in 3-D processing, traces are collected into common-cell gathers. A common-cell gather coincides with a CMP gather for swath shooting when the lines are straight. Sorting into common-cell gathers introduces special problems. For a dipping reflector, there is the problem of azimuthal variations of the normal moveout (NMO) within the cell for most land data and for marine data with significant cable feathering. In this section, important aspects of marine and land 3-D processing are discussed. Section 6.5 is devoted to 3-D migration.

6.4.1 Marine Processing

After preprocessing, the data are ready for common-cell sorting. As illustrated in Figure 6-10, a grid is superimposed on the survey area. This grid consists of cells with dimensions half the receiver group spacing in the in-line direction (equivalent to the CMP spacing in 2-D processing) and the nominal line spacing in the cross-line direction. Traces that fall within a cell make up a common-cell gather. Not all of these traces are from the same shot line because of cable feathering. Sorting the data into cells is called *binning*.

Figure 6-16 shows a single cell that is associated with the field geometry in the same figure. The cell is 12.5 m in the in-line (shot-line) direction and 50 m in the cross-line direction. Different symbols represent the midpoints that are associated with different shot lines. This cell contains midpoints from six different shot lines. In Figure 6-16, the midpoint distribution corresponds to the ideal case of constant feathering angle over the entire survey. The shooting direction is assumed to be the same for all the shot lines that contribute midpoints to this cell.

In reality, the midpoint distribution within a cell is not necessarily uniform since cable shape varies from shot to

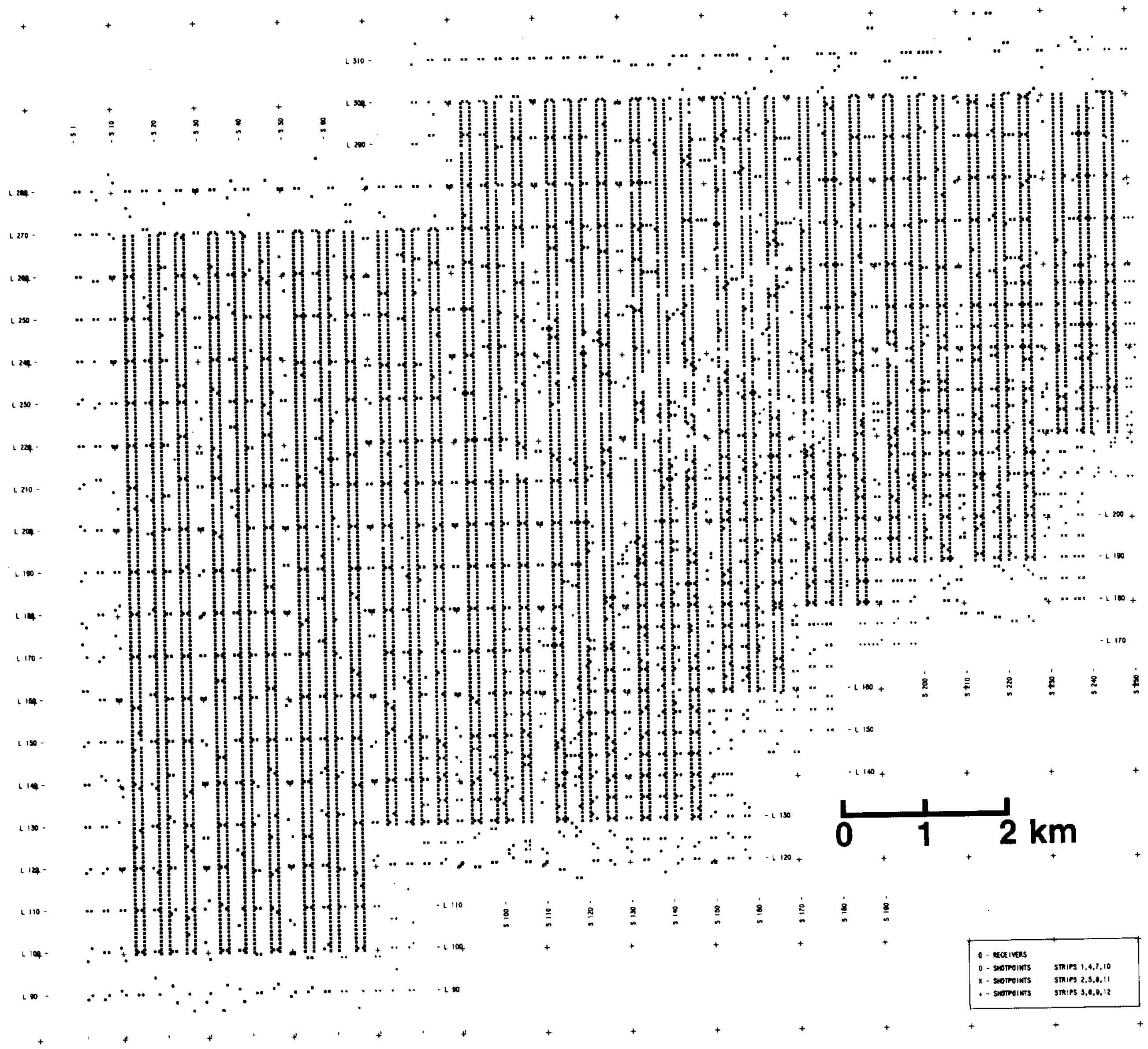
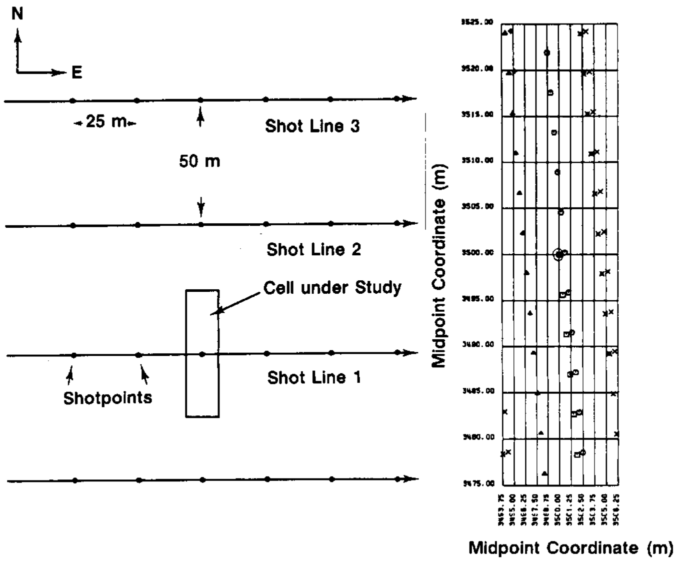


FIG. 6-15. Shot and receiver locations for an entire land 3-D survey. There are 12 swaths like the one in Figure 6-14. Coverage varies over the survey between 12-fold and 24-fold.



(a) (b)

FIG. 6-16. (a) A marine recording geometry (line shooting) and (b) an individual cell with midpoint locations. Panel (b) illustrates midpoint scattering in cells (Bentley and Yang, 1982).

shot and line to line (Figure 6-11). Midpoints may be clustered at one part of the cell; that is, the centroid of the midpoints is not necessarily at the center of the cell. Midpoint distribution also can vary from cell to cell. Some cells may contain more traces than others, while some cells may have less uniformly distributed midpoints than others.

Figure 6-17 shows the fold of coverage in the cells superimposed on the midpoints whose actual locations are plotted in Figure 6-13. Note that the fold of coverage is not uniform over the survey area. Lack of uniformity causes an inconsistency in the accuracy of velocity estimation from one analysis location to another and also causes variations in the stacked amplitudes. Long offsets are necessary for accurate velocity estimation and effective multiple suppression.

Consider some modifications to the common-cell sorting (binning) described above. A slight translation and rotation of the grid imposed on the survey area sometimes significantly reduces problems associated with binning. This type of grid optimization may yield a more uniform midpoint distribution within each cell and even improve the uniformity of the fold of coverage over the survey area.

It is common not to restrict the gridding to cells of equal size. Expanding the cell in the cross-line direction as needed can yield a uniform fold of coverage. More midpoints are included in the cell from the neighboring cells. Although the same midpoint could be used in more than one cell, restrictions often are imposed on the offset range and multiplicity to prevent an excessive use of each midpoint.

The common-cell sorting problem is not completely solved by restoring uniformity in the fold of coverage. As mentioned earlier, the centroid of the midpoints may not coincide with the center of the cell. If the centroid departs significantly from the center, then placing the stacked trace at the centroid, rather than at the cell center, must be considered. This destroys equal spacing of the stacked traces, primarily in the cross-line direction. Nevertheless, trace interpolation can be used to create equally spaced traces from data stacked to centroids. Trace interpolation often is needed for 3-D migration in any case (Section 6.5.4).

With common-cell sorting, there is an additional problem. The traveltimes of the arrivals in a single cell may not follow a single hyperbolic moveout curve (Levin, 1983, 1984). To see why this problem arises, assume that the cable shape is straight with a constant feathering angle. Consider the single cell and field geometry in Figure 6-16. Assume a simple earth model with a single, 30-degree dipping event. Also assume a constant velocity medium above the dipping interface. The midpoint distribution shown within this cell corresponds to a 10-degree constant feathering angle.

Figure 6-18 shows the traveltime curves for the cell for three different shooting directions:

a. Strike line shooting—no dip perceived along the in-line direction (maximum cross-dip case).
b. Shot line at a 45-degree azimuth with respect to dip direction.
c. Dip line shooting—no dip perceived along the cross-line direction.

Note that data from different shot lines contribute to different portions of the traveltime curves. For a single dipping event, the 2-D recording geometry normally yields a hyperbolic moveout curve. As the line orientation becomes more parallel to the reflector strike, cross-dip increases and traveltimes deviate from the ideal hyperbolic moveout curve. This ideal hyperbola corresponds to the case in which all midpoints in the cell coincide with the cell center; i.e., no cable feathering. The traveltime deviation is worse when shooting in the strike direction (Figure 6-18a). The traveltime deviation increases with increasing feathering angle, cross-dip, and cell dimension in the cross-line direction. It is also more significant at lower velocities and shallow depths. If a common-cell stack were done along the best-fit hyperbolic path (Figure 6-18a), then a loss of high frequencies would be expected.

How significant is the high-cut action due to midpoint scatter in the cross-line direction? Measure the time differences between the ideal hyperbolic path and the actual moveout curve in Figure 6-18a to derive the stacking operator plotted with its amplitude spectrum in Figure 6-19. Note the high-cut filtering action of the stacking operator; the -6 dB amplitude level is at about 70 Hz. Whether or not the effects of midpoint scatter are significant on a particular data set depends on the amount of cross dip, cable feathering, and desired bandwidth of

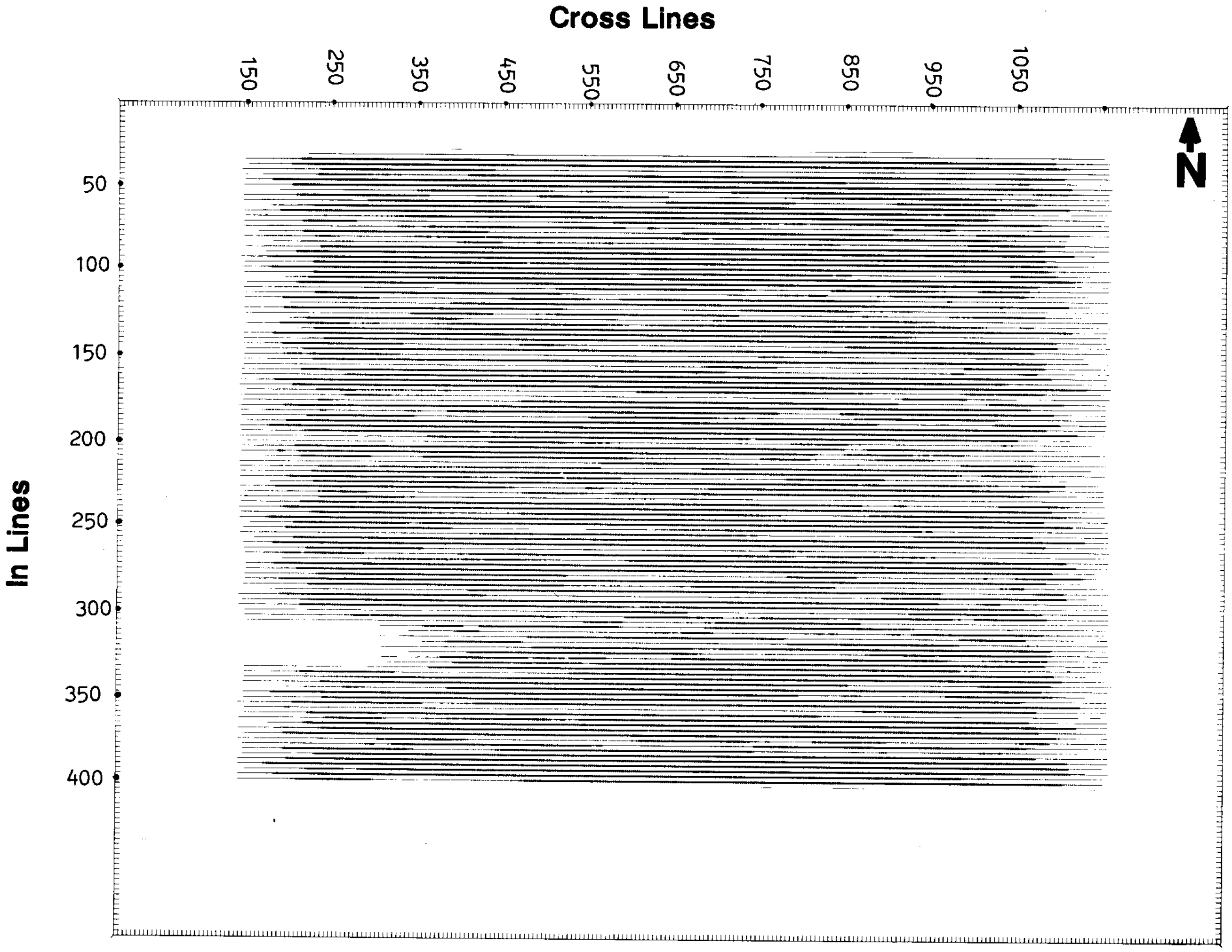


FIG. 6-17. Map of fold of coverage in each cell after sorting into common-cell gathers. Figure 6-13 shows the location of the midpoints. Where lines are thick, cells contain more than 48 midpoints; where lines are thin, cells contain between 24 and 48 midpoints. Blank spots correspond to fold of coverage less than 24-fold. Note the nonuniformity of the fold in several parts of the survey.

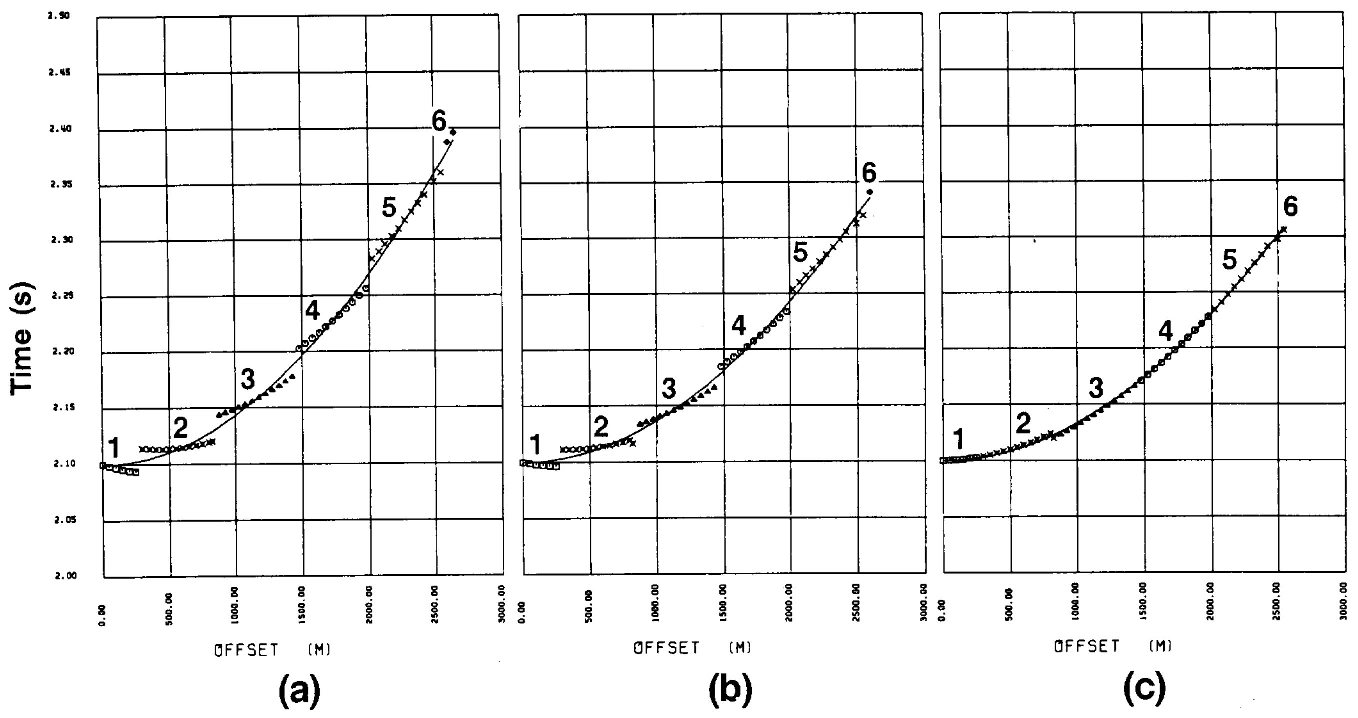


FIG. 6-18. Traveltimes and the least-squares-fit hyperbolic moveout curves associated with (a) strike-line shooting, (b) shooting in the 45-degree direction from the downdip azimuth, and (c) dip-line shooting. These traveltimes were derived from a single planar interface with a 30-degree dip in a constant-velocity medium. The feathering angle is 10 degrees and midpoint distribution in the cell is shown in Figure 6-16. Numbers along the moveout curves denote the lines that contribute midpoints to the cell under study (Bentley and Yang, 1982).

the stacked data. The recording conditions and bandwidth objectives often are such that cross-line smear is not an issue.

In summary, crosscurrents cause cable feathering, which makes the midpoints scatter within a cell in the cross-line direction. If the shooting direction is such that a dipping interface has a cross-dip component, then the traveltimes associated with that interface in a common-cell gather deviate from a single hyperbolic moveout curve. This causes amplitude smearing during stacking, which acts as a high-cut filter. The cutoff frequency primarily is a function of the amount of cross dip, reflection time, and velocity.

If cross-line smear is significant, there are a number of ways to minimize it. The simplest, but most expensive, way to avoid cross-line smear is to have the cell dimension in the cross-line direction sufficiently small, which means shooting with close line spacing. Cross-line smearing can be handled at less cost (but perhaps not as well) in processing by applying a CMP location correction. This correction involves mapping traveltimes associated with all the midpoints within the cell onto a specified midpoint location, normally cell center (Meinardus and McMahon, 1981). The procedure requires local dip and velocity estimates that can be based directly on the seismic data or on a predefined geologic model. Once these estimates are known, it is a simple matter to correct traveltimes to account for the distance of a trace from the desired midpoint. Use of this correction is demonstrated with the model example from the above reference (Figure 6-20).

Figure 6-20a shows a single cell with 48 midpoints from 5 different source lines shot in the same direction. The earth model consists of a single dipping reflector with 30-degree cross-line dip and 45-degree in-line dip components. Note the discontinuities along the traveltime curve shown in the common-cell gather data in Figure 6-20b. Each of the five lines contributes to different parts of the moveout curve, A, B, C, D, and E. If we fit the best hyperbolic moveout curve to this gather and apply the appropriate moveout correction, then the gather in Figure 6-20c results. Stacking would cause high-frequency attenuation as illustrated in Figure 6-19. The midpoint-location correction amounts to mapping the traveltimes in Figure 6-20b to those in Figure 6-20d. After the NMO correction (Figure 6-20e), we expect a better quality stacked trace without cross-line smearing.

Cross-line smearing may be compounded by a large azimuthal range of shots and receivers. This is the case in Figure 6-21a, where the cell contains midpoints from five lines: two shot in one direction (traces 29 through 48) and three shot in the opposite direction (traces 1 through 28). Shooting in the opposite direction makes the effective azimuth range twice the feathering angle. Although midpoint scattering can be corrected (as in Figure 6-20), there still is a discontinuity along the corrected moveout curve (Figure 6-21d) from the source-receiver azimuthal variation. Segments A and B along the moveout curve are associated with the two shooting directions that are opposite each other. From Figure 3-17, NMO depends on dip and the source-receiver azimuth measured from the dip-line direction. Therefore, different velocities should be used to do the NMO correction for segments A and B. Figure 6-21e shows the result of correcting for azimuth. Azimuth-dependent velocity estimation is discussed in Section 6.4.2.

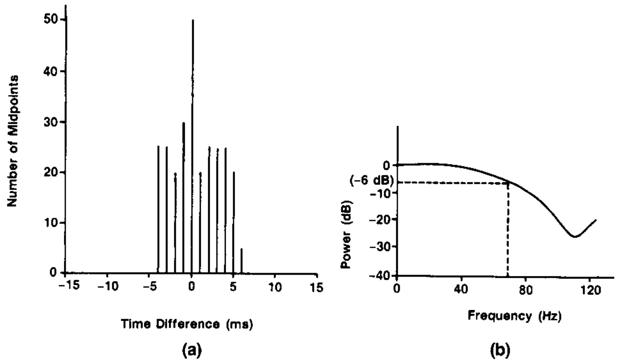


FIG. 6-19. (a) The stacking operator associated with midpoint scatter along the cross-line direction. This operator was derived from traveltime deviations from the best-fit hyperbolic moveout curve in Figure 6-18a. (b) The amplitude spectrum suggests that high frequencies are attenuated (causing smearing of stacked amplitudes) because of midpoint scatter (Bentley and Yang, 1982).

Once the data are sorted into common-cell gathers, velocity analysis is performed. In this respect, there is no difference between 2-D and 3-D data processing. For 2-D data processing, a number of neighboring CMP gathers are included in the velocity analysis to increase the S/N ratio. Similarly, a number of common-cell gathers are included in the 3-D velocity analysis. This number can be, say, 5 in the in-line and 5 in the cross-line direction, for a total of 25 common-cell gathers. For the 3-D example, velocity analyses are performed at certain intervals (e.g., 0.5 km) along selected in lines that may be as much as 0.5 km apart. The results of velocity analyses at selected control points are used to derive the 3-D velocity field for all common-cell gathers over the entire survey. This is achieved by performing a 3-D interpolation of the velocity functions between the control points.

6.4.2 Land Processing

The swath shooting geometry commonly used in land 3-D surveys produces common-cell gathers in which all midpoints coincide with the cell center. However, this technique also can result in large variations in the source-receiver azimuths and, thus, can result in traveltime deviations similar to or greater than those caused by midpoint scattering in marine surveys. Figure 6-22 is a recording geometry that comprises a range of shot-receiver azimuths. Reflections from a dipping interface do not align along a single hyperbolic moveout curve (Figure 6-22b). Instead, there are some traveltime deviations. Using a single velocity for the NMO correction

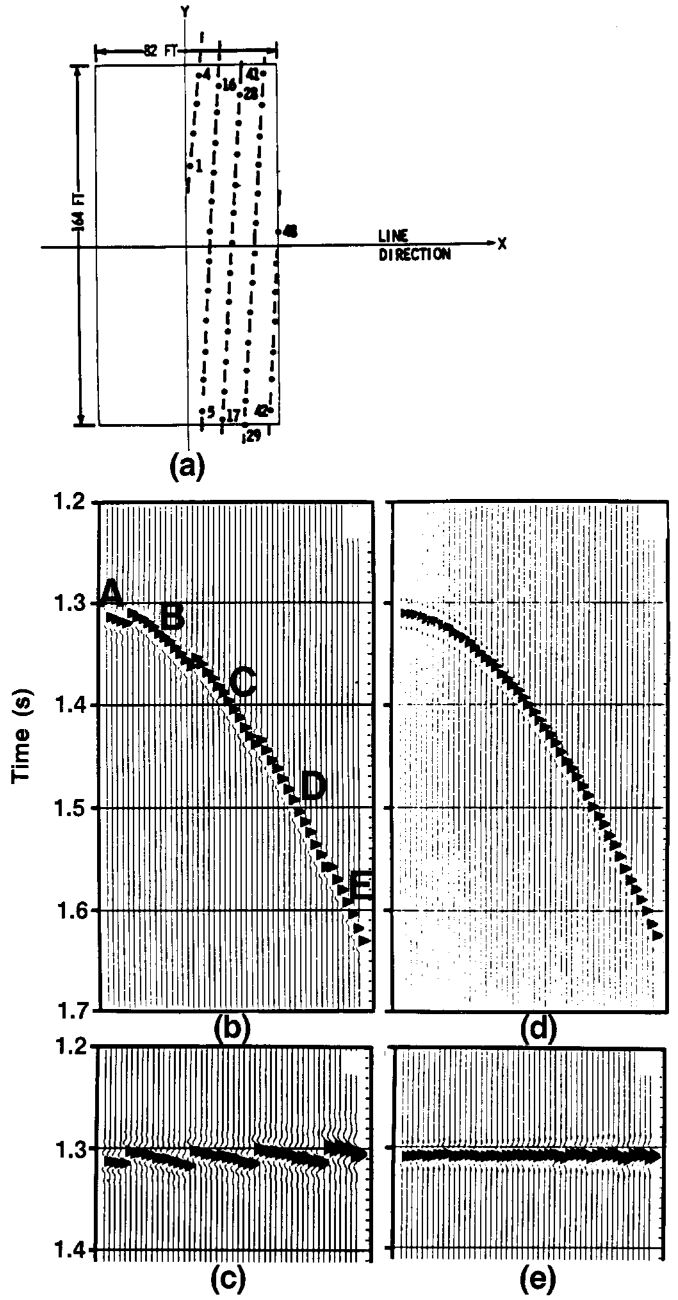


FIG. 6-20. (a) A cell containing midpoints from five different shot lines in the same direction, (b) the corresponding common-cell gather data for a dipping reflector with a nonzero cross-dip angle. Segments A, B, C, D, and E come from midpoints 1-4, 5-16, 17-28, 26-41, and 42-48, respectively. Traveltimes can be fitted to a hyperbola and moveout correction can be applied as in (c). Stacking will cause loss of high frequencies as in Figure 6-19 (cross-line smear). After midpoint-location correction, all midpoints in (a) are mapped onto the cell center and the moveout curve is simpler (d). Stacking after midpoint-location and moveout correction (e) will preserve all frequencies (Meinardus and McMahon, 1981; courtesy Geophysical Service Inc.).

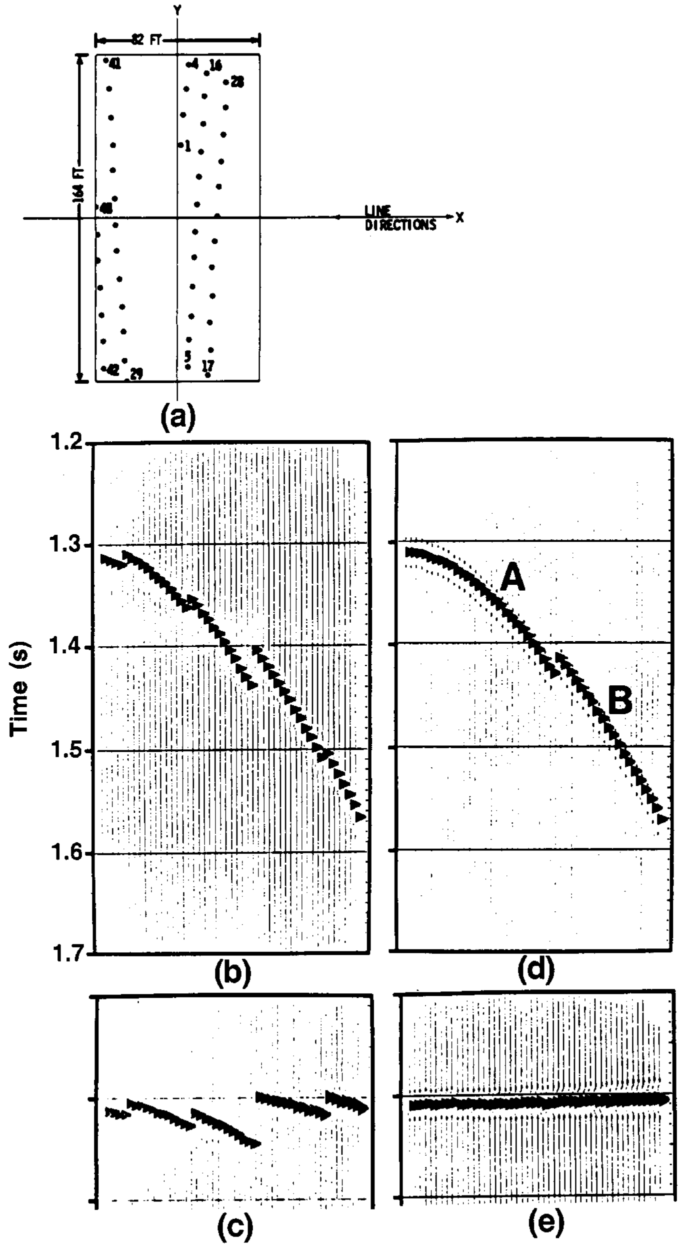


FIG. 6-21. (a) A common-cell gather containing midpoints from five different shots. Midpoints 1-28 come from lines shot in one direction, while midpoints 29-48 come from lines shot in the opposite direction. (b) The corresponding common-cell gather data for a dipping reflector with nonzero cross-dip angle. (c) The same gather after NMO correction. (d) The gather in (b) after midpoint-location correction. (e) The gather in (d) after NMO correction using azimuth-dependent velocities (one velocity for segment A and another for segment B). Note that even after midpoint-location correction, the moveout curve is composed of two segments, A and B, which are associated with two shooting directions opposite each other (Meinardus and McMahon, 1981; courtesy Geophysical Service Inc.).

(Figure 6-22c) yields a stacked trace with high frequencies attenuated as in cable feathering. Stack attenuation increases with increasing shot-receiver azimuth range. If this attenuation is serious, correction for the shot-receiver azimuth before stacking is needed (Figure 6-22d).

Levin (1971) shows that the moveout velocity for a dipping reflector is dip-dependent and a function of the source-receiver azimuth. Figure 3-17 is the 3-D geometry for a dipping event, while equation (3.9) is the associated moveout velocity. Azimuth is measured from the dip-line direction. Figure 6-22b suggests a way to correct for the source-receiver azimuth. The traces in a CMP gather can be grouped into various ranges of azimuths and different velocities can be used for moveout correction of each group of traces.

Levin's relation [equation (3.9)] is the equation of an ellipse in polar coordinates. The radial coordinate is the NMO velocity, while the polar angle is the azimuth. Orientation of the major axis of the velocity ellipse is in the true dip direction. The velocity ellipse can be constructed based on moveout velocities measured in three different directions (Figure 6-23). Lehmann and Houba (1985) discuss several practical aspects of this measurement. By subgrouping traces from a CMP gather into three different azimuths (α_1, α_1, α_3), the stacking velocities (v_{s1}, v_{s2}, v_{s3}) can be estimated along those azimuths. If the stacking velocities in three different directions are known (Figure 6-23a), then the semimajor and semiminor axes (a and b, respectively), and the orientation (downdip azimuth angle, ε) of the velocity ellipse can be defined. Once the velocity ellipse is constructed for a dipping reflector at a CMP location, then the following parameters are determined (Figure 6-23b):

1. The true dip of the reflector, $\phi = \cos^{-1} (b/a)$.
2. The downdip azimuth ε (Figure 6-23b).
3. The stacking velocity at any azimuth, $v_s = v/(1 - \sin^2 \phi \cos^2 \theta)$, where $v = b$ (medium velocity above the reflector) and θ is the azimuth angle measured from the major axis a.

This is called the three-parameter velocity analysis. Each trace in the CMP gather is moveout corrected using the velocity along the shot-receiver azimuth that is associated with that midpoint.

Although azimuthal variations in velocity may be observed on velocity spectra (Figure 6-24), they may not impact stacking. Some differences between the three velocity functions at a single CMP location (Figure 6-24) may result from the nonuniform azimuthal coverage and offset distribution. Ideally, all offsets at all source-receiver azimuths are needed to define the velocity ellipse accurately at a CMP location. In this ideal situation and by using the three-parameter approach, moveout velocity at any source-receiver azimuth, true dip, and the downdip azimuth of a reflector can in principle be determined.

There are some other practical considerations with regard to 3-D velocity estimation. Land 3-D data often are of low fold; so partitioning the data into azimuth ranges with lower folds of coverage may not yield good velocity estimates. Azimuth variations can be minimized on land swath recording by minimizing perpendicular offset of sources from the receivers. Azimuth variation also often is largest on the near offsets where total moveout is small. Thus, the effect on stacking is smaller than normally thought. Another issue is offset distribution. Ideally, it is desirable to have a wide selection of offsets for each common-cell gather. Otherwise small errors in moveout correction may cause trace-to-trace chatter on stacked sections.

In this section, we discussed a method to estimate azimuth-dependent moveout velocities. In Section 4.4.1, we saw that the dip-moveout (DMO) process corrects the moveout velocity of a dipping reflector for the cosine of the true dip angle. We concluded that for 3-D, DMO should correct for the dip and the azimuth (measured from the dip-line direction). The 2-D DMO operator maps the moveout-corrected amplitudes on a common-offset section along truncated elliptical trajectories that narrow at increasingly late times (Deregowski and Rocca, 1981). The 3-D DMO operator maps the amplitudes on the plane of source-receiver azimuths over an ellipsoid. Just as it is true for the three-parameter velocity analysis, the fidelity of this process depends on source-receiver azimuthal coverage; in particular, how well the ellipsoid is defined.

A final topic in land 3-D processing is residual statics corrections. The surface-consistent residual statics model (Section 3.3) does not restrict shot and receiver stations to a survey line; they can be situated anywhere. Hence, equation (3.25) can be used to model the static shifts at all shot and receiver stations over the entire 3-D survey area. As it is for the 2-D case, the critical step is in picking (estimating) the traveltime deviations from the common-cell gathers. For 3-D seismic data, an initial pilot trace is built from the common-cell gather of a selected line where the S/N ratio is good. Pilot traces for

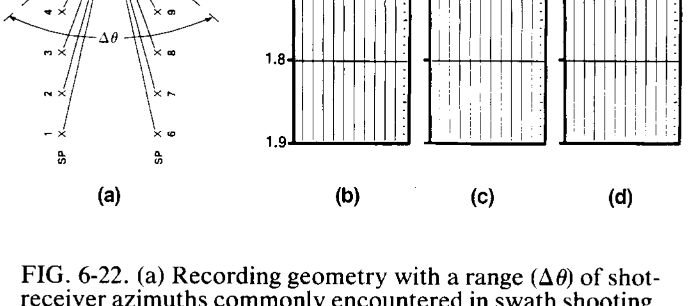


FIG. 6-22. (a) Recording geometry with a range ($\Delta\theta$) of shot-receiver azimuths commonly encountered in swath shooting. (b) A CMP gather consisting of 10 traces. (c) Moveout correction using the velocity derived from the best-fit hyperbola. (d) Moveout correction using azimuth-dependent velocities (Meinardus and McMahon, 1981; courtesy Geophysical Service Inc.).

other common-cell gathers are computed from both local input traces and neighboring pilot traces. Note that land recording geometry must accommodate some overlap between adjacent swaths to ensure statics coupling between the swaths.

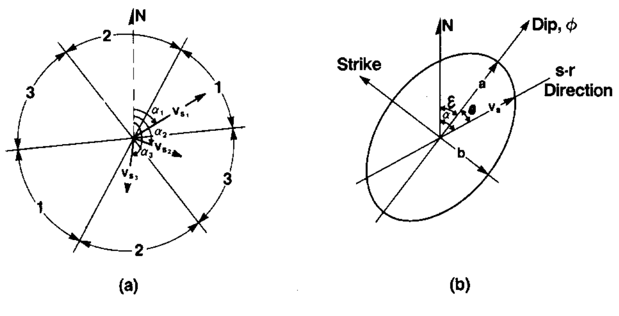
6.5 3-D MIGRATION

To understand 3-D migration, consider a point scatterer that is buried in a constant-velocity medium. The zero-offset traveltime curve in two dimensions is a hyperbola. We imagine that the response in three dimensions is a hyperboloid. Migration in two dimensions amounts to summing amplitudes along the diffraction hyperbola, then placing the result at its apex. This idea can be extended to three dimensions. Migration in three dimensions amounts to summing amplitudes over the surface of the hyperboloid. In Section 4.2.1, we learned that the simple diffraction summation technique for migration can be improved by making some amplitude and phase corrections based on the Kirchhoff integral solution to the scalar wave equation before summation [equation (4.5)]. Schneider (1978) provides a detailed theoretical discussion on 3-D Kirchhoff migration with some field data examples. Basically, in 3-D migration the integration given by equation (4.5) is done over the surface of a 3-D survey area instead of along a 2-D seismic line. For 2-D migration, as many as 300 traces may be included in the summation. In three dimensions, this implies that as many as 70,000 traces may need to be included in the summation. Problems arise when such large number of traces needs to be handled efficiently in the computer.

6.5.1 Two-Pass versus One-Pass 3-D Time Migration

We now discuss some practical alternatives to summing over the surface of the hyperboloid. The most complete account of 3-D imaging techniques is provided by Ristow (1980). The diffraction hyperboloid, which is strictly associated with a constant-velocity medium, has a hyperbolic cross section in any azimuthal direction. Sum along the hyperbolic cross sections in the in-line direction and place the summed amplitudes at the local apexes of these hyperbolas. The hyperboloid now is collapsed to a hyperbola that is in the plane perpendicular to the direction of the first summation. This hyperbola comprises the first summed amplitudes at the local apexes and is contained in the plane of the cross-line direction. Next, sum the energy along this hyperbola and place it to its apex, which is also the apex of the original hyperboloid and is where the image should be placed. Gardner et al. (1978) proposed this two-pass approach to do 3-D migration that involves two successive 2-D migrations of the in lines and cross lines.

The underlying mathematical notion for the two-pass approach is *full separation* of extrapolation operators in the in-line and cross-line directions (Claerbout, 1985). Another method of 3-D migration is based on the numerical procedure called *splitting* (Claerbout, 1985). Here, extrapolation operators are independently applied to the data in both the in-line and cross-line directions at each downward continuation step. In contrast, full separation calls for complete migration in one direction (say, in-line) before any operation is applied in the other direction. The sequence of operations in the splitting and separation methods is shown in Figure 6-25.



(a) (b)

FIG. 6-23. (a) Azimuth-dependent velocity estimation involves dividing traces in a common-cell gather into different compartments (in this case, three) and estimating stacking velocities for each compartment. (b) The NMO velocity relationship derived from the geometry in Figure 3-17 is the equation of an ellipse in polar coordinates [equation (3.9)]. The radial coordinate represents the NMO velocity at a given azimuth, which is the polar angle α. (Refer to the text for further details.)

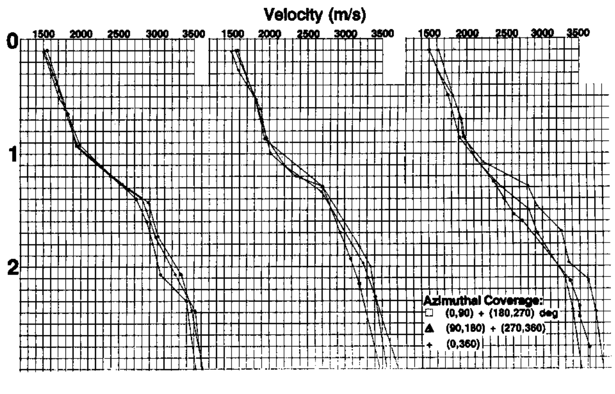


FIG. 6-24. Azimuth-dependent stacking velocities at three locations from a land 3-D survey.

As noted earlier, 3-D migration based on full separation is known as the two-pass approach. Three-dimensional migration based on splitting is considered a one-pass approach. The efficiency of the two-pass method stems from the fact that data management in the computer is much easier and hence there are fewer input/output operations.

Since the need for 3-D imaging of the subsurface becomes especially significant in areas with steep dips and lateral velocity variations, the 2-D omega-*x* algorithm (Appendix C.3) implemented for one-pass 3-D migration and used in the examples shown in this section. A concise theoretical description of 3-D migration is provided in Appendix C.8.

Figures 6-26 and 6-27 show the 3-D 15-degree and 45-degree operator impulse responses for the one-pass approach. The corresponding 2-D operators are shown in Figure 4-90. Selected lines and time slices are shown to better illustrate the shape of the responses. There are 101 in lines and 101 cross lines. The ideal 3-D impulse response is a hollow hemisphere. The 15-degree response (Figure 6-26) appears to be symmetric with respect to the center midpoint (note the nearly circular shapes on the time slices), although the traveltimes are not quite correct. The 45-degree operator (Figure 6-27) produces more accurate traveltime responses (especially in the in-line and cross-line directions), but the error in traveltimes does not have azimuthal symmetry (see Appendix C.8). The largest errors occur at the two diagonal directions. Ristow (1980) discusses further splitting of the 3-D migration operator in the two diagonal directions, in addition to the in-line and cross-line directions, as a way to achieve a more circular time slice. Ristow (1980) also gives a least-squares approximation (for a specified dip angle) to the 3-D migration operator by two 2-D operators.

The two-pass approach for 3-D migration is valid strictly for a constant-velocity medium (Ristow, 1980; Jakubowicz and Levin, 1983). Since we never have a constant-velocity medium, it is reasonable to question the practical value of the two-pass 3-D migration. Now consider the case of the vertically varying velocity model shown in Figure 4-61.

Compute the 3-D zero-offset response of the three point scatterers buried in the subsurface below the center midpoint (32) of the center in line (32) at the layer boundaries (denoted by the asterisks in the velocity model in Figure 4-61). Figure 6-28a shows four selected in lines from this 3-D zero-offset data that consist of 63 in lines and 63 cross lines. Obviously the traveltime responses are circularly symmetric, the shallow one being an exact hyperboloid, while the two deeper ones are approximately hyperboloid.

One-pass 3-D migration collapses the energy to the apexes of the three diffraction traveltime surfaces at the center midpoint (32) of the center line (32) (Figure 6-28b). Now consider two-pass 3-D migration. First migrate in the in-line direction (Figure 6-28c). The center in line (32) is migrated properly, but the lines farther from the center

Separation

```
3-D Stacked Data
Loop 1 Over In Lines
    Loop 2 over depth step.
    Downward continue and
    Apply imaging principle.

    Close Loop 2.
Close Loop 1
    Sort data into cross lines.
Loop 3 Over Cross Lines
    Loop 4 over depth step.
    Downward continue and
    Apply imaging principle.
    Close Loop 4.
Close Loop 3.
```

Splitting

```
3-D Stacked Data
Loop 1 Over Depth Step
    Downward continue in the in-line direction.
    Downward continue in the cross-line
    direction and
    Apply imaging principle.

Close Loop 1.
```

FIG. 6-25. Algorithms for 3-D migration based on the splitting and separation techniques of implementing migration operators.

line are increasingly overmigrated. Moreover, within one line, say, line 2, the shallower the diffraction, the more the overmigration. This is easy to understand once we refer to the velocity model (Figure 4-61).

On lines farther from the center line, the 3-D diffraction events arrive at increasingly late times; for example, on line 2 in Figure 6-28a. Because the velocity varies with depth, diffractions on line 2 get migrated with velocities higher than the true velocities associated with these diffractions. The diffractions on line 2 are sideswipes from the point scatterers situated beneath the center line (32). Because velocities corresponding to late times are used to migrate energy generated by shallow diffractors, these sideswipes are overmigrated. Now sort the data into cross lines (Figure 6-28d). It is necessary to display only those lines closer to the center line. Note that most of the energy has collapsed to the hyperbolas on the center line (32); however, because of the error induced by the first pass due to the effect of *v(z)*, some energy still remains on the neighboring lines. Also note that the error is more significant at shallow diffraction apexes. Finally, migration in the cross-line direction collapses most of the energy to the center midpoint (32) (Figure 6-28e). Sort the data back to in lines (Figure 6-28f) and compare them with the one-pass result (Figure 6-28b).

Energy spreading to the neighboring midpoints is due to the error of the two-pass approach in the presence of vertical velocity variations. The test results shown in Figure 6-28 also suggest that the larger the vertical velocity gradient, the more error induced by the two-pass approach. Another important point is that with the two-pass approach and for a *v(z)* function, it makes a difference in which direction migration is done first. Smearing is greatest in the direction in which the first pass of 3-D migration is performed.

Now consider the reality. Theoretically, it is true that the

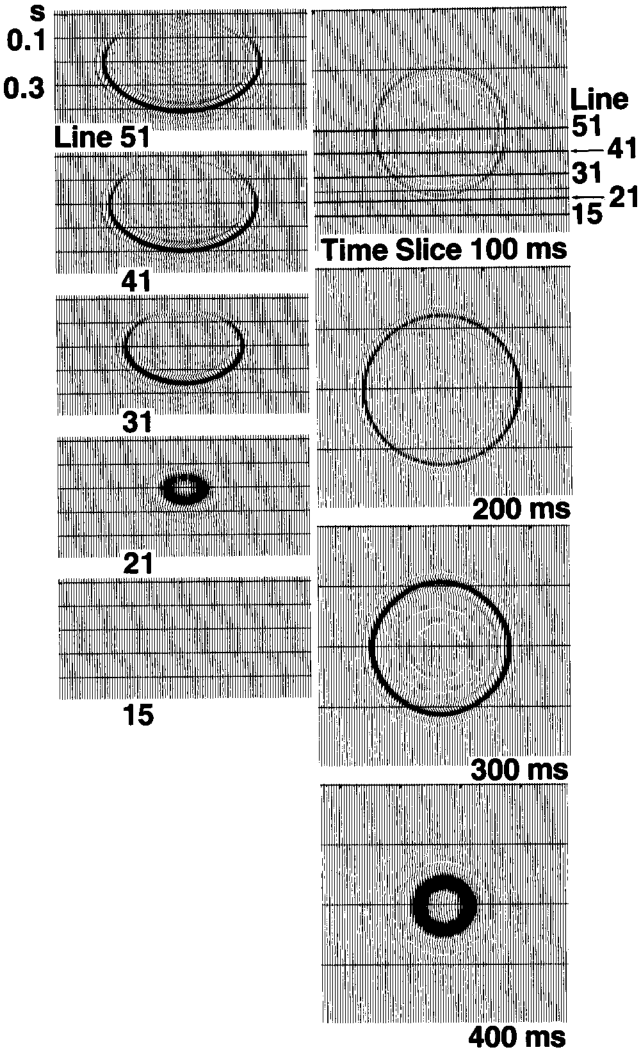


FIG. 6-26. A 15-degree split 3-D migration operator impulse response. (See Figure 4-90 for the 2-D equivalent response.)

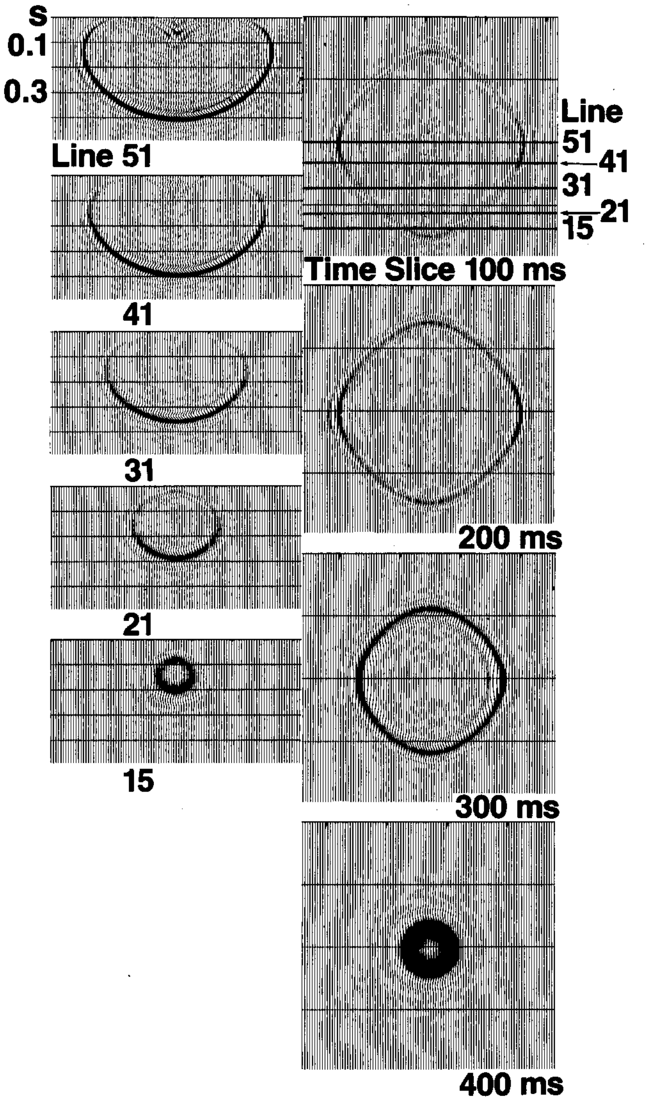


FIG. 6-27. A 45-degree split 3-D migration operator impulse reponse. (See Figure 4-90 for the 2-D equivalent response.)

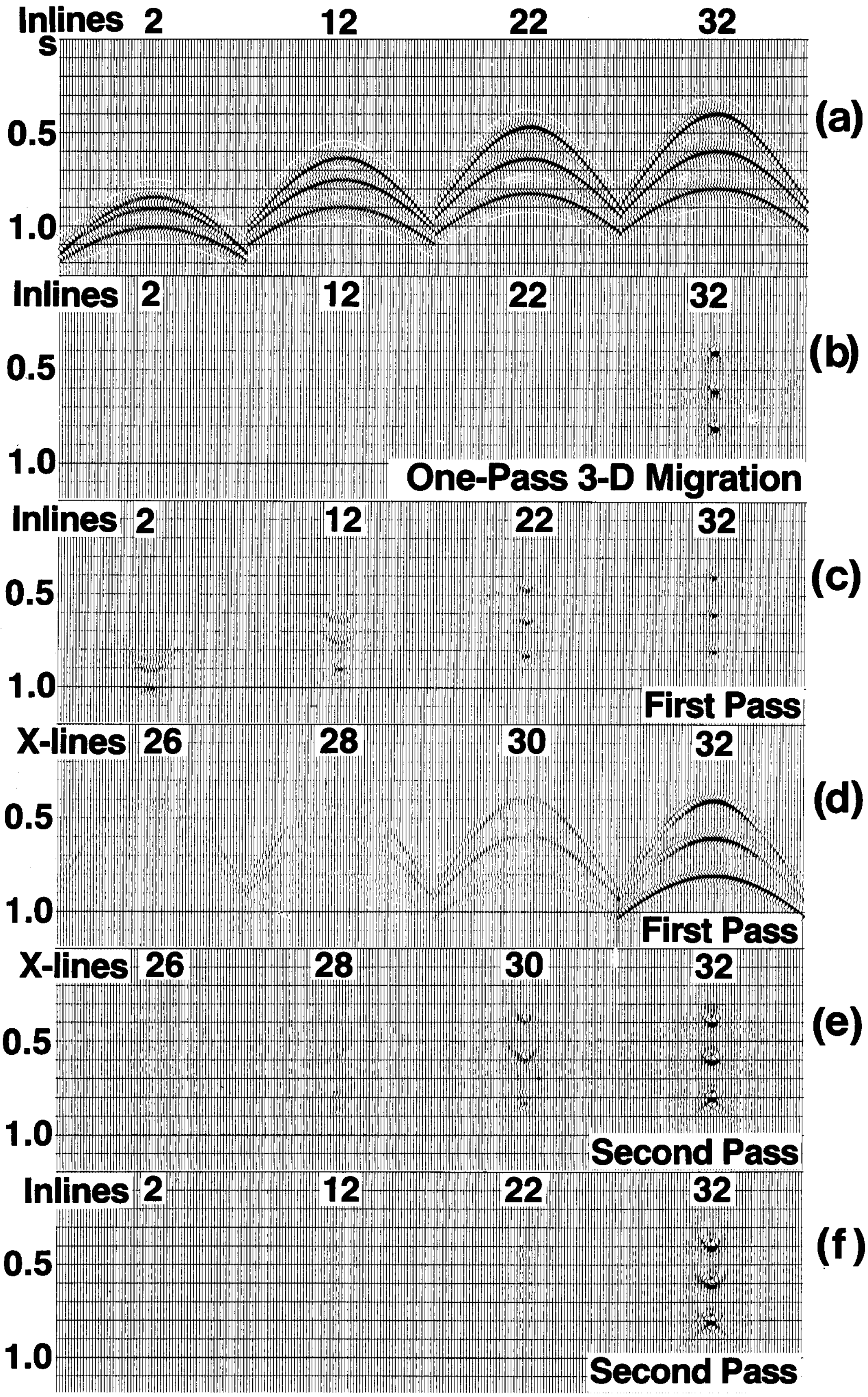


FIG. 6-28. Two-pass versus one-pass (full) 3-D migration. (Refer to the text for details.)

two-pass method is only good for velocities that vary slowly in the vertical direction and do not vary laterally. On the other hand, the cost advantages of the two-pass method may be appreciable, particularly for large-sized 3-D surveys. There is also the the question of migration algorithm accuracy versus accuracy in the velocity model used for migration. In many practical situations, error induced by the two-pass approach probably is less than the possible errors associated with uncertainty in the velocity model. Hence, it may be difficult on real data to tell whether smearing (imperfect imaging) results from the velocity errors or the two-pass approach. In practice, we often find that the two-pass approach yields acceptable results in areas in which dips are small, vertical velocity variations are moderate, and lateral velocity variations are within the limits of time migration.

Now consider the field data example in Figure 6-29. The top left is an in-line and the top right is a cross-line stacked section from a land 3-D survey. Migration in the in-line direction yields the sections shown in the center row. Note the partially collapsed diffraction energy on the in-line section (center left) and a pronounced enhancement of the diffraction energy on the cross-line section (center right). Sorting the data and migrating in the cross-line direction yields the results shown in the bottom row. These are the two-pass 3-D migrated sections.

Figure 6-30 shows the comparison between the two-pass (center row) and one-pass (bottom row) 3-D migrations of the same in-line and cross-line sections (top row). The apparent overmigration (just below A) seen on the two-pass result may be attributed to the error of the two-pass approach resulting from large vertical velocity gradients associated with these data.

Note that the above analysis is entirely within the framework of time migration. Two-pass 3-D migration algorithms cannot even be considered approximately accurate in situations where lateral velocity variations are large enough to warrant depth migration.

6.5.2 3-D Time versus Depth Migration

Consider a 3-D survey over a hypothetical salt-dome structure. The base map is shown in Figure 6-31. The synthetic 3-D survey consists of 481 in lines and 481 cross lines with 25-m trace spacing in both directions. Selected cross sections of the 3-D velocity-depth model and the associated zero-offset in-line sections are shown in Figure 6-32. The salt dome has a circular symmetry with a flat base. The base map shown in Figure 6-31 actually is a time slice at 1200 ms to show the circular symmetry.

Figure 6-33 shows the 3-D (one-pass) and 2-D time migrations of the 3-D zero-offset data shown in Figure 6-32. The migration velocity field is the true subsurface velocity-depth model used in building the zero-offset data. Note the following observations:

1. The top of the salt has been imaged properly with 3-D time migration, while the base of the salt has not. This is because the salt diapir acts as a complex overburden.
2. The 2-D time migration produced the correct result for the top of the salt only along the center in line (I241), because there are no sideswipes on this line; therefore, there is no need for 3-D migration. However, on a line away from the center line, say I181, even the top of the salt has not been imaged properly by 2-D time migration, let alone the base of the salt. This is because the line contains sideswipes off the flank of the salt dome.
3. Do you now have second thoughts about 2-D seismic exploration? Fortunately most structures have dominant strike and dip directions. Properly oriented 2-D migrated lines often yield an acceptably accurate structural picture. However, one lesson to be learned from comparing the 3-D and 2-D time migrations of line I181 shown in Figure 6-33 is that the migration velocities that yield an acceptable 2-D migrated section may be quite different from the true subsurface velocity model required by 3-D migration.

Finally, Figure 6-34 shows 3-D and 2-D depth migrations of the 3-D zero-offset data shown in Figure 6-32. The velocity field is the true subsurface velocity-depth model used in building the zero-offset data. Note the following observations:

1. The top and base of the salt now have been imaged properly with 3-D depth migration. The complex overburden problem has been solved. (Compare the left column in Figure 6-34 with that in Figure 6-32.)
2. The 2-D depth migration produced the correct subsurface model only along the center in line (I241) because there are no sideswipes on this line. However, on a line away from the center line, say I181, neither the top nor the base of the salt have been imaged properly. This is because the line contains sideswipes off the flank of the salt dome.
3. By performing 2-D depth migration iteratively to converge to assumingly correct depth model, we may be forcing the model to converge to something very different from the truth. This stems from the treatment of the sideswipes as events that are in the plane of profile.

Figures 6-35 and 6-36 show a field data example of 3-D depth migration. The in-line (top left) and cross-line (top right) stacked sections in both figures show a structural high, which is associated with culminations in an overthrust belt. Figure 6-37a is an in-line cross section of the 3-D velocity-depth model. The deepest horizon on the velocity model (horizon 8) corresponds to the event below 2 s on the in-line stacked section in Figure 6-35. The true geometry of this horizon has been distorted by the structure above acting as a complex overburden.

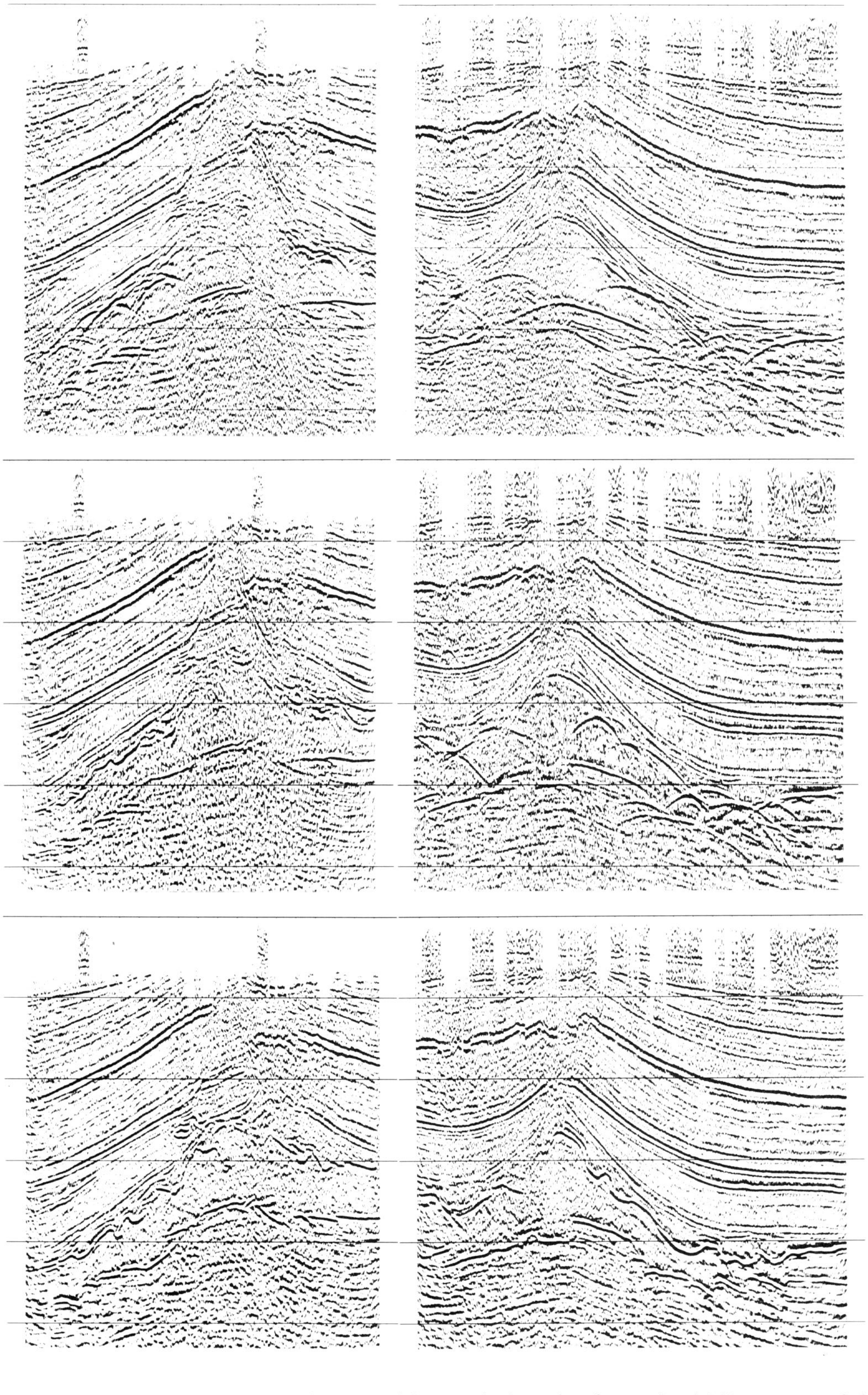


FIG. 6-29. An in-line (top left) and a cross-line (top right) stacked section from a land 3-D survey and the results from the first pass (center row) and second pass (bottom row) of a two-pass 3-D migration process. (Data courtesy Nederlandse Aardole Maatschappij B.V.)

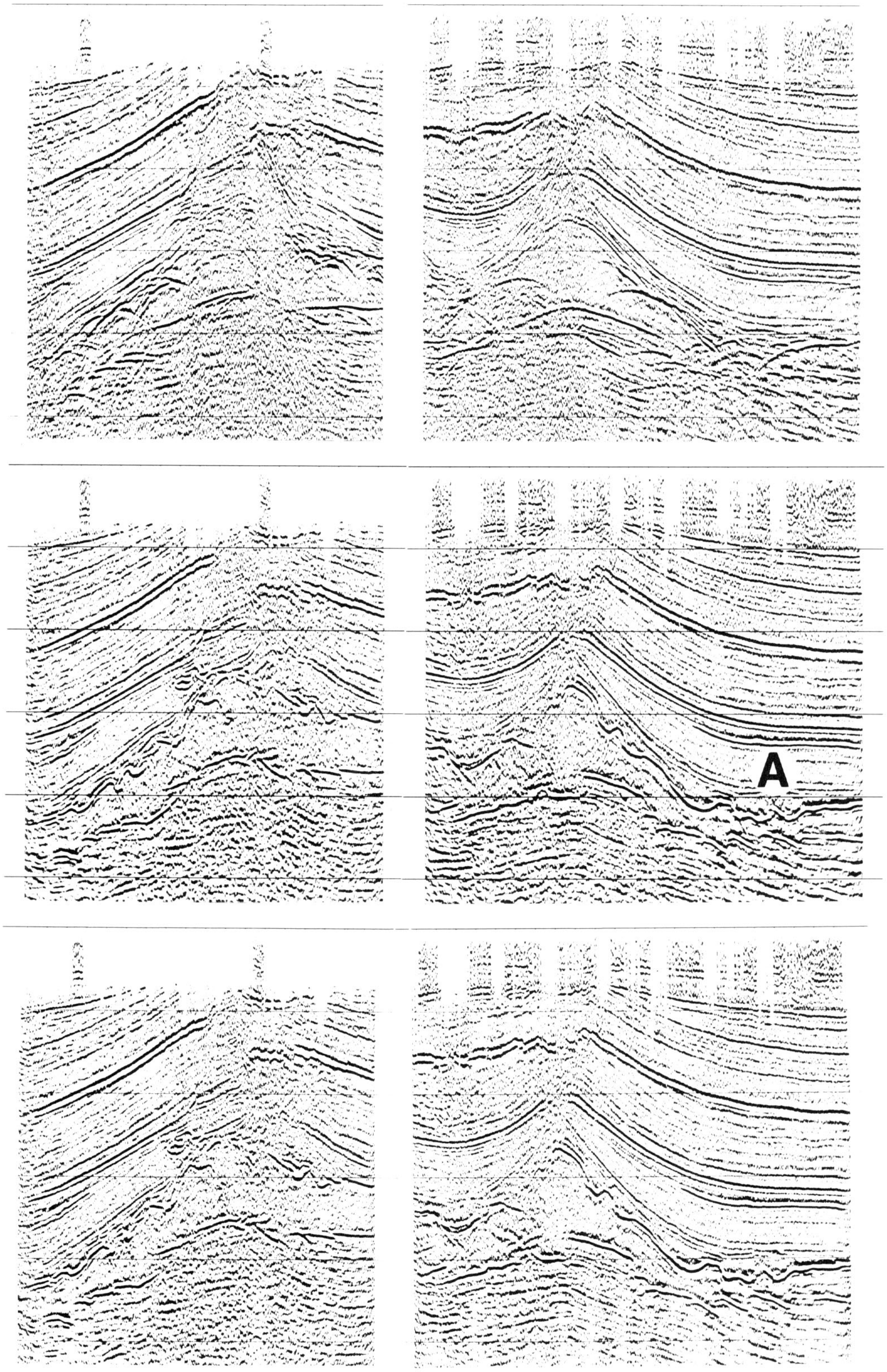


FIG. 6-30. An in-line (top left) and a cross-line (top right) stacked section from a land 3-D survey and two-pass (center row) and one-pass (bottom row) 3-D migration results. (Data courtesy Nederlandse Aardole Maatschappij B.V.)

Although in two dimensions, the image ray plot (Figure 6-37b) through the velocity-depth model (Figure 6-37a) verifies the presence of a complex overburden above horizon 8. (See Section 5.2 for a discussion on image rays.)

Note the similarity between the in-line velocity-depth model (Figure 6-37a) and line B after 3-D depth migration (bottom left, Figure 6-35). Despite this similarity, there are parts of the survey in which the data are overmigrated (on line A, bottom left, Figure 6-36) and parts in which the data are undermigrated (not shown). Differences between the output from depth migration and the velocity-depth model used requires an iterative modification of the velocity-depth model where it departs from the output of depth migration (Section 5.2). For comparison, Figures 6-35 and 6-36 show 2-D depth migrations of the selected lines. Note the obvious difference between 2-D and 3-D depth migrations of line D (Figure 6-36). Although the 2-D depth-migrated section contains an abundance of reflections as compared to the 3-D depth migrated section, that reflection energy does not belong on line D. From line A or B, note that energy should be migrated in the updip direction from line D to C. Hence, after 3-D migration, line D is depleted of reflection energy (Figure 6-36), while line C is enriched (Figure 6-35).

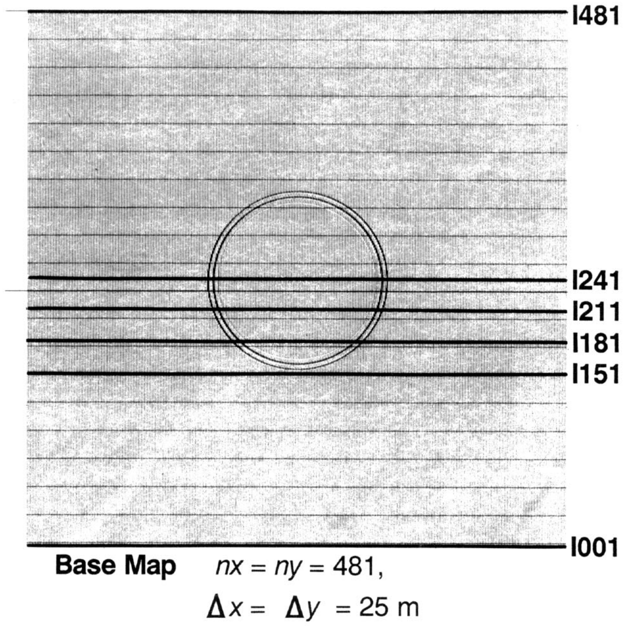


FIG. 6-31. A time slice used as a base map for a synthetic 3-D survey. The velocity-depth model and the 3-D zero-offset data for selected in lines are shown in Figure 6-32.

6.5.3 3-D Datuming

As a byproduct of the particular 3-D migration routine used here, 3-D stacked data can be datumed from the surface to a specified depth or time level in a 3-D sense. This capability can be particularly useful in reservoir studies. Basically, the surface wave field is downward continued to a desired depth without invoking the imaging principle along the way. Figure 6-38 shows the 3-D zero-offset data (Figure 6-32) datumed from the surface ($z = 0$) to a 1000-m depth. The following conclusions can be made:

1. Since there are no sideswipes on the center line, I241, datuming in a 2-D or 3-D sense is identical.
2. However, there is a significant difference between 2-D datuming and 3-D datuming for lines with sideswipe energy; i.e., those that are increasingly farther from the center line, say, Line I181.

The wave-equation datuming described here takes the input wave field from one flat constant horizontal datum to another. Prestack and poststack wave-equation datuming using arbitrary 2-D datum interfaces is discussed in Section 5.3. The constant datum level should not be a limitation, particularly in reservoir studies. For example, the 3-D stacked data can be datumed to the top of the reservoir level followed by detailed imaging of the target zone only (Berkhout, 1985).

6.5.4 Trace Interpolation

A typical 3-D survey has a trace spacing in the cross-line direction that normally is coarser than the trace spacing in the in-line direction, perhaps by as much as four times. This coarse spacing can cause spatial aliasing in the cross-line direction (Section 4.3.5). This problem can be overcome by trace interpolation.

A typical trace interpolation procedure involves determining dominant dip directions in the data based on trace crosscorrelations, say between 4 and 10 traces, within a sliding time gate. The dominant dip direction corresponds to the largest correlation value. A simple midpoint interpolation (Bracewell, 1965) between the time samples along that dip direction yields the amplitude value at the center of the gate (Rothman et al., 1981). Noise, variations in the waveform from trace to trace, and the complexity of structure affect the quality of the output from the interpolation procedure. Note that trace interpolation does not create data, it merely unwraps the spectrum (on the basis of the dominant dip pick) so that aliased frequencies are mapped to the correct quadrant in the frequency-wavenumber plane. Finally, data do not necessarily need to be trace-interpolated in the cross-line direction down to the trace spacing of the in-line direction. Instead, signal bandwidth and subsurface dip can be taken into consideration to compute the optimum trace spacing to avoid spatial aliasing [equation (4.17)].

(Text continued on page 416)

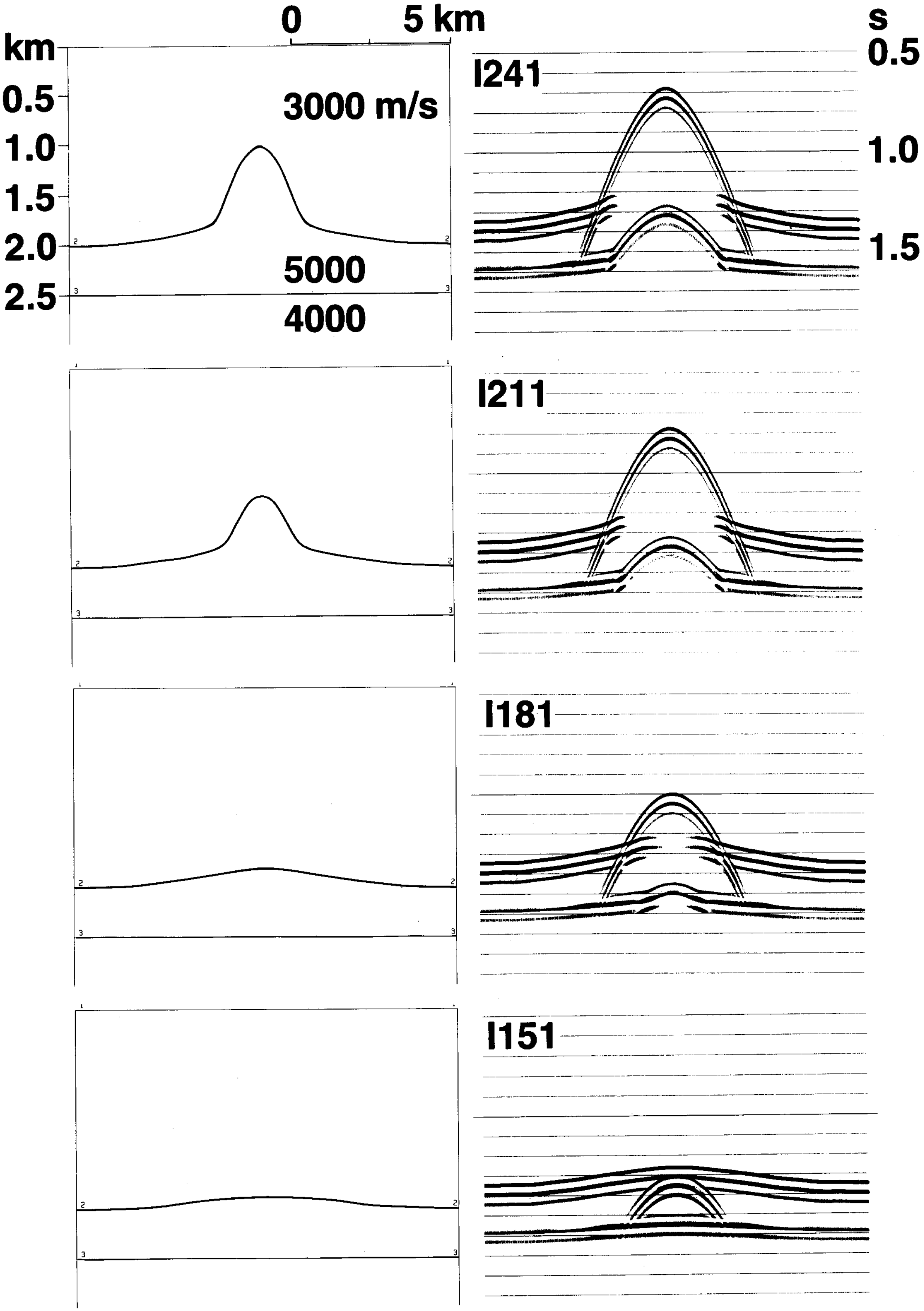


FIG. 6-32. Velocity-depth model (left column) and zero-offset 3-D synthetic data (right column) for a circularly symmetric salt dome. (The base map is shown in Figure 6-31.)

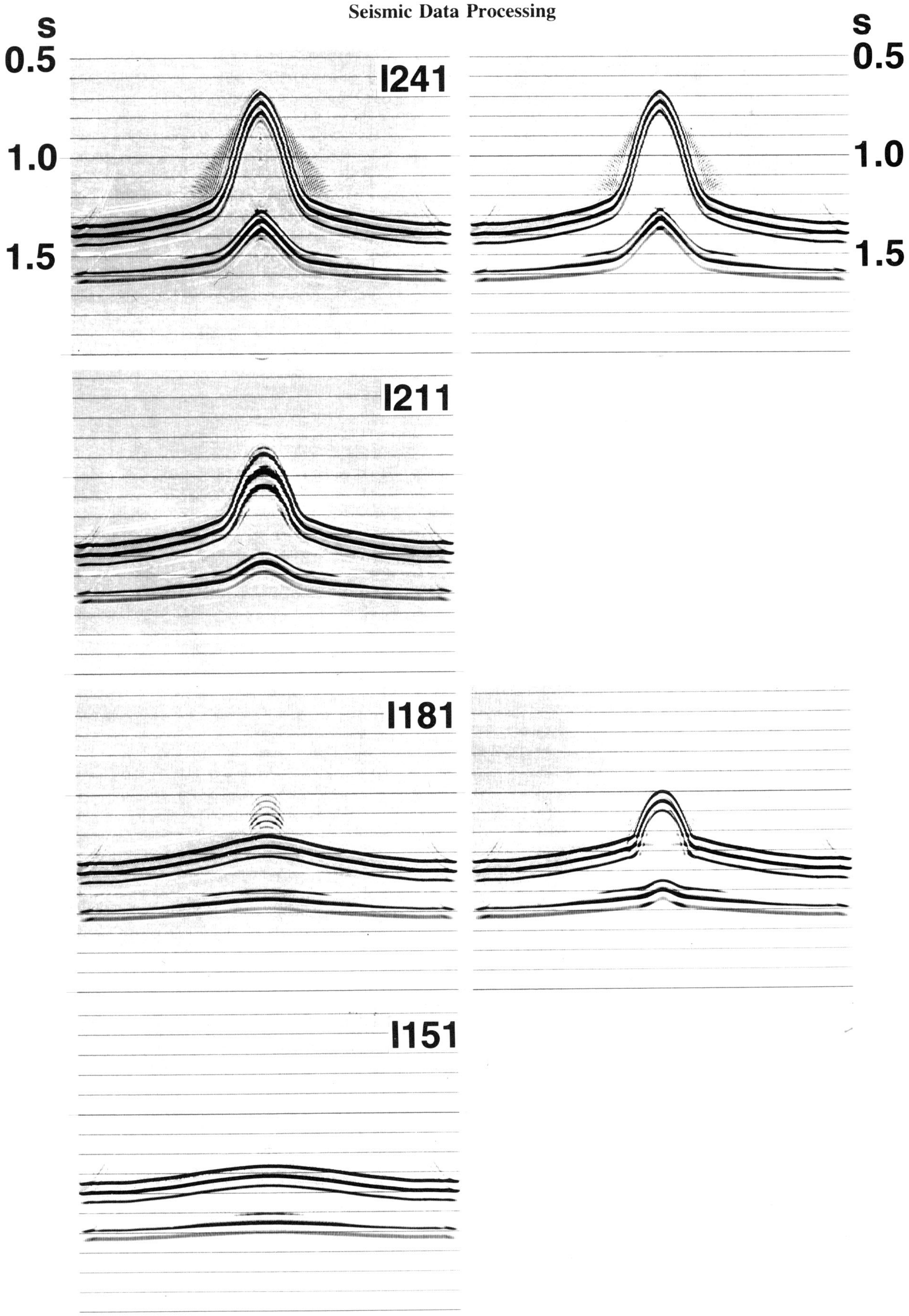


FIG. 6-33. A 3-D time migration (left column) versus a 2-D time migration (right column) of the synthetic data shown in Figure 6-32.

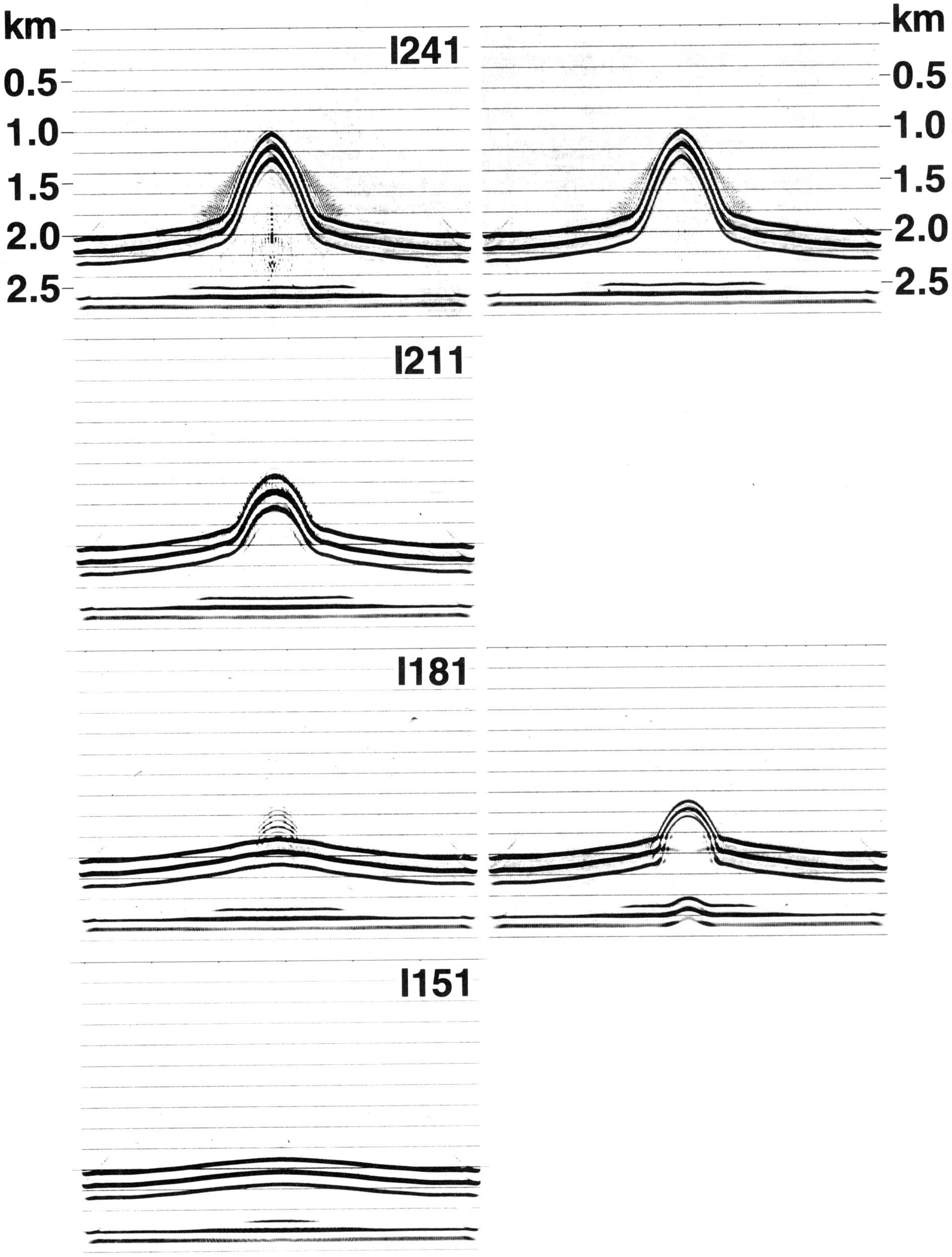


FIG. 6-34. A 3-D depth migration (left column) versus a 2-D depth migration (right column) of the synthetic data shown in Figure 6-32.

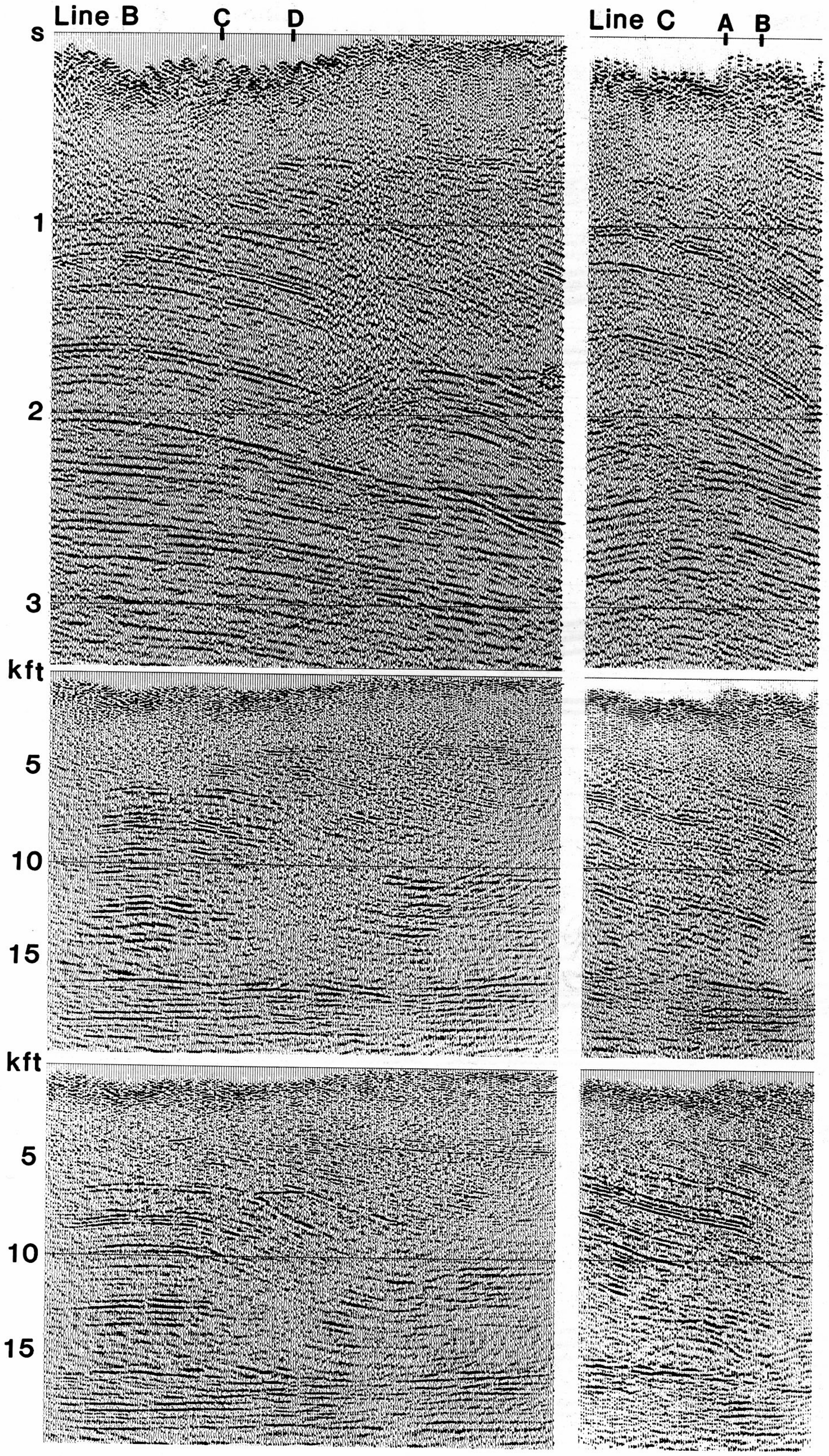


FIG. 6-35. An in-line (top left) and a cross-line stacked section (top right) from a land 3-D survey and the corresponding 2-D (center) and 3-D (bottom) depth migration results. (Data courtesy Chevron USA, Inc.)

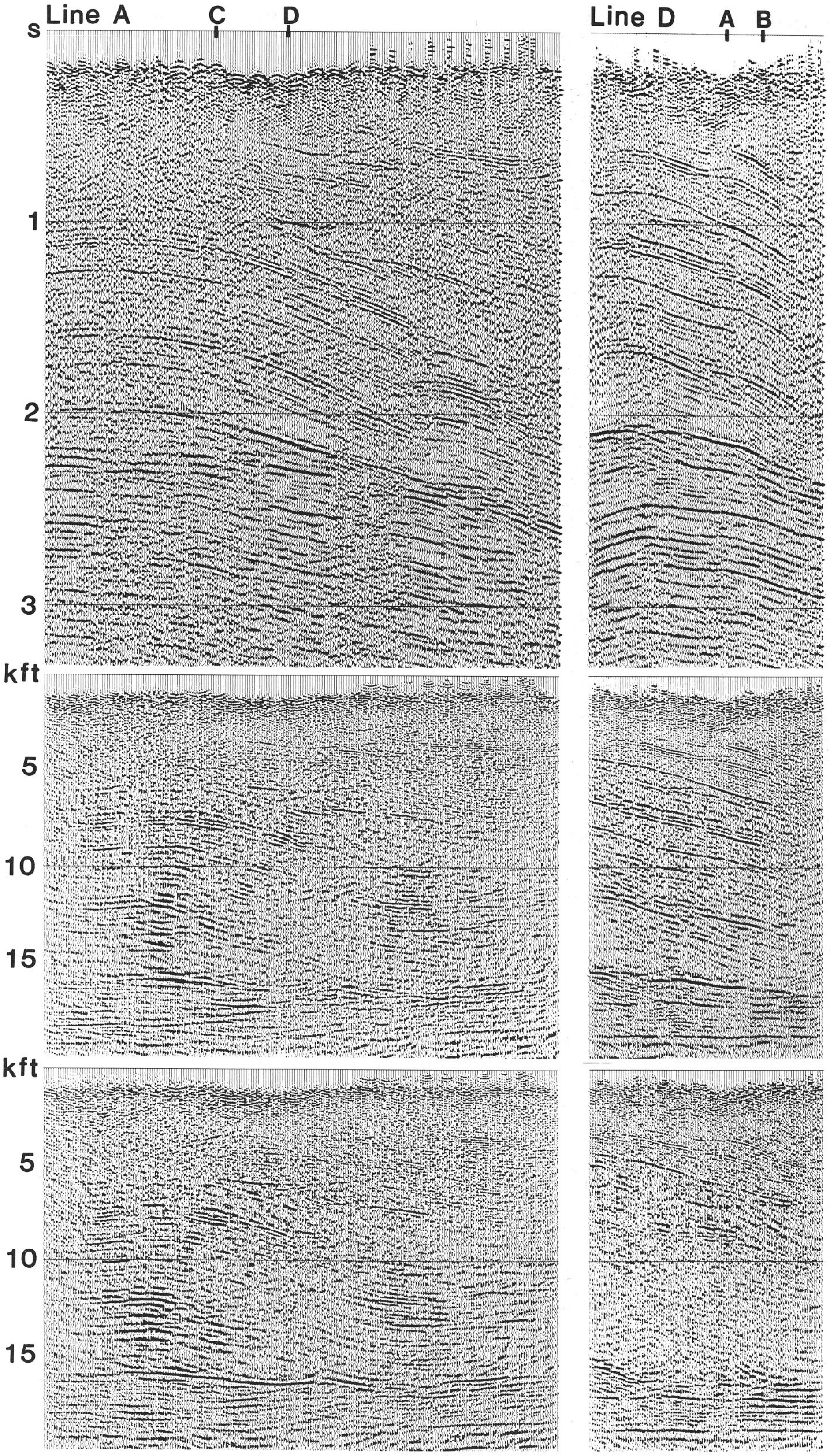


FIG. 6-36. Another in-line (top left) and cross-line stacked section (top right) from the same land 3-D survey as in Figure 6-35 with their 2-D (center) and 3-D (bottom) depth migrations. (Data courtesy Chevron USA, Inc.)

6.6 INTERPRETATION OF 3-D SEISMIC DATA

After 3-D migration, a 3-D data volume is suitable for derivation of the 3-D subsurface geological model. Because of the completeness of data within that volume, the interpreter has more information than he would from a 2-D reconnaissance-type data set. Although more information leads to less uncertainty in deriving the geological model, interpretation of abundant data can be exhausting. Use of an interactive interpretation workstation is a common way to deal with large 3-D data volumes. Moreover, an interactive environment is versatile in viewing the 3-D data volume. For example, we can examine vertical sections in in-line, cross-line, or any arbitrary direction, as well as horizontal sections, namely, time slices. An interactive environment also can provide the capability to improve interpretation. For example, horizon flattening, correlation of marker horizons across faults, and some image processing tools to enhance certain features within the data volume can be implemented. While 3-D interactive workstations are used heavily in 3-D interpretations, most final maps result from a combination of both interactive and conventional paper-oriented interpretations.

6.6.1 Time Slices

A time slice contains events from more than one reflecting horizon at the same time level. Time slices are used to generate structural contour maps. A (spatially) high-frequency event on a time slice is either a steeply dipping event or a high-frequency event in time. Thus, from the time slice in Figure 6-39, we can infer steep dip at location H from the high-frequency character of the event, and gentle dip at location L from the low-frequency character. If contours narrow between a shallow and deep time slice, then the feature is a structural low (Figure 6-39). Conversely, if contours widen between a shallow and deep time slice, then the feature is a structural high (Figure 6-40).

Some 2-D smoothing can be applied to time slices to ease contouring (Figure 6-41b). Edge enhancement can be used to better delineate zero crossings (Figure 6-41c). Some image processing tools also can be used to enhance subtle features on time slices. Illumination of a time slice by a pseudo-sun is an example of such a technique. The angle of illumination and its direction can be manipulated rapidly in an interactive environment. Figure 6-42 is an example of pseudo-sun illumination. The main structure here is a salt dome. Note the presence of tensional faults associated with salt diapirism. Bone et al. (1975) and Brown et al. (1982) offer extensive discussions on the use of time slices in 3-D interpretation.

Besides contouring, time slices also are useful in quality control. Figure 6-43 is a time slice before and after residual statics corrections were applied to the data. Note the improvement in the S/N ratio and the continuity of events in Figure 6-43b.

One common artifact that is observable on time slices is the presence of horizontal streaks (striations). The pseudo-sun illumination has enhanced the streaks in Figure 6-42. These striations typically are oriented along the shooting direction. Although there are several causes for time-slice striations, one cause is the positioning errors present in the navigation data. Schultz and Lau (1984) identify these striations as time shifts from one in line to another (cross-line statics). They also suggest a poststack procedure to estimate and eliminate these time shifts in the in-line/cross-line wavenumber domain.

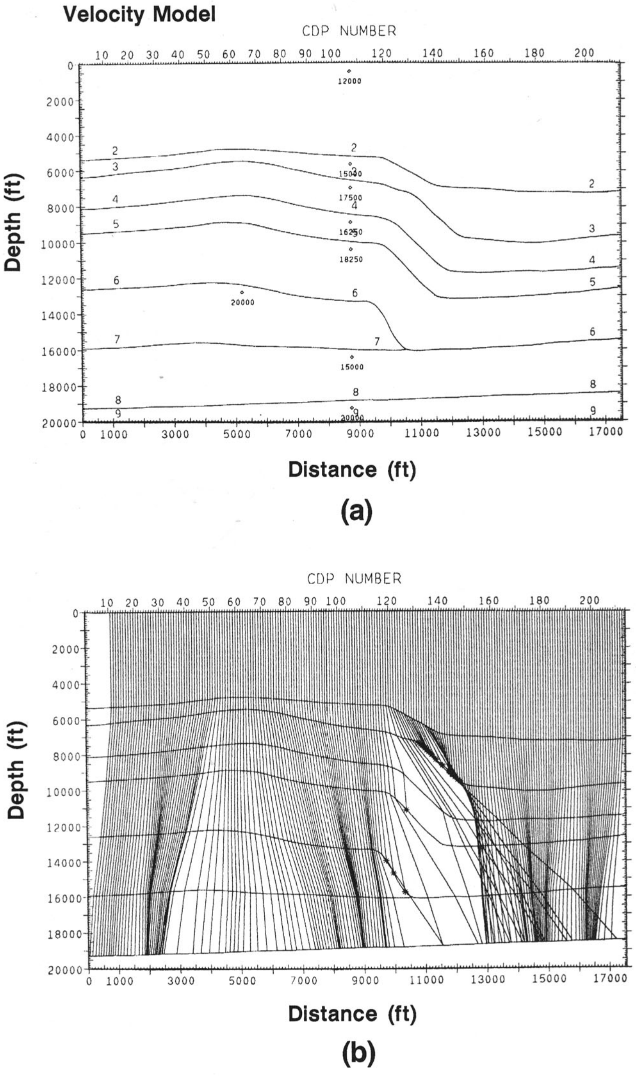
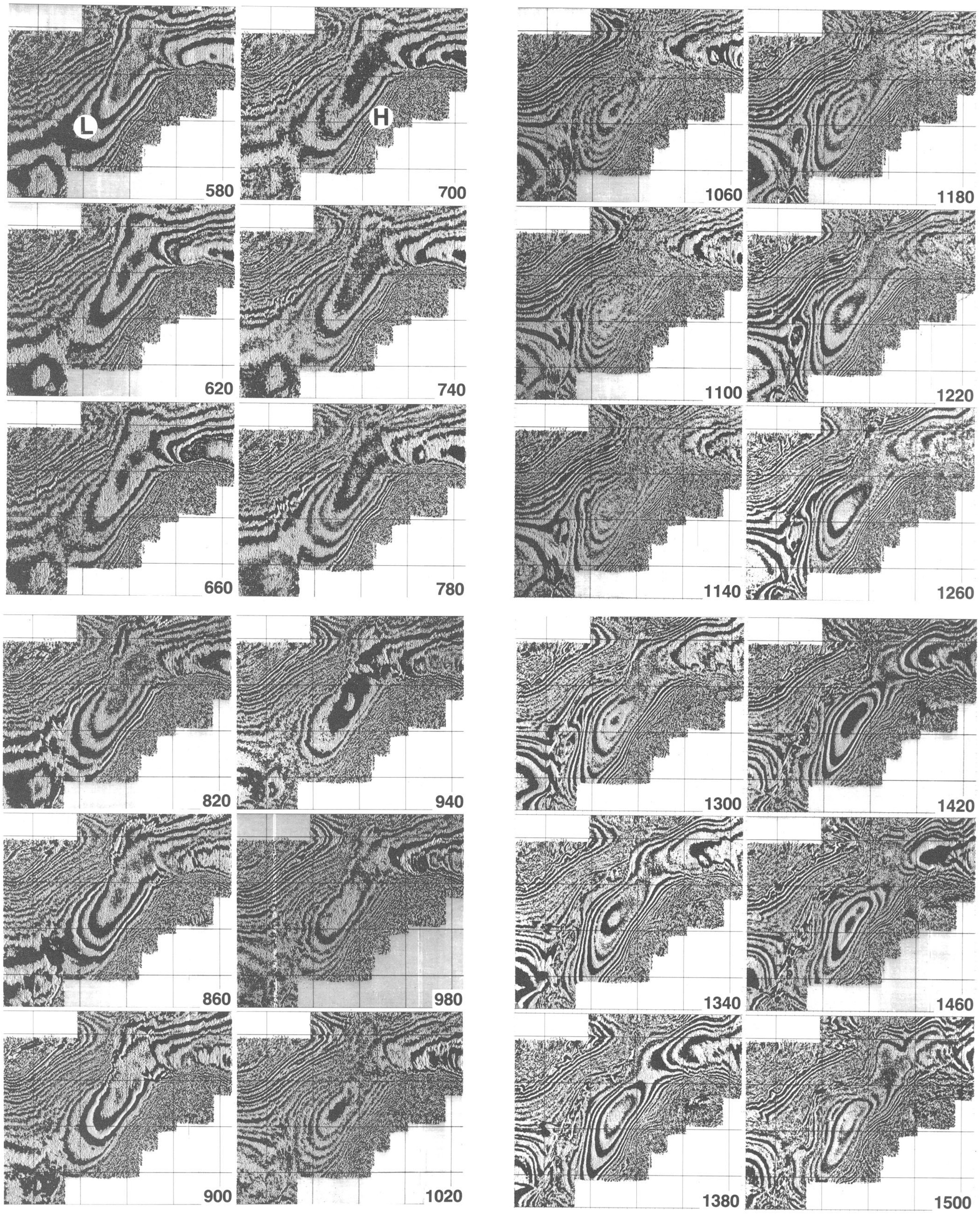


FIG. 6-37. (a) An in-line cross section of a 3-D velocity-depth model that was used in 3-D depth migration of the data shown in Figures 6-35 and 6-36. (b) Image rays through this velocity-depth model.


L
H
580
700
620
740
660
780
820
940
860
980
900
1020
1060
1180
1100
1220
1140
1260
1300
1420
1340
1460
1380
1500

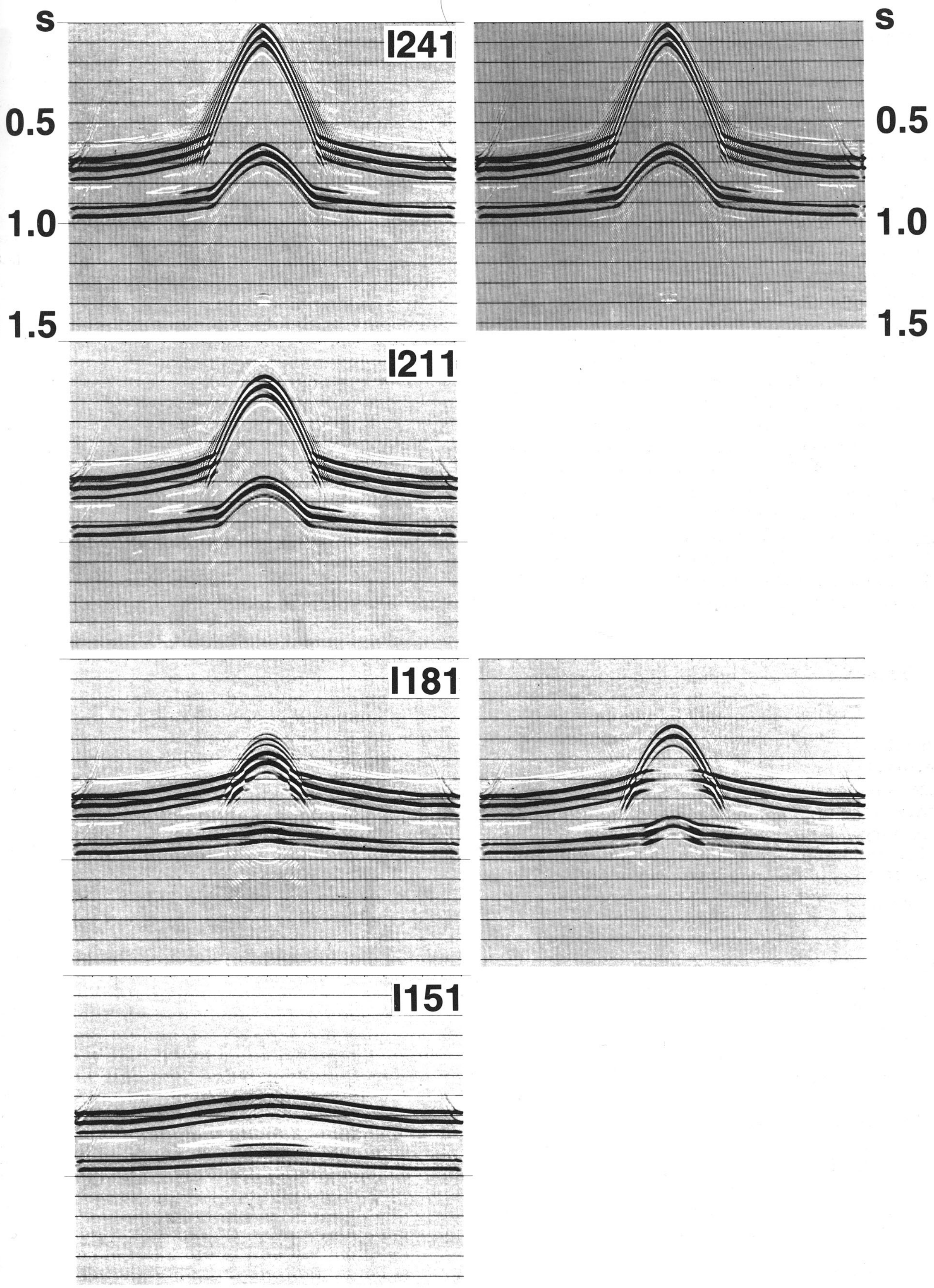


FIG. 6-38. Three-dimensional wave-equation datuming (left column) versus 2-D wave-equation datuming (right column) from the surface down to a 1000-m depth of the synthetic data shown in Figure 6-32.

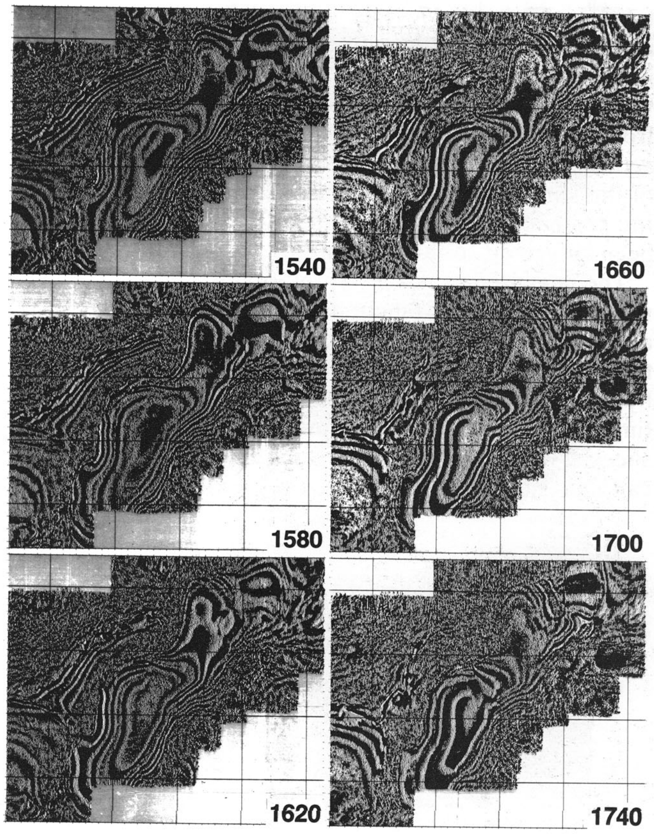


FIG. 6-39. Selected time slices from a land 3-D survey starting at 580 ms and moving down to 1740 ms at 40-ms intervals. The northeast-southwest extending feature is a small basin between two salt domes. (Data courtesy Nederlandse Aardolie Maatschappij B.V.)

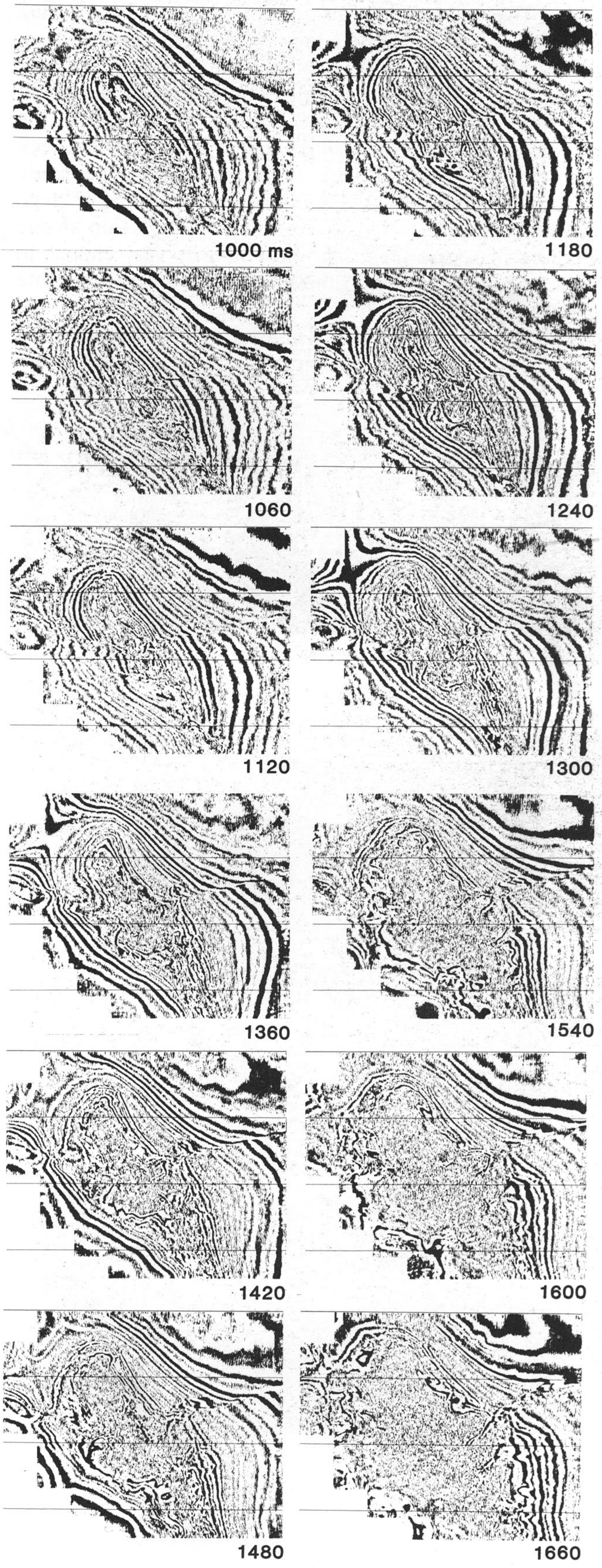


FIG. 6-40. Selected time slices from a land 3-D survey starting at 1000 ms and moving down to 1660 ms at 60-ms intervals. The irregular feature is the top of a salt dome. (Data courtesy Nederlandse Aardolie Maatschappij B.V.)

6.6.2 Interactive Interpretation Session

A 3-D interpretation session may begin with viewing selected in-line and cross-line sections to acquire a regional understanding of the subsurface geology. Other orientations, such as vertical sections along a dominant dip direction, also may be needed to determine the structural pattern. Time slices then are studied to check the structural pattern. These previews may be made dynamic in an interactive environment; vertical or horizontal sections can be viewed in rapid succession as one would a film strip from a motion picture. Any change in structure in space and time can thus be grasped quite effectively.

Figure 6-44 shows selected cross lines from a marine 3-D survey. For reference, in lines and cross lines are annotated in Figure 6-17. The structural configuration was inferred best from cross lines; therefore, only these are included. Every 50th line is displayed, starting from the west side of the survey and moving east. As we move from west to east, a model for the regional structural setting emerges. There is a structural high in the middle of the survey area. The time slices in Figure 6-45 verify the presence of this structural high and show that it has been subjected to intense faulting.

The next step in 3-D interpretation is to mark the main structural features (such as faults and synclinal and anticlinal axes) on the vertical sections, cross-check, then project these onto the time slices. This is followed by preliminary contouring of the marker horizons using time slices. The cross lines run vertically across the time slices. Two marker horizons were interpreted using vertical and horizontal sections. Horizons A and B are

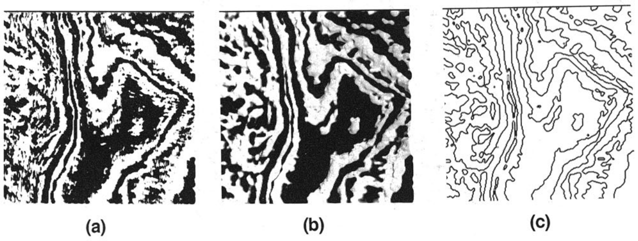


FIG. 6-41. Processing time-slice data (a) for improved interpretation, (b) 2-D smoothing, and (c) edge enhancement. (Data courtesy Shell Oil Company and Esso.)

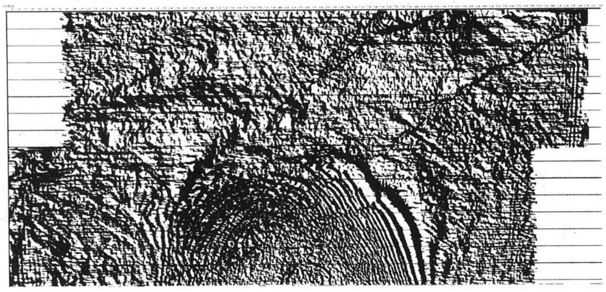


FIG. 6-42. Processing time-slice data for improved interpretation. Pseudo-sun illumination to enhance subtle structural features. (Data courtesy Sohio Petroleum Company.)

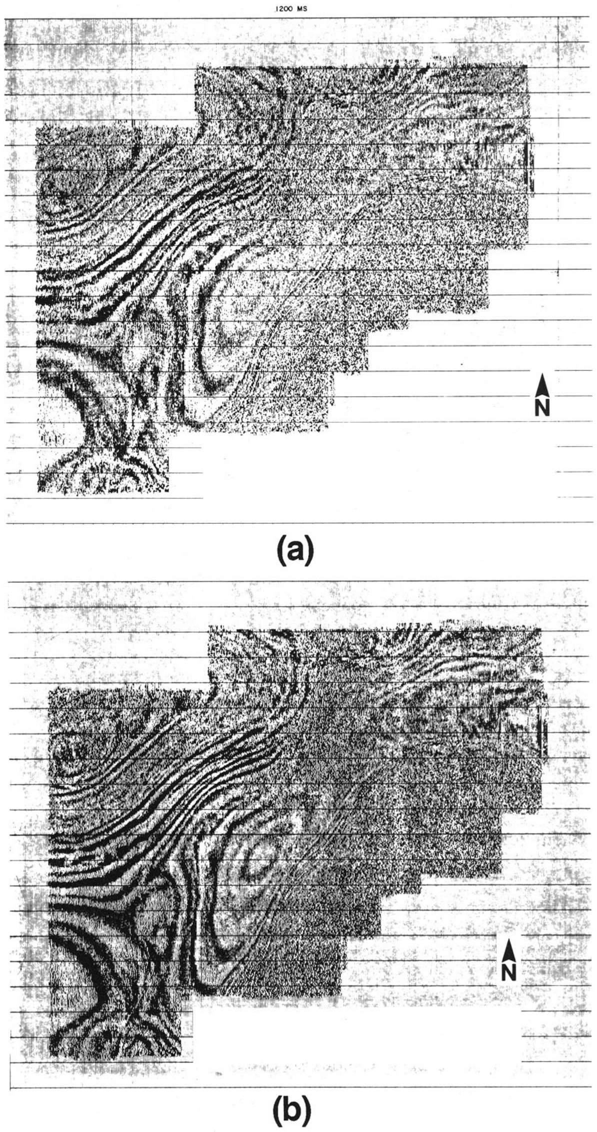


FIG. 6-43. A time slice (a) before and (b) after residual statics corrections. (Data courtesy Nederlandse Aardolie Maatschappij B.V.)

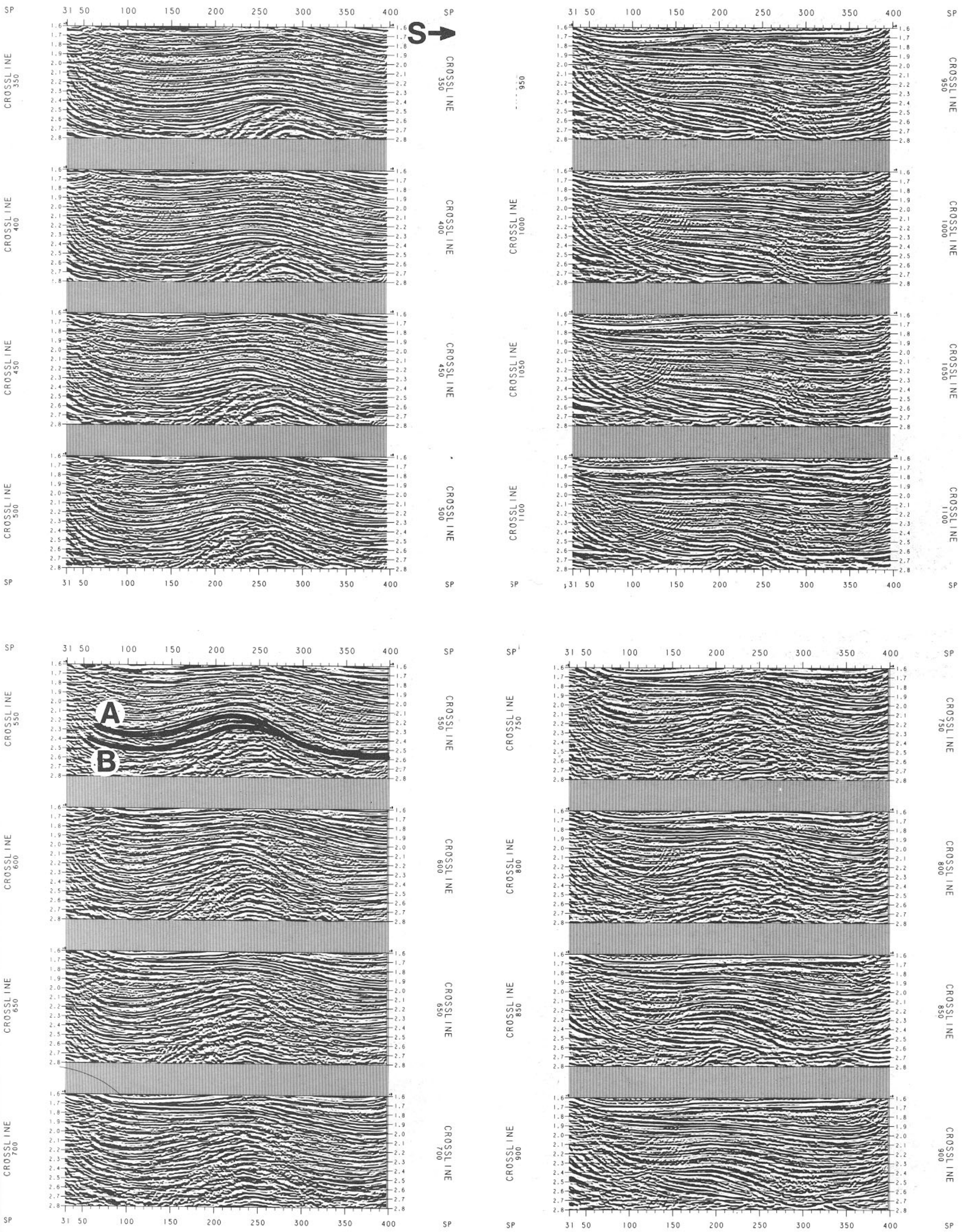


FIG. 6-44. Selected cross lines from a marine 3-D survey. In lines and cross lines are denoted in Figure 6-17.

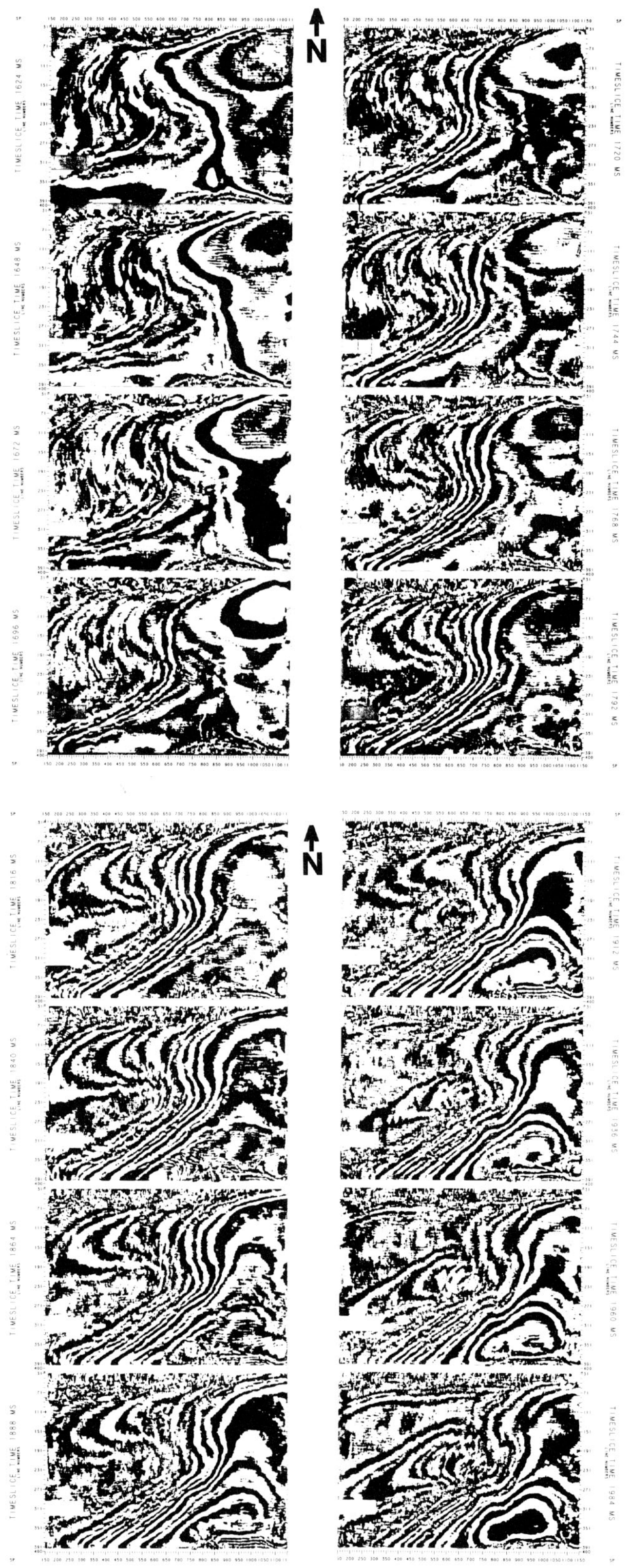
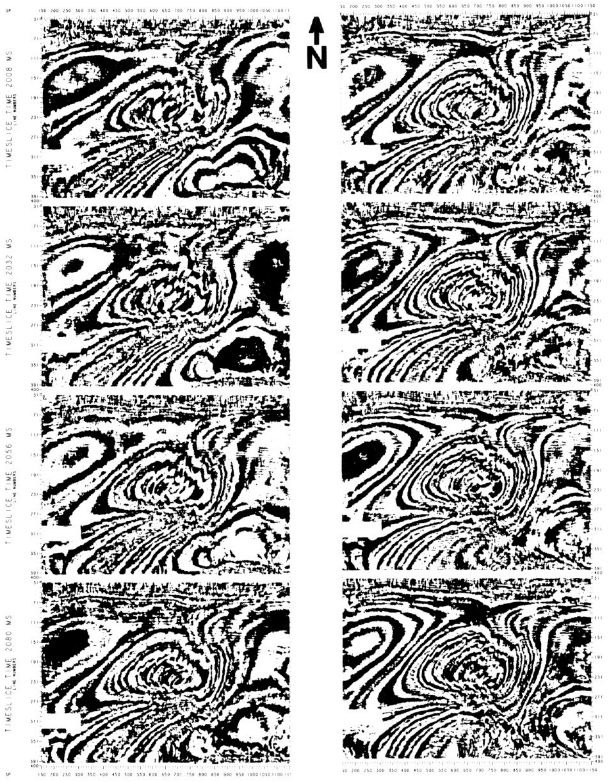
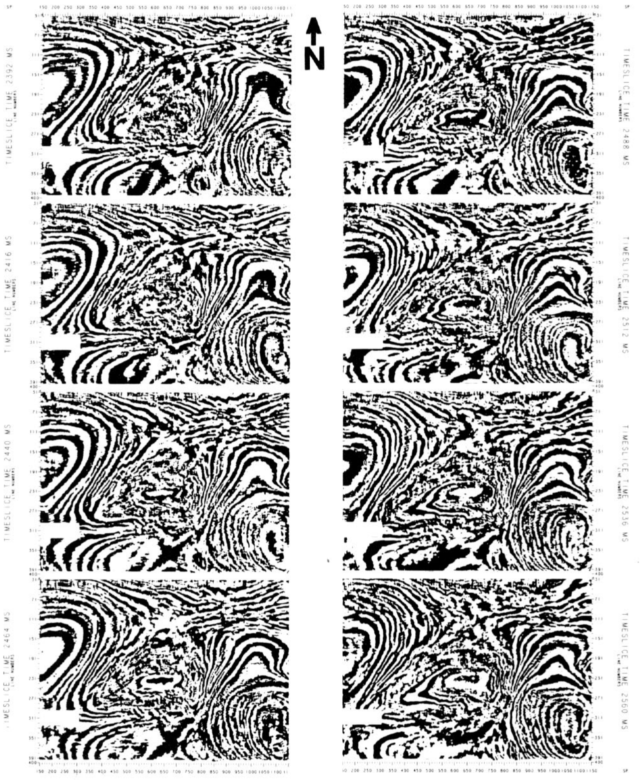
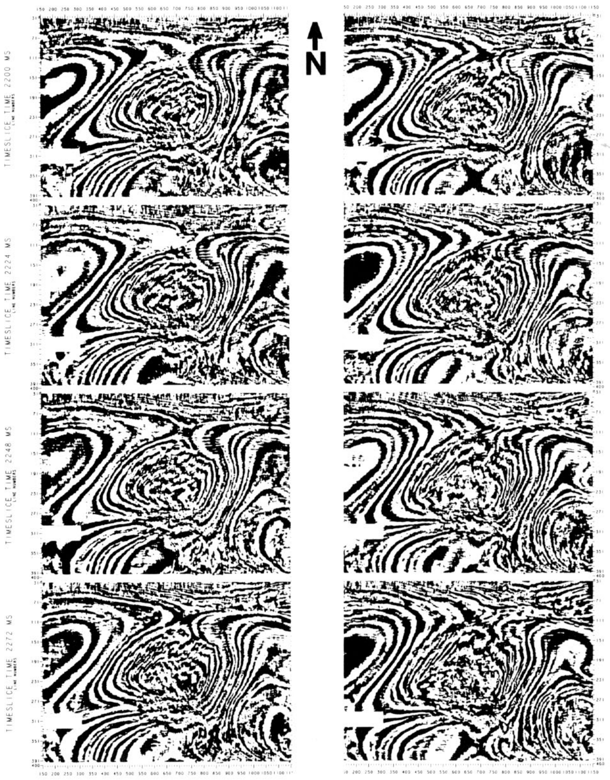
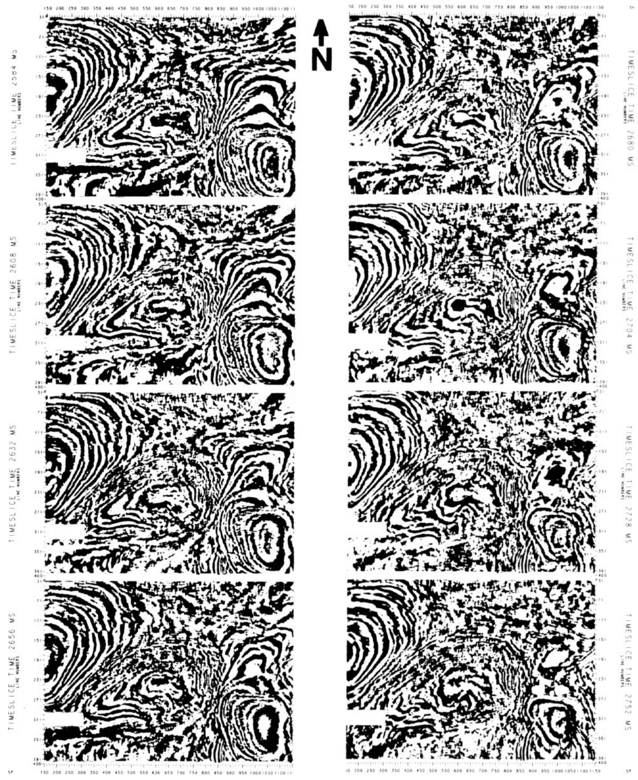


FIG. 6-45. Selected time slices from a marine 3-D survey starting at 1624 ms and moving down to 2752 ms at 24-ms intervals. Selected cross lines are shown in Figure 6-44.


TIMESLICE TIME 2008 MS
TIMESLICE TIME 2032 MS
TIMESLICE TIME 2056 MS
TIMESLICE TIME 2080 MS
LINE NUMBERS
SP
N


TIMESLICE TIME 2392 MS
TIMESLICE TIME 2416 MS
TIMESLICE TIME 2440 MS
TIMESLICE TIME 2464 MS
TIMESLICE TIME 2488 MS
TIMESLICE TIME 2512 MS
TIMESLICE TIME 2536 MS
TIMESLICE TIME 2560 MS
LINE NUMBERS
SP
N


TIMESLICE TIME 2200 MS
TIMESLICE TIME 2224 MS
TIMESLICE TIME 2248 MS
TIMESLICE TIME 2272 MS
LINE NUMBERS
SP
N


TIMESLICE TIME 2584 MS
TIMESLICE TIME 2608 MS
TIMESLICE TIME 2632 MS
TIMESLICE TIME 2656 MS
TIMESLICE TIME 2680 MS
TIMESLICE TIME 2704 MS
TIMESLICE TIME 2728 MS
TIMESLICE TIME 2752 MS
LINE NUMBERS
SP
N

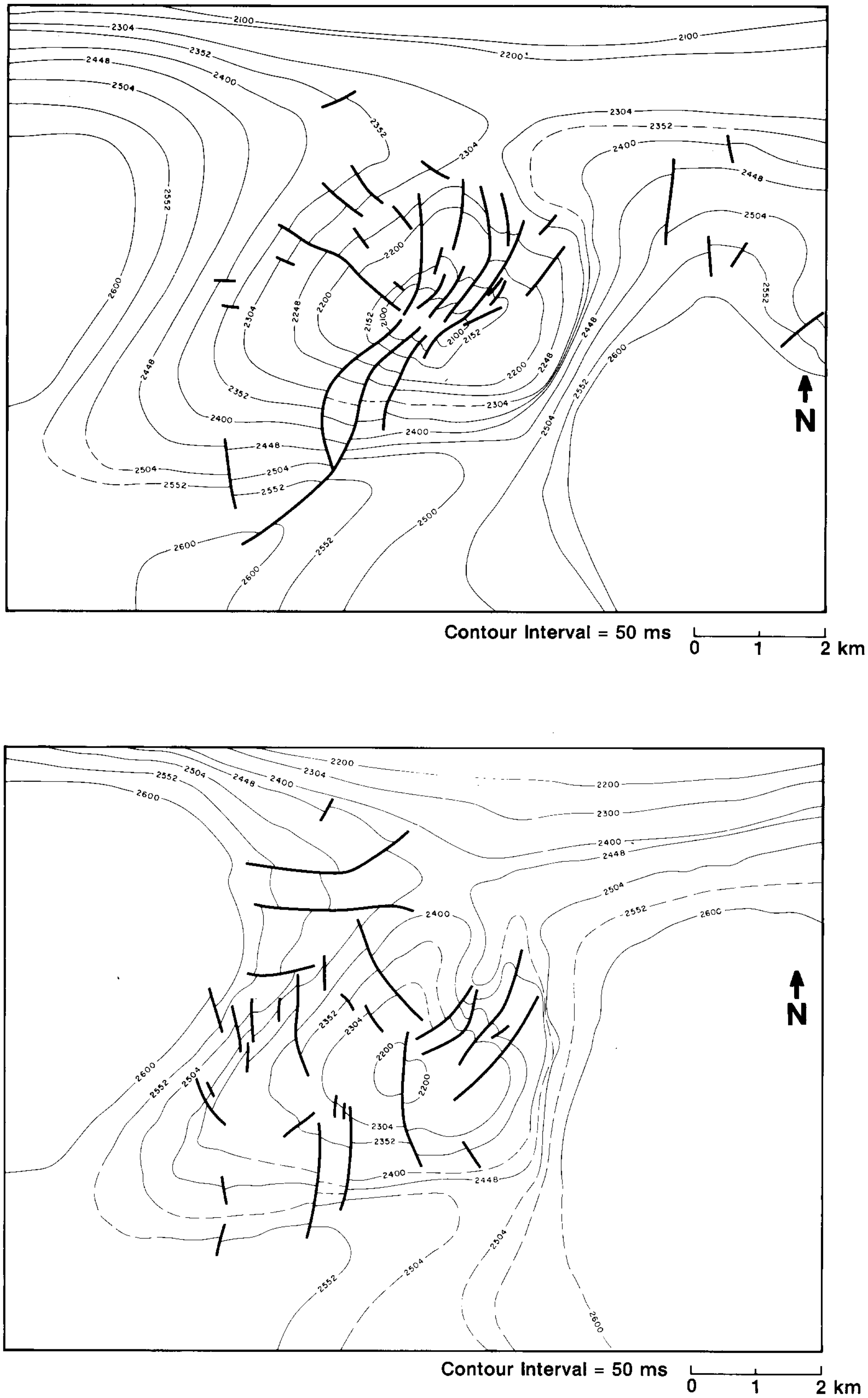


FIG. 6-46. Preliminary structural contour maps of horizons A and B as indicated in Figure 6-44 (cross line 550) from the marine 3-D survey. Selected vertical and horizontal sections from this survey are shown in Figures 6-44 and 6-45, respectively.

indicated for reference on cross line 550 (Figure 6-44). The time structure maps (Figure 6-46) for horizons A and B were obtained from the preliminary contouring of time slices 50 ms apart. The final stage of interpretation involves continuously tracking marker horizons along all of the vertical sections, then correlating across the faults. The interpreted horizon and fault information is stored in a data base and later retrieved to construct a complete 3-D subsurface image.

The following publications provide more information on 3-D case histories: Blake et al., 1982; Bone, 1978; Bone et al., 1975; Brown, 1978; Brown et al., 1981; Brown et al., 1982; Brown and McBeath, 1980; Dahm and Graebner, 1982; French, 1974; Galbraith and Brown, 1982; Hautefeuille and Cotton, 1979; Johnson and Bone, 1980; Kurfess et al., 1977; Leflaive and Paturet, 1981; Saeland and Simpson, 1982; Sanders and Steel, 1982; and Tegland, 1977.

REFERENCES

Bentley, L. and Yang, M., 1982, Scatter of midpoints grouped in cells and its effects on stacking: unpublished technical document, Western Geophysical Company.

Berkhout, A. J., 1985, 3-D seismic processing with an eye to the future: World Oil, **201**, 5, 91-94.

Blake, B. A., Jennings, J. B., Curtis, M. P., and Philipson, R. M., 1982, Three-dimensional seismic data reveals the finer structural details of a piercement salt dome: Offshore Technology Conference, Paper 4258, 403-406.

Bone, M. R., 1978, A case history of 3D seismic application in the Gulf of Mexico: Offshore Technology Conference, Paper 3176, 1106-1118.

Bone, M. R., Giles, B. F., and Tegland, E. R., 1975, Analysis of seismic data using horizontal cross-sections: Presented at the 45th Ann. Int. Soc. Explor. Geophys. Mtg.

Bracewell, R., 1965, The Fourier transform and applications: McGraw-Hill Book Company.

Brown, A. R., 1978, 3-D seismic interpretation methods: Presented at the 48th Ann. Int. Soc. Explor. Geophys. Mtg.

Brown, A. R. and McBeath, R. G., 1980, 3-D seismic surveying for field development comes of age: Oil and Gas J., **78**, *46*, 63-65.

Brown, A. R., Dahm C. G., and Graebner, R. J., 1981, A stratigraphic case history using three-dimensional seismic data in the Gulf of Thailand: Geophys. Prosp., **29**, 327-349.

Brown, A. R., Graebner, R. J., and Dahm, C. G., 1982, Use of horizontal seismic sections to identify subtle traps, in The Deliberate Search for the Subtle Trap: Am. Assoc. Petr. Geol. Memoir **32**, 47-56.

Claerbout, J. F., 1985, Imaging the earth's interior: Blackwell Scientific Publications.

Dahm, C. G. and Graebner, R. J., 1982, Field development with three-dimensional seismic methods in the Gulf of Thailand—a case history: Geophysics, **47**, 146-176.

Deregowski, S. and Rocca, F., 1981, Geometrical optics and wave theory for constant-offset sections in layered media: Geophys. Prosp., **29**, 374-387.

French, W. S., 1974, Two-dimensional and three-dimensional migration of model-experiment reflection profiles: Geophysics, **39**, 265-277.

Galbraith, R. M. and Brown, A. R., 1982, Field appraisal with three-dimensional seismic surveys in offshore Trinidad: Geophysics, **47**, 177-195.

Gardner, G. H. F., McDonald, J. A., Watson, T. H., and Kotcher, J. S., 1978, An innovative 3-D marine seismic survey: Presented at the 40th Ann. Eur. Assoc. Explor. Geophys. Mtg.

Hautefeuille, A. and Cotton, W. R., 1979, Three-dimensional seismic surveying aids exploration in the North Sea: Oil and Gas J., **77**, *45*, 72-79.

Homabrook, J. T., 1974, Seismic re-interpretation clarifies North sea structures: Petr. Int., April-May.

Jakubowicz, H. and Levin, S., 1983, A simple exact method of 3-D migration— theory: Geophys. Prosp., **31**, 34-56.

Johnson, J. P. and Bone, M. R., 1980, Understanding field development history utilizing 3D seismic: Offshore Technology Conference, Paper 3849, 473-475.

Kurfess, J. A., Giles, B. F., and Bone, M. R., 1977, Field development with 3-D seismic methods: A case history: Presented at the 47th Ann. Int. Soc. Explor. Geophys. Mtg.

Leflaive, R. and Paturet, D., 1981, Three-dimensional seismic structural interpretation of Parentis oil field, France: Presented at the 51st Ann. Int. Soc. Explor. Geophys. Mtg.

Lehmann, H. J. and Houba, W., 1985, Practical aspects of determination of 3-D stacking velocities: Geophys. Prosp., **33**, 155-163.

Levin, F. K., 1971, Apparent velocities from dipping interface reflections: Geophysics, **36**, 510-516.

Levin, F. K., 1983, The effects of streamer feathering on stacking: Geophysics, **48**, 1165-1171.

Levin, F. K., 1984, The effect of binning on data from a feathered streamer: Geophysics, **49**, 1386-1387.

Meinardus, H. A. and McMahon, I. T., 1981, Velocity modeling for 3-D NMO and stack: Paper presented at the 51st Ann. Int. Soc. Explor. Geophys. Mtg.

Ristow, D., 1980, 3-D downward extrapolation of seismic data in particular by finite difference methods: PhD thesis, University of Utrecht, The Netherlands.

Rothman, D., Larner, K. L., and Chambers, R., 1981, Trace interpolation and design of 3-D surveys: Paper presented at the 39th Ann. Eur. Assoc. Explor. Geophys. Mtg.

Saeland, G. T. and Simpson, G. S., 1982, Interpretation of 3-D data in delineating a subunconformity trap in Block 34-10, Norwegian North Sea: Am. Assoc. Petr. Geol. Memoir **32**, 217-235.

Sanders, J. I. and Steel, G., 1982, Improved structural resolution from 3D surveys in Australia: Austral. Petr. Explor. Assoc. J., **22**, 17-41.

Schneider, W. A., 1978, Integral formulation for migration in two and three dimensions: Geophysics, **43**, 49-76.

Schultz, P. S. and Lau, A., 1984, Poststack estimation of three-dimensional crossline statics: Geophysics, **49**, 227-236.

Tegland, E. R., 1977, 3-D seismic techniques boost field development: Oil and Gas J., **75**, *37*, 76-82.

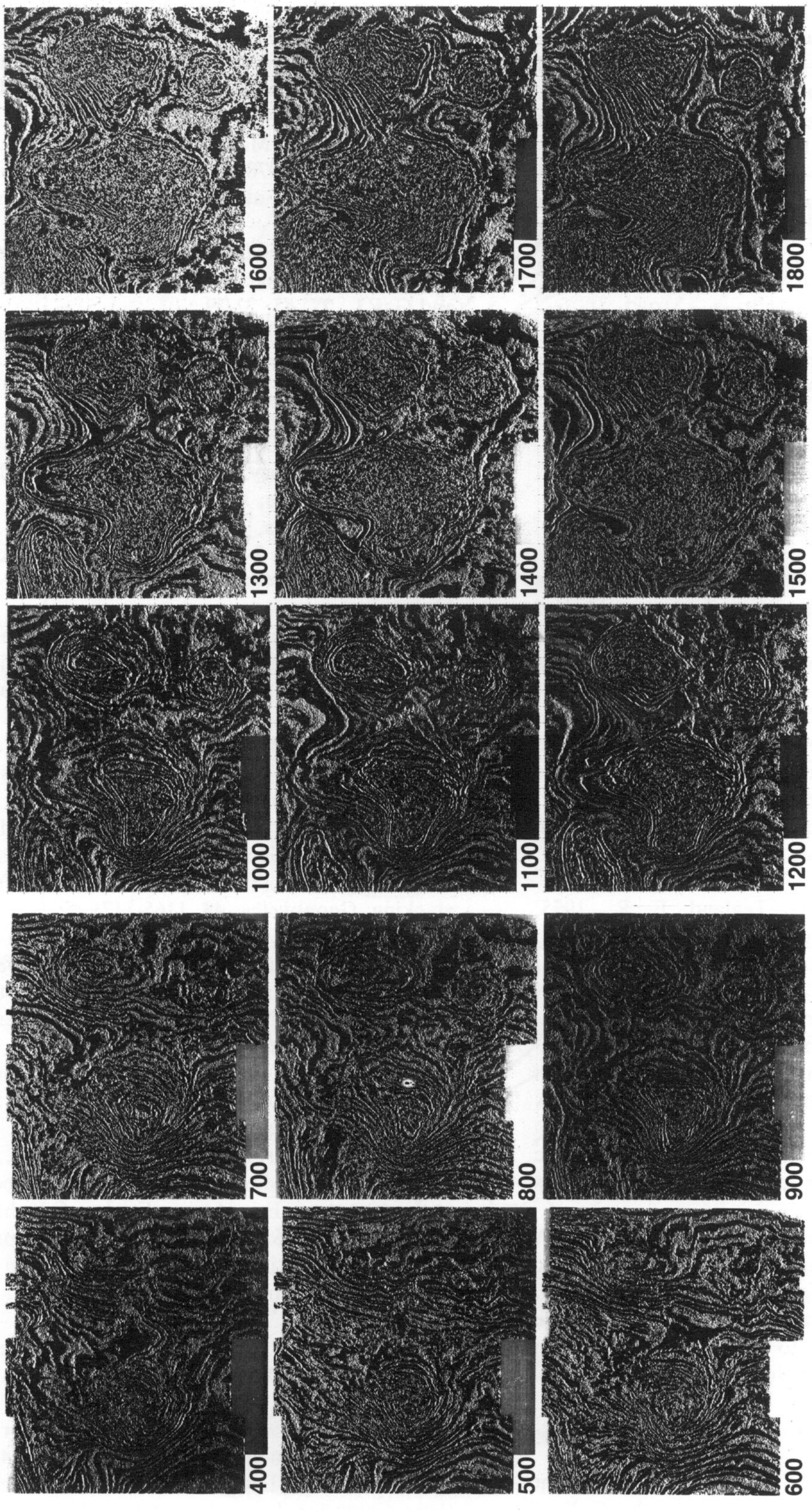


FIG. 6-47. Time slices from a marine 3-D survey (see Exercise 6.1). (Data courtesy Amoco Europe and West Africa, Inc.)

Workman, R., 1984, Marine 3-D acquisition and processing: unpublished technical document, Western Geophysical Company.

EXERCISES

Exercise 6.1. From the time slices in Figure 6-47, infer the structural setting of the subsurface.

Exercise 6.2. At what angle is the pseudo-sun illuminating the time slice in Figure 6-42?

Exercise 6.3. What are the advantages and disadvantages of dip-line and strike-line shooting in terms of (a) cross-line smear, (b) DMO, (c) spatial aliasing, and (d) velocity estimation?

Exercise 6.4. Suppose you did a 2-D survey over the salt dome structure shown in Figure 6-32 and applied 2-D imaging to your data. Then you did a 3-D survey and 3-D imaging. Consider mapping the top of the salt. Which one, 2-D or 3-D seismic, will imply a salt dome structure with a larger spatial extent and structural closure?

Exercise 6.5. Identify the prominent sideswipe energy on Lines I181 and I151 in Figure 6-32 (right column).

Exercise 6.6. If you use the method of Stolt with stretch (Section 4.2.3) in 3-D time migration, do you have to do it in one pass?

Exercise 6.7. In areas with complex overburden, which one is more important: (a) 3-D poststack depth migration or (b) 2-D prestack depth migration.

Exercise 6.8. Consider a dual-boat operation in a 3-D survey such that line spacings alternate between 112.5 and 37.5 m. Is this equivalent to a single-boat operation with 75-m line spacing?

Exercise 6.9. How would you obtain 3-D migration of a cross-line section without having to do 3-D migration of the entire 3-D data set?

7

Slant Stack and Applications

7.1 INTRODUCTION

In Section 1.6, we learned that the 2-D Fourier transform is one way to decompose a wave field into its plane-wave components, each with a unique frequency and each traveling at a unique angle from the vertical direction. In this chapter, the domain of ray parameter is discussed and another way to decompose the wave field into its plane-wave components is presented. Plane-wave decomposition of a wave field [such as a common-shot gather (CSG)] can be achieved by applying linear moveout (LMO) and summing amplitudes over the offset axis. This procedure is called slant stacking. An underlying assumption of slant stacking is that of a horizontally layered earth. Conventional processing primarily is done in midpoint-offset coordinates. Slant stacking replaces the offset axis with the ray parameter p axis. (The ray parameter is the inverse of the horizontal phase velocity.) A family of traces with a range of p values is called a slant-stack gather.

Several processing techniques have been devised in midpoint-ray-parameter coordinates. Examples include trace interpolation (Section 7.2), dip filtering (Section 7.4), multiple suppression (Section 7.5), refraction inversion (Appendix D), and migration and velocity analysis. Taner (1977) was the first to introduce the midpoint-ray-parameter coordinates. He discussed the interpretive use of plane-wave stacks, where several constant-p sections are superpositioned over a restricted range of p values to enhance dipping events. Other processing methods were investigated later, such as migration (Ottolini, 1982) and velocity analysis (Schultz and Claerbout, 1978; Diebold and Stoffa, 1981; and Gonzalez-Serrano, 1982). Alam and Lasocki (1981a) and Alam and Austin (1981b) discussed possible applications in trace interpolation and multiple suppression, respectively. Clayton and McMechan (1981) devised a method to invert a refracted wave field, which involves downward continuation in the slant-stack domain. McMechan and Yedlin (1981) devised a method to obtain phase velocity curves for dispersive waves using slant-stack transformation. Based on downward continuation of a slant-stack gather, Schultz (1982) developed a technique to estimate interval velocities.

We now examine the physical aspects of constructing a slant-stack gather, commonly referred to as tau-p-gather or just p-gather. Each trace in this gather represents a plane wave that propagates at a certain angle from vertical. In reality, when a dynamite source explodes, the energy propagates at all angles (Figure 7-1). The reflected energy arrives at different receiver groups at different angles because of the offset between source and receiver locations. The farther the offset or the shallower the reflecting interface, the more oblique the angle of the upcoming wavefront.

As an aid in defining a scheme for constructing slant-stack gathers, first consider how plane waves can be generated. Figure 7-2 shows a line of point sources. Assume that this line of sources is activated so that all points on the line are excited simultaneously, and each point generates a spherical wave field. Some distance from the surface, the spherical wavefronts superimpose and result in a *plane wave* that travels vertically downward. This plane wave reflects from an interface and is recorded by a receiver at the surface. (There are source types, such as Geoflex* and Primacord,** which approximate short line sources.)

A plane wave that travels at a desired angle from vertical can be generated using the same line of point sources (Figure 7-3). To do this, the point sources must be activated in succession starting at one end of the line with an equal time delay between them. When a particular point source is activated, the wavefront generated from the previous source location already will have traveled a certain distance into the earth. When all the spherical wavefronts generated by the various sources superimpose, the result is a tilted plane wavefront (Figure 7-3). This plane wave then propagates, reflects from an interface, and is recorded by a receiver at the surface.

The amount of tilt of the wavefront (or the angle of propagation of the plane wave) can be controlled. Consider the geometry in Figure 7-4. By the time the wavefront generated at source location S_1 reaches point A in the subsurface, the point source at location S_2 should be excited so that the desired angle is attained. Define the distance between S_1 and S_2 as Δx, and the medium velocity with which the wave travel as v. If it takes Δt time for the wavefront to go from S_1 to A, then the dip angle θ of the plane wave is,

$$\sin \theta = v\Delta t/\Delta x. \tag{7.1}$$

The active source location must travel with speed $\Delta x/\Delta t = v/\sin\theta$ along the horizontal direction and the point source at location S_2 must be excited so that we can catch the wavefront at S_1 as it reaches point A in the subsurface. The velocity with which the source location must move ($v/\sin\theta$) is called the *horizontal phase velocity.*

From the experiments in Figures 7-2 and 7-3, note that a plane wave propagating at an angle from vertical can be generated by:

1. Placing a line of point sources on the earth's surface.
2. Exciting the point sources in succession with a time delay.
3. Superimposing the responses that are in the form of spherical wavefronts.

The superimposed response is recorded on a single receiver (Figure 7-3). This response is in the form of a plane wave that is reflected from an interface. Superposition means summing over the shot axis for a given receiver location. Using the reciprocity principle, summation also can be performed over the receiver axis for a given shot location.

The above discussion indicates how a CSG, as a single wave field, can be decomposed into its plane-wave

---

*Geoflex is a registered trademark of Imperial Chemical Industries.

**Primacord is a registered trademark of Ensign-Bickford.

components. By replacing the shot axis in Figure 7-4 with the receiver axis, the raypath geometry in Figure 7-5 results. The time delay associated with the plane wave that travels at angle θ from the vertical is given by:

$$\Delta t = (\sin \theta / v)\Delta x. \tag{7.2}$$

Snell's law says that the quantity $\sin \theta / v$, which is the inverse of the horizontal phase velocity, is constant along a raypath in a layered medium (Figure 7-6). This constant is called the ray parameter p. Equation (7.2) then is rewritten as:

$$\Delta t = p\Delta x. \tag{7.3}$$

The angle of propagation of the plane wave is controlled by adjusting the p value. Setting $p = 0$ corresponds to a plane wave that travels vertically. Given p and the velocity model for the layered earth, the family of raypaths associated with a particular p value can be traced as shown in Figure 7-7. A plane wave that travels in a layered earth is called a Snell wave (Claerbout, 1978). This type of plane wave changes its direction of propagation at each layer boundary according to Snell's law (Figure 7-6). For a single p value, note that the signal is recorded at many offsets (Figure 7-7). In general, receivers at all offsets record plane waves of many p values. To decompose the offset gather into plane-wave components, all the trace amplitudes in the gather must be summed along several slanted paths, each with a unique time delay defined by equation (7.3).

We have discussed decomposing a CSG wave field into its plane-wave components. As long as there is no dip, the traveltime curves in CSG and CDP gathers are indistinguishable (Figure 7-8). Since a CDP gather is not a single wave field, plane-wave decomposition would not seem to apply to CDP gathers. However, the equivalence of CDP gathers and CSG in a horizontally layered earth provides a rationale for applying plane-wave decomposition to both types of gathers.

7.2 CONSTRUCTION OF SLANT STACKS

Two steps typically are used in synthesizing plane waves by summing amplitudes in the offset domain along slanted

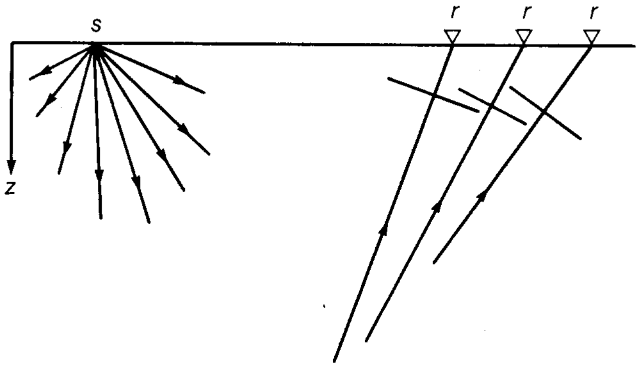


FIG. 7-1. A seismic source generates waves that propagate in all directions; waves traveling in different directions are recorded at different receiver locations.

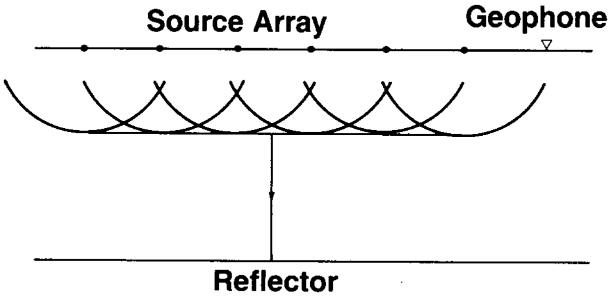


FIG. 7-2. A vertically incident plane wave is generated by setting off many shots in unison.

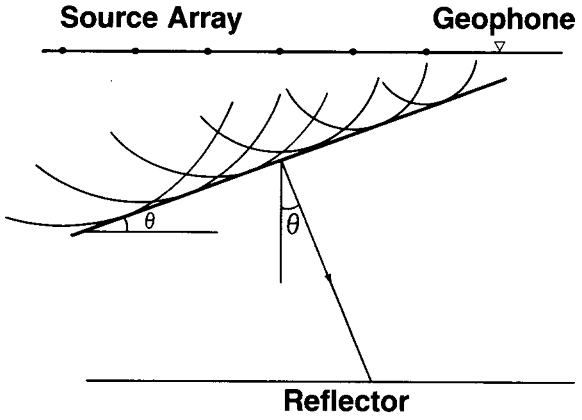


FIG. 7-3. A plane wave that travels at angle θ from vertical is generated by setting off many shots (starting from left) at appropriate time intervals.

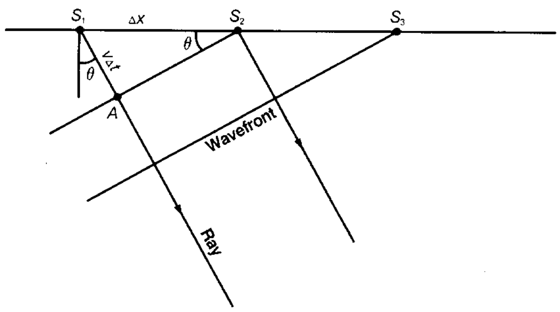


FIG. 7-4. Computation of time interval between shots (S) to generate the oblique plane wave in Figure 7-3.

paths. First, a linear moveout (LMO) correction is applied to the data through a coordinate transformation (Claerbout, 1978):

$$\tau = t - px, \tag{7.4}$$

where p is the ray parameter, x is the offset, t is the two-way traveltime, and τ is the linearly moved out time. After LMO, an event with slope p on input is flat. Next, the data are summed over the offset axis to obtain:

$$S(p, \tau) = \sum_x P(x, \tau + px). \tag{7.5}$$

Here, $S(p, \tau)$ represents a plane wave with ray parameter $p = \sin \theta / v$. By repeating the LMO for various values of p and performing the sum [equation (7.5)], the complete slant-stack gather (or p-gather), which consists of all the dip components in the original offset data, is constructed.

There is a distinction between slant stack and exact plane-wave decomposition of a wave field. Treitel et al. (1982) mathematically analyzed the plane-wave decomposition process and distinguished between conventional slant stack described here and the proper slant stack. A conventional slant stack yields an exact plane-wave decomposition when we deal with line sources; a proper slant stack yields an exact plane-wave decomposition when we deal with point sources. A proper slant stack is generated using the same steps described for a conventional slant stack except that a convolution of the linearly moved out wave field by a filter operator is performed

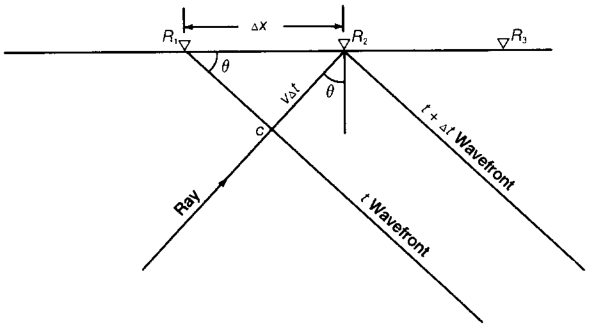


FIG. 7-5. The reciprocity principle applied to the geometry of Figure 7-4 to replace shots (S) with receivers (R).

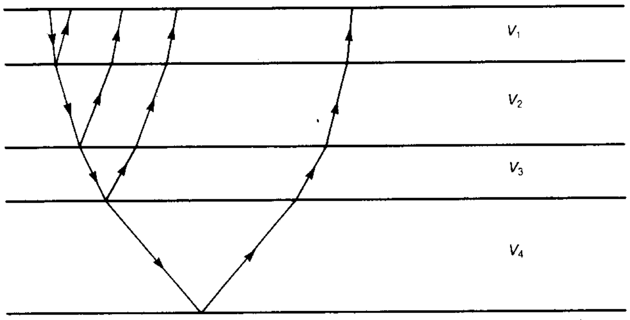
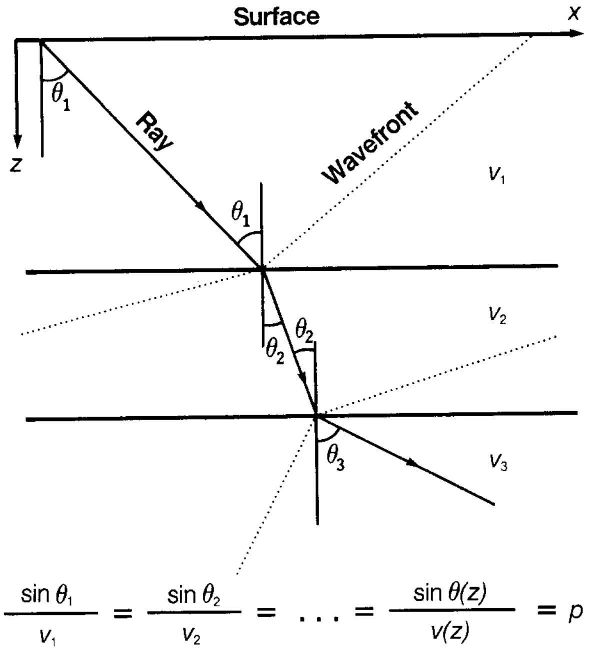


FIG. 7-7. Some raypaths for a given p value, corresponding to a single trace in the (p, τ) plane.



$$\frac{\sin\theta_1}{V_1} = \frac{\sin\theta_2}{V_2} = \ldots = \frac{\sin\theta(z)}{v(z)} = p$$

FIG. 7-6. If the ray parameter p is specified, then the ray can be traced in a horizontally layered earth model with a known velocity function $v(z)$.

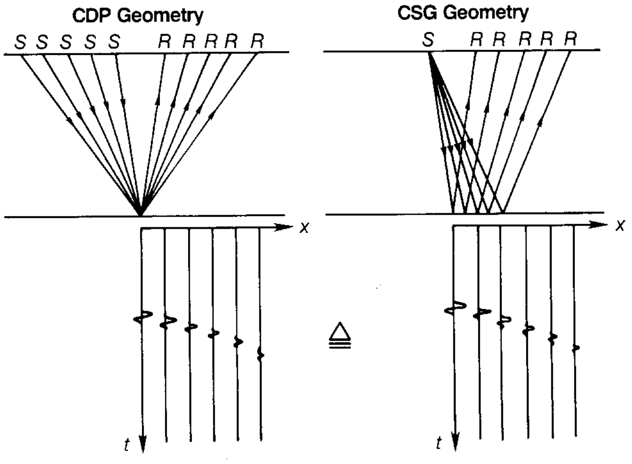


FIG. 7-8. Raypaths and traveltimes associated with CDP and CSG geometries.

before summation. This operator corrects for 3-D effects by converting a wave field that was obtained from a point source into a wave field that was obtained from a line source. As far as kinematics is concerned, the two types of slant stacking are equivalent. It is only when treating amplitudes that they differ (Treitel, personal communication).

A schematic description of the plane-wave mapping described by equations (7.4) and (7.5) is shown in Figure 7-9. We start by summing amplitudes in the offset domain along the horizontal path, $p = 0$. This line intersects the reflection hyperbola in the vicinity of apex A. Thus, point A maps onto point A′ on the (p, τ) plane. By tilting the line of summation, the hyperbola is intersected at location B, which maps onto B′. Note that a major contribution to summation along the slanted path comes in the area of the tangential point B. This zone of tangency is called the Fresnel zone. The Fresnel zone gets broader for higher velocities and deeper events. In fact, the summation in equation (7.5) can be confined to the Fresnel zone. The steepest necessary path of summation is along $p = 1/v$, which is the asymptote to the hyperbola. This path corresponds to rays that are 90 degrees to the vertical. The energy along the asymptote maps to C′ on the p-axis. By using the mapping previously described, the hyperbolic trajectories in the (x,t) domain are mapped to elliptical trajectories in the (p, τ) domain (Schultz and Claerbout, 1978; refer to Exercise 7.1). In reality, we never record a hyperbola with infinite extent nor a zero-offset trace. Therefore, the elliptical path in the slant-stack domain never is complete from A′ to C′.

A more complicated situation is shown in Figure 7-10. Subcritical reflections A and D (those with an angle of incidence smaller than the critical angle) map into the region of lower p values, while supercritical reflections C (wide-angle reflections) map into the region of higher p values. Ideally, a linear event in the offset domain, such as a refraction arrival B, maps to a point in the slant-stack domain. Conversely, a linear event in the slant-stack domain maps to a point in the offset domain (Exercise 7.2).

Figure 7-11 shows a field data example that contains predominantly water-bottom and peg-leg multiples. Besides the water-bottom reflection W, there are two distinct primaries, P_1 and P_2. Multiple reflections map along the elliptical trajectories that converge at $p = (1/1500)$ s/m, the inverse of the water velocity. A field data example containing linear events is shown in Figure 7-12. Note the strong amplitudes on the slant-stack gather that correspond to the guided waves observed in the offset data. In both field data examples, slant-stack gathers were constucted using only positive p values. Thus, the backscattered energy in the offset data in Figure 7-12 is not represented in the slant-stack gather.

The interrelations between various domains used in seismic data processing now are examined. Consider a band-limited dipping event in offset domain (t,x) as shown in Figure 7-13. The offset range is from 250 to 5000 m with a trace spacing of 50 m. This event is mapped along a distinct radial line in the f-k domain (ω,k_x). The slope of the radial line is related to the horizontal phase velocity by the relationship:

$$\omega/k_x = v/\sin\theta. \tag{7.6}$$

Substitute for $p = \sin\theta/v$ to find the relationship between the variables in the transform domain:

$$k_x = p\omega. \tag{7.7}$$

Figure 7-13 also shows mapping of the dipping event to the slant-stack domain. A 1-D Fourier transform of the slant-stack traces in the time direction gives the amplitude spectrum (p, ω), which also is shown in Figure 7-13. This plane describes the frequency dependence of horizontal phase velocity and is exploited in analyzing guided

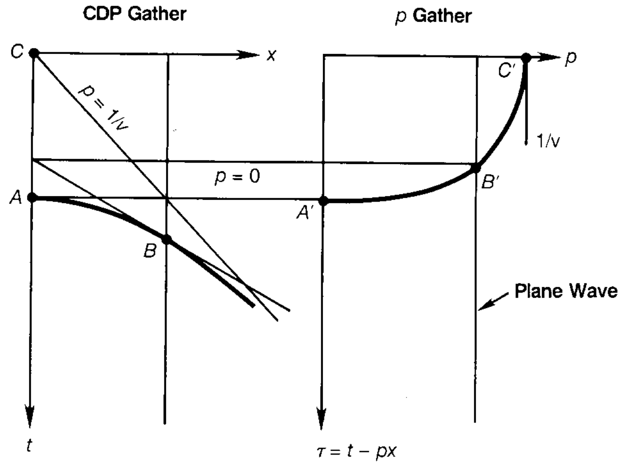


FIG. 7-9. A hyperbola on a CDP gather maps onto an ellipse on the p-gather.

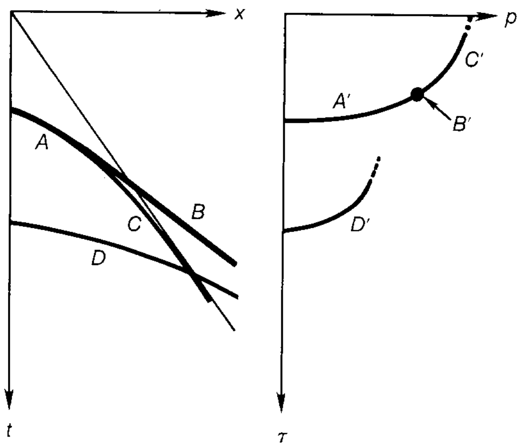


FIG. 7-10. Various arrivals on a CDP gather mapped onto the corresponding p-gather. Events A, B, C, and D map onto A′, B′, C′, and D′.

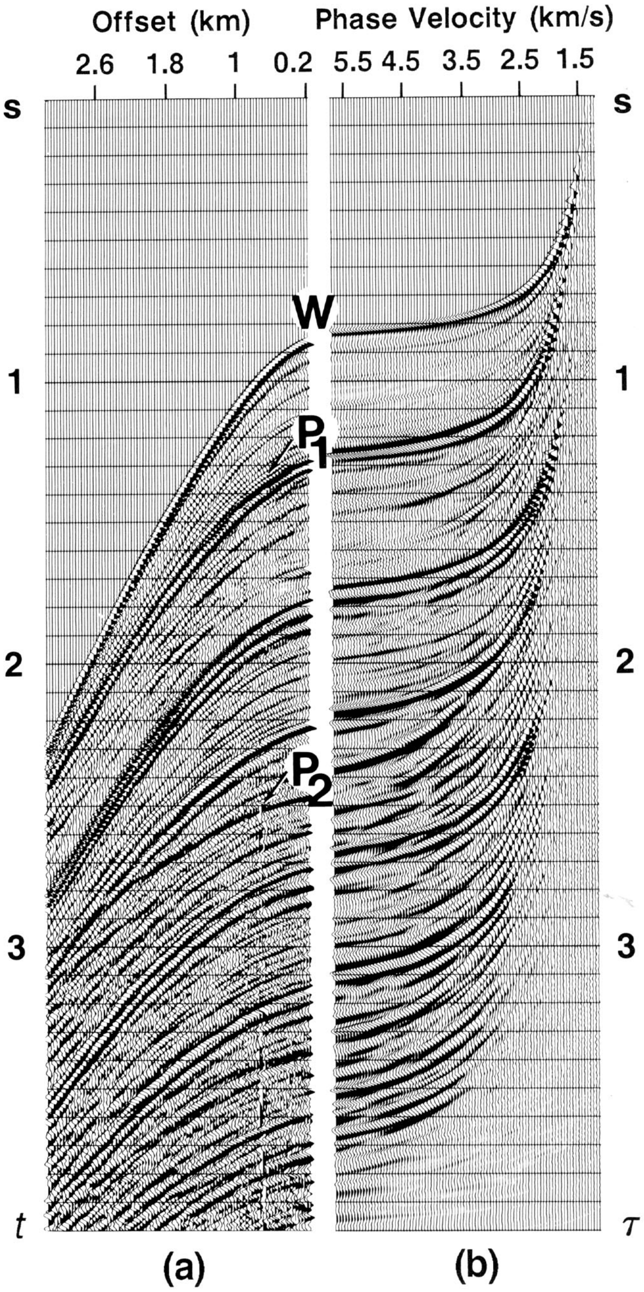


FIG. 7-11. (a) A shot gather containing strong multiples, (b) the corresponding p-gather. The horizontal axis in (b) is horizontal phase velocity ($1/p$). (Data courtesy Shell and Esso.)

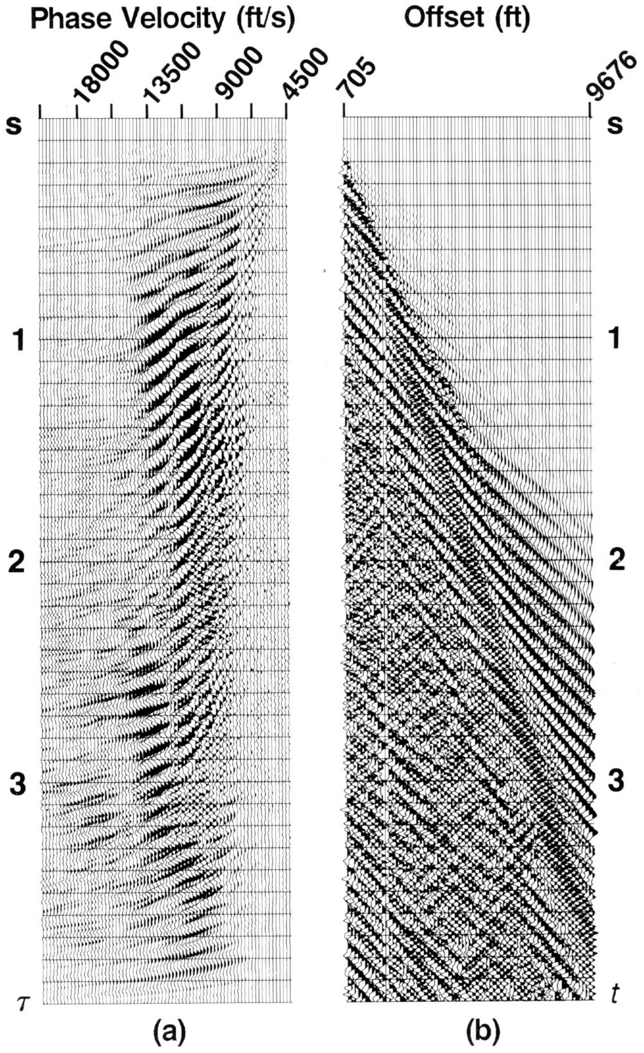


FIG. 7-12. (b) A shot gather and (a) its slant stack. The horizontal axis in (b) is horizontal phase velocity ($1/p$).

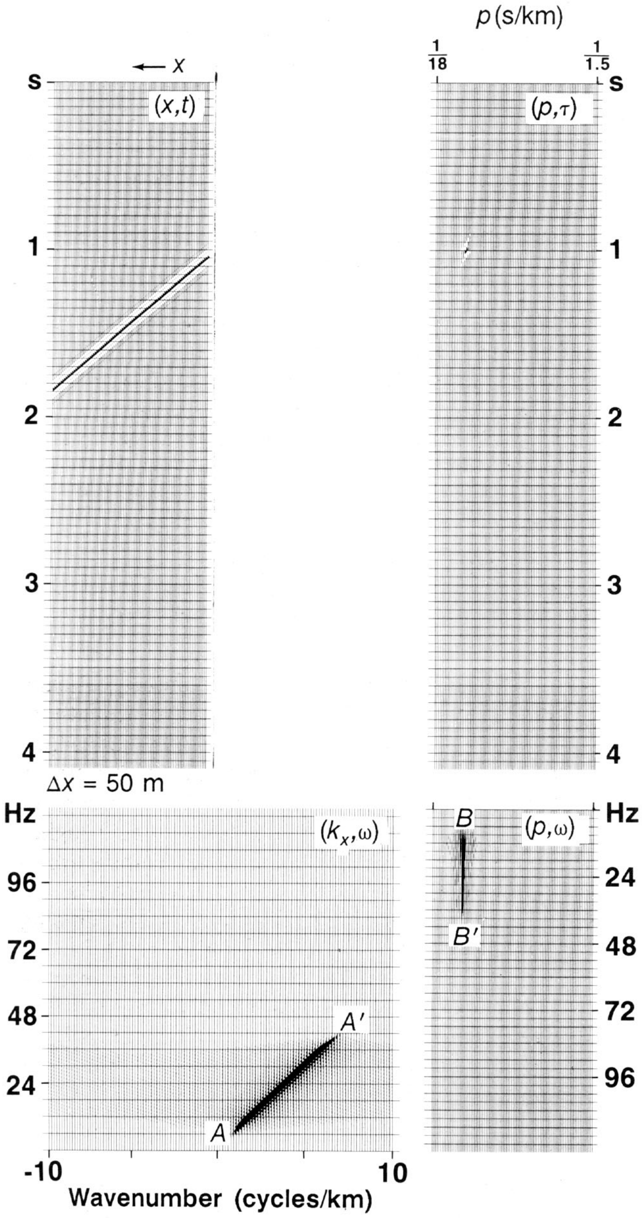


FIG. 7-13. A single dipping event in various domains.

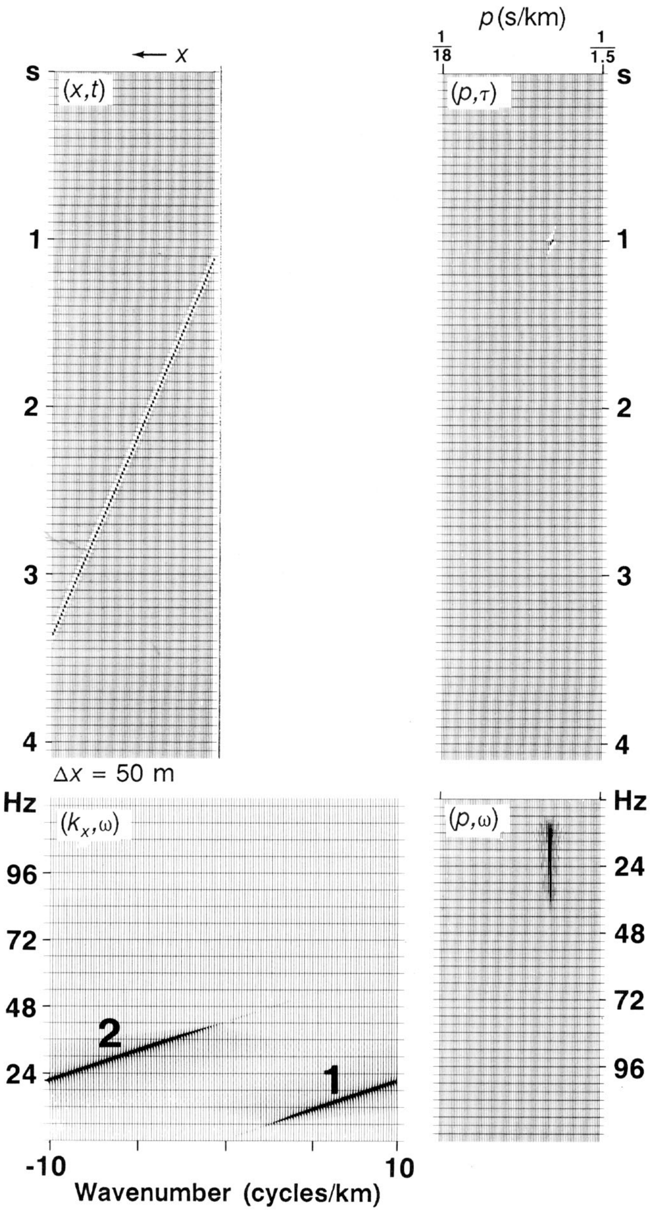


FIG. 7-14. A spatially aliased single dipping event in various domains.

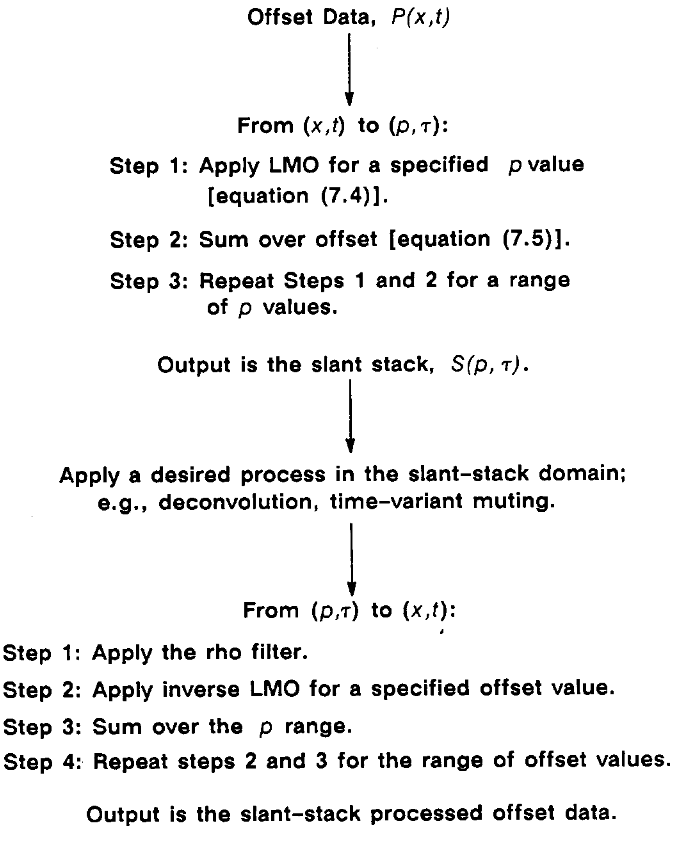
waves (Section 7.3). The energy along the radial direction AA′ on the (ω,k_x) plane is equivalent to that along the vertical direction BB′ on the (p, ω) plane.

Figure 7-14 shows a spatially aliased dip component. The wraparound observed in the (ω,k_x) plane results from inadequate spatial sampling of the event. Note that both the unaliased and aliased components (segments 1 and 2, respectively) map onto a single p trace. We expect the spatially aliased part to map onto a number of negative p traces. However, if this were the case, then the aliased frequency range (21 to 42 Hz) would be absent from the (p, ω) plane in which only the positive p values were included. (Figure 7-14 is for a single dip; reconstruction of a range of dips is discussed in Figure 7-17.)

Once a particular process is carried out in the slant-stack domain, inverse mapping is used to reconstruct the data in the offset domain. Thorson (1978) provides details of the reconstruction procedure. To restore amplitudes properly, rho filtering is applied before inverse mapping. This is accomplished by multiplying the amplitude spectrum of each slant-stack trace by the absolute value of the frequency. This is somewhat analogous to differentiation of the wave field before the summation that is involved in the integral formulation of migration [equation (4.5)].

Figure 7-15 is a flowchart of slant-stack processing.

Figure 7-16 shows a synthetic offset gather, corresponding slant stack, and reconstructed offset gather without any process applied, except the rho-filter. The x-to-p mapping is reversible (Thorson, 1978). The linear streaks (CT) on the slant-stack gather are caused by finite cable length. Tapering of the offset gather on both sides helps suppress these cable truncation effects.



Offset Data, $P(x,t)$

From (x,t) to (p,τ):

Step 1: Apply LMO for a specified p value [equation (7.4)].

Step 2: Sum over offset [equation (7.5)].

Step 3: Repeat Steps 1 and 2 for a range of p values.

Output is the slant stack, $S(p,\tau)$.

Apply a desired process in the slant-stack domain; e.g., deconvolution, time-variant muting.

From (p,τ) to (x,t):

Step 1: Apply the rho filter.

Step 2: Apply inverse LMO for a specified offset value.

Step 3: Sum over the p range.

Step 4: Repeat steps 2 and 3 for the range of offset values.

Output is the slant-stack processed offset data.

FIG. 7-15. Flowchart for slant-stack processing.

During reconstruction of the (x,t) gather, we do not have to use the same trace spacing that was used for the original (x,t) gather. Consider the synthetic gather in Figure 7-17a. The 2-D amplitude spectrum shows that frequencies above 48 Hz are spatially aliased (Figure 7-17b). This gather can be mapped to the slant-stack domain (Figure 7-17c) and reconstructed using a finer trace spacing (Figure 7-17d). The original trace spacing is 25 m; the reconstructed gather has a trace spacing of 12.5 m. The 2-D amplitude spectrum of the trace-interpolated gather shows that no frequencies are spatially aliased (Figure 7-17e). Nevertheless, note the missing high-frequency energy beyond 60 Hz. This energy mainly is along the steep direct arrival path in the input gather (Figure 7-17a) and is absent in the output gather (Figure 7-17d). We see that reconstruction can be successful even for spatially aliased data, provided dips do not have a wide range of variation.

7.2.1 Optimum Selection of Slant-Stack Parameters

A single dipping event in the (x,t)-domain theoretically maps onto a single p trace that represents the dip of that event (Figure 7-13). However, because of sampling along the p-axis and because only a finite number of p traces are spanned from a finite number of offset traces, an imperfect mapping results. When plotted with a higher gain, the slant-stack gather of Figure 7-13 seems surprisingly different (Figure 7-18b). The linear streaks are contributions from end points E and F of the dipping event in the (x,t) domain. More specifically, point E maps onto A and B when p is set to its minimum and maximum values, respectively. For any intermediate value of p, point E maps along AB. Similarly, the other end point F maps along CD. Linear streaks that result from end effects (cable truncation) are only one type of artifact encountered when constructing slant stacks. Another type of artifact is the high-frequency wave train that is especially apparent on traces with large p values. It occurs because the dipping event is sampled along steep slanted paths.

Several practical considerations affect the artifact level in slant stacks. A short cable length in the (x,t) domain enhances the end effects and, thus, causes poor reconstruction as demonstrated in Figure 7-18. Start with an offset gather that contains a single dipping event EF in panel (a). Panel (b) is the p-gather and panel (c) is the reconstructed offset gather from it. To emphasize the artifacts, the last two panels of the sets of three are displayed at a somewhat higher gain when compared to the original. Using two-thirds of the offset gather, panel (d), the p gather, and the reconstruction from it were obtained as shown in panels (e) and (f), respectively.

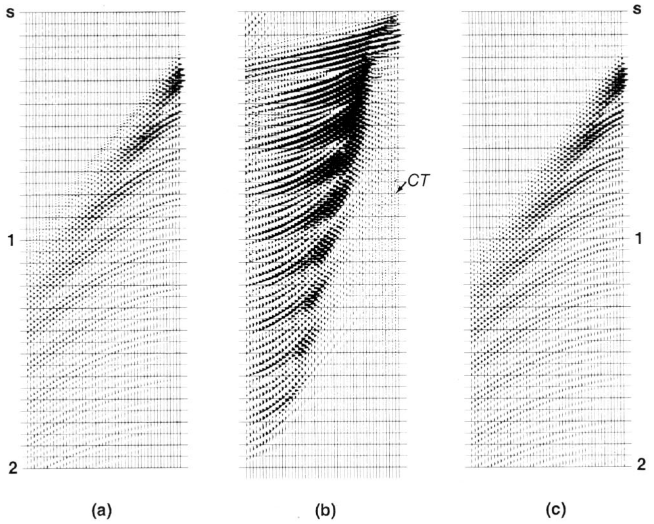


FIG. 7-16. Slant stacking is invertible. (a) An offset gather is mapped onto the (p, τ) domain (b), from which the original gather can be reconstructed (c).

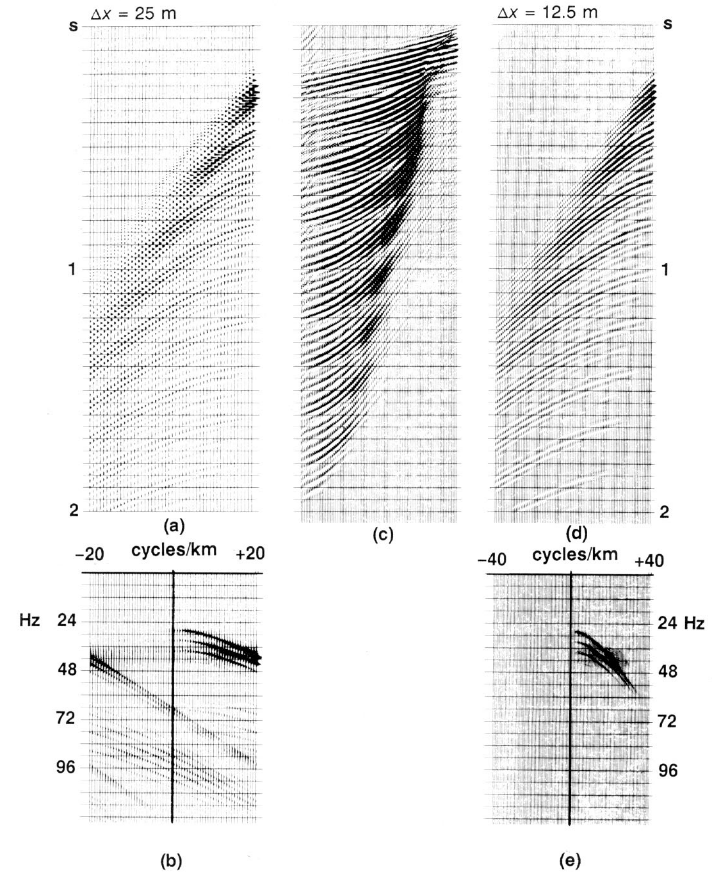


FIG. 7-17. Slant stack can be used for trace interpolation. (a) an offset gather is mapped onto the (p, τ) domain (c) and is reconstructed using a finer trace spacing (d). The corresponding f,k spectra show spatial aliasing in the original gather (b), which was eliminated after reconstruction (e).

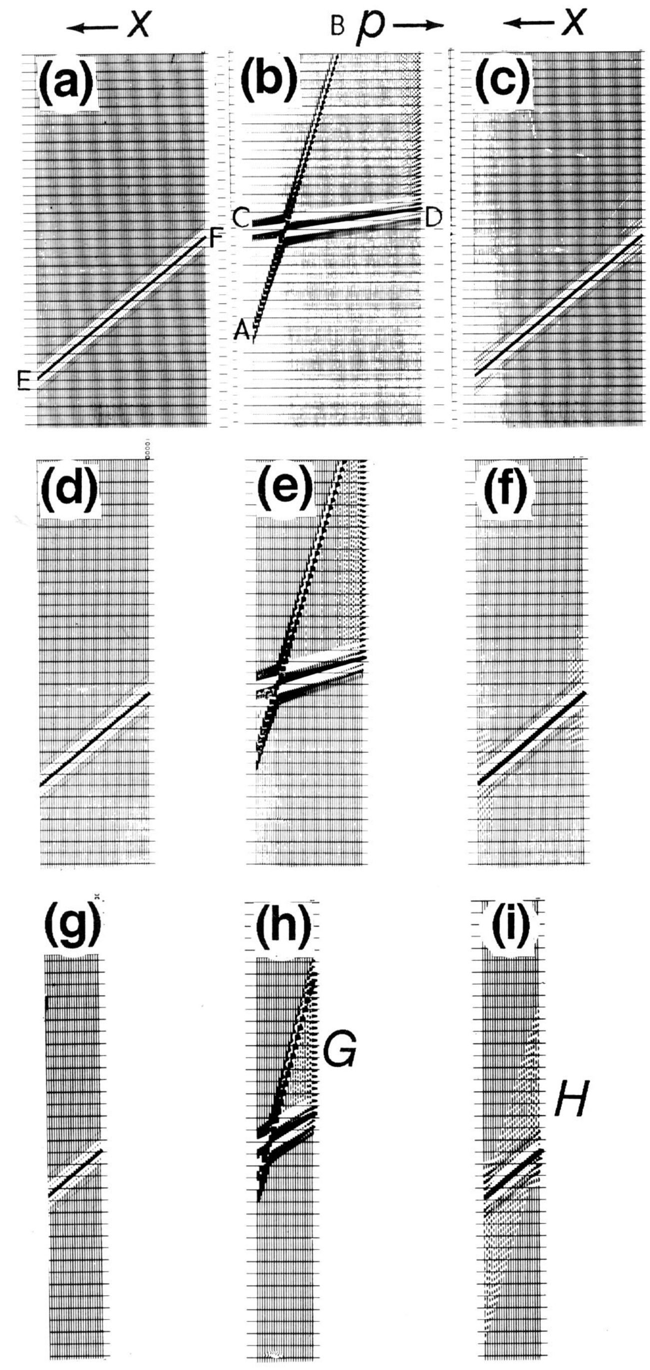


FIG. 7-18. Panels (a), (d), and (g) are the input offset gathers, which contain a single dipping event EF. Panels (b), (e), and (h) are the corresponding slant-stack gathers. Panels (c), (f), and (i) are the reconstructed offset gathers. The slant-stack and reconstructed gathers are displayed at a higher gain than the input gathers.

Finally, using only one-third of the original gather, panel (g), panels (h) and (i) are obtained. Note that short cables produce artifacts G and H on the slant-stack and reconstructed gathers. Accurate construction of slant-stack gathers usually requires cable configurations (i.e., cable length and number of channels) found only in more recently recorded data.

To study the sampling along the p-axis and the range of p values used in constructing a slant-stack gather, consider the synthetic gather in Figure 7-19, panel (a), which consists of hyperbolic events. These events are mapped along the ellipses in the slant-stack gather panel (b). The following values were chosen: the number of p traces (n_p) equal to the number of x traces (n_x); the minimum p value, $p_{\min} = 0$; and the maximum p value, $p_{\max}$ = the largest dip present in the data. Reconstruction using these parameters produced an accurate result [panel (c)]. There is some difference in the 2-D amplitude spectra of the original, panel (d), and the reconstructed gathers panel (e), because $p_{\min}$ was set to zero.

What happens when the p-axis is undersampled? Figure 7-19 shows the slant-stack gather, panel (f), and the reconstructed gather panel (g) that is obtained by setting $n_p = n_x/2$ and keeping p_{min} and p_{max} the same as in panel (b); thus, the p-increment is twice as large as in panel (b). The input gather is the same as it was in panel (a). Note that undersampling along the p-axis introduces some noise into the reconstructed gather [location A in panel (g)].

Consider the opposite situation of oversampling along the p-axis as in panel (h). Here, $n_p = 2n_x$ and the ($p_{\min}$, $p_{\max}$) range is the same as in panel (b). Note that oversampling in the p-axis does no harm, but gains nothing either [panel (i)]. Although not shown here, further experiments show that regardless of spread length, oversampling in the p-domain does not improve the quality of the reconstructed gather.

In practice, we may encounter an inappropriate choice of the ($p_{\min}$, $p_{\max}$) range, meaning that $p_{\max}$ corresponds to a larger dip than is present in the input gather [Figure 7-19, panel (j)]. Here, $n_p = n_x$, $p_{\min} = 0$, $p_{\max}$ is twice as large as the value chosen in panel (b), and the p-increment is the same as in panel (f). Thus, the right half of the p-gather contains no dip components that were present in the input data panel (a). Instead, the right half contains noise resulting from cable truncation and sampling along steep slanted paths with p values associated with dips not contained in the offset data. Creating nonexisting dip components in the p-domain produces some noise in reconstruction [location B in panel (k)]. In practice, suitable muting in the p-domain can eliminate the artifacts caused by spurious p traces [i.e., the right half of panel (j)].

Figure 7-20 shows panels that are equivalent to those in Figure 7-19, except that the input gather [panel (a)] has spatially aliased frequency components. It is clear that the artifacts observed in Figure 7-19 are more pronounced in the latter case. However, note that if the ($p_{\min}$, $p_{\max}$) range n_p and the p-increment are chosen properly [panel (b)], then reconstruction is quite accurate, even with spatially aliased data. The amplitude spectra of the original panel (d) and the reconstructed data panel (e) are essentially replicas of each other, except that the latter does not contain unaliased energy for $p<0$, which was not included in constructing panel (b).

This experimental study and other similar studies of the parameters involved in slant-stack processing lead to the following empirical statements:

1. Here, $n_p = n_x$ is a good, general rule (Figures 7-20b and 7-20c).
2. The ($p_{\min}$, $p_{\max}$) range should only span the dip components of interest in the data (Figure 7-20b). For example, for marine CMP data, $p_{\min} = 0$, $p_{\max} = (1/1500)$ s/m.
3. The p-increment then is $(p_{\max} - p_{\min})/n_x$. Sampling along the p-axis also can be done with equal increment in $(1/p)$, horizontal phase velocity (see Exercise 7.3).
4. The end effects generated by finite spread lengths are linear streaks on the slant-stack gather (see Figure 7-18b). The end effect is more pronounced with smaller cable length (Figure 7-18h) and spatially aliased data (Figure 7-20j).
5. Undersampling along the p-axis, where the p-increment is larger than recommended in item 3, causes noise buildup in reconstruction (Figure 7-19g). On the other hand, oversampling does no harm (Figure 7-19i). Finally, constructing slant-stack traces for which no p value exists in the offset data causes noise in the slant-stack and reconstructed gathers (Figures 7-19j and 7-19k).

7.3 ANALYSIS OF GUIDED WAVES

Marine data often are contaminated by guided waves that travel horizontally within the water layer or in the layers beneath the water layer. These waves exhibit characteristics that depend on water depth and on the geometry and material properties of the substrata. Modeling these pressure waves traveling within the water layer can lead to a better understanding of certain aspects of the field data and sometimes may even result in inferences about the strata below the water layer.

The well-known normal mode theory provides a way to laterally extrapolate acoustic and elastic waves (Pekeris, 1948; Press and Ewing, 1950). In this section, a normal mode procedure is applied to model shot profiles recorded over a water layer on top of a homogeneous elastic half space. Raypaths corresponding to multiple reflections, direct arrivals, refracted arrival, and its multiples are included in the normal mode theory.

The seismic waveguide effect of a surface layer is well known. Wave propagation in a surface layer can be described by using the normal mode theory (Pekeris, 1948). Pekeris' model consists of a liquid layer over an acoustic (liquid) half space. More general models, which consist of a liquid layer on top of an elastic half space,

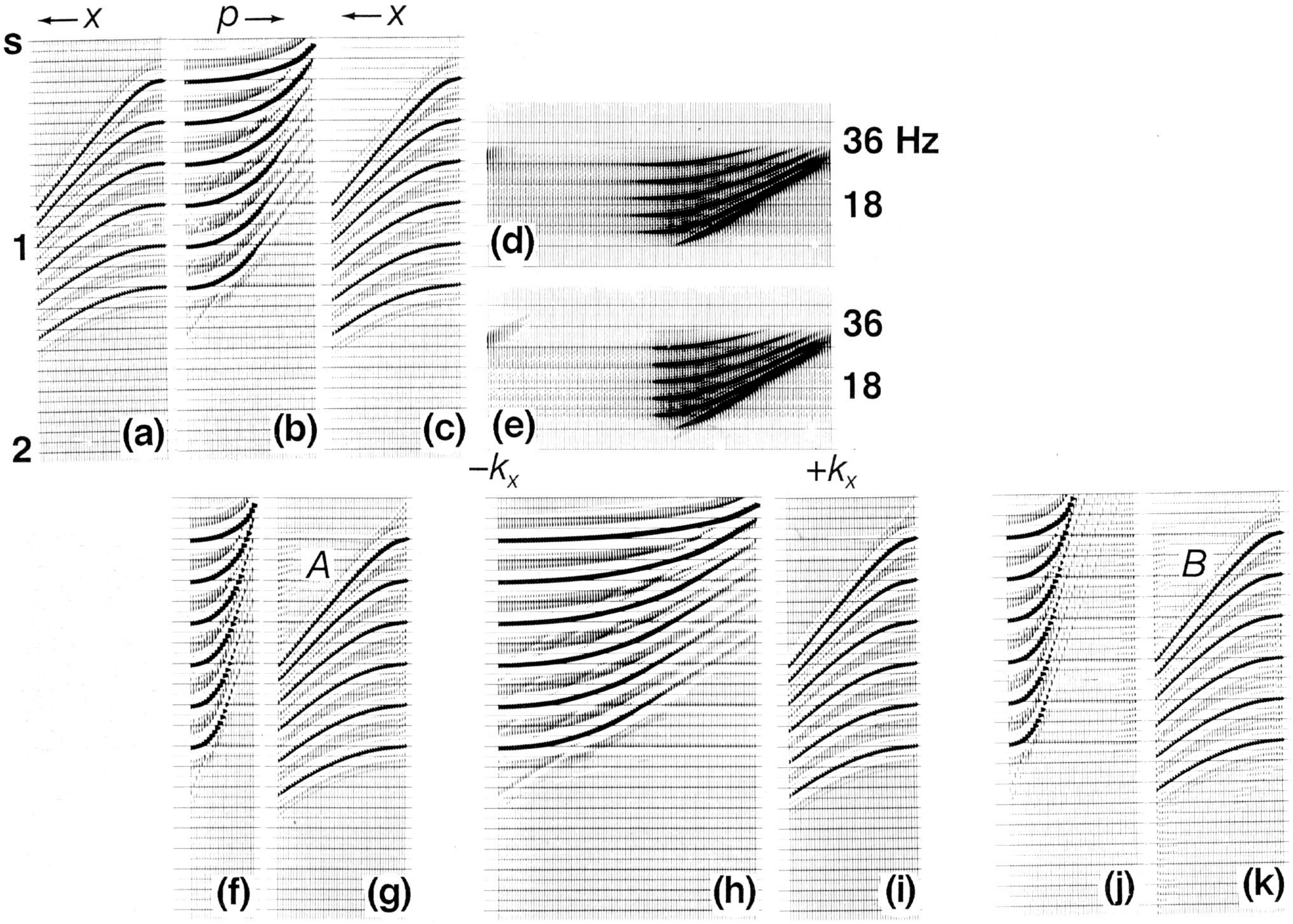


FIG. 7-19. (a) Input gather, (b) slant-stack gather, (c) reconstructed offset gather, (d) *f-k* spectrum of panel (a), (e) *f-k* spectrum of panel (c). Panels (f), (h), and (j) are the slant-stack gathers derived from the input gather (a) using different numbers of p-values and ranges, while panels (g), (i), and (k) are reconstructions from them. Input is the same for all [panel (a)]. See text for details.

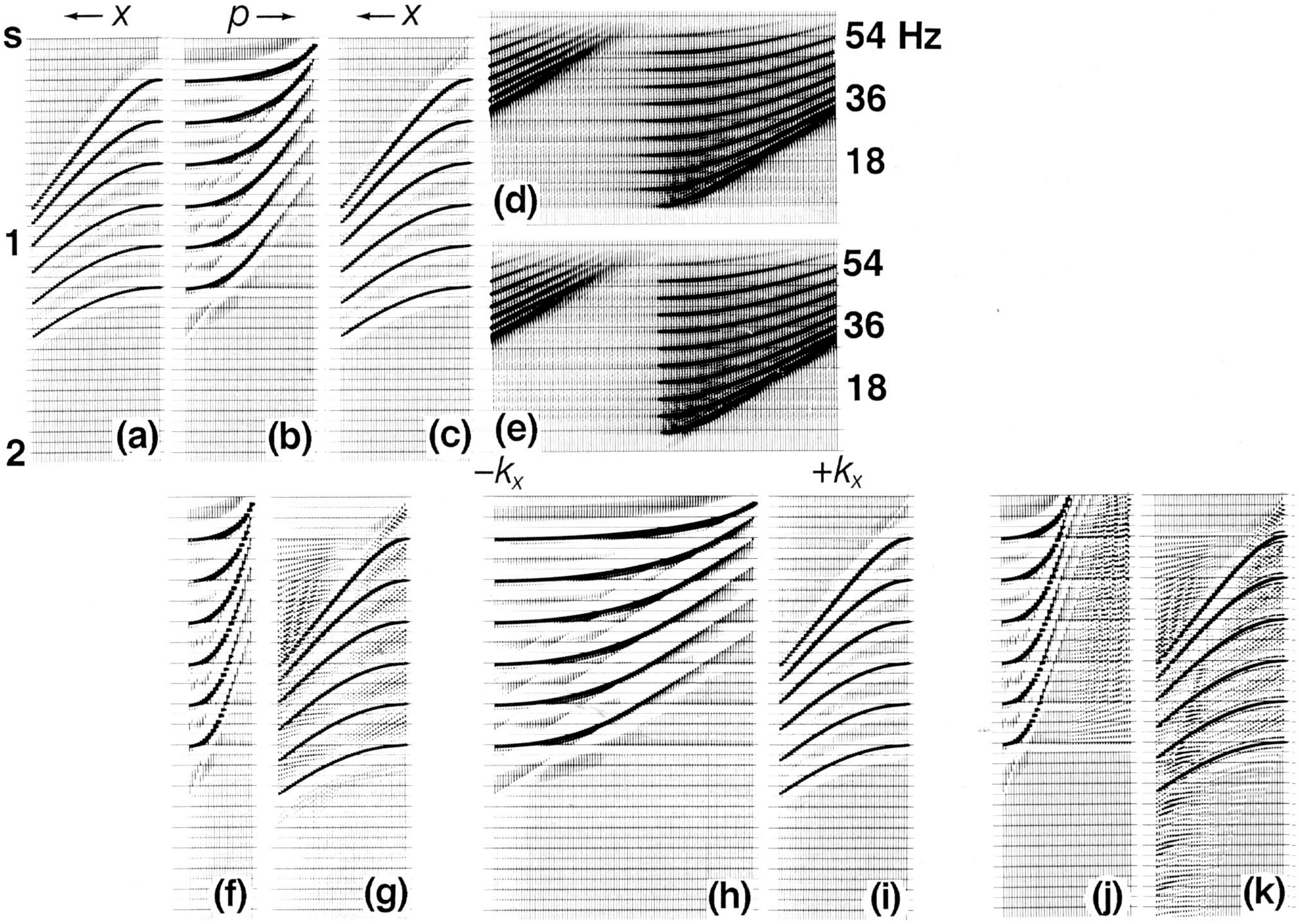


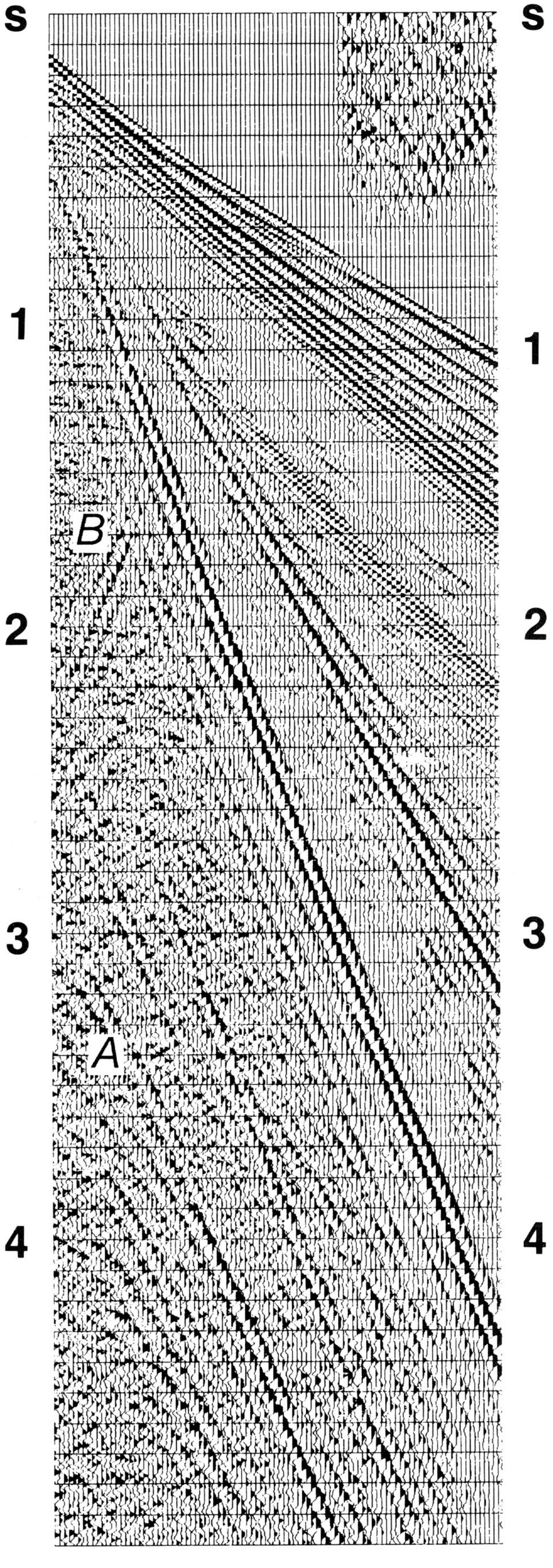
FIG. 7-20. The same sequence of panels as in Figure 7-19, except that the input gather contains spatially aliased frequency components. Note wraparound in the *f-k* spectrum (d).

were investigated by Press and Ewing (1950). The most complete summary of work in the field is contained in Ewing et al. (1957).

Guided waves are dispersive. This means each frequency component travels at a different speed; namely, the horizontal phase velocity. An excellent example of guided waves is seen on the field data in Figure 7-21 (between 1 and 3 s on the far offset). The first part of the wave package contains low frequencies. High frequencies ride along the direct arrival path, followed by moderate frequencies. Also seen on this record is the backscattered guided wave (Zone B) with reverse linear moveout, which indicates the presence of irregularities over the ocean bottom. These irregularities also cause hyperbolic arrivals (Zone A) that represent point scatterers.

The dispersive character of guided waves is most pronounced in shallow water environments (less than 100 m). Depending on various water-bottom conditions, such as a mud layer with variable thickness or a hard bottom, the character of these waves may vary from shot to shot (Figure 7-22). They also can cause linear noise on stacked data (Figure 1-86a) and are easily confused with the linear noise that is associated with side scatterers (Figure 1-88a).

McMechan and Yedlin (1981) proposed a way to obtain phase velocity information from field data. Their approach is based on a wave field transformation. The shot record first is transformed into the slant-stack domain. Fourier transforming (in time) each trace of the slant-stack gather then yields phase velocity as a function of frequency. This two-step process is demonstrated with the field data example in Figure 7-23a. The slant-stack gather is shown in Figure 7-23b; its 1-D amplitude spectrum is shown in Figure 7-23c. Note that the horizontal axis in the slant-stack domain is the ray parameter that is



←

FIG. 7-21. A shot gather containing predominantly guided waves. See text for a description of the labeled events.

the inverse of the horizontal phase velocity. Therefore, in Figure 7-23c, we see the variation of the horizontal phase velocity as a function of frequency. Each curve corresponds to a particular *normal mode* propagating in the water layer. The phase velocities of the normal mode components asymptotically approach that of the water velocity v_w at the high-frequency end of the spectrum.

The normal mode theory of Ewing et al. (1957) provides an analytic expression for phase velocity as a function of frequency (the so-called characteristic equation or dispersion relation) for a given surface-layer model. This expression is used to model guided waves. Consider the recording geometry and subsurface model that consists of a water layer on top of an elastic half space in Figure 7-24. The source is at a certain depth below the water surface, so the two raypaths, primary and ghost, must be considered. The characteristic equation for this geometry is given by Ewing et al. (1957):

$$\tan(k_x H r_1) = \frac{\rho_2 \beta_2^4 r_1}{\rho_1 c^4 r_2}\left[4 r_2 s_2 - (1 + s_2^2)^2\right] = B, \quad (7.8)$$

where k_x is the horizontal wavenumber [which is $p\omega$ via equation (7.7)], H is water depth, ρ_1 and ρ_2 are the water and substratum densities, respectively, β_2 is the S-wave velocity of the substratum, and c is the phase velocity of the guided waves in the water layer. The normalized variables are:

$$r_1 = (c^2/\alpha_1^2 - 1)^{1/2},$$

$$r_2 = (c^2/\alpha_2^2 - 1)^{1/2},$$

$$s_2 = (c^2/\beta_2^2 - 1)^{1/2},$$

where α_1 = P-wave velocity in the water layer and α_2 = P-wave velocity in the substratum. Because of the periodic nature of the tangent, equation (7.8) has a multiple-valued function on the left side and a single-valued function on the right side. To explicitly state the ambiguity, equation (7.8) can be rewritten as

$$k_x H r_1 + n\pi = \tan^{-1} B, \quad (7.9)$$

where the integer, $n = 0,1,2, \ldots$, defines the mode number. By careful examination of equations (7.8) and (7.9), note that the phase velocity is a function of frequency, hence, the guided waves are dispersive. Equation (7.9) yields real values of k_x for $\alpha_1 \leq c \leq \alpha_2$. (Assume that $\alpha_1 < \beta_2 < \alpha_2$). Table 7-1 is a summary of the phase velocity regions and types of rays associated with each region.

Only in the supercritical region, $\alpha_1 < c < \beta_2$, are waves totally trapped within the water layer. These waves often form the major contribution to normal-mode propagation at longer offsets as in the field data example in Figure 7-23a. (The supercritical region is to the right of line CC′.) In the subcritical region, energy leaks into the substratum (thus, the name *leaky modes*). The contribution of this region to energy arriving at longer offsets is relatively weak.

The recorded pressure wave field for the supercritical region at various receiver locations (Figure 7-24) is given by (Ewing et al., 1957):

$$P(x, z = h_r, t) = 4 \sum_n \int d\omega \, \{\omega^2 A(\omega) \sin(k_x r_1 h_s)$$

$$\cdot \sin(k_x r_1 h_r) \cdot \exp[i\omega(t - x/c)]\}. \quad (7.10)$$

where $A(\omega)$ is the amplitude spectrum of the source. Yilmaz (1981) modified this expression to account for the ghost effect. Note that reciprocity is satisfied here; i.e.,

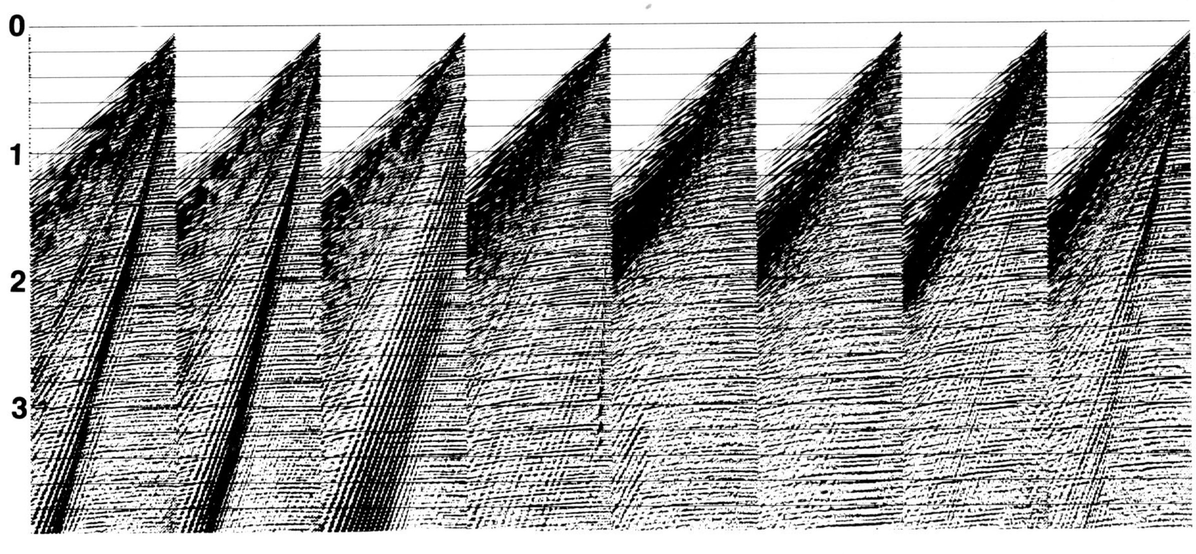


FIG. 7-22. Shot gathers (not sequential) containing guided waves with varying strength. The CMP stack obtained from these gathers is shown in Figure 1-86a. (Data courtesy Deminex Petroleum Company.)

Table 7-1. Regions of phase velocity and associated ray types.

Phase Velocity	Ray Type
$\alpha_1 < c < \alpha_2$	Supercritical, totally reflected P-waves; i.e., wide angle reflections.
$\beta_2 < c < \alpha_2$	Supercritical, but only partially reflected P-waves.
$c > \alpha_2$	Subcritical P-waves.
$c = \beta_2$	Critically refracted S-wave.
$c = \alpha_2$	Critically refracted P-wave.

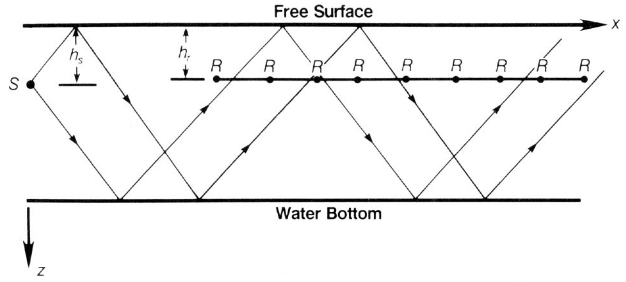


FIG. 7-24. The geometry for normal-mode modeling of the guided waves illustrated in Figure 7-25. Here, S = source, R = receivers, h_s = source depth, and h_r = receiver depth.

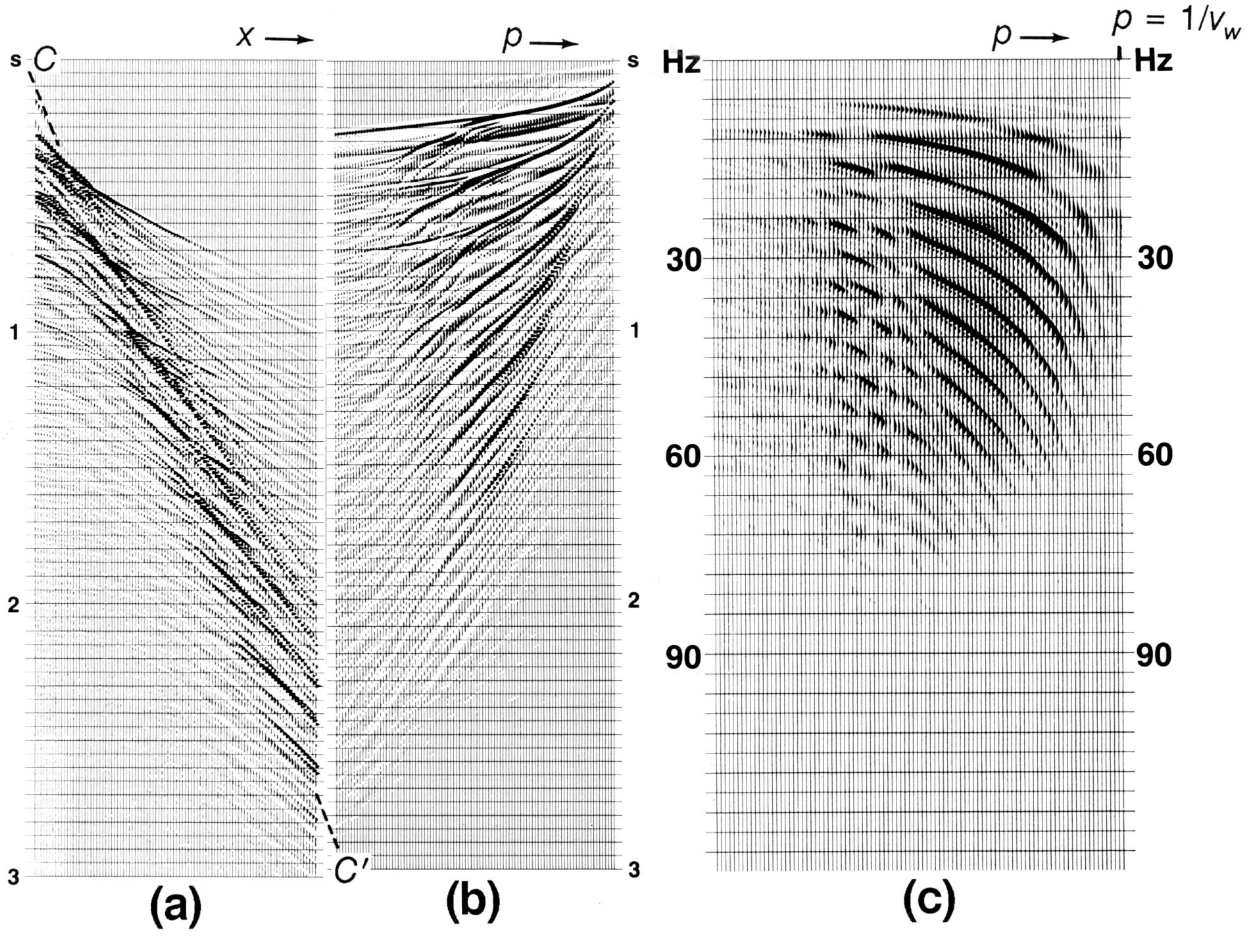


FIG. 7-23. (a) A shot gather containing the strong reflected and refracted multiples associated with hard water-bottom conditions. Here, CC′ = critical-angle energy. (b) The slant-stack gather derived from this shot gather. (c) The (p, ω) gather derived from the (p, τ) plane in panel (b). The inverse of p is the horizontal phase velocity. This figure demonstrates the dispersive nature of guided waves; that is, phase velocity is a function of frequency for all propagating normal mode components. These modes are represented by the curved trajectories on panel (c).

the product of the sinusoidal factors that modulate the source spectrum A(ω) is unchanged if h_r and h_s are interchanged.

Figure 7-25 shows normal-mode propagation computed from equation (7.10) (including the ghost effect) for a range of water depths. The model parameters are $\alpha_1 =$ 1500 m/s, $\beta_2 = 2\alpha_1$, $\alpha_2 = 1.6\ \beta_2$, and $\rho_2/\rho_1 = 2.2$. All the experimental results represent impulse responses of guided waves; i.e., $A(\omega) = 1$ in equation (7.10). Guided waves manifest themselves with a complex interfering wave pattern at shallow water, then gradually separate into simple water-bottom multiples at increasing water depths. The dispersive character of the guided waves is prominent, especially for shallow water depths. In Figure 7-25, the guided waves are simulated in the supercritical region, where RP is the refracted arrival, RM is its multiple, and M1, M2, and M3 are the water-bottom multiples. The phase velocity curves in Figure 7-26 verify the existence of a number of propagating modes for each case. The elastic substratum, which is equivalent to the hard water-bottom case, supports the early refraction energy RP, its multiples RM, and the reflected water-bottom multiples M1, M2, and M3. The acoustic substratum ($\beta_2 = 0$, equivalent to the soft water-bottom case), yields only the reflected water-bottom multiples. Acoustic behavior of the substratum implies that no P-to-S conversion occurs.

7.4 TIME-VARIANT DIP FILTERING

The slant-stack domain is convenient for implementing dip filtering. To illustrate this, consider the problem of suppressing the strong ground roll in the field record in Figure 7-27a, which was obtained from a walk-away noise test. For simplicity, ignore the backscattered Rayleigh waves, since removing them would mean computing

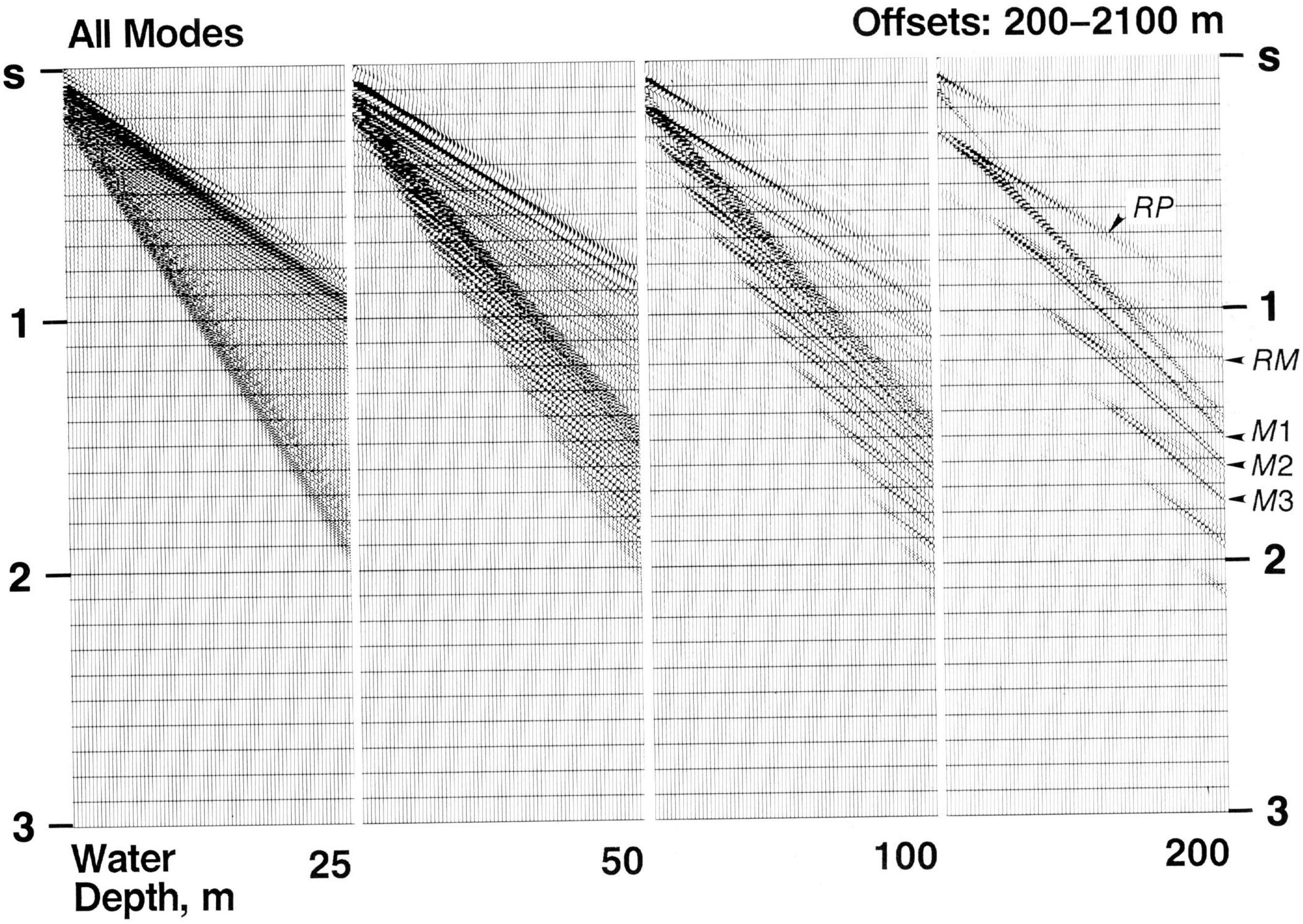


FIG. 7-25. Superposition of all modes in water layers of different thicknesses. The depth model is shown in Figure 7-24. Ghost effects are included in the modeling. (Refer to the text for a description of the labeled events.)

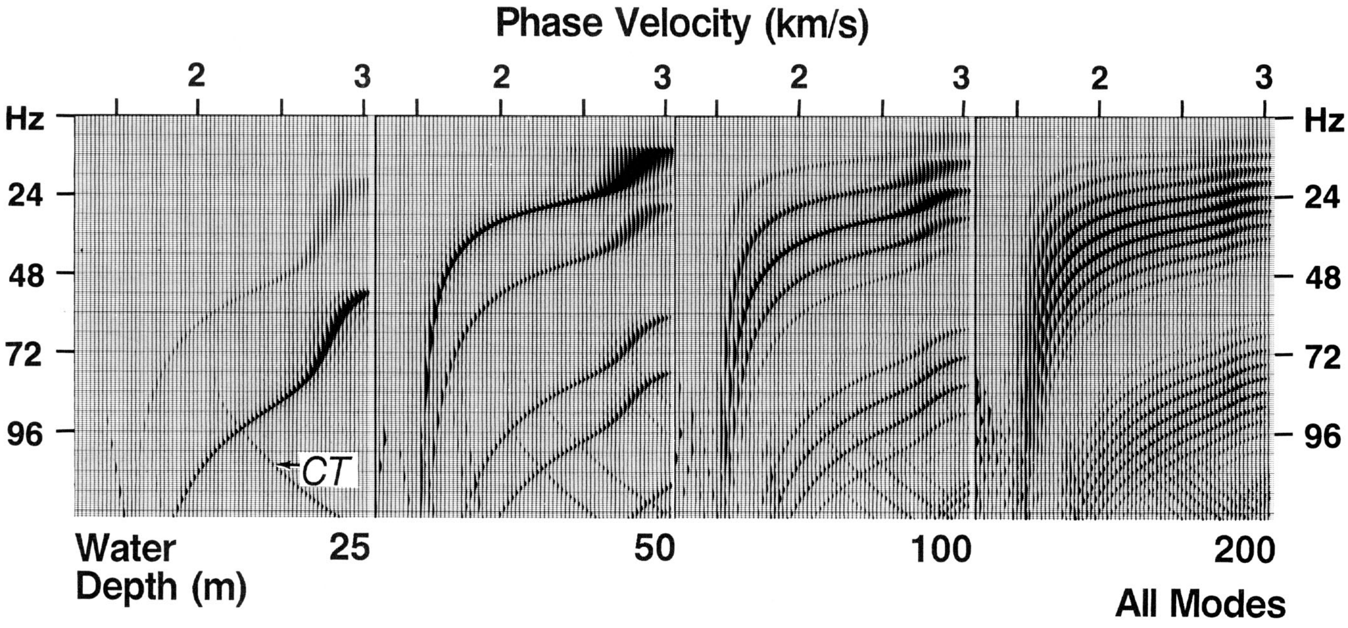


FIG. 7-26. Phase velocity as a function of frequency for the cases shown in Figure 7-25. Here, *CT* = cable truncation artifact.

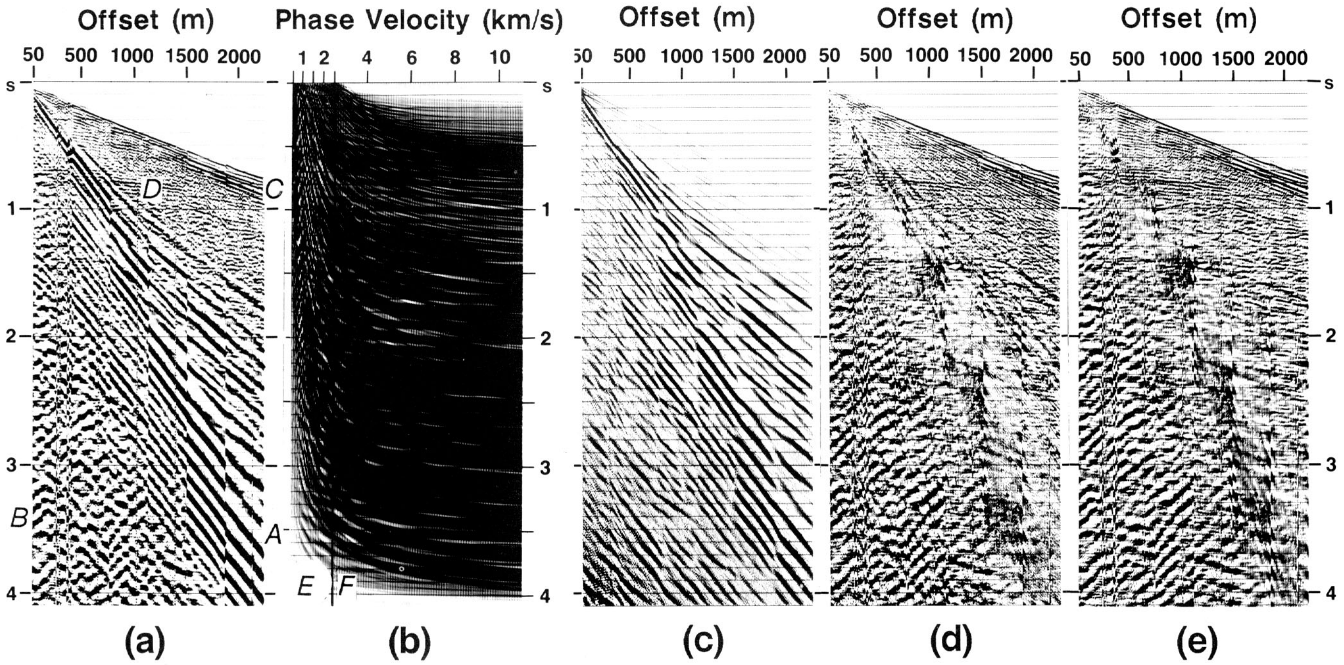


FIG. 7-27. (a) A field data set with strong ground roll energy *A*, its backscattered component *B*, guided waves *C*, and a strong reflection *D*; (b) *p*-gather obtained from this field data set; (c) reconstruction of the field record using the portion to the left of the solid vertical line in (b) (zone E); (d) dip-filtered data obtained by subtracting the gather in (c) from the original data in (a); (e) the original data set (a) after *f-k* dip filtering. (Data courtesy Turkish Petroleum Corporation.)

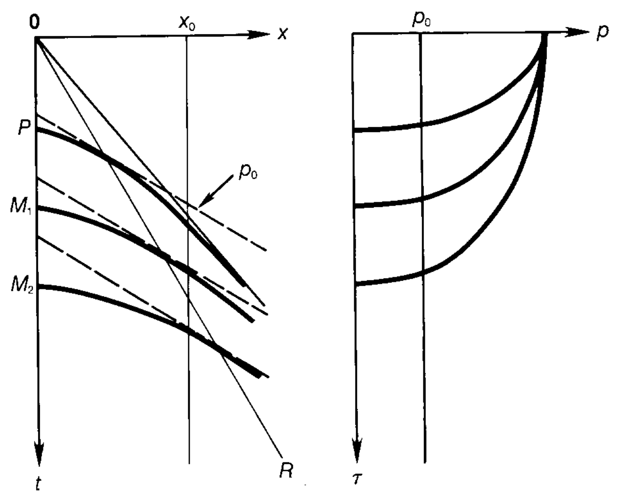


FIG. 7-28. The periodicity of multiples along radial trace OR and the p traces.

negative p traces. Figure 7-27b shows the slant-stack gather obtained from this field data set. Phase velocity values from 500 m/s to well over 10 km/s are spanned. The ground roll generally has very low phase velocity and is mapped to the left of the solid line at 2500 m/s (zone E).

Suppose that the slant-stack gather traces that contain the ground roll (zone E in Figure 7-27b) are used to reconstruct (inverse slant stack) traces at the original offsets. The reconstructed gather shown in Figure 7-27c contains only the dips that we want to remove from the original wave field. If this reconstructed gather is subtracted from the original (Figure 7-27a), the result is the dip-filtered shot record in Figure 7-27d. An alternate way to compute the dip-filtered shot record is to use the slant-stack traces that are only within zone F for reconstruction. With either approach, the amplitudes at the boundary between the pass and reject zones, the solid line at 2500 m/s in Figure 7-27b, must be tapered to reduce artifacts in the reconstructed offset gather.

Dip filtering in the slant-stack domain should be nearly equivalent to the well-known *f-k* dip filtering process described in Section 1.6.2. Figure 1-81b shows the 2-D amplitude spectrum of the original field record of Figure 1-81a. (This is the same data set as in Figure 7-27a). The reject zone is defined by the fan in Figure 1-81c. This zone is equivalent to the part of the p-gather that is to the left of the vertical line, zone E in Figure 7-27b.

When compared with the slant-stack output (Figure 7-27d), the result of *f-k* dip filtering of the field data set in Figure 7-27e suggests basically no difference in performance. However, with the slant-stack approach, dip filtering can be applied in a naturally time-variant way. This means that the boundary between the pass and reject zones need not be vertical (Figure 7-27b).

Also, with the slant-stack technique, we can work with data that are unequally spaced along the offset axis. This is not the case for the *f-k* method of dip filtering, since the fast Fourier transform requires data with equal trace spacing. On occasion, dip filtering also is incorporated into multiple suppression in the slant-stack domain to further eliminate multiples.

7.5 MULTIPLE SUPPRESSION

Multiple suppression techniques are based on one of the following characteristics of multiples:

1. The moveout difference between primaries and multiples (velocity discrimination).
2. The dip difference between primaries and multiples on the CMP stack.
3. The difference in frequency content between primaries and multiples.
4. The periodicity of multiples.

Techniques based on velocity discrimination are described in Section 8.2. The slant-stack multiple suppression technique, which is based on predictive criterion (characteristic 4) is discussed here. Alam and Austin (1981b) and Treitel et al. (1982) investigated application of predictive deconvolution in the slant-stack domain for multiple suppression. The application of predictive deconvolution to multiple suppression (in the case of vertical incidence and zero offset) was discussed in Section 2.7.5. However, multiples are not periodic in time for a given nonzero offset. Figure 7-28 shows a shot gather with primary P (water-bottom reflection) and its multiples M_1, M_2 with the corresponding slant-stack gather. The time separations between the multiple arrivals at a particular offset x_0 are only equal if $x = 0$.

Taner (1980) first recognized that the time separations between the arrivals are equal along a radial direction OR. A trace can be constructed by pulling out the samples along one of these radial directions. This trace, along which the angle of propagation is constant, is called a radial trace (Ottolini, 1982). A radial trace in a layered medium is called a Snell trace (Claerbout, 1985). In a layered medium, the Snell trace would not follow a straight path as in Figure 7-28, since its angle of propagation changes at layer boundaries according to Snell's law (Figure 7-6).

Taner (1980) applied predictive deconvolution along radial traces to successfully eliminate long-period multiples. Note that the magnitude of the time separations between multiples is different from one radial trace to another (Figure 7-28). However, the time separations are equal along each of the slanted paths of summation. Therefore, a predictive deconvolution operator can be designed from the autocorrelogram of each p trace (p_0) and applied to suppress multiples. This is shown in Figure 7-29. Here we see a shot gather (a) and its autocorrelogram (b). The data contain nothing but the water-bottom reflection and its multiples. Note that the repetitive nature of the multiples is not apparent on the autocorrelogram. Therefore, predictive deconvolution should not be expected to do well

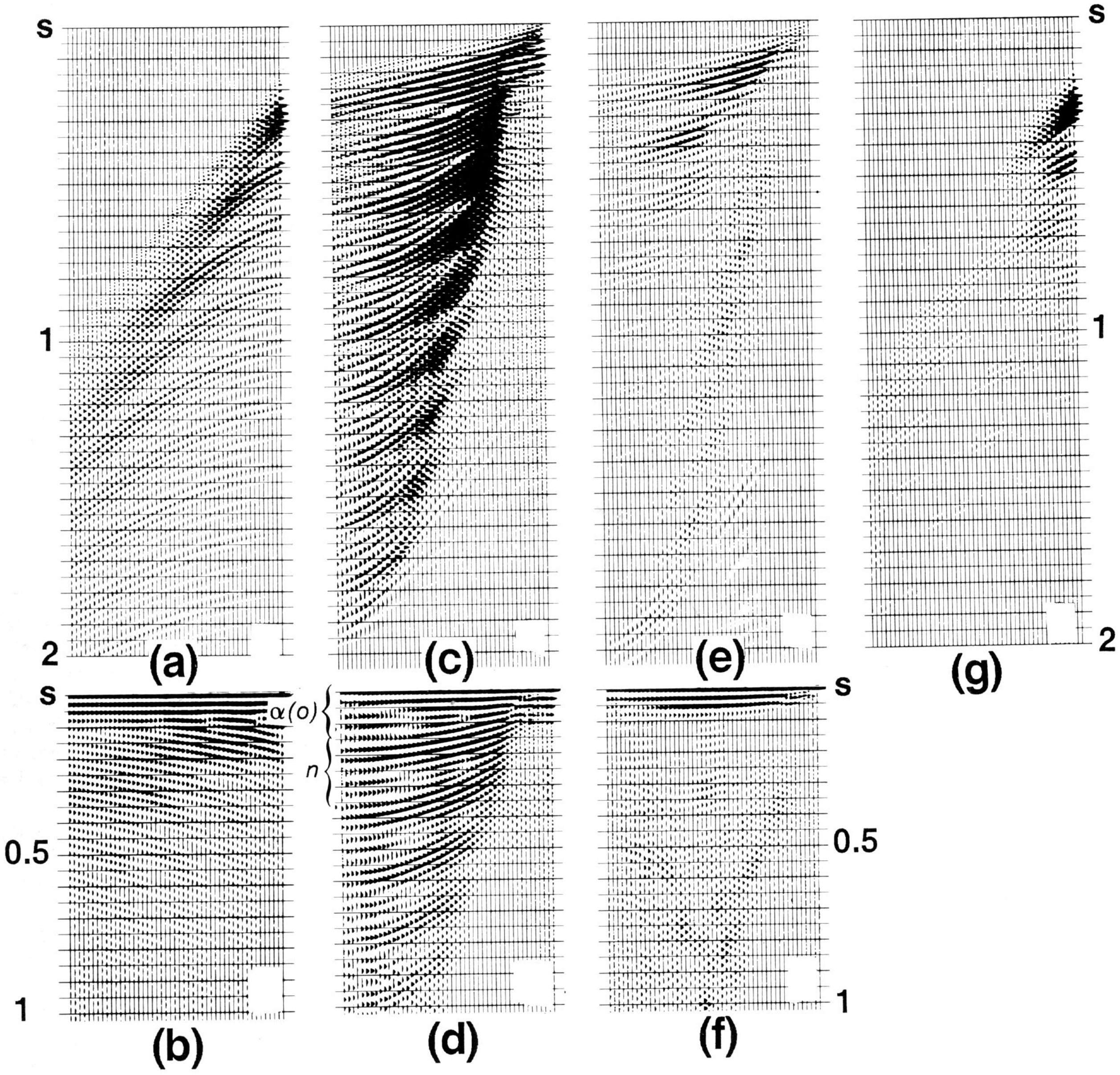


FIG. 7-29. Multiple suppression in the slant-stack domain. (a) A shot gather; (b) its autocorrelogram; (c) the slant-stack gather; (d) the autocorrelogram of (c); (e) the slant-stack gather after predictive deconvolution, where operator length = 240 ms and prediction lag (at $p = 0$) = 120 ms; (f) the autocorrelogram of (e); (g) reconstruction of the shot gather from (e).

in suppressing these multiples when applied to the shot gather.

This shot gather now is examined in the slant-stack domain. Figures 7-29c and 7-29e show the slant-stack gather before and after predictive deconvolution was applied. Figure 7-29g shows reconstruction of the shot gather from Figure 7-29e. Autocorrelograms before and after deconvolution in the slant-stack domain are shown beneath the respective panels. Unlike the autocorrelogram of the shot gather in Figure 7-29b, the autocorrelogram of the slant-stack gather well illustrates the periodic nature of the multiples in the data (Figure 7-29d). Note that the periodicity of multiples changes from one p trace to the next. The largest period occurs along the trace that corresponds to the minimum p value. The autocorrelogram after predictive deconvolution shows that the energy in the lags less than the prediction lag is retained, while the multiple energy is suppressed (Figure 7-29f).

Prediction lag α and operator length n must be specified by examining the autocorrelogram of the slant-stack gather. These two parameters are specified for the trace corresponding to the lowest p value, as indicated in Figure 7-29d. Operator length is kept constant, while prediction lag is adjusted based on the p value across the gather (Alam and Austin, 1981b):

$$\alpha(p) = \alpha(0)(1 - p^2 v_w^2)^{1/2}, \qquad (7.11)$$

where α (0) = prediction lag at $p = 0$ and v_w is multiple velocity (typically water velocity). At higher p values, the prediction lag decreases. The reconstructed offset gather (Figure 7-29g) should be compared with the input gather (Figure 7-29a). Note that the output contains the water-bottom reflection (the only one present in the input data) and a residual of the first multiple.

Multiple suppression in the slant-stack domain is demonstrated further by the model data in Figure 7-30a. These data are the output of an attempt at normal-mode modeling of the shot gather in Figure 7-23a. Several arrivals are identified: C is the direct arrival; A is the refracted arrival associated with the hard water bottom; B is the water-bottom reflection; M1, M2, and M3 are the refracted multiples; and m1, m2, and m3 are the reflected multiples. [D is an artifact of the normal-mode modeling technique (Section 7.3).]

Figure 7-30c is the slant-stack gather of the model data. This gather should be compared with the slant-stack gather of the field data in Figure 7-23b. Refraction A and its multiples M1, M2, and M3 map to points in the slant-stack domain. Figure 7-30d is the autocorrelogram of the p-gather. Unlike the autocorrelogram of the offset data (Figure 7-30b), note that it shows the periodic nature of the multiples in the data. After applying predictive deconvolution, the slant-stack gather in Figure 7-30e results. Only the refracted arrival A and the water-bottom reflection B remain. The nearly linear streaks, which also are present in the unprocessed slant-stack gather (Figure 7-30c), are artifacts caused by the finite cable length. The autocorrelogram after deconvolution is free from the multiple energy (Figure 7-30f). The prediction lag and operator length for the minimum p value are indicated in Figure 7-30d. Adjustment for the prediction lag was made across the gather using equation (7.11). Finally, reconstruction of the shot gather is shown in Figure 7-30g. When compared with Figure 7-30a, note that both the refracted and reflected primaries are retained, while their associated multiples are largely suppressed.

The performance of slant-stack multiple suppression on field data now is examined. Figure 7-31a shows a shot gather that contains a strong water-bottom reflection A, two distinct primaries B and C, the water-bottom multiples D and E, and the peg-leg F, which is associated with primary event B. The slant-stack gathers before and after predictive deconvolution are shown in Figures 7-31b and 7-31d with their autocorrelograms (Figures 7-31c and 7-31e). Note that multiples are significantly attenuated in the reconstructed gather (Figure 7-31f).

Choice of the prediction lag and operator length is tricky for this particular data set. From the autocorrelogram in Figure 7-31c, note the energy G, which is due to correlation of two primaries, A and B in Figure 7-31a. Energy H is due to correlation of the water-bottom multiples. Prediction lag is chosen to bypass the primary energy G. Operator length is chosen to include the multiple energy H. Note that multiple energy is significantly suppressed in Figure 7-31e.

Since slant stack is a plane-wave decomposition, and since plane waves do not have spherical divergence, input to the slant stack must not be compensated for by geometric spreading. Preserving correct amplitude relationships is essential for the effectiveness of slant-stack multiple suppression. The geometric spreading correction is applied to offset data by using a primary velocity function. This enhances the multiples in the data and destroys the amplitude relationship between them. Predictive deconvolution in the offset domain then would not suppress these multiples effectively.

After multiple suppression in the slant-stack domain, reconstruction of the offset data is performed, the geometric spreading correction is applied, and processing is continued. Figures 7-32a and 7-32b show the shot gathers in Figures 7-31f and 7-31a after the geometric spreading correction. Note that strong multiples are significantly suppressed after slant-stack processing. Again, the autocorrelogram (Figure 7-32c) of the shot gather without slant-stack processing does not have a clear indication of the presence of strong multiples (compare with Figure 7-31c).

Figure 7-33 shows that when multiples are in the form of short-period reverberations, the autocorrelogram of the offset data seems to adequately represent the periodic nature. Hence, predictive deconvolution of the offset data often can remove reverberations. On the other hand, long-period multiples are poorly represented by the autocorrelogram of the offset data as in Figure 7-32c and are better defined in the slant-stack domain (Figure 7-31c).

Figure 7-34 shows the stacked section that corresponds to the field data in Figure 7-32b. The slant-stack processed section, which corresponds to the data in Figure 7-32a, is shown in Figure 7-35. The slant-stack processed

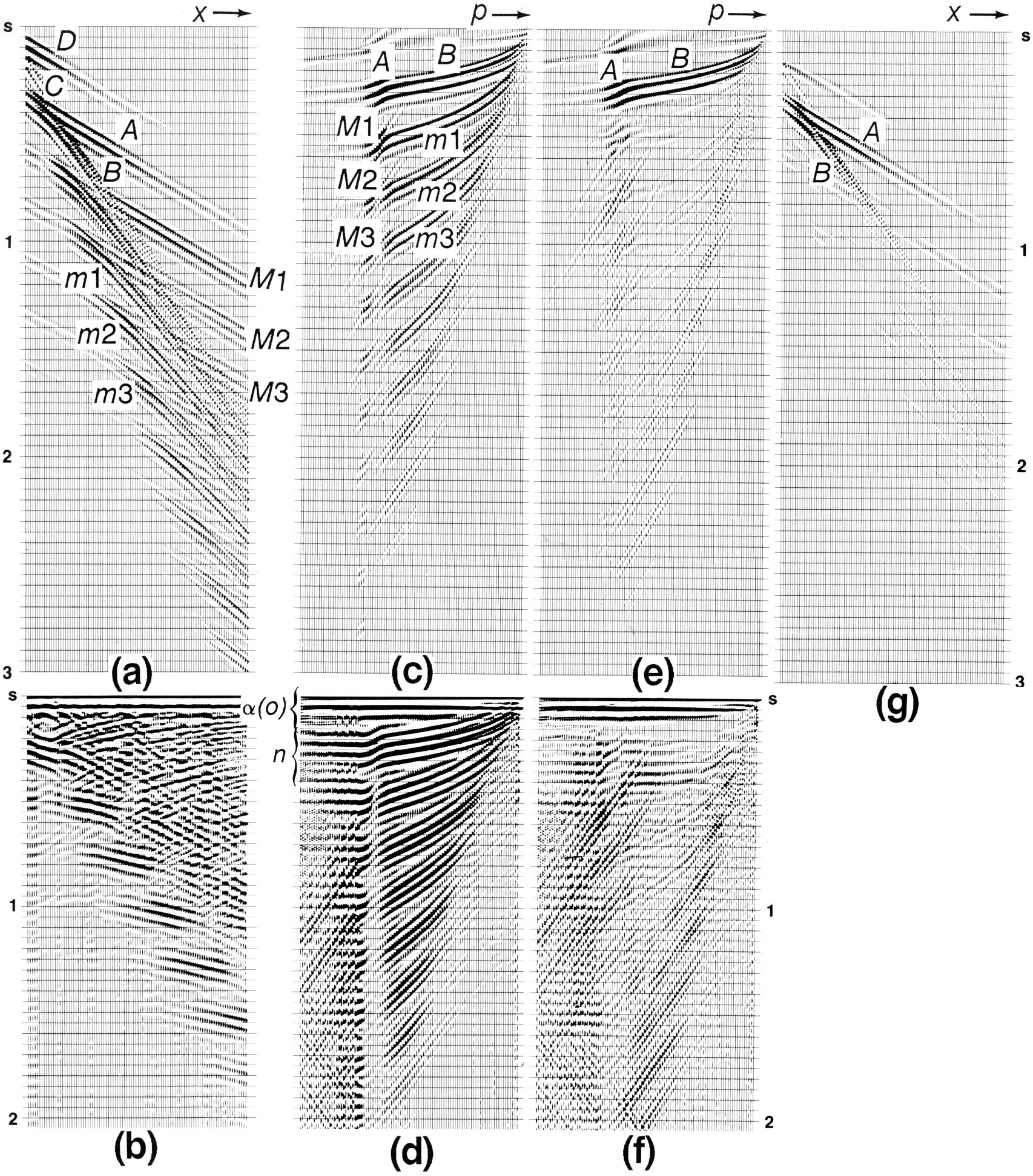


FIG. 7-30. (a) Simulation of the shot gather shown in Figure 7-23a by normal-mode modeling; (b) autocorrelogram of this synthetic gather; (c) slant stack of the synthetic gather; (d) autocorrelogram of the slant-stack gather; (e) the slant-stack gather in (c) after predictive deconvolution; (f) the autocorrelogram of (e); (g) reconstruction of the synthetic gather from the slant-stack gather in (e). (Refer to the text for a description of the labeled events.)

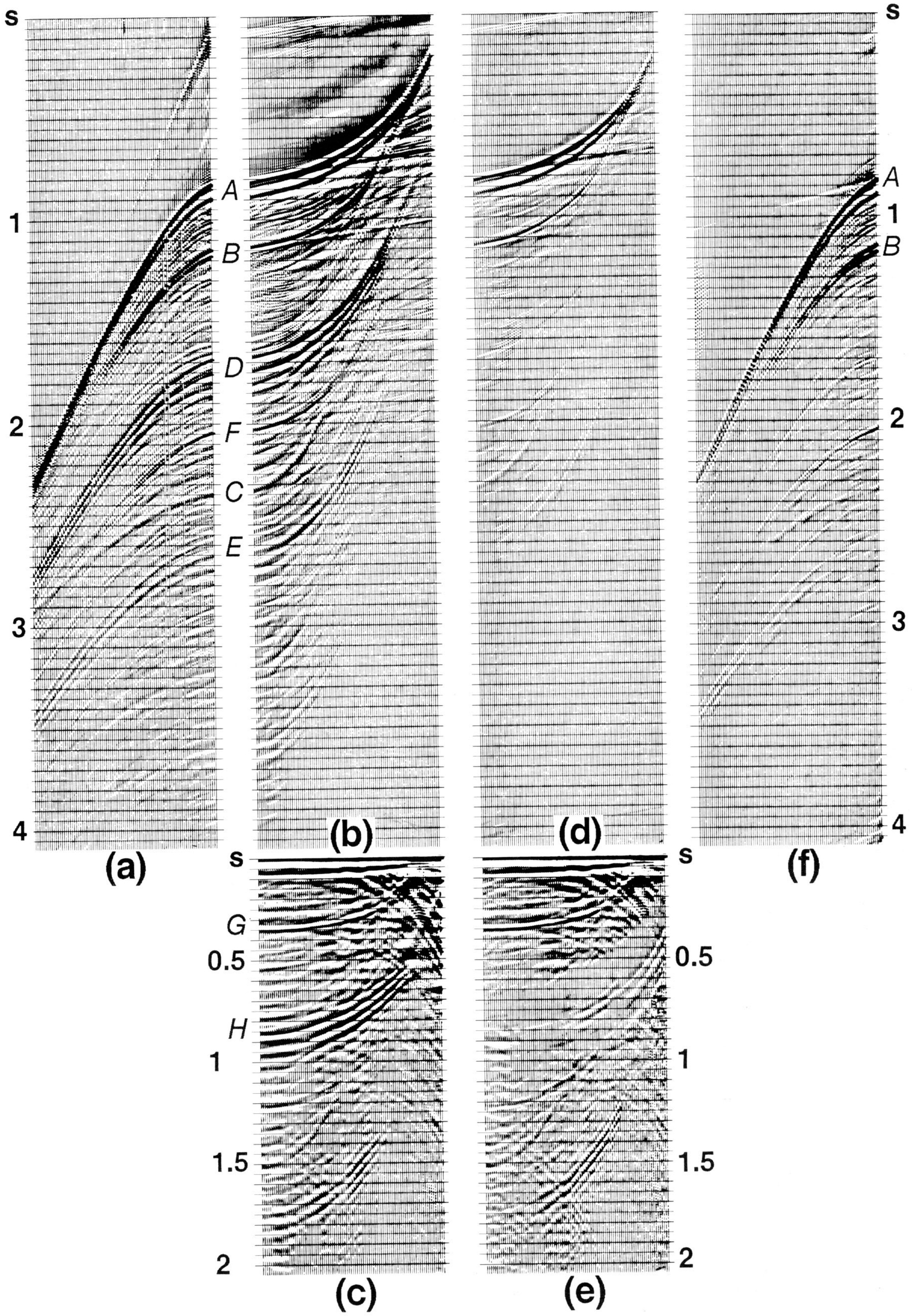
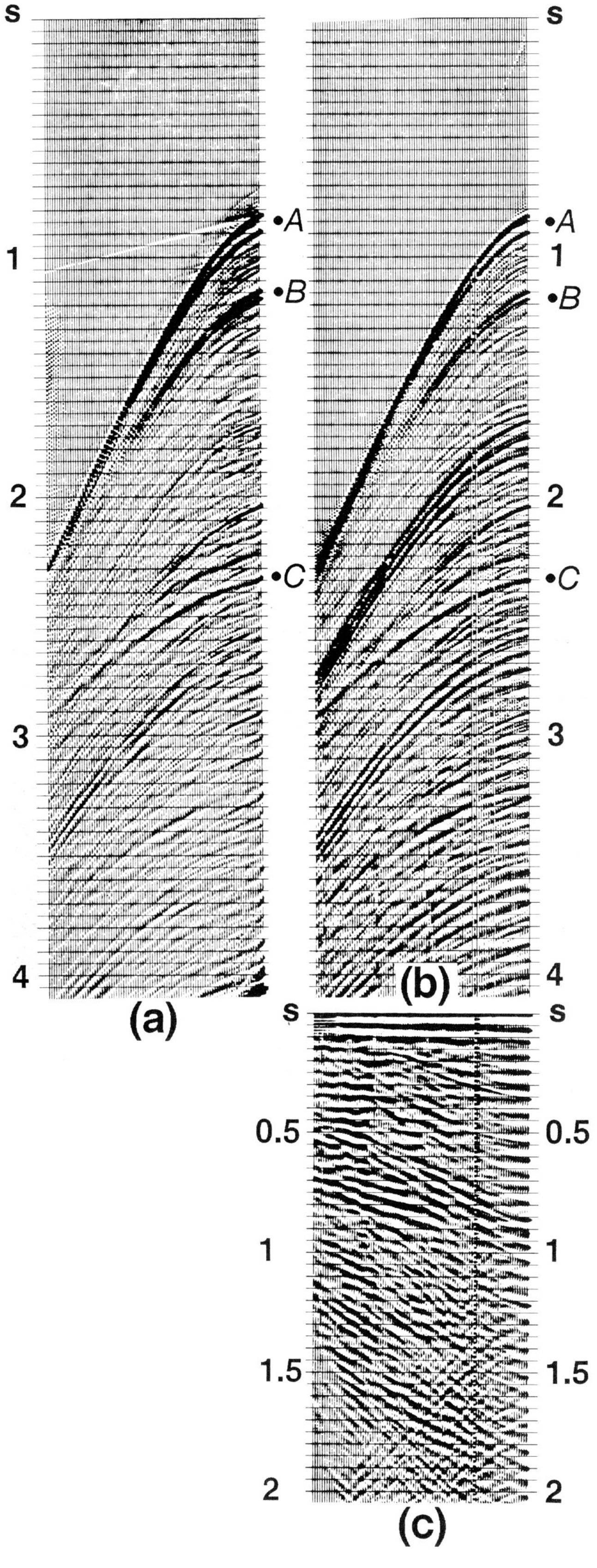


FIG. 7-31. Multiple suppression in the slant-stack domain. (a) A field record without geometric spreading correction; (b) the slant-stack gather obtained from it; (c) the autocorrelogram of (b); (d) the slant-stack gather after predictive deconvolution, where operator length = 400 ms and prediction lag (at $p = 0$) = 700 ms; (e) the autocorrelogram of (d); (f) reconstruction of the shot gather from (d). (Data courtesy Shell.)



◄ FIG. 7-32. (a) Shot gather in Figure 7-31f after geometric spreading correction; (b) shot gather in Figure 7-31a after geometric spreading correction; (c) autocorrelogram of (b). (Data courtesy Shell.)

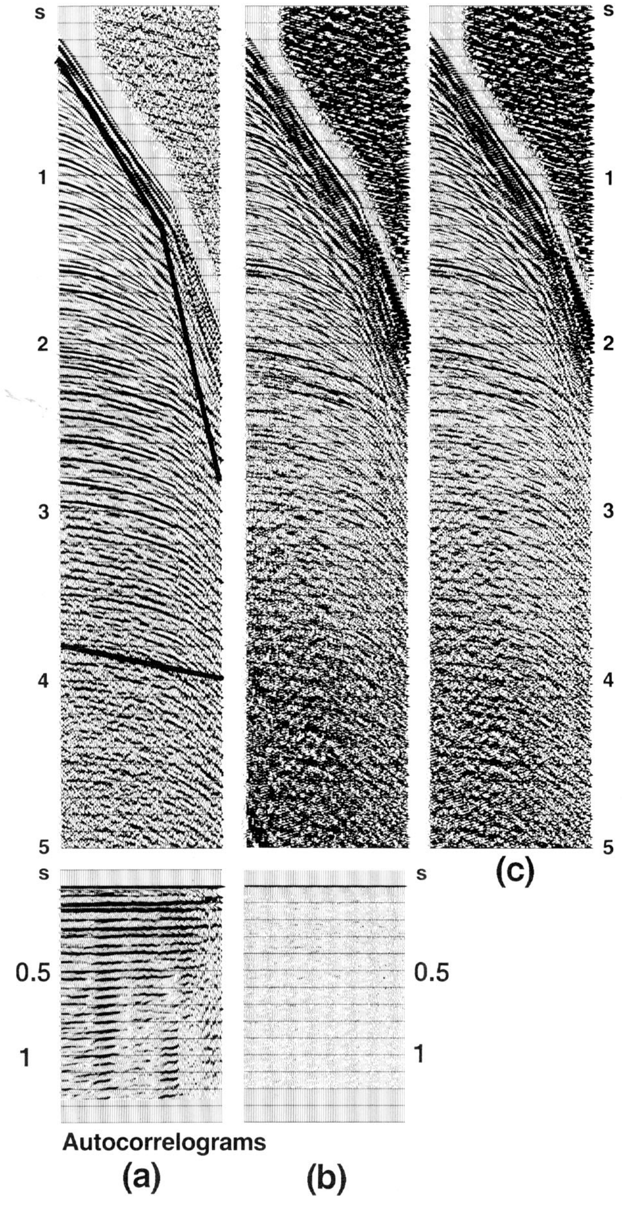


FIG. 7-33. A field record containing short-period reverberations before (a) and after deconvolution (b), followed by band-pass filtering (c). The solid lines represent the start and end times for the autocorrelation estimation windows. The spiking deconvolution operator length is 160 ms.

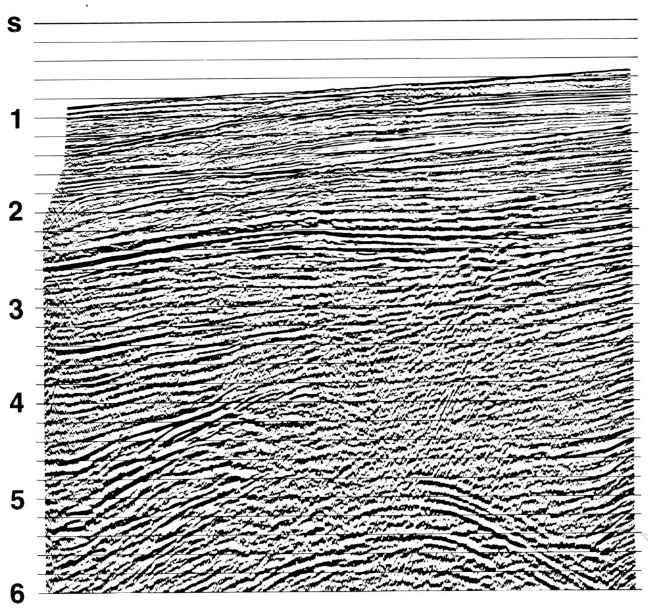


FIG. 7-34. A CMP stack associated with the gather shown in Figure 7-32b. (Data courtesy Shell.)

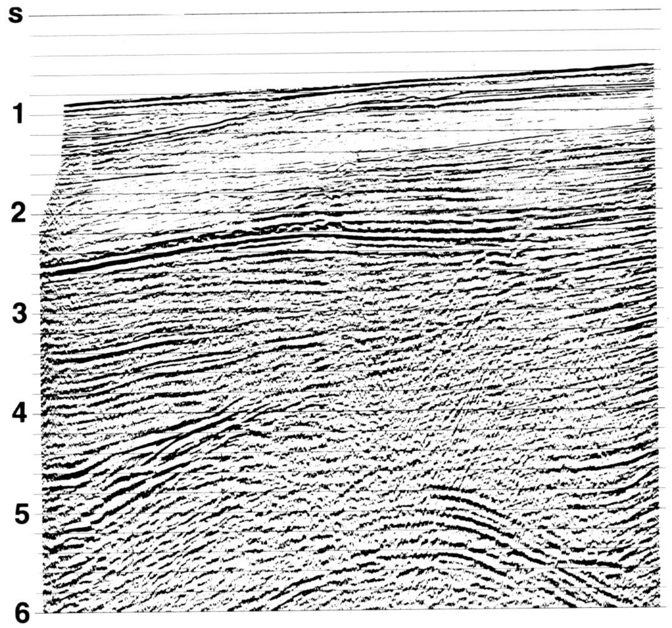


FIG. 7-35. A CMP stack associated with the gather shown in Figure 7-32a (after slant-stack multiple suppression). Compare with Figure 7-34. (Data courtesy Shell.)

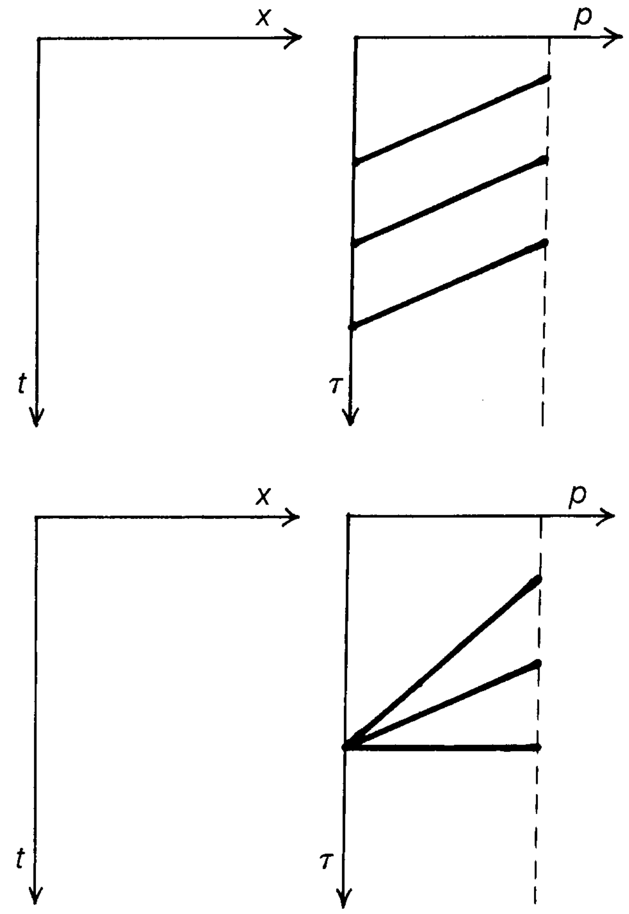


FIG. 7-36. Linear events in the p-gather are mapped onto points in the (x,t) domain (see Exercise 7.2).

section obviously provides clearer definition of the true structure. Enhancement is most apparent on the left side of the section between 3 and 5 s.

REFERENCES

Alam, A. and Lasocki, L., 1981a, Slant stack and applications: Presented at the 41st Ann. Eur. Assoc. Explor. Geophys. Mtg.

Alam, A. and Austin, J., 1981b, Multiple suppression using slant stacks: Tech. Rep., Western Geophysical Company.

Claerbout, J. F., 1978, Snell waves: Stanford Exploration Project Report No. 15, 57-72, Stanford University.

Claerbout, J. F., 1985, Imaging the earth's interior: Blackwell Scientific Publications.

Clayton, R. W. and McMechan, G., 1981, Inversion of refraction data by wave-field extrapolation: Geophysics, **46**, 860-868.

Diebold, J. B. and Stoffa, P. L., 1981, The traveltime equation, tau-*p* mapping and inversion of common midpoint data: Geophysics, **46**, 238-254.

Ewing, M., Jardetsky, W. S., and Press, F., 1957, Elastic waves in layered media: McGraw-Hill Book Co.

Gonzalez-Serrano, A., 1982, Wave-equation velocity analysis: Ph.D. thesis, Stanford University.

McMechan, G. and Yedlin, M., 1981, Analysis of dispersive waves by wave-field transformation: Geophysics, **46**, 869-879.

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

Pekeris, C. L., 1948, Theory of propagation of explosive sounds in shallow water: Geol. Soc. Am. Mem. **27**.

Press, F. and Ewing, M., 1950, Propagation of explosive sound in a liquid layer overlying a semi-infinite elastic solid: Geophysics, **15**, 111-148.

Schultz, P. S. and Claerbout, J. F., 1978, Velocity estimation by wavefront synthesis: Geophysics, **43**, 691-712.

Schultz, P. S., 1982, A method for direct estimation of interval velocities: Geophysics, **47**, 1657-1671.

Taner, M. T., 1977, Simplan—plane-wave stacking: Presented at the 37th Ann. Eur. Assoc. Explor. Geophys. Mtg.

Taner, M. T., 1980, Long-period sea-floor multiples and their suppression: Geophys. Prosp., **28**, 30-48.

Thorson, J., 1978, Reconstruction of slant stacks: Stanford Exploration Project Report No. 14, 83-91, Stanford University.

Treitel, S., Gutowski, P. R., and Wagner, D. E., 1982, Plane-wave decomposition of seismograms: Geophysics, **47**, 1375-1401.

Yilmaz, O., 1981, Modeling guided waves in an irregular water layer: Presented at the 51st Ann. Int. Soc. Explor. Geophys. Mtg.

EXERCISES

Exercise 7.1. Prove that a hyperbola in offset domain (x,t) maps onto an ellipse in slant-stack domain (p, τ).

Exercise 7.2. Refer to Figure 7-36. What would the (x,t) domains look like?

Exercise 7.3. Consider constructing the slant-stack gather from offset data that consists of a reflection hyperbola. Does equal increment in p, the ray parameter, cause undersampling or oversampling of the steep dips? Of the gentle dips? What happens when an equal increment in $(1/p)$ is used? What happens when an equal increment in θ is used, where θ is related to p by $p = \sin \theta / v$?

8

Special Topics

8.1 INTRODUCTION

This chapter discusses a variety of essentially unrelated topics: multiple attenuation, seismic resolution, seismic modeling, synthetic sonic logs, instantaneous attributes, vertical seismic profiling (VSP), and 2-D surface data processing.

Multiple attenuation is the only topic included in a conventional processing flow. Sections 2.7.5 and 7.5 discussed multiple attenuation based on prediction. Section 8.2 examines multiple attenuation based on velocity discrimination between primaries and multiples in both the *f-k* and *t-x* domains.

The remaining topics are interpretational aids that involve manipulation and simulation of seismic data. Seismic resolution is discussed in Section 8.3. Resolution is the ability to separate two events that are very close together. There are two aspects of seismic resolution: vertical (or temporal) and lateral (or spatial). Seismic resolution becomes especially important in mapping small structural features, such as subtle sealing faults, and in delineating thin stratigraphic features that may have limited areal extent.

Seismic forward modeling (Section 8.4) involves generating the traveltime response of a subsurface reflectivity model associated with a velocity-depth model. Forward modeling has a variety of applications. An understanding of the seismic responses of subsurface structural and stratigraphic features is made possible by forward modeling. Modeling also is used to generate test data for use in evaluating processing algorithms. Recording parameters, such as group interval and spread length, sometimes are chosen, at least partially, on the basis of forward models. Forward modeling also is useful for determining whether the reflection response of an interpreted geologic model is consistent with the CMP stacked sections used in the interpretation.

The synthetic sonic log (Section 8.5) is a simple 1-D inversion of reflection data. While the middle-and high-frequency components of the synthetic sonic log are derived from the seismic trace, the low-frequency trend must be obtained from separate sources of information such as conventional velocity analysis or actual sonic-log data.

Instantaneous attributes (Section 8.6) can help emphasize the continuity of reflectors and delineate seismic depositional sequence boundaries. When displayed in color, reflectivity strength, instantaneous phase, and instantaneous frequency are useful tools in stratigraphic exploration.

Vertical seismic profiling is reviewed in Section 8.7. The basic processing sequence and the use of VSP data are outlined. Finally, Section 8.8 reviews 2-D map data processing.

8.2 MULTIPLE SUPPRESSION

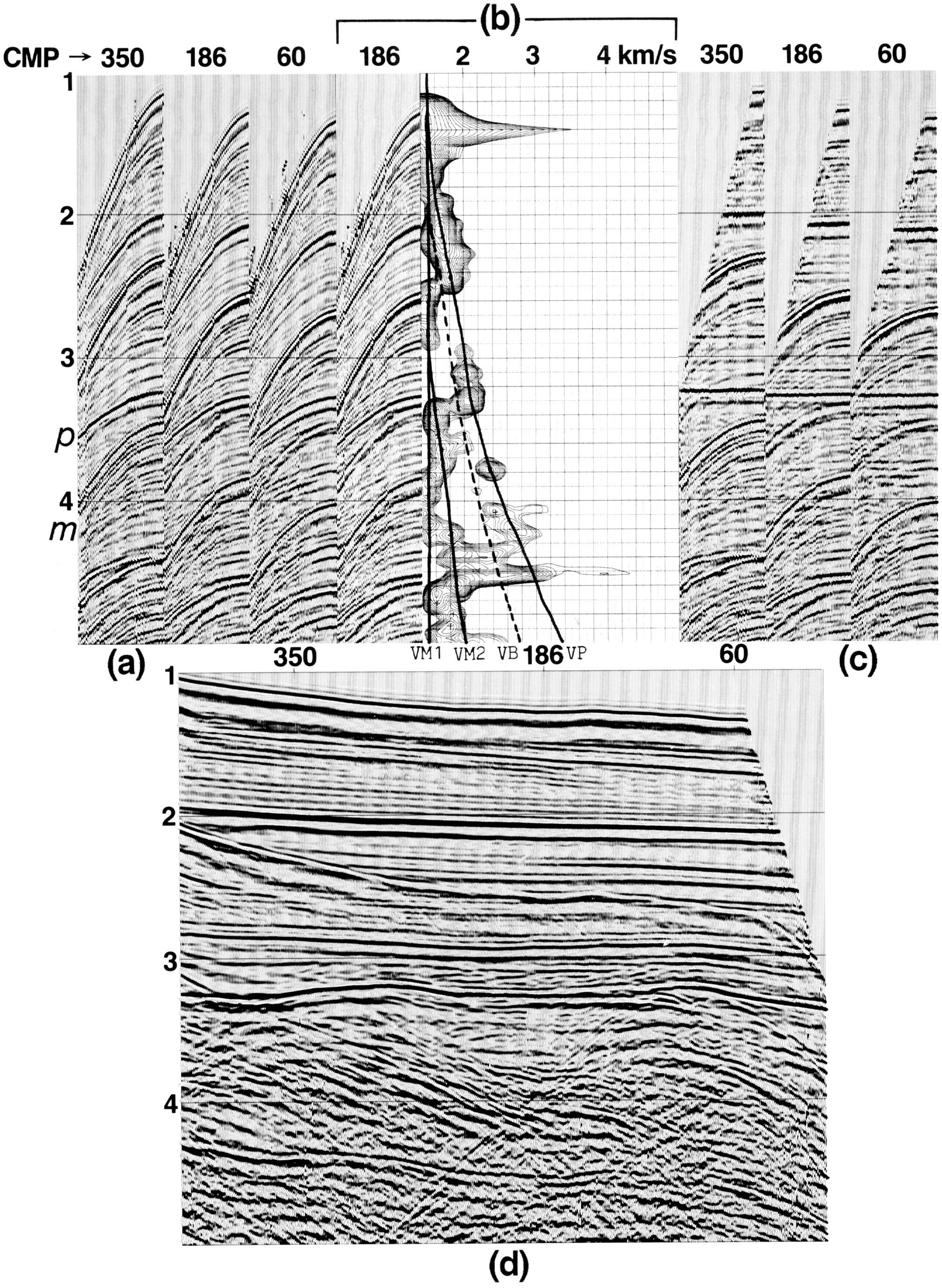
Section 7.5 discussed multiple attenuation based on the periodicity of multiples in the slant-stack domain. The two techniques presented now are based on the moveout difference between primaries and multiples. The CMP gathers in Figure 8-1a clearly illustrate this moveout difference. A primary *p* typically has less moveout than a multiple *m*. From the velocity spectrum in Figure 8-1b, note the difference between the velocity trends associated with primaries VP and multiples VM1 and VM2. The VM1 and VM2 velocity functions represent the water-bottom and peg-leg multiples, respectively. If NMO correction is applied using the primary velocities, as is normally done to generate final stacks, then the primaries are aligned while the multiples are undercorrected (Figure 8-1c). This suggests that CMP stacking itself is a viable method of multiple suppression. The CMP stack derived from the gathers in Figure 8-1c is shown in Figure 8-1d.

The synthetic CMP gather in Figure 8-2c contains five primaries, including the water-bottom reflection *W* and the multiples associated with the water-bottom reflection. Note that the water-bottom multiples have higher amplitudes than the primaries. The velocity spectrum shows a significant separation between the velocity functions for multiples VM and primaries VP. Stacking with the primary velocity function should, to a large extent, discriminate against the multiples and result in a section that contains essentially the primary energy as shown in Figure 8-3. The stack trace is repeated to better examine the amplitudes.

Stacking far offsets work to suppress multiples. However, stacking near offsets works against multiple suppression, since the moveout difference between the primaries and multiples is negligibly small on those offsets (see, for example, Figure 8-1c). The simplest way around this problem is to apply an inside mute to the CMP gathers before stacking. Another problem then emerges: the outside mute. The severity of this mute governs the amount of far-offset data left at early times for velocity discrimination (Figure 8-1c). If there is a severe multiple problem, an effort must be made to preserve the maximum amount of far-offset data associated with target events. The stacked section with inside mute applied is shown in Figure 8-4a. When compared with Figure 8-1d, note that the deeper peg-leg multiple below 4 s has been further attenuated somewhat. The difference between the conventional CMP stack (Figure 8-1d) and the inside mute stack (Figure 8-4a) indicates the amount of energy, mostly multiples, that was removed by the inside mute (Figure 8-4b). Sometimes more sophisticated muting schemes in which weights (between 0 and 1) are assigned to each offset (the smaller weights assigned to the near offsets) during stacking can produce better results than those shown here.

8.2.1 Velocity Discrimination in the *f-k* Domain

Figure 8-5a shows the synthetic CMP gather (top) from Figure 8-2c and its 2-D amplitude spectrum (bottom). The primary and multiple energy can be separated into two different quadrants in the *f-k* plane. This is achieved by NMO correcting the gather using a velocity function


(b)
CMP → 350
186
60
186
2
3
4 km/s
350
186
60
1
2
3
p
4
m
VM1
VM2
VB
186
VP
(a)
350
60
(c)
1
2
3
4
(d)

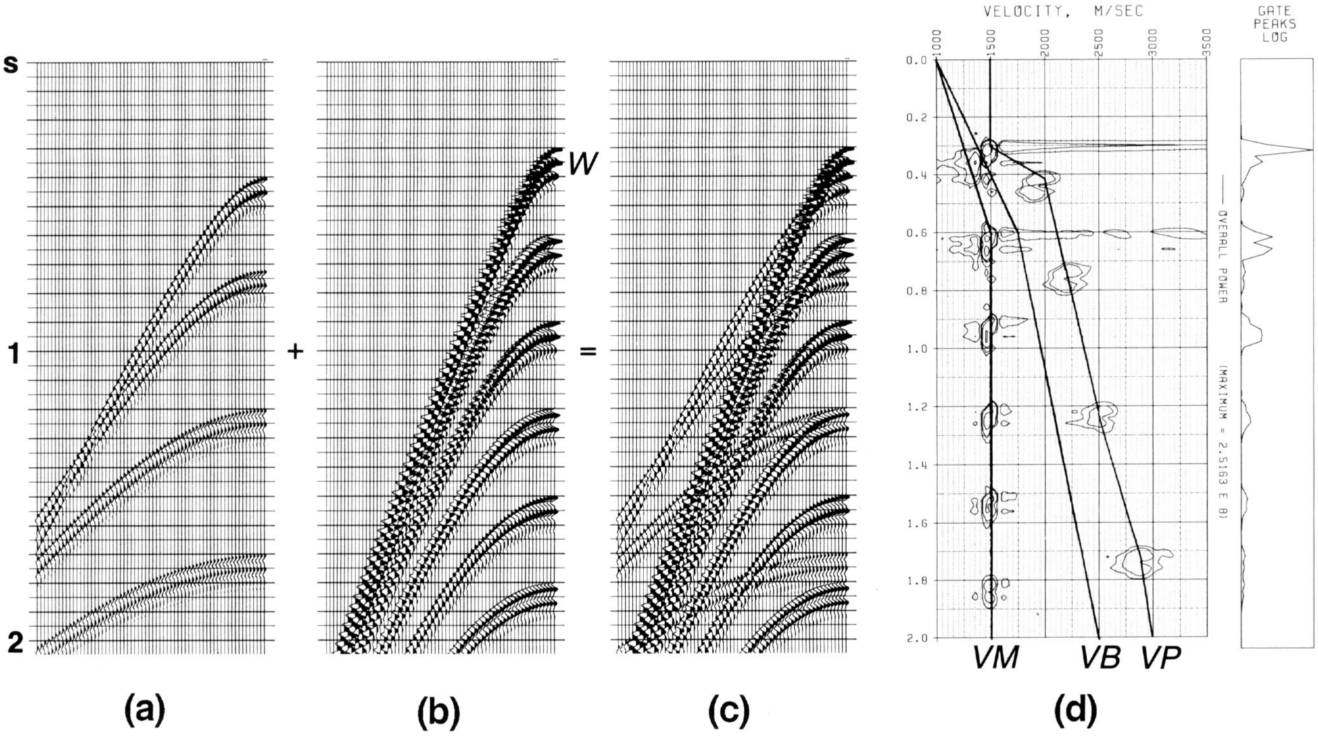


FIG. 8-2. Synthetic CMP gathers containing (a) primaries, (b) water-bottom multiples, (c) superposition of (a) and (b). (d) The velocity spectrum derived from (c). Here, W = water-bottom primary, VM = velocity function for multiples, VP = velocity function for primaries, VB = a velocity function between VM and VP used in generating Figure 8-5b.

FIG. 8-1. (a) Three CMP gathers with strong multiples; (b) velocity analysis at CMP 186, where VP = primary velocity trend, VM1 = slow (water-bottom) multiples, and VM2 = fast (peg-leg) multiples. (VB is the velocity function used in generating Figure 8-8a.) For reference, the CMP gather is displayed next to the velocity spectrum. (c) The same CMP gathers as in (a) after NMO correction using the primary velocities. (d) CMP stack using the gathers as in (c). (Data courtesy Petro-Canada Resources.)

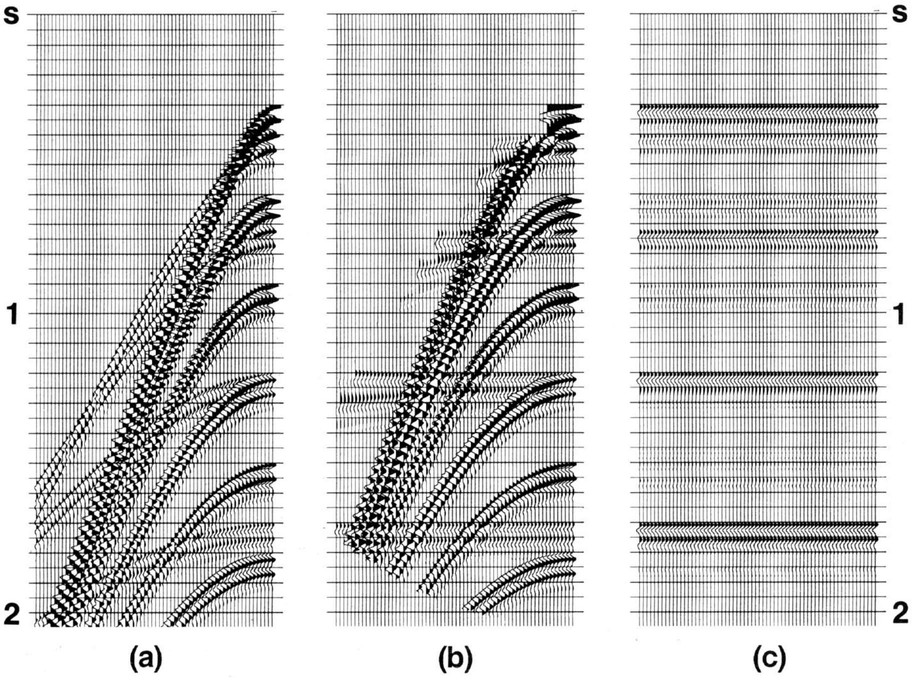
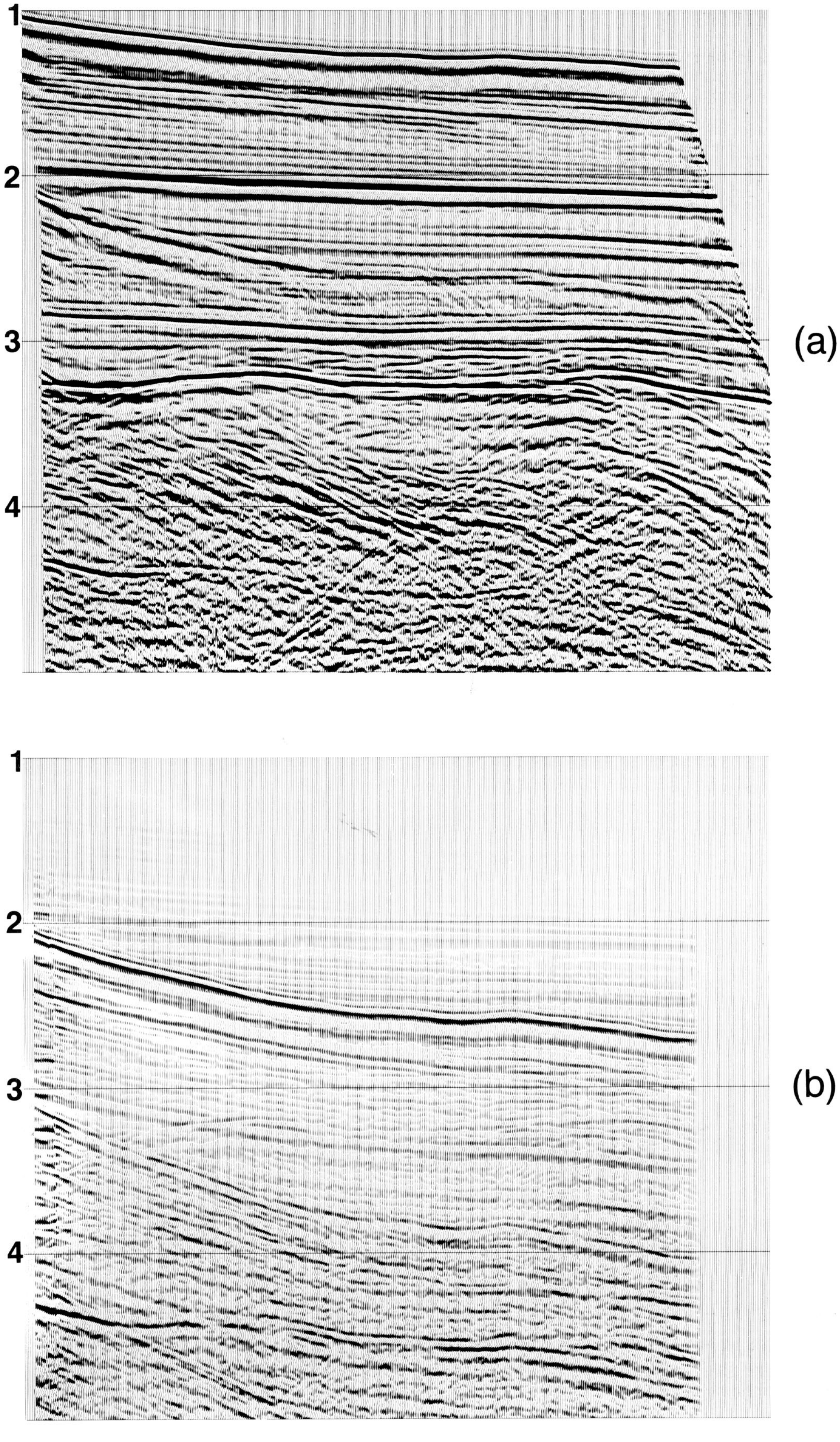


FIG. 8-3. (a) The CMP gather in Figure 8-2c and (b) after NMO correction using the primary velocity function (VP in Figure 8-2d). (c) The stack trace derived from (b) repeated to emphasize the strong events.

FIG. 8-4. (a) The CMP stack derived from the CMP gathers in Figure 8-1c with inside mute applied. Compare this with Figure 8-1d. (The inside mute pattern can be recognized on the left edge of the section.) (b) The difference between the conventional CMP stack (Figure 8-1d) and the inside mute stack (a). (Data courtesy Petro-Canada Resources.)


1
2
3
4
(a)
1
2
3
4
(b)

(denoted by VB in Figure 8-2d) that is between the primary and multiple velocities. The resulting NMO-corrected gather and its 2-D amplitude spectrum are shown in Figure 8-5b. The multiples are undercorrected, while the primaries are overcorrected. On the *f-k* plane, the multiples and primaries map largely onto two different quadrants (indicated by *P* for primaries and *M* for multiples). The exception to this separation is the near-offset energy (both primary and multiple), which almost entirely maps along the frequency axis. This is because multiples and primaries have no significant moveout difference at near offsets. Aliased energy (such as *A*) is wrapped around and mapped to the wrong quadrant. (Spatial aliasing is discussed in detail in Section 1.6.1.)

Multiples can be suppressed by zeroing the quadrant corresponding to multiple energy in the *f-k* domain (Figure 8-5c) (Ryu, 1980; Sengbush, 1983). Inverse NMO correction (Figure 8-5d) using the same intermediate velocity function VB restores the original moveout of the primaries. However, note that spatially aliased multiple energy remains in the gather (A in Figure 8-5). Besides zeroing out the multiple quadrant, a reject zone can be imposed in the primary quadrant (Figure 8-6a, where the reject zone is denoted by *R*.) The *f-k* filtered CMP gather (Figure 8-6b) now also is free of the aliased energy. (Compare this with Figure 8-5d). Following this proce-

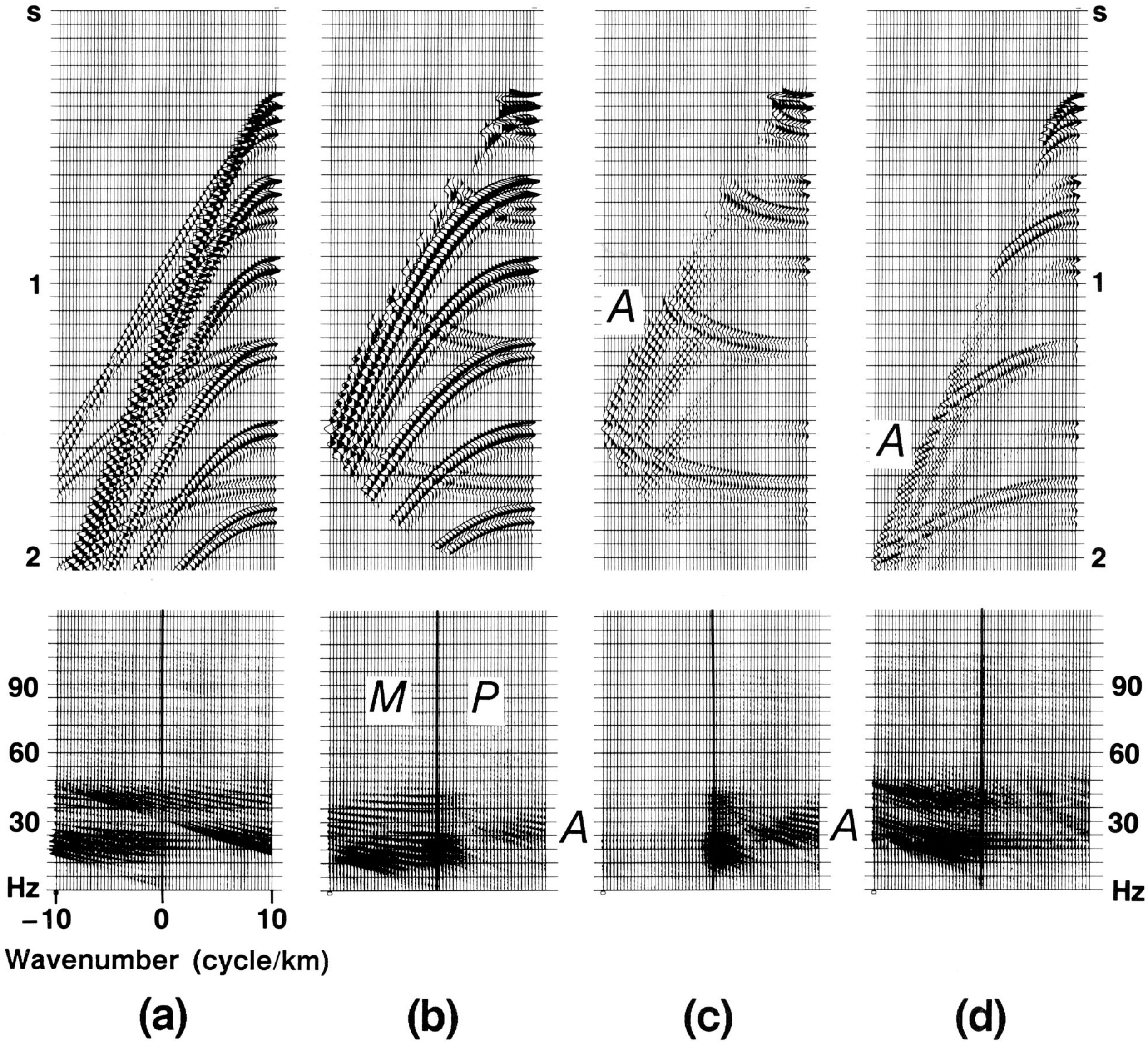


FIG. 8-5. (a) The CMP gather in Figure 8-2c, (b) after NMO correction using a velocity function (VB in Figure 8-2d) between the multiple and primary trend, (c) the result of zeroing the *f-k* quadrant associated with multiples, (d) the gather in (c) after inverse NMO using the same velocity function as in (b). The bottom panels are the corresponding *f-k* spectra.

dure, apply NMO correction using the primary velocity function VP and stack (Figures 8-6c and d). Figure 8-7 is a flowchart for *f-k* multiple attenuation.

Now consider the field data example shown in Figure 8-1. The moveout difference between the primaries and multiples is apparent in the CMP gathers (Figure 8-1a) and the velocity spectrum (Figure 8-1b). To the left of solid line VP, all peaks are associated with water-bottom and peg-leg multiples. Apply a moveout correction using a velocity function VB that is somewhere between the primary and multiple velocities. As a result, the multiples are undercorrected and the primaries are overcorrected as shown in Figure 8-8a. Now consider the moveout of primaries and multiples in the *f-k* domain. The primary and multiple indicated in Figure 8-1a map to the same quadrant (say positive quadrant) in the *f-k* domain before moveout correction. The same events after moveout correction using an intermediate velocity map into two different quadrants; in particular, the multiple maps into the positive quadrant and the primary maps into the negative quadrant. Thus, by zeroing one quadrant in which the multiples are clustered, the primaries can be enhanced.

A CMP gather and the associated velocity spectrum following *f-k* multiple suppression are shown in Figure 8-8b. When compared with Figure 8-1b, Figure 8-8b

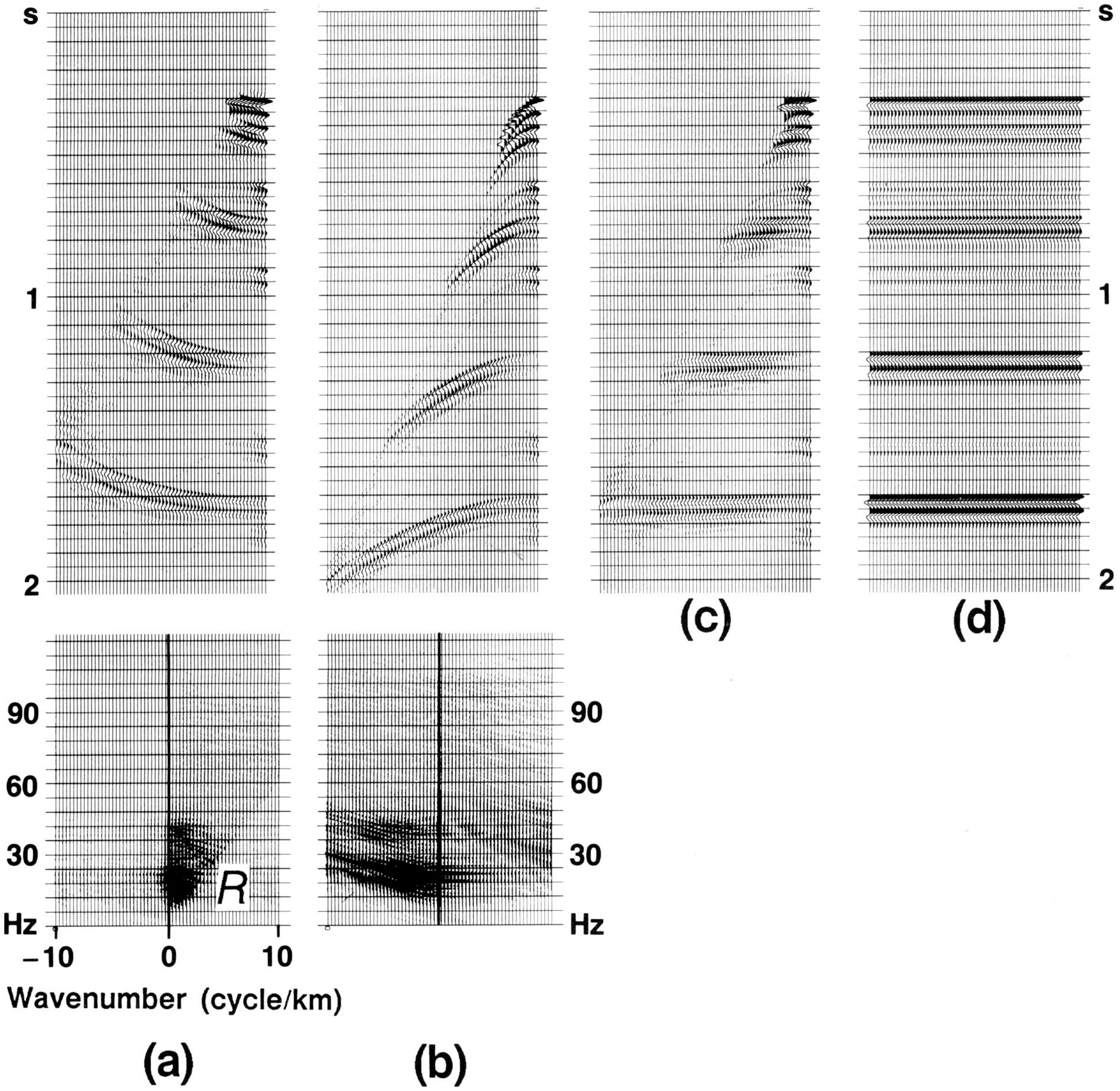


FIG. 8-6. (a) The same figure as in Figure 8-5c, except that in addition to zeroing the left quadrant, a portion of the right quadrant in the *f-k* spectrum (denoted by R) also is zeroed to suppress aliased energy: (b) the result of applying inverse NMO to (a) using velocity function VB in Figure 8-2d; (c) the result of applying NMO correction to (b) using primary velocity function VP in Figure 8-2d; (d) stack of (c) repeated to emphasize the strong events. The bottom panels show the corresponding *f-k* spectra.

shows that the multiples region in the velocity spectrum was suppressed, while the primary velocity trend was enchanced. Finally, Figure 8-8c shows selected CMP gathers after moveout correction using the primary velocities, picked from the velocity spectrum in Figure 8-8b. Stacking of these gathers yields the section in Figure 8-8d. This section should be compared with Figures 8-1d and 8-4a.

In practice, there are variations in selecting the moveout velocity in the flowchart of Figure 8-7. Some prefer to apply NMO correction using the multiple velocity, then zeroing the energy along the frequency axis and the multiple quadrant of the *f-k* spectrum. Others prefer to apply NMO correction using primaries and placing a tight pass-zone around the frequency axis. Finally, note that this *f-k* method of velocity discrimination is one type of *f-k* filtering. Thus, we must deal with the same practical issues discussed in Section 1.6.2; in particular, wrap-around, spatial aliasing, and tapering over the boundary between the pass/reject zones.

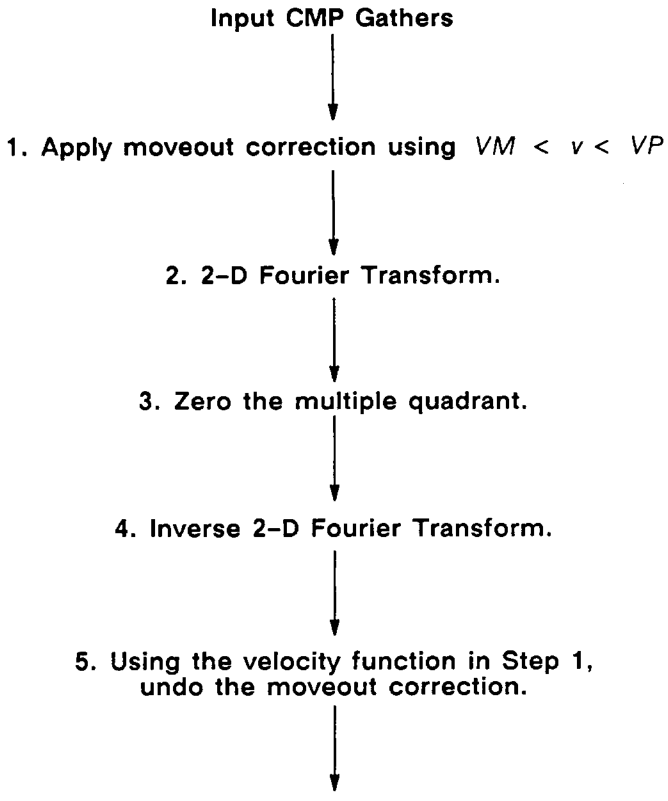


FIG. 8-7. Flowchart for *f-k* multiple suppression.

8.2.2 Velocity Discrimination in the *t-x* Domain

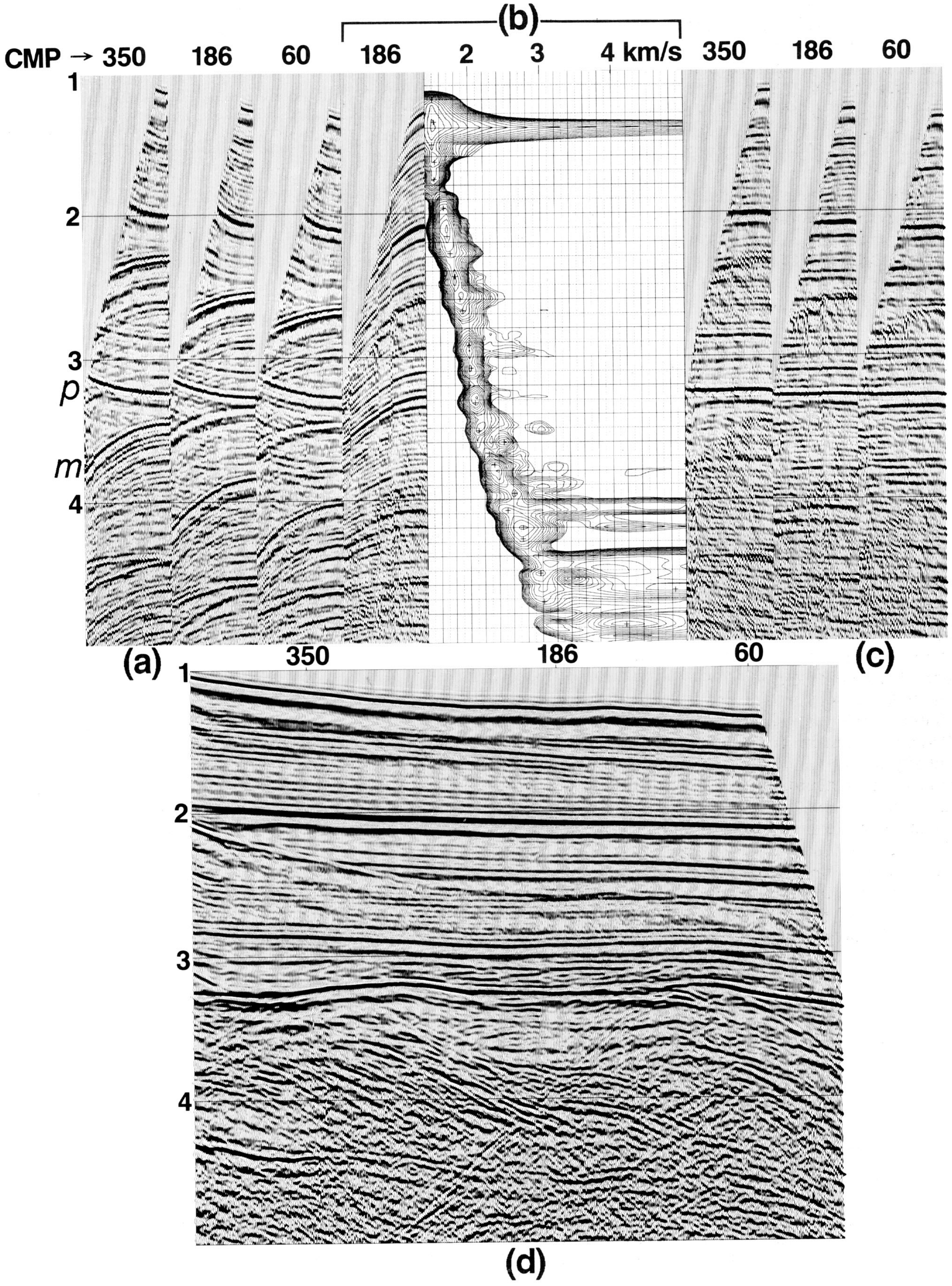
Another approach to multiple suppression operates in the *t-x* domain. Again, consider the synthetic CMP gather in Figure 8-9a, which is the same as that in Figure 8-2c. Apply NMO correction, this time using the multiple velocity function (VM in Figure 8-2d). The result is shown in Figure 8-9b, while the stack trace is shown in Figure 8-9c. Figure 8-9c, which is called the model trace for multiples, contains almost entirely the multiple energy. Subtract this trace from the individual traces of the NMO-corrected gather (Figure 8-9b). The resulting traces essentially should contain only primary energy. Note that this model-based approach applies to one multiple velocity function at a time.

The main problem with this technique is constructing a model trace that contains only multiples. Because of slight waveform changes and variation of the moveout differential between primaries and multiples with offset, the model trace for multiples will not represent multiples equally well for each offset. Better representations of multiple energy can be obtained by constructing individual model traces for each offset by stacking only a few traces on both sides of the trace associated with that offset.

Even if individual model traces were used, it still is difficult to generate model traces that do not contain some primary energy. Good suppression of primary energy in the model trace ultimately depends on the moveout differential between primaries and multiples being a substantial fraction of the period of the seismic wavelet. At lower temporal frequencies, this usually is not the case and, hence, the model trace often includes some of the low-frequency components of the primaries. Consequently, subtraction of the model trace from the moveout-corrected traces often leads to suppression of the multiples and the low-frequency components of the primaries. Exclusion of the low-frequency end of the spectrum in building the model traces is a way to deal with this latter problem.

To study the results of this subtraction technique on field data, consider the selected CMP gathers in Figure 8-1a. From the velocity spectrum in Figure 8-1b, note that the

FIG. 8-8. (a) The CMP gathers in Figure 8-1a after NMO correction using a velocity function (VB in Figure 8-1b) between the primary and multiple velocities; (b) velocity spectrum at CMP 186 estimated from the *f-k* dip-filtered gather shown to the left of the spectrum. (Compare this with Figure 8-1b.) (c) The same CMP gathers as in (a) after *f-k* multiple attenuation followed by NMO correction using the primary velocities derived from velocity spectrum (b). (d) The CMP stack derived from the CMP gathers as in (c) after *f-k* multiple suppression. Compare this with Figures 8-1d and 8-4a. (Data courtesy Petro-Canada Resources.).


(b)
CMP → 350
186
60
186
2
3
4 km/s
350
186
60
1
2
3
p
m
4
(a)
(c)
350
186
60
1
2
3
4
(d)

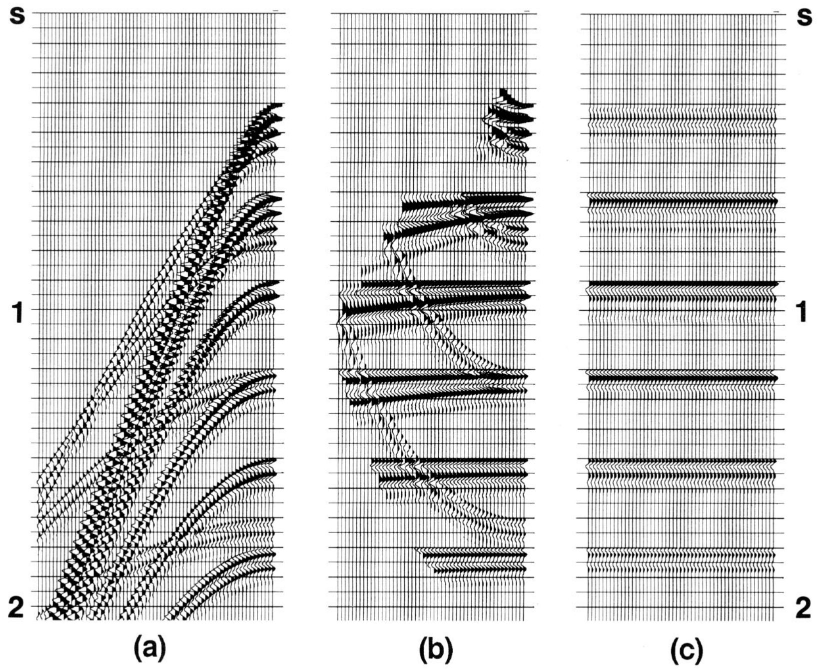


FIG. 8-9. (a) The CMP gather in Figure 8-2c, (b) after NMO correction using the multiple velocity function (VM in Figure 8-2d). (c) The stack trace is repeated to emphasize the strong events.

multiples can have more than one velocity trend (the lines denoted as VM1 and VM2). The NMO-corrected CMP gathers in Figure 8-10a are obtained by using one of the velocity trends (VM1). The primaries are overcorrected, while the multiples associated with the velocity trend VM1 are flattened. The crucial task is to recognize and pick multiple trends on the velocity spectrum. The velocity spectrum after multiple attenuation (Figure 8-10b) shows an enhanced primary velocity trend. Also note the removal of the multiple trend (VM1) from the velocity spectrum. The selected CMP gathers following moveout correction using the primary velocities from Figure 8-10b are shown in Figure 8-10c. The stacked section after applying the multiple attenuation procedure is shown in Figure 8-10d. This section should be compared with Figures 8-1d, 8-4a, and 8-8d.

This model-based approach can be cascaded to attenuate more than one class of multiples present in the data. Using the multiple velocities VM2, indicated in Figure 8-1b, yields the results shown in Figure 8-11. Input CMP gathers to the second pass (Figure 8-11a) are the output CMP gathers from the first pass. Note the attenuation of the multiple trend VM2 from the velocity spectrum (Figure 8-11b). The deeper peg-leg multiple below 4 s has been suppressed further (compare Figures 8-10d and 8-11d).

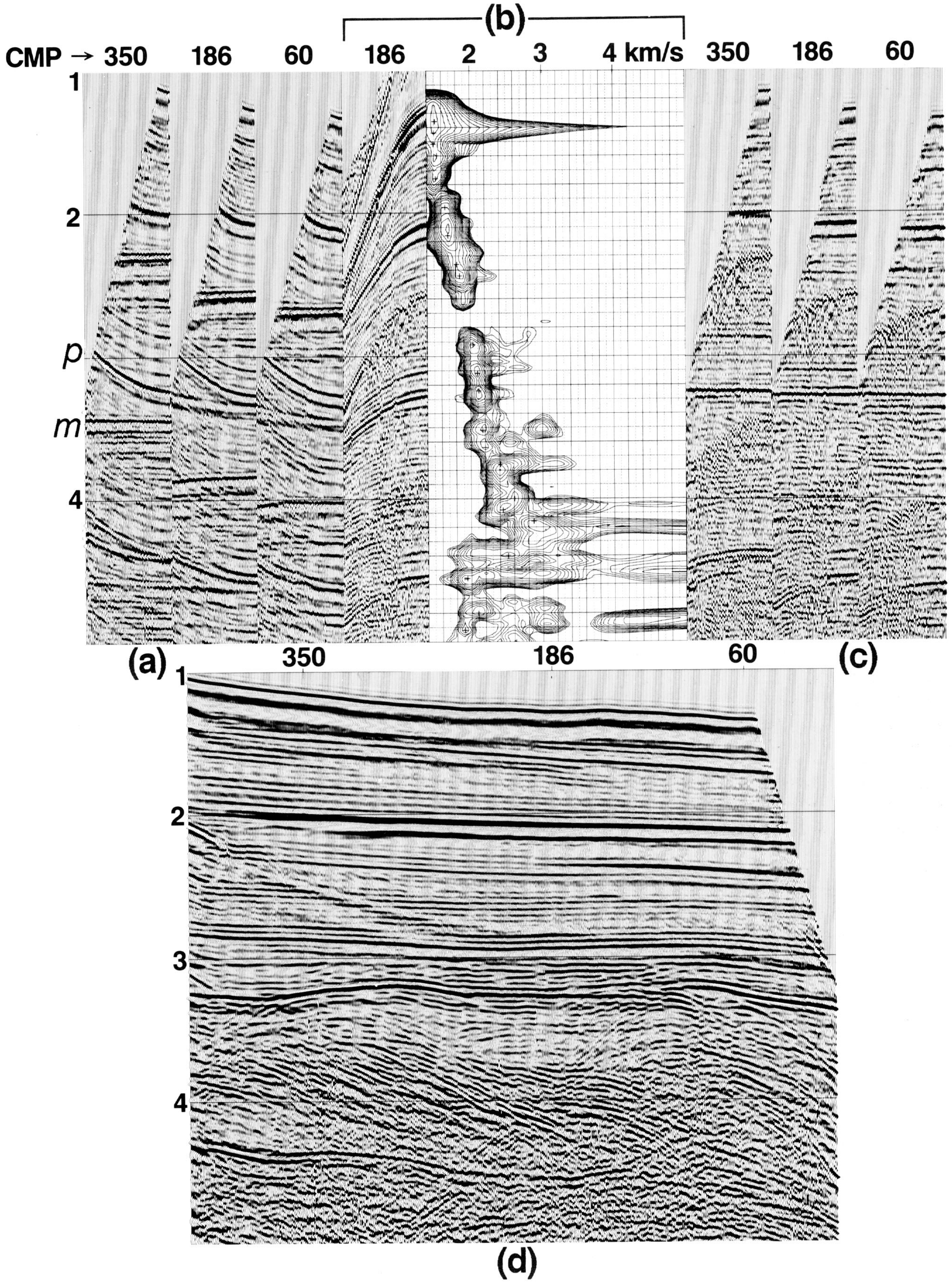
The stacked sections resulting from the first pass (Figure 8-10d) and the second pass (Figure 8-11d) have a high-frequency character compared to the conventional CMP stack (Figure 8-1d). As indicated earlier, this effect can be suppressed by excluding the low frequencies from the model traces. Multiple attenuation using the filtered versions of the model traces yields the stacked sections in Figure 8-12.

Finally, because there is relatively less moveout differential between the primaries and multiples in the near-offset range, the inside mute (or some kind of weighted stacking) helps suppress the multiples. Therefore, it may help to cascade any one of the multiple suppression techniques described here (and in Section 7.5) with the inside mute during stacking.

We discussed multiple suppression techniques based on (a) velocity discrimination between multiples and primaries (this section), and (b) periodicity of multiples (Section 7.5). While these techniques seem to have a good conceptual basis, their performance on field data often is

→

FIG. 8-10. (a) The CMP gathers in Figure 8-1a after NMO correction using slow multiple velocities (VM1 in Figure 8-1b). (b) The velocity spectrum at CMP 186 after single-pass *t-x* multiple attenuation. For reference, the CMP gather after multiple attenuation is shown to the left of the velocity spectrum. (Compare this with Figure 8-1b.) (c) The same CMP gathers as in (a) after the single-pass multiple attenuation, followed by NMO correction using primary velocities derived from velocity spectrum (b). (d) The CMP stack derived from the CMP gathers as in (c) after multiple suppression. Compare this with Figures 8-1d, 8-4a, and 8-8d. (Data courtesy Petro-Canada Resources.)


(b)
CMP → 350
186
60
186
2
3
4 km/s
350
186
60
1
2
p
m
4
(a)
350
186
60
(c)
1
2
3
4
(d)

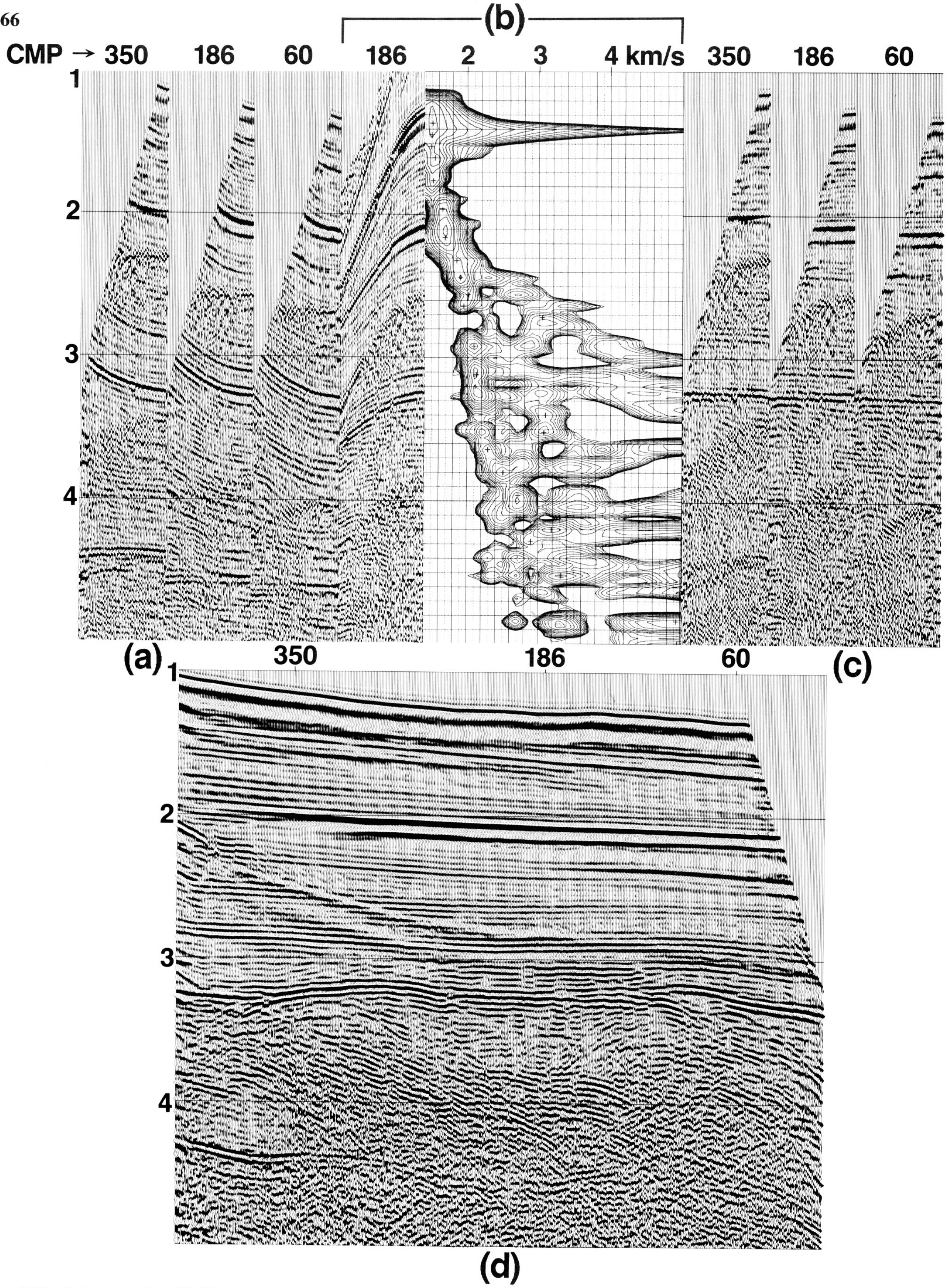


FIG. 8-11. (a) The CMP gathers from the first-pass *t-x* multiple attenuation (Figure 8-10) after NMO correction using fast multiple velocities (VM2 in Figure 8-1b). (b) The velocity spectrum at CMP 186 after second-pass *t-x* multiple attenuation. For reference, the CMP gather after multiple attenuation is shown to the left of the velocity spectrum. (c) The same CMP gathers as in (a) after the second-pass multiple attenuation, followed by NMO correction using primary velocities from (b). (d) The CMP stack derived from the CMP gathers as in (c) after the second-pass multiple attenuation. (Data courtesy Petro-Canada Resources.)

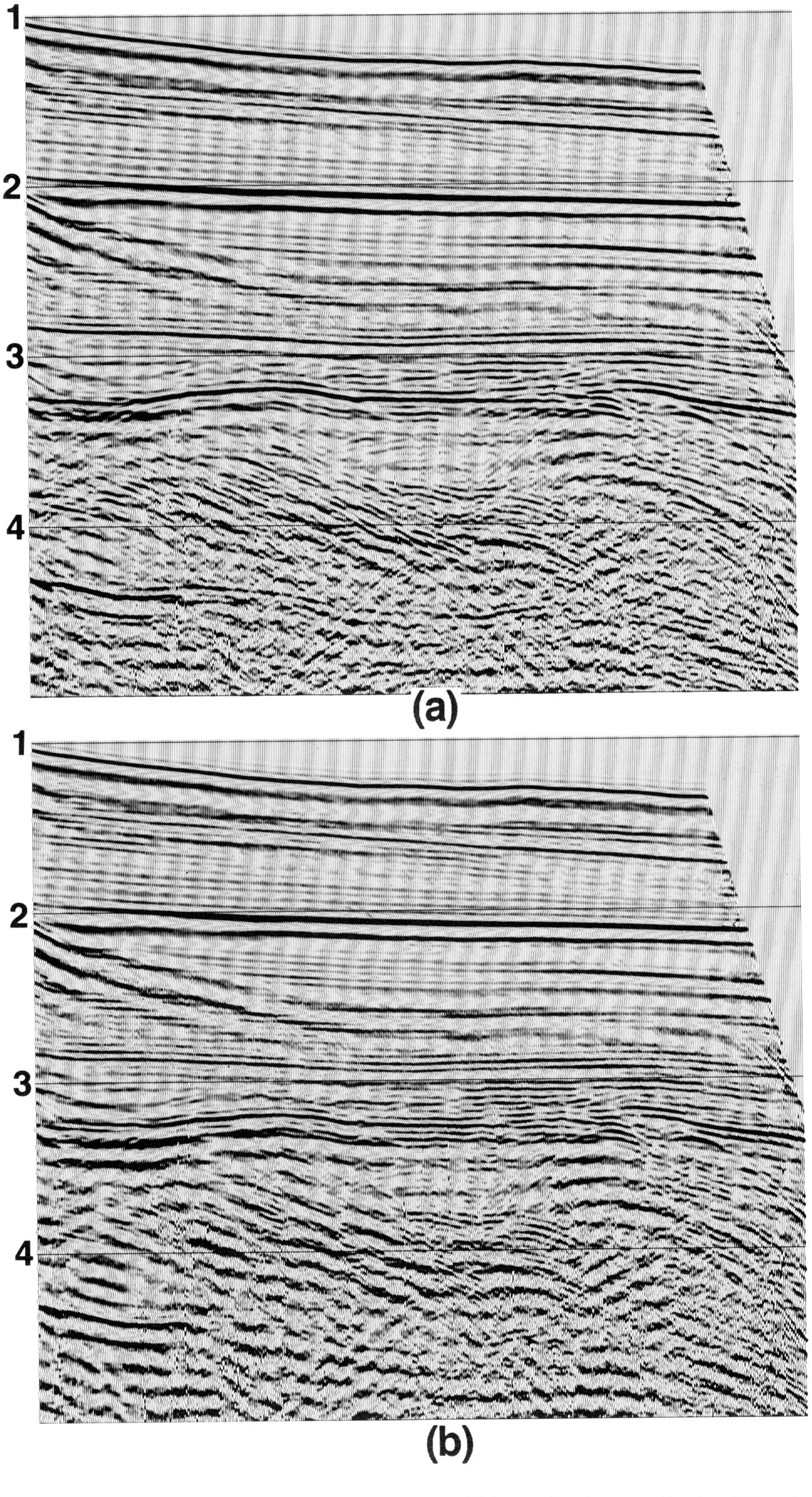


FIG. 8-12. The CMP stacks after *t-x* multiple attenuation, which was implemented using filtered model traces. (a) First pass using multiple velocites VM1 and (b) second pass using multiple velocities VM2, as shown in Figure 8-1b. Compare these with Figures 8-10d and 8-11d. (Data courtesy Petro-Canada Resources.)

disappointing. There are several possible explanations for this. First, for velocity discrimination techniques to be effective, significant moveout differences must exist between the primaries and multiples. However, the inability to exploit the large moveout differences between the primaries and multiples in the mute zone works against the methods based on velocity discrimination. There is also a problem caused by the application of geometric spreading correction (Section 1.5), which is applied using the primary velocity function. This type of correction usually results in enhancement of multiple amplitudes. The slant-stack approach (Section 7.5) is implemented before the geometric spreading correction, so there is no danger of boosting the multiple energy. Nevertheless, there is no guarantee that periodicity of multiples is well preserved across the entire profile. The layered-earth assumption, which is required by the slant-stack approach, often is violated even by small structural irregularities in the lateral direction (such as a rugged water bottom).

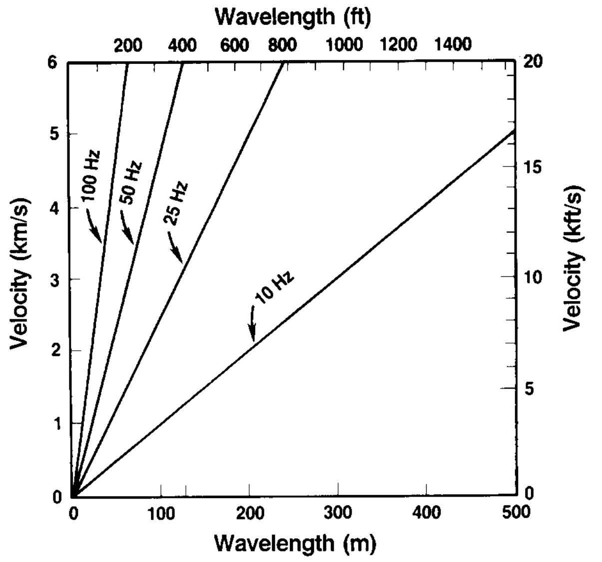


FIG. 8-13. The relationship between velocity, dominant frequency, and wavelength. Here, wavelength = velocity/frequency. (Adapted from Sheriff, 1976; courtesy American Association of Petroleum Geologists.)

Table 8-1. Threshold for vertical resolution.

	$\lambda/4 = v/4f$	
v (m/s)	f (Hz)	$\lambda/4$ (m)
2000	50	10
3000	40	18
4000	30	33
5000	20	62

8.3 SEISMIC RESOLUTION

Resolution relates to how close two points can be, yet still be distinguished. Two types of resolution are considered: vertical and lateral. Both are controlled by spectral bandwidth. The yardstick for vertical resolution is the dominant wavelength, which is wave velocity divided by dominant frequency. Deconvolution tries to increase the vertical resolution by broadening the spectrum, thereby compressing the seismic wavelet. The yardstick for lateral resolution is the Fresnel zone, a circular area on a reflector whose size depends on the depth to the reflector, the velocity above the reflector and, again, the dominant frequency. Migration improves the lateral resolution by decreasing the width of the Fresnel zone, thus separating features that are blurred in the lateral direction.

8.3.1 Vertical Resolution

For two reflections, one from the top and one from the bottom of a thin layer, there is a limit on how close they can be, yet still be separable. This limit depends on the thickness of the layer and is the essence of the problem of vertical resolution.

The dominant wavelength of seismic waves is given by:

$$\lambda = v/f, \tag{8.1}$$

where v is velocity and f is the dominant frequency. Seismic wave velocities in the subsurface range between 2000 and 5000 m/s and generally increase in depth. On the other hand, the dominant frequency of the seismic signal typically varies between 50 and 20 Hz and decreases in depth. Therefore, typical seismic wavelengths range from 40 to 250 m and generally increase with depth. Since wavelength determines resolution, deep features must be thicker than the shallow features to be resolvable. A graph of wavelength as a function of velocity for various values of frequency is plotted in Figure 8-13. The wavelength is easily determined from this graph, given the velocity and dominant frequency.

The acceptable threshold for vertical resolution generally is a quarter of the dominant wavelength. This is subjective and depends on the noise level in the data. Sometimes the quarter-wavelength criterion is too generous, particularly when the reflection coefficient is too small and no reflection event is attainable. Sometimes the criterion may be too stringent, particularly when events do exist and their amplitudes can be picked with ease.

Table 8-1 contains the wavelength threshold values for vertical resolution, considering the realistic velocity and frequency ranges. For example, a shallow feature with 2000 m/s velocity and 50-Hz dominant frequency potentially can be resolved if it is as thin as 10 m. A thinner feature cannot be resolved. Similarly, for a deep feature with velocity as high as 5000 m/s and dominant frequency

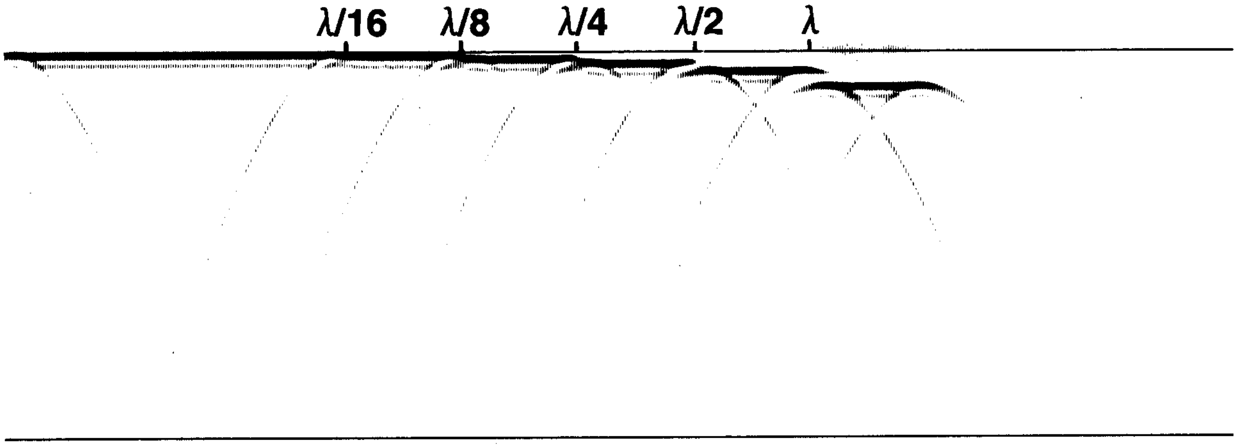


FIG. 8-14. Faults with different amounts of vertical throws expressed in fractions of the dominant wavelength.

as low as 20 Hz, the thickness must be at least 62 m for it to be resolvable.

It is now appropriate to ask whether a thin stratigraphic unit must be resolved to be mapped. The answer is no. Resolution as defined here and in the geophysical literature implies that reflections from the top and bottom of a thin bed are seen as separate events or wavelet lobes. Using this definition, resolution does not consider amplitude effects. The thickness and areal extent of beds below the resolution limit often can be mapped on the basis of amplitude changes. This amplitude-based analysis can be especially precise when used for mapping gas-generated bright spots in tertiary rocks. Thus, in many stratigraphic plays, resolution in the strict sense is not an issue. For these plays, detection, not resolution, is the problem.

Vertical resolution is a concern when discontinuities are inferred along a reflection horizon because of faults. Figure 8-14 shows a series of faults with vertical throws that are equal to 1, 1/2, 1/4, 1/8, and 1/16th of the dominant wavelength. It is when the throw is equal to or greater than one-fourth of the dominant wavelength that the presence of the fault can be inferred easily. Perhaps a smaller throw can be inferred by using diffractions from faults along the reflection horizon, provided the noise level is low in the data.

Clearly the ability to resolve or detect small targets can be increased by increasing the dominant frequency of the stacked data. The dominant frequency of a stacked section from a given area is governed by the physical properties of the subsurface, processing quality, and recording parameters. Since we cannot control the subsurface properties, the high-frequency signal level can only be influenced by the effort put into recording and processing.

The emphasis in recording should be to preserve high frequencies and suppress noise. The sampling rate and antialiasing filters should be adequate to record the desired frequencies. Receiver arrays should be small enough to prevent significant loss of high-frequency signal due to intragroup moveout and statics. However, the arrays should not be too small, since small arrays are not as effective at suppressing random, high-frequency noise (wind noise) as large arrays. Finally, the source effort should be high enough to provide adequate signal level relative to noise level within the desired frequency band. Unless the signal-to-noise ratio of the field data is above some minimal level, say 0.25, processing algorithms have difficulty in recovering the signal. The signal has to be detectable before it can be enhanced.

The emphasis in processing should be to preserve and display the high-frequency signal present in the input data. Filters with good high-frequency response should be used for interpolation processes such as NMO removal, datum and statics corrections, and multiplex skew corrections. Extra care should be taken to ensure that small-scale residual moveout or statics, which might cause loss of high-frequency signal during stacking, are removed before stack. Nonsurface-consistent alignment programs (often called trim statics programs) sometimes are used for this purpose. Finally, care must be taken to ensure that all the high-frequency signal is displayed on the final stack. Poststack deconvolution is a useful tool for this purpose.

8.3.2 Lateral Resolution

Lateral resolution refers to how close two reflecting points can be situated horizontally, yet be recognized as two separate points rather than one. Consider the spherical wavefront that impinges on the horizontal planar reflector AA′ in Figure 8-15. This reflector can be visualized as a continuum of point diffractors. For a coincident

Table 8-2. Threshold for lateral resolution (first Fresnel zone).

$r = (v/2)\,(t/f)^{1/2}$

t (s)	v (m)	f (Hz)	r (m)
1	2000	50	141
2	3000	40	335
3	4000	30	632
4	5000	20	1118

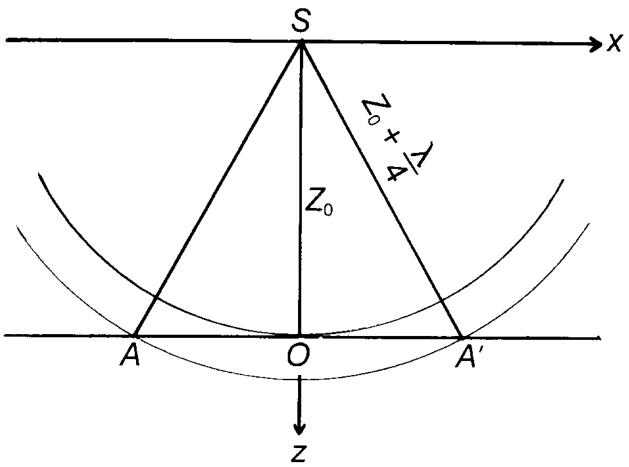


FIG. 8-15. Fresnel zone AA′ in *(x,z)* space.

source and receiver at the earth's surface (location S), the energy from the subsurface point (O) arrives at $t_0 = 2z_0/v$. Now let the incident wavefront advance in depth by the amount $\lambda/4$. Energy from subsurface location A, or A′, will reach the receiver at time $t_1 = 2(z_0 + \lambda/4)/v$. The energy from all the points within the reflecting disk with radius OA′ will arrive sometime between t_0 and t_1. The total energy arriving within the time interval $(t_1 - t_0)$, which equals half the dominant period $(T/2)$, interferes constructively. The reflecting disk AA′ is called a half-wavelength Fresnel zone (Hilterman, 1982) or the first Fresnel zone (Sheriff, 1984). Two reflecting points that fall within this zone generally are considered indistinguishable as observed from the earth's surface.

Since the Fresnel zone depends on wavelength, it also depends on frequency. For example, if the seismic signal riding along the wavefront is relatively high frequency, then the Fresnel zone is relatively narrow. The smaller the Fresnel zones, the easier it is to differentiate between two reflecting points. Hence, the Fresnel zone width is a measure of lateral resolution. Besides frequency, lateral resolution also depends on velocity and the depth of the reflecting interface (the radius of the wavefront) (Exercise 8.1):

$$r \approx (z\lambda/2)^{1/2} = (v/2)(t/f)^{1/2} \cdot \qquad (8.2)$$

Table 8-2 shows the Fresnel zone radius, where r = OA′ in Figure 8-15 for a range of frequency and velocity combinations at various depths ($t = 2z/v$).

From Table 8-2, note that the shallower the event (and the higher the dominant frequency), the smaller the Fresnel zone. Since the Fresnel zone generally increases with depth, spatial resolution also deteriorates with depth.

Figure 8-16 shows reflections from four interfaces, each with four nonreflecting segments. The actual sizes of these segments are indicated by the solid bars on top. On the seismic section, the reflections appear to be continuous across some of these segments. This is because the size of some of the nonreflecting segments is much less than the width of the Fresnel zone; they are beyond the lateral resolution limit.

Spatial resolution is better understood in terms of diffractions. Note that in Figure 8-16, the diffraction energy is smeared across the nonreflecting segments on the deeper reflectors. Since migration is the process that collapses diffractions, it is reasonable to think that migration increases spatial resolution. Remember that migration can be achieved by downward continuation of receivers from the surface to the reflecting horizons. As a result of downward continuation, the observation points get closer and closer to the reflection points and, therefore, the Fresnel zone gets smaller and smaller. A smaller Fresnel zone means a higher spatial resolution [equation (8.2)].

Migration tends to collapse the Fresnel zone to approximately the dominant wavelength [equation (8.1)] (Stolt and Benson, 1986). Therefore, we anticipate that migration will not resolve the horizontal limits of some of the nonreflecting segments along the deeper reflectors in Figure 8-16. Tables 8-1 and 8-2 can be used to estimate the potential resolution improvement that may result from migration. Unless 3-D migration (Section 6.5) is performed, actual resolution will be less than that indicated. Two-dimensional migration only shortens the Fresnel zone in the direction parallel to the line orientation. Resolution in the perpendicular direction is not affected by 2-D migration.

Figure 8-17 indicates how vertical and lateral resolution problems are interrelated. We want to determine the edge of the pinchout. The basis of the pinchout model is a wedge of material represented at a given midpoint location by a two-term reflectivity sequence, one term associated with the top and one with the bottom of the wedge. The true thickness of the wedge at various locations is indicated on top of Figure 8-17a. The velocity within the wedge is 2500 m/s.

We first consider the reflectivity sequence with two

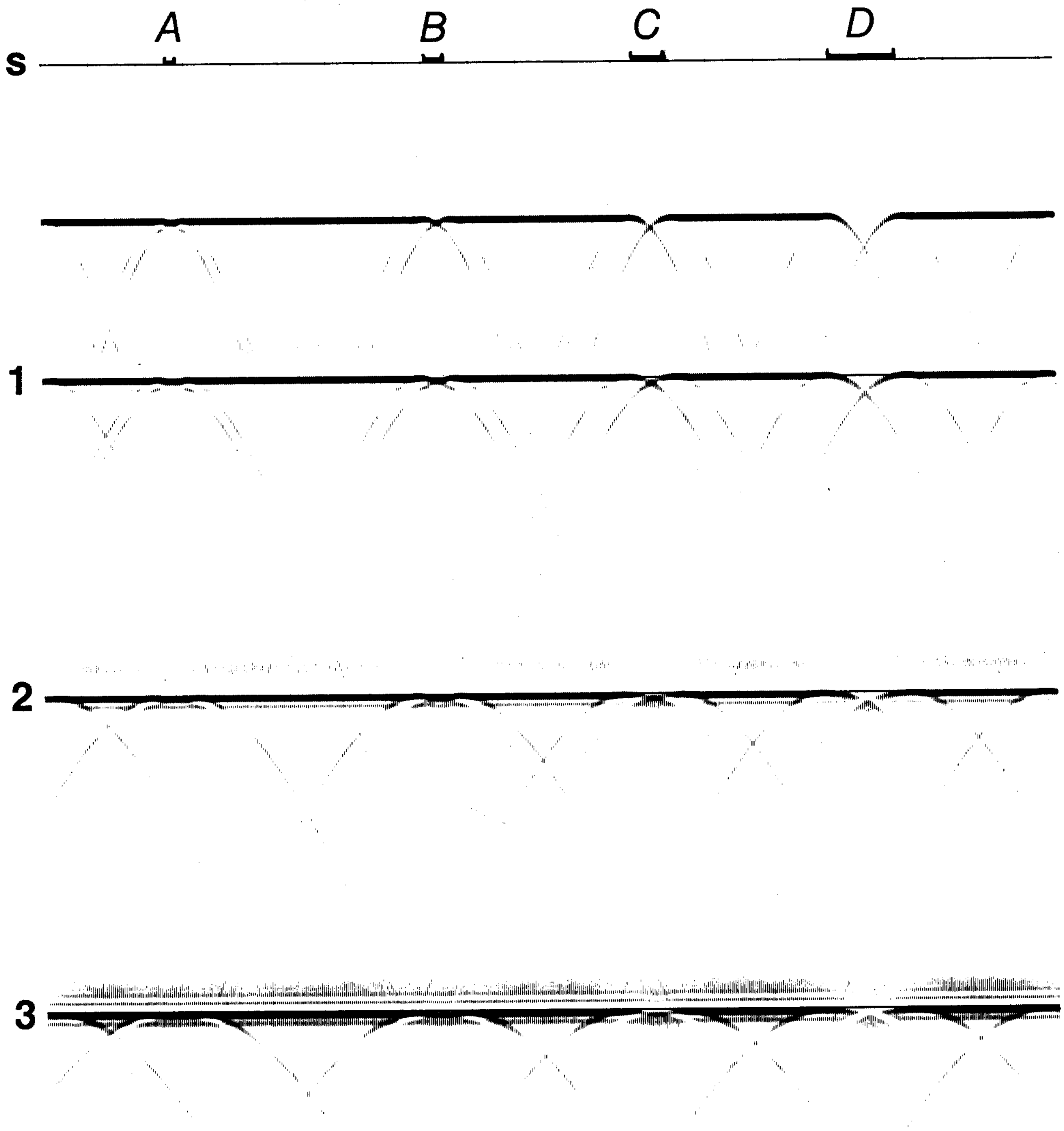


FIG. 8-16. A constant-velocity zero-offset section from an earth model that consists of four reflectors, each with four nonreflecting segments A, B, C, and D. Lateral resolution is governed by the size of the Fresnel zone. The lateral extent of each gap is indicated by the solid bars on top. Note that A is hardly recognizable on any of the four horizons; B can be inferred on the shallow horizon at 0.5 s; C is difficult to infer after 2 s, while D is recognizable at all depths. All of these observations depend on noise level and how easily the diffractions can be recognized.

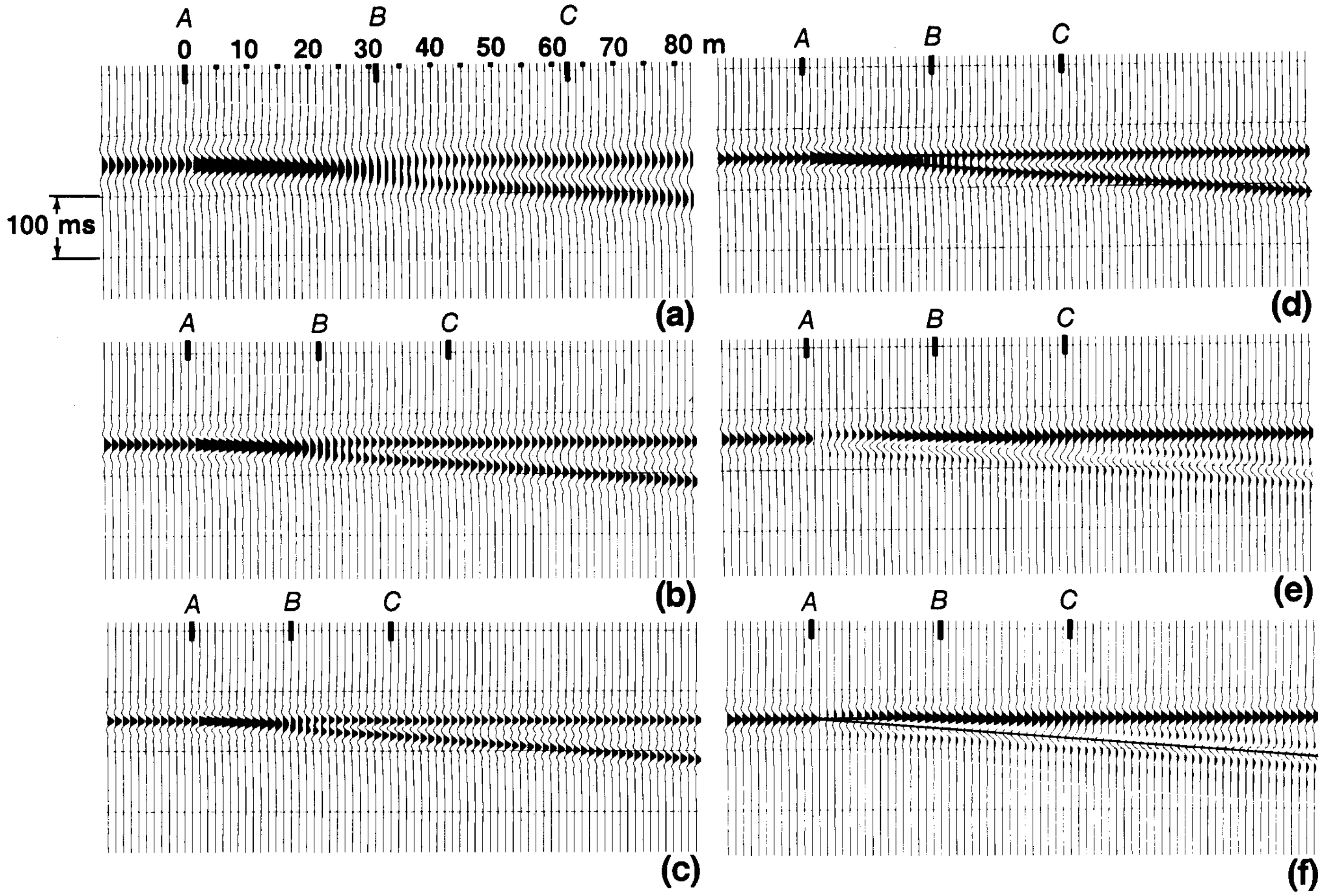


FIG. 8-17. (a) The result of convolving a zero-phase wavelet of 20-Hz dominant frequency with a wedge reflectivity model. The reflection coefficients associated with the top and bottom of the wedge are of equal amplitude and identical polarity. The true edge of the wedge is beneath location A and the true thickness of the wedge is indicated by the numbers on top; (b) same as (a) except the dominant frequency of the wavelet is 30 Hz; (c) same as (a) except the dominant frequency of the wavelet is 40 Hz; (d) same as (b) with the actual geometry of the wedge superimposed on the seismic response; (e) same as (b) except the reflection coefficients from the top and bottom of the wedge have opposite polarity; (f) same as (e) with the actual geometry of the wedge superimposed on the seismic response.

spikes of equal amplitude and identical polarity. The vertical-incidence seismic response (Figure 8-17a) is obtained by convolving the sequences with a zero-phase wavelet with 20-Hz dominant frequency. (The zero-phase response simplifies event tracking from the top and bottom of the wedge.) Based on this response, the edge of the wedge can be inferred as left of location B, where the waveform reduces to a single wavelet (Figure 8-17a). From the resolution threshold criterion, the smallest thickness that can be resolved is (2500 m/s)/(4 · 20 Hz) = 31.25 m. Figures 8-17a, b, and c show the same pinchout modeled using three different zero-phase wavelets with increasing dominant frequency (20, 30, and 40 Hz). Separation between the true location of pinchout A and the position of the minimum resolvable wedge thickness B decreases with increasing wavelet bandwidth.

While the resolution threshold criterion allows us to say only that the thickness of the wedge is less than 31.25 m left of B, an amplitude-based criterion can provide substantially more accurate location for the edge of the wedge. Again, refer to Figure 8-17a and observe the sudden change in amplitude at A where the true edge of the pinchout is located. Hence, the edge still can be reliably detected, even though it may not be resolved, provided the signal-to-noise (S/N) ratio is favorable. Assuming that the relative size of the top and bottom reflection coefficients is known, amplitudes also can be used to estimate the wedge thickness between B and A.

Figures 8-17a, b, and c show an apparent lateral variation in layer thickness. To see the difference between the true thickness and the apparent thickness (peak-to-peak time), refer to Figure 8-17d. This figure shows the data of Figure 8-17b with the actual geometry of the wedge superimposed on the seismic response. Since the composite wavelet has only one positive peak, the apparent thickness between locations A and B is nearly zero. At location B, the composite wavelet has a flat top. Immediately to the right of location B, the flat top disappears and the composite wavelet splits. The flat-top character can be identified as the limit of vertical resolution (Ricker, 1953). A short distance to the right of the point at which the composite wavelet first splits into two peaks, the apparent thickness becomes equal to the true thickness. This thickness is called the tuning thickness and is equal to peak-to-trough separation (one half the dominant period) of the convolving wavelet (Kallweit and Wood, 1982). Beyond the point of tuning thickness, note the apparent thickening of the layer between locations B and C. To the right of location C, the apparent and true thicknesses become equal.

Besides apparent thickness, the maximum absolute amplitude of the composite wavelet along the pinchout also changes (Kallweit and Wood, 1982). To the left of location A in Figure 8-17b, note the single isolated zero-phase wavelet. Immediately to the right of location A, the response of the two closely spaced spikes with identical polarity results in the maximum absolute amplitude. This amplitude gradually decreases to a minimum exactly at the tuning thickness. It then increases and reaches the amplitude value of the original single wavelet to the right of location C.

Maximum amplitude and apparent thickness change in reverse when the reflectivity model consists of reflection coefficients with equal amplitude and opposite polarity (Figure 8-17e). The composite waveform resulting from this reflectivity model is discussed by Widess (1973). Two spikes of opposite polarity with a small separation between them act as a derivative operator. When applied to a zero-phase wavelet, this operator causes a 90-degree phase shift. This phase shift can be seen in Figure 8-17e on the wavelets between locations A and B. Widess (1973) observed that the composite wavelet within this zone basically retains its shape while its amplitude changes.

Figure 8-17f shows data of Figure 8-17e with the actual geometry of the wedge superimposed on the seismic response. Note that the wedge appears thicker than it actually is between locations A and B. Also note the apparent thinning of the layer between locations B and C. Beyond location C, the apparent and true thicknesses become equal. Immediately to the right of location A in Figure 8-17e, the response of the two closely spaced spikes with opposite polarity results in cancellation of the amplitudes. The largest absolute amplitude of the composite wavelet gradually increases to a maximum immediately to the right of location B. It gradually decreases and reaches the amplitude value of the original single wavelet to the right of location C.

From the above discussion, we see that peak-to-peak time measurements and amplitude information can aid in detecting pinchouts that may otherwise be unresolvable. If the size of the reflection coefficients were known, then the amplitudes could be used to map the thickness below the resolution limit.

Nevertheless, the reliability of the analysis depends to some extent on the S/N ratio. Finally, the deceptive changes in amplitude and apparent thickness must be noted during mapping of the top and bottom of the pinchout.

8.4 SEISMIC MODELING

Seismic modeling involves generating traveltimes and amplitudes of seismic waves propagating through a specified subsurface reflectivity model that is associated with a specified velocity-depth model. It can be done either in a computer or in a laboratory. The most successful example of the latter is the model tank of the Seismic Acoustic Laboratory at the University of Houston. Throughout this book, we use seismic modeling to examine the critical parameters associated with deconvolution, velocity analysis, and migration. Examples from a few of the many available modeling techniques are provided.

In theory, any migration algorithm can be driven in the opposite direction to perform modeling. In particular, we can think of migration and modeling as extrapolations in depth and time, respectively:

$$\begin{array}{ccc} & \text{Migration} & \\ & \xrightarrow{\hspace{2cm}} & \\ \text{Observations} & & \text{Earth} \\ z = 0 & & t = 0 \\ \text{At All } t & & \text{At All } z \\ & \xleftarrow{\hspace{2cm}} & \\ & \text{Modeling} & \end{array}$$

The input to migration is the wave field $P(x,z = 0,t)$ recorded at surface $z = 0$ along spatial axis x. Consider $P(x,z = 0,t)$ as a zero-offset wave field for which the stacked section is a good substitute. The migrated section $P(x,z,t = 0)$ is the earth's cross section along the line given at all depth and imaged at $t = 0$. To compute the migrated section, we extrapolate the surface wave field in depth and apply the imaging principle. On the other hand, the input to modeling is the earth's reflectivity cross section $P(x,z,t = 0)$. The output of modeling is the simulation of the corresponding zero-offset section $P(x,z = 0,t)$, as if it were recorded at the earth's surface $z = 0$. The wave equation is solved in both migration and modeling.

Just as there are several approaches to solving the wave equation for migration, there also are several types of modeling techniques. There are modeling techniques based on the Kirchhoff integral (Hilterman, 1970), finite-difference (Kelly et al., 1976), and *f-k* domain (Sherwood et al., 1983) solutions to the wave equation. The algorithms based on the scalar (acoustic) wave equation, which describes *P*-wave propagation, are suitable for structural modeling in which amplitudes are not as important as traveltimes. The algorithms based on the elastic wave equation, which describes both *P*- and *S*-wave

8.1 INTRODUCTION

This chapter discusses a variety of essentially unrelated topics: multiple attenuation, seismic resolution, seismic modeling, synthetic sonic logs, instantaneous attributes, vertical seismic profiling (VSP), and 2-D surface data processing.

Multiple attenuation is the only topic included in a conventional processing flow. Sections 2.7.5 and 7.5 discussed multiple attenuation based on prediction. Section 8.2 examines multiple attenuation based on velocity discrimination between primaries and multiples in both the *f-k* and *t-x* domains.

The remaining topics are interpretational aids that involve manipulation and simulation of seismic data. Seismic resolution is discussed in Section 8.3. Resolution is the ability to separate two events that are very close together. There are two aspects of seismic resolution: vertical (or temporal) and lateral (or spatial). Seismic resolution becomes especially important in mapping small structural features, such as subtle sealing faults, and in delineating thin stratigraphic features that may have limited areal extent.

Seismic forward modeling (Section 8.4) involves generating the traveltime response of a subsurface reflectivity model associated with a velocity-depth model. Forward modeling has a variety of applications. An understanding of the seismic responses of subsurface structural and stratigraphic features is made possible by forward modeling. Modeling also is used to generate test data for use in evaluating processing algorithms. Recording parameters, such as group interval and spread length, sometimes are chosen, at least partially, on the basis of forward models. Forward modeling also is useful for determining whether the reflection response of an interpreted geologic model is consistent with the CMP stacked sections used in the interpretation.

The synthetic sonic log (Section 8.5) is a simple 1-D inversion of reflection data. While the middle-and high-frequency components of the synthetic sonic log are derived from the seismic trace, the low-frequency trend must be obtained from separate sources of information such as conventional velocity analysis or actual sonic-log data.

Instantaneous attributes (Section 8.6) can help emphasize the continuity of reflectors and delineate seismic depositional sequence boundaries. When displayed in color, reflectivity strength, instantaneous phase, and instantaneous frequency are useful tools in stratigraphic exploration.

Vertical seismic profiling is reviewed in Section 8.7. The basic processing sequence and the use of VSP data are outlined. Finally, Section 8.8 reviews 2-D map data processing.

8.2 MULTIPLE SUPPRESSION

Section 7.5 discussed multiple attenuation based on the periodicity of multiples in the slant-stack domain. The two techniques presented now are based on the moveout difference between primaries and multiples. The CMP gathers in Figure 8-1a clearly illustrate this moveout difference. A primary *p* typically has less moveout than a multiple *m*. From the velocity spectrum in Figure 8-1b, note the difference between the velocity trends associated with primaries VP and multiples VM1 and VM2. The VM1 and VM2 velocity functions represent the water-bottom and peg-leg multiples, respectively. If NMO correction is applied using the primary velocities, as is normally done to generate final stacks, then the primaries are aligned while the multiples are undercorrected (Figure 8-1c). This suggests that CMP stacking itself is a viable method of multiple suppression. The CMP stack derived from the gathers in Figure 8-1c is shown in Figure 8-1d.

The synthetic CMP gather in Figure 8-2c contains five primaries, including the water-bottom reflection *W* and the multiples associated with the water-bottom reflection. Note that the water-bottom multiples have higher amplitudes than the primaries. The velocity spectrum shows a significant separation between the velocity functions for multiples VM and primaries VP. Stacking with the primary velocity function should, to a large extent, discriminate against the multiples and result in a section that contains essentially the primary energy as shown in Figure 8-3. The stack trace is repeated to better examine the amplitudes.

Stacking far offsets work to suppress multiples. However, stacking near offsets works against multiple suppression, since the moveout difference between the primaries and multiples is negligibly small on those offsets (see, for example, Figure 8-1c). The simplest way around this problem is to apply an inside mute to the CMP gathers before stacking. Another problem then emerges: the outside mute. The severity of this mute governs the amount of far-offset data left at early times for velocity discrimination (Figure 8-1c). If there is a severe multiple problem, an effort must be made to preserve the maximum amount of far-offset data associated with target events. The stacked section with inside mute applied is shown in Figure 8-4a. When compared with Figure 8-1d, note that the deeper peg-leg multiple below 4 s has been further attenuated somewhat. The difference between the conventional CMP stack (Figure 8-1d) and the inside mute stack (Figure 8-4a) indicates the amount of energy, mostly multiples, that was removed by the inside mute (Figure 8-4b). Sometimes more sophisticated muting schemes in which weights (between 0 and 1) are assigned to each offset (the smaller weights assigned to the near offsets) during stacking can produce better results than those shown here.

8.2.1 Velocity Discrimination in the *f-k* Domain

Figure 8-5a shows the synthetic CMP gather (top) from Figure 8-2c and its 2-D amplitude spectrum (bottom). The primary and multiple energy can be separated into two different quadrants in the *f-k* plane. This is achieved by NMO correcting the gather using a velocity function

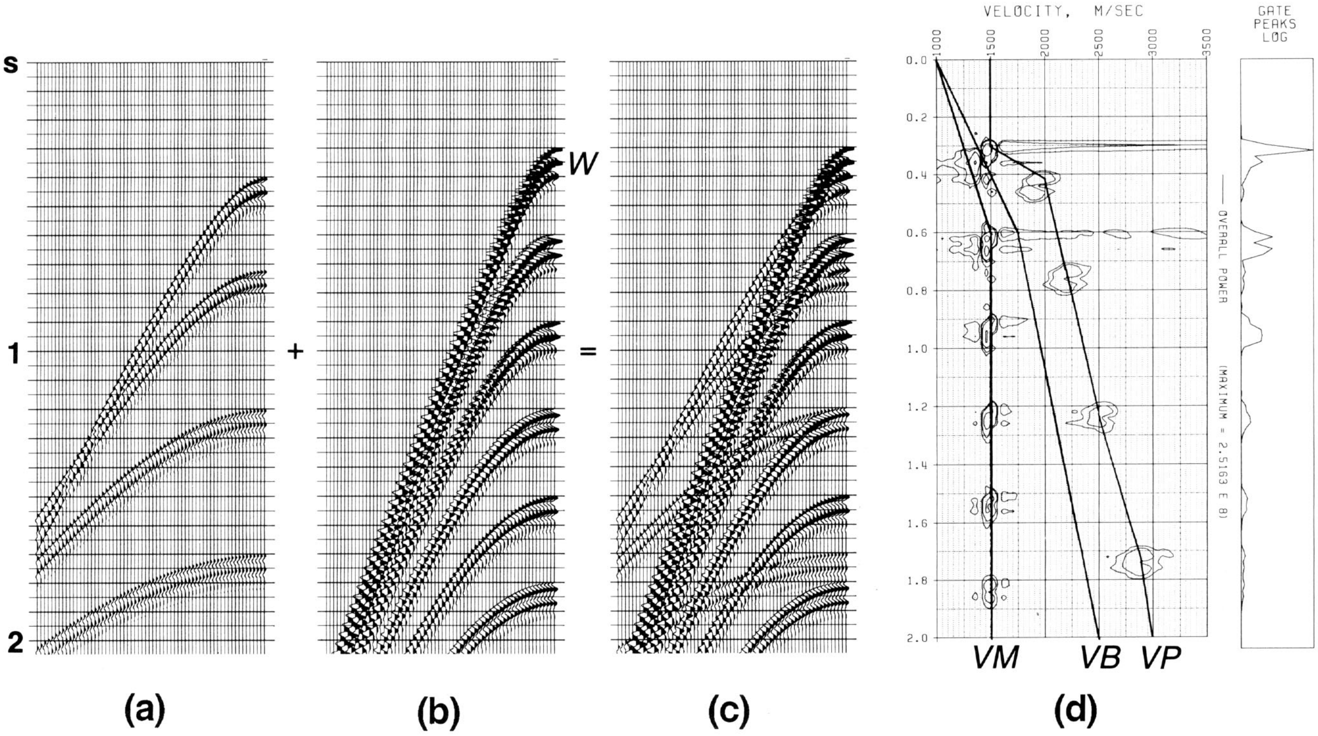
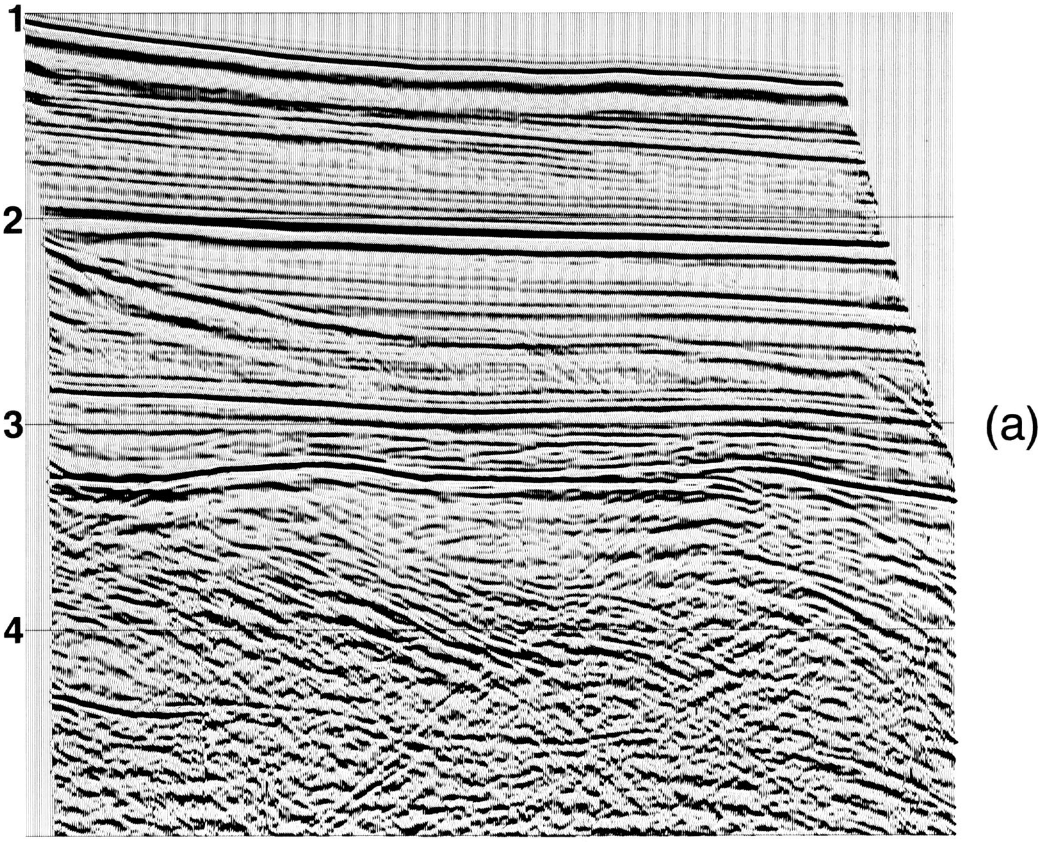
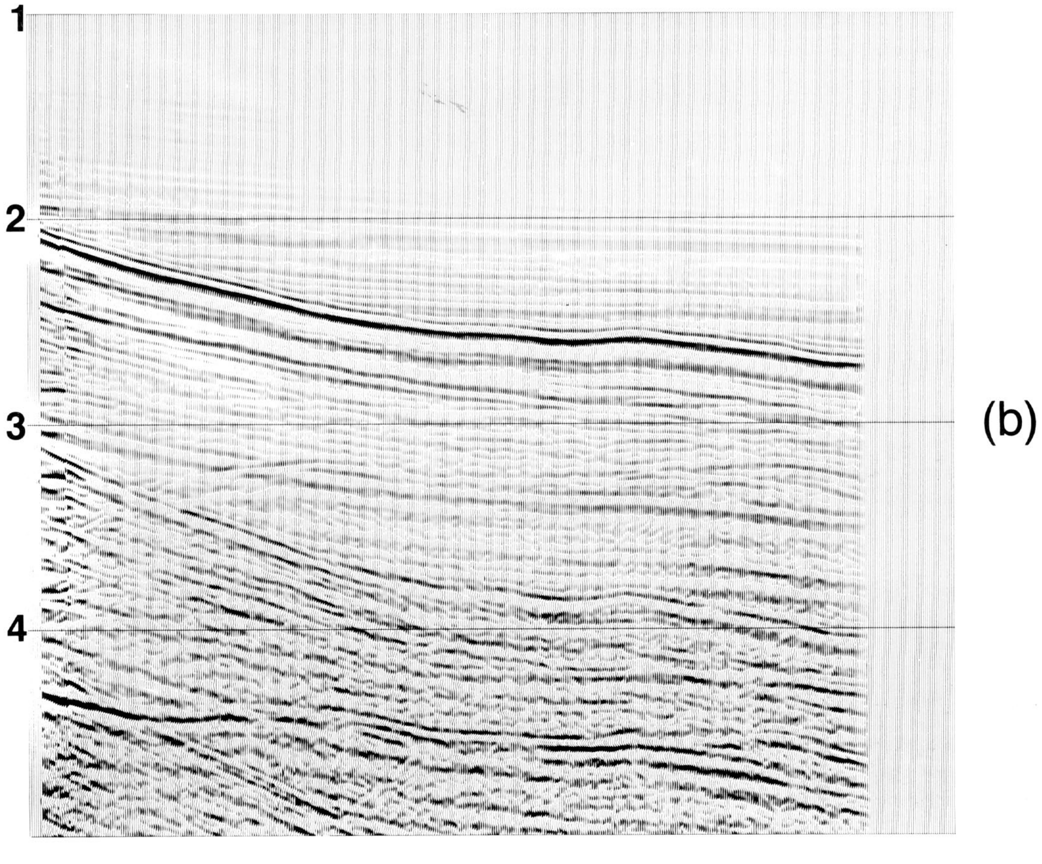


FIG. 8-2. Synthetic CMP gathers containing (a) primaries, (b) water-bottom multiples, (c) superposition of (a) and (b). (d) The velocity spectrum derived from (c). Here, W = water-bottom primary, VM = velocity function for multiples, VP = velocity function for primaries, VB = a velocity function between VM and VP used in generating Figure 8-5b.

FIG. 8-1. (a) Three CMP gathers with strong multiples; (b) velocity analysis at CMP 186, where VP = primary velocity trend, VM1 = slow (water-bottom) multiples, and VM2 = fast (peg-leg) multiples. (VB is the velocity function used in generating Figure 8-8a.) For reference, the CMP gather is displayed next to the velocity spectrum. (c) The same CMP gathers as in (a) after NMO correction using the primary velocities. (d) CMP stack using the gathers as in (c). (Data courtesy Petro-Canada Resources.)


1
2
3
4
(a)


1
2
3
4
(b)

dure, apply NMO correction using the primary velocity function VP and stack (Figures 8-6c and d). Figure 8-7 is a flowchart for *f-k* multiple attenuation.

Now consider the field data example shown in Figure 8-1. The moveout difference between the primaries and multiples is apparent in the CMP gathers (Figure 8-1a) and the velocity spectrum (Figure 8-1b). To the left of solid line VP, all peaks are associated with water-bottom and peg-leg multiples. Apply a moveout correction using a velocity function VB that is somewhere between the primary and multiple velocities. As a result, the multiples are undercorrected and the primaries are overcorrected as shown in Figure 8-8a. Now consider the moveout of primaries and multiples in the *f-k* domain. The primary and multiple indicated in Figure 8-1a map to the same quadrant (say positive quadrant) in the *f-k* domain before moveout correction. The same events after moveout correction using an intermediate velocity map into two different quadrants; in particular, the multiple maps into the positive quadrant and the primary maps into the negative quadrant. Thus, by zeroing one quadrant in which the multiples are clustered, the primaries can be enhanced.

A CMP gather and the associated velocity spectrum following *f-k* multiple suppression are shown in Figure 8-8b. When compared with Figure 8-1b, Figure 8-8b

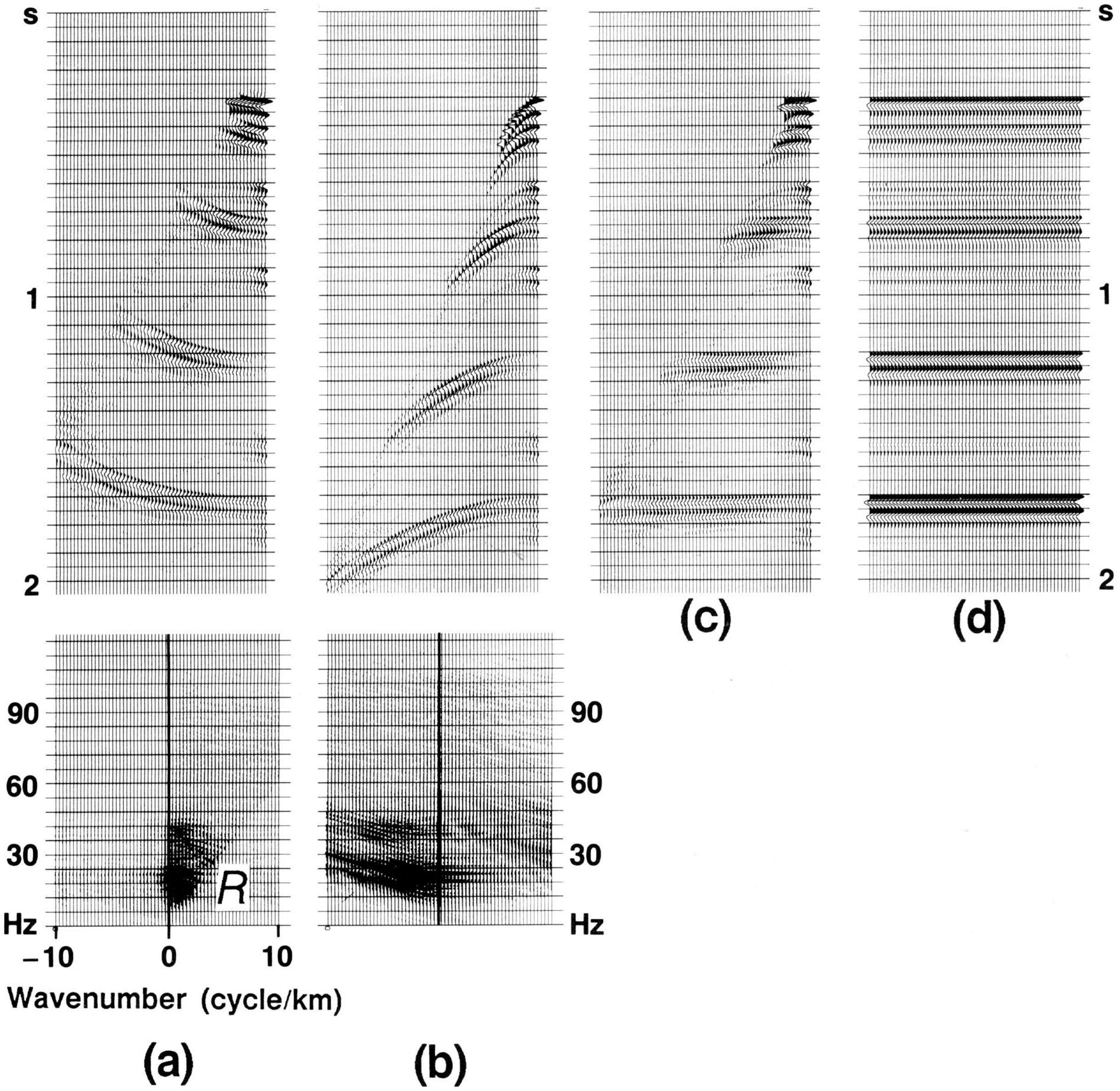
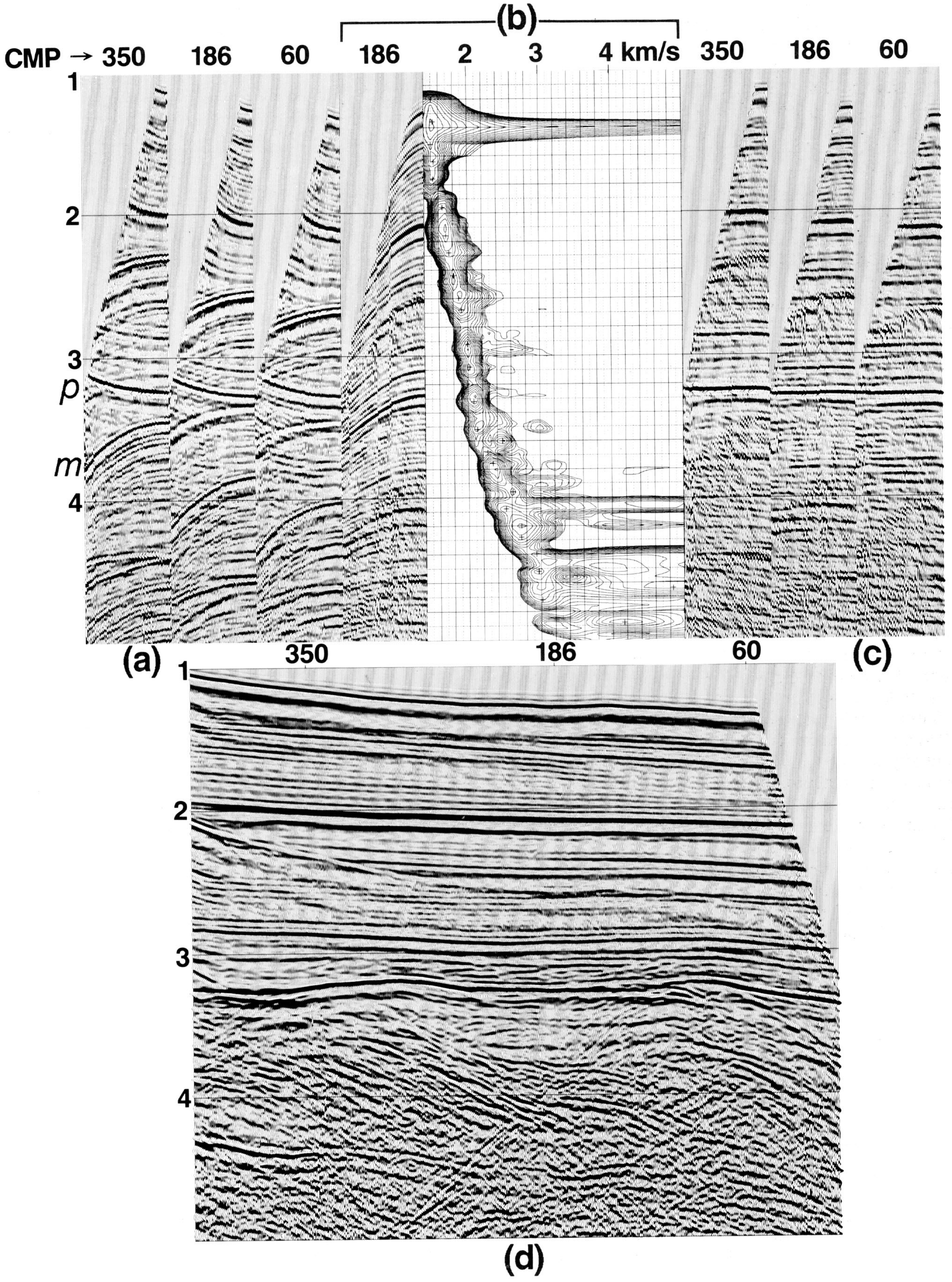
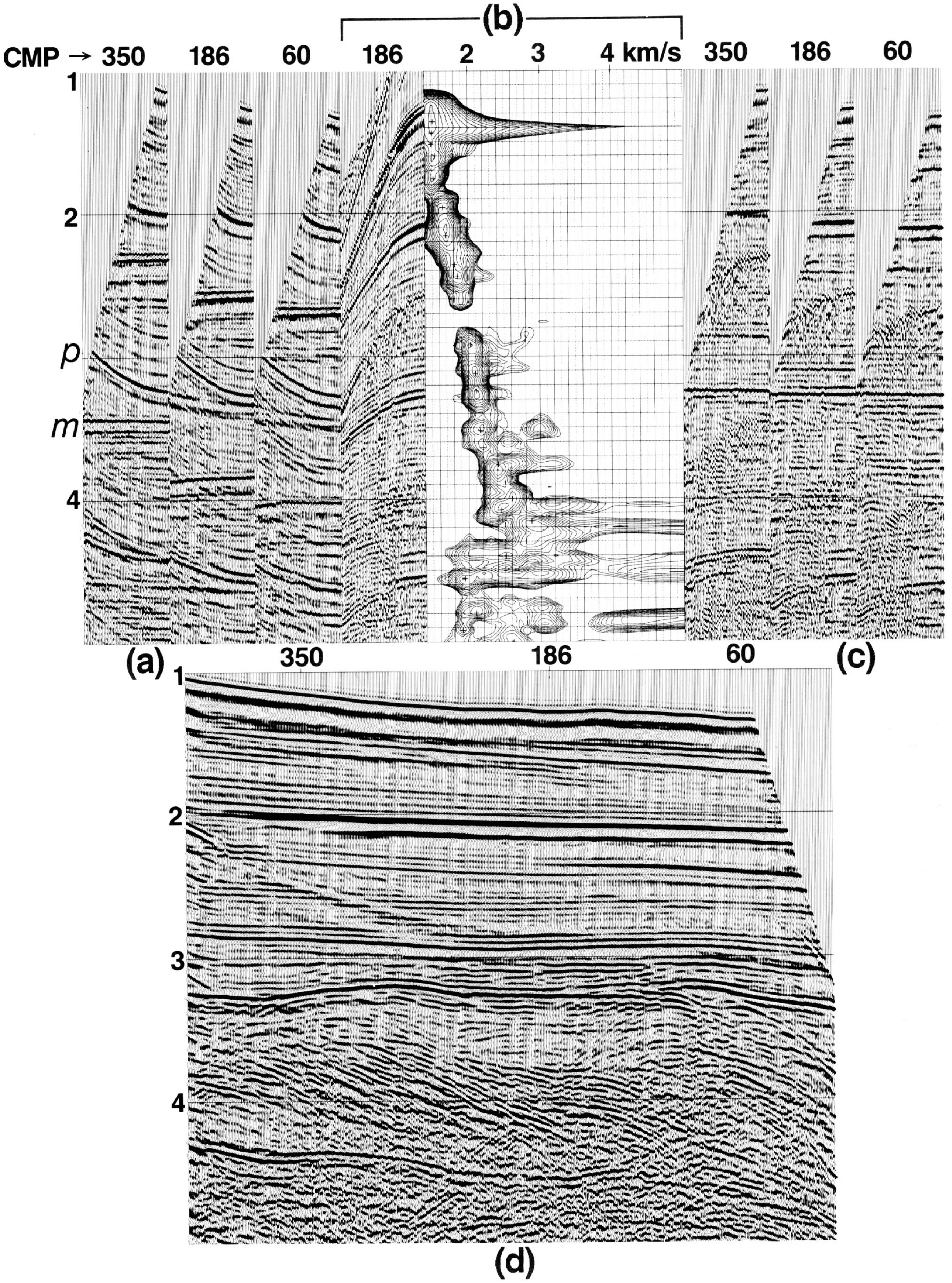


FIG. 8-6. (a) The same figure as in Figure 8-5c, except that in addition to zeroing the left quadrant, a portion of the right quadrant in the *f-k* spectrum (denoted by R) also is zeroed to suppress aliased energy: (b) the result of applying inverse NMO to (a) using velocity function VB in Figure 8-2d; (c) the result of applying NMO correction to (b) using primary velocity function VP in Figure 8-2d; (d) stack of (c) repeated to emphasize the strong events. The bottom panels show the corresponding *f-k* spectra.


(b)
CMP → 350
186
60
186
2
3
4 km/s
350
186
60
1
2
3
p
m
4
(a)
350
186
60
(c)
1
2
3
4
(d)


(b)
CMP → 350
186
60
186
2
3
4 km/s
350
186
60
1
2
p
m
4
(a)
350
186
60
(c)
1
2
3
4
(d)

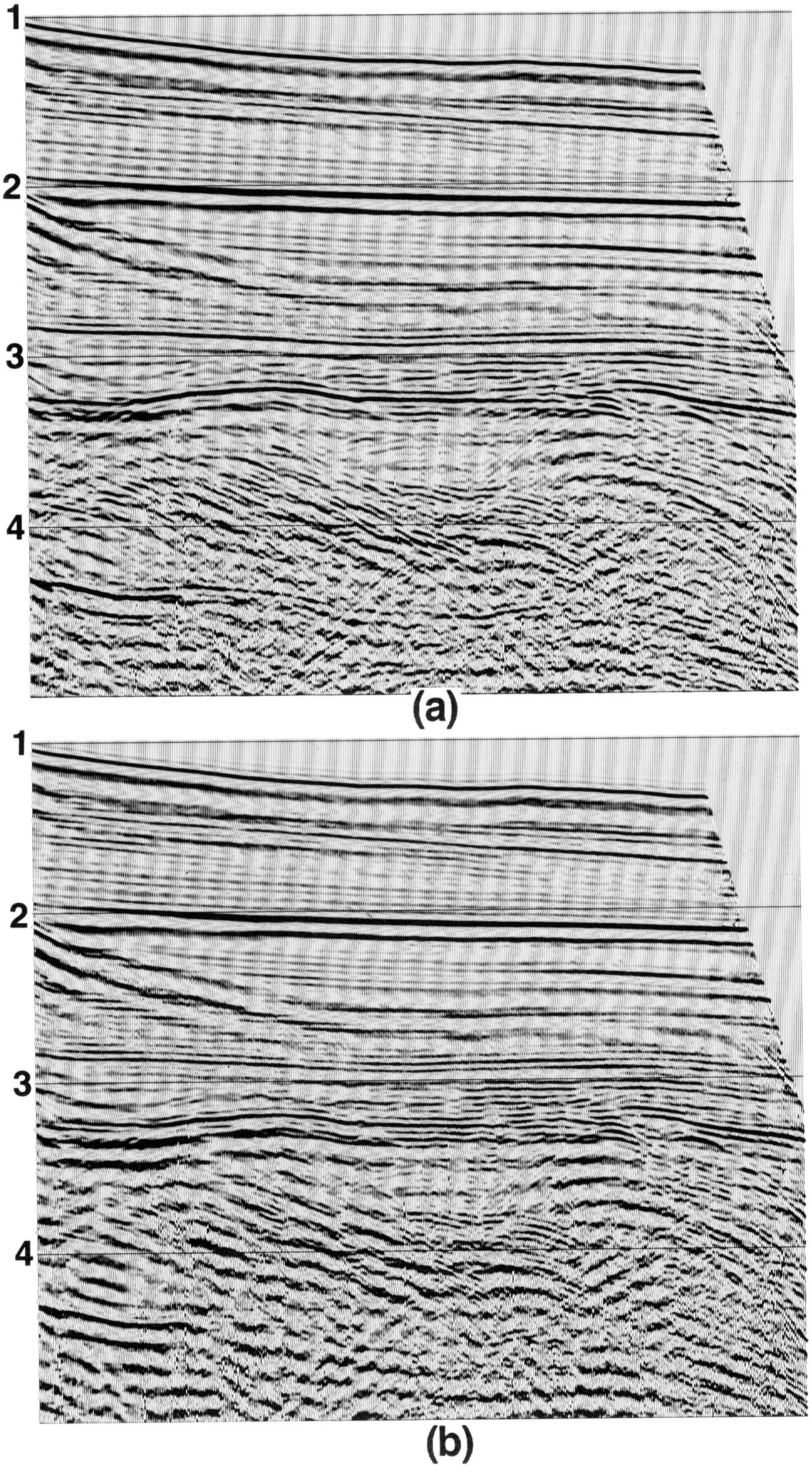


FIG. 8-12. The CMP stacks after t-x multiple attenuation, which was implemented using filtered model traces. (a) First pass using multiple velocites VM1 and (b) second pass using multiple velocities VM2, as shown in Figure 8-1b. Compare these with Figures 8-10d and 8-11d. (Data courtesy Petro-Canada Resources.)

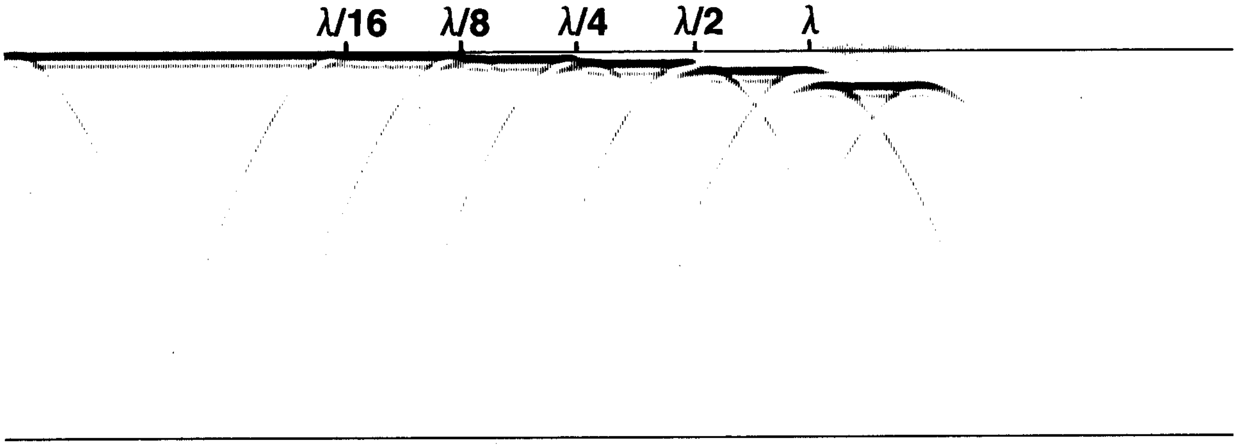


FIG. 8-14. Faults with different amounts of vertical throws expressed in fractions of the dominant wavelength.

as low as 20 Hz, the thickness must be at least 62 m for it to be resolvable.

It is now appropriate to ask whether a thin stratigraphic unit must be resolved to be mapped. The answer is no. Resolution as defined here and in the geophysical literature implies that reflections from the top and bottom of a thin bed are seen as separate events or wavelet lobes. Using this definition, resolution does not consider amplitude effects. The thickness and areal extent of beds below the resolution limit often can be mapped on the basis of amplitude changes. This amplitude-based analysis can be especially precise when used for mapping gas-generated bright spots in tertiary rocks. Thus, in many stratigraphic plays, resolution in the strict sense is not an issue. For these plays, detection, not resolution, is the problem.

Vertical resolution is a concern when discontinuities are inferred along a reflection horizon because of faults. Figure 8-14 shows a series of faults with vertical throws that are equal to 1, 1/2, 1/4, 1/8, and 1/16th of the dominant wavelength. It is when the throw is equal to or greater than one-fourth of the dominant wavelength that the presence of the fault can be inferred easily. Perhaps a smaller throw can be inferred by using diffractions from faults along the reflection horizon, provided the noise level is low in the data.

Clearly the ability to resolve or detect small targets can be increased by increasing the dominant frequency of the stacked data. The dominant frequency of a stacked section from a given area is governed by the physical properties of the subsurface, processing quality, and recording parameters. Since we cannot control the subsurface properties, the high-frequency signal level can only be influenced by the effort put into recording and processing.

The emphasis in recording should be to preserve high frequencies and suppress noise. The sampling rate and antialiasing filters should be adequate to record the desired frequencies. Receiver arrays should be small enough to prevent significant loss of high-frequency signal due to intragroup moveout and statics. However, the arrays should not be too small, since small arrays are not as effective at suppressing random, high-frequency noise (wind noise) as large arrays. Finally, the source effort should be high enough to provide adequate signal level relative to noise level within the desired frequency band. Unless the signal-to-noise ratio of the field data is above some minimal level, say 0.25, processing algorithms have difficulty in recovering the signal. The signal has to be detectable before it can be enhanced.

The emphasis in processing should be to preserve and display the high-frequency signal present in the input data. Filters with good high-frequency response should be used for interpolation processes such as NMO removal, datum and statics corrections, and multiplex skew corrections. Extra care should be taken to ensure that small-scale residual moveout or statics, which might cause loss of high-frequency signal during stacking, are removed before stack. Nonsurface-consistent alignment programs (often called trim statics programs) sometimes are used for this purpose. Finally, care must be taken to ensure that all the high-frequency signal is displayed on the final stack. Poststack deconvolution is a useful tool for this purpose.

8.3.2 Lateral Resolution

Lateral resolution refers to how close two reflecting points can be situated horizontally, yet be recognized as two separate points rather than one. Consider the spherical wavefront that impinges on the horizontal planar reflector AA′ in Figure 8-15. This reflector can be visualized as a continuum of point diffractors. For a coincident

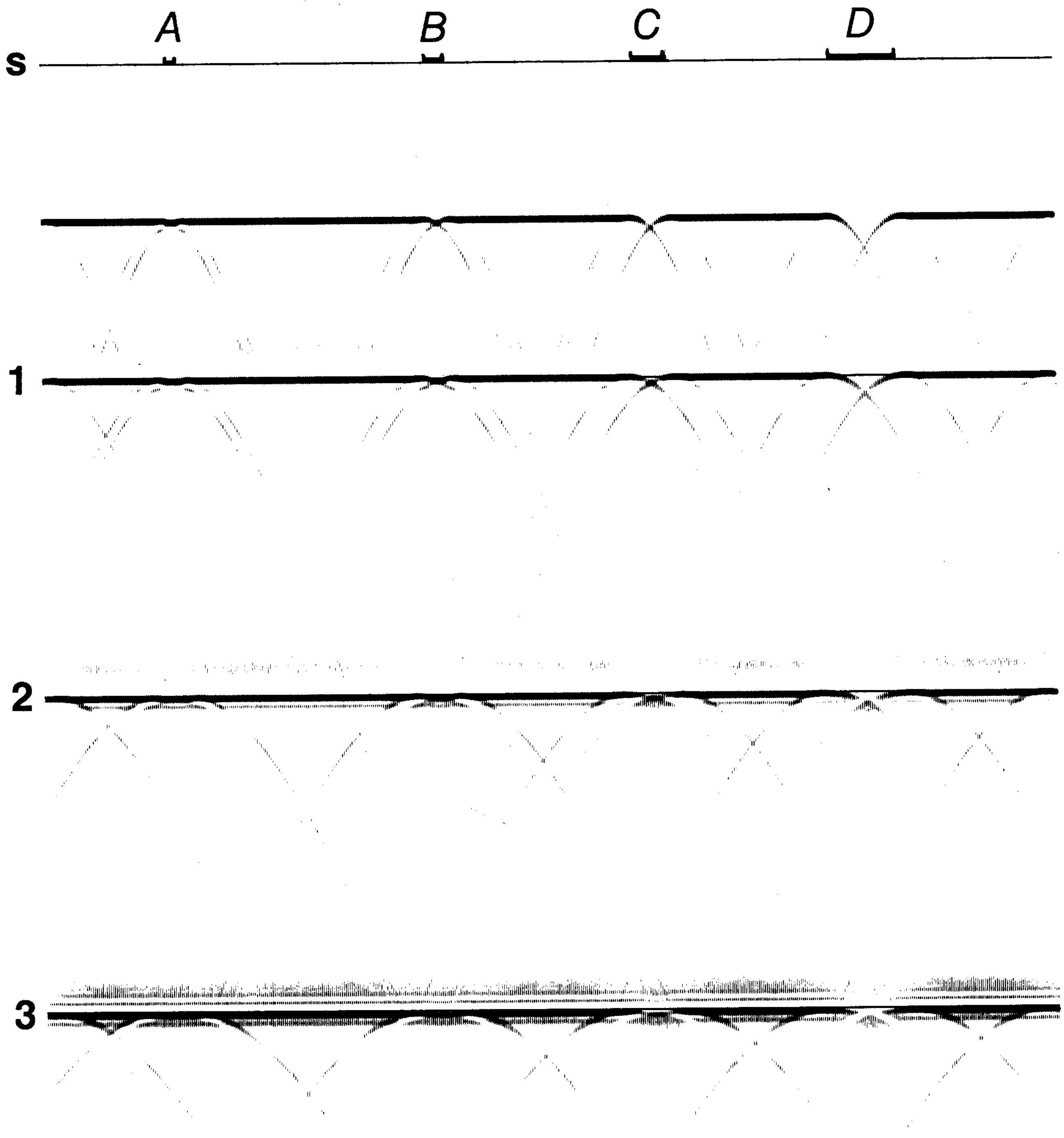


FIG. 8-16. A constant-velocity zero-offset section from an earth model that consists of four reflectors, each with four nonreflecting segments A, B, C, and D. Lateral resolution is governed by the size of the Fresnel zone. The lateral extent of each gap is indicated by the solid bars on top. Note that A is hardly recognizable on any of the four horizons; B can be inferred on the shallow horizon at 0.5 s; C is difficult to infer after 2 s, while D is recognizable at all depths. All of these observations depend on noise level and how easily the diffractions can be recognized.

location B, the composite wavelet has a flat top. Immediately to the right of location B, the flat top disappears and the composite wavelet splits. The flat-top character can be identified as the limit of vertical resolution (Ricker, 1953). A short distance to the right of the point at which the composite wavelet first splits into two peaks, the apparent thickness becomes equal to the true thickness. This thickness is called the tuning thickness and is equal to peak-to-trough separation (one half the dominant period) of the convolving wavelet (Kallweit and Wood, 1982). Beyond the point of tuning thickness, note the apparent thickening of the layer between locations B and C. To the right of location C, the apparent and true thicknesses become equal.

Besides apparent thickness, the maximum absolute amplitude of the composite wavelet along the pinchout also changes (Kallweit and Wood, 1982). To the left of location A in Figure 8-17b, note the single isolated zero-phase wavelet. Immediately to the right of location A, the response of the two closely spaced spikes with identical polarity results in the maximum absolute amplitude. This amplitude gradually decreases to a minimum exactly at the tuning thickness. It then increases and reaches the amplitude value of the original single wavelet to the right of location C.

Maximum amplitude and apparent thickness change in reverse when the reflectivity model consists of reflection coefficients with equal amplitude and opposite polarity (Figure 8-17e). The composite waveform resulting from this reflectivity model is discussed by Widess (1973). Two spikes of opposite polarity with a small separation between them act as a derivative operator. When applied to a zero-phase wavelet, this operator causes a 90-degree phase shift. This phase shift can be seen in Figure 8-17e on the wavelets between locations A and B. Widess (1973) observed that the composite wavelet within this zone basically retains its shape while its amplitude changes.

Figure 8-17f shows data of Figure 8-17e with the actual geometry of the wedge superimposed on the seismic response. Note that the wedge appears thicker than it actually is between locations A and B. Also note the apparent thinning of the layer between locations B and C. Beyond location C, the apparent and true thicknesses become equal. Immediately to the right of location A in Figure 8-17e, the response of the two closely spaced spikes with opposite polarity results in cancellation of the amplitudes. The largest absolute amplitude of the composite wavelet gradually increases to a maximum immediately to the right of location B. It gradually decreases and reaches the amplitude value of the original single wavelet to the right of location C.

From the above discussion, we see that peak-to-peak time measurements and amplitude information can aid in detecting pinchouts that may otherwise be unresolvable. If the size of the reflection coefficients were known, then the amplitudes could be used to map the thickness below the resolution limit.

Nevertheless, the reliability of the analysis depends to some extent on the S/N ratio. Finally, the deceptive changes in amplitude and apparent thickness must be noted during mapping of the top and bottom of the pinchout.

8.4 SEISMIC MODELING

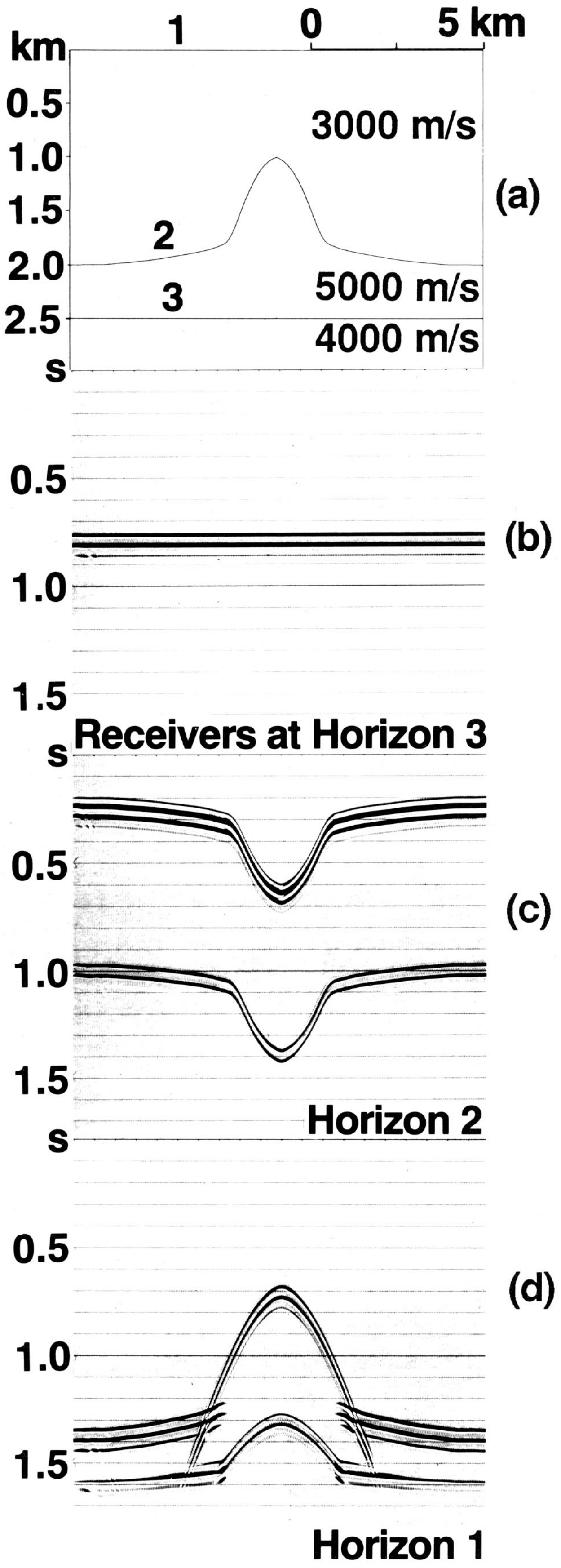
Seismic modeling involves generating traveltimes and amplitudes of seismic waves propagating through a specified subsurface reflectivity model that is associated with a specified velocity-depth model. It can be done either in a computer or in a laboratory. The most successful example of the latter is the model tank of the Seismic Acoustic Laboratory at the University of Houston. Throughout this book, we use seismic modeling to examine the critical parameters associated with deconvolution, velocity analysis, and migration. Examples from a few of the many available modeling techniques are provided.

In theory, any migration algorithm can be driven in the opposite direction to perform modeling. In particular, we can think of migration and modeling as extrapolations in depth and time, respectively:

$$\begin{tikzcd}
\begin{matrix}\text{Observations}\\ z=0\\ \text{At All } t\end{matrix} \arrow[r, shift left, "\text{Migration}"] & \begin{matrix}\text{Earth}\\ t=0\\ \text{At All } z\end{matrix} \arrow[l, shift left, "\text{Modeling}"]
\end{tikzcd}$$

The input to migration is the wave field $P(x,z = 0,t)$ recorded at surface $z = 0$ along spatial axis x. Consider $P(x,z = 0,t)$ as a zero-offset wave field for which the stacked section is a good substitute. The migrated section $P(x,z,t = 0)$ is the earth's cross section along the line given at all depth and imaged at $t = 0$. To compute the migrated section, we extrapolate the surface wave field in depth and apply the imaging principle. On the other hand, the input to modeling is the earth's reflectivity cross section $P(x,z,t = 0)$. The output of modeling is the simulation of the corresponding zero-offset section $P(x,z = 0,t)$, as if it were recorded at the earth's surface $z = 0$. The wave equation is solved in both migration and modeling.

Just as there are several approaches to solving the wave equation for migration, there also are several types of modeling techniques. There are modeling techniques based on the Kirchhoff integral (Hilterman, 1970), finite-difference (Kelly et al., 1976), and *f-k* domain (Sherwood et al., 1983) solutions to the wave equation. The algorithms based on the scalar (acoustic) wave equation, which describes *P*-wave propagation, are suitable for structural modeling in which amplitudes are not as important as traveltimes. The algorithms based on the elastic wave equation, which describes both *P*- and *S*-wave



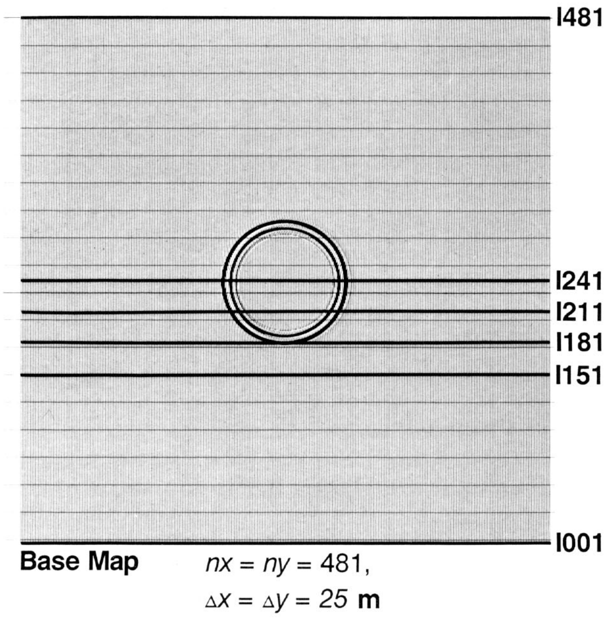


FIG. 8-19. Base map for a hypothetical 3-D survey. The 3-D velocity-depth and reflectivity model cross sections are shown in Figure 8-20 and the modeled 3-D zero-offset wave field cross sections are shown in Figure 8-21.

◂ FIG. 8-18. Wave-equation datuming (Section 5.3) used as a modeling tool. The zero-offset wave field at the surface (d) has been obtained by upward continuation [(b) and (c)] through the velocity-depth model (a).

propagation, are suitable for detailed stratigraphic modeling in which amplitudes are as important as traveltimes. Modeling based on one-way wave equations does not include multiples, while modeling based on full wave equations includes multiples in the solution.

Wave-equation datuming (Berryhill, 1979), which was described in Section 5.3, has applications in zero-offset and nonzero-offset modeling. In particular, the datuming approach can propagate a wave field from one irregular interface to another. Consider zero-offset modeling using the datuming technique of the velocity-depth model shown in Figure 8-18a. Horizons 2 and 3 are the top and base of a salt dome. Start with the receivers situated along horizon 3. The zero-offset section (Figure 8-18b) contains the reflection from the bottom of the velocity-depth model at z = 4000 m (not shown in Figure 8-18a). Take this wave field and extrapolate it to a new datum, horizon 2, using the salt velocity (5000 m/s). The zero-offset section (Figure 8-18c) contains the reflection (the deeper one) from the bottom of the velocity-depth model (z = 4000 m) and the reflection (the shallow one) from the base of the salt (horizon 3). Finally, extrapolate this wave field (Figure 8-18c) from horizon 2 to the surface (horizon 1 at z = 0) using the overburden velocity (3000 m/s) to get the 2-D zero-offset section in Figure 8-18d. This section contains reflections from both the top and base of the salt. (The reflection from the bottom of the model arrives after the latest time shown on this section.) Note the velocity pull-up along the reflection from the base of the salt dome. Proper imaging of the top of the salt dome can be achieved by time migration (Section 4.2), while proper imaging of the base of the salt requires depth migration (Section 5.2).

We now use an example to demonstrate a finite-difference modeling technique that is the reverse of the omega-x migration algorithm (Appendix C.3). The original 2-D implementation is by Kjartansson (1979) (see also, Kelamis and Kjartansson, 1985). The technique used for this example is a 3-D zero-offset omega-x modeling algorithm based on the splitting approach discussed in Appendix C.8. Consider a hypothetical 3-D survey with the base map shown in Figure 8-19. Figure 8-20 shows cross sections of the input 3-D velocity-depth (left column) and reflectivity (right column) models along the four selected in lines indicated on the base map. The base map in Figure 8-19 actually is a depth slice of the 3-D reflectivity model at z = 1700 m. Cross sections of the modeled 3-D zero-offset wave field along the same in lines are shown in Figure 8-21 (left column).

Figure 8-21 also shows 2-D modeling along two of the in lines (right column). Note that the 2-D and 3-D zero-offset wave fields are the same (except for the 2-D versus 3-D amplitude effects) along the center in line 1241. However, the 2-D and 3-D zero-offset wave fields along the off-center in line 1181 are quite different. What we actually record in the field is a section that contains sideswipe energy like the 3-D modeling result. However, the 2-D modeling result does not contain any sideswipe energy and implies a false time response of the subsurface reflectivity.

Modeling results from finite-difference schemes can be expected to exhibit artifacts similar to those encountered on finite-difference migrated sections (Exercise 8.9). Close inspection of Figure 8-21 reveals the expected dispersive noise (Sections 4.3.2 and 4.3.4). Boundary effects also are just as much a problem with modeling as they are for migration. A good modeling program needs to implement all the standard noise suppression features found in the better migration schemes. Besides absorbing boundaries (Claerbout, 1985), the most effective way to avoid boundary effects is to model a longer line or a larger area than actually is needed. In Figure 8-21, the lines were trimmed on both ends to disguise the boundary reflections and only the uncontaminated portions were retained.

Now consider some nonzero-offset modeling examples. A finite-difference technique for modeling acoustic and elastic wave fields is described by Kelly et al. (1976). Figure 8-22 shows an example of acoustic modeling. A seismic line is simulated over a 2-D complex structure (Figure 8-22a). Selected common shot and CMP gathers from this simulation clearly show the many complexities in the arrivals. Since this is a full (two-way) acoustic solution, the modeled gathers contain not only primaries, but also multiples. The zero-offset and stacked sections associated with this nonzero-offset data are shown in Figure 8-23. Note the broad traveltime responses associated with the tight imbricate structures in the velocity-depth model (Figure 8-22a).

As an example of nonzero-offset modeling, an application of wave-equation datuming is provided in Figure 8-24. The shot gathers in Figure 8-22b are computed to a flat datum level z = 0. Better simulation of the actual field conditions requires that the gathers be computed using an irregular topography. To do this, we can upward continue the shots and receivers to the new irregular datum represented by the topography shown in Figure 8-24a, then compute the shot gathers in Figure 8-24b. Compare Figures 8-22b and 8-22c with Figures 8-24b and 8-24c, and note the traveltime distortions.

We now examine some examples of elastic modeling. Sherwood et al. (1983) developed an f-k method for nonzero-offset modeling of elastic waves in a 2-D horizontally layered medium. Figure 8-25 shows five shot gathers derived from a $v(z)$ model beneath five different water layers. The water depths are 5, 10, 15, 20, and 50 m. Note the guided wave energy, which is especially prominent in gathers corresponding to water depths of 5, 10, and 15 m. These gathers contain all primaries, both P-waves and S-waves, as well as all possible multiples and converted modes. By examining such modeled data, we can better understand the nature of coherent noise (guided waves and multiples) in both land and marine environments.

A more interpretive application of the elastic modeling program by Sherwood et al. (1983) can be seen in Figure 8-26. The synthetic shot gather on the left in Figure 8-26a is from a clastic section with a shallow layer of low P-wave velocity. This layer has been replaced with a fast-velocity limestone for the record on the right. Note the invasion of large-offset primary reflection data with coherent noise that is associated with this limestone layer.

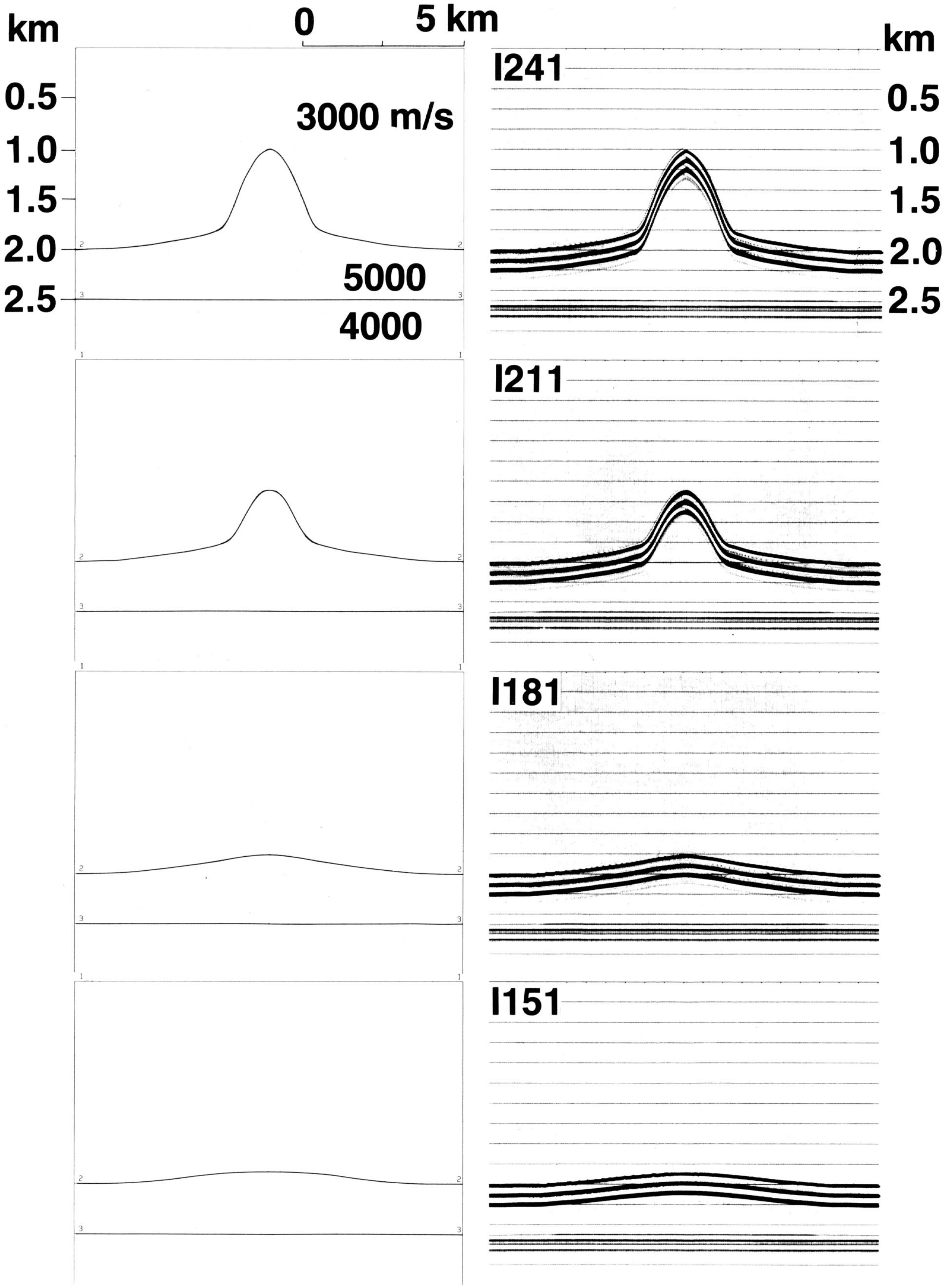


FIG. 8-20. Cross sections along four in lines (indicated on the base map in Figure 8-19) of the 3-D velocity-depth (left) and reflectivity (right) models. The 3-D and 2-D modeling results are shown in Figure 8-21.

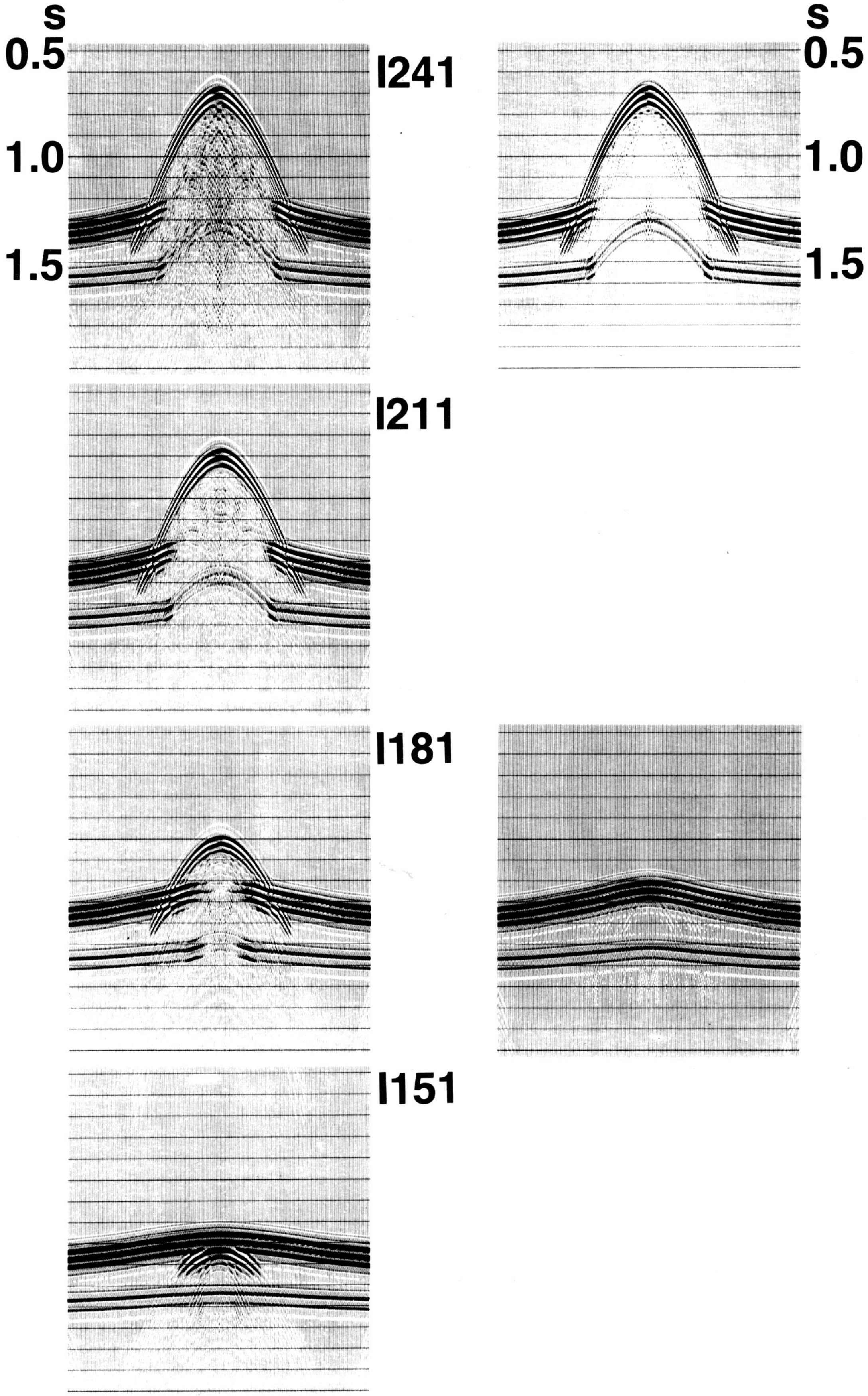


FIG. 8-21. Cross sections along four in lines (indicated on the base map in Figure 8-19) of the modeled 3-D (left) and 2-D (right) zero-offset wave fields associated with the 3-D velocity-depth and reflectivity models shown in Figure 8-20.

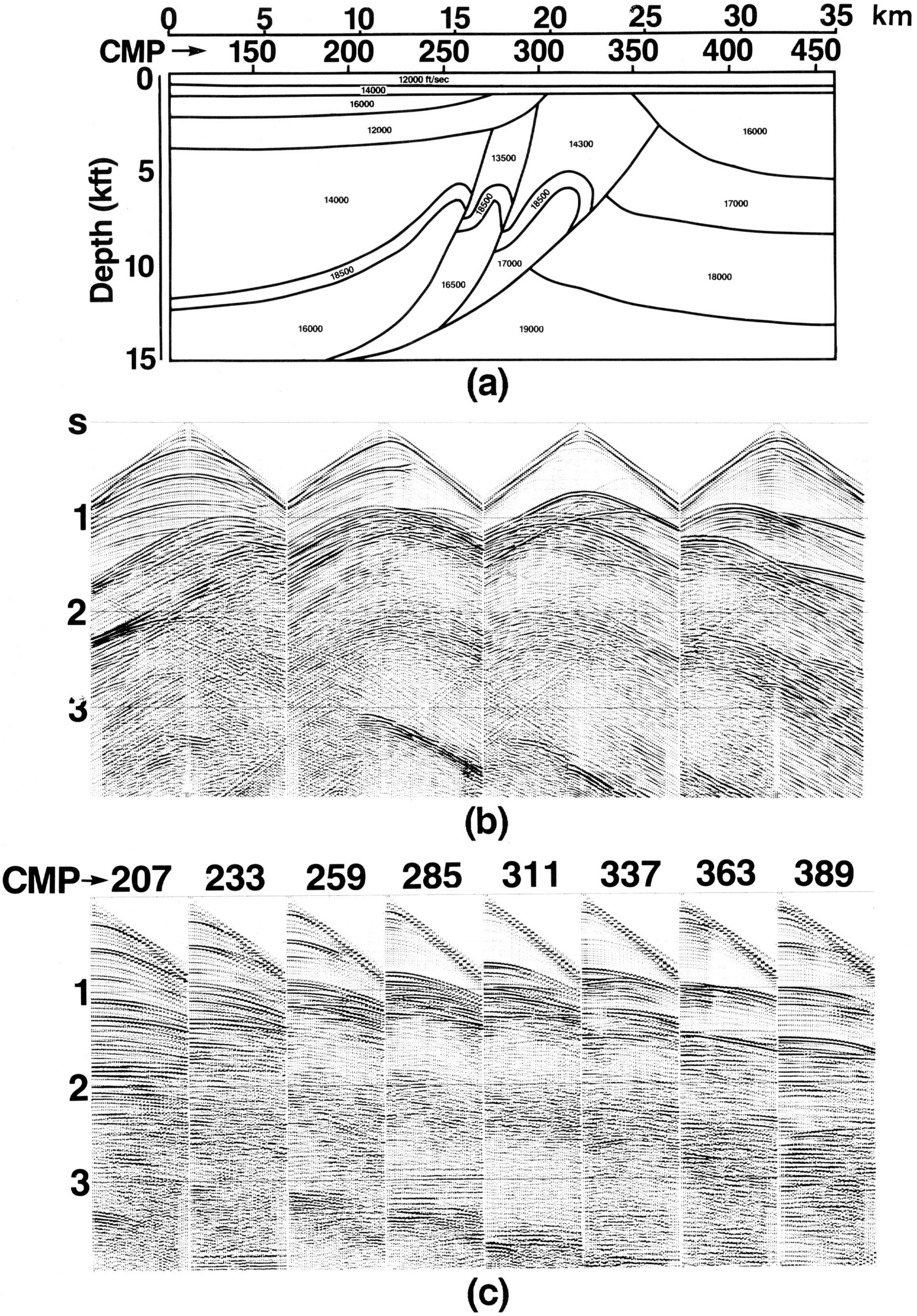


FIG. 8-22. (a) A 2-D velocity-depth model, (b) full acoustic modeling of shot gathers, (c) selected CMP gathers. (Modeling courtesy Amoco Production Company.)

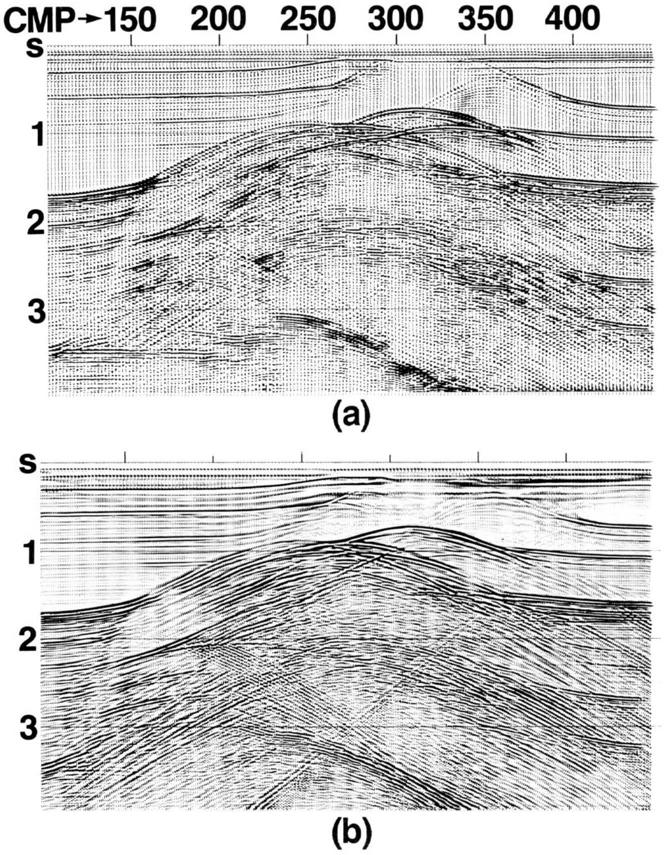


FIG. 8-23. (a) Zero-offset, (b) stacked sections from the nonzero-offset model data shown in Figure 8-23. (Data courtesy Amoco Production Company.)

In the field, limestone on the surface often generates a large amount of coherent noise.

Figures 8-26b and 8-26c show two synthetic shot gathers for which the upper part of the depth model consists of a 50-ft water layer underlain by a 1020-ft shale section with a *P*-wave velocity of 5500 ft/s. The deeper portion of the model for Figure 8-26b is an all-shale section with *P*-wave velocity increasing from 5600 ft/s on the top to 7700 ft/s on the bottom of the layer. The primary reflections *PP* between 380 and 850 ms are associated with this all-shale sequence. Figure 8-26c shows the effects of including a 30 percent sand layer between 660 and 850 ms. The *P*-wave reflections *PP* in Figure 8-26c show stronger amplitudes at large offsets. An analysis of amplitudes as a function of offset can provide hints for determining the sand-shale ratio, as well as fluid content, in some cases. Because the effects are complex, this type of modeling can be helpful in analyzing amplitude variations with offset. Also note the relatively strong converted *PS* and *SP* waves on the record from the sand-shale model. Modeling of this type also is useful when analyzing converted waves in multicomponent reflection data.

8.5 SYNTHETIC SONIC LOGS

The purpose of seismic data processing is to solve the inverse problem; that is, given the recorded wave field, we want to determine the geologic picture. If the processing has been careful and complete, and if the amplitude information has been properly restored, then the assumption that the traces in a stacked section represent the broadband reflection coefficient series at various CMP locations usually is reasonable. Since the geology (and not the reflection coefficients) is our final goal, we need to go one step further and derive the detailed velocity structure from this section. The signal bandwidth of the CMP stack determines the degree of detail in the velocity information that can be extracted in this final step.

From the convolutional model of a seismic trace described in Section 2.2, note that the starting point was the sonic log (Figure 2-8a). The reflection coefficient is defined as the ratio of the reflection amplitude to the incident wave amplitude. In terms of acoustic impedance, $I = \rho v$, where ρ is rock density and v is velocity, the reflection coefficient c is given by:

$$c = (I_2 - I_1)/(I_2 + I_1), \tag{8.3}$$

where I_1 and I_2 are the acoustic impedances of the medium above and below the reflecting interface. From equation (8.3), the reflection coefficient is interpreted as the ratio of the change in acoustic impedance (ΔI) to twice the average acoustic impedance I:

$$c = \Delta I/2I. \tag{8.4}$$

If we assume that density is constant, then equation (8.3) takes the form:

$$c = (v_2 - v_1)/(v_2 + v_1). \tag{8.5}$$

Therefore, reflection coefficients can be computed by differentiating the sonic log. The inverse process of getting the interval velocities from the reflection coefficients, which are assumed proportional to the stacked trace amplitudes, involves integration. In practice, only the high-frequency component of the interval velocity function can be obtained from this inversion. The low-frequency trend must be obtained from other information such as conventional velocity analysis or nearby sonic logs. In many practical situations, there is insufficient well control and the bands of the low-and high-frequency information derived from the seismic data do not overlap. In these cases, problems arise in merging the various types of information; hence, the fidelity of the synthetic sonic logs is degraded.

Figure 8-27 is a flowchart of synthetic sonic log construction. Figure 8-28 illustrates a stacked section with a bright-spot prospect. The corresponding synthetic sonic log section is shown in Figure 8-29, where the seismic data are superimposed as wiggle traces.

Lindseth (1979) first introduced the concept of synthetic sonic logs and used the concept to interpret stratigraphic

(text continued on page 484)

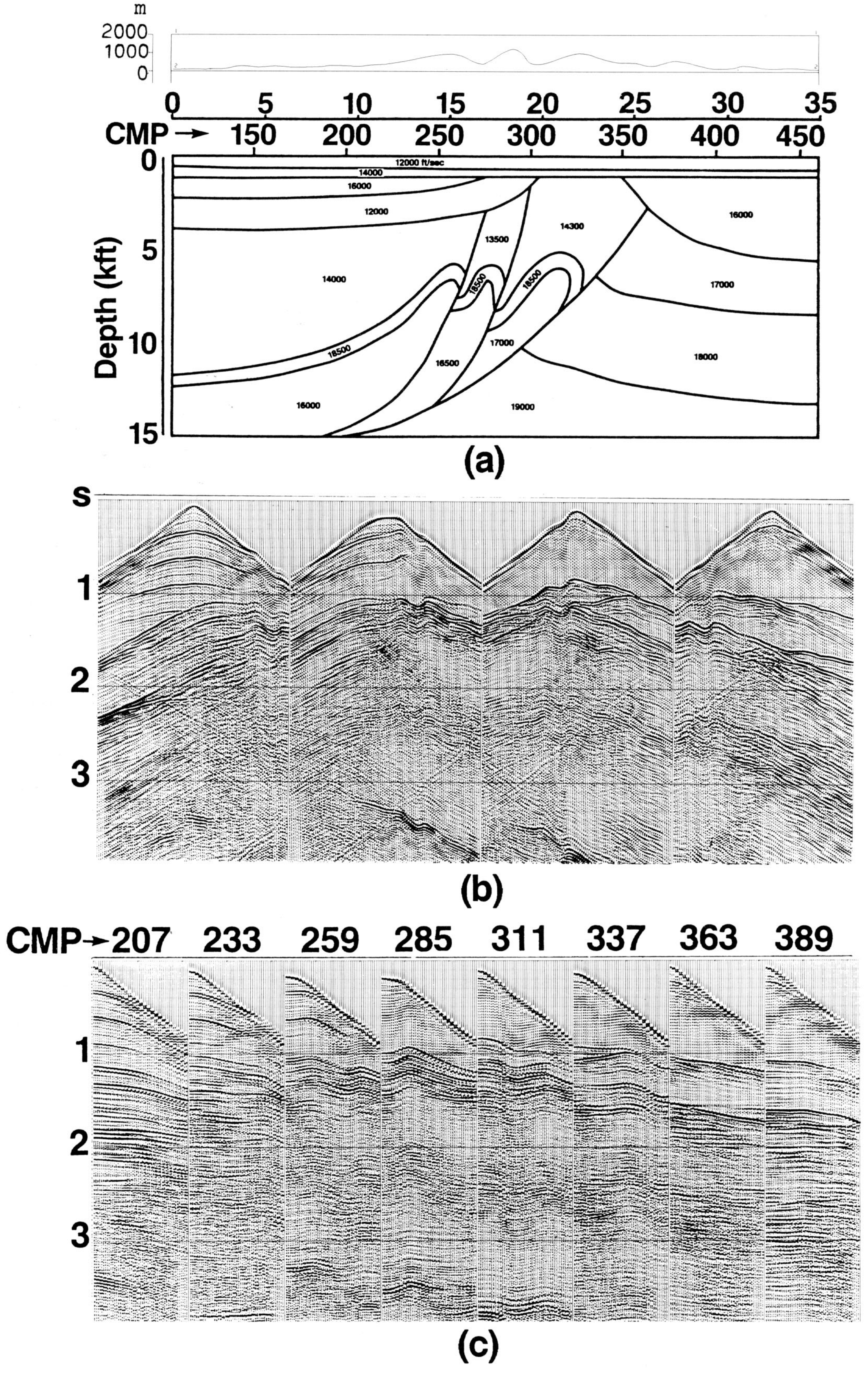


FIG. 8-24, Upward continuation using wave-equation datuming of shots and receivers from a flat datum at $z = 0$ to an irregular topography (a). Input is the shot gathers as in Figure 8-22b. (b) Shot gathers and (c) CMP gathers along the irregular datum. (Data courtesy Amoco Production Company.)

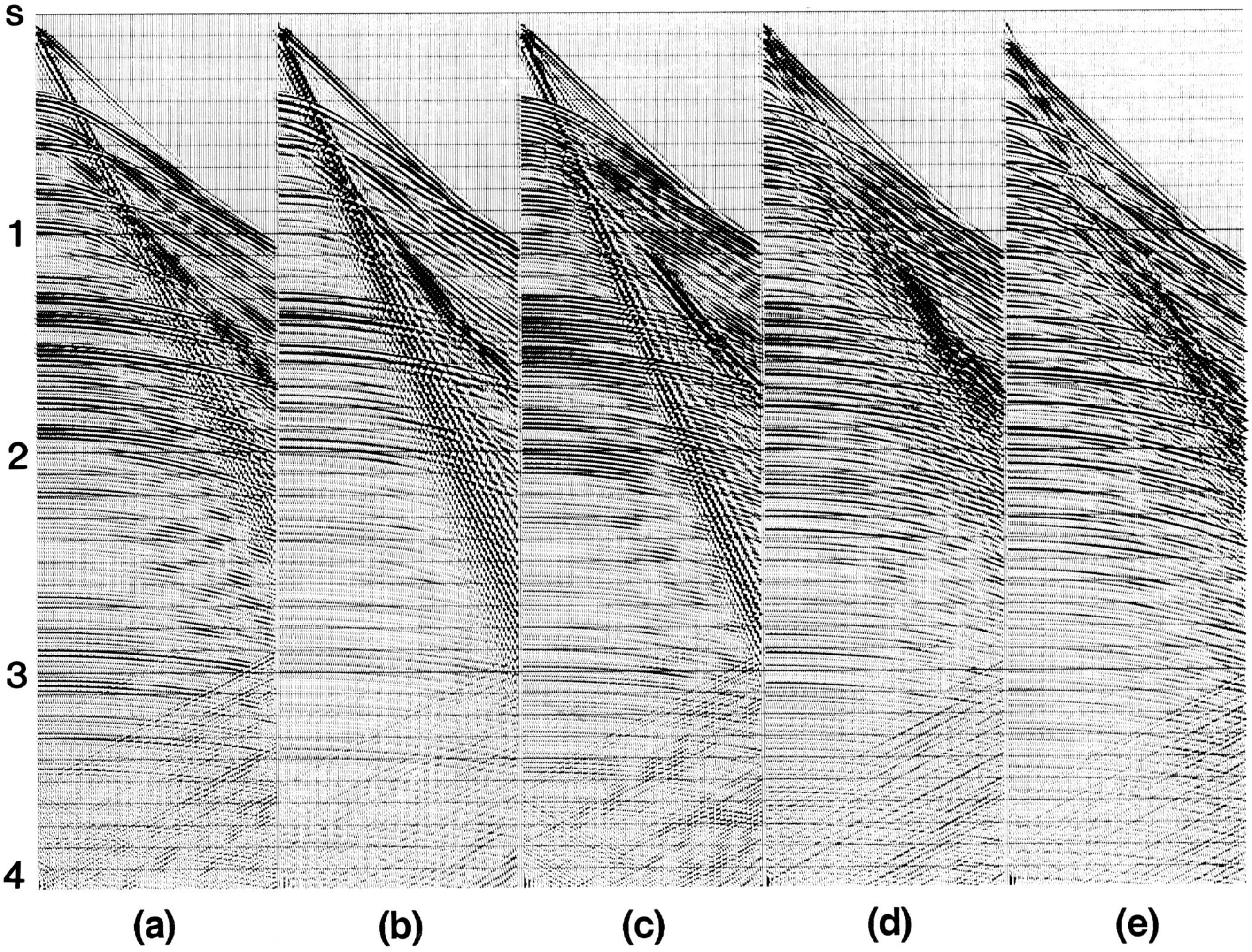


FIG. 8-25. Full elastic modeling of a water layer on top of a $v(z)$-layered earth. The water depths are (from left to right) 5, 10, 15, 20, and 50 m. Identify multiples (both reflected and refracted) and guided waves. The linear features below 3 s are artifacts of the modeling.

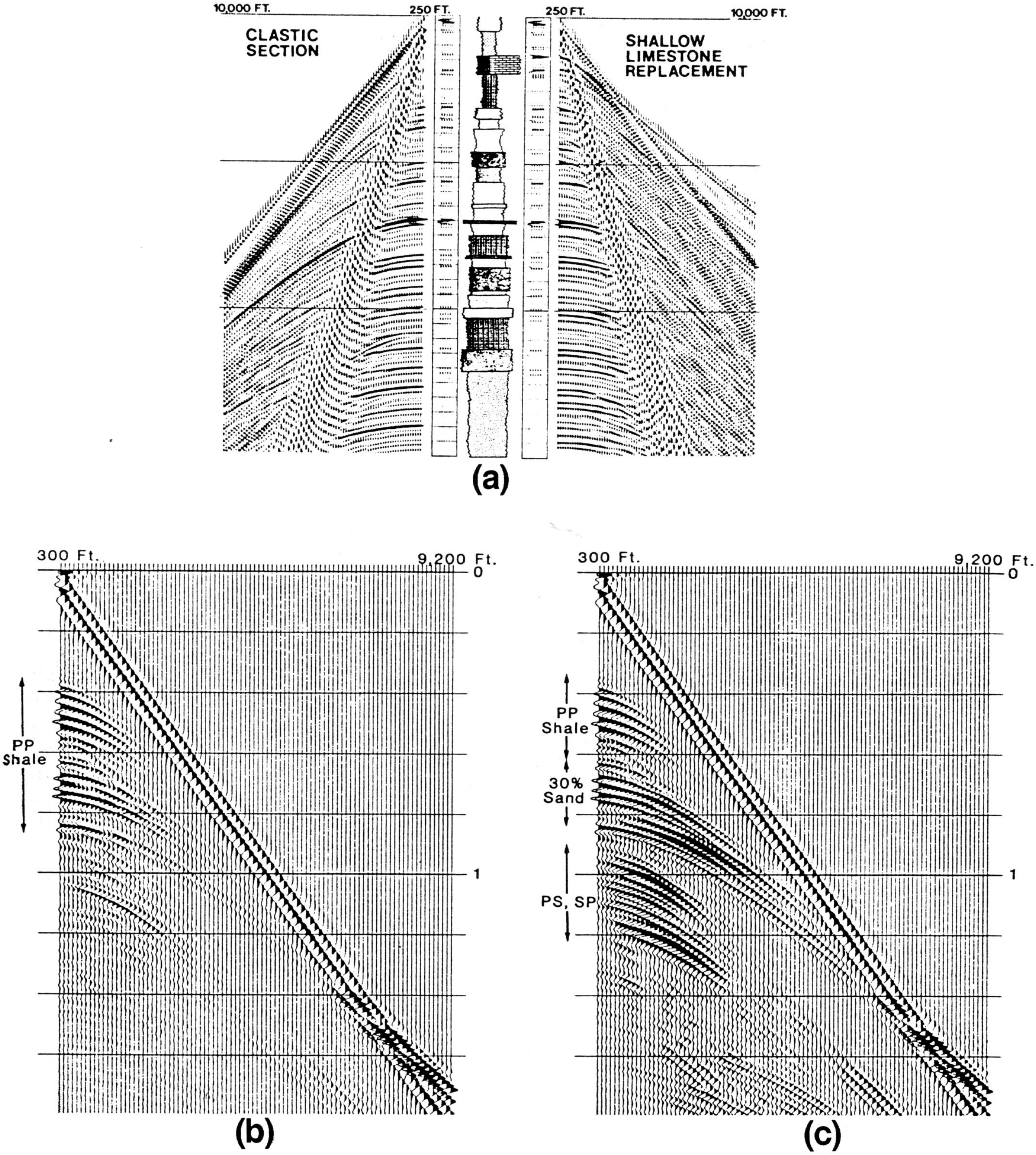


FIG. 8-26. Full elastic modeling examples: (a) the response of a clastic section with a low-velocity shallow layer on the left and the response of the same clastic section with a fast-velocity shallow layer on the right. What is the low-frequency low-group velocity dispersive energy? (b) Seismic response of an all-shale section, (c) seismic response of a sand-shale section (see text for details; Sherwood et al., 1983).

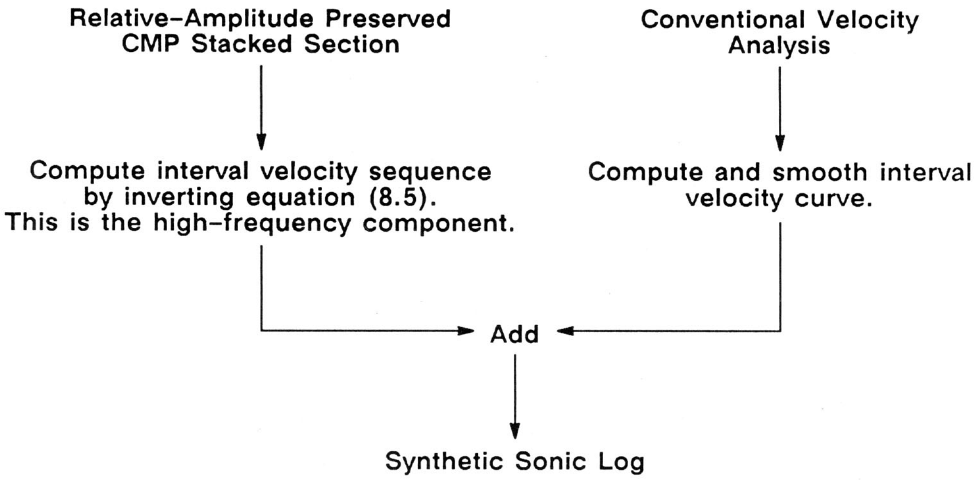


FIG. 8-27. Flowchart for synthetic sonic log construction.

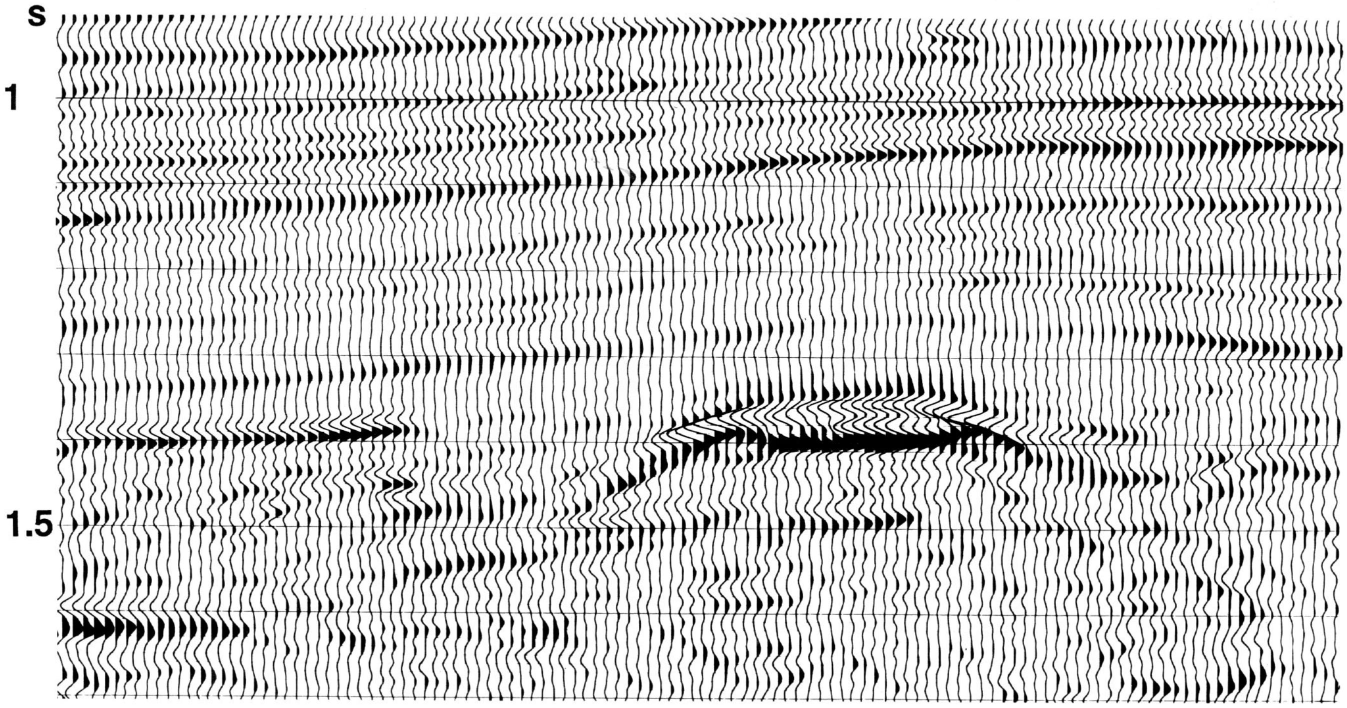


FIG. 8-28. Part of a CMP stack containing a bright spot.

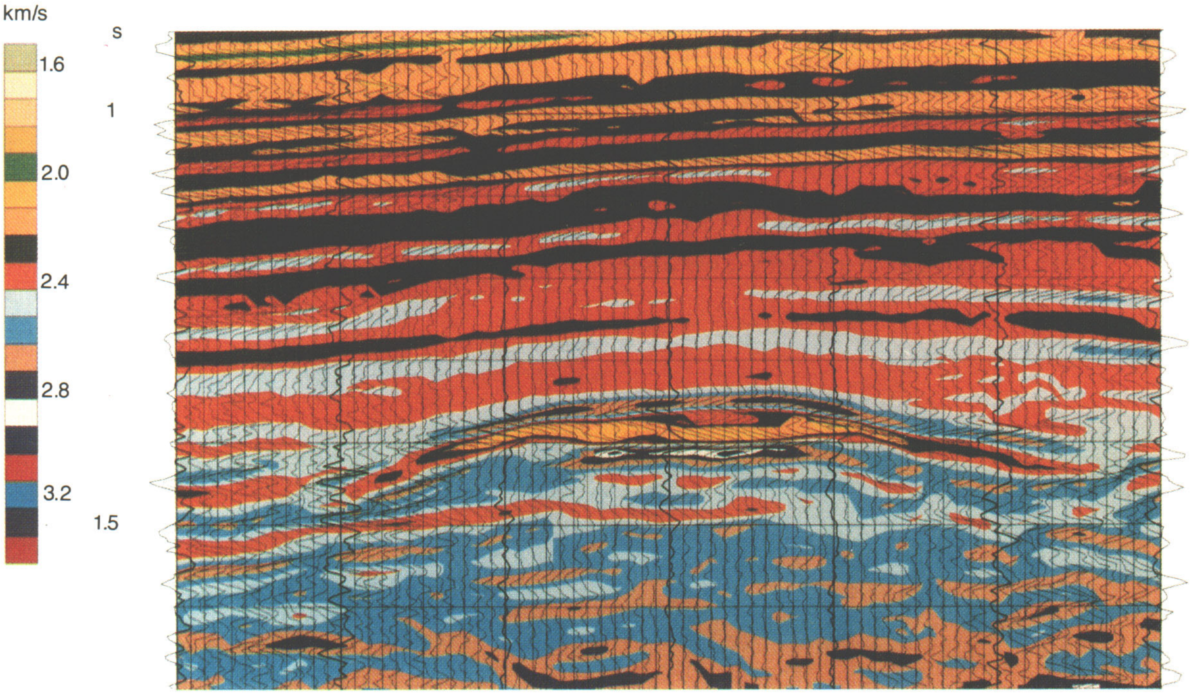


FIG. 8-29. A synthetic sonic log section derived from the stacked section in Figure 8-28.

prospects. Measured sonic logs generally contain much higher frequencies than seismic data. Integration of seismic traces to get synthetic sonic logs implies a further lowering of the frequency content. A synthetic sonic log and a measured log can only be compared when a high-cut filter is applied to the measured log to make the two appear to have equivalent bandwidths. On the other hand, a VSP provides essentially the same signal bandwidth as the surface seismic data (Section 8.7). Thus, comparison of the two is easier.

The inversion described here is based on the recursive relation in equation (8.5). Specifically, given values for c_i and v_i, v_{i+1} can be computed. Model-based, iterative inversion schemes also exist (Cooke and Schneider, 1983). These begin with an initial impedance function (specified at a CMP location) from which a synthetic seismogram is computed. This seismogram is crosscorrelated with the actual stack trace at that CMP location. The initial impedance model then is perturbed until the best match is attained between the estimated and actual seismic traces.

Finally, inversion techniques, which are based on a certain type of characterization of the reflection coefficient series estimated from stacked data, also exist. For example, the estimated reflection coefficient series can be characterized by a sparse spike series. The sparseness requirement can be satisfied by the condition that the sum of the absolute values of the estimated reflection coefficients is minimum (Oldenburg et al., 1983).

Note that regardless of the method, any inversion process suffers from the uncertainty (nonuniqueness) associated with the frequency components outside the dominant signal bandwidth. Thus, the results of the inversion may be questionable for the low-and high-frequency ends of the spectrum at which noise dominates the signal. Constraints (information other than the seismic data being inverted, such as geologic or well control) often are incorporated into inversion schemes used to handle the low-and high-frequency ends of the spectrum.

8.6 INSTANTANEOUS ATTRIBUTES

When considered as an analytic signal (in the mathematical sense), a seismic trace can be expressed as a complex function (Taner, 1978). The real part is the recorded seismic signal itself. The imaginary part is the quadrature, which is simply the 90-degree phase-shifted version of the real part. The quadrature trace is the Hilbert transform of the real part (Bracewell, 1965). After obtaining the complex seismic trace, we can easily compute the so-called instantaneous attributes associated with the seismic signal. (Mathematical details of computing the instantaneous attributes are given in Appendix E.)

The instantaneous amplitude measures the reflectivity strength, which is proportional to the square root of the total energy of the seismic signal at an instant of time. The instantaneous phase is a measure of the continuity of events on a seismic section. The temporal rate of change of the instantaneous phase is the instantaneous frequency. The instantaneous frequency may have a high degree of variation, which may be related to stratigraphy. However, it also may be difficult to interpret all this variation.

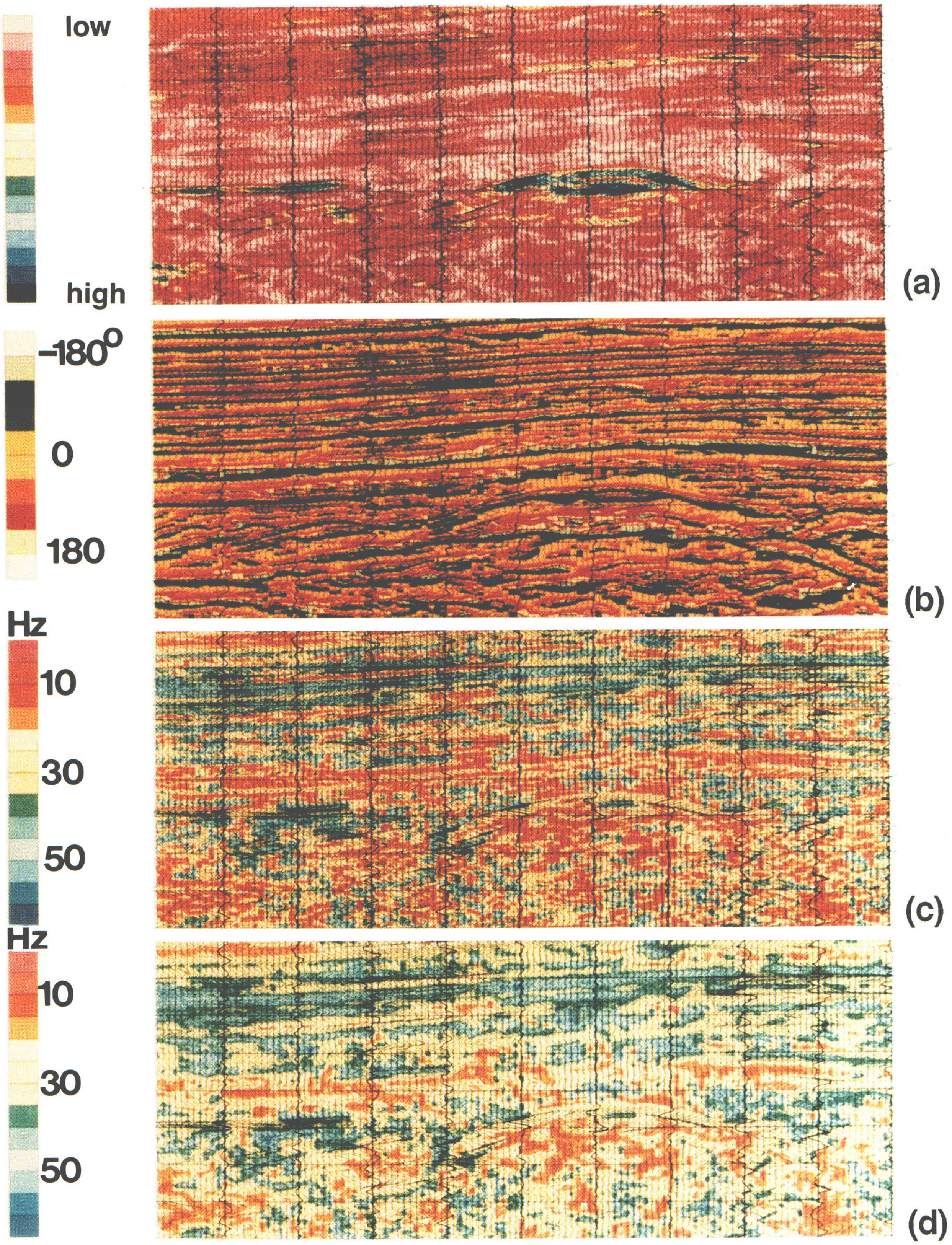


FIG. 8-30. Instantaneous attributes derived from the stacked section in Figure 8-28: (a) reflectivity strength, (b) phase, (c) frequency, and (d) smoothed frequency.

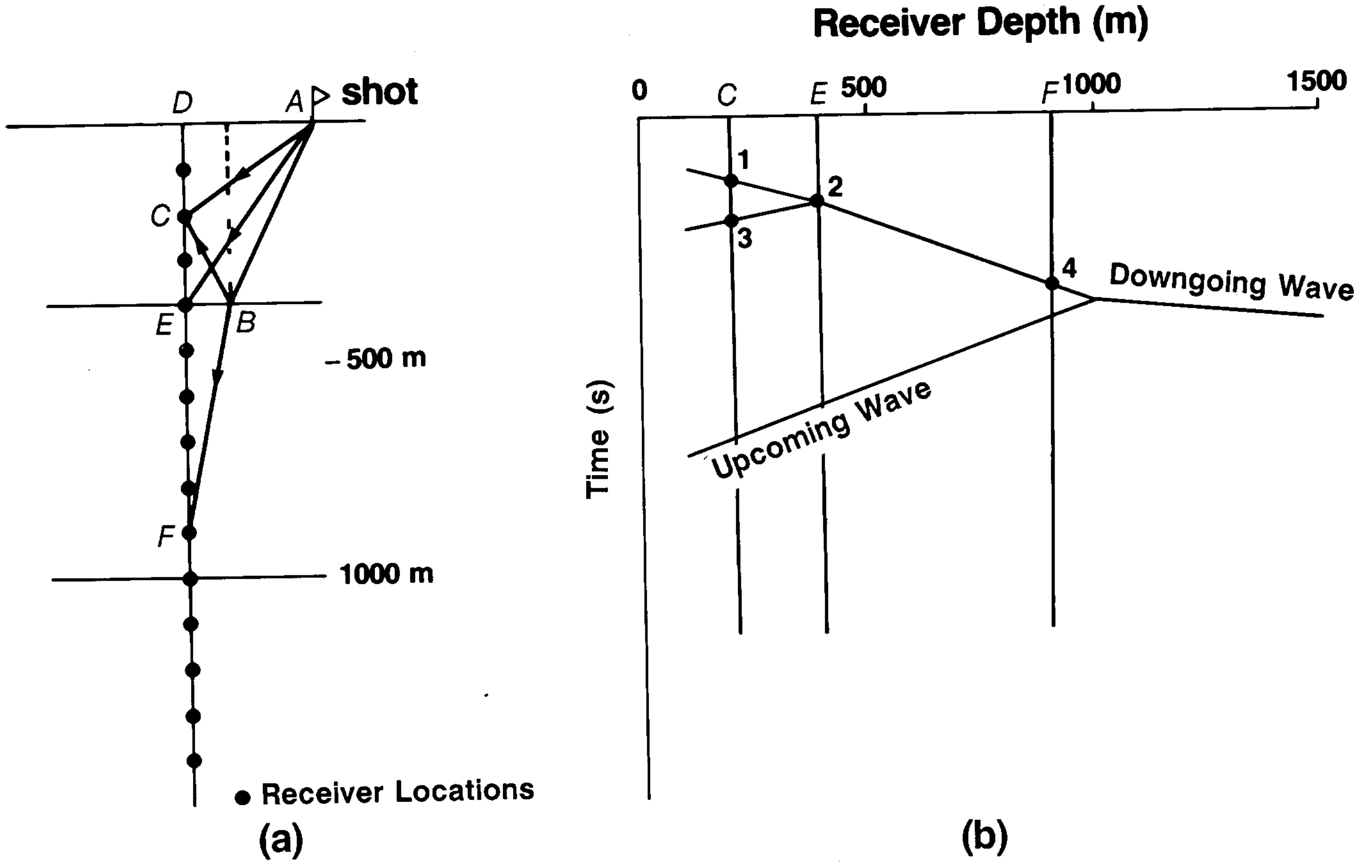


FIG. 8-31. Vertical seismic profiling geometry. (a) Raypaths and (b) associated traveltimes (see text for details). Static correction amounts to mapping traveltime associated with raypath ABC to traveltime associated with raypath ABC + CD; the NMO correction amounts to mapping traveltime associated with raypath ABCD to traveltime associated with raypath 2DE.

Therefore, the instantaneous frequency values often are smoothed in time.

The instantaneous measurements that are related to an analytic signal are associated with an instant of time, rather than an average over a time interval. These measurements are reliable when the seismic signal is recorded and processed so that the CMP stack closely represents the subsurface. In other words, to deduce any stratigraphic meaning from the seismic data before estimating the instantaneous parameters, the amplitude and frequency content of the seismic signal must be preserved in each processing step. Any variation in the shape of the basic waveform that is not attributable to the subsurface geology must be eliminated. Multiples and all types of random noise limit the reliability of the results.

Reflectivity strength is an effective tool to identify bright and dim spots. Phase information is useful in delineating such interesting features as pinchouts, faults, onlaps, and prograding reflections. Finally, instantaneous frequency information has helped to identify some condensate reservoirs, which tend to attenuate high frequencies.

Instantaneous attributes often are displayed in color for interpretation. Figure 8-30 shows the instantaneous attributes that correspond to the seismic section in Figure 8-28. The seismic data are displayed on top of the attributes as wiggle traces. Note the distinctive anomaly on the instantaneous amplitude section and enhancement of the continuity of reflections on the instantaneous phase section.

8.7 VERTICAL SEISMIC PROFILING

Vertical seismic profiling data often provide more reliable correlation of well control to seismic data than synthetic seismograms derived from sonic logs. There are two reasons for this. First, the VSP data have signal bandwidth closer to the seismic data than do sonic logs. More importantly, VSP data usually are not as sensitive to borehole conditions such as washouts.

Acquisition of VSP data involves a surface source that is located either close to the well head (zero-offset case) or away from the well head (offset VSP) and a geophone in the well bore. Several traces are recorded at the same geophone depth, then edited and summed. (Repeatable sources such as vibrators or air guns are used.) The geophone then is moved to a new depth location and recording is repeated. The resulting section is a profile displayed in depth and time. A comprehensive reference on VSP is the book by Hardage (1983).

Figure 8-31a shows VSP acquisition geometry. Consider a few of the raypaths that are intercepted by the receivers down the borehole: the direct arrivals from shot to receivers AC and AE, reflection ABC, and refraction ABF. Each receiver location yields a trace on the VSP that is in the depth-time domain (Figure 8-31b). Note that trace C has both the direct arrival (1) and the reflection from the first interface (3). The direct arrival (1 on trace C) and the refraction (4 on trace F) have only downgoing paths; therefore, they are called *downgoing waves*. On the other hand, reflection path ABC has a final upcoming segment BC (Figure 8-31a) and therefore constitutes an *upcoming wave* (Figure 8-31b). Note that direct arrival 2 on trace E coincides with a reflection arrival, provided the receiver is situated on the interface that causes the reflection. At this location, the upcoming and downgoing waves coincide (arrival 2 on trace E in Figure 8-31b).

We now discuss the basic steps in VSP processing. After some trace editing, VSP processing starts with separation of the downgoing waves from the upcoming waves (reflections). One separation technique is based on *f-k* filtering. Study of the zero-offset VSP in Figure 8-32a shows that the upcoming and downgoing waves have opposite dips. Because of this, each wave type should map in a different half plane in the *f-k* domain. Hence, downgoing waves can be suppressed with an *f-k* dip filter, thereby leaving only the reflection and associated multiples that constitute the upcoming waves (Figure 8-32b).

A VSP data set may not always have uniform receiver sampling in depth—a prerequisite for *f-k* filtering. The problems of edge effects and amplitude smearing often are observed on *f-k*-filtered data (Section 1.6.2).

An alternate approach to extracting upcoming waves is to use median filtering (Hardage, 1983). To start, apply first-arrival time shifts to traces in a VSP data set $VSP(z,t)$ to flatten the downgoing waves. (First arrivals now are at $t = 0$.) Then apply a median filter to each horizontal array of samples, $VSP(z,t = \text{constant})$. Median filtering is best explained by an example. Consider the following array of numbers: $(-1,2,1,2.5,1.5)$. Reorder its elements from small to large values: $(-1,1,1.5,2,2.5)$. The median of the series then is the middle sample, 1.5. Median filtering rejects noise bursts and any event that is not flat. Apply the median filter to the VSP data set with flattened downgoing waves to yield the downgoing waves. This result can be subtracted from the input to obtain the upcoming waves. The last step involves unflattening the data.

The next VSP processing step involves datuming all receivers to the well head (D in Figure 8-31a). From Figure 8-31, this static correction is the same as correcting each trace by an amount that is equal to the one-way traveltime down to the corresponding receiver location. For example, trace C is corrected by an amount equal to the traveltime associated with raypath DC.

The static corrections are followed by deconvolution and filtering (Figure 8-32c). In principle, the deconvolution operators can be designed from downgoing or upcoming waves. These deconvolution operators then are applied to traces of the upcoming wave profile. The common practice is to use the downgoing waves to design the deconvolution operators. This is because downgoing waves on a VSP record are much stronger than upcoming waves. Hence, designing deconvolution operators from downgoing waves has an advantage in that the operators are based on stronger signal, with more emphatically represented multiples (Hardage, 1983).

The last step involves stacking the traces in Figure 8-32c. Stacking normally includes a narrow corridor along the region in which upcoming and downgoing waves coincide. The resulting trace, repeated a few times, is shown in Figure 8-32d. To a large degree, corridor stacking prevents the multiples that do not merge with the downgoing wave path from being stacked in. The trace in Figure 8-32d can be considered an alternative to a zero-offset synthetic seismogram derived from the sonic log; thus, it can be compared to the CMP stack (not shown here) at the well location.

Figure 8-33a shows raw data from an offset VSP data set. By using the median filtering scheme described earlier, downgoing waves are extracted from the raw data (Figure 8-33b). Subtraction of downgoing waves from original data (followed by another median filtering for suppression of noise bursts) yields upcoming waves (Figure 8-33c). The upcoming waves then are deconvolved (Figure 8-33d) using operators designed from the downgoing waves of Figure 8-33b. This step normally is followed by static corrections. For nonzero-offset data, we also must correct for moveout resulting from the offset separation between the well head and the shot location. From Figure 8-31, this dynamic correction involves mapping the traveltime associated with raypath ABCD to 2DE. An NMO-corrected upcoming-wave profile can be compared with the surface seismic at the well location, provided the subsurface consists of horizontal layers with no lateral velocity variations.

When there are dipping interfaces, the upcoming-wave profile needs to be migrated; that is, the energy must be mapped to the actual subsurface reflection points. This is true even for zero-offset VSP data. A ray-tracing procedure for reflector mapping is illustrated in Figure 8-34. Note that reflection points D, E, and F have different lateral displacements OA, OB, and OC, respectively, from borehole Oz (Figure 8-34a). However, upcoming wave energy from all three reflection points is recorded on the same VSP trace at the receiver location R. The reflection times RG, RH, and RK (Figure 8-34b) are associated with raypaths SDR, SER, and SFR, respectively. Mapping this energy to reflection points involves a coordinate transformation (Wyatt and Wyatt, 1981; Cassell et al., 1984). In this transformation, the amplitudes on a single VSP trace are mapped onto several traces on the (x,t) plane, where x is the lateral distance of reflection points from the borehole (Figure 8-34b). The event times RG, RH, and RK are mapped onto two-way vertical times AL, BM, and CN, respectively. These vertical times are associated with raypaths AD, BE, and CF in

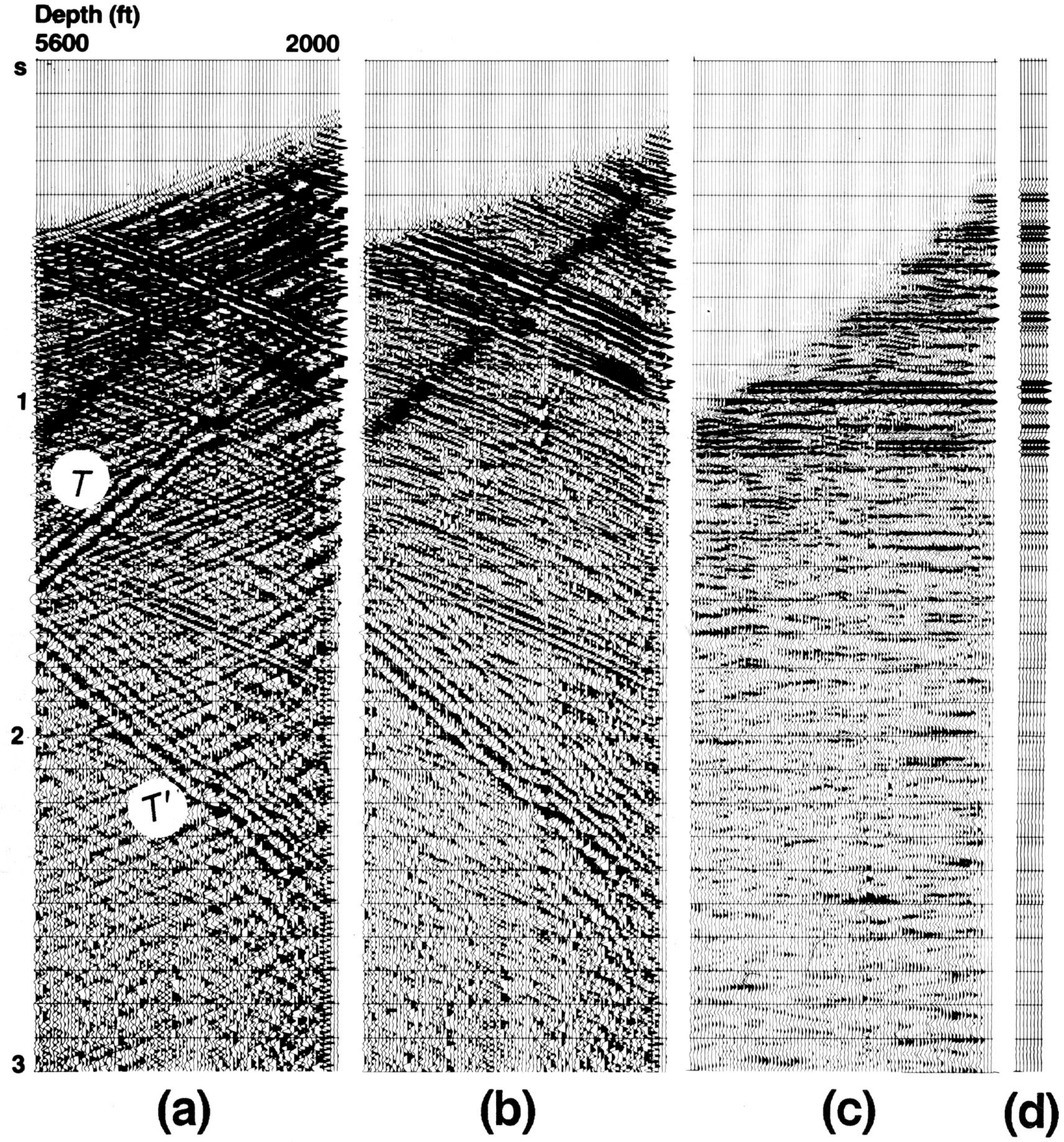


FIG. 8-32. Zero-offset VSP data at various stages of the processing sequence from left to right: (a) raw data, (b) upcoming waves, (c) after statics corrections followed by deconvolution and band-pass filtering, (d) corridor stack. Here, TT′ is a tube wave traveling through the borehole. (Data courtesy Amoco Europe and West Africa, Inc.)

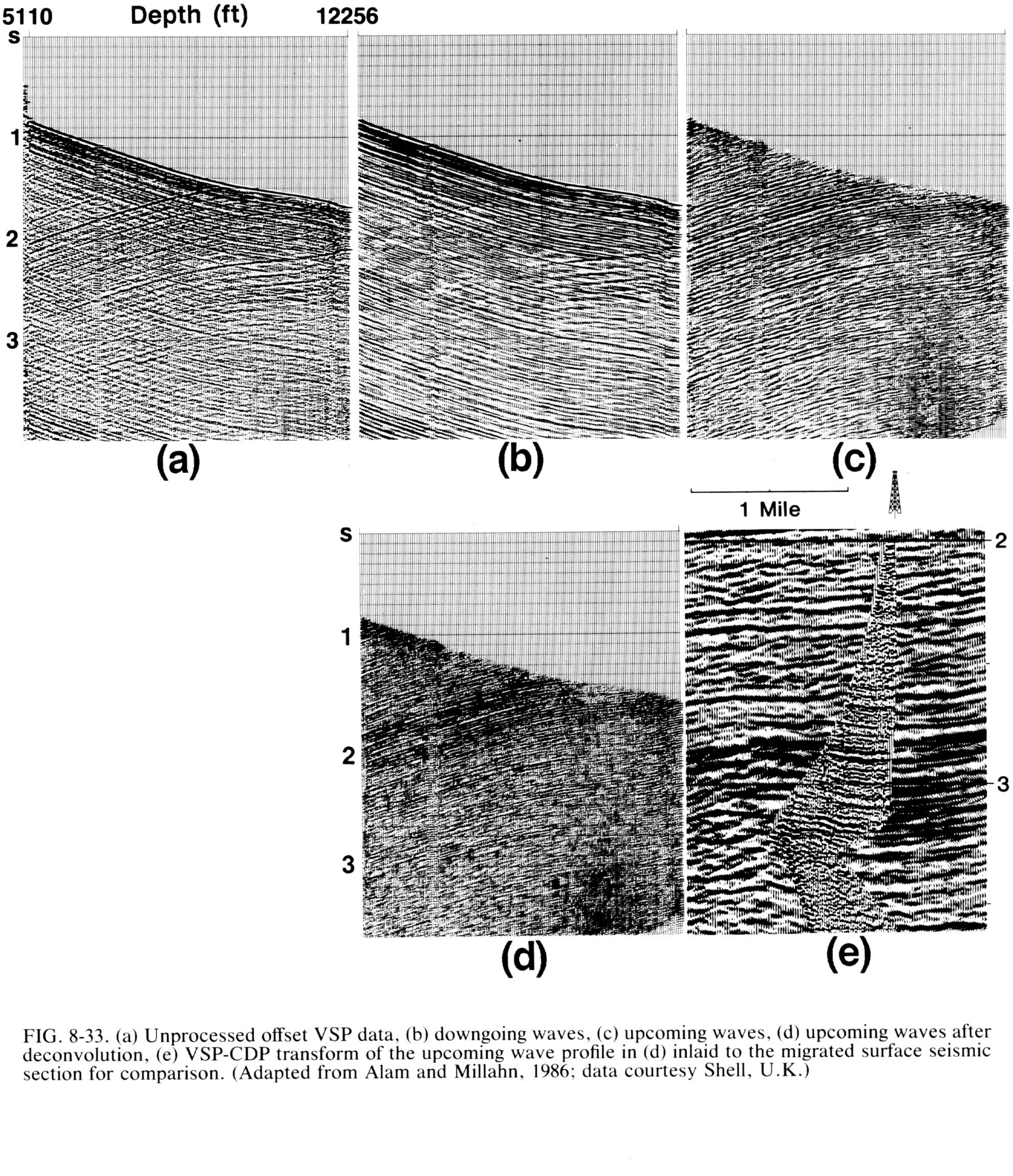


FIG. 8-33. (a) Unprocessed offset VSP data, (b) downgoing waves, (c) upcoming waves, (d) upcoming waves after deconvolution, (e) VSP-CDP transform of the upcoming wave profile in (d) inlaid to the migrated surface seismic section for comparison. (Adapted from Alam and Millahn, 1986; data courtesy Shell, U.K.)

Figure 8-34a. The resulting *(x,t)* section consists of traces similar to the traces of a migrated zero-offset section. Hence, the described ray-tracing procedure often is called VSP-CDP transformation.

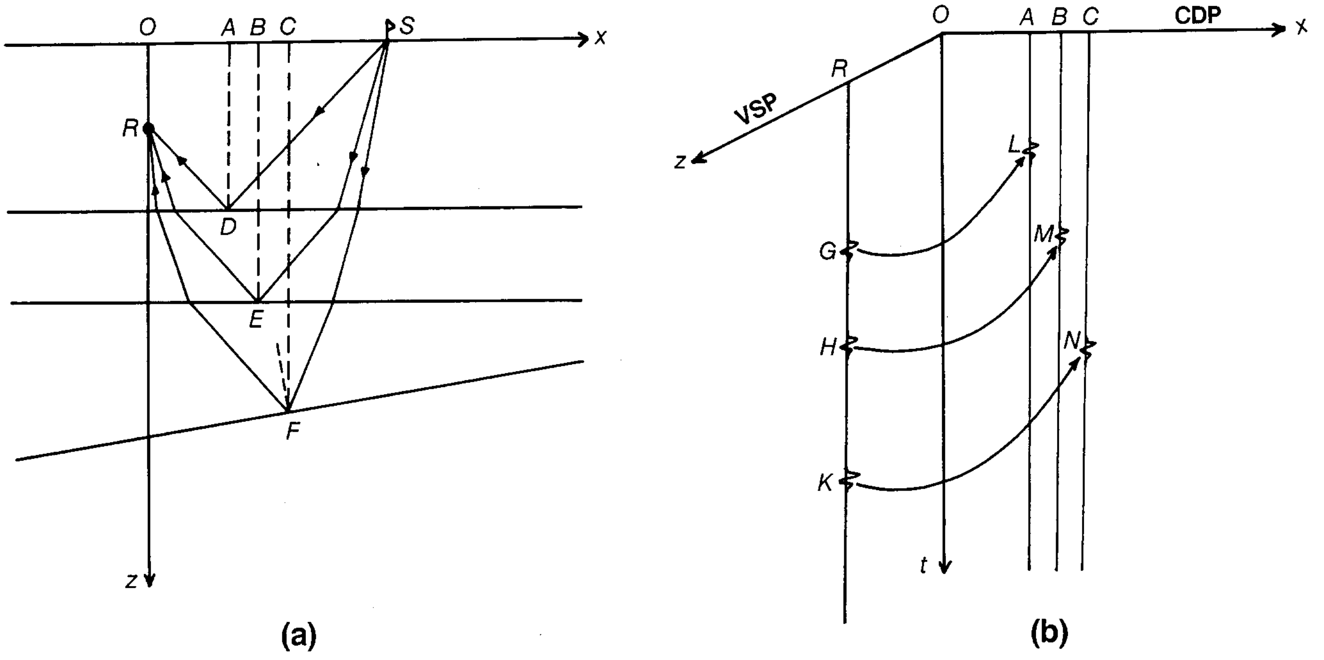
The VSP-CDP transform of the data in Figure 8-33d is shown in Figure 8-33e (Alam and Millahn, 1986). Comparison with the migrated surface seismic section at the well location indicates a good correlation of events. The difference in frequency content is partly attributed to differences in processing these two sections and is partly attributed to less high-frequency attenuation effects due to the shorter traveltimes associated with VSP recording.

The VSP-CDP transformation requires knowledge of the velocity-depth model around the borehole, since we must determine the location of the reflection points in the subsurface to perform mapping. The velocity-depth model can be derived using an iterative approach (Cassell et al., 1984). Starting with an initial velocity-depth model and the recording geometry for the VSP data, traveltimes for upcoming waves are computed. When these estimated traveltimes are compared with the observed traveltimes, discrepancies are noted and the velocity-depth model is modified accordingly. The process is repeated until a good match is attained between the estimated and observed traveltimes.

Finally, note that the VSP-CDP transform is not exactly a migration process. It handles neither diffractions nor curved interfaces. To handle these features, VSP data must be migrated (Dillon and Thomson, 1983). The VSP geometry is like the geometry of a common-shot gather, except the shot axis is perpendicular to the receiver axis. Migration of VSP data can be viewed as mapping amplitudes along semielliptical trajectories with their focal points being the source and receiver locations. Superposition of all these trajectories yields the migrated section. The aperture width for VSP data often is inadequate to obtain a migrated section without much smearing.

8.8 2-D SURFACE DATA PROCESSING

A *map* is defined as a 2-D surface *g(x,y)*. Depending on the quantity being mapped, *g* may have many types of units; for example, gravitational attraction (mgal), magnetic intensity (gamma), elevation, or even times picked along marker horizons from seismic data. Here, we set the convention that the positive *x* axis points eastward and the positive *y* axis points northward. For many types of mapping, *g(x,y)* is a smooth function and such a map can be analyzed in the Fourier transform domain. However, there are situations (as for isochron and structure maps) in which the map function has discontinuities that represent faulting. A discrete map function is represented



(a)

(b)

FIG. 8-34. (a) Source-receiver geometry for offset VSP; (b) VSP-CDP transformation: Traveltimes RG, RH, and RK, associated with raypaths SDR, SER, and SFR are mapped to traveltimes AL, BM, and CN, associated with the two-way vertical raypaths 2AD, 2BE, and 2CF. Also note that the amplitudes on the VSP trace are positioned on traces after transformation with x-coordinates the same as those of the reflection points OA, OB, and OC.

by a grid of mesh points over the (x,y) plane. These mesh points commonly are spaced at equal intervals in the x and y directions.

Maps usually are created from irregularly spaced observation values. Thus, the map function at a particular grid point must be computed by some fitting procedure. A local plane surface fitting procedure is described in Appendix F.

Before map creation, some reduction usually is applied to observed data, followed by various types of editing. Figure 8-35 shows the contour map of the test surface used in this study. The void grid points are filled with the arithmetic mean of the map function.

The main topic of this section is the next processing stage, which involves various 2-D processing techniques, each of which is aimed at achieving a particular interpretive goal. To clarify the objectives in digital processing, consider the properties of the quantity represented by a map function. In particular, consider the gravity problem. Ultra-long wavelength anomalies generally are associated with variations in crustal thickness. Moderately long wavelengths usually are due to variations in the basement topography. Medium to short wavelength anomalies are closely related to local tectonic disturbances, such as folding. Finally, very short wavelength anomalies have a variety of sources; some are due to small, near-surface features, while others are just noise of various types. Note that gradation in wavelength variations is related somewhat to depth of the source that causes the anomaly. The longer the wavelength, the deeper the anomaly; the shorter the wavelength, the shallower the anomaly.

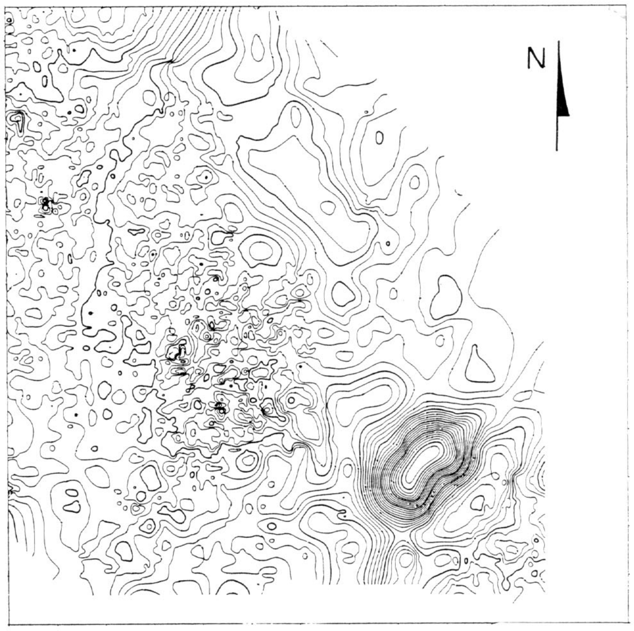


FIG. 8-35. A time map of a seismic horizon. The faults have been smoothed out.

Similar physical interpretations are possible with other map functions, such as the seismic (time) horizon map in Figure 8-35. Some of the very short wavelengths here are due to near-surface effects that occur as residual statics. Some of the short wavelengths may represent true structure, in which case the problem is subtle. Also note in Figure 8-35 that some moderately long wavelengths correspond to structural undulations that exist in the area. The most important observation made from this map is that most of the features with different wavelengths are not spatially isolated, but are superimposed. This characteristic is common for all types of map functions, whether gravity, magnetic, elevation, or time. To separate the effects of different features from each other, we must analyze them in terms of wavelength. The wavelength analysis also is a way to discriminate the depth of the source for some types of data such as potential field measurements.

The basic motivation behind digital processing of maps is to separate anomalies. Simple 2-D smoothing and wavelength filtering are techniques for separating anomalies. Vertical derivatives and analytic continuation also are useful for enhancing certain anomalies so that they appear more pronounced on the map. Some techniques for separating anomalies by their wavelengths are described next.

Before applying digital processing techniques, it is best to study the map in terms of wavelength composition. The 2-D amplitude spectrum of a map is an excellent tool for recognizing not only the wavelength content, but also the orientation of various components. The most useful display is the colored contour plot of the amplitude spectrum (Figure 8-36) from which various bands of wavelengths are clearly distinguished. Pink represents long, beige represents moderate, and yellow represents short wavelength anomalies. For a 2-D real function, such as a map, the amplitude spectrum is antisymmetric. Thus, only a pair of quadrants (e.g., the first and second) of the amplitude spectrum needs to be displayed.

8.8.1 Separation of Regional and Residual Anomalies

A simple 2-D smoothing operation is the easiest way to obtain a map that represents the regional anomaly. Figure 8-37 shows the map in Figure 8-35 after 2-D smoothing was applied. Most of the very short wavelength features present in the center of the original map have been eliminated. Depending on the interpreter's goal, this output might be a satisfactory regional map. (However, a smoother regional map often is desired.)

Smoothing basically is done by computing the average value of the grid points that fall onto a ring of some desired radius. The center of the ring coincides with the output point. There is no reason why more than one ring should not be considered. For n concentric rings with m_i points over the ith ring, the average value g_i of the quantity g_{ij} that is being mapped is:

$$\bar{g}_i = \frac{1}{m_i}\sum_{j=1}^{m_i} g_{ij}. \tag{8.6}$$

Thus, the cumulative average g over n rings is:

$$\bar{g} = \frac{1}{n}\sum_{i=1}^{n} \bar{g}_i. \tag{8.7}$$

Weighting factors, which depend on the distance from the center of the rings, often are used in smoothing algorithms. In these cases, equation (8.7) becomes:

$$\bar{g} = \sum_{i=1}^{n} w_i \bar{g}_i, \tag{8.8}$$

where w_i are the weights. The residual anomaly is defined by:

$$g = g_0 - \bar{g}, \tag{8.9}$$

where g_0 is the grid value at the center of the rings. Figure 8-38 shows the regional anomaly obtained by using 15 rings. In general, the more rings, the more smoothing of the data. Note that the output of the ring method is relative. Depending on the nature of the objectives of map processing, Figure 8-38 may or may not be a good estimate of the regional anomaly. In most cases, determining what is regional and what is residual is a matter of interpretive judgment.

8.8.2 2-D Wavelength Filtering

The hypothetical map in Figure 8-39a has 10 different anomalies of various shapes and orientations. Figure 8-39b is the 2-D amplitude spectrum of this map with (k_x, k_y) as the Fourier duals of (x,y). Anomaly 3, which has infinite wavelength in the NS direction (y axis) and a wavelength of 2AA′ in the EW direction (x axis), lies on the k_x axis in the transform domain. Anomaly 7 also lies on the k_x axis, but farther from the origin since it has a shorter wavelength component (2BB′) in the x direction. Anomaly 8 tilted counterclockwise by an angle θ degrees from the y direction has an apparent wavelength component 2CC′ in the y direction and 2DD′ in the x direction. Note the following relations:

$$\tan\theta = \frac{DD'}{CC'} = \lambda_x/\lambda_y, \tag{8.10}$$

where λ_x and λ_y are the wavelength components. By definition,

$$\lambda_x = 2\pi/k_x, \qquad \lambda_y = 2\pi/k_y,$$

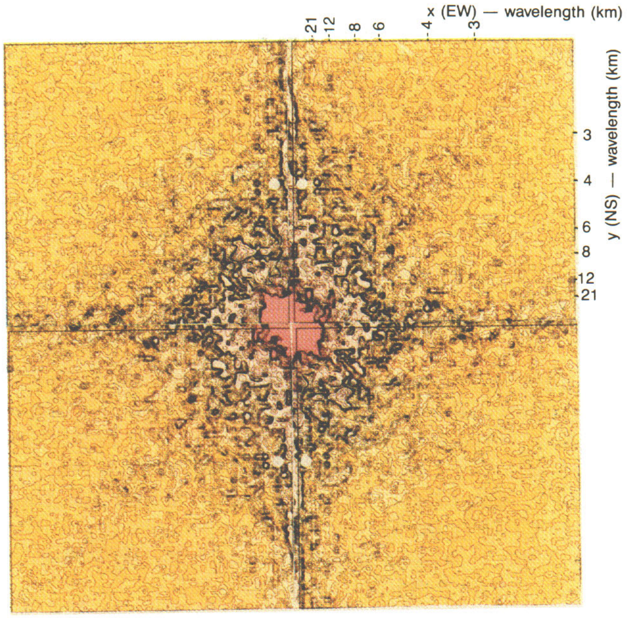


FIG. 8-36. 2-D amplitude spectrum of the map in Figure 8-35.

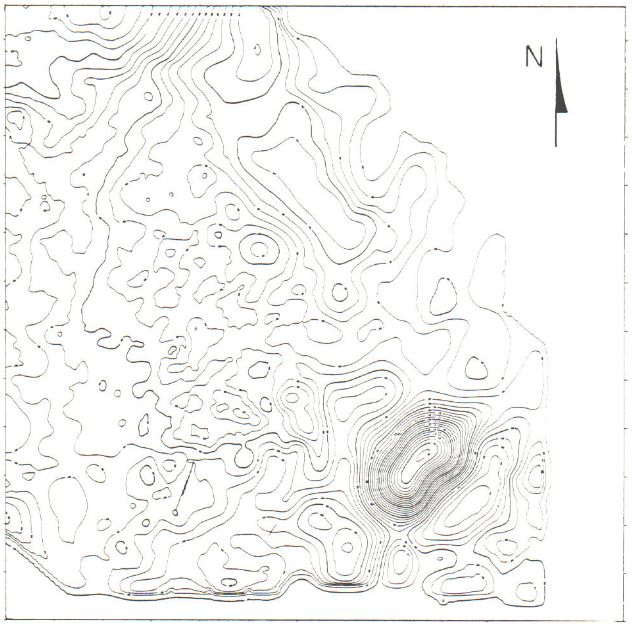


FIG. 8-37. A smoothed version of the contour map in Figure 8-35.

which, after substitution into equation (8.10), gives

$$\tan \theta = k_y/k_x. \tag{8.11}$$

Equations (8.10) and (8.11) imply that map trend MM′ is orthogonal to transform trend TT′.

Unlike the simple, linear anomalies 3, 7, and 8, rounded anomalies 1, 2, and 10 map onto the entire transform plane. Moreover, because of their equal size, anomalies 1 and 2 cannot be distinguished on the plane of 2-D amplitude spectrum, even though they are isolated in the space domain. However, the spatial extent of these two anomalies is much wider than that of anomaly 10. Thus, they are confined to lower wavenumbers. Finally, elongated features like anomalies 5, 6, and 9 map onto the transform plane at certain dominant directions that are orthogonal to their respective spatial trends.

Unlike the hypothetical anomalies in Figure 8-39a, real data consist of anomalies of various shapes and orientations that are superimposed on each other. However, in the transform domain, the real anomalies can be separated in terms of their wavelength contents and orientations; something that cannot be achieved in the space domain. Thus, the transform domain provides a way to apply various filtering operations to a map. A band of wavelengths can be passed regardless of orientation, as shown by the radial filter transfer function in Figure 8-40a. Anomalies that are oriented in the range (θ_1, θ_2) from the N direction can be passed regardless of their size, using a directional filter whose transfer function is given by Figure 8-40b. Finally, a band of wavelengths with a directional orientation can be passed. This is achieved by a filter with a transfer function as shown in Figure 8-40c. In practice, as in other filter design techniques, the transfer functions shown must be tapered in the neighborhood of cutoff wavelengths for optimal performance.

Once the transfer function is designed to suit the purpose, the filter itself can be applied to the map in the transform or the space domain. In the transform domain, the transfer function is multiplied with the 2-D Fourier transform of the map. Subsequent inverse transformation yields the filtered map. To apply the filtering in the space domain, first inverse Fourier transform the filter's transfer function to obtain its 2-D impulse response. Two-dimensional convolution of this impulse response with the map yields the filtered map. Note that filters whose transfer functions are depicted in Figure 8-40 do not cause any phase shift; i.e., anomalies are not displaced in the space domain after filtering.

Before applying the 2-D filters, the possible bands of the wavelengths present on the amplitude spectrum of the map to be filtered must be determined. Four regions were distinguished from the color·plot (Figure 8-36) of the amplitude spectrum of the test surface. We have the regional component down to 21-km wavelengths. A subregional part of the spectrum between the 21- and 12-km wavelengths is defined. A moderate range of wavelengths is between 12 and 8 km, and the residual components are beyond the 8-km cutoff wavelength. This residual region can be subdivided; anything less than 4 km corresponds to very short wavelength anomalies. These anomalies may be caused by various types of noise.

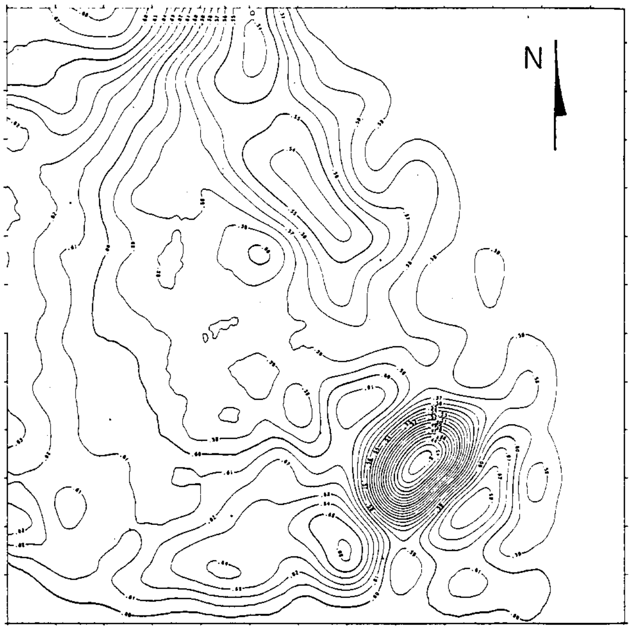


FIG. 8-38. A regional anomaly map based on equation (8.8). The input is the contour map in Figure 8-35.

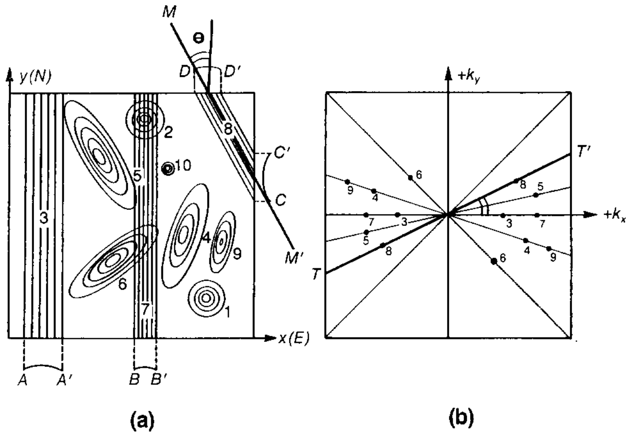


FIG. 8-39. Anomalies of various shapes in (a) the space and (b) the wavenumber domains.

The bands that were just determined now are used as cutoff wavelengths of the radial filter transfer function. A given map can be scanned with a suite of low-pass filters and several filtered maps can be produced, each with a potentially unique interpretational value. A result from such filtering operation is shown in Figure 8-41. Note the resemblance between Figure 8-41 and the regional map (Figure 8-38) that was obtained from the ring method described in Section 8.8.1. Although not demonstrated here, note that as the bandwidth of the filter is increased, more and more of the short wavelength anomalies are included, making the output less and less regional in character.

A filter scan is not limited to low-pass filtering only. High-pass, band-pass, and band-reject filters can be applied to maps to achieve a particular interpretational goal. For example, we may want to obtain a residual map that is free of the very short wavelength anomalies that usually are considered noise. This is accomplished by properly selecting the band-pass filter. A high-pass filter is the proper choice to retain all of the short wavelength region of the spectrum.

Directional filters scan the map of interest at various angles to emphasize a particular trend that may exist in the data. Figure 8-42 shows the result of one of the several directional filters applied to the test surface in Figure 8-35. A low-pass radial filter was cascaded with each directional filter so that any regional trend could be delineated in the area. The results of directional filtering indicate the existence of a prominent regional trend

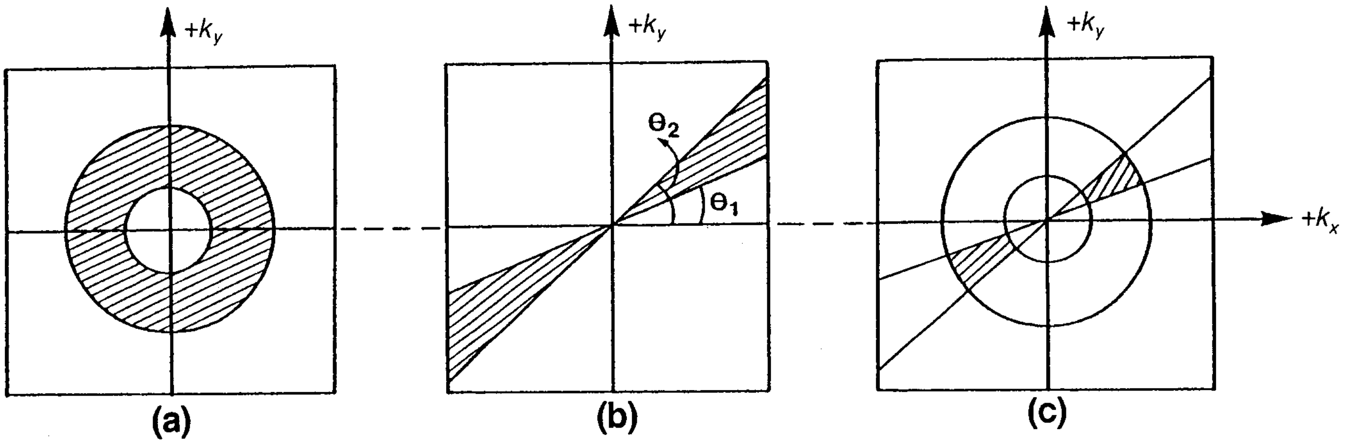


FIG. 8-40. Transfer functions for (a) band-pass, (b) directional, and (c) directional band-pass wavenumber filters.

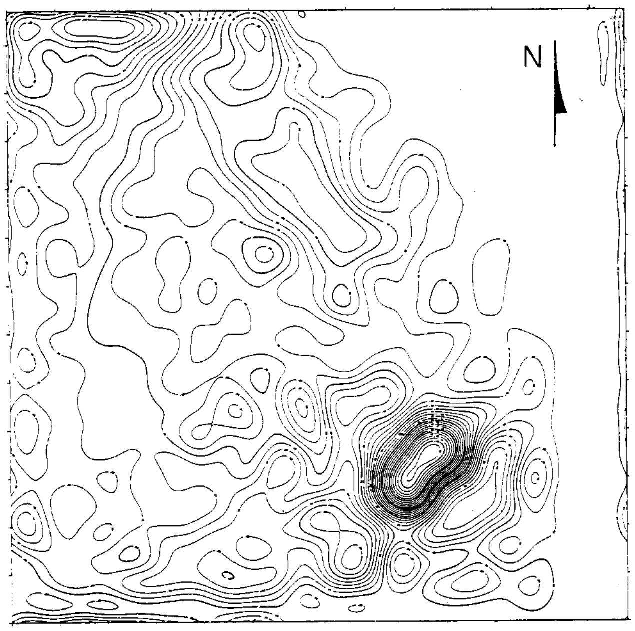


FIG. 8-41. The contour map in Figure 8-35 after low-pass filtering.

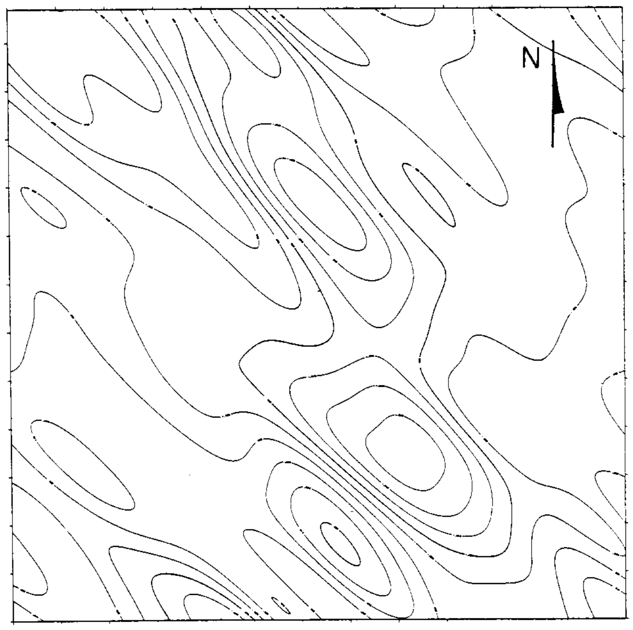


FIG. 8-42. The contour map in Figure 8-35 after directional low-pass filtering.

approximately N30°W (Figure 8-42). In some cases, a certain band of wavelengths may have one dominant trend that is different from that of another band of wavelengths. This situation may imply a change in the tectonic setting over the geologic history in the area.

REFERENCES

Alam, A. and Millahn, K., 1986, Interactive model-based VSP-CDP transform: Presented at the Symp. on Practical Aspects of Modeling in Exploration and Development, Kristiansand, Norway, March 3-5.

Berryhill, J., 1979, Wave-equation datuming: Geophysics, **44**, 1329-1344.

Bracewell, R., 1965, The Fourier transform and its applications: McGraw-Hill Book Co.

Cassell, B., Alam, A., and Millahn, K., 1984, Model-based interactive VSP-CDP transformation: Presented at the 54th Ann. Int. Soc. Explor. Geophys. Mtg.

Claerbout, J. F., 1985, Imaging the earth's interior: Blackwell Scientific Publications.

Cooke, D. A. and Schneider, W. A., 1983, Generalized linear inversion of reflection seismic data: Geophysics, **48**, 665-676.

Dillon, P. B. and Thomson, R. C., 1983, Image reconstruction for offset source VSP surveys: Presented at the 45th Ann. Eur. Assoc. Explor. Geophys. Mtg.

Hardage, B. A., 1983, Vertical seismic profiling: Geophysical Press, London-Amsterdam.

Hilterman, F. J., 1970, Three-dimensional seismic modeling: Geophysics, **35**, 1020-1037.

Hilterman, F. J., 1982, Interpretive lessons from three-dimensional modeling: Geophysics, **47**, 784-808.

Kallweit, R. S. and Wood, L. C., 1982, The limits of resolution of zero-phase wavelets: Geophysics, **47**, 1035-1046.

Kelamis, P. and Kjartansson, E., 1985, Forward modeling in frequency-space domain: Geophys. Prosp., **33**, 252-262.

Kelly, K. R., Ward, R. W., Treitel, W., and Alford, R. M., 1976, Synthetic seismograms: A finite-difference approach: Geophysics, **41**, 2-27.

Kjartansson, E., 1979, Modeling and migration by the monochromatic 45-deg equation: Stanford Exploration Project Report No. 16, Stanford University.

Lindseth, R., 1979, Synthetic sonic logs—a process for stratigraphic interpretation: Geophysics, **44**, 3-26.

Oldenburg, D. W., Scheuer, T., and Levy, S., 1983, Recovery of the acoustic impedance from reflection seismograms: Geophysics, **48**, 1318-1337.

Ricker, N., 1953, Wavelet contraction, wavelet expansion and the control of seismic resolution: Geophysics, **18**, 769-792.

Ryu, J. V., 1980, Decomposition of seismic gathers into velocity components by a space-time filter: Presented at the 50th Ann. Int. Soc. Explor. Geophys. Mtg.

Sengbush, R.L., 1983, Seismic exploration methods: International Human Resources Development Corp., Boston.

Sheriff, R. E., 1976, Inferring stratigraphy from seismic data: Am. Assoc. Petr. Geol. Bull., **60**, 528-542.

Sheriff, R. E., 1984, Encyclopedic dictionary of exploration geophysics: Soc. Explor. Geophys., Tulsa.

Sherwood, J. W. C., Hilterman, F. J., Neale, R. N., and Chen, K. C., 1983, Synthetic seismograms with offset for a layered elastic medium: Offshore Technology Conference, Paper 4508.

Stolt, R. H. and Benson, A. K., 1986, Seismic migration—theory and practice: Geophysical Press, London-Amsterdam.

Taner, M. T., 1978, Complex seismic trace analysis: Geophysics, **44**, 1041-1063.

Widess, M. B., 1973, How thin is a thin bed: Geophysics, **38**, 1176-1180.

Wyatt, K. and Wyatt, S.B., 1981, Determination of subsurface structural information using the vertical seismic profile: Presented at the 51st Ann. Int. Soc. Explor. Geophys. Mtg.

EXERCISES

Exercise 8.1. Derive the Fresnel zone expression [equation (8.2)] using Figure 8-15.

Exercise 8.2. Refer to the CMP stack in Figure 8-1d. Follow the first water-bottom multiple starting at 2 s on the left and ending at about 3 s on the right. Note the sharp increase in the apparent dip on the right within the mute pattern. What is the cause of the increase in the apparent dip?

Exercise 8.3. Invert equation (8.5) to derive the expression for computing interval velocities from reflection coefficients. Assume a starting velocity value of v_0.

Exercise 8.4. Refer to Figure 8-31a. Consider a surface multiple from the first reflecting interface. Trace the traveltime on the VSP diagram in Figure 8-31b. Multiples do not reach the downgoing wave path; thus, they can be eliminated by corridor stacking.

Exercise 8.5. Refer to Figure 8-31b. Should the slopes of the downgoing and upcoming waves associated with a layer be the same in magnitude?

Exercise 8.6. By examining the velocity spectrum following multiple suppression, can you tell which of the techniques has been used, *f-k* (Section 8.2.1) or *t-x* (Section 8.2.2)? (Hint: refer to Figures 8-1b, 8-8b, and 8-10b.)

Exercise 8.7. What procedure does CMP stacking correspond to in the *f-k* domain?

Exercise 8.8. Sketch the traveltime response for a point scatterer on a zero-offset VSP record.

Exercise 8.9. (a) Consider the reflectivity model for line I241 in Figure 8-20 (right column). How would the side boundary effects appear on the zero-offset forward model section if you made no attempt to suppress them in your modeling program? (b) Consider the zero-offset section for line I241 in Figure 6-32 (right column). How would the side boundary effects appear on the migrated section if you made no attempt to suppress them in your migration program?

Appendixes

Appendix A
MATHEMATICAL FOUNDATION OF THE FOURIER TRANSFORM

Given a continuous function $x(t)$ of a single variable t, its Fourier transform is defined by the following integral:

$$X(\omega) = \int_{-\infty}^{\infty} x(t) \exp(-i\omega t)\, dt, \qquad \text{(A.1)}$$

where ω is the Fourier dual of the variable t. If t signifies time, then ω is angular frequency. The temporal frequency f is related to angular frequency by $\omega = 2\pi f$. The Fourier transform is reversible; that is, given $X(\omega)$, the corresponding time function is:

$$x(t) = \int_{-\infty}^{\infty} X(\omega) \exp(i\omega t)\, d\omega. \qquad \text{(A.2)}$$

Throughout this book, the following sign convention is used for the Fourier transform. For the forward transform, the sign of the argument in the exponent is negative if the variable is time and positive if the variable is space. Of course, the inverse transform has the opposite sign used in the respective forward transform. For convenience, the scale factors in equations (A.1) and (A.2) are omitted.

Generally, $X(\omega)$ is a complex function. By using the properties of the complex functions, $X(\omega)$ is expressed as two other functions of frequency:

$$X(\omega) = A(\omega) \exp[i\phi(\omega)], \qquad \text{(A.3)}$$

where $A(\omega)$ and $\phi(\omega)$ are the amplitude and phase spectra, respectively. These are computed by the following equations:

$$A(\omega) = [X_r^2(\omega) + X_i^2(\omega)]^{1/2} \qquad \text{(A.4a)}$$

$$\phi(\omega) = \arctan[X_i(\omega)/X_r(\omega)], \qquad \text{(A.4b)}$$

where $X_r(\omega)$ and $X_i(\omega)$ are the real and imaginary parts of the Fourier transform $X(\omega)$. When $X(\omega)$ is expressed in terms of its real and imaginary components:

$$X(\omega) = X_r(\omega) + iX_i(\omega) \qquad \text{(A.5)}$$

and is compared with equation (A.3), note that:

$$X_r(\omega) = A(\omega) \cos \phi(\omega) \qquad \text{(A.6a)}$$

and

$$X_i(\omega) = A(\omega) \sin \phi(\omega). \qquad \text{(A.6b)}$$

We now consider functions $x(t)$ and $y(t)$. Table A-1 contains some basic theorems that are useful in various applications of the Fourier transform.

Table A-1. Fourier transform theorems (Bracewell, 1965).

Operation	Time Domain		Frequency Domain
(1) Addition	$x(t) + y(t)$	↔	$X(\omega) + Y(\omega)$
(2) Multiplication	$x(t)y(t)$	↔	$X(\omega) * Y(\omega)$
(3) Convolution	$x(t) * y(t)$	↔	$X(\omega)Y(\omega)$
(4) Autocorrelation	$x(t) * x(-t)$	↔	$\lvert X(\omega)\rvert^2$
(5) Derivative	$dx(t)/dt$	↔	$i\omega X(\omega)$
(6) Parseval's Theorem	$\int \lvert x(t)\rvert^2\, dt$	↔	$\int \lvert X(\omega)\rvert^2\, d\omega$

A discrete time function is called a *time series*. When digitized, the continuous function $x(t)$ takes the form:

$$x(t) = \sum_k x_k \delta(t - k\Delta t), \quad k = 0, 1, 2, \ldots, \qquad \text{(A.7)}$$

where Δt is the sampling interval and $\delta(t\text{-}k\Delta t)$ is the Dirac delta function. The discrete equivalent of the Fourier integral, equation (A.1), is written as a summation:

$$X(\omega) = \sum_k x_k \exp(-i\omega\, k\Delta t), \quad k = 0, 1, 2, \ldots. \qquad \text{(A.8)}$$

A new variable, $z = \exp(-i\omega\Delta t)$, now is defined. By substituting into equation (A.8) and explicitly writing the summation, we get:

$$X(z) = x_0 + x_1 z + x_2 z^2 + \cdots. \qquad \text{(A.9)}$$

Function $X(z)$ is called the z-transform of $x(t)$. It is a polynomial of the z variable. The power of z represents the time delay of the discrete samples in the time series $x(t)$.

The 2-D Fourier transform of wave field $P(x,t)$ is given by

$$P(k_x, \omega) = \iint P(x, t) \exp(ik_x x - i\omega t)\, dx\, dt. \qquad \text{(A.10)}$$

Function $P(x,t)$ can be reconstructed from $P(k_x,\omega)$ by the 2-D inverse Fourier transform:

$$P(x, t) = \iint P(k_x, \omega) \exp(-ik_x x + i\omega t)\, dk_x\, d\omega. \qquad \text{(A.11)}$$

The integral given by equation (A.10) is evaluated in two steps. First, by Fourier transforming in t,

$$P(x, \omega) = \int P(x, t) \exp(-i\omega t)\, dt, \qquad \text{(A.12)}$$

then by Fourier transforming in x, we get the 2-D transform:

$$P(k_x, \omega) = \int P(x, \omega) \exp(ik_x x)\, dx. \qquad \text{(A.13)}$$

REFERENCE

Bracewell, R. N., 1965, The Fourier transform and its applications: McGraw-Hill Book Co.

Appendix B

MATHEMATICAL FOUNDATION OF DECONVOLUTION

The material presented in this appendix is based on Robinson (1967), Treitel and Robinson (1966), Robinson and Treitel (1980), Claerbout (1976), and Yilmaz (1974).

B.1 Synthetic Seismogram

Consider an earth model that consists of homogeneous horizontal layers with thicknesses corresponding to the sampling interval. The seismic impedance associated with a layer is defined as $I = \rho v$, where ρ is density and v is the compressional wave velocity within the layer. The instantaneous value of seismic impedance for the kth layer is given by

$$I_k = \rho_k v_k. \tag{B.1}$$

For a vertically incident plane wave, the pressure amplitude reflection coefficient associated with an interface is given by

$$c_k = (I_{k+1} - I_k)/(I_{k+1} + I_k). \tag{B.2}$$

Assume that change of density with depth is negligible compared to the change of velocity with depth. Equation (B.2) then takes the form:

$$c_k = (v_{k+1} - v_k)/(v_{k+1} + v_k). \tag{B.3}$$

If the amplitude of the incident wave is unity, then the magnitude of the reflection coefficient corresponds to the fraction of amplitude reflected from the interface.

With knowledge of the reflection coefficients, we can compute the impulse response of the horizontally layered earth model using the Kunetz method (Claerbout, 1976). The impulse response contains not only the primary reflections, but also all the possible multiples. Finally, convolution of the impulse response with a source wavelet yields the synthetic seismogram. Although random noise can be added to the seismogram, the convolutional model used here to establish the deconvolution filters does not include random noise. We also assume that the source waveform does not change as it propagates down into the earth; hence, the convolutional model does not include intrinsic attenuation.

The convolution of a seismic wavelet *w(t)* with the impulse response *e(t)* yields the seismogram *x(t)*:

$$x(t) = w(t) * e(t). \tag{B.4}$$

By Fourier transforming both sides, we get:

$$X(\omega) = W(\omega)E(\omega), \tag{B.5}$$

where $X(\omega)$, $W(\omega)$, and $E(\omega)$ represent the complex Fourier transforms of the seismogram, the source waveform, and the impulse response, respectively. In terms of amplitude and phase spectra, the Fourier transforms are expressed as:

$$X(\omega) = A_x(\omega) \exp [i\phi_x(\omega)], \tag{B.6a}$$

$$W(\omega) = A_w(\omega) \exp [i\phi_w(\omega)], \tag{B.6b}$$

and

$$E(\omega) = A_e(\omega) \exp [i\phi_e(\omega)]. \tag{B.6c}$$

By substituting equations (B.6) into equation (B.5), we have

$$A_x(\omega) = A_w(\omega)A_e(\omega) \tag{B.7a}$$

and

$$\phi_x(\omega) = \phi_w(\omega) + \phi_e(\omega). \tag{B.7b}$$

In convolving the seismic wavelet with the impulse response, their phase spectra are added, while the amplitude spectra are multiplied. We now assume that the impulse response represents a *white process*; i.e., its amplitude spectrum is flat:

$$A_e(\omega) = A_0 = \text{constant}. \tag{B.8}$$

By substituting into equation (B.7a), we obtain:

$$A_x(\omega) = A_0 A_w(\omega). \tag{B.9}$$

This implies that the amplitude spectrum of the seismogram is a scaled version of the amplitude spectrum of the source wavelet.

We now examine the autocorrelation functions r of x, w, and e. Autocorrelation is a measure of similarity between the events on a time series at different time positions. It is a running sum as given by the expression:

$$r_e(\tau) = \frac{1}{N} \sum_{t=0}^{N-1} e_t e_{t+\tau},$$

$$\tau = 0, \mp 1, \mp 2, \ldots, \mp(N-1), \tag{B.10}$$

where τ is time lag. A random time series is an uncorrelated series. (Strictly, it is uncorrelated when it is continuous and infinitely long.) Therefore,

$$r_e(\tau) = 0, \qquad \tau \neq 0, \tag{B.11a}$$

and

$$r_e(0) = r_0 = \text{constant}. \tag{B.11b}$$

Equation (B.11) states that the autocorrelation of a perfect random series is zero at all lags except at zero lag.

The zero-lag value actually is the cumulative energy contained in the time series:

$$r_0 = e_0^2 + e_1^2 + \cdots + e_{N-1}^2. \tag{B.12}$$

Consider the z-transform of the convolutional model in equation (B.4):

$$X(z) = W(z)E(z). \tag{B.13}$$

By putting $1/z$ in place of z and taking the complex conjugate, we get

$$\bar{X}(1/z) = \bar{W}(1/z)\bar{E}(1/z), \tag{B.14}$$

where the bar denotes the complex conjugate. By multiplying both sides of equations (B.13) and (B.14), we get:

$$X(z)\bar{X}(1/z) = [W(z)\bar{W}(1/z)][E(z)\bar{E}(1/z)]. \tag{B.15}$$

By rearranging the right side,

$$X(z)\bar{X}(1/z) = [W(z)E(z)][\bar{W}(1/z)\bar{E}(1/z)]. \tag{B.16}$$

Finally, by definition, equation (B.16) yields:

$$r_x = r_w * r_e, \tag{B.17}$$

where r_x, r_w, and r_e are the autocorrelations of the seismogram, seismic wavelet, and impulse response, respectively. Based on the white reflectivity series assumption, we have:

$$r_x = r_0\, r_w. \tag{B.18}$$

Equation (B.18) says that the autocorrelation of the seismogram is a scaled version of that of the seismic wavelet. We will see that conversion of the seismic wavelet into a zero-lag spike requires knowledge of the wavelet's autocorrelation. Equation (B.18) suggests that the autocorrelation of the seismogram can be used in lieu of the seismic wavelet, since the latter often is not known.

B.2 Inverse of the Source Wavelet

A basic purpose of deconvolution is to compress the source waveform into a zero-lag spike so that closely spaced reflections can be resolved. Assume that a filter operator $f(t)$ exists such that

$$w(t) * f(t) = \delta(t), \tag{B.19}$$

where $\delta(t)$ is the Kronecker delta function. The filter $f(t)$ is called the inverse filter for $w(t)$. Symbolically, $f(t)$ is expressed in terms of the seismic wavelet $w(t)$ as:

$$f(t) = 1/w(t). \tag{B.20}$$

The z-transform of the seismic wavelet with a finite length $m + 1$ is (Appendix A):

$$W(z) = w_0 + w_1 z + w_2 z^2 + \cdots + w_m z^m. \tag{B.21}$$

The z-transform of the inverse filter can be obtained by polynomial division:

$$F(z) = 1/W(z). \tag{B.22}$$

The result is another polynomial whose coefficients are the terms of the inverse filter:

$$F(z) = f_0 + f_1 z + f_2 z^2 + \cdots + f_n z^n + \cdots. \tag{B.23}$$

Note that the polynomial $F(z)$ in equation (B.23) has only positive powers of z; this means $f(t)$ is causal. If the coefficients of $F(z)$ asymptotically approach zero as time goes to infinity, so that the filter has finite energy, we say that the filter $f(t)$ is realizable. If the coefficients increase without bound, we say that the filter is not realizable. In practice, we prefer to work with a causal and realizable filter. Such a filter, by definition, also is minimum-phase. If $f(t)$ is minimum-phase, then the seismic wavelet $w(t)$ also must be minimum-phase. Finally, to apply the filter with a finite length $n + 1$, the polynomial $F(z)$ must be truncated. Truncation of the filter operator induces some error in spiking the seismic wavelet.

The inverse of the seismic wavelet also can be computed in the frequency domain. By Fourier transforming equation (B.19), we get

$$W(\omega)F(\omega) = 1. \tag{B.24}$$

By substituting equation (B.6b), we get

$$F(\omega) = 1/\{A_w(\omega) \exp\,[i\phi_w(\omega)]\}. \tag{B.25}$$

We express the Fourier transform of the inverse filter $F(\omega)$ as:

$$F(\omega) = A_f(\omega) \exp\,[i\phi_f(\omega)], \tag{B.26}$$

and compare it with equation (B.25) to get:

$$A_f(\omega) = 1/A_w(\omega), \tag{B.27a}$$

and

$$\phi_f(\omega) = -\phi_w(\omega). \tag{B.27b}$$

Equations (B.27) show that the amplitude spectrum of the inverse filter is the inverse of that of the seismic wavelet, and the phase spectrum of the inverse filter is negative of that of the seismic wavelet.

B.3 The Inverse Filter

Instead of the polynomial division procedure described by equation (B.22), consider a different approach to derive the inverse filter. Start with the z-transform of the autocorrelation of the seismic wavelet:

$$R_w(z) = W(z)\bar{W}(1/z), \tag{B.28}$$

and the z-transform of equation (B.19):

$$W(z)F(z) = 1 \tag{B.29}$$

from which we get:

$$W(z) = 1/F(z). \tag{B.30}$$

By substituting into equation (B.28):

$$R_w(z)F(z) = \bar{W}(1/z). \tag{B.31}$$

Since the wavelet is a real-time function,

$$r_w(\tau) = r_w(-\tau). \tag{B.32}$$

Consider the special case of a three-point inverse filter (f_0, f_1, f_2). We will assume that the seismic wavelet $w(t)$ is minimum-phase (causal and realizable); hence, its inverse $f(t)$ also is minimum-phase. The z-transform of $f(t)$ is:

$$F(z) = f_0 + f_1 z + f_2 z^2. \tag{B.33a}$$

The z-transform of its autocorrelogram $r_w(\tau)$ is [by way of equation (B.32)]:

$$R_w(z) = \cdots + r_2 z^{-2} + r_1 z^{-1} + r_0 + r_1 z + r_2 z^2 + \cdots \tag{B.33b}$$

The z-transform of $w(t)$ is:

$$W(z) = w_0 + w_1 z + w_2 z^2 + \cdots + w_m z^m,$$

therefore,

$$\bar{W}(1/z) = \bar{w}_0 + \bar{w}_1 z^{-1} + \bar{w}_2 z^{-2} + \cdots + \bar{w}_m z^{-m}. \tag{B.33c}$$

By substituting equations (B.33a), (B.33b), and (B.33c) into equation (B.31), we obtain:

$$\begin{aligned}(r_2 z^{-2} + r_1 z^{-1} + r_0 + r_1 z + r_2 z^2)(f_0 + f_1 z + f_2 z^2)\\ = \bar{w}_0 + \bar{w}_1 z^{-1} + \bar{w}_2 z^{-2}.\end{aligned} \tag{B.34}$$

To solve for (f_0, f_1, f_2), we identify coefficients of powers of z. The coefficient of z^0 is:

$$r_0 f_0 + r_1 f_1 + r_2 f_2 = \bar{w}_0.$$

The coefficient of z^1 is:

$$r_1 f_0 + r_0 f_1 + r_1 f_2 = 0,$$

while the coefficient of z^2 is:

$$r_2 f_0 + r_1 f_1 + r_0 f_2 = 0.$$

When put into matrix form, the equations for the coefficients yield:

$$\begin{bmatrix} r_0 & r_1 & r_2 \\ r_1 & r_0 & r_1 \\ r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} \bar{w}_0 \\ 0 \\ 0 \end{bmatrix}. \tag{B.35}$$

Note that $\bar{w}_0$, which equals w_0 for the usual case of real source wavelet, is the amplitude of the wavelet at $t = 0$. There are four unknowns and three equations. By normalizing with respect to f_0, we get:

$$\begin{bmatrix} r_0 & r_1 & r_2 \\ r_1 & r_0 & r_1 \\ r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} L \\ 0 \\ 0 \end{bmatrix}, \tag{B.36}$$

where $a_1 = f_1/f_0$, $a_2 = f_2/f_0$, and $L = w_0/f_0$. We now have three unknowns, a_1, a_2, and L, and three equations. The square matrix elements on the left side of the equation represent the autocorrelation lags of the seismic wavelet, which we do not know. However, the autocorrelation lags from equation (B.18) of the seismogram that we do know can be substituted.

The autocorrelation matrix in equation (B.36) is special. First, it is a symmetric matrix; second, its diagonal elements are identical. This type of matrix is called a Toeplitz matrix. For n number of normal equations, the standard algorithms require a memory space that is proportional to n^2 and CPU time that is proportional to n^3. Because of the special properties of the Toeplitz matrix, Levinson devised a recursive scheme (Claerbout, 1976) that requires a memory space and CPU time proportional to n and n^2, respectively.

B.4 Frequency-Domain Deconvolution

We want to estimate a minimum-phase wavelet from the recorded seismogram. The inverse of the wavelet is the spiking deconvolution operator. We start with autocorrelation of the seismic wavelet $w(t)$ in the frequency domain:

$$R_w(\omega) = W(\omega)\bar{W}(\omega), \tag{B.37}$$

where $\overline{W}(\omega)$ is the complex conjugate of the Fourier transform of $w(t)$. Since $w(t)$ normally is not known, $R_w(\omega)$ is not known. However, based on an assumption of white reflectivity series [equation (B.18)], we can substitute the autocorrelation function of the seismogram in equation (B.37).

Define a new function $U(\omega)$:

$$U(\omega) = \ln\,[R_w(\omega)]. \tag{B.38}$$

When both sides of equation (B.38) are exponentiated:

$$R_w(\omega) = \exp\,[U(\omega)]. \tag{B.39}$$

Suppose that another function, $\phi(\omega)$, is defined and that equation (B.39) is rewritten as (Claerbout, 1976):

$$R_w(\omega) = \exp\,\{\tfrac{1}{2}[U(\omega) + i\phi(\omega)]\}\,\exp\,\{\tfrac{1}{2}[U(\omega) - i\phi(\omega)]\}. \tag{B.40}$$

When equation (B.40) is compared with equation (B.37), we see that

$$W(\omega) = \exp\,\{\tfrac{1}{2}[U(\omega) + i\phi(\omega)]\}. \tag{B.41}$$

We know $R_w(\omega)$ from equation (B.18) and therefore know $U(\omega)$ from equation (B.38). To estimate $W(\omega)$ from equation (B.41), we also need to know $\phi(\omega)$. If we make the minimum-phase assumption, $\phi(\omega)$ turns out to be the Hilbert transform of $U(\omega)$ (Claerbout, 1976). To perform the Hilbert transform, first inverse Fourier transform $U(\omega)$ back to the time domain. Then, double the positive time values, leave the zero-lag alone, and set the negative time values to zero. This operation yields a time function $u^+(t)$, which vanishes before $t = 0$. Then return to the transform domain to get

$$U^+(\omega) = \tfrac{1}{2}[U(\omega) + i\phi(\omega)], \tag{B.42}$$

where $U^+(\omega)$ is the Fourier transform of $u^+(t)$. Finally, exponentiating $U^+(\omega)$ yields the Fourier transform $W(\omega)$ of the minimum-phase wavelet $w(t)$ [equation (B.41)].

Once the minimum-phase wavelet $w(t)$ is estimated, its Fourier transform $W(\omega)$ is rewritten in terms of its amplitude and phase spectra,

$$W(\omega) = A(\omega)\exp\,[i\phi(\omega)]. \tag{B.43}$$

The deconvolution filter in the Fourier transform domain is:

$$F(\omega) = 1/W(\omega). \tag{B.44}$$

By substituting equation (B.43), we obtain the amplitude and phase spectra of this filter:

$$A_f(\omega) = 1/A(\omega), \tag{B.45a}$$

and

$$\phi_f(\omega) = -\phi(\omega). \tag{B.45b}$$

Since the estimated wavelet $w(t)$ is minimum phase, it follows that the deconvolution filter [whose Fourier transform is given by equation (B.44)] also is minimum phase. By inverse Fourier transforming, equation (B.44) yields the deconvolution operator. The frequency-domain method of deconvolution is outlined in Figure B-1.

Figure B-2 illustrates the frequency-domain deconvolution that is described in Figure B-1. The panels in Figure B-2 should be compared with the corresponding results of the Wiener-Levinson (time-domain) method in Figure 2-20. As expected, there is virtually no difference between the two results.

From the discussion so far, we see that the spiking deconvolution operator is the inverse of the minimum-phase equivalent of the seismic wavelet. (The process of estimating the minimum-phase equivalent of a seismic wavelet is called *spectral decomposition.*) The minimum-phase wavelet, once computed, can easily be inverted to get the deconvolution operator. To avoid dividing by zero in equation (B.45a) and to ensure that the filter is stable, a small number often is added to the amplitude spectrum before division. (This is called *prewhitening.*) Equation (B.45a) then takes the form:

$$A_f(\omega) = 1/[A(\omega) + \varepsilon]. \tag{B.46}$$

The amplitude spectrum also can be smoothed to get a more stable operator. Smoothing the spectrum is analogous to shortening the equivalent time-domain deconvolution operator.

Zero-phase deconvolution can be conveniently implemented in the frequency domain. To do this, the phase spectrum given by equation (B.45b) is set to zero and thus yields a deconvolution operator that flattens the amplitude spectrum of the input seismogram, but does not alter the phase.

B.5 Optimum Wiener Filters

The following concise discussion of optimum Wiener filters is based on Robinson and Treitel (1980). Consider the general filter model in Figure B-3. Wiener filtering involves designing the filter $f(t)$ so that the least-squares error between the actual and desired outputs is minimum. Error L is defined as

$$L = \sum_t (d_t - y_t)^2. \tag{B.47}$$

The actual output is the convolution of the filter with the input:

$$y_t = f_t * x_t. \tag{B.48}$$

When equation (B.48) is substituted into equation (B.47),

$$L = \sum_t \left(d_t - \sum_\tau f_\tau x_{t-\tau}\right)^2. \tag{B.49}$$

The goal is to compute the filter coefficients $(f_0, f_1, \ldots, f_{n-1})$ so that the error is minimum. Filter length n must be predefined. The minimum error is attained by setting the variation of L with respect to f_i to zero.

$$\frac{\partial L}{\partial f_i} = 0, \qquad i = 0, 1, 2, \ldots, (n-1). \tag{B.50}$$

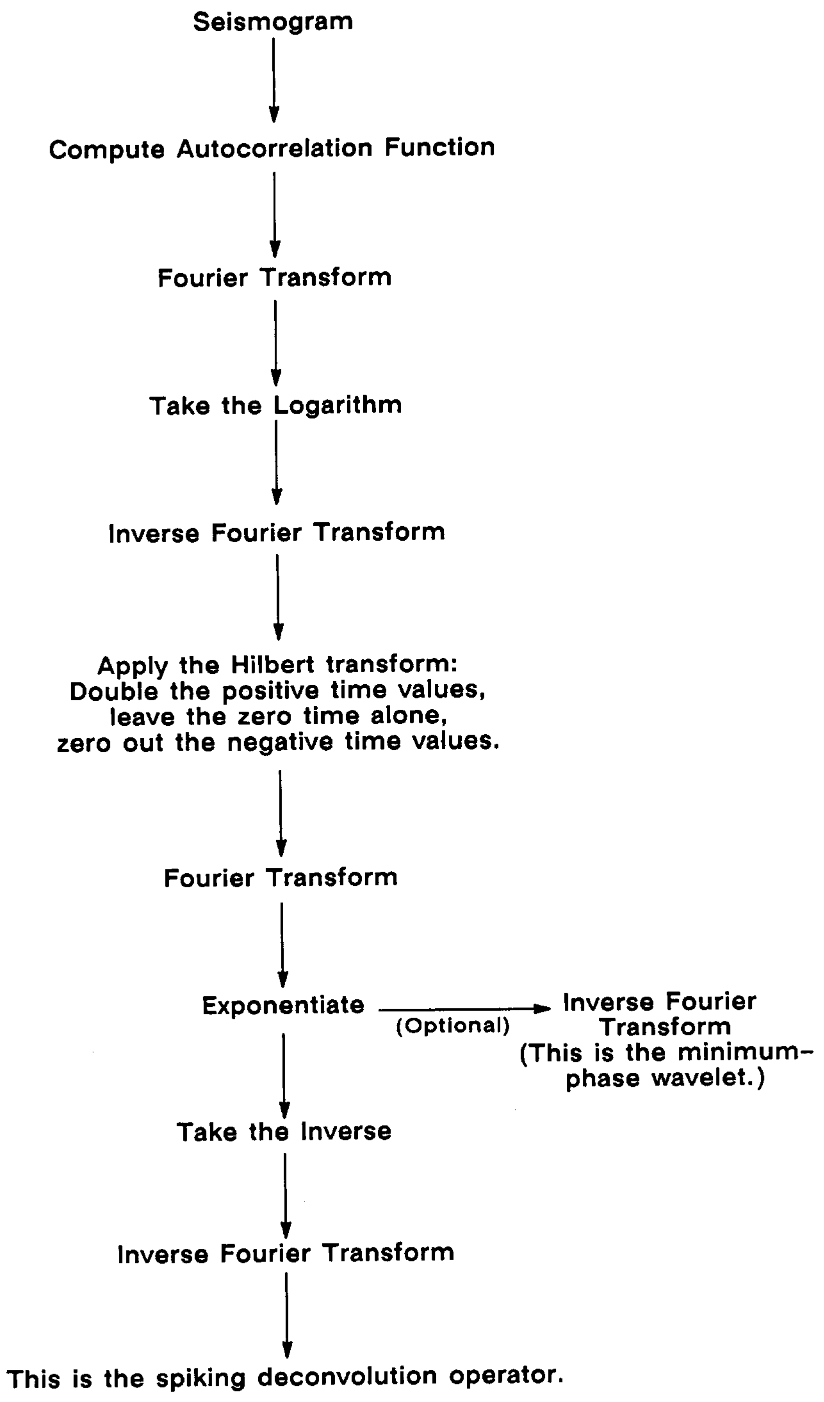


FIG. B-1. Flowchart for frequency-domain deconvolution.

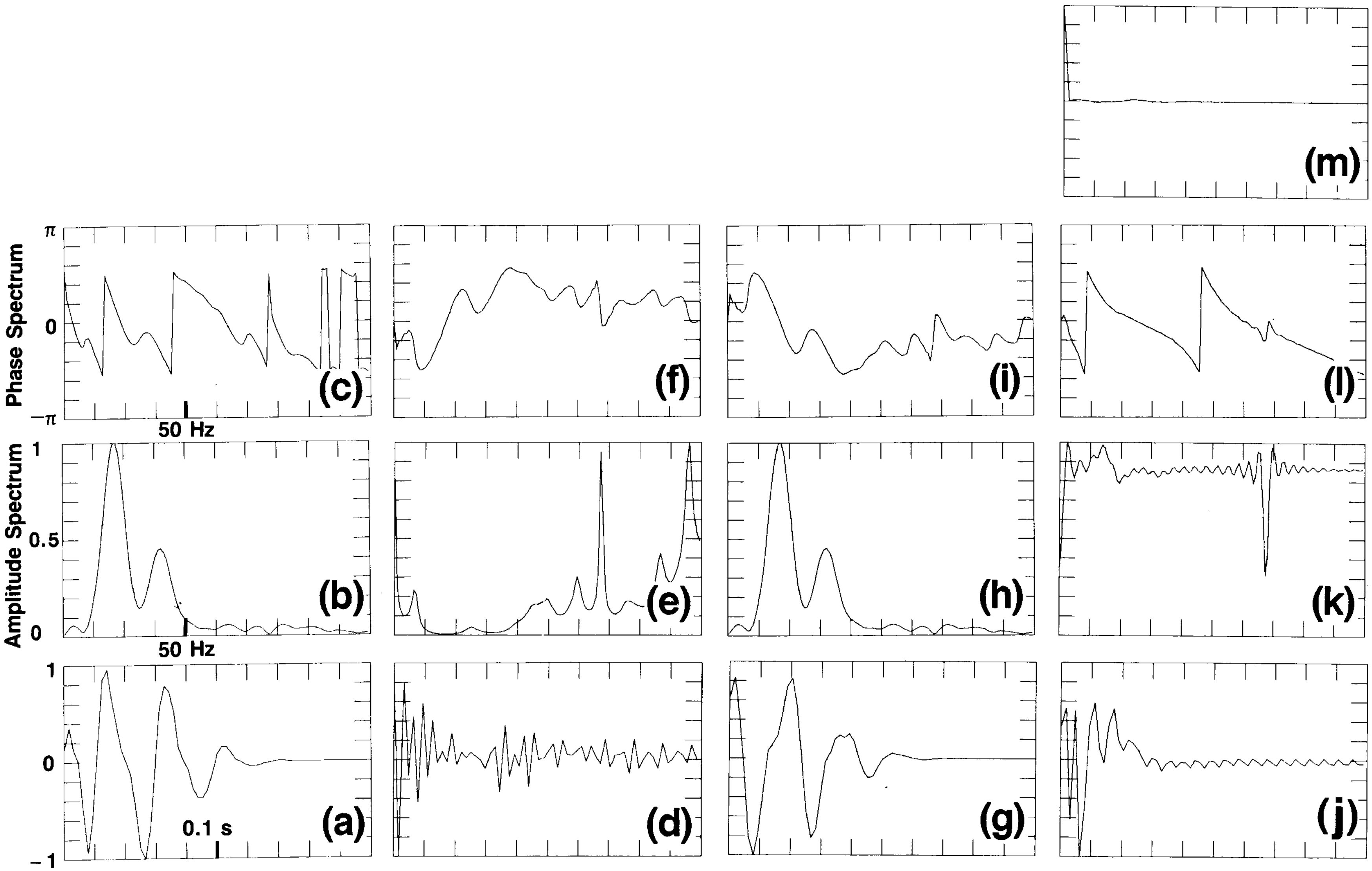


FIG. B-2. Frequency domain spiking deconvolution. (a) Starting with the mixed-phase wavelet (a), the spiking deconvolution operator (d) is derived via the procedure in Figure B-1. This operator is minimum phase; its inverse (g) also is minimum phase. The inverse has the same amplitude spectrum as that of the original wavelet [compare (h) with (b)]. Its phase spectrum is a sign-reversed phase spectrum of the deconvolution operator [compare (i) with (f)]. Convolve operator (d) with original wavelet (a) to yield output (j). Although the spectrum of the output is nearly flat (k), the output (j) is far from being a spike at $t = 0$. The desired output (m) is obtained only with a wavelet that is minimum phase (g), rather than mixed phase (a). Compare these results with those obtained by the time-domain approach for deconvolution as shown in Figure 2-20.

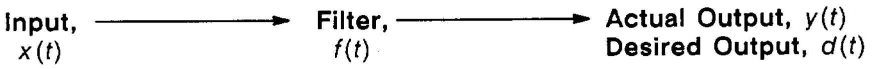


FIG. B-3. Wiener filter model.

By expanding the square term in equation (B.49), we have

$$L = \sum_t d_t^2 - 2\sum_t d_t \sum_\tau f_\tau x_{t-\tau} + \sum_t \left(\sum_\tau f_\tau x_{t-\tau}\right)^2. \tag{B.51}$$

By taking the partial derivatives and setting them to zero, we get

$$\frac{\partial L}{\partial f_i} = -2\sum_t d_t x_{t-i} + 2\sum_t \left(\sum_\tau f_\tau x_{t-\tau}\right) x_{t-i} = 0, \tag{B.52}$$

or

$$\sum_\tau f_\tau \sum_t x_{t-\tau} x_{t-i} = \sum_t d_t x_{t-i}, \quad i = 0, 1, 2, \ldots, (n-1). \tag{B.53}$$

By using

$$\sum_t x_{t-\tau} x_{t-i} = r_{i-\tau} \tag{B.54a}$$

and

$$\sum_t d_t x_{t-i} = g_i \tag{B.54b}$$

for each *i*th term, we get

$$\sum_\tau f_\tau r_{i-\tau} = g_i, \quad i = 0, 1, 2, \ldots, (n-1). \tag{B.55}$$

When put into matrix form, equation (B.55) becomes:

$$\begin{bmatrix} r_0 & r_1 & r_2 & \cdots & r_{n-1} \\ r_1 & r_0 & r_1 & \cdots & r_{n-2} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ r_{n-1} & r_{n-2} & r_{n-3} & \cdots & r_0 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_{n-1} \end{bmatrix} = \begin{bmatrix} g_0 \\ g_1 \\ \cdot \\ \cdot \\ \cdot \\ g_{n-1} \end{bmatrix}. \tag{B.56}$$

Here, r_i are the autocorrelation lags of the input and g_i are the lags of the crosscorrelation between the desired output and input. Since the autocorrelation matrix is Toeplitz, the optimum Wiener filter coefficients f_i can be computed by using Levinson recursion (Claerbout, 1976).

The least-squares error involved in this process now is computed. Expressed again, equation (B.51) becomes:

$$L_{\min} = \sum_t d_t^2 - 2\left(\sum_t \sum_\tau d_t x_{t-\tau} f_\tau\right) + \sum_t \left(\sum_\tau f_\tau x_{t-\tau}\right)^2. \tag{B.57}$$

By substituting the following relationships:

$$\sum_t d_t x_{t-\tau} = g_\tau \tag{B.58a}$$

and

$$\sum_t x_t x_{t-\tau} = r_\tau \tag{B.58b}$$

into equation (B.57), we get

$$L_{\min} = \sum_t d_t^2 - 2\sum_\tau f_\tau g_\tau + \sum_\tau f_\tau \sum_i f_i \sum_t x_{t-\tau} x_{t-i}, \tag{B.59}$$

or

$$L_{\min} = \sum_t d_t^2 - 2\sum_\tau f_\tau g_\tau + \sum_\tau f_\tau \sum_i r_{\tau-i} f_i. \tag{B.60}$$

Finally, by using equation (B.55), we get

$$L_{\min} = \sum_t d_t^2 - \sum_\tau f_\tau g_\tau. \tag{B.61}$$

Suppose that the desired output in the filter model of Figure B-3 is a time-advanced version of the input, $d(t) = x(t+\alpha)$. We want to design a Wiener filter $f(t)$ that predicts $x(t+\alpha)$ from the past values of the input $x(t)$. In this special case, the crosscorrelation function g becomes

$$g_\tau = \sum_t d_t x_{t-\tau} = \sum_t x_{t+\alpha} x_{t-\tau} = \sum_t x_t x_{t-(\alpha+\tau)}. \tag{B.62}$$

By definition, we have

$$r_\tau = \sum_t x_t x_{t-\tau}. \tag{B.63}$$

For the $\alpha + \tau$ lag, equation (B.63) becomes

$$r_{\alpha+\tau} = \sum_t x_t x_{t-(\alpha+\tau)} = g_\tau. \tag{B.64}$$

By substituting into equation (B.56), we get the set of normal equations that must be solved to find the *prediction filter* $(f_0, f_1, \ldots, f_{n-1})$:

$$\begin{bmatrix} r_0 & r_1 & r_2 & \cdots & r_{n-1} \\ r_1 & r_0 & r_1 & \cdots & r_{n-2} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ r_{n-1} & r_{n-2} & r_{n-3} & \cdots & r_0 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_{n-1} \end{bmatrix} = \begin{bmatrix} r_\alpha \\ r_{\alpha+1} \\ \cdot \\ \cdot \\ \cdot \\ r_{\alpha+n-1} \end{bmatrix}. \tag{B.65}$$

For a unit prediction lag, $\alpha = 1$, equation (B.65) takes the form:

$$\begin{bmatrix} r_0 & r_1 & r_2 & \cdots r_{n-1} \\ r_1 & r_0 & r_1 & \cdots r_{n-2} \\ \cdot & \cdot & \cdot & \cdots\cdot \\ \cdot & \cdot & \cdot & \cdots\cdot \\ \cdot & \cdot & \cdot & \cdots\cdot \\ r_{n-1} & r_{n-2} & r_{n-3} & \cdots r_0 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_{n-1} \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ \cdot \\ \cdot \\ \cdot \\ r_n \end{bmatrix}. \qquad \text{(B.66)}$$

By augmenting the right side to the square matrix on the left side, we have

$$\left[\begin{array}{c|cccc} -r_1 & r_0 & r_1 & \cdots & r_{n-1} \\ -r_2 & r_1 & r_0 & \cdots & r_{n-2} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ -r_n & r_{n-1} & r_{n-2} & \cdots & r_0 \end{array}\right] \begin{bmatrix} 1 \\ f_0 \\ f_1 \\ \cdot \\ \cdot \\ f_{n-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}. \qquad \text{(B.67)}$$

By adding one row, then putting the negative sign to the filter column, we obtain:

$$\left[\begin{array}{c|cccc} r_0 & r_1 & \cdot & \cdot\ \cdot & r_n \\ \hline r_1 & r_0 & \cdot & \cdot\ \cdot & r_{n-1} \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ r_n & r_{n-1} & \cdot & \cdot\ \cdot & r_0 \end{array}\right] \left[\begin{array}{c} 1 \\ \hline -f_0 \\ \cdot \\ \cdot \\ \cdot \\ -f_{n-1} \end{array}\right] = \left[\begin{array}{c} L \\ \hline 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{array}\right]. \qquad \text{(B.68)}$$

We now have $n + 1$ equations and $n + 1$ unknowns; e.g., $(f_0, f_1, \ldots, f_{n-1}, L)$. From equation (B.68)

$$L = r_0 - r_1 f_0 - r_2 f_1 - \cdots - r_n f_{n-1}. \qquad \text{(B.69)}$$

By using equation (B.61), the minimum error associated with the unit prediction-lag filter can be computed as follows. Start with

$$d_t = x_{t+1} \qquad \text{(B.70)}$$

so that

$$\sum_t d_t^2 = r_0 \qquad \text{(B.71a)}$$

and

$$g_t = r_{t+1}. \qquad \text{(B.71b)}$$

Substituting into equation (B.61), we get:

$$L_{\min} = r_0 - (r_1 f_0 + r_2 f_1 + \cdots + r_n f_{n-1}), \qquad \text{(B.72)}$$

which is identical to quantity L given by equation (B.69). Therefore, when we solve equation (B.68), we compute both the minimum error and the filter coefficients.

Equation (B.65) gives us the prediction filter with prediction lag α. The desired output is $x_{t+\alpha}$. The actual output is $\hat{x}_{t+\alpha}$, which is an estimate of the desired output. This desired output is the predictable component of the input series; i.e., periodic events such as multiples. The error series contains the unpredictable component of the series and is defined as:

$$e_{t+\alpha} = x_{t+\alpha} - \hat{x}_{t+\alpha} = x_{t+\alpha} - \sum_\tau f_\tau x_{t-\tau}. \qquad \text{(B.73)}$$

We assume that the unpredictable component e_t is the uncorrelated reflectivity we want to extract from the seismogram. By taking the z-transform of equation (B.73), we have

$$z^{-\alpha}E(z) = z^{-\alpha}X(z) - F(z)X(z), \qquad \text{(B.74)}$$

or

$$E(z) = [1 - z^\alpha F(z)]X(z). \qquad \text{(B.75)}$$

Now define a new filter $a(t)$ whose z-transform is

$$A(z) = 1 - z^\alpha F(z) \qquad \text{(B.76)}$$

so that, when substituted into equation (B.75), we have

$$E(z) = A(z)X(z). \qquad \text{(B.77)}$$

The corresponding time-domain relationship is

$$e(t) = a(t) * x(t). \qquad \text{(B.78)}$$

After defining $e(t)$ as reflectivity, equation (B.78) says that by applying filter $a(t)$ to the input seismogram $x(t)$, we obtain the reflectivity series. Since computing the reflectivity series is a goal of deconvolution, prediction filtering can be used for deconvolution. The time-domain form $a(t)$ of the filter defined by equation (B.76) is:

$$a_t = (1, \overbrace{0, 0, \cdots, 0}^{\alpha - 1}, -f_0, -f_1, \ldots, -f_{n-1}). \qquad \text{(B.79)}$$

The filter $a(t)$ is obtained from the prediction filter $f(t)$, which is the solution to equation (B.65). We call $a(t)$ the *prediction error filter*. For unit-prediction lag, the filter coefficients given by equation (B.79) are of the form:

$$a_t = (1, -f_0, -f_1, \ldots, -f_{n-1}). \qquad \text{(B.80)}$$

This is the same as the solution of equation (B.68). Moreover, equation (B.68) is equivalent to equation (B.36) for $n = 2$. Thus, we can conclude that a prediction error filter with unit-prediction lag and with $n + 1$ length is equivalent to an inverse filter of the same length except for a scale factor.

B.6 Surface-Consistent Deconvolution

Deconvolution can be formulated as a surface-consistent spectral decomposition (Taner and Coburn, 1981). In such a formulation, the seismic trace is decomposed into the convolutional effects of source, receiver, offset, and the earth's impulse response, thus explicitly accounting for variations in wavelet shape due to near-source and near-receiver conditions and source-receiver separation. Decomposition is followed by inverse filtering to recover the earth's impulse response. The assumption of surface-consistency implies that the basic wavelet shape depends only on the source and receiver locations, not on the details of the raypath from source to reflector to receiver.

The convolutional model discussed in Section 2.2 was given by:

$$x(t) = w(t) * e(t) + n(t), \tag{B.81}$$

where *x(t)* is the recorded seismogram, *w(t)* is the source waveform, *e(t)* is the earth's impulse response that we want to estimate, and *n(t)* is the noise component. A postulated surface-consistent convolutional model is given by:

$$x_{ij}(t) = s_j(t) * h_{(i-j)/2}(t) * e_{(i+j)/2}(t) * q_i(t) + n(t), \tag{B.82}$$

where $x_{ij}(t)$ is the seismogram, $s_j(t)$ is the waveform component associated with source location *j*, $q_i(t)$ is the component associated with receiver location *i*, and *h(t)* is the component associated with offset dependence of the waveform. As in equation (B.81), *e(t)* represents the earth's impulse response at the source-receiver midpoint location, $(i + j)/2$. By comparing equations (B.81) and (B.82), we infer that *w(t)* represents the combined effects of *s(t)*, *h(t)*, and *q(t)*.

To illustrate a method of computing *s*, *h*, *e*, and *q*, assume $n(t) = 0$ and Fourier transform equation (B.82):

$$X(\omega) = S(\omega)H(\omega)E(\omega)Q(\omega). \tag{B.83}$$

This equation can be separated into the following amplitude and phase spectral components:

$$A_x(\omega) = A_s(\omega)A_h(\omega)A_e(\omega)A_q(\omega), \tag{B.84a}$$

and

$$\phi_x(\omega) = \phi_s(\omega) + \phi_h(\omega) + \phi_e(\omega) + \phi_q(\omega). \tag{B.84b}$$

If the minimum-phase assumption is made, only the amplitude spectra need to be considered. Equation (B.84a) now can be linearized by taking the logarithm of both sides:

$$\ln A_x = \ln A_s + \ln A_h + \ln A_e + \ln A_q. \tag{B.85}$$

The left side is the logarithm of the amplitude spectrum of the modeled input trace. The individual filters now can be computed by least-squares error minimization. The least-squares error energy is defined as:

$$L = \sum_{i,j,\omega} (\ln A_x - \ln \hat{A}_x)^2, \tag{B.86}$$

where $\hat{A}_x$ is the actual amplitude spectrum.

Quantity *L* is minimized by requiring that

$$\frac{\partial L}{\partial(\ln A_s)} = \frac{\partial L}{\partial(\ln A_h)} = \frac{\partial L}{\partial(\ln A_e)} = \frac{\partial L}{\partial(\ln A_q)} = 0. \tag{B.87}$$

This requirement yields a set of normal equations whose solution provides the individual spectral components that are associated with the source and receiver locations, offset dependency, and earth's impulse response. The surface-consistent deconvolution operator is then the minimum-phase inverse of *s(t)*q(t)*.

REFERENCES

Claerbout, J. F., 1976, Fundamentals of geophysical data processing: McGraw-Hill Book Co.

Robinson, E. A., 1967, Multichannel z-transforms and minimum-delay: Geophysics, **31**, 482-500.

Robinson, E. A. and Treitel, S., 1980, Geophysical signal analysis: Prentice-Hall Book Co.

Taner, M. T. and Coburn, K., 1981, Surface-consistent deconvolution: Presented at the 51st Ann. Int. Soc. Explor. Geophys. Mtg.

Treitel, S. and Robinson, E. A., 1966, The design of high resolution filters: Inst. Electr. Electron. Eng., **GE-4**, *1*.

Yilmaz, O., 1974, The problems of resolution and reverberations in reflection seismology: J. Geophys. Soc. Turkey, 5, *2*.

Appendix C
MATHEMATICAL FOUNDATION OF MIGRATION

C.1 Wave Field Extrapolation and Migration

A fundamental equation of reflection seismology is the double square-root (DSR) equation. This equation describes downward continuation of both shots and receivers into the earth. It is exact in that it can handle all ranges of dip and offset angles. Neglecting the velocity gradient $dv(z)/dz$ makes the DSR equation also applicable to a stratified earth. The DSR equation can be extended, with some approximation, to treat weak lateral velocity variations. A comprehensive mathematical development of the DSR equation is found in Claerbout (1985).

The basic two-dimensional (2-D) theory for wave field extrapolation is presented here. Then, using the DSR equation, a rigorous analysis of conventional seismic data processing is made. We show that conventional implementation of the DSR equation requires zero-dip and zero-offset assumptions.

Before discussing the DSR equation, we review the basic theory for wave field extrapolation. Once the extrapolation equations are developed, they can be used with the imaging principle to migrate 2-D or three-dimensional (3-D) prestack and poststack data. Start with the 2-D scalar wave equation, which describes propagation of compressional wave fields $P(x,z,t)$ in a medium with constant material density and compressional wave velocity $v(x,z)$:

$$\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial z^2}-\frac{1}{v^2}\frac{\partial^2}{\partial t^2}\right)P(x, z, t) = 0, \qquad \text{(C.1)}$$

where x is the horizontal spatial axis, z is the depth axis (positive downward), and t is time. Given the upcoming seismic wave field $P(x,0,t)$, which is recorded at the surface, we want to determine reflectivity $P(x,z,0)$. Determining reflectivity requires extrapolating the surface wave field to depth z, then collecting it at $t = 0$. (Collecting the wavefield at $t = 0$ is equivalent to invoking the imaging principle.)

It is advantageous to decompose the wave field into monochromatic plane waves with different angles of propagation from the vertical. Therefore, we will work in the Fourier transform domain whenever possible. The wave field can always be Fourier transformed over time t. If there is no lateral velocity variation, then the wave field also can be Fourier transformed over the horizontal axis x. Thus,

$$P(k_x, z, \omega) = \iint P(x, z, t)\exp(ik_x x - i\omega t)\,dx\,dt, \qquad \text{(C.2a)}$$

and inversely,

$$P(x, z, t) = \iint P(k_x, z, \omega)\exp(-ik_x x + i\omega t)\,dk_x\,d\omega. \qquad \text{(C.2b)}$$

When the differential operator in equation (C.1) is applied to equation (C.2b), we get

$$\frac{\partial^2}{\partial z^2}P(k_x, z, \omega) + \left(\frac{\omega^2}{v^2} - k_x^2\right)P(k_x, z, \omega) = 0. \qquad \text{(C.3)}$$

Although v can be varied with depth z in equation (C.3), for now we assume the constant velocity case. The stratified earth case is considered later in this appendix. Equation (C.3) has two solutions, one for upcoming waves, the other for downgoing waves. The upcoming wave solution to equation (C.3) is recognized as

$$P(k_x, z, \omega) = P(k_x, 0, \omega)\exp\left[-i\left(\frac{\omega^2}{v^2} - k_x^2\right)^{1/2} z\right]. \qquad \text{(C.4)}$$

Equation (C.4) also is the solution to the following one-way wave equation:

$$\frac{\partial}{\partial z}P(k_x, z, \omega) = -i\left(\frac{\omega^2}{v^2} - k_x^2\right)^{1/2} P(k_x, z, \omega). \qquad \text{(C.5)}$$

This solution can be verified by substituting equation (C.4) into equation (C.5). We define the vertical wavenumber as

$$k_z = \frac{\omega}{v}\left[1 - \left(\frac{vk_x}{\omega}\right)^2\right]^{1/2}. \qquad \text{(C.6)}$$

Equation (C.6) often is called the *dispersion relation* of the one-way scalar wave equation. By using this expression, equation (C.4) takes the simple form

$$P(k_x, z, \omega) = P(k_x, 0, \omega)\exp(-ik_z z). \qquad \text{(C.7)}$$

To determine the reflectivity $P(x,z,0)$ from the wave field recorded at the earth's surface $P(x,0,t)$, proceed as follows: First, do a 2-D Fourier transform over x and t to get $P(k_x,0,\omega)$. Then multiply by the all-pass filter $\exp(-ik_z z)$ to obtain the wave field $P(k_x,z,\omega)$ at depth z. Subsequent summation over ω and inverse Fourier transformation over k_x yields the earth's image $P(x,z,0)$ at that depth. For the constant velocity case $P(k_x,k_z,0)$ can be computed by a direct mapping in the transform domain from (k_x,ω) to (k_x,k_z) using equation (C.6) (Stolt, 1978).

The main objective here is to interpret equation (C.7) as a tool for downward extrapolating wave fields given at the surface. While the mathematical development of the process presented is simple, its physical basis is not obvious. In an effort to develop a physical motivation for equation (C.7), a simpler derivation follows.

Given the upcoming wave field recorded at surface $P(x,0,t)$, we can decompose this wave field into monochromatic plane waves, each traveling at a different angle from the vertical. We identify these plane waves by attaching them to a unique (k_x,ω) pair of numbers. This plane-wave decomposition is equivalent to Fourier transforming the wave field to yield $P(k_x,0,\omega)$.

Now consider one of these plane waves as shown in Figure C-1. Imagine that this plane wave passed point P at $t = 0$, traveled upward, and was recorded by a receiver at surface point G at time t. To map the reflectors, we need to take the energy located at point G on the wavefront at time t, back to its position at $t = 0$; i.e., to reflection point P. To return the energy to P, it makes sense to follow the same ray path used for outward propagation from P. The fact that the same path is used means that downward continuation does not alter the horizontal wavenumber k_x. Suppose that the wavefront is moved to a depth $\Delta z = GG'$ beneath the receiver at G so that the waveform at G now is at G″. If a receiver were buried at G″, it would have recorded the plane wave at $t - \Delta t$, where Δt is the traveltime between G and G″. In other words, moving the receiver at G vertically down a distance Δz to a new location G′ changes the traveltime along the raypath from G to P by $-\Delta t$.

From the geometry of Figure C-1, we have

$$\Delta t = \frac{\Delta z}{v} \cos \theta, \qquad \text{(C.8)}$$

where $v/\cos\theta$ is the vertical phase velocity. We know the k_x and ω values for the plane wave. Suppose the distance between G and G″ is one wavelength, λ. At time $t - \Delta t$, the wavefront intersects the x axis at distance λ_x from G.

From the geometric relation in Figure C-1,

$$\lambda/\lambda_x = \sin \theta. \qquad \text{(C.9a)}$$

By using the definitions $\lambda = 2\pi/(\omega/v)$, $\lambda_x = 2\pi/k_x$, and equation (C.9a), we obtain

$$\sin \theta = v k_x/\omega, \qquad \text{(C.9b)}$$

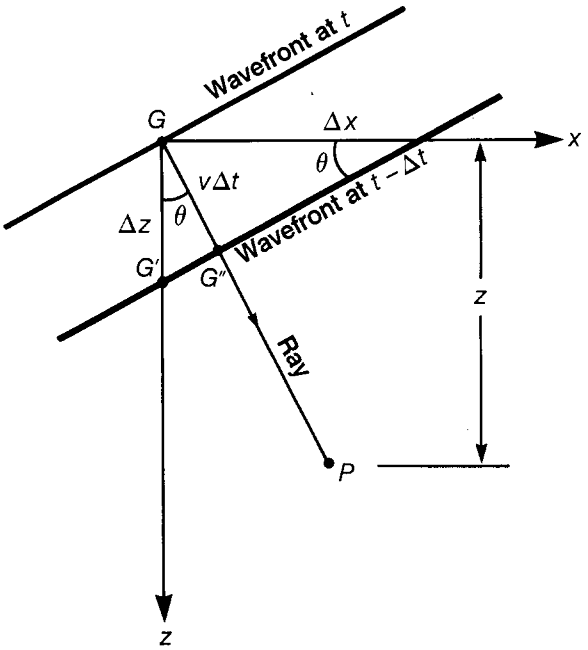


FIG. C-1. Geometry for wave field extrapolation (Yilmaz, 1979).

and

$$\cos \theta = \left[1 - \left(\frac{v k_x}{\omega}\right)^2\right]^{1/2}, \qquad \text{(C.9c)}$$

where ω/v is the wavenumber along the raypath. By substituting equation (C.9c) into equation (C.8), we have

$$\Delta t = \frac{1}{v}\left[1 - \left(\frac{v k_x}{\omega}\right)^2\right]^{1/2} \Delta z. \qquad \text{(C.10a)}$$

As we move down, we do not want to change the amplitude of the plane wave. Given the change in traveltime, $-\Delta t$ by equation (C.10a), the corresponding phase shift is $-\omega\Delta t$. At each Δz step of descent, we may propagate the plane wave with a different velocity $v(z)$. The total phase shift to which the waveform was subjected with arrival at P is $-\int \omega dt$. To compute the wave field at P, we use equation (C.10a) and multiply the transformed surface wave field $P(k_x, 0, \omega)$ by

$$\exp\left(-i\int_G^P \omega\, dt\right) = \exp\left\{-i\int_0^z \frac{\omega}{v(z)} \cdot \left[1 - \left(\frac{v(z) k_x}{\omega}\right)^2\right]^{1/2} dz\right\}. \qquad \text{(C.10b)}$$

Equation (C.10b) is the same operator used in equation (C.7), except that equation (C.7) was derived for constant v.

We now return to the more mathematical discussion and consider the stratified earth with a velocity $v(z)$. Since we have not Fourier transformed P over z, the one-way wave equation [equation (C.5)] also is valid for $v(z)$:

$$\frac{\partial}{\partial z} P(k_x, z, \omega) = -i\left[\frac{\omega^2}{v^2(z)} - k_x^2\right]^{1/2} P(k_x, z, \omega), \qquad \text{(C.11)}$$

in which case equation (C.6) becomes

$$k_z(z) = \frac{\omega}{v(z)}\left\{1 - \left[\frac{v(z) k_x}{\omega}\right]^2\right\}^{1/2}. \qquad \text{(C.12)}$$

Substitution verifies that equation (C.11) has the following solution:

$$P(k_x, z, \omega) = P(k_x, 0, \omega) \exp\left[-i\int_0^z k_z(z)\, dz\right]. \qquad \text{(C.13)}$$

We must check whether this solution satisfies the two-way scalar wave equation, equation (C.3). From equations (C.11) and (C.12),

$$\frac{\partial^2}{\partial z^2} P = -i \frac{dk_z(z)}{dv} \frac{dv(z)}{dz} P - i k_z(z) \frac{\partial}{\partial z} P,$$

where $P = P(k_x, z, \omega)$. By substituting equation (C.11) for

$\partial P/\partial z$, we get:

$$\frac{\partial^2}{\partial z^2} P = -i \frac{dk_z(z)}{dv} \frac{dv(z)}{dz} P - k_z^2(z)P.$$

If the velocity gradient $dv(z)/dz$ is ignored, then the first term on the right side (the amplitude term) drops out. The final expression then is

$$\frac{\partial^2}{\partial z^2} P + k_z^2(z)P = 0.$$

When equation (C.12) is substituted into this expression,

$$\frac{\partial^2}{\partial z^2} P + \left[\frac{\omega^2}{v^2(z)} - k_x^2\right] P = 0, \qquad \text{(C.14)}$$

which is identical to equation (C.3) in which velocity can be varied with depth z.

So far, we have shown that a 2-D wave field recorded at the earth's surface can be extrapolated downward using the phase-shift operator given by equation (C.10b). Wave extrapolation can be done through either a constant-velocity medium [equation (C.7)] or a vertically varying velocity medium [equation (C.13)]. Seismic imaging is not complete until a stopping condition is imposed on downward continuation. The process of downward continuation is terminated when the clock, which measures $t - \int dt$, reads zero traveltime.

The concepts described can be used to downward continue a complete seismic experiment that involves many shots and receivers. This downward continuation is achieved by the DSR operator. A derivation of the DSR operator is found in Claerbout (1985). By using the DSR operator instead of the single square-root expression for the vertical wavenumber given by equation (C.6), we have

$$k_z = \frac{\omega}{v} \text{DSR}(G, S), \qquad \text{(C.15)}$$

where v is the medium velocity that can be varied in depth z, and

$$\text{DSR}(G, S) = (1 - G^2)^{1/2} + (1 - S^2)^{1/2}, \qquad \text{(C.16)}$$

where G and S are the normalized receiver and shot wavenumbers, k_g and k_s, respectively:

$$G = vk_g/\omega, \qquad \text{(C.17a)}$$

and

$$S = vk_s/\omega. \qquad \text{(C.17b)}$$

The newly defined vertical wavenumber, equation (C.15), is inserted into the extrapolation equation (C.7) in which the input wavefield is $P(s,g,t,z = 0)$, the prestack data in the shot-receiver coordinates. This new extrapolation equation then can be used to downward continue common-shot gathers, and (by way of reciprocity) common-receiver gathers.

The DSR operator given by equation (C.16) is separable in terms of shot and receiver wavenumbers. This separation means that we can start with the wave fields recorded at the surface as common-shot gathers and use the first part of the DSR operator to downward continue the receivers to depth Δz. After this, we can reorganize the already downward-continued wave fields into common-receiver gathers and use the second part of DSR to downward continue the shots to depth Δz. By alternating between common-receiver and common-shot gathers, the entire seismic experiment (whole line) can be downward continued until imaging is accomplished.

Although no approximation (besides the stratified earth assumption) is made in this alternating downward continuation scheme, it is computationally exhausting. In fact, most of today's seismic data processing is done in midpoint-(half) offset (y,h) coordinates, rather than in shot-receiver (s,g) coordinates. Therefore, we will put DSR as defined by equation (C.16) into the (y,h) coordinates. The following coordinate transformation is required (Figure 1-40):

$$y = \tfrac{1}{2}(g + s), \qquad \text{(C.18a)}$$

and

$$h = \tfrac{1}{2}(g - s). \qquad \text{(C.18b)}$$

After the transformation (Claerbout, 1985), we obtain

$$G = Y + H, \qquad \text{(C.19a)}$$

and

$$S = Y - H, \qquad \text{(C.19b)}$$

where Y and H are the normalized midpoint and offset wavenumbers, respectively:

$$Y = vk_y/\omega, \qquad \text{(C.20a)}$$

and

$$H = vk_h/\omega. \qquad \text{(C.20b)}$$

By substituting equations (C.19a) and (C.19b) into equation (C.16), the following DSR operator in midpoint-offset coordinates results:

$$\text{DSR}(Y, H) = [1 - (Y + H)^2]^{1/2} + [1 - (Y - H)^2]^{1/2}. \qquad \text{(C.21)}$$

The vertical wavenumber [equation (C.15)] now is expressed in terms of normalized midpoint-offset wavenumbers Y and H:

$$k_z = \frac{\omega}{v} \text{DSR}(Y, H). \qquad \text{(C.22)}$$

Figure C-2 shows the response characteristics of the DSR operator, equation (C.22). Note the semielliptical wavefronts in the (y,z) plane for a single frequency; while in the (y,t) plane, note the table-top traveltime curves.

Is equation (C.21) an improvement over equation (C.16)? Note that in equation (C.16), the terms with different spatial wavenumbers are separable. However, we have lost the property of separation in equation (C.21), for the operators in Y and H are strongly coupled. As a result, the Taylor series expansions of the square roots in equation (C.21) yield terms that contain cross-products of the two wavenumbers. The penalty for processing in the conventional coordinate system (y,h,t) is that strong coupling in the extrapolation operator requires the entire prestack data set to be handled at the same time for each depth step.

Is there a connection between DSR processing and conventional processing in (y,h,t) space? The conventional processing sequence is comprised of two major components. First, the data are organized into common-midpoint (CMP) gathers and normal-moveout (NMO) correction is applied to each gather. The time shift associated with the NMO correction is given by equation (C.23):

$$\Delta t = t(h) - t(0) = t(0)\left\{\left[1 + \left(\frac{2h}{vt(0)}\right)^2\right]^{1/2} - 1\right\}, \qquad \text{(C.23)}$$

where $t(h)$ is the two-way traveltime for a given (half) offset h and $t(0)$ is the corresponding two-way zero-offset time. The total moveout correction is the difference between $t(h)$ and $t(0)$. Here, v is the root-mean-square (rms) velocity at t(0). Equation (C.23) is based on the stratified earth (zero-dip) assumption. After NMO correction, traces of the CMP gather are stacked. This not only reduces data volume, but also enhances the signal-to-noise (S/N) ratio.

Second, the CMP stack is migrated as if it were the zero-offset wave field generated by exploding reflectors (Section 4.1). The equation used for the downward extrapolation portion of migration is the one-way wave equation, equation (C.11). To account for the one-way traveltime of the exploding reflectors model, the velocity used in extrapolation is taken as half the medium velocity. Thus, the vertical wavenumber given by equation (C.12) is expressed as:

$$k_z = \frac{2\omega}{v}\left[1 - \left(\frac{vk_y}{2\omega}\right)^2\right]^{1/2}. \qquad \text{(C.24)}$$

where v can be varied with depth z. (Since migration is done in midpoint space, k_x has been replaced with k_y.) Some of the current migration techniques based on wave extrapolation use certain rational approximations to equation (C.24), while some implement the exact form in the frequency-wavenumber domain.

By now it is clear that conventional processing has an advantage over the exact theory represented by the DSR equation in midpoint offset space [equation (C.21)]. Unlike the latter, the conventional approach is composed of two separable operators; i.e., the NMO + stack applied in offset space and migration applied in midpoint space. However, this advantage is based on zero-dip and zero-offset assumptions. Where do we go from here? On the one hand, we have an exact theory that can handle all dips and offset angles, but is difficult to implement. On the other hand, we have a conventional approach that has the convenient property of separation, but is based on two assumptions, which may not be valid particularly in a region with a complex structural setting. To examine the relationship between the two approaches, we return to the exact theory and make the same two assumptions that underlie the conventional approach.

The zero-dip assumption implies that the earth model is stratified [in (y,t) space]. The seismic energy recorded over such an earth is concentrated completely at the zero midpoint wavenumber $k_y = 0$. This suggests that we set the normalized wavenumber Y equal to zero in DSR as defined by equation (C.21). The resulting operator here is defined as the stacking (St) operator:

$$\text{St}\,(H) = 2(1 - H^2)^{1/2} - 2. \qquad \text{(C.25)}$$

The -2 factor at first appears to be a *post facto* addition to St(H). The first part of St(H) takes each shot-receiver pair on a CMP gather down to a reflecting point following raypaths with $\cos\theta = (1 - H^2)^{1/2}$. Because normal moveout correction is designed to correct to the zero-offset time and not to $t = 0$, the two-way traveltime associated with the vertical path from the reflector to the surface must be explicitly included in St(H). The second term, -2, accounts for this path.

The NMO shift given by equation (C.23) is a stationary-phase approximation to equation (C.25) (Clayton, 1978). The St(H) operator condenses primary information on a CMP gather down to zero offset. After applying this operator on a CMP gather, we may keep the zero-offset trace and abandon all other offsets. Since a CMP stack can be regarded as a zero-offset wave field, equation (C.25) is a zero-dip NMO + stack-type operator.

Incorporation of the zero-offset ($h = 0$) assumption into the DSR operator is more subtle. On a CMP gather at and near $h = 0$, energy essentially is concentrated at zero value of the offset wavenumber, $k_h = 0$. In fact, NMO correction tries to push the primary energy on a CMP gather toward $k_h = 0$. Therefore, by setting the normalized offset wavenumber $H = 0$ in equation (C.21), the exploding reflectors (ER) migration operator can be expressed as:

$$\text{ER}\,(Y) = 2(1 - Y^2)^{1/2}. \qquad \text{(C.26)}$$

Equation (C.26) is an approximation, because setting $H = 0$ is not exactly the same as setting $h = 0$. By setting $H = 0$ in equation (C.22), we have the zero-offset vertical wavenumber:

$$k_z = \frac{2\omega}{v}(1 - Y^2)^{1/2}. \qquad \text{(C.27)}$$

By substituting the definition for $Y = vk_y/2\omega$ into equation (C.27), we obtain

$$k_z = \frac{2\omega}{v}\left[1 - \left(\frac{vk_y}{2\omega}\right)^2\right]^{1/2}, \qquad \text{(C.28)}$$

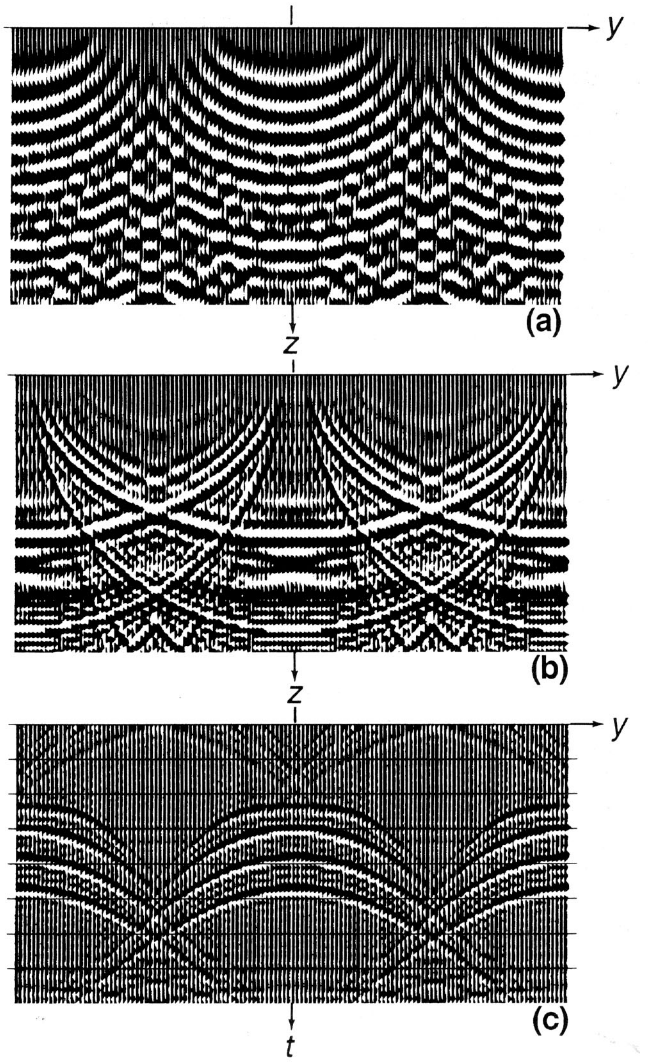


FIG. C-2. The response characteristics of the DSR operator [equation (C.22)] (Yilmaz, 1979). (a) Real part of the (y,z) plane at 16 Hz and $h = 400$ m. Note the semielliptical wavefronts. (b) Real part of the (y,z) plane at $t = 1024$ ms, $h = 400$ m. Because of the wraparound in h, we observe two wavefronts, one for $h = 400$ m and one for $h = 0$. (c) Real part of the (y,t) plane at $z = 200$, 400, 600, and 800 m superimposed. These are the table-top trajectories for $h = 400$ m. The loci of the arrival times are determined by a stationary-phase approximation to DSR (Clayton, 1978). (Periodicity in y and t result from approximating Fourier integrals by sums.)

which is identical to equation (C.24). We conclude that the zero-offset migration operator ER(Y) derived from the DSR equation is identical to the migration operator that is based on the exploding reflectors model of conventional processing.

Figure C-3 shows the response characteristics of the exploding reflectors operator, equation (C.28). Note the semicircular wavefronts in the (y,z) plane and the hyperbolic traveltime curves in the (y,t) plane. Compare these with the response of the complete DSR operator for the nonzero-offset case in Figure C-2.

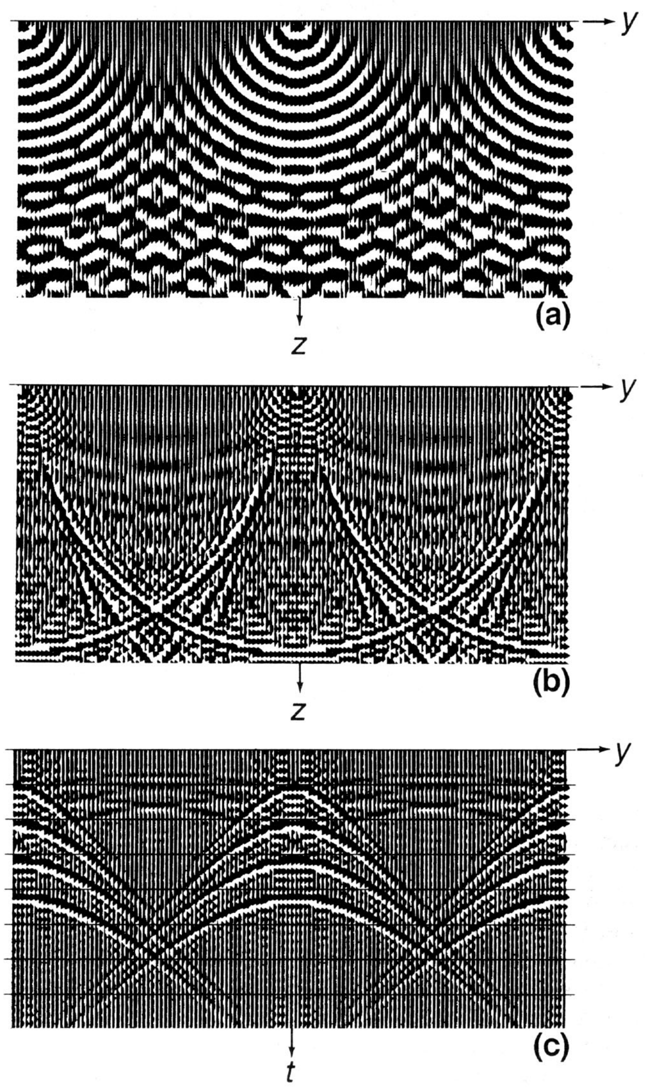


FIG. C-3. The response characteristics of the exploding reflectors operator ER(Y) [equation (C.27)] (Yilmaz, 1979). (a) Real part of the (y,z) plane at 16 Hz. Note the circular wavefronts. (b) Real part of the (y,z) plane at $t = 1024$ ms. (c) Real part of the (y,t) plane at $z = 200$, 400, 600, and 800 m superimposed. These are the hyperbolic trajectories. The loci of the arrival times are determined via the stationary-phase approximation to ER(Y) (Clayton, 1978). (Periodicity in y and t result from approximating Fourier integrals by sums.)

C.2 The Parabolic Approximation

When the 15-degree (parabolic) approximation of equation (C.26) is made by the Taylor expansion, we get

$$\text{DSR } (Y, H = 0) \approx 2(1 - Y^2/2), \qquad \text{(C.29)}$$

where Y is the normalized midpoint wavenumber defined by equation (C.20a). By substituting equations (C.29) and

(C.20a) into equation (C.22), we get the dispersion relation associated with the parabolic equation:

$$k_z = \frac{2\omega}{v} - \frac{vk_y^2}{4\omega}. \tag{C.30}$$

Operating on the wave field P and replacing $-ik_zP$ with $\partial P/\partial z$, we write the corresponding differential equation:

$$\frac{\partial P}{\partial z} = -i\left(\frac{2\omega}{v} - \frac{vk_y^2}{4\omega}\right)P. \tag{C.31}$$

Derivation of equation (C.30) is based on the constant-velocity assumption. Nevertheless, just as we did for the 90-degree one-way wave equation [equation (C.11)], the 15-degree one-way wave equation [equation (C.31)] can be recast using a vertically varying velocity function $v(z)$. Going one step further, once equation (C.31) is inverse Fourier transformed from hoizontal wavenumber k_y to horizontal axis y, we will replace $v(z)$ with a laterally varying velocity function $v(y,z)$. Theoretically, this may not be permissible, but in practice, it has a well-established validity.

The effect of translation is removed by retardation (Figure 4-30b):

$$\tau = 2\int_0^z \frac{dz}{\bar{v}(z)}, \tag{C.32}$$

where $\bar{v}(z)$ generally is chosen to be a horizontal average of $v(y,z)$. The time shift defined by equation (C.32) is equivalent to a phase shift in the frequency domain. Therefore, the actual wave field P is related to the time-shifted wave field Q by:

$$P = Q \exp(-i\omega\tau). \tag{C.33}$$

By differentiating with respect to z, we get:

$$\frac{\partial P}{\partial z} = \left(\frac{\partial}{\partial z} - i\frac{2\omega}{\bar{v}(z)}\right)Q \exp(-i\omega\tau). \tag{C.34}$$

By substituting equations (C.33) and (C.34) into equation (C.31), we get:

$$\frac{\partial Q}{\partial z} = i\frac{vk_y^2}{4\omega}Q + i2\omega\left[\frac{1}{\bar{v}(z)} - \frac{1}{v}\right]Q. \tag{C.35}$$

The first term on the right side of this equation is called the *diffraction term*; the second term is called the *thin-lens term*.

Consider the special case of $v = \bar{v}(z)$. The thin-lens term then vanishes and we are left with:

$$\frac{\partial Q}{\partial z} = i\frac{vk_y^2}{4\omega}Q. \tag{C.36}$$

After inverse Fourier transforming, we obtain the parabolic differential equation:

$$\frac{\partial^2 Q}{\partial z \partial t} = \frac{v}{4}\frac{\partial^2 Q}{\partial y^2}, \tag{C.37}$$

where, practice,
well as vertically.

In some practical implementations of equation (C.37), downward continuation is performed in τ rather than z. Migrated section then is displayed in time, τ (time migration). These two variables are related by equation (C.32). The dispersion relation in equation (C.28) for the scalar wave equation in terms of $\omega_\tau = vk_z/2$, the Fourier dual of τ, takes the form (Exercise 4.10):

$$\omega_\tau = \omega\left[1 - \left(\frac{vk_y}{2\omega}\right)^2\right]^{1/2} \tag{C.38a}$$

By squaring both sides, we get the equation of an ellipse in the (ω_τ, k_y) plane. The dispersion relation in equation (C.30) for the parabolic equation, also expressed in terms of ω_τ, takes the form

$$\omega_\tau = \omega - \frac{v^2k_y^2}{8\omega}, \tag{C.38b}$$

which is the equation of a parabola in the (ω_τ, k_y) plane. The first term is associated with a vertical time shift that can be removed by retardation. To obtain the differential equation [equivalent to equation (C.37)], apply the second term on the right side of equation (C.38b) to the retarded wave field Q and inverse Fourier transform:

$$\frac{\partial^2 Q}{\partial \tau \partial t} = \frac{v^2}{8}\frac{\partial^2 Q}{\partial y^2}. \tag{C.39}$$

This equation is the basis for the 15-degree time migration algorithms.

The dispersion relations, equations (C.38a) and (C.38b), are plotted in Figure C-4a for constant velocity v and input frequency $\omega = DE$. Curve 1 is associated with the 90-degree scalar wave equation and is the same as shown in Figure 4-34, except that the latter is in terms of k_z. Point A is mapped onto point B after migration with the 90-degree equation (C.38a). The same point A is mapped onto point C after migration with the 15-degree equation (C.38b). The dispersion curve for the parabolic equation (2) increasingly departs from the dispersion curve for the exact wave equation (1) as dip gets steeper. (The dip before migration is measured as the angle between the vertical axis and radial direction DA.) Thus, the parabolic equation causes more and more undermigration as dip increases.

From Figure C-4b, note the effect of velocity errors on the performance of the exact and parabolic equations. Correct mapping is from A to B for a given input frequency, $\omega = DE$. A list of mappings of point A by exact and parabolic equations using lower and higher velocities follows:

	$0.8v$	v	$1.2v$
Exact	L1	B	H1
Parabolic	L2	C	H2

Using too low a velocity causes undermigration, $AL1 < AB$, $AL2 < AB$, while using too high a velocity causes

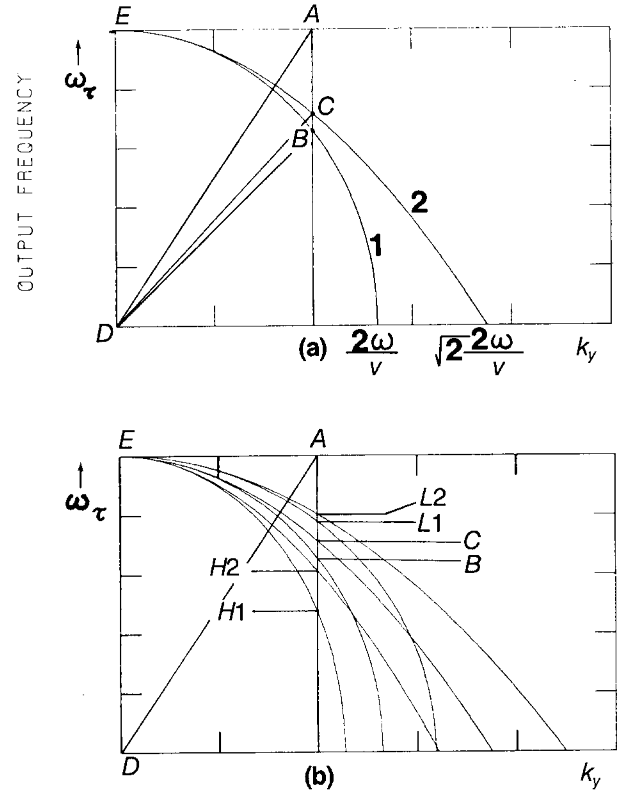


FIG. C-4. (a) Dispersion relations for the exact [equation (C.38a), Curve 1] and the parabolic [equation (C.38b), Curve 2] equations, plotted on the (ω_t, k_y) plane for a given input frequency, where ω = DE and velocity is v. (b) The curves in (a) are plotted for three different velocity values 0.8v ($L1$,$L2$), v (B,C), and 1.2v ($H1$,$H2$).

overmigration, $AH1 > AB$, $AH2 > AB$). The undermigration effect is more pronounced for the parabolic equation ($BL2 > BL1$), while the overmigration effect is more pronounced for the exact equation ($BH1 > BH2$). Refer to Section 4.2.2 for the practical aspects of the 15-degree finite-difference migration.

C.3 Steep-Dip Finite-Difference Migration

Higher-order approximations to the dispersion relation of the one-way equation can be found by continued fraction expansion (Claerbout, 1985). When equation (C.27) is rewritten, we have:

$$k_z = \frac{2\omega}{v}\,\mathrm{R}, \tag{C.40}$$

where

$$R = (1 - Y^2)^{1/2}. \tag{C.41}$$

The various orders of approximations to equation (C.41) are defined by the following recurrence relation:

$$R_{n+1} = 1 - \frac{Y^2}{1 + R_n} \tag{C.42}$$

with the initial value $R_0 = 1$. By setting $n = 0$ in equation (C.42), we have

$$R_1 = 1 - Y^2/2. \tag{C.43}$$

Then, substitution of equation (C.43) into equation (C.40) yields:

$$k_z = \frac{2\omega}{v}(1 - Y^2/2), \tag{C.44}$$

which is the same equation obtained with the parabolic approximation, equation (C.30). The next higher-order expansion is obtained by setting $n = 1$ in equation (C.42) and using equation (C.43):

$$R_2 = 1 - \frac{Y^2}{2 - \frac{Y^2}{2}}, \tag{C.45}$$

which is the 45-degree approximation to equation (C.41).

Ma (1981) discovered that the recurrence relation for the continuous fraction expansion can be expressed as ratios of two polynomials for the even-ordered expansions. He also showed that the expression can be split into the following partial fractions:

$$R_{2n} = 1 - \sum_{i=1}^{n} \frac{\alpha_i Y^2}{1 - \beta_i Y^2}. \tag{C.46}$$

For example, when $n = 1$,

$$R_2 = 1 - \frac{\alpha_1 Y^2}{1 - \beta_1 Y^2}, \tag{C.47}$$

which is equivalent to the 45-degree expansion in equation (C.45) when $\alpha_1 = 0.5$ and $\beta_1 = 0.25$. Note that expansions $R_4, R_6, R_8, \ldots$ are made up on sums of the 45-degree term, each with a different set of coefficients, α_i and β_i. Lee and Suh (1985) minimized the difference (in the least-squares sense) between R of equation (C.41) and R_{2n} of equation (C.46) for a specified dip angle and derived optimal coefficients (α_i, β_i) for up to the 10th order (Table C-1).

We now derive the differential equation associated with the 45-degree dispersion relation. By substituting equation (C.47) and the definition for $Y = vk_y/2\omega$ into equation (C.40), we get:

$$k_z = \frac{2\omega}{v}\left(1 - \frac{a_1 v^2 k_y^2/4\omega^2}{1 - \beta_1 v^2 k_y^2/4\omega^2}\right). \tag{C.48}$$

The first term is a vertical shift that can be removed by retardation, just as it was for the 15-degree approximation [equations (C.31) through (C.35)]. By simplifying the remaining terms of equation (C.48), we get:

$$\frac{\beta_1}{\alpha_1}\frac{v}{2\omega}k_z k_y^2 - k_y^2 - \frac{1}{\alpha_1}\frac{2\omega}{v}k_z = 0. \tag{C.49}$$

Table C-1. Coefficients of optimized, fractioned, one-way wave equations (Lee and Suh, 1985).

Order, $2n$	Degree of Accuracy	α_i	β_i
2	45	0.5	0.25
2	65	0.478242060	0.376369527
4	80	0.040315157	0.873981642
		0.457289566	0.222691983
6	87	0.004210420	0.972926132
		0.081312882	0.744418059
		0.414236605	0.150843924
8	90−	0.000523275	0.994065088
		0.014853510	0.919432661
		0.117592008	0.614520676
		0.367013245	0.105756624
10	90	0.000153427	0.997370236
		0.004172967	0.964827992
		0.033860918	0.824918565
		0.143798076	0.483340757
		0.318013812	0.073588213

By operating on the retarded wave field $Q(y, \omega, z)$ [equation (C.33)], we get:

$$i \frac{\beta_1}{\alpha_1 m} \frac{\partial^3 Q}{\partial z \partial y^2} - \frac{\partial^2 Q}{\partial y^2} + i \frac{m}{\alpha_1} \frac{\partial Q}{\partial z} = 0, \qquad \text{(C.50)}$$

where $m = 2\omega/v$. Kjartansson (1979) derived this equation differently and programmed it for 45-degree modeling and migration in the frequency-space domain. Migration in the frequency-space domain (commonly known as the omega-x algorithm) involves two interleaved operations: (a) a time shift based on equation (C.33), which is velocity-independent for time migration and velocity-dependent for depth migration, and (b) focusing the diffraction energy using equation (C.50).

Once you have a code for the basic 45-degree operator, it is easy to implement the higher-order approximations that are given by equation (C.46) with the associated coefficients in Table C-1. Note that the difference between the 45-degree and 65-degree algorithms is the values used for coefficients (α_1, β_1). Also note that the 15-degree approximation is obtained from equation (C.50) by setting $\alpha_1 = 0.5$ and $\beta_1 = 0$.

Figure 4-90 shows the impulse responses of various approximations to the one-way dispersion relation based on equation (C.46). Note that the wavefronts become increasingly closer to a semicircle as the higher-order terms are included in equation (C.46). The 15-degree approximation yields an elliptical wavefront. The 45-degree approximation yields an impulse response of a heart shape. Refer to Section 4.3.4 for the practical aspects of frequency-space steep-dip time migration, Section 5.1 for depth migration, and Section 6.5 for 3-D migration.

C.4 *F-K* Migration

A brief discussion of the mathematical formulation of the *f-k* migration techniques follows. We start with the solution of the scalar wave equation for the zero-offset wave field as given by equation (C.7) and assume a horizontally layered earth model $v(z)$. By inverse Fourier transforming equation (C.7), where k_x is replaced with k_y, we have:

$$P(y, z, t) = \iint P(k_y, 0, \omega) \exp(-ik_z z) \cdot \exp(-ik_y y + i\omega t)\, dk_y\, d\omega, \qquad \text{(C.51)}$$

where k_z is defined by equation (C.28). The imaging principle $t = 0$ then is applied to get the migrated section $P(y,z,t = 0)$,

$$P(y, z, t = 0) = \iint P(k_y, 0, \omega) \cdot \exp(-ik_y y - ik_z z)\, dk_y\, d\omega. \qquad \text{(C.52)}$$

This is the equation for phase-shift method (Gazdag, 1978). Equation (C.52) involves an integration over frequency and inverse Fourier transformation along midpoint axis y. (Refer to Figure 4-38 for a flow diagram of phase-shift migration.)

We now consider the special case of $v(z) = v =$ constant. Stolt (1978) devised a migration techniue that involves an efficient mapping in the 2-D Fourier transform domain from temporal frequency ω to the vertical wavenumber k_z. We rewrite equation (C.28) to get:

$$\omega = \frac{v}{2} (k_y^2 + k_z^2)^{1/2}. \qquad \text{(C.53)}$$

By differentiating and keeping the horizontal wavenumber k_y unchanged, we get:

$$d\omega = \frac{v}{2} \frac{k_z}{(k_y^2 + k_z^2)^{1/2}}\, dk_z. \qquad \text{(C.54)}$$

When equations (C.53) and (C.54) are substituted into equation (C.52), we get:

$$P(y, z, t = 0) = \iint \left[\frac{v}{2} \frac{k_z}{(k_y^2 + k_z^2)^{1/2}} \right] \cdot P\left[k_y, 0, \frac{v}{2} (k_y^2 + k_z^2)^{1/2} \right] \cdot \exp(-ik_y y - ik_z z)\, dk_y\, dk_z. \qquad \text{(C.55)}$$

This is the equation for constant-velocity Stolt migration. It involves two operations in the *f-k* domain. First, the temporal frequency ω is mapped onto the vertical wavenumber k_z via equation (C.53). This is the same as mapping point B′ onto point B in Figure 4-34. Second, the amplitudes are scaled by

$$\frac{v}{2} \frac{k_z}{(k_y^2 + k_z^2)^{1/2}}, \qquad \text{(C.56)}$$

which is equivalent to the obliquity factor associated with Kirchhoff migration (Section 4.2.1). (Refer to Figure 4-39 for a flow diagram of constant-velocity Stolt migration.)

To extend the algorithm to the variable-velocity case, yet retain efficiency, Stolt (1978) did a coordinate transformation that involves stretching the time axis to make the scalar wave equation velocity independent. A summary of the theoretical procedure is given here. Consider wave field $P(y,z,t)$ and the transformed wave field $P(y,d,T)$:

$$P(y, z, t) = P(y, d, T),$$

where T is the stretched time axis and d is the output variable (equivalent of z) for migration in the stretched coordinate system. Here, the y variable is identical in both coordinate systems. This coordinate transformation basically is equivalent to stretching the data using the rms velocities:

$$T(t) = \frac{1}{c}\left[2\int_0^t dt'\, v_{\text{rms}}^2(t')t'\right]^{1/2}, \tag{C.57}$$

where

$$v_{\text{rms}}^2(t) = \frac{1}{t}\int_0^t v^2(t')\, dt',$$

and c is an arbitrary reference velocity used to maintain the vertical axis as time after the coordinate transformation from t to T. After some tedious algebra, the time-retarded scalar wave equation takes the following form in the stretched coordinates (Stolt, 1978):

$$\frac{\partial^2 P}{\partial y^2} + W\frac{\partial^2 P}{\partial d^2} + \frac{2}{c}\frac{\partial^2 P}{\partial d \partial T} = 0, \tag{C.58}$$

where W is a complicated function of velocity and coordinate variables. In practice, it normally is set to a constant between 0 and 1. The procedure for Stolt migration with stretch follows:

(0) Start with the stacked section, which is assumed to be a zero-offset section $P(y,z = 0,t)$.
(1) Convert this time section to stretched section $P(y,d = 0,T)$ by the coordinate transformation [equation (C.57)].
(2) 2-D Fourier transform the stretched section $P(k_y, \mathrm{d} = 0, \omega_T)$.
(3) Apply the following mapping function to perform the migration:

$$k_d = \left(1 - \frac{1}{W}\right)\frac{\omega_T}{c} - \frac{1}{W}\left(\frac{\omega_T^2}{c^2} - Wk_y^2\right)^{1/2}. \tag{C.59}$$

This equation is based on the dispersion relation of the retarded wave equation in the stretched coordinates [equation (C.58)].

The expression for the wave field after migration is (omitting the heavy algebra):

$$\left\{\frac{c}{2-W}\left[(1 - W) + \frac{1}{[1 + (2 - W)k_y^2/k_d^2]^{1/2}}\right]\right]\right\}$$
$$\cdot P\left\{k_y, 0, \left[\frac{ck_d}{2-W}\left((1-W)+[1+(2-W)k_y^2/k_d^2]^{1/2}\right)\right]\right]\right\}. \tag{C.60}$$

(4) 2-D inverse Fourier transform the migrated section in the stretched coordinates $P(y,d,T = 0)$.
(5) Convert back to the familiar space-time coordinates $P(y,z,t = 0)$. This is the final migrated section.

When $W = 1$, equation (C.59) takes the simple form

$$k_d = \left(\frac{\omega_T^2}{c^2} - k_y^2\right),$$

which makes mapping of equation (C.59) equivalent to the constant-velocity Stolt algorithm.

Note that Stolt migration with stretching tries to handle velocity variations. However, it is no substitute for depth migration. Stolt only tried to accommodate for velocity variations that can be handled by time migration. See Section 4.3.3 for the practical aspects of *f-k* migration.

C.5 Residual Migration

The relationship between the output and input temporal frequencies for a 90-degree time migration operator [rewritten from equation (C.28)] is given by:

$$\omega_\tau = \left(\omega^2 - \frac{v^2k_y^2}{4}\right)^{1/2}, \tag{C.61}$$

where $\omega_\tau = vk_z/2$ is the output frequency, ω is the input frequency, v is the true medium velocity, and k_y is the midpoint wavenumber. Equation (C.61) refers to single-pass migration if the migration velocity is the same as the medium velocity. Consider a two-pass migration, first with velocity v_1, followed by a second migration with velocity v_2. The output frequency from the first pass is:

$$\omega_1 = \left(\omega^2 - \frac{v_1^2k_y^2}{4}\right)^{1/2}, \tag{C.62a}$$

which is the input frequency for the second pass:

$$\omega_2 = \left(\omega_1^2 - \frac{v_2^2k_y^2}{4}\right)^{1/2}. \tag{C.62b}$$

The horizontal wavenumber k_y is fixed, since it is invariant under migration. If the output frequency from the single-pass migration given by equation (C.61) is set equal to that from the two-pass migration given by equation (C.62b), then the necessary relationship between the residual migration velocity v_2 and the medium velocity v can be established:

$$v^2 = v_1^2 + v_2^2. \tag{C.63}$$

One practical scheme for residual migration is implemented as follows:

(1) The constant-velocity Stolt migration with a stretch factor of $W = 1$ is used for the first pass. Velocity [v_1 in equation (C.63)] is chosen as the minimum value in the actual velocity field.
(2) For the second pass, a (dip-limited) finite-difference

migration is used. The first-pass migration brings the dips down to within the values that the second-pass finite-difference migration can handle accurately. Remember that the velocity field for the second pass is computed by the the relation:

$$v(\text{second-pass}) = [v(\text{original})^2 - v(\text{first-pass})^2]^{1/2}.$$

Perhaps a relatively coarser depth step size can be used, since the perceived dip for the second migration pass is lowered because of the first pass migration. (Refer to Section 4.3.3 for the practical aspects of residual migration.)

C.6 Migration Velocity for the Parabolic Equation

From theory, we expect optimum migration velocities to be independent of dip for 90-degree algorithms such as Kirchhoff migration. This is not the case for dip-limited algorithms, such as the 15-degree finite-difference method. A simple frequency-domain derivation of an expression for dip-dependent migration velocities follows. It can be used to optimize the response of the 15-degree equation.

Start with the dispersion relation for the 90-degree equation [modified from equation (C.61)]:

$$\omega_\tau = \omega\left[1 - \left(\frac{v_{\text{med}}k_y}{2\omega}\right)^2\right]^{1/2}. \qquad \text{(C.64)}$$

The 15-degree equation is obtained by a Taylor series expansion of the square root:

$$\omega_{\tau,15^\circ} = \omega\left[1 - \frac{1}{2}\left(\frac{v_{\text{med}}k_y}{2\omega}\right)^2\right]. \qquad \text{(C.65)}$$

To match the output frequency $\omega_{\tau,15^\circ}$ with the desired output frequency ω_τ, we first replace v_{med} with $v_{\text{mig},15^\circ}$ in equation (C.65). We also substitute

$$\frac{k_y}{2\omega} = \frac{\sin\theta}{v_{\text{med}}},$$

where θ is the true dip. Equation (C.65) then becomes:

$$\omega_{\tau,15^\circ} = \omega\left[1 - \left(\frac{v_{\text{mig},15^\circ}}{v_{\text{med}}}\right)^2 \frac{\sin^2\theta}{2}\right]. \qquad \text{(C.66)}$$

To find the optimum velocity $v_{\text{mig},15^\circ}$ for equation (C.66), set the output frequencies from equations (C.64) and (C.66) equal:

$$v_{\text{mig},15^\circ} = \left(\frac{2}{1+\cos\theta}\right)^{1/2} v_{\text{med}}. \qquad \text{(C.67)}$$

For a 45-degree dip, the scale factor in equation (C.67) is 1.08, which suggests that when migrating a 45-degree dipping event with the 15-degree equation, you should use a velocity that is 8 percent greater than the medium velocity.

Equation (C.67) does not explicitly consider the errors and effects of finite-difference approximations. In practice, optimum migration velocities depend on the details of the finite-difference approximations used in a particular implementation. (Refer to the results of the velocity tests in Section 4.3.2.)

C.7 Migration Velocity Analysis

A method of migration velocity analysis based on wave field extrapolation is described in Section 4.5. The main computational steps of this method (Yilmaz and Chambers, 1984) are outlined below. We work with seismic data in midpoint-(half) offset (y,h) coordinates. We want to obtain a volume of focused energy at zero offset in (y,v,τ) coordinates from a prestack data set in (y,h,t) coordinates. For a midpoint location y, migration velocity function then can be picked from the corresponding (v,τ) plane.

First, a 3-D Fourier transformation is applied to the upcoming wave field $P(y,h,\tau = 0,t)$, which is recorded at the surface:

$$P(k_y, k_h, \tau = 0, \omega) = \iiint P(y, h, \tau = 0, t) \cdot \exp\,(ik_y y + ik_h h - i\omega t)\, dk_y\, dk_h\, d\omega, \qquad \text{(C.68)}$$

where t is the two-way traveltime and

$$\tau = 2\int \frac{dz}{v(z)} \qquad \text{(C.69)}$$

is the two-way vertical time equivalent of downward-continuation depth z in a medium with velocity $v(z)$. The variables (k_y, k_h, ω) are the Fourier duals of (y,h,t).

The surface wave field given by equation (C.68) then is extrapolated down to depth τ by

$$P(k_y, k_h, \tau, \omega) = P(k_y, k_h, \tau = 0, \omega)\exp\left(-i\frac{\omega}{2}\tau\text{DSR}\right), \qquad \text{(C.70)}$$

where

$$\text{DSR} \equiv [1 - (Y + H)^2]^{1/2} + [1 - (Y - H)^2]^{1/2} - 2, \qquad \text{(C.71)}$$

and Y and H are the normalized midpoint and offset wavenumbers given by equations (C.20a) and (C.20b). The -2 term puts the expression in retarded time form. [This term was not included in the previous definition of DSR given by equation (C.21).] Equation (C.70) is used recursively to extrapolate the wave field from one depth to another in steps of $\Delta\tau$.

Next, we transform the extrapolated wave field $P(k_y, k_h,$

τ, ω) into the space-time domain. In doing so, we only need to obtain the zero-offset information ($h = 0$). By summing equation (C.70) over k_h, we get the wave field at zero offset, $P(k\ , h = 0, \tau, \omega)$. By doing the 2-D inverse transform over (k_y, ω), we get

$$P(y, h = 0, \tau, t) = \iint P(k_y, h = 0, \tau, \omega) \cdot \exp(-ik_y y + i\omega t)\, dk_y\, d\omega. \qquad \text{(C.72)}$$

Here, $P(y, h = 0, \tau, t)$ is the zero-offset section at various depth levels from which we want to extract velocity information.

Suppose that velocity v_e were used to extrapolate the surface wave field down to depth τ. Equation (C.70) is written with v_e and τ as:

$$P(k_y, k_h, \tau, \omega) = P(k_y, k_h, \tau = 0, \omega) \cdot \exp\left[-i\frac{\omega}{2}\tau\mathrm{DSR}\,(v_e)\right]. \qquad \text{(C.73)}$$

Now suppose that the true medium velocity v were used to extrapolate the surface wave field down to depth $\tau = t$. By rewriting equation (C.70) with v and t, we have:

$$P(k_y, k_h, t, \omega) = P(k_y, k_h, \tau = 0, \omega) \cdot \exp\left[-i\frac{\omega}{2}t\mathrm{DSR}\,(v)\right]. \qquad \text{(C.74)}$$

Match the two extrapolated wave fields in equations (C.73) and (C.74) to get a relationship between v_e, τ, v, and t:

$$\tau\mathrm{DSR}\,(v_e) = t\mathrm{DSR}\,(v). \qquad \text{(C.75)}$$

Because of the complexity of DSR [equation (C.71)], equation (C.75) does not provide an explicit expression for v in terms of the other three variables τ, t, and v_e. However, we can get an approximate expression by expanding the square roots in Taylor series. The first-order expansion of equation (C.71) yields

$$\mathrm{DSR} \approx -Y^2 - H^2. \qquad \text{(C.76)}$$

By substituting equation (C.76) into equation (C.75) with the definitions in equations (C.20a) and (C.20b), then simplifying, the following approximate relationship results:

$$\tau v_e^2 = t v^2. \qquad \text{(C.77)}$$

This expression suggests that downward continuation [using the 15-degree approximate form of DSR, equation (C.76)] with the correct (medium) velocity to a wrong depth is equivalent to downward continuation to the correct depth with the wrong velocity (Doherty and Claerbout, 1974).

The derivation of equation (C.77) assumes that v_e is constant. When v_e is depth-variable, then the relationship in equation (C.77) still holds because equation (C.70) is valid for a stratified earth model. However, quantity v_e in that equation is replaced by the rms velocity:

$$v_{\mathrm{rms}}^2 = \frac{1}{\tau}\sum_k \Delta\tau v_e^2(k\Delta\tau). \qquad \text{(C.78)}$$

Because the approximation [equation (C.76)] is best for small ratios of offset-to-reflector depth, the accuracy of the mapping procedure based on equation (C.77) degrades at very shallow depths. (Refer to Section 4.5 for the practical considerations of the migration velocity estimation technique described in this appendix.)

C.8 3-D Migration

Although 3-D migration is discussed in Section 6.5, the mathematical development is given in this appendix. Equation (C.7) is rewritten for the 3-D case as follows:

$$P(k_x, k_y, z, \omega) = P(k_x, k_y, 0, \omega)\exp(-ik_z z), \qquad \text{(C.79)}$$

where

$$k_z = \frac{2\omega}{v}(1 - X^2 - Y^2)^{1/2}, \qquad \text{(C.80)}$$

$$X = vk_x/2\omega, \qquad \text{(C.81a)}$$

and

$$Y = vk_y/2\omega, \qquad \text{(C.81b)}$$

with x and y being the in-line and cross-line axes, respectively. Note that the in-line and cross-line normalized wavenumbers X and Y are coupled in equation (C.80).

First consider the constant-velocity case and rewrite equation (C.80) as for the 2-D case [equation (C.61)] to get:

$$\omega_\tau = \left(\omega^2 - \frac{v^2k_x^2}{4} - \frac{v^2k_y^2}{4}\right)^{1/2}, \qquad \text{(C.82)}$$

where $\omega_\tau = vk_z/2$ is the output temporal frequency.

Suppose we perform 2-D migration on all the in lines by mapping the input frequency ω to an output frequency ω_1 using the relation:

$$\omega_1 = \left(\omega^2 - \frac{v^2k_x^2}{4}\right)^{1/2}. \qquad \text{(C.83)}$$

We now sort the already migrated in lines in the cross-line direction. Then suppose that we again perform 2-D migration on all the cross lines. This second-pass migration is performed by mapping ω_1 of equation (C.83) to an output frequency ω_2 using the relation:

$$\omega_2 = \left(\omega_1^2 - \frac{v^2k_y^2}{4}\right)^{1/2}. \qquad \text{(C.84)}$$

When equation (C.84) is substituted into equation (C.83), we get:

$$\omega_2 = \left(\omega^2 - \frac{v^2k_x^2}{4} - \frac{v^2k_y^2}{4}\right)^{1/2}, \qquad \text{(C.85)}$$

which is the same as equation (C.82). This means that 3-D migration can be performed in two passes: 2-D migration in the in-line direction followed by 2-D migration in the cross-line direction. The above derivation of a 90-degree 3-D migration operator (Jakubowicz and Levin, 1983) is based on the constant-velocity assumption.

Mathematically, the two-pass approach is based on the idea of *full separation* (Claerbout, 1985) of migration operators in the in-line and cross-line directions. The separation is strictly valid for the constant-velocity medium. When velocities vary spatially, *splitting* (Claerbout, 1985) of the in-line and cross-line operators is necessary (Figure 6-25). A rigorous mathematical treatment of separation and splitting of extrapolation operators can be found in a paper by Brown (1983).

To implement separation or splitting, the in-line and cross-line wavenumbers need to be decoupled in the 3-D dispersion relation [equation (C.80)]. By Taylor series expansion of the square root and by keeping the first three terms, we get:

$$k_z \approx \frac{2\omega}{v}\left(1 - \frac{X^2}{2} - \frac{Y^2}{2}\right). \qquad \text{(C.86)}$$

When the cross-dip component is zero ($Y = 0$), equation (C.86) reduces to its 2-D form [equation (C.30)]. The in-line and cross-line terms in equation (C.86) are separable when velocity is slowly variable in z or independent of x and y (Brown, 1983). The impulse response of the 2-D 15-degree migration operator [equation (C.30)] is an ellipse (Figure 4-90). The impulse response of the 3-D 15-degree migration operator [equation (C.86)] is an ellipsoid (Figure 6-26). The 3-D 15-degree impulse response departs from the ideal 3-D impulse response (a hollow hemisphere) at dips greater than approximately 30 degrees.

Additional problems arise when more accuracy is desired; the X and Y terms in equation (C.80) no longer are decoupled. For example, the 45-degree approximation (Appendix C.3) to equation (C.80) introduces cross terms. Another approximation is to represent the single square root in equation (C.80) by two square roots (Ristow, 1980):

$$k_z \approx \frac{2\omega}{v}\left[(1 - X^2)^{1/2} + (1 - Y^2)^{1/2} - 1\right]. \qquad \text{(C.87)}$$

Note that equation (C.87) has the same form as equation (C.15). By Taylor series expansion of the two square roots, equation (C.87) reduces to equation (C.86). Assuming a small cross-dip component, more accuracy in the in-line direction is achieved by making the 45-degree approximation:

$$k_z \approx \frac{2\omega}{v}\left(1 - \frac{X^2}{2 - \frac{X^2}{2}} - \frac{Y^2}{2}\right). \qquad \text{(C.88)}$$

When more accuracy is required in both directions, we can use the following 45-degree expansion to equation (C.87) in both the in-line and cross-line directions:

$$k_z \approx \frac{2\omega}{v}\left(1 - \frac{X^2}{2 - \frac{X^2}{2}} - \frac{Y^2}{2 - \frac{Y^2}{2}}\right). \qquad \text{(C.89)}$$

The impulse response of the 3-D 45-degree migration operator [equation (C.89)] is shown in Figure 6-27. Note that the traveltimes are accurate in both the in-line and cross-line directions, and that the largest error (undermigration) occurs in the two diagonal directions. Horizontal cross sections (time slices) of the impulse responses for equations (C.86) and (C.87) are circular and diamond shaped, respectively.

In summary, to implement separation or splitting, the in-line and cross-line terms of the 3-D migration operator need to be decoupled. This can be achieved by making rational approximations [equation (C.86) or (C.89)] to the square root in equation (C.80). Separation works for 90-degree, constant-velocity (i.e., Stolt migration) or 15-degree, slowly varying $v(z)$ cases. When there are large vertical velocity gradients or strong lateral velocity variations, splitting must be used. (Refer to Section 6.5 for the practical aspects of 3-D migration.)

REFERENCES

Brown, D. L., 1983, Applications of operator separation in reflection seismology: Geophysics, **48**, 288-294.

Claerbout, J. F., 1985, Imaging the earth's interior: Blackwell Scientific Publications.

Clayton, R., 1978, Common midpoint migration: Stanford Exploration Project Report No. 14, Stanford University.

Doherty, S. M. and Claerbout, J. F., 1974, Velocity analysis based on the wave equation: Stanford Exploration Project Report No. 1, Stanford University.

Gazdag, J., 1978, Wave-equation migration by phase-shift: Geophysics, **43**, 1342-1351.

Jakubowicz, H. and Levin, S., 1983, A simple, exact method of 3-D migration-theory: Geophys. Props., **31**, 34-56.

Kjartansson, E., 1979, Modeling and migration by the monochromatic 45-degree equation: Stanford Exploration Project Report No. 15, Stanford University.

Lee, M. W. and Suh, S. Y., 1985, Optimization of one-way wave equations: Geophysics, **50**, 1634-1637.

Ma, Z., 1981, Finite-difference migration with higher order approximation: Presented at the 1981 Joint Mtg. Chinese Geophys. Soc. and Soc. Explor. Geophys.

Ristow, D., 1980, 3-D downward extrapolation of seismic data in particular by finite difference methods: Ph.D. thesis, University of Utrecht, The Netherlands.

Stolt, R. H., 1978, Migration by Fourier transform: Geophysics, **43**, 23-48.

Yilmaz, O., 1979, Prestack partial migration: Ph.D. thesis, Stanford University.

Yilmaz, O. and Chambers, R., 1984, Migration velocity analysis by wave field extrapolation: Geophysics, **49**, 1664-1674.

Appendix D

WAVE FIELD EXTRAPOLATION IN SLANT-STACK DOMAIN

Section 7.1 discussed transforming a wave field from midpoint-offset to midpoint-ray-parameter coordinates. This transformation is done by applying the appropriate linear moveout and summing over the offset axis for each value of the ray parameter (slant stacking). From the results in Appendix C.1, the double square-root operator can be specialized for migration before stack in midpoint-ray-parameter coordinates (Ottolini, 1983). To derive the extrapolation equation in the slant-stack domain, we start with the relationship between the transform-domain variables:

$$k_h/\omega = 2p, \qquad \text{(D.1)}$$

where k_h is the offset wavenumber, ω is the temporal frequency, and p is the ray parameter. The normalized offset wavenumber H is defined as:

$$H = vk_h/2\omega. \qquad \text{(D.2)}$$

Combining the two expressions yields:

$$H = pv. \qquad \text{(D.3)}$$

By substituting into the double square-root operator [equation (C.21) in Appendix C.1], we get:

$$\begin{aligned}\mathrm{DSR}\ (Y, H = pv) = {} & [1 - (Y + pv)^2]^{1/2} \\ & + [1 - (Y - pv)^2]^{1/2}. \qquad \text{(D.4)}\end{aligned}$$

Ottolini (1983) used this operator for migration before stack in midpoint-ray-parameter coordinates. The procedure is described in Figure D-1. Details on this approach to migration before stack are not given here. Instead, equation (D.4) is adapted to the zero-dip case. By setting $Y = 0$, we have:

$$\mathrm{DSR}\ (Y = 0, H = pv) = 2(1 - p^2v^2)^{1/2}. \qquad \text{(D.5)}$$

Clayton and McMechan (1982) used this operator to downward continue the refraction energy on CMP or shot gathers. Their procedure is outlined in Figure D-2.

The final step gives a profile of the horizontal phase velocity ($1/p$) as a function of depth. Two issues must be kept in mind. First, the procedure is based on layered earth assumption. Second, the procedure requires knowledge of the medium velocity to extrapolate the wave field in depth. To get around this second problem, the process must be iterated until the phase velocity profile converges to the velocity function used in extrapolation. If the signal-to-noise ratio is fair, it may take up to three iterations to achieve convergence.

REFERENCES

Clayton, R. and McMechan, G., 1982, Inversion of refraction data by wave field extrapolation: Geophysics, **46**, 860-868.

Ottolini, R., 1983, Migration of seismic data in angle-midpoint coordinates: Ph.D. thesis, Stanford University.

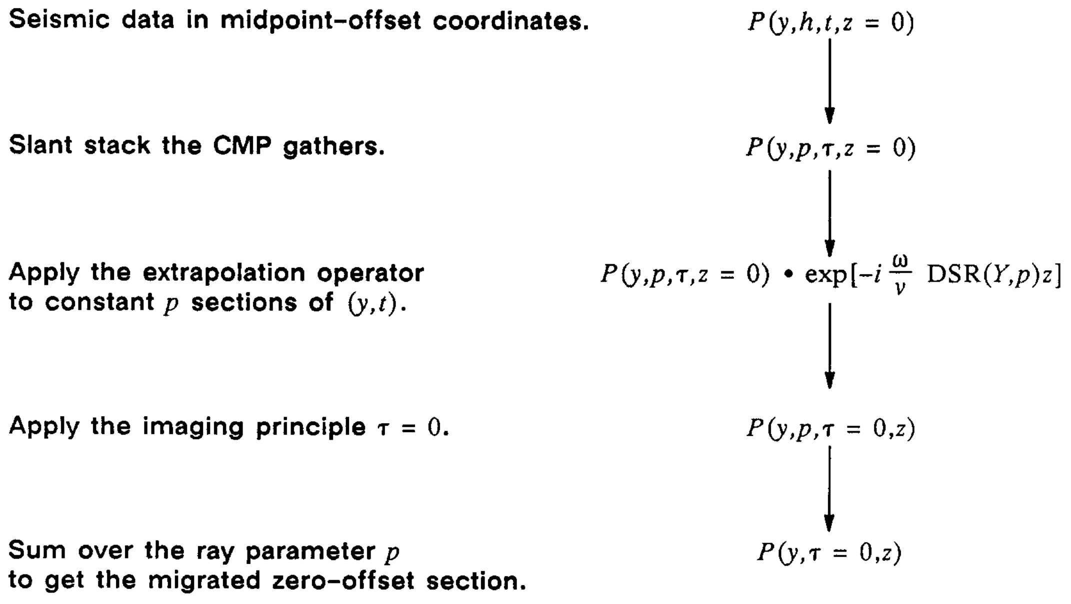


FIG. D-1. Flowchart for migration before stack in midpoint-ray-parameter coordinates.

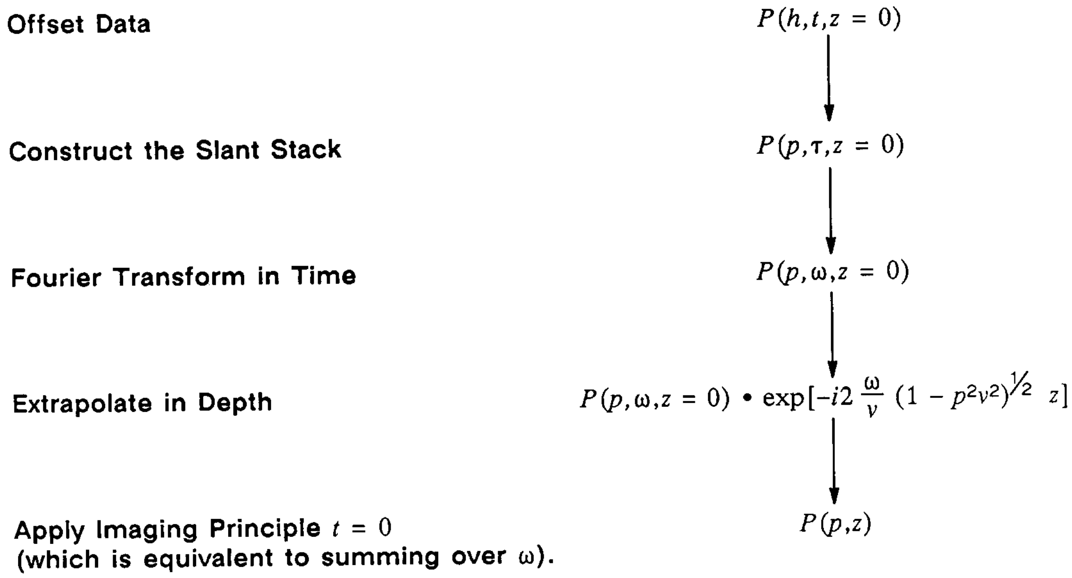


FIG. D-2. Processing flow for refraction inversion.

Appendix E

INSTANTANEOUS ATTRIBUTES

Section 8.6 gives data examples for instantaneous attributes. Here, the mathematical details of computing instantaneous parameters are provided. An analytic signal is expressed by a time-dependent complex variable *u(t)*:

$$u(t) = x(t) + iy(t), \tag{E.1}$$

where *x(t)* is the signal itself and *y(t)* is its quadrature. The quadrature is a 90-degree phase-shifted version of the recorded signal. It is obtained by taking the Hilbert transform of *x(t)* (Bracewell, 1965):

$$y(t) = \frac{1}{\pi t} * x(t). \tag{E.2}$$

When substituted into equation (E.1), we have

$$u(t) = x(t) + i\,\frac{1}{\pi t} * x(t) \tag{E.3}$$

or

$$u(t) = \left[\delta(t) + \frac{i}{\pi t}\right] * x(t). \tag{E.4}$$

Thus, to get the analytic signal *u(t)* for a seismic trace *x(t)*, the following operator is applied to the trace:

$$\delta(t) + \frac{i}{\pi t}. \tag{E.5}$$

When analyzed in the Fourier transform domain, this operator is zero for negative frequencies. Therefore, the complex trace *u(t)* does not contain negative frequency components.

Once *u(t)* is computed, it can be expressed in polar form as:

$$u(t) = R(t) \exp\,[i\phi(t)], \tag{E.6}$$

where

$$R(t) = [x^2(t) + y^2(t)]^{1/2} \tag{E.7}$$

and

$$\phi(t) = \arctan\,[y(t)/x(t)]. \tag{E.8}$$

Here, *R(t)* represents instantaneous amplitude and $\phi(t)$ represents instantaneous phase. Instantaneous phase can be computed with the following alternative approach. Taking the logarithm of both sides of equation (E.6), we have

$$\ln u(t) = \ln R(t) + i\phi(t) \tag{E.9}$$

so that

$$\phi(t) = \operatorname{Im}\,[\ln u(t)], \tag{E.10}$$

where Im is *imaginary*.

The instantaneous frequency *ω(t)* is the temporal rate of change of the instantaneous phase function:

$$\omega(t) = \frac{d\phi(t)}{dt}. \tag{E.11}$$

When the derivative of equation (E.10) is taken with respect to time,

$$\frac{d\phi(t)}{dt} = \operatorname{Im}\left[\frac{1}{u(t)}\,\frac{du(t)}{dt}\right]. \tag{E.12}$$

For practical implementation, equation (E.12) is written as the difference equation:

$$\omega_t = \operatorname{Im}\left[\frac{1}{(u_t + u_{t-\Delta t})/2}\;\frac{u_t - u_{t-\Delta t}}{\Delta t}\right]. \tag{E.13}$$

Finally, by simplifying equation (E.13), we have:

$$\omega_t = \frac{2}{\Delta t}\operatorname{Im}\left[\frac{u_t - u_{t-\Delta t}}{u_t + u_{t-\Delta t}}\right]. \tag{E.14}$$

REFERENCE

Bracewell, R., 1965, The Fourier transform and its applications: McGraw-Hill Book Co.

Appendix F

PLANE SURFACE FITTING

Consider a plane-surface fit in the least-squares sense:

$$\tilde{g}(x, y) = a_0 + a_1 x + a_2 y. \tag{F.1}$$

The least-squares error is

$$L = \sum_{i=1}^{M} (g_i - \tilde{g}_i)^2, \tag{F.2}$$

where M is the number of observations and g is the observed value at the grid point (x,y). We want to find a set of (a_0, a_1, a_2) for which L is minimum; i.e.,

$$\frac{\partial L}{\partial a_0} = \frac{\partial L}{\partial a_1} = \frac{\partial L}{\partial a_2} = 0. \tag{F.3}$$

By substituting equation (F.1) into equation (F.2)

$$L = \sum_{i=1}^{M} (g_i - a_0 - a_1 x_i - a_2 y_i)^2, \tag{F.4}$$

then by doing the differentiation [equation (F.3)], we get the following set of simultaneous equations:

$$\begin{aligned} \sum a_0 \quad &+ \sum a_1 x \;\; + \sum a_2 y \;\; = \sum g \\ \sum a_0 x \;\; &+ \sum a_1 x^2 + \sum a_2 xy = \sum xg. \\ \sum a_0 y \;\; &+ \sum a_1 xy + \sum a_2 y^2 \;\, = \sum yg \end{aligned}$$

When put into matrix form, we obtain

$$\begin{bmatrix} M & \sum x & \sum y \\ \sum x & \sum x^2 & \sum xy \\ \sum y & \sum xy & \sum y^2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum g \\ \sum xg \\ \sum yg \end{bmatrix} . \tag{F.5}$$

Equation (F.5) then is solved for the set of coefficients (a_0, a_1, a_2). This algorithm is used to locally fit M (usually 8) observations around a grid point at which the map function is to be evaluated.

INDEX