APPLICATIONS OF THE KARHUNEN-LOÈVE TRANSFORM IN REFLECTION SEISMOLOGY

by

IAN FREDERICK JONES

M.Sc. The University of Western Ontario, 1980
B.Sc.(Hon) The University of Manchester, 1976

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Department Of Geophysics And Astronomy)

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

April 1985

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of GEOPHYSICS & ASTRONOMY

The University of British Columbia
1956 Main Mall
Vancouver, Canada
V6T 1Y3

Date 19/4/85

ABSTRACT

The Karhunen-Loève transform, which optimally extracts coherent information from multichannel data, has been applied to several problems in reflection seismic data processing. The transform is derived by a least-squares construction of an orthogonal set of principal components and eigenvectors, with corresponding eigenvalues. Data are reconstructed as a linear combination of the principal components.

The mathematical properties of the Karhunen-Loève transform which render it applicable to problems in seismic data processing are reviewed, and a number of new algorithms developed. Most algorithms are tested on synthetic and real data examples, and 'production-line' industrially viable versions of some of the programs have been developed.

A new signal-to-noise ratio enhancement technique, based on reconstruction of stacked seismic sections, has proved to be successful on real data. Reconstruction of less coherent information to emphasize anomalous features in stacked seismic data ("misfit" reconstruction) shows some promise. Diffraction hyperbolae isolated by misfit reconstruction are used to estimate residual migration velocities with some success. And, the ability of the transform to segregate coherent information is used successfully as the basis of a new multiple suppression technique. An anomaly identification scheme, based on cluster

analysis of the eigenvectors of the transform, works well on the synthetic data used, and gives promising results when applied to real data. A new velocity analysis method, utilizing a ratio of the eigenvalues, works well for good data at early travel times, and offers a potential for high resolution velocity inversion studies. Use of the eigenvalues in evaluation of a constant phase approximation to dispersion for synthetic data provides promising results, leading to quantification of dispersion in terms of relative phase shifts. As part of this development, an analysis of the effect of dispersion on Vibroseis© data acquisition, which represents an original investigation, is presented.

Table of Contents

List of Figures

ACKNOWLEDGEMENTS.

I express my deepest thanks to my supervisors Drs. R.M. Clowes and R.M. Ellis, for their continued support during my time at U.B.C, and to Drs. T.J. Ulrych, D.W. Oldenburg, and M.J. Yedlin, for their considerable input and advice during this project. My thanks also to Drs. P. H. LeBlond and K. Larner for their participation as committee member and external examiner, respectively.

My special thanks goes to Shlomo Levy, who initially suggested this topic, and who's continued input guided me past many hurdles during its execution, leading to its present form.

The friendships of Ron Clowes, Mat Yedlin, Don Plenderleith, Brad Prager, Julian Cabrera, and David Waldron, have all been instrumental in helping me through periods of turmoil whilst striving to keep my head above the waves during the past four years.

This work was supported in part by the Natural Sciences and Engineering Research Council (grants A7707, A2617, A4270, and A1804), Petro-Canada Ltd., and Mobil Oil Canada Ltd. Inverse Theory and Applications (IT&A) Inc., generously made available the computing facilities, data, and programming expertise necessary for the completion of much of the work presented here. My thanks also to Dr. J. den Boer, of Mobil Oil Canada, and to Mr. J. Muir, of Chevron Canada Resources, for kindly supplying data for this project.

DEDICATION

GEMMA:

A causa de este trabajo,
Y por tiempo mal pasado,
con ensueños no realizados,
Yo digo ahora:

Lo demas es para ti.

CHAPTER 1.

INTRODUCTION.

SECTION 1.I: GENERAL INTRODUCTION.

The Karhunen-Loève transform has long been known and well utilized in the image processing field (Ahmed and Rao, 1975); however, in geophysical data processing, applications to date have been scant. In this thesis, my objective is to introduce and exemplify the Karhunen-Loève transformation to the seismic data processing community as a viable method for addressing several problems in multichannel seismic data processing. The basic tenet behind this work is that using the Karhunen-Loève transform, seismic data possessing linear trace-to-trace coherency may be projected into a space where the coherent information is compressed into the smallest possible number of 'alternative data' traces. This compression separates the correlated from uncorrelated parts of the input data. I exploit this separability feature to improve existing seismic processing techniques, and to form the basis of new multichannel seismic data processing techniques. Several of the topics in this thesis are presented here for the first time, while others have been mentioned in existing literature, or in recent articles on which I am co-author.

My attention was originally directed to this topic by S. Levy and T.J. Ulrych, who with D.W. Oldenburg and myself have been investigating the usefulness of the Karhunen-Loève transform in the framework of a more general inversion approach to seismic data processing. Initially, my interest in the Karhunen-Loève transform (referred to in the remainder of this work as the KL transform, or the KLT) focussed on its ability to quantify similarity between waveforms. This interest arose as I was interested in the problem of the quantification of dispersion in seismic body-wave data, and the subsequent correction of dispersive effects. Subsequently, my interest shifted more towards the application of KLT to various other problems. However, evidence of the initial thrust of my thesis is to be found in the coverage given to that problem in the Q-inversion section of Chapter 4, and in the developments concerning the effects of dispersion on the choice of bandwidth for Vibroseis® data (Appendix 2).

As the majority of interest in this work is within the industrial seismic data processing community, one of my goals has been to produce a viable 'production-line' processing package for each of the applications investigated. However, some of the applications, although they proved to be promising on synthetic data, have not proved as powerful for real data examples.

Some applications of the Karhunen-Loève transform were presented by myself on behalf of my co-authors at the 1983 Annual

International Meeting of the Society of Exploration Geophysicists (SEG) in Las Vegas, Nevada, USA. A manuscript in preparation on multichannel applications (Jones and Levy, 1985) was presented at the 1984 Annual International Meeting of the European Association of Exploration Geophysicists in London, England, by my co-author, S. Levy. A brief review of some of the work considered here and elsewhere by co-workers and myself was given by D.W. Oldenburg, at the 1984 joint meeting of the Canadian Society of Exploration Geophysicists (CSEG) and Canadian Society of Petroleum Geologists (CSPG), in Calgary, Alberta, Canada, and also at the SEG Seismic Deconvolution Workshop in July 1984, in Vail, Colorado, USA.

To introduce the reader to the background of the KLT, the problems I consider, and the application of the KLT to those problems, the body of this thesis is divided into four major sections:

1. An introduction to previous work using the KLT in other fields, and to the subsequent sections considered here (Chapter 1);

2. A theoretical introduction to the various aspects of the KLT outlining the well established theory from the literature, and introducing the physical insights behind each of the applications considered here (Chapter 2);

3. Applications of the method to synthetic multichannel seismic data sets, and to multichannel field data, with reference to the relevant theoretical development (Chapters 3 and 4); and

4. An assesment of the usefulness and limitations of the method in each of the applications discussed (Chapter 5).

SECTION 1.II: REVIEW OF PREVIOUS WORK.

In this section, I review the previous work in the literature on orthogonal transforms, and more specifically, the expansion of multichannels of data in terms of orthogonal transforms.

a. Orthogonal Expansion of a Function.

The best known, and earliest treatise on orthogonal transforms was that of Fourier, who asserted that a function which met certain broad conditions (specified by Dirichlet; see Kanasewich, 1982) could be expressed as a linear combination of sinusoids. This approach was later generalized to the representation of an arbitary function by linear combination of orthogonal non-sinusoidal functions. For example, Walsh (1923) introduced a set of square-wave-like orthogonal functions, and several other functions were also devised (e.g. the Rademacher functions, Beauchamp, 1975).

Hotelling (1933) was one of the earliest workers to begin to utilize the more general approach in the guise of factor, or principal component analysis. His interest was in classifying psychological test scores; hence his data consisted of

multichannels of discrete values.

These methods first appeared in the signal processing literature when Karhunen (1947) expanded replications of single-channel stellar line-spectral data into a suite of orthogonal functions. His approach was formalized by Loève (1948, 1955), who treated the derivation of the transform in a rigorous statistical sense, and the term Karhunen-Loève transform then appeared in the literature to complement the terms Hotelling and principal component transforms.

b. Orthogonal Expansion of Multichannel Data.

The Karhunen-Loève transform was defined to statistically describe the expansion of a suite of realizations of a single function as an infinite sum of orthogonal functions. As a corollary to this, several workers explored the possibility of simultaneously expanding multichannels of data in terms of the same set of orthogonal functions. In the multichannel case, the covariance matrix of the data is decomposed, and for n channels of data (continuous or discrete) we need (at most) n orthogonal functions for an exact representation. If the original functions comprising the multichannel input are not linearly independent, then they could presumably be represented by a subset of the orthogonal functions.

The first workers to demonstrate the development of the transform for multichannel data using a matrix algebra approach were Kramer and Mathews (1956) who applied their procedures to speech analysis and data compression prior to telegraphic transmission. In fact, the speech processing and telegraphic communications communities were the first groups to fully explore and exploit the usefulness of the KL transform.

Young and Huggins (1962, 1963) and Christensen and Hirschman (1979) subsequently utilized the multichannel data approach to extract the underlying signal forms from a suite of electrocardiograph signals. Young and Huggins (1962) and Young and Calvert (1974) also refer to the method in its multichannel form as Intrinsic Component Analysis. Also, Gubbins et al. (1971) investigated the use of Haar and Walsh functions in filtering techniques.

Watanabe (1965) gave a comprehensive overview of the transform, showing how it could be derived not only in a least-squares sense, as was done independently by Kramer and Mathews (1956), but also in a way which minimized the entropy of an objective function. He also elucidated the salient differences between factor, or principal component analysis (which was developed with discrete data in mind) and the KL transform, which was derived for a continuous function.

A succinct definition of the transform coined by Watanabe reads as follows:

> "The K-L-expansion is usually known as one which minimizes the average error committed by taking only a finite number of terms in the infinite series of an expansion when a given collection of functions is to be expressed as a series in terms of some complete set of orthogonal functions."

In the same year, Papoulis (1965) presented the method as that which, in a least-squares sense, 'optimally' transformed a vector or suite of vectors into a complete orthogonal set.

After further work on pattern recognition and feature selection (Chien and Fu, 1967) and a proposed modification to the method (Fukunaga and Koontz, 1969), use of the method in the image processing field became widespread (Pratt, 1970; Ready and Wintz, 1973; Pelat, 1974; Ahmed and Rao, 1975; Andrews and Patterson, 1976 a, b, and c; Jain, 1976, 1977; Lowitz, 1978; Mallick and Murthy, 1984). According to Young and Calvert (1974), the application of the KL transform to feature extraction was due initially to Chien and Fu (1967) and Watanabe (1965).

Ahmed and Rao (1975), in their book on the applications of orthogonal transforms to digital image processing, review most of the above mentioned techniques and applications. They use the definition of the Karhunen-Loève transform as applied to the

decomposition of the covariance matrix of an image when described as a series of row vectors.

Throughout this work I will be dealing with multichannels of data, and will use the term 'Karhunen-Loève' (KL) transform to mean the expansion of such data via diagonalization of its covariance matrix.

c. Geophysical Applications.

The first geophysical application of the method was the investigation by Hemon and Mace (1978) of the ability of the transform to extract a common signal from a set of move-out corrected CDP traces as an alternative to conventional stacking. They demonstrated, with synthetic examples, how the effects of Gaussian noise and of trace-to-trace static time shifts were less troublesome to the KL approach than to the conventional stacking method. However, their paper was published in French which unfortunately led to their work being overlooked by a large segment of the geophysical community.

Tjostheim and Sandvin (1979) employed the KL transform, in conjunction with auto-regressive methods, to distinguish between underground nuclear explosions and large earthquakes. They claimed success on the basis of the difference in number of orthogonal functions needed to represent the records of an earthquake as opposed to a nuclear explosion.

Milligan et al. (1978) analysed 140,000 acoustic 'pinger' records in an attempt to discriminate water-bottom sediment type on the basis of seismic pulse character. Using 1.8 ms of data, the aligned bottom 'ping' records were decomposed into their representative orthogonal vectors. Depending upon which subset of orthogonal vectors could be used to represent the ping from a given locale, they were able to group locales into regions which proved to be representative of bottom-sediment type. In a later paper, the same group (LeBlanc and Middleton, 1980) used the method to investigate the similarities between underwater sound velocity profiles on an ocean-wide scale.

Under the guise of principal components analysis, Hagen (1982), working with the instantaneous phase (Taner et al., 1979) of his data, used the technique of cluster analysis (on the basis of the transformation's amplitude coefficients) to group seismic traces into 'natural' clusters. His hope was that in conjunction with well-log control, these clusters could be related to porosity. Hence, a glance at a trace clustering 'map' for adjacent undrilled areas would enable an interpreter to discern the likely locations for oil and gas accumulations.

d. The Work of the U.B.C. Group.

More in line with the signal extraction and coherency aspects of the method, Ulrych et al. (1983) presented a series of applications in the geophysical signals processing field. They

concerned themselves with wavelet extraction from both move-out hyperbolae and single channels of data. They also carried out stacking and velocity analysis using the eigenvalues of the transform to construct a similarity measure to replace the conventional semblance criterion (Neidell and Taner, 1971).

Levy et al. (1983) in a companion paper extended the KL method to deal with complex signals. This allowed the authors to address the problem of phase changes in the seismic wavelet. They produced enhanced stacks of move-out corrected CDP gathers for synthetic data, and also parallelled the new velocity analysis approach introduced by Ulrych et al. (1983). Utilizing the diagnostic features of the transformation's amplitude coefficients (as in cluster analysis), they developed a 'dead-trace' detection routine to discriminate against unusually different seismic traces to avoid degradation of the final results when, for example, stacking or performing velocity analysis. Extending the versatility of the complex KL method even further, they extracted phase information from sets of common signals for synthetic data, and in the case of super-critical reflections from a fluid-fluid interface, were able to invert phase and velocity information to recover density contrasts across an interface.

In a more recent work, Chapman et al. (in press) utilized these procedures to extract density contrast information from the relative phase changes observed in super-critical reflections

within Arctic abyssal plains sediments.

On the papers by Chapman et al. (in press), Ulrych et al. (1983), and Levy et al. (1983), I am a co-author. The velocity analysis and constant phase approximation for dispersion components from the latter two papers comprise part of this thesis. It is from my work on these two papers and further discussions with S. Levy, T.J. Ulrych and D.W. Oldenburg, that this thesis evolved.

As mentioned previously, the original thrust of this thesis was an investigation of dispersion and its quantification using sensitive similarity measures. From this, the KL method grew to be the central theme of this work. A significant amount of time was devoted to the possibility of estimating the seismic quality factor Q, assuming the constant-Q model (Futtermann, 1962) and utilizing the eigenvalues of the KL method to define a similarity measure. An overview of the quantification of Q values, the constant-Q model, and the effects of dispersion on seismic source signals is given in Section 4.III and Appendices 1 and 2, respectively.

SECTION 1.III: SEISMIC DATA RECONSTRUCTION.

a. Reconstruction of Coherent Information.

Given a stacked seismic section from a region of relatively flat-lying or parallel dipping layers, the changes seen from trace to trace across the section are in part structurally related and in part random noise. I will assume that the contributions from random events are not correlated from trace to trace, whereas the data representing the geological horizons will be. In view of these assumptions, I enquire whether we can separate those parts of the data which display trace-to-trace coherency from those parts which do not.

As will be shown in the next chapters, the KLT enables us to separate the most coherent parts of the data from the less coherent parts. The method replaces a set of seismic traces with an equal number of alternative (and orthogonal) data traces (known as the principal components), which are arranged in order of decreasing variance (or energy content). The first principal component contains that signal which is most common across the stacked seismic section. The subsequent principal components can either be viewed as the next most common signals, or as the 'correction terms' to be added to the common signal to reproduce the input signals. Reconstructing the original stacked seismic section from those principal components which account for the

greatest variance will produce a seismic section from which the least coherent information has been omitted.

This truncated seismic data reconstruction method has not been presented previously in the seismic literature, although it forms the basis of many image enhancement techniques.

b. Misfit Reconstruction.

As a natural adjunct to the reconstruction of coherent information, we may ask the following question: What do we have if we perform a reconstruction using the least coherent information present? The answer is simply that we then have those parts of the data which are less common from trace to trace.

For example, in the case where we have a background of similar flat-lying (or parallel dipping) events upon which is superimposed some small-scale structure such as a sand lens, or pinch-out, we may try to remove the common features in order to highlight the anomalies. A technique such as this would draw attention to regions which may warrant further investigation. For display purposes, the anomalous parts of the data could be plotted in grey-shades on top of a variable area plot of a reconstruction of the original data in order to produce an 'anomaly-highlighted' section.

This technique, which I call anomaly or misfit reconstruction, has not been presented previously in the

seismological or image processing literature.

c. Diffractions.

Whilst performing reconstructions of stacked seismic data, I noticed a most interesting feature. In areas with much faulting, the misfit reconstruction highlighted the diffraction hyperbolae associated with fault edges. The fact that a migrated stacked section still incorporates some diffraction energy indicates that residual migration may be in order. Further, migrating the diffraction section (a misfit reconstruction showing diffractions) itself may produce a section with events concentrated near the diffraction sources: this could serve for example as a fault-edge indicator.

This serendipitous application is new, although the separation of diffraction events from stacked seismic sections by other means is not (Levin et al., 1983, Harlan et al., 1983).

d. Multiple Suppression.

In shallow marine seismic data, multiples often pose a problem. The multiple events all have approximately the same move-out velocity, and often distort the information of interest. Performing a constant velocity move-out correction on a multiple infested common depth point (CDP) or common shot point (CSP) gather using the velocity of the multiple, we produce a gather in

which the 'multiple' events are essentially flattened. In other words they will then appear to be the most coherent information in the gather.

A KL reconstruction omitting the most common information will produce a gather free of multiple arrivals. The gather would then be un-move-out corrected with the same constant velocity. Naturally, the processing would commence just after the primary event so as not to suppress it.

This novel approach to multiple suppression is presented here for the first time. It constitutes an alternative approach to that of Ryu (1982) who proposed a procedure using frequency-wavenumber filtering of move-out altered gathers.

SECTION 1.IV: SIMILARITY MEASURES.

In the previous section it was mentioned that the KLT produces a set of orthogonal principal components from the input data traces. As shown in the theory sections (Chapter 2), these principal components are constructed as a linear combination of the input data traces (and vice versa for data reconstruction).

In this section, I will introduce the physical significance of the weighting coefficients used in these linear combinations. As we will see, these weights are the elements of the eigenvectors of the covariance matrix of the input data traces.

a. Cluster Analysis.

When constructing, for example, the first principal component, we will have a set of weights which will tell us how much of each input trace goes into making the first principal component.

If the first dozen weights had the same value, then we would conclude that the first dozen input traces contributed equally to the first principal component. If the second dozen weights were similar to each other but differed radically from the first dozen, we could infer that we had two 'natural' similarity groups of input data traces, characterized by the first and second dozen input data traces, respectively.

By studying the groupings of the weighting elements associated with the KLT, we can form a basis for identifying groups, or clusters, of input signals which were similar. One of the major aims of such an analysis would be to link the resulting groups of input traces to a physical parameter, eg. porosity, noted from attendant well log information.

I review the approach taken by Hagen (1982) and critically assess the applicability of cluster analysis in the context of seismic exploration.

b. Velocity Analysis.

The energy of the j.th principal component is given by the j.th eigenvalue of the input data covariance matrix. Further, if that signal most common to the input data has a comparatively large energy, then we can say that the input data traces are all very similar.

Consequently, a ratio of the first eigenvalue to the sum of the remaining eigenvalues gives a measure of the distribution of the total variance between common and uncommon parts of the input data. This eigenvalue ratio, which I later modify to be the ratio of a sum of the first m eigenvalues to the sum of the remaining eigenvalues, can be used as a similarity measure.

Under certain circumstances, this similarity measure proves to be more reliable than the conventional semblance criterion (Neidell and Taner, 1973), especially in the presence of small time shifts between the members of the suite of input signals under consideration. I present the modified eigenvalue ratio as an alternative to the semblance criterion used in velocity analysis. The initial results of this investigation were presented in Ulrych et al. (1983) and Levy et al. (1983).

c. Q-Inversion.

In a procedure similar to the application to velocity analysis, I utilize the eigenvalue ratio to quantify the similarity between a dispersed signal and a reference signal. Utilizing the Futtermann (1962) relationship (which relates the change in phase velocity to the Q-value for a medium; Knopoff (1964)), the dispersed signal is iteratively undispersed. That is, at each iteration, the signal is first subjected to a frequency dependent stretching designed to negate the effects of dispersion. This undispersed signal is used, in conjunction with a reference signal, to compute the value of the eigenvalue ratio for the pair of signals. The change in magnitude of the eigenvalue ratio will give an indication of how similar to the reference signal the processed signal has become. By identifying a minimum in the function of eigenvalue ratio versus undispersing Q-value, I try to identify the Q-value of the medium in which the seismic wavelet travelled.

In addition to the application of the KL eigenvalue ratio to similarity quantification, an overview of the constant Q model, and of the effects of dispersion on Vibroseis® data is presented (Appendices 1 and 2). In Section 4.III, I present an iterative dispersion removal technique for synthetic data from an impulsive source function. This extends the work of Robinson (1979, 1982)

who attempted to remove the effects of dispersion from seismic data in an ad hoc fashion.

Within the body of this thesis, all figures are collated at the end of the pertinent section, so as not to interrupt the text.

CHAPTER 2.

THEORY AND BACKGROUND.

SECTION 2.I: THE KARHUNEN-LOÈVE TRANSFORMATION.

a. Introduction.

In this chapter, I present the theory from the literature, and outline the physical interpretation of this theory as applied to problems in reflection seismology. I will be dealing with the KL transformation as applied to multichannels of data (Ahmed and Rao (1975), Ulrych et al. (1983), Levy et al. (1983), Jones and Levy (1984)) and begin by defining the terminology to be used in this work.

b. Theory.

Given a set of n real signals $x_i(t)$, $0 \leq t \leq T$, and an $[n \times n]$ transformation matrix A with elements a_{ij}, we may construct a set of alternative data as a linear combination of the input data:

$$\psi_j(t) = \sum_{i=1}^{n} a_{ij}\, x_i(t) \qquad j=1,\ldots,n;\ 0 \leq t \leq T \qquad (1)$$

Conversely, we may reconstruct the input signals as a linear combination of the vectors $\psi_j(t)$, with weights b_{ij} of a matrix B:

$$x_i(t) = \sum_{j=1}^{n} b_{ij}\, \psi_j(t) \qquad i=1,\ldots,n \qquad (2)$$

We may derive the coefficients a_{ij} and b_{ij} in a number of ways (Ulrych et al., 1985, in prep.). However, use of the least-squares method gives rise to a set of orthogonal vectors $\psi_j(t)$, arranged in order of decreasing energy content, i.e. the signal with the largest variance will appear on $\psi_1(t)$, and so on. This approach was taken by Kramer and Mathews (1956) who showed that the coefficient matrices A and B for that transform which packs the greatest possible amount of coherent information into the smallest possible number of vectors $\psi_j(t)$, are simply $A = R$ and $B = R$, where R is the eigenvector matrix derived from the covariance matrix of the input data (in the notation of Kramer and Mathews, my A would be transposed).

Their result can be shown as follows. Defining a truncated data reconstruction as:

$$\tilde{x}_i(t) = \sum_{j=1}^{m} b_{ij}\, \psi_j(t) \qquad i = 1, \ldots, n;\ m \leq n \tag{3}$$

we have an associated misfit (or truncation error):

$$\phi(m) = \int_0^T \sum_{i=1}^{n} (x_i(t) - \tilde{x}_i(t))^2\, dt \tag{4}$$

$$= \int_0^T \sum_{i=1}^{n} \left(x_i(t) - \sum_{j=1}^{m} b_{ij}\, \psi_j(t)\right)^2 dt$$

$$= \int_0^T \sum_{i=1}^{n} \left(x_i(t) - \sum_{j=1}^{m} \sum_{k=1}^{n} b_{ij}\, a_{kj}\, x_k(t)\right)^2 dt$$

Demanding that $\phi(m)$ be a minimum for $1 \leq m \leq n$, we set the partial derivatives of $\phi(m)$ with respect to a_{ij} and b_{ij}, to zero. For

$\partial\phi(m)/\partial b_{pq}$, we have:

$$\partial\phi(m)/\partial b_{pq} = -2 \int_0^T \{x_p(t) - \sum_{j=1}^{m} \sum_{k=1}^{n} b_{pj}\, a_{kj}\, x_k(t)\} \cdot \{\sum_{l=1}^{n} a_{lq}\, x_l(t)\}\, dt$$

$$= -2 \int_0^T [\{\sum_{l=1}^{n} a_{lq}\, x_p(t) x_l(t)\} - \{\sum_{j=1}^{m} \sum_{k=1}^{n} \sum_{l=1}^{n} a_{lq}\, b_{pj}\, a_{kj}\, x_k(t)\, x_l(t)\}]\, dt$$

$$= 0$$

But the integral over pairs of input vectors gives the symmetric covariance matrix, i.e.:

$$\gamma_{ij} = \int_0^T x_i(t)\, x_j(t)\, dt$$

So, we have:

$$\partial\phi(m)/\partial b_{pq} = -2 [\{\sum_{l=1}^{n} \gamma_{pl}\, a_{lq}\} - \{\sum_{j=1}^{m} \sum_{k=1}^{n} \sum_{l=1}^{n} a_{kj}\, b_{pj}\, \gamma_{kl}\, a_{lq}\}]$$

Summing over l gives:

$$\partial\phi(m)/\partial b_{pq} = -2 [(\Gamma A)_{pq} - \{\sum_{j=1}^{m} \sum_{k=1}^{n} a_{kj}\, b_{pj}\, (\Gamma A)_{kq}\}]$$

And summing over k and j gives:

$$\partial\phi(m)/\partial b_{pq} = -2\ [\ (\Gamma\ A)_{pq} - (B\ A^{T}\ \Gamma\ A)_{pq}\ ]$$

$$= 0$$

Dropping the element subscripts and transposing gives:

$$A^{T}\ \Gamma = A^{T}\ \Gamma\ A\ B^{T} \qquad (5)$$

Likewise, from setting $\partial\phi(m)/\partial a_{pq} = 0$, we obtain:

$$\Gamma\ B = \Gamma\ A\ B^{T}\ B \qquad (6)$$

To investigate the properties of A and B, we perform the following steps:. Setting $C = B\ A^{T}$, *equations (5)* and *(6)* become:

$$C\ \Gamma = C\ \Gamma\ C^{T}$$

$$\Gamma\ C = \Gamma\ C^{T}\ C$$

From the latter equation, we see that:

$C = C^{T}C$, or transposing:

$C^{T} = C^{T}C$, thus:

$C = C^{T}$, and hence:

$C = C^{2}$

In other words, C is diagonal with either ones or zeroes arranged in some as yet unknown order along the diagonal.

Proceeding with a parallel argument for the spectral decomposition of Γ in terms of R and Λ, we have the following:

$$\Gamma = R \Lambda R^{T} \qquad (7)$$

where R is the matrix of orthogonal column eigenvectors $\bar{r}_j$, and the diagonal matrix Λ contains the corresponding eigenvalues λ_1, $\lambda_2, \ldots, \lambda_n$, arranged in decreasing size.

As C is also symmetric, we can write $C = R Q R^{T}$ where Q is a diagonal matrix with rank equal to the rank of C (i.e. rank $\leq m$).

Now:

$$Q^2 = R^{T} C R R^{T} C R$$

$$= R^{T} C^2 R$$

$$= R^{T} C R$$

$$= Q$$

Consequently, the elements of the diagonal matrix Q must either be unity or zero.

In addition, we can also write *equation (4)* in terms of the deviation of a forward and inverse transform from the original data, i.e.:

$$\phi(m) = Trace[(I - B A^{T}) (I - A B^{T}) \Gamma]$$

$$= Trace[(I - C) \Gamma]$$

but as this only looks at the diagonal elements, we can equally well rotate the argument into a diagonal form:

$$\phi(m) = Trace[R^T (I - C) R R^T \Gamma R]$$

$$= Trace[(I - R^T C R) \Lambda]$$

$$= Trace[(I - Q)\Lambda]$$

$$= \sum_{j=m+1}^{p} \lambda_j$$

where p is the rank of Γ. But this representation of $\phi(m)$ will give a minimum when the sum extends over the smallest $(p-m)$ eigenvalues, and as the ordered eigenvalues correspond to the ordered eigenvectors of R, we must have that:

$$Q = M$$

where M is an identity matrix of rank m, but:

$$C = R Q R^T$$

$$= R M R^T$$

$$= B A^T$$

In other words, the first m columns of A or B correspond to the first m columns of R. Hence the result of Kramer and Mathews (1956) that A = B = R for m=n, and A = B = the first m columns of R for m<n.

Consequently, a convenient way of solving this least-squares problem is to first compute the outer-product covariance matrix Γ, of the data $x_i(t)$:

$$\Gamma = X X^T \qquad (8)$$

where $X = \{x_i(t),\ i=1,\ldots,n\}$,

and proceed to diagonalize *equation (8)* via *equation (7)* to obtain the eigenvectors R. (Throughout this work it is understood that a function $x_i(t)$ may be discretized so as to form the elements of the $i.th$ row of a matrix X).

With A and B thus chosen, the transformation pair *equations (1)* and *(2)* define the multichannel Karhunen-Loève (KL) transform:

$$\psi_j(t) = \sum_{i=1}^{n} r_{ij} x_i(t), \qquad j=1,\ldots,n, \tag{9}$$

or $\Psi = R^{T} X$

and

$$x_i(t) = \sum_{j=1}^{n} r_{ij} \psi_j(t), \qquad i=1,\ldots,n, \tag{10}$$

or $X = R\,\Psi$

where r_{i1} is the $i.th$ element of the first column vector of R, $\Psi=\{\psi_j(t),\ j=1,\ldots,n\}$, and the vectors $\psi_j(t)$ are known as the principal components of the transformation.

Further, the $j.th$ eigenvalue λ_j gives the energy content of the $j.th$ principal component, i.e.:

$$\lambda_j = \int_0^T \psi_j(t)^2 \, dt \qquad (11)$$

In other words, the eigenvalue matrix Λ is simply the covariance matrix of the orthogonal alternative data $\psi_j(t)$:

$$\Lambda = \Psi \, \Psi^T \qquad (12)$$

Also, as the transformation is merely a spatial rotation, the total input energy is conserved and we have:

$$Trace[\Gamma] = \int_0^T \sum_{i=1}^{n} x_i(t)^2 \, dt = Trace[\Lambda] \qquad (13)$$

Due to the way in which the covariance matrix is calculated, the principal component with the largest variance contains that signal which is most highly correlated from trace-to-trace. In other words, the first principal component represents the 'common signal'.

The underlying assumption in this work is that there is a signal which is common to each input trace: our problem is to extract that signal. In matrix terms, we assert that X has a predominant signal component with a rank much smaller than n, the number of input vectors.

If there were no trace-to-trace similarity (i.e. if the $x_i(t)$ were orthogonal), Γ would be diagonal, and we would have:

$$\psi_j(t) = x_i(t), \qquad \text{and} \qquad \lambda_j = \int_0^T x_i(t)^2 \, dt.$$

where the

$\underset{i}{x}(t)$ are re-arranged in order of decreasing energy content (as $\psi_j(t)$) to meet the requirement that $\lambda_1 > \lambda_2 > \ldots > \lambda_n$.

c. Singular Value Decomposition.

For most seismological applications we have many more points per trace than we have traces. In this case it is simple and convenient to proceed as outlined above. However, in situations where the number of traces far exceeds the number of points per trace (Milligan et al., 1978) we utilize the relationship between the KL transform and singular-value decomposition (SVD) to reduce the computational expense.

Given the data matrix X, we may write:

$$X = R\ \Omega\ V^{\mathsf{T}} \qquad (14)$$

where Ω is in general a rectangular diagonal matrix containing the singular values of this decomposition, and V is the post matrix, which is later seen to contain the normalized principal components.

Now, we note from *equation (8)* that the covariance matrix Γ for our data X is given by the outer-product:

$$\Gamma = X\ X^{\mathsf{T}} \qquad (15)$$

$$= R\ \Omega\ V^{\mathsf{T}}\ V\ \Omega\ R^{\mathsf{T}},$$

and as both R and V are orthogonal, this becomes:

$$\Gamma = R\ \Omega^2\ R^T,$$

so we see that from *equation (7)* $\Omega^2 = \Lambda$.

The matrices R and Λ are both $[nxn]$. However, for our n traces each of N points, we have that $X = [nxN]$, $\Omega = [nxN]$, and $V = [NxN]$, but for $n < N$, the system is overdetermined and we only have at most n non-zero eigenvalues in the matrix Ω. Hence $\Lambda^{1/2}$ which is $[nxn]$ fully represents Ω which is $[nxN]$.

Now, if we had a case where the number of traces exceeded the number of points in a trace, i.e. $n > N$, we may wish to form the following inner-product covariance matrix:

$$\underline{\Gamma} = X^T\ X \qquad (16)$$

$$= V\ \Omega^2\ V^T$$

Here V is the eigenvector matrix of $\underline{\Gamma}$ (both V and $\underline{\Gamma}$ are $[NxN]$); recall that R contained the eigenvectors derived from Γ. We constructed the principal components Ψ (*equation (9)*) as:

$$\Psi = R^T\ X \qquad (17)$$

where Ψ is $[nxN]$. Substituting *equation (14)* into *equation (17)* yields:

$$\Psi = R^T\ R\ \Omega\ V^T$$

$$= \Omega\ V^T \qquad (18)$$

So we see that the principal component matrix Ψ, of Γ is simply a transposed scaled version of the eigenvector matrix V, of $\underline{\Gamma}$. Given V and Ω we may derive the eigenvector matrix R, of Γ: equating *equations (17)* and *(18)* gives:

$$R^T X = \Omega V^T,$$

which when rearranged becomes:

$$R = X V \Omega^{-1} \qquad (19)$$

So, if we simply want the principal components, or if $n >> N$, we decompose $\underline{\Gamma}$, whereas if we are interested in the eigenvectors R, or if $N >> n$, we decompose Γ.

d. Outer-Product Image Summation.

Here I present a reconstruction as a linear combination of principal components with appropriate weights. However, following the SVD approach, we may use the terminology of Andrews (1970) and Andrews and Patterson (1976 a, b, c) and think of the data as constituting an image matrix X (Davenport and Root, 1958; Huang and Narendra, 1975; Hunt and Kubler, 1984; or Fukunaga and Koontz (1969)). This image can be decomposed into a sum of rank-one images, each of which is formed by an outer-product expansion. That is, from *equation (14)*:

$$X = \sum_{j=1}^{n} \lambda_j^{1/2} \bar{r}_j \bar{v}_j^T$$

$$= \lambda_1^{1/2}\ \bar{\mathbf{r}}_1\bar{\mathbf{v}}_1^T + \ldots + \lambda_n^{1/2}\ \bar{\mathbf{r}}_n\bar{\mathbf{v}}_n^T$$

where each outer-product $\bar{\mathbf{r}}_j\bar{\mathbf{v}}_j^T$ forms a rank-one image of the same size as X. Photographs of several such rank-one decomposition images can be found in Andrews (1970).

SECTION 2.II: RECONSTRUCTION AND ASSOCIATED ERRORS.

a. Identical Signals.

When we have a suite of signals which are identical to within a constant scale factor, the first principal component has an important characteristic. If $x_i(t)=c_i s(t)$ where the c_i are real constants, and $s(t)$ is a given signal, then the first principal component will be:

$$\psi_1(t) = \left(\sum_{i=1}^{n} c_i^2 \right)^{1/2} s(t) \qquad (20)$$

That is, the first principal component will be a scaled version of the signal $s(t)$, and the complete set of input signal vectors $\{x_i(t),\ i=1,\ldots,n\}$, can be reconstructed exactly from this one principal component and the appropriate weights (the elements of the first row of R). The remaining principal components $\{\psi_j(t),\ j=2,\ldots,n\}$, will have zero eigenvalues and are not needed in the reconstruction.

b. Similar Signals.

If the suite of signals had random noise or other contaminants added to them, we no longer have the ideal case where the suite can be exactly represented by a single principal component with appropriate scale factors. Rather, we may still represent the suite with a single principal component and accept the introduction of some error.

In *equation (11)* we noted that the $j.th$ eigenvalue gives the energy associated with the $j.th$ principal component. With this in mind, I define a similarity measure:

$$\chi(m) = \sum_{j=1}^{m} \lambda_j \; / \; \sum_{j=m+1}^{n} \lambda_j \qquad (21)$$

For the 'identical' case with $\lambda_2, \lambda_3, \ldots, \lambda_n = 0$ we will have:

$$\chi(m) = \infty, \qquad m = 1, 2, \ldots, n-1.$$

Of more physical significance is the case when the waveforms are similar, and contaminated with noise. In this case, $\chi(m)$ will be large. Conversely, for inherently dissimilar signals (or white noise) with the same mean energy, we would have:

$$\chi(m) \simeq O\{m/(n-m)\}$$

In other words, if all the signals were dissimilar, then the covariance matrix would be mostly diagonal, and all the

eigenvalues would be of the same magnitude, hence *equation (21)* would yield a result of order $\{m/(n-m)\}$.

This modified eigenvalue ratio (EVR) $\chi(m)$, is analagous to the condition number (Strang, 1980), and is used throughout this work as a measure of signal coherency.

c. Winnowing and Truncation Error.

For the truncated inverse transformation (*equation (3)*) we must decide upon a criterion to determine how many principal components to use in order to satisfactorily reconstruct (or represent) the data. We may discard principal components $m+1$ to n given that the bulk of the energy is packed into the first m principal components. However, this approximation gives rise to an associated misfit (or truncation) error, $\phi(m)$ (*equation (4)*), which tells us how much of the input energy we have omitted from the reconstruction.

Given that the $j.th$ eigenvalue λ_j simply gives a measure of the energy content of the $j.th$ principal component $\psi_j(t)$ (*equation (11)*), and a sum of the eigenvalues yields the total input energy (*equation (13)*), we may rewrite this as:

$$\phi(m) = \sum_{j=m+1}^{n} \lambda_j \qquad (22)$$

So, to decide upon the number of principal components to use in the reconstruction, I simply request that a specified amount of

the input energy be present in the output. I express this requirement in terms of the reconstruction energy:

$$\eta(m) = 100 \sum_{j=1}^{m} \lambda_j \; / \sum_{j=1}^{n} \lambda_j \qquad (23)$$

Conversely, we may speak of the percentage misfit:

$$\mu(m) = 100 \sum_{j=m+1}^{n} \lambda_j \; / \sum_{j=1}^{n} \lambda_j = 100 \; \phi(m) \; / \; \phi(0) \qquad (24)$$

In Figure 2.1, we see a plot of the cumulative percentage energy versus number of eigenvalues, for a stacked seismic section. We see manifest here the energy packing property of the KL transform: virtually all the signal energy is contained in the first 15 of the 100 principal components.

Furthermore, due to the optimal packing of information by the KL transformation, this percentage of the input energy will be that associated with features which are most similar across the traces.

SECTION 2.III: APPLICATION TO SEISMIC DATA.

a. Introduction

From the simple case of similar waveforms, I now progress to the description of waveforms within a set of seismic data. For the most part, I will discuss stacked seismic sections, which often incorporate flat-lying or dipping events with a particular

waveform across the data section.

b. Flat Lying Events.

From the way in which the covariance matrix Γ is defined, a waveform which occurs from trace-to-trace at the same arrival time will be seen as most coherent. It is this similarity of signal character across the traces, which constitutes the dominant and most coherent characteristic. If the data are highly coherent, as for example in a flat lying stacked seismic section, then we will need only a few principal components to adequately reconstruct the data.

c. Dipping Events.

The flat lying events are seen as 'most common' due to the way in which the covariance matrix Γ is defined. However, if we wished to emphasize a set of parallel dipping events, we could adjust the covariance matrix calculation in such a way as to represent preferred lags, corresponding to some desired dip direction, rather than the zero-lag which corresponds to horizontal events.

This end is achieved most easily by introducing a wedge of zeroes into the start of the data. The wedge is chosen to make an event of the specified dip appear to be horizontal in the new, shifted data field.

In effect, we are producing a new data set Y, such that:

$$y_i(t) = \delta(t-\nabla(i-1)) * x_i(t) \qquad (25)$$

where δ is the Dirac delta function, and ∇ is the dip of the beds expressed in time-samples per trace; '*' denotes convolution. ∇ can be positive or negative, depending on the dip direction. I refer later to a KL transformation on Y as a 'slant-KL' by analogy to the slant-stack.

I consider the usefulness of data reconstruction from a limited number of principal components in Chapter 3, Section 3.I.

d. Less Coherent Information.

It is of interest at this stage to consider again what is left over when we have reconstructed a data set from the first m principal components. As noted in Section 1.IIIb, the answer is that we are left with a residue which represents that which is least common in our data. This residue will consist primarily of anomalous data, described below, and random (white) noise.

If we utilized very few of the principal components in a reconstruction of a stacked seismic section, then we would construct a section which incorporated only the most flat lying coherent events. Consequently, events which were dipping, or exhibited an undulating structure, would be omitted. A reconstruction forged from the next few principal components

would comprise the dipping or undulating events. The remaining principal components would represent incoherent noise.

So, we see that a misfit seismic section, defined as:

$$x\dagger_i(t) = \sum_{j=m+1}^{p} r_{ij}\psi_j(t) \qquad i=1,\ldots,n;\ m < p \leq n \qquad (26)$$

can yield useful insight into the subtler characteristics of the data. The effectiveness of misfit reconstruction in outlining anomalous parts of a seismic section is demonstrated in Chapter 3, Section II.

Misfit reconstruction also forms the basis for Sections 3.III and 3.IV. In the former, I separate diffraction from reflection events, while in the latter I separate multiples (which have been aligned by constant-velocity normal moveout correction: our common signal) from primary events (our desired signal) in common-depth-point data.

e. The Significance of the Eigenvector Elements.

Useful insight can be gained from an investigation of the physical properties of the eigenvectors in R. The elements of the eigenvector $\bar{\mathbf{r}}_j$, which correspond to the $j.th$ principal component, denote the contribution of each seismic trace to that principal component. Consequently, if all the input traces except the third were similar, then all the elements of $\bar{\mathbf{r}}_1$ would be of similar size except for r_{31} which would be much smaller (see Levy et al.,

1983, Figure 4).

We may also note that when all the input traces are identical, the 1.st principal component alone will carry all the information necessary for signal reconstruction and the eigenvalues λ_2, λ_3, ..., λ_n will be zero.

The eigenvector elements r_{ij} tell us how much of the $i.th$ seismic trace is projected onto the $j.th$ principal component. In light of this, I use the following similarity measure to examine the 'distance' between the projections of the input traces onto a given group of principal components:

$$\Delta_{ik} = \sum_{j=1}^{m} (a_{ij} - a_{kj})^2 \qquad (27)$$

where $a_{ij} = r_{ij} / \sqrt{\lambda_j}$

The significance of this measure can be seen as follows: for $m = 1$, say, i.e. using only the first eigenvector, Δ_{ik} will be the square of the difference between the amplitudes of the projections of the $i.th$ and $k.th$ seismic traces onto the 1.st principal component. So, if $\Delta_{ik} = 0$, we infer that traces i and k have the same contributions to the 1.st principal component. If Δ_{ik} is relatively large, then the converse is true.

If we now sum over all m eigenvectors associated with the coherent energy, and find that Δ_{ik} is still near zero, then we infer that traces i and k have similar projections onto all m

principal components, and further, that traces i and k are very similar. We would determine m from the requirement that we incorporate the effects of the most significant percentage of the input data energy (using *equation (22)*). This method was used by Tjostheim and Sandvin (1979) to discriminate between seismic events and nuclear blasts on the basis of the eigenstructure of groups of events.

SECTION 2.IV: THE COMPLEX KARHUNEN-LOÈVE TRANSFORMATION.

a. Introduction.

The important characteristics of the KL transformation enumerated by Kramer and Mathews (1956) for real signals carry over directly to the case when complex signals are used. The rationale for considering complex signals is that we will be able to address the problem of phase changes in the signal (Levy et al., 1983).

Most of the applications considered in this work using the KLT can be given a further degree of freedom (i.e. that of phase change) by utilizing the complex KL transformation. Whenever the complex KLT is used as well as the real KLT, I will draw a comparison between the results.

In general, the complex KL (or CKL) transform will achieve a greater compression of data than the real KL (or RKL) transform.

This is due to the ability of a phase change to make certain signals even more similar. Further, the complex technique allows us to deal directly with frequency domain data, from which a signal common to within a phase shift, at each frequency in the band of interest, can be extracted.

Given that we record a real signal, we are faced initially with the problem of constructing a complex analogue in order to proceed. This is most readily accomplished by considering the complex trace (or analytic signal) of the data $x_i(t)$ (eg. Taner et al., 1977; Bracewell, 1978).

b. The Complex Trace.

Given a wavelet $w(t)$, its complex trace $\hat{w}(t)$ is defined as:

$$\hat{w}(t) = w(t) - i\underline{w}(t) \qquad (28)$$

where $\underline{w}(t) = H[w(t)]$ is the Hilbert transform (eg. Bracewell, 1978; Aki and Richards, 1980; Levy and Oldenburg, 1982) of the initial wavelet.

To introduce a pure phase change to a wavelet, we multiply its complex trace by a complex exponential. Consequently, a wavelet phase shifted by an amount ϵ, is given by:

$$w(t;\epsilon) = Re[\hat{w}(t)\ exp(-i\epsilon)]$$

$$= cos(\epsilon)\ w(t) + sin(\epsilon)\ \underline{w}(t) \qquad (29)$$

For complex signals, the covariance matrix is Hermitian and positive semi-definite and hence a unitary matrix U, is required for diagonalization (Strang, 1980), i.e.:

$$\hat{\Gamma} = \hat{X}\ \hat{X}^H = U\ \Lambda\ U^H,$$

where $\hat{X}$ is complex; 'H' denotes the Hermitian, or conjugate, transpose.

The eigenvalues will still be real, but the eigenvectors are complex. Nevertheless, the truncation error associated with a reconstruction is still given by *equation (22)*.

c. Identical Signals.

As in Section 2.IIa, if we consider the case where $\hat{x}_i(t) = c_i\ \hat{s}(t)$, but where the c_i are now complex constants, and $\hat{x}_i(t)$ and $\hat{s}(t)$ are also complex, then the first principal component will be:

$$\hat{\psi}_1(t) = \left(\sum_{i=1}^{n} |c_i|^2 \right)^{1/2} \hat{s}(t) \qquad (30)$$

This latter equation shows that complex signals which differ only by a complex scale factor can be represented exactly by a single complex principal component and associated weights (the complex elements of the first row of U).

As an illustration, we consider the following simple example. Let $\hat{x}_1(t) = \hat{s}(t)$ and $\hat{x}_2(t) = exp(-i\epsilon)\ \hat{s}(t)$. The energy

in the signal is:

$$||s||^2 = \int_0^T \hat{s}(t)\ \hat{s}(t)^*\ dt$$

where '*' denotes the complex conjugate. Thus:

$$\hat{\Gamma} = ||s||^2 \begin{bmatrix} 1 & exp(-i\epsilon) \\ exp(i\epsilon) & 1 \end{bmatrix}$$

The eigenvalues of $\hat{\Gamma}$ are $\lambda_1=2||s||^2$, $\lambda_2=0$. The unitary eigenvector matrix is:

$$U = 1/\sqrt{2} \begin{bmatrix} 1 & exp(-i\epsilon) \\ exp(i\epsilon) & 1 \end{bmatrix}$$

The first principal component is:

$$\hat{\psi}_1(t) = u_{11}\ \hat{x}_1(t) + u_{21}\ \hat{x}_2(t) = \sqrt{2}\ \hat{s}(t) \qquad (31)$$

where u_{ij} are the elements of $\bar{u}_1$, the first column eigenvector of U.

d. Phase Recovery.

Let us now consider the two signals:

$x_1(t) = w(t)$, and

$x_2(t) = w(t;\epsilon)$.

Their corresponding analytic signals are:

$$\hat{x}_1(t) = \hat{w}(t)$$

$$\hat{x}_2(t) = \hat{w}(t)\exp(i\epsilon)$$

and hence these signals are like those considered in the preceding example. Application of the CKL transformation will produce a complex first principal component equal to $\hat{w}(t)$. Furthermore, substitution of the expressions for $\hat{x}_1(t)$ and $\hat{x}_2(t)$ into *equation (31)* demands that the constants u_{ij} are such that the phase shift ϵ is cancelled. In other words, *equation (31)* shows that the phase rotation in the second signal is recoverable directly from the eigenvector associated with the first principal component. That is:

$$\epsilon = \tan^{-1}(u_{21}) \qquad (32)$$

e. The Complex KL Transformation and Time Shifts.

Adjacent seismic signals differ not only because of phase shifts of the wavelet but also because of time shifts, that is, 'residual statics'. If these time shifts are small they can be accommodated by a phase shift of the waveform. To see this we consider a signal $x(t)$ and its Fourier transform $X(f)$. If:

$$x(t) \text{ <----> } X(f)$$

Then:

$$x(t-t_0) \quad <----> \quad X(f)\exp(i\epsilon),$$

where $\epsilon = 2\pi f t_0$, and t_0 is the time shift.

Although this shows that the phase shift cannot truly be regarded as a constant, for a sufficiently band limited signal with center frequency f_c, the frequency dependence can be neglected and using the notation of *equation (29)*, $x(t-t_0)$ can be written as $x(t;\epsilon)$ with ϵ redefined as:

$$\epsilon = 2\pi f_c t_0$$

When presented with two signals offset by a static time shift, the complex KL routine will rotate the displaced wavelet in an attempt to minimize the least-squares misfit between the real part of the two signals. In Figure 2.2, we see a series of plots of the phase rotation angle ϵ (returned by the complex KL routine) versus the static time shift t_0 (in sample points) between two Ricker wavelets. Each plot is for a different wavelet centre frequency (ergo bandwidth). The figures shown are for centre frequencies: 5, 15, 25, 35, 45, and 55 Hz, respectively. The time shift over which the locus is a straight line decreases for higher freqencies, as expected, since the waveform is then narrower and the static shift has a more pronounced effect. However, the linearity of this relationship is not diagnostic of the usefulness of the phase rotation. As the waveforms shift progressively, so the phase shift increases. We are maximizing

the commonality of the real parts of the complex wavelets, but this does not necessarily make them look similar. Thus, the linearity does not indicate whether the shifted wavelet, when rotated, will resemble the reference. To ascertain this, we look at the actual rotated waveforms, or at the percentage of energy present in the first principal component. This will equal 100% when the identical waveforms are coincident, and fall-off to 50% when there is no overlap at all.

Up to a time shift prescribed by *equation (33)* below, all rotated-wavelet distortions are small enough to produce a good stacked result, and the slope of the straight line segments of ϵ versus t_0 (Figure 2.2) changes uniformally with centre frequency down to about 5 Hz. The actual waveforms are shown in Figure 2.3. Here we see (1) the reference, (2) the time shifted wavelet, and (3) the shifted wavelet from (2) after phase rotation by ϵ, for three different time shifts and centre frequencies. As can be seen visually as well as from the large $\chi(1)$ values, the rotated time shifted wavelets are very similar to the reference pulse, and are more likely to stack contructively than are their unrotated counterparts.

For the Ricker wavelets of centre frequency f_c I found empirically that the phase shift versus static time shift approximation was valid over time shifts given by:

$$t_0 < (0.4/f_c) \text{ seconds} \qquad (33)$$

and that the relationship between centre frequency and phase rotation per millisecond shift was:

$$\epsilon_{ms} = 0.006\ f_c$$

Equation 33 has a simple origin related to the shape of a Ricker wavelet. The side lobe to side lobe breadth of a Ricker pulse, b, is given by:

$$b = \sqrt{6}/(\pi f_c)$$

and for visually acceptable phase rotations, which keep 85 - 90 % of the energy in the first principal component, I noted that the time shift should be $\leq b/2$, as exemplified by *equation (33)*. This translates into phase shifts of ≤ 2.0 radians.

f. Complex Eigenvectors.

I introduce here an extension of *equation (27)* to deal with the complex eigenvectors generated by the complex KL transform (Levy et al., 1983). In this case we would be seeing differences in phase structure as reflected in the 'phase' of the complex eigenvector elements (by which I mean $atan\{Im(r_{ij})/Re(r_{ij})\}$), as well as in the amplitudes of the projections.

For the complex case *equation (27)* becomes:

$$\Delta_{ik} = \sum_{j=1}^{n} (a_{ij} - a_{kj})\ (a_{ij} - a_{kj})^* \qquad (34)$$

where '*' denotes the complex conjugate, and the a's are now

complex, i.e.:

$$a_{ij} = r_{ij} \; / \; \sqrt{\lambda_j}$$

for the complex r_{ij}.

I utilize both *equations (27)* and *(34)* in Section 4.I, where I review the approach of Hagen (1982) and assess the limitations of the applicability of cluster analysis to seismic data.

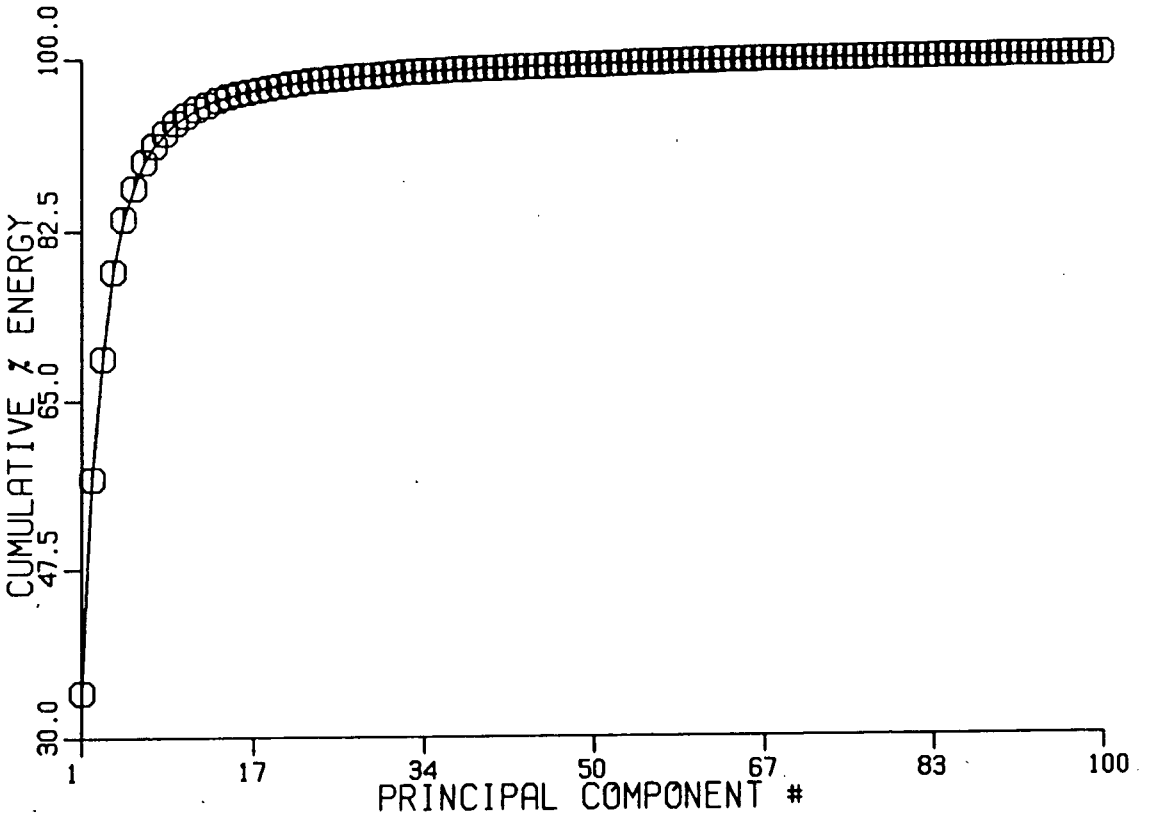


FIGURE 2.1

The cumulative percentage energy from the RKL decomposition of the 100 traces of seismic data shown in Figure 3.5a. Note how virtually all of the signal energy is contained in the first 15 principal components.

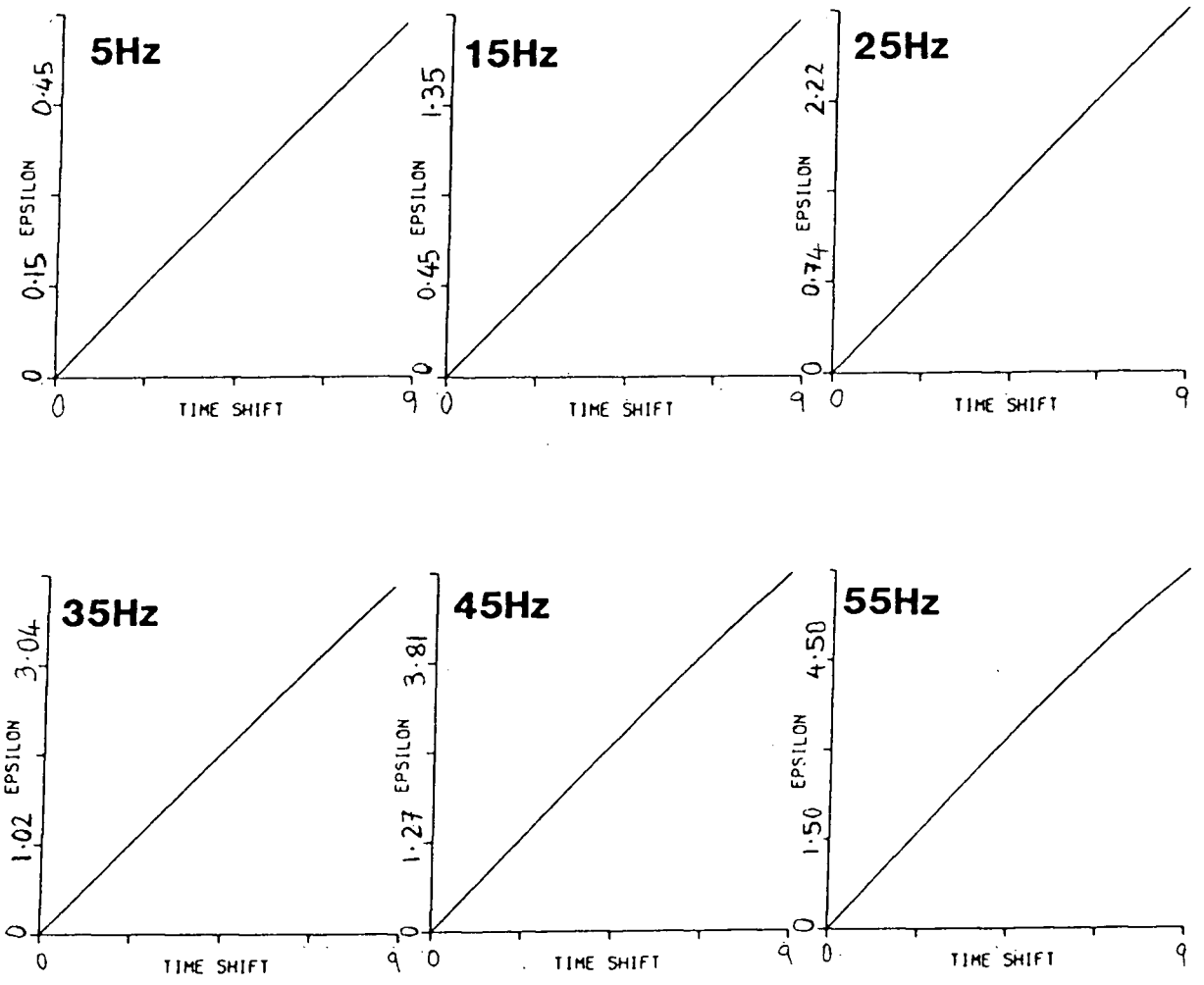


FIGURE 2.2

Eight plots each showing the requisite phase shift necessary to minimize the least-squares misfit between the real part of time displaced signals. In this case, the signals are pairs of Ricker wavelets with centre frequency 5, 15, 25, 35, 45, 55, 65, and 75 Hz, respectively, sampled at 2 ms. The time shift axis is given in sample points.

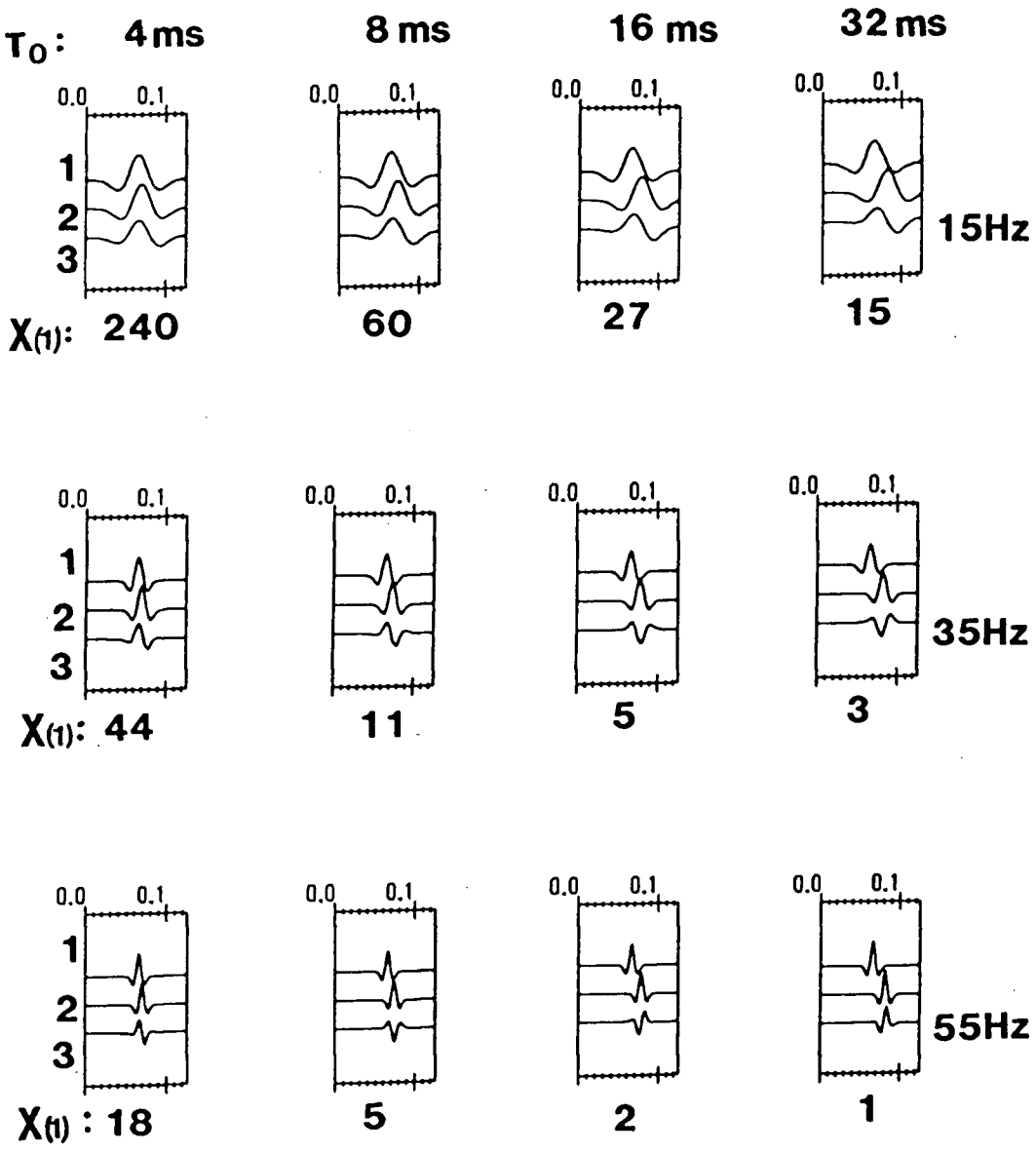


FIGURE 2.3

Three sets of plots each showing (1) the reference Ricker wavelet, (2) the time shifted wavelet, and (3) the shifted wavelet after phase rotation by ϵ, for three time shifts. Indicated below each are t_0, $\chi(1)$, and ϵ. In this case, the signals are pairs of Ricker wavelets with centre frequency 15, 35, and 55 Hz, respectively.

CHAPTER 3.

DATA RECONSTRUCTION.

SECTION 3.I: RECONSTRUCTION OF COHERENT INFORMATION.

a. Introduction.

Here I employ the energy-packing property of the KL transform. Recall that for a set of similar input traces, there will be one principal component that contains more energy than any other. The principal components, when ordered in decreasing energy content, afford the identification and subsequent elimination of incoherent energy present in the input data. Furthermore, the method also allows reconstruction of data to within a desired accuracy, from a relatively small number of principal components.

Before progressing to examples of reconstruction of stacked seismic sections, I introduce an example of stacking in the presence of time and phase shifts by representing a stack by a sum of the most significant principal components resulting from the decomposition of a move-out corrected gather. This stacking method may be performed using either the real or complex KL transformation.

b. KL Stack Versus the Mean Stack.

A complex KL analysis of a set of signals having time shifts which are random and have zero mean will have a first principal component that is like the true signal. This suggests an immediate application in stacking signals contaminated with static shifts (Levy et al., 1983). The same is true to a lesser extent for a real KL analysis; Hemon and Mace (1978) and Ulrych et al. (1983) both demonstrate how the first principal component from a real KL decomposition of a moveout corrected CDP gather can be used as an estimate of the stack. The assertion is that after moveout correction, there will be an unknown trace which is common to all the seismic traces: this is usually estimated by extracting the mean of the input ensemble. For the KL case we estimate that common trace by constructing the first principal component, which is a least squares estimate of the common signal.

A further step, which was not fully explored in the above papers, is the ability to reconstruct the zero-offset representation using a small number of principal components which are associated with the largest eigenvalues. This possibility is particularly interesting for cases where a phase shift cannot fully account for the observed phenomenon (as would be the case with residual statics problems). In effect, I replace the mean

stack:

$$\bar{x}(t) = (n^{-1}) \sum_{i=1}^{n} x_i(t) \qquad (35)$$

with a sum of the most energetic principal components:

$$\bar{\bar{x}}(t) = (n^{-1}) \sum_{i=1}^{n} \sum_{j=1}^{m} r_{ij} \psi_j(t) \qquad (36)$$

i.e., I have simply replaced $x_i(t)$ in *equation (35)* with its truncated reconstruction from *equation (3)*.

c. Synthetic Data Examples.

In the presence of small time (or phase) shifts, the first principal component should provide a better representation of the true signal than the mean stack. This is illustrated in the examples in Figure 3.1. In Figure 3.1(a) 15 signals have been randomly shifted by amounts up to ± 20 ms. The mean stack of these signals has been computed and it is seen (Figure 3.1(a) bottom) to be a poor representation of the initial signal. On the other hand, the RKL and CKL stacks more closely represent the initial signal. In Figure 3.1(b) I show analagous results when the signals differ only because of phase shifts. The mean stack has very small amplitude because of cancellation effects resulting from the phase shifts. The CKL stack of course reproduces the input signal exactly, while the RKL stack fares somewhat better than the mean stack. Figure 3.1(c) shows the mean stack, RKL, and CKL stacks of signals that are phase shifted and

contaminated by statics.

In all cases, the KL stack yields a better representation of the input signal than does the mean stack. This should not surprise us since the CKL algorithm possess two (and the RKL stack one) additional degrees of freedom which are not utilized in the mean stack. In the construction of the principal components, each trace is weighted according to the degree of correlation it exhibits with the other input signals; also, for the CKL stack before the construction of the principal components, each trace is further rotated by a constant phase shift so that the total correlated energy in the real part is increased. Unless otherwise specified, for a complex KL analysis, rotation of the signals to the phase of the uppermost ('near offset') traces was carried out throughout this work under the assumption that we are generally interested in the reconstruction of the zero-offset trace.

d. Reconstruction of Stacked Data.

Moving on to the situation where we have stacked seismic sections, I consider the following synthetic examples to demonstrate how reconstruction is able to segregate different events. It was found that the CKL method tended to obliterate the small dips and undulations in seismic sections, replacing them with flat lying phase shifted waveforms. Consequently, all reconstructions shown in this Chapter are derived from the RKL

transform which produced excellent results.

It should be noted that in plotting all the following figures I normalize to the largest value; hence the maximum amplitude in all figures is arbitarily set to unity.

e. Synthetic Data Examples.

1. Twenty four synthetic seismic traces representing shallowly dipping events superimposed on steeply dipping coherent 'noise' events, such as those introduced by marine streamer cable motion, or ground roll, into a final stacked seismic section (Larner et al., 1983) were constructed. I introduced a progressive phase change (from 0 to $\pi/3$ radians) to the wavelets across the events representing the geological horizons, and introduced a vertical fault of offset 36 ms at the 10.th trace. I then added 10% white noise (all noise levels are expressed here as a percentage of the maximum trace amplitude).

Figure 3.2a shows the seismic representation of the basic geological model: phase shifted wavelets in shallowly dipping horizons offset by a vertical fault. Figure 3.2b is the 'noise' section, i.e. steeply-dipping events plus random noise. Figure 3.2c shows the contaminated synthetic seismic section described above (i.e. the sum of 2a and 2b). In Figure 3.2d we see a 75% reconstruction of the data: the criterion governing reconstruction was the reconstruction energy, defined in *equation*

(23). We note that the events which deviate from the flat-lying character of the model have been severely attenuated, leaving the representation of the underlying 'geological' structure basically intact.

The phase character of the signals has been preserved, as has the distinctiveness of the 'fault' edge. However, we note an increase in the background noise in comparison to Figure 3.2a. This is a residual effect brought about by discarding principal components which contained contributions from both the dipping noise and the flat lying waveforms. The 24 principal components corresponding to the data in Figure 3.2c are seen in Figure 3.2e. The first 5 principal components are characteristic of flat lying events, and combine with the appropriate weights to produce the data of Figure 3.2d. The next few principal components (6 - 12) show a sawtooth pattern (shaded in the figure) which combines to form the dipping waveforms of the input data. The principal components have been normalized for plotting, otherwise only the first few would have a discernable amplitude.

2. The second example I consider is similar to the first, but with 50% white noise added to the first 7 and last 8 traces. Note also that the fault location is between traces 7 and 8 in this example. In Figure 3.3a and b, we see the seismic data and noisy section after bandpass filtering (0 - 50 Hz). Attempting to isolate the model from the noise, I reconstructed the seismograms to 72% of the total energy. As is seen in Figure 3.3c, this

attempt has been reasonably successful: the noise level has been considerably reduced, and the phase character and fault edge have been well preserved. For the most part, the fault edge is not smeared over adjacent traces but remains distinct.

3. If the geological model had comprised more steeply-dipping parallel (or sub-parallel) beds, then I could have used the 'slant KL' procedure described earlier (Section 2.IIIc). This will bias the segregation of dipping events in favour of those with the specified dip (*equation (5)*). This procedure is demonstrated in the following example.

A data set similar to that used in example 1 was constructed, but the dips on the 'geological' horizons were increased by a factor of 5. Following the same procedure as before, we see in Figures 3.4a, b, c, d, and e, the seismic section; the section plus noise and 'streamer noise'; the data after a dip of 8 ms per trace has been removed; the 75% reconstruction depicting the underlying structure, and finally, the reconstructed data with the dip re-instated. The phase character and fault definition are well preserved, whereas the dipping noise events have been severely attenuated.

f. Real Data Examples.

The following figures exemplify the ability of KL reconstruction to enhance coherency in real stacked data. In all

reconstructions we will immediately see how the uncorrelated background noise level has been reduced. Trace-to-trace coherency has been somewhat enhanced, as the eye is no longer distracted by a background of incoherent energy. The storage requirements associated with reconstructions of different energies were considerably reduced and are mentioned later.

In most of the real data examples supplied to me, no information concerning the trace spacing, section location, geology, or hydrocarbon potential was made available. Consequently, the description of the data presented here is sketchy.

For processing large seismic sections, I break the data into overlapping trace and time segments. This is necessary for two reasons. First is the underlying assumption that we have parallel flat (or dipping) events, an assumption which breaks down unless we look at small windows (eg. 100 traces by 1 second). Second is the computational expense. The time for decomposition of a matrix using SVD is proportional to the cube of the number of traces (Strang (1982)). Consequently, we do not desire to decompose segments with more than about 100 traces. For example, decomposition of 530 traces of data, each of 500 points, broken into overlapping segments of 85 traces by 150 points, took 70 minutes of CPU time on a Perkin Elmer 3220. The corresponding reconstruction took a few minutes. I refer to reconstruction of large sections of data by this overlapping technique as compound

reconstruction.

1. In Figure 3.5a, I have selected a window of 100 traces from a conventionally processed stacked section. The traces were decomposed into their principal components: the first five principal components alone accounted for 85% of the input energy, and the first 12 principal components for 95% of the energy. In other words, for this data I need only 12% of the principal components to almost perfectly reconstruct the input data. Figures 3.5b and 3.5c respectively show the 95% and 85% reconstructions. Note the increase in coherency in the zone 1.1 - 1.5 seconds, which consists of interbedded sand-shale sequences.

In the 95% reconstruction, we note the general increase in coherency, while at the same time small-scale features are preserved. For example, the small lens-like feature depicted in the box at about 1.1 seconds between traces 55 - 75, is enhanced in the 95% reconstruction. However, discarding more principal components as is done for the 85% reconstruction leaves the section looking very smooth and continuous, especially over areas with small scale features. The lens-like feature is obliterated in this reconstruction. The phase structure of the major horizons remains intact in both reconstructions; this can be seen by examining the waveforms along the edges of the section, for example.

2. The second real example (Figure 3.6a) is also from a conventionally processed stacked section. The 96 traces were decomposed into their principal components. The first 12 principal components account for 85% of the input energy; the first 32 principal components for 95% of the energy. Here I need only about 30% of the principal components to almost perfectly reconstruct the input data set. Figures 3.6b and 3.6c show the 95% and 85% reconstructions, respectively.

In this example, I focus attention on two features. Firstly, in box A, we see a negative trough under a strong positive reflection. This sharp contrast is indicative of a localized high reflection coefficient, and is sometimes indicative of hydrocarbon potential. This positive-negative feature is well preserved in both reconstructions; at the same time the background noise level is greatly reduced.

The second feature noted is the bifurcation of the reflector just above 1.7 seconds in the centre of box B. This feature is preserved in the 95% reconstruction, but, by 85% we see a loss of definition of the event, indicative of over winnowing.

3. As an example of geological interest, we see in Figure 3.7 part of a section of data incorporating a buried caldera, identified by interpreters, and in Figure 3.8, the corresponding window of the 95% compound reconstruction of the data, as well as my interpretive sketch of the figure. The crater, seen in the

sketch, spans over 80 traces, and is barely visible in the original data. However, after a 95% compound reconstruction, the incoherent noise level has fallen off drastically, and the features of interest, the fine-scale structure demonstrating the draping of sedimentary strata over the edge of the crater, are more clearly seen in the reconstruction. Also of interest is the severe normal block faulting in the central part of the section. This is probably related to the meteorite impact. It is such a structure which is of interest as a potential trap for hydrocarbons, and this enhanced presentation may enable better interpretation of the data.

4. To demonstrate the slant-KL procedure, in Figure 3.9 I consider an enlarged section of the data shown in Figure 3.6. The events in this section dip to the left about 0.1 s over the 96 traces, and the computation of the covariance matrix was adjusted to accommodate this dip. The strong pair of arrivals seen above 1.0 s in the input data (a) are clarified markedly in the 95% reconstruction (b). However, the 85% reconstruction (c) loses much of the finer detail. A further example of loss of resolution is the small 'pull-up' type event in the box A. This feature, perhaps indicative of a velocity anomaly associated with a reef, is well preserved in the 95% reconstruction, but has been lost in the 85% reconstruction.

g. Data Compression.

As a corollary to parsimonious reconstruction, I note the following important point. In all the examples considered, we saw how the data could be almost perfectly reconstructed from a small subset of principal components. In other words, the required storage space will be drastically reduced, as will the data transmission speed. This latter point is of concern when data are being shunted from a main-frame computer to an auxiliary machine for processing. In the data examples processed to date, I find that we typically dispense with 70 - 90% of our required storage space (this point is exemplified in Figure 2.1, which shows the cumulative percentage energy associated with the principal components). Further, it is noted that any processing which assumes the conditions of the linear convolutional model may be performed on the principal components themselves rather than on the original section.

h. Discussion.

The main thrust of this section has been to introduce the application of the KLT to the recovery of coherent information for image enhancement of stacked seismic sections (presented here for the first time). The stacking of seismic data (first applied to moveout corrected gathers by Hemon and Mace (1978) for the

real KLT alone) was mentioned for the RKL and CKL transforms, as an introduction to the signal extraction capabilities of the method.

The stacked section reconstruction technique has proved very successful so far, and has been applied extensively to real data in an industrial processing environment. Its main advantage is in its ability to greatly reduce the background level of incoherent noise, which is often prevalent at earlier times in a seismic section. The choice of a suitable reconstruction energy is subjective, but for most of the data I have examined, 90 - 95% reconstructions seem to give reasonable results. Dropping the reconstruction energy below 90% results in loss of resolution, as seen in many of the 85% reconstructions presented here. When disc storage is a problem during intermediate processing stages, the data compression aspect of the transform is also valuable.

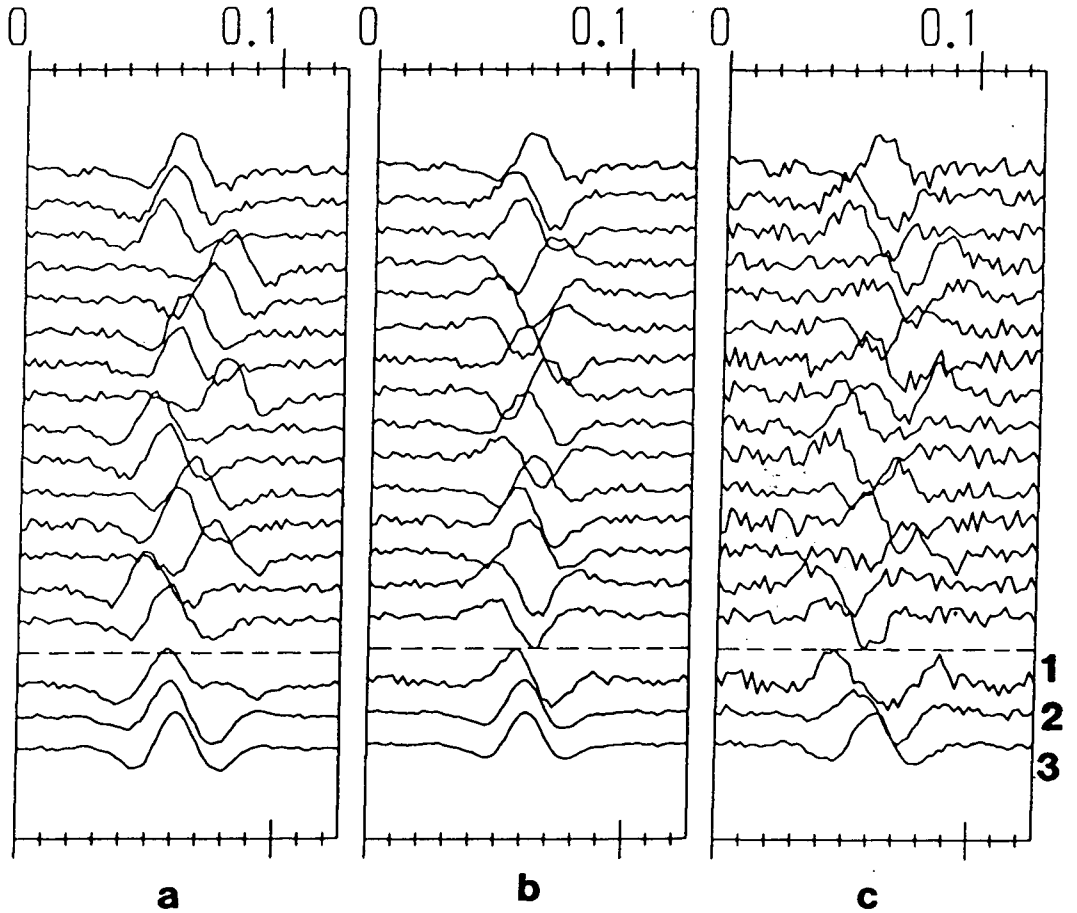


FIGURE 3.1

(a) Fifteen Ricker wavelets, plus noise (up to 40% of the maximum amplitude) time shifted by up to ± 20 ms. The lower panel shows: (1) the mean, (2) RKL, and (3) CKL stacks (the RKL and CKL stacks used one principal component).
(b) The same Ricker wavelets with phase shifts of up to $\pm\pi$. Below are: (1) the mean, (2) RKL, and (3) CKL stacks
(c) The wavelets with ≤100% noise, and up to ±20 ms static shifts, AND up to $\pm\pi$ phase shifts. Below are: (1) the mean, (2) RKL, and (3) CKL stacks

Note how in all three examples the CKL stack is better able to reproduce the underlying wavelet, and how the RKL stack performs well unless large phase shifts are involved.

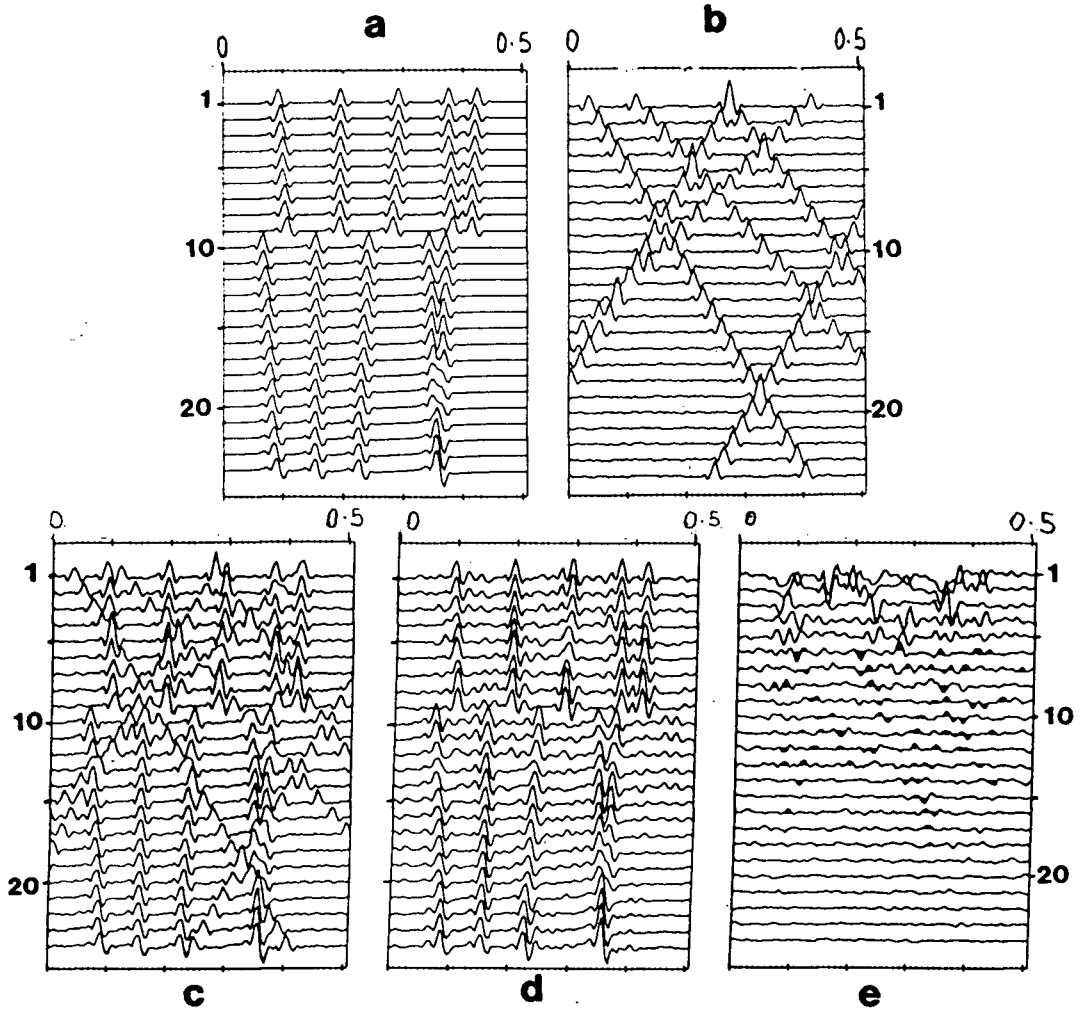


FIGURE 3.2

(a) Seismic representation of the geological model: phase shifted wavelets in shallowly dipping horizons (up to 2 ms per trace dip), offset by a vertical fault at the 10th trace.
(b) The "noise" section: steeply dipping coherent events (dips between 16 and 24 ms per trace, such as could be produced by marine streamer noise of ground roll) and 10% random noise.
(c) A sum of the previous two data sets: this is the input for the processing.
(d) A 75% reconstruction of 2c, requiring the first 5 of the 24 principal components. Note the preservation of phase information, and the clarity of the fault edge.
(e) The 24 principal components corresponding to 2c. The first 5 principal components are dominated by energetic peaks characteristic of flat lying structure. Note the smaller amplitude peaks in the band of principal components 6 - 17 (shaded in the figure). It is these waveforms which combine to form the steeply dipping events.

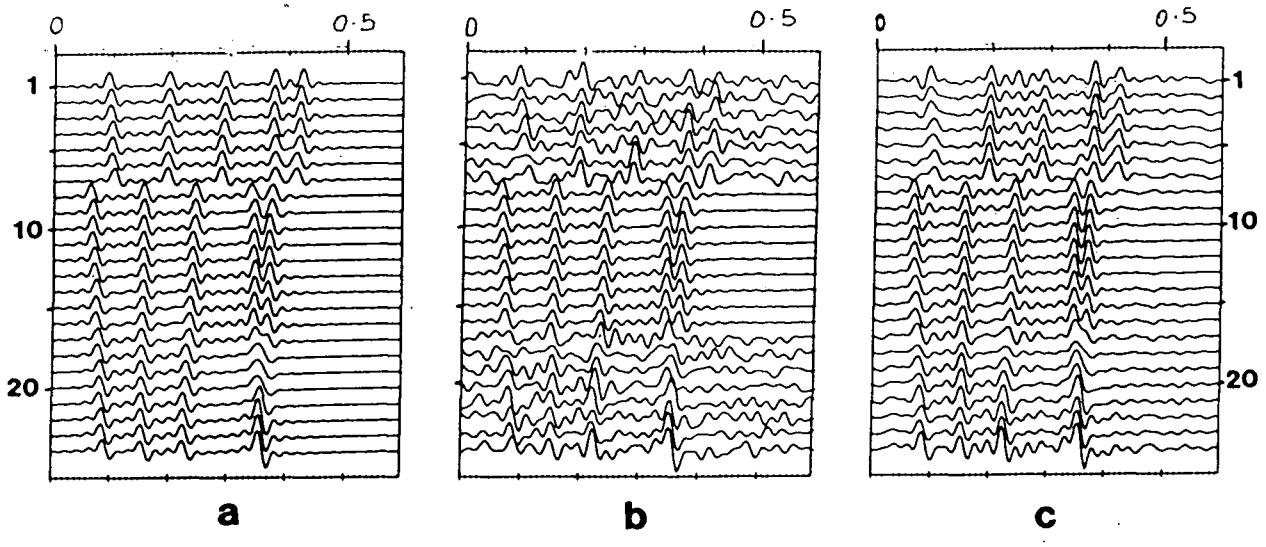


FIGURE 3.3

(a) Synthetic data similar to that of Figure 3.2a, with a vertical fault between traces 7 and 8, after bandpass filtering (10 - 50 Hz).

(b) The data after addition of 50% random noise to the first 7 and last 8 traces, and bandpass filtering as above.

(c) A 72% reconstruction requiring 4 of the 24 principal components. Note the marked reduction in the background noise level and the preservation of the essential features.

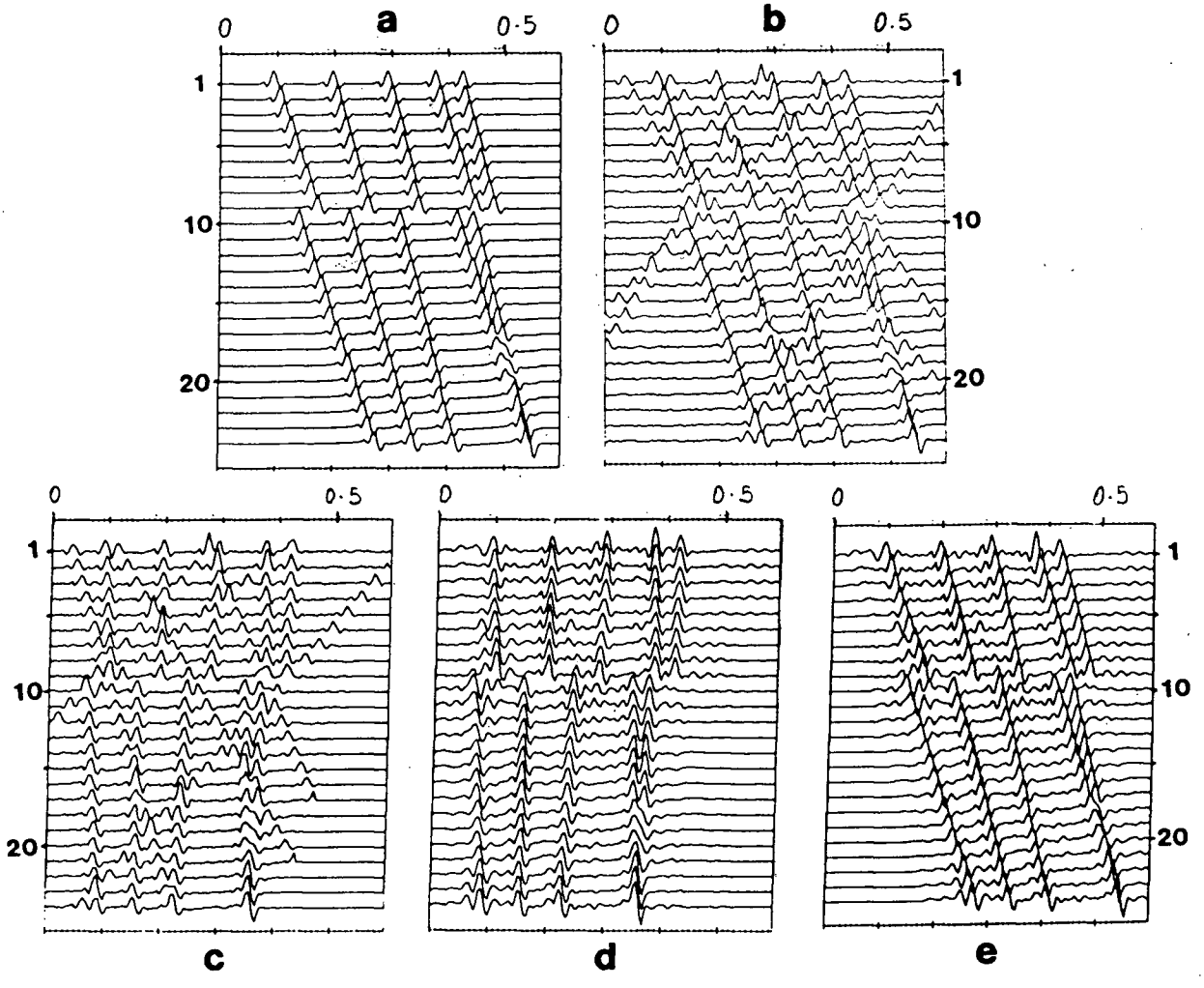


FIGURE 3.4

(a) Steeply dipping (between 7 and 10 ms per trace) sub-parallel phase shifted events with a vertical fault.
(b) The data plus the dipping "noise" events of Figure 3.2b.
(c) The flattened contaminated data, i.e. 3.4b after a dip of 8 ms per trace has been removed.
(d) A 75% reconstruction requiring 5 of the 24 principal components.
(e) The reconstruction after the dip has been reinstated. Again, the phase character and fault definition are well preserved, whereas the dipping noise events have been severely attenuated. The background noise level in places has increased as we have omitted some principal components which contained both the desired signal and the dipping noise events.

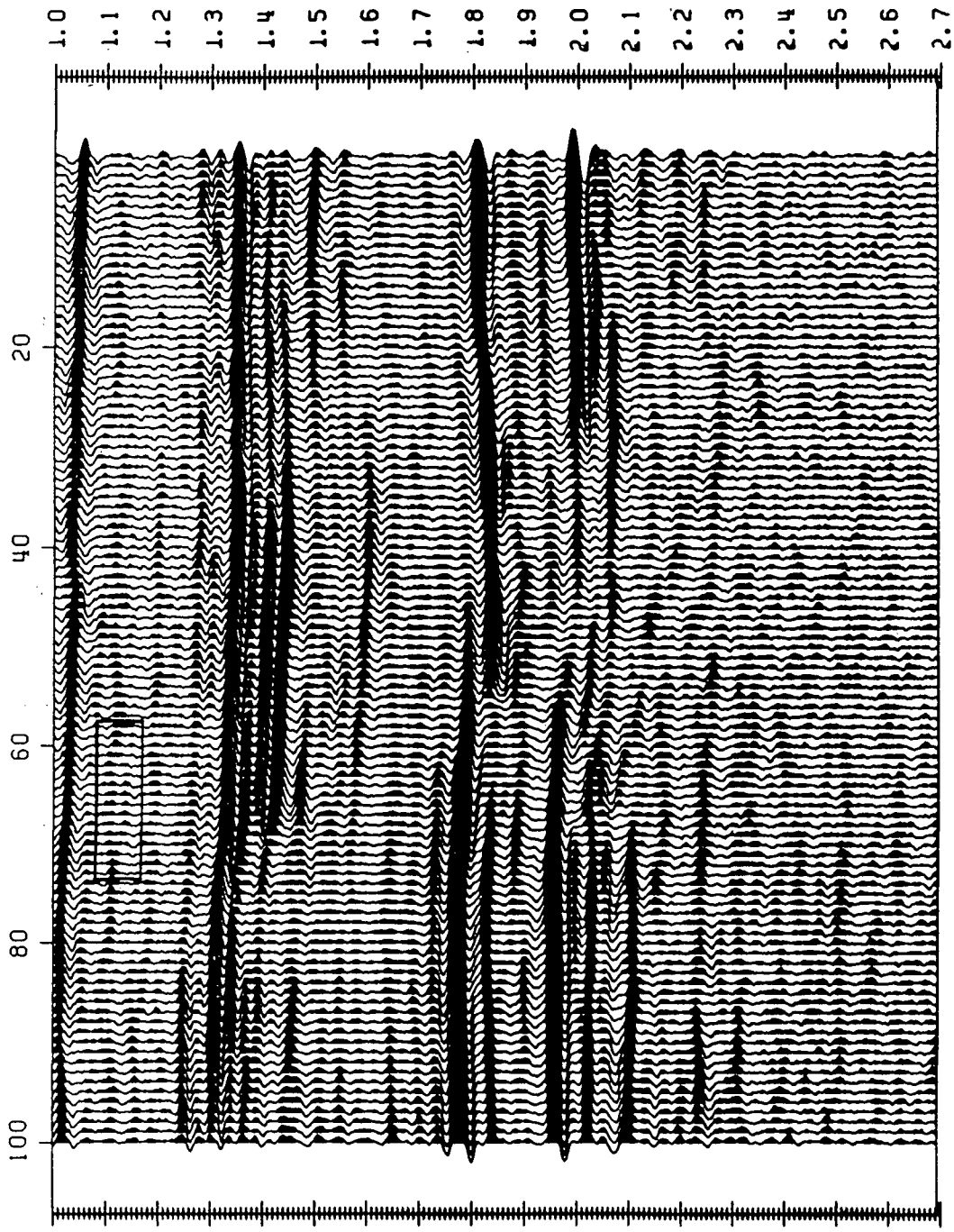


FIGURE 3.5a

One hundred traces from a conventionally processed stacked section. Note the discontinuous events in the lower part of the section (2.1 - 2.7 seconds) and the small lens-like feature in the box at 1.1 s.

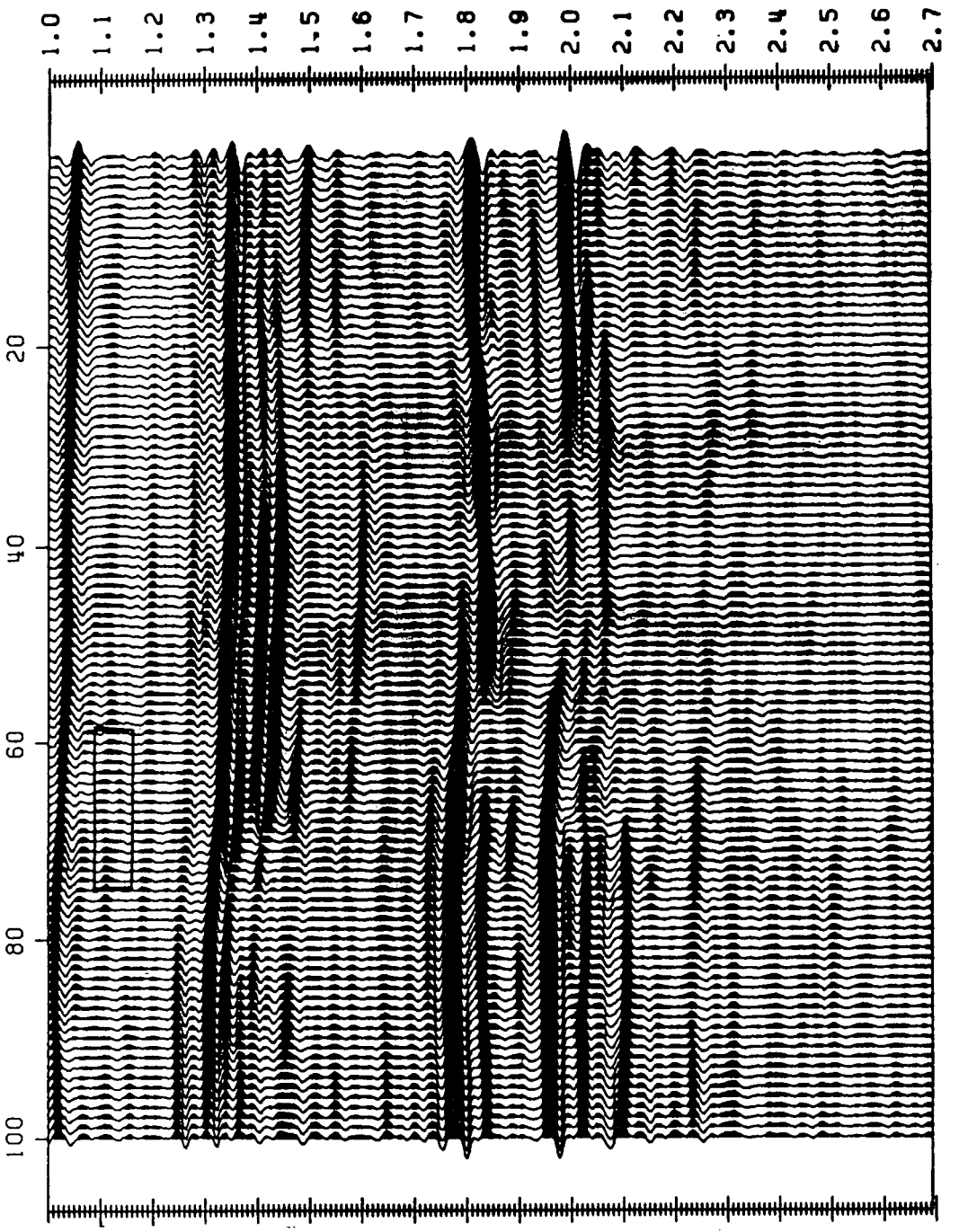


FIGURE 3.5b

The 95% reconstruction requiring 12 of the 100 principal components. In this reconstruction, the background of incoherent energy has been reduced greatly, and small scale features such as the lens highlighted in the box are preserved.

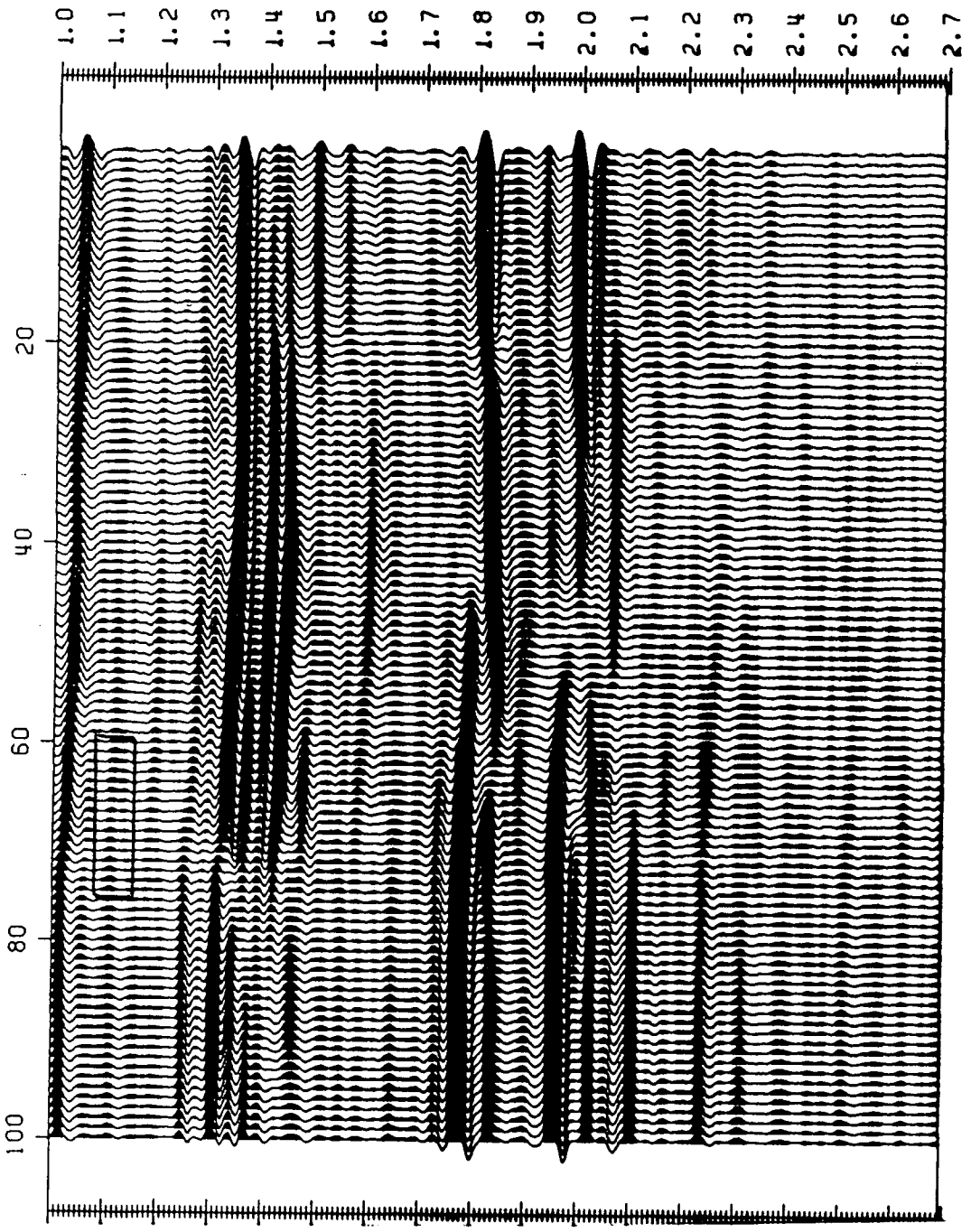


FIGURE 3.5c

The 85% reconstruction requiring 5 of the 100 principal components. In this reconstruction, the background of incoherent energy has also been reduced greatly, but small scale features such as the lens highlighted in the box have been obliterated, because by this reconstruction energy we have left only the flattest horizons.

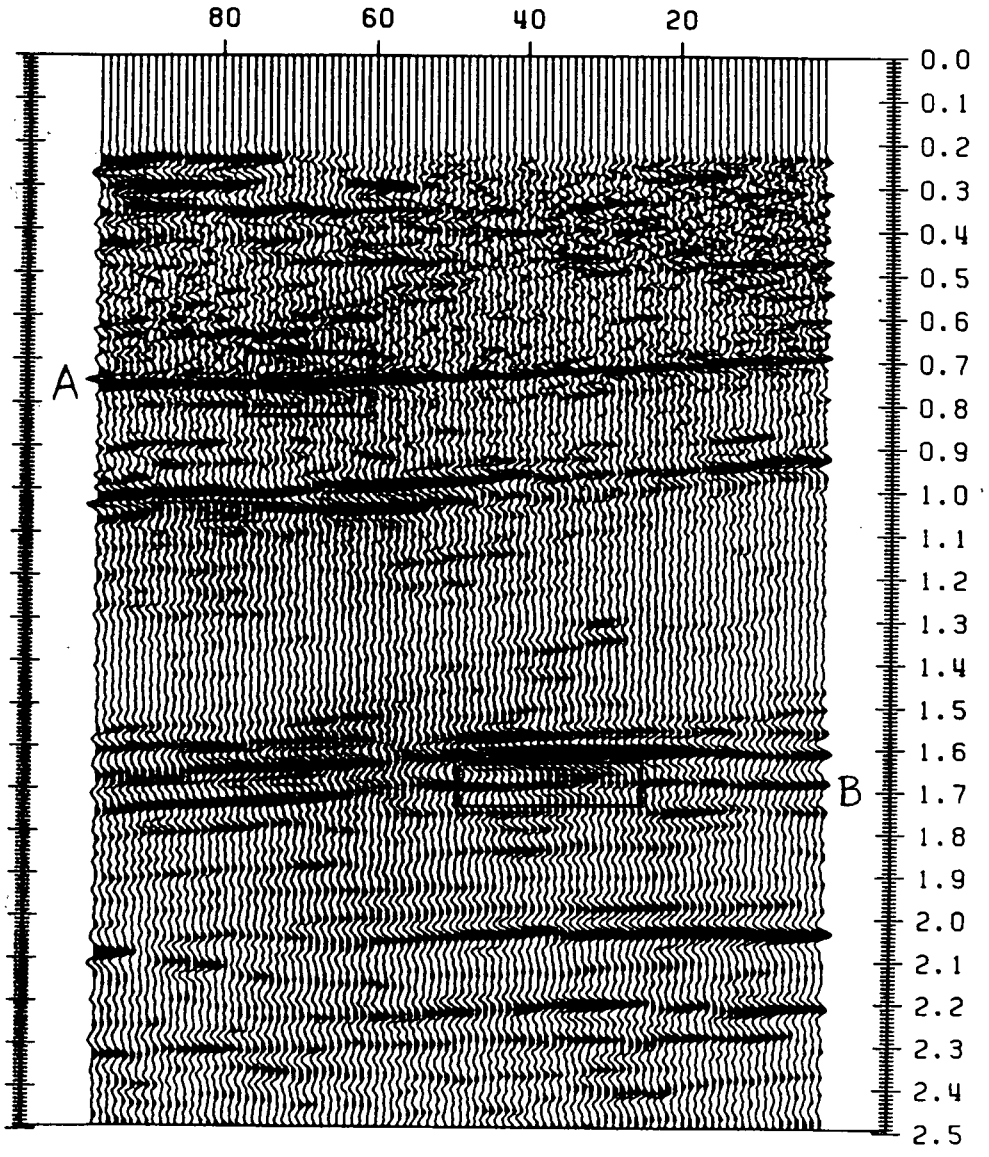


FIGURE 3.6a

Ninety six traces from a conventionally processed stacked section. Note the strong negative trough below the prominent reflection in box A, and the bifurcation of the horizon in box B.

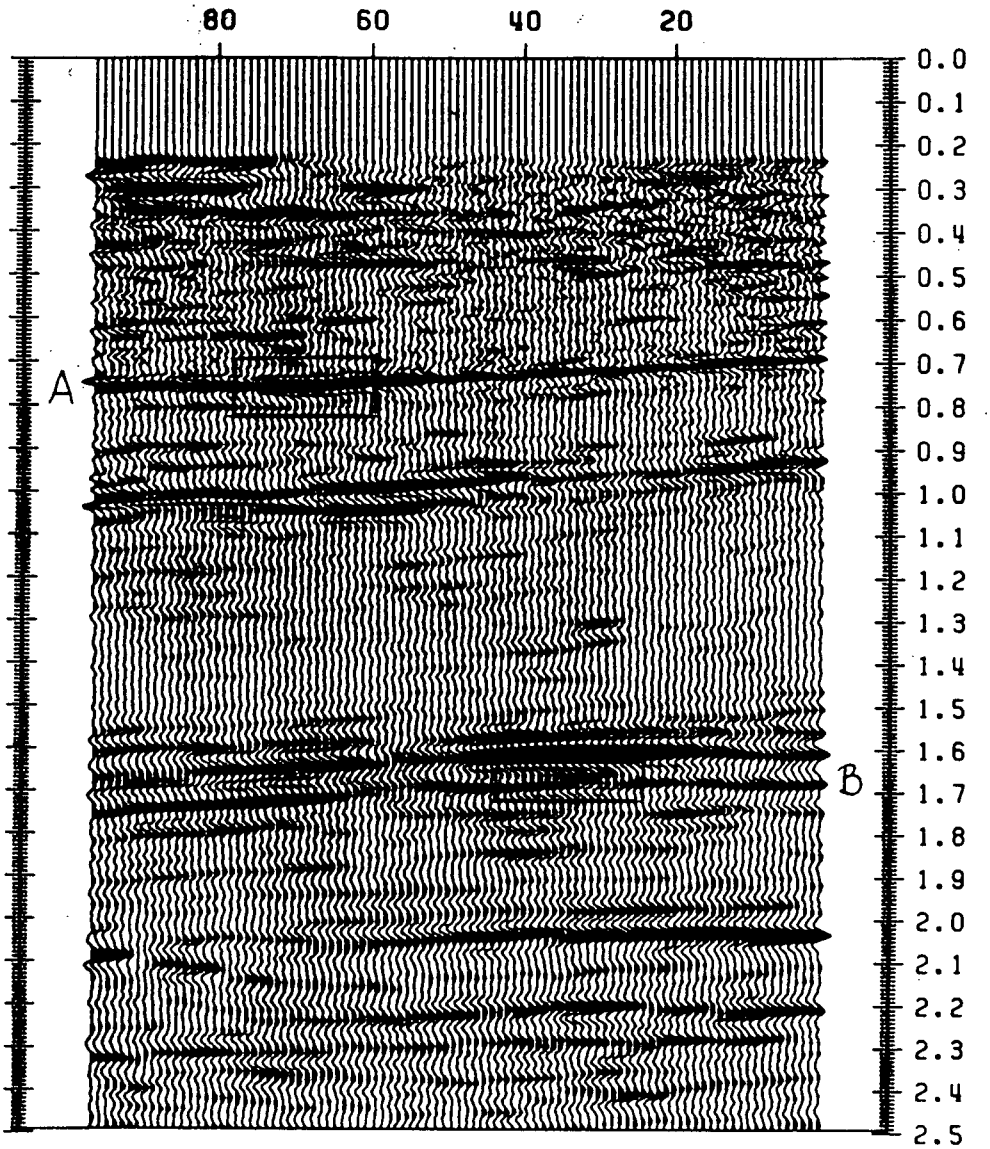


FIGURE 3.6b

The 95% reconstruction requiring 32 of the 96 principal components. In this reconstruction the positive-negative contrast in box A is well preserved, while the background noise is severely attenuated. Also, the bifurcation of the horizon at 1.7 s central to box B is well preserved.

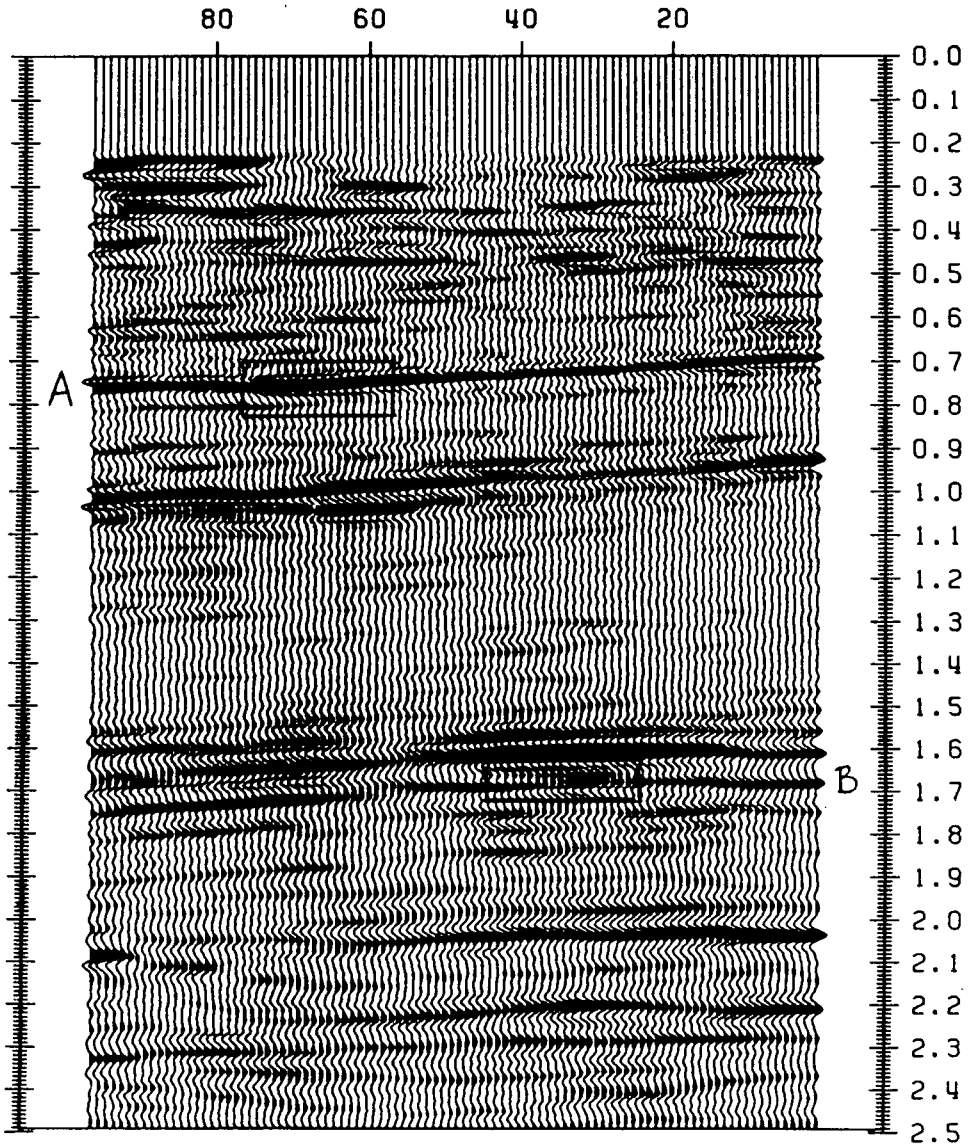


FIGURE 3.6c

The 85% reconstruction requiring 12 of the 96 principal components. Also in this reconstruction, the positive-negative contrast in box A is well preserved, while the background noise is severely attenuated. However, the bifurcation of the horizon at 1.7 s central to box B has lost definition.

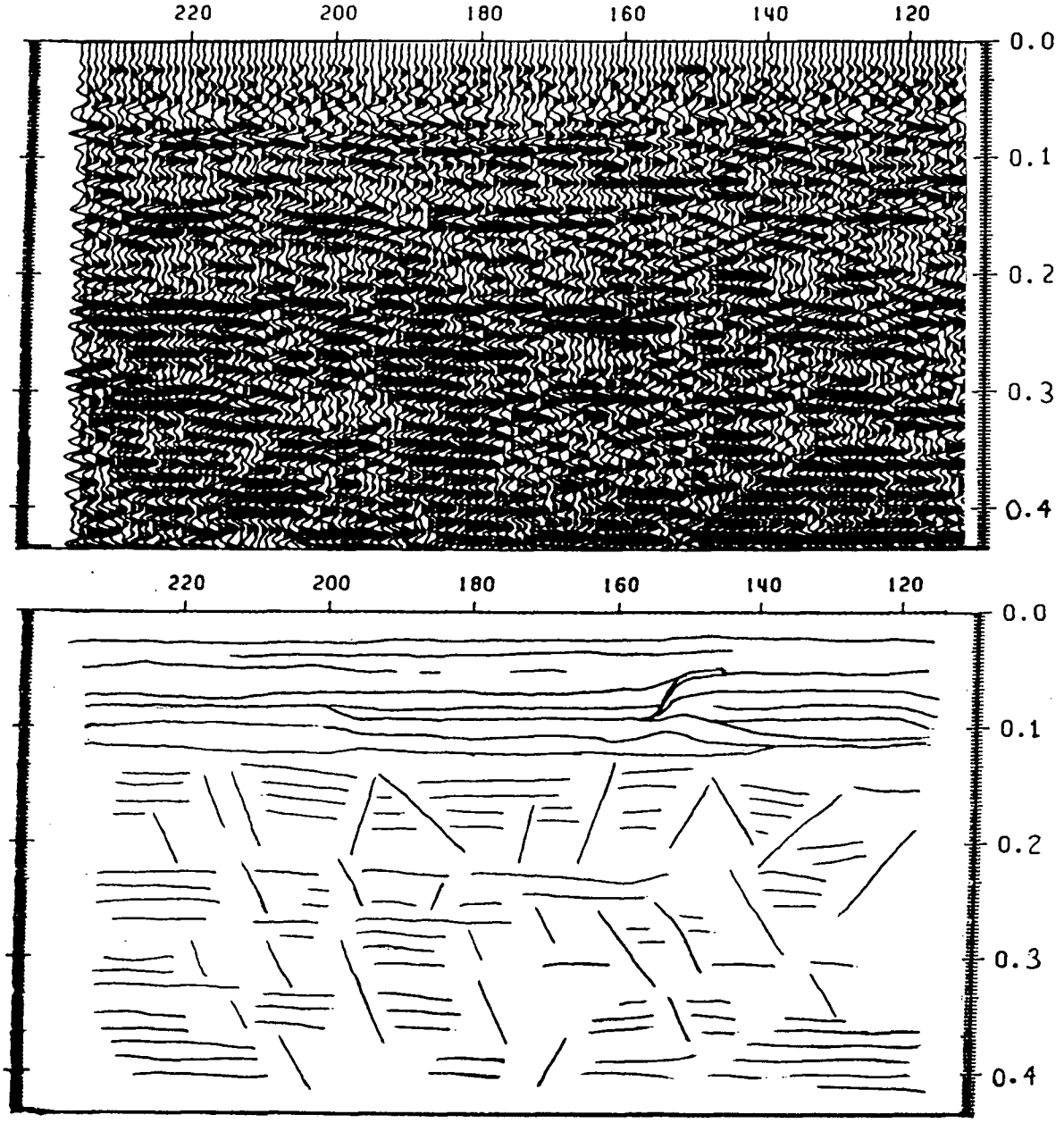


FIGURE 3.7

(a). A window of 140 traces from a section with a caldera. This feature, picked by interpreters, is not particularly clear in the input data, which is noisy, and exhibits severe faulting perhaps associated with the impact feature.

(b). Below is an interpretive sketch based on the reconstructed data.

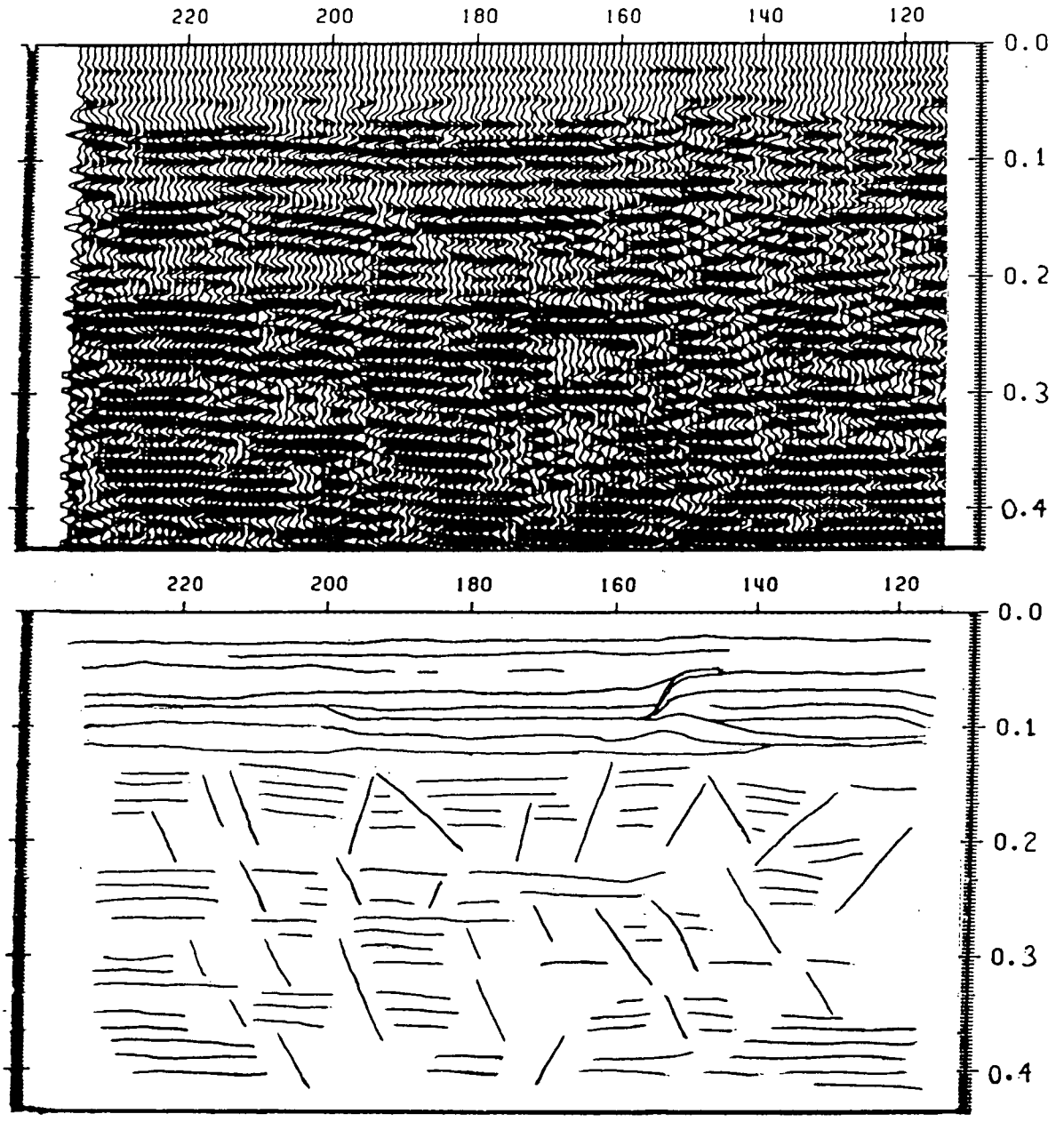


FIGURE 3.8
(a). A window of 140 traces from a section with a caldera. In the 95% compound reconstruction (see text) the crater and sedimentary fill are much clearer. Also preserved and clarified are the normal block fault edges clearly visible in the section.
(b). The sketch is reproduced below.

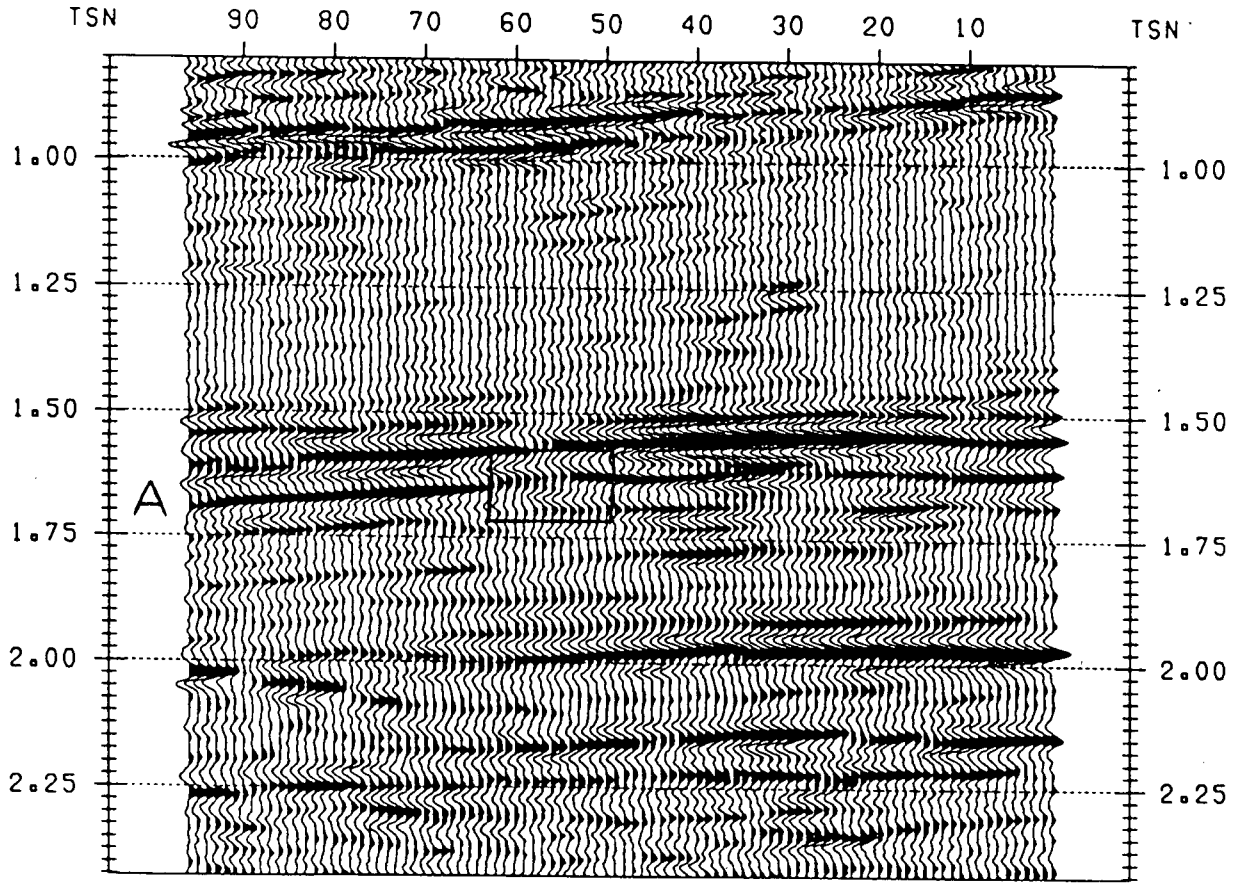


FIGURE 3.9a

Ninety six traces from a conventionally processed stacked seismic section. Note the background noise surrounding the pair of dipping events above 1.0 s. Also note the reef-like structure in box A.

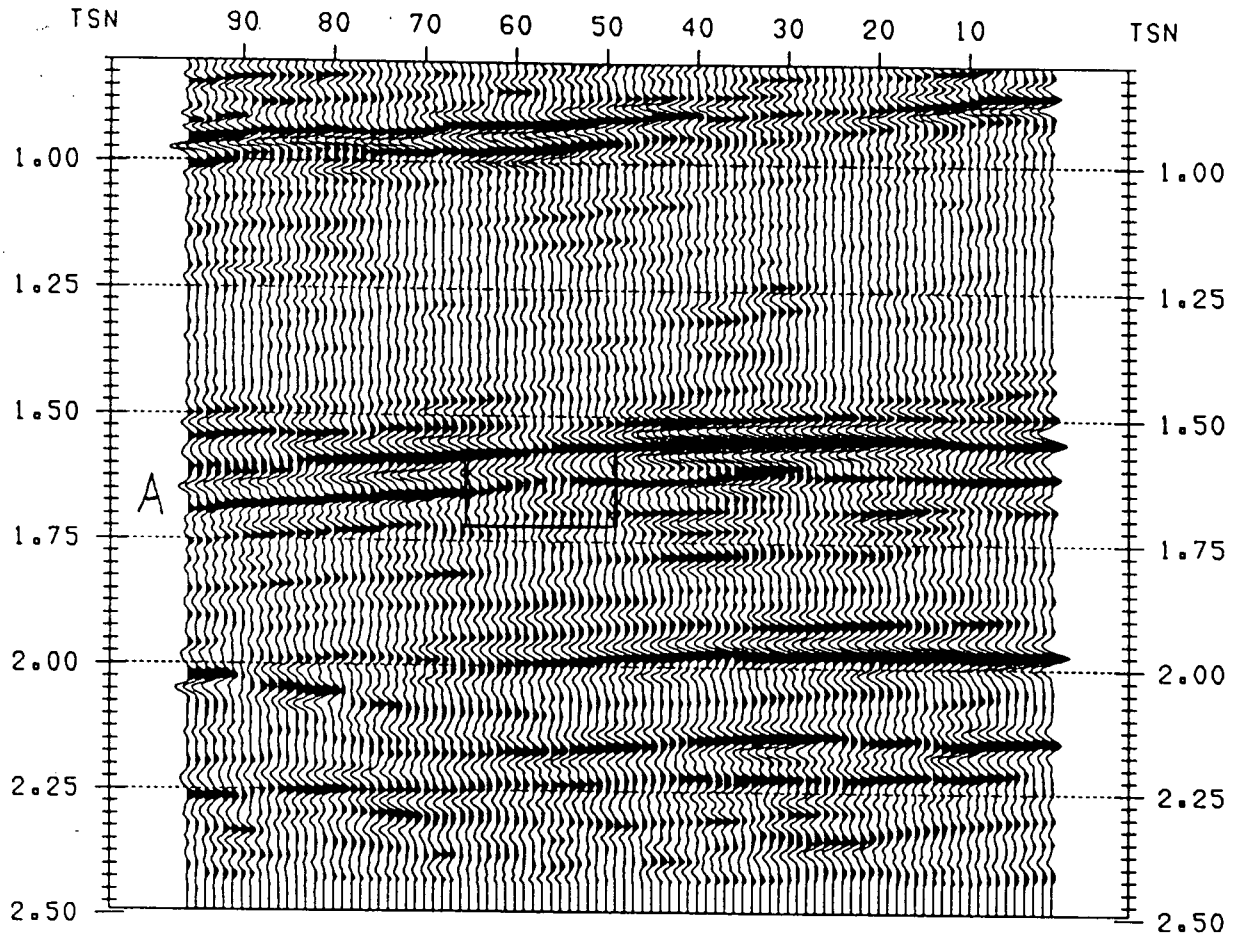


FIGURE 3.9b

Slant-KL 95% reconstruction: the dipping events have been clarified, and the reef-like structure has been left intact.

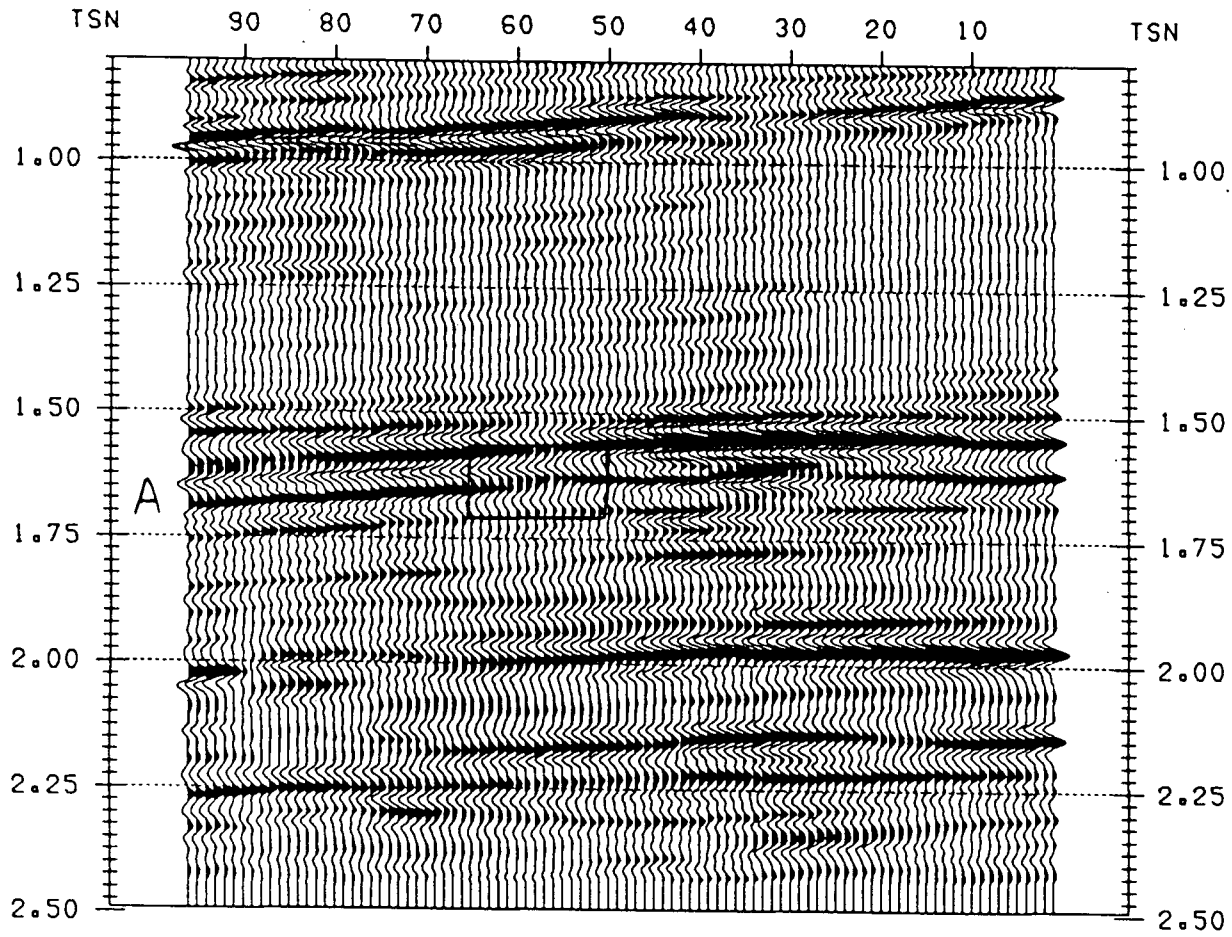


FIGURE 3.9c

Slant-KL 85% reconstruction: resolution has been lost, and the reef-like structure has been smeared-out.

SECTION 3.II: MISFIT RECONSTRUCTION.

a. Introduction.

Using *equation (23)*, I may discard the most common part of a data set and investigate the appearance and significance of the less obvious, perhaps diagnostic features which remain.

b. Synthetic Data Examples.

1. In the reconstruction of the synthetic data example (Figure 3.2d), I used 5 principal components; there, I was able to separate the steeply dipping coherent noise events from the seismic representation of the geological structure. In Figure 3.10a, I construct a misfit seismic section of Figure 3.2c by leaving-out the first 75% of the energy, and including the remainder (i.e. I sum principal components 6 to 24). A sum of the data in Figures 3.2d and 3.10a (before plotting normalization) would yield exactly the data shown in Figure 3.2c.

The same is true of the example with more steeply dipping events shown in Figure 3.4. Figure 3.10b shows the misfit reconstruction which is left over when the first 5 principal components have been omitted. A sum of Figures 3.4d and 3.10b, before plotting normalization, would yield exactly the data shown in Figure 3.4c.

2. The second example shows parallel flat-lying layers with

a central lens-like feature. In Figures 3.11a, b, and c, we see the data, a 94% reconstruction, and a misfit reconstruction, respectively. The 94% reconstruction essentially shows the basic features, i.e. flat lying events, and discriminates against the structure. However, the misfit reconstruction clearly shows the lens-like feature, as this was not a major contributing factor to the first 3 principal components. A display of this kind would serve to highlight anomalous features.

c. Real Data Examples.

1. In Figure 3.12 (a) I show data from an area with a steeply dipping fault. In the box at 1.45 s we see a strong negative event which terminates at about trace 68. In box B, we see a sinuous and discontinuous horizon extending to the right from trace 85 to its termination by the fault near trace 50. In (b) I show the misfit overlay produced by plotting a misfit reconstruction (omitting the first 70% of the energy) as grey shades over the wiggle trace plot of the original data. In box A, we note a strong black-white contrast highlighting the events at the fault edge. The whole of the upper zone tends to give large misfit values as it dips quite steeply to the right, due to the fault. In box B, we again see the black-white contrast, this time over the hump in the centre of the box. The white patch terminates to the right at the fault.

2. An example from a braided stream system, of sand and

gravel type deposits, forms the next case. In Figure 3.13(a) we see a discontinuous horizon at about 0.7 s. Individual events fade in and out, as the sand lenses terminate and begin again. The misfit overlay (b) highlights the entire central region as having many unusual features. As a whole the picture is confusing; however, we do see some features of interest. In box A, we see a white area corresponding to the pinch-out of the thin positive anomaly just above 0.7 s (arrow). A similar black-white feature (box B) corresponds to the apex of the hump between traces 1 and 20 at 0.73 s.

d. Discussion.

In this section, I have introduced misfit reconstruction as a tool for highlighting areas of anomalous structure. However, as noted from the examples, it is questionable as to whether this technique is generally applicable. Because the misfits are expressed relative to flat lying (or parallel) events (by virtue of the way in which the covariance matrix is computed) any event which deviates from a flat lying, or parallel model, will contribute to the misfit. Hence data with many small 'kinks' in it will have a lot of background incoherent energy which will tend to produce 'misfits' everywhere. With such data, the misfit overlay may be of little use. However, in data where pinch-outs occur, we would expect to see a noticeable misfit in the overlay presentation.

This method of presentation, although in itself perhaps not diagnostic of hydrocarbon-trapping structures, may tend to draw the interpreter's eye quickly to zones which are not typical of the stacked seismic section. However, without the input of an experienced interpreter and specific knowledge of the data areas, I am unable to determine the significance or usefulness of these and other specific examples of misfit overlays. The usefulness of this particular technique would best be assessed in an industrial seismic interpretation department where access to many diverse data sets and their geological background are readily available.

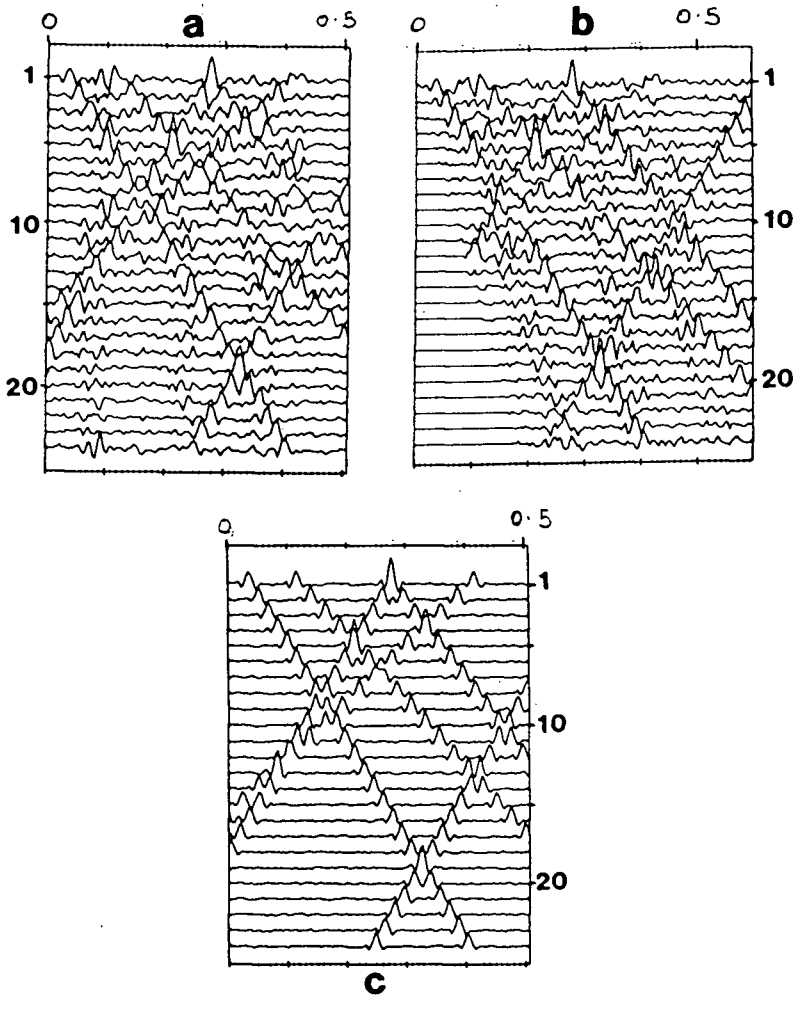


FIGURE 3.10
(a) A misfit reconstruction (principal components 6 - 24) of the data shown in Figure 3.2c, isolating the dipping noise events (c).
(b) A misfit reconstruction (principal components 6 - 24) of the data shown in Figure 3.4c, isolating the dipping noise events (c).

Again, because I have included some of the principal components which contain information from the horizontal structure, the general "noise" level has been increased.

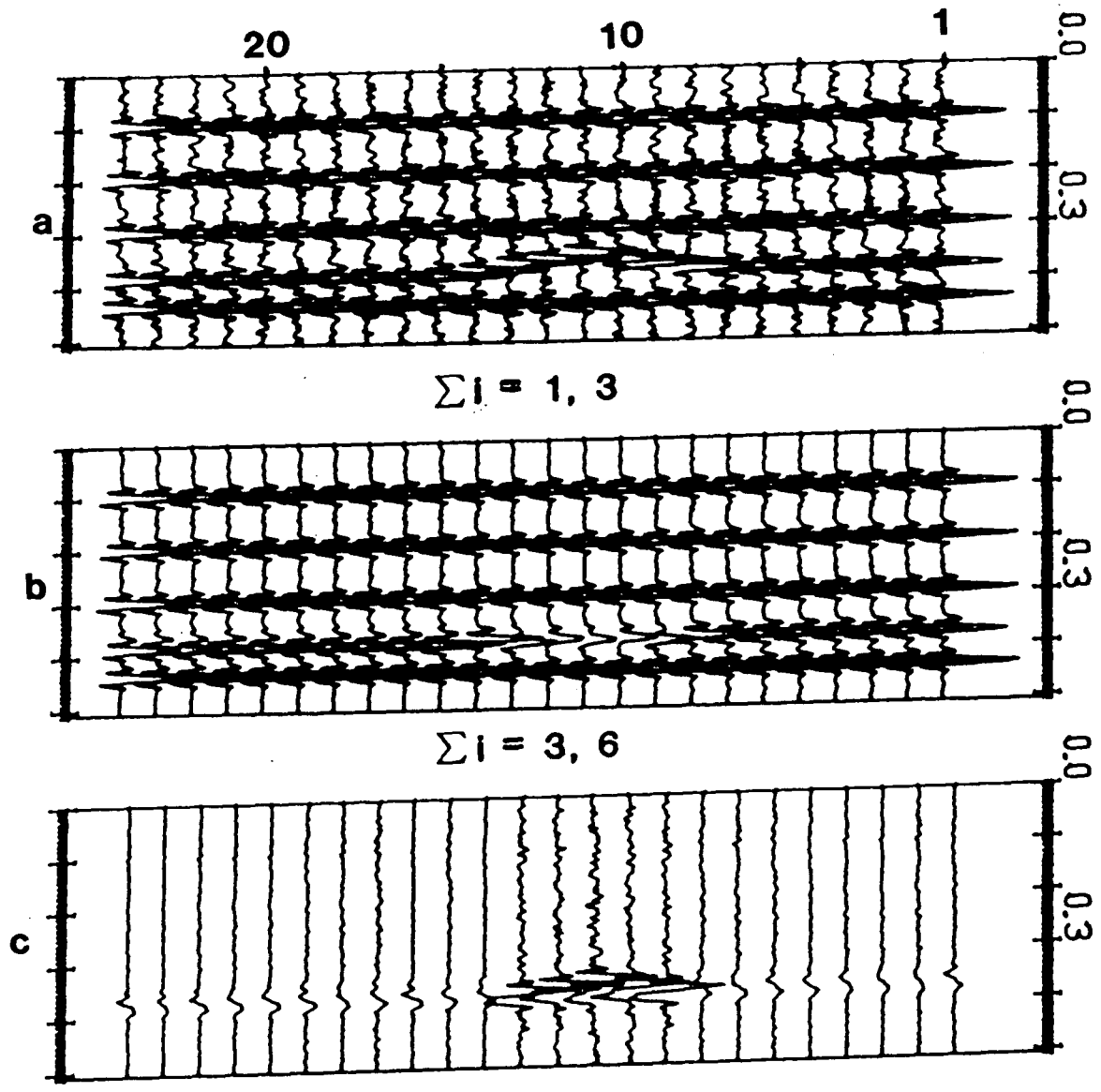


FIGURE 3.11
(a) A synthetic seismic section depicting flat horizons with a lens-like structure.
(b) The 94% reconstruction (requiring 3 of the 24 principal components) isolating the predominant flat-lying events.
(c) The misfit reconstruction (principal components 3 - 24) highlighting the lens-like feature; renormalized for plotting.

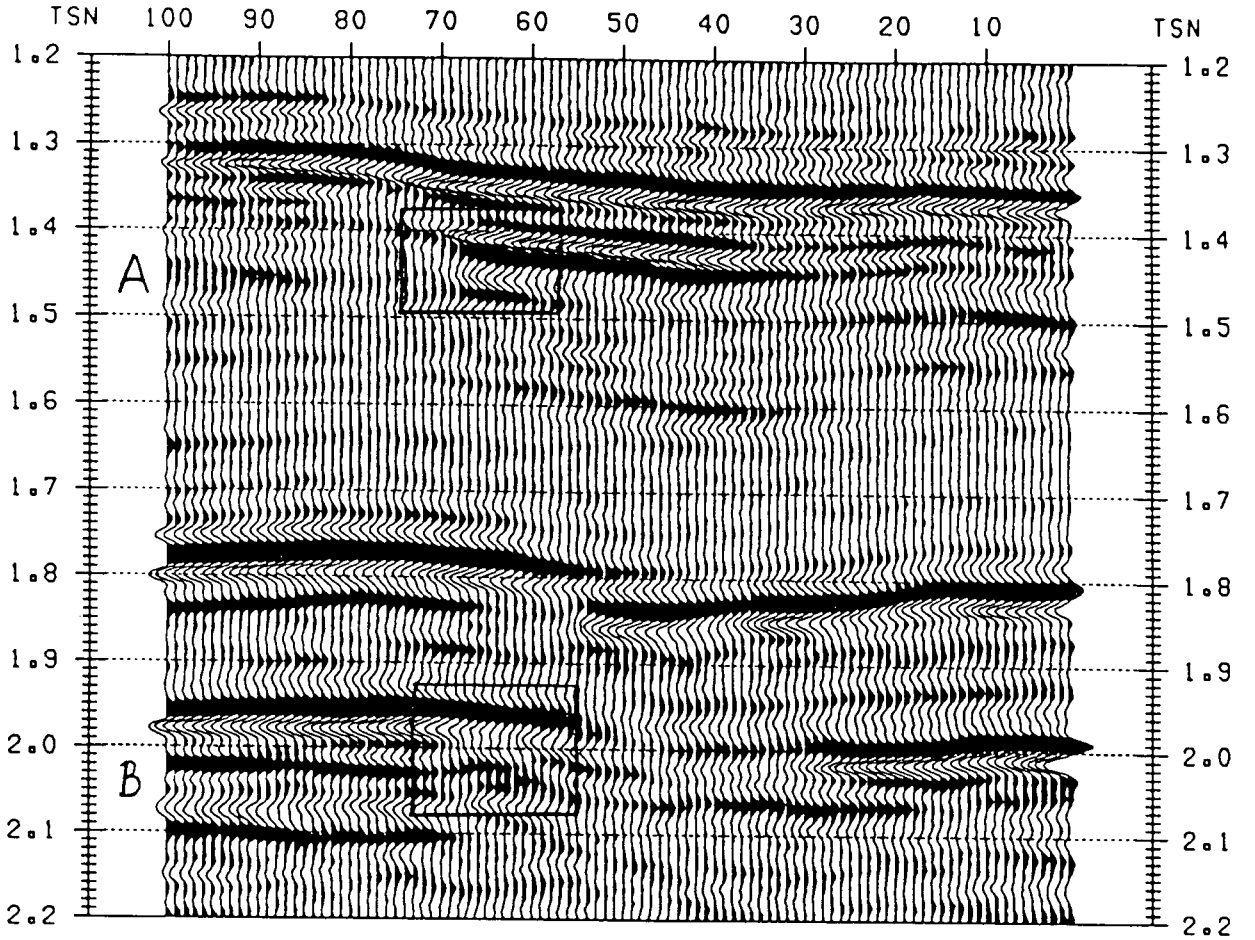


FIGURE 3.12a

Input data from a faulted region. Note the large negative anomaly which terminates to the left at the fault near trace 68 (box A), and the discontinuous horizon in box B, which terminates to the right at the fault.

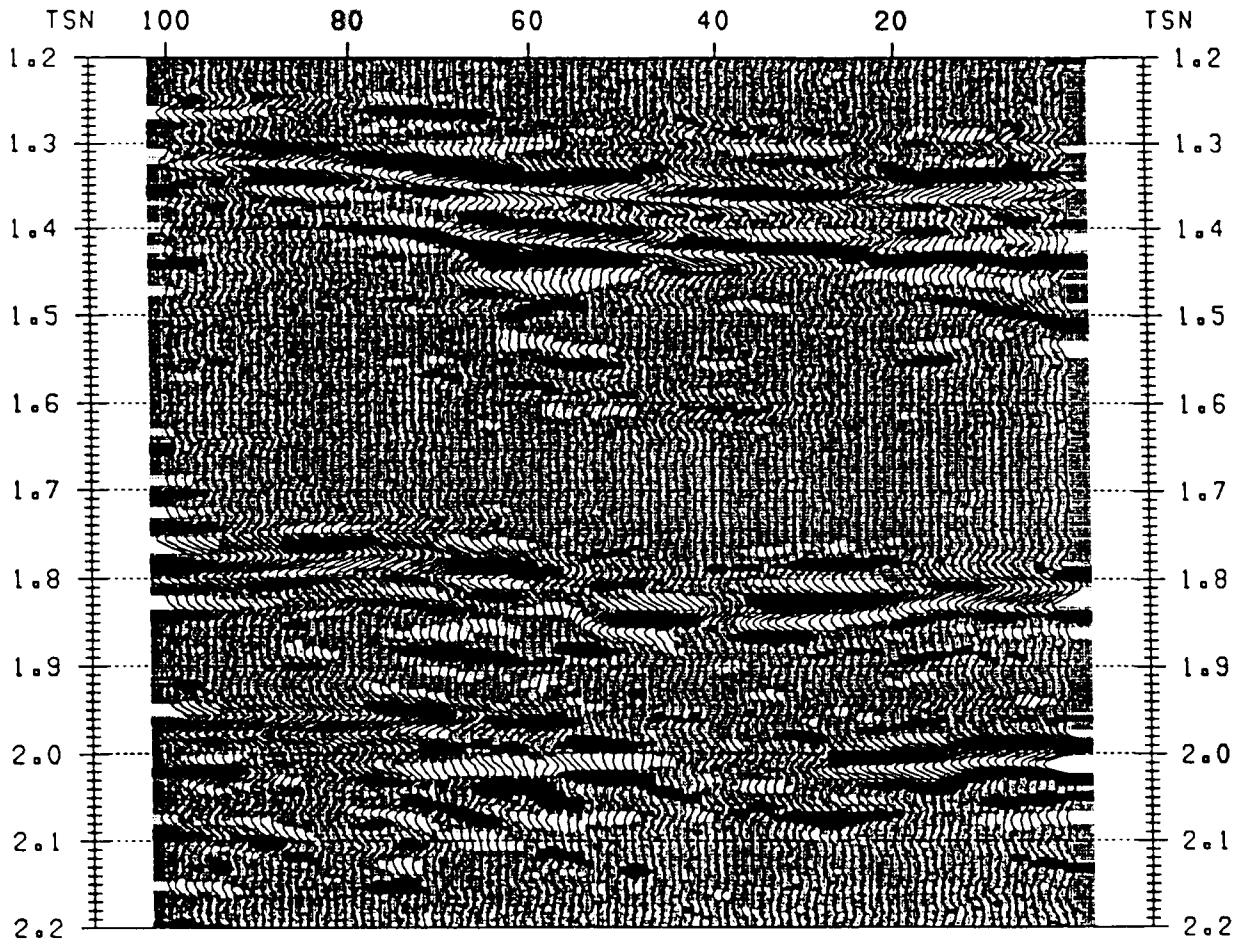


FIGURE 3.12b

Misfit overlay (omitting the first 70% of the energy; i.e. the first four principal components). Note the strong black-white contrasts which highlight the events terminating near the fault (see boxes A and B).

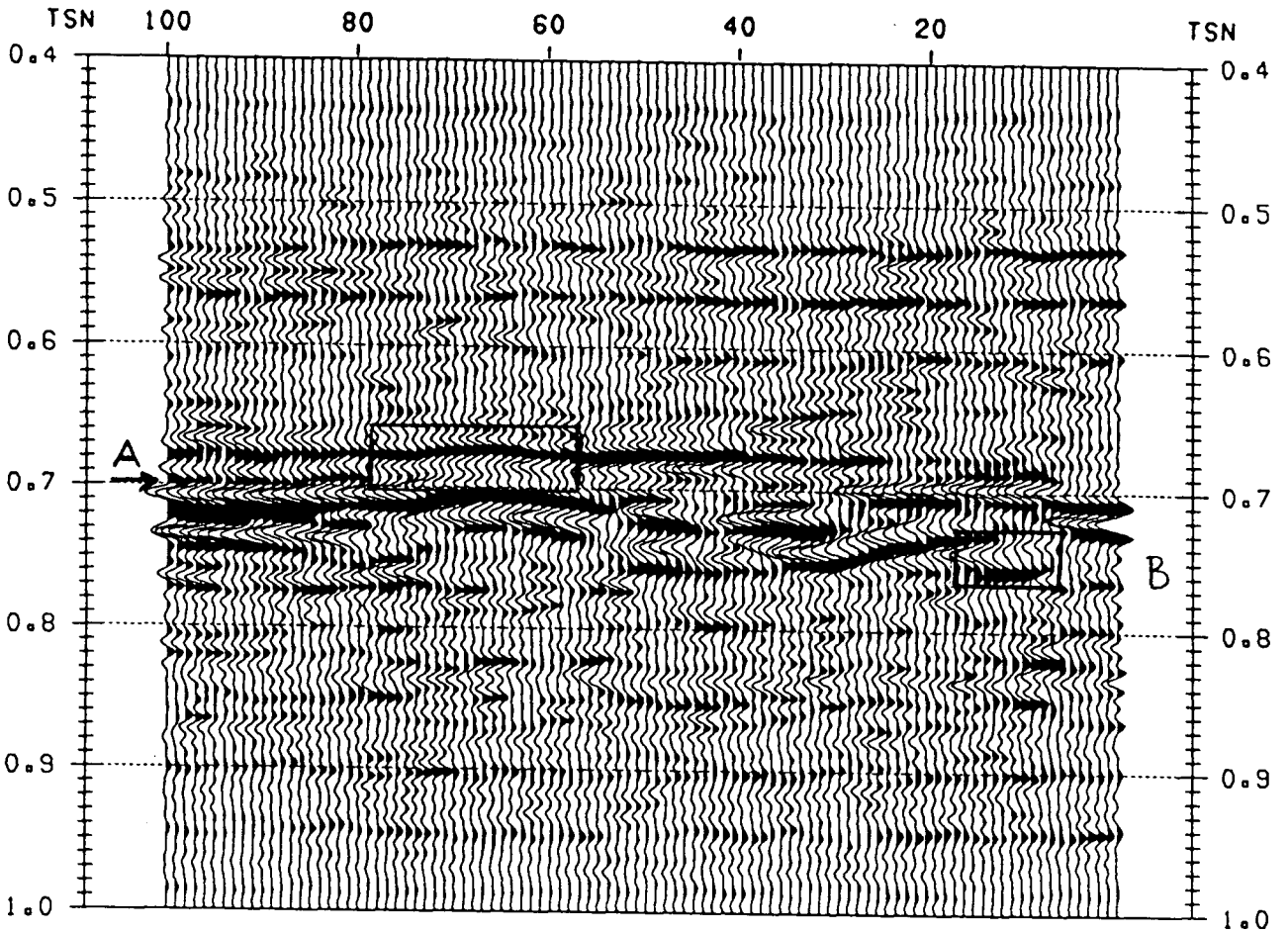


FIGURE 3.13a
Data from an area of braided streams. Note the discontinuity in the event in box A, denoted by the arrow, and the hump between traces 1 and 20 at 0.73 s in box B.

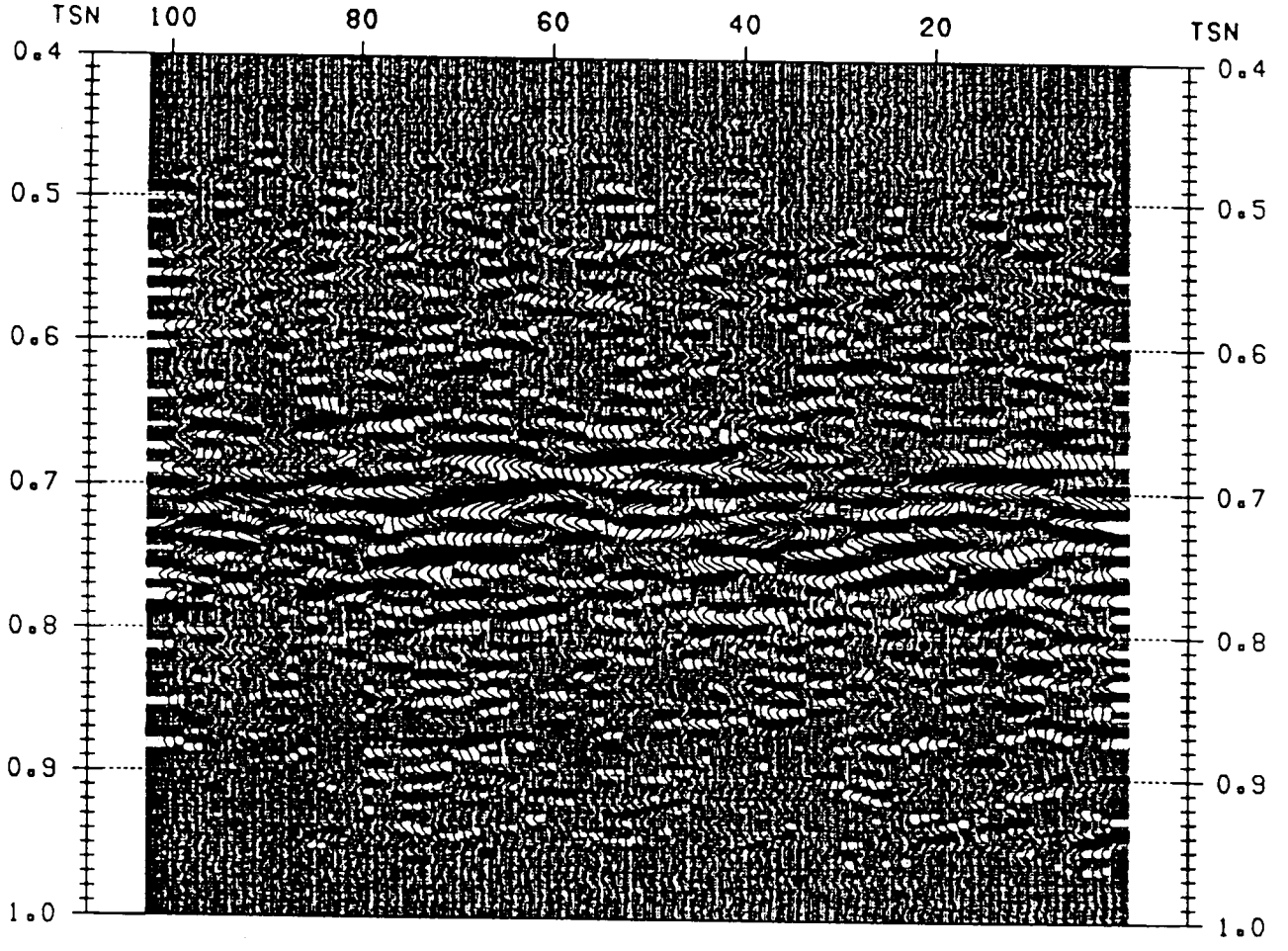


FIGURE 3.13b

Misfit overlay (omitting the first 60% of the energy; i.e. the first four principal components). Note the strong black-white contrasts which highlight the events in the boxes. In box A, we see the negative trough which replaces the thin event marked with the arrow, while in box B we see a black-white pair highlighting the apex of a hump.

SECTION 3.III: SEPARATION OF DIFFRACTIONS.

a. Introduction.

As an offshoot of the misfit reconstruction procedure, I noted that in areas of severe faulting or with numerous 'point' diffractors, diffraction events could readily be separated from the flat or parallel lying horizons.

It is difficult to perform residual migration or to estimate velocities from diffraction hyperbolae without isolating those diffraction events from the reflection events. Levin et al. (1983) employed a slant-stack (τ-p) filtering procedure to eliminate the reflection events, leaving only the diffraction events and noise. I propose here to utilize the misfit reconstruction to achieve this goal.

b. Synthetic Data Examples.

1. For a model consisting of a single horizontal sheet terminating at trace 10, buried in a medium with velocity 5 km/s, 21 traces separated by 50 m, each of 1.0 second duration were produced using a Fresnel integral synthetic seismogram program (M. Yedlin, pers. comm., 1984). The source function was a 16 Hz Ricker wavelet sampled at 4 ms. The seismic response of this model yielded a series of reflection events, and a hyperbolic diffraction arrival centered about the sheet edge (Figure 3.14b). This synthetic example is particularly problematical, as to a large extent the right hand of the diffraction event is buried in

the reflection.

The moveout was removed from the data to flatten the reflection event, (Figure 3.14c), and an isolated diffraction event (produced separately for comparison) is shown after the same moveout correction in Figure 3.14d. In Figures 3.15a and b, we see the 98% reconstruction and the misfit reconstruction for the real KL decomposition. We still see residual waveforms where the flat lying event was; this is due to the fact that the waveform in the reflection event changed slightly across the section and could not totally be reproduced as a linear combination of the first two principal components. However, the amplitudes of these residual events is only comparable to the diffraction event energy (i.e. about 10% of the amplitude of the input reflection waveforms). On the other hand, the diffraction event has been faithfully reproduced both in phase and amplitude, except where the diffraction and reflection events ran together, as occurs between traces 11 and 21. In these regions the two waveforms are not separable, as they are essentially scaled versions of one another.

2. The second synthetic example (M. Yedlin, pers. comm., 1984) shows a zero offset seismic section over the same truncated buried horizontal sheet (Figure 3.14a). This results in a series of reflection events which terminate over the sheet edge, and a set of diffraction arrivals centred about the termination point (Figure 3.16a). The diffraction event itself, produced separately for comparison and normalized, is shown in Figure 3.16b. Figures

3.17a and b show the 97% KL reconstruction and misfit reconstruction. As is seen, the misfit reconstruction is well able to separate out the diffraction event whilst preserving the phase and amplitude information. Some residual features remain in the vicinity of the section where the horizontal reflection feature was, as some contribution from that feature exists on principal components other than the first.

c. Migration of Diffraction Sections.

For the model discussed earlier, with a single sheet buried in a medium of constant velocity, I created a zero-offset section with trace separation 12.5 m and velocity 4 km/s (Figure 3.18a). The diffraction event, produced separately for comparison with subsequent results, is shown in Figure 3.18b.

Using the frequency-wavenumber migration algorithm of Stolt (1978), I migrated the diffraction section (Figures 3.18c) derived from Figure 3.18a. In this case, as the diffraction curvature is less that in Figure 3.16b, there is more distortion of the waveforms along the hyperbola, as more energy is segregated on to the first principal component. The best result, in terms of collapsing the event, was achieved with the correct velocity of 4 km/s. The events are seen after migration in Figure 3.18d.

In the case of incorrectly migrated data, velocity thus inferred could be used for residual migration of an original migrated stacked section showing residual diffraction events

(Levin et al., 1983, Harlan et al., 1983) or simply to ascertain the velocity of the materials about the diffractor. Further, in regions of severe faulting, this procedure may possibly serve as a fault-edge indicator.

d. Real Data Examples.

1. Figure 3.19a shows 96 traces of stacked data, which I was informed were time migrated. Events between traces 45 and 65 are less continuous, and at about 1.1 s we see the right hand limb of a diffraction hyperbola. Given that the section had been previously migrated, but still shows evidence of diffractions, we may conclude that the data were incorrectly migrated (probably due to an inadequate velocity function). This opens the question of the need for residual migration (Levin et al. 1983) and also points to the possibility of using such diffractions to estimate seismic velocity (Harlan et al. 1983). Figure 3.19b shows a 92% reconstruction which has eliminated all evidence of diffraction events leaving only the gross, common features of the data. However, producing a misfit reconstruction (excluding the first 96%) clearly shows a series of diffraction events centred throughout the zone between traces 65 and 75 (Figure 3.19c). Now that the diffraction events have been segregated from the flat lying events, we may perform a series of migrations on these data in order to collapse the hyperbolae at each depth. This was done with two specific hyperbolae in mind (indicated in Figure 3.19c). Event A was best collapsed with a velocity of 3 km/s (Figure

3.19d). Event B was best collapsed with a velocity of 4 km/s (Figure 3.19e). In both these figures, I show the window of data used for migration, and three migration results: i.e. the best visual result, and the result from using a velocity 1 km/s higher and lower than this. The overall appearance of the migrated difraction sections is that of migrated noise (Berkhout, 1984). This is not surprising, since the majority of the data in the diffraction section is noise. However, focussing one's attention on the zones where the diffractions were centred can allow selection of a suitable migration velocity. The fact that these velocities are high raises the posibility that these particular data were not in fact previously migrated.

e. Discussion.

Misfit reconstruction with a view to separating diffraction events offers a new way of isolating diffractions from stacked seismic data. The method of Harlan et al. (1983,1984) is considerably more protracted. Once the diffraction events are segregated, the residual migration technique proposed by Levin et al. (1983) may be used to estimate correction velocities for iterative residual migration of stacked seismic data.

As a corollary to residual migration in areas with severe faulting, such as the North Sea, a display of the diffractions may serve as a fault-edge indicator. This may be of use in helping to delineate fault controlled hydrocarbon trapping structures.

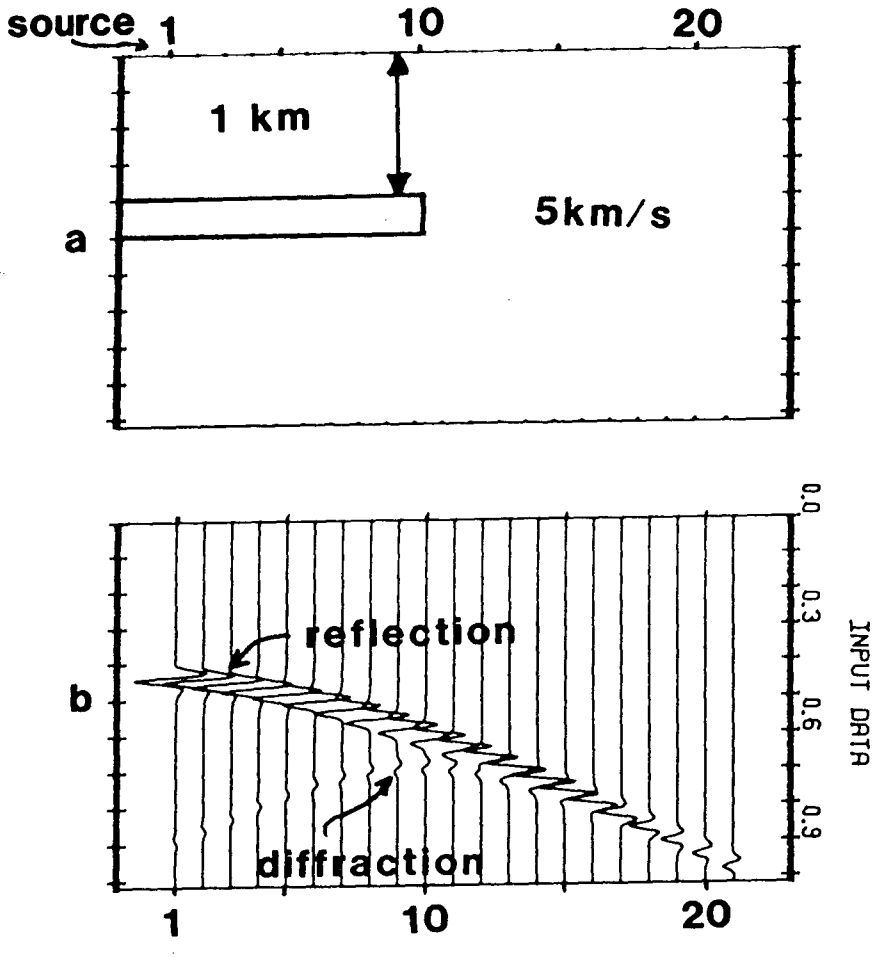


FIGURE 3.14a
(a). The model: a thin sheet buried at 1 km depth, terminating at trace 10
(b). The reflection and diffraction arrivals for a 16 Hz Ricker source wavelet, propagating in the model (a).

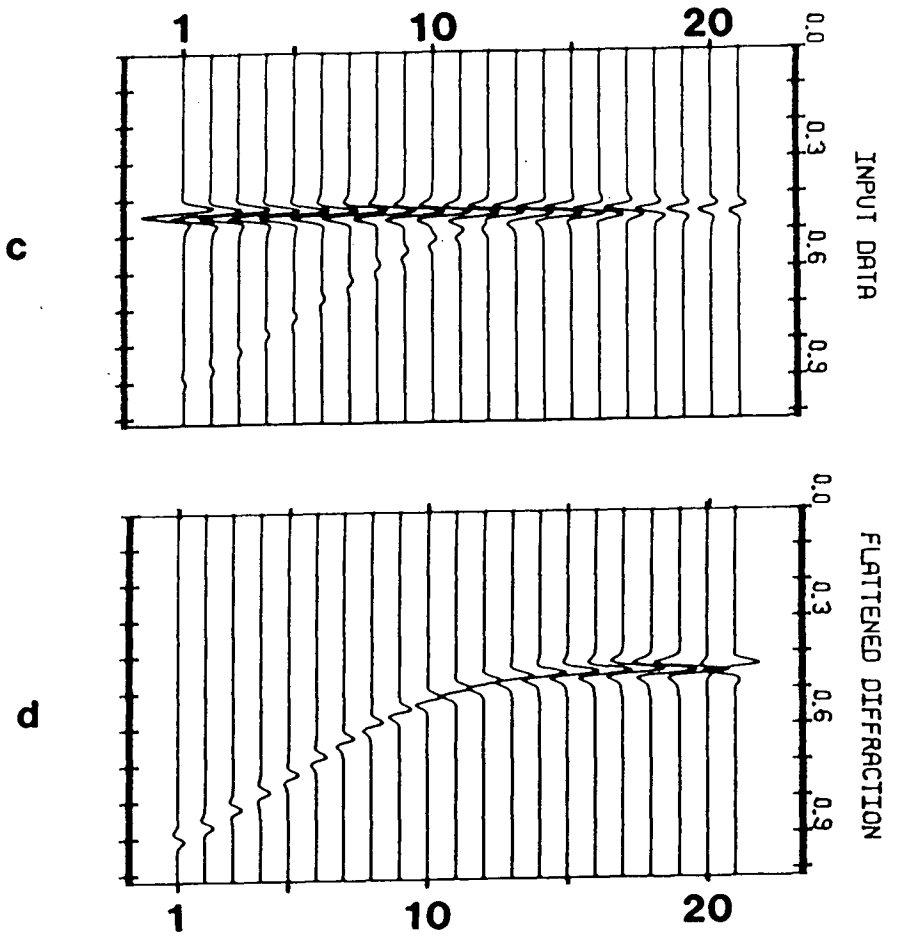


FIGURE 3.14c
(c) The data after flattening on the reflection event.
(d) The flattened diffraction event produced separately for comparison with later results, renormalized for plotting.

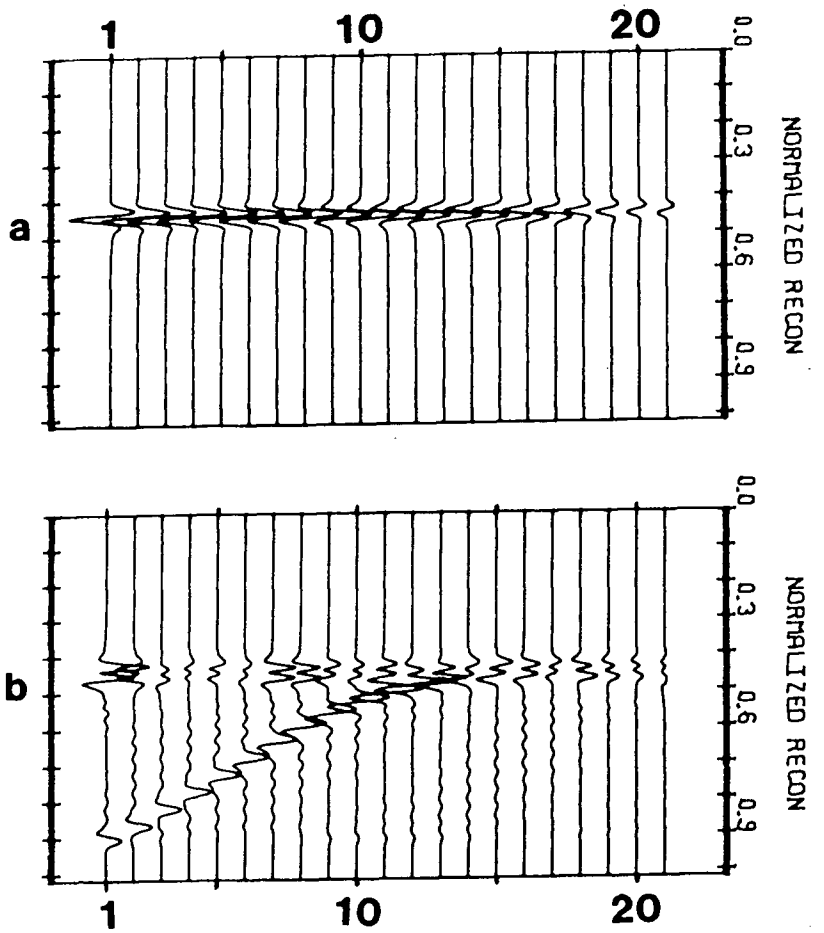


FIGURE 3.15

(a) A 98% reconstruction of the data in 14(b) (requiring 2 of the 21 principal components). To the eye this appears to have recovered the flattened reflection event.

(b) The misfit reconstruction of the data in 14(b) (principal components 3 - 21) renormalized for plotting. For the most part, the isolated diffraction limb is faithfully reproduced. However, where interference had distorted the original limb, the misfit reconstruction is subject to distortion. Where the diffraction runs into the reflection, they are inseparable, as the waveforms are the same.

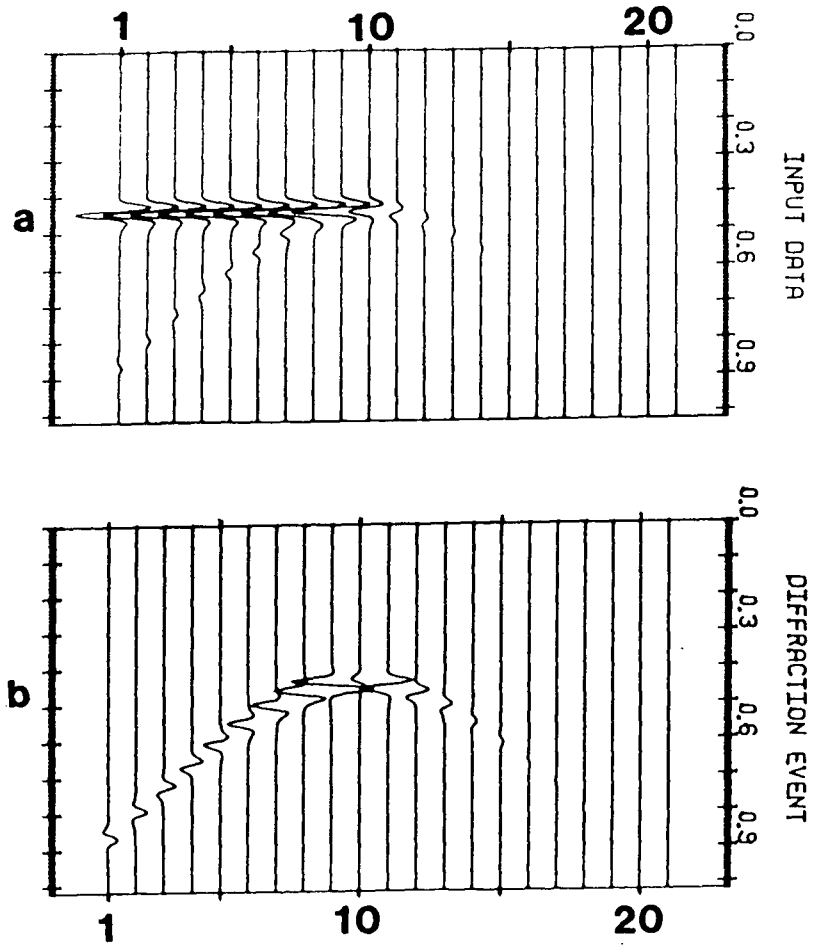


FIGURE 3.16
(a) A zero offset seismic section over a buried horizontal sheet (see sketch in Figure 3.14(a)) which terminates at trace 10, giving rise to reflection events, and diffraction events centred about trace 10.
(b) The diffraction events produced separately for comparison to subsequent results; renormalized for plotting.

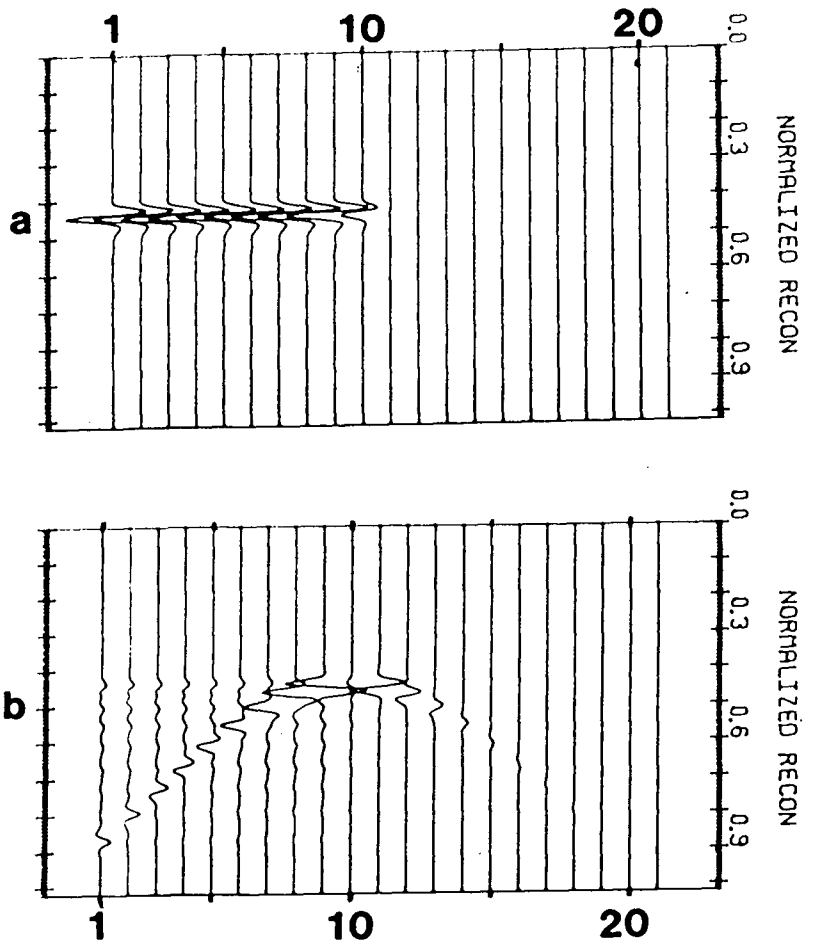


FIGURE 3.17

(a) The 97% reconstruction of the data in 16(a) (requiring 1 of the 21 principal components).
(b) The misfit reconstruction of the data in 16(a) (principal components 2 - 21) has successfully isolated the diffraction event. Phase and amplitude information are well conserved, as is seen in comparison with 16(b). Renormalized for plotting.

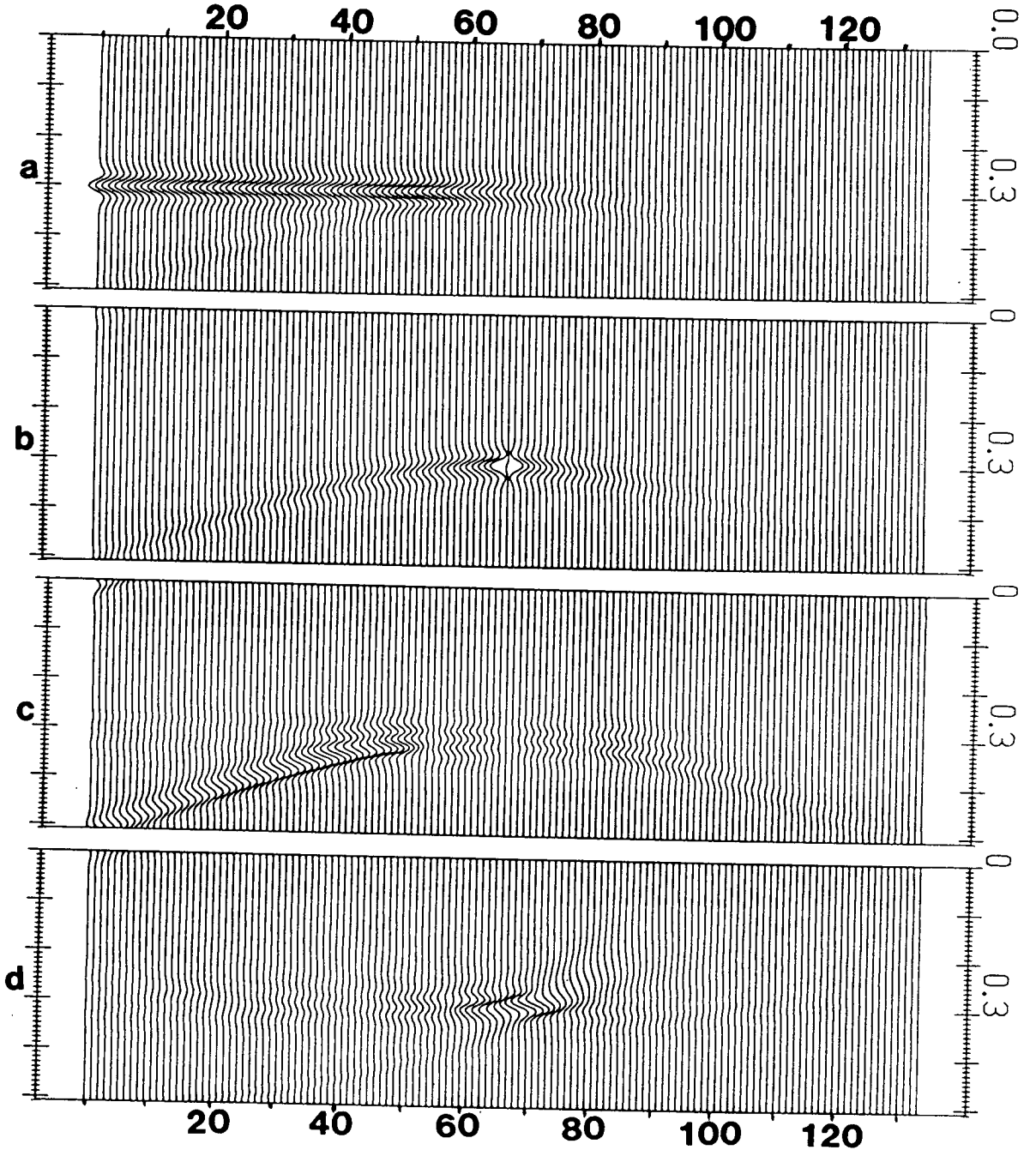


FIGURE 3.18
(a). The input section (see text). (b). The diffraction produced independently for comparison. (c). The misfit reconstruction using principal components 2-128. (d). The diffraction section, (c) above, after migration with velocity 4 km/s.

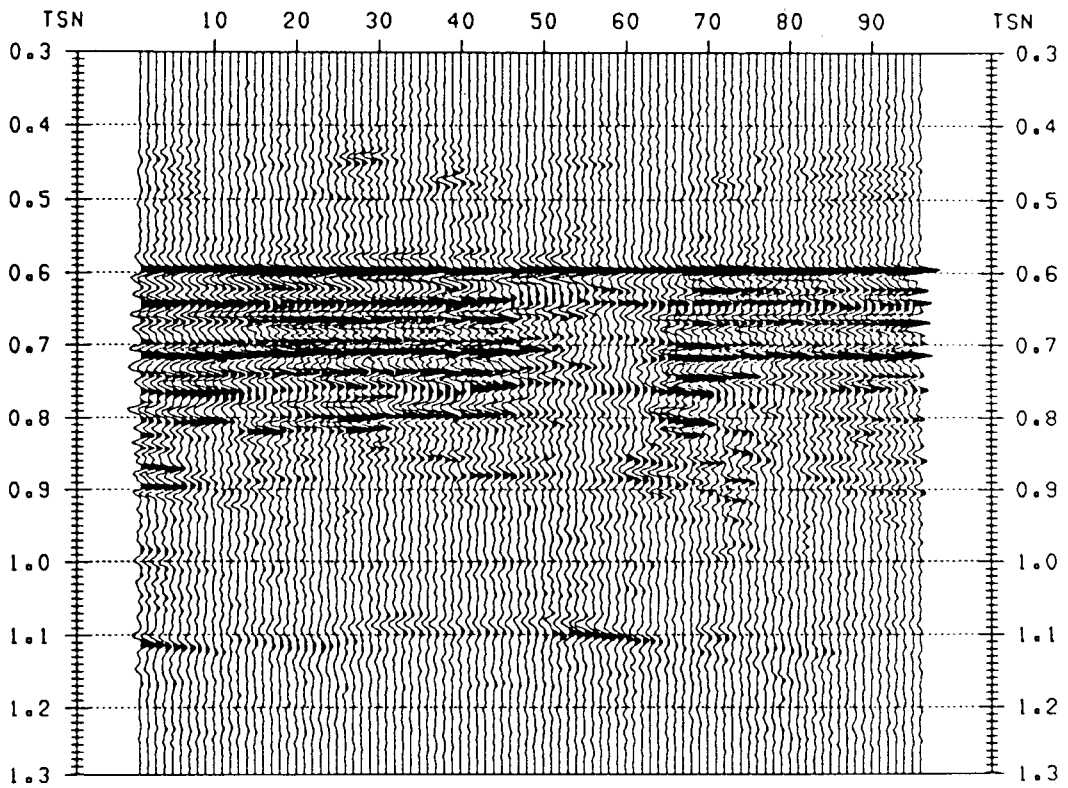


FIGURE 3.19a
Ninety six traces of stacked time migrated data in a region of faulting. Note the discontinuous events in the area from traces 45 - 65 between 0.6 and 0.9 s.

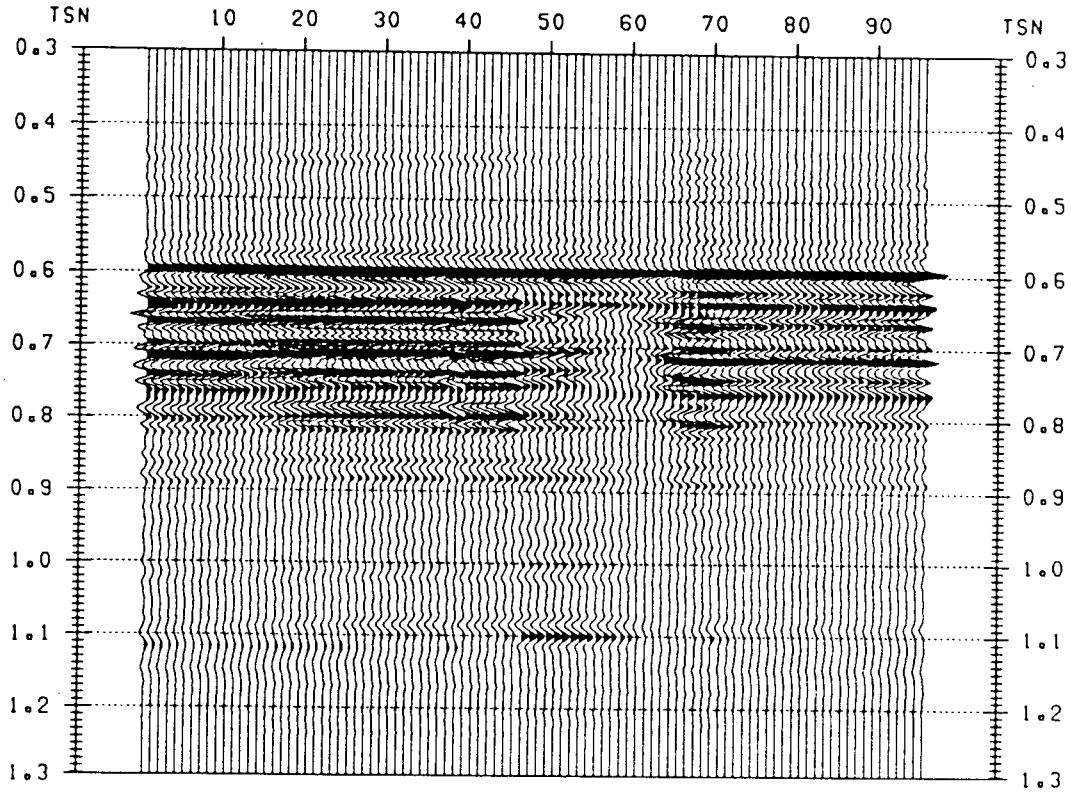


FIGURE 3.19b

The 92% reconstruction (requiring 3 of the 96 principal components) which highlights the gross structural features of the section. Note the loss of continuity of the events between traces 45 - 65.

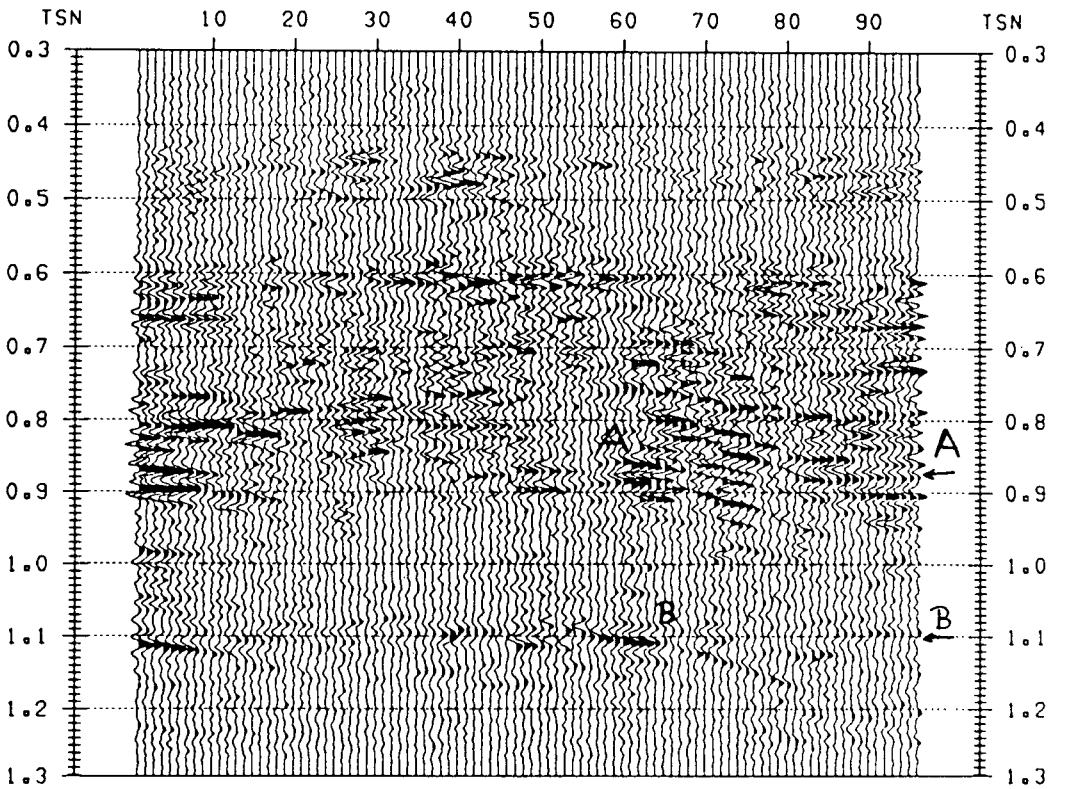


FIGURE 3.19c

The misfit reconstruction (principal components 6 to 96, excluding the first 96%) showing a series of diffraction events in the region of interest. Hyperbolic events "A" and "B" are the target events for migration. The data have been renormalized for plotting.

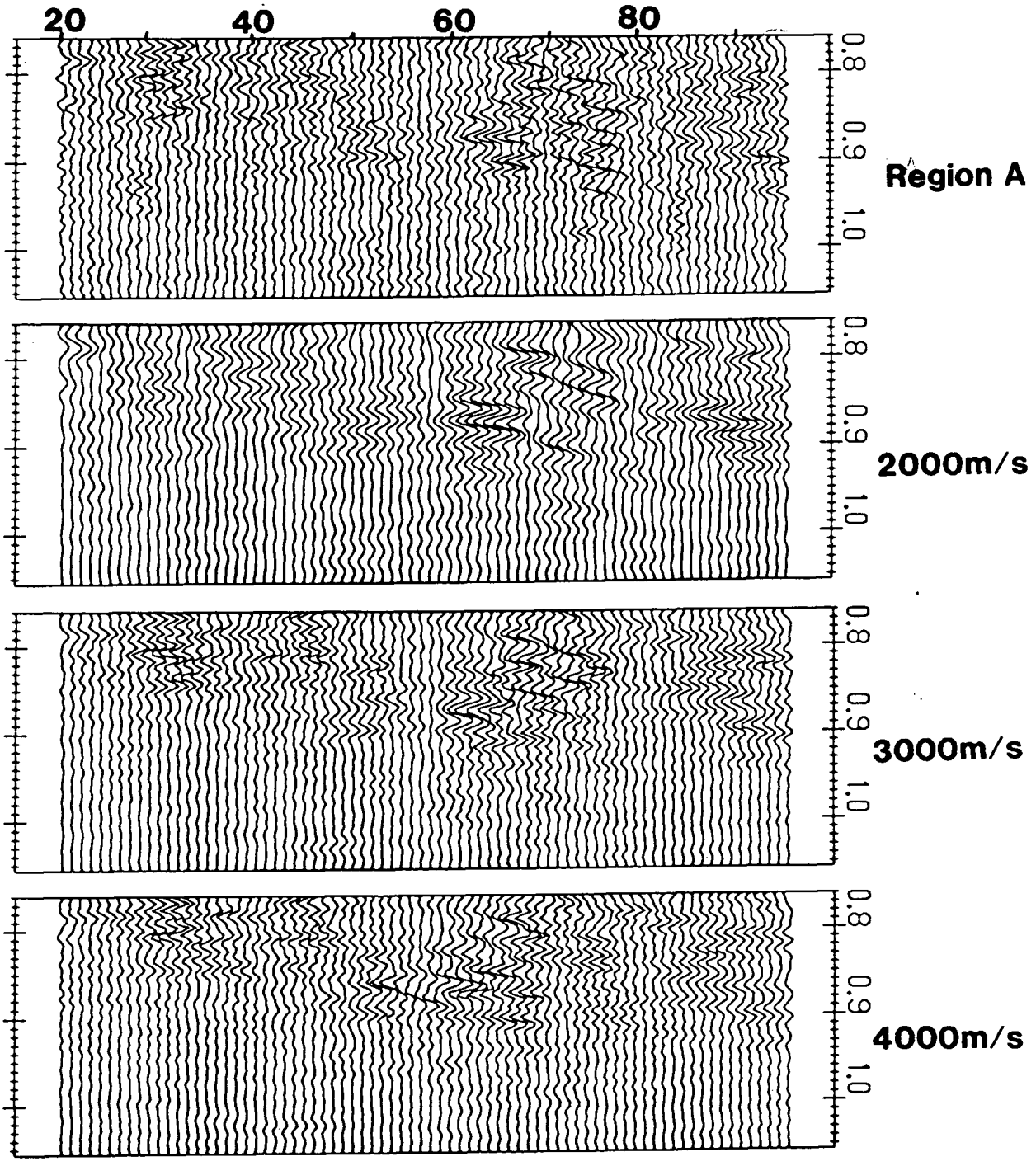


FIGURE 3.19d

Region A of the misfit reconstruction (Figure 3.19c). Migrations with velocity: 2000, 3000, and 4000 m/s. The hyperbolic event "A" has been flattened best using 3000 m/s. The data have been renormalized for plotting.

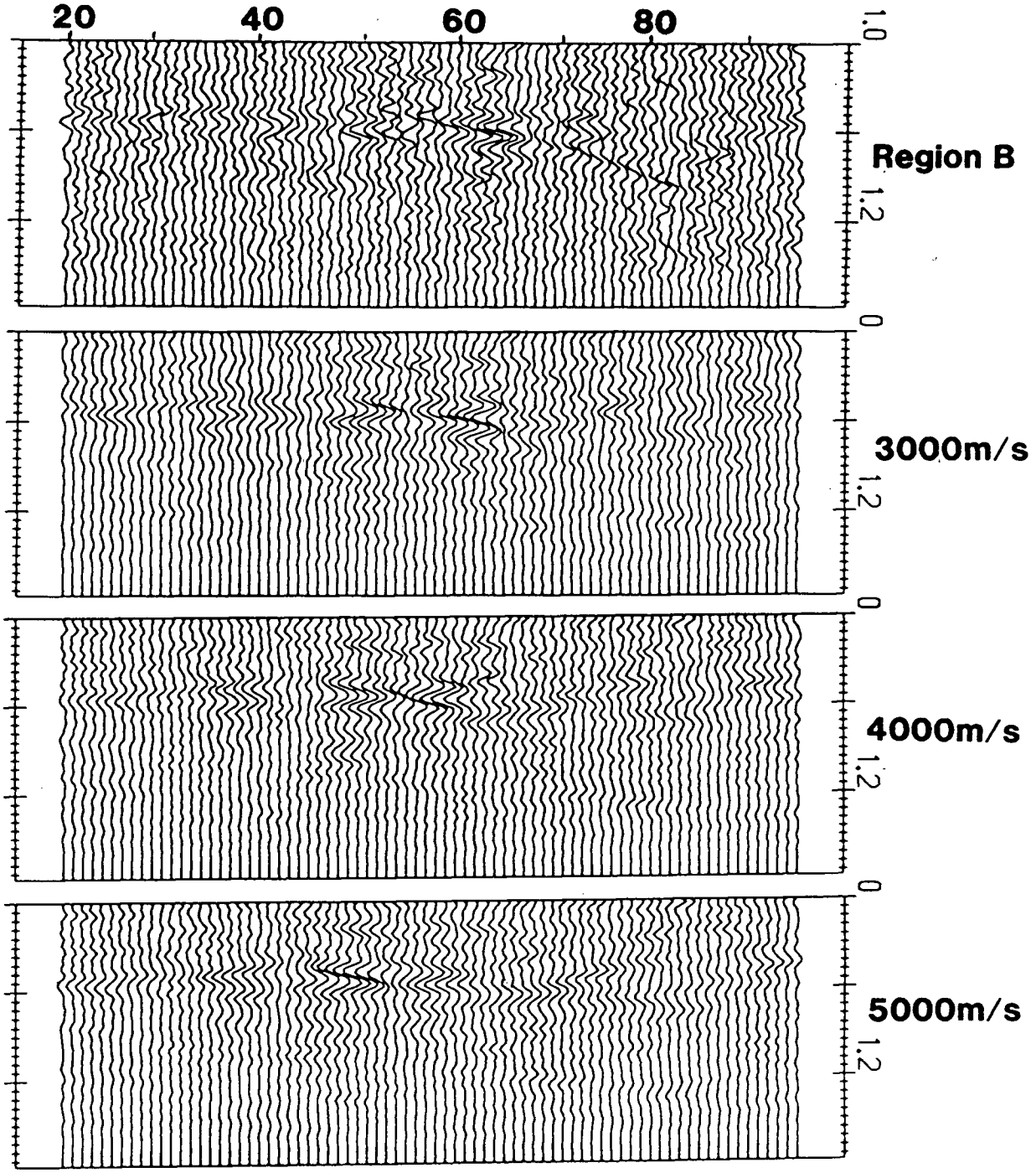


FIGURE 3.19e
Region B of the misfit reconstruction (Figure 3.19c). Migrations with velocity: 3000, 4000, and 5000 m/s. The hyperbolic event "B" has been flattened best using 4000 m/s. The data have been renormalized for plotting.

SECTION 3.IV: MULTIPLE SUPPRESSION.

a. Introduction.

For the purpose of multiple suppression, I utilize the energy packing property of the real KL transform. The idea here is to segregate the energy associated with the multiples onto a single principal component. A data reconstruction omitting that principal component should be essentially multiple-free.

I proceed in five basic steps:

i. From the standard velocity analysis, identify the velocity and onset time for a multiple.

ii. Using the RMS velocity associated with the multiple, perform a constant velocity moveout correction on the seismic data. We note at this stage that the arrivals due to multiples will have been more or less flattened, whereas the primary events are under or over corrected, and will have increased curvature in the section.

iii. Compute the KL transform of the constant velocity moveout corrected data set. Correlated energy in the seismogram which had a moveout velocity equal to that of the multiple's velocity will now appear predominantly on the first principal component. This is because coherent energy having the velocity associated with the multiple has been flattened, or aligned across the section, and will now appear to be the most highly

correlated energy.

iv. Reconstruct the constant velocity moveout corrected seismogram from the principal components but omit the first principal component, i.e. leave out any correlated energy associated with the velocity of the multiple. It may also be advantageous to omit the second principal component when the multiple arrival's waveform has been severely distorted by interfering primary events.

v. Remove the moveout stretching from the reconstructed data using the same velocity as in step ii.

The procedure is repeated for all multiples for which suppression is desired, and then the usual NMO correction to align the data prior to stacking is applied. Ryu (1982) addressed the problem of multiple suppression using the FK transformation in conjunction with data stretching via moveout correction. However, the approach described here is a new and alternative one, and will not be as prone to the aliasing problems encountered with the FK method when dealing with a small number of traces (D. Hampson, pers. comm., 1985).

b. Synthetic Data Examples.

1. In Figure 3.20a, I show a simple synthetic seismic section representing reflection events from 9 flat layers over a half space, all overlain by water. Included are two events due to 'multiple' travel paths in the surficial water layer (indicated

by arrows). Figure 3.20b shows the semblance velocity analysis (VA) (Neidell and Taner, 1971) of this data. We see 'multiple' energy at 1.1 and 1.65 seconds with a characteristic velocity of 1450 m/s. Figure 3.21a shows the data after a constant velocity NMO correction. Note how this procedure has flattened the multiple events at 1.1 and 1.65 seconds. Starting just after the primary event at 1.0 s (Figure 3.21a), a real KL decomposition of the section is carried out. Reconstruction omitting the first and second principal components gives the results in Figure 3.21b. The effect of moveout stretching is then removed from the data using the velocity 1450 m/s (Figure 3.22a). Figure 3.22b shows a VA of the data after multiple suppression processing. We note by comparison with Figure 3.20 that the multiples have been effectively removed from the data, and that the VA shows only the primary bottom reflection with the water velocity. Figure 3.23a shows the multiple suppressed data after a conventional NMO prior to stack. In Figure 3.23b, we see a comparison of the trace resulting from: (1) the stack of synthetic data which never had the multiple arrivals; (2) the conventional stack of the data with multiples; and (3) the conventional stack of the data after multiple suppression using the KL technique.

A comparison of Figures 3.23b (2) and (3) shows that the multiples at 1.1 and 1.65 s have been successfully suppressed. The noise introduced to the section after removal of the first two principal components (Figure 3.21b) did not stack

constructively.

With a view to automation, the algorithm was modified to automatically detect the multiples, i.e. their RMS velocity and onset time. To this end, a stack of overlapping time segments of the VA map was made. In Figure 3.20b we saw that multiple energy is reasonably vertical along a locus of constant velocity, whereas the stacking velocity trajectory is not. Consequently, we expect a stack of time segments of the VA map to display maxima at the location of the velocity of each multiple. These maxima are detected, and the processing is started for a given multiple just after the onset of that multiple so as not to suppress the primary. Figure 3.20c shows a stack of the first half of Figure 3.20b. We see a peak at 1450 m/s which corresponds to the observed trend in the VA map.

In Figure 3.20c a maximum appears at the velocity of the multiple (1450 m/s). However, the stack also shows maxima at 1700 m/s corresponding to the primary events between 0.85 and 1.10 s, and at 2000 m/s corresponding to the primary event at 1.28 s. Care must be taken so as not to allow the program to suppress primary events on the basis of these maxima. On the whole, the automatic mode must be treated cautiously, but works well if we only wish to suppress multiples with one characteristic velocity: usually water bottom multiples.

c. Real Data Examples.

Of greater interest is the performance of the algorithm on real seismic data when multiples pose a problem.

1. Figure 3.24 shows a CDP gather of marine seismic data shot off Sable Island, Canada. Each gather has 60 traces sampled at 4 ms, but to reduce expense, only every third trace was used here. Figure 3.25a shows a broad velocity band VA of the gather. We see a dominant trend of multiple energy at 1620 m/s starting at 0.5 seconds. Figure 3.25b shows a stack of the VA map, highlighting the presence of the multiples: this information is used to detect the multiples for the program's automatic multiple suppression mode. Figure 3.25c shows the VA of Figure 3.24 after multiple suppression using an automatically picked velocity (1620 m/s). Note the absence of the band of multiple energy between 0.5 and 3.4 s. Note also how the event at 2.2 s and 2400 m/s has been enhanced in the multiple suppressed VA, as it is no longer obscured by arrivals due to multiple travel paths.

Twenty gathers of data were processed in this way, and a comparison of the conventionally stacked data with a stack of the multiple suppressed data is made in Figure 3.26a and b. The most noticeable features which differ between the multiple suppressed stack and the conventional stack are: the absence of the event at 0.91 s (marked 'A') in the multiple suppressed section, which

appears as part of a 'doublet' in the unprocessed data, and the changes at 1.35 s (marked 'B'). Several events are clearer in the multiple suppressed section, such as that at 0.86 s ('C'), the trough at 1.10 s ('D'), and the pair at 2.45 and 2.55 s ('E'). Subtracting these two sections yields a difference section (Figure 3.26c) which emphasizes the location and nature of the differences between the processed and unprocessed data. Bands of multiple related energy are seen with 0.44 s spacing.

d. Discussion.

Multiple arrivals often pose a serious problem in shallow water and certain land environments. Simple and effective methods for the elimination of multiple events of known stacking velocity are of use in alleviating these problems. Here I present an intuitively simple method for isolating and removing multiple events prior to final moveout correction and stack. On simple layered earth model synthetic data, the method has worked well. On the small sample of real data analysed, the VA maps highlight the almost complete absence of multiple related energy after processing. In addition, primary events which were of small amplitude in the VA map before processing were greatly enhanced, because the actual data hyperbola corresponding to the primary event was no longer masked by multiple energy arriving at the same time.

The results on the basis of 'before and after' VA maps look very promising, and initial tests on twenty gathers show several noticeable differences, their VA maps showing that a significant number of multiple events have been removed. In addition, the actual waveforms are essentially uncorrupted by the processing: an important point to note.

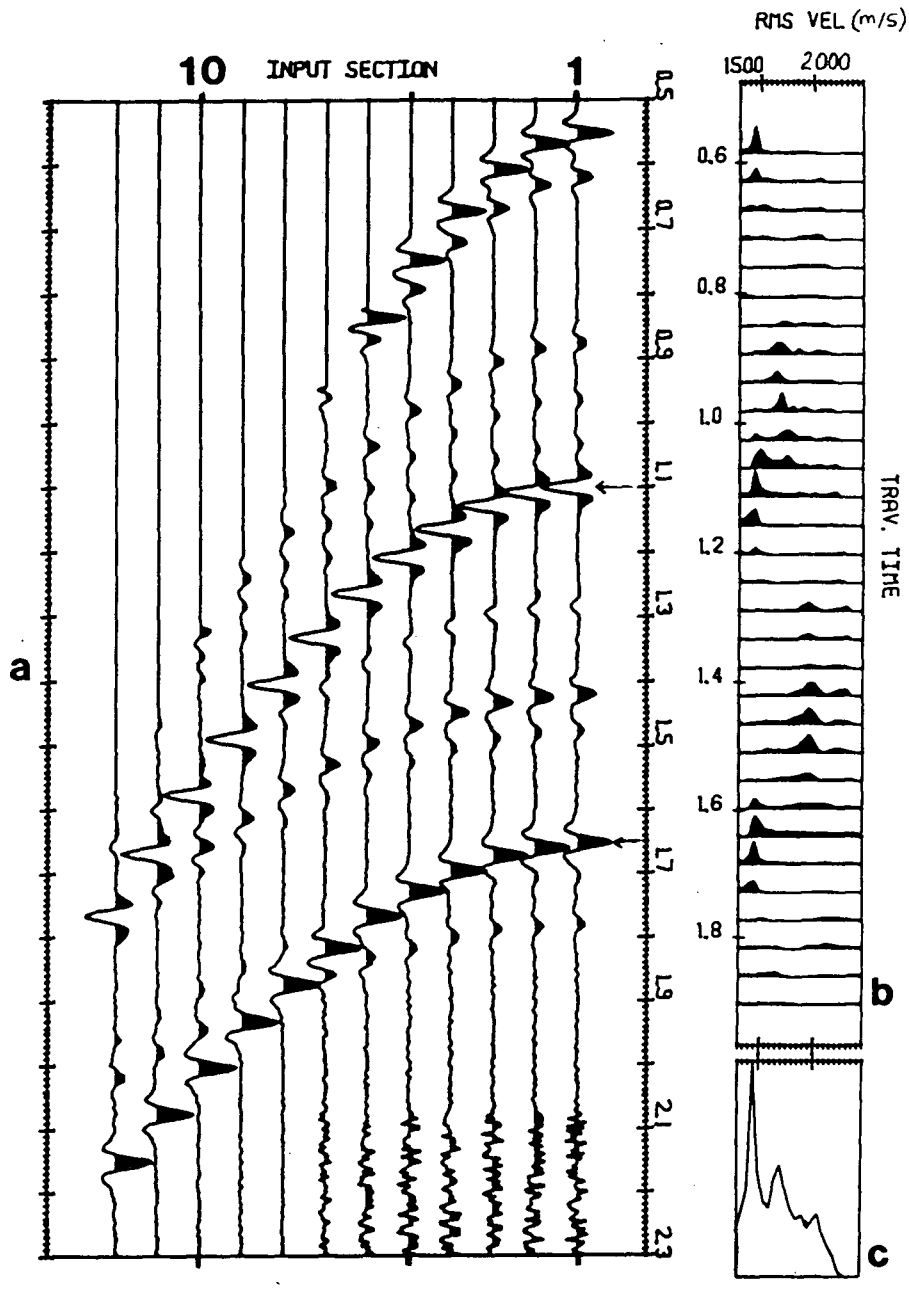


FIGURE 3.20

(a) The seismic representation of reflections from 9 layers overlying a half space, all overlain by water. The data have been muted and AGC'd. Two water bottom multiple events can be seen: the first, polarity reversed at 1.1 s, the second at 1.65 s.
(b) The velocity analysis (VA) of the data. Note the peaks due to the multiples with a velocity of 1450 m/s which dominate the lower portion of the VA.
(c) Stack of the first half of the VA map (b). As multiple events tend to lie vertically in the VA map, they will stack to produce a maximum, whereas comparatively, primary events will not.

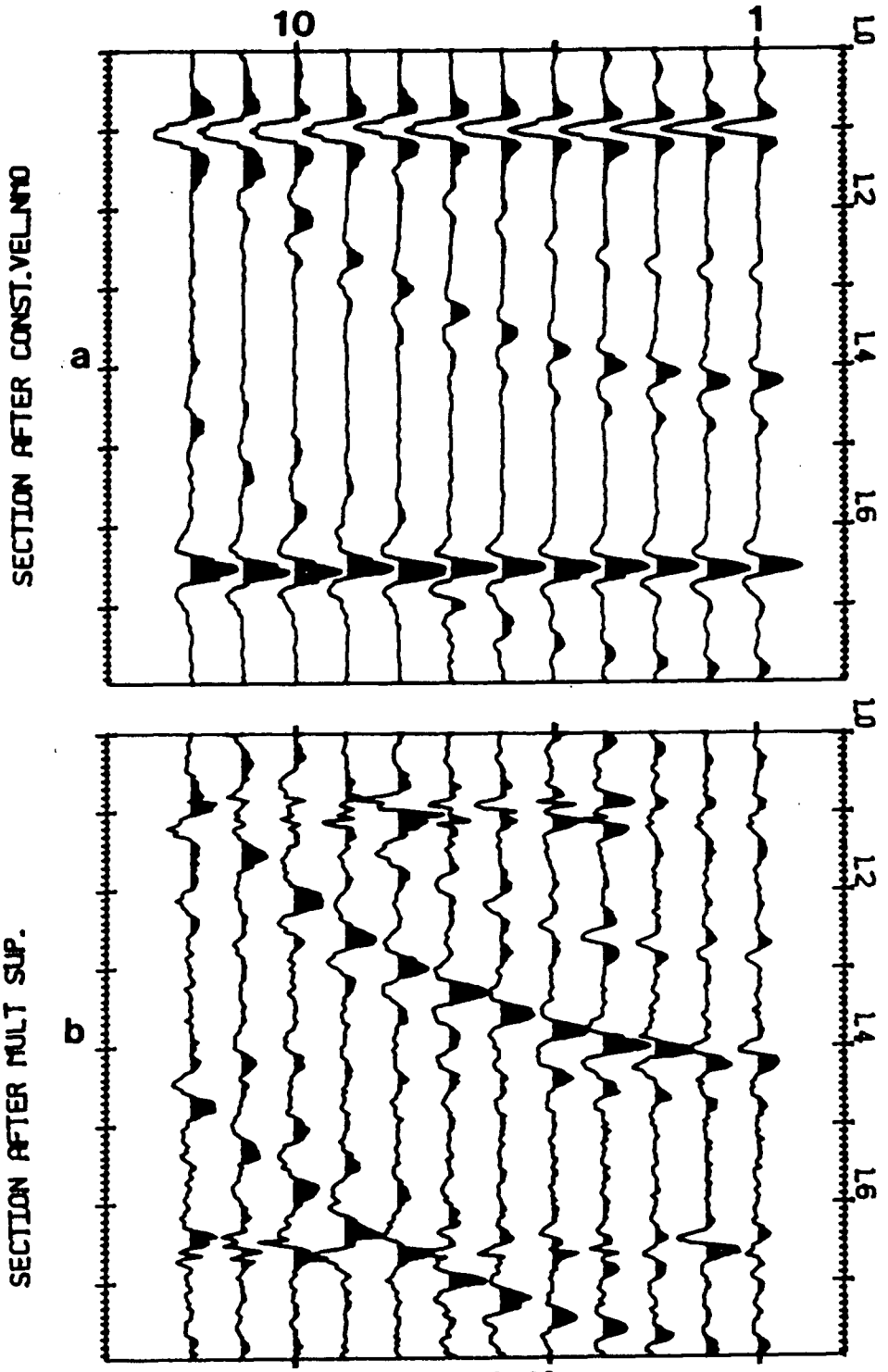


FIGURE 3.21

(a) A segment of the data after move-out correction with a constant velocity of 1450 m/s. Notice how the two events due to multiple arrivals have been flattened.
(b) A misfit reconstruction of the flattened data omitting the first 2 principal components. Some residual noise remains in the locations formerly occupied by the multiple events, but this is largely incoherent, and will not stack to produce a noticeable effect.

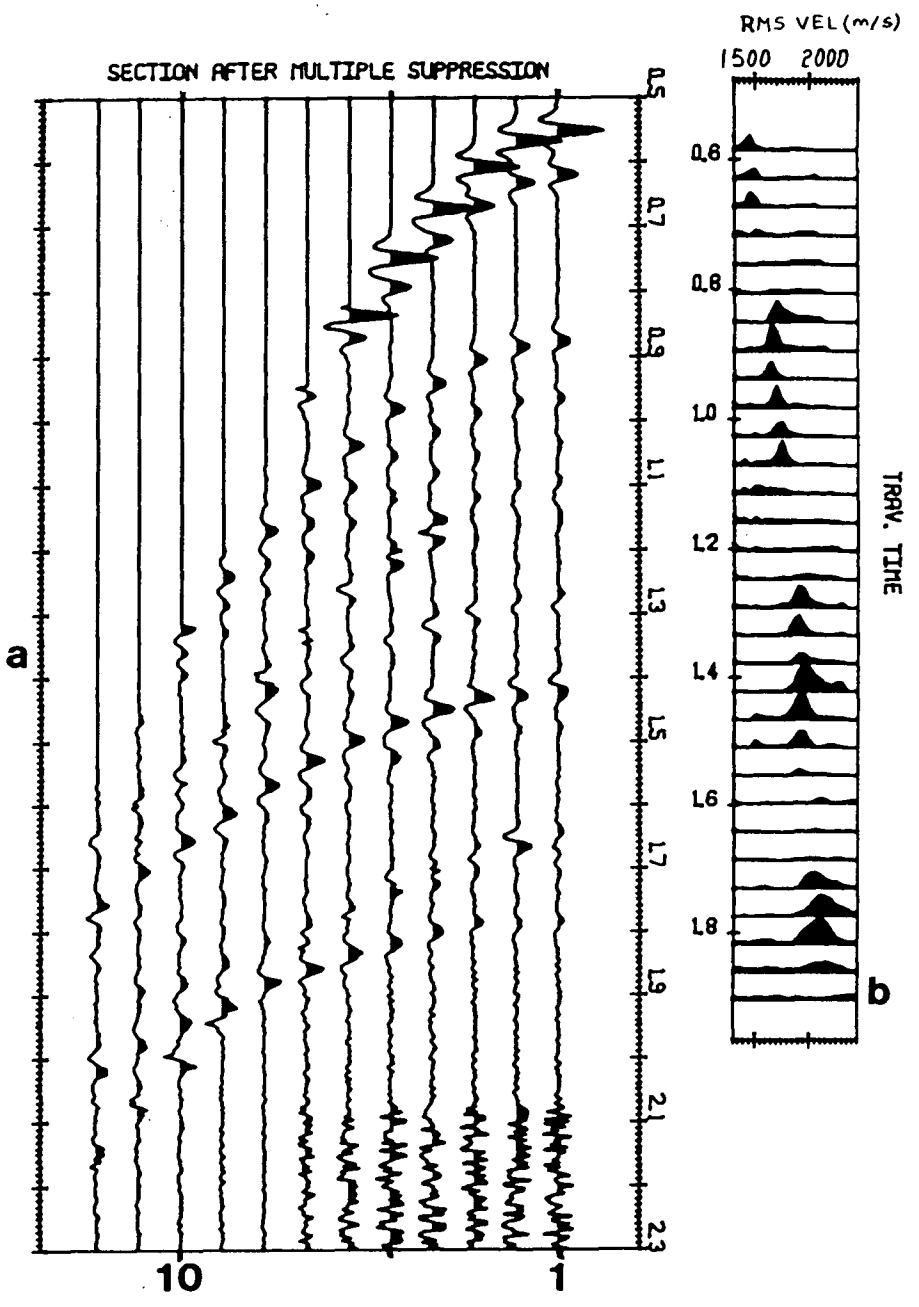


FIGURE 3.22

(a) The reconstructed data, after the effects of the constant velocity moveout correction have been removed, embedded back into the original data, replacing the segment which contained the multiples.

(b) The VA of the data after multiple suppression. Note that in the absence of multiples (which were energetic events) the correct VA trend can be more readily discerned than in 3.20(b).

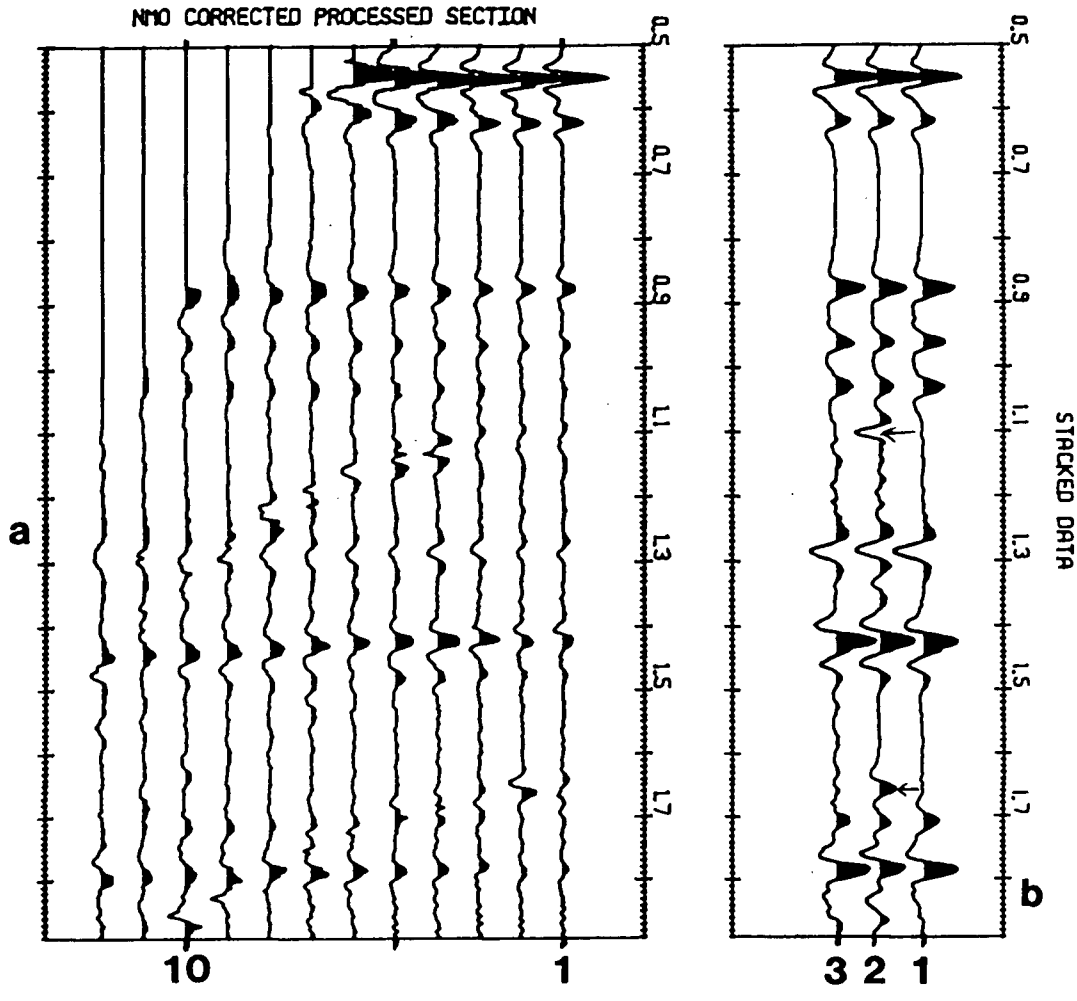


FIGURE 3.23

(a) The multiple suppressed data after normal moveout correction using a velocity function picked from 3.22b. The residual events remaining from the multiples can still be seen, but these do not stack constructively.

(b) Three versions of a stack of the data: (1) the stack of a synthetic data set produced without multiples: this is our desired, or optimum result; (2) the stack of the data in 3.20(a) after NMO correction. Notice the multiple arrivals at 1.1 and 1.65 s which stacked constructively; and (3) the stack of the multiple suppressed data in 3.22(a) after NMO correction. In comparison with (2) we note the absence of multiple events.

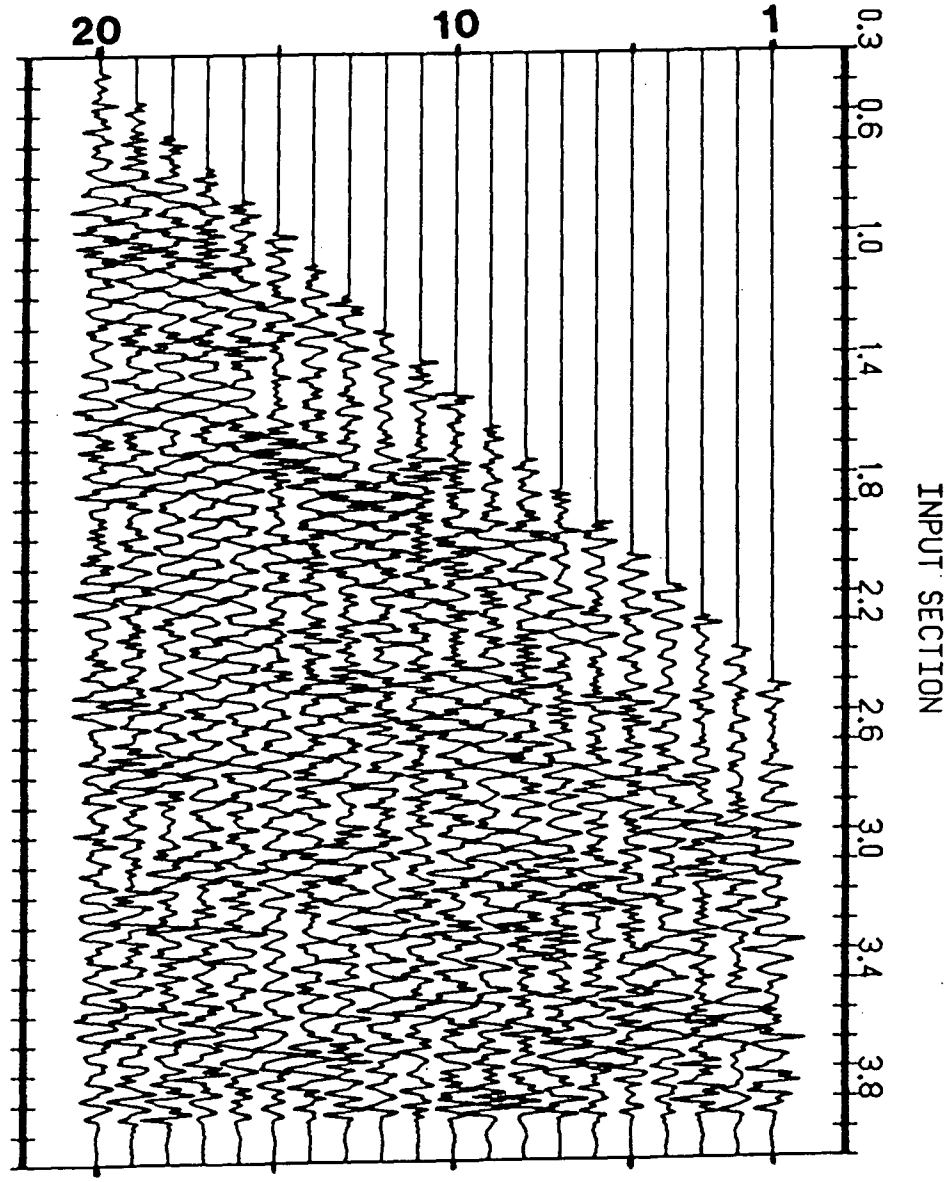


FIGURE 3.24

Every third trace from a marine gather of 60 traces, after bandpass (5 - 55 Hz), muting, and application of an AGC.

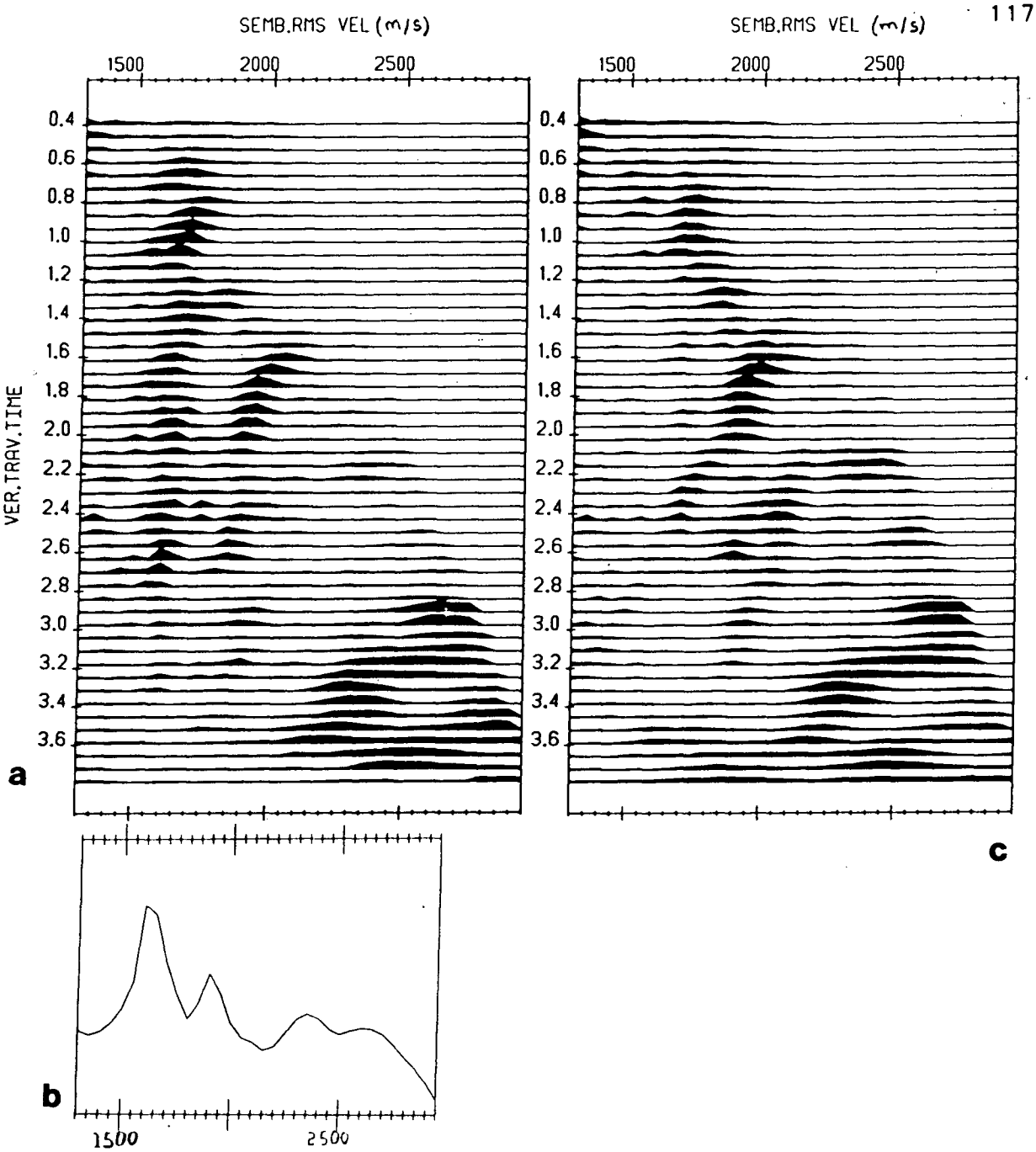


FIGURE 3.25

(a) VA of the data in 3.24. Note the vertical trend of energy at 1620 m/s due to water bottom related multiples.
(b) The stack of the VA map above. The maximum at 1620 m/s was detected by the algorithm and subsequently used as the multiple suppression velocity.
(c) The VA map of the section after multiple suppression. Note the absence of energy at 1620 m/s after 0.5 s, and the enhancement of the primary arrival peak at 2.2 s with velocity 2400 m/s.

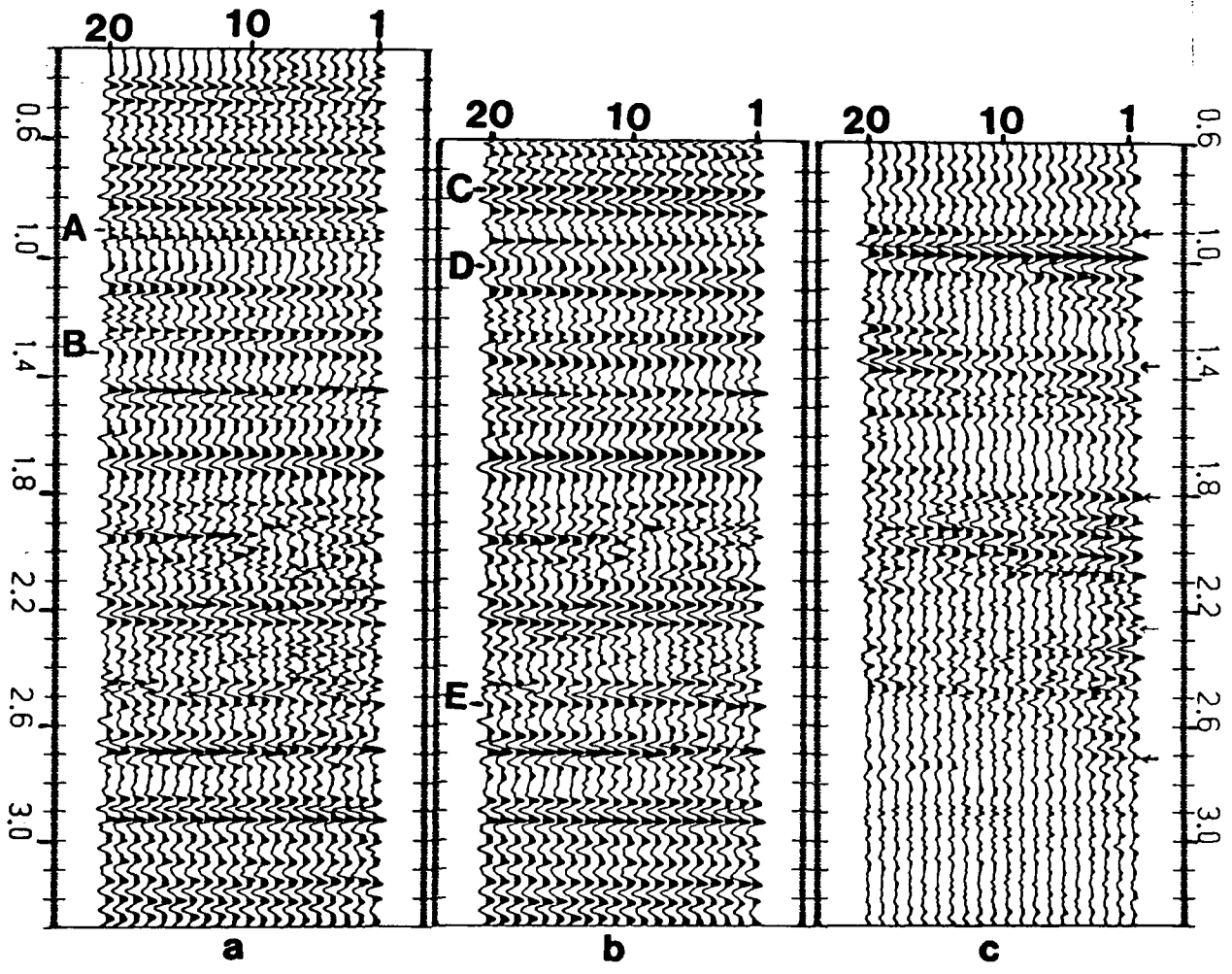


FIGURE 3.26

(a) The conventional stack of 20 gathers.
(b) The same gathers after multiple suppression using the automatically located multiple velocity of 1620 m/s. Note the absence of the event at 0.91 s (A) and changes at 1.35 s (B). Also note the enhancement of events at 0.86 s (C), 1.10 s (D), and the pair at 2.45 and 2.55 s (E). Only the processed section of the gathers has been plotted here.
(c) The difference section, to emphasize the location and nature of the differences between the procesed and unprocessed data. Bands of energy associated with the multiples appear at about 0.44 s intervals, indicated by the arrows.

CHAPTER 4.

SIMILARITY MEASURES.

SECTION 4.I: TRACE CLUSTER ANALYSIS.

a. Introduction.

Recall that the eigenvector elements from *equation (5)* are simply the weights applied to each input trace to construct a given principal component. We infer that if a group of traces (group 1, say) has the same weight for the first principal component, then those traces are similar to each other. If we then see another group of traces (group 2, say) which had weightings which differed from those of group 1, but were similar within group 2, we infer that the traces in group 2 are similar to each other but different from the members of group 1.

In this way one may strive to recognise natural groupings within a multichannel data set. To be meaningful, such a procedure must only consider a narrow time window centred about a particular event. In effect, we will be searching for both changes in character and repetitions of certain characteristics along the horizon in the window.

Milligan et al. (1978) considered the above problem for the case of 'acoustic pinger' data whilst mapping the distribution of bottom sediments in a shallow bay. They analysed the bottom

reflection pulse, and proceeded to correlate the pulse character with known bottom sediment type with considerable success. They implemented the KL transform via *equation (6)*, because they had many more traces than time samples, and derived the eigenvector elements by correlation.

Hagen (1982) tried to adapt their method for application to seismic reflection data to facilitate the recognition of stratigraphic character change. He noted that the 'instantaneous phase' (Taner et al., 1977) was sometimes indicative of porosity (a large instantaneous phase corresponding to a large value of porosity) and decided to perform the KL decomposition of the Instantaneous Phase representation of his data. He used the dominant eigenvectors associated with the KL transform as the input for a cluster analysis routine, and searched for natural groupings of eigenvector elements, as well as for groupings about user defined locations. With a single real KL transform, he was limited to obtaining the eigenstructure of the Instantaneous Phase data alone. However, since the introduction of the complex KL transform (Levy et al., 1983) we may obtain the eigenstructure of the complex trace of the original data.

In synthetic data examples, I found that the features which dominated seismic trace clustering were the gross structural changes in the data, e.g. fault offsets or steep dips. Subtler variations in the data such as phase drifts, or other changes in wavelet character could only be isolated when gross structural

features were absent.

A problem with the example considered by Hagen (1982, Figure 2; reproduced in Figure 4.1a) is that the upper event (positive peaks at about 1.66 seconds) pinches-out or joins with the lower event near traces 155 and 225. Also, the data were not levelled on the layer in the zone analysed. Consequently there is a two fold gross character perturbation to an otherwise flat lying structure. As expected, we see in Figure 4.1b, the first and second eigenvectors roughly in antiphase (Hagen 1982, Figure 6; reproduced in Figure 4.1b) and showing gross character changes near traces 170 and 215. So, to a first approximation, Hagen's user defined clustering about traces 214 and 230 will simply tell us which traces lie in the flat central region of the data (containing two horizons), and which traces lie in the extremal regions (containing only one horizon). However, in essence the method Hagen presented was sound, and I proceeded to follow his lead and examine data sets using the real and complex KL transforms.

I introduce here an alternative approach to the isolation of subtle character change from the dominant structures. By analogy with misfit reconstruction (defined by *equation (26)*), I perform cluster analysis on the eigenvectors separately, or in groups, and concentrate on determining the groupings displayed by these analyses. This procedure would be expected to be of use when we had subtle phase changes in an otherwise flat lying structure.

The horizon's structure would dominate the first principal component and be reflected in the makeup of the first eigenvector. The deviation from the horizontal structure would be characterized by subsequent principal components and may be seen in their eigenvectors.

I also diverge slightly from the approach of Hagen in that I compute a membership probability for each of the traces with respect to each of the groups, on the basis of the statistical mode of the given group. That is, for the $i.th$ seismic trace $x_i(t)$, the probability of belonging to the $j.th$ group G_j, on the basis of information obtained from the $k.th$ principal component is given by:

$$P\{x_i(t) \epsilon G_j\} = [1.0 \; / \; D\{r_{ik}; w_{jk}\}]/NORM$$

where:

$$D\{r_{ik}; w_{jk}\} = |r_{ik} - w_{jk}|$$

is the 'distance' of the eigenvector element r_{ik}, from the mode w_{jk}, of the eigenvector elements of the members of the $j.th$ cluster group G_j, and:

$$NORM = \sum_{j=1}^{nc} [1.0 \; / \; D\{r_{ik}; w_{jk}\}]$$

where nc is the number of cluster groups formed.

b. Synthetic Data Examples.

To demonstrate the application of this method, I considered four basic synthetic data examples. The first involved using the first two layers of Figure 3.2a, i.e. gently dipping layers, over which the wavelet is progressively phase shifted, offset with a vertical fault. Figure 4.2 shows these data, the first and second eigenvectors (Figure 4.2b), and the cluster groupings determined separately from them (Figure 4.2c). In this case only two groups were needed to highlight the fault, but in general I found that one should request more groups than one expects, so that traces which are completely dissimilar ('bad' traces) can fall into their own group without adversely affecting the overall results. The results from the complex KL option were identical, whereas the results of a run using the instantaneous phase of the data were slightly inferior. I have displayed the cluster groupings from the first two eigenvectors separately so as to indicate how each side of the fault dominates a particular eigenvector. The side of the fault with most traces dominates the system, and appears on the first eigenvector as the most 'energetic' event. The compound clustering mode, which combines the effects of several eigenvectors as per *equation (27)*, yielded these same groupings when combining these first two eigenvectors (corresponding to 73% of the total energy).

In the second example, Figure 4.3, we see two flat lying closely spaced events with a phase drift increasing to $\pi/2$ at the centre, and falling-off again to zero at the last trace. The dominant structure is that of two parallel horizontal layers: consequently, the first eigenvector shows no particular structure. However, the second eigenvector is readily able to indicate the underlying, subtler features of the phase drift, and yields groupings symmetric about the central traces. In this instance, the compound clustering was dominated by the second eigenvector, as it contained the most pronounced structure, and the groupings from the combination of the first two eigenvectors (corresponding to 97% of the total energy) also delineated the phase drift. Again, the results from the complex KL option were identical, and the instantaneous phase results were slightly inferior.

In the case of gently dipping layers (Figure 4.4) we expect and do see a tendency for clusters to contain traces in groups moving progressively across the section. If the data had been flattened, these groups would not exist and no particular grouping patterns would be evident. Consequently, groupings which change systematically and progressively across the section will be indicative of a dip (in the absence of phase changes). These results are evident in the groupings for the first eigenvector. However, the second eigenvector shows the symmetric distribution of clusters similar to the example with a phase drift into the

centre of the group of traces. This is because the second eigenvector is telling us about the residual structure, which for a constantly dipping layer is a series of antisymmetric 'static' time shifts. This antisymmetry in the residual structure is seen as a symmetric pattern in the cluster groupings for the second eigenvector, as the algorithm uses the squares of misfits to determine the groupings. Similar results would be obtained if the data comprised flat layers with a constant phase drift in the waveforms across the event. Two such cases could not be distinguished by this method. Again, both the real and complex KL methods yield essentially the same results.

Finally, I consider the case of a gently dipping horizon with an anomalous zone (in this case a hump representing a reef or sand lens: see Figure 4.5a). Because the central hump deviates so much from the flat lying events, the energy associated with it is spread over several principal components. Consequently, the groupings containing the traces in the hump can be seen in the cluster analysis of the first 4 eigenvectors. Figure 4.5b and c, shows the results for the first and fourth eigenvectors, respectively. Both the real and complex KL methods clearly delineate this feature, but the instantaneous phase results were poor.

c. Real Data Examples.

1. The first real data example (Figure 4.6) is of 100 traces from a predominantly flat lying zone. The strong upper event (1) dominates the structural character, while the weak intermittent event (2) and the lower discontinuous event (3) are less coherent. The first eigenvector showed no particularly significant groupings. However, the second eigenvector (Figure 4.6b), shows three main groups, which correspond to (A) the predominant flat lying structure between traces 1 and 45, (B) the zones containing the deep trough at 0.61 s (traces 18 - 28, and 38 - 54), and (C) the fairly uniform zone between traces 73 and 94. The third and fourth eigenvectors (Figure 4.6c and d) also emphasize the location of the disrupted central zone.

2. The second real data example comprises events offset by a fault near trace 58 (Figure 4.7): to the right of the fault the horizons drop slightly. The cluster group membership, based on the first three eigenvectors, is shown below (Figure 4.7b, c, and d). For the first eigenvector, the first group, A, contains most of the traces to the left of the fault. However, a small anomaly (B) clearly shows up in the groupings; namely traces 33 - 40: this corresonds to the anomaly at 1.66 s. To the right of the fault, a third group C brackets traces 60 to 76, and may delineate a second fault to the right near trace 80 at 1.6 s.

The groupings determined on the basis of the second eigenvector show a symmetric pattern belying the symmetry about the central fault. These groupings are symmetric rather than antisymmetric, as the clustering criterion was based on the squares of differences. The third eigenvector again clearly denotes the central fault.

3. The next example, of 96 traces, is of an event which undergoes a marked character change across the section (Figure 4.8). This can be seen in the two eigenvectors (Figure 4.8b and c). Also of interest here is the appearance of the third eigenvector (Figure 4.8d), which clearly denotes the slight discontinuities at 2.04 s near trace 20 and again near trace 36, and the truncation of the event near trace 85 (these discontinuities are marked in the figure).

4. In the final example (Figure 4.9), I show a 0.1 s window of data from a sand bar. The bar is seen at about 0.68 s, between traces 29 and 70. In this case, the first eigenvector clusterings delineate a group in the region of the bar (B) bracketed by members of a different group (A) (Figure 4.9b). In the second eigenvector (Figure 4.9c), we also note a discontinuity between traces 65 - 85, where the reflection character of the upper event (at 0.67 s) changes, becoming stronger. The third eigenvector (Figure 4.9d) resembles the first in that it denotes a central group flanked by members of a second group.

d. Discussion.

For data in which the dominant structures are parallel, cluster analysis may provide a useful tool to aid the explorationist in locating unusual features. Much depends on the presentation of these data. A colour overlaying technique of presentation may prove quite useful in conveying the information clearly.

Features such as fault boundaries, pinch-outs, and lenses, all fall in the category of structures for which the method has the potential of identifying as being 'different'. However, as with many novel techniques, much interactive study is necessary in order to assess the full potential (if any) of the method. Comprehensive analysis of well studied areas is needed to give an outline of the reliability and meaning of the results seen in the cluster groups. This can most readily be accomplished in an environment where the appropriate data sets are available.

The major pitfall with this technique is likely to be its sensitivity to non-parallelism in the input data. In this regard, cluster analysis on the basis of eigenvector elements suffers from the same problem as does the misfit reconstruction technique presented earlier.

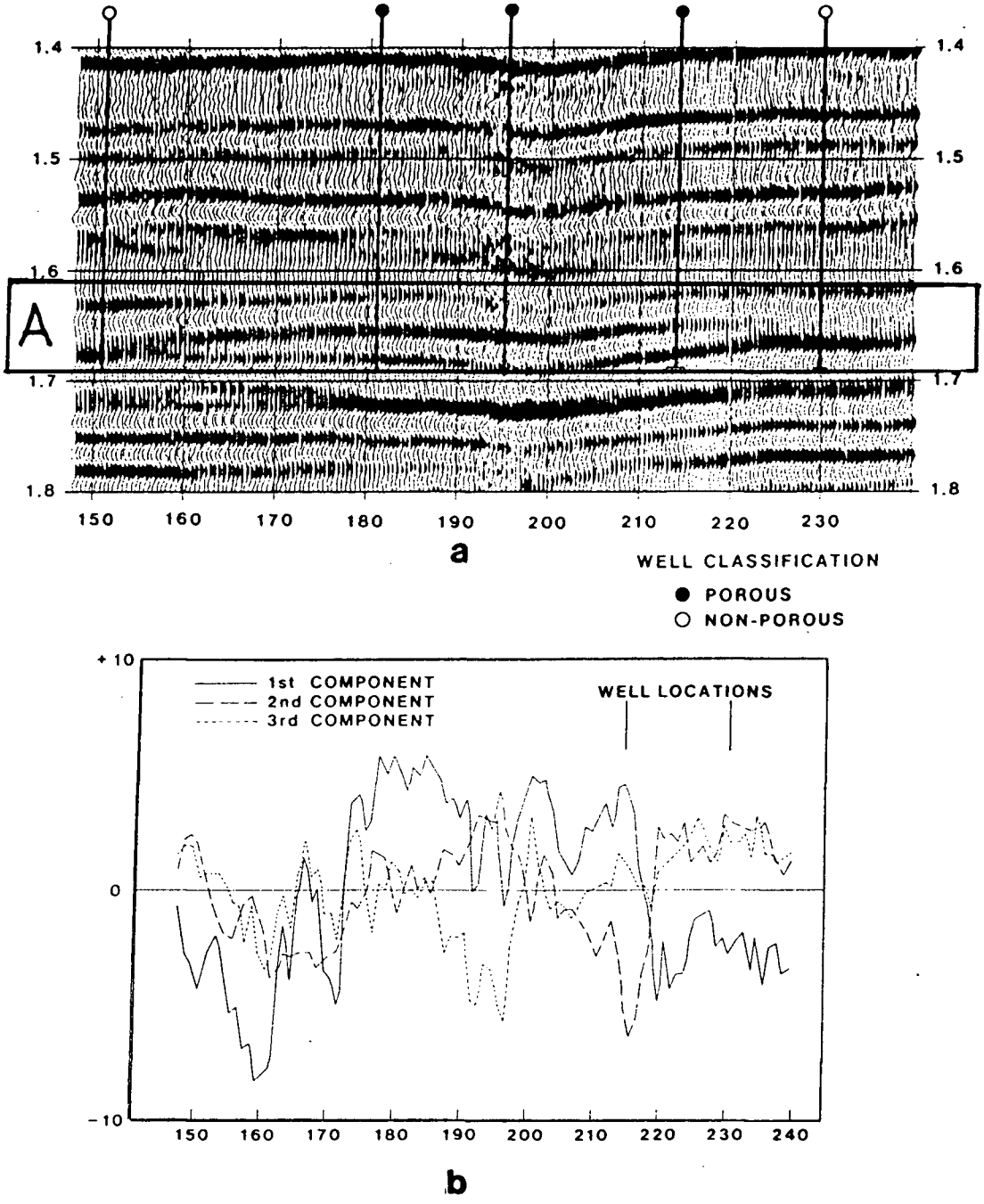


FIGURE 4.1

(a) The data used by Hagen (1982) for his cluster analysis. Note the central pair of events which are flanked by single horizons (A).

(b) The first three eigenvectors for decomposition of the data in (a). Note the antiphase relationship between the first and second eigenvectors, and how the character change coincides with the pinch-outs of the flat central region of the data (diagrams from Hagen, 1982).

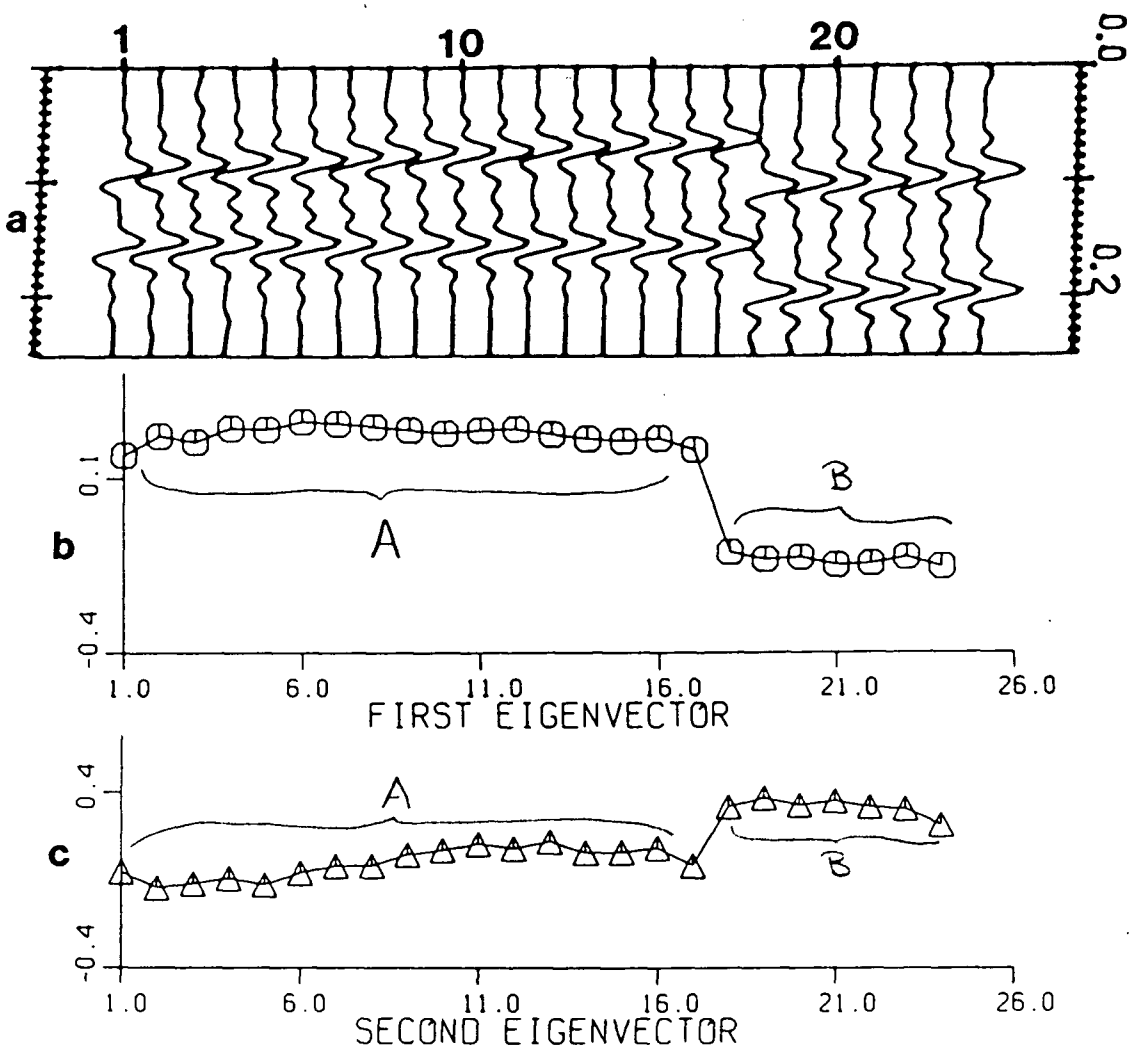


FIGURE 4.2

(a) The seismic representation of the upper two layers dipping at about 2 ms per trace, offset by a vertical fault.
The first two eigenvectors (a and b) with the cluster group memberships denoted on them. Note how the traces on either side of the fault fall into separate groups.

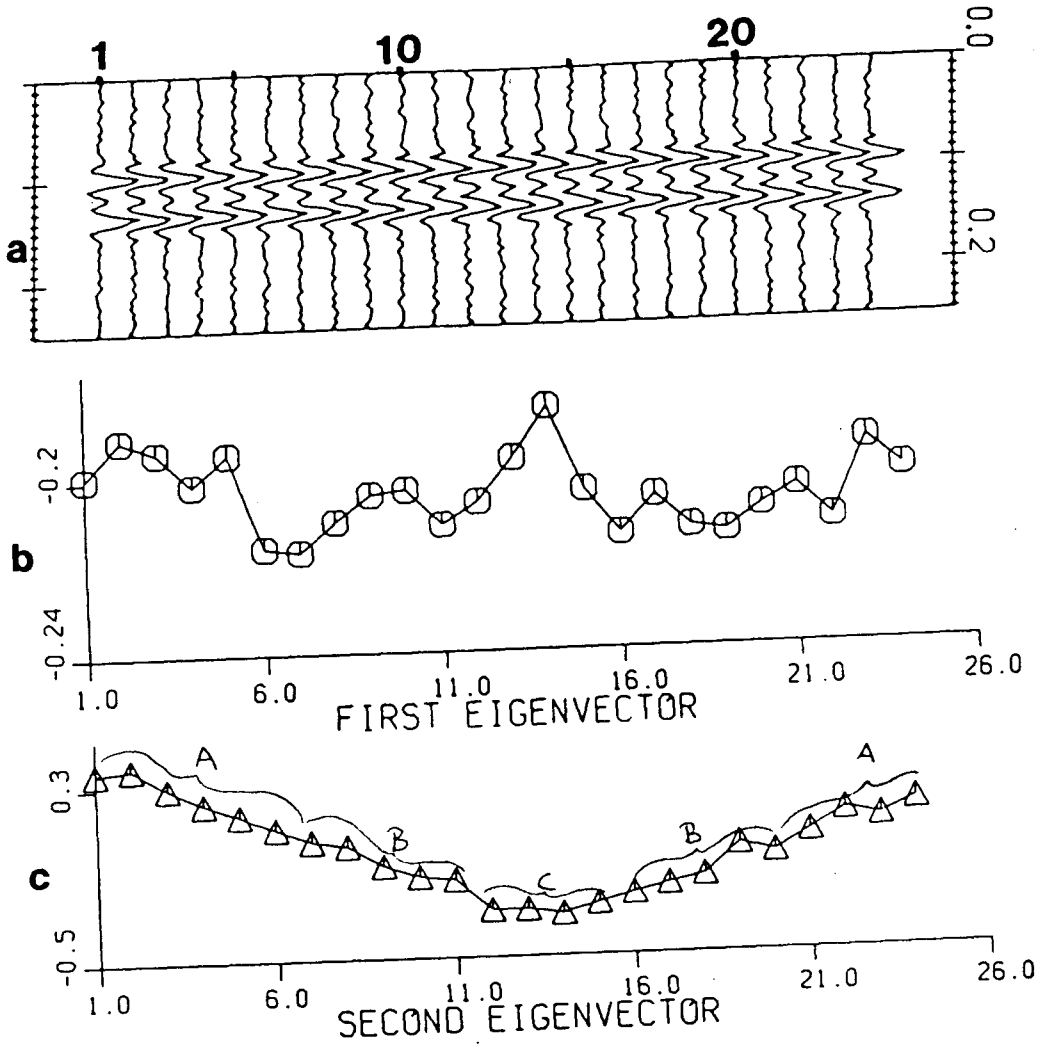


FIGURE 4.3

(a) The seismic representation of two flat layers with a phase drift increasing to $\pi/2$ at the centre and decreasing back to zero at the far side.
(b) The clusterings of the first eigenvector show no particular patterns, as the dominant structure is flat.
(c) The clusterings for the second eigenvector display distinctive groupings delineating the phase drift into the centre and out again.

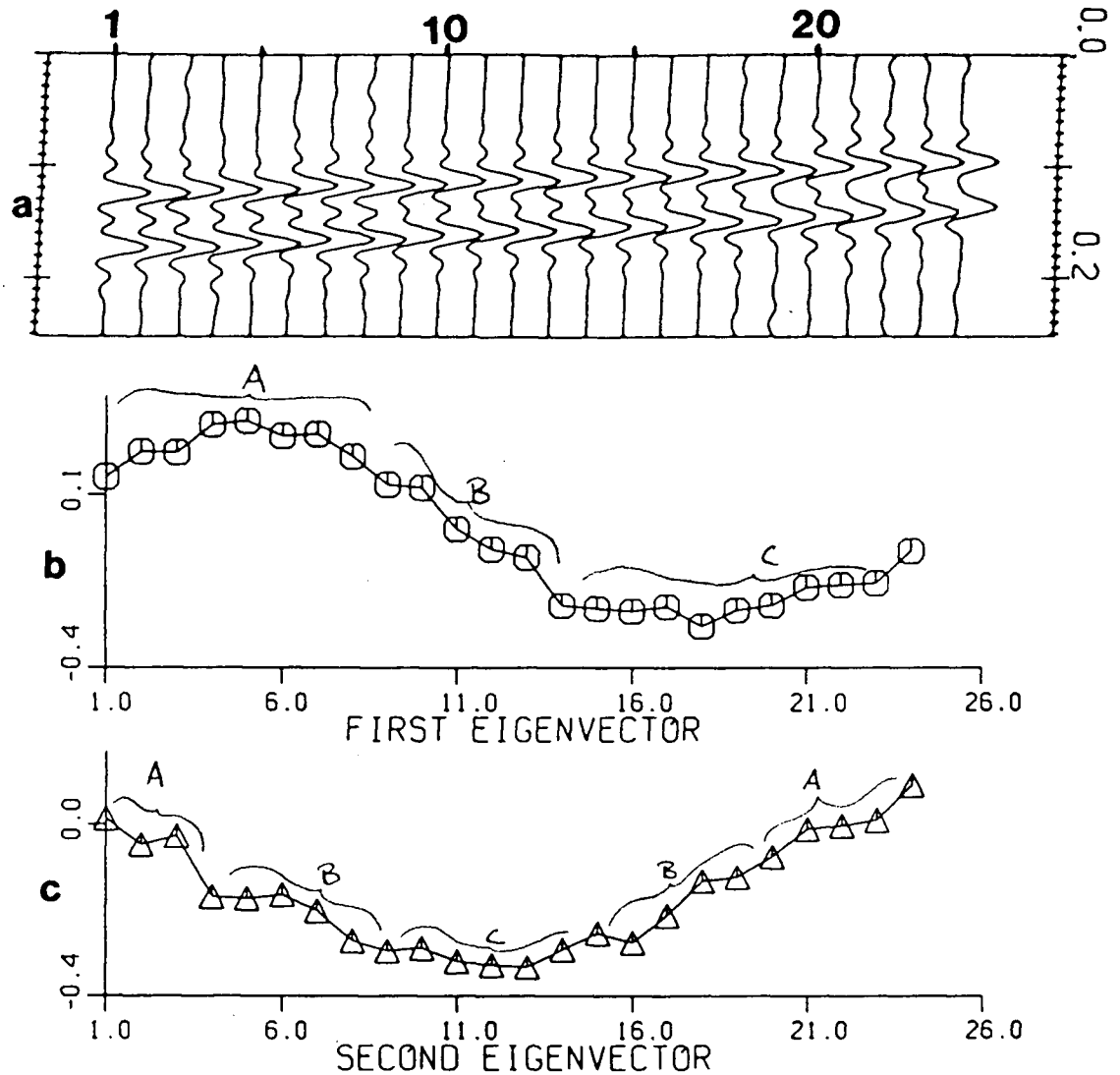


FIGURE 4.4

(a) The seismic representation of two layers dipping at 2 ms per trace.

(b) The first eigenvector delineates groupings with trace membership moving progressively across the section. This is characteristic of a dip, or of a constant phase drift over flat layers.

(c) The second eigenvector delineates the residual structure, which for a constantly dipping layer is an antisymmetric feature.

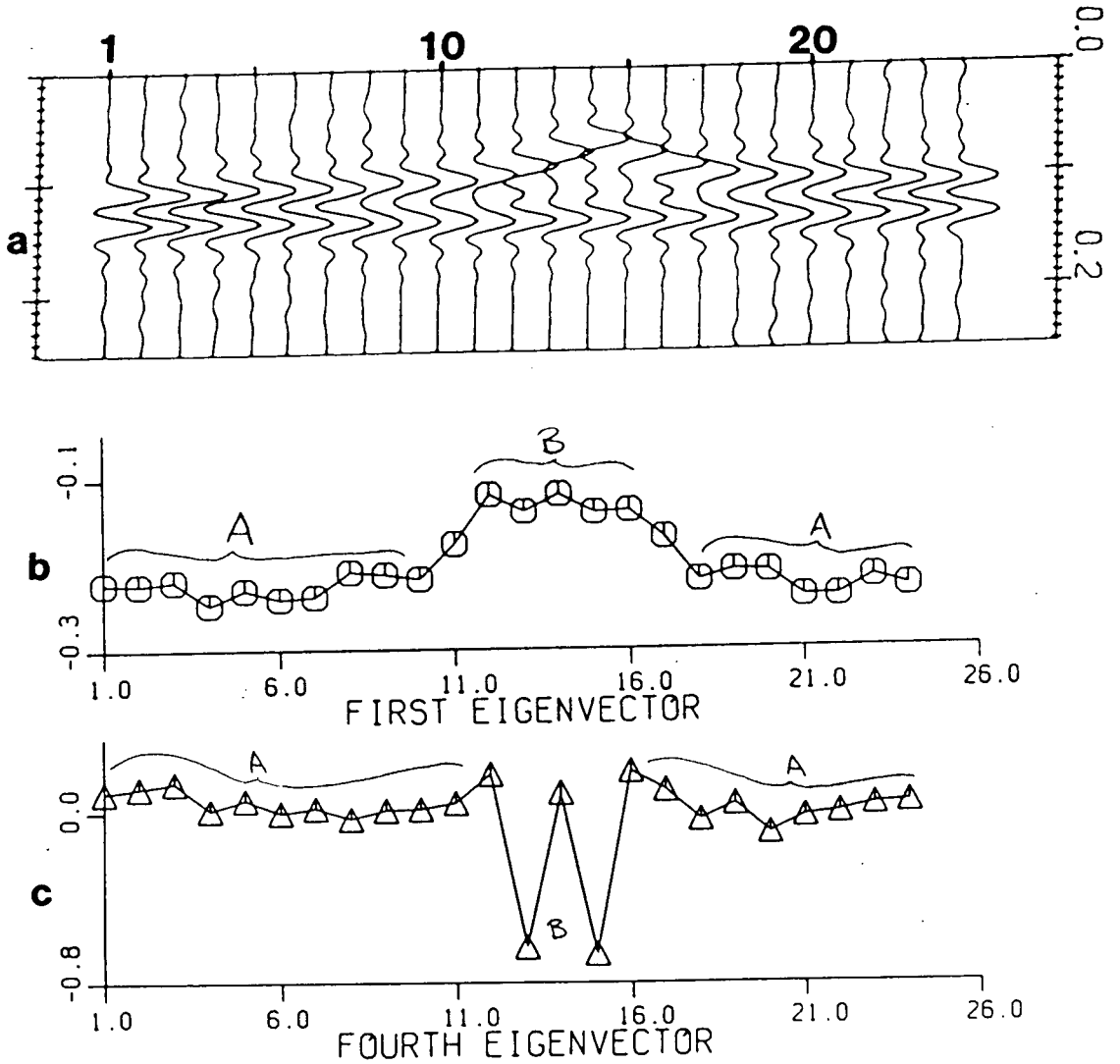


FIGURE 4.5

(a) The seismic representation of two flat layers with a central hump.
The first (b) and fourth (c) eigenvectors clearly delineate this feature.

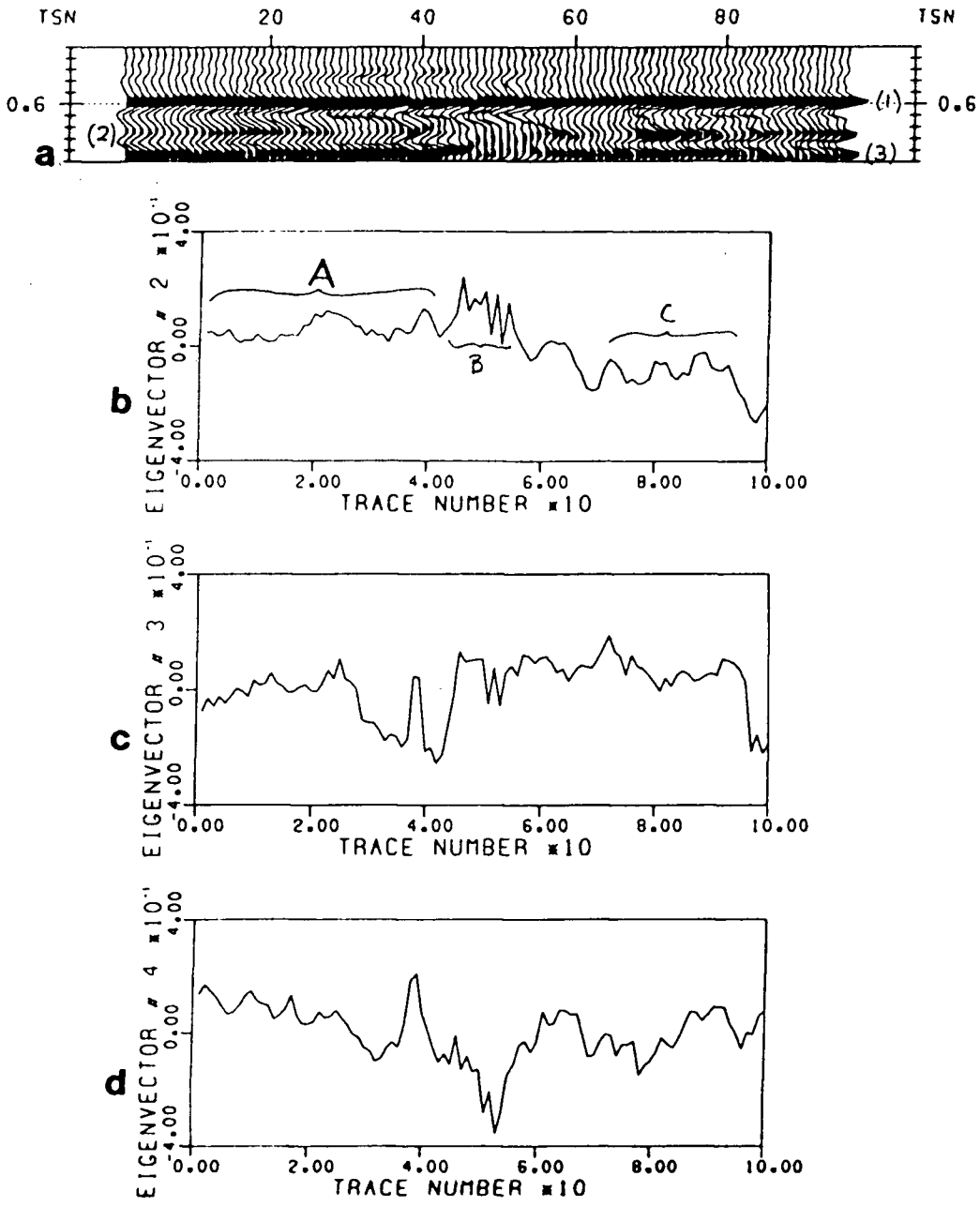


FIGURE 4.6

(a) One hundred traces from a real stacked seismic section, with three characteristic events: the strong upper event (1), the weak intermittent event (2), and the lower event (3). The first eigenvector is fairly flat, hence shows no clear groupings and is not shown.

(b) The second eigenvector shows three major groups which correspond to (A) the flat lying structure between traces 1 and 45, (B) the deep trough at 0.61 s (traces 18-28 and 38-54), and (C) the fairly uniform zone between traces 73 and 94.

(c) The third eigenvector appears to reflect character change in the trough below the first event (1), while the fourth eigenvector (d), again demarcates events to the right and left of the central discontinuous zone.

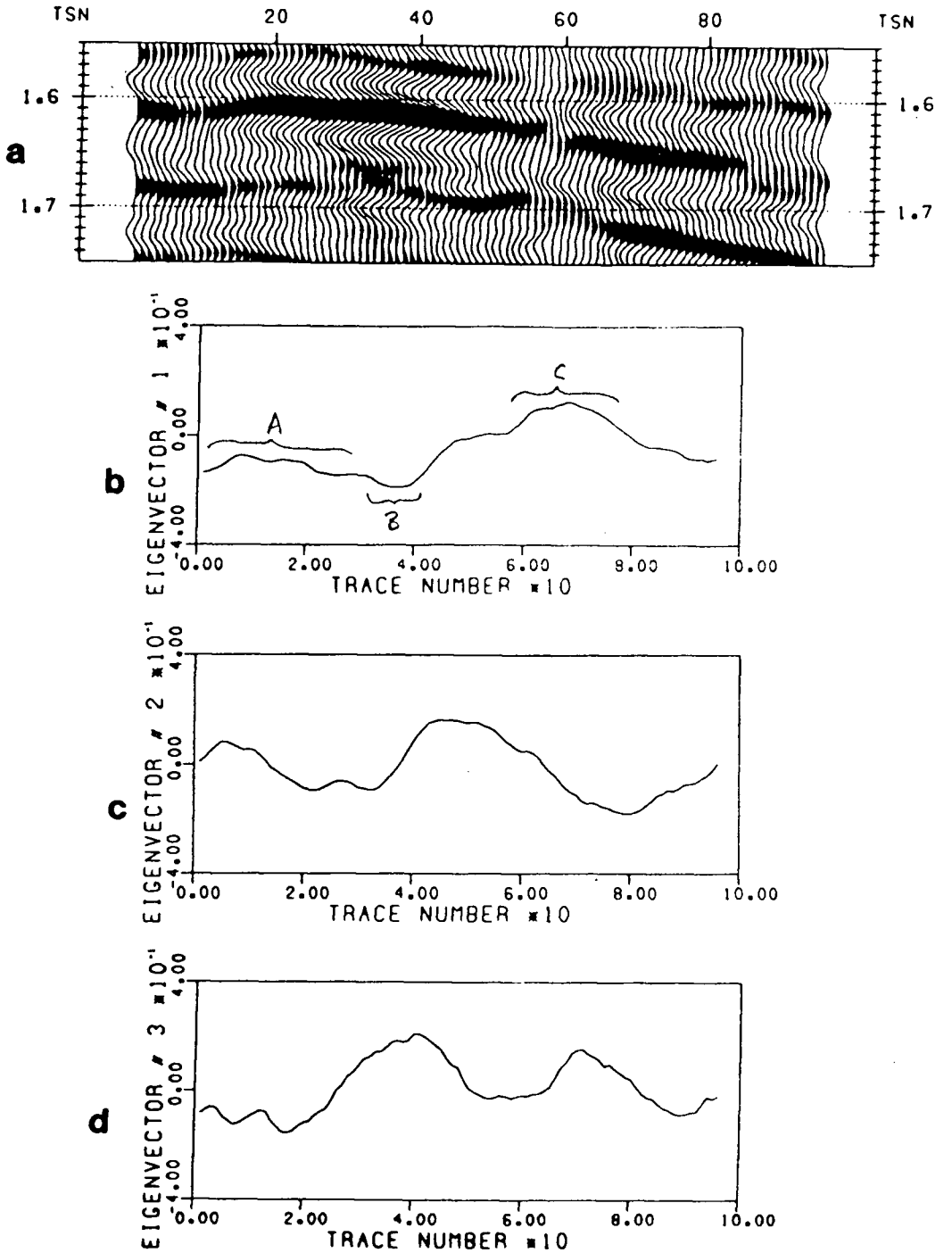


FIGURE 4.7

(a) Ninety six traces from a faulted zone.
(b) For the first eigenvector (1), group (A) contains most of the traces to the left of the fault. A small anomaly (B) can be seen between traces 33 to 40 at 1.66 s. To the right, a third group may delineate a second fault (C) near trace 76.
(c) The groupings from the second eigenvector show a symmetric pattern belying the symmetry about the central fault, as does the third eigenvector (d).

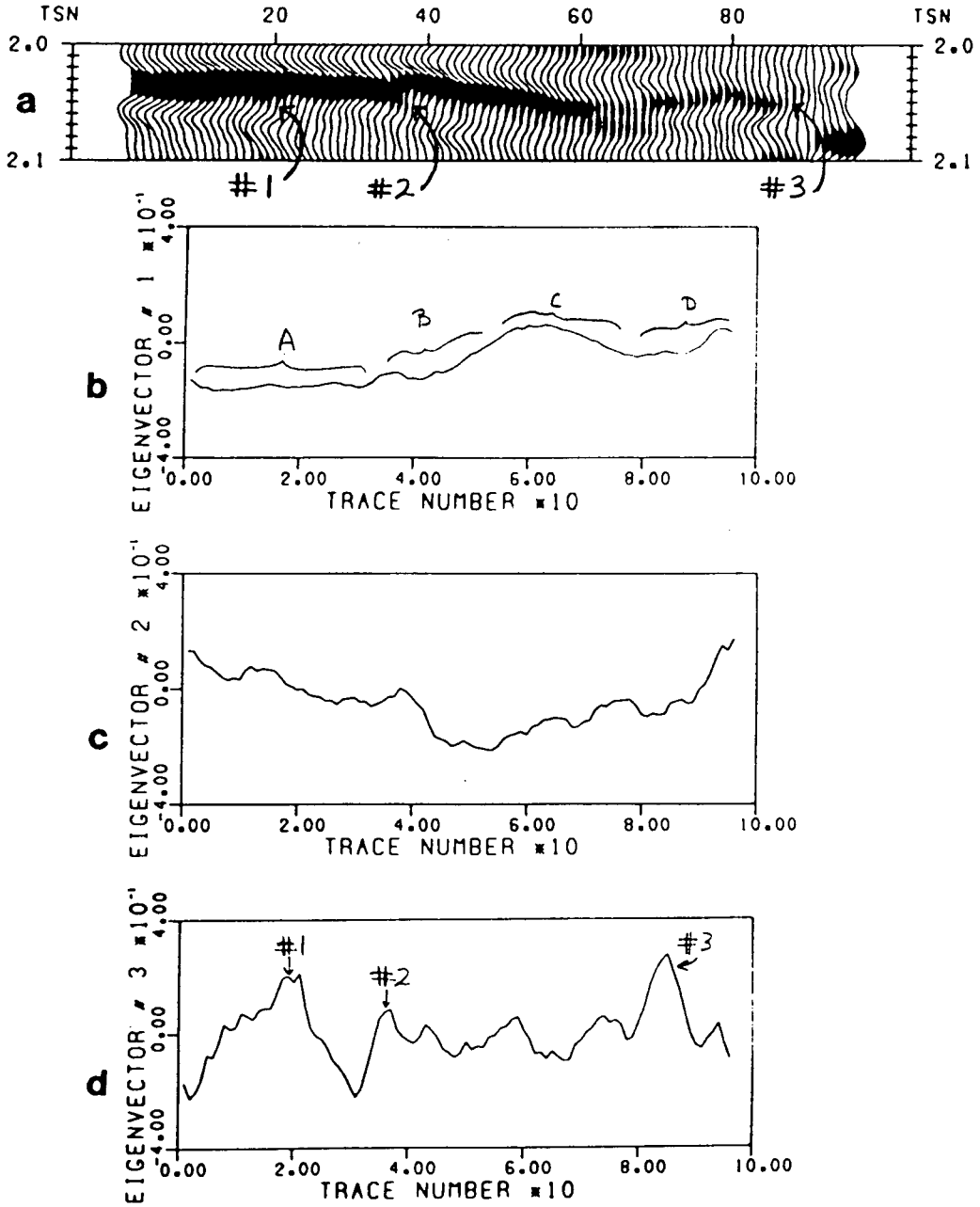


FIGURE 4.8

(a) Ninety six traces which change noticeably in character across the event. The first two eigenvectors (b and c) belie this change.

(d) The clusterings from the third eigenvector clearly denote subtle changes in the data: namely discontinuities near trace 20 and trace 36 at 2.04 s, the truncation of the event near trace 85 (marked #1, #2, and #3, respectively).

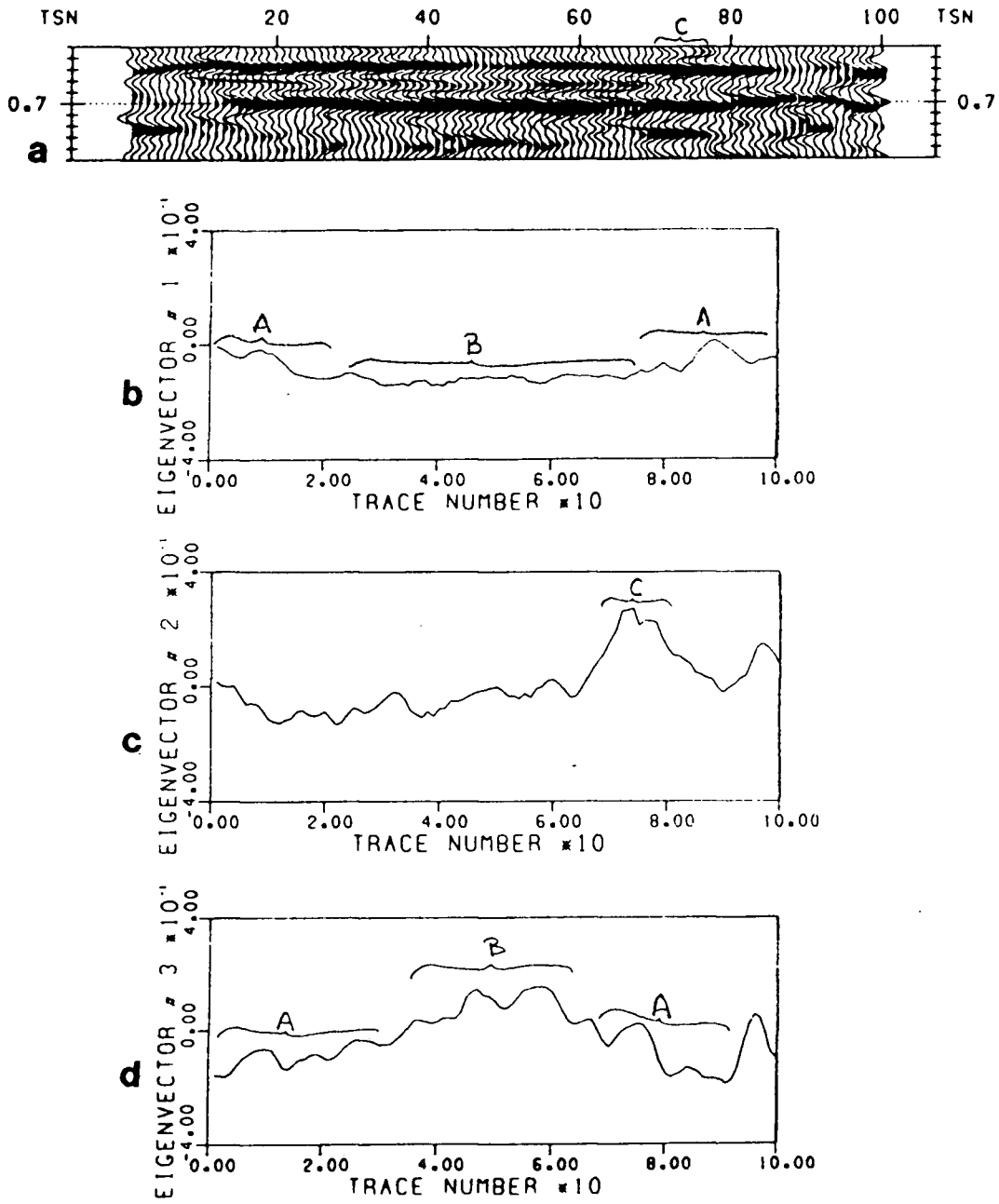


FIGURE 4.9

(a) One hundred trace of 0.1 s of data from a region of braided sand lenses, in a potential production zone. Note the persistance of a thin positive event at 0.68 s, between traces 29 and 70. In this case, the cluserings from both the first and second eigenvectors (b and c) delineate a group in the region of the bar (B) bracketed by members of a different group (A). Also in the second eigenvector we note a discontinuity between traces 65-85 where the reflection character near 0.67 s changes, becoming stronger (C). The third eigenvector (d) also demonstrates the presence of a central group, flanked by members of a second group.

SECTION 4.II: VELOCITY ANALYSIS.

a. Introduction.

In this section I show how the eigenvalues derived from the covariance matrix can be used to establish a correlation criterion, which under certain conditions may be superior to the commonly used semblance criterion (eg. Neidel and Taner, 1971). With $\chi(m)$ defined in *equation (21)* we note for normalized input traces that if $\chi(1) = \infty$ the signals are perfectly correlated while if $\chi(1) \simeq 1/(n-1)$ then the signals are nearly orthogonal.

b. The Effects of Static Shifts.

For wavelets which have been offset by a static shift we also expect $\chi(1)$ to provide a good indication of the correlation if the amount of static shift is less than that prescribed by *equation (33)*. As the static shift increases however, it will be impossible to represent the signal energy in terms of a single principal component; more principal components will be required. So, for the CKL velocity analysis (VA), a wavelet displaced by a small static time shift will be rotated so as to appear more coherent, and the VA should be less sensitive to the degrading effects of the time shifts.

c. Synthetic Data Examples.

In the following trials, the width of the VA time window was set to 16 samples, about 2/3 of the wavelet width. As the window was narrowed or broadened, the results for both semblance and KL VA methods deteriorated, as expected, because with a window which is too wide most events appear dissimilar, while with a very narrow window most events look similar, in that they are fairly constant and will have a large common mean component.

To demonstrate the performance of semblance versus the real and complex KL modified eigenvalue ratio, I consider the case of a single event (at 1 s, with an associated velocity of 1500 m/s). Twelve traces (Figure 4.10) were progressively subjected to static time shifts, ranging from zero to a maximum of $\leq \pm 8$ time samples. In Figure 4.11 we see in the left column the semblance VA for static time shifts of $\leq \pm 0$, ± 2, ± 4, ± 6, and ± 8 time samples. The peak (at 1500 m/s, and 1.0 s) is quickly smeared as the static shift increases. Also note the vertical smearing characteristic of semblance VA results. On the right of the figure are shown the results of an RKL VA for each of several $\chi(m)$ values for each static time shift. Note how the peak of the RKL VA can still be identified even at large static shifts by increasing the number of eigenvalues in the numerator of *equation (21)* (i.e. increasing m in $\chi(m)$). This effectively includes more

of the signal related energy in the numerator of *equation (21)*.

This excercise was repeated for the CKL VA, and the comparison with semblance results made again (Figure 4.12). Again, the KL VA method is readily able to denote the location of the peak as static time shifts are introduced.

Similar trials were performed for the case of increasing white noise levels. In this instance nothing was gained by increasing the number of eigenvalues in the numerator of *equation (21)*. This is because white noise adds energy uniformally to the diagonal of the covariance matrix and increasing the number of elements on the top line merely includes more noise energy in the numerator. A comparison of the semblance, RKL and CKL VA results are shown for 5 noise levels (40, 55, 70, 85, and 100%) in Figure 4.13). Although the semblance results degrade severely with increase in noise level, they remain superior to both the KL VA results. In other words, the KL VA method is more sensitive to high noise levels than is the semblance. However, at moderate noise levels (<30% by maximum amplitude) the KL methods are superior.

To examine the resolution of closely spaced 'thin beds', I repeated the above example with a four layer model. In the previous example, the velocity was quite low and hence the moveout was pronounced and the algorithms all worked well. Increasing the velocity reduces the curvature, and we expect to

see a smearing of the peaks in the VA map. Figure 4.14 shows the 12 synthetic seismograms, sampled at 4 ms, representing the three interface reflections from four layers (with thickness and velocity: 1800 m, 1800 m/s; 50 m, 1850 m/s; 50 m, 1900 m/s; and 50 m, 1950 m/s respectively). Note that in such a model, the stacking velocities change very little. The results from the semblance, RKL and CKL VA's are shown in Figure 4.15 (the first three $\chi(m)$ criteria VA's for the KL runs are shown). The RKL result for $\chi(2)$ separates the three events, whereas the CKL $\chi(1)$ VA does less well and the CKL $\chi(2)$ VA is somewhat smeared, although the three events can be seen. The semblance VA fails completely, and smears the triplet into a vague zone of high amplitude. This indicates that the KL method should be better able to resolve proximal events. However, this conclusion is not confirmed by the majority of subsequent results.

In the following synthetic data examples, I present the case of a multilayered reflection sequence. The input seismograms, sampled at 4 ms, Figure 4.16a correponding to the adjacent layered earth model (b), were contaminated by a small amount of white noise (up to 5% of the maximum signal amplitude), primarily to stabilize the performance of the algorithm at zones where no signal is present. The results of the semblance, RKL, and CKL velocity analyses are shown after gain control in Figure 4.17a, b, and c, respectively.

We see that the KL results (presented for $\chi(1)$, $\chi(2)$, and $\chi(3)$, respectively) give better temporal resolution of events than does the semblance. This has important advantages when stacking velocities are to be inverted to recover interval velocities. However, the semblance results are fairly evenly balanced and will also yield a good stacking velocity curve. None of the VA maps in this example was able to delineate the central peak of the triplet at 1.06 s from its adjacent peaks (at about 1.00 and 1.12 s, respectively).

In the next example I demonstrate the performance of the VA algorithms when statics are present. Static shifts with values as large as ±12 ms are introduced (by deleting or adding zeroes to the start of each trace). Consequently more than one principal component is needed to represent the signal energy, and $\chi(3)$ was visually chosen as the best correlation criterion. The data, contaminated with statics are shown in Figure 4.18. The VA results (Figure 4.19a, b, and c) are considerably worse than those in Figure 4.17, but the peaks exhibited by the $\chi(2)$ and $\chi(3)$ plots for both the RKL and CKL VA's are at the correct velocities. On the other hand, the semblance results are smeared and multiple peaks appear for certain times (e.g. at 1.34 seconds where the semblance power is large).

d. Real Data Examples.

The 16 traces shown in Figure 4.20a are taken from a CSP gather shot in northern Alberta by Chevron Canada Resources Ltd. The offsets range from 582 to 1662 m, with a geophone spacing of 72 m, and the basement is at 1.6 s. For this examination, I have selected data so as to provide 1.2 s of zero offset time in the VA map.

The semblance results (Figure 4.20b) show a smeared trajectory, but picks can readily be made down to 1 s. In an attempt to determine a good $\chi(m)$ to use in the KL VA, I selected two small windows: one from 0.0 - 0.2 s, and one from 0.9 - 1.1 s. A comparison of results for semblance, RKL, and CKL VA's is shown in Figure 4.21 for the two windows repectively. Both the RKL and CKL results suggested using $\chi(1)$, $\chi(2)$, or $\chi(3)$. The RKL and CKL VA maps for the complete data sample are shown for all three $\chi(m)$ values, after application of an AGC operator, in Figures 4.22 and 4.23, respectively.

A major problem that arises with both the KL VA methods is the lack of consistency down the VA map. A $\chi(1)$ criterion works well at the start of the map, but not later. None work well in the region near 0.4 s; the $\chi(2)$ and $\chi(3)$ criteria work well at greater times. The general trends are those expected, i.e. we need to increase m in $\chi(m)$ with increasing time, as the moveout

hyperbolae are becoming flatter and static shifts will be more noticable.

The second real data set I examine, was collected during the UBC-RECOPE project in Costa Rica in 1982 (von Breymann and Clowes, pers. comm., 1984). It consists of wide angle reflection data shot with an airgun in deep water (>2 km) (see Figure 4.24a). The offsets for the data subset I consider range from 1630 to 4415 m with an average receiver spacing of 120 m. The VA maps span the zero offset two way travel time window 2.9 to 4.4 s (the water bottom reflection comes in at about 3.0 s).

To determine which $\chi(m)$ criterion to use for the KL VA's, two small time windows were chosen (at 3.0 and 4.0 s: see Figure 4.25). However, these proved to be misleading, as the best overall results were obtained using $\chi(1)$ and $\chi(2)$, and not with $\chi(4)$ as suggested by the sample windows. The results from the semblance, RKL, and CKL VA's are shown after AGC in Figures 4.24b, 4.26, and 4.27, respectively. For the KL VA's, results for both $\chi(1)$ and $\chi(2)$ are shown.

From the semblance map, picks can be made for:

τ = 2.95 s, V_{RMS} = 1500 m/s;

τ = 3.50 s, V_{RMS} = 1600 m/s; and

τ = 4.25 s, V_{RMS} = 1700 m/s.

From the RKL maps ($\chi(1)$ and $\chi(2)$), picks can be made for:

τ = 2.95 s, V_{RMS} = 1500 m/s;

τ = 3.04 s, V_{RMS} = 1550 m/s;

τ = 3.18 s, V_{RMS} = 1580 m/s;

τ = 3.45 s, V_{RMS} = 1650 m/s; and

τ = 4.23 s, V_{RMS} = 1710 m/s;

And from the CKL maps ($\chi(1)$ and $\chi(2)$), picks can be made for:

τ = 2.95 s, V_{RMS} = 1500 m/s;

τ = 3.04 s, V_{RMS} = 1550 m/s;

τ = 3.18 s, V_{RMS} = 1580 m/s;

τ = 3.45 s, V_{RMS} = 1640 m/s;

τ = 3.90 s, V_{RMS} = 1690 m/s;

τ = 4.45 s, V_{RMS} = 1750 m/s;

In the RKL VA maps, some double arrivals are seen (e.g. at 3.20 and 4.43 s) which are probably due to the internal structure of the wavelet being resolved, i.e. the map shows the reflections from the leading and trailing edges of the downgoing wavelet. All the KL VA's were dissapointing in their general appearance, as I had expected greater clarity and definition of arrivals. Decreasing ΔV (the VA velocity increment) from 50 to 25 m/s improved the semblance, but not the KL results. Band passing the data from its original 2 - 85 Hz down to 5 - 50 Hz again did not improve the KL VA's. Bandpassing should reduce the amount of uncorrelated high frequency energy, and thus reduce the degradation of the $\chi(m)$ criterion. Uncorrelated high frequency

energy would act as white noise, adding power uniformally to the covariance matrix diagonal and degrading $\chi(m)$.

e. Discussion.

The eigenvalue ratio similarity criterion constitutes a highly sensitive similarity measure. However, with real data the measure seems to fail severely for noisy data. In a complicated CDP gather with many interfering moveout hyperbolae, the CKL method will be able to rotate most waveforms and make them so similar that the VA will have maxima at spurious τ and V_{RMS} locations. Therefore, this method must be treated cautiously.

For early arrivals, such as those seen in the real data example of Figure 4.21, the KL method works very well in comparison to the semblance VA. However, for the real data example with higher noise levels (Figure 4.25) the semblance result is clearly superior. The technique may thus prove to be of use for high resolution near surface velocity inversion problems.

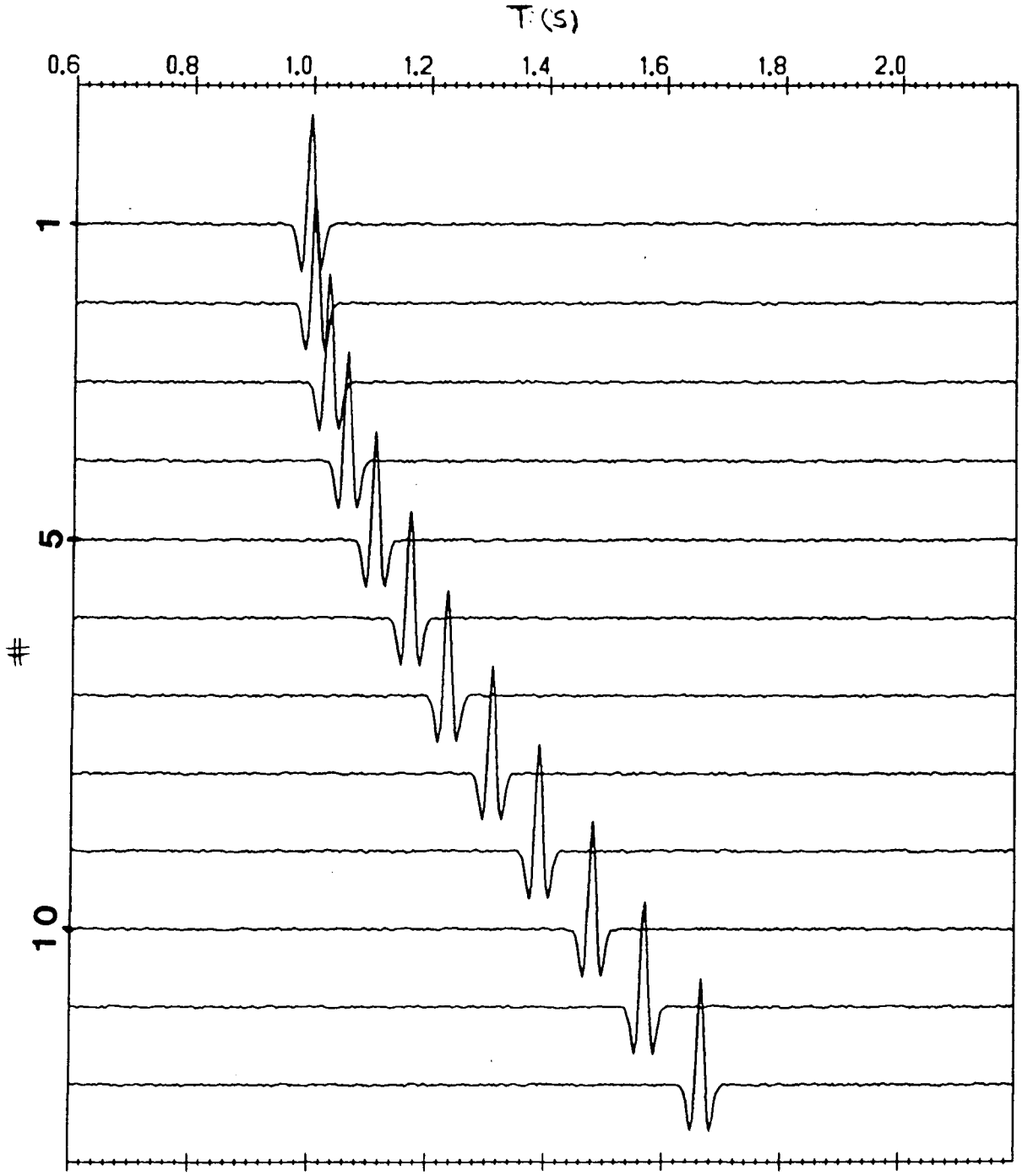


FIGURE 4.10
Twelve traces representing reflections from the base of a layer 750 m thick with velocity 1500 m/s. Offsets range from 0 to 2000 m.

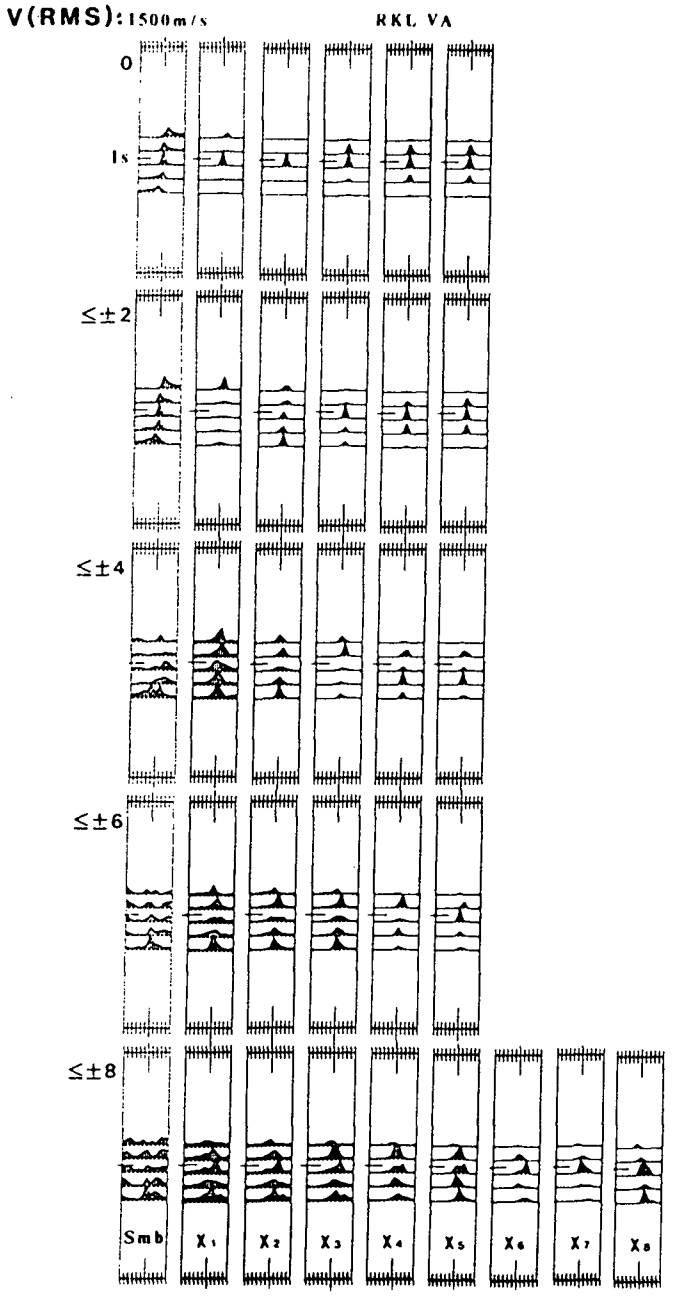


FIGURE 4.11

Left column: Semblance criterion VA for progressively increasing static time shifts; down the column: up to 0, ±2, ±4, ±6, ±8 sample points (i.e. 0, ±8, ±16, ±24, ±32 ms).
Second left to far right: RKL eigenvalue ratio criterion; columns for $\chi(1)$, $\chi(2)$, ..., $\chi(8)$.
As the static scatter increases, the semblance VA breaks down and is unable to reliably locate the peak corresponding to the arrival. For the RKL VA, we note how a progressing increase in m, in $\chi(m)$, counteracts the degrading effect of an increase in static scatter. The velocity tick marks are separated by 50 m/s.

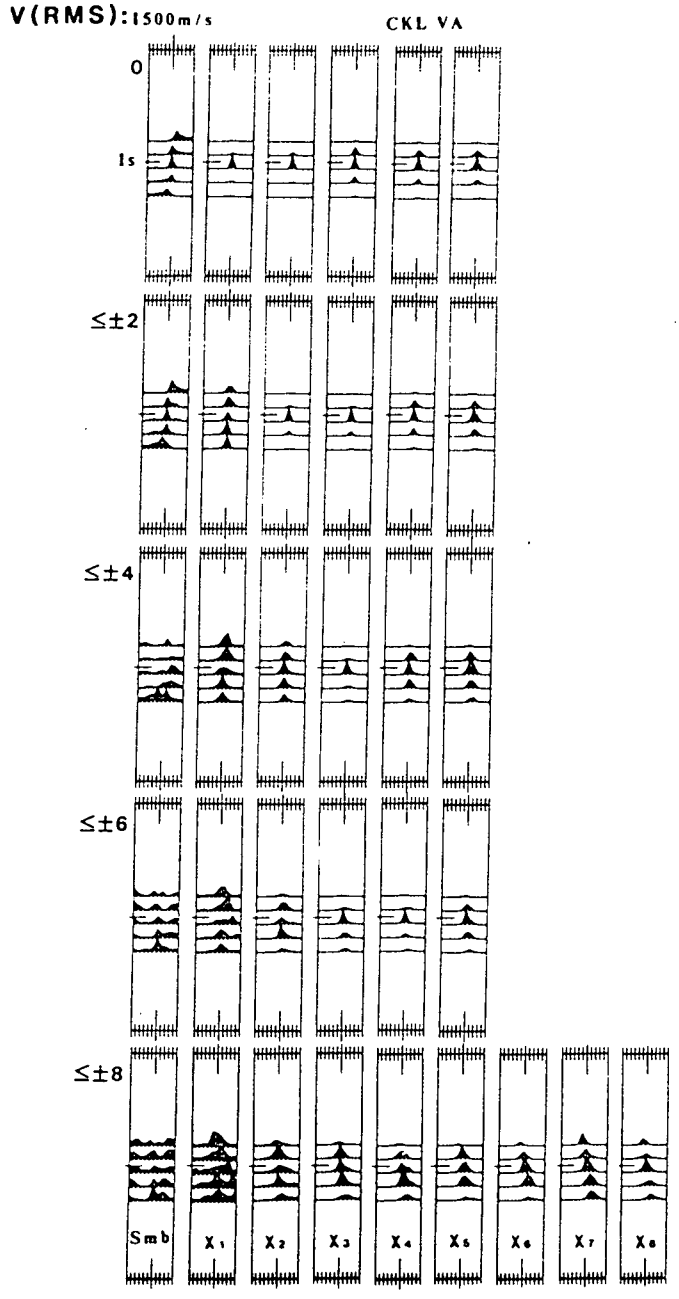


FIGURE 4.12

Left column: Semblance criterion VA for progressively increasing static time shifts; down the column: up to 0, ±2, ±4, ±6, ±8 sample points (i.e. 0, ±8, ±16, ±24, ±32 ms).
Second left to far right: CKL eigenvalue ratio criterion; columns for $\chi(1)$, $\chi(2)$, ..., $\chi(8)$.
As the static scatter increases, the semblance VA breaks down and is unable to reliably locate the peak corresponding to the arrival. For the CKL VA, we note how a progressing increase in m, in $\chi(m)$, counteracts the degrading effect of an increase in static scatter. The velocity tick marks are separated by 50 m/s.

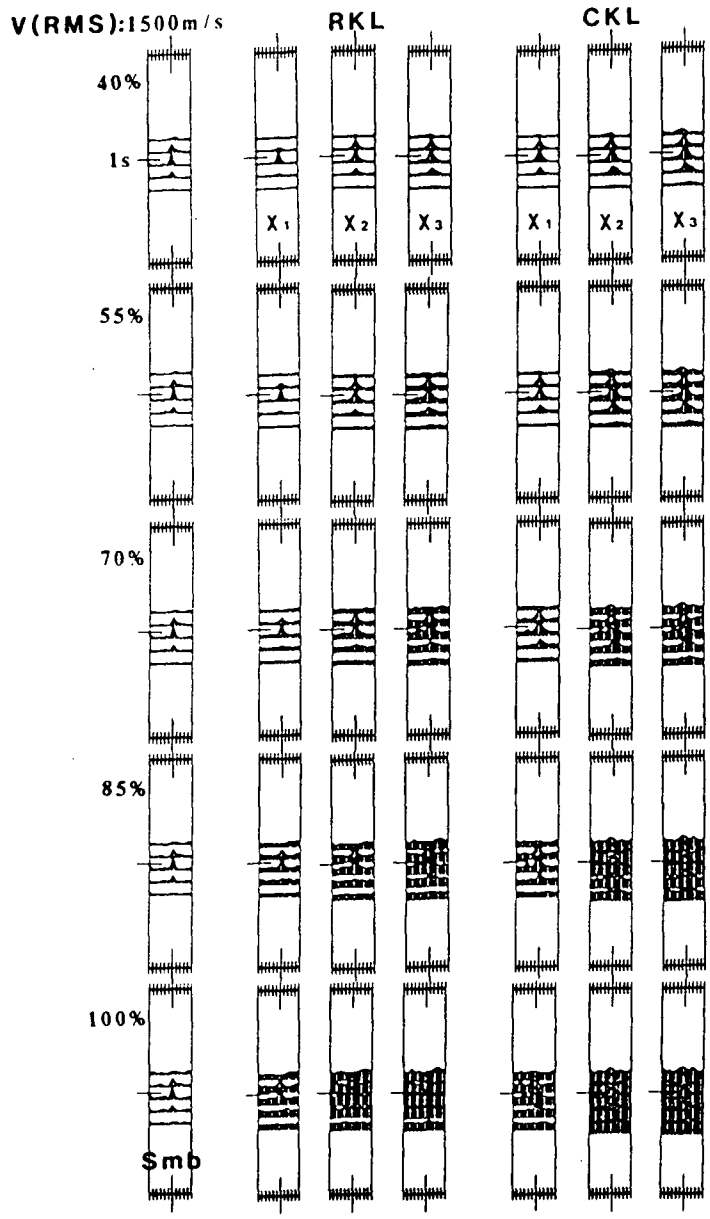


FIGURE 4.13

The effect of increasing random noise levels (down the columns: 40, 55, 70, 85, and 100% of the maximum amplitude) on semblance (far left) and KL VA's. Centre are the RKL results, far right are the CKL results. An increase in m in $\chi(m)$ does not improve the result for added random noise. This is because white noise adds energy to the diagonal of the covariance matrix uniformally, degrading the similarity measure for all m (see text). As the noise level increases, the semblance VA results are able to reliably locate the maximum, but, due to the increase in energy all along the covariance diagonal, the KL VA maxima are lost in the background noise. The velocity tick marks are at 50 m/s.

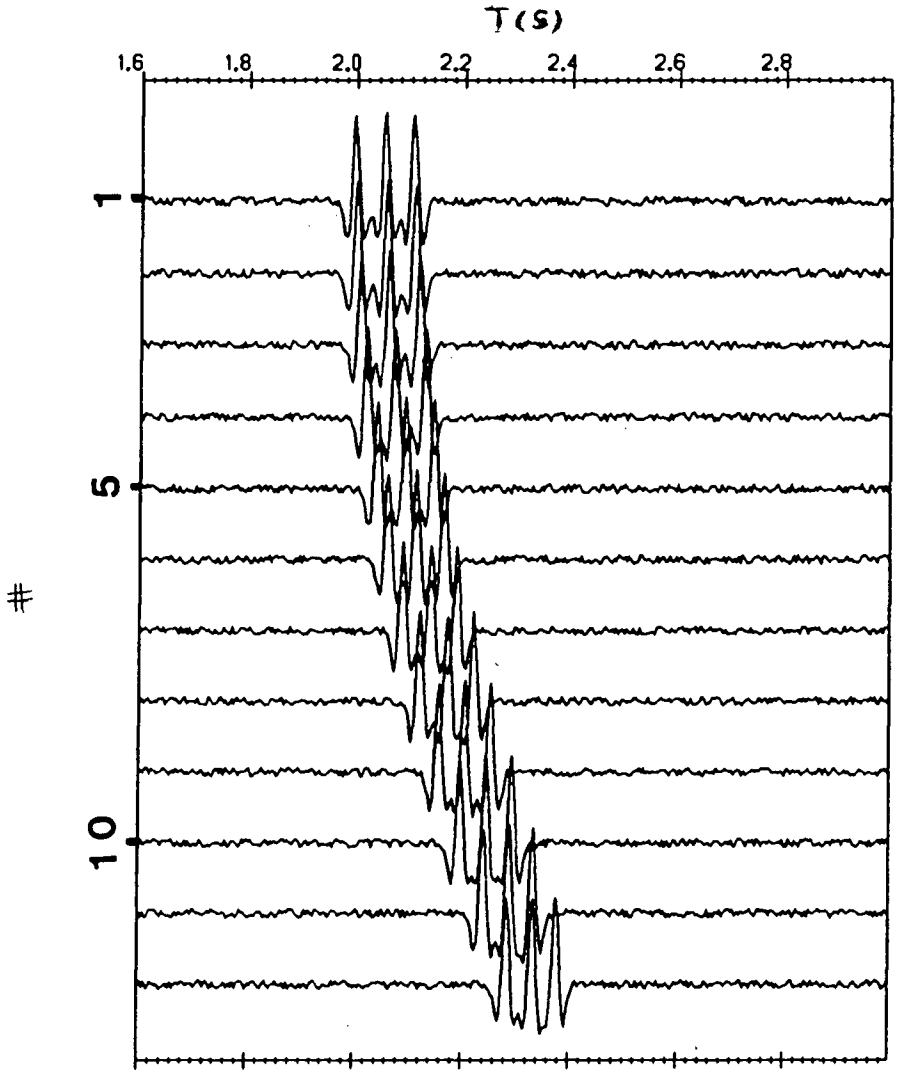


FIGURE 4.14

Twelve synthetic seismograms representing reflections from three closely spaced interfaces (50 m apart). Offsets range from 0 to 2000 m.

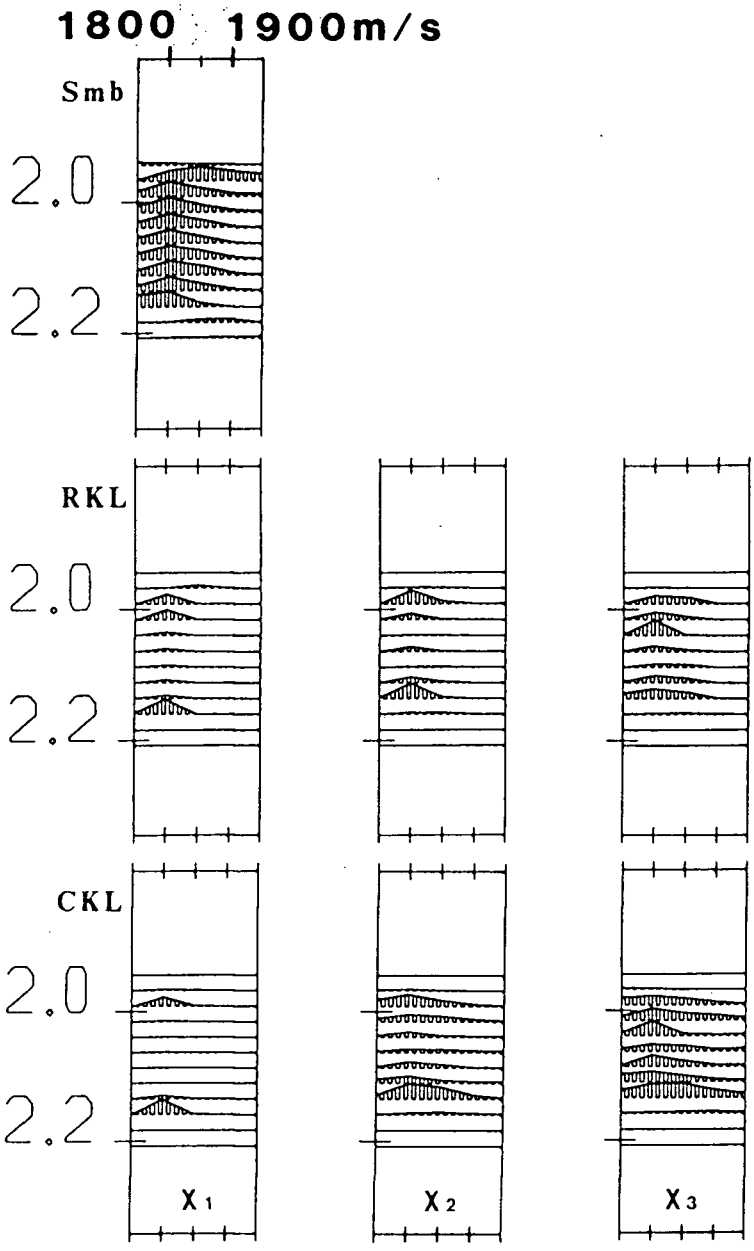
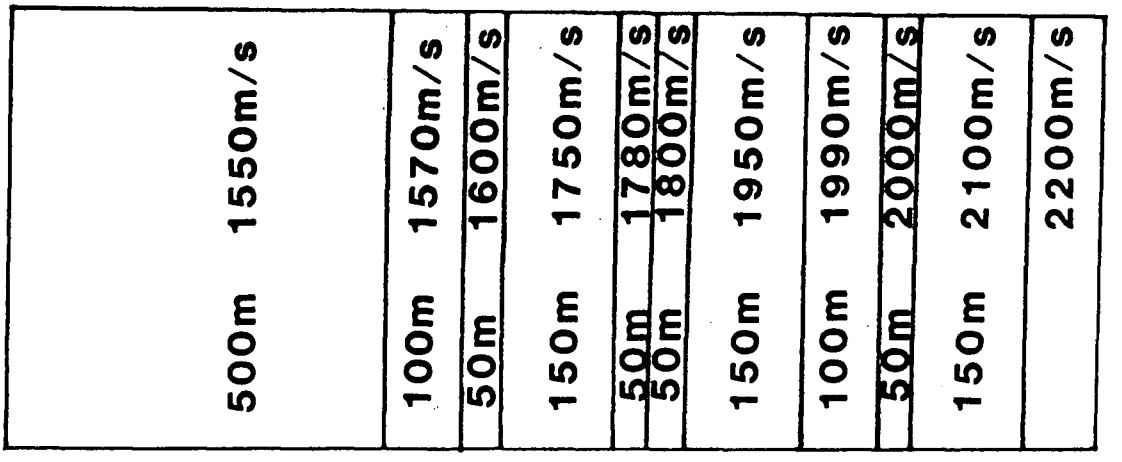


FIGURE 4.15

VA results for the three reflection events. The semblance VA smears the events and is unable to resolve them. However, both KL VA's are well able to distinguish the outer two events, and the RKL $\chi(2)$ and CKL $\chi(3)$ VA's indicate the possibility of a third central event. In both these VA's, the outer event appears at about 2.125 s, which is a bit late. True τ and V_{RMS} values are: 1800 m/s at 2.0 s, 1801 m/s at 2.06, and 1804 m/s at 2.11 s. The velocity tick marks are at 50 m/s.



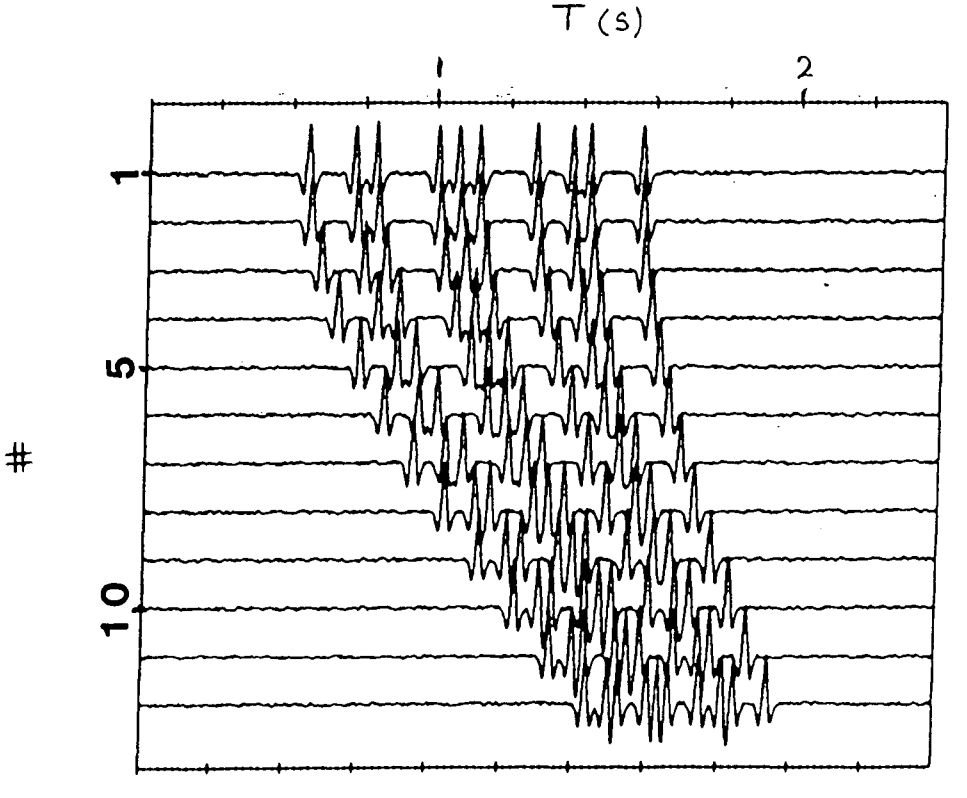


FIGURE 4.16
Twelve synthetic seismograms (a) representing reflections from ten flat layers, as shown in (b). Offsets range from 0 to 2000 m.

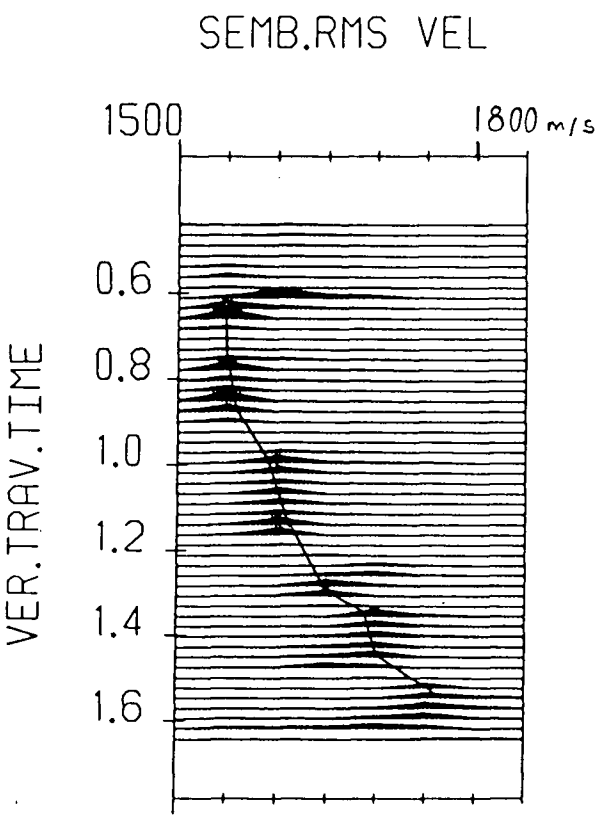


FIGURE 4.17a

For the data of Figure 4.16, the semblance VA (a) gives a fairly consistent picture of smeared arrivals, however, the stacking velocity trajectory can readily be picked. The solid line is the true τ-V_{RMS} locus. The velocity tick marks are at 50 m/s.

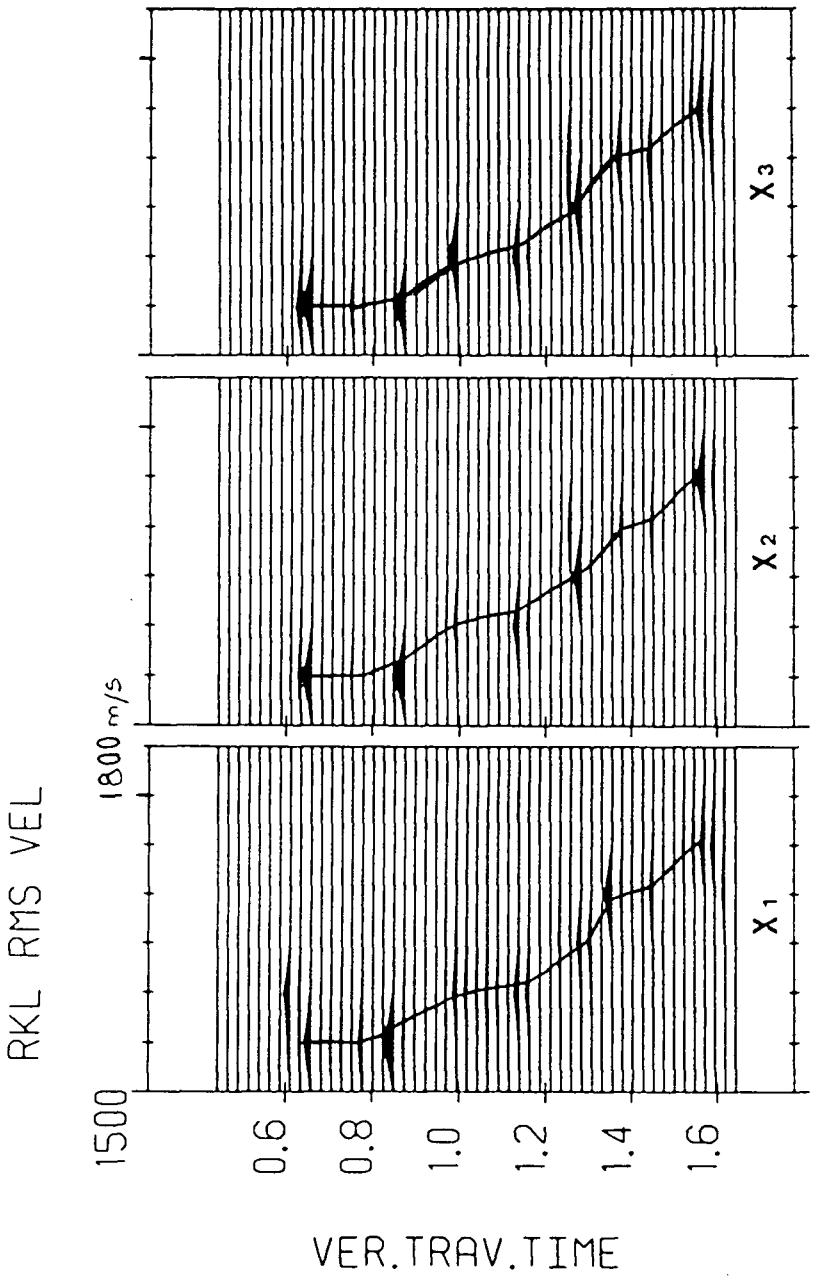


FIGURE 4.17b

The RKL method (for $\chi(1)$, $\chi(2)$, and $\chi(3)$) gives a much sharper definition of events, and all except $\chi(1)$ have given a clear delineation of the very first event (at 0.6 s, and 1550 m/s), whereas the semblance VA gives a spurious value of 1600 m/s for this arrival. However, the semblance VA gives a clear indication of the presence of the event at 1.05 s, whereas the RKL VA does not. The solid line is the true τ-V_{RMS} locus. The velocity tick marks are at 50 m/s.

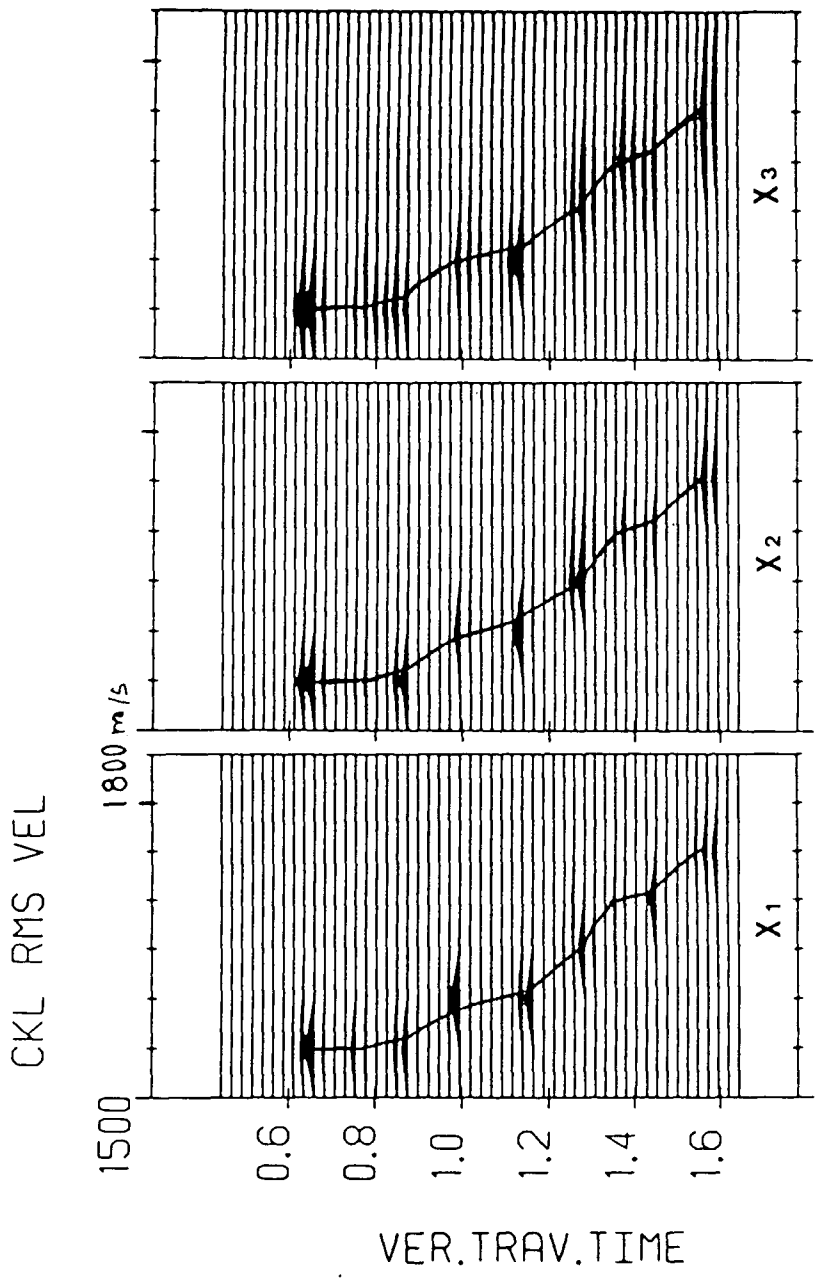


FIGURE 4.17c

The CKL method (for $\chi(1)$, $\chi(2)$, and $\chi(3)$) gives a much sharper definition of events, and all have given a clear delineation of the very first event (at 0.6 s, and 1550 m/s), whereas the semblance VA gives a spurious value of 1600 m/s for this arrival. However, the semblance VA gives a clear indication of the presence of the event at 1.05 s, whereas the CKL VA does not. The solid line is the true τ-V_{RMS} locus. The velocity tick marks are at 50 m/s.

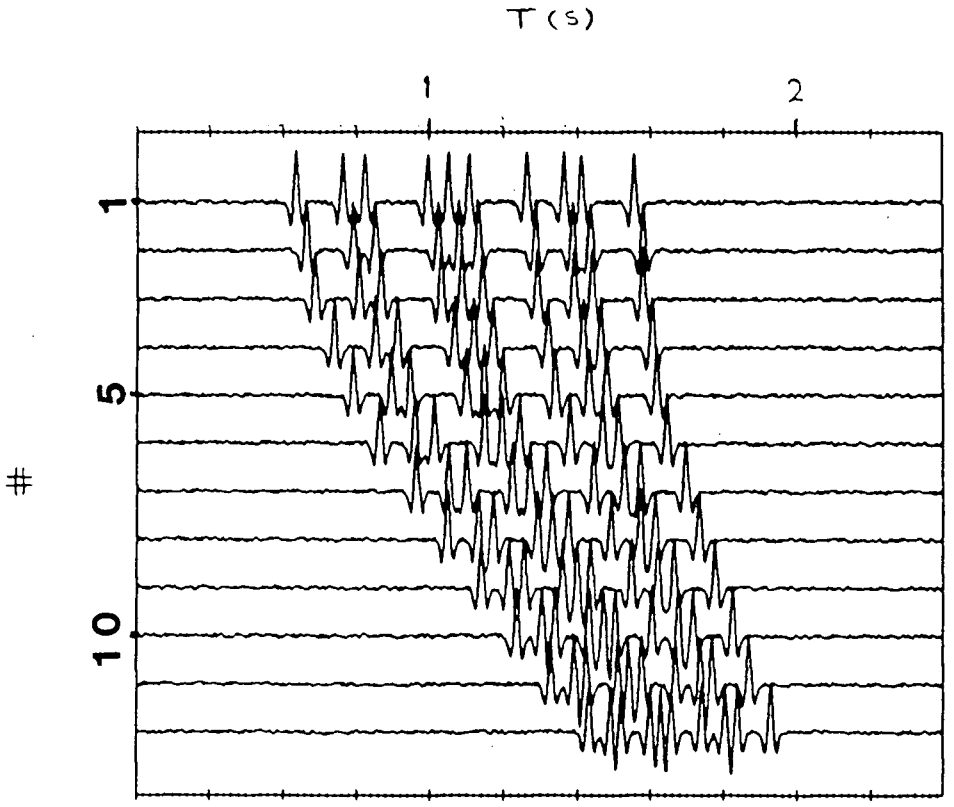


FIGURE 4.18

Twelve synthetic seismograms as per Figure 4.16a, but with time shifts of up to ± 12 ms introduced to each trace. Offsets range from 0 to 2000 m.

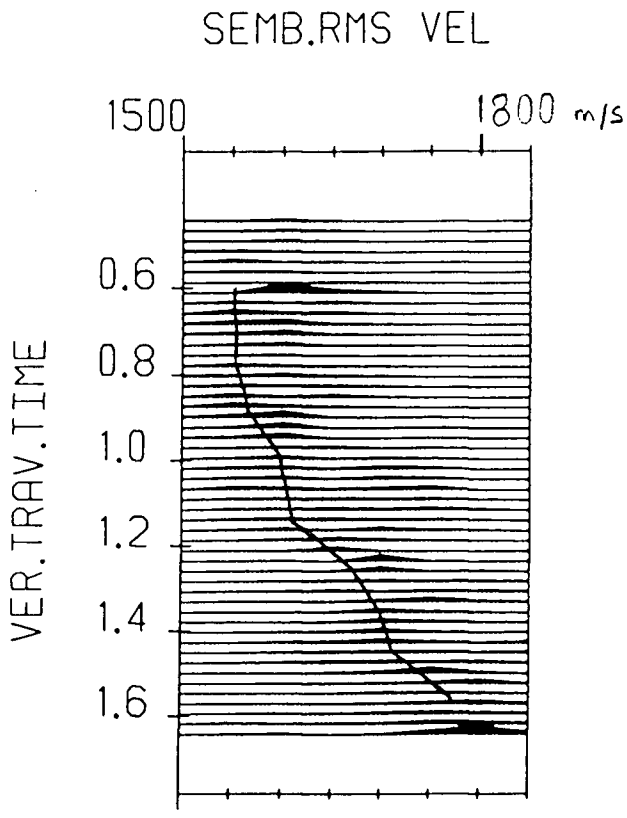


FIGURE 4.19a
In the case with static time shifts, the semblance VA of the data (Figure 4.18) performs poorly. The solid line is the true τ-V_{RMS} locus. The velocity tick marks are at 50 m/s.

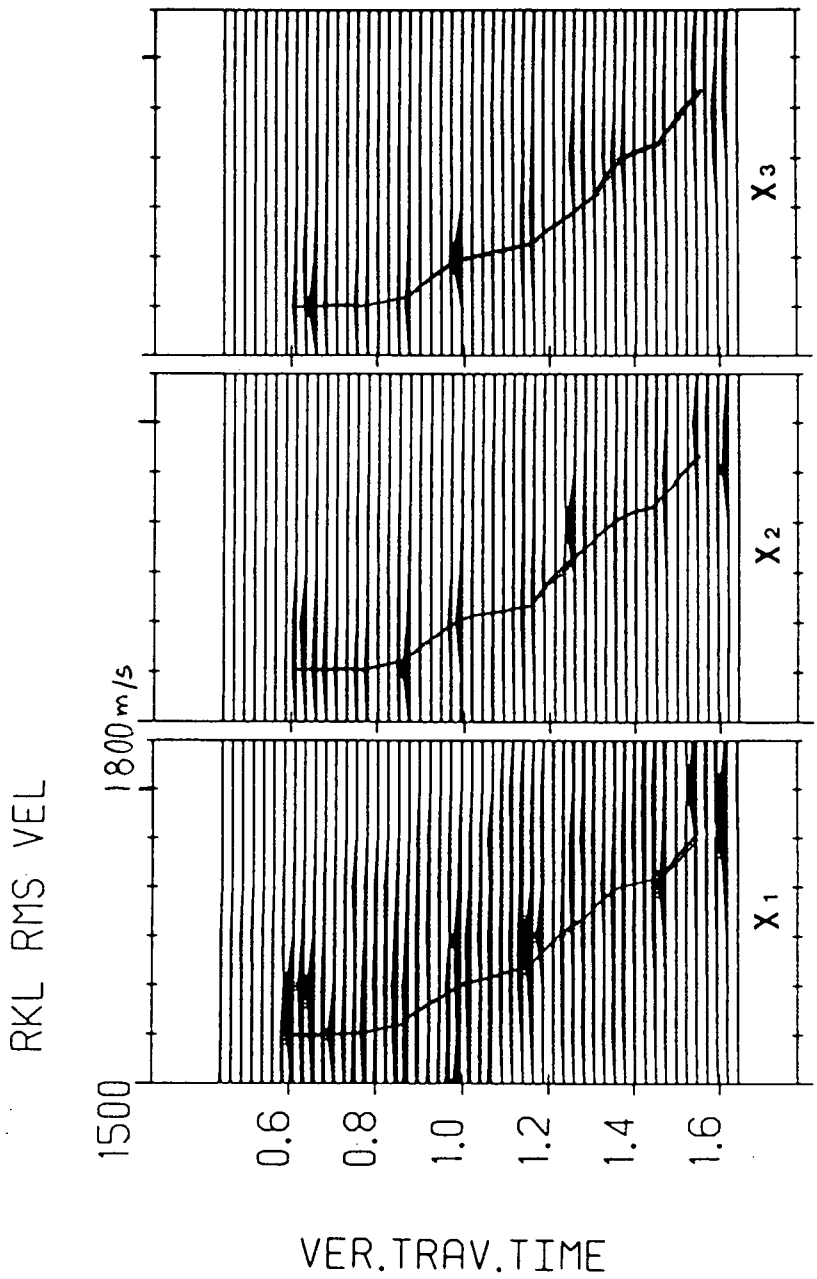


FIGURE 4.19b

For the RKL VA of the data (Figure 4.18), the VA maps for $\chi(2)$ and $\chi(3)$ are able to resolve the majority of the reflection events. The solid line is the true τ-V_{RMS} locus. The velocity tick marks are at 50 m/s.

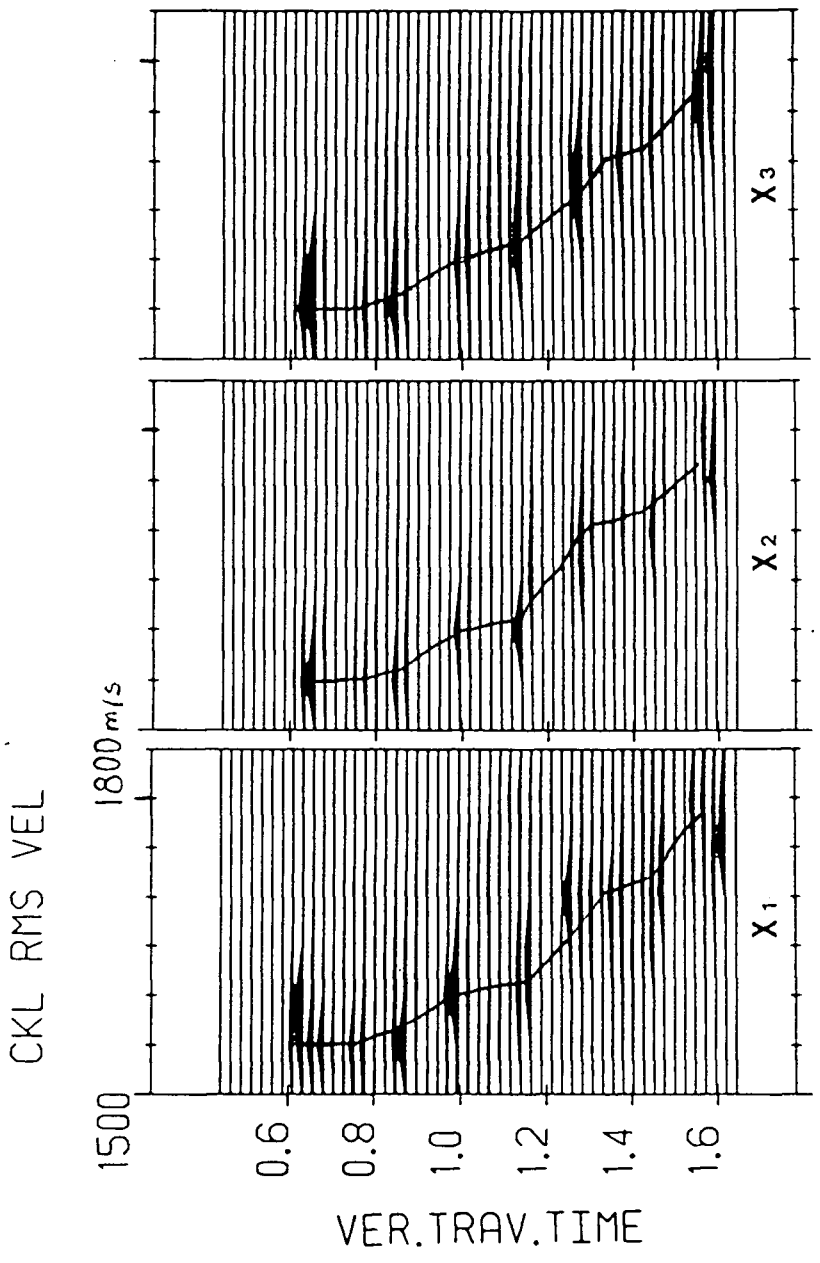


FIGURE 4.19c

For the CKL VA of the data (Figure 4.18), all three VA maps, are able to resolve the majority of the reflection events. The solid line is the true τ-V_{RMS} locus. The velocity tick marks are at 50 m/s.

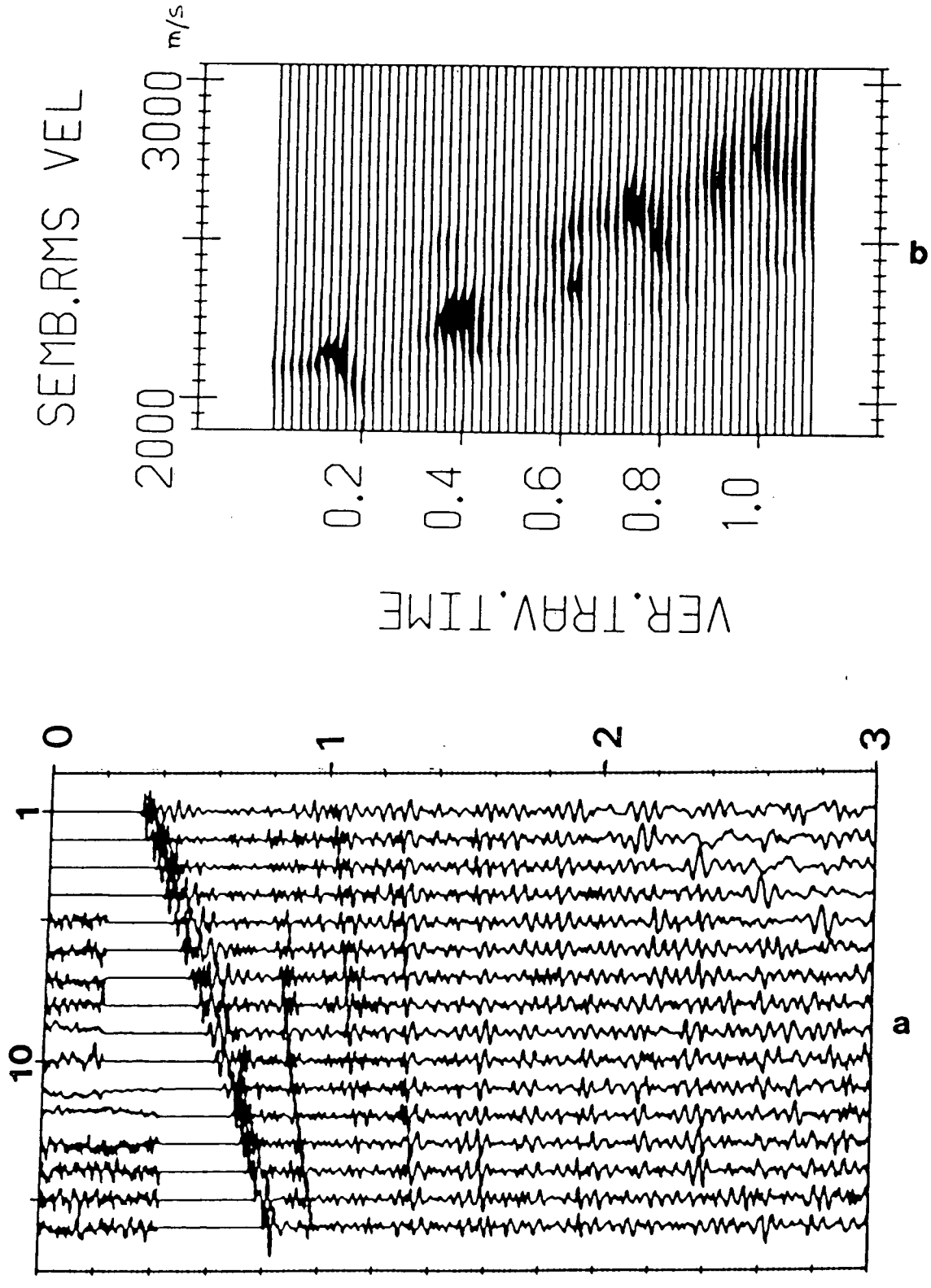


FIGURE 4.20

(a). Three seconds of 2 ms reflection seismic data for 16 offsets (582 to 1662 m) after bandpass filtering (8 to 55 Hz), mute, and application of an AGC. The basement is at about 1.62 s. (b). The semblance VA of this data: a stacking velocity locus can be picked for several points. The velocity tick marks are at 50 m/s.

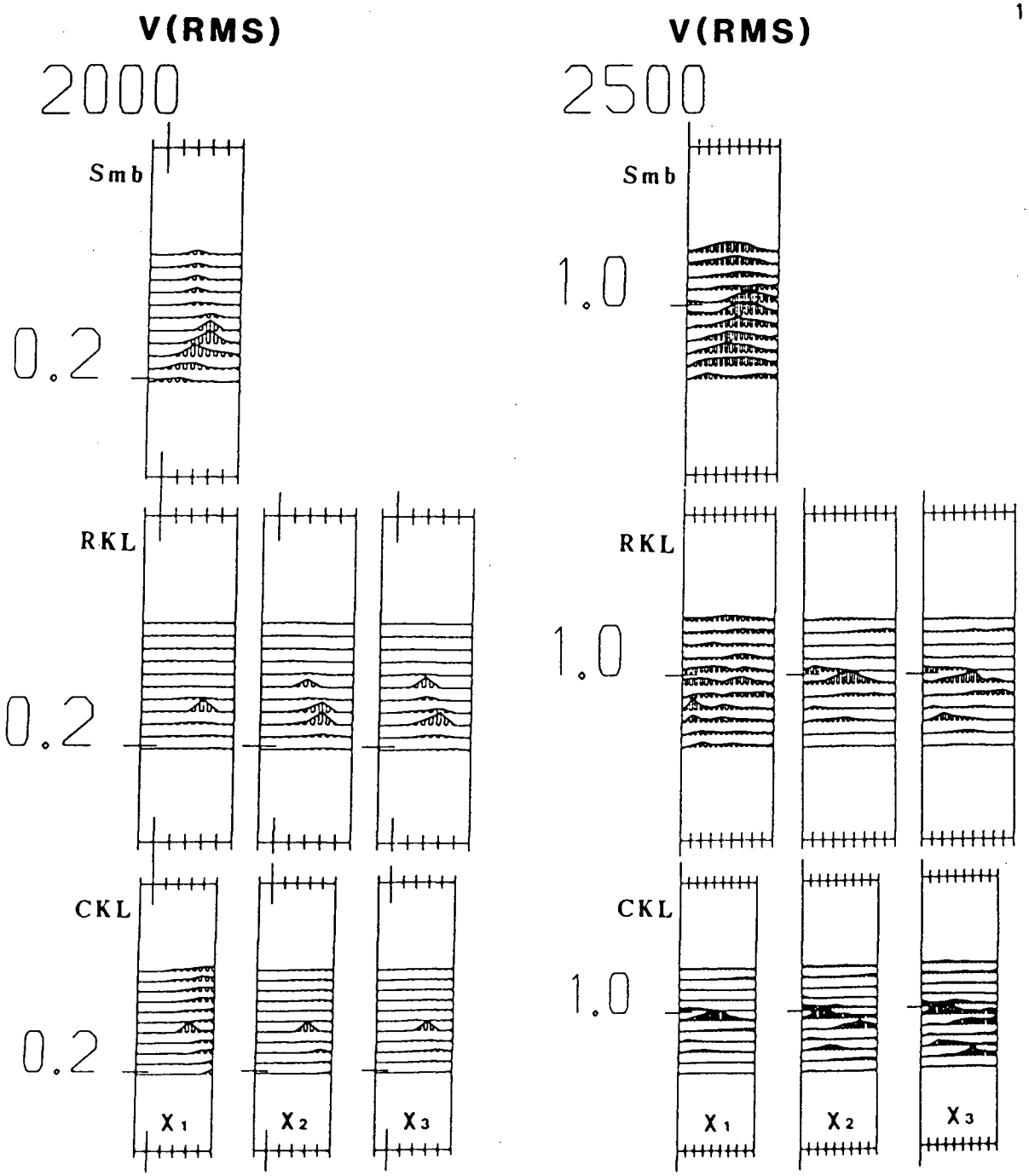


FIGURE 4.21

Sample VA maps for 0.2 s of data in two time windows. These results were used to determine a good choice of $\chi(m)$ to use in the KL VA's. At 0.2 s (left), the RKL VA for $\chi(2)$ and $\chi(3)$ picks-out two distinct events corresponding to a smeared equivalent in the semblance VA map. The CKL VA only delineates one event. For the window at 1.0 s (right), RKL $\chi(2)$ and CKL $\chi(1)$ delineate a single event, whereas the semblance is again smeared. The velocity tick marks are at 50 m/s.

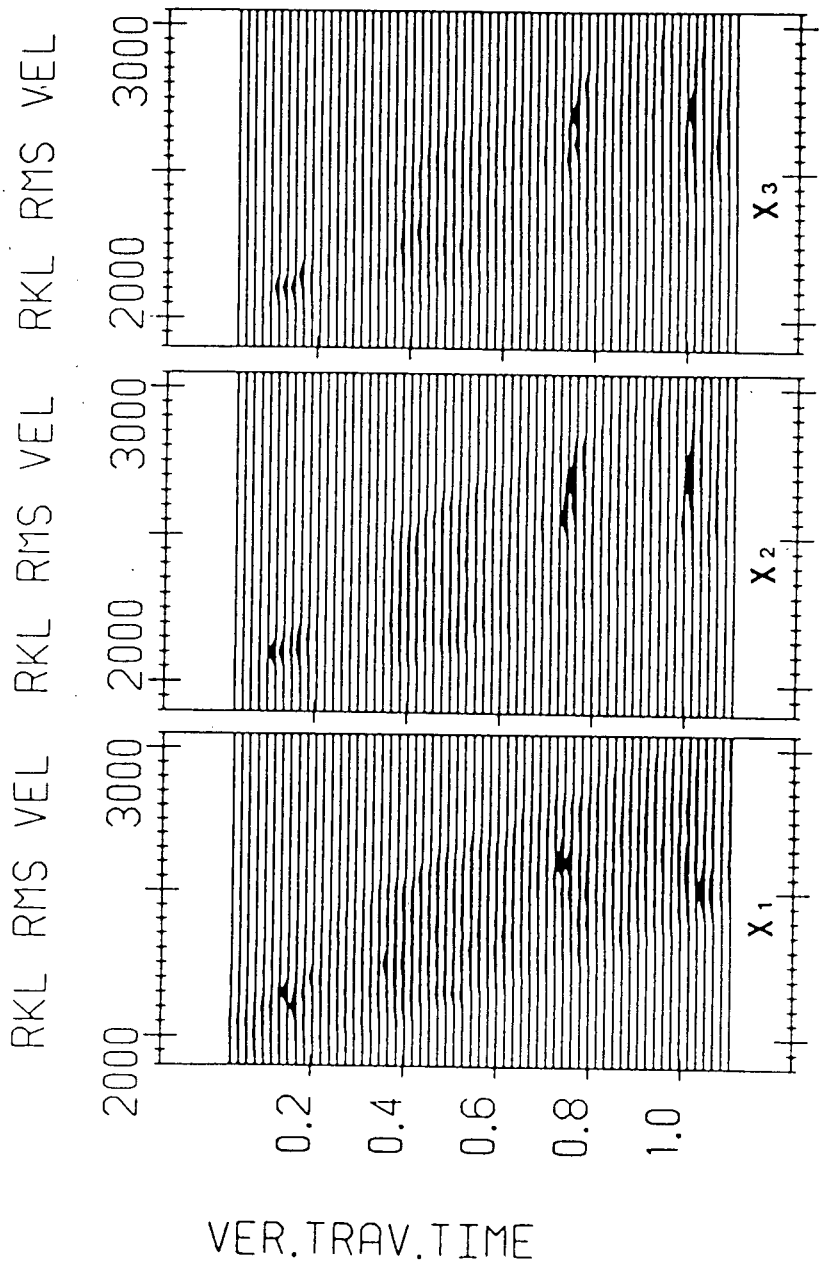


FIGURE 4.22

RKL results for $\chi(1)$, $\chi(2)$, and $\chi(3)$ for the data seen in Figure 4.20. Although individual events are sharper than those of the semblance VA map, the general picture is less convincing. Events are well resolved at early times, but by 0.4 s, events have started to disappear. The $\chi(3)$ map gives a clear pick at 1 s, but the preceding events seen in the semblance VA map at 0.95 s, is absent.

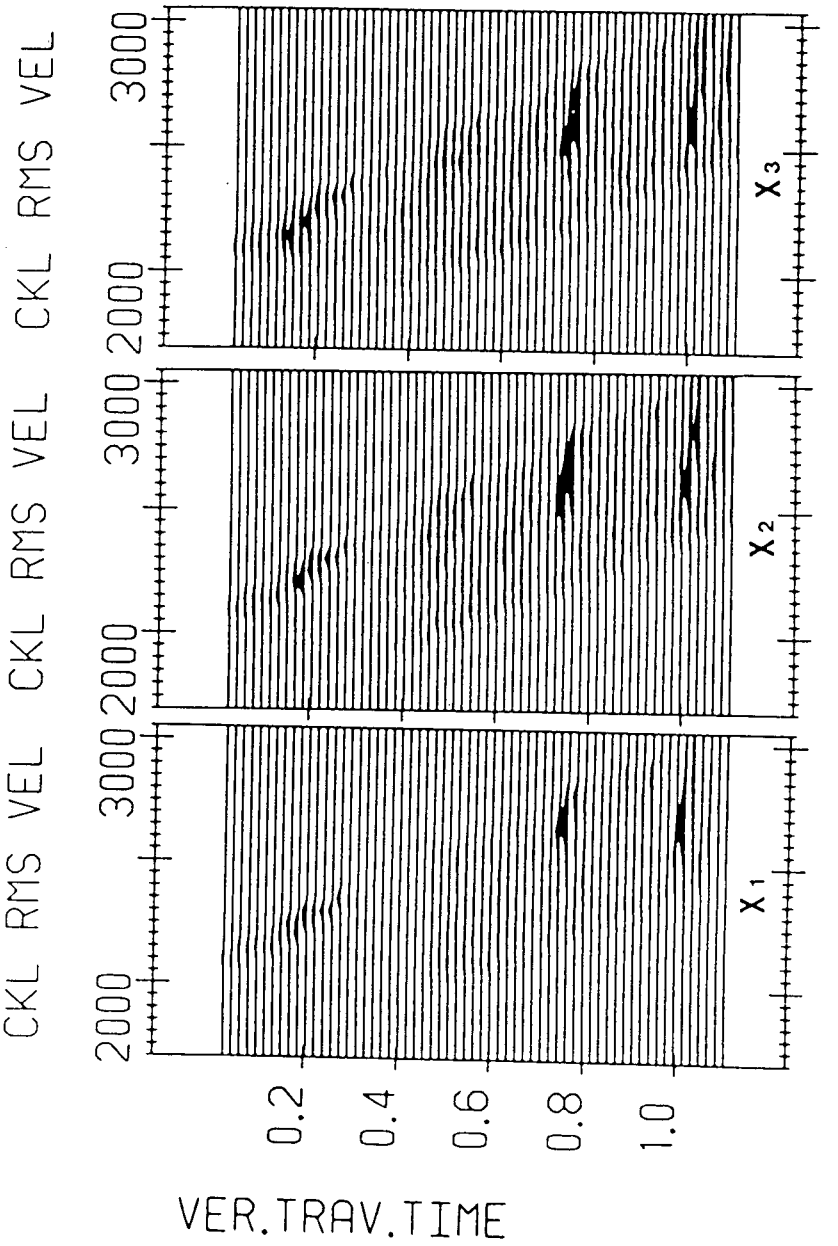
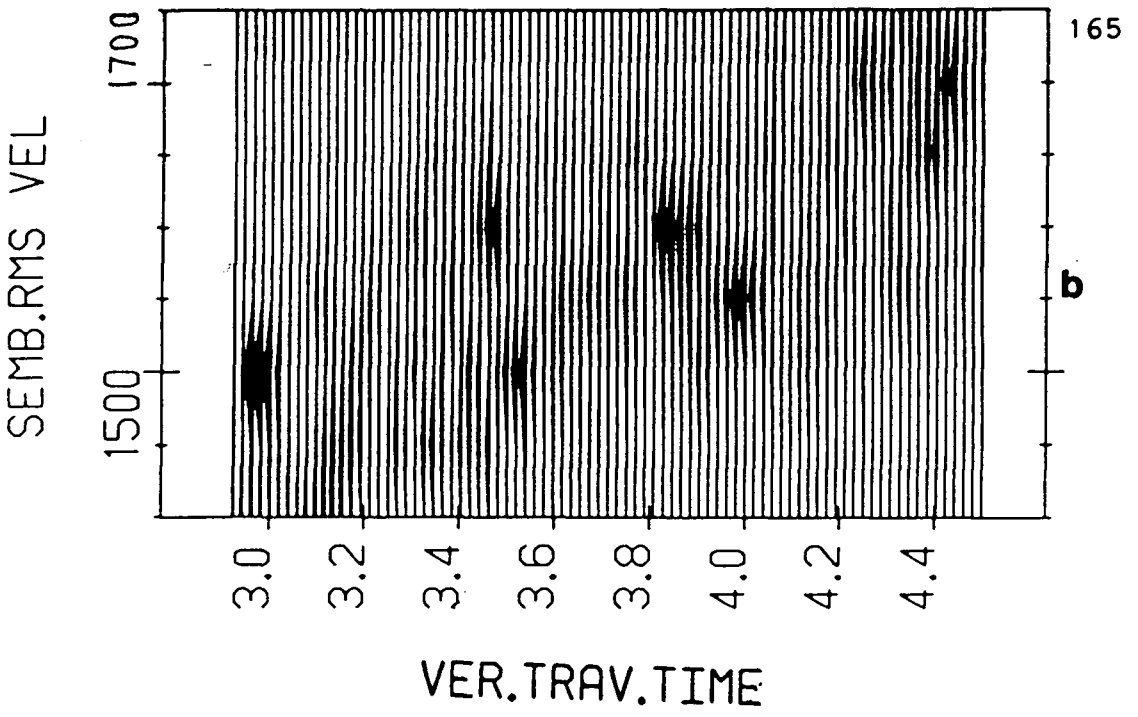


FIGURE 4.23

CKL results for $\chi(1)$, $\chi(2)$, and $\chi(3)$ for the data of Figure 4.20. As with the RKL VA results, these VA maps are poorly balanced. Resolution at early times is good, but deteriorates markedly past 0.3 s. Overall, these results seen less plausible than those of the semblance VA map. The velocity tick marks are at 50 m/s.



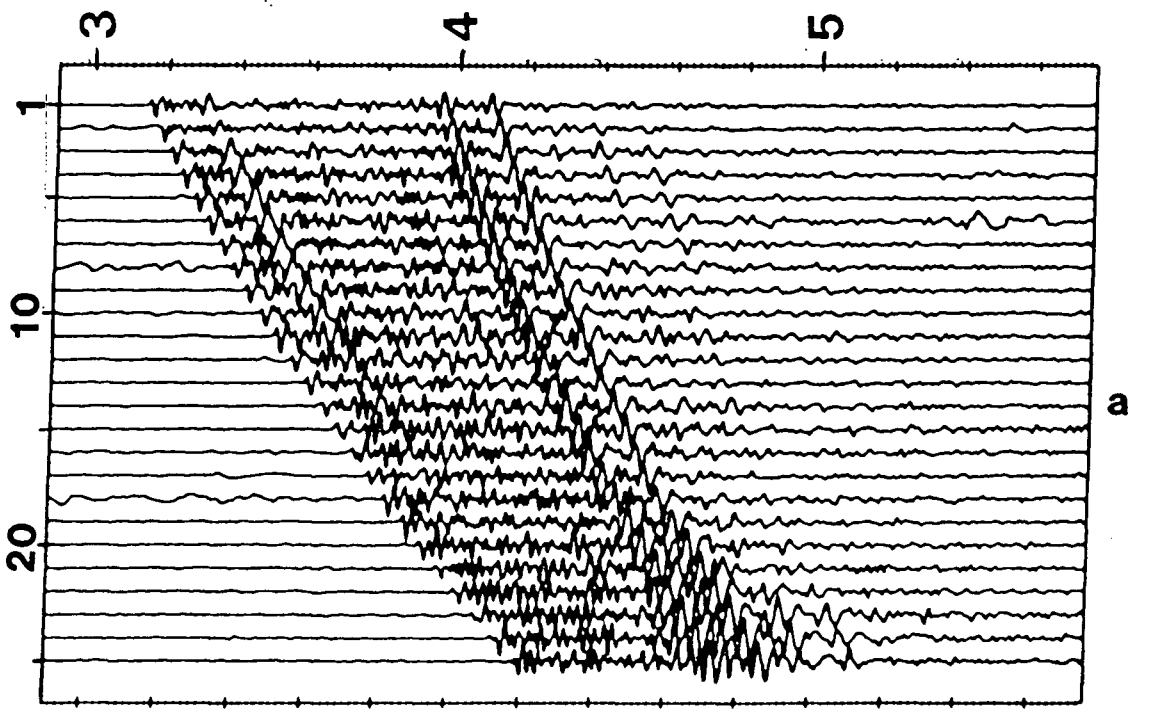


FIGURE 4.24

(a). A sample of the wide-angle seismic reflection data collected during the UBC-RECOPE (Costa Rican National Oil Co.) joint seismic project (1982). Offsets range from 1630 to 4415 m. (b). Results of a semblance VA. Although smeared, picks can readily be made for three primary arrivals.

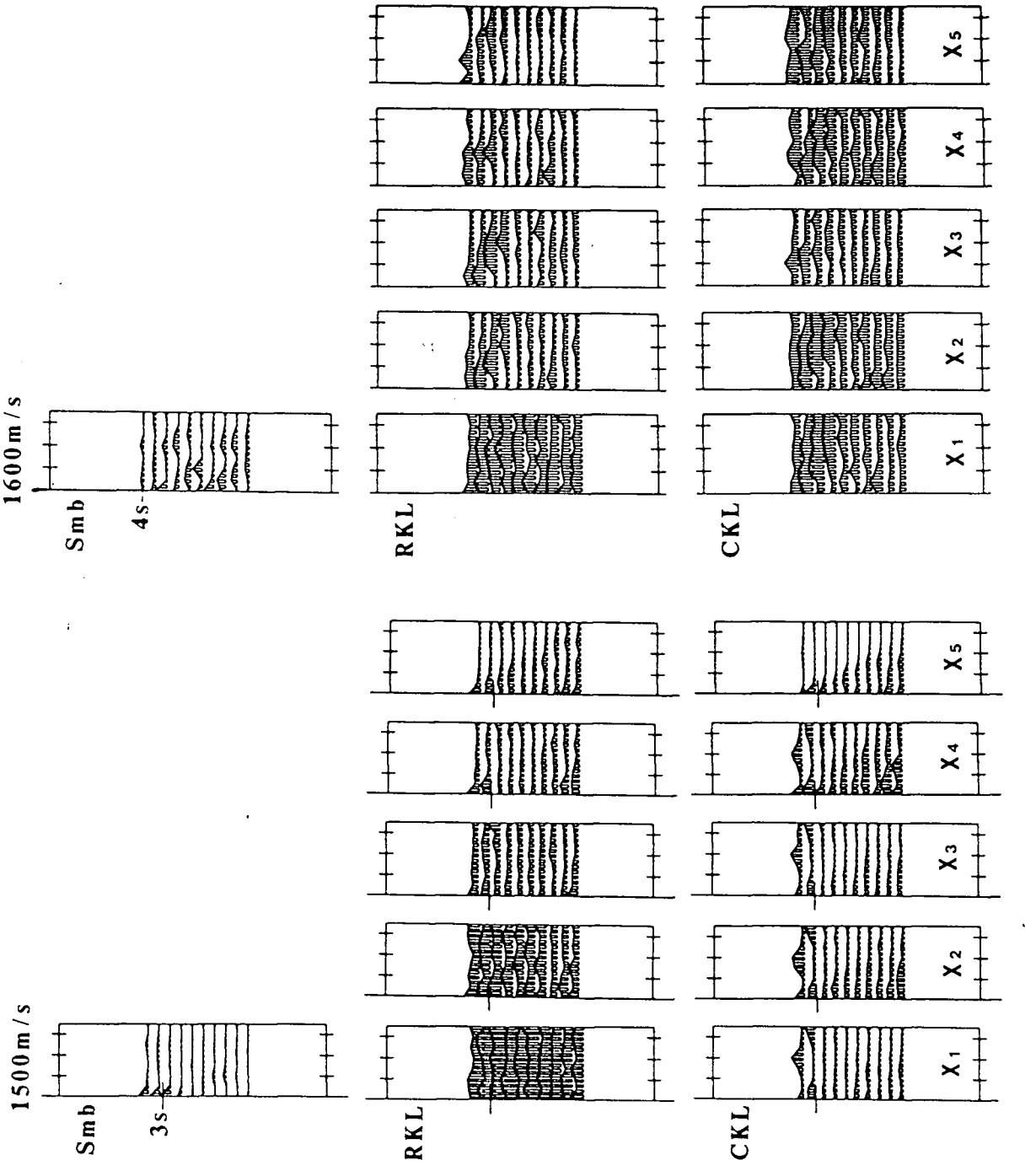


FIGURE 4.25

Two sample time windows to assist in choosing an appropriate $\chi(m)$ criterion for the KL VA of the data. The results at 3 s (left) suggested using $\chi(4)$ for both the RKL and CKL VA's. For the 4 s window (right), $\chi(3)$ and $\chi(4)$ look promising. However, in practice the $\chi(1)$ and $\chi(2)$ VA's proved to be of greater use. The velocity tick marks are at 50 m/s.

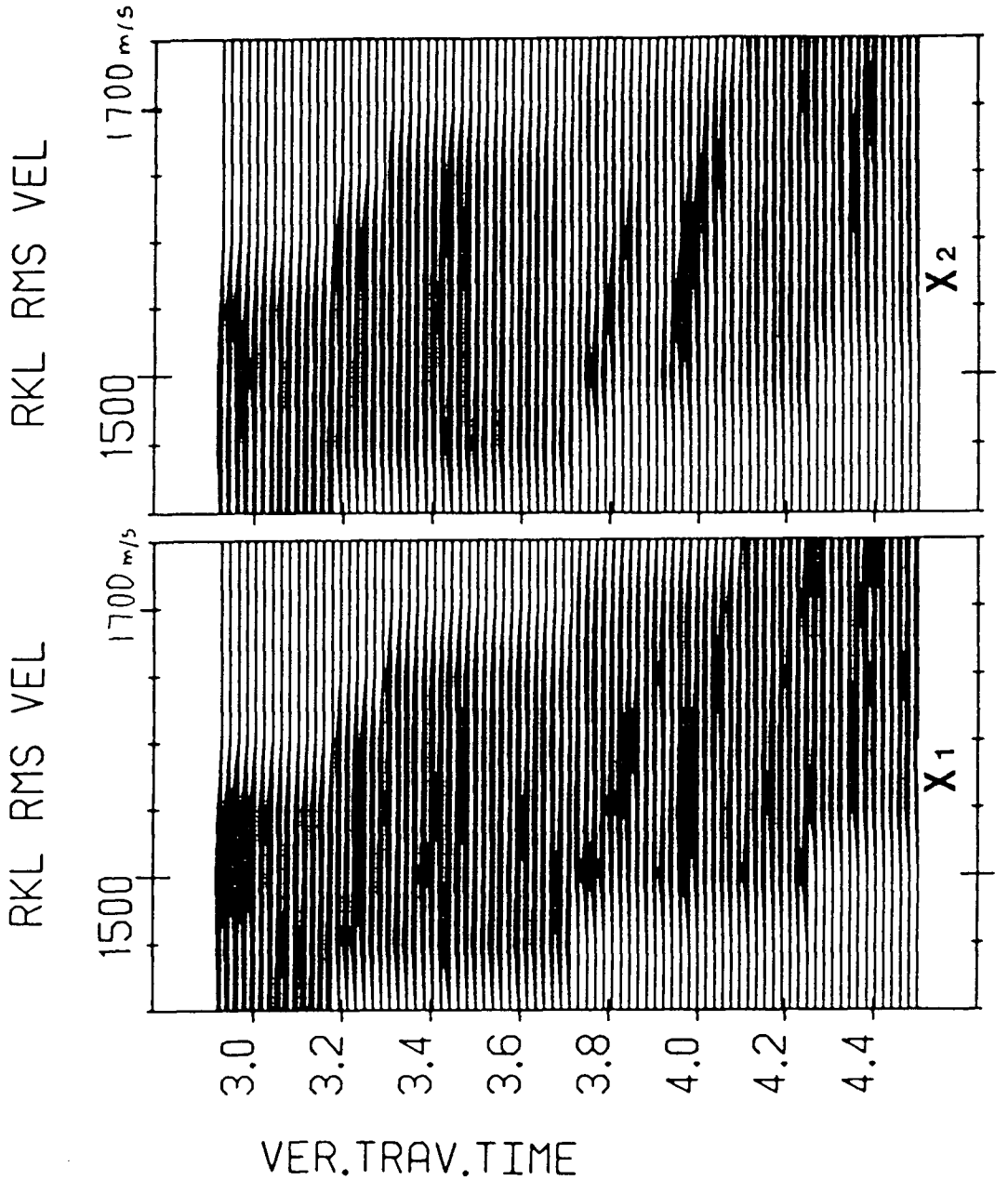


FIGURE 4.26

RKL VA results for $\chi(1)$ and $\chi(2)$. After much scrutiny, picks for five RMS velocities can be made: an improvement over the semblance VA map (see text). However, the overall appearance of these maps is not helpful to the picking. The velocity tick marks are at 50 m/s.

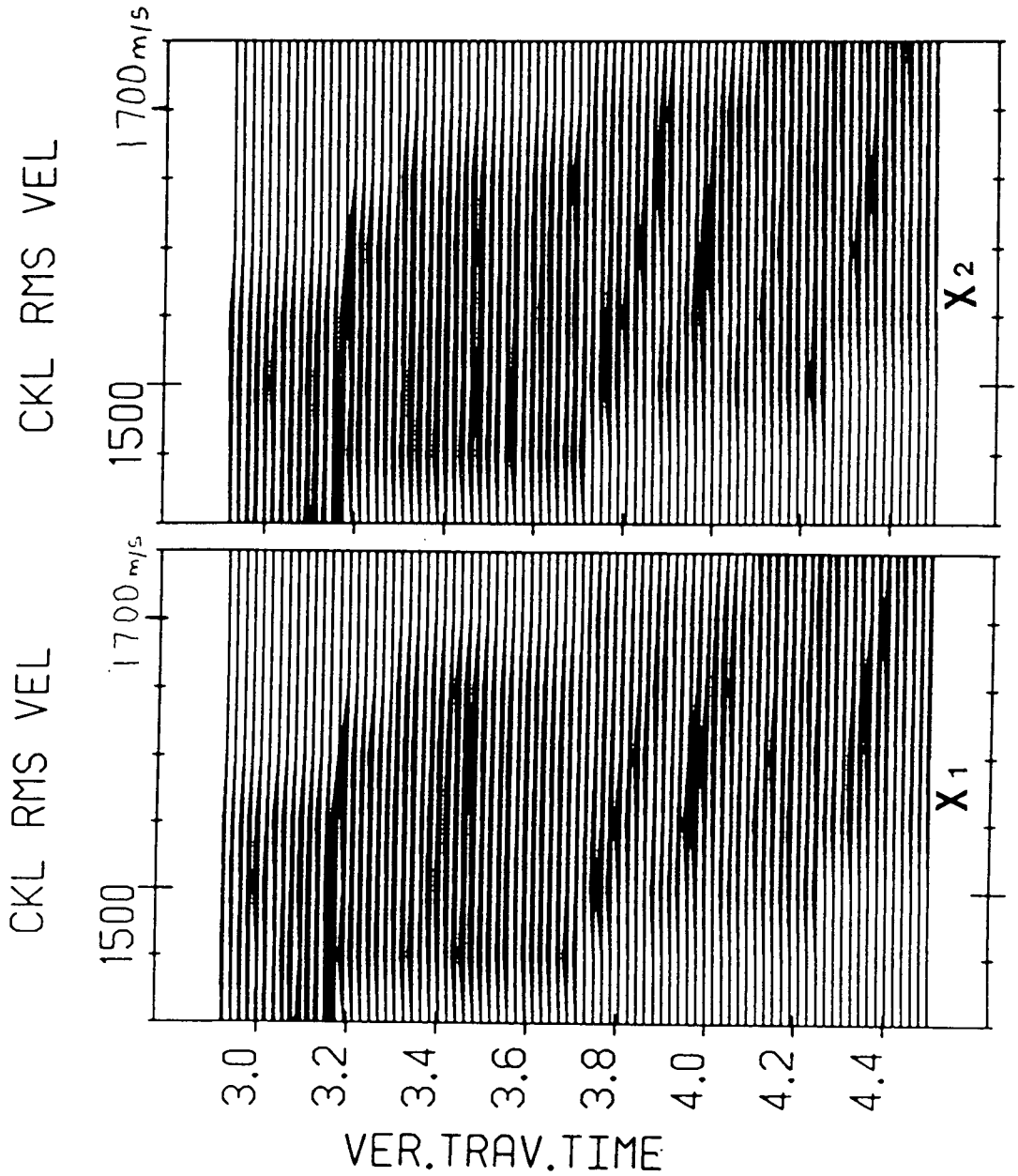


FIGURE 4.27

CKL VA results for $\chi(1)$ and $\chi(2)$. Six RMS velocities can be made from these maps (see text). The velocity tick marks are at 50 m/s.

SECTION 4.III: Q INVERSION.

a. Introduction.

Both Robinson (1979, 1982) and Beresford-Smith and Mason (1980) developed techniques to facilitate the removal of dispersive effects from seismic reflection data. Robinson (1979) utilised the theoretical developments of Futtermann (1962) to produce synthetic seismograms for dispersed data, and proceeded to show how the effects of dispersion may be removed from both synthetic and real data for an homogeneous medium with a known Q value (Appendix 1). Robinson (1982) extended the work in his 1979 paper to remove the effects of dispersion in vertically inhomogeneous media (i.e. Q varying with depth). However, he assumed that the Q structure of the reflection section was known a priori. Beresford-Smith and Mason (1980) relied on a measure of the degree of dispersion to estimate the dispersive properties of the medium. Subsequent minimization of this measure in an iterative procedure was used to remove the correct amount of distortion from each section of the data. Both these works attacked the problem directly and should prove most useful in the treatment of existing conventionally collected seismic reflection data.

Ganley and Kanasewich (1979), rather than attempting to remove the dispersive effects from the data, sought an estimate

of the Q structure for the seismic reflection section. They assumed that the input signal and the dispersion relation (again a constant Q model) were known. Using the complex spectral ratio technique, they divided-out the non-dispersed input signal from the dispersed data to obtain the cumulative, and thence interval, Q values.

Attention was also directed to this problem at the Stanford Q Conference (1980), and by Smith and Neidel (1976), Toksoz and Johnson (1981), McMechan and Yedlin (1981), and Rafipour (1981).

b. An Iterative Dispersion Removal Scheme Using an Effective-Q Value

Following the approach of Robinson (1979) outlined in Appendix 1, the constant Q model was employed to undisperse a segment of data using a specific Q value. Here I diverge from his approach and repeat the procedure for a suite of Q values, at each iteration computing a similarity measure $\chi(1)$, between a reference wavelet and the undispersed signal, or between the members of a suite of undispersed signals. Inspection of the function of undispersing-Q-value versus $\chi(1)$ should yield a maximum when the signals have been undispersed with the Q value with which they were originally dispersed (by the earth). An equivalent and alternative approach would be to disperse the reference signal prior to comparison with the dispersed data.

For wavelets with large travel times, the expense of Fourier transforming many long time segments poses a problem. I introduce here an approximation to facilitate adequate representation of a pulse at a large time T_L which travelled through a medium with quality factor Q_0, with a pulse located at some small time T_s. I effect this approximation by introducing an effective-Q value, Q_{eff}. For a wavelet of duration less than T_w centred about arrival time T_L, I replace the time series with the same wavelet centred in a window of width T_w. After propagating for a time T_L, the waveform will suffer from a loss in amplitude given by:

$$A(t) = A_0 \exp(-\omega t / 2Q),$$

which is *equation (A1.1)*, where ω is angular frequency. To produce the same loss of energy, while inserting a small apparant arrival time T_s, where $T_s = T_w/2$, we must choose a lower 'effective' Q value, hence:

$$A(t) = A_0 \exp(-\omega T_s / 2Q_{eff})$$

so:

$$T_s / 2Q_{eff} = T_L / 2Q$$

or:

$$Q_{eff} = QT_w / 2T_L \qquad (37)$$

It is this 'effective' Q value that is used in most subsequent processing.

We may think of the reasoning behind this approximation as follows: for a constant Q model, a sinusoid loses a fixed proportion of energy per cycle travelled. Consequently, allowing a sinusoid to propagate for 100 cycles with Q = 50, incurs the same energy loss as allowing it to propagate for 50 cycles with Q = 25. I translate this into a shift in arrival time in conjunction with an effective Q value.

However, as noted in Appendix 1, the Futterman relation only holds for Q values down to 2π, due to the approximations involved. Consequently, if Q_{eff} falls below 2π for a given window, the length of the sample window is doubled (with the wavelet again centred) and Q_{eff} recalculated. The sample window length is increased in this way until $Q_{eff} \geq 4\pi$.

When a reliable reference signal is available, for example a clean near offset early arrival, then the subsequent, and more dispersed signals may be compared to it. When one has a suite of signals with similar travel paths (as along a moveout hyperbola for example), then the entire suite of signals may be undispersed and the suite as a whole subsequently compared to the reference. In this latter case, each wavelet in the suite is undispersed with its appropriate arrival time and effective Q value. The suite as a whole will be best undispersed, hence most similar, when the actual cumulative Q value has been used. Consequently, the Q structure derived in this way would be absolute. If the horizontal component of the travel paths varied greatly, the net

effect would be to average any lateral changes in Q.

c. The Constant Phase Approximation.

The iterative undispersion scheme may be used with high quality data in an attempt to recover the cumulative (thence interval) Q structure of the data. However, when the object of the exercise is simply to alleviate the effects of dispersion, we may seek a simpler and more cost effective method.

Here, I use the complex KL method to determine whether, and under what conditions, dispersion can be approximated by a constant phase shift, viz. $x_i(t)=Re[\hat{x}_i(t)exp(i\epsilon)]$, where $x_i(t)$ is a measured (and dispersed) seismogram, and $\hat{x}_i(t)$ is the complex trace of the underlying signal (Levy and Oldenburg, 1982). If applicability can be determined the first eigenvector will be used to estimate the shift angle ϵ which will rotate the dispersed trace to its undispersed form.

The changes of ϵ over a seismic section would be indicative of the degree of distortion of the underlying signal. This distortion would be brought about by phenomena such as interference in thin layers, attenuation, dispersion, and out-of-plane scattering. A section depicting ϵ values would highlight the rapidity of change in the propagating waveform. With the appropriate quality of data, such a map of ϵ values could be used as a direct hydrocarbon indicator: a large ϵ value

being indicative of a large change in the signal. Such a change may be brought about by an increase in the dispersive properties of the medium, such as occurs in gas saturated media.

However, here I will simply demonstrate the applicability of the constant phase shift approximation to dispersion by considering a set of synthetic data examples and completing the following steps:

(1) select a time window which contains a dispersed pulse;

(2) calculate the envelope of the analytic signal for the dispersed pulse and reference pulse (initial wavelet), aligning the peaks of the envelopes so there is no time discrepancy;

(3) apply the complex KL transformation;

(4) evaluate $\chi(1)$ and ϵ (*equation (32)*).

If $\chi(1)$ is large, then the dispersed signal is approximately a phase shifted verson of the original. If $\chi(1) \simeq 1$ then the constant phase shift assumption is not valid.

The first series of examples covers the constant phase approximation. In each of the figures presented, I show (1) the reference wavelet (a Ricker wavelet of centre frequency 35 Hz, not dispersed); (2) the dispersed wavelet (which simulates an arrival from a specified time); (3) the real part of the first principal component; (4) the real part of the second principal component, principal components from the CKL decomposition of the two input waveforms, and (5) the dispersed waveform (as in (2))

rotated by the phase angle ϵ (from *equation (32)*). Figure 4.28 shows four panels, each as described above for a wavelet centred at 0.5 seconds dispersed with Q values 25, 55, 85, and 115 respectively. Figure 4.29 shows the analagous results for the case where the effects of attenuation are included. In all examples, the waveforms were aligned on the basis of complex trace envelope.

In general, we note that the CKL algorithm is well able to extract a signal common to the reference (1) and the dispersed wavelet (2), with a large value of $\chi(1)$. The results improve with increasing Q value, as the dispersed waveform becomes more like a purely phase rotated wavelet. Including the effects of attenuation degrades the result. The phase shift required to 'correct' for the dispersion also decreases with increasing Q value, as the dispersive distortion decreases.

The examples shown were for a Ricker wavelet of centre frequency 35 Hz. In Figure 4.30, I show the results of varying the centre frequency between 15 and 55 Hz, on the locus of $Ln(|\epsilon|)$ versus $Ln(Q_{eff})$. For a fixed window length, for wavelets with centre frequency above 10 Hz, I determined the following empirical relationship:

$$Ln(|\epsilon|) = 3.7 - 1.0\ Ln(Q_{eff}/T_w)$$

Here I have invoked my assumption that a shift in time can be duplicated by a change in Q, hence the relation involves only the

variable Q_{eff}, rather than time and Q explicitly. For the lowest centre frequency considered (15 Hz), the relationship is erratic below Q = 85; this is due to wavelet alignment problems (manually shifting the wavelet into alignment by one time sample moved most of the points onto the common locus). Also, the 'constant' for the intercept does change with f_c, but for the mid-range of centre frequencies is about 3.7 (the dependence of the intercept on f_c is discussed in some detail later). The results were consistent for the case with attenuation (Figure 4.31). Here, the wavelet alignment was more of a problem, and the locus jumped down to a subparallel trajectory for each sample point of misalignment. Changing the length of the wavelet window T_w, did not affect the slope or intercept of the curves.

The form of this relationship is not surprising, since it gives ϵ as being inversely proportional to the effective Q value, Q_{eff}, or:

$$|\epsilon| = C(2T_L/Q) \qquad (38)$$

where $C = exp(3.7) = 45$

This formulation also raises the possibility of estimating Q directly from the estimated phase shift ϵ.

To see the form of this relationship from an analytical viewpoint, I consider the problem in the frequency domain. For a wavelet $X(f)$, which was centred about time zero, and which can be

dispersed according to *equation (A1.9)* and *equation (A1.10)*, I assert that the dispersion can be represented by a single phase shift, ϵ. That is:

$$X(f)exp(i\epsilon) = |1 + Ln(f/f_b)/\pi Q|\ X(f[1 + Ln(f/f_b)/\pi Q])$$

where f_b is the reference frequency discussed in Appendix 1.

For an analytic signal, we consider only positive frequencies, and for a Ricker wavelet, the signal is real, other than for the phase component introduced by centring the wavelet at T_L in the time domain. Assuming that $X(f)$ is a Ricker wavelet (the second derivative of a Gaussian pulse, Ricker, 1953):

$$X(f) = (f/f_c)^2 exp(-(f/f_c)^2)$$

we have, for a wavelet centred about T_L:

$$(f/f_c)^2 exp(-(f/f_c)^2)\ exp(i\epsilon)\ exp(2\pi i f T_L) =$$

$$exp(2\pi i f T_L\ [1 + Ln(f/f_b)/\pi Q])\ |1 + Ln(f/f_b)/\pi Q|$$

$$(f/f_c)^2\ [1 + Ln(f/f_b)/\pi Q]^2\ exp(-(f/f_c)^2[1 + Ln(f/f_b)/\pi Q]^2)$$

Equating the real parts, and rearranging gives:

$$Cos(\epsilon) = Cos(2\pi f T_L[1 + Ln(f/f_b)/\pi Q] - 2\pi f T_L)\,|1 + Ln(f/f_b)/\pi Q|^3$$

$$exp([-(f/f_c)^2 Ln(f/f_b)/\pi Q]\ [2 + Ln(f/f_b)/\pi Q])$$

or:

$$Cos(\epsilon) = Cos(2T_L f Ln(f/f_b)/Q) \; |1 + Ln(f/f_b)/\pi Q|^3 \; exp([-(f/f_c)^2 Ln(f/f_b)/\pi Q] \; [2 + Ln(f/f_b)/\pi Q]) \qquad (39)$$

In the problem as posed here, I am extracting that part of the dispersed signal which is most similar to a non-dispersed Ricker wavelet. The ϵ value returned by the complex KLT simply tells us by how much we must rotate this signal to best resemble the dispersed waveform. A component of the dispersed waveform is discarded and appears on the second principal component. Hence, we should not expect the above equation to exactly reproduce the empirical result. However, by separating the phase distortive component of the dispersive process (that associated with the complex part of the signal) from the amplitude distortive component, we should be able to better mimic the phase relationship noted from *equation (38)*.

In this case, the phase related factors are those within the cosine terms, while the amplitude related terms are the latter two on the right hand side of *equation (39)*. Dropping the amplitude related terms and equating the arguments of the cosines gives:

$$\epsilon = (2T_L/Q) f Ln(f/f_b) \qquad (40)$$

which is of the same form as *equation (38)*.

The frequency dependent proportionality term in *equation (40)* may be estimated by comparison with that found in *equation (38)*. The dominant frequency in the band of interest is f_c, the centre frequency of the Ricker wavelet, and substituting this into *equation (40)* for f, we have:

$$C(f_c) = f_c Ln(f_c/f_b) \qquad (41)$$

The base frequency f_b in this case was fixed at the Nyquist frequency, so that all comparisons between wavelets with different centre frequencies would be consistent. In Figure 4.32, I show plots of the magnitude of the observed $C(f_c)$ values from *equation (38)* plotted against f_c, and superimposed on this the locus of $C(f_c)$ values calculated from *equation (40)*. The agreement is very good, however, as the bandwidth increases, the approximation degrades.

This result and that relating ϵ to Q_{eff} also hold true for a cosine-Gaussian wavelet (i.e. a cosine in a Gaussian envelope). Results similar to those shown for a Ricker wavelet are shown for this waveform in Figures 4.33 and 4.34 respectively. Problems of alignment persist here also, resulting in the deviation from a linear trend in Figure 4.33. However, as the band width of this signal is narrower than that of a Ricker wavelet, the fit of $C(f_c)$ to the theoretical curve is much better at high frequencies.

In conclusion, the constant phase approximation is summarized by:

$$\epsilon = 2T_L f_c Ln(f_c/f_b)/Q$$

where ϵ is the constant phase shift which best mimics the dispersive effect on a wavelet of dominant frequency f_c, travelling for T_L seconds in a medium with seismic quality factor Q.

d. Dispersion Quantification Objective Functions.

I deal next with an attempt to estimate Q using the iterative scheme described earlier. First I present the simple case of a known reference trace and a single dispersed pulse. All wavelets are aligned on the basis of complex trace envelopes prior to subsequent processing.

In the first set of figures (Figure 4.35), two similarity measures, or objective functions, are shown. The first is the eigenvalue ratio (EVR) i.e. $\chi(1)$; the second is the phase difference ϵ, between the reference trace and the undispersed wavelet, or a sum of the phase differences in the case of a suite of signals. If the signals were made to look identical by the undispersing procedure, then no phase difference would exist between them. However, when the signals are dispersed with respect to one another, the differences between them take on the appearance of a constant phase rotation. This rotation increases

with relative dispersion.

In Figure 4.35a, I show the two objective functions versus undispersing Q value. In this case, the wavelet (a 35 Hz Ricker at 1.0s) was originally dispersed with Q=30. As is seen, the EVR peaks at Q=30, and the ϵ value has a minimum there. In this simple case the algorithm has sucessfully located the Q value which optimally removes the effects of dispersion from the dispersed wavelet. In Figure 4.35b are shown the analagous results for an attenuated and dispersed wavelet: there is essentially no difference in the results. Adding random noise (20% of the maximum amplitude) to the attenuated and dispersed wavelet, severely degrades the EVR, but ϵ still has a well defined minimum at the correct Q value.

In the following trials, I use a suite of signals representing the members of a hyperbolic move-out curve. Each wavelet is undispersed using the Q_{eff} value (*equation (37)*) appropriate for its arrival time, and the members of the suite of signals compared to each other simultaneously. The similarity measure in this case will still be $\chi(1)$. However, when the data are members of moveout hyperbolae, certain conditions must be met for the method of comparisons within hyperbolae to work. Firstly, we must see a significant difference in the dispersion of each wavelet along a hyperbola. Otherwise all wavelets will be similar from the outset, and the method will fail. A wavelet to wavelet difference can be ensured by having sufficient moveout, such that

the individual travel times vary significantly. However, if the offset range is too great, the travel times in the individual earth layers vary greatly, and the cumulative Q value seen by wavelets propagating along ray paths to different receivers from a given reflector will also vary greatly. In this case the assumption that each ray path samples the same cumulative Q value breaks down, and the suite of wavelets picked from an individual moveout hyperbola cannot be used to estimate a common cumulative Q value. For the simple model used here (Figure 4.36a) offsets of up to 3 km produce little change in cumulative Q over the various ray paths for a given reflector. The synthetic seismograms in this example spanned 2 km, and the data were contaminated with 5% random noise.

In Figure 4.37, I show the objective function (the real KL EVR, $\chi(1)$) versus undispersing Q value for the first three reflecting horizons. In 4.37a we see a well defined peak at Q=30, which is the correct Q value for this layer. The maxima for the next two layers give cumulative Q values of 42 and 51, respectively (both these values are too high-see Table 1 later). After the third moveout hyperbola, the results degrade severely, as there is insufficient moveout to produce a noticeable difference in the travel times (hence appearance) of the waveforms. In Figure 4.38, I show the wavelets for the first reflecting horizon, from a sequence of undispersing Q values. As the arrival time of each wavelet within the window (picked from

the moveout hyperbola) increases slightly down the window, the processing will remove proportionally more dispersive effect for each successive wavelet. The correct Q value for the first horizon is 30. Consequently, the early panels (Q=20 and 25) are over corrected, and the later panels (Q= 35 and 45) are under corrected. At Q=30, the wavelets within the panel have all been restored to the form of a non-dispersed Ricker wavelet.

Of greater practical interest in this case is the result of picking one wavelet from each arrival time for a single trace. In this case I compute the relative dispersion between the first wavelet and each subsequent wavelet along the trace. This procedure is repeated for each trace in turn, and the results summed to estimate the cumulative Q values. This assumes that Q is not varying laterally over the region of the gather. The results for a comparison of the wavelets within each trace (Figure 4.36) are shown in Figure 4.39. The first comparison yields the interval Q value for the first layer, as the reflection event from the base of layer 1 is being compared to that from layer 2. Subsequent results give the cumulative Q value between the top of layer 2 and subsequent reflection horizons. This procedure was repeated for the case including attenuative effects. In this instance the objective functions are less distinct (Figure 4.40). Increasing the noise levels above 10% severely degraded these results, and for all examples, the complex KL results were essentially the same.

From *equation (A1.12)*, we derive that the interval Q value is related to the cumulative Q value by:

$$Q_j = t_j \{ (\sum_{i=1}^{j} t_i)/Qcum_j - (\sum_{i=1}^{j-1} t_i)/Qcum_{j-1} \}^{-1}$$

This inversion scheme holds for zero offset data, but was found to be adequate for the offsets considered here, as there was little change in the cumulative Q values. For the results based on the data shown in Figures 4.37 and 4.39, we note in Table 1 that:

TABLE 1

j	Layer Thickness	Interval Velocity	Q(True)	Q(Disp)		Q(Atten)	
				$Qcum$(est)	Q(est)	$Qcum$(est)	Q(est)
1	500	1700	30	--	30	--	30
2	500	1850	50	53	53	51	51
3	400	2000	80	62	80	66	109
4	500	2150	90	69	90	70	80
5	300	2350	100	78	280	75	124

Where the columns show: layer number, layer thickness (m), interval velocity (m/s), actual Q value used to generate the synthetic data, the cumulative Q values picked from the plotted results, and the Q value estimated from the inversion procedure (for the dispersed synthetic data), and finally, the cumulative Q values and estimated Q values for the attenuated and dispersed synthetic data. The Q value for the first layer was estimated

from the associated moveout hyperbola, whereas subsequent values were estimated from the down-trace comparisons.

The agreement, especially for the earlier arrivals is very good (the layer thickness and interval velocity are shown for interest). The interval travel times for use in the inversion scheme were picked from an RKL velocity analysis of the data, and not assumed a priori. For the dispersed-only synthetic data, the objective function estimates for all 12 traces were combined. However, for the attenuated and dispersed data only the first five traces were used (using more degraded the objective function).

e. Discussion.

The eigenvalue ratio derived from the complex or real KL transformation can be used to construct a sensitive similarity measure. Here I utilized this measure to ascertain when, and under what conditions, dispersive signal distortion can be approximated by a constant phase change. With simple synthetic examples, the constant phase approximation was shown to hold quite well, and a relationship between effective Q value, centre frequency, and representative phase shift was derived empirically and analytically.

Further, I used the measure to quantify the degree of success in removing dispersive effects from signals. In the case

of a synthetic multichannel common shot point gather, I estimated the cumulative and interval Q structure for a layered earth model. This was achieved by comparing waveforms along moveout hyperbolae and down traces within a loop iterating over Q values. Results for Q structure were degraded by including attenuative effects, and also by including random noise (above 10% by amplitude).

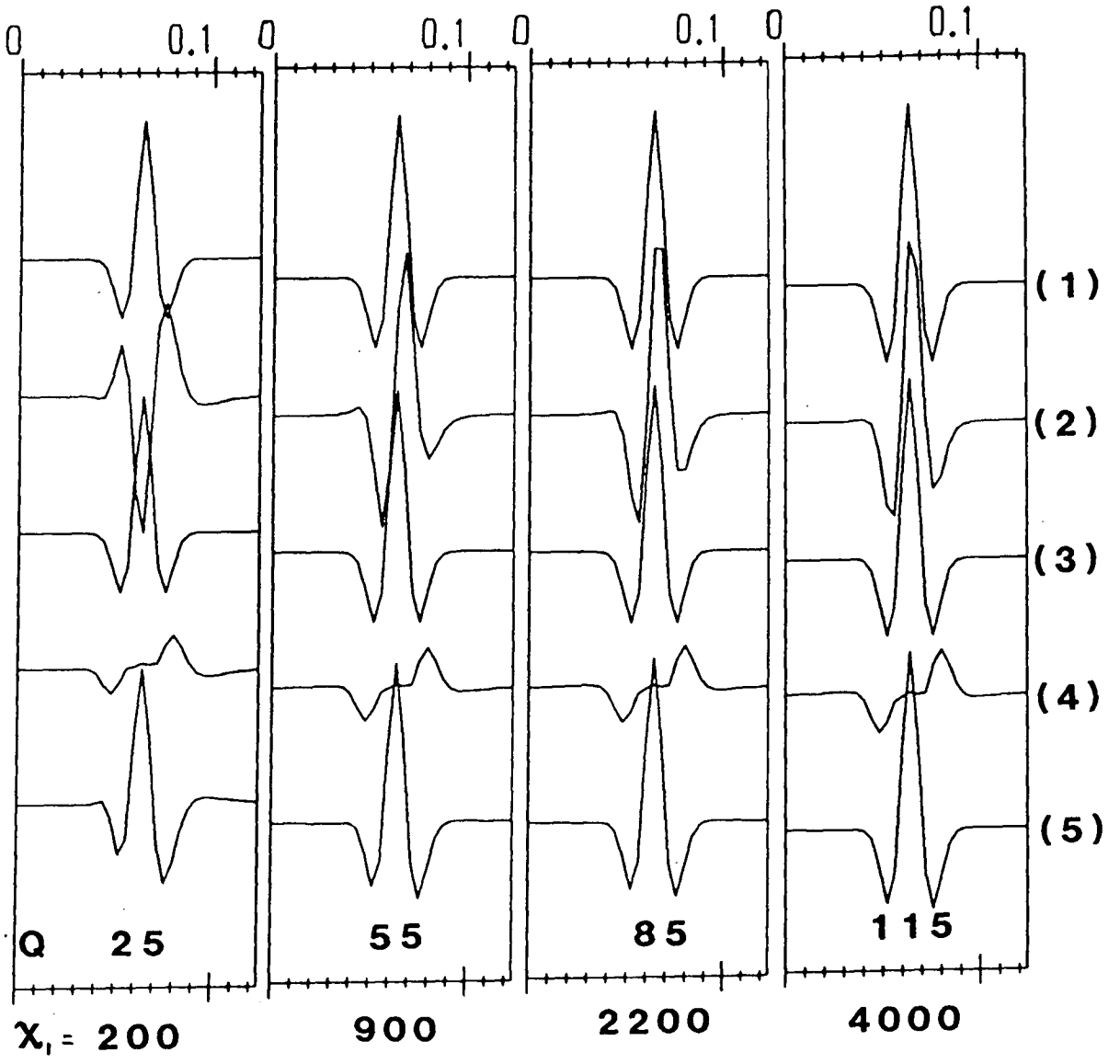


FIGURE 4.28

Each panel shows: (1) the reference wavelet, (2) the dispersed wavelet, (3) the real part of the CKL first principal component, (4) the real part of the CKL second principal component (with amplitude exaggerated for plotting), and (5) the dispersed waveform as in (2) rotated by ϵ. Panels are shown for Q= 25 to 115

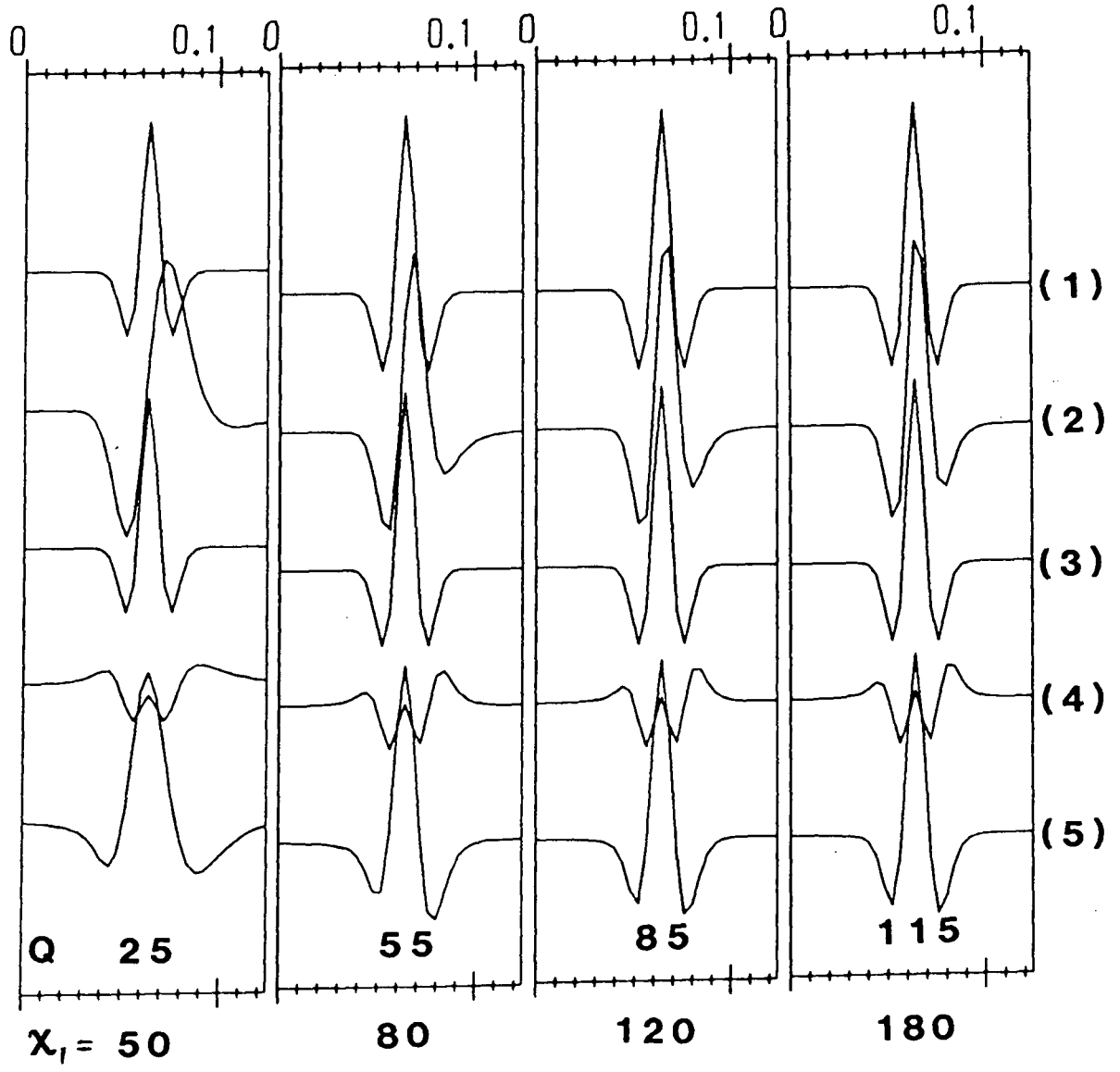


FIGURE 4.29

Each panel shows: (1) the reference wavelet, (2) the attenuated dispersed wavelet, (3) the real part of the CKL first principal component, (4) the real part of the CKL second principal component, (with amplitude exaggerated for plotting), and (5) the dispersed waveform as in (2) rotated by ϵ. Panels are shown for Q= 25 to 115

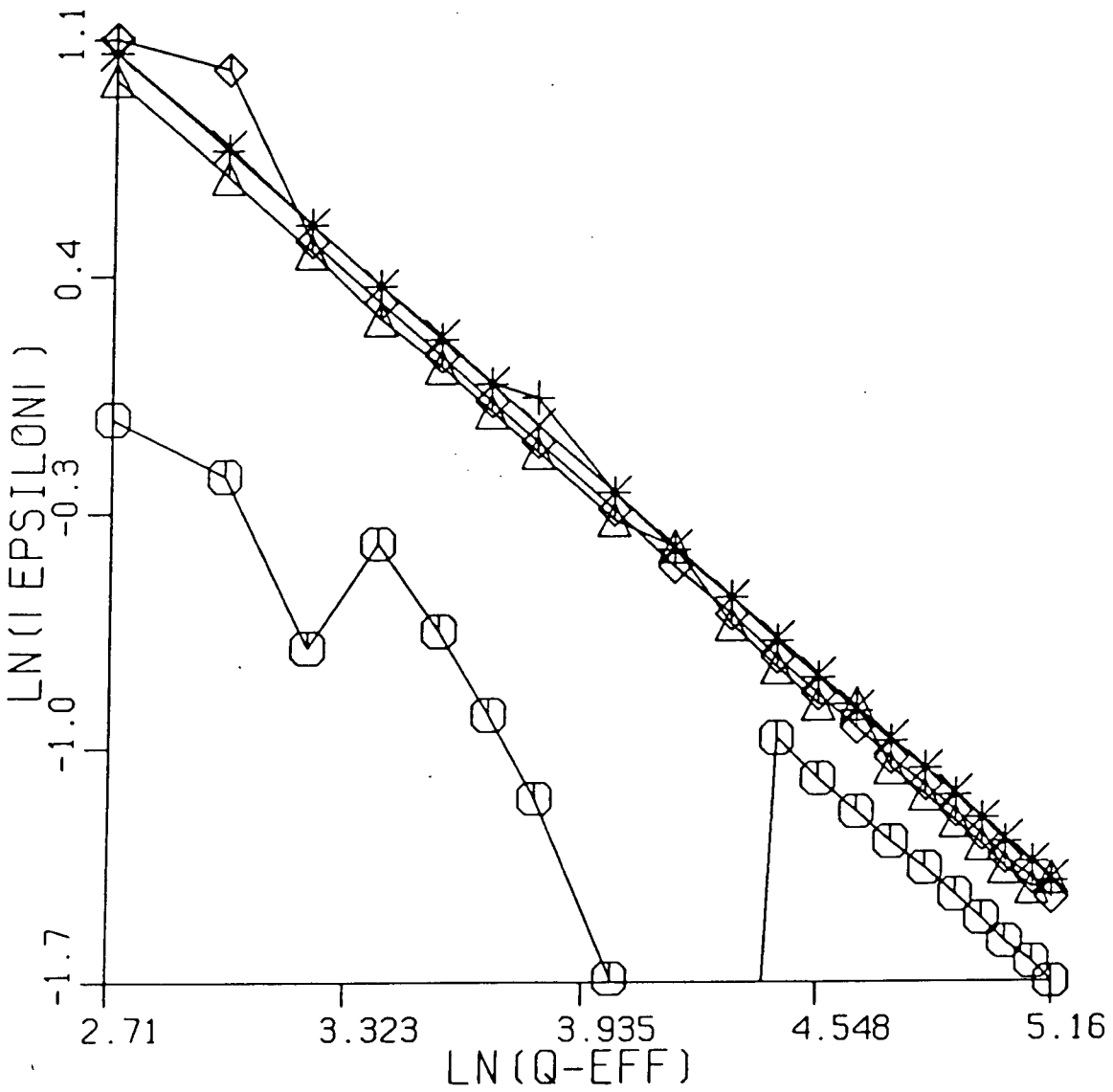


FIGURE 4.30
The effect of varying centre frequency on the locus of $\log|\epsilon|$ versus $\log(Q_{eff})$, for dispersed Ricker wavelets. Deviations from the locii are caused by alignment problems (see text). Circles=15Hz, triangles=25Hz, +=35Hz, X=45Hz, and diamonds=55Hz

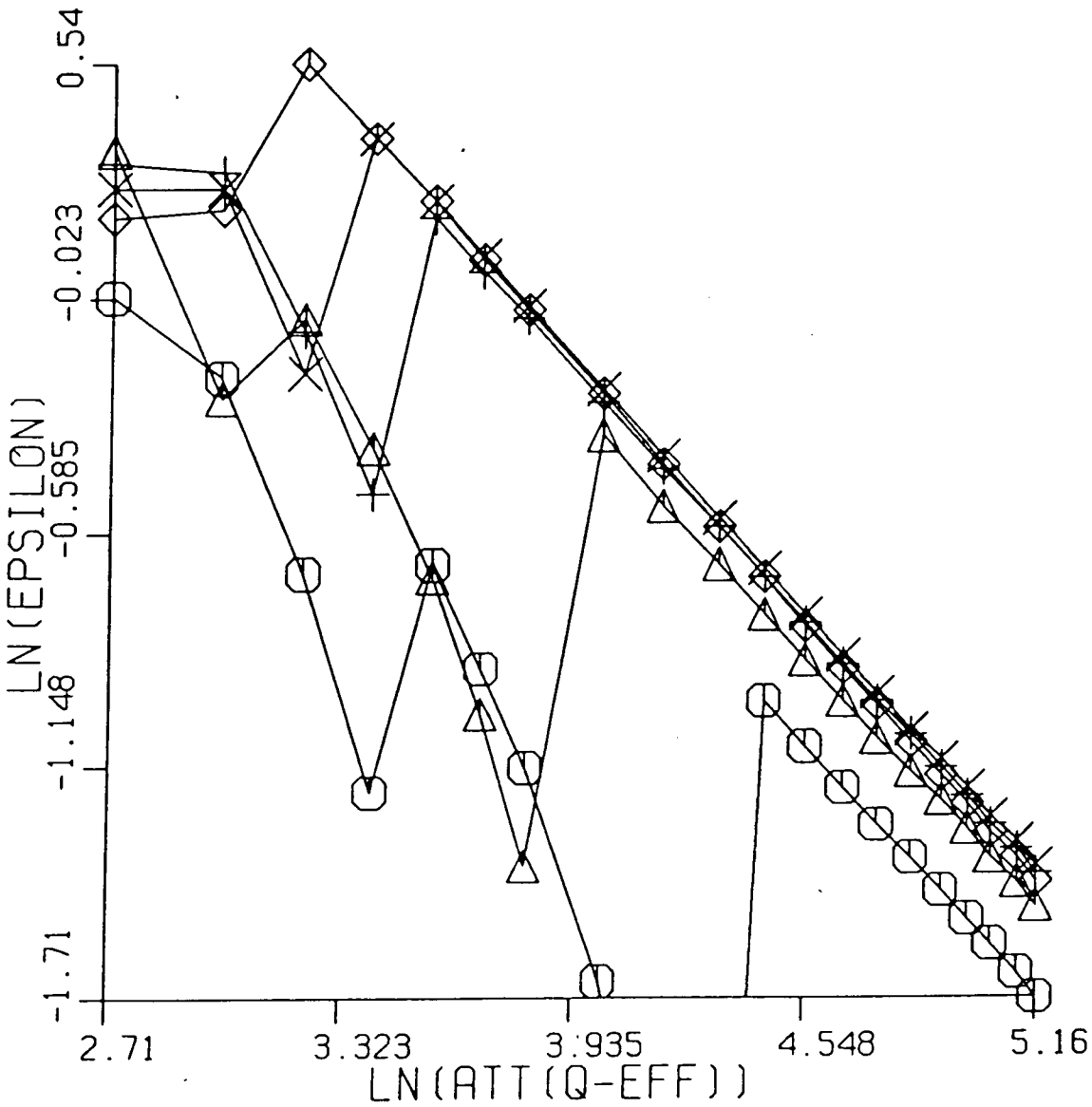


FIGURE 4.31

The effect of varying centre frequency on the locus of $\log|\epsilon|$ versus $\log(Q_{eff})$, for attenuated and dispersed Ricker wavelets. Deviations from the locii are caused by alignment problems (see text). Circles=15Hz, triangles=25Hz, +=35Hz, X=45Hz, and diamonds=55Hz

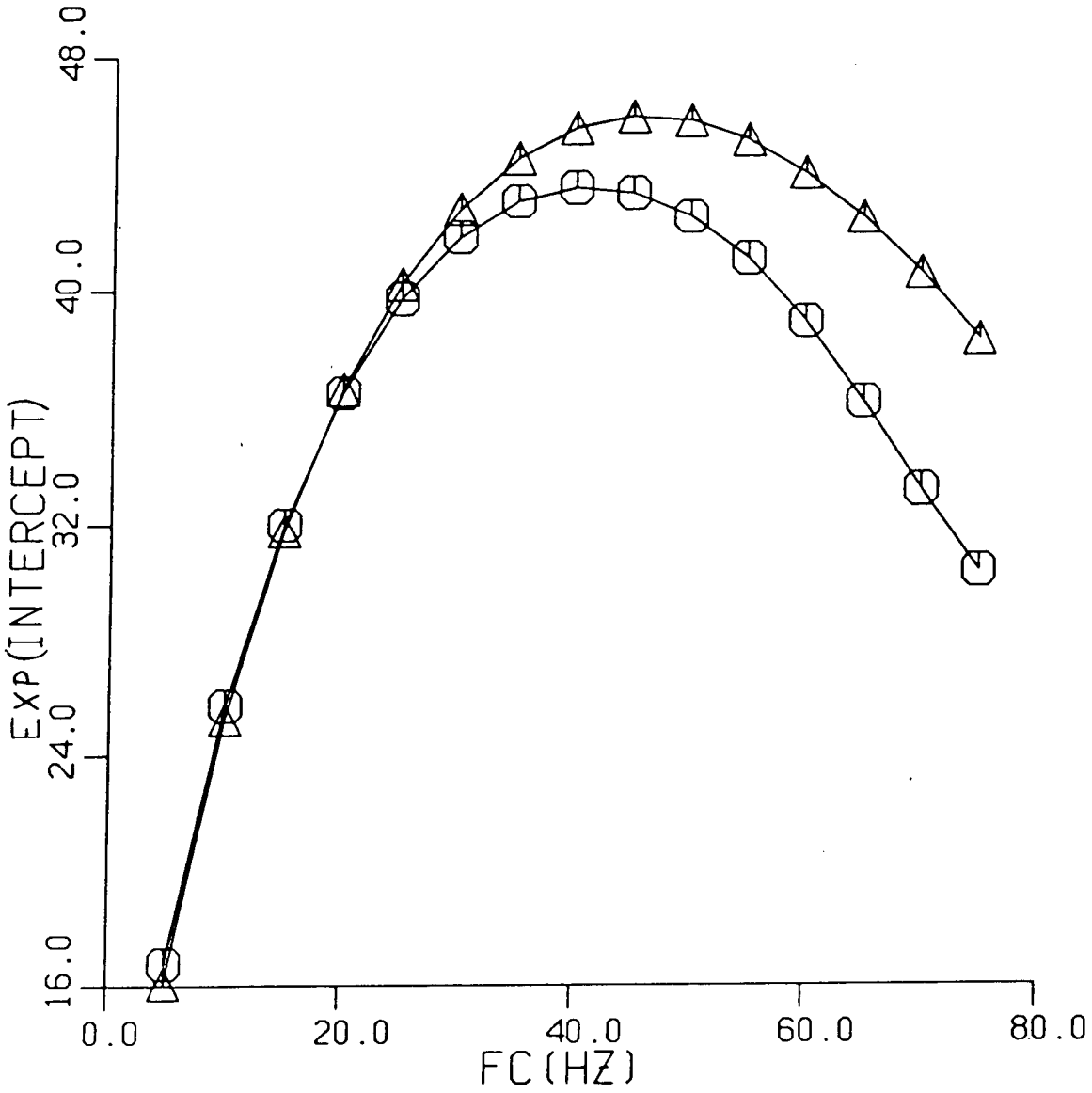


FIGURE 4.32
The variation of the intercept in Figure 4.30, as a function of centre frequency (shown as circles). Triangles represent the analytical results.

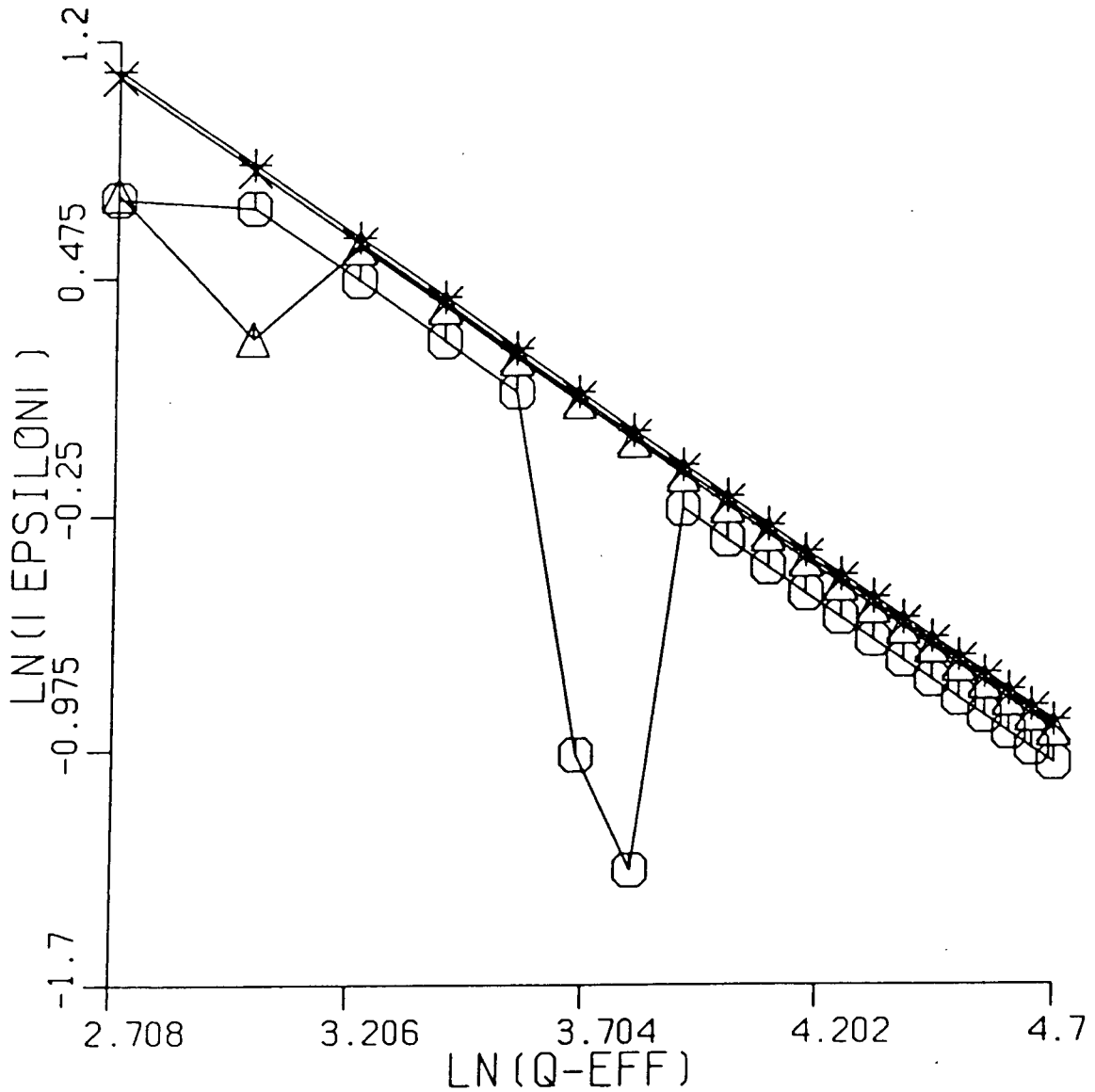


FIGURE 4.33
The effect of varying centre frequency on the locus of $\log|\epsilon|$ versus $\log(Q_{eff})$, for dispersed cosine-Gaussian wavelets. Deviations from the locci are caused by alignment problems. Circles=25Hz, triangles=35Hz, +=45Hz, X=55Hz, and diamonds=65Hz

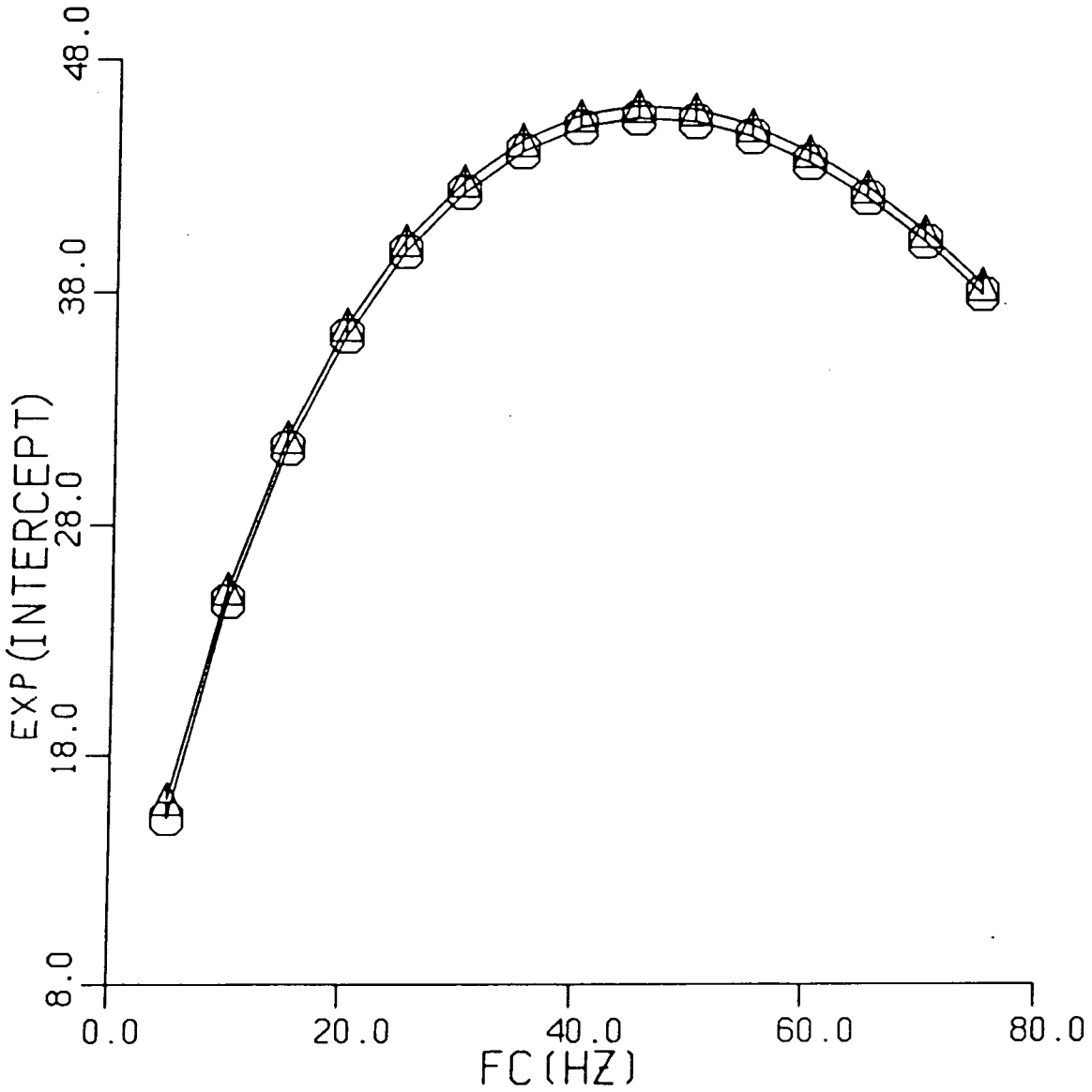


FIGURE 4.34
The variation of the intercept in Figure 4.33, as a function of centre frequency (shown as circles). Triangles represent the analytical results.

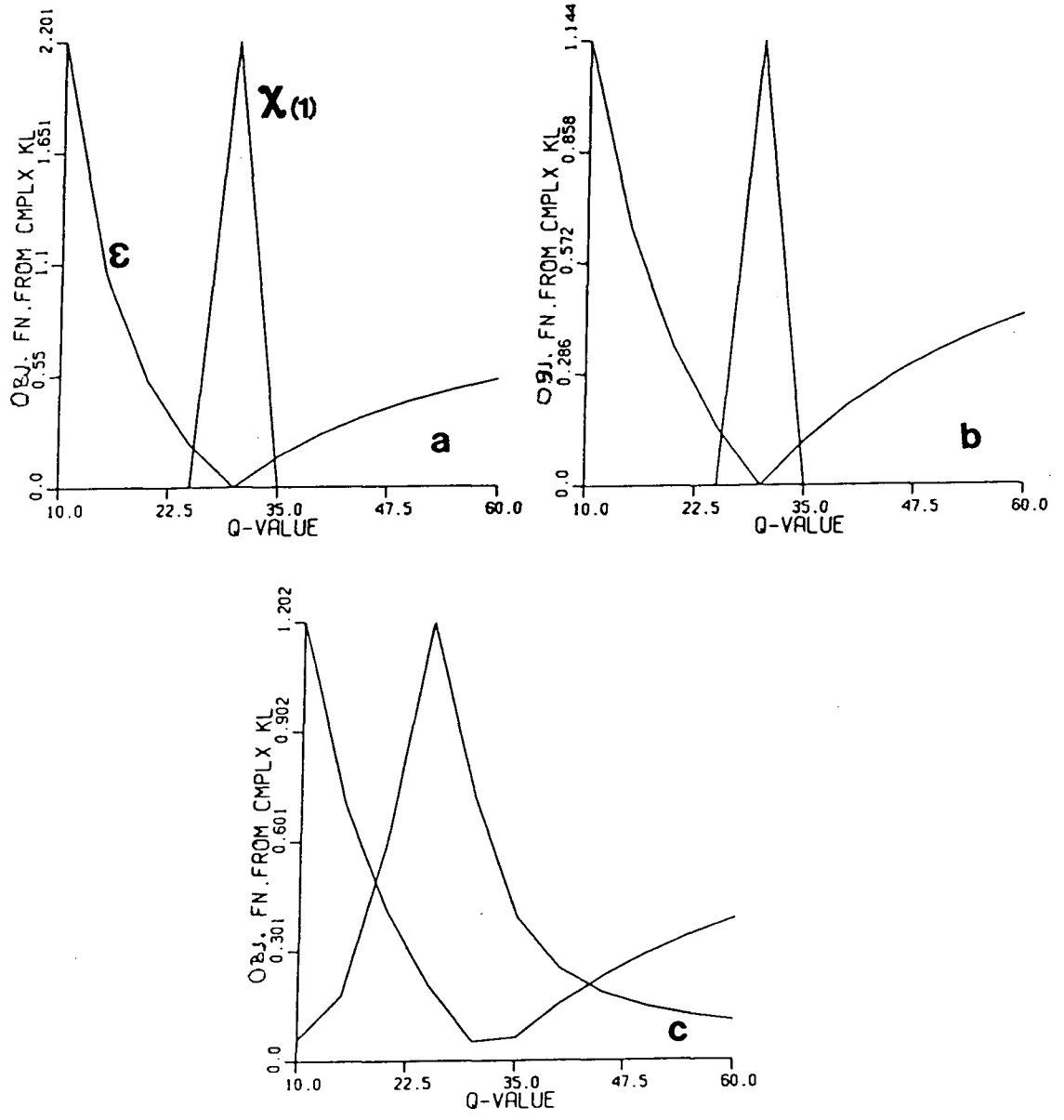


FIGURE 4.35

(a) The EVR and ϵ objective functions from a CKL dispersion removal procedure on a dispersed Ricker wavelet.
(b) The results for attenuated and dispersed data, and
(c) degraded results from noisy attenuated dispersed data.

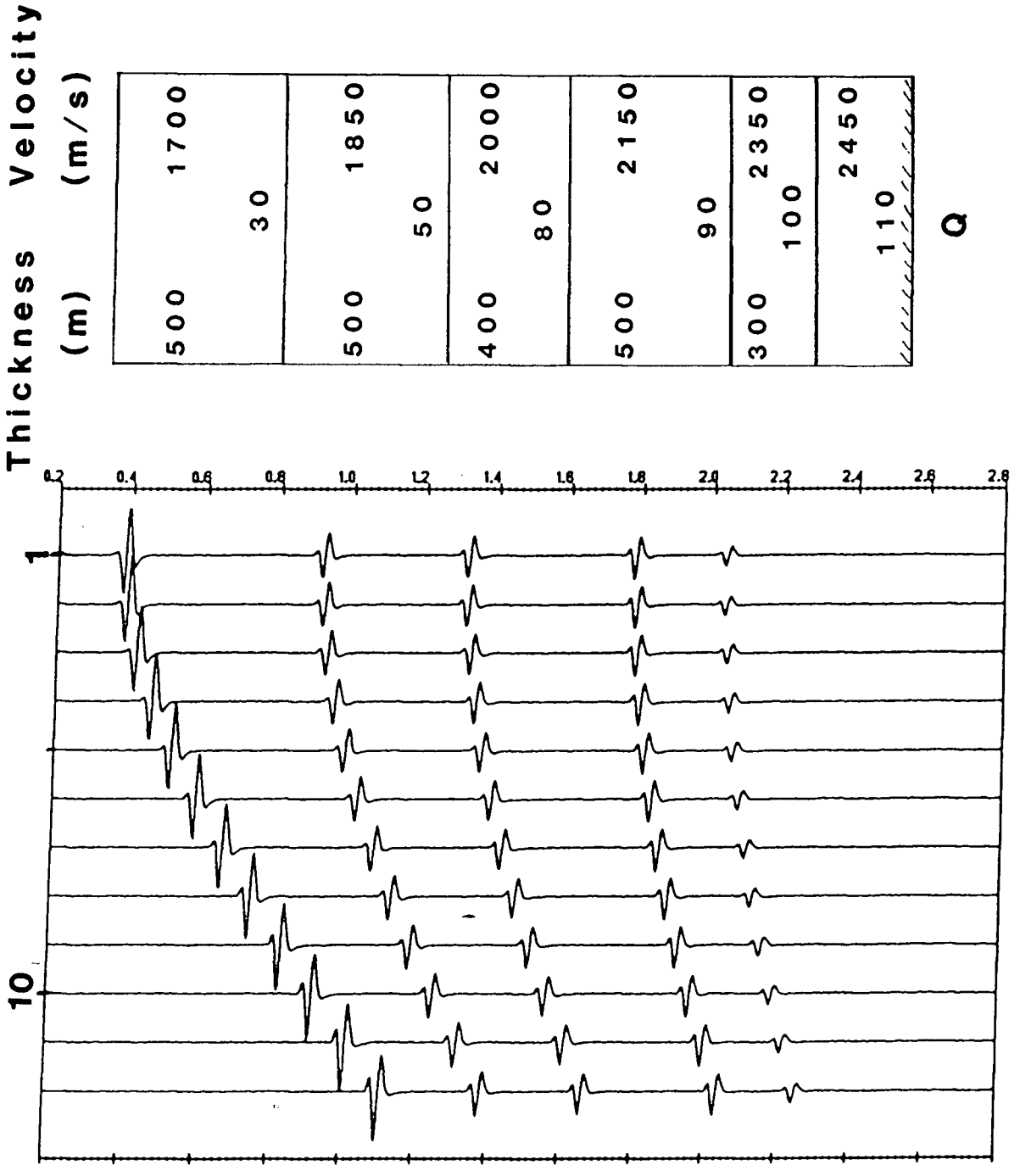


FIGURE 4.36
(a) The layered earth model used for Q inversion.
(b) Dispersed synthetic seismograms, with 5% noise. Offsets range from 0 to 2000 m.

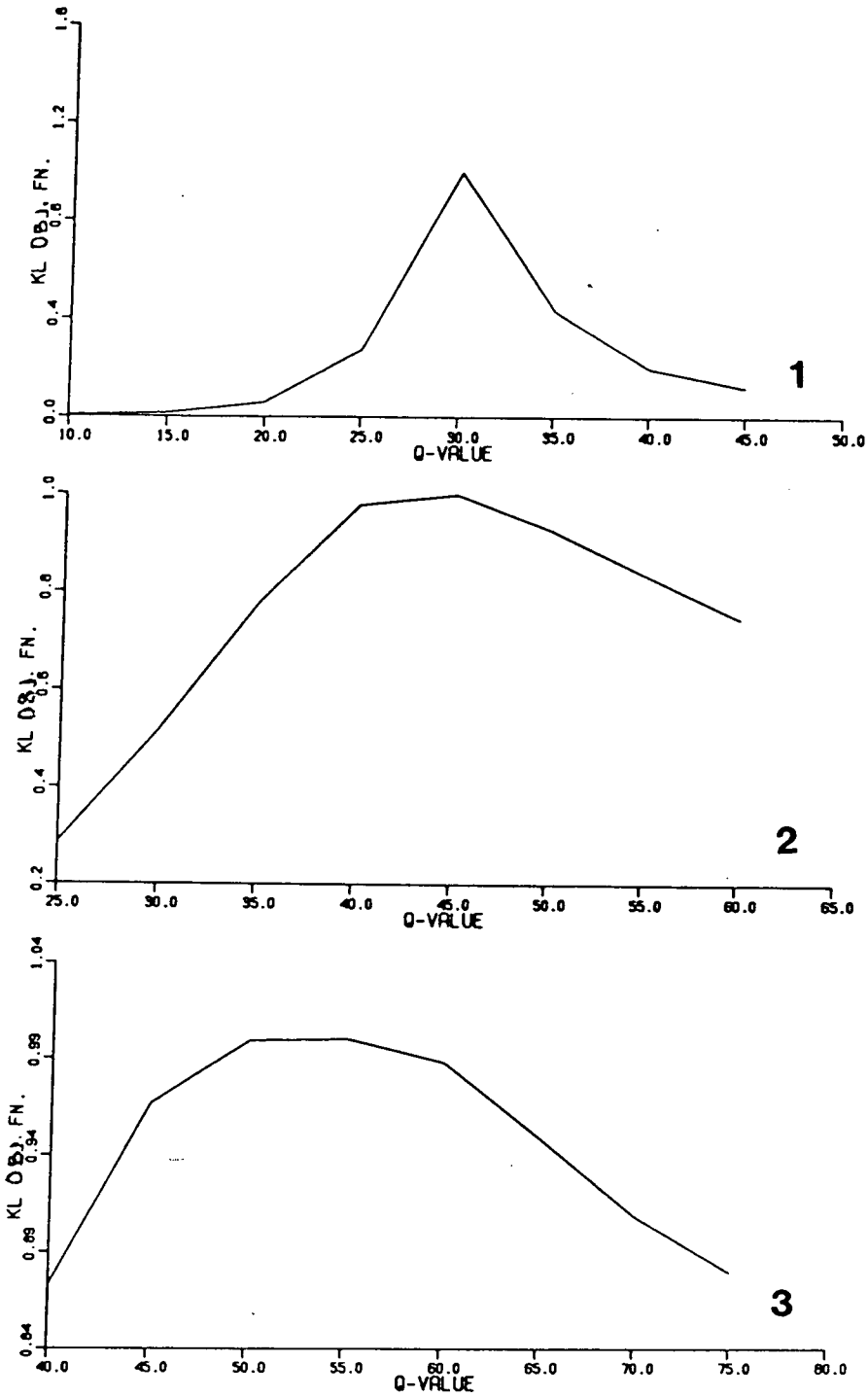


FIGURE 4.37

The RKL objective function for the first three moveout hyperbolae. The similarity is computed between the members of a hyperbola as a whole. After the first result, the cumulative Q values are too high.

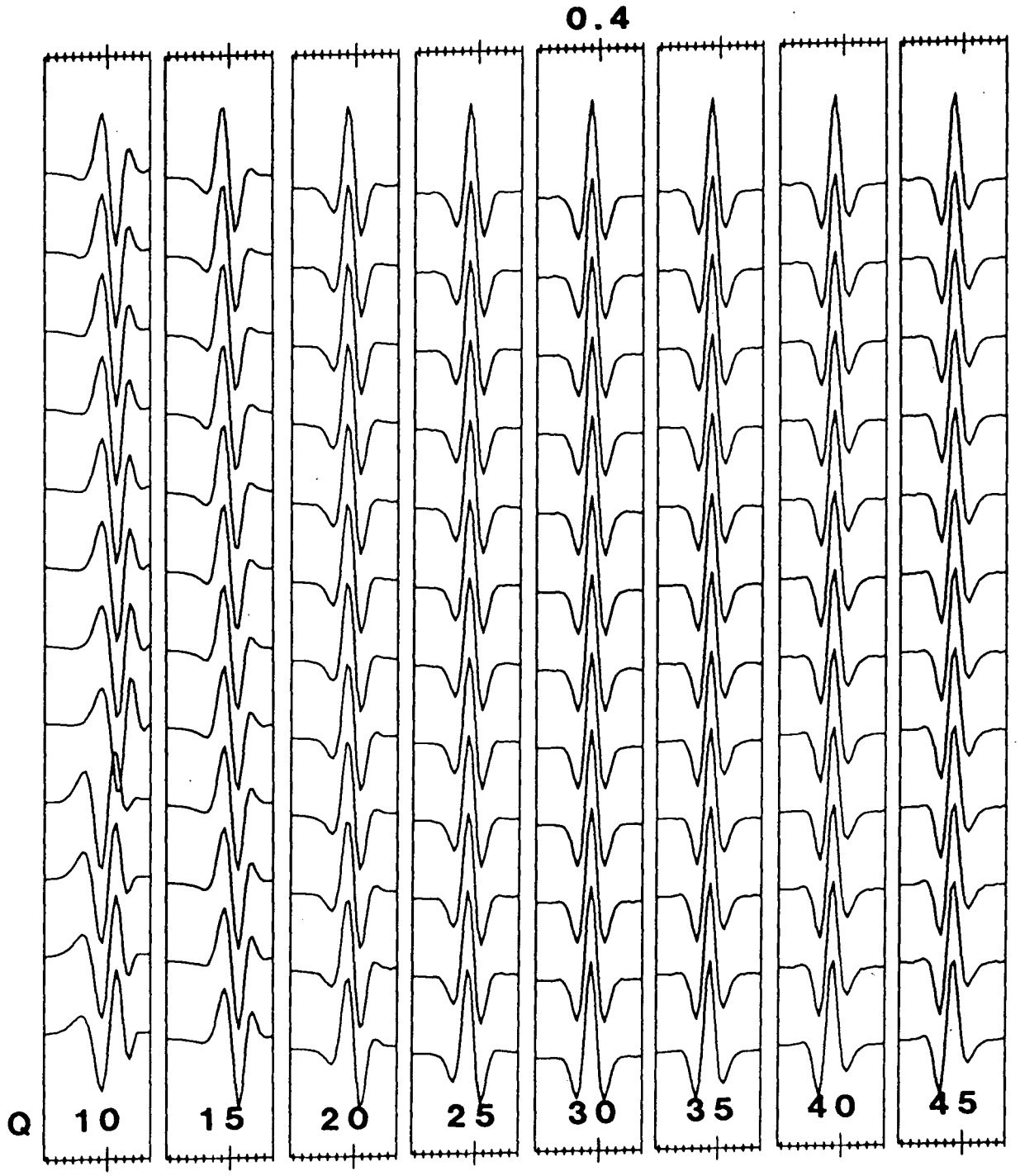


FIGURE 4.38
Twelve wavelets picked from the first moveout hyperbola (at 0.4 s in Figure 4.36), shown at different stages in the iterative process. At Q=30, the wavelets are correctly undispersed, and the corresponding value of the objective function exhibits a maximum.

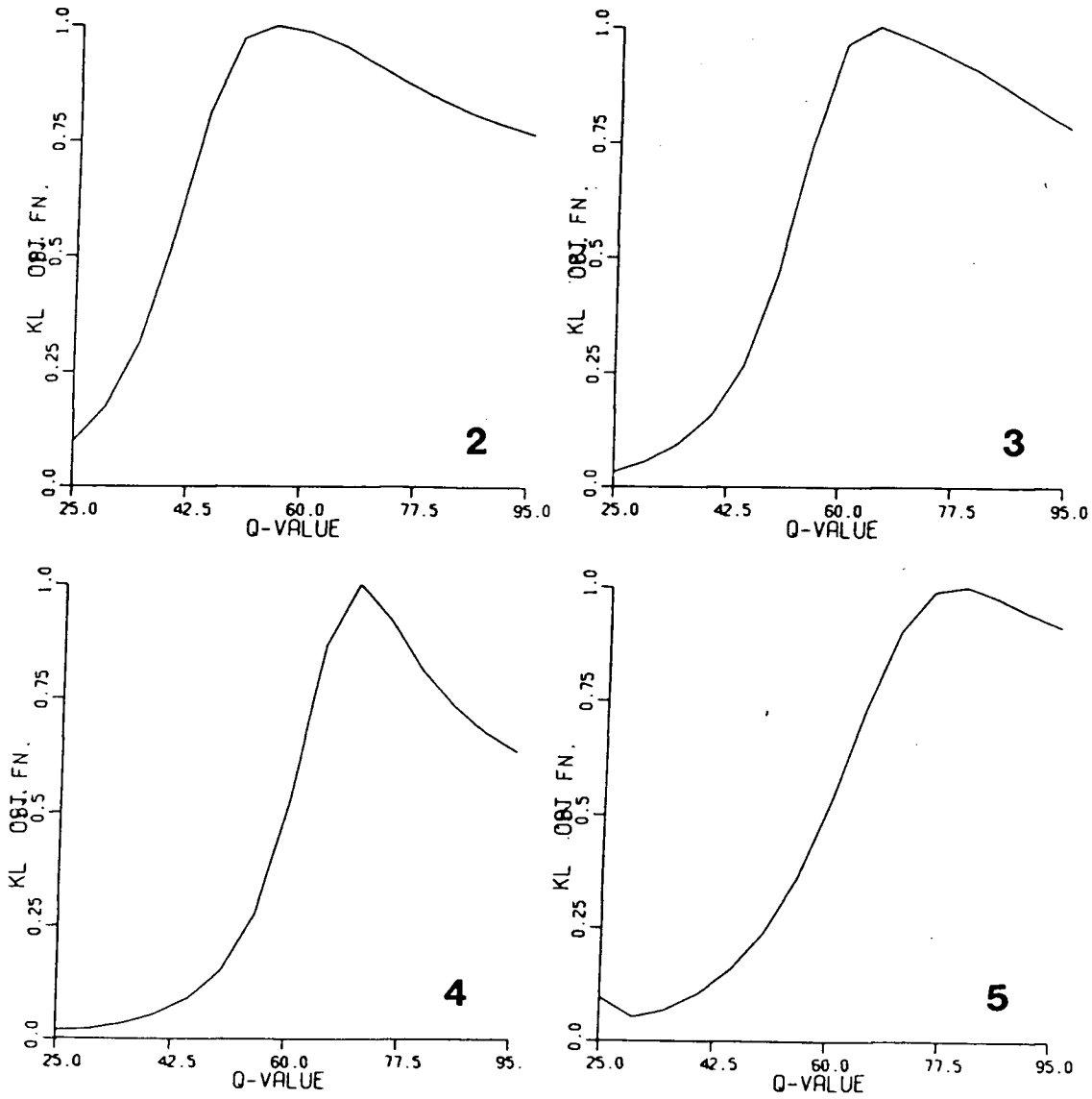


FIGURE 4.39

The RKL objective function for the latter four reflection events (2, 3, 4, and 5), for dispersed data. Here the similarity is computed pair-wise between the members of a single trace. The results from each of the 12 traces were combined to produce a single compound objective function for a given reflection event.

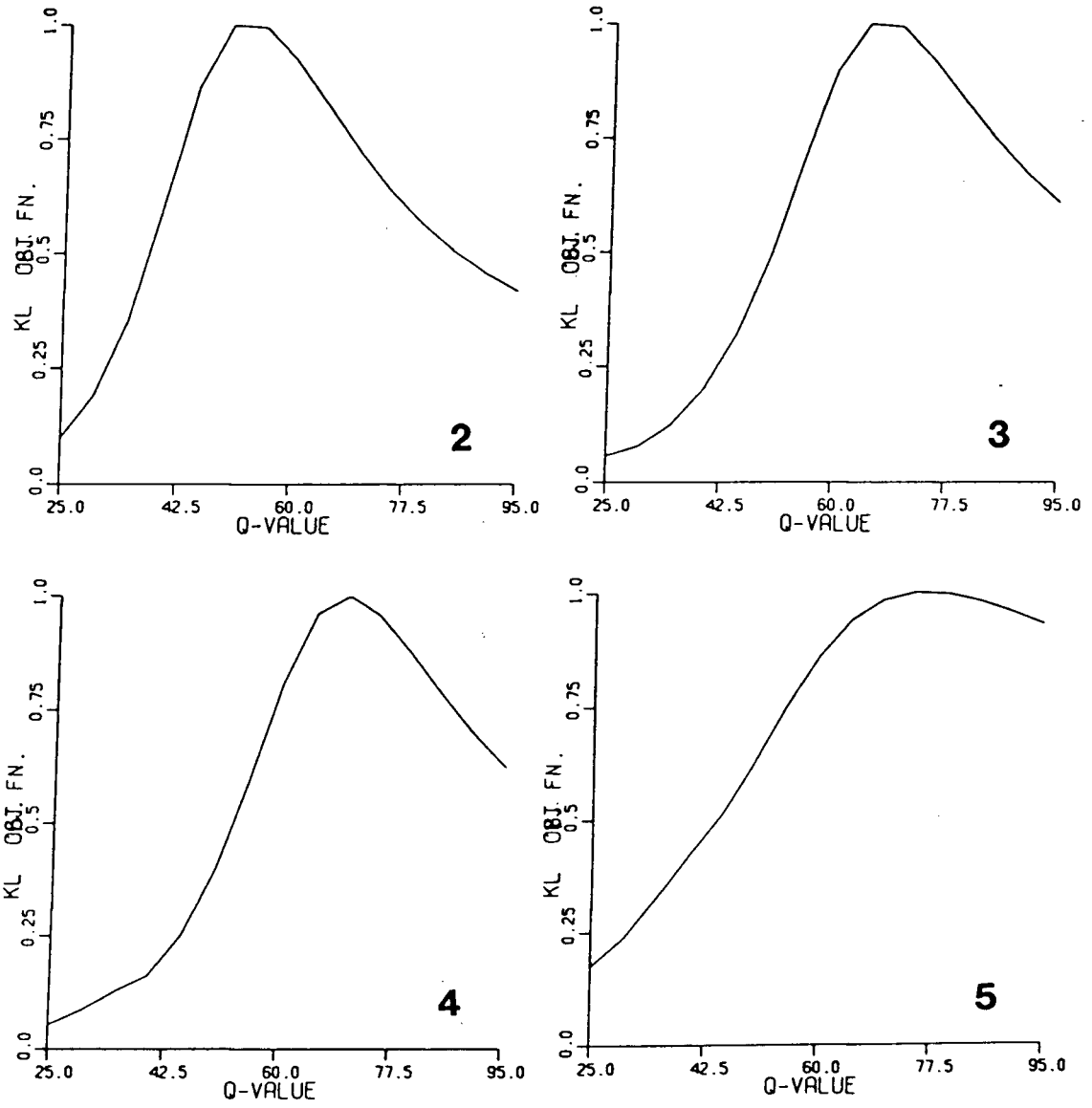


FIGURE 4.40

The RKL objective function for the latter four reflection events (2, 3, 4, and 5), for attenuated and dispersed data. Here the similarity is computed pair-wise between the members of a single trace. The results from each of the 12 traces were combined to produce a single compound objective function for a given reflection event.

CHAPTER 5.

CONCLUSIONS.

SECTION 5.I: REVIEW OF THE GOALS OF THIS WORK.

My initial interest in the Karhunen-Loève transform concerned its ability to quantify the similarity between a pair, or a suite of signals. This interest arose from my attempts to quantify dispersion in seismic data. Extending from that work, my interest shifted to the more general applications of this transform to image processing problems in seismic data processing.

I first used the transformation to update existing processing techniques with a view to extracting additional information from the data. From these endeavours the applications to stacked data reconstruction (following the lead of the satellite image processing community), misfit reconstruction, diffraction separation, and multiple suppression have evolved.

Applying insights gained from work on 'dead trace detection' (Levy et al., 1983), I considered the problem of grouping seismic traces on the basis of eigenvector weightings. Subsequently I found the paper by Hagen (1982) which dealt with the same problem. Noting some of the pitfalls inferred from this paper prompted me to investigate that problem further. The work by LeBlanc and Middleton (1980) and Milligan et al. (1978) dealing

with acoustic pinger records, had showed conclusively that reliable inferences could be made about bottom sediment type on the basis of eigenvector cluster analysis in conjunction with measurements of physical parameters. The application of this technique to seismic stratigraphic character change recognition has potential significance.

The work on velocity analysis also evolved from my involvement in the Levy et al. (1983) study. At that time I was optimistic of the potential of the velocity analysis technique, as the results on synthetic examples were extremely encouraging. The disappointing results with real data point to the limitation of the technique in the presence of noise.

Finally, I present my initial work. This began as a search for an 'objective function' akin to that devised by Beresford-Smith and Mason (1980), for the quantification of dispersive effects. I began by examining the phasor diagrams for the complex trace of the seismic waveform (i.e. a plot of the real versus the complex part of the analytic trace). These results (not shown here) were picturesque, but of little practical use. They dealt with geometric dispersion, which produces a more profound effect on data than does attenuation-related dispersion. As a result, I abandoned this specific topic, and enroute to use of the KL eigenvalue ratio procedure, I investigated the usefulness of the instantaneous phase and instantaneous frequency of the data. Neither of the

latter transformations proved to be of much help, and finally I tried the eigenvalue ratio method described in Section 2.II.

While addressing the dispersion problem, I noted the phase shifted appearance of dispersed signals, and attempted to model dispersive wavelet character change with a simple constant phase change. This procedure proved reasonably acceptable, and may even obviate the procedure of iterative dispersion removal, as an estimate of ϵ for a reference and dispersed signal yields a direct estimate of Q (relative to the reference signal: Section 4.III).

SECTION 5.II: REMARKS.

a. Image processing.

The main thrust of this work has been to introduce new applications of image processing techniques based on the real Karhunen-Loève transformation (Section 2.I). Recovery of coherent information for image enhancement of stacked seismic sections and depiction of anomalous information in those stacked sections constituted the core of Chapter 3.

The stacked section reconstruction technique (Section 3.I) has been the most successful to date, and has aroused some interest in the industrial community, and been applied extensively to real data in an industrial processing environment.

Its main advantage is its ability to greatly reduce the background level of incoherent noise. Of the topics considered in this thesis, I consider this particular application to be the most significant. It is simple to implement, and with appropriate data, generally yields good results.

Misfit reconstruction (Section 3.II) is probably not as widely applicable, but for data which are appropriate for this method, a misfit overlay presentation may be of assistance in drawing attention to anomalous zones. The input from an experienced interpreter and specific knowledge of the data area would enable more meaningful evaluation of the significance or usefulness of the misfit overlay. The usefulness of this particular technique would best be assessed in an industrial seismic interpretation environment where access to many diverse data sets and their geological background are readily available.

An interesting corollary is misfit reconstruction with a view to isolating diffractions from stacked seismic data (Section 3.III). Once segregated, the diffraction hyperbolae can form the basis for residual migration. In addition, in areas with severe faulting, the segregated diffractions may indicate the location of fault edges. This may prove useful in delineating structurally controlled hydrocarbon trapping features. This application is an interesting one, but will probably not be of general use. More likely it will best be used with specific data from severely faulted regions.

Finally, I present a simple and effective method for the elimination of multiple events of known stacking velocity (Section 3.IV). As an aside, an attempt to estimate the multiple velocity by stacking the VA map also proved successful. On real data, the VA maps highlight the almost complete absence of multiple related energy after processing. In addition, primary events which were of small amplitude in the VA map before processing were greatly enhanced, as the actual data hyperbola corresponding to the primary event was no longer distorted by multiple energy arriving at the same time. Also of importance with this method is its ability to preserve wavelet shape after stacking. That is, the processing technique does not distort the waveforms such that the stacked result is degraded. The initial tests of this technique produced stacked sections which looked promising. I consider this technique to be the second major contribution of this thesis. Again, it is simple enough to be readily incorporated into a standard processing stream, and produces encouraging results.

b. Similarity.

In Chapter 4, I investigated the properties of the eigenvectors and eigenvalues of the transformation. This was a natural extension of the work on reconstruction, as a more detailed knowledge of the vagaries of the method was necessary to appreciate its pitfalls.

Application of cluster analysis to seismic problems (Section 4.I) showed the same strengths and weaknesses as did misfit reconstruction. Cluster analysis may provide a useful tool to aid the explorationist in locating unusual features, but much depends on the presentation of this data, and on how non-uniform a particular stacked section may be. As with the misfit reconstruction, this technique would best be investigated further in an industrial interpretation environment.

The eigenvalue ratio criterion constitutes a highly sensitive similarity measure. However, in the context of velocity analysis (Section 4.II) with real data, the measure seems to fail except when applied to small segments of good data. High levels of background noise seem to be the cause for this failure. Working with data after deconvolution may increase the probability of success. However, as demonstrated in Figure 4.21, the RKL and CKL methods clearly out-performed the conventional semblance VA in resolving arrivals at early times. This points to the application of the KL VA to high resolution VA and RMS velocity inversion, for detailed study of the near surface from high quality data.

In addition to velocity analysis, I utilized the eigenvalue ratio criterion to ascertain when, and under what conditions, dispersive signal distortion can be approximated be a constant phase change, for synthetic data (Section 4.III). Further, I used the measure to quantify the degree of success in removing

dispersive effects from signals. With simple synthetic examples, the constant phase approximation was shown to hold, and a relationship between effective Q value, centre frequency, and representative phase shift was derived empirically. A theoretical treatment of the problem for Ricker wavelets was presented, and a result derived which corroborated the empirical findings.

SECTION 5.III: RECOMMENDATIONS.

a. Stacking.

The application of the Karhunen-Loève transform to signal extraction in stacking of moveout corrected CDP gathers has not been explored fully in this work. I introduced this application (explored in detail by Hemon and Mace, 1978) as a preamble to the reconstruction of stacked seismic data. It has been noted (Levy et al., 1983, Ulrych et al., 1983) that both the RKL and CKL transforms optimally extract coherent signal from a pre-stacked gather. Although I was a co-author in those studies, I restricted the scope of this thesis by excluding the stacking problem due to limitations of time.

b. Reconstruction.

The reconstruction of stacked seismic sections has been applied extensively to data, with good results in the industrial environment. Further modifications to the existing algorithm

would include the incorporation of the slant KL procedure (Section 2.III) to compound reconstruction (Section 3.I(f)). At present, the computation of a 'dip dependent' covariance matrix is limited to decomposition of a single image. The program which combines overlapping segments of large seismic sections (i.e. the 'compound reconstruction' technique) does not yet incorporate a slant KL option.

Under the category of misfit reconstruction, I introduced the misfit seismic section in an attempt to highlight both anomalous features and diffractions in stacked seismic data; and also the new multiple suppression technique. Of these three topics, the latter is of the greatest immediate value. An intuitively simple scheme, it works well even when only a few traces per gather are available, whereas the commonly used FK technique of Ryu (1982) suffers from wrap-around problems (D. Hampson, pers. comm. 1985).

For the other two applications, the real data examples I have found so far have not been as convincing. The misfit stacked section for anomaly recognition will probably work best in areas with generally parallel structures, whereas the diffraction separation scheme may produce useful results in regions of severe faulting.

c. Similarity Criteria.

In the fourth chapter of this thesis, I considered the eigenvectors of the KLT and their associated eigenvalues. The cluster groups determined from the eigenvectors (Section 4.I) were shown to belie structural changes, and also to denote changes in phase character for synthetic data examples. For real data examples however, the situation was not as clear. Groupings were seen which correlated with discernable features in the data. The question remains as to the significance or usefulness of these correlations. If groupings can be found which match the inferences of other data, obtained from well logs for example, then the technique may have great value. Further development would best be performed through interaction with experienced seismic interpreters.

Using the eigenvalue ratio to constitute a similarity measure (Section 4.III), I expanded the work of Levy et al. (1983), to appraise the KL VA technique. This gave good results for early arrivals in clean data: the KL VA events being better resolved than those of a corresponding semblance VA. The KL VA technique may be of use for high resolution RMS velocity inversion studies for near surface sediments. However, for data which had significant levels of noise, as evidenced by the example of Figure 4.25, the KL VA technique failed. This

inconsistency in performance renders it inappropriate as a standard processing technique.

The final section of Chapter 4 dealt with the effects of dispersion in synthetic data, and the ability of the KL eigenvalue ratio to aid in quantification of signal similarity. I found that the eigenvalue ratio $\chi(m)$, was a good indicator of signal similarity. Further, using the CKL technique to estimate the phase difference between a dispersed signal and an undispersed reference, I was able to derive a relationship between the measured phase shift angle ϵ, and the dispersing Q value for the data, both theoretically and empirically for synthetic seismograms. Further investigation and corroboration of the synthetic data results could be performed using high quality waveform recordings, such as those presented by Ricker (1953).

d. Conclusion.

In this work I have demonstrated the applicability of the Karhunen-Loève transform to several topics in multichannel seismic reflection data processing. The image reconstruction techniques are essentially an extension of applications from the picture processing and satellite data transmission fields. However, as applied to the separation of residual data (misfit reconstructions), the techniques are new. The algorithms developed during this work are currently being tested in an industrial environment, where additional data and expertise are

available to judge the usefulness of the procedures.

Hopefully, further work on these and related topics will be given consideration by industrial processing groups, with a view to fully exploiting the potential of the Karhunen-Loève transform in the seismic data processing field.

REFERENCES.

1. Ahmed, N., and Rao, K.R., 1975, Orthogonal transforms for digital image processing: Springer, Berlin, p. 189-224.

2. Aki, K., and Richards, P.G., 1980, Quantitative seismology: W.H. Freeman, New York, p167-180.

3. Andrews, H.C., 1970, Computer techniques in image processing: Academic Press, New York.

4. Andrews, H.C., and Patterson, C.L.(III), 1976(a), Outer product expansions and their uses in digital image processing: IEEE Trans. Comput., v.25, p. 140-148.

5. Andrews, H.C., and Patterson, C.L.(III), 1976(b), Singular value decomposition (SVD) image coding: IEEE Trans. Commun., v.24, p. 425-432.

6. Andrews, H.C., and Patterson, C.L.(III), 1976(c), Singular value decomposition and digital image processing: IEEE Trans. Acoustics, Speech, and Signal Processing, v.24, p. 26-53.

7. Beauchamp, K.G, 1975, Walsh functions and their applications: Academic Press, London and New York.

8. Beresford-Smith, G., and Mason, I.M., 1980, A parametric approach to the compression of seismic signals by frequency transformation: Geoph. Prosp., v.28, p. 551-571.

9. Berkhout,A., J., 1984, Seismic migration: imaging of acoustic energy by wavefield extrapolation. B. practical aspects: Elsevier, New York.

10. Bracewell, R.N., 1978, The Fourier transform and its applications: second edition, McGraw-Hill, New York.

11. Chapman, N.R, Levy, S., Stinson, K., Jones, I.F., Prager, B.T., and Oldenburg, D.W., 1984, Inversion of sound speed and density profiles in deep ocean sediments: submitted to the Journal of the Acoustical Society of America.

12. Chapman, W.L., Brown, G.L., and Fair, D.W. 1981, The Vibroseis® system: a high frequency tool: Geophysics, v.46, p. 1657-1666.

13. Chien, Y.T., and Fu, K.S., 1967, On the generalized Karhunen-Loève expansion: IEEE Trans. Information Theory, v.IT-13, p.518-520.

14. Christensen, R.A., and Hirschman, A.D. 1979, Automatic phase alignment for the Karhunen-Loève expansion: IEEE Trans. Biomed. Eng., v.BME-26.

15. Davenport, W.B.,Jr., and Root, W.L., 1958, An introduction to the theory of random signals and noise: McGraw-Hill, New York.

16. Dix, C.H., 1955, Seismic velocities from surface measurements: Geophysics, v.20, p. 68-86.

17. Edelmann, H.A.K., and Werner, H., 1982, The encoded sweep technique for Vibroseis®: Geophysics, v.47, p. 809-818.

18. Fukunaga, K., and Koontz, W.L.G., 1969, Application of the Karhunen-Loève expansion to feature selection and ordering. IEEE Trans. Comput., v.C-19, p. 311-318.

19. Futterman, W.I., 1962, Dispersive body waves: J. Geophys. Res., v.67, p. 5279-5291.

20. Ganley, D.C., and Kanasewich, E.R., 1980, Measurement of absorption and dispersion in check shot surveys: J. Geophys. Res., v.85, p. 5219-5226.

21. Gubbins, D., Schollar, I., and Wisskirchen, P., 1971, Two dimensional digital filtering with HAAR and WALSH

transforms: Annales de Geophysique, v.27, p. 85-104.

22. Gurbuz, B.M., 1982, Upsweep signals with high-frequency attenuation and their use in the construction of Vibroseis® synthetic seismograms: Geoph. Prosp., v.30, p. 432-443.

23. Hagen, D.C., 1982, The application of principal components analysis to seismic data sets: Geoexploration, v.20, p. 93-111.

24. Harlan, W.S., Claerbout, J.F., and Rocca, F., 1983, Extracting velocities from diffractions: Proceedings of the 53rd Annual International SEG Meeting Las Vegas, paper S22.2, p. 574-577.

25. Harlan, W.S., Claerbout, J.F., and Rocca, F., 1984, Signal/noise separation and velocity estimation: Geophysics, v.49, p. 1869-1895.

26. Hemon, Ch., and Mace, D., 1978, Use of the Karhunen-Loève transformation in seismic data processing: Geoph. Prosp., v.26, p. 600-626.

27. Hotelling, H., 1933, Analysis of complex statistical variables into principal components: J. Educ. Psychol., v.24, p. 417-438 and p. 498-520.

28. Huang, T.S., and Narendra, P.M., 1975, Image restoration by singular value decomposition: Appl. Opt., v.14, p.2213-2216.

29. Hunt, B.R., and Kubler, O., 1984, Karhunen-Loève multispectral image restoration, part 1: theory: IEEE Trans Accous., Speech, Signal Proc., v.Assp-32, p.592-599.

30. Jain, A.K., 1976, A fast Karhunen-Loève transform for a class of random processes: IEEE Trans. Commun., v.COM-24, p.1023-1029.

31. Jain, A.K., 1977, A fast Karhunen-Loève transform for digital restoration of images by white and colored noise: IEEE Trans. Comput., v.C-26, p.560-571.

32. Jones, I.F., and Levy, S., 1985, On the separation of signal and coherent noise: for submission to Geoph. Prosp.

33. Kanamori, H., and Anderson, D.L., 1977, Importance of physical dispersion in surface wave and free oscillation problems: review: Rev. Geophys. Space Phys., v. 15, p. 105-112.

34. Kanasewich, E.R., 1982, Time sequence analysis in geophysics: University of Alberta Press, Edmonton.

35. Karhunen, K., 1947, Uber lineare methoden in der Wahrscheinlichkeitsrechnung: Ann. Acad. Sci. Fenn., (Suomalainen Tiedeakatemia), v.37, p. 1-79. (Translation by I. Selin, 1960: "On linear methods in probability theory.": T-131 RAND Corp., Santa Monica, California).

36. Kjartansson, E., 1979, Constant Q wave propagation and attenuation: J. Geophys. Res., v.84, p. 4737-4748.

37. Knopoff, L., 1964, Q: Rev. Geophys. Space Phys., v.2, p. 625-660.

38. Kramer, H.P., and Mathews, M.V., 1956, A linear coding for transmitting a set of correlated signals: IRE Trans. Inform. Theo., v.IT-2, p. 41-46.

39. Larner, K., Chambers, R., Yang, M., lynn, W., Wai, W., 1983, Coherent noise in marine seismic data: Geophysics, v.48, p. 854-886.

40. LeBlanc, L.R., and Middleton, F.H., 1980, An underwater acoustic sound velocity data model: J. Acoust. Soc. Am., v.67, p. 2055-2062.

41. Levin, S.A., Rothman, D., and Rocca, F., 1983, Residual

migration: applications and limitations: Proceedings of the 53rd Annual International SEG Meeting, Las Vegas, paper S10.7, p. 393-395.

42. Levy, S., and Oldenburg, D.W., 1982, Deconvolution of phase shifted wavelets: Geophysics, v.47, p. 1285-1294.

43. Levy, S., Ulrych, T.J., Jones, I.F., and Oldenburg, D.W., 1983, Applications of complex common signal analysis in exploration seismology: Proceedings of the 53rd Annual International SEG Meeting in Las Vegas, paper S6.6, p. 325-328.

44. Loève, M., 1948, Functions aleatories de second ordre: Chapter 8, p.299-352, Hermann, Paris.

45. Loève, M., 1955, Probability theory: D. van Nostrand, New York.

46. Lowitz, G.E., 1978, Stability and dimensionality of Karhunen-Loève multispectral image expansions: Pattern Recognition, v.10, p. 359-363.

47. Mallick, K., and Murthy, Y.V.S., 1984, Pattern of Landsat MSS data over Zawar lead-zinc mines, Rajasthan, India: First Break, v.2, p.16-21.

48. McMechan, G., and Yedlin, M., 1981, Analysis of dispersive waves by wavefield transformation: Geophysics, v.46, p. 869-874.

49. Milligan, S.D., LeBlanc, L.W., and Middleton, F.H., 1978, Statistical grouping of acoustic reflection profiles: J. Acoust. Soc. Am., v.64, p. 759-807.

50. Neidell, N.S., and Taner, M.T., 1971, Semblance and other coherency measures for multichannel data: Geophysics, v.36, p. 482-497.

51. Papoulis, A., 1965, Probability, random variables, and stochasic processes: McGraw-Hill, New York.

52. Pelat, D., 1974, Karhunen-Loève series expansion: a new approach for studying astrophysical data: Astron. and Astrophys., v.33, p. 321-329.

53. Pratt, W.K., 1970, Karhunen-Loève transform coding of images: Proceedings of the 1970 IEEE Int. Symp. Inform. Theory.

54. Rafipour, B.J., 1981, Phase and attenuation studies using phase-matched filtering technique: Ph.D thesis, Southern Methodist University, Dallas.

55. Ready, R.J., and Wintz, P.,A., 1973, Information extraction, SNR improvement, and data compression in multispectral imagery: IEEE Trans. Commun., v.COM-21, p. 1123-1130.

56. Ricker, N.H., 1953, The form and laws of propagation of seismic wavelets: Geophysics, v.18, p. 10-40.

57. Rietsch, E., 1981, Reduction of harmonic distortion in vibratory source records: Geoph. Prosp., v.29, p. 178-188.

58. Robinson, J.C., 1979, A technique for the continuous representation of dispersion in seismic data: Geophysics, v.44, p. 1345-1351.

59. Robinson, J.C., 1982, Time-variable dispersion processing through the use of "phased" sinc functions: Geophysics, v.47, p. 1106-1110.

60. Rosenbaum, J.H., and Boudreaux, G.F., 1981, Rapid convergence of some seismic processing algorithms: Geophysics, v.46, p. 1667-1672.

61. Ryu, J.V., 1982, Decomposition (DECOM) approach to wave field analysis with seismic reflection records: Geophysics, v.47, p. 869-883.

62. Smith, T.A., and Neidell, N.S., 1976, Seismic dispersion estimation in laboratory and marine reflection data:

Proceedings of the 50th Annual International SEG Meeting, New Orleans.

63. Stanford Q Conference 1980, J. Geophys. Res., v.85, p.5171-5256.

64. Strang, G., 1980, Linear algebra and its applications: Academic Press, New York.

65. Strick, E., 1970, A predicted pedestal effect for pulse propagation in constant-Q solids: Geophysics, v.35, p. 387-403.

66. Taner, M.T., Koehler, F., and Sheriff, R.E., 1979, Complex seismic trace analysis: Geophysics, v.44, p. 1041-1057.

67. Tjostheim, D., and Sandvin, O, 1979, Multivariate autoregressive feature extraction and the recognition of multichannel waveforms: IEEE Transactions on Pattern Analysis and Machine Intelligence, v.PAMI-1, p. 80-86.

68. Toksoz, M.N., and Johnston, D.H., (editors) 1981, Seismic wave attenuation: Society of Exploration Geophysicists, Geophysics Reprint Series no.2.

69. Ulrych, T.J., Levy, S., Oldenburg, D.W., and Jones, I.F., 1983, Applications of the Karhunen-Loève transformation in reflection seismology: Proceedings of the 53rd Annual International SEG Meeting in Las Vegas, paper S6.5, p. 323-325.

70. Walsh, J.L., 1923, A closed set of orthogonal functions: Am. Journ. Math., v.55, p. 5-24.

71. Watanabe, S., 1965, Karhunen-Loève expansion and factor analysis. Theoretical remarks and applications. Reprinted in: Pattern Recognition, J. Sklansky, ed., Stroudsburg, Pennsylvania, 1973 (p. 146-171).

72. Werner, H., and Krey, T., 1979, Combisweep - a contribution to sweep techniques: Geoph. Prosp., v.27, p. 78-105.

73. Wuenschel, P.C., 1965, Dispersive body waves - an experimental study: Geophysics, v.30, p.539-551.

74. Young, T.Y., and Huggins, W.H., 1962, The intrinsic component theory of electrocardiograms: IRE Trans. Biomedical Elec., v.BME-9, p. 214-221.

75. Young, T.Y., and Huggins, W.H., 1963, On the representation of electrocardiograms: IEEE Trans. Biomedical Elec., v.BME-10, p.86-95.

76. Young, T.Y., and Calvert, T.W., 1974, Classification estimation and pattern recognition: Elsevier, New York.

APPENDIX 1.

ATTENUATION RELATED DISPERSION.

a. Theory.

For a material which obeys a linear stress-strain relationship, the Q factor for a monochromatic plane wave is defined as the ratio of the amplitude A to the decrease in amplitude δA, over a travel path of one wavelength (Aki and Richards, 1980, p. 168.

For a given angular frequency ω, we have:

$$Q(\omega) = -\pi A/\delta A$$

From this definition, we see that the amplitude of this monochromatic wave at some time t, is given by:

$$A(t) = A_0(1-\pi/Q)^n,$$

where n is the number of wavelengths λ, travelled in time t.

Using the relation $t = n\lambda/v = 2\pi n/\omega$, where v is the velocity, and assuming that Q is constant for all frequencies, we obtain:

$$A(t,\omega) = A_0(1-\omega t/2Qn)^n,$$

which for large n becomes:

$$A(t,\omega) = A_0 exp(-\omega t/2Q) \qquad (A1.1)$$

For a plane wave which undergoes attenuation, we have from *equation (A1.1)*:

$$A(t) = A_0 exp(i\omega t) exp(-\omega t/2Q) \qquad (A1.2)$$

Now, for causality to remain inviolate, we require:

$$A(t) = 0, \text{ for } t < 0$$

Satisfying this condition requires that waves be dispersed, as well as attenuated. Futtermann (1962), using *equation (A1.2)* proceeded to invoke the causality condition to obtain three possible absorption-dispersion pairs. Robinson (1979) used the dispersion relation from the third of these pairs:

$$V(f) = V_0/\{1-Ln(\epsilon f/f_0)/\pi Q_0\} \qquad (A1.3)$$

where $V(f)$ is the phase velocity at frequency f,

f_0 the low frequency cut-off,

V_0 the phase velocity at f_0,

Q_0 the value of the seismic quality factor at f_0,

$Ln(\epsilon)$ Euler's constant, and

$$Q_0V_0 = Q(f)V(f) \qquad (A1.4)$$

To render this expression independent of the low frequency limit Robinson chose an arbitary base frequency f_b and redefined the dispersion relation *equation (A1.3)* with respect to f_b. For some frequency

$f=f_b+\delta f$, we have:

$$V(f_b+\delta f) = V(f_b) + \delta V(f, f_b) \qquad (A1.5)$$

Substituting *equation (A1.4)* and *equation (A1.5)* into *equation (2.3)* and subtracting *equation (A1.3)* from the result yields:

$$\delta V(f, f_b)/\{V(f_b) + \delta V(f, f_b)\} = Ln(f/f_b)/\{\pi Q(f_b)\} \qquad (A1.6)$$

Under the assumption that Q is independent of frequency, which is justified for relatively narrow band seismic signals (Futtermann, 1962; Wuenschel, 1965; Kanamori & Anderson, 1977; Kjartansson, 1979; Aki & Richards, 1980, p.167) *equation (A1.6)* yields a simple and useful basis for dispersion modelling or for the removal of dispersive effects.

Now, consider a pair of plane waves of frequencies f and f_b initiated simultaneously at some arbitary source location X_s. The arrival times of these waves at some receiver location X_r, is given by:

$T(f) = \delta X/V(f)$, and

$T(f_b) = \delta X/V(f_b)$,

where $\delta X = X_r - X_s$.

thus, $T(f_b) - T(f) = T(f_b) - \delta X/V(f)$

$= T(f_b) - T(f_b)V(f_b)/V(f)$

$$= \delta T(f, f_b), \text{ say} \qquad (A1.7)$$

By analogy with *equation (A1.5)* let $V(f)=V(f_b)+\delta V(f, f_b)$, so:

$$\delta T(f, f_b) = T(f_b) - T(f_b)V(f_b)/\{V(f_b) + \delta V(f, f_b)\}$$

$$\delta T(f, f_b)/T(f_b) = 1 - V(f_b)/\{V(f_b) + \delta V(f, f_b)\}$$

$$= \delta V(f, f_b)/\{V(f_b) + \delta V(f, f_b)\}$$

Equating this with *equation (A1.6)* we have:

$$\delta T(f, f_b)/T(f_b) = Ln(f/f_b)/\{\pi Q\} \qquad (A1.8)$$

From *equation (A1.7)* we see that frequencies above f_b will have a positive δT, i.e. from *equation (A1.3)* the higher frequencies arrive first, and vice versa.

Essentially, the relation specified in *equation (A1.3)* states that the effects of a frequency dependent velocity are equivalent to those obtained from a simple linear scaling of the time axis for each frequency. This scaling operation is summarized by the Fourier scaling or similarity theorem (Bracewell, 1978, p.101).

For our signal $x(t)$, we introduce this frequency dependent time rescaling factor:

$$\gamma(f, f_b) = (T(f_b) + \delta T(f, f_b))/T(f_b)$$

$$= 1 + \delta T(f, f_b)/T(f_b)$$

which from *equation (A1.8)* becomes:

$$\gamma(f, f_b) = 1 + Ln(f/f_b)/\{\pi Q\} \qquad (A1.9)$$

Therefore, upon dispersion we have:

$$x(t) \text{ ------> } x(t/\gamma(f, f_b)),$$

i.e. the contribution to $x(t)$ at time t from frequency f, for $f > f_b$ say, now arrives in a shorter time than it would in an undispersed signal.

In the context of this scaling theorem, the time signal $x(t/\gamma)$ is given by:

$$x(t/\gamma) = \int \gamma X(f\gamma) exp(i 2\pi f t)\ df$$

that is:

$$x(t/\gamma(f, f_b)) \text{ <---> } \gamma(f, f_b) X(\gamma(f, f_b) f) \qquad (A1.10)$$

It is this expression upon which Robinson's (1979), and our, programs for dispersion were based, while an attenuative component in our synthetic data could be introduced via *equation (A1.1)*, if desired.

b. Numerical Methods.

As the complex spectrum $X(f)$ is rescaled to $\gamma X(\gamma f)$ but only sampled at $X(n\delta f)$ (where δf is the frequency interval and $n=0, 1, 2, \ldots$) we must at some stage introduce an interpolation. Our

approach was to ask the following question: given the discrete locations $n\delta f$, where did the spectral component, with its current phase and amplitude, come from (assuming that it was shifted by some phase change due to dispersion)? Knowing that:

$$f_c = \gamma f_0$$

$$= [1 + Ln(f_0/f_b)/\{\pi Q\}]f_0,$$

where f_c is the current frequency, and f_0 is the frequency that was shifted, we develop a Newton-Raphson iterative approach to finding f_0 given f_c. To do this, we need expressions for both the function to be minimized, and its derivative:

$$\Phi_1 = \gamma f_0 - f_c$$

$$= [1 + Ln(f_0/f_b)/\{\pi Q\}]f_0 - f_c$$

$$\Phi_2 = \partial\Phi_1/\partial f_0$$

$$= 1 + [1 + Ln(f_0/f_b)]/\{\pi Q\}$$

Having found the location f_0, at which we wish to evaluate the complex function $X(f)$, we utilize an interpolation algorithm based on integration-by-parts (after Rosenbaum and Boudreaux, 1981).

To invoke attenuation satisfying *equation (A1.1)*, we sampled the data with a window of width W, centred about the current time T_c. The data in the window were Fourier transformed and

multiplied by $exp(-\omega T_c/2Q)$. The window was moved along the data in steps of $W/2$ and the resulting set of attenuated complex spectra transformed back into the time domain. The data segments were then combined to construct the attenuated time series.

c. Cumulative to Interval Q Inversion.

In the same way that interval velocities are averaged to produce RMS velocities (Dix, 1955), the layer Q values are seen in an averaged sense in the measured data. For a layered earth with interval Q values Q_i corresponding to the $i.th$ layer with interval travel time t_i, we have:

$$1/Qcum_j = \{ \sum_{j=1}^{i} t_i/Q_i \} / \sum_{j=1}^{i} t_i \qquad (A1.11)$$

A sequence of measured cumulative Q values may be iteratively treated in the usual 'layer stripping' manner, or inverted using Backus-Gilbert inversion methods, to yield the interval Q structure.

APPENDIX 2.

THE EFFECT OF DISPERSION ON VIBROSEIS DATA PROCESSING.

a. Introduction.

The material in this Appendix is essentially the text of a manuscript in preparation, and is a continuation of work initiated by S. Levy in 1980.

Here we present a background to the problem of dispersion as it affects a signal with regard to bandwidth and, in light of the constant Q model, produce synthetic seismograms to demonstrate how the problems of dispersion manifest themselves. Our adoption of the constant Q model is meant to be instructive, not definitive: we believe that it yields useful insights into the nature of the problem. The constant Q model (Futtermann, 1962; Wuenschel, 1965; Kanamori & Anderson, 1977; Kjartansson, 1979; Aki & Richards, 1980, p167) states that attenuation increases linearly with frequency over the band of interest: i.e. Q, the fractional energy loss per cycle, is constant for a given medium.

Given the assumption of constant Q, we proceed to show how a judicial choice of band-width and band location of the input signal can partially overcome the deleterious effects of dispersion. Such choice is only possible with tuned sources, such as those applied in the Vibroseis® technique. Through the use of

synthetic data, we consider in some detail the effect of dispersion on Vibroseis® data processing. We develop the rationale for a method of separating the usual multi-octave Vibroseis® sweep (hereafter referred to as the sweep) into non-overlapping individual octave sweeps. After recording the separate data sets and cross correlating with the appropriate sweep, the correlograms from the different sweeps are weighted and then stacked. It is our contention that data gathered and compiled in this manner may suffer less from the effects of dispersion than would a single, multi-octave sweep, or a stack of overlapping multi-octave frequency band sweeps (e.g. the Combisweep of Werner & Krey (1979)).

b. Theory.

Dispersion can be conceptualised as a frequency dependent stretching of the time axis (Appendix 1). For a signal $x(t)$ which becomes dispersed, one has, following *equation (A1.9)*:

$$x(t) \text{------>} x(t/\gamma(f,f_b)),$$

where $\gamma(f,f_b)=1 + Ln(f/f_b)/\{\pi Q\}$ *(A2.1)*

f_b is some base, or reference frequency,

and Q is the seismic quality factor.

From this expression we note that frequencies above f_b will be phase advanced, frequencies below f_b phase delayed.

We now consider the effect of band width on the absolute degree of distortion due to dispersion. Taking $\gamma(f,f_b)$ from *equation (A2.1)* at frequencies f_1 and f_2 $(f_1 < f_2)$, we have:

$$\gamma(f_1,f_b) - \gamma(f_2,f_b) = \{Ln(f_1) - Ln(f_2)\}/\{\pi Q\}$$

$$= Ln(f_1/f_2)/\{\pi Q\} \qquad (A2.2)$$

$$= \delta\gamma(f_1,f_2), \text{ say}$$

For a band one octave in width, regardless of frequency, we have

$$f_1 = 2f_2, \text{ thus } \delta\gamma(2f,f) = Ln(2)/\{\pi Q\} = 0.22/Q.$$

Therefore, for n octaves we have:

$$\delta\gamma(2^n f/f) = nLn(2)/\{\pi Q\} = 0.22n/Q \qquad (A2.3)$$

The difference $\delta\gamma(f_1,f_2)$ in the frequency dependent time-rescaling factor $\gamma(f,f_b)$, over a given band is a linear function of the number of octaves in that band.

Thus, a signal with a spectrum several octaves wide will have frequencies at either end of the band with large relative phase shifts. This leads to considerable distortion of the received dispersed signal. Subsequent deconvolution or cross correlation with a known or estimated input signature gives rise to even greater problems, as the signature used in these procedures is not dispersed whilst its replications within the recorded signal are increasingly dispersed, and attenuated, with

greater travel time.

c. The Vibroseis® Technique.

One of the problems with the Vibroseis® method is the degradation of signal due in part to dispersion and attenuation. Several methods have been suggested whereby the sweep is modified or encoded in some way so as to minimize the noise in, and increase the resolution of the recorded signal (Gurbuz 1982; Edelmann & Werner 1982; Chapman et al. 1981; Rietsch 1981; Werner & Krey 1979).

Following from the insights gleened from analysis of the theory, we propose a modification of the existing methods. Recalling that a multi-octave sweep is prone to more distortion than a single octave sweep, we suggest that the sweep be confined to one octave. Repeating the sweep successively for three different octaves would yield three sets of data, each data set having succumbed to a minimal amount of dispersive effects. Subsequent cross correlation with the respective sweeps, and stacking, would produce a multi-octave data set with less distortion than data of comparable band-width that had been collected and processed conventionally. Repeating the higher frequency octave sweeps several times and including these data sets in the subsequent stack would partially alleviate attenuative loss by initial boosting of the high frequency content. This procedure could be paralleled with existing data by

bandpassing the pre-correlated data and wavelet, and stacking the correlograms from several such band limited replications.

d. Synthetic Data Examples.

The objective of the following examples is to demonstrate the effects of dispersion and of attenuation on the phase and amplitude structure of the input data. To further this end, the synthetic data were kept simple.

Four synthetic seismograms, each comprising five spikes of alternating polarity spread over four seconds, were created and correlated with a particular sweep. For this demonstration, sweeps of only one half second duration were used (in practice the sweep is at least several seconds in duration) with a sampling interval of 4 ms. The bandwidths chosen for the four synthetic sweeps were:

a. 10.0 - 21.5 Hz

b. 18.5 - 43.0 Hz

c. 37.0 - 80.0 Hz

d. 10.0 - 80.0 Hz.

All spectra were tapered and bandwidths a, b, & c were overlapped slightly to avoid gaps in the spectrum of the stacked correlogram. The spectra shown in Figure A.2, are of one half second of data centred about the pulse near 3.6 s.

Following the approaches of Robinson (1979) and Strick(1970), programs were developed to disperse and attenuate the data given the seismic quality factor, Q, the frequency band, and the base frequency, f_b (Appendix 1). In each case the base frequency f_b was chosen to lie at the mid-point of the respective frequency band. In Figure A.1 the correlogram of the multi-octave sweep (a), is compared to the correlogram formed from the stack of the three single octave sweep correlograms (b). The stack was weighted to balance the energies in the three octaves (Figure A.2 a and b). As expected, these results are similar. The lobes in the conventional band spectrum (Figure A.2a) are due to Gibb's phenomenon (Bracewell, 1978).

The four data sets were dispersed with a constant Q of 150, and cross-correlated with the respective non-dispersed sweeps. This mimics the actual field procedure, i.e. the recorded signal, complete with the various 'earth effects', is cross correlated with the known input sweep. The three individual octave sweeps were again stacked. Figure A.3 (a and b) demonstrates the marked disparity between the two procedures. The multi-octave results show significant distortion after the first second of two way travel time. Polarity begins to be lost as the dispersive effects increasingly distort the phase structure of the signal. However, the result from the stack, is much better, although it also is somewhat effected.

Repeating the above procedure with Q=125 and introducing attenuation prior to dispersion we have, as expected, an increasing loss of resolution with depth as the high frequencies are attenuated. Figure A.4 (a and b) shows once more the comparison of the conventional approach to octave stacking. Polarity and position are now severely degraded in the standard multi-octave sweep. In the spectra (Figure A.2 d and e) we note that the high frequencies are severely attenuated.

The distortion due to the time-dependent component of attenuation can be partially compensated by the implementation of automatic gain control. However, the recovery of attenuated frequencies poses a greater problem. In our case, with the well defined separate octave sweeps, we proceed to implement a form of spectral normalization. The mean energy in each octave was equalized so as to boost the high frequencies. The normalization is only implemented within the respective frequency band so as not to amplify noise outside the band. Figure A.5 (a and b) compares the attenuated and dispersed data of Figure A.4b before and after high frequency recovery. The wavelets sharpen-up considerably, as noted in the respective spectra (Figure A.5 c and d). This procedure assumed no prior knowledge of the Q value used to introduce attenuation in the original synthetic seismograms.

This theoretical procedure could be mimiced in the field by repeating the high frequency octave sweep several times and

including these repetitions in the stack with the other octave sweeps. In effect we boost the high frequency content of the compound spectrum. In order to implement high frequency recovery in a more exact manner, we would need a detailed knowledge of the absorptive and scattering contributions to the attenuation process.

e. Comparison with Previous Work.

Overlapping frequency bands, as suggested by Werner and Krey (1979) provide an improvement over single multi-octave procedure, but the method is still inferior to the separate octave approach. To exemplify this point, we created three more synthetic traces, dividing the 10 - 80 Hz range into three overlapping bands. They were:

e. 10 - 52 Hz,

f. 24 - 66 Hz, and

g. 38 - 80 Hz.

All bands are 42 Hz wide, and each subsequent band is shifted by 14 Hz, this choice of frequency bands approximately mimics those of Werner and Krey (1979). In Figures A.1c & 2c respectively, the stacked correlogram for the overlapping bands and its spectrum are shown. Following our earlier examples we disperse the seismograms using Q=150. Figure A.3 compares the correlograms from the conventional method (a), octave stack (b), and overlapping band stack (c), for dispersed data. The improvement

of the octave stack over the overlapping band stack is about the same as that of the latter over a conventional multi-octave correlogram.

Attenuating and dispersing the data with Q=125 produces results similar to those in Figure A.3 only now we see the drop-off of the high frequencies (Figure A.2f) and the loss of signal strength with time (Figure A.4).

Chapman et al. (1981), in their paper on the use high frequency Vibroseis® sweeps show several real data sections comparing conventional to high frequency sweeps. Their data elegantly show the vast improvement of the high frequency method over the conventional sweep. However, their higher frequency sweeps span less than two octaves, while their low sweeps span more than two octaves. Referring back to *equations (A2.2)* and *(A2.3)* we note that this improvement would be expected given the reduction in the number of octaves encompassed in the high frequency range.

In order to demonstrate this contention more clearly, we constructed one more set of synthetic traces:

h. 10 - 32 Hz, and

i. 24 - 80 Hz.

Both these sweeps constitute about 1.4 octaves, i.e. 10 - 28 and 28 - 80 Hz with some overlap to avoid gaps in the compound spectrum, and span the same frequency range as did our first

triad of uni-octave sweeps. This pair of sweeps was correlated and stacked in a similar manner to the earlier triad. Without a knowledge of the implications of the constant Q model we would expect these two examples, when dispersed, to give similar results, as the overall frequency range spanned by the stacked data sets is the same. The three component stack should have slightly lower noise. However, after dispersion with Q=100, the result from the stacked pair of 1.4 octave sweeps is markedly inferior to that of the stacked uni-octave triad (Figures A.6).

f. Conclusions.

Dispersion, which plays a role in the degradation of long seismic records, produces effects related to the number of octaves present in a signal. Given the applicability of the constant Q model to seismic reflection data, the deleterious effects of dispersion may be partially avoided by judicial choice of the band width and band location of the input signal. Such choice is possible with tuned sources, such as Vibroseis®, and may prove to be worth the additional recording and processing costs involved in the proposed method. Systems already designed to run a Combisweep survey could be readily adjusted to the octave-stack method.

Regardless of the absolute validity of a constant, or nearly constant, Q model, analyses such as the one presented here can offer useful insights into the processes affecting and afflicting

the propagating seismic wavelet.

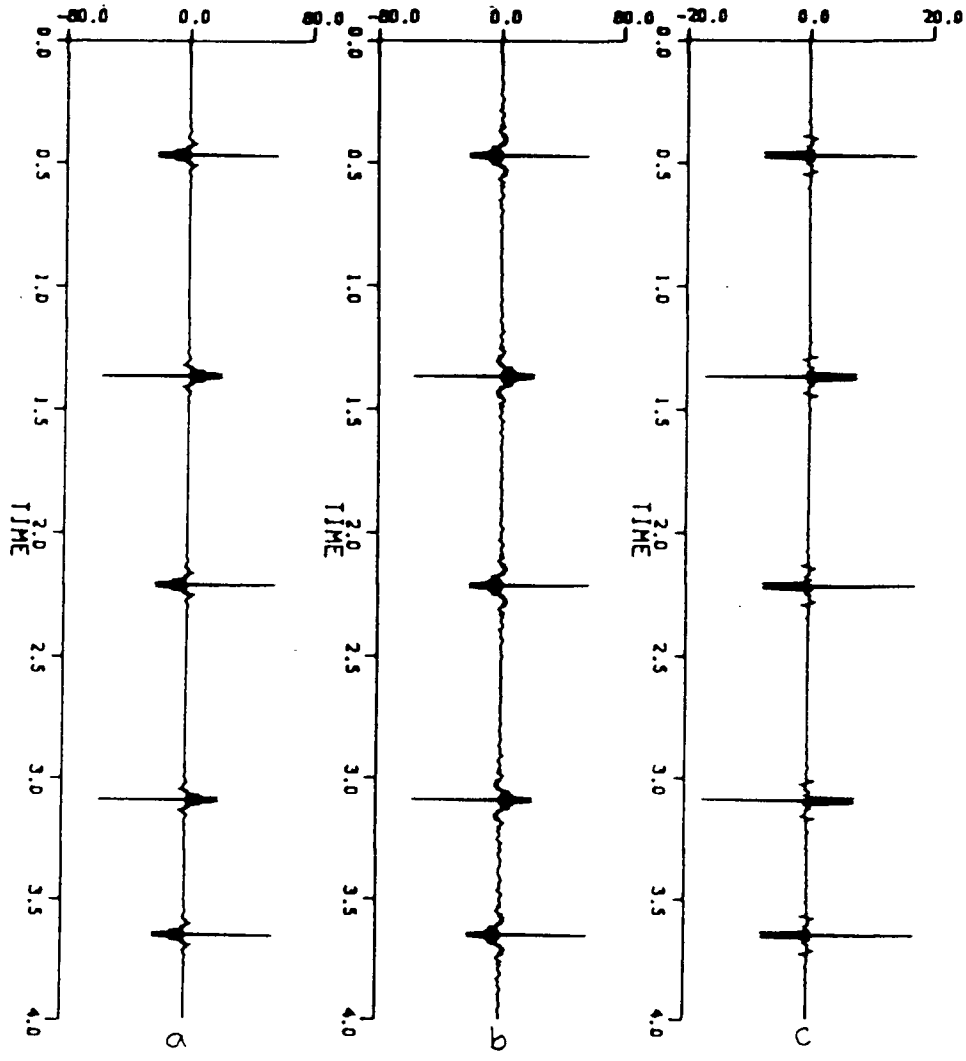


FIGURE A.1

Correlograms for synthetic seismograms:
(a) produced by a single multi-octave sweep,
(b) produced by stacking three single octave sweeps,
(c) produced by stacking three overlapping band sweeps,
all data are of the same final bandwidth. All methods result in a symmetric waveform with well defined polarity

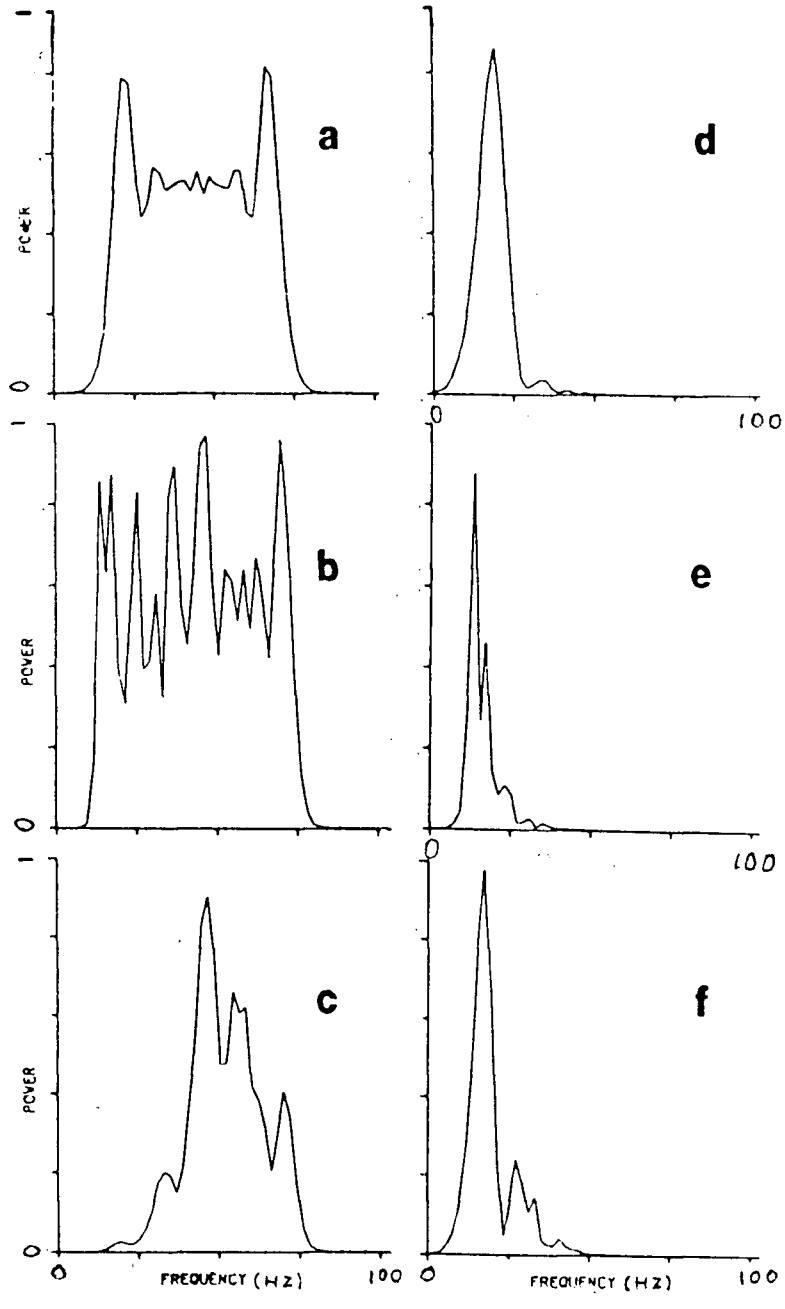


FIGURE A.2

The first three spectra correspond to the data of Figure A.1:
(a) produced by a single multi-octave sweep,
(b) produced by stacking three single octave sweeps,
(c) produced by stacking three overlapping band sweeps,
The lobes in (a) are due to Gibb's effect. In (b) the spectra were tapered and overlapped slightly to avoid gaps, and in (c) we note the pyramidal effect of stacking Combisweep data
The second triad of spectra ((d) - (e)), correspond to the attenuated versions of the seismograms which produced the first three spectra

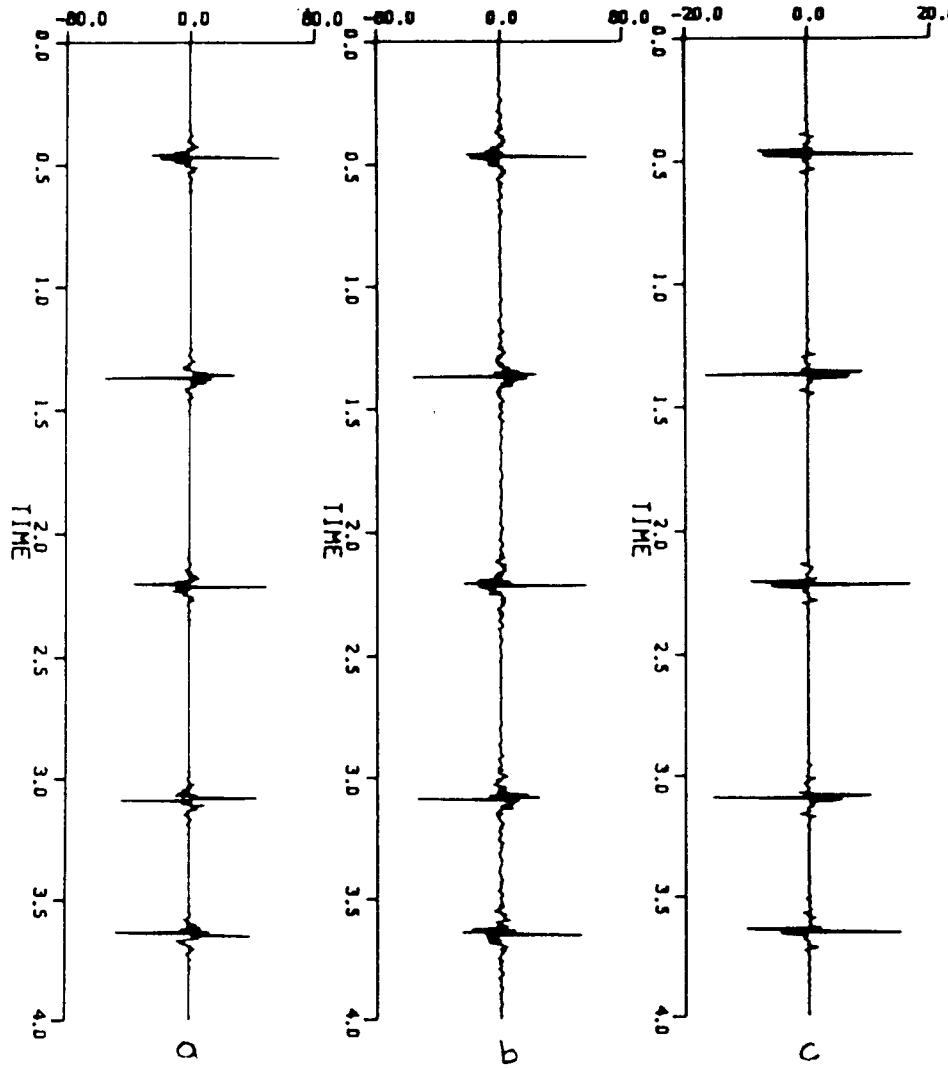


FIGURE A.3

Correlograms for dispersed synthetic seismograms (with Q=150):
(a) produced by a single multi-octave sweep,
(b) produced by stacking three single octave sweeps,
(c) produced by stacking three overlapping band sweeps,
all data are of the same final bandwidth. Polarity begins to be lost after about 1.5 s, however, the effect is much worse for the single octave sweep. The best result is for the single octave stack (b).

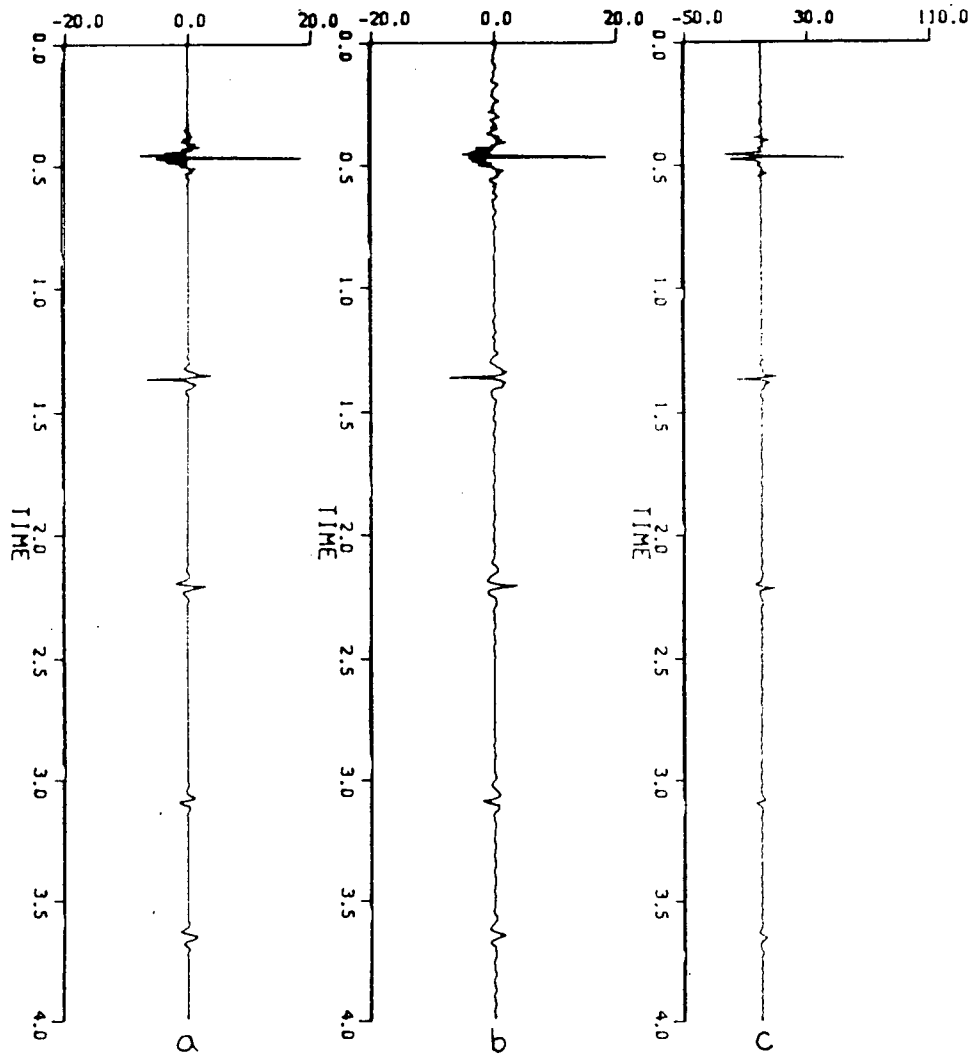


FIGURE A.4

Correlograms for attenuated dispersed synthetic seismograms (with Q=125):
(a) produced by a single multi-octave sweep,
(b) produced by stacking three single octave sweeps,
(c) produced by stacking three overlapping band sweeps,
all data are of the same final bandwidth. Polarity begins to be lost after about 1.5 s, however, the effect is much worse for the single octave sweep. The best result is for the single octave stack (b). Due to the attenuation, the energy dies down rapidly with time, and the high frequencies are lost.

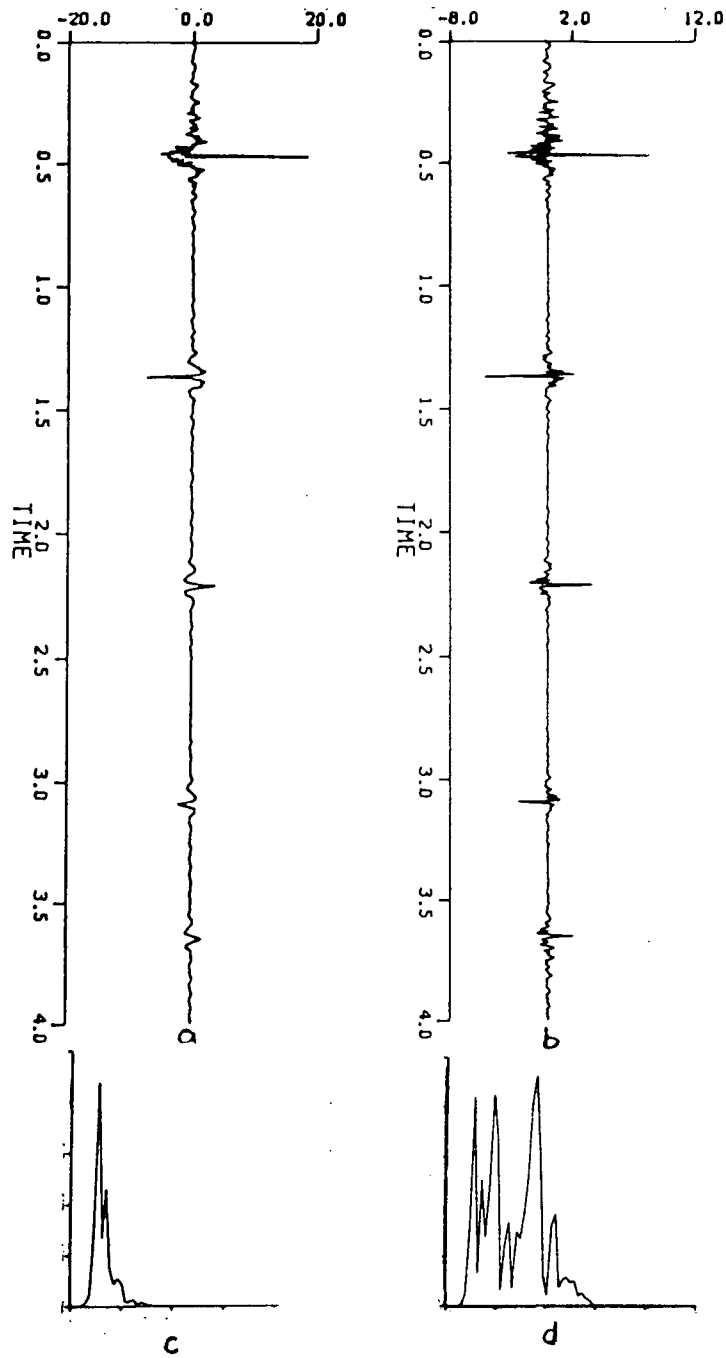


FIGURE A.5

(a) The correlogram for an attenuated dispersed synthetic seismogram produced using the single octave stack method (as per Figure A.4b)

(b) The same after recovery of the high frequencies (as described in the text). This effect may be mimicked in the field by repeating the higher frequency sweeps and then stacking the correlograms.

(c) The spectrum of (a).

(d) The spectrum of (b).

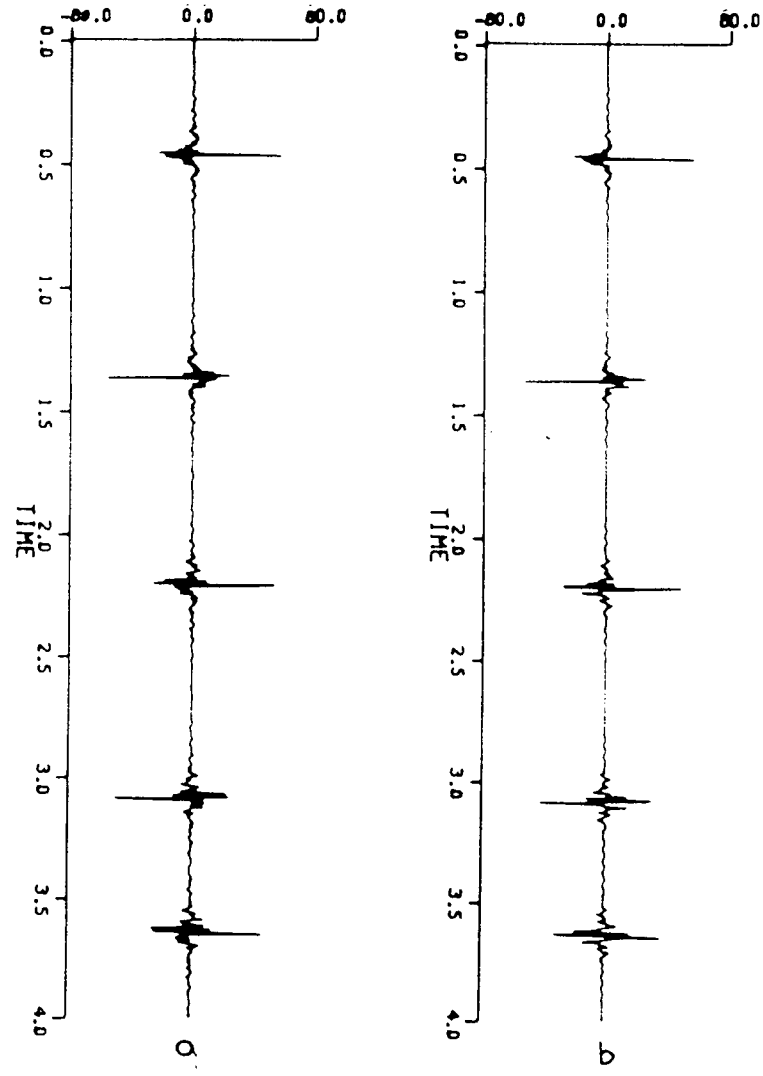


FIGURE A.6
Correlograms for dispersed synthetic seismograms (with Q=100):
(a) produced by stacking three single octave sweeps,
(b) produced by stacking two 1.4 octave sweeps,
The single octave method (a) produces a better final product, even though both data are of the same final bandwidth.

REFEREED PUBLICATIONS:

1. Jones, I.F., Mansinha, L, and Shen, P.Y., 1982, On the double exponential frequency-magnitude relation of earthquakes: Bulletin of the Seismological Society of America, v.72, p. 2373-2375.

2. Ellis, R.M., Spence, G.D., Clowes, R.M., Waldron, D.A., Jones, I.F., Green, A.G., Forsyth, D.A., Mair, J.A., Berry, M.J., Mereu, R.F., Kanasewich, E.R., Cumming, G.L., Hajnal, Z., Hyndman, R.D., McMechan, G.A., and Loncarevic, B.D., 1983, The Vancouver Island seismic project: a CO-CRUST onshore-offshore study of a convergent margin: Canadian Jounal of Earth Sciences, v.20, p. 719-741.

3. Clowes, R.M., Ellis, R.M., Hajnal, Z., and Jones, I.F., 1983, Seismic reflections from the subducting lithosphere?: Nature, v.303, p. 668-670.

4. Chapman, N.R., Levy, S., Jones, I.F., Stinson, K., Oldenburg, D.W., and Prager, B.T., 1984, Inversion of deep ocean marine seimic data to recover velocity and density information for abysaal plain sediments. Submitted to the Journal of the Acoustical Society of America.

5. Jones, I.F., and Levy, S., 1984, The application of the Karhunen-Loeve transformation in in multichannel seismic data pocessing: (invited paper for Geophysical Prospecting).

6. Levy, S., Jones, I.F., Ulrych, T.J., and Oldenburg, D.W., 1984, Complex Common Signal Analysis. In prep..

7. Ulrych, T.J., Levy, S., Oldenburg D.W., and Jones, I.F., 1984, Applications of the complex Karhunen-Loeve transformation in exploration seismology: in prep..

OTHER PUBLICATIONS:

1. Jones, I.F., 1980, Aspects of global seismicity: Master's thesis, Univ. Western Ontario.

2. Ulrych, T.J., Levy, S., Oldenburg D.W., and Jones, I.F., 1983, Applications of the Karhunen-Loeve transformation in exploration geophisics: paper S6.5, proceedings of the 53 rd. annual meeting of the Society of Exploration Geophysicists.

3. Levy, S., Ulrych, T.J., Jones, I.F., and Oldenburg, D.W., 1983, Applications of complex common signal analysis in exploration seismology: paper S6.6, proceedings of the 53 rd. annual meeting of the Society of Exploration Geophysicists.