Geophys. J. R. astr. Soc. (1976) **47**, 151–177

Source shape estimation and deconvolution of teleseismic bodywaves

Rob W. Clayton and Ralph A. Wiggins* *Department of Geophysics and Astronomy, University of British Columbia, Vancouver, British Columbia, Canada, V6T 1W5*

Received 1976 April 8

Summary. We consider the deconvolution of a suite of teleseismic recordings of the same event in order to separate source and transmission path phenomena. The assumption of source uniformity may restrict the range of azimuths and distances of the seismograms included in the suite. The source shape is estimated by separately averaging the log amplitude spectra and the phase spectra of the recordings. This method of source estimation uses the redundant source information contained in secondary arrivals. The necessary condition for this estimator to resolve the source wavelet is that the travel times of the various secondary arrivals be evenly distributed with respect to the initial arrivals. The subsequent deconvolution of the seismograms is carried out by spectral division with two modifications. The first is the introduction of a minimum allowable source spectral amplitude termed the waterlevel. This parameter constrains the gain of the deconvolution filter in regions where the seismogram has little or no information, and also trades-off arrival time resolution with arrival amplitude resolution. The second modification, designed to increase the time domain resolution of the deconvolution, is the extension of the frequency range of the transmission path impulse response spectrum beyond the optimal passband (the passband of the seismograms). The justification for the extension lies in the fact that the impulse response of the transmission path is itself a series of impulses which means its spectrum is not band-limited. Thus, the impulse response is best represented by a continuous spectrum rather than one which is set to zero outside the optimal passband. This continuity is achieved by a recursive application of a unit-step prediction operator determined by Burg's maximum entropy algorithm. The envelopes of the deconvolution are used to detect the presence of phase shifted arrivals.

1 Introduction

This paper examines the problem of the deconvolution of source functions from teleseismic recordings. The first step in any deconvolution process is the estimation of the source time function. The next section of this paper deals with this problem.

* Present address: Western Geophysical Company, PO Box 2469, Houston, Texas 77001, USA.

A standard technique in exploration seismology is to decompose the source from the seismogram autocorrelation function. This method assumes that the source is minimum phase and that the impulse sequence behaves like white noise (Robinson 1967). Since neither of these assumptions appeared to be generally valid for teleseismic recordings, the technique was not used. Source estimation by homomorphic transformation has also been suggested (Ulrych 1971), and this method is discussed in some detail. However, the low quefrency assumption of this method was found to not be generally valid, and there is also a problem of phase instabilities with the homomorphic transform itself.

Both of the above methods attempt to estimate the source from a single seismogram. We decided that on the basis of a single recording of an event it would be difficult, if not impossible, to devise a general method of estimating the source time function. For example, if an earthquake source has more than one distinct motion, then it is difficult to decide whether the effect as observed on a single seismogram belongs to the source or to the transmission path. For this reason, we decided to restrict the problem to the case where a suite of seismograms from the same event is available. The redundant source information contained in the suite will allow us to estimate the source with greater confidence than if the estimation were done on the basis of a single seismogram. Restrictions have to be placed on the distance and azimuthal ranges of the suite of seismograms so that approximate source uniformity can be assumed.

The actual deconvolution of the seismograms with a given source estimate is discussed in Section 3. Techniques such as Wiener deconvolution, which convolve the seismogram with an inverse of the estimated source, were not used because of the problems encountered when the source is not minimum phase (Robinson 1967). In the frequency domain where deconvolution is the division of the seismogram by the source, the problems with Wiener deconvolution are seen to arise from the spectral locations where the source amplitude becomes vanishingly small. The method of deconvolution by spectral division can be easily modified to overcome this problem by constraining the minimum allowable source amplitude level (Helmberger & Wiggins 1971; Dey Sarkar & Wiggins 1976). A second modification which allows increased time domain resolution of the impulse response is also outlined in that section.

Finally, the deconvolution of a real example which employs the methods outlined is presented in Section 4.

2 Source estimations

2.1 GENERAL ASSUMPTIONS

For teleseismic recordings there are no independent observations of the source time function, so it is necessary that it be estimated directly from the data. This, in general, will present some problems because the source wavelet will usually have some characteristics which overlap with those of the transmission path impulse response. Some assumptions will be necessary to resolve such problems of non-uniqueness.

The general model that will be considered for the form of each seismogram recorded from the same event is:

$$x_j(t) = s(t) * h_j(t) + n_j(t) \tag{1}$$

where $x_j(t)$ is the time series recorded at the j-th station, $s(t)$ is the source wavelet, $h_j(t)$ is the transmission path impulse response to the j-th station, $n_j(t)$ is noise, and $*$ denotes convolution.

The assumption implicit in this model is that the source time function is uniform in the sense that it is independent of both azimuth and takeoff angle. In general, this uniformity assumption is not true, as the kinematic dislocation models of radiation indicate (Savage 1966), but its validity can be improved by restricting the azimuthal angles, epicentral distances, and time ranges of the suite of seismograms. Problems with this assumption may be expected when surface reflected phases such as *pP* are mixed with main arrivals or if the total radiated source shape is radically different for different takeoff angles.

We will assume that over a suite of seismograms the various arrivals are minor perturbations of a uniform source. We seek to estimate an 'average' source which can be used in a deconvolution process to remove the major source effects of all the arrivals. The procedure usually produces visual enhancement of the recordings, which is one of the purposes of deconvolution, even though the result may not be a true representation of the Earth's impulse response. The criterion we used in judging the source estimators was the visual enhancement that deconvolution produced. There is no adequate source-shape estimation theory at present to allow the shape of a complicated estimated teleseismic source wavelet to be related to the various observed earthquake parameters.

The source uniformity assumption is not sufficient to solve equation (1) for the source wavelet. Additional assumptions about the features of either the impulse response or the source wavelet are necessary. One common assumption that is made in exploration seismology is that the source is minimum phase. This allows a unique decomposition of the source from its autocovariance function (Robinson 1967). If this assumption is coupled with the assumption that the impulse response sequence is white, then the source may be obtained by the decomposition of the estimated seismogram autocovariance (Robinson 1967). The minimum phase assumption may be true for explosive sources, but it is not generally true for earthquake sources. The example used in this paper appears to have a non-minimum phase source. The assumption of the whiteness of the impulse sequence is also invalid for a single teleseismic recording because, within a particular time window of interest, there are usually only a few arrivals.

For the case when a particular arrival on the record is separated from its neighbours by a time which is greater than the source length, then an obvious method of source estimation is simply to pick this arrival as the source. Seismograms recorded near a distance of 30 degrees often have this feature (DeySarkar 1974; Dey Sarkar & Wiggins 1976). However, one has to be confident that there is only one arrival in the particular time window, and that the window contains all of the source. Also, if the source wavelet is not uniform, then this estimate may not be the 'best-fitting' one.

A simple extension of this 'simplest arrival' method is the averaging of a particular arrival over several traces after suitable time alignment. The essential assumption of this method is that all other arrivals present within the time window defined by the source wavelet length must average to zero. In other words, the other arrivals must have sufficiently different phase velocities (moveouts) compared to the arrival which is to be enhanced. This method uses the redundant source information in a particular arrival over the suite but not in the secondary arrivals on the seismograms. The resulting deconvolution will necessarily be biased toward the particular arrival on which the source estimate was based.

Another method of source estimation, which uses homomorphic transforms, has been proposed (Ulrych 1971; Stoffa, Buhl & Bryan 1974). This method replaces the minimum phase source and white impulse response assumptions by a low quefrency source assumption. This assumption also appears to be unrealistic for complex earthquake sources. However, this method will be discussed in some detail because it forms the basis of the estimator proposed in this paper.

2.2 SOURCE ESTIMATION BY HOMOMORPHIC TRANSFORMS

The concept of a homomorphic transform from a convolutional space to an additive space was first proposed by Oppenheim (1967), and was elaborated upon by Schafer (1969). Ulrych (1971) has suggested that it is a suitable method for seismic source estimation.

The basic idea of the method is to transform the convolution operator of equation (1) into an additive operator so that the system may be treated by linear filtering theory. The transformation that accomplishes this and its inverse are given below (Oppenheim 1967).

$$\hat{x} = H(x) = F^{-1}\{\log[F(x)]\} \tag{2}$$

$$x = H^{-1}(\hat{x}) = F^{-1}\{\exp[F(\hat{x})]\} \tag{3}$$

where $H(\cdot)$ denotes homomorphic transformation, $F(\cdot)$ denotes Fourier transform, and $\hat{x}$ is the homomorphic transform of x.

The cepstrum $\hat{x}$ defined here is usually termed the complex cepstrum (denoting the usage of the complex logarithm), to distinguish it from the cepstrum defined by Bogert *et al.* (1963) which uses only the real part of the logarithm. Since the 'complex cepstrum' is the only one used in this paper, it will be referred to as simply the cepstrum.

The phase information of x is preserved under the transform because the complex logarithm is used. The homomorphic transform of equation (1), for the case when there is no additive noise is:

$$\hat{x} = H(x) = H(s * h) = H(s) + H(h) = \hat{s} + \hat{h} \tag{4}$$

Additive noise causes special problems under this transform because there is no simple expression for the homomorphic transform of the addition operator. Additive noise predominately affects the phase part of the transform. This aspect will be discussed later.

To separate $\hat{s}$ and $\hat{h}$ in equation (4) an assumption is needed about their cepstral relationship. The usual assumption is that the source cepstrum $\hat{s}$ has a lower quefrency content than the impulse response cepstrum $\hat{h}$ (Schafer 1969; Ulrych 1971). Quefrency is the independent variable of the cepstral domain (i.e. it is the 'frequency' content of the log spectra) and has the dimension of time. If this assumption is true, then the source may be approximately recovered by filtering out the higher order quefrency terms. This assumption is illustrated by the example of Fig. 1. In this example the source is approximately recoverable from its cepstral representation by retaining only the lower quefrency components, while the impulse response is approximately recoverable by retaining the higher quefrency terms of its cepstrum.

However, the low quefrency assumption is not considered to be sufficiently valid for source estimation of teleseismic events. If the z-transform of the source function happens to contain a zero near the unit circle, then the amplitude and phase spectra will change quite rapidly in the vicinity of this zero (Claerbout 1976). In the cepstral domain this translates into higher order quefrency terms for the source. The amplitude spectra of a suite of 16 recordings of an event from Western Australia were averaged together to produce the amplitude spectrum shown in Fig. 2. The presence of the hole (marked by an arrow) indicates that there is a zero of the z-transform near the unit circle. Since this feature appears in all of the amplitude spectra it is assumed to be a source effect and not part of the impulse response. This hole makes contributions to the source cepstrum at quefrencies of 1.1 s and higher. Since the impulse cepstrum is expected to have contributions in this range, the source cepstrum cannot be effectively recovered by just filtering out higher quefrencies.

Several practical problems exist in the computation of homomorphic transforms. The

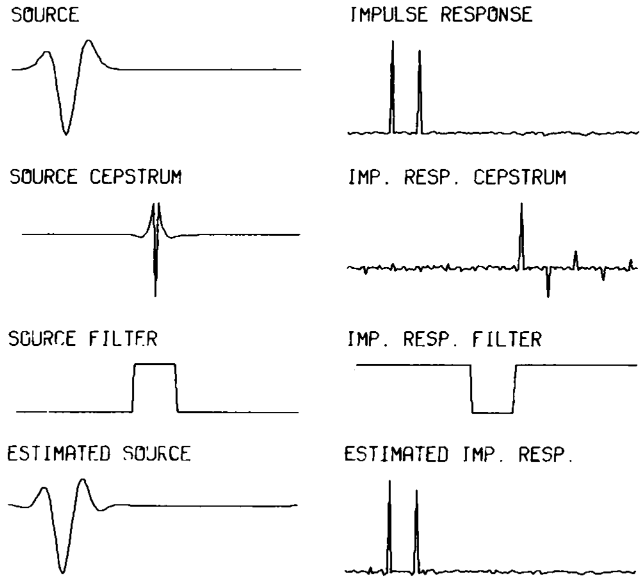

Figure 1. An illustration of the low-quefrency source assumption. The cepstra of a simple source and impulse are shown. The source cepstrum is low-passed with the filter shown while the impulse response cepstrum is high-passed with the complement of the source filter. The result shows, that for this example, the source is approximately a low quefrency process while the impulse response is a high quefrency process.

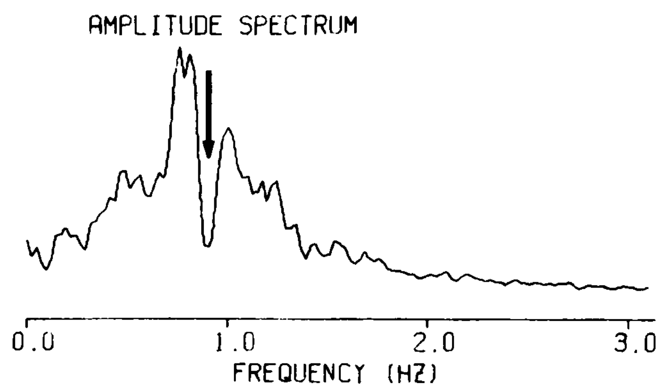

Figure 2. An example of a hole in the source amplitude spectrum. The amplitude spectrum shown is the average of 16 seismograms of the West Australia event of 1970 March 24 recorded in the range 113° to 150°. The hole (marked by an arrow) is assumed to be part of the source rather than a transmission path effect because it appears in all amplitude spectra of the suite at the same frequency position. The assumption illustrated in Fig. 1 is invalid for this case because the hole contributes to the high quefrency terms to the source capstrum.

main problem is the unwrapping of the phase curve. The complex logarithm is not unique with respect to its imaginary component (the phase) so a factor of $\pm 2n\pi$, $n = 0, 1, \ldots$, must be added at each point to make the phase curve continuous (Schafer 1969). The usual procedure for unwrapping the phase curve is to compare the difference between the phase at a particular lag with the phase at the preceding lag. If this difference is greater than π, then a $\pm 2\pi$ factor is added to the phase to make the difference less than π (Schafer 1969). This procedure is started at zero frequency where the true phase is known to be zero. If the time series has a negative DC component, then it is inverted before the transform is taken (Schafer 1969). A linear phase component, which corresponds to a shift in time of the signal, is removed from the unwrapped phase curve to prevent the linear component from dominating the cepstrum.

Additive noise can make the unwrapping procedure unstable for two reasons. The first reason is that the phase deviation due to additive noise can become large when the signal amplitude is low. To illustrate this, suppose that a particular sample of the signal spectrum has Gaussian noise added to its real and imaginary parts. Fig. 3 illustrates two cases. When the signal amplitude is high, the phase deviations remain at a reasonable level. However, when the signal amplitude drops, the phase deviation becomes quite large. In the case where the signal amplitude drops to zero, the amplitude deviations assume a Rayleigh distribution and the phase has a uniform distribution over $[-\pi, \pi]$ – Bracewell (1965, p. 335). Thus, for low amplitude regions of the signal spectrum, the phase can have a large variance.

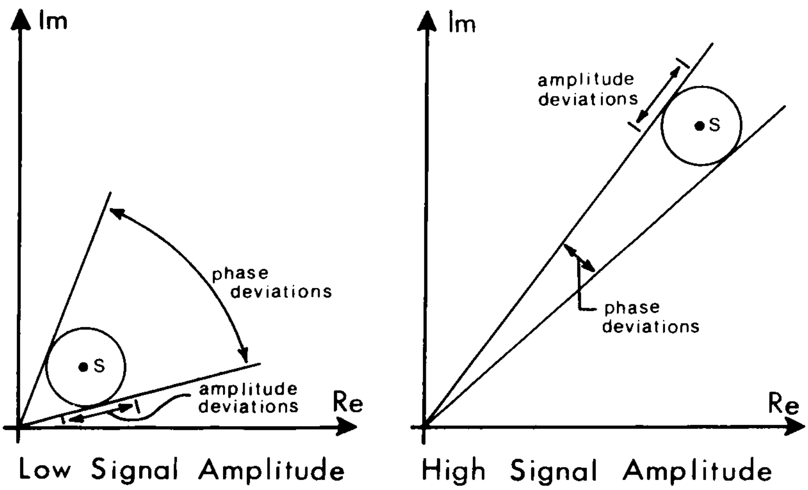

Figure 3. The effect of additive noise on amplitude and phase. The circle surrounding the signal point S approximates one standard deviation of additive noise. The corresponding amplitude and phase deviations are shown. As the amplitude of S becomes small, the amplitude deviations remain constant but the phase deviations become large.

The second problem also occurs at low amplitude regions of the spectrum. A hole in the amplitude spectrum indicates the presence of a zero in the z-transform near the unit circle. Suppose that the true signal has a zero just inside the unit circle and that additive noise forces it just outside the unit circle. The phase of the true signal would increase by π near the zero, but when the noise is added the phase decreases by π at this point. Fig. 4 shows this effect on the unwrapped phase as well as on the unwrapped phase with the linear trend removed. It is obvious that a change in location of zeros near the unit circle can drastically alter the shape of the unwrapped phase curve with the linear component removed.

Several methods were tried in an attempt to remove the phase instability at low amplitudes. Although none of them worked sufficiently well to be deemed useful, they are listed here because the problem is usually ignored in the literature.

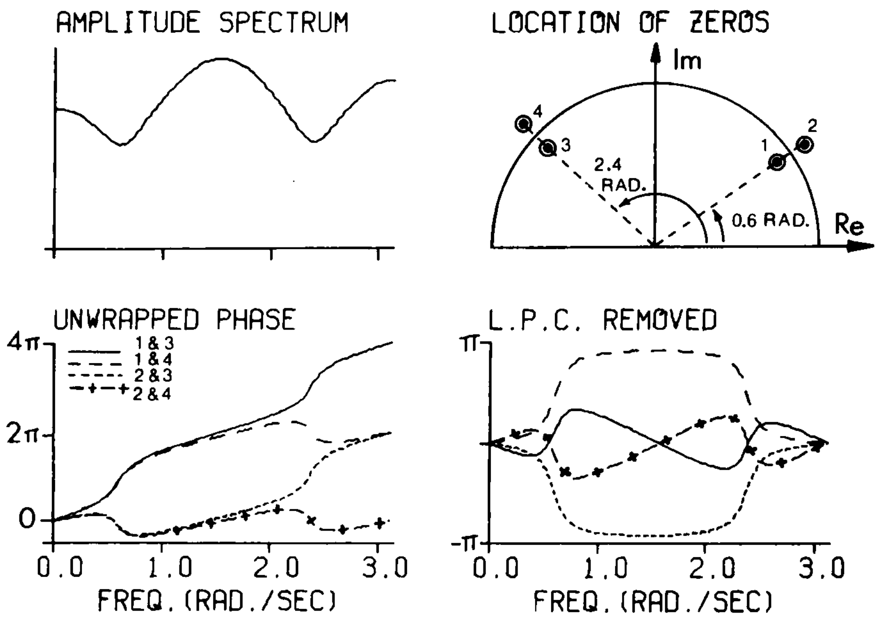

Figure 4. The effect of changing a zero location on the phase curve. The presence of a hole in the amplitude spectrum of a function indicates there is a zero near the unit circle in the z-transform of the function. If additive noise causes one of these zeros to move across the unit circle line, then the phase deviations become large. This figure illustrates the four different phase curves for a pair of zeros moved in this fashion. The amplitude spectrum of the four phase curves is the same. The phase deviations are more pronounced when the linear phase component (LPC) is removed as is done with homomorphic transforms.

(1) The phase was estimated by interpolation from its neighbouring points when the amplitude dropped below a certain level. Two problems make this method impractical. First, it is difficult to define a cutoff level which rejects all poorly-estimated phase terms while retaining a sufficient number of well estimated ones to make interpolation worthwhile. Secondly, the method tends to remove all π jumps, even the ones that should be present.

(2) Stoffa *et al.* (1974) suggested computing the derivative of the phase directly from the real and imaginary parts of the Fourier transform and then integrating to obtain the unwrapped phase curve. If the Fourier transform is $X(\omega) + iY(\omega)$, then the phase derivative is: (Stoffa *et al.* 1974).

$$\Phi'(\omega) = [Y(\omega)X'(\omega) - X(\omega)\,Y(\omega')]/A(\omega)^2 \qquad (5)$$

where $A(\omega)$ is the amplitude spectrum. The phase derivative also becomes unstable at low amplitudes because of the $A^2(\omega)$ term in the denominator. Also, integrating the phase derivative in regions of rapid phase changes can be a problem, because of poor sampling of the phase derivative. Interpolating the phase derivative over low amplitude frequencies tends to remove all π jumps, as with the first method.

Essentially, the problems with the phase unwrapping techniques reduce to the following. While it is possible to unwrap the phase in regions of high spectral amplitude, there appears to be no way of connecting the unwrapped sections of the phase curve across regions of low spectral amplitude.

We conclude that the filtered cepstrum is not a good source estimator when it is used on a single recording. The phase estimate is unstable in the presence of noise and there is doubt of the general validity of the low quefrency assumption.

More successful results are obtained by averaging the log amplitude and phase spectra of a number of records from the same event. This tends to make the phase more stable and it frees the method from the low quefrency assumption. This estimator will be discussed in the next section.

2.3 SOURCE ESTIMATION BY AVERAGED LOG AMPLITUDE AND PHASE SPECTRA

A method of averaged log amplitude and phase spectra is proposed as a technique for freeing the single trace homomorphic transform from the low quefrency assumption and for improving the phase stability. Since the expectation operator is commutative with the Fourier transform, averaging the log amplitude and phase spectra is equivalent to averaging the cepstra. When the phase is averaged, the raw phase and not the unwrapped phase is used, so that the instabilities of the unwrapping procedure can usually be avoided. The object of the averaging is to enhance the source part of the cepstrum while diminishing the impulse response part. The fundamental assumption made here is that although an individual seismogram has non-random arrival times, the arrival times in a suite of seismograms considered together are sufficiently random for cepstral averaging to reduce the impulse response effect. If the averaging succeeds in eliminating most of the impulse cepstrum, then the low quefrency assumption need not be applied to recover the source.

The impulse response of an ideal earth (a non-attenuating, non-dispersive, and non-phase-shifting earth) is a sum of Dirac impulses.

$$h_t = \sum_{n=1}^{N} C_n \, \delta(t - T_n) \tag{6}$$

where C_n is the amplitude of the n-th arrival and T_n is its arrival time.

Attenuation and dispersion preferentially modify the frequency components of the source wavelet as it travels through the Earth. Any source estimation method which attempts to find an 'average' source wavelet over a suite of seismograms will out of necessity also include an 'average' attenuation and dispersion factor. We will consider attenuation and dispersion phenomena, apart from the 'average' effect which will be included with the source wavelet, to be minor perturbations on the form of equation (6). If phase-shifting were included in the impulse response model, then a phase term dependent on the summation index would have to be included in the sinusoidal terms of the following three equations. The omission of this effect does not affect the arguments based on these equations. The principal effect of variations in the phase of various arrivals is that the estimated source may have a uniform phase delay. The subsequent use of the envelope of the deconvolved traces nullifies the effects of uniform phase shifts.

The Fourier transform of the impulse response model is

$$H(\omega) = \sum_{n=1}^{N} C_n \exp(-i\omega T_n) \tag{7}$$

The log amplitude ($\log A(\omega)$) and phase ($\Phi(\omega)$) of $H(\omega)$ are

$$\log A(\omega) = \tfrac{1}{2} \log \left\{ \sum_{n=1}^{N} C_n^2 + 2 \sum_{n=1}^{N} \sum_{m>n}^{N} C_n C_m \cos \omega (T_n - T_m) \right\} \tag{8}$$

$$\Phi(\omega) = \tan^{-1}\left[\frac{\sum_{n=1}^{N} C_n \sin \omega T_n}{\sum_{n=1}^{N} C_n \cos \omega T_n}\right] \tag{9}$$

The log amplitude will have its main quefrency contributions at all combinations of the differences of the various travel times. The average of the log amplitude spectra over a suite of impulse responses will tend to a constant, as the number of members in the suite is increased, if the combinations of travel time differences are sufficiently uniformly distributed among the impulse responses. In the time domain, this means that the various arrivals must have differing phase velocities (moveouts). The constant term

$$\sum_{n=1}^{N} C_n^2$$

of the log amplitude is the average of the energy in each impulse response. This will introduce a multiplicative constant into the source estimate which is not important for deconvolution. The phase spectra will also have contributions which correspond to the various travel times so that an average of the phase over a suite of impulse responses will tend to zero under the same conditions as the amplitude spectra. In general the impulse response components of the seismogram cepstra will not average to zero, but if they are reduced sufficiently, the source wavelet will be enhanced when the cepstrum is transformed to the time domain. It is usually necessary to apply a time window to the source wavelet to remove any remaining effects of the impulse response.

To form the source estimate, the log amplitude and phase spectra of a number of records are averaged together. We found that the best results were obtained, as would be expected, when noisy records or records with obviously different sources were excluded. A sufficient number of 'good' recordings must be retained to make the averaging process effective. Some care is necessary when the phase is averaged because of the arbitrary branch cut in the arctangent function at $\pm\pi$. This break would only be suitable if the expected value of the source phase is zero. However, if the expected value is not zero, then the branch should be placed π radians above and below this expected value. An initial estimate of the phase is necessary to establish the correct interval over which to average. Each phase value is then placed in this interval by adding the appropriate $\pm 2\pi$ if its difference from the initial estimate is greater than π.

The initial phase estimate is taken as the phase value whose corresponding spectrum is closest to the average amplitude spectrum at each frequency sampling point. This is the phase that is most likely to be independent of the phase spectrum of the impulse response.

It is necessary that scale factor differences among the amplitude spectra be removed before the averaging in order that the initial phase estimate is not biased by energy differences in the seismograms. Rather than normalizing the trace amplitudes by the peak trace height or the peak spectral amplitude (both of which could be strongly affected by the impulse response) a multiplicative constant for each trace was determined by least squares. The first trace was picked as a reference and the following least squares problem was solved for the $\{C_j\}$.

$$e^2 = \sum_{\omega} \{A_1(\omega) - C_j A_j(\omega)\}^2 \tag{10}$$

where $A_j(\omega)$ is the amplitude spectrum of the j-th recording.

The summation is taken over the frequency passband of the recordings. The solution of C_j which minimizes the error is

$$C_j = \sum_\omega A_1(\omega) A_j(\omega) \Big/ \sum_\omega A_j^2(\omega) \tag{11}$$

The amplitude spectra of all other traces are then normalized by the C_j factors.

$$A'_j(\omega) = A_j(\omega) \cdot C_j \qquad j = 2, \ldots, N \tag{12}$$

After averaging the log amplitude spectra and phase spectra, the low quefrency source assumption may be applied, if it is desired. To do this, the final inverse Fourier transform of the homomorphic transform is taken and the resulting cepstrum is zeroed above some quefrency level. However, as mentioned earlier, the low quefrency assumption is a dubious one for complex earthquake sources. The assumption, if it used at all, must therefore be applied with caution. Furthermore, since the impulse response is diminished by the averaging process, there will generally be no clear indication of the correct upper quefrency limit.

If there is a secondary arrival on the seismograms that does not move out significantly with respect to the primary signal, then this method has trouble separating the source from the combination of the two arrivals. If a source is estimated and it appears to have two separate parts, then a decision must be made whether the source estimate actually contains two separate arrivals or whether both parts belong to the source. The easiest method of deciding this question is to check for the presence of holes (*cf.* Fig. 2) in the amplitude spectrum of each recording. The frequency of a hole corresponds to the time separation of the two apparent components. If this feature is present on all records and its frequency position does not change, then it may be concluded that, within the resolving power of the suite of seismograms, both parts belong to the source. Examining the amplitude spectra in this manner is more reliable than looking directly at the records themselves for apparent moveout because; first, the feature is usually more obvious in the amplitude spectra, and second, the amplitude spectra uses the redundant information of all the arrivals on the record.

Several similar source estimators can be formulated along the same lines as the one given above. For example, a simple average of the amplitude spectra could be substituted for the average of the log amplitude spectra. The median of the amplitude spectra could also be used. The arguments given earlier that the impulse response contributions will cancel out if the various arrivals on the seismograms have different phase velocities also apply to these estimators. If the source wavelet shape is quite uniform over the whole suite of seismograms then the performances of the estimator outlined in this paper, and the two mentioned above are similar. However, we have found that when a non-uniform source is present, deconvolution with the source estimated by average log spectra and phase spectra gives a superior enhancement of the seismograms.

3 Deconvolution of the seismograms

A method will now be outlined for the deconvolution of the individual seismograms with a given source estimate. The desirable properties of a deconvolution method are: (1) it resolves the arrival times sharply, and estimates the arrival amplitudes accurately, (2) it should be stable with respect to small source estimation errors or source non-stationarity, and (3) it should be robust with respect to the rejection of random noise.

The basic method that will be used is that of divisional deconvolution in the frequency domain, because it does not require assumptions about the form of the source wavelet (i.e.

that it be minimum phase). Two modifications will be made to improve the stability and resolution of this method.

The first modification will be the introduction of a minimum allowable source amplitude level (Helmberger & Wiggins 1971; Dey Sarkar 1974; Dey Sarkar & Wiggins 1976), termed the waterlevel, to reduce spurious noise components and the effect of small errors in source estimation. The second modification is the bandwidth extension of the impulse response beyond its optimum passband, to improve the time domain resolution of the deconvolution.

The last part of this section deals with the envelope of deconvolution, which provides a useful tool for the detection of phase-shifted arrivals.

3.1 THE WATERLEVEL PARAMETER

The frequency domain form of the seismogram model – equation (1) – is

$$X_j = S \cdot H_j + N_j \tag{13}$$

where the capitals denote the Fourier transform pairs of the quantities in equation (1). To obtain the estimate of the impulse response $\tilde{H}_j$, the spectrum of the seismogram is divided by the estimated source $\tilde{S}$.

$$\tilde{H}_j = \frac{X_j}{\tilde{S}} = \left[\frac{S\tilde{S}^*}{|\tilde{S}|^2}\right] H_j + \left[\frac{\tilde{S}^*}{|\tilde{S}|^2}\right] N_j \tag{14}$$

where * denotes conjugation.

As the estimated source amplitude becomes small ($|\tilde{S}| \to 0$), the factor multiplying H_j is $0(1)$, assuming that $\tilde{S}$ does not deviate too far from S. However, the factor multiplying the noise component N_j is $0(1/|\tilde{S}|)$. Therefore, it is essential to establish a minimum amplitude level for the source to prevent the noise term from becoming too large. This level can be thought of as a limit on the gain of the filter $1/\tilde{S}$ in the spectral regions where the seismogram has little or no information. The minimum source amplitude is termed the waterlevel and is conveniently expressed as a fraction of the maximum source amplitude. With this modification the estimator of H becomes

$$\tilde{H}_j = \frac{S\tilde{S}^* H_j + N_j\tilde{S}^*}{\max\{|\tilde{S}|^2, (k|\tilde{S}|_{\max})^2\}} \tag{15}$$

where k is the waterlevel parameter ($0 \leqslant k \leqslant 1$) and $|\tilde{S}|_{\max}$ is the maximum source amplitude.

The deconvolution will be stable with respect to small errors in the source estimation if the factor

$$\frac{S\tilde{S}^*}{\max\{|\tilde{S}|^2, (k|\tilde{S}|_{\max})^2\}} \tag{16}$$

does not become unstable as a function of frequency. To show the effect of the waterlevel parameter on this type of error let the real source be

$$S = \tilde{S} + cR \tag{17}$$

where R is an arbitrary function of unit maximum amplitude and c is a scale factor which indicates the degree of error in estimation. The factor (16) now becomes

$$\frac{(\tilde{S} + cR)\tilde{S}^*}{\max\{|\tilde{S}|^2, k^2|\tilde{S}|^2_{\max}\}} = \frac{|\tilde{S}|^2 + cR\tilde{S}^*}{\max\{|\tilde{S}|^2, k^2|\tilde{S}|^2_{\max}\}} \tag{18}$$

As with the noise, the factor multiplying cR is $O(1/|\tilde{S}|)$. Therefore, the introduction of the waterlevel also prevents this factor from becoming too large, if c is sufficiently small.

An optimum choice for k would be one which makes $k|S|_{\max}$ greater than both the noise level and the deviations of the estimated source from the real source. Since these factors cannot be determined in practice, an alternate procedure is suggested. Rather than attempt to pick an optimum k, we suggest that the deconvolution be performed for a range of $k \in [0, 1]$. The stability of the deconvolution can be checked by comparing the impulse response for the various waterlevels. This is the procedure that is used for the examples in this paper.

The waterlevel parameter has another interesting interpretation. As k approaches zero, the estimator becomes an unrestricted deconvolution of X_j by $\tilde{S}$. As k approaches unity, the estimator is just a scale factor times the crosscorrelation of X_j and $\tilde{S}$. The unrestricted deconvolution attempts to remove all of the source effects from the seismogram, thus making it the best estimator of the true impulse response. This form of the estimator will be the best for resolving the travel times. The crosscorrelation is the least squares estimate of the arrival amplitudes (Helmberger & Wiggins 1971). If a particular phase arrives at time T, then the squared error due to amplitude differences between the arrival and the estimated source is

$$e^2 = \sum_t \{x(t - T) - A\,s(t)\}^2 \tag{19}$$

The amplitude factor A which minimizes this error is

$$A = \sum_t s(t)\,x(t - T) \Big/ \sum_t s(t)^2 \tag{20}$$

which is just a scale factor times the crosscorrelation of $x(t)$ and $s(t)$, evaluated at time T. If $s(t)$ is normalized to unit maximum amplitude, then A is the arrival amplitude. To obtain A from the divisional deconvolution with $k = 1$, the peaks are multiplied by the maximum source spectral amplitude squared and divided by the energy of the source. The waterlevel can therefore, also be interpreted as a parameter which trades off amplitude resolution with arrival time resolution. This is added incentive to perform the deconvolution for a range of waterlevels.

3.2 EXTENSION OF THE IMPULSE RESPONSE SPECTRUM

The second modification that was introduced into the divisional deconvolution method was an impulse response spectral extension scheme designed to increase the time domain resolution of the estimated impulse response. This scheme consists of assuming a model for the impulse spectrum and predicting the unknown spectral components on the basis of this model. In the section on source estimation it was shown that the assumption of source stationarity led to an impulse response which was a sum of Dirac impulses for a non-dispersive, non-phase shifting earth – equation (6). The Fourier transform of this relation is

$$H(\omega) = \sum_{n=1}^{N} C_n \exp\{-i\omega T_n\} \tag{21}$$

which is just a sum of N complex sinusoids.

When divisional deconvolution is performed, it is necessary to filter the impulse response to a passband defined by the band of significant energy of the particular seismogram. If this were not done, then spurious noise information could be included in the impulse estimate as is shown in the example of Fig. 5. The large spikes at 1.1 Hz and 1.4 Hz are noise components which are enhanced by divisional deconvolution. The ringing nature of the unfiltered impulse response is due to these spikes. When the impulse response is filtered with the indicated passband, the result is much more stable. This approach, however, makes a zero extension assumption about the nature of the impulse response outside the passband and, in view

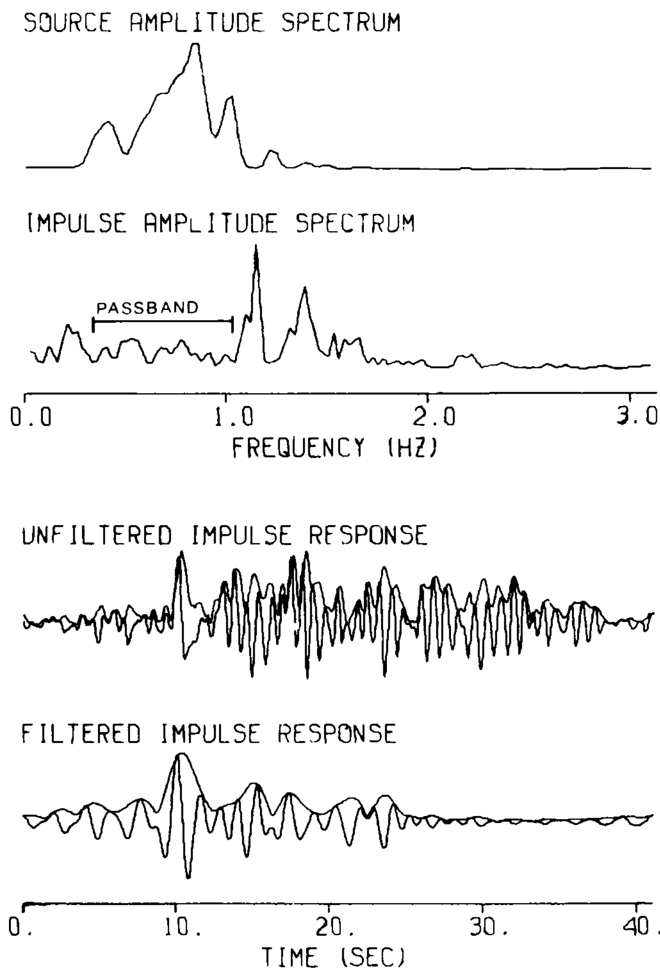

Figure 5. The effect of choice of passband on the impulse response. This example is the AYSD recording of the Kern County event. When the deconvolution is not restricted to a passband of significant source energy, the estimated impulse response is dominated by the enhanced noise spikes. The result is much more stable when the deconvolved impulse response is filtered by the indicated passband.

of the model of the impulse spectrum – equation (21) – this does not seem to be a particularly good assumption. A better assumption would be one which fills in the impulse spectrum in accordance with the continuous nature of this model.

A procedure for extending short realizations of a continuous process by the maximum entropy method (Burg 1967), has been given by Smylie, Clarke & Ulrych (1973) and Ulrych & Bishop (1975). A prediction filter is found by fitting the following autoregressive process to the known data

$$x_j = \sum_{k=1}^{p} a_k x_{j-k} + z_j \tag{22}$$

where $\{a_k\}$ are the coefficients of the prediction operator, p is the order of the autoregressive process (the length of the prediction operator), and z_j is a white noise process termed the innovation.

An alternate way of writing the autoregressive model is in terms of the prediction error operator:

$$\sum_{k=0}^{p} \gamma_k x_{j-k} = z_j \tag{23}$$

where $\gamma_k = -a_k$, $k = 1, \ldots, p$, and $\gamma_0 = 1$.

For Burg's maximum entropy algorithm the process is fitted by minimizing the variance of the innovation when the prediction error operator is applied in both the forward and reverse directions (Smylie *et al.* 1973). The method will always produce a minimum phase prediction error operator (Burg 1967; Claerbout 1976).

For the problem of extending the frequency range of the impulse response, the $\{x_j\}$ of equations (22) and (23) is the Fourier spectrum of the impulse response. The minimum phase property of the prediction error operator is ideal for extending the impulse spectrum because the energy of the predicted part of the impulse spectrum will not increase away from the known passband. Since the impulse spectrum is complex, the complex form of Burg's algorithm will be required to compute the prediction operator. The derivation of the complex Burg algorithm follows along the same lines as the real version, except that it must be realized that the prediction operator is conjugated for reverse prediction (Smylie *et al.* 1973).

It could be argued that since the impulse spectrum model is a sum of sinusoids, it would be better to fit a sinusoidal model rather than an autoregressive model to the known data. The problem with this approach is that the sinusoidal model can only be assumed to be approximately valid, due to attenuation and phase shifts of the arrivals. The autoregressive model makes the more general assumption that the given impulse passband is part of a continuous process which covers the entire spectrum. If the spectrum is truly a sum of sinusoids, the autoregressive model does an excellent job, as long as there is some additive noise present (Ulrych & Bishop 1975).

Once the prediction operator has been determined by the maximum entropy method, it is applied in a unit step prediction fashion, until the entire spectrum is filled in. An example of spectral extension is shown in Fig. 6. The trace shown in Fig. 6(a) was generated by band-passing an impulse sequence consisting of three impulses. The real part of the Fourier transform of this trace is shown in Fig. 6(b). The real part and its associated complex part of the Fourier transform inside the passband were used to construct a prediction operator which was then used to extend the spectrum. The real part of the extended spectrum is shown in

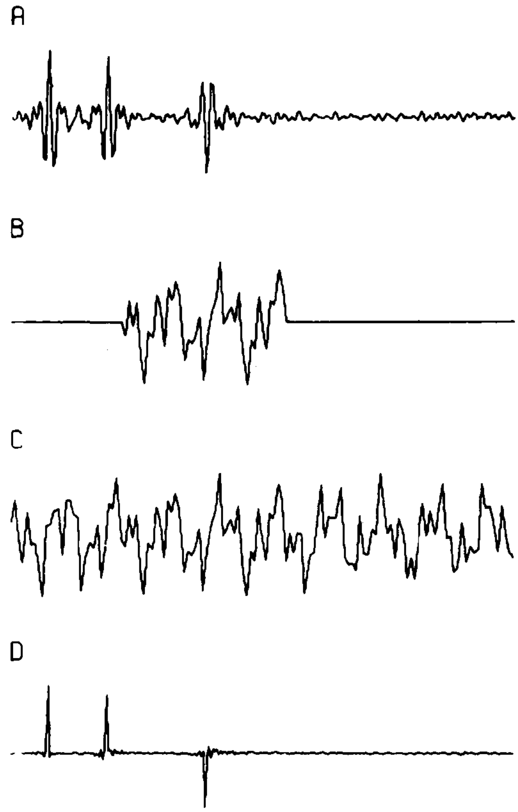

Figure 6. Spectral extension applied to a simple impulse response. (A) A synthetic trace consisting of three impulses which was bandpassed with an ideal filter to reduce their resolution. (B) The real part of the Fourier transform of (A). A complex prediction operator which was fitted to the part of the transform within the passband, was used to extend the spectrum by unit-step prediction. (C) The real part of the extended spectrum. (D) The result of transforming the extended spectrum back into time domain. A comparison of (A) and (D) indicates the increase in resolution that is achieved.

Fig. 6(c). When the extended spectrum is transformed back into the time domain (Fig. 6(d)), the increase in resolution of the impulse sequence is obvious.

The amplitude of the predicted oscillations of the spectrum always decreases away from the optimal impulse passband because: first, the prediction error operator is minimum phase, and second, the innovation of the autoregressive model is not included in the prediction.

The one parameter that has to be chosen for this scheme is the length of the prediction operator. The same problem occurs in the estimation of power spectra by the maximum entropy method (Ulrych & Bishop 1975). One criterion that has been suggested for optimally choosing this parameter is that of the final prediction error (Akaike 1969; Fryer, Odegard & Sutton 1975; Ulrych & Bishop 1975). This criterion constructs a trade-off curve between the minimum of the innovation variance and the variance of the estimate of the prediction coefficients (Akaike 1969). The minimum of this curve is taken as the optimal autoregressive order. However, observational experience indicates that for sinusoidal type processes, significant resolution can be obtained by increasing the autoregressive order beyond that indicated by the final prediction error criterion. The penalty that one pays for doing so is the possible introduction of spurious noise components into the estimate. Instead of attempting to find an optimum autoregressive order, we decided to use the same procedure as with the waterlevel parameter. That is, the deconvolution was performed for a number of prediction lengths and the stability of the deconvolution is checked by comparing the various results.

The impulse response passband for the calculation of the prediction coefficients was determined by first performing a preliminary deconvolution in a passband determined from the main spectral energy of the source estimate. If any large noise spikes were found to be present, the passband was reduced to exclude them. Fig. 5 shows an example of these noise spikes and the passband chosen to exclude the large spikes.

3.3 ENVELOPES OF DECONVOLUTION

Several authors have demonstrated that phase shifting of arrivals can introduce systematic errors into travel time and amplitude observations by changing the relative peak and trough positions of the arrival wavelets (Choy & Richards 1975; Hill 1974; Helmberger & Wiggins 1971). A simple procedure to overcome this problem is to compute the envelope of the deconvolution (Helmberger & Wiggins 1971; Farnbach 1975).

The envelope is defined as the modulus of the analytic signal (Bracewell 1965, pp. 267–272) and can easily be shown to be independent of phase shifts (Clayton, McClary & Wiggins 1976). A useful way of displaying the envelope to show the phase properties of the signal is to superimpose it on the deconvolution. Fig. 7 shows the typical relationships between an impulse and its envelope for various phase shifts. A 90 degree phase shift, for example, is characterized by a trough followed by a peak of equal amplitude.

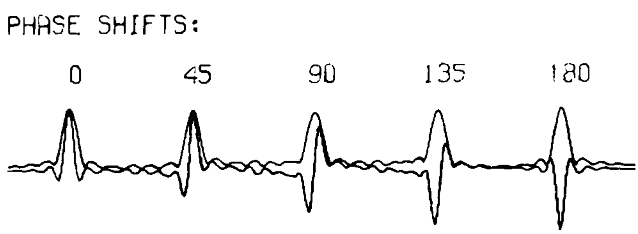

Figure 7. The relationship between a phase-shifted impulse and its envelope. Several band-passed, phase-shifted impulses are shown with their superimposed envelopes, for various choices of phase shift angle. The envelope may be used as a reference for gauging the phase shift angle of a particular impulse because it is independent of the phase shift angle.

4 An example of deconvolution

The source estimation and deconvolution techniques outlined in the previous two sections have been applied to some of the LRSN recordings of the Kern County, California earthquake of 1962 September 16. This event was used by Helmberger & Wiggins 1971 and Wiggins & Helmberger 1973. The suite of seismograms is shown in Fig. 8.

This example has an interesting source wavelet which invalidates the assumption of low quefrency sources. It is apparent on most records, particularly those beyond 29° epicentral distance, that the source has a precursor feature about 5 s before the main event. If the low quefrency assumption were applied to this source, then the precursor would be taken as part of the impulse response, and not part of the source. It is also apparent by examining HTMN and HNME that some relative phase shifting takes place between these two parts of the source, or that the Earth's impulse response contains a phase shift.

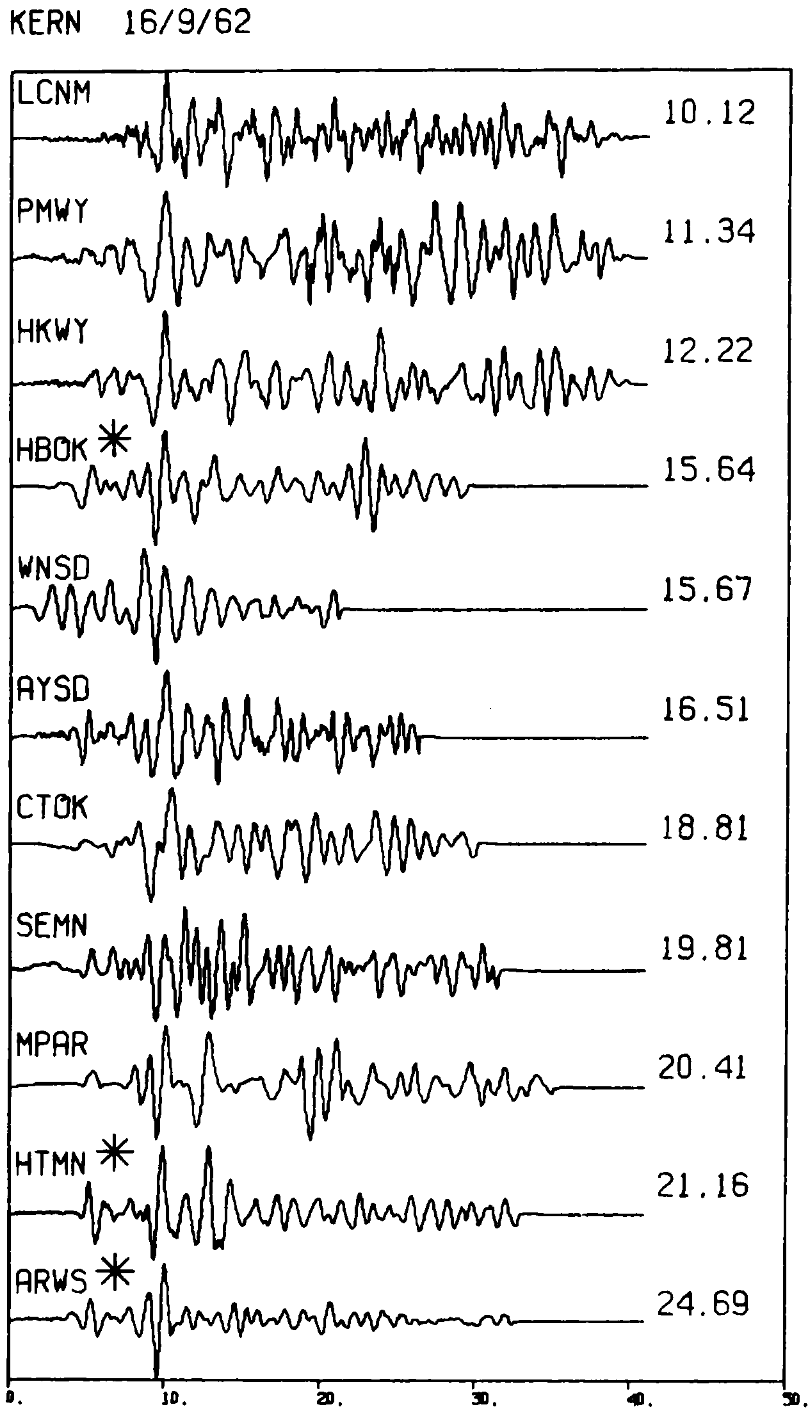

Figure 8a.

Figure 8. The seismogram suite for the deconvolution example. This suite of seismograms is a set of LRSN recordings of the Kern County event of 1962 September 16. The numbers to the right refer to the epicentral distances of the stations, and the asterisks indicate the seismograms that where used in the source estimate.

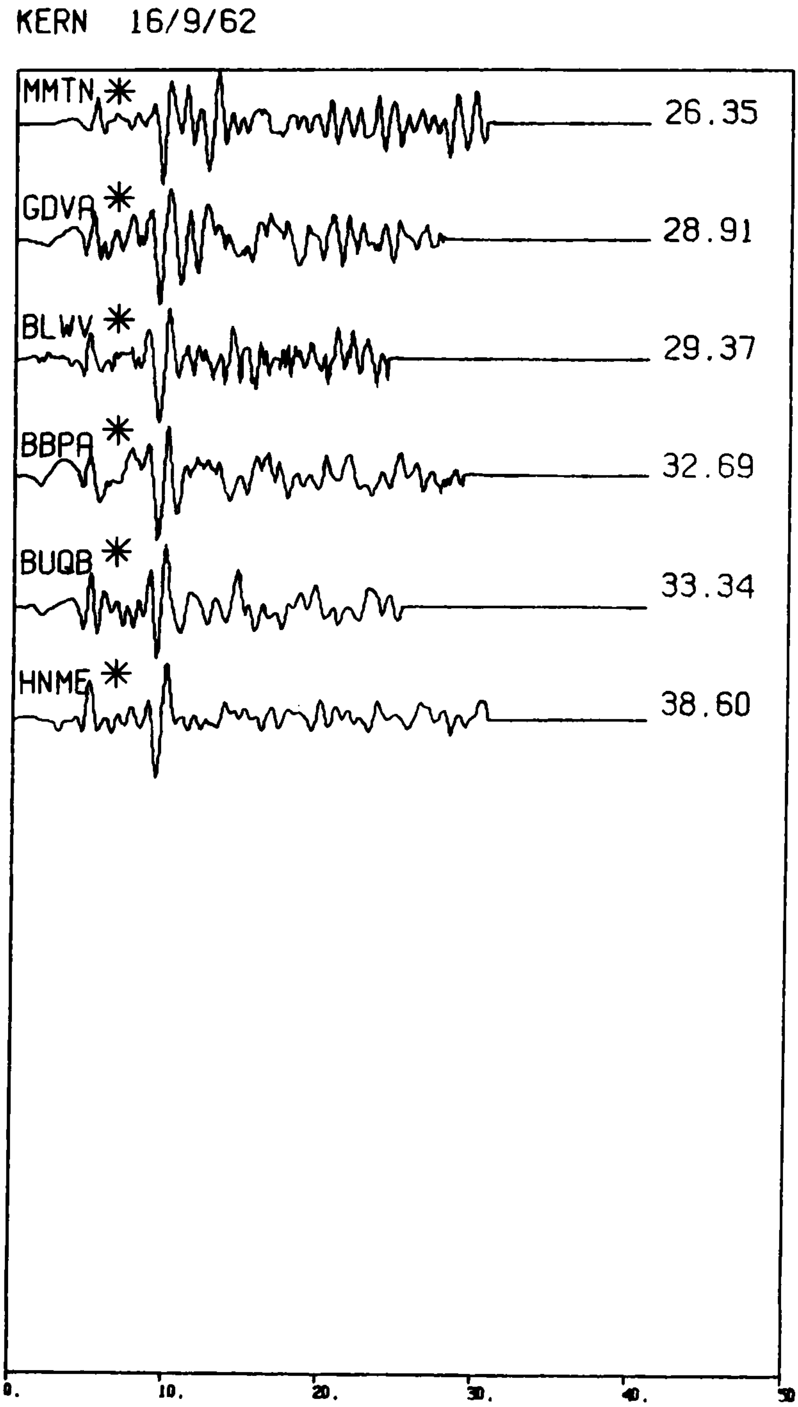
KERN 16/9/62
MMTN *
26.35
GDVA *
28.91
BLWV *
29.37
BBPA *
32.69
BUQB *
33.34
HNME *
38.60
0.
10.
20.
30.
40.
50.

Figure 8b.

A source was estimated for this event by averaging the log amplitude and phase spectra of a suite of nine stations which showed the clearest *P* phases. The seismograms marked with asterisks in Fig. 8 are the ones which were used in the source estimation. The source wavelet was windowed by a 20-s window.

Several of the more interesting deconvolutions of the Kern County suite are presented in the figures following Fig. 8. Each of the seismic traces was delayed so that the first obvious arrival occurs at the 10 s mark on the arbitrary time scale given with the plots. The impulse that corresponds to this arrival should appear at the 0 s mark, but to make a clearer presentation, the impulse response was delayed by 10 s.

The enhancement achieved by deconvolution was judged to be good for this event. The results are consistent with the work of Helmberger & Wiggins (1971) and Wiggins & Helmberger (1973). For every recording, the initial *P* wave onset is easily discernible, and for the stations HTMN and MMTN (Figs 12 and 13), the second arrival is very well spiked.

Two problems with this deconvolution method become apparent in these examples. The first occurs with the waterlevel parameter near unity, i.e. the deconvolution is approaching a crosscorrelation. For this case, both parts of the source tend to be 'spiked' because the autocorrelation of the source itself contains three spikes. Thus, it is important to perform the deconvolution for a number of waterlevels in order that this problem can be detected.

The second problem is over-resolution by the impulse spectrum extension scheme. In particular, the deconvolution of WNSD may suffer from this problem. It is not clear what is causing the resonance in this recording. The deconvolution indicates the arrival of two interfering phases, but we believe this result should be treated with suspicion.

5 Conclusions

In this paper we have presented a method for the deconvolution of teleseismic recordings. The first step in the deconvolution is source estimation. Since earthquakes tend to produce complex source time functions, the ability to separate the source and transmission path effects on the basis of a single seismogram is limited. For example, if the earthquake has two or more successive motions it is difficult, if not impossible, to decide whether the observed effect on a particular recording is due to the source or the transmission path. Source estimation by homomorphic transforms of single seismograms (Ulrych 1971) was found to be an unsatisfactory source estimator for earthquakes because of the invalidity of the low quefrency assumption and the phase instabilities of the transform itself. The obvious remedy for the source estimation problem is to use the redundant source information available from a suite of recordings of the same event, whose distance, azimuthal, and time ranges have been restricted sufficiently to allow the assumption of source uniformity.

The specific method of source estimation that is suggested is to separately average the log amplitude and phase spectra of the recordings of the suite together. This method is equivalent to averaging the cepstra of the seismograms together. However, it is not necessary to unwrap the phase curve with the result that some of the inherent phase instabilities of the homomorphic transform are avoided. The low quefrency assumption is replaced by the assumption that averaging sufficiently reduces the cepstral contributions of the impulse response. This method has the advantage over simple time domain averaging that it uses the redundant information of secondary arrivals on the recordings. The necessary condition for this method to resolve the source time function is that the various arrivals on the recordings have different phase velocities.

The actual deconvolution of the seismograms is accomplished by spectral division. Two modifications to this method were made. The first is the introduction of a waterlevel parameter which constrains the minimum allowable source spectral amplitude level (Helmberger

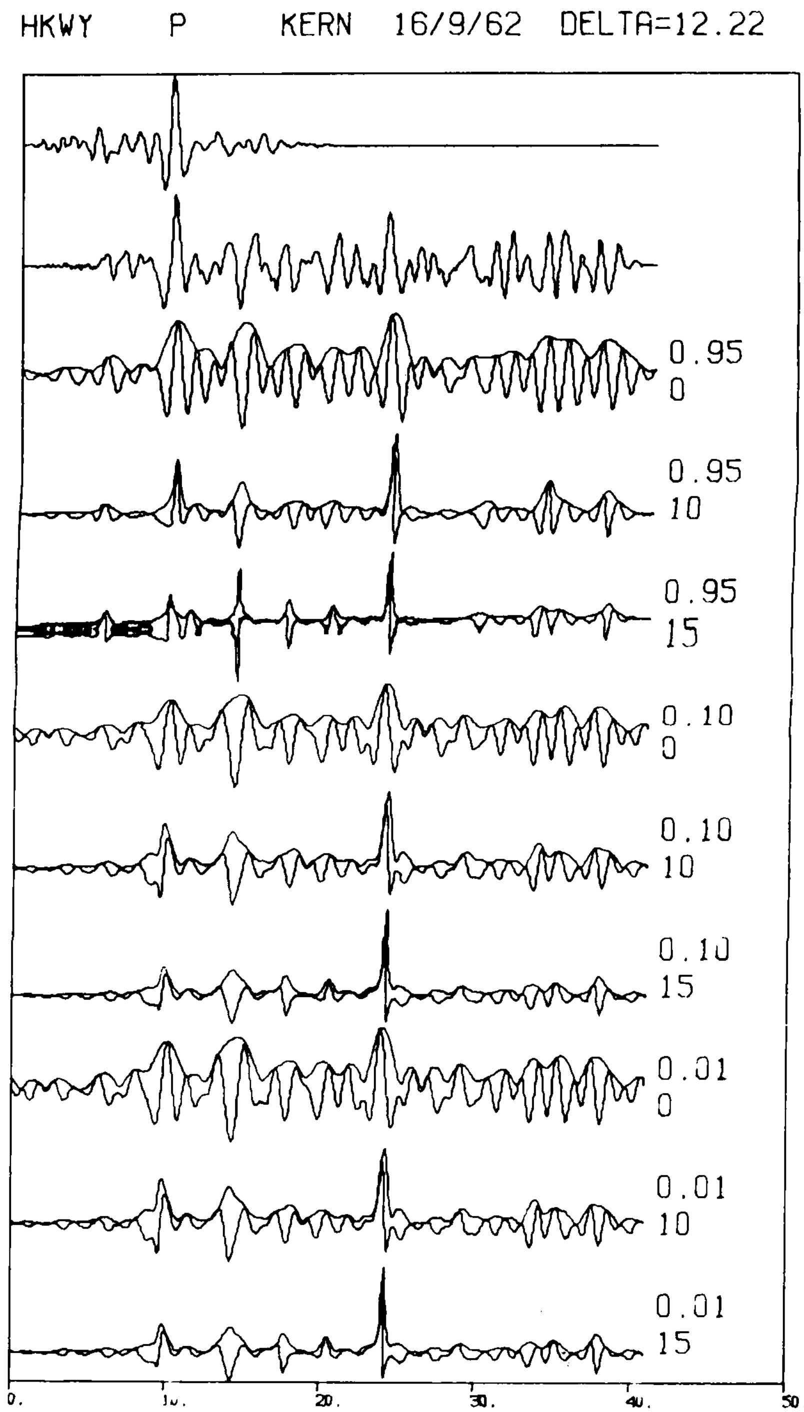

Figure 9. The deconvolution of HKWY. The format of Fig. 9–14 is the same. The station name, the primary phase on the record, the name of the event, and the source–receiver separation in degrees are given in the title. The first trace is the source estimate for the suite, the second trace is the particular seismogram, and the remaining traces are the various deconvolutions of the seismogram. The decimal number to the right of each deconvolution is the value of the waterlevel parameter that was used, and the integer is the autoregressive order that was used in the spectral extension (zero means no extension).

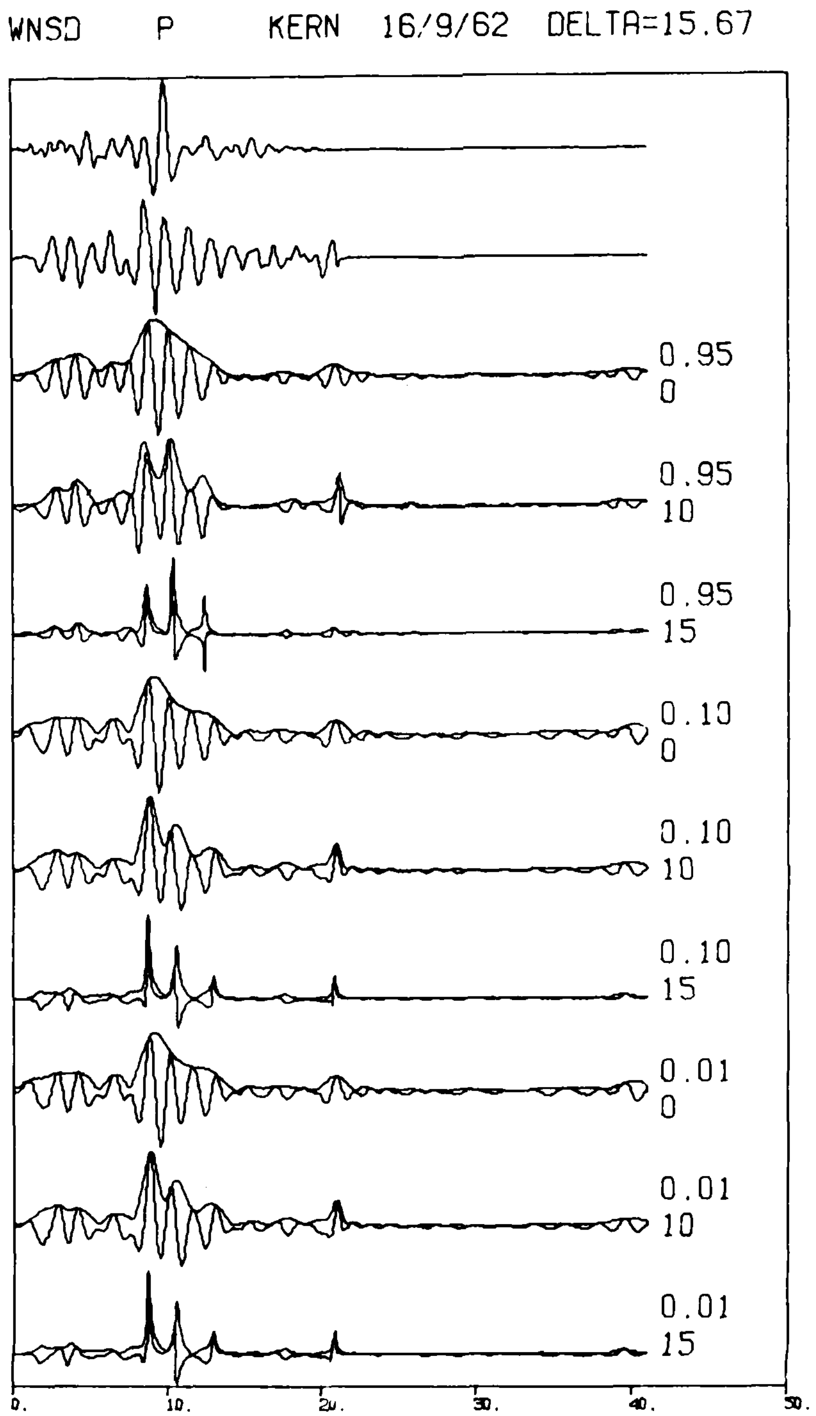

Figure 10. The deconvolution of WNSD. See Fig. 9 for a description of the plot format. This is a difficult seismogram to deconvolve because of the strong resonance effects. The three main spikes in the deconvolution at high autoregressive order appears to be a case of over-resolution by the spectral extension scheme.

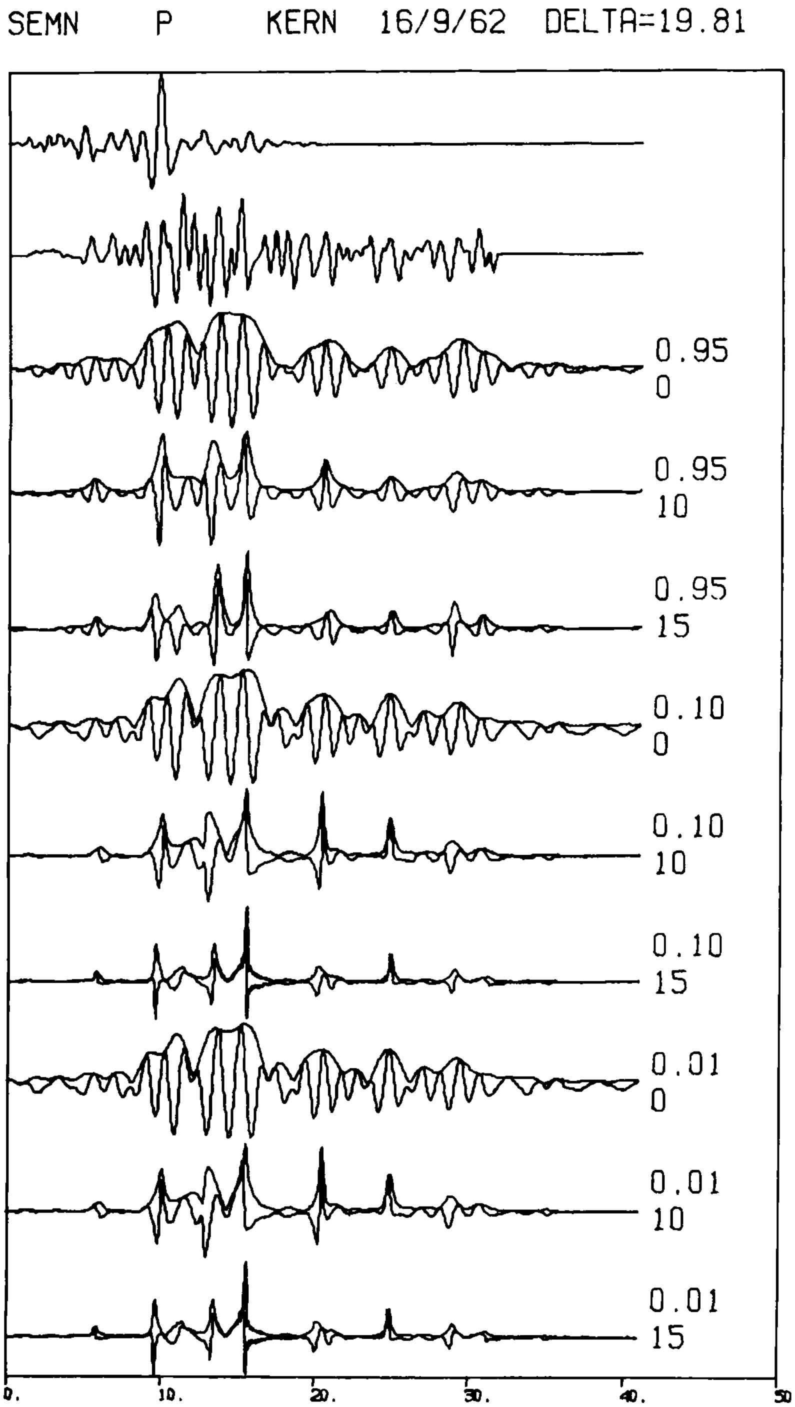

Figure 11. The deconvolution of SEMN, see Fig. 9. The several arrivals that appear in the 10–20 s time window of the seismogram are well separated in the deconvolution with spectral extension.

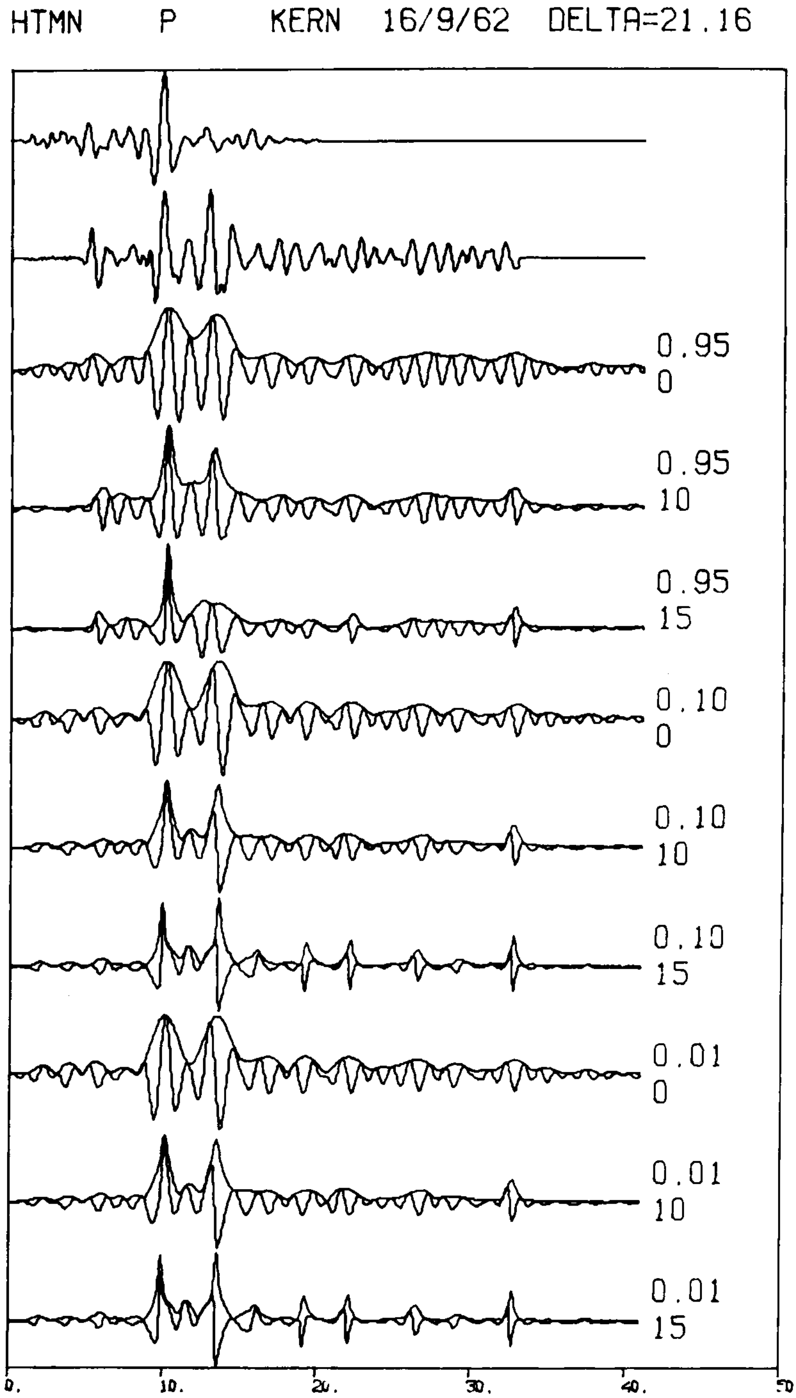

Figure 12. The deconvolution of HTMN, see Fig. 9. The two main arrivals are well resolved in the deconvolutions. Notice the change in phase between these arrivals.

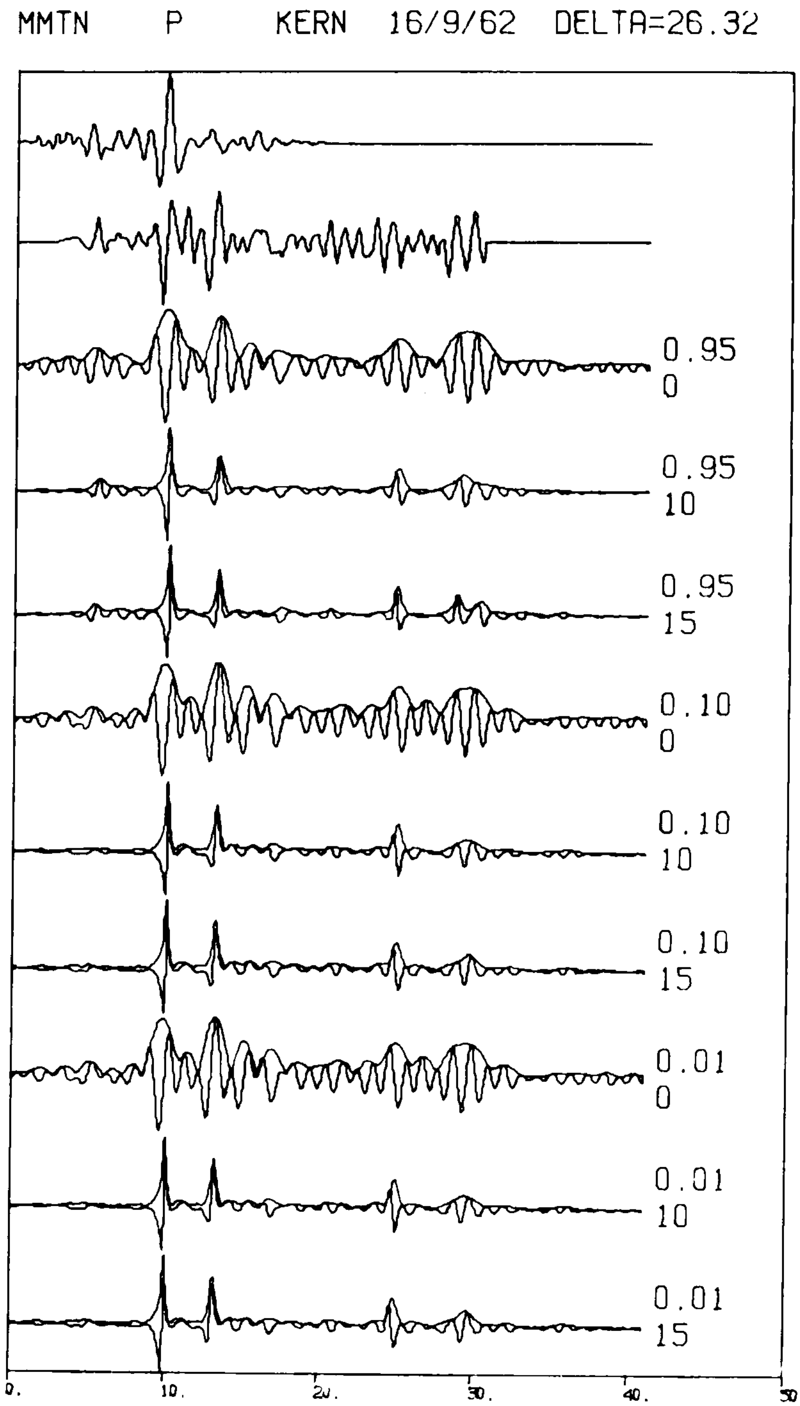

Figure 13. The deconvolution of MMTN, see Fig. 9. The two main arrivals are again well separated in the final deconvolution.

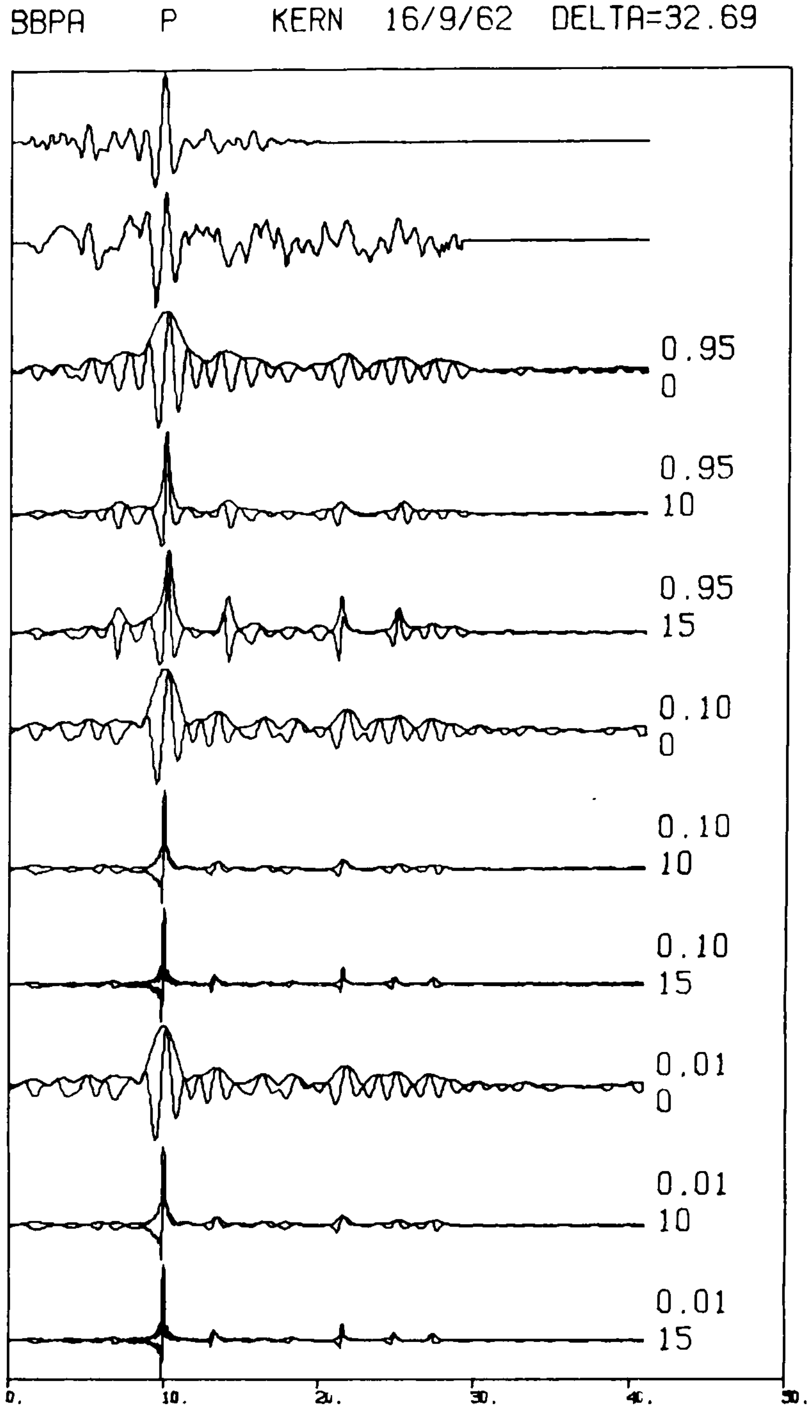

Figure 14. The deconvolution of BBPA, see Fig. 9. The single main arrival is well resolved despite the high noise level of the seismogram. In the third deconvolution which is actually a crosscorrelation of the source and the seismogram, the peaks on either side of the main peak do not correspond to arrivals. For this case both parts of the source are spiked because the autocorrelation of the source itself has three spikes. The relationship between the envelope and the deconvolution indicates that the average source differs somewhat from the phase of the arrival at 32.7°.

& Wiggins 1971). The waterlevel parameter is defined as a fraction of the maximum source spectral amplitude, and is useful for limiting the gain of the deconvolution in regions where the seismograms contain little or no information. The parameter also trades off arrival time resolution with arrival amplitude resolution.

The second modification is the extension of transmission path impulse response beyond its optimal spectral passband to increase the time domain resolution of the deconvolution method. This is accomplished by predicting the unknown spectral regions with a prediction operator determined by Burg's maximum entropy algorithm from the estimated impulse response spectrum in the optimal passband. The length of the prediction operator and the waterlevel mentioned above are two parameters which have to be chosen for this method. Since in practice neither can be chosen optimally, we suggest that the deconvolution be performed for a range of these parameter values. The stability of the deconvolution can be checked by comparing the results for the various parameter values.

The appearance of a deconvolved seismogram can be substantially improved by superimposing the envelope of deconvolution on it (Helmberger & Wiggins 1971). This envelope is independent of the arrival phase shifts and consequently the relationship between the envelope and the deconvolution indicates the presence of any arrivals phase shifted relative to the estimated source.

Acknowledgment

This research was supported by the National Research Council of Canada under operating grant A8854.

References

Akaike, H., 1969. Fitting autoregressive models for prediction, *Ann. Inst. Stat. Math.*, **21**, 261–265.

Bogert, B. P., Healy, M. J. R. & Tukey, J. W., 1963. The quefrency analysis of time series for echoes: cepstrum, pseudo-autocovariance, cross-cepstrum, and saphe cracking, *Proceedings of the symposium on time series analysis*, pp. 206–243, ed. M. Rosenblatt, Wiley and Sons, Inc., New York.

Bracewell, R., 1965. *The Fourier transform and its applications*, McGraw-Hill, New York.

Burg, J. P., 1967. Maximum entropy spectral analysis, *37th Annual International SEG Meeting*, Oklahoma City, 1967 October 31.

Choy, G. L. & Richards, P. G., 1975. Pulse distortion and Hilbert transformation in multiply reflected and refracted body waves, *Bull. seism. Soc. Am.*, **65**, 55–70.

Claerbout, J. F., 1976. *Fundamentals of geophysical data processing*, p. 274, McGraw-Hill, New York.

Clayton, R. W., McClary, B. & Wiggins, R. A., 1976. Comments on the paper 'Phase-distortion and Hilbert transformation in multiply reflected and refracted body waves', by G. L. Choy & P. G. Richards, *Bull. seism. Soc. Am.*, **66**, 325–327.

Dey Sarkar, S. K., 1974. Upper mantle *P*-wave velocity distributions beneath Western Canada, *PhD thesis*, University of Toronto.

Dey Sarkar, S. K. & Wiggins, R. A., 1976. Source deconvolution of teleseismic *P*-wave arrivals between 14° and 40°, *J. geophys. Res.*, in press.

Farnbach, J. S., 1975. The complex envelope in seismic signal analysis, *Bull. seism. Soc. Am.*, **65**, 951–962.

Fryer, G. J., Odegard, M. E. & Sutton, G. H., 1975. Deconvolution and spectral estimation using final prediction error, *Geophysics*, **40**, 411–425.

Helmberger, D. & Wiggins, R. A., 1971. Upper mantle structure of the Midwestern United States, *J. geophys. Res.*, **76**, 3229–3245.

Hill, D. P., 1974. Phase shift and pulse distortion in body waves due to internal caustics, *Bull. seism. Soc. Am.*, **64**, 1733–1742.

Oppenheim, A. V., 1967. Generalized superposition, *IEEE, Information and Control*, **11**, 528–536.

Robinson, E. A., 1967. *Statistical communication and detection*, Hafner Publishing Co., New York.

Savage, J. C., 1966. Radiation from a realistic model of faulting, *Bull. seism. Soc. Am.*, **56**, 577–592.

Schafer, R. W., 1969. Echo removal by discrete generalized linear filtering, Research lab. of Electronics, *MIT, Tech. Report*, 466.

Smylie, D. E., Clarke, G. K. C. & Ulrych, T. J., 1973. Analysis of irregularities in the Earth's rotation, *Methods in computational physics*, vol. 13, pp. 391–430, Academic Press, New York.

Stoffa, P. L., Buhl, P. & Bryan, G. M., 1974. The application of homomorphic deconvolution to shallow water marine seismology Part I: Models, *Geophysics*, **39**, 401–416.

Ulrych, T J., 1971. Application of homomorphic deconvolution to seismology, *Geophysics*, **36**, 650–660.

Ulrych, T. J., Smylie, D. E., Jensen, O. G. & Clarke, G. K. C., 1973. Predictive filtering and smoothing of short records by using maximum entropy, *J. geophys. Res.*, **78**, 4959–4964.

Ulrych, T. J. & Bishop, T. N., 1975. Maximum entropy spectral analysis and autoregressive decomposition, *Rev. Geophys.*, **13**, 183–200.

Wiggins, R. A. & Helmberger, D., 1973. Upper mantle structure of the Western United States, *J. geophys. Res.*, **78**, 1870–1880.