The Annals of Statistics
1986, Vol. 14, No. 2, 559–578

STATISTICAL ESTIMATION OF THE PARAMETERS OF A MOVING SOURCE FROM ARRAY DATA[1]

BY SHEAN-TSONG CHIU

Rice University

This paper is concerned with the problem of estimating the variable time delays of a signal arriving at an array of sensors. A procedure to estimate the parameters of a linear time delay model is proposed. The procedure compares the Fourier transforms at different frequencies (thereby taking the Doppler effect into consideration). Under regularity conditions, the estimate obtained is shown to be consistent and asymptotically normal. Simulations were carried out and the results were found to agree well with the theoretical results. The procedure was applied to the records of the Imperial Valley earthquake of October 15, 1979, as recorded by the El Centro differential array.

1. Introduction. The situation we are interested in is that of a signal emitted by a moving source, such as the leading point of an earthquake rupture, as received by an array of sensors. The particular data we are working with come from sensors located near the source of a major earthquake with a lengthy source rupture. Ground motion obtained by such sensors is referred to as strong motion records.

Strong-motion seismology is a relatively new discipline. It is however an active one because of its importance in understanding the source properties, the wave properties in near field, and the effect of strong-motion on engineering structures [see, for example, Bolt (1981)].

Most earthquakes appear to be caused by faulting. The associated theory, the elastic-rebound theory, of earthquakes was first outlined by Reid (1910) in his study of the great San Francisco earthquake of 1906. According to the elastic-rebound theory [see Boore (1977)], rocks are elastic, and mechanical energy can be stored in them as in a compressed spring. When the two blocks forming the opposite sides of the fault move by a small amount, the motion elastically strains the rocks near the fault. When the stress becomes larger than the frictional strength of the fault, the frictional bond fails at its weakest point. (The point of initial rupture is called the hypocenter or focus.) From the hypocenter, the rupture rapidly propagates along the surface of the fault, causing the rocks on opposite sides of the fault to slip past each other. A portion of the elastic strain the rocks had stored before the rupture is suddenly released. The rocks along the

Received December 1984; revised July 1985.

[1]This research was support in part by National Science Foundation Grants MCS 80-02698 and CEE 79-01642

AMS 1980 *subject classifications.* Primary 62M99; secondary 60G35.

Key words and phrases. Time delay, Fourier transform, periodogram, spectrum, coherence, Doppler effect.

fault rebound to an equilibrium position in a matter of seconds. The elastic energy stored in the rocks is released as heat (generated by friction) and as seismic waves. The seismic waves are radiated by a moving source, so the waves may be expected to show Doppler-like effects.

The faults of natural earthquakes appear to rupture in quite complex ways, and a reasonable representation might be an erratic motion superimposed on a generally smooth slip [Aki (1967)]. The ground motion produced by this kind of rupture will tend to look like a stochastic process.

Available information for studying source rupture processes including records from seismometers, located either close to or far away from the source. In this paper, we study the records from an array of sensors which are close to an earthquake source. In general, the model of interest can be described as

$$X_i(t) = S_0(h(t, \mathbf{r}_i)) + \varepsilon(t, \mathbf{r}_i). \tag{1.1}$$

Here $X_i(t)$ is the observation recorded by the sensor located at $\mathbf{r}_i$, $S_0(t)$ is the signal emitted by the source, and $h(t, \mathbf{r}_i)$ is the time the signal which arrives at time t at location $\mathbf{r}_i$ was emitted by the source. Thus, each sensor receives the common signal, but with different time shift, together with the noise, $\varepsilon(t, \mathbf{r}_i)$, on that station.

In the following, we consider the case of two stations. We have the model

$$\begin{aligned} X_1(t) &= S(t) + \varepsilon_1(t), \\ X_2(t) &= S(h(t)) + \varepsilon_2(t). \end{aligned} \tag{1.2}$$

2. Estimation of the time delay for a fixed source. When the source is fixed, the time delay between the stations is constant and the model can be simplified to

$$\begin{aligned} X_1(t) &= S(t) + \varepsilon_1(t), \\ X_2(t) &= S(t + \tau_0) + \varepsilon_2(t). \end{aligned} \tag{2.1}$$

In this case the Fourier transforms of the signals have the relationship,

$$\mathbf{d}_s^T(\tau_0, \lambda) \approx \mathbf{d}_s^T(\lambda)\exp(i\tau_0\lambda), \tag{2.2}$$

where

$$\mathbf{d}_s^T(\lambda) = \sum_{t=0}^{T-1} S(t)\exp(-i\lambda t) \tag{2.3}$$

is the Fourier transform of the signal $S(t)$ at frequency λ, and

$$\mathbf{d}_s^T(\tau, \lambda) = \sum_{t=0}^{T-1} S(t + r)\exp(-i\lambda t) \tag{2.4}$$

is the Fourier transform of the delayed signal, $S(t + r)$. With similar definitions for $\mathbf{d}_{\varepsilon_1}^T(\lambda)$ and $\mathbf{d}_{\varepsilon_2}^T(\lambda)$, the Fourier transforms of the observed series have the approximate relation

$$\mathbf{d}_{X_2}^T(\lambda) \approx \mathbf{d}_{X_1}^T(\lambda)\exp(i\lambda\tau_0) + \mathbf{d}_{\varepsilon_2}^T(\lambda) - \mathbf{d}_{\varepsilon_1}^T(\lambda)\exp(i\lambda\tau_0). \tag{2.5}$$

This suggests estimating τ_0 by that $\hat{\tau}$ which maximizes the criterion

$$Q_T(\tau) = \sum_{|s|<T/2} \psi(\omega_s)\mathbf{d}_{X_2}^T(\omega_s)\mathbf{d}_{X_1}^T(-\omega_s)\exp(-i\tau\omega_s). \tag{2.6}$$

Here $\omega_s = 2\pi s/T$, $s = 0, \pm 1, \ldots,$ are the Fourier frequencies. Since the spectra of the signal and of the noise are usually not constant, the function ψ was introduced to put different weights on different frequencies. Hannan (1975) proved that, under regularity conditions, such an estimate is strongly consistent and asymptotically normal. Thomson (1982) extended the results to vector observations. The case of dispersive waves (different frequency components traveling at different speeds) was studied in Hannan (1975). We remark that the problem of estimating the constant time delay has also been considered extensively in the study of passive sonar signal processing [see Carter (1981) and references therein].

3. Estimation of the linear time delay parameters. When the source is moving, the time delays between sensors are time dependent. Due to the complexity, little research concerning a moving source has been done. The problem was considered in Knapp and Carter (1977) and in Schultheiss and Weinstein (1979).

It was shown in Chiu (1984) that, for some pertinent cases, the time delay can be well approximated by a linear function. In these cases we have the approximate model

$$\begin{aligned} X_1(t) &= S(t) + \varepsilon_1(t), \\ X_2(t) &= S(\alpha_0 + \beta_0 t) + \varepsilon_2(t), \end{aligned} \tag{3.1}$$

and the problem of estimating the variable time delay is simplified to the problem of estimating the parameters α_0 and β_0. In practice, β_0 is quite close to 1 and it seems reasonable to write $\beta_0^{-1} = 1 + c_0/T$. Then we have

$$\mathbf{d}_s^T((\alpha_0, \beta_0), \lambda) \approx \mathbf{d}_s^T(\lambda/\beta_0)\exp(i\alpha_0\lambda), \tag{3.2}$$

where $\mathbf{d}_s^T((\alpha, \beta), \lambda)$ is the Fourier transform of the shifted signal, $S(\alpha + \beta t)$. Therefore, the Fourier transform of the second series can be written as

$$\begin{aligned} \mathbf{d}_{X_2}^T(\lambda) &= \mathbf{d}_s^T((\alpha_0, \beta_0), \lambda) + \mathbf{d}_{\varepsilon_2}^T(\lambda) \\ &\approx \mathbf{d}_{X_1}^T(\lambda/\beta_0)\exp(i\lambda\alpha_0) - \mathbf{d}_{\varepsilon_1}^T(\lambda/\beta_0)\exp(i\lambda\alpha_0) + \mathbf{d}_{\varepsilon_2}^T(\lambda). \end{aligned} \tag{3.3}$$

This suggests consideration of the estimate $(\hat{\alpha}, \hat{\beta})$ which maximizes the criterion

$$Q_T(\alpha, \beta) = \sum_{|s|<T/2} \psi(\omega_s)\mathbf{d}_{X_2}^T(\omega_s)\mathbf{d}_{X_1}^T(-\omega_s/\beta)\exp(-i\lambda\alpha). \tag{3.4}$$

In the following, we only discuss the case of stochastic signals. Similar results for a deterministic signal can be found in Chiu (1984). In the next section we will state the assumptions under which the estimate is strongly consistent and asymptotically normal.

4. Assumptions and results. We first describe the model of interest in Assumption 1.

ASSUMPTION 1. Let $S(t)$, $-\infty < t < \infty$, be a real valued stationary process. Let the observations $X_1(t)$, $X_2(t)$ have the structure

$$(4.1)\qquad \begin{aligned} X_1(t) &= S(t) + \varepsilon_1(t), \\ X_2(t) &= S(\alpha_0 + \beta_0 t) + \varepsilon_2(t). \end{aligned}$$

Further suppose that $t = 0, \dots, T-1$ and $\beta_0^{-1} = 1 + c_0/T$. Also assume that α_0 and c_0 are contained in the interiors of compact sets A and C, respectively, in $\mathbf{R}$.

The requirement that (α_0, c_0) is contained in the interior of a compact set seems not to be a strict one, for in usual physical situations it is clearly possible to limit the extent of the delay. We next assume the series, $\varepsilon_1(t)$, $\varepsilon_2(t)$, and $S(t)$ to be statistically independent Gaussian processes and to satisfy mixing conditions.

ASSUMPTION 2. $S(t)$, $\varepsilon_1(t)$, and $\varepsilon_2(t)$ are stationary independent Gaussian processes with mean zero. Also suppose $\Sigma(1 + |u|)|c_i(u)| < \infty$, where $c_i(u)$, $i = 1, 2$, are the autocovariance functions of $\varepsilon_i(t)$, $i = 1, 2$.

Under this assumption, the noises $\varepsilon_i(t)$, $i = 1, 2$, have power spectra $f_{\varepsilon_i}(\lambda)$, $i = 1, 2$, respectively. In practice, the sampling interval usually has been chosen so that the signal has very little spectral mass beyond π. Therefore, it seems reasonable to assume that the signal has no spectral mass above frequency π.

ASSUMPTION 3. Let $c(u) = E[S(t)S(t+u)]$, $-\infty < t, u < \infty$, the autocovariance function of $S(t)$, satisfy

$$(4.2)\qquad \sum_{v=-\infty}^{\infty} (1 + |v|) \sup_{v \le u \le v+1} |c(u)| < \infty,$$

and $f_s(\lambda) = 0$ for $|\lambda| > \pi$, where $f_s(\lambda)$, $-\infty < \lambda < \infty$, is the power spectrum of $S(t)$.

Because the signal in practice often has significant magnitude only in some frequency intervals, we would like the weighting function to satisfy the following conditions.

ASSUMPTION 4. $\psi(\lambda) = I_\Omega(\lambda)\phi(\lambda)$, where $\phi(\lambda)$ is a nonnegative continuous function of λ with period 2π, $I_\Omega(\lambda)$ is here the indicator function of a set $\Omega \subset (-\pi, \pi)$, Ω is a finite collection of intervals, and $\psi(\lambda)$ is symmetric about 0. Further suppose $\int \psi(\lambda) f_s(\lambda)\, d\lambda > 0$.

The condition, $\int \psi(\lambda) f_s(\lambda)\, d\lambda > 0$, requires the intervals, Ω, to contain some frequency intervals in which the signal has spectral mass (otherwise, we will not be able to estimate the parameters).

Under Assumptions 1–4, the estimate obtained by maximizing the criterion function $Q_T(\theta)$ of (3.4) is strongly consistent and asymptotically normal. We state the theorems here and postpone the proofs to the last section.

THEOREM 1. *Suppose Assumptions* 1–4 *hold. Let* $\hat{\theta}_T = (\hat{\alpha}_T, \hat{c}_T)$ *be a sequence of estimates which maximizes* $Q_T(\theta)$ *of* (3.4). *Then* $\hat{\theta}_T$ *converges to* $\theta_0 = (\alpha_0, c_0)$ *almost surely as* $T \to \infty$. *Here* $\beta^{-1} = 1 + c/T, \theta = (\alpha, c)$.

THEOREM 2. *Under the conditions of Theorem* 1, *let* $\hat{\theta}_T = (\hat{\alpha}_T, \hat{c}_T)$ *be a sequence of estimates which maximizes* $Q_T(\theta)$. *Then* $T^{1/2}(\hat{\alpha}_T - \alpha_0, \hat{c}_T - c_0)$ *is asymptotically normal with mean zero and covariance matrix*

$$a\begin{pmatrix} 4 & 6 \\ 6 & 12 \end{pmatrix}. \tag{4.3}$$

Here a equals

$$\frac{2\pi\int_{-\pi}^{\pi}\lambda^2\psi^2(\lambda)\{f_s(\lambda)[f_{\varepsilon_1}(\lambda) + f_{\varepsilon_2}(\lambda)] + f_{\varepsilon_1}(\lambda)f_{\varepsilon_2}(\lambda)\}\,d\lambda}{\left(\int_{-\pi}^{\pi}\lambda^2\psi(\lambda)f_s(\lambda)\,d\lambda\right)^2}. \tag{4.4}$$

From the covariance matrix, we saw that the correlation coefficient between $\hat{\alpha}$ and $\hat{c}$ is quite high (0.866). This was confirmed in the simulation study. It should be noted that Figure 2 is the plot of $\hat{\alpha}$ and $\hat{\beta}$. Since $\beta^{-1} = 1 + \hat{c}/T$, as we assigned before, Figure 2 shows a negative correlation between $\hat{\alpha}$ and $\hat{\beta}$.

5. Estimation of the spectra and selection of the weighting function. From Theorem 2 we see that the asymptotic covariance matrix of the estimate depends on the spectra of the signal and the noises. In practice, however, these spectra are unknown and we need to estimate them. After getting the spectrum estimates we can estimate the covariance matrix by substituting the estimated spectra for the true ones. We discuss a method of estimating the spectra in this section.

We note that

$$\begin{aligned} &\mathbf{d}_{X_2}^T(\lambda)\mathbf{d}_{X_1}^T(-\lambda/\beta_0)\exp(-i\alpha_0\lambda) \\ &\quad \approx \mathbf{d}_s^T(\theta_0,\lambda)\mathbf{d}_s^T(\theta_0,-\lambda) + \mathbf{d}_s^T(\theta_0,\lambda)\mathbf{d}_{\varepsilon_1}^T(-\lambda/\beta_0)\exp(-i\alpha_0\lambda) \\ &\qquad + \mathbf{d}_{\varepsilon_2}^T(\lambda)\mathbf{d}_s^T(-\lambda/\beta_0)\exp(-i\alpha_0\lambda) \\ &\qquad + \mathbf{d}_{\varepsilon_2}^T(\lambda)\mathbf{d}_{\varepsilon_1}^T(-\lambda/\beta_0)\exp(-i\alpha_0\lambda). \end{aligned} \tag{5.1}$$

On the right-hand side the expected value of the first term is equal to $2\pi T f_s(\lambda)$ and the expected values of the second, third, and fourth terms are zero. So

$$\mathbf{d}_{X_2}^T(\lambda)\mathbf{d}_{X_1}^T(-\lambda/\beta_0)\exp(-i\alpha_0\lambda) \tag{5.2}$$

has expected value approximately equal to $2\pi T f_s(\lambda)$. These values are asymptotically independent at different Fourier frequencies, $\omega_s = 2\pi s/T$. This suggests estimating $f_s(\lambda_0)$, the spectrum of the signal at frequency λ_0, by averaging the

$\mathbf{d}^T_{X_2}(\omega_s)\mathbf{d}^T_{X_1}(-\omega_s/\beta_0)\exp(-i\alpha_0\omega_s)$ at the Fourier frequencies near λ_0. Though we do not know the true value of θ_0, we expect that replacing θ_0 with $\hat{\theta}$, the estimate of θ_0, will give us a useful estimate of the signal's spectrum. We establish the following result.

THEOREM 3. *Under Assumptions* 1–4, *let* $\theta_T \to \theta_0$ *and* $M_T \to \infty$, $M_T/T \to 0$, *then* $\lim_{T\to\infty} E|\hat{f}_s(\lambda_0) - f_s(\lambda_0)|^2 = 0$, *where*

$$\hat{f}_s(\lambda_0) = \frac{1}{2\pi}\frac{1}{M_T T}\sum_{\omega_s \in I_{M_T}} \mathbf{d}^T_{X_2}(\omega_s)\mathbf{d}^T_{X_1}(-\omega_s/\beta_T)\exp(-i\alpha_T\omega_s) \tag{5.3}$$

and I_{M_T} *is the set containing* M_T *Fourier frequencies which are closest to* λ_0. [*Here* $\beta_T^{-1} = 1 + c_T/T$ *and* $\theta_T = (\alpha_T, c_T)$.]

Having an estimate of the signal spectrum, we may proceed to estimate $f_{\varepsilon_1}(\lambda_0)$, the spectrum of the noise $\varepsilon_1(t)$, by $\max(0, \hat{f}_{X_1}(\lambda_0) - \hat{f}_s(\lambda_0))$, where $\hat{f}_{X_1}(\lambda_0)$ is an estimate of $f_{X_1}(\lambda_0) = f_s(\lambda_0) + f_{\varepsilon_1}(\lambda_0)$, the spectrum of $X_1(t)$. $f_{\varepsilon_2}\lambda_0)$, the spectrum of the noise $\varepsilon_2(t)$, can be estimated in a similar way.

We have applied a weighting function in computing the criterion function $Q_T(\theta)$ of (3.4). We should now discuss a method for selecting a pertinent weighting function. Some specific weighting functions have been suggested. Knapp and Carter (1977) have reviewed some of these weighting functions. The optimal choice of $\psi(\lambda)$ is

$$\psi(\lambda) = \frac{f_s^2(\lambda)}{f_s(\lambda)f_{\varepsilon_1}(\lambda) + f_s(\lambda)f_{\varepsilon_2}(\lambda) + f_{\varepsilon_1}(\lambda)f_{\varepsilon_2}(\lambda)}. \tag{5.4}$$

This was obtained by using a quasi-maximum likelihood procedure [see Hannan and Robinson (1973)]. This weighting function minimizes the value of a of (4.3) in Theorem 2 among all weighting functions. In practice, however, the spectra of the signal and the noises are unknown. This may cause problems when one substitutes the estimated values for the true ones, since when the signal to noise ratio is low the spectrum estimate of the signal tends to be bigger than the true value. This gives too much weight to that frequency. Hannan and Thomson (1981) assumed a finite parameter linear model, and then estimated the spectrum of the linear model.

Since the spectrum of the signal often has significant magnitude only in some frequency intervals, we would want the summation of (3.4) to extend only over these intervals. This suggests an alternative approach, namely, to discard the information at frequencies where the signal to noise ratio is low. So we select $\Omega = \Lambda \cup (-\Lambda)$, where Λ is a finite collection of intervals in $(0, \pi)$, and use a weighting function which can be written as $\psi(\lambda) = \phi(\lambda)I_\Omega(\lambda)$. $\phi(\lambda)$ is an even, positive, and continuous function over $(-\pi, \pi)$, and $I_\Omega(\lambda)$ is the indicator function of Ω.

Intuitively, we would like to select these intervals which have high power from the signal and low power from the noise. Under our model it is equivalent to choosing the intervals which have high coherence. Thus, we can use coherence to select the intervals. Brillinger (1975) discussed estimating coherence from the periodogram and gave the asymptotic distribution of the coherence estimate. We should note that this theory is for stationary series, and the model we discuss does not satisfy this condition. However since β_0 is quite close to 1 we might still get a reasonable estimate of the coherence.

After getting the estimate of the coherence, we can calculate, for example, the 90% quantile of the null distribution of the estimate under the hypothesis that the coherence is zero. We shall choose those intervals inside which the coherence estimate is higher than that quantile. Having chosen the intervals, we select a weighting function, for example, by the quasi-maximum likelihood method. Then we compute the criterion function $Q_T(\theta)$ and estimate $\theta_0 = (\alpha_0, c_0)$ by $\hat{\theta} = (\hat{\alpha}, \hat{c})$ which maximizes $Q_T(\theta)$.

Now, we reestimate the coherence. This time we substitute $\mathbf{d}_{X_1}^T(-\omega_s/\hat{\beta}_1)$ for $\mathbf{d}_{X_1}^T(\omega_s)$ in the estimation procedure, as before $\hat{\beta}_1^{-1} = 1 + \hat{c}_1/T$. This should give us an improved estimate of the coherence. We then use these estimates to choose intervals and the weighting function to get the final estimate of the time delay parameters.

In practice some other considerations may affect the choice of the weighting function. The sensors may receive several signals at the same time, and different signals might have different frequency contents. For this case if the signal we are interested in has low power in some frequency intervals, then we should not choose those intervals even when they have high coherence. We also note that the estimated coherence tends to be bigger than the true value when both spectra of the observation series have small power. Therefore, we should be careful in employing the intervals which have small signal power.

6. Simulation results and application to seismic data. In this section we present simulation results to evaluate the performance of the estimation procedure discussed in the previous sections. The signal used in the simulations is a band-limited stationary process. The frequency content is limited in the interval $(60\pi/512, 120\pi/512)$. The signal is

$$S(t) = \frac{3}{200}\sum_{j=1}^{2000}\left(z_1(j)\cos(\lambda_j) + z_2(j)\sin(\lambda_j)\right), \tag{6.1}$$

where $z_1(j)$, $z_2(j)$ are independent Gaussian random variables with means zero and variances 1, and

$$\lambda_j = \frac{60\pi}{512} + \frac{60j\pi}{512 \cdot 2000}. \tag{6.2}$$

We first generate the signals, $S(t)$ and $S(\alpha_0 + \beta_0 t)$, of length 512 with $\alpha_0 = 0.25$, $\beta_0 = 1.02$. We then form the observed series by adding noise to the signals, that is,

$$\begin{aligned} X_1(t) &= S(t) + \varepsilon_1(t), \\ X_2(t) &= S(0.25 + 1.02t) + \varepsilon_2(t). \end{aligned} \tag{6.3}$$

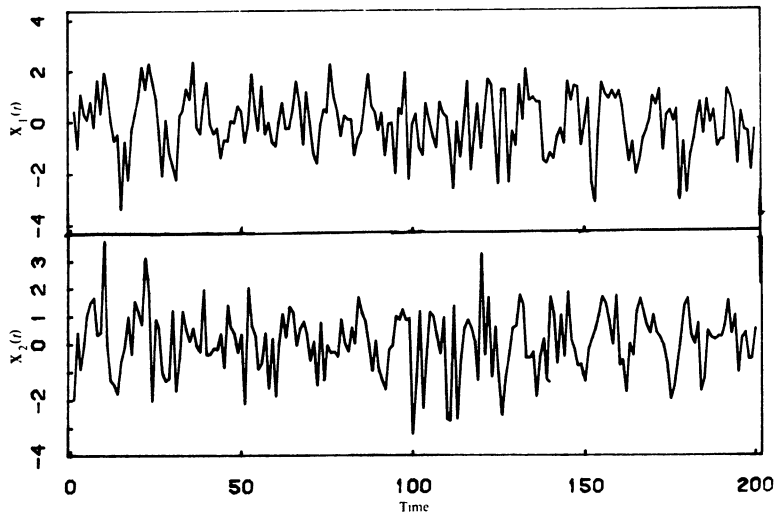

FIG. 1. *Sample series of the simulation.*

The noises are independent Gaussian white series with mean zero and variance 1. Figure 1 is the plot of a pair of sample series. We show only the first 200 points of each series.

The criterion function used in the simulations is

$$Q_T(\theta) = \operatorname{Re} \sum_{s=29}^{59} \mathbf{d}_{X_2}^T(\omega_s)\mathbf{d}_{X_1}^T(-\omega_s/\beta)\exp(-i\alpha\omega_s), \tag{6.4}$$

where $\omega_s = 2\pi s/T$ and $\theta = (\alpha, \beta)$. Figure 2 is the plot of the estimates which maximize $Q_T(\theta)$ above. Figures 3 and 4 are the normal probability plots of the estimates. The result of 75 simulations is summarized in Table 1.

The estimation procedure was applied in analyzing the accelerograms from the El Centro differential array in Imperial Valley. The data are records from the October 15, 1979, Imperial Valley earthquake. This earthquake has been studied extensively. The information about the array can be found in Bycroft (1982).

We analyze the records of stations 1 and 4, which are 420 feet apart. Station 4 is on the north of station 1. We take 512 time points (5.12 seconds), beginning at 23:17:05.15, from each of the north–south components of stations 1 and 4.

In order to examine the frequency content of the signal and choose a proper weighting function, $\psi(\lambda)$, we first estimate the power spectra and the cross-spectrum. We estimate the spectra and the cross-spectrum by smoothing the periodograms at the Fourier frequencies $\omega_s = 2\pi s/T$, that is, we smooth the functions

$$\begin{aligned} I_{11}(\omega_s) &= \mathbf{d}_{X_1}^T(\omega_s)\mathbf{d}_{X_1}^T(-\omega_s)/2\pi T, \\ I_{22}(\omega_s) &= \mathbf{d}_{X_2}^T(\omega_s)\mathbf{d}_{X_2}^T(-\omega_s)/2\pi T, \\ I_{21}(\omega_s) &= \mathbf{d}_{X_2}^T(\omega_s)\mathbf{d}_{X_1}^T(-\omega_s)/2\pi T. \end{aligned} \tag{6.5}$$

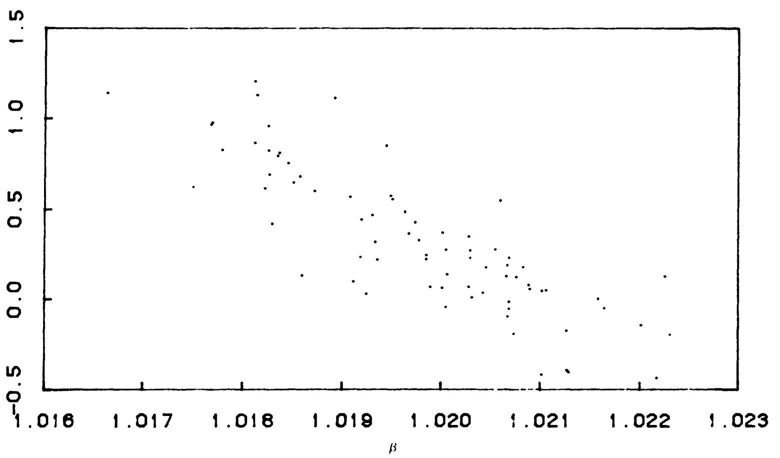

FIG. 2. *Estimates of the 75 simulations.*

Here $\mathbf{d}^T_{X_1}(\lambda), \mathbf{d}^T_{X_2}(\lambda)$ are the Fourier transforms of the data. From the estimated spectra, we see that most of the frequency content lies below 0.1 (10 Hz) for both series. Figure 5 is the plot of the estimated coherence. It suggests that the frequency contents below 0.07 (7 Hz) of these two series are highly correlated.

The criterion function we used in deriving the estimates is the real part of

$$\sum_{s=3}^{31} \mathbf{d}^T_{X_2}(\omega_s)\mathbf{d}^T_{X_1}(-\omega_s/\beta)\exp(-i\omega_s\alpha). \tag{6.6}$$

The maximum of $Q_T(\theta)$ is located at $(-0.10, 1.0163)$. Now we reestimate the

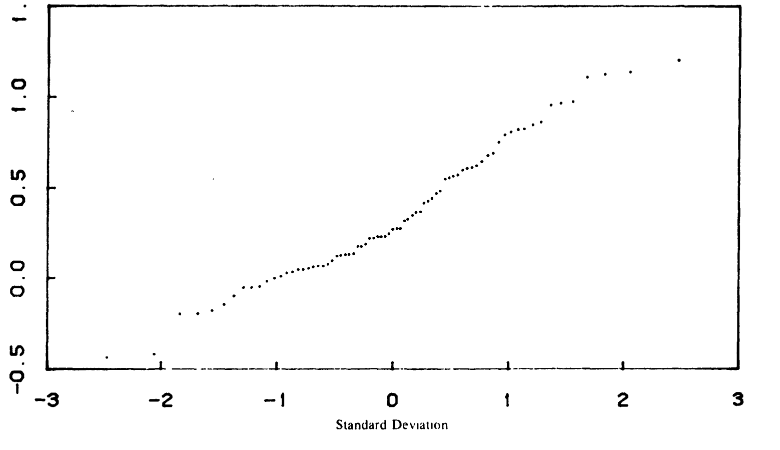

FIG. 3. *Normal probability plot of the 75 estimates of* $\alpha_0 = 0.25$ *from the simulations.*

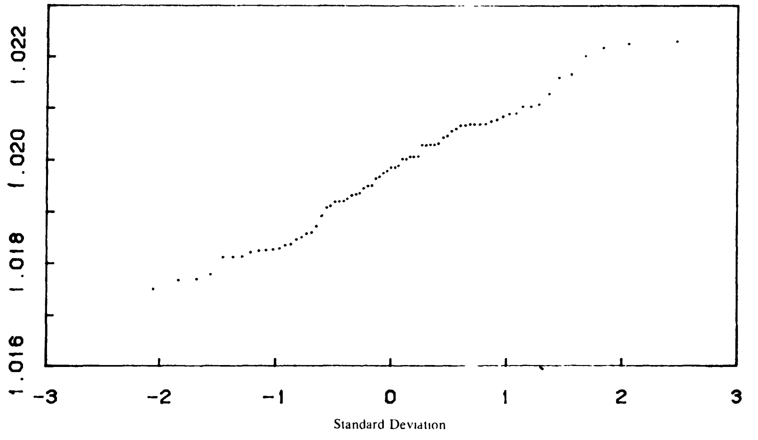

FIG. 4. *Normal probability plot of the* 75 *estimates of* $\beta_0 = 1.02$ *from the simulations.*

TABLE 1
Simulation Results

	Theoretical	Sample
Mean ($\hat{\alpha}$)	0.25	0.349
S.D. ($\hat{\alpha}$)	0.352	0.377
Mean ($\hat{\beta}$)	1.02	1.0197
S.D. ($\hat{\beta}$)	0.00119	0.00123
Corr. ($\hat{\alpha}, \hat{\beta}$)	−0.866	−0.838

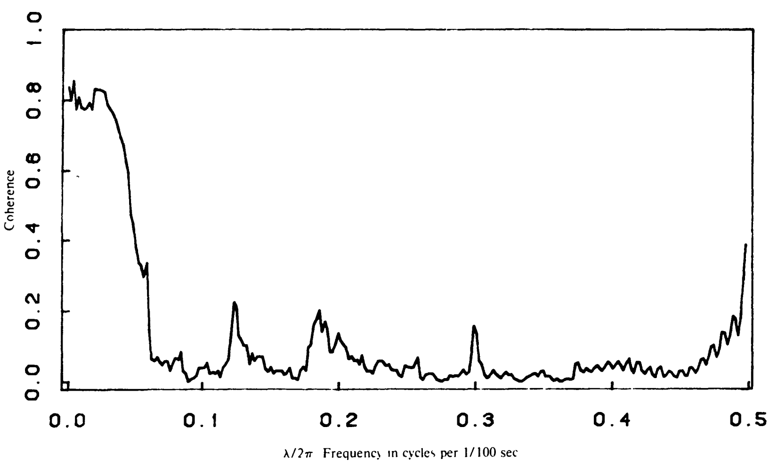

FIG. 5. *Estimated coherence of* $X_1(t)$ *with* $X_2(t)$.

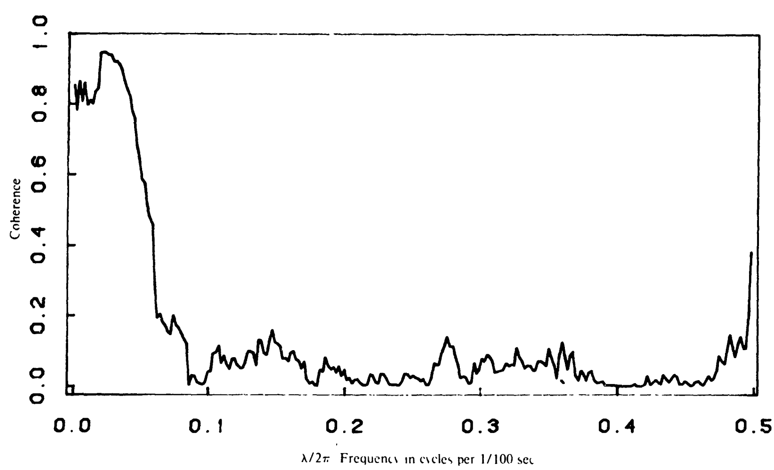

FIG. 6. *Modified coherence of $X_1(t)$ with $X_2(t)$.*

cross-spectrum, this time smoothing the modified periodograms

$$\tilde{I}_{21}(\omega_s) = \mathbf{d}^T_{X_2}(\omega_s)\mathbf{d}^T_{X_1}(-\omega_s/1.0163)/2\pi T. \tag{6.7}$$

As before $\omega_s = 2\pi s/T$ and $T = 512$.

Figure 6 is the plot of the modified coherence. This plot shows correlation only for frequencies below 0.07 (7 Hz). In comparing with Figure 5, we see that the estimated coherence has risen from 0.83 to 0.95 for frequencies near 0.03 (3 Hz). This gives evidence of the existence of the moving source in the data. Because the modified coherence does not suggest that frequency components above 0.07 are correlated, we will not change the weighting function. The final estimates are those above. We then estimate the variances of the estimates by substituting the estimated spectra for the true ones in Theorem 2. The variances of $\hat{\alpha}$ and $\hat{\beta}$ are estimated at 0.0608 and 0.696×10^{-6}, respectively. Thus the estimate of α_0 is -0.10 with a standard deviation 0.25 and the estimate of β_0 is 1.0163 with a standard deviation 0.0008

This set of data has also been analyzed by Spudich and Cranswick (1984). They estimated the variable time delays between the stations. The method they used is to move a time window of fixed length (the length is 127 sample points in their analysis) along the series. For each time window, constant time delay estimate was obtained by finding the location of the maximum of the cross-covariance function of the filtered series. (This method is equivalent to the method we discussed in Section 2.) Then, the time delay estimates were plotted as a function of time. Figure 7, taken from Spudich and Cranswick (1984), shows the results of the north–south components between station 1 and station 3. A variable time delay can be observed from the figure. The vertical unit in the figure is slowness; the slowness is obtained by dividing the time delay (in seconds) by the distance (in km) between the stations. The period of the data we analyzed is from 5.8 seconds to 10.9 seconds, and the major part of the seismic waves

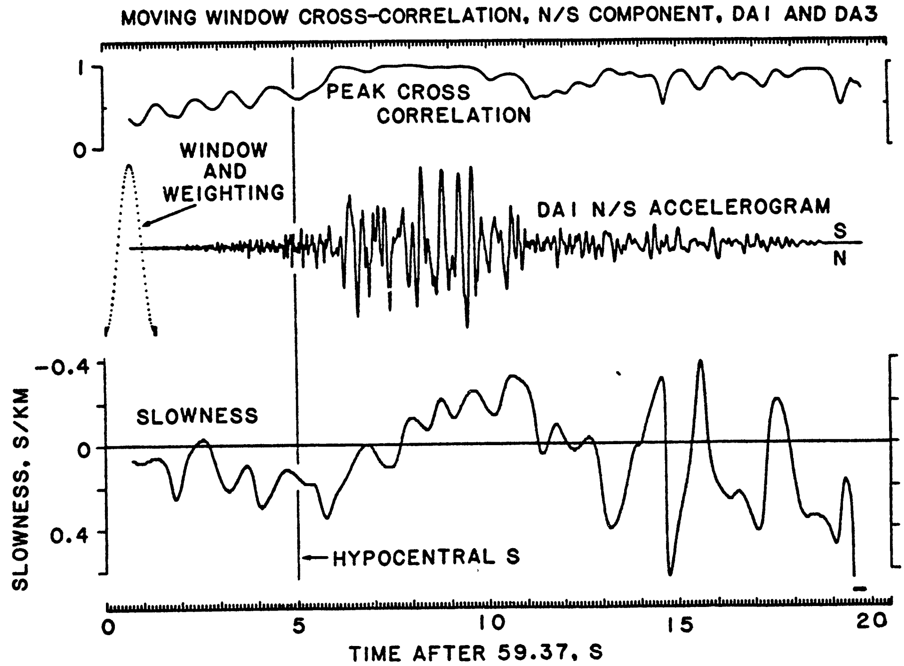

FIG. 7. *Moving-window cross-correlation of the north–south component accelerograms from station 1 and station 3. The vertical line indicates the theoretical arrival time of the S wave from the hypocenter. The upper trace shows the peak value of correlation as a function of time, the middle trace is the north–south accelerogram, and the lowest is the slowness as a function of time through the record. Also indicated are the width and weighting function of the moving window [from Spudich and Cranswick (1984)].*

arrived in this period. In Figure 7, the slowness changed 0.6 (from 0.3 to -0.3) in this period. This result is quite consistent with ours. From our estimate of β, we find the slowness changed 0.65. Since the common time of these records was lost and different time origins were used, we cannot compare the value of $\hat{\alpha}$ with the estimate they obtained at 5.8 seconds.

For the situation of variable time delays, the moving-window has two main disadvantages: (1) The constant time delay estimates are unstable when the window length is small. (2) When the window length is big, the time delay cannot be approximated by a constant time delay model; therefore, it will be difficult to find the statistical properties of the estimates. In this case a linear function, or a piecewise linear function, provides a better approximation to the variable time delay function.

7. Proofs. We proceed to prove the results in this section. The following lemma allows us to replace $\mathbf{d}_{s_t}^T(\omega_s/\beta)\exp(i\alpha\omega_s)$ by $\mathbf{d}_{s_t}^T(\theta, \omega_s)$. This makes it easier to calculate various cumulants.

LEMMA 1. *Let* $S(t)$, $-\infty < t < \infty$, *satisfy Assumption* 3 *and* $S_i(t) = (t/T)^i S(t)$, $i = 0, 1$. *Then, for any bounded function* $g(\lambda)$ *and* $i, j = 0, 1$,

$$E\left\{ \sum_{|s|<T/2} g(\omega_s) \mathbf{d}^T_{s_i}(\theta_1, \omega_s) \mathbf{d}^T_{s_j}(-\omega_s/\beta_2) \exp(-i\alpha_2\omega_s) \right\} = E\left\{ \sum_{|s|<T/2} g(\omega_s) \mathbf{d}^T_{s_i}(\theta_1, \omega_s) \mathbf{d}^T_{s_j}(\theta_2, -\omega_s) \right\} + O(T \log T) \tag{7.1}$$

uniformly on $\Theta \times \Theta$. *Here* $\omega_s = 2\pi s/T$, $\beta_i^{-1} = 1 + c_i/T$, *and* $\theta_i = (\alpha_i, c_i)$.

PROOF OF LEMMA 1. For any $\lambda_1, \lambda_2 \in (-\pi, \pi)$

$$\begin{aligned} E\left\{ \mathbf{d}^T_{s_i}(\theta_1, \lambda_1) \mathbf{d}^T_{s_j}(\lambda_2) \right\} &= \sum_{t_1} \sum_{t_2} (t_1/T)^i (t_2/T)^j c(\alpha_1 + \beta_1 t_1 - t_2) \exp(-i\lambda_1 t_1) \exp(-i\lambda_2 t_2) \\ &= \sum_{t_1} \sum_{t_2} (t_1/T)^i (t_2/T)^j \exp(-i\lambda_1 t_1) \exp(-i\lambda_2 t_2) \\ &\quad \times \int_{-\pi}^{\pi} \exp(i\eta(\alpha_1 + \beta_1 t_1 - t_2)) f_s(\eta)\, d\eta \\ &= \int_{-\pi}^{\pi} H_i^T(\lambda_1 - \eta\beta_1) H_j^T(\lambda_2 + \eta) \exp(i\eta\alpha_1) f_s(\eta)\, d\eta, \end{aligned} \tag{7.2}$$

where $H_i^T(\lambda) = \sum_{t=0}^{T-1} (t/T)^i \exp(-i\lambda t)$, $i = 0, 1, 2$. So

$$\begin{aligned} E&\left\{ \sum_{|s|<T/2} g(\omega_s) \mathbf{d}^T_{s_i}(\theta_1, \omega_s) \mathbf{d}^T_{s_j}(-\omega_s/\beta_2) \exp(-i\alpha_2\omega_s) \right\} \\ &= \sum_{|s|<T/2} g(\omega_s) \int_{-\pi}^{\pi} H_i^T(\omega_s - \eta\beta_1) H_j^T(-\omega_s/\beta_2 + \eta) \exp(i\eta\alpha_1) \\ &\quad \times \exp(-i\alpha_2\omega_s) f_s(\eta)\, d\eta \\ &= \sum_{|s|<T/2} g(\omega_s) \int_{-\pi}^{\pi} H_i^T(\omega_s - \eta\beta_1) H_j^T(-\omega_s + \eta\beta_2) \exp(i\eta\alpha_1) \\ &\quad \times \exp(-i\alpha_2\eta) f_s(\eta)\, d\eta + O\left(\int_{-\pi}^{\pi} \sum_{|s|<T/2} \left| H_i^T(\omega_s - \eta\beta_1) \right| d\eta \right). \end{aligned} \tag{7.3}$$

The last equality holds because $H_i^T(\beta\lambda) = H_i^T(\lambda) + O(1)$ and $\exp(i\alpha\omega) = \exp(i\alpha\eta) + O(|\omega - \eta|)$. Note that the first term of (7.3) is equal to

$$E\left\{ \sum_{|s|<T/2} \psi(\omega_s) \mathbf{d}^T_{s_i}(\theta_1, \omega_s) \mathbf{d}^T_{s_j}(\theta_2, -\omega_s) \right\}, \tag{7.4}$$

and the second term is of order $T \log T$. This completes the proof of the lemma. □

Analogous to the proof of Theorem 4.5.1 of Brillinger (1975), we have the following lemma. The details of the proof can be found in Chiu (1984).

LEMMA 2. *Let $\varepsilon_1(t)$, $\varepsilon_2(t)$ satisfy Assumption 2 and let $\psi(\lambda)$ satisfy Assumption 4. Then*

$$\lim_{T\to\infty} T^2 \sum_{|s|<T/2} \psi(\omega_s)\mathbf{d}^T_{\varepsilon_1}(\omega_s)\mathbf{d}^T_{\varepsilon_2}(\omega_s/\beta)\exp(i\alpha\omega_s) = 0 \quad a.s. \tag{7.5}$$

and uniformly on $(\alpha, c) \in \Theta = C \times A$ (C, A compact sets in $\mathbf{R}$). Here $\beta^{-1} = 1 + c/T$ and $\omega_s = 2\pi s/T$, $s = 0, \pm 1, \dots$.

In proving the strong consistency of the estimate and deriving its asymptotic distribution, we need to calculate the limits of

$$T^{-2} \sum_{|s|<T/2} \psi(\omega_s)\mathbf{d}^T_{s_i}(\theta_1, \omega_s)\mathbf{d}^T_{s_j}(-\omega_s/\beta_2)\exp(-i\alpha_2\omega_s),$$

where $S_i(t) = (t/T)^i S(t)$ for $i = 0, 1, 2$. Lemma 3 gives us the values of these limits.

LEMMA 3. *Let $S(t)$ satisfy Assumption 3 and $\psi(\lambda)$ satisfy Assumption 4. Let $S_i^T(t) = (t/T)^i S(t)$, $i = 0, 1, 2$. Then, uniformly on $\Theta \times \Theta$, and for i, j = 0, 1, 2,*

$$\lim_{T\to\infty} T^{-2} \sum_{|s|<T/2} \psi(\omega_s)\mathbf{d}^T_{s_i}(\theta_1, \omega_s)\mathbf{d}^T_{s_j}(-\omega_s/\beta_2)\exp(-i\alpha_2\omega_s) = Q_{ij}(\theta_1, \theta_2) \tag{7.6}$$

almost surely. Here $\omega_s = 2\pi s/T$, $\beta_i^{-1} = 1 + c/T$, $\theta_i = (\alpha_i, c_i)$, and

$$\begin{aligned} Q_{ij}(\theta_1, \theta_2) = &\int_{-\pi}^{\pi} \psi(\eta)\exp(i\eta(\alpha_1 - \alpha_2))f_s(\eta) \\ &\times \int_0^1 \exp[i\eta(c_1 - c_2)x]x^{i+j}\,dx\,d\eta. \end{aligned} \tag{7.7}$$

PROOF OF LEMMA 3. By using the same arguments as in the lemma of Hannan and Robinson (1973), one can show that

$$\sum_{|s|<T/2} \psi(\omega_s)H_i^T(\omega_s - \beta_1\eta)H_j^T(-\omega_s + \beta_2\eta) \tag{7.8}$$

converges to

$$\psi(\eta)\int_0^1 x^{i+j}\exp(i\eta(c_2 - c_1)x)\,dx. \tag{7.9}$$

Here $H_i^T(\lambda) = \sum_{t=0}^{T-1}(t/T)^i\exp(-i\lambda t)$, $i = 0, 1, 2$, as defined in the proof of Lemma 1. From this and Lemma 1, we have

$$\lim_{T\to\infty} T^{-2}E\left\{\sum_{|s|<T/2} \psi(\omega_s)\mathbf{d}^T_{s_i}(\theta_1, \omega_s)\mathbf{d}^T_{s_j}(\theta_2, -\omega_s)\right\} = Q_{ij}(\theta_1, \theta_2). \tag{7.10}$$

Then, similar to the proof of Lemma 2, one can show that

$$T^{-2} \sum_{|s|<T/2} \psi(\omega_s)\mathbf{d}^T_{s_i}(\theta_1, \omega_s)\mathbf{d}^T_{s_j}(\theta_2, -\omega_s) \tag{7.11}$$

converges, uniformly, to $Q_{ij}^T(\theta_1, \theta_2)$ almost surely. □

Next we prove Theorem 1 concerning the strong consistency of the estimate.

PROOF OF THEOREM 1. Note that $Q_T(\theta)$ can be separated into four terms, namely,

$$
(7.12)\quad
\begin{aligned}
Q_T(\theta) &= T^{-2}\sum_s \psi(\omega_s)\mathbf{d}_s^T(\theta_0,\omega_s)\mathbf{d}_s^T(-\omega_s/\beta)\exp(-i\omega_s\alpha)\\
&\quad + T^{-2}\sum_s \psi(\omega_s)\mathbf{d}_{\varepsilon_2}^T(\omega_s)\mathbf{d}_s^T(-\omega_s/\beta)\exp(-i\omega_s\alpha)\\
&\quad + T^{-2}\sum_s \psi(\omega_s)\mathbf{d}_s^T(\theta_0,\omega_s)\mathbf{d}_{\varepsilon_1}^T(-\omega_s/\beta)\exp(-i\omega_s\alpha)\\
&\quad + T^{-2}\sum_s \psi(\omega_s)\mathbf{d}_{\varepsilon_2}^T(\omega_s)\mathbf{d}_{\varepsilon_1}^T(-\omega_s/\beta)\exp(-i\omega_s\alpha).
\end{aligned}
$$

From Lemma 2 the second, third, and fourth terms converge to 0, uniformly on Θ and almost surely. Further, from Lemma 3 the first term converges uniformly to

$$
(7.13)\quad Q(\theta) = \int_{-\pi}^{\pi} \exp(i\eta(\alpha_0-\alpha))\frac{\exp(i\eta(c_0-c))-1}{i\eta(c_0-c)}\psi(\eta)f_s(\eta)\,d\eta.
$$

Because $\psi(\eta)$ and $f_s(\eta)$ are symmetric functions, $Q(\theta)$ is a real-valued function.

The theorem can be proven by following a classical argument [Jennrich (1969)], if we can show that $Q(\theta)$ has a unique maximum at $\theta=\theta_0$.

Since $|\exp(i\eta(c_0-c))-1|$ is the distance between 1 and $\exp(i\eta(c_0-c))$, and $|\eta(c_0-c)|$ is the length of the arc between 1 and $\exp(i\eta(c_0-c))$ on the unit circle, $Q(\theta)$ will not attain the maximum when $c\neq c_0$.

Consider

$$
(7.14)\quad
\begin{aligned}
Q((\alpha,c_0)) &= \int_{-\pi}^{\pi}\exp(i\eta(\alpha_0-\alpha))\psi(\eta)f_s(\eta)\,d\eta\\
&= 2\int_{-\pi}^{\pi}\cos(\eta(\alpha_0-\alpha))\psi(\eta)f_s(\eta)\,d\eta.
\end{aligned}
$$

If $\alpha_0\neq\alpha$, then we can find an interval J contained in the set, $\{\omega\colon \psi(\omega)f_s(\omega)>0\}$ such that $\cos(\eta(\alpha_0-\alpha))<1$ for $\eta\in J$. Therefore the only maximum is located at θ_0. □

Next, we prove Theorem 2 concerning the asymptotic distribution of the estimate.

PROOF OF THEOREM 2. When T is large enough, the estimate $\hat{\theta}$ will be in the interior of Θ, and $\partial Q_T(\hat{\theta})/\partial\alpha = \partial Q_T(\hat{\theta})/\partial c = 0$. Therefore, we can find points θ' and θ'' between $\hat{\theta}$ and θ_0, such that

$$
(7.15)\quad -T^{3/2}\begin{pmatrix}\dfrac{\partial Q_T(\theta_0)}{\partial\alpha}\\[2mm] \dfrac{\partial Q_T(\theta_0)}{\partial c}\end{pmatrix} = T^{-2}\begin{pmatrix}\dfrac{\partial^2 Q_T(\theta')}{\partial\alpha\,\partial\alpha} & \dfrac{\partial^2 Q_T(\theta')}{\partial\alpha\,\partial c}\\[2mm] \dfrac{\partial^2 Q_T(\theta'')}{\partial\alpha\,\partial c} & \dfrac{\partial^2 Q_T(\theta'')}{\partial c\,\partial c}\end{pmatrix}\sqrt{T}\begin{pmatrix}\hat{\alpha}-\alpha_0\\ \hat{c}-c_0\end{pmatrix}.
$$

Firstly, we should find the asymptotic distributions of

$$(7.16)\qquad \frac{\partial Q_T(\theta)}{\partial \alpha} = \sum_{|s|<T/2} -i\omega_s\psi(\omega_s)\mathbf{d}^T_{X_2}(\omega_s)\mathbf{d}^T_{X_1}(-\omega_s/\beta)\exp(-i\alpha\omega_s)$$

and of

$$(7.17)\qquad \frac{\partial Q_T(\theta)}{\partial c} = \sum_{|s|<T/2} i\omega_s\psi(\omega_s)\mathbf{d}^T_{X_2}(\omega_s)\tilde{d}^T_{X_1}(-\omega_s/\beta)\exp(-i\alpha\omega_s).$$

Here

$$(7.18)\qquad \tilde{d}^T_{X_1}(\lambda) = \sum_{t=0}^{T-1}\frac{t}{T}X_1(t)\exp(-i\lambda t),$$

and as before $\beta^{-1} = 1 + c/T$. [We define $\tilde{d}^T_S(\lambda)$ and $\tilde{d}^T_{\varepsilon_1}$ similarly.]

From Lemma 1 we find that the expected values of

$$(7.19)\qquad \sum_{|s|<T/2} \omega_s\psi(\omega_s)\mathbf{d}^T_s(\theta_0,\omega_s)\tilde{d}^T_S(-\omega_s/\beta_0)\exp(-i\alpha_0\omega_s)$$

and

$$(7.20)\qquad \sum_{|s|<T/2} \omega_s\psi(\omega_s)\mathbf{d}^T_s(\theta_0,\omega_s)\mathbf{d}^T_s(-\omega_s/\beta_0)\exp(-i\alpha_0\omega_s)$$

are of order $T\log T$. Therefore

$$(7.21)\qquad \lim_{T\to\infty} T^{-3/2}E\left(\frac{\partial Q_T(\theta_0)}{\partial\alpha}\right) = \lim_{T\to\infty} T^{-3/2}E\left(\frac{\partial Q_T(\theta_0)}{\partial c}\right) = 0.$$

In order to derive the covariance matrix of $T^{-3/2}(\partial Q_T(\theta_0)/\partial\alpha)$ and $T^{-3/2}(\partial Q_T(\theta_0)/\partial c)$, we need to compute the values of the cumulants.

$$(7.22)\qquad \begin{aligned}
&\operatorname{cum}\Big[\sum_s i\omega_s\psi(\omega_s)\mathbf{d}^T_s(\theta_0,\omega_s)\mathbf{d}^T_s(\theta_0,-\omega_s),\\
&\qquad \sum_s i\omega_s\psi(\omega_s)\mathbf{d}^T_s(\theta_0,\omega_s)\mathbf{d}^T_s(\theta_0,-\omega_s)\Big],\\
&\operatorname{cum}\Big[\sum_s -i\omega_s\psi(\omega_s)\mathbf{d}^T_s(\theta_0,\omega_s)\mathbf{d}^T_s(\theta_0,-\omega_s),\\
&\qquad \sum_s i\omega_s\psi(\omega_s)\mathbf{d}^T_s(\theta_0,\omega_s)\tilde{d}^T_S(\theta_0,-\omega_s)\Big],\\
&\operatorname{cum}\Big[\sum_s i\omega_s\psi(\omega_s)\mathbf{d}^T_s(\theta_0,\omega_s)\tilde{d}^T_S(\theta_0,-\omega_s),\\
&\qquad \sum_s i\omega_s\psi(\omega_s)\mathbf{d}^T_s(\theta_0,\omega_s)\tilde{d}^T_S(\theta_0,-\omega_s)\Big].
\end{aligned}$$

It is easy to see that the first one and the second one are of order $o(T^3)$. The third one is equal to

$$\begin{aligned}
&\sum_{s_1}\sum_{s_2} - \omega_{s_1}\omega_{s_2}\psi(\omega_{s_1})\psi(\omega_{s_2})\mathrm{cum}\big[\mathbf{d}_s^T(\theta_0,\omega_{s_1})\tilde{d}_S^T(\theta_0,-\omega_{s_1}),\\
&\qquad\qquad \mathbf{d}_s^T(\theta_0,\omega_{s_2})\tilde{d}_S^T(\theta_0,-\omega_{s_2})\big]\\
&= \sum_{s_1}\sum_{s_2} - \omega_{s_1}\omega_{s_2}\psi(\omega_{s_1})\psi(\omega_{s_2})\mathrm{cum}\big[\mathbf{d}_s^T(\theta_0,\omega_{s_1}),\tilde{d}_S^T(\theta_0,-\omega_{s_2})\big]\\
&\quad\times \mathrm{cum}\big[\tilde{d}_S^T(\theta_0,-\omega_{s_1}),\mathbf{d}_s^T(\theta_0,\omega_{s_2})\big]\\
&\quad+ \sum_{s_1}\sum_{s_2} - \omega_{s_1}\omega_{s_2}\psi(\omega_{s_1})\psi(\omega_{s_2})\mathrm{cum}\big[\mathbf{d}_s^T(\theta_0,\omega_{s_1}),\mathbf{d}_s^T(\theta_0,\omega_{s_2})\big]\\
&\quad\times \mathrm{cum}\big[\tilde{d}_S^T(\theta_0,-\omega_{s_1}),\tilde{d}_S^T(\theta_0,-\omega_{s_2})\big].
\end{aligned} \tag{7.23}$$

The second term of (7.23) is equal to $(2\pi T^3/3)\int_{-\pi}^{\pi}\lambda^2\psi^2(\lambda)f_s(\lambda) + O(T^2\log T)$. The first term of (7.23) is equal to

$$\begin{aligned}
&\sum_{s_1}\sum_{s_2} - 4\pi^2\omega_{s_1}\omega_{s_2}\psi(\omega_{s_1})\psi(\omega_{s_2})\\
&\qquad\times H_1^T(\omega_{s_1}-\omega_{s_2})H_1^T(\omega_{s_2}-\omega_{s_1})f_s^2(\omega_{s_1}) + O(T^2).
\end{aligned} \tag{7.24}$$

Because

$$\begin{aligned}
H_1^T(\lambda) &= \sum_{t=0}^{T-1} t/T\exp(-i\lambda t)\\
&= \frac{\exp(-i\lambda)-\exp(-i\lambda T)}{T[1-\exp(-i\lambda)]^2} - \frac{(T-1)\exp(-i\lambda T)}{T(1-\exp(-i\lambda))},
\end{aligned} \tag{7.25}$$

which is equal to $iT/2\pi s$ at $\lambda = \omega_s = 2\pi s/T$, $s \neq 0$, expression (7.24) is equal to

$$\begin{aligned}
&\sum_s - \pi^2\omega_s^2\psi^2(\omega_s)f_s^2(\omega_s) + \sum_{s_1\neq s_2}\sum \frac{4\pi^2 s_1 s_2}{(s_1-s_2)^2}f_s^2(\omega_{s_1}) + O(T^2)\\
&= -\left(\frac{2\pi T^3}{4} + \frac{2\pi T^3}{12}\right)\int_{-\pi}^{\pi}\lambda^2\psi^2(\lambda)f_s^2(\lambda)\,d\lambda + o(T^3)\\
&= \frac{-2\pi T^3}{3}\int_{-\pi}^{\pi}\lambda^2\psi^2(\lambda)f_s^2(\lambda)\,d\lambda + o(T^3).
\end{aligned} \tag{7.26}$$

The first equality of (7.26) holds since $\sum_{s=1}^{\infty}1/s^2 = \pi^2/6$.

It may now be seen that the covariance matrix of $(\partial Q_T(\theta_0)/\partial\alpha, \partial Q_T(\theta_0)/\partial c)$ converges to

$$p\begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{3} \end{pmatrix} \tag{7.27}$$

with

$$p = 2\pi \int_{-\pi}^{\pi} \lambda^2 \psi^2(\lambda) f_s(\lambda) \left[f_{\varepsilon_1}(\lambda) + f_{\varepsilon_2}(\lambda) \right] d\lambda \tag{7.28}$$
$$+ 2\pi \int_{-\pi}^{\pi} \lambda^2 \psi^2(\lambda) f_{\varepsilon_1}(\lambda) f_{\varepsilon_2}(\lambda)\, d\lambda.$$

From Lemma 3 it may be seen that if $\theta_T \to \theta_0$ as $T \to \infty$, then $T^{-2}\, \partial^2 Q_T(\theta_T)/\partial \alpha^2$, $T^{-2}\, \partial^2 Q_T(\theta_T)/\partial c^2$, and $T^{-2}\, \partial^2 Q_T(\theta_T)/\partial \alpha\, \partial c$ converge, almost surely, to $-d$, $d/2$, and $-d/3$, respectively, with

$$d = \int_{-\pi}^{\pi} \lambda^2 \psi(\lambda) f_s(\lambda)\, d\lambda. \tag{7.29}$$

Note that

$$\begin{pmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{3} \end{pmatrix}^{-1} = \begin{pmatrix} 4 & 6 \\ 6 & 12 \end{pmatrix}, \tag{7.30}$$

giving the covariance matrix indicated.

Next, since the kth-order cumulants of $T^{-3/2}(\partial Q_T(\theta_0)/\partial \alpha, \partial Q_T(\theta_0)/\partial c)$ are of order $T^{1-k/2}(\log T)^k$, these converge to zero for $k > 2$. Therefore, $T^{-3}(\partial Q_T(\theta_0)/\partial \alpha, \partial Q_T(\theta_0)/\partial c)$ is asymptotically normal and the proof is finished. □

We next prove Theorem 3.

THEOREM 3. *Under Assumptions* 1–4, *let* $\theta_T \to \theta_0$ *and* $M_T \to \infty$, $M_T/T \to 0$. *Then* $\lim_{T\to\infty} E|\hat{f}_s(\lambda_0) - f_s(\lambda_0)|^2 = 0$, *where*

$$\hat{f}_s(\lambda_0) = \frac{1}{2\pi} \frac{1}{M_T T} \sum_{\omega_s \in I_M} \mathbf{d}_{X_2}^T(\omega_s) \mathbf{d}_{X_1}^T(-\omega_s/\beta_T) \exp(-i\alpha_T \omega_s) \tag{7.31}$$

and I_{M_T} *is the set containing* M_T *Fourier frequencies which are closest to* λ_0. [*Here* $\beta_T^{-1} = 1 + c_T/T$ *and* $\theta_T = (\alpha_T, c_T)$.]

PROOF OF THEOREM 3. First we note that, as in the proof of Lemma 1,

$$E\left[\mathbf{d}_s^T(\theta_0, \omega_s)\mathbf{d}_s^T(-\omega_s/\beta_T)\right] \tag{7.32}$$
$$= E\left[\mathbf{d}_s^T(\omega_s/\beta_0)\mathbf{d}_s^T(-\omega_s/\beta_T)\right]\exp(i\alpha_0\omega_s) + O(1).$$

Now, we calculate the expected value of $\mathbf{d}_s^T(\omega_s/\beta_0)\mathbf{d}_s^T(-\omega_s/\beta_T)$. It is

$$2\pi\Delta^T\left(\frac{\omega_s}{\beta_0} - \frac{\omega_s}{\beta_T}\right) f_s(\lambda_0) + O(1), \tag{7.33}$$

where

$$\Delta^T(\lambda) = \sum_{t=0}^{T-1} \exp(-i\lambda t) = \frac{1 - \exp(-i\lambda T)}{1 - \exp(-i\lambda)}. \tag{7.34}$$

Since $\Delta^T(\omega_s/\beta_0 - \omega_s/\beta_T) = T + o(T)$, we get

$$E\big(\mathbf{d}_s^T(\omega_s/\beta_0)\mathbf{d}_s^T(-\omega_s/\beta_T)\big) = 2\pi T f_s(\lambda_0) + o(T). \tag{7.35}$$

We therefore have

$$E(\hat{f}_s(\lambda_0)) = \frac{1}{2\pi}\frac{1}{MT}\left[\sum_{\omega_s \in I_M} 2\pi T f_s(\lambda_0) + o(T)\right] = f_s(\lambda_0) + o(1). \tag{7.36}$$

Next we want to find the variance of $\hat{f}_s(\lambda)$. We first show that the covariance of $\mathbf{d}_s^T(\omega_{s1}/\beta_0)\mathbf{d}_s^T(-\omega_{s_1}/\beta_T)$ and $\mathbf{d}_s^T(\omega_{s_2}/\beta_0)\mathbf{d}_s^T(-\omega_{s_2}/\beta_T)$ for $\omega_{s_1}, \omega_{s_2} \in I_{M_T}$ and $\omega_{s_1} \neq \omega_{s_2}$ can be neglected. This can be seen from the following argument. We note that

$$\begin{aligned}
&\operatorname{cum}\big(\mathbf{d}_s^T(\omega_{s_1}/\beta_0)\mathbf{d}_s^T(-\omega_{s_1}/\beta_T), \mathbf{d}_s^T(-\omega_{s_2}/\beta_0)\mathbf{d}_s^T(\omega_{s_2}/\beta_T)\big)\\
&= \operatorname{cum}\big(\mathbf{d}_s^T(\omega_{s_1}/\beta_0), \mathbf{d}_s^T(-\omega_{s_2}/\beta_T)\big)\operatorname{cum}\big(\mathbf{d}_s^T(-\omega_{s_1}/\beta_T), \mathbf{d}_s^T(\omega_{s_2}/\beta_0)\big)\\
&\quad + \operatorname{cum}\big(\mathbf{d}_s^T(\omega_{s_1}/\beta_0), \mathbf{d}_s^T(\omega_{s_2}/\beta_0)\big)\operatorname{cum}\big(\mathbf{d}_s^T(-\omega_{s_1}/\beta_T), \mathbf{d}_s^T(-\omega_{s_2}/\beta_T)\big).
\end{aligned} \tag{7.37}$$

The second term is bounded. For the first term, we have

$$\begin{aligned}
\operatorname{cum}\big(\mathbf{d}_s^T(\omega_{s_1}/\beta_0), \mathbf{d}_s^T(-\omega_{s_2}/\beta_T)\big) &= 2\pi\Delta^T\left(\frac{\omega_{s_1}}{\beta_0} - \frac{\omega_{s_2}}{\beta_T}\right) f_s(\lambda_0) + O(1)\\
&= o\big((\omega_{s_1} - \omega_{s_2})^{-1}\big).
\end{aligned} \tag{7.38}$$

Hence we have

$$\begin{aligned}
&\operatorname{cum}\big(\mathbf{d}_s^T(\theta_0, \omega_{s_1})\mathbf{d}_s^T(-\omega_{s_1}/\beta_T), \mathbf{d}_s^T(\theta_0, -\omega_{s_2})\mathbf{d}_s^T(\omega_{s_2}/\beta_T)\big)\\
&= o\big((\omega_{s_1} - \omega_{s_2})^{-2}\big).
\end{aligned} \tag{7.39}$$

Then, it can be seen that the variance of $\hat{f}_s(\lambda_0)$ is

$$M^{-1}\{f_s(\lambda_0)[f_{\varepsilon_1}(\lambda_0) + f_{\varepsilon_2}(\lambda_0) + f_s(\lambda_0)] + f_{\varepsilon_1}(\lambda_0)f_{\varepsilon_2}(\lambda_0)\}. \tag{7.40}$$

This finishes the proof of the theorem. □

Acknowledgments. This paper is part of the author's doctoral dissertation which was written at the University of California, Berkeley under the supervision of Professor David R. Brillinger. Professor Brillinger's guidance, encouragement, and support are gratefully acknowledged. The author would also like to thank Professors Bruce Bolt and Kjell Doksum for their valuable comments. The author also appreciates the helpful suggestions given by Drs. Norman Abrahamson, Ross Ihaka, and George Terrell. The valuable comments from the associate editor and the referees are gratefully appreciated. Finally, the author would like to thank Dr. Paul Spudich for providing the seismic data.

REFERENCES

AKI, K. (1967). Scaling law of seismic spectrum. *J. Geophys. Res.* **73** 1217–1231.

BOLT, B. A. (1981). Interpretation on strong ground motion records. Miscellaneous Paper S-73-1, Report 17, U.S. Army Engineers Waterways Experiment Station, Vicksburg, Miss.

BOORE, D. M. (1977). The motion of the ground in earthquakes. *Sci. Amer.* **237** 68–78.

BRILLINGER, D. R. (1975). *Time Series Data Analysis and Theory.* Holt, Rinehart, and Winston, New York.

BYCROFT, G. N. (1982). El Centro differential ground motion array. In *The Imperial Valley, California, Earthquake of October 15, 1979, U.S. Geol. Surv. Prof. Pap. 1254.* 351–356.

CARTER, G. C. (1981). Time delay estimation for passive sonar signal processing. *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-29** 463–470.

CHIU, S. T. (1984). Statistical estimation of the parameters of a moving source from array data. Ph.D. dissertation, Univ. California, Berkeley.

HANNAN, E. J. (1975). Measuring the velocity of a signal. In *Perspectives in Probability and Statistics* (J. Gani, ed.) 227–237. Academic, New York.

HANNAN, E. J. and ROBINSON, P. M. (1973). Lagged regression with unknown lags. *J. Roy. Statist. Soc. Ser. B* **35**, 252–267.

HANNAN, E. J. and THOMSON, P. J. (1981). Delay estimation and the estimation of coherence and phase. *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-29** 485–490.

JENNRICH, R. I. (1969). Asymptotic properties of non-linear least squares estimators. *Ann. Math. Statist.* **40** 633–643.

KNAPP, C. H. and CARTER, G. C. (1977). Estimation of time delay in the presence of source or receiver motion. *J. Acoust. Soc. Amer.* **61** 1545–1549.

REID, H. F. (1910). Mechanics of the earthquake. In *The California Earthquake of April 18, 1906* **2**. Carnegie Inst. of Washington, D.C.

SCHULTHEISS, P. M. and WEINSTEIN, E. (1979). Estimation of differential Doppler shifts. *J. Acoust. Soc. Amer.* **66** 1412–1419

SPUDICH, P. and CRANSWICK E. (1984). Direct observation of rupture propagation during the 1979 Imperial Valley earthquake using a short baseline accelerometer array. *Bull. Seism. Soc. Amer.* **74** 2083–2114.

THOMSON, P. J. (1982). Signal estimation using an array of recorders. *Stochastic Process. Appl.* **13** 201–214.

DEPARTMENT OF MATHEMATICAL SCIENCES
RICE UNIVERSITY
HOUSTON, TEXAS 77251