To my wife Cheryl

THE MEASUREMENT OF VELOCITY

by

John D. Henstridge

A thesis submitted to the

Australian National University

for the degree of

Doctor of Philosophy

January 1977

ACKNOWLEDGEMENTS

I thank Professor E.J. Hannan who supervised my work for his continual guidance, suggestions and encouragement.

I also thank Professor P.A. Moran who supervised this work in Professor Hannan's absence, and Dr Cheryl E. Praeger and Professor T.P. Speed who read drafts of this thesis. Special thanks also go to Mr M. Asten whose advice on the geophysical application was invaluable.

This work was carried out at the Australian National University with the financial support of a Commonwealth Postgraduate Scholarship and a Public Service Board Financial Assistance Award. All of these I gratefully acknowledge.

Finally I wish to thank Marian Currie for her excellent typing.

ABSTRACT

In the past forty years in many fields of physics and geophysics, signals, usually in the form of waves, have been observed by using arrays of receivers or detectors. The practical problems of estimating the velocity and direction of a signal observed with such an array, in the presence of noise not correlated between receivers, have been solved by many people and the statistical theory has been covered by E.J. Hannan. This thesis firstly extends this statistical theory to the situation where the noise is correlated between receivers.

Many *ad hoc* attempts have been made to cover situations where more than one signal is present. We derive several consistent estimators of velocity and direction for this situation and discuss their asymptotic properties. These estimators are further examined using simulations and it is shown that they can separate signals which the array cannot resolve in the conventional sense. Also several methods of preliminary analysis are discussed using particular designs of arrays.

NOTATION

Where possible the notation of Hannan (1970) has been followed.

The following matrix notation is used

A^{-1}	Inverse of A.
A^{*}	Conjugate transpose of A.
A'	Transpose of A.
A_{jk}	The (j,k) element of A.
$A \otimes B$	Kronecker product of A and B.

Within chapters, items such as lemmas and theorems are numbered sequentially regardless of type. The end of such items is denoted by //.

TABLE OF CONTENTS

page

INTRODUCTION

In the past forty years in many fields ranging from radio astronomy to oceanography it has been found useful to observe signals, usually in the form of waves, by using several receivers or observation points simultaneously. The set of receivers and its associated processing equipment is generally called an *array*. Until recently two different approaches dominated the field. The first approach was that used in early radar work: arrays sensitive only to a narrow band of frequencies were used and the relative phase of the signal was measured at the receivers. Averages over a short period of time were used to obtain accuracy. The second approach was that used in seismology where only transient signals (earthquake tremors) were observed and the time of first arrival at each receiver was measured.

In recent years attempts have been made to extend the phase measurement methods to much wider ranges of frequencies. This has only become possible with the development of electronic computers. It was first applied in oceanography where the frequencies and hence the rates of data collection were low enough for early computers (Munk *et al.*, 1962). More recently similar work has been done in seismology and sonar studies.

The aims of these measurements are to estimate the velocity or direction from which the signals arrive, either to gain information about the signals themselves or about the medium through which they are passing. Of secondary interest are problems of signal detection

and estimation of signal spectra.

The situation of one signal observed with noise which is not correlated between receivers has been solved in practice by many people since about 1955 and its statistical theory has been treated by Hannan (1975). One signal observed with noise which is correlated between receivers has been discussed by Widrow *et al.* (1967) in a signal extraction context where the signal velocity and direction are known.

Many attempts have been made to treat the case where more than one signal is present. Some *ad hoc* methods were suggested by Munk *et al.* (1962). Hinich and Shaman (1972) obtained some asymptotic results for regular arrays by increasing the size of the array. Various approximate methods are commonly used in practice. The most notable of these is the "maximum likelihood method" of Capon (1969).

This thesis firstly extends, in Chapter 2, the estimation theory developed by Hannan to the situation where the noise is correlated between receivers. It is shown that this is really only possible if the correlation structure is known. Several consistent estimators of velocity and direction are derived.

The next chapter (Chapter 3) derives and discusses several estimators for the case of more than one signal and uncorrelated noise. It is shown that the maximum likelihood estimator is not likely to be very useful in practice since it requires an accurate knowledge of the spectra of all the processes involved. To overcome this several other less efficient estimators are proposed. Asymptotic results are obtained for these with a fixed array size. In Chapter 4 these asymptotic results are compared with the results of simulation studies. Asymptotic variances are found to be reasonable guides to small sample variances. Comparisons are also

made with Capon's method.

A different approach is developed in Chapter 5 where a method first proposed by Aki (1957, 1961) is used to obtain preliminary estimates of velocity, avoiding the nonlinear optimization problems of other methods. Finally, in Chapter 6, some illustrations of applying the methods of this thesis to some seismic data are given.

CHAPTER 1

PRELIMINARY DISCUSSION

1.1 REPRESENTATION OF SIGNALS AND NOISE

We consider an array of p receivers at locations $\xi(1)$, $\xi(2)$ . . ., $\xi(p)$ in Euclidean space. Usually this space will be R^2 corresponding to part of the earth's surface or R^3 corresponding to real physical space.

At the origin of this space (or at some fixed reference point) the signal can have the spectral representation

$$z(t) = \int_{-\infty}^{\infty} e^{-i\omega t}\, d\xi(\omega) \quad , \qquad [1]$$

where $\xi(\omega)$ is the *spectral process* with orthogonal increments. The signal will pass each receiver at some time before or after it passes the origin. This delay could vary with frequency. A general representation of the signal at the a^{th} receiver would be

$$z(t,a) = \int_{-\infty}^{\infty} e^{-i\{\omega t - \theta(\omega,a)\}}\, d\xi(\omega) \ .$$

$\theta(\omega,a)$ is called the *phase delay* at frequency ω and receiver a.

There are several common models for $\theta(\omega,a)$. Consider, as in Figure 1.1, a plane wave travelling in direction ϕ (a unit direction vector) with frequency ω and phase velocity $c(\omega)$. As can be seen from Figure 1.1, after passing the origin the plane of the wave must

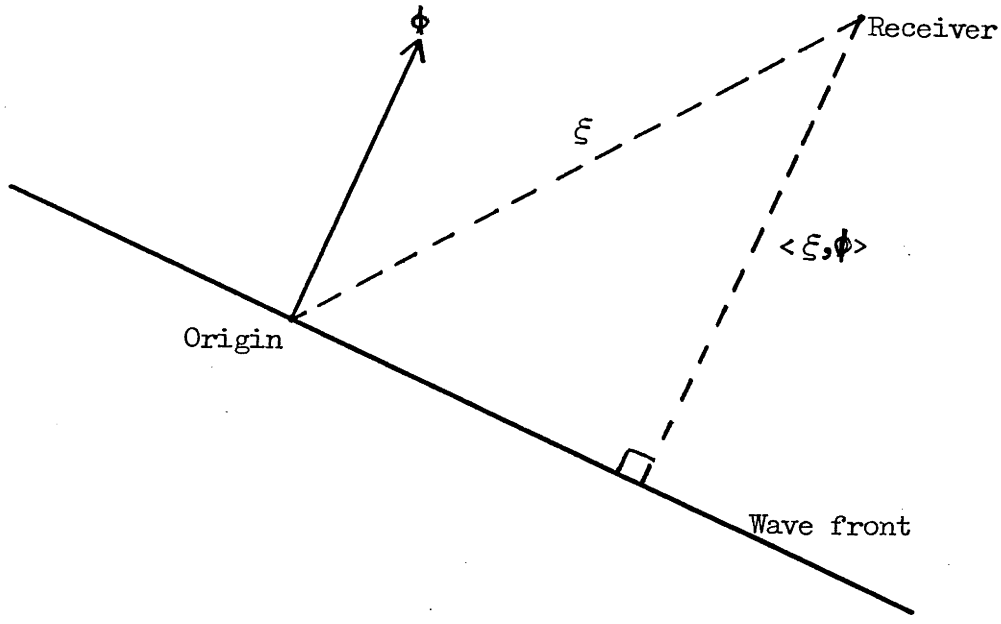


Figure 1.1

travel a further distance of $\langle\xi(a),\phi\rangle$ before passing $\xi(a)$. The time taken to do this is $\langle\xi(a),\phi\rangle/c(\omega)$ which corresponds to $\omega\langle\xi(a),\phi\rangle/c(\omega)$ radians of a sinoid of frequency ω. Hence

$$\theta(\omega,a) = \omega\langle\xi(a),\phi\rangle/c(\omega) \quad .$$

It is often useful to combine the directional and velocity information into one vector quantity, the *slowness* vector $\mu(\omega) = \phi/c(\omega)$. Then

$$\theta(\omega,a) = \omega\langle\xi(a),\mu(\omega)\rangle \quad .$$

Each component of the slowness vector gives the reciprocal of the velocity component in each basis direction. In a non-dispersive situation where the phase velocity is constant with frequency, a single slowness vector would fully describe the propagation of the signal.

Because of its interpretation as spatial frequency, the *wavenumber* vector, given by $\kappa(\omega) = \omega\mu(\omega)$ is often used. Then

$$\theta(\omega,a) = \langle \xi(a), \kappa \rangle \tilde{}$$

Throughout this thesis the wavenumber will be as defined here, with units of radians per unit distance, not cycles per unit distance as is sometimes done. It can be seen that even in non-dispersive situations the wavenumber varies with frequency, a clear disadvantage when compared with slowness. We shall use the scalar wavenumber $k(\omega) = \|\kappa(\omega)\| = \omega/c(\omega)$.

It is inherent in the above discussion that since we are using phase delay methods, it is the phase velocity which is important. Methods based on first arrival times usually give group velocities $u(\omega)$ where $u(\omega) = \frac{d\omega}{dk}$. However phase techniques can be used to estimate group velocities if desired (Hannan and Thomson, 1973).

Finally we note the following relations

$$k = \frac{2\pi}{\lambda} \quad , \quad T = \frac{2\pi}{\omega} \quad , \quad \frac{\lambda}{T} = c(\omega)$$

where λ is the wavelength and T the period of a wave.

We shall be considering s stationary signals observed with noise at each of the receivers. The assumption that the signals and the noise are independent of each other shall be made throughout. The observed process is

$$y(t,a) = x(t,a) + \sum_{j=1}^{s} z_j(t,a) \quad , \qquad [2]$$

$$a = 1, 2, \ldots, p \quad ,$$

and

$$z_j(t,a) = \int_{-\infty}^{\infty} e^{-i\{\omega t - \theta_j(\omega,a)\}} \, d\xi_j(\omega) \qquad [3]$$

If we let $z_j(t) = (z_j(t,1), z_j(t,2), \ldots, z_j(t,p))'$ and $\theta_j(\omega) = (e^{i\theta_j(\omega,1)}, e^{i\theta_j(\omega,2)}, \ldots, e^{i\theta_j(\omega,p)})'$ then [3] can be

expressed in vector form

$$z_j(t) = \int_{-\infty}^{\infty} e^{-i\omega t}\, \theta_j(\omega)\, d\xi_j(\omega)$$

In this form it can be seen that this vector process is simply a linearly filtered version of the signal process observed at the origin.

If the vector processes $y(t) = (y(t,1), y(t,2), \ldots, y(t,p))'$ and $x(t) = (x(t,1), x(t,2), \ldots, x(t,p))$ then [2] reduces to

$$y(t) = x(t) + \sum_{y=1}^{s} z_j(t)$$

If both the signal and noise processes have absolutely continuous spectra and $F_x(\omega)$ is the spectral density matrix of the noise, $f_j(\omega)$ is the spectral density of the j^{th} signal then the spectral density of matrix $F_y(\omega)$ of the observed process is

$$F_y(\omega) = F_x(\omega) + \sum_{j=1}^{s} f_j(\omega)\, \theta_j(\omega)\, \theta_j(\omega)^*$$

This can be further simplified if we let $F(\omega) = \text{diag}\,(f_1(\omega), f_2(\omega), \ldots, f_s(\omega))$, that is a diagonal matrix with $f_j(\omega)$ in the j^{th} position and $\theta(\omega) = (\theta_1(\omega), \theta_2(\omega), \ldots, \theta_s(\omega))$. Then

$$F_y(\omega) = F_x(\omega) + \theta(\omega)\, F(\omega)\, \theta(\omega)^*$$

The matrix $F(\omega)$ can be considered to be the spectral density of the vector process $\tilde{z}(t) = (\tilde{z}_1(t), \tilde{z}_2(t), \ldots, \tilde{z}_s(t))'$ where $\tilde{z}_j(t)$ is the j^{th} signal as observed at the origin. The assumption that $F(\omega)$ is diagonal corresponds to assuming that the signals are not correlated with each other.

The vectors $\theta_j(\omega)$ shall be called *phase vectors* and the matrix $\theta(\omega)$ the *phase matrix*.

When reference is made to correlated noise, it will be meant

that the noise at different receivers is correlated and that $F_x(\omega)$ is an arbitrary positive definite matrix. The noise will still be considered independent of the signals. References to uncorrelated noise will mean that $F_x(\omega)$ is diagonal. Often it will be assumed that the noise spectra at each of the receivers is identical and hence $F_x(\omega) = f_x(\omega)\, I_p$.

1.2 DISCRETE TIME EFFECTS AND ESTIMATION PROBLEMS

In practice the process is not observed in continuous time but rather the output of each receiver is sampled at regular discrete time intervals. The frequency of sampling must be high enough to retain all the information in the frequency range desired without significant aliasing. With no loss of generality the sampling interval can be assumed to be unity, giving spectra defined on the interval $[-\pi,\pi]$.

Aliasing of noise presents no problem since aliased noise still looks like noise. If the higher frequency components of a signal are aliased then these can be modelled as extra signals. A signal with a frequency range up to $k\pi$ can be represented as k separate signals. The spectral representation theorem implies that these components will not be correlated with each other.

Alternatively aliasing may only interfere with the signal over a limited frequency range. In this case such frequencies can simply be ignored. Due to this and other forms of interference which could be present our analysis shall always be restricted to a set B of frequencies such that (a) $B \subsetneq [-\pi,\pi]$, (b) B is symmetric about zero and (c) B is a finite union of open sets.

Instead of using the data directly, we shall be using the

discrete Fourier transforms, given (for a sample of size N) by

$$Y_a(\omega) = (2\pi N)^{-\frac{1}{2}} \sum_{n=1}^{N} y(n,a)\, e^{in\omega}$$

or in vector form

$$Y(\omega) = (2\pi N)^{-\frac{1}{2}} \sum_{n=1}^{N} y(n)\, e^{in\omega} \quad .$$

This transforms the data into "frequency space" where it can be directly compared with the spectral models. The Fourier transform is evaluated only at the so-called *fundamental frequencies*

$$\omega_j = \frac{2\pi j}{N} \quad , \quad j = -\frac{N}{2} \quad , \ldots, \quad \frac{N}{2} \quad .$$

It will be assumed that the phase delays for each signal at each receiver are all functions of a single *parameter vector* τ. This vector will contain all the unknown direction and velocity parameters for all the signals and it is this which we wish to estimate. The true value of τ we shall denote by τ_0. It will be assumed that τ is confined to a compact set with non-empty interior in Euclidean space.

This approach can readily cover all the models discussed in the previous section. By having a single parameter vector instead of one for each signal it is possible to model, for example, several signals with unknown directions and common but unknown velocity.

1.3 NARROWBAND ARRAYS

To appreciate the relevance of the results in the following chapters, it is necessary to discuss the construction and properties of real-time arrays which are used in fields such as radar.

Most receivers consist of an oscillator physically coupled to the medium. When the medium vibrates due to the passage of a wave,

the oscillator also vibrates and in a band of frequencies close to its resonant frequency it can greatly amplify the vibrations. The less the oscillator is dampened, the greater is this amplification, although over a narrower band of frequencies.

The ease of constructing such receivers meant that early arrays tended to be constructed with them.

A signal arriving at the array can be extracted simply by introducing a phase shift at each receiver to compensate for the signals phase and then after appropriate weighting, summing the outputs. If the phase shifts are not correct for that signal then the power at the summed output will be reduced. The procedure is known as phase and sum *beamforming* and is shown schematically in Figure 1.2.

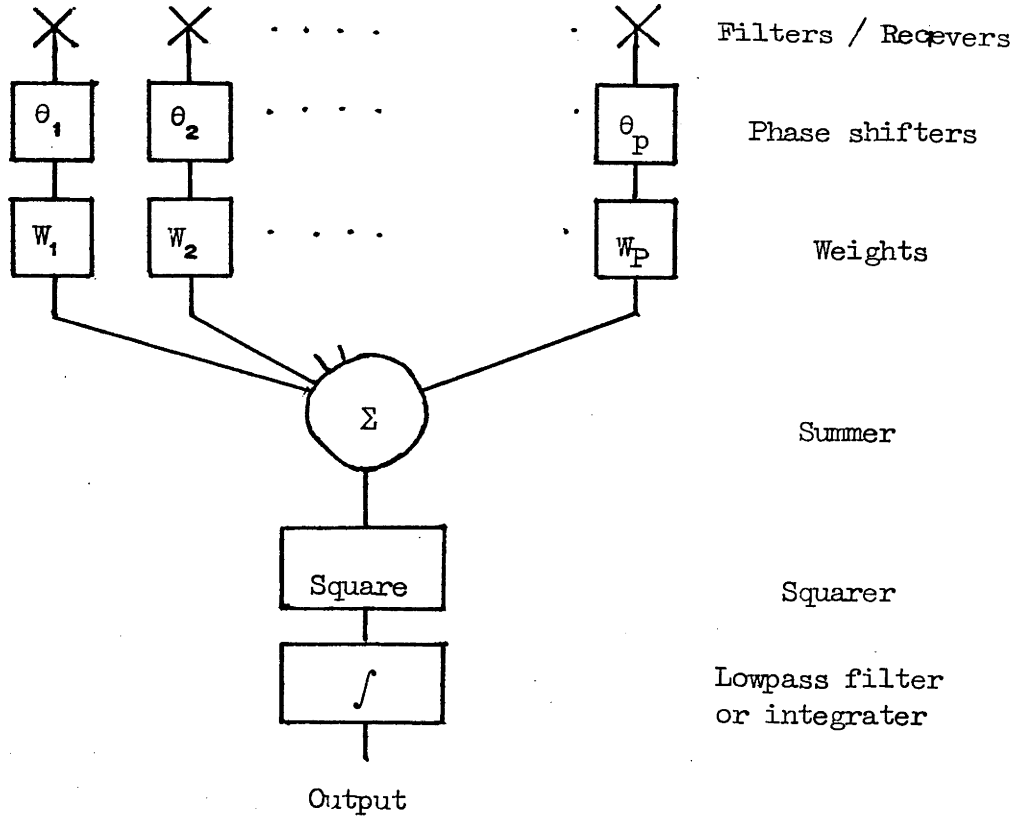


Figure 1.2

If the bandwidth is too wide, the phase shifts are not so effective (Knittel, 1974) and must be replaced by time delays. Time delay units have always created practical problems and the only real successes have been with digital delays (Anderson, 1960).

In operation the phase shifters are set so that they are correct for a signal arriving from a particular direction (or velocity). This direction is then varied so that the array can "look for" signals. This creates immediate problems with transient signals since they could be missed if the array is pointed in the wrong direction when they arrive. Also when the array is pointed in one direction, it is still sensitive to signals from other directions, a problem if more than one signal is present.

This sensitivity of the array to other directions is described by the *array function*. In the case where the signals are modelled in terms of their wavenumber, this is

$$S(\kappa) = \left| \sum_{a=1}^{p} W_a \, e^{i\langle \xi(a), \kappa \rangle} \right|^2$$

where the weights W_a are real and $\sum_{a=1}^{p} W_a = 1$. If the array is regarded as a spatial filter (Barber, 1961), or Schell (1969) then S is the power gain function.

S achieves its maximum of one when k = 0. This maximum which is called the *main beam* or *mainlobe*, corresponds to the direction in which the array is looking. Secondary maxima, called *sidelobes*, indicate significant sensitivity away from the main beam. A good array is considered to be one with small sidelobes and a narrow main beam. A rough measure of the narrowness of the main beam is given by the second derivatives of S at zero

$$\frac{\partial^2}{\partial \kappa_i \partial \kappa_j} S(0) = \sum_a \sum_b W_a W_b (\xi_i(a) - \xi_i(b))(\xi_j(a) - \xi_j(b)) \quad [4]$$

This is maximized by making the array as large as is possible. Unfortunately if the spacing between receivers is too great the sidelobes become dominant.

1.4 FIELDS OF APPLICATION

In most applications the signals we are concerned with are waves although this need not be the case. For instance the methods of this thesis could be used to study travelling nerve impulses.

It was mentioned earlier that radar and radio astronomy were the first fields of application of arrays. Due to the very high frequencies involved (1 MH_z - 10 GH_z) observation times can be very short and little is lost by operating in real time. Also it is almost impossible to sample fast enough to apply digital processing techniques.

Sonar is the acoustical equivalent to radar but due to the significantly lower frequencies (1 H_z to 20 kH_z) and much lower velocity (1.2 Km/sec) it is relatively simple to sample and record receiver outputs. In some applications where fast processing is a necessity, special hardwired electronic units are used for parts of the computation such as fast Fourier transforming. The arrays are usually working in three dimensions. The field is reviewed by Winder (1975).

In seismology arrays must be used since no worthwhile directional receiver exists. The frequencies are low (.01 H_z to 20 H_z) and wavelengths large (.05 Km to 1000 Km). Therefore arrays are very large and until recently receivers were not directly linked together so that the outputs had to be recorded and processed later. Usually only small sections of the record are of interest.

Seismic waves can be divided into two types--body waves which propagate in three dimensions through the depths of the earth, and surface waves which are confined to the surface of the earth. Both of these types can be further classified by the mode of vibration they involve. Body waves do not undergo significant dispersion. However, arrays used to observe them are confined to the two-dimensional earth's surface and body waves arriving from different angles appear to be travelling across the surface at different velocities. Surface waves are greatly dispersed and measurements of their velocity are very important methods of examining the earth's crust (Dziewonski and Hales, 1972). Further details can be found in Davies (1973) or Båth (1975).

Ocean waves have also been studied using arrays, notably by Munk *et al.* (1962). These are surface waves with predictable speed so that their directions are the main interest. They are one means of studying their sources, ocean storm centres.

CHAPTER 2

ONE SIGNAL ESTIMATES

2.1 ONE SIGNAL WITH UNCORRELATED NOISE

The case of one signal with noise uncorrelated between receivers was considered by Hannan (1975). In this case s=1 and θ is a vector of dimension p. Hannan made the following assumptions which we also shall use.

CONDITION 2.1

(a) The process y(n) is stationary and ergodic.

(b) Both the signal and noise processes have absolutely continuous spectra, and spectral density functions with a finite number of discontinuities.

(c) The parameter τ is confined to a compact subset of Euclidean space.

(d) $\theta(\omega,\tau)$ is continuous in ω and twice differentiable in τ in a neighbourhood of τ_0. //

The condition of ergodicity is made without cost since from one realization of the process it is not possible to decide whether the process is ergodic or not. If ergodicity is not assumed then the following results are conditional upon being in a particular ergodic class. It is impossible in practice to tell if other

ergodic classes exist or not.

Hannan defined the following function (which we shall call the *estimating function*),

$$q_N(\tau) = N^{-1} \sum_{\omega \in B} \sum_{a \neq b} W_{ab}(\omega)\, I_{ab}(\omega)\, e^{i\{\theta(\omega,a,\tau)-\theta(\omega,b,\tau)\}} \qquad [1]$$

where the summation is over those fundamental frequencies in the set B (see Section 1.2) and over all pairs of receivers. The periodogram $I(\omega)$ is defined by $I(\omega) = Y(\omega)\, Y(\omega)^*$, and $W_{ab}(\omega)$ are real valued weights. The only relevant weight functions are those for which $W_{ab}(\omega) = W_a(\omega)\, W_b(\omega)$. If such a weight function is used and the terms for which a=b are included (these do not depend upon τ) then [1] reduces to

$$q_N(\tau) = N^{-1} \sum_{\delta \in B} \left| \sum_a W_a\, Y_a(\omega)\, e^{-i\theta(\omega,a,\tau)} \right|^2 . \qquad [2]$$

This may be seen to imitate completely the functions of the processor discussed in Section 1.3. The Fourier transform of $y_a(n)$ can be thought of as a narrowband filtered output of the a^{th} receiver. It is rephased, weighted and these are then summed over a. The power from each frequency is combined to give the beam output. Changing the value of τ is equivalent to adjusting the array for a different velocity or direction.

As we would expect from this, our estimate for τ_0 is that value $\hat{\tau}_N$ which maximizes the estimating function. An identification condition is required to ensure that this global maximum is unique.

CONDITION 2.2

There exists no allowable $\tau \neq \tau_0$ such that for all $\omega \in B$ and a=1,2,..,p

$$\theta(\omega,a,\tau) - \theta(\omega,1,\tau) = \theta(\omega,a,\tau_0) - \theta(\omega,1,\tau_0) \quad .//$$

This corrects the condition as stated by Hannan.

Hannan gave two principal results. The first theorem states that if Conditions 2.1 and 2.2 are satisfied then the estimate $\hat{\tau}_N$ is consistent--that is $\underset{N\to\infty}{\ell im}\ \hat{\tau}_N = \tau_0$ almost surely. The second theorem also assumed a fourth moment condition and a mixing condition to obtain a central limit result. This stated that $N^{\frac{1}{2}}(\hat{\tau}_N - \tau_0)$ tended in distribution to a normal variable with zero mean and given variance matrix Σ.

For the case where $W_{ab}(\omega) = W(\omega)$, that is the receivers are weighted equally, and $F_x(\omega) = f_x(\omega)\, I_p$, then it follows from Hannan that $\Sigma = A^{-1}CA^{-1}$ where A,C are matrices with elements given by

$$A_{jk} = \frac{1}{\pi} \int_B W(\omega) \sum_a \sum_b \left(\frac{\partial}{\partial\tau_j}\theta(\omega,a,\tau_0) - \frac{\partial}{\partial\tau_k}\theta(\omega,b,\tau_0)\right)$$

$$\left(\frac{\partial}{\partial\tau_k}\theta(\omega,a,\tau_0) - \frac{\partial}{\partial\tau_k}\theta(\omega,b,\tau_0)\right) f_z(\omega)\, d\omega$$

and

$$C_{jk} = \frac{1}{\pi} \int_B W(\omega)^2 \sum_a \sum_b \left(\frac{\partial}{\partial\tau_j}\theta(\omega,a,\tau_0) - \frac{\partial}{\partial\tau_j}\theta(\omega,\ ,\tau_0)\right)$$

$$\left(\frac{\partial}{\partial\tau_k}\theta(\omega,a,\tau_0) - \frac{\partial}{\partial\tau_k}\theta(\omega,b,\tau_0)\right) f_x(\omega)(f_x(\omega) + pf_z(\omega))d\omega\,.$$

If the maximum likelihood weight function is used we have

$$W(\omega) = f_z(\omega)\, f_x^{-1}(\omega)\, (f_x(\omega) + pf_z(\omega))^{-1}$$

and

$$(\Sigma^{-1})_{jk} = \frac{1}{\pi} \int_B \frac{R(\omega)^2}{1+pR(\omega)} \sum_a \sum_a \left(\frac{\partial}{\partial\tau_j}\theta(\omega,a,\tau_0) - \frac{\partial}{\partial\tau_j}\theta(\omega,b,\tau_0)\right)$$

$$\left(\frac{\partial}{\partial\tau_k}\theta(\omega,a,\tau_0) - \frac{\partial}{\partial\tau_k}\theta(\omega,b,\tau_0)\right) d\omega$$

where $R(\omega) = f_z(\omega)/f_x(\omega)$ is the signal to noise ratio.

It is useful to consider this asymptotic variance with one of the models from Section 1.1. If we let

$$\theta(\omega,a,\mu) = \omega\langle\xi(a),\mu\rangle \quad ,$$

which corresponds to no dispersion, then

$$(\Sigma^{-1})_{jk} = \frac{1}{\pi} \int_B \frac{R(\omega)^2}{1+pR(\omega)} \omega^2 \sum_a \sum_b (\xi_j(a) - \xi_j(b))$$

$$(\xi_k(a) - \xi_k(b))\, d\omega \quad .$$

The similarities between this and [1.4] are obvious. An array with a narrow main beam will give estimates with a low asumptotic variance.

2.2 *MAXIMUM LIKELIHOOD ESTIMATING FUNCTIONS*

Here we derive an approximate likelihood function which shall be used in later sections. As several approximations are made estimators based on this should more properly be described as pseudo maximum likelihood estimators.

Consider the observations $y(1),y(2),\ldots,y(N)$ of a p dimensional vector valued Gaussian stationary process, with autocorrelation function $\Gamma(n)$ and spectral density matrix $G(\omega)$. Let

$$y_N = \mathrm{vec}(y(1),y(2),\ldots,y(N))$$

$$= \begin{pmatrix} y(1) \\ y(2) \\ \cdot \\ \cdot \\ \cdot \\ y(N) \end{pmatrix}$$

Then y_N is a Gaussian random variable of dimension Np. Its variance matrix Σ_N is block Toeplitz.

$$\Sigma_N = \begin{pmatrix} \Gamma(0) & \Gamma(1) & \Gamma(2) & \cdot & \cdot & \cdot & \Gamma(N-1) \\ \Gamma(1) & \Gamma(0) & \Gamma(1) & & & & \\ \Gamma(2) & \Gamma(1) & \Gamma(0) & & & & \\ \cdot & & & & & & \\ \cdot & & & & & & \\ \cdot & & & & & & \\ \Gamma(N-1) & & & & & & \Gamma(0) \end{pmatrix}$$

Suppose that $\Gamma(n)$ and hence Σ_N depends upon a parameter τ to be estimated. Then the log-likelihood function is

$$L_N(\tau) = -\tfrac{1}{2}\log(\det \Sigma_N(\tau)) - \tfrac{1}{2}N \log 2\pi - \tfrac{1}{2} y_N^* \Sigma_N^{-1}(\tau)\, y_N .$$

This is exact but almost useless in practice.

If P_N is a $N \times N$ unitary matrix with $N^{-\frac{1}{2}} e^{i\omega_j n}$, where $\omega_j = 2\pi j N^{-1}$, in the (j,n) position, and $\tilde{P}_N = P_N \otimes I_p$ then $\tilde{P}_N \Sigma_N \tilde{P}_N^*$ is approximately block diagonal with $\frac{1}{2\pi} G(\omega_j,\tau)$ as the j^{th} $p \times p$ block. Also $\tilde{P}_N y_N$ is a vector of Fourier transforms, that is $\sqrt{2\pi}\ \mathrm{vec}(Y(\omega_1), Y(\omega_2), \ldots, Y(\omega_N))$. Hence

$$\begin{aligned} L_N(\tau) = & -\tfrac{1}{2}\log \det(\tilde{P}_N \Sigma_N(\tau) \tilde{P}_N^*) - \tfrac{1}{2}\log 2\pi \\ & - \tfrac{1}{2} y_N^* \tilde{P}_N^* (\tilde{P}_N \Sigma_N(\tau)^- \tilde{P}_N^*) \tilde{P}_N y_N \\ \approx & \ \tfrac{1}{2} \sum_{j=1}^{N} \{ -\log(\det G(\omega_j,\tau)) - Y^*(\omega_j)\, G(\omega_j,\tau)^{-1} Y(\omega_j) \}. \end{aligned}$$

It is in this form that the likelihood is most useful. In one sense we are acting as if the Fourier transform evaluated at the fundamental frequencies is a set of uncorrelated Gaussian random variables with zero mean and variance $G(\omega_j)$. If certain frequencies are not used (as when we restrict the sum to the set B) we are merely ignoring some of these variables. The resulting function is still the approximate likelihood for the reduced set of data.

This likelihood function will be used even when the processes are non Gaussian. The proofs given for consistency and limiting distributions will not be assuming that the processes are Gaussian.

2.3 ONE SIGNAL WITH CORRELATED NOISE

We now set $\theta_0(\omega) = \theta(\omega,\tau_0)$ and we omit reference to ω and τ where this can be done without confusion.

The estimator of Section 2.1 essentially uses the fact that the expression

$$\theta^*(f_x I_p + f_z \theta_0 \theta_0^*)\theta = f_x p + f_z |\theta_0^* \theta|^2$$

is maximized when $\theta = \theta_0$. If a diagonal noise spectral density matrix is replaced by an arbitrary positive definite matrix F_x, then almost definitely the maximum will occur for $\theta \neq \theta_0$. In this section we shall consider the case of an arbitrary positive definite noise spectral density matrix F_x.

It is only possible to derive estimators for correlated noise if we have prior knowledge of the structure of the noise correlation. Otherwise a signal can be found in any positive definite noise matrix since if

$$A(\varepsilon) = F_x - \varepsilon \theta \theta^*$$

is not positive definite for any $\varepsilon > 0$ then F_x cannot be positive definite because the eigenvalues of $A(\varepsilon)$ are continuous in ε.

We first consider the model

$$F_y = F_x + f_z \theta \theta^* ,$$

where both F_x and f_z are assumed known. Then

$$F_y^{-1} = F_x^{-1} - (f_z/(1 + f_z \theta^* F_x^{-1} \theta)) F_x^{-1} \theta \theta^* F_x^{-1}$$

and $$\det F_y = (\det F_x)(1 + f_z \theta^* F_x^{-1} \theta) .$$

This gives the likelihood function

$$q_N(\tau) = N^{-1} \sum_{\omega \in B} \{ - \log(\det F_x) - \log(1 + f_z \theta^* F_x^{-1} \theta) - Y^* F_x^{-1} Y + f_z (1 + f_z \theta^* F_x^{-1} \theta)^{-1} |Y^* F_x^{-1} \theta|^2 \} \qquad [3]$$

The first and third terms do not depend on τ and would be ignored in practice. As before we estimate τ_0 by that value $\hat{\tau}_N$ which maximizes [3].

Before proving consistency, we require the following lemma

which is well known in the field of multivariate analysis. An alternative proof using calculus is given by Anderson (1958).

LEMMA 2.3

If A and C are positive definite $p \times p$ hermitian matrices, then the function

$$g(A) = \log(\det A) + \operatorname{trace}(A^{-1}C)$$

is minimized by A=C and no other value of A.

PROOF

$$\begin{aligned} g(A) &= \log(\det A) + \operatorname{trace}(A^{-\frac{1}{2}}A^{-\frac{1}{2}*}C) \\ &= \log(\det A) + \operatorname{trace}(A^{-\frac{1}{2}*}CA^{-\frac{1}{2}}) \\ &= \log(\det A) + \sum_{i=1}^{p} \lambda_i \end{aligned}$$

where λ_i is the i^{th} eigenvalue of $A^{-\frac{1}{2}*}CA^{-\frac{1}{2}}$. Then using the inequality for arithmetic and geometric means we have that

$$\begin{aligned} g(A) &\geq \log(\det A) + p\left(\prod_{i=1}^{p} \lambda_i\right)^{1/p} \\ &= \log(\det A) + p(\det(A^{-\frac{1}{2}*}CA^{-\frac{1}{2}}))^{1/p} \\ &= \log(\det A) + p(\det A)^{-1/p}(\det C)^{1/p} \\ &\geq \log(\det C) + p \end{aligned}$$

The last step comes from considering the minimum of the function $x \to \log x + p(\frac{c}{x})^{1/p}$. Clearly this bound is attained when A=C. If it is achieved for any other values of A, then the two inequalities above must be equalities. Hence

$$\frac{1}{p}\sum_{i=1}^{p} \lambda_i = \left(\prod_{i=1}^{p} \lambda_i\right)^{1/p},$$

and $\det A = \det C$.

The first of these implies that all the eigenvalues of $A^{-\frac{1}{2}*}CA^{-\frac{1}{2}}$ are equal. Since $A^{-\frac{1}{2}*}CA^{-\frac{1}{2}}$ is also hermitian, it must be a multiple of the identity and thus $C=\lambda A$. But the second equality

implies that $\lambda=1$. //

THEOREM 2.4

If Conditions 2.1 and 2.2 are satisfied then $\underset{N\to\infty}{\text{lim}} \hat{\tau}_N = \tau_0$ almost surely.

PROOF

As in Hannan and Robinson (1973) it can be shown that $q_N(\tau) \to q(\tau)$ almost surely, uniform in τ, where

$$q(\tau) = (2\pi)^{-1}\int_B - \log(\det F_y(\omega,\tau)) - \text{trace}(F_y(\omega,\tau)^{-1}F_y(\omega,\tau_0))d\omega$$

since $F_y(\omega,\tau_0)$ is the true spectrum. By Lemma 2.3 the integrand is maximized when $F_y(\omega,\tau) = F_y(\omega,\tau_0)$. Condition 2.2. ensures that this occurs only when $\tau=\tau_0$. Then for any $\varepsilon>0$ there exists N sufficiently large such that

$$q(\tau_0) - \varepsilon \le q_N(\tau_0) \le q_N(\hat{\tau}_N) \le q(\hat{\tau}_N) + \varepsilon \le q(\tau_0) + \varepsilon.$$

Since $\varepsilon>0$ is arbitrary, this holds only if $\hat{\tau}_N \to \tau_0$ almost surely. //

A limiting variance result may be obtained for this estimator, using methods to be developed in the next chapter. However, the asymptotic variance is not very meaningful so we do not present it here.

To use the above estimator both the noise and the signal spectra must be known. If the noise spectrum is known, then for a given value of τ, the signal spectrum f_z may be estimated by maximizing [3] with respect to f_z at each fundamental frequency. In fact

$$\frac{\partial q_N}{\partial f_z} = N^{-1}\{-\theta^* F_x^{-1}\theta(1+f_x\theta^* F_x^{-1}\theta)^{-1} + |Y^* F_x^{-1}|^2(1+f_z\theta^* F_x^{-1}\theta)^{-2}\}.$$

Equating this to zero we have an estimate $\hat{f}_z$:

$$\hat{f}_z = (|Y^* F_x^{-1}\theta|^2/(\theta^* F_x^{-1}\theta) - 1)/(\theta^* F_x^{-1}\theta) \quad . \qquad [4]$$

It would be inefficient to use [4] directly since the value of $\hat{f}_z$

it gives can even be negative. A better estimate of f_z in practice would be the average of [4] over a narrow band of fundamental frequencies, as is commonly done in spectral estimation.

To obtain a preliminary value of τ, we may substitute [4] into [3] directly to give a new estimating function:

$$q_N(\tau) = N^{-1} \sum_{\omega \epsilon B} \{-\log(\det F_x) - \log (|Y^* F_x^{-1}\theta|^2/\theta^* F_x^{-1}\theta) - Y^* F_x^{-1} Y + |Y^* F_x^{-1}\theta|^2/\theta^* F_x^{-1}\theta\} \quad . \qquad [5]$$

As before the first and third terms would not be used in practice since they are independent of τ. After using this estimating function to obtain a preliminary estimate of τ, the signal spectrum could be estimated by averaging [4] over bands and then the maximum likelihood estimate could be used. The last two steps could be iterated if necessary.

It can be seen that [5] can be expressed as

$$q_N(\tau) = N^{-1} \sum_{\omega \epsilon B} [R(\tau) - \log(R(\tau))] + \text{constants} \ ,$$

where $R(\tau) = |Y^* F_x^{-1}\theta|^2/\theta^* F_x^{-1}\theta$. When $R(\tau)$ is significantly larger than one, little would be lost by maximizing $N^{-1} \sum_{\omega \epsilon B} R(\tau)$. This leads to the possibility of treating the case where $F_x = f_x V_x$ and V_x is known but f_x is not. In some real situations theoretical models could be used to predict V_x(Eckart, 1953; Cron and Sherman, 1962). Hence we define a new estimating function

$$q_N(\tau) = N^{-1} \sum_{\omega \epsilon B} W(\omega) |Y^* V_x^{-1}\theta|^2/\theta^* V_x^{-1}\theta \qquad [6]$$

and we shall call estimates based on this function *adaptive estimates* due to the similarities to adaptive beamforming methods (Widdrow *et al.*, 1967).

THEOREM 2.5

If Conditions 2.1 and 2.2 are satisfied then the adaptive estimate based on [6] is consistent.

PROOF

As in Theorem 2.4 we have, uniformly in τ, $q_N(\tau) \to q(\tau)$ where

$$q(\tau) = (2\pi)^{-1} \int_B W(\omega)\theta^* V_x^{-1} F_y(\omega,\tau_0) V_x^{-1}\theta/\theta^* V_x^{-1}\theta d\omega \quad .$$

But since $F_y(\omega,\tau_0) = f_x V_x + f_z \theta_0\theta_0^*$ we have

$$q(\tau) = (2\pi)^{-1} \int_B W(\omega)\{f_x(\omega) + f_z|\theta_0^* V_x^{-1}\theta|^2/\theta^* V_x^{-1}\theta\}d\omega \quad .$$

However, by the Cauchy Schwartz inequality

$$|\theta_0^* V_x^{-1}\theta|^2/\theta^* V_x^{-1}\theta \leq \theta_0^* V_x^{-1}\theta_0$$

and this bound is attained if and only if $\theta=\lambda\theta_0$ where λ is a (complex) scalar. Condition 2.2 precludes this for any $\lambda \neq 1$, so that $q(\tau)$ is maximized at τ_0. The remainder of the proof follows as for Theorem 2.4. //

2.4 *SOME NON-QUADRATIC ESTIMATING FUNCTIONS*

The principle estimating functions used so far have all been linear functions of the periodogram and therefore quadratic functions of the data. Several non-quadratic functions have been proposed amongst which are Capon's "maximum likelihood" function (Capon, 1969), the entropy function (Burg, 1967) and the cophase function (Posmentier and Herman, 1971). Of these, only Capon's method is widely used in the estimation context. The maximum entropy method arose in spectral estimation problems and does not appear to have been applied to irregular arrays, despite suggestions that it could be (McDonough, 1974). The cophase method has greater applications

in the signal detection field.

We shall slightly change our notation for this section since these estimators are usually used when several signals are present and the true phase vectors are of the form $\theta_j = \Phi(\tau_j)$. For example, each signal may have a wavenumber or slowness vector associated with it and this parameter must be estimated for each signal. It is then hoped that the S largest maxima of the estimating function correspond to the S signals present.

Capon's estimator uses estimates $\hat{F}_y$ of the spectral density matrix at a single frequency, formed by averaging the periodogram over a narrow band of fundamental frequencies. The estimating function is then

$$q(\tau) = (\Phi(\tau)^* \hat{F}_y^{-1} \Phi(\tau))^{-1} \quad . \qquad [7]$$

Care must be taken that $\hat{F}_y$ is not ill conditioned or close to being singular. It must be formed by an average over at least p fundamental frequencies and significantly many more if a taper has been used or if the signals are transient. It can be seen that [7] is not related to the likelihood function and the name "maximum likelihood" which has become attached to Capon's estimator is a misnomer. Simulations in Chapter 4 will show that it can be significantly inferior to true maximum likelihood estimators.

If the conventional one signal estimator of Hannan is also to be based on an estimate of the spectral density matrix, its estimating function is

$$q(\tau) = \Phi(\tau)^* \hat{F}_y \Phi(\tau) \quad . \qquad [8]$$

Piserenko (1971) noted that both of these estimating functions could be fitted into a wider class as follows. Consider the eigenvalue-eigenvector decomposition of $\hat{F}_y$: that is

$$\hat{F}_y = \sum_{i=1}^{p} \lambda_i U_i U_i^*$$

where $\lambda_1, \lambda_2, \ldots, \lambda_p$ are the eigenvalues and $U_1, U_2, \ldots, U_p$ are the associated orthonormal eigenvectors. Then the estimating function is defined by

$$q_\varepsilon(\tau) = \begin{cases} \left(\sum_{i=1}^{p} \lambda_i^{\varepsilon} |\Phi(\tau)^* U_i|^2 \right)^{1/\varepsilon} & \text{for } \varepsilon \neq 0 \\ \exp\left(\sum_{i=1}^{p} (\log \lambda_i) |\Phi(\tau)^* U_i|^2 \right) & \text{for } \varepsilon = 0 \end{cases} .$$

With $\varepsilon = -1$ we obtain [7] and with $\varepsilon = 1$ we obtain [8].

To obtain asymptotic results, we shall consider this class of functions with $\hat{F}_y$ replaced by F_y.

THEOREM 2.6

If $F_y = f_x I_p + f_z \theta\theta^*$, then the estimating functions $q_\varepsilon(\tau)$ are monotonically increasing functions of $|\Phi(\tau)^* \theta|^2$ for all ε.

PROOF

If $F_y = f_x I_p + f_z \theta\theta^*$

then $\det(F_y - \lambda I_p) = (f_x - \lambda)^{p-1}(f_x + pf_z - \lambda)$.

Hence $\lambda_1 = f_x + pf_z$ and $\lambda_i = f_x$, $i = 2,3,\ldots p$.

Also $U_1 = \theta/\sqrt{p}$ and $U_2, U_3, \ldots, U_p$ cannot be uniquely specified but are orthogonal to θ. Let $\Phi = \alpha\,\theta + b$ where α is a complex scalar and b is a vector in the span of $U_2, U_3, \ldots, U_p$. Then

$$p = ||\Phi||^2 = |\alpha|^2 \, ||\theta||^2 + ||b||^2 ,$$

and $$||b||^2 = p(1 - |\alpha|^2) .$$

Also we have that

$$|\Phi^* U_1|^2 = |(\alpha\theta + b)^* \theta|^2 / p = p|\alpha|^2$$

and hence

$$\sum_{i=2}^{p} |\Phi^* U_i|^2 = ||b||^2 = p - |\Phi^* \theta|^2 / p .$$

Then for $\varepsilon \neq 0$

$$q_\varepsilon(\tau) = \{(f_x+pf_z)^\varepsilon |\Phi(\tau)^* \theta|^2/p + f_x^\varepsilon (p-|\Phi(\tau)^*\theta|^2)/p\}^{1/\varepsilon}$$

$$= \{pf_x^\varepsilon + |\Phi(\tau)^*\theta|^2 (f_x+pf_z)^\varepsilon - f_x^\varepsilon)/p\}^{1/\varepsilon}$$

and for $\varepsilon = 0$

$$q_0(\tau) = \exp\{p \log f_x + |\Phi(\tau)^*\theta|^2 (\log(f_x+pf_z) - \log f_x)/p\} .$$

In both cases $q_\varepsilon(\tau)$ is a monotonic function of $|\Phi(\tau)^*\theta|^2$.

Since $|\Phi(\tau)^*\theta|^2$ is maximized when $\Phi(\tau)=\theta$, all of these estimating functions should give asymptotically identical results. //

In the case of correlated noise, the eigenvalues of F_y are no longer simple expressions as above and a general result for this class has not been obtained. However, in the case $\varepsilon = -1$, that is Capon's estimator, the known form of F_y^{-1} can be used. Then

$$F_y^{-1} = F_x^{-1} - f_z(1+f_z\theta^* F_x^{-1}\theta)^{-1} F_x^{-1}\theta\theta^* F_x^{-1}$$

and

$$q_{-1}(\tau) = \{\Phi(\tau)^* F_x^{-1}\Phi(\tau) - f_z(1+f_z\theta^* F_x^{-1}\theta)^{-1} |\Phi(\tau)^* F_x^{-1}\theta|^2\}^{-1}$$

$$= (\Phi(\tau)^* F_x^{-1}\Phi(\tau)) \sum_{k=o}^{\infty} \left(\frac{f_z}{1+f_z\theta^* F_x^{-1}\theta} \; \frac{|\Phi(\tau)^* F_x^{-1}\theta|^2}{\Phi(\tau)^* F_x^{-1}\Phi(\tau)} \right)^k . \qquad [9]$$

The last term in this is identical to that used in the adaptive estimator. However, the first factor $(\Phi(\tau)^* F_x^{-1}\Phi(\tau))^{-1}$ will usually bias the estimator. If $F_x = f_x I_p$, then [9] reduces to

$$q_{-1}(\tau) = (pf_x)^{-1} \sum_{k=o}^{\infty} \{ \left(\frac{pf_z f_x^{-1}}{1+pf_z f_x^{-1}} \right) |\Phi(\tau)^*\theta|^2/p^2 \}^k .$$

This is a positive average of powers of $|\Phi(\tau)\theta|^2$. When the signal to noise ratio, f_z/f_x is large the higher powers would dominate. This explains the very sharp peaks observed with Capon's estimating function.

CHAPTER 3

MULTIPLE SIGNAL ESTIMATORS

The estimators of the previous chapter will fail to be consistent when faced with more than one signal. In fact, relative to one signal, the other signals are equivalent to correlated noise of unknown structure. In this chapter, consistent estimates for more than one signal and uncorrelated noise are developed.

3.1 THE SPECTRAL DENSITY MATRIX

The spectral density matrix corresponding to S signals and uncorrelated noise with identical spectra at each receiver is

$$F_y = F_x + \theta F \theta^* \quad . \qquad [1]$$

As before we shall only refer to ω and τ when necessary to do so to avoid confusion.

If $p \geq S$ and θ has rank S then F_y^{-1} has the following forms:

$$F_y^{-1} = f_x^{-1}(I_p - \theta H \theta^*) \qquad [2]$$

where $H^{-1} = f_x F^{-1} + \theta^* \theta$

and secondly

$$F_y^{-1} = f_x^{-1}(I_p - \theta(\theta^*\theta)^{-1}\theta^*) + \theta(\theta^*\theta)^{-1}(f_x(\theta^*\theta)^{-1} + F)^{-1}(\theta^*\theta)^{-1}\theta^* \qquad [3]$$

These expressions can be verified by multiplying by F_y.

LEMMA 3.1

If Θ is of rank S then

$$\det F_y = f_x^{p-S} \det(\Theta^*\Theta) \det(f_x(\Theta^*\Theta)^{-1} + F)$$

$$= f_x^{p-S} \det F \det(H^{-1}) .$$

PROOF

Let $\Theta = P^*\tilde{\Lambda}Q$ where P is a p×p unitary matrix, Q is an S×S unitary matrix and $\tilde{\Lambda} = \begin{pmatrix}\Lambda\\0\end{pmatrix}$ where Λ is a nonsingular S×S diagonal matrix. Also let $\tilde{F} = QFQ^*$.

Then

$$\det F_y = \det(f_x P^*P + P^*\tilde{\Lambda}QFQ^*\tilde{\Lambda}P)$$

$$= \det\{P^*(f_x I_p + \tilde{\Lambda}F\tilde{\Lambda}^*)P\}$$

$$= \det(f_x I_p + \tilde{\Lambda}F\tilde{\Lambda}^*) .$$

This matrix is zero except for the diagonal and the upper left S×S block. Hence

$$\det F_y = f_x^{p-S} \det(f_x I_S + \Lambda\tilde{F}\Lambda^*)$$

$$= f_x^{p-S} \det\{\Lambda(f_x\Lambda^{-1}\Lambda^{-1*} + \tilde{F})\Lambda^*\}$$

$$= f_x^{p-S} \det(\Lambda^*\Lambda) \det(f_x(\Lambda^*\Lambda)^{-1} + \tilde{F})$$

$$= f_x^{p-S} \det(Q^*\tilde{\Lambda}^*PP^*\tilde{\Lambda}Q) \det(f_x Q(Q^*\tilde{\Lambda}^*PP^*\tilde{\Lambda}Q)^{-1}Q^* + Q\tilde{F}Q^*$$

$$= f_x^{p-S} \det(\Theta^*\Theta) \det(f_x(\Theta^*\Theta)^{-1} + F),$$

which is the first of the forms required. From this we have

$$\det F_y = f_x^{p-S} \det \Theta^*\Theta(f_x(\Theta^*\Theta)^{-1} + F)$$

$$= f_x^{p-S} \det(f_x I_S + \Theta^*\Theta F)$$

$$= f_x^{p-S} \det (f_x F^{-1} + \theta^* \theta) F$$

$$= f_x^{p-S} \det F \det H^{-1} \quad . \quad //$$

It is important to consider the extent to which the process is identified by the spectrum; that is, the extent to which the representation [1] is unique. This will obviously depend on the model which is chosen for $\theta(\omega,\tau)$ but the problem is difficult to solve in general.

If the spectra f_x and F are fixed, we need only consider the term $\theta F \theta^*$. The eigenvectors of this expression are in the space spanned by the columns of θ and this leads to the following condition.

CONDITION 3.2

There exists no allowable $\tau \neq \tau_0$ such that for all $\omega \varepsilon B$ the columns of $\theta(\omega,\tau)$ are all contained in the span of the columns of $\theta(\omega,\tau_0)$; that is, $\theta(\omega,\tau) = \theta(\omega,\tau_0)C(\omega)$ for some matrix $C(\omega)$ and all $\omega \varepsilon B$ implies $\tau=\tau_0$. //

This is a sufficient condition for identification but it seems unnecessarily strong if F is known. Note that it implies that $S<p$.

A simple case to consider is one where there is no dispersion and hence

$$\theta_u(\omega, a, \tau) = \omega \langle \mu_u, \xi(a) \rangle \quad .$$

(In fact, by suitable non linear stretching of the frequency scale many dispersive models can be reduced to this form.) Then the (a,b) element of $\theta F \theta^*$ is

$$\sum_{u=1}^{S} f_u(\omega)\, e^{i\omega \langle \mu_u, \xi(a)-\xi(b) \rangle} \quad .$$

If the spectra are constant then this is an almost periodic function and hence the representation will be unique. Also for S=2

it can readily be shown to be unique if the spectra are known and not identical. However there does not seem to be any general theory to cover this.

If the spectra are not kept fixed (that is, are not known, then it appears that Condition 3.2 cannot be improved upon. Let us consider an example where p=s=2. First suppose that

$$f_x(\omega) = 3 - 2\cos \omega/2,$$

$$f(\omega) = f_2(\omega) = \cos \omega/2$$

and $$\theta(\omega) = \begin{pmatrix} 1 & 1 \\ e^{i2\omega/4} & e^{-i2\omega/4} \end{pmatrix} .$$

Then $$F_y(\omega) = \begin{pmatrix} 3 & 2\cos \omega/2 \operatorname{cose} 2\ \omega/4 \\ 2\cos \omega/2 \cos 2\ \omega/4 & 3 \end{pmatrix} .$$

Next consider

$$f_x(\omega) = 3 - 2\cos 2\ \omega/4$$

$$f_1(\omega) = f_2(\omega) = \cos 2\ \omega/4$$

and $$\theta(\omega) = \begin{pmatrix} 1 & 1 \\ e^{i\omega/2} & e^{-i\omega/2} \end{pmatrix} .$$

This gives exactly the same spectrum F_y as the previous case, at all frequencies. In both cases the phase matrices are physically reasonable.

In what follows, we shall assume that "the process is identifiable", meaning that the spectrum can uniquely identify τ_0. This implies that each signal obeys Condition 2.2. In practice, even when $p \leq S$ and the spectra are specified, identification does not seem to be a real problem.

3.2 *THE MAXIMUM LIKELIHOOD ESTIMATOR*

We now derive a maximum likelihood estimator, on the assumption that both the signal and the noise spectra are known. Using the results of Section 2.2 we define the estimating function

$$q_N(\tau) = N^{-1} \sum_{\omega \in B} -\log(\det F_y) - Y^* F_y^{-1} Y \quad .$$

By Lemma 3.1 and [2], we have

$$q_N(\tau) = N^{-1} \sum_{\omega \in B} \{-\log(f_x^{p-S} \det F) - \log(\det H^{-1}) - f_x^{-1} Y^* Y + f_x^{-1} Y^* \theta H \theta^* Y\} \quad . \qquad [4]$$

In practice the first and third terms would be ignored since they are independent of τ. For the case S=2, if we ignore these terms, [4] reduces to

$$q_N(\tau) = N^{-1} \sum_{\omega \in B} \{-\log\Delta + (f_x \Delta)^{-1} [f_1(f_x + pf_2)|Y^* \theta_1|^2 + f_2(f_x + pf_1)|Y^* \theta_2|^2 - f_1 f_2(\psi \theta_2^* Y Y^* \theta_1 + \bar{\psi} \theta_1^* Y Y^* \theta_2)] \qquad [5]$$

where $\Delta = f_x^2 + pf_x f_1 + pf_x f_2 + f_1 f_2(p^2 - |\psi|^2)$

and $\psi = \theta_1^* \theta_2$.

The terms $|Y^* \theta_1|^2$ and $|Y^* \theta_2|^2$ can be thought of as corresponding to beamforming in two directions, so that the last term in [5] takes into account interaction between the two beams. $|\psi|^2$ is a measure of this interaction, being significant when the two beams are close to each other or when one beam is on a sidelobe of the other.

As before τ_0 is estimated by that value τ_N which maximizes $q_N(\tau)$.

THEOREM 3.3

If Condition 2.1 is satisfied and the process is identified then the maximum likelihood estimator is consistent.

PROOF

The proof follows exactly as the proof of Theorem 2.4. //

Before considering the asymptotic variance, the following lemmas are required.

LEMMA 3.4

If $A(\tau)$ is a non signular matrix function of the vector τ, $A^{(j)}(\tau) = \frac{\partial}{\partial \tau_j} A(\tau)$ and

$$g(\tau) = \log(\det A(\tau)) + \operatorname{trace}(A(\tau)^{-1}A(\tau_0))$$

then $$\frac{\partial^2}{\partial \tau_j \partial \tau_k} g(\tau)\Big|_{\tau_0} = \frac{\partial^2}{\partial \tau_j \partial \tau_k} g(\tau_0)$$

$$= \operatorname{trace}(A(\tau_0)^{-1}A(\tau_0)^{(j)}A(\tau_0)^{-1}A(\tau_0)^{(k)}) .$$

PROOF

We have

$$\frac{\partial}{\partial \tau_j} g(\tau) = \operatorname{trace}(A(\tau)^{(j)}A(\tau)^{-1}) - \operatorname{trace}(A(\tau)^{-1}A(\tau)^{(j)}A(\tau)^{-1}A(\tau_0))$$

and

$$\begin{aligned}\frac{\partial^2}{\partial \tau_j \partial \tau_k} g(\tau) = {} & \operatorname{trace}(A(\tau)^{(j,k)}A(\tau)^{-1} - A(\tau)^{(j)}A(\tau)^{-1}A(\tau)^{(k)}A(\tau)^{-1}) \\ & - \operatorname{trace}(-A(\tau)^{-1}A(\tau)^{(k)}A(\tau)^{-1}A(\tau)^{(j)}A(\tau)^{-1}A(\tau_0) \\ & + A(\tau)^{-1}A(\tau)^{(j,k)}A(\tau)^{-1}A(\tau_0) \\ & - A(\tau)^{-1}A(\tau)^{(j)}A(\tau)^{-1}A(\tau)^{(k)}A(\tau)^{-1}A(\tau_0)).\end{aligned}$$

Hence

$$\frac{\partial^2}{\partial \tau_j \partial \tau_k} g(\tau_0) = \operatorname{trace}(A(\tau_0)^{-1}A(\tau_0)^{(j)}A(\tau_0)^{-1}A(\tau_0)^{(k)}) .$$

The above steps use the identities

$$\frac{\partial}{\partial \tau_j} \log(\det A(\tau)) = \operatorname{trace}(A(\tau)^{(j)} A(\tau)^{-1})$$

and $$\frac{\partial}{\partial \tau_j} (A(\tau)^{-1}) = A(\tau)^{-1} A(\tau)^{(j)} A(\tau)^{-1} \quad . \quad //$$

CONDITION 3.5

A stationary process shall be said to satisfy Condition 3.5 if its fourth cumulant function is the Fourier transform of a function $\hat{\beta}(\omega_1,\omega_2,\omega_3,\omega_4)$ which is continuous in some neighbourhood of the hyperplane $\omega_1+\omega_2 = \omega_3+\omega_4 = 0$. //

LEMMA 3.6

Let $\Psi(\omega)$ and $\chi(\omega)$ be p×p hermitian matrix valued functions and let $Y(\omega)$ be the Fourier transform of a p dimensional vector process which satisfied Condition 3.5. Let

$$u = N^{-1} \sum_{\omega \in B} Y(\omega)^* \Psi(\omega) Y(\omega)$$

and $$v = N^{-1} \sum_{\omega \in B} Y(\omega)^* \chi(\omega) Y(\omega) \quad .$$

Then N Cov(u,v) converges to

$$(2\pi)^{-1} \int_B \operatorname{trace}(\Psi(\omega) F_y(\omega) \chi(\omega) F_y(\omega))\, d\omega$$

$$+ (2\pi)^{-1} \int_B \int_B \sum_{abcd} \Psi_{ab}(\omega) \chi_{dc}(\omega') \hat{\beta}_{abcd}(\omega,-\omega,\omega',-\omega')\, d\omega d\omega' .$$

PROOF

This follows directly from Theorem 3, Hannan (1973). //

THEOREM 3.7

If the process y(n) satisfies Condition 3.5 and the noise

is independent between receivers then the variance matrix of $N^{\frac{1}{2}}(\hat{\tau}_N - \tau_0)$ tends to Σ where

$$\Sigma = A^{-1} + A^{-1} C A^{-1} ,$$

and $$A_{jk} = (2\pi)^{-1}\int_B \text{trace}(F_y(\omega,\tau_o)^{-1}F_y(\omega,\tau_0)^{(j)}F_y(\omega,\tau_0)^{-1} F_y(\omega,\tau_0)^{(k)})d\omega ,$$

$$C_{jk} = (2\pi)^{-1}\int_B\int_B \sum_{abcd}(F_y(\omega,\tau_0)^{-1}F_y(\omega,\tau_0)^{(j)}F_y(\omega,\tau_0)^{-1})_{ab} (F_y(\omega',\tau_0)^{-1}F_y(\omega',\tau_0)^{(k)}F_y(\omega',\tau_0)^{-1})_{dc} \hat{\beta}_{abcd}(\omega,-\omega,\omega',-\omega')d\omega d\omega' .$$

PROOF

Consider the Taylor series expansion of $\frac{\partial}{\partial\tau_j} q_N(\tau_0)$ about $\hat{\tau}_N$:

$$\frac{\partial}{\partial\tau_j} q_N(\tau_0) - \frac{\partial}{\partial\tau_j} q_N(\hat{\tau}_N) = \sum_k (\tau_{0_k} - \hat{\tau}_{N_k}) \frac{\partial^2}{\partial\tau_j\partial\tau_k} q_N(\bar{\tau}_N)$$

where $||\bar{\tau}-\tau_0|| \leq ||\hat{\tau}_N-\tau_0||$.

Since $\frac{\partial}{\partial\tau_j} q_N(\hat{\tau}_N) = 0$ then

$$N^{\frac{1}{2}} \frac{\partial}{\partial\tau_j} q_N(\tau_0) = N^{\frac{1}{2}} \sum_k (\hat{\tau}_{N_k} - \tau_{0_k})(-\frac{\partial^2}{\partial\tau_j\partial\tau_k} q_N(\bar{\tau}_N) . \qquad [6]$$

It is then sufficient to consider the variance of the left hand side of [6]. By Lemma 3.6, the asymptotic variance matrix of $(N^{\frac{1}{2}} \frac{\partial}{\partial\tau_j} q_N(\tau_0))$ is A+C. $\frac{\tau^2}{\partial\tau_j\partial\tau_k} q_N(\bar{\tau})$ converges almost surely to $\frac{\partial^2}{\partial\tau_j\partial\tau_k} q(\tau_0)$ since $\bar{\tau}_N \to \tau_0$ almost surely. Hence $\frac{\partial^2}{\partial\tau_j\partial\tau_k} q_N(\bar{\tau}_N) \to A_{jk}$ by Lemma 3.4, and then the asymptotic variance of $\hat{\tau}_N$ is $A^{-1}(A+C)A^{-1}$ as required. //

The asymptotic variance is given in this form since it is easily computable. In practice the fourth cumulant term would usually be ignored since the cumulant spectra would almost definitely be unknown. The estimation of high order cumulant spectra was studied by Brillinger and Rosenblatt (1966) but is usually very difficult. Much geophysical data has roughly Gaussian distributions so the cumulant spectra can be assumed very small.

If p is large and S is reasonably small, a form for A not involving p×p matrices would be convenient. First, we note that

$$I_S = H H^{-1} = H(f_x F^{-1} + \theta^* \theta) = H(f_x I + \theta^* \theta F)F^{-1}$$

and hence

$$H\theta^* \theta F = F - f_x H = F\theta^* \theta H \quad . \qquad [7]$$

If we let $\theta^{(j)} = \frac{\partial}{\partial \tau_j} \theta$, then

$$F_y^{-1} F_y^{(j)} = f_x^{-1}(I_p - \theta H \theta^*)(\theta^{(j)} F \theta^* + \theta F \theta^{(j)*})$$

$$= f_x^{-1}(\theta^{(j)} F \theta^* + F\theta^{(j)} - \theta H \theta^* \theta^{(j)} F \theta^* - \theta H \theta^* \theta F \theta^{(j)*}) \quad .$$

Using [7] to simplify the last term, we have

$$F_y^{-1} F_y^{(j)} = f_x^{-1}(\theta^{(j)} F \theta^* - \theta H \theta^* \theta^{(j)} F \theta^* + f_x \theta H \theta^{(j)}) \quad .$$

Then trace $(F_y^{-1} F_y^{(j)} F_y^{-1} F_y^{(j)})$

$$= f_x^{-2} \operatorname{trace}(\theta^{(j)} F \theta^* \theta^{(k)} F \theta^* - \theta^{(j)} F \theta^* \theta H \theta^* \theta^{(k)} F \theta^*$$

$$+ f_x \theta^{(j)} F \theta^* \theta H \theta^{(k)} - \theta H \theta^* \theta^{(j)} F \theta^* \theta^{(k)} F \theta^*$$

$$+ \theta H \theta^* \theta^{(j)} F \theta^* \theta H \theta^* \theta^{(k)} F \theta^* - f_x \theta H \theta^* \theta^{(j)} F \theta^* \theta H \theta^{(k)*}$$

$$+ f_x \theta H \theta^{(j)*} \theta^{(k)} F \theta^* - f_x \theta H \theta^{(j)*} \theta H \theta^* \theta^{(k)} F \theta^*$$

$$+ f_x^2 \theta H \theta^{(j)*} \theta H \theta^{(k)*}) \quad .$$

Using [7] repeatedly and cycling the order of the products, this becomes:

$$
\begin{aligned}
& f_x^{-2}\,\mathrm{trace}(\theta^*\theta^{(j)}F\theta^*\theta^{(k)}F - \theta^*\theta^{(j)}F\theta^*\theta^{(k)}F + f_x\theta^*\theta^{(j)}H\theta^*\theta^{(k)} \\
& + f_x\theta^{(k)*}\theta^{(j)}F - f_x^2\theta^{(k)*}\theta^{(j)}H - F\theta^*\theta^{(j)}F\theta^*\theta^{(k)} \\
& + f_x H\theta^*\theta^{(j)}F\theta^*\theta^{(k)} + F\theta^*\theta^{(j)}F\theta^*\theta^{(k)} - f_x F\theta^*\theta^{(j)}H\theta^*\theta^{(k)} \\
& - f_x H\theta^*\theta^{(j)}F\theta^*\theta^{(k)} + f_x^2 H\theta^*\theta^{(j)}H\theta^*\theta^{(k)} - f_x\theta^{(k)*}\theta H\theta^*\theta^{(j)}F \\
& + f_x^2\theta^{(k)*}\theta H\theta^*\theta^{(j)}H + f_x F\theta^{(j)*}\theta^{(k)} - f_x^2 H\theta^{(j)*}\theta^{(k)} \\
& - f_x F\theta^{(j)*}\theta H\theta^*\theta^{(k)} + f_x^2 H\theta^{(j)*}\theta H\theta^*\theta^{(k)} + f_x^2 H\theta^{(j)*}\theta H\theta^{(k)*}\theta) \\
= \; & \mathrm{trace}\{(\theta^{(j)*}\theta+\theta^*\theta^{(j)})H(\theta^{(k)*}\theta+\theta^*\theta^{(k)})H - (\theta^{(j)*}\theta^{(k)}+\theta^{(k)*}\theta^{(j)})H\} \\
& + f_x^{-1}\mathrm{trace}\{(\theta^{(j)*}\theta^{(k)}+\theta^{(k)*}\theta^{(j)})F - \theta^*\theta^{(j)}F\theta^{(k)*}\theta H \\
& - \theta^*\theta^{(k)}F\theta^{(j)*}\theta H\} \; .
\end{aligned}
$$

Hence

$$
\begin{aligned}
A_{jk} = \; & \frac{1}{2\pi}\int_B \mathrm{trace}\{(\theta^{(j)*}\theta+\theta^*\theta^{(j)})H(\theta^{(k)*}\theta+\theta^*\theta^{(k)})H \\
& - (\theta^{(j)*}\theta^{(k)}+\theta^{(k)*}\theta^{(j)})H\} + f_x^{-1}\,\mathrm{trace}\{(\theta^{(j)*}\theta^{(k)}+\theta^{(k)*}\theta^{(j)})F \\
& - \theta^*\theta^{(j)}F\theta^{(k)*}\theta H - \theta^*\theta^{(k)}F\theta^{(j)*}\theta H\}d\omega,
\end{aligned}
$$

where this is evaluated at τ_0.

Experience with approximate maximum likelihood estimates leads us to expect that this estimator would have minimal variance. However its requirement that the spectra be known severely limits its usefulness. It cannot be readily simplified to remove this requirement.

3.3 THE EUCLIDEAN ESTIMATOR

The maximum likelihood estimator could be considered to minimize at each frequency a measure of distance between the observed

periodogram and the modelled spectral density matrix: this measure for the distance between matrices C and D is $\log(\det C) + \operatorname{trace}(C^{-1}D)$ and is minimized when C = D(Lemma 2.3). A different measure of distance could give a different class of estimators.

One possibility is to use the squared Euclidean distance between $F_y(\omega,\tau)$ and $I(\omega) = Y(\omega)Y^*(\omega)$. This gives the estimating function

$$q_N(\tau) = -N^{-1} \sum_{\omega\in B} W(\omega) \sum_a \sum_b |F_y(\omega,\tau)_{ab} - I_{ab}(\omega)|^2$$

$$= -N^{-1} \sum_{\omega\in B} W(\omega) \operatorname{trace}(F_y(\omega,\tau) - I(\omega))^2 \quad . \qquad [8]$$

As usual our estimate $\hat{\tau}_N$ maximizes [8]. We shall call this the *Euclidean estimator*, to emphasize that it is not a least squares estimator in the usual sense.

Note that

$$\operatorname{trace}(F_y(\omega,\tau)-I(\omega))^2 = \operatorname{trace} F_y(\omega,\tau)^2 - 2 \operatorname{trace}(F_y(\omega,\tau)I(\omega))^2 + \operatorname{trace} I(\omega)^2 \ .$$

If we consider the individual terms of this,

$$\operatorname{trace} F_y^2 = \operatorname{trace}(f_x f_p + \theta F \theta^*)^2$$

$$= f_x^2 \operatorname{trace} I_p + 2 f_x \operatorname{trace}(\theta^*\theta F) + \operatorname{trace}(\theta^*\theta F)^2$$

$$= pf_x^2 + 2f_x \sum_{u=1}^{S} f_u + 2 \sum_{u>v} \sum f_u f_v |\theta_u^* \theta_v|^2 + p \sum_{u=1}^{S} f_u^2$$

of which only the third terms depends on τ, and:

$$\operatorname{trace}(F_y I(\omega)) = Y^*(f_x I_p + \theta F \theta^*)Y$$

$$= f_x ||Y||^2 + Y^*\theta F \theta^* Y$$

$$= f_x ||Y||^2 + \sum_u f_u |Y^*\theta_u|^2$$

of which only the second term depends on τ. Hence if we neglect terms which are independent of τ, we obtain a simpler estimating function:

$$q_N(\tau) = N^{-1} \sum_{\omega\in B} W(\omega)\{ \sum_{u=1}^{S} f_u |Y^* \theta_u|^2 - \sum_{u>v} \sum f_u f_v |\theta_u^* \theta_v|^2\} . \quad [9]$$

In the case S=2 this is particularly simple:

$$q_N(\tau) = N^{-1} \sum_{\omega\in B} W(\omega)\{f_1 |Y^* \theta_1|^2 + f_2 |Y^* \theta_2|^2 - f_1 f_2 |\psi|^2\} ,$$

where ψ is as in the previous section.

THEOREM 3.8

If Condition 2.1 is satisfied and the process is identified, then the Euclidean estimator is consistent.

PROOF

Note that [9] can be expressed as

$$q_N(\tau) = N^{-1} \sum_{\omega\in B} W(\omega)\{\text{trace}(\theta F \theta^* I(\omega)) - \sum_{u>v} \sum f_u f_v |\theta_u^* \theta_v|^2\} .$$

As before we have $q_N(\tau) \to q(\tau)$ almost surely where

$$q(\tau) = (2\pi)^{-1} \int_B W(\omega)\{\text{trace}(\theta F \theta^* F_y) - \sum_{u>v} \sum f_u f_v |\theta_u^* \theta_v|^2\} \, d\omega$$

$$= -(2\pi)^{-1} \int_B W(\omega)\{\tfrac{1}{2} \text{trace}(F_y(\omega,\tau) - F_y(\omega,\tau_0))^2$$

$$-\tfrac{1}{2} \text{trace } F_y(\omega,\tau_0)^2 \quad -\tfrac{1}{2} p \sum_u f_u^2\} \, d\omega \quad .$$

Only the term $\text{trace}(F_y(\omega,\tau) - F_y(\omega,\tau_0))^2$ depends on τ and is obviously minimized if and only if $\tau=\tau_0$. //

THEOREM 3.9

If the process y(n) satisfies Condition 3.5 and the noise is

independent between receivers, then the variance matrix of $N^{\frac{1}{2}}(\hat{\tau}_N - \tau_0)$ tends to Σ where

$$\Sigma = A^{-1}(C+D)A^{-1}$$

$$\text{and } A_{jk} = (2\pi)^{-1}\int_B W(\omega)\ \text{trace}(\theta^{(j)}F\theta^{*}+\theta F\theta^{(j)*})(\theta^{(k)}F\theta^{*}+\theta F\theta^{(k)*})d\omega$$

$$C_{jk} = (2\pi)^{-1}\int_B W(\omega)^2\ \text{trace}\{(\theta^{(j)}F\theta^{*}+\theta F\theta^{(j)*})(f_x I_p+\theta F\theta^{*})$$
$$(\theta^{(k)}F\theta^{*}+\theta F\theta^{(k)*})(f_x I_p+\theta F\theta^{*})\}d\omega$$

$$\text{and } D_{jk} = (2\pi)^{-1}\int_B\int_B W(\omega)W(\omega')\sum_{u=1}^{S}\{\theta_u(\omega)^{*}(\theta(\omega)^{(j)}F(\omega)\theta(\omega)^{*}$$
$$+ \theta(\omega)F(\omega)\theta^{(j)}(\omega)^{*})\theta_u(\omega)\}$$
$$\{\theta_u(\omega')^{*}(\theta(\omega')^{(k)}F(\omega')\theta(\omega')^{*} + \theta(\omega')F(\omega')\theta(\omega')^{(k)*})\theta_u(\omega')\}$$
$$\hat{\beta}_u(\omega,-\omega,\omega',-\omega')d\omega d\omega'$$

where $\hat{\beta}_u$ is the fourth cumulant density of $z_u(n)$ and all expressions are evaluated at τ_0.

PROOF

This follows the proof of Theorem 3.7. That A_{jk} is $\frac{\partial^2}{\partial\tau_j\partial\tau_k}q(\tau_0)$ is most easily shown by considering [8].

The expressions for C_{jk} and D_{jk} are derived from Lemma 3.6. //

This variance would be almost definitely greater than that of the maximum likelihood estimator. An indication of the inefficiency of the Euclidean estimator is that its variance does not vanish as $f_x(\omega)\to 0$ on B.

The only advantages it has compared with the maximum likelihood estimator is that it is computationally easier and it does not assume prior knowledge of the noise spectrum.

3.4 *THE PHASE ESTIMATOR*

Various attempts have been made to use only the phase information in the periodogram. At first sight this may suggest a method of overcoming the problem of unknown spectra. One such *ad hoc* estimator will now be considered.

If we let $\langle z \rangle = \frac{z}{|z|}$ for all complex numbers $z \neq 0$, we may form an estimating function

$$q_N(\tau) = N^{-1} \sum_{\omega \in B} W(\omega) \sum \sum_{a \neq b} I_{ab}(\omega) \overline{\langle F_y(\omega,\tau)_{ab} \rangle} \quad . \qquad [10]$$

Let $\hat{\tau}_N$ be, as usual, the value which maximizes [10]. We shall call this the *phase estimator*.

THEOREM 3.10

If Condition 2.1 is satisfied and the process is identified, then the phase estimator is consistent.

PROOF

As before, $q_N(\tau) \to q(\tau)$ almost surely where

$$q(\tau) = (2\pi)^{-1} \int_B W(\omega) \sum \sum_{a \neq b} F_y(\omega,\tau_0)_{ab} \overline{\langle F_y(\omega,\tau)_{ab} \rangle} \, d\omega$$

$$= \pi^{-1} \int_B W(\omega) \sum \sum_{a<b} |F_y(\omega,\tau_0)_{ab}| \text{Real} \langle F_y(\omega,\tau_0)_{ab} \overline{F_y(\omega,\tau)_{ab}} \rangle d\omega \quad .$$

This achieves its maximum only when $F_y(\omega,\tau_0) = F_y(\omega,\tau)$.

The remainder of the proof follows as before. //

Even though consistent, this estimator can only partially solve the problem. While independent of the absolute level of the signal spectra, it requires prior knowledge of the ratios of the spectra.

Few situations can be imagined where these would be known without the actual spectra being known. An asymptotic variance could be computed as in the previous section, but due to the complexity of the derivatives of the estimating function, this would be very complex.

3.5 THE DETERMINISTIC ESTIMATOR

The previous estimators were derived assuming that the signals were stationary stochastic processes. They all required knowledge of the spectra and this would not usually be available in practice. If, for the purpose of derivation only, the signals are assumed to be deterministic and hence without spectra, this problem could be avoided. We shall do this, even though later on we shall discuss the behaviour of the estimate thus derived in terms of stochastic signals. Consider

$$y(n,a) = x(n,a) + \sum_{u=1}^{S} z_u(n,a)$$

where

$$z_u(n,a) = N^{-\frac{1}{2}} \sum_{\omega_r} Z_{u,a}(\omega_r)\, e^{-i\omega_r n}$$

$$Z_{u,a}(\omega_r) = Z_u(\omega_r)\, e^{i\theta_u(\omega_r,a,\tau)} \qquad [11]$$

and

$$\omega_r = 2\pi r/N \quad , \quad r = -[N/2], \ldots, [N/2] .$$

It is not assumed that [11] holds at any other frequencies.

Let $Z(\omega_r) = (Z_1(\omega_r), Z_2(\omega_r), \ldots, Z_S(\omega_r))'$.

Treating $Z(\omega_r)$ as an unknown but non random quantity, the maximum

likelihood approach coincides with the weighted least squares approach, since the "errors" are roughly uncorrelated in the frequency domain. Hence we minimize

$$T = N^{-1} \sum_{\omega \epsilon B} f_x^{-1} \sum_a \Big| Y_a - \sum_{u=1}^{S} Z_u(\omega_r)\, e^{i\theta_u(\omega,a,\tau)} \Big|^2 .$$

$$= N^{-1} \sum_{\omega \epsilon B} f_x^{-1} \, ||Y - \theta Z||^2 .$$

For each τ, we can replace $Z(\omega_r)$ by that vector which minimizes T. Differentiating,

$$\frac{\partial}{\partial Z_u(\omega_r)} T = 2N^{-1} \, \text{Real}\{f_x^{-1}(\theta_u^*(\omega_r)Y(\omega_r) - \sum_{v=1}^{S} Z_v(\omega_r)\theta_u^*(\omega_r)\theta_v(\omega_r))\} .$$

Equating the derivatives to zero, we obtain

$$\theta^*\theta Z = \theta^* Y ,$$

$$\text{and hence} \quad Z = (\theta^*\theta)^{-1}\theta^* Y . \qquad [12]$$

Then

$$T = N^{-1} \sum_{\omega \epsilon B} f_x^{-1} Y^* (I_p - \theta(\theta^*\theta)\theta^*) Y .$$

Thus we shall use as our estimating function

$$q_N(\tau) = N^{-1} \sum_{\omega \epsilon B} W(\omega)\, Y^*\theta(\theta^*\theta)^{-1}\theta^* Y \qquad [13]$$

and the estimate $\hat{\tau}_N$ obtained by maximizing this shall be called the *deterministic estimator*. An arbitrary weight function $W(\omega)$ is used since if there is incomplete prior knowledge of the spectra, the weight f_x^{-1} might not be optimal. We now proceed to discuss the properties of this estimator when used with stochastic signals.

THEOREM 3.11

If Conditions 2.1 and 3.2 are satisfied, then the deterministic

estimator is consistent.

PROOF

Let $E = \theta(\theta^*\theta)^{-1}\theta^*$. Then E is an ~~eigen~~idem potent matrix (that is $E^2=E$) and furthermore

$$\begin{aligned} \text{trace } (E) &= \text{trace}[(\theta^*\theta)^{-1}\theta^*\theta] \\ &= \text{trace } I_S \\ &= S \quad . \end{aligned}$$

Also, it is clear that any eigenvector of E, associated with a non zero eigen~~vector~~value must be in the span of the columns of θ and hence of the form $\theta\phi$ where ϕ is an S dimensional vector. But

$$E\theta\phi = \theta(\theta^*\theta)^{-1}\theta^*\theta\phi = \theta\phi \ .$$

Therefore there are S eigenvalues of E equal to one, with associated eigenvectors in the span of $\theta_1,\theta_2, \ldots ,\theta_S$, and the other p-S eigenvalues are zero.

As in previous consistency theorems, $q_N(\tau)\to q(\tau)$ almost surely where

$$q(\tau) = (2\pi)^{-1}\int_B W(\omega)\ \text{trace}\{E(f_x I_p + \theta_0 F \theta_0)\}\ d\omega \ . \qquad [14]$$

As before, θ_0 denotes the true value of θ.

First we observe that

$$\text{trace } E f_x I_p = f_x S$$

and second

$$\text{trace}(E\theta_0 F \theta_0^*) = \sum_{u=1}^{S} f_u\, \theta_{u_0}^* E \theta_{u_0} \ .$$

Since the largest eigenvalue of E is one,

$$\theta_{u_0}^* E \theta_{u_0} \le \|\theta_{u_0}\|^2 = p$$

and this bound is attained when θ_{u_0} is in the span of $\theta_1,\theta_2, \ldots,\theta_S$. Hence the integrand in [14] is maximized when $\theta_{1_0},\theta_{2_0},\ldots,\theta_{S_0}$ are all

in the span of $\theta_1,\theta_2,\ldots,\theta_S$. According to Condition 3.2, this occurs at all frequencies in B if and only if $\tau=\tau_0$.

The remainder of the proof proceeds as before. //

THEOREM 3.12

If Conditions 2.1, 3.1 and 3.5 are satisfied, and the noise is independent between receivers then the variance matrix of $N^{\frac{1}{2}}(\hat{\tau}_N-\tau_0)$ tends to Σ where,

$$\Sigma = A^{-1}(C+D)A^{-1} .,$$

$$A_{jk} = (2\pi)^{-1}\int_B W(\omega)\ \mathrm{trace}[(\theta^{(j)*}(I_p-E)\theta^{(k)}+\theta^{(k)*}(I_p-E)\theta^{(j)})F]\ d\omega$$

$$C_{jk} = (2\pi)^{-1}\int_B W(\omega)^2\mathrm{trace}[(\theta^{(j)*}(I_p-E)\theta^{(k)}+\theta^{(k)*}(I_p-E)\theta^{(j)})(\theta^{*}\theta)^{-1}]f_x^{\,2}\ d\omega$$

$$D_{jk} = (2\pi)^{-1}\int_B W(\omega)^2\mathrm{trace}[(\theta^{(j)*}(I_p-E)\theta^{(k)}+\theta^{(k)*}(I_p-E)\theta^{(j)})F]f_x\ d\omega$$

and these expressions are evaluated at τ_0.

PROOF

As before we must consider $\frac{\partial^2}{\partial\tau_j\partial\tau_k}q(\tau)$ and the variance of $N^{\frac{1}{2}}\frac{\partial}{\partial\tau_j}q_N(\tau)$:

$$\frac{\partial^2}{\partial\tau_j\partial\tau_k}q(\tau) = -(2\pi)^{-1}\int_B W(\omega)\ \mathrm{trace}[\theta_0{}^{*}(\frac{\partial^2}{\partial\tau_j\partial\tau_k}E)\theta_0 F]\ d\omega \quad .$$

First, we note that

$$(I_p-E)\theta = 0$$

and $\theta^{*}(I_p-E) = 0$.

Also:

$$\frac{\partial}{\partial\tau_j} E = \theta^{(j)}(\theta^*\theta)^{-1*} - \theta(\theta^*\theta)^{-1}(\theta^*\theta)^{(j)}(\theta^*\theta)^{-1\theta^*} + \theta(\theta^*\theta)^{-1}\theta^{(j)*}$$

$$= (I_p - E)\theta^{(j)}(\theta^*\theta)^{-1}\theta + \theta(\theta^*\theta)^{-1}\theta^{(j)*}(I_p - E)$$

and hence

$$\theta^* \frac{\partial}{\partial\tau_j} E\,\theta = 0$$

and

$$\frac{\partial}{\partial\tau_j} E\,\theta = (I_p - E)\theta^{(j)} .$$

Then

$$\frac{\partial^2}{\partial\tau_j\partial\tau_k} E = -\frac{\partial}{\partial\tau_k} E\,\theta^{(j)}(\theta^*\theta)^{-1}\theta + (I_p - E)\frac{\partial}{\partial\tau_k}(\theta^{(j)}(\theta^*\theta)^{-1}\theta)$$

$$+ \frac{\partial}{\partial\tau_k}(\theta(\theta^*\theta)^{-1}\theta^{(j)*})(I_p - E) - \theta(\theta^*\theta)^{-1}\theta^{(j)*}\frac{\partial}{\partial\tau_k} E$$

and therefore

$$\theta^* \frac{\partial^2}{\partial\tau_j\partial\tau_k} E\,\theta = -\theta^{(k)*}(I_p - E)\theta^{(j)} - \theta^{(j)*}(I_p - E)\theta^{(k)} .$$

Thus we have

$$\frac{\partial^2}{\partial\tau_j\partial\tau_k} q(\tau_0) = -(2\pi)^{-1}\int_B W(\omega)\,\mathrm{trace}[(\theta^{(j)*}(I_p - E)\theta^{(k)} + \theta^{(k)*}(I_p - E)\theta^{(j)})F]d\omega$$

which gives the expression for A_{jk}. Now consider asymptotic covariance of $N^{\frac{1}{2}}\frac{\partial}{\partial\tau_j} q_N(\tau_0)$ and $N^{\frac{1}{2}}\frac{\partial}{\partial\tau_k} q_N(\tau_0)$. By Lemma 3.6 this is

$$(2\pi)^{-1}\int_B W(\omega)^2\,\mathrm{trace}\{\frac{\partial}{\partial\tau_j} E(F_x I_p + \theta F\theta^*)\frac{\partial}{\partial\tau_k} E(f_x I_p + \theta F\theta^*)\}\,d\omega$$

$$+ (2\pi)^{-1}\int_B\int_B W(\omega)W(\omega')\sum_{abcd}\sum\sum\sum (\frac{\partial}{\partial\tau_j} E(\omega))_{ab}(\frac{\partial}{\partial\tau_k} E(\omega'))_{dc}$$

$$\hat{\beta}_{abcd}(\omega,-\omega,\omega',-\omega')\,d\omega\,d\omega' ,$$

where $\hat{\beta}$ is the fourth order cumulant spectrum of $y(n)$.

Now

$$\text{trace}\{\frac{\partial}{\partial\tau_j} E \frac{\partial}{\partial\tau_k} E\} = \text{trace}\{((I_p-E)\theta^{(j)}(\theta^*\theta)^{-1}\theta^* + \theta(\theta^*\theta)^{-1}\theta^{(j)*}$$

$$(I_p-E))(I_p-E)\theta^{(k)}(\theta^*\theta)^{-1}\theta^*$$

$$+ \theta(\theta^*\theta)^{-1}\theta^{(k)*}(I_p-E))\}$$

$$= \text{trace}\{(\theta^{(j)*}(I_p-E)\theta^{(k)} + \theta^{(k)*}(I_p-E)\theta^{(j)})$$

$$(\theta^*\theta)^{-1}\}$$

since I_p-E is idempotent. This gives the term C_{jk}.

We also have

$$\text{trace}\{\frac{\partial}{\partial\tau_j} E \frac{\partial}{\partial\tau_k} E\theta F\theta^*\} =$$

$$= \text{trace}\{\theta^* \frac{\partial}{\partial\tau_j} E \frac{\partial}{\partial\tau_k} E\theta F\}$$

$$= \text{trace}\{\theta^{(j)*}(I_p-E)(I_p-E)\theta^{(k)}F\}$$

$$= \text{trace}\{\theta^{(j)*}(I_p-E)\theta^{(k)}F\}$$

and similarly

$$\text{trace}\{\frac{\partial}{\partial\tau_k} E \frac{\partial}{\partial\tau_j} E\theta F\theta^*\}$$

$$= \text{trace}\{\theta^{(k)*}(I_p-E)\theta^{(j)}F\} .$$

Together these give the term D_{jk}.

Now consider the signal with signal term in the variance.

$$\text{trace}\{\frac{\partial}{\partial\tau_j} E\theta F\theta^* \frac{\partial}{\partial\tau_k} E\theta F\theta^*\}$$

$$= \text{trace}\{(\theta^* \frac{\partial}{\partial\tau_j} E\theta)F(\theta^* \frac{\partial}{\partial\tau_k} E\theta)F\}$$

$$= 0 .$$

Finally we must show that the fourth cumulant term is zero. Note that

$$\hat{\beta}_{abcd}(\omega,-\omega,\omega',-\omega') = \sum_{u=1}^{S} e^{-i\{\theta_u(\omega,a)-\theta_u(\omega,b)+\theta_u(\omega',c)-\theta_u(\omega',d)\}} \hat{\beta}_u(\omega,-\omega,\omega',-\omega')$$

where $\hat{\beta}_u$ is the fourth cumulant spectrum of $z_u(n)$. Thus the fourth cumulant term in the variance is

$$(2\pi)^{-1}\int_B\int_B W(\omega)W(\omega') \sum_{u=1}^{S} (\theta_u^*(\omega) \frac{\partial}{\partial\tau_j} E(\omega)\theta_u(\omega))(\theta_u(\omega')^* \frac{\partial}{\partial\tau_k} E(\omega')\theta_u(\omega')) \hat{\beta}_u(\omega,-\omega,\omega',-\omega')\, d\omega d\omega'$$

which is zero. //

The terms C and D in the variance may be regarded as noise with noise, and signal with noise terms respectively. The absence of a signal with signal term indicates that, if the signal to noise ratios are high enough, the deterministic estimator must give smaller variances than the Euclidean estimator. In fact in this situation the estimating functions for the deterministic estimator and the maximum likelihood estimator are very similar. If the noise with noise term is neglected and $W(\omega) = f_x(\omega)^{-1}$ then

$$(\Sigma^{-1})_{jk} = (2\pi)^{-1}\int_B \sum_{u=1}^{S} (\theta_u^{(j)*}(I_p-E)\theta_u^{(k)*}+\theta_u^{(k)*}(I_p-E)\theta_u^{(j)})f_u/f_x\, d\omega.$$

Thus parameters which affect different signals will have uncorrelated estimates when the signal to noise ratios are high.

3.6 METHODS OF APPLICATION

In most situations neither the number of signals present nor their spectra would be known. A practical procedure to follow would be to "fit" the signals one at a time until one is satisfied that no more

are present. This is the reverse of the normal statistical method of treating hierarchical models. It is obvious that the deterministic estimator is the one to use first.

Two problems arise with this. First the number of parameters (or the dimension of τ) is usually proportional to the number of signals and when this is large there are significant computational problems in finding the maximum of the estimating function. Normal optimization algorithms are not suitable due to the large number of local maxima usually present. A search must be made, and in high dimensions this is a time-consuming task. Second, a decision rule is required to decide how many signals are present.

To overcome the first problem we propose the following procedure:

1. Assume one signal is present and estimate the parameters associated with it, using the deterministic estimator.

2. Keeping the parameters for the first signal fixed, search for a second signal, again using the deterministic estimator and obtain a preliminary estimate of the parameters of the second signal.

3. Now we are close to the maximum of the estimating function (for S=2). An appropriate optimization algorithm can then be used to accurately locate this.

4. Repeat the last two steps to find further signals (S=3,4,5,...).

This procedure reduces the searches to an absolute minimum and appears to be effective in practice (see Chapter 6). The following result assists in the computations.

Let $\theta_{(S)} = (\theta_1,\theta_2,\ldots,\theta_S)$, $E_{(S)} = \theta_{(S)}(\theta_{(S)}^{*}\theta_{(S)})^{-1}\theta_{(S)}^{*}$, and

$$q_{(S)}(\tau_{(S)}) = N^{-1} \sum_{\omega\epsilon B} W(\omega) Y^{*} E_{(S)} Y \ . \qquad [15]$$

We define $\hat{\tau}_{(S)}$ to be that value which maximizes [15]. The dimension

of $\tau_{(S)}$ and $\hat{\tau}_{(S)}$ obviously depends upon S.

LEMMA 3.13

$$E_{(S+1)} = E_{(S)} + \Delta^{-1}(I_p - E_{(S)})\theta_{S+1}\theta_{S+1}^{*}(I_p - E_{(S)})$$

where $\Delta = p - \theta_{S+1}^{*}E_{(S)}\theta_{S+1} = \theta_{S+1}^{*}(I_p - E_{(S)})\theta_{S+1}$.

PROOF

Observe that $\theta_{(S+1)} = (\theta_{(S)}, \theta_{S+1})$, and

$$\theta_{(S+1)}^{*}\theta_{S+1} = \begin{pmatrix} \theta_{(S)}^{*}\theta_{(S)} & \theta_{(S)}^{*}\theta_{S+1} \\ \theta_{S+1}^{*}\theta_{(S)} & p \end{pmatrix}$$

Therefore

$$\left\{\theta_{(S+1)}^{*}\theta_{(S+1)}\right\}^{-1} = \begin{pmatrix} A & b^{*} \\ b & \Delta^{-1} \end{pmatrix}$$

where

$$A = (\theta_{(S)}^{*}\theta_{(S)})^{-1} + \Delta^{-1}(\theta_{(S)}^{*}\theta_{(S)})^{-1}\theta_{(S)}^{*}\theta_{S+1}\theta_{S+1}^{*}\theta_{(S)}(\theta_{(S)}^{*}\theta_{(S)})^{-1}$$

and

$$b = -\Delta^{-1}\theta_{S+1}^{*}\theta_{(S)}(\theta_{(S)}^{*}\theta_{(S)})^{-1}$$

and

$$E_{(S+1)} = \theta_{(S+1)}(\theta_{(S+1)}^{*}\theta_{(S+1)})^{-1}\theta_{(S+1)}^{*}$$

$$= E_{(S)} + \Delta^{-1} E_{(S)}\theta_{S+1}\theta_{S+1}^{*}E_{(S)} - \Delta^{-1} E_{(S)}\theta_{S+1}\theta_{S+1}^{*}$$

$$- \Delta^{-1}\theta_{S+1}\theta_{S+1}^{*}E_{(S)} + p\,\Delta^{-1}$$

$$= E_{(S)} + \Delta^{-1}(I_p - E_{(S)})\theta_{S+1}\theta_{S+1}^{*}(I_p - E_{(S)}) \ . \ //$$

Hence we have

$$q_{(S+1)}(\tau_{(S+1)} = N^{-1} \sum_{\omega \in B} W(\omega)\, Y^{*}E_{(S+1)}Y$$

$$= N^{-1} \sum_{\omega \epsilon B} W(\omega) Y^{*} E_{(S)} Y + N^{-1} \sum_{\omega \epsilon B} W(\omega) \Delta^{-1} \left| Y^{*} (I_p - E_{(S)}) \theta_{S+1} \right|^2$$

$$= q_{(S)}(\bar{\tau}_{(S+1)}) + N^{-1} \sum_{\omega \epsilon B} W(\omega) \left| Y^{*} (I_p - E_{(S)}) \theta_{S+1} \right|^2 (p - \theta_{S+1}^{*} E_{(S)} \theta_{S+1})^{-1}$$

[16]

where $\bar{\tau}_{(S+1)}$ is the vector $\tau_{(S+1)}$ with the components associated with the S+1 signal omitted. At Step 2 of the above procedure we keep $\bar{\tau}_{(S+1)}$ fixed at $\hat{\tau}_{(S)}$ and only the last term of [16] is then maximized with respect to the parameters of the S+1 th signal. These parameters affect only θ_{S+1}, not $E_{(S)}$.

The second problem, of determining the number of signals present, is very difficult to solve completely. However, the following reasoning can be used to derive a decision rule.

The deterministic model [11] can be written as

$$Y(\omega_r) = \theta_{(S)}(\omega_r) Z_{(S)}(\omega_r) + X(\omega_r) \qquad [17]$$

$$\omega_r = 2\pi \, r \, N^{-1} \, , \; r = 0,1,2,\ldots \left[\frac{N}{2}\right]$$

where $Y(\omega_r)$ and $X(\omega_r)$ are the vector Fourier transforms of y(n) and x(n) respectively and $Z_{(S)}(\omega_r) = (Z_1(\omega_r), \ldots, Z_S(\omega_r))'$.

Since the Fourier transforms at the fundamental frequencies are approximately uncorrelated, [17] corresponds to the standard linear model. The hypothesis that only S-1 signals are present is equivalent to the hypothesis that $Z_S(\omega_r) = 0$, $r = 0,1,2,\ldots[\frac{N}{2}]$. Furthermore if $W(\omega) = f_x^{-1}(\omega)$, then $q_{(S)}(\tau_{(S)})$ is $2N^{-1}$ times the log likelihood for the model. Thus, by the standard theory of likelihood ratio tests (see for example Wilks, 1962, p. 408) we would expect that

$$N \left(q_{(S)} (\hat{\tau}_{(S)}) - q_{(S-1)} (\hat{\tau}_{(S-1)}) \right) \qquad [18]$$

would have an approximate χ^2_N distribution under the null hypothesis.

This argument ignores the change in the design matrix $\theta_{(S)}$ so that it must be applied with care. Still one would expect that the statistic given by [18] would be reasonable.

3.7 ESTIMATION OF SPECTRA

Once a reasonable estimate $\hat{\tau}_N$ has been obtained by use of the deterministic estimator, the estimation of signal and noise spectra is an obvious problem. What follows is an informal discussion of how this may be done.

Consider the situation where $\hat{\tau}_N$ differs from τ_0 by an insignificant amount. If B is a narrow band of width ε, centred on ω_0 and $-\omega_0$, then

$$N^{-1} \sum_{\omega \varepsilon B} Y^*(I_p - E)Y$$

will tend almost surely to

$$(2\pi)^{-1} \varepsilon \operatorname{trace}\{(I_p - E)(f_x(\omega_0)I_p + \theta F \theta^*)\}$$

$$= (2\pi)^{-1} \varepsilon (p-S) f_x(\omega_0) \quad .$$

Hence we can use $2\pi(\varepsilon N(p-S))^{-1} \sum_{\omega \varepsilon B} Y^*(I_p - E)Y$ as an estimate of $f_x(\omega_0)$. This is very convenient since the terms Y^*EY would have already been calculated when estimating τ_0.

The estimation of signal spectra is more difficult. Methods based simply on moments tend to be unstable, particularly at frequencies where $\theta^*\theta$ is almost singular. An approximate maximum likelihood estimate could be used. If we use [3] for F_y^{-1} and Lemma 3.1, we have as the log likelihood

$$\sum_{\omega\in B} \{-(p-S)\log f_x - \log(\det \theta^*\theta) - \log\det(f_x(\theta^*\theta)^{-1}+F)$$

$$-f_x^{-1}Y^*(I_p-A)Y - Y^*\theta(\theta^*\theta)^{-1}(f_x(\theta^*\theta)^{-1}+f)^{-1}(\theta^*\theta)^{-1}\,{}^*Y\}\ .$$

If f_x and τ_0 are known or estimated sufficiently accurately we may ignore all but the third and last terms. Setting $(\theta^*\theta)^{-1}\theta^*Y = U$, we have

$$\sum_{\omega\in B} -\log\det(f_x(\theta^*\theta)^{-1}+F) - U^*(f_x(\theta^*\theta)^{-1}+F)^{-1}U$$

which involves only S×S matrices. It is simple enough to be minimized at a reasonable number of frequencies to estimate F.

3.8 NON QUADRATIC ESTIMATING FUNCTIONS

It seems appropriate to discuss the behaviour of the non quadratic estimators of Section 2.4 when more than one signal is present. Again, we will only consider the true spectral density matrix. Let $S = \text{span}(\theta_1,\theta_2,\ldots\theta_S)$.

The first results are of interest in low noise level situations.

THEOREM 3.14

If $F_y = f_x I_p + \theta F\theta^*$ then for $\varepsilon < 0$

$$\lim_{f_x\to 0} q_\varepsilon(\tau) = 0 \quad \text{if} \quad \Phi(\tau) \notin S$$

and $$\lim_{f_x\to 0} q_\varepsilon(\tau) > 0 \quad \text{if} \quad \Phi(\tau) \in S\ .$$

PROOF

The eigenvalues are solutions of

$$\det(F_y - \lambda I_p) = 0\ .$$

But $\det(F_y - \lambda I_p) = \det((f_x - \lambda)I_p + \theta F \theta^*)$

$$= (f_x - \lambda)^{p-S} \det(\theta^* \theta F + (f_x - \lambda) I_S).$$

Thus we have as eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_S > f_x$ with associated orthonormal eigenvectors $U_1, U_2, \ldots, U_S \in S$ and $\lambda_{S+1}, \lambda_{S+2}, \ldots, \lambda_p = f_x$ with eigenvectors $U_{S+1}, U_{S+2}, \ldots, U_p$, each orthogonal to S. As $f_x \to 0, \lambda_1, \lambda_2, \ldots, \lambda_S$ will tend to the eigenvalues of $\theta^* \theta F$ which are positive. Hence $\lim_{f_x \to 0} \sum_{i=1}^{S} \lambda_i^{\varepsilon} |\Phi(\tau)^* U_i|^2$ will be finite, and, if $\Phi(\tau)$ is not orthogonal to S, it will be non zero.

However $\lim_{f_x \to 0} \sum_{i=S+1}^{p} \lambda_i^{\varepsilon} |\Phi(\tau)^* U_i|^2$

$$= \lim_{f_x \to 0} f_x^{-\varepsilon} \sum_{i=S+1}^{p} |\Phi(\tau)^* U_i|^2$$

is infinity for $\varepsilon < 0$, unless $\Phi(\tau)^* U_i = 0$, $i = S+1, \ldots, p$; that is, unless $\Phi(\tau) \in S$. Hence

$$\lim_{f_x \to 0} \left(\sum_{i=1}^{p} \lambda_i^{\varepsilon} |\Phi(\tau)^* U_i|^2 \right)^{1/\varepsilon} = 0 \text{ if } \Phi(\tau) \notin S.$$

If $\Phi(\tau) \in S$, the result is obvious. //

Note that this result is true even if F is not diagonal: that is if the signals are correlated with each other. In the case of Capon's estimator ($\varepsilon = -1$) we have a stronger result.

THEOREM 3.15

Let $F_y = cF_x + \theta F \theta^*$. Then

$$\lim_{c \to 0} q_{-1}(\tau) = 0 \qquad \text{if } \Phi(\tau) \notin S$$

and $$\lim_{c \to 0} q_{-1}(\tau) > 0 \qquad \text{if } \Phi(\tau) \in S$$

PROOF

From Rao (1973)

$$F_y^{-1} = c^{-1}\{F_x^{-1} - F_x^{-1}\theta(\theta^* F_x^{-1}\theta)^{-1}\theta^* F_x^{-1}$$

$$+ c\, F_x^{-1}\theta(\theta^* F_x^{-1}\theta)^{-1}(c(\theta^* F_x^{-1}\theta)^{-1}+F)^{-1}(\theta^* F_x^{-1}\theta)^{-1}\theta^* F_x^{-1}\}.$$

Let $\Phi = \theta\alpha+b$ where α is now an S dimensional vector and $\theta^* F_x^{-1} b = 0$. This corresponds to decomposing Φ into two vectors, one in S and the other in the space orthogonal to this, using the inner product $\langle u,v\rangle = u^* F_x^{-1} v$. Note that

$$b^* F_y^{-1} b = c^{-1}\, b^* F_x^{-1}\, b\ .$$

Hence

$$q_{-1}(\tau) = c\{b^* F_x b + c\alpha^*(c(\theta^* F_x^{-1}\theta)^{-1} + F)^{-1}\alpha\}^{-1}\ .$$

If $\Phi(\tau) \in S$, then $b = 0$, so

$$q_{-1}(\tau) = \{\alpha^*(c(\theta^* F_x^{-1}\theta)^{-1} + F)^{-1}\alpha\}^{-1}$$

and

$$\lim_{c\to 0} q_{-1}(\tau) = (\alpha^* F^{-1}\alpha)^{-1} > 0\ .$$

If $b \neq 0$ (that is $\Phi(\tau) \notin S$), then

$$\lim_{c\to 0} q_{-1}(\tau) = (\lim_{c\to 0} c)\{b^* F_x^{-1} b + \lim_{c\to 0} c\alpha^*(c(\theta^* F_x\theta)^{-1} + F)^{-1}\alpha\}^{-1}$$

$$= 0\ . \ //$$

This result, again for arbitrary F_x and F, is an extension of a result of Woods and Lintz (1973).

Several authors (for example, Der and Flinn, 1975) have suggested using the eigenvectors of an estimate of F_y directly.

It was noted in the proof of Theorem 3.14 that F_y has S eigenvalues greater than f_x and the other p-S equal to f_x. This could be used as a method of estimating the number of signals actually present. A decision would have to be made as to how many of the smaller eigenvalues belong to the set associated with f_x.

The number remaining would give an estimate of S.

If $\theta_1,\theta_2,\ldots,\theta_S$ are orthogonal, then each of the eigenvectors $U_1, U_2,\ldots,U_S$ would correspond to a particular phase vector. If they are approximately orthogonal (that is, the matrix $\theta^*\theta$ has small off diagonal entries) then each eigenvector would be dominated by a single phase vector. This domination would be the greatest for the strongest signal. As the eigenvectors are not in general allowable phase vectors, the phase vector of best fit is found by maximizing

$$g_i(\tau) = |\Phi(\tau)^* U_i|^2 \quad .$$

The first nontrivial case is when S=2. Then

$$\lambda_1 = f_x + \tfrac{1}{2}(p(f_1+f_2) + [p^2(f_1-f_2)^2 + 4f_1f_2|\psi|^2]^{\frac{1}{2}}),$$

$$\lambda_2 = f_x + \tfrac{1}{2}(p(f_1+f_2) - [p^2(f_1-f_2)^2 + 4f_1f_2|\psi|^2]^{\frac{1}{2}}).$$

Let $U_1 = \alpha_1\theta_1 + \alpha_2\theta_2$ and $U_2 = \beta_1\theta_1 + \beta_2\theta_2$. Then it is readily shown that if $\psi \neq 0$ then

$$\frac{\alpha_2}{\alpha_1} = \frac{f_2\bar{\psi}}{pf_2-\lambda_1} = \frac{\lambda_1-pf_1}{f_1\psi} \quad .$$

If f_1 is very much greater than f_2, this ratio is small indicating that U_1 is dominated by θ_1. However, in this case β_1/β_2 approximates $\bar{\psi}/p$ which need not be small and U_2 will not be dominated by θ_2. If f_1 and f_2 are of the same magnitude, neither eigenvector would be dominated by a single phase vector unless ψ is very small.

To summarize, we can assume that estimators like Capon's should give reasonable, although not consistent, results when the noise level is low. Methods which use the eigenvectors individually would be far less predictable.

CHAPTER 4

SIMULATION STUDIES

4.1 THE AIM

In the previous two chapters only asymptotic results were presented. Analogous results for a finite sample size are virtually impossible to obtain by analytic techniques.

It is hoped that the asymptotic results give an indication of small sample behaviour and to investigate this a series of simulations were carried out. The primary aims were to investigate the bias of small sample estimates, and to compare the asymptotic variances with the small sample variances. However several other properties could also be examined, including:

1. The structure of the sidelobes and the problems caused by them. The asymptotic results are concerned only with the estimating function in the neighbourhood of the maximum, whereas in practice secondary maxima can detract from the usefulness of an estimator.

2. The behaviour relative to other estimators including Capon's. In fact simulations appear to be the main way of obtaining variance results for many of these estimators.

3. Robustness.

The simulations do not give much information as to how the estimating functions should be maximized in practice since in the

simulations, our prior knowledge of the position of the maximum was used to advantage. A practical example with real data is given in Section 6.2.

It is apparent that the problem simulated involves so many possible variables that any study must be rather limited. However, the results presented here appear to be typical of the general case and are quite instructive.

4.2 PROGRAMMING CONSIDERATIONS AND THE MODEL USED

All computing was performed on a Univac 1108 computer at the Australian National University, mainly using demand terminals. Programs were written in Fortran V using single precision (9 decimal digits). The only library routine used was one for Fast Fourier transforms. Attempts were made to use library optimization routines but these invariably behaved badly with single precision.

A four element array was used, with receivers having position coordinates

0.54	0.84
-0.84	0.54
-0.54	-0.84
0.54	-0.54

The velocity was assumed known and as a function of frequency was given by $c(\omega) = 0.5 - 0.05\omega$. Two signals were modelled and the parameters to be estimated were directions of the signals. Hence τ was a two-dimensional vector giving the direction from which each signal arrived, measured in degrees (from 0° to 360°).

4.3 GENERATION OF SIGNALS AND NOISE

The starting point of the simulations was a sequence of pseudo random numbers, generated by standard programs, giving uncorrelated, approximately normally distributed variables with zero mean and unit variance. These were transformed into correlated sequences with the desired spectra by autoregressive filters. That is, if $\varepsilon(1),\varepsilon(2),\ldots$ was the initial uncorrelated sequence, we formed

$$x(n) = (\varepsilon(n) - \sum_{j=1}^{k} a_j\, x(n-j))/a_0 \qquad [1]$$

where $a_0, a_1,\ldots,a_k$ are constants. The x(n) process then has the spectrum

$$f_x(\omega) = (2\pi)^{-1} \left| \sum_{j=0}^{k} a_j\, e^{ij\omega} \right|^{-2}$$

Second order autoregressions were used almost exclusively (that is k=2). To start each sequence, the initial values of x(n) were set to zero and then the first 1000 generated were discarded so that the section used could reasonably be assumed stationary.

For the noise sequences this was all that was required. With the signals a greater problem arose since they had to be delayed appropriately at each receiver. Simple delay methods were not used since they cannot model dispersion and only a limited choice of delays could be generated. Instead, a discrete version of the defining equation [1.3] was used. If $Z_j(n)$ is the undelayed signal (at some fixed reference point) we formed

$$Z_j(\omega_r) = (N')^{-\frac{1}{2}} \sum_{n=1}^{N'} Z_j(n)\, e^{in\omega_r} ,$$

$$\omega_r = 2\pi r (N')^{-1} , \quad r = 0,1\ldots[\frac{N'}{2}]$$

and

$$Z_j(n,a) = (N')^{-1} \sum_{\omega_r} Z_j(\omega_r)\, e^{-i\{n\omega_r - \theta_j(\omega_n a)}$$

where $N' > 2N$. The sequence on which these operations were carried out was made longer than absolutely necessary so that the data would not be generated using precisely the same frequencies as would be used later to analyse it. To minimize leakage effects in the frequency domain, a partial cosine taper was applied to the long sequence and then only the centre of the transformed sequence was actually used.

4.4 EVALUATION OF THE ESTIMATES

We examined the behaviour of the maximum likelihood, Euclidean, deterministic, Capon's and the "conventional" one signal estimates. The estimating functions of the first three are functions of two variables and hence created the main problem. For efficiency it was convenient to evaluate these simultaneously on a grid, using the fact that all three involve several common terms. Furthermore, using a grid instead of irregular points reduced calculations by an order of magnitude. The Euclidean and deterministic estimators used weighting inversely proportional to the noise: that is $W(\omega) = f_x(\omega)^{-1}$.

In the case of the replicated simulations where a quick but accurate method was required for finding the maxima, the estimating functions were evaluated on a small grid around τ_0. Then $\hat{\tau}_N$ was found by a combination of taking the maximum on the grid and fitting a quadratic surface. This introduced slight errors, but these were an order of magnitude less than the standard deviation of the estimates.

Capon's estimator and the conventional estimator, being functions of only one variable, presented no such problems, and simple search

procedures could be used to maximize them. However a decision had to be made as to how to evaluate Capon's estimating function, so that it could be compared with the others properly. Capon's estimator, unlike the others, requires an estimate of the spectral density matrix, formed by averaging the periodogram over a band of fundamental frequencies. Neither the number of bands into which we should partition the interval $[0,\pi]$, nor the way in which we should combine the results from each band (for example arithmetic average or harmonic average) are at all obvious. Approximately $N^{\frac{1}{2}}$ bands were used as this should be a good compromise between frequency resolution and stability. Then the estimating functions from each band were combined by an arithmetic average, weighted inversely to the noise spectrum. Slightly better results might be obtained by altering this procedure but we doubt whether the changes would be significant.

4.5 RESULTS OF REPLICATED SIMULATIONS

Six different situations with N = 256 were modelled and replicated 50 times each, so that we could obtain reasonable information about the variances. The results are presented in Tables 1 to 6. Each table is reasonably self-explanatory, giving:

1. The spectra of the signals and noise.
2. The variance or power levels of the signals and noise.
3. The deviations of the means $\bar{\tau}$ of the estimates from the true values.
4. The sample variances and covariances.
5. The predicted variances and covariances using the results of Chapter 3.
6. $S = (\bar{\tau}-\tau_0)'\Sigma^{-1}(\bar{\tau}-\tau_0)$ where Σ is the sample variance matrix.

SIMULATION 1

True Signal Direction

100° *180°*

Spectra

Frequency	0	π/4	π/2	3π/4	π
Noise	.159	.159	.159	.159	.159
1st signal	.569	.687	.591	.295	.210
2nd signal	.341	.545	1.173	.577	.341

Variances

Noise	1.
1st signal	3.
2nd signal	4.

	ML	Euclidean	Determin-istic	Capon	Conven-tional
$\bar{\tau}_1 - \tau_{10}$	- .174	.011	- .172	.497	2.985
$\bar{\tau}_2 - \tau_{20}$	- .454	- .568	- .433	.125	.380
Observed					
$\text{Var}(\hat{\tau}_1)$	.481	.943	.533	.850	1.111
$\text{Var}(\hat{\tau}_2)$	.272	1.239	.247	.453	.707
$\text{Cov}(\hat{\tau}_1, \hat{\tau}_2)$	.055	- .119	.049	.027	- .265
Predicted					
$\text{Var}(\hat{\tau}_1)$	.2742	1.037	.404	-	-
$\text{Var}(\hat{\tau}_2)$	.180	.765	.189	-	-
$\text{Cov}(\hat{\tau}_1, \hat{\tau}_2)$	- .008	- .127	.008	-	-
S	.762	.265	.762	.317	9.866
MS error	.987	2.507	.994	1.566	10.966

SIMULATION 2

True Signal Direction

120° *150°*

Spectra

Frequency	0	π/4	π/2	3π/4	π
Noise	.109	.256	.198	.041	.027
1st signal	.163	.385	.298	.062	.041
2nd signal	.218	.513	.397	.083	.054

Variances

Noise	1.0
1st signal	1.5
2nd signal	2.0

	ML	Euclidean	Determin-istic	Capon	Conven-tional
$\bar{\tau}_1-\tau_{10}$	- .176	- .358	- .138	1.223	3.567
$\bar{\tau}_2-\tau_{20}$	- .268	- .108	- .301	- .920	-1.913
Observed					
$\mathrm{Var}(\hat{\tau}_1)$	.398	4.185	.449	1.390	2.067
$\mathrm{Var}(\hat{\tau}_2)$	.346	2.259	.361	.791	.708
$\mathrm{Cov}(\hat{\tau}_1,\hat{\tau}_2)$	- .039	- .718	- .083	- .083	- .614
Predicted					
$\mathrm{Var}(\hat{\tau}_1)$	.268	3.254	.411	-	-
$\mathrm{Var}(\hat{\tau}_2)$	.180	1.608	.277	-	-
$\mathrm{Cov}(\hat{\tau}_1,\hat{\tau}_2)$	.010	- .304	.000	-	-
S	.3163	.044	.350	1.987	7.538
MS error	.848	6.584	.919	4.520	19.155

SIMULATION 3

True Signal Directions

120° *150°*

Spectra

Frequency	0	π/4	π/2	3π/4	π
Noise	.2102	.3123	.151	.048	.035
1st signal	.1946	.2966	.165	.052	.037
2nd signal	.3160	.5008	.413	.134	.092

Variances

Noise	1.
1st signal	1.
2nd signal	2.

	ML	Euclidean	Determin-istic	Capon	Conven-tional
$\bar{\tau}_1-\tau_{10}$	- .193	- .335	- .150	1.202	8.7639
$\bar{\tau}_2-\tau_{20}$	- .301	- .258	- .330	- .589	-1.172
Observed					
$\mathrm{Var}(\hat{\tau}_1)$	.510	4.751	.593	2.492	17.201
$\mathrm{Var}(\hat{\tau}_2)$	.133	.640	.139	.330	.189
$\mathrm{Cov}(\hat{\tau}_1,\hat{\tau}_2)$	.038	.083	.054	.199	- .387
Predicted					
$\mathrm{Var}(\hat{\tau}_1)$	.429	4.01	.601	-	-
$\mathrm{Var}(\hat{\tau}_2)$	.121	.427	.201	-	-
$\mathrm{Cov}(\hat{\tau}_1,\hat{\tau}_2)$	.011	- .195	.000	-	-
S	.704	.1234	.782	2.074	14.812
MS error	.771	5.569	.864	4.612	95.569

SIMULATION 4

True Signal Direction

120° *160°*

Spectra

Frequency	0	π/4	π/2	3π/4	π
Noise	.1404	.1579	.1804	.1579	.1404
1st signal	.1382	.3323	.1448	.0325	.0218
2nd signal	.2102	.3123	.1508	.0494	.0347

Variances

Noise	1.
1st signal	1.
2nd signal	1.

	ML	Euclidean	Determin-istic	Capon*	Conven-tional
$\bar{\tau}_1-\tau_{10}$	.3205	.2526	.4686	6.4356	7.0430
$\bar{\tau}_2-\tau_{20}$	- .1179	- .1028	- .1143	-5.3186	-7.5148
Observed					
$\text{Var}(\hat{\tau}_1)$	2.7421	10.3578	2.8086	40.7607	9.0387
$\text{Var}(\hat{\tau}_2)$	2.1357	10.0863	2.6001	27.9836	15.4576
$\text{Cov}(\hat{\tau}_1,\hat{\tau}_2)$	.2718	- .4268	.3274	-3.6131	- .0483
Predicted					
$\text{Var}(\hat{\tau}_1)$	1.7421	7.2177	2.9237	-	-
$\text{Var}(\hat{\tau}_2)$	1.3939	6.9528	2.2777	-	-
$\text{Cov}(\hat{\tau}_1,\hat{\tau}_2)$	.0338	- .3946	.1154	-	-
S	.0480	.0075	.0893	1.8331	9.1049
MS error	5.0023	20.5194	5.6414	138.448	130.5724

* Calculated for 47 replications.

SIMULATION 5

True Signal Direction

200° *250°*

Spectra

Frequency	0	π/4	π/2	3π/4	π
Noise	.059	.080	.153	.251	.223
1st signal	.097	.151	.094	.029	.019
2nd signal	.136	.211	.131	.040	.026

Variances

Noise	1.0
1st signal	.5
2nd signal	.7

	ML	Euclidean	Determin-istic	Capon*	Conven-tional
$\bar{\tau}_1-\tau_{10}$	-1.071	-1.142	-1.397	8.280	10.980
$\bar{\tau}_2-\tau_{20}$	- .377	- .246	- .508	-2.761	-4.409
Observed					
$\text{Var}(\hat{\tau}_1)$	3.362	10.875	12.511	71.309	35.446
$\text{Var}(\hat{\tau}_2)$	3.275	5.455	4.674	16.685	8.568
$\text{Cov}(\hat{\tau}_1,\hat{\tau}_2)$	.238	-1.891	-1.403	-18.075	-8.555
Predicted					
$\text{Var}(\hat{\tau}_1)$	3.405	10.918	6.163	-	-
$\text{Var}(\hat{\tau}_2)$	2.151	3.861	3.252	-	-
$\text{Cov}(\hat{\tau}_1,\hat{\tau}_2)$	- .013	-1.315	- .375	-	-
S	.3688	.158	.251	.998	11.064
MS error	7.926	17.693	19.361	163.1811	184.001

* Calculated from 43 replications.

SIMULATION 6

True Signal Direction

120° *150°*

Spectra

Frequency	0	π/4	π/2	3π/4	π
Noise	.140	.159	.1803	.157	.140
1st signal	.027	.087	.068	.010	.006
2nd signal	.048	.209	.055	.010	.006

Note: 1st signal has narrow peak of .500 at .38π
2nd signal has narrow peak of .696 at .34π

Variances

Noise	1.0
1st signal	.5
2nd signal	.7

	ML	Euclidean	Determin-istic	Capon	Conven-tional
$\bar{\tau}_1-\tau_{10}$	.315	-1.337	1.750	.108	-
$\bar{\tau}_2-\tau_{20}$	- .550	- .545	- .078	-2.133	-
Observed					
$\text{Var}(\hat{\tau}_1)$	7.124	38.222	24.245	74.742	-
$\text{Var}(\hat{\tau}_2)$	5.250	18.185	19.167	64.052	-
$\text{Cov}(\hat{\tau}_1,\hat{\tau}_2)$	.010	-10.227	1.776	19.583	-
Predicted					
$\text{Var}(\hat{\tau}_1)$	4.212	28.431	17.835	-	-
$\text{Var}(\hat{\tau}_2)$	3.312	15.138	13.347	-	-
$\text{Cov}(\hat{\tau}_1,\hat{\tau}_2)$	.691	1.617	.038	-	-
S	.071	.049	.128	.080	-
MS error	12.78	58.49	46.48	143.36	-

Note: *1.* Capon's estimator data based on 17 our of 50 replications.
2. Conventional estimator failed to resolve two signals.

This is a measure of the significance of the bias--the 5% level is 0.14.

7. The total mean square error.

In all cases trimmed means and variances were used: observations more than three sample standard deviations from the mean were rejected. This gave robustness to the means and variances while introducing only a slight 3% bias in the variances. In Simulations 4, 5 and 6, Capon's estimator required rather harsher treatment due to very non normal distributions.

Note that the simulations are presented in order of decreasing signal to noise ratio. In Simulation 6, the signals are also very narrowband which creates extra problems. We expect all estimators to behave best on Simulation 1 and worst on Simulation 6. The following are comments on the results.

1. Not surprisingly, the maximum likelihood estimator gave best results in all situations, in terms of both low bias and low variance. In Simulation 6 it was the only estimator which could be considered reasonably stable.

2. The Euclidean estimator was of doubtful use, always having much larger variance than the maximum likelihood estimator, and comparable bias. Considering that it requires prior knowledge of the spectra, this is not very impressive. It would only be used in practice for its computational simplicity.

3. The deterministic estimator gave surprisingly good results, usually being better than the Euclidean estimator and only marginally worse than the maximum likelihood estimator when the signal to noise ratio is high.

4. Capon's estimator behaved reasonably well in the high signal to noise ratio situation (as the results of Section 3.8 would

lead us to expect) but very poorly when the signal to noise ratio is low. In terms of total mean square error it was superior to the Euclidean estimator in Simulations 1, 2 and 3. In Simulation 6 it behaved very poorly--possibly the estimated spectral density matrix was almost singular in some replications.

5. As expected, the conventional estimator showed very significant bias, and moderate variance. In Simulation 6 it failed to resolve the two signals.

Note that in all situations, the predicted variances were close to the observed ones.

4.6 GRAPHICAL RESULTS

We now give the results of some unreplicated simulations to illustrate the estimating functions of the five estimators studied.

First, three situations were modelled, all with identical spectra but different signal directions. The signal directions are given by

Case	1st direction	2nd direction
A	70°	150°
B	180°	100°
C	130°	150°

The velocity function is given in Figure 4.1 and the spectra in Figure 4.2. The variances or power levels of the noise and two signals were 1.0, 5.0 and 1.5 respectively.

Figure 4.3 gives the result of applying the conventional beamforming method. The signals are resolved in Cases A and B but not in C. The sidelobes are somewhat disturbing as they could be mistaken for other signals. In Figure 4.4 we have the result of applying Capon's method. Case C is still not resolved. The peaks are sharper

than before but the sidelobes have a very similar structure even though they are slightly smaller.

Figures 4.5, 4.6 and 4.7 are contour diagrams of the maximum likelihood estimating function. In all cases the maximum,denoted by ∇, corresponds to the true combination of signal directions, even in Case C. The function surfaces are rather complex, but are dominated by ridges corresponding to the true signal directions. These ridges could be used in practice to help find the maximum.

The corresponding diagrams for the Euclidean estimator are found in Figures 4.8, 4.9 and 4.10. These surfaces are even more strongly dominated by the ridges corresponding to the true directions, and are somewhat smoother. Figures 4.11, 4.12 and 4.13 give the surfaces for the deterministic estimator. These are symmetric about the diagonal since, unlike the previous two estimators, the deterministic estimator does not distinguish between the two signals in terms of their spectra. The surfaces are again rather complex with many local maxima.

In order to demonstrate the effect of looking for too many signals, one more simulation was carried out with only one signal, direction 340°. The same velocity function as before was used and the spectra are given in Figure 4.14. The estimating functions for both the conventional beamforming method and Capon's method are shown in Figure 4.15. Both are dominated by the peak in the true direction but both also have a second peak at about 230°, which could very easily be mistaken for a second signal. The estimating function of the deterministic estimator does not have this problem. As seen in Figure 4.16, the ridges corresponding to the true signal direction are level, indicating the absence of further significant signals.

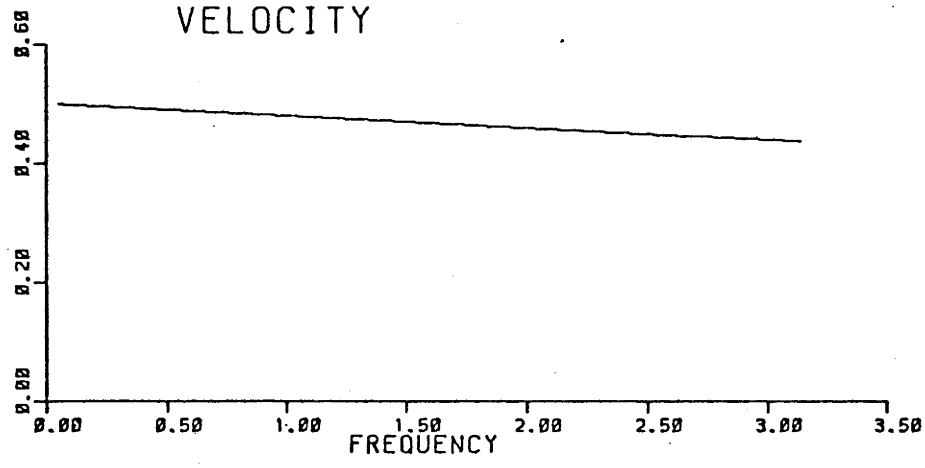


Figure 4.1

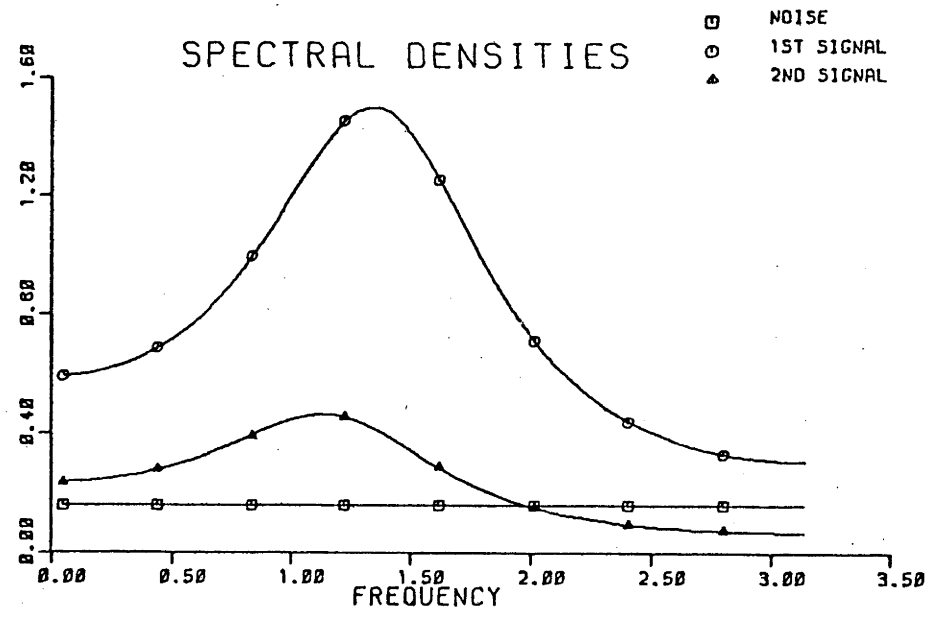
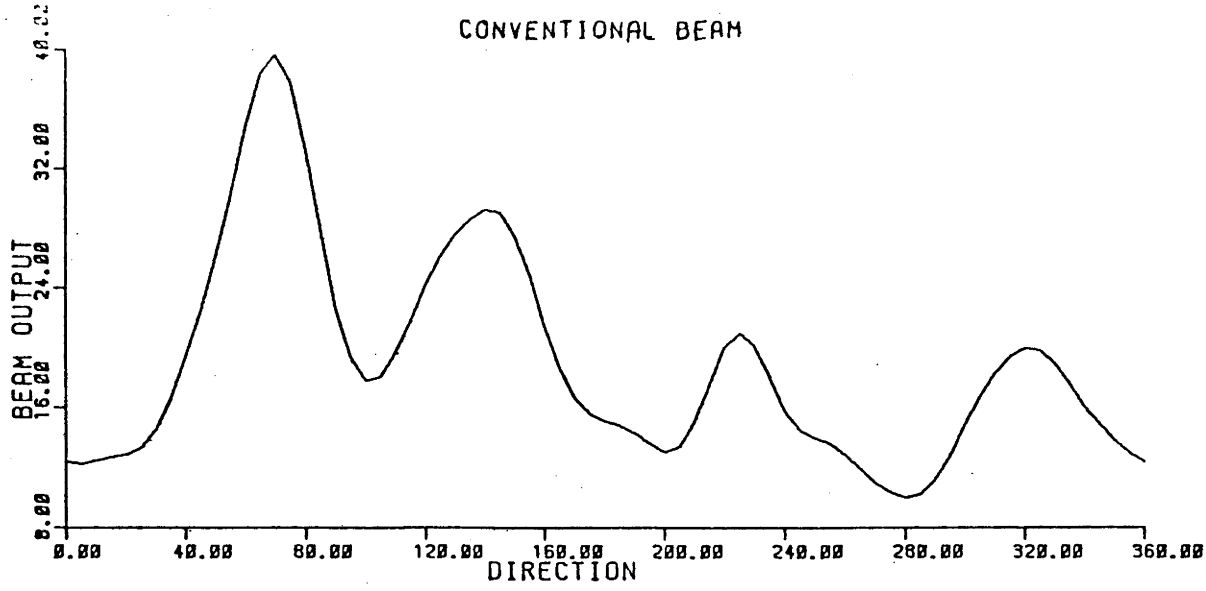
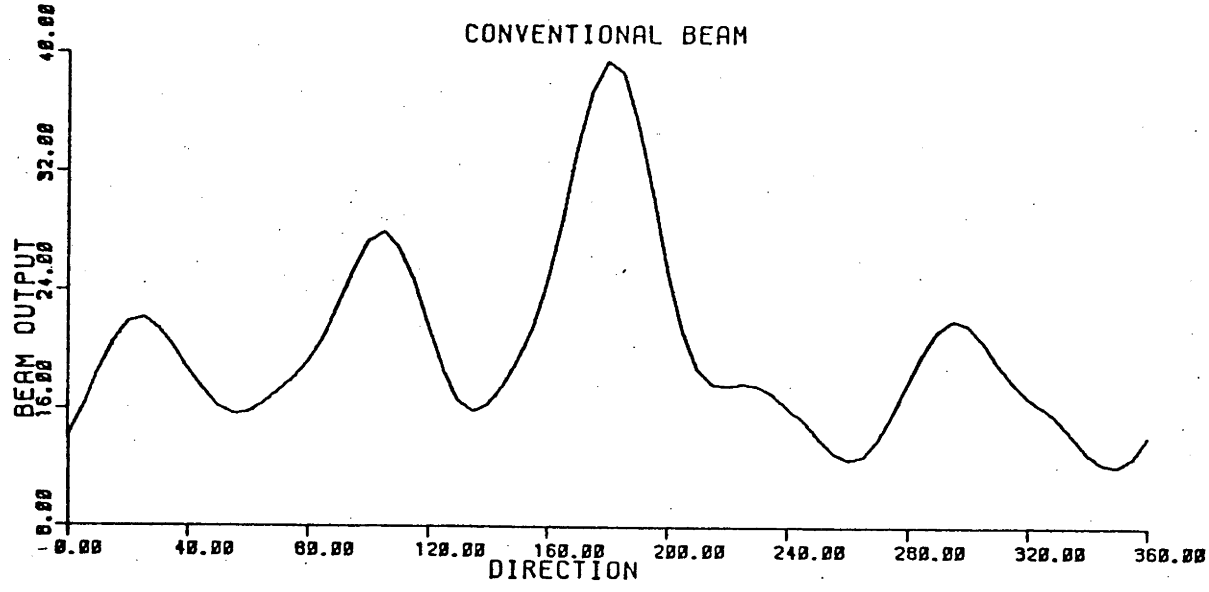


Figure 4.2


CONVENTIONAL BEAM
BEAM OUTPUT
8.00
16.00
24.00
32.00
0.00
40.00
80.00
120.00
160.00
200.00
240.00
280.00
320.00
360.00
DIRECTION


CONVENTIONAL BEAM
BEAM OUTPUT
8.00
16.00
24.00
32.00
40.00
0.00
40.00
80.00
120.00
160.00
200.00
240.00
280.00
320.00
360.00
DIRECTION

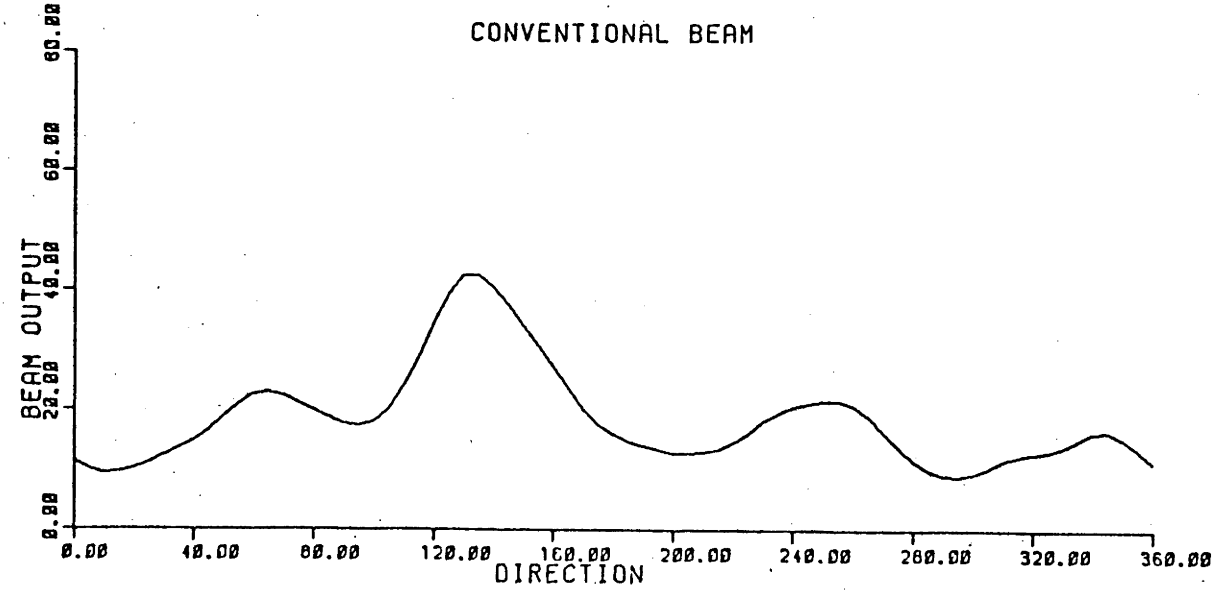

CONVENTIONAL BEAM
BEAM OUTPUT
8.00
20.00
40.00
60.00
80.00
0.00
40.00
80.00
120.00
160.00
200.00
240.00
280.00
320.00
360.00
DIRECTION

Figure 4.3

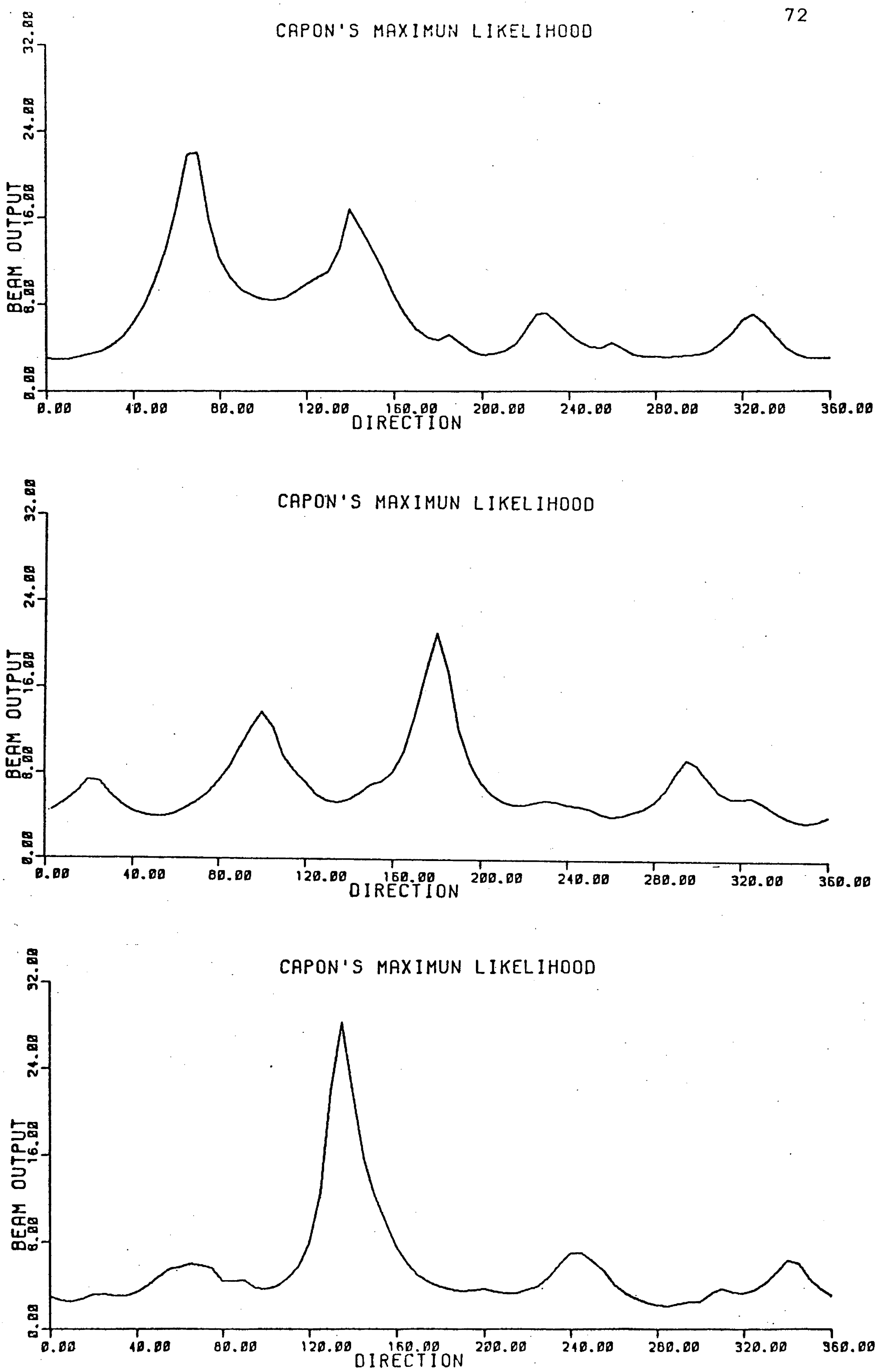

CAPON'S MAXIMUN LIKELIHOOD
BEAM OUTPUT
0.00
8.00
16.00
24.00
32.00
0.00
40.00
80.00
120.00
160.00
200.00
240.00
280.00
320.00
360.00
DIRECTION
CAPON'S MAXIMUN LIKELIHOOD
BEAM OUTPUT
DIRECTION
CAPON'S MAXIMUN LIKELIHOOD
BEAM OUTPUT
DIRECTION

Figure 4.4

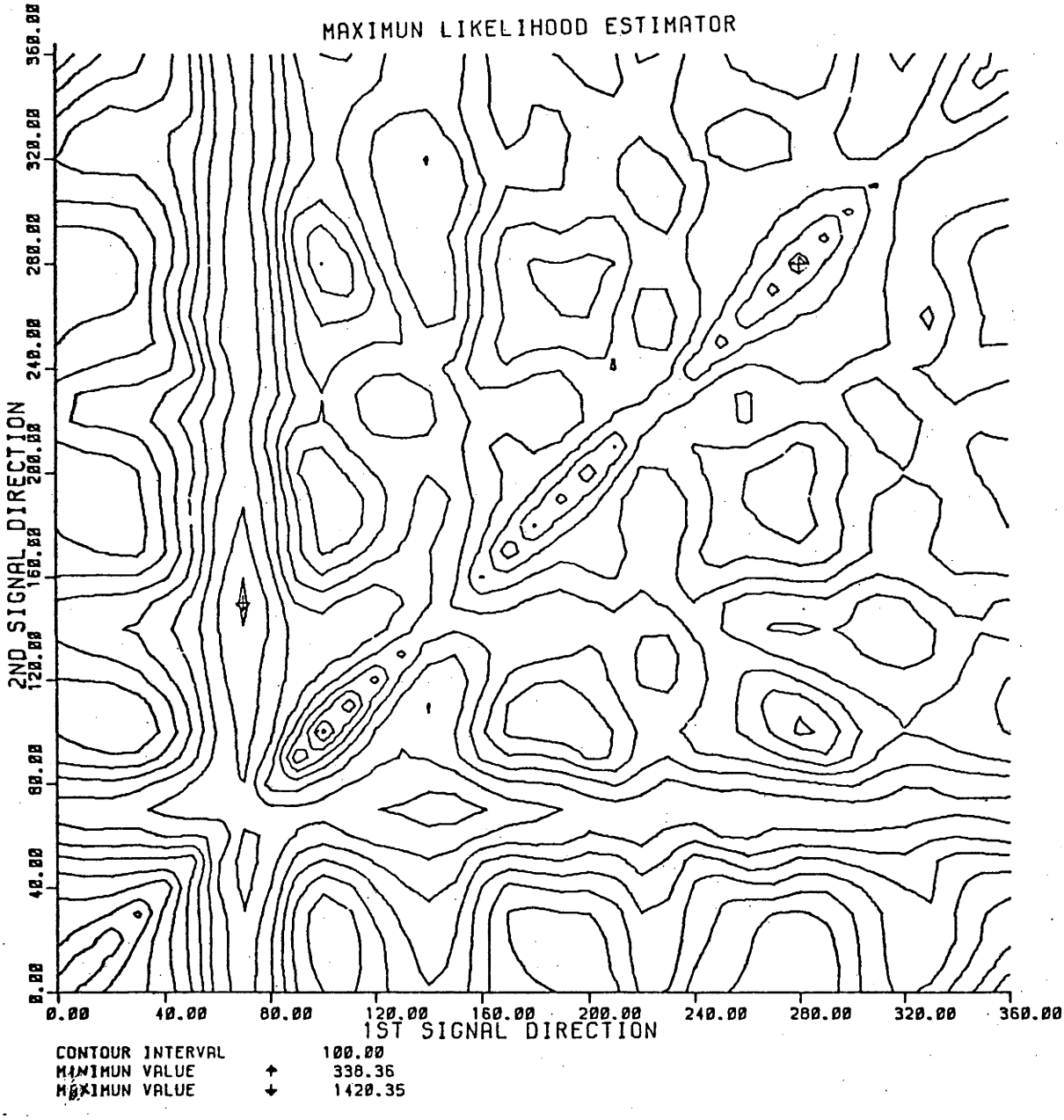

MAXIMUN LIKELIHOOD ESTIMATOR
2ND SIGNAL DIRECTION
0.00
40.00
80.00
120.00
160.00
200.00
240.00
280.00
320.00
360.00
1ST SIGNAL DIRECTION
CONTOUR INTERVAL 100.00
MINIMUN VALUE 338.36
MAXIMUN VALUE 1420.35

Figure 4.5

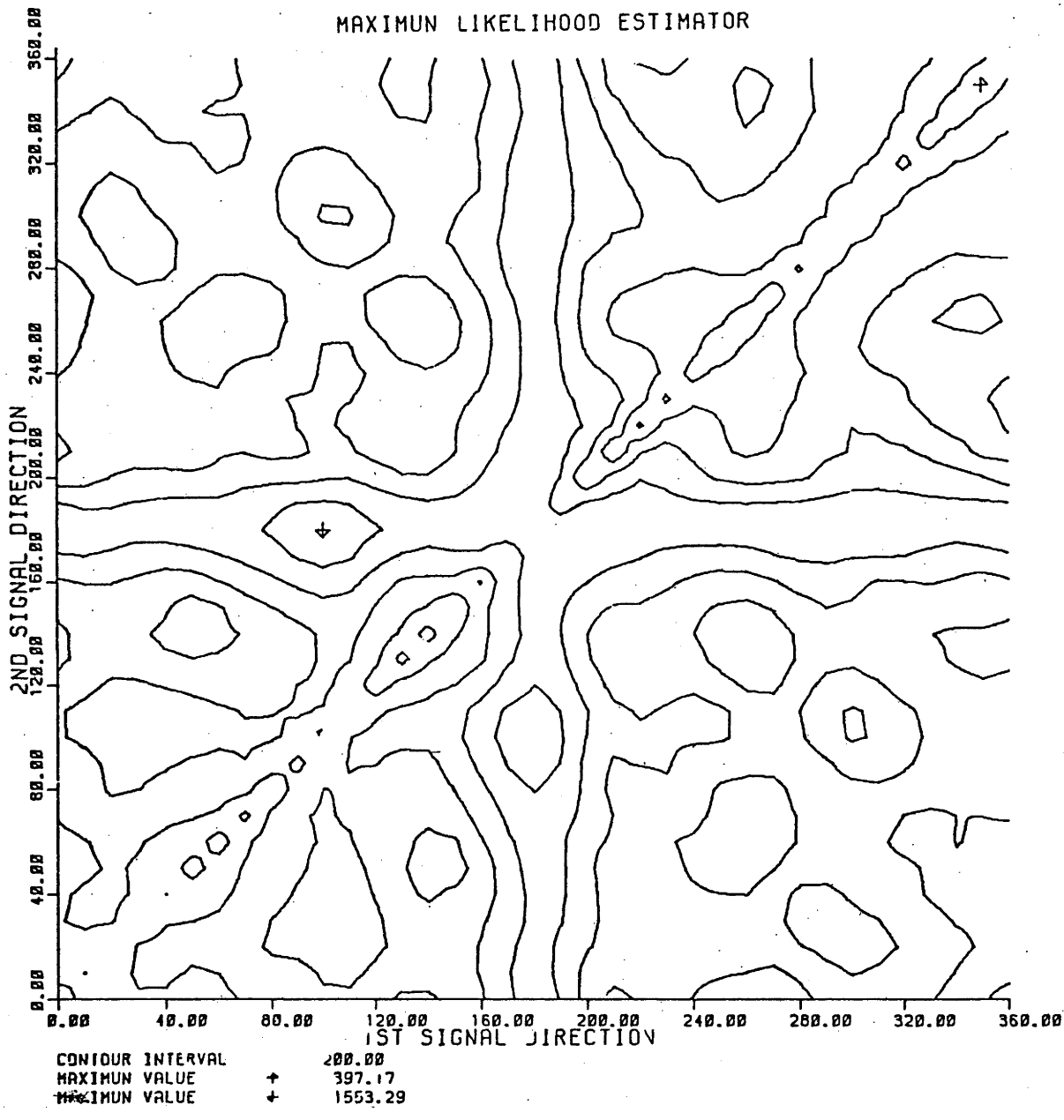

MAXIMUN LIKELIHOOD ESTIMATOR
2ND SIGNAL DIRECTION
0.00
40.00
80.00
120.00
160.00
200.00
240.00
280.00
320.00
360.00
1ST SIGNAL DIRECTION
CONTOUR INTERVAL 200.00
MAXIMUN VALUE 397.17
MAXIMUN VALUE 1553.29

Figure 4.6

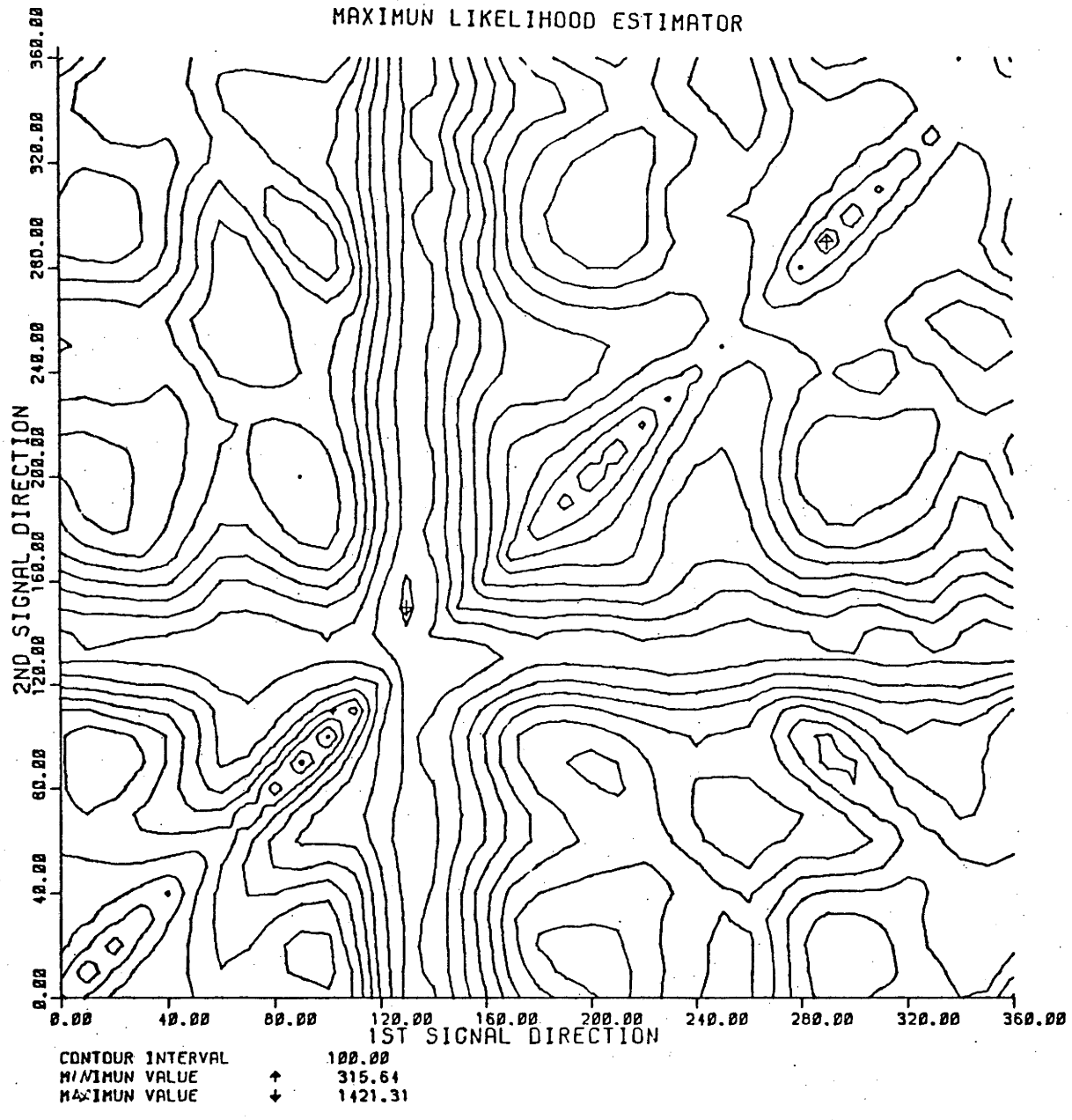

MAXIMUN LIKELIHOOD ESTIMATOR
2ND SIGNAL DIRECTION
0.00
40.00
80.00
120.00
160.00
200.00
240.00
280.00
320.00
360.00
1ST SIGNAL DIRECTION
CONTOUR INTERVAL 100.00
MINIMUN VALUE 315.64
MAXIMUN VALUE 1421.31

Figure 4.7

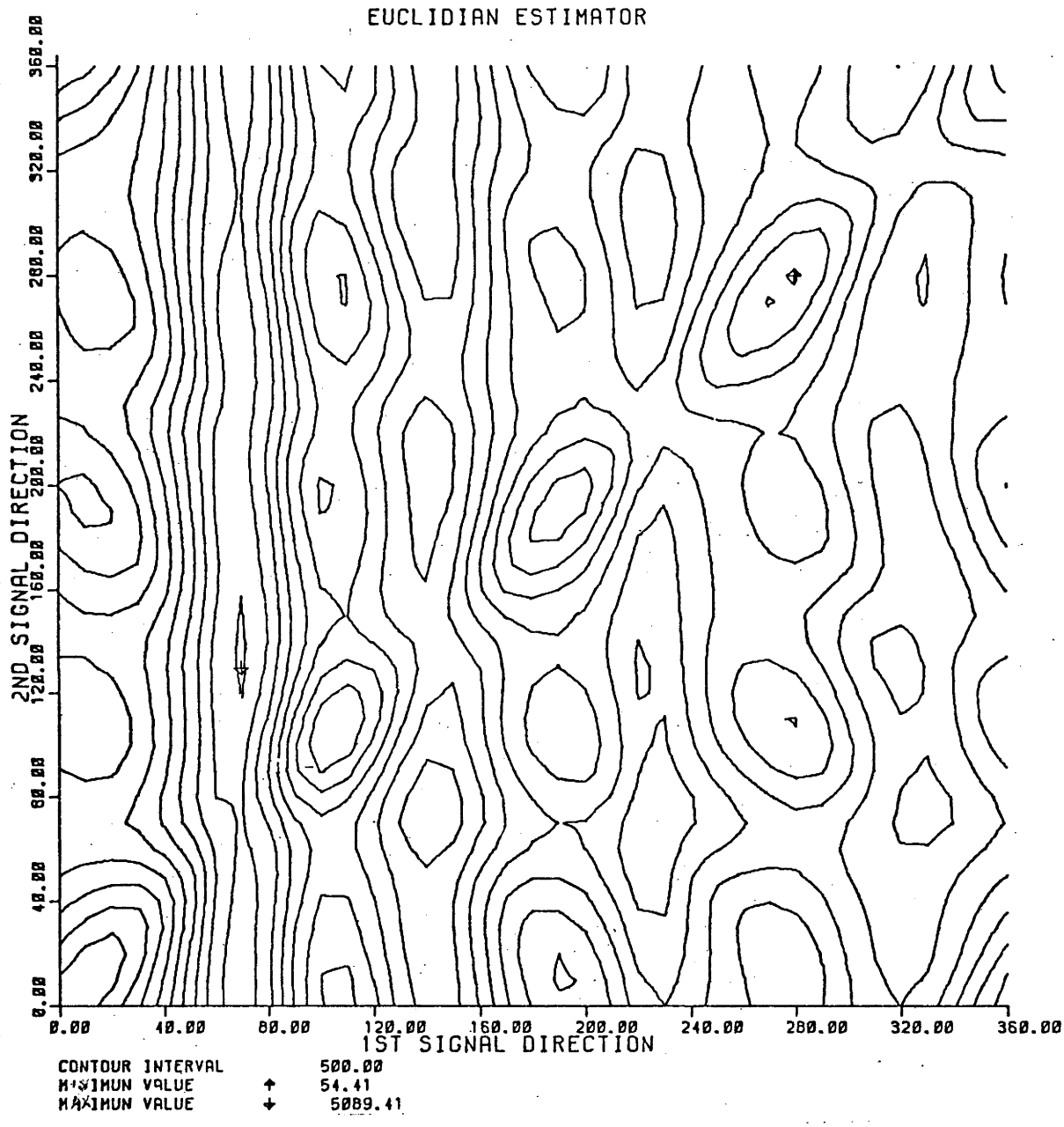

EUCLIDIAN ESTIMATOR
2ND SIGNAL DIRECTION
0.00
40.00
80.00
120.00
160.00
200.00
240.00
280.00
320.00
360.00
1ST SIGNAL DIRECTION
CONTOUR INTERVAL 500.00
MINIMUN VALUE 54.41
MAXIMUN VALUE 5089.41

Figure 4.8

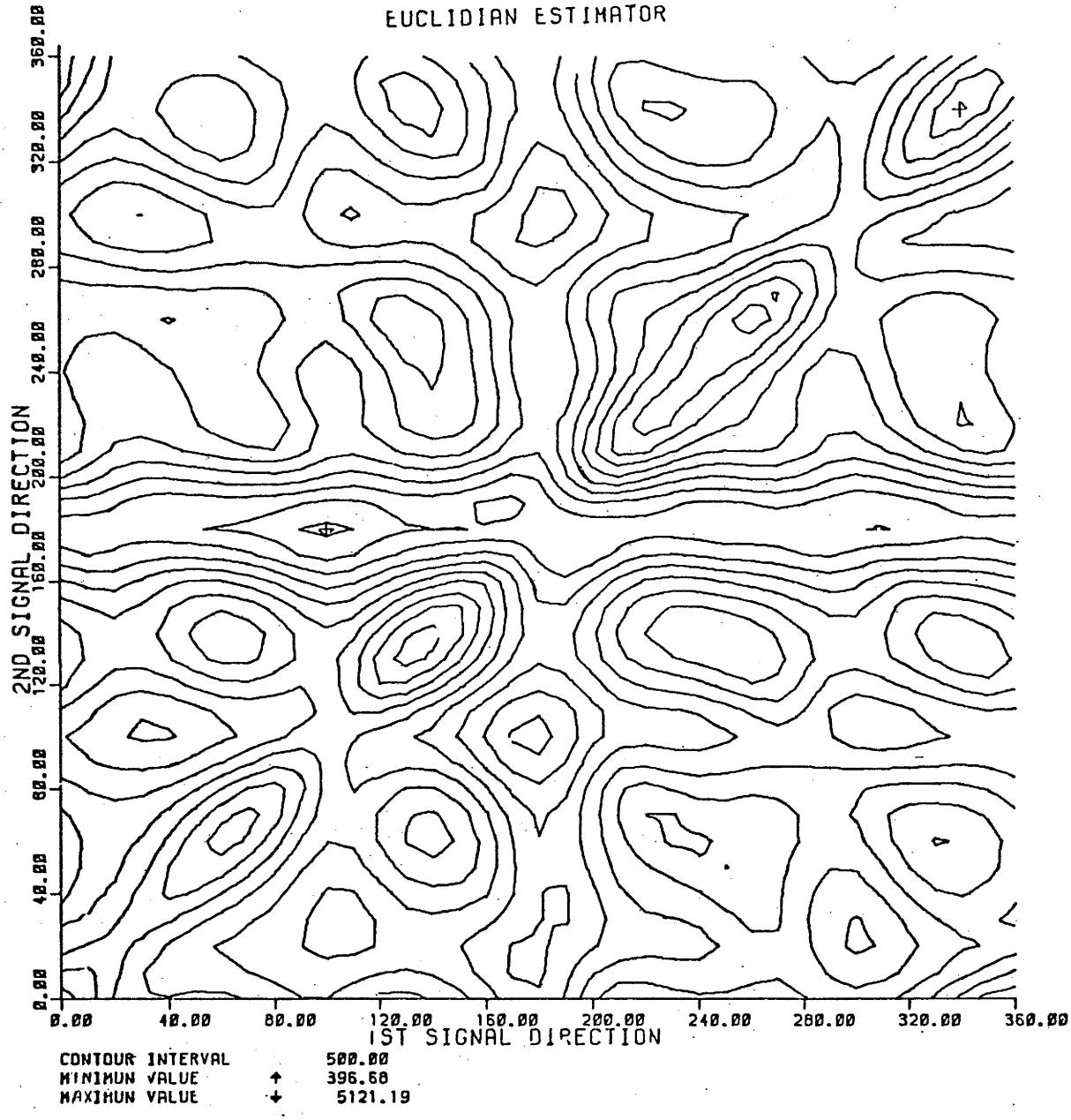

EUCLIDIAN ESTIMATOR
2ND SIGNAL DIRECTION
0.00
40.00
80.00
120.00
160.00
200.00
240.00
280.00
320.00
360.00
1ST SIGNAL DIRECTION
CONTOUR INTERVAL 500.00
MINIMUN VALUE 396.68
MAXIMUN VALUE 5121.19

Figure 4.9

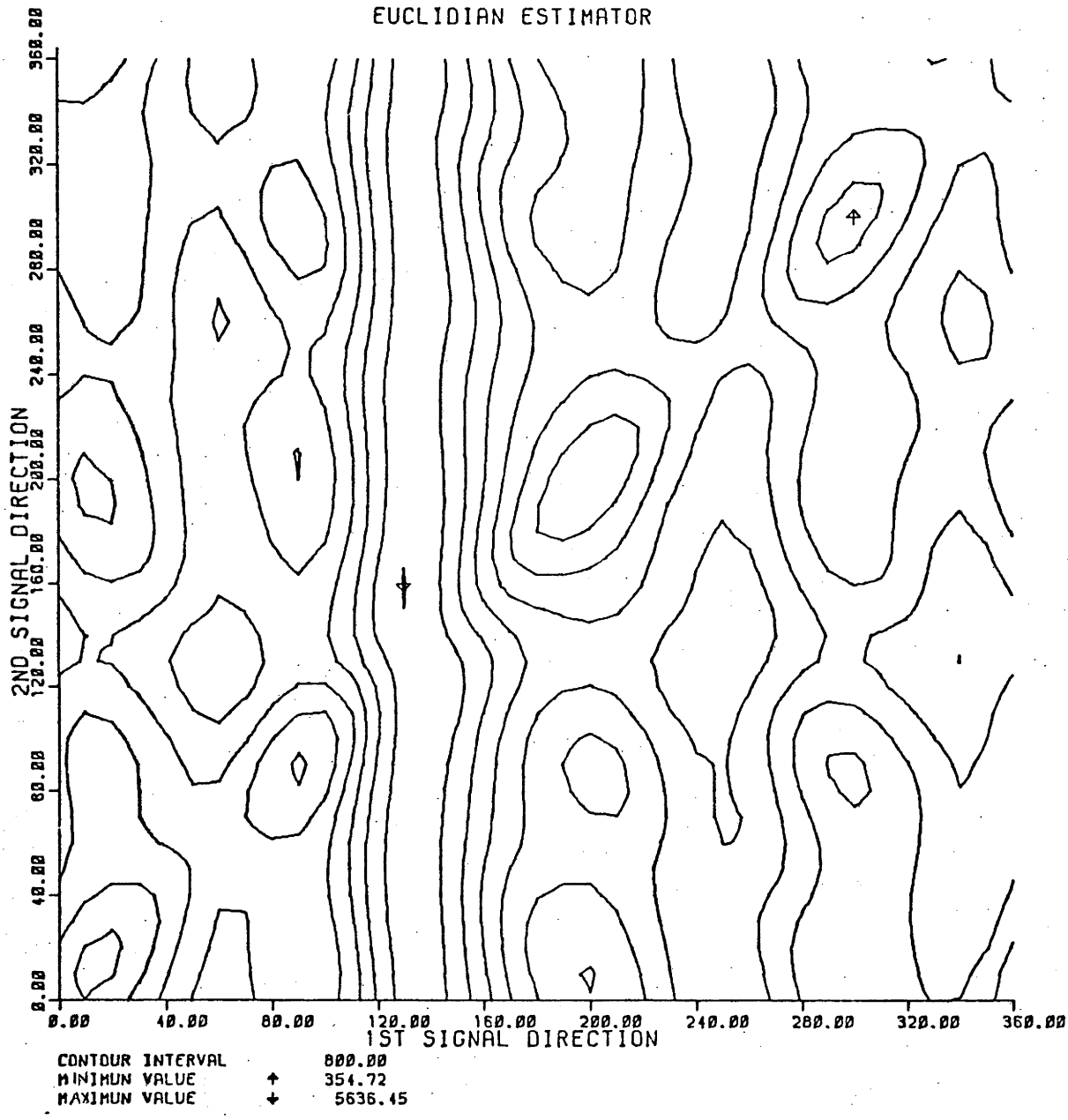

EUCLIDIAN ESTIMATOR
2ND SIGNAL DIRECTION
0.00
40.00
80.00
120.00
160.00
200.00
240.00
280.00
320.00
360.00
1ST SIGNAL DIRECTION
CONTOUR INTERVAL 800.00
MINIMUN VALUE 354.72
MAXIMUN VALUE 5636.45

Figure 4.10

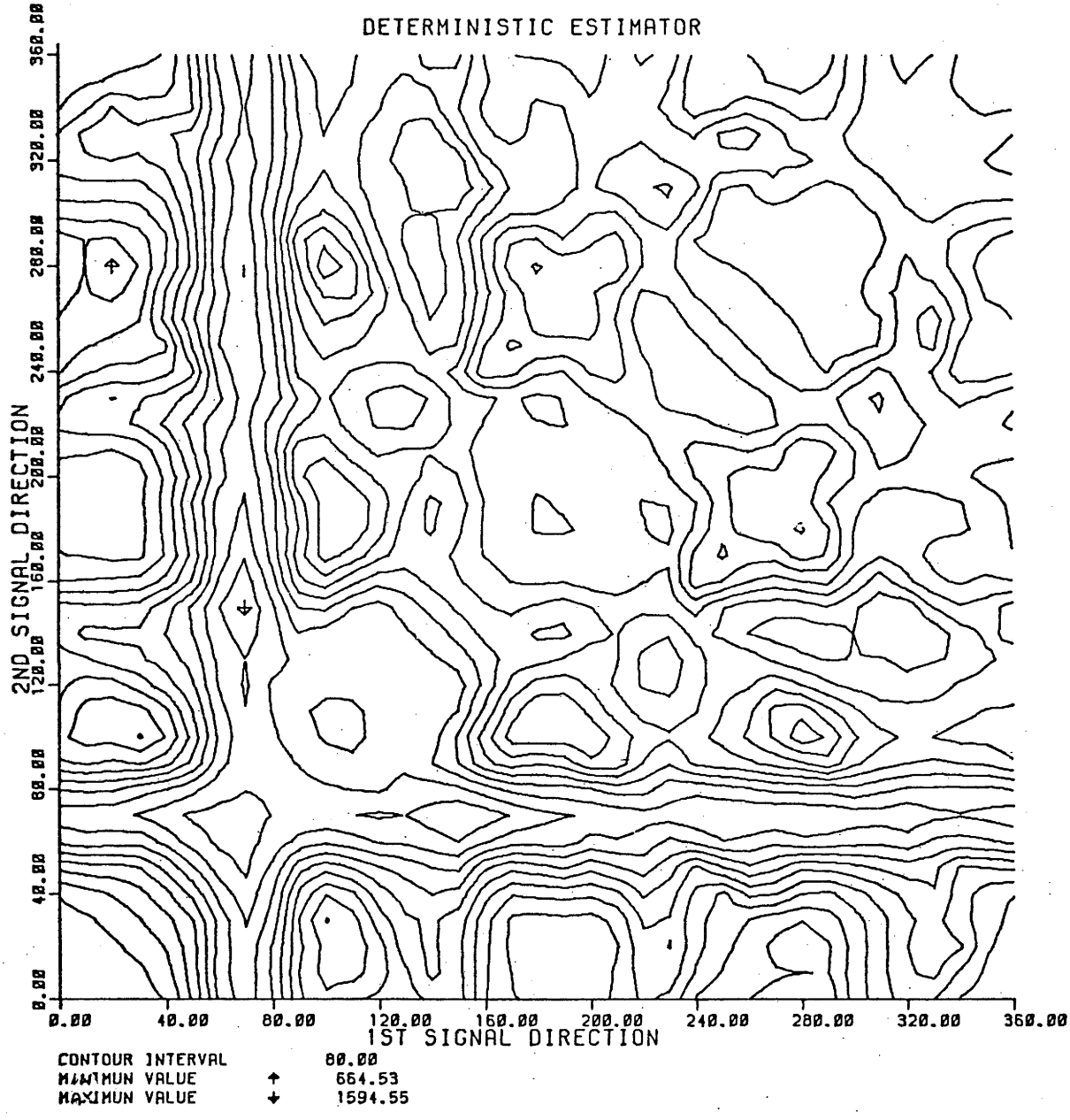

DETERMINISTIC ESTIMATOR
2ND SIGNAL DIRECTION
0.00
40.00
80.00
120.00
160.00
200.00
240.00
280.00
320.00
360.00
1ST SIGNAL DIRECTION
CONTOUR INTERVAL 80.00
MINIMUN VALUE 664.53
MAXIMUN VALUE 1594.55

Figure 4.11

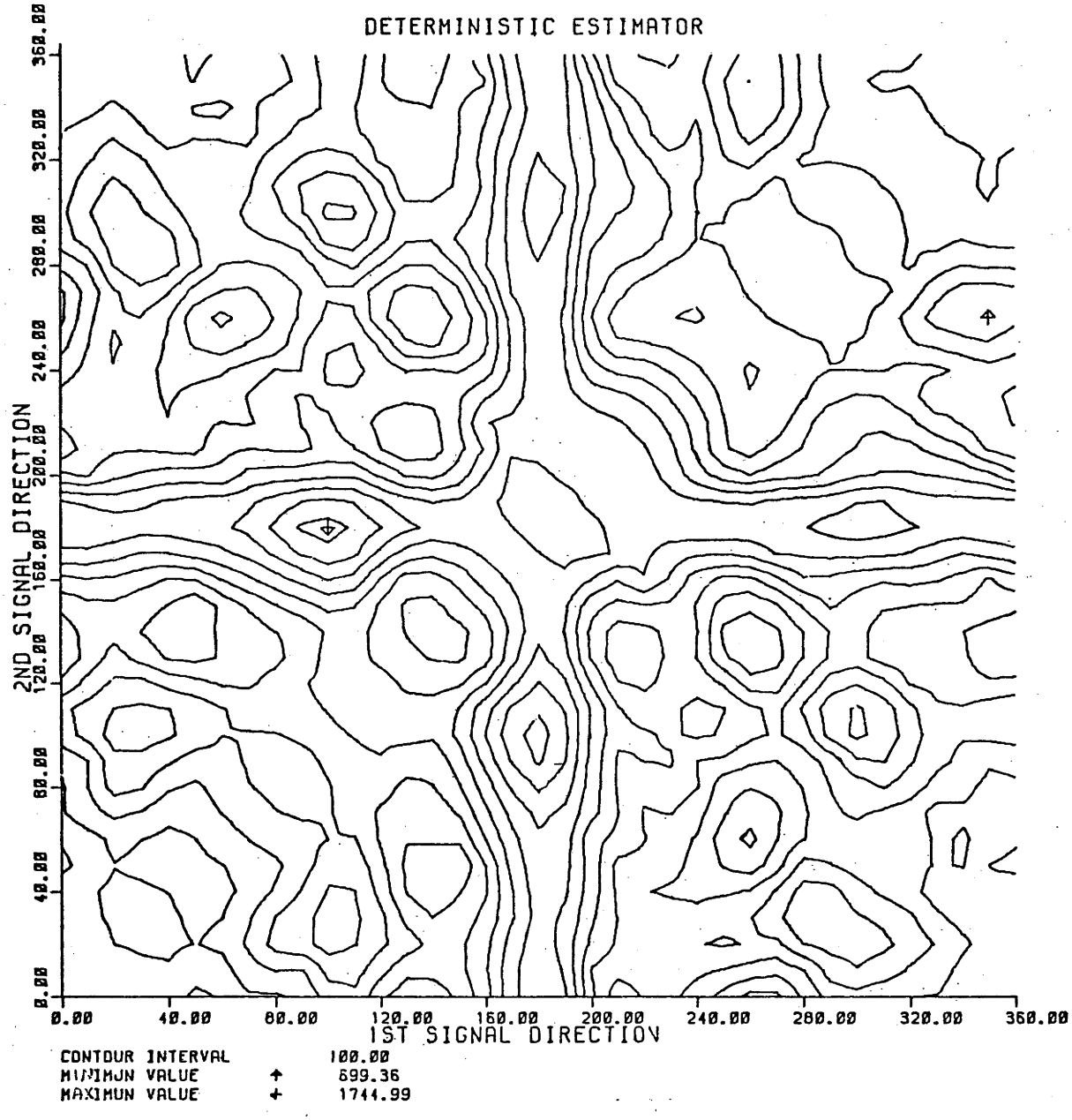

DETERMINISTIC ESTIMATOR
2ND SIGNAL DIRECTION
0.00
40.00
80.00
120.00
160.00
200.00
240.00
280.00
320.00
360.00
1ST SIGNAL DIRECTION
CONTOUR INTERVAL 100.00
MINIMUN VALUE 699.36
MAXIMUN VALUE 1744.99

Figure 4.12

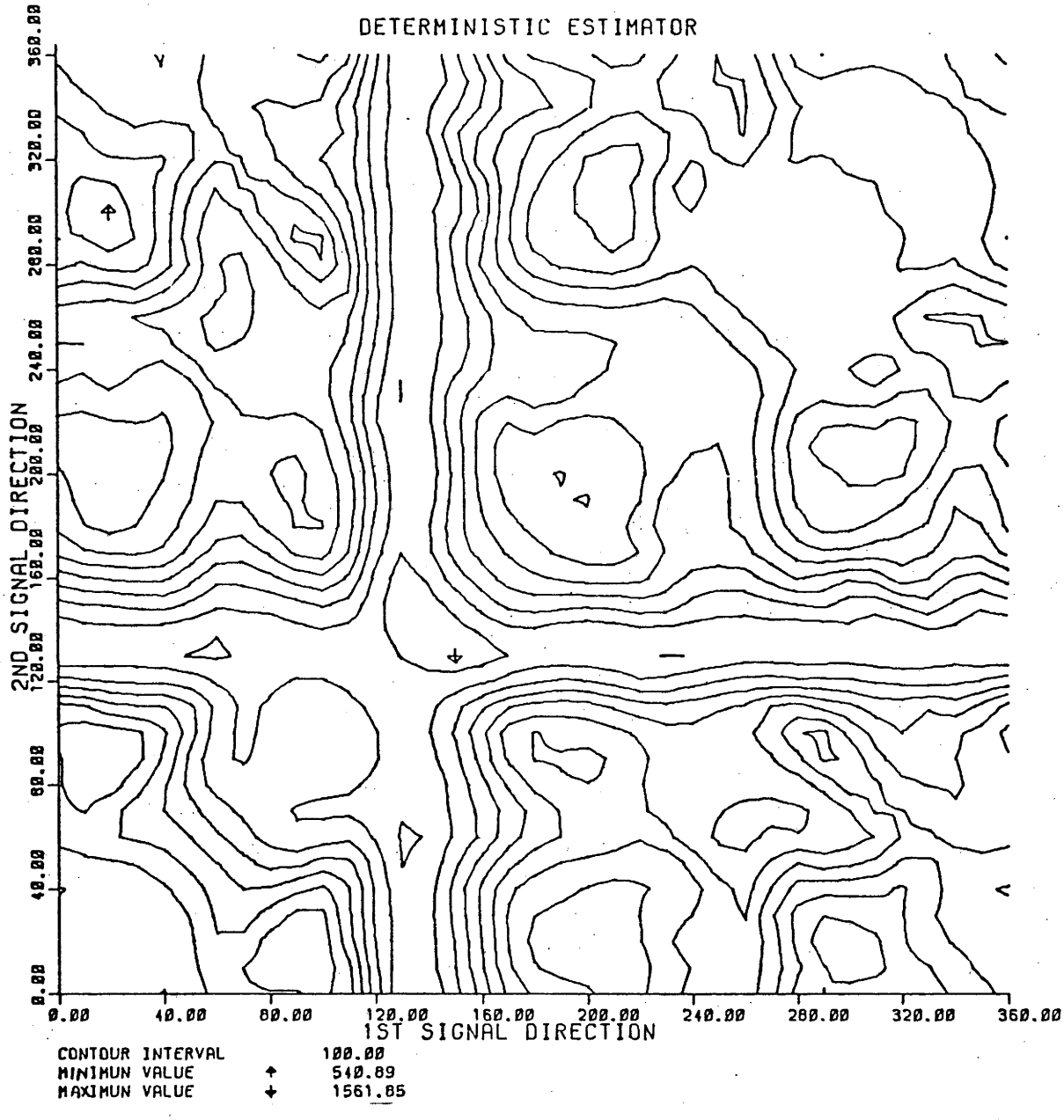

DETERMINISTIC ESTIMATOR
2ND SIGNAL DIRECTION
1ST SIGNAL DIRECTION
0.00
40.00
80.00
120.00
160.00
200.00
240.00
280.00
320.00
360.00
CONTOUR INTERVAL 100.00
MINIMUN VALUE 540.89
MAXIMUN VALUE 1561.85

Figure 4.13

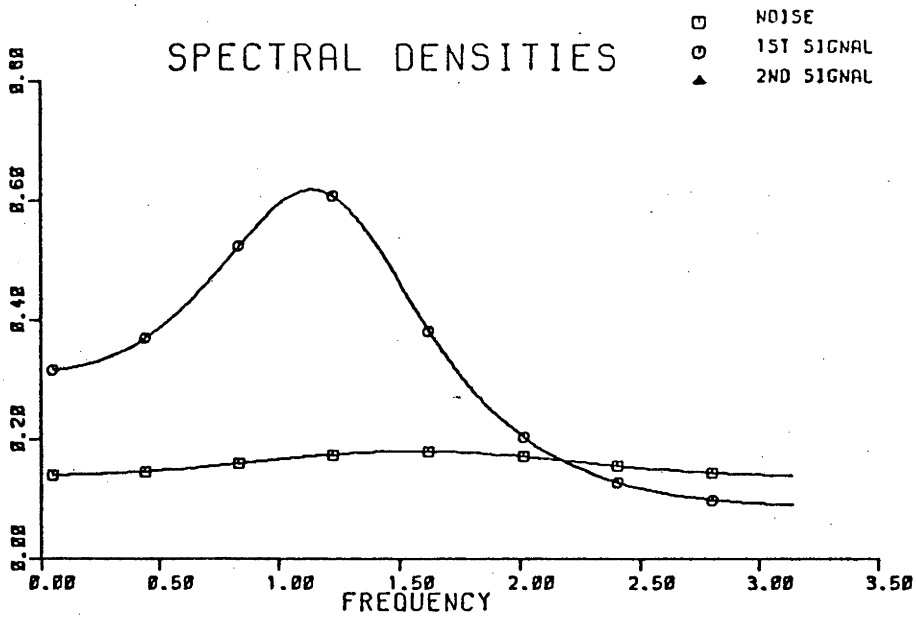
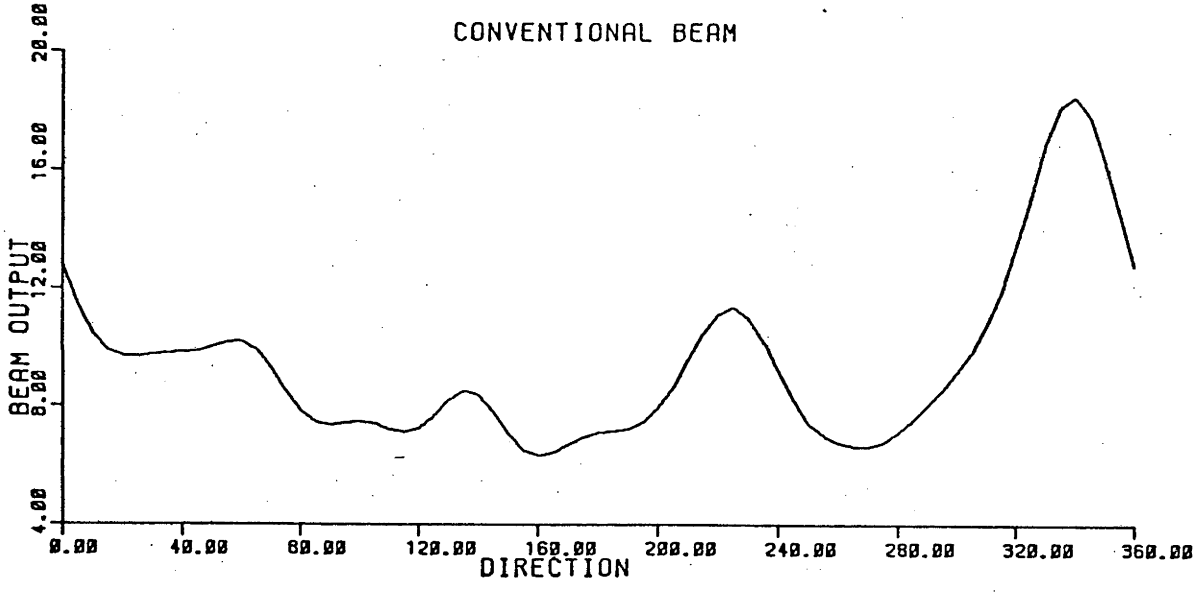


Figure 4.14



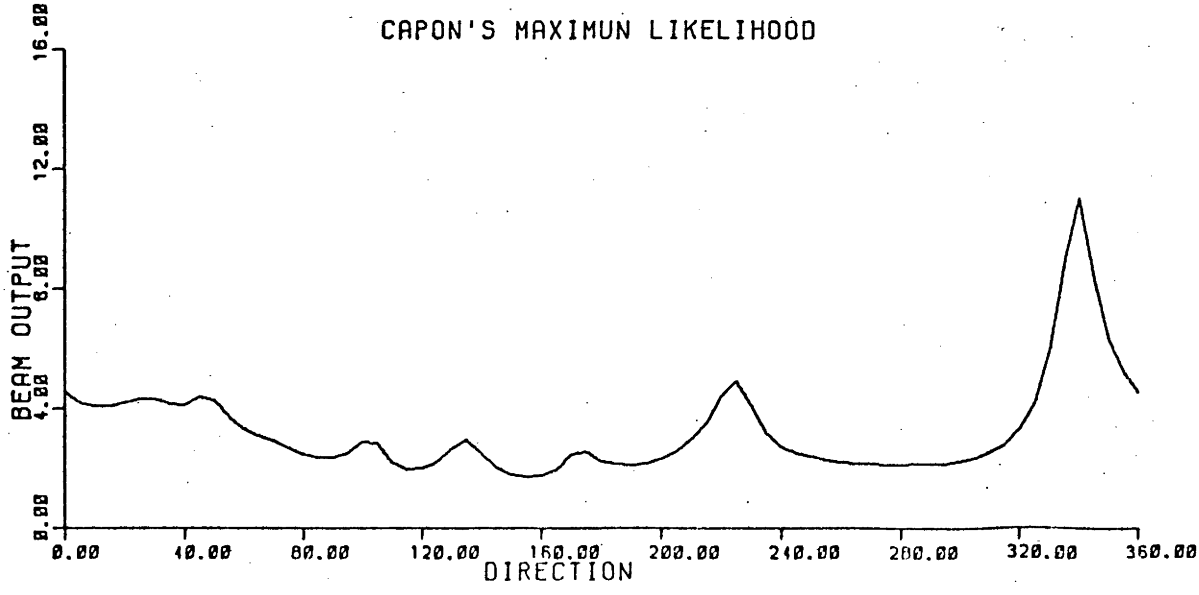


Figure 4.15

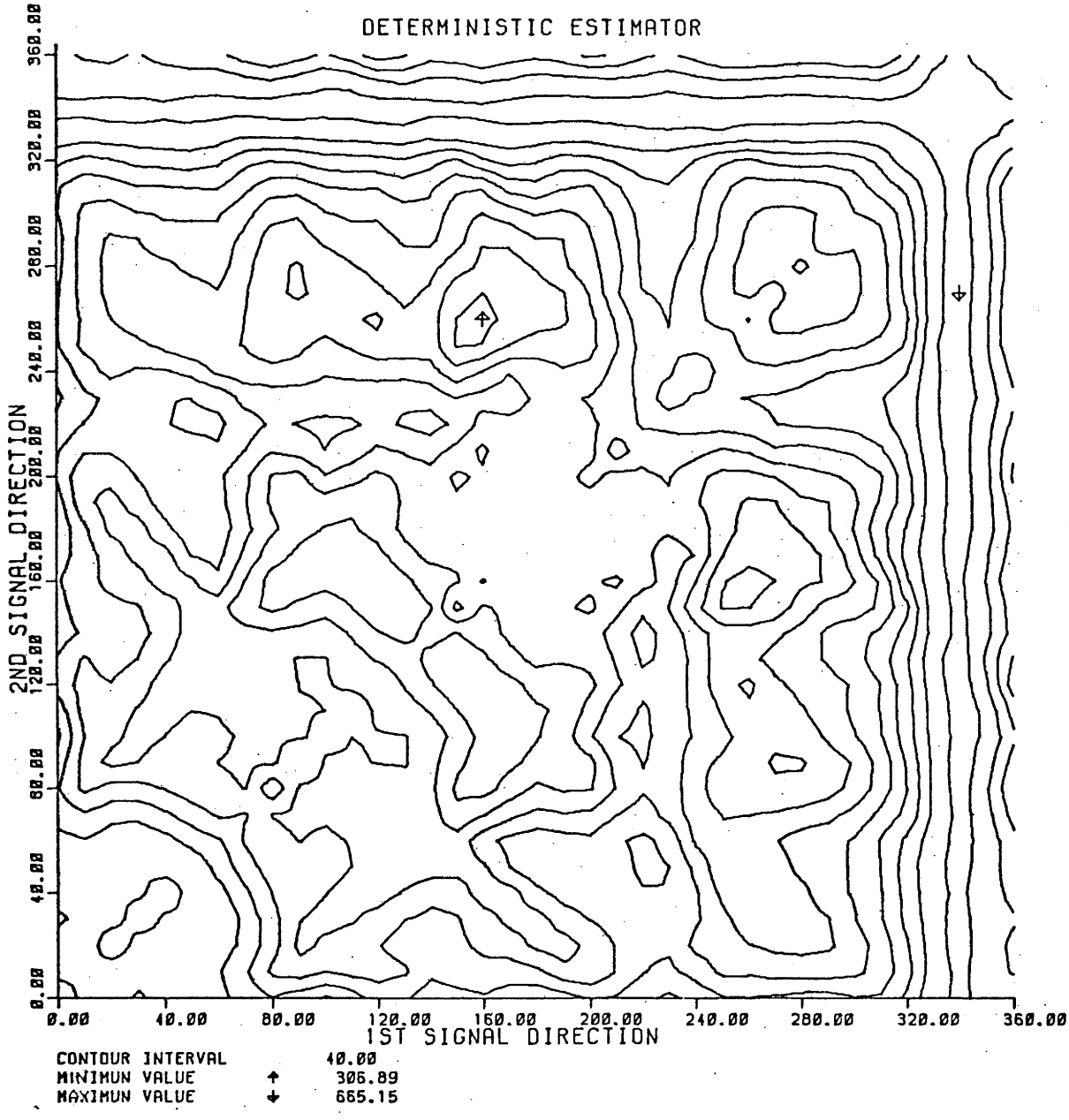

DETERMINISTIC ESTIMATOR
2ND SIGNAL DIRECTION
1ST SIGNAL DIRECTION
0.00
40.00
80.00
120.00
160.00
200.00
240.00
280.00
320.00
360.00
CONTOUR INTERVAL 40.00
MINIMUN VALUE 306.89
MAXIMUN VALUE 665.15

Figure 4.16

CHAPTER 5

NON DIRECTIONAL METHODS

5.1 MOTIVATION AND MODEL

Sometimes, only the velocity of the signals is of interest and, if so, the direction can only be regarded as a nuisance parameter. Alternatively, if we do require estimates of both velocity and direction it would be convenient if one of these could be estimated roughly using an explicit estimator, for in using the non linear estimation procedures of Chapter 3, any extra parameters greatly increase the computation involved. We shall discuss some methods of estimating velocity independently of direction. These methods will not be as efficient as those of Chapter 3, for the emphasis will be on obtaining explicit estimates suitable for preliminary analysis.

Two main approaches will be taken. The first is concerned with observing processes which have no directional properties, that is isotropic processes. The second approach develops a method proposed by Aki (1957, 1961). In both cases only waves in the plane will be considered.

We make the following assumptions:

1. At each frequency there is only one velocity of propagation; that is only one scalar wavenumber $k(\omega)$.

2. The processes are in continuous time, with spectra on $(-\infty,\infty)$.

3. The spectra and directional energy distributions are sufficiently smooth (for example, piecewise continuous) that their Fourier series are convergent almost everywhere.

We shall attempt to estimate the wavenumber function $k(\omega)$. Assumption 1 is necessary since the presence of more than one velocity makes the problem very confused.

Any stationary process defined in continuous time and the plane can be represented by

$$z(t,\xi) = \int_{-\infty}^{\infty} \int_{R^2} e^{-i\{\omega t+<\kappa,\xi>\}} \zeta'(d\omega,d\kappa)$$

where ξ is the position in space, κ is the vector wavenumber and ζ' is a random process on $R\times R^2$ with orthogonal increments (Yaglom, 1961). It will be convenient to consider this in polar coordinate form: if $\xi = r(\cos\ \theta,\sin\ \theta)$ and $\kappa = k(\cos\ \theta,\sin\ \theta)$ then

$$z(t,r,\theta) = \int_{-\infty}^{\infty} \int_{0}^{\infty} \int_{-\pi}^{\pi} e^{-i\{\omega t+rk\ \cos(\theta-\phi)\}} \zeta(d\omega,dk,d\phi)$$

Our assumption of only one velocity at each frequency implies that the measure ξ is concentrated almost surely on a curve $(\omega,k(\omega))$ and therefore

$$z(t,r,\theta) = \int_{-\infty}^{\infty} \int_{-\pi}^{\pi} e^{-i\{\omega t+rk(\omega)\cos(\theta-\phi)\}} \zeta(d\omega,d\phi) \ . \qquad [1]$$

A physical interpretation of this is that the process is a summation of uncorrelated waves at different frequencies and from different directions.

We define a *frequency-directional spectrum* F by

$$E(|\zeta(d\omega,d\phi)|^2) = F(d\omega,d\phi) \ .$$

We shall always assume that this measure is absolutely continuous, with density f. Using this a spatial covariance function may be defined thus:

$$\Gamma(\omega,r,\theta) = \int_{-\pi}^{\pi} e^{irk(\omega)\cos(\theta-\phi)} f(\omega,\phi)d\phi \quad . \qquad [2]$$

This shall be used in the following sections.

5.2 *ISOTROPIC PROCESSES*

A process is called isotropic if its probability structure is invariant under rotations. In our model this means that frequency directional spectral density $f(\omega,\phi)$ is independent of ϕ. Then

$$\Gamma(\omega,r,\theta) = \int_{-\pi}^{\pi} e^{irk(\omega)\cos(\theta-\phi)} f(\omega)d\phi$$

$$= J_0(rk(\omega))\, f(\omega)$$

where J_0 is the zero order Bessel function of the first kind. If this process is observed at points $\xi(1),\xi(2),\ldots\xi(p)$ in the plane to give the vector process $y(t)$, then $y(t)$ will have spectral density matrix

$$F_y(\omega) = f(\omega)\, \mathcal{J}\,(k(\omega)),$$

where

$$(\mathcal{J}(k(\omega)))_{ab} = J_0\,(\,||\xi(a) - \xi(b)||\, k(\omega)) \quad .$$

We may assume that the function $k(\omega)$ is parametrized by a vector τ, the true value of which is τ_0. Then the maximum likelihood estimating function is

$$q_N(\tau) = N^{-1} \sum_{\omega\varepsilon B} -\log\det(f(\omega)\,\mathcal{J}\,(k(\omega,\tau)))-f(\omega)^{-1}\, Y(\omega)^{*}\,\mathcal{J}\,(k(\omega,\tau))^{-1}Y(\omega) \qquad [32]$$

where, following the previous conventions, $Y(\omega)$ is the vector of Fourier transforms of N points sampled at unit intervals and the sum is over the fundamental frequencies in the set B. The following identification

condition is required.

CONDITION 5.1

There is no allowable $\tau \neq \tau_0$ such that for some r and for all ω, $J_0(rk(\omega,\tau)) = J_0(rk(\omega,\tau_0))$. //

Observe that if $k(\omega,\tau)$ is restricted to polynomials in ω, with the coefficients being the components of τ, then this condition is satisfied. (This is easily shown by considering the derivatives of $J_0(rk(\omega,\tau))$ at $\omega = 0$.)

With Condition 5.1 holding we have both consistency and a central limit result.

THEOREM 5.2

If Conditions 2.1 and 5.1 are satisfied then the estimate $\hat{\tau}_N$ obtained by maximizing [3] is consistent.

PROOF

The proof follows that of Theorem 2.4. //

THEOREM 5.3

If the process is Gaussian and Conditions 2.1 and 5.1 are satisfied, then the variance matrix of $N^{\frac{1}{2}}(\hat{\tau}_N - \tau_0)$ tends to where

$$J(\Sigma^{-1})_{j,k} = J(2\pi)^{-1} \int_B \text{trace}\Big\{ J(k(\omega,\tau_0))^{-1} J(k(\omega,\tau_0))^{(j)} J(k(\omega,\tau_0))^{-1} J(k(\omega,\tau_0))^{(k)} \Big\} d\omega .$$

PROOF

This follows from Lemmas 3.4 and 3.6. The assumption of Gaussianity eliminates the fourth cumulant term. //

The assumption of Gaussianity is reasonable in this situation

since the observed process is a combination or average of waves from all directions. However in practice this result is not as useful as it may at first appear, for in the next section we shall see that it is possible to construct an array such that the spectral density matrix F_y is almost singular. Then the asymptotic variances are extremely small and consequently even a small amount of independent noise at the receivers will change the variance greatly. Of course in that case we ought to use the model

$$F_y(\omega) = f_x(\omega)I_p + f(\omega)\, J\,(k(\omega)).$$

However, unless the noise spectrum is known, the maximum likelihood approach with this model is too cumbersome.

Before leaving this topic, consider the optimal estimation of $f(\omega)$. By differentiating [3] with respect to $f(\omega)$ we obtain an estimate

$$\hat{f}(\omega) = Y(\omega)^{*}\, J\,(k(\omega,\tau))^{-1}\, Y(\omega)/p \qquad [4]$$

If p is large, this in itself may be satisfactory but usually an average of [4] over a narrow band of fundamental frequencies should be used.

5.3 THE METHOD OF AKI

Aki (1957, 1961) suggested that the use of a special array consisting of a circle of equally spaced receivers and one receiver at the centre could be used to simplify estimation of velocity. The idea is very simple. Observe that if the spatial covariance function is averaged over all directions, we obtain

$$\gamma(\omega,r) = (2\pi)^{-1} \int_{-\pi}^{\pi} \Gamma(\omega,r,\theta)\, d\theta$$

$$= \int_{-\pi}^{\pi} (2\pi)^{-1} \int_{-\pi}^{\pi} e^{irk(\omega)\cos(\theta-\phi)}\, d\theta f(\omega,\phi) d\phi$$

$$= J_0(rk(\omega)) \int_{-\pi}^{\pi} f(\omega,\phi) d\phi$$

$$= J_0(rk(\omega))\ f_0(\omega)$$

where $f_0(\omega) = \int_{-\pi}^{\pi} f(\omega,\phi) d\phi$. The average covariance function is no longer dependent upon the directional properties of the process. The spectrum of the process observed at one receiver is $f_0(\omega)$ so that the averaged complex coherency will be $J_0(rk(\omega))$.

Aki suggested obtaining the complex coherency between each receiver on the circle and the centre receiver and then averaging this to obtain a quantity say $\hat{\rho}(\omega,r)$. The wavenumber estimate would then be given by

$$J_0(r\hat{k}(\omega)) = \text{Real}(\hat{\rho}(\omega,r)) \qquad [5]$$

By using complex coherencies throughout, Aki avoided the problems of the receivers having different responses and was able to implement the technique in 1957, before the use of digital computers.

Ideally in practice the value of r (that is, the radius of the array) is kept fixed; this limits the useful range of the array. From [5] we have (approximately)

$$\text{Var}\,(k(\omega)) = \text{Var}(\hat{\rho}(\omega,r))(rJ_1(rk(\omega))^{-2}\ .$$

For small wavenumbers, any errors or uncertainties in the estimation of $\hat{\rho}(\omega,r)$ are going to be greatly magnified.

If several different values of r are used, it in theory should

be possible to treat the situation where more than one velocity is significant. If the two wavenumber functions are $k_1(\omega)$ and $k_2(\omega)$, and the energy is divided between them in the ratio $\alpha(\omega)$ to $1 - \alpha(\omega)$, then

$$\gamma(\omega,r) = (\alpha(\omega)J_0(rk_1(\omega)) + (1-\alpha(\omega))J_0(rk_2(\omega))f_0(\omega).$$

Three different values of r would be required to make it possible to estimate k_1, k_2 and α. This has been attempted by Asten (1976). However, we shall not pursue this further.

Rather we shall investigate further methods of estimating one wavenumber function, using a fixed value of r. We define the centre process by

$$x_0(t) = z(t,0,0)$$

and the circle process by

$$y_0(t) = (2\pi)^{-1} \int_{-\pi}^{\pi} z(t,r,\theta)d\theta$$

Expressed in terms of the spectral process, these are

$$x_0(t) = \int_{-\infty}^{\infty} e^{-i\omega t} \int_{-\pi}^{\pi} \zeta(d\omega,d\phi)$$

$$y_0(t) = \int_{-\infty}^{\infty} \int_{-\pi}^{\pi} (2\pi)^{-1} \int_{-\pi}^{\pi} e^{-i\{\omega t+rk(\omega)\cos(\theta-\phi)} \, d\theta\zeta(d\omega,d\phi)$$

$$= \int_{-\infty}^{\infty} e^{-i\omega t} J_0(rk(\omega)) \int_{-\pi}^{\pi} \zeta(d\omega,d\phi)$$

If we let $\int_{-\pi}^{\pi} \zeta(\omega,d\phi) = \zeta_0(\omega)$, then

$$x_0(t) = \int_{-\infty}^{\infty} e^{-i\omega t} \, d\zeta_0(\omega)$$

$$y_0(t) = \int_{-\infty}^{\infty} e^{-i\omega t} \, J_0(rk(\omega)) \, d\zeta_0(\omega) \quad .$$

The circle process can then be regarded as a linearly filtered form of the centre process, with gain function $J_0(rk(\omega))$. Except when $J_0(rk(\omega)) = 0$, the coherence between the two processes will be one. Furthermore, the introduction of either noise or another significant velocity will reduce this coherence to a value less than one.

There are two practical outcomes of this. First by estimating the coherence between the circle and centre processes a check can be made on the validity of the assumptions necessary to apply these techniques. Second, if the gain function is estimated by standard procedures (for example, Brillinger, 1975), it can be used as an alternative method of estimating the wavenumber function.

If this circular array is used to observe an isotropic process as in the previous section, it can be seen that an almost singular spectral density matrix is the result.

5.4 FOURIER BESSEL ANALYSIS USING A CIRCULAR ARRAY

The terms $f_0(\omega)$ and $\zeta_0(\omega)$ may be regarded as the zero order terms of the Fourier series of $f(\omega,\phi)$ and $\xi(\omega,d\phi)$ respectively. The higher order terms will be defined by

$$f_m(\omega) = \int_{\pi}^{\pi} e^{-im\phi} f(\omega,\phi)d\phi \ ,$$

$$\text{and} \quad \zeta_m(\omega) = \int_{\pi}^{\pi} e^{-im\phi} \xi(\omega,d\phi) \ .$$

Then

$$f(\omega,\phi) = (2\pi)^{-1} \sum_{m=\infty}^{\infty} e^{im\phi} f_m(\omega)$$

$$\xi(\omega,d\phi) = (2\pi)^{-1} \sum_{m=\infty}^{\infty} e^{im\phi} \zeta_m(\omega)d\phi$$

and using the representation [1]

$$z(t,r,\theta) = \int_{-\infty}^{\infty} \sum_{m=-\infty}^{\infty} e^{-i\omega t}(2\pi)^{-1} \int_{-\pi}^{\pi} e^{-i\{rk(\omega)\cos(\theta-\phi)-m\phi\}} d\phi d\zeta_m(\omega)$$

$$= \int_{-\infty}^{\infty} \sum_{m=-\infty}^{\infty} e^{-i\{\omega t-m\theta\}} J_m(rk(\omega)) d\zeta_m(\omega)$$

Now we define what we shall call the m^{th} order Fourier process

$$y_m(t) = (2\pi)^{-1} \int_{-\pi}^{\pi} e^{-im\theta} z(t,r,\theta) d\theta$$

$$= \int_{-\infty}^{\infty} e^{-i\omega t} J_m(rk(\omega)) d\zeta_m(\omega)$$

Hence the Fourier processes provide a method of isolating the Fourier components of the spectral process. Before examining this further a slight generalization is made in order to cover the case where the energy arriving from different directions is not uncorrelated. It would be useful to be able to consider the case where waves from a single source are scattered by irregularities in the medium and then arrive at the observer from several directions simultaneously. This can be introduced into the model by defining a scattering function $W(\omega,\psi)$ and a new spectral generator ζ^0 by

$$\zeta^0(\omega,\phi) = \int_{-\pi}^{\pi} W(\omega,\phi-\psi)\ \zeta(\omega,d\psi) \ . \qquad [6]$$

Assume that $W(\omega,\psi)$ has a Fourier series with respect to the parameter ψ, that is

$$W(\omega,\psi) = (2\pi)^{-1} \sum_{m=-\infty}^{\infty} e^{im\psi} C_m(\omega) \ .$$

Since [6] is a convolution we have

$$\zeta^0(\omega,d\phi) = (2\pi)^{-1} \sum_{m=-\infty}^{\infty} e^{im\phi}\ C_m(\omega)\ \zeta_m(\omega)d\phi \ ,$$

$$z(t,r,\theta) = \int_{-\infty}^{\infty} \sum_{m=-\infty}^{\infty} e^{-i\{\omega t - m\theta\}} J_m(rk(\omega))\, C_m(\omega)\, d\zeta_m(\omega)$$

and $$y_m(t) = \int_{-\infty}^{\infty} e^{-i\omega t} J_m(rk(\omega))\, C_m(\omega) d\zeta_m(\omega) \quad .$$

Now we can proceed to examine the properties of these Fourier processes. Note that

$$E\ (\zeta_\ell(d\omega)\ \overline{\zeta_m(d\omega)})$$

$$= \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} e^{-i\{\ell\phi - m\phi'\}} E(\zeta(d\omega d\phi)\ \overline{\zeta(d\omega d\phi')})$$

$$= \int_{-\pi}^{\pi} e^{-i(\ell-m)} f(\omega,\phi) d\phi\, d\omega$$

$$= f_{\ell-m}(\omega)\, d\omega \quad .$$

As before the centre process is

$$x_0(t) = \int_{-\infty}^{\infty} e^{-i\omega t} C_0(\omega)\, d\zeta_0(\omega).$$

Then the correlation relations between the Fourier processes and the centre process are

$$E(y_\ell(S)\ \overline{y_m(S+t)}) = \int_{-\infty}^{\infty} e^{i\omega t} J_\ell(rk(\omega)) J_m(rk(\omega)) C_\ell(\omega)\ \overline{C_m(\omega)}\ f_{\ell-m}(\omega) d\omega$$

$$E(x_0(S)\ \overline{y_m(S+t)}) = \int_{-\infty}^{\infty} e^{i\omega t} J_m(rk(\omega)) C_0(\omega)\ \overline{C_m(\omega)}\ f_{-m}(\omega) d\omega$$

$$E(x_0(S)\ \overline{x\ (S+t)}) = \int_{-\infty}^{\infty} e^{i\omega t}\ |C_0(\omega)|^2\ f_0(\omega) d\omega \ .$$

As in the previous section, it seems most useful to compare the centre process with each of the Fourier processes. The bivariate process $(x\ (t), y_m(t))'$ has a spectral density matrix $G(\omega)$ given by

$$G(\omega) = \begin{pmatrix} |C_0(\omega)|^2 f_0(\omega) & J_m(rk(\omega))C_0(\omega)\overline{C_m(\omega)}f_{-m}(\omega) \\ J_m(rk(\omega))\,\overline{C_0(\omega)}\,C_m(\omega)f_m(\omega) & J_m(rk(\omega))^2|C_m(\omega)|^2 f_0(\omega) \end{pmatrix}$$

In the next section we shall consider practical uses of observing this bivariate process.

5.5 *PRACTICAL CONSIDERATIONS*

In practice an array with p receivers equally spaced on a circle of radius r, and one receiver at the centre would be used. While the centre process is directly observable, the Fourier processes are not. Instead of the integral definition, a finite sum is used; that is

$$\tilde{y}_m(t) = p^{-1} \sum_{a=1}^{p} z(t,r,2\pi ap^{-1})e^{-im2\pi ap^{-1}}$$

This need be defined only for m = -[p/2],...[p/2].
A form of directional aliasing will occur and

$$\tilde{y}_m(t) = \sum_{k=-\infty}^{\infty} y_{m+kp}(t)$$

$$= \int_{-\infty}^{\infty} \sum_k e^{-i\omega t} J_{m+pk}(rk(\omega))\, C_{m+pk}(\omega)\, d\zeta_{m+pk}(\omega) .$$

However, $\lim_{n\to\infty} J_n(x) = 0$ and this convergence is very rapid so that this aliasing will not in general be significant for values of p not too small.

Considering the bivariate process $(x_0(t),\tilde{y}_m(t))'$ we may form an estimate $\hat{G}(\omega)$ of the spectral density matrix and define the following functions: the m^{th} order Fourier coherence

$$\sigma_m(\omega) = |\hat{G}_{12}(\omega)|\,(\hat{G}_{11}(\omega)\,\hat{G}_{22}(\omega))^{-\frac{1}{2}}$$

the m^{th} order Fourier phase

$$\sigma_m(\omega) = \arg \hat{G}_{12}(\omega)$$

and the m^{th} order scatter function

$$S_m(\omega) = \hat{G}_{22}(\omega)/\hat{G}_{11}(\omega) .$$

Observe that asymptotically $\sigma_m(\omega)$ approaches $|f_m(\omega)|/f_0(\omega)$, $\phi_m(\omega)$ approaches $\arg(C_o(\omega)\ \overline{C_m(\omega)\ f_{-m}(\omega)})$ and S_m approaches $J_m(rk(\omega))^2\ |C_m(\omega)/C_0(\omega)|^2$.

We may always assume that $C_0(\omega) = 1$ since the unscattered process is never observed. However, it does not seem possible to estimate the phases of $f_m(\omega)$ and $C_m(\omega)$ separately--only their absolute values can be estimated.

By these procedures we obtain a rough indication of the nature of the observed waves. If $C_m(\omega)$ is small for $m = 0$, a high degree of scatter is indicated, $|f_m(\omega)|/f_0(\omega)$ being small from $m \neq 0$ indicates a reasonable number of sources present. If this ratio is close to one we could reasonably expect that only one signal is present.

These results are rather inconclusive but they are useful in providing a preliminary analysis of the data simply by the evaluation of explicitly defined functions.

CHAPTER 6

AN APPLICATION TO SEISMIC DATA

6.1 DESCRIPTION OF THE DATA

This chapter contains some results of applying the methods of Chapters 3 and 5 to some seismic data.

The data was obtained by Michael Asten of Macquarie University, New South Wales, and contains records of microseismic events in the Sydney basin area. The events are of very low energy, some being generated by local traffic, and are very non stationary.

Arrays comprised up to seven short period Willmore vertical seismometers with a resonance at 0.7 Hz and a useful range of 0.1 Hz to 10 Hz. The outputs of the seismometers were amplified, filtered to remove energy above 10 Hz and relayed by cable to a truck mounted processor. There they were again filtered, converted to digital form and stored on nine-track magnetic tape using an Interdata 70 minicomputer as a buffer. A sampling rate of 200 samples per second for each seismometer gave a Nyquist frequency of 100 Hz. This rather high rate and the prefiltering were used to overcome interference from the 50 Hz local electricity supply.

The purpose of using this array was to determine velocities and dispersion of surface waves, using microseisms--not controlled explosions as is usual in exploration seismology. Vertical seismometers

are primarily sensitive to Rayleigh waves and if the velocity of these waves is known as a function of frequency, certain inferences can be made about the geological structure beneath the array. The problem is compounded slightly by the fact that Rayleigh waves can propagate in different modes, these having different velocities.

Before applying the methods of this thesis, Fourier transforms had to be calculated. Due to the great variation of energy levels at adjacent frequencies, a 100% cosine taper was used to reduce leakage. Prewhitening with a time domain filter followed by using a less severe taper would have been preferable but this would have required the design of a different filter for each set of data. After fast Fourier transforming in blocks of 4096 points, the transforms were corrected for the phase and amplitude characteristics of each receiver.

A thirteen-second record from five receivers is shown in Figure 6.1.

6.2 APPLYING THE DETERMINISTIC ESTIMATOR

In the absence of any useful finite parameter model for the velocity, it was decided to examine the data in successive frequency bands, the bands being sufficiently narrow that the assumption of constant velocity within bands was approximately correct. Within a band each signal was parametrized by its two slowness coordinates in the plane.

The programms used are reproduced in full in the last section of this chapter. They are designed for interactive use and produce contour plots in the two slowness coordinates using a conventional single beam method, Capon's method or the deterministic estimator. The technique of Section 3.6 was used for the deterministic estimator,

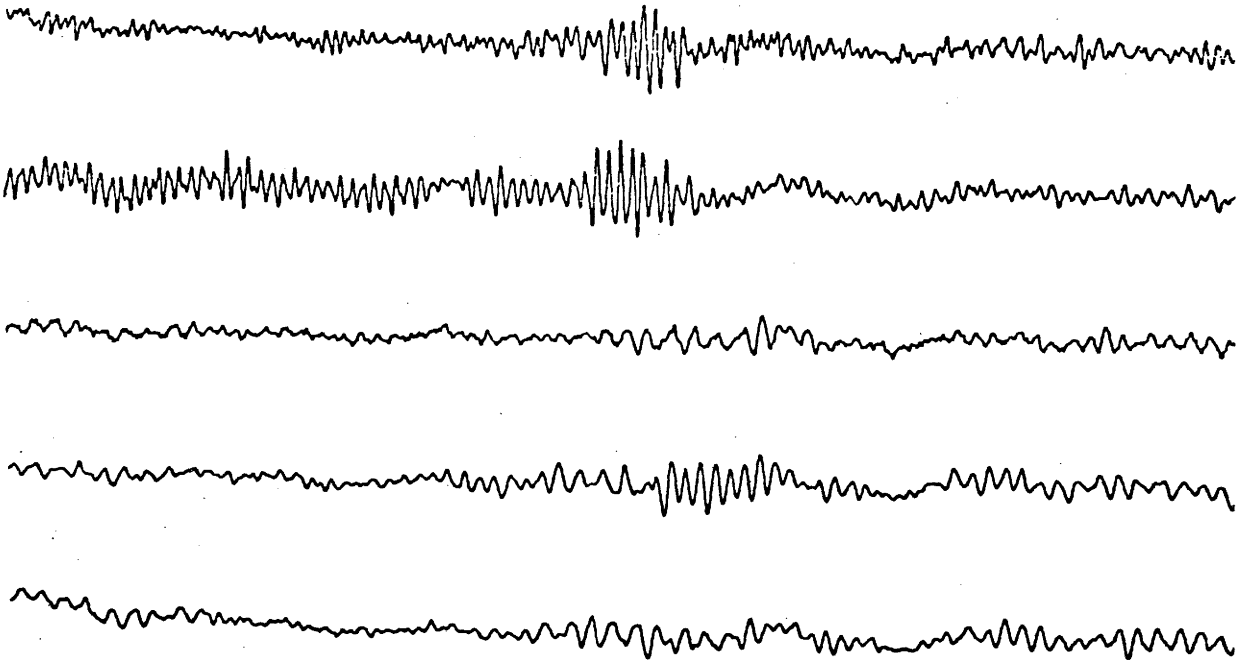


Figure 6.1

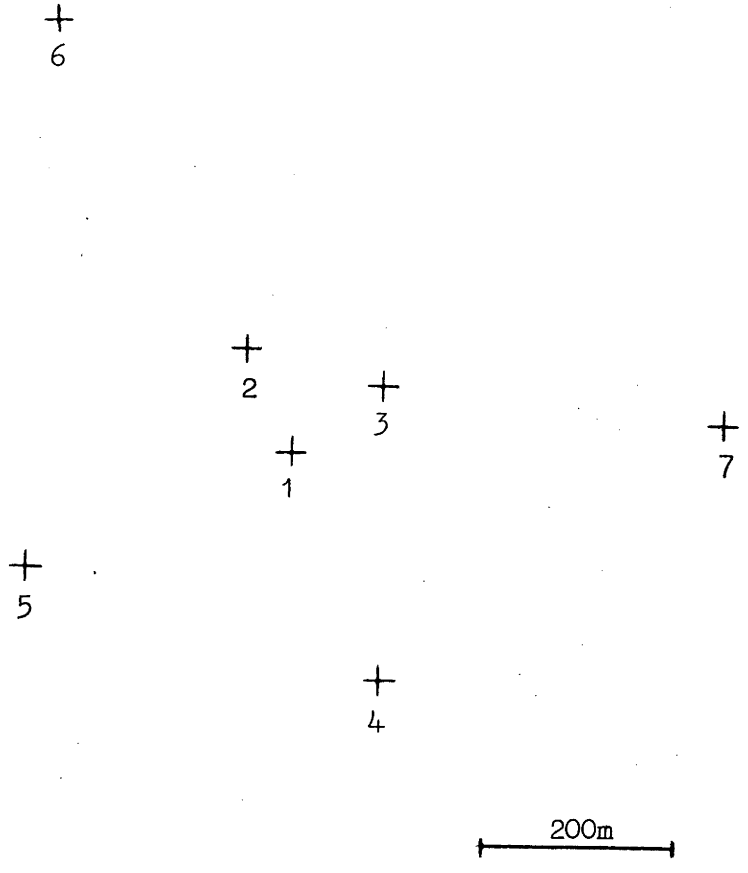


Figure 6.2

keeping the parameters of all but one signal fixed at any one time.

In Figures 6.3 to 6.8 we give an example of the use of this program. A set of data from a seven-receiver array was used (Asten's file T6F1), the receivers having the positions as shown in Figure 6.2.

For this example the band 1.0 Hz to 1.6 Hz was chosen. This contained thirteen fundamental frequencies. Velocities were expected to be greater than 1 Km/sec so the slowness parameters were restricted to the range [-1.0,1.0].

Figure 6.3 shows the result of Capon's method. It exhibits a rather sharp peak with a maximum (denoted by the asterisk) at (0.25,-0.15). This corresponds to a signal arriving from the south east with a velocity of 3.4 Km/sec. Furthermore the asymmetrical shape of the peak suggests that more signals may be present--one from the north east and a faint one from the south west. Velocities cannot be assigned to these and it is not clear whether or not they are "real".

The conventional single beam analysis (Figure 6.4) has a similar maximum but it has a very broad peak which gives no indication of other signals. The height of the peak indicates that one signal could explain 75% of the energy observed in that band whereas no such interpretation can be given to the height of the peak produced by Capon's method.

This one signal provides a starting point for the deterministic estimator. If we fix one signal with parameters (0.4,-0.2) and look for a second signal we obtain Figure 6.5, which shows a signal present in the north east quadrant, parameters (0.15,0.25). To carry out the iterative method we fix the second signal parameters at (0.15,0.25) and attempt to revise the estimates for the first signal, and so on until the estimates stabilize. Figures 6.6 and 6.7 show the first two steps

of this. If we continue we quickly obtain

1st signal	(0.3087,-0.1612)	Velocity 2.87 Km/sec
2nd signal	(0.1038, 0.1869)	Velocity 4.68 Km/sec .

At this stage we may proceed to consider a third signal. Fixing two signals with the above parameters, we obtain Figure 6.8 as the estimating function for the third signal. It clearly indicates a signal in the south west and if again we iterate this procedure we obtain

1st signal	(0.3308,-0.1444)	Velocity 2.77 Km/sec
2nd signal	(0.2509, 0.2214)	Velocity 2.99 Km/sec
3rd signal	(-0.2231,-0.6031)	Velocity 1.56 Km/sec .

The estimate of velocity for the third signal must be treated with some caution. As can be seen in Figure 6.8, the maximum corresponding to this lies on a long narrow ridge directed away from the origin, indicating a high variance in this direction.

Other studies by Asten indicate that velocities of 2 Km/sec and 2.8 Km/sec could be expected for the fundamental and first higher order Rayleigh waves respectively in this region and frequency (Asten, 1976).

For a narrow band of frequencies as used in this example, we may reasonably assume that the spectra are roughly constant and we may test for the significance of a signal by using an analysis of variance procedure. The relevant sums of squares are

Total	29610
Explained by one signal	22190
Explained by two signals	26870
Explained by three signals	28660 .

The total number of degrees of freedom is 182 (13 fundamental frequencies, 7 channels, real and imaginary parts) but this must be

```
ENTER LOWER AND UPPER FREQUENCIES (2F10.4)
1.         1.6
LOWER AND UPPER FUNDEMENTAL FREQUENCIES
       21        33
SINGLE BEAM PLOT?
YES
ENTER BOUND FOR PLOT (F10.4)
1.
CPONS METHOD? (OTHERWISE CONVENTIONAL)
YES
CYCLE THROUGH FREQUENCIES?
NO
MIN AND MAX VALUES      0.3637E+00      0.5705E+02
```

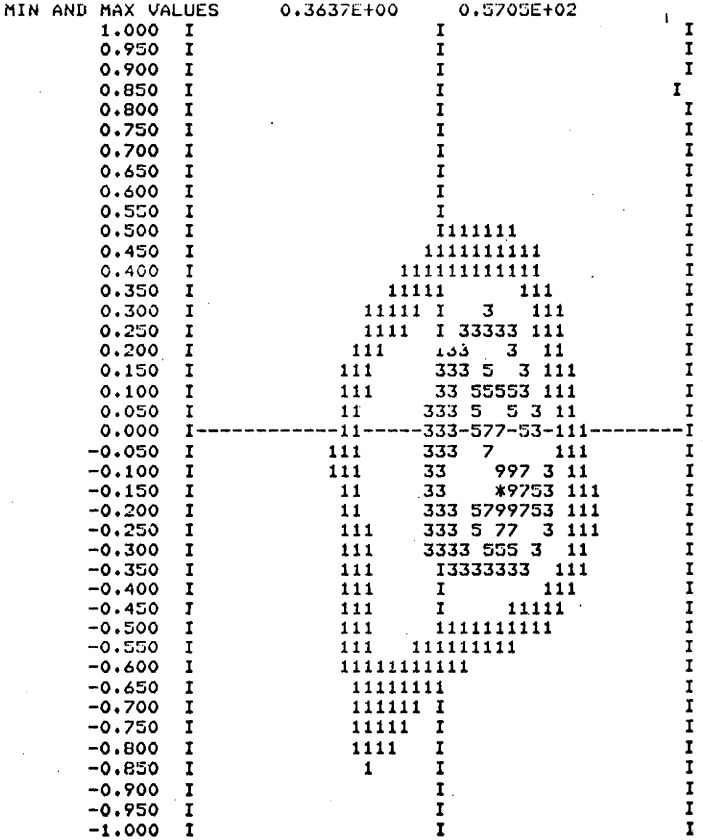


Figure 6.3

```
SINGLE BEAM PLOT?
YES
ENTER BOUND FOR PLOT (F10.4)
1.
CAPONS METHOD? (OTHERWISE CONVENTIONAL)
NO
CYCLE THROUGH FREQUENCIES?
NO
MAX. POSIBLE VALUE      0.2073E+06
MIN AND MAX VALUES      0.5268E+04      0.1553E+06
```

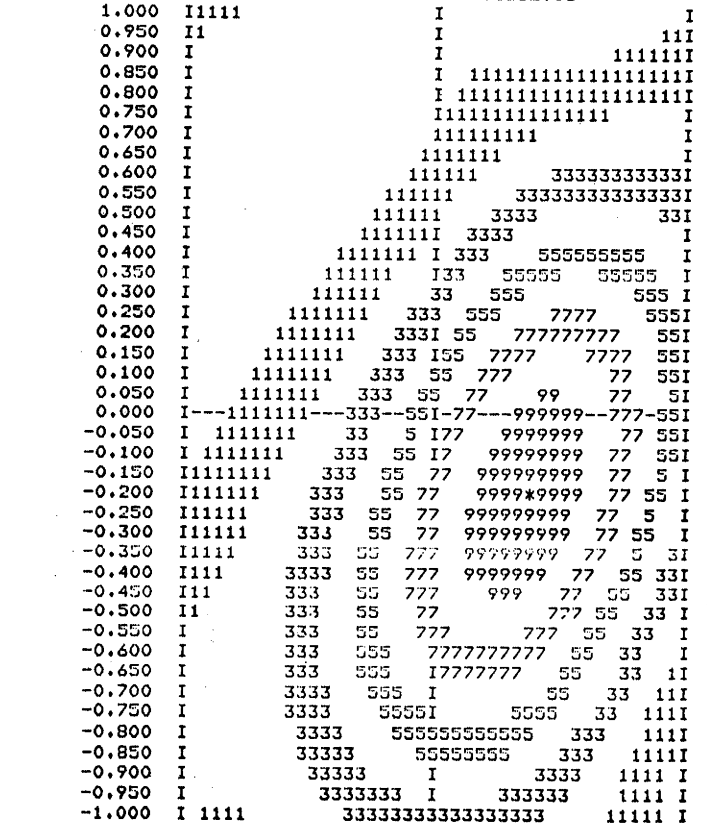


Figure 6.4

```
MULTI BEAM PLOT?
YES
ENTER NUMBER OF FIXED SIGNALS (I1)
1
ENTER FIXED SIGNAL PARAMETERS (2F10.4)
.39        -.21
ENTER BOUND FOR PLOT (F10.4)
1.
TOTAL VARIATION      0.2961E+05
RESIDUAL VARIATION   0.7418E+04
MIN AND MAX VALUES      0.4533E+03      0.4514E+04
        1.000  I                  111I  333        3333     I
        0.950  I                1111 I333          3333    I
        0.900  I              1111   33   5555555    333    I
        0.850  I            1111   33I  5555555555  333    I
        0.800  I          1111   333 I555       555  333    I
        0.750  I        111   333  555           55   33    I
        0.700  I      111   333  555I  777777    55   33    I
        0.650  I    111   33   55  I777777777   55 33   1I
        0.600  I  1111  33   55   777        77  55 33  11I
        0.550  I 1111   33  555  777I         77 55 33  11I
        0.500  I1111  333  55   77  I  999    77 5  3  111I
        0.450  I111  33   55  777   I999999  77 5 33  111I
        0.400  I1   33  555  777   99999999  77 5 3  111 I
        0.350  I   33  555  77    9999999999 7  5 3 111  I
        0.300  I  33   55  777   99999999999 7 5 3  11   I
        0.250  I 333  55  777   9999999*9999 7 5 3 11    I
        0.200  I333  55   77    999999999999 7 53  1     I
        0.150  I33   55  777    999999999999 75 3 11     I
        0.100  I33  555  777      9999999999  53 11      I
        0.050  I3   55   7777       9999999    3 1       I
        0.000  I3---555--77777------I---999-7-3-1--------I
       -0.050  I3   555   7777777777I        75 11       I
       -0.100  I3   5555    77777777777777  731111111111I
       -0.150  I3    5555          I     7  1111111111I
       -0.200  I33    5555         555555551 3     11111I
       -0.250  I333    55555555555555   33  75  33   1111I
       -0.300  I3333     555555555  I333    7 5  33   111I
       -0.350  I 33333             333    35  55 33   111I
       -0.400  I   33333          3333    3 5  55 33   111I
       -0.450  I     333333    33333I   33 5 555 33   111I
       -0.500  I     333333333333I  33 55555  33  1111I
       -0.550  I11      33333333333 I 333 5555   33  111 I
       -0.600  I1111     333333333333333   55   33   111 I
       -0.650  I11111      3333333333333       333  111  I
       -0.700  I111111       3333333333        33   111  I
       -0.750  I111111        333333333      333   111   I
       -0.800  I 111111       333333333     3333  111    I
       -0.850  I 111111       3333333333 333333   111    I
       -0.900  I 111111       33333333333333    111     I
       -0.950  I 1111111      3333333333333    1111     I
       -1.000  I 1111111      333333333333    1111      I
```

Figure 6.5

```
MULTI BEAM PLOT?
YES
ENTER NUMBER OF FIXED SIGNALS (I1)
1
ENTER FIXED SIGNAL PARAMETERS (2F10.4)
.14        .245
ENTER BOUND FOR PLOT (F10.4)
1.
TOTAL VARIATION      0.2961E+05
RESIDUAL VARIATION   0.1803E+05
MIN AND MAX VALUES      0.7955E+03      0.1524E+05
        1.000  I              1111 I          111          I
        0.950  I                1111  I 3333   111         I
        0.900  I              1111   3333333  111          I
        0.850  I            1111   33333333  11           I
        0.800  I           1111   3333  333  11           I
        0.750  I          1111   333I    33  1            I
        0.700  I         1111   333 I    33 11            I
        0.650  I        1111   333  I55  33 1            1I
        0.600  I        111   333  55555 3 11           11I
        0.550  I       1111   33   55555 3 1           1111I
        0.500  I      1111   333  55I  5 31          111111I
        0.450  I      111   333   55I  5  1      1111111  I
        0.400  I     1111   333   55I 7 31     1111        I
        0.350  I    1111    333   55I77     111           I
        0.300  I    1111    3333   55 75 1   3333333333331
        0.250  I   11111    333333  I 513              3333I
        0.200  I   11111     333333 I11 77  555555       3I
        0.150  I  111111          11 3   777    5555     I
        0.100  I  111111          111 I 5 99   777   555   I
        0.050  I  1111111     1111 113  9999  777   555   I
        0.000  I--1111111111111111111---99999---77---55--I
       -0.050  I  11111111111111111 35  999999   77  555 I
       -0.100  I   111111111111111 I57 9999999  77   55 I
       -0.150  I   1111111111111 3I 7 999*999   77  55 I
       -0.200  I   11111111111111 335 7 9999999   77 55  I
       -0.250  I    111111111111  3 5 7 9999999   7  55  I
       -0.300  I    111111111111 33 5 7 9999999  77  55 3I
       -0.350  I    11111111111  3 55 7  999999  77 55  3I
       -0.400  I    1111111111  33 5I 7  99999  77  55 33I
       -0.450  I   1111111111   33 5I 7   999   77 55  33I
       -0.500  I  11111111111  33  5I 77       77  5  33 I
       -0.550  I 1111111111   33  5I 77      77  55 33  I
       -0.600  I111111111    33  55  77777777  55  3  1I
       -0.650  I1111111111    333  55  7777777  55  33  1I
       -0.700  I1111111111    333  55    77    55  33  111
       -0.750  I111111111     333  555       555  33  111I
       -0.800  I11111111      333   5555   5555  33  111 I
       -0.850  I111111111     3333  I55555555   33  111  I
       -0.900  I   11111      3333  I   5     333  111   I
       -0.950  I    11111      3333 I       3333  1111   I
       -1.000  I    111111      333333  333333   1111    I
```

Figure 6.6

```
MULTI BEAM PLOT?
YES
ENTER NUMBER OF FIXED SIGNALS (I1)
1
ENTER FIXED SIGNAL PARAMETERS (2F10.4)
.345         -.153
ENTER BOUND FOR PLOT (F10.4)
1.
TOTAL VARIATION      0.2961E+05
RESIDUAL VARIATION   0.7651E+04
MIN AND MAX VALUES      0.3937E+03      0.4875E+04
        1.000  I                 111 I 333       3333     11I
        0.950  I               111   333         333      1I
        0.900  I             111   333    5555    333     1I
        0.850  I           1111   33 I 555555555  333     1I
        0.800  I         111   33   5555    555   33     1I
        0.750  I       111   33   555        555  33    11I
        0.700  I      111  333  555 I 77777   55  33    11I
        0.650  I     111  33   55   77777777  55  33   111I
        0.600  I   111   33  55   7777    777  5  3   111I
        0.550  I  111   33  55  7777I      77  5  3  1111I
        0.500  I1111  33  555  777  I       7  5 33  1111I
        0.450  I111  33  55   77    999999  7  5 3  1111 I
        0.400  I11  33  55   77    9999999  7 55 3 1111  I
        0.350  I1  333  55  77    999999999 7 5 3  111   I
        0.300  I   33  55  777   9999999999 7 5 3 111    I
        0.250  I  33  55  777    9999999999 7 53  11     I
        0.200  I 333  55  777    999999*999 7  3 11      I
        0.150  I333  55   777    9999999999 75  111     I
        0.100  I333  55   777     999999999   3 11   1111I
        0.050  I33   55   7777      I999999    1111111111I
        0.000  I33---555---77777----I---999-5-11111111111I
       -0.050  I33   555    7777777777    97 11111  11111I
       -0.100  I33    555        777 I 7777 711          11I
       -0.150  I333   55555         I555555    3333     1I
       -0.200  I3333    5555555555555   3   7 5  333    1I
       -0.250  I33333     5555555   I33    77  5  333   1I
       -0.300  I 33333             333   35777 55  33   1I
       -0.350  I   33333          333I   3  77  55  33   1I
       -0.400  I     3333333333333 I  335     55  33   1I
       -0.450  I       33333333333  I 33 5    55  333  11I
       -0.500  I11       33333333   I33  55   55  33   11I
       -0.550  I1111      33333    333  555555   33  111I
       -0.600  I11111           3333   555555  33   111I
       -0.650  I111111          33333   5555   333  111 I
       -0.700  I1111111        33333    5    333  111  I
       -0.750  I11111111      333333        333   11   I
       -0.800  I 1111111      333333       333   111   I
       -0.850  I 1111111     3333333      3333   111    I
       -0.900  I11111111     33333333   33333   111     I
       -0.950  I111111111    333333333333333   111     I
       -1.000  I111111111    33333333333333    1111    I
```

Figure 6.7

```
MULTI BEAM PLOT?
YES
ENTER NUMBER OF FIXED SIGNALS (I1)
2
ENTER FIXED SIGNAL PARAMETERS (2F10.4)
.3087        -.1612
.1038        .1869
ENTER BOUND FOR PLOT (F10.4)
1.
TOTAL VARIATION      0.2961E+05
RESIDUAL VARIATION   0.2736E+04
MIN AND MAX VALUES      0.3029E+03      0.1701E+04
        1.000  I          11111       I   111       111      I
        0.950  I         1111111      I   111        111     I
        0.900  I1       111111111     I   111        111     I
        0.850  I1      1111111111     I   11   33     111    I
        0.800  I11  11111     1111    I   11  3333    111    I
        0.750  I111111111       111   I   11 333333    11    I
        0.700  I1111111          111  I   1 333  333   11    I
        0.650  I1111111      3    11  I   1 3     33   11    I
        0.600  I11111     333333  11 I  1 3      33   11    I
        0.550  I11111   33333333  1 1  1 3 5555  33 11    I
        0.500  I1111   3333   333  11  13 55555  33 11    I
        0.450  I1111   33       33 11    5   55  33 11    I
        0.400  I111   33     5555  3 1   5     5  3  11    I
        0.350  I111   33  55555555 3I 13 777  5  3  11    I
        0.300  I11   33   55       531157777  5  3 11     I
        0.250  I11   33  55   777777I317  77  5 33 11     I
        0.200  I11   33  5  777     I  9  77  5 3  1      I
        0.150  I11  33  55  77  999999357 77 55 3 11      I
        0.100  I11  33  55 77  9999997315777 5  3 11      I
        0.050  I1   33  55 77  99999I7 1  7  5 3  1       I
        0.000  I1---33--5--77--9999-I-5-1----5-3-11------I
       -0.050  I1   33  5  77  9999 77531 5 5 3  11      I
       -0.100  I1   33  5  77  999  I7 531  5 3 111      I
       -0.150  I1   33  5  77  999  I77 5 15 3  111      I
       -0.200  I1   33  5  77  9999 I 7777 1   1111     I
       -0.250  I1   33  5  7   9999 I  7775 1111111      I
       -0.300  I    33  5  7  9999999   7   11111111     I
       -0.350  I    33 55 77  9999999   75  11111111     I
       -0.400  I    33 55 7  99999999  7 53  11111111    I
       -0.450  I   33  5  7  99999999  7  3  11111111    I
       -0.500  I   33  5 77 999999999 7 5 3 111111111    I
       -0.550  I   33 55 7  9999*999I 7  3  1111111111  I
       -0.600  I   33 5  7  9999999917 5 3 11111111111  I
       -0.650  I  33  5 77  9999999 77  3  11111111111  I
       -0.700  I  33  5 77 9999999  7 5 3 111111111111  I
       -0.750  I  33 55 77  99999  7I5 3  111111111111  I
       -0.800  I  33 55 77  9999  77I5 3 1111111111111  I
       -0.850  I  33 55 77       7 5 33 1111111111111  I
       -0.900  I  33 55 77      7  5 3  1111111111111  I
       -0.950  I  3  55  77    777 5I 3 111  111111111  I
       -1.000  I  33 55  77777777  5I3  11    11111111  I
```

Figure 6.8

corrected for the cosine taper used. A correction factor of 18/35 is required (Koopmans, 1974, p. 300). Hence the effective degrees of freedom for the total is 93.6. Similarly the effective degrees of freedom for each signal is 13.4.

The analysis of variance is presented in Table 6.1.

For the test of the significance of the third signal the test statistic is 7.5. This is obviously very significant compared with the F distribution with 13 and 52 degrees of freedom so that in that sense the third signal is "real".

If the statistic was not so clearly significant, such an analysis of variance would have to be treated with care as it ignores the changes in the design matrix θ.

6.3 NON-DIRECTIONAL METHODS

Asten also obtained data in the same region using an array consisting of six seismometers equally spaced on a circle of radius 50 metres and a seventh at the centre. The smaller size of this array indicates that it would be most useful at shorter wavelengths and hence higher frequencies than the previous one. The processing of the data up to the point of having the Fourier transforms is the same as before.

All the results which follow are based on spectral estimates obtained by averaging over eleven fundamental frequencies at a time.

Firstly the average of the complex coherencies between the centre seismometer and each one on the circle was calculated. Its real and imaginary parts are shown in Figure 6.9, in the frequency range 0 Hz to 10 Hz. Except for the "kink" at 5 Hz it is readily interpreted as giving a slowly decreasing velocity since it follows the shape of a

Table 6.1

	Sum of Squares	e.d.f.	Mean Square
2 signals	26870	26.8	1004.8
Difference between 2 signals and 3	1787	13.4	133.4
Residual	952	53.5	17.8
Total	29610	93.6	

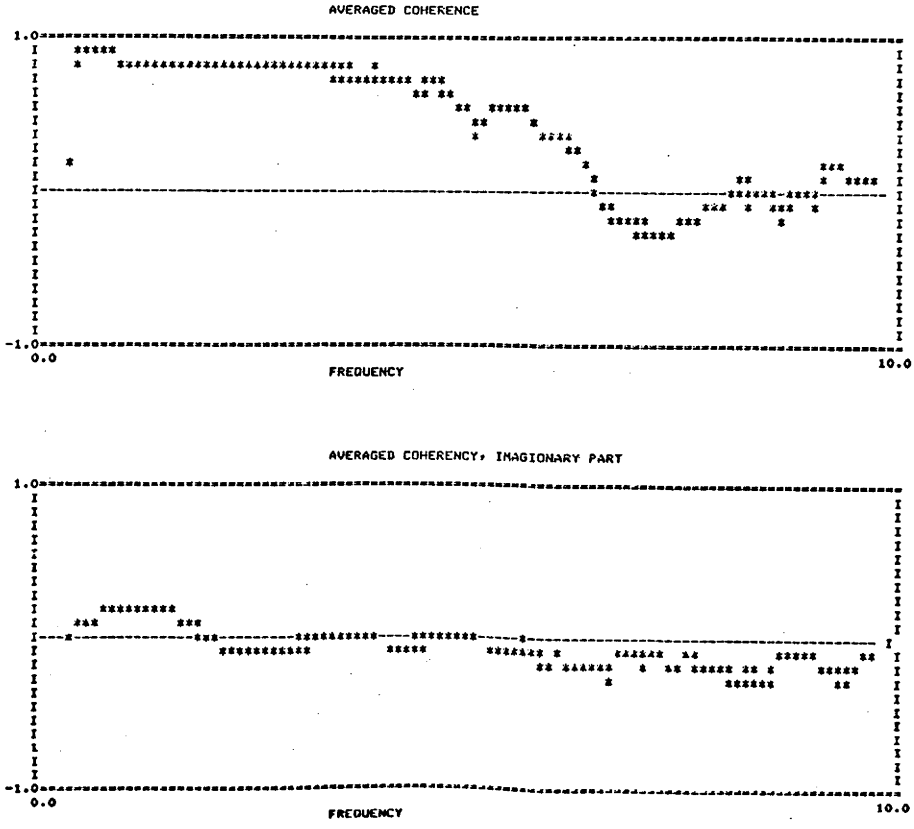


Figure 6.9

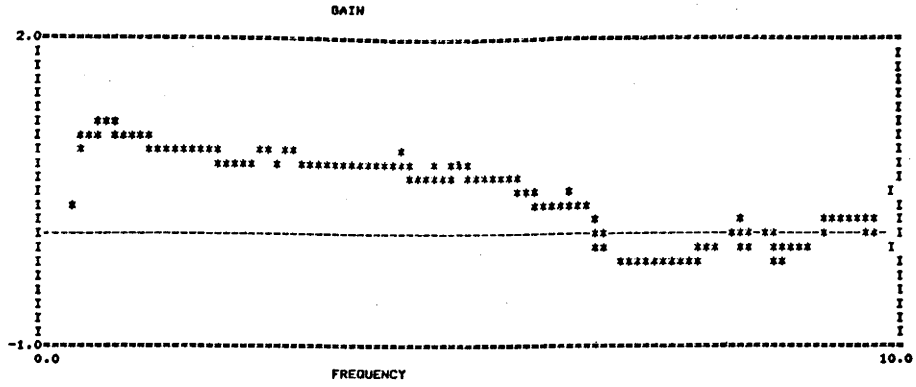


Figure 6.10

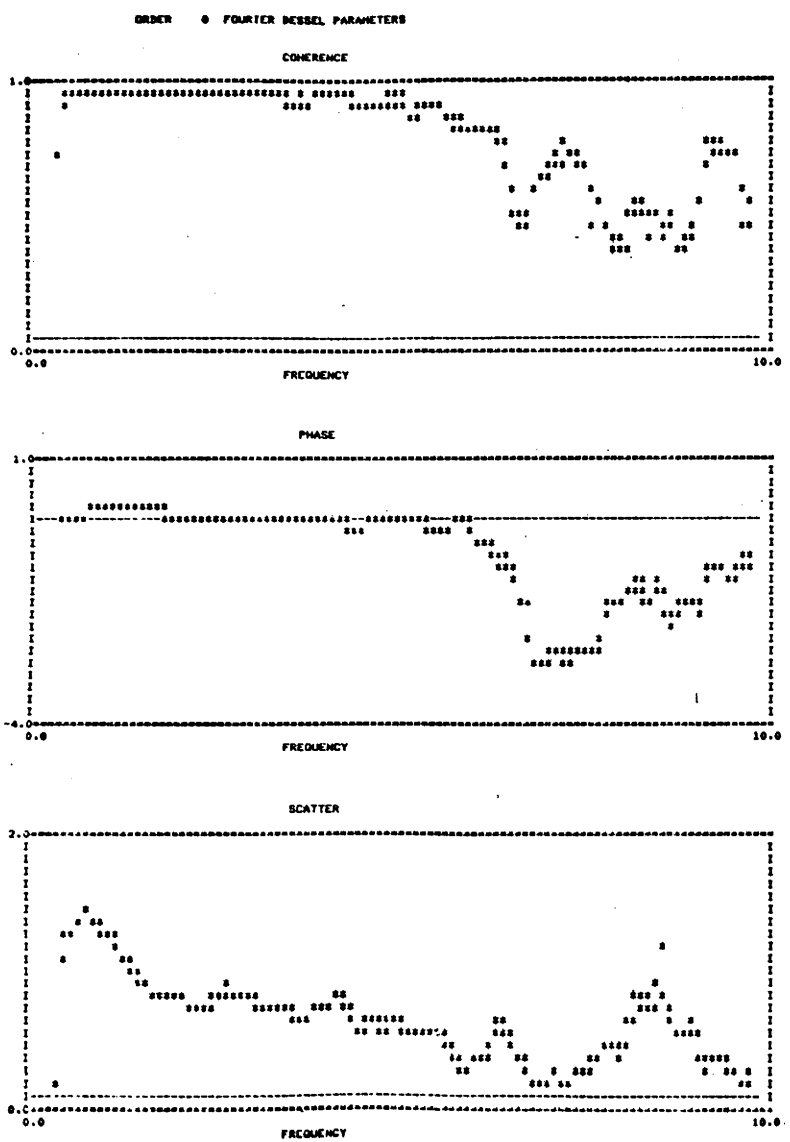


Figure 6.11

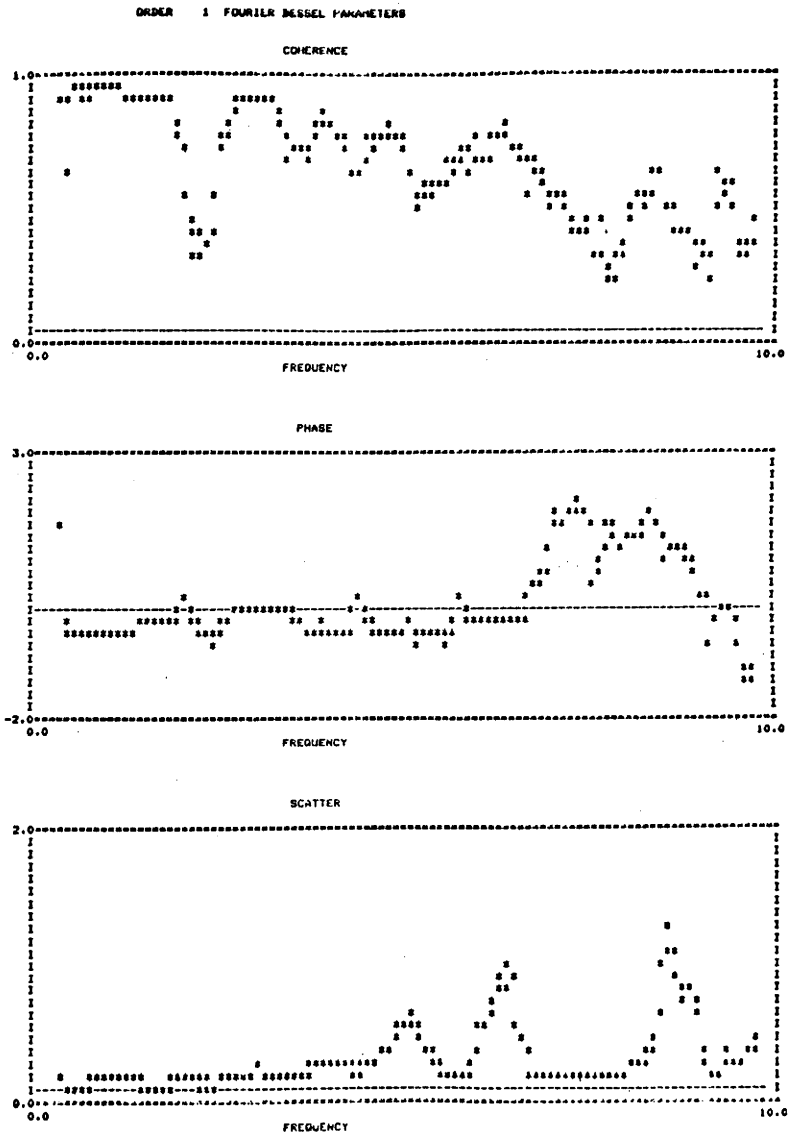


Figure 6.12

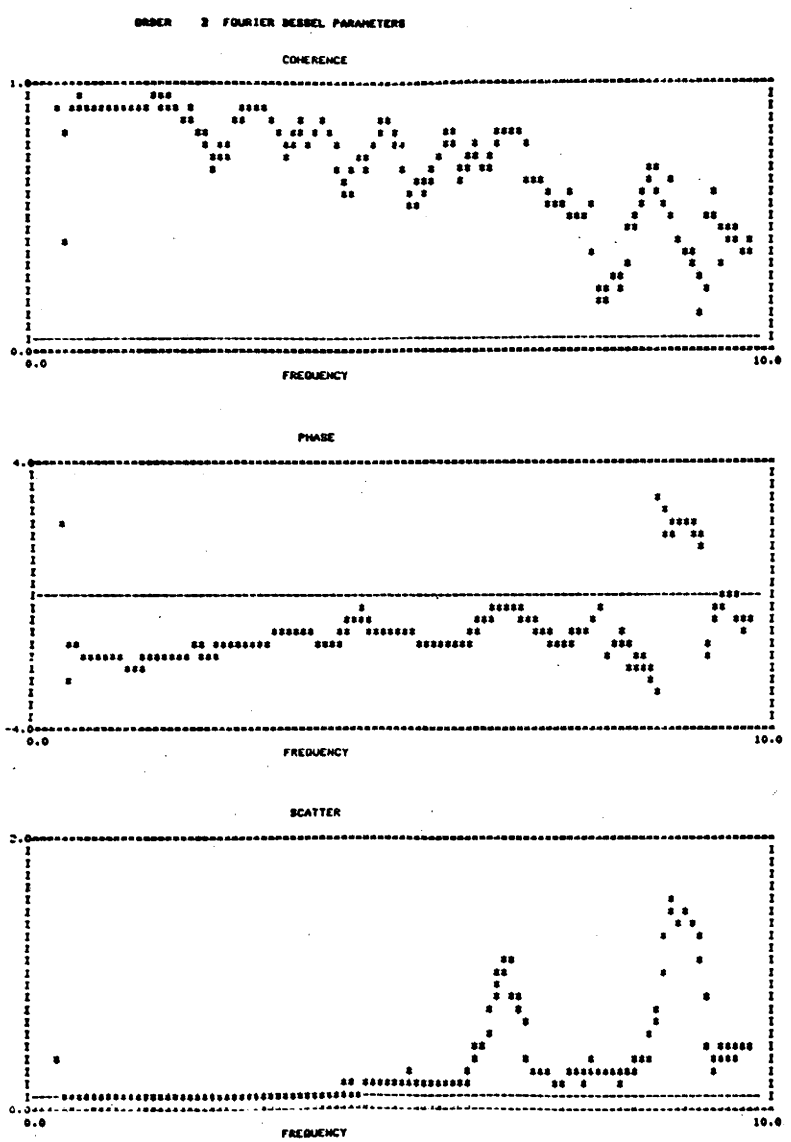


Figure 6.13

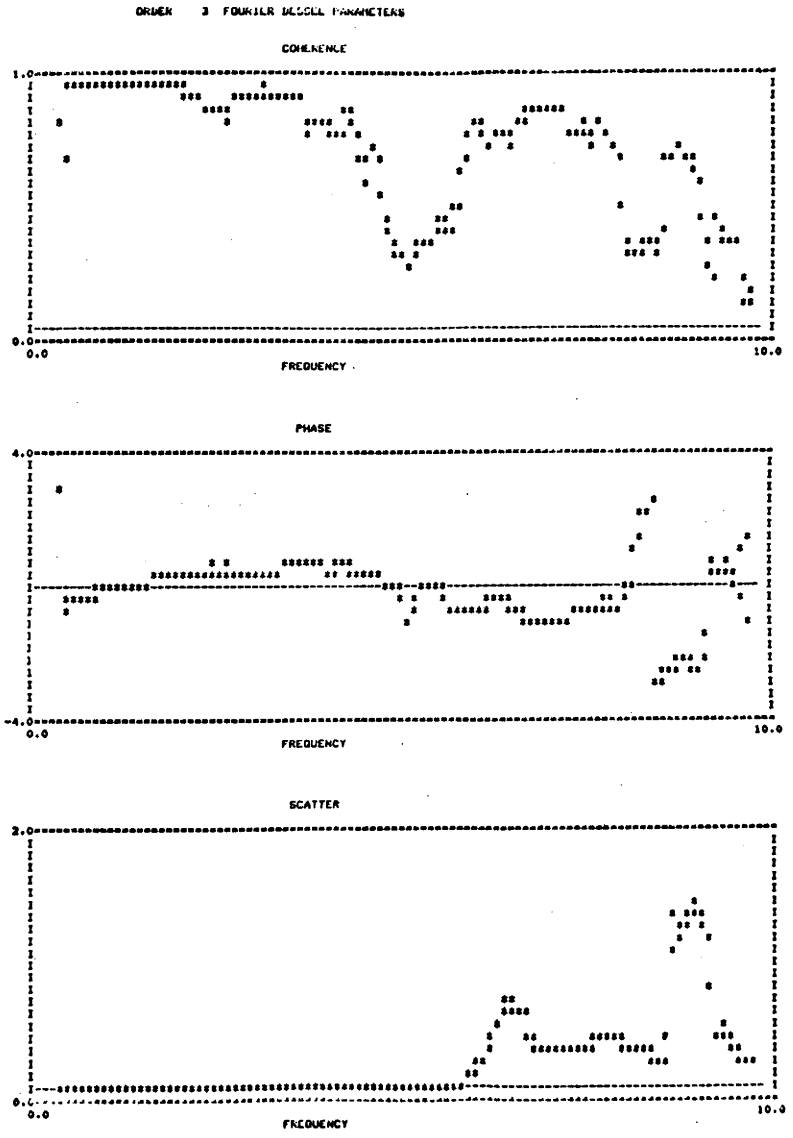


Figure 6.14

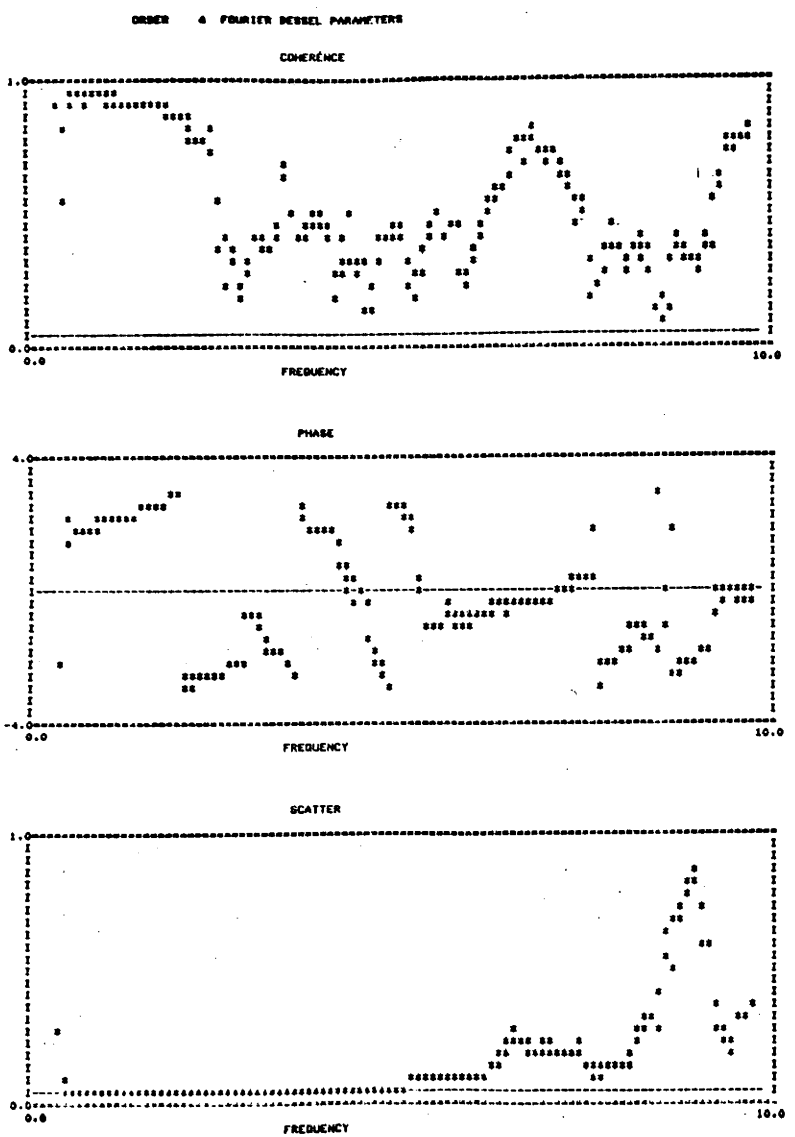


Figure 6.15

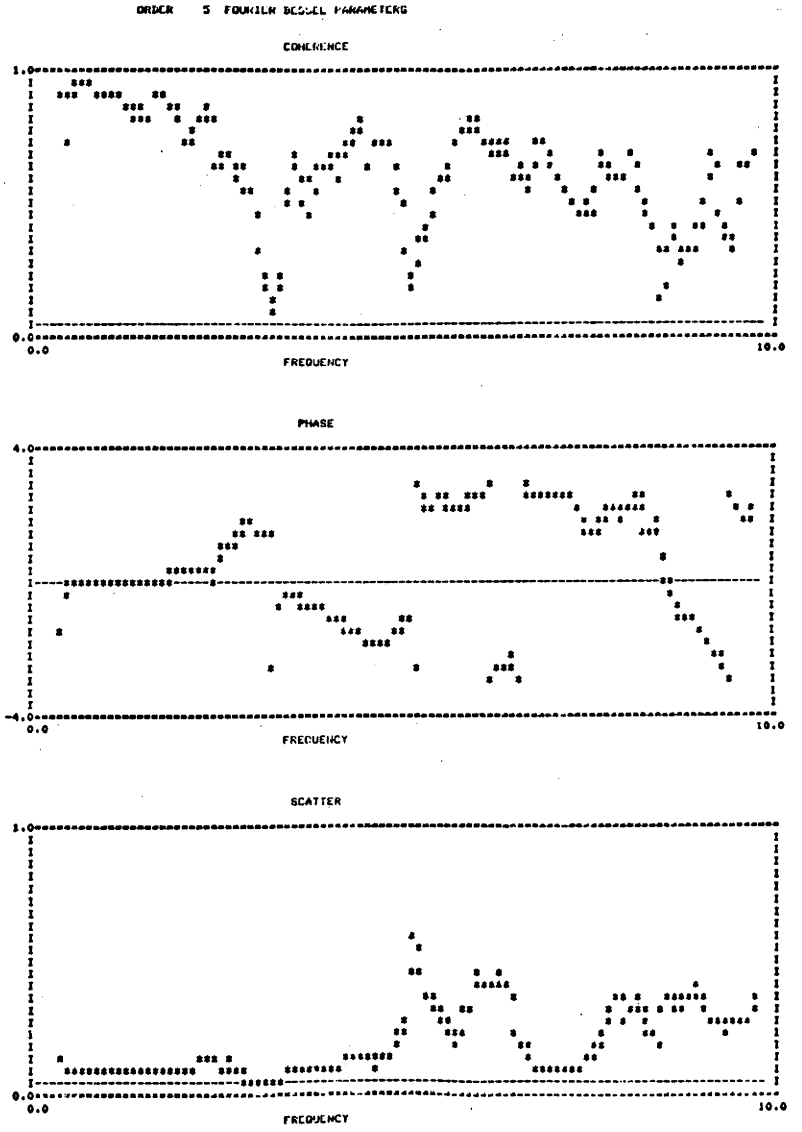


Figure 6.16

J_0 curve "squeezed" at the higher frequencies. The imaginary part is close to zero as expected.

If we look at the gain between the average of the outputs on the circle and the centre output the result is much less impressive. The real part of this estimated gain function is shown in Figure 6.10. It is much more irregular and is even greater than one in parts. This could be due to noise local to one receiver (very unlikely) or differences in sensitivity between receivers. It is an advantage of Aki's original methods that they compensate for differing sensitivities.

Figures 6.11 to 6.16 give Fourier coherence phase and scatter functions. Since there were six receivers on the circle of the array these functions are calculated for orders 0, 1, 2, 3(or -3), 4 (or -2) and 5 (or -1). The zero order functions are easiest to interpret. Below 6 Hz the coherence is close to one and the phase is close to zero, indicating the presence of only one significant velocity. The sharp minimum at 6.6 Hz almost definitely implies that $J_0(kr) = 0$ at this frequency, especially since the phase changes from zero to π. Above this frequency the coherence is low, indicating either noise at the receivers or more than one significant velocity.

The interpretation of the higher order results is not so clear. However the following remarks may be made.

1. Below 2Hz, the coherency functions are almost one, possibly indicating that only one signal is present. (The Fourier coefficients of a delta function have constant modulus.)

2. Above 6Hz the scatter functions have unexpected large peaks. They can only be explained by interference.

6.4 APPENDIX

The following programs were written for interactive use on a DEC System-10 computer and consequently contain some input and output statements which are not standard FORTRAN. The main controlling program gives self-explanatory prompts to the user. The printer plots have a linear scale but this can be altered to a logarithmic scale my altering the call to GPLOT. The iterative program BFAST, which uses BEAM to implement the deterministic estimator is not particularly efficient and could be improved.

```
C
C ***MAINLINE***
C
      DIMENSION S1(41),S2(41),X(41,41),S(2,3),TITLE(48)
      COMPLEX F(220,7)
      COMMON /POS/POSI( 7,2)
C *****    F IS ARRAY OF FOURIER TRANSFORMS OF DATA
C *****    S1 AND S2 ARE SLOWNESS COORDINATES USED IN GRID
C *****    POSI IS ARRAY OF RECEIVER POSITIONS
      DATA IN,IY/'N','Y'/
C ***         ENTER NUMBER OF INPUT FILE
      TYPE69
   69 FORMAT(' ENTER NUMBER OF INPUT FILE (I2)')
      ACCEPT63,INFILE
C ***     READ IN DATA
      READ(INFILE)(TITLE(I),I=1,48),NR,((POSI(I,J),J=1,2),I=1,NR)
     1 ,((F(I,J),I=1,220),J=1,NR)
    6 FORMAT(A1)
      TYPE3,(TITLE(I),I=1,48)
    3 FORMAT(5X,8A5)
    4 FORMAT(2F10.4)
      TYPE5,(I,(POSI(I,J),J=1,2),I=1,NR)
    5 FORMAT(//7(I5,2F15.4/) )
C
C ***       ENTER FREQUENCY LIMITS
   11 TYPE12
   12 FORMAT(' ENTER LOWER AND UPPER FREQUENCIES (2F10.4)')
      ACCEPT4,FMIN,FMAX
      NMIN=FMIN/.04883+1
      NMAX=FMAX/.04883+1
      TYPE13,NMIN,NMAX
   13 FORMAT(' LOWER AND UPPER FUNDEMENTAL FREQUENCIES '/2I
     1 10)
C
C ***       SINGLE BEAM PLOT
      TYPE21
   21 FORMAT(' SINGLE BEAM PLOT?')
      ACCEPT6,ICON
      IF(ICON.EQ.IN)GO TO 31
C ***       ENTER BOUNDS FOR PLOT
      TYPE45
      ACCEPT4,B
      B=B/20
      DO 22 I=1,41
      S1(I)=(I-21)*B
   22 S2(I)=S1(I)
      TYPE23
   23 FORMAT(' CAPONS METHOD? (OTHERWISE CONVENTIONAL)')
      ACCEPT6,ICON
C ***         CYCLE THROUGH FREQUENCIES? (SINGLE BEAM ONLY)
C ***       THIS ALLOWS COMPUTING OF A SERIES OF SLOWNESS PLOTS FOR
C ***       SUCCESSIVE FREQUENCY BANDS
      TYPE61
   61 FORMAT(' CYCLE THROUGH FREQUENCIES?')
      NC=0
      ACCEPT6,IC
      IF(IC.NE.IY)GO TO 26
      TYPE62
   62 FORMAT(' ENTER CYCLES,BASE FREQUENCY,STEP,BANDWIDTH (I2,',
     1  '3F10.2)')
      ACCEPT63,NC,FB,FS,BW
   63 FORMAT(I2,3F10.2)
   64 TYPE65,FB,BW
   65 FORMAT(//' CENTRE FREQUENCY',F10.2,'  BANDWIDTH',F10.2)
      NMAX=(FB+BW/2.)/.04883
      NMIN=(FB-BW/2.)/.04883
      FB=FB+FS
      NC=NC-1
   26 IF(ICON.EQ.IN)GO TO 24
      CALL CAPON(F,NR,NMIN,NMAX,S1,S2,X,41,41)
      GO TO 25
   24 CALL CON(F,NR,NMIN,NMAX,S1,S2,X,41,41)
   25 CALL GPLOT(X,41,41,S1,S2,0)
   31 IF(NC.GT.0)GO TO 64
C
C ***       MULTI BEAM PLOT
      TYPE32
   32 FORMAT(' MULTI BEAM PLOT?')
      ACCEPT6,ICON
      IF(ICON.EQ.IY)GO TO 36
      TYPE 71
      ACCEPT6,ICON
      IF(ICON.EQ.IY)GO TO 79
      GO TO 11
   36 TYPE33
   33 FORMAT(' ENTER NUMBER OF FIXED SIGNALS (I1)')
      ACCEPT34,NS
   34 FORMAT(I1,I2)
C ***    ENTER FIXED SIGNAL PARAMETERS
      TYPE41
   41 FORMAT(' ENTER FIXED SIGNAL PARAMETERS (2F10.4)')
      ACCEPT4,((S(I,J),I=1,2),J=1,NS)
C ***       ENTER BOUNDS FOR PLOT
      TYPE 45
   45 FORMAT(' ENTER BOUND FOR PLOT (F10.4)')
      ACCEPT4,B
      B=B/20.
      DO 46 I=1,41
      S1(I)=(I-21)*B
   46 S2(I)=S1(I)
      CALL BEAM(F,NR,NMIN,NMAX,S,NS,S1,S2,X,41,41)
      CALL GPLOT(X,41,41,S1,S2,0)
      TYPE50
   50 FORMAT(' REPEAT?')
      ACCEPT6,ICON
      IF(ICON.NE.IN)GO TO 31
   70 TYPE71
   71 FORMAT(' ITERATE MULTIBEAM?')
      ACCEPT6,ICON
      IF(ICON.NE.IY)GO TO 31
   79 TYPE72
   72 FORMAT(' ENTER NUMBER OF SIGNALS AND ITERATES (I1,I2)')
      ACCEPT34,NS,IMAX
      TYPE73
   73 FORMAT(' ENTER SIGNAL PARAMETERS (2F10.4)')
      ACCEPT4,((S(I,J),I=1,2),J=1,NS)
      TYPE74
   74 FORMAT(' ENTER GRID RANGE (F10.4)')
      ACCEPT4,D
      CALL FAST(F,NR,NMIN,NMAX,S,NS,IMAX,D)
      TYPE75,((S(I,J),I=1,2),J=1,NS)
   75 FORMAT(' ITERATED VALUES'/(2F10.4))
```

```
      TYPE76
   76 FORMAT(' EXIT?')
      ACCEPT6,ICON
      IF(ICON.NE.IY)GO TO 71
      STOP
      END
C
C ******************************************************************
C
C ***BEAM***
C
      SUBROUTINE BEAM(F,NR,NMIN,NMAX,S,NS,S1,S2,X,M1,M2)
C ***    THIS PROGRAM EVALUATES ESTIMATING FUNCTION FOR DETERMINISTIC
C ***    ESTIMATER, KEEPING ALL BUT ONE BEAM FIXED
C ***    F IS ARRAY OF FOURIER TRANSFORMS
C ***    S IS ARRAY OF FIXED SIGNALS
C ***    NS IS NUMBER OF FIXED SIGNALS
C ***    S1,S2 ARE SLOWNESS COORDINATES OF GRID
C ***    X IS OUTPUT ARRAY (M1 BY M2)
      COMPLEX F(220,7),THETA(7,3),YTH(7),TT(3,3),R2(50,7),
     1 R1(7),TT2(3),DET,D,THETS(7)
      DIMENSION S(2,3),X(M1,M2),DIS(3,7),S1(M1),S2(M2)
      COMMON /POS/POSI(7,2)/F400/THETA,YTH,TT,R2,R1,THETS,DIS
C ***    INITIALISE VARIABLES
      DO 11 I=1,M1
      DO 11 J=1,M2
      X(I,J)=0.
   11 CONTINUE
      ANR=NR
C ***    CALCULATE DISTANCES OF TRAVEL FOR FIXED SIGNALS
      DO 12 I=1,NS
      DO 12 J=1,NR
   12 DIS(I,J)=POSI(J,1)*S(1,I)+POSI(J,2)*S(2,I)
      A=0.
      B=0.
C ***    BEGIN LOOP THROUGH FREQUENCIES
      DO 50 I=NMIN,NMAX
      OM=(I-1)*.04883*6.2832
C ***     CALCULATE PHASE MATRIX AND INVERT FOR FIXED SIGNALS
      DO 21 J=1,NS
      DO 21 K=1,NR
   21 THETA(K,J)=CEXP(CMPLX(0.,OM*DIS(J,K)))
      TT(1,1)=ANR
      IF(NS.EQ.1)GO TO 24
      DO 23 J=2,NS
      JJ=J-1
      TT(J,J)=ANR
      DO 23 K=1,JJ
      TT(K,J)=0.
      DO 22 L=1,NR
   22 TT(K,J)=TT(K,J)+CONJG(THETA(L,K))*THETA(L,J)
   23 TT(J,K)=CONJG(TT(K,J))
      CALL MATINC(TT,NS,D,0,DET,3)
      GO TO 26
   24 TT(1,1)=1./ANR
C ***    SET UP PHASE DELAYS FOR EXTRA SIGNAL IN 2ND COMPONENT
   26 DO 25 J=1,M2
      DO 25 K=1,NR
   25 R2(J,K)=CEXP(CMPLX(0.,OM*S2(J)*POSI(K,2)))
      DO 27 J=1,NR
      YTH(J)=CONJG(F(I,J))
      DO 27 K=1,NR
      DO 27 L=1,NS
      DO 27 M=1,NS
   27 YTH(J)=YTH(J)-CONJG(F(I,K))*THETA(K,L)*TT(L,M)*CONJG(THETA
     1 (J,M))
      DO 29 J=1,NR
      A=A+CABS(F(I,J))**2
   29 B=B+CABS(YTH(J))**2
C ***    BEGIN CYCLE THROUGH FIRST COMPONENT
      DO 50 J=1,M1
C ***    SET UP PHASE DELAYS FOR 1ST COMPONENT
      DO 31 K=1,NR
   31 R1(K)=CEXP(CMPLX(0.,OM*S1(J)*POSI(K,1)))
C ***    BEGIN CYCLE THROUGH SECOND COMPONENT
      DO 50 K=1,M2
      D=0.
      DO 41 L=1,NR
      THETS(L)=R1(L)*R2(K,L)
   41  D=D+THETS(L)*YTH(L)
      R=CABS(D)**2
C ***    CALCULATE CROSS PRODUCTS AND INVERT
      DO 42 L=1,NS
      TT2(L)=0.
      DO 42 M=1,NR
   42 TT2(L)=TT2(L)+CONJG(THETA(M,L))*THETS(M)
      D=ANR
      DO 44 L=1,NS
      DO 44 M=1,NS
   44 D=D-CONJG(TT2(L))*TT(L,M)*TT2(M)
      XX=R/REAL(D)
      X(J,K)=X(J,K)+XX
   50 CONTINUE
      TYPE51,A,B
   51 FORMAT(' TOTAL VARIATION',E15.4/' RESIDUAL VARIATION'
     1 ,E12.4)
      RETURN
      END
C
C ******************************************************************
C
***CON***
C
      SUBROUTINE CON(F,NR,NMIN,NMAX,S1,S2,X,M1,M2)
C ***    THIS PROGRAM EVALUATES THE ESTIMATING FUNCTION FOR THE
C ***    SINGLE BEAM ESTIMATOR
      COMPLEX F(220,7),Z,R2(50,7),R1(7)
      DIMENSION X(M1,M2),S1(M1),S2(M2)
      COMMON /POS/POSI(7,2)/F400/R1,R2,ANR,Z
C ***    INITIALISE ARRAY
      ANR=NR
      DO 11 J=1,M1
      DO 11 K=1,M2
   11 X(J,K)=0.
      AL=0.
C ***    BEGIN CYCLE THROUGH FREQUENCIES
      DO 50 I=NMIN,NMAX
      OM=(I-1)*.04883
      DO 12 J=1,NR
   12 AL=AL+CABS(F(I,J))**2
C ***    SET UP PHASE DELAYS FOR 2ND COMPONENT
      DO 21 J=1,M2
```

```
      DO 21 K=1,NR
   21 R2(J,K)=CEXP(CMPLX(0.,OM*6.2832*S2(J)*POSI(K,2)))
C ***    BEGIN CYCLE THROUGH FIRST COMPONENT
      DO 50 J=1,M1
C ***    SET UP PHASE DELAY FOR 1ST COMPONENT
      DO 31 K=1,NR
   31 R1(K)=CEXP(CMPLX(0.,OM*6.2832*S1(J)*POSI(K,1)))
C ***    BEGIN CYCLE THROUGH SECOND COMPONENT
      DO 50 K=1,M2
      Z=0.
      DO 42 L=1,NR
   42 Z=Z+CONJG(F(I,L))*R1(L)*R2(K,L)
   50 X(J,K)=X(J,K)+CABS(Z)**2
      AL=AL*ANR
      TYPE60,AL
   60 FORMAT(' MAX. POSIBLE VALUE',E15.4)
      RETURN
      END
C
C ***************************************************************
C
C ***CAPON***
C
      SUBROUTINE CAPON(F,NR,NMIN,NMAX,S1,S2,X,M1,M2)
C ***    THIS PROGRAM EVALUATES CAPONS 'MAXIMUN LIKELIHOOD
C ***    FUNCTION ON A GRID
      COMPLEX F(220,7),FS(7,7),DET,Z,R2(50,7),R1(7),THETA(7)
      DIMENSION X(M1,M2),S1(M1),S2(M2)
      COMMON /POS/POSI(7,2)/F400/DET,Z,R2,R1,THETA
C ***    FORM AND INVERT SPECTRAL DENSITY MATRIX
      R=NMAX-NMIN+1.
      DO 12 J=1,NR
      DO 12 K=J,NR
      Z=0.
      DO 11 I=NMIN,NMAX
   11 Z=Z+F(I,J)*CONJG(F(I,K))
      FS(J,K)=Z/R
   12 FS(K,J)=CONJG(FS(J,K))
      CALL MATINC(FS,NR,Z,0,DET,7)
      DO 13 K=1,NR
   13 FS(K,K)=FS(K,K)/2.
      OM=(NMAX+NMIN+2)/2.*.04883
C ***    SET UP PHASE DELAYS FOR 2ND COMPONENT
      DO 21 J=1,M2
      DO 21 K=1,NR
   21 R2(J,K)=CEXP(CMPLX(0.,OM*6.2832*S2(J)*POSI(K,2)))
C ***    BEGIN CYCLE THROUGH FIRST COMPONENT
      DO 50 J=1,M1
C ***    SET UP PHASE DELAY FOR 1ST COMPONENT
      DO 31 K=1,NR
   31 R1(K)=CEXP(CMPLX(0.,OM*6.2832*S1(J)*POSI(K,1)))
C ***    BEGIN CYCLE THROUGH SECOND COMPONENT
      DO 50 K=1,M2
      DO 41 L=1,NR
   41 THETA(L)=R1(L)*R2(K,L)
      Z=0.
      DO 42 L=1,NR
      DO 42 M=L,NR
   42 Z=Z+CONJG(THETA(L))*FS(L,M)*THETA(M)
   50 X(J,K)=.5/REAL(Z)
      RETURN
      END
C
C ***************************************************************
C
C ***GPLOT***
C
      SUBROUTINE GPLOT(X,M,N,S1,S2,ILOG)
C ***    THIS PROGRAM CREATES PRINTER CONTOUR PLOTS
      DIMENSION X(M,N),IX(50),IA(5),S1(M),S2(N)
      COMMON /F400/IX
      DATA IB,IM,II,IAS,IA/' ','-','I','*','1','3','5','7','9'/
      XMAX=X(1,1)
      XMIN=X(1,1)
      DO 2 I=1,M
      DO 2 J=1,N
      IF(X(I,J).GT.XMAX)XMAX=X(I,J)
      IF(X(I,J).LT.XMIN)XMIN=X(I,J)
    2 CONTINUE
C ***    TEST IF PLOT IS TRIVIAL AND SET UP RANGE
      TYPE31,XMIN,XMAX
   31 FORMAT(' MIN AND MAX VALUES',2E15.4)
      IF(ILOG.EQ.0)GO TO 14
      IF(XMIN.GT.0.)GO TO 12
      TYPE11,XMIN
   11 FORMAT(' XMIN='E15.4,'LOG PLOT NOT POSSIBLE')
      RETURN
   12 XMIN=ALOG10(XMIN)
      R=ALOG10(XMAX)-XMIN
      GO TO 16
   14 R=XMAX-XMIN
      IF(R.GT.0.)GO TO 16
      TYPE15
   15 FORMAT(' RANGE ZERO, PLOT TRIVIAL')
      RETURN
   16 CONTINUE
C ***    BEGIN OUTER LOOP
      DO 23 JJ=1,N
      J=N-JJ+1
      IL=IB
      IF(J.EQ.21)IL=IM
      DO 35 I=1,M
      IF(ILOG.NE.0)GO TO 21
      IXX=(X(I,J)-XMIN)/R*10
      GO TO 22
   21 IXX=(ALOG10(X(I,J))-XMIN)/R*10
   22 IK=IL
      IF(I.EQ.21)IK=II
      IX(I)=IK
      IT=IXX/2
      IF(IT*2.EQ.IXX)GO TO 32
      IX(I)=IA(IT+1)
      GO TO 35
   32 IF(IXX.EQ.10)IX(I)=IAS
   35 CONTINUE
   23 TYPE24,S2(J),(IX(I),I=1,M),II
   24 FORMAT(F14.3,'  I',100A1)
      RETURN
      END
C
C ***************************************************************
C
```

```
C ***MATINC***
C
      SUBROUTINE MATINC(A,N,B,M,DETERM,NMAX)
C ***     THIS PROGRAM INVERTS COMPLEX MATRICES
      COMPLEX A(NMAX,N),B(NMAX,M)
      COMPLEX PIVOT,DETERM,SWAP,AMAX,T
      COMMON /F402/PIVOT(20),INDEX(20)
      DETERM=1.
      DO 20 I=1,N
      PIVOT(I)=0.
   20 INDEX(I)=0
      DO 550 I=1,N
      AAMAX=0.
      DO 105 J=1,N
      IF(CABS(PIVOT(J)).NE.0.)GO TO 105
      DO 100 K=1,N
      IF(CABS(PIVOT(K)).NE.0.)GO TO 100
      TEMP=CABS(A(J,K))
      IF(TEMP.LT.AAMAX)GO TO 100
      IROW=J
      ICOLUM=K
      AAMAX=TEMP
  100 CONTINUE
  105 CONTINUE
      INDEX(I)=4096*IROW+ICOLUM
      J=IROW
      AMAX=A(J,ICOLUM)
      DETERM=AMAX*DETERM
      IF(DETERM.EQ.0.)GO TO 600
      PIVOT(ICOLUM)=AMAX
      IF(IROW.EQ.ICOLUM)GO TO 260
      DETERM=-DETERM
      DO 200 K=1,N
      SWAP=A(J,K)
      A(J,K)=A(ICOLUM,K)
      A(ICOLUM,K)=SWAP
  200 CONTINUE
      IF(M.LE.0)GO TO 260
      DO 250 K=1,M
      SWAP=B(J,K)
      B(J,K)=B(ICOLUM,K)
      B(ICOLUM,K)=SWAP
  250 CONTINUE
  260 K=ICOLUM
      A(ICOLUM,K)=1.
      DO 350 K=1,N
      A(ICOLUM,K)=A(ICOLUM,K)/AMAX
  350 CONTINUE
      IF(M.LE.0)GO TO 380
      DO 370 K=1,M
      B(ICOLUM,K)=B(ICOLUM,K)/AMAX
  370 CONTINUE
  380 DO 550 J=1,N
      IF(J.EQ.ICOLUM)GO TO 550
      T=A(J,ICOLUM)
      A(J,ICOLUM)=0.
      DO 450 K=1,N
      A(J,K)=A(J,K)-A(ICOLUM,K)*T
  450 CONTINUE
      IF(M.LE.0)GO TO 550
      DO 500 K=1,M
      B(J,K)=B(J,K)-B(ICOLUM,K)*T
  500 CONTINUE
  550 CONTINUE
  600 DO 710 I=1,N
      I1=N+1-I
      K=INDEX(I1)/4096
      ICOLUM=INDEX(I1)-4096*K
      IF(K.EQ.ICOLUM)GO TO 710
      DO 705 J=1,N
      SWAP=A(J,K)
      A(J,K)=A(J,ICOLUM)
      A(J,ICOLUM)=SWAP
  705 CONTINUE
  710 CONTINUE
      RETURN
      END
C
C ***********************************************************
C
C ***FAST***
C
      SUBROUTINE FAST(F,NR,NMIN,NMAX,S,NS,IMAX,D)
C ***     THIS PROGRAM ITERATES BEAM ON SMALL GRIDS
C ***     ROUTINE ITERATES BEAM PROGRAM ON SMALL GRIDS
      COMPLEX F(220,7)
      DIMENSION S(2,3),X(9,9),S1(9),S2(9)
      IN=0
      DO 30 I=1,IMAX
      DO 11 J=1,9
      S1(J)=(J-5.05)*D/8.+S(1,NS)
   11 S2(J)=(J-5.05)*D/8.+S(2,NS)
      CALL BEAM(F,NR,NMIN,NMAX,S,NS-1,S1,S2,X,9,9)
C ***     FIND MAXIMUN ON GRID
      XMAX=X(1,1)
      JM=1
      KM=1
      DO 12 J=1,9
      DO 12 K=1,9
      IF(X(J,K).LT.XMAX)GO TO 12
      XMAX=X(J,K)
      JM=J
      KM=K
   12 CONTINUE
C ***     ROLL DOWN
      DO 13 JJ=2,NS
      J=NS-JJ+1
      S(1,J+1)=S(1,J)
   13 S(2,J+1)=S(2,J)
      S(1,1)=S1(JM)
      S(2,1)=S2(KM)
      IN=IN+1
      IF(JM.LE.2.OR.JM.GE.8.OR.KM.LE.2.OR.KM.GE.8)IN=0
      IF(IN.LT.NS)GO TO 14
      D=D/2.
      IN=0
   14 CONTINUE
   30 CONTINUE
      RETURN
      END
C
```

BIBLIOGRAPHY

Aki, K. (1957) Space and time spectra of stationary stochastic waves with special reference to microtremors. *Bull. Earthquake Res. Inst.* 35, 415-456.

Aki, K. (1961) A note on the use of microseisms in determining the shallow structure of the earth, *Geophysics* 30, 665-666.

Anderson, T.W. (1958) *An Introduction to Multivariate Statistical Analysis*, Wiley, New York.

Anderson, V.C. (1960) Digital array phasing, *J. Acoustical Soc. Amer.* 32, 867-870.

Asten, M. (1976) Use of Microseisms in Exploration Geophysics, Ph.D. Thesis, Macquarie University, Sydney.

Barber, N.F. (1961) The directional resolving power of an array of wave detectors; in *Ocean Wave Spectra*, Prentice Hall, 137-150.

Båth, M. (1975) *Introduction to Seismology*, Birkhauser Verlag, Basel-Stuttgart.

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

Brillinger, D.R. and Rosenblatt, M. (1966) Asymptotic theory of k^{th} order spectra; in *Advanced Seminar on Spectral Analysis of Time Series*, ed. by B. Harris, Wiley, New York, 153-188.

Burg, J.P. (1967) *Maximum Entropy Spectral Analysis*; presented at the 37th Annual International meeting of the Society of Exploration Geophysicists.

Capon, J. (1969) High resolution frequency wavenumber spectra,

Proc. I.E.E.E. 57, 1408-1418.

Cron, B.F. and Sherman, C. (1962) Theoretical derivation of spatial correlation functions for various noise models, *J. Acoustical Soc. Amer.* 34, 1732-1736.

Davies, D. (1973) Seismology with large arrays, *Reports Progress Phys.* 36, 1223-1283.

Der, Z.A. and Flinn, E.A. (1975) The applicability of principal component analysis to the separation of multiple planewave signals, *Bull. Seism. Soc. Amer.* 65, 627-635.

Dziewonski, A.M. and Hales, A.L. (1972) Numerical analysis of dispersed seismic waves; in *Methods of Computational Physics, Vol.II, Seismology, Surface Waves and Earth Oscillations*, ed. by B. Bolt, Academia, New York, 39-85.

Eckart, C. (1953) The theory of noise in continuous media, *J. Acoustical Soc. Amer.* 25, 195-199.

Hannan, E.J. (1973) Multivariate time series analysis, *J. Multivariate Anal.* 3, 395-407.

Hannan, E.J. (1975) Measuring the velocity of a signal; in *Perspectives in Probability and Statistics*, ed. by J. Gani, Applied Probability Trust, Sheffield, 227-237.

Hannan, E.J. and Robinson, P. (1973) Lagged regression with unknown lags, *J. Royal Stat. Soc. B* 35, 252-267.

Hannan, E.J. and Thompson, P.J. (1973) Estimating group delay, *Biometrika* 60, 241-253.

Hinnich, M.J. and Shamon, P. (1972) Parameter estimation for an R dimensional plane wave observed with independent Gaussian errors, *Ann. Math. Stat.* 43, 153-169.

Knittel, G.H. (1974) Relation of radar range resolution and signal to noise ratio to phase array bandwidth, *I.E.E.E. Trans. Anten.*

Hannan, E.J. (1970) Multiple Time Series, John Wiley, New York

AP-22, 418-426.

Koopmans, L.H. (1974) *Spectral Analysis of Time Series*, Academic Press, New York.

McDonough, R.N. (1974) Maximum entropy spatial processing of array data, *Geophysics* 39, 843-851.

Munk, W.H., Miller, G.R., Snodgrass, F.E., and Barber, N.F. (1963) Directional recording of swell from distant storms, *Phil. Trans. Royal Soc. A*, 225, 505-585.

Pisarenko, V.F. (1971) On the estimation of spectra by means of non-linear functions of the covariance matrix, *Geophys. J. Royal Astr. Soc.* 28, 511-531.

Posmentier, E.S. and Herman, R.W. (1971) Cophase: an ad hoc array processor, *J. Geophys. Res.* 76, 2194-2201.

Rao, C.R. (1973) *Linear Statistical Inference and its Applications*, Wiley, New York.

Schell, A.C. (1969) The antenna array as a spatial filter; in *Antenna Theory, Part 2*; ed. by R. Collin and F. Zucker, McGraw Hill, New York.

Widdrow, B., Montey, P.E., Griffiths, L.J., and Goode, B.B. (1967) Adaptive antenna systems, *Proc. I.E.E.E.* 55, 2143-2159.

Wilks, S.S. (1962) *Mathematical Statistics*, Wiley, New York.

Winder, A.A. (1975) Sonar system technology, *I.E.E.E. Trans. Sonics Ultrasonics*, SU-22, 291-332.

Woods, J.W. and Lintz, P.R. (1973) Plane waves at small arrays, *Geophysics* 38, 1023-1041.

Yaglom, A.N. (1961) Second order homogeneous random fields; in *Proc. 4th Berkeley Symposium on Probability and Statistics*; ed. by J. Neyman, University of California Press, Berkeley, 593-622.