LIDS-P-1620

AN EFFICIENT MAXIMUM ENTROPY TECHNIQUE FOR 2-D ISOTROPIC RANDOM FIELDS

Ahmed H. Tewfik Bernard C. Levy
Alan S. Willsky

Laboratory for Information and Decision Systems and
The Departement of Electrical Engineering and Computer Science
Massachusetts Institute of Technology, Cambridge, MA 02139

November 30, 1986

Abstract

In this paper, we present a new linear MEM algorithm for 2-D isotropic random fields. Our procedure differs from previous 2-D MEM algorithms by the fact that we take maximal advantage of the symmetries implied by isotropy, which is the natural generalization to several dimensions of the 1-D notion of stationarity. Unlike general 2-D covariances, isotropic covariance functions which are positive definite on a disk are known to be extendible. Here, we develop a computationally efficient procedure for computing the MEM isotropic spectral estimate corresponding to an isotropic covariance function which is given over a finite disk of radius $2R$. We show that the isotropic MEM problem has a linear solution and that it is equivalent to the problem of constructing the optimal linear filter for estimating the underlying isotropic field at a point on the boundary of a disk of radius R given noisy measurements of the field inside the disk. Our procedure is based on Fourier series expansions in both the space and wave-number domains of the inverse of the MEM spectral estimate. Furthermore, our method is guaranteed to yield a valid isotropic spectral estimate and it is computationally efficient since it requires only $O(BRL^2)$ operations where L is the number of points used to discretize the interval $[0, R]$, and where B is the bandwidth in the wave-number plane of the spectrum that we want to estimate. We also present examples to illustrate the behavior of our algorithm and its high resolution property.

---

*This work was supported by the National Science Foundation under Grant No. ECS-83-12921 and by the Army Research Office under Grant No. DAAG-84-K-0005.

1 INTRODUCTION

The need for efficient power spectral estimation techniques arises in a number of practical applications, such as speech processing [1], radar [2], sonar [3], image processing [4] and seismic signal processing [5], to mention a few. For one-dimensional signals, the maximum entropy spectral estimation method (MEM) has become very popular due to the facts that it can provide excellent frequency resolution, and that it can be implemented in a computationally efficient way [6]. Because of the multidimensional nature of the signals arising in many applications (e.g. in geophysical problems, imaging, sonar, etc.) a number of maximum entropy algorithms have been developed over the past ten years ([7]-[10]) for estimating two-dimensional spectra. These algorithms are very general and do not attempt to exploit any special structure of the power spectrum to be estimated. Since 2-D polynomials do not possess in general a quarter-plane factorization [11],[12], most of the known 2-D MEM algorithms involve solving a *non-linear* optimization problem that cannot be reduced to a linear prediction problem as in the 1-D case [13]. Furthermore, 2-D covariance functions which are positive definite on a subspace of the plane $\mathbf{R}^2$ do not necessarily have a positive-definite extension to the whole plane [10], [14]. Thus, for any given set of stationary covariance data, the 2-D MEM problem is not guaranteed in general to have a solution. This can constitute a major problem in practice, since the covariance values that are usually used as an input to the direct 2-D MEM spectral estimation algorithms are estimates, rather than exact values, of the true covariance values, and thus may not correspond to an extendible positive-definite 2-D function. A good review of the various 2-D MEM algorithms and of the extendibility issue can be found in [13].

In this paper, by contrast, we present a new *linear* MEM algorithm for 2-D isotropic random fields. Isotropic fields are characterized by the fact that their mean value is a constant independent of position and their autocovariance function is invariant under all rigid body motions, i.e. under translations and rotations. In some sense, isotropy is the natural extension of the notion of stationarity in one dimension. Isotropic fields arise in a number of physical problems of interest among which we can mention

the modeling of background noises in seismology [15] and ocean acoustics [16], [17], the investigation of temperature and pressure distributions in the atmosphere at a constant altitude [18], the analysis of turbulence in statistical fluid mechanics [19] and the representation of rainfall structure in hydrology [20].

An important property of isotropic covariance functions which are positive-definite over a disk is that they *always* have *positive-definite isotropic extensions* to the whole plane [21]. Here, we develop a computationally efficient linear procedure for computing the maximum entropy isotropic power spectral estimate corresponding to a covariance function that is given over a disk of radius $2R$. The maximum entropy power spectral estimate is the one that maximizes the entropy of the underlying random field. Our 2-D isotropic MEM algorithm is similar in spirit to the 1-D MEM procedure as will become clear from what follows. By using a nonsymmetric half plane spectral factorization and the properties of radially symmetric functions which are zero outside a disk in the space domain, we show that the isotropic MEM problem is equivalent to the problem of constructing the optimal linear filter for estimating the value of the underlying isotropic field at a point on the boundary of a disk of radius R given noisy observations of the field inside the disk. We then present a computationally efficient and robust procedure for computing the isotropic MEM spectral estimate. Our procedure is based on a Fourier expansion of the optimal linear estimation filter in terms of the angle θ in a polar representation of the underlying 2-D space, and on the fast recursions that were derived in [22] for solving filtering problems for isotropic random fields. These recursions are very similar to the Levinson's equations of one-dimensional prediction. The computational complexity of our procedure is $O(BRL^2)$ where B is the bandwidth in the wave-number plane of the spectrum that we want to estimate, and where L is the number of points used to discretize the interval $[0, R]$. Note that our results show that the isotropic MEM spectral estimation problem has a linear solution. It was previously shown in the 2-D discrete space case that the MEM spectral estimation problem has a linear solution whenever the underlying field is Gauss-Markov [23]. However, there is no contradiction between our results and those of [23], since the condition of [23] is only sufficient but not necessary.

This paper is organized as follows. In Section 2 we review some properties of isotropic random fields. In particular, we discuss Fourier expansions of such fields in terms of the angular coordinate θ in a polar representation of the underlying 2-D space. Such expansions will be later used to develop an efficient procedure for constructing the MEM spectrul estimate. In Section 3, we derive an expression for the isotropic MEM estimate. The MEM spectral estimation problem is then related to the problem of finding the best linear filter for estimating an isotropic field on the boundary of a disk given noisy observations of the field inside the disk. By using Fourier expansions of the optimal linear estimation filter and the efficient recursions of [22], a fast and robust method for computing the MEM estimate is developed in Section 4. The numerical implementation of our procedure is described is Section 5. Particular attention is given in this section to the issues of numerical stability and convergence of our implementation. Finally, several examples are presented in Section 6 to illustrate the behavior of our algorithm and particularly to demonstrate its high resolution property.

2 FOURIER SERIES FOR ISOTROPIC FIELDS

In this section, we review some of the properties of isotropic random fields. Specifically, we focus our attention on Fourier series representations of such fields with respect to the angle θ in a polar coordinate representation of the underlying 2-D space.

The covariance function

$$K(\vec{r}) = E[z(\vec{v})z(\vec{v}+\vec{r})] \tag{2.1}$$

of any zero-mean isotropic random field $z(\vec{r})$ [1], is a function of r only, so that, by abuse of notation we can write

$$K(\vec{r}) = K(r). \tag{2.2}$$

Since $K(r)$ is a function of r only, it is straightforward to show that the power spectrum $S(\vec{\lambda})$ of the field $z(\vec{r})$, i.e. the 2-D Fourier transform of $K(\vec{r})$, is actually a function of $\lambda = |\vec{\lambda}|$ only [24], and with a slight abuse of notation we will write this as $S(\lambda)$. Furthermore, it can be shown that $S(\lambda)$ is 2π times the Hankel transform of $K(r)$ viewed as a function of the scalar $r = |\vec{r}|$, i.e.

$$S(\lambda) = 2\pi \int_0^\infty dr\, r\, J_0(\lambda r) K(r), \tag{2.3}$$

where $J_0(\cdot)$ denotes the Bessel function of order 0. By using (2.3) and the addition theorem for Bessel functions [25], we can write

$$K(|\vec{r}-\vec{s}|) = \sum_{n=-\infty}^{\infty} k_n(r,s) e^{jn(\theta-\phi)}, \tag{2.4}$$

where

$$k_n(r,s) = \frac{1}{2\pi} \int_0^\infty J_n(\lambda r) J_n(\lambda s) S(\lambda) \lambda \, d\lambda, \tag{2.5}$$

and where $\vec{r} = (r,\theta)$ and $\vec{s} = (s,\phi)$. In (2.5) $J_n(\cdot)$ is the Bessel function of order n. Alternatively, $k_n(r,s)$ can be computed from $K(\cdot)$ as

$$k_n(r,s) = \frac{1}{2\pi} \int_0^{2\pi} K((r^2+s^2-2rs\cos\theta)^{1/2})\, e^{-jn\theta}\, d\theta. \tag{2.6}$$

[1] Throughout this paper we use $\vec{r}$ to denote a point in 2-D Cartesian space. The polar coordinates of this point are denoted by r and θ.

Note that since $K(\cdot)$ is a real and even function of θ then

$$k_n(r,s) = k_{-n}(r,s). \tag{2.7}$$

Alternatively, (2.7) can be derived by using (2.5) and the fact that

$$J_n(x) = (-1)^n J_{-n}(x). \tag{2.8}$$

Equation (2.7) will prove useful in Section 4 where we develop an efficient method for computing the MEM spectral estimate.

Observe that (2.4) is just an eigenfunction expansion of the positive-definite symmetric function $K(|\vec{r} - \vec{s}|)$ viewed as a function of the scalar variables θ and ϕ. Hence, by using the Karhunen-Loeve theorem [26] we can expand $z(\vec{r})$ as [27]

$$z(\vec{r}) = \sum_{n=-\infty}^{\infty} z_n(r)\, e^{jn\theta}, \tag{2.9}$$

$$z_n(r) = \frac{1}{2\pi}\int_0^{2\pi} z(\vec{r}) e^{-jn\theta}\, d\theta, \tag{2.10}$$

where

$$E[z_n(r) z_m(s)] = k_n(r,s)\delta_{n,m}, \tag{2.11}$$

and where $\delta_{n,m}$ is a Kronecker delta function. Equation (2.9) is very interesting since it can also be interpreted as a Fourier series expansion of the field $z(\vec{r})$ in terms of the coordinate angle θ. In particular, the relations (2.9)-(2.10) indicate that the Fourier coefficient processes $z_n(r)$ in a Fourier series expansion of $z(\vec{r})$ in terms of the angle θ are *independent*. This observation plays a key role in a number of works dealing with isotropic random fields (e.g. [22],[27]) and we shall use it to relate the MEM spectral estimation problem to the filtering problem considered in [22]. Finally, observe that although $z(\vec{r})$ is isotropic, the process $z_n(r)$ is *not* stationary since $k_n(r,s)$ is *not* a function of $r - s$.

3 ISOTROPIC MEM SPECTRAL ESTIMATE

Consider now the following spectral estimation problem. Suppose that we are given the value of the covariance function $K_y(|\vec{r}-\vec{s}|) = E[y(\vec{r})y(\vec{s})]$, of an *isotropic* random field $y(\vec{r})$ for $|\vec{r}-\vec{s}| \leq 2R$, and suppose that we wish to estimate the power spectrum of the "most random" *isotropic* field $y(\cdot)$ whose covariance function is consistent with the set of known values of $K_y(r)$. Furthermore, assume that $y(\vec{r})$ is given by

$$y(\vec{r}) = z(\vec{r}) + v(\vec{r}), \quad \vec{r} \in \mathbf{R}^2 \tag{3.1}$$

where $z(\vec{r})$ is an isotropic zero-mean Gaussian random field with a covariance function $K_z(|\vec{r}-\vec{s}|) = E[z(\vec{r})z(\vec{s})]$, and where $v(\vec{r})$ is a two-dimensional white Gaussian noise of strength P which is uncorrelated with $z(\vec{r})$.

Our problem is really that of extending a radial positive definite function given its values inside a disk of radius $2R$. It is well known [10], [14] that in general, 2-D positive definite functions defined over some finite domain, do not always have a positive definite extension on $\mathbf{R}^2$. However, it was shown in [21] that every radial positive definite function $K(r)$ defined over a disk is *extendable*. Specifically, it is proved in [21] that for every radial positive definite function $K(r)$ defined on a disk of radius $2R$ there exists *radially symmetric positive definite* functions $\hat{K}(r)$ on $\mathbf{R}^2$ such that $\hat{K}(r) = K(r)$ for $r \leq 2R$. Among all such extensions $\hat{K}_y(r)$ of $K_y(r)$, we are looking here for the one whose 2-D Fourier transform $\hat{S}_y(\vec{\lambda})$ maximizes the entropy H of the field $y(\cdot)$ where

$$\begin{aligned} H &= \frac{1}{4\pi^2}\int_{\mathbf{R}^2} d\vec{\lambda}\, \ln \hat{S}_y(\vec{\lambda}) \\ &= \frac{1}{2\pi}\int_0^\infty d\lambda\, \lambda\, \ln \hat{S}_y(\lambda), \end{aligned} \tag{3.2}$$

and where we have used the fact that $\hat{S}_y(\vec{\lambda}) = \hat{S}_y(\lambda)$ since $y(\cdot)$ is an isotropic random field (cf. (2.3)). The exact form of the power spectrum $\hat{S}_y(\lambda)$ that we seek is given in the following theorem.

Theorem 3.1 *The estimated power spectrum $\hat{S}_y(\lambda)$ which maximizes H in (3.2) subject to the positive definiteness constraint*

$$\hat{S}_y(\lambda) \geq 0 \quad \forall \lambda \geq 0 \tag{3.3}$$

and the correlation matching constraint

$$\begin{aligned} \frac{1}{2\pi}\int_0^\infty \hat{S}_y(\lambda)\, J_0(\lambda r)\, \lambda\, d\lambda &= \hat{K}_y(r) \\ &= K_y(r) \quad \textit{for } r \leq 2R \end{aligned} \tag{3.4}$$

is given by

$$\hat{S}_y(\lambda) = \frac{P}{|e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})|^2} \tag{3.5}$$

where $G(R,\vec{\lambda})$ is the 2-D Fourier transform of the function $g(R,\vec{r})$ defined by the integral equation

$$K_z(|\vec{R}_0 - \vec{r}|) = \int_{u\leq R} d\vec{u}\, K_z(|\vec{r} - \vec{u}|) g(R,\vec{u}) + Pg(R,\vec{r}), \quad r \leq R \tag{3.6}$$

and where $\vec{R}_0 = (R,0)$.

The proof of Theorem 3.1 is based on the Lagrange multiplier method for solving constrained optimization problems [28], and on a non-symmetric half plane (NSHP) factorization that we obtain for power spectra corresponding to positive-definite radially symmetric functions that are zero outside a disk of radius $2R$ in the space domain. However, unlike in the 2-D discrete space case [11], [29], the NSHP spectral factor that we find has a bounded support in the space domain, and its spatial support is in fact a disk of radius R. The details of the proof of Theorem 3.1 can be found in Appendix A.

Several comments have to be made at this point. First, note that even though $e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})$ in (3.5) is a function of $\vec{\lambda}$, its magnitude $|e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})|$ is by construction a function of $\lambda = |\vec{\lambda}|$ only, which is consistent with the fact that $\hat{S}_y(\lambda)$ is an isotropic power spectrum. (See Appendix A for details.) Second, if we assume that $K_z(|\vec{r}-\vec{s}|)$ is square integrable then it can be shown that the solution of (3.6) exists and is unique

[30]. Third, observe that $g(R,\vec{r})$ is just the *optimal linear filter for estimating* $z(\vec{R}_0)$ given the observations $y(\vec{r})$ on the disk of radius R centered at the origin. Given the observations $y(\vec{r})$ of (3.1) for $0 \leq r \leq R$, we can express the linear least-squares estimate of $z(\vec{R}_0)$ as

$$\hat{z}(\vec{R}_0) = \int_{u \leq R} d\vec{u}\, y(\vec{u}) h(\vec{u}). \tag{3.7}$$

Using the orthogonality principle of linear least squares estimation [31], we find that the optimal filter $h(\vec{u})$ satisfies the integral equation

$$K_z(|\vec{r} - \vec{R}_0|) = \int_{u \leq R} K_z(|\vec{r} - \vec{u}|) h(\vec{u})\, d\vec{u} + Ph(\vec{r}) \quad r \leq R. \tag{3.8}$$

It then follows from the uniqueness of the solution to (3.6) that

$$h(\vec{u}) = g(R,\vec{u}). \tag{3.9}$$

Hence, solving the 2-D isotropic MEM spectral estimation problem is equivalent to solving a *filtering* problem for the underlying signal field. This is analogous to the 1-D *continuous* time case where the MEM spectral estimation problem is equivalent to a filtering problem for the underlying signal process [32]. In contrast, the MEM spectral estimation problem is equivalent to a *prediction* problem for the underlying signal process [6] in the 1-D *discrete* time case. In the next section we use (3.9) to compute $g(R,\vec{r})$, and hence $\hat{S}_y(\lambda)$, *recursively* as a function of R via the efficient recursions developed in [22] to solve a filtering problem for 2-D isotropic random fields. The notation $g(R,\vec{r})$ (as opposed to $g(\vec{r})$) is used here to stress the dependence of the filter $g(R,\vec{r})$ on the radius $2R$ of the disk over which $K_y(\cdot)$ is given. It is this dependence that will be exploited below to compute $g(R,\vec{r})$ recursively for increasing values of R. In this respect, our method is similar to the 1-D MEM algorithms that use the Levinson equations of 1-D prediction [6] to compute spectral estimates recursively as a function of the size of the interval over which correlation lags are given. Observe also that the choice of the point $\vec{R}_0$ in (3.6) is not restrictive. In fact, we can chose $\vec{R}_0$ to be any point on the boundary of the disk of radius R centered at the origin. Specifically, by using the fact that $z(\cdot)$ is an isotropic random field and the theory of [24], it can be shown [30] that $\hat{S}(\lambda)$ in (3.5) is invariant under rotations

of the vector $\vec{R}_0$. Finally, note that as mentioned earlier, it was previously shown in the 2-D discrete space case that the MEM extension problem has a linear solution whenever the underlying field is a Gauss-Markov random field [23]. According to Theorem 1, the highly non-linear MEM covariance extension problem has a *linear* solution whenever the underlying field is a Gaussian *isotropic* random field regardless of whether it is Gauss-Markov or not. This is not inconsistent with the results of [23], since the condition of [23] is sufficient but not necessary. The existence of a linear solution to the isotropic MEM covariance extension problem should not come as a surprise given that the 1-D MEM stationary covariance extension problem is known to have a linear solution and that isotropy in higher dimensions is the natural extension of stationarity in 1-D. Finally, note that we have so far assumed that $K_y(r)$ is known exactly for $r \leq 2R$. In practice, one is given the observations $y(\cdot)$ over a finite disk, rather than exact values of $K_y(r)$ itself. However, $K_y(r)$ can be estimated directly from the observed data $y(\vec{r})$ by using the procedure of [33] which is summarized in Appendix A. The use of this procedure is illustrated in the second part of Example 6.1 where we compute MEM isotropic power spectral estimates starting from the observations $y(\cdot)$.

4 A FAST ALGORITHM FOR COMPUTING $\hat{S}_y(\lambda)$

In order to use (3.5) to compute the MEM spectral estimate $\hat{S}_y(\lambda)$, we need to know the optimal linear estimation filter $g(R, \vec{r})$ and the noise intensity P. In the 1-D discrete time case where the MEM spectral estimation problem is equivalent to a prediction problem, the constant P appearing in the numerator of the MEM spectral estimate is equal to the variance of the prediction error, and can be computed directly from the known lags of the signal covariance function [6]. In contrast, in the continuous 2-D isotropic case the constant P in (3.5) is equal to the intensity of the observation white Gaussian noise process and cannot be reliably computed from the known values of $K_y(r)$.

A Estimation of the Noise Intensity P

Given the measurements $y(\vec{r})$, the noise intensity P can be estimated by passing $y(\vec{r})$ through a 2-D filter whose wave-number response is zero, or almost zero, within the region of the wave-number plane that contains the spectral support of $z(\vec{r})$. The noise intensity can then be computed from the knowledge of the wave-number response of the filter and from the total power of P_f of the filtered signal. Specifically, if $F(\vec{\lambda})$ is the wave-number response of the filter then P can be estimated as

$$P = \frac{P_f}{\int_{\mathbf{R}^2} |F(\vec{\lambda})|^2 \, d\vec{\lambda}}. \tag{4.1}$$

Note that this approach is analogous to estimating the intensity of a 1-D additive white Gaussian noise process by passing the noisy measurements of a signal of interest through a 1-D bandpass filter followed by an output power measurement stage, with the bandpass filter specifically designed to block the signal.

B Efficient Computation of $g(R,\vec{r})$

Next, the filter $g(R,\vec{r})$ can obviously be computed by discretizing the integral equation (3.6) using any of the rules outlined in ([34], Chapter 5) and by solving the resulting linear equation. Such an approach has two major drawbacks. The first is that it is *computationally* very *expensive* since it requires $O(M^3N^3)$ operations, where M and N are the number of discretization steps used to approximate the integral (3.6) in the angular variable and the radial variable respectively. Secondly, the accumulation of rounding errors and approximation errors made during the numerical computation of $g(R,\vec{r})$ and of its 2-D Fourier transform $G(R,\vec{\lambda})$ can destroy the circular symmetry of the quantity $|e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})|^2$, so that the estimated power spectrum $\hat{S}_y(\lambda)$ can turn out to be *non-isotropic*. Let us now present a computationally efficient procedure for computing $\hat{S}_y(\lambda)$ that has the additional feature of guaranteeing that $\hat{S}_y(\lambda)$ is an isotropic power spectrum. As mentioned earlier, our procedure exploits the relationship between the 2-D isotropic MEM problem and a filtering problem for isotropic random fields to compute $\hat{S}_y(\lambda)$ recursively as a function of the radius $2R$ of the disk over which $K_y(r)$ is known, much in the same spirit as the 1-D MEM algorithms that compute 1-D spectral estimates recursively as a function of the number of the known covariance lags. Our approach is based on a Fourier series expansion of $g(R,\vec{r})$ in the space domain as

$$g(R,\vec{r}) = \sum_{n=-\infty}^{\infty} g_n(R,r)\, e^{jn\theta}, \tag{4.2}$$

and on a corresponding Fourier series expansion for $e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})$ in the wave-number plane. In the remainder of this section we shall show how to compute the coefficients $g_n(R,r)$ efficiently and then use the Hankel transform of those coefficients to compute $\hat{S}_y(\lambda)$ in a robust fashion.

B.1 Interpretation of the Fourier coefficients $\mathbf{g_n(R,r)}$

Substituting (4.2) and (2.4) into (3.6) and equating the Fourier coefficients on both sides of the resulting equation yields the countably infinite

set of integral equations

$$k_n(r,R) = 2\pi \int_0^\infty k_n(r,u)\, g_n(R,u)\, u\, du + P g_n(R,u) \quad 0 \le r \le R. \qquad (4.3)$$

Equation (4.3) is quite interesting because it also arises in the context of filtering for isotropic random fields [22]. In particular, the Fourier series expansions (2.9) are used in [22] to convert the 2-D problem of estimating the value of $z(R,\theta)$ on the boundary of a disk of radius R given the observations $y(\vec{r})$ inside the disk, into a countably infinite number of 1-D estimation problems where the objective is to estimate each of the signal Fourier coefficient processes $z_n(R)$ given the corresponding observations Fourier coefficient processes $y_n(r)$ on the interval $0 \le r \le R$. By comparing (4.3) with equation (2.21) of [22], it becomes clear that the coefficient $g_n(R,r)$ is only a scaled version of the optimum linear filter for estimating $z_n(R)$ given $\{y_n(s) : 0 \le s \le R\}$. Furthermore, it is shown in [22] that the optimum linear filter for estimating $z_n(R)$ given $\{y_n(s) : 0 \le s \le R\}$ obeys a quasilinear hyperbolic system [35], [36] of partial differential equations which when properly scaled take the form

$$\begin{aligned} (\frac{\partial}{\partial R} - \frac{n}{R}) g_n(R,r) + (\frac{\partial}{\partial r} + \frac{(n+1)}{r}) g_{n+1}(R,r) &= \\ -\rho_n(R) g_n(R,r) & \qquad (4.4) \\ (\frac{\partial}{\partial r} - \frac{n}{r}) g_n(R,r) + (\frac{\partial}{\partial R} + \frac{(n+1)}{R}) g_{n+1}(R,r) &= \\ \rho_n(R) g_n(R,r), & \qquad (4.5) \end{aligned}$$

with

$$\rho_n(R) = \frac{R}{2\pi}(g_n(R,R) - g_{n+1}(R,R)) \qquad (4.6)$$

and with the initial conditions

$$\begin{aligned} \frac{\partial}{\partial r} g_0(R,r)|_{r=0} &= 0, & (4.7) \\ g_n(R,0) &= 0, \quad \text{for } n \neq 0. & (4.8) \end{aligned}$$

Note that as claimed earlier the coefficients $g_n(R,r)$, and hence the filter $g(R,\vec{r})$ and the power spectral estimate $\hat{S}_y(\lambda)$, can be computed recursively

as a function of the radius $2R$ of the disk over which $K_y(r)$ is given via (4.4)-(4.5). In this respect, equations (4.4)-(4.5) are similar to the Levinson recursions of 1-D prediction. Equations (4.4)-(4.5) can be derived by exploiting the special structure of $k_n(r,s)$ as displayed by equation (2.5), and by using the properties of Bessel function (see [22] for details). The numerical computation of $g_n(R,r)$ via (4.3)-(4.5) has to be performed with some care. In particular, one has to study carefully the stability and convergence properties of any numerical method used to solve the coupled partial differential equations (4.4)-(4.5) [35], [36]. In Section 5 we present a stable and convergent numerical method for computing $g_n(R,r)$. Our method is computationally very efficient and requires $O(L^2)$ operations where L is the number of discretization points in the interval $[0,R]$ where we want to compute $g_n(R,r)$.

B.2 Fourier expansions in the wave-number plane

Next, we expand $e^{-j\vec{\lambda}\cdot\vec{R_0}} - G(R,\vec{\lambda})$ in terms of the angle ϕ defined by $\vec{\lambda} = (\lambda, \frac{\pi}{2} - \phi)$ in a polar representation of the wave-number space. Then, by using the theory of [24, Chapter 5] and the expansion

$$e^{-j\vec{\lambda}\cdot\vec{R_0}} = \sum_{n=-\infty}^{\infty} J_n(\lambda R)\, e^{-jn\phi} \tag{4.9}$$

we can write

$$e^{-j\vec{\lambda}\cdot\vec{R_0}} - G(R,\vec{\lambda}) = \sum_{n=-\infty}^{\infty} \left(J_n(\lambda R) - 2\pi G_n(R,\lambda)\right) e^{-jn\phi}. \tag{4.10}$$

In (4.10), $G_n(R,\lambda)$ is the nth- order Hankel transform of $g_n(R,r)$ [24], i.e.

$$G_n(R,\lambda) = \int_0^{\infty} g_n(R,r) J_n(\lambda r)\, r\, dr. \tag{4.11}$$

Since the magnitude of $e^{-j\vec{\lambda}\cdot\vec{R_0}} - G(R,\vec{\lambda})$ is a function of λ only, it follows from (4.10) that

$$|e^{-j\vec{\lambda}\cdot\vec{R_0}} - G(R,\vec{\lambda})|^2 = \sum_{n=-\infty}^{\infty} |J_n(\lambda R) - 2\pi G_n(R,\lambda)|^2. \tag{4.12}$$

Equation (4.12) is a little surprising at first sight because it claims that the square magnitude of a function of the variable ϕ is equal to the sum of the square magnitudes of its Fourier coefficients in a Fourier expansion in terms of ϕ. However, the functions that we consider here have a very special structure since their magnitude is *not* a function of ϕ by construction. An example of such functions is provided by the function $\lambda \sin\phi - j\sqrt{\lambda^2 \cos^2\phi + 1}$ whose squared magnitude λ^2+1 depends on λ only. Observe also that (4.12) implies that $e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})$ is an "all-pass" function of the variable ϕ.

A further simplification of (4.12) is possible by noting that (2.7) together with the uniqueness of the solution of (4.3) [22] imply that

$$g_n(R,r) = g_{-n}(R,r) \tag{4.13}$$

Hence, it follows from the fact that

$$J_{-n}(\lambda r) = (-1)^n J_n(\lambda r) \tag{4.14}$$

that

$$G_{-n}(R,\lambda) = (-1)^n G_n(R,\lambda). \tag{4.15}$$

By combining (4.12) and (4.14)-(4.15) we can rewrite (4.12) as

$$\begin{aligned} |e^{-j\vec{\lambda}\cdot\vec{R}_0} - G(R,\vec{\lambda})|^2 &= |J_0(\lambda R) - 2\pi G_0(R,\lambda)|^2 \\ &+ 2\sum_{n=1}^{\infty} |J_n(\lambda R) - 2\pi G_n(R,\lambda)|^2. \end{aligned} \tag{4.16}$$

In practice, of course, one would compute only a finite number $N+1$ of the coefficient functions $G_n(R,\lambda)$ and one would obtain an approximation to the estimated power spectrum $\hat{S}_y(\lambda)$ as

$$\hat{S}_y(\lambda) \approx \frac{P}{|C_N(R,\lambda)|^2}, \tag{4.17}$$

where

$$\begin{aligned} |C_N(R,\lambda)|^2 &= |J_0(\lambda R) - 2\pi G_0(R,\lambda)|^2 \\ &+ 2\sum_{n=1}^{N} |J_n(\lambda R) - 2\pi G_n(R,\lambda)|^2. \end{aligned} \tag{4.18}$$

The number N can be determined by noting that

$$J_n(x) \approx 0 \quad \text{for } x \gg 1 \text{ and } n > x. \tag{4.19}$$

Hence, if we are interested in computing $\hat{S}_y(\lambda)$ over the disk $\lambda \leq B$ in the wave-number plane, we can take $N = BR$ provided that $BR \gg 1$, and in this case (4.17)-(4.18) give a very good approximation to $\hat{S}_y(\lambda)$.

Let us now make a few comments. First, note that (4.17)-(4.18) guarantee that $\hat{S}_y(\lambda)$ is isotropic since (4.18) involves a sum of positive terms that depend on λ only. Second, observe that the nth-order Hankel transforms in (4.11) can be implemented efficiently by using any of the existing fast Hankel transform algorithms [37], [38], [39]. These techniques require $O(L \ln L)$ operations, where L is the number of discretization points at which $g_n(R,r)$ is available. Hence, our procedure for constructing $\hat{S}_y(\lambda)$ requires $O(L^2)$ per coefficient and its complexity in practice is $O(BRL^2)$ operations. Finally, note that in our procedure, the coefficients $g_n(R,r)$ are computed recursively as a function of R via (4.4)-(4.5), so that the spectral estimate $\hat{S}_y(\lambda)$ can be easily updated whenever new measurements become available, i.e. as the disk radius R is increased.

C Summary

The procedure for computing $\hat{S}_y(\lambda)$ approximately can therefore be summarized as follows:

1. Estimate $K_y(r)$ for $0 \leq r \leq 2R$ and $k_n(r,s)$ for $0 \leq r,s \leq R$ and for $|n| \leq N$, from the given data using the procedure outlined in [28] and summarized in Appendix B.

2. Use a stable and convergent numerical method, such as the one appearing in the next section, to compute $g_n(R,r)$ recursively from equations (4.3)-(4.5) for $n \leq N$ and for a suitably chosen N.

3. Evaluate the nth-order Hankel transforms $G_n(R,\lambda)$ by using a fast Hankel transform method.

4. Compute an approximation to $\hat{S}_y(\lambda)$ via equations (4.17)-(4.18).

5 NUMERICAL COMPUTATION OF THE COEFFICIENTS $g_n(R,r)$

Recall that the fast algorithm that we proposed in the last section for computing $\hat{S}_y(\lambda)$ involves the solution of the quasilinear hyperbolic system of partial differential equations (4.4)-(4.5). It is quite possible to discretize a system of partial differential equations in an apparently natural way and yet obtain completely erroneous computational results. This is especially true for propagation problems described by parabolic and hyperbolic equations. The reason for this numerical ill-behavior is that round-off and other computational errors coupled with a bad choice of a discretization scheme may lead to both numerical instability and convergence problems. In this section, we present a stable and convergent method for computing $g_n(R,r)$ via (4.3)-(4.5). Our approach is based on the *method of characteristics* for solving hyperbolic partial differential equations [35], [36]. The basic idea is to replace the original system of hyperbolic partial differential equations with an equivalent system of differential equations each involving differentiation in only one of the variables of an appropriate coordinate system. The resulting system can then be solved in a well-behaved, stable and convergent manner. Specifically, let us consider a new coordinate system α, β defined by

$$\alpha = R + r \tag{5.1}$$

$$\beta = R - r \tag{5.2}$$

Equations (4.4)-(4.5) can now be rewritten in the new coordinate system as

$$\begin{aligned} \frac{\partial}{\partial\alpha}g_n(\alpha,\beta) + \frac{\partial}{\partial\alpha}g_{n+1}(\alpha,\beta) &= \left(\frac{4\alpha n}{\alpha^2-\beta^2} - \rho_n\left(\frac{\alpha+\beta}{2}\right)\right)g_n(\alpha,\beta) \\ &+ \left(\rho_n\left(\frac{\alpha+\beta}{2}\right) - \frac{4\alpha(n+1)}{\alpha^2-\beta^2}\right)g_{n+1}(\alpha,\beta) \end{aligned} \tag{5.3}$$

$$\frac{\partial}{\partial\beta}g_n(\alpha,\beta) - \frac{\partial}{\partial\beta}g_{n+1}(\alpha,\beta) = -(\frac{4\beta n}{\alpha^2-\beta^2} + \rho_n(\frac{\alpha+\beta}{2}))g_n(\alpha,\beta) - (\rho_n(\frac{\alpha+\beta}{2}) + \frac{4\beta(n+1)}{\alpha^2-\beta^2})g_{n+1}(\alpha,\beta) \tag{5.4}$$

Note that in the new coordinate system each partial differential equation involves differentiation with respect to only one of the independent variables α and β. Refering to Fig. 1, we see that given the values of $g_n(R,r)$ and $g_{n+1}(R,r)$ on the line AB we can compute $g_n(R,r)$ and $g_{n+1}(R,r)$ within the triangle ABC by integrating (5.3) and (5.4) along the *characteristic* directions α = constant (for (5.4)) and β = constant (for (5.3)), or equivalently along lines of slope $\pm 45°$ in the (R,r) plane. Specifically, if the values of $g_n(R,r)$ and $g_{n+1}(R,r)$ have been computed inside the triangle OAB (see Fig. 1), and in particular on the line AB, then by integrating equation (5.3) along β = constant lines starting on AB, we can compute the sum $g_n(R,r) + g_{n+1}(R,r)$ inside the parallelogram ABGF. Similarly, by integrating (5.4) along α = constant directions starting on AB, we can compute the difference $g_n(R,r) - g_{n+1}(R,r)$ inside the region ABED. Thus, $g_n(R,r)$ and $g_{n+1}(R,r)$ can be uniquely determined within the triangle ABC (the intersection of regions ABED and ABGF). The values of $g_n(R,r)$ and $g_{n+1}(R,r)$ which are outside triangle ABC, will have to be computed using the integral equation (4.3). Our numerical procedure is based on equations (4.3) and (5.3)-(5.4). To compute $g_n(R^*,r)$ and $g_{n+1}(R^*,r)$ for $0 \le r \le R^*$, we divide the interval $[0,R^*]$ into L subintervals of length $\Delta = R^*/L$. If we denote by $G_n(k,l) = g_n(k\Delta,l\Delta)$, and if at stage k we assume that $G_n(k,l)$ and $G_{n+1}(k,l)$ have been computed for $0 \le l \le k$ (i.e. on the line AB of Fig. 1), then $G_n(k,l)$ and $G_{n+1}(k+1,l)$ can be evaluated for $0 < l \le k-1$ by integrating equations (5.3)-(5.4) along the characteristic directions R = constant $\pm\, r$. For $l = k, k+1$ (i.e. outside of the triangle ABC), $G_n(k+1,l)$ and $G_{n+1}(k+1,l)$ can be computed by solving a two by two linear system obtained by discretizing the integral equation (4.3) (see Fig. 1). Specifically, if we use a simple Euler difference method to integrate (5.3)-(5.4), and solving for $G_n(k+1,l)$ and $G_{n+1}(k+1,l)$, we obtain the following recursions for $0 < l \le k-1$,

$$
\begin{aligned}
G_n(k+1,l) = \quad & [\frac{1}{2} + \frac{n}{\sqrt{2}k} + \frac{n}{\sqrt{2}(l-1)} - \frac{\Delta}{\sqrt{2}}\rho_n(k)]G_n(k,l-1)) \\
+ \; & [\frac{1}{2} - \frac{n+1}{\sqrt{2}k} - \frac{n+1}{\sqrt{2}(l-1)} + \frac{\Delta}{\sqrt{2}}\rho_n(k)]G_{n+1}(k,l-1)) \\
+ \; & [\frac{1}{2} + \frac{n}{\sqrt{2}k} - \frac{n}{\sqrt{2}(l+1)} - \frac{\Delta}{\sqrt{2}}\rho_n(k)]G_n(k,l+1)) \\
+ \; & [-\frac{1}{2} + \frac{n+1}{\sqrt{2}k} - \frac{n+1}{\sqrt{2}(l+1)} - \frac{\Delta}{\sqrt{2}}\rho_n(k)]G_{n+1}(k,l+1))
\end{aligned}
\tag{5.5}
$$

$$
\begin{aligned}
G_{n+1}(k+1,l) = \quad & [\frac{1}{2} + \frac{n}{\sqrt{2}k} + \frac{n}{\sqrt{2}(l-1)} - \frac{\Delta}{\sqrt{2}}\rho_n(k)]G_n(k,l-1)) \\
+ \; & [\frac{1}{2} - \frac{n+1}{\sqrt{2}k} - \frac{n+1}{\sqrt{2}(l-1)} + \frac{\Delta}{\sqrt{2}}\rho_n(k)]G_{n+1}(k,l-1)) \\
+ \; & [-\frac{1}{2} - \frac{n}{\sqrt{2}k} + \frac{n}{\sqrt{2}(l+1)} + \frac{\Delta}{\sqrt{2}}\rho_n(k)]G_n(k,l+1)) \\
+ \; & [\frac{1}{2} - \frac{n+1}{\sqrt{2}k} + \frac{n+1}{\sqrt{2}(l-1)} + \frac{\Delta}{\sqrt{2}}\rho_n(k)]G_{n+1}(k,l+1))
\end{aligned}
\tag{5.6}
$$

where

$$
\rho_n(k) = \frac{k\Delta}{2\pi}(G_n(k,k) - G_{n+1}(k,k)). \tag{5.7}
$$

Similarly, if we discretize equation (4.3) using the trapezoidal rule, we obtain for $l = k,\ k+1$

$$
\begin{aligned}
PG_n(k+1,l) = k_n((k+1)\Delta, l\Delta) \quad & - \quad 2\pi \sum_{i=1}^{k} k_n(l\Delta, i\Delta) G_n(k+1,i)\, i\, \Delta^2 \\
& - \quad \pi k_n(l\Delta,(k+1)\Delta) G_n(k+1,k+1)(k+1)\Delta^2.
\end{aligned}
\tag{5.8}
$$

Other integration rules can be used as well, instead of the ones we have chosen. Note that our algorithm involves only numerical integration of ordinary differential equations and thus can be implemented in a well behaved, stable and convergent manner. Furthermore, it can be checked that this approach requires $O(L^2)$ operations per Fourier coefficient $g_n(R,r)$.

6 EXAMPLES

In this section we present three examples to illustrate the behavior of our 2-D isotropic MEM procedure. The first example is meant to illustrate the high resolution property of our procedure using both exact and estimated covariance data for the case of a signal power spectrum consisting of two cylindrical impulses in an additive white Gaussian noise. In particular, we use exact covariance values in the first part of this example to demonstrate the high resolution property of our algorithm and to study the effect of increasing the radius of the disk over which the covariance function is given on the spectral estimates that we obtain. In the second part of this example, we generate a random field with the desired covariance function and use the procedure of Appendix A and the method of Sections 4 and 5 to compute MEM spectral estimates. The results that we obtain show that our procedure does not seem to suffer from the spectral line splitting problem observed in 1-D MEM spectral estimates ([6], Section 2.E). In the second example we use exact covariance data corresponding to a smooth signal spectrum to study the effect of varying the number $N+1$ of terms used in (4.17)-(4.18) to compute the 2-D isotropic MEM estimate. Finally, the third example illustrates the behavior of our algorithm when dealing with a signal that has a power spectrum consisting of both a smooth and an impulsive component.

Example 6.1

To demonstrate the resolution capability of our algorithm let us consider a signal power spectrum consisting of two cylindrical impulses which are spaced closer than the classical Fourier resolution limit of π/R, where $2R$ is the radius of the disk over which the covariance function is given. Specifically, consider the signal covariance function

$$K_z(r) = 10J_0(0.2r) + 10J_0(0.4r) \tag{6.1}$$

given over a disk of radius 20 meters. Covariance functions of the form $AK_0(Br)$ are often used in seismology [15] and in ocean acoustics [16], [17] to model some types of background noise fields. Note that $K_z(r)$ corresponds to a power spectrum consisting of two cylindrical impulses at 0.2

rad/m and 0.4 rad/m, i.e.

$$S_z(\lambda) = 50\delta(\lambda - 0.2) + 25\delta(\lambda - 0.4). \tag{6.2}$$

Observe also that the separation between the two cylindrical impulses is smaller than the resolution limit of any classical spectral estimation method, which is of the order of 0.3 rad/m in this case. Furthermore, let us assume that the additive white noise intensity P is 3 Watt.m^2 . Thus the total noise power in the wave-number band [0,0.5] rad/m is 6.27 dB lower than that of either impulses. The true power spectrum of the observations (i.e. of the signal plus noise field) is shown in Fig. 2. Fig. 3, 4 and 5 show the estimated power spectra that we obtain when the order N of the highest Fourier coefficients that we use in (4.18) is 1, 5 and 10 respectively. Note that even with $N = 1$ we can clearly see two peaks at the correct impulse locations (Fig. 3). Note also that the power spectrum that we obtain with $N = 1$ is more 'peaky' than those obtained with $N = 5$ and $N = 10$ (Figs. 4 and 5). In general, we have observed that when we used only two terms in (4.18) the power spectra that we obtained were highly peaked and that sometimes spurious peaks tended to appear (e.g. in Example 6.2). Hence, even if one is interested in estimating a highly peaked spectrum consisting of pure cylindrical impulses in an additive white Gaussian noise, one should use a higher value of N to make sure that none of the peaks that appear when $N = 1$ are spurious in nature. Next note that in Fig. 4 where we used $N = 5$, the amplitude of the peak around 0.4 rad/m is larger than that of the peak around 0.2 rad/m. By looking at the form of the true signal power spectrum (eq. (6.2)) one would expect the amplitude of the peak at 0.2 rad/m to be twice as large as the one of the peak at 0.4 rad/m. As seen in Fig.5, this actually happens when we use a value of N which is large enough to make the approximate spectrum that we compute via (4.17)-(4.18) very close to the true MEM spectral estimate. Finally, observe that the estimated spectrum of Fig. 5 is relatively smooth. This is to be expected since the MEM power spectral estimate is the smoothest of all possible spectra that satisfy the correlation matching constraint. While MEM does a good job of resolving the peaks of the power spectrum of this example, one might prefer to use the method of [33] if the spectra of interest are exclusively of the impulsive form of (6.2). This corresponds to

using Pisarenko's method [40] or the MUSIC method [41], [42] rather than the MEM method in the 1-D case to estimate power spectra corresponding to a sum of sinusoids in a white Gaussian noise.

Next to study the effect of the radius $2R$ of the disk over which the covariance $K_y(r)$ is assumed to be given we double the value of $2R$ from 20 meters to 41 meters. The power spectrum that we obtain in this case using 21 terms in (4.18) is plotted in Fig. 6. Note that this spectral estimate is quite peaky and that the peak at 0.2 rad/m is twice as large as the one at 0.4 rad/m. This improvement is quite natural and in fact as $2R$ tends to infinity, one expects to be able to reconstruct the power spectrum exactly.

Finally, to study the behavior of our algorithm when data measurements, rather than exact correlation measurements, are given, we synthesized an isotropic random field with a power spectrum of the form (6.2) using the method described in [43]. We then added to the resulting field a white Gaussian noise field of intensity 3 Watt.m^2. Using the value of the observations $y(\cdot)$ over disks of various radii, we obtained estimates of the covariances $K_y(r)$ and $k_n(r,s)$ using the spatial averaging procedure of Appendix B. Particular attention was given in this step to the numerical computation of an estimate of $k_n(r,s)$ via (B.3) to avoid the possible errors that may have resulted from the highly oscillatory nature of the integrand. In our experiments we used an integration rule based on Filon's procedure [34] to implement (B.3) numerically. Fig. 7 is a plot of the power spectrum that we obtain when we use the observations available over a disk of radius 100 m to estimate $K_y(r)$ for $0 \leq r \leq 20$ and $k_n(r,s)$ for $0 \leq r,s \leq 10$ and for $0 \leq n \leq N = 5$ and then feed those estimates as an input to our algorithm. Note the small bias in the position of the spectral peaks which are now located at 0.215 rad/m and 0.40 rad/m respectively. Fig. 8 shows the power spectrum that we obtain when we use the observations available over a disk of radius 30 m to estimate $K_y(r)$ for $0 \leq r \leq 20$ and $k_n(r,s)$ for $0 \leq r,s \leq 10$ and for $0 \leq n \leq N = 5$. Observe that the peak at 0.4 rad/m is now barely visible and that the peak at 0.2 rad/m is displaced to about 0.185 rad/m. This degradation in the quality of our spectral estimate is not surprising since we are now using less accurate estimates of $K_y(r)$ as an input to our procedure.

We conclude this example by computing a power spectral estimate using

the values of $K_y(r)$ for $0 \leq r \leq 20$ estimated from the observations inside a disk of radius 30 when the noise intensity is only 0.0001 Watt.m^2 instead of 3 Watt.m^2. Note that the total noise power in the interval [0,0.5] rad/m is now 51 dB lower than that of each of the two cylindrical impulses. The spectrum that we obtain in this case is plotted in Fig. 9. Note the definite presence of the two peaks which are now displaced to about 0.18 rad/m and 0.408 rad/m respectively. However, no line splitting is observed. In the 1-D case, MEM algorithms have been observed to yield two close peaks where only one is present whenever the underlying signal is a pure sinusoid with a weak additive noise component [6]. This phenomenon is called the line splitting problem and is more pronounced when the initial phase of the sinusoid is an odd multiple of $\pi/4$ and when the signal-to-noise ratio is high. In the 2-D isotropic case, there is no corresponding initial phase effect. Furthermore, our procedure is based on the computation of the filters $g_n(R,r)$ and thus conceptually involves minimizing the estimation error in all possible directions. Hence, our procedure corresponds to the 1-D MEM algorithms based on minimizing the forward and backward prediction errors [6]. Such approaches are known to alleviate the line splitting problem in the 1-D case.

Example 6.2

In this example we study the effect of varying the number of terms in the series (4.18). Consider a signal covariance function of the form

$$K_z(r) = 0.25rK_1(0.25r) \tag{6.3}$$

where $K_1(x)$ is a modified Bessel function of first order [25], and assume that the noise intensity is 1 Watt.m^2. Covariance functions of the form $ArK_1(Ar)$ have been used in hydrology to model the correlation structure of rainfall [20]. The power spectrum corresponding to such covariance functions is smooth and is given by

$$S_z(\lambda) = \frac{2A^2}{(\lambda^2 + A^2)^2} \tag{6.4}$$

The true power spectrum corresponding to the observations (i.e. the signal plus noise fields) for this example is shown in Fig. 10.

To study the effect of N in (4.18) on the shape of the estimated power spectrum we fixed $2R$ to be 20 meters. Fig. 11 shows the power spectrum that we obtain when we pick $N = 1$. Note the presence of ripples in this case. Such ripples can easily be mistaken for cylindrical impulses of the type discussed in Example 6.1 . With $N = 3$ we obtain the power spectrum of Fig. 12. Note that this estimate is quite smooth. However, a spurious small and broad peak is still visible around 0.56 rad/m. If we pick $N = 10$ we obtain the power spectrum shown in Fig. 13. Comparing Fig. 10 and 13 we see that this estimate is good except around the origin of the wave-number plane. Experimental results indicate that the quality of our spectral estimates close to the origin improves with the number of discretization points used. In this example we used 100 discretization points and the quality of the spectral estimate that we obtained is good for $\lambda > 0.1$ rad/m. To get better spectral estimates close to the origin one needs to use a very large number of discretization points. For example, when we increased the number of discretization points from 100 to 150 we obtained only a slight improvement over the case that we show here. Finally, note that in this case $B = 1$ and $R = 10$ so that $N = BR = 10$.

To conclude, one should compute 2-D isotropic MEM spectral estimates via (4.17)-(4.18) by gradually increasing the number $N + 1$ of terms used until the resulting estimates stop changing noticeably as N is increased. In general, this requires computing roughly $BR + 1$ terms as mentioned in Section 4.

Example 6.3

In this last example, we illustrate the behavior of our algorithm when the signal power spectrum consists of both a smooth component and an impulsive component. Assume that we are given the values of the observations covariance function

$$K_y(r) = 0.25rK_1(0.25r) + AJ_0(0.6r) + \delta(\vec{r}) \tag{6.5}$$

over the disk $r \leq 20$. Note that the noise intensity is equal to 1 Watt.m ^{2}

and that the signal power spectrum consists of both a smooth part of the form (6.4) and a cylindrical impulse at 0.6 rad/m. The true power spectrum of the observations is shown in Fig. 14.

With $A = 1$, i.e. with the total power in the smooth part of the signal power spectrum equal to that of the impulsive part, we obtain the estimate shown in Fig. 15 when we pick $N = 10$ in (4.18). Note that the presence of the impulsive component around 0.6 rad/m is barely visible.

When $A = 10$, i.e. when the total power in the impulsive component of the signal power spectrum is 10 dB higher than that of the smooth part, we obtain the estimate plotted in Fig. 16 when we choose $N = 10$ in (4.18). Observe that the presense of the impulsive component is now well marked.

In conclusion, these experimental results, and others we have obtained, indicate that the spectral estimate computed via the technique that we propose depends strongly on the size of the interval over which the observations covariance function $K_y(r)$ is known and on the accuracy of the estimates of $K_y(r)$ that are used. The number of terms that have to be used in (4.18) is on the order of $BR + 1$ where $2R$ is the radius of the disk over which $K_y(r)$, or its estimate is known, and where B is the bandwidth in the wave-number plane of the spectrum that we want to estimate. Finally, our procedure does not seem to suffer from the line splitting problem observed with 1-D MEM algorithms

7 CONCLUSION

In this paper, we have presented a new linear MEM algorithm for 2-D isotropic random fields. Our procedure differs from previous 2-D MEM algorithms by the fact that we take maximal advantage of the symmetries implied by isotropy which is the natural generalization to several dimensions of the 1-D notion of stationarity. Unlike general 2-D covariances, isotropic covariance functions which are positive definite on a disk are known to be extendible. Here, we have developed a computationally efficient procedure for computing the MEM isotropic spectral estimate corresponding to an isotropic covariance function which is given over a finite disk of radius $2R$. We have shown that the isotropic MEM problem has a linear solution which can be obtained by constructing the optimal linear filter for estimating the underlying isotropic field at a point on the boundary of a disk of radius R given noisy measurements of the field inside the disk. Our procedure is based on Fourier series expansions in both the space and wave-number domains of the inverse of the MEM spectral estimate. Furthermore, our method is guaranteed to yield a valid isotropic spectral estimate and it is computationally efficient since it requires only $O(BRL^2)$ operations where L is the number of points used to discretize the interval $[0, R]$, and where B is the bandwidth of the spectrum that we want to estimate. Finally, we have presented examples to illustrate the behavior of our algorithm and its high resolution property.

There are several directions in which one can try to generalize this work. For example, 2-D covariance functions which are constant along ellipses rather than along circles, arise in some cases of practical interest. Such covariance functions become radially symmetric under an appropriate scaling and rotation of the underlying coordinate axes, and the techniques of this paper can then be used to estimate a warped version of the power spectrum of the underlying random field. A challenging problem here is to develop an algorithm for finding the correct scaling and rotation operations to be performed given limited measurements of the random field. More generally, another interesting problem is to extend some of the ideas that appear throughout this paper to homogeneous, but not necessarily isotropic, covariance functions which are defined continuously over the plane. This will

require a study of filtering problems for homogeneous fields aimed at determining whether the homogeneity property can be exploited in higher dimensional spaces to develop fast filtering algorithms.

APPENDIX A

Proof of Theorem 3.1.

The problem that we consider in Section 3 is mathematically the problem of finding the $\hat{S}_y(\lambda)$ that maximizes the entropy H

$$H = \frac{1}{2\pi}\int_0^\infty d\lambda\, \lambda\, \ln S_y(\lambda), \tag{A.1}$$

subject to the positive definiteness and correlation matching constraints

$$\text{(i)} \quad \hat{S}_y(\lambda) \geq 0 \quad \text{for } \lambda \geq 0, \tag{A.2}$$

$$\text{(ii)} \quad \frac{1}{2\pi}\int_0^\infty \hat{S}_y(\lambda)\, J_0(\lambda r)\, \lambda\, d\lambda = K_y(r) \quad \text{for } r \leq 2R. \tag{A.3}$$

By using the approach outlined in [28] for solving optimization problems with global pointwise inequality constraints, we find that the MEM power spectral estimate $\hat{S}_y(\lambda)$ is given by

$$\hat{S}_y(\lambda) = \frac{1}{A(2R,\lambda)} \tag{A.4}$$

where

$$A(2R,\lambda) = \int_0^{2R} a(2R,r)\, J_0(\lambda r)\, r\, dr, \tag{A.5}$$

and where $a(2R,r)$ is the Lagrange multiplier function associated with the constraints (A.2)-(A.3). Observe that $A(2R,\lambda)$ can be interpreted as being the zeroth-order Hankel transform of the function $a(2R,r)$ which is zero outside the disk $r \leq 2R$. Note also that

$$\hat{S}_y(\lambda) = P + \hat{S}_z(\lambda) \tag{A.6}$$

where $\hat{S}_z(\lambda)$ is the estimated power spectrum of the process $z(\cdot)$. Hence, if we assume that $K_z(0)$ is finite we must have

$$\lim_{\lambda\to\infty} \hat{S}_z(\lambda) = 0, \tag{A.7}$$

for otherwise the integral

$$\int_0^\infty \hat{S}_z(\lambda)\,\lambda\,d\lambda = K_z(0) \tag{A.8}$$

would fail to converge. Taking (A.7) into account, we can rewrite (A.4) as

$$\hat{S}_y(\lambda) = \frac{P}{1 - B(2R,\lambda)} \tag{A.9}$$

where

$$\lim_{\lambda\to\infty} B(2R,\lambda) = 0, \tag{A.10}$$

and

$$B(2R,\lambda) < 1 \tag{A.11}$$

since $\hat{S}_y(\lambda)$ is strictly positive for all λ. Note that (A.4)-(A.5) and (A.9) imply that $B(2R,\lambda)$ is the Hankel transform of a function $b(2R,\vec{r})$ that is zero outside the disk of radius $2R$ centered at the origin of the plane. Now let $\vec{\lambda} = (\lambda_1,\lambda_2)$ in a Cartesian representation of the wave-number plane and consider the function $1 - B(2R,\sqrt{\lambda_1^2+\lambda_2^2})$ viewed as a function of λ_1 only (i.e. with λ_2 fixed). Then (A.11) implies that $1 - B(2R,\sqrt{\lambda_1^2+\lambda_2^2})$ is strictly positive for all values of λ_1, so that we can use the results of [44] to factor $1 - B(2R,\lambda)$ as

$$(1 - B(2R,\lambda)) = (1 - F(R,\lambda_1,\lambda_2))(1 - F^*(R,\lambda_1,\lambda_2)) \tag{A.12}$$

where $F(R,\lambda_1,\lambda_2)$ is the 2-D Fourier transform of a real function $f(R,\vec{r})$ [1] that is causal in the Cartesian coordinate r_1, where $\vec{r} = (r_1,r_2)$, i.e. where

$$f(R,\vec{r}) = 0, \quad \text{for } r_1 < 0. \tag{A.13}$$

Substituting (A.12) into (A.9) we obtain

$$\hat{S}_y(\lambda) = \frac{P}{(1 - F(R,\lambda_1,\lambda_2))(1 - F^*(R,\lambda_1,\lambda_2))}. \tag{A.14}$$

[1]The reason for the notation $f(R,\vec{r})$ as opposed to $f(2R,\vec{r})$ will become clear in the sequel.

Equation (A.14) is the continuous space version of the well known result in the discrete space case [11], [29] that any power spectral density function $S(e^{ju}, e^{jv})$ which is strictly positive for all $(u, v) \in [-\pi, \pi]^2$ can be written in factored form as

$$S(z_1, z_2) = \frac{\sigma^2}{A(z_1, z_2)A^*(z_1, z_2)} \qquad (A.15)$$

where the filter $A(z_1, z_2)$ has a nonsymmetric half plane support. In fact (A.14) could have been derived by using (A.4), (A.15) and the transformations

$$z_1 = \frac{1 + j\lambda_1}{1 - j\lambda_1} \qquad (A.16)$$

$$z_2 = \frac{1 + j\lambda_2}{1 - j\lambda_2} \qquad (A.17)$$

where λ_1 and λ_2 are allowed to take complex values. The transformations (A.16)-(A.17) are analogous to those which are used in the 1-D context to map the continuous time case into the discrete time case and vice-versa. However, unlike in the discrete space case [11], [29], where $A(z_1, z_2)$ often corresponds to a filter with an unbounded spatial support, the filter $f(R, \vec{r})$ has a finite support in the spatial domain. According to Theorem 3.4.2 in [45] and which is originally due to Plancherel and Polya [46], $B(2R, |\vec{\mu}|)$ where $\vec{\mu} \in \mathbf{C}^2$, must be an entire function of exponential type since it is equal to the Fourier transform of a function that is zero outside the disk of radius $2R$ in $\mathbf{R}^2$. Hence, $F(R, \vec{\mu})$ must also be an entire function of exponential type and must therefore be the Fourier transform of a function that is zero outside a bounded domain by the above mentioned theorem. Let us now study the spatial support $\mathcal{D}$ of the filter $f(R, \vec{r})$. Equation (A.12) implies that

$$b(2R, \vec{r}) = f(R, \vec{r}) + f(R, -\vec{r}) - \int_{\mathcal{D}} f(R, \vec{r'}) f(R, \vec{r} + \vec{r'}) d\vec{r'}. \qquad (A.18)$$

Since $b(2R, \vec{r})$ is zero for $r > 2R$, and since $f(R, \vec{r})$ and $f(R, -\vec{r})$ appear on the right hand side of (A.18), then $\mathcal{D}$ must lie inside the half disk $\{\vec{r} : \quad r < 2R \quad \text{and} \quad \frac{-\pi}{2} < \theta < \frac{\pi}{2}\}$. Equation (A.18) implies also that the convolution

$$\int_{\mathcal{D}} f(R, \vec{r'}) f(R, \vec{r} + \vec{r'}) d\vec{r'} \qquad (A.19)$$

has to be zero outside the disk $\{\vec{r}\ :\ \ r < 2R\}$. Hence, the product $f(R,\vec{r'})f(R,\vec{r}+\vec{r'})$ must vanish identically for all $\vec{r}\ :\ \ r > 2R$, except maybe on a set of measure zero and on which it must remain finite. From the above discussion, we conclude that $\mathcal{D}$ must satisfy the following two constraints:

(i) $\mathcal{D} \subset \{\vec{r}:\ \ r < 2R \ \text{ and } \ \frac{-\pi}{2} < \theta < \frac{\pi}{2}\}$

(ii) $\mathcal{D} \bigcap \{\vec{s}:\ \ \vec{s}+\vec{r} \in \mathcal{D} \ \text{ and } \ r > 2R\} = \phi.$

A simple geometrical picture shows that the only subset of $\mathbf{R}^2$ that satisfies the above two constraints is a disk of radius R centered at the point $\vec{R_0} = (R,0)$, i.e.

$$\mathcal{D} = \{\vec{r}:\ \ |\vec{r} - \vec{R_0}| < R\}. \qquad (A.20)$$

Identity (A.20) can also be derived using the theory of [45, Chapter 3]. (See [30] for details.) Next denote by $\mathcal{C}$ the causal space of functions of $\vec{\lambda}$ which are the Fourier transforms of functions that are zero for $r_1 < 0$ in a Cartesian coordinate representation of the spatial domain (i.e. where $\vec{r} = (r_1, r_2)$), and denote by $\mathcal{A}$ the anticausal space of functions of $\vec{\lambda}$ which are the Fourier transforms of functions that are zero for $r_1 > 0$. Since $F^*(R,\vec{\lambda})$ is the Fourier transform of a function that is zero for $r_1 > 0$, then

$$D(\vec{\lambda}) = (1 - F^*(R,\vec{\lambda}))^{-1} - 1 \qquad (A.21)$$

must also correpond to the Fourier transform of a function that is zero for $r_1 > 0$. To see why this has to be true, factor $D(\vec{\lambda})$ as [47]

$$D(\vec{\lambda}) = D_c(\vec{\lambda}) + D_a(\vec{\lambda}), \qquad (A.22)$$

where $D_c(\vec{\lambda})$ and $D_a(\vec{\lambda})$ belong respectively to $\mathcal{C}$ and $\mathcal{A}$. Equations (A.21)-(A.22) imply that

$$F^*(R,\vec{\lambda}) = D_c(\vec{\lambda}) + D_a(\vec{\lambda}) - D_c(\vec{\lambda})F^*(R,\vec{\lambda}) - D_a(\vec{\lambda})F^*(R,\vec{\lambda}) \qquad (A.23)$$

and since $F^*(R,\vec{\lambda})$ is the Fourier transform of a function that is zero for $r_1 > 0$, we must have

$$D_c(\vec{\lambda}) = D_c(\vec{\lambda})F^*(R,\vec{\lambda}) \qquad (A.24)$$

or

$$D_c(\vec{\lambda}) = 0 \tag{A.25}$$

which proves our assertion. Combining (A.21) with (A.14) we obtain

$$\int_{-\infty}^{\infty} \hat{S}_y(\vec{\lambda})(1 - F(R, \vec{\lambda}))e^{j\lambda_1 r_1}\, d\lambda_1 = P\delta(r_1) \quad \text{for} \quad r_1 > 0. \tag{A.26}$$

Furthermore, if we take the inverse Fourier transform of (A.26) with respect to λ_2 we get

$$\int_{\mathcal{D}} \hat{K}_y(|\vec{r} - \vec{s}|)(\delta(\vec{s}) - f(R, \vec{s}))\, d\vec{s} = P\delta(\vec{r}) \quad \text{for} \quad r_1 > 0. \tag{A.27}$$

To compute $f(R, \vec{r})$ from the above integral equation, we note that (A.20) implies that for any $\vec{r} \in \mathcal{D}$ and any $\vec{s} \in \mathcal{D}$ we have $|\vec{r} - \vec{s}| < 2R$, so that

$$\hat{K}_y(|\vec{r} - \vec{s}|) = K_y(|\vec{r} - \vec{s}|) \quad \forall \vec{r} \in \mathcal{D} \quad \text{and} \quad \forall \vec{s} \in \mathcal{D} \tag{A.28}$$

by the correlation matching constraint (A.3). Since $K_y(\vec{r}) = P\delta(\vec{r}) + K_z(|\vec{r}|)$ is known by assumption for $r < 2R$ then $f(R, \vec{r})$ can be computed as the solution of the following integral equation

$$K_z(r) = \int_{\mathcal{D}} K_z(|\vec{r} - \vec{s}|) f(R, \vec{s})\, d\vec{s} + Pf(R, \vec{s}) \quad \forall \vec{r} \in \mathcal{D}. \tag{A.29}$$

Once $f(R, \vec{r})$ has been computed via (A.29), then (A.27) can be used to extend $K_z(r)$ beyond the disk $r < 2R$. Finally, if we make the change of variables

$$\vec{r'} = \vec{R}_0 - \vec{r} \tag{A.30}$$

$$\vec{s'} = \vec{R}_0 - \vec{s} \tag{A.31}$$

we obtain from (A.29)

$$K_z(|\vec{R}_0 - \vec{r'}|) = \int_{s' < R} K_z(|\vec{r'} - \vec{s'}|) g(R, \vec{s'})\, d\vec{s'} + Pg(R, \vec{r'}) \quad \forall r' < R, \tag{A.32}$$

where we defined

$$g(R, \vec{s'}) = f(R, \vec{R}_0 - \vec{s'}). \tag{A.33}$$

Note that (A.32) is just (3.6) with $\vec{r}$ and $\vec{u}$ replaced by $\vec{r'}$ and $\vec{s'}$ respectively.Furthermore, observe that (A.33) implies that

$$G(R,\vec{\lambda}) = F^*(R,\vec{\lambda})e^{-j\vec{\lambda}\cdot\vec{R}_0} \tag{A.34}$$

where $G(R,\vec{\lambda})$ is the 2-D Fourier transform of $g(R,\vec{r})$. Hence,

$$\begin{aligned} |1-F(R,\vec{\lambda})|^2 &= |1-F^*(R,\vec{\lambda})|^2 \\ &= |1-G(R,\vec{\lambda})e^{j\vec{\lambda}\cdot\vec{R}_0}|^2 \\ &= |e^{-j\vec{\lambda}\cdot\vec{R}_0}-G(R,\vec{\lambda})|^2. \end{aligned} \tag{A.35}$$

Combining (A.14) with (A.35) we obtain

$$\hat{S}_y(\lambda) = \frac{P}{|e^{-j\vec{\lambda}\cdot\vec{R}_0}-G(R,\vec{\lambda})|^2} \tag{A.36}$$

which is equation (3.5).

APPENDIX B

Estimation of the covariance functions

The algorithm that we presented in Sections 3 and 4 for computing $\hat{S}_y(\lambda)$ is based on the knowledge of $k_n(r_i, r_j)$, the covariance function of the nth order-Fourier coefficient process corresponding to the measurements $y(\vec{r})$. However in practice, one is given measurements of the field itself rather than $k_n(r_i, r_j)$. In this appendix, we summarize a procedure developed in [33] to compute an *unbiased* and *consistent* estimate of the *non-stationary* covariance function $k_n(r_i, r_j)$ directly from the measurements. Let us start by assuming that measurements of the field $y(\vec{r})$ are available at all the points inside the disk $D_{R^*} = \{\vec{r} : \ 0 \leq r \leq R^*\}$. Then to estimate $k_n(r_i, r_j)$, we can use a two step procedure. In the first step we estimate $K(r)$ using the given data. In the second step we substitute our estimate of $K(r)$ into (2.6) to obtain $k_n(r_i, r_j)$.

$K(r)$ can be estimated by using a simple extension of the 1-D techniques that were developed to estimate the covariance function of ergodic stationary processes. Observe that along any line $\phi = \phi_0$ in a tomographic coordinate system[1], $y(\vec{r})$ is stationary. Hence, given the measurements $\{y(t, \phi_0) : \ -R^* \leq t \leq R^*\}$ along this line we can estimate $K(r)$, using a simple extension of the 1-D techniques, as

$$\hat{K}(r : \phi_0) = \frac{1}{R^{*2}} \int_{-R^*}^{R^*} y(t, \phi_0) y(r + t, \phi_0) \, |t| \, dt. \qquad (B.1)$$

Since measurements of $y(\vec{r})$ are assumed to be available all over the disk D_{R^*}, we can compute $\hat{K}(r : \phi_0)$ for all ϕ_0, $0 \leq \phi_0 \leq \pi$, and take $\hat{K}(r)$ to be the average of the $\hat{K}(r : \phi_0)$ over all ϕ_0. In other words, we can estimate $K(r)$ as

$$\hat{K}(r) = \frac{1}{\pi R^{*2}} \int_0^{R^*} ds \int_0^{2\pi} d\theta \, s \, y(s, \theta) y(r + s, \theta). \qquad (B.2)$$

Note that we have used the weight function $w(t) = |t|$ in (B.1) to guarantee that $\hat{K}(r)$ corresponds to a spatial average.

[1] A tomographic coordinate system (t, ϕ) is a modified polar coordinate system where t takes both positive and negative real values, and where ϕ varies from 0 to π.

Next, we can use $\hat{K}(r)$ to obtain an estimate of $k_n(r_i, r_j)$ by simply substituting $\hat{K}(r)$ for $K(r)$ into (2.6). Thus, we take as our estimate of $k_n(r_i, r_j)$ the quatity

$$\hat{k}_n(r_i, r_j) = \frac{1}{2\pi} \int_0^{2\pi} d\theta \, \hat{K}((r_i^2 + r_j^2 - 2r_i r_j \cos\theta)^{1/2}) \, e^{-jn\theta}. \qquad (B.3)$$

Note that according to (B.3) one needs to estimate $K(r)$ for $0 \leq r \leq 2r^*$ in order to be able to estimate $k_n(r_i, r_j)$ for $0 \leq r_i, r_j \leq r^*$.

It is shown in [33] that $\hat{K}(r)$ is an unbiased estimate of $K(r)$. Furthermore, since $k_n(r_i, r_j)$ and $\hat{k}_n(r_i, r_j)$ are related linearly to $K(r)$ and $\hat{K}(r)$ respectively, then it follows immediately from the unbiasedness and consistency properties of $\hat{K}(r)$ that $\hat{k}_n(r_i, r_j)$ is an unbiased and consistent estimate of $k_n(r_i, r_j)$. Thus, by using (B.2) and (B.3) we are able to obtain an unbiased and consistent estimate of the non-stationary covariance function $k_n(r_i, r_j)$.

REFERENCES

[1] J. Makhoul, "Spectral Analysis of Speech Estimation," *IEEE Trans. Audio Electroacoust*, vol. AU-21, pp. 140-148, June 1973.

[2] J. H. McClellan and R. J. Purdy, "Applications of Digital Signal Processing to Radar," in *Applications of Digital Signal Processing*, A. V. Oppenheim, Ed., Prentice-Hall, Englewood Cliffs, N.J., 1978.

[3] A. B. Baggeroer, "Sonar Signal Processing," in *Applications of Digital Signal Processing*, A. V. Oppenheim, Ed., Prentice-Hall, Englewood Cliffs, N.J. 1978.

[4] S. J. Wernecke and L. R. D'Addario, "Maximum Entropy Image Reconstruction," *IEEE Trans. Comput.*, vol. C-26, pp. 351-364, April 1977.

[5] E. A. Robinson and S. Treitel, "Maximum Entropy and the Relationship of the Partial Autocorrelation to the Reflection Coefficients of a Layered System," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-28, no. 2, pp. 224-235, April 1980.

[6] S. M. Kay and S. L. Marple, "Spectrum Analysis - A Modern Perspective," *Proc. IEEE*, vol. 69, no. 11, pp. 1380-1419, Nov. 1981.

[7] S. E. Roucos and D. G. Childers, "A Two Dimensional Maximum Entropy Spectral Estimator," *IEEE Trans. Inform. Theory*, vol. IT-26, pp. 554-560, Sept. 1980.

[8] J. S. Lim and N. A. Malek, "A New Algorithm for Two-Dimensional Maximum Entropy Power Spectrum Estimation," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-29, pp. 401-413, June 1981.

[9] A. K. Jain and S. Raganath, "Two-Dimensional Spectral Estimation," in *Proc. RADC Spectrum Estimation Workshop*, Rome, N. Y. May 1978, pp. 217-225.

[10] S. W. Lang and J. H. McClellan, "Multidimensional MEM Spectral Estimation," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-30, no. 6, pp. 880-887, Dec. 1982.

[11] M. P. Ekstrom and J. W. Woods, "2-D Spectral Factorization with Applications in Recursive Digital Filtering," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-24, no. 2, pp. 115-178, Apr. 1976.

[12] J. J. Murray, "Spectral Factorization and Quarter-plane Digital Filters," *IEEE Trans. Circuits and Systems*, vol. CS-25, no. 8, pp. 586-592, Aug. 1978.

[13] J. H. McClellan, "Multidimensional Spectral Estimation," *Proc. IEEE*, vol. 70, pp. 1029-1039, Sept. 1982.

[14] W. Rudin, "The Extension Problem for Positive-Definite Functions," *Ill. J. Math.*, vol. 7, pp. 532-539, 1963.

[15] J. Capon, "Maximum-Likelihood Spectral Estimation," in *Nonlinear Methods of Spectral Analysis*, S. Haykin, Ed., Springer-Verlag, New York, N.Y. 1979.

[16] A. B. Baggeroer, *Space/Time Random Processes and Optimum Array Processing*, Naval Undersea report NUC TP 506, San Diego, CA 1976.

[17] N. S. Burdic, *Underwater Acoustic System Analysis*, Prentice Hall Inc., Englewood Cliffs, N. J., 1984.

[18] P. R. Julian and A. Cline, "The Direct Estimation of Spatial Wavenumber Spectra of Atmospheric Variables," *Jour. Atmospheric Sci.*, vol. 31, pp. 1526-1539, 1976.

[19] A. S. Monin and A. M. Yaglom, *Statistical Fluid Mechanics: Mechanics of Turbulence*, Volume 2, MIT Press, Cambridge, MA., 1975.

[20] I. Rodriguez-Iturbe and J. M. Mejia, "The Design of Rainfall Networks in Space and Time," *Water Resour. Res.*, vol. 10, no. 4, Aug. 1974.

[21] W. Rudin, "An Extension Theorem for Positive-Definite Functions," *Duke Math. J.*, vol. 37, pp. 49-53, 1970.

[22] B. C. Levy and J. N. Tsitsiklis, "A fast Algorithm for Linear Estimation of Two-Dimensional Isotropic Random Fields," *IEEE Trans. Inform. Theory*, vol. IT-31, no. 5, pp 635-644, Sept. 1985.

[23] G. Sharma and R. Chellappa, "A Model Based Approach for Estimation of Two-Dimensional Maximum Entropy Power Spectra," *IEEE Trans. Inform. Theory*, vol. IT-31, no. 1, pp 90-99, Jan. 1985.

[24] A. Papoulis, *Systems and Transforms with Applications in Optics*, MacGraw-Hill, New York, N. Y., 1968.

[25] H. Bateman, *Higher Transcendal Functions*, Volume II, McGraw-Hill, New-York, N.Y. 1953.

[26] H. L. Van Trees, *Detection, Estimation and Modulation Theory*, vol. 1, John Wiley, New York, NY, 1968.

[27] M. I. Yadrenko, *Spectral Theory of Random Fields*, Optimization Software Inc., New York, N.Y. 1986.

[28] D. R. Smith, *Variational Methods in Optimization*, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1974.

[29] T. L. Marzetta, "Two-Dimensional Linear Prediction: Autocorrelation Arrays, Minimum-Phase Prediction Error Filters, and Reflection Coefficient Arrays," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. ASSP-28, pp. 726-733, Dec. 1980.

[30] A. H. Tewfik, *Recursive Estimation and Spectral Estimation for 2-D Isotropic Random Fields*, Sc.D. thesis, in preparation.

[31] T. Kailath, *Lectures on Wiener and Kalman Filtering*, Springer-Verlag, New York, N. Y., 1981.

[32] H. Dym and I. Gohberg, "On an Extension Problem, Generalized Fourier Analysis, and an Entropy Formula," *Integral Equations and Operator Theory*, vol. 3, pp. 143-215, 1980.

[33] A. H. Tewfik, B. C. Levy and A. S. Willsky, "An Eigenstructure Approach for the Retrieval of Cylindrical Harmonics from 2-D Isotropic Covariance Data," *Report LIDS-P-1477 (revised)*, LIDS, MIT, Cambridge, MA, Oct. 1986; also submitted to *Signal Processing.*

[34] P. J. Davis and P. Rabinowitz, *Methods of Numerical Integration*, Academic Press, Orlando, FL. 1984.

[35] L. Lapidus and G. F. Pinder, *Numerical Solution of Partial Differential Equations in Science and Engineering*, John Wiley, New York, N. Y., 1982.

[36] W. F. Ames, *Numerical Methods for Partial Differential Equations*, Academic Press, New York, N. Y., 1977.

[37] S. M. Candel, "Simultaneous Calculation of Fourier-Bessel Transforms up to Order N," *J. Comput. Phys.*, vol. 23, pp. 343-353, 1981.

[38] D. R. Mook, "An Algorithm for the Numerical Evaluation of the Hankel and Abel Transforms," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. ASSP-31, pp. 979-985, August 1983.

[39] E. W. Hansen, "Fast Hankel Transform Algorithm," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. ASSP-33, no. 3, pp. 666-671, June 1985.

[40] V. F. Pisarenko, "The Retrieval of Harmonics from a Covariance Function," *Geophys. J. Roy. Astron. Soc.*, Vol. 33, pp. 347-366, 1973.

[41] R. O. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," in *Proc. RADC Spectrum Estimation Workshop*, Griffths AFB, N.Y., pp. 243-258, Oct. 1979.

[42] G. Bienvenu and L. Kopp, "Adaptivity to Background Noise Spatial Coherence for High Resolution Passive Methods," in *Proc. ICASSP 80*,

[43] M. Shinozuka and C.-M. Jan, "Digital Simulation of Random Processes and its Applications," *J. Sound Vib.*, vol. 25, no. 1, pp. 111-128, 1972.

[44] I. C. Gohberg and M. G. Krein, *Theory and Applications of Volterra Operators in Hilbert Space*, Trans. Math. Monographs, vol. 24, Amer. Math. Soc., Providence, R.I., 1970.

[45] L. I. Ronkin, *Introduction to the Theory of Entire Functions of Several Variables*, Trans. Math. Monographs, vol. 44, Amer. Math. Soc., Providence, R.I., 1974.

[46] M. Plancherel and G. Polya, "Fonctions Entieres et Integrales de Fourier Multiples," *Comment. Math. Helv.*, vol. 9, pp 224-248, 1937.

[47] G. F. Carrier, M. Krook and C. E. Pearson, *Functions of a Complex Variable; Theory and Technique*, McGraw-Hill, New York, NY, 1966.

FIGURE CAPTIONS

Fig. 1 Discretization scheme and numerical implementation of the recursions for $g_n(R,r)$.

Fig. 2 True observation power spectrum for Example 6.1.

Fig. 3 Plot of the estimated power spectrum in Example 6.1 when exact covariance data is given over a disk of radius 20 m and with $N = 1$.

Fig. 4 Plot of the estimated power spectrum in Example 6.1 when exact covariance data is given over a disk of radius 20 m and with $N = 5$.

Fig. 5 Plot of the estimated power spectrum in Example 6.1 when exact covariance data is given over a disk of radius 20 m and with $N = 10$.

Fig. 6 Plot of the estimated power spectrum in Example 6.1 when exact covariance data is given over a disk of radius 40 m and with $N = 20$.

Fig. 7 Plot of the estimated power spectrum in Example 6.1 when estimates of the covariance function are computed over a disk of radius 20 m given the data over a disk of radius 100 m and with $N = 5$.

Fig. 8 Plot of the estimated power spectrum in Example 6.1 when estimates of the covariance function are computed over a disk of radius 20 m given the data over a disk of radius 30 m and with $N = 5$.

Fig. 9 Plot of the estimated power spectrum in Example 6.1 when estimates of the covariance function are computed over a disk of radius 20 m given the data over a disk of radius 30 m and when $P = 0.0001$ Watt.m^2 and $N = 5$.

Fig. 10 True observation power spectrum for Example 6.2.

Fig. 11 Plot of the estimated power spectrum in Example 6.2 when exact covariance data is given over a disk of radius 20 m and with $N = 1$.

Fig. 12 Plot of the estimated power spectrum in Example 6.2 when exact covariance data is given over a disk of radius 20 m and with $N = 3$.

Fig. 13 Plot of the estimated power spectrum in Example 6.2 when exact covariance data is given over a disk of radius 20 m and with $N = 10$.

Fig. 14 True observation power spectrum for Example 6.3.

Fig. 15 Plot of the estimated power spectrum in Example 6.3 when exact covariance data is given over a disk of radius 20 m and with $A = 1$ and $N = 10$.

Fig. 16 Plot of the estimated power spectrum in Example 6.3 when exact covariance data is given over a disk of radius 20 m and with $A = 10$ and $N = 10$.

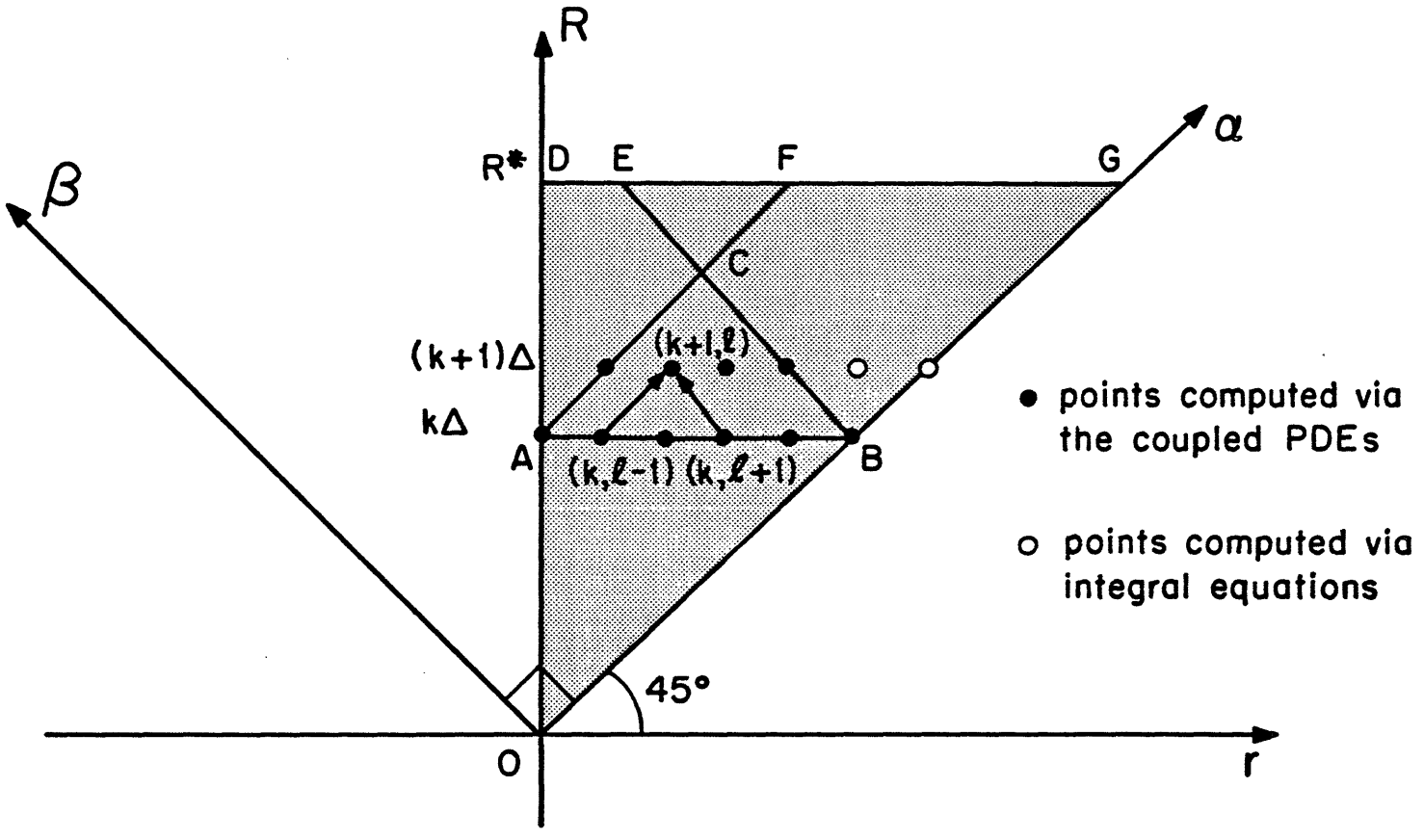

R
α
β
R*
D
E
F
G
C
(k+1)Δ
kΔ
(k+1,ℓ)
A
B
(k,ℓ-1)
(k,ℓ+1)
45°
O
r
• points computed via the coupled PDEs
o points computed via integral equations

Figure 1

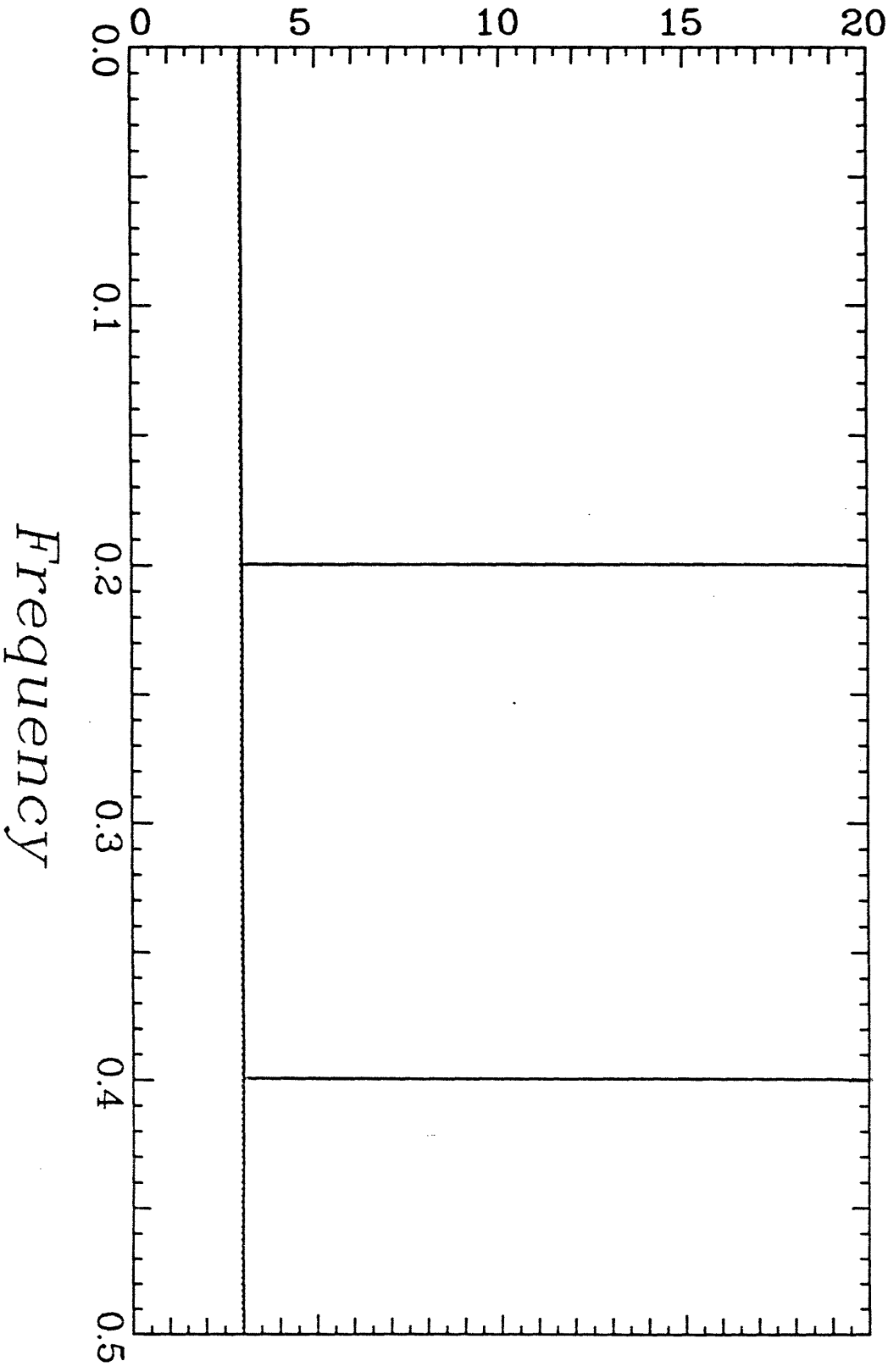


Figure 2

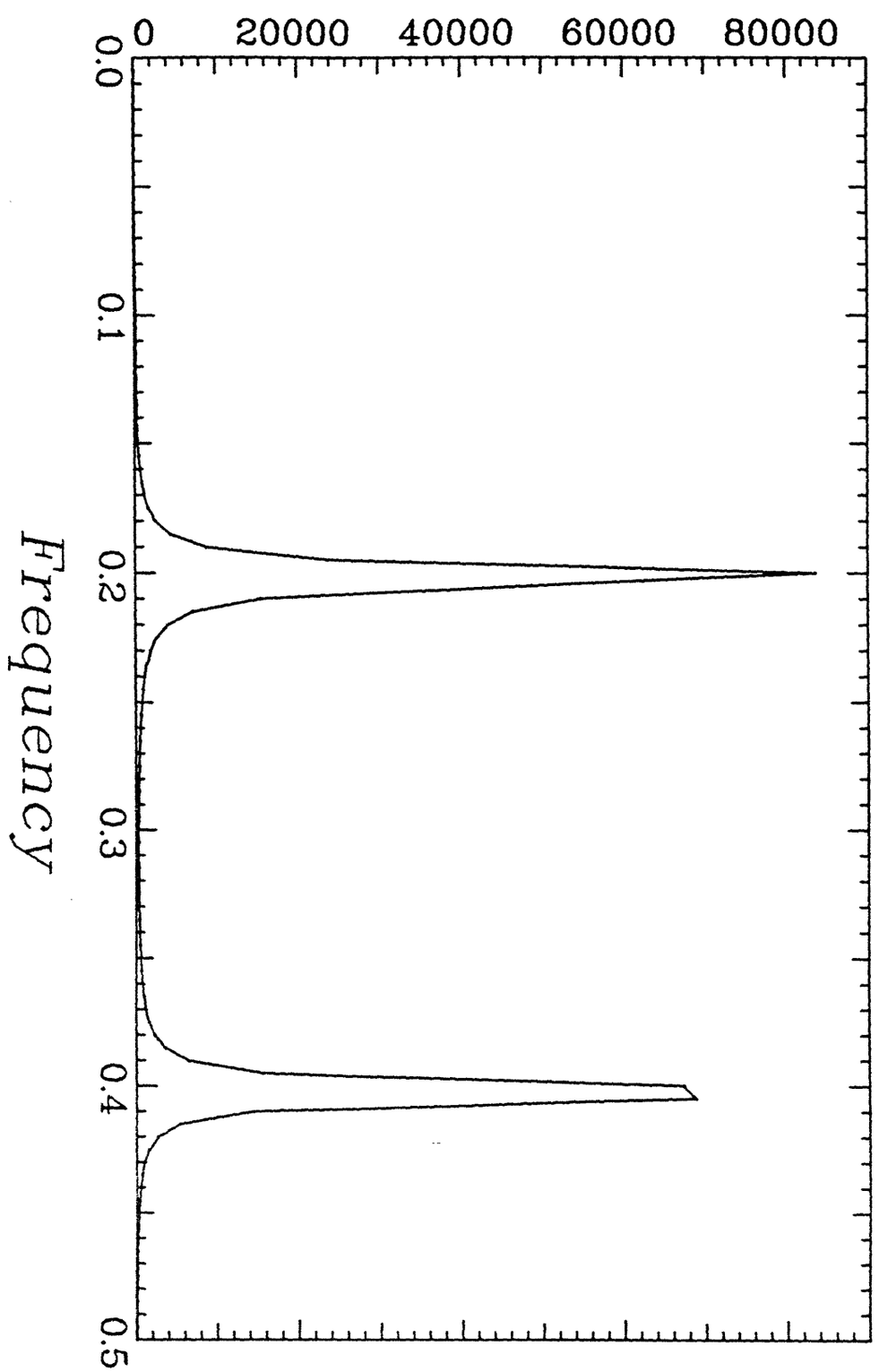
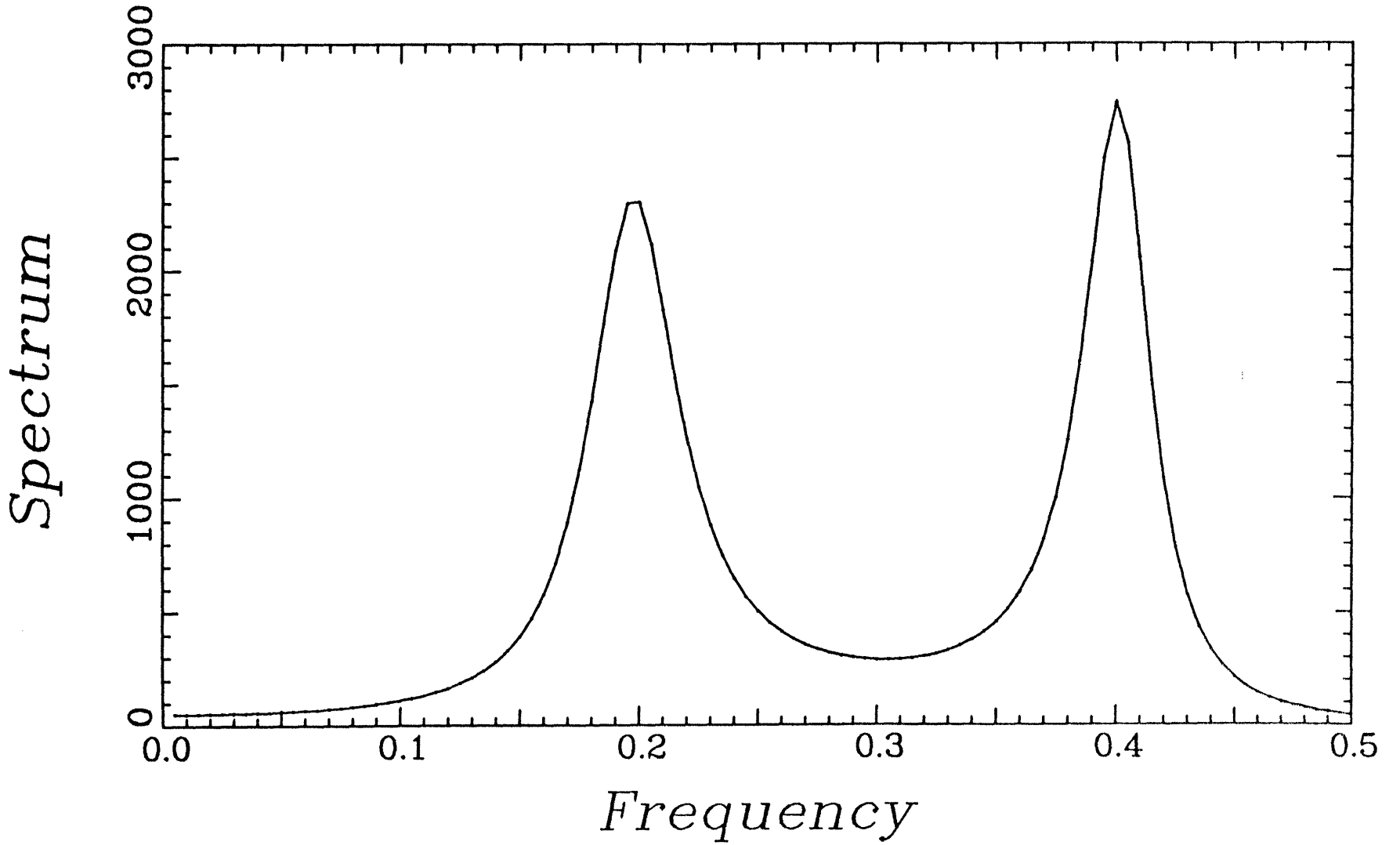

Spectrum
0
20000
40000
60000
80000
0.0
0.1
0.2
0.3
0.4
0.5
Frequency

Figure 3

Figure 4



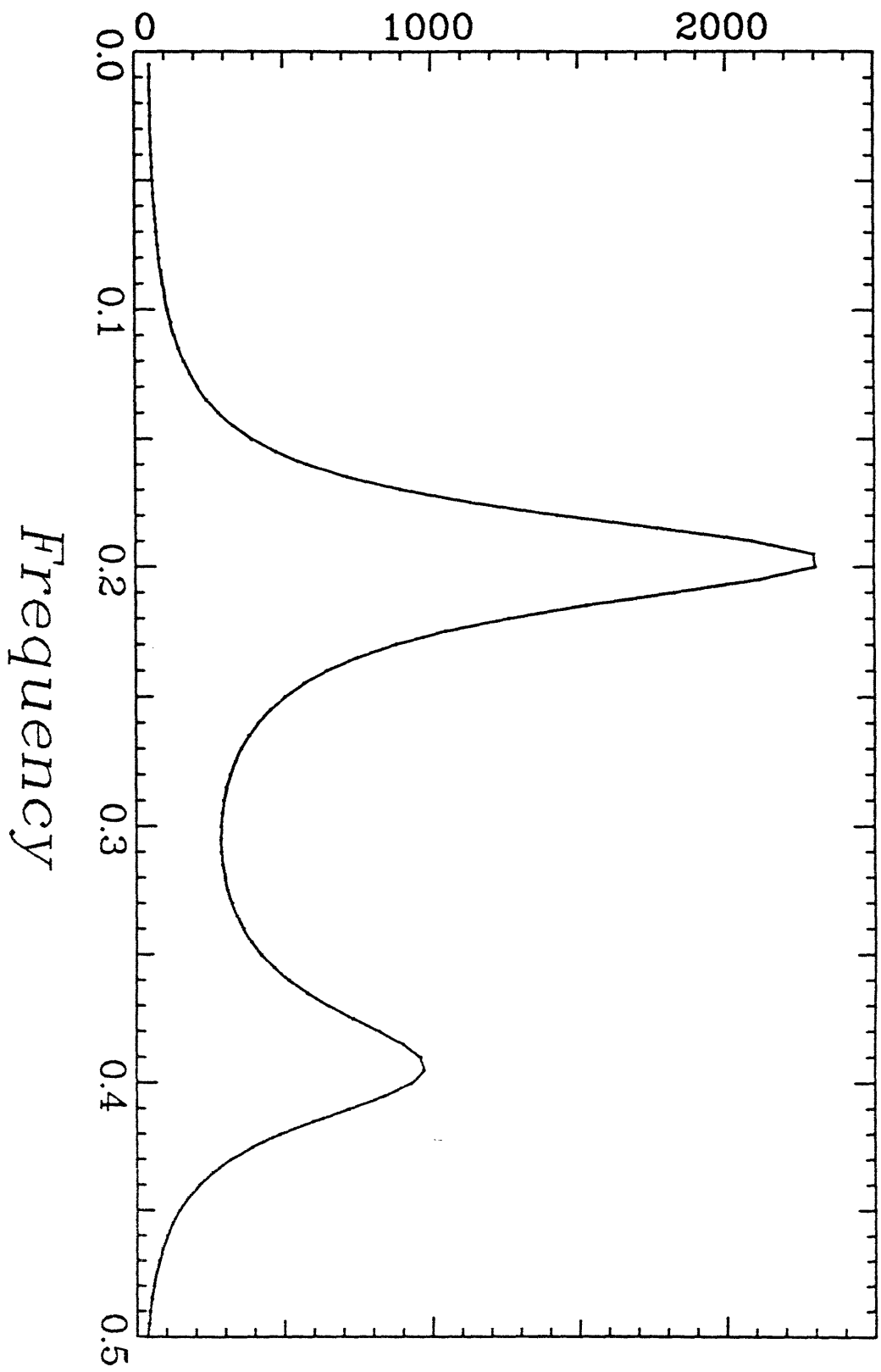
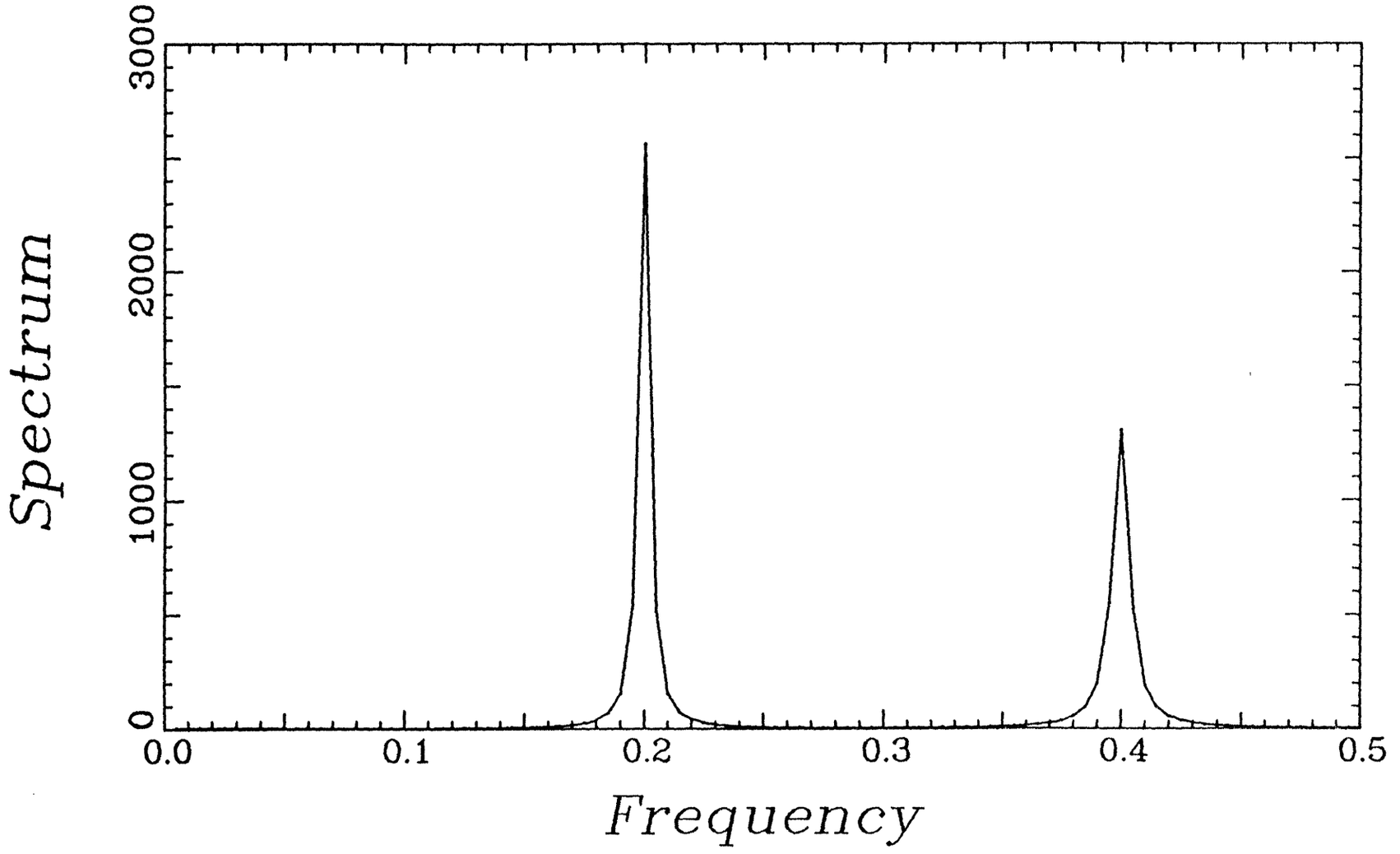


Figure 5

Figure 6



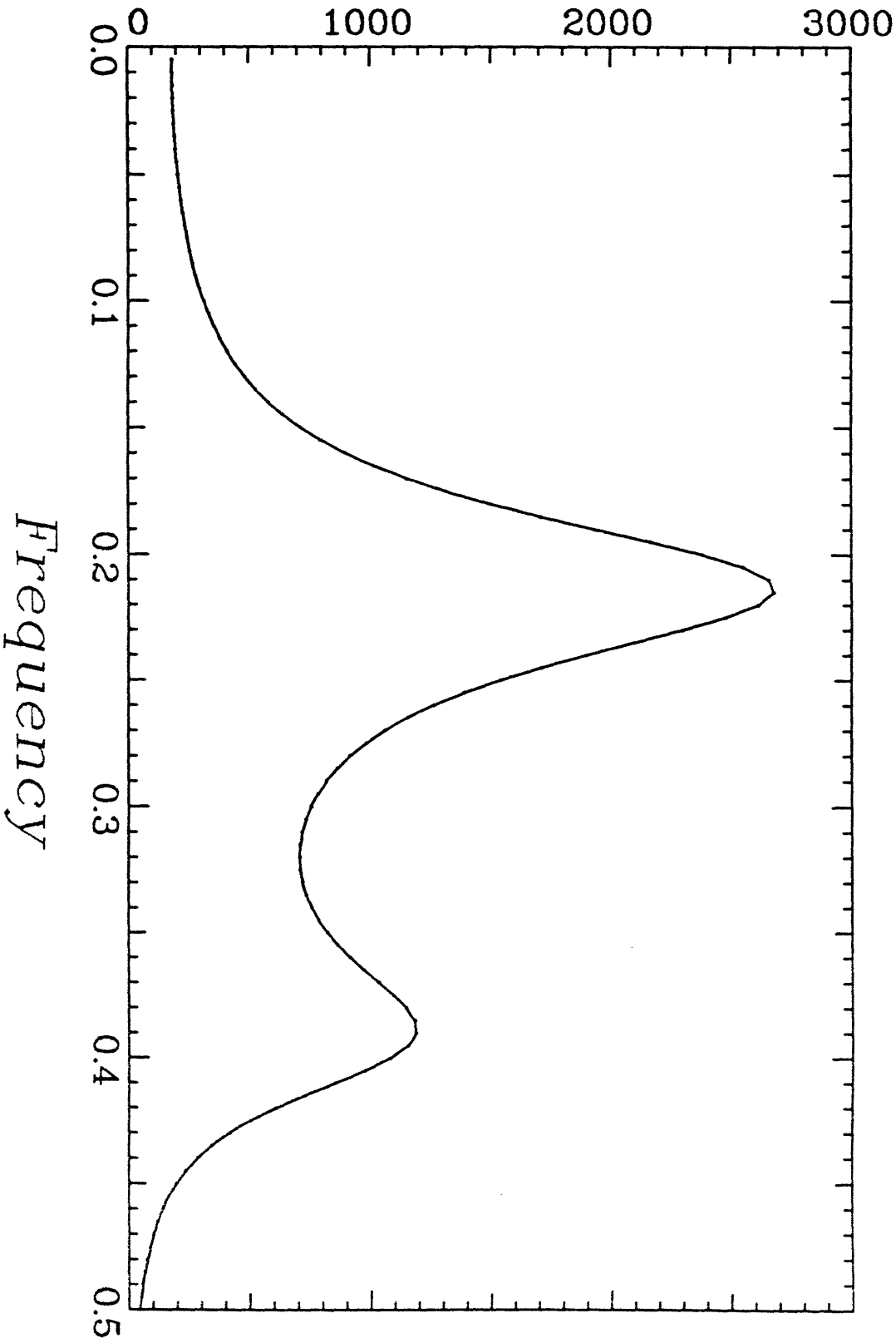


Figure 7

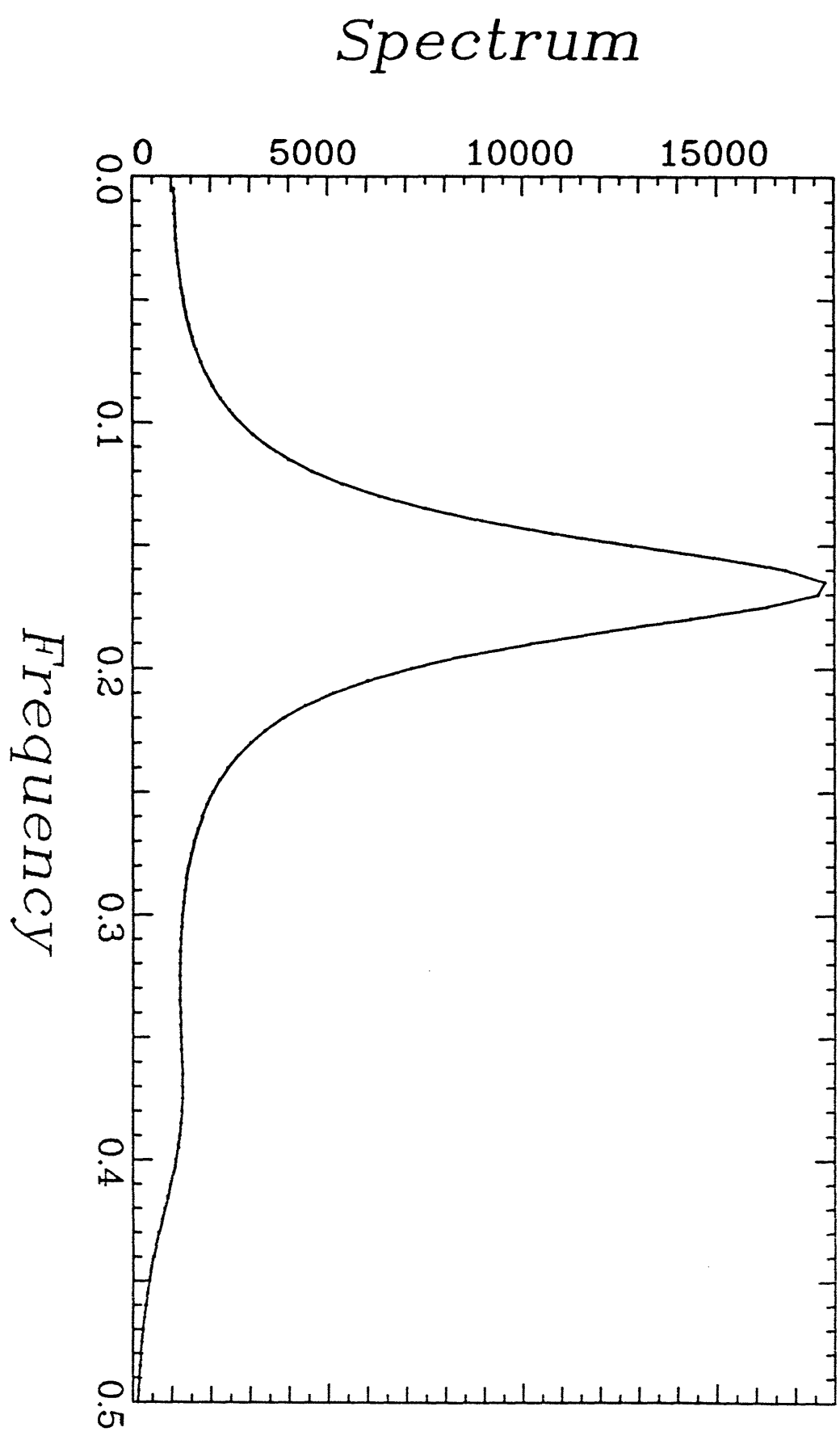
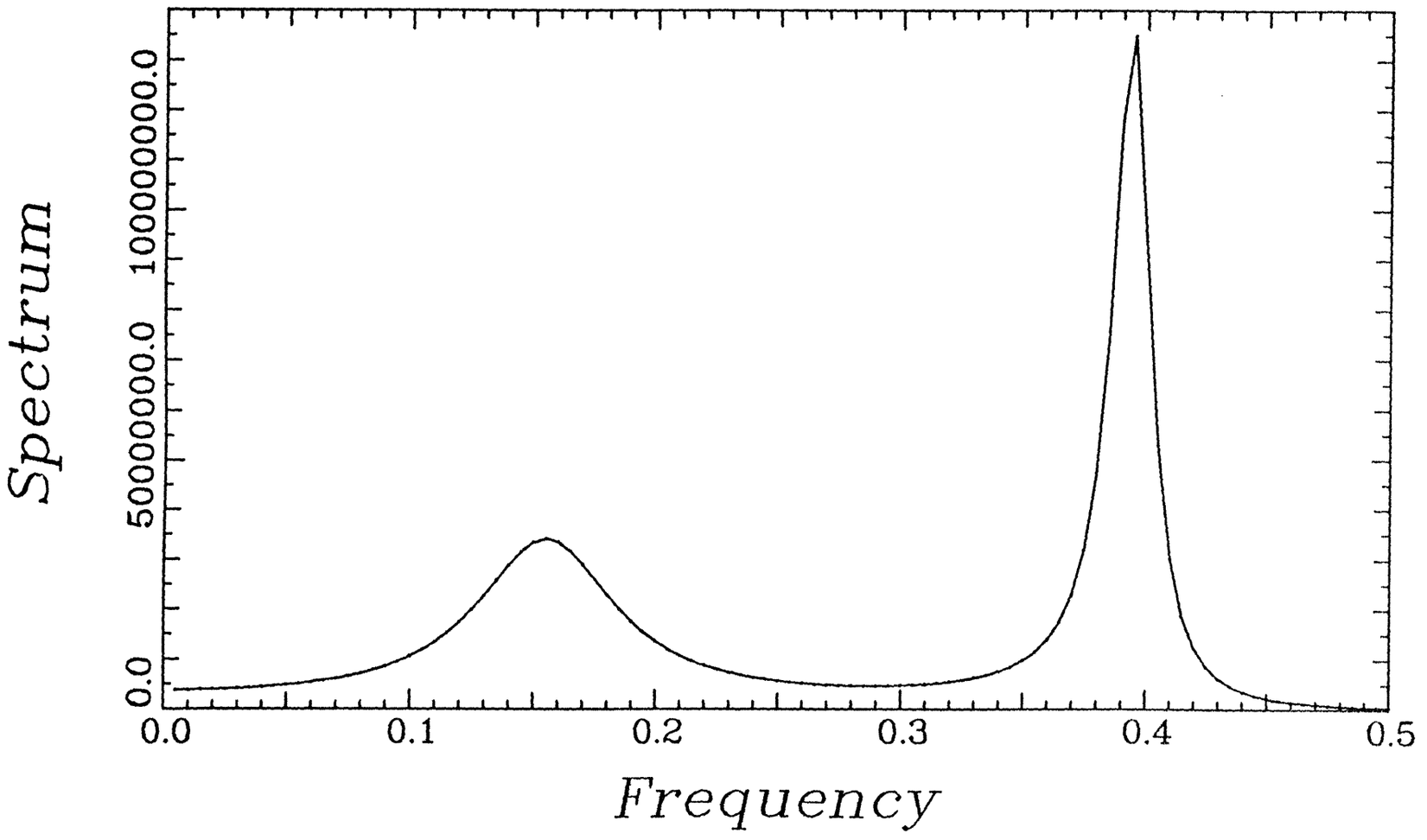


Figure 8

Figure 9


Spectrum
0.0
50000000.0
100000000.0
0.0
0.1
0.2
0.3
0.4
0.5
Frequency

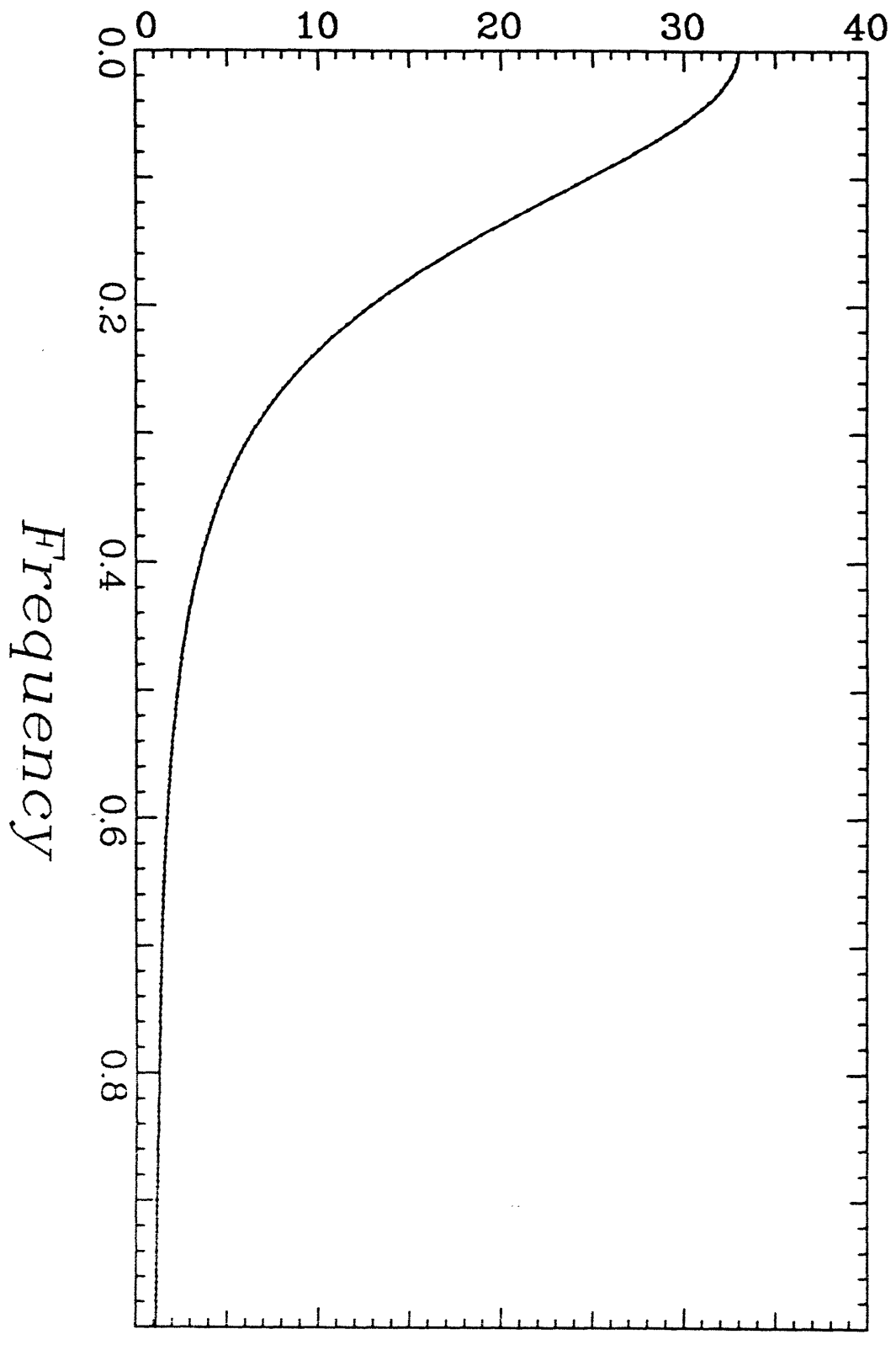


Figure 10

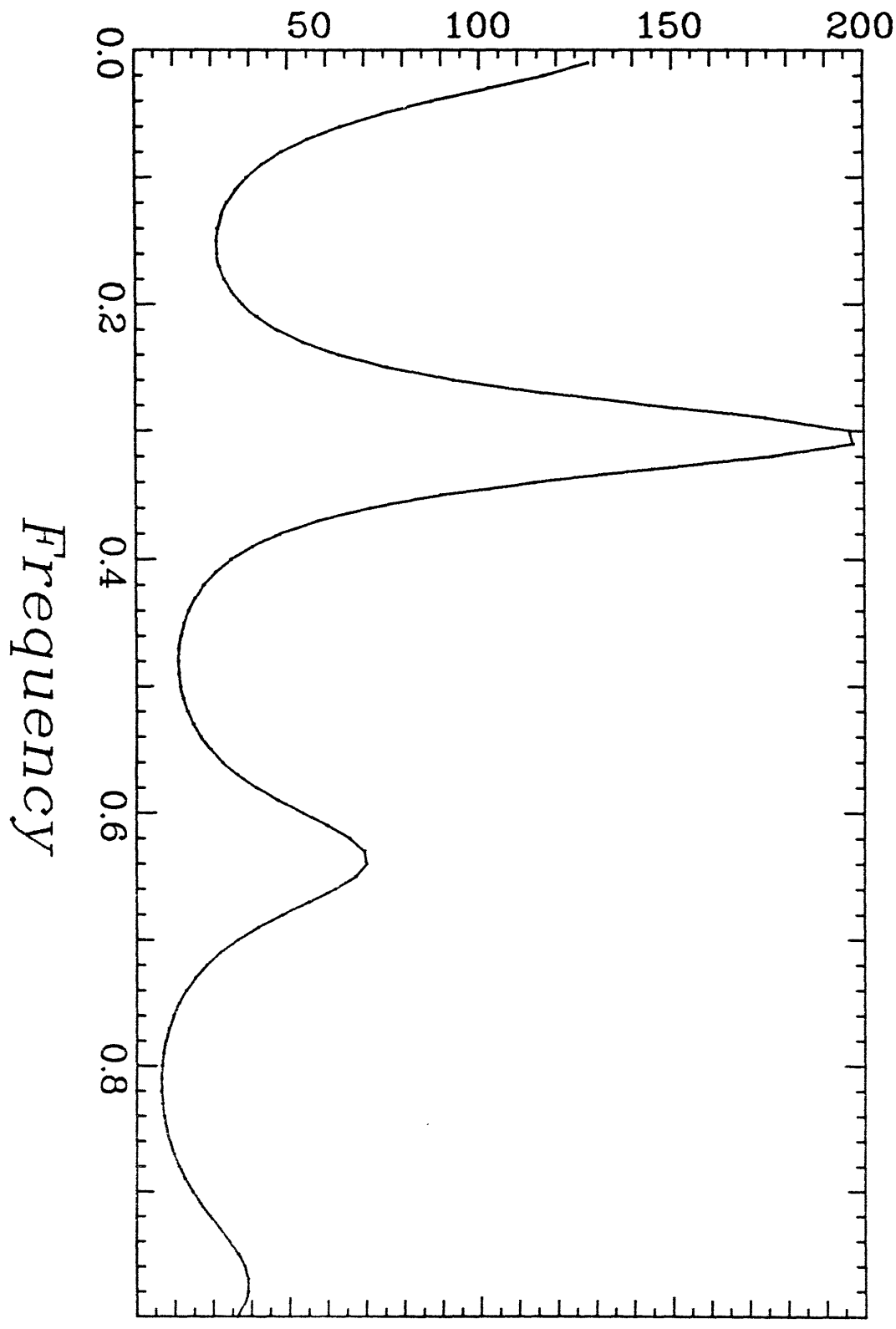


Figure 11

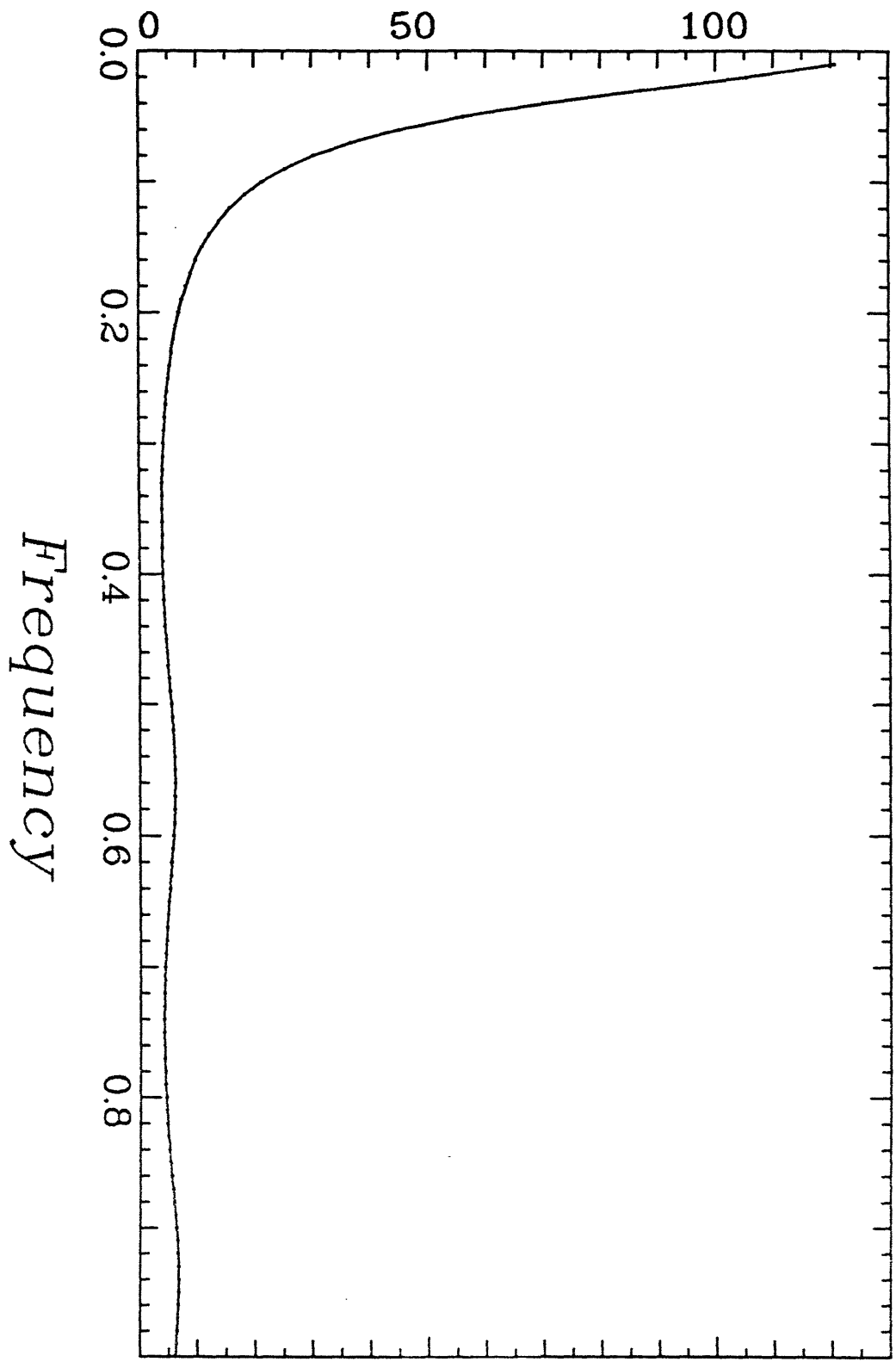


Figure 12

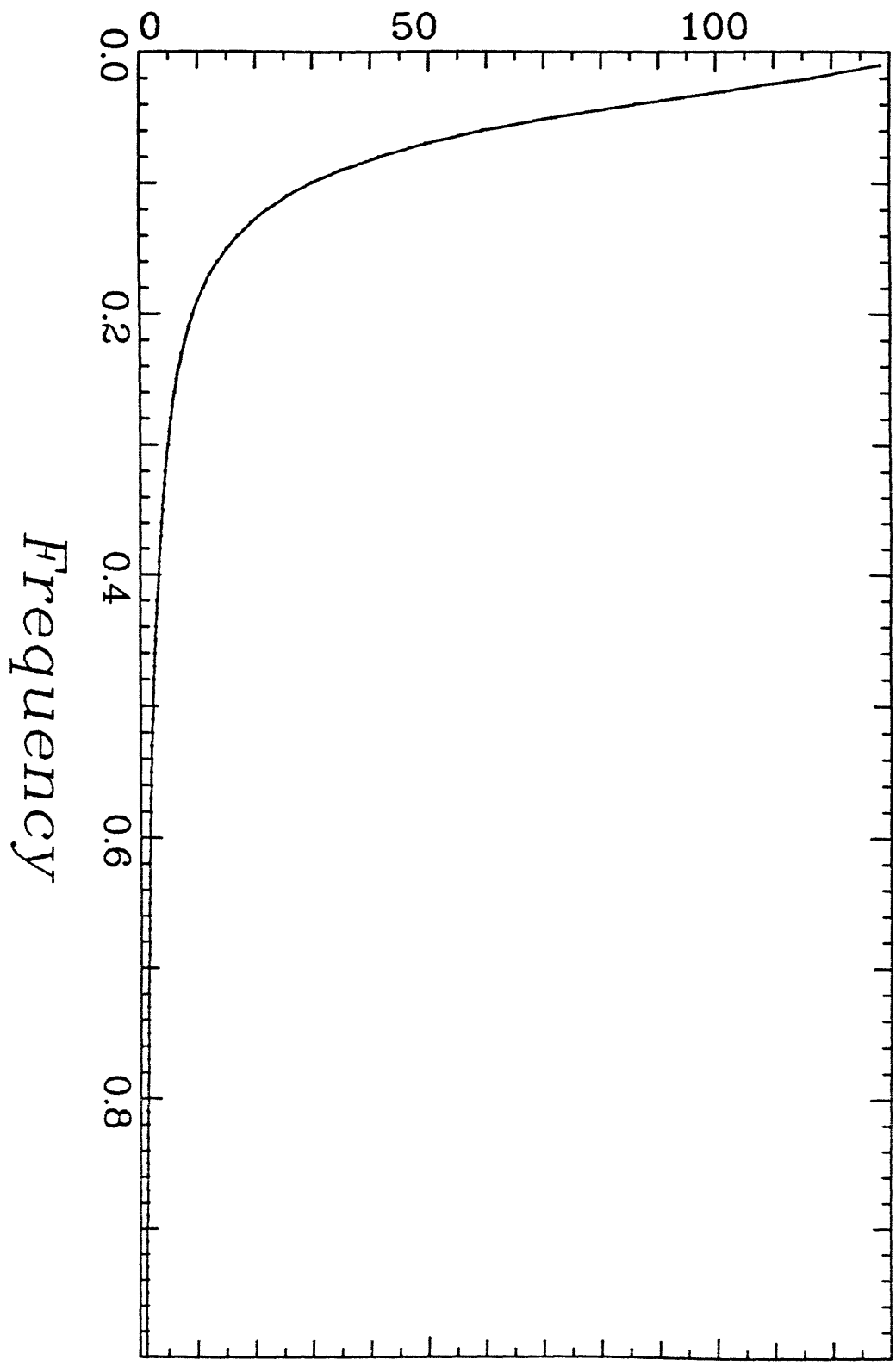


Figure 13

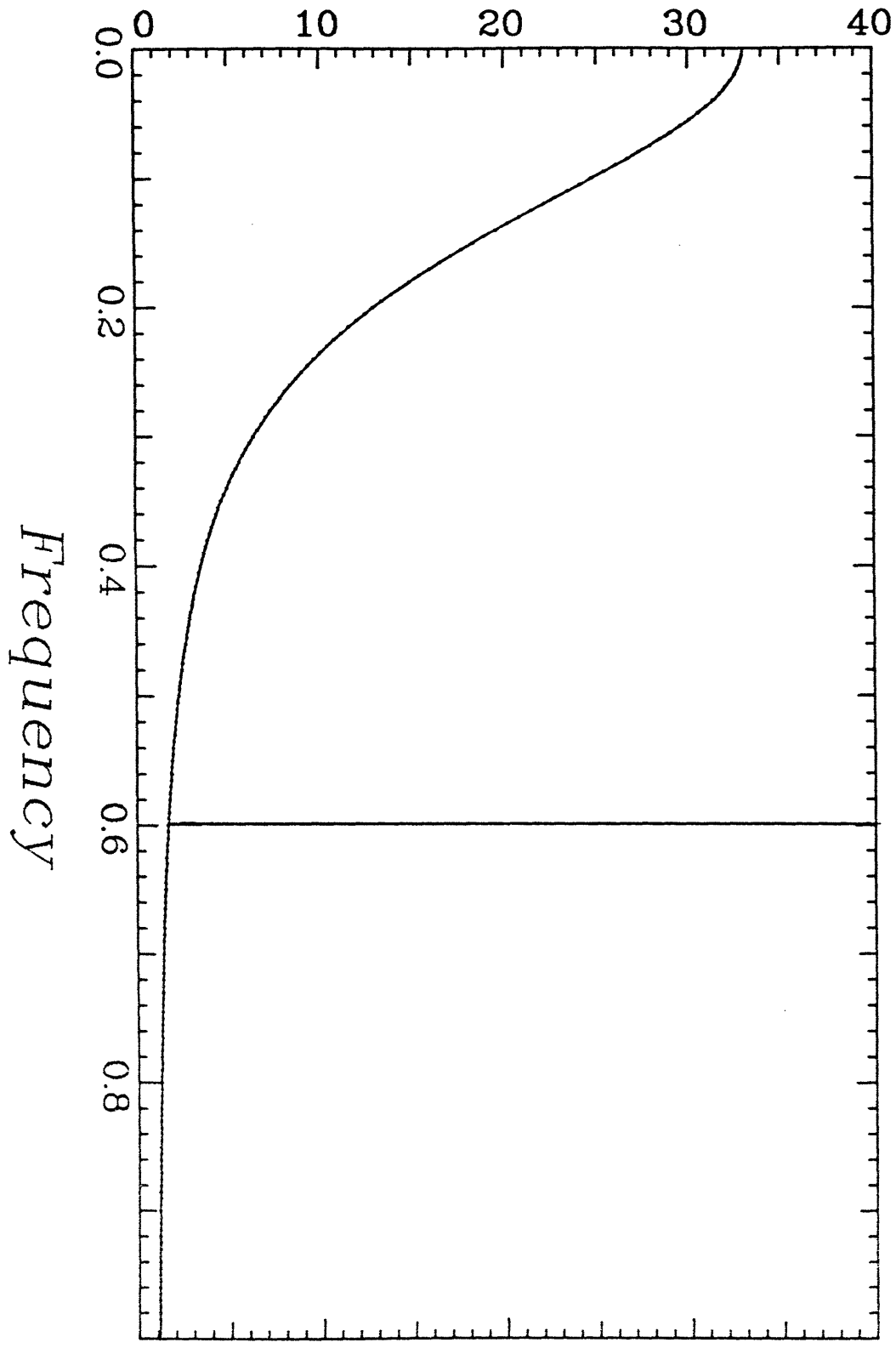


Figure 14

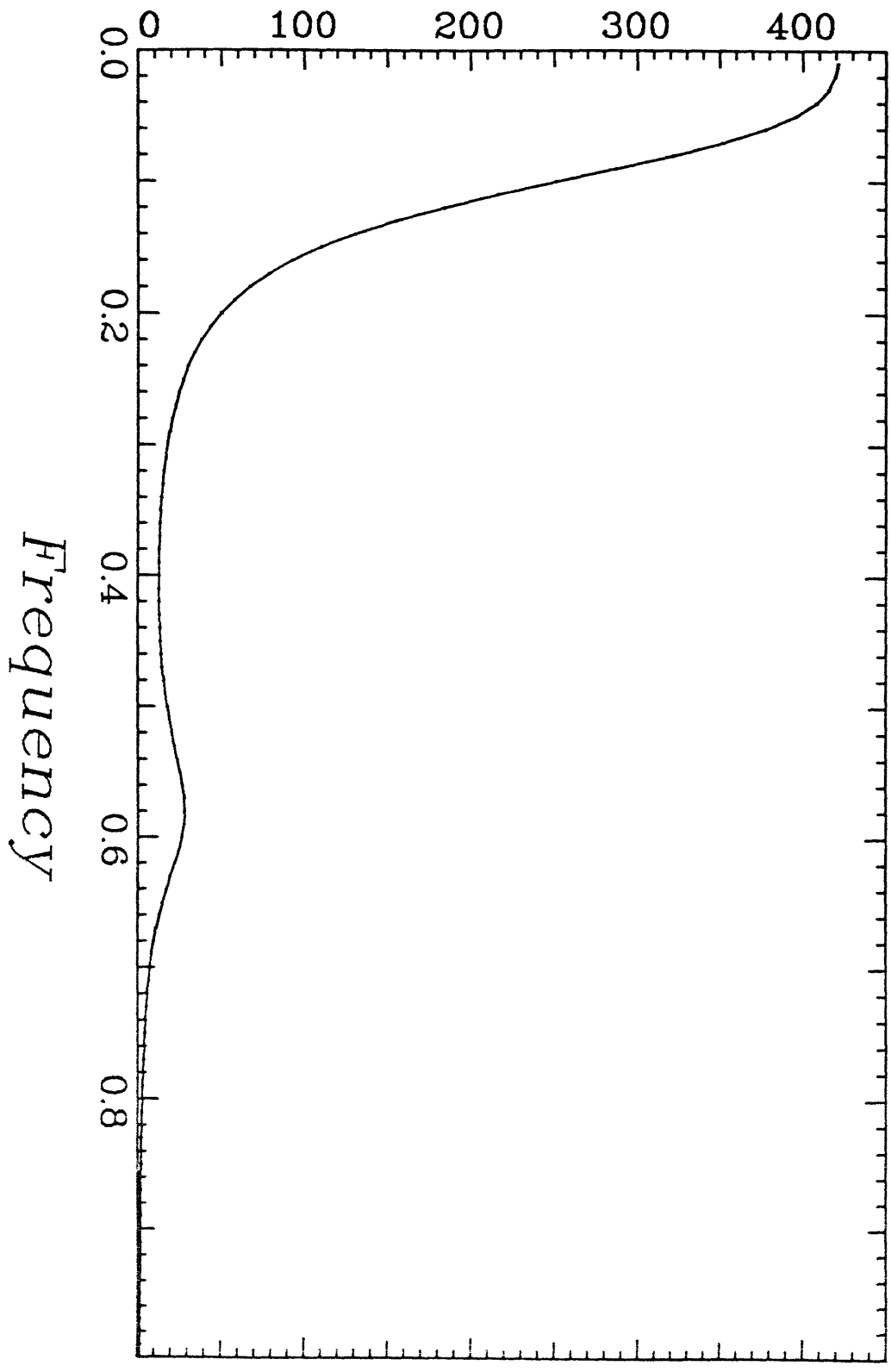


Figure 15

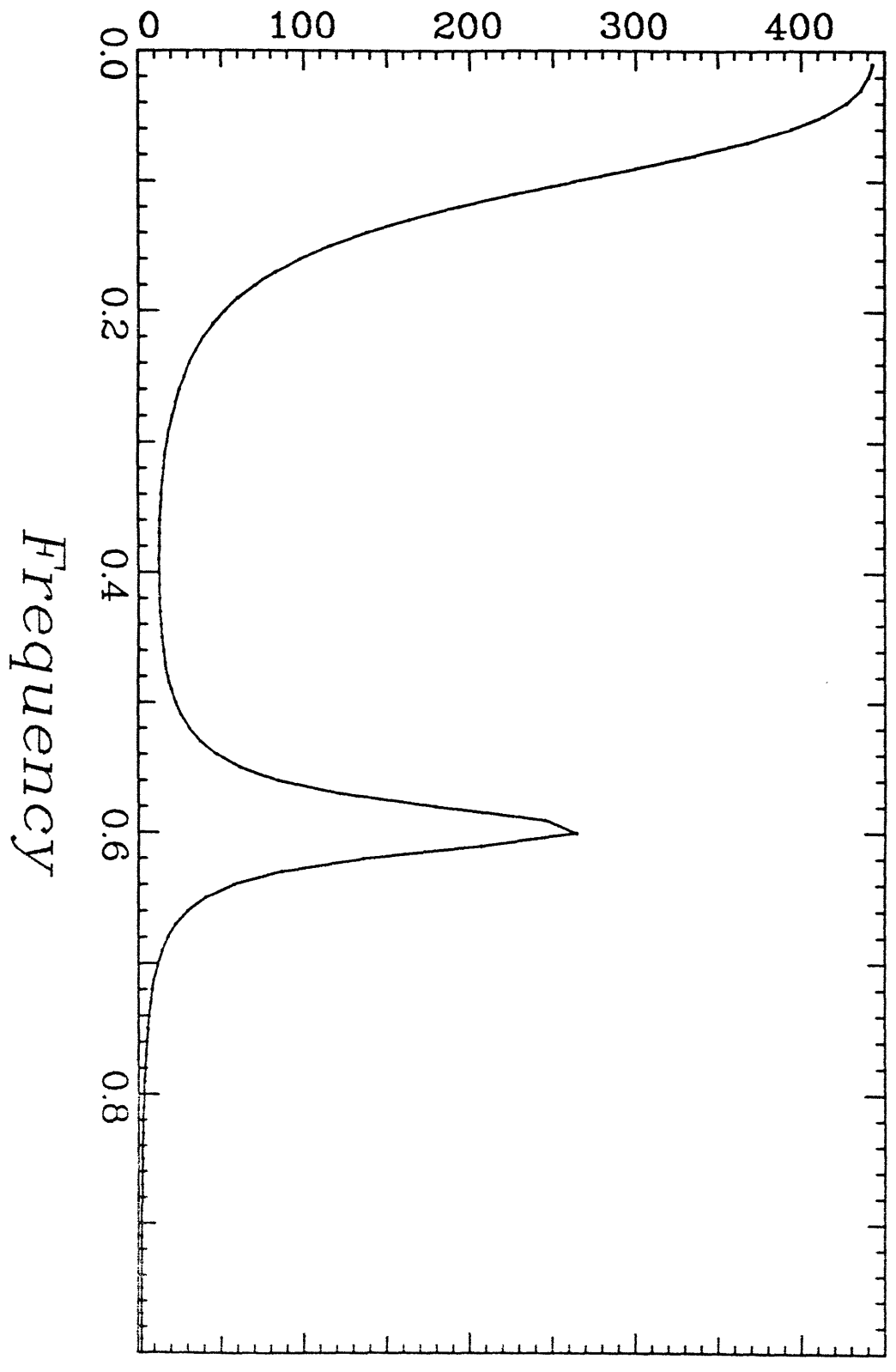


Figure 16