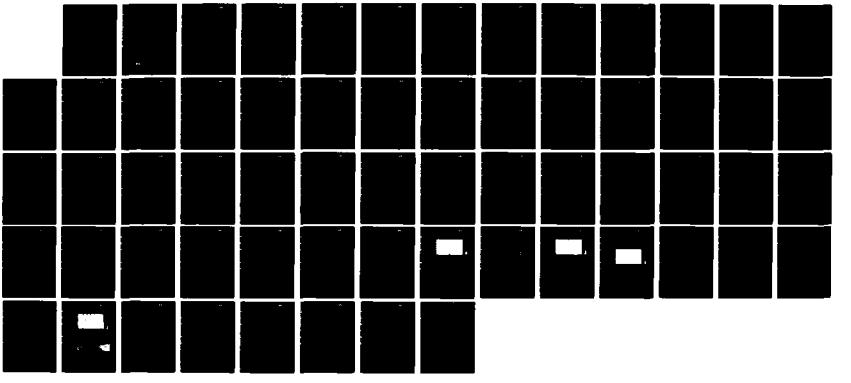
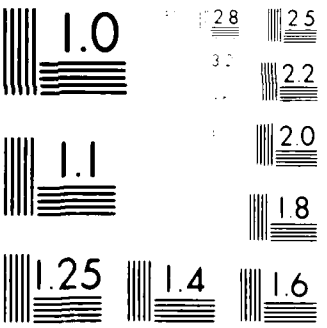
AD-A170 394 HIGH RESOLUTION PASSIVE SONAR IMAGING BY THE PHASE CLOSURE TECHNIQUE PHAS.. (U) ROCKWELL INTERNATIONAL THOUSAND OAKS CA SCIENCE CENTER K A NARSH JUN 86 SC5427.FR N00014-85-C-0182 F/G 17/1 1/1

UNCLASSIFIED NL




1.0
2.8
2.5
3.2
2.2
1.1
2.0
1.8
1.25
1.4
1.6

(12)

SC5427.FR

Copy No. ___

SC5427.FR

HIGH RESOLUTION PASSIVE SONAR IMAGING BY THE PHASE CLOSURE TECHNIQUE

PHASE I: ESTABLISHMENT OF FEASIBILITY

FINAL TECHNICAL REPORT FOR THE PERIOD
February 15, 1985 through May 31, 1986

CONTRACT NO. N00014-85-C-0182

Prepared for

Office of Naval Research
Department of the Navy
800 N. Quincy Street
Arlington, VA 22217

and

Defense Advanced Research Projects Agency
1400 Wilson Boulevard
Arlington, VA 22202

K.A. Marsh
Principal Investigator

JUNE 1986

Approved for public release; distribution unlimited

Rockwell International
Science Center

TABLE OF CONTENTS

A-1

LIST OF FIGURES

LIST OF FIGURES

PREFACE

This final Technical Report, covering the period from May 6, 1985 to May 31, 1986, was prepared by the Rockwell International Science Center under DARPA Contract N00014-85-C-0182. The Program Manager was J.F. Martin and the Principal Investigator was K.A. Marsh. Other major contributors to the program at the Science Center were J.M. Richardson, R.M. Panos, W.A. Guthmiller and B.L. Pierson.

We are grateful to the personnel of the Autonetics Marine Systems Division of Rockwell International for many useful discussions, and for supplying the BQQ-9 data. We particularly wish to thank R. Fairbanks, T. Moore, and W. Wilson for thir support and encouragement.

EXECUTIVE SUMMARY

High Resolution Passive Sonar Imaging by the Phase Closure Technique

Phase I - Establishment of Feasibility

This is a summary of the final report for the first year of effort in the project entitled "High Resolution Passive Sonar Imaging by the Phase Closure Technique". The long-term objective of this effort is to develop techniques for overcoming propagation anomalies and instrumental errors in the location and characterization of noise sources in the ocean using passive sonar.

In the problem of detection, location, and classification of noise sources in the ocean, a number of complicating factors arise. These include refractive index variations due to ocean inhomogeneities in density, temperature, or salinity, drifting of hydrophones in a towed array, and reflection effects from the sea surface, sea floor, or submerged objects. These factors all reduce the spatial coherence of sonar signals. An extreme example is the problem of interpretation of multipath signals in the ocean channel beneath Arctic ice. Even in the open sea, there is still a number of effects which significantly reduce the coherence of received signals. The result is that with conventional beamforming techniques there is a limit to the maximum length of a passive sonar array over which coherence can be maintained at a given frequency. The higher the frequency, of course, the shorter the array which can function successfully. This sets a limit to the resolution obtainable, and hence sets a limit to the detectability and locatability of sources.

The problem is not unique to passive sonar; a similar problem occurs in radio astronomy in the imaging of celestial sources by cross correlating the signals from widely separated receiving

dishes. The distorting medium in that case is the Earth's atmosphere. A technique known as phase closure has, however, been developed to overcome this problem, and over the past few years we have developed this technique for application in sonar. A major difference between the acoustic case and radio astronomy, of course, is that for the latter case the distorting medium represents a thin layer in front of the detectors, whereas for underwater acoustics the inhomogeneities are distributed continuously between the source and receivers. An additional problem in the passive sonar case is that the signal-to-noise ratio is extremely low, and it was not obvious at the outset whether the phase closure technique could function under such conditions. Finally, in a typical sonar patrol situation there are many interfering sources, for example distant shipping, marine life, seismic and wind effects.

Bearing these effects in mind, our objectives in this first phase of the program were to:

1) Extend the regime of validity of the phase closure method to the case of distributed inhomogeneities;

2) Determine the effect of noise on the phase closure technique in a rigorous manner; and

3) Apply the technique to real sonar patrol data.

In the case of the first objective, we arrived at a simple modification of the phase closure technique, a procedure to which we refer as "cross-correlated subarrays". This procedure involves processing the transducer signals in groups, corresponding to operating the array as a series of smaller subarrays. The beamwidth of a subarray is chosen to restrict the response of the system to the angular field over which the assumptions of phase closure are valid. If we let Δx_c be the coherence scale of the system (a function of

the parameters of the medium, together with range R and wavelength λ), then for phase closure to be valid (with or without cross-correlated subarrays) we require $\Delta x_c > (\lambda R)^{1/2}$. Cross-correlated subarrays are required explicitly if the maximum source separation, s, is such that $s > \Delta x_c$. These expectations were tested by means of numerical experiments in which synthetic data were generated for a group of point sources as viewed through an inhomogeneous medium, simulated by 4 random phase screens interspersed between sources and receivers. The object of the experiments was to determine the conditions under which the original source distribution could be reconstructed using phase closure. The results were found to be in agreement with prediction, giving us confidence that the main features of our algorithm are correct.

In the investigation of the behavior of the phase closure technique in the presence of noise, it was found that an improved estimate of the image (bearing as a function of azimuth) must result from the incorporation of phase closure, regardless of the signal-to-noise ratio. The behavior of the phase closure technique is independent of the actual form of the statistical distribution of measurement noise, as a result of the central limit theorem.

The phase closure technique was applied to data from the BQQ-9 towed array in the frequency range 100-1000 Hz. The profile of time delay errors as a function of hydrophone position was estimated from the data, and the result is shown in Fig. 1.

It can be seen that a gradual trend of time delays is present, consistent with either hydrophone drift or oceanic inhomogeneities. An additional effect is the large deviations of hydrophones 39 and 40, which we attribute to some type of instrumental error.

When these time delays were used to correct the data, it was found that an improvement in source intensity of about 25% at 850 Hz resulted.

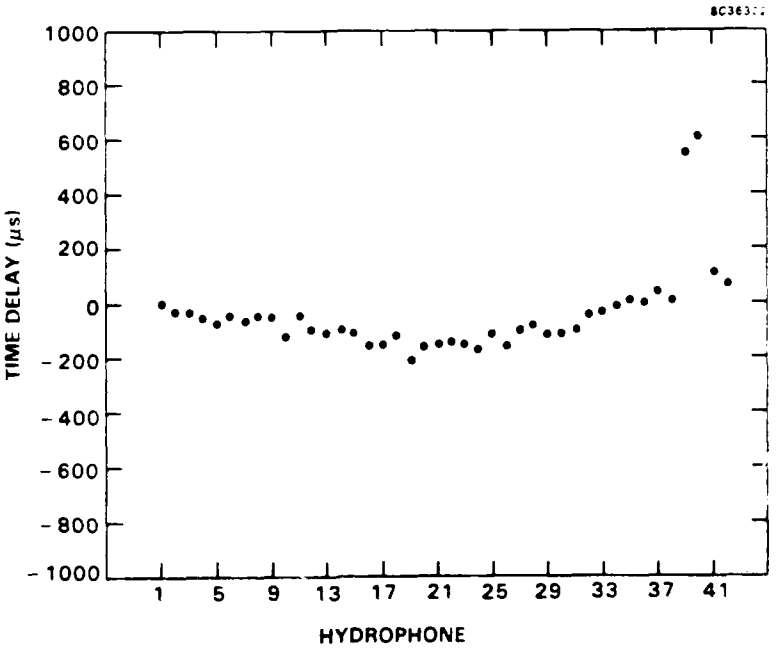


Fig. 1 Plot of experimentally obtained time delay as a function of hydrophone location for the BQQ-9 array.

We then used the computed time delay profile to predict the improvement expected for longer arrays, assuming that for these longer arrays, the phase curvature (2nd derivative of time delay error as a function of hydrophone position) was the same as we measured for the BQQ-9 array. Specifically, pairs of 1-d images (bearing records) were made using synthetic data for a point source at 750 Hz, both without phase errors, and with phase errors applied. The results are shown in Fig. 2.

This figure implies that although only a modest improvement would be expected when phase closure is applied to an array whose length is similar to that of BQQ-9, the improvement for longer arrays should be quite dramatic.

Our results show that the phase closure technique will make it possible to maintain phase coherence over longer arrays than was previously possible, and that source detectability and locatability will be substantially improved. An additional consideration is that

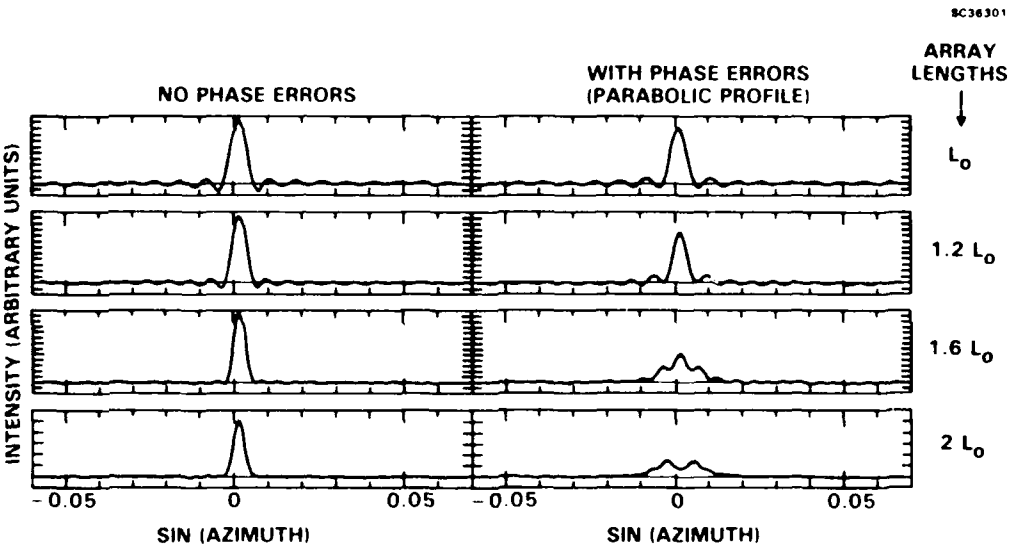


Fig. 2 Synthetically determined bearing records for a point source at 750 Hz for arrays of various lengths, whereby L_o represents the length of the BOQ-9 array. The plots on the left are for data with no phase errors, while the plots on the right were made assuming a parabolic phase error profile with the same phase curvature as measured for the BOQ-9 array.

in recent years, a substantial amount of research effort has been devoted to the improvement of imaging performance by the use of superresolution techniques. However, these techniques depend critically on accurate phase information, and are unstable to phase errors as small as ±5°. Phase closure would thus appear to be an important ingredient for the practical application of any superresolution technique.

1.0 INTRODUCTION

1.1 Background

The spatial resolution obtainable with passive sonar arrays is limited both by propagation anomalies and instrumental phase errors due for example to the drifting of towed elements. Both of these effects set limits to the length over which coherence can be maintained in a towed array. It may, however, be possible to restore the lost coherence, and thereby improve array performance, by the use of suitable signal processing techniques, one candidate being the phase closure technique, developed originally for radio astronomy (Readhead and Wilkinson 1978). An important distinction between this case and passive sonar, however, is that for radio astronomy, the distorting medium represents a thin sheet at one end of the propagation path, whereas for the sonar case, the inhomogeneities responsible for the phase distortion are distributed continuously between the source and receiver.

A closely related technique for overcoming phase distortion effects has recently been discussed by Paulraj and Kailath (1985). Also closely related is the concept of phase recovery from bispectra (Bartelt, Lohmann, and Wirnitzer 1984). As with the basic phase closure technique, both of these methods would be useful in overcoming instrumental phase errors, but would not be able to handle the case of distributed inhomogeneities without some modification.

Preliminary calculations based on available oceanic data indicated that with a sufficiently long array, the phase closure technique might make it possible to obtain spatial resolution high enough to obtain spatial signatures of individual submarines at distances of about 10 km. This was the basis for our original proposal "High Resolution Passive Sonar Imaging by the Phase Closure Technique - Phase 1: Establishment of Feasibility". It was envisaged that this technique could be used with towed arrays, especially in

situations where the measurements could be combined with those of a conformal hull array in order to provide the long baselines necessary for high angular resolution. The feasibility demonstration was to have taken the form of an appropriately scaled experiment in our ultrasonic testbed. However, during the time interval of about one year between the submission of our proposal and the start of the project, our ideas evolved considerably, both as a result of work performed on our own Independent Research and Development funding, and also as a result of discussions with personnel at the Autonetics and Marine Systems Division (AMSD) of Rockwell International in Anaheim, California.

On the theoretical side, we investigated the question of how the unmodified phase closure technique would behave in the presence of distributed inhomogeneities, and performed some numerical experiments to test our ideas. This work is described by Marsh, Richardson, and Martin (1985).

On the practical side, we recall that at the time of submission of the original proposal, our understanding was that Soviet submarines had reduced their narrow-band (line) emission down to the point where broadband noise from the wake was often the only detectable signal remaining. If this were the case, then high resolution spatial imaging would provide an indispensible clue for identification. However, discussions with AMSD personnel indicated that wake detection is very difficult, and that the detection range for broad band noise from a Soviet submarine might only be about 3 km. This may, of course, be due to the poor spatial resolution of existing beamforming techniques, in which case the extra spatial resolution provided by the phase closure technique might facilitate the detection (and imaging) of this broadband feature.

AMSD personnel also mentioned that the narrow-band noise (pumps, motors, reduction gears, and the blades themselves) comes entirely from the aft section of the submarines, with a maximum

extent of 30 m, even for the largest vessels. If we compare this figure with the resolution expected from a reasonably large towed array (assuming a length of 300 m and a frequency of 1000 Hz, which gives a diffraction-limited resolution of 25 m for a source at 10 km) it is apparent that without further modification, this system would be insufficient to obtain a spatial signature. However, the necessary resolution could in principle be obtained either by increasing the baseline length (for example by combining the towed array measurements with those of a conformal hull array) and/or by making use of superresolution techniques based on the prior knowledge that the sources are compact. Several algorithms are available for this latter purpose, for example those discussed by Johnson (1982), as well as the CLEAN algorithm (Hogbom 1974) which is capable of substantial superresolution in the case of pointlike sources.

We also considered whether the phase closure technique would be of benefit for the more modest goal of improving the quality of images obtained from conventional beamformers used in regular sonar patrols. The objective in such cases is to detect and locate point sources in the face of various types of noise and interference. The question then arises as to the principal mechanisms responsible for degrading the currently obtained images. If, for example, the image has been degraded as a result of propagation anomalies or by the drifting of towed elements, then phase closure would be able to improve the resolution. If, on the other hand the principal factors degrading the image are background noise or diffraction, then phase closure would not contribute anything. One can estimate the order of magnitude of the effects of propagation anomalies using the appropriate expression derived by Chernov (1960) together with oceanic parameters given by Flatte et al. (1979). If for example we assume a horizontal towed array of length 300 m and a frequency of 1000 Hz, then using the data of Flatte et al., we find that propagation anomalies would cause serious image degradation for sources at ranges of 4 km and beyond. This suggests that phase

closure would substantially improve the resolution of arrays whose lengths are of the order of 300 m or more. At larger ranges, shorter arrays would benefit also, especially when the data are distorted by the additional effect of the drifting of the towed elements.

1.2 Objectives of This Program

From the considerations discussed above, we concluded that the key issues to be addressed in the present program were:

1) Can the phase closure technique be extended to deal with the case of distributed inhomogeneities?

2) Under what physical conditions does the phase closure technique break down?

3) How will the technique perform in the presence of actual sonar patrol conditions, bearing in mind the extremely low signal-to-noise ratio, equipment malfunctions, and the presence of large numbers of sources?

We therefore drew up a revised plan, designed to address these issues. It did not include the originally proposed ultrasonic testbed experiment, since we felt that although such an experiment would have the merit of providing real data under controlled conditions, it would not be capable of simulating some of the important phenomena encountered in the real world of undersea sonar. Our revised plan, as we discussed with our contract monitor, Dr. R. Fitzgerald, at the beginning of this program, was:

1) Perform theoretical work to extend the range of validity of the phase closure algorithm.

2) Test the new algorithm with synthetic data for a wide range of simulated physical conditions.

3) Perform a theoretical investigation of how the phase closure technique would be affected by large amounts of measurement noise.

4) Test the phase closure algorithm with real sonar patrol data obtained from AMSD.

1.3 Format of This Report

This report will first discuss the basic principles of interferometric imaging, and describe the phase closure technique as it was developed for radioastronomy. This material is presented for introductory purposes. The sections which follow will present the results of work done under the present DARPA program, which will include:

(1) A modification to the phase closure technique designed to extend its validity to the case of distributed inhomogeneities, and an investigation of the physical regime under which the extension will be valid; this will be followed by some numerical experiments designed to test these concepts.

(2) An investigation of the noise question.

(3) The results of our work with real sonar patrol data, together with the implications of our results and the proposed directions of future research.

2.0 THEORETICAL BACKGROUND

2.1 Principles of Interferometric Imaging

Consider a group of hydrophones situated near the origin of a rectangular Cartesian coordinate system x, y, z, whose orientation is such that the z-axis points in the vicinity of a spatially uncorrelated noise source in the far field, as shown in Fig. 3. We further suppose that the received intensity of this noise source (power spectrum per unit solid angle at the origin) at frequency f is $I(\theta,\phi)$ where θ,ϕ are angular coordinates representing small angular offsets in the x and y directions, respectively. Consider a pair of hydrophones i and j whose locations are (x_i, y_i, z_i) and (x_j, y_j, z_j), respectively. The baseline joining i to j can be represented as a vector whose x and y components in units of wavelengths are:

$$u_{ij} = (x_i - x_j)f/c \tag{1}$$

$$v_{ij} = (y_i - y_j)f/c \tag{2}$$

where c is the velocity of sound. The quantities u_{ij} and v_{ij} represent the components of the projected baseline on a plane perpendicular to the line of sight to the source. Suppose the complex frequency-domain signals received at the two hydrophones at frequency f are S_i and S_j, respectively. We define a basic measurable quantity V_{ij} known as the visibility by:

$$V_{ij} = E\ S_i\ S_j^*\ e^{-2\pi i(z_i - z_j)f/c} \tag{3}$$

where E is the expectation operator. With an ergodicity assumption, the visibility corresponds to the time-averaged cross-correlation of the signals from the two transducers in an appropriately narrow frequency band. It can be shown (Devaney 1979) that the measured visibility defined above is related to the far-field intensity distri-

bution $I(\theta,\phi)$ by:

$$V_{ij} = \int\int V(u,v)\ e^{2\pi i(u\theta+v\phi)}du\ dv \quad . \tag{5}$$

Therefore, it would appear that with a sufficient number of hydrophone pairs, we could completely sample the spatial frequency function out to the desired resolution and obtain $I(\theta,\phi)$ using Eq. (5). A serious problem which arises in practice, however, is that inhomogeneities in the medium cause unknown phase shifts and thus destroy the phase information in $V(u,v)$. In addition, the amplitudes may be distorted as a result of interference between micromultipaths. We discuss these effects in the next section.

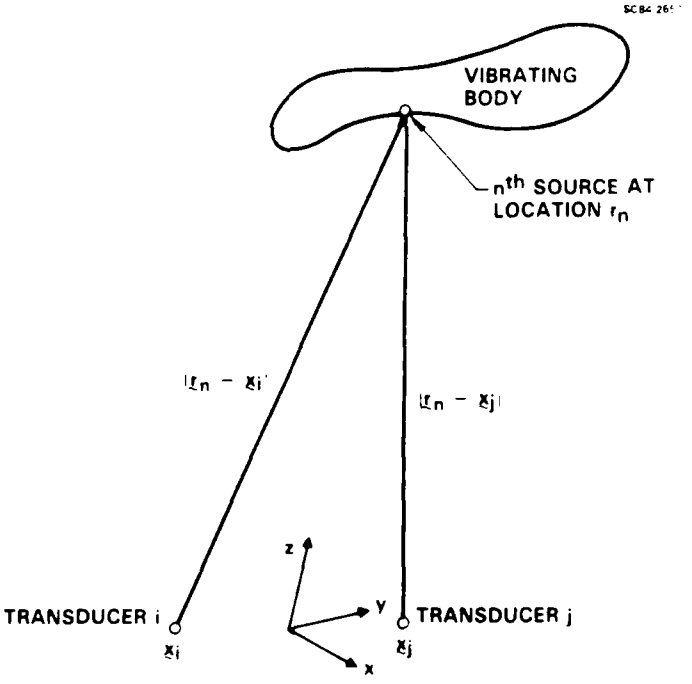


Fig. 3 The geometry of source and receivers.

2.2 The Effect of Propagation Anomalies

In order to represent the effect of propagation anomalies in the medium it is convenient to decompose the source into a large (but finite) number of small sources, as indicated in Fig. 3, the n^{th} source subtending a solid angle $\Delta\Omega_n$ at the center of the receiving array. The effect of the medium can then be expressed in terms of a set of complex gains g_{in}, representing the effect on the signal received at the i^{th} hydrophone due to the n^{th} source component. The measured visibility on baseline i-j is then given by

$$V_{ij}^{meas} = \sum_n g_{in}\, g_{jn}^{*}\, I(\underset{\sim}{e}_n)\, e^{-2\pi i\, \underset{\sim}{u}_{ij}\cdot\underset{\sim}{e}_n}\, \Delta\Omega_n \qquad (6)$$

where $\underset{\sim}{e}_n$ is a unit vector in the direction of the n^{th} source component and $\underset{\sim}{u}_{ij}$ is the baseline vector in wavelengths.

If we assume that the complex gains can be factored into source-adjacent and receiver-adjacent parts, i.e.,

$$g_{in} = g_i g_n \qquad (7)$$

and provided attenuation in the vicinity of the source can be neglected, i.e., $|g_n| = 1$, then we can express the measured visibility in terms of the true visibility V_{ij}^{true} as

$$V_{ij}^{meas} = g_i\, g_j^{*}\, V_{ij}^{true} \quad . \qquad (8)$$

The assumption concerning the factorizability of g_{in} will be discussed in more detail in Section 3.1.1.

It is often convenient to consider the amplitude and phase distortion G_i and ψ_i, separately, hence we will write

$$g_i = G_i e^{i\psi_i} \qquad (9)$$

where G_i is real and positive. We will confine our attention in this report to the case where absorption and interference between multipaths can be neglected. Under these conditions, the principal effect on the signal received at the i^{th} hydrophone is the phase error ψ_i. We can then express the measured phase ϕ_{ij}^{meas} in terms of the true phase ϕ_{ij}^{true} as

$$\phi_{ij}^{meas} = \phi_{ij}^{true} + \psi_i - \psi_j \quad . \tag{10}$$

A useful property of Eq. (10) is that if we add up the measured phases around a closed loop of three baselines corresponding to three transducers, then the ψ's (and hence the effect of the propagation medium) cancel out exactly. Such a sum of phase is called a closure phase (Pearson and Readhead, 1984), and for the transducer set i, j, k is given by:

$$\Phi_{ijk} = \phi_{ij}^{meas} + \phi_{jk}^{meas} + \phi_{ki}^{meas} \tag{11}$$

which from Eq. (10) can be re-expressed as:

$$\Phi_{ijk} = \phi_{ij}^{true} + \phi_{jk}^{true} + \phi_{ki}^{true} \quad . \tag{12}$$

For a symmetrical source, Φ_{ijk} is easily shown to be zero. For all other source geometries, Φ_{ijk} is nonzero in general, and reflects the structure of the source.

A set of N receiving transducers yields (N-1)(N-2)/2 independent closure phases, compared with N(N-1)/2 absolute phases. For a general array, the set of measured closure phases can be used to reconstruct the source intensity distribution by the iterative technique described by Readhead and Wilkinson (1978). If, however, redundant spacings are available, then further simplifications are possible. In the extreme case of a regularly sampled array, a solu-

tion in closed form is possible, and the appropriate algorithm will be discussed in Section 3.3.

It should be emphasized that whereas the phase closure technique is rigorously applicable to the radio astronomical case in which the distorting medium represents a thin sheet at one end of the propagation path, its success in the acoustic case depends on the assumption that g_{in} can be factored according to (7). In the next section, we will discuss the validity of this assumption in the case where the inhomogeneities are distributed throughout the propagating medium.

3.0 DESCRIPTION OF PROGRESS

3.1 Dealing with Distributed Inhomogeneities

3.1.1 Theoretical Considerations

As discussed in Section 2.2, the principal assumption underlying the phase closure technique is that the complex gains g_{in} can be factored into source-adjacent and receiver-adjacent parts, as expressed by Eq. (7). Physically this means the medium is being modelled in terms of a pair of phase screens, one in front of the source and the other in front of the receivers. In the general case of a medium with spatially distributed inhomogeneities, this model may not constitute a good description of the medium itself, but it provides a substantial number of degrees of freedom within which to represent the effects of the medium on acoustic signals. Specifically, these degrees of freedom correspond to the number of receiving elements plus the number of source elements.

For any given set of physical conditions, it is possible to define an angular field-of-view $\Delta\theta_{pc}$ over which (7) represents a valid approximation. Provided all of the acoustic sources are located within this narrow field, the standard phase closure technique is valid. Unfortunately, this is unlikely to be true in practice, since sources occur typically over a wide range of azimuths. Recent work has been devoted to the problem of extending the phase closure technique to handle this situation. A conceptually simple way of accomplishing this is to divide the array up into a number of identical subarrays, and operate each as a small phased array, with the beams all pointed in the same direction. The beam pattern of each subarray will, of course, be much broader than the resolution of the full array. Following astronomical terminology, we will refer to the beam pattern of each subarray as the primary beam. We will treat each subarray as an individual element in a larger scale array, and cross-correlate the signals from each subarray as if we

were dealing with an array of single hydrophones. The effect of this is that the angular response of the larger array will have been multiplied by the primary beam pattern corresponding to an individual subarray. By an appropriate choice of the number of elements in a subarray, we could, provided certain conditions are met, restrict the primary beamwidth $\Delta\theta_{pri}$ to the field $\Delta\theta_{pc}$ over which (7) is satisfied. In the case of a homogeneous ocean, $\Delta\theta_{pri}$ can be obtained simply from the Fourier transform of the aperture distribution for a single subarray. In the case of an inhomogeneous ocean, however, the phase distortions will broaden the response, ultimately setting a lower limit $\Delta\theta_b$ on the primary beamwidth obtainable.

Both $\Delta\theta_{pc}$ and $\Delta\theta_b$ can be related to the phase structure function $D(x - x')$, defined by Flatte et al. (1979) as:

$$D(x - x') = E[\psi(x) - \psi(x')]^2 \tag{13}$$

where x and x' represent the positions of two receivers located at a distance R from a point source and displaced perpendicular to the direction of propagation, $\psi(x)$ represents the spatial variation of the phase of the received signal and E is the expectation operator. We can then express $\Delta\theta_{pc}$ and $\Delta\theta_b$ in terms of some convenient criterion limiting the amount of permissible phase variation $\Delta\psi$ over the source and receiver, respectively. Provided the inhomogeneities have the same statistical character throughout the region between source and receiver, and taking $\Delta\psi$ to be 1 radian, we can then write

$$R\ \Delta\theta_{pc} = \frac{\lambda}{\Delta\theta_b} = \Delta x_c \tag{14}$$

where λ is the wavelength and Δx_c is defined by

$$D(\Delta x_c) = 1 \quad . \tag{15}$$

Provided $\Delta\theta_{pc} > \Delta\theta_b$, the primary beamwidth can be restricted to $\Delta\theta_{pc}$, enabling the phase closure technique to produce an image over the restricted field $\theta_r - \Delta\theta_{pc}/2$ to $\theta_r + \Delta\theta_{pc}/2$ where θ_r represents the center of the restricted field. A mosaic of subimages could then be produced, corresponding to various values of θ_r, and the final image produced by a combination of these subimages. In order to produce the r^{th} subimage, the phase closure technique would be applied to the set of visibilities $V_{ij}^{(r)}$ defined

$$V_{ij}^{(r)} = \sum_{k=k_o(i)+1}^{k_o(i)+n_s} \sum_{\ell=\ell_o(j)+1}^{\ell_o(j)+n_s} a_k(i)\, a_\ell(j)\, E(S_k S_\ell^*)\, e^{-2\pi i u_{k\ell} \sin\theta_r} \qquad (16)$$

where

$$k_o(i) = (i - 1)n_s$$

$$\ell_o(j) = (j - 1)n_s$$

and S_k, S_ℓ represent the signals received by the k^{th} and ℓ^{th} hydrophones; $a_k(i)$, $a_\ell(j)$ are weights representing the apodizing function for each subarray, and n_s is the number of elements in each subarray.

The new algorithm sketched above would enable phase-closure imaging over a wide field-of-view, and should provide a substantial improvement in imaging performance over that obtained by conventional beamforming. We will refer to it subsequently as "phase closure with cross-correlated subarrays." We now consider the question as to the physical regime over which this algorithm will be valid. A useful quantity to bear in mind in this regard is Δx_c defined by (15), which represents the coherence scale for the system at the particular range and frequency. The condition that $\Delta\theta_{pc} > \Delta\theta_b$ is equivalent to

$$\Delta x_c > (\lambda R)^{1/2} \quad . \qquad (17)$$

Equation (17) thus represents a criterion by which one can determine whether or not phase closure (with cross-correlated subarrays) is capable of improving a distorted image. In order to relate it to more fundamental physical parameters, however, requires a knowledge of the form of the phase structure function D, which in turn is related to the spatial correlation function of the refractive index. If we assume that

$$E\psi(x)\ \psi(x') = \Phi^2 \exp\left(-(x-x')^2/L^2\right) \tag{18}$$

then provided $|x - x'| \ll L$,

$$D(x - x') \approx 2\Phi^2 (x - x')^2/L^2 \tag{19}$$

and hence

$$\Delta x_c \approx \frac{L}{\sqrt{2}\,\Phi} \quad . \tag{20}$$

If the innomogeneities are statistically uniform and isotropic, then from Flatte et al., (1979):

$$\Phi^2 \approx 0.4 \left(\frac{2\pi}{\lambda}\right)^2 \langle\mu^2\rangle\, RL \tag{21}$$

where $\langle\mu^2\rangle$ represents the variance of the refractive index fluctuations. Equations (20) and (21) then give

$$\Delta x_c \approx 0.2\ \lambda \left(\frac{L}{\langle\mu^2\rangle\, R}\right)^{1/2} \quad . \tag{22}$$

Thus from (17), we find that the maximum range over which the phase closure technique (with cross-correlated subarrays) can be applied to a distributed medium is

$$R_{max} = 0.2 \left(\frac{\lambda L}{\langle\mu^2\rangle}\right)^{1/2} \quad . \tag{23}$$

In the case of horizontal imaging in the ocean at 500 Hz (assuming L = 5 km and $\langle\mu^2\rangle^{1/2} = 5 \times 10^{-4}$) we obtain R_{max} = 49 km, which suggests that substantial improvement in sonar images from horizontal arrays should be possible over ranges up to this value.

3.1.2 Numerical Experiments

In order to test the concepts discussed in the previous section, a series of numerical experiments was performed. In these experiments, an assumed source consisting of 4 pointlike components spread over a distance of 200 m was observed at a distance of 1 km and a frequency of 300 Hz, using a regularly spaced array of length 200 m containing 21 hydrophones. The sound velocity was assumed to be 1500 m/s. Ocean inhomogeneities were simulated by placing four equally spaced phase screens between the source and receivers. The direction of propagation was taken to be the z-direction. Along each phase screen (taken as the x-direction), the phase was made to vary randomly, as a Gaussian process with a correlation length L, and a standard deviation such that the rms phase deviation through the entire system was approximately 1 radian, so that L corresponds to Δx_c. Various values of L were assumed, in some cases chosen purposely to violate the assumptions for phase closure. The propagation was calculated according to the Fresnel approximation, as discussed by Marsh, Richardson and Martin (1985). In these calculations, the refractive index variations were assumed to be independent of the y coordinate, for reasons of computational economy, and hence the propagation calculations were confined to the xz-plane. In the Fresnel approximation, the complex pressure $\Psi(x,z)$ is then related to that at distance z+d by:

$$\Psi(x,z+d) = \int \Psi(x,z)\, T(x-x',\, d)\, dx' \qquad (24)$$

where

$$T(x-x',d) = e^{iKd} \left(\frac{K}{2\pi id}\right)^{1/2} \exp\left(\frac{-K(x-x')^2}{2id}\right) \qquad (25)$$

and

$$K = \frac{2\pi f}{c} \quad . \qquad (26)$$

For computational purposes, the phase variation along each phase screen was sampled spatially at intervals of one tenth of a wavelength over a distance of 500 m, and the integration in (24) performed numerically using the trapezoidal rule.

In the case of the cross-correlated subarrays algorithm, each subarray had 5 elements, weighted with a rectangular function, i.e., no apodizing was used. A mosaic of 3 subimages was produced in each case, spaced by the half-power widths of the corresponding primary beams (sinc functions). Simple addition was used in order to construct the final image from the 3 subimages.

Since in these experiments $L \approx \Delta x_c$, Eq. (17) implies that phase closure would be invalid if $L < (\lambda R)^{1/2}$. In the case of $L > (\lambda R)^{1/2}$, phase closure would be applicable, but cross-correlated subarrays would be required if the source components are separated by a distance s greater than L. These two criteria form a rather crucial test of the theoretical basis of our algorithm, and values of L were chosen to test them. In these experiments, $(\lambda R)^{1/2}$ corresponds to 71 m. We now discuss the results.

Case a: $L > (\lambda R)^{1/2}$, $s = L$.

The assumed value of L was 200 m. In this case we expect phase closure to be valid, and cross-correlated subarrays should not be necessary. The results are shown in Fig. 4, which fulfills our expectations. The use of phase closure has substantially improved the image over that obtained by conventional beamforming, but the

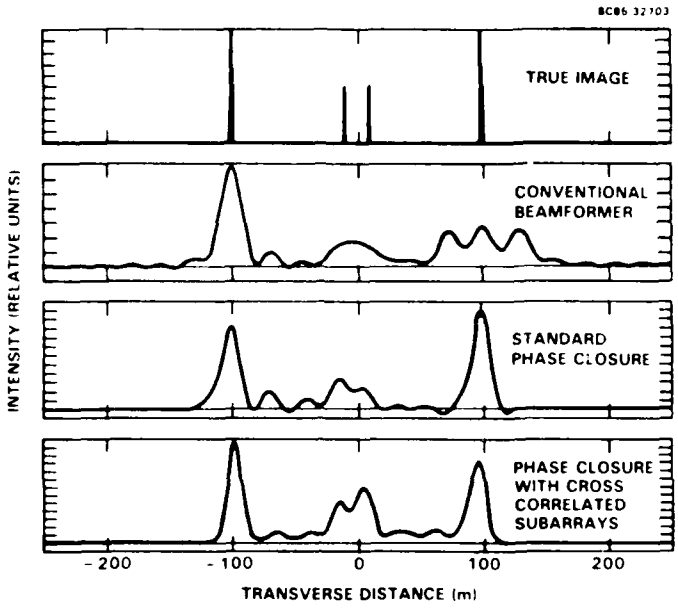


Fig. 4 Imaging results for the case of mild phase distortion. The length scale of the inhomogeneities, L, was chosen such that $L > (\lambda R)^{1/2}$ where λ is the wavelength and R is the range, thus phase closure was expected to be valid. In addition, the maximum source separation, s, did not exceed L, and hence, standard phase closure performs satisfactorily without the need for cross-correlated subarrays.

use of cross-correlated subarrays has produced only a marginal improvement over standard phase closure.

Case b: $L > (\lambda R)^{1/2}$, $s = 2L$.

The assumed value of L was 100 m. We expect phase closure to be valid, but the sources are now so widely spaced in comparison to the length scale of the inhomogeneities that cross-correlated subarrays should be required. The results, shown in Fig. 5 support these expectations. Standard phase closure gave a very poor result, whereas cross-correlated subarrays gave a reasonable reconstruction.

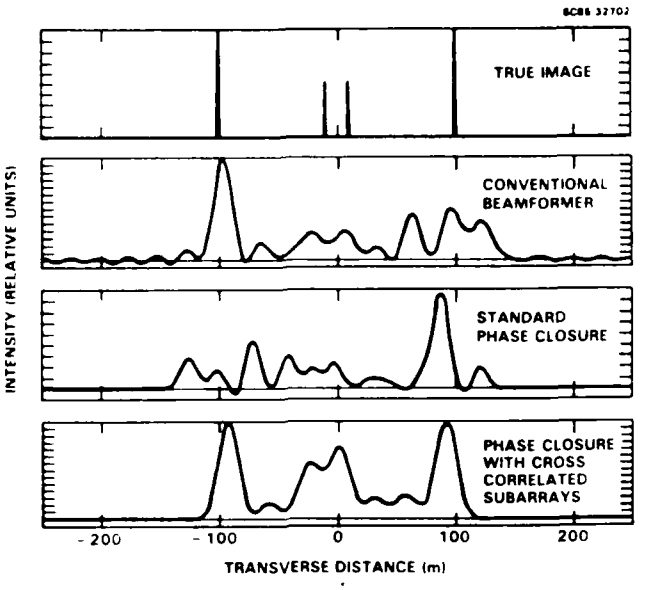


Fig. 5 Imaging results for the case of moderate phase distortion. Parameters are the same as for Fig. 4, except that s = 2L, i.e., the sources are now so widely spaced in comparison to the length scale of the inhomogeneities that cross-correlated subarrays are required.

Case c: $L < (\lambda R)^{1/2}$.

The assumed value of L was 50 m. In this case we expect phase closure to be completely invalid, and the results in Fig. 6 show this to be true. It is interesting, however, that the cross-correlated subarrays image bears some resemblance to the assumed source, altho gh the separation between components is incorrect.

3.1.3 Summary

In this section we have discussed an improvement that allows the phase closure technique to deal more correctly with the case of distributed inhomogeneities. Our modification of the phase closure technique, to which we refer as "cross-correlated subarrays" involves processing the hydrophone signals in groups, corresponding to operating the array as a series of smaller subarrays. The beam-

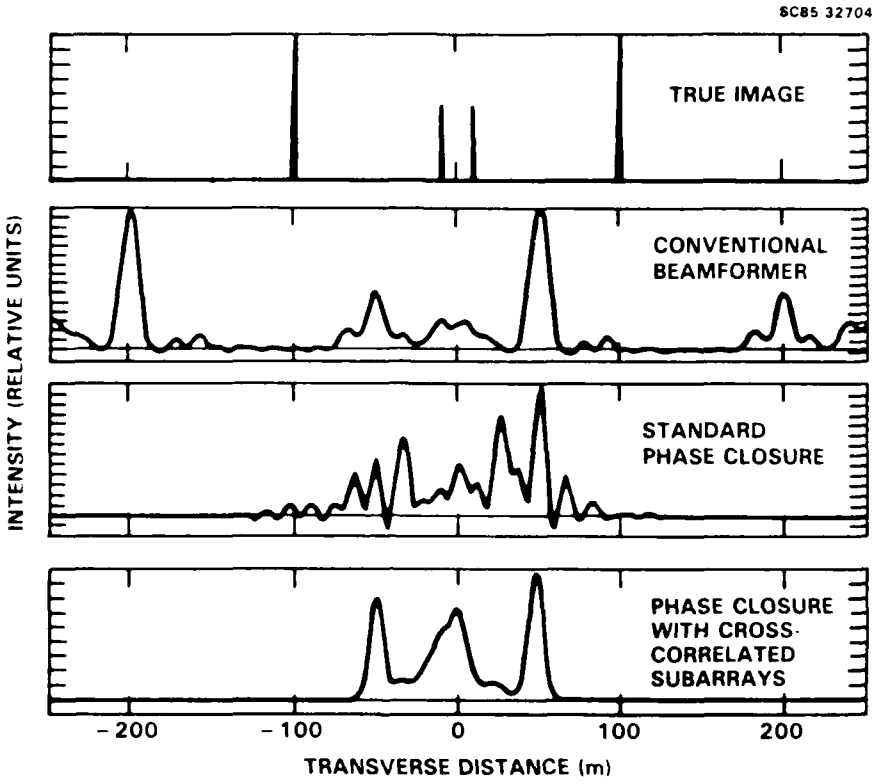


Fig. 6 Imaging results for the case of strong phase distortion. In this case $L < (\lambda R)^{1/2}$, and hence, phase closure is not able to restore the image.

width of a subarray is chosen to restrict the response of the system to the angular field over which the assumptions of phase closure are valid. If we let Δx_c be the coherence scale of the system (a function of oceanic parameters, together with range R and wavelength λ), then for phase closure to be valid (with or without cross-correlated subarrays) we require $\Delta x_c > (\lambda R)^{1/2}$. Cross-correlated subarrays are required explicitly if the maximum source separation, s, is such that $s > \Delta x_c$. These expectations have been tested by means of a series of numerical experiments. The results are consistent with the predictions, giving us confidence that the main features of our

algorithm are correct. An important future step is to apply the algorithm to actual sonar patrol data. Before doing so, however, one other major question must be addressed, and that is the behavior of the phase closure technique in the presence of noise. This will be the subject of the next section.

3.2 Phase Closure in the Presence of Severe Measurement Noise

In order to examine the effect of noise on the phase closure technique in a rigorous manner, it will be necessary to start from first principles. We begin by stating our basic model for the signal distortion introduced by ocean inhomogeneities, and then include the effect of noise, and finally discuss the considerations involved in the inversion itself.

3.2.1 Model for Signal Distortion

Consider a point source at location $\underset{\sim}{r}_o$, whose complex strength at frequency ω is $f_o(\omega)$, being monitored by the i^{th} hydrophone at location $\underset{\sim}{x}_i$. Suppose the intervening medium is inhomogeneous, and that in general there are multiple propagation paths between source and receiver, the ℓ^{th} such path having attenuation $a_{i\ell}$ and time delay $\tau_{i\ell}$ with respect to that expected in a homogeneous ocean. In the absence of noise, the received signal can then be expressed

$$S_i(\omega) = (g_i)_o \frac{f_o(\omega)}{|\underset{\sim}{r}_o - \underset{\sim}{x}_i|} e^{-\frac{i\omega}{c}|\underset{\sim}{r}_o - \underset{\sim}{x}_i|} \tag{27}$$

where $(g_i)_o$ is a complex gain factor representing the effect of ocean inhomogeneities, and is given by

$$(g_i)_o = \sum_{\ell=1}^{L_i} a_{i\ell} e^{-i\omega\tau_{i\ell}} \tag{28}$$

where L_i is the number of propagation paths. The quantities $a_{i\ell}$ and $\tau_{i\ell}$ are , or course, functions of source position, and hence in the

case of multiple sources, $(g_i)_o$ is in general different for each source. Let the k^{th} source have complex strength $f_k(\omega)$ and position $\underset{\sim}{r}_k$, and suppose that the complex propagation gain between this source and the i^{th} hydrophone is g_{ik}. The received signal is then

$$S_i(\omega) = \sum_k g_{ik} \frac{f_k(\omega)}{|\underset{\sim}{r}_k - \underset{\sim}{x}_i|} e^{-\frac{i\omega}{c}|\underset{\sim}{r}_k - \underset{\sim}{x}_i|} \quad . \tag{29}$$

As discussed in Sectin 3.1.1, one may, over an appropriately small region R_g, make the approximation

$$g_{ik} \simeq g_i g_k \quad . \tag{30}$$

If the sources are confined to R_g, or if the angular response of the array can be restricted as discussed in 3.1.1, then we can rewrite (29) as

$$S_i(\omega) = g_i \tilde{S}_i(\omega) \tag{31}$$

where $\tilde{S}_i(\omega)$ is given by

$$\tilde{S}_i(\omega) = \sum_k g_k \frac{f_k(\omega)}{|\underset{\sim}{r}_k - \underset{\sim}{x}_i|} e^{-\frac{i\omega}{c}|\underset{\sim}{r}_k - \underset{\sim}{x}_i|} \quad . \tag{32}$$

Equations (31) and (32) form the basis for the measurement model used in the estimation of intensity as a function of direction. The complete measurement model must, of course, include the effect of noise, and we now discuss this aspect. For brevity we will omit explicit reference to the frequency dependence in quantities such as $S_i(\omega)$, and simply write S_i.

3.2.2 Effect of Noise

We represent the received frequency-domain signal on the i^{th} hydrophone as

$$S_i = g_i \tilde{S}_i + \nu_i \tag{33}$$

where ν_i represents the noise. The factor g_i represents the effect of ocean inhomogeneities as discussed above. It may also include instrumental errors in gain or delay.

We now define an important quantity known as the visibility (also called the coherence function) with respect to the pair of hyrophones i and j, by:

$$V^o_{ij} = E\, S_i S^*_j \tag{34}$$

where E is the expectation operator. In practice, we assume ergodicity and replace the expectation operator with a time average. The result is the "measured visibility" V_{ij}, given by

$$V_{ij} = \frac{1}{M} \sum_{m=1}^{M} (S_i S^*_j)_m \tag{35}$$

where $(S_i S^*_j)_m$ is the m^{th} time sample of $S_i S^*_j$.

We can relate V_{ij} to V^o_{ij} by

$$V_{ij} = V^o_{ij} + \eta_{ij} \tag{36}$$

where

$$E\, \eta_{ij} = 0 \quad . \tag{37}$$

By the central limit theorem, if M is sufficiently large, and the time samples uncorrelated, then η_{ij} represents a Gaussian random process, even if ν_i is non-Gaussian. If we assume that ν_i is statistically independent of $\tilde{S}_i$, then from (33) we have

$$V^{\circ}_{ij} = g_i g_j^{*} E(\tilde{S}_i \tilde{S}_j^{*}) + (C_v)_{ij} \tag{38}$$

where $(C_v)_{ij}$ is the spatial covariance of the noise, given by

$$(C_v)_{ij} = E\ v_i v_j^{*} \quad . \tag{39}$$

If we assume that the sources are spatially uncorrelated, then from (32) we can show that in the far field

$$E\ \tilde{S}_i \tilde{S}_j^{*} = \frac{1}{R^2} \sum_k |g_k|^2\ E|f_k|^2\ e^{-2\pi i \underset{\sim}{u}_{ij} \cdot \underset{\sim}{e}_k} \tag{40}$$

where $\underset{\sim}{e}_k$ is a unit vector in the direction of the k^{th} source, R is the range, and

$$\underset{\sim}{u}_{ij} = (\underset{\sim}{x}_i - \underset{\sim}{x}_j)\omega/(2\pi c) \quad . \tag{41}$$

We wish to express (40) in terms of the intensity profile observed at the receivers. For this purpose it will be convenient to divide the angular field-of-view into a large number of elements, each of solid angle $\Delta\Omega$, and let the index k refer to the k^{th} such element. We define the true source intensity $I(\underset{\sim}{e}_k)$ as the power spectrum per unit area per unit solid angle in direction $\underset{\sim}{e}_k$ which would have been received in the absence of inhomogeneities. If we neglect attenuation in the vicinity of the source (i.e., $|g_k| = 1$), then (40) becomes

$$E\ \tilde{S}_i \tilde{S}_j^{*} = \sum_k I(\underset{\sim}{e}_k)\ e^{-2\pi i \underset{\sim}{u}_{ij} \cdot \underset{\sim}{e}_k}\ \Delta\Omega \quad . \tag{42}$$

If attenuation in the vicinity of the source cannot be neglected, then the intensity in (42) should be regarded as an effective intensity.

Combining (36), (38) and (42) we obtain the following equation, which may be considered as our formal measurement model:

$$V_{ij} = g_i g_j^* \sum_k I(\underset{\sim}{e}_k)\, e^{-2\pi i \underset{\sim}{u}_{ij} \cdot \underset{\sim}{e}_k}\, \Delta\Omega + (C_\nu)_{ij} + \eta_{ij} \qquad (43)$$

which can be rewritten in matrix notation as

$$V = Ax + C_\nu + \eta \qquad (44)$$

where A is a matrix whose components are

$$A_{ij,k} = g_i g_j^*\, e^{-2\pi i \underset{\sim}{u}_{ij} \cdot \underset{\sim}{e}_k}\, \Delta\Omega \qquad (45)$$

and x is the "image" vector whose components are

$$x_k = I(\underset{\sim}{e}_k) \quad . \qquad (46)$$

3.2.3 Inversion Considerations

If we take the most probable value of x as our optimal estimate, then our problem can be stated in terms of maximizing $P(x,g|V)$, the probability of x and g conditioned on the measurements V, where g represents the set of g_i values. We can express this conditional probability as

$$P(x,g|V) = P(V|x,g) \cdot P(x)P(g)/P(V) \qquad (47)$$

where we have assumed x and g to be statistically independent *a priori*. The quantity P(V) may be treated as an ignorable constant for optimization purposes. As stated earlier, η will be Gaussian according to the central limit theorem, and hence

$$\ln P(V|x,g) = -(V - Ax - C_\nu)^\dagger\, C_\eta^{-1} (V - Ax - C_\nu) + \text{const} \quad . \qquad (48)$$

where

$$C_\eta = E\eta\eta^\dagger \quad . \qquad (49)$$

and the symbol "†" represents Hermitian transpose.

We will assume a flat prior distribution for g, i.e.,

$$\ln P(g) = \text{const} \quad . \tag{50}$$

The only missing ingredient at this stage is the a priori statistical distribution for x, i.e., P(x). Various forms for P(x) can be used, depending on our prior knowledge. One piece of such knowledge is that intensity is always positive, i.e.,

$$P(x) = 0 \qquad \text{if} \quad \text{Ik for which } x_k < 0 \quad . \tag{51}$$

In addition, one could make use of the fact that with currently used arrays, the targets of interest are spatially unresolved, and hence x consists of a limited number of point sources. This fact is exploited implicitly in many of the so-called superresolution techniques (see for example Johnson, 1982). In the context of P(x), it could be expressed as:

$$P(x) = \prod_k [\alpha\delta(x_k) + (1 - \alpha)\, Q(x_k)] \tag{52}$$

where Q is a slowly varying function such that

$$\int_{-\infty}^{\infty} Q(\xi)\, d\xi = 1 \quad . \tag{53}$$

An alternate philosophy preferred by some investigators is that of maximum entropy, in which P(x) takes the form:

$$\ln P(x) = -\beta \sum_k x_k \ln x_k \quad . \tag{54}$$

Having selected an appropriate form for P(x), one is then in a position to obtain the optimal estimate of x by maximizing

$$\ln P(x,g|V) = -(V - Ax - C_\nu)^\dagger C_\eta^{-1} (V - Ax - C_\nu) + \ln P(x) + \text{const} \quad .$$

with respect to x and g, where V is related to the raw hydrophone signals by (35) and A is given by (45).

The difference between the phase closure algorithm and the standard approach is that in the latter case, the g_i's are all assumed to be unity. Since in general this assumption will be inconsistent with the measurements, the standard approach will give a suboptimal estiamte for x. On the other hand, the phase closure technique allows the g_i's to take values consistent with the measurements, and hence should yield an improved estimate for x. Strictly speaking, the term "phase closure" refers to an algorithm with which we incorporate the effect of phase distortion only. In the present discussion we are using the term in a more general sense to include the effect of amplitude distortion also.

3.2.4 Conclusion

The above analysis has shown that an improved estimate of x results from the incorporation of phase closure, regardless of the statistical distribution of measurement noise. It is also true regardless of the actual signal-to-noise ratio, and thus the above considerations should be applicable under typical sonar conditions whereby the signal-to-noise ratio on a single hydrophone may be as low as -40 dB.

3.3 Practical Method of Solution

So far, we have discussed the problem of phase errors ϕ_i in terms of the single frequency case. The ϕ_i's are, of course, frequency dependent. In a practical situation, one is dealing with broadband data, and hence one could reduce considerably the number of variables by solving instead for the set of time delays τ_i, related to ϕ_i by

$$\psi_i = 2\pi f \tau_i \quad . \tag{56}$$

The ψ_i's are related to the components of g in Section 3.2.3 by (9).

In principle, one could obtain the most probable set of ψ_i conditioned on the data by maximizing ln $P(x,g|V)$ from Eq. (55). This, however, is a difficult procedure since it involves finding the global maximum of a non-convex function. Two possible approaches are available:

1) Write the single-frequency version of the measurement model in terms of phase, as in Eq. (10), with additive Gaussian noise. The problem then reduces to simple linear estimation. Difficulties arise in cases of low signal to noise, however, due to the lobe ambiguity problem which will be elaborated upon below.

2) By using a sufficiently wide range of frequencies, ln $P(x,g|V)$ may be approximately convex with respect to $\{\tau_i\}$, and it may therefore be possible to maximize it using the gradient method provided suitable precautions are taken to ensure that the local maximum is indeed the global one.

In the present analysis we opted for the first possibility, and we now discuss the essential details of the algorithm. We will restrict this discussion to the case of the regularly spaced array, since this was appropriate for the patrol data analyzed in the next section. A considerable simplification results from the regularly-spaced case, in that there is sufficient redundancy in the measurements that the set of τ_i can be obtained independent of any knowledge of the source. This is the principal difference between the

algorithm discussed below and that used by Readhead and Wilkinson (1978).

We assume a measurement model such that the measured phase $\tilde{\phi}_{ij}(\omega_m)$ on baseline ij at frequency ω_m is related to the true phase $\phi_{ij}(\omega_m)$ by:

$$\phi_{ij}(\omega_m) = \tilde{\phi}_{ij}(\omega_m) + \omega_m (\tau_i - \tau_j) + \varepsilon_{ij}(\omega_m) \qquad (57)$$

where $\varepsilon_{ij}(\omega_m)$ represents the noise, assumed to be Gaussian. The indices i and j take values between 1 and N, where N is the number of hydrophones. There are N (N-1)/2 independent relations of the form (57).

The redundancy in a regularly spaced array gives us the additional set of relations

$$\tilde{\phi}_{i,i+k}(\omega_m) = \tilde{\phi}_{1,1+k}(\omega_m) \qquad k = 1,\ldots N-1 \quad . \qquad (58)$$

Since it is only the time delay differences which are important, one can arbitrarily set

$$\tau_1 = 0 \quad . \qquad (59)$$

If M frequencies are present in the data, then by combining (57), (58) and (59) we end up with a system of M(N-1)(N-2)/2 equations and (M+1)(N-1) unknowns. The unknowns are the N-1 values of τ_i, together with the N-1 values of $\tilde{\phi}_{1,1+k}$ for each frequency.

This constitutes a rather formidable system of linear equations. One way of reducing it to tractable proportions is to divide (57) by ω_m, and sum over m. One is then left with (N-1)(N-2)/2 equations in 2(N-1) unknowns. A major problem with this is that each measured phase is modulo 2π, and hence each of the (N-1)(N-2)/2

equations would have an additive term corresponding to an unknown multiple of 2π. This is the lobe ambiguity problem mentioned earlier. One could, in principle, overcome it by working at sufficiently low frequencies that the time delays τ_i were each a small fraction of a period. However, in cases of low signal-to-noise, even this would be invalid because of lobe ambiguities due to the cumulative effects of $\varepsilon_{ij}(\omega_n)$.

The solution adopted in the present investigation was to first average the complex visibilities over the appropriate frequency band, and then use the phases of these averaged visibilities together with the "single-frequency" formalism, corresponding to the mid-frequency of the band, ω_0. There are two benefits of the averaging procedure:

1) The phases of the averaged visibilities have considerably less noise than the phases of the raw visibilities and hence there is less likelihood of noise-related lobe ambiguity problems.

2) When averaging over a frequency band $\Delta\omega$, the response to a source at angular distance θ from the phase center is reduced by a factor

$$F(\theta) = \text{sinc}\left(\frac{\ell\ \Delta\omega\ \theta}{2c}\right) \tag{60}$$

where ℓ is the baseline length. This is known as the "bandwidth effect." It can be used to advantage in the suppression of interfering signals. In order to do this, one must apply appropriate time delays in order to position the phase center in the desired direction, and this must take place before the averaging procedure. Sources far from the phase center are then suppressed.

Our measurement model then becomes:

$$\phi_{i,i+k} = \tilde{\phi}_{1,1+k} + \omega_0(\tau_i - \tau_{i+k}) + \varepsilon_{i,i+k} \qquad \begin{matrix} k = 1,\ldots N-1 \\ i = 1,\ldots N-k \end{matrix} \tag{61}$$

subject to (59).

An additional constraint is required in order to make possible a unique solution. A convenient constraint for this purpose is obtained by assuming that one of the true phases is known, for example $\tilde{\phi}_{12}$. We will thus express this constraint as:

$$\tilde{\phi}_{12} = \frac{\alpha}{\omega_0} \tag{62}$$

and say more about our choice for α later.

Consider now the solution procedure for $\{\tau_i\}$. We can express (59), (61) and (62) as a single matrix equation of the form:

$$y = Ax + v \tag{63}$$

where y, b and v are vectors of dimensionality $n = (N-1)(N-2)/2$, x is a vector of dimensionality $n = 2N-3$, and A is an m x n matrix. The first N-1 components of x represent $\tau_2,\ldots\tau_N$, while the remaining components represent the unknown $\tilde{\phi}$'s. The vector v represents the noise, while the components of A and y are given by:

$$A_{ik,j} = \begin{cases} 1 & \text{if } j = i-1 \\ -1 & \text{if } j = i+k-1 \\ 0 & \text{otherwise} \end{cases} \qquad j = 1, \ldots N-1$$

$$\begin{cases} 1-\delta_{k1} & \text{if } j = k+N-2 \\ 0 & \text{otherwise} \end{cases} \qquad j = N, \ldots n \tag{64}$$

$$y_{ik} = \frac{\phi_{i,i+k}}{\omega_o} - \alpha\,\delta_{k1} \quad . \tag{65}$$

The ranges of i and k in (64) and (65) are:

$k = 1, \ldots N-1$

$i = 1, \ldots N-k$.

If we assume that ν is Gaussian, with a diagonal covariance matrix, then the maximum likelihood estimate of x, and hence $\{\tau_i\}$, is given by:

$$\hat{x} = (A^T A)^{-1} A^T y \quad . \tag{66}$$

The covariance of x is given by:

$$C_x = \frac{(y - Ax)^T (y - Ax)}{n - m} (A^T A)^{-1} \quad . \tag{67}$$

In order to complete the algorithm, it remains to specify α, or equivalently $\tilde{\phi}_{12}$. Since for a regularly spaced array of N hydrophones, the phases on the N-1 shortest baselines are all equal to $\tilde{\phi}_{12}$ plus error terms, one approach might be to average the visibilities on these N-1 baselines and consider the phase of this averaged visibility to be our best estimate of $\tilde{\phi}_{12}$. Alternately, one can select α on the basis of an a priori bias against excessive time delays, i.e., one can select the value of α which minimizes the sum of squares of $\tau_2, \ldots \tau_N$. This was the procedure used for the present work.

3.4 Application of the Phase Closure Technique to Real Sonar Patrol Data

3.4.1 Considerations Involved in Data Selection

One of the primary goals of this project was to test the concepts of phase closure using real sonar patrol data from a towed array. Our objectives were to answer the following questions:

1) Is there evidence for systematic time delay errors associated with individual hydrophone elements?

2) If so, when the data are corrected, is source detectability improved?

In order to provide the best possible test of the phase closure technique, it was necessary to use the longest possible array at the highest possible frequency. High frequency data would be most susceptible to time delay errors due both to the drifting of the hydrophones and to ocean inhomogeneities, while the use of a long array would enhance the latter effect. However, currently used towed arrays are designed so as to be short enough not to suffer too much coherence loss as a result of propagation anomalies and hydrophone drift. This means that long arrays are used only at low frequencies while short arrays are used at high frequencies. In the absence of techniques such as phase closure, this is a sensible thing to do. However, if arrays were made longer, the phase closure technique could restore the lost coherence and thus enable greater spatial resolution and detectability. Thus currently existing data are far from ideal for our purpose. Fortunately, however, our enquiries indicated that some data did exist for moderately long arrays, with time sampling sufficiently fine to enable high frequency analysis. The drawback was that the arrays from which these data were taken were designed for low frequency use, and at high

frequencies the spatial sampling would be highly inadequate. Nevertheless, it was felt that with a proper understanding of these limitations, something useful could still be learned from an application of phase closure principles to such data.

The data for this purpose were obtained from Autonetics and Marine Systems Division (AMSD) of Rockwell International, in Anaheim, California.

3.4.2 Description of the Data

The data were taken from a linear array consisting of 42 hydrophones equally spaced, whose signals were sampled at a rate of 3072 Hz. The recording took place in the Atlantic Ocean off Charleston, S.C., during a "pre-OPEVAL" exercise for the TASPE sonar system in late October and early November in 1983. The particular data analysed by us were recorded during day 304, hour 2. During this time period the ship towing the array was moving at about 20 knots at a depth of 350 ft; no significant heading changes occurred during this time.

The sonar array used during this exercise was the AN/BQQ-9 (SPALT 9080) system. The full system consisted of a long low frequency subarray together with a shorter high frequency subarray. Each subarray consisted of 42 elements, some elements of which were in common. The low frequency subarray was used for our analysis, for reasons discussed in Section 3.4.1. Frequencies of up to 1000 Hz were used, considerably above the nominal design frequency of the subarray, although within the passband of individual hydrophones. This procedure incurred certain penalties, which will be discussed in detail. In this report, for purposes of brevity, we will use the term "BQQ-9" to refer to the low frequency subarray only.

The original data were recorded on high density tapes. The selected data were transferred to 9-track 1600 b.p.i. tapes by AMSD

personnel. The analysis was performed on our VAX 11/780 at the Science Center.

3.4.3 Method of Analysis

Before subjecting the data to our phase closure algorithm, it was necessary to apply some conventional methods of analysis in order to see what was there, and to select a subset of the data for more detailed analysis. The first step was to construct a broadband bearing-time record (BTR) for all of the data. The second was to use this bearing-time record to select time intervals of interest, and calculate a frequency-azimuth (FRAZ) plot for each interval. The procedures for calculating these were:

1) Bearing-time record: This was accomplished by a simple delay-and-sum beamformer, whereby the output of the m-th beam at the j-th time is given by:

$$y_{im} = \frac{1}{2\Delta\ell} \sum_{i=1}^{N} \sum_{\ell=-\Delta\ell+1}^{\Delta\ell} v_{i,j+k+\ell} \qquad (68)$$

where

$$k = \text{int}\left[- \frac{(i-1)\Delta x \sin\theta_m}{c\Delta t}\right] \qquad (69)$$

such that int [A] means "nearest integer to A", v_{ij} is the j-th time sample of the signal from the i-th hydrophone, Δt is the time sampling increment, c is the sound speed, Δx is the spacing between adjacent elements of the array, and θ_m is the azimuth, defined as zero broadside to the array, and $-\pi/2$ in the direction of the towing ship. The purpose of the sum over ℓ in (68) is to create a low-pass filter, in order to eliminate the aliasing which would occur in a simple bearing-time record due to

spatial undersampling. The low-pass filter used here is a rectangular averaging window in the time domain, corresponding to a sinc function in the frequency domain. The averaging window was designed to give a cutoff frequency corresponding to the nominal design frequency of the array.

The power in the m-th beam is then given by:

$$P_m = \frac{1}{J} \sum_{j=1}^{J} y_{jm}^2 \qquad (70)$$

where J represents the integration time in units of Δt. A total of 42 beams was used, covering the range sin θ = -1 to +1.

Before displaying the bearing record, two further processing steps were employed, namely background substraction and scaling. These operations have been found to enhance the detectability of sources in BTR's, and are part of the standard algorithm used at AMSD. The assumption is that the background represents noise, and the presence of a real signal is indicated by a significant deviation from the local distribution. This is discussed in detail in a report by Rockwell International (1980), the full reference for which is given in 4.0

In our analysis, the first approximation to the background level, b_m, for the m^{th} direction was calculated by:

$$(b_m)_{approx} = \frac{\sum_{m-4}^{m-1} P_m + \sum_{m+1}^{m+4} P_m}{8} \quad . \qquad (71)$$

Having calculated $(b_m)_{approx}$ as a function of m, a list was made of P_m values which were more than 5 stand-

and deviations from $(b_m)_{approx}$. These were regarded as potential sources, and excluded from the final calculation of b_m.

The final step was to calculate the ratio

$$R_m = \frac{P_m - b_m}{\sigma_m} \tag{72}$$

where σ_m is the local standard deviation given in this instance by:

$$\sigma_m^2 = \frac{\sum_{m-4}^{m-1} (P_m - b_m)^2 + \sum_{m+1}^{m+4} (P_m - b_m)^2}{7} \quad . \tag{73}$$

The RTR was then displayed as a greyscale plot of R_m as a function of m (horizontally), and as a function of time (vertically).

2) Frequency-Azimuth Plot. To obtain FRAZ plots, the first step was to form temporal Fourier transforms of each hydrophone signal, perform cross correlations to form the visibilities (the set of which constitutes the spatial covariance matrix), and then to do the 1-dimensional imaging at each frequency. In the present case, the Fourier transforms were calculated from sets of 512 time samples, and the products averaged over various integration times to form the visibilities according to (35).

At a given frequency the power in the m-th beam is then given by:

$$P_m = \sum_{\ell} V(u_\ell)\, e^{2\pi i u_\ell \sin\theta_m}\, \delta u \tag{74}$$

where $V(u_\ell)$ represents the visibility function sampled on a regular grid at intervals of δu, such that

$$\delta u = \frac{1}{\sin \theta_{max} - \sin \theta_{min}} \qquad (75)$$

in which $(\theta_{min}, \theta_{max})$ represents the azimuth range to be imaged. For the present calculations, $V(u_\ell)$ was estimated from the measured $\{V_{ij}\}$ by simple box convolution in the spatial frequency domain.

FRAZ plots were generated for the range -1 to +1 in sin(azimuth), and 100 to 1000 Hz in frequency. Two final processing steps designed to aid in the display of the plot are to subtract a uniform background (in the same way as for the BTR), and to renormalize individual scans according to the rms of neighboring scans. The latter step is designed to overcome the problem whereby the high frequencies have a much lower amplitude, and hence are difficult to display on a plot with limited dynamic range.

As mentioned above, the BQQ-9 array is spatially undersampled at the frequencies we are using, i.e. the element spacing is greater than a half wavelength. The effect of this undersampling is to introduce a regular pattern of grating lobes, which become more and more closely spaced as the frequency increases. This will, of course, introduce artifacts into the FRAZ plot. On the plus side, however, the spatial resolution increases with increasing frequency, and this property causes the main lobe of the beam to decrease with frequency. In order to demonstrate these effects, a FRAZ was generated using synthetic data corresponding to a point source of white noise at an azimuth of zero. The results are shown in Fig. 7, which shows the characteristic pattern of frequency-dependent grating lobes and the narrowing of the main lobe with increasing frequency. This figure will aid in the interpretation of FRAZ plots made with real data.

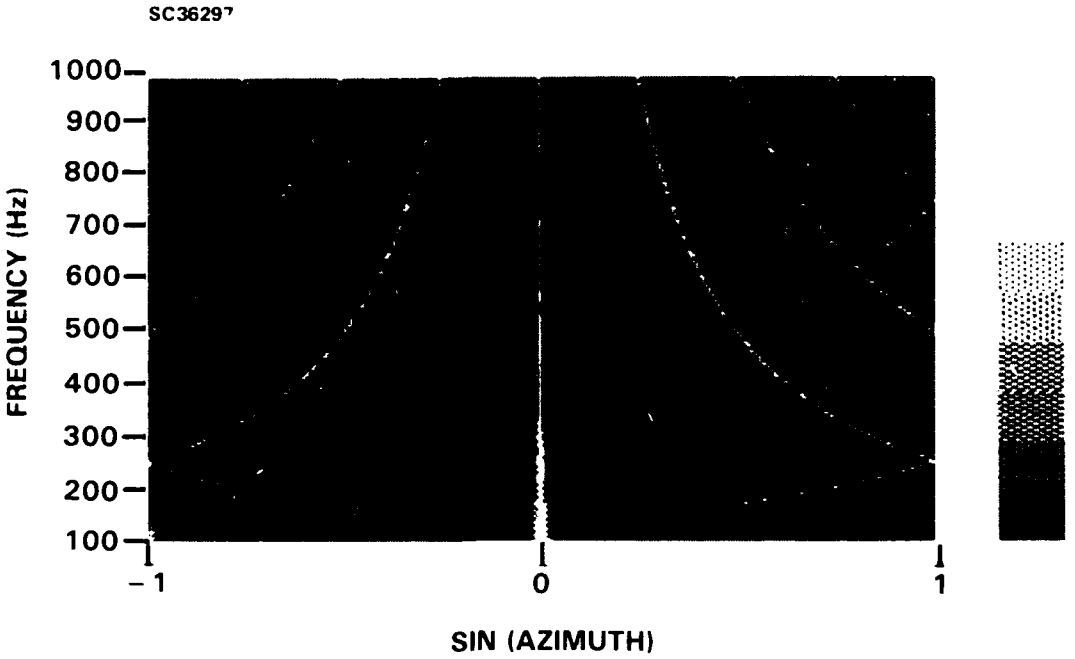


Fig. 7 FRAZ (frequency vs azimuth) plot for simulated point source of white noise.

3.4.4 Results of Preliminary Analysis

Bearing-time records were constructed for all of the data using an integration time of 40 s. It was found, however, that a large increase occurred in the amplitude of the data after about 24 min from the beginning, and this was accompanied by a certain amount of erratic behavior in amplitude. Apparently there were some equipment malfunctions. For this reason, the analysis was restricted to the first 24 min of data which appeared to quite clean. A BTR for these data is shown in Fig. 8. Several sources are apparent, some of which exhibit apparent motion. The discontinuity which occurred after 12 min had elapsed is presumably due to a time gap in the data. The source near the far left of the plot, corresponding to sin(azimuth) near -1, is presumably due to own-ship noise, and is displaced from the direction corresponding to exactly

SC36305 SC5427.FR

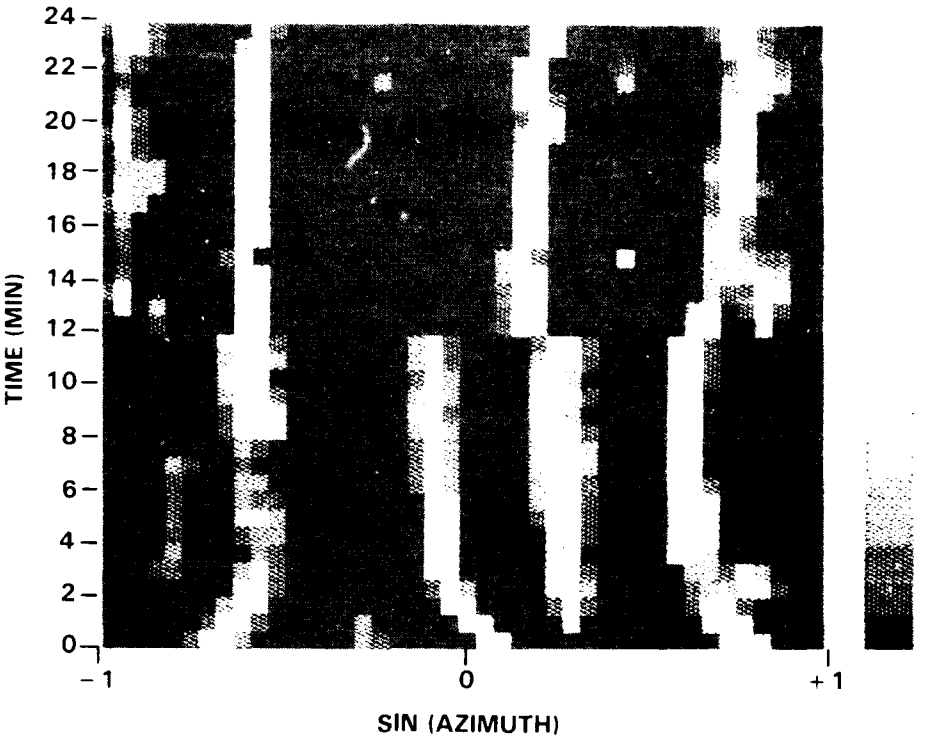


Fig. 8 BTR (bearing-time record) made using broadband data.

-1 because of the inclination of the towed array with respect to the horizontal.

The time interval between 18 and 21 minutes was selected for further analysis. From the BTR, the sources appeared to be sufficiently stable during this time period that it was decided to generate a FRAZ on the basis of a 3-minute time average. The result is shown in Fig. 9. On this figure, vertical streaks correspond to real sources, while the curved streaks are grating lobes of the same type as in Fig. 7. The positions of the real sources (most prominent at the low frequencies) can be identified with those on the broadband BTR plot of Fig. 8, during the time interval specified above.

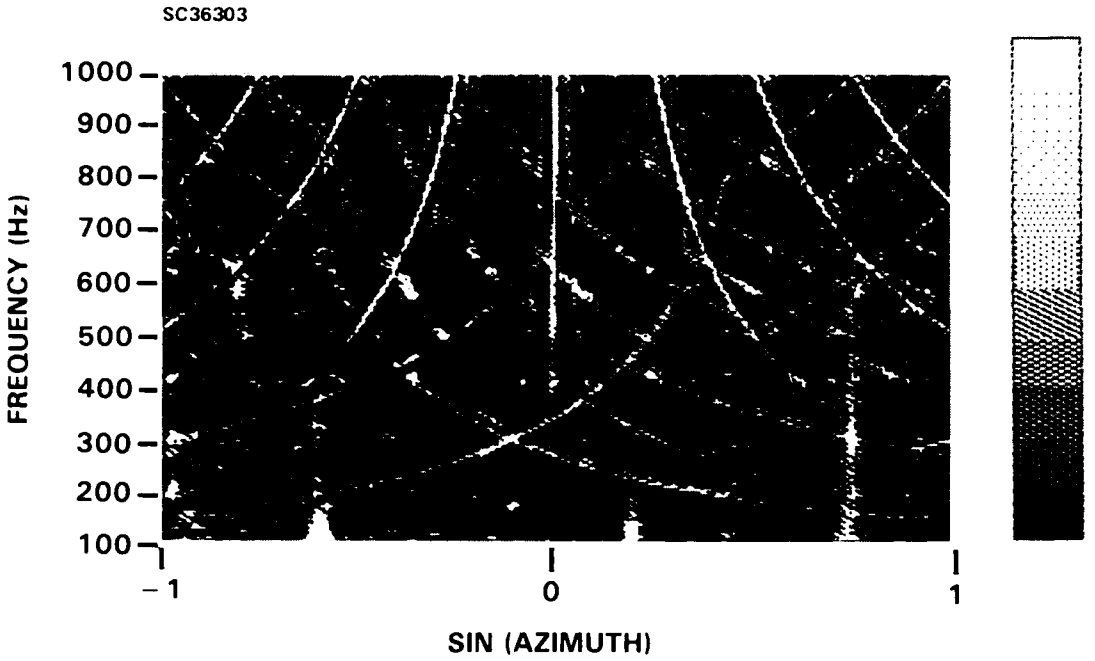


Fig. 9 FRAZ plot for a 3-min time average of data.

Besides the pattern of grating lobes, whose origin is well understood, two additional spurious effects are present in Fig. 9:

a) The vertical streak at precisely zero azimuth, and its corresponding pattern of grating lobes.

b) Horizontally arranged row of blobs at 400 Hz (there are others also, but the one at 400 Hz is by far the most intense, even though this is difficult to discern from the limited dynamic range of Fig. 9).

Before examining possible causes of interference in the signals themselves, it was decided first to check the possibility that these effects were an artifact of the algorithm used to generate the FRAZ plots. To do this, we generated a FRAZ using synthetic hydrophone data corresponding to Gaussian uncorrelated noise. The

result is shown in Fig. 10. This figure shows the expected random pattern, free from the above effects. We therefore conclude that the effects (a) and (b) are due to contamination of the hydrophone signals themselves, and not an artifact of the algorithm.

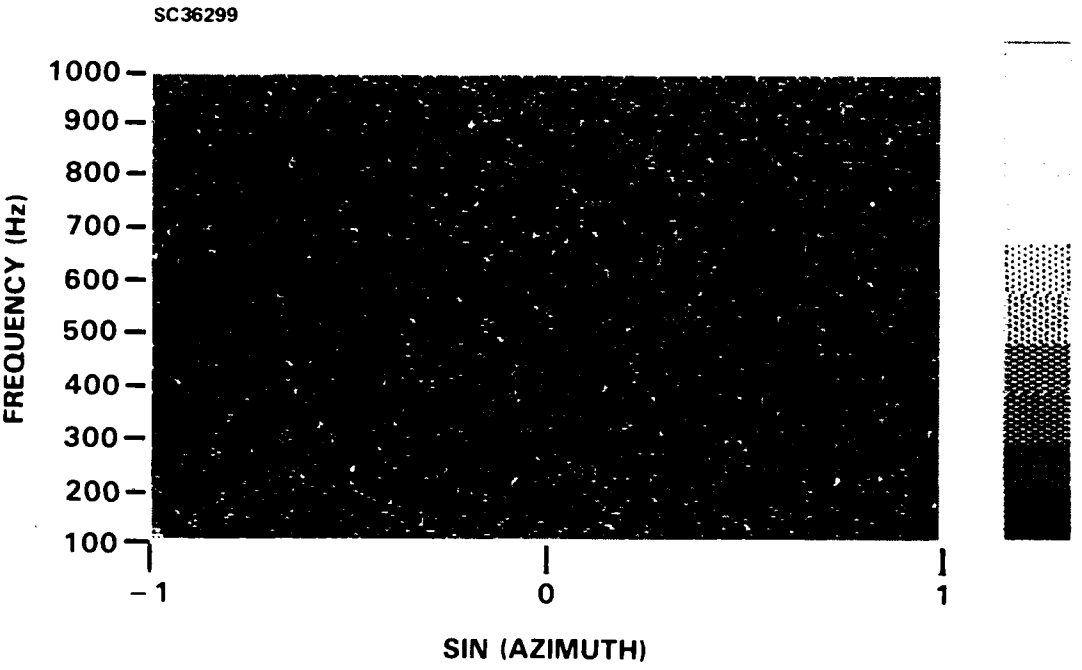


Fig. 10 FRAZ plot generated from synthetic hydrophone data consisting of uncorrelated white noise.

With regard to (a), the observed "central streak" is symptomatic of broadband correlated noise at zero delay, with respect to all of the hydrophones. It could not easily be explained as an acoustic effect since the interfering signal would have to be received by all hydrophones simultaneously. Unless the interfering source really were present in the far field at zero azimuth, this would require a disturbance propagating at about two orders of magnitude faster than the sound speed in water. This would rule out explanations involving acoustic disturbances in the towing line. It would thus seem to be due to some type of electrical effect.

With regard to (b), this is obviously a narrow band effect. There are two distinct sets of blobs present at 400 Hz, one set being coincident with grating lobes of the "central streak", and the other set being coincident with the set of grating lobes from a source at the far left of the plot, corresponding to the direction of the towing ship. The frequency of 400 Hz actually turns out to be consistent with one of the line frequencies for electrical equipment on the ship.

The above results therefore suggest that noise is getting into the hydrophones in two ways:

(i) Vibration of the towing ship at 400 Hz is being picked up acoustically by the hydrophones.

(ii) Electrical noise, both broadband and 400 Hz, is somehow getting into the data recording system.

3.4.5 Application of the Phase Closure Technique

The selected 3-minute portion of data was used to solve for the systematic time delay errors τ_i associated with each hydrophone by the procedure discussed in Section 3.3.

Because of the spurious effects discussed above, however, it was necessary to take some precautions in this procedure. Since the contamination becomes severe only at high frequencies, it was decided to restrict the range of data used in the solution to an upper frequency of 300 Hz. This, of course, avoids contamination by the 400 Hz line. The low frequencies were excluded since they are not particularly sensitive to time delay errors. The frequency interval selected for the solution was therefore 200-300 Hz.

The bandwidth effect (discussed in Section 3.3) was used to advantage, by shifting the phase center of the visibilities to the angular position corresponding to the prominent source at

sin(azimuth) = 0.75 (see Fig. 9), and averaging the visibilities over the frequency range 200-300 Hz. From (60), this reduces the contamination from the "central streak" by a factor of 0.05. The contribution from the real source, however, will be unaffected by this procedure. The results of the time delay calculations are shown in Fig. 11, which consists of a plot of time delay versus hydrophone number. A measure of the uncertainty in these results can be obtained from the a posteriori covariance matrix of the time delay vector. The interpretation is, however, complicated by the fact that it is highly nondiagonal, a consequence of the form of the measurement model (61), in which the measured phases are used to infer the difference in time delay between pairs of hydrophones rather than the absolute time delays themselves. A meaningful measure of uncertainty is therefore provided by the a posteriori standard deviation of time delay differences between neighboring hydro-

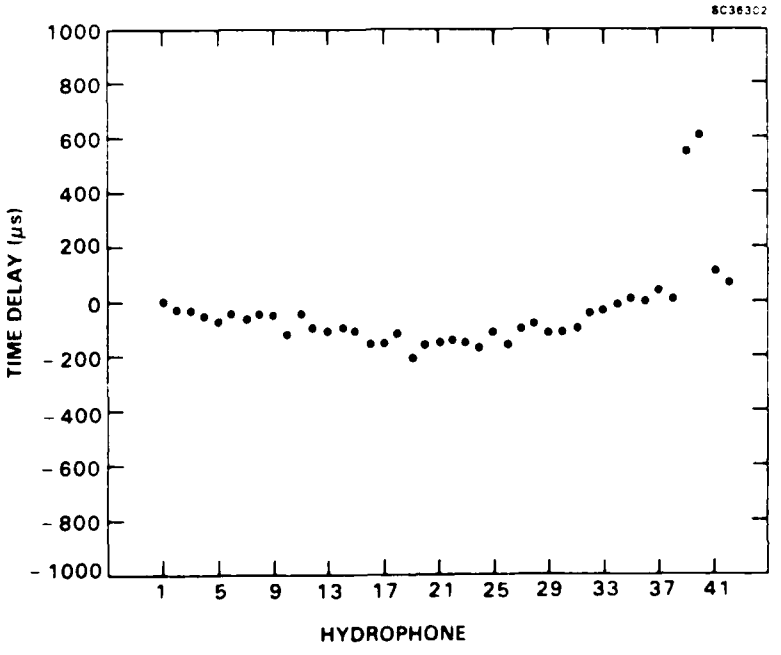


Fig. 11 Plot of experimentally obtained time delay as a function of hydrophone location for the BQQ-9 array.

phones, which turns out to be approximately 40 μs in all cases. With this in mind, the statistically significant properties of Fig. 11 are:

a) The gradual curvature of the time delay profile
b) The large deviations in the case of hydrophones 39 and 40.

In addition, there are marginally significant (2 sigma) deviations from the smooth profile in a few cases. The gradual variation of time delay as a function of hydrophone location is probably due to either hydrophone drift or propagation anomalies, or a combination of both. The large discontinuities observed for hydrophones 39 and 40, however, are unlikely to be due to either of these effects, and we attribute them to an instrumental problem.

The time delays were also estimated using the refinement discussed in Section 3.1.1, namely the "cross-correlated subarrays" algorithm. It will be recalled that this technique was proposed in order to restrict the overall angular response of the array to the region in which phase closure was valid. For the present analysis, the array was divided into 21 groups of 2 elements, and also 14 groups of 3 elements. The results in both cases were consistent with Fig. 11, in that they showed the gradual trend characterized by zero delay at each end of the array and approximately -200 μs in the middle. The effect of the large discontinuity in the time delay profile caused by hydrophones 39 and 40 was also evident, although somewhat diluted by the averaging of the hydrophone signals within each subgroup. The maximum positive delays were 440 μs and 300 μs for the 2-element and 3-element subarrays, respectively, as compared to 600 μs when no division into subarrays was used. The consistency between these results indicates that standard phase closure was satisfactory for these particular data, and that cross-correlated subarrays were not required. This may at first seem surprising,

since the time delay errors, whether due to ocean inhomogeneities or hydrophone drift, must be direction-dependent. This would mean that some reduction in angular response would be required in order that Eq. (57) be a valid error model. A likely explanation is that the necessary reduction in angular response has already been achieved by the bandwidth effect as expressed by Eq. (60). We would not expect this always to be the case. For example, if signal-to-noise considerations dictated the use of a narrower bandwidth, or else if the phase distortion due to the medium were more severe, then cross-correlated subarrays would be required.

Having estimated the time delays, the next step was to use them to apply phase corrections to the data, and generate another FRAZ plot. This was carried out, and the results are shown in Fig. 12, which is to be compared with Fig. 9, which it will be recalled was made without any phase corrections. Unfortunately, Figs. 9 and 12 are so dominated by the frequency-dependent grating lobes discussed earlier that it is difficult to see whether phase closure has made any improvement to the detectability of sources. In order to examine this question, the plots were examined more closely in the vicinity of the known real sources, particularly the source at sin(azimuth) = 0.75, since the latter source extends to high frequencies.

Figure 13 shows a comparison of the FRAZ plots in the vicinity of this source, before and after phase closure, in the frequency range 600-1000 Hz, and the sin(azimuth) range 0.5 to 1.0. It can be seen that the vertical stripe indicating the presence of the source is more prominent in the plot made with phase closure. In order to better estimate the improvement obtained, a pair of bearing records (with and without phase closure) were computed using data averaged over the frequency range 800-900 Hz. The results are presented in graphical form in Fig. 14. The source of interest is indicated by the arrow. The other sources on this plot are due to grating lobes, as can be judged by comparison with the FRAZ plots.

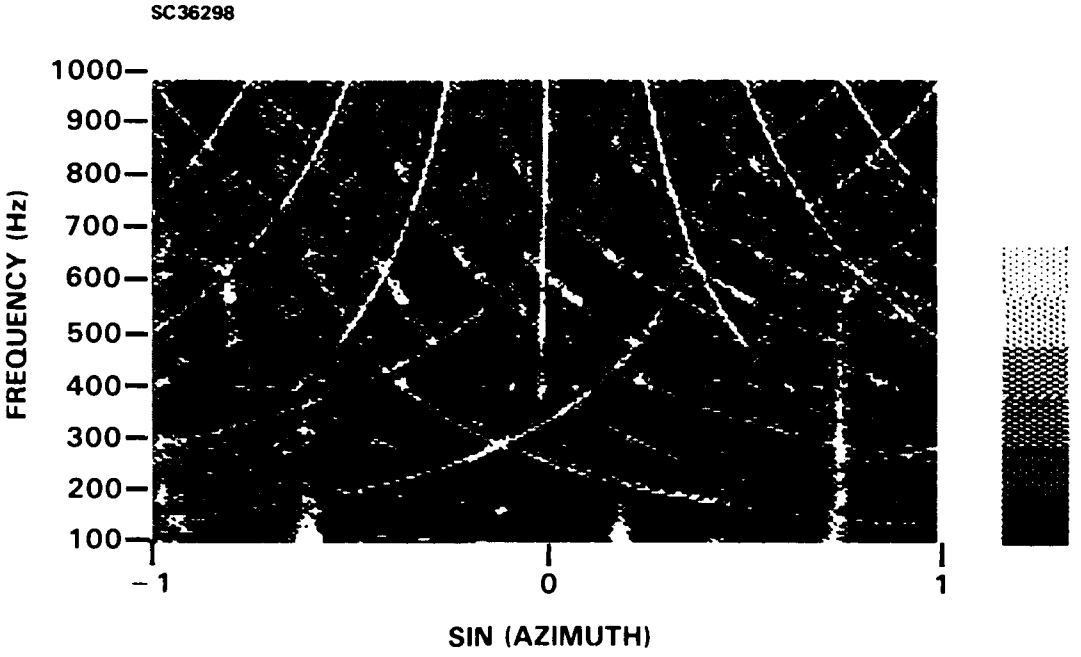


Fig. 12 FRAZ plot for using the same data as for Fig. 7, except with phase corrections applied according to the experimentally determined time delays.

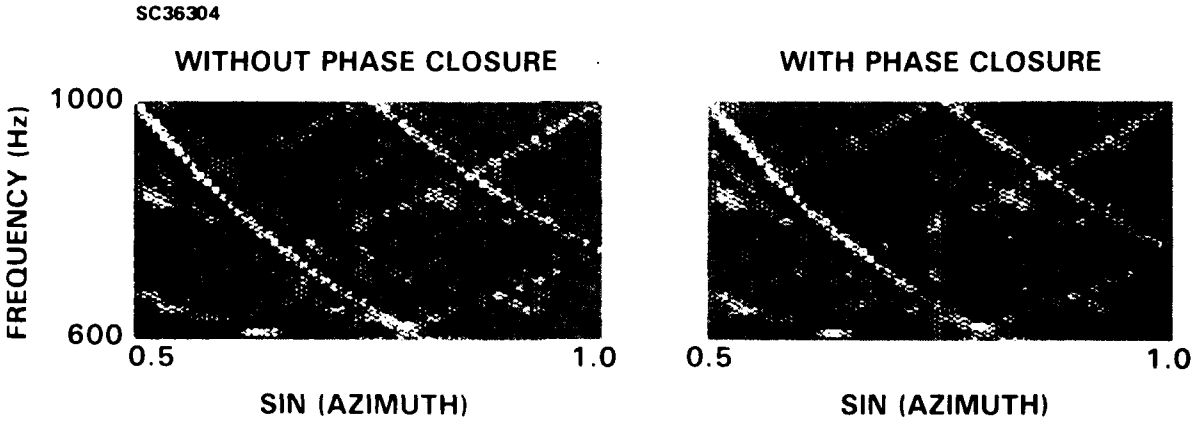


Fig. 13 Comparison of FRAZ plots with and without phase corrections. These plots represent portions of Figs. 7 and 10, in the sin(azimuth) range 0.5 to 1.0 and frequency range 600-1000 Hz. The vertical streak down the center of each plot represents the source of interest.

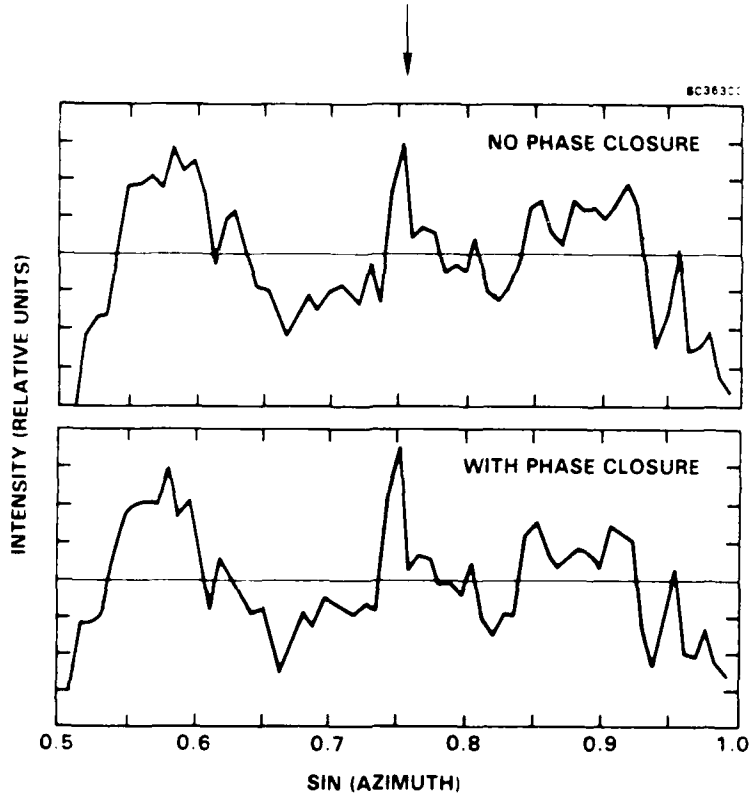


Fig. 14 Bearing records calculated from data averaged over the frequency range 800-900 Hz, without phase corrections (upper plot) and with phase corrections (lower plot). The source of interest is indicated by an arrow.

It can be seen that the source under consideration has been boosted in amplitude by the phase closure technique, the improvement being a factor of about 25%. It is interesting to compare this figure with what one would have expected from the magnitudes of the time delays calculated above. The r.m.s. time delay for all of the hydrophones was 155 μs, corresponding to a phase error of 47° at a frequency of 850 Hz. If the phases were completely random, the amplitude of a point source would be reduced by a factor of cos 47° = 0.68, and hence application of the phase closure technique should produce an improvement of 32%. Although this calculation was crude, it does suggest that the improvement obtained was consistent with that expected.

3.4.6 Implications of These Results

If the observed behavior of time delay as a function of hydrophone location for the BQQ-9 array is typical of towed arrays, then we can use the results of the previous section to make some predictions about the effect of phase closure on longer arrays. In order to do this, we fitted the observed profile of time delay versus hydrophone location (Fig. 11) with a parabola and made the assumption that the phase curvature of this parabola (2nd derivative of time delay as a function of position) is typical of towed arrays. On this basis, the profile of time delay $\tau(\mu s)$ as a function of hydrophone location h for an array of length L was taken to be:

$$\tau = 924\ h(h-L)/L_o^2 \qquad (76)$$

where L_o is the length of the BQQ-9 array.

Synthetic data for a point source at a frequency of 750 Hz were generated both with and without the effect of these phase errors, and used to make 1-d images (bearing records) for a series of array lengths. The results are shown in Fig. 15. It can be seen that for an array of length L_o, there is only a small degradation of source amplitude. For an array of length 1.2 L_o it is somewhat larger, and for arrays of length 1.6 L_o and $2L_o$ it is substantial. In the case of an array of length $2L_o$, the degradation is a factor of 4 (6 dB). In addition, the spatial resolution has been degraded severely. It would be unreliable to continue the extrapolation beyond a factor of 2 in array length without further data, but these results do indicate that for longer arrays, the use of phase closure should produce a rather dramatic improvement in source detectability.

These results show that the phase closure technique will make it possible to maintain phase coherence over longer arrays than was previously possible, and that source detectability and locatability will be substantially improved. An additional consideration concerns the use of superresolution techniques. In principle, such

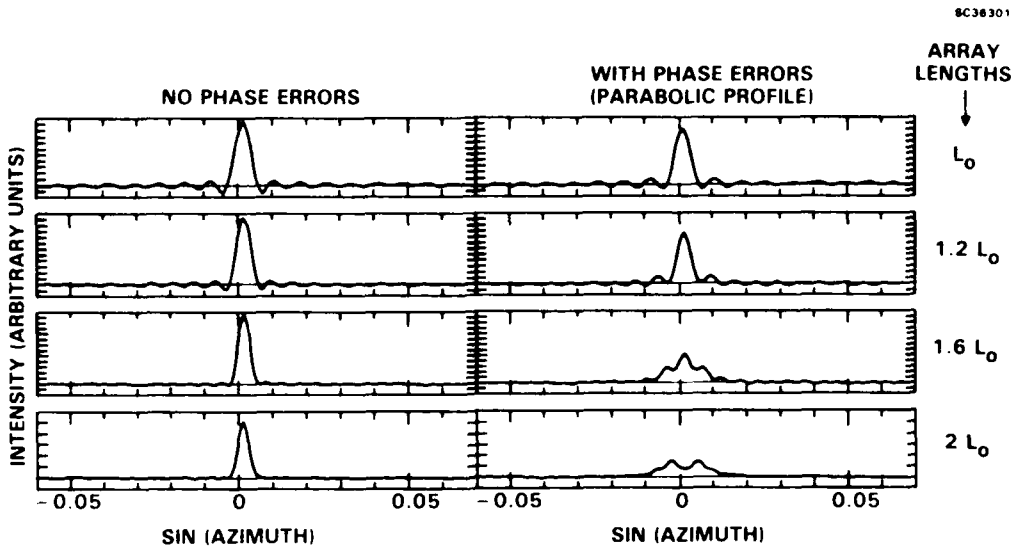


Fig. 15 Synthetically determined bearing records for a point source at 750 Hz for arrays of various lengths, whereby L_o represents the length of the BQQ-9 array. The plots on the left are for data with no phase errors, while the plots on the right were made assuming a parabolic phase error profile with the same phase curvature as measured for the BQQ-9 array.

techniques can increase the accuracy of bearing estimation to beyond the diffraction limit of the system (see for example Gabriel 1981, Johnson 1982, Byrne and Fitzgerald 1983, Byrne, Fitzgerald and Steele 1984). In practice, however, these techniques are unstable to even small phase errors of ±5° (Byrne and Steele, 1985). Phase closure would thus appear to be an important ingredient for the practical application of any superresolution technique.

The next phase of this research should:

1) Combine the phase closure technique with superresolution techniques and assess the improvement obtained in the face of various degrees of phase errors; and

2) Extend phase closure concepts to the problem of multipathing, particularly with regard to source location in the channel under Arctic ice.

4.0 REFERENCES

Bartelt, H., Lohmann, A.W., and Wirnitzer, B. 1984, Applied Optics 23, 3121.

Byrne, C.L. and Fitzgerald, R.M. 1983, J. Acoust. Soc. Am. 74, 1224.

Byrne, C.L., Fitzgerald, R.M., and Steel, A.K. 1984, "On the Role of Prior Information in Nonlinear Bearing Estimation," in Proc. NATO Advanced Study Institute on Adaptive Methods in Underwater Acoustics (Luneburg, Germany).

Byrne, C.L. and Steele A.K. 1985, IEEE Oceanic Engineering (in press).

Chernov, L.A. 1960, "Wave Propagation in a Random Medium," Eng. transl. R.A. Silverman (New York, McGraw-Hill).

Devaney, A.J. 1979, J. Math. Phys. 20, 1687.

Flatte, S.M., Dashen, R., Munk, W.H., Watson, K.M., and Zachariasen, F. 1979, "Sound Transmission Through a Fluctuating Ocean," (New York, Cambridge Univ. Press).

Gabriel, W.F. 1981, Proc. ASSP Workshop on Spectral Estimation, pp 5.1.1.-5.1.7 (McMaster Univ., Hamilton, Ont., Canada).

Hogbom, J.A. 1974, Astron. Astrophys. Suppl 15, 417.

Johnson, D.H. 1982, Proceedings of the IEEE 70, 1018.

Marsh, K.A., Richardson, J.M., and Martin, J.M. 1985, "Application of the Phase Closure Technique to Passive Acoustic Imaging Through Inhomogeneous Media," Acoustical Imaging 14, 133.

Paulraj, A. and Kailath, T. 1985, "Direction of Arrival Estimation by Eigenstructure Methods with Unknown Sensor Gain and Phase," to appear in IEEE ASSP.

Pearson, T.J. and Readhead, A.C.S. 1984, Ann. Rev. Astron. Astrophys. 22, 97.

Readhead, A.C.S. and Wilkinson, P. 1978, Astrophy. J. 223, 25.

Rockwell International 1980, "Automatic Acoustic Detection Algorithm Development," Document No. C80-395.1/301.

END

DTIC

8-86