Theoretical and Computational Acoustics
D. Lee, Y-H Pao, M.H. Schultz, and Y-C Teng (Editors)


# Range-Dependent Propagation: A State-Space Approach

E.J. Sullivan
Code 82101, Naval Undersea Warfare Center, Newport, RI 02841

G.W. Hedstrom
University of California, Lawrence Livermore National Laboratory
P.O. Box 808, L-321, Livermore, CA 94550

J.V. Candy
University of California, Lawrence Livermore National Laboratory
P.O. Box 808, L-495, Livermore, CA 94550

## Abstract

The use of a state-space representation for propagation models has several advantages. First, it provides a recursive framework whereby the solution of the equations can be stepped through range and depth, thus providing the so-called innovations sequence that allows the performance of the model to be monitored on-line. Second, it allows the introduction of a Gauss-Markov model, which permits both measurement and dynamic model noise structures to be included in the modeling scheme. Third, it permits the inclusion of Kalman filter structures, thereby providing estimators for the model parameters to be introduced in a self-consistent manner. It has been shown that this approach can be highly effective for the normal mode propagation model. In this work we apply these ideas to the range-dependent parabolic equation, develop the corresponding state space forms, and construct a model-based processor for the estimation of the range-dependent sound speed profile.

## INTRODUCTION

Recent work in model-based processing (MBP).[1-5] has shown that it has many useful applications in the field of ocean acoustic signal processing. Its effectiveness lies in the fact that MBP casts the problem into a state-space representation. This has several advantages. First, it provides a recursive framework whereby the solution of the equations can be stepped through range and depth, thus providing the so-called innovations sequence that allows the performance of the model to be monitored on-line. Second, it allows the introduction of a Gauss-Markov model, which permits both measurement and dynamic model noise structures to be included in the modeling scheme. Third, it permits the inclusion of Kalman filter structures, thereby providing estimators for the model parameters to be introduced in a self-consistent manner.

There are basically two sets of equations in the MBP representation: the state equations themselves, and the measurement equations, which relate the states to the actual measured quantities. In the stochastic case, a Gauss-Markov representation evolves, which allows the inclusion of both the measurement noise and the state or "process" noise as second-order statistical models. Once this framework is established, it is possible to investigate ocean acoustic signal processing problems in a very general way, since full advantage can be taken of existing systems theory (e.g., observability, identifiability) and techniques (Kalman state estimators) to characterize a recursive processor which constitutes a minimum-variance, adaptive estimator.[6] That is, the so-called innovation sequence, which is automatically generated by the Kalman filter, allows the performance of the model to be monitored, while at the same time Kalman estimators can be utilized on-line to continuously update model parameters, such as the sound speed profile, thus making great inroads into the so-called *mismatch problem.*[7,8]

Recent studies[1,2,4] have shown this approach to be highly effective for the normal mode propagation model. In this work we apply these ideas to the range-dependent parabolic equation and develop the corresponding state space forms. In the next section we present an overview of the state space approach and in Section 2 a recursive form of a finite-difference solution to the parabolic equation (PE) based on the Crank-Nicholson method[9] will be developed. Section 3 describes a model-based processor for the parabolic equation and Section 4 develops an estimation scheme for the sound speed profile based on the development in Section 3. Finally there is a short discussion contained in Section 5.

## I. STATE-SPACE APPROACH

The state-space form of a differential equation is basically a method of reducing an *nth* order equation to $n$ first order equations. As an example, consider the following ordinary second-order differential equation.

$$\frac{d^2U}{du^2} + \alpha^2 U = 0 \tag{1}$$

Defining the state vector as $\mathbf{U} = [U \quad \dot{U}]^T$, the state space representation of Equation 1 is

$$\frac{d}{du}\begin{bmatrix} U \\ \dot{U} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\alpha^2 & 0 \end{bmatrix}\begin{bmatrix} U \\ \dot{U} \end{bmatrix} \tag{2}$$

or, with the definition $\mathbf{G} = \begin{bmatrix} 0 & 1 \\ -\alpha^2 & 0 \end{bmatrix}$,

$$\frac{d}{du}\mathbf{U} = \mathbf{A}\mathbf{U} \tag{3}$$

The state can be connected to the vector of observations $\mathbf{y}$ (measurements) by the so-called measurement equation, viz.,

$$\mathbf{y} = \mathbf{C}[\mathbf{U}] \tag{4}$$

Once we have this structure in hand we are in a position to generate a Kalman filter structure wherein we can provide a recursive processor for the estimates of the state $\mathbf{U}$. This consists of the following set of equations. The first is the recursion relation for the state estimates

$$\hat{\mathbf{U}}(p+1|p) = (\mathbf{G}+\mathbf{I})\hat{\mathbf{U}}(p|p) \tag{5}$$

Here the notation signifies that this is the predicted estimate at (space or time) index $p+1$ based on all the data up to index $p$ and the hat signifies estimate. This is then compared to the measurement at index $p+1$, resulting in the *innovation* sequence, i.e.,

$$\mathbf{e}(p) = \mathbf{y}(p+1) - \hat{\mathbf{y}}(p+1|p) = \mathbf{y}(p+1) - \mathbf{C}[\hat{\mathbf{U}}(p+1|p) \tag{6}$$

The next step is the computation of the *Kalman gain* given by

$$\mathbf{K}(p) = \hat{\mathbf{P}}(p+1|p)\mathbf{C}^T\mathbf{R}_e^{-1}(p) \tag{7}$$

$\mathbf{P}(p)$ and $\mathbf{R}_e(p)$ are the state covariance and innovations covariance estimations, respectively.
The state estimate $\hat{\mathbf{U}}(p+1|p)$ is now corrected providing the *filtered* state estimate.

$$\hat{\mathbf{U}}(p+1|p+1) = \hat{\mathbf{U}}(p+1|p) + \mathbf{K}(p)\mathbf{e}(p) \tag{8}$$

## II. STATE-SPACE FORM FOR THE PARABOLIC EQUATION

We begin with the standard form of the PE which is given by

$$2ik_0\frac{\partial\psi}{\partial r} + \frac{\partial^2\psi}{\partial z^2} + k_0^2(n^2-1)\psi = 0 \tag{9}$$

where the solution for the total pressure field is given by

$$p(r,z) = \psi(r,z)H_0^{(1)}(k_0 r) \tag{10}$$

Here, $k_0 = \omega/c_0$, $\omega$ is the radian frequency, $c_0$ is a reference sound speed, $H_0^{(1)}(k_0 r)$ is the Hankel function of the first kind, and $n = c_0/c$ is the acoustic index of refraction. That is, $c = c(r,z)$ is the actual sound speed. In practice, the Hankel function is usually replaced by its large argument approximation given by

$$H_0^{(1)}(k_0 r) \approx \left(\frac{2}{\pi k_o r}\right)^{1/2} e^{i(k_0 - \frac{\pi}{4})}; \qquad k_0 r >> 1 \tag{11}$$

A comprehensive derivation of equation 9 can be found in Reference 10.

A finite difference approximation for the first derivative in $r$ gives

$$2ik_0 \frac{\partial \psi}{\partial r} \approx 2ik_0 \left[\frac{\psi_{p+1,q} - \psi_{p,q}}{\Delta r}\right] \tag{12}$$

where $p$ is the range index and $q$ is the depth index. Thus, $\Delta r = r_{p+1} - r_p$. We next construct the finite difference approximation for the second derivative in $z$. This results in

$$\frac{\partial^2 \psi}{\partial z^2} \approx \left[\frac{\psi_{p,q+1} - 2\psi_{p,q} + \psi_{p,q-1}}{(\Delta z)^2}\right] \tag{13}$$

Here, $\Delta z = z_{q+1} - z_q$. This is referred to as the explicit finite-difference method (EFDM) where the *implicit* finite difference method (IFDM) is the average of the EFDM over one step in $r$. That is, the IFDM is given by

$$\frac{\partial^2 \psi}{\partial z^2} \approx \frac{1}{2}\left[\frac{(\psi_{p,q+1} - 2\psi_{p,q} + \psi_{p,q-1}) + (\psi_{p+1,q+1} - 2\psi_{p+1,q} + \psi_{p+1,q-1})}{(\Delta z)^2}\right] \tag{14}$$

The PE will now be written in terms of both the EFDM and IFDM. Beginning with the EFDM, we find that

$$\frac{2ik_0}{\Delta r}\left[\psi_{p+1,q} - \psi_{p,q}\right] = -\frac{1}{(\Delta z)^2}\left[\psi_{p,q+1} - 2\psi_{p,q} + \psi_{p,q-1}\right] - k_0^2(n^2 - 1)\psi_{p,q} \tag{15}$$

It will be convenient to define

$$\nu_{p,q} = n_{p,q}^2 - 1 \tag{16}$$

where we have shown the dependence of $n$ on $p$ and $q$ for clarity. Equation 15 now becomes

$$\frac{2ik_0}{\Delta r}\left[\psi_{p+1,q} - \psi_{p,q}\right] = -\frac{1}{(\Delta z)^2}\left[\psi_{p,q+1} - 2\psi_{p,q} + \psi_{p,q-1}\right] - k_0^2 \nu_{p,q}\psi_{p,q} \tag{17}$$

We now define the (range and depth independent) parameter $\beta$ as

$$\beta = \frac{\Delta r}{(\Delta z)^2 2ik_0} \tag{18}$$

and solve Equation 17 explicitly for $\psi_{p+1,q}$. This results in

$$\psi_{p+1,q} = \psi_{p,q} - \beta\left[\psi_{p,q+1} - 2\psi_{p,q} + \psi_{p,q-1}\right] - \beta(\Delta z)^2 k_0 \nu_{p,q} \tag{19}$$

Here, we are assuming a constant depth, range-independent bottom with an absorbing layer. The boundary conditions are that $p = 0$ on the top and bottom (see Fig. 1). Equation 19 can be written in the following tri-diagonal matrix form.

$$\begin{bmatrix} \psi_{p+1,1} \\ \psi_{p+1,2} \\ \vdots \\ \psi_{p+1,N-1} \\ \psi_{p+1,N} \end{bmatrix} = \begin{bmatrix} \phi_{p,1} & -\beta & & & \\ -\beta & \phi_{p,2} & -\beta & & \\ & \ddots & \ddots & \ddots & \\ & & -\beta & \phi_{p,N-1} & -\beta \\ & & & -\beta & \phi_{p,N} \end{bmatrix} \begin{bmatrix} \psi_{p,1} \\ \psi_{p,2} \\ \vdots \\ \psi_{p.N-1} \\ \psi_{p,N} \end{bmatrix} \tag{20}$$

where $N$ is the size of the grid in depth and

$$\phi_{p,q} = 1 + 2\beta - \beta(\Delta z)^2 k_0 \nu_{p,q} \tag{21}$$

This leads to the range recursion relation for the state $\Psi$ given by

$$\Psi(p+1) = D_1 \Psi(p) \tag{22}$$

where

$$\Psi(p) = \begin{bmatrix} \psi_{p,1} \\ \psi_{p,2} \\ \vdots \\ \psi_{p,N-1} \\ \psi_{p,N} \end{bmatrix} \tag{23}$$

and

$$D_1 = \begin{bmatrix} \phi_{p,1} & -\beta & & & \\ -\beta & \phi_{p,2} & -\beta & & \\ & \ddots & \ddots & \ddots & \\ & & -\beta & \phi_{p,N-1} & -\beta \\ & & & -\beta & \phi_{p,N} \end{bmatrix} \tag{24}$$

Writing the PE in terms of the IFDM results in

$$\frac{2ik_0}{\Delta r}\left[\psi_{p+1,q} - \psi_{p,q}\right] =$$

$$-\frac{1}{2(\Delta z)^2}\left[\psi_{p,q+1} - 2\psi_{p,q} + \psi_{p,q-1} - \psi_{p+1,q+1} + 2\psi_{p+1,q} - \psi_{p+1,q-1}\right] - k_0^2 \nu_{p,q}\psi_{p,q} \quad (25)$$

The matrix form of Equation 25 is

$$B\Psi(p+1) = D_2\Psi(p) \quad (26)$$

where

$$D_2 = \begin{bmatrix} \phi'_{p,1} & -\beta/2 & & & \\ -\beta/2 & \phi'_{p,2} & -\beta/2 & & \\ & \ddots & \ddots & \ddots & \\ & & -\beta/2 & \phi'_{p,N-1} & -\beta/2 \\ & & & -\beta/2 & \phi'_{p,N} \end{bmatrix} \quad (27)$$

with

$$\phi'_{p,q} = 1 + \beta - \beta(\Delta z)^2 k_0^2 \nu_{p,q} \quad (28)$$

and

$$B = \begin{bmatrix} 1+\beta & -\beta/2 & & & \\ -\beta/2 & 1+\beta & -\beta/2 & & \\ & \ddots & \ddots & \ddots & \\ & & -\beta/2 & 1+\beta & -\beta/2 \\ & & & -\beta/2 & 1+\beta \end{bmatrix} \quad (29)$$

This leads to the range recursion

$$\Psi(p+1) = B^{-1}D_2\Psi(p) \quad (30)$$

Note that B does not depend upon $p$ or $q$.

There are more accurate versions of the PE than the one we are using but we point out that the approach we are using here is quite general in that it can be applied to all of them.

In the next section we will build a processor based on the state-space recursion relation from Equations 22 and 30 for the estimation of the range-dependent sound speed profile.

## III. A MODEL-BASED PROCESSOR

We are now in a position to construct a model-based processor for the range dependent estimation problem. As an example of the application of such a processor, we shall consider the problem of sound speed profile (SSP) estimation in a range-dependent environment.

In this example, range dependent means that the SSP is a function of range and all other parameters (bottom parameters, depth etc.) are not.

In order to implement our processor, we will need two sets of equations, the state propagation equations and the measurement equations. A Gauss-Markov representation for the state propagation equations evolves from Equation 30 as

$$\hat{\mathbf{\Psi}}(p+1) = (\mathbf{A}\Delta r + \mathbf{I})\hat{\mathbf{\Psi}}(p) + \Delta r \mathbf{w} \tag{31}$$

The actual form of $A$ used at this point would be determined by the form of the PE we are using. For our example, $A$ is either $D_1$ or $B^{-1}D_2$, depending upon whether we choose to use the EFDM or the IFDM, respectively, and $\mathbf{w}$ is the process noise. We will assume that the measurements are made by a set of vertical arrays, each containing $L$ equally spaced elements. The measurement equation then takes the form

$$\mathbf{P}(p) = C[\mathbf{\Psi}(p)] + \mathbf{v} = \mathbf{\Psi}(p)H_0^{(1)}(k_0 r_p) + \mathbf{v} \tag{32}$$

Here, $\mathbf{P}(p)$ is the $L \times 1$ vector of hydrophone measurements from a vertical array with $\mathbf{z} = [z_1, z_2, \ldots, z_L]^T$ being the vector of hydrophone coordinates and $\mathbf{v}$ is the measurement noise.

We now model the SSP at range point $p$ using empirical orthogonal eigenfunctions (EOF) as a basis set.[12] This gives

$$c(p,q) = c(r_p, z_q) = a_1(p)E_1(q) + a_2(p)E_2(q) + \ldots + a_K(p)E_K(q) \tag{33}$$

Defining

$$\alpha(p) = \begin{bmatrix} a_1(p) \\ a_2(p) \\ \vdots \\ a_K(p) \end{bmatrix} \tag{34}$$

and

$$\mathbf{E}^T(q) = [E_1(q), E_2(q), \ldots, E_K(q)] \tag{35}$$

we have

$$c(p,q) = \mathbf{E}^T(q)\alpha(p) \tag{36}$$

## IV. SOUND SPEED PROFILE ESTIMATION

In order to obtain an estimate of the range-dependent SSP, we shall assume that we have a known SSP and starting field at $r = r_p$ and a set of pressure and SSP measurements from a vertical array at $r = R = r_{p+2}$ (see Fig. 2). As a first example, we shall attempt

to estimate the SSP at $r = R/2 = r_{p+1}$. For our approach, we follow a modified Kalman filter estimation technique.

(1) We estimate the field at $r = R/2$ based on the known (starting) field and SSP at $r = r_p$ using a state equation based on the Gauss-Markov model given by Equation 31, i.e.,

$$\hat{\Psi}(p+1|p) = A[c(p,q)]\hat{\Psi}(p|p) \tag{37}$$

In Equation 37 $\hat{\Psi}(p+1|p)$ is the starting field and $c(p,q)$ is the(known) SSP at $r = r_p$.

(2) Since we do not have a measurement at $r = R/2$ we now provide a "pseudo measurement" based on the *unknown* SSP at $r = R/2$. That is,

$$\mathbf{P}_\alpha(p+1) = H_0^{(1)}(k_0 r_{p+1})A[c(\alpha)]\hat{\Psi}(p|p) \tag{38}$$

where $\alpha$ represents the unknown parameters of the SSP at $r = R/2$.

(3) It is now possible to compute a "pseudo residual" in terms of the unknown parameter vector $\alpha$. This is given by

$$\mathbf{e}_\alpha(p) = [\mathbf{P}_\alpha(p+1) - H_0^{(1)}(k_0 r_{p+1})\hat{\Psi}(p+1|p)] \tag{39}$$

(4) We now "correct" the estimate $\hat{\Psi}(p+1|p)$, yielding

$$\hat{\Psi}_\alpha(p+1|p+1) = \hat{\Psi}(p+1|p) + \mathbf{K}\mathbf{e}_\alpha(p) \tag{40}$$

where we make explicit the fact that the corrected state depends upon the unknown parameter vector $\alpha$.

(5) A new prediction is now made.

$$\hat{\Psi}_\alpha(p+2|p+1) = A[c(\alpha)]\hat{\Psi}_\alpha(p+1|p+1) \tag{41}$$

(6) The SSP parameter vector at $R/2 = r_{p+1}$ is now estimated by finding the minimum of the innovation based on the measurement at $r_{p+2} = R$. That is, the estimate of $\alpha$ is found as the solution of

$$H_0^{(1)}(k_0 r_{p+2})[\Psi(p+2) - \hat{\Psi}_\alpha(p+2|p+1)] = min \tag{42}$$

Although the Kalman gain is computed on the basis of the estimates of $\mathbf{P}$ and $\mathbf{R}_e$, it is clear from the above that we really do not have access to these quantities since we have made only two iterations. However, it can be shown that in the case of vanishing measurement noise, the expression for the Kalman gain reduces to

$$\mathbf{K} = \mathbf{C}^{+} \tag{43}$$

where $\mathbf{C}^{+}$ is the pseudoinverse, i.e.,

$$\mathbf{C}^{+} = [\mathbf{C}^{T}\mathbf{C}]^{-1}\mathbf{C}^{T} \tag{44}$$

## V. DISCUSSION

The success of this technique depends upon whether the SSP is observable in the systems theoretic sense. Here, we have made the assumption that it is. Intuitively, one would expect that if the SSP at $r = R/2$ were changed, then the field at $r = R$ would be impacted. Reversing this argument implies that the field at $r = R/2$ (and the field at $r = 0$) contains information about the SSP at $r = R/2$.

A remaining question is that of uniqueness. That is, is there more than one SSP that can produce the same field values? The answer to this is not clear. It is probably the case that there are situations where it is true and situations where it is not. A situation of non uniqueness would be an issue if the goal of the calculation were the SSP itself. However if the intent of the calculation were to be able to reproduce the fields, then the issue is moot.

What we have presented here is a technique for determining the SSP in a sound channel by remote measurements. Although there are other tomographic techniques for the remote determination of SSP, we feel that, at least for the scenario described herein, the method is quite simple and straightforward. Also, it should be added that the approach presented here is not limited to the determination of the SSP, but can be generalized to include the bottom properties as well. The success of this approach, as well as for any model-based calculation clearly depends on the validity and accuracy of the model.

## REFERENCES

[1] J. V. Candy and E. J. Sullivan, "Ocean acoustic signal processing: A model-based approach," *J. Acoust. Soc. Am.*, vol. 92, no. 6, pp 3185-3201, 1992.

[2] J. V. Candy and E. J. Sullivan, "Sound velocity estimation: A system theoretic approach," *IEEE Oceanic Engr.*, vol. 18, no 3, pp 240-252,1993.

[3] E. J. Sullivan and J. V. Candy, "Passive synthetic aperture processing as a Kalman filter problem," *J. Acoust Soc. Am.*, vol. 95, no 5, Part 2, page 2953, 1994.

[4] J. V. Candy and E. J. Sullivan,"Model-based processor design for a shallow water experiment," *J. Acoust. Soc. Am.*, vol. 95, no.4, pp 2038-2051, 1994.

[5] J. V. Candy and E. J. Sullivan, "Passive localization in ocean acoustics: A model-based approach," *J. Acoust. Soc. Am.*, vol. 98, no.3, 1995.

[6] J. V. Candy, *Signal Processing: The Model-based approach*, McGraw-Hill, New York, 1986.

[7] A. Tolstoy, *Matched field processing for Ocean Acoustics*, World Scientific Publishing Co., River Edge, NJ, 1993.

[8] E. J. Sullivan and David Middleton, "Estimation and detection issues in matched-field processing," *IEEE J. Ocean Eng*, vol. 18, no. 3, 1993.

[9] D. Lee and S. T. McDaniel, *Ocean Acoustic Propagation by Finite Difference Methods* (pergammon, New York, 1988).

[10] F. B. Jensen, et al, *Computational Ocean Acoustics*, AIP Series in Modern Acoustics and Signal Processing, AIP Press, New York, 1994.

[11] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, Academic Press, New York, 1970

[12] P. K. Kundu, J. S. Allen, and R. L. Smith, *Modal Decomposition of the Velocity field Near the Oregon Coast*, J. Phys. Oceanogr. vol 5, pp 683-704, 1975.

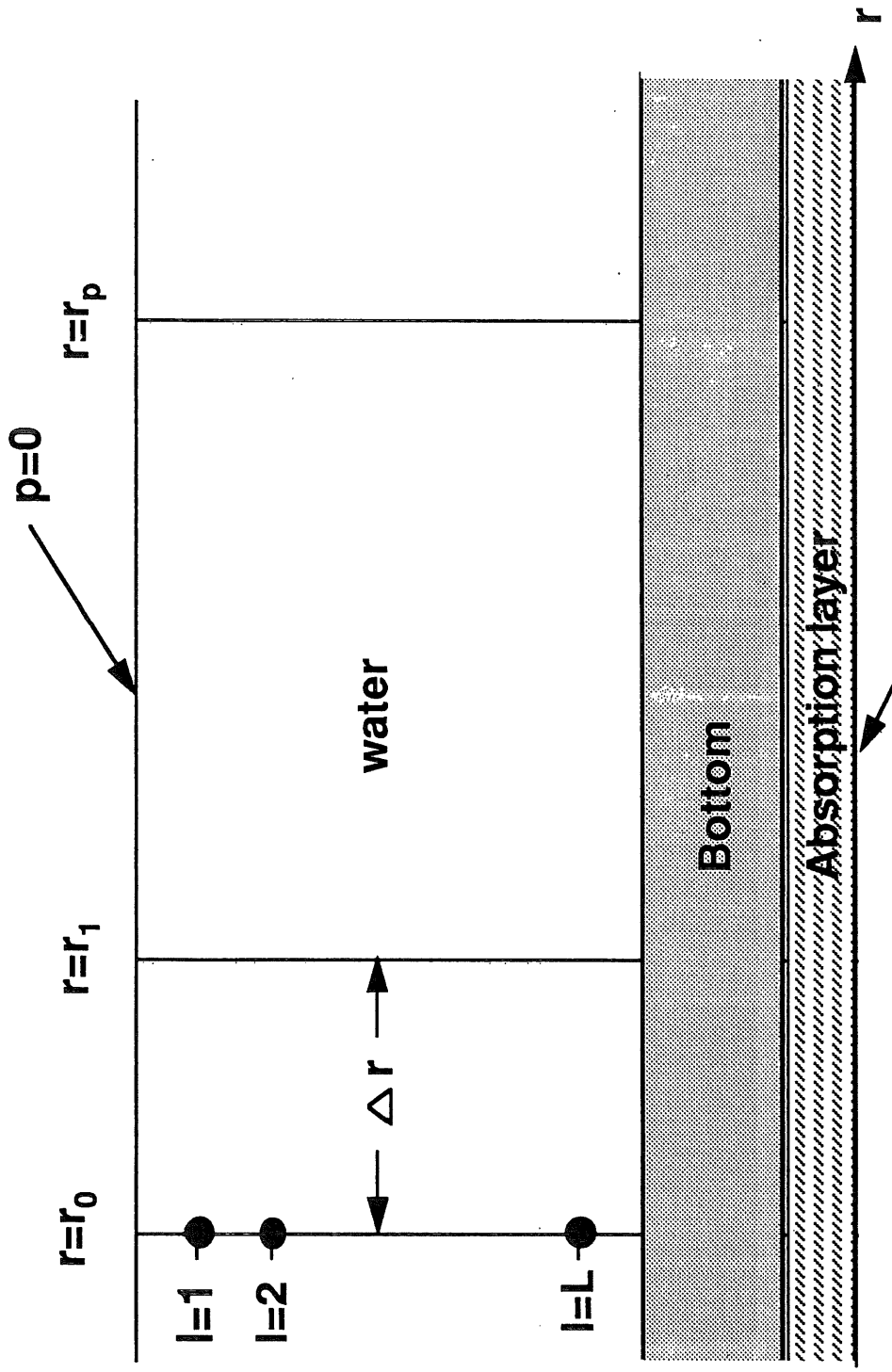



FIG. 1

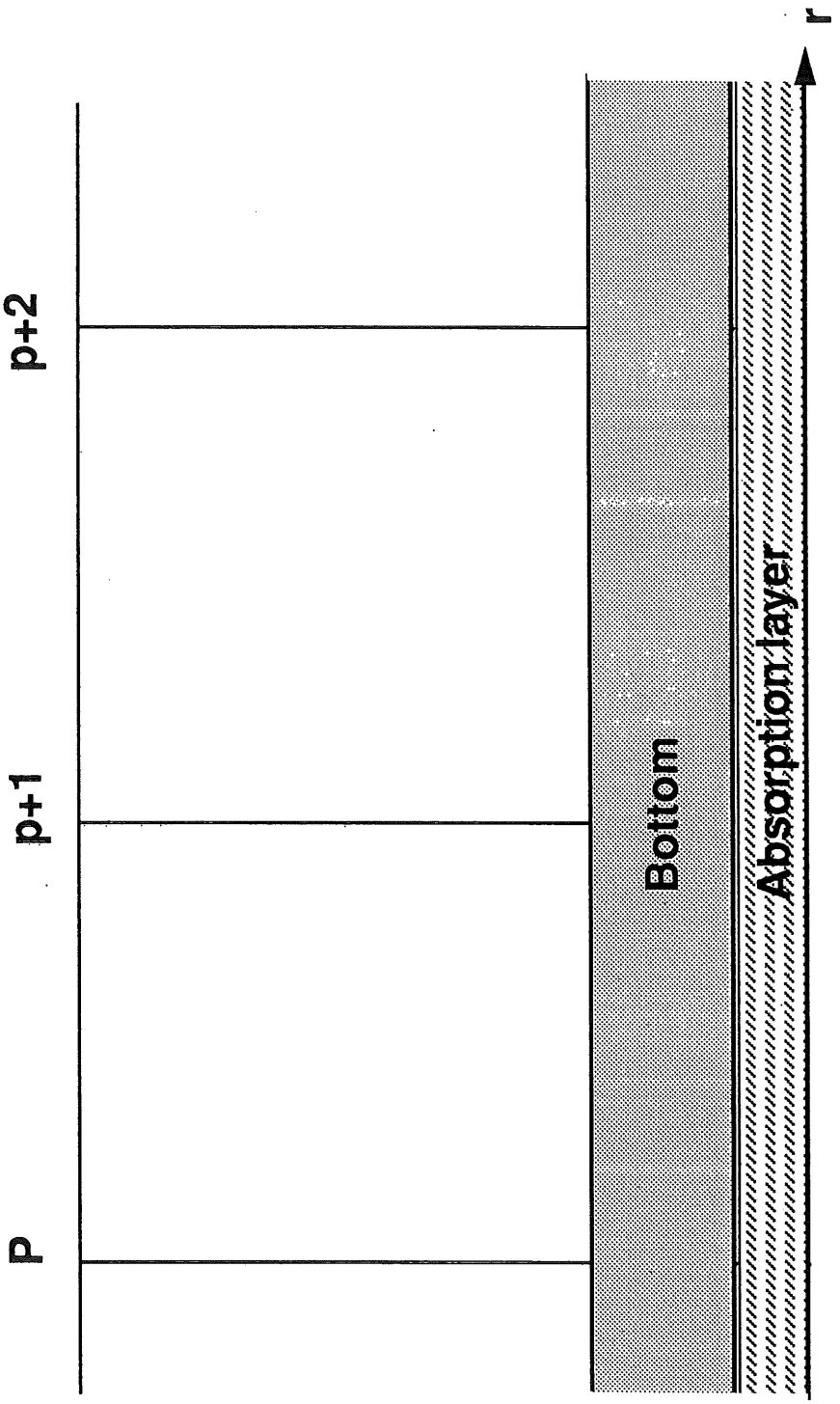



FIG. 2

Theoretical and Computational Acoustics
D. Lee, Y-H Pao, M.H. Schultz, and Y-C Teng (Editors)


# Acoustic Solitary Waves in a Tunnel with an Array of Helmholtz Resonators

N. Sugimoto
Department of Mechanical Engineering, Faculty of Engineering Science
University of Osaka, Toyonaka, Osaka 560, Japan

## 1. Introduction

It has recently been demonstrated theoretically that a soliton† is possible even in propagation of nonlinear acoustic waves through the air [1-3]. This acoustic soliton is predicted for propagation in a tunnel with a periodic array of Helmholtz resonators, if their natural frequency is set high enough and dissipative effects are made negligible. Since nonlinear acoustic waves evolve usually into shock waves, it is the new finding. Physical mechanism to bear the soliton is the dispersion of Bloch wave type due to the spatially periodic structure [4].

If the condition of the high natural frequency is relaxed, the soliton cannot generally exist but there still remains the possibility that other solitary waves can be propagated in place of it. This article clarifies the point by examining the steady propagation of nonlinear acoustic waves in the tunnel. Analysis is made on the basis of the equations derived previously under the continuum approximation for the periodic array of Helmholtz resonators [1].

## 2. Basic equations for nonlinear acoustic waves in the tunnel

Supposing the lossless and one-dimensional propagation along the tunnel, the basic equations are given as follows:

$$\frac{\partial f}{\partial X} - f\frac{\partial f}{\partial \theta} = -K\frac{\partial g}{\partial \theta}\,, \tag{1}$$

$$\frac{\partial^2 g}{\partial \theta^2} + \Omega g = \Omega f\,, \tag{2}$$

where $f$ and $g$ denote the dimensionless excess pressure appropriately normalized, respectively, in the tunnel and in the cavity of resonators, while $X$ and $\theta$ denote, respectively, the dimensionless far-field axial coordinate and the retarded time in a frame moving with the linear sound speed. Here the parameters $K$ and $\Omega$ measure the ratio of the smallness

†By a soliton, we mean a solitary wave as a steady-wave solution to the Korteweg-de Vries equation.

of the resonator $\kappa$ to the small nonlinearity $\varepsilon$, and the ratio of the resonator's natural angular frequency $\omega_0$ to a typical frequeny of the acoustic waves $\omega$, defined, respectively, as follows:

$$K = \frac{\kappa}{2\varepsilon} \quad \text{and} \quad \Omega = \left(\frac{\omega_0}{\omega}\right)^2 , \tag{3}$$

where $\kappa$ ($\ll 1$) is the ratio of the cavity's volume relative to the tunnel's volume per axial spacing between the neighboring resonators and $\varepsilon$ ($\ll 1$) designates the order of the excess pressure relative to the atmospheric. Incidentally if $\Omega$ is large enough, $g$ is approximated by (2) to be

$$g = f - \frac{1}{\Omega}\frac{\partial^2 g}{\partial \theta^2} = f - \frac{1}{\Omega}\frac{\partial^2 f}{\partial \theta^2} + O\left(\frac{1}{\Omega^2}\right) . \tag{4}$$

Substituting (4) into (1), we derive immediately the Korteweg-de Vries equation (called simply K-dV equation hereafter).

The parameters $K$ and $\Omega$ can be removed from (1) and (2) by replacement of $(f, g)$ and $(X, \theta)$, respectively, with $(Kf, Kg)$ and $(X/K\sqrt{\Omega}, \theta/\sqrt{\Omega})$. Thus they are set equal to unity without any loss of generality in the following analysis.

## 3. Propagation of solitary waves

We seek steady-wave solutions to (1) and (2) by assuming that $f$ and $g$ depend on $X$ and $\theta$ only through a combination $\theta - sX$ ($\equiv \zeta$), $s$ being a parameter. Here we simply note that $s$ is a parameter to measure the deviation of physical propagation velocity $a_0(1 - \varepsilon K s)$ from $a_0$. By imposing the boundary condition that the state far ahead of propagation is undisturbed, there exist the solutions satisfying the undisturbed condition far behind as well as far ahead. For $0 < s < 1$, in fact, we have for $0 \le f \le f_+$

$$4\tan^{-1}\sqrt{\frac{f_+ - f}{f - f_-}} - \frac{2s}{\sqrt{-f_+ f_-}} \log\left|\frac{[\sqrt{-f_-(f_+ - f)} - \sqrt{f_+(f - f_-)}]^2}{(f_+ - f_-)f}\right| = \pm\zeta , \tag{5}$$

where $f_+$ and $f_-$ ($-4/3 < f_- < 0 < f_+ < 8/3$) are constants determined by $s$ and the $\zeta$ axis is chosen so that $f$ takes the maximum value $f_+$ at $\zeta = 0$. When $f$ is available, $g$ is obtained by the relation $g = f^2/2 + sf$.

Figure 1 shows the explicit profiles of $f$ for some values of $s$. As can be seen, they are localized in $\zeta$ (i.e. spatially and temporally) and represent the solitary waves. It is found that the solitary waves are *compressive* and that they are propagated with speed slower than the usual sound speed $a_0$ (i.e. *subsonic*), but faster than $a_0(1 - \kappa/2)$ in the linear long-wave limit. As the propagation speed approaches the upper bound ($s \to 0+$), the

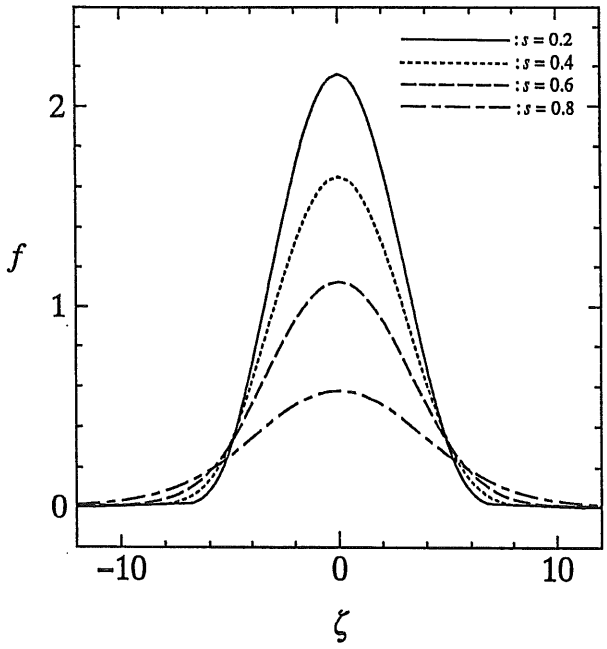



Figure 1: Profiles of the solitary-wave solutions $f$ for $s = 0.2, 0.4, 0.6$ and $0.8$ where $K$ and $\Omega$ are set equal to unity.

solitary waves tend to the limiting solitary wave given by

$$f = \frac{8}{3}\cos^2\frac{\zeta}{4}\,, \tag{6}$$

where $-2\pi \leq \zeta \leq 2\pi$ and $f$ vanishes in the outside of this interval, and therefore the height of the solitary waves approaches the limiting height $8/3$. The excess pressure $\Delta p$ corresponding to this height is given in reference to the atmospheric $p_0$ by

$$\frac{\Delta p}{p_0} = \frac{8\gamma}{3(\gamma+1)}\,\kappa\,, \tag{7}$$

where $\gamma$ is the ratio of the specific heats.

As the speed approaches the lower bound ($s \to 1-$), on the other hand, the solitary waves tend to be the soliton solution of the K-dV equation [2,3]:

$$f = \alpha\,\mathrm{sech}^2\sqrt{\frac{\alpha}{12}}\zeta\,, \tag{8}$$

where $\alpha$ ($0 < \alpha \ll 1$) is arbitrary and $s = 1 - \alpha/3$. Hence it is found that the soliton is included as the special case. For the speed below the lower bound, no steady-wave solutions exist, while for the speed above the upper bound (i.e. *supersonic*), the discontinuous

shock waves are allowed in the solutions. Unlike usual shock waves in the air, there appears behind the discontinuity a nonlinearly oscillatory and periodic wavetrain extending downstream. As the speed increases, the discontinuity also increases but the amplitude of the wavetrain approaches a finite value, while as the speed decreases to approach the linear sound speed, the shock waves tend to diminish of course but the wavetrain remains.

## 4. Conclusions

It has been demonstrated that the acoustic solitary waves can be propagated in a tunnel with a periodic array of Helmholtz resonators. Their explicit profiles have been obtained analytically. It is revealed that while the height of the solitary waves is limited by the limiting solitary wave, solitary waves of small height are well described by the soliton. If a typical frequency of the solitary wave is defined as $\pi/D$, $D$ being its temporal half width, it is found that the solitary waves are possible for $\Omega > 4$. In the case of the soliton, $D$ becomes long in proportion to the inverse square root of the height. Therefore its typical frequency is so low that the assumption $\Omega \gg 1$ is satisfied.

## References

[1] Sugimoto, N. 1992 Propagation of nonlinear acoustic waves in a tunnel with an array of Helmholtz resonators. *J. Fluid Mech.* **244**, 55-78.

[2] Sugimoto, N. 1993 On generation of 'acoustic soliton'. *Advances in Nonlinear Acoustics* (ed. H. Hobaek), World Scientific, 545-550.

[3] Sugimoto, N. 1995 The generation of an acoustic soliton and a soliton tube. *Proc. Estonian Acad. Sci. Phys. Math.* 44, 56-72.

[4] Sugimoto, N. & Horioka, T. 1995 Dispersion characteristics of sound waves in a tunnel with an array of Helmholtz resonators. *J. Acoust. Soc. Am.* **97**, 1446-1459.

Theoretical and Computational Acoustics
D. Lee, Y-H Pao, M.H. Schultz, and Y-C Teng (Editors)


# Two-Dimensional Acoustic Wave Propagation in a Medium Containing Rigid Cracks

Guus J.H. Muijres and Gérard C. Herman
Faculty of Technical Mathematics and Computer Science
Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands

## Abstract

Wave propagation through media containing a large number of inclusions or cracks is a computationally intensive problem, especially if one is interested in dynamic properties of the medium like apparent absorption or dispersion due to multiple scattering caused by the inclusions. In the present paper, an integral-equation method is presented that is particularly efficient due to the use of appropriately chosen expansion functions. In a number of model studies, the method is compared with a perturbative solution and it is found that the latter can also be accurate provided the inclusions are small enough. For the special case of a monochromatic incident plane wave, the perturbative solution can then be used to replace the actual cracks by a much smoother 'apparent' medium which accounts for multiple-scattering effects in terms of a frequency-dependent and angle-dependent dispersion.

## Introduction

The variations in the earth's subsurface range from scales much larger than seismic wavelengths down to scales that are much smaller. These small-scale variations can have a significant effect on the amplitude and phase of the transmitted wave field, as was shown for plane-stratified subsurface models [O'Doherty and Anstey, 1971]. Both stochastic and deterministic methods have been developed for studying this transmission problem in more detail. The deterministic approaches developed so far (see, for instance, [Burridge et al., 1988]) are mainly limited to plane-stratified models. Present-day theories for the propagation of acoustic waves through media containing small-scale inclusions are mostly stochastic (see, for a recent overview, [Hudson and Knopoff, 1989]).
In this paper we discuss a fully deterministic method for computing the wave field transmitted through a 2-D medium containing a large number of small-scale rigid cracks. A rigid crack is characterized by a vanishing normal component of the pressure gradient at the cracks. For a similar treatment of compliant cracks and circular heterogeneities we refer to [Muijres and Herman, 1994].
Starting from an integral representation for the pressure, an integral equation is obtained

for the unknown jump in the pressure across the cracks. By choosing adequate expansion functions, an efficient set of equations is derived with only one unknown coefficient per crack. In this respect, the method compares favorably with methods based on discretization of the wave equation.
We have compared the numerical solution of the resulting system of equations to solutions based on a Neumann series expansion taking scattering processes up to second-order into account. We have found that this expansion is accurate provided the cracks are small enough. This Neumann series expansion can then be used to replace the actual heterogeneous medium by a much simpler 'apparent' one [Herman, 1994].

## Formulation of the problem

We consider two-dimensional, acoustic scattering from a large number of small-scale rigid cracks, embedded in a homogeneous medium. The $n^{th}$ crack, of width $2a_n$, occupies the region $C_n$:

$$C_n = \{ (x,z) \ : \ |x - x_n| < a_n \ \wedge z = z_n \ \}. \tag{1}$$

In Eq. (1), $x_n$ and $z_n$ are the horizontal and vertical coordinate of the center of the crack, respectively. All cracks are assumed to be horizontally aligned.
The total pressure field, $p$, can be written as a superposition of the incident field, $p^{inc}$, which is the field in the absence of cracks, and the scattered field, $p^{sc}$, which accounts for the presence of the cracks:

$$p(x,z;\omega) = p^{inc}(x,z;\omega) + p^{sc}(x,z;\omega), \tag{2}$$

where $\omega$ is the angular frequency. For brevity we omit the explicit $\omega$-dependence in the remainder. Outside the cracks, $p$ satisfies the Helmholtz equation

$$\nabla^2 \, p(x,z) + \frac{\omega^2}{c_0^2} \, p(x,z) = -s(x,z), \tag{3}$$

where $c_0$ is the velocity of the embedding medium and $s(x,z)$ is the source that generates the incident field. If the source is a point source (i.e. $s(x,z) = \delta(x - x', z - z')$) then the solution of Eq. (3) is the Green's function for the embedding medium, given by

$$p^G(x,z;x',z') = \frac{i}{4} H_0^{(1)}(k_0 r), \tag{4}$$

in which $H_0^{(1)}(k_0 r)$ is the zeroth order Hankel function of the first kind, $r$ is given by

$$r(x,x',z,z') = \sqrt{(x-x')^2 + (z-z')^2}, \tag{5}$$

and $k_0$ is the wavenumber ($k_0 = \omega / c_0$). In the following, the Green's function is used for deriving an integral equation formulation from which the scattered field can be determined.

## Integral equation for the rigid crack

The presence of rigid cracks is accounted for by the Neumann boundary condition:

$$\forall_n \quad \lim_{z \to z_n} \frac{\partial p}{\partial z}(x,z) = 0, \quad |x - x_n| < a_n. \tag{6}$$

From the Helmholtz Eq. (3), the following integral representation can be derived for the scattered field outside the crack [van den Berg, 1981]:

$$p^{sc}(x,z) = \sum_{n=1}^{N} \int_{x_n - a_n}^{x_n + a_n} dx' \, \frac{\partial p^G}{\partial z'}(x,z;x',z'=z_n) \, \phi_n(x'), \quad (x,z) \notin C_n, \; n = 1, \cdots, N \tag{7}$$

where $N$ denotes the number of cracks and $\phi_n$ represents the jump in the pressure across the crack,

$$\phi_n(x) = \lim_{z \downarrow z_n} p(x,z) - \lim_{z \uparrow z_n} p(x,z), \quad |x - x_n| < a_n. \tag{8}$$

From Eq. (7) and (8) it is seen that the jump $\phi_n$ in field at the cracks has to be known to compute the field outside the cracks. In order to determine $\phi_n$, first a Fredholm integral equation of the first kind has to be derived from Eq.(7) [van den Berg, 1981]. To this aim, we take the partial derivative of Eq. (7) with respect to $z$ and let the point of observation $(x,z)$ approach crack $C_m$. With the aid of boundary condition (6), we then obtain

$$\forall_m \quad \frac{\partial p^{inc}}{\partial z}(x, z=z_m) = - \lim_{z \to z_m} \sum_{n=1}^{N} \int_{x_n - a_n}^{x_n + a_n} dx' \, \frac{\partial^2 p^G}{\partial z \partial z'}(x,z;x',z'=z_n) \, \phi_n(x'), \quad |x - x_m| < a_m. \tag{9}$$

Eq. (9) is a Fredholm integral equation of the first kind in the unknown functions $\phi_n$.

## Discretisation

In order to solve the integral in Eq.(9), $\phi_n$ is expanded in terms of an appropriately chosen sequence of functions. As the size of the cracks is much smaller than the wavelength, these expansion functions are chosen such that $\phi_n$ is accurately represented by only one expansion function per crack. This implies that we have

$$\phi_n(x) = b_n \, \psi_n(x), \quad |x - x_n| < a_n \, . \tag{10}$$

The choice for $\psi_n$ will be specified later on.
To obtain a linear system of equations we multiply Eq. (9) with an appropriately chosen weight function $w_m(x)$, and integrate the result over $C_m$, to arrive at

$$\forall_m \quad \left( \frac{\partial p^{inc}}{\partial z} \right)_m = \sum_{n=1}^{N} G_{mn} \, b_n \tag{11}$$

where

$$\left( \frac{\partial p^{inc}}{\partial z} \right)_m = \int_{x_m - a_m}^{x_m + a_m} dx \; w_m(x) \, \frac{\partial p^{inc}}{\partial z}(x, z_m) \, , \tag{11a}$$

and

$$G_{mn} = -\lim_{z\to z_m} \int_{x_m-a_m}^{x_m+a_m} dx\; w_m(x) \int_{x_n-a_n}^{x_n+a_n} dx'\; \frac{\partial^2 p^G}{\partial z \partial z'}(x,z;x',z'=z_n)\; \psi_n(x')\,. \tag{11b}$$

Once this $N \times N$ linear system is solved the solutions $b_n$ can be used to compute the field outside the cracks. Methods for solving Eq. (11) include explicit solvers for smaller values of $N$ ($< 500$) and iterative schemes for larger crack numbers.

## Choices for weight and expansion functions

For both the expansion and weight functions we make the following choice:

$$\begin{aligned} \psi_n(x) &= \sqrt{2}\,\sqrt{a_n^2-(x-x_n)^2} \\ w_n(x) &= \sqrt{2}\,\sqrt{a_n^2-(x-x_n)^2} \end{aligned}\quad , \tag{12}$$

The motivation for this choice is that, for a single crack, it results in the same leading-order term as for the case of scattering by a small crack or slit [de Hoop, 1955b]. This implies that one expansion function is sufficient to represent the wave field provided the crack is small enough. With choices (12) the matrix elements of the kernel function $\frac{\partial^2 p^G}{\partial z \partial z'}$ are calculated in Appendix A and are given by:

$$G_{mn} = \begin{cases} \frac{\pi a_n^2}{2}\left(1+\left(\frac{k_0 a_n}{2}\right)^2\left(\log\left(\frac{k_0 a_n \gamma_e}{4}\right) - \frac{i\pi}{2} - \frac{3}{4}\right)\right) & m = n \\ -i\frac{\pi^2 a_m^2 a_n^2}{8} H_{mn}, & m \neq n \end{cases}\quad , \tag{13}$$

where $H_{mn}$ is given by

$$\begin{aligned} H_{mn} \;=\; & \frac{k_0}{r_{mn}^3}\Big\{ (z_m-z_n)^2\, k_0\, |x_m-x_n|\, H_0^{(1)}(k_0 r_{mn}) + \\ & \Big((x_m-x_n)^2 + (z_m-z_n)^2\Big) H_1^{(1)}(k_0 r_{mn})\Big\}. \end{aligned} \tag{14}$$

Here $H_1^{(1)}$ is the first order Hankel function of the first kind and

$$r_{mn} = r(x_m, x_n, z_m, z_n). \tag{15}$$

Using Eq. (10) in Eq. (7), the following expression for the pressure outside the cracks results:

$$p(x,z) = p^{inc}(x,z) + \sum_{n=1}^{N} i\,\frac{\pi a_n^2}{4\sqrt{2}}\, k_0\, \frac{(z-z_n)}{r_n} H_1^{(1)}(k_0 r_n)\; b_n, \tag{16}$$

where $r_n = r(x, x_n, z, z_n)$. To compute the field, the solutions $b_n$ from the system (11) are substituted. Eq. (16) shows an angle-dependent scattering because a rigid crack has an acoustic dipole radiation pattern.

## Neumann series expansion

An alternative approach for solving Eq.(11) is the expansion of $b_n$ in terms of a Neumann series (see, for instance, [Courant and Hilbert, 1931] (p. 119). The use of this approximate solution is motivated by the observation that, under the assumption of a monochromatic incident plane wave and periodically ordered cracks, this Neumann series expansion, if convergent, can be used to replace the actual cracked medium by an much simpler 'apparent' medium which accounts for multiple-scattering effects in terms of a frequency-dependent and angle-dependent dispersion and attenuation [Herman, 1994]. The accuracy of the Neumann series (and also of the apparent medium) can be checked against the numerical solution discussed in the previous section. To first order, this Neumann expansion for the coefficients $b_m$ is given by

$$b_m = \frac{1}{G_{mm}} \Big( \frac{\partial p^{inc}}{\partial z} \Big)_m - \frac{1}{G_{mm}} \sum_{\substack{n=1 \\ n \neq m}}^{N} \frac{G_{mn}}{G_{nn}} \Big( \frac{\partial p^{inc}}{\partial z} \Big)_n . \tag{17}$$

To compute the pressure outside the cracks these coefficients are substituted in Eq. (16). This yields

$$\begin{aligned} p(x,z) \;=\; & p^{inc}(x,z) - \sum_{n=1}^{N} \frac{1}{G_{nn}} \Big( \frac{\partial p^G}{\partial z}(x,z) \Big)_n \Big( \frac{\partial p^{inc}}{\partial z} \Big)_n \\ & + \sum_{\substack{m,n=1 \\ m \neq n}}^{N} \frac{1}{G_{nn}} \Big( \frac{\partial p^G}{\partial z}(x,z) \Big)_n \frac{G_{nm}}{G_{mm}} \Big( \frac{\partial p^{inc}}{\partial z} \Big)_m , \end{aligned} \tag{18}$$

where $\Big( \frac{\partial p^G}{\partial z}(x,z) \Big)_n$ is given by

$$\Big( \frac{\partial p^G}{\partial z}(x,z) \Big)_n = \int_{x_n - a_n}^{x_n + a_n} dx' \, \frac{\partial p^G(x,z;x',z'=z_n)}{\partial z'} \, \psi_n(x'). \tag{19}$$

In expression (18), we recognize, from left to right, three contributions to the total wave field corresponding to different scattering processes: zeroth-order (incident field), first-order (single scattering at all cracks) and a second-order (double scattering involving all pairs of different cracks). In contrast to the numerical solution which takes all reflections into account, the Neumann expansion we use, only accounts for scattering processes up to second order.

## Examples

We present the computed wave field for two distinct cases. First, we consider scattering from a single crack in order to compare our method (I) with a different method (II) developed by [Thorbecke, 1991]. This method (II) is based on an integral equation for the case of a single, large crack, which is solved numerically with a preconditioned conjugate

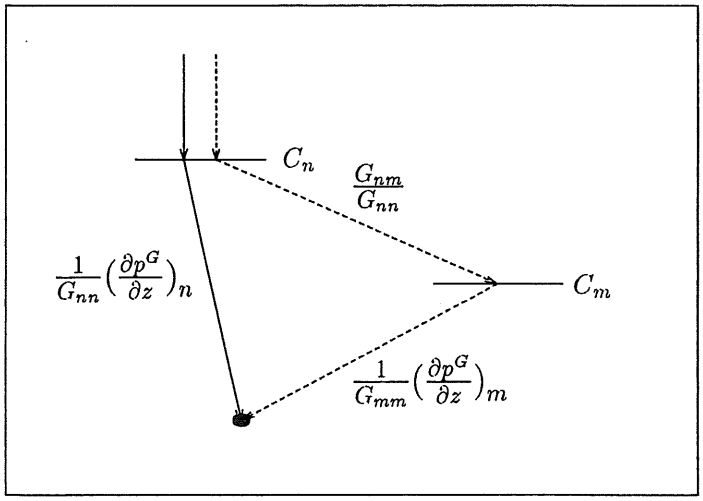



Figure 1: Physical interpretation of the Neumann series.

gradient scheme. The comparison is limited to intermediate crack-sizes as method (I) is most accurate for small cracks and method (II) for larger cracks. Therefore we choose a crack of half-width $a = 7.5\ m$, which is relatively large for our method (I), but relatively small for method (II). The configuration is shown in Fig. 2. The incident field is propagating in the negative $z$-direction. It has a dominant wavelength of about 50 $m$ and it contains frequencies from 0 $Hz$ to 40 $Hz$. The embedding velocity is 1500 $m/s$. The scattered field is computed for a receiver located 200 $m$ above the origin. The results for both methods are shown together in Fig. 3. We observe a good agreement, especially if one realizes that the crack size is such, that both methods (I) and (II) have a limited accuracy. We now consider a plane wave propagating in the positive $z$-direction, through a medium containing 1000 randomly located cracks. The crack half-width is now chosen as $a = 1\ m$. The cracks are confined to a region of $250 \times 250\ m$ centered around the origin. The embedding velocity equals $c_0 = 3000\ m/s$. We calculate the transmitted field for a receiver at a depth of 400 $m$ below the origin. The geometry is sketched in Fig. 4. The spectrum of the incident field contains frequencies between 5 $Hz$ and 60 $Hz$ and has a dominant wavelength of about 100 $m$. In Fig. 5 we observe that the presence of the cracks gives rise to dispersion of the waveform of the direct field.

To solve system (11) we have used a conjugate gradient method, in which all previous search directions are taken into account (see, for instance, [Vuik et al.]). Next we used the second order Neumann series expansion to compute the transmitted field at the receiver. Upon comparison of this solution and the previous we conclude that for small cracks ($a \leq 1\ m$) the Neumann solution is accurate. For larger cracks, the multiple-scattered energy is present in the coda and the Neumann solution is no longer accurate.

From further numerical experiments, we have found that, as the number of cracks increases, the transmitted field becomes more and more coherent and the main effect of the presence of the cracks seems to be an apparent dispersion. This suggests that it is possible to construct an apparent medium for media containing large numbers of rigid

cracks. This has already been shown by [Herman, 1994] for the case of small-scale velocity heterogeneities.

## Conclusions

Based on an integral-equation formulation, an efficient forward modeling scheme is derived. For small cracks and carefully chosen expansion functions, each crack can be accurately accounted for by only one expansion function per crack. By using iterative techniques to solve for the unknown expansion coefficients, wave propagation through media containing large numbers of rigid cracks can be considered.
Comparisons of iterative solutions with solutions based upon a second-order Neumann series expansion show that the Neumann series is accurate provided the cracks are small enough. For the special case of a monochromatic incident plane wave and periodically ordered heterogeneities, this Neumann series expansion can then be used to replace the actual cracked medium by an much simpler 'apparent' medium which accounts for multiple-scattering effects in terms of a frequency-dependent and angle-dependent dispersion and attenuation [Herman, 1994]. This is the topic of our current research.

## Appendix A: Calculation of the matrix elements $G_{mn}$

In the following we use the Fourier integral representation of the Hankel function $H_0^{(1)}$:

$$H_0^{(1)}\Big(k_0\sqrt{(x-x')^2+(z-z')^2}\Big) = \frac{1}{\pi}\int\limits_{-\infty}^{\infty} dk_x \frac{e^{ik_x(x-x')+i\gamma(k_x)|z-z'|}}{\gamma(k_x)} \tag{20}$$

with $\gamma(k_x)^2 + k_x^2 = k_0^2$.
Partial derivatives with respect to $z$ or $z'$ are taken outside the integrals because they do not interfere with integrations over $x$ or $x'$.
**I. The diagonal elements ($m = n$).**
In Eq. (11b) we use the integral representation Eq. (20). Then we apply twice the following property [van den Berg, 1981]

$$\int\limits_{-a}^{a} dx\, \sqrt{1-(\frac{x}{a})^2}\, \exp\left(-ik_x x\right) = \pi\, \frac{J_1(k_x a)}{k_x} \tag{21}$$

to obtain

$$\begin{aligned} G_{nn} &= \lim_{\substack{z\to z_n \\ z'\to z_n}} -\frac{i\pi a_n^2}{2}\,\frac{\partial^2}{\partial z \partial z'}\left(\int\limits_{-\infty}^{\infty} dk_x\, \frac{\exp\left(i\gamma(k_x)|z-z'|\right)}{k_x^2\gamma(k_x)}\, J_1^2(k_x a_n)\right) \\ &= -\frac{i\pi a_n^2}{2}\int\limits_{-\infty}^{\infty} dk_x\, \frac{\gamma(k_x)\, J_1^2(k_x a_n)}{k_x^2} \equiv \pi a_n^2\, I(k_0 a_n). \end{aligned} \tag{22}$$

The last integral can be approximated to third order in $k_0a$ [de Hoop, 1955a] by

$$\begin{aligned} I(k_0a) &= \frac{1}{2} - \frac{1}{4\pi}\sum_{p=1}^{\infty} \frac{\Gamma(p+\frac{1}{2})\Gamma(p-\frac{1}{2})(-k_0a^2)^p}{\Gamma(p)\Gamma(p+1)\Gamma(p+1)\Gamma(p+2)}\Big[2\log(k_0a) - i\pi + \\ &\qquad \psi(p+\frac{1}{2}) + \psi(p-\frac{1}{2}) - \psi(p) - 2\psi(p+1) - \psi(p+2)\Big] \\ &= \frac{1}{2}\left(1 + \left(\frac{k_0a}{2}\right)^2\left(\log(\frac{k_0a\gamma_e}{4}) - \frac{i\pi}{2} - 3/4\right)\right) + \mathcal{O}(k_0a^4) \end{aligned}$$

where $\Gamma$ and $\psi$ are the gamma and psi functions (see [Abramowitz]) and $\log(\gamma_e) \approx 0.577215$ is Euler's constant. Using this result in Eq. (22) we get the following approximation for the diagonal elements

$$G_{nn} = \frac{\pi a_n^2}{2}\left(1 + \left(\frac{k_0a_n}{2}\right)^2\left(\log(\frac{k_0a_n\gamma_e}{4}) - \frac{i\pi}{2} - \frac{3}{4}\right)\right). \tag{23}$$

**II. The off-diagonal elements ($m \neq n$).**

We perform a Taylor series expansion of $H_0^{(1)}(k_0|r-r'|)$ in $x$ and $x'$ around $x_m$ and $x_n$, respectively. To lowest order this yields

$$H_0^{(1)}(k_0\sqrt{(x-x')^2+(z-z')^2}\,) \approx H_0^{(1)}(k_0\sqrt{(x_m-x_n)^2+(z-z')^2}\,). \tag{24}$$

After substitution in Eq. (11b) and calculating the resulting integrals we find for the off-diagonal elements

$$G_{mn} = \lim_{\substack{z\to z_m \\ z'\to z_n}} -\frac{i\pi^2a_m^2a_n^2}{8}\frac{\partial^2}{\partial z\partial z'}\left(H_0^{(1)}(k_0\sqrt{(x_m-x_n)^2+(z-z')^2}\,)\right). \tag{25}$$

The partial derivatives can be calculated using [Abramowitz]

$$\frac{\partial}{\partial u}H_0^{(1)}(u) = -H_1^{(1)}(u), \tag{26}$$

which leads to the expressions (13) for the off-diagonal elements.

## Appendix B: Calculation of the vector elements $\left(\frac{\partial p^{inc}}{\partial z}\right)_m$

We expand $\frac{\partial p^{inc}}{\partial z}(x, z_m)$ to lowest order in a Taylor series around $x = x_m$ and obtain from Eq. (11a)

$$\left(\frac{\partial p^{inc}}{\partial z}\right)_m = \lim_{z\to z_m}\frac{\pi a_m^2}{\sqrt{2}}\frac{\partial p^{inc}}{\partial z}(x_m, z). \tag{27}$$

The partial derivative can be calculated once the incident field is specified.

## References

[Abramowitz] Abramowitz, M. and Stegun, I.A., Handbook of Mathematical Functions: Dover Publications.

[van den Berg, 1981] van den Berg, P.M., 1981, Transition matrix in acoustic scattering by a strip: Journ. Acoust. Soc. Am., **70**, 615-619.

[Burridge et al., 1988] Burridge, R., Papanicolaou, G.S., and White, B.S., 1988, One-dimensional wave propagation in a highly discontinuous medium: Wave Motion, **10**, 19-44.

[Courant and Hilbert, 1931] Courant, R. and Hilbert, D., Methoden der Mathematischen Physik I: Springer Verlag (Berlin).

[de Hoop, 1955a] de Hoop, A.T., 1955, On integrals occuring in the variational formulation of diffraction problems: Proc. Kon. Ned. Akad. Wet. **B58**, 325-330.

[de Hoop, 1955b] de Hoop, A.T., 1955, Variational formulation of two-dimensional diffraction problems with application to diffraction by a slit: Proc. Kon. Ned. Akad. Wet. **B58**, 401-411.

[Herman, 1994] Herman, G.C., 1994, Transmission of acoustic waves through strongly heterogeneous, two-dimensional velocity models: Wave Motion, **20**, 111-130.

[Hudson and Knopoff, 1989] Hudson, J.A., and Knopoff, L., 1989, Predicting the overall properties of composite materials with small-scale inclusions or cracks: PAGEOPH, **131**, 551-576.

[Muijres and Herman, 1994] Muijres, Guus J.H. and Herman, Gerard C., Two-dimensional acoustic waves in cracked media: 64nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1310-1313.

[O'Doherty and Anstey, 1971] O'Doherty, R.F., and Anstey, N.A., 1971, Reflections on amplitudes: Geophysical Prospecting **19**, 430-458.

[Thorbecke, 1991] Thorbecke, J.W., Scattering by a strip in a homogeneous medium, 1991, MSc-thesis.

[Vuik et al.] Vuik, K., Sevink, A.G.J., and Herman, G.C., A preconditioned Krylov subspace method for the solution of least squares problems: submitted to the Journal of Computational Physics.

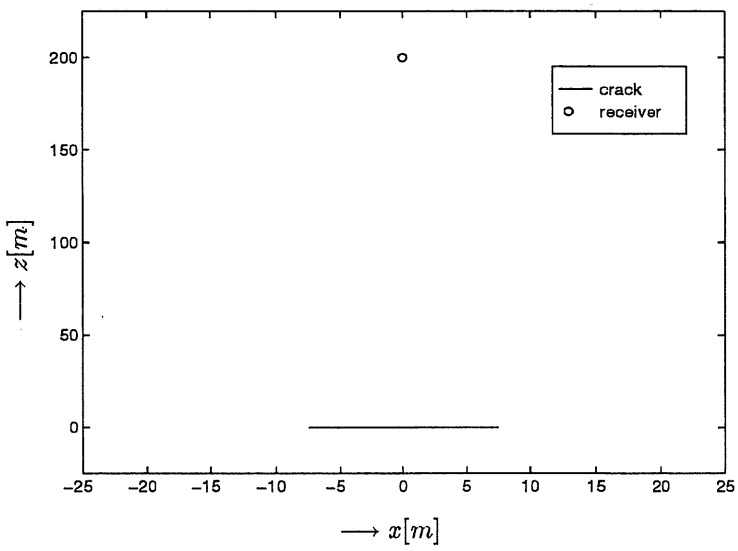



Figure 2: Crack ($a = 7.5\ m$) and receiver location $(0, 200)$ indicated with o.

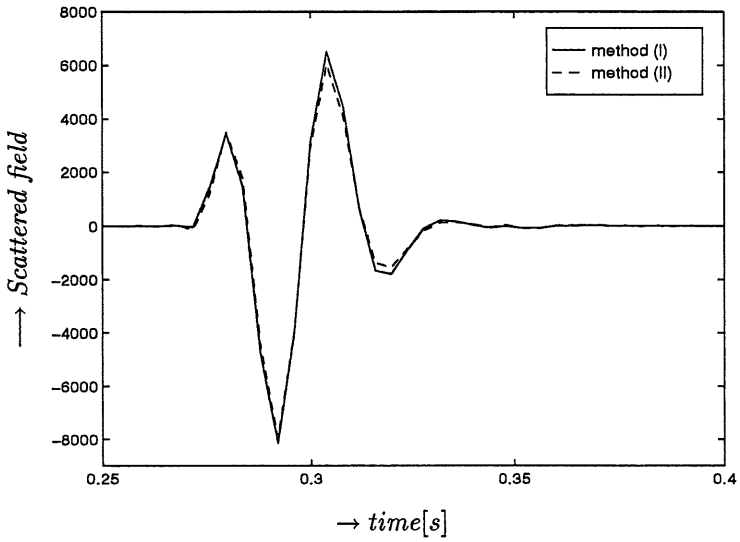



Figure 3: Comparison of scattered field from a single crack for 2 different methods (I) and (II) (See text). Crack size: $a = 7.5\ m$.

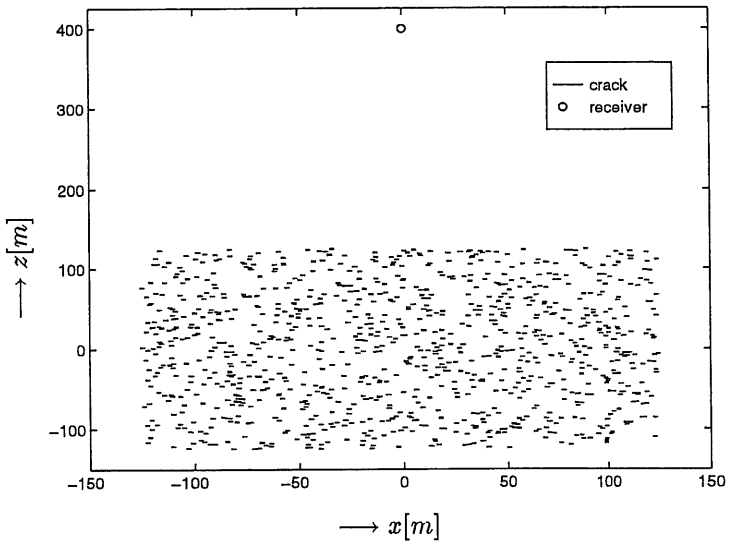



Figure 4: Crack distribution (1000 cracks, $a = 1\ m$) and receiver location $(0, 400)$ indicated with ∘.

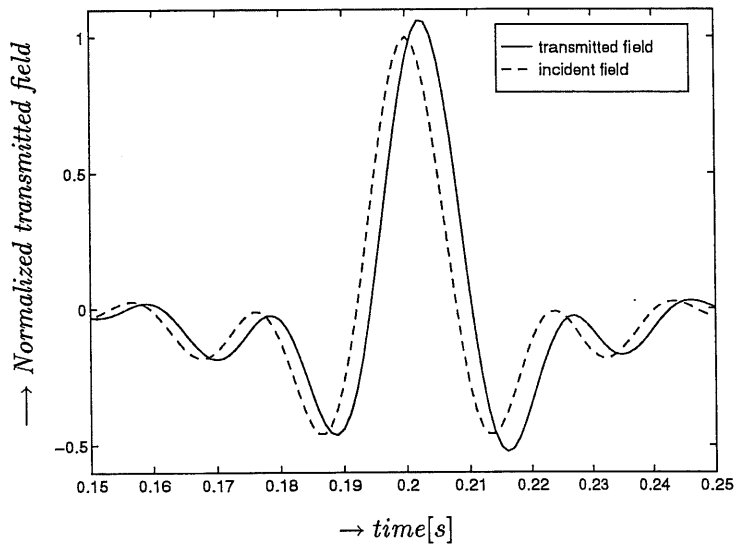



Figure 5: Comparison of the incident field with the transmitted field. Number of cracks: $N = 1000$, crack size: $a = 1\ m$.

Theoretical and Computational Acoustics
D. Lee, Y-H Pao, M.H. Schultz, and Y-C Teng (Editors)


# Neural Network Inversion and Applications to Estimate Some Parameters of a Submerged Moving Body

Dongyu Fei, Yu-Chiung Teng, and John T. Kuo
Aldridge Laboratory of Applied Geophysics, Henry Krumb School of Mines
Columbia University, New York, NY 10027, USA

## ABSTRACT

**This paper presents the concept, principle, and method of neural network inversion. The neural network refers to a particular multi-layered, paralleled data processing system which can adapt the weight matrix to match the changes of the environment. The neural network inversion means to implement the mapping from a multi-dimensional space of observation field to another multi-dimensional parameter's space by this particular system. As one of the applications to the ocean science, the research discusses the detection of some parameters for a moving body that can be seen as a moving acoustic source. A trained neural network has a capability to find the parameters of a submerged moving body with a real time solution.**

## 1 INTRODUCTION

The detection of a submerged moving source has attracted attentions for the last decade. Basically, there are two measurement means: acoustics and electromagnetics. As a acoustic approach a source that emits a constant tone and moves at a constant speed can be localized by measurement of the Doppler shifted frequencies (DSF). The acoustic signal is received by spatially separated sensors. There must be five sensors to give a determinant solution. Weinstein (1982) gives an exact solution. Weistein and Levanon (1980) and Statman and Rodemich (1987) gives iterative solutions, and Chan and Jardines (1990) and Chan and Towers (1992) gives the grid search solution. The high dimensionality leads to a large computational time. Chan (1994) developed a one dimensional grid search solution which requires three sensors and greatly reduces the computational time.

As an electromagnetic approach it is possible to detect the submerged body itself or its wake. Robindon (1992), Marshall (1988) indicated that the sensitivity of electromagnetic field sensors has been improved to allow non-negligible detection ranges. However, the electromagnetic fields produced by the eddy currents in its metal parts which are rotating in the earth magnetic field may not expect to observe more then several hundred meters away from the moving body. Tuck (1994) presented a theoretical study and shown that the motion of sea water due to the wake of a submerged moving body induces significant magnetic signal as far away as ten kilometers along its path. Earlier, it was the general belief that only shot-period waves in a moderately rough sea could

produce a measurable magnetic field. However, Weaver (1965) showed that long-wavelength ocean swell can be as important as local wind waves of great amplitude.

In this paper we present a new approach based on the finite element method (Teng, 1989; Teng, 1993) and neural network inversion (Fei, 1995; Fei, Kuo and Teng, 1995) to localize the submerged moving source. Although geophysicists have successfully applied neural network to detect unknown underground inclusions, designing and training a network are still more of arts than science (Poulton and et al., 1992). Fei (1995) and Fei, Kuo and Teng (1995) present the concept and principle of neural network inversion to detect the spatial parameters of unknown inclusions (Fei, Teng and Kuo, 1994). Moreover, the basic theorem-three-layer neural network existence theorem (Lorentz, 1976; Kolomogorov, 1957; Hecht-Nielsen, 1987; Hecht-Nielsen, 1989; Carrol and Dickinson, 1989; Cybenko, 1989) and multi-layer neural network existence theorem (Fei, Kuo and Teng, 1995) provide a solid theoretical foundation for the neural network inversion.

An acoustic field distribution for a moving source that emits a series pulse with special forms and moves at a constant course and speed can be computed by finite element method. The finite element modeling can simulate unstratified fluids and complex sea bottom structure, this method can generate real data to train the neural nets. The trained network can be used to find a real time solution based the neural network inversion principle. Using the 3-D finite element method to solve problems presented by realistic sea shore structures still requires considerable memory and execution time in computation. In this research, we only provide the numerical data from the 1-D and 2-D finite element modeling.

## 2 FINITE ELEMENT MODELING

We employ finite element method as a training data generator. The wave equation of pressure field $p$ for a moving source with a constant velocity $v_0$ is

$$\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} + S\delta(x - x_0 + v_0 t)\delta(z - z_0)$$

$$c^2 = k/\rho$$

where $\delta$ is Dirac delta function, $x_0, z_0$ is the source point, $c$ is the sound speed of acoustic wave, $k$ is the bulk modulus, $\rho$ is mass density and $S$ is the applied source function.

The finite element equation in matrix form is

$$[M]\{\ddot{p}\} + [K]\{p\} = \{f\}$$

where $[M]$ is the global mass matrix, $[K]$ is the global pseudo stiffness matrix, $\{\ddot{p}\}$ is the column matrix for pressure field. $\{f\}$ is the applied source matrix.

The time integration-explicit central difference scheme is

$$\{p(t + \Delta t)\} = \{\dot{p}\,(t) + \{p(t)\}\,\Delta t$$

$$\{\dot{p}\,(t + \Delta t)\} = \{\dot{p}\,(t) - [K][M]^{-1}\{p(t + \Delta t)\}\,\Delta t$$

The Figure 1 indicates the source function for this moving source. While in the conversional methods the moving source emits a constant frequent tone. The two kinds of sources can be converted by Fourier transformation.

## 3 NEURAL NETWORK AND NEURAL NETWORK INVERSION

The term neural network (or artificial neural network), is derived from its resemblance to the biological interconnection of neurons. A typical neural network consists of many processing elements,

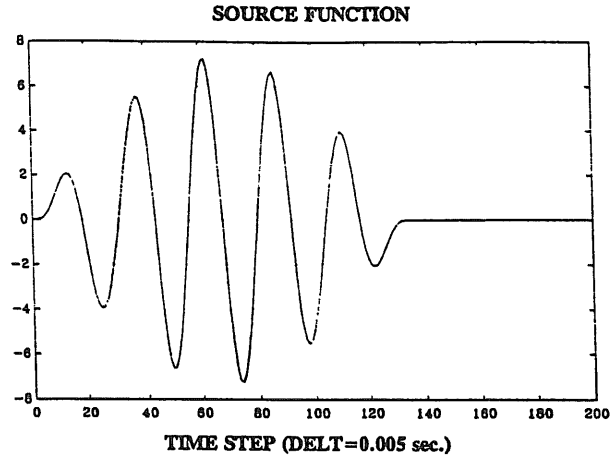



**Figure 1: The source function of a moving source.**

each of which is tightly connected to another by the modifiable weights. Neural network is expected to learn by representative sets of examples from the given condition.

The neural network may be defined as:

**Definition** The neural network is defined as an overly simplified human-brain system with a parallel, distributed information processing structure, which consists of a number of layered processing elements. Each element in a given layer has a single output connection into all elements in the next layer. When the information is sensed by analogy to, say, the eyes, ears or skin, it is passed on to the multi-layer nerve system. Comparison with the prior information as the trained network keeps the learning results as a fixed weigh matrix, the neural network thus discriminates the output of the results, just as if the human-brain system has the capability of judgment.

In this paper we simply regard "neural network" as a parallel, three-layer data-processing machine without real biological meaning attached. A three-layer neural network configuration consists of one input, one hidden, and one output layer. Every element (or neuron) is connected to the next layer by the so called "weights." The weights are calculated during a supervised training process, in which a representative set of input and a desired output is presented to the network. The network undergoes a learning procedure to register all changes of weights through the feed-forward and backpropagation neural networks cyclically. One cycle (or one iteration) is to complete a feed-forward and backpropagation training procedure.

For a given sea water model with a moving source, the finite-element modeling method can simulate the distribution of the wave-field on the sea water as

$$\mathbf{p}(\mathbf{m}, x_r, t, n)$$

where $\mathbf{m}$ denotes the source position, $x_r$ denotes the recording point, $t$ is the time, and $n$ denotes environment parameters of the sea.

The meaning of inversion here is to seek a set of parameters $\mathbf{m}$ that minimizes some norm of the difference between the observed data $\mathbf{d}$ and the predicted data $\mathbf{p}$. In general, the $L_2$ norm of the data mismatch is minimized, leading to the following objective function,

$$F(\mathbf{m}) = |\mathbf{d} - \mathbf{p}(\mathbf{m})|^2 = |\mathbf{e}|^2 .$$

The solution for an inverse problem is defined as a set of parameters $\mathbf{m}$ for which $F(\mathbf{m})$ is the global minimum of the function $F$.

The objective function is generally minimized interactively by updating the subsequent initial guess of the position parameter $\mathbf{m}_0$.

Assume the unknown true value is $\mathbf{m}_t$. Set up an n-dimensional Euclidean space with coordinates $M_1$, $M_2$, ...,$M_n$. The initial guess $\mathbf{m}_0$ and the unknown true value $\mathbf{m}_t$ are mapped into the n-dimensional space. The inversion procedure can be seen as a point movement from $\mathbf{m}_0$ toward $\mathbf{m}_t$. We may use the finite-element method to forward-model $\mathbf{p}(\mathbf{m}_0)$, and calculate the error function such that

$$|\mathbf{d} - \mathbf{p}(\mathbf{m}_0)|^2 = |\mathbf{e}|^2$$

Therefore, we adopt an interactive inversion method to move the point $\mathbf{m}_0$ through $m_{1...}m_n$. The distance and direction of the movement are related to the adopted method. By means of minimization, the error function approaches a minimum $e_0 \geq e_1 \geq e_2 ... \geq e_n \rightarrow 0$. When $|\mathbf{e}|$ is smaller than a given positive value, this interactive procedure terminates. This global minimum is empirically assured in a conventional inversion method, when the initial point $\mathbf{m}_0$ is in the valley of the global minimum, the search will successfully reach the point $\mathbf{m}_t$. When the initial point $\mathbf{m}_0$ is trapped in the valley of the local minima, the search fails.

We choose $k$ sampling points $\mathbf{m}_1, \mathbf{m}_2, ..., \mathbf{m}_k$, which are uniformly distributed in the sphere. Then we establish a matrix $\mathbf{p}_{tm}$ to train the neural network, where

$$\mathbf{p}_{tm} = \mathbf{p}(\mathbf{m}, x_r, t, ).$$

The training procedure can be viewed as the confirmation of the mapping between $\mathbf{p}_{tm}$ and $\mathbf{m}$. After learning, the neural network assumes the capability to inverse the real observed data $\mathbf{p}_{tm}$ to find the $\mathbf{m}_t$.

# 4 NEURAL NETWORK TRAINING

In the feed-forward neural network, The input data points of the vector $x_{i_1}^{(1)}$ is first presented to the $l_1$ input layer of the neural units so that the after-sigmoidal output at the $l_2$ later (or hidden layer) is

$$y_{i_2}^{(2)}(\mathbf{x}) = \sigma(\sum_{i_1=1}^{I_1} W_{i_1 i_2}^{(1,2)} x_{i_1}^{(1)}) = \sigma(S_{i_2}^{(2)})$$

which, in turn, is used as the input to the $l_2$ layer.

The after-sigmoidal output at the $l_3$ output layer is

$$y_{i_3}^{(3)}(\mathbf{x}) = \sigma(S_{i_3}^{(3)}(\mathbf{x}) = \sigma(\sum_{i_2=1}^{I_2} W_{i_2 i_3}^{(2,3)} \sigma(\sum_{i_1=1}^{I_1} W_{i_1 i_2}^{(1,2)} x_{i_1}^{(1)})).$$

In above two equations, $W_{i_1 i_2}^{(1,2)}$ and $W_{i_2 i_3}^{(2,3)}$ are the weights, linking the $l_1$ and $l_2$ layer, and the $l_2$ and $l_3$ layer, respectively, $\sigma$ is a sigmoidal function, $i_1$=1, 2, ..., $I_1$, and $i_2$ = 1, 2, ...$I_2$ are the neuron's numbers in the $l_1$ and $l_2$ layer, respectively.

In the feed-forward neural network, $x_{i_1}^{(1)}$ is known, and in the initial feed-forward (or the first iteration), $W_{i_1 i_2}^{(1,2)}$ and $W_{i_2 i_3}^{(2,3)}$ are made to be random, and $\sigma$ is the sigmoidal function.

$$\sigma(s) = \frac{1}{1 + e^{-S}}.$$

The backpropagation is essentially a generalization of the least-squares procedure for the neural network between the input and output layers. The change of the weight is calculated starting on

the output layer and going backward calculating updated weights for each layer. In the process, the so-call "delta-rule" is applied to increase the speed of convergence to stable state.

In training, the post-sigmoidal error $e_{i_3}^{(3)}[t]$ at the $l_3$ layer is then backpropagated to the $l_3$ layer,

$$e_{i_3}^{(3)}[t] \leftarrow y_{i_3}^{(3)'} - y_{i_3}^{(3)}$$

where $y_{i_3}^{(3)'}$ is the known desired output.

The post-sigmoidal error in the $l_2$ (or hidden) layer through backpropagation is then:

$$e_{i_2}^{(2)}[t] \leftarrow \sum_{i_3=1}^{I_3} W_{i_2 i_3}^{(2,3)}[t]\delta_{i_3}^{(3)}[t],$$

where $\delta_{i_3}^{(3)}[t]$ is obtained by:

$$\delta_{i_3}^{(3)}[t] \leftarrow \frac{\partial\sigma(p_{i_3}^{(3)}[t])}{p_{i_3}^{(3)}[t]} e_{i_3}^{(3)}[t],$$

and

$$p_{i_3}^{(3)}[t] \leftarrow \sum_{i_2=1}^{I_2} W_{i_2 i_3}^{(2,3)}[t] q_{i_2}^{(2)}[t].$$

The after-sigmoidal input $q_{i_2}^{(2)}[t]$ at the $l_2$ layer is simply

$$q_{i_2}^{(2)}[t] = \sigma(S_{i_2}^{(2)}) = \frac{1}{1+e^{-S_{i_2}^{(2)}}},$$

and, $W_{i_2 i_3}^{(2,3)}$ is obtained from the initial feed-forward in being random.

Again, here $W_{i_2 i_3}^{(2,3)}[t]$ is the random weights assigned in the initial feed-forward network.

At this stage the change of the weights between neuron $i_3$ in the $l_3$ layer and every neuron $i_2$ in the $l_2$ layer can be calculated by:

$$\Delta W_{i_2 i_3}^{(2,3)}[t] \leftarrow \beta\delta_{i_3}^{(3)}[t] q_{i_2}^{(2)}[t]$$

and the modified weight with the addition of the momentum term

$$W_{i_2 i_3}^{(2,3)}[t+1] = W_{i_2 i_3}^{(2,3)}[t] + \beta\delta_{i_3}^{(3)}[t] q_{i_2}^{(2)}[t] + \alpha \Delta W_{i_2 i_3}^{(2,3)}[t-1],$$

where $\beta$ is the learning rate and $\alpha$ is the momentum rate.

Following the same procedure, we are able to backpropagate to obtain the post-sigmoidal error $e_{i_1}^{(1)}[t]$, the change of the weight $\Delta W_{i_1 i_2}^{(1,2)}[t-1]$ and the modified weight $W_{i_1 i_2}^{(1,2)}[t+1]$ linking the $l_2$ to $l_1$ layer.

For the second iteration (or the second cycle), the second feed-forward output at the $l_3$ layer $y_{i_3}^{(3)}(\mathbf{x})$ can be obtained by the feed-forward process, using the weights $W_{i_1 i_2}^{(1,2)}$ and $W_{i_2 i_3}^{(2,3)}$ having been updated in the first iteration, and the input data points of the vector $x_{i_1}^{(1)}$. By the same token the post-sigmoidal error $e_{i_3}^{(3)}$ at the $l_3$ layer is backpropagated through the $l_3$, $l_2$, and $l_1$ layer to complete the second iteration (or the second cycle), and to obtain the newly updated modified weights $W_{i_1 i_2}^{(1,2)}$ and $W_{i_2 i_3}^{(2,3)}$.

Such an iterative feed-forward and backpropagation procedure is repeatedly carried out to obtain the final optimized weights.

If the network is properly trained with a representative set of acoustic data, the final inversion is accomplished by the feed-forward network to obtain the after-sigmoidal output at the $l_3$ layer that is the most optimized output in the sense of least-square-error.

**Table 1. The velocity inversion for 1-D modeling.**

| No | NNT inversion unit(km) | True value unit(km) | Accuracy % | NNT inversion unit(km) | True value unit(km) | Accuracy % |
|---|---|---|---|---|---|---|
| 1 | 20.23 | 20 | 1.13 | 19.76 | 20 | -1.2 |
| 2 | 25.03 | 25 | 0.11 | 25.05 | 25 | 0.22 |
| 3 | 29.92 | 30 | -0.25 | 30.16 | 30 | 0.53 |
| 4 | 35.04 | 35 | 0.11 | 35.03 | 35 | 0.07 |
| 5 | 39.98 | 40 | -0.05 | 39.97 | 40 | -0.06 |
| 6 | 44.92 | 45 | -0.18 | 45.02 | 45 | 0.04 |
| 7 | 49.94 | 50 | -0.12 | 49.92 | 50 | -0.16 |
| 8 | 54.99 | 55 | -0.01 | 54.99 | 55 | -0.02 |
| 9 | 60.03 | 60 | 0.05 | 59.94 | 60 | -0.1 |
| Average | | | 0.22 | | | 0.26 |
| | Position=12.5km | | | Position=13km | | |

| No | NNT inversion unit(km) | True value unit(km) | Accuracy % | NNT inversion unit(km) | True value unit(km) | Accuracy % |
|---|---|---|---|---|---|---|
| 1 | 20.06 | 20 | 0.31 | 19.81 | 20 | -0.97 |
| 2 | 25.1 | 25 | 0.41 | 25.09 | 25 | 0.37 |
| 3 | 29.82 | 30 | -0.59 | 29.98 | 30 | -0.08 |
| 4 | 34.95 | 35 | -0.13 | 35.04 | 35 | 0.11 |
| 5 | 39.99 | 40 | -0.02 | 39.99 | 40 | -0.03 |
| 6 | 45.05 | 45 | 0.12 | 44.98 | 45 | -0.04 |
| 7 | 49.84 | 50 | -0.32 | 50 | 50 | 0.01 |
| 8 | 55.03 | 55 | 0.06 | 55.08 | 55 | 0.15 |
| 9 | 59.99 | 60 | -0.01 | 60.06 | 60 | 0.1 |
| Average | | | 0.21 | | | 0.21 |
| | Position=13.5km | | | Posision=14km | | |

| No | NNT inversion unit(km) | True value unit(km) | Accuracy % | NNT inversion unit(km) | True value unit(km) | Accuracy % |
|---|---|---|---|---|---|---|
| 1 | 20.24 | 20 | 1.18 | 20.11 | 20 | 0.54 |
| 2 | 24.94 | 25 | -0.22 | 25.05 | 25 | 0.2 |
| 3 | 29.87 | 30 | -0.43 | 29.89 | 30 | -0.36 |
| 4 | 35 | 35 | -0.01 | 35.05 | 35 | 0.13 |
| 5 | 39.96 | 40 | -0.09 | 39.87 | 40 | -0.32 |
| 6 | 45.04 | 45 | 0.09 | 45.05 | 45 | 0.1 |
| 7 | 49.96 | 50 | -0.08 | 50 | 50 | 0 |
| 8 | 55.11 | 55 | 0.2 | 54.95 | 55 | -0.08 |
| 9 | 60.06 | 60 | 0.09 | 59.98 | 60 | -0.03 |
| Average | | | 0.26 | | | 0.2 |
| | Position=14.5km | | | Position=15km | | |

# 5 APPLICATIONS

Figure 2 shows the 1-D finite element modeling of a moving source. The acoustic velocity of the sea water is 1.5 km/s. The x-coordinate of the receiver is 0 km. The x-coordinate of the start point is from 12 to 18 km at an interval 0.5 km. The velocity of the moving source is 5, 10, ... 60 km/h.

Figure 3 illustrates the training data. The training model number 1 to 12 denotes the velocities of the moving source from 5 to 60 km/h at an interval 5 km/h. Figure 3(a). The start point position of the moving source is 12 km. Figure 3(b). The start point position of the moving source is 14.5 km. Figure 3(c). The start point position of the moving source is 17 km.

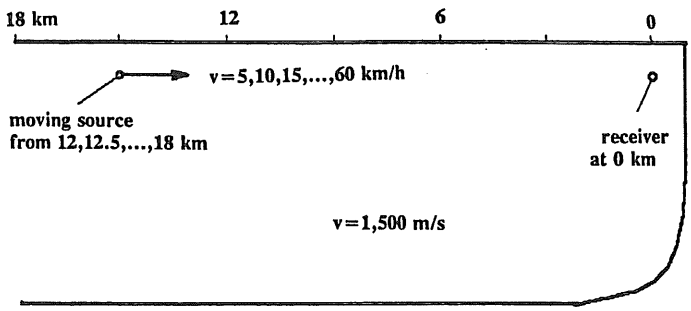



**Figure 2: 1-D finite-element model. The acoustic velocity of the sea water is 1.5 km/s. The x-coordinate of the receiver is 0 km. The x-coordinate of the start point is from 12 to 18 km at an interval 0.5 km. The velocity of the moving source is 5, 10, ... 60 km/h.**

Table 1 indicates the result of neural network inversion for the 1-D moving source for the position of 12.5, 13, 13.5, 14, 14.5, 15 km. The average accuracy is 0.22%.

Figure 4 is 2-D finite-element modeling for a moving source. The acoustic velocity of sea water is 1.5 km/s. The thickness of the layer is 1000m. The second layer's velocity is 2.7 km/s. The depth of the moving source is 200 m, and the start position is 0 km. The x-coordinate of the receivers on the surface of the sea-air is 0, 0.5,...6.0 km. The velocity of the moving source is from 5 km/h to 60 km/h at an interval 5 km.

Figure 5 shows the finite element synthetic seismogram. Figure 5(a). The velocity of the moving source is 20 km/h. Figure 5(b). The velocity of the moving source is 40 km/h. Figure 5(c). The velocity of the moving source is 60 km/h.

Figure 6 shows the training data, which is collected from the results of the finite element method. The training model number 1 to 12 denotes the velocities of the moving source from 5 to 60 km/h at an interval 5 km/h. Figure 6(a). The velocity of the moving source is 5 km/h. Figure 6(b). The velocity of the moving source is 30 km/h. Figure 6(c). The velocity of the moving source is 55 km/h.

Table 2 shows the detection of velocity for 2-D finite element modeling. For the position of 5, 10, 15, 25 km, the average accuracy of the neural network inversion is 2.7%, 3.7%, 2.2%, 2.36%, respectively.

# 6 CONCLUSIONS

1) A combination of the finite-element forward modeling method and the neural network inversion method proves to be an efficient tool to deal with the detection of speed and position of a moving source in the sea water. The sea water can be unstratified fluids and the host medium can be a complex structure. The finite-element forward modeling is essential to provide accurate, reliable and representative sets of data in the neural network training.

2) Although this paper deals with the moving source in 1-D and 2-D sea water environment, the treatment is equally applicable to 3-D resistively problems.

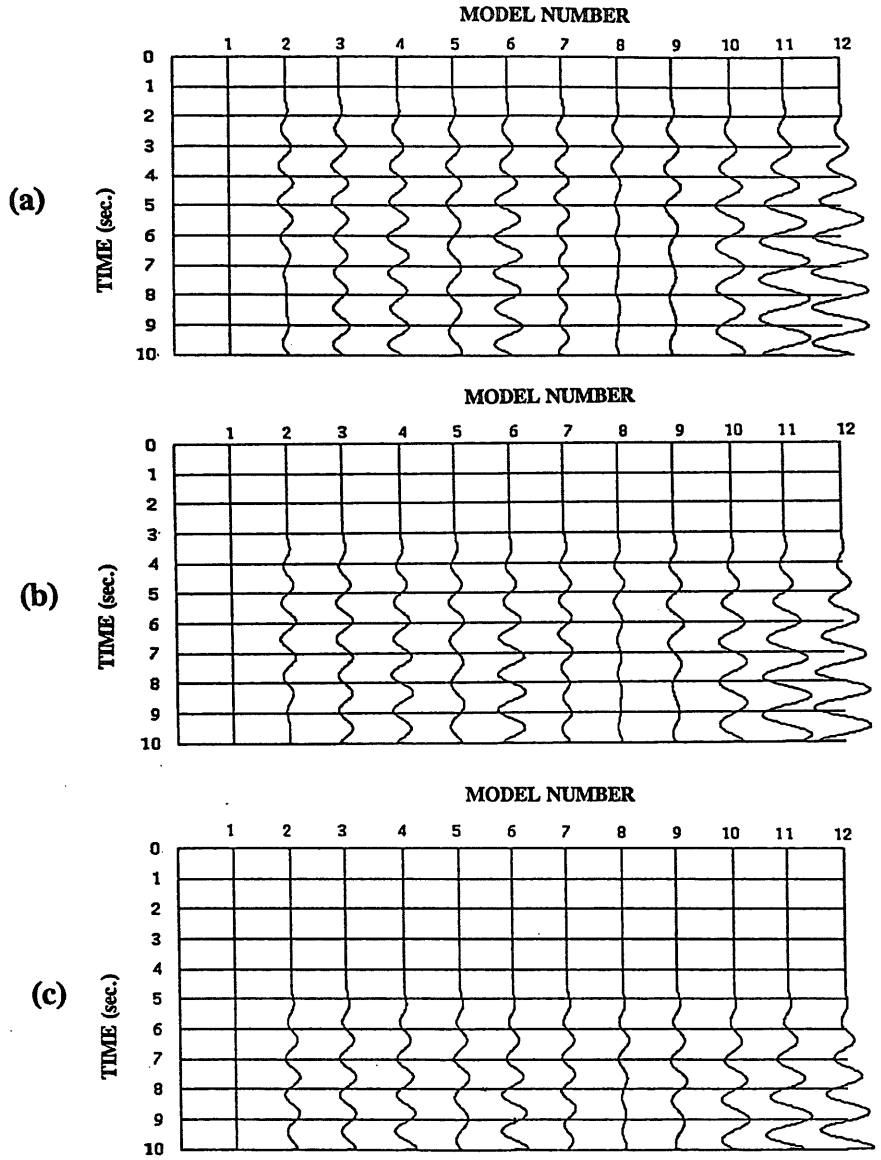



Figure 3: The training data. The training model number 1 to 12 denotes the velocities of the moving source from 5 to 60 km/h at an interval 5 km/h. a) The start point position of the moving source is 12 km. b) The start point position of the moving source is 14.5 km. c) The start point position of the moving source is 17 km.

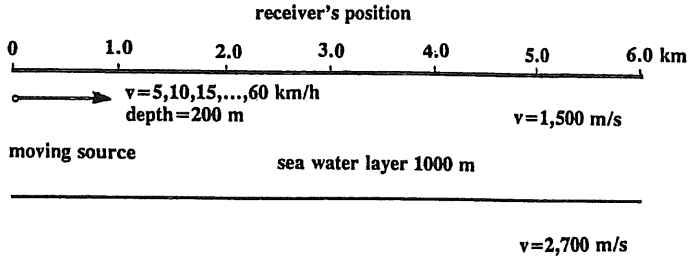



Figure 4: 2-D finite-element model. The acoustic velocity of the sea water is 1.5 km/s. The thickness of the first layer is 1000 m. The second layer's velocity is 2.7 km/s. The depth of the moving source is 200 m, and the start position is 0 km. The x-coordinate of the receivers on the surface of the sea-air is 0, 0.5,...6.0 km. The velocity of the moving source is from 5 km/h to 60 km/h at an interval 5 km.

Table 2. The velocity inversion for 2-D modeling.

| No | NNT inversion unit(km) | True value unit(km) | Accuracy % | NNT inversion unit(km) | True value unit(km) | Accuracy % |
|---|---|---|---|---|---|---|
| 1 | 21.75 | 20 | 8.75 | 17.13 | 20 | -14.36 |
| 2 | 24.1 | 25 | -3.6 | 24.19 | 25 | -3.22 |
| 3 | 30.88 | 30 | 2.95 | 28.69 | 30 | -4.36 |
| 4 | 35.17 | 35 | 0.48 | 34.38 | 35 | -1.76 |
| 5 | 39.19 | 40 | -2.03 | 42.26 | 40 | 5.66 |
| 6 | 45.44 | 45 | 0.97 | 45.35 | 45 | 0.77 |
| 7 | 48.68 | 50 | -2.65 | 49.56 | 50 | -0.88 |
| 8 | 56.33 | 55 | 2.42 | 54.21 | 55 | -1.43 |
| 9 | 59.72 | 60 | -0.47 | 60.96 | 60 | 1.61 |
| Average | | | 2.7 | | | 3.7 |
| | Position=5km | | | Position=10km | | |

| No | NNT inversion unit(km) | True value unit(km) | Accuracy % | NNT inversion unit(km) | True value unit(km) | Accuracy % |
|---|---|---|---|---|---|---|
| 1 | 21.17 | 20 | 5.85 | 14.22 | 20 | *** |
| 2 | 24.52 | 25 | -1.92 | 24.49 | 25 | -2.06 |
| 3 | 30.03 | 30 | 0.1 | 29.59 | 30 | -1.36 |
| 4 | 33.85 | 35 | -3.29 | 34.91 | 35 | -0.25 |
| 5 | 39.43 | 40 | -1.44 | 40.07 | 40 | 0.19 |
| 6 | 46.92 | 45 | 4.26 | 46.02 | 45 | 2.26 |
| 7 | 50.84 | 50 | 1.69 | 53.16 | 50 | 6.32 |
| 8 | 54.39 | 55 | -1.11 | 52.11 | 55 | -5.25 |
| 9 | 60.06 | 60 | 0.1 | 60.72 | 60 | 1.21 |
| Average | | | 2.19 | | | 2.36 |
| | Position=15km | | | Position=25km | | |

# 7 ACKNOWLEDGEMENTS

This research is partially supported by the Air Force Office of Scientific Research under Contract F49620-93-0073.

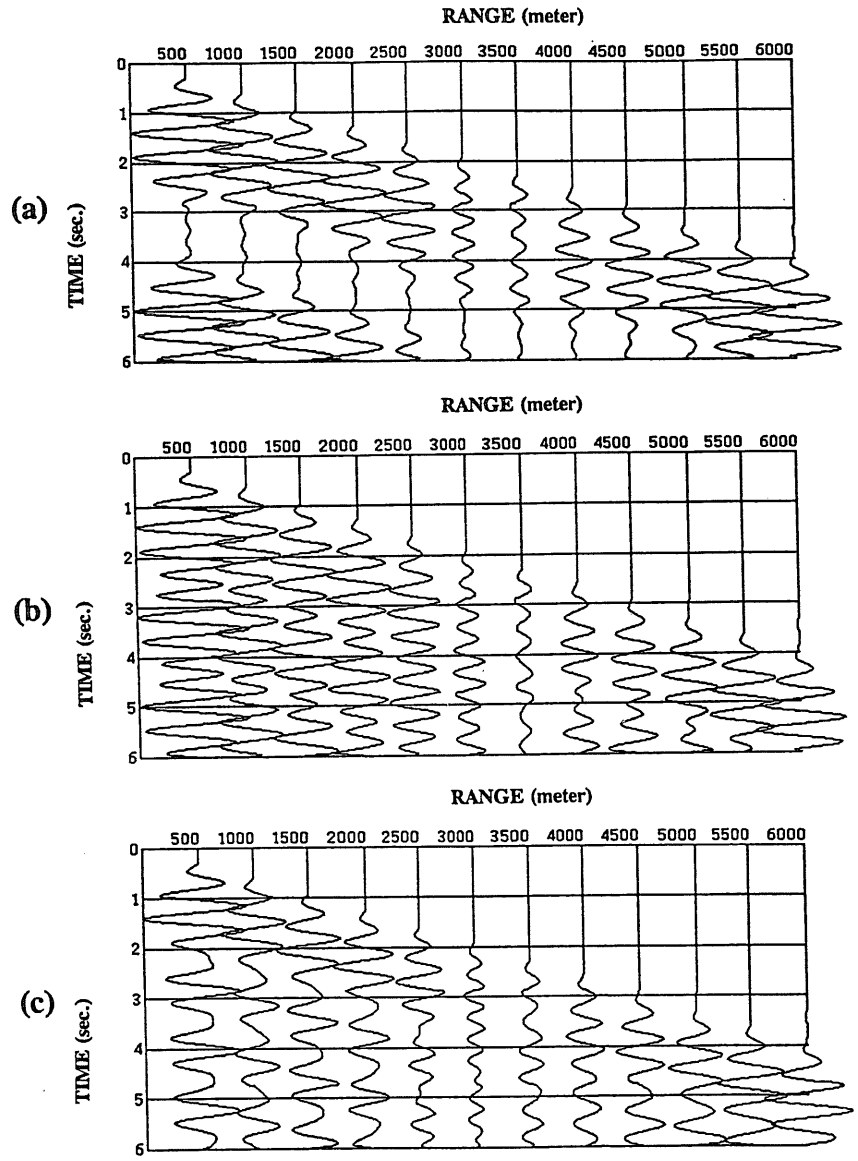



Figure 5: The finite element synthetic seismogram. a) The velocity of the moving source is 20 km/h. b) The velocity of the moving source is 40 km/h. c) The velocity of the moving source is 60 km/h.

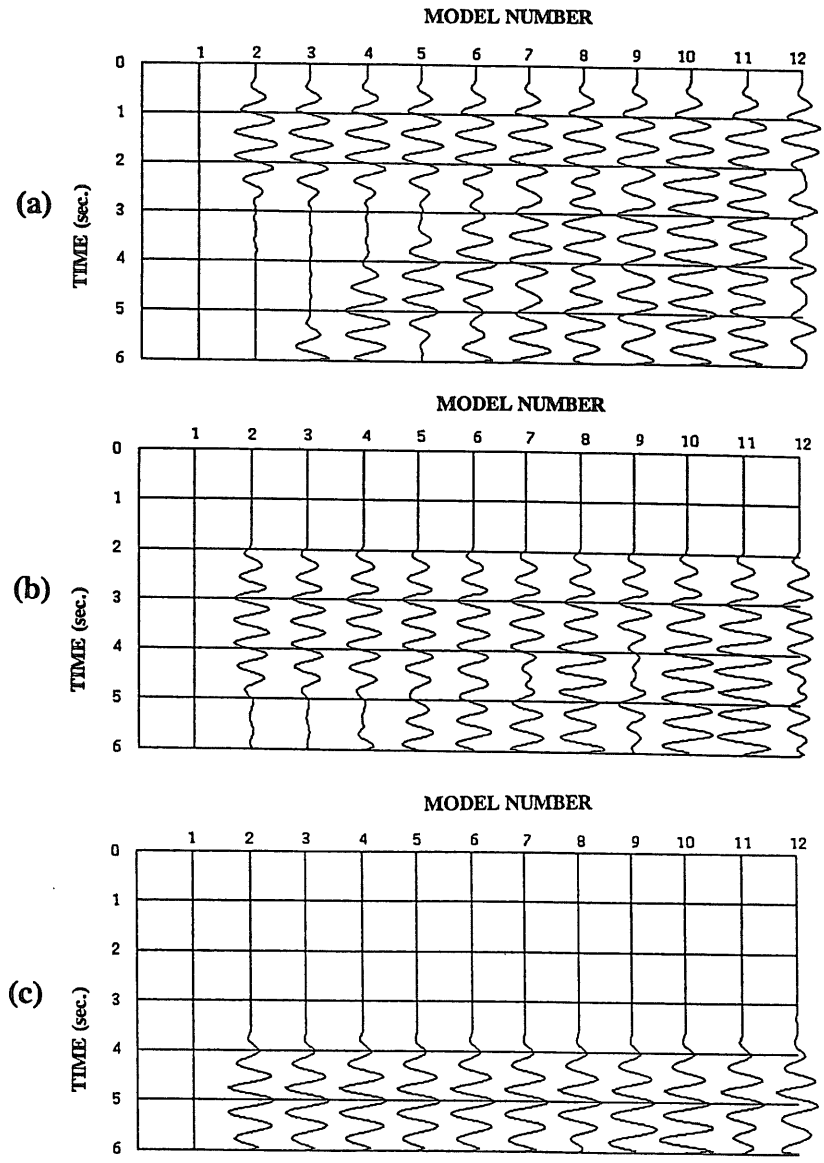



Figure 6: The training data collected from the computation results of the finite element method. The training model number 1 to 12 denotes the velocities of the moving source from 5 to 60 km/h at an interval 5 km/h. a) The velocity of the moving source is 5 km/h. b) The velocity of the moving source is 30 km/h. c) The velocity of the moving source is 55 km/h.

# 8 REFERENCES


1. Chan, Y. T. and Jardines, F. L., 1990, Target localization and tracking from Doppler-shifted frequency measurements: IEEE J. Ocean. Eng., vol. 15, pp. 251-257.
2. Chan, Y. T. and Towers, J. J., 1992, Passive localization from Doppler shifted frequency measurements: IEEE Trans. Signal Processing, Vol. 40, pp. 2594-259.
3. Chan, Y. T., 1994, A 1-d search solution for localization from frequency measurements: IEEE J. of oceanic eng., vol. 19, no. 3.
4. Carrol, S. M. and Dickinson, B. W., 1989, Construction of neural nets using the radon transform: IEEE INNS Internat. Joint Conf. , vol.1, pp. 607-611.
5. Cybenko, G. 1989, Approximation by superpositions of sigmoidal function: 2, Math. control systems signals. pp. 303-314.
6. Fei, Dongyu, 1995, Neural network inversion - theory, principle and applications: Ph.D. Thesis, Columbia University.
7. Fei, Dongyu, Kuo, J. T., and Teng, Yu-Chiung, 1995, Waveform inversion and multi-layer neural network: J. of Computational acoustics. vol 3, pp. 175-202.
8. Fei, Dongyu, Teng, Yu-Chiung, and Kuo, J. T., 1994, Detection of conductive thick plate based on finite element method and neural networks: 64th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, pp. 636-639.
9. Hecht-Nielsen, R., 1987, Kolmogorov's mapping and neural network existence theorem: First international Conference on Neural networks, IEEE. pp.11-14.
10. Hecht-Nielsen, R., 1989, Theory of the backpropagation neural network: IEEE INNS International Joint Conference on Neural networks, vol.1, pp. 593-606.
11. Kolmogorov, A. N., 1957, On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition: Dokalady Akademi Nauk SSSR, 114, pp. 953-956.
12. Lorentz, G. G., 1976, The thirteenth problem of Hilbert: Proceedings of symposia in pure mathematics: Providence, RI, American Math. Society, vol 28, pp. 419-30.
13. Marshall, B. J., 1988, Undersea measurement ranges-The impact of technology: Oct. pp. 681-686.
14. Poulton, M. M., Sternberg, B. K. and Glass, C. E., 1992, Location of subsurface targets in geophysical data using neural networks: Geophysics, vol 57, pp.1534-1544.
15. Robindon, C. A., 1992, Radar detection challenges submarine warfare shroud: Signal, pp. 31-34, Mar.
16. Statman, J. K. and Rodemich, E. R., 1987, Parameter estimation based on Doppler frequency shifts: IEEE Trans. Aerosp. Electron. Syst., vol. AES-16, pp. 31-39.
17. Teng, Yu-Chiung, 1989, Three-dimensional finite element analysis of waves in acoustic media with inclusion: J. Acoust. Soc. Am., vol 86, pp. 414-422.
18. Teng, Yu-Chiung, 1993, Scattering of transient waves by finite cracks in an anti-plane strain elastic solid: J. of computational acoustics, vol. 1, no. 1, pp.101-116.
19. Tuck, E. O. 1994, The induced electromagnetic field associated with submerged moving bodies in an unstratified conducting field: IEEE J. on oceanic eng., vol. 19, no. 2.
20. Weaver, T. J., 1965, Magnetic variations associated with ocean waves and swell: J. Geophys. Res., Vol. 70, no. 8, pp. 1921-1928.
21. Weinstein, E., (1982), Optimal source localization and tracking from passive array measurements: IEEE Trans. Acoust., Speesh, Signal Processing, vol. ASSP-30, pp. 69-76.
22. Weinstein, E. and Levanon, N., 1980, Passive array tracking of a continuous wave transmitting projectile: IEEE Trans. Aerosp. Electron. Syst., vol. AES-16, pp. 721-726.

Theoretical and Computational Acoustics
D. Lee, Y-H Pao, M.H. Schultz, and Y-C Teng (Editors)


# Three Neural Inverse Methods for Ocean Acoustic Tomography

Yann Stephan
Centre Militaire d'Océanographie, EPSHOM, BP426, F-29275 Brest Cedex

Sylvie Thiria, Fouad Badran
Centre National des Arts et Métiers, 292, rue Saint-Martin, F-75003 Paris

Francois Régis Martin-Lauzer
Groupe d'Études Sous-Marine de l'Atlantique, DCN, BP 42, F-29240 Brest Naval

**Abstract**

Ocean Acoustic Tomography (OAT) is an oceanographic method for the global observation of the ocean. All the concepts rely on the ability of the ocean as a multipath sound guiding system on long ranges. Acoustical data are used to map the structure of the sound velocity field in both time and space. Usually, linear inversions are performed. They use a linearization around a reference environment i.e. the real sound velocity field is considered to be the sum of the reference field and a first order perturbation field. The difference between travel times of rays computed in the reference environment and the measured travel times (travel time anomaly) of the same rays is inverted to retrieve the sound speed perturbations. We present in this paper three non linear inversion methods based upon the use of neural networks: direct inversion, network (or adjoint) inversion and distal inversion. The methods are applied in the case of simulated data in a Mediterranean environment. They are compared with a usual linear method (based upon perturbations around a reference environment).

# 1 INTRODUCTION

Ocean Acoustic Tomography (OAT) is an oceanographic method for the global observation of the ocean [1]. All the concepts rely on the ability of the ocean as a multipath sound guiding system on long ranges. Acoustical data are used to map the structure of the sound velocity field in both time and space. More precisely, the estimation of the sound-velocity is obtained by inverting the arrival times of a given signal emitted from a fixed source to a fixed receiver at a fixed range. Usually, linear inversions are performed. They use a linearization around a reference environment, i.e. the real sound velocity field is considered to be the sum of the reference field and a first order perturbation field. The difference between travel times of rays computed in the reference environment and the measured travel times (travel time anomaly) of the same rays is inverted to retrieve the sound speed perturbations. For these inversions, the construction of a model mapping the sound speed perturbation onto the travel time anomaly is required and two major hypothesis are assumed. First, good a prior information about the reference environment has to be known to track and identify arrivals. Second, it is assumed that the paths of the rays in the reference environment are not altered by the perturbation of the sound speed environment. These hypothesis can yield to errors and biases in inversion results.

A new approach, avoiding the latter assumption has been proposed recently using neural networks [2]. It relies on the ability of neural networks to approximate non-linear functions. Given a model for the forward problem, a neural network is used to approximate the inverse solution by simply reversing the role of input and output parameters (direct inversion). We propose in this paper alternative solutions to perform indirect inversion with neural nets. The first approach is referred to the adjoint inversion method). It consists of approximating the forward function with a multilayer Perceptron, to freeze weights and to apply a backpropagation-to-input algorithm (instead of backpropagation-to-weights as in the direct method). This scheme can be related to adjoint-based inversion methods used in Meteorology and Oceanography. The second approach is a combination of the two previous ones. It consists in composing both the inverse and forward problem to retrieve the identity function. It is referred to the distal inversion.

The paper is organized as follows: Section 2 gives some background on acoustic tomography and describes the inverse problem to be solved. Section 3 describes the three inverse methods. Section 4 presents the results in the case of simulated data in single slice tomography. The results are compared to those obtained by a usual linear scheme.

# 2 THE INVERSE PROBLEM IN SINGLE-SLICE ACOUSTIC TOMOGRAPHY

Neglecting the oceanic current effect, the travel time $t_i$ of the acoustic signal along the path $R_i$ is given by:

$$t_i = \int_{R_i} \frac{ds}{c(r,z)} \tag{1}$$

where $r$ and $z$ denotes respectively the horizontal and vertical axis, $c(r,z)$ is the sound velocity field between the source and the receiver and $s$ is the curvilinear abscissa. The relationship between an arrival time and the sound speed field is denoted $G_i$. It is non-linear and takes the form of:

$$G_i(c(r,z)) = \int_{R_i} \frac{ds}{c(r,z)} \tag{2}$$

Denoting by N the number of rays, the problem can be written in term of a set $G = (G_1, G_2, ..., G_N)$ of non-linear equations so that:

$$t_i = G_i(c(r, z)) \tag{3}$$

The general inverse problem in OAT consists in retrieving an estimation of the sound-speed environment from the time measurements $\mathbf{t}=(t_i)_{i=1,..,N}$. The usual inversion methods are based upon a linearization around a reference environment :The real sound velocity field is considered to be the sum of the reference field and a first order perturbation field. The difference between travel times of rays computed in the reference environment and the measured travel times (travel time anomaly) of the same rays is inverted to retrieve the sound speed perturbations [3, 4, 5].

# 3 THREE NEURAL SCHEMES FOR INVERSION

To avoid the linearity hypothesis, we propose in this paper "by learning" inversion methods. They rely on the ability of neural networks to approximate non-linear functions.

## 3.1 Multilayer Perceptron (MLP) for multidimensional approximation

Connectionist models rely upon the behavior of a set of interconnected elementary units (neurons). The global properties of so-called Artificial Neural Networks depend on the internal structure of the model. The main properties used in function approximation are learning and retrieving ability: a neural net is able to learn a relationship from a given set of input-output patterns and to generalize (i.e approximate) the global relationship for unlearned input patterns. Multilayer Perceptrons are a particular class of neural networks in which neurons are connected in layers. One layer is receiving the system input (input layer). One layer is providing the network response to the input (output layer). One or more layers are internal to the network (hidden layers). The layers are fully connected (each neuron of a layer is connected to every neuron of the next layer) and the signals propagate from input to output (no backward information). Biases (additional weights connected to a constant input) can be added on the output and hidden neurons in order to reinforce the non-linear behavior of the network. In this work, the transfer function of hidden units is sigmoidal and the transfer function of the output units is linear. The problem of multidimensional approximation can be posed as follows [6]: Given a real function $F$:
$F : \mathcal{R}^Q \rightarrow \mathcal{R}^P$
$\mathbf{x} = (x_1, .., x_Q) \rightarrow \mathbf{y} = (y_1, .., y_P) = F(\mathbf{x})$,
and a set of example behaviors of $F$ (i.e a set $\mathcal{L}$ of N examples $(x^0, y^0)$), the problem is to find a function $G(W, \mathbf{x})$ which gives the best approximation of $F$ on $\mathcal{L}$. W is a vector of internal parameters of the approximating function $G$. This can be achieved by finding a configuration $W^*$ such as:

$$||G(W^*, \mathbf{x}) - F(\mathbf{x})|| \leq ||G(W, \mathbf{x}) - F(\mathbf{x})|| \tag{4}$$

for all W. The notation $||...||$ stands for an arbitrary norm in $\mathcal{R}^Q$. This definition is general and is not specially related to the neural net concept. However, it can be easily interpreted in the neural framework: the approximation function $G(W, \mathbf{x})$ generalizes the function $F$ from a set of learned examples $\mathcal{L}$ ($\mathcal{L}$ is called the learning base). For this approximation task, the MLP is built as follows. Each x-component is encoded on one neuron of the input layer. Each y-component is encoded on one neuron of the output layer. A cost function, measuring the sum-squared error between the output of the network and the values of F is defined as follows:

$$E(W) = \frac{1}{2} \sum_{m=1}^{N} \sum_{j=1}^{P} [y_{j,m}(W, x_m^0) - y_{j,m}^0]^2 \quad (5)$$

The optimal weight configuration is found by minimizing the cost function using a stochastic backpropagation algorithm [7]. Theoretical works have been done to demonstrate the ability of MLP to approximate continuous multidimensional function defined on compact subsets [8, 9, 10]. Given a function $F$, there exists a multilayer neural network with one or more hidden layers able to approximate $F$ in the uniform convergence sense. Nevertheless, there are no absolute rules to determine the optimal architecture of the network (number of hidden layers, number of cells in each layer). It must be done heuristically. Approximating a function by increasing the number of hidden units is always possible. However, to be statistically relevant, the approximation has to be done on a number of examples "much greater" than the number of weights. Generally, arbitrarily increasing the number of hidden unit (i.e. the number of internal parameters in the net) generally leads to poor performances in generalization.

## 3.2 Multilayer Perceptrons for tomographic inversions

### 3.2.1 Direct inverse modeling

The general scheme of Direct Inverse Modeling (DIM) is given in figure 1. This approach simply consists of reversing the input and output of the problem. A propagation model computes a series of ray pattern for a given set of sound-velocity profiles. Then, a multilayer Perceptron is taught to learn the inverse mapping i.e. the relation between the time pattern in input and the sound velocity profiles in output. Taking notation of 3.1. and introducing $c$ for sound velocity in output and $t$ for time in input, the cost function can be write as follows:

$$E(W) = \frac{1}{2} \sum_{m=1}^{N} \sum_{j=1}^{P} [c_{j,m}(W, t_m^0) - t_{j,m}^0]^2 \quad (6)$$

The optimal weight configuration W* is found by minimizing the cost function using a stochastic backpropagation algorithm [7]. After each presentation of an input-output pattern (step $s$), the weights are modified following a gradient rule:

$$w_{k,l}^s = w_{k,l}^{s-1} - \alpha \left( \frac{\partial E}{\partial w_{k,l}} \right)^s \quad (7)$$

with $\alpha > 0$.

### 3.2.2 Adjoint inverse modeling

The general scheme of adjoint inverse modeling (AIM) is given in figure 2. This approach consists of approximating the forward function with a multilayer Perceptrons. i.e. the relation between the sound velocity profiles and the time pattern. This is achieved following a standard backpropagation algorithm. The cost function to minimize becomes:

$$E(W) = \frac{1}{2} \sum_{m=1}^{N} \sum_{j=1}^{P} [t_{j,m}(W, c_m^0) - c_{j,m}^0]^2 \quad (8)$$

The optimal weight configuration $W*$ is computed following rule (7). Once the network in trained on the forward problem, the weights of the network are frozen. An initial solution is presented in input of the network. The network computes its ouputs which are an estimate of the temporal pattern associated to the initial sound velocity. The output error is backpropagated

in the network and the initial solution is iteratively updated so that the output error decreases. The input adaptation is done following a gradient rule:

$$w_{k,l}^{s} = w_{k,l}^{s-1} - \alpha \left( \frac{\partial E}{\partial x_l} \right)^{s} \tag{9}$$

Note that the goal is to adjust the input (backpropagation-to-input) but weights remain unchanged. Such a scheme can be referred to "network inversion" [11]. The term "adjoint" inversion needs to be clarified. Given a model (in our case a ray tracing model), the purpose of the adjoint model is to give a convenient solution to compute the sensitivities of the output parameters of the model with respect to the input parameters. In words, the adjoint inversion model has to solve the problem: for a given perturbation of the output arount a target value (so, a local error), what is the perturbation to add on the inputs to minimize the error?

### 3.2.3 Distal inverse modeling

The distal inverse modeling (DIM) is a combination of the two previous ones. It has been proposed by Jordan and Rumelhart [12] for control inverse problems. Basically, it consists of composing both the inverse and the forward problem to retrieve the identity function. The forward problem is learned by a multilayer Perceptrons out of context. The weights are frozen. The distal scheme is used in coupling two nets. The time patterns to be inverted are presented to a randomly initialized net. The outputs of this network (which give a first estimate of the sound velocitiy profile) are presented to the forward network which computes an estimation of the temporal patterns. The error (i.e the difference between the estimation of the time patterns -outputs of the system- and the real time patterns -inputs of the system-) is backpropagated in the weight-frozen network (backpropagation-to-input) and consecutively backpropagated in the first network (backpropagation-to-weight). When the iteration scheme is completed, the first network achieves an estimation of the inverse problem. The general scheme is given in figure 3.

### 3.2.4 Remarks

The simplest approach to invert data is to reverse the role of inputs and outputs. This is basically the idea of the direct inverse modeling algorithm. For bijective inverse functions, one can always find a network which realizes the approximation on a given compact subset. The problem stems from the fact that inverse functions are often non univoque (e.g. inverse of constant functions). The philosophy of the adjoint inversion technique is to answer the following question: Given a frozen network, what input must be presented to the network to produce a set of target output values? This method does not enable to approximate multivoque functions but enables to control the solution (or to choose one among all solutions) respect to given "constraints" (assimilation of observations, initial guess, data assimilation,...). In addition, the adjoint inversion can be "goal-directed" in the research of the solution since it relies on a local estimate of a cost function and an iterative perturbation of an initial guess whereas the direct inverse modeling is a global searching method. As far as the distal inversion scheme is concerned, it is a combination of the two preceeding schemes. Its major advantage is that it enables to approximate multivoque functions (several output possible for a given input) by automatically choosing one among all solutions. It is goal directed and is convenient in case of "many-to-one" mapping, in particular in case of non convex domain for wich a linear combination of possible solutions (as the average for example) may not be acceptable.

# 4 RESULTS

## 4.1 Numerical experiment

We simulate a tomography experiment in the Western Mediterranean Sea. A source is moored at 100 meters depth and a receiver is moored at 400 meters depth (figure 4). We simulate the propagation on the abyssal plain (flat bottom) for a 50 kilometers propagation range. The forward problem is solved by a ray tracing model. An arbitrary set of sound velocity profiles $c(z)$ has been generated from typical situations in the Mediterranean Sea taken from historical data). The profiles are defined by 4 points at the respective depths of 0, 60, 300 and 2500 meters. The sound velocity variables are c(0), c(60), c and c(300). The sound velocity at the bottom is assumed to be constant (the oceanic variability is negligible). Each sound velocity values can vary in a given interval whose extension is assumed from oceanographic consideration. The set of sound velocity profiles is uniformly distributed in S=[1506.6,1508.6] x[1506.6,1507.6] x [1511.3,1511.8] (units in m/s). The increments in the intervals were respectively 0.2, 0.2, 0.1(m/s) for the learning base and 0.15, 0.07, 0,08 (m/s) for the testing base. Six identified and resolved rays were used for inversion. The profiles for wich the ray tracing algorithm did not predict the six rays have been arbitrarly removed (to avoid tedious manual refined computation). At this step, 1162 profiles were left. Among this set, 282 profiles were used to build the learning base. The performance of the inverse methods were tested on the 880 remaining profiles and are given in table 1. An example of profile inversion is given in figure 5 in the case of direct inverse modeling. A comparison is given with a usual linear method based upon the canonical decomposition of the sound velocity profiles on a basis of triangular functions.

## 4.2 Discussion

The performance of the three neural methods is good. A good accuracy is obtained (relative error lower than 3 p.c.). The absolute error is greater on the surface parameter which is expected to be less influent on the temporal pattern. Nevertheless, the relative error is greater on the shallow parameter (sound velocity at 60 meters). All neural methods compare with favor to the linear scheme. The performance of the linear scheme is maximum near the reference profile. The results get worse when increasing the "distance" to the reference profile. In order to keep the linear assumption we have restricted the sound velocity perturbation values. Even in favorable cases for linear inversions, all neural methods are more accurate (figure 6, table 1) We have compared the linear method and the direct inverse method for other types of mediterranean profiles [13]. Conversely to linear schemes, the neural performances do not depend on the distance to a reference profile. Compared together, the performances of the three neural methods are of the same higher of magnitude even though the direct approach seems to give better results. In the example presented in this work there is no advantage to work with one or another method (except for the computation time, which is smaller in the direct approach). The reason is that it is a well-posed problem with a quite regular function. Future work will be concerned with ill-posed problem (weakly influent parameters, non-univocity of the function,...).

# 5 CONCLUSION

Three neural methods to invert tomographic data have been proposed in this paper. All rely on the ability of Multilayered Perceptrons to approximate multidimensional functions. The inverse mapping between sound-velocity profiles and corresponding time patterns is learned by a neural net on the basis of a given set of behavior examples. Then, the performance in generalization (ability to retrieve sound velocity profiles from unknown time patterns) is used to perform inversions. The results show the efficiency of the neural approach which gives a very accurate

retrieval of the sound velocity profiles. Furthermore, the methods compare with favor to a usual linear schemes. As it stands, our approach is restricted to range-independent media and would have to be extended to all media. The main problem is an exponential increase of the number of parameters to be estimated and so an increase of the number of examples in the learning base. The neural approach may appear very useful to invert the unresolved part of the signal. This is particularly interesting in double SOFAR channel environments [14] where ray identification and estimation is difficult or even impossible.

## References


[1] Munk, W. H. and Wunsch, C. (1979) "Ocean acoustic tomography: a scheme for large scale monitoring", Deep-Sea Research, Vol. 26A, pp 123-161.

[2] Stephan, Y and Thiria, S. (1994) "Neural inversion for ocean acoustic tomography", in Inverse Problems in engineering Mechanics, Edited by Bui, Tanaka and al. Balkema Publishers, pp 55-60.

[3] Howe, B.M., Worcester, P.F. and Spindel, R.C. (1987) "Ocean acoustic tomography: mesoscale velocity", Journal of Geophysical research, Vol. 92, pp 3785-3805.

[4] Gaillard, F. (1992) "Evaluating the content of tomographic data: Application to mesoscale observations", Journal of Geophysical Research, Vol. 97, no C10, pp 15,490-15,505.

[5] Stephan, Y. and Martin-Lauzer, F.R. (1994) "GASTOM 90 tomography experiment. Data inversion", Proceedings of the Int. Conf. on Acoustics, Signal and Speech Processing, Vol 2, pp 301-304.

[6] Poggio, T. and Girosi, F. (1990) "Networks for approximation and learning", Proc. IEEE, Vol. 78,no 9.

[7] Rumelhart, D.E, Hinton, G.E. and Williams, R.J. (1986) "Learning representations by back-propagating errors", Nature, Vol 323, pp 533-536.

[8] Cybenko, G. (1989) "Approximation by superposition of sigmoidal function", Math. Control Signal Systems, Vol. 2, pp 303-314.

[9] Funahashi, K.I. (1989) "On the approximate realization of continuous mappings by neural networks", Neural Networks, Vol. 2, pp 183-192.

[10] Hornik, K., Stinchcomb, M., and White, H. (1989) "Multi-layer feedforward networks are universal approximators," Neural Networks, Vol. 2, pp 359-366.

[11] Linden, A. and Kinderman, J. (1989) "Inversion of multilayer nets", Proceedings of the International Joint Conference on Neural Networks, Vol II., pp 425-430.

[12] Jordan, M. and Rumelhart, D. (1992) "Forward models: supervised learning with a distal teacher", Cognitive Science, Vol. 16, pp 307-354.

[13] Stéphan, Y., Thiria, S. and Badran, F. (1995) "Inverting tomographic data with neural nets", Proceedings of OCEANS'95, San Diego.

[14] Stéphan, Y., Evennou, F. and Martin-Lauzer, F.R. (1995) "Acoustic Modeling and measurements in the Bay of Biscay in a double SOFAR channel", in Full Field Inversion Methods in Ocean and Seismo-Acoustics, O. Diachok et Al. (eds), Kluwer Academic Publishers.

| Method | error | c(0) | c(60) | c(300) |
|---|---|---|---|---|
| DIM | mean error (m/s) | 0.151 | 0.032 | 0.003 |
| | standart deviation | 0.113 | 0.021 | 0.002 |
| | relative mean error (p.c.) | 0.52 | 2.28 | 0.30 |
| AIM | mean error (m/s) | 0.184 | 0.040 | 0.002 |
| | standart deviation | 0.127 | 0.028 | 0.001 |
| | relative mean error (p.c.) | 0.83 | 2.86 | 0.20 |
| DLM | mean error (m/s) | 0.171 | 0.040 | 0.002 |
| | standart deviation | 0.126 | 0.029 | 0.001 |
| | relative mean error (p.c.) | 0.77 | 2.86 | 0.02 |
| LM | mean error (m/s) | 0.512 | 0.200 | 0.032 |
| | standart deviation | 0.372 | 0.144 | 0.022 |
| | relative mean error (p.c.) | 2.33 | 14.86 | 3.2 |

Table 1: Results of the Direct Inverse Method (DIM), the Adjoint Inverse Method (AIM), the DIstal Learning Method (DLM) and a Linear Method (LM) for 880 Mediterranean profiles.

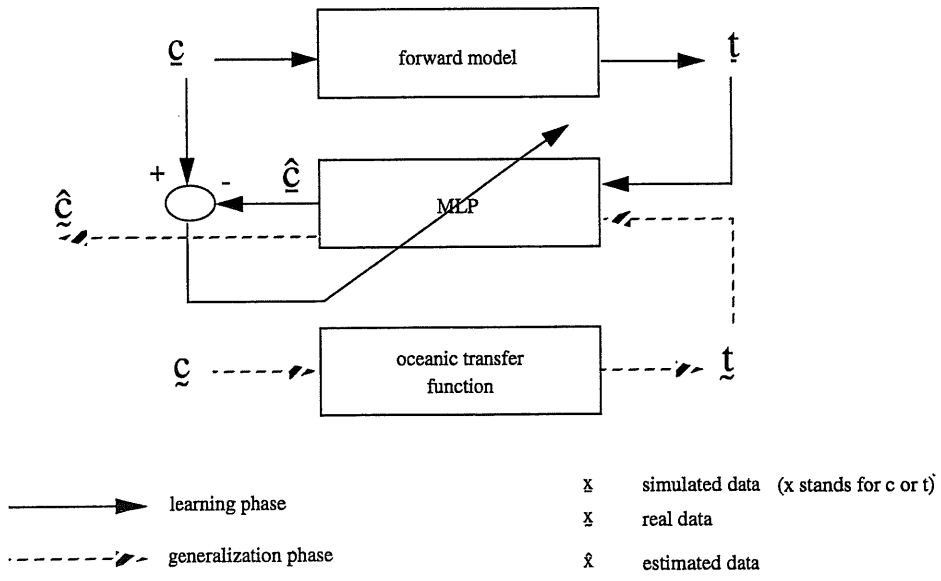



Figure 1: General scheme for neural direct inversion.

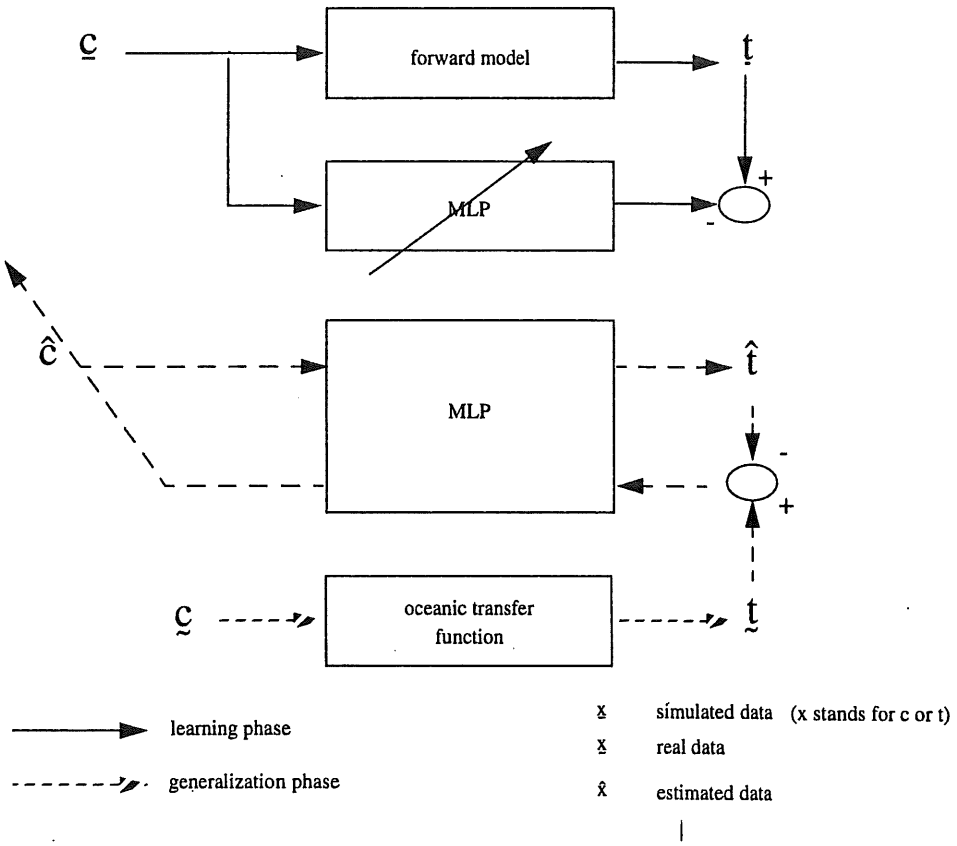



Figure 2: General scheme for neural adjoint inversion.

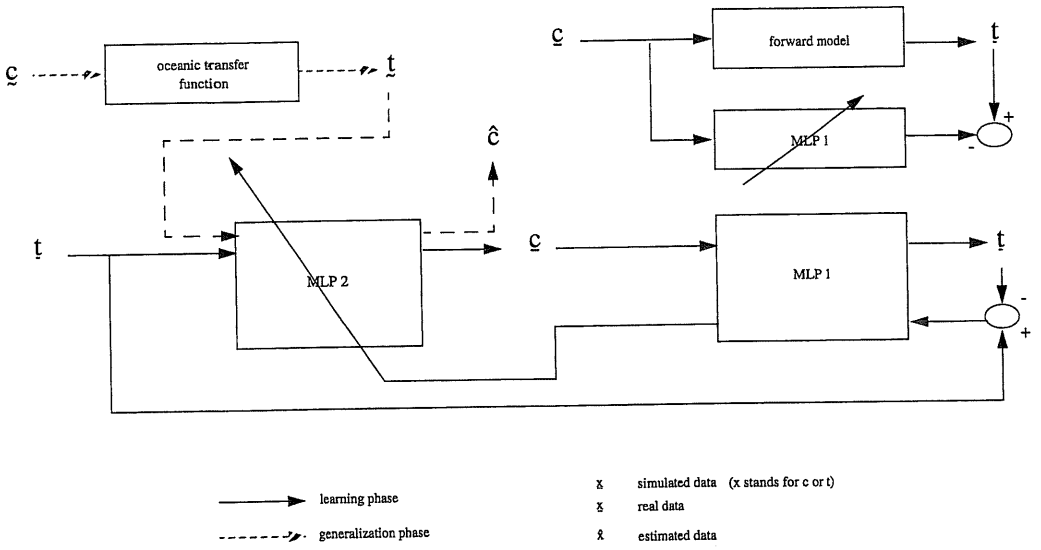



Figure 3: General scheme for neural distal inversion.

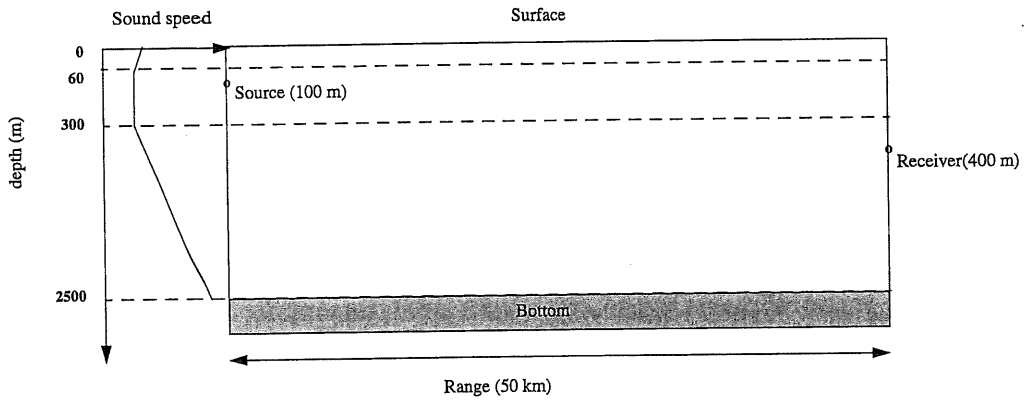



Figure 4: Design of the experiment for simulations.

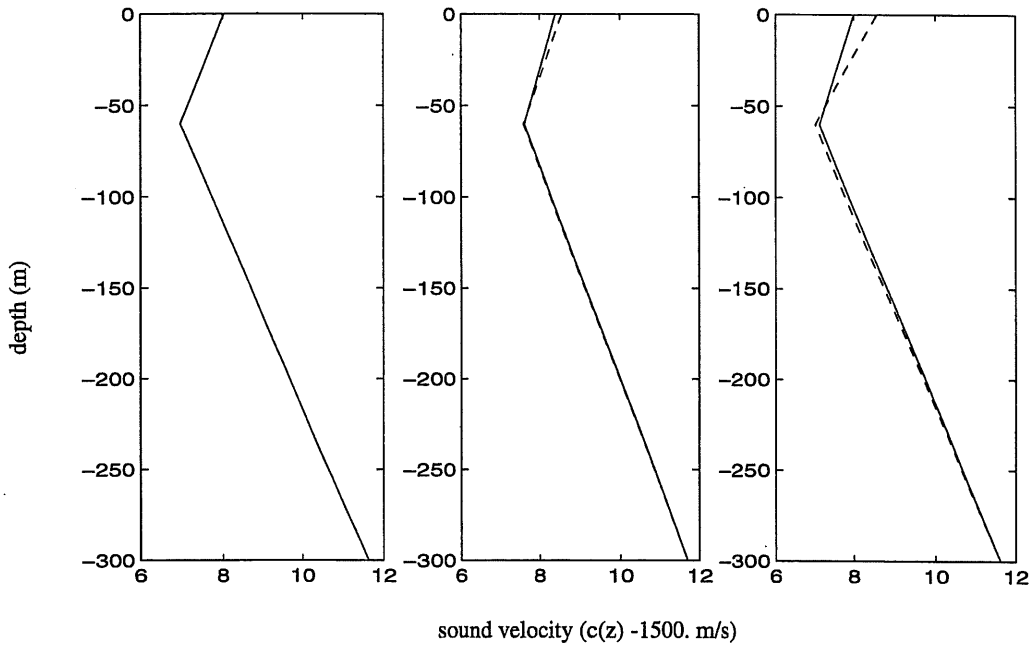



Figure 5: Results of the neural direct inversions on three characteristic profiles:
Left: Best approximated profile,
Middle: Inverted profile for which the error is closest to the mean error on the whole test set,
Right: Worse inverted profile.

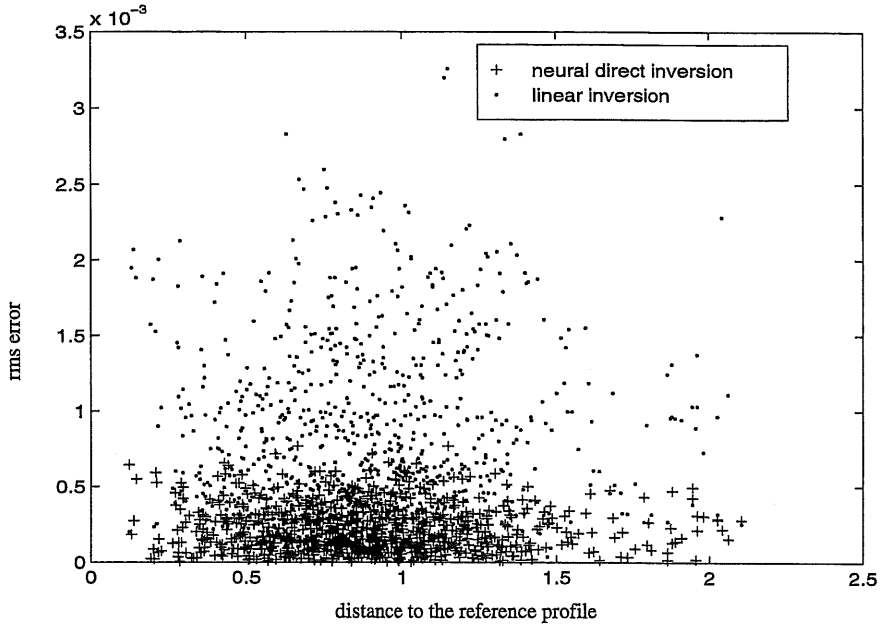



Figure 6: Comparisons between the neural Direct Inverse Method (DIM) and a Linear method(LM):
RMS error versus the distance to the reference profile (used for the linear scheme).

Theoretical and Computational Acoustics
D. Lee, Y-H Pao, M.H. Schultz, and Y-C Teng (Editors)


# On the Detection and Identification of Marine Mammals Using Artificial Neural Networks

Christopher M. DeAngelis, Kurt R. Almquist and Kelly J. Ross
NUWC Division Newport, Code 2253
Newport, Rhode Island 02841

**Abstract:** A technique for detecting and identifying marine mammals is presented. This technique utilizes a wavelet based preprocessor to extract features from a 1/3 of a second marine mammal sound clip. The dominant features are then used by an artificial neural network to identify the species of the marine mammal. In addition, a technique for unwrapping the internal mappings of the artificial neural network is presented. A combination of backpropagation, neural net weight decay and pruning are used to assist in forming a minimum neural network architecture. This resulting architecture can then be decomposed and analyzed, determining internal mappings, hidden node functionality, as well as input node importance and utilization. Furthermore, the actual features needed to determine and discriminate the classes can be identified as a function of classification accuracy. Results show that modest classification can be achieved with only a minimum number of features, and a higher degree of classification requires the non-linear combination of more features. The key is selecting the correct ones. The presented technique provides this ability.

## 1. INTRODUCTION

This research developed a technique for distinguishing between species of marine mammals. Initially, the technique uses a wavelet-based preprocessor for feature extraction, followed by an Artificial Neural Network (ANN) to perform the differentiation between species of marine mammals. It is planned to expand the ANN to recognize more classes of marine mammals in the future.

The overall goal of this research is to develop a technique for unwrapping and understanding the internal mappings of a backpropagation ANN. This will make it possible to validate the operations of the network, and determine the bounding envelopes under which it will correctly operate. By determining how the inputs are being combined and processed when inputs are partly obscured or completely degraded, it is possible to attain insight into the ANN's performance under severe operating conditions.

### 1.1 Bounds of Research

This research focuses on classification applications with boolean outputs in order to simplify the network decomposition process. For example, it is easier to determine which nodes contribute to a boolean output node turning *on* or *off*, than it is to determine the cause of a 5% increase in a continuous valued output.

To further simply the network decomposition process, the research began by examining only single output architectures, thus all of the hidden nodes contribute to the value of a single

output. If multiple outputs were processed by a single ANN, then each node could influence multiple outputs, and a scheme would have to be developed to separate and decouple the contribution of each node to each of the output nodes. Once a technique is developed for understanding mappings for the single boolean output case, it can be expanded to accommodate multiple, as well as continuous valued, outputs.

### 1.2 Technique Overview

To facilitate the generation of minimum sized neural networks, a number of algorithms were exploited. First, wavelets[1] are used to preprocess the sounds. This allows for accurate representation of the signals with a minimum number of terms. These terms are used as inputs to the neural based classifier. Next, a weight decay algorithm[2] was incorporated with a quasi-Newton weight update algorithm in the backpropagation ANN, encouraging unnecessary terms to fade to zero and enhancing error surface minimization. Finally, a network pruning algorithm[3] is used to remove nodes and interconnections that are not required for the mapping, with any minimal or redundant contributions redistributed to other active nodes. Collectively, these algorithms promote the reduction of a signal to a representative set of features that can then be combined by a minimally sized ANN to form the class mappings. The process of decomposing and interpreting the underlying mapping is greatly simplified by the use of a minimal-sized ANN architecture.

Since the sounds being classified contain little periodic information, a wavelet preprocessor was selected instead of the Fast Fourier transform (FFT) or the windowed FFT. Unlike the FFTs, wavelets more naturally capture the time-varying characteristics of transient signals, allowing the signal's energy to be accurately represented with a minimum number of coefficients[4,5].

Neural network based data processing systems have a number of favorable characteristics[6,7]. In contrast with many classical systems, they are adaptable, fast in execution, require no a priori statistics, can accommodate non-stationary data, and have the ability to approximate system functions that are not well definable or are even unknown[8,9]. These characteristics make them well suited for pattern completion and pattern matching applications[10,11].

If an Artificial Neural Network (ANN) can be reduced to a minimum architecture, then the weights of the resulting sparsely connected nodes can be more easily examined to determine node functionality. In this paper, a minimum architecture is defined as one providing acceptable performance with a minimum number of nodes and interconnections. Node functionality refers to the contribution each node is making to the overall mapping. By analyzing the connections from the input layer to the hidden layer, the importance and use of each input can be determined. The importance of an input node is reflected in the number of hidden nodes actually using that input. Determining the usage of an input is a bit more tenuous, as this information must be extracted from the weights assigned to each connection between the input and hidden nodes.

The weights of a network contain the information on how the inputs are combined to generate the desired output. An advantage of having a minimum architecture network is that all of the nodes are being used to their full capacity. One attribute of full capacity refers to a node having a dynamic output value which varies as the node's inputs change. Full capacity also means that the nodes have their own identity and are not mimicking or duplicating other nodes. When each node is used to its full capacity, one can determine whether it is adding to or detracting from the evidence that class "A" or class "B" exists by looking at the polarity of a hidden node's connections along with that of its bias term.

## 2. ANALYSIS

### 2.1 The Data Set

The data set being used consists of multiple samples of underwater acoustic signals. These signals have been sampled at 25KHz and each contain nearly a million data points. The four classes contained in these signals are: Carpenter Fish (also known as a Sperm Whale), Dolphin, Porpoise, and an underwater vehicle. It should be noted that in the case of the marine biological signals, a number of different animals were recorded for each species. It is hoped that this will promote the classification of the species as a whole and not just the individual members in each class. Additional data is needed to validate this premise since there is no indication in the data at hand of when each individual is speaking. Numerous individuals were present during each recording.

800 sound clips of each class were extracted from the recorded samples using a 8192 point window, resulting in approximately one-third of a second of information. Half of these clips were used for training the neural network and the other half for testing. The clips were not extracted in any particular order in that a random window placement was performed.

### 2.2 Preprocessing The Data Using A Wavelet Transform

Each of the sound clips was preprocessed using a Daubechies 4 wavelet. Due to the nature of the wavelet transform, the number of wavelet translations and dilations is equal to the length of the vector being analyzed which was $2^{13}$. Therefore, the 8192 point vector is translated into 8192 wavelet coefficients, the majority of which are zero. These coefficients are then normalized such that the largest is 1. When the coefficients were sorted from largest to smallest, a rapid exponential decay in the coefficient magnitudes was noticed, meaning that most of the signal's energy could be represented by a few coefficients. In order to retain the time and frequency information associated with each coefficient, an additional vector is created containing the respective wavelet matrix entry locations. The largest 20 coefficients along with their 20 respective time and frequency locating indexes (40 terms in all) are retained and used as inputs to the neural network.

### 2.3 Classification Of The Data With An Artificial Neural Network

In the first stage of this research two species were chosen to classify: porpoise and carpenter fish. The respective assigned class values are 0 (zero) and 1 (one), providing the single output architecture needed to simplify node analysis. The follow-on to this work will expand analysis to the multiple output case where each output will represent an individual class.

The backpropagation supervised learning neural paradigm[12] was used. This network "learns by doing" which means when the ANN is repeatedly presented with input-output training vectors, its internal connections adapt according to a given cost function until the correct response is given for a respective stimuli[13,14]. When analyzing the performance of this network, a classification threshold of 0.5 is used. Network output values greater than 0.5 are said to be in the 1 class and the others in the 0 class.

Since the sizing of the network is ad hoc, a single hidden layer of 10 nodes was chosen, resulting in a 40 input, 10 hidden, 1 output neural net architecture (40x10x1). This gave 421 degrees of freedom (the number of adjustable weights and biases) to separate the two classes. A number of networks were trained, each starting with a different set of random weights. All converged and correctly classified over 95% of the sounds for each class.

## 2.4 Neural Network Minimization

To aid in the ANN unwrapping process, several techniques were used to reduce the size of the ANN from the original 40x10x1 size. When analyzing the activation levels of the trained hidden nodes to various inputs, a great deal of duplication was discovered. The levels indicated that only 4 of the hidden nodes were actively contributing to the mapping. Thus, new ANNs with a 40x5x1 structure were then created. These networks were trained and gave similar, but inferior, classification results. Although there were still enough degrees of freedom, it was not surprising that the 40x5x1 networks did not perform as well as the 40x10x1 networks since fewer degrees of freedom means less of a chance to converge to a good solution.

An automated method to reduce the network's architecture was needed. The implemented minimization technique used a weight decay algorithm[2] incorporated into the quasi-Newton weight update scheme followed by an Optimum Brain Surgeon[15] based pruning algorithm[3]. The pruning algorithm was utilized to remove the non-needed nodes having constant activation levels, as well as nodes contributing little to the overall mapping. The contributions of any deleted nodes were redistributed over the remaining nodes. In addition to the nodes removed by the pruning algorithm, any of the remaining interconnections between nodes that do not contribute to the mapping are also removed. Some additional training is required after connections are pruned in order to refine the mapping.

Test set classification increased to 97% when this iterative process of decaying, pruning, and training was employed on a 40x10x1 network. When analyzing the remaining 102 interconnections, it was noticed that a considerable number of connections to the inputs corresponding to the wavelet coefficients were removed, leaving the majority of the connections to the time-frequency input nodes. This suggests that the time at which various frequencies are present is more important in separating the two classes than the magnitude of the frequencies. It was also noticed that the separation between the classes was within 0.1 of the extremes far from the 0.5 threshold value (i.e. the output of the network was less than 0.1 for all of class "0" and greater than 0.9 for all of class "1". This indicates that there is a strong separation between the species of marine mammals classified by the neural network.

## 2.5 Unwrapping The ANN Internal Mappings

Once the minimally sized network that provides the desired class separation was found, we can proceed to determine which inputs are important for classification, and how they are combined to provide a mapping for each class. The size and architecture of the network suggests the complexity of classification, with the number of connections to each input reflecting that input's importance and the polarity of the interconnections encouraging or inhibiting a hidden node's activation.

In order to determine which inputs are fundamental to class separation, one of the correct classification 100% networks (on the training set) was pruned until 50% correct classification was no longer achievable. Pruning was performed in stages with the number and location of pruned interconnections saved along the way. The smallest network that provided better than 50% classification contained 6 parameters (from 102), connecting 2 hidden nodes to 3 inputs. The resulting classification was better than 72% combining only 1 coefficient with 2 time-frequency indexes. More analysis needs to be conducted, but it appears that the output node is biased *on* with the two hidden nodes providing evidence to turn it *off*. Furthermore, the inputs need to be un-normalized and analyzed to determine the respective time-frequency magnitude and locations. This may suggest that in this two class application, a simple band-pass filter could be used to separate the classes; looking for energy in a specific frequency band.

Remember that more interconnections are needed to improve this classifiers performance, but only 72 of the 102 are needed for 92% classification. As connections are added back to the network, their polarity and what connecting nodes give insight into the rebuilding of the network's internal mapping. The functionality of the individual nodes should not change, only the accuracy in the mapping.

## 3. FUTURE WORK

Additional network analysis needs to be conducted in order to determine the behavior of the network (i.e. What each node is doing and how it contributes to the overall network mapping) which will be done by examining node activation levels for each class. This will allow the determination of which inputs are used for classifying each class, and which inputs are taken collectively to represent a class or discriminate between classes. Once this is done, these results will be used to determine what functionality the nodes are providing in terms of known functions, such as band-pass filters or differentiators.

As parameters are added back into the network we can continue to refine our notion of how inputs are being put together to represent or separate the classes: Which nodes are acting as band-pass, low-pass, high-pass, or feature-pass filters? Are there feature or class differentiating nodes? These questions can be answered as the network is built back up to the acceptable classification level.

## 4. CONCLUSIONS

A technique for identifying several classes of marine mammals has been developed along with a technique for unwrapping the internal mappings of a backpropagation neural network for classification applications. This involves the use of a wavelet transform to extract the most representative features from various classes of signals. The largest 20 wavelet coefficients and their respective time-frequency indexes are used as the inputs to a neural network classifier trained to separate the signals of interest.

An iterative process of training the neural network, decaying the weights, and pruning the non-important nodes and interconnections from the neural network was employed to form the smallest possible network providing the desired level of classification. In all cases, it was apparent that the wavelet indexes are more important in providing class separation than the wavelet coefficients themselves. In addition, as the network was pruned further it was determined that the output node was initially biased toward one class and the hidden nodes provided evidence suggesting the other.

Additional research needs to be conducted to compare the actions of hidden nodes to more understandable functions, such as band-pass filters. In addition, once the boolean output case is understood, it is hoped that this technique can be extended to multiple, continuous-valued outputs.

**Acknowledgment:** This research was sponsored by ONR (Dr. J. Davis) and is part of an ongoing effort at NUWC Division Newport, Newport, Rhode Island. The data and its wavelet transformation was provided by Quyen Huynh and Walter Greene, Weapons Technology and Undersea Systems Department at NUWC. The weight decay and pruning algorithms were provided by Robert Green, Computer and Information Science Department, University of Massachusetts Dartmouth. The principal investigator for the technique of decomposing and unwrapping the neural network is Christopher DeAngelis, Combat Control Systems Department at NUWC.

## REFERENCES

[1] C.K. Chui, *An Introduction to Wavelets*, Academic Press, San Diego, Ca, 1992.

[2] A. Krogh and J. Hertz, "A Simple Weight Decay Can Improve Generalization", Advances in Neural Information Processing 4, 950-957, Morgan-Kaufmann, 1992.

[3] D. Weigand, D. Rumelhart and B. Huberman, "Generalization by Weight-elimination with Application to Forecasting", Advances in Neural Information Processing 3, 875-882, Morgan-Kaufmann, 1991.

[4] B. Jawerth and W. Sweldens, "An Overview of Wavelet Based Multiresolution Analysis", SIAM Review 36, No. 3, 377-412, September, 1994.

[5] J. Benedetto and M. Frazier, Eds., *Wavelets: Mathematics and Applications*, CRC Press, Boca Raton, 1993.

[6] M. Minsky and S. Papert, "Perceptrons: An introduction to Computational Geometry," MIT Press, Cambridge, Massachusetts, 1969.

[7] R.P. Lippman, "An Introduction to Computing with Neural Networks," IEEE ASSP Magazine Vol. 4, No. 2, 4-22, April, 1987.

[8] W.T. Illingworth, "Beginners guide to neural networks," IEEE AES Magazine September, 44-49, 1989.

[9] S. Haykin, *Neural Networks: A Comprehensive Foundation*, IEEE Press, Macmillan College Publishing Company, New York, 1994.

[10] R.P. Lippman, "Pattern Classification Using Neural Networks," IEEE Communications Magazine, November, 47-64, 1989.

[11] E. Wilson, "Transient Detection and Feature Extraction Using Neural Networks," University of Rhode Island PhD Thesis, 1993.

[12] J. L. McClelland and Rumelhart D. E., *Explorations in Parallel Distributed Processing,* MIT Press, Cambridge, Massachusetts, 1988.

[13] D.E. Rumelhart, J.L. McClelland and the PDP Research Group, *Parallel Distributed Processing, Volume 1: Foundations* , MIT Press, Cambridge, Massachusetts, 1989.

[14] P. Werbos, "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences," Harvard University Ph.D. Thesis, dated 1974.

[15] B. Hassibi and D. Stork, "Second order derivatives for network pruning: Optimal Brain Surgeon", Advances in Neural Information Processing, Morgan-Kaufmann, 164-171, 1992.

Theoretical and Computational Acoustics
D. Lee, Y-H Pao, M.H. Schultz, and Y-C Teng (Editors)


# Analysis of Transient Acoustic Wavefields in Continuously Layered Media with Attenuation: An Approach Based on Symbolic Manipulation

Martin D. Verweij
Laboratory of Electromagnetic Research, Department of Electrical Engineering
Centre for Technical Geoscience, Delft University of Technology
P.O. Box 5031, 2600 GA Delft, The Netherlands

**Abstract:** An integral transformation-type method is presented for the analysis of the space-time domain acoustic wavefield – and the associated Green's function – in a continuously layered, lossy, isotropic (equivalent) fluid. Application of a vertically varying compliance memory function makes it possible to model a large class of depth-dependent loss properties. The method combines higher-order WKBJ-like asymptotics with the Cagniard-De Hoop method of inverse transformation. The coefficients that occur in the WKBJ asymptotics follow from a recurrence scheme that is easy to evaluate by means of a symbolic manipulation program. The form of the transform domain expressions leads to a very fast inversion process. Numerical results are presented for reflections from continuously layered halfspaces with depth-dependent losses.

## 1. INTRODUCTION

There are a number of areas where continuously layered fluids – fluids with medium properties that change in a continuous way in the vertical direction – may be applied for modeling purposes. For ocean acoustics and underwater acoustics, their usefulness is easily recognized. In seismics, continuously layered equivalent fluid models may be employed when only the slow vertical changes of the properties of the Earth are relevant – e.g., in computational backgrounds – while at the same time it suffices to represent the solid medium by an equivalent fluid. Unlike most fluids, the internal losses in many solids are significant. For seismic applications it is therefore important that the loss behavior of a viscoelastic solid is carried over to its equivalent fluid representation. The compliance memory function (Boltzmann [1], Ben-Menahem and Singh [2]) is well suited for modeling the losses in fluids, since it can directly be incorporated in the basic acoustic equations (Verweij [3]).

For the analysis of the acoustic wavefield in media with temporal and/or spatial invariances, various integral transformation-type methods exist (Chapman and Orcutt [4]). In this paper we will develop an integral transformation-type method for the determination of the space-time domain acoustic wavefield in a continuously layered, isotropic fluid with depth-dependent losses. Its main ingredients are: a temporal Laplace transformation – with a real and positive transform parameter – followed by horizontal spatial Fourier transformations; higher-order WKBJ-like asymptotic approximations of the transform domain solution in inverse powers of the Laplace parameter; and the Cagniard-De Hoop method of inverse transformation. Since the asymptotic representations are valid for large values of the Laplace parameter, the resulting space-time domain approximations will be most accurate near the arrival time. The occurrence

of the same exponential function in all terms of the WKBJ-like asymptotic representations will enable us to organize the inverse transformation in a very efficient way. By applying higher-order asymptotics our method will often yield more accurate results – especially somewhat away from the arrival time – than with the more common zero-order or first-order asymptotic methods. As always, there will also be another side of the coin. Due to the nature of the applied asymptotics, for any finite value of the Laplace transform parameter there exists a certain order beyond which the accuracy of the transform domain asymptotic representations will no longer improve. In the space-time domain this may lead to divergence of the higher-order approximations after a configuration-dependent time instant, i.e., the method may loose its significance for later time instants. When the first arrival at the point of observation is associated with a turning ray, it will not be accounted for by the exponential function in the WKBJ-like asymptotics, and the method can not be applied.

## 2. CONFIGURATION AND BASIC ACOUSTIC EQUATIONS

In this paper we will determine, for a known point source at $x_i^S$ in a known continuously layered, lossy fluid, the acoustic wavefield that is present at a point of observation at $x_i^{\rm obs}$. We assume that the wavefield in this medium satisfies the linearized basic acoustic equations

$$\partial_k v_k(x_i,t) + \kappa(x_3)\,\partial_t p(x_i,t) + a(x_3,t) * p(x_i,t) = q(x_i,t), \tag{1}$$

$$\partial_k p(x_i,t) + \rho(x_3)\,\partial_t v_k(x_i,t) + b(x_3,t) * v_k(x_i,t) = f_k(x_i,t). \tag{2}$$

Here we have introduced the subscript notation and the summation convention, where the lower-case Latin subscripts (except $t$) range from 1 to 3. The state quantities of the acoustic wavefield are the particle velocity $v_k$ and the acoustic pressure $p$. The symbols $\partial_k$ and $\partial_t$ denote a differentiation with respect to $x_k$ and $t$, respectively. The symbol $*$ indicates a temporal convolution. The source action is described by the volume density of volume injection rate $q$ and the volume density of volume force $f_k$. Without loss of generality we may assume that

$$q(x_i,t) = Q^S(t)\,\delta(x_1,x_2,x_3-x_3^S), \tag{3}$$

$$f_k(x_i,t) = F_k^S(t)\,\delta(x_1,x_2,x_3-x_3^S), \tag{4}$$

with $Q^S(t) = 0$ and $F_k^S(t) = 0$ for $t < 0$. The compressibility $\kappa(x_3)$ and the mass density $\rho(x_3)$ represent the instantaneous reaction of the medium, and the memory functions $a(x_3,t)$ and $b(x_3,t)$ represent the non-instantaneous reaction of the medium, i.e., the losses. For the after-effects in the compliance, the relation $a(x_3,t) = \partial_t^2\phi(x_3,t)$ holds, where $\phi(x_3,t)$ is the (reduced) creep function. We omit memory effects in the inertia behavior and set $b(x_3,t) = 0$, as is often done in the literature. Further we assume that $\kappa(x_3)$, $\rho(x_3)$ and $a(x_3,t)$ are continuous functions of $x_3$ that are as often differentiable as required by our analysis.

## 3. FORWARD TRANSFORMATION

We first subject the space-time domain quantities to the temporal Laplace transformation

$$\hat{p}(x_i,s) = \int_{0-}^{\infty} p(x_i,t)\,\exp(-st)\,{\rm d}t. \tag{5}$$

Reversely, for a given function $\hat{p}(x_i,s)$, the solution $p(x_i,t)$ of this integral equation is unique and causal if $\hat{p}(x_i,s)$ is bounded for the infinite set of points $s_n = s_0 + n\ell$ with $s_0$ real, positive and sufficiently large, $n = 0,1,2,\cdots$, and $\ell$ positive and real (Lerch's theorem, see Widder [5]).

In the Cagniard-De Hoop method only these values of $s$ play a role. We further apply the double horizontal spatial Fourier transformation of the Radon type

$$\tilde{p}(\alpha_1, \alpha_2, x_3, s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \hat{p}(x_i, s) \exp[is(\alpha_1 x_1 + \alpha_2 x_2)] \, \mathrm{d}x_1 \, \mathrm{d}x_2. \tag{6}$$

The inverse of this Fourier transformation is

$$\hat{p}(x_i, s) = \left(\frac{s}{2\pi}\right)^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{p}(\alpha_1, \alpha_2, x_3, s) \exp[-is(\alpha_1 x_1 + \alpha_2 x_2)] \, \mathrm{d}\alpha_1 \, \mathrm{d}\alpha_2. \tag{7}$$

To keep the expressions readable, from now on we will omit most arguments of the functions.

Application of the forward transformations to Eqs. (1) and (2), followed by the elimination of $\tilde{v}_1$ and $\tilde{v}_2$, results in the transform domain state vector differential equation

$$\partial_3 \tilde{b}_I + sA_{IJ}(x_3) \, \tilde{b}_J = K_{IJ}(x_3, s) \, \tilde{b}_J + \tilde{u}_I. \tag{8}$$

The upper case Latin subscripts range from 1 to 2. Here, $\tilde{b}_I = (\tilde{v}_3, \tilde{p})^{\mathsf{T}}$ is the transform domain state vector, and $\tilde{u}_I = \delta(x_3 - x_3^S)(\hat{Q}^N, \hat{F}^N)^{\mathsf{T}}$ is a notional source vector with

$$\begin{pmatrix} \hat{Q}^N \\ \hat{F}^N \end{pmatrix} = \begin{pmatrix} \hat{Q}^S + i(\alpha_1 \hat{F}_1^S + \alpha_2 \hat{F}_2^S)/\rho(x_3) \\ \hat{F}_3^S \end{pmatrix}. \tag{9}$$

Further,

$$A_{IJ}(x_3) = \begin{pmatrix} 0 & \gamma(x_3) Y(x_3) \\ \gamma(x_3)/Y(x_3) & 0 \end{pmatrix} \tag{10}$$

forms the system matrix, in which $\gamma(x_3) = [c(x_3)^{-2} + \alpha_1^2 + \alpha_2^2]^{1/2}$ is the vertical slowness – with $c(x_3) = [\rho(x_3) \, \kappa(x_3)]^{-1/2}$ being the acoustic wavespeed – and $Y(x_3) = \gamma(x_3)/\rho(x_3)$ is the vertical acoustic wave admittance. The fact that $\gamma(x_3)$ is real and greater than zero for any of the values of $\alpha_1$ and $\alpha_2$ that occur in the inverse Fourier transformation, is important for the validity of the WKBJ-like asymptotics that will be introduced in the next section. Moreover,

$$K_{IJ}(x_3, s) = \begin{pmatrix} 0 & -\hat{a}(x_3, s) \\ 0 & 0 \end{pmatrix} \tag{11}$$

is the memory matrix representing the losses.

## 4. WKBJ-LIKE ASYMPTOTIC REPRESENTATIONS

As the first step in finding approximate solutions of the state vector differential equation, we introduce the wavevector $\tilde{w}_I$ through the composition relation $\tilde{b}_I = N_{IJ}(x_3) \, \tilde{w}_J$, where

$$N_{IJ}(x_3) = \tfrac{1}{2}\sqrt{2} \begin{pmatrix} Y^{1/2}(x_3) & -Y^{1/2}(x_3) \\ Y^{-1/2}(x_3) & Y^{-1/2}(x_3) \end{pmatrix}. \tag{12}$$

The wavevector components $\tilde{w}_1$ and $\tilde{w}_2$ represent waves that travel in the positive and the negative $x_3$-direction, respectively. Substitution of the composition relation in Eq. (8) yields the wavevector differential equation, which we subsequently recast into the corresponding integral equation (cf. Verweij [3], [6])

$$\tilde{w}_I = \mathsf{L}_{IJ} \, \tilde{w}_J + \tilde{h}_I. \tag{13}$$

The term

$$\mathsf{L}_{IJ}\,\tilde{w}_J = \begin{pmatrix} \displaystyle\int_{-\infty}^{x_3} [\Theta_{11}(x_3',s)\,\tilde{w}_1(x_3') + \Theta_{12}(x_3',s)\,\tilde{w}_2(x_3')]\,\exp(-s\textstyle\int_{x_3'}^{x_3}\gamma\,\mathrm{d}\zeta)\,\mathrm{d}x_3' \\ \\ \displaystyle -\int_{x_3}^{\infty} [\Theta_{21}(x_3',s)\,\tilde{w}_1(x_3') + \Theta_{22}(x_3',s)\,\tilde{w}_2(x_3')]\,\exp(-s\textstyle\int_{x_3'}^{x_3}\gamma\,\mathrm{d}\zeta)\,\mathrm{d}x_3' \end{pmatrix}, \tag{14}$$

with $\Theta_{IJ}(x_3,s) = [(-1)^I\,\hat{a}(x_3,s) + (1-\delta_{IJ})\,\partial_3 Y(x_3)]/2Y(x_3)$, stems from the losses and the inhomogeneity of the medium. The term

$$\tilde{h}_I = \begin{pmatrix} \frac{1}{2}\sqrt{2}\,\tilde{a}_1\,H(x_3 - x_3^S)\exp(-s\textstyle\int_{x_3}^{x_3^S}\gamma\,\mathrm{d}\zeta) \\ \\ -\frac{1}{2}\sqrt{2}\,\tilde{a}_2\,H(x_3^S - x_3)\exp(-s\textstyle\int_{x_3^S}^{x_3}\gamma\,\mathrm{d}\zeta) \end{pmatrix}, \tag{15}$$

with $\tilde{a}_i = \hat{F}_3^S\,Y^{1/2}(x_3^S) - (-1)^i\hat{Q}^S\,Y^{-1/2}(x_3^S)$, originates from the source.

Now we have to solve the wavevector integral equation. Repeated substitution of Eq. (13) into itself yields, under very general conditions, a convergent Neumann series solution. For a general inhomogeneous, lossy medium, the analytical and numerical evaluation of the higher-order terms of this series will be virtually impossible. However, repeated integration by parts of the terms of the Neumann series solution will result in WKBJ-like asymptotic expansions of the wavevector in inverse powers of $s$ (cf. Verweij [6]). This has inspired us to approximate the wavevector by

$$\tilde{w}_I \;\sim\; \exp(-s\textstyle\int_{x_3^S}^{x_3}\gamma\,\mathrm{d}\zeta)\displaystyle\sum_{n=0}^{N} s^{-n}\,P_I^{(n)}(x_3,s), \qquad (x_3 > x_3^S), \tag{16}$$

$$\tilde{w}_I \;\sim\; \exp(-s\textstyle\int_{x_3}^{x_3^S}\gamma\,\mathrm{d}\zeta)\displaystyle\sum_{n=0}^{N} s^{-n}\,Q_I^{(n)}(x_3,s), \qquad (x_3 < x_3^S). \tag{17}$$

These are $N$th-order WKBJ-like asymptotic representations of the wavevector for large values of $s$. The exponential parts are supposed to give the propagation of the wavefront from the source to the point of observation, and the summation parts describe the behavior of the wavefield at the point of observation after the arrival of the wavefront. The evaluation of the coefficients $P_1^{(n)}$, $P_2^{(n)}$, $Q_1^{(n)}$ and $Q_2^{(n)}$ forms the subject of the next section.

## 5. RECURRENCE SCHEME FOR THE COEFFICIENTS

Substitution of Eqs. (16) and (17) into Eq. (13), followed by differentiation of the equations for $\tilde{w}_2$ with respect to $x_3$ and collection of the terms with equal powers of $s$, results in an *implicit* (i.e., with the unknown quantities being defined in terms of themselves) recurrence scheme for the components of $P_I^{(n)}$ and $Q_I^{(n)}$. We find for $n = 0$

$$P_1^{(0)} \;=\; \tfrac{1}{2}\sqrt{2}\tilde{a}_1 + \int_{x_3^S}^{x_3} \Theta_{11}\,P_1^{(0)}\,\mathrm{d}x_3', \tag{18}$$

$$P_2^{(0)} \;=\; 0, \tag{19}$$

$$Q_1^{(0)} \;=\; 0, \tag{20}$$

$$Q_2^{(0)} \;=\; -\tfrac{1}{2}\sqrt{2}\tilde{a}_2 - \int_{x_3}^{x_3^S} \Theta_{22}\,Q_2^{(0)}\,\mathrm{d}x_3', \tag{21}$$

and for $n \geq 1$

$$P_1^{(n)} = \int_{x_3^S}^{x_3} [\Theta_{11} P_1^{(n)} + \Theta_{12} P_2^{(n)}] \, dx_3' + Q_1^{(n)}(x_3^S), \tag{22}$$

$$P_2^{(n)} = \frac{-1}{2\gamma} [\Theta_{21} P_1^{(n-1)} + \Theta_{22} P_2^{(n-1)} - \partial_3 P_2^{(n-1)}], \tag{23}$$

$$Q_1^{(n)} = \frac{1}{2\gamma} [\Theta_{11} Q_1^{(n-1)} + \Theta_{12} Q_2^{(n-1)} - \partial_3 Q_1^{(n-1)}], \tag{24}$$

$$Q_2^{(n)} = \int_{x_3}^{x_3^S} [\Theta_{21} Q_1^{(n)} + \Theta_{22} Q_2^{(n)}] \, dx_3' + P_2^{(n)}(x_3^S). \tag{25}$$

We have maintained the exact functions $\Theta_{IJ}(x_3, s)$ in this scheme because in our case it will be difficult to obtain accurate approximations of the functions $\Theta_{IJ}(x_3, s)$ in inverse powers of $s$ within a reasonably low number of terms. This will be the case since we will employ memory functions with characteristic times that are small on the time scale that will be used for the wavefield quantities. Since the summation parts in Eqs. (16) and (17) will not only contain inverse powers of $s$ – as with WKBJ asymptotics – but also the function $\hat{a}(x_3, s)$ and/or its derivatives $\partial_3^k \hat{a}(x_3, s)$, we have used the notion WKBJ-*like* asymptotics in this paper.

To evaluate the implicit recurrence scheme, we assume that on the closed interval between $x_3^S$ and $x_3^{\text{obs}}$ the functions $\Theta_{IJ}(x_3, s)$, $1/2\gamma(x_3)$ and $\Theta_{IJ}(x_3, s)/2\gamma(x_3)$ may be approximated by the following Taylor polynomials

$$\Theta_{IJ}(x_3, s) \approx \sum_{m=0}^{M} \Theta_{IJ}^{(m)}(s) \, (x_3 - x_3^S)^m, \tag{26}$$

$$\frac{1}{2\gamma(x_3)} \approx \sum_{m=0}^{M} \gamma^{(m)} \, (x_3 - x_3^S)^m, \tag{27}$$

$$\frac{\Theta_{IJ}(x_3, s)}{2\gamma(x_3)} \approx \sum_{m=0}^{M} D_{IJ}^{(m)}(s) \, (x_3 - x_3^S)^m. \tag{28}$$

Now our goal is to find approximate expressions for $P_I^{(n)}$ and $Q_I^{(n)}$ in the form of Taylor polynomials like

$$P_I^{(n)}(x_3, s) = \sum_{\ell=0}^{L} P_I^{(n,\ell)}(s) \, (x_3 - x_3^S)^\ell. \tag{29}$$

Substitution of Eqs. (26) - (29) into Eqs. (18) - (25) and collection of the terms with equal powers of $(x_3 - x_3^S)$ leads to an *explicit* (i.e., with the unknown quantities being defined only in terms of known quantities) recurrence scheme for the coefficients $P_1^{(n,\ell)}$, $P_2^{(n,\ell)}$, $Q_1^{(n,\ell)}$ and $Q_2^{(n,\ell)}$. For $n = 0$ we obtain

$$P_1^{(0,\ell)} = \left\{ \begin{array}{ll} \frac{1}{2}\sqrt{2}\, \tilde{a}_1, & (\ell = 0) \\ \\ \frac{1}{\ell} \sum_{m=0}^{\ell-1} \Theta_{11}^{(m)} P_1^{(0,\ell-m-1)}, & (\ell \geq 1) \end{array} \right\}, \tag{30}$$

$$P_2^{(0,\ell)} = 0, \qquad (\ell \geq 0), \tag{31}$$

$$Q_1^{(0,\ell)} = 0, \qquad (\ell \geq 0), \tag{32}$$

$$Q_2^{(0,\ell)} = \left\{ \begin{array}{ll} -\frac{1}{2}\sqrt{2}\,\tilde{a}_2, & (\ell = 0) \\ \frac{-1}{\ell} \sum_{m=0}^{\ell-1} \Theta_{22}^{(m)}\, Q_2^{(0,\ell-m-1)}, & (\ell \geq 1) \end{array} \right\}, \tag{33}$$

and for $n \geq 1$ we get

$$P_1^{(n,\ell)} = \left\{ \begin{array}{ll} Q_1^{(n,0)}, & (\ell = 0) \\ \frac{1}{\ell} \sum_{m=0}^{\ell-1} \Theta_{11}^{(m)}\, P_1^{(n,\ell-m-1)} + \Theta_{12}^{(m)}\, P_2^{(n,\ell-m-1)}, & (\ell \geq 1) \end{array} \right\}, \tag{34}$$

$$P_2^{(n,\ell)} = \sum_{m=0}^{\ell} [-D_{21}^{(m)}\, P_1^{(n-1,\ell-m)} - D_{22}^{(m)}\, P_2^{(n-1,\ell-m)} + (\ell - m + 1)\,\gamma^{(m)}\, P_2^{(n-1,\ell-m+1)}], \qquad (\ell \geq 0), \tag{35}$$

$$Q_1^{(n,\ell)} = \sum_{m=0}^{\ell} [D_{11}^{(m)}\, Q_1^{(n-1,\ell-m)} + D_{12}^{(m)}\, Q_2^{(n-1,\ell-m)} - (\ell - m + 1)\,\gamma^{(m)}\, Q_1^{(n-1,\ell-m+1)}], \qquad (\ell \geq 0), \tag{36}$$

$$Q_2^{(n,\ell)} = \left\{ \begin{array}{ll} P_2^{(n,0)}, & (\ell = 0) \\ \frac{-1}{\ell} \sum_{m=0}^{\ell-1} \Theta_{21}^{(m)}\, Q_1^{(n,\ell-m-1)} + \Theta_{22}^{(m)}\, Q_2^{(n,\ell-m-1)}, & (\ell \geq 1) \end{array} \right\}. \tag{37}$$

It is easy to evaluate this recurrence scheme using a symbolic manipulation program.

## 6. INVERSE TRANSFORMATION

To elucidate the application of the Cagniard-De Hoop method, as an example we will show here how to find the space-time domain acoustic pressure below a source of volume injection. First we remark that the coefficients $\Theta_{IJ}^{(m)}$, $\gamma^{(m)}$ and $D_{IJ}^{(m)}$ consist of terms that contain integer inverse powers of $\gamma(x_3^S)$. In many of these terms the memory function $\hat{a}(x_3^S, s)$ and/or its derivatives $\partial_3^k \hat{a}(x_3^S, s)$ occur as well. Further, in the present case we find that $\tilde{a}_1 = -\tilde{a}_2 = Y^{-1/2}(x_3^S)\,\hat{Q}^S$. From Eqs. (30) - (37) it then follows that $P_1^{(n,\ell)}$ and $P_2^{(n,\ell)}$ are composed of terms of the form $C_1\,\hat{Q}^S\,\hat{\Pi}(s)\,Y^{-1/2}(x_3^S)\,\gamma^{-\eta}(x_3^S)$. Here $C_1$ indicates a real constant, the function $\hat{\Pi}(s)$ stands for a certain product of the memory function and/or several of its derivatives, and $\eta$ is a nonnegative integer. Upon applying the composition relation and Eqs. (16) and (29), we find that the terms of $\tilde{p}$ are of the form

$$\tfrac{1}{2} C_2\,\hat{Q}^S\,(x_3^{\rm obs} - x_3^S)^\ell\, s^{-n}\,\hat{\Pi}(s)\,Y^{-1/2}(x_3^S)\,Y^{-1/2}(x_3^{\rm obs})\,\gamma^{-\eta}(x_3^S)\,\exp(-s\textstyle\int_{x_3^S}^{x_3^{\rm obs}} \gamma\,{\rm d}\zeta), \tag{38}$$

in which $C_2$ is a real constant and $\ell$ and $n$ are nonnegative integers. At this stage it is convenient to write $\tilde{p} = s^2 \hat{Q}^S \tilde{G}$. The Green's function $\tilde{G}$ is the acoustic pressure that is generated by a

source with a unit ramp signature $Q^S$. This Green's function is composed of a number of terms of the form $C_3\, s^{-n}\, \hat{\Pi}(s)\, \tilde{\Upsilon}(\eta)$, with

$$\tilde{\Upsilon}(\eta) = \tfrac{1}{2} s^{-2}\, Y^{-1/2}(x_3^S)\, Y^{-1/2}(x_3^{\rm obs}) \gamma^{-\eta}(x_3^S) \exp(-s \int_{x_3^S}^{x_3^{\rm obs}} \gamma \, {\rm d}\zeta). \tag{39}$$

The constant $C_3$ contains all products $C_2(x_3^{\rm obs} - x_3^S)^\ell$ of the terms of the form (38) with corresponding $n$, $\hat{\Pi}(s)$ and $\eta$. When higher-order asymptotics are involved, the task of finding the constants $C_3$ and each of the possible combinations of $n$, $\hat{\Pi}(s)$ and $\eta$, is almost impossible to carry out by hand, but forms no problem for a symbolic manipulation program.

The function $\tilde{\Upsilon}(\eta)$ that shows up in each term of $\tilde{G}$ is the only factor that depends on $s$ as well as on $\alpha_1$ and $\alpha_2$, and only for the inversion of this part we need to employ the Cagniard-De Hoop method (see Verweij and De Hoop [7] for a description of this method in case of continuously layered media). In all cases in which our WKBJ-like asymptotics are useful, the result of this process is

$$\Upsilon(\eta) = \frac{1}{2\pi^2} H(t - T_{\rm arr}) \int_0^{Q(t)} {\rm Im} \left\{ \bar{Y}^{-1/2}(x_3^S) \bar{Y}^{-1/2}(x_3^{\rm obs}) \bar{\gamma}^{-\eta}(x_3^S)\, \partial_t p \right\} {\rm d}q, \tag{40}$$

Here, $p$ denotes a (complex) horizontal slowness that satisfies

$$t = pr + \int_{x_3^S}^{x_3^{\rm obs}} \bar{\gamma}(\zeta) \, {\rm d}\zeta = {\rm real}. \tag{41}$$

In both foregoing equations the bar indicates that the vertical slowness has been changed into $\bar{\gamma}(x_3) = [c^{-2}(x_3) - p^2 + q^2]^{1/2}$. The points $p$ that are relevant in view of the Cagniard-De Hoop method form a so-called Cagniard contour in the complex $p$-plane. The value of $t$ associated with the point where a Cagniard contour crosses the real $p$-axis is indicated by $T(q)$; the unique inverse of this function for $q \geq 0$ is denoted by $Q(t)$. The quantity $T_{\rm arr} = T(0)$ is the arrival time of the wave front.

The other factors in the terms of $\tilde{G}$ solely depend on $s$. The inversion of $s^{-n}$ is obtained by inspection. The time domain function $\Pi(t)$ is recognized as the convolution of the relevant time domain memory functions $a(x_3^S, t)$ and/or its derivatives $\partial_3^k a(x_3^S, t)$, which are known from the start. Convolution of the space-time domain counterparts of the factors of each term of $\tilde{G}$ yields the corresponding contribution to the space-time domain Green's function $G$. Once this function has been found, the space-time domain acoustic pressure follows as $p = \partial_t^2 [Q^S * G]$.

## 7. NUMERICAL RESULTS

In this section we will present results for the space-time domain acoustic wavefield in a simplified version of a marine seismic configuration. The quantity of interest is the acoustic pressure, and the point source is of the volume injection type. The configuration consist of a homogeneous, lossless upper halfspace (water; the effects of the surface are discarded) and an inhomogeneous and/or lossy lower halfspace (subsurface geology). This type of configuration leads to an analysis that is somewhat simpler than the one presented in the previous sections [we may replace the Taylor polynomials in Eqs. (26) - (29) by Taylor polynomials in $(x_3 - x_3^{\rm int})$, and it suffices to set $M = N - 1$ in Eqs. (26) - (28) and $L = 0$ in Eq. (29)]. The acoustic wave is generated by a source with $x_3^S = 50$ m, it reflects from the interface between the halfspaces at $x_3^{\rm int} = 100$ m, and it is observed at a receiver with $x_3^{\rm obs} = 0$ m and with the same horizontal position as the source. The medium parameters are given in Fig. 1. To describe the losses of

the equivalent fluid in the lower halfspace, we introduce the causal creep function

$$\phi(x_3,t) = \kappa(x_3)\Delta(x_3)[\gamma_{\text{euler}} + \mathrm{E}_1(\omega_0 t) + \ln(\omega_0 t)]\, H(t), \tag{42}$$

in which $\Delta(x_3) = 2/\pi Q_0(x_3)$. The symbol $\mathrm{E}_1$ denotes the exponential integral. The amount of loss at a certain level $x_3$ is inversely proportial to the value of $Q_0(x_3)$. For angular frequencies lower than $\omega_0$, the value of the quality factor $Q$ of the medium is almost frequency independent and approximates $Q_0(x_3)$. This frequency behavior of $Q$ makes that the applied creep function is well-suited for modeling the losses occurring in many types of rock. Further, a source signature is applied for which it suffices to take $\omega_0 = 200$ rad/s.

Several approximate Green's functions are depicted in Figure 2, in which the trivial step-function contribution of the direct wave has been omitted. The value of $N$ indicates the order of the asymptotic representation. We see that before $t = 1.5$ s (unshaded region) the differences between the subsequent higher-order approximations decrease: for $N \geq 3$ they are even indistinguishable. We assume that as soon as the latter is the case, the exact Green's function is arrived at. However, beyond $t = 1.5$ s (shaded region) the subsequent approximations diverge. The fact that useful results are generated only up till a specific time instant is a consequence of the application of WKBJ-like asymptotics and determines the applicability of the method in each specific case. When there are turning rays traveling below the interface that arrive earlier than the rays reflecting at the interface, the first arrival will not be accounted for by the exponential function in the asymptotics. If this happens, divergence will occur for any time instant.

We can define two other media by either assuming the losses to be absent in the lower halfspace ($Q_0 = \infty$), or by taking the values of the wavespeed and the mass density in the lower halfspace equal to those of the upper halfspace. The Green's functions for the media thus defined are given in Fig. 3. We observe that near the arrival time the effect due to the inhomogeneity of the wavespeed and mass density is opposite to the effect due to the losses. This also follows from Fig. 4, which shows the reflected acoustic pressures in the three media for a source with a four-point optimum Blackmann signature of 0.1 s duration and unit amplitude.

## 8. CONCLUSIONS

In this paper we have derived a method for the determination of the space-time domain acoustic wavefield – and the corresponding Green's function – in a continuously layered, lossy, isotropic fluid or equivalent fluid. By using a depth dependent compliance memory function, intricate losses may be modeled. The form of the transform domain WKBJ-like asymptotic representations has enabled us to employ the Cagniard-De Hoop method in a very efficient way. We have tackled the problem of performing tasks that are almost impossible to carry out by hand, like the generation of the coefficients of the asymptotic representations, by invoking a symbolic manipulation program.

Typical numerical results have been generated for acoustic wavefields that are reflected by continuously layered halfspaces with a depth-dependent loss behavior. For a zero horizontal offset we have observed that there is an interval, beginning with the arrival time, on which the differences between the approximate Green's functions with increasing orders become invisibly small. In general, for later time instants the subsequent approximations diverge. The method is applicable for nonzero horizontal offsets, but becomes invalid when the first arrival at the point of observation is associated with a turning ray. We have shown numerically that near the arrival time the effect of the losses can counteract the effect of the inhomogeneity of the wavespeed and mass density.

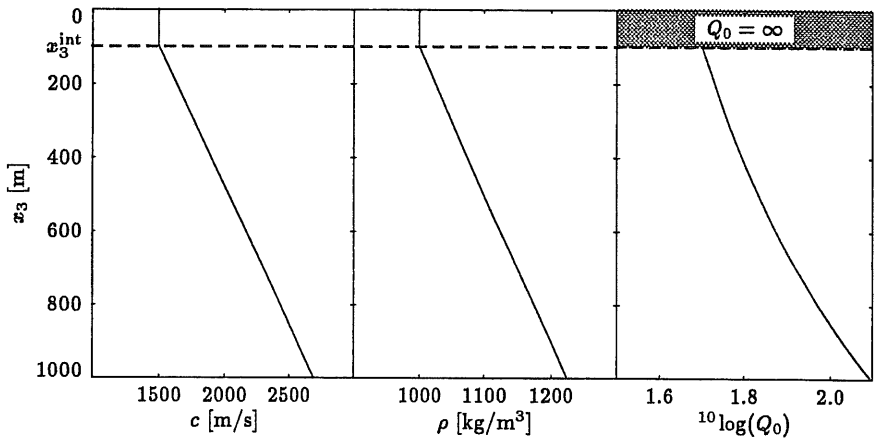



Figure 1: The depth dependence of the wavespeed $c$, the mass density $\rho$, and the parameter $Q_0$, as used in the example configuration.

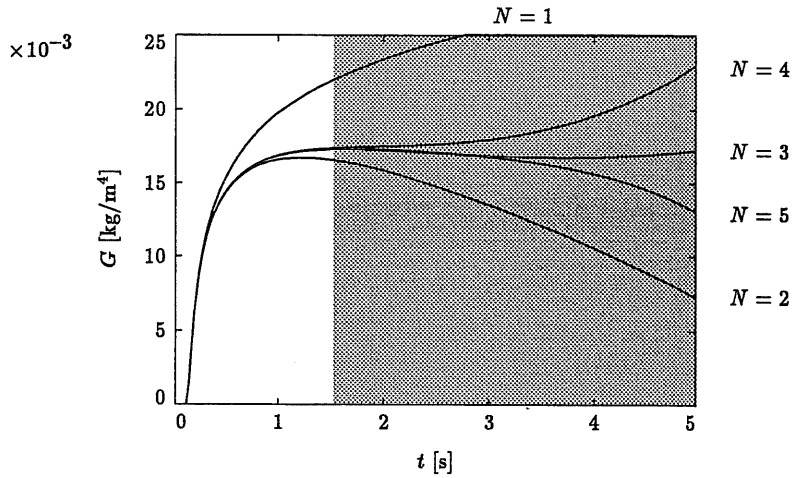



Figure 2: Several approximate Green's functions $G$. For $t \leq 1.5$ s the differences between the subsequent higher-order approximations decrease (unshaded region), while for $t > 1.5$ s the subsequent approximations diverge (shaded region).

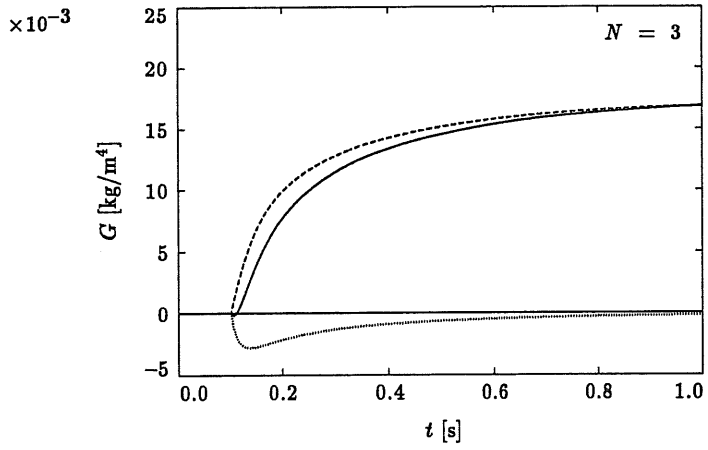



Figure 3: The Green's functions $G$ for the original medium (solid line), a medium with the same wavespeed and mass density as the original medium but without losses (dashed line), and a medium with the same losses as the original medium but with the constant wavespeed $c = 1500$ m/s and the constant mass density $\rho = 1000$ kg/m$^3$ (dotted line).

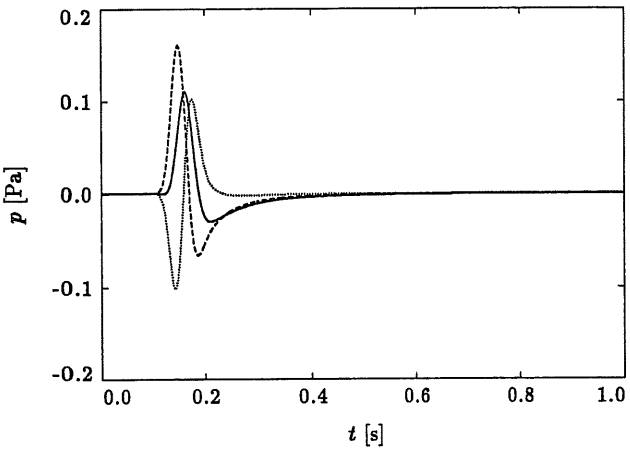



Figure 4: The acoustic pressures $p$ for the original medium (solid line), a medium with the same wavespeed and mass density as the original medium but without losses (dashed line), and a medium with the same losses as the original medium but with the constant wavespeed $c = 1500$ m/s and the constant mass density $\rho = 1000$ kg/m$^3$ (dotted line).

## ACKNOWLEDGMENTS

The research of the author has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences. The research reported in this paper has been financially supported through research grants from the Stichting Fund for Science, Technology and Research (a companion organization to the Schlumberger Foundation in the U.S.A.).

## REFERENCES

[1] Boltzmann, L., Zur Theorie der elastischen Nachwirkung, Sitzber. Kaiserl. Akad. Wiss. Wien, Math.-Naturw. Kl., vol. 70(II), 1874, pp. 275-306.

[2] Ben-Menahem, A., and Singh, S. J., Seismic waves and sources, Springer-Verlag New York, Inc., New York, 1981.

[3] Verweij, M. D., Modeling space-time domain acoustic wave fields in media with attenuation: The symbolic manipulation approach, J. Acoust. Soc. Am., vol. 97, 1995, pp. 831-843.

[4] Chapman, C. H., and Orcutt, J. A., The computation of body wave synthetic seismograms in laterally homogeneous media, Rev. Geophys., vol. 23(2), 1985, pp. 105-163.

[5] Widder, D. V., The Laplace transform, Princeton Univ. Press, Princeton, 1946.

[6] Verweij, M. D., Transient acoustic wave modeling: Higher-order Wentzel-Kramers-Brillouin-Jeffreys asymptotics and symbolic manipulation, J. Acoust. Soc. Am., vol. 92, 1992, pp. 2223-2238.

[7] Verweij, M. D., and De Hoop, A. T., Determination of seismic wavefields in arbitrarily continuously layered media using the modified Cagniard method, Geophys. J. Int., vol. 108, 1990, pp. 731-754.

Theoretical and Computational Acoustics
D. Lee, Y-H Pao, M.H. Schultz, and Y-C Teng (Editors)


# A Simulation Study of the Variability of Acoustic Transmissions from Hawaii to Monterey

Robert A. Staten, Ching-Sang Chiu, Albert J. Semtner
Department of Oceanography, Naval Postgraduate School
Monterey, CA 93943, USA

**Abstract**

As part of the modeling effort in the Acoustic Thermometry of Ocean Climate (ATOC) project, the influence of mesoscale, gyrescale and seasonal ocean variability on three-dimensional (3D) acoustic ray paths from the planned Hawaii to Monterey site is investigated. Ray paths and signal arrival structure over a two year period are simulated at a six-day interval using a 3D ray-based acoustic model. The input sound speed fields are interpolated from gridded (1/4 degree, 20 level) temperature and salinity output data from the Semtner-Chervin eddy-resolving Parallel Ocean Climate Model. Based on the simulated acoustic multipath arrival structure, the issues of stability and travel time variability are addressed. Bottom topography in the Moonless Mountains region is found to have occasional mild effects on steeper rays. Arrival structure is found to be strongly dependent on depth and quasi-stable over time with a 0.6 correlation between arrival patterns at different times. Travel time variability estimates are 0.42 s rms for steep rays, due to a combination of seasonal and mesoscale ocean variability, and 0.28 s rms for near-axial rays, due primarily to seasonal variability.

## 1 INTRODUCTION

We have observed the concentrations of "greenhouse gases" increasing at a significant rate, yet debate still exists over several key issues:

1. the amount of global warming that will result from the build-up of these gases,
2. the length of time over which global warming will take place, and
3. the discernability and predictability of warming trends, given our current data, networks and models.

Difficulties arise in trying to determine warming trends from atmospheric measurements, as they have similar spatial patterns to the background air temperature variability. In addition, many land stations are contaminated by the "urban heat-center" effect and are further biased by the northerly distribution of most observation stations. [1]

Munk and Forbes proposed "Underwater Acoustic Thermometry" in 1989 as a method for the early detection of global warming. Using a network of projectors and receivers distributed worldwide, acoustic signal travel time variability would be measured over a ten year period and theoretically yield a quantitative estimate of global warming. An advantage to utilizing the ocean rather than the atmosphere to detect warming trends is that the warming signal is significantly different than the pattern of inherent variability of the ocean. [1]

In 1991 the Heard Island Feasibility Test (HIFT) sought to answer questions concerning

transmission and reception of trans-global acoustic signals. Coded acoustic transmissions originating from Heard Island in the southern Indian Ocean were monitored at 16 sites in the Atlantic, Pacific, Indian, and Southern Oceans. The successful test established that low-frequency sound from modulated sources could be transmitted reliably over great distances with sufficient intensity.

The effects of background ocean variability on acoustic variability for the Heard-to-California path was studied by Chiu *et al.* [2] utilizing a similar global ocean circulation model to that used for this study. An important conclusion was that excessive background travel time fluctuation of 2.5 s rms would preclude detection of the greenhouse signal in ten years along this 18-megameter path, since the rate of travel time increase due to potential warming was estimated by Munk and Forbes [1] to be 1 s per decade along trans-global paths. This significant travel time fluctuation was caused by the highly variable Antarctic Circumpolar Front, through which the Heard-to-California path traversed.

The HIFT was prerequisite to the current Scripps Institution of Oceanography's program, Acoustic Thermometry of Ocean Climate (ATOC). Avoiding transmissions through the Antarctic Polar Front, ATOC calls for a planned network of sites located throughout the Pacific, including two acoustic sources to be deployed near California and Hawaii. Receiver sites include existing U.S. Navy facilities, several fixed Vertical Line Arrays (VLA's), and drifting Surface Suspended Acoustic Receivers (SSAR's).

In addition to measuring ocean climate change, another benefit of a cross-basin acoustic thermometry system would be enhancement of ocean nowcast/forecast capability. Basin specific information on gyre scale and mesoscale processes could be obtained by travel time fluctuations along paths through specific basins. Travel time variability information can be used to validate/calibrate ocean nowcast/forecast models.

There are three acoustic issues that are of great concern to and must be addressed by ATOC. The first is acoustic reliability. It must be known whether a site will be ensonified all year round. The location of the sound source determines locations of ensonified and shadow zones. These zones fluctuate as a function of time due to the variability of the ocean.

The next issue is acoustic stability. It must be determined whether significant portions of the acoustic arrival patterns change over an extended period of time. Only stable arrivals can be used to construct useful travel-time series.

The final issue deals with geophysical noise. Frontal effects, mesoscale eddies and seasonal cycles cause travel time fluctuations that can be interpreted as "noise" relative to the greenhouse signal. The size of these fluctuations dictates the length of observation needed before a statistically significant trend can be revealed. Different acoustic paths traverse different ocean features of varying strengths. Consequently, arrival time fluctuations vary from location to location. The most ideal receiving locations are those which are reliable, with stable arrivals, and those having minimal geophysical noise in the travel times.

This paper focuses on the Hawaii-to-California path, examining the effects of mesoscale, gyrescale and seasonal variability and single-phone arrival structure. Issues of acoustic reliability and stability are addressed and travel time variability is quantified.

The analysis of acoustic variability was conducted by interfacing an acoustic model with a circulation model. A three-dimensional raytracing computer program called the Hamiltonian Acoustic Raytracing Program for the Ocean (HARPO) was used for the acoustic model. The

Semtner-Chervin Parallel Ocean Climate Model (POCM) served as the ocean circulation model, providing simulated salinity and temperature fields. HARPO requires a continuous sound-speed field as input and an interpolation procedure was used to interface HARPO with the ocean model. Fans of acoustic rays and the corresponding parameters were computed for every six days in a two year model period. For every six-day "ocean snapshot" eigenrays were identified and arrival structures were constructed.

Time series of arrival structure for two different receiver depths were analyzed. The depth of 700 m corresponds to the approximate depth of the sound channel axis off the coast of California, and the 1359 m depth corresponds to the approximate depth of a single hydrophone of an existing array located near Point Sur.

The remainder of this paper consists of three sections. Section 2 contains a description of the approach, including brief descriptions of HARPO, the Semtner-Chervin model, and the ocean-acoustic interface procedure. Also included is a description of the arrival structure generation procedure. In Section 3, results are presented and analyzed. Section 4 presents the conclusions and recommendations of this paper.

## 2 DESCRIPTION OF APPROACH

### 2.1 HAMILTONIAN ACOUSTIC RAYTRACING

The acoustic model used to trace raypaths is the Hamiltonian Acoustic Raytracing Program for the Ocean (HARPO). Originating from NOAA [3], HARPO is a computer program that numerically integrates Hamilton's equations of motion in spherical polar coordinates for three-dimensional acoustic ray paths through a model ocean. The input ocean model is required to have a continuous three-dimensional sound speed field with a continuous representation of both the sea surface and bathymetry. An updated version of HARPO that contains an interpolating interface to gridded data [2] was used for this study to perform all raytracing on a workstation.

The initial conditions for the integrations consist of vertical and azimuthal launch angles of the rays. The effect of the Earth's curvature is accounted for since HARPO performs calculations in polar coordinates, but geodesic veering due to the elliptical shape of the Earth is not accounted for.

### 2.2 OCEAN MODEL

The ocean environment model used is the Semtner-Chervin Parallel Ocean Climate Model (POCM) [4], a descendent of the eddy-resolving Semtner-Chervin Global Ocean Circulation Model (SCGCM) [5, 6]. Improvements can be attributed in part to the model's ability to harness the parallel processors of the CRAY Y-MP/8 supercomputer at the National Center for Atmospheric Research (NCAR). The model is gridded with 1/4° average horizontal resolution and has a domain extending from 75°S to 65°N. In the vertical plane there are 20 levels with an approximate resolution of 25 m in the upper 100 m, with decreasing resolution with increasing depth. Model ocean depths were obtained from NCAR for 1/12° latitude and longitude points and averaged to correspond to local model resolution. Coastlines were also prescribed at the

local model resolution and not smoothed and simplified as in the earlier model.

The uppermost levels are forced on a monthly basis to values from the Levitus climatological data set. Momentum fluxes are derived from smoothly interpolated monthly values of European Center for Medium Range Weather Forecasting (ECMWF) wind stress at 1000 mb for the period of 1985-1989. Subsurface restoring of temperature and salinity is only used north of 58°S, south of 68°N, and in the vicinity of Gibraltar.

Initialization of the model was accomplished by using the full 33 year spin-up of the SCGCM to generate instantaneous fields, which were then interpolated to the 1/4° grid spacing. The model was resumed, starting with the ECMWF winds of 1985. At the end of each 3-day period, a "snapshot" was taken of the global ocean and archived at NCAR. For this thesis every other archived data set was used for years 1987 and 1988 resulting in a six-day resolution over a period of two years.

## 2.3 OCEAN-ACOUSTIC INTERFACE

HARPO requires continuous data fields as its input for computations. An earlier procedure to interface HARPO to gridded ocean model data [7] was modified by Chiu *et al.* in 1994 [2] and used for this interface. This method makes computations possible on local workstations with limited hard disk space and memory. Since the global data sets archived at NCAR are very large, it was desirable to subsample only the data pertaining to our region of interest.

Temperature, salinity, and bathymetry data were extracted for the specific region from the global data set. Sound speed was then calculated from temperature and salinity data using the Mackenzie empirical formula [8]. To limit the computational demand, a 5° by 5° window was used to move along with the traced ray. Inside of the window sound speed was expressed as a sum of the mean speed and perturbed speed. The perturbation was further expressed as a linear combination of empirical orthogonal functions (EOF's). All perturbation EOF's, mean profiles, and bathymetry data were splined to produce smooth and continuous input fields for HARPO. After each ray was traced to a point near the edge of the window, the window was shifted to center at that point and new sound speed and bathymetry data were subsampled and processed, and raytracing continued.

## 2.4 MULTIPATH ARRIVAL STRUCTURE GENERATION

HARPO does not perform eigenray searches and compute signal amplitudes and phase shifts. An external MATLAB™ program called *ray3d.m* was developed to perform these functions based on the HARPO outputs. The program first determines the location where each wavefront (i.e., surface normal to a ray) intercepts the receiver as a function of launch angle by using a local tangent-normal analysis. As sketched in Figure 1, HARPO outputs the position vectors as well as local ray angles of points along each traced ray (i.e., ray points). These HARPO outputs are used to calculate displacement vectors $\vec{d}$ pointing to the receiver and unit directional vectors $\hat{t}$ tangent to the ray at those ray points. As the wavefront of the ray crosses the receiver, a sign change in $\vec{d} \cdot \hat{t}$ occurs and thus the "cross-segment" is determined. The exact interception point within the ray cross-segment is found by fitting a polynomial to the positions

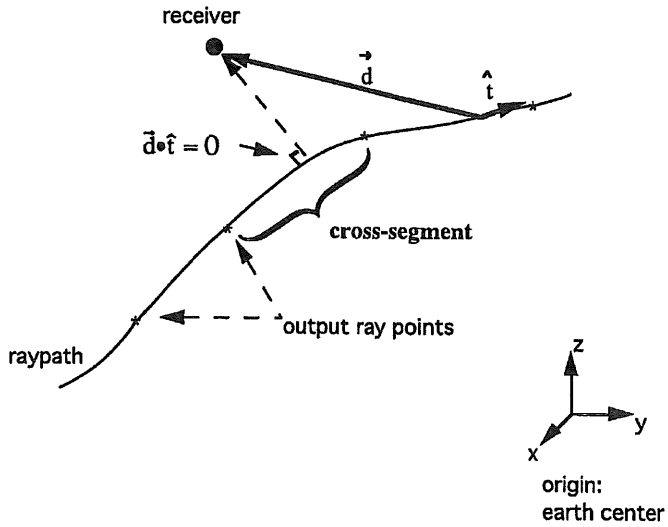



**Figure 1.** Determination of the position along a ray where the corresponding wavefront crosses the receiver using a local tangent-normal coordinate analysis. $\vec{d}$ is a displacement vector defining the distance and direction from any position on the ray to the receiver and $\hat{t}$ is a tangent directional vector along the ray.

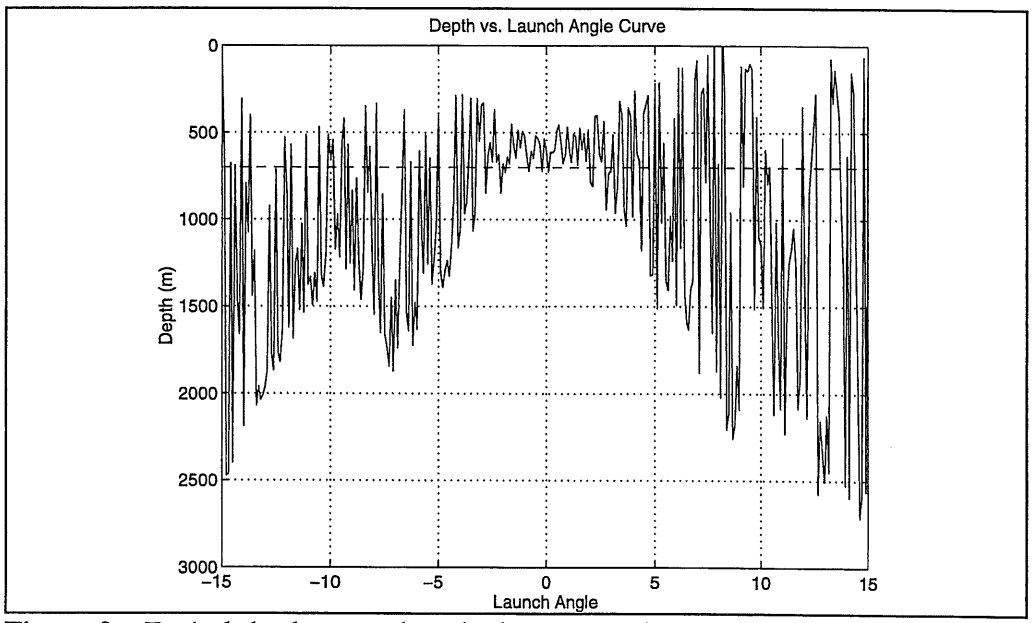



**Figure 2.** Typical depth-versus-launch elevation angle curve (solid line). A receiver depth at 700 m is also plotted (dashed line). The intersections give the launch elevation angles of the eigenrays.

and ray angles of the two local ray points and then solving for $\vec{d} \bullet \hat{t} = 0$.

Based on the calculated interception points of the traced ray fan, a depth-versus-launch elevation angle curve can be constructed. As shown in Figure 2, the points where such a curve intersects the receiver depth line gives the launch angles of the eigenrays. For these intersection points, eigenray parameters are interpolated for, including initial angles, arrival times, number of turning points, etc.

Eigenrays with five or more bottom or surface interactions are not included and would not be considered reliable ray paths. For each individual eigenray arrival, spreading loss, boundary loss and phase changes are accounted for. Then the time-translated, phase-shifted and amplitude-modified signal arrivals are coherently summed.

## 3 ANALYSIS AND RESULTS

### 3.1 ENVIRONMENT

#### 3.1.1 Fronts/Circulation

The general pattern of circulation in the North Pacific has a concentrated north-eastward current flow in the western half with a widely dispersed southern return flow in the eastern half. The most notable circulation feature is the sub-tropical front, a region of convergence located just north of Hawaii and stretching east-west in direction. In this region warmer surface water tends to be forced downward by the converging circulation at the center of the Pacific gyre. This feature is well represented in the Semtner-Chervin POCM and can be observed in both the horizontal and vertical slices of the sound speed fields (to be displayed in upcoming figures). The deep sound channel axis does not vary a great deal, ranging from 850 m near Hawaii to about 700 m near the coast of California.

#### 3.1.2 Bathymetry

In comparison to other parts of the Pacific basin the topography of the northeastern region tends to be somewhat smooth; however, sections underlying the transmission paths can occasionally provide an obstacle (i.e., a complication) to some steeper acoustic rays.

Several fracture zones cover the eastern Pacific basin and they are characterized by extremely rough and elevated terrain relative to the surrounding area. As can be seen in Figure 3, the Murray Fracture Zone stretches east-west for several thousand kilometers, crossing the trajectory of the raypaths approximately half way. In this particular region the Moonless Mountains rise to within nearly 2000 m of the sea surface, thereby affecting in some way deeper cycling raypaths.

### 3.2 OCEAN MODEL VARIABILITY

#### 3.2.1 Spatial Variability

Figure 4 displays a horizontal section of an instantaneous sound speed field simulated by the

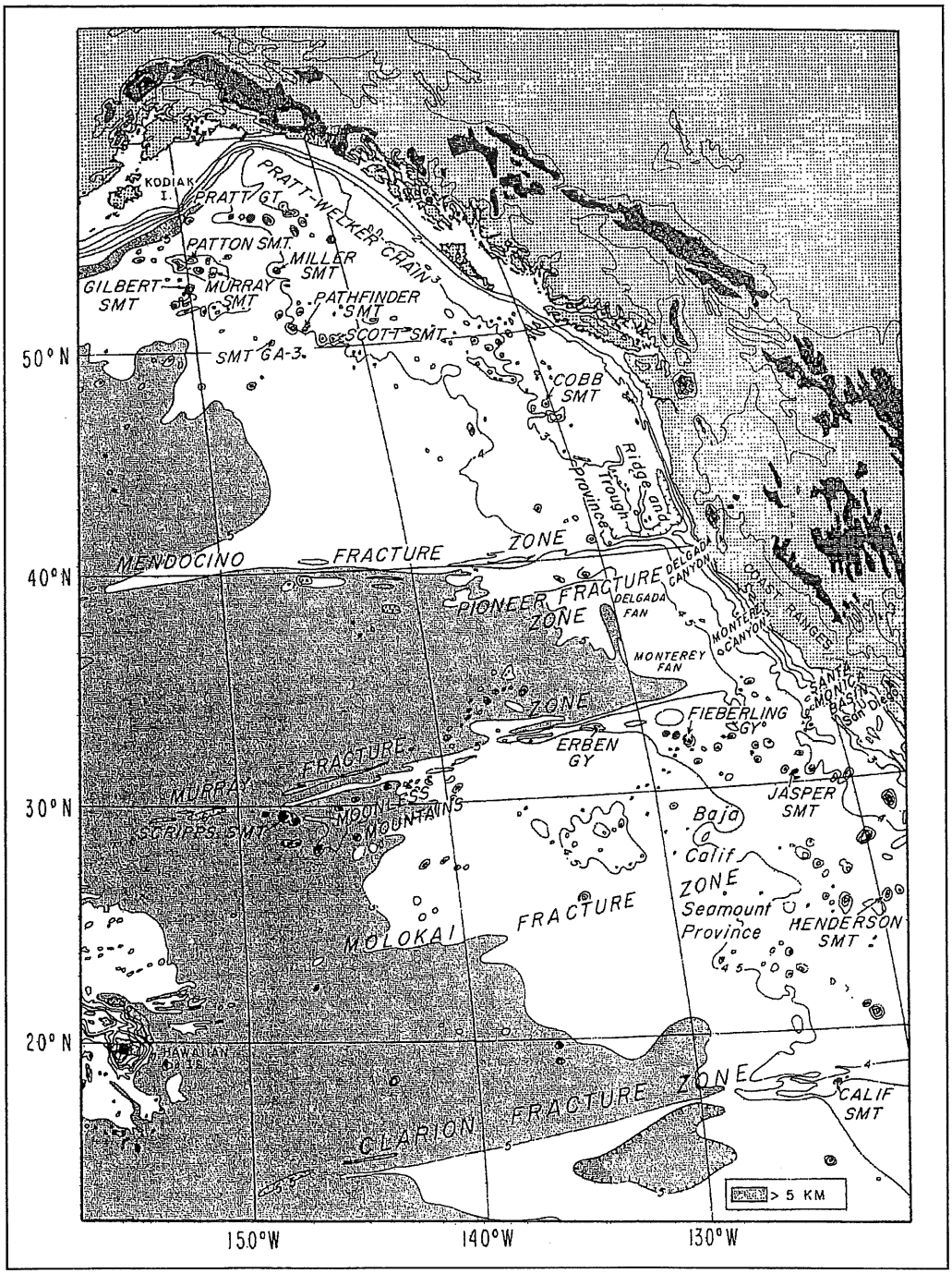



**Figure 3.** Bathymetry of the Pacific Basin (from reference [9]).

POCM at a depth of 435 m. Substantial horizontal variability exists between Hawaii and California at this level, and approximate locations of significant mesoscale features including the subtropical front can be discerned. In contrast, Figure 5 displays sound speed at 847 m, with noticeably reduced variability (note color axis scaling).

Figure 6 presents a vertical slice of an instantaneous sound speed field along a path between the source and the receivers. Vertical gradients in sound speed are significant in the upper 500 m of the water column, especially in the vicinity of the subtropical front. Below 500 m changes are relatively gradual.

### 3.2.2 Temporal Variability

In Figure 7 the standard deviation of sound speed in a vertical slice along the transmission path, calculated from the POCM for the two year period of this study, is shown. A region of increased variability which appears to be associated with mesoscale phenomenon lies just beyond the 1500 km range and seems to correspond to the average position of the subtropical front. Again note the zone of decreased variability below approximately 500 m.

## 3.3 RAYPATH ANALYSIS

As indicated in Section 1, the examination of three acoustic issues were the thrust of this simulation analysis: reliability, stability and travel time variability.

### 3.3.1 Acoustic Reliability

Rays were traced for a single azimuth in a vertical fan for launch angles between -15° (downward) and +15° (upward) per six-day time step. Out of these rays an average of 110 eigenrays per time slice were found for the deep receiver and 131 eigenrays were found for the axial receiver. The reduced number of eigenrays for the deep receiver is expected due to its decreased proximity to the sound channel axis.

Figure 8 presents a typical vertical geometry of a near-axial raypath. Generally the width of the vertical envelope for small-angle rays is no more than 300-400 m and, more importantly, the depth of the ray trajectory remains well below the area of greatest variability in the water column (see Figure 7).

A typical deep cycling ray path is shown in Figure 9. The width of the vertical envelope is far greater than the near axial, increasing to over 2000 m. It is important to note that the upper turning points are located within the region of high variability within the water column. The ray in this figure has been impacted by a portion of the Moonless Mountains. Depending on the angle of incidence on the seamount, some rays could be reflected into much steeper paths that have numerous surface and bottom interactions. These significant boundary interactions and losses would render them unreliable. However, some other impacted rays, such as the one shown in Figure 9, were merely reflected into slightly steeper paths that do not encounter further boundary interactions beyond the seamount. These alterations in ray trajectories were rare in occurrence and only rays with launch elevation angles of magnitude 12° and higher would be altered.

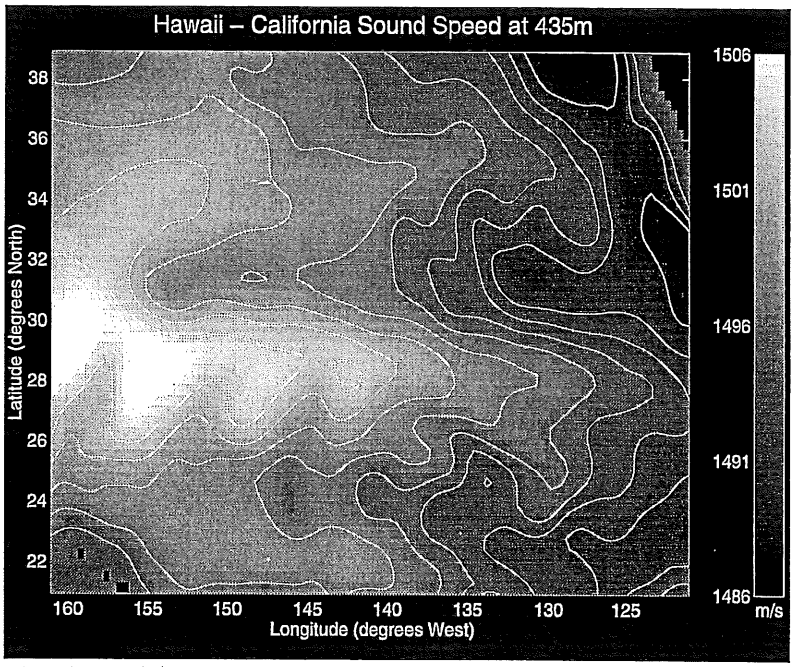



**Figure 4.** Horizontal layer of sound speed at a depth of 435 m between Hawaii and California, a snapshot. Contour interval is 2 m/s.

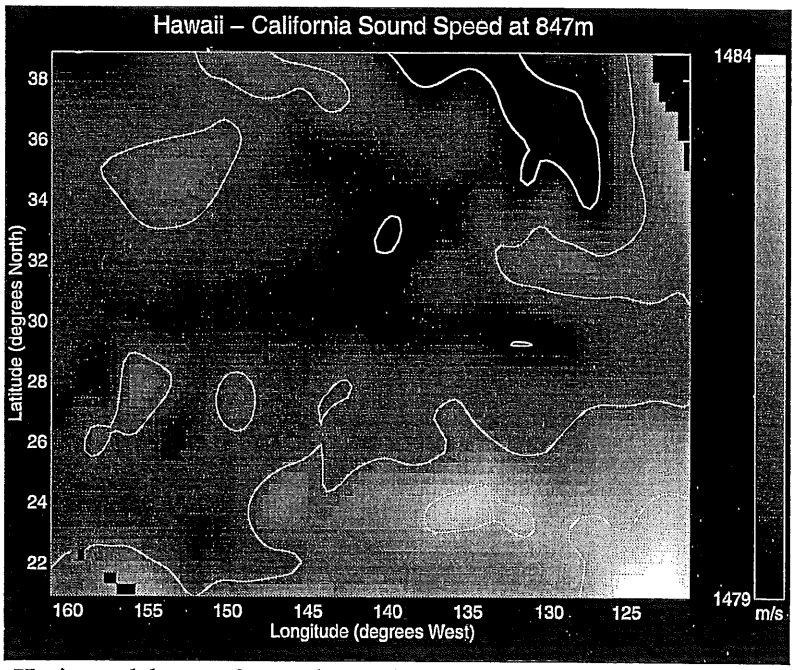



**Figure 5.** Horizontal layer of sound speed at a depth of 847 m between Hawaii and California, a snapshot. Contour interval is 2 m/s.

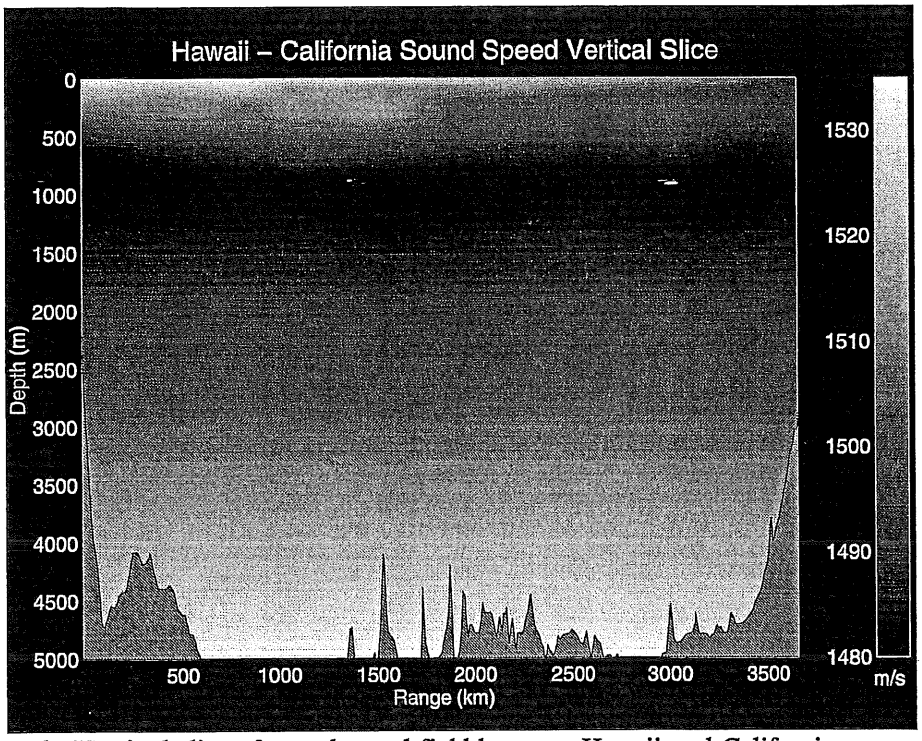



**Figure 6.** Vertical slice of sound speed field between Hawaii and California, a snapshot.

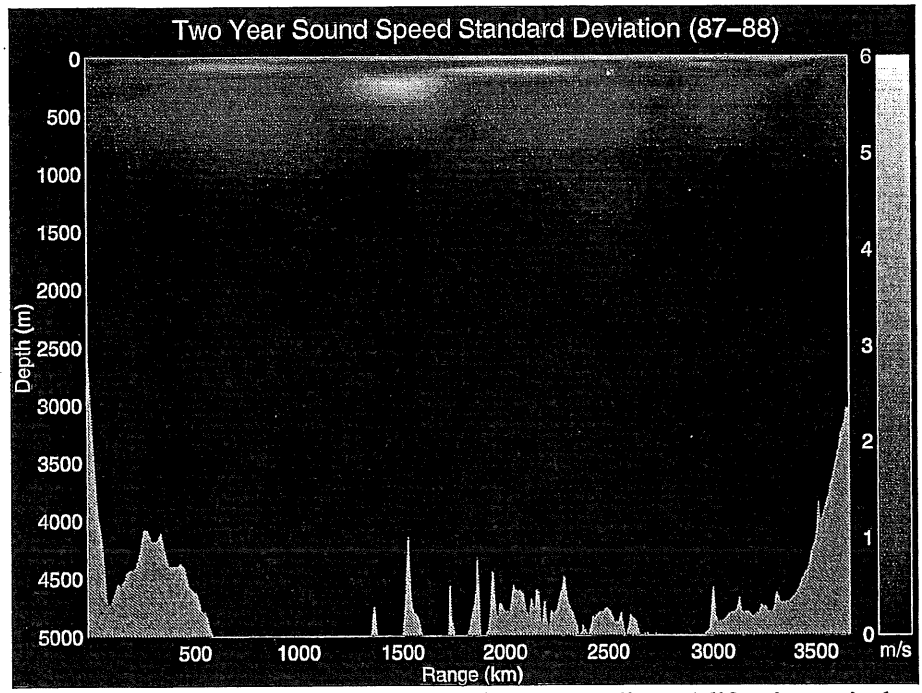



**Figure 7.** Standard deviation of sound speed in a Hawaii-to-California vertical section over a two-year time period.

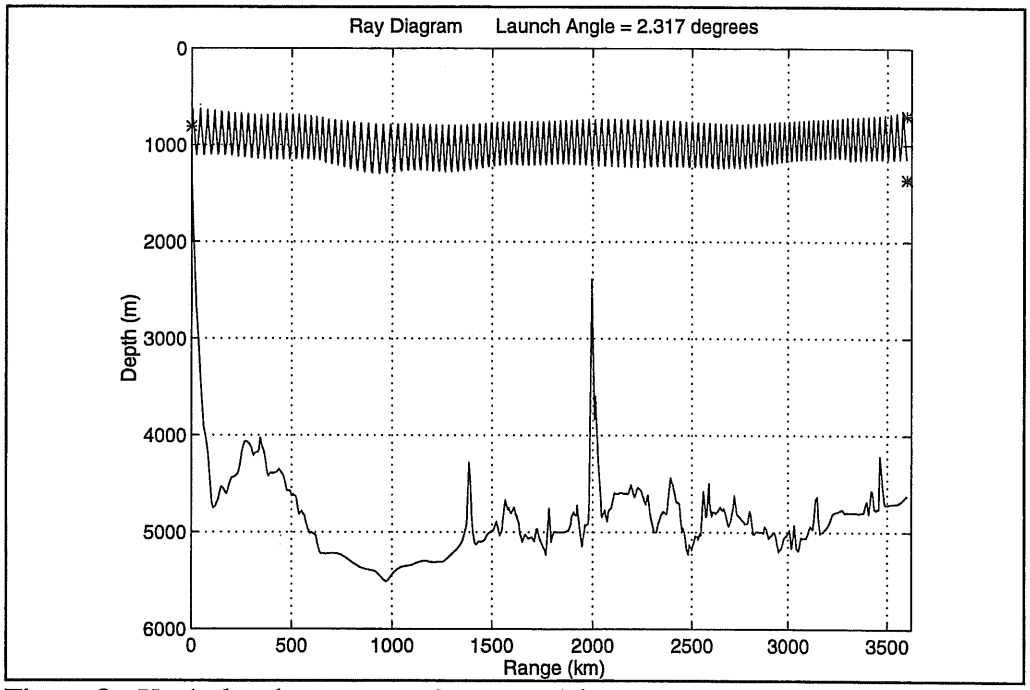



**Figure 8.** Vertical path geometry of a near-axial ray.

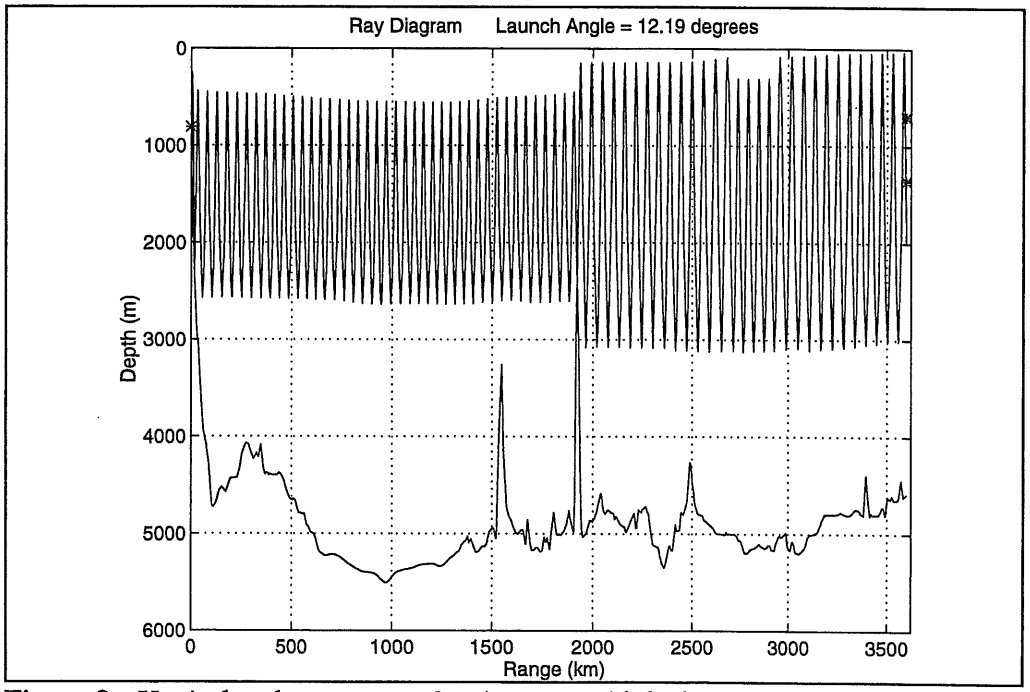



**Figure 9.** Vertical path geometry of a steep ray with bathymetry interaction.

Horizontal refraction is considerable. Figure 10 displays the horizontal path geometry of both a near-axial (southern ray) and a steep raypath (northern ray) along with contoured topography. Both rays were launched with the same azimuth angle. Refractive effects are obvious as steeper ray paths are influenced more by the north-south temperature gradient, refracting northward toward cooler temperatures (and lower sound speeds). A great circle path with an initial azimuth angle of 58.1° (same initial azimuth as traced rays) is also shown and appears concurrent with the near-axial ray path. Horizontal separations between near-axial paths and deep-cycling paths launched with the same azimuthal angle off Hawaii can be as much as 100 km near California.

In addition to the effect of the topography in the Murray Fracture Zone, interaction with the bottom near the source in Hawaii is also observed. Figure 11 shows a dot plot of travel time versus vertical launch angle with a receiver depth of 700 m for all eigenrays calculated over the two-year study period. The general shape of decreasing arrival time with increasing magnitude of angle is expected as steeper rays travel through higher sound speeds and thus arrive sooner. However, two anomalies exist, one between -7° and -10° and the other one between +8° and +10°. Rays launched at these two angular ranges strike the sloping bottom near Hawaii and are reflected out into more axial paths, causing the corresponding travel times to increase as slower sound speeds are encountered. Similar features can be seen in Figure 12 for the 1359-m receiver. The lack of near-axial arrivals is apparent for the deeper receiver, which is located over 500 m below the sound channel axis.

### 3.3.2 Arrival Structure Stability

The simulated time series of arrival structures provides a means for analyzing both stability and travel time variability. First we look at stability.

The ATOC sound source is planned to transmit a phase-modulated signal at a carrier frequency of 75 Hz. The modulation is an m-sequence of 1023 digits, with each digit containing two cycles. This gives a total sequence period of T = 1/75 Hz * 1023 digits * 2 cycles/digit = 27.28 seconds. The duration of the multipath arrival structure will last approximately eight seconds, so the entire arrival structure will be observed without lane ambiguity. The matched filter output is effectively a pulse having a duration of 26.6667 ms, but in generating a simulated arrival structure the pulse length was roughly doubled in an attempt to account for dispersion effects.

The arrival structure was determined from the arrival time, amplitude and phase shift associated with each eigenray. Typical simulated arrival structures for the 700-m receiver and the 1359-m receiver are shown in Figures 13 and 14, respectively. Amplitudes ramp up to peak at later arrival times for the 700-m receiver, signifying the strength of the more dominant axial eigenrays. The arrival structure for the 1359-m receiver noticeably lacks the higher amplitude later arrivals and has a less banded pattern than the 700-m receiver. Already we can see the effect of receiver depth on arrival structure, as the upper turning points of steeper eigenrays pick up more of the natural variability of the upper water column, resulting in a less organized arrival structure.

Arrival structures were computed for each six-day "snap-shot" for both the 700-m and 1359-m receiver depths. The time series of the arrival patterns for the latter depth is shown in a waterfall plot displayed in Figure 15. In viewing the arrival structure it is difficult to "visualize" stable

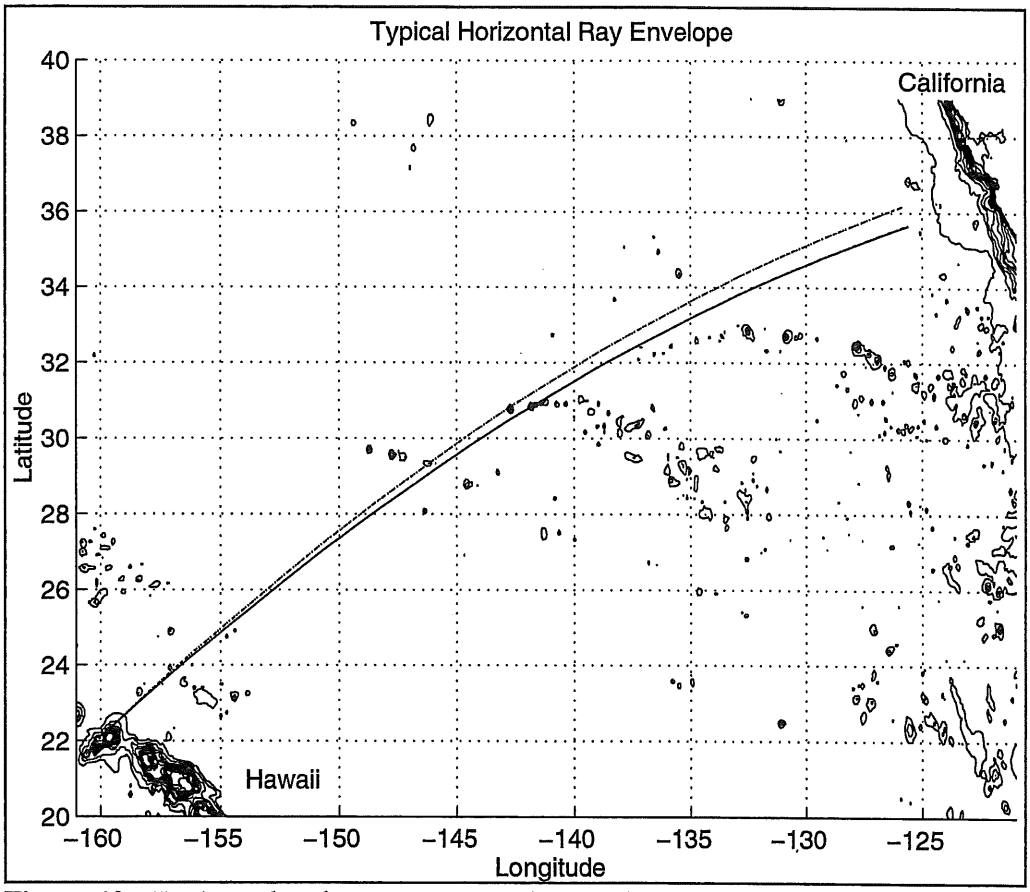



**Figure 10.** Horizontal path geometry of a deep-cycling (upper curve) and a near-axial ray (lower curve), both having the same initial azimuth angle. The horizontal path geometry of the near-axial ray coincides with the great circle path.

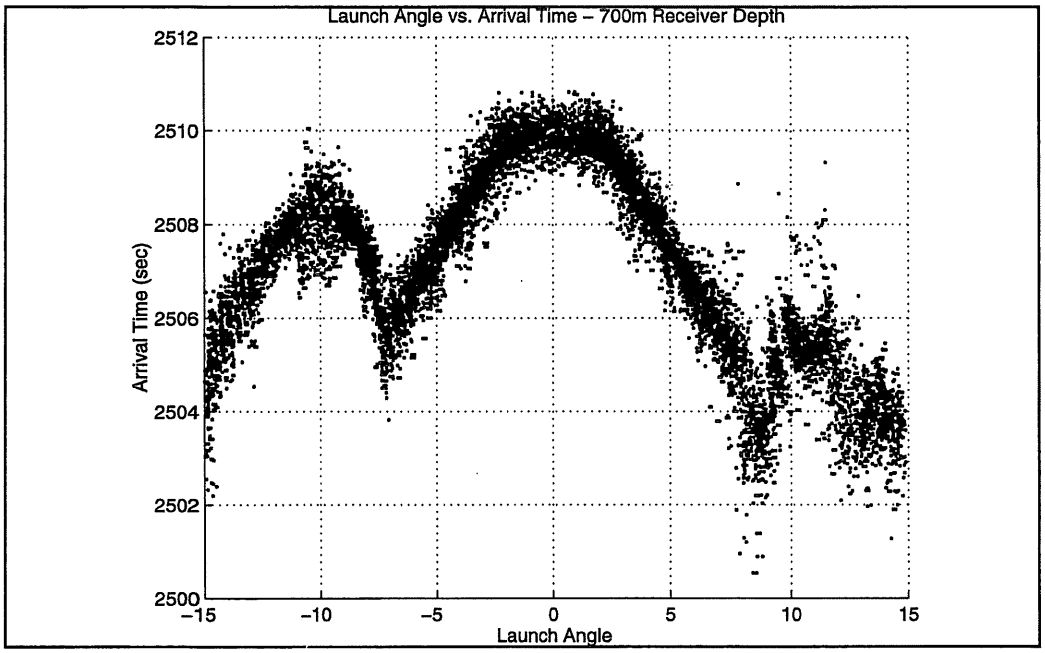



**Figure 11.** Dot plot of travel time versus vertical launch angle with a receiver depth of 700 m for all eigenrays calculated over a two-year study period.

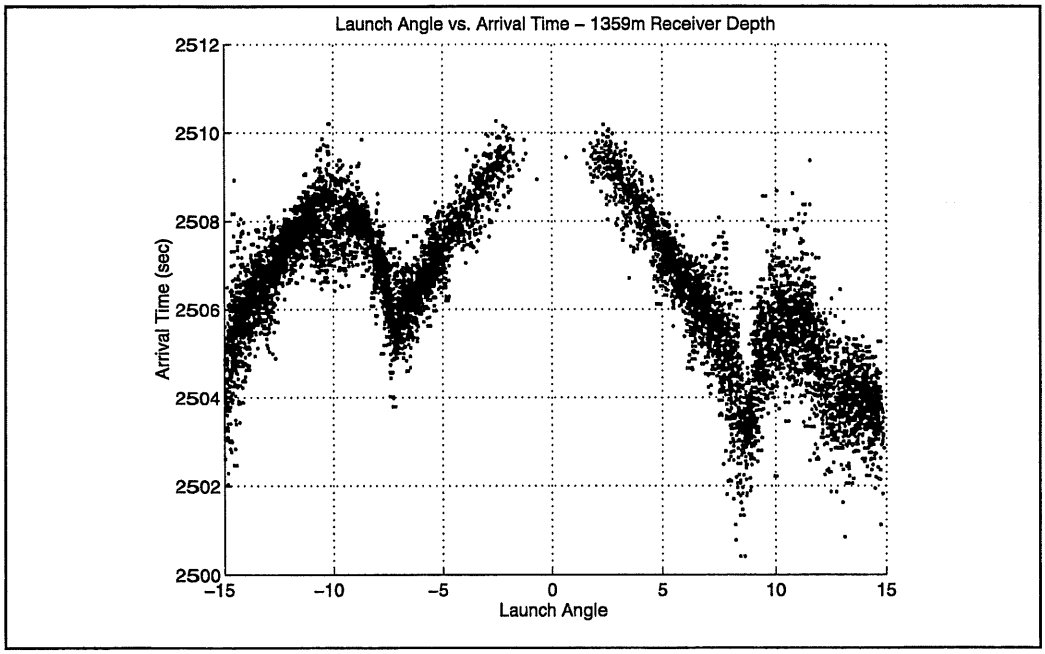



**Figure 12.** Dot plot of travel time versus vertical launch angle with a receiver depth of 1359 m for all eigenrays calculated over a two-year study period.

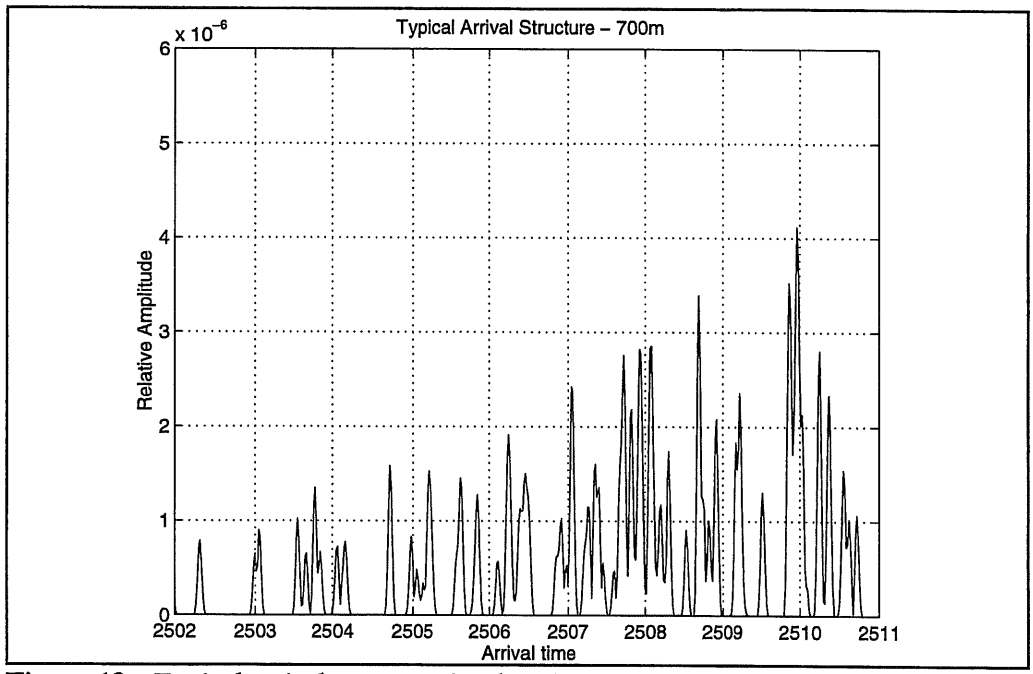



**Figure 13.** Typical arrival structure for the 700-m receiver.

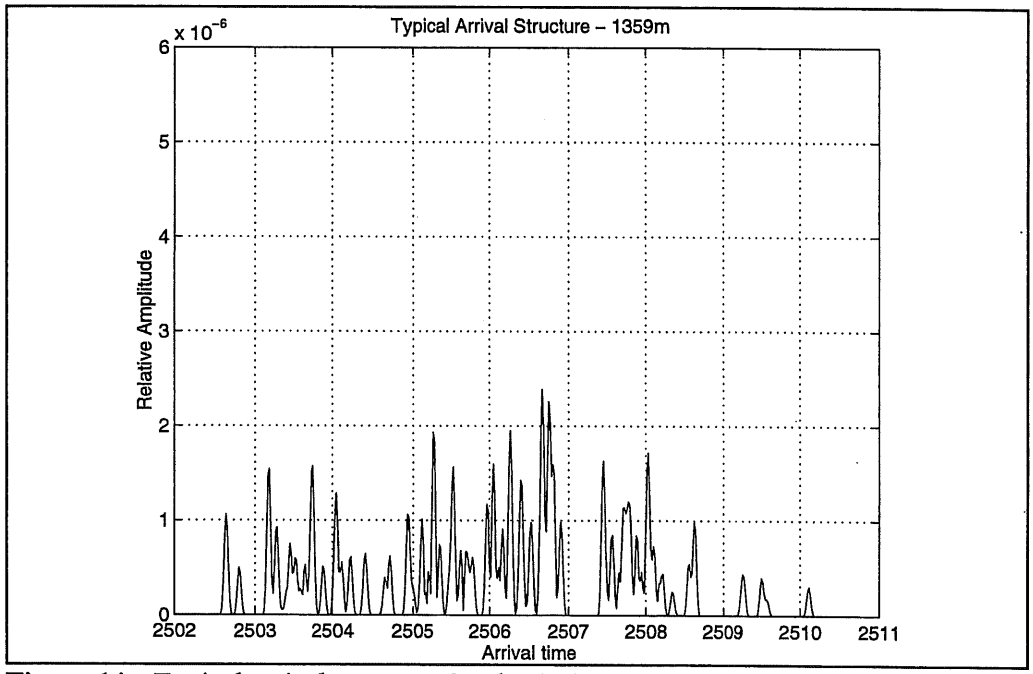



**Figure 14.** Typical arrival structure for the 1359-m receiver.

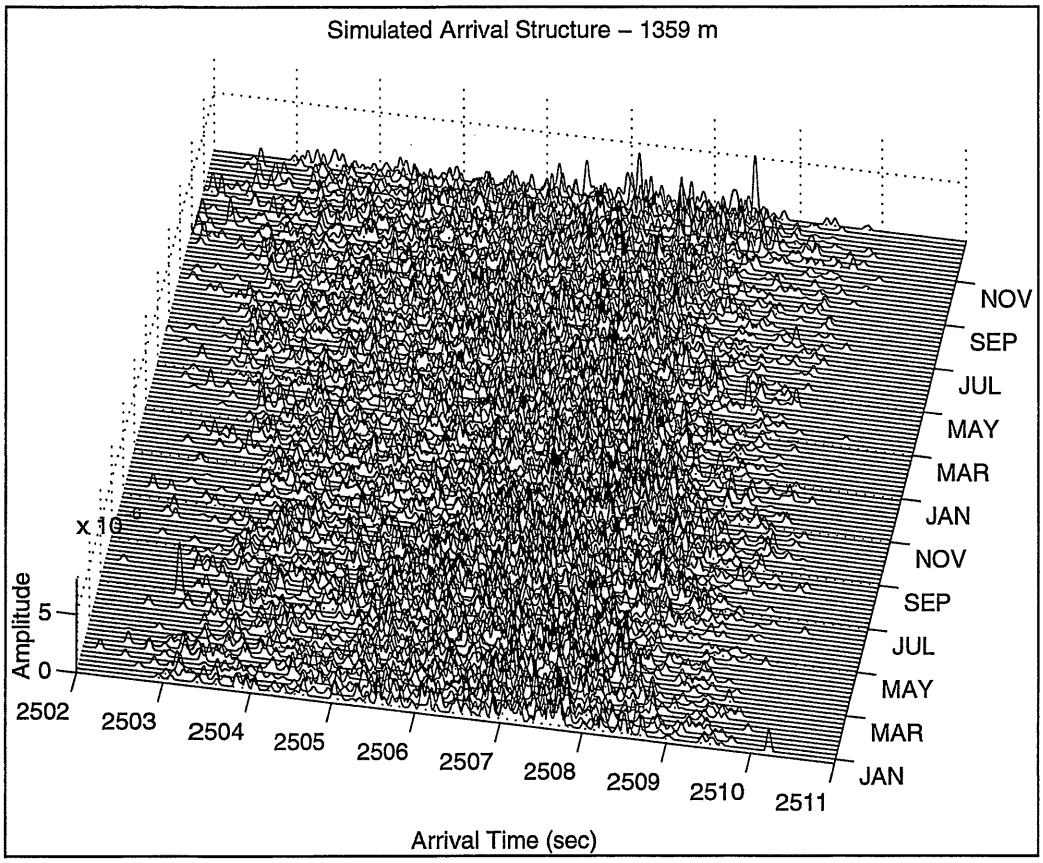



**Figure 15.** Arrival structures over two year time period for the 1359-m receiver depth.

patterns of arrivals or any correlation over the two year period.

The correlations and the lag times of the correlation peaks between the first arrival pattern and all the later arrival patterns for each receiver depth were calculated. The results as shown in Figure 16(a) for the deep receiver indicate a consistent 0.6 correlation, which is perhaps higher than one would expect from a visual inspection of the waterfall plot. In addition, time shifts from zero lag are plotted, showing a pronounced mesoscale oscillation and a less defined seasonal oscillation. Rms lag time variability for the deep receiver is estimated to be 0.51 s.

Figure 16(b) displays results for the near-axial receiver, yielding a 0.6 correlation also. In contrast to the deep receiver lag time plot, seasonal oscillations are very pronounced and mesoscale oscillations are very subdued. Rms lag time variability for the near-axial receiver is estimated to be 0.33 s.

In order to facilitate visual tracking of arrivals in the waterfall plots, correlation lag times were subtracted from their respective arrival structures and the results are plotted in Figure 17 for the 1359-m receiver depth. Obvious bands of consistent arrivals can be seen, indicating a certain degree of stability over time and also identifying which arrivals are the most stable.

### 3.3.3 Travel Time Variability

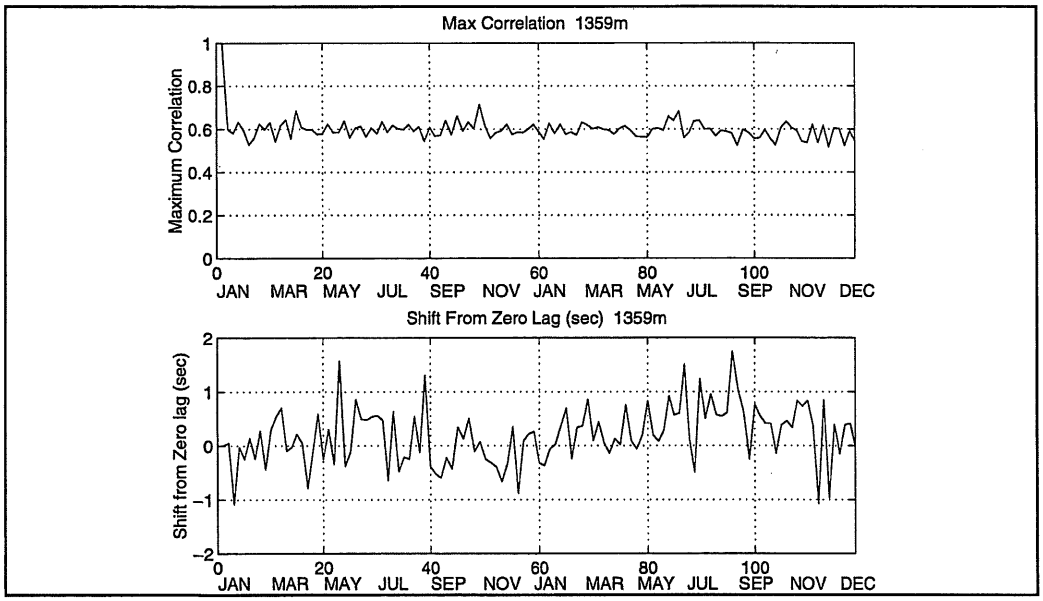
To characterize and quantify travel time variability over the entire two-year time period, the frequency spectra of the time series of the shift-from-zero-lag times were calculated for the two receiver depths. These two spectra of lag times are shown in Figure 18. Note frequency is plotted in cycles per year. The dominant frequency of variability for both receiver depths is seasonal (one cycle per year) with roughly similar spectral values for each, but variability at mesoscale frequencies are clearly more significant for the deeper receiver.

To reinforce the results of the lag time analysis, individual eigenrays were selected for similar analysis. Two reasonably stable eigenrays were isolated, one near-axial with a launch elevation angle of 5° and the other steep with a 13° angle. Figure 19 shows the spectra of the respective time series of arrival times. For the near-axial eigenray, the dominant frequency is seasonal, while for the steep eigenray seasonal and mesoscale variability appear nearly equal in size. The distinct peak at three cycles per year is interesting in that it could be an indication of "sensitivity" to a particular scale of mesoscale phenomenon. This "sensitivity" might be inherent to the particular eigenray selected for analysis, and other eigenrays may be responsive to other corresponding scales of variability.

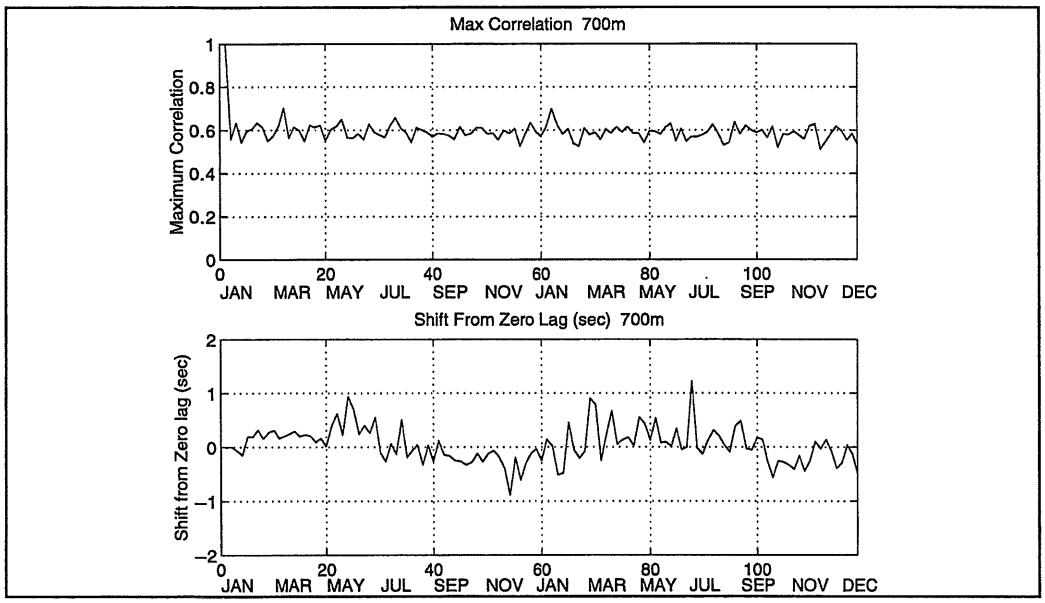
The overall rms travel time variability, or "geophysical noise", is estimated to be 0.28 s for near-axial rays and 0.42 s for steep rays. These estimates are consistent with the lag time values presented in the previous subsection of 0.33 s and 0.51 s for the near-axial and deep receivers, respectively. The frequency spectra indicate that travel time changes for near-axial rays are primarily due to seasonal changes, while for steep rays a combination of seasonal and mesoscale changes contribute to the overall variability of travel time.

## 4 CONCLUSIONS

In this research, a planned ATOC transmission from Hawaii to Monterey was simulated for a two-year period using a ray-based acoustic model and a global ocean circulation model. The




(a)




(b)

**Figure 16.** Correlation coefficients and the lag times of the correlation peaks between the reference (first) arrival pattern and all the later arrival patterns for (a) a receiver depth of 1359 m and (b) a receiver depth of 700 m.

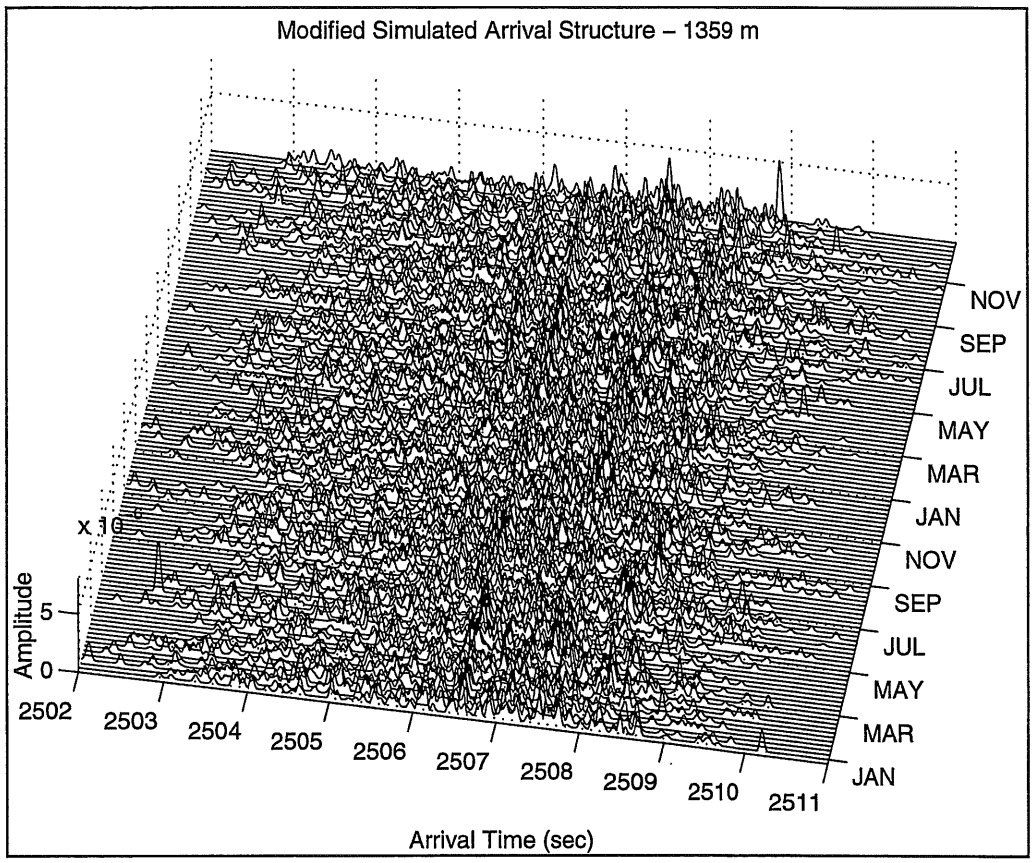



**Figure 17.** Aligned arrival structures for the receiver at a depth of 1359 m with correlation lag times subtracted, resulting in an improved visualization of stable arrivals.

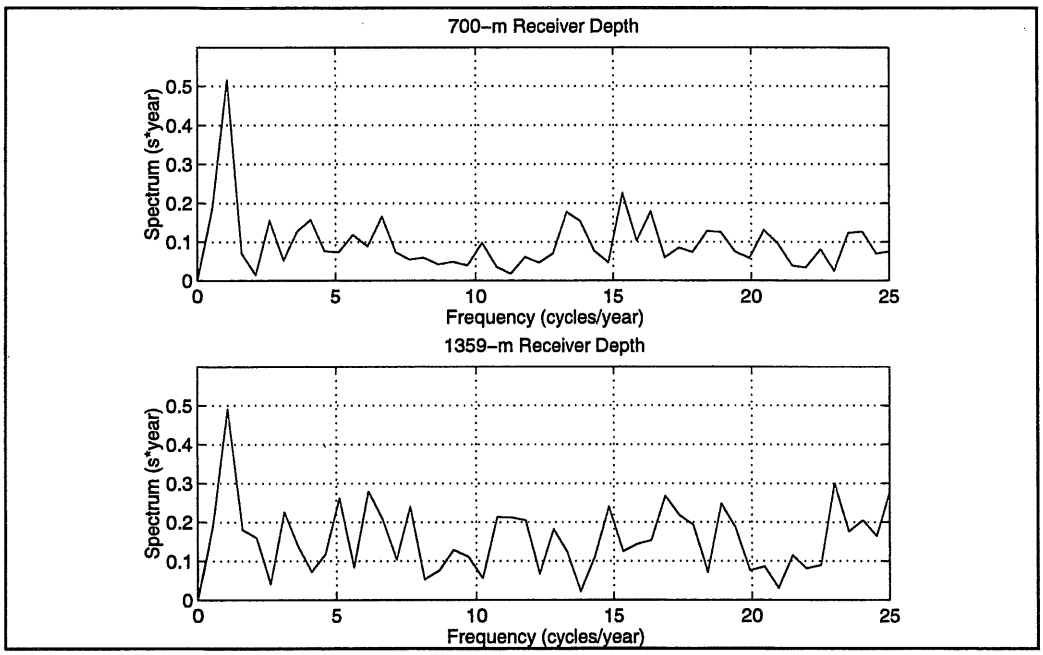



**Figure 18.** Frequency spectra of time series of the correlation peak lag times for receiver depths of 700 m (upper panel) and 1359 m (lower panel).

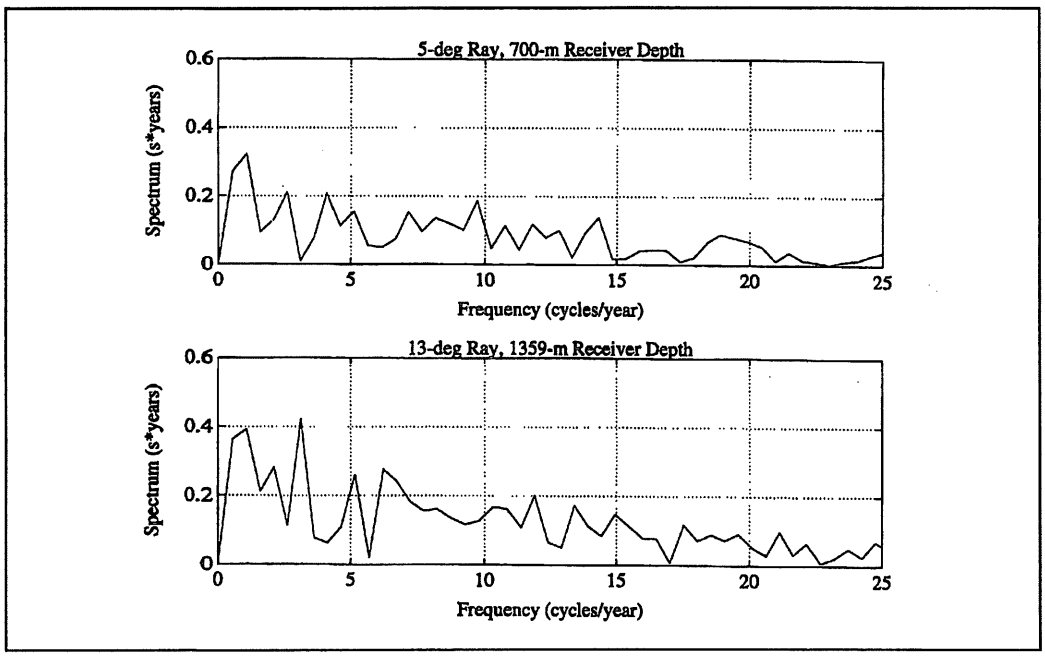



**Figure 19.** Frequency spectra of time series of travel times for a near-axial eigenray (upper panel) and a deep-cycling eigenray (lower panel).

simulated raypath geometries, arrival structures, and the effects of mesoscale, gyre and seasonal variability were analyzed.

Major conclusions based on this study are as follows:

1. Multipath arrival structure strongly dependent on receiver depth.
2. Multipath arrival structure quasi-stable with a correlation of 0.6 between arrival patterns at different times.
3. It is difficult to "visually" track individual arrivals over time without alignment.
4. Subtraction of correlation peak times allows for alignment to facilitate visual tracking of individual arrivals.
5. The temporal change of the time lag of correlation peaks is consistent with individual ray travel-time variations.
6. Bottom topography in the vicinity of the Moonless Mountains has an occasional mild effect on rays with launch angles larger than approximately 12°.
7. Travel time variability ("noise") estimates are 0.42 seconds rms for steep rays and 0.28 seconds rms for near-axial rays.
8. Near-axial ray variability is due primarily to seasonal changes, while steep ray variability is due to a combination of seasonal and mesoscale changes.
9. A longer duration simulation study is needed to appropriately characterize "noise" due to inter-annual cycles.

## ACKNOWLEDGEMENTS

The authors wish to acknowledge those whose contributions made this study possible: Stefan Hudson, Chris Miller, Robin Tokmakian, and Rob Bourke. This work is sponsored by the ATOC Project Office, Scripps Institution of Oceanography. This work relates to the Advanced Research Projects Agency Grant MDA972-93-1-0003 issued by the Contracts Management Office.

## REFERENCES

[1] W.H. Munk and A.M.G. Forbes, "Global Acoustic Warming: An Acoustic Measure?," *Journal of Physical Oceanography,* Vol. 19, pp. 1765-1778, 1989.

[2] C-S. Chiu, A.J. Semtner, C.M. Ort, J.H. Miller and L.L. Ehret, "A Ray Variability Analysis of Sound Transmission from Heard Island to California," *The Journal of the Acoustical Society of America,* Vol. 96:4, pp. 2380-2388, 1994.

[3] R.M. Jones, J.P. Riley and T.M. Georges, "HARPO: A Versatile Three-Dimensional Hamiltonian Ray Tracing Program for Acoustic Waves in an Ocean with Irregular Bottom," Wave Propagation Laboratory, National Oceanic and Atmospheric Administration, Boulder, 457 pp., 1986.

[4] A.J. Semtner, "Very High-resolution Estimates of Global Ocean Circulation, Suitable For Carbon-cycle Modeling," *Proceedings of the Snowmass Global Change Institute on the Global Carbon Cycle,* Office of Interdisciplinary Earth Studies, Boulder, in press, 1995.

[5] A.J. Semtner, and R.M. Chervin, "A Simulation of the Global Ocean Circulation with Resolved Eddies," *Journal of Geophysical Research,* Vol. 93:C12, pp. 15502-15522 and 15767-15775, 1988.

[6] A.J. Semtner, and R.M. Chervin, "Ocean General Circulation from a Global Eddy Resolving Ocean Model," *Journal of Geophysical Research,* Vol. 97:C4, pp. 5493-5550, 1992.

[7] A.E. Newhall, J.F. Lynch, C-S. Chiu, and J.R. Daugherty, "Improvements in Three Dimensional Ray Tracing Codes for Underwater Acoustics," *Computational Acoustics: Ocean Acoustic Models and Supercomputing,* North-Holland, Amsterdam, 1990.

[8] K.V. Mackenzie, "Nine-term Equation for Sound Speed in the Ocean," *Journal of the Acoustical Society of America,* Vol. 70:4, pp. 807-812, 1981.

[9] H.W. Menard, "Marine Geology of the Pacific," McGraw-Hill, Inc., New York, 1964.

Theoretical and Computational Acoustics
D. Lee, Y-H Pao, M.H. Schultz, and Y-C Teng (Editors)


# Simulated Wave Propagation in a Waveguide by a Modified Finite Length Ping Time Dependent Ray Trace Technique

Elmer White
130 Moonraker Drive, Slidell, LA 70458, USA

**Abstract**
A finite length ping ray trace program has been developed which is capable of simulating time dependent wave motion in a complicated bathymetric waveguide. One objective is to compare predictions of a finite length ping ray trace model against an accepted parabolic equation (PE) model to determine how well a time dependent ray trace agrees with the PE calculation of the transmission loss. Another objective is to compare the prediction of wave motion in a waveguide by normal mode theory to that generated by the finite length ping ray trace approach. With the finite length ping ray trace time dependent (FLPRT) program, it has been shown that backscattering is the primary cause of lowering of the transmission loss (TL) at the initial entry area of a wedge.

## 1. Introduction

FLPRT has been benchmarked against the accepted Navy standard PE code. The benchmark was made by comparing the TL values computed by both FLPRT and PE to determine the accuracy of the prediction from FLPRT to that of PE. The solution comparing the two data is taken from the PE workshop 2 report [1] and shown in figure 29. A description of the new ray structure and logic is shown in figures 1 and 2. A constant frequency ray is depicted as type 1 and a variable frequency ray is depicted as type 2 in figure 1. These finite length ping rays have the referenced angular orientation as shown in figure 2. The number of finite points that make up the structure of a ray is a function of the ping length and the time step which is automatically computed in the program FLPRT. The logic for selecting the number of points making up a single ray is arbitrary but consistent. The uniqueness of FLPRT is demonstrated in figures 4 to 28 which shows the relative positions of a wave traversing a wedge submerged in a shallow water column. The concept is that of a water column over a half space. The solution approximating test case 2 (TCASE2) in the report [1] in a range dependent environment demonstrates forward and backscatter characteristics which agree with both PE and normal mode theory in its behavior pattern. The decimation of the unique wave pattern on

reaching the wedge pinnacle is unusual in that the wave is split with a reduced part being transmitted in the forward direction over the wedge while the main wave recedes toward the starting point. The ray program contains a correction factor for caustics, displacement and critical angle in its logic and/or the complex pressure computation. The acoustic pressure field is computed in the path of the ray and converted into TL data at pre-selected depths.

## 2. Modified Ray Theory (FLPRT Approach)

The basic approach to generating the formal ray theory follows that by Officer [2] and will not be rehashed in this paper. The approach is to begin by defining an element of ray in a non-gradient field as shown in field 3. Based on the geometry shown

$$\delta z = \delta s \, sin \, \Phi \tag{1}$$

where $\delta s$ is an element of path in the wave length of a ray at some frequency. In a velocity field containing a gradient,

$$\delta z = c / g \, ( \, cos \, \Phi_2 - cos \, \Phi_1 \, ) \tag{2}$$

where $c$ is the velocity at the initial point and $g$ is the velocity gradient at that point. It can be seen that

$$sin \, \Phi = \delta z / \delta s = \delta \lambda / ( c / g ) \tag{3}$$

where $\delta \lambda$ is an element of ray path length. Therefore;

$$arcsin \, [ \, \delta \lambda / ( \, c / g \, ) \, ] = \delta \Phi \tag{4}$$

and the wave length is

$$\lambda = c / f \tag{5}$$

where $f$ is the frequency. In the established geometry, the angle at the end of $\delta \Gamma$ is defined as

$$\Phi_2 = \Phi_1 + \delta \Phi \tag{6}$$

If $\Phi$ is less than the absolute value of 90 degrees, then the element of range is seen to be

$$\delta x = ( \delta s^2 - \delta z^2 )^{1/2} \quad (7)$$

otherwise

$$\delta x = -\delta x \quad (8)$$

If no velocity gradient is present,

$$\delta x = \delta s \cos \Phi, \quad (9)$$

therefore,

$$x = x_0 + \delta x \quad (10)$$

and

$$z = z_0 + \delta z. \quad (11)$$

Since

$$\delta s = ( \delta x^2 + \delta z^2 )^{1/2} = \delta \lambda \quad (12)$$

$$s = s_0 + \delta s. \quad (13)$$

By definition

$$\delta t = \delta s / c \quad (14)$$

then

$$t = t_0 + \delta t. \quad (15)$$

If the ping length ($pl$) has been specified, the number of points in a ping can be determined by

$$np = ( pl / \delta t + 0.5). \quad (16)$$

### 3. Complex Pressure In Path of Ray

The complex pressure in the path of a ray is based on the formalism developed by Westwood [3]. The equations are rotated to correspond to the geometries shown in describing the ray path concepts in figure 3. The complex pressure in the path of a ray is defined as

$$p = -1^{ns} G(R) K(R) e^{i\pi/4} e^{kR(R1 + iR2)} (\Phi_1 - \Phi_2) \quad (17)$$

where

$$R1 = (1/kR)[nb \log_{10} v + \log_{10} w] \quad (18)$$

$$R2 = \cos(\Theta - \Theta_0) - [z_r / R](\cos\Theta - n\cos\Theta_1) \quad (19)$$

$$\Theta = \pi/2 - \Phi_1 \quad (20)$$

$$\Theta_1 = \pi/2 + \Phi_1 \quad (21)$$

$$\Theta_0 = \pi/2 - \Phi_i. \quad (22)$$

For a flat wave guide,

$$K(r) = (k / 2\pi r)^{1/2} \quad (23)$$

For the wedge,

$$K(r) = (k / 2\pi)^{1/2} \quad (24)$$

where

$$k = 2\pi f / c \quad (25)$$

and

$$v = \sum_{j=1}^{nb} (r_a \cos\Phi_j + a) \quad (26)$$

and the density ratio

$$r_a = d_b / d_w. \quad (27)$$

The relationship between the wedge and the ray becomes

$$\Phi_j = \Theta - 2(j\text{-}1)\alpha \quad (28)$$

Where $\alpha$ = wedge face angle.

$$w = 2r_a \cos \Theta_n / r_a \cos \Theta_n + a ) \quad (29)$$

$$\Theta_n = \Theta - 2nb\ \alpha \quad (30)$$

$$a = ( n^2 - \sin^2\Theta )^{1/2} \quad (31)$$

$$G(R) = 1 / R \quad (32)$$

$R$ = distance traversed in the path of the ray.

For a flat waveguide,

$$G(R) = 1 \quad (33)$$

but

$$w = 2\, r_d \cos \Theta_1 \ / \ ( r_d \cos \Theta_1 + a ) \quad (34)$$

and

$$v = ( r_d \cos \Theta_1 - a ) / ( r_d \cos \Theta_1 + a ). \quad (35)$$

The symbol $a$ is defined the same for both the wedge and the flat waveguide. Snell's Law is the same in each case as defined by

$$n = \sin \Theta / \sin \Theta_1 . \quad (36)$$

The symbol definitions are indicated below.

$\Theta$ = wave angle incident on the interface.

$\Theta_1$ = wave angle transmitted into the bottom.

$w$ = transmission coefficient for fluid-fluid interface.

$v$ = Rayleigh reflection coefficient for fluid-fluid interface.

$nb$ = number of bottom bounces.

$ns$ = number of surface bounces.

## 4. PE Workshop Test Case 2 Wedge Problem

Figure 4 represent the wedge assembly benchmarked against by the modified ray model. Note the wave expansion which is beginning between the depths 0.08 and 0.12 kilometers. The scale between the range and depth is approximately 25 to 1. Figure 5 details the acoustic properties in the water column and the wedge half space. In the test case, $z_r$ is the receiver depth and $z_s$ the source depth. The wedge angle $\alpha$ is shown to be 2.86 degrees. In viewing figures 6 and 7 (ref. times 0.06 to 0.08 seconds), it can be seen that the wave is bouncing off the interfaces. Advancing to times 0.46 to 0.56 seconds, the wave is more defined with observable oscillations bouncing from one interface to the other as shown in figures 8 to 13. Note that the wave is progressing up the side of the wedge. From times 1.96 to 2.06 seconds the well defined oscillation goes through one cycle as shown in figures 14 to 19 and is near the top of the wedge. Between times 2.52 to 2.64 seconds (figures 20 to 25), the oscillatory part of the wave has advanced to the top of the wedge where it is beginning to be decimated. The advanced part of the wave has crossed the peak and progressed down the far side as shown in figure 25. As shown in figures 26 and 27 (times 3 to 3.02 seconds), the wave has split and begins to recede to the origin while the advanced part continues forward. At time 4.5 seconds (figure 28), the wave is shown in full recession with the forward section relatively sparse near the boundary.

## 5. Transmission Loss Curve Comparison Between FLPRT and PE Predictions

Figure 29 is taken from the book [1] edited by Chin-Bing et al, 1993 which was generated by the PE model. The overlay of the TL data generated by FLPRT on the PE prediction is shown in figure 30. The FLPRT TL represent a two way transmission from source to target. The TL curve drop off on the downhill side of the wedge is due to the loss of energy crossing the wedge peak. The FLPRT loss curve recovers near the boundary. This can be explained in that the rays refocus after passing the peak of the wedge. It has been observed that the initial peak near range zero is primarily due to

backscatter which begins immediately as the wave enters the wedge area. For one way transmission, the TL curve tends to remain essentially flat indicating a continuous loss rate for the full transmission process.

## 6. Conclusion

The accuracy of the finite length ping process depends on the sampling rate and correctly optimizing the complex pressure at the depth of interest due to many paths being summed in a specific area or range. By performing a summation of the rays within an element of range and averaging in the complex domain the TL values tend to increase below the PE curves. Backscatter raises the TL values near the source but lowers the TL values near the peak but recovers on the downhill side due to refocussing.

## 7. References

1. Chin-Bing, S., Davis, J.A., Evans, R.B. and King, D.B., PE Workshop II, Proceedings of the Second Parabolic Equation Workshop, Naval Research Laboratory, Stennis Space Center, (MS) 39529-5004, May 1993.

2. Officer, C.B., Introduction to the Theory of Sound Transmission With Application to the Ocean (McGraw-Hill Book Co., Inc.), New York (NY) 1958.

3. Westwood, E.K., Acoustic Propagation Modeling in Shallow Water Using Ray Theory, Applied Research Laboratory, The University of Texas at Austin, Austin, Texas 78713-8029, Technical Report ARL-TR-89-6, 21 February 1989.

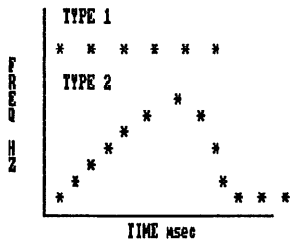



Figure 1 Ping Structure Type

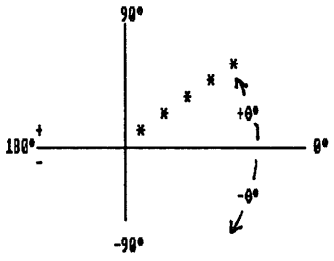



Figure 2 Ray Orientation

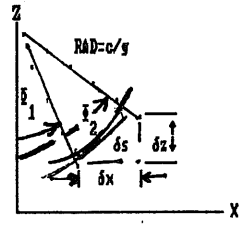



Figure 3 Ping Geometry

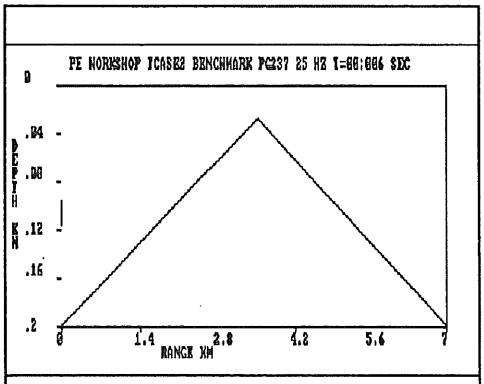



Figure 4 FLPRT T=0.006 SEC

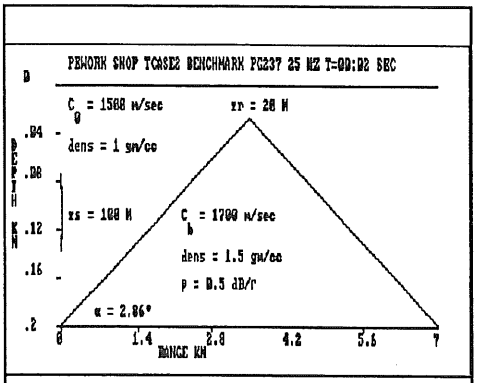



Figure 5 FLPRT T=0:02 SEC

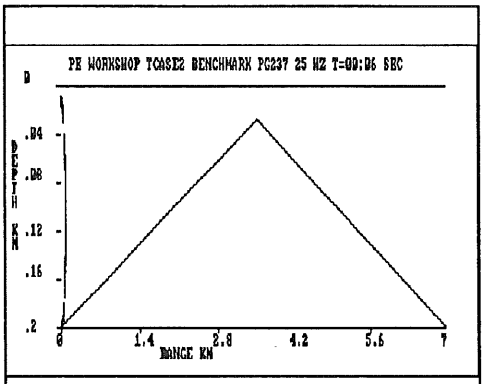



Figure 6 FLPRT T=0:06 SEC

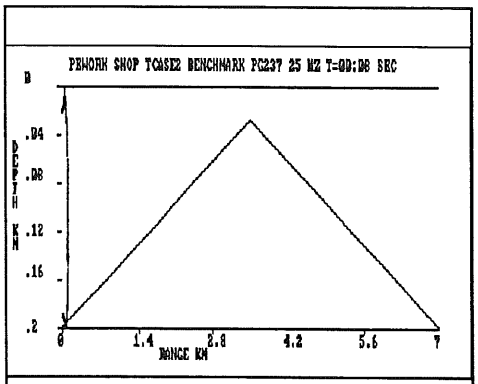



Figure 7 FLPRT T= 0:08 SEC

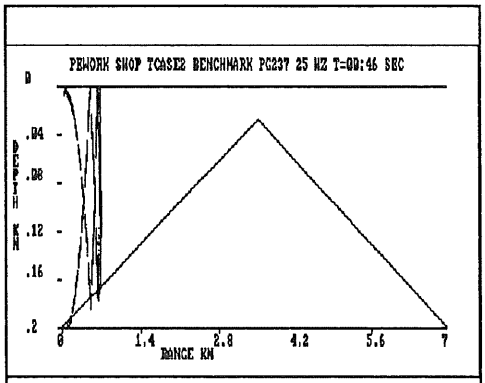



Figure 8 FLPRT T=0:46 SEC

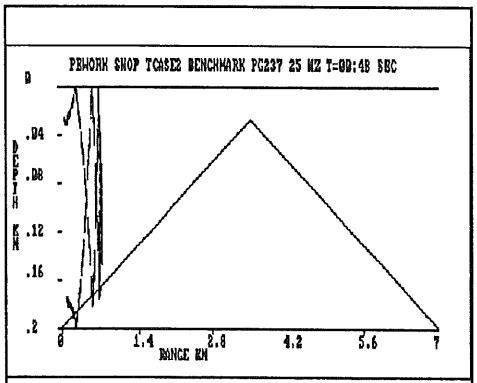



Figure 9 FLPRT T=0:48 SEC

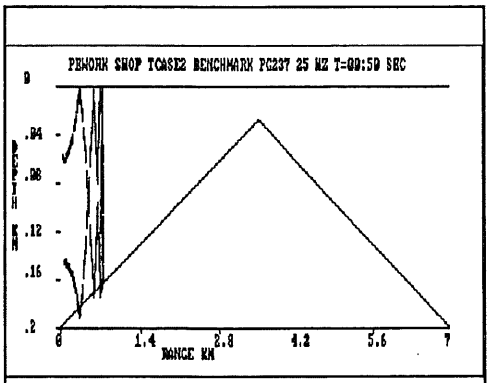



**Figure 10 FLPRT T=0:05 SEC**

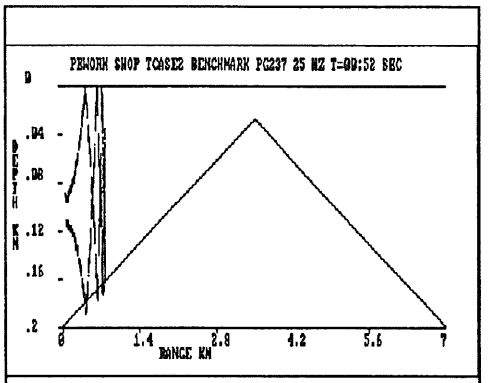



**Figure 11 FLPRT T=0:52 SEC**

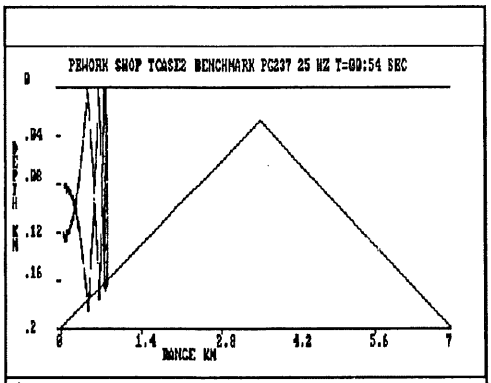



**Figure 12 FLPRT T=0:54 SEC**

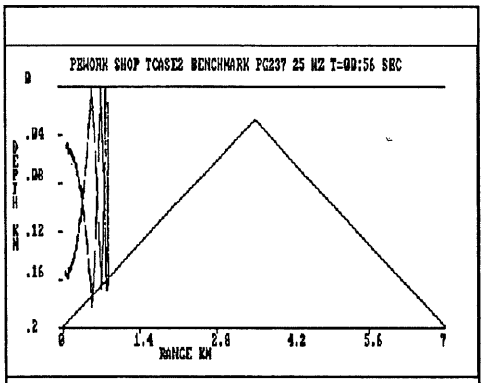



**Figure 13 FLPRT T=0:56 SEC**

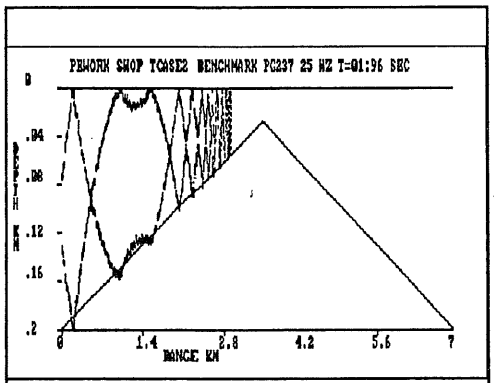



**Figure 14 FLPRT T=1:96 SEC**

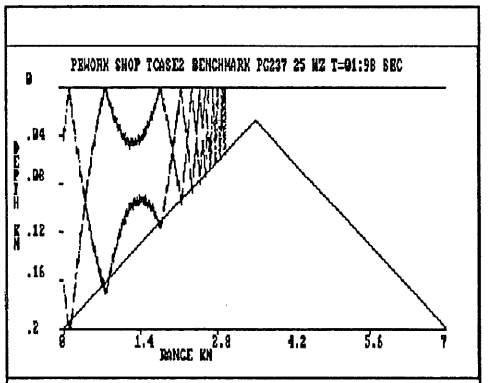



**Figure 15 FLPRT T=1:98 SEC**

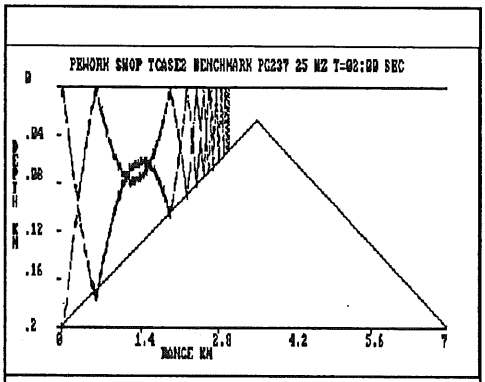



Figure 16 FLPRT T=2:00 SEC

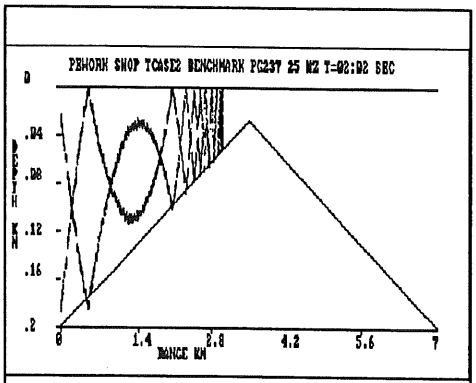



Figure 17 FLPRT T=2:02 SEC

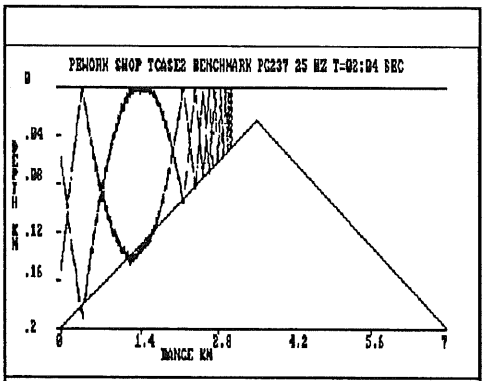



Figure 18 FLPRT T=2:04 SEC

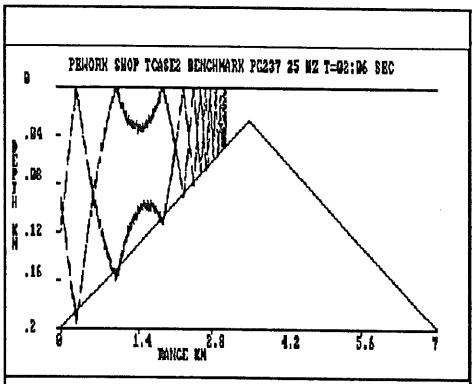



Figure 19 FLPRT T=2:06 SEC

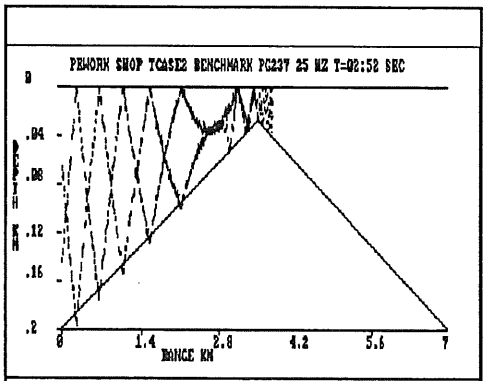



Figure 20 FLPRT T=2:52 SEC

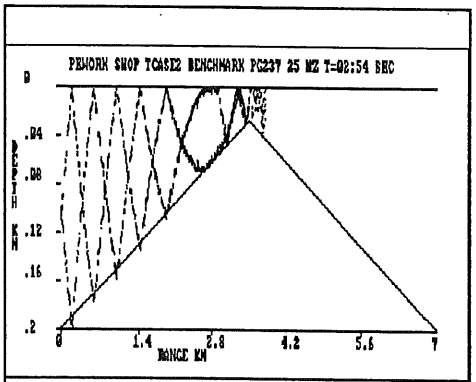



Figure 21 FLPRT T=2:54 SEC

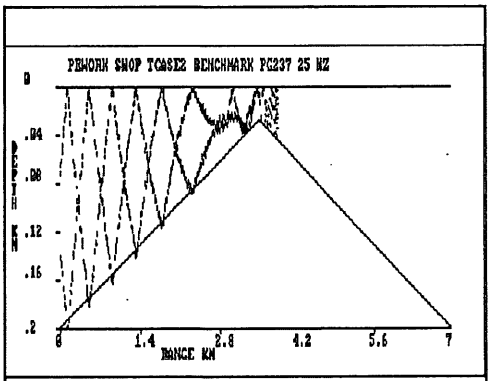



Figure 22 FLPRT T=2:56 SEC

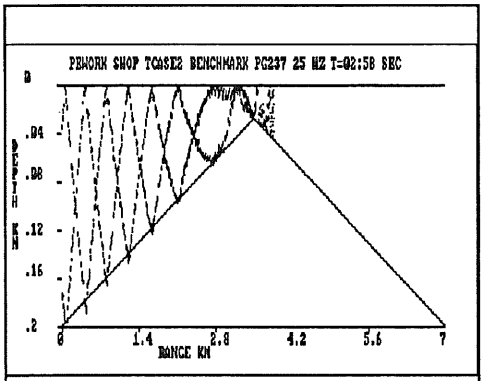



Figure 23 FLPRT T=2:58 SEC

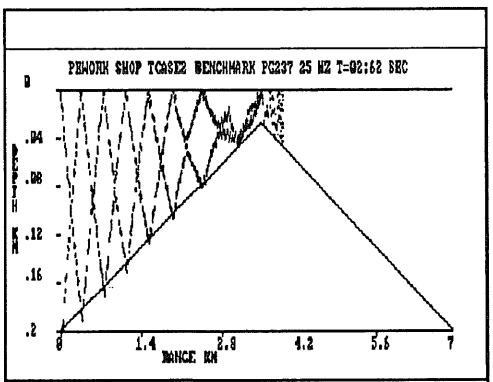



Figure 24 FLPRT T=2:62 SEC

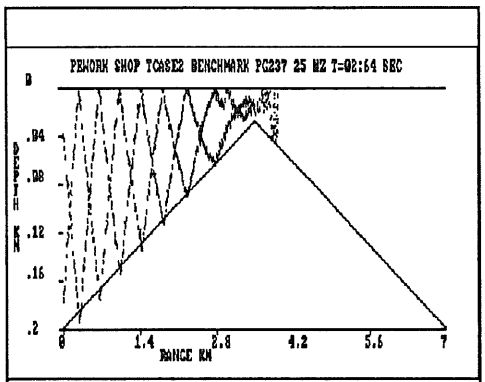



Figure 25 FLPRT T=2:64 SEC

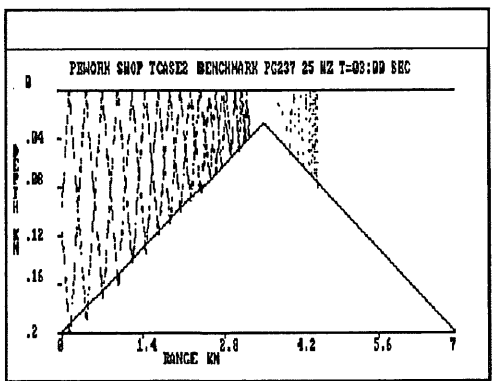



Figure 26 FLPRT T=3:00 SEC

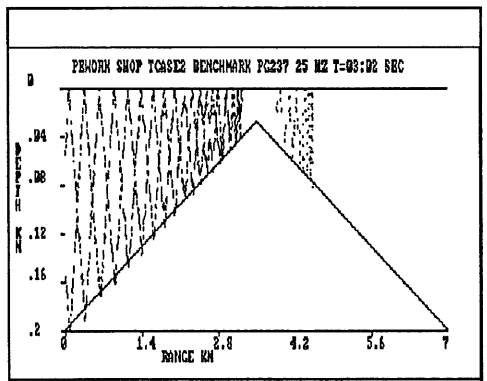



Figure 27 FLPRT T=3:02 SEC

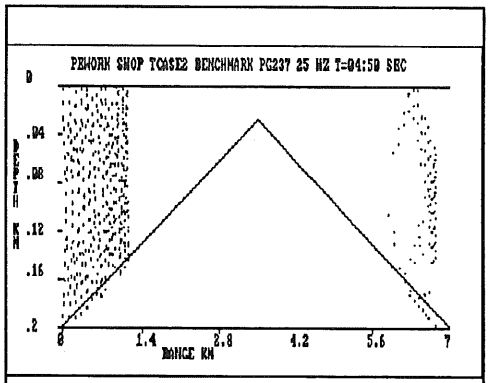



Figure 28 FLPRT T=4:50 SEC

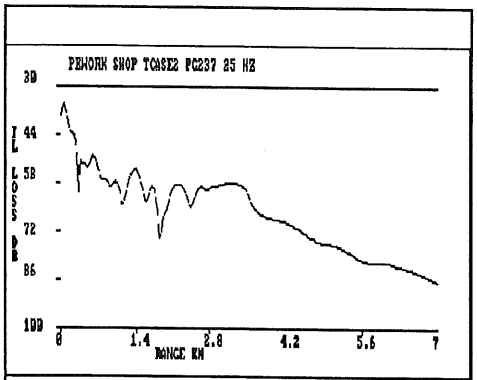



Figure 29 PE TESTCASE 2

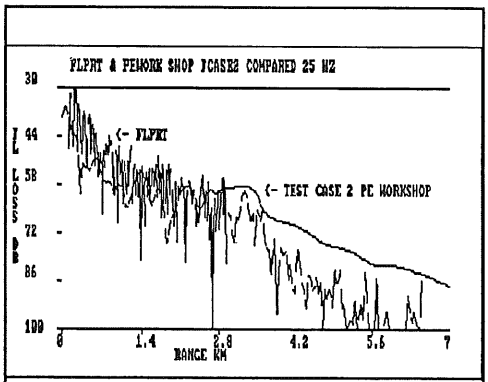



Figure 30 PE vs FLPRT

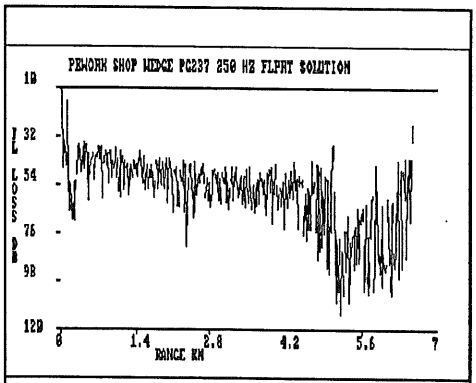



Figure 31 FLPRT 250 HZ

Theoretical and Computational Acoustics
D. Lee, Y-H Pao, M.H. Schultz, and Y-C Teng (Editors)


# Noise Cancellation by Arrays of Active Noise Sources in Enclosures with Impedance Surfaces

G. Rosenhouse and E. Saski
Faculty of Civil Engineering, Technion-Israel Institute of Technology
Haifa, 32000, Israel

## Abstract

The paper presents a new computer program for calculation and graphical display of sound fields caused by sound sources in enclosed spaces. For the simulation purposes the sources include monopoles, dipoles, quadrupoles etc as primary and secondary sources. We describe the sound field affected by rigid and absorbing boundaries, introducing the mathematical complexity due to the complex spherical wave reflection over highly absorbing areas. Next, impedance boundary effects are analyzed by an exact integral solution where necessary and in other cases by an approximate formulation based on "image" sources using complex reflection coefficients. The simulation program is then used to describe sound fields within enclosures containing a desired number of absorbing walls. Engineering and physical conclusions summarize the present work, including descriptions of the sound fields of interest.

## 1. Introduction

Active noise control within enclosures is involved not only with the superposition of sound fields caused by various sources, but also with reflections which add a diffusive field to the direct one. This combination increases considerably the complexity of the sound fields of both primary and secondary sources that serve the Active Noise Control (ANC). ANC is involved with multi-reflections effects not only because of the boundaries in the domain, but also due to all other obstacles in the examined environment that can be influential. Such sound reflections strongly distort the sound field in a way that might be understood only by a global view of the domain, as will be demonstrated in this paper.

There are many analysis methods for simulation of sound fields within enclosures. Basic reference books on room acoustics in this context are those of Cremer, Mueller and Schultz (1978) and Kuttruff (1976).

These methods include:

**Ray tracing methods**: The physical fundamentals of the methods and historical background are given in Pierce (1981). Specific publications on the use of these methods in acoustical analyses of auditoria were presented for example by Schroeder (1970), Santon (1976), Krokstad et al. (1968, 1983) and Sekiguchi, et al. (1985).

**Image sources methods**: Juricic and Santon (1973) and Gensane and Santon (1979). A more advanced theory, but with the deficiency of too many image sources at high order reflections was presented by Sakurai (1987), who superposed the response from reflected sounds from rigid plane panels, using the line integral formulation based on

Kirchhoff's formula and Kirchhoff's boundary conditions. Rosenhouse and Saski (1995) modified the image source method to enclosures with several sound absorbing impedance surfaces of high absorption coefficients.
**Numerical techniques**: Problems with complicated domains, such as obstacles within the enclosure, are more reasonably solved by numerical techniques like the Finite Element Method (FEM) combined with modal analysis, Migeot (1993) and the Boundary Element Method (BEM) for steady state and transient fields respectively - Terai, T., Kawai in the book of Ciskowski and Brebbia (1991).
**Modal analysis combined with minimization**: See Nelson and Elliott (1992).
In the present paper we apply the image sources method in empty rooms, where the direct sound and the effect of reflections from the wall surfaces within the enclosures are taken into account and some surfaces are highly sound absorbing. Various case studies show the possibility of using the method as an aid for design based on general understanding of source locations effect and the enclosure shape and properties. It brings us to concepts of modern control. Toady's control methodologies enable the computer "see" by guess technologies more than people can see. The computer can find today the clear picture of hidden phenomena by techniques of eliminating noise and observing meaningful clues and dots. The human brain can also draw conclusions by observing trends and clues, and even to some extent "intuition". The rest of the paper develops the technology and illustrates some sound fields created by a combination of primary and secondary sources, in various configurations, to show possibilities of active noise control in enclosures, saving a lot of computational minimization effort.

## 2. The choice between modal analysis based on rays reflected from impedance surfaces

A main difference between sound radiation in an enclosure and sound radiation in a free space is the existence of multi-reflections in enclosures due to the bounding surfaces, that should be considered carefully. Hence, an essential element in the analysis of a sound field in an enclosure is the reflection of sound from a sound absorbing surface. The strength and the directivity of the reflections and the natural modes of the enclosure depend on the boundary conditions that are attributed to the bounding surfaces. Both modal analysis and ray theories of room acoustics are possible and useful for different analysis purposes.
Generally, modal analysis is not common in room acoustics because of the overwhelming number of eigen frequencies and normal modes involved (about $10^9$ within the audio frequency domain). This situation does not enable an adequate acoustic estimation for design purposes. To this fact one should add the difficulties due to possible complexities in the room's geometry, including coupling between rooms and surface irregularities.
The approach based on sound rays and their reflection may be considered relatively simple, and we use it to illustrate sound fields in enclosures for Active Noise Control (ANC) design. The effect of multi-reflections of sound waves from impedance surfaces and their influence on ANC was discussed in detail by Rosenhouse and Saski (1995 - Active 95) and Chapter 3. It appears that when the sound absorption coefficient of a surface is small, it is possible to use an image source, taking into account complex surface impedance and complex absorption coefficient (Ingrad and Morse, 1968), which is simple to use. On the other hand, the reflection from highly absorbing surfaces

is much more complicated and necessitates an exact solution (Thomasson, 1977). The two types of solutions will be used simultaneously in the next section in order to find a solution which is within a prescribed margin of error and also reasonably easy and convenient for use.

### 3. Sound fields in enclosures with impedance surfaces

A typical room consists mostly of one or two highly absorbing surfaces (walls, ceilings, floors) while the other surfaces absorb very little acoustic energy - See figure 1 for the scheme. After establishing the sound reflection model of the various kinds of surfaces of the room, the comprehensive model of the sound field analysis in enclosure can be built. That includes all the reflecting surfaces, using geometrical acoustics and application of image sources. The following assumptions on which this model is based include:

* Using the principles of geometrical acoustics, reflections of sound from rigid surfaces are considered mirror reflections. It means that instead of considering the whole reflector, image sources represent its effect on the internal sound field.
* Reflections from low absorption surfaces are included in the cases where mirror images are considered. However, under such circumstances the reflection coefficient becomes complex.
* Calculation of reflections from a highly absorbing surface can be correctly derived only by using an exact solution. The pattern of reflection from such surfaces is usually very scattered, which justifies the application of the exact solution under those conditions in spite of the significant increase in the complexity of the mathematical expressions involved.
* After a sound ray hits a highly absorbing surface, in our model the higher order reflections from the neighboring reflecting surfaces are ignored, since only an insignificant error occurs This assumption is also a condition for terminating the series of reflections. Last limitation is schematically illustrated in figure 2.

As a result of the last assumption, the sound field of a room will be derived considering first and second order image sources, in accordance with its geometry. The first model of analysis of the velocity potential $\psi$ in the room assumes that the room surfaces are of a low absorption coefficient, which allows for application of an "approximate method" which will be described latter in this paper. Next we assume that one of the surfaces (number 1), is highly absorbent, which necessitates application of the "exact model". In our examples, the source which is located at $(x_0, y_0, z_0)$, radiates a pure tone of 300 Hz. We assume that the velocity of sound propagation in air is 340 m/s. The room has a box shape of the dimensions X, Y, Z and a receiver is located in it at (x, y, z).

The calculation initiates with estimation of the direct radiation of sound:

$$\psi = \frac{\exp(ikR)}{4\pi R}; \quad R = \sqrt{(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2}$$

Next, reflections of the first order from the 6 reflecting surfaces n=1 are calculated:

$$\psi'_n = C_m \frac{\exp(ikR'_n)}{4\pi R'_n}; \quad n = 1,2,\ldots,6$$

$$C_m = \frac{\varsigma_n \cos(\theta_n) - 1}{\varsigma_n \cos(\theta_n) + 1}; \quad \cos(\theta_n) = \frac{h_n}{R'_n};$$

$$h_1 = z + z_0;\ h_2 = y + y_0;\ h_3 = x + x_0;\ h_4 = 2Z - z_0 - z;\ h_5 = 2Y - y_0 - y;\ h_6 = 2X - x_0 - x$$

$$R'_1 = \sqrt{(x - x_0)^2 + (y - y_0)^2 + h_1^2}; \quad R'^1_2 = \sqrt{(x - x_0)^2 + h_2^2 + (z - z_0)^2};$$

$$R'_3 = \sqrt{h_3^2 + (y - y_0)^2 + (z - z_0)^2}; \quad R'_4 = \sqrt{(x - x_0)^2 + (y - y_0)^2 + h_4^2};$$

$$R'_5 = \sqrt{(x - x_0)^2 + h_5^2 + (z - z_0)^2}; \quad R'_6 = \sqrt{h_1^2 + (y - y_0)^2 + (z - z_0)^2}$$

Now the second order reflections are calculated as already given in Rosenhouse and Saski (1995 - Active 95) and chapter 3 for reflections from two adjacent reflecting surfaces. However, when the surfaces are parallel we have two possible reflections instead of one. See figure 3. The resulting formulae are:

$$\psi''_{mn} = C'_{mn} C'_m \frac{\exp(ikR''_{mn})}{4\pi R''_{mn}}; \quad n = 1,2,\ldots,6$$

$$C'_{mn} = \frac{\varsigma_m \cos(\theta'_m) - 1}{\varsigma_m \cos(\theta'_m) + 1}; \quad C'_m = \frac{\varsigma_n \cos(\theta'_n) - 1}{\varsigma_n \cos(\theta'_n) + 1}$$

For parallel surfaces we have the following source receiver distances:

$h_{14} = 2Z - z + z_0;\ h_{25} = 2Y - y + y_0;\ h_{36} = 2X - x + x_0;$

$h_{41} = 2Z - z_0 + z;\ h_{52} = 2Y - y_0 + y;\ h_6 = 2X - x_0 + x$

The corresponding distances are:

$$R''_{12} = \sqrt{(x - x_0)^2 + h_2^2 + h_1^2}; \qquad R''_{13} = \sqrt{h_3^2 + (y - y_0)^2 + h_1^2};$$

$$R''_{14} = \sqrt{(x - x_0)^2 + (y - y_0)^2 + h_{14}^2}; \quad R''_{15} = \sqrt{h_6^2 + (y - y_0)^2 + h_1^2};$$

$$R''_{23} = \sqrt{h_3^2 + h_2^2 + (z - z_0)^2}; \qquad R''_{24} = \sqrt{(x - x_0)^2 + h_2^2 + h_4^2}$$

$$R''_{25} = \sqrt{(x - x_0)^2 + h_{25}^2 + (z - z_0)^2}; \quad R''_{26} = \sqrt{h_6^2 + h_2^2 + (z - z_0)^2};$$

$$R''_{34} = \sqrt{h_3^2 + (y - y_0)^2 + h_4^2}; \qquad R''_{35} = \sqrt{h_3^2 + h_5^2 + (z - z_0)^2};$$

$$R''_{36} = \sqrt{h_{36}^2 + (y - y_0)^2 + (z - z_0)^2}; \quad R''_{41} = \sqrt{(x - x_0)^2 + h_5^2 + h_4^2}$$

$$R''_{45} = \sqrt{(x - x_0)^2 + h_5^2 + h_4^2}; \qquad R''_{46} = \sqrt{h_6^2 + (y - y_0)^2 + h_4^2};$$

$$R''_{52} = \sqrt{(x - x_0)^2 + h_{52}^2 + (z - z_0)^2}; \quad R''_{56} = \sqrt{h_6^2 + h_5^2 + (z - z_0)^2};$$

$$R''_{63} = \sqrt{h_{63}^2 + (y - y_0)^2 + (z - z_0)^2};$$

The angles to be calculated depend on the order of reflection of the rays. This implies that we have to compare for each case the geometrical measures, $D_{mn}$ and $D^0_{mn}$. Using those quantities we have for $D_{mn} > D^0_{mn}$:

$$\cos(\theta_{m'}) = \frac{h_m}{R''_{mn}} \quad \text{and} \quad \cos(\theta_n) = \frac{h_n}{R''_{mn}}$$

Otherwise:

$$\cos(\theta_{m'}) = \frac{h_n}{R''_{mn}} \quad \text{and} \quad \cos(\theta_n) = \frac{h_m}{R''_{mn}}$$

The definitions of $D_{mn}$ and $D^0_{mn}$ are respectively:

$$D_{12} = \frac{h_2}{h_1} z_0; D_{13} = \frac{h_3}{h_1} z_0; D_{15} = \frac{h_5}{h_1} z_0; D_{16} = \frac{h_6}{h_1} z_0$$

$$D_{23} = \frac{h_3}{h_2} z_0; D_{24} = \frac{h_4}{h_2} z_0; D_{26} = \frac{h_6}{h_2} z_0; D_{34} = \frac{h_4}{h_3} z_0$$

$$D_{35} = \frac{h_5}{h_3} z_0; D_{45} = \frac{h_5}{h_4}(Z - z_0); D_{46} = \frac{h_6}{h_4}(Z - z_0); D_{46} = \frac{h_6}{h_4}(Z - z_0); \quad \text{and}$$

$$D^0_{12} = y_0; D^0_{13} = x_0; D^0_{15} = Y - y_0; D^0_{16} = X - x_0; D^0_{23} = x_0; D^0_{24} = Z - z_0;$$

$$D^0_{26} = X - x_0; D^0_{34} = Z - z_0; D^0_{35} = Y - y_0; D^0_{45} = Y - y_0; D^0_{46} = X - x_0; D^0_{56} = X - x_0;$$

For parallel surfaces we have:

$$\cos(\theta'_m) = \cos(\theta'_n) = \frac{h_{mn}}{R''_{mn}}$$

The solution for a highly absorbent boundary (surface number 1 in figure 1 follows the exact representation - see Rosenhouse and Saski (1995 - Active 95). From figures 4, 5 we chose for illustration the situation given in figure 4. This situation yields the following expressions for calculating the resulting sound field. The total sound radiation is involved with four components of the velocity potential:

$$\psi = \psi_0 + \psi'_1 + \psi'_2 + \psi'_3$$

The first term stands for the direct radiation of sound:

$$\psi_0 = \frac{\exp(ikR)}{4\pi R}$$

The second expression is due to the first order reflection from surface 1:

$$\psi'_1 = \frac{\exp(ikR'_1)}{4\pi R'_1} + \psi_{SD} + \psi_B$$

The third expression is the first order reflection from the second surface:

$$\psi'_2 = \frac{\exp(ikR'_2)}{4\pi R'_2}$$

The last expression represents the reflection of the second order, namely:

$$\psi''_{12} = \left\{ \frac{\exp(ikR''_{12})}{4\pi R''_{12}} + \psi'_{SD} + \psi'_B \right\} C'_{r2}$$

To summarize, the reflection of sound from surface 1 is calculated following the "exact solution". In its expression, $\psi_{SD}$ and $\psi_B$ are terms from the solution shown by Thomasson, taking into account the effects of the real source S and its image $S'_1$. The reflection from surface 2 is based on the reflection coefficient $C_{r2}$ for the incidence angle $\theta_2$. In addition, $\psi'_{SD}$ and $\psi'_B$, are the corresponding terms of the solution

presented by Thomasson, taking into account the image sources $S_2'$ and $S_{12}''$. $S_2'$ is the real source suitable for surface 1, after being reduced by the reflection coefficient $C_{r2}'$ for the incidence angle $\theta_2'$.

## 4. Examples

The aforementioned theory is used now for calculation of some examples of reflections up to the second order, as discussed below. The active control simulation is represented here by the set of secondary sources and the superposition of their fields with that of the primary source. Surface 1 in figure 1 has the following specific impedance components:

$$\Re(\varsigma) = \Im(\varsigma) = 14$$

while those surfaces of low sound absorption acquire the values:

$$\Re(\varsigma) = \Im(\varsigma) = 1000$$

Figure 6 shows the sound field of a monopole within the enclosure. A radial sound radiation pattern is observed.It is seen that the decay of sound from the source outwards is not smooth any more, but some interference obstructions appear close to the location of the source, and practically disappear at a certain distance.

Now we introduce an additional secondary monopole source in anti-phase with the primary source, as depicted in figures 7, 8, 9. In figure 7 we observe that the sound level is reduced in a strip-like domain which is normal to the axis between the sources. If the additional auxiliary source is moved closer to the primary source, as in figure 8, stronger attenuation is achieved at points which are far from both sources, and the domain of higher noise level in the vicinity of the sources becomes smaller. If the additional monopole is located apart from the primary source, as in figure 9, the sound field mapping changes significantly, but no general contribution to the noise reduction is obtained by the secondary control unit.

An introduction of secondary dipole and quadrupole sources as shown in figures 10, 11 yields a slight reduction of the levels of the primary source field at regions far from the source.

Next, we consider an addition of two secondary monopole sources in the vicinity of the primary source as in figure 12, each of half the strength of the primary source and in anti-phase to it. Once again we see that the sound field is reduced significantly.

Finally, a test of five anti-phased monopoles each with fifth the strength of the primary source were placed in the vicinity of the primary source, as in figure 13. We see that in this case the sound field is greatly attenuated. This last result is in agreement with the theoretical approach of canceling the sound field created by a primary noise source. This theory suggests that a large number of anti-phased sources of exactly the same strength as of the primary source can cancel the

## 5. Summary

This paper studies active noise cancellation in a typical room. Such a room consists of one or two absorbing boundaries, while the rest of the boundaries are practically not absorbing. The model is based on the following assumptions:

* Reflection from a rigid boundary is taken as an "optical" mirror reflection. Here, the common "image source" is used.

* A reflection from a boundary with a low absorption coefficient is involved with a complex reflection coefficient. Yet, a usual "image source" can replace the reflecting surface.
* A reflection from a boundary of high sound absorption coefficient necessitates an application of the "exact" solution, which is characterized by a dispersed reflection.

Some of the results are:

- When an acoustic monopole acts in a room, the radial pattern fits an omni-directional source, but it is not characterized by the ever decreasing circular domains, emanating from the source. First, the levels decrease, and then overcome a number of disturbances, to become virtually constant at a certain distance from the source.
- adding secondary sources in an anti-phase leads to the following results:
  a. The sound levels are reduced in a strip normal to the axis between the sources.
  b. If the auxiliary sources are moved to the vicinity of the primary source a significant attenuation is observed in the far sound field, and the domain of high level in the near field is significantly reduced in size.
- an introduction of a secondary dipole source and a secondary quadrupole source in the vicinity of the primary source, leads to only a slight reduction in the levels in regions which are far from the source.
- A significant reduction in the sound field is obtained by adding two secondary monopole sources which are close to the location of the primary source and are half its strength.
- five monopole sources surrounding the primary source, each with a fifth of the strength of the primary source, cause a large attenuation of the sound field. This result is close to the theoretical approach of annihilating the sound field caused by a source of noise by surrounding it with a large number of auxiliary anti-phased sources.

## 6. References


1. . Antes, H., Applications in environmental noise, Chapter 11 in: Ciskowski, R.D., Brebbia, C.A. (editors), Boundary Element Methods in Acoustics, Computational Mechanics Publications- Southampton and Boston and Elsevier Applied Science - London and New-York, 1991
2. Ciskowski, R.D., Brebbia, C.A. (editors), Boundary Element Methods in Acoustics, Computational Mechanics Publications- Southampton and Boston and Elsevier Applied Science - London and New-York, 1991
3. Cremer, L., Mueller, H.A., Schultz, T.J., Principles and Applications of Room Acoustics, Applied Science Publishers, London & NY, 1982, volumes 1 and 2
4. Gensane, M., Santon, F., Predicting of sound fields in rooms of arbitrary shape: validity of the image sources method, Jr. Sound Vib., 63, 1979, 97-108
5. Juricic, H., Santon, F., Images et rayons sonores dans le calcul numerique des echogrammes, Acustica, 28, 1973, 77-89
6. Krokstad, A., Strom, S., Sorsdal, S., Calculating the acoustical room response by the use of ray tracing technique, Jr. Sound Vib., 8, 1968, 118-125
7. Krokstad, A., Strom, S., Sorsdal, S., Fifteen years' experience with computerized ray tracing, Appl. Acoust., 16, 1983, 291-312
8. Kuttruff, H., Room Acoustics, Applied Science Publishers, London, 1976
9. Migeot, J.-L., Active noise control of a car interior, SYSNOISE application note 26, 19.10.1993
10. Morse, P.M., Ingrad, K.U., Theoretical Acoustics, McGraw-Hill, NY, 1968, 259 -263

11.Nelson, P.A., Elliott, S.J., Active Control of Sound, Academic Press, London, 1992
12.Pierce, A.D., Acoustics, McGraw-Hill, NY, 1981. Ch. 8: Ray acoustics
13.Rosenhouse, G., Saski, E., Noise cancellation by use of arrays of active noise sources in the vicinity of impedance surfaces, Active 95, Newport beach, CA, USA, 6.81995, 685-696
14.Santon, F., Numerical prediction of echograms and of the intelligibility of speech in rooms, Jr. Acoust. Soc. Am., 59, 1976, 1399-1405
15.Sakurai, Y., The early reflections of the impulse response in an auditorium, J. Acoust. Soc. Jpn. (E), 8, 1987, 127-138
16.Schroeder, M.R., Digital simulation of sound transmission in reverberant spaces, Jr. Acoust. Am., 47, 1970, 424-431
17.Sekiguchi, K., Kimura, S., Sugiyama, T., Approximation of impulse response through computer simulation based on finite sound ray integration, J. Acoust. Soc. Am. Jpn. (E), 6, 1985, 103-115
18.Terai, T., Kawai, Y., BEM applications in room acoustics, Chapter 10 in: Ciskowski, R.D., Brebbia, C.A. (editors), Boundary Element Methods in Acoustics, Computational Mechanics Publications- Southampton and Boston and Elsevier Applied Science - London and New-York, 1991
20..Thomasson, S.I., Reflection of waves from a point source by an impedance boundary, Jr. Acoust. Soc. Am., 59, 4, 1976, 780-785

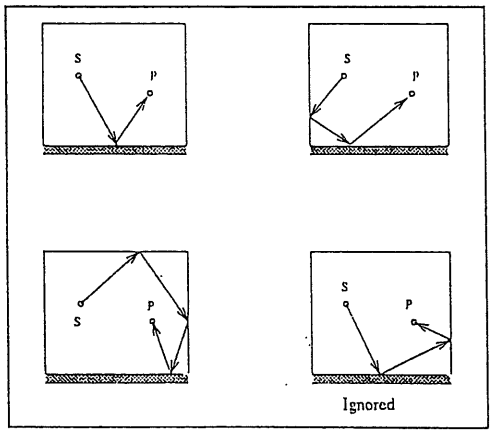



Figure 2. A model of a highly absorbing surface in a room and the relevant reflections

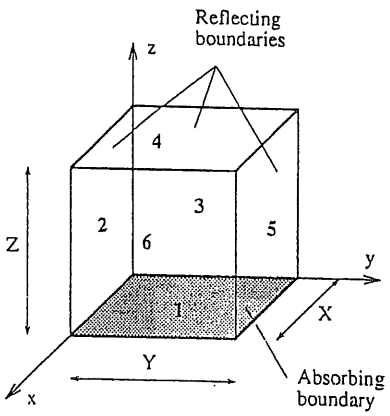



Figure 1. Configuration of a typical room

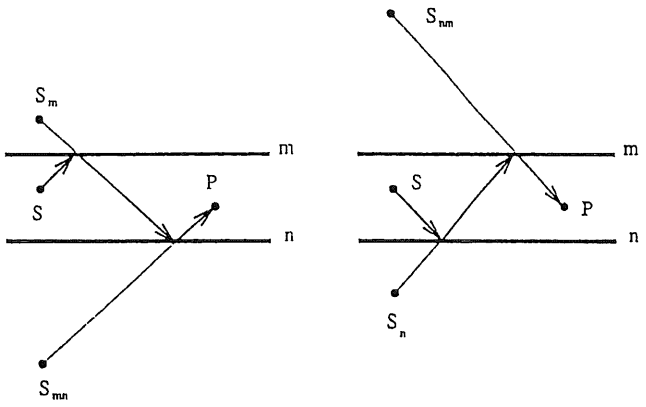



Figure 3. Two types of possible reflections between parallel walls

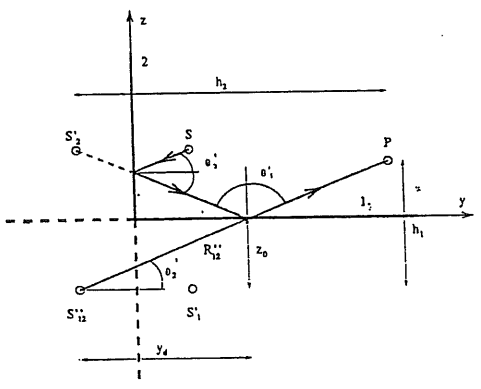



Figure 4. Second order effects: Case A

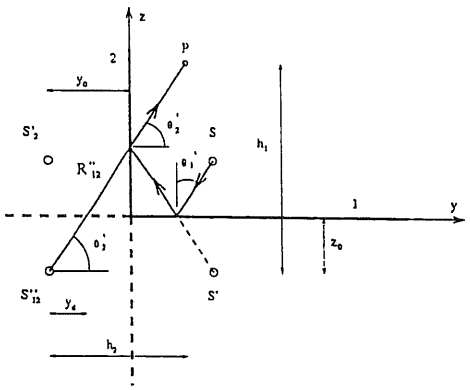



Figure 5 Second order effects: Case B

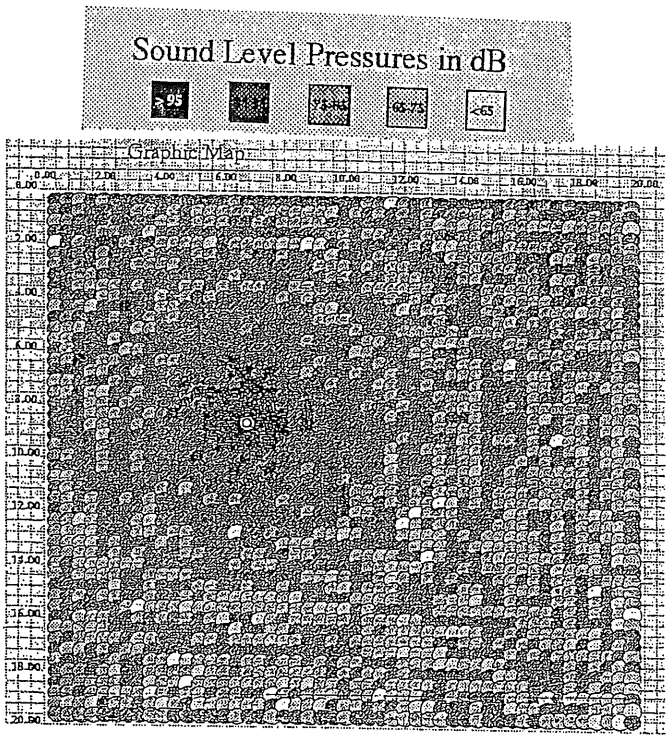



Figure 6. Primary sound field in an enclosure

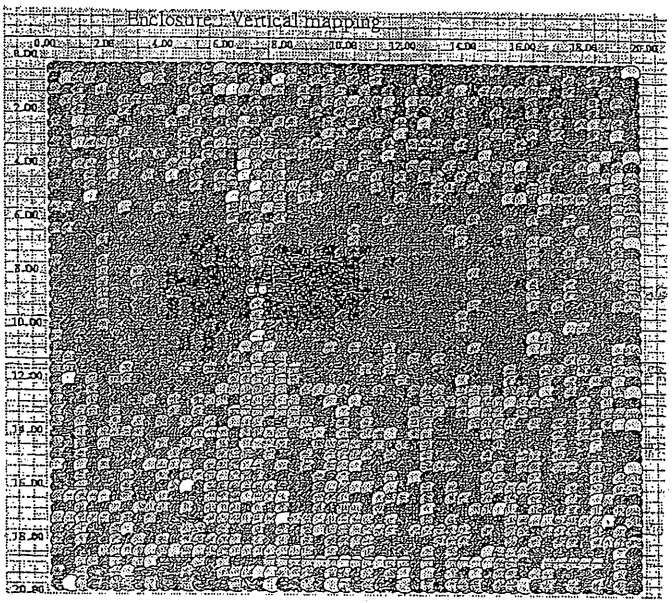



Figure 7. Sound field when a secondary monopole is close to the primary source

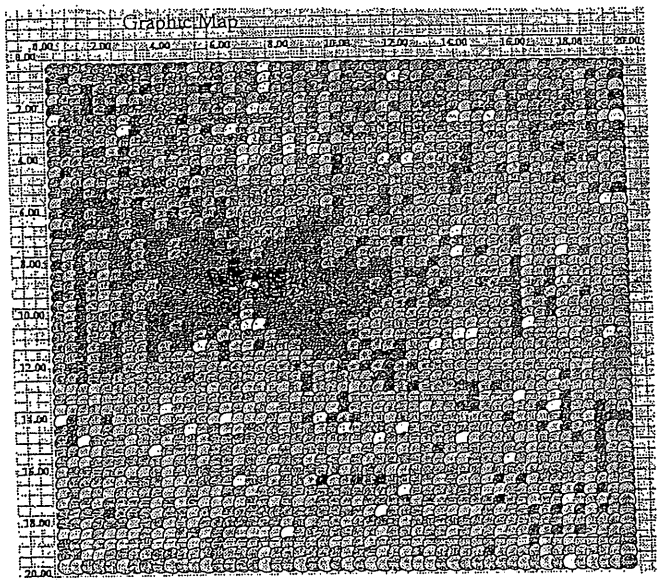



Figure 8. Sound field when a secondary monopole is very close to the primary source

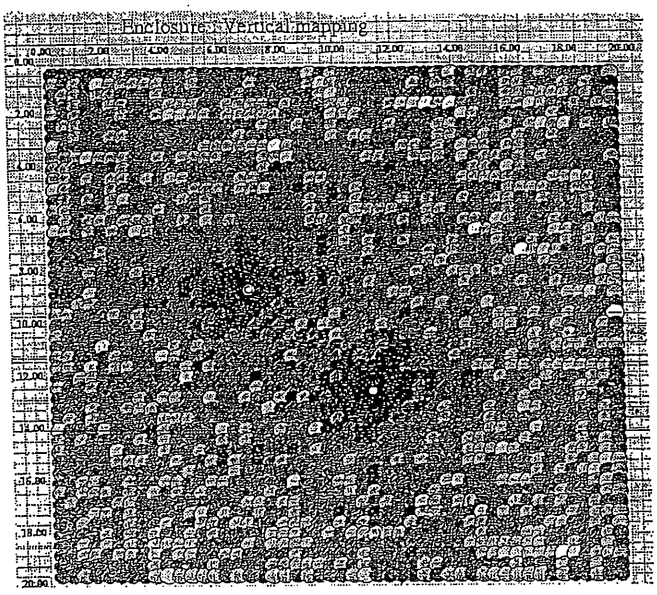



Figure 9. Sound field when a secondary monopole is far from the primary source

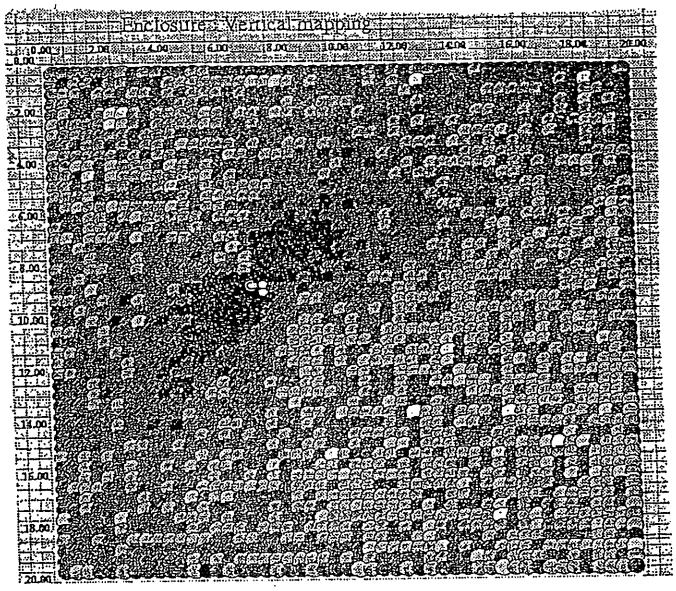



Figure 10. Sound field when a secondary dipole is close to the primary source

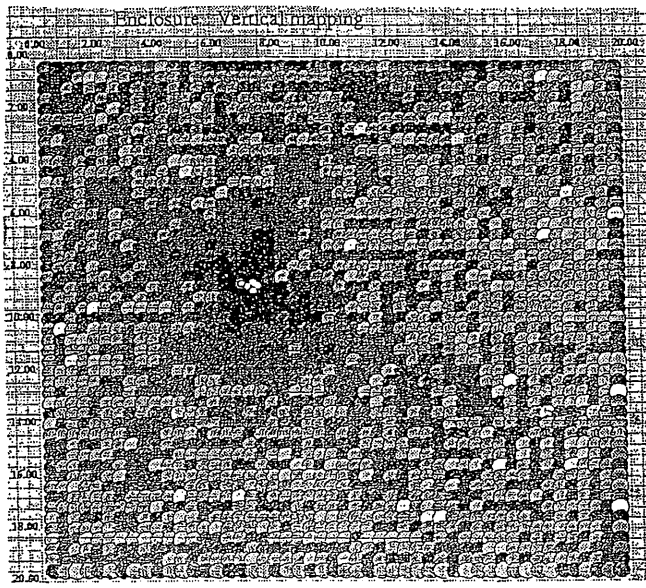



Figure 11. Sound field when a secondary quardrupole is close to the primary source

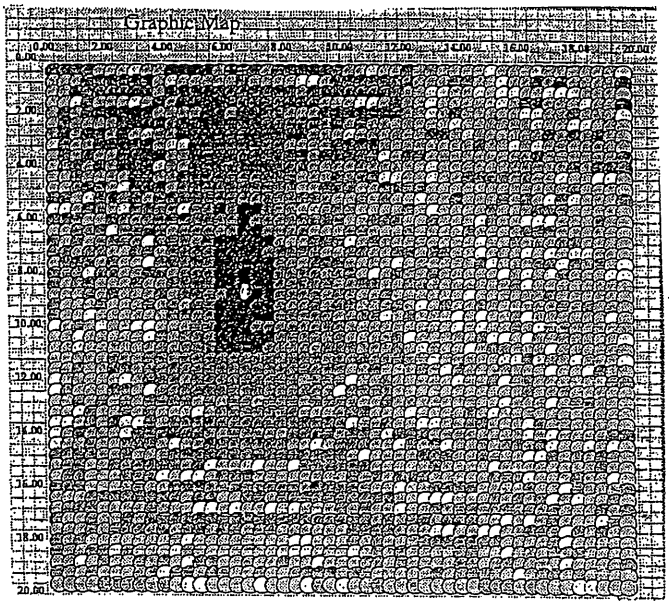



Figure 12. Sound field when two secondary monopoles of half the strength of the primary source each are close to the primary source

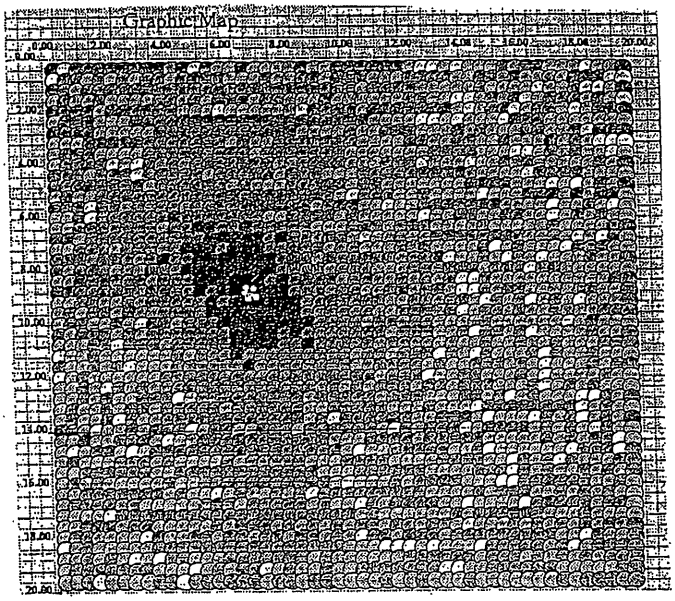



Figure 13. Sound field when five secondary monopoles of fifth the strength of the primary source each are close to the primary source

Theoretical and Computational Acoustics
D. Lee, Y-H Pao, M.H. Schultz, and Y-C Teng (Editors)


# Model-Based Internal Wave Processing

J.V. Candy and D.H. Chambers
Lawrence Livermore National Laboratory

## MODEL-BASED INTERNAL WAVE PROCESSING

*J. V. Candy* *D. H. Chambers*
*Lawrence Livermore National Laboratory*

A model-based approach is proposed to solve the oceanic internal wave signal processing problem that is based on state-space representations of the normal-mode vertical velocity and plane wave horizontal velocity propagation models. It is shown that these representations can be utilized to spatially propagate the modal (depth) vertical velocity functions given the basic parameters (wave numbers, Brunt-Vaisala frequency profile etc.) developed from the solution of the associated boundary value problem as well as the horizontal velocity components. Based on this framework, investigations are made of model-based solutions to the signal enhancement problem for internal waves.

## I. INTRODUCTION

When operating in a stratified environment like the upper ocean with relatively sharp density gradients, then any excitation that disturbs the pycnocline (density profile) will generate *internal waves* that propagate away from this region (Apel,[1]). Internal waves are volume gravity waves having maximum vertical displacement typically at a plane where the density gradient are largest and are detectable far above and below this interface (Clay,[2]). They can be generated from tidal flow against islands, sea mounts and continental shelf edges or surface/internal wave interactions created displacements in the pyncocline. For instance, a ship traveling along the surface of a stratified ocean creates various visible wakes: the turbulent or centerline wake, the Kelvin wake and, of most interest in this work, the surface generated internal waves. Internal waves have been measured experimentally both in controlled environments as well as the open ocean (Garrett,[3]) and observed using synthetic aperture radar processing techniques from satellite imagery (Alpers,[4]; Thompson,[5]). From the scientific viewpoint, it is of high interest to understand the effect of internal waves on acoustic propagation in the ocean [2] as well as the ability to measure their effect directly using current sensor technology. Military applications are obvious, since a submerged body moving through the ocean environment disturbing the pyncocline generates internal wave signatures.

The inclusion of a propagation model in any oceanic signal processing scheme provides a means of introducing environmental information in a self-consistent manner. Recent work in ocean acoustics (Candy and Sullivan,[6]) has shown that a propagation model can be imbedded into a signal processing scheme to solve various enhancement, localization and detection problems. In this paper, we propose a model-based approach to the internal wave signal processing problem founded on a state-space representation of the normal mode and plane wave models of propagation. Specifically, using the normal mode model of the wave velocity field, the vertical velocity modal functions and the horizontal velocity can be estimated from noisy sensor array measurements in the following way. First, the propagation model is cast into state-space form. It is shown that this representation can be used to propagate the modal functions, given the basic parameters (wave numbers etc.) developed from the solution of the associated boundary value problem. There are basically two sets of equations in this representation: the state equation and the measurement

equation. The state equation describes the evolution in space of the vertical velocity modal and horizontal velocity functions, whereas the measurement equation relates the states to the actual array measurements. In the stochastic case, an approximate Gauss-Markov representation evolves. Once this framework is established, as will be shown in what follows, we investigate model-based solutions to the signal enhancement. We characterize a realizable, recursive processor as shown in Figure 1 (Candy,[7]; Jazwinski,[8]).

In the next section the state space representation of the internal wave propagation model is developed for both vertical and horizontal propagation. In the vertical, the so-called "forward propagator" is defined, while in the horizontal, plane wave propagators adequately characterize the propagation. Next the measurement equations are developed for both a vertical and then horizontal sensor array. The state-space model is then explicitly formulated for the case of all model parameters known where estimates of the velocity field and the modal functions are made and various noise models can be exercised in order to simulate different sources of modal and measurement noise. Next a dispersive internal wave model-based processor is designed which uses a horizontal array of sensors modeling the configuration employed in the Loch Linnhe experiments performed in Scotland during the summer of 1994. It is shown that the internal wave can be successfully estimated from noisy measurements. The final section summarizes the results.

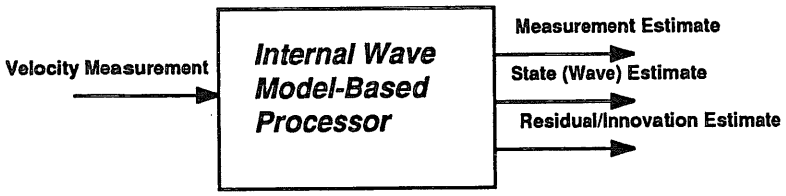



**Figure 1.** Basic Internal Wave Model-Based Processor.

## II. INTERNAL WAVE STATE-SPACE PROPAGATION MODELS

In this section we investigate the feasibility of developing state-space propagation models from the corresponding internal wave dynamics. The development of the wave equation associated with internal wave propagation evolves from the small perturbation momentum equations under the assumptions of linearity, incompressibility, zero mean shear, a Boussinesq fluid, flat or slowly varying ocean bottom, and a horizontally homogeneous or slowly varying density field. Under these assumptions, the small perturbation component momentum equations governing the propagation of the *vector* velocity field defined are by $\chi(x,y,z,t) := [u(x,y,z,t)v(x,y,z,t)w(x,y,z,t)]'$ where $u$, $v$, $w$ represent the respective *on-track velocity*, *cross-track velocity*, and *vertical velocity* components of the vector velocity field.

The solution to the homogeneous wave equation is accomplished by using the separation of variables approach (see Apel,[1]) with velocity function given by $w(x,y,z,t) = \mu(x)\nu(y)\phi(z)\tau(t)$.

where plane wave propagation in the respective horizontial dimensions, $\mu$ and $\nu$ are identical, $\phi$ is the vertical distribution and $\tau$ is the corresponding temporal function. Substituting, collecting terms, etc., it is possible to show that the function associated with each independent variable can be separated yielding a set of coupled ordinary differential equations with the corresponding separation constants shown as

$$\begin{aligned} \frac{d^2}{dx^2}\mu(x) + \kappa_x^2\mu(x) &= 0 \\ \frac{d^2}{dy^2}\nu(y) + \kappa_y^2\nu(y) &= 0 \\ \frac{d^2}{dz^2}\phi(z) + \kappa_h^2\left(\frac{N^2(z)}{\omega^2} - 1\right)\phi(z) &= 0 \\ \frac{d^2}{dt^2}\tau(t) + \omega^2\tau(t) &= 0 \end{aligned} \tag{1}$$

where $\kappa_x$ and $\kappa_y$ are the horizontal wave numbers with $\kappa_h^2 = \kappa_x^2 + \kappa_y^2$, and $\omega^2$ the temporal frequency (a function of $\kappa_h$). These relations describe the dynamics of the vertical velocity component of an internal wave in a constant depth ocean. It is clear that the first two and last equations above admit harmonic solutions, that is, $\mu(x) = e^{i\kappa_x x}$ $\nu(y) = e^{i\kappa_y y}$ $\tau(t) = e^{-i\omega t}$ while the depth relation is an eigen-equation in $z$ with

$$\frac{d^2}{dz^2}\phi_m(z) + \kappa_h^2\left(\frac{N^2(z)}{\omega_m^2} - 1\right)\phi_m(z) = 0, \qquad m = 1, \cdots, M \tag{2}$$

whose eigen-solutions $\{\phi_m(z)\}$ are the so-called *modal functions* associated with the vertical velocity component, $\kappa_h$ is the *horizontal wavenumber*, $N(z)$ is the *Brunt-Vaisala frequency* profile (BVP) and $\omega_m$ is the corresponding *eigenvalue* associated wtih the $m^{th}$-modal function. State-space models will eventually be employed as "forward" propagators in model-based signal processing schemes (Candy,[7]). Note that this approach does *not* offer a new solution to the resulting boundary value problem, but, in fact, it requires that solution be available *a-priori* in order to propagate the normal-modes recursively in an initial value scheme.

Since Eq. 2 is a linear, space-varying coefficient (for each layer) differential equation, it can easily be placed in state-space form where the state-vector for each mode is defined as $\mathbf{x}_m := [\phi_m(z) \quad \frac{d}{dz}\phi_m(z)]'$. Examining mode propagation in more detail, we see that each mode is characterized by a set of ordinary differential equations which can be written as

$$\frac{d}{dz}\mathbf{x}_m(z) = A_m(z)\mathbf{x}_m(z) + B_m(z)\mathbf{u}_m(z) = \begin{bmatrix} 0 & 1 \\ -\kappa_h^2\left(\frac{N^2(z)}{\omega_m^2} - 1\right) & 0 \end{bmatrix}\mathbf{x}_m(z) + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\mathbf{u}_m(z) \qquad (3)$$

where $\mathbf{u}(z) = \frac{ds(z)}{dz}$ is the source driving function (a scalar in our problem). With the parameters $[\{\kappa_h\}, N(z), \{\omega_m\}]$ known, we see that we have constructed a *forward propagator* for this problem driven by the source and initial condition vector, $\mathbf{x}_m(z_0)$ for each mode. From linear systems theory (see Chen,[9]), it is well-known that the general solution to this equation is governed by the state-transition matrix, $\Phi(z, z_0)$, where the state equation is solved by

$$\mathbf{x}_m(z) = \Phi_m(z, z_0)\mathbf{x}_m(z_0) + \int_{z_0}^{z} \Phi_m(z, \alpha)B_m(\alpha)\mathbf{u}_m(\alpha)d\alpha, \qquad m = 1, \ldots, M \qquad (4)$$

and the state transition matrix satisfies

$$\frac{d}{dz}\Phi_m(z, z_0) = A_m(z)\Phi_m(z, z_0), \qquad m = 1, \ldots, M \qquad (5)$$

with $\Phi_m(z_0, z_0) = I$.

If we include all of the $M$-modes in the model, then we obtain

$$\frac{d}{dz}\mathbf{x}_w(z) = \begin{bmatrix} A_1(z) & \cdots & O \\ \vdots & & \vdots \\ O & \cdots & A_M(z) \end{bmatrix}\mathbf{x}_w(z) + \begin{bmatrix} B_1(z) \\ \vdots \\ B_M(z) \end{bmatrix}\mathbf{u}(z) \qquad (6)$$

or simply

$$\frac{d}{dz}\mathbf{x}_w(z) = A(z)\mathbf{x}_w(z) + B(z)\mathbf{u}(z) \qquad (7)$$

The general solution to these modal state-space vertical velocity propagation equations are given by Eq. 10 over all the modes as

$$\mathbf{x}(z) = \Phi(z, z_0)\mathbf{x}(z_0) + \int_{z_0}^{z} \Phi(z, \alpha)B(\alpha)\mathbf{u}(\alpha)d\alpha \qquad (8)$$

and as before the state transition matrix satisfies

$$\frac{d}{dz}\Phi(z,z_0) = A(z)\Phi(z,z_0), \qquad \Phi(z_0,z_0) = I \tag{9}$$

with $\Phi(z,z_0) = diag[\Phi_1(z,z_0),\cdots,\Phi_M(z,z_0)]$.

The solutions in the other horizontal dimensions, $x$, $y$ are also possible in state-space form, if desired, and take on the same form for each relation with the separation constants interchanged. For instance, the component $\nu(y)$ satisfies the following state equations with the state vector defined as $\mathbf{x}(y) := [\nu(y) \quad \frac{d}{dy}\nu(y)]'$

$$\frac{d}{dy}\mathbf{x}(y) = A_y\mathbf{x}(y) \tag{10}$$

where

$$A_y := \begin{bmatrix} 0 & 1 \\ -\kappa_y^2 & 0 \end{bmatrix} \tag{11}$$

which gives the plane wave solution of Eq. 7 above. We note that in this formulation the *strain rate*, which is an important quantity for radar images, appears as a component of the state vector. Thus, we see that using the state-space formalism enables us to characterize the propagation (forward) of the vertical internal wave velocity component which will prove useful for processing. Next we consider characterizing the corresponding measurement system.

Sensor technology enables us to measure components of internal wave cross-track velocity at depth $z$, $v(x,y,z,t)$ as well as on-track velocity $u(x,y,z,t)$. Let us investigate the deployment of a horizontal array of current meter sensors positioned to measure the cross-track velocity. Actual sensors sample the cross-track velocity in time at given spatial positions. To obtain the sensor output we must specify the functions $\mu(x)$, $\nu(y)$, and $\tau(t)$ as well as $\phi_v(z)$. The measurement at the $n^{th}$ horizontal sensor at the location ($x_n$,$y_n$,$z_\ell$) is given by

$$v(x_n,y_n,z_\ell,t) = \mu(x_n)\nu(y_n)\phi_v(z_\ell)\tau(t) = A\left(\frac{\kappa_y}{\kappa_h^2}\frac{d\phi(z_\ell)}{dz}e^{i(\kappa_x x_n+\kappa_y y_n-\omega(\kappa_h)t+\pi/2)}\right) \tag{12}$$

where $A$ is a complex amplitude, and $\omega(\kappa_h)$ is the angular frequency which is a function of the wave

number magnitude given by, $\kappa_h = \sqrt{\kappa_x^2 + \kappa_y^2}$. The sensor measures the real part of the velocity field so we can write (using the magnitude phase form of $A$)

$$v(x_n, y_n, z_\ell, t) = -|A| \frac{\kappa_y}{\kappa_h^2} \frac{d\phi(z_\ell)}{dz} \sin(\kappa_x x_n + \kappa_y y_n - \omega(\kappa_h)t + \theta) \tag{13}$$

Since we have $M$-modal velocities, then (as before)

$$v(x_n, y_n, z_\ell, t) = \mathbf{C_y^T x}(z_\ell) \tag{14}$$

with $\mathbf{C_y^T} = \frac{\kappa_y}{\kappa_h^2}[0 \;\; \sigma_1(x,y,t) \;\; |\cdots| \;\; 0 \;\; \sigma_M(x,y,t)]$, and $\sigma_m(x_m, y_m, t) = |A_m| \sin(\kappa_x x_n + \kappa_y y_n - \omega_m(\kappa_h)t + \theta_m)$.

Incorporating *both* the vertical and horizontal state-space representations simultaneously leads to the set of coupled equations:

$$\begin{aligned} \tfrac{d}{dz}\mathbf{x}_w(z) &= A(z)\mathbf{x}_w(z) + B(z)\mathbf{u}(z) \\ \phi(z_\ell) &= \mathbf{C}_\phi^T \mathbf{x}_w(z_\ell) \end{aligned} \tag{15}$$

where $\mathbf{C}_\phi^T = \frac{\kappa_y}{\kappa_h^2} e^{i\pi/2}[0 \;\; 1 \;\; |\cdots| \;\; 0 \;\; 1]$

$$\begin{aligned} \tfrac{d}{dy}\mathbf{x}(y) &= A_y \mathbf{x}(y) \\ \nu(y_n) &= \mathbf{C}_\nu^T \mathbf{x}(y_n) \end{aligned} \tag{16}$$

with $\mathbf{C}_\nu^T = [1 \;\; 0 \;\; |\cdots| \;\; 1 \;\; 0]$ and the overall velocity measurment is

$$\begin{aligned} v(x, y_n, z_\ell, t) &= \nu^T(y_n)\mathbf{C}_v(x,t)\phi_v(z_\ell) \\ \mathbf{C}_v(x,t) &= diag[A_1\mu_1(x)\tau_1(t) \cdots A_M\mu_M(x)\tau_M(t)] \end{aligned} \tag{17}$$

which is the equivalent of a two-dimensional measurement taking into account both vertical and horizontal velocity measurements. So we see that the state-space representations enables alot of flexibility in modeling both the phenomenology (internal wave dynamics) as well as the accompanying measurement systems.

This constitutes a complete deterministic representation of the normal-mode and plane wave models in state-space form. However, since propagation in the ocean is effected by inhomogeneities in the water, slow variations in the BVP and motion of the surface, the model must be modified in order to include these effects. This can be done in a natural way by placing the model into a Gauss-Markov representation which includes the second order statistics of the measurement as well as the velocity noise. The measurement noise can represent the near-field velocity noise field, flow noise on the current meter and electronic noise. The modal noise can represent BVP errors, distant shipping noise, errors in the boundary conditions, sea state effects and ocean inhomogeneities. Besides the ability to lump the various noise terms into the model, the Gauss-Markov representation provides a framework in which the various statistics associated with the model such as the means and their associated covariances can be computed. This completes the state-space respresentations of internal wave dynamics. Next we investigate horizontal solutions using dispersive plane wave representations in the next section, develop the model-based processor, perform various simulations and apply the results to actual measurement data.

## III. CROSS-TRACK VELOCITY ESTIMATION

In this section we use a model developed for the nondispersive case [12] and extend it to the dispersive by including more sophisticated dispersion models based on some empirical results [13] and approximating the appropriate temporal derivative. To extend these results (approximately) to the dispersive case, that is, the case where the temporal frequency is no longer constant (narrowband) but varies temporally, we return to the original development of the plane wave processor of the previous section. Recall that the states were defined as, $\mathbf{x}(t) := [v(t)\frac{d}{dt}v(t)]^T$ and now from Eq. 45 with both frequency and wave number temporal functions, we have

$$v(y,t) = \cos(\kappa_y(t)y - \omega(t)t) = \cos\kappa_y(t)y\ \cos\omega(t)t + \sin\kappa_y(t)y\ \sin\omega(t)t \tag{18}$$

From our definition of the temporal state vector, we have that $\tau(t) := \cos\omega(t)t$ and approximate its corresponding derivative using

$$\frac{d}{dt}\tau(t) = -\omega(t)\left[1 + \frac{\dot{\omega}(t)}{\omega(t)}t\right]\sin\omega(t)t \approx -\omega(t)\sin\omega(t)t \tag{19}$$

Defining the notation for the dispersive case, we have for the $\ell^{th}$-sensor

$$\alpha_\ell(t) := \cos\kappa_y(t)y_\ell, \qquad \beta_\ell(t) := \frac{\sin\kappa_y(t)y_\ell}{\omega(t)} \tag{20}$$

therefore, the measurement becomes

$$v(y_\ell, t) = \alpha_\ell(t)\cos\omega(t)t - \beta_\ell(t)\sin\omega(t)t = \alpha_\ell(t)\tau_1(t) - \beta_\ell(t)\omega(t)\tau_2(t) \tag{21}$$

Next we must develop the relations for the dispersive case. For a plane wave propagating in the y-direction, the value of wave number at time $t$ is a function of *group speed* $c_g(\kappa_y)$, that is,

$$y_\ell = c_g\left(\kappa_y(t)\right) t \tag{22}$$

with corresponding temporal frequency $\omega(\kappa_y(t), t)$. Using the Barber approximation [13] for internal wave dispersion and group speed, we have

$$\omega(\kappa_y(t), t) = \frac{\kappa_y(t) C_o}{1 + \kappa_y(t) C_o / N_o} \tag{23}$$

and

$$c_g(\kappa_y(t)) = \frac{C_o}{(1 + \kappa_y(t) C_o / N_o)^2} \tag{24}$$

Substituting Eq. 24 for $c_g$ in Eq. 22 and solving for the wave number at the $\ell^{th}$-sensor, we obtain

$$\kappa_y(t) = \frac{N_o}{C_o}\left(\sqrt{\frac{C_o t}{y_\ell}} - 1\right) \tag{25}$$

Substituting this result into Eq. 23, the corresponding temporal frequency is given by

$$\omega_\ell(t) = N_o\left(1 - \sqrt{\frac{y_\ell}{C_o t}}\right) \qquad t > \frac{y_\ell}{C_o} \tag{26}$$

Further assuming a bandlimited pulse with minimum frequency $\omega_o$ and observation time at the $\ell^{th}$-sensor given by

$$t_{o\ell} = \frac{y_\ell}{C_o(1 - \omega_o / N_o)^2} \tag{27}$$

and defining the time index as $t = t_{o\ell} + t_k$, $t_k > 0$, then we can obtain an expression for the temporal frequency as

$$\omega_\ell(t_k) = N_o \left( 1 - \frac{1 - \omega_o/N_o}{\sqrt{1 + \frac{C_o t_k}{y_\ell} (1 - \omega_o/N_o)^2}} \right) \qquad t_k > 0 \tag{28}$$

It should be denoted that this dispersive model is based on the fact that the onset has been detected, that is, the appropriate time delays estimated. In our case we use Eq. 27 to calculate the appropriate delays once an onset is established at the specified sensor. When we show the measured experimental data, it should be noted that we selected this onset time and then applied the processor. A more sophisticated approach would be to find the "optimal" onset based on these models, but this is left for future work. Now if we return to the central difference form for the equations, we obtain the approximate *spatially-constrained, dispersive* state-space representation

$$\mathbf{x}(t_{k+1}) := \begin{bmatrix} 1 - \omega(t_k)^2 \Delta t^2 & \Delta t & | & & \\ -\omega(t_k)^2 \Delta t & 1 & | & O & \\ - & - & & - & - \\ & & \ddots & & \\ - & - & & - & - \\ & O & | & 1 - \omega(t_k)^2 \Delta t^2 & \Delta t \\ & & | & \omega(t_k)^2 \Delta t & 1 \end{bmatrix} \mathbf{x}(t_k), \tag{29}$$

and with corresponding discrete cross-track velocity array measurements (in this coordinate system)

$$\mathbf{v}(t_k) := \begin{bmatrix} \alpha_1(t_k) & -\beta_1(t_k) & | & O & \\ - & - & & - & - \\ & & \ddots & & \\ - & - & & - & - \\ & O & | & \alpha_L(t_k) & -\beta_L(t_k) \end{bmatrix} \mathbf{x}(t_k) \tag{30}$$

In this coordinate system the initial condition vector is given by

$$\mathbf{x}(t_o) = \begin{bmatrix} 1 \\ 0 \\ - - - \\ \vdots \\ - - - \\ \cos(\kappa_y(L-1)\Delta) + \omega(t_k)\Delta t \sin(\kappa_y(L-1)\Delta) \\ \omega(t_k) \sin(\kappa_y(L-1)\Delta) \end{bmatrix}, \tag{31}$$

With this approximate dynamic dispersive model, we developed a Gauss-Markov simulation with the corresponding frequency and wave numbers varying temporally according to the dispersion relation. This presents no particular problem in the Gauss-Markov formulation, since it can support nonlinear dynamics as well as nonstationary statistics. We performed the simulation with the same basic set of parameters and the results for a $-13dB$ SNR were quite successful indicating an optimal design.

Next we consider the application of the dispersive model-based processor to an internal wave field experiment performed in Loch Linnhe, Scotland in September of 1994 (Mantrom,14]; Robey,[15]). The general objective of this experiment is to examine the relationship between modulations observed in radar images and ship generated, internal waves by fusing the data with that measured by an array of current meter sensors. Loch Linnhe is a narrow, salt water estuary located on the west coast of Scotland which possesses favorable subsurface environmental conditions (stratification) which are conducive for the generation and propagation of internal waves. The internal waves are generated by surface ships with concurrent oceanographic measurements (temperature, salinity, BV frequency, meterology, etc.) performed by other vessels in the Loch. Internal waves are typically imaged by airborne radar systems employing synthetic aperture radar techniques; however, this experiment concentrated on real aperture radar images of the waves from low-grazing angle marine radars positioned on the surrounding hillsides and an array of current meter sensors which provide measurements of hydrodynamic currents associated with the internal waves. The Lawrence Livermore National Laboratory (LLNL) current meter array (CMA) measurement platform consists of ten (10) standard S4 current meters uniformly spaced at $3.75m$, longitudinally, creating a $34m$ aperture, positioned at a depth of $2m$ and sampled at a $10sec$ rate. The CMA is positioned to measure both cross-track and on-track velocities. A typical measurement is shown in Figure 3a, where we see the output of each sensor channel typifying the phase front of an internal wave and subsequent ambient noise.

We processed the Loch Linnhe expermental data using the dispersive rather than plane wave processor and the results are shown in Figure 2. Here we see the differences in using the dispersive over the nondispersive approach. First, the processing doesn't really begin until the "onset" has begun and then the MBP utilizes the dispersive model, until that time a simple random walk model is employed (constant + gaussian noise). Next we note that each internal wave component temporal signal is clearly visible especially those which meet the onset criterion established by the dispersion relation. Those non-propagating wavefronts are attenuated quite heavily. We see the enhanced internal wave propagating across the array based on the onset and all other events attenuated significantly. We note the overall selectivity of the MBP to events that "match" both its dispersive signal model as well as onset. This completes the development of the dispersive MBP for internal wave enhancement and its application to experimental data, we summarize our results and place them in perspective.

## IV. DISCUSSION

We have developed the model-based approach to processing internal waves. It was shown

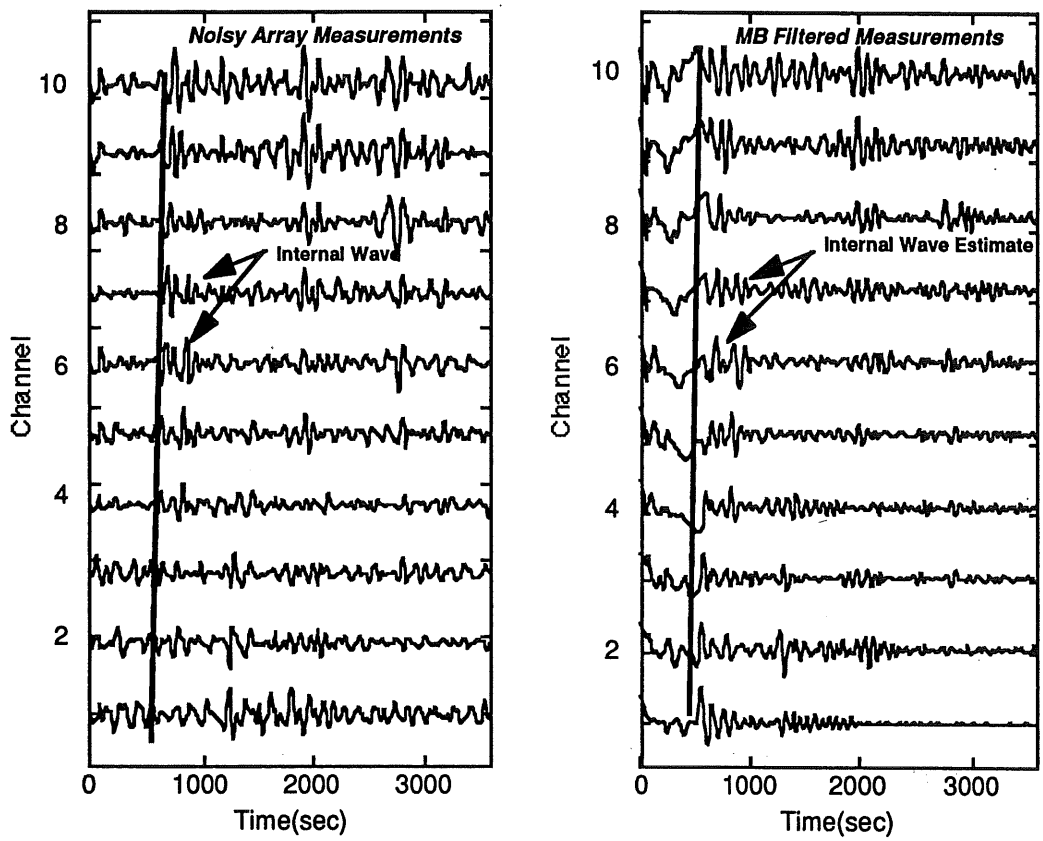



**Figure 2.** Dispersive Wave MBP Design for Loch Linnhe Experimental Array (10-elements): (a) Preprocessed Sensor Signals (s7r1). (b) Enhanced Sensor Estimates.

how to develop the underlying state-space representations of the internal wave propagation model. In the vertical, a normal-mode propagation model resulted, while in the corresponding horizontal plane wave propagation models were developed. Since our primary motivation was to process measurements from a horizontal array of current meter sensors, we developed the horizontal (plane wave) processors for both the nondispersive and dispersive cases. After using the model to develop Gauss-Markov stochastic simulations, the model-based processors were designed using the extended Kalman filter (EKF) algorithm [7] for signal enhancement. After designing the MBP, the minimum variance estimates were shown capable of extracting the desired wavefronts from simulated data quite effectively.

We then applied the various MBP to sets of data gathered from experiments performed in Loch Linnhe, Scotland during the summer of 1994 [14,15] and the results are quite promising. It was shown that the nondispersive plane wave MBP was capable of selecting all wave fronts that matched the correct temporal frequency and prescribed onset arrival times while attenuating all others. The dispersive MBP design approximated the true dispersive solution for slow changes in temporal frequency and appears to be a more effective approach, since it depends heavily on the pre-specified wavefront onset arrivals at the array and therefore is more discriminating then the nondispersive design.

## References


[1] J. R. Apel, *Principles of Ocean Physics*, New York:Academic Press, 1987.

[2] C. S. Clay, and H. Medwin, *Acoustical Oceanography.* New York:Wiley, 1977.

[3] C. Garrett and W. Munk, "Internal waves in the ocean." *Ann. Rev. Fluid Mech.*, **11**, 339-369, 1979.

[4] W. Alpers, "Theory of radar imaging of internal waves." *Nature*, **314**, 245-247, 1985.

[5] D. R. Thompson and R. F. Gasparovic, "Intensity modulation in SAR images of internal waves". *Nature*, **320**, 345-348, 1986.

[6] J. V. Candy and E. J. Sullivan. "Ocean acoustic signal processing: a model-based approach." *J. Acoust. Soc. Am.*, **92**, (12), 3185-3201, 1992.

[7] J. V. Candy, *Signal Processing: The Model-Based Approach.* New York:McGraw-Hill, 1986.

[8] A. Jazwinski, *Stochastic Processes and Filtering Theory.* New York:Academic Press, 1970.

[9] O. M. Phillips, *The Dynamics of the Upper Ocean*, Cambridge:Cambridge Univ. Press, 1977.

[10] C. T. Chen, *Introduction to Linear System Theory.* New York:Holt, Rinehart and Winston, 1970.

[11] G. J. Bierman, *Factorization Methods of Discrete Sequential Estimation.* New York:Academic Press, 1977.

[12] E. J. Sullivan and J. V. Candy. "Passive synthetic aperture processing as a Kalman filtering problem." *J. Acoust. Soc. Am.*, **95**, (5), Part 2, 2953, 1994.

[13] J. V. Candy and P. M. Candy, "SSPACK_PC: A model-based signal processing package on personal computers," *DSP Applic.*, **2**,(3), 33-42, 1993.

[14] B. C. Barber, "On the dispersion relation for trapped internal waves," *J. Fluid Mechanics.*, **252**, 31-49, 1993.

[15] D. D. Mantrom, "Loch Linnhe '94: Test Operations Description and On-Site Analysis US Activities," *LLNL Report.*,UCRL-ID-119197, 1994.

[16] H. F. Robey and D. L. Ravizza, "Loch Linnhe Experiment 1994: Background Stratification and Shear Measurements Part 1: Profile Summary and Dispersion Relations", *LLNL Report.*,UCRL-ID-119352, 1994.

**Acknowledgements:** We would like to acknowledge the motivation and support of Dr. R. Twogood, Program Leader LLNL Advanced Applications Program and Mr. James Brase, Project Leader for Radar Ocean Imaging. We also gratfully acknowledge our colleagues for the many clarifications and technical discussions: Dr. D. Holliday, Mr. H. Jones, Mr. D. Mantrom and Dr. H. Robey. This work was performed under the auspices of the Department of Energy by the Lawrence Livermore National Laboratory under contract W-7405-Eng-48 and supported by the Office of the Secretary of Defense.