Properties for inverse analysis of sound propagation in simple oceanic waveguides

by Carl Wunsch[1]

ABSTRACT

For a source and receiver on the axis of a deep sound channel we solve the two point boundary value problem for the arriving sound pulse using the classical ray theory approximation. The procedure is a generalization of that of Munk (1974) and we emphasize the determination of parameter changes in the sound channel by inverse techniques. The 'canonical' profile of Munk is shown to give an arrival structure qualitatively different from an observed profile. The range of parameters for which the axial ray will appear last (as is conventional), and first, is found; we also show that it is possible to have two rays with the same identifiers appear. These and other possibilities must be understood for the inverse problem (tomography).

1. Introduction

Because the ocean is opaque to electromagnetic radiation the possible use of sound as an alternative method for remote measurement and monitoring has been a long discussed idea amongst oceanographers. In a recent paper, Munk and Wunsch (1979) outlined a complete system for remotely measuring the ocean by acoustic methods, which they call "acoustic tomography". Their procedure involves "inverting" temporal perturbations in travel time between a multiplicity of sources and receivers to solve for temporal changes in sound speed (and by inference, density) in the regions through which the sound has travelled. But inverse problems as a class are deeply dependent for their solution upon a detailed understanding of the "forward" problem (see for example, Parker, 1977). In the present context, the forward problem is the determination of the arrival of sound at a hydrophone when generated by an impulsive source at a distance of order 10^3 km.

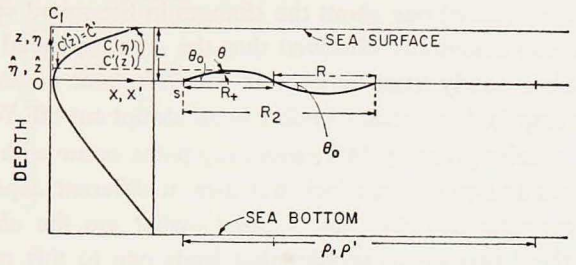
The acoustic properties of the ocean are complex and a solution to the complete forward problem is not available. As a framework for their discussion Munk and Wunsch (1978) (hereafter MW) used a model of the deep sound channel (SOFAR) channel proposed by Munk (1974). Munk (1974) refers to his model as the "canonical" sound channel; here we will call it the Munk profile. It is based upon the as-

1. Department of Earth and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, 02139, U.S.A.

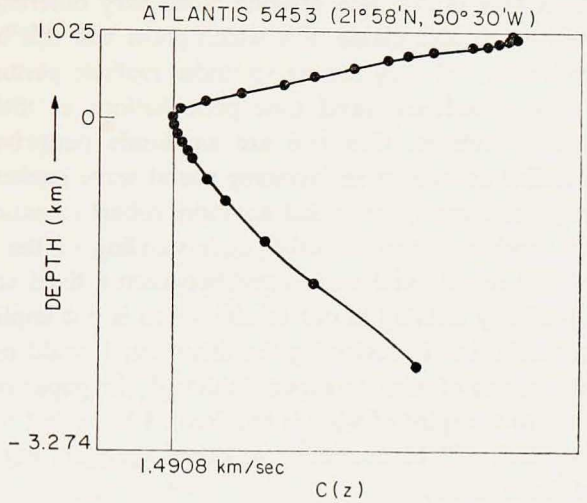
sumption that the buoyancy frequency in the ocean $N(z)$, is a simple exponential. The predicted sound arrival structure from this model, is, as Munk (1974) points out, dissimilar in many ways to what is observed. Examination of Munk's results shows that many of the predicted features of the sound field are very sensitive to the particular numerical values he obtains. The exponential $N(z)$ assumption is of course often a poor one, but it is difficult to understand the effects of deviations from the model because of the selection of specific numerical values.

A number of investigators of sound propagation have resorted to numerical ray tracing techniques in order to determine the characteristics of sound propagation, and others have studied analytically a variety of special profile shapes (e.g. Hirsch and Carter, 1965, Pedersen and Gordon, 1967). But these results are difficult to apply toward ultimate use in the inverse problem. What one needs is an understanding of the sound field as received at a fixed receiver when transmitted from a fixed source, and the dependence of the arriving signal upon the parameters of the ocean as they vary with oceanographic conditions. It is possible to use the numerical models for this purpose, but until one has an analytic understanding of simple models, the numerical ones are cumbersome, expensive to use, and possibly misleading. The analytical models in the literature tend to be special ones, chosen for convenience rather than as models that can easily be varied as one needs in inverse theory. Numerical calculations (some are displayed in MW) often give surprisingly different results for sound speed profiles that appear very similar.

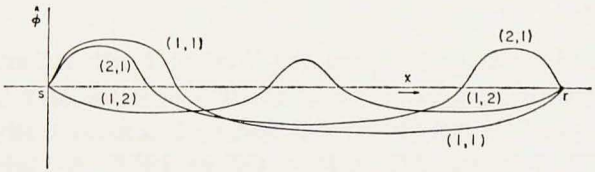
There are a number of questions. Let a sound source and receiver of arbitrary depths be separated by a distance ρ' in an ocean of constant depth D. We suppose that the sound speed C' is a function of the vertical coordinate z only. The two point initial-boundary value problem we need to consider is to find all those angles (measured from the horizontal) at which an acoustic ray can leave the source and subsequently intersect the receiver. We need to know the depth z at which the rays "turn", the time at which each arrives at the receiver, and their intensity (see Fig. 1) assuming a transmitted signal which is a delta function in time. In the special case in which source and receiver are both located on the sound channel axis, some rays travel almost horizontally from source to receiver (axially). Others make a number of loops well away from the sound axis before arriving at the receiver (we ignore rays reflecting from ocean surface or bottom and consider only the purely refracted—or SOFAR—rays). Now it is characteristic of propagation in the observed SOFAR channel that the axial rays, despite having a shorter path to travel, arrive *later* than those travelling much further by looping away from the sound channel axis. The increased sound speed away from the channel axis compensates for the longer path. But what sound channels have this property, i.e. what particular characteristics makes the axial arrival come in last? Is it possible the real sound channel could change from one with a terminal axial arrival to one in which the axial arrival was first or intermediate? Precisely how would one have to change the channel to ar-



a)



b)



c)

Figure 1a. $C'(z)$ represents Munk sound speed profile. The various coordinates are defined in the figure. $\hat{z}$ is the turning depth of a ray that leaves the source, s, at an angle θ. R is defined as the distance between which a given ray crosses the $z = 0$ axis.

1b. $C'(z)$ computed from Atlantis Station 5453 at (21°58′N, 50°30′W) in the North Atlantic.

1c. Three types of paths connecting a source and receiver located on sound channel axis: equal number of upper and lower loops; one more upper loop than lower loop; one more lower loop than upper loop. Equal number of loops case is degenerate to sign of initial ray.

range this? A special case of the inverse problem would occur if one observed such a reversal—what would it tell one about the changes in the sound speed profile and their causes? It is also commonly observed that the arriving signal is such that as rays arrive from more nearly axial paths, that the difference in propagation times decreases, and the amplitude increases, ending in an abrupt cut-off. Why?

In one of the examples given by MW, some ray paths occur with the same number of loops between source and receiver, but turn at different depths (see Fig. 2), but in the other example this does not happen—what are the characteristic differences between these two sound profiles that leads one to this multiple solution for the same number of loops, and the other having only one solution? One of the major assumptions in MW is that rays arriving along very different paths are resolvable in time and hence identifiable. For which paths will this be true? If two such paths are resolvable, will they remain so under realistic perturbations of the sound channels? MW considered travel time perturbations as their fundamental datum for the inverse problem. Can one use amplitude perturbations as well? Finally, is it the detailed arrivals in an incoming sound wave packet that are most useful, or are there gross characteristics that are more robust or easier to use? Fundamentally, we are seeking a comprehensive understanding of the qualitative behavior of a broad-band sound pulse transmitted betweeen a fixed source and fixed receiver. There is probably nothing in our results which is not implicitly known to most workers in the field, but in reviewing the literature, I could not find an adequate qualitative discussion of what to expect (although the paper of Pedersen and White (1970) is close to the spirit of what is required). I hope that the result, which is somewhat pedagogical, will be accessible to oceanographers not normally conversant with sound propagation.

2. Sound channels

An actual sound speed profile from the Western North Atlantic is shown in Figure 1b. The minimum in sound speed at 1025 m is the SOFAR channel discovered by Ewing and Worzel (1948) and discussed in many textbooks (e.g. Clay and Medwin, 1977, Tolstoy and Clay, 1966, Officer, 1958). A minimum occurs because of the competing effects of decreasing temperature which lowers the sound speed, and of increasing pressure which tends to increase it.

By Snell's law, an acoustic ray travelling from a region of low sound speed to one of high sound speed will be refracted toward the horizontal. Eventually, the ray will be turned horizontally and reverse its vertical motion, thus remaining trapped within some finite distance of the sound speed minimum (unless it first intersects the top or bottom boundaries). It is this property that leads to a waveguide effect and makes possible long-range SOFAR propagation.

An enormous range of problems must be dealt with in fully understanding oceanic

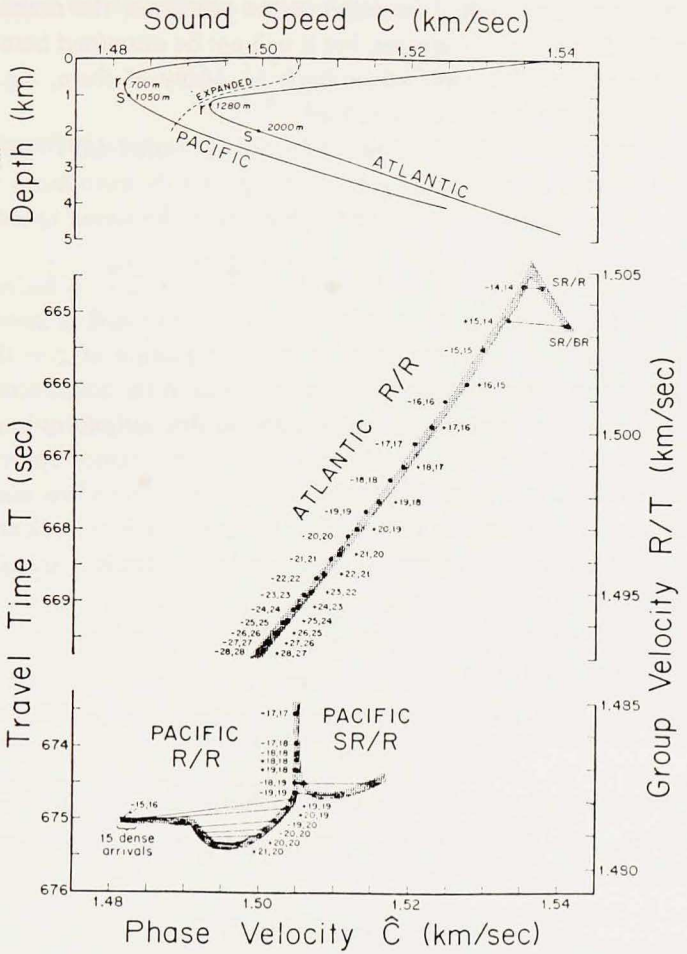


Figure 2. Numerically computed travel time curves for Atlantic and Pacific profiles (taken from Munk and Wunsch, 1978). Notice that in Pacific profile the travel time curve is double valued in $\hat{C}$ (i.e. in ψ) for some rays having same ray identifier (m,n). SR and BR denote surface reflected and bottom reflected rays which are not considered in the present paper. $\pm$ in front of ray identifier states whether ray is initially upward (+) or downward (−) at source; rays of form $(p+1,p)$ always have + sign; $(p,p+1)$ takes − sign. $(\pm p,p)$ cannot be distinguished for the pure axial case. For the numerical cases displayed, sources and receivers were *not* axial.

sound propagation (Urick, 1975; Flátte *et al.*, 1978)—absorption, diffraction from bottom features, small scale inhomogeneities like internal waves and fine-structure, eddies, rings, ambient noise, failure of the ray approximation, etc. Here we will ignore most of these effects and confine ourselves to the classical geometrical optics limit of a smooth, one-dimensional nondissipative profile of sound speed $C'(z)$. The

other classical approach to initial-boundary value problems, the normal mode approach is one that has many advantages, but it will not be examined here. Ultimately, we must put the neglected complications back in. Many of them, e.g. the eddies, are the signal we eventually need to determine.

Our general procedure will follow that adopted by Munk (1974) (hereafter M) which is based upon a mathematical formalism apparently introduced by Pedersen (1969). Some of the details are different here. We define the sound speed profile as

$$C'(z) = C_1 \{1+\epsilon'(z^2+\beta' z^3+\gamma' z^4+\delta' z^5+\zeta' z^6+\ldots\} \tag{1}$$

The absence of a term proportional to z insures that $dC/dz = 0$ at $z = 0$. Thus z is measured positive upward from the sound channel minimum at $z = 0$. Profiles of the form of (1) are analytic at the origin. One can find in the literature special profiles which are not analytic and which lead to peculiar discontinuities in ray behavior (e.g. Pedersen, 1969). Because the applicability of the ray optics approximation itself is doubtful in those cases, we will confine ourselves to functions analytic at $z = 0$, a plausible physical assumption in any case. Whether such complications as fine-structure can be dealt with by assuming nonanalytic profiles is a separate question beyond the scope of the present paper.

Let x be the radial coordinate normal to z (Fig. 1a). Snell's law states that for any ray crossing the x axis at an angle θ_o

$$\frac{C_1}{\cos\theta_o} = \frac{C'(z)}{\cos\theta} = C'(\hat{z}) \equiv \hat{C}' \tag{2}$$

(see Fig. 1a) where θ is the angle made with the horizontal at any other depth z on the ray trajectory, and $C'(\hat{z}) = \hat{C}'$ is the value of C' at the depth $\hat{z}$ where the ray turns horizontal ($\theta = 0$). It follows from (2) that

$$\cot\theta = C'(z)/(\hat{C}'^2 - C'^2)^{1/2} \tag{3}$$

and we can write

$$x' = \int_o^z \cot\theta \, dz = \int_o^z \frac{C'(z)}{(\hat{C}'^2 - C'^2)^{1/2}} dz \tag{4}$$

where x' is the horizontal distance traversed by the ray measured from the point at which it crosses the sound channel axis $z = 0$. Then the distance between the points at which the ray crosses $z = 0$ is

$$R' = 2\int_o^{\hat{z}} \frac{C'(z)}{(\hat{C}'^2 - C'^2)^{1/2}} dz \tag{5}$$

Let the differential arc length along the ray be ds. The differential time of travel along the ray is

$$dt' = \frac{ds}{C'(z)} = \frac{dz}{C'(z)\sin\theta} = \frac{\hat{C}'\,dz}{C'(z)(\hat{C}'^2 - C^2)^{1/2}}$$

and so the time of travel from the axis to depth z is

$$t' = \int_0^z \frac{\hat{C}'\, dz}{C'(z)\,(\hat{C}'^2 - C'^2)^{1/2}} \tag{6}$$

The time T' required for the ray to cross the axis and return is then

$$T' = 2\int_0^{\hat{z}} \frac{\hat{C}'\, dz}{C'(z)\,(\hat{C}'^2 - C'^2)^{1/2}} \tag{7}$$

It is convenient to nondimensionalize the variables. Let L be the distance from the axis of the sound channel to the ocean surface and let $L' = L/2\sqrt{2})$. We write $R = R'/L$, $x = x'/L$, $\eta = z/L'$, $T = T'L/C_1$, $C = C'/C_1$. Notice R and T are scaled by L, but z is scaled by L'. This minor inconvenience suppresses many factors of $2\sqrt{2}$ in the results. The sound speed profile is then

$$\begin{aligned} C(\eta) &= 1 + \epsilon(\eta^2 + \beta\eta^3 + \gamma\eta^4 + \delta\eta^5 + \zeta\eta^6 + \ldots); \\ \epsilon &= \epsilon' L'^2,\ \beta = \beta' L',\ \gamma = \gamma' L'^2, \ldots \end{aligned} \tag{8}$$

We will make the convention that ϵ, $\beta > 0$. The first condition assures that $z = \eta = 0$ is a sound velocity minimum and is thus focussing, rather than a velocity maximum which is defocussing. The condition $\beta > 0$ reflects a general qualitative observation that real sound profiles have a greater curvature above the sound channel than below. This condition can be relaxed easily but doing so will add no insights to our eventual results.

Following M, introduce the new variable

$$\phi^2 = (C-1)/\epsilon;\ \phi = \pm\left(\frac{C-1}{\epsilon}\right)^{1/2} \tag{9a,b}$$

so

$$\psi^2 = (\hat{C}-1)/\epsilon \tag{9c}$$

where the sign of ϕ, ψ corresponds to $\mathrm{sgn}(\eta)$. Then (5) is

$$R = \sqrt{2}\int_0^{\psi} \frac{(1+\epsilon\psi^2)}{\{(1+\epsilon\psi^2)^2 - (1+\epsilon\phi^2)^2\}^{1/2}} \frac{d\eta}{d\phi}\, d\phi \tag{10}$$

We can, using the tables given by Abramowitz (1964), obtain the series square root for (9b),

$$\begin{aligned} \phi &= \eta + A_2\eta^2 + A_3\eta^3 + A_4\eta^4 + \ldots \\ A_2 &= \beta/2 \\ A_3 &= (\gamma/2 - \beta^2/8) \\ A_4 &= (\delta/2 - \beta\gamma/4 + \beta^3/16) \\ A_5 &= \iota/2 - \beta\delta/4 - \gamma^2/8 + 3\beta^2\delta/16 - 5\beta^4/128 \end{aligned} \tag{11a}$$

and from the series reversion formula

$$\eta = \phi + b\phi^2 + c\phi^3 + d\phi^4 + e\phi^5 + \ldots$$

$$b = -\beta/2,\ c = \{2(\beta/2)^2 - (\gamma/2 - \beta^2/8)\} \tag{11b}$$

$$d = (5A_2A_3 - A_4 - 5A_2^3),\ e = \{6A_2A_4 + 3A_3^2 + 14A_2^4 - A_5 - 21A_2^2A_3\}$$

Computation of additional terms is tedious; Morse and Feshbach (1953) discuss it. We now have

$$R = \sqrt{2}\epsilon^{-1/2} \int_0^{\psi} \frac{(1+\epsilon\phi^2)\,(1/2+b\phi+3c\phi^2/2+2d\phi^3+5e\,\phi^4/2+\ldots}{(\psi^2-\phi^2)^{1/2}(2+\epsilon(\psi^2+\phi^2))^{1/2}}\,d\phi$$

$$\cong \epsilon^{-1/2} \int_0^{\psi} \frac{(1/2+b\phi+3c\phi^2/2+\ldots)}{(\psi^2-\phi^2)^{1/2}}\,d\phi$$

$$+ \epsilon^{1/2} \int_0^{\psi} \frac{(3\phi^2/4 - \psi^2/4)\,(1/2 + b\phi+3c\phi^2/2+\ldots)}{(\psi^2-\phi^2)^{1/2}}\,d\phi$$

The integrals can be obtained from formulas 3.248 of Gradshteyn and Ryzhik (1965) and we have

$$R = \epsilon^{-1/2}[\pi/4+b\psi+3\pi c\psi^2/8 + 4d\psi^3/3 + 15\pi e\psi^4/32 + 0(\psi^5)]$$

$$+ \epsilon^{1/2}[\pi\psi^2/32 + b\psi^3/4 + 15\pi c\psi^4/128 + 0(\psi^5)]$$

$$+ 0(\epsilon^{3/2}\psi^4) \tag{12}$$

This is a double expansion in ψ and ϵ; I have no proof of the convergence of the series except empirical calculation of additional terms. In general we should not retain in (12) terms in ψ which are $0(\epsilon)$ or smaller. When we examine real sound profiles below we will find that the solution sometimes will be valid only for uncomfortably small values of ψ. Because of the slow convergence of the series for realistic sound profiles, the results of this paper are probably *not* a good way to carry out the numerical procedures actually involved in performing an inverse calculation. But the analytical insight obtained nonetheless seems an important prologue to the requisite numerical model.

As a reference point, we will use the profile given in M to obtain initial numerical values. In our notation, Munk's profile is

$$C_m(\eta) = 1 + 4\epsilon_m(-\eta/\sqrt{2} + \exp(\eta/\sqrt{2}) - 1) \tag{13}$$

which, when expanded about $\eta = 0$ is

$$C_m(\eta) = 1 + \epsilon_m(\eta^2+\beta_m\eta^3+\gamma_m\eta^4+\delta_m\eta^5+\zeta_m\eta^6+\ldots) \tag{14a}$$

$$\phi_m(\eta) = \eta + A_{2m}\eta^2 + A_{3m}\eta^3 + \ldots \tag{14b}$$

$$\eta = \phi + b_m\phi^2 + c_m\phi^3 + d_m\phi^4 + e_m\phi^5 \ldots \tag{14c}$$

where

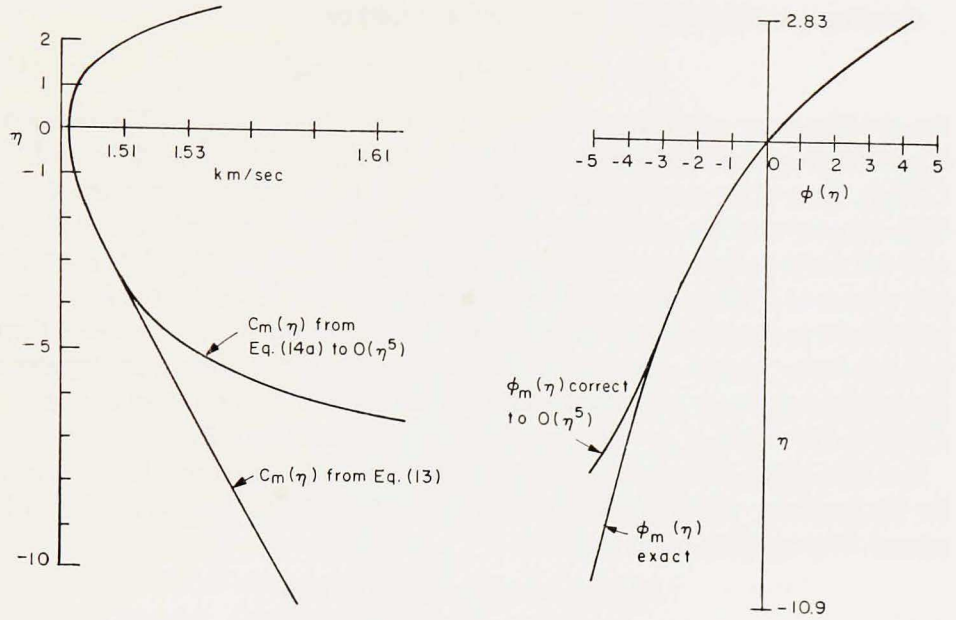


Figure 3a. "Exact" Munk profile $C(\eta)$ and its Taylor series approximation carried through terms of $0(\eta^5)$. Approximation is clearly not valid for rays turning in very deep water.
3b. $\phi(\eta)$ for exact Munk profile and Taylor series approximation through terms of $0(\eta^5)$.

$$\epsilon_m = 1.85 \times 10^{-3}$$

$$\beta_m = .235,\ \gamma_m = .042,\ \delta_m = .0059,\ \zeta_m = 6.9 \times 10^{-4}$$

$$A_{2m} = .12,\ A_{3m} = .014,\ A_{4m} = .0013,\ A_{5m} = 9.4 \times 10^{-5}$$

$$b_m = -.12,\ c_m = .014,\ d_m = -.0013,\ e_m = 1.0 \times 10^{-4} \qquad (14d)$$

If the ocean is 5 km deep, and the sound channel at 1.3 km, the range of η and ϕ (from (14)) is

$$-4.3 \leqslant \phi \leqslant 4.2,\ -8.05 \leqslant \eta \leqslant 2.82$$

MW carried the expansion in (14) to $0(\eta^3)$. In Figure 3a is displayed $C_m(n)$ and $\phi_m(\eta)$ from (13) and its approximation from (14) to $0(\eta^5)$. It can be seen from the figure that only rays that turn shallower than $\eta \cong -3.5$ (dimensionally $z \cong -1.6$ km) will be properly computed by an expansion to this order. Evidently it takes many terms in the Taylor series to represent (13) over the full water depth. This limitation should be kept firmly in mind, and we can anticipate that our results will be valid strictly only for rays turning at small values of $|\eta|$ (small $|\phi|$). Let η_c and ψ_c be the limits of validity. The domain of validity of our expansions in ψ is such that $\psi_B \leqslant \psi \leqslant \psi_s$ where ψ_S, ψ_B are the values of ψ at the surface and bottom respectively, or, $|\psi| \leqslant \psi_c$ whichever range is *smaller*.

Consider now the value of R correct to $0(\epsilon^{-1/2},\psi^2)$ or

$$R = \epsilon^{-1/2}(\pi/4 + b\psi + 3\pi/8\, c\psi^2) \tag{15}$$

For $\psi = 0$ the ray will travel a distance $R_0 = \epsilon^{-1/2}\pi/4$ or $R_0 = 18.3$ in our non-dimensional units using the values of (14d) (dimensionally $R_0 = 23.4$ km if $L = 1.3$ km). Thus if the source is on the axis $\eta = 0$, the axial ray crosses the axis a finite distance from the source; this is the ray which is defined to leave the source at $\theta = 0$ and it is slightly puzzling that the ray supposedly travelling right along the axis crosses at a finite distance. R_0 is, of course, obtained by taking the limit $\psi \to 0$ in (15) and is the value of R at the removable singularity in (10) as $\hat{z} \to 0$. Rays that turn below the axis correspond to values of $\psi = \psi_- \leqslant 0$ and those turning above the axis have $\psi = \psi_+ > 0$. Thus with the convention $\beta > 0$ we have $b < 0$ and $R(\psi_+) < R(\psi_-)$ with $\psi_- = -\psi_+$, to this order of approximation if $c > 0$.

For both upper and lower loops R is a monotonically decreasing function of ψ for the particular values (14d) and small ψ (see Fig. 4). But it need not be so in general. The condition for a reversal is

$$\partial R/\partial\psi = \epsilon^{-1/2}(b + 3\pi c\psi/4) = 0$$

or

$$\psi = -4b/3\pi c$$

and will occur above or below the axis depending upon the sign of c (note $b < 0$). For sufficiently large values of $|c|$ this reversal will occur within the range of validity of our expansions.

A complete double loop is obtained from

$$R_2(\psi) = R(\psi_+) + R(\psi_-) = 2\epsilon^{-1/2}(\pi/4 + 3\pi c\psi^2/8) \tag{16}$$

independent of b (a ray turning at ψ above the axis, turns at $-\psi$ below the axis. In R_2 we make the convention $\psi > 0$). Notice that if $c = 0$, all rays will be brought to a focus (to this order of approximation) at $R_2 = 2R_0$. The condition that $c = 0$ is

$$\gamma = 5/4\ \beta^2 \tag{17}$$

We now compute the travel time in a fashion analogous to the computation of R. We have

$$T = R + \epsilon^{1/2}(-b\psi^3/3 - 3\pi c\psi^4/16 + 0(\psi^5)) \tag{18}$$

as the time required for a ray to intersect the axis $\eta = \phi = 0$ twice. (To obtain this result, one must compute R to $0(\epsilon^{1/2})$, not order $0(\epsilon^{-1/2})$ as implied by M). Thus we have the very important result that the difference in time of arrival between a ray arriving axially and one arriving from off axis will differ by $0(\epsilon^{1/2})$ at most. It is the smallness of ϵ which leads to the practical difficulties at sea.

The time of travel for a double loop is

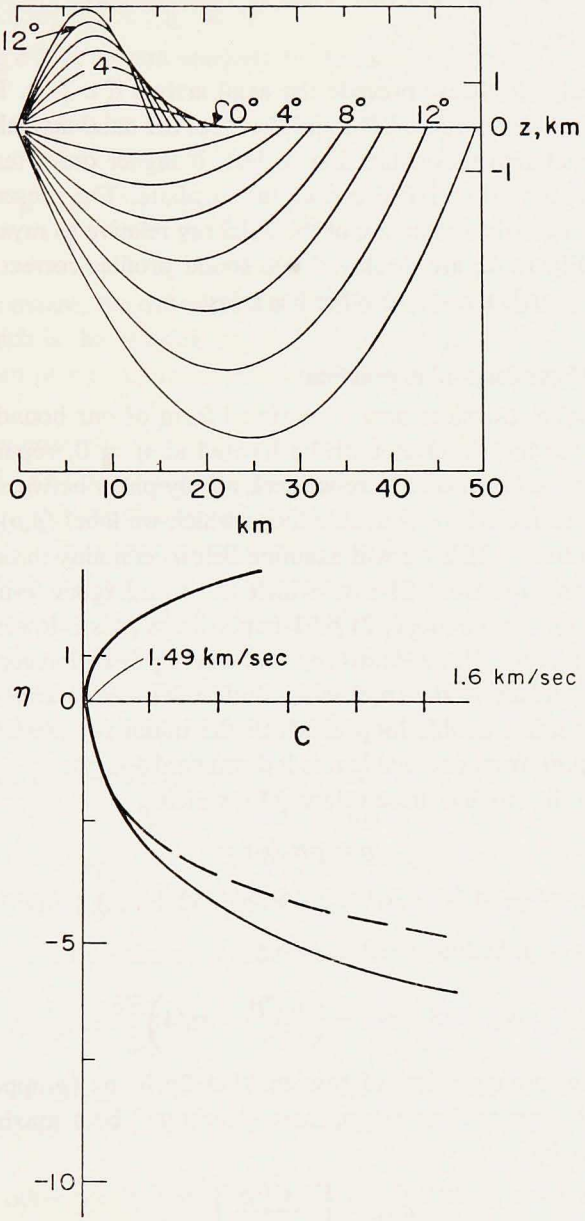


Figure 4a. Adapted from Munk (1974), showing ray trajectories as a function of angle. Notice monotonic decrease in range R as ψ increases.

4b. Two sound speed profiles, correct to $0(\eta^5)$, in which one has axial arrival last, and in other (dashed) axial arrival is *first*.

$$T_2 = T(\psi_+) + T(\psi_-) = R_2 + 2\epsilon^{1/2}(-3\pi c\psi^4/16)$$

$$\psi_- = -\psi_+ \tag{19}$$

An off-axis arrival will always precede the axial arrival if $c > 0$. But if $c = 0$, all rays arrive together (at the focus $2R_0$) and if $c < 0$, the axial arrival is *first*. (To this order in ψ the axial arrival is either first or last. If higher order terms are retained in the expansions, the axial arrival can be intermediate. The present results should be interpreted as the arrival structure of the axial ray relative to rays turning at small values of ψ). In Figure 4b are displayed two sound profiles correct to $0(\eta^5)$ one of which has the axial arrival first; the other has it last.

3. The restricted boundary value problem

a. Range and timing. Consider now a restricted form of our boundary value problem. Let the source and receiver both be located at $\eta = 0$, separated by a non-dimensional horizontal distance. Here we seek all ray paths between the source and receiver which have undergone p double loops which we label (p,p) (notice that the more general solution, which we will examine below, contains three kinds of paths between source and receiver: 1) rays which have undergone p-upper loops and p-lower loops, i.e. p-double loops; 2) $p+1$-upper loops plus p-lower loops $(p+1,p)$; 3) p-upper loops and $p+1$-lower loops $(p,p+1)$ where p is an integer; see Figure 1c).

There is a degeneracy in the (p,p) case. With axial source and receiver there is nothing to distinguish a double loop in which the initial ray leaves the source upward, written $(+p,p)$, from one that leaves it downward $(-p,p)$.

The problem is thus to find those values ψ for which

$$\rho = pR_2(\psi) =$$

$$(2p)\,\epsilon^{-1/2}(\pi/4 + 3\pi c\psi^2/8 + 15\pi e\psi^4/32 + \ldots) + 0(\epsilon^{1/2}\psi^2) \tag{20}$$

for integral values of p. Setting $e \cong 0$, we have

$$\psi_{2p} \cong (8/3\pi c)^{1/2}\left(\frac{\epsilon^{1/2}\rho}{2p} - \pi/4\right)^{1/2} \tag{21}$$

where ψ_{2p} are the turning points of rays making $2p$ loops (p-upper, p-lower) between source and receiver. For $c > 0$, there clearly will be a maximum number of double loops given by

$$p_{\max} = \left[\frac{2\epsilon^{1/2}\rho}{\pi}\right] \tag{22}$$

As p ranges from $p_{\min}$ to $p_{\max}$ the ray paths change from near surface turning to near-axial at $p = p_{\max}$. If we restrict ψ to be less than ψ_s where ψ_s is the value at the surface, on the basis that surface scattered rays will be attenuated over long distances, then there is a minimum p given by

$$p_{\min} \cong \left[\frac{\epsilon^{1/2}\rho}{2} \Big/ (\pi/4 + 3\pi c\psi_s^2/8) \right] \tag{23}$$

Still restricting $c > 0$, we can compute the times of arrival as

$$T_{2p} = \rho + \tau_{2p} = \rho + \epsilon^{1/2}(2p)\left(- \frac{3\pi}{16} c\psi^4 2p \right) \tag{24}$$

$$\rho - \epsilon^{1/2}(3\pi c/16)\,(2p)\,(8/3\pi c)^2 \left(\frac{\epsilon^{1/2}\rho}{2p} - \pi/4 \right)^2$$

$$p_{\min} \leqslant p \leqslant p_{\max}$$

Because ρ is a constant, the differences in time of arrival of successive rays *decrease* as $p \to p_{\max}$ which is the classical rapid SOFAR cut-off burst. As Munk notes, all times are less than $\rho (= \rho'/C_1$ in dimensional form) and thus all rays come in *sooner* than the axial.

If we can now observe changes in the times of arrivals of the different rays in a SOFAR channel like this we are in a position to perform one example of an inversion procedure on this "vertical slice". An equivalent inverse problem occurs at the "initializing" stage. One would normally have an initial estimate of the mean sound speed profile which will be used to compute estimates of the arrival times of the various rays. Observations will almost certainly differ from the calculations. One is then faced with the problem of modifying the initial estimate to bring it into conformity with the data. Schematically, the process is as follows. For an arrival, identified as corresponding to $p = p_o$, we compute from (24) the partial derivatives

$$\frac{\partial T_{2p_o}}{\partial \epsilon} \; ; \; \frac{\partial T_{2p_o}}{\partial \beta} = \frac{\partial T_{2p_o}}{\partial c}\frac{\partial c}{\partial \beta} + \frac{\partial T_{2p_o}}{\partial \psi}\frac{\partial \psi}{\partial \beta} \; ; \tag{25}$$

$$\frac{\partial T_{2p_o}}{\partial \gamma} = \frac{\partial T_{2p_o}}{\partial c}\frac{\partial c}{\partial \gamma} + \frac{\partial T_{2p_o}}{\partial \psi}\frac{\partial \psi}{\partial \gamma}$$

Then the changes in travel times (or differences from the initializing calculations) are

$$\Delta T_{2po} = \frac{\partial T_{2p_o}}{\partial \epsilon}\Delta\epsilon + \frac{\partial T_{2p_o}}{\partial \beta}\Delta\beta + \frac{\partial T_{2p_o}}{\partial \gamma}\Delta\gamma + \ldots \tag{26}$$

which, with measured values of ΔT_2 for several p_o, can be solved for $\Delta\epsilon$, $\Delta\beta$, . . . and hence the parameters of the modified sound channel found as $\epsilon + \Delta\epsilon$, $\beta + \Delta\beta$, . . .

But there are other possibilities too. $p_{\max}$ depends directly upon $\epsilon^{1/2}$. A change in the number of received arrivals, if detectable, would immediately give a datum for determining a new value of $\epsilon^{1/2}$, etc.

This procedure is a different one than described by MW for solving the vertical slice; there the perturbations in travel time were ascribed to perturbations in sound speed in a series of vertical layers. Obviously with a sound speed profile given as

here in terms of a Taylor series in the vertical coordinate we could compute its mean over any set of vertical distances we chose. The two procedures are thus entirely equivalent and one should use whichever is more convenient. As noted above however, the Taylor series approach is unlikely to work well in practice because of its slow convergence—unless one is willing to use only the near axial rays—a serious restriction. (Obtaining the partial derivatives in (25) is equivalent to finding the Fréchet derivative in the Backus-Gilbert formulation of inverse theory).

We have seen above that it is possible to have $c < 0$. If that is the case, then the condition that a solution to (20) exists is now

$$\psi_{2p} = \left(\frac{-8}{3\pi c}\right)^{1/2} \left(\pi/4 - \frac{\epsilon^{1/2}\rho}{2p}\right)^{1/2} \tag{27}$$

where

$$p \geqslant p_{\min} \geqslant \left[\frac{\epsilon^{1/2}2\rho}{\pi}\right]$$

To prevent ψ_{2p} from exceeding ψ_s, we obtain

$$p_{\max} \leqslant \left[\left(\frac{\epsilon^{1/2}\rho}{\pi^2}\right)(\pi/4 + 3\pi c_s\psi^2/8)\right]$$

and the axial arrivals correspond to *small* p rather than large p as in the prior case. The times of arrival for these rays at a distance ρ is

$$T_{2p} = \rho + \tau_{2p} = \rho + \epsilon^{1/2}2p(3\pi\psi_{2p}{}^4/16\,|c|)$$

$$= \rho + \epsilon^{1/2}\left(\frac{3}{16}\,|c|\,\right)\,2p\left(\frac{8}{3\pi|c|}\right)^2\left(\pi/4 - \frac{\epsilon^{1/2}\rho}{2p}\right)^2$$

and as anticipated, the axial arrival is *first*, with an abrupt onset rather than termination.

Clearly, if $c = 0$ all rays arrive together (assuming that the receiver is at a focus) and the only measureable parameter change would be in C_1 which controls the overall timing and which has been suppressed here by scaling it out. Should one be dealing with an ocean in which c passed through zero, there would be great difficulty in ray identification (but it would not be hopeless because the characteristic *spacing* between arrivals could be used to identify the axial rays).

If the sound channel is such that we should keep higher terms in ψ in our expansions then the situation is even more complicated. Let us retain the quartic term in (20). We find for our restricted problem that

$$\psi_{2p} = \left\{-(2/5)c/e \pm 2/5(c/e)\left(1 - \frac{25}{4}\,e/c^2\left(\pi/4 - \frac{\epsilon^{1/2}\rho}{2p}\right)(32/15\pi)\right)^{1/2}\right\}^{1/2} \tag{29}$$

There is now the possibility of obtaining *two* values of ψ_{2p} for fixed ρ and p. Suppose $e/c << 1$. Then

$$\psi_{2p} \cong -2/5(c/e)\left\{1 \mp \left[1-(25/8)\,(e/c^2)\,Q\,(32/15\pi) - \frac{e^2}{8c^4}(25/4)^2\,Q^2\,(32/15\pi)^2 + \ldots\right]\right\}^{1/2}$$

$$Q = -\epsilon^{1/2}\,\rho/2p + \pi/4$$

or

$$\psi_{2p} \cong [-(5/4c)\,Q\,(32/15\pi)]^{1/2} + 0(e/c^3)^{1/2} \tag{30a}$$

and

$$\psi_{2p} \cong [-4/5(c/e) + (5/4c)\,Q\,(32/15\pi)]^{1/2} + 0(e/c^3) \tag{30b}$$

The first root is to lowest order the same as (21). The second root is valid only if $c/e < 0$ and if $\psi_{2p} < \psi_c$. If both of these conditions are fulfilled, one can have two rays, with the same identifier (p,p) covering the vertical column in different ways. The travel time to this higher order is

$$T_{2p} = \rho + \epsilon^{1/2} 2p\,[-3\pi c\psi_{2p}^4 - 15\pi e\psi_{2p}^6/48]$$

For the first root, the value is almost (24). For the second root, we haxe approximately

$$T_{2p} \cong \rho + 2p\epsilon^{1/2}[-3\pi c/16(-3\pi c/16(4c/5e)^4 - (15\pi e/48)\,(4c/5e)^6)] \tag{31}$$

which is very different in value from (24). Evidently, this is one possible explanation for the numerically computed double-valued arrivals in the Pacific displayed in Figure 2.

b. Ray intensity. Thus far, we have discussed using changes in arrival time as data for the inverse problem. But each ray will arrive with a characteristic amplitude, which is also a function of the sound speed profile parameters. There is the possibility of using changes in amplitude as an additional datum; in any event, knowledge of the expected intensity of a particular ray is vital to experiment design.

We need to find the ray bundle intensity, as it spreads cylindrically, relative to the intensity emitted at the source. Officer (1958) and other texts provide the necessary details which we will only summarize here.

Let P be an arbitrary emission intensity/unit solid angle at the source and let θ_o be the angle, as before, of the ray emitted by the source on the channel axis measured relative to the horizontal. With the receiver also on the channel axis, the ray must be received at this same angle (or $\pi - \theta_o$).

Then from Officer (1958) the received intensity is

$$I = \frac{P\cos\theta_o}{\rho\sin\theta_o\, d\rho/d\theta_o} \tag{32}$$

We use

$$\rho \cong 2p\epsilon^{-1/2}(\pi/4 + 3\pi c\psi^2/8)$$

and

$$d\rho/d\theta_o = (\partial\rho/\partial\psi)\partial\psi/\partial\theta_o;\ \partial\psi/\partial\theta_o = \frac{\sin\theta_o}{2\psi\cos^2\theta_o}\,\epsilon^{-1}$$

from Snell's law and (9). We have approximately then

$$I = \frac{\epsilon^{1/2}P}{4p\,\rho\left(\dfrac{\epsilon^{1/2}\rho}{2p} - \pi/4\right)^{1/2}} \tag{33}$$

In Figure 5a we plot the relative intensity from (33) (but carried through $0(e)$) as a function of p at a (dimensional) distance of 1000 km. Notice the rising amplitude for increasing p until the axial cut-off-the classical SOFAR reception effect (Ewing and Worzel, 1948). Only rays turning at $|\psi| < 4$ have been included because the expansions are very doubtful at larger values. Rays turning at $|\psi| > 4$ would come in *earlier* than those displayed. To the extent that one can indeed measure *changes* in received intensities they provide another datum for channel inversion through $\partial I/\partial c$, etc.

If we choose to regard p as a continuous parameter (possibly by introducing some suitable interpolation procedure) then we can speak of the envelope of the arriving wave packet. Because this smooth curve is also a function of the sound speed profile parameters, we could contemplate using its gross shape as our description of the arriving signals. In practice, one could least-square fit a curve to the envelope and study its variation in time. This procedure could be a more robust one than attempting to resolve the individual arrivals in a situation of high noise level or inadequate bandwidth.

4. Complete axial boundary value problem

We now will complete the axial case by considering those rays which have one more upper loop than lower, and vice versa.

a. Additional upper loop. Suppose that there are p lower loops and $p+1$ upper loops denoted $(p+1, p)$. Then

$$\rho = 2p\,\epsilon^{-1/2}\left(\pi/4 + 3\pi c\psi^2/8 + \frac{15\pi}{32}\,e\psi^4 + \ldots\right) + \epsilon^{-1/2}\,(\pi/4 + b\psi + 3\pi c\psi^2/8 + 4d\psi^3/3 + 15\pi e\psi^4/32 + \ldots) + 0(\epsilon^{1/2}\psi^2) \tag{34}$$

Retaining terms to order ψ^2 we have

$$\rho = (2p+1)\,\epsilon^{-1/2}\,\pi/4 + b\,\epsilon^{-1/2}\psi + (2p+1)\,\epsilon^{-1/2}\,3\pi c\psi^2/8 \tag{35}$$

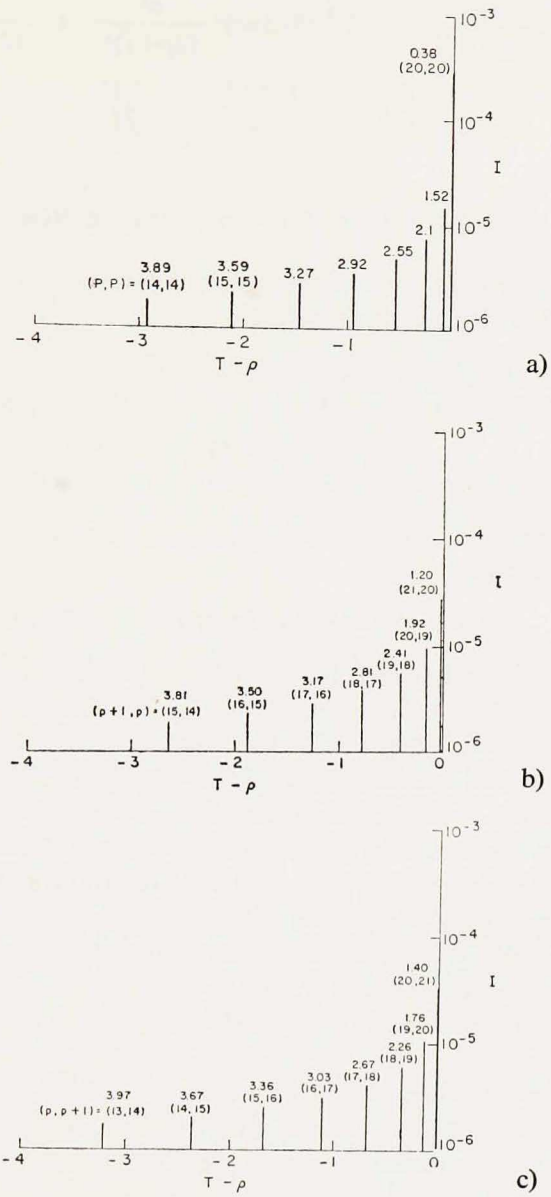


Figure 5a. Arrival sequence of double loop rays in Munk profile from a source 10^3 km from the receiver. Both are axial. No rays turning at $|\psi| \geqslant 4$ are displayed as solutions are increasingly doubtful for large $|\psi|$. Upper number is ψ, lower pair is ray identifier. Units are all non-dimensional. Time axis can be dimensionalized by multiplying by $L'/C_1 \sim 1$. I is relative intensity.

5b. Same as 5a, except for rays with one more upper loop than lower.

5c. Same as 5a except for rays with one more lower loop than upper.

$$\psi = \frac{-4b}{3\pi c(2p+1)} \pm 1/2 \left\{ (8/3\pi c)^2 \frac{b^2}{(2p+1)^2} + \frac{32}{(2p+1)3\pi c} \cdot \left(\rho\epsilon^{1/2} - \frac{(2p+1)}{4} \pi \right) \right\}^{1/2}$$

The travel time is

$$T = \rho + \epsilon^{1/2}(2p+1)(-3\pi c\psi^4/16) - \epsilon^{1/2}(b\psi^3/3) \tag{36}$$

b. Additional lower loop. Now suppose there are p upper loops and $p+1$ lower loops $(p, p+1)$. We obtain to this same order

$$\psi = \frac{4b}{3\pi c(2p+1)} \pm 1/2 \left\{ (8/3\pi c)^2 \frac{b^2}{(2p+1)^2} + \frac{32}{(2p+1)(3\pi c)} \left(\epsilon^{1/2}\rho - \frac{(2p+1)}{4} \right) \pi \right\}^{1/2} \tag{37}$$

where ψ must be positive. The expressions (35, 37) differ only in the sign of b.

The travel time for this case is

$$T = \rho + \epsilon^{1/2}(2p+1)(-3\pi c\psi^4/16) + \epsilon^{1/2}(b\psi^3/3) \tag{38}$$

Because b is always negative, for an additional upper loop and $c > 0$ we can have two solutions ψ for the same value of p. For the additional lower loop one can have only one solution, unless $c < 0$. The number of possible solutions and hence the number of arrivals is clearly highly dependent upon c. Solutions still exist for these cases even if $c = 0$. Because ψ must be positive, there will be a $p_{\min}$ for an extra upper loop and a $p_{\max}$ for an additional lower loop. Because all double loops focus at fixed points for the case $c = 0$, and if the receiver is not at one of these foci, we will see then only the odd number loop arrivals, determined by the appropriate value of p. Obviously, a complete description of arrival structure is a complex function of the parameters of the sound channel.

The intensity of arrivals may be obtained analogously to (33). These are graphed in Figures 5b, c.

For the numerical values of (14d) we display in Figure 6 the complete arrival structure for all loops at a receiver at a horizontal distance of 10^3 km. The (p,p) arrival dominates the sequence because as noted above, with the source and receiver both on the channel axis, there are two rays which arrive at the same time and reinforce each other. One ray leaves the source directed upward; the second ray leaves it directed downward. The detailed form of this solution is dependent in an intricate way upon the values of ϵ, β, γ etc. But numerically, it is not difficult to vary the coefficients and build up a variational table from the resulting partial derivatives as is done in studying the free oscillations of the earth (e.g. Wiggins, 1968).

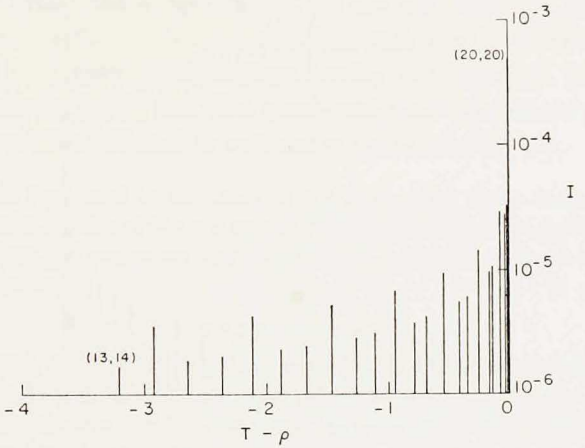


Figure 6. Complete arrival sequence for axial source and receiver in Munk profile. This figure is the sum of 5a-c except (p,p) arrivals are multiplied by 2 and dominate, because up-down degeneracy gives 2 rays arriving simultaneously.

5. Oceanic profiles

If we do a polynomial fit (a hazardous procedure) to the observed profile displayed in Figure 1b, we obtain the values: $L = 1.025$, $\epsilon = 4.7 \times 10^{-3}$, $\beta = 1.15$, $\gamma = 0.239$, $\delta = -.409$, $\zeta = -.195$. The very large value of β which is the coefficient of the cubic term is the result of the marked asymmetry of the sound speed profile. Substituting these values into (11a,b) we obtain

$$b = -.58,\ c = .72,\ d = -.93,\ e = 1.24,$$

and the expansions will be poorly convergent except for very small ψ.

The final arrival structure is quite different from that of Munk's profile. It is displayed in Figure 7. The increased values of $\epsilon^{1/2}$ and of β increase the value of $p_{\max}$ at this range from 21 to 41; the final cut-off burst thus consists of many more arrivals, extremely tightly spaced. The large value of β for this profile gives substantial differences between the arrival of $(p+1, p)$ and $(p,p+1)$ even for small value of ψ. The marked difference between this arrival structure and that of the Munk profile is encouraging for the inverse problem in that it suggests a *qualitative* sensitivity of observables—e.g. the number and spacing of arrivals—to comparatively small changes in sound channel parameters. Whether deep sound channels with $c < 0$ and an early axial arrival actually exist in nature is not clear.

6. Some final remarks

One can proceed to solve the more interesting case of off-axial sources and receivers by the same methods. The horizontal distance traveled by a ray traveling

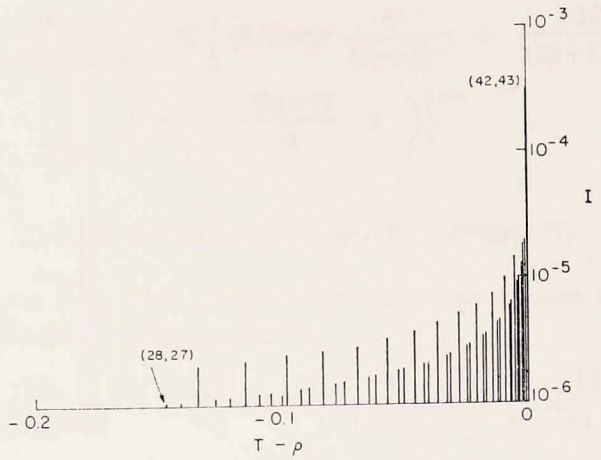


Figure 7. Complete axial arrival sequence for profile of Figure 1b in North Atlantic at range 10^3 km. There are many more arrivals here than in Munk profile and they are very tightly bunched near termination (some earlier arrivals are not displayed).

an arbitrary vertical distance from the axis $\eta = 0$ may be obtained from equation (10) (replacing the upper limit of integration by an arbitrary value of ϕ) and the formulas of Gradshteyn and Ryzhik (1965). The results are not displayed here; a referee has suggested (rightfully) that the analytical complexity is so great that no insight over-and-above that obtained from an ordinary numerical ray-trace is likely to be obtained. We will thus leave the problem restricted as it is to the purely axial source and receiver. It is possible that some expansion procedure more powerful than the Taylor series used here would permit a truely analytical approach to the general case. But that is something for another time.

From the point of view of the inverse analysis required of the tomographic system, the present exercise is of limited value even for the axial case: the series expansions required to solve the direct problem with sufficient accuracy are cumbersome to compute and slowly convergent.

Observational evidence suggests that major perturbations of the acoustic structure of the ocean can be described by the lowest dynamical baroclinic mode—or, in other words that the ocean tends to change by shifting the entire thermocline up and down. For a source and receiver initially on the sound channel axis, the major result of the perturbation will be to shift them off. To deal with this case we must discuss the acoustic propagation between an off-axis source and receiver. Distance from the axis must then be treated as a variable in the inverse calculation.

Perhaps the chief virtue of what has been accomplished here is that we obtain some understanding of the complexity of the two-point boundary value problem. The expressions for turning depths, number of arriving rays, and their intensity permit one to anticipate potential difficulties in an operational tomographic system

and to understand their cause. Specifically, we can anticipate the possibility of rays arriving too close together to be resolved (although they may have very different intensities); rays with the same identifiers having very different paths, changes in the number of arriving paths, etc. The quasi-analytical expressions also permit one to contemplate inversion schemes as outlined in equations (25, 26).

Acknowledgments. Supported by the Office of Naval Research under Contract N00014-75-C-0291. The very careful reading of the manuscript by Peter Worcester is appreciated. I am grateful to Charmaine King for computing help. A continuing acoustic dialogue with Walter Munk has been a great stimulus.

REFERENCES

Abramowitz, M. 1964. Elementary analytical methods, *in* Handbook of Mathematical Functions, Abramowitz, M., and I. A. Stegun, eds., Dover, New York, 9–63.

Clay, C. S., and H. Medwin. 1977. Acoustical Oceanography: Principles and Applications, Wiley-Interscience, New York, 544 pp.

Ewing, M. and J. L. Worzel. 1948. Long-range sound transmission, Geol. Soc. Amer. Mem., *27,* 1–35.

Flatté, S., ed., R. Dashen, W. H. Munk, K. M. Watson and F. Zachariasen. 1978. Sound Transmission Through a Fluctuating Ocean, Cambridge University Press, 299 pp.

Gradshteyn, I. S. and I. M. Ryzhik. 1965. Table of Integrals, Series, and Products, 4th edition, Academic Press, New York, 1086 pp.

Hirsch, P. and A. H. Carter. 1965. Mathematical models for the prediction of SOFAR propagation effects, J. Acoust. Soc. Amer., *37,* 90–94.

Morse, P. M., and H. Feshbach. 1953. Methods of Theoretical Physics, Vol. I, McGraw-Hill, New York, 997 pp.

Munk, W. H. 1974. Sound channel in an exponentially stratified ocean, with application to SOFAR, J. Acoust. Soc. Amer., *55,* 220–226.

Munk, W., and C. Wunsch. 1979. Ocean acoustic tomography: a scheme for large-scale monitoring, Deep-Sea Res., *26A,* 123–161.

Officer, C. B. 1958. Introduction to the Theory of Sound Transmission with Application to the Ocean, McGraw-Hill, New York, 284 pp.

Parker, R. 1977. Understanding inverse theory, Ann. Rev. Earth Planet. Sci., *5,* 35–64.

Pedersen, M. A. 1969. Theory of the axial ray, J. Acoust. Soc. Amer., *45,* 157–176.

Pedersen, M. A. and D. F. Gordon. 1967. Comparison of curvilinear and linear profile approximation in the calculation of underwater sound intensities by ray theory, J. Acoust. Soc. Amer., *41,* 419–438.

Pedersen, M. A. and DeW. White. 1970. Ray theory for sources and receivers on an axis of minimum velocity, J. Acoust. Soc. Amer., *48,* 1219–1248.

Tolstoy, I. and C. S. Clay. 1966. Ocean Acoustics: Theory and Experiment in Underwater Sound, McGraw-Hill, New York, 293 pp.

Urick, R. J. 1975. Principles of Underwater Sound, McGraw-Hill, New York, 384 pp.

Wiggins, R. A. 1968. Terrestrial variational tables for the periods and attenuation of the free oscillations, Phys. Earth Planet. Interiors, *1,* 201–266.

Received: *19 June, 1979;* revised: *18 March, 1980.*