CHAPTER 1

THE TOMOGRAPHY PROBLEM

The problem of ocean acoustic tomography is to infer from precise measurements of travel time, or of other properties of acoustic propagation, the state of the ocean traversed by the sound field. The tomographic method[1] was introduced by Munk and Wunsch (1979) in direct response to the demonstration in the 1970s that about 99% of the kinetic energy of the ocean circulation is associated with features that are only about 100 km in diameter, called the mesoscale.[2] Measuring and understanding the behaviors of both the mesoscale and the larger-scale features associated with the general circulation present a formidable sampling task. Not only are the flow elements very compact spatially, but also they have long time scales (order 100 days). To produce statistically significant measurements of the fluid behavior, even in an area as compact as 1 Mm × 1 Mm (1 megameter = 1000 km), about 1% of an ocean basin, requires several full-time vessels or several hundred fixed moorings. One is accordingly led to the technology of sound propagation to measure the properties of the fluid *between* moorings.

Ocean acoustic tomography takes advantage of the facts that (i) travel time and other measurable acoustic parameters are functions of temperature, water velocity, and other parameters of oceanographic interest and can be interpreted to provide information about the intervening ocean using inverse methods, and (ii) the ocean is nearly transparent to low-frequency sound, so that signals can be transmitted over distances of many thousands of kilometers. There is some analogy with classical seismology, in which the properties of the Earth's interior are inferred from travel times of earthquake waves. (However, in ocean tomography the emphasis is not on the *mean* field of sound-speed, but on its space-time *variability*.) One advantage of the oceanographic problem over the seismological and medical ones is that the interior of the ocean is generally accessible. In the long run, that access may have been a liability, for it delayed

[1] See Appendix A for an account of early ocean acoustic tomography.
[2] Figs. 1.5 and 1.6 show realizations of mesoscale-dominated oceans.

the development of indirect methods, which, in our prejudiced view, are made inevitable by the magnitude of the ocean sampling task.

There are numerous attractive features to ocean acoustic tomography. In common with other methods of remote sensing, it permits the monitoring of regions that are difficult to observe directly. Meanders of the Gulf Stream have been monitored with sources and receivers moored in the relatively sluggish waters to both sides of the stream. Another advantage has to do with the speed of sound (3000 knots) exceeding that of research vessels and permitting the construction of synoptic fields. The geometry of measuring the oceans *between* moorings can be exploited; with M moorings, conventional techniques yield M "spot" measurements. But M moorings consisting of S sources and R receivers yield $S \times R$ pieces of information, rather than $S + R = M$. This quadratic information growth is an attractive feature (though the quadratic loss with instrument failures is not).

A key attribute of tomographic measurements is that they are spatially *integrating*. The potential for forming horizontal and vertical averages over large ranges, up to global scales, is an attractive (but unfamiliar) tool; for example, the heat content in a vertical section across an ocean basin could be rapidly and repeatedly measured using tomographic techniques. The integrals suppress unwanted small scales that contaminate the conventional spot measurements, leading to aliasing. Transmissions over a few hundred kilometers subdue the internal wave "noise," and transmissions over a few thousand kilometers subdue the mesoscale noise. Such integral data test the skill of dynamic models and provide powerful model constraints.

Construction of a practical system, deployable at sea, accompanied by the mathematical apparatus to handle the resulting data, requires a working knowledge of elements of oceanography, acoustics, and mathematics, which are normally discussed in isolation. None of these elements is by itself particularly difficult, but in combination they present a formidable challenge. To sustain the reader's interest through much of the necessarily technical material that follows, this chapter is devoted to a heuristic summary account of ocean acoustic tomography. Most of the details are swept aside, to be taken up again in later chapters.

1.1. Ocean Acoustics

The sound-speed profile. Over much of the world ocean, the speed of sound, C, is characterized by a distinct minimum, at depths between 800 and 1200 m, called the sound channel axis. The minimum owes its existence to the dependence of the sound-speed on temperature and pressure. Sound-speed increases upward from the axis with increasing temperature, and increases downward

from the axis (where the temperature gradients are small) with increasing pressure.

At high latitudes (north and south), the surface waters are increasingly colder, and the sound-channel axis shoals. During winter, convective overturning and the resulting mixing lead to near-adiabatic[3] conditions. Sound-speed then increases downward (negative z) at approximately the adiabatic rate: $-(1/C)\, dC/dz = \gamma_a = 0.0113\,\mathrm{km}^{-1}$.

We have found it useful to define two idealized model profiles for which most of the propagation characteristics can be derived analytically: the "temperate" (or canonical) sound-speed profile, and the polar winter (adiabatic) profile (fig. 1.1). Formulae are given in sections 2.17 and 2.18. In general, real profiles differ significantly from the idealized models, and change from place to place (see propagation atlas, Appendix B).

For many important dynamic purposes, the ocean is characterized by the vertical gradient of the potential density field, written as

$$N^2 = \frac{-g}{\rho}\frac{\partial \rho_p}{\partial z}\,, \tag{1.1.1}$$

where z is the vertical coordinate (upward from the surface), g is the force due to gravity, and ρ_p is the potential density (*i.e.*, the density corrected for adiabatic vertical displacements). N has the units of a frequency and is called the "buoyancy" or (less desirably) the Brunt-Väisälä frequency. The implications of N for ocean dynamics have been discussed by Turner (1973) and Gill (1982). In chapter 2, this quantity is related explicitly to the sound-speed gradient: under quite general conditions,

$$\frac{dC/dz}{C_A} = \gamma_a \frac{N^2 - N_A^2}{N_A^2}\,, \tag{1.1.2}$$

where C_A and N_A are the axial sound-speed and buoyancy frequency. For the two idealized models previously introduced,

$$\text{polar:} \quad N = 0, \qquad \text{temperate:} \quad N = N_0\, e^{z/h},$$

respectively. In polar winter, $N \to 0$ and $dC/dz \to -\gamma_a C_A$, the adiabatic profile. For the temperate profile at great depth, $N \to 0$ as in the polar profile. At the axis, $N(z) = N_A$ and $dC/dz = 0$. This depth of minimum sound-speed C_A (about 1 km in temperate latitudes) is where the acoustic waveguide is centered. Above the axis, $N(z)$ and hence dC/dz increase sharply up to a surface-mixed layer.

[3]Here *adiabatic* refers to vertical gradients in temperature, density, and sound-speed associated with the hydrostatic pressure gradient. The term is also used to refer to horizontal gradients so gradual that there is no significant acoustic scattering.

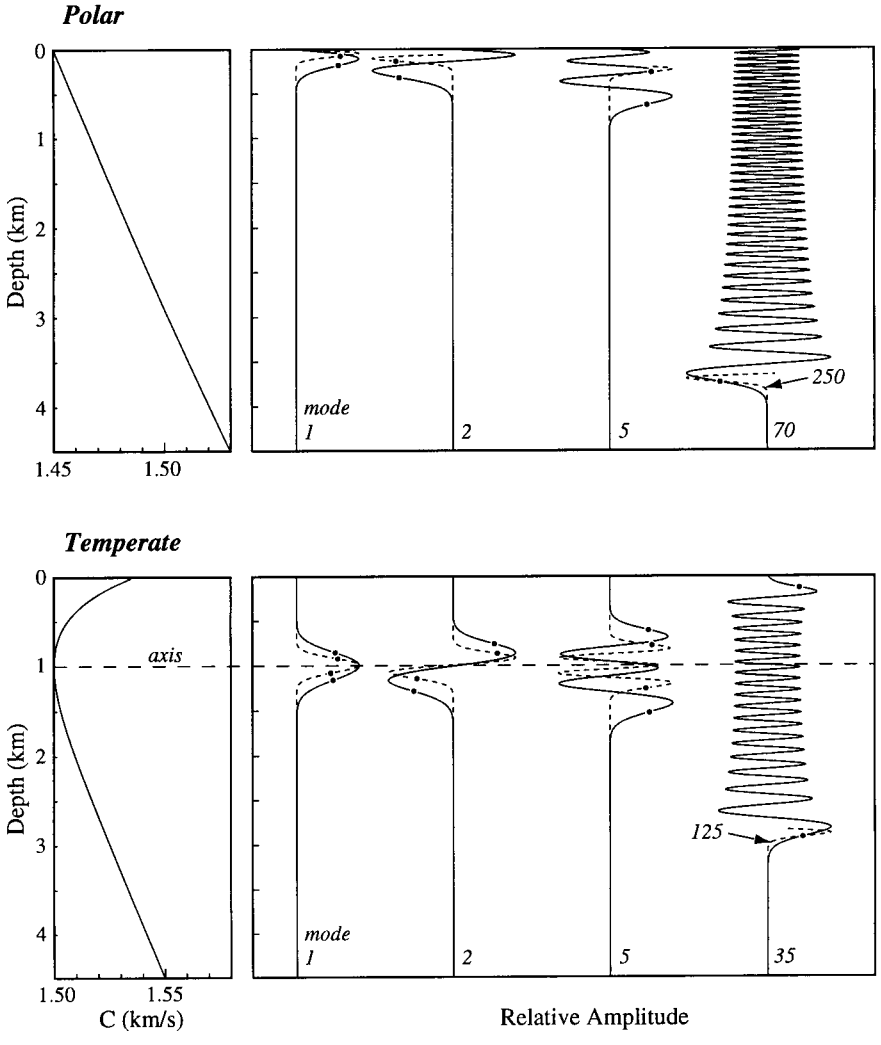

Fig. 1.1. Panels on left show the sound-speed profiles for the polar and temperate models. Polar rays are drawn for an (unrealistic) surface source and surface receiver. All rays are refracted surface-reflected (RSR); ray identifiers -7, -9, ... designate the number of (upper plus lower) turning points, including surface reflection, with minus signs indicating downward launch angles. Temperate rays $+8$, $+9$, -9, $+10$ are for an axial source and receiver. The vertical exaggeration in ray diagrams (right) is 25 to 1; a "small-inclination" approximation is generally valid. Selected modes $m = 1, 2, \ldots$ are shown at two frequencies, 70 Hz (solid) and 250 Hz (dashed), with WKBJ turning points indicated. Polar profile layers $j = 1$ to 6 serve as an example for the application of inverse theory.

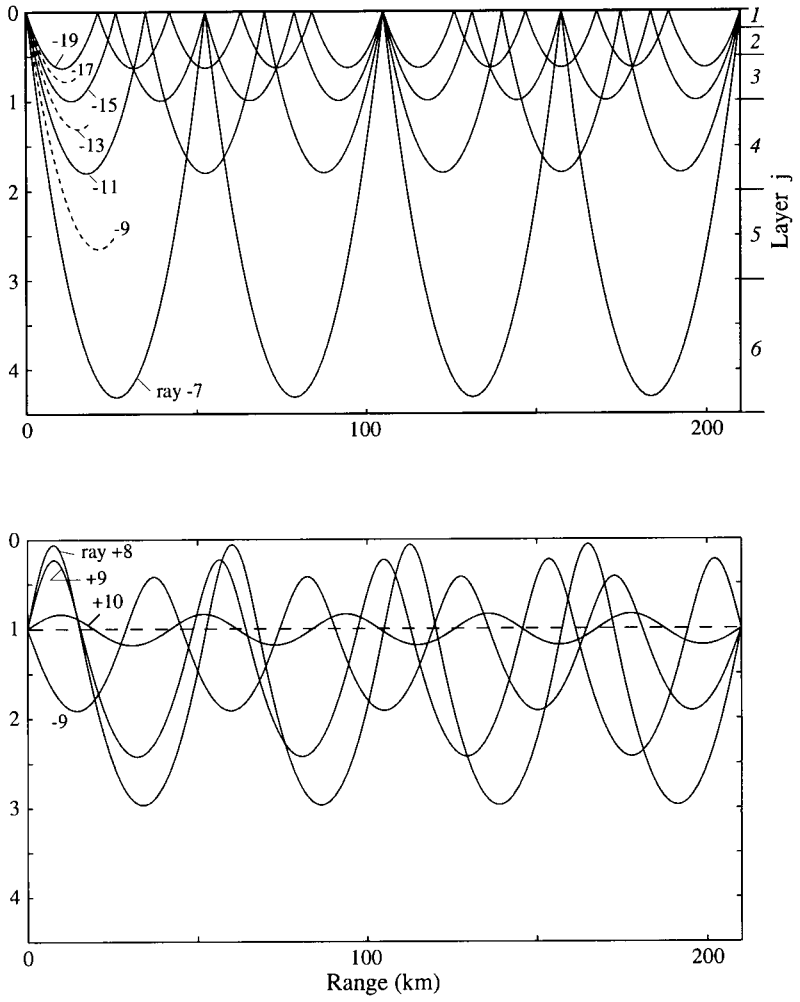

Fig. 1.1 (*cont.*)

Acoustic rays. A useful way of understanding the propagation of sound waves in the ocean is afforded by the geometric optics approach. Acoustic energy tends to propagate along arcs called "rays." In the presence of gradients in the sound-speed, Snell's law shows that the rays bend (refract) away from, or are "repelled" by, regions of higher speeds; for the polar profile, all rays are then refracted upward and subsequently reflected at the surface (fig. 1.1). They are called RSR (for refracted surface-reflected). At a specified range, rays can be designated $\pm p$ by the upward/downward launch direction and the number

of turning points (including surface reflections), as shown. For late arrivals, $p \to \infty$. Surprisingly, the steep rays are the first to arrive; the higher velocity at depth more than makes up for the longer path length.

In the temperate profile, rays are refracted downward from the region above the axis, and upward from the region below the axis. Such rays are called RR (refracted refracted), to denote that they turn owing to refraction at both upper and lower excursions. Typically the *ray wavelength* (upper plus lower loop) is of order 50 km. Steep rays may intersect the surface (RSR), the bottom (RBR), or both (SRBR). For a fixed range there is a discrete number of RR rays, of which four are shown in fig. 1.1. The steepest rays ($p = \pm 8$) are the first to arrive, and the flatter (near-axial) rays ($p = \pm 10$) are the last to arrive.

Ray travel times. The travel time of a nearly horizontal eigenray in a moving, range-independent ocean can be written

$$\tau_n^{\pm} = \int_{\Gamma_n^{\pm}} \frac{ds}{C(z) \pm u(z)} \tag{1.1.3}$$

for a transmission in the positive/negative x direction, respectively. A transceiver (source and receiver) is located at both the starting point and end point; u is the flow velocity component along the ray in the positive x direction. The paths of integration $\Gamma_n^{\pm}$ are along the trajectories of the nth ray and are generally functions of $C(z)$ and $u(z)$. It will be shown that the path geometry is reciprocal to order $u/C \ll 1$, hence $\Gamma^+ \approx \Gamma^- \equiv \Gamma$. The sum and difference of reciprocal travel times are defined by

$$s_n = \tfrac{1}{2}(\tau_n^+ + \tau_n^-) = \int_{\Gamma} ds \frac{C}{C^2 - u^2}, \tag{1.1.4a}$$

$$d_n = \tfrac{1}{2}(\tau_n^+ - \tau_n^-) = -\int_{\Gamma} ds \frac{u}{C^2 - u^2}. \tag{1.1.4b}$$

C and u are of order 10^3 and 10^{-1} m/s, respectively, so u^2 can be neglected in the denominator. The difference travel time is a small fraction of the one-way travel time τ. Hence C is well determined by one-way travel times in either direction.

Here we are more interested in the *perturbation* $\Delta\tau$ from a previous measurement, or from that inferred for the climatic ocean mean. Linearizing (1.1.4) yields

$$s_n = \int_{\Gamma} ds \frac{1}{C}, \quad \Delta s_n = -\int_{\Gamma} ds \frac{\Delta C}{C^2}, \quad d_n = -\int_{\Gamma} ds \frac{u}{C^2}. \tag{1.1.5a, b, c}$$

ΔC is of order 10 m/s and is still large compared with u. Hence, the perturbation ΔC is well determined by one-way travel time in either direction; measurements of u, however, require travel-time differences.[4]

Useful measurements of current profiles, $u(z)$, have been made by difference tomography (chapter 3). The method is particularly useful for separating barotropic (depth-independent) from baroclinic (depth-variable) tidal currents. Ocean current meter records are found to have comparable contributions from the barotropic and baroclinic components. Steep rays produce good estimates of vertically averaged horizontal velocities and are thus responsive predominantly to the barotropic component.

The determination of the sound-speed $C(x, y, z)$ is not of particular interest to oceanographers, except for its relation to temperature T and fluid density ρ. Both C and ρ are functions of T and salinity Sa (for fixed pressure p). We show that the effect of salinity is relatively small. An equation for C in terms of T and Sa is a complicated function that can be linearized to

$$\Delta C/C = \alpha\, \Delta T + \beta\, \Delta Sa\,,$$

with $\alpha \sim 3 \times 10^{-3}/°\mathrm{C}$, $\beta = 1 \times 10^{-3}/\,‰$. For a locally linear temperature–salinity relation

$$Sa = Sa(T_0) + \mu\, \Delta T, \quad \Delta T = T - T_0,$$

we have

$$\Delta C/C = \alpha\, \Delta T(1 + \mu\beta/\alpha)\,. \tag{1.1.6}$$

A typical value is $\mu\beta/\alpha = 0.03$, and a determination of ΔC is, to first order, a determination of the temperature field.

Acoustic modes. An alternative sound-field representation is in terms of acoustic modes. Application of modal theory is a straightforward way of solving the propagation problem in the range-independent case. For some analytical profiles, including the polar and temperate models, exact solutions exist.

Modes are designated by $m = 1, 2, \ldots$, having $0, 1, \ldots, m-1$ zero crossings of the vertical wave function, as shown in fig. 1.1. The scale of the mode function depends on frequency, f. Higher frequencies are more concentrated near the axis. (For the polar ocean, the axis is at the surface.) The inflection points farthest from the axis (the WKBJ "turning points") are measures of the

[4] We can regard u as a perturbation from zero flow, and write Δd_n for d_n.

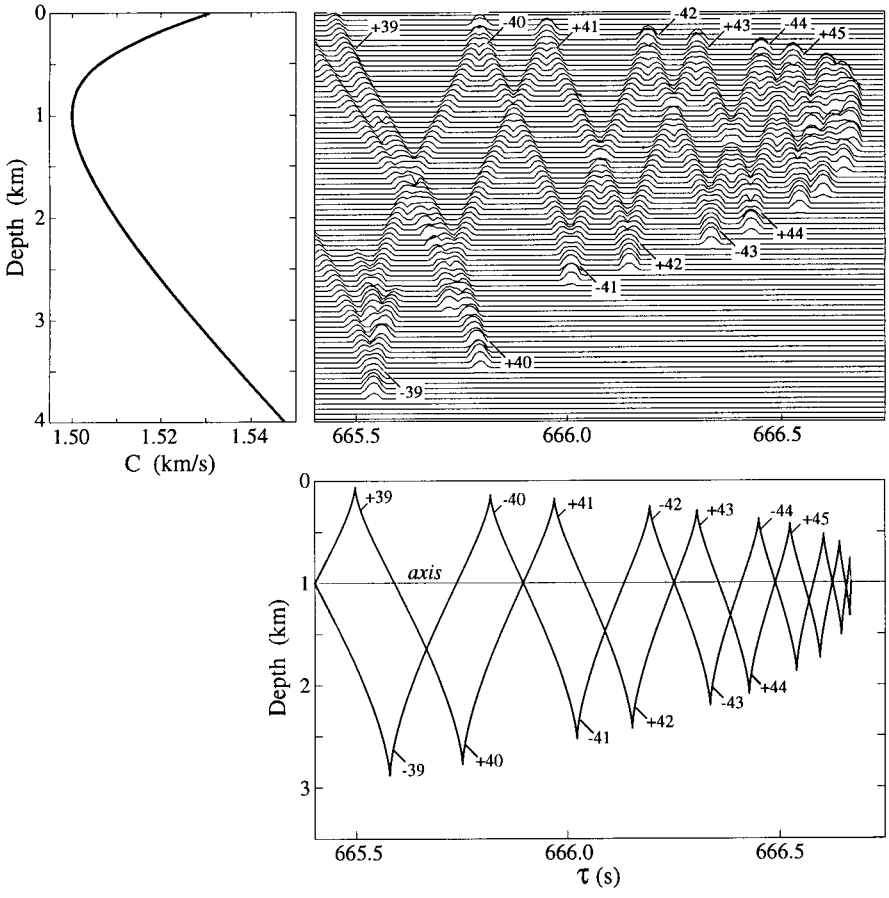

Fig. 1.2. Computed arrival pattern in a temperate ocean at 1 Mm range according to ray theory (bottom) and mode theory (top). The ray axial arrival pattern has a peak whenever one of the ray fronts crosses the line $z = z_{\text{axis}}$; the off-axis pattern is similarly derived using z_{receiver}. The modal pattern is derived by summing modes 1 to 250 at 250 Hz with bandwidth 100 Hz. It is seen that all but the latest interference ridges are ray-like. (The splitting of the ±39 ridges is a surface effect.)

penetration of the modes into the ocean away from the axis. The figure shows only modes with turning points well above the bottom boundary.

An acoustic point source generates all these modes, with amplitudes proportional to the vertical wave function at the depth of the source. Each mode propagates with group velocity c_g (group slowness s_g), which is a known function of mode number m and frequency f. An important parameter is the "action,"

$$A = (m - \tfrac{1}{2})/f, \quad A = (m - \tfrac{1}{4})/f\,, \tag{1.1.7}$$

for non-surface-interacting and surface-interacting modes, respectively. Modes with the same value of A have the same turning point. For example, in a temperate-ocean mode, $m = 125$ at frequency $f = 250$ Hz and $m = 35$ at $f = 70$ Hz produce nearly the same $A \approx \frac{1}{2}$ s.

The composite arrival pattern at a receiver is obtained by summing over individual mode arrivals. Fig. 1.2 shows the result in a temperate ocean at a 1-Mm range for a 250-Hz source with 100-Hz bandwidth. Such a broadband source generates many modes with the same A and accordingly the same turning depth; these modes interfere constructively to produce the prominent accordion-like feature in fig. 1.2. For comparison, the "ray fronts" obtained from the ray theoretical approximation are shown. The patterns are in close agreement;[5] upper and lower ray turning depths occur at the two modal WKBJ turning depths. For both rays and modes, travel time at range r is conveniently written

$$\tau = s_g r\,, \tag{1.1.8}$$

but the expressions for phase and group slowness s_g are different (see section 2.11).

Under normal operating conditions, the ray-like constructive interference pattern is the most prominent and robust feature of the recorded signal. Nearly all tomography to date has been done using this feature, which is more easily interpreted by ray theory than by summing over many modes. Accordingly, in chapter 2 we first discuss the forward problem in terms of ray theory, and subsequently derive the modal interpretation, leading to some duplication.

1.2. The Forward and Inverse Problems

The "forward" or "direct" problem can be stated as follows: given $C(x, y, z)$ and $u(x, y, z)$, together with the characteristics of the sound source, compute the detailed structure of the signal at the receiver. This is the classical problem of finding solutions to the wave equation. The "inverse" problem demands calculation of the ocean properties, $C(x, y, z)$ and/or $u(x, y, z)$, given the properties of the transmitted and received signals. The inverse problem is of considerable interest in oceanography.

Application of an inverse method requires an understanding of the forward problem. When Munk and Wunsch (1979) proposed ocean acoustic tomography, they thought that the forward problem had been "solved." That notion

[5]Ray theory fails for the final, axial phase. A slight discrepancy in the early arrivals evidently has to do with the fact that RR rays do not feel the surface at all, no matter how close they are, whereas modes reaching near to the surface have some slight surface interaction.

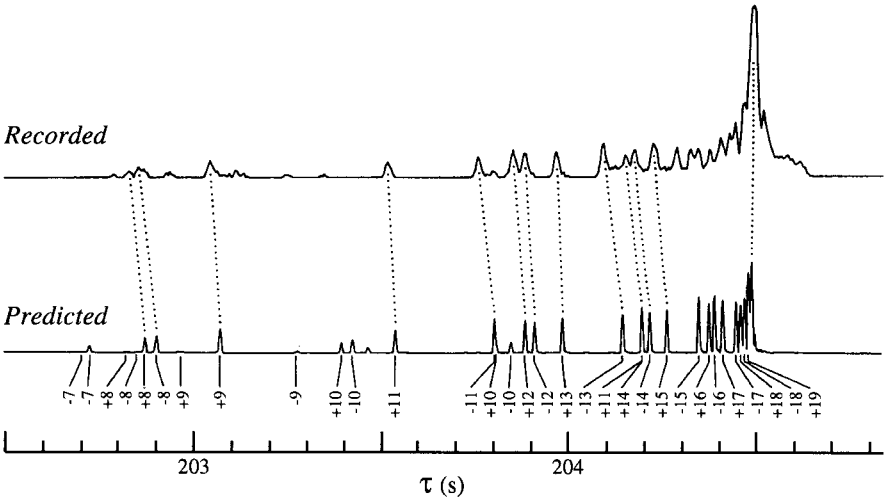

Fig. 1.3. Recorded and predicted (from historical data) axial arrival patterns for 1983, day 217, over a 300-km path west of Bermuda (Howe *et al.*, 1987). The predictions are based on ray theory, modified by Brown (1981). Recorded arrivals are generally earlier, suggesting that temperatures were above their climatological mean.

turned out to be false. Accordingly, much effort in the first decade following 1979 was devoted to the acoustic forward problem, leading to, among other things, a revision of the equation for sound-speed in sea water. Not until 1989 was the ocean adequately sampled to produce a binding comparison between prediction and measurement (Worcester *et al.*, 1994). Some problems remain, of course. For example, referring to fig. 1.2, the wedge, determined by the turning points converging onto the axis, is observed to be broader than computed, and the preceding ray-like pattern is more diffuse.[6]

Returning now to the tomography problem, fig. 1.3 compares the measured arrival pattern to that computed for the climatological-mean ocean at a temperate latitude site. Early arriving rays, which integrate over the entire ocean column, arrive slightly earlier than predicted, indicating that instantaneous temperatures were above their climatological mean. The late arrival, which "sees" only the axial ocean, is close to the prediction. Differences between the observed and predicted patterns thus contain information about the entire sound-speed profile, even though source and receiver are both located at only one depth (here they are both at the sound-channel axis). Inverse methods seek to exploit this information in a systematic way. For a successful inversion, one wishes the

[6] A possible explanation is offered by internal-wave-induced acoustic scattering (section 4.4).