Bulletin of the
Seismological Society of America

Vol. 73 October 1983 No. 5

THE COMPUTATION OF SEISMIC TRAVEL TIMES

By Ray Buland and C. H. Chapman

Abstract

We have developed a method for estimating travel times as a function of epicentral range and hypocentral depth which is faster than ray tracing or direct evaluation of ray integrals yet more compact and general than the interpolation of traditional travel-time tables. In this method, delay or intercept time (tau) as a function of ray parameter and source depth is tabulated. We show that direct manipulation of delay time yields travel time as an explicit function of range eliminating the iteration required to determine the proper ray parameter in ray tracing or the evaluation of ray integrals. Travel-time versus range tables exhibit the same explicit dependence by virtue of having eliminated the common independent variable, ray parameter. However, delay-time branches are monotonic, single-valued, and have fixed ray parameter limits. In contrast, travel-time branches may be nonmonotonic and multi-valued and generally have range limits which vary with source depth. Thus, delay-time tables are simpler to generate and interpolate than are travel-time tables and so result in travel-time estimates which are both more reliable and more precise for a given table size. Further, by retaining the explicit ray parameter dependence and hence a more direct relationship with the ray integrals, the delay-time method offers more flexibility in organizing suites of related phases and in tailoring the algorithm to meet specific computational requirements than does the travel-time table approach. Additionally, we show how the delay-time method can be extended to laterally inhomogeneous media.

Introduction

A problem commonly encountered in wave propagation problems may be formulated as follows: given the positions of a source and a receiver point, and a model of the medium in which they are embedded, estimate the time required for a wave group to propagate from the former to the latter. For seismological purposes, the medium is the Earth, the source is generally an earthquake, an explosion or a rockburst, and the receiver is usually an accelerometer. In most seismic problems, several simplifying assumptions may be made. First, travel time may be defined as the time required for an infinite frequency wave front to propagate from the source point to the receiver point. That is, the geometrical optics approximation is adequate. Second, as the Earth is spherically symmetric to zeroth order and the receivers are always positioned very near the free surface, the source-receiver geometry may conveniently be specified by a hypocentral depth (or radius) and an epicenter-receiver angular distance (or range).

The best-known seismological application of travel times occurs in the problem of locating earthquakes using Geiger's method (Geiger, 1910). In Geiger's algorithm, the location is determined by successive nonlinear least-squares corrections to an assumed source position. Traditionally, travel time is computed either by ray tracing (or equivalently by direct evaluation of the range and travel-time integrals) (Lee

and Lahr, 1975) or by interpolating tables of travel time in both source depth and range (Engdahl and Gunst, 1966). The former technique is computationally unreasonable for all but the simplest Earth models. Not only are the rays (or range and travel-time integrals) expensive to evaluate, but they are functions of a third variable, ray parameter, which is related to ray take-off angle at the source. Thus, for each arrival of interest, one must trace the ray (or evaluate the range integral) a number of times in order to determine the proper take-off angle before the travel time may be evaluated. The latter (table) method generally requires a great deal of mass storage since travel time as a function of epicentral distance can be discontinuous and multi-valued making interpolation awkward. This trouble is aggravated by the fact that the ranges delimiting travel-time branches (segments where travel time is a continuous, monotonic, single-valued function of range) vary with hypocentral depth. In practice, these difficulties serve to limit the flexibility of travel-time tables as each phase of each model may require special treatment.

In this paper, we propose a new method of estimating travel times. Computational efficiency is gained by retaining a table-driven approach. However, ease of table generation is gained by saving neither travel time nor range, but the theoretically more desirable delay or intercept time (tau function) as a function of source depth and ray parameter. The simple behavior of the tau function is well known and has been exploited in inversion studies (Johnson and Gilbert, 1972; Bessonova *et al.*, 1976). Not only is each delay-time branch monotonic and single-valued (even when there is a caustic in the corresponding travel time), but also ray parameters delimiting each branch are simple, invariant functions of the model. Further, tau is simply related to the Earth model and the singular behavior of tau and its derivatives is known analytically. These facts have been used to derive "natural" schemes for interpolating delay time in both source depth and range resulting in sparse ray parameter and model depth sampling and so in relatively compact tables.

Ray Integrals

Consider waves propagating in the half-space $z \leqq 0$, in which velocity, $v(z)$, is a function of depth alone. $v(z)$ may be taken to be either compressional velocity, $\alpha(z)$, or shear velocity, $\beta(z)$, depending upon the type of ray desired. For convenience, assume the ray of interest is propagating in the $x - z$ plane and in the $+x$ direction. Define medium slowness

$$u(z) = v^{-1}(z); \tag{1}$$

horizontal ray slowness (or ray parameter)

$$p = \frac{\sin(i(z))}{v(z)}, \tag{2}$$

where $i(z)$ is the (acute) angle the direction of ray propagation makes with the vertical and vertical ray slowness

$$q(p, z) = (u^2(z) - p^2)^{1/2}. \tag{3}$$

It is well known that ray parameter is invariant along the ray path under these circumstances. The ray may propagate where $p \leqq u(z)$ and will turn (change from

down- to up-going or vice-versa) when $p = u(z)$ (vertical slowness is zero). No geometrical ray may propagate where $p > u(z)$. Thus, both p and $q(p, z)$ are always real and (taking the principal value of the square root) nonnegative. The ray slowness vector is then

$$\boldsymbol{p}(x, z) = p\hat{x} \pm q\hat{z} \tag{4}$$

where + is taken for an up-going ray segment and − for a down-going segment. $\boldsymbol{p}(x, z)$ has magnitude $u(z)$ and points in the direction of ray propagation.

Travel time is given by

$$T(p) = \oint \frac{u^2(z)\, dz}{q(p, z)}, \tag{5}$$

where $\oint$ indicates the integral over the complete ray path. The contribution of every segment of the ray path must be taken as positive. The range integral is given by

$$X(p) = \oint \frac{p\, dz}{q(p, z)}. \tag{6}$$

The delay-time function, $\tau(p)$, is related to $T(X)$ by the Legendre transformation (Goldstein, 1964, section 7.1)

$$\tau(p) = T(p) - pX(p) \tag{7}$$

and

$$p = \frac{dT}{dX}$$

$$X = -\frac{d\tau}{dp}. \tag{8}$$

In analogy with Hamilton's canonical equations, ray parameter corresponds to a generalized momentum. Thus, the invariance of ray parameter is simply a consequence of the fact that the differential equations are cyclic in (independent of) the horizontal coordinates (because medium velocity is a function of depth only). Delay time may also be written as a ray integral,

$$\tau(p) = \oint q(p, z)\, dz. \tag{9}$$

The Earth Flattening Transformation

In seismic problems, it is often necessary to explicitly account for the nearly spherical nature of the Earth. It can be shown that for kinematic purposes, there is an exact conformal mapping between planar and spherical geometries (Gerver and Markushevich, 1966; Müller, 1971). Therefore, equations for the planar geom-

etry may be used without loss of generality. The transformation is

$$x = r_0\Delta$$
$$z = r_0 \ln\left[\frac{r}{r_0}\right]$$
$$v_f(z) = \frac{r_0 v_s(r)}{r}, \tag{10}$$

where Δ is angular distance on the sphere in radians, the subscripts s and f denote spherical and flat geometries, respectively, and r_0 is some reference radius. If r_0 is chosen to be the radius of the Earth then $-z$ corresponds to depth below the free surface and x is identical to arc length on the sphere. Equation (10) maps the spherical coordinates (r, Δ, φ) onto the cylindrical polar coordinates (z, x, φ). The azimuthal angle, φ, is the same in both coordinate systems.

The exact nature of the transformation may be demonstrated by reconsidering the tau integral. Rewriting equation (9),

$$\tau(p_f) = \oint \left[\frac{1}{v_f^2(z)} - p_f^2\right]^{1/2} dz \tag{11}$$

and substituting equation (10) yields

$$\tau(p_f) = \oint (\eta_s^2(r) - p_s^2)^{1/2}\, dr, \tag{12}$$

where

$$\eta_s(r) = \frac{r}{v_s(r)}$$
$$p_s = \frac{r \sin(i_s(r))}{v_s(r)} = r_0 p_f. \tag{13}$$

Equation (12) yields delay time in the spherical coordinate system in terms of angular slowness, η_s, and angular ray parameter, p_s, as defined by Bullen (1963, chapter 7). Note that $i_f(z) = i_s(r)$ as the mapping is conformal.

In spite of the exact nature of the transformation, one note of caution is desirable. As $r \to 0$ near the center of a spherical model, $z \to -\infty$ in the corresponding flat model and the flat model velocity behaves exponentially,

$$v_f(z) \sim v_s(0)e^{-z/r_0}. \tag{14}$$

While this behavior is not a problem in a mathematical sense, it can lead to numerical difficulties. The simplest solution is to choose a model interpolation in this region which gives sensible results for the spherical model and which allows analytical evaluation of equation (9) over a semi-infinite interval.

The Theta Function

Consider the function

$$\vartheta(p, x) = \tau(p) + px \tag{15}$$

or substituting equation (7)

$$\vartheta(p, x) = T(p) + p(x - X(p)). \tag{16}$$

The theta function represents the arrival time of a wave front with horizontal slowness p at range x. This wave front arrives at range $X(p)$ after time $T(p)$. The time required to propagate the extra distance $(x - X(p))$ with horizontal slowness p is $p(x - X(p))$. This property of the theta function is the basis of the WKBJ synthetic seismogram theory (Chapman, 1978). Stationary points of the theta function correspond to geometrical wave group arrivals.

$$\frac{\partial \vartheta(p, x)}{\partial p} = \frac{d\tau(p)}{dp} + x = x - X(p) \tag{17}$$

so

$$\left.\frac{\partial \vartheta(p, x)}{\partial p}\right|_{p_0} = 0 \tag{18}$$

when

$$x = X(p_0). \tag{19}$$

From equation (16), the value of the theta function at the stationary point is just the travel time

$$T(p_0) = \vartheta(p_0, X(p_0)). \tag{20}$$

The stationarity of the theta function at the geometrical arrivals is, of course, an expression of Fermat's principal.

In other words, given $\tau(p)$ and a range x at which the travel times of a suite of phases is desired, it is sufficient to examine the stationary points of the corresponding theta function. Neither $X(p)$ nor $T(p)$ are required. Caustics simply result in multiple stationary points allowing *all* geometrical arrivals to be easily determined. From equation (18) we see that the ray parameter corresponding to each arrival will be determined as a byproduct of the travel-time calculation. Due to stationarity, we expect that travel-time estimates derived from equation (20) will be more accurate than the ray parameter estimates or indeed than the tau function itself (provided it is unbiased on the average).

Additionally, the "travel times" of certain nongeometric diffracted signals may also be determined from the theta function. These include head waves and diffracted waves related to caustics and to shadows:

(a) Head waves are caused by the branch cut in the reflection/transmission coefficient from an interface. If a ray interacts with an interface, any velocity at

the interface (α or β on either side of the interface) which is greater than the actual ray velocity (on either side of the interface) can result in a head wave. Call any such velocity v_{crit} and the corresponding horizontal slowness p_{crit}. If $x \geqq X(p_{crit})$, then (Chapman, 1978)

$$\begin{aligned} T_{head} &= \vartheta(p_{crit}, x) \\ &= T(p_{crit}) + p_{crit}(x - X(p_{crit})). \end{aligned} \tag{21}$$

Although T_{head} can also be computed when $x < X(p_{crit})$, it does not correspond to an arrival in this case. Note that in general, the head wave is accompanied by an interference head wave resulting from waves transmitted through the interface at near critical ray parameter and multiply reflected from the under side of the discontinuity. When the distance that the wave is diffracted along the interface is not too large, T_{head} is a good approximation for the travel time of the interference head wave as well.

(b) On the illuminated side of a caustic, two arrivals exist close together. These correspond to a maximum and minimum in the theta function. At the caustic, these stationary points coalesce and in the nonilluminated region, no geometrical arrival exists. However, the gradient of the theta function will still become very small, giving rise to a diffracted arrival (Chapman, 1978). The wave fronts *almost* add up in a stationary fashion. This is an example of classical Airy diffraction. The travel time can be estimated from

$$T_{Airy} = \vartheta(p_0, x) \tag{22}$$

where

$$\left| \frac{\partial \vartheta(p, x)}{\partial p} \right|_{p_0} = \text{minimum} \tag{23}$$

or equivalently

$$\left. \frac{dX(p)}{dp} \right|_{p_0} = 0. \tag{24}$$

The condition for what constitutes a significant minimum is frequency dependent and is best answered by examining synthetic seismograms. An example of such a caustic is the point B in PKP.

(c) The travel time of the diffracted arrival that exists in a shadow caused by an interface can also be predicted from the theta function. The shadow is caused by the wave front of a turning ray being cut off by the interface. The diffracted signal is partly due to this incomplete wave front (Fresnel diffraction) and partly due to interaction with the interface (an interface mode). The travel-time branch for the turning ray must have a minimum p value, p_{end} (and similarly the reflection from the interface has a maximum p value). The travel time of the diffracted signal is

$$\begin{aligned} T_{shad} &= \vartheta(p_{end}, x) \\ &= T(p_{end}) + p_{end}(x - X(p_{end})). \end{aligned} \tag{25}$$

This only applies in the shadow ($x \geqq X(p_{end})$). Although T_{shad} is defined in the illuminated zone as well, it does not correspond to an arrival. An example of such a shadow is P_{diff} beyond the P/PcP core shadow.

Note that in all cases, the travel-time estimates for these diffracted signals are appropriate for very high-frequency waves. However, it is well known that real diffracted wave groups have a lower frequency content than refracted or reflected waves which have traveled a similar distance. The further the waves have diffracted along the interface in cases *a* and *c* and the further away from the caustic distance the waves arrive in case *b*, the lower the mean frequency content. As a result, the greater the diffraction, the more important the dispersion of the finite frequency wave group becomes. The reliability of the high-frequency travel-time estimates is correspondingly diminished. To study the frequency-dependent apparent travel times of these diffracted signals, it will be necessary to construct synthetic body-wave seismograms. A very convenient method is to use the WKBJ formalism of Chapman (1978) and Dey-Sarkar and Chapman (1978). Not only is this method computationally efficient, but the theta function used for travel-time estimation represents a major portion of the computational labor required by the algorithm. Although the WKBJ approximation is less accurate for diffracted waves than it is for refracted or reflected waves, it can still be quite useful. WKBJ seismograms are accurate for the head wave, the Airy caustic, and Fresnel diffraction, but do not include the interference head wave or interface modes. If more accurate synthetic seismograms are required, several other methods are available, although the computational labor increases dramatically. For a discussion of the relative merits of various techniques, see Burdick and Orcutt (1979) and Choy *et al.* (1980).

Tau Branches

In order to understand how the wide variety of observed seismic phases may be modeled in the theta function, it is useful to examine how the multiplicity of tau branches arise. For purposes of illustration, we have used the PEMC model of Dziewonski *et al.* (1975) throughout. The PEMC angular slowness for compressional and shear waves as a function of radius are shown in Figure 1. Figure 2 shows all the tau branches for down-going compressional waves from a surface source which are turned once by either refraction or reflection. Beginning at the right-hand (high ray parameter) end, each tau branch increases monotonically as ray take-off angle (and so ray parameter) decreases.

Each tau branch is associated with one layer in the model, where a layer is taken to be a region of continuous, nonzero slowness gradient. Thus, a layer may be delimited either by a first-order discontinuity is slowness or by a discontinuity or reversal in slowness gradient (e.g., the lip of a low velocity zone). The largest ray parameter associated with each branch corresponds to the largest take-off angle for which a ray can just penetrate into the relevant layer. If the top of the layer is due to a discontinuous velocity increase, the tau branch begins with a refracted (or "transmitted") segment where rays turn within the layer (solid lines in Figure 2). If the bottom of the layer is also due to a discontinuous velocity increase, the transmitted segment will be followed by a segment where the rays are totally internally reflected (dotted lines in Figure 2). That is, the ray parameter is sufficient to penetrate through the layer, but not through the lower interface. Finally, for even smaller take-off angles, the tau curve bifurcates. The partially reflected portion (dashed lines in Figure 2) we associate with the same tau branch. The partially

transmitted portion is the beginning of the tau branch associated with the next deeper layer.

If the top of the layer is due to a discontinuous velocity decrease, then tau for the layer will be offset from the tau branch for the layer just above. That is, the ray which just grazes the discontinuity divides transmitted and partially reflected regions for the preceding tau branch (there is, of course, no totally reflected portion). This same ray is also the first to penetrate into the low velocity layer. As the ray traverses a finite path in the low velocity layer, a tau offset corresponding to the geometrical shadow results. Depending on the geometry of the low velocity layer, the resulting tau branch may not have a transmitted portion. That is, the first ray to penetrate into the layer may also penetrate right through the layer to be reflected from the lower boundary.

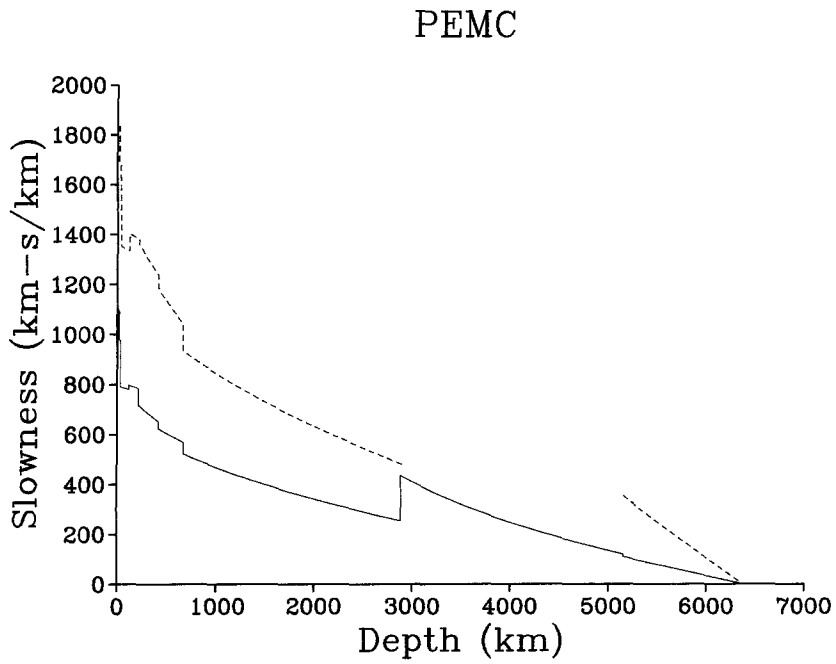

FIG. 1. Angular medium slowness as a function of (radial) depth for Earth model PEMC. The solid line is compressional slowness and the dashed line is shear slowness. Infinite shear slowness in the outer core has been suppressed.

If a low velocity layer is delimited by a second-order discontinuity (i.e., slowness continuous, slowness gradient discontinuous), then the behavior is similar to that for a first-order discontinuity and a partial low-frequency reflection will still result. If a low velocity layer is delimited by reversals in slowness gradient with no discontinuity, then no reflections are possible. However, in this case, trapped waves develop as described in the next section.

Many of these phenomena may be observed in Figure 2. From top to bottom, the tau branches correspond to the inner core, outer core, lower mantle, lower transition zone (420- to 670-km depth), upper transition zone (220- to 420-km depth), low velocity zone, high velocity lid, lower crust, and upper crust. The crustal branches are partially obscured by the abscissa. For more detail, see the expanded view in Figure 4. Low velocity zone behavior is best seen in the core shadow. Note that the

core caustic is not visible in the tau diagram although it is quite clear in the range diagram (Figure 3). This is one of the greatest strengths of the tau approach.

If the source point is placed at a depth below the free surface, the tau diagram is modified somewhat. Tau branches for the layer in which the source is embedded and for deeper layers are qualitatively the same as those in Figure 2. However, branches due to shallower layers are replaced by a single branch corresponding to up-going rays (if we neglect reflections from the under side of discontinuities). At the highest allowable ray parameter, the up-going branch joins the shallowest down-going branch. A minor complication arises when the source is placed in a low velocity zone. In this case, there will be a range of the highest allowable ray parameters for which both up- and down-going rays are trapped by the low velocity channel. In the high-frequency approximation employed here, these trapped waves are never observable. Neglecting them entirely leads to no algorithmic difficulties

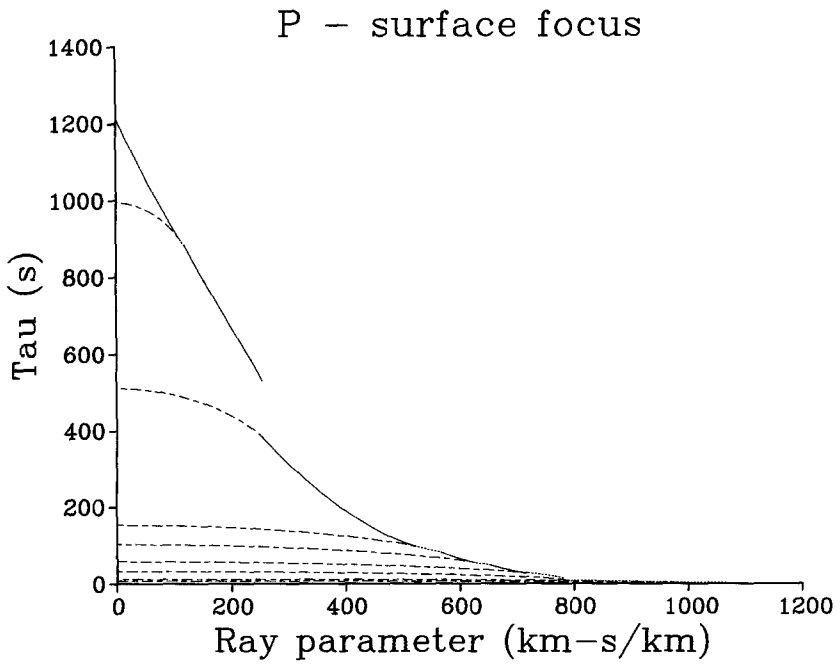

FIG. 2. Compressional wave delay-time branches as a function of ray parameter for all rays which originate from a surface source are turned once in Earth model PEMC and arrive at a surface receiver. Solid lines denote rays turned by refraction. Dotted and dashed lines denote rays turned by reflection. Dotted lines correspond to total reflections and dashed lines to partial reflections.

in estimating the travel times of observable waves. Theoretically, such trapped waves may be observed due to finite frequency effects. However, we know of no observation of this phenomena.

SINGULAR BEHAVIOR

As we have seen, the travel times of whole suites of related body-wave arrivals may be estimated if $\tau(p)$ is known for all relevant branches. From a theoretical point of view this is all that is needed. However, from a computational point of view many other issues must be addressed before a practical algorithm can be developed. For instance, it will be necessary to find methods of calculating tau at selected ray parameters. Further, since it will not be feasible to compute tau for all necessary

ray parameters, some method of estimating it between computed values will also be needed. In both cases, a careful study of singularities in the derivatives of tau is required to insure qualitative correctness for all desired source depths and epicentral ranges for the widest possible range of models.

Both $T(p)$ and $X(p)$ have a singularity in their integrands [see equations (5) and (6), respectively] at any turning point. Provided that the zero in $q^2(p, z)$ is first-order [i.e., $q^2(p, z) = 0$ but $\partial q^2(p, z)/\partial z \neq 0$], the singularity is integrable. If the zero is higher order, then the singularity is not integrable and both the range and the travel time become infinite. For example, if the zero is second-order, then the range behaves logarithmically (Hron and Chapman, 1974). Keeping only the

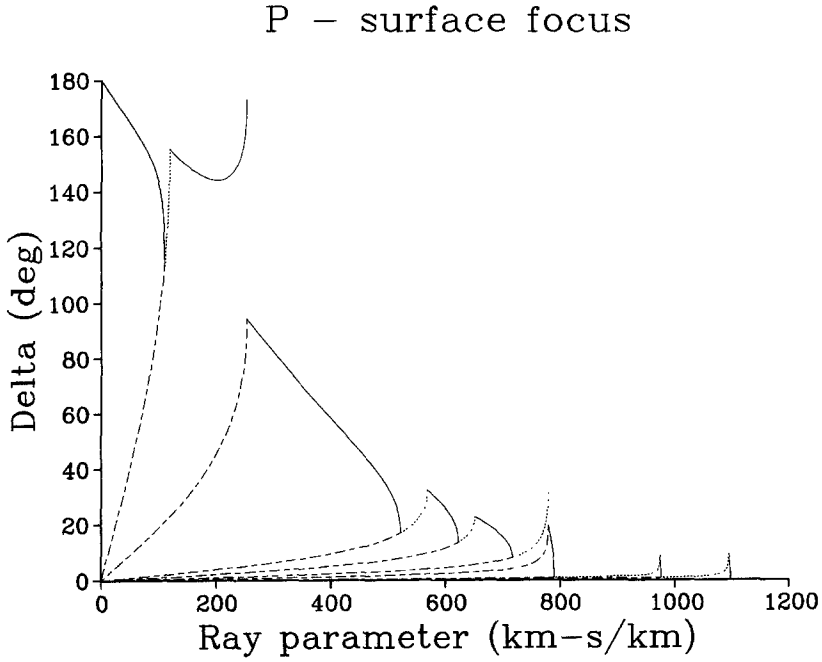

FIG. 3. Compressional wave range as a function of ray parameter for the delay-time branches shown in Figure 2. The solid, dotted, and dashed lines correspond to refracted, totally reflected, and partially reflected rays, respectively.

dominant term

$$X(p) \sim \begin{cases} -\left[\dfrac{u}{u''}\right]^{1/2} \ln(p - u) & p \geqq u \\ -2\left[\dfrac{u}{u''}\right]^{1/2} \ln(u - p) & p \leqq u, \end{cases} \tag{26}$$

where u and u'' are evaluated at the stationary point and prime means differentiation with respect to the argument. From equation (3), it is clear that a second-order zero is $q^2(p, z)$ implies $u'(z) = 0$. This case would arise, for instance, in a model possessing a low velocity zone whose lid has a smooth lip. The qualitative behavior

of such a model is quite different than that of a similar model whose lid has a continuous slowness, but a discontinuous slowness gradient at the lip. In the latter case, range and travel time are bounded, turning rays exhibit a shadow zone, and a weak, low frequency reflection results from the discontinuity in slowness gradient. In the former case, there is neither a shadow zone nor a reflection. However, the waves arriving in place of the shadow zone must be quite weak as waves with a small range of ray parameter are spread over an infinite distance range [equation (26)]. In fact, neither model is realistic. Finite frequency wave groups traversing finite ray tubes respond to averaged model properties and result in seismograms in which differences between these models may not be observable given Earth noise. Lacking observational constraints and given the computational complexity of accommodating second and higher order zeroes in $q^2(p, z)$, we have chosen to restrict the class of acceptable Earth models to those for which reversals in model slowness gradient occur only at depths where $u'(z)$ is discontinuous. In practice, it will be necessary to select discrete model depths and to perform the ray integrals over some interpolation of slowness. A proper selection of interpolation formulae will guarantee that the model used in travel-time computations will always satisfy the above criteria even if the conceptual model does not.

This case aside, $T(p)$ and $X(p)$ [and so $\tau(p)$] are well behaved everywhere. However, the same cannot be said of the second derivative of tau (or equivalently the first derivative of range). To see this, we develop explicit formulas for $X'(p)$. For reflected rays

$$X'(p) = \oint \frac{u^2\, dz}{q^3(p, z)}. \tag{27}$$

For turning rays, however, equation (27) breaks down due to the integrable singularity in $X(p)$ at the turning point. For illustrative purposes, assume that it is a down-going ray which is turned once. Rewrite equation (6) as

$$X(p) = 2 \int_{z_p}^{z^*} \frac{p\, dz}{q(p, z)} + \oint_{z^*} \frac{p\, dz}{q(p, z)}, \tag{28}$$

where z_p is the depth of the turning point, z^* is a shallower depth in the neighborhood of z_p, and $\oint_{z^*}$ is the complete ray integral excluding the depth interval $I_z = (z_p, z^*)$.

Let the interval I_z be sufficiently small that

$$g(z) = \frac{u(z)}{u'(z)} \tag{29}$$

is continuous and positive over I_z. The second integral in equation (28) is well behaved everywhere. For the first integral, we change the variable of integration to u

$$X(p) = 2 \int_{p}^{u(z^*)} \frac{gp\, du}{uq(p, u)} + \oint_{z^*} \frac{p\, dz}{q(p, z)} \tag{30}$$

and integrate by parts

$$X(p) = 2g \sec^{-1}\left[\frac{u}{p}\right]\Big|_p^{u(z^*)} - 2\int_p^{u(z^*)} g' \sec^{-1}\left[\frac{u}{p}\right] du + \oint_{z^*} \frac{p\, dz}{q(p, z)}. \tag{31}$$

Now, both the constant term and the first integrand of equation (31) are zero at the turning point. Finally, we may write (Wesson, 1970; Chapman, 1971)

$$X'(p) = -\frac{2g(z^*)}{q(p, z^*)} + 2\int_p^{u(z^*)} \frac{g'\, du}{q(p, u)} + \oint_{z^*} \frac{u^2(z)\, dz}{q^3(p, z)}. \tag{32}$$

Using equation (32), the behavior of $X'(p)$ may be investigated for each tau branch.

For a reflected branch, the lower p limit is always $p = 0$. Since $X(p) \to 0$ smoothly as $p \to 0$, there should be no particular difficulty in interpolating the tau function in this region. At the upper p limit, $X'(p)$ is singular corresponding to the grazing reflection. Expanding the integral of $X'(p)$ about this upper limit, we obtain

$$X'(p) \sim \frac{(2u)^{1/2}}{u'} (u - p)^{-1/2} \tag{33}$$

(excluding the nonsingular terms) where u and u' are evaluated just above the reflecting interface (the expression assumes that a down-going ray is reflected once from the interface). An example of this behavior is *PcP* at the core shadow.

Again for a turning ray, the lower p limit is not singular. Normally, it corresponds to a grazing ray (e.g., *P* at the core shadow). A special case occurs for the ray in a spherical model which passes through the center of the model. In this case, we have $r = 0$ ($z = -\infty$) and for the limiting turning ray $\Delta = \pi (X = r_0 \pi)$ despite the fact that $p = 0$ (e.g., *PKIKP*). Even at the center of the Earth though, $X'(0)$ is well behaved. However, as for reflections, $X'(p)$ will be singular at the upper p limit of each transmitted ray branch. Two cases may occur. For a ray turning below an interface at which the velocity increases, the upper limit is at the critical point where the turning ray is horizontal and traverses an infinitesimal path just below the interface. Taking z^* just below the interface, we find

$$X'(p) \sim \frac{2g}{q} \sim -\frac{(2u)^{1/2}}{u'} (u - p)^{-1/2} \tag{34}$$

where u, u', g, and q are evaluated just below the interface. An example is point *D* of *PKIKP*. A similar result applies for the turning rays in the surface layer (e.g, *Pg*). The second case occurs when a ray turns in a layer below an interface at which the velocity decreases. Now the turning ray has a significant portion in the lower layer, and the singularity occurs because of propagation through the higher velocity lid. The singular behavior is

$$X'(p) \sim \frac{(2u)^{1/2}}{u'} (u - p)^{-1/2} \tag{35}$$

where u and u' are evaluated just above the interface. An example is point *A* of *PKP*.

Thus, at the upper limit of one segment (either refracted or reflected) of any tau branch, we always have an inverse square root singularity in $X'(p)$

$$X'(p) \sim a(u - p)^{-1/2} \tag{36}$$

where a can be positive or negative, and u is the upper p limit of the branch segment. Therefore,

$$\begin{aligned} X(p) &\sim X(u) - 2a(u - p)^{1/2} \\ \tau(p) &\sim \tau(u) + X(u)(u - p) - \tfrac{4}{3}a(u - p)^{3/2}. \end{aligned} \tag{37}$$

The predicted square root behavior is quite apparent in the range diagram. In Figure 3, range is shown for all down-going compressional wave branches from a surface source which are turned once by either refraction or reflection. The large ray parameter end of the reflected portion (dashed and dotted lines) of each branch exhibits the behavior predicted by equation (33) [$a > 0$ in equation (36)]. The large ray parameter end of the transmitted portion (solid lines) of each branch (save in the outer core) exhibits the behavior predicted by equation (34) [$a < 0$ in equation (36)]. The outer core transmitted branch (*PKP*) exhibits the behavior predicted by equation (35) [$a > 0$ in equation (36)]. The core shadow and *PKP* caustic are quite clear in this domain. Note that the low velocity zone (fourth branch up counting from the abscissa) exhibits the same shadow behavior, but has no transmitted segment. All rays which penetrate this low velocity layer are reflected from the bottom discontinuity. In practice, any scheme for interpolating $\tau(p)$ must explicitly account for these singularities in $X'(p)$ in order to avoid excessive loss of precision and spurious triplications or caustics.

Discretization

The very nature of digital computation requires the discretization of each function needed in this problem. Indeed, efficiency dictates that the theta function be actually calculated at as few discrete ray parameters as possible in order to minimize the number of ray parameter intervals which must be searched for extrema. On the other hand, theta must be sampled densely enough that interpolation over each interval is sufficiently accurate. Although not strictly necessary, it is convenient to employ the same strategy in representing the Earth model itself. In fact, choosing discretizations for ray parameter and for model depth are not necessarily independent problems. It is convenient to sample the model at depths corresponding to slownesses which are identical to discrete ray parameters. Thus, a ray with one of the discrete ray parameters will always bottom at a model depth sample.

In choosing the discretization, certain critical slownesses (and hence model depths and ray parameters) must be sampled exactly. In particular, the slowness just above and below each first-order discontinuity and any local slowness extrema must be sampled to insure that the model is properly represented. As we have seen, these critical points also delimit the transmitted and reflected segments of various tau branches as well, insuring that the resulting travel-time branches will have the proper ranges and will join smoothly onto one another.

Once the critical points have been selected, our problem becomes how to sample both ray parameter and model depth simultaneously. Unfortunately, the criteria

that $\tau(p)$ and $u(z)$ are both well sampled is not sufficient to ensure adequate reconstruction of $T(X)$ from $\vartheta(p, x)$. Figure 1 illustrates the difficulty. Most of the complexity in the model resides in the upper 700 km ($\sim\frac{1}{10}$ of the radius). This portion of the model accounts for about $\frac{1}{6}$ of the distance range for both compressional and shear waves, but nearly $\frac{1}{2}$ of the slowness range.

A practical strategy is to sample the ray parameter between critical points such that range, $X(p)$, is sampled at approximately equal intervals. The form of equation (37) suggests that a quadratic ray parameter spacing,

$$p_j = p_{end} - j^2 \cdot \delta p, \qquad j = 0, 1, 2, \cdots, k \tag{38}$$

where p_{end} and p_k are critical points, should do a reasonable job. In practice, as sampling is determined only once per model, it is worthwhile to fine tune equation (38) by making δp a slowly varying function of ray parameter. This avoids undersampling crustal and upper mantle reflectors and oversampling mantle and core-transmitted segments. In the latter case, relatively small changes in δp over a very long transmitted segment can save oversampling by nearly a factor of two. Unfortunately, making δp nonconstant results in an implicit algorithm for sample determination. In our implementation, for example, given the critical slownesses, a preliminary depth sampling is constructed. Then, between successive critical ray parameters, equation (38) is applied. For each tentative ray parameter sample, $X(p)$ is evaluated using equation (6) and δp is modified as needed. Although messy, this process is not particularly slow as high precision is not required in evaluating $X(p)$.

This procedure must be modified in two cases. For crustal and upper mantle branches, ray parameter intervals can become so large that the interpolation of theta is not sufficiently accurate. This can be controlled by specifying a maximum ray parameter interval. Also, the criteria of equal range intervals breaks down in the neighborhood of a caustic. This case can be treated by detecting the caustic in advance and fixing δp at a suitable value until the caustic is passed. Additionally, care must be taken in the core and lower mantle that the maximum interval between successive depth samples does not become too large. In examples we have tried, this condition has always been satisfied *a priori*.

Once ray parameter and hence model slowness sampling have been determined, all that remains is to determine model depth sampling. Some care must be taken as each slowness sample may correspond to more than one depth sample due to low velocity zones. However, slowness is guaranteed to be a monotonic function of depth between successive critical slownesses. Given such a monotonic interval, an Aiken iteration (Isaacson and Keller, 1966, chapter 3) has been used to determine the depth corresponding to a given slowness sample. This iteration is required because Earth models are generally specified by velocity as a function of depth.

Interpolation of the Model

For discrete calculations such as we are proposing, a complete model specification requires a knowledge of the base model (PEMC in our examples), the discretization, and the interpolation scheme. In our experience, the model interpolation is not critical. Therefore, we have usually used simple two-parameter interpolants for

which the ray integrals may be evaluated analytically. If we further restrict the class of interpolants to methods for which the flat model slowness is monotonic in each depth interval, singularities due to a smooth reversal in slowness gradient are avoided.

For a flat model, some well-known possibilities are plane homogeneous layers

$$\begin{aligned} \Delta T_{ij} &= \Delta z_i \, u_i^2 \, (u_i^2 - p_j^2)^{-1/2} \\ \Delta X_{ij} &= \Delta z_i p_j (u_i^2 - p_j^2)^{-1/2} \end{aligned} \tag{39}$$

and plane layers with constant velocity gradients

$$\begin{aligned} \Delta T_{ij} &= \frac{\Delta z_i}{\Delta v_i} \ln\left[\frac{u_i + (u_i^2 - p_j^2)^{1/2}}{u_{i+1} + (u_{i+1}^2 - p_j^2)^{1/2}}\right] \\ \Delta X_{ij} &= \frac{\Delta z_i}{\Delta v_i}\left[\frac{(u_i^2 - p_j^2)^{1/2}}{u_i p_j} - \frac{(u_{i+1}^2 - p_j^2)^{1/2}}{u_{i+1} p_j}\right], \end{aligned} \tag{40}$$

where the layer is bounded by depths z_i and z_{i+1} with corresponding slownesses $u_i = u(z_i) \geqq u(z_{i+1}) = u_{i+1}$, $\Delta z_i = |z_i - z_{i+1}|$, and $\Delta v_i = u_i^{-1} - u_{i+1}^{-1}$. ΔT_{ij} and ΔX_{ij} are just the travel-time and range increments which result when a ray of parameter p_j passes through the ith depth interval. The corresponding delay-time increment is given by $\Delta \tau_{ij} = \Delta T_{ij} - p_j \Delta X_{ij}$. These definitions guarantee that the contribution of each depth interval is always nonnegative for both travel time and range whether slowness decreases or increases with increasing depth. Note that both formulations have the property that depth and slowness contributions are separable. One could take advantage of this property to modify a model without reevaluating any special functions by changing discrete depth samples, leaving the corresponding discrete slownesses unchanged.

Neither of these formulations is particularly successful when interpolating planar models derived from spherical models via the Earth-flattening transformation. This is because slow velocity variations with radius in the spherical model transform into exponential variations with depth in the planar case. This is, of course, particularly noticeable near the center of the spherical model [equation (14)]. In these cases, it is better to use an interpolation which is sensible in the spherical geometry. An example is the Mohorovičić or Bullen law $v_s(r) = Ar^B$ which transforms to $v_f(z) = ae^{bz}$[with $a = Ar_0^B$ and $b = (B - 1)/r_0$]. This choice yields

$$\begin{aligned} \Delta T_{ij} &= \frac{\Delta z_i}{\ln\left(\dfrac{u_i}{u_{i+1}}\right)}[(u_i^2 - p_j^2)^{1/2} - (u_{i+1}^2 - p_j^2)^{1/2}] \\ \Delta X_{ij} &= \frac{\Delta z_i}{\ln\left(\dfrac{u_i}{u_{i+1}}\right)}\left[\cos^{-1}\left(\frac{p_j}{u_i}\right) - \cos^{-1}\left(\frac{p_j}{u_{i+1}}\right)\right]. \end{aligned} \tag{41}$$

Another example is linear interpolation of the angular slowness $u_s(r) = Ar + B$

which transforms to $u_f(z) = Ae^{z/r_0} + b$. This choice yields

$$\Delta T_{ij} = r_0 \left\{ (u_{i+1}^2 - p_j^2)^{1/2} - (u_i^2 - p_j^2)^{1/2} + b \cdot \ln\left[\frac{u_{i+1} + (u_{i+1}^2 - p_j^2)^{1/2}}{u_i + (u_i^2 - p_j^2)^{1/2}} \right] \right\} + \frac{b^2}{p_j} \Delta X_{ij}$$

$$\Delta X_{ij} = \frac{r_0 p_j}{(b^2 - p_j^2)^{1/2}} \cdot \left\{ \frac{(u_{i+1} - b)[(b^2 - p_j^2)^{1/2}(u_i^2 - p_j^2)^{1/2} + bu_i - p_j^2]}{(u_i - b)[(b^2 - p_j^2)^{1/2}(u_{i+1}^2 - p_j^2)^{1/2} + bu_{i+1} - p_j^2]} \right\}, \tag{42}$$

where

$$B = r_0 b = r_0 \left[u_i - \frac{u_{i+1} - u_i}{(e^{\Delta z_i / r_0} - 1)} \right].$$

Notice that in this and more complicated cases, the depth and slowness contributions are no longer separable.

The Bullen law [equation (41)] can conveniently be used in the interval containing the center of a spherical model if velocity is constrained to be constant. In this special case $z_{i+1} = -\infty$ and $u_{i+1} = 0$. Equation (41) becomes

$$\Delta T_{ij} = r_0 (u_i^2 - p_j^2)^{1/2}$$

$$\Delta X_{ij} = r_0 \cos^{-1}\left(\frac{p_j}{u_i} \right)^{1/2}. \tag{43}$$

Because a constant velocity yields a linear angular slowness, equation (42) gives the same result in this interval. Because our discretization of both ray parameter and slowness are the same, equation (43) will always be evaluated with $p_j = u_{i+1} = 0$ yielding

$$\Delta T_{i,i+1} = r_0 u_i$$

$$\Delta X_{i,i+1} = r_0 \frac{\pi}{2}. \tag{44}$$

In other words, the range integral for zero ray parameter accumulates no distance except when passing through the center of the spherical model.

Several other two-parameter interpolants for $v(z)$ which result in analytical expressions for the ray integrals are known. At least one three-parameter formula with this property is also known (Azbel and Yanovskaya, 1972), but we are not aware of any higher order methods. It is now well known that an alternative is to interpolate $z(u)$. This technique has been widely used in the Soviet Union (e.g., Gurvitch and Nomokonov, 1966) although the first application in western literature appears to be Woodhouse (1974). Although reversals in slowness gradient and plane

constant velocity layers must be treated as special cases, there are an infinite number of interpolation formulas for which the ray integrals may be evaluated analytically.

To see this, rewrite equations (5) and (6), changing the variable to slowness and the range to one slowness interval

$$\Delta T_{ij} = \left| \int_{u_{i+1}}^{u_i} \frac{u^2}{q(p_j, u)} \frac{dz}{du} du \right|$$

$$\Delta X_{ij} = \left| \int_{u_{i+1}}^{u_i} \frac{p_j}{q(p_j, u)} \frac{dz}{du} du \right| . \tag{45}$$

Representing $z(u)$ as a polynomial in u

$$z(u) = \sum_{n=0} a_n u^n \tag{46}$$

it can be shown that

$$\Delta T_{ij} = \left| \sum_{n=1} n a_n p_j^{n+1} \left\{ I_{n+1}\left(\frac{u_i}{p_j}\right) - I_{n+1}\left(\frac{u_{i+1}}{p_j}\right) \right\} \right|$$

$$\Delta X_{ij} = \left| \sum_{n=1} n a_n p_j^{n} \left\{ I_{n-1}\left(\frac{u_i}{p_j}\right) - I_{n-1}\left(\frac{u_{i+1}}{p_j}\right) \right\} \right| , \tag{47}$$

where

$$I_n(x) = \int^x \frac{x^n}{(x^2 - 1)^{1/2}} dx \tag{48}$$

can be evaluated recursively since

$$n I_n(x) = (n - 1) I_{n-2} + (x^2 - 1)^{1/2} x^{n-1} \tag{49}$$

with

$$I_0(x) = \cosh^{-1}(x)$$

$$I_1(x) = (x^2 - 1)^{1/2} . \tag{50}$$

For spherical models, equation (46) is equivalent to

$$\ln(r) = \sum_{n=0} A_n \eta_s^n(r) \tag{51}$$

due to the Earth-flattening transformation, where

$$a_n = \begin{cases} (A_0 - \ln r_0) r_0 & n = 0 \\ A_n r_0^{n+1} & n > 0 \end{cases} \tag{52}$$

This formulation has been used by Jansky and Červený (1980) with even powers. Because of the logarithmic dependence on radius [equation (51)], this formulation is not suitable near the center of a spherical Earth model. An alternative is to include the exponential behavior required by the Earth-flattening transformation [equation (14)] in the expansion of $z(u)$

$$z(u) = r_0 \ln u + \sum_{n=0} b_n u^n \tag{53}$$

which is equivalent to

$$\ln(v) = \sum_{n=0} \left[\frac{b_n}{r_0^{n+1}}\right] \eta_s^n(r) \tag{54}$$

in the spherical model, as suggested by Woodhouse (personal communication, 1981). This formulation leads to

$$\begin{aligned}
\Delta T_{ij} = \Bigg| & r_0 p_j \left\{ I_1\left[\frac{u_i}{p_j}\right] - I_1\left[\frac{u_{i+1}}{p_j}\right] \right\} + \sum_{n=1} n b_n p_j^{n+1} \left\{ I_{n+1}\left[\frac{u_i}{p_j}\right] \right. \\
& \left. - I_{n+1}\left[\frac{u_{i+1}}{p_j}\right] \right\} \Bigg| \\
\Delta X_{ij} = \Bigg| & r_0 \left\{ \cos^{-1}\left[\frac{p_j}{u_i}\right] - \cos^{-1}\left[\frac{p_j}{u_{i+1}}\right] \right\} + \sum_{n=1} n b_n p_j^{n} \left\{ I_{n-1}\left[\frac{u_i}{p_j}\right] \right. \\
& \left. - I_{n-1}\left[\frac{u_{i+1}}{p_j}\right] \right\} \Bigg| .
\end{aligned} \tag{55}$$

Interpolation of Theta

In order to determine travel times from the extremal values of the theta function [equations (15) to (20)], it will be necessary to find zeros in the derivative of some interpolant of $\vartheta(p, x)$ for each range, x, of interest. Clearly, it is inconvenient to interpolate theta for each new range. However, from equation (15), it is apparent that if we can find an interpolation for tau which has a linear term in ray parameter, then the same interpolation will serve for theta for any range. In fact, theta need never be explicitly constructed at all.

For the extremal search to be efficient, it is desirable for the tau interpolant to be both easily differentiated and for zeros in the derivative to be simple to evaluate. Piecewise cubic spline interpolation satisfies all of these requirements: there is a linear term in ray parameter, the polynomial is easy to differentiate, and zeros in the derivative may be found by evaluating a quadratic equation. Further, two extrema per interval are possible permitting realistic modeling near a caustic. Unfortunately, both H splines (Hermite splines fit both the function and its first derivative exactly at each discrete point) and B splines (smooth splines fit the function exactly at each discrete point, the first derivative exactly at the first and last discrete points, and have continuous first and second derivatives everywhere) become unstable near the singular points in $X'(p)$. In both cases, the second derivative of the interpolant, becomes so large that numerical caustics sometimes develop.

Fortunately, an alternative which preserves all of the advantages of piecewise cubic spline interpolation and explicitly accounts for the singular behavior of $X'(p)$ can be constructed. First, divide each tau branch into segments such that the extremal ray parameters of each segment are either zero or singular points of $X'(p)$. Each tau segment will then have exactly one square root singularity in its second derivative at its highest ray parameter. For each segment in turn, call this singular point p_{end}. Then τ splines, as we will call them, are defined by

$$\tau(p) = a_i + b_i(p_{end} - p) + c_i(p_{end} - p)^2 + d_i(p_{end} - p)^{3/2}, \quad p_i \leqq p \leqq p_{i+1}. \tag{56}$$

The range (the derivative of tau) is given by

$$X(p) = b_i + 2c_i(p_{end} - p) + \tfrac{3}{2}\, d_i(p_{end} - p)^{1/2}; \quad p_i \leqq p \leqq p_{i+1} \tag{57}$$

and the first derivative of range is given by

$$X'(p) = -2c_i - \tfrac{3}{4}\, d_i(p_{end} - p)^{-1/2}; \quad p_i \leqq p \leqq p_{i+1} \tag{58}$$

which has a square root singularity at p_{end} as required. Further, theta will have an extremum when

$$2c_i(p_{end} - p) + \tfrac{3}{2}\, d_i(p_{end} - p)^{1/2} + (b_i - x) = 0 \tag{59}$$

which is quadratic in $(p_{end} - p)^{1/2}$.

As for piecewise cubic splines, piecewise τ splines may be constructed in one of two different ways which are completely analogous to H splines and to B splines. The analog of H splines requires less computation to construct than does the B spline analog but nearly twice as much information must be saved [$X(p)$ at all discrete points rather than at the end points of each tau segment]. In either case $X(p)$ at the end points of each segment is fit exactly, guaranteeing that each travel-time branch will cover the proper distance range and will connect smoothly with its neighbors. Surprisingly, experimentation indicates that both methods have almost exactly the same precision in recovering $T(X)$. However, the H spline analog breaks down in the neighborhood of a broad caustic, such as point B of *PKP*. Because there are only weak constraints on the second derivative, which is small, several adjacent ray parameter intervals each yield a numerical caustic.

Thus the B spline analog of the τ splines is the interpolation method of choice. While its construction is more complicated than that for the H spline analog, the required computation is still not excessive. Our experience has been that proper τ interpolation is crucial to the success of this method. As the only completely successful method known has been invented here for this purpose, details of the algorithm are given in Appendix. As an aside, it is worth commenting on the comparable precision of these methods in estimating $T(X)$. Because the H spline analog satisfies nearly twice as many constraints, one would expect that it is the more accurate interpolation formula. In fact, this is true for interpolating $\tau(p)$. The apparently contradictory result for $T(X)$ is due to the stationarity of the theta function at each geometrical arrival. Thus, errors in $T(p)$ and $X(p)$ are correlated in such a way that errors in $T(X)$ are, to first order, independent of errors in ray parameter, p. Consequently, in choosing the B spline rather than the H spline

analog of τ splines, we sacrifice precision in the ray parameter estimates and some computational efficiency for improved interpolation stability and a 50 per cent reduction in table size.

In practice, one additional τ spline stability problem must be addressed. We permit only one discretization of ray parameter per model, and it has been designed to adequately sample once refracted and once totally reflected portions of all down-going tau branches. Thus, up-going and partially reflected down-going branch segments, for instance, are often heavily and very nonuniformly sampled. Rapid fluctuations in ray parameter sample interval sometimes induce oscillations in the τ spline interpolation. The same phenomenon is also exhibited by both H and B spline interpolations, although the amplitude of the oscillations is considerably smaller. Fortunately, this problem may be eliminated by decimating the ray parameter sampling for affected branches. Simply insuring that range is not sampled too finely has been adequate in all cases we have encountered.

Finally, notice that once the interpolation is performed, both $\tau(p)$ and $X(p)$ are known everywhere. This fact greatly reduces the computational effort required to find extrema of theta, since the range interval to which each ray parameter interval contributes is known in advance. With a little care, entire branches may be excluded from a given search with a single test. For contributing branches, ray parameter intervals with no arrivals may also be excluded so that the quadratic [equation (59)] need be solved only when desired solutions are known to exist.

Interpolation of Tau with Source Depth

The tau tables are discretized in both ray parameter and model depth. The results for interpolating $\tau(p)$ may be applied only if the desired source depth coincides with one of the discrete model depths. Although $\tau(p)$ may be interpolated between two successive model depths, subtle inconsistencies between the specified Earth model and derived travel-time branches may arise. While these inconsistencies are negligible for most purposes, they may be avoided completely with only a modest increase in computational labor.

The desired result is achieved by taking advantage of the simplicity of the analytical partial ray integral formulas presented above [e.g., equations (39) to (55)]. In other words, given the model interpolation employed in the initial construction of the tau tables, $\tau(p)$ at any desired source depth is derived by adding (or subtracting) partial ray integral contributions to (or from) $\tau(p)$ at the nearest discrete model depth. In addition, it will be necessary to compute both tau and range for the ray parameter which bottoms at the desired source depth. This is an extremal point for both the up-going and the shallowest down-going tau branches. The total labor involved is slightly less than computing both range and delay time for a ray which passes straight through the flat model or through the center of a spherical model. For applications involving many travel-times estimates per source depth, the cost of this additional overhead is negligible.

Organization

Because the theta-function technique depends on mathematics rather than on "brute-force" as in traditional travel-time tables, an unusual degree of flexibility is possible in tailoring the algorithm for a particular problem or computational environment. The organization of the tau tables governs the most important of these trade-offs. Several possibilities which we have examined in detail serve to illustrate this point.

Maximum versatility is gained by saving $\Delta\tau_{ij}$ in the table. Because no assumptions have been made about user requirements, the tau branch corresponding to *any* type of geometrical ray path can be constructed by performing the appropriate sums. If one is willing to restrict the class of model interpolants to those which are separable in depth and slowness [e.g., equations (39) to (41)], then the table might be reduced substantially. This is due to the fact that the slowness portion of the partial ray integrals can be identical for more than one depth interval if a velocity reversal has occurred. Note that one may take advantage of the separability in this way only if the model discretization is done by finding depths corresponding to discrete slownesses. As pointed out before, separability also offers the advantage that the model may be perturbed by modifying only the discrete depths which leaves the table containing the slowness contributions to $\Delta\tau_{ij}$ unchanged. In our experience, proper discretization of ray parameter is so model dependent that major alterations in a model may create difficulties in the interpolation of tau.

Unfortunately, versatility is expensive. Saving $\Delta\tau_{ij}$ in the table requires large table mass storage and, as the entire table is required to construct the tau branches, considerable input time and program address space are required as well. Further, the computational time and algorithmic complexity required to construct the tau branches is substantial. If one is willing to sacrifice the ability to perturb the model, an obvious solution is to save the partial ray integrals over larger segments of the ray paths. For example, if one saved the tau contribution due to each model layer then the table would be greatly reduced in size, but it would still be possible to construct the tau branch corresponding to any type of geometrical ray path from a surface source. In order to generalize this procedure to any source depth, it is also necessary to save the tau contribution due to a possible up-going branch (between the free surface and a discrete model depth) for all model depths shallower than the deepest source of interest. All possible up-going branches will usually comprise the bulk of the table, but at most one such branch is required by the algorithm for any given source depth. In extreme cases, the table itself may rival the $\Delta\tau_{ij}$ table in size. However, program input time, address space, and computational time needed to construct tau branches of interest are all greatly reduced.

If one is willing to restrict the class of desired ray paths, it is possible to do even better. For example, in earthquake location work, it is rare that branches other than up-going, once reflected or refracted down-going, and surface reflected down-going are required. By examining the relevant tau and range diagrams (Figures 2 to 4), it is clear that for these branches tau is a multi-valued function of ray parameter only because of partial reflections (dashed lines). If all partial reflections are discarded, the storage required for all remaining down-going contributions may be collapsed into a single vector greatly simplifying algorithmic logic. The table in this case is comprised only of the two-way surface focus tau "branch" and the up-going contribution from each discrete depth of interest. One sum and one difference of surface focus tau and one up-going branch are sufficient to construct up-going, down-going, and surface reflected down-going branches.

An Example

The problem of generating travel times for teleseismic earthquake location provides an illustration of many of the capabilities of the tau-table method. For this example, we have used the third (least flexible-most efficient) organization discussed. The spherical Earth model PEMC was used with the piecewise linear slowness interpolation. Equation (42) was used to evaluate all ray integrals. All

ranges in the interval (0, 180) degrees were allowed as well as source depths in the interval (0, 700) km. Requiring that range be sampled approximately every 200 km or less and that the maximum ray parameter interval not greatly exceed 11 km-sec/km resulted in 215 discrete ray parameter samples for the compressional velocity structure. The maximum sample interval in depth was less than 115 km and occurred in the deep outer core.

A surface focus P-wave travel-time diagram (Figure 5) was constructed by sampling all branches at 10-km intervals in range. P_{diff} (the dotted line) was generated using equation (25) and was arbitrarily truncated at a range of 120°. The multiplicity of branches at ranges less than 30° is due to the numerous upper mantle discontinuities in this model (Figure 1). The upper mantle low velocity zone produces

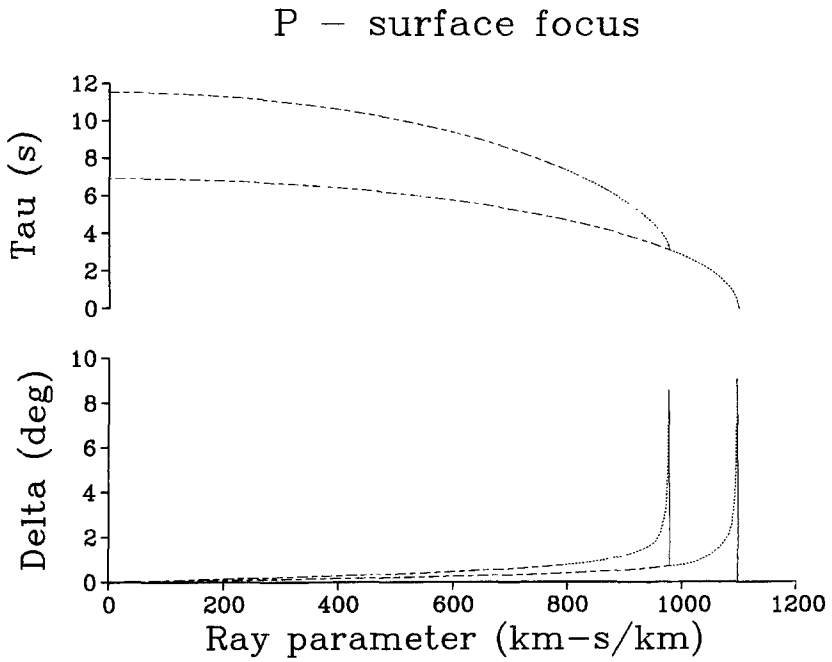

FIG. 4. Compressional delay time (*upper panel*) and range (*lower panel*) as functions of ray parameter for the crustal delay-time branches of Earth model PEMC. These branches have been reproduced on an expanded scale from Figure 2 and 3 where they were partially obscured by the abscissas. The solid, dotted, and dashed lines correspond to refracted, totally reflected, and partially reflected rays, respectively.

a shadow which is masked by the reflection from the bottom of the layer. Figure 6 illustrates the error in $T(X)$ resulting from the τ spline interpolation of $\tau(p)$, for all branches (excluding p_{diff}). In other words, the $T(X)$ derived from the theta function extrema are compared with the results of direct integration through the same model using the same depth sampling and slowness interpolation. As the maximum error is less than 7 msec, the interpolation may be considered to be perfect in the context of teleseismic earthquake location where observations are usually made only to the nearest tenth of a second. Notice that interpolation errors are similar to observational errors in being roughly independent of range. The error curves are apparently oscillatory because the τ splines provide range estimates which are, on the average, correct. Also for some of the core branches, the curve looks like a square wave. The error is either zero or 0.12 msec corresponding,

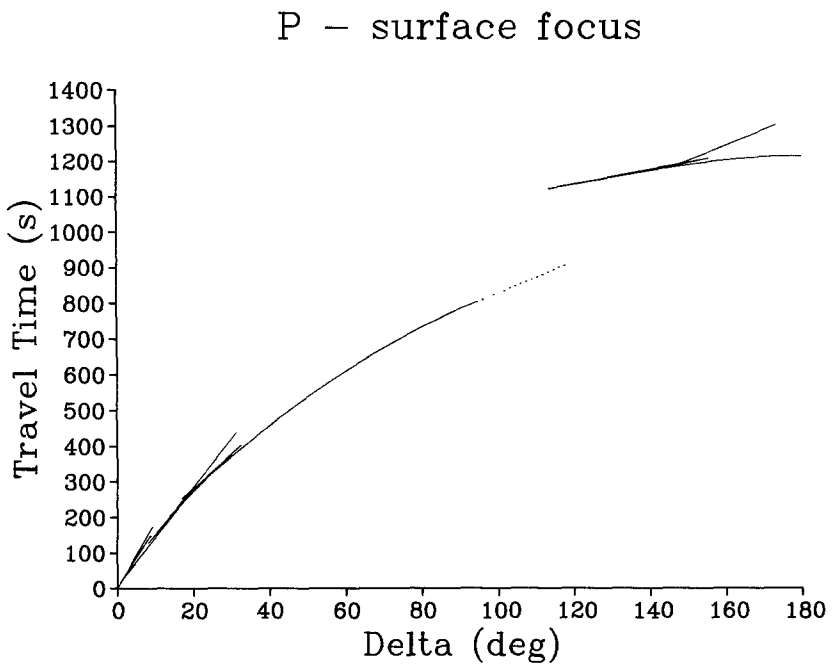

FIG. 5. Compressional wave travel time as a function of angular range from a surface source to surface receivers. The solid lines denote rays which have been refracted or reflected. The dotted line denotes diffraction along the core-mantle interface. All rays which are partially reflected, converted, or turned more than once have been neglected.

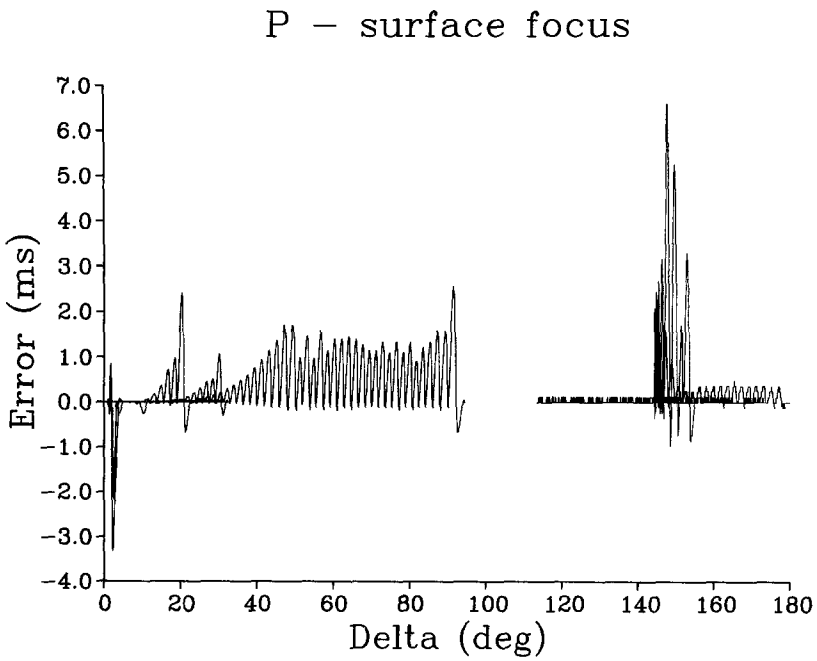

FIG. 6. Interpolation error in milliseconds as a function of angular range for all (geometrical) arrivals shown in Figure 5. Details of the curves are somewhat obscured as travel time and so travel-time error are multi-valued functions of range.

probably, to a least significant bit error in calculating some parameter internal to the error estimation. The fact that the errors are predominantly above or below zero for each branch segment is due to the stationarity in estimating travel time. Figure 6 implies that the number of ray parameter samples might be reduced by factors of 2 to 5 before errors in $T(X)$ became unacceptable. However, care must be taken that depth is sampled sufficiently often that the model interpolation is adequate.

Figure 7 shows the compressional wave travel-time diagram for a source depth of 700 km. The dashed lines correspond to surface reflected down-going rays (e.g., *pP* and *pPKP*). Travel-time errors for the 700-km source depth are shown in Figure 8. Curves for down-going (solid lines) and surface reflected down-going (dashed lines)

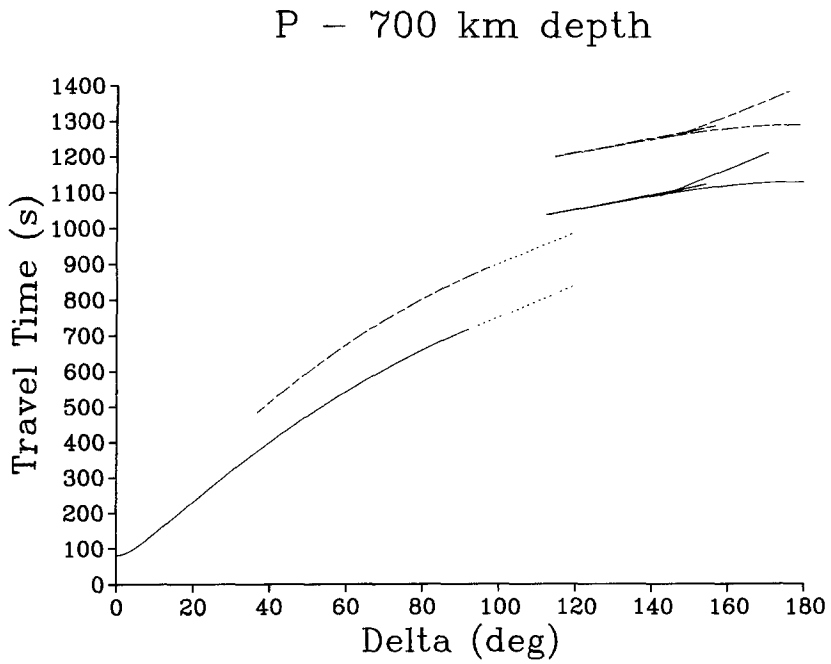

FIG. 7. Compressional wave travel time as a function of angular range from a 700-km deep source to surface receivers. Solid lines show direct refracted or reflected arrivals, dashed lines show surface reflected arrivals, and dotted lines show diffracted arrivals.

branches overlap. The error distribution for 700 km and surface focus source depths are qualitatively the same. Furthermore, the surface reflected branches have error curves that closely mirror the error curve for the analogous down-going branch, but shifted in distance. That is, the error characteristics seem to depend only on ray parameter.

Figures 9 and 10 show similar diagrams for surface focus shear waves. In generating these diagrams, the PEMC shear structure was used in the crust and mantle, but the compressional structure was used throughout the core. The ray parameter discretization was created under the same conditions as for the compressional wave structure, but resulted in 319 samples. The low velocity zone shadow is more pronounced in shear than in compression. Further, a small caustic is faithfully reproduced in the branch which turns below the low velocity zone, but is not visible

on this scale. Figure 10 shows that maximum interpolation errors for S waves are of the same order as those for P waves, although the distribution is, of course, different. Note that there is no problem in principal in treating converted phases (e.g., *SKS*). As in any other ray-interface interaction, a P to S conversion (or vice versa) is treated by simply conserving ray parameter. For these examples, mantle compressional and shear travel times were generated from separate tables. If phases are desired for which both compressional and shear velocity structures in one depth range are required (e.g., *sP* or *SKP*), then it would be more convenient to merge the two tables in some manner. A minimum requirement is that the P-wave ray parameter samples be a subset of the S-wave ray parameter samples. While it appears that there is no difficulty in principle, such a scheme has yet to be tried.

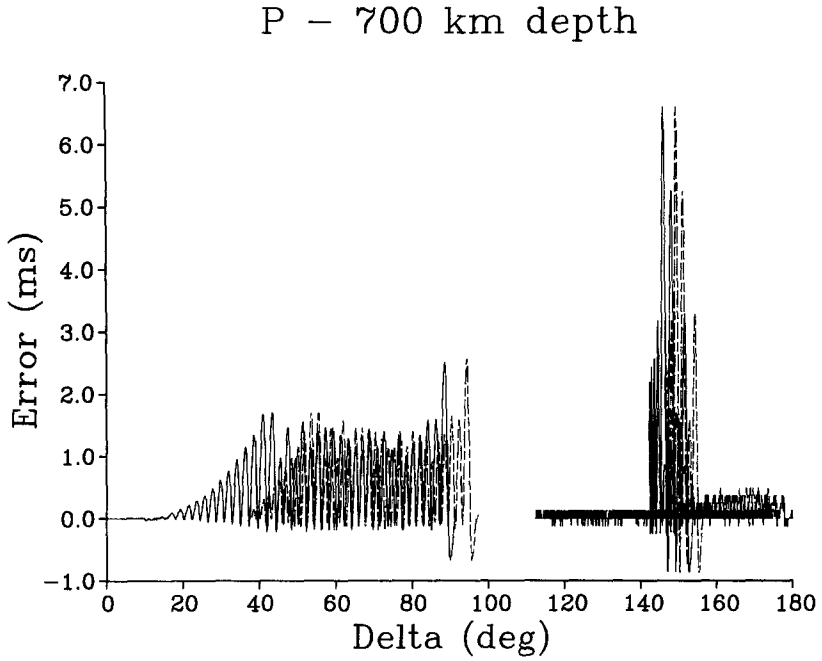

FIG. 8. Interpolation error as a function of angular range for all (geometrical) arrivals shown in Figure 7. Solid lines correspond to direct and dashed lines to surface-reflected phases.

Figures 5 to 10 show quite clearly that even the modest table dimensions used are more than adequate for the teleseismic earthquake location problem. From our experiments, we estimate that given suitable tuning this algorithm is fast enough for even high volume applications such as the generation of the National Earthquake Information Service's "Preliminary Determination of Epicenters." Furthermore, as tables can be generated quickly and easily for any model, there is no reason that this approach could not be used on a regional or even a local scale. In these cases, there is, of course, no need to include the Earth model below the bottoming point of the deepest rays of interest which will greatly decrease the size of the required tables. The tau table algorithm has a number of significant advantages over homogeneous, plane layer techniques currently in use including proper modeling of down-going rays, of later arriving phases, of low velocity zones, and of surface reflected phases.

Laterally Varying Models

The concepts of the tau and theta functions, and many of their properties can be extended to laterally inhomogeneous media. The problem of interpolating travel-time tables is more serious than in laterally homogeneous models as more types of travel-time curves and discontinuities are possible. Additionally, the ray integrals are more expensive to evaluate. We shall not discuss the solution of the ray equations here, but restrict ourselves to the use of tables evaluated by the ray-shooting method. For simplicity, initially, we will restrict our discussion to two-dimensional models in which the rays remain in a plane.

As before, we can evaluate the range and travel-time for an initial ray direction. The ray parameter (horizontal slowness) is no longer conserved along the ray, but

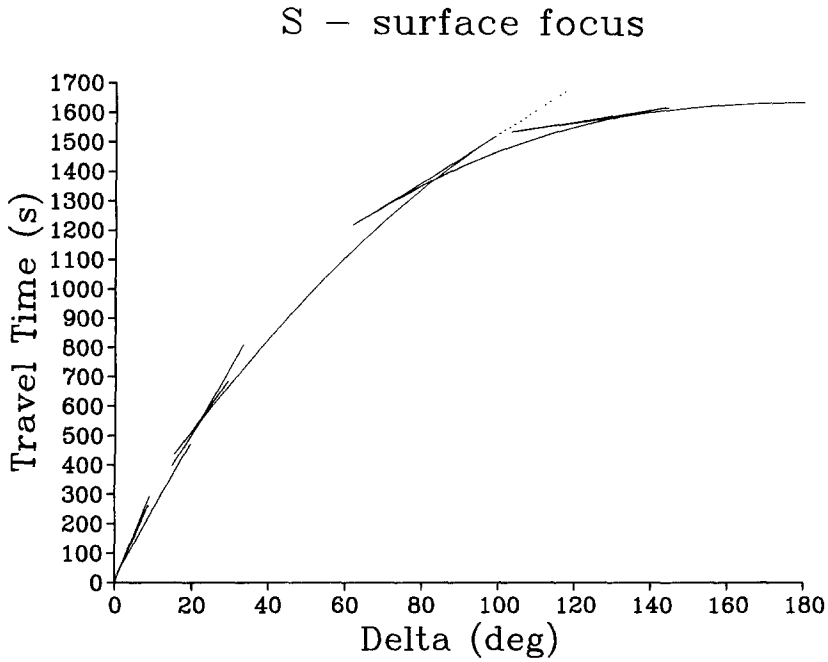

FIG. 9. Shear wave travel time as a function of angular range from a surface source to surface receivers. Compressional velocities were used in the core so all core phases are twice converted (e.g., SKS and SKIKS). Solid lines show direct refracted or reflected arrivals, and the dotted line shows diffraction along the core-mantle interface.

we can still define the horizontal slowness at any point on the ray,

$$p = \frac{\sin i}{v}, \tag{60}$$

where i is the (acute) angle the ray makes with the vertical at a point where the medium velocity is v. We may equally well regard the travel-time and range integrals as functions of p at the source, p_0, or at the receiver, p_r. As in laterally homogeneous models, the gradient of the travel-time curve gives the horizontal slowness at the receiver

$$p_r = \frac{dT}{dX}. \tag{61}$$

This being the case, we will consider all relevant functions to be functions of p_r in the following. As no other horizontal slowness need be considered we will drop the subscript and p will always mean p_r. Proceeding, we can also define the intercept or delay time

$$\tau(p) = T(p) - pX(p) \tag{62}$$

and the theta function

$$\vartheta(p, x) = \tau(p) + px \tag{63}$$

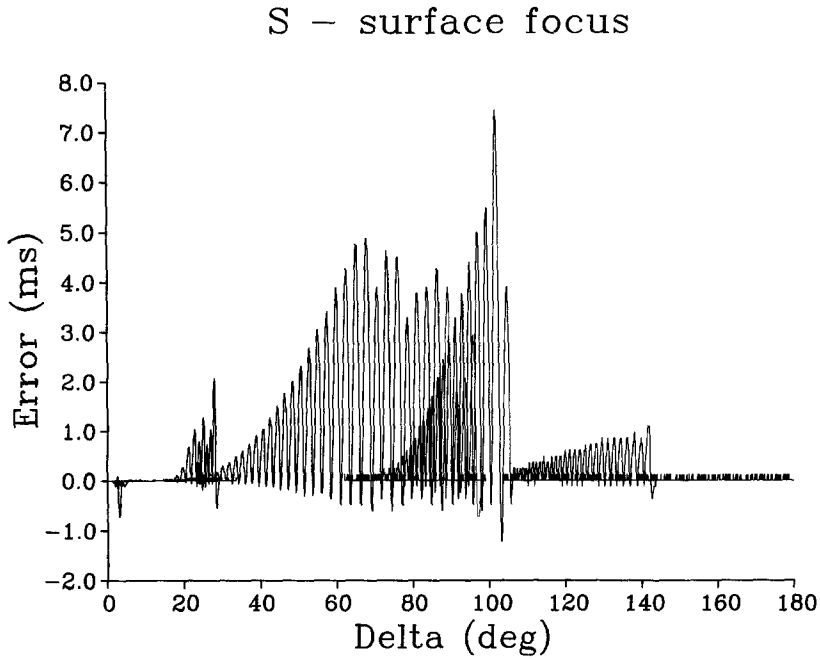

FIG. 10. Interpolation error as a function of angular range for all (geometrical) arrivals shown in Figure 9.

as before. Also

$$\frac{d\tau(p)}{dp} = \frac{dT(p)}{dp} - p\frac{dX(p)}{dp} - X(p). \tag{64}$$

Applying the chain rule to equation (61) and substituting into equation (64) yields

$$\frac{d\tau(p)}{dp} = -X(p) \tag{65}$$

as before. Hence, the basic property that $T(X)$ and $\tau(p)$ are related by the Legendre transformation is preserved and geometrical arrivals will correspond to extrema of the theta function in exactly the same ways as for homogeneous media.

The added complication in laterally inhomogeneous media is that the tau function

need no longer be a single-valued function of p, although the various functions are necessarily single-valued functions of p_0. In inhomogeneous media, the derivative $\partial p/\partial p_0$ can change sign (it is, of course, unity in a laterally homogeneous model). Points at which

$$\frac{\partial p}{\partial p_0} = 0 \tag{66}$$

correspond to reversals or "caustics" in $\tau(p)$. Frazer and Phinney (1980) have called these *telescopic points*. Regarding tau and theta as functions of the source horizontal slowness, p_0, they are single-valued functions, but tau is no longer monotonically decreasing, and theta has stationary points when $\partial p/\partial p_0 = 0$ which do not correspond to geometrical arrivals as

$$\frac{\partial \tau(p_0)}{\partial p_0} = -X(p_0)\frac{\partial p}{\partial p_0} \tag{67}$$

and

$$\frac{\partial \vartheta(p_0, x)}{\partial p_0} = (x - X(p_0))\frac{\partial p}{\partial p_0}. \tag{68}$$

These points can be easily detected when locating geometrical arrivals. Maslov (1965, 1972) has developed the ideas behind transformations between the spatial and slowness domains in more generality. Kravtsov (1968) has provided a concise review of Maslov's asymptotic theory that is available in English, and Chapman and Drummond (1982) have used this theory to extend WKBJ synthetic seismograms to laterally varying models.

These ideas are easily extended to three-dimensional models. We consider Cartesian coordinates where, for notational simplicity, we call the horizontal coordinates $\boldsymbol{x} = (x_1, x_2)$. We define a horizontal slowness vector $\boldsymbol{p}$ (p_1, p_2) which varies along the ray, and call the value at the source $\boldsymbol{p}_0$. The horizontal slowness vector is related to the travel time by

$$\boldsymbol{p} = \bar{\nabla} T(\boldsymbol{X}) \tag{69}$$

where the gradient operator, $\bar{\nabla}$, acts in the horizontal direction only. We define the tau function

$$\tau(\boldsymbol{p}) = T(\boldsymbol{p}) - \boldsymbol{p} \cdot \boldsymbol{X}(\boldsymbol{p}) \tag{70}$$

which can be regarded as a function of the initial slowness, $\boldsymbol{p}_0$, or the receiver slowness, $\boldsymbol{p}$. T and $\boldsymbol{X}$ are the travel-time and range functions to a fixed, horizontal receiver surface (i.e., z = constant) and again can be regarded as functions of either $\boldsymbol{p}_0$ or $\boldsymbol{p}$. Letting the various functions be functions of $\boldsymbol{p}$, in analogy with equation (65)

$$\bar{\nabla}\tau(\boldsymbol{p}) = -\boldsymbol{X}(\boldsymbol{p}). \tag{71}$$

Defining the theta function

$$\vartheta(\boldsymbol{p}, \boldsymbol{x}) = \tau(\boldsymbol{p}) + \boldsymbol{p} \cdot \boldsymbol{x} = T(\boldsymbol{p}) + \boldsymbol{p} \cdot (\boldsymbol{x} - \boldsymbol{X},(\boldsymbol{p})). \tag{72}$$

where $\boldsymbol{x}$ is the receiver position, we have

$$\nabla\vartheta(\boldsymbol{p}, \boldsymbol{x}) = \boldsymbol{x} - \boldsymbol{X},(\boldsymbol{p}) \tag{73}$$

which again is zero for geometrical arrivals. As before when $\boldsymbol{x} = \boldsymbol{X}$, $\vartheta = T$. As for the two-dimensional case, τ and ϑ can be multi-valued when regarded as functions of $\boldsymbol{p}$. Regarded as functions of $\boldsymbol{p}_0$, they are necessarily single-valued, but the theta function is stationary at telescopic points where the Jacobian is zero

$$\frac{\partial(p_1, p_2)}{\partial(p_{0_1}, p_{0_2})} = 0 \tag{74}$$

which do not correspond to geometrical arrivals.

Discussion

In this paper, we have presented a new method for computing the travel times of high-frequency body waves in spherically symmetric Earth models. We have first developed a complete theoretical framework. Second, we have restricted the class of phenomena of interest so as to simplify the algorithm and improve computational performance with no detriment for most purposes. Third, we have presented "cook book" procedures and formulas for discretizing model depth and wave ray parameter, evaluating ray integrals, interpolating $\tau(p)$, locating extrema of $\vartheta(p, x)$, and organizing the tau tables. Fourth, we have demonstrated the precision of the method and its ability to correctly represent complex geophysical phenomena including shadow zones, caustics, and "triplications" due to model discontinuities even at teleseismic distances. Finally, we have shown how this method may be extended to treat laterally inhomogeneous structures. We feel that this algorithm has many applications in seismology by virtue of its precision and computational efficiency, its flexibility in meeting specific computational requirements, its ability to be easily tailored to determine as many or as few secondary arrivals as desired, and its robustness in being able to treat a wide range of geophysically interesting Earth models.

A side effect of this technique is the easy availability of estimates of $\tau(p)$ and $X(p)$ for all ray parameters, source depths, and branches of interest. Thus, the same tables and τ spline interpolation can provide a basis for other types of calculation as well. One example has already been mentioned. Evaluating the theta function is a major portion of the computational labor required to construct WKBJ synthetic body-wave seismograms. For some purposes (e.g., phase identification), approximate body-wave amplitude is of more interest than the waveform itself. We suggest that such an amplitude might be derived from the WKBJ theory. Although such an estimate can be misleading in certain special cases, it will always be far more reliable than the infinite frequency estimate. Another possibility is to extract incremental range from the tables at each model depth for which tau information is saved. In this way, an approximation of the ray path geometry (or at least the

segment which is shallower than the deepest allowable source depth) can be constructed. Buland (1982) has suggested that integration of a velocity perturbation along such a ray path segment might be used to include the travel-time contribution of a three-dimensional structure in an otherwise one-dimensional model by means of first-order perturbation theory. As we have seen, three-dimensional structures can be treated directly. However, for certain purposes (most notably earthquake location), the computational speed of the perturbation approach is attractive. Further, elimination of infinite frequency shadows and multipathing can stabilize an earthquake location procedure with little or no loss of precision in the solution.

ACKNOWLEDGMENTS

This research was partly supported by a Natural Science and Engineering Council of Canada Operating Grant, A-9130. We thank Jim Dewey and George Choy for critically reading the manuscript.

REFERENCES

Azbel', I. Ya and T. B. Yanovskaya (1972). Approximation of velocity distributions for calculation of *P*-wave times and amplitudes. In *Computational Seismology*, V. I. Keilis-Borok, Editor, Consultants Bureau, New York, 62–69.

Bessonova, E. N., V. M. Fishman, M. G. Shnirman, G. A. Sitnikova, and L. R. Johnson (1976). The tau method for inversion of travel-times. II. Earthquake data, *Geophys. J. R. astr. Soc.* **46,** 87–103.

Buland, R. (1982). Towards locating earthquakes in a laterally heterogeneous medium, *Phys. Earth Planet. Interiors* **30,** 157–160.

Bullen, K. E. (1963). *An Introduction to the Theory of Seismology*, Cambridge University Press, Cambridge, England.

Burdick, L. J. and J. A. Orcutt (1979). A comparison of the generalized ray and reflectivity methods of waveform synthesis, *Geophys. J. R. astr. Soc.* **58,** 261–278.

Chapman, C. H. (1971). On the computation of seismic ray travel times and amplitudes, *Bull. Seism. Soc. Am.* **61,** 1267–1274.

Chapman, C. H. (1978). A new method for computing synthetic seismograms, *Geophys. J. R. astr. Soc.* **54,** 481–518.

Chapman, C. H. and R. Drummond (1982). Body wave seismograms in inhomogeneous media using Maslov asymptotic theory, *Bull. Seism. Soc. Am.* **72,** S277–S317.

Choy, G. L., V. F. Cormier, R. Kind, G. Müller, and P. G. Richards (1980). A comparison of synthetic seismograms of core phases generated by the full wave theory and by the reflectivity method, *Geophys. J. R. astr. Soc.* **61,** 21–39.

Dey-Sarkar, S. K. and C. H. Chapman (1978). A simple method for computing synthetic seismograms, *Bull. Seism. Soc. Am.* **68,** 1577–1593.

Dziewonski, A. M., A. L. Hales, and E. R. Lapwood (1975). Parametrically simple Earth models consistent with geophysical data, *Phys. Earth Planet. Interiors* **10,** 12–48.

Engdahl, E. R. and R. H. Gunst (1966). Use of a high speed computer for the preliminary determination of earthquake hypocenters, *Bull. Seism. Soc. Am.* **56,** 325–336.

Frazer, L. N. and R. A. Phinney (1980). The theory of finite frequency body wave synthetic seismograms in inhomogeneous elastic media, *Geophys. J. R. astr. Soc.* **63,** 691–718.

Geiger, L. (1910). Herdbestimmung bei Erdbeben aud den Ankunftzeiten (in German), *K. Gessell. Wiss. Goett.* **4,** 331–349.

Gerver, M. and V. Markushevich (1966). Determination of a seismic wave velocity from the travel-time curve, *Geophys. J. R. astr. Soc.* **11,** 165–173.

Goldstein, H. (1964). *Classical Mechanics*, Addison-Wesley, Reading, Pennsylvania.

Gurvitch, I. I. and V. P. Nomokonov (Editors) (1966). *Geophysical Handbook* (in Russian), Publish House NEDRA, Moscow, USSR.

Hron, M. and C. H. Chapman (1974). The 'Shadow' from a velocity reversal, *Bull. Seism. Soc. Am.* **64,** 25–32.

Isaacson, E. and H. B. Keller (1966). *Analysis of Numerical Methods*, John Wiley and Sons, New York.

Jansky, J. and V. Červený (1980). Fast and stable computation of ray integrals and ray amplitudes in radially symmetric media, ESC Meeting, Abstract.

Johnson, L. E. and F. Gilbert (1972). Inversion and inference for teleseismic ray data, *Meth. Comp. Phys.* **12,** 231–266.

Kravtsov, Yu. A. (1968). Two new asymptotic methods in the theory of wave propagation in inhomogeneous media (review), *Societ Physics-Acoustics* **14,** 1–17.

Lee, W. H. K. and J. C. Lahr (1975). HYP071 (revised): a computer program for determining hypocenter, magnitude, and first motion patterns of local earthquakes, *U.S. Geol. Surv., Open-File Rept. 75–311.*

Maslov, V. P. (1965). *Theory of Perturbations and Asymptotic Methods* (in Russian), Izd. MGU, Moscow, USSR.

Maslov, V. P. (1972). *Théories des Perturbations et Méthodes Asymptotiques*, Dunod, Paris, France

Müller, G. (1971). Exact ray theory and its application to the reflection of elastic waves from vertically inhomogeneous media, *Geophys. J. R. astr. Soc.* **21,** 261–283.

Wesson, R. L. (1970). A time integration method for computation of the intensities of seismic rays, *Bull. Seism. Soc. Am.* **60,** 307–316.

Woodhouse, J. H. (1974). Aspects of high frequency seismic wave propagation, *Ph.D. Thesis*, University of Cambridge, Cambridge, England.

U.S. GEOLOGICAL SURVEY
BRANCH OF GLOBAL SEISMOLOGY AND GEOMAGNETISM
DENVER FEDERAL CENTER
DENVER, COLORADO 80225 (R.B.)

DEPARTMENT OF PHYSICS
UNIVERSITY OF TORONTO
TORONTO, ONTARIO, CANADA M5S 1A7 (C.H.C.)

Manuscript received 11 January 1983

APPENDIX

Consider one segment of a tau branch which is sampled at the N ray parameters

$$p_0 < p_1 < \cdots < p_{N-1} = p_{end}. \tag{A1}$$

The segment is defined so that $X'(p)$ is singular only at p_{N-1}. As in equation (56), we define $4(N-1)$ τ spline coefficients such that

$$\tau(p) = a_i + b_i(p_{end} - p) + c_i(p_{end} - p)^2 + d_i(p_{end} - p)^{3/2}; \quad p_i \leqq p \leqq p_{i+1}. \tag{A2}$$

In defining these coefficients, we require that $\tau(p)$ be fit exactly at each sample point which places two constraints per interval or $2(N-1)$ constraints total. Also, we require that $X(p_0)$ and $X(p_{end})$ be fit exactly for two more constraints. The remaining $2(N-2)$ constraints arise from requiring continuity in the first and second derivatives of $\tau(p)$ at each internal sample point.

We could simply determine all coefficients by solving the $4(N-1) \times 4(N-1)$ system of equations. Even though the resulting matrix is banded, a more stable and compact algorithm can be designed by employing basis functions. Consider a basis function, $\psi_i(p)$, which is nonzero only on the interval (p_{i-2}, p_{i+2}) and is comprised of four (one per interval) piecewise tau splines. The 16 coefficients defining ψ_i can be determined by constraining ψ_i and its first and second derivatives to all be zero at p_{i-2} and p_{i+2} and to all be continuous at p_{i-1}, p_i and p_{i+1}. As the final constraint, we normalize $\psi_i(p_i)$ to unity. Neglecting, for the moment, difficulties in constructing basis functions in the vicinity of p_0 and p_{end}, we may rewrite equation (A2) as

$$\tau(p) = \sum_{i=1}^{N} \gamma_i \psi_i(p). \tag{A3}$$

The $N + 2$ γ's are determined by fitting the N tau values and two ranges as before. Continuity conditions are automatically satisfied due to the definition of the basis functions.

Because each $\psi_i(p)$ interacts only with its nearest neighbors, there are only $N + 2$ basis functions which can contribute to the interval $[p_0, p_{end}]$. For the same reason, the resulting $(N + 2) \times (N + 2)$ system of equations will be tridiagonal and thus trivial to solve. Finally, the $4(N - 1)$ coefficients required in equation (A2) could be determined as linear combinations of the $16(N + 2)$ coefficients defining the basis functions in equation (A3). An equivalent procedure which is more compact, is to use linear combinations of the basis functions to determine $X(p)$ at the $N - 2$ internal samples. The τ spline coefficients in equation (A3) may then be determined simply by fitting $\tau(p)$ and $X(p)$ at both ends of each interval, a purely local scheme which is identical to the construction of H splines. In effect, the basis functions are used to differentiate $\tau(p)$. An advantage of this latter scheme is that it is not necessary to save the $16(N + 2)$ basis function coefficients. Rather, three basis function contributions to $\tau(p)$ and three contributions to $X(p)$ are needed at each sample for a total of $6N$ values.

Determining ψ_{-1}, ψ_0, and ψ_1 require defining ray parameter samples $p_{-3} < p_{-2} < p_{-1} < p_0$. As the details are not critical, we have done so by requiring that $p_{-2} - p_{-3} = p_{-1} - p_{-2} = p_0 - p_{-1} = p_1 - p_0$. The same trick cannot be used at p_{end}. Because of the singularity in $X'(p)$ at p_{end}, special definitions are required for ψ_{N-3}, ψ_{N-2}, ψ_{N-1}, and ψ_N. We have found it convenient to truncate ψ_{N-2}, ψ_{N-1}, and ψ_N at p_{end}. For ψ_{N-3}, we have substituted the condition that the first derivative at p_{N-3} be zero rather than the second derivative at p_{end} be zero. For the truncated ψ_{N-2} there are 12 coefficients and 10 usable constraints. The difference is made up by constraining $\psi_{N-2}(p_{N-1})$ to be $\frac{1}{4}$ and forcing the first derivative to be zero at p_{N-2}. For ψ_{N-1}, there are 8 coefficients and 7 usable constraints. Again, we constrain the first derivative to be zero at p_{N-1}. Finally, for ψ_N, there are 4 coefficients and 3 usable constraints. In this case, we simply constrain $\psi_N(p_{N-1})$ to be $\frac{1}{4}$.