INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

H4.SMR/942-27

Third Workshop on
3D Modelling of Seismic Waves Generation
Propagation and their Inversion

4 - 15 November 1996

Ray Perturbation Theory for Traveltimes and Ray Paths in 3-D Heterogeneous Media

(1) R. Snieder, (2) Malcolm Sambridge

(1) Dept. of Theoretical Geophysics
University of Utrecht
Urecht, The Netherlands

(2) Institute of Theoretical Geophysics
University of Cambridge
Cambridge, UK

Geophys. J. Int. (1992) 109, 294–322

Ray perturbation theory for traveltimes and ray paths in 3-D heterogeneous media

Roel Snieder[1] and Malcolm Sambridge[2]
[1]Department of Theoretical Geophysics, University of Utrecht, PO Box 80.021, 3508 TA Utrecht, The Netherlands
[2]Institute of Theoretical Geophysics, University of Cambridge, Cambridge CB2 3EQ, UK

Accepted 1991 November 8. Received 1991 October 21; in original form 1991 July 26

SUMMARY

Both paraxial ray tracing and two-point ray tracing are powerful tools for solving wave propagation problems. When a slowness model is mildly perturbed from a reference model, one can use perturbation theory for the determination of the ray positions and the traveltimes. An extension of Fermat's theorem is presented, which states that the traveltime is stationary with respect to the perturbations in the ray position provided that the endpoints of the ray are perturbed along the wavefront of the unperturbed ray. It is shown that when the ray perturbation satisfies this condition the second-order traveltime perturbation can be computed from the first-order ray perturbation. A perturbation analysis of the equation of kinematic ray tracing leads to a simple second-order differential equation for the ray deflection expressed in ray coordinates. This constitutes a perturbation method based on a Lagrangian formulation, and leads to a first-order expression for the ray deflection and a second-order expression for the traveltime perturbation. This is of relevance to non-linear traveltime tomography because it leads to an efficient method for evaluating the lowest order ray deflection and the non-linear effect this has on the traveltimes. The theory is applicable both to two-point ray tracing and to the determination of paraxial rays. The derivations in this paper are completely self-contained. All expressions, including the transformation to ray coordinates, are derived from first principles. In this way one obtains insight in the approximations that are actually made. A scale analysis leads to dimensionless numbers that give an indication whether the theory is applicable to a specific problem. For the special case of a layered reference medium the final equations are particularly simple. Plane discontinuities in the reference model and the slowness perturbation are incorporated in the theory. The final expressions for the ray deflection and the traveltime perturbation can be implemented numerically in a simple way. It is indicated how applications to very large-scale problems can be achieved. Several examples, including the propagation of waves through a quasi-random model of the earth's mantle illustrate the theory.

Key words: Fermat's theorem, paraxial rays, perturbation theory, ray tracing, tomography.

1 INTRODUCTION

Ray theory plays an important role both in exploration seismics and large-scale seismology. Green's functions derived from ray theory form the basis of several algorithms for computing synthetic seismograms. Traveltime tomography is an important technique for the determination of seismic velocities on a variety of scales ranging from cross-borehole tomography to tomography of the earth's mantle. There is a vast amount of literature on ray tracing methods. Traditionally one uses either shooting methods (e.g. Červený 1987) or bending methods (e.g. Julian & Gubbins 1977; Peyrera, Lee & Keller 1980). Recently, new algorithms have been developed which either solve the Eikonal equation with a finite difference method (Vidale

1988) or search for the shortest path on a graph which approximates all possible ray paths emanating from a given source (Moser 1991). The bending method has recently been further developed by Moser, Nolet & Snieder (1992). In order to use the full potential of the steadily increasing data sets it is important to have an efficient method for tracing rays in heterogeneous media.

This is especially important in traveltime tomography. When one linearizes the inverse problem for this case one can use Fermat's theorem, which implies that one can use the rays of the reference medium and assume that the observed traveltimes are due to slowness perturbations along the unperturbed ray (e.g. Nolet 1987). In reality the inverse problem is non-linear, because the slowness perturbations perturb the rays themselves. The true rays are curves along which the traveltime is a minimum. This means that if one uses the rays in the reference medium for the inversion, rather than the true rays, that one overestimates the time needed for the propagation from the source to the receiver. In order to obtain a fit of the observed traveltimes one will obtain a velocity model that is too fast. This means that neglecting the non-linearity of the inverse problem leads to a bias in the reconstructed model.

The fact that a linearized approach has been relatively successful in traveltime tomography indicates that the effects of ray bending on the inversion may not be too dramatic. In fact, a radially stratified earth accounts for a large amount for the slowness variations within the earth. This suggests that perturbation theory could be used to account for the effect of the slowness perturbation on the ray positions and the traveltimes. It is shown by Snieder (1990, 1991) how one can set up an inversion method based on a perturbation expansion of the forward problem. The theory presented here could be used for this.

The use of perturbation methods in ray theory is not new. The equations of dynamic ray tracing can be cast in a Hamiltonian form. Perturbation methods for Hamiltonian systems are well developed, and a number of authors have used this to develop a perturbation theory for ray tracing problems (Chapman 1985; Wunsch 1987; Farra & Madariaga 1987; Farra, Virieux & Madariaga 1989; Virieux 1991). As an alternative to Hamiltonian perturbation theory, one can also apply perturbation theory directly to the equation of kinematic ray tracing (Moore 1991). In the jargon of classical mechanics this corresponds to a Lagrangian formulation of the perturbation theory. It is *a priori* not evident that the different perturbation methods lead to the same result. Furthermore, if one sets up a perturbation problem in two different coordinate systems that are related through a non-linear transformation, then perturbation theory in general leads to different physical results in the two formulations.

Another ambiguity in setting up a perturbation treatment is the choice of the independent parameter and the employed coordinate system. Wunsch (1987) uses the horizontal distance of a ray as independent parameter and solves for the depth of a ray. Farra & Madariaga (1987) use the distance along the unperturbed ray as independent parameter and employ ray coordinates. In later studies Farra *et al.* (1989) and Virieux (1991) use Cartesian coordinates. Moore (1991) uses the distance along the perturbed ray as independent parameter. This last choice leads to conceptual problems since the independent parameter depends on the ray perturbation itself, which is a serious inconsistency in the theory. The approach of Moore (1991) also has the disadvantage that in two-point ray tracing the receiver is not at a known position of the independent coordinate since the arc length of a perturbed ray joining a given source and receiver is not known *a priori*.

Treating ray perturbation theory with a Hamiltonian formalism has the advantage that the derivations are algebraically simple, and that one can change relatively easy from one coordinate system to another. However, the derivation of the perturbation equations from a Hamiltonian formalism gives little physical insight and it is often not trivial to derive the conditions of validity for Hamiltonian perturbation theory. In contrast to this, Lagrangian perturbation theory gives direct physical insight because the Lagrangian is the traveltime, and the Euler–Lagrange equation is the equation of kinematic ray tracing. The disadvantage of Lagrangian perturbation theory over Hamiltonian perturbation methods is that the algebra usually is more complicated, and that the algebra has to be redone whenever one changes to a new coordinate system.

The aim of this paper is to present a perturbation solution of the equation of kinematic ray tracing which is conceptually simple. In contrast to earlier perturbation methods, the theory leads to an explicit expression for the traveltime that is correct to second order in the slowness perturbation. The theory leads to equations that can be applied well to extremely large data sets, because the problem can be set up in such a way that the perturbations in the ray position and traveltimes can be reduced to vector operations. Two dimensionless numbers are derived that give an indication whether the method is applicable to a certain problem.

This paper has a tutorial flavour and is completely self-contained. Since the perturbation theory is applied directly to the equation of kinematic ray tracing, rather than to Hamilton's equations, one can understand how the perturbation in the slowness affects the ray positions and the traveltimes. In Section 2 general expressions are derived for the perturbation of the traveltime, and it is established under which conditions one can obtain the second-order traveltime perturbation from the first-order ray deflection. The equation of kinematic ray tracing is perturbed in Section 3. A transformation to ray coordinates, as shown in Section 4, simplifies this expression considerably. All expressions needed to make the transformation to ray coordinates are derived. The relation with the Frenet equations is shown in Appendix A. In Section 5, the second-order traveltime perturbation is related in three alternative ways to the ray perturbation derived in ray coordinates. Dimensionless numbers are derived in Section 6 which indicate whether the perturbation theory can be applied to a specific problem. In Section 7 the equations for the ray perturbation are simplified for the important case of a stratified reference medium. The

theory is illustrated with several examples. In Section 8 the theory is applied to a model with a linear velocity gradient for which the exact solution is known in closed form. The theory is applied in Section 9 to rays propagating through a model of the earth's mantle that is perturbed by quasi-random slowness fluctuations. The effect of slowness discontinuities is introduced in a heuristic way in Section 10 for the case of a two-layer model. In Section 11 it is shown in a more rigorous way how the theory can be adapted in order to handle plane discontinuities in the reference slowness and/or the slowness perturbation. Finally, the application to rays propagating through a mantle model where both the reference slowness and the slowness perturbation are discontinuous at the 670 discontinuity is shown in Section 12.

2 PERTURBATION THEORY FOR TRAVELTIMES

In geometrical optics, a ray is defined as a curve along which the traveltime is stationary. The Euler–Lagrange equation corresponding to this variational problem is the equation of kinematic ray tracing

$$\frac{d}{ds}\left(u\frac{d\mathbf{r}}{ds}\right) = \nabla u, \tag{1}$$

where the slowness is given by $u(\mathbf{r})$, s is the arc length along the ray and $\mathbf{r}$ is the position vector. Suppose that the slowness can be written as a reference slowness $u_0(\mathbf{r})$ and a perturbation $\varepsilon u_1(\mathbf{r})$

$$u(\mathbf{r}) = u_0(\mathbf{r}) + \varepsilon u_1(\mathbf{r}). \tag{2}$$

The reference slowness $u_0(\mathbf{r})$ may for instance denote the slowness in a flat layered earth model, and the perturbation $\varepsilon u_1(\mathbf{r})$ may denote lateral variations in the slowness. The parameter ε is used for bookkeeping purposes and facilitates a systematic perturbation approach of the ray tracing problem. It is assumed momentarily that $u_0(\mathbf{r})$ and $u_1(\mathbf{r})$ are continuous functions of the space coordinates, but there are no further restrictions.

The case where the slowness is perturbed is relevant for two-point ray tracing problems. In paraxial ray tracing one perturbs the initial conditions of the ray, and one may or may not perturb the slowness too, depending on the application. For paraxial ray tracing problems with a fixed slowness one may simply set $u_1 = 0$ in ensuing expressions.

Let a ray in the reference medium $u_0(\mathbf{r})$ be denoted by $\mathbf{r}_0(s_0)$, where s_0 is the arc length along this reference ray. The ray in the perturbed medium is a non-linear function of the slowness perturbation. Except in the vicinity of caustics, the ray perturbation can be expanded in a regular perturbation series

$$\mathbf{r}(s_0) = \mathbf{r}_0(s_0) + \varepsilon\mathbf{r}_1(s_0) + \varepsilon^2\mathbf{r}_2(s_0) + \cdots. \tag{3}$$

The ray perturbation is parametrized with the arc length along the reference ray. Perturbations in the ray position along the reference ray are irrelevant because these perturbations don't change the position of the ray. Therefore without loss of generality the ray perturbations can be restricted to the plane perpendicular to the direction of the reference ray

$$(\mathbf{r}_i \cdot \dot{\mathbf{r}}_0) = 0 \quad \text{for} \quad i \geq 1. \tag{4}$$

The derivative along the reference ray is denoted with a dot:

$$\dot{F} \equiv \frac{dF}{ds_0}. \tag{5}$$

In two-point ray tracing the perturbed ray joins the source and receiver of the reference ray; the ray perturbation therefore vanishes at the endpoints of the ray

$$\mathbf{r}_i(0) = \mathbf{r}_i(S_0) = 0 \quad \text{for} \quad i \geq 1 \quad \text{(two-point ray tracing)}, \tag{6}$$

where the source is at location $s_0 = 0$ and the unperturbed ray has total arc length S_0. Since s_0 measures the arc length along the reference ray $\mathbf{r}_0(s_0)$, the vector $\dot{\mathbf{r}}_0$ is of unit length:

$$|\dot{\mathbf{r}}_0| = \left|\frac{d\mathbf{r}_0}{ds_0}\right| = 1. \tag{7}$$

The deflection of the ray due to the slowness perturbation has the effect that the perturbed ray has a different arc length than the unperturbed ray. By inserting (3) in the relation

$$\left(\frac{\partial s}{\partial s_0}\right)^2 = \left(\frac{d\mathbf{r}}{ds_0} \cdot \frac{d\mathbf{r}}{ds_0}\right)$$

and by using (4) and (7) one finds after some algebra that to second order

$$\frac{\partial s}{\partial s_0} = 1 + \varepsilon(\dot{\mathbf{r}}_0 \cdot \dot{\mathbf{r}}_1) + \varepsilon^2[\tfrac{1}{2}(\dot{\mathbf{r}}_1 \cdot \dot{\mathbf{r}}_1) - \tfrac{1}{2}(\dot{\mathbf{r}}_0 \cdot \dot{\mathbf{r}}_1)^2 + (\dot{\mathbf{r}}_0 \cdot \dot{\mathbf{r}}_2)], \tag{8}$$

and its reciprocal expression

$$\frac{\partial s_0}{\partial s} = 1 - \varepsilon(\dot{\mathbf{r}}_0 \cdot \dot{\mathbf{r}}_1) - \varepsilon^2[\tfrac{1}{2}(\dot{\mathbf{r}}_1 \cdot \dot{\mathbf{r}}_1) - \tfrac{3}{2}(\dot{\mathbf{r}}_0 \cdot \dot{\mathbf{r}}_1)^2 + (\dot{\mathbf{r}}_0 \cdot \dot{\mathbf{r}}_2)]. \tag{9}$$

Note that Moore (1991) ignores the distinction between the unperturbed and the perturbed arc length. In that case the $O(\varepsilon)$ and $O(\varepsilon^2)$ terms in (8) and (9) are ignored and the resulting perturbation scheme is not consistent.

Now consider the traveltime T

$$T = \int u[\mathbf{r}(s)]\, ds. \tag{10}$$

The perturbation in the slowness and the ray position has two effects. First, the perturbed ray samples different regions than the unperturbed ray. By inserting (3) into (2) and performing a Taylor expansion of u_0 and u_1 about $\mathbf{r}_0$, one finds that to second order

$$u(\mathbf{r}) = u_0(\mathbf{r}_0) + \varepsilon[u_1(\mathbf{r}_0) + \mathbf{r}_1 \cdot \nabla u_0(\mathbf{r}_0)] + \varepsilon^2[\mathbf{r}_1 \cdot \nabla u_1(\mathbf{r}_0) + \tfrac{1}{2}\mathbf{r}_1\mathbf{r}_1 : \nabla\nabla u_0(\mathbf{r}_0) + \mathbf{r}_2 \cdot \nabla u_0(\mathbf{r}_0)]. \tag{11}$$

In this expression the symbol ':' denotes a double contraction. Second, the perturbation of the ray path leads to a perturbed arc length. Using $ds = \partial s/\partial s_0\, ds_0$ and (11) in (10) this leads with (8) to a perturbation expansion of the traveltime

$$T = T_0 + \varepsilon T_1 + \varepsilon^2 T_2 + O(\varepsilon^3), \tag{12}$$

where

$$T_0 = \int_0^{S_0} u_0(\mathbf{r}_0)\, ds_0, \tag{13a}$$

$$T_1 = \int_0^{S_0} [u_1(\mathbf{r}_0) + (\mathbf{r}_1 \cdot \nabla u_0) + (\dot{\mathbf{r}}_0 \cdot \dot{\mathbf{r}}_1)u_0]\, ds_0, \tag{13b}$$

$$T_2 = \int_0^{S_0} \left(\frac{u_0}{2}(\dot{\mathbf{r}}_1 \cdot \dot{\mathbf{r}}_1) - \frac{u_0}{2}(\dot{\mathbf{r}}_1 \cdot \dot{\mathbf{r}}_0)^2 + [u_1 + (\mathbf{r}_1 \cdot \nabla u_0)](\dot{\mathbf{r}}_0 \cdot \dot{\mathbf{r}}_1) + (\mathbf{r}_1 \cdot \nabla u_1) + \tfrac{1}{2}(\mathbf{r}_1\mathbf{r}_1 : \nabla\nabla u_0) + (\mathbf{r}_2 \cdot \nabla u_0) + (\dot{\mathbf{r}}_0 \cdot \dot{\mathbf{r}}_2)u_0\right) ds_0, \tag{13c}$$

where it is understood that the slowness and its derivatives are to be evaluated on the reference ray $\mathbf{r}_0(s_0)$.

The expressions for the first- and second-order traveltime perturbation can be simplified by using the fact that $\mathbf{r}_0(s_0)$ is a ray in the reference medium. Consider the following integral:

$$I = \int_0^{S_0} [(\boldsymbol{\xi} \cdot \nabla u_0) + (\dot{\mathbf{r}}_0 \cdot \dot{\boldsymbol{\xi}})u_0]\, ds_0, \tag{14}$$

where $\boldsymbol{\xi}(s_0)$ is an arbitrary vector function along the reference ray which at the endpoints of the reference ray is perpendicular to the reference ray

$$(\boldsymbol{\xi} \cdot \dot{\mathbf{r}}_0)(0) = (\boldsymbol{\xi} \cdot \dot{\mathbf{r}}_0)(S_0) = 0. \tag{15}$$

It follows from (1) that the reference ray satisfies

$$\nabla u_0 = u_0\ddot{\mathbf{r}}_0 + (\dot{\mathbf{r}}_0 \cdot \nabla u_0)\dot{\mathbf{r}}_0. \tag{16}$$

In deriving this it is used that

$$\frac{dF}{ds_0} = (\dot{\mathbf{r}}_0 \cdot \nabla F). \tag{17}$$

Using integration by parts and the condition (15) one may derive the general result

$$\int_0^{S_0} u_0(\boldsymbol{\xi} \cdot \ddot{\mathbf{r}}_0)\, ds_0 = -\int_0^{S_0} u_0(\dot{\boldsymbol{\xi}} \cdot \dot{\mathbf{r}}_0)\, ds_0 - \int_0^{S_0} (\dot{\mathbf{r}}_0 \cdot \nabla u_0)(\boldsymbol{\xi} \cdot \dot{\mathbf{r}}_0)\, ds_0. \tag{18}$$

By inserting (16) and (18) in (14) one finds

$$\int_0^{S_0} [(\boldsymbol{\xi} \cdot \nabla u_0) + (\dot{\mathbf{r}}_0 \cdot \dot{\boldsymbol{\xi}})u_0]\, ds_0 = 0. \tag{19}$$

Using this expression in (13b, c), the first- and second-order traveltime perturbation have the simplified form

$$T_1 = \int_0^{S_0} u_1(\mathbf{r}_0)\, ds_0, \tag{20a}$$

$$T_2 = \int_0^{s_0} \left(\frac{u_0}{2}(\dot{\mathbf{r}}_1 \cdot \dot{\mathbf{r}}_1) - \frac{u_0}{2}(\dot{\mathbf{r}}_1 \cdot \dot{\mathbf{r}}_0)^2 + [u_1 + (\mathbf{r}_1 \cdot \nabla u_0)](\dot{\mathbf{r}}_0 \cdot \dot{\mathbf{r}}_1) + (\mathbf{r}_1 \cdot \nabla u_1) + \tfrac{1}{2}(\mathbf{r}_1\mathbf{r}_1 : \nabla\nabla u_0) \right) ds_0. \tag{20b}$$

Equation (19) is nothing but Fermat's theorem. It has the consequence that the first-order traveltime perturbation is the integral of the slowness perturbation along the reference ray (see 20a). This expression forms the basis of linearized tomographic inversions. Another consequence of Fermat's theorem (19) is that the second-order traveltime perturbation does not depend on the second-order ray perturbation $\mathbf{r}_2$ (see 20b). This means that in order to obtain expressions of the traveltime perturbation which are correct to second order, it suffices to know the ray perturbation to first order. For this reason, only the first-order ray perturbation is derived in this paper.

Note that in the derivation it is not used that the ray perturbation vanishes at the endpoints of the ray, but that it is only required that (15) is satisfied. Since the wavefront is perpendicular to the ray, this expression states that Fermat's theorem is valid when the endpoint of the ray is perturbed along the wavefront of the reference ray. This means that (20b) is not only applicable to two-point ray tracing problems, but also to perturbed paraxial rays provided one works in a coordinate system where (15) is satisfied.

3 FIRST-ORDER PERTURBATION THEORY FOR THE RAY DEFLECTION

The first-order ray deflection can be derived by applying a perturbation analysis to the equation of kinematic ray tracing (1). With (17) this equation can be written as

$$u(\mathbf{r})\frac{d^2\mathbf{r}}{ds^2} + \left(\frac{d\mathbf{r}}{ds} \cdot \nabla u\right)\frac{d\mathbf{r}}{ds} = \nabla u. \tag{21}$$

In evaluating the derivatives in this expression one should take into account that the derivative along the perturbed ray and the unperturbed ray are different:

$$\frac{d}{ds} = \frac{\partial s_0}{\partial s}\frac{d}{ds_0}, \tag{22}$$

with $\partial s_0/\partial s$ given by (9). For the gradient of the slowness one finds with (2), (3) and a Taylor expansion that

$$\nabla u(\mathbf{r}) = \nabla u_0(\mathbf{r}_0) + \varepsilon[\nabla u_1(\mathbf{r}_0) + \mathbf{r}_1 \cdot \nabla\nabla u_0(\mathbf{r}_0)]. \tag{23}$$

Inserting (11), (22) and (23) in (21), and differentiating carefully, leads to a linear differential equation for $\mathbf{r}_1$:

$$u_0\ddot{\mathbf{r}}_1 - u_0(\ddot{\mathbf{r}}_1 \cdot \dot{\mathbf{r}}_0)\dot{\mathbf{r}}_0 - u_0(\ddot{\mathbf{r}}_0 \cdot \dot{\mathbf{r}}_1)\dot{\mathbf{r}}_0 - 2u_0(\dot{\mathbf{r}}_0 \cdot \dot{\mathbf{r}}_1)\ddot{\mathbf{r}}_0 + (\dot{\mathbf{r}}_1 \cdot \nabla u_0)\dot{\mathbf{r}}_0 - 2(\dot{\mathbf{r}}_0 \cdot \dot{\mathbf{r}}_1)(\dot{\mathbf{r}}_0 \cdot \nabla u_0)\dot{\mathbf{r}}_0 + \dot{\mathbf{r}}_0(\dot{\mathbf{r}}_0\mathbf{r}_1 : \nabla\nabla u_0) + \dot{\mathbf{r}}_1(\dot{\mathbf{r}}_0 \cdot \nabla u_0)$$
$$+ (\mathbf{r}_1 \cdot \nabla u_0)\ddot{\mathbf{r}}_0 - \mathbf{r}_1 \cdot \nabla\nabla u_0 = \nabla u_1 - (\dot{\mathbf{r}}_0 \cdot \nabla u_1)\dot{\mathbf{r}}_0 - u_1\ddot{\mathbf{r}}_0, \tag{24}$$

where it is understood that the slowness and its derivatives are to be evaluated on the reference ray.

Equation (24) contains a large number of terms, but it is possible to simplify this expression considerably. The second derivatives $\ddot{\mathbf{r}}_0$ can be eliminated by writing (16) in the form

$$\ddot{\mathbf{r}}_0 = \frac{1}{u_0}[\nabla u_0 - (\dot{\mathbf{r}}_0 \cdot \nabla u_0)\dot{\mathbf{r}}_0]. \tag{25}$$

Differentiating (4), using (25) to eliminate $\ddot{\mathbf{r}}_0$ and invoking the orthogonality of $\mathbf{r}_1$ and $\dot{\mathbf{r}}_0$ gives

$$(\dot{\mathbf{r}}_1 \cdot \dot{\mathbf{r}}_0) = -\frac{1}{u_0}(\mathbf{r}_1 \cdot \nabla u_0). \tag{26}$$

Inserting (25) and (26) in (24) leads after some algebra to

$$u_0\ddot{\mathbf{r}}_1 - u_0(\dot{\mathbf{r}}_0 \cdot \ddot{\mathbf{r}}_1)\dot{\mathbf{r}}_0 + \dot{\mathbf{r}}_1(\dot{\mathbf{r}}_0 \cdot \nabla u_0) + \frac{1}{u_0}(\mathbf{r}_1 \cdot \nabla u_0)[3\nabla u_0 - 2(\dot{\mathbf{r}}_0 \cdot \nabla u_0)\dot{\mathbf{r}}_0] - [(\mathbf{r}_1 \cdot \nabla\nabla u_0) - (\dot{\mathbf{r}}_0\mathbf{r}_1 : \nabla\nabla u_0)\dot{\mathbf{r}}_0]$$
$$= u_0\nabla\left(\frac{u_1}{u_0}\right) - u_0\left[\dot{\mathbf{r}}_0 \cdot \nabla\left(\frac{u_1}{u_0}\right)\right]\dot{\mathbf{r}}_0. \tag{27}$$

It follows that the slowness perturbation affects the first-order ray deflection only through the gradient of the relative slowness perturbation $\nabla(u_1/u_0)$. Indeed, one finds from (1) that multiplying the slowness with a constant $[u(\mathbf{r}) \rightarrow Cu(\mathbf{r})]$ leaves the rays unaffected. For such a perturbation $u_1(\mathbf{r}) = (C-1)u_0(\mathbf{r})$, hence $\nabla(u_1/u_0) = 0$ and the ray perturbation vanishes.

Because of the orthogonality condition (4), only the component of the ray deflection perpendicular to the reference ray is relevant. For an arbitrary vector $\boldsymbol{\xi}$ the component perpendicular to the reference ray may be denoted by

$$\boldsymbol{\xi}_\perp \equiv \boldsymbol{\xi} - (\dot{\mathbf{r}}_0 \cdot \boldsymbol{\xi})\dot{\mathbf{r}}_0, \tag{28}$$

where it is used that $\dot{\mathbf{r}}_0$ is of unit length (see 7). Using the definition (28) gives

$$3\nabla u_0 - 2(\dot{\mathbf{r}}_0 \cdot \nabla u_0)\dot{\mathbf{r}}_0 = 3\nabla_\perp u_0 + (\dot{\mathbf{r}}_0 \cdot \nabla u_0)\dot{\mathbf{r}}_0. \tag{29}$$

Since $\mathbf{r}_1$ is perpendicular to the reference ray one has

$$(\mathbf{r}_1 \cdot \xi) = (\mathbf{r}_1 \cdot \xi_\perp). \tag{30}$$

From the definition (28), and the relations (26) and (30) it follows that

$$\dot{\mathbf{r}}_1 = \dot{\mathbf{r}}_{1\perp} - \frac{1}{u_0}(\mathbf{r}_1 \cdot \nabla u_0)\dot{\mathbf{r}}_0. \tag{31}$$

Using (29) and (31) in (27) gives with the definition (28):

$$u_0\ddot{\mathbf{r}}_{1\perp} + (\dot{\mathbf{r}}_0 \cdot \nabla u_0)\dot{\mathbf{r}}_{1\perp} + \frac{3}{u_0}(\mathbf{r}_1 \cdot \nabla u_0)\nabla_\perp u_0 - (\mathbf{r}_1 \cdot \nabla\nabla u_0)_\perp = u_0\nabla_\perp\left(\frac{u_1}{u_0}\right). \tag{32}$$

Note from this expression that we only need the perpendicular components of $\dot{\mathbf{r}}_1$ and $\ddot{\mathbf{r}}_1$.

4 THE TRANSFORMATION TO RAY COORDINATES

Expression (32) is considerably simpler than the original perturbation equation (24). Nevertheless, it is not the most convenient form to use in numerical computations. First, the simplicity of equation (32) is deceptive, because the relation between the perpendicular components of $\mathbf{r}_1$ and its derivatives is in general not trivial. For example, in general one has $(d\mathbf{r}_1/ds_0)_\perp \neq d\mathbf{r}_{1\perp}/ds_0$ because the direction $\dot{\mathbf{r}}_0$ of the unperturbed ray varies with s_0. Second, the first-order ray deflection $\mathbf{r}_1$ is a 3-D vector. However, the constraint (4) reduces the number of independent components of $\mathbf{r}_1$ by one, so that (32) should be solved under the constraint (4). These complications are both handled when a transformation to ray coordinates is made. An additional advantage of using ray coordinates is that the condition (15) for the validity of Fermat's theorem is automatically satisfied.

Consider two mutually orthogonal unit vectors $\hat{\mathbf{q}}_1$ and $\hat{\mathbf{q}}_2$ that are orthogonal to the reference ray. This implies that

$$(\hat{\mathbf{q}}_1 \cdot \hat{\mathbf{q}}_2) = (\hat{\mathbf{q}}_1 \cdot \dot{\mathbf{r}}_0) = (\hat{\mathbf{q}}_2 \cdot \dot{\mathbf{r}}_0) = 0, \tag{33}$$

and with (7)

$$(\hat{\mathbf{q}}_1 \cdot \hat{\mathbf{q}}_1) = (\hat{\mathbf{q}}_2 \cdot \hat{\mathbf{q}}_2) = (\dot{\mathbf{r}}_0 \cdot \dot{\mathbf{r}}_0) = 1. \tag{34}$$

The direction $\dot{\mathbf{r}}_0$ of the reference ray changes along the reference ray, so that the unit vectors $\hat{\mathbf{q}}_i$ also change direction along the reference ray: $\hat{\mathbf{q}}_i = \hat{\mathbf{q}}_i(s_0)$. Since the ray perturbation is perpendicular to the reference ray, it can be expanded as

$$\mathbf{r}_1 = q_1\hat{\mathbf{q}}_1 + q_2\hat{\mathbf{q}}_2. \tag{35}$$

It is the aim of this section to convert the differential equation (32) to a differential equation for the components q_1 and q_2.

In order to do so one needs the derivatives of the unit vectors $\hat{\mathbf{q}}_i$. Since the basis $(\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2, \dot{\mathbf{r}}_0)$ is orthonormal, one can expand $\dot{\hat{\mathbf{q}}}_1$ as

$$\dot{\hat{\mathbf{q}}}_1 = (\dot{\hat{\mathbf{q}}}_1 \cdot \hat{\mathbf{q}}_1)\hat{\mathbf{q}}_1 + (\dot{\hat{\mathbf{q}}}_1 \cdot \hat{\mathbf{q}}_2)\hat{\mathbf{q}}_2 + (\dot{\hat{\mathbf{q}}}_1 \cdot \dot{\mathbf{r}}_0)\dot{\mathbf{r}}_0. \tag{36}$$

The vectors $\hat{\mathbf{q}}_i$ are normalized, it follows by differentiation of (34) that

$$(\dot{\hat{\mathbf{q}}}_1 \cdot \hat{\mathbf{q}}_1) = (\dot{\hat{\mathbf{q}}}_2 \cdot \hat{\mathbf{q}}_2) = 0. \tag{37}$$

The quantity $(\dot{\hat{\mathbf{q}}}_1 \cdot \hat{\mathbf{q}}_2)$ describes the rotation of the unit vectors $\hat{\mathbf{q}}_1$ and $\hat{\mathbf{q}}_2$ along the reference ray. This quantity is completely arbitrary, since one is free in choosing the direction of the unit vector $\hat{\mathbf{q}}_1$ in any direction within the plane perpendicular to the reference ray. This choice may vary along the reference ray. Denote the rotation rate of the unit vectors $\hat{\mathbf{q}}_1$ and $\hat{\mathbf{q}}_2$ along the reference ray by $\Omega(s_0)$:

$$(\dot{\hat{\mathbf{q}}}_1 \cdot \hat{\mathbf{q}}_2) \equiv \Omega(s_0). \tag{38}$$

In order to find the term $(\dot{\hat{\mathbf{q}}}_1 \cdot \dot{\mathbf{r}}_0)$ in (36), differentiate the expression $(\hat{\mathbf{q}}_i \cdot \dot{\mathbf{r}}_0) = 0$ with respect to s_0, and eliminate $\ddot{\mathbf{r}}_0$ with (25); this gives

$$(\dot{\hat{\mathbf{q}}}_i \cdot \dot{\mathbf{r}}_0) = -\frac{1}{u_0}(\hat{\mathbf{q}}_i \cdot \nabla u_0). \tag{39}$$

Note that $(\hat{\mathbf{q}}_i \cdot \nabla u_0)$ is nothing but the derivative of u_0 in the normal direction $\hat{\mathbf{q}}_i$. Inserting (37), (38) and (39) in (36) one finds

$$\dot{\hat{\mathbf{q}}}_1 = \Omega\hat{\mathbf{q}}_2 - \frac{1}{u_0}(\hat{\mathbf{q}}_1 \cdot \nabla u_0)\dot{\mathbf{r}}_0. \tag{40a}$$

A similar expression follows for $\dot{\hat{\mathbf{q}}}_2$. By differentiating the first term of (33) one finds that

$$(\hat{\mathbf{q}}_1 \cdot \dot{\hat{\mathbf{q}}}_2) = -(\dot{\hat{\mathbf{q}}}_1 \cdot \hat{\mathbf{q}}_2),$$

so that $\dot{\hat{\mathbf{q}}}_2$ is given by

$$\dot{\hat{\mathbf{q}}}_2 = -\Omega \hat{\mathbf{q}}_1 - \frac{1}{u_0}(\hat{\mathbf{q}}_2 \cdot \nabla u_0)\dot{\mathbf{r}}_0. \tag{40b}$$

It is convenient to use the summation convention for summing over the two unit vectors $\hat{\mathbf{q}}_1$ and $\hat{\mathbf{q}}_2$ and their related components. Let ε_{ij} denote the second-order Levi–Cevita tensor:

$$\varepsilon_{11} = \varepsilon_{22} = 0, \qquad \varepsilon_{12} = -\varepsilon_{21} = 1.$$

The expressions (40a) and (40b) can then be written as

$$\dot{\hat{\mathbf{q}}}_i = \Omega \varepsilon_{ij} \hat{\mathbf{q}}_j - \frac{1}{u_0}(\hat{\mathbf{q}}_i \cdot \nabla u_0)\dot{\mathbf{r}}_0. \tag{43}$$

It is shown in Appendix A how equation (43) is related to the Frenet equations. The second derivative of $\hat{\mathbf{q}}_i$ follows by differentiating (43) and eliminating $\ddot{\mathbf{r}}_0$ with (25) and $\dot{\hat{\mathbf{q}}}_i$ with (43); this gives

$$\ddot{\hat{\mathbf{q}}}_i = -\Omega^2 \hat{\mathbf{q}}_i + \dot{\Omega} \varepsilon_{ij} \hat{\mathbf{q}}_j - \frac{1}{u_0^2}(\hat{\mathbf{q}}_i \cdot \nabla u_0)(\hat{\mathbf{q}}_j \cdot \nabla u_0)\hat{\mathbf{q}}_j + \alpha_i \dot{\mathbf{r}}_0, \tag{44}$$

where α_i is used to represent the coefficient of the $\dot{\mathbf{r}}_0$ terms and where we have expanded the vector ∇u_0 using $\nabla u_0 = (\nabla u_0 \cdot \hat{\mathbf{q}}_1)\hat{\mathbf{q}}_1 + (\nabla u_0 \cdot \hat{\mathbf{q}}_2)\hat{\mathbf{q}}_2 + (\nabla u_0 \cdot \dot{\mathbf{r}}_0)\dot{\mathbf{r}}_0$.

In order to convert (32) to ray coordinates it is necessary to express the derivatives $\dot{\mathbf{r}}_1$ and $\ddot{\mathbf{r}}_1$ in the components q_i and their derivatives. It follows from the definition (28) that for any vector $\boldsymbol{\xi}$

$$\boldsymbol{\xi}_\perp = (\hat{\mathbf{q}}_i \cdot \boldsymbol{\xi})\hat{\mathbf{q}}_i.$$

Differentiate (35) with respect to s_0 and use (43) to give

$$\dot{\mathbf{r}}_1 = (\dot{q}_i - \Omega \varepsilon_{ij} q_j)\hat{\mathbf{q}}_i - \frac{1}{u_0} q_i (\hat{\mathbf{q}}_i \cdot \nabla u_0)\dot{\mathbf{r}}_0. \tag{46}$$

Differentiate this expression once more with respect to s_0; with (45) one finds that

$$\ddot{\mathbf{r}}_{1\perp} = (\ddot{q}_i - \Omega^2 q_i - 2\Omega \varepsilon_{ij}\dot{q}_j - \dot{\Omega}\varepsilon_{ij} q_j)\hat{\mathbf{q}}_i - \frac{1}{u_0^2}(\hat{\mathbf{q}}_i \cdot \nabla u_0)(\hat{\mathbf{q}}_j \cdot \nabla u_0) q_j \hat{\mathbf{q}}_i. \tag{47}$$

Insert (46) and (47) in (32), and take the inner product of the resulting expression with the unit vectors $\hat{\mathbf{q}}_1$ and $\hat{\mathbf{q}}_2$. Using the identities

$$\nabla\nabla u_0 - \frac{2}{u_0}(\nabla u_0)(\nabla u_0) = -u_0^2 \nabla\nabla \frac{1}{u_0}$$

and

$$u_0 \ddot{q}_i + (\dot{\mathbf{r}}_0 \cdot \nabla u_0)\dot{q}_i = \frac{d}{ds_0}(u_0 \dot{q}_i) \tag{49}$$

one finds after some algebra that the components q_i satisfy

$$\frac{d}{ds_0}(u_0 \dot{q}_i) - u_0 \Omega^2 q_i + u_0^2 \left(\hat{\mathbf{q}}_i \hat{\mathbf{q}}_j : \nabla\nabla \frac{1}{u_0}\right) q_j - 2u_0 \Omega \varepsilon_{ij} \dot{q}_j - \frac{d}{ds_0}(u_0 \Omega)\varepsilon_{ij} q_j = u_0 \hat{\mathbf{q}}_i \cdot \nabla\left(\frac{u_1}{u_0}\right). \tag{50}$$

This equation needs to be supplemented with two boundary conditions. For two-point ray tracing these are given by

$$q_i(0) = q_i(S_0) = 0 \quad \text{(two-point ray tracing)};$$

for initial value ray tracing one prescribes $q_i(0)$ and $\dot{q}_i(0)$. (51)

The system (50) for the ray perturbation constitutes a pair of coupled linear second-order differential equations and may be solved efficiently using a variety of standard numerical techniques (see Press *et al.* 1986). After discretization, (50) with the appropriate boundary conditions can be reduced to a linear system of equations, where only the main diagonal and four other bands of the involved matrix have non-zero elements. For the important application of layered reference media (see Section 7)

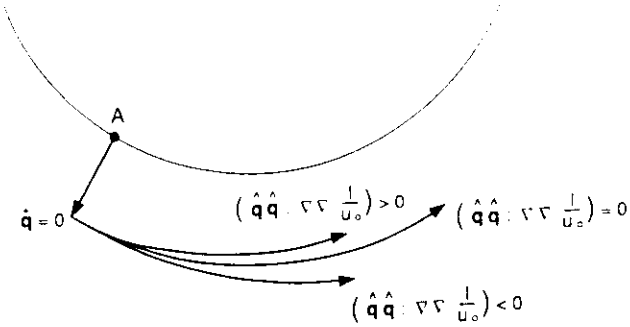

Figure 1. Paraxial rays that are parallel to the reference ray at point A for several values of the transverse curvature of the velocity $\hat{\mathbf{q}}\hat{\mathbf{q}} : \nabla\nabla(1/u_0)$.

the equations for q_1 and q_2 decouple, and the resulting matrix system is tridiagonal. Such a system can be solved extremely efficiently (e.g. Press *et al.* 1986).

Alternatively, one can solve (50) using a Green's function $G_{ij}(s_0, s)$ which satisfies

$$\frac{d}{ds_0}(u_0\dot{G}_{ij}) - u_0\Omega^2 G_{ij} + u_0^2\left(\hat{\mathbf{q}}_i\hat{\mathbf{q}}_k : \nabla\nabla\frac{1}{u_0}\right)G_{kj} - 2u_0\Omega\varepsilon_{ik}\dot{G}_{kj} - \frac{d}{ds_0}(u_0\Omega)\varepsilon_{ik}G_{kj} = \delta_{ij}\delta(s_0 - s), \tag{52}$$

with appropriate boundary conditions. For example, the solution of (50) subject to the boundary condition (51) for two-point ray tracing is given by

$$q_i(s_0) = \int_0^{s_0} G_{ij}(s_0, s)\left[u_0\hat{\mathbf{q}}_j \cdot \nabla\left(\frac{u_1}{u_0}\right)\right](s)\, ds \quad \text{(two-point ray tracing)}. \tag{53}$$

This representation of the solution may be useful when one wants to compute the ray deflection for a very large number of rays. In that case one can compute the Green's function $G_{ij}(s_0, s)$ once and store this function in tabulated form. The numerical computation of the solution (53) can then be reduced to efficient array processing operations. This may be useful in very large-scale tomographic inversions where one wants to take ray bending effects into account.

Equation (50) is a second-order differential equation for the ray deflection which bears a remarkable resemblance to the equation of motion of a particle in classical mechanics. When one considers u_0 as the mass of a particle, then $u_0\dot{q}_i$ denotes the momentum. When s_0 is associated with time, then the first term in (50) is the time derivative of the momentum, which is present in Newton's law. The terms containing Ω in (50) can be associated respectively with the centrifugal force and the Coriolis force which operate because of the rotation of the coordinate system $(\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2)$ around the reference ray. The term with $d(u_0\Omega)/ds_0$ is due to the variation in the rate of rotation along the reference ray. The right-hand side of (50) is equivalent to the external forces in Newton's law. This term is proportional to the gradient of the relative slowness perturbation. The term $u_0^2(\hat{\mathbf{q}}_i\hat{\mathbf{q}}_j : \nabla\nabla 1/u_0)q_j$ arises because of the change in direction of the reference ray. It acts in (50) as a linear restoring force that forces the ray deflection towards the reference ray or repels the perturbed ray from the reference ray depending on the term $(\hat{\mathbf{q}}_i\hat{\mathbf{q}}_j : \nabla\nabla 1/u_0)$. This effect can be understood as follows. Consider momentarily a 2-D medium in which in a certain region the reference velocity increases linearly with depth $(\hat{\mathbf{q}}\hat{\mathbf{q}} : \nabla\nabla 1/u_0) = 0$ and where the slowness perturbation vanishes. Consider a perturbed ray that at a certain position A has a finite displacement from the reference ray and that runs parallel to the reference ray at that point (see Fig. 1). In that case the reference ray and the perturbed ray are concentric circles, and the ray displacement q is constant. One sees in (50) that for this reference medium the linear restoring force vanishes because $\nabla\nabla 1/u_0 = 0$. Now consider the same situation, but let the velocity increase faster than linear with depth: $(\hat{\mathbf{q}}\hat{\mathbf{q}} : \nabla\nabla 1/u_0) > 0$. In that case the perturbed ray is curved more strongly than the reference ray (see Fig. 1), which means that the perturbed ray is attracted towards the reference ray. For a medium where the velocity increases slower than linear with depth $(\hat{\mathbf{q}}\hat{\mathbf{q}} : \nabla\nabla 1/u_0) < 0$ the reference ray has a smaller curvature than in the case of a constant velocity gradient, and the perturbed ray is effectively repelled from the reference ray. This behaviour is described by the linear restoring force in equation (50).

5 THE SECOND-ORDER TRAVELTIME PERTURBATION

Once the ray deflection is known, one can compute the change in the traveltime due to the ray deflection by integrating the slowness along the perturbed ray. However, it is possible to simplify expression (20b) for the second-order traveltime perturbation. This leads to compact expressions for the bias in the traveltime that can be evaluated efficiently.

In order to simplify the second-order traveltime perturbation, insert (35) and (46) in (20b). Using

$$\varepsilon_{ij}\varepsilon_{ik} = \delta_{jk}, \tag{54}$$

with (48) one obtains an expression for T_2 in terms of the deflection q_i in ray coordinates:

$$T_2 = \int_0^{S_0} \left[\frac{u_0}{2} \dot{q}_i \dot{q}_i - u_0 \Omega \varepsilon_{ij} \dot{q}_i q_j + \frac{u_0}{2} \Omega^2 q_i q_i - \frac{u_0^2}{2} q_i q_j \left(\hat{\mathbf{q}}_i \hat{\mathbf{q}}_j : \nabla\nabla \frac{1}{u_0} \right) + u_0 \hat{\mathbf{q}}_i \cdot \nabla \left(\frac{u_1}{u_0} \right) q_i \right] ds_0. \tag{55}$$

This expression can be simplified further by using equation (50) for the ray deflection. Multiply (50) with q_i, sum over i and integrate the result along the reference ray; this gives

$$\int_0^{S_0} \left(q_i \frac{d}{ds_0} (u_0 \dot{q}_i) - u_0 \Omega^2 q_i q_i + u_0^2 \left(\hat{\mathbf{q}}_i \hat{\mathbf{q}}_j : \nabla\nabla \frac{1}{u_0} \right) q_i q_j - 2 u_0 \Omega \varepsilon_{ij} q_i \dot{q}_j - \frac{d}{ds_0} (u_0 \Omega) \varepsilon_{ij} q_i q_j \right) ds_0 = \int_0^{S_0} u_0 q_i \hat{\mathbf{q}}_i \cdot \nabla \left(\frac{u_1}{u_0} \right) ds_0. \tag{56}$$

Now use that

$$\varepsilon_{ij} q_i q_j = \varepsilon_{12} q_1 q_2 + \varepsilon_{21} q_2 q_1 = q_1 q_2 - q_2 q_1 = 0. \tag{57}$$

By changing dummy indices one finds

$$\varepsilon_{ij} q_i \dot{q}_j = \varepsilon_{ji} q_j \dot{q}_i = -\varepsilon_{ij} \dot{q}_i q_j. \tag{58}$$

Furthermore, using an integration by parts it follows that

$$\int_0^{S_0} q_i \frac{d}{ds_0} (u_0 \dot{q}_i) \, ds_0 = -\int_0^{S_0} u_0 \dot{q}_i \dot{q}_i \, ds_0 + [u_0 q_i \dot{q}_i]_0^{S_0}. \tag{59}$$

Inserting the results (57–59) in (56) gives

$$\int_0^{S_0} \left[\frac{u_0}{2} \dot{q}_i \dot{q}_i - u_0 \Omega \varepsilon_{ij} \dot{q}_i q_j + \frac{u_0}{2} \Omega^2 q_i q_i - \frac{u_0^2}{2} \left(\hat{\mathbf{q}}_i \hat{\mathbf{q}}_j : \nabla\nabla \frac{1}{u_0} \right) q_i q_j \right] ds_0 = \frac{-1}{2} \int_0^{S_0} u_0 q_i \hat{\mathbf{q}}_i \cdot \nabla \left(\frac{u_1}{u_0} \right) ds_0 + [\tfrac{1}{2} u_0 q_i \dot{q}_i]_0^{S_0}. \tag{60}$$

Using this in (55) leads to a compact expression for the second-order traveltime perturbation

$$T_2 = \frac{1}{2} \int_0^{S_0} u_0 q_i \hat{\mathbf{q}}_i \cdot \nabla \left(\frac{u_1}{u_0} \right) ds_0 + [\tfrac{1}{2} u_0 q_i \dot{q}_i]_0^{S_0}. \tag{61}$$

Because of (51) the boundary term $u_0 q_i \dot{q}_i$ vanishes for two-point ray tracing problems. Once the ray deflection q_i is computed one can determine the bias in the traveltime with the simple integral (61) along the reference ray.

Expression (61) contains both the ray deflection and the slowness perturbation. For some applications it may be advantageous to have an expression for the second-order traveltime perturbation which contains either the ray deflection or the slowness perturbation. The ray deflection q_i can be eliminated in a simple way from (61) using the Green's function. For example, for two-point ray tracing one finds from (53) that

$$T_2 = \frac{1}{2} \int_0^{S_0} \int_0^{S_0} G_{ij}(s, s') \left(u_0 \hat{\mathbf{q}}_i \cdot \nabla \frac{u_1}{u_0} \right)(s) \left(u_0 \hat{\mathbf{q}}_j \cdot \nabla \frac{u_1}{u_0} \right)(s') \, ds \, ds' \quad \text{(two-point ray tracing)}, \tag{62}$$

where the Greens function $G_{ij}(s, s')$ satisfies (52) with the boundary conditions (51). This expression is useful for analytical computations where one specifies the slowness perturbation explicitly. If one has computed and stored the Green's function in tabulated form, one can obtain the second-order traveltime perturbation from (62) by performing a double integral along the reference ray. This computation can be reduced to efficient array-processing operations. In this way it may be possible to incorporate second-order traveltime effects in very large-scale tomographic inversions.

For completeness an expression is derived for T_2 which contains the ray deflection, but which does not contain the slowness perturbation explicitly. This is achieved by eliminating the term $\int u_0 \hat{\mathbf{q}}_i \nabla (u_1/u_0) q_i \, ds_0$ from (55) using (60); this gives

$$T_2 = \int_0^{S_0} \left[-\frac{u_0}{2} \dot{q}_i \dot{q}_i + u_0 \Omega \varepsilon_{ij} \dot{q}_i q_j - \frac{u_0}{2} \Omega^2 q_i q_i + \frac{u_0^2}{2} \left(\hat{\mathbf{q}}_i \hat{\mathbf{q}}_j : \nabla\nabla \frac{1}{u_0} \right) q_i q_j \right] ds_0 + [u_0 q_i \dot{q}_i]_0^{S_0}. \tag{63}$$

As mentioned in Section 4, the differential equation for the ray deflection can be closely associated with the motion of a particle in classical mechanics. This analogy can also be used for the second-order traveltime. As noted by Wunsch (1987) and Farra and Madariaga (1987), one can consider the traveltime in ray tracing problems as the Lagrangian of an equivalent problem in classical mechanics. (The basic connection is of course the variational principle of the action respectively the traveltime.) The Lagrangian is the difference of the kinetic and the potential energy. One can see that expression (63) is indeed of this form. The term $\frac{1}{2} u_0 \dot{q}_i \dot{q}_i$ denotes the change in the deflection formulated in ray coordinates (the velocity). The terms containing Ω describe the change in ray deflection due to the rotation of the coordinate system along the reference ray. These terms together describe the absolute rate of change in the ray deflection, and are equivalent to the kinetic energy in classical mechanics. The last term within the integral in (63) can be associated with the potential energy due to the linear restoring force term in (50).

6 CONDITIONS FOR THE APPLICABILITY OF FIRST-ORDER PERTURBATION THEORY

When the slowness perturbation is large, the first-order perturbation method of this paper cannot be expected to give useful results. In this section estimates for the domain of the applicability of the first-order ray perturbation theory are derived. First, the change in the length of the ray induced by the perturbation may not be too severe, because the relations (8) and (9) are only useful when $|\partial s/\partial s_0 - 1| \ll 1$. This condition is satisfied when

$$|\dot{\mathbf{r}}_1| \ll 1. \tag{64}$$

Furthermore a Taylor expansion of u_1 is used (up to order $\mathbf{r}_1 \cdot \nabla u_1$), and of u_0 (up to order $\mathbf{r}_1\mathbf{r}_1 : \nabla\nabla u_0$). Let u_0 vary on a length-scale L_0, and let u_1 vary on a length-scale L. The truncated Taylor expansions represent the true slowness variations with an acceptable accuracy only when

$$|\mathbf{r}_1| \ll L_0 \quad \text{and} \quad |\mathbf{r}_1| \ll L. \tag{65}$$

One can check the conditions (64) and (65) after one has computed the perturbed ray, and one can thus verify *a posteriori* whether it was justified to use first-order perturbation theory.

This is not very satisfactory from the practical point of view, where one would like to know *a priori* whether the first-order perturbation theory is sufficiently accurate. For this one needs to relate the slowness perturbation to the total ray deflection. As a simple model, consider a homogenous reference slowness u_0, and let the slowness perturbation have a constant derivative perpendicular to the reference ray: $\hat{\mathbf{q}} \cdot \nabla u_1 = u_1/L$. The solution of (50) for the ray deflection in the direction of the gradient of the slowness perturbation is given by

$$q(s_0) = -\frac{u_1}{2u_0 L} s_0(S_0 - s_0). \tag{66}$$

The condition (64) implies that for the first-order theory to be valid one must have $|\partial q/\partial s_0| \ll 1$. The angle between the perturbed ray and the reference ray is largest for $s_0 = 0$. Using this, the condition (64) implies that

$$C_1 \equiv \frac{1}{2}\frac{u_1}{u_0}\frac{S_0}{L} \ll 1. \tag{67}$$

The ray deflection of the solution (66) is largest when $s_0 = S_0/2$, so that the condition $|\mathbf{r}_1| \ll L$ is satisfied when

$$C_2 \equiv \frac{1}{8}\frac{u_1}{u_0}\left(\frac{S_0}{L}\right)^2 \ll 1. \tag{68}$$

The simple model discussed here cannot be used to investigate the criterion $|\mathbf{r}_1| \ll L_0$, because u_0 is constant. However, for most practical problems the slowness perturbation varies on a shorter length-scale than the reference slowness, so that the condition $|\mathbf{r}_1| \ll L$ usually ensures that $|\mathbf{r}_1| \ll L_0$.

Note that the criteria (67) and (68) depend on two dimensionless numbers; a factor u_1/u_0 (which is usually small), and a factor S_0/L (which may be large). The first-order perturbation theory is only valid when the weakness of the relative slowness perturbation (u_1/u_0) outweighs the factor S_0/L. The criteria (67) and (68) depend critically on the length S_0 of the unperturbed ray. The reason for this is that for a given slowness perturbation, a longer ray is deflected more than a short ray, so that the condition that the deflection is small is more stringent for long rays than for short rays.

7 RAY PERTURBATIONS IN A LAYERED REFERENCE MEDIUM

For the important case of a layered reference medium the differential equation (50) simplifies considerably. Consider first the case where the reference slowness is a function of the depth coordinate z only: $u_0 = u_0(z)$. In that case the reference ray defines a vertical plane (see Fig. 2). Let the unit vector vector $\hat{\mathbf{q}}_2$ be perpendicular to this plane, and let $\hat{\mathbf{q}}_1$ be the unit vector in the

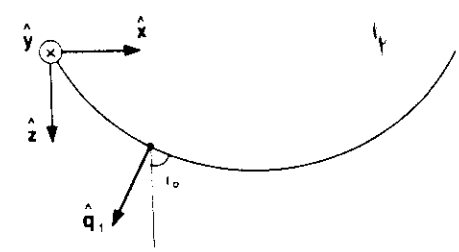

Figure 2. Definition of the geometric variables for a layered reference medium. The unit vector $\hat{\mathbf{y}}$ points out off the diagram.

plane of the reference ray (Fig. 2). In that case, the unit vectors $\hat{\mathbf{q}}_1$ and $\hat{\mathbf{q}}_2$ do not rotate around the reference ray, hence

$$\Omega = 0. \tag{69}$$

Define a Cartesian coordinate system with the z-direction pointing downwards, and with the y-direction perpendicular to the vertical plane of the reference ray. In terms of the angle of incidence i_0 of the reference ray, the unit vectors for defining the ray coordinates are given by

$$\hat{\mathbf{r}}_0 = \sin i_0 \hat{\mathbf{x}} + \cos i_0 \hat{\mathbf{z}}, \tag{70a}$$

$$\hat{\mathbf{q}}_1 = -\cos i_0 \hat{\mathbf{x}} + \sin i_0 \hat{\mathbf{z}}, \tag{70b}$$

$$\hat{\mathbf{q}}_2 = \hat{\mathbf{y}}. \tag{70c}$$

Since the reference slowness depends only on depth, one has

$$\nabla\nabla \frac{1}{u_0} = \hat{\mathbf{z}}\hat{\mathbf{z}}\, \partial_{zz} \frac{1}{u_0}, \tag{71}$$

and

$$\nabla\left(\frac{u_1}{u_0}\right) = \hat{\mathbf{x}}\, \partial_x\left(\frac{u_1}{u_0}\right) + \hat{\mathbf{y}}\, \partial_y\left(\frac{u_1}{u_0}\right) + \hat{\mathbf{z}}\, \partial_z\left(\frac{u_1}{u_0}\right). \tag{72}$$

The relations (69–72) can readily be inserted in equation (50) for the ray deflection; this gives

$$\frac{d}{ds_0}(u_0 \dot{q}_1) + u_0^2 \sin^2 i_0 \left(\partial_{zz} \frac{1}{u_0}\right) q_1 = -\cos i_0 u_0\, \partial_x\left(\frac{u_1}{u_0}\right) + \sin i_0 u_0\, \partial_z\left(\frac{u_1}{u_0}\right), \tag{73a}$$

$$\frac{d}{ds_0}(u_0 \dot{q}_2) = \partial_y u_1. \tag{73b}$$

Since the coordinate system does not rotate around the reference ray ($\Omega = 0$), and since the cross terms $(\hat{\mathbf{q}}_i \hat{\mathbf{q}}_j : \nabla\nabla(1/u_0))$ in (50) vanish (because u_0 does not depend on the y-coordinate), the equations for q_1 are q_2 decoupled. The trigonometric terms in (73a,b) can be eliminated by introducing the ray parameter p of the reference ray defined by

$$p \equiv u_0 \sin i_0. \tag{74}$$

This parameter is constant along the reference ray (Aki & Richards 1980). Using this parameter, the equations (73a,b) can be written as

$$\frac{d}{ds_0}(u_0 \dot{q}_1) + p^2 \left(\partial_{zz} \frac{1}{u_0}\right) q_1 = -\sqrt{u_0^2 - p^2}\, \partial_x\left(\frac{u_1}{u_0}\right) + p\, \partial_z\left(\frac{u_1}{u_0}\right), \tag{75a}$$

$$\frac{d}{ds_0}(u_0 \dot{q}_2) = \partial_y u_1. \tag{75b}$$

The equation for the transverse ray deflection q_2 can be integrated in closed form. For example, for two-point ray tracing one finds with the boundary conditions (51) that

$$q_2(s_0) = \frac{1}{h(S_0)} [F(s_0) h(S_0) - F(S_0) h(s_0)] \quad \text{(two-point ray tracing)}, \tag{76}$$

where

$$h(s_0) \equiv \int_0^{s_0} \frac{ds}{u_0(s)}, \tag{77}$$

and

$$F(s_0) \equiv \int_0^{s_0} ds \frac{1}{u_0(s)} \int_0^s ds'\, \partial_y u_1(s'). \tag{78}$$

Equation (75a) with the boundary conditions for q_1 reduces after discretization to a tridiagonal system of linear equations that can be solved efficiently. Alternatively, one can use a Green's function solution analogous to (53) for the solution of (75a). Note that in the special case where the reference velocity varies linearly with depth $[\partial_{zz}(1/u_0) = 0]$, one can integrate (75a)

analytically, and the solution for q_1 is of the form (76); the only difference is that the term $\partial_x u_1$ in (78) should be replaced by

$$-\sqrt{u_0^2 - p^2}\,\partial_x\left(\frac{u_1}{u_0}\right) + p\,\partial_z\left(\frac{u_1}{u_0}\right).$$

For a reference model with a spherical geometry, a simple form of the differential equation for the ray deflection is obtained by choosing a system of spherical coordinates with the North Pole ($\theta = 0$) at the source. Let $(\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}})$ denote the unit vectors in the vertical, θ and ϕ directions respectively. As in the case of Cartesian coordinates, choose $\hat{q}_2$ perpendicular to the vertical plane of the reference ray. This means that the unit vectors do not rotate around the reference ray, condition (69). When i_0 denotes the angle between the reference ray and the vertical, one has

$$\hat{\mathbf{q}}_1 = -\cos i_0 \hat{\boldsymbol{\theta}} - \sin i_0 \hat{\mathbf{r}}, \tag{79a}$$

$$\hat{\mathbf{q}}_2 = \hat{\boldsymbol{\phi}}. \tag{79b}$$

The spherical symmetry of the reference velocity implies that

$$\nabla\nabla\frac{1}{u_0} = \hat{\mathbf{r}}\hat{\mathbf{r}}\,\partial_{rr}\frac{1}{u_0}. \tag{80}$$

Furthermore, one has

$$\nabla\left(\frac{u_1}{u_0}\right) = \hat{\mathbf{r}}\,\partial_r\left(\frac{u_1}{u_0}\right) + \frac{\hat{\boldsymbol{\theta}}}{r}\partial_\theta\left(\frac{u_1}{u_0}\right) + \frac{\hat{\boldsymbol{\phi}}}{r\sin i_0}\partial_\phi\left(\frac{u_1}{u_0}\right). \tag{81}$$

For a spherically symmetric slowness model the ray parameter defined by

$$p \equiv r u_0 \sin i_0, \tag{82}$$

is constant along the reference ray (Aki & Richards 1980). Inserting the expressions (79–81) in (50) and using (82) to eliminate i_0 gives the following differential equations for the ray deflection

$$\frac{d}{ds_0}(u_0\dot{q}_1) + \frac{p^2}{r}\left(\partial_{rr}\frac{1}{u_0}\right)q_1 = -\frac{1}{r}\sqrt{u_0^2 - (p^2/r^2)}\,\partial_\theta\left(\frac{u_1}{u_0}\right) - \frac{p}{r}\partial_r\left(\frac{u_1}{u_0}\right), \tag{83a}$$

$$\frac{d}{ds_0}(u_0\dot{q}_2) = \frac{1}{r\sin\theta}\partial_\phi u_1. \tag{83b}$$

For two-point ray tracing the solution of (83b) is given by (76), where one should replace the derivative ∂_y by $(r\sin\theta)^{-1}\partial_\phi$.

8 EXAMPLE 1: THE LINEAR VELOCITY GRADIENT

As an example of the ray perturbation theory, consider a homogeneous reference medium with velocity c_0, where the velocity is perturbed with a linear gradient

$$c(z) = c_0\left(1 + \frac{z}{L}\right). \tag{84}$$

The slowness perturbation $(1/c - 1/c_0)$ is given by

$$\varepsilon u_1 = -u_0\frac{z/L}{1 + z/L}. \tag{85}$$

In this section the two-point ray tracing problem is considered. Since the relative slowness perturbation has a vanishing gradient in the horizontal direction, one needs only to consider the ray deflection q_1 of Fig. 2. Now consider the situation where the endpoints of the ray are located at $z = 0$, and the reference ray is horizontal. In this particular case equation (73a) reduces to

$$\ddot{q}_1 = -\frac{1}{L}, \tag{86}$$

where u_0 is the slowness of the reference medium, $u_0 = 1/c_0$. Equation (86), with the boundary condition (51), has the solution

$$q_1(s_0) = \frac{1}{2L}s_0(S_0 - s_0). \tag{87}$$

The distance along the reference ray is just the Cartesian coordinate x, while the $\hat{\mathbf{q}}_1$ direction is the vertical. The ray deflection

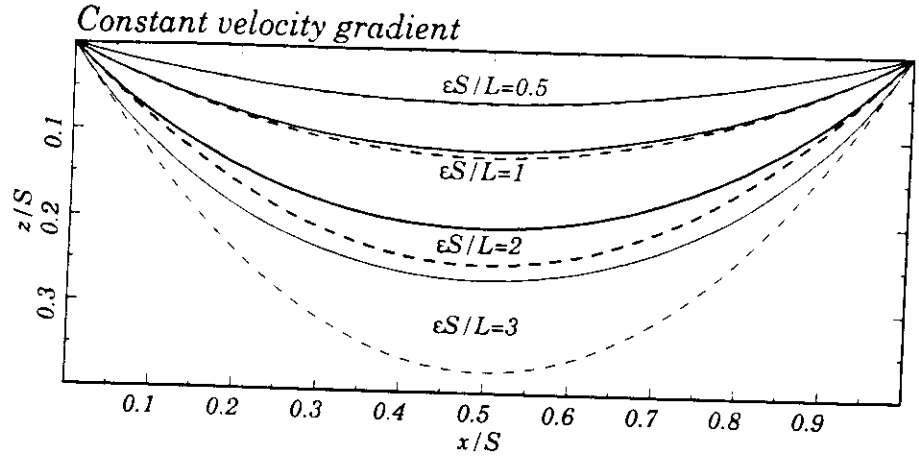

Figure 3. True rays (solid lines) and perturbed rays (dashed lines) for a homogeneous reference model and a slowness perturbation corresponding to a linear increase of the velocity with depth, for several values of the perturbation $\varepsilon S/L$.

then denotes the depth of the perturbed ray; according to (87) the perturbed ray is given by

$$z_{\text{pert}}(x) = \frac{\varepsilon}{2L} x(S_0 - x), \tag{88}$$

where it is used that the depth of the perturbed ray is $z_{\text{pert}} = \varepsilon q_1$. For this example, the true ray is the arc of a circle (e.g. Červený 1987), whose equation is given by

$$\left(x - \frac{S_0}{2}\right)^2 + \left(z_{\text{exact}} + \frac{L}{\varepsilon}\right)^2 = \left(\frac{S_0}{2}\right)^2 + \left(\frac{L}{\varepsilon}\right)^2, \tag{89}$$

or

$$z_{\text{exact}} = \frac{L}{\varepsilon}\left\{-1 + \left[1 + \left(\frac{\varepsilon}{L}\right)^2 x(S_0 - x)\right]^{1/2}\right\}. \tag{90}$$

Expanding this result in ε gives

$$z_{\text{exact}} = \frac{\varepsilon}{2L} x(S_0 - x) - \frac{\varepsilon^3}{8L^3} x^2 (S_0 - x)^2 + O(\varepsilon^5). \tag{91}$$

The first-order term in (91) agrees with the ray deflection obtained from the ray perturbation theory. Note that in the perturbed solution (88) the circle segment of the true ray is approximated with a parabola.

The ray deflection depends critically on the non-dimensional parameter $\varepsilon S_0/L$. The true rays and the perturbed rays are shown for several values of this parameter in Fig. 3. In agreement with the criterion (67) one sees that the ray perturbation deviates significantly from the true rays when the parameter $\varepsilon S_0/L > 2$. For the case $\varepsilon S_0/L = 1$ the error in the ray deflection is only 5 per cent, despite the fact that the angle between the reference ray ($z = 0$) and the perturbed ray is about 30° at the source.

The true traveltime, the traveltime from the second-order perturbation expression (61), and the traveltime from numerical integration along the perturbed ray are shown in Fig. 4 as a function of $\varepsilon S_0/L$. The traveltime by integrating numerically along the perturbed ray is always larger than the true traveltime; this reflects the fact that the true ray is a curve of minimal traveltime. A relative error in the traveltime of 0.1 per cent is reached for the traveltime from perturbation theory for $\varepsilon S_0/L \approx 0.7$, while for the traveltime computed from the numerical integration along the reference ray this error level is reached for $\varepsilon S_0/L \approx 1.5$.

The traveltime from the numerical integration along the perturbed ray is much more accurate than the traveltime computed from second-order perturbation theory. The reason for this is that in the perturbation calculation one uses expression (8) to relate the path lengths along the perturbed ray and the reference ray. This expression is sensitive to discrepancies in the directions of the reference ray and the perturbed ray. It is striking that for large values of $\varepsilon S_0/L$ (say around 3) where the ray deflection is not described well by the perturbation theory, the traveltime computed by numerical integration along the perturbed ray is relatively accurate. This is a consequence of Fermat's theorem; since the true ray renders the traveltime stationary, the traveltime is relatively insensitive to perturbations in the ray position.

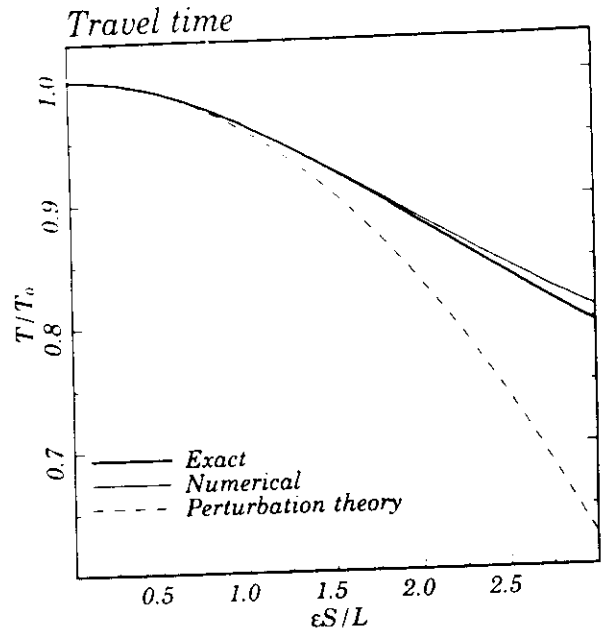

Figure 4. The exact traveltime, the traveltime obtained from perturbation theory using (61) and the traveltime from numerical integration along the perturbed ray for the model of Fig. 3 as a function of the perturbation $\varepsilon S/L$.

9 EXAMPLE 2: APPLICATION TO THE EARTH'S MANTLE

In solid earth geophysics, large-scale linearized inversions of traveltimes have been used to construct 3-D models of the P velocity in the upper mantle (e.g. Aki, Christofferson & Husebye 1977; Spakman 1990). The linearization in these inversions is based on Fermat's theorem, and it is assumed that the observed traveltimes are caused by slowness perturbations integrated over the rays in a laterally homogeneous reference model. For a fixed slowness perturbation the true traveltime, computed along the true ray, is always smaller than the traveltime along the reference ray. The resulting models may thus suffer from a bias due to the neglect of ray bending effects. In this section, the application of ray perturbation theory to rays propagating through the earth's mantle is shown.

The model used in this example is relatively crude. The Jeffreys–Bullen model (Jeffreys & Bullen 1940; Bullen & Bolt 1985) is used as a reference model $u_0(\mathbf{r})$. By applying an earth-flattening transformation (Gerver & Markushevich 1966; Buland & Chapman 1983) the true rays and the perturbed rays can be computed in a Cartesian coordinate system. For the slowness perturbation a 3-D quasi-random field with the statistics of a Gaussian distribution is used. The method described by Frankel & Clayton (1986) is used for generating the slowness perturbation. The statistical properties of the slowness perturbation are assumed to be the same at every depth in the (earth-flattened) model. Scale lengths between 100 and 500 km have been used; the peak value of the relative slowness variation is normalized to 3 per cent, which corresponds to rms variations of about 1 per cent. A statistical analysis of ISC traveltimes suggests a slowness perturbation of about 1 per cent in the upper mantle, and dominant scale length of approximately 300 km (Gudmundsson, Davies & Clayton 1990). For the lower mantle, larger scale lengths and weaker inhomogeneities are probably more appropiate. Because of the earth-flattening transformation one would have to use even larger scale lengths in the lower mantle. In this sense, the results in this section may underestimate the applicability of ray perturbation theory to mantle tomography.

True rays are computed with the two-point ray tracing algorithm of Sambridge & Kennett (1990). In the calculations presented here the traveltime accuracy is approximately 0.0001 s. The perturbed rays are computed by constructing a numerical solution of (75). A simple first-order finite difference approximation for the second derivatives leads to a pair of tridiagonal linear systems for the values of q_1 and q_2 at a series of N points along the ray. The coefficients in each system require the values of the reference slowness u_0 at points along the ray which are in between the points where the ray deflection is obtained. This means that the reference ray is sampled at $2N$ points. The degree to which the size of N controls the accuracy of the numerical solution of (75) depends on the complexity of the slowness perturbation u_1 and the length of the reference ray. Since tridiagonal linear systems can be solved efficiently using standard techniques (Press *et al.* 1986) one is able to set N large and ensure numerical accuracy. Computations are performed for epicentral distances of 13.6°, 39.2° and 86.5°.

For an assessment of the usefulness of ray perturbation theory to this problem one needs to consider the accuracy of the ray positions. For every point along the reference ray, a vertical plane perpendicular to the source–receiver line can be defined. The deviation between rays is defined as the distance between the points of intersection within this plane (see Fig. 5).

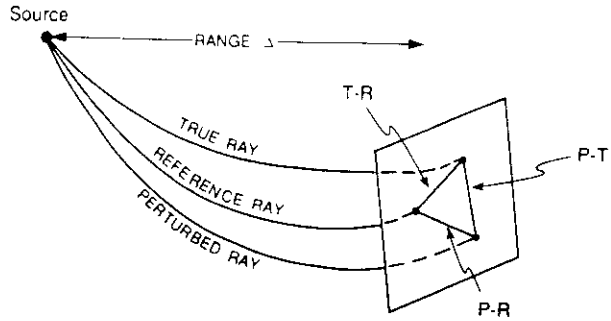

Figure 5. Definition of the distances between the true ray, the perturbed ray, and the ray in the reference medium. The distances are measured in the vertical plane.

For the medium length ray, with an epicentral distance of 39.2°, the distance between the true ray and the reference ray (T − R), the perturbed ray and the reference ray (P − R) and the perturbed ray and the true ray (P–T) is shown in Fig. 6 for various scale lengths L. Note that the slowness model is three-dimensional, so that the difference of T − R and P − R is not equal to T − P. Ray perturbation theory accounts for the ray deflection accurately when the error in the perturbed ray (P − T) is much less than the true deflection (T − R). For scale lengths of 200 km and larger, the error in the ray deflection computed with perturbation theory is much smaller than the total ray deflection. For a scale length of 100 km, the accuracy of the perturbation theory is greatly reduced. The reason for this is that in this case the ray deflection, which is about 70 km, is of the same order of magnitude as the scale length of the slowness perturbation. The condition (65) for the validity of the perturbation theory is therefore not satisfied, and the ray deflection determined from perturbation theory has an error of approximately 40 per cent. One can see from Table 1 that the constant C_2, defined in (68) is larger than unity for $L = 100$ km.

For the long ray, with an epicentral distance of 86.5°, the ray deflection is less accurately described by perturbation theory. In this case the error in the ray deflection (P − T) is about 30 per cent of the true deflection (T − R) for a correlation length of 400 km; this error increases with a decreasing correlation length. The reason for this is that the conditions for the validity of perturbation theory depend critically on the ratio S_0/L (see equations 67–68). However, it should be noted that the employed model is probably too rough for the lower mantle. This is aggravated by the use of the earth-flattening transformation because this transformation extends the deeper structures in the horizontal direction. Rays are most sensitive slowness perturbations near the middle of the ray. For the long ray, the turning point is deep in the lower mantle. For this reason the results for the long ray may be overly pessimistic. For the short ray, with an epicentral distance of 13.6°, the ray deflection is described very well by the perturbation theory (see the error P − T in Table 1). This error is between 4 and 10 per cent of the total ray deflection. Indeed one finds that for this ray the constant C_1 and C_2 are much smaller than unity for all employed scale lengths.

Note that the parameters C_1 and C_2 used in the criteria (67) and (68) are derived for a homogeneous reference slowness and a slowness perturbation with a constant gradient perpendicular to the reference ray. However, a comparison of the error in the ray deflection (P − T) with the total ray deflection (T − R) in Table 1 shows that for more complicated ray geometries the parameters C_1 and C_2 are good indicators for the accuracy of the ray deflection from perturbation theory.

The traveltimes in the ISC data set have an accuracy of about 0.1 s. The theory used for performing the tomographic inversions should match this accuracy. In Table 2, the total traveltime is shown for various length-scales of the inhomogeneity. The difference between the true traveltime and the traveltime in the reference velocity is shown, together with the error in the first- and second-order approximations of the traveltime computed with (20a) and (61) respectively. The error in the traveltime by integrating the slowness perturbation numerically over the perturbed ray is also shown.

For the long ray, the error in the traveltime computed using Fermat's theorem (Error 1st) is about four times as large as the required accuracy of 0.1 s. Note that the error in the first-order traveltime perturbation and from the numerical integration is always positive. This is due to the fact that both traveltimes are computed by integration along a curve which is not the true ray. Since the true ray is a curve of minimum traveltime this implies that the error induced by integrating over another curve is always positive. In contrast to this, the error in the traveltime computed with second-order perturbation theory has an alternating sign. Using the traveltime computed from second-order perturbation theory will thus reduce the bias in the models obtained from tomographic inversions. For the long ray, accounting for ray bending effects with perturbation theory reduces the error in the traveltime for the long ray by a factor of 10, so that the error in the traveltime is about 0.04 s (see the entries 'Error 2nd' and 'Error int' in Table 2).

It may be surprising that for the long ray the traveltime computed with ray perturbation theory is very accurate, while the ray deflection has a relative error between 10 and 50 per cent (see Table 1). As shown in the example of Fig. 4, relatively large errors in the ray deflection do not necessarily produce large errors in the traveltime computed by numerical integration along the perturbed ray. In contrast to the situation in Fig. 4, the angle between the reference ray and the true ray in the examples in this section is extremely small. This is reflected in the small value of C_1 in Table 1; this quantity measures to what extent the

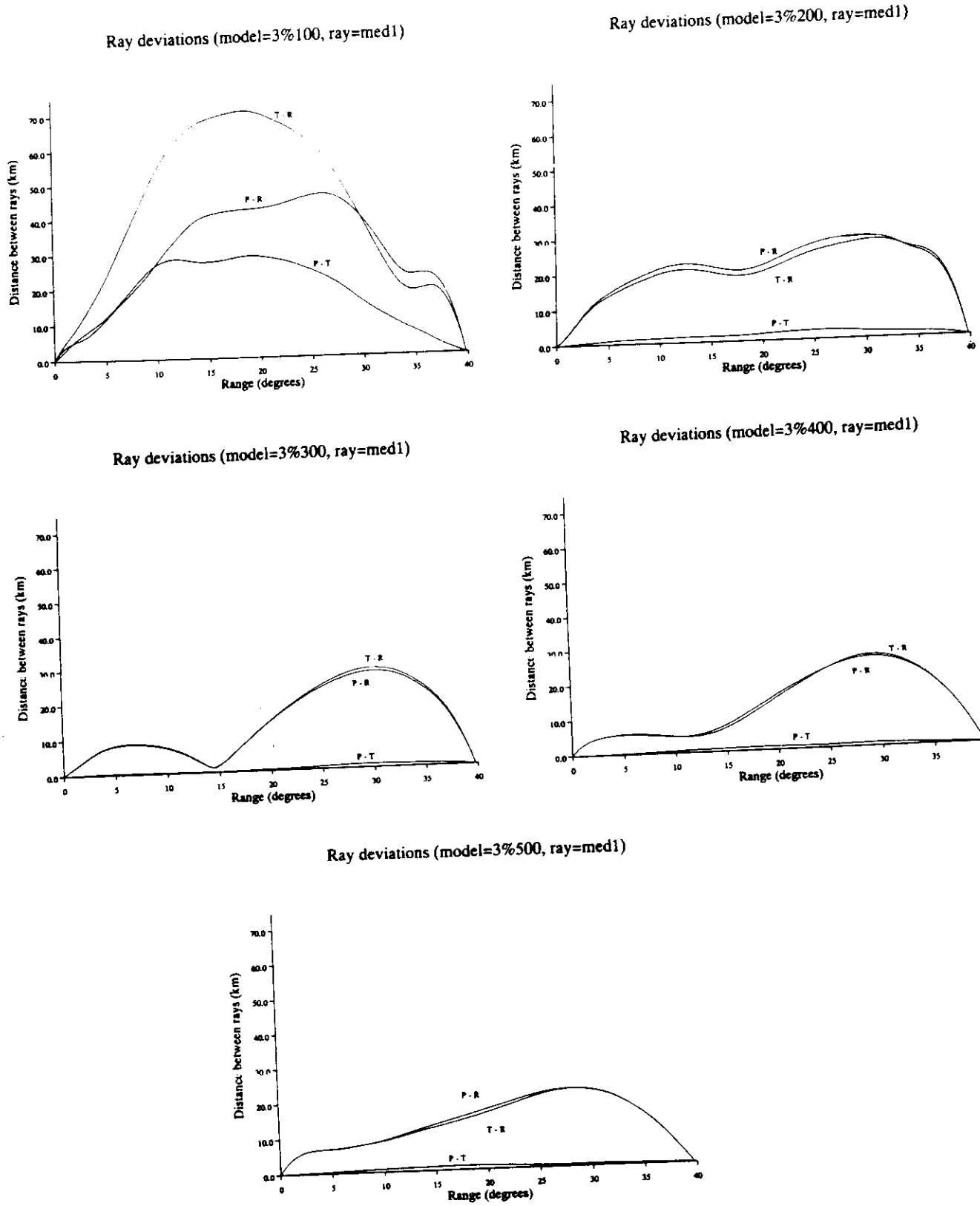

Figure 6. The distance between the true ray, the perturbed ray and the reference ray as defined in Fig. 5 for the continuous quasi-random model of the earth's mantle used in Section 9 for a ray of epicentral distance of 39.2°. The correlation length is indicated above each figure.

Table 1. The distance (in km) as defined in Fig. 5 between the true ray and the reference ray (T − R), the perturbed ray and the reference ray (P − R), and the perturbed ray and the true ray (P − T) averaged along rays for different epicentral distances and correlation lengths L. The constants C1 and C2 are defined in (67) and (68).

L	Dist.	T-R	P-R	P-T	C1	C2
200	86.5	45.49	33.05	22.11	0.230	3.548
300	86.5	43.43	38.46	16.01	0.181	1.864
400	86.5	47.50	37.69	15.90	0.167	1.290
500	86.5	50.84	36.07	6.69	0.165	1.018
100	39.2	35.40	20.42	20.52	0.171	2.174
200	39.2	17.75	16.37	1.86	0.076	0.484
300	39.2	12.83	11.07	2.34	0.072	0.304
400	39.2	11.69	9.98	2.26	0.074	0.234
500	39.2	11.52	9.99	1.88	0.072	0.184
100	13.6	6.56	7.24	0.70	0.071	0.274
200	13.6	3.64	3.51	0.14	0.041	0.079
300	13.6	1.84	1.77	0.08	0.028	0.036
400	13.6	2.89	2.66	0.23	0.019	0.019
500	13.6	3.47	3.15	0.32	0.015	0.011

requirement (64) is satisfied and to what extent the quantity $\partial s/\partial s_0$ is close to unity. The reason that the traveltime computed from second-order perturbation theory is so close to the traveltime computed by numerical integration along the perturbed ray is thus a consequence of the fact that the angle between the reference ray and the perturbed ray is extremely small. For the short ray and the medium ray, the error in the traveltime reduces by a factor of between 5 and 10 by taking ray bending effects into account using perturbation theory for the ray deflection (see Table 2).

Equation (75) for the ray deflection and equation (61) for the second-order traveltime perturbation are implemented by using a discretization along the reference ray. The number of discretization points that is required depends on the length of the ray and the length-scale of the slowness fluctuations. In Table 3, the error in the traveltime is shown for the ray with an epicentral distance of 39.2° for various numbers of discretization points and a correlation length $L = 300$ km of the slowness perturbation. The same number of discretization points N was used for the numerical solution of the ray deflection and for the integration of the traveltime. The error in the first-order traveltime perturbation is only sensitive for the number of discretization points when one uses less than 58 points. The reason for this is that the first-order traveltime perturbation is independent of the ray deflection, so that the discretization of (75) does not affect the first-order traveltime. The error in the second-order traveltime computed from (61) does not change significantly with N for values of N between 58 and 2243. For larger values of N the cumulative effect of the round-off errors in the discretization of (75) are apparent. For smaller values of N the discretization is simply too rough. The results in Table 3 indicate that for the media used here a relative small number of discretization points (around 200) is sufficient for the computation of the traveltime perturbation.

Table 2. The total traveltime (Tref) of the reference ray, the true perturbation in the traveltime (True-ref), the error in the first-order traveltime computed with (20a) (Error 1st), the error in the second-order traveltime computed with (61) (Error 2nd), and the error in the traveltime computed by numerical integration along the perturbed ray (Error int), for varying epicentral distances and correlation lengths L. The time units are seconds.

L	Dist.	Tref	True-ref	Error 1st	Error 2nd	Error int
200	86.5	756.9	-2.0738	0.4223	0.0305	0.0576
300	86.5	756.9	-3.0359	0.4021	-0.0398	0.0399
400	86.5	756.9	-3.8291	0.3678	-0.0020	0.0209
500	86.5	756.9	-4.1722	0.3344	0.0461	0.0191
100	39.2	448.3	-0.4365	0.3462	0.1235	0.1120
200	39.2	448.3	-1.6070	0.1332	0.0077	0.0026
300	39.2	448.3	-3.0264	0.0938	0.0143	0.0069
400	39.2	448.3	-4.4218	0.0703	0.0122	0.0030
500	39.2	448.3	-5.4943	0.0513	0.0094	0.0068
100	13.6	188.2	-1.3414	0.0355	-0.0033	0.0002
200	13.6	188.2	-1.5993	0.0116	0.0003	0.0003
300	13.6	188.2	-1.6044	0.0063	0.0015	0.0015
400	13.6	188.2	-1.6279	0.0061	0.0007	0.0001
500	13.6	188.2	-1.7131	0.0084	0.0023	0.0024

Table 3. Accuracy of the traveltime computations for the medium length ray, for the Jeffreys–Bullen reference model and the continuous quasi-random slowness perturbation with a correlation length $L = 300$ km, as a function of the number of discretization points N. The errors are defined in Table 2.

N.	Tref	True-ref	Error 1st	Error 2nd	Error int
22418	448.3	-3.026	0.0938	0.0143	0.0044
2243	448.3	-3.026	0.0938	0.0051	0.0023
226	448.3	-3.026	0.0938	0.0050	0.0056
114	448.3	-3.026	0.0937	0.0050	0.0155
58	448.3	-3.026	0.0937	0.0048	0.0526
22	448.3	-3.026	0.0967	0.0087	0.3784

The model used here for the slowness perturbation in the earth's mantle is rather crude. It does not take into account that the lower mantle may be smoother than the upper mantle. The length-scale of the inhomogeneity is also not corrected for the effects of the earth-flattening transformation. Furthermore, a quasi-random slowness perturbation is not completely satisfactory for studying ray bending effects in mantle tomography, because the earth's mantle contains significant organized structures such as subduction zones. The validity of ray perturbation theory for mantle tomography is presently further investigated.

10 EXAMPLE 3: A DISCONTINUOUS SLOWNESS PERTURBATION

The theory for the ray perturbation was developed for continuous slowness perturbations, because Taylor expansions have been used in the derivation of the perturbed traveltime (11) and the ray perturbation (23). However, discontinuities in the slowness perturbation can be handled by representing the discontinuities by Heaviside functions and the gradient of the slowness perturbation by Dirac delta functions. An example of this is shown for the geometry of Fig. 7, where a homogeneous reference medium is perturbed with a constant slowness perturbation in the half-space $z > 0$:

$$u_1(z) = -u_0 H(z), \tag{92}$$

where $H(z)$ is the Heaviside function.

The ray is not deflected out of the plane of Fig. 7, so that only the deflection in the plane of the diagram needs to be considered. For this reason the subscript on the coordinate q_1 is deleted in this section. In this case the differential equation (73a) for the ray deflection becomes

$$\ddot{q} = \sin i_0 \, \partial_z \left(\frac{u_1}{u_0} \right). \tag{93}$$

From (92) it follows that

$$\partial_z \left(\frac{u_1}{u_0} \right) = -\delta(z), \tag{94}$$

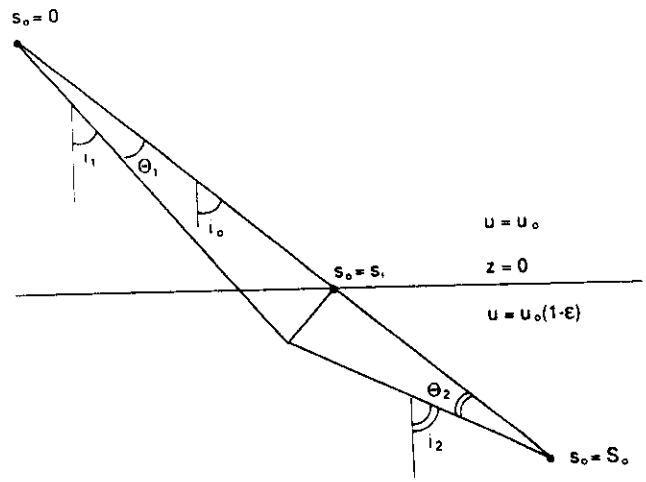

Figure 7. Definition of the geometric variables for the two-layer model.

where $\delta(z)$ is the Dirac delta function. The derivatives in equation (93) are with respect to the parameter s_0. The delta function in the right-hand side of (94) needs to be expressed in s_0. This can be achieved using

$$\delta[z(s_0)] = \frac{1}{\left|\frac{\partial z}{\partial s_0}\right|_{s_0=s_i}} \delta(s_0 - s_i) = \frac{1}{\cos i_0} \delta(s_0 - s_i), \tag{95}$$

where s_i denotes the distance between the source and the point of intersection of the straight reference ray with the slowness discontinuity.

Inserting (95) in (93) gives

$$\ddot{q} = -\tan i_0 \, \delta(s_0 - s_i). \tag{96}$$

Integrating this equation leads with the boundary conditions (51) to the solution

$$q(s_0) = \tan i_0 \left(1 - \frac{s_i}{S_0}\right) s_0 \quad \text{for} \quad 0 < s_0 < s_i, \qquad q(s_0) = \tan i_0 s_i \left(1 - \frac{s_0}{S_0}\right) \quad \text{for} \quad s_i < s_0 < S_0. \tag{97}$$

Note that this solution breaks down when $i_0 = \pi/2$. This is the case when the source and the receiver are on the same side of the slowness discontinuity. This situation is not consistent with the assumption that the straight reference ray intersects the slowness perturbation. It will be clear that the solution of (93) will not give an accurate description of rays that are refracted along interfaces.

However, the solution (97) gives for small values of ε a good description of the perturbation of a reference ray that does intersect the slowness discontinuity. In Fig. 8 the true rays and the solution (97) from the perturbation theory are shown for several values of ε. For small values of ε the true rays and the rays obtained from perturbation theory are close, but for values of ε of the order 0.4 the solutions start to diverge. As shown in Fig. 7, the solution (97) leads to piecewise straight rays. Note that the kink in the representation of the perturbed ray does not occur exactly at the slowness discontinuity, which is an artifact of the use of ray coordinates. Since this is a second-order effect in the ray displacement this does not distract from the use of first-order perturbation theory for the ray perturbation. It is shown in Appendix B that Snell's law is satisfied to first order by the solution (97).

The second-order traveltime perturbation computed from (61) while treating the slowness gradient as a Dirac delta function is correct to second order, despite the fact that the perturbation theory is developed for continuous slowness perturbations. This is shown here for the case that the source and the receiver are separated by the same distance $D/2$ from the slowness discontinuity. The horizontal distance between the endpoints of the ray is denoted by X (see Fig. 9). In this case we

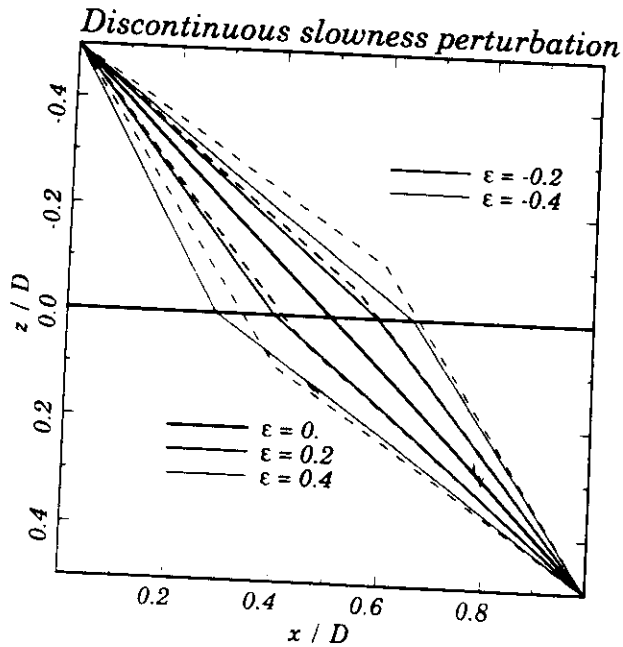

Figure 8. The true rays (solid lines) and the perturbed ray (dashed lines) for the two-layer model for various values of the relative slowness contrast ε.

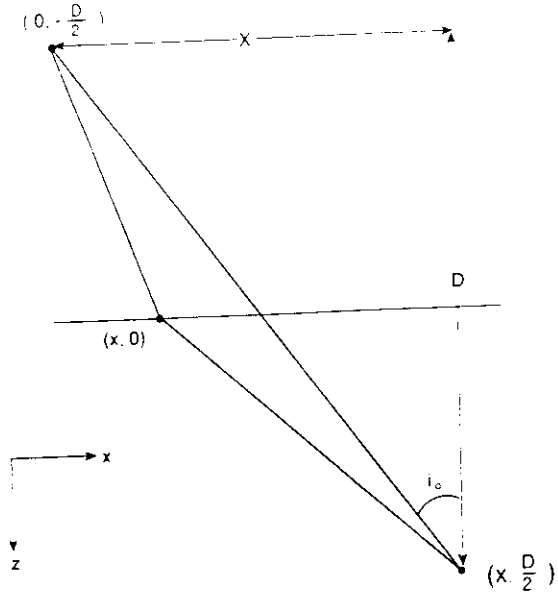

Figure 9. Definition of geometric variables for the compution of the second-order traveltime perturbation for the two-layer model.

have $s_i = S_0/2$. The second-order traveltime follows by combination of (61) and (94–95) and is given by

$$T_2 = -\frac{1}{2}\int_0^{S_0} u_0 q(s_0) u_0 \tan i_0 \delta\left(s_0 - \frac{S}{2}\right) ds_0. \tag{98}$$

Carrying out the integration over s_0, using the solution (97) and using that $\tan i_0 = X/D$ (see Fig. 9) gives

$$T_2 = -\frac{1}{8}\frac{X^2}{D^2} T_0, \tag{99}$$

where $T_0 = u_0 S_0$ is the traveltime of the unperturbed ray. In this case of a straight reference ray the second-order traveltime perturbation is always negative, regardless of the sign of the slowness perturbation. This reflects the fact that ray bending

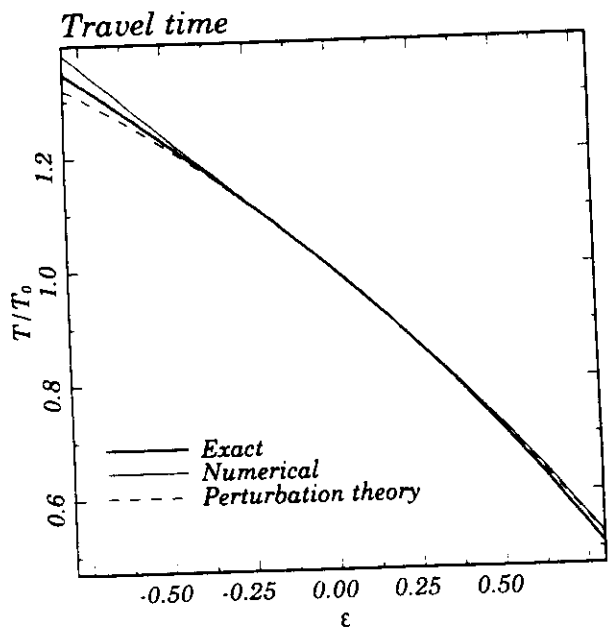

Figure 10. The exact traveltime, the traveltime obtained from perturbation theory using (61) and the traveltime from numerical integration along the perturbed ray for the two-layer model as a function of the relative slowness contrast ε.

effects reduce the traveltime. Note also that for $X = 0$ the second-order traveltime perturbation vanishes, which is due to the fact that a ray perpendicular to the slowness discontinuity is not affected by the slowness discontinuity. The second-order traveltime in (99) is based on perturbation theory for continuous slowness perturbations, where a discontinuity is treated as a Dirac delta function in the slowness gradient. As shown in Appendix C the result (99) agrees with the second-order traveltime perturbation computed for a discontinuous slowness model.

The true traveltime, and the traveltime from the integration along the perturbed ray are shown in Fig. 10 as a function of ε. As in Fig. 4, the traveltime along the perturbed ray is always larger than the true traveltime. Note that the traveltime perturbation is dominated by a linear trend, and that the second-order effects on the traveltime are rather small. Indeed, for values of the slowness perturbation as large as $\varepsilon = 0.2$ the relative error in the traveltime is less than 0.1 per cent.

This example shows that one can use the ray perturbation theory of this paper for slowness perturbations with discontinuities, by describing the gradient of the slowness perturbation in terms of Dirac delta functions.

11 DISCONTINUITIES IN THE REFERENCE MODEL AND THE SLOWNESS PERTURBATION

Up to this point, the reference model $u_0(\mathbf{r})$ and the slowness perturbation $u_1(\mathbf{r})$ were assumed to be continuous. This was used in the expansions (11) and (23) where Taylor expansions of $u_0(\mathbf{r})$ and $u_1(\mathbf{r})$ were used. In fact, the ray tracing equation (1), which forms the basis of the employed perturbation theory, breaks down at slowness discontinuities. In this situation the ray tracing equation (1) should be replaced by matching conditions at the slowness discontinuity. These matching conditions, the continuity of the ray position and Snell's law, can be used to derive a perturbation scheme for the ray deflection and the traveltime. In this paper the simplest case of a plane discontinuity is considered for the special case where the reference slowness does not vary along the discontinuity. For the generalization to curved interfaces and variations of the reference slowness along the interface, an extension of the theory similar to the work of Farra *et al.* (1989) is needed.

Consider a plane discontinuity in the slowness. Both the reference slowness u_0 and the slowness perturbation u_1 may be discontinuous across the interface. Let the unit vector perpendicular to the interface be denoted by $\hat{\mathbf{n}}$; see Fig. 11 for a definition of the variables. Quantities on one side of the interface are denoted with superscript $(-)$, and the corresponding quantity on the other side of the interface is denoted with the superscript $(+)$. The change of a quantity ξ across the interface is denoted by

$$[\xi]_-^+ \equiv \xi^{(+)} - \xi^{(-)}. \tag{100}$$

The reference ray is continuous across the interface, so that

$$[\mathbf{r}_0]_-^+ = 0. \tag{101}$$

The continuity condition of the perturbed ray requires some attention. The slowness discontinuity affects the ray perturbation in two ways. First, a discontinuity in the relative slowness perturbation (u_1/u_0) leads to a kink in the perturbed ray. Second, when the reference model $u_0(\mathbf{r})$ is discontinuous, the reference ray has a kink at the interface. As shown in Fig. 11, the

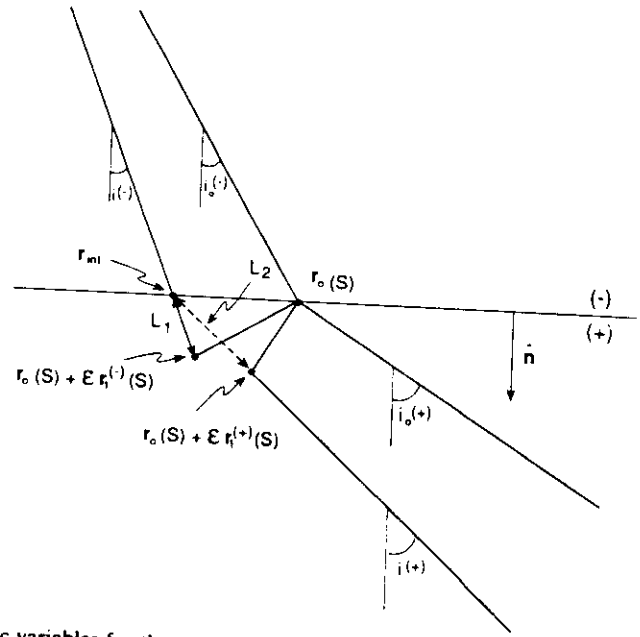

Figure 11. Definition of the geometric variables for the analysis of the ray perturbation at a discontinuity.

definition of the ray coordinates produces in this situation a discontinuity in the representation of the perturbed ray. This is an artifact of the use of ray coordinates that needs to be corrected. This can be achieved by determining the point of intersection of the perturbed rays (or their extensions) with the interface. By imposing an appropriate continuity condition on the ray deflection on the two sides of the interface, one can obtain a continuous perturbed ray by replacing the broken representation of the perturbed ray by a new representation which is continuous at the interface (see the dashed line in Fig. 11). The condition to be imposed on the perturbed rays is therefore that the perturbed rays (or their extensions) intersect the interface in the same position $\mathbf{r}_{int}$.

Consider for the moment a coordinate system with the z-axis perpendicular to the interface, and let the interface be located at $z = z_0$. Let the interface be located at arc length S along the unperturbed ray. On either side of the interface the ray can in the vicinity of the interface be parametrized as

$$\mathbf{r}(s_0) = \mathbf{a} + \mathbf{b}s_0, \tag{102}$$

with

$$\mathbf{a} = \mathbf{r}_0(S) + \varepsilon\mathbf{r}_1(S), \tag{103a}$$

$$\mathbf{b} = \dot{\mathbf{r}}_0(S) + \varepsilon\dot{\mathbf{r}}_1(S). \tag{103b}$$

Let the point of intersection of the perturbed ray with the plane of the discontinuity occur for $s_0 = s_{int}$; this point is defined by the condition

$$z_0 = (\hat{\mathbf{n}} \cdot \mathbf{r}_{int}) = (\mathbf{a} \cdot \hat{\mathbf{n}}) + (\mathbf{b} \cdot \hat{\mathbf{n}})s_{int}. \tag{104}$$

Solving this expression for s_{int}, inserting the result in (102) and using the definitions (103a,b) gives to first order

$$\mathbf{r}_{int} = \mathbf{r}_0 + \varepsilon\left(\mathbf{r}_1 - \frac{(\hat{\mathbf{n}} \cdot \mathbf{r}_1)}{(\hat{\mathbf{n}} \cdot \dot{\mathbf{r}}_0)}\dot{\mathbf{r}}_0\right). \tag{105}$$

In this expression and the ensuing matching conditions, all quantities should be evaluated at $s_0 = S$. Since the perturbed rays (or their extensions) must intersect the interface in the same position, the quantity derived in (105) must be continuous across the interace. Because of the continuity of $\mathbf{r}_0$ (see 101), this implies that

$$\left[\mathbf{r}_1 - \frac{(\hat{\mathbf{n}} \cdot \mathbf{r}_1)}{(\hat{\mathbf{n}} \cdot \dot{\mathbf{r}}_0)}\dot{\mathbf{r}}_0\right]_-^+ = 0. \tag{106}$$

The second continuity condition comes from the requirement that the perturbed rays satisfy Snell's law, which states that (Aki & Richards 1980)

$$\left[\hat{\mathbf{n}} \times u\frac{d\mathbf{r}}{ds}\right]_-^+ (s_{int}) = 0. \tag{107}$$

Note that this continuity condition holds at the point of intersection $\mathbf{r}_{int}$ of the perturbed ray with the interface, hence the argument s_{int} rather than $s_0 = S$. Expanding (107) around the point $s_0 = S$ gives

$$\left[\hat{\mathbf{n}} \times u\frac{d\mathbf{r}}{ds}\right]_-^+ (S) + \varepsilon(s_{int} - S)\frac{d}{ds_0}\left[\hat{\mathbf{n}} \times u\frac{d\mathbf{r}}{ds}\right]_-^+ (S) = 0. \tag{108}$$

The factor ε has been introduced because $(s_{int} - S)$ is linear in the ray deflection. Consider the last term in (108). Since only the first-order ray deflection is desired one can replace all quantities in the term

$$\frac{d}{ds_0}\left[\hat{\mathbf{n}} \times u\frac{d\mathbf{r}}{ds}\right]_-^+$$

by their zeroth-order approximations. Carrying out the differentiation with respect to s_0, using that the discontinuity is plane so that $\hat{\mathbf{n}}$ is a constant, and using (17) for the differentiation of u_0, one finds that to $O(\varepsilon)$

$$\varepsilon\frac{d}{ds_0}\left[\hat{\mathbf{n}} \times u\frac{d\mathbf{r}}{ds}\right]_-^+ (S) = \varepsilon\hat{\mathbf{n}} \times [\dot{\mathbf{r}}_0(\dot{\mathbf{r}}_0 \cdot \nabla u_0) + u_0\ddot{\mathbf{r}}_0]_-^+. \tag{109}$$

With (16) this implies that

$$\varepsilon\frac{d}{ds_0}\left[\hat{\mathbf{n}} \times u\frac{d\mathbf{r}}{ds}\right]_-^+ (S) = \varepsilon\hat{\mathbf{n}} \times [\nabla u_0]_-^+. \tag{110}$$

When u_0 does not vary along the interface, ∇u_0 and $\hat{\mathbf{n}}$ are parallel, and the right-hand side of (110) vanishes. (This condition

can in fact be slightly relaxed, since only the discontinuity in the gradient of u_0 along the interface needs to vanish.) Using this in (108) implies that

$$\left[\hat{\mathbf{n}} \times u \frac{d\mathbf{r}}{ds}\right]_-^+ (S) = 0. \tag{111}$$

From this point on all quantities are evaluated for $s_0 = S$; the variable S is therefore omitted. The unperturbed ray also satisfies Snell's law at the point $s_0 = S$, so that

$$[\hat{\mathbf{n}} \times u_0 \dot{\mathbf{r}}_0]_-^+ = 0. \tag{112}$$

Insert (2) and (22) in (111), with (112) this implies that to first order

$$[\{u_1 - u_0(\dot{\mathbf{r}}_1 \cdot \dot{\mathbf{r}}_0)\}\hat{\mathbf{n}} \times \dot{\mathbf{r}}_0 + u_0 \hat{\mathbf{n}} \times \dot{\mathbf{r}}_1]_-^+ = 0. \tag{113}$$

The condition (106) imposes a relation on $[\mathbf{r}_1]_-^+$, while Snell's law (113) constrains the kink in the ray perturbation $[\dot{\mathbf{r}}_1]_-^+$.

The matching conditions (106) and (113) are expressed in terms of the ray perturbation $\mathbf{r}_1$. However, as in Section 4 it is advantageous to re-express these conditions in ray coordinates. This results in matching conditions for q_1, q_2 and their derivatives. It is convenient to define two unit vectors $\hat{\mathbf{e}}_\parallel$ and $\hat{\mathbf{e}}_\perp$ in the plane of the interface (see Fig. 12). The unit vector $\hat{\mathbf{e}}_\parallel$ lies along the intersection of the plane of the interface, with the plane spanned by the incoming and the outgoing rays. The unit vector $\hat{\mathbf{e}}_\perp$ is the unit vector in the interface perpendicular to $\hat{\mathbf{e}}_\parallel$. The orthogonality of $\hat{\mathbf{e}}_\parallel$ and $\hat{\mathbf{e}}_\perp$ implies that

$$\hat{\mathbf{e}}_\parallel \cdot (\hat{\mathbf{n}} \times \dot{\mathbf{r}}_0) = 0. \tag{114}$$

The conditions (106) and (113) can be recast as conditions for q_1 and q_2 by using (35) and (46) and dotting the results with both $\hat{\mathbf{e}}_\parallel$ and $\hat{\mathbf{e}}_\perp$. This gives using (114):

$$\left[q_1(\hat{\mathbf{e}}_\parallel \cdot \hat{\mathbf{q}}_1) + q_2(\hat{\mathbf{e}}_\parallel \cdot \hat{\mathbf{q}}_2) - \frac{(\hat{\mathbf{e}}_\parallel \dot{\mathbf{r}}_0)}{(\hat{\mathbf{n}} \cdot \dot{\mathbf{r}}_0)}[q_1(\hat{\mathbf{n}} \cdot \hat{\mathbf{q}}_1) + q_2(\hat{\mathbf{n}} \cdot \hat{\mathbf{q}}_2)]\right]_-^+ = 0, \tag{115a}$$

$$\left[q_1(\hat{\mathbf{e}}_\perp \cdot \hat{\mathbf{q}}_1) + q_2(\hat{\mathbf{e}}_\perp \cdot \hat{\mathbf{q}}_2) - \frac{(\hat{\mathbf{e}}_\perp \cdot \dot{\mathbf{r}}_0)}{(\hat{\mathbf{n}} \cdot \dot{\mathbf{r}}_0)}[q_1(\hat{\mathbf{n}} \cdot \hat{\mathbf{q}}_1) + q_2(\hat{\mathbf{n}} \cdot \hat{\mathbf{q}}_2)]\right]_-^+ = 0, \tag{115b}$$

$$[u_0\{(\dot{q}_1 - \Omega q_2)\hat{\mathbf{e}}_\parallel \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{q}}_1) + (\dot{q}_2 + \Omega q_1)\hat{\mathbf{e}}_\parallel \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{q}}_2)\}]_-^+ = 0, \tag{115c}$$

$$[u_1 \hat{\mathbf{e}}_\perp \cdot (\hat{\mathbf{n}} \times \dot{\mathbf{r}}_0) + u_0\{(\dot{q}_1 - \Omega q_2)\hat{\mathbf{e}}_\perp \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{q}}_1) + (\dot{q}_2 + \Omega q_1)\hat{\mathbf{e}}_\perp \cdot (\hat{\mathbf{n}} \times \hat{\mathbf{q}}_2)\}]_-^+ = 0. \tag{115d}$$

All unit vectors in these expressions are known, so they provide explicit matching conditions for q_1, q_2, $\dot{q}_1$ and $\dot{q}_2$.

Note that the unit vectors $\hat{\mathbf{q}}_i$ may be chosen completely differently on either side of the interface. In fact, if the reference ray has a kink at least one of the $\hat{\mathbf{q}}_i$ must be discontinuous across the interface. As long as one uses the $\hat{\mathbf{q}}_i$ appropriate for the two sides of the interfaces in (115) this poses no problems. Similarly, the rate of rotation Ω of the unit vectors can be chosen independently on the two sides of the interface.

At the interface of the slowness discontinuity, the matching conditions (115) are to be used rather than the differential equation (50). In the regions between the slowness discontinuities one should of course use (50) for the computation of the ray perturbation. When one discretizes the ray perturbation, one can express the matching conditions as linear equations with a tridiagonal structure. The resulting linear equations thus have the same algebraic structure as the linear equations that result from the discretization of expression (50) for the regions between the slowness discontinuities. This means that incorporating discontinuities in the reference slowness does not alter the structure of the algorithm for the computation of the first-order ray deflection.

As an example of the matching conditions, consider a layered reference medium with horizontal interfaces $z = constant$.

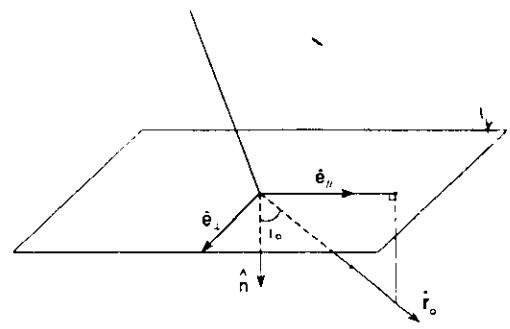

Figure 12. Definition of the unit vectors $\hat{\mathbf{e}}_\parallel$ and $\hat{\mathbf{e}}_\perp$ in the plane of the discontinuity.

For this situation it is convenient to assume that the $\hat{\mathbf{q}}$ vectors do not rotate around the reference ray, condition (69), and to use the expressions (70) for the $\hat{\mathbf{q}}_i$. Referring to Fig. 2 we have in this case

$$\hat{\mathbf{e}}_{\parallel} = \hat{\mathbf{x}}, \tag{116a}$$

$$\hat{\mathbf{e}}_{\perp} = \hat{\mathbf{y}}, \tag{116b}$$

$$\hat{\mathbf{n}} = \hat{\mathbf{z}}. \tag{116c}$$

Using these relations in (115) and using (74) to eliminate the $\sin i_0$ and $\cos i_0$ terms in favour of the ray parameter p one obtains the following matching conditions:

$$\left[\frac{u_0}{\sqrt{u_0^2 - p^2}} q_1\right]_-^+ = 0, \tag{117a}$$

$$[q_2]_-^+ = 0, \tag{117b}$$

$$\left[p\frac{u_1}{u_0} - \sqrt{u_0^2 - p^2}\dot{q}_1\right]_-^+ = 0, \tag{117c}$$

$$[u_0 \dot{q}_2]_-^+ = 0. \tag{117d}$$

As an example, for the slowness perturbation (92) of the two-layer model of Section 10, one finds from (117) and the definition (74) that $[\dot{q}_1]_-^+ = -\tan i_0$. This result is equivalent to expression (96) obtained by replacing the gradient of (u_1/u_0) by a delta function when this quantity is discontinuous.

For the computation of the second-order traveltime perturbation one can either integrate numerically over the perturbed ray using the dashed interval L_2 in Fig. 11, or one can use the expressions of Section 5 for T_2. To see that these expressions can be used when slowness discontinuities are present one should realize that the derivations in the Sections 2 and 5 hold for any subsection of the ray between the discontinuities. There is no reason why the interval $0 < s_0 < S_0$ in the Sections 2 and 5 should denote the whole ray, it may also denote a subsection of the ray between the interfaces. The generalized condition (15) for the validity of Fermat's theorem allows for this flexibility. Because of the use of ray coordinates, the boundary condition (15) is automatically satisfied at all points along the reference ray, hence also at the points where the reference ray intersects the interface. This implies that when one uses for example (61) for the computation of T_2, that one should take the boundary term $\frac{1}{2}u_0 q_i \dot{q}_i$ in (61) into account at the interface because this quantity is not necessarily continuous across the interface. One can verify easily that the contribution of this boundary term in (61) from subsections of the ray at the opposite sides of the interface leads to a contribution $[\frac{1}{2}u_0 q_i \dot{q}_i]_-^+$ to the second-order traveltime perturbation. Ignoring this contribution in the example of Section 12 for a mantle model with realistic discontinuities leads to a tenfold increase in the error in the second-order traveltime. For the special case of a layered reference medium with horizontal interfaces, as in expression (117), one can eliminate the q_2 contribution to the interface term: $[\frac{1}{2}u_0 q_i \dot{q}_i]_-^+ = [\frac{1}{2}u_0 q_1 \dot{q}_1]_-^+$.

12 EXAMPLE 4: DISCONTINUOUS SLOWNESS MODELS FOR THE EARTH'S MANTLE

As an illustration of the effect of discontinuities, it is shown in this section what the effect of discontinuities of the reference model and the slowness perturbation are for waves propagating through the earth's mantle. The PREM model (Dziewonski & Anderson 1981) is used as a reference model. This model has a pronounced discontinuity in the slowness at a depth of about 670 km. For the slowness perturbation, the quasi-random model used in Section 9 is used. The only difference is that the slowness perturbations are independent in the upper and lower mantle, so that the slowness perturbation is also discontinuous at a depth of 670 km.

The ray deviations for an epicentral distance of 39.2° are shown in Fig. 13 for various values of the correlation length L. The ray deviations, as defined in Fig. 5, exhibit a kink at the location where the ray intersects the 670 km discontinuity. This is due to the fact that the ray deviations are measured in a vertical plane perpendicular to the source–receiver line. When the rays have a kink this implies that the ray deviation has a discontinuity.

A pronounced difference with the ray deviations of Fig. 6 for the continuous slowness model is that the ray deviation in Fig. 13 are largest at the slowness discontinuity. At these locations the ray has a kink; apparently this increases the sensitivity of the ray deflection to perturbations in the slowness. As in Fig. 6, the perturbed ray is not accurate for the shortest correlation length ($L = 100$ km). For the continuous slowness model used for Fig. 6 the error in the perturbed ray steadily decreases when the correlation length L is increased and the model becomes smoother. In that case the error in the perturbed ray is negligible when $L = 500$ km. This is not the case in Fig. 13 for the discontinuous slowness model. The reason for this is that for the discontinuous slowness model the ray deflection is caused both by the slowness variations with correlation length L in the upper and lower mantle, and by the discontinuity in the slowness perturbation $[u_1/u_0]_-^+$ at a depth of 670 km. When L increases, the effect of the quasi-random slowness perturbation on the ray deflection decreases, but the effect of the slowness discontinuity remains constant. One can conclude that the error in the ray deflection in Fig. 13 for $L > 300$ km is due to the

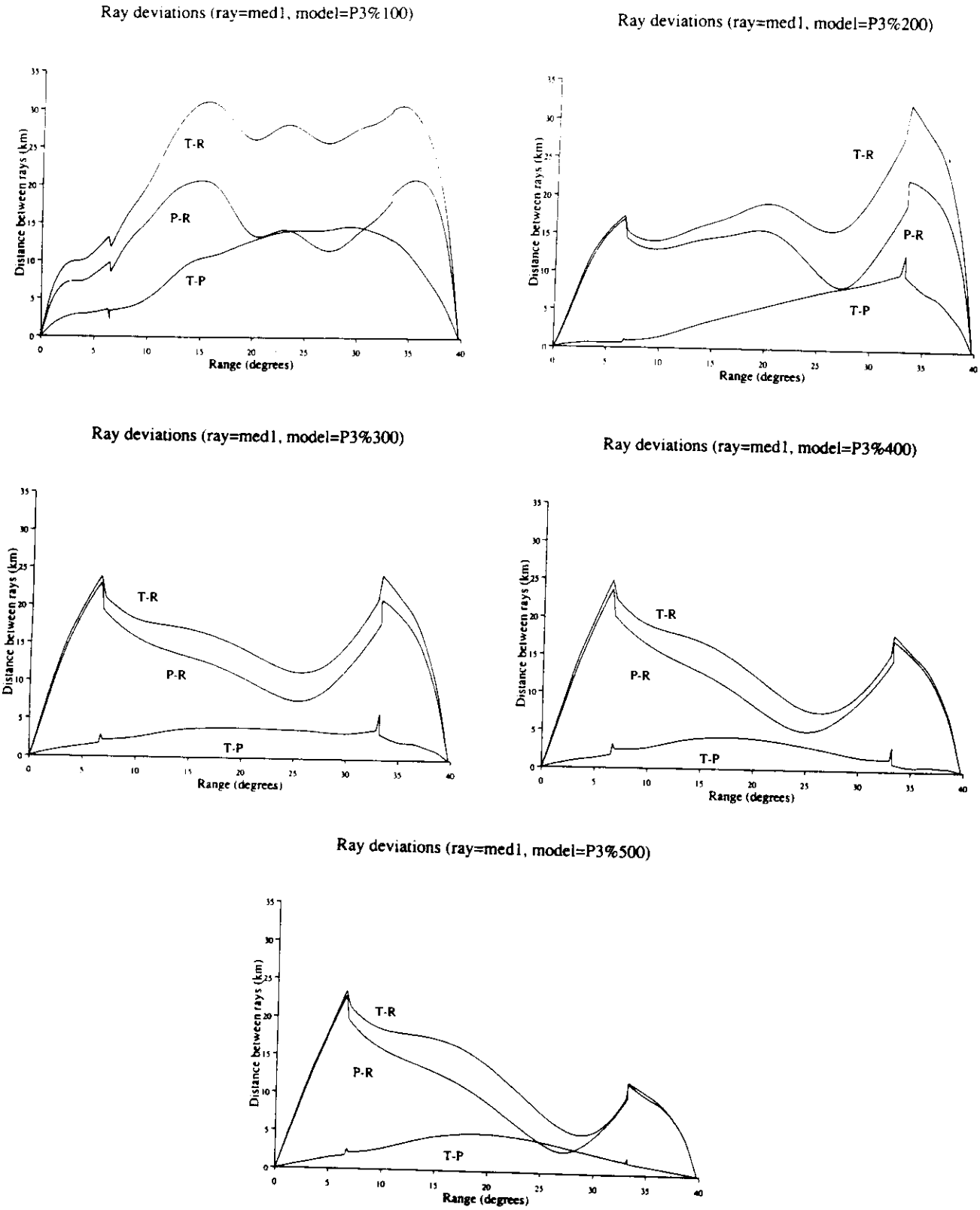

Figure 13. The distance between the true ray, the perturbed ray and the reference ray as defined in Fig. 5 for the quasi-random model of the earth's mantle with discontinuities used in Section 12 for a ray of epicentral distance of 39.2°. The correlation length is indicated above each figure.

Table 4. Errors in the traveltime for a ray with an epicentral distance of 39.2° where PREM is the reference model and where the slowness perturbation is independent in the upper and lower mantle. The errors are defined as in Table 2.

L	Dist.	Tref	True-ref	Error 1st	Error 2nd	Error int
100	39.2	444.7	1.0616	0.1666	0.0514	0.0197
200	39.2	444.7	1.3109	0.1028	0.0165	0.0061
300	39.2	444.7	1.1690	0.0881	0.0070	0.0031
400	39.2	444.7	0.6871	0.0765	0.0051	0.0042
500	39.2	444.7	0.1433	0.0635	0.0007	0.0014

slowness discontinuity at a depth of 670 km rather than the continuous slowness perturbation. For these correlation lengths the error in the ray deflection is about 20 per cent of the maximum ray deflection.

The errors in the traveltime for the discontinuous slowness model are shown in Table 4 for a ray with an epicentral distance of 39.2° for various values of the correlation length. The numerical integration of the slowness along the perturbed ray was performed along the dashed interval L_2 in Fig. 11. The second-order traveltime perturbation was computed with (61), including the contribution of the boundary term at the discontinuities. A comparison with Table 2 reveals that the error in the first- and second-order traveltimes are roughly the same for the continuous and the discontinuous slowness models. The entries 'Error-int' in the Tables 2 and 4 show that the traveltime computed by numerical integration along the perturbed ray has the same accuracy for the continuous and the discontinuous slowness models. (The error is about 20 times smaller than the error in the first-order traveltime.) This example shows that plane discontinuities in the slowness model can be incorporated in the ray perturbation theory.

13 DISCUSSION

The ray perturbation method presented here provides an efficient method for computing the change in ray positions and traveltimes when the slowness model is perturbed. The theory can be applied when the slowness perturbation is sufficiently smooth. The conditions (67) and (68) provide simple criteria for the applicability of the theory. The equation for the ray perturbation can be applied both to initial value ray tracing and for two-point ray tracing. Initial value ray tracing is used in the computation of paraxial rays (e.g. Červeńy, Klimeš & Pšenčík 1984). The paraxial rays obtained can be used for the computation of amplitudes (e.g. Farra & Madariaga 1987). The main application of two-point ray tracing is non-linear traveltime tomography, where the theory can be used to update the rays with the slowness model during an inversion.

The ray perturbation equation (50) was derived from Lagrangian perturbation theory formulated in ray coordinates. One can show that for non-rotating unit vectors ($\Omega = 0$) expression (50) is equivalent to the perturbation equations derived by Farra & Madariaga (1987) using a Hamiltonian formalism. This can most easily be seen by using the explicit form of their Hamiltonian perturbation equations given by Nowack & Lutter (1988). [The derivation of this result involves a non-trivial differentiation of the geometric term h_σ in the Hamiltonian of Farra & Madariaga (1987).] One can incorporate unit vectors $\hat{\mathbf{q}}_i$ that rotate around the reference ray in the theory of Farra & Madariaga (1987) by adding a term $\Omega\varepsilon_{ij}p_iq_j$ to their Hamiltonian (4) (Farra, personal communication). Alternatively, one can derive equation (50) by projecting the Hamiltonian perturbation equations for a Cartesian coordinate system (Chapman 1985; Virieux 1991) on the ray coordinate system $(\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2)$ using the transformation equations shown in Section 4. Equation (50) is not equivalent to the perturbation equation (18) derived by Moore (1991). This discrepancy is due to the fact that Moore (1991) did not properly handle the change in the ray length due to the slowness perturbation. Note finally that the differential equation (50) closely resembles equation (C22) of Moser *et al.* (1992) for linearized ray bending where the slowness is fixed and where one seeks the perturbation that one needs to make to a reference curve to deform it towards the true ray.

In contrast to the perturbation schemes for ray perturbation presented by Farra *et al.* (1989), Virieux (1991) and Moore (1991), the theory presented here employs ray coordinates. Using ray coordinates reduces the number of unknowns at the expense of some additional bookkeeping. However, a major advantage of using ray coordinates is that when the ray perturbation is perpendicular to the unperturbed ray one can obtain in a natural fashion an expression for the travel that is correct to second order in the slowness perturbation from the first-order ray deflection (see 20b). This property is lost when one uses Cartesian coordinates for the ray perturbation and one has a component in the ray perturbation parallel to the reference ray because this invalidates the condition (15). The fact that for the second-order traveltime perturbation one only needs the first-order ray deflection relies on an extension of Fermat's theorem, which states that the traveltime is stationary in the ray perturbation, provided the endpoints of the ray are perturbed along the wavefront of the reference ray (condition 15). This condition is trivially satisfied when applied to two-point ray tracing. For paraxial rays it can be satisfied provided one ensures that the ray perturbation is perpendicular to the reference ray. With the ray coordinates used here this condition is trivially satisfied. The rotation rate Ω of the employed ray coordinates gives an additional degree of freedom in the formulation

of ray perturbation problem. This may be convenient for the generalization of the theory to weakly anisotropic media where it is crucial to incorporate the polarization of the waves in the theory.

The shooting algorithm used in Sections 9 and 12 was designed to be efficient and was optimized. The computations for the ray perturbation were obtained from (75) by using a simple and straightforward discretization of this equation. Without making any attempts to optimize the algorithm, the computations for the ray perturbation require two to three orders less CPU time than the employed shooting method of Sambridge & Kennett (1990). One can also compute the ray deflection and the traveltime perturbation from a Green's function (see 53 and 62). By computing this Green's function once and storing the result in tabulated form one may speed up the computations with several orders of magnitude. This may make it possible to apply the theory to mantle tomography where over a million rays are being used (e.g. Spakman 1990).

ACKNOWLEDGMENTS

Discussions with Jean Virieux and Robert Nowack have been helpful and illuminating. The critical and constructive comments of Veronique Farra are greatly appreciated.

REFERENCES

Aki, K. & Richards, P. G., 1980. *Quantitative Seismology*, vol. 1, Freeman, San Francisco.
Aki, K., Christofferson, A. & Husebye, E. S., 1977. Determination of the three-dimensional structure of the lithosphere, *J. geophys. Res.*, **82**, 277–296.
Buland, R. & Chapman, C. H., 1983. The computation of seismic travel times, *Bull. seism. Soc. Am.*, **68**, 1577–1593.
Bullen, K. E. & Bolt, B. A., 1985. *An Introduction to the Theory of Seismology*, 4th edn, Cambridge University Press, Cambridge, UK.
Červený, V., 1987. Ray tracing algorithms in three-dimensional laterally varying layered structures, in *Seismic Tomography, with Applications in Global Seismology and Exploration Geophysics*, ed. Nolet, G., Reidel, Dordrecht.
Červený, V., Klimeš, L. & Pšenčík, I., 1984. Paraxial ray approximations in the computation of seismic wavefields in inhomogeneous media, *Geophys. J. R. astr. Soc.*, **79**, 89–104.
Chapman, C. H., 1985. Ray theory and its extensions: WKBJ and Maslov seismograms, *J. Geophys.*, **58**, 27–43.
Dziewonski, A. M. & Anderson, D. L., 1981. Preliminary Reference Earth Model, *Phys. Earth. planet Inter.*, **25**, 297–356.
Farra, V. & Madariaga, R., 1987. Seismic waveform modelling in heterogeneous media by ray perturbation theory, *J. geophys. Res.*, **92**, 2697–2712.
Farra, V., Virieux, J. & Madariaga, R., 1989. Ray perturbation theory for interfaces, *Geophys. J. Int.*, **99**, 377–390.
Frankel, A. & Clayton, R. W., 1986. Finite difference simulations of seismic scattering: Implications for the propagation of short-period seismic waves in the crust and models of crustal heterogeneity, *J. geophys. Res.* **91**, 6465–6489.
Gerver, M. & Markushevich, M., 1966. Determination of a seismic wave velocity from the travel time curve, *Geophys. J. R. astr. Soc.*, **11**, 165–173 1990.
Gudmundsson, O., Davies, J. H. & Clayton, R. W., 1990. Stochastic analysis of global traveltime data: mantle heterogeneity and random errors in the ISC data, *Geophys. J. Int.*, **102**, 25–43.
Jeffreys, H. & Bullen, K. E., 1940. *Seismological Tables*, Brit. Ass., Grey-Milne Trust.
Julian, B. R. & Gubbins, D., 1977. Three-dimensional seismic ray tracing, *J. Geophys.*, **43**, 95–113.
Moore, B. J., 1991. Seismic rays in media with slight variation in velocity, *Geophys. J. Int.*, **105**, 213–227.
Moser, T. J., 1991. Shortest path calculations of seismic rays. *Geophysics*, **56**, 56–67.
Moser, T. J., Nolet, G. & Snieder, R., 1992. Ray bending revisited, *Bull. seism. Soc. Am.*, in press.
Nolet, G., 1987. Seismic wave propagation and seismic tomography, in *Seismic Tomography, with Applications in Global Seismology and Exploration Geophysics*, ed. Nolet, G., Reidel, Dordrecht.
Nowack, R. L. & Lutter, W. J., 1988. Linearized rays, amplitude and inversion, *Pageoph*, **128**, 401–421.
Pereyra, V., Lee, W. H. K. & Keller, H. B., 1980. Solving two-point seismic ray tracing problems in a heterogeneous medium, part 1, A general adaptive finite difference method, *Bull. seism. Soc. Am.*, **70**, 79–99.
Press, A. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T., 1986. *Numerical Recipes*, Cambridge University Press, Cambridge, UK.
Sambridge, M. S. & Kennett, B. L. N., 1990. Boundary value ray tracing in a heterogeneous medium; a simple and versatile algorithm, *Geophys. J. Int.*, **101**, 157–168.
Snieder, R., 1990. A perturbative analysis of non-linear inversion, *Geophys. J. Int.*, **101**, 545–556.
Snieder, R., 1991. An extension of Backus–Gilbert Theory to nonlinear inverse problems, *Inverse Problems*, **7**, 409–433.
Spakman, W., 1990. Tomographic images of the upper mantle below central Europe and the Mediterranean, *Terra Nova*, **2**, 542–553.
Vidale, J., 1988. Finite-difference calculation of travel times, *Bull. seism. Soc. Am.*, **78**, 2062–2076.
Virieux, J., 1991. Fast and accurate ray tracing by Hamiltonian perturbation, *J. geophys. Res.*, **96**, 579–594.
Wunsch, C., 1987. Acoustic tomography by Hamiltonian methods including the adiabatic approximation, *Rev. Geophys.* **25**, 41–53.

APPENDIX A

The relation between the expressions (40a,b) and the Frenet equations

In the notation of Aki & Richards (1980) the Frenet equations for the unit vectors perpendicular to the ray are given by

$$\dot{\hat{\mathbf{n}}} = -\kappa\dot{\mathbf{r}}_0 + \tau\hat{\mathbf{b}}, \qquad \dot{\hat{\mathbf{b}}} = -\tau\hat{\mathbf{n}}. \tag{A1}$$

In this expression $\hat{\mathbf{n}}$ is the unit vector perpendicular to the reference ray in the osculating plane of the ray. The unit vector $\hat{\mathbf{b}}$ is perpendicular to both $\hat{\mathbf{n}}$ and $\dot{\mathbf{r}}_0$. Since the two bases $(\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2)$ and $(\hat{\mathbf{n}}, \hat{\mathbf{b}})$ are both perpendicular to the reference ray, they are related through a rotation:

$$\hat{\mathbf{q}}_1 = \cos\phi\hat{\mathbf{n}} + \sin\phi\hat{\mathbf{b}}, \tag{A2a}$$

$$\hat{\mathbf{q}}_2 = -\sin\phi\hat{\mathbf{n}} + \cos\phi\hat{\mathbf{b}}, \tag{A2b}$$

where the angle of rotation may depend on the position along the reference ray $[\phi = \phi(s_0)]$. Differentiation of (A2) with respect to s_0 gives with (A1)

$$\dot{\hat{\mathbf{q}}}_1 = (T + \dot{\phi})\hat{\mathbf{q}}_2 - \cos\phi\kappa\dot{\mathbf{r}}_0, \tag{A3a}$$

$$\dot{\hat{\mathbf{q}}}_2 = -(T + \dot{\phi})\hat{\mathbf{q}}_1 + \sin\phi\kappa\dot{\mathbf{r}}_0. \tag{A3b}$$

A comparison with (43) shows that

$$\Omega = T + \dot{\phi}; \tag{A4}$$

the rate of rotation of the $\hat{\mathbf{q}}$-vectors (Ω) is the rate of rotation of the vectors in the Frenet equation (T) plus the relative rate of rotation between the two coordinate systems ($\dot{\phi}$). The Frenet equations constitute a special choice for the coordinate system $(\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2)$ where these unit vectors are aligned with the osculating plane of the reference ray.

The curvature κ follows by comparing (A3a,b) with (40a,b), this gives

$$\cos\phi\kappa = \frac{1}{u_0}(\hat{\mathbf{q}}_1 \cdot \nabla u_0), \tag{A5a}$$

$$\sin\phi\kappa = \frac{-1}{u_0}(\hat{\mathbf{q}}_2 \cdot \nabla u_0). \tag{A5b}$$

Using (A2a,b) it follows from this expression that

$$\kappa = \frac{1}{u_0}[(\cos\phi\hat{\mathbf{q}}_1 - \sin\phi\hat{\mathbf{q}}_2) \cdot \nabla u_0] = \frac{1}{u_0}(\hat{\mathbf{n}} \cdot \nabla u_0). \tag{A6}$$

APPENDIX B

A proof that the perturbed ray (97) satisfies Snell's law to first order

Referring to Fig. 7, the angle between the perturbed ray and the reference ray in the upper half-space satisfies

$$\tan\theta_1 = \varepsilon\left(\frac{dq}{ds_0}\right)_{s_0<s_i} = \varepsilon\left(1 - \frac{s_i}{S_0}\right)\tan i_0, \tag{B1}$$

hence

$$i_1 = i_0 - \theta_1 = i_0 - \arctan\left[\varepsilon\left(1 - \frac{s_i}{S_0}\right)\tan i_0\right]. \tag{B2}$$

This means that

$$u_0 \sin i_1 = u_0 \sin i_0 - \varepsilon u_0\left(1 - \frac{s_i}{S_0}\right)\sin i_0 + O(\varepsilon^2). \tag{B3}$$

Likewise, one has for the lower half-space

$$\tan\theta_2 = -\varepsilon\left(\frac{dq}{ds_0}\right)_{s_0>s_i} = \varepsilon\frac{s_i}{S_0}\tan i_0, \tag{B4}$$

and

$$i_2 = i_0 + \theta_2 = i_0 + \arctan\left(\varepsilon\frac{s_i}{S_0}\tan i_0\right), \tag{B5}$$

so that

$$u_0(1-\varepsilon)\sin i_2 = u_0(1-\varepsilon)\left(\sin i_0 + \varepsilon\frac{s_i}{S_0}\sin i_0\right) = u_0\sin i_0 - \varepsilon u_0\left(1 - \frac{s_i}{S_0}\right)\sin i_0 + O(\varepsilon^2). \tag{B6}$$

Comparing this with (B3) proves that Snell's law is satisfied to first order:

$$u_0 \sin i_1 = u_0(1-\varepsilon)\sin i_2 + O(\varepsilon^2). \tag{B7}$$

APPENDIX C

Derivation of the second-order traveltime perturbation for the two-layer medium

Let $T(x)$ denote the traveltime of a ray which intersects the slowness discontinuity in the point $(x, 0)$ (see Fig. 9). This quantity is given by

$$T(x) = u_0\left(x^2 + \frac{D^2}{4}\right)^{1/2} + u_0(1-\varepsilon)\left((X-x)^2 + \frac{D^2}{4}\right)^{1/2}. \tag{C1}$$

For the ray the traveltime is stationary; the condition $\partial T/\partial x = 0$ gives

$$(-2\varepsilon+\varepsilon^2)(x-X)^2x^2 + [(1-\varepsilon)^2(x-X)^2 - x^2]\frac{D^2}{4} = 0. \tag{C2}$$

The point of intersection of the ray with the slowness discontinuity can be expanded in a perturbation series

$$x = x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \cdots. \tag{C3}$$

Inserting this in (C2), and equating the contribution of $O(\varepsilon^0)$ and $O(\varepsilon^1)$ gives

$$(2x_0 - X)X\frac{D^2}{4} = 0, \tag{C4a}$$

$$(x_0 - X)^2 x_0^2 + [(x_0-X)^2 + x_1 X]\frac{D^2}{4} = 0. \tag{C4b}$$

These equations have the solution

$$x_0 = X/2, \tag{C5a}$$

$$x_1 = -\frac{X}{4}\frac{S_0^2}{D^2} \quad \text{with} \quad S_0^2 = X^2 + D^2. \tag{C5b}$$

In order to obtain a second-order perturbation expansion for the traveltime, insert (C3) with the solutions (C5a,b) in (C1). The second-order deflection terms x_2 cancel in the resulting expression. This is analogous to the situation in Section 2, where the second-order terms $\mathbf{r}_2$ in expression (13c) for the traveltime cancel. It is for this reason that the solution x_2 is not computed in the equations (C4) and (C5). The resulting perturbation expansion of the traveltime is given by

$$T = T_0 - \frac{\varepsilon}{2}T_0 - \frac{\varepsilon^2}{8}\frac{X^2}{D^2}T_0 + O(\varepsilon^3), \tag{C6}$$

with $T_0 = u_0 S_0$.

A comparison with equation (99) reveals that the second-order traveltime perturbation computed from the ray perturbation theory applied to discontinuous slowness perturbations leads to the correct second-order traveltime perturbation.