Geophys. J. Int. (1989) **99**, 377–390

Ray perturbation theory for interfaces

Véronique Farra*,† Jean Virieux* and Raúl Madariaga*

* *Laboratoire de Sismologie, Institut de Physique du Globe, 4 Place Jussieu, Tour 14, 75252 Paris Cedex 05, France*

† *Institut Français du Pétrole, 1–4 Avenue du Bois Préau, 92506 Rueil-Malmaison, France*

Accepted 1989 March 5. Received 1989 January 15; in original form 1988 October 3.

SUMMARY

We propose a new formalism for the calculation of perturbations of ray trajectories and amplitudes in laterally heterogeneous medium. A Hamiltonian technique leads to a unified approach for the calculation of paraxial rays and rays perturbed by small changes of velocity distribution and interface shape. Instead of using ray centred coordinates as in the classical approach to dynamic ray tracing, we use straightforward cartesian coordinates. This has the advantage that paraxial rays may be referred to the unperturbed ray in a very flexible way. We first study perturbation of initial conditions or paraxial ray tracing. With this technique an ensemble of rays propagating in the vicinity of a central ray is traced with the help of the so-called paraxial ray propagator. This ray propagator is the basis of all the techniques discussed in this paper. Its efficient determination is discussed and we propose a finite element approach in which the medium is divided into a set of trapezoidal elements with simple velocity distribution. We propose that the simpler results are obtained when a constant gradient of the square of the slowness is adopted in each element. In the second part of the paper we calculate the effect of perturbations of the velocity structure and interfaces upon ray trajectories, amplitudes and waveforms. Our results can be easily adapted for the calculation of Fréchet derivatives for the linearized inversion of travel times, amplitudes and waveforms. Finally, we present an example of the calculation of synthetic seismograms in a simple medium with a perturbed interface. Simplified expressions for the calculation of perturbed rays in a few typical reference media are given.

Key words: amplitude, Hamiltonian formulation, perturbation, ray tracing

1 INTRODUCTION

The paraxial ray method is a technique to approximately trace rays in the vicinity of a given reference (central) ray, by a first-order perturbation technique. These paraxial rays are an essential ingredient of many applications of ray theory to seismological problems, for instance, in the computation of ray amplitudes (Popov & Pšenčík 1978), in Gaussian beam summation (Popov 1982; Červený, Popov & Pšenčík 1982) or in Maslov's method as proposed by Chapman & Drummond (1982) or Thomson & Chapman (1985). They are also very useful for solving two-point ray tracing, as well as for interpolating travel times.

When the central ray hits a discontinuity of zeroth- or first-order, the paraxial rays have to satisfy specific continuity conditions. Two equivalent methods have been used to derive these continuity conditions: the phase matching method (see Červený 1985) and the perturbation of the Snell's law (see Chapman 1985). In the former technique the phase, expanded up to second-order terms, is matched across the discontinuity. The first-order term gives Snell's law at the hitting point, while the second-order term gives the paraxial ray continuity conditions. In Chapman's (1985) approach the perturbation of Snell's law to first order gives the paraxial ray continuity conditions.

The purpose of this paper is to develop the paraxial ray continuity conditions when the ray field is affected by small perturbations in the velocity (or slowness) and interface shapes. The problem of slowness perturbation has already been studied by Farra & Madariaga (1987) in generalized coordinates; Farra (1987) solved in her thesis the problem of interface perturbations in generalized coordinates. Generalized coordinates were required because paraxial rays have traditionally been traced using the ray centred coordinate system proposed by Popov & Pšenčík (1978). The solution of continuity conditions in this curvilinear coordinate system is extremely complex, requiring several canonical transformations between ray centred and interface centred coordinates. In Virieux, Farra & Madariaga (1988) we proposed that ray and paraxial ray tracing be performed directly in Cartesian coordinates. This increases the size of the ray tracing equation system but simplifies considerably the calculations. We are presently convinced that coupled with a finite element discretization of the slowness field, this approach is the most efficient for the solution of ray and paraxial ray problems.

In this paper we will focus our attention on the construction of paraxial boundary conditions in global Cartesian coordinates, and how to perturb them when the shape of interfaces is modified slightly. The solution of this problem may be expressed in terms of Fréchet derivatives and has numerous applications. For instance in the study of the effect of small lateral heterogeneity on travel times, ray amplitudes, in the perturbation of two-point ray tracing by interface changes, and in non-linear tomography of interfaces (Nowack & Lyslo 1989). In the perturbed medium, central rays, as well as their paraxial rays, will be obtained by first-order perturbation of Snell's law at the interfaces. Two specific reference media of interest in seismology will be explicitly worked out. Illustrations on a simple synthetic example will hopefully demonstrate the possibilities of the method.

2 RAY AND PARAXIAL RAY THEORY FOR ELASTIC MEDIA

Let us briefly recall the Hamiltonian formulation used by Virieux *et al.* (1988). In ray theory we assume that the high-frequency asymptotic form of a scalar or vector wave field $\phi(\mathbf{x}, \omega)$ is:

$$\phi(\mathbf{x}, \omega) = s(\omega)A(\mathbf{x})e^{i\omega\theta(\mathbf{x})}, \tag{1}$$

where $A(\mathbf{x})$ is the first term of the expansion of $A(\mathbf{x}, \omega)$ in inverse powers of ω, $\theta(\mathbf{x})$ is the eikonal or traveltime function, ω the circular frequency and $s(\omega)$ the source function. From the wave equation, one obtains the eikonal equation $(\nabla\theta)^2 = u^2(\mathbf{x})$, where u^2 is the square of the slowness: compressional slowness for P-waves, shear slowness for S-waves (see, e.g. Červený, Molotkov & Pšenčík 1977). In order to perform ray tracing, we introduce the slowness vector $\mathbf{p} = \nabla\theta$, which is perpendicular to the surfaces of equal phase θ, or wavefronts. Let s be the arclength and τ the sampling parameter along the ray defined by $u\,d\tau = ds$ (Chapman 1985). Let us underline that τ has units $\mathrm{km^2\,s^{-1}}$. Because the rays are everywhere tangent to the slowness vector, position along the ray $\mathbf{x}$ is related to the slowness vector by:

$$\mathbf{p} = u\frac{d\mathbf{x}}{ds} = \frac{d\mathbf{x}}{d\tau}. \tag{2}$$

Introducing the Hamiltonian proposed by Burridge (1976):

$$H(\mathbf{x}, \mathbf{p}, \tau) = \tfrac{1}{2}[\mathbf{p}^2 - u^2(\mathbf{x})] \tag{3}$$

we observe that the eikonal equation implies that $H \equiv 0$ along a ray. From Hamilton's canonical equations, we find the ray tracing equations:

$$\begin{aligned} \dot{\mathbf{x}} &= \nabla_{\mathbf{p}} H = \mathbf{p} \\ \dot{\mathbf{p}} &= -\nabla_{\mathbf{x}} H = \tfrac{1}{2}\nabla_{\mathbf{x}} u^2, \end{aligned} \tag{4}$$

where dot denotes differentiation with respect to τ; $\nabla_{\mathbf{x}}$ and $\nabla_{\mathbf{p}}$ denote the gradients with respect to the vectors $\mathbf{x}$ and $\mathbf{p}$, respectively. Let us recall (Červený *et al.* 1977) that the six equations in (4) are not independent since at least one of them may be eliminated by using the fact that $\mathbf{p}$ should satisfy the eikonal equation (3). System (4) may be reduced to four equations as shown by Farra & Madariaga (1987), but only at the expense of a more complicated curvilinear geometry. Our present feeling is that using Cartesian coordinates (Virieux *et al.* 1988) is much simpler than reducing the system.

Suppose a ray has been traced in the medium with slowness distribution $u_0(\mathbf{x})$. Around this ray, called the central ray, we can obtain neighbouring rays by means of first-order perturbation theory, as explained by Farra and Madariaga (1987). Let $\mathbf{x}_0(\tau)$ and $\mathbf{p}_0(\tau)$ be the position and the slowness vector of the central ray. For conciseness we will sometimes use the notation $\mathbf{y}_0(\tau) = [\mathbf{x}_0(\tau), \mathbf{p}_0(\tau)]$, the so-called *canonical* vector of the central ray. This is a 6-vector in *phase space*, the space of position and slowness. The position of a paraxial ray and its slowness vector are given by:

$$\mathbf{x}(\tau) = \mathbf{x}_0(\tau) + \delta\mathbf{x}(\tau) \qquad \mathbf{p}(\tau) = \mathbf{p}_0(\tau) + \delta\mathbf{p}(\tau). \tag{5}$$

The perturbation of position and slowness vector $\delta\mathbf{y}(\tau) = (\delta\mathbf{x}, \delta\mathbf{p})$ satisfies the paraxial ray tracing equations deduced from (4):

$$\delta\dot{\mathbf{y}} = A_0\,\delta\mathbf{y}, \tag{6}$$

with

$$A_0 = \begin{pmatrix} 0 & I \\ U_0 & 0 \end{pmatrix} \tag{7}$$

where I is the identity matrix and U_0 is the matrix of second-order partial derivatives of the square of slowness defined by:

$$U_{0_{ij}} = \frac{1}{2}\frac{\partial^2 u_0^2}{\partial x_i\,\partial x_j}. \tag{8}$$

Solutions to the linear system (6) may be found by standard propagator techniques. Given the initial value $\delta\mathbf{y}(\tau_0)$ the subsequent evolution of the canonical vector in phase space is given by:

$$\delta\mathbf{y}(\tau) = \mathscr{P}_0(\tau, \tau_0)\,\delta\mathbf{y}(\tau_0), \tag{9}$$

where $\mathscr{P}_0(\tau, \tau_0)$ is the propagator matrix of system (6).

Just as with the full non-linear ray tracing system (4), the six equations of the paraxial system (6) are not really independent. In other words, not every solution of the system (6) represents a paraxial ray trajectory. In fact, $\delta\mathbf{y}$ should satisfy a condition derived from the perturbation of the eikonal equation $H \equiv 0$. To first-order, when position and slowness are perturbed as in (5), the perturbation of the Hamiltonian should satisfy

$$\delta H(\tau) = \mathbf{p}_0 \cdot \delta\mathbf{p} - \tfrac{1}{2}\nabla_{\mathbf{x}} u_0^2 \cdot \delta\mathbf{x} = 0. \tag{10}$$

Since δH is constant along any solution of the system (6),

$$\delta H(\tau) = \delta H(\tau_0) \tag{11}$$

it is sufficient to enforce $\delta H \equiv 0$ at the source in order to satisfy (10) everywhere.

3 REFLECTION AND TRANSMISSION OF PARAXIAL RAYS

We define the inner product of two vectors $\mathbf{u}$ and $\mathbf{v}$ by $\langle\mathbf{u} \mid \mathbf{v}\rangle$ with $\langle\mathbf{u}|$ applied to $|\mathbf{v}\rangle$. An operator A applied to $|\mathbf{v}\rangle$

will give $A\,|\mathbf{v}\rangle$. This will allow us to use the operator notation in the following transformations.

Let us consider a reference ray whose canonical vector is $\mathbf{y}_0(\tau) = (\mathbf{x}_0, \mathbf{p}_0)$. This reference ray hits an internal velocity discontinuity at point O of coordinate $\mathbf{x}_0(\tau_i)$ with a local slowness vector $\mathbf{p}_0(\tau_i)$ (Fig. 1a). We consider a paraxial ray of this reference ray that intersects the same discontinuity at point I of position $\mathbf{x}(\tau_i')$, with sampling parameter τ_i' and local slowness $\mathbf{p}(\tau_i')$. We denote $d\mathbf{x} = \mathbf{x}(\tau_i') - \mathbf{x}_0(\tau_i)$, $d\mathbf{p} = \mathbf{p}(\tau_i') - \mathbf{p}_0(\tau_i)$ and $d\mathbf{y} = (d\mathbf{x}, d\mathbf{p})$. Paraxial rays are traced in the incidence medium using the perturbation relations (5). Let the paraxial canonical vector at sampling parameter τ_i be $\delta\mathbf{y}(\tau_i) = [\delta\mathbf{x}, \delta\mathbf{p}]$. Vector $\delta\mathbf{x}(\tau_i)$ defines the position of point Q in Fig. 1a. In general Q and I do not coincide so that $d\mathbf{y} \neq \delta\mathbf{y}$. Using the ray equations (4) and referring to the vector diagrams in Fig. 1(b) and (c) we find that $d\mathbf{y}$ and $\delta\mathbf{y}$ are related by;

$$\begin{aligned} d\mathbf{x} &= \delta\mathbf{x}(\tau_i) + \nabla_{\mathbf{p}} H\, d\tau \\ d\mathbf{p} &= \delta\mathbf{p}(\tau_i) - \nabla_{\mathbf{x}} H\, d\tau, \end{aligned} \tag{12}$$

where $d\tau = (\tau_i' - \tau_i)$ and the gradients of H are computed at $\mathbf{y}_0(\tau_i)$.

Let the interface be defined by the relation $f(\mathbf{x}) = 0$ and denote by ∇f_0 the local normal to the interface at point O (see Fig. 1). The condition that point I belongs to the interface $f[\mathbf{x}_0(\tau_i) + d\mathbf{x}] = 0$, yields $\langle d\mathbf{x} \mid \nabla f_0 \rangle = 0$ to first order in $d\mathbf{x}$. Taking the inner product of the first of equations (12) with ∇f_0 and imposing this condition yields the following expression for the sampling parameter increment $d\tau$:

$$d\tau = -\frac{\langle \nabla f_0 \mid \delta\mathbf{x} \rangle}{\langle \nabla f_0 \mid \nabla_{\mathbf{p}} H \rangle}, \tag{13}$$

which follows also directly from Fig. 1(b). The paraxial vector at point I with respect to the central ray at the intersection point O, $d\mathbf{y} = (d\mathbf{x}, d\mathbf{p})$, is obtained then as a linear transformation Π of $\delta\mathbf{y}(\tau_i)$:

$$d\mathbf{y} = \Pi\, \delta\mathbf{y}(\tau_i), \tag{14}$$

with:

$$\Pi = \begin{bmatrix} \pi_1 & 0 \\ \pi_2 & I \end{bmatrix}, \tag{15}$$

where the submatrices are given by:

$$\pi_1 = I - \frac{|\nabla_{\mathbf{p}} H\rangle \langle \nabla f_0|}{\langle \nabla f_0 \mid \nabla_{\mathbf{p}} H \rangle} \qquad \pi_2 = \frac{|\nabla_{\mathbf{x}} H\rangle \langle \nabla f_0|}{\langle \nabla f_0 \mid \nabla_{\mathbf{p}} H \rangle}. \tag{16}$$

The notation $|1\rangle\langle 2|$ represents a matrix obtained by the tensor product of the vectors $\langle 1|$ and $\langle 2|$. For the corresponding formulation in generalized coordinates, we refer the reader to Farra (1987).

3.1 Continuity conditions for paraxial rays

Let us now construct the continuity conditions for paraxial rays across the interface. We will denote variables in the reflected/transmitted medium with a caret above them. The new Hamiltonian will be, for example, $\hat{H}$. The continuity of position of paraxial rays at the interface gives the following

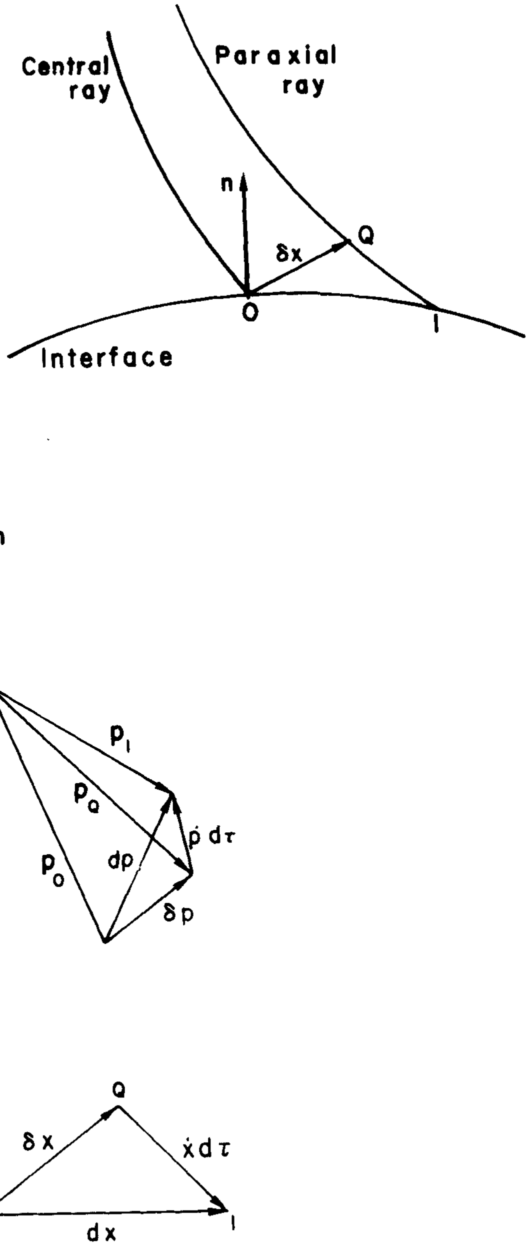

Figure 1. Geometry of the interaction of a ray and one of its paraxial rays with an interface. The central ray intersects the interface at O with slowness vector $\mathbf{p}_O = \mathbf{p}_0(\tau_i)$, while the paraxial ray arrives at I with slowness vector $\mathbf{p}_I = \mathbf{p}(\tau_i')$. At the interface, paraxial vectors $\delta\mathbf{x}$ and $\delta\mathbf{p}$ have to be transformed into $d\mathbf{x}$ and $d\mathbf{p}$, respectively. $\mathbf{n} = \nabla f_0$ is the local normal to the interface at point O.

simple relation:

$$d\hat{\mathbf{x}} = d\mathbf{x}. \tag{17}$$

The continuity condition for slowness perturbation is more difficult to obtain because we have to take into account the local curvature of the interface. We need three conditions to continue the vector $d\mathbf{p}$ into $d\hat{\mathbf{p}}$. A first relation comes from the continuity of $dH = d\hat{H}$ at the interface:

$$\langle \nabla_{\mathbf{p}} \hat{H} \mid d\hat{\mathbf{p}} \rangle + \langle \nabla_{\mathbf{x}} \hat{H} \mid d\hat{\mathbf{x}} \rangle = \langle \nabla_{\mathbf{p}} H \mid d\mathbf{p} \rangle + \langle \nabla_{\mathbf{x}} H \mid d\mathbf{x} \rangle. \tag{18}$$

This relation is an extension of (11) to a medium with zero- or first-order discontinuities. The two other conditions come from the local perturbation of Snell's law along the interface. In vector notation, Snell's law at point I is:

$$(\hat{\mathbf{p}}_0 + d\hat{\mathbf{p}}) \times \nabla f = (\mathbf{p}_0 + d\mathbf{p}) \times \nabla f \tag{19}$$

where the cross-product has been noted by $\times$. The normal to the interface at point I (Fig. 1), is given to first-order by:

$$\nabla f = \nabla f_0 + \nabla\nabla f_0 \,|d\mathbf{x}\rangle, \tag{20}$$

where $\nabla\nabla f_0$ stands for the matrix of second derivatives (curvature) of the interface at $\mathbf{x}_0(\tau_i)$. Inserting (20) in (19), we obtain to first-order:

$$d\hat{\mathbf{p}} \times \nabla f_0 = d\mathbf{p} \times \nabla f_0 + (\mathbf{p}_0 - \hat{\mathbf{p}}_0) \times (\nabla\nabla f_0 \,|d\mathbf{x}\rangle). \tag{21}$$

We are now ready to obtain $d\hat{\mathbf{p}}$ from $d\mathbf{p}$ and $d\mathbf{x}$ using relations (17), (18) and (21). Let us develop $d\hat{\mathbf{p}}$ along the normal and the tangent plane to $f(\mathbf{x}) = 0$ at the point O (Fig. 1):

$$d\hat{\mathbf{p}} = \frac{\langle d\hat{\mathbf{p}} \,|\, \nabla f_0\rangle \nabla f_0 - (d\hat{\mathbf{p}} \times \nabla f_0) \times \nabla f_0}{\langle \nabla f_0 \,|\, \nabla f_0\rangle}. \tag{22}$$

Inserting (22) into (18) and using (17) we may express the inner product $\langle d\hat{\mathbf{p}} \,|\, \nabla f_0\rangle$ in terms of $d\mathbf{x}$ and the cross-product $d\hat{\mathbf{p}} \times \nabla f_0$. The cross-product can in turn be deduced from (21) in terms of $d\mathbf{p}$ and $d\mathbf{x}$.

After some heavy but straightforward algebra, we express the difference $d\hat{\mathbf{p}} - d\mathbf{p}$ as the sum of a term along the normal ∇f_0 and another one parallel to the tangent plane:

$$d\hat{\mathbf{p}} = d\mathbf{p} - a \frac{\nabla f_0}{\langle \nabla_{\mathbf{p}} \hat{H} \,|\, \nabla f_0\rangle} - b \nabla\nabla f_0 \,|d\mathbf{x}\rangle \tag{23}$$

with

$$\begin{aligned} a &= \langle \nabla_{\mathbf{x}} \hat{H} - \nabla_{\mathbf{x}} H \,|\, d\mathbf{x}\rangle + \langle \nabla_{\mathbf{p}} \hat{H} - \nabla_{\mathbf{p}} H \,|\, d\mathbf{p}\rangle \\ &\quad - \frac{\langle \nabla_{\mathbf{p}} H - \nabla_{\mathbf{p}} \hat{H} \,|\, \nabla f_0\rangle}{\langle \nabla f_0 \,|\, \nabla f_0\rangle} \langle \nabla_{\mathbf{p}} \hat{H} |\, \nabla\nabla f_0 \,|d\mathbf{x}\rangle \\ b &= \frac{\langle \nabla_{\mathbf{p}} H - \nabla_{\mathbf{p}} \hat{H} \,|\, \nabla f_0\rangle}{\langle \nabla f_0 \,|\, \nabla f_0\rangle}. \end{aligned} \tag{24}$$

Finally the continuity condition for the canonical paraxial vectors $d\hat{\mathbf{y}}$ and $d\mathbf{y}$ across the interface may be expressed in terms of the transformation matrix T defined by:

$$d\hat{\mathbf{y}} = T\, d\mathbf{y} \tag{25}$$

and which may be written in the form:

$$T = \begin{bmatrix} I & 0 \\ T_1 & T_2 \end{bmatrix} \tag{26}$$

where the submatrices T_1 and T_2 are given by:

$$\begin{aligned} T_1 &= -\frac{|\nabla f_0\rangle \langle \nabla_{\mathbf{x}} \hat{H} - \nabla_{\mathbf{x}} H|}{\langle \nabla_{\mathbf{p}} \hat{H} \,|\, \nabla f_0\rangle} \\ &\quad - \frac{\langle \nabla_{\mathbf{p}} H - \nabla_{\mathbf{p}} \hat{H} \,|\, \nabla f_0\rangle}{\langle \nabla f_0 \,|\, \nabla f_0\rangle} \left\{ I - \frac{|\nabla f_0\rangle \langle \nabla_{\mathbf{p}} \hat{H}|}{\langle \nabla_{\mathbf{p}} \hat{H} \,|\, \nabla f_0\rangle} \right\} \nabla\nabla f_0 \end{aligned} \tag{27}$$

and

$$T_2 = I - \frac{|\nabla f_0\rangle \langle \nabla_{\mathbf{p}} \hat{H} - \nabla_{\mathbf{p}} H|}{\langle \nabla_{\mathbf{p}} \hat{H} \,|\, \nabla f_0\rangle}.$$

All the quantities appearing in (27) are calculated on the reference ray. Without any further transformation the canonical perturbation vector $d\hat{\mathbf{y}}$ may be used as the new initial condition $\delta\hat{\mathbf{y}}$ to propagate the reflected-transmitted paraxial ray in the new medium. Therefore, the complete transformation of the paraxial ray vectors at the discontinuity is given by:

$$\delta\hat{\mathbf{y}} = T\Pi\, \delta\mathbf{y}(\tau_i) \tag{28}$$

with

$$T\Pi = \begin{pmatrix} \pi_1 & 0 \\ T_1\pi_1 + T_2\pi_2 & T_2 \end{pmatrix}. \tag{29}$$

When this transformation is written in local Cartesian coordinates with axis z along the normal to the discontinuity, we recover the results of Virieux *et al.* (1988) obtained with the phase matching procedure. It is worth noting that Červený, Langer & Pšenčík (1974) have already introduced similar transformations at interfaces using differential geometry approach.

3.2 An example of paraxial ray tracing

In order to illustrate paraxial ray tracing in Cartesian coordinates, we consider two media with constant vertical velocity gradient, separated by a curved interface interpolated by B-splines. The expression of the velocity is $(4.0 + 0.004z)\ \mathrm{km\,s^{-1}}$ in the upper layer and $(4.5 + 0.1z)\ \mathrm{km\,s^{-1}}$ in the lower medium. Fig. 2 presents a reference ray traced by a Runge–Kutta solver, as well as one of its paraxial rays for a point source. The paraxial vectors $\delta\mathbf{x}$ are explicitly drawn at several positions along the ray. We remark that unlike in ray-centred coordinates (e.g. Červený 1985) the paraxial vectors $\delta\mathbf{x}$ are not required to be perpendicular to the central ray. When the reference ray intersects the boundary, the transformation Π is activated in order to obtain $d\mathbf{x}$ tangent to the interface. The transformation Π affects both $\delta\mathbf{x}$ and $\delta\mathbf{p}$ and is consistent with the eikonal equation, i.e. it is a linearized canonical transformation in the language of analytical mechanics (Goldstein 1980). Now that $d\mathbf{x}$ is parallel to the interface we apply the transformation T in order to obtain the initial perturbation vector $\delta\mathbf{y}$ for the propagation of the paraxial ray transmitted into the lower medium. Transformation T affects only the slowness perturbation, not $d\mathbf{x}$. Actually, it contains the perturbation of the take-off angle of the paraxial ray from the interface. Because of the complex geometry of the interface the ray passes through a caustic at an offset of 7.2 km. At that point the paraxial vector crosses the central ray and, as expected, $\delta\mathbf{x}$ changes sign. Finally the ray is reflected back to the surface. Fig. 2 illustrates several of the transformations that have to be introduced at interfaces in order to continue paraxial rays across interfaces.

4 PERTURBATION OF RAYS

In this section we will develop ray perturbation theory for central rays when either the medium or the interfaces are slightly perturbed from their values in the reference medium. Farra & Madariaga (1987) presented ray perturbation theory for slowness modification in orthogonal

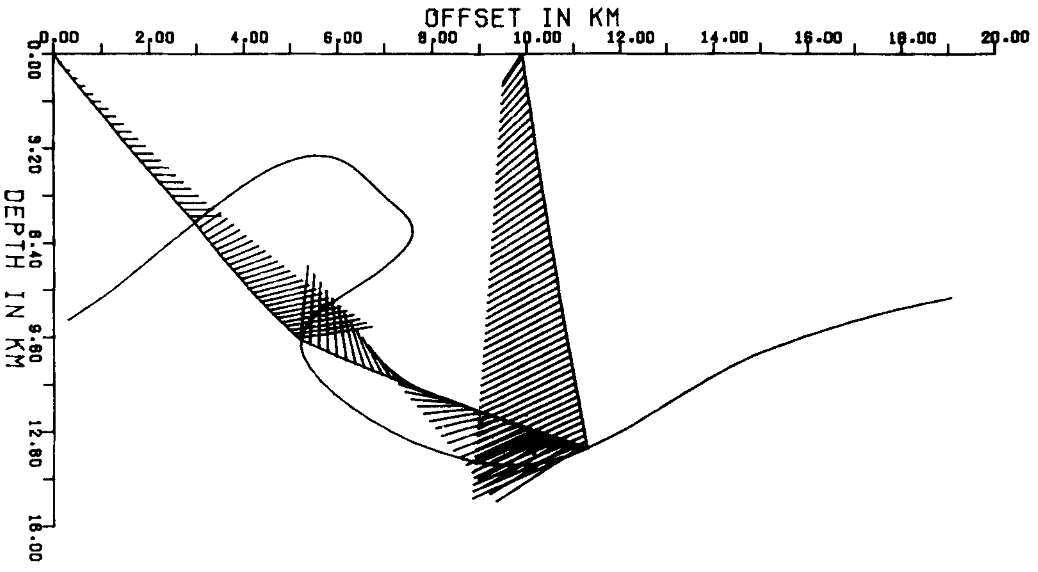

Figure 2. Illustration of paraxial rays and their transformation at the intersection with an interface. A paraxial ray is obtained by drawing its vector $\delta\mathbf{x}$ from the central ray. Every time the ray crosses an interface the paraxial ray $\delta\mathbf{x}$ and the associated slowness vector $\delta\mathbf{p}$ have to be redefined by a transformation defined by (28). The ray goes through a caustic in its third leg. A final transformation of the paraxial ray has to be performed at the free surface in order to get a horizontal paraxial vector $d\mathbf{x}$.

curvilinear coordinates. This formalism is necessary when using the ray-centred coordinates of Popov & Pšenčík (1978). As proposed by Virieux *et al.* (1988) it is very likely that using Cartesian coordinates both for ray tracing and paraxial ray tracing simplifies considerably the calculation of the effect of interfaces. In this section we will briefly adapt Farra & Madariaga (1987) results for Cartesian coordinates and we will then tackle the interface perturbation problem. A complete treatment of interfaces in curvilinear coordinates may be found in Farra (1987); the expressions are so unwieldy that we prefer to present the much simpler results in Cartesian coordinates here. To our knowledge Nowack & Lyslo (1988) are the only authors that have studied the effect of interface perturbation on central rays in a seismological context. Our results will extend theirs to arbitrary gradients in the media and, in a later section, to paraxial rays.

4.1 Ray perturbation due to slowness change

Let us consider, as in Fig. 3, a smooth perturbation of the model such that the slowness is slightly changed from u_0 to $u = u_0 + \Delta u$. The perturbation in slowness produces a corresponding perturbation of the Hamiltonian: $H = H_0 + \Delta H$, where H_0 is the Hamiltonian (3) for the reference slowness $u_0(\mathbf{x})$ and $\Delta H = -\frac{1}{2}\Delta u^2 = -u_0\,\Delta u$.

We assume that a ray has already been traced in the reference medium with unperturbed slowness distribution $u_0(x)$. To first-order in Δu, it is possible to obtain rays of the perturbed medium that deviate slightly from this reference ray. Following Farra & Madariaga (1987), we introduce the perturbed canonical vector $\mathbf{y}(\tau) = \mathbf{y}_0(\tau) + \Delta\mathbf{y}(\tau)$ of these rays as defined in Fig. 3. Inserting it and the slowness perturbation in the ray equations (4), we get:

$$\Delta\dot{\mathbf{y}} = A_0\,\Delta\mathbf{y} + \Delta B, \tag{30}$$

where

$$\Delta B = \begin{pmatrix} 0 \\ \frac{1}{2}\nabla(\Delta u^2) \end{pmatrix} \tag{31}$$

and all the derivatives are calculated on the reference ray. Equations (30) form a linear system which has the same form as that of paraxial rays in the unperturbed medium (6), except for the source term ΔB derived from $\Delta u^2(\mathbf{x})$.

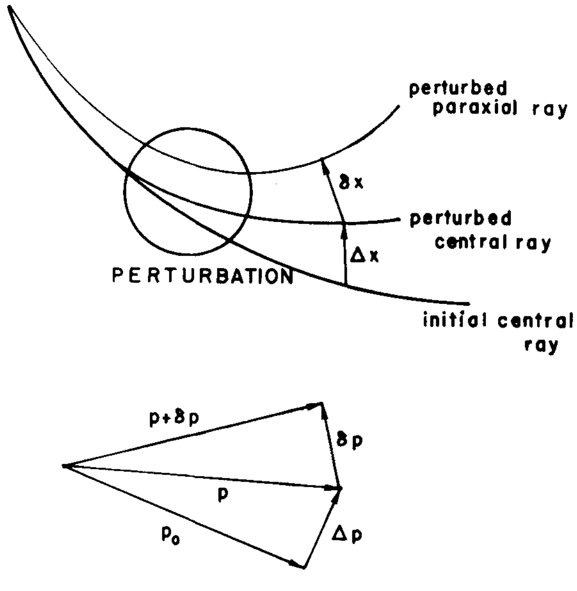

Figure 3. Geometry of the effect of a smooth slowness perturbation upon a ray and one of its paraxial rays. At the top we show the ray geometry and at the bottom the perturbation in slowness vector. The perturbation of the central ray is given by $\Delta\mathbf{x}$, $\Delta\mathbf{p}$, the position of a paraxial ray with respect to the *perturbed central ray* is given by $\delta\mathbf{x}$, $\delta\mathbf{p}$.

Solutions to (30) may be found by standard propagator techniques:

$$\Delta\mathbf{y}(\tau) = \mathscr{P}_0(\tau, \tau_0)\, \Delta\mathbf{y}(\tau_0) + \int_{\tau_0}^{\tau} \mathscr{P}_0(\tau, \tau')\, \Delta B(\tau')\, d\tau', \tag{32}$$

where $\Delta\mathbf{y}(\tau_0)$ is the initial perturbation and $\mathscr{P}_0$ is the propagator of the paraxial system (6) in the unperturbed medium.

The perturbation solution (32) may be used to solve a number of initial and boundary value problems. For instance, if we want to trace a perturbed ray with the same initial conditions as the reference ray we would take $\Delta\mathbf{y}(\tau_0) = (0, \Delta\mathbf{p}_0)$, where $\Delta\mathbf{p}_0 = \mathbf{p}_0\, \Delta u^2/2u_0^2$. The perturbation of initial slowness is necessary in order to satisfy the perturbed eikonal equation $\langle \mathbf{p}_0 \mid \Delta\mathbf{p} \rangle + \Delta H = 0$ at the source. $\Delta\mathbf{p}_0 = 0$ if the medium is not perturbed at the source. One of the most interesting and straightforward applications of (32) is to ray continuation. Suppose we have solved the two-point ray tracing problem between a source and receiver. The perturbed ray passing through the same source (sampling parameter τ_s) and receiver (sampling parameter τ_r) as the unperturbed ray, is obtained using

$$\begin{aligned} &\Delta\mathbf{x}(\tau_s) = 0 \\ &\langle \mathbf{p}_0(\tau_s) \mid \Delta\mathbf{p}(\tau_s) \rangle + \Delta H(\tau_s) = 0 \\ &\pi_1^r\, \Delta\mathbf{x}(\tau_r) = 0 \end{aligned} \tag{33}$$

π_1^r is an element of the transformation matrix (15) that extrapolates $\Delta\mathbf{y}(\tau_r)$ on a local plane passing through the receiver. This extrapolation is necessary because the perturbed ray can arrive at the receiver with a different sampling parameter τ. Because of the linear relation (32) between $\Delta\mathbf{y}(\tau_s)$ and $\Delta\mathbf{y}(\tau_r)$, inserting (33) in (32), we easily find the initial conditions $\Delta\mathbf{y}(\tau_s)$.

4.2 Ray perturbation due to interface change

Let us now consider a smooth perturbation of an interface. We denote by $f_0(\mathbf{x}) = 0$ the reference interface and $f(\mathbf{x}) = f_0(\mathbf{x}) + \Delta f(\mathbf{x}) = 0$ the perturbed interface. To first-order in Δf, it is possible to linearize the problem considering rays that deviate only slightly from the reference ray traced in the unperturbed medium. Because of linearity, perturbation of interfaces may be considered independently of perturbation of slowness.

Consider as before a reference ray with canonical vector $\mathbf{y}_0(\tau)$, and a ray in the perturbed medium that propagates in the neighbourhood of this reference ray. Its perturbation vector measured from the reference ray is $\Delta\mathbf{y}(\tau)$. The reference ray intersects the initial interface at $\mathrm{O}[\mathbf{x}_0(\tau_i)]$ and the perturbed ray intersects the perturbed interface at $\mathrm{O}'[\mathbf{x}(\tau_i')]$ (Fig. 4). Denoting $d\mathbf{x} = \mathbf{x}(\tau_i') - \mathbf{x}_0(\tau_i)$ and $d\mathbf{p} = \mathbf{p}(\tau_i') - \mathbf{p}_0(\tau_i)$ we find from Fig. 4 that to first order:

$$\begin{aligned} d\mathbf{x} &= \Delta\mathbf{x} + \nabla_{\mathbf{p}} H_0\, d\tau \\ d\mathbf{p} &= \Delta\mathbf{p} - \nabla_{\mathbf{x}} H_0\, d\tau, \end{aligned} \tag{34}$$

where $d\tau$ is the increment $(\tau_i' - \tau_i)$. In (34), the gradients of H_0 are calculated at $\mathbf{y}_0(\tau_i)$. These expressions are the same as (12) but their geometrical interpretation is quite different.

Because O′ belongs to the perturbed interface $f[\mathbf{x}(\tau_i')] = 0$

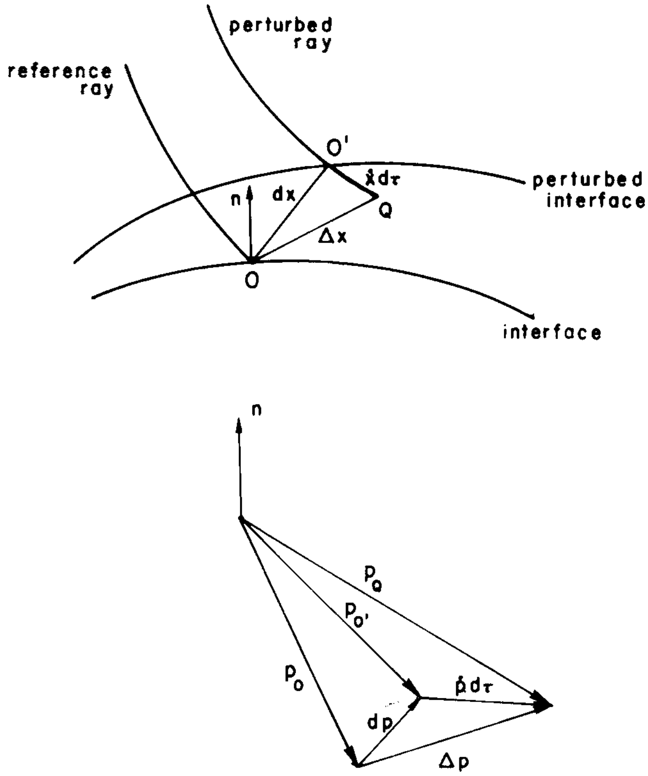

Figure 4. Geometry of the interaction of a ray with a perturbed interface. The reference ray intersects the reference interface at O, while the perturbed ray intersects the perturbed interface at O′.

and at O, $\mathbf{f}_0[\mathbf{x}_0(\tau_i)] = 0$, we get to first-order:

$$\langle \nabla f_0 \mid d\mathbf{x} \rangle + \Delta f = 0, \tag{35}$$

where the gradient and Δf are calculated at $\mathbf{x}_0(\tau_i)$. Using (34) and (35), we obtain the sampling parameter increment $d\tau$:

$$d\tau = -\frac{\langle \Delta\mathbf{x}(\tau_i) \mid \nabla f_0 \rangle + \Delta f}{\langle \nabla_{\mathbf{p}} H_0 \mid \nabla f_0 \rangle}. \tag{36}$$

Finally, the rotated paraxial vector $d\mathbf{y}$ is given by the canonical transformation:

$$d\mathbf{y} = \Pi_0\, \Delta\mathbf{y}(\tau_i) + \Delta\mathbf{y}_1^I, \tag{37}$$

with

$$\Pi_0 = \begin{pmatrix} \pi_1 & 0 \\ \pi_2 & I \end{pmatrix}, \tag{38}$$

where π_1 and π_2 are matrices defined as:

$$\pi_1 = I - \frac{|\nabla_{\mathbf{p}} H_0\rangle\langle \nabla f_0|}{\langle \nabla_{\mathbf{p}} H_0 \mid \nabla f_0 \rangle} \qquad \pi_2 = \frac{|\nabla_{\mathbf{x}} H_0\rangle\langle \nabla f_0|}{\langle \nabla_{\mathbf{p}} H_0 \mid \nabla f_0 \rangle} \tag{39}$$

and

$$\Delta y_1^I = -\begin{pmatrix} \nabla_{\mathbf{p}} H_0 \\ -\nabla_{\mathbf{x}} H_0 \end{pmatrix} \frac{\Delta f}{\langle \nabla_{\mathbf{p}} H_0 \mid \nabla f_0 \rangle}. \tag{40}$$

In order to propagate the transmitted and reflected rays away from the interface, we have to change the reference unperturbed ray. We choose as the new reference ray, the reflected-transmitted ray of the unperturbed medium corresponding to the reference incident ray. We denote by $\hat{\mathbf{y}}_0(\tau)$ and $\hat{\mathbf{y}}(\tau)$ the canonical vectors of the new reference ray and the perturbed ray in the reflected-transmitted medium.

Let us now construct the canonical perturbation $d\hat{\mathbf{y}} = (d\hat{\mathbf{x}}, d\hat{\mathbf{p}})$ of the perturbed transmitted/reflected ray. The perturbed ray has to satisfy $\hat{H} \equiv 0$. Thus, to first-order, the perturbations of position and slowness vector satisfy

$$d\hat{H} = \langle \nabla_\mathbf{x} \hat{H}_0 \mid d\hat{\mathbf{x}} \rangle + \langle \nabla_\mathbf{p} \hat{H}_0 \mid d\hat{\mathbf{p}} \rangle + \Delta \hat{H} = 0, \tag{41}$$

where $\hat{H}_0$, $\hat{H} = \hat{H}_0 + \Delta \hat{H}$ are the Hamiltonians in the reference and the perturbed medium, respectively. In the incidence medium we have the corresponding relation:

$$dH = \langle \nabla_\mathbf{x} H_0 \mid d\mathbf{x} \rangle + \langle \nabla_\mathbf{p} H_0 \mid d\mathbf{p} \rangle + \Delta H = 0. \tag{42}$$

The continuity of the perturbed ray gives the following relation:

$$d\hat{\mathbf{x}} = d\mathbf{x}. \tag{43}$$

Moreover, the perturbed ray satisfies the Snell's law

$$(\hat{\mathbf{p}}_0 + d\hat{\mathbf{p}}) \times \nabla f = (\mathbf{p}_0 + d\mathbf{p}) \times \nabla f, \tag{44}$$

The normal to the perturbed interface at $\mathbf{x}(\tau_i')$ is given to first-order by:

$$\nabla f = \nabla f_0 + \nabla\nabla f_0 \, | d\mathbf{x} \rangle + \nabla(\Delta f), \tag{45}$$

where the vectors ∇f_0 and $\nabla(\Delta f)$ and the matrix $\nabla\nabla f_0$ are computed at $\mathbf{x}_0(\tau_i)$. Thus, we obtain the perturbed Snell's law:

$$\begin{aligned} d\hat{\mathbf{p}} \times \nabla f_0 = d\mathbf{p} \times \nabla f_0 + (\nabla_\mathbf{p} H_0 - \nabla_\mathbf{p} \hat{H}_0) \\ \times [\nabla\nabla f_0 \mid d\mathbf{x} \rangle + \nabla(\Delta f)]. \end{aligned} \tag{46}$$

As in (22) we express $d\hat{\mathbf{p}}$ in the form:

$$d\hat{\mathbf{p}} = \frac{\langle d\hat{\mathbf{p}} \mid \nabla f_0 \rangle \nabla f_0 + (\nabla f_0 \times d\hat{\mathbf{p}}) \times \nabla f_0}{\langle \nabla f_0 \mid \nabla f_0 \rangle}. \tag{47}$$

Following the same procedure as before we write $d\hat{\mathbf{p}}$ as the sum of three terms:

$$d\hat{\mathbf{p}} = T_1 \, d\mathbf{x} + T_2 \, d\mathbf{p} + \Delta \mathbf{p}_2^I \tag{48}$$

with

$$\begin{aligned} T_1 &= -\frac{|\nabla f_0\rangle \langle \nabla_\mathbf{x} \hat{H}_0 - \nabla_\mathbf{x} H_0|}{\langle \nabla_\mathbf{p} \hat{H}_0 \mid \nabla f_0 \rangle} \\ &\quad - \frac{\langle \nabla_\mathbf{p} H_0 - \nabla_\mathbf{p} \hat{H}_0 \mid \nabla f_0 \rangle}{\langle \nabla f_0 \mid \nabla f_0 \rangle} \left\{ I - \frac{|\nabla f_0\rangle \langle \nabla_\mathbf{p} \hat{H}_0|}{\langle \nabla_\mathbf{p} \hat{H}_0 \mid \nabla f_0 \rangle} \right\} \nabla\nabla f_0 \\ T_2 &= I - \frac{|\nabla f_0\rangle \langle \nabla_\mathbf{p} \hat{H}_0 - \nabla_\mathbf{p} H_0|}{\langle \nabla_\mathbf{p} \hat{H}_0 \mid \nabla f_0 \rangle} \\ \Delta \mathbf{p}_2^I &= \frac{\Delta H - \Delta \hat{H}}{\langle \nabla_\mathbf{p} \hat{H}_0 \mid \nabla f_0 \rangle} \nabla f_0 \\ &\quad - \frac{\langle \nabla_\mathbf{p} H_0 - \nabla_\mathbf{p} \hat{H}_0 \mid \nabla f_0 \rangle}{\langle \nabla f_0 \mid \nabla f_0 \rangle} \left\{ I - \frac{|\nabla f_0\rangle \langle \nabla_\mathbf{p} \hat{H}_0|}{\langle \nabla_\mathbf{p} \hat{H}_0 \mid \nabla f_0 \rangle} \right\} \nabla(\Delta f). \end{aligned} \tag{49}$$

Finally from (43) and (48), we obtain the continuity conditions for the perturbed ray across interfaces:

$$d\hat{\mathbf{y}} = T_0 \, d\mathbf{y} + \Delta \mathbf{y}_2^I \tag{50}$$

with

$$\Delta \mathbf{y}_2^I = \begin{pmatrix} 0 \\ \Delta \mathbf{p}_2^I \end{pmatrix} \tag{51}$$

$$T_0 = \begin{pmatrix} I & 0 \\ T_1 & T_2 \end{pmatrix}. \tag{52}$$

The new canonical vector $d\hat{\mathbf{y}}$ is used as the initial condition $\Delta\hat{\mathbf{y}}$ to propagate the reflected-transmitted perturbed ray away from the interface. Using (37) and (50), we obtain:

$$\Delta\hat{\mathbf{y}}(\hat{\tau}_i) = T_0 \Pi_0 \, \Delta\mathbf{y}(\tau_i) + T_0 \, \Delta\mathbf{y}_1^I + \Delta\mathbf{y}_2^I. \tag{53}$$

This transformation contains three terms. The first one is the same as the linear transformation (28) connecting the incident and reflected-transmitted paraxial rays at the interface. This term takes into account perturbations in initial conditions and in slowness between the source and the interface. $\Delta\mathbf{y}(\tau_i)$ is given by (32). The next two terms include the effect of interface perturbation. $\Delta\mathbf{y}_1^I$ is due to the displacement Δf of the interface. $\Delta\mathbf{y}_2^I$ is a perturbation of the slowness of the ray that emerges from the interface. It contains two terms as shown in (49). The first one is due to slowness perturbation in the vicinity of the interface; this term is due to the change in Snell's law produced by velocity perturbation in the vicinity of the interface. The second term in $\Delta\mathbf{y}_2^I$ is due to the rotation of the normal to the interface at point O.

5 PERTURBATION OF PARAXIAL RAYS

In this section we consider the more difficult problem of the propagation of paraxial rays in the perturbed medium. These are rays that propagate in the vicinity of the perturbed ray $\mathbf{y}(\tau) = \mathbf{y}_0(\tau) + \Delta\mathbf{y}(\tau)$. As shown in (6) paraxial rays are solutions of the following linear system of equations:

$$\delta\dot{\mathbf{y}} = A \, \delta\mathbf{y}, \tag{54}$$

where $\delta\mathbf{y}$ is the paraxial canonical vector measured from the perturbed reference ray $\mathbf{y}(\tau)$ (see Fig. 3 for a definition of $\delta\mathbf{x}$ and $\delta\mathbf{p}$). Matrix A is of the form (7), where matrix U_0 is replaced by matrix U, which contains the second-order partial derivatives of the square of slowness $u^2(\mathbf{x})$ computed on the perturbed central ray. To first order A may be expanded in the form $A(\tau) = A_0(\tau) + \Delta A(\tau)$, where

$$\Delta A = \begin{pmatrix} 0 & 0 \\ \langle \Delta\mathbf{x} \mid \nabla_\mathbf{x} \rangle U_0 + \Delta U & 0 \end{pmatrix}. \tag{55}$$

Matrices $\langle \Delta\mathbf{x} \mid \nabla_\mathbf{x} \rangle U_0$ and ΔU are defined by:

$$[\langle \Delta\mathbf{x} \mid \nabla_\mathbf{x} \rangle U_0]_{ij} = \langle \Delta\mathbf{x} \mid \nabla_\mathbf{x} [U_0]_{ij} \rangle \quad \text{and} \quad \Delta U_{ij} = \frac{1}{2} \frac{\partial^2 \Delta u^2}{\partial x_i \, \partial x_j}.$$

In (55), all the quantities are calculated on the original unperturbed ray $\mathbf{y}_0$. The gradient of U_0 comes from the perturbation $\Delta\mathbf{y}$ of the reference central ray and the matrix ΔU is a term due to perturbations in the slowness. The paraxials of the perturbed ray are given by

$$\delta\mathbf{y}(\tau) = \mathscr{P}(\tau, \tau_0) \, \delta\mathbf{y}(\tau_0), \tag{56}$$

where $\mathscr{P}(\tau, \tau_0)$ is the propagator of (54). To first-order in the slowness perturbation, $\mathscr{P}(\tau, \tau_0)$ is given by its Born approximation (Farra & Madariaga 1987):

$$\mathcal{P}(\tau, \tau_0) = \mathcal{P}_0(\tau, \tau_0) + \int_{\tau_0}^{\tau} \mathcal{P}_0(\tau, \tau') \, \Delta A(\tau') \mathcal{P}_0(\tau', \tau_0) \, d\tau'. \quad (57)$$

Let us now consider the interaction of the paraxials of the perturbed ray with a perturbed interface. The continuity conditions of the reflected or transmitted paraxial rays are obtained from the transformation (28) at the interface:

$$\delta\hat{\mathbf{y}}(\hat{\tau}_i) = T\Pi \, \delta\mathbf{y}(\tau_i'), \quad (58)$$

where matrices Π and T are computed at $\mathbf{y} = \mathbf{y}_0(\tau_i) + d\mathbf{y}$ and $\hat{\mathbf{y}} = \mathbf{y}_0(\hat{\tau}_i) + d\hat{\mathbf{y}}$, in the perturbed medium. τ_i' is the sampling parameter of the reference perturbed ray at its incident point. To first-order, the matrix $T\Pi$ can be expanded in a Taylor series in $d\mathbf{y}$ and $d\hat{\mathbf{y}}$:

$$T\Pi = T_0\Pi_0 + \langle d\mathbf{x} \,|\, \nabla_\mathbf{x} \rangle T_0\Pi_0 + \langle d\mathbf{p} \,|\, \nabla_\mathbf{p} \rangle T_0\Pi_0 + \langle d\hat{\mathbf{x}} \,|\, \nabla_{\hat{\mathbf{x}}} \rangle T_0\Pi_0 + \langle d\hat{\mathbf{p}} \,|\, \nabla_{\hat{\mathbf{p}}} \rangle T_0\Pi_0 + \Delta(T\Pi). \quad (59)$$

The derivatives of $T_0\Pi_0$ come from the perturbations of the reference central ray and $\Delta(T\Pi)$ is a term due to perturbations in the Hamiltonian and in the interface. Explicit expressions for the different terms in (59) are given in the Appendix.

Expression (57) gives the paraxial vector $\delta\mathbf{y}(\tau_i)$ at τ_i, where τ_i is the sampling parameter of the reference unperturbed ray at its incident point on the interface. We obtain from (54):

$$\delta\mathbf{y}(\tau_i') = (I + A_0 \, d\tau) \, \delta\mathbf{y}(\tau_i), \quad (60)$$

where the increment $d\tau = (\tau_i' - \tau_i)$ is given by (36) and A_0 is calculated at $\mathbf{x}_0(\tau_i)$ from (7). The paraxial vector $\delta\hat{\mathbf{y}}(\hat{\tau}_i)$ determined from (58) and (60) can be used as the new initial condition to propagate the reflected-transmitted paraxial ray away from the perturbed interface.

6 TRAVELTIMES AND AMPLITUDES

In order to construct a continuous wavefront in the vicinity of a central ray we have to impose an additional condition to the paraxial rays calculated with equation (6). Without this condition the paraxial rays would cross each other in random ways. Following a notation introduced by Popov (1982), we require that

$$\delta\mathbf{p}(\tau_0) = M_0 \, \delta\mathbf{x}(\tau_0), \quad (61)$$

where M_0 is a 3×3 matrix that determines the initial shape of the ray beam. Equation (61) is a linear relation between the components of slowness and position perturbation vectors for a given value τ_0. For Snell-waves, M_0 has the following form

$$M_0 = \begin{pmatrix} 0 & 0 & m_1 \\ 0 & 0 & m_2 \\ m_1 & m_2 & m_3 \end{pmatrix},$$

where

$$m_1 = \frac{1}{2p_z} \frac{\partial u^2}{\partial x}$$

$$m_2 = \frac{1}{2p_z} \frac{\partial u^2}{\partial y}$$

and

$$m_3 = \frac{1}{2p_z} \left(\frac{\partial u^2}{\partial z} - \frac{p_x}{p_z} \frac{\partial u^2}{\partial x} - \frac{p_y}{p_z} \frac{\partial u^2}{\partial y} \right).$$

For a point source, matrix M_0 is singular and is given by

$$M_0 = \lim_{\epsilon \to 0} \frac{u_0}{\epsilon} \left[I - \frac{|\mathbf{p}_0\rangle\langle\mathbf{p}_0|}{\langle\mathbf{p}_0 \,|\, \mathbf{p}_0\rangle} \right]$$

where $\mathbf{p}_0$ is the initial slowness vector of the central ray.

Let us introduce the now classical notation for the submatrices of the paraxial ray propagator (see Červený 1985):

$$\mathcal{P}(\tau, \tau_0) = \begin{pmatrix} Q_1 & Q_2 \\ P_1 & P_2 \end{pmatrix}, \quad (62)$$

so that the paraxial solution is written

$$\begin{pmatrix} \delta\mathbf{x}(\tau) \\ \delta\mathbf{p}(\tau) \end{pmatrix} = \begin{pmatrix} Q_1 & Q_2 \\ P_1 & P_2 \end{pmatrix} \begin{pmatrix} \delta\mathbf{x}(\tau_0) \\ \delta\mathbf{p}(\tau_0) \end{pmatrix}. \quad (63)$$

Using the initial condition (61), (63) gives

$$\begin{aligned} \delta\mathbf{x}(\tau) &= (Q_1 + Q_2 M_0) \, \delta\mathbf{x}(\tau_0) \\ \delta\mathbf{p}(\tau) &= (P_1 + P_2 M_0) \, \delta\mathbf{x}(\tau_0) \end{aligned} \quad (64)$$

and the perturbed position and slowness are linearly related by

$$\delta\mathbf{p}(\tau) = M(\tau) \, \delta\mathbf{x}(\tau), \quad (65)$$

where $M(\tau)$ is the 3×3 matrix:

$$M(\tau) = (P_1 + P_2 M_0)(Q_1 + Q_2 M_0)^{-1}. \quad (66)$$

Matrices $M(\tau)$ for increasing values of τ may then be obtained from their initial value at the source

$$M(\tau_0) = M_0. \quad (67)$$

With this relationship, we can write the second-order expansion of the traveltime around a central ray in the form:

$$\theta(\mathbf{x} + \delta\mathbf{x}) = \theta(\mathbf{x}) + \mathbf{p} \cdot \delta\mathbf{x} + \tfrac{1}{2} \delta\mathbf{x}' M(\tau) \, \delta\mathbf{x}, \quad (68)$$

where $\mathbf{x}(\tau)$ and $\mathbf{p}(\tau)$ are the position and slowness vector of the central perturbed ray. To first-order in the slowness perturbation, the traveltime $\theta(\mathbf{x})$ along the perturbed reference ray is

$$\begin{aligned} \theta(\mathbf{x}) &= \theta_0(\mathbf{x}) \\ &+ \left[\frac{1}{2} \int_{\tau_0}^{\tau} \Delta u^2 \, d\tau - \sum_{\text{interfaces}} \frac{\langle \mathbf{p}_0(\tau_i) - \hat{\mathbf{p}}_0(\hat{\tau}_i) \,|\, \nabla f_0 \rangle}{\langle \nabla f_0 \,|\, \nabla f_0 \rangle} \Delta f \right] \\ &+ [\mathbf{p}_0(\tau) \cdot \Delta\mathbf{x}(\tau) - \mathbf{p}_0(\tau_0) \cdot \Delta\mathbf{x}(\tau_0)] \end{aligned} \quad (69)$$

with the obvious notation that $\theta_0(\mathbf{x}_0)$ is the travel time along the unperturbed reference ray $\mathbf{x}_0(\tau)$. The second term in (69) comes from the slowness and interface perturbations, the third one is due to perturbations in position of the source and receiver. This expression is comparable with the one used by Bishop *et al.* (1985) in reflection tomography of interfaces.

We may now determine the amplitude. Consider the Jacobian

$$D = \det \left(\frac{\partial \mathbf{x}}{\partial \mathbf{x}_0} \right), \quad (70)$$

where $\mathbf{x_0}$ is an initial point on a ray and $\mathbf{x}$ defines a ray path. Then, following Thomson & Chapman (1985), we can write

$$A(\mathbf{x}) = \mathcal{R}\frac{A(\mathbf{x_0})}{\sqrt{\mathbf{D}}}, \tag{71}$$

where $\mathcal{R}$ is the product of reflection and transmission coefficients at the interfaces. Using the first relation of (64), we obtain

$$D = \det(Q_1 + Q_2 M_0). \tag{72}$$

This expression has to be modified for point sources since in this case M_0 is singular. This difficulty is solved incorporating ϵ in the excitation function $s(\omega)$ (see equation 1).

Using (68) for θ and (71) for the amplitude, we have the general expression for a beam in the vicinity of a reference ray. These expressions are valid both in unperturbed and perturbed media. In the former case, Q_i and P_i are obtained partitioning the unperturbed propagator $\mathcal{P}_0(\tau, \tau_0)$, in the latter from the perturbed propagator (57).

7 EXAMPLE: CELL RAY TRACING

Virieux *et al.* (1988) proposed a finite element method for ray tracing in 3-D media. In this method the medium is discretized into triangles or tetrahedra with a linear distribution of the square of the slowness. Let the linear distribution of the square of slowness in one of the cells be

$$u_0^2(\mathbf{x}) = \alpha_0^2 + \langle \boldsymbol{\gamma_0} \mid \mathbf{x} \rangle, \tag{73}$$

where $\boldsymbol{\gamma_0}$ is the gradient. In order to obtain exact rays, we put the slowness distribution (73) into the ray equations (4). Solving the corresponding system, we find the simple expressions (Virieux *et al.* 1988):

$$\begin{aligned} \mathbf{p_0}(\tau) &= \tfrac{1}{2}(\tau - \tau_0)\boldsymbol{\gamma_0} + \mathbf{p_0}(\tau_0) \\ \mathbf{x_0}(\tau) &= \tfrac{1}{4}(\tau - \tau_0)^2\boldsymbol{\gamma_0} + (\tau - \tau_0)\mathbf{p_0}(\tau_0) + \mathbf{x_0}(\tau_0), \end{aligned} \tag{74}$$

where $\mathbf{x_0}(\tau_0)$ and $\mathbf{p_0}(\tau_0)$ are the initial conditions.

Paraxial rays in the unperturbed medium are obtained by a small perturbation of the initial conditions $\delta\mathbf{y}(\tau_0)$. They satisfy the linear differential system (6) where the matrix A_0 for the slowness distribution (73) is:

$$A_0 = \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix}. \tag{75}$$

The solutions of (6) are:

$$\delta\mathbf{y}(\tau) = \mathcal{P}_0(\tau, \tau_0)\, \delta\mathbf{y}(\tau_0),$$

where $\mathcal{P}_0$ is the propagator matrix given by:

$$\mathcal{P}_0(\tau, \tau_0) = \begin{pmatrix} I & (\tau - \tau_0)I \\ 0 & I \end{pmatrix}. \tag{76}$$

Let us assume that the central ray has been traced inside one of the cells up to one of its plane boundaries. Let, for instance, $f(\mathbf{x}) = \langle \mathbf{n_0} \mid \mathbf{x} \rangle - q = 0$ be the equation of this plane of unit normal $\mathbf{n_0}$. The continuity conditions for the paraxial rays transmitted across this boundary are given by:

$$\delta\hat{\mathbf{y}} = T_0 \Pi_0\, \delta\mathbf{y}(\tau_i) \tag{77}$$

The matrix Π_0 has the following submatrices:

$$\pi_1 = I - \frac{|\mathbf{p_0}\rangle\langle \mathbf{n_0}|}{\langle \mathbf{n_0} \mid \mathbf{p_0}\rangle} \qquad \pi_2 = -\frac{1}{2}\frac{|\boldsymbol{\gamma_0}\rangle\langle \mathbf{n_0}|}{\langle \mathbf{n_0} \mid \mathbf{p_0}\rangle}, \tag{78}$$

where $\mathbf{p_0}$ is the slowness vector of the central ray on the interface. Moreover, because of the continuity of the velocity field through the boundary, the submatrix T_2 reduces to the identity matrix, while the vector $T_1 \Pi_1\, \delta\mathbf{x}(\tau_i)$ is zero. Then, transformation T_0 acts simply as the identity for this case. Fig. 5 shows the results of ray tracing using Virieux *et al.* (1988) technique in a medium with a vertical gradient of the velocity. The medium is divided in rectangular elements subdivided in triangles. The square of the slowness is given at each node of the grid. Assuming a gradient of the square of the slowness in each triangle gives a continuous u^2 field. This prevents any reflection or refraction of the central ray. But, for the paraxial ray or the canonical vector, this simple distribution requires that the operator Π_0 be applied at each boundary, giving either a vector $\delta\hat{\mathbf{x}}$ along the axis x or the axis z or the anti-diagonal. In Fig. 5 we plot several rays shot from a point source at a depth of 16 km. The medium has a constant velocity gradient $v(z) = (3.0 + 0.7z)$ km s^{-1}. A paraxial vector $\delta\hat{\mathbf{x}}$ is drawn at each intersection of the central ray with the sides of the triangular mesh. The widening of a ray tube might be estimated from this paraxial information.

7.1 Perturbation of the medium

Let us now consider the following perturbed slowness distribution $u^2(\mathbf{x}) = u_0^2(\mathbf{x}) + \Delta u^2(\mathbf{x})$, where the perturbation $\Delta u^2(\mathbf{x})$ is continuous at internal boundaries and has a constant gradient inside each cell:

$$\Delta u^2(\mathbf{x}) = \Delta\alpha^2 + \langle \Delta\boldsymbol{\gamma} \mid \mathbf{x} \rangle. \tag{79}$$

Perturbed rays are solutions of (30) with

$$\Delta B(\tau) = \begin{pmatrix} 0 \\ \tfrac{1}{2}\Delta\boldsymbol{\gamma} \end{pmatrix}. \tag{80}$$

The perturbation $\Delta\mathbf{y}(\tau)$ of these rays is given by (32), which can be explicitly written as:

$$\begin{aligned} \Delta\mathbf{x}(\tau) &= \Delta\mathbf{x}(\tau_0) + (\tau - \tau_0)\, \Delta\mathbf{p}(\tau_0) + \tfrac{1}{4}(\tau - \tau_0)^2\, \Delta\boldsymbol{\gamma} \\ \Delta\mathbf{p}(\tau) &= \Delta\mathbf{p}(\tau_0) + \tfrac{1}{2}(\tau - \tau_0)\, \Delta\boldsymbol{\gamma}. \end{aligned} \tag{81}$$

Using expression (74) for the reference ray, the perturbed rays $\mathbf{y}(\tau) = \mathbf{y}_0(\tau) + \Delta\mathbf{y}(\tau)$ are given by:

$$\begin{aligned} \mathbf{x}(\tau) &= \mathbf{x}(\tau_0) + (\tau - \tau_0)\mathbf{p}(\tau_0) + \tfrac{1}{4}(\tau - \tau_0)^2\boldsymbol{\gamma} \\ \mathbf{p}(\tau) &= \mathbf{p}(\tau_0) + \tfrac{1}{2}(\tau - \tau_0)\boldsymbol{\gamma}, \end{aligned} \tag{82}$$

where $\boldsymbol{\gamma} = \boldsymbol{\gamma_0} + \Delta\boldsymbol{\gamma}$ is the gradient of the square of slowness in the perturbed medium. The first-order solutions (82) reduce to the exact ray expressions (74) for the perturbed medium. Thus, first-order perturbation gives the exact ray equations in media with constant gradient of the square of the slowness.

Paraxial rays with respect to the perturbed ray $\mathbf{y}(\tau) = \mathbf{y}_0 + \Delta\mathbf{y}$ are solutions of (54) with $A(\tau) = A_0(\tau)$. The perturbed propagator $\mathcal{P}$ reduces then to the unperturbed propagator $\mathcal{P}_0$ given in (76). At internal boundaries, the continuity conditions for perturbed rays are given by (53)

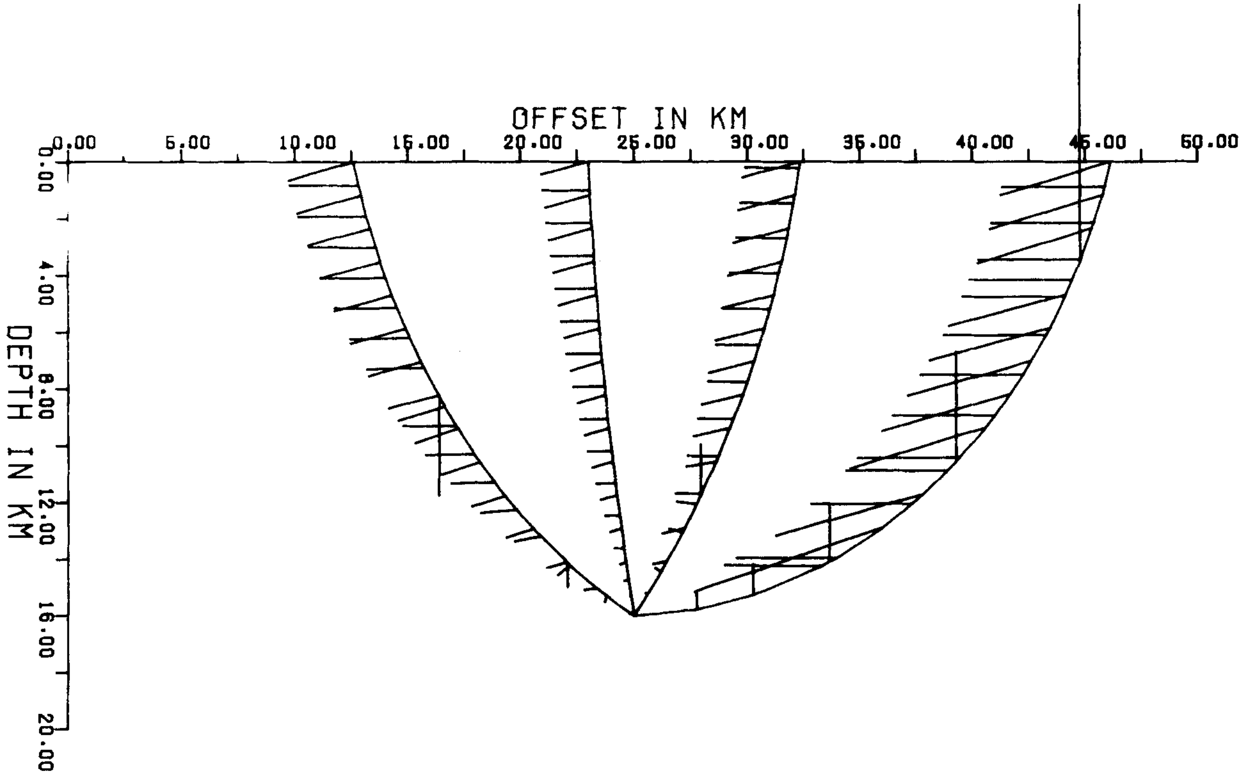

Figure 5. Example of ray tracing in a triangular mesh. Four central rays radiated from a point source at 16 km depth are plotted. A paraxial ray is drawn for each of the central rays. The paraxial rays are represented by the relative position vectors $\delta\hat{\mathbf{x}}$ at each intersection of the central ray with the edge of one of the triangular elements, which can be horizontal, vertical or anti-diagonal.

which becomes:

$$\Delta\hat{\mathbf{y}}(\hat{\tau}_i) = \Pi_0\, \Delta\mathbf{y}(\tau_i). \quad (83)$$

The initial conditions of the perturbed paraxial rays are given by (58) and (60) which can be written for the simple reference medium

$$\delta\hat{\mathbf{y}}(\hat{\tau}_i) = [\Pi + \Pi_0 A_0\, d\tau]\, \delta\mathbf{y}(\tau_i), \quad (84)$$

where the matrix Π is given to first-order by

$$\Pi = \Pi_0 + \langle d\mathbf{p} \mid \nabla_{\mathbf{p}} \rangle \Pi_0 + \Delta\Pi, \quad (85)$$

where

$$\langle d\mathbf{p} \mid \nabla_{\mathbf{p}} \rangle \Pi_0 = \begin{pmatrix} -\pi_1 \dfrac{|\Delta\hat{\mathbf{p}}\rangle\langle \mathbf{n}_0|}{\langle \mathbf{n}_0 \mid \mathbf{p}_0 \rangle} & 0 \\ -\pi_2 \dfrac{|\Delta\hat{\mathbf{p}}\rangle\langle \mathbf{n}_0|}{\langle \mathbf{n}_0 \mid \mathbf{p}_0 \rangle} & 0 \end{pmatrix}, \quad (86)$$

$$\Delta\Pi = \begin{pmatrix} 0 & 0 \\ -\dfrac{1}{2}\dfrac{|\Delta\gamma\rangle\langle \mathbf{n}_0|}{\langle \mathbf{n}_0 \mid \mathbf{p}_0 \rangle} & 0 \end{pmatrix}.$$

The matrix $[\Pi_0\ A_0\ d\tau]$ is given by:

$$[\Pi_0\ A_0\ d\tau] = -\frac{\langle \Delta\mathbf{x} \mid \mathbf{n}_0 \rangle}{\langle \mathbf{p}_0 \mid \mathbf{n}_0 \rangle} \begin{pmatrix} 0 & \pi_1 \\ 0 & \pi_2 \end{pmatrix} \quad (87)$$

The perturbation of the propagator, which allows estimation of Fréchet derivatives for amplitudes, may be used to include wave amplitudes in the inversion of seismic velocity by the method of Aki, Christofferson & Husebye (1977).

8 A REFERENCE MEDIUM WITH HOMOGENEOUS LAYERS

The general results obtained above take a very simple form when the reference medium consists of homogeneous layers with plane interfaces. In the following we discuss only the case where the interfaces are perturbed.

Inside each layer, a reference ray is of course a straight line. Its expression is given by (74) with $\gamma_0 = 0$;

$$\begin{aligned} \mathbf{x}_0(\tau) &= (\tau - \tau_0)\mathbf{p}_0(\tau_0) + \mathbf{x}_0(\tau_0) \\ \mathbf{p}_0(\tau) &= \mathbf{p}_0(\tau_0) \end{aligned} \quad (88)$$

where $\mathbf{x}_0(\tau_0)$ and $\mathbf{p}_0(\tau_0)$ are the initial conditions. Paraxial rays in the unperturbed medium are given by (9) with the propagator $\mathscr{P}_0$ of equation (76).

Let us assume that the reference ray has been traced in the layer up to one of its plane boundaries. Let, for instance, $f_0(\mathbf{x}) = \langle \mathbf{n}_0 \mid \mathbf{x} \rangle - q = 0$ be the equation of this plane of unit normal $\mathbf{n}_0$. The continuity conditions for reflected-transmitted paraxial rays are:

$$\delta\hat{\mathbf{y}} = T_0 \Pi_0\, \delta\mathbf{y}(\tau_i), \quad (89)$$

where matrices Π_0 and T_0 have the following submatices:

$$\pi_1 = I - \frac{|\mathbf{p}_0\rangle\langle \mathbf{n}_0|}{\langle \mathbf{n}_0 \mid \mathbf{p}_0 \rangle} \qquad \pi_2 = 0 \quad (90)$$

$$T_1 = 0 \qquad T_2 = I - \frac{|\mathbf{n}_0\rangle\langle \hat{\mathbf{p}}_0 - \mathbf{p}_0|}{\langle \hat{\mathbf{p}}_0 \mid \mathbf{n}_0 \rangle},$$

where $\mathbf{p}_0$ and $\hat{\mathbf{p}}_0$ are the slowness vectors of the reference ray on the interface in the incident and the reflected-transmitted medium, respectively.

Because we consider only interface perturbations, between two interfaces perturbed rays behave like paraxial rays in the original medium. The perturbation $\Delta\mathbf{y}(\tau)$ of the reference ray is given by:

$$\Delta\mathbf{y}(\tau) = \mathscr{P}_0(\tau, \tau_0)\, \Delta\mathbf{y}(\tau_0), \tag{91}$$

where $\mathscr{P}_0$ is the propagator (76).

The continuity conditions for perturbed reflected-transmitted rays are given by (53) which becomes:

$$\Delta\hat{\mathbf{y}}(\hat{\tau}_i) = T_0\Pi_0\, \Delta\mathbf{y}(\tau_i) + \Delta\mathbf{y}^I, \tag{92}$$

with

$$\Delta\mathbf{y}^I = \begin{pmatrix} \Delta\mathbf{x}^I \\ \Delta\mathbf{p}^I \end{pmatrix},$$

where

$$\begin{aligned} \Delta\mathbf{x}^I &= -\frac{\Delta f}{\langle \mathbf{p}_0 \mid \mathbf{n}_0 \rangle}\mathbf{p}_0 \\ \Delta\mathbf{p}^I &= -\frac{\langle \mathbf{p}_0 - \hat{\mathbf{p}}_0 \mid \mathbf{n}_0 \rangle}{\langle \mathbf{n}_0 \mid \mathbf{n}_0 \rangle}\left\{ I - \frac{|\mathbf{n}_0\rangle\langle \hat{\mathbf{p}}_0|}{\langle \hat{\mathbf{p}}_0 \mid \mathbf{n}_0 \rangle}\right\}\nabla(\Delta f). \end{aligned} \tag{93}$$

The continuity conditions $\delta\hat{\mathbf{y}}(\hat{\tau}_i)$ of the reflected or transmitted paraxial rays of the perturbed central ray are obtained from (58) and (60) which can be written

$$\delta\hat{\mathbf{y}}(\hat{\tau}_i) = [T\Pi + T_0\Pi_0 A_0\, d\tau]\, \delta\mathbf{y}(\tau_i). \tag{94}$$

The matrix $T\Pi$ is given to first-order by

$$\begin{aligned} T\Pi = {} & T_0\Pi_0 + \langle d\mathbf{p} \mid \nabla_{\mathbf{p}} \rangle T_0\Pi_0 + \langle d\hat{\mathbf{p}} \mid \nabla_{\hat{\mathbf{p}}} \rangle T_0\Pi_0 \\ & + \Delta(T\Pi). \end{aligned} \tag{95}$$

We obtain from the Appendix:

$$\langle d\mathbf{p} \mid \nabla_{\mathbf{p}} \rangle T_0\Pi_0 = \begin{pmatrix} -\pi_1 \dfrac{|\Delta\mathbf{p}\rangle\langle \mathbf{n}_0|}{\langle \mathbf{n}_0 \mid \mathbf{p}_0 \rangle} & 0 \\ 0 & \dfrac{|\mathbf{n}_0\rangle\langle \Delta\mathbf{p}|}{\langle \hat{\mathbf{p}}_0 \mid \mathbf{n}_0 \rangle} \end{pmatrix} \tag{96}$$

$$\langle d\hat{\mathbf{p}} \mid \nabla_{\hat{\mathbf{p}}} \rangle T_0\Pi_0 = \begin{pmatrix} 0 & 0 \\ 0 & -\dfrac{|\mathbf{n}_0\rangle\langle \Delta\hat{\mathbf{p}}|}{\langle \hat{\mathbf{p}}_0 \mid \mathbf{n}_0 \rangle} T_2 \end{pmatrix}.$$

Matrix $\Delta(T\Pi)$ is given by:

$$\Delta(T\Pi) = \begin{pmatrix} \Delta\pi_1 & 0 \\ \Delta T_1 \pi_1 & \Delta T_2 \end{pmatrix}, \tag{97}$$

with

$$\Delta\pi_1 = -\frac{|\mathbf{p}_0\rangle\langle \nabla\Delta f|}{\langle \mathbf{p}_0 \mid \mathbf{n}_0 \rangle}\pi_1 \qquad \Delta T_2 = -\hat{T}_2 \frac{|\nabla\Delta f\rangle\langle \hat{\mathbf{p}}_0 - \mathbf{p}_0|}{\langle \hat{\mathbf{p}}_0 \mid \mathbf{n}_0 \rangle} \tag{98}$$

and

$$\Delta T_1 = -\frac{\langle \mathbf{p}_0 - \hat{\mathbf{p}}_0 \mid \mathbf{n}_0 \rangle}{\langle \mathbf{n}_0 \mid \mathbf{n}_0 \rangle}\hat{T}_2 \nabla\nabla\Delta f \qquad \hat{T}_2 = I - \frac{|\mathbf{n}_0\rangle\langle \hat{\mathbf{p}}_0|}{\langle \hat{\mathbf{p}}_0 \mid \mathbf{n}_0 \rangle}. \tag{99}$$

Matrix $[T_0\Pi_0 A_0\, d\tau]$ is given by:

$$[T_0\Pi_0 A_0\, d\tau] = -\frac{\langle \Delta\mathbf{x} \mid \mathbf{n}_0 \rangle + \Delta f}{\langle \mathbf{p}_0 \mid \mathbf{n}_0 \rangle}\begin{pmatrix} 0 & \pi_1 \\ 0 & 0 \end{pmatrix}. \tag{100}$$

We have now all the elements to calculate the ray and paraxial ray fields in the perturbed medium. These results will be used in the following section.

9 A SIMPLE EXAMPLE

We consider a simple 2-D acoustic wave propagation problem. Previous work by Farra & Madariaga (1987) gave examples in which they computed seismograms for a perturbed velocity structure using only rays traced in an initial velocity structure. We investigate here the perturbation of amplitude due to an interface perturbation. The reference medium consists of two homogeneous layers with velocities of 3 and 5 km s^{-1}, respectively. We perturb the shape of the interface between the two layers, transforming it into a small basin. The acoustic source was located on the free surface. As shown in Fig. 6, we compare the rays traced using the perturbation method with the result of exact ray tracing. The rays obtained by perturbation theory are shown by the dotted lines. The same rays calculated using exact ray tracing are shown with continuous lines. The error of the perturbation method is larger for the rays that intersect the perturbed interface at points where the curvature was strongly perturbed.

We calculated synthetics at a number of receivers distributed every 1 km on the horizontal line $z = 0$ km. For the calculation of exact seismograms we have to solve a series of two point ray tracing problems in order to trace rays from source to receiver. We did this interpolating the traveltimes and amplitudes obtained from neighbouring rays.

All the calculations were carried out in two dimensions. However, in order to simulate a 3-D medium, we chose a source time function of the following form:

$$s(t) = \frac{d}{dt}\left[\frac{H(t)}{\sqrt{t}}\right] * \dot{g}(t),$$

where $H(t)$ is the Heaviside function and $\dot{g}(t)$ is the derivative of the Ricker's function $g(t) = \exp -\frac{1}{2}(t/\Delta t_1)^2$ with $\Delta t_1 = 0.02$ s. In Figs 7 and 8 we show the synthetics calculated by classical ray theory in the reference and perturbed medium, respectively. They were computed using complex reflection coefficients at the interface. One can see

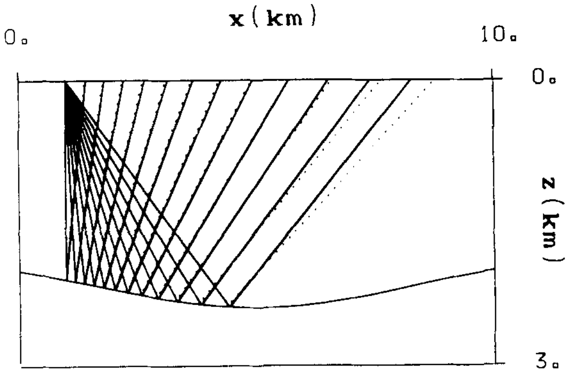

Figure 6. Geometry of a simple layered structure where the lower interface has been transformed into a slightly concave basin. The source is located at (1., 0.). Exact rays (solid lines) may be compared with rays calculated by perturbation theory (dotted lines).

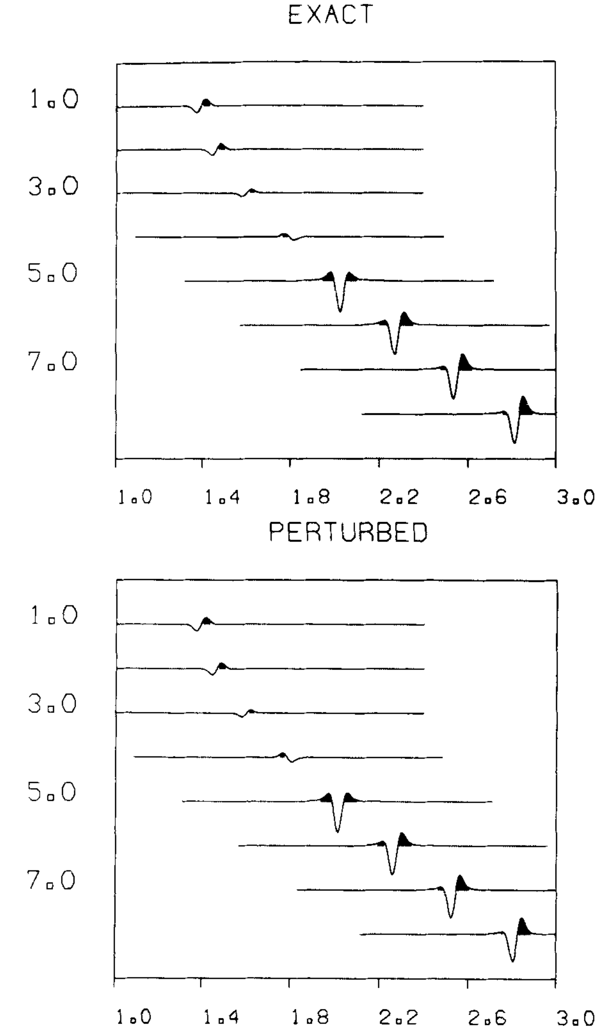

Figure 7. Synthetic seismograms calculated at a set of receivers regularly distributed along the surface of the structure. At the top we show the seismograms calculated exactly, while at the bottom are those obtained by perturbation theory.

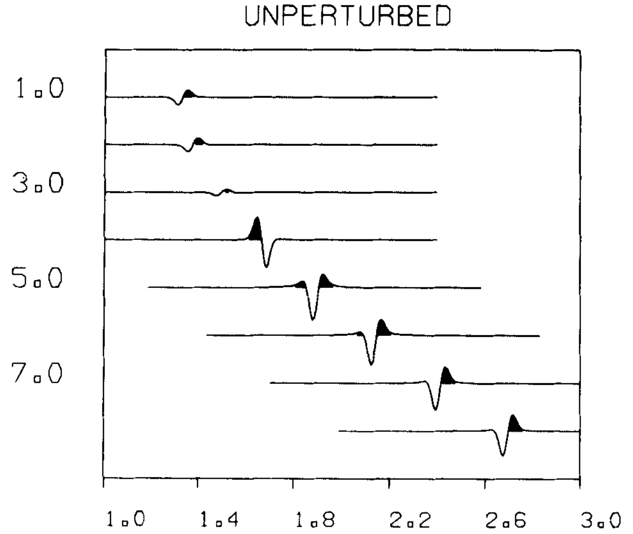

Figure 8. Synthetic seismograms calculated at a set of receivers regularly distributed along the free surface of the unperturbed layered model of Fig. 6.

the important effect of the perturbation on some of the synthetics. Since we are close to critical angle, the reflection coefficient is very sensitive to changes in incidence angle. At the bottom of Fig. 7, we show the synthetics calculated by the perturbation method. The reflection coefficients were computed for the perturbed incidence angle obtained from $\mathbf{p}_0 + d\mathbf{p}$ and $\nabla \Delta f$. The comparison with the results given by classical ray theory is excellent.

10 CONCLUSION

In this paper we extended previous work by Farra *et al.* (1987) on ray perturbation theory in order to include small changes in the position and shape of interfaces. A systematic approach to perturbation theory based on a simple Hamiltonian formulation was adopted in order to simplify the treatment of interfaces. An important innovation is the use of Cartesian coordinates in order to perform paraxial ray tracing instead of the most commonly used ray-centred coordinate system. This simplifies the equations of dynamic (paraxial) ray tracing considerably. Paraxial ray tracing in Cartesian coordinates is entirely equivalent to dynamic ray tracing as introduced by Popov & Pšenčík (1978) and by Červený (1985) in the case of an unperturbed medium. An interesting application of these results is to the study of simple velocity distributions that admit analytic ray and paraxial ray tracing. Among many such structures, Virieux *et al.* (1988) proposed a medium with constant gradient of square slowness. In such a medium all standard ray quantities (ray tracing and paraxial propagator) can be calculated exactly. This medium is ideal for a finite element approach to ray tracing. We subdivide the continuous medium into large triangles with constant gradient of the square of slowness. Inside the elements rays and paraxial rays are traced analytically. Once the elements are assembled, ray and paraxial ray tracing reduce to the solution of a series of algebraic continuity conditions at the intersection of the rays with the sides of the triangles in the mesh. Finally, we studied the effect of small perturbations of the velocity structure or of interface shape on central and paraxial rays. The expressions obtained in this paper using Cartesian coordinates are much simpler than those of Farra & Madariaga (1987) who used ray-centred coordinates.

The results for the perturbation of ray and paraxial rays were finally used to develop an approach for the calculation of synthetic seismograms when the velocity or the interfaces of the medium are slightly perturbed. Complete albeit lengthy expressions are given for the effect of perturbations of slowness and interface shape on synthetic seismograms.

The results obtained in this paper should be useful in several problems of seismology and applied geophysics. The most obvious application is to the study of the effect of small changes in a model upon synthetic seismograms. All that is required to calculate these effects is to know the paraxial propagator matrix along the ray trajectory. A typical example of this type of application is the perturbation of a simple vertically stratified model for which we give expressions both for ray tracing and the calculation of the paraxial ray propagator. Another possibility is to use the finite-element approach previously described in order to compute rays and the paraxial ray propagator.

Other applications are, for instance, to continuation methods for the solution of two-point ray tracing. In these methods a ray is traced through the source and receiver in a simpler medium than that in which we want to perform ray tracing. Then the two-point ray in the more complex medium is found by iterative perturbation of the rays in the simpler medium. Perturbation techniques provide a simple guess for the perturbation of the initial conditions of the ray when the medium properties change. Another very interesting application of perturbation theory is to the calculation of Fréchet derivatives for the inversion of waveforms and amplitudes of seismic waves. This approach has recently been used by Nowack & Lyslo (1989) for the inversion of interfaces and velocities. Another application that will be the subject of further work is the calculation of ray tracing and amplitudes in slightly anisotropic media. Anisotropy may be calculated as a perturbation with respect to a simpler related anisotropic medium.

ACKNOWLEDGMENTS

We thank I. Pšenčík and two anonymous reviewers for very stimulating reviews. This work was sponsored by Institut National des Sciences de l'Univers under its program ECORS. A fellowship from Institut Français du Pétrole is gratefully acknowledged by V. Farra. IPG Publication No. 1050.

REFERENCES

Aki, K., Christofferson, A. & Husebye, E. S., 1977. Determination of the three-dimensional seismic structure of the lithosphere. *J. geophys. Res.*, **82**, 277–296.

Bishop, T. N., Bube, K. P., Cutler, R. T., Langan, R. T., Love, P. L., Resnick, J. R., Shuey, R. T., Spindler, D. A. & Wyld, H. W., 1985. Tomographic determination of velocity and depth in laterally varying media, *Geophysics*, **50**, 903–923.

Burridge, R., 1976. *Some Mathematical Topics in Seismology*, Courant Institute of Mathematical Sciences, New York University, New York.

Červený, V., 1985. The application of ray tracing to the numerical modelling of seismic wave fields in complex structures, in *Handbook of Geophysical Exploration, Section 1, Seismic exploration*, Vol. 15A, pp. 1–119, Geophysical Press, London.

Červený, V., Langer, J. & Pšenčík, I., 1974. Computation of geometric spreading of seismic body waves in laterally inhomogeneous media with curved interfaces, *Geophys. J. R. astr. Soc.*, **38**, 9–19.

Červený, V., Molotkov, I. A. & Pšenčík, I., 1977. *Ray Method in Seismology*, Universita Karlova, Praha.

Červený, V., Popov, M. M. & Pšenčík, I., 1982. Computation of wave fields in inhomogeneous media-Gaussian beam approach, *Geophys. J. R. astr. Soc.*, **70**, 109–128.

Chapman, C. H., 1985. Ray theory and its extensions: WKBJ and Maslov seismograms, *J. Geophys.*, **58**, 27–43.

Chapman, C. H. & Drummond, R., 1982. Body-wave seismograms in inhomogeneous media using Maslov asymptotic theory, *Bull. seism. Soc. Am.*, **72**, S277–S317.

Farra, V., 1987. Approche hamiltonienne de la théorie des rais: application à l'étude de perturbations, inversion des temps de parcours en sismique réflexion, *Thèse parcours de l'Université Paris 7*, Paris.

Farra, V. & Madariaga, R., 1987. Seismic waveform modeling in heterogeneous media by ray perturbation theory, *J. geophys. Res.*, **92**, 2697–2712.

Goldstein, H., 1980. *Classical Mechanics*, Addison-Wesley, Reading, Massachussetts.

Nowack, R. L. & Lyslo, J. A., 1989. Frechet derivatives for curved interfaces in the ray approximation, *Geophys. J.*, **97**, 497–509.

Popov, M. M., 1982. A new method of computation of wave fields using Gaussian beams, *Wave Motion*, **4**, 85–97.

Popov, M. M. & Pšenčík, I., 1978. Computation of ray amplitudes in inhomogeneous media with curved interfaces, *Studia geophys. geod.*, **22**, 248–258.

Thompson, C. J. & Chapman, C. H., 1985. An introduction to Maslov's asymptotic method, *Geophys. J. R. astr. Soc.*, **83**, 143–168.

Virieux, J., Farra, V. & Madariaga, R., 1988. Ray tracing in laterally heterogeneous media for earthquake location, *J. geophys. Res.*, **93**, 6585–6599.

APPENDIX

Let us denote the matrix of second partial derivatives $\nabla\nabla f_0$ by F_0, while U_0 is the matrix of second partial derivatives of the square of the slowness.

$$\Pi_0 = \begin{pmatrix} \pi_1 & 0 \\ \pi_2 & I \end{pmatrix},$$

where

$$\pi_1 = I - \frac{|\nabla_{\mathbf{p}} H_0\rangle\langle\nabla f_0|}{\langle\nabla f_0 \,|\, \nabla_{\mathbf{p}} H_0\rangle}$$

$$\pi_2 = \frac{|\nabla_{\mathbf{x}} H_0\rangle\langle\nabla f_0|}{\langle\nabla f_0 \,|\, \nabla_{\mathbf{p}} H_0\rangle}.$$

Derivatives:

$$\langle d\mathbf{x} \,|\, \nabla_{\mathbf{x}}\rangle \Pi_0 = \begin{pmatrix} \langle d\mathbf{x} \,|\, \nabla_{\mathbf{x}}\rangle \pi_1 & 0 \\ \langle d\mathbf{x} \,|\, \nabla_{\mathbf{x}}\rangle \pi_2 & 0 \end{pmatrix}$$

with

$$\langle d\mathbf{x} \,|\, \nabla_{\mathbf{x}}\rangle \pi_1 = -\frac{|\mathbf{p}_0\rangle\langle d\mathbf{x}|}{\langle\nabla f_0 \,|\, \mathbf{p}_0\rangle} F_0 \pi_1$$

$$\langle d\mathbf{x} \,|\, \nabla_{\mathbf{x}}\rangle \pi_2 = -U_0 \frac{|d\mathbf{x}\rangle\langle\nabla f_0|}{\langle\nabla f_0 \,|\, \mathbf{p}_0\rangle} + \frac{|\nabla_{\mathbf{x}} H_0\rangle\langle d\mathbf{x}|}{\langle\nabla f_0 \,|\, \mathbf{p}_0\rangle} F_0 \pi_1$$

$$\langle d\mathbf{p} \,|\, \nabla_{\mathbf{p}}\rangle \Pi_0 = \begin{pmatrix} \langle d\mathbf{p} \,|\, \nabla_{\mathbf{p}}\rangle \pi_1 & 0 \\ \langle d\mathbf{p} \,|\, \nabla_{\mathbf{p}}\rangle \pi_2 & 0 \end{pmatrix}$$

with

$$\langle d\mathbf{p} \,|\, \nabla_{\mathbf{p}}\rangle \pi_1 = -\pi_1 \frac{|d\mathbf{p}\rangle\langle\nabla f_0|}{\langle\nabla f_0 \,|\, \mathbf{p}_0\rangle}$$

$$\langle d\mathbf{p} \,|\, \nabla_{\mathbf{p}}\rangle \pi_2 = -\pi_2 \frac{|d\mathbf{p}\rangle\langle\nabla f_0|}{\langle\nabla f_0 \,|\, \mathbf{p}_0\rangle}$$

$$T_0 = \begin{pmatrix} I & 0 \\ T_1 & T_2 \end{pmatrix},$$

with

$$T_2 = I - \frac{|\nabla f_0\rangle\langle\nabla_{\mathbf{p}}\hat{H}_0 - \nabla_{\mathbf{p}} H_0|}{\langle\nabla_{\mathbf{p}}\hat{H}_0 \,|\, \nabla f_0\rangle}$$

$$T_1 = -\frac{|\nabla f_0\rangle\langle\nabla_{\mathbf{x}}\hat{H}_0 - \nabla_{\mathbf{x}} H_0|}{\langle\nabla_{\mathbf{p}}\hat{H}_0 \,|\, \nabla f_0\rangle} - \frac{\langle\nabla_{\mathbf{p}} H_0 - \nabla_{\mathbf{p}}\hat{H}_0 \,|\, \nabla f_0\rangle}{\langle\nabla f_0 \,|\, \nabla f_0\rangle} \hat{T}_2 F_0$$

with

$$\hat{T}_2 = I - \frac{|\nabla f_0\rangle\langle\nabla_{\mathbf{p}}\hat{H}_0|}{\langle\nabla_{\mathbf{p}}\hat{H}_0 \,|\, \nabla f_0\rangle}.$$

Derivatives:

$$\langle d\mathbf{x} \mid \nabla_\mathbf{x}\rangle T_0 = \begin{pmatrix} 0 & 0 \\ \langle d\mathbf{x} \mid \nabla_\mathbf{x}\rangle T_1 & \langle d\mathbf{x} \mid \nabla_\mathbf{x}\rangle T_2 \end{pmatrix}$$

$$\langle d\mathbf{p} \mid \nabla_\mathbf{p}\rangle T_0 = \begin{pmatrix} 0 & 0 \\ \langle d\mathbf{p} \mid \nabla_\mathbf{p}\rangle T_1 & \langle d\mathbf{p} \mid \nabla_\mathbf{p}\rangle T_2 \end{pmatrix}$$

$$\langle d\mathbf{x} \mid \nabla_\mathbf{x}\rangle T_2 = -\hat{T}_2 F_0 \frac{|d\mathbf{x}\rangle\langle \hat{\mathbf{p}}_0 - \mathbf{p}_0|}{\langle \hat{\mathbf{p}}_0 \mid \nabla f_0\rangle}$$

$$\langle d\mathbf{p} \mid \nabla_\mathbf{p}\rangle T_2 = \frac{|\nabla f_0\rangle\langle d\mathbf{p}|}{\langle \hat{\mathbf{p}}_0 \mid \nabla f_0\rangle}$$

$$\langle d\hat{\mathbf{p}} \mid \nabla_{\hat{\mathbf{p}}}\rangle T_2 = -\frac{|\nabla f_0\rangle\langle d\hat{\mathbf{p}}|}{\langle \hat{\mathbf{p}}_0 \mid \nabla f_0\rangle} T_2$$

$$\langle d\mathbf{x} \mid \nabla_\mathbf{x}\rangle \hat{T}_2 = -\hat{T}_2 F_0 \frac{|d\mathbf{x}\rangle\langle \hat{\mathbf{p}}_0|}{\langle \hat{\mathbf{p}}_0 \mid \nabla f_0\rangle}$$

$$\langle d\hat{\mathbf{p}} \mid \nabla_{\hat{\mathbf{p}}}\rangle \hat{T}_2 = -\frac{|\nabla f_0\rangle\langle d\hat{\mathbf{p}}|}{\langle \hat{\mathbf{p}}_0 \mid \nabla f_0\rangle} \hat{T}_2$$

$$\langle d\mathbf{p} \mid \nabla_\mathbf{p}\rangle T_1 = -\frac{\langle d\mathbf{p} \mid \nabla f_0\rangle}{\langle \nabla f_0 \mid \nabla f_0\rangle} \hat{T}_2 F_0$$

$$\langle d\hat{\mathbf{p}} \mid \nabla_{\hat{\mathbf{p}}}\rangle T_1 = -\frac{|\nabla f_0\rangle\langle d\hat{\mathbf{p}}|}{\langle \hat{\mathbf{p}}_0 \mid \nabla f_0\rangle} T_1 + \frac{\langle d\hat{\mathbf{p}} \mid \nabla f_0\rangle}{\langle \nabla f_0 \mid \nabla f_0\rangle} \hat{T}_2 F_0$$

$$\begin{aligned}\langle d\mathbf{x} \mid \nabla_\mathbf{x}\rangle T_1 = &-\hat{T}_2 F_0 \frac{|d\mathbf{x}\rangle\langle \nabla_\mathbf{x}\hat{H}_0 - \nabla_\mathbf{x} H_0|}{\langle \hat{\mathbf{p}}_0 \mid \nabla f_0\rangle} \\ &+ \frac{\langle \mathbf{p}_0 - \hat{\mathbf{p}}_0 \mid F_0 \mid d\mathbf{x}\rangle}{\langle \nabla f_0 \mid \nabla f_0\rangle} \hat{T}_2 F_0 - \frac{\langle \mathbf{p}_0 - \hat{\mathbf{p}}_0 \mid \nabla f_0\rangle}{\langle \nabla f_0 \mid \nabla f_0\rangle} \\ &\{\langle d\mathbf{x} \mid \nabla_\mathbf{x}\rangle \hat{T}_2 F_0 + \hat{T}_2 \langle d\mathbf{x} \mid \nabla_\mathbf{x}\rangle F_0\} \\ &- \frac{|\nabla f_0\rangle\langle d\mathbf{x}|}{\langle \hat{\mathbf{p}}_0 \mid \nabla f_0\rangle} U_0\end{aligned}$$

$$\langle d\hat{\mathbf{x}} \mid \nabla_{\hat{\mathbf{x}}}\rangle T_1 = \frac{|\nabla f_0\rangle\langle d\hat{\mathbf{x}}|}{\langle \hat{\mathbf{p}}_0 \mid \nabla f_0\rangle} \hat{U}_0$$

$$\Delta \Pi = \begin{pmatrix} \Delta\pi_1 & 0 \\ \Delta\pi_2 & 0 \end{pmatrix},$$

where

$$\Delta\pi_1 = -\frac{|\mathbf{p}_0\rangle\langle \nabla \Delta f|}{\langle \mathbf{p}_0 \mid \nabla f_0\rangle} \pi_1$$

$$\Delta\pi_2 = \frac{|\nabla_\mathbf{x} \Delta H\rangle\langle \nabla f_0|}{\langle \mathbf{p}_0 \mid \nabla f_0\rangle} + \frac{|\nabla_\mathbf{x} H_0\rangle\langle \nabla \Delta f|}{\langle \mathbf{p}_0 \mid \nabla f_0\rangle} \pi_1$$

$$\Delta T = \begin{pmatrix} 0 & 0 \\ \Delta T_1 & \Delta T_2 \end{pmatrix},$$

where

$$\Delta T_2 = -\hat{T}_2 \frac{|\nabla \Delta f\rangle\langle \hat{\mathbf{p}}_0 - \mathbf{p}_0|}{\langle \hat{\mathbf{p}}_0 \mid \nabla f_0\rangle}$$

$$\Delta \hat{T}_2 = -\hat{T}_2 \frac{|\nabla \Delta f\rangle\langle \hat{\mathbf{p}}_0|}{\langle \hat{\mathbf{p}}_0 \mid \nabla f_0\rangle}$$

$$\begin{aligned}\Delta T_1 = &-\hat{T}_2 \frac{|\nabla \Delta f\rangle\langle \nabla_\mathbf{x}\hat{H}_0 - \nabla_\mathbf{x} H_0|}{\langle \hat{\mathbf{p}}_0 \mid \nabla f_0\rangle} - \frac{|\nabla f_0\rangle\langle \nabla_\mathbf{x} \Delta\hat{H} - \nabla_\mathbf{x} \Delta H|}{\langle \hat{\mathbf{p}}_0 \mid \nabla f_0\rangle} \\ &+ \frac{\langle \mathbf{p}_0 - \hat{\mathbf{p}}_0 \mid \nabla \Delta f\rangle}{\langle \nabla f_0 \mid \nabla f_0\rangle} \hat{T}_2 F_0 \\ &- \frac{\langle \mathbf{p}_0 - \hat{\mathbf{p}}_0 \mid \nabla f_0\rangle}{\langle \nabla f_0 \mid \nabla f_0\rangle} \{\Delta \hat{T}_2 F_0 + \hat{T}_2 \nabla\nabla\, \Delta f\}.\end{aligned}$$

All these expressions are computed for the reference ray on the reference interface.