Sensitivity of $qP$-wave traveltimes and polarization vectors to heterogeneity, anisotropy and interfaces

Véronique Farra and Soazig Le Bégat

Laboratoire de Sismologie, Institut de Physique du Globe de Paris, 4 Place Jussieu, 75252 Paris cedex 05, France

Accepted 1994 September 20. Received 1994 September 1: in original form 1994 March 28

SUMMARY

Using ray perturbation theory, a linear relationship is derived that describes the first-order perturbation in slowness vectors due to elastic parameter and interface perturbations as well as to source and receiver position perturbations. For layered media, a propagator matrix, which has symplectic properties, is defined in Cartesian coordinates. The method is valid for heterogeneous isotropic or anisotropic media with interfaces, and allows the efficient calculation of Frechet derivatives of the polarization vectors. This makes it possible to use polarization data in tomographic inversion. Explicit integral expressions of the Frechet derivatives of the traveltimes and the polarization vector are given for the $qP$ wave in a transverse isotropic medium. The reference medium is a heterogeneous medium with elliptical anisotropy. A sensitivity study shows that the $qP$-wave polarization vectors provide information on the isotropic component of the slowness.

Key words: polarization vector, ray perturbation theory, transverse isotropy.

1 INTRODUCTION

In recent years, traveltime inversion techniques have been applied to crosshole or VSP geometry (Bois et al. 1972; Dines & Lytle 1979; Wong, Hurley & West 1983; Fehler & Pearson 1984; Ivaness 1985; Peterson, Paulsson & McEvilly 1985 and Bregman, Bailey & Chapman 1989). The first-arrival times are interpreted in terms of inhomogeneous structures, generally 2-D. Most seismic data sets have been interpreted assuming isotropy. Chapman & Pratt (1992) and Pratt & Chapman (1992) developed the theory in order to include weak anisotropy in crosshole tomography experiments. The aperture limitation of these experiments resulted in data that were not sufficient to specify uniquely the medium parameters. Regularization of the system of equations using appropriate constraints (for example smoothing techniques) was essential (Pratt & Chapman 1992). Another approach to regularization is the use of additional data sets that carry independent information on the parameter distribution. Polarization vectors should provide independent information, being sensitive to the gradient of the slowness field along the ray path. Hu & Menke (1992) showed the potential use of the polarization vector in tomographic procedures for smooth isotropic media. The computation of the Frechet derivatives of the polarization vector is more complicated than for traveltimes. In this paper, expressions of the Frechet derivatives for traveltimes and slowness vectors are given for a heterogeneous anisotropic medium with interfaces. We consider perturbations in elastic parameters and interface positions. The perturbation of the canonical vector of a ray admits a general expression using a propagator matrix. We give an expression for the propagator matrix in Cartesian coordinates satisfying the symplectic property, which is very useful for numerical applications. This property is used to perform efficient computations of the Frechet derivatives. Moreover, explicit expressions of the Frechet derivatives are given for the $qP$ wave in a transverse isotropic medium. It is extremely convenient to use a factorized anisotropic inhomogeneous (FAI) medium (Cerveny 1989): in such a medium, the density-normalized elastic parameters $a_{ijkl}$ share the same spatial variations. The concept of the FAI medium not only reduces the number of parameters describing the model, but also simplifies considerably the ray computations (see Cerveny 1989). We use an FAI medium with elliptical anisotropy as the reference medium, and consider transverse isotropic perturbations. Transverse isotropy is suitable for studying most forms of anisotropy that have been obtained in the Earth: periodic thin layering (Backus 1962), aligned cracks (Hudson 1980, 1981) and preferred orientation of a single crystal axis. The transverse isotropy is the only symmetry considered explicitly here, although the present approach could be used for any symmetry.

2 HAMILTONIAN FORMULATION OF RAY TRACING

In the first part of this paper, we recall the Hamiltonian formulation used for ray tracing in anisotropic media (see
Ray and paraxial ray tracings are performed in Cartesian coordinates (see Cerveny 1972), whereas, for ray tracing in isotropic media, the ray-centred coordinate system proposed by Popov & Psenick (1978) is most commonly used. Use of the Cartesian coordinate system increases the size of the system of equations but simplifies considerably the calculations, even in isotropic media. Boundary equations at interfaces can be introduced very easily in this formulation, as well as simple expressions of Frechet derivatives. Moreover, for a layered medium, we derive an expression of the propagator matrix which satisfies symplectic properties, which was not the case for the expression given by Farra (1989) or Gajewsky & Psenick (1990). All the derived results are valid for heterogeneous isotropic or anisotropic media.

2.1 Ray-tracing equations

In the high-frequency approximation, the elastodynamic equation yields a non-linear first-order partial differential equation for the traveltime (the eikonal equation):

$$H(x, p) = 0. \quad (1)$$

The function $H$ is called the Hamiltonian, $p = \nabla T$ is the slowness vector and $T$ is the traveltime. Many suitable forms of the Hamiltonian can be used (see Cerveny 1989 and Farra 1993 for a discussion). For example, in isotropic media, a useful form of the Hamiltonian was proposed by Burridge (1976):

$$H(x, p) = \frac{1}{2} [p^2 - u^2(x)].$$

where the slowness $u(x)$ is the reciprocal of the velocity. The theory developed in this paper is independent of the chosen form of the Hamiltonian and is valid in isotropic as well as anisotropic media. The most common way of solving the eq. (1) is to use the ray-tracing method. A ray is defined by its canonical vector $y(\tau) = (x(\tau), p(\tau))$, where $x(\tau)$ is the position along the ray, $p(\tau)$ is the slowness vector of the wavefront at position $x(\tau)$ and $\tau$ is a sampling parameter which depends on the chosen form of the Hamiltonian (Cerveny 1989). The canonical vector of the rays satisfies Hamilton’s equations:

$$\dot{x} = \nabla_x H,$$

$$\dot{p} = -\nabla_p H, \quad (2)$$

where $\nabla_x$ and $\nabla_p$ denote the gradients with respect to the vectors $x$ (position vector) and $p$ (slowness vector), respectively; dots indicate derivatives with respect to the sampling parameter $\tau$. The six equations in (2) are not independent, since at least one of them may be eliminated by using the fact that the slowness vector $p$ should satisfy the eikonal equation (1).

The Lagrangian can be derived from the Hamiltonian by means of the Legendre transform:

$$L(x, \dot{x}) = \langle p | \dot{x} \rangle - H(x, p),$$

where $\langle a | b \rangle$ denotes the scalar product of vectors $a$ and $b$.

The traveltime is obtained by simple integration of the Lagrangian along the ray path:

$$T(x(\tau), x(\tau)) = \int_{\tau_0}^{\tau_1} L(x, \dot{x}) d\tau = \int_{\tau_0}^{\tau_1} \left[ \langle p | \dot{x} \rangle - H(x, p) \right] d\tau = \int_{\tau_0}^{\tau_1} \langle p | \dot{x} \rangle d\tau. \quad (3)$$

2.2 Paraxial ray-tracing equations

Suppose a ray has been traced in the medium. Around this ray, we can obtain neighbouring rays by means of first-order perturbation theory (Farra & Madariaga 1987). Let $y_\delta(\tau) = (x(\tau), p(\tau))$ be the canonical vector of the central ray. The position of a paraxial ray and its slowness vector are given by

$$x(\tau) = x_0(\tau) + \delta x(\tau),$$

$$p(\tau) = p_0(\tau) + \delta p(\tau). \quad (4)$$

These perturbations are due to small changes $\delta x(\tau_0)$ and $\delta p(\tau_0)$ in the initial values of $x$ and $p$.

The paraxial perturbation $\delta y(\tau) = (\delta x(\tau), \delta p(\tau))$ satisfies the linear system (Farra, Virieux & Madariaga 1989)

$$\delta y = A(\tau) \delta y,$$

where

$$A = \left[ \begin{array}{cc} \nabla_x \nabla_p H & \nabla_x \nabla_p H \\ -\nabla_x \nabla_p H & -\nabla_p \nabla_x H \end{array} \right].$$

Solutions to the linear system (5) may be found by standard propagator techniques (Gilbert & Backus 1966). Given the initial value $\delta y(\tau_0)$, the subsequent evolution of the paraxial canonical vector $\delta y(\tau)$ is given by

$$\delta y(\tau) = P(\tau, \tau_0) \delta y(\tau_0), \quad (7)$$

where $P(\tau, \tau_0)$ is the so-called propagator matrix of the system (5). However, in order to be a paraxial ray, the paraxial solution has to satisfy a condition derived from the perturbation of the eikonal equation $H(x, p) = 0$ (see eq. 1):

$$\delta H = \langle \nabla_x H | \delta x \rangle + \langle \nabla_p H | \delta p \rangle = 0. \quad (8)$$

We can remark that $\delta H$ is constant along any solution of system (5), so that it is sufficient to enforce $\delta H = 0$ at the source in order to satisfy (8) everywhere.

The propagator matrix has very important properties, which are due to the special form of matrix $A$ (see Thomson & Chapman 1985). One can write

$$\det P(\tau, \tau_0) = \det P(\tau_0, \tau) = 1 \quad (9)$$

and

$$P(\tau, \tau_0) = P(\tau, \tau') P(\tau', \tau_0). \quad (10)$$

Moreover, the propagator matrix has symplectic properties (see Thomson & Chapman 1985):

$$P^T J P = J. \quad (11)$$

where $J$ is the antisymmetric matrix

$$J = \left[ \begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array} \right].$$

(12)
and \( I \) is the \( 3 \times 3 \) identity matrix.

Introducing the following notation,

\[
P(\tau, \tau_0) = \begin{bmatrix} P_{01} & P_{12} \\ P_{21} & P_{22} \end{bmatrix},
\]

one can write the inverse of the propagator matrix in a simple way:

\[
P^{-1}(\tau, \tau_0) = P(\tau_0, \tau) = \begin{bmatrix} P_{12} & -P_{12} \\ -P_{21} & P_{22} \end{bmatrix}.
\]

These properties may be very useful in numerical applications (see, for example, Cerveny & Soares 1992; Hubral, Schleicher & Tygel 1993; Coates & Chapman 1990, 1991) and will be used to perform efficient computation of Frechet derivatives.

### 2.3 Transformation across an interface

In order to take into account more complex velocity distributions, the model can be divided into individual layers separated by interfaces of zero order (elastic parameters discontinuities). The presence of these interfaces requires the introduction of appropriate boundary conditions for ray tracing and paraxial ray tracing. Following Farra et al. (1989), we will denote variables associated with the reflected/transmitted ray with a circumflex above them. The new Hamiltonian will be, for example, \( \tilde{H} \). Let the interface location \( x \) be defined by the general relation \( f(x) = 0 \); \( \nabla f \) defines the local normal to the interface. Let us consider a reference ray whose canonical vector is \( \mathbf{y} \). This reference ray hits the interface at \( \mathbf{x} \). At the interface, the ray satisfies the following equations (Farra et al. 1989; Farra 1989):

\[
\begin{aligned}
\mathbf{x}_i &= \mathbf{x}, \\
\mathbf{p}_i \times \nabla f &= \mathbf{p}_i \times \nabla f, \\
\tilde{H}(\mathbf{x}_i, \mathbf{p}_i) &= H(\mathbf{x}_i, \mathbf{p}_i),
\end{aligned}
\]

where the cross-product has been denoted by \( \times \). These relations give the initial conditions of the reflected/transmitted ray in the new medium.

Let us now consider the continuity relations of the paraxial rays. The transformation of the canonical vector \( \delta \mathbf{y} = (\delta \mathbf{x}, \delta \mathbf{p}) \) at the discontinuity is given by (see equation 16, Appendix A)

\[
\delta \mathbf{y} = \mathbf{M} \delta \mathbf{y},
\]

where the elements of the transformation matrix \( \mathbf{M} \) (given in Appendix A) are computed on the reference ray at \( \mathbf{x}_i \). To first order, rays in the neighbourhood of the reference ray satisfy relation (7), where \( \mathbf{P}(\tau, \tau_i) \) is the generalized propagator taking into account transformation matrices at the interfaces crossed by the central ray:

\[
\mathbf{P}(\tau, \tau_0) = \mathbf{P}(\tau_0, \tau_0) \prod_{i=1}^{N} \mathbf{M}_i \mathbf{P}(\tau_i, \tau_{i-1}),
\]

where \( N \) is the number of interfaces crossed by the reference ray between \( \tau_0 \) and \( \tau \), and \( \tau_i \) is the value of the \( \tau \) parameter of the central ray at the intersection point with the interface \( i \). \( \mathbf{M}_i \) is the interface transformation matrix. Let us remark that the generalized propagator (17) satisfies the properties (9–14) for the expression of matrix \( \mathbf{M} \) given in Appendix A. This is not the case for the transformation matrix given by Farra et al. (1989) or Gajewski & Psencik (1990).

### 2.4 Ray perturbation theory

Let us assume that we have a reference model characterized by the Hamiltonian \( H \). Moreover, we assume that a ray has been defined by its canonical vector \( \mathbf{y} = (\mathbf{x}, \mathbf{p}) \). Let us consider a perturbation of the model such that the Hamiltonian is changed from \( H \) to \( H + \Delta H \), where \( \Delta H \) is the Hamiltonian for the reference medium. Capital \( \Delta \) will be used to denote perturbations due to the structure. To first order, it is possible to obtain rays in the perturbed medium that deviate slightly from the reference ray \( \mathbf{y} \). (Farra & Madariaga 1987). We introduce the perturbed canonical vector \( \mathbf{y} = \mathbf{y} + \Delta \mathbf{y} \) of these rays. The equations of the perturbed rays (Farra 1989) are given by

\[
\delta \mathbf{y} = \mathbf{P}(\tau, \tau_i) \Delta \mathbf{y}(\tau) + \int_{\tau_i}^{\tau} \mathbf{P}(\tau, \tau') \Delta \mathbf{B}(\tau') d\tau',
\]

where \( \Delta \mathbf{B} \) is the initial perturbation. \( \mathbf{B}(\tau, \tau) \) is the paraxial propagator computed along the reference ray in the reference medium, and

\[
\Delta \mathbf{B} = \begin{bmatrix} \nabla \Delta H \\ -\nabla \Delta H \end{bmatrix}.
\]

Moreover, the perturbation \( \Delta \mathbf{y} = (\Delta \mathbf{x}, \Delta \mathbf{p}) \) should satisfy a condition derived from the eikonal equation (1):

\[
H = (\nabla H, H) + (\nabla H, \Delta \mathbf{p}) + \Delta H = 0,
\]

where the Hamiltonian \( H \), its partial derivatives and \( \Delta H \) are computed for \( \mathbf{y} \). We remark that \( H \) is constant along any solution (18), so that it is sufficient to enforce \( H = 0 \) at the source in order to satisfy (20) everywhere.

Let us now consider a perturbation of the interface. We denote by \( f(x) \) the reference interface and by \( f(x) + \Delta f(x) \) the perturbed interface. Consider a reference ray with canonical vector \( \mathbf{y}(\tau) \) and a ray in the perturbed medium that propagates in the neighbourhood of this reference ray. In order to propagate the transmitted or reflected perturbed ray away from the interface, we choose as the new reference ray, the reflected/transmitted ray in the unperturbed medium corresponding to the reference incident ray. The canonical vectors of the two reference rays are connected by relations (15) at the interface in the reference medium. We denote by \( \tilde{\mathbf{y}}(\tau) \) and \( \mathbf{y}(\tau) \) the canonical vectors of the reflected/transmitted reference ray and the reflected/transmitted perturbed ray.

The general transformation of the perturbed canonical vector at the interface is then (Appendix B)

\[
\Delta \mathbf{y} = \mathbf{M}_0 \Delta \mathbf{y} + \Delta \mathbf{y},
\]

where \( \mathbf{M}_0 \) is the paraxial transformation matrix (16) computed in the reference medium and \( \Delta \mathbf{y} \) is a perturbation term given in Appendix B. The transformation (21) contains two terms. The first one is the same as the linear transformation (16) connecting the incident and reflected/
transmitted paraxial rays at the interface. This term takes into account perturbations in initial conditions and in medium parameters between the source and the interface. \( \Delta y \) is due to the displacement \( \Delta f \) of the interface and to the perturbation of Snell's law.

Given the initial perturbation \( \Delta y(\tau) \), the subsequent evolution of the perturbed canonical vector \( \Delta y(\tau) \) is given by

\[
\Delta y(\tau) = P_0(\tau, \tau)\Delta y(\tau) + \Delta y_0(\tau),
\]

(22)

where \( P_0(\tau, \tau) \) is the generalized propagator (17) computed along the reference ray in the reference medium. The perturbation \( \Delta y_0 \) is the solution of the correction problem (25) (see 23).

The perturbations \( \Delta y_0 \) and \( \Delta y_0 \) are related by the equation

\[
\Delta y_0(\tau) = P_0(\tau, \tau)\Delta y(\tau),
\]

(25)

so one can easily obtain the perturbation vectors \( \Delta y_0 \) and \( \Delta y_0 \) knowing the propagator \( P_0 \) and the perturbation \( \Delta y_0 \). In expression (24), both the integrated term and the terms in the sum are independent of the parameter \( \tau \). The perturbation \( \Delta y_0(\tau) \) is modified only when the reference ray crosses a perturbed region of the medium. This is not the case for the perturbation vector \( \Delta y_0(\tau) \): because of its explicit dependence on the parameter \( \tau \) (see 23), one has to compute the integral along the whole ray, even if the perturbed region is localized in the vicinity of the source.

2.5 Boundary conditions

Let us assume that a ray with canonical vector \( y(\tau) \) has been defined in the phase space such that the points \( x_0 \) and \( x_1 \), correspond to \( x(\tau) \) and \( x(\tau) \), respectively. First-order perturbation theory can be used in order to find, in the vicinity of this reference ray, the ray of the perturbed medium that connects the points \( x' = x + dx \), and \( x' = x + dx \). This ray is defined in the phase space by its canonical vector \( y(\tau) = (x(\tau), p(\tau)) \) with \( x(\tau) = x \) and \( x(\tau) = x \). We remark that the sampling parameter \( \tau \) is generally different from the parameter \( \tau \). To first order, the ray between \( x' \) and \( x' \) can be described by the perturbation of its canonical vector \( \Delta y = y - y_0 \), which satisfies the expressions (20) and (22).

Let us define the perturbation \( d y = y(\tau) - y_0(\tau) \) of the ray canonical vector at \( x' \). The perturbation \( d y = (d x, d p) \) is linearly related to \( \Delta y(\tau) = y(\tau) - y_0(\tau) \) by the projection matrix \( P' \) that extrapolates \( \Delta y(\tau) \) on an arbitrary plane containing the points \( x_0 \) and \( x_1 \) (see Farra et al. 1989; Farra 1989, 1993). For convenience, the normal vector \( n_0 \) to this plane may be chosen such that vectors \( d x, \dot{x}_0 = \nabla y_0(\tau) \) and \( \dot{n}_0 \) are coplanar and \( (n_0 | dx) = 0 \). In the case \( \Delta x = 0 \), this plane is chosen such that the normal vector is \( n_0 = \nabla y_0(\tau) \).

Thus, the perturbed canonical vector \( \Delta y \) should satisfy the boundary conditions

\[
\Delta x(\tau) = dx_x,
\]

(26)

\[
\Delta x(\tau) = dx_y,
\]

where \( \Delta x, \Delta x, \) is one submatrix of the projection matrix \( P' \):

\[
\Delta x, = I - \frac{\nabla y_0(\tau)}{\nabla y_0(\tau) | n_0},
\]

(27)

where the notation \([a/b]\) denotes the outer product of vectors \( a \) and \( b \). Because of the relation (22) between \( \Delta y(\tau) \) and \( \Delta y(\tau) \), the two-point boundary problem corresponding to boundary conditions (26) can be written as the following system of four equations and three unknowns [the three components of \( \Delta p(\tau) \)]:

\[
\begin{align*}
\Delta p', P'_{11} \Delta p(\tau) &= dx_x - \Delta x, P'_1 dx_x + \Delta x, \\
&= \Delta x, P'_1 dx_x - \Delta x, \\
\n\n\n&= \langle \nabla y_0(\tau) | dx_x, \Delta x, \rangle - \Delta H',
\end{align*}
\]

(28a)

where we used the partition (13) of the propagator \( P_0(\tau, \tau) \) and \( \Delta y_0(\tau) = (\Delta x, \Delta p_0) \) is given by (23). Eq. (28b) is the eikonal equation (20) written at \( \tau \). In (28b), the subscript 'r' or 'l' indicates that the function is computed at \( \tau \) and \( \tau \), respectively. Noting that the projection matrix \( \Delta p' \) is of rank 2, eqs (28a) provide two independent equations for the determination of \( \Delta p(\tau) \). Therefore, the linear system (28b) is not overdetermined and is equivalent to three independent equations.

We can therefore obtain the perturbation \( \Delta p \), of the slowness vector at \( x' \):

\[
d p' = \Delta p(\tau) + \Delta p', \Delta x(\tau)
\]

(29)

where \( \Delta p' \) is another submatrix of the projection matrix \( P' \) (see Farra et al. 1989 and Appendix A). In this way, we can obtain the slowness vector perturbation due to initial and final perturbations, \( dx_x \), and \( dx_y \), as well as medium perturbations.

In order to compute the Frechet derivatives of the slowness vector, one has to solve eqs (28) and (29) with \( dx_x = dx_x = 0 \):

\[
\begin{align*}
\Delta p', P'_{11} \Delta p(\tau) &= -\Delta p', P_1 \Delta x', + \Delta p', P_1 \Delta p', \\
&= \langle \nabla y_0(\tau) | \Delta p(\tau) \rangle = -\Delta H'
\end{align*}
\]

(30)
Thus the traveltime perturbation can be computed to first order in the parameters perturbation by simple integration of perturbation terms \( \nabla_{\phi} \Delta H \), \( \nabla_{\psi} \Delta H \) and addition of interface terms \( \Delta \gamma \).

### 3 TRAVELTIME PERTURBATION IN ANISOTROPIC MEDIA

The expression for the traveltime perturbation in smooth anisotropic media has been obtained by Cerveny (1982) and Cerveny & Jech (1982) (see also Jech & Psencik 1989 and Nowack & Psencik 1991). In this section, the first-order perturbation of traveltime is derived, whatever the form of the Hamiltonian may be. Perturbations of the medium as well as perturbations of interface position are considered.

Suppose a ray has been traced in the unperturbed medium characterized by the Hamiltonian \( H_0 \). Let \( y_0(\tau) = (x_0(\tau), p_0(\tau)) \) be the canonical vector of this reference ray. We consider a perturbation of the model such that the Hamiltonian is changed from \( H_0 \) to \( H = H_0 + \Delta H \). For rays in the perturbed medium that deviate slightly from the reference ray, we can write, to first order,

\[
\langle p | \dot{x} \rangle = (p_0 | \dot{x}_0) + (p_0 | \Delta \dot{x}) + (\Delta p | \dot{x}_0) = H_0 + \langle \nabla_{\phi} H_0 | \Delta x \rangle + \langle \nabla_{\psi} H_0 | \Delta p \rangle + \Delta H.
\]

where \( H_0 \), its first derivatives and \( \Delta H \) are computed at \( y_0 = (x_0, p_0) \).

From expressions (3) and (33), we can write the traveltime along the perturbed ray as

\[
\langle x(\tau), x(\tau) \rangle_T = \int_{\tau}^{\tau_f} [\langle p_0 | \dot{x}_0 \rangle - H_0] \, d\tau
\]

\[
- \int_{\tau}^{\tau_f} \Delta H \, d\tau + [\langle p_0 | \Delta \dot{x} \rangle]_i
\]

\[
= T_0(x_0(\tau), x_0(\tau)) - \int_{\tau}^{\tau_f} \Delta H \, d\tau
\]

\[
+ \langle p_0(\tau) | \Delta \dot{x}(\tau) \rangle - \langle p_0(\tau) | \dot{x}(\tau) \rangle.
\]

where we use the ray equations (2) and integration by parts. Thus the traveltime perturbation can be computed to first order in the parameters perturbation by simple integration along the reference ray.

For fixed end points, the traveltime perturbation is given by

\[
\Delta T(x, x) = -\int_{\tau}^{\tau_f} \Delta H \, d\tau.
\]

where \( \tau, \) and \( \tau_f \), are the sampling parameters of the reference ray at positions \( x_0 \) and \( x_0 \), respectively. This expression, which is valid in anisotropic media, is a generalization of Fermat’s principle.

Moreover, if we consider interface perturbations, we can write to first order in the perturbation (Appendix C):

\[
\Delta T(x, x) = -\int_{\tau}^{\tau_f} \Delta H \, d\tau
\]

\[
+ \sum_{i=1}^{N} \left[ \beta_i \cdot \frac{p_0 | \nabla f_i \rangle}{\langle \nabla f_i | p_0 \rangle} \right] \Delta f_{i}.
\]

where \( \Delta f_i \) is the perturbation of the interface number \( i \). For rays in the perturbed medium that deviate slightly from the reference ray, we can write, to first order,

\[
\langle p | \dot{x} \rangle = (p_0 | \dot{x}_0) + (p_0 | \Delta \dot{x}) + (\Delta p | \dot{x}_0) = H_0 + \langle \nabla_{\phi} H_0 | \Delta x \rangle + \langle \nabla_{\psi} H_0 | \Delta p \rangle + \Delta H.
\]

where \( \Delta H \) is the perturbation of the Hamiltonian.

4 FRECHET DERIVATIVES OF qP-WAVE TRAVELTIMES AND POLARIZATION VECTORS IN TRANSVERSE ANISOTROPIC MEDIA

The expressions obtained in the preceding sections are valid for any heterogeneous isotropic or anisotropic media with interfaces. The explicit expressions of perturbed rays in a slightly anisotropic medium without interfaces have been given by Nowack & Psencik (1991). We consider here a medium with transverse isotropy (TI medium) for which the partial derivatives of the traveltime and the polarization vector have very simple expressions. For such a medium, we use the Thomsen (1986) parameter set, which simplifies the Hamiltonian expression (Farra 1989, 1990). Only \( qP \) waves are considered. The parameters are \( \eta, \) (the square of \( qP \) phase slowness along the symmetry axis), \( \nu = \eta/\mu (\text{the square of the ratio of } qP \text{ to } qS \text{ phase slowness along the symmetry axis}) \), and \( \Xi = \delta - \tau \) (non-dimensional parameters which describe the amount of anisotropy and the anellipticity of the slowness sheets, respectively), and two angles that describe the orientation of the axis of symmetry. Following Farra (1989, 1990), the \( qP \)-wave Hamiltonian (1) for small values of \( \Xi \) is given by

\[
H(x, p) = \frac{1}{2} \left( p^T A p - u_0^2(x) + \Xi p^T M_p p + M_p p \right).
\]
For an axis of symmetry along the z-direction, the 3 × 3 matrices \( A, M, B, \) and \( M \), are given by

\[
M_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad M_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{38}
\]

where \( \xi = \epsilon (1 - \nu^2)^{-1} \) is a parameter proportional to \( \epsilon \). In the general case, for an arbitrary symmetry axis orientation, these matrices can be obtained by using rotation matrices (see Farra 1990).

We will assume that the medium is a factorized TI medium (FA1 medium). In a FA1 medium (see Cerveny 1989), the density-normalized elastic parameters \( a_{ik} \) share the same spatial variations. Therefore, in each layer, the parameters \( \nu, \epsilon \) and \( \Xi \) are constants and the symmetry-axis direction is fixed. However, these parameters may change from layer to layer. Moreover, we use an FA1 elliptical medium as the reference medium \( (\Xi = 0) \). For such a medium, rays and propagator matrices are no more complicated to compute than in an isotropic medium, which corresponds to the case \( \epsilon_i = 0 \) (see Farra 1989, 1990).

Let us assume that the perturbed model is described by the set of parameters \( m = (u^2, \nu, \epsilon, \Xi) \), where \( u^2 \) are the coefficients of the B-spline interpolation of the square of slowness \( \nu^2 \) (see Farra & Madariaga 1988 for a description of the B-spline interpolation):

\[
\nu^2(x) = \sum_i u_i B_i(x). \tag{39}
\]

One of the advantages of the B-spline interpolation is the local support of the blending functions \( B_i \).

For a reference elliptical anisotropic model, the partial derivatives of the Hamiltonian (37) are given by the expressions

\[
\frac{\partial H}{\partial u^2} = -\frac{1}{2} B_i, \quad \frac{\partial H}{\partial \nu^2} = 0, \quad \frac{\partial H}{\partial \epsilon} = \frac{1}{2} p^T M(p), \quad \frac{\partial H}{\partial \Xi} = \frac{1}{2} p^T M(p) p. \tag{40}
\]

The partial derivative of the traveltime with respect to parameter \( m \), is given by the relation

\[
\frac{\partial T}{\partial m} = -\int_t \frac{\partial H}{\partial u^2} dt. \tag{41}
\]

The partial derivative of the traveltime with respect to the position of the interfaces may be easily obtained from relation (36). Generally, a B-spline representation is also used for the reflector depths.

We remark that the traveltime perturbation will be most sensitive to the perturbation of the parameter \( \epsilon \) for propagation orthogonal to the symmetry axis, and to the perturbation of the parameter \( \Xi \) for propagation close to 45° from this axis. The \( qP \) traveltime is not sensitive to perturbation of the parameter \( \nu^2 \), the corresponding partial derivative being zero.

The polarization vector of the \( qP \) wave is the normalized, i.e. unit, eigenvector of the Christoffel matrix, denoted \( \Gamma \), corresponding to the eigenvalue given to first order in \( \Xi \) by

\[
G_p(x, p) = u^2 \gamma(x) \left( p^T A_p + \Xi b^T M_p p^T M_p p \right), \tag{42}
\]

with \( G_p(x, p) = 1 \) (see Farra 1989).

The polarization vector, denoted \( g_p \), satisfies the Christoffel equations:

\[
(\Gamma - G_p) g_p = 0 \tag{43}
\]

where \( \Gamma, G_p \) and \( g_p \) are functions of position \( x \), slowness vector \( p \), and of the model \( m \).

For TI media with a symmetry axis along the \( z \)-direction, the matrix \( \Gamma \) has a simple form (see Daley & Hron 1977). Exact expressions for the polarization vector may be obtained in this case. Let us denote

\[
k_\nu = p \left[ 1 + \frac{\epsilon}{1 - \nu^2} + \frac{\Xi}{1 - \nu^2} \left( 1 + \xi \right) p^2 + p^2 \right].
\]

\[
k_\epsilon = p \sqrt{1 + \frac{\epsilon}{1 - \nu^2}}.
\]

where \( p = \sqrt{p_1^2 + p_2^2} \). The \( qP \)-polarization vector \( g_p \) is in the plane containing the symmetry axis and the slowness vector \( p \). Its components are given to first order in \( \Xi \) by

\[
g_\nu = \frac{k_\nu}{\sqrt{k_\nu^2 + k_\epsilon^2}} p_1, \quad g_\epsilon = \frac{k_\epsilon}{\sqrt{k_\nu^2 + k_\epsilon^2}} p_2, \quad g_\Xi = \frac{k_\nu}{\sqrt{k_\nu^2 + k_\epsilon^2}}.
\]

where \( p = \sqrt{p_1^2 + p_2^2} \). For a TI medium with an arbitrary orientation of the symmetry axis, the polarization vector can be obtained by using a rotation matrix.

The perturbation \( d g_p \) of the polarization vector \( g_p \) due to a perturbation \( dm \), is obtained by differentiation at fixed position \( x \):\n
\[
d g_p = \left( \nabla_p g_p \cdot \frac{\partial p}{\partial m} + \frac{\partial g_p}{\partial m} \right) dm. \tag{46}
\]

Let us assume that the model is described by a set of parameters \( m = (m_i) \). \( m_i \), may represent the B-spline coefficients of the square of phase slowness \( \nu^2 \), the parameter \( \nu^2 \), the anisotropic parameters \( \epsilon \) and \( \Xi \), and the reflector B-spline coefficients. In order to compute the Frechet derivatives \( \partial P/\partial m \), of the slowness vector, one has to solve the eqs (30) and (31) for every perturbation \( dm \), with \( \Delta H = (\partial H/\partial m_i) dm_i \). We note that the use of B-splines accelerates the computation of the traveltime partial derivatives and the computation of the perturbation vector \( \Delta y_p \) because most of the splines are equal to zero except for those that have knots in the neighbourhood of the ray. Only the corresponding partial derivatives are computed when
parameter perturbations by analysing the singular values (SVD) problem (48). We use the singular-value decomposition sensitivity of traveltimes and polarization vectors to parameter perturbations, where 

\[ \mathbf{D} \text{ of } \mathbf{S} \text{ and } \mathbf{g} \text{ respectively. In addition, it may be appropriate to constrain the solution by introducing a priori information (see Farra & Madariaga 1988). Additional information is needed because tomography is intrinsically unstable due to the aperture limitation of the experiments, as well as to the poor resolution near the borders of the model (see Pratt & Chapman 1992).}

It is not our purpose in this paper to discuss the techniques available to solve this inverse problem. Many methods exist for obtaining a solution and they have been extensively discussed in the literature. The non-linear least-squares problem (47) can be solved iteratively by the Gauss-Newton method, which linearizes expression (47) around the current model to obtain a linear least-squares problem:

\[ E(\mathbf{m}) = \sigma_T^{-2} \| \mathbf{T}(\mathbf{m}) - \mathbf{g}_o \| + \sigma_g^{-2} \| \mathbf{g}_p - \mathbf{g}_p(\mathbf{m}) \|^2. \]  

(47)

where \( \mathbf{T} \) and \( \mathbf{g}_o \) are the observed \( qP \) traveltimes and polarization vectors. \( \sigma_T \) and \( \sigma_g \) represent the estimated uncertainties of observed traveltimes and polarization data, respectively. In addition, it may be appropriate to constrain the solution by introducing a priori information (see Farra & Madariaga 1988). Additional information is needed because tomography is intrinsically unstable due to the aperture limitation of the experiments, as well as to the poor resolution near the borders of the model (see Pratt & Chapman 1992).

The least-squares solution of the linearized problem (48) is

\[ \Delta \mathbf{m} = (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \begin{bmatrix} \sigma_T^{-1} \Delta T \\ \sigma_g^{-1} \Delta \mathbf{g}_p \end{bmatrix}. \]  

(50)

In order to study the properties of the linearized inverse problem (48), we use the singular-value decomposition (SVD) approach (Jackson 1972). We will study the sensitivity of traveltimes and polarization vectors to parameter perturbations by analysing the singular values and the associated eigenvectors of the Jacobian matrix \( \mathbf{D} \) for simple models. The singular values \( \lambda_i \) of the Jacobian matrix and the corresponding eigenvectors \( \mathbf{w}_i \) of the parameter space define the parameter combinations that are well determined or not by the data. A perturbation proportional to \( \mathbf{w}_i \) will be well determined if the corresponding \( \lambda_i \) is large.

In the first example that we consider, the reference model is isotropic with a constant velocity. The simulated experiment is as follows. Two sources were located on the surface at 0.2 km and 0.9 km away from the borehole. For each source, nine receivers were regularly distributed every 100 m in the borehole, the first one being located at 100 m depth. Fig. 1 shows the ray paths we consider in the following model study. One can see that there is a good distribution of the ray angles in the experiment.

The perturbation of the square of slowness \( \Delta u_p \) is defined by a 1-D B-spline expansion in depth:

\[ \Delta u_p(z) = \sum_i \Delta u_{p_i} B(i, z), \]

(51)

with six knot points \( z = (z_1, z_2, \ldots, z_6) \) distributed evenly at every 200 m. Parabolic B-splines were used in expansion (51): if we take a value \( z \) in some interval \([z_i, z_{i+1}]\), the sum in (51) contains only three terms \( i \leq j \leq i + 2 \), because of the locality properties of B-splines. Moreover, B-spline coefficients model the function they represent (see de Boor 1978, Chap. 1X): we have the important result that

\[ \Delta u_p^2 = \Delta u_{p_1}^2 \left( \frac{z_i + z_{i+1}}{2} \right). \]

We thus have seven parameters to represent \( \Delta u_p \) in the medium. Moreover, we consider constant perturbations of the anisotropic parameters \( \varepsilon \) and \( \Xi \). The symmetry axis was assumed to be vertical and we did not consider any perturbation of parameter \( v^2 \). Thus we used a total of nine parameters to represent the model. For each source-receiver pair, we computed the partial derivatives of the traveltimes and polarization vector with respect to the parameters \( \varepsilon \), \( \Xi \) and \( u_p \). Because of the symmetry of the problem, the \( qP \) polarization vector remains in the vertical

---

**Figure 1.** The ray geometry used to generate the linear system analysed in Figs 2, 3 and 4. The reference model is an isotropic model with constant velocity.
plane containing the rays. Hence, only the vertical component of the polarization vector is used as data.

First, we consider the case $\sigma_0^2 = 0$ (only traveltimes are used). Fig. 2 shows the corresponding nine singular values normalized to the largest one. The corresponding eigenvectors in the parameter space are plotted above. The computed normalized singular values spread over an interval between 1 and $2 \times 10^{-3}$. One can see that the largest singular values correspond to combinations of 'isotropic parameters' $\Delta u^2_\phi$; the singular values corresponding to higher frequency slowness variations are smaller than those corresponding to lower frequency variations. Parameter combinations corresponding to anisotropic parameters are associated with intermediate singular values; since the parameter $\varepsilon$ is associated with a larger singular value, it should be better resolved than the parameter $\Xi$. One can see the poor resolution at the borders of the model: the eigenvectors associated with the two smallest singular values correspond to the parameters $\gamma^2_1$ and $\gamma^2_3$, respectively. If the number of nodes in the B-spline expansion is increased, the general pattern of the SVD is not modified: the singular values corresponding to the eigenvectors that are sensitive to anisotropic parameters keep the same values; the higher the frequency of the slowness variations, the smaller the corresponding singular value. If the sampling interval of the function $\Delta u^2_\phi$ is smaller than the distance between two successive receivers, very small singular values (less than $10^{-4}$) appear which are associated with very high frequency variations of $\Delta u^2_\phi$.

We now consider the case $\sigma_0^1 = 0 \text{ s}^{-1}$ (only the vertical component of polarization vectors is used). Fig. 3 shows the corresponding SVD. One can see that a homogeneous perturbation of all the parameters $\gamma^2_j$ has no influence on the polarization vector. For such a perturbation, ray paths and polarization vectors are not modified. However, the polarization vector is sensitive to local variations of parameter $\gamma^2_4$ which correspond to the largest singular values. Anisotropic parameters are associated with normalized singular values of the order of $10^{-2}$ and have less influence on the polarization data. If the number of nodes in the B-spline expansion is increased, the general pattern of the SVD is not modified. If the sampling interval of the function $\Delta u^2_\phi$ is smaller than the distance between two successive receivers, very small singular values (less than $10^{-4}$) appear which are associated with very high-frequency variations of $\Delta u^2_\phi$.

Figure 4 shows the SVD obtained when $\sigma_0^1/\sigma_0 = 0.1$ s (the two sets of data are used). Compared with Fig. 2, introducing the polarization data increased the singular values corresponding to eigenvectors that are combinations of the parameters $\Delta u^2_j$. The singular values corresponding to eigenvectors that are sensitive to parameters $\varepsilon$ and $\Xi$ are not modified by the introduction of the polarization vector information. Thus the polarization data provide mostly information on the isotropic component of the slowness.

The second model that we consider has two isotropic layers with a constant velocity. The reflector is horizontal and is at 1 km depth. The simulated experiment is as follows. Two sources were located on the surface, at 0.2 and 0.9 km away from the borehole. For each source, 19 receivers were regularly distributed every 100 m in the borehole, the first one being located at 100 m depth. Fig. (5) shows the ray paths we consider in the following model study.

In every layer, the perturbation of the square of slowness $\Delta u^2_\phi$ is defined by the 1-D B-spline expansion (51) with nodes distributed every 200 m. We have seven parameters to
receiver pair. We computed the partial derivatives of the parameters to represent the model. For each source different from layer to layer. The symmetry axis considering constant perturbations of the anisotropic parameters values are normalized to the largest one. The parameters are \( \varepsilon, \Xi, u_x^2 \) and seven parameters describing the B-spline interpolation of \( u_x(z) \).

![Eigenvectors](image)

**Figure 4.** Eigenvectors and associated singular values for traveltimes and polarization vector tomography using the experiment shown in Fig. 1. The two sets of data are used simultaneously. The singular values are normalized to the largest one. The parameters are \( \varepsilon, \Xi \) and seven parameters describing the B-spline interpolation of \( u_x(z) \).

represent \( \Delta u_x^2 \) in each layer. Moreover, in each layer, we consider constant perturbations of the anisotropic parameters \( \varepsilon \) and \( \Xi \). However, these perturbations may be different from layer to layer. The symmetry axis is assumed to be vertical. The reflector depth perturbation is defined by a constant \( \Delta Z \). Thus we used a total number of 19 parameters to represent the model. For each source-receiver pair, we computed the partial derivatives of the traveltime and polarization vector with respect to the parameters \( \varepsilon, \Xi, u_x^2 \) in each layer, as well as the partial derivatives with respect to the interface depth \( Z \). The units of \( u_x^2 \) and \( Z \) are \( s^2 \) km \(^2\) and km, respectively.

We consider the case \( \sigma_{\varepsilon, \Xi} = 0 \) (only traveltimes are used) and the case \( \sigma_{\varepsilon, \Xi} = 0 \) (only the vertical component of the polarization vector is used). Figs 6 and 7 show the corresponding singular values and eigenvectors for each case. Considering Fig. 6, \( \sigma_{\varepsilon, \Xi} = 0 \), one can see that the largest singular values correspond to combinations of 'isotropic parameters' \( \Delta u_x^2 \) and \( \Delta Z \). The singular values corresponding to higher frequency slowness variations are smaller than those corresponding to lower frequency variations. Parameter combinations corresponding to anisotropic parameters in the upper part of the model are associated with intermediate eigenvalues. The eigenvectors associated with the smallest singular values are related to edge effects and to anisotropic parameters in the lower layer because of the aperture limitation of the experiment in this part of the medium (see Fig. 5).

Analysing the SVD obtained for polarization data alone (see Fig. 7), one can see that a homogeneous perturbation of all the parameters \( u_x^2 \) has no influence on the polarization vectors. For such a perturbation, rays and polarization vectors are not modified. The largest singular values correspond to local perturbations of the square of slowness \( u_x^2 \). Parameter combinations that contain interface perturbations are associated with intermediate singular values. The smallest singular values correspond to eigenvectors sensitive to parameters \( \varepsilon \) and \( \Xi \) in the lower layer. The parameter \( \varepsilon \) in the lower layer is associated with a very small singular value (of the order of \( 10^{-4} \)). This is due to the very poor distribution of ray angles in the lower layer. Introducing the polarization vector information will therefore provide information mostly on the isotropic component of the slowness.

**6 CONCLUSION**

In this paper, we extended previous works by Farra et al. (1989) and Farra (1989, 1990) on ray perturbation theory. The approach to ray theory is based on a Hamiltonian formulation. The formulation is independent of the chosen form of the Hamiltonian and is valid in isotropic or anisotropic media with interfaces. An expression of the generalized propagator matrix, which has symplectic properties, was obtained in Cartesian coordinates. Propagator matrices are essential in many applications of ray theory to seismological problems (see Cerveny, Klimes & Psencik 1988). They are very useful for solving two-point ray tracing (Virieux & Farra 1991) and for computing amplitudes as well as for obtaining rays in a perturbed medium (Farra & Madariaga 1987). The symplectic property can be very useful in numerical applications (see, for example, Cerveny & Soares 1992; Hubral et al. 1993; Coates & Chapman 1990, 1991), especially in the computation of rays in a perturbed medium.

An application of the perturbation approach is the calculation of sensitivity operators for tomographic studies in heterogeneous media. In this paper, we proposed the use of polarization data in tomography. We showed how to compute the Frechet derivatives of the polarization vector in
Figure 6. Eigenvectors and associated singular values for traveltimc tomography using the experiment shown in Fig. 5. The parameters are $\varepsilon$, $\Xi$ and seven parameters describing the B-spline interpolation of $u_i^2(z)$ in each layer plus the depth of the reflector $Z^1$.

a general anisotropic medium with interfaces. We considered perturbations of source and receiver positions, of elastic parameters and of interface position. The method presented clearly shows that the polarization data are sensitive to the gradient of the perturbation terms and therefore add independent information to traveltime data. Moreover, we gave explicit expressions of partial derivatives of the $qP$-wave traveltime and polarization vector in a factorized (FAI) TI medium. The concept of the FAI medium not only reduces the number of parameters describing the model, but also simplifies considerably the ray computations (see Cerveny 1989; Farra 1989, 1990). A factorized TI medium may be described by a function of coordinates $u_i^2(x)$ (the square of $qP$ phase slowness along the symmetry axis), and by four parameters which are independent of position $x$. We have assumed that the orientation of the symmetry axis is known and is the same in the reference and perturbed media. The Frechet derivatives of the $qP$-wave traveltime and polarization vector are easy to compute if the reference (background) medium is an FAI

Figure 7. Eigenvectors and associated singular values for polarization vector tomography using the experiment shown in Fig. 5. Only the vertical component of the polarization vector is used. The parameters are $\varepsilon$, $\Xi$ and seven parameters describing the B-spline interpolation of $u_i^2(z)$ in each layer plus the depth of the reflector $Z^1$. 
medium with elliptical anisotropy: the $qP$-wave traveltimes and slowness vectors are only sensitive to perturbations of the parameter $\beta^2$ and of the two anisotropic parameters $\theta$ and $\Xi$. For more general TI media with a known axis of symmetry, the Frechet derivatives are easy to obtain by using the same form of the Hamiltonian and the same parametrization.

We studied the sensitivity of traveltimes and polarization vectors to parameter perturbations for simple models. Using singular-value decomposition, we found that the traveltimes and the polarization vectors are mostly sensitive to the ‘isotropic component’ of the medium. Long-wavelength components of the parameter $\beta^2$ have a greater influence on the traveltimes, whereas the polarization data are mostly sensitive to local variations of the parameter $\beta^2$. Traveltime data of reflected waves, as well as polarization data, are also sensitive to the position of the reflecting interface. In traveltime tomography, as well as in polarization tomography, the singular values associated with the anisotropic parameters $\theta$ and $\Xi$ are strongly dependent on the distribution of ray angles. The polarization data can be used to invert for medium parameters simultaneously with traveltimes and can improve the quality of the model.

REFERENCES


APPENDIX A: PARAXIAL TRANSFORMATION MATRIX AT INTERFACE

Let the interface be defined by the general relation $f(x) = 0$. Let us consider a reference ray whose canonical vector is
This reference ray hits the interface at point 0 of coordinates $(x_0, y_0)$ with a local slowness vector $p_0$. The reflected/transmitted reference ray is given by its canonical vector $\delta y = (\delta x, \delta p)$ which is related to $y_0$ at the interface by relations (15). We consider a paraxial ray of this reference ray that intersects the same discontinuity at point 1 of position $(x_1, y_1)$, with sampling parameter $\tau$. Let the paraxial canonical vector at sampling parameter $\tau$ be $\delta y = (\delta x, \delta p)$. A first-order linear transformation denoted by $T$ gives the canonical vector $d y_{1} = (d x, d p)$ of any paraxial ray along the interface (Farra et al. 1989; Farra 1989):

$$dy = T \delta y(\tau),$$

with

$$T = \begin{bmatrix} \pi_1 & 0 \\ \pi_2 & 1 \end{bmatrix},$$

where the $3 \times 3$ submatrices are given by

$$\pi_1 = \frac{(\nabla \cdot n)(\nabla f)}{\nabla \cdot n},$$

$$\pi_2 = \frac{(\nabla \cdot n)(\nabla f)}{\nabla \cdot n}.$$

The notation $[a](b)$ represents a $3 \times 3$ matrix obtained by the outer product of the vectors $a$ and $b$. The elements of this matrix are given by $[a](b)_i = a_i b_j$. $\nabla f$ is the normal vector to the interface at the intersection point of the central ray.

To obtain relation (A1), Farra et al. (1989) just write that the paraxial ray satisfies ray theory developed to first order:

$$d y = \delta y + \frac{d y}{d \tau} d \tau,$$

where $d \tau = \tau' - \tau$. Moreover, as $d x$ is tangent to the interface to first order, one has

$$d \tau = - \frac{\langle \nabla f | \delta x \rangle}{\langle \nabla y | \nabla f \rangle}.$$  

Let us now construct the continuity conditions for the paraxial rays across the interface. Let $d \tilde{x} = \tilde{x}(\tau') - \tilde{x}(\tau)$ and $d \tilde{p} = \tilde{p}(\tau') - \tilde{p}(\tau)$, the perturbations of position and slowness vector of the reflected/transmitted ray along the interface. The canonical vector $d y$ is transformed by a linear transformation $T$ into the paraxial vector $d y = (d x, d p)$ of the converted paraxial ray along the interface (see Farra 1989):

$$d y = T d y_{1}.$$  

The $6 \times 6$ matrix $T$ is given by

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

where the $3 \times 3$ submatrices $T_1$ and $T_2$ are defined as

$$T_1 = \frac{(\nabla f)(\nabla \cdot n)(\nabla f)}{(\nabla \cdot n)(\nabla f)} - \frac{(\langle \nabla f | \delta x \rangle)}{(\langle \nabla y | \nabla f \rangle)} \frac{(\nabla \nabla f)}{(\nabla \cdot n)(\nabla f)};$$

$$T_2 = I - \frac{(\nabla f)(\nabla \cdot n)(\nabla f)}{(\nabla \cdot n)(\nabla f)}.$$  

The notation $n = \nabla f$ is the local normal to the interface at point $O$.

[Figure A1. 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 $\mathbf{p}_0 = p_0(\tau)$, while the paraxial ray arrives at $I$ with slowness vector $\mathbf{p}_1 = p_1(\tau)$. Vector $\delta y(\tau)$ defines the position of point $Q$ where the paraxial ray arrives with slowness vector $\mathbf{p}_2 = p_2(\tau)$. At the interface, paraxial vectors $(\delta \mathbf{x}, \delta \mathbf{p})$ are transformed into $(d \mathbf{x}, d \mathbf{p})$ by matrix $T$. $n = \nabla f$ is the local normal to the interface at point $O$.]

One can remark that matrix $T$ has a null determinant, so that the transformation matrix $T$ does not have the properties (9, 11, 14). The continuation procedure at interfaces gives the same extrapolated value $d y$ at the interface for different canonical perturbations describing the same paraxial ray. This does not limit the usefulness of the corresponding generalized propagator, except if one needs to use the properties (9–14).

Let us remark that, if we consider the sampling parameter $\tau$ as continuous at the interface, we can write the
perturbations \( \delta \tilde{y} = \tilde{y}(\tau') - \tilde{y}_0(\tau) \) and \( \delta \tilde{y} = \tilde{y}(\tau) - \tilde{y}_0(\tau) \) so that to first order
\[
\delta \tilde{y} = \delta \tilde{y}_0 \quad \frac{d \delta \tilde{y}_0}{d \tau} \, d \tau = \delta \tilde{y}_0 \left[ \begin{array}{c} \nabla_{\tilde{y}} \hat{H} \\ \nabla_{\tilde{y}} \hat{f} \end{array} \right] \, d \tau.
\]

Then
\[
\delta \tilde{y} = M \delta y,
\]

where
\[
M = TP + \left[ \begin{array}{c}
\nabla_{\tilde{y}} \hat{H} (\nabla f) \\
\nabla_{\tilde{y}} \hat{f} (\nabla f)
\end{array} \right] = \left[ \begin{array}{cc}
M_{11} & M_{12} \\
M_{21} & M_{22}
\end{array} \right],
\]

where the submatrices are given by
\[
M_{11} = I - \frac{\nabla_{\tilde{y}} \hat{H} (\nabla f)}{(\nabla_{\tilde{y}} \hat{H} (\nabla f)}
\]
\[
M_{12} = 0
\]
\[
M_{21} = \frac{\nabla_{\tilde{y}} \hat{H} (\nabla f)}{(\nabla_{\tilde{y}} \hat{H} (\nabla f)} - \frac{\nabla_{\tilde{y}} \hat{f} (\nabla f)}{(\nabla_{\tilde{y}} \hat{f} (\nabla f)}
\]
\[
M_{22} = I - \frac{\nabla_{\tilde{y}} \hat{H} (\nabla f)}{(\nabla_{\tilde{y}} \hat{H} (\nabla f)}
\]

Matrix \( M \) has the properties (9, 11, 14) so that the generalized propagator (17) has the properties (9–14).

The transformation matrix obtained by Gajewski & Psencik (1990) is quite similar to our matrix \( M \) computed for a Hamiltonian corresponding to the sampling parameter \( \tau = T \). However, some terms are missing in their submatrices \( M_{21} \) and \( M_{22} \). In fact, the application of their transformation matrix to a paraxial ray perturbation \( \delta y = (\delta x, \delta p) \) gives the same result, because the missing terms correspond to the perturbation of the eikonal equation (8), which is zero. However, matrix \( M \) obtained by Gajewski & Psencik (1990) does not verify the properties (9–14), which may be a problem in some applications of perturbation ray theory.

APPENDIX B: TRANSFORMATION OF PERTURBED RAY CANONICAL VECTOR AT AN INTERFACE

Let us consider a perturbation of the interface. We denote by \( f_0(x) \) the reference interface and by \( f(x) = f_0(x) + \Delta f(x) \) the perturbed interface. Consider a reference ray with canonical vector \( y_0(\tau) \), 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 \( \Delta y(\tau) \). The reference ray intersects the reference interface at \( x_0(\tau) \) and the perturbed ray intersects the perturbed interface at \( x(\tau') \). In order to propagate the transmitted or reflected perturbed ray away from the interface, we choose, as the new reference ray, the reflected/transmitted ray in the unperturbed medium corresponding to the reference incident ray. The canonical vectors of the two reference rays are connected by relations (15) at the interface in the reference medium. We denote by \( \tilde{y}_0(\tau) \) and \( \tilde{y}(\tau) \) the canonical vectors of the reflected/transmitted reference ray and the reflected/transmitted perturbed ray. We will assume that the sampling parameter \( \tau \) is continuous at the interface.

Following the approach developed by Farra et al. (1989) in isotropic media, we obtain the perturbation \( \delta \tilde{y} = \tilde{y}(\tau') - \tilde{y}_0(\tau) \).

\[
d\tilde{y} = T \partial y + \Delta y_1 + \Delta \gamma_2
\]

where \( \partial y_1 \) and \( \partial y_2 \) are matrices (A2) and (A7) computed on the reference ray and

\[
\Delta y_1 = -\frac{\Delta f}{(\nabla_{\tilde{y}} \hat{H}_0 (\nabla f_0)} \left[ \begin{array}{c}
\nabla_{\tilde{y}} \hat{H}_0 \\
\nabla_{\tilde{y}} \hat{f}_0
\end{array} \right],
\]
\[
\Delta y_2 = \left[ \begin{array}{c}
0 \\
\Delta p_2
\end{array} \right],
\]
\[
\Delta \gamma_2 = \Delta H - \frac{\Delta H}{(\nabla_{\tilde{y}} \hat{H}_0 (\nabla f_0)} \nabla f_0
\]
\[
- \frac{\Delta H}{(\nabla_{\tilde{y}} \hat{H}_0 (\nabla f_0)} \left[ \begin{array}{c}
\nabla_{\tilde{y}} \hat{H}_0 \\
\nabla_{\tilde{y}} \hat{f}_0
\end{array} \right],
\]
\[
\Delta \gamma_1 = \left[ \begin{array}{c}
\Delta f \\
\Delta p_1
\end{array} \right].
\]

All the quantities appearing in (B2) are calculated on the reference ray at the reference interface point with the reference interface.

The new canonical vector \( \Delta \tilde{y} = \tilde{y}(\tau') - \tilde{y}_0(\tau) \) is related to first order to \( \delta \tilde{y} = \tilde{y}(\tau') - \tilde{y}_0(\tau) \) by the relation

\[
d\tilde{y} = \delta \tilde{y}_0 \quad \frac{d \delta \tilde{y}_0}{d \tau} \, d \tau = \delta \tilde{y}_0 \left[ \begin{array}{c} \nabla_{\tilde{y}} \hat{H}_0 \\
\nabla_{\tilde{y}} \hat{f}_0
\end{array} \right] \, d \tau,
\]

where \( d \tau = \tau' - \tau \), is given by (Farra et al. 1989)

\[
d\tau = \frac{\Delta \gamma_1}{(\nabla_{\tilde{y}} \hat{H}_0 (\nabla f_0)}.
\]

The general transformation is then

\[
\Delta \tilde{y} = M_0 (\Delta \gamma_1 + \Delta y_1).
\]

where \( M_0 \) is the paraxial transformation matrix given by expression (A10) computed in the reference medium and

\[
\Delta y_1 = T_\partial y_1 + \Delta y_2 + \Delta y_1,
\]

\[
\Delta y_1 = \frac{\Delta f}{(\nabla_{\tilde{y}} \hat{H}_0 (\nabla f_0)} \left[ \begin{array}{c}
\nabla_{\tilde{y}} \hat{H}_0 \\
\nabla_{\tilde{y}} \hat{f}_0
\end{array} \right],
\]

This transformation contains four terms. The first one is the same as the linear transformation (16) connecting the incident and reflected/transmitted paraxial rays at the interface. This term takes into account perturbations in initial conditions and in medium parameters between the source and the interface. \( \Delta y_2 = \) and \( \Delta y_3 \) are due to the displacement \( \Delta H \) of the interface. \( \Delta y_1 \) as explained in Farra et al. (1989) is due to the perturbation of Snell's law. Expression (B5) is valid for isotropic as well as anisotropic media.
APPENDIX C: TRAVELTIME PERTURBATIONS DUE TO INTERFACE PERTURBATIONS

Let us consider a medium with two layers separated by an interface. We assume that a ray has been traced between two points with coordinates \( x_1 \) and \( x_2 \). The intersection point of this reference ray has coordinates \( x_i \). Let us now consider a perturbation of the interface. We denote by \( f_0(x) \) the reference interface and by \( f(x) = f_0(x) + \Delta f(x) \) the perturbed interface. Let us consider the perturbed ray between points \( x_1 \) and \( x_2 \). The intersection point of this perturbed ray has coordinates \( x_i \). To first order, the perturbation of the traveltime due to the interface perturbation is given by

\[
\Delta T = (p_1(\tau) - \bar{p}_1(\tau) | x'_i - x_i) \tag{C1}
\]

Moreover, one can write the equations of the interfaces at \( x_i \) and \( x'_i \), respectively:

\[
f_0(x_i) = 0, \quad f(x'_i) = 0. \tag{C2}
\]

Developing equations (C2) to first order, one obtains

\[
\langle \nabla f_0 | x'_i - x_i \rangle + \Delta f(x_i) = 0. \tag{C3}
\]

From the Snell–Descartes law (15), one can write the following relation:

\[
\begin{align*}
\mathbf{p}_0(\tau) - \mathbf{p}_0(\tau) &= \frac{\langle \mathbf{p}_0(\tau) - \mathbf{p}_0(\tau) | \nabla f_0 \rangle}{\langle \nabla f_0 | \nabla f_0 \rangle} \nabla f_0, \\
\end{align*} \tag{C4}
\]

From eqs (C1), (C3) and (C4), the traveltime perturbation due to interface perturbation may be written

\[
\Delta T = \frac{\langle \mathbf{p}_0(\tau) - \mathbf{p}_0(\tau) | \nabla f_0 \rangle}{\langle \nabla f_0 | \nabla f_0 \rangle} \Delta f. 
\]

This expression, which is valid for the perturbation of an interface separating two anisotropic media is the same as that obtained by Farra et al. (1989) for isotropic media.