Paraxial ray methods for anisotropic inhomogeneous media

T.J. Moser[1] and V. Červený[2]

accepted for publication in Geophysical Prospecting

[1] Zeehelden Geoservices, van Alkemadelaan 550A, 2597 AV 's-Gravenhage, The Netherlands.
E-mail: mosertj@hotmail.com

[2] Charles University, Faculty of Mathematics and Physics, Department of Geophysics, Ke Karlovu 3, 121 16 Praha 2, Czech Republic.
E-mail: vcerveny@seis.karlov.mff.cuni.cz

Abstract

A new formalism of surface-to-surface paraxial matrices allows a very general and flexible formulation of the paraxial ray theory, equally valid in anisotropic and isotropic inhomogeneous layered media. The formalism is based on conventional dynamic ray tracing in Cartesian coordinates along a reference ray. At any user-selected pair of points of the reference ray, a pair of surfaces may be defined. These surfaces may be arbitrarily curved and oriented, and may represent structural interfaces, data recording surfaces, or merely formal surfaces. A newly obtained factorization of the interface propagator matrix allows to transform the conventional 6×6 propagator matrix in Cartesian coordinates into a 6×6 surface-to-surface paraxial matrix. This matrix defines the transformation of paraxial ray quantities from one surface to another. The redundant non-eikonal and ray-tangent solutions of the dynamic ray-tracing system in Cartesian coordinates can be easily eliminated from the 6×6 surface-to-surface paraxial matrix, and it can be reduced to 4×4 form. Both the 6×6 and 4×4 surface-to-surface paraxial matrices satisfy useful properties, particularly the symplecticity. In their 4×4 reduced form, they can be used to solve important boundary-value problems of a four-parametric system of paraxial rays, connecting the two surfaces, similarly as the well-known surface-to-surface matrices in isotropic media in ray-centred coordinates. Applications of such boundary-value problems include the two-point eikonal, relative geometrical spreading, Fresnel zones, the design of migration operators, and more.

Seismic Waves in Complex 3-D Structures, Report 16 (Department of Geophysics, Charles University, Prague 2006), pp. 115-144

Keywords

Dynamic ray tracing, ray propagator matrices, two-point problems, seismic anisotropy, inhomogeneity, structural interfaces

1. Introduction

The main principles of kinematic and dynamic ray tracing in inhomogeneous anisotropic layered media are well known. Conventional ray tracing is used to compute the ray as a space trajectory, together with the travel time, the slowness vector and ray velocity vector at any point of the ray. Dynamic ray tracing (also called paraxial ray tracing) consists in the solution of an additional system of linear ordinary differential equations along the ray, and can be used to determine the geometrical spreading, amplitude and wavefront curvature along the ray. Moreover, it provides kinematic quantities not only along the ray, but also in some vicinity of the ray. We call this vicinity of the ray its paraxial vicinity, and speak of paraxial rays, paraxial travel times, paraxial slowness vectors, and so on. The dynamic ray tracing is very suitable for the solution of various boundary value problems for orthonomic (two-parametric) systems of paraxial rays. Also, dynamic ray tracing is essential in various extensions of the ray method, such as method of Gaussian beams and Gaussian packets, the Maslov-Chapman and Kirchhoff-Helmholtz methods, and more. For more details on kinematic and dynamic ray tracing in inhomogeneous anisotropic media we refer to Červený (1972, 2001), Gajewski and Pšenčík (1987, 1990), Norris (1987), Shearer and Chapman (1989), Farra (1989), Kendall, Guest and Thomson (1992), Klimeš (1994), Farra and Le Bégat (1995), Bakker (1996), Gjøystdal et al. (2002), Iversen (2004), Moser (2004), and Chapman (2004), where many other references can be found.

The dynamic ray-tracing system along a selected reference ray can be expressed in many forms and in various coordinate systems. The most common coordinate systems are the ray-centred coordinate system, where the reference ray serves as a coordinate axis, and the global Cartesian coordinate system. Since the dynamic ray-tracing system consists of linear ordinary differential equations, it is possible to introduce the relevant propagator matrices. Here, we refer to the propagator matrices corresponding to the dynamic ray-tracing system as the ray propagator matrices, or briefly ray propagators. For ray-centred coordinates these matrices are 4×4, for global Cartesian coordinates they are 6×6, corresponding to systems of four and six differential equations, respectively. In a smooth medium without interfaces, the ray propagators in both coordinate systems satisfy the important symplectic property. Symplecticity imposes an algebraic structure on the ray system, as a result of which several quantities are immediately available without the need of additional numerical computations (most notably the inverse of the ray propagator). At intersections with structural interfaces, interface propagator matrices are needed to ensure continuity and to preserve symplecticity (Červený, 2001; Moser, 2004).

Both global Cartesian and ray-centred coordinates have certain advantages and disadvantages. The main advantage of dynamic ray tracing in Cartesian coordinates is that the six dynamic ray-tracing equations and their transformations at interfaces have a general and simple structure. However, from the six independent solutions two are redundant: the so-called ray-tangent and non-eikonal solution. Dynamic ray tracing in ray-centred coordinates, on the other hand, consists of only four equations and does not involve the redundant solutions, but for anisotropic media the equations are more complicated. Moreover, it requires to solve an additional differential equation along the ray to determine the basis vectors of the ray-centred coordinate system (see Klimeš, 1994; Bakker, 1996; Červený, 2001).

The ray propagator matrices in Cartesian or ray-centred coordinates are not quite suitable to solve the boundary-value problems of a four-parametric system of rays. In Cartesian coordinates, the redundant solutions have to be removed by some transformation or reduction. In ray-centred coordinates, the coordinate plane is rigidly connected to the reference ray, usually tangent to the wavefront or perpendicular to the ray trajectory. As this plane generally does not coincide with surfaces of interest intersecting the ray, such as structural interfaces or datum surfaces, additional transformations are needed.

In this paper, we present a formalism for paraxial ray methods in anisotropic inhomogeneous layered media that combines the advantages of both mentioned coordinate systems: simple and general expressions, valid even in anisotropic media, and removal of the redundant solutions. We start with a new expression for the 6×6 interface propagator matrix in Cartesian coordinates (Moser, 2004) and its factorization (see eqs. (28)-(30) and (32) of this paper). Based on the specific forms of the components in the factorization, we introduce surface-to-surface paraxial matrices for anisotropic inhomogeneous layered media. They are connected with surfaces, crossing the reference ray, and connect the paraxial ray field on an initial surface, called the anterior surface, to the paraxial ray field on a final surface, called the posterior surface. We refer to these matrices as the surface-to-surface paraxial matrices, or briefly surface-to-surface matrices. All surfaces may be arbitrarily curved, shaped and oriented in space, and represent structural interfaces, data recording surfaces or merely formal surfaces. For isotropic layered media, surface-to-surface matrices were derived from the dynamic ray-tracing system in ray-centred coordinates by Hubral et al. (1992, 1993, 1995) and Červený (2001, sec. 4.4.7), and called the surface-surface propagator matrices. Here, we prefer the term surface-to-surface paraxial matrices, as these matrices are not the propagators of any system of differential equations. For anisotropic media, however, the derivation of surface-to-surface paraxial matrices in Cartesian or ray-centred coordinates has not yet been reported in ray-theoretical literature. The only known approach to derive them for anisotropic media is based on dynamic ray tracing in wavefront orthonormal coordinates, see Červený (2001, sec. 4.14.10). The wavefront orthonormal coordinate system is a local Cartesian system, with its origin at the reference ray, two axes tangent to the wavefront and the third parallel to the slowness vector (not to the ray). The reference ray is not a coordinate axis of the system, as in ray-centred coordinates. Surface-to-surface paraxial matrices in both Cartesian coordinates (as presented in this paper) and wavefront orthonormal coordinates can be used alternatively. Without any doubt, dynamic ray tracing is more commonly used in Cartesian than in wavefront orthonormal coordinates. For this reason, we expect that the surface-to-surface paraxial matrices in Cartesian coordinates will find broader application in the seismic ray method.

The primary advantage of the surface-to-surface formalism is that it allows the ray propagator to be transformed to surface-to-surface matrices between every pair of arbitrary surfaces crossing the reference ray that suits our needs. Algorithmically, the kinematic and dynamic ray equations are solved in the conventional global Cartesian coordinates; the transformation to surface-to-surface matrices is made only where needed.

The relevant surface-to-surface paraxial matrices can be introduced both in 6×6 and 4×4 versions. The 4×4 versions are physically particularly appealing, because they can be directly applied to the solution of important boundary-value ray-tracing problems of four-parametric system of paraxial rays. Among other things, they can be applied to deriving the expression for the paraxial travel time from a point on one surface to a point an another surface. Such a function is called the Hamiltonian point characteristic, and plays a fundamental role in the theory of seismic systems (see Bortfeld, 1989). Consequently, the surface-to-surface paraxial matrices provide a direct link between dynamic ray tracing and the theory of seismic systems. For details on the theory of seismic systems in media composed of homogeneous and inhomogeneous isotropic layers see Bortfeld (1989), Hubral et al. (1992, 1993, 1995), Schleicher et al. (1993), and Červený (2001, sec. 4.4.7). This paper offers a generalization of the theory of seismic systems for anisotropic media. Certain results derived here were already known, from the pioneering work by Schleicher et al. (2001). An analogous theory of optical systems is well known from optics, and is described in most textbooks on optics. See, for example, Herzberger (1958), Luneburg (1964), Buchdahl (1970), Guenther (1990). The paraxial theory of optical systems has found broad applications in the design of optical systems. In the theory of seismic systems, however, the seismic system (structure) is fixed, and we study the properties of the system (structure) from travel-time measurements along the anterior and posterior surfaces.

Briefly to the content of the paper. In Section 2, we review the theory of dynamic ray tracing in Cartesian coordinates for inhomogeneous anisotropic layered media. The basic result is the expression for the interface propagator matrix (28)-(30), and for its factorization (32). It is shown that both factors are symplectic. In Section 3, 6×6 surface-to-surface paraxial matrices are constructed from conventional 6×6 ray propagator matrices in Cartesian coordinates. It is shown that surface-to-surface matrices satisfy certain properties of ray propagator matrices (in particular symplecticity). It is also shown how 6×6 surface-to-surface paraxial matrices can be reduced to 4×4 surface-to-surface matrices, by removing the entries (columns and rows) related to the ray-tangent and non-eikonal solutions. In Section 4, some possible applications of the derived results are presented. In Sections 4.1 and 4.2, two important boundary-value problems of the four-parametric system of rays between two surfaces are solved, and the

Hamiltonian point characteristic for an anisotropic inhomogeneous layered medium is derived. Section 4.3 explains how the surface-to-surface paraxial matrix can be fully determined from the travel-time measurements along the anterior and posterior surfaces. The remaining sections are devoted to the definition and determination of relative geometrical spreading of the ray-theory Green's function, to the factorization of relative geometrical spreading, and to the computation of Fresnel zones. The presented results, particularly the Hamiltonian point characteristic, also offer useful applications in seismic imaging and inversion in inhomogeneous anisotropic layered structures. To keep the article to an admissible length, we will not discuss such applications here.

The notation in this paper follows Červený (2001) and Moser (2004), apart from a few exceptions. Most importantly, since we deal extensively with matrices of different sizes, we refrain from using the hat "^" to denote 3×3 matrices (as in Červený (2001)). Instead, all 3×1 vectors are denoted in roman boldface lowercase $\mathbf{a}$, 3×3 and 6×6 matrices are denoted in roman boldface uppercase $\mathbf{A}$. All 2×1 vectors are denoted in italic boldface lowercase $\boldsymbol{a}$, 2×2 and 4×4 matrices are denoted in italic boldface uppercase $\boldsymbol{A}$. Whenever there may be reason for confusion, the dimensions of the matrix will be indicated. The notation $\mathbf{a}^T$, $\boldsymbol{a}^T$, $\mathbf{A}^T$, $\boldsymbol{A}^T$ denotes the transpose of a vector or matrix, $\mathbf{A}^{-T}$, $\boldsymbol{A}^{-T}$ denotes the inverse transpose of a matrix. Greek indices $\mu, \nu, \cdots$ run over 1 and 2, Latin indices $i, j, k, l, \cdots$ over 1, 2, and 3. The Kronecker symbol δ_{ij} equals 1 for $i = j$ and 0 for $i \neq j$. The summation convention is assumed for repeated indices in the same product. Derivatives or partial derivatives are denoted by subscripts as indicated when they are introduced. The expression "across the interface" denotes the situation after reflection or transmission ("R/T"), even although the reflected ray lies on the same side as the incident ray; the tilde ˜ denotes ray quantities across the interface. In the whole paper, the medium is tacitly assumed to be a general, three-dimensional, anisotropic, inhomogeneous, perfectly elastic, layered medium with smooth interfaces; if this is not the case, it is indicated.

2. 6×6 propagator $\boldsymbol{\Pi}$ in Cartesian coordinates

2.1 Dynamic ray tracing in Cartesian coordinates

Let us start by reviewing some well-known material relevant to the following sections. Consider a one-parametric ray field

$$\mathbf{x}(\gamma, \tau), \mathbf{p}(\gamma, \tau) \tag{1}$$

in a 3D smooth medium, where $\mathbf{x}$ is the ray location vector, $\mathbf{p}$ the slowness vector, both in Cartesian coordinates, and τ a monotonically increasing sampling parameter along the ray. Each ray is specified by the ray parameter γ and satisfies the kinematic ray equations

$$\mathbf{x}_\tau = H_{\mathbf{p}}, \quad \mathbf{p}_\tau = -H_{\mathbf{x}} \ , \tag{2}$$

where $H(\mathbf{x}, \mathbf{p})$ is the Hamiltonian, or in matrix form

$$\begin{bmatrix} \mathbf{x}_\tau \\ \mathbf{p}_\tau \end{bmatrix} = \mathbf{J} \begin{bmatrix} H_\mathbf{x} \\ H_\mathbf{p} \end{bmatrix} \ , \ \mathbf{J} = \begin{bmatrix} \mathbf{O} & \mathbf{I} \\ -\mathbf{I} & \mathbf{O} \end{bmatrix} \ , \tag{3}$$

where $\mathbf{J}$ is a 6×6 matrix, and $\mathbf{I}$ and $\mathbf{O}$ are the 3×3 identity and null matrix. Here, and in the following, subscripts $\mathbf{x}$ and $\mathbf{p}$ denote first-order partial derivatives with respect to these quantities; double subscripts $\mathbf{xx}$, $\mathbf{xp}$, and $\mathbf{pp}$ denote second-order partial derivatives. The subscript τ denotes the first-order ordinary derivative with respect to τ, $\mathbf{x}_\tau = \mathrm{d}\mathbf{x}/\mathrm{d}\tau$, $\mathbf{p}_\tau = \mathrm{d}\mathbf{p}/\mathrm{d}\tau$.

The Hamiltonian comprises all medium characteristics, such as inhomogeneity and anisotropy. For an isotropic inhomogeneous smooth medium with a velocity distribution $c(\mathbf{x})$, it may take the form

$$H(\mathbf{x}, \mathbf{p}) = \tfrac{1}{2} \left(c(\mathbf{x})^2 \mathbf{p}^T \mathbf{p} - 1 \right) \ , \tag{4}$$

or one of many equivalent forms. In an anisotropic medium with density-normalized elastic moduli $a_{ijkl}(\mathbf{x})$ $(i, j, k, l = 1, 2, 3)$, it may take the form

$$H(\mathbf{x}, \mathbf{p}) = \tfrac{1}{2} \left(G_m(\mathbf{x}, \mathbf{p}) - 1 \right), \tag{5}$$

where $G_m(\mathbf{x}, \mathbf{p})$ $(m = 1, 2, 3)$ is one of the three eigenvalues of the 3×3 matrix $a_{ijkl}(\mathbf{x}) p_j p_l$.

The kinematic ray equations (2) form a system of six coupled non-linear ordinary differential equations, usually to be solved by numerical integration. Initial conditions for (2) have to satisfy the eikonal equation

$$H(\mathbf{x}, \mathbf{p}) = 0 \ , \tag{6}$$

which is then satisfied along the whole ray. The general independent parameter τ is related to the travel time T along the ray through the specific form of the Hamiltonian by

$$\mathrm{d}T/\mathrm{d}\tau = \mathbf{p}^T H_\mathbf{p} \ . \tag{7}$$

A wavefront is defined as a surface of equal travel time , and the slowness vector $\mathbf{p} = \nabla T$ is perpendicular to it. In the isotropic case, the phase velocity vector, $\mathbf{p}/(\mathbf{p}^T\mathbf{p})$, and the ray velocity vector, $\mathbf{x}_T$, are identical, but in the general anisotropic case they have different magnitudes and directions, so that the rays are generally not normal to the wavefronts.

The *dynamic ray equations*, also called *paraxial ray equations*, for the ray field (1) are obtained by taking the partial derivative of (2) with respect to the ray parameter γ:

$$\begin{bmatrix} \mathbf{x}_{\gamma\tau} \\ \mathbf{p}_{\gamma\tau} \end{bmatrix} = \mathbf{J}\,\mathbf{H} \begin{bmatrix} \mathbf{x}_\gamma \\ \mathbf{p}_\gamma \end{bmatrix} \ , \ \mathbf{H} = \begin{bmatrix} H_\mathbf{xx} & H_\mathbf{xp} \\ H_\mathbf{px} & H_\mathbf{pp} \end{bmatrix} \ , \tag{8}$$

where $H_\mathbf{xx}$, $H_\mathbf{xp}$, $H_\mathbf{px}$, and $H_\mathbf{pp}$ are 3×3 second-order partial derivative matrices of $H(\mathbf{x}, \mathbf{p})$, $\mathbf{x}_{\gamma\tau} = \mathrm{d}\mathbf{x}_\gamma/\mathrm{d}\tau$ and $\mathbf{p}_{\gamma\tau} = \mathrm{d}\mathbf{p}_\gamma/\mathrm{d}\tau$.

The dynamic ray equations (8) form a system of six linear ordinary differential equations for $\mathbf{x}_\gamma$ and $\mathbf{p}_\gamma$. These equations can be solved simultaneously with the kinematic ray tracing equations (2), or along a known ray. The computed paraxial ray quantities can be used to approximate rays in the neighbourhood of a reference ray (paraxial rays), without the need to trace each of them individually. Initial conditions to (8) have to satisfy the eikonal constraint relation , which follows from the requirement that paraxial quantities satisfy the eikonal equation (6) to first order:

$$0 = H_\gamma(\mathbf{x}, \mathbf{p}) = H_{\mathbf{x}}^T \mathbf{x}_\gamma + H_{\mathbf{p}}^T \mathbf{p}_\gamma \quad . \tag{9}$$

As a result of this constraint, one of the six general independent solutions to (8) is non-physical. This solution is referred to as the non-eikonal solution. Another solution to (8) is trivial, namely the ray-tangent solution, where $(\mathbf{x}_\gamma, \mathbf{p}_\gamma)$ and $(\mathbf{x}_\tau, \mathbf{p}_\tau)$ are parallel (linearly dependent). The remaining four solutions are physical and non-trivial, and are called standard paraxial solutions. Dynamic ray tracing in two independent ray-centred coordinates is designed to exclude the non-eikonal and ray-tangent solutions, in order to reduce the number of equations to be solved (for isotropic media see Popov and Pšenčík (1978a,b); for anisotropic media, see Klimeš (1994), Červený (2001)). On the other hand, dynamic ray tracing in three independent Cartesian coordinates includes all six solutions, redundant or not, but arrives at simpler and more concise expressions, which are generally valid in inhomogeneous anisotropic media.

Let us now consider a two-parametric (or orthonomic) ray field

$$\mathbf{x}(\gamma_1, \gamma_2, \tau), \mathbf{p}(\gamma_1, \gamma_2, \tau) \quad , \tag{10}$$

where γ_1 and γ_2 are parameters specifying the ray, for example take-off angles from a point source, or the coordinates on an initial surface. Both γ_1 and γ_2 can play the role of γ in the above equations (1)-(9). It is useful to combine the ray parameters into one vector $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \tau)$ and the paraxial quantities into 3×3 matrices

$$\mathbf{Q} = \partial \mathbf{x} / \partial \boldsymbol{\gamma} = (\mathbf{x}_{\gamma_1}, \mathbf{x}_{\gamma_2}, \mathbf{x}_\tau), \quad \mathbf{P} = \partial \mathbf{p} / \partial \boldsymbol{\gamma} = (\mathbf{p}_{\gamma_1}, \mathbf{p}_{\gamma_2}, \mathbf{p}_\tau) \quad , \tag{11}$$

where each column of $\mathbf{Q}$ and $\mathbf{P}$ satisfies the paraxial ray equations (8); the third column $\mathbf{x}_\tau, \mathbf{p}_\tau$ represents the ray-tangent solution discussed above. The ray parameters γ_1, γ_2 combined with τ are also called *ray coordinates*.

Dynamic ray tracing is a very useful tool to study orthonomic ray bundles, and to investigate their geometrical properties such as divergence, convergence, caustics, wavefront curvatures, as well as the paraxial travel-time field. In particular, $|\det \mathbf{Q} \, \mathrm{d}\tau / \mathrm{d}s|^{1/2}$ is the geometrical spreading (for arclength s), which is inversely proportional to the ray-theoretical amplitude along the reference ray. At caustics, $\det \mathbf{Q} = 0$. For $\det \mathbf{Q} \neq 0$, the 3×3 matrix of the second-order derivatives of the travel-time field is given by

$$\mathbf{M} = T_{\mathbf{xx}} = \mathbf{p}_{\mathbf{x}} = (\partial \mathbf{p} / \partial \boldsymbol{\gamma}) \, (\partial \boldsymbol{\gamma} / \partial \mathbf{x}) = \mathbf{P} \mathbf{Q}^{-1} \quad . \tag{12}$$

The matrix $\mathbf{M}$ is symmetric ($\mathbf{M} = \mathbf{M}^T$) and permits a second-order expansion of the travel time around a reference ray point $\mathbf{x}$:

$$T(\mathbf{x}+\delta\mathbf{x}) = T(\mathbf{x}) + \mathbf{p}^T\delta\mathbf{x} + \tfrac{1}{2}\,\delta\mathbf{x}^T\mathbf{M}\delta\mathbf{x} + O(\|\delta\mathbf{x}\|^3) \ . \tag{13}$$

From the matrix $\mathbf{M}$ the principal wavefront curvatures can be easily derived.

2.2 6×6 *ray propagator* $\mathbf{\Pi}$

The 6×6 ray propagator $\mathbf{\Pi}(\tau,\tau_0)$ is composed of the fundamental solutions to the paraxial ray equations $(\mathrm{d}/\mathrm{d}\tau)\mathbf{\Pi}(\tau,\tau_0) = \mathbf{JH\Pi}(\tau,\tau_0)$ subject to initial conditions $\mathbf{\Pi}(\tau_0,\tau_0) = \mathbf{I}$ (where $\mathbf{I}$ is the 6×6 identity). It is a linear mapping relating any particular solution $(\mathbf{x}_\gamma(\tau),\mathbf{p}_\gamma(\tau))$ at τ to $(\mathbf{x}_\gamma(\tau_0),\mathbf{p}_\gamma(\tau_0))$ at τ_0 through

$$\begin{bmatrix}\mathbf{x}_\gamma(\tau)\\ \mathbf{p}_\gamma(\tau)\end{bmatrix} = \mathbf{\Pi}(\tau,\tau_0)\begin{bmatrix}\mathbf{x}_\gamma(\tau_0)\\ \mathbf{p}_\gamma(\tau_0)\end{bmatrix}\ , \ \text{or}\ \begin{bmatrix}\mathbf{Q}(\tau)\\ \mathbf{P}(\tau)\end{bmatrix} = \mathbf{\Pi}(\tau,\tau_0)\begin{bmatrix}\mathbf{Q}(\tau_0)\\ \mathbf{P}(\tau_0)\end{bmatrix}\ . \tag{14}$$

The propagator $\mathbf{\Pi}$ has the form

$$\mathbf{\Pi}(\tau,\tau_0) = \begin{bmatrix}\mathbf{Q}_1(\tau,\tau_0) & \mathbf{Q}_2(\tau,\tau_0)\\ \mathbf{P}_1(\tau,\tau_0) & \mathbf{P}_2(\tau,\tau_0)\end{bmatrix} = \begin{bmatrix}\partial\mathbf{x}(\tau)/\partial\mathbf{x}(\tau_0) & \partial\mathbf{x}(\tau)/\partial\mathbf{p}(\tau_0)\\ \partial\mathbf{p}(\tau)/\partial\mathbf{x}(\tau_0) & \partial\mathbf{p}(\tau)/\partial\mathbf{p}(\tau_0)\end{bmatrix}\ , \tag{15}$$

where the 3×3 submatrices $\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{P}_1$, and $\mathbf{P}_2$ are defined by the right-hand side of (15). The propagator has a number of useful properties. It satisfies the chain rule:

$$\mathbf{\Pi}(\tau_n,\tau_0) = \mathbf{\Pi}(\tau_n,\tau_{n-1})\cdots\mathbf{\Pi}(\tau_i,\tau_{i-1})\cdots\mathbf{\Pi}(\tau_2,\tau_1)\mathbf{\Pi}(\tau_1,\tau_0) \tag{16}$$

for any set of arbitrarily ordered points τ_i $(i = 0,\cdots,n)$ along the ray. For any two points τ_0 and τ, $\mathbf{\Pi}(\tau,\tau_0)^{-1} = \mathbf{\Pi}(\tau_0,\tau)$, and for every single point along the ray $\mathbf{\Pi}(\tau,\tau) = \mathbf{\Pi}(\tau_0,\tau_0) = \mathbf{I}$. A very important property is symplecticity:

$$\mathbf{\Pi}^T\mathbf{J\Pi} = \mathbf{J}\ . \tag{17}$$

Symplecticity implies that the inverse ray propagator $\mathbf{\Pi}^{-1}$ is directly available as

$$\mathbf{\Pi}^{-1} = \mathbf{J}^T\mathbf{\Pi}^T\mathbf{J} = \begin{bmatrix}\mathbf{P}_2^T & -\mathbf{Q}_2^T\\ -\mathbf{P}_1^T & \mathbf{Q}_1^T\end{bmatrix}\ , \tag{18}$$

so that its computation does not require numerical inversion of $\mathbf{\Pi}$ or tracing the ray back to its origin. The exact inverse (18) is an advantage for the case of an ill-conditioned $\mathbf{\Pi}$. The symplectic property automatically applies for Hamiltonian ray systems (2). Phase space incompressibility or Liouville's theorem implies $\det\mathbf{\Pi} = 1$. The symplectic ray propagator $\mathbf{\Pi}$ preserves the quantity H_γ, given by (9), along the ray, even if it is non-zero; therefore, it encompasses a complete set of solutions to the paraxial equations (8).

2.3 Interface propagator

Let us now consider a surface given by the implicit relation $\Sigma(\mathbf{x}) = 0$, and separating two parts of the medium, where $\Sigma > 0$ and $\Sigma < 0$, respectively. The scalar function $\Sigma(\mathbf{x})$ has first and second-order derivatives, given by the vector $\Sigma_\mathbf{x}$ and matrix $\Sigma_{\mathbf{xx}}$. The vector $\Sigma_\mathbf{x}$ is normal to the surface and we denote the unit normal by $\mathbf{n}^{\Sigma} = \Sigma_{\mathbf{x}} / (\Sigma_{\mathbf{x}}^T \Sigma_{\mathbf{x}})^{1/2}$. The matrix $\Sigma_{\mathbf{xx}}$ is related to the curvatures of the surface. The implicit definition of the surface by $\Sigma(\mathbf{x}) = 0$ is useful in numerical ray tracing, in order to compute its intersection with the ray. For the purpose of transforming paraxial quantities across Σ, we will also use a parametric description given by

$$\mathbf{x} = \mathbf{g}(u_1, u_2) \ , \tag{19}$$

where u_1, u_2 are Gaussian coordinates of the surface, and $\Sigma(\mathbf{g}(u_1, u_2))$ vanishes identically for all u_1, u_2. In the following, we shall call u_1, u_2 surface coordinates and denote $\boldsymbol{u} = (u_1, u_2)^T$. The vectorial function $\mathbf{g}(u_1, u_2)$ has first- and second-order partial derivatives given by $\mathbf{g}_\mu = \partial \mathbf{g} / \partial u_\mu$ and $\mathbf{g}_{\mu\nu} = \partial^2 \mathbf{g} / \partial u_\mu \partial u_\nu$, $(\mu, \nu = 1, 2)$; $\mathbf{g}_\mu$ are (generally non-orthogonal and non-unit) tangent vectors to the surface, and the vectors $\mathbf{g}_{\mu\nu}$ are related to the curvatures. We assume that the surface Σ is regular: $\det[\mathbf{g}_1, \mathbf{g}_2, \Sigma_{\mathbf{x}}] \neq 0$. The $\mathbf{g}$ derivatives and Σ derivatives are connected by

$$\Sigma_{\mathbf{x}}^T \mathbf{g}_\mu = 0 \ , \quad \mathbf{g}_\nu^T \Sigma_{\mathbf{xx}} \mathbf{g}_\mu + \Sigma_{\mathbf{x}}^T \mathbf{g}_{\mu\nu} = 0 \ \ (\mu, \nu = 1, 2) \ . \tag{20}$$

The relations (20) leave some freedom, and at any point $\mathbf{x}$ the surface $\Sigma(\mathbf{x}) = 0$ can be locally approximated to second order by

$$\mathbf{x} = \mathbf{g}_0 + \mathbf{g}_\mu u_\mu + \tfrac{1}{2}\, \mathbf{g}_{\mu\nu} u_\mu u_\nu, \tag{21}$$

in such a way that $\mathbf{g}_{\mu\nu}$ is perpendicular to the surface tangent vectors $\mathbf{g}_\mu$ $(\mu, \nu = 1, 2)$ and $\mathbf{g}_\mu^T \mathbf{g}_\nu = \delta_{\mu\nu}$ (where $\delta_{\mu\nu}$ is the Kronecker symbol and $\mathbf{g}_0 = \mathbf{g}(0, 0)$). To achieve this, choose $\mathbf{g}_1$ and $\mathbf{g}_2$ orthonormal to $\Sigma_{\mathbf{x}}$ (by reparametrization of u_1, u_2), take an arbitrary vector $\mathbf{g}$ out of the surface tangent plane, project it on $\Sigma_{\mathbf{x}}$

$$\mathbf{g} \rightarrow (\mathbf{g}^T \Sigma_{\mathbf{x}} / \Sigma_{\mathbf{x}}^T \Sigma_{\mathbf{x}})\ \Sigma_{\mathbf{x}} \ \ , \tag{22}$$

and stretch it to satisfy (20)

$$\mathbf{g}_{\mu\nu} = \ -\ (\mathbf{g}_\mu^T \Sigma_{\mathbf{xx}} \mathbf{g}_\nu / \Sigma_{\mathbf{x}}^T \mathbf{g})\ \mathbf{g} \ \ . \tag{23}$$

The surface coordinates u_μ then represent local Cartesian coordinates in the surface tangent plane, and the quantities $\mathbf{g}_{\mu\nu}^T \mathbf{n}^{\Sigma}$ are equal to the elements of the 2×2 curvature matrix $\boldsymbol{D}$, given by Červený (2001, eq. 4.4.14). We note in passing that the same procedure can be used to derive wavefront curvatures from the travel-time function $T(\mathbf{x})$ (13).

We will need these surface relations in the following sections. For now, let us suppose that the surface $\Sigma(\mathbf{x}) = 0$ represents a structural interface, and that it separates two parts of the medium with different smooth distributions of the elastic moduli. We consider the reflection or transmission (R/T) of a ray at the interface, and denote ray quantities across the interface with a tilde ˜ - "across" meaning after the reflection or transmission, even although the reflected ray lies on the same side of the interface as the incident ray. Across the interface, the ray field is governed by the appropriate Hamiltonian and eikonal equation:

$$\tilde{H}(\tilde{\mathbf{x}}, \tilde{\mathbf{p}}) = 0 \ . \tag{24}$$

Let two points τ and $\tilde{\tau}$ be located immediately (infinitesimally) before and across the interface. The kinematic boundary conditions at the interface are ray continuity

$$\tilde{\mathbf{x}}(\gamma, \tilde{\tau}) = \mathbf{x}(\gamma, \tau) \tag{25}$$

and Snell's law

$$(\tilde{\mathbf{p}}(\gamma, \tilde{\tau}) - \mathbf{p}(\gamma, \tau))^T \mathbf{g}_\mu = 0 \ , \ \ (\mu = 1, 2) \ . \tag{26}$$

Together with (24), (26) has generally six, possibly complex-valued, solutions for $\tilde{\mathbf{p}}$ at any side of the interface Σ. Selecting the solutions corresponding to waves propagating away from the interface, we obtain three reflected $(\mathrm{P}, \mathrm{S}_1, \mathrm{S}_2)$ and three transmitted waves $(\mathrm{P}, \mathrm{S}_1, \mathrm{S}_2)$. From those, we consider here any wave with real-valued $\tilde{\mathbf{p}}$.

The transformation of paraxial quantities across the interface is given by the interface propagator

$$\begin{bmatrix} \tilde{\mathbf{Q}} \\ \tilde{\mathbf{P}} \end{bmatrix} = \mathbf{\Pi}(\tilde{\tau}, \tau) \begin{bmatrix} \mathbf{Q} \\ \mathbf{P} \end{bmatrix} \ , \tag{27}$$

and can be derived by systematic, but laborious, differentiation of (25) and (26) with respect to γ. A particularly concise form is given by Moser (2004):

$$\mathbf{\Pi}(\tilde{\tau}, \tau) = \begin{bmatrix} \tilde{\mathbf{X}}\mathbf{X}^{-1} & \mathbf{O} \\ \tilde{\mathbf{X}}^{-T}\mathbf{R}\mathbf{X}^{-1} & \tilde{\mathbf{X}}^{-T}\mathbf{X}^T \end{bmatrix} \ , \tag{28}$$

where

$$\mathbf{X} = (\mathbf{g}_1, \mathbf{g}_2, \mathbf{x}_\tau) \ , \ \ \tilde{\mathbf{X}} = (\mathbf{g}_1, \mathbf{g}_2, \tilde{\mathbf{x}}_\tau) \ , \tag{29}$$

and

$$\mathbf{R} = \begin{bmatrix} -\mathbf{g}_{\mu\nu}^T \Delta\mathbf{p} & \mathbf{g}_\mu^T \Delta\mathbf{p}_\tau \\ \mathbf{g}_\nu^T \Delta\mathbf{p}_\tau & \tilde{\mathbf{p}}_\tau{}^T \tilde{\mathbf{x}}_\tau - \mathbf{p}_\tau{}^T \mathbf{x}_\tau \end{bmatrix} \ . \tag{30}$$

Quantities with Δ denote the differences across the interface: $\Delta\mathbf{p} = \tilde{\mathbf{p}} - \mathbf{p}$, $\Delta\mathbf{p}_\tau = \tilde{\mathbf{p}}_\tau - \mathbf{p}_\tau$. All quantities in (28) can be evaluated, as soon as the kinematic interface transform (25)-(26) has been completed. It can be shown that the interface propagator in the form given by (28) shares the properties of ray propagators in smooth parts of the medium. Most importantly, it is easy to show that 1. $\mathbf{\Pi}(\tilde{\tau}, \tau)$ is symplectic, 2. $\det \mathbf{\Pi}(\tilde{\tau}, \tau) = 1$, 3. it preserves the eikonal constraint relation H_γ (whether it is zero or not), and 4. for a given interface, $\mathbf{x}$, $\mathbf{p}$, and $\tilde{\mathbf{p}}$, $\mathbf{\Pi}(\tilde{\tau}, \tau)$ is unique and does not depend on the interface parametrization. It provides, therefore, an invertible transformation (with the inverse given by (18)) of the complete set of solutions to the paraxial ray equations (8). As a result, piecewise smooth media with smooth interfaces can be formally treated in the same way as smooth media. A derivation of (28) is given in Moser (2004), an independent proof is given in the appendix. In the following section, we give a factorization and provide a physical interpretation of the interface propagator in the form

(28).

2.4 Factorization of interface propagator

Putting

$$\tilde{\mathbf{W}} = \begin{bmatrix} -\mathbf{g}_{\mu\nu}^T\tilde{\mathbf{p}} & \mathbf{g}_{\mu}^T\tilde{\mathbf{p}}_{\tau} \\ \mathbf{g}_{\nu}^T\tilde{\mathbf{p}}_{\tau} & \tilde{\mathbf{p}}_{\tau}{}^T\tilde{\mathbf{x}}_{\tau} \end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix} -\mathbf{g}_{\mu\nu}^T\mathbf{p} & \mathbf{g}_{\mu}^T\mathbf{p}_{\tau} \\ \mathbf{g}_{\nu}^T\mathbf{p}_{\tau} & \mathbf{p}_{\tau}{}^T\mathbf{x}_{\tau} \end{bmatrix}, \quad \text{and } \mathbf{R} = \tilde{\mathbf{W}} - \mathbf{W}\ , \tag{31}$$

$(\mu, \nu = 1, 2)$, we obtain the following factorization of the interface propagator $\mathbf{\Pi}(\tilde{\tau}, \tau)$

$$\mathbf{\Pi}(\tilde{\tau}, \tau) = \tilde{\mathbf{Y}}^{-1}(\tilde{\tau})\mathbf{Y}(\tau)\ , \tag{32}$$

where

$$\mathbf{Y}(\tau) = \begin{bmatrix} \mathbf{X}^{-1} & \mathbf{O} \\ -\mathbf{W}\mathbf{X}^{-1} & \mathbf{X}^T \end{bmatrix}, \quad \text{and} \quad \tilde{\mathbf{Y}}^{-1}(\tilde{\tau}) = \begin{bmatrix} \tilde{\mathbf{X}} & \mathbf{O} \\ \tilde{\mathbf{X}}^{-T}\tilde{\mathbf{W}} & \tilde{\mathbf{X}}^{-T} \end{bmatrix} \tag{33}$$

are 6×6 matrices. The factorization (32) is analogous to the factorization of the 4×4 propagator in ray-centred coordinates in isotropic media, presented by Červený (2001, eq. 4.4.77). The matrices $\mathbf{Y}$ and $\tilde{\mathbf{Y}}$ are symplectic (this follows simply from (17)). Contrary to $\mathbf{\Pi}(\tilde{\tau}, \tau)$, $\mathbf{Y}$ and $\tilde{\mathbf{Y}}$ do depend on the interface parametrization (19).

Let us further analyze the matrix $\mathbf{Y}$. We define the 3×3 surface paraxial matrices $\mathbf{Q}^{(u)}$ and $\mathbf{P}^{(u)}$ as follows:

$$\begin{bmatrix} \mathbf{Q}^{(u)} \\ \mathbf{P}^{(u)} \end{bmatrix} = \mathbf{Y} \begin{bmatrix} \mathbf{Q} \\ \mathbf{P} \end{bmatrix}, \tag{34}$$

or

$$\mathbf{Q}^{(u)} = \mathbf{X}^{-1}\mathbf{Q}\ , \quad \mathbf{P}^{(u)} = -\mathbf{W}\mathbf{X}^{-1}\mathbf{Q} + \mathbf{X}^T\mathbf{P}\ . \tag{35}$$

The factorization (32) implies that $\tilde{\mathbf{Q}}^{(u)} = \mathbf{Q}^{(u)}$ and $\tilde{\mathbf{P}}^{(u)} = \mathbf{P}^{(u)}$, and $\tilde{\mathbf{Q}}^{(u)}$ and $\tilde{\mathbf{P}}^{(u)}$ are transformed back to Cartesian coordinates by applying $\tilde{\mathbf{Y}}^{-1}$:

$$\tilde{\mathbf{Q}} = \tilde{\mathbf{X}}\mathbf{Q}^{(u)}, \quad \tilde{\mathbf{P}} = \tilde{\mathbf{X}}^{-T}\tilde{\mathbf{W}}\mathbf{Q}^{(u)} + \tilde{\mathbf{X}}^{-T}\mathbf{P}^{(u)}\ . \tag{36}$$

Since the 3×3 matrices $\mathbf{Q}$ and $\mathbf{P}$ still contain the non-eikonal and ray-tangent solutions to the dynamic ray-tracing equations (8), it is not straightforward to give a physical interpretation to the transformed matrices $\mathbf{Q}^{(u)}$ and $\mathbf{P}^{(u)}$.

To obtain a physical interpretation, we first expand the travel time of the incident ray field along the interface, around the reference ray point $\mathbf{x}$. We do this by inserting the interface parametrization (21) in the second-order travel-time expansion around $\mathbf{x}$ (13):

$$T(u_1, u_2) = T(0,0) + \mathbf{p}^T\mathbf{g}_{\mu}u_{\mu} + \tfrac{1}{2}(\mathbf{p}^T\mathbf{g}_{\mu\nu} + \mathbf{g}_{\mu}^T\mathbf{M}\mathbf{g}_{\nu})u_{\mu}u_{\nu} + O(\|\boldsymbol{u}\|^3)\ . \tag{37}$$

Next, we define 2×2 surface paraxial matrices $\boldsymbol{Q}^{(u)}$ and $\boldsymbol{P}^{(u)}$ as follows:

$$\boldsymbol{Q}^{(u)} = \frac{\partial(u_1, u_2)}{\partial(\gamma_1, \gamma_2)} \ , \ \ \boldsymbol{P}^{(u)} = \frac{\partial(p_1^{(u)}, p_2^{(u)})}{\partial(\gamma_1, \gamma_2)} \ . \tag{38}$$

Here, γ_1, γ_2 are again two parameters specifying an orthonomic ray field (10), and u_1, u_2 are surface coordinates specifying position of a point on the surface (19). The quantities $p_1^{(u)}$, $p_2^{(u)}$ are tangential components of the slowness vector along the interface, which are given by $p_\mu^{(u)} = \partial T(u_1, u_2) / \partial u_\mu = \mathbf{p}^T \mathbf{g}_\mu$ $(\mu = 1, 2)$.

It is now easy to show that the 2×2 matrices $\boldsymbol{Q}^{(u)}$ and $\boldsymbol{P}^{(u)}$ are the upper left submatrices of the 3×3 matrices $\mathbf{Q}^{(u)}$ and $\mathbf{P}^{(u)}$. For $\mathbf{Q}^{(u)}$, we have

$$\mathbf{Q}^{(u)} = \mathbf{X}^{-1}\mathbf{Q} = \frac{\partial(u_1, u_2, \tau)}{\partial(x_1, x_2, x_3)} \frac{\partial(x_1, x_2, x_3)}{\partial(\gamma_1, \gamma_2, \tau)} = \frac{\partial(u_1, u_2, \tau)}{\partial(\gamma_1, \gamma_2, \tau)} \ . \tag{39}$$

The upper left submatrix of the right-hand side of equation (39) equals $\boldsymbol{Q}^{(u)}$, as defined by equation (38). By rewriting the first equation of (35), we get

$$(\mathbf{g}_1, \mathbf{g}_2, \mathbf{x}_\tau)\mathbf{Q}^{(u)} = (\mathbf{x}_{\gamma_1}, \mathbf{x}_{\gamma_2}, \mathbf{x}_\tau) \ . \tag{40}$$

If the incident ray is not tangent to the interface, then the vectors $\mathbf{x}_{\gamma_1}$, $\mathbf{x}_{\gamma_2}$, and $\mathbf{x}_\tau$ are linearly independent, so we have:

$$\mathbf{Q}^{(u)} = \begin{bmatrix} \boldsymbol{Q}^{(u)} & 0 \\ & 0 \\ \times \ \times & 1 \end{bmatrix} . \tag{41}$$

In (41), the crosses "×" denote possibly non-zero entries.

In order to get a similar result for $\mathbf{P}^{(u)}$, we define the 2×2 matrix of second-order travel-time derivatives along the interface:

$$\boldsymbol{M}^{(u)} = T_{\mu\nu} = \frac{\partial^2 T}{\partial u_\mu \partial u_\nu} = \frac{\partial(p_1^{(u)}, p_2^{(u)})}{\partial(u_1, u_2)} = \boldsymbol{P}^{(u)} \boldsymbol{Q}^{(u)^{-1}} \ (\mu, \nu = 1, 2) \ . \tag{42}$$

According to the travel-time expansion (37), this is equal to

$$T_{\mu\nu} = \mathbf{p}^T \mathbf{g}_{\mu\nu} + \mathbf{g}_\mu^T \mathbf{M} \mathbf{g}_\nu \ (\mu, \nu = 1, 2) \ . \tag{43}$$

Now, we form the (symmetric) 3×3 matrix $\mathbf{X}^T\mathbf{M}\mathbf{X}$ and note that $\mathbf{g}_\mu^T \mathbf{M} \mathbf{g}_\nu$ is its upper left submatrix. The third column and row of $\mathbf{X}^T\mathbf{M}\mathbf{X}$ are as follows

$$\begin{bmatrix} \mathbf{x}_\tau^T \mathbf{M} \mathbf{g}_\mu \\ \mathbf{x}_\tau^T \mathbf{M} \mathbf{x}_\tau \end{bmatrix} = \begin{bmatrix} \mathbf{p}_\tau{}^T \mathbf{g}_\mu \\ \mathbf{p}_\tau{}^T \mathbf{x}_\tau \end{bmatrix} \ , \begin{bmatrix} \mathbf{x}_\tau^T \mathbf{M} \mathbf{g}_\nu, \mathbf{x}_\tau^T \mathbf{M} \mathbf{x}_\tau \end{bmatrix} = \begin{bmatrix} \mathbf{p}_\tau{}^T \mathbf{g}_\nu, \mathbf{p}_\tau{}^T \mathbf{x}_\tau \end{bmatrix} . \tag{44}$$

We define the 3×3 matrix $\mathbf{M}^{(u)}$ as follows, and obtain, combining (31), (35), (42) and (44),

$$\mathbf{M}^{(u)} = \mathbf{P}^{(u)}\mathbf{Q}^{(u)^{-1}} = \mathbf{X}^T\mathbf{M}\mathbf{X} - \mathbf{W} = \begin{bmatrix} \boldsymbol{M}^{(u)} & 0 \\ & 0 \\ 0 \; 0 & 0 \end{bmatrix} . \tag{45}$$

From (41), (42), and (45) it is easy to see that the 2×2 matrix $\boldsymbol{P}^{(u)}$ is the upper left submatrix of $\mathbf{P}^{(u)}$ and that the other entries of $\mathbf{P}^{(u)}$ are zero:

$$\mathbf{P}^{(u)} = \begin{bmatrix} \boldsymbol{P}^{(u)} & 0 \\ & 0 \\ 0 \; 0 & 0 \end{bmatrix} . \tag{46}$$

Let us make several comments on these results (41) and (46).

1. The 2×2 matrices $\boldsymbol{Q}^{(u)}$ and $\boldsymbol{P}^{(u)}$ allow a direct and straightforward physical interpretation (unlike their 3×3 counterparts). For a given orthonomic ray field, with ray parameters γ_1 and γ_2, $\boldsymbol{Q}^{(u)}$ denotes the partial derivatives of the surface coordinates u_1, u_2 with respect to γ_1, γ_2 (see first equation of (38)); its determinant is directly related to $\det \mathbf{Q}$ (by $\det \boldsymbol{Q}^{(u)} = \det \mathbf{Q}^{(u)} = \det \mathbf{Q} / \det \mathbf{X}$), which is needed in the computation of the geometrical spreading. The 2×2 matrix $\boldsymbol{P}^{(u)}$ is completely determined by the travel-time field along the surface through (42).

2. For given 2×2 matrices $\boldsymbol{Q}^{(u)}$, $\boldsymbol{P}^{(u)}$, and $\boldsymbol{M}^{(u)} = \boldsymbol{P}^{(u)}\boldsymbol{Q}^{(u)^{-1}}$, and given vectors $\mathbf{p}$, $\mathbf{x}_\tau$, $\mathbf{p}_\tau$, $\mathbf{g}_\mu$, $\mathbf{g}_{\mu\nu}$, the full 3×3 matrices $\mathbf{Q}$, $\mathbf{P}$, and $\mathbf{M} = \mathbf{P}\mathbf{Q}^{-1}$ can be reconstructed. Suppose the surface Σ given by (19) is regular and a two-parametric ray field is parametrized by the surface coordinates $(\gamma_1, \gamma_2) = (u_1, u_2)$. Then (39) and (45) yield

$$\mathbf{Q} = \mathbf{X} \ , \quad \text{and} \quad \mathbf{P} = \mathbf{X}^{-T}(\mathbf{M}^{(u)} + \mathbf{W}) \ . \tag{47}$$

For a more general parametrization of the ray field, a reparametrization can be applied using surface-to-surface paraxial matrices (see Section 3.1).

3. The factorization (32) and definition of 3×3 surface paraxial matrices $\mathbf{Q}^{(u)}$ and $\mathbf{P}^{(u)}$, see (34), and 2×2 surface paraxial matrices $\boldsymbol{Q}^{(u)}$ and $\boldsymbol{P}^{(u)}$, see (38), are equally valid for a structural interface as for an arbitrary surface, which does not necessarily represent a contrast in elastic moduli. In fact, the formulation for arbitrary surfaces plays an important role in the introduction of surface-to-surface paraxial matrices, see next section.

4. Across the interface, we can expand the travel time of the R/T field similarly as in equation (37):

$$\tilde{T}(u_1, u_2) = \tilde{T}(0,0) + \tilde{\mathbf{p}}^T \mathbf{g}_\mu u_\mu + \tfrac{1}{2} (\tilde{\mathbf{p}}^T \mathbf{g}_{\mu\nu} + \mathbf{g}_\mu^T \tilde{\mathbf{M}} \mathbf{g}_\nu) u_\mu u_\nu + O(\| \boldsymbol{u} \|^3) \ . \tag{48}$$

The quantities on the right-hand side of (48) follow from the phase matching approach: $T(\mathbf{x})$ and $\tilde{T}(\mathbf{x})$ should match exactly along the interface. The phase matching argument has been used to derive the 4×4 interface propagator in ray-centred coordinates for isotropic media (Červený, 2001). The treatment for the 6×6 propagator in Cartesian coordinates equally applies for the isotropic and anisotropic cases using relatively simple algebra, but carries an intentional redundancy in the form of the ray-tangent and non-eikonal solutions. The matching of the first and second terms on the right-hand sides of (37) and (48) are trivial and represent ray continuity and Snell's law. The matching of the third terms leads to $\tilde{\boldsymbol{M}}^{(u)} = \boldsymbol{M}^{(u)}$ and, therefore, to $\tilde{\mathbf{M}}^{(u)} = \mathbf{M}^{(u)}$ (using the matrices defined in (42) and (45) and their versions across the interface). The matrix $\tilde{\mathbf{M}}^{(u)}$ is transformed back to Cartesian coordinates by applying $\tilde{\mathbf{Y}}^{-1}$ as in equation (36). The result is

$$\tilde{\mathbf{M}} = \tilde{\mathbf{X}}^{-T}(\mathbf{M}^{(u)} + \tilde{\mathbf{W}})\tilde{\mathbf{X}}^{-1} = \mathbf{P}_1\mathbf{P}_2^T + \mathbf{P}_2\mathbf{M}\mathbf{P}_2^T \ , \tag{49}$$

with $\mathbf{P}_1$ and $\mathbf{P}_2$ defined by (15) and (28). Equation (49) is identical to Moser (2004, eq. 52). Finally, we note that the expression $\mathbf{g}_\mu^T\mathbf{M}\mathbf{g}_\nu$ in (43) can be related to equation (4.14.58) for the 2×2 matrix $\tilde{\mathbf{M}}$ in wavefront orthonormal coordinates in Červený (2001).

3. Surface-to-surface paraxial matrices

In analogy to 4×4 surface-to-surface paraxial matrices in ray-centred coordinates in isotropic media (Hubral et al., 1992, Červený, 2001, Section 4.4.7) and in wavefront orthonormal coordinates in anisotropic media (Červený, 2001, Section 4.14.10), we introduce 6×6 surface-to-surface paraxial matrices derived from ray propagators in Cartesian coordinates, and their reduced 4×4 form. The surface-to-surface paraxial matrices yield a direct link between kinematic and dynamic ray tracing and the theory of seismic systems (Bortfeld, 1989), in analogy to optical systems (Herzberger, 1958; Buchdahl, 1970, and more). They provide a framework for transforming kinematic and paraxial ray quantities from one surface to another, and solving various important boundary value ray-tracing problems of four-parametric systems of rays. The surfaces may be geological interfaces, data recording surfaces, or arbitrary auxiliary surfaces in the Earth. The theory of seismic systems for isotropic media was formulated by Bortfeld (1989) and Hubral et al. (1992). The treatment presented here is equally applicable to isotropic and anisotropic inhomogeneous layered media.

3.1 Definition and properties

We consider a general stack of $N+1$ layers, separated by N interfaces, and a ray connecting a source S in the first layer to a receiver R in the $N+1$-st layer (Figure 1, left). The chain rule for the ray propagator $\mathbf{\Pi}(R,S)$ reads:

$$\begin{aligned} \mathbf{\Pi}(R,S) = \mathbf{\Pi}(R,\tilde{\tau}_N)\mathbf{\Pi}(\tilde{\tau}_N,\tau_N)\mathbf{\Pi}(\tau_N,\tilde{\tau}_{N-1})\mathbf{\Pi}(\tilde{\tau}_{N-1},\tau_{N-1}) \\ \cdots \mathbf{\Pi}(\tau_2,\tilde{\tau}_1)\mathbf{\Pi}(\tilde{\tau}_1,\tau_1)\mathbf{\Pi}(\tau_1,S) \ . \end{aligned} \tag{50}$$

Equation (50) is formally the same as (16), but it is applicable for media with structural interfaces. The N interface intersections between S and R are denoted by $\tau_N, \tilde{\tau}_N$, where τ_N is located infinitesimally before the intersection, and $\tilde{\tau}_N$ infinitesimally across the interface ("across" as defined in Section 2.3). The matrices $\mathbf{\Pi}(\tilde{\tau}_k, \tau_k)$ $(k = 1, 2, \ldots, N)$ represent the interface propagator matrices. The notation S and R is shorthand for τ_S and τ_R. It is important to note that the interface transformations in (50) may represent any type: transmission, reflection, and any type of mode conversion. Inserting the factorized expression (32) for interface propagators and multiplying (50) by $\mathbf{Y}(R)$ from the left and by $\tilde{\mathbf{Y}}^{-1}(S)$ from the right yields

$$\mathbf{T}(R, S) = \mathbf{Y}(R)\mathbf{\Pi}(R, S)\tilde{\mathbf{Y}}^{-1}(S) = \mathbf{T}(R, \tilde{\tau}_N) \cdots \mathbf{T}(\tau_k, \tilde{\tau}_{k-1}) \cdots \mathbf{T}(\tau_1, S) \; , \tag{51}$$

where the 6×6 matrix $\mathbf{T}(R, S)$ is the *surface-to-surface paraxial matrix* from S to R and $\mathbf{T}(\tau_k, \tilde{\tau}_{k-1})$ are the individual surface-to-surface paraxial matrices from $\tilde{\tau}_{k-1}$ to τ_k given by

$$\mathbf{T}(\tau_k, \tilde{\tau}_{k-1}) = \mathbf{Y}(\tau_k)\mathbf{\Pi}(\tau_k, \tilde{\tau}_{k-1})\tilde{\mathbf{Y}}^{-1}(\tilde{\tau}_{k-1}) \; . \tag{52}$$

The 6×6 surface-to-surface paraxial matrix $\mathbf{T}(\tau_k, \tilde{\tau}_{k-1})$ given by (52) is symplectic, since all the three factors $\mathbf{Y}(\tau_k)$, $\mathbf{\Pi}(\tau_k, \tilde{\tau}_{k-1})$, $\tilde{\mathbf{Y}}^{-1}(\tilde{\tau}_{k-1})$ are symplectic themselves. Consequently, $\mathbf{T}(R, S)$ in (51) is symplectic and $\det \mathbf{T}(\tau_k, \tilde{\tau}_{k-1}) = \det \mathbf{T}(R, S) = 1$ $(k = 1, 2, \ldots, N)$. We denote

$$\mathbf{T}(R, S) = \begin{bmatrix} \mathbf{A}(R, S) & \mathbf{B}(R, S) \\ \mathbf{C}(R, S) & \mathbf{D}(R, S) \end{bmatrix} \; , \tag{53}$$

where $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ are 3×3 matrices. The inverse of $\mathbf{T}$ is then given by

$$\mathbf{T}^{-1}(R, S) = \mathbf{T}(S, R) = \begin{bmatrix} \mathbf{D}^T(R, S) & -\mathbf{B}^T(R, S) \\ -\mathbf{C}^T(R, S) & \mathbf{A}^T(R, S) \end{bmatrix} \; . \tag{54}$$

The 6×6 surface-to-surface paraxial matrix$\mathbf{T}(R, S)$ depends on the shape and orientation of the surfaces Σ^a and Σ^p and on the choice of their parametrizations, through the matrices $\tilde{\mathbf{Y}}^{-1}(S)$ and $\mathbf{Y}(R)$. It obeys the chain rule, as given by the right-hand side of (51). We emphasize that the anterior and posterior surfaces, together with their parametrizations, can be chosen completely freely.

The case of a ray of zero length between S and R $(S = R)$ deserves special attention. In this case, the propagator property $\mathbf{T}(S, S) = \mathbf{I}$ applies only if 1. the anterior and posterior surfaces are identical, and 2. their parametrizations are identical. The fact that these conditions are not generally satisfied implies that the surface-to-surface paraxial matrix is not a propagator matrix in a strict mathematical sense. For this reason, we prefer the term surface-to-surface paraxial matrix, rather than surface-to-surface propagator matrix (Hubral et al., 1992, 1993, 1995, and Červený, 2001, sec. 4.4.7). In fact, the case of two different surfaces Σ^a and Σ^p intersecting each other at $S = R$ still allows to set up a surface-to-surface paraxial matrix, even although the ray between S and R has zero length (but with a given slowness vector $\mathbf{p}$). Such a construction is only valid within the paraxial approximation, and applies only to formal surfaces, as far as they do not intersect with structural interfaces. The construction is useful in redatuming, or applying topographic corrections. Even the case of two identical surfaces Σ^a and Σ^p, but with different parametrizations, allows to set up a surface-to-surface paraxial matrix. Such a construction is valid for any type of surfaces (formal surfaces or structural interfaces) and is useful for reparametrization of the two-parametric ray field.

In Červený (2001, sec. 4.4.7) only the case $\mathbf{T}(S,S) = \mathbf{I}$ is mentioned. Since, however, the relation $\mathbf{T}(S,S) = \mathbf{I}$ is not further needed there, all equations in Section 4.4.7 are still valid, even for the case $\mathbf{T}(S,S) \neq \mathbf{I}$.

3.2 Reduction to $\mathbf{4 \times 4}$ *surface-to-surface paraxial matrix* $\boldsymbol{T}$

Let us recall that the ray propagator $\mathbf{\Pi}(R,S)$ from the source S to the receiver R

$$\begin{bmatrix} \mathbf{Q}(R) \\ \mathbf{P}(R) \end{bmatrix} = \mathbf{\Pi}(R,S) \begin{bmatrix} \mathbf{Q}(S) \\ \mathbf{P}(S) \end{bmatrix} \tag{55}$$

is a linear mapping between the paraxial matrices in Cartesian coordinates at S and at R. Left-multiplying of (55) by $\mathbf{Y}(R)$ gives:

$$\mathbf{Y}(R) \begin{bmatrix} \mathbf{Q}(R) \\ \mathbf{P}(R) \end{bmatrix} = \mathbf{Y}(R)\mathbf{\Pi}(R,S)\mathbf{Y}^{-1}(S)\mathbf{Y}(S) \begin{bmatrix} \mathbf{Q}(S) \\ \mathbf{P}(S) \end{bmatrix} = \mathbf{T}(R,S)\mathbf{Y}(S) \begin{bmatrix} \mathbf{Q}(S) \\ \mathbf{P}(S) \end{bmatrix}. \tag{56}$$

Combining (56) with (34) yields

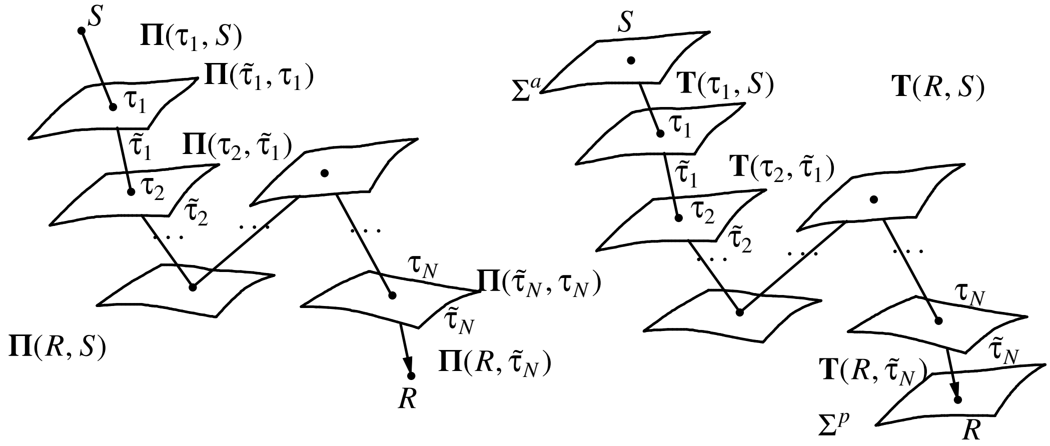

Fig. 1. The ray propagator matrix $\mathbf{\Pi}$ (left) and surface-to-surface paraxial matrix $\mathbf{T}$ (right) - Σ^a and Σ^p denote the anterior and posterior surfaces. The interface ray propagators should be explicitly included in the construction of the ray propagator $\mathbf{\Pi}$. For surface-to-surface paraxial matrices, they are implicit.

$$
\begin{bmatrix} \mathbf{Q}^{(u)}(R) \\ \mathbf{P}^{(u)}(R) \end{bmatrix} = \mathbf{T}(R,S) \begin{bmatrix} \mathbf{Q}^{(u)}(S) \\ \mathbf{P}^{(u)}(S) \end{bmatrix} . \tag{57}
$$

The surface-to-surface paraxial matrix $\mathbf{T}(R,S)$ is therefore a linear mapping between the 3×3 surface paraxial matrices $\mathbf{Q}^{(u)}$ and $\mathbf{P}^{(u)}$, for source and receiver points embedded in the anterior surface Σ^a and posterior surface Σ^p, respectively. However, it has been pointed out before (see Section 2.4) that these matrices still contain the non-eikonal and ray-tangent solutions, and therefore do not allow a straightforward physical interpretation. Using the 2×2 matrices $\boldsymbol{Q}^{(u)}$ and $\boldsymbol{P}^{(u)}$ and inserting the relations (41) and (46) into (57), we obtain the following structure formula

$$
(57) \Rightarrow \begin{bmatrix} \boldsymbol{Q}^{(u)}(R) & & 0 \\ & & 0 \\ \times & \times & 1 \\ \boldsymbol{P}^{(u)}(R) & & 0 \\ & & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{A} & & \times & \boldsymbol{B} & & \times \\ & & \times & & & \times \\ \times & \times & \times & \times & \times & \times \\ \boldsymbol{C} & & \times & \boldsymbol{D} & & \times \\ & & \times & & & \times \\ \times & \times & \times & \times & \times & \times \end{bmatrix} \begin{bmatrix} \boldsymbol{Q}^{(u)}(S) & & 0 \\ & & 0 \\ \times & \times & 1 \\ \boldsymbol{P}^{(u)}(S) & & 0 \\ & & 0 \\ 0 & 0 & 0 \end{bmatrix} , \tag{58}
$$

where the crosses "$\times$" denote possibly non-zero entries, and $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, and $\boldsymbol{D}$ as yet unspecified 2×2 matrices. Several of these entries can be identified. Taking the third columns on the right and left hand side of (58) yields $(0,0,1,0,0,0)^T$ for the third column of $\mathbf{T}(R,S)$. Also, since the inverse of $\mathbf{T}(R,S)$ should have zeroes and ones on the same locations, we can use symplecticity and the formula for the inverse of symplectic matrices (18) to obtain $(0,0,0,0,0,1)$ for the sixth row of $\mathbf{T}(R,S)$. In summary,

$$
\mathbf{T}(R,S) = \begin{bmatrix} \boldsymbol{A} & & 0 & \boldsymbol{B} & & \times \\ & & 0 & & & \times \\ \times & \times & 1 & \times & \times & \times \\ \boldsymbol{C} & & 0 & \boldsymbol{D} & & \times \\ & & 0 & & & \times \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} . \tag{59}
$$

Upon inserting (59) in (58), it is easy to see that the equations for $\boldsymbol{Q}^{(u)}$ and $\boldsymbol{P}^{(u)}$ are decoupled:

$$
\begin{bmatrix} \boldsymbol{Q}^{(u)}(R) \\ \boldsymbol{P}^{(u)}(R) \end{bmatrix} = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} \begin{bmatrix} \boldsymbol{Q}^{(u)}(S) \\ \boldsymbol{P}^{(u)}(S) \end{bmatrix} . \tag{60}
$$

Consequently, we may denote

$$
\boldsymbol{T}(R,S) = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} \tag{61}
$$

as the 4×4 surface-to-surface paraxial matrix. It has similar properties as the 6×6 surface-to-surface paraxial matrix: from expanding the third and sixth columns of $\mathbf{T}(R,S)$ it follows that $\det \boldsymbol{T}(R,S) = \det \mathbf{T}(R,S) = 1$, and the symplecticity of $\boldsymbol{T}(R,S)$ can be easily verified by applying (17) to (59) and multiplying. Therefore, for the inverse of $\boldsymbol{T}$ we have a relation analogous to (18):

$$\boldsymbol{T}(R,S)^{-1} = \boldsymbol{T}(S,R) = \begin{bmatrix} \boldsymbol{D}^T & -\boldsymbol{B}^T \\ -\boldsymbol{C}^T & \boldsymbol{A}^T \end{bmatrix}. \tag{62}$$

For the purpose of computing paraxial rays (see Section 4.1), it is useful to cast (60) in the alternative form

$$\begin{bmatrix} \delta\boldsymbol{u}(R) \\ \delta\boldsymbol{p}^{(u)}(R) \end{bmatrix} = \boldsymbol{T}(R,S) \begin{bmatrix} \delta\boldsymbol{u}(S) \\ \delta\boldsymbol{p}^{(u)}(S) \end{bmatrix}. \tag{63}$$

Here, $\delta\boldsymbol{u}$ denotes small perturbations of $\boldsymbol{u}$ around the reference ray along the relevant surface, and $\delta\boldsymbol{p}^{(u)} = \delta\partial T/\partial\boldsymbol{u}$ small perturbations of $\boldsymbol{p}^{(u)}$ along the same surface.

3.3 Discussion

We summarize the results of the Sections 3.1-3.2. We consider a reference ray and two surfaces intersecting it. Along the reference ray, we perform dynamic ray tracing in Cartesian coordinates. The 6×6 matrices $\mathbf{Y}$ in equation (33) enable the definition of 6×6 surface-to-surface matrices $\mathbf{T}$ in equation (52). Through the matrices $\mathbf{Y}$, full flexibility exists in selecting the surfaces, and in transforming paraxial matrices forward and backward from Cartesian coordinates to surface coordinates, and also between any pair of surfaces. Due to the symplecticity of $\mathbf{Y}$ all these transformations are well-defined and invertible.

Once a pair of anterior and posterior surfaces intersecting the reference ray has been selected, a corresponding 6×6 surface-to-surface matrix $\mathbf{T}$ can be constructed. From this matrix the non-eikonal and ray-tangent solutions can be easily removed by deleting its third and sixth row and columns. Similarly, the 2×2 submatrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ and $\boldsymbol{D}$ of the 4×4 surface-to-surface matrix $\boldsymbol{T}$ can be easily constructed from the 3×3 submatrices of the 6×6 matrix $\mathbf{T}$, by deleting their third row and column. The removal of non-eikonal and ray-tangent solutions implies that the remaining system consists only of physical and non-trivial solutions (the standard paraxial solutions). The reduced 4×4 surface-to-surface matrix $\boldsymbol{T}$ is therefore a convenient framework for many applications, particularly in the solution of various boundary-value ray-tracing problems in a complete four-parametric system of rays (for example, two-point ray tracing). In this way, the surface-to-surface paraxial matrices proposed here generalize the theory of seismic systems (Bortfeld, 1989), and the surface-to-surface paraxial matrices based on dynamic ray tracing in isotropic media (Hubral et al., 1992) to inhomogeneous anisotropic layered structures. The approach proposed here fully relies on dynamic ray tracing in Cartesian coordinates. Only at points where we really need to know the surface-to-surface paraxial matrix $\mathbf{T}(R,S)$, or its 4×4 reduced form, we transform the results obtained in Cartesian coordinates to surface representation.

4. Applications

This section is devoted to the application of 4×4 surface-to-surface paraxial matrices in inhomogeneous anisotropic layered structures, derived in Section 3. The computation of these matrices is based on the conventional dynamic ray tracing in Cartesian coordinates along a reference ray, and on consequent transformation of ray propagator matrices to surface coordinates at surfaces under consideration. We are interested mainly in the solution of boundary value problems of four-parametric system of paraxial rays, connecting a point on a selected anterior surface Σ^a with a point on a selected posterior surface Σ^p. Such four-parametric boundary-value problems cannot be solved by dynamic ray tracing solely, as the dynamic ray tracing considers only orthonomic (two-parametric) system of paraxial rays.

We focus our attention to the two-point eikonal (paraxial travel time from any point on Σ^a to any point on Σ^p), to the initial slowness vectors of paraxial rays at relevant surfaces, to the Fresnel region, and to relative geometrical spreading and its factorization. We show that most equations of this section can be expressed in a similar (or the same) form as derived in the theory of seismic systems, see Bortfeld (1989). The theory of seismic systems, however, does not propose how to relate the 2×2 matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ and $\boldsymbol{D}$ to a known inhomogeneous anisotropic structure. Hubral et al. (1992) showed how these matrices can be calculated by dynamic ray tracing in ray-centred coordinates for isotropic media. Similarly, Červený (2001) showed how they can be calculated in wavefront orthonormal coordinates, both in isotropic and anisotropic media. In this paper, we show how they can be calculated by conventional dynamic ray tracing in Cartesian coordinates, for both isotropic and anisotropic inhomogeneous layered media.

We expect that the surface-to-surface paraxial matrices proposed here may find important applications in seismic exploration for oil and in seismology.

4.1 Paraxial slowness vector

We consider two arbitrary surfaces in a general inhomogeneous, anisotropic medium: the anterior surface Σ^a and the posterior surface Σ^p (see Section 3), and a reference ray connecting the points S on Σ^a and R on Σ^p. We further consider a complete, four-parametric system of paraxial rays, connecting any point of the anterior surface with any point of the posterior surface. We parametrize the anterior surface by $u_\mu(S)$, and similarly the posterior surface by $u_\mu(R)$ ($\mu = 1, 2$). Equation (63) can be used for initial-value ray tracing of paraxial rays, assuming $\delta\boldsymbol{u}(S)$ and $\delta\boldsymbol{p}^{(u)}(S)$ are known at Σ^a.

However, we often wish to solve the *boundary-value ray-tracing problem*: to find the paraxial ray which connects specified points $\delta\boldsymbol{u}(S)$ and $\delta\boldsymbol{u}(R)$ on the anterior and posterior surface. Equation (63) is not directly applicable in this case, as we know $\delta\boldsymbol{u}(S)$ and $\delta\boldsymbol{u}(R)$, but not $\delta\boldsymbol{p}^{(u)}(S)$ or $\delta\boldsymbol{p}^{(u)}(R)$. We can, however, simply compute both $\delta\boldsymbol{p}^{(u)}(S)$

and $\delta \boldsymbol{p}^{(u)}(R)$ from known $\delta \boldsymbol{u}(S)$ and $\delta \boldsymbol{u}(R)$, assuming the 2×2 submatrices $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{D}$ are known. (As shown below, the knowledge of $\boldsymbol{C}$ is not necessary).

In explicit form (upon inserting the submatrices $\boldsymbol{A}(R,S)$, $\boldsymbol{B}(R,S)$, $\boldsymbol{C}(R,S)$, and $\boldsymbol{D}(R,S)$ from (61)), equation (63) reads

$$\delta \boldsymbol{u}(R) = \boldsymbol{A}\delta \boldsymbol{u}(S) + \boldsymbol{B}\delta \boldsymbol{p}^{(u)}(S) \ , \ \ \delta \boldsymbol{p}^{(u)}(R) = \boldsymbol{C}\delta \boldsymbol{u}(S) + \boldsymbol{D}\delta \boldsymbol{p}^{(u)}(S) \ . \tag{64}$$

The equations (64) can be solved for $\delta \boldsymbol{p}^{(u)}(R)$ and $\delta \boldsymbol{p}^{(u)}(S)$:

$$\begin{aligned} \delta \boldsymbol{p}^{(u)}(S) &= -\boldsymbol{B}^{-1}\boldsymbol{A}\ \delta \boldsymbol{u}(S) + \boldsymbol{B}^{-1}\ \delta \boldsymbol{u}(R) \ , \\ \delta \boldsymbol{p}^{(u)}(R) &= (\boldsymbol{C} - \boldsymbol{D}\boldsymbol{B}^{-1}\boldsymbol{A})\ \delta \boldsymbol{u}(S) + \boldsymbol{D}\boldsymbol{B}^{-1}\ \delta \boldsymbol{u}(R) \ . \end{aligned} \tag{65}$$

These equations can be simplified using the symplecticity relation $\boldsymbol{D}\boldsymbol{B}^{-1} - \boldsymbol{C}\boldsymbol{A}^{-1} = \boldsymbol{B}^{-T}\boldsymbol{A}^{-1}$. The final equations are simple:

$$\begin{aligned} \delta \boldsymbol{p}^{(u)}(S) &= -\boldsymbol{B}^{-1}\boldsymbol{A}\ \delta \boldsymbol{u}(S) + \boldsymbol{B}^{-1}\ \delta \boldsymbol{u}(R) \ , \\ \delta \boldsymbol{p}^{(u)}(R) &= -\boldsymbol{B}^{-T}\ \delta \boldsymbol{u}(S) + \boldsymbol{D}\boldsymbol{B}^{-1}\ \delta \boldsymbol{u}(R) \ . \end{aligned} \tag{66}$$

Now consider a reference ray connecting the points $S \in \Sigma^a$ and $R \in \Sigma^p$, and a paraxial ray connecting the points $S' \in \Sigma^a$ and $R' \in \Sigma^p$ (see Fig. 2). The equations (66) then yield:

$$\begin{aligned} \boldsymbol{p}^{(u)}(S') &= \boldsymbol{p}^{(u)}(S) + \delta \boldsymbol{p}^{(u)}(S) = \boldsymbol{p}^{(u)}(S) - \boldsymbol{B}^{-1}\boldsymbol{A}\ \delta \boldsymbol{u}(S) + \boldsymbol{B}^{-1}\delta \boldsymbol{u}(R) \ , \\ \boldsymbol{p}^{(u)}(R') &= \boldsymbol{p}^{(u)}(R) + \delta \boldsymbol{p}^{(u)}(R) = \boldsymbol{p}^{(u)}(R) - \boldsymbol{B}^{-T}\ \delta \boldsymbol{u}(S) + \boldsymbol{D}\boldsymbol{B}^{-1}\delta \boldsymbol{u}(R) \ . \end{aligned} \tag{67}$$

Equation (67) allows the computation of the paraxial slowness vector, and hence the ray velocity vector, at any point S' on Σ^a (close to S) and R' on Σ^p (close to R), corresponding to the paraxial ray connecting the two points S' and R'. We assume that $\boldsymbol{p}^{(u)}(S)$, $\boldsymbol{p}^{(u)}(R)$, $\boldsymbol{A}(R,S)$, $\boldsymbol{B}(R,S)$ and $\boldsymbol{D}(R,S)$, connected with the reference ray, are known.

4.2 Two-point eikonal

We now wish to determine the paraxial travel time $T(R',S')$ from the point $S' \in \Sigma^a$ to the point $R' \in \Sigma^p$, assuming that the travel time $T(R,S)$ along the reference ray is known. For the travel time $T(S')$ we can write

$$T(S') = T(S) + \delta \boldsymbol{u}^T \boldsymbol{p}^{(u)}(S) + \tfrac{1}{2}\,\delta \boldsymbol{u}^T \boldsymbol{M}^{(u)}(S)\delta \boldsymbol{u} \ , \tag{68}$$

and consequently, for $\boldsymbol{p}^{(u)}(S')$,

$$\boldsymbol{p}^{(u)}(S') = \boldsymbol{p}^{(u)}(S) + \boldsymbol{M}^{(u)}(S)\delta \boldsymbol{u} \ . \tag{69}$$

Combining (68) and (69) gives

$$T(S') = T(S) + \tfrac{1}{2}\,\delta \boldsymbol{u}^T(S)(\boldsymbol{p}^{(u)}(S) + \boldsymbol{p}^{(u)}(S')) \ . \tag{70}$$

Let us emphasize that (70) represents a quadratic approximation for $T(S')$ in $\delta \boldsymbol{u}$, not a linear approximation. Using (67), (70) and an analogous equation for R' we obtain:

$$T(R') = T(R) + \delta\boldsymbol{u}^T(R)\boldsymbol{p}^{(u)}(R) + \tfrac{1}{2}\,\delta\boldsymbol{u}^T(R)\delta\boldsymbol{p}^{(u)}(R)\,,$$
$$T(S') = T(S) + \delta\boldsymbol{u}^T(S)\boldsymbol{p}^{(u)}(S) + \tfrac{1}{2}\,\delta\boldsymbol{u}^T(S)\delta\boldsymbol{p}^{(u)}(S)\,. \tag{71}$$

Denoting $T(R', S') = T(R') - T(S')$, $T(R, S) = T(R) - T(S)$, and using (66) we obtain the final expression for $T(R', S')$:

$$T(R', S') = T(R, S) + \delta\boldsymbol{u}^T(R)\boldsymbol{p}^{(u)}(R) - \delta\boldsymbol{u}^T(S)\boldsymbol{p}^{(u)}(S) + \tfrac{1}{2}\,\delta\boldsymbol{u}^T(R)\boldsymbol{D}\boldsymbol{B}^{-1}\delta\boldsymbol{u}(R)$$
$$+ \tfrac{1}{2}\,\delta\boldsymbol{u}^T(S)\boldsymbol{B}^{-1}\boldsymbol{A}\delta\boldsymbol{u}(S) - \delta\boldsymbol{u}^T(R)\boldsymbol{B}^{-T}\delta\boldsymbol{u}(S)\,. \tag{74}$$

Note that expression (74), formulated in local Cartesian coordinates at S and R, is also called the *Hamiltonian point characteristic*, see Bortfeld (1989, eq. 24). The last, "mixed" term plays a great role in various applications. An analogous equation is known from the theory of seismic systems (Bortfeld, 1989, eq. 24), and from the surface-to-surface paraxial theory in ray-centred coordinates in isotropic inhomogeneous media (Hubral et al. , 1992; Červený, 2001). Thus, this well known expression can be used even if we use dynamic ray tracing in Cartesian coordinates and consider inhomogeneous anisotropic layered structures. The 2×2 matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ are, however, different in this case, and are derived in Section 3.

Using (74), we can easily express the 2×2 matrices of the second-order derivatives of the travel-time field with respect to $\delta\boldsymbol{u}(R)$ and $\delta\boldsymbol{u}(S)$, in terms of the matrices $\boldsymbol{A}(R, S)$, $\boldsymbol{B}(R, S)$, and $\boldsymbol{D}(R, S)$:

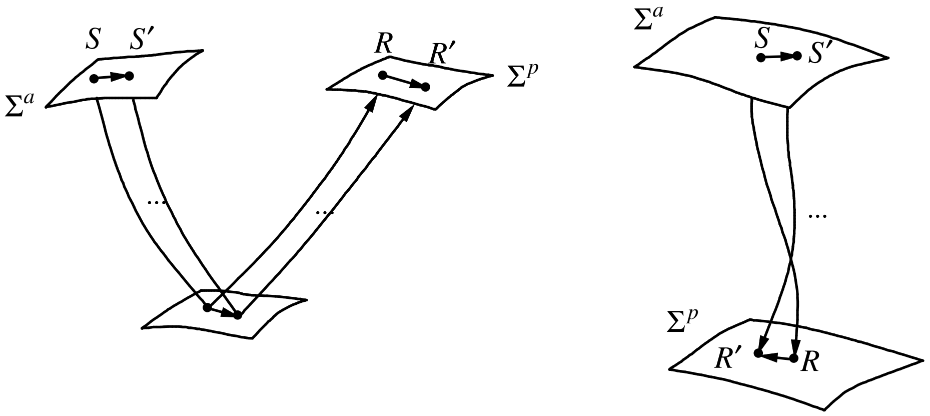

Fig. 2. Paraxial ray computation. Reflection (left) and transmission geometry (right). Dots denote arbitrary sequences of layers between the displayed surfaces.

$$
\begin{aligned}
\boldsymbol{M}^{(u)}(R,S) &= \frac{\partial^2 T(R',S')}{\partial\delta\boldsymbol{u}(R)\partial\delta\boldsymbol{u}(R)} = \boldsymbol{D}\boldsymbol{B}^{-1} \ , \\
\boldsymbol{M}^{(u)}(S,R) &= -\frac{\partial^2 T(R',S')}{\partial\delta\boldsymbol{u}(S)\partial\delta\boldsymbol{u}(S)} = -\boldsymbol{B}^{-1}\boldsymbol{A} \ , \\
\boldsymbol{M}^{(u)\mathrm{mix}}(R,S) &= \frac{\partial^2 T(R',S')}{\partial\delta\boldsymbol{u}(S)\partial\delta\boldsymbol{u}(R)} = -\boldsymbol{B}^{-1} \ .
\end{aligned}
\tag{75}
$$

Here, $\boldsymbol{M}^{(u)}(R,S)$ represents the 2×2 matrix of second-order derivatives of the paraxial travel-time field at R, due to a point source S, and $\boldsymbol{M}^{(u)}(S,R)$ the analogous matrix at S, due to a point source at R. A very important role is played by $\boldsymbol{M}^{(u)\mathrm{mix}}(R,S)$, which represents the 2×2 matrix of mixed second-order derivatives of the paraxial travel-time field, with respect to $\delta\boldsymbol{u}(S)$ and $\delta\boldsymbol{u}(R)$.

4.3 Determination of the surface-to-surface paraxial matrix $\boldsymbol{T}(R,S)$ from travel time along Σ^a and Σ^p

The 2×2 submatrices $\boldsymbol{A}(R,S)$, $\boldsymbol{B}(R,S)$, $\boldsymbol{C}(R,S)$ and $\boldsymbol{D}(R,S)$ of the 4×4 surface-to-surface paraxial matrix $\boldsymbol{T}(R,S)$ can be determined without dynamic ray tracing, from travel-time measurements along the anterior and posterior surfaces, see also Vanelle and Gajewski (2002). For this purpose, we have to determine the three 2×2 matrices $\boldsymbol{M}^{(u)}(R,S)$, $\boldsymbol{M}^{(u)}(S,R)$, and $\boldsymbol{M}^{(u)\mathrm{mix}}(R,S)$. The third equation of (75) yields $\boldsymbol{B}(R,S)$:

$$
\boldsymbol{B}(R,S) = -\,[\,\boldsymbol{M}^{(u)\mathrm{mix}}(R,S)\,]^{-1} \ . \tag{76}
$$

Using the first and second equations of (75), we get $\boldsymbol{D}(R,S)$ and $\boldsymbol{A}(R,S)$:

$$
\boldsymbol{D}(R,S) = \boldsymbol{M}^{(u)}(R,S)\boldsymbol{B}(R,S) \ , \quad \boldsymbol{A}(R,S) = -\,\boldsymbol{B}(R,S)\boldsymbol{M}^{(u)}(S,R) \ . \tag{77}
$$

Finally, $\boldsymbol{C}(R,S)$ can be determined from the symplecticity relation $\boldsymbol{D}^T\boldsymbol{A} - \boldsymbol{B}^T\boldsymbol{C} = \boldsymbol{I}$:

$$
\boldsymbol{C}(R,S) = \boldsymbol{B}(R,S)^{-T}(\boldsymbol{D}(R,S)^T\boldsymbol{A}(R,S) - \boldsymbol{I}) \ . \tag{78}
$$

4.4 Relative geometrical spreading

For a point source, the determinant $\det\boldsymbol{B}(R,S)$ represents one possible form of the relative geometrical spreading, which may be used to compute amplitudes of the ray-theory elastodynamic Green's functions of elementary seismic body waves. It follows from the symplectic properties of the surface-to-surface paraxial matrix $\boldsymbol{T}(R,S)$ given by (61), that $\boldsymbol{B}(R,S) = -\,\boldsymbol{B}^T(S,R)$, see (62). Hence $\det\boldsymbol{B}(R,S)$ is reciprocal: $\det\boldsymbol{B}(R,S) = \det\boldsymbol{B}(S,R)$. The reciprocity of the relative geometrical spreading plays an important role in the studies of the reciprocity of the elementary Green's function.

To derive alternative expressions for $\det\boldsymbol{B}(R,S)$ for a point source at S, we shall use a simple approach, based on the equations (34) and (60). We emphasize our notation: bold roman for 3×3 matrices and bold italic for 2×2 matrices. We understand $\boldsymbol{B}(R,S)$ without superscript to be consistently related to surface coordinates u_μ $(\mu = 1,2)$. Equation (60) yields

$$\boldsymbol{Q}^{(u)}(R) = \boldsymbol{A}(R,S)\boldsymbol{Q}^{(u)}(S) + \boldsymbol{B}(R,S)\boldsymbol{P}^{(u)}(S) \ . \tag{79}$$

At a point source, $\boldsymbol{Q}^{(u)}(S)$ equals the 2×2 null matrix, and we have

$$\det \boldsymbol{B}(R,S) = \det \boldsymbol{Q}^{(u)}(R) / \det \boldsymbol{P}^{(u)}(S) \ . \tag{80}$$

The determinants of $\boldsymbol{Q}^{(u)}(R)$ and $\boldsymbol{P}^{(u)}(S)$ can be simply obtained from (35). The first equation of (35), together with (41), yields

$$\det \boldsymbol{Q}^{(u)}(R) = \det \mathbf{Q}(R) / \det \mathbf{X}(R) \ . \tag{81}$$

The determinant of $\boldsymbol{P}^{(u)}(S)$ for a point source at S is also simple. We use the second equation of (35) and (81) for $\mathbf{Q}$, and obtain

$$\det \boldsymbol{P}^{(u)}(S) = \det \begin{bmatrix} \mathbf{g}_1^T \mathbf{p}_{\gamma_1} & \mathbf{g}_1^T \mathbf{p}_{\gamma_2} \\ \mathbf{g}_2^T \mathbf{p}_{\gamma_1} & \mathbf{g}_2^T \mathbf{p}_{\gamma_2} \end{bmatrix}_S = (\mathbf{g}_1 \times \mathbf{g}_2)_S^T (\mathbf{p}_{\gamma_1} \times \mathbf{p}_{\gamma_2})_S \ , \tag{82}$$

or, alternatively

$$\det \boldsymbol{P}^{(u)}(S) = \|\mathbf{g}_1 \times \mathbf{g}_2\|_S \|\mathbf{p}_{\gamma_1} \times \mathbf{p}_{\gamma_2}\|_S \cos\alpha(S) \ . \tag{83}$$

Here, $\alpha(S)$ is the acute angle between $\mathbf{g}_1 \times \mathbf{g}_2$ and $\mathbf{p}_{\gamma_1} \times \mathbf{p}_{\gamma_2}$ at S. Note that the direction of $\mathbf{p}_{\gamma_1} \times \mathbf{p}_{\gamma_2}$ at a point source S is the same as the direction of $\mathbf{x}_\tau$. Inserting (81) and (83) into (80), yields

$$\det \boldsymbol{B}(R,S) = \frac{[\,\mathbf{t}^T(\mathbf{x}_{\gamma_1} \times \mathbf{x}_{\gamma_2})\,]_R}{\|\mathbf{p}_{\gamma_1} \times \mathbf{p}_{\gamma_2}\|_S \|\mathbf{g}_1 \times \mathbf{g}_2\|_S \|\mathbf{g}_1 \times \mathbf{g}_2\|_R \cos\alpha(S)\cos\alpha(R)} \ , \tag{84}$$

where $\mathbf{t} = \mathbf{x}_\tau / \|\mathbf{x}_\tau\|$ is the unit tangent along the ray (and $\alpha(R)$ is the angle between $\mathbf{g}_1 \times \mathbf{g}_2$ and $\mathbf{x}_\tau$ at R).

Now we shall consider two simple, but important reference cases of (84).

1) The surfaces Σ are perpendicular to the ray. We take $\mathbf{g}_\mu = \mathbf{e}_\mu^\perp$ $(\mu = 1,2)$ to be two mutually perpendicular unit vectors, perpendicular to the ray, so that $\|\mathbf{e}_1^\perp \times \mathbf{e}_2^\perp\| = 1$ and $\cos\alpha = 1$. We use the notation $\boldsymbol{B}^\perp(R,S)$ for this special case of $\boldsymbol{B}(R,S)$. Equation (84) then yields

$$\det \boldsymbol{B}^\perp(R,S) = [\,\mathbf{t}^T(\mathbf{x}_{\gamma_1} \times \mathbf{x}_{\gamma_2})\,]_R / \|\mathbf{p}_{\gamma_1} \times \mathbf{p}_{\gamma_2}\|_S \ , \tag{85}$$

Consequently, equation (84) for arbitrary surfaces Σ at S and R can also be expressed in terms of $\det \boldsymbol{B}^\perp(R,S)$ by:

$$\det \boldsymbol{B}(R,S) = \frac{\det \boldsymbol{B}^\perp(R,S)}{\|\mathbf{g}_1 \times \mathbf{g}_2\|_S \|\mathbf{g}_1 \times \mathbf{g}_2\|_R \cos\alpha(S)\cos\alpha(R)} \ . \tag{86}$$

Here, $\mathbf{g}_\mu$ and $\cos\alpha$ have the same meaning as in (84).

2) Surfaces Σ representing wavefronts. We consider $\tau = T$, take $\mathbf{g}_\mu = \mathbf{e}_\mu$ $(\mu = 1,2)$ as two vectors perpendicular to the slowness vector (tangent to the wavefront), and use the notation $\boldsymbol{B}^{(w)}(R,S)$. Equation (86) then yields

$$\det \boldsymbol{B}^{(w)}(R,S) = \frac{\det \boldsymbol{B}^{\perp}(R,S)}{\| \mathbf{e}_1 \times \mathbf{e}_2 \|_S \, \| \mathbf{e}_1 \times \mathbf{e}_2 \|_R \cos\chi(S)\cos\chi(R)} \; . \tag{87}$$

Here, χ is the angle between the ray and the slowness vector, $\cos\chi = \mathcal{C}/\mathcal{U}$, where $\mathcal{C}$ and $\mathcal{U}$ are the phase and ray velocities, respectively (Červený, 2001, eq. 3.6.18). Using (86) and (87), we can express $\det \boldsymbol{B}(R,S)$ in terms of $\det \boldsymbol{B}^{(w)}(R,S)$:

$$\det \boldsymbol{B}(R,S) = \frac{\cos\chi(R)\cos\chi(S)}{\cos\alpha(R)\cos\alpha(S)} \det \boldsymbol{B}^{(w)}(R,S) \, \frac{\| \mathbf{e}_1 \times \mathbf{e}_2 \|_S \, \| \mathbf{e}_1 \times \mathbf{e}_2 \|_R}{\| \mathbf{g}_1 \times \mathbf{g}_2 \|_S \, \| \mathbf{g}_1 \times \mathbf{g}_2 \|_R} \; . \tag{88}$$

These equations (86)-(88) simplify if we use individual basis vectors as mutually perpendicular unit vectors. For example, equation (86) for $\| \mathbf{g}_1 \times \mathbf{g}_2 \| = 1$ yields

$$\det \boldsymbol{B}(R,S) = \det \boldsymbol{B}^{\perp}(R,S) / \cos\alpha(R)\cos\alpha(S) \; . \tag{89}$$

Similarly, equation (87) for $\| \mathbf{e}_1 \times \mathbf{e}_2 \| = 1$ reads

$$\mathcal{C}(S)\mathcal{C}(R) \det \boldsymbol{B}^{(w)}(R,S) = \mathcal{U}(R)\mathcal{U}(S) \det \boldsymbol{B}^{\perp}(R,S) \; . \tag{90}$$

Finally, we show how the relative geometrical spreading can be expressed in terms of travel-time measurements along the anterior and posterior surfaces Σ^a and Σ^p, namely in terms of the 2×2 matrix of second-order mixed derivatives of the travel-time field, $\boldsymbol{M}^{(u)\mathrm{mix}}(R,S)$. The basic relation for $\det \boldsymbol{B}(R,S)$ in terms of $\det \boldsymbol{M}^{(u)\mathrm{mix}}(R,S)$ follows from the last equation of (75), $\det \boldsymbol{B}(R,S) = 1/\det \boldsymbol{M}^{(u)\mathrm{mix}}(R,S)$. This immediately yields alternative expressions for the relative geometrical spreading, given here. For example, (89) yields

$$\det \boldsymbol{B}^{\perp}(R,S) = \cos\alpha(R)\cos\alpha(S) / \det \boldsymbol{M}^{(u)\mathrm{mix}}(R,S) \; . \tag{91}$$

A note to the computation of the ray theory elementary Green's function in anisotropic inhomogeneous media, using the relative geometrical spreading. For $\tau = T$, several alternative expressions for the Green's function have been published in the seismological literature, which use different expressions for the relative geometrical spreading. Červený (2001, eq. 5.4.24) uses $\det \boldsymbol{B}^{(w)}(R,S)$ (denoted $\boldsymbol{Q}_2(R,S)$ by him), determined from the dynamic ray-tracing system in wavefront orthonormal coordinates. The ray-theory Green's function is expressed there in terms of $|\mathcal{C}(R)\mathcal{C}(S) \det \boldsymbol{Q}_2(R,S)|^{1/2}$, see (91). Schleicher et al. (2001, eqs. 5 and 9) introduce $\det \boldsymbol{B}^{\perp}(R,S)$ (denoted $\det \boldsymbol{Y}(R,S)$ by them), and express the Green's function in terms of $|\mathcal{U}(R)\mathcal{U}(S) \det \boldsymbol{Y}(R,S)|^{1/2}$, see (91). They determine both $\boldsymbol{B}^{\perp}(R,S)$ and $\boldsymbol{B}^{(w)}(R,S)$ from the 3×3 version of (91) and derive (90). Finally, Chapman (2004, eq. 5.4.19) derived equation (85) for $\boldsymbol{B}^{\perp}(R,S)$ (denoted $S^{(3)}(R,S)$ by him), directly from the dynamic ray-tracing system in Cartesian coordinates. He then expresses the ray-theory Green's function in terms of $|\mathcal{U}(R)\mathcal{U}(S)S^{(3)}(R,S)|^{1/2}$. Chapman (2004) also gives a simple link of his relative geometrical spreading to the one given by Kendall et al. (1992), see Chapman (2004, eq. 5.4.23). Our treatment shows that all these expressions for the relative geometrical spreading are consistent and yield the same ray-theory Green's function, although derived in different ways.

Consider a reference ray Ω, connecting a point $S \in \Sigma^a$ and $R \in \Sigma^p$. In addition, we consider one R/T surface Σ^m between Σ^a and Σ^p. The R/T surface may represent a structural interface, or merely a formal surface in a smooth medium. Assume that the reference ray Ω intersects Σ^m at a point M, specified by coordinates $u_\mu(M)$ (specular point of reflection/transmission, see Figure 3). We also consider some other point M', situated on Σ^m close to M, specified by the coordinates $u_\mu(M') = u_\mu(M) + \delta u_\mu(M)$.

We introduce the 2×2 matrices $\boldsymbol{A}_1 = \boldsymbol{A}(M,S)$, $\boldsymbol{B}_1 = \boldsymbol{B}(M,S)$, $\boldsymbol{C}_1 = \boldsymbol{C}(M,S)$, and $\boldsymbol{D}_1 = \boldsymbol{D}(M,S)$, corresponding to the incident part of the ray (from S to M), and matrices $\boldsymbol{A}_2 = \boldsymbol{A}(R,M)$, $\boldsymbol{B}_2 = \boldsymbol{B}(R,M)$, $\boldsymbol{C}_2 = \boldsymbol{C}(R,M)$, and $\boldsymbol{D}_2 = \boldsymbol{D}(R,M)$, corresponding to the R/T part of the ray (from M to R). The notation $\boldsymbol{B}$ has the standard meaning $\boldsymbol{B} = \boldsymbol{B}(R,S)$.

We now wish to determine the Fresnel zone along Σ^m. For a given frequency f, the Fresnel region at Σ^m consists of points M' satisfying the inequality

$$|T(M',S) + T(M',R) - T(R,S)| \le (2f)^{-1} , \tag{92}$$

see Červený (2001, eq. (3.1.46)). It is simple to calculate $T(M',S)$ and $T(M',R)$ using (74). For S fixed, we obtain

$$T(M',S) = T(M,S) + \delta\boldsymbol{u}^T(M)\boldsymbol{p}^{(u)}(M) + \tfrac{1}{2}\,\delta\boldsymbol{u}^T(M)\boldsymbol{D}_1\boldsymbol{B}_1^{-1}\delta\boldsymbol{u}(M) .$$

Similarly, for R fixed, we obtain

$$T(M',R) = T(M,R) - \delta\boldsymbol{u}^T(M)\boldsymbol{p}^{(u)}(M) + \tfrac{1}{2}\,\delta\boldsymbol{u}^T(M)\boldsymbol{B}_2^{-1}\boldsymbol{A}_2\delta\boldsymbol{u}(M) .$$

Combining these two equations into (92), with $T(M,S) + T(M,R) = T(R,S)$, yields

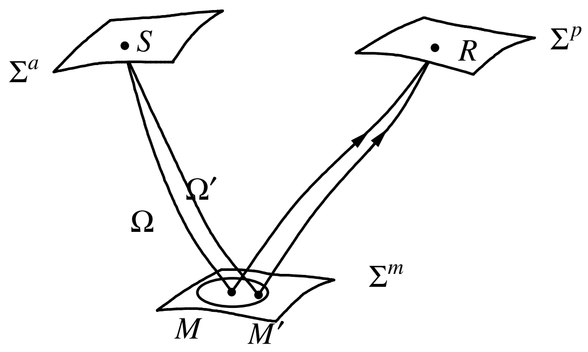

Fig. 3. Fresnel region construction along surface Σ^m.

$$|\,\delta\boldsymbol{u}^T(M)\,(\boldsymbol{D}_1\boldsymbol{B}_1^{-1}+\boldsymbol{B}_2^{-1}\boldsymbol{A}_2)\,\delta\boldsymbol{u}(M)\,|\le f^{-1}\ . \tag{93}$$

Taking into account the chain property $\boldsymbol{B}=\boldsymbol{A}_2\boldsymbol{B}_1+\boldsymbol{B}_2\boldsymbol{D}_1$, we obtain $\boldsymbol{D}_1\boldsymbol{B}_1^{-1}+\boldsymbol{B}_2^{-1}\boldsymbol{A}_2=\boldsymbol{B}_2^{-1}(\boldsymbol{A}_2\boldsymbol{B}_1+\boldsymbol{B}_2\boldsymbol{D}_1)\boldsymbol{B}_1^{-1}=\boldsymbol{B}_2^{-1}\boldsymbol{B}\boldsymbol{B}_1^{-1}$. Consequently, (93) can be condensed to

$$|\,\delta\boldsymbol{u}^T(M)\,\boldsymbol{F}\,\delta\boldsymbol{u}(M)\,|\le f^{-1}\ , \tag{94}$$

where $\boldsymbol{F}$ is given by

$$\boldsymbol{F}=\boldsymbol{B}_2^{-1}\ \boldsymbol{B}\ \boldsymbol{B}_1^{-1}\ . \tag{95}$$

The 2×2 matrix $\boldsymbol{F}$ is called the *Fresnel zone matrix*, and (94) represents a *Fresnel zone* at M. The boundary of the Fresnel zone is obtained by turning the inequality (94) into an equality: replace " $\le$ " by " = ". Note that the Fresnel zone matrix depends on the points S, M, R and that it is symmetric with respect to R and S: $\boldsymbol{F}=\boldsymbol{F}(M;R,S)=\boldsymbol{F}(M;S,R)$. For isotropic media and dynamic ray tracing in ray-centred coordinates, equation (95) was first derived by Hubral et al. (1992, 1993). It is shown here that the equation remains valid even for anisotropic media and dynamic ray tracing in Cartesian coordinates. The 2×2 matrices $\boldsymbol{B}$, $\boldsymbol{B}_1$, and $\boldsymbol{B}_2$, are, however, different in this case, and are derived in Section 3 and in this section.

4.6 Factorization of relative geometrical spreading

Equation (95) offers an important result, related to the factorization of the matrix $\boldsymbol{B}$:

$$\boldsymbol{B}(R,S)=\boldsymbol{B}(R,M)\ \boldsymbol{F}\ \boldsymbol{B}(M,S)\ . \tag{96}$$

Thus, the Fresnel zone matrix $\boldsymbol{F}$ can be used to connect the matrices $\boldsymbol{B}_2=\boldsymbol{B}(R,M)$ and $\boldsymbol{B}_1=\boldsymbol{B}(M,S)$ calculated along the two branches of the ray. For example, for reflected waves, the matrix $\boldsymbol{B}$ can be calculated independently for the incident and the reflected branch. The complete $\boldsymbol{B}(R,S)$ for the whole reflected ray is thus obtained by (96), using the Fresnel zone matrix $\boldsymbol{F}$ at the point of incidence.

In determinants, the equation (96) reads

$$\det\boldsymbol{B}(R,S)=\det\boldsymbol{B}(R,M)\ \det\boldsymbol{F}\ \det\boldsymbol{B}(M,S) \tag{97}$$

This equation can be used in the theory of Kirchhoff-Helmholtz integrals and in true-amplitude migration, and in some other applications. For isotropic media, see Hubral et al. (1995) and Vanelle and Gajewski (2002), for anisotropic media see Schleicher at al. (2001).

5. Concluding remarks

A very general class of problems in paraxial ray tracing can be suitably solved in terms of surface-to-surface paraxial matrices. These include all types of initial and boundary value problems, such as the computation of paraxial slowness vectors in the vicinity of the refererence ray, the two-point eikonal, the evaluation and factorization of relative geometrical spreading, the construction of Fresnel regions, and many more. The surface-to-surface paraxial matrix provides a linear mapping of the paraxial ray matrices between any pair of arbitrary surfaces. The surfaces may represent any surface arising in particular situations: structural interfaces, data recording surfaces, wavefronts or isochron surfaces, or any type of formal auxiliary surface.

The propagation between the surfaces is taken care of by dynamic ray tracing in Cartesian coordinates in the Hamilton formalism. As such it applies equally to isotropic and anisotropic inhomogeneous layered media. At the surfaces themselves the ray propagator matrix is transformed from Cartesian coordinates to surface coordinate form. This form allows full flexibility in the choice of the surfaces at any point of the ray. The redundant ray-tangent and non-eikonal solution which are contained in the dynamic ray tracing in global Cartesian coordinates, can be easily identified and removed. All matrices - the 6×6 ray propagator in global Cartesian coordinates, the 6×6 and 4×4 surface-to-surface paraxial matrix (with and without the redundant solutions) - satisfy the symplectic property. Additionally, in a layered medium the interface propagator ensures symplecticity at ray intersections with structural interfaces. Therefore, the paraxial ray computations can benefit from the algebraic structure provided by symplecticity, as a result of which several quantities are directly available in analytic form (most importantly the inverse of the propagators). The surface-to-surface paraxial matrices allow derivation of the two-point eikonal or Hamiltonian point characteristic, and therefore generalize the theory of seismic (and optical) systems to fully general anisotropic and inhomogeneous layered media.

We expect the surface-to-surface paraxial matrix to have implications even beyond the applications mentioned, in any forward modeling problem in which paraxial ray tracing plays a role, and in particular in a general anisotropic inhomogeneous layered medium.

Acknowledgments

The authors are greatly indebted to Chris Chapman, Luděk Klimeš, Dirk Kraaijpoel and Ivan Pšenčík for helpful suggestions and valuable discussions. They also thank Einar Iversen and Peter Bakker for very stimulating reviews of the paper. This research has been supported by the Grant Agency of the Czech Republic under the Contracts 205/01/D097 and 205/04/1104, and by the Consortium "Seismic Waves in Complex 3-D Structures". Part of this work has been done during TJM's stays at the Charles University, sponsored by the EC Project EVG3-CT-2002-80006 (MAGMA).

References

Bakker, P.M., 1996. Theory of Anisotropic Dynamic Ray Tracing in Ray-centred Coordinates, *Pure and Applied Geoph.*, **148**, 583-589.

Bortfeld, R., 1989. Geometrical ray theory: Rays and traveltimes in seismic systems (second-order approximation of the traveltimes), *Geophysics*, **54**, 342-349.

Buchdahl, H.A., 1970. *An introduction to Hamiltonian optics*, Cambridge Univ. Press.

Chapman, C.H., 2004. *Fundamentals of Seismic Wave Propagation*, Cambridge Univ. Press.

Červený, V., 1972. Seismic rays and ray intensities in inhomogeneous anisotropic media, *Geophys. J. R. astr. Soc.*, **29**, 1-13.

Červený, V., 2001. *Seismic Ray Theory*, Cambridge University Press.

Farra, V., 1989. Ray perturbation theory for heterogeneous hexagonal anisotropic medium, *Geophys. J. Int.*, **99**, 723-738.

Farra, V., and Le Bégat, S., 1995. Sensitivity of qP-wave traveltimes and polarization vectors to heterogeneity, anisotropy and interfaces, *Geophys. J. Int.*, **121**, 371-384.

Gajewski, D., and Pšenčík, I., 1987. Computation of high-frequency seismic wavefields in 3-D laterally inhomogeneous anisotropic media, *Geophys. J. R. astr. Soc.*, **91**, 383-411.

Gajewski, D., and Pšenčík, I., 1990. Vertical seismic profile synthetics by dynamic ray tracing in laterally varying layered anisotropic structures, *J. Geophys. Res.*, **95**, 11301-11315.

Gjøystdal, H., Iversen, E., Laurain, R., Lecomte, I., Vinje, V., and Åstebøl, K., 2002. Review of ray theory applications in modelling and imaging of seismic data, *Stud. geoph. geod.*, **46**, 113-164.

Guenther, R., 1990. *Modern Optics*, John Wiley & Sons, New York.

Herzberger, M., 1958. *Modern Geometrical Optics*, Interscience Publ. Inc.

Hubral, P., Schleicher, J., and Tygel, M., 1992. Three-dimensional paraxial ray properties: Part I. Basic relations. *J. of Seismic Exploration*, **1**, 265-279. Part II. Applications. *J. of Seismic Exploration*, **1**, 347-362.

Hubral, P., Schleicher, J., Tygel, M., and Hanitzsch, C., 1993. Determination of Fresnel zones from traveltime measurements, *Geophysics*, **58**, 703-712.

Hubral, P., Tygel, M., and Schleicher, J., 1995. Geometrical-spreading and ray-caustic decomposition of elementary seismic waves, *Geophysics*, **60**, 1195-1202.

Iversen, E., 2004. Reformulated kinematic and dynamic ray tracing systems for arbitrary anisotropic media, *Stud. geoph. geod.*, **48**, 1-20.

Kendall, J-M., Guest, W.S., and Thomson, C.J., 1992. Ray theory Green's function reciprocity and ray-centered coordinates in anisotropic media, *Geophys. J. Int.*, **108**, 364-371.

Klimeš, L., 1994. Transformations for dynamic ray tracing in anisotropic media, *Wave Motion*, **20**, 261-272.

Luneburg, R.K., 1964. *Mathematical Theory of Optics*, Berkeley, Univ. of California Press.

Moser, T.J., 2004. Review of the anisotropic interface ray propagator: symplecticity, eigenvalues, invariants and applications, *Stud. geoph. geod.*, **48**, 47-73.

Norris, A.N., 1987. A theory of pulse propagation in anisotropic elastic solids, *Wave Motion*, **9**, 509-532.

Popov, M.M., and Pšenčík, I., 1978a. Ray amplitudes in inhomogeneous media with curved interfaces, in *Geofyzikální sborník*, vol. 24, ed. A. Zátopek, pp. 111-129. Praha: Academia.

Popov, M.M., and Pšenčík, I., 1978b. Computation of ray amplitudes in inhomogeneous media with curved interfaces, *Stud. geoph. geod.*, **22**, 248-258.

Schleicher, J., Tygel, M., and Hubral, P., 1993. 3-D true-amplitude finite-offset migration, *Geophysics*, **58**, 1112-1126.

Schleicher, J., Tygel, M., Ursin, B., and Bleistein, N., 2001. The Kirchhoff-Helmholtz integral for anisotropic elastic media, *Wave Motion*, **34**, 353-364.

Shearer, P.M., and Chapman, C.H., 1989. Ray tracing in azimuthally anisotropic media. I. Results for models of aligned cracks in the upper crust, *Geophys. J.*, **96**, 51-64.

Vanelle, C., and Gajewski, D., 2002. True amplitude migration weights from travel times, *Pure and Applied Geophys.*, **159**, 1583-1599.

Appendix - proof of (28)

For a proof of (28) and (30) the following identities are helpful:

$$\mathbf{Q}^{-1}\mathbf{x}_\tau = \tilde{\mathbf{Q}}^{-1}\tilde{\mathbf{x}}_\tau = \mathbf{P}^{-1}\mathbf{p}_\tau = \tilde{\mathbf{P}}^{-1}\tilde{\mathbf{p}}_\tau = (0, 0, 1)^T \ . \tag{A1}$$

Furthermore, we define

$$\Delta\mathbf{P} = \tilde{\mathbf{P}} - \mathbf{P} \ , \ \ \Delta\mathbf{p}_\tau = \tilde{\mathbf{p}}_\tau - \mathbf{p}_\tau \ , \ \ \Delta\mathbf{p} = \tilde{\mathbf{p}} - \mathbf{p} \ . \tag{A2}$$

Differentiating the ray continuity condition (25) along the interface gives

$$\mathbf{0} = \partial(\tilde{\mathbf{x}} - \mathbf{x})/\partial u_\mu = (\tilde{\mathbf{Q}} - \mathbf{Q})\mathbf{Q}^{-1}\mathbf{g}_\mu \ , \ \ \mu = 1, 2 \ . \tag{A3}$$

Differentiating Snell's law (26) gives

$$0 = \partial\mathbf{g}_\mu^T\Delta\mathbf{p}/\partial u_\nu = \mathbf{g}_{\mu\nu}^T\Delta\mathbf{p} + \mathbf{g}_\mu^T\Delta\mathbf{P}\mathbf{Q}^{-1}\mathbf{g}_\nu \ , \ \ \mu, \nu = 1, 2 \ . \tag{A4}$$

Preservation of the eikonal constraint (9) across the interface gives

$$\mathbf{x}_\tau^T\mathbf{P} - \mathbf{p}_\tau{}^T\mathbf{Q} = \tilde{\mathbf{x}}_\tau^T\tilde{\mathbf{P}} - \tilde{\mathbf{p}}_\tau{}^T\tilde{\mathbf{Q}} \ . \tag{A5}$$

Now consider the following expression:

$$\begin{aligned}
&\tilde{\mathbf{X}}^T\tilde{\mathbf{P}}\mathbf{Q}^{-1}\mathbf{X} - \mathbf{X}^T\mathbf{P}\mathbf{Q}^{-1}\mathbf{X} = \\
&= \begin{bmatrix} \mathbf{g}_1^T \\ \mathbf{g}_2^T \\ \tilde{\mathbf{x}}_\tau^T \end{bmatrix} \tilde{\mathbf{P}}\,(\mathbf{Q}^{-1}\mathbf{g}_1, \mathbf{Q}^{-1}\mathbf{g}_2, \mathbf{Q}^{-1}\mathbf{x}_\tau) - \begin{bmatrix} \mathbf{g}_1^T \\ \mathbf{g}_2^T \\ \mathbf{x}_\tau^T \end{bmatrix} \mathbf{P}\,(\mathbf{Q}^{-1}\mathbf{g}_1, \mathbf{Q}^{-1}\mathbf{g}_2, \mathbf{Q}^{-1}\mathbf{x}_\tau) = \\
&= \begin{bmatrix} \mathbf{g}_\mu^T\Delta\mathbf{P}\mathbf{Q}^{-1}\mathbf{g}_\nu & \mathbf{g}_\mu^T\Delta\mathbf{p}_\tau \\ (\tilde{\mathbf{x}}_\tau^T\tilde{\mathbf{P}} - \mathbf{x}_\tau^T\mathbf{P})\mathbf{Q}^{-1}\mathbf{g}_\nu & \tilde{\mathbf{x}}_\tau^T\tilde{\mathbf{p}}_\tau - \mathbf{x}_\tau^T\mathbf{p}_\tau \end{bmatrix} = \\
&= \begin{bmatrix} -\mathbf{g}_{\mu\nu}^T\Delta\mathbf{p} & \mathbf{g}_\mu^T\Delta\mathbf{p}_\tau \\ (\tilde{\mathbf{p}}_\tau{}^T\tilde{\mathbf{Q}} - \mathbf{p}_\tau{}^T\mathbf{Q})\mathbf{Q}^{-1}\mathbf{g}_\nu & \tilde{\mathbf{x}}_\tau^T\tilde{\mathbf{p}}_\tau - \mathbf{x}_\tau^T\mathbf{p}_\tau \end{bmatrix} = \\
&= \begin{bmatrix} -\mathbf{g}_{\mu\nu}^T\Delta\mathbf{p} & \mathbf{g}_\mu^T\Delta\mathbf{p}_\tau \\ \Delta\mathbf{p}_\tau{}^T\mathbf{g}_\nu & \tilde{\mathbf{x}}_\tau^T\tilde{\mathbf{p}}_\tau - \mathbf{x}_\tau^T\mathbf{p}_\tau \end{bmatrix} \\
&= \mathbf{R} \ ,
\end{aligned} \tag{A6}$$

as defined by (30). (A1) and (A3) yield

$$\tilde{\mathbf{Q}} = \tilde{\mathbf{X}}\mathbf{X}^{-1}\mathbf{Q} \ , \tag{A7}$$

and (A6) yields

$$\tilde{\mathbf{P}} = \tilde{\mathbf{X}}^{-T}\mathbf{R}\mathbf{X}^{-1}\mathbf{Q} + \tilde{\mathbf{X}}^{-T}\mathbf{X}^T\mathbf{P} \ , \tag{A8}$$

which proves (28). The symplecticity of (28) follows from the symmetry of **R** and the symplecticity relation (17).