Gaussian beams in inhomogeneous anisotropic layered structures using dynamic ray tracing in Cartesian coordinates

Vlastislav Červený [1]) and Ivan Pšenčík [2])

1) 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.
2) Institute of Geophysics, Acad. Sci. of Czech Republic, Boční II, 141 31 Praha 4, Czech Republic. E-mail: ip@ig.cas.cz.

Summary

Gaussian beams, approximate solutions of elastodynamic equation concentrated close to rays of high-frequency seismic body waves, propagating in inhomogeneous anisotropic layered structures, are studied. They have Gaussian amplitude distribution along any straight line profile intersecting the ray. At any point of the ray, the Gaussian distribution of amplitudes is controlled by the 2×2 complex-valued symmetric matrix $\mathbf{M}^{(y)}$ of the second derivatives of the travel-time field with respect to wavefront orthonormal coordinates y_1, y_2, local Cartesian coordinates in a plane tangential to the wavefront with its origin at the central ray. Matrix $\mathbf{M}^{(y)}$ can be simply determined along the ray if the real-valued propagator matrix of the dynamic ray tracing equations (ray propagator matrix) is known and if the value of $\mathbf{M}^{(y)}$ is specified at an initial point of the ray. The ray propagator matrix can be calculated along the ray by solving twice the dynamic ray tracing system: once for the real-valued initial plane-wave conditions and once for the real-valued initial point-source conditions. Alternatively, matrix $\mathbf{M}^{(y)}$ can be determined along the ray by solving the dynamic ray tracing system only once, but for complex-valued initial conditions. The dynamic ray tracing can be performed in various coordinate systems (global Cartesian x_i, local Cartesian y_i, ray-centred q_i, etc.).

Here we use three alternative variants of dynamic ray tracing in Cartesian coordinates: the global Cartesian system x_i, the local Cartesian (wavefront orthonormal) coordinate system y_i, and the simplified version of the DRT system in global Cartesian coordinates x_i. In all these variants, the 2×2 matrix $\mathbf{M}^{(y)}$ may be used to specify suitably the initial conditions for the dynamic ray tracing. We also present a simple local transformation of 2×2 matrix $\mathbf{M}^{(y)}$ to 3×3 matrix of second order derivatives of travel times $\mathbf{M}^{(x)}$ in global Cartesian coordinates. This 3×3 matrix simplifies considerably the computation of Gaussian beams at any paraxial observation point. The paper is self-contained and presents all the equations needed in computing Gaussian beams. The proposed expressions for Gaussian beams are applicable to general 3-D inhomogeneous layered structures of

Seismic Waves in Complex 3-D Structures, Report 20, Charles University, Faculty of Mathematics and Physics, Department of Geophysics, Praha 2010, pp. 127-168.

arbitrary anisotropy (specified by upto 21 independent position-dependent elastic moduli). Possible simplifications are outlined.

Keywords: Body waves, seismic anisotropy, theoretical seismology, wave propagation.

1 Introduction

The method of summation of Gaussian beams is a powerful extension of the ray method. The Gaussian beams are approximate solutions of elastodynamic equation concentrated close to rays, called central, of high-frequency seismic body waves. The amplitudes of Gaussian beams decrease exponentially with the square of the distance from the central ray along any straight-line profile intersecting the ray. This is the reason why these beams are called Gaussian beams. The equations for the Gaussian beams are valid along the whole central ray and have no singularity at caustics.

The Gaussian beams discussed here are not exact solutions of the elastodynamic equation. Exact Gaussian beams can be computed only exceptionally, e.g., for a point source in a homogeneous medium. Such *exact Gaussian beams* are obtained by moving the point source into the complex space (Felsen, 1976). If we consider Gaussian beams propagating in inhomogeneous media, such an approach cannot be used. We can, however, compute exactly or approximately the complex-valued travel time in the vicinity of the central ray. If we determine the travel times in the vicinity of the central ray by exact solution of the eikonal equation and use them in the expressions for Gaussian beams, we speak of *strict Gaussian beams* (Červený, Klimeš and Pšenčík, 2007, p.76). In the vicinity of the central ray, called the paraxial vicinity, the complex-valued travel time of the beam is usually approximated by its Taylor expansion to quadratic terms at any point of the central ray. To distinguish the beams with approximate complex-valued travel times from exact and strict Gaussian beams, we call them *paraxial Gaussian beams*. In seismological literature, however, it is common to call them Gaussian beams, without emphasizing their approximate validity in the paraxial vicinity of the central ray. We adopt the same terminology in this paper.

The theory of Gaussian beams in isotropic heterogeneous layered structures has been described in many papers. For the scalar wave equation, refer to Babich (1968), Babich and Popov (1981), Popov (1982), Červený, Popov and Pšenčík (1982). For the elastodynamic isotropic wave equation, see Červený and Pšenčík (1983a,b), Klimeš (1984), Červený (1985), George, Virieux and Madariaga (1987), White, Norris, Bayliss and Burridge (1987), Popov (2002), Bleistein (2007), Červený, Klimeš and Pšenčík (2007), Kravtsov and Berczynski (2007), Leung, Qian and Burridge (2007). The Gaussian beam summation method has been successfully applied in migration in seismic exploration. A description of the theory of Gaussian beam migration, of its algorithms and excellent results can be found in Hill (1990, 2001), Gray (2005), Vinje, Roberts and Taylor (2008), Gray and Bleistein (2009), Bleistein and Gray (2010). The references to the theory and applications of Gaussian beams propagating in inhomogeneous anisotropic media are not as common. We have to refer to the excellent mathematical treatment by Ralston (1983), devoted to Gaussian beams and propagation of singularities, to Hanyga (1986), to Červený (2001, sec. 5.8) and to Červený and Pšenčík (2010). To seismic migration in anisotropic media, the Gaussian beam summation method was applied by Alkhalifah (1995) and by Zhu, Gray and Wang (2007).

In most of the above publications, beams were used as building elements in the computation of wavefields based on the method of summation of Gaussian beams. The method removes problems of the standard ray method with caustics since Gaussian beams are regular everywhere including caustics. The method of summation of Gaussian beams is not only regular at a caustic point of the central ray, but yields there approximately correct amplitudes, similarly as the ray method at other points of the central ray. The method of summation of Gaussian beams also avoids need for time consuming two-point ray tracing because evaluation of the superposition integral can be performed at any point of the medium sufficiently illuminated by beams. Although reliable two-point ray-tracing procedures are available (see, e.g., Bulant, 1996), which make possible detection of not only first, but also later arrivals, the method of summation of Gaussian beams, combined with the controlled initial-value ray tracing (Bulant, 1996), may be more efficient in retaining later arrivals. In this article, we concentrate, however, on study of individual Gaussian beams only, not on the superposition integrals. The discussion of superposition integrals would increase the length of the article inadmissibly.

Commonly, Gaussian beams have been derived as asymptotic high-frequency one-way solutions of the elastodynamic equation, concentrated close to a ray of a selected seismic body wave. In the vicinity of this ray, the elastodynamic equation is reduced to a parabolic equation, which further leads to a non-linear differential equation of the Riccati type for the 2×2 complex-valued matrix $\mathbf{M}^{(y)}$ of the second derivatives of the travel-time field with respect to wavefront orthonormal (or ray-centred) coordinates and to the transport equation for the complex-valued amplitude factor along the central ray. The non-linear matrix Riccati equation for the 2×2 matrix $\mathbf{M}^{(y)}$ can be linearized by specifying $\mathbf{M}^{(y)} = \mathbf{P}^{(y)}(\mathbf{Q}^{(y)})^{-1}$, where $\mathbf{P}^{(y)}$ and $\mathbf{Q}^{(y)}$ are again 2×2 matrices. Matrices $\mathbf{P}^{(y)}$ and $\mathbf{Q}^{(y)}$ are then solutions of the well-known system of linear ordinary differential equations of the first order, called the dynamic ray tracing (DRT) system.

Actually, the same approach has been used in the paraxial ray method, in which the real-valued travel time is computed not only along the central ray, but also in its "paraxial vicinity". The only difference is that 2×2 matrices $\mathbf{Q}^{(y)}, \mathbf{P}^{(y)}$ and $\mathbf{M}^{(y)}$ are real-valued in the paraxial ray method, but complex-valued for Gaussian beams. Consequently, we can calculate expressions for Gaussian beams using the paraxial ray method. We merely replace the real-valued matrices $\mathbf{Q}^{(y)}, \mathbf{P}^{(y)}, \mathbf{M}^{(y)}$ by complex-valued matrices. The method of deriving Gaussian beams, based on the paraxial ray approach, is also used in this paper.

Other approaches introducing Gaussian beams have also been proposed recently in literature. See, for example, the so-called Eulerian Gaussian beams, introduced by Leung et al. (2007). These approaches, however, are not discussed in this paper.

The basic procedure in the computation of Gaussian beams is dynamic ray tracing (DRT). The dynamic ray tracing system, and, consequently, the expressions for Gaussian beams, can be expressed in various coordinate systems (global Cartesian, local Cartesian, ray-centred, etc.). The Gaussian beams in inhomogeneous anisotropic media using the dynamic ray tracing in ray-centered coordinates were investigated in detail in Červený and Pšenčík (2010). The mentioned paper is self-contained and includes all equations needed to calculate the Gaussian beams in ray-centred coordinates. Here we use the DRT system in two kinds of Cartesian coordinates, namely in global Cartesian coordinates x_i and in local Cartesian wavefront orthonormal coordinates y_i. In addition, we also discuss a simplified version of the DRT system in global Cartesian coordinates. It will be shown that the 2×2 matrix $\mathbf{M}^{(y)}$ of second derivatives of travel time with respect to wavefront orthonormal coordinates y_1, y_2 at an initial point of the central ray may be used in all the three cases to specify suitably the initial conditions for the DRT system. This is a very important fact, as the 2×2 initial matrix $\mathbf{M}^{(y)}$ may be also complex-valued and the dynamic ray tracing then produce complex-valued $\mathbf{M}^{(y)}$ for Gaussian beams along the ray. Thus, varying $\mathbf{M}^{(y)}$ at an initial point, we can produce any amplitude profile of Gaussian beams we wish. The advantages and limitations of individual DRT systems are discussed, but have not yet been numerically studied.

To make this paper self-contained, we include all equations needed to calculate Gaussian beams. For this reason, the paper also includes certain parts which are similar or even the same as those in Červený and Pšenčík (2010). We believe that this will considerably simplify its reading and understanding the content.

The DRT system may be used for the computation of the matrix $\mathbf{M}^{(y)}$ in two alternative ways. a) In one, we first use DRT to determine the real-valued ray propagator matrix. The ray propagator matrix is obtained by solving the DRT system in matrix form twice: once for the initial plane-wave conditions and once for the initial point-source conditions. Once the ray propagator matrix is known, the 2×2 matrix $\mathbf{M}^{(y)}$ can be determined at any point of the central ray from the initial matrix $\mathbf{M}^{(y)}$ by a single matrix multiplication. b) In the other, the DRT is solved directly, using specific complex-valued initial matrix $\mathbf{M}^{(y)}$. In isotropic media, the first alternative has been preferred in computing Gaussian beam synthetic seismograms, and the second alternative in Gaussian beam migration.

To increase the flexibility of the Gaussian beam computation in wavefront orthonormal coordinates at a paraxial (observation) point, it is useful to transform locally the 2×2 matrix $\mathbf{M}^{(y)}$ in wavefront orthonormal coordinates y_1, y_2 to the analogous 3×3 matrix $\mathbf{M}^{(x)}$ in global Cartesian coordinates x_i. We could obtain this matrix directly by using DRT in global Cartesian coordinates as shown later. See also Červený (1972), Ralston (1983), Leung et al. (2007), Tanushev (2008). However, if we know the 2×2 matrix $\mathbf{M}^{(y)}$, we can also determine $\mathbf{M}^{(x)}$ locally by transforming 2×2 matrix $\mathbf{M}^{(y)}$ into 3×3 matrix $\mathbf{M}^{(x)}$. Use of $\mathbf{M}^{(x)}$ is particularly suitable for evaluation of the wavefield at an observation point by the method of the summation of Gaussian beams, connected with the fan of central rays. Let us note that the local use of Cartesian coordinates for the evaluation of Gaussian beams, computed in ray-centred coordinates in isotropic media, was proposed long time ago (Klimeš, 1984, Sec.4.6). Since then, Klimeš' (1984) procedure has been successfully used in numerical modelling of wave propagation in isotropic media using the Gaussian beam summation method. The procedure is simple, efficient, and removes problems with construction of normals from the observation point to the central ray. As the Klimeš' (1984) approach plays even more important role in anisotropic media than in isotropic media, we pay a special attention to it in the following.

To make the paper more readable and understandable, we try to explain the main procedures based on the ray method as simply as possible, and shift the mathematics to appendices. In the appendices, however, we include all ray-theory prerequisities needed in the computations of Gaussian beams in inhomogeneous anisotropic layered media. In Section 2, we present several basic equations of the ray method for inhomogeneous anisotropic layered media, discuss ray tracing and DRT procedures in Cartesian coordinates, and describe the computation of ray-theory amplitudes. In anisotropic media, both ray tracing and dynamic ray tracing are usually more time-consuming than in isotropic media, particularly for media of lower anisotropic symmetry. The computation and applications of the ray propagator matrix of the DRT system are explained. In Sec. 2, we also present an important transformation formula allowing conversion of 2×2 matrix $\mathbf{M}^{(y)}$ in wavefront orthonormal coordinates into the 3×3 matrix $\mathbf{M}^{(x)}$ in global Cartesian coordinates. Section 3 is devoted to paraxial ray approximations. The expressions for paraxial travel times and for the paraxial approximation of the displacement vector are presented there. Finally, Section 4 is devoted to a simple derivation of Gaussian beams in layered inhomogeneous anisotropic media, various aspects of their computation and description of their properties. Section 5 is devoted to additional discussion of computing Gaussian beams in global Cartesian coordinates and to some additional remarks.

To express the equations of the paper in a concise form, we use alternatively the component and matrix notation for vectors and matrices. In the component notation, the upper-case indices (I, J, K,...) take the values 1 or 2, and the lower-case indices (i, j, k,...) the values 1, 2, or 3. The Einstein summation convention is used throughout the paper. The matrices and vectors are denoted by bold upright symbols. The vectors are considered as column matrices. In this way, the scalar product of vectors $\mathbf{a}$ and $\mathbf{b}$ reads $\mathbf{a}^T\mathbf{b}$, the dyadic reads $\mathbf{a}\mathbf{b}^T$. The components of vectors are expressed in global Cartesian coordinates, unless otherwise stated. For matrices, the dimensions are mostly explicitly indicated, particularly in places of possible confusion. The dot above the letter denotes the partial derivative with respect to time, the index following the comma in the subscript indicates a partial derivative with respect to the relevant Cartesian coordinate.

2 Ray tracing, dynamic ray tracing and ray-theory amplitudes in inhomogeneous anisotropic layered structures

In this section, we discuss basic techniques of computing ray-theory quantities of an arbitrary high-frequency seismic body wave propagating in an inhomogeneous anisotropic layered structure.

2.1 Basic equations of the ray method

The Gaussian beams represent an extension of the ray concepts. For this reason, it is useful to introduce some of these concepts briefly.

Let us consider an inhomogeneous anisotropic perfectly elastic medium. The source-free elastodynamic equation for this medium reads:

$$(c_{ijkl}u_{k,l})_{,j} = \rho\, \ddot{u}_i \ . \tag{1}$$

Here $u_i(x_n)$ are the Cartesian components of the displacement vector $\mathbf{u}(x_n)$, $c_{ijkl}(x_n)$ are real-valued elastic moduli, $\rho(x_n)$ the density and x_n the Cartesian coordinates. In the zero-order approximation of the ray method, the time-harmonic solution of (1) for any high-frequency seismic body wave is usually expressed in the following form:

$$\mathbf{u}(x_i, t) = \mathbf{U}(x_i) \exp[-\mathrm{i}\omega(t - T(x_j))] \ . \tag{2}$$

Here $T(x_i)$ is the travel time, $\mathbf{U}(x_i)$ the complex-valued vectorial ray-theory amplitude, ω the circular frequency, and t time. Inserting (2) into (1), and considering only the terms with the highest power of ω (specifically ω^2), we obtain the system of three equations for U_k:

$$(\Gamma_{ik} - \delta_{ik})U_k = 0 \ , \qquad i = 1, 2, 3 \ . \tag{3}$$

The 3×3 matrix $\mathbf{\Gamma}(x_m, p_n)$ is given by the relation

$$\Gamma_{ik}(x_m, p_n) = a_{ijkl}p_j p_l \ , \tag{4}$$

and is usually called the generalized Christoffel matrix. Here a_{ijkl} are the density-normalized elastic moduli

$$a_{ijkl}(x_n) = c_{ijkl}(x_n)/\rho(x_n) \ . \tag{5}$$

The quantities $p_i = \partial T/\partial x_i$ are Cartesian components of the slowness vector $\mathbf{p}$.

The Christoffel matrix with elements (4) has three eigenvalues $G_m(x_i, p_j)$ and three corresponding eigenvectors $\mathbf{g}^{(m)}(x_i, p_j)$, $m = 1, 2, 3$. They correspond to the three relevant elementary waves, propagating in heterogeneous anisotropic media, specifically P, S1 and S2. Since matrix $\mathbf{\Gamma}$ is symmetric and positive definite, all the three eigenvalues G_1, G_2 and G_3 are real-valued and positive. Moreover, they are homogeneous functions of the second degree in p_i. For simplicity, below we consider that all three eigenvalues are different.

Let us consider the m-th elementary wave. It follows from (3) that the eigenvalue G_m of this wave satisfies the relation

$$G_m(x_i, p_j) = 1 \ . \tag{6}$$

Equation (6) is a non-linear partial differential equation of the first order for the travel time function $T(x_i)$. It is usually called the eikonal equation for a heterogeneous anisotropic medium. It can be expressed in Hamiltonian form

$$\mathcal{H}(x_i, p_j) = \tfrac{1}{2} G_m(x_i, p_j) = \tfrac{1}{2} \ . \tag{7}$$

The Hamiltonian $\mathcal{H}(x_i, p_j)$ is used in ray tracing and DRT, see Appendices A and B, from which the travel time and spreading are usually computed.

The vectorial ray-theory amplitude $\mathbf{U}$ of any elementary wave can be expressed in terms of the unit real-valued eigenvector $\mathbf{g}^{(m)}$ of the Christoffel matrix with elements (4) as follows:

$$\mathbf{U}(x_i) = A(x_i)\mathbf{g}^{(m)}(x_i) \ . \tag{8}$$

Here $A(x_i)$ is a complex-valued frequency-independent scalar ray-theory amplitude. See Section 2.4. Equation (8) shows that eigenvector $\mathbf{g}^{(m)}$ specifies the polarisation of the wave under consideration. For this reason, we call $\mathbf{g}^{(m)}(x_i)$ the polarisation vector.

2.2 Ray tracing

Let us consider an arbitrary high-frequency seismic body wave (P, S1, S2; direct, reflected, transmitted, multiply reflected/transmitted, etc.) propagating in a layered medium specified by smooth structural interfaces and by smooth spatial distribution of upto 21 density-normalized elastic moduli inside layers. We can use *ray tracing* with proper initial conditions to compute any ray Ω of the two-parametric (orthonomic) system of rays corresponding to a selected wave, and denote its ray parameters γ_1 and γ_2. The ray tracing system consists of generally non-linear ordinary differential equations of the first order. We can introduce a monotonically increasing sampling parameter γ_3 along ray Ω, which uniquely specifies the position of a point on ray Ω. Sampling parameter γ_3 may be chosen in various ways. In inhomogeneous anisotropic media, it is most convenient to take $\gamma_3 = \tau$, where τ is the travel time T along ray Ω of the wave under consideration. The ray tracing equations for inhomogeneous anisotropic media and the initial conditions for these equations are given in Appendix A. The transformation equations of the ray at a structural interface are also given there.

Alternatively, it is possible to specify a layered anisotropic structure by smooth structural interfaces and by smooth spatial distribution of Thomsen's (1986) parameters (see Zhu, Gray and Wang, 2005, 2007) or weak-anisotropy (WA) parameters (see Pšenčík and Farra, 2005).

From ray tracing, we obtain the coordinates $\mathbf{x}(\tau)$ of the points on the ray trajectory Ω and slowness vectors $\mathbf{p}(\tau)$ at these points. As a by-product of ray tracing, we can determine several other useful quantities, which we shall need in the following: the ray-velocity vector $\boldsymbol{\mathcal{U}}(\tau) = d\mathbf{x}(\tau)/d\tau$, the unit vector $\mathbf{t}(\tau) = \boldsymbol{\mathcal{U}}(\tau)/|\boldsymbol{\mathcal{U}}(\tau)|$ tangent to the ray Ω, the unit vector $\mathbf{N}(\tau) = \mathbf{p}(\tau)/|\mathbf{p}(\tau)|$ perpendicular to the wavefront, the vector $\boldsymbol{\eta}(\tau) = d\mathbf{p}(\tau)/d\tau$, which represents the variations of the slowness vector along the ray, polarization vector $\mathbf{g}(\tau)$, phase velocity $\mathcal{C}(\tau) = 1/|\mathbf{p}(\tau)|$, and ray velocity $\mathcal{U} = |\boldsymbol{\mathcal{U}}(\tau)|$. The ray-velocity vector is also sometimes called the energy-velocity vector or the group-velocity vector. In non-dissipative media, the latter terms have the same meaning. The travel time $T(\tau)$ along ray is determined automatically as it equals the sampling parameter along the ray, $T(\tau) = \tau$. In the following, we consider the so-called initial-valued rays specified by the initial conditions $\tau = \tau_0$, $\mathbf{x}(\tau_0) = \mathbf{x}_0$, $\mathbf{p}(\tau_0) = \mathbf{p}_0$, where τ_0 is an arbitrarily chosen initial time, and $\mathbf{p}_0 = [\mathcal{C}^{-1}(\tau)\mathbf{N}(\tau)]_{\tau=\tau_0}$.

The ray tracing can be used to compute all the above-mentioned quantities only on the considered ray Ω, not in its vicinity. This is, however, not sufficient to calculate the ray-theory amplitudes and/or Gaussian beams concentrated to ray Ω. This is because the ray-theory amplitudes depend on geometrical spreading, which is a function of the ray field, not of a single ray. In the case of Gaussian beams, we also need to compute complex-valued paraxial travel times (the complex-valued travel times in the vicinity of the ray Ω). To be able to compute the quantities related to the ray field, it is necessary to supplement the ray tracing by an additional procedure called *dynamic ray tracing*.

2.3 Dynamic ray tracing in Cartesian coordinates

The DRT is a basic procedure for the computation of geometrical spreading and for the computation of second derivatives of the travel time field with respect to spatial coordinates along the ray. Its results can be also used to compute the ray-theory amplitudes along the ray. Therefore, we speak of dynamic ray tracing in order to distinguish it from the standard kinematic ray tracing. Dynamic ray tracing consists in the solution of a system of linear ordinary differential equation of the first order along the ray Ω. The system may be solved together with ray tracing, or along an already known ray Ω.

The DRT systems can be expressed in various coordinate systems (global Cartesian x_i, local Cartesian y_i, ray-centred q_i, etc.). The DRT system in ray-centred coordinates, with its application in the computation of Gaussian beams was explained in detail in Červený and Pšenčík (2010). This paper is devoted to three versions of DRT in Cartesian coordinate systems. First we briefly discuss the DRT system in global Cartesian coordinate system x_i, and then in local Cartesian (wavefont orthonormal) coordinates y_i. To distinguish the 2×2 matrices in local Cartesian coordinates y_1, y_2 from the analogous 3×3 matrices in global Cartesian coordinates x_1, x_2, x_3, we use superscripts (y) or (x) over relevant symbols. Since the components of vectors are expressed in Cartesian coordinates, we use superscripts (x) and (y) over them only in cases which may lead to misunderstanding. The third version of the DRT is the simplified DRT in global Cartesian coordinates, in which the 2×3 matrices instead of 3×3 matrices are considered. It may represent a more efficient procedure.

An alternative version of the simplified DRT in global Cartesian coordinates, called surface-to-surface DRT, is not treated here. For anisotropic inhomogeneous media, the details of surface-to-surface DRT are described in Červený (2001) and Moser and Červený (2007).

2.3.1 Dynamic ray tracing system in global Cartesian coordinates

Let us consider six paraxial quantities

$$Q_i^{(x)} = \partial x_i / \partial \gamma \ , \qquad P_i^{(x)} = \partial p_i / \partial \gamma \ , \tag{9}$$

where $i = 1, 2, 3$. They describe the changes of the trajectory $\mathbf{x}(\tau)$ and of the relevant slowness vector $\mathbf{p}(\tau)$, caused by the change of γ. The quantity γ is any of the initial conditions x_{i0} or p_{i0}, or any ray coordinate. Paraxial quantities (9) can be computed by solving the DRT system in global Cartesian coordinates, see equations (B-1)-(B-2) in Appendix B.

The dynamic ray tracing system in global Cartesian coordinates is often used to seek 3×3 matrices $\mathbf{Q}^{(x)}$, $\mathbf{P}^{(x)}$, with elements

$$Q_{ij}^{(x)} = \partial x_i / \partial \gamma_j \ , \qquad P_{ij}^{(x)} = \partial p_i / \partial \gamma_j \ . \tag{10}$$

Here $\gamma_i = x_{i0}$, or $\gamma_i = p_{i0}$, or γ_i are the ray coordinates. The DRT system in a matrix form is given by (B-5).

The 3×3 matrices $\mathbf{Q}^{(x)}$ and $\mathbf{P}^{(x)}$ can be used to compute the second-order partial derivatives of the travel time $T(x_m)$ with respect to spatial derivatives x_i:

$$M_{ij}^{(x)} = \partial^2 T(x_m) / \partial x_i \partial x_j \ . \tag{11}$$

It is easy to show that the 3×3 matrix $\mathbf{M}^{(x)}$, with nine components $M_{ij}^{(x)}$, can be simply expressed in terms of 3×3 matrices $\mathbf{Q}^{(x)}$ and $\mathbf{P}^{(x)}$, obtained from the DRT. As

$$M_{ij}^{(x)} Q_{jk}^{(x)} = \frac{\partial^2 T}{\partial x_i \partial x_j} \frac{\partial x_j}{\partial \gamma_k} = \frac{\partial p_i}{\partial \gamma_k} = P_{ik}^{(x)} \ , \tag{12}$$

we obtain

$$\mathbf{M}^{(x)} = \mathbf{P}^{(x)} (\mathbf{Q}^{(x)})^{-1} \ . \tag{13}$$

Actually, the 3×3 matrix $\mathbf{M}^{(x)}$ could be also computed directly from a Riccati equation solved along the ray (Ralston, 1983). The Riccati equation is, however, ordinary differential non-linear and is not as suitable for the computation as the system of linear DRT equations. Moreover, DRT system can be used to construct the powerful 6×6 ray propagator matrix $\mathbf{\Pi}^{(x)}(\tau, \tau_0)$, which plays an important role in the computation of Gaussian beams. The Riccati equation cannot be used for this purpose, as it is nonlinear.

The applications of the DRT are much broader than those described above. As the travel time $T = \tau$ along the ray Ω and its first derivatives $p_i = \partial T / \partial x_i$ are known from ray tracing, and as the second derivatives $M_{ij}^{(x)}$ can be determined using the DRT, we can use a quadratic expansion of the travel time at an arbitrary point of the ray. Consequently, we can determine approximately the travel time $T(x_m)$ even at points x_m situated in the vicinity of the ray Ω. We speak of paraxial travel time and of quadratic (paraxial) vicinity of the ray Ω. We can also construct linear expansion of the slowness vector and compute paraxial rays in the paraxial vicinity of the ray Ω. Actually, the DRT system itself represents the approximate ray tracing system for paraxial rays in a vicinity of the central ray. For this reason, the dynamic ray tracing is often called paraxial ray tracing. Here, we are not interested in paraxial rays and speak of dynamic ray tracing.

It is possible to show that the DRT system in global Cartesian coordinates may be also used in a simplified way. It is sufficient to compute only two columns of 3×3 matrices $\mathbf{Q}^{(x)}$ and $\mathbf{P}^{(x)}$. For more details see Section 2.3.3.

2.3.2 Dynamic ray tracing in wavefront orthonormal coordinates

Instead of the DRT in global Cartesian coordinates x_i, we can alternatively use the DRT system in local Cartesian coordinates y_i, called wavefront orthonormal coordinate system. It can be also used to find suitable initial conditions for $Q_{iJ}^{(x)}$ and $P_{iJ}^{(x)}$ in (B-5). For this reason, we first introduce the wavefront orthonormal coordinates in Section 2.3.2, and after this we return to the computation of $Q_{iJ}^{(x)}$ and $P_{iJ}^{(x)}$ in Section 2.3.3.

We introduce the wavefront orthonormal coordinates only briefly here. For more detailed information see Červený (2001, sec. 4.2.1). An analogous local Cartesian coordinate system was also used by Bakker (1996), who called it the ray-centred coordinate system. Here, however, we strictly keep the term ray-centred coordinate system for system in which one coordinate axis is the ray.

The wavefront orthonormal coordinates y_1, y_2, y_3 are local Cartesian coordinates with their origin $O(\tau)$ moving along the ray with the wavefront, as the wave progresses. Thus, there is infinity of wavefront orthonormal coordinate systems along the ray with the origins situated along the ray Ω. The y_1- and y_2- axes are tangential to the wavefront and mutually perpendicular at $O(\tau)$. The y_3-axes is perpendicular to the wavefront at $O(\tau)$, so that it is parallel to the slowness vector at $O(\tau)$. The advantage of the wavefront orthonormal coordinate systems is that they are locally orthonormal. The disadvantage with respect to ray-centred coordinates is that the ray itself is not one coordinate axis of the system and that the origin of the system varies from a point to a point along the ray.

We now specify the unit basis vectors $\mathbf{e}_1(\tau)$, $\mathbf{e}_2(\tau)$, $\mathbf{e}_3(\tau)$ of the local wavefront orthonormal coordinate system y_1, y_2, y_3 at any point $O(\tau)$ of the ray. The unit basis vector $\mathbf{e}_3(\tau)$ is given by the simple relation $\mathbf{e}_3(\tau) = \mathcal{C}(\tau)\mathbf{p}(\tau)$, where $\mathcal{C}(\tau)$ is the phase velocity and $\mathbf{p}(\tau)$ the slowness vector. The basic vectors $\mathbf{e}_1(\tau)$ and $\mathbf{e}_2(\tau)$ can be introduced using the ordinary differential equation

$$\mathrm{d}\mathbf{e}_I/\mathrm{d}\tau = -(\mathbf{p}^T\mathbf{p})(\mathbf{e}_I^T\boldsymbol{\eta})\mathbf{p} \; . \tag{14}$$

Here $\boldsymbol{\eta} = \mathrm{d}\mathbf{p}/\mathrm{d}\tau$ is known from ray tracing, see (A-3). A great advantage of the differential equation (14) applied along the ray Ω is that the unit basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ form a right-handed system of orthonormal vectors along the whole ray, once they form such a triplet at arbitrary selected initial point $O(\tau_0)$ of the ray.

Consequently, if we use wavefront orthonormal coordinates, we have to solve (14). Actually, it is not necessary to compute both $\mathbf{e}_1(\tau)$ and $\mathbf{e}_2(\tau)$ using (14); it is sufficient to compute only one of them, e.g. $\mathbf{e}_1(\tau)$. The second basis vector $\mathbf{e}_2(\tau)$ can be then determined from $\mathbf{e}_1(\tau)$ and $\mathbf{e}_3(\tau)$ using the condition of orthonormality, $\mathbf{e}_2(\tau) = \mathbf{e}_3(\tau) \times \mathbf{e}_1(\tau)$.

In addition to basis vectors $\mathbf{e}_1(\tau)$ and $\mathbf{e}_2(\tau)$, it is also useful to compute at initial and termination points of the ray the vectors $\mathbf{f}_1(\tau)$ and $\mathbf{f}_2(\tau)$, perpendicular to the ray. They are given by simple relations (Červený and Pšenčík, 2010),

$$\mathbf{f}_1 = \mathcal{C}^{-1}(\mathbf{e}_2 \times \boldsymbol{\mathcal{U}}) \ , \qquad \mathbf{f}_2 = \mathcal{C}^{-1}(\boldsymbol{\mathcal{U}} \times \mathbf{e}_1) \ . \tag{15}$$

The vectors $\mathbf{f}_1$, $\mathbf{f}_2$ are generally non-unit and not perpendicular. They satisfy the following relation

$$(\mathbf{f}_1, \mathbf{f}_2, \mathbf{p})^T (\mathbf{e}_1, \mathbf{e}_2, \boldsymbol{\mathcal{U}}) = \mathbf{I} \ , \tag{16}$$

where $\mathbf{I}$ is the 3×3 identity matrix.

We now introduce the 3×3 transformation matrix $\mathbf{H}(\tau)$ from local wavefront orthonormal coordinate system y_i to the global Cartesian coordinate system x_i,

$$H_{kl} = \partial x_k / \partial y_l = \partial y_l / \partial x_k \ . \tag{17}$$

The 3×3 transformation matrix $\mathbf{H}$ with elements H_{kl} is unitary, $\mathbf{H}^{-1} = \mathbf{H}^T$, and is related to $\mathbf{e}_i$ as follows:

$$H_{iK} = e_{Ki} \ , \qquad H_{i3} = e_{3i} = \mathcal{C} p_i \ . \tag{18}$$

Thus, H_{iK} is the i-th Cartesian component of the unit basis vector $\mathbf{e}_K$.

The paraxial quantities $\partial y_i / \partial \gamma_K$, $\partial p_i^{(y)} / \partial \gamma_K$ in wavefront orthonormal coordinates in the plane tangent to the wavefront at $O(\tau)$ can be denoted as follows:

$$Q_{iK}^{(y)} = \partial y_i / \partial \gamma_K \ , \qquad P_{iK}^{(y)} = \partial p_i^{(y)} / \partial \gamma_K \ . \tag{19}$$

Here γ_K are the ray-parameters, and $p_n^{(y)} = H_{in} p_i^{(x)}$. The partial derivatives in (19) with respect to γ_K are taken for constant τ. It is not difficult to show that the paraxial quantities $Q_{3K}^{(y)}$ and $P_{3K}^{(y)}$ can be calculated from known quantities $Q_{IK}^{(y)}$ and $P_{IK}^{(y)}$:

$$Q_{3K}^{(y)} = 0 \ , \qquad P_{3K}^{(y)} = \mathcal{C}^{-1} (\eta_I^{(y)} Q_{IK}^{(y)} - \mathcal{U}_I^{(y)} P_{IK}^{(y)}) \ , \tag{20}$$

where $\eta_I^{(y)} = H_{iI} \eta_i^{(x)}$ and $\mathcal{U}_i^{(y)} = H_{iI} \mathcal{U}_i^{(x)}$ are the components of vectors $\boldsymbol{\eta}$ and $\boldsymbol{\mathcal{U}}$ in the wavefront orthonormal coordinates. Consequently, we do not need to use the DRT system in wavefront orthonormal coordinates to compute all 6 quantities $Q_{iK}^{(y)}$, $P_{iK}^{(y)}$, but only the 2×2 matrices

$$\mathbf{Q}^{(y)} = \begin{pmatrix} Q_{11}^{(y)} & Q_{12}^{(y)} \\ Q_{21}^{(y)} & Q_{22}^{(y)} \end{pmatrix} \ , \qquad \mathbf{P}^{(y)} = \begin{pmatrix} P_{11}^{(y)} & P_{12}^{(y)} \\ P_{21}^{(y)} & P_{22}^{(y)} \end{pmatrix} \ . \tag{21}$$

These matrices do not depend on the coordinate y_3, and depend only on y_1, y_2 and τ. The DRT system in wavefront orthonormal coordinates y_1, y_2 for the 2×2 matrices $\mathbf{Q}^{(y)}$ and $\mathbf{P}^{(y)}$ is presented in Appendix B, see eq. (B-20).

As soon as the 2×2 matrices $\mathbf{Q}^{(y)}$ and $\mathbf{P}^{(y)}$ are known, we can use them to determine the 2×2 matrix of second derivatives of the travel time field with respect to wavefront orthonormal coordinates y_1 and y_2:

$$M_{IJ}^{(y)} = \frac{\partial^2 T}{\partial y_I \partial y_J} \; . \tag{22}$$

Using the same procedure as in the derivation of (13), we obtain

$$\mathbf{M}^{(y)} = \mathbf{P}^{(y)}(\mathbf{Q}^{(y)})^{-1} \; . \tag{23}$$

As we can see from (B-20) and (B-18), the dynamic ray tracing system in wavefront orthonormal coordinates is more complicated than the dynamic ray tracing system in general Cartesian coordinates see (B-1) and (B-2), but consists only of eight equations (not of eighteen as the DRT in general Cartesian coordinates).

It was shown by Červený (2007) that the DRT in wavefront orthonormal coordinates gives the same 2×2 matrices $\mathbf{Q}^{(y)}$, $\mathbf{P}^{(y)}$ and $\mathbf{M}^{(y)}$ as the DRT in ray-centred coordinates. This is simple to understand, as the wavefront orthonormal coordinates y_1, y_2 are introduced exactly in the same way as the ray-centred coordinates q_1, q_2. This is the reason why most of results of DRT in wavefront orthonormal coordinates y_I are the same as the results of DRT in ray-centred coordinates. For example, we can introduce the 4×4 ray-propagator matrix $\mathbf{\Pi}^{(y)}(R, S)$ in the same way as $\mathbf{\Pi}^{(q)}(R, S)$. Similarly, the 2×2 matrix of second derivatives of the travel-time field $\mathbf{M}^{(y)}(\tau)$ in wavefront orthonormal coordinates is equivalent to the 2×2 matrix of second derivatives in ray-centred coordinates $\mathbf{M}^{(q)}(\tau)$. Thus, the DRT in wavefront orthonormal coordinates y_1, y_2 can be used as a full alternative of DRT in ray-centred coordinates q_1, q_2, described in detail in Červený and Pšenčík (2010). The numerical efficiency of both DRT systems should be also practically the same. Only the transformation matrix $\mathbf{H}$, which is unitary in wavefront orthonormal coordinates, and thus $\mathbf{H}^{-1} = \mathbf{H}^T$, is replaced by the transformation matrix, which must be inverted numerically in ray centred coordinates. The geometrical difference is that the ray-centred coordinate system q_1, q_2, q_3 is curvilinear, but global, and the wavefront orthonormal coordinate system y_1, y_2, y_3 is Cartesian, but local.

2.3.3 Simplified DRT in global Cartesian coordinates

The disadvantage of the DRT system in the wavefront orthonormal coordinates with respect to global Cartesian coordinates is that the DRT system in wavefront orthonormal coordinates is more complicated than in global Cartesian coordinates. The reason is that we have to rotate appropriately the whole system at any step of computation. The disadvantage of the DRT system in global Cartesian coordinates is the number of its equations. We can, however, simplify the DRT in global Cartesian coordinates, and compute only the first two columns of 3×3 matrices $\mathbf{Q}^{(x)}$ and $\mathbf{P}^{(x)}$, namely the elements $Q_{iJ}^{(x)}$ and $P_{iJ}^{(x)}$. These four columns are quite sufficient for the computation of the 2×2 matrices $\mathbf{M}^{(y)}(\tau)$ and $\mathbf{Q}^{(y)}(\tau)(\mathbf{Q}^{(y)}(\tau_0))^{-1}$ at an arbitrary point τ of the central ray.

Consequently, they are sufficient for computing paraxial ray approximations and Gaussian beams, at any point τ of the central ray. The great advantage of the simplified DRT system is that the DRT system itself remains as simple as in global Cartesian coordinates. The disadvantage with respect to complete global Cartesian coordinates is that we have to evaluate the basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ along the ray, similarly as in wavefront orthonormal coordinates. The disadvantage with respect to wavefront orthonormal coordinates is that we must solve four additional equations for $Q^{(x)}_{3I}$ and $P^{(x)}_{3I}$, which are not computed in wavefront orthonormal coordinates. The procedure is as follows:

a) At the initial point τ_0 of the ray, we compute $Q^{(x)}_{iN}(\tau_0)$ and $P^{(x)}_{iN}(\tau_0)$ from $Q^{(y)}_{IN}(\tau_0)$ and $P^{(y)}_{IN}(\tau_0)$.

b) We perform the DRT in global Cartesian coordinates, but only for 2×3 matrices $Q^{(x)}_{iN}(\tau)$ and $P^{(x)}_{iN}(\tau)$, see (B-32).

c) At any point of the ray, considered to be the termination point, we may transform $Q^{(x)}_{iN}(\tau)$ and $P^{(x)}_{iN}(\tau)$ to $Q^{(y)}_{IN}(\tau)$ and $P^{(y)}_{IN}(\tau)$.

The necessary transformation equations were derived by Červený (2001, eq. 4.2.50). They read:

$$Q^{(x)}_{iN} = e_{Ji} Q^{(y)}_{JN} \; , \qquad P^{(x)}_{iN} = e_{Ij} p_i \eta_j Q^{(y)}_{IN} + (e_{Ii} - e_{Ij} p_i \mathcal{U}_j) P^{(y)}_{IN} \; . \tag{24}$$

The inverse transform read:

$$Q^{(y)}_{JK} = e_{Ji} Q^{(x)}_{iK} \; , \qquad P^{(y)}_{JK} = e_{Ji} P^{(x)}_{iK} \; . \tag{25}$$

Here the components of vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{p}, \boldsymbol{\eta}$ and $\boldsymbol{\mathcal{U}}$ are expressed in global Cartesian coordinates x_i. The mutually perpendicular vectors $\mathbf{e}_1, \mathbf{e}_2$ can be chosen arbitrarily at the initial point τ_0 of the ray, in the plane perpendicular to $\mathbf{p}$.

The simplified DRT system in global Cartesian coordinates is presented in Appendix B, see eq. (B-31). Appendix B also explains how to choose initial conditions $P^{(x)}_{iN}(\tau_0)$ and $Q^{(x)}_{iN}(\tau_0)$ when we wish to use (B-32) to compute the 4×4 ray propagator matrix in wavefront orthonormal coordinates.

Let us note that the simplified DRT in Cartesian coordinates can be used very efficiently for the Gaussian beam computations with such codes like ANRAY (Gajewski and Pšenčík, 1991). In fact, the simplified Cartesian DRT system with point source initial conditions $Q^{(y)}_{IN} = 0$ inserted to (24) has been already used in ANRAY, see Pšenčík and Teles (1996). In order to generalize ANRAY for Gaussian beam computations, it will be sufficient to solve the simplified DRT also for plane wave initial conditions $P^{(y)}_{IN}$ inserted to (24)).

2.4 Ray-theory amplitudes

Let us again consider a harmonic high-frequency seismic body wave propagating in a laterally varying, anisotropic layered structure and a ray Ω corresponding to this wave.

The zero-order ray-theory displacement vector $\mathbf{u}(\tau)$ at any point τ on ray Ω is given by the expression resulting from (2) and (8):

$$\mathbf{u}(\tau) = A(\tau)\mathbf{g}(\tau)\exp[-\mathrm{i}\omega(t - T(\tau))] \ . \tag{26}$$

Here $A(\tau)$ is a scalar ray-theory amplitude, which is generally complex-valued, $\mathbf{g}(\tau)$ is the real-valued unit polarization vector, and $T(\tau) = \tau$ is the travel time along ray Ω. The travel time and the polarization vector are determined during the ray tracing. It remains to discuss the computation of the scalar ray-theory amplitude $A(\tau)$. For this, we need the DRT.

The scalar ray-theory amplitude can be determined using the continuation relation along ray Ω, which follows from the transport equation. We consider two points on the ray specified by τ and τ_0, and assume that $A(\tau_0)$ and $\det \mathbf{Q}^{(y)}(\tau_0)$ are known and $\det \mathbf{Q}^{(y)}(\tau_0) \neq 0$. Then the continuation formula reads:

$$A(\tau) = \left[\frac{\rho(\tau_0)\mathcal{C}(\tau_0)}{\rho(\tau)\mathcal{C}(\tau)}\right]^{1/2} \left[\frac{\det \mathbf{Q}^{(y)}(\tau_0)}{\det \mathbf{Q}^{(y)}(\tau)}\right]^{1/2} \mathcal{R}^C A(\tau_0) \ . \tag{27}$$

In equation (27), ρ is the density, $\mathcal{C}$ the phase velocity (known from ray tracing), $\mathcal{R}^C$ is the complete energy reflection/transmission coefficient along ray Ω from τ_0 to τ, and $\mathbf{Q}^{(y)}$ is the 2×2 matrix which can be determined from DRT and whose elements are defined in eq.(21). Matrix $\mathbf{Q}^{(y)}$ is often called the matrix of geometrical spreading, and $|\det \mathbf{Q}^{(y)}|^{1/2}$ the geometrical spreading. In the ray theory, the term $[\det \mathbf{Q}^{(y)}(\tau_0)/\det \mathbf{Q}^{(y)}(\tau)]^{1/2}$ is usually expressed in terms of modulus (related to geometrical spreading) and phase (phase shift due to caustics). The phase shift can be specified by the so-called KMAH index. Here, we use the continuation formula in the form (27), which is convenient in the study of Gaussian beams. Continuation relation (27) is regular as long as $\det \mathbf{Q}^{(y)}(\tau)$ is non-zero. The complete energy R/T coefficient $\mathcal{R}^C$ along Ω from τ_0 to τ is a product of the plane-wave energy R/T coefficients determined at all points of incidence of ray Ω on the structural interfaces between τ_0 and τ. Mode conversions at individual interfaces are automatically included in $\mathcal{R}^C$. The algorithms for computing the plane-wave energy R/T coefficients at structural interfaces are well known, see, e.g., Červený (2001), where other references can also be found. For this reason, we do not discuss these algorithms here. Let us emphasize that in the zero-order ray approximation, reflection/transmission of any high-frequency seismic wave at a curved interface separating two inhomogeneous media is described by plane-wave R/T coefficients. They do not depend on the curvatures of the considered interface and and of the wavefront of the incident wave at the point of incidence at the interface. They also do not depend on the gradients of the density and the density-normalized elastic moduli at the point of incidence on both sides of the interface. Consequently, the R/T coefficients in the zero-order ray approximation depend only on local values of the density and the density-normalized elastic moduli at the point of incidence (on both sides of the interface) and on the angle of incidence. We further emphasize that $\mathcal{R}^C$ is a product of the energy R/T coefficients, not of the displacement R/T coefficients. With the displacement R/T coefficients, eq.(27) would include an additional multiplication factor.

We emphasize the important assumption made in the derivation of the continuation relation (27), namely that $\det \mathbf{Q}^{(y)}(\tau_0) \neq 0$. This means that (27) cannot be used for individual Gaussain beams and individual paraxial ray approximations of the displacement vector, if a point source is situated at the initial point τ_0. In the method of Gaussian beam summation, the point source is approximately simulated by a weighted superposition of Gaussian beams computed along central rays shot to different directions from the point source. Although on each of these rays $\det \mathbf{Q}^{(y)}(\tau_0) \neq 0$, the superposition represents approximately the wavefield generated by a point source see, for example, Hill (2001; App.B).

If we use the simplified DRT (B-31) in global Cartesian coordinates, we can easily calculate $\det \mathbf{Q}^{(y)}(\tau)$ using (25). The same equation can be used even in standart DRT in global Cartesian coordinates.

There is another important aspect related to the use of eq.(27) for the scalar ray-theory amplitude $A(\tau)$, which corresponds to the zero-order approximation of the ray method. The equation is valid along any ray situated in a smoothly varying inhomogeneous medium with smooth structural interfaces. With increasing complexity of the medium (with its decreasing smoothness), one must expect decrease of the accuracy of the scalar ray-theory amplitude $A(\tau)$.

It is useful to express (26) with (27) in the following form:

$$\mathbf{u}(\tau) = \mathbf{U}^{\Omega}(\tau) \left[\frac{\det \mathbf{Q}^{(y)}(\tau_0)}{\det \mathbf{Q}^{(y)}(\tau)} \right]^{1/2} \exp[-\mathrm{i}\omega(t - T(\tau))] \ , \tag{28}$$

where $\mathbf{U}^{\Omega}$ is called the vectorial spreading-free amplitude. It is given by the relation

$$\mathbf{U}^{\Omega}(\tau) = \left[\frac{\rho(\tau_0)\mathcal{C}(\tau_0)}{\rho(\tau)\mathcal{C}(\tau)} \right]^{1/2} \mathcal{R}^C A(\tau_0)\mathbf{g}(\tau) \ . \tag{29}$$

The vectorial spreading-free amplitude depends on the quantities determined only at points on the ray Ω. It does not depend on the paraxial ray field.

2.5 Ray propagator matrix

In the ray method and its various modifications and extensions, particularly in the paraxial ray method and in the computation of Gaussian beams, a very important role is played by the ray propagator matrix $\mathbf{\Pi}(\tau, \tau_0)$. Under the ray propagator matrix $\mathbf{\Pi}(\tau, \tau_0)$ we understand the propagator matrix of the system of linear ordinary differential equations of the first order representing the DRT system along central ray Ω. Two basic properties of the DRT system, which allow us to construct and exploit the powerful propagator matrix concept are the linearity of the system and the fact that it consists of ordinary differential equations of the first-order.

In this section, we discuss the 4×4 real-valued ray propagator matrix $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$ defined as a solution of the DRT system (B-19) in wavefront orthonormal coordinates:

$$\mathrm{d}\mathbf{\Pi}^{(y)}(\tau, \tau_0)/\mathrm{d}\tau = \mathbf{S}^{(y)}(\tau)\mathbf{\Pi}^{(y)}(\tau, \tau_0) \ , \tag{30}$$

which satisfies, at $\tau = \tau_0$, the initial condition:

$$\mathbf{\Pi}^{(y)}(\tau_0, \tau_0) = \mathbf{I} \ . \tag{31}$$

The ray propagator matrix $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$ is a function of τ, for fixed τ_0. Here $\mathbf{S}^{(y)}(\tau)$ is the 4×4 system matrix of the DRT system and $\mathbf{I}$ is the 4×4 identity matrix. Note that the ray propagator matrix is a fundamental matrix composed of four linearly independent solutions of the DRT system. The linearly independent solutions are guaranteed by the initial conditions (31). The 4×4 system matrix $\mathbf{S}^{(y)}(\tau)$ of the DRT systems in wavefront orthonormal coordinates, is shown in Appendix B.

In global Cartesian coordinates x_i, the above equations remain the same; only the dimensions of the relevant matrices are higher. The propagator matrix $\mathbf{\Pi}^{(x)}(\tau, \tau_0)$ and the system matrix $\mathbf{S}^{(x)}(\tau)$ are 6×6, and relevant submatrices are 3×3.

Let us make two important comments to eq.(30) with the initial condition (31). Firstly, all four linearly independent 4×1 solutions of (30) are sought, in fact, for the price not too much higher than the price of one 4×1 solution. The most expensive part in the computation of $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$ using eq.(30) with (31) consists in the determination of the 4×4 system matrix $\mathbf{S}^{(y)}(\tau)$ along the central ray. The system matrix $\mathbf{S}^{(y)}(\tau)$ is, however, the same for all four linearly independent solutions forming $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$. Consequently, there is no great difference between seeking one of the linearly independent columns of $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$ and the whole $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$. Secondly, the solution of DRT system (30) for $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$ is much faster than standard ray tracing. We remind the reader that the DRT system is linear while the ray-tracing system is generally nonlinear. Except for the second-order derivatives of the Hamiltonian $\mathcal{H}(x_i, p_j)$, see (A-2), only the terms already used to determine the right-hand sides of the ray-tracing equations are used to solve (30).

It is common to express the 4×4 ray propagator matrix $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$ in the wavefront orthonormal coordinates in the following form:

$$\mathbf{\Pi}^{(y)}(\tau, \tau_0) = \begin{pmatrix} \mathbf{Q}_1^{(y)}(\tau, \tau_0) & \mathbf{Q}_2^{(y)}(\tau, \tau_0) \\ \mathbf{P}_1^{(y)}(\tau, \tau_0) & \mathbf{P}_2^{(y)}(\tau, \tau_0) \end{pmatrix} \ . \tag{32}$$

Here $\mathbf{Q}_1^{(y)}(\tau, \tau_0)$, $\mathbf{Q}_2^{(y)}(\tau, \tau_0)$, $\mathbf{P}_1^{(y)}(\tau, \tau_0)$ and $\mathbf{P}_2^{(y)}(\tau, \tau_0)$ are 2×2 real-valued matrices. For the orthonomic system of rays, these matrices have a very simple physical meaning following from (31):

a) $\mathbf{Q}_1^{(y)}(\tau, \tau_0)$ and $\mathbf{P}_1^{(y)}(\tau, \tau_0)$ are solutions of the DRT system in matrix form for the initial conditions at $\tau = \tau_0$:

$$\mathbf{Q}_1^{(y)}(\tau_0, \tau_0) = \mathbf{I} \ , \quad \mathbf{P}_1^{(y)}(\tau_0, \tau_0) = \mathbf{0} \ . \tag{33}$$

In (33), $\mathbf{I}$ is the 2×2 identity matrix and $\mathbf{0}$ is the 2×2 null matrix. It is easy to see that these initial conditions correspond to a plane wavefront at $\tau = \tau_0$.

b) $\mathbf{Q}_2^{(y)}(\tau, \tau_0)$ and $\mathbf{P}_2^{(y)}(\tau, \tau_0)$ are solutions of the DRT system in matrix form for the initial conditions at $\tau = \tau_0$:

$$\mathbf{Q}_2^{(y)}(\tau_0, \tau_0) = \mathbf{0} \ , \quad \mathbf{P}_2^{(y)}(\tau_0, \tau_0) = \mathbf{I} \ . \tag{34}$$

It is easy to see that in this case the initial conditions correspond to a point source at $\tau = \tau_0$.

Thus, the ray propagator matrix is determined if the DRT system in a matrix form (B-20) is solved twice: once with the initial plane-wave conditions corresponding to (33) and once with the initial point-source conditions corresponding to (34). Alternatively, we can say that the DRT system (B-17) should be solved four times, with the initial conditions specified by the four columns of the 4×4 identity matrix.

In global Cartesian coordinate system x_i, equations similar to (33) and (34) hold, but for matrices 3×3. In this case, however, equations with the matrices $\mathbf{Q}_1^{(x)}(\tau_0, \tau_0)$, $\mathbf{Q}_2^{(x)}(\tau_0, \tau_0)$, $\mathbf{P}_1^{(x)}(\tau_0, \tau_0)$, and $\mathbf{P}_2^{(x)}(\tau_0, \tau_0)$ do not have the physical meaning of plane wave and point-source initial conditions.

Once we know the ray propagator matrix $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$, we can determine the solution of the DRT system for arbitrary initial conditions, specified at the point $\tau = \tau_0$ of the central ray Ω by the 2×2 matrices $\mathbf{Q}^{(y)}(\tau_0)$ and $\mathbf{P}^{(y)}(\tau_0)$ using a single matrix operation:

$$\begin{pmatrix} \mathbf{Q}^{(y)}(\tau) \\ \mathbf{P}^{(y)}(\tau) \end{pmatrix} = \mathbf{\Pi}^{(y)}(\tau, \tau_0) \begin{pmatrix} \mathbf{Q}^{(y)}(\tau_0) \\ \mathbf{P}^{(y)}(\tau_0) \end{pmatrix} . \tag{35}$$

The ray propagator matrices have many important and interesting properties, which can be conveniently used in the computation of Gaussian beams. Some of these properties are briefly described in Appendix C.

Let us mention that we can seek the solution of the DRT system, without constructing the ray propagator matrix, by only specifying the specific initial conditions of the DRT system. This may reduce computational efforts, but considerably reduces the flexibility, which the use of the ray propagator matrix offers. For more details, see Sections 3.1 and 4.1.2.

The ray propagator matrix $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$ can be used to construct a simple analytical relation for the continuation of the 2×2 matrix $\mathbf{M}^{(y)}$ of the second spatial derivatives of the travel-time field along the ray Ω. Using (23), (32) and (35), we obtain:

$$\begin{aligned} \mathbf{M}^{(y)}(\tau) &= \mathbf{P}^{(y)}(\tau)\mathbf{Q}^{(y)-1}(\tau) = [\mathbf{P}_1^{(y)}(\tau, \tau_0) + \mathbf{P}_2^{(y)}(\tau, \tau_0)\mathbf{M}^{(y)}(\tau_0)] \\ &\times \; [\mathbf{Q}_1^{(y)}(\tau, \tau_0) + \mathbf{Q}_2^{(y)}(\tau, \tau_0)\mathbf{M}^{(y)}(\tau_0)]^{-1} . \end{aligned} \tag{36}$$

This is a very important relation in the theory of Gaussian beams. Once we know the ray propagator matrix $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$ for a given τ_0 and matrix $\mathbf{M}^{(y)}(\tau_0)$ at $\tau = \tau_0$, equation (36) provides a simple way of determining $\mathbf{M}^{(y)}(\tau)$ for arbitrary τ along the whole ray Ω. If we use the interface propagator matrix (B-25) in $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$, see (B-24), equation (36) can also be used along a ray of a wave reflected or transmitted at a structural interface.

Another important application of the ray-propagator matrix $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$ is in the transformation of the factor $[\det \mathbf{Q}^{(y)}(\tau_0)/ \det \mathbf{Q}^{(y)}(\tau)]^{1/2}$ appearing in the formula for ray-theory amplitudes, see eq.(27). We obtain:

$$[\det \mathbf{Q}^{(y)}(\tau_0)/ \det \mathbf{Q}^{(y)}(\tau)]^{1/2} = [\det \Big(\mathbf{Q}_1^{(y)}(\tau, \tau_0) + \mathbf{Q}_2^{(y)}(\tau, \tau_0)\mathbf{M}^{(y)}(\tau_0)\Big)]^{-1/2} . \tag{37}$$

Similarly as eq.(36), eq.(37) plays an important role in the computation of Gaussian beams. The great advantage of the expressions on the r.h.s. of eqs.(36) and (37) is that they are expressed in terms of known submatrices $\mathbf{Q}_1^{(y)}(\tau,\tau_0)$, $\mathbf{Q}_2^{(y)}(\tau,\tau_0)$, $\mathbf{P}_1^{(y)}(\tau,\tau_0)$ and $\mathbf{P}_2^{(y)}(\tau,\tau_0)$, of the propagator matrix and of physically well understandable initial conditions, namely in terms of the matrix $\mathbf{M}^{(y)}(\tau_0)$ of the second spatial derivatives of the travel time field. No other initial values are required.

2.6 Transformation of $\mathbf{M}^{(y)}(\tau)$ to global Cartesian coordinates

For evaluation of paraxial travel times, it is useful to transform the 2×2 matrix $\mathbf{M}^{(y)}(\tau)$ in wavefront orthonormal coordinates to 3×3 matrix $\mathbf{M}^{(x)}(\tau)$ in global Cartesian coordinates. The appropriate relation was derived by Červený and Klimeš (2010) and reads:

$$\mathbf{M}^{(x)} = \mathbf{f}\ \mathbf{M}^{(y)}\ \mathbf{f}^T + \mathbf{p}\boldsymbol{\eta}^T + \boldsymbol{\eta}\mathbf{p}^T - \mathbf{p}\mathbf{p}^T(\boldsymbol{\mathcal{U}}^T\boldsymbol{\eta})\ . \tag{38}$$

Here $\mathbf{M}^{(x)}$ denotes the 3×3 matrix of the second derivatives of the travel-time field with respect to global Cartesian coordinates, with elements $M_{ij}^{(x)} = \partial^2\tau/\partial x_i\partial x_j$, $\mathbf{M}^{(y)}$ is the 2×2 matrix of the second derivatives of the travel-time field with respect to wavefront orthonormal coordinates, with elements $M_{IJ}^{(y)} = \partial^2\tau/\partial y_I\partial y_J$. The quantities $\mathbf{p}$, $\boldsymbol{\eta}$ and $\boldsymbol{\mathcal{U}}$ are slowness, eta and ray-velocity vectors, which are known from ray tracing. Finally, $\mathbf{f}$ is a 3×2 matrix, $\mathbf{f} = (\mathbf{f}_1, \mathbf{f}_2)$, where $\mathbf{f}_I$ are given by (15). The transformation equation (38) can be used for simple computation of the 3×3 matrix $\mathbf{M}^{(x)}$ from the 2×2 matrix $\mathbf{M}^{(y)}$ at any point of central ray Ω.

Note that the 3×3 matrix $\mathbf{M}^{(x)}$ in global Cartesian coordinates can be also computed directly by dynamic ray tracing in global Cartesian coordinates. Consequently, equation (38) is useful mainly when the DRT system is solved in wavefront orthonormal coordinates y_i, or when the simplified DRT in Cartesian coordinates is used.

3 Paraxial ray approximation

The solution of the DRT system and the computations of the ray-propagator matrix considerably extend the possibilities of the ray method. The most important extension consists in the possibility to compute approximately the travel time field T in the "paraxial" vicinity of the ray Ω, not only on the ray Ω itself. We call such travel times the paraxial travel times. The knowledge of the real-valued paraxial travel times is a necessary prerequisite of many other extensions of the ray method including Gaussian beams. For this reason, we first derive useful expressions for the real-valued paraxial travel times. After this, we use the expressions for paraxial travel times in equations for the relevant displacement vector and obtain the so-called paraxial approximation of the displacement vector. We pay attention mostly to paraxial ray approximation expressed in wavefront orthonormal coordinates y_1, y_2. It is not difficult to express the following equations alternatively in terms of the solution of the DRT in global Cartesian coordinates and in simplified global Cartesian coordinates. See Section 5.

3.1 Paraxial travel times

Paraxial travel time is usually specified by Taylor expansion of travel time T to quadratic terms at point R_Ω situated on ray Ω. The wavefront orthonormal coordinates $\mathbf{y} \equiv (y_1, y_2)^T$ are Cartesian coordinates in the plane tangent to the wavefront at R_Ω, with the origin at R_Ω. Then the paraxial travel time $T(R)$ at the point R, situated in the plane tangent to the wavefront at R_Ω, can be expressed in terms of $T(R_\Omega)$ as follows:

$$T(R) = T(R_\Omega) + \tfrac{1}{2}\mathbf{y}^T(R)\mathbf{M}^{(y)}(R_\Omega)\mathbf{y}(R) \; . \tag{39}$$

The linear term with respect to $\mathbf{y}(R)$ is absent in (39), as the wavefront orthonormal components of the slowness vector, $p_1^{(y)}(R_\Omega)$ and $p_2^{(y)}(R_\Omega)$ at R_Ω on Ω are zero (the slowness vector is perpendicular to the wavefront). Equation (39) for the paraxial travel time in wavefront orthonormal coordinates is very simple. It only requires knowledge of $T(R_\Omega)$ and of the 2×2 matrix $\mathbf{M}^{(y)}(R_\Omega) \equiv \mathbf{M}^{(y)}(\tau)$.

There are two approaches to compute the 2×2 matrix $\mathbf{M}^{(y)}(R_\Omega)$ of the second derivatives of the travel-time field at an arbitrary point R_Ω on the central ray Ω. In both approaches, we need to know the 2×2 matrix $\mathbf{M}^{(y)}(\tau_0)$ at an "initial" point τ_0 on Ω. The matrix $\mathbf{M}^{(y)}(\tau_0)$ must be real-valued, symmetric and finite. The first approach requires computation of the 4×4 ray propagator matrix $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$ in wavefront orthonormal coordinates along the central ray Ω, see Sec.2.5. Then we can obtain the 2×2 matrix $\mathbf{M}^{(y)}(\tau)$ using the simple relation (36). The other approach is to compute $\mathbf{M}^{(y)}(\tau)$ from $\mathbf{M}^{(y)}(\tau) = \mathbf{P}^{(y)}(\tau)(\mathbf{Q}^{(y)}(\tau))^{-1}$, where $\mathbf{P}^{(y)}(\tau)$ and $\mathbf{Q}^{(y)}(\tau)$ are calculated directly from the DRT system with specific initial conditions. We can choose the initial conditions $\mathbf{Q}^{(y)}(\tau_0)$ and $\mathbf{P}^{(y)}(\tau_0)$ for the dynamic ray tracing system (B-20) in the following way:

$$\mathbf{Q}^{(y)}(\tau_0) = \mathbf{A} \; , \qquad \mathbf{P}^{(y)}(\tau_0) = \mathbf{M}_0^{(y)}\mathbf{A} \; . \tag{40}$$

Here $\mathbf{A}$ is an arbitrary constant real-valued finite 2×2 matrix, for which $\det \mathbf{A} \neq 0$. The 2×2 matrix $\mathbf{M}_0^{(y)}$ is real-valued, finite and symmetric. Note that $\mathbf{M}^{(y)}(\tau_0) = \mathbf{P}^{(y)}(\tau_0)(\mathbf{Q}^{(y)}(\tau_0))^{-1} = \mathbf{M}_0^{(y)}$. Solving the DRT system (B-20) with initial conditions (40), we can obtain $\mathbf{Q}^{(y)}(\tau)$ and $\mathbf{P}^{(y)}(\tau)$ at an arbitrary point of central ray Ω. From $\mathbf{Q}^{(y)}(\tau)$ and $\mathbf{P}^{(y)}(\tau)$ we then get:

$$\mathbf{M}^{(y)}(\tau) = \mathbf{P}^{(y)}(\tau)(\mathbf{Q}^{(y)}(\tau))^{-1} \qquad [\det \mathbf{Q}^{(y)}(\tau_0)/\det \mathbf{Q}^{(y)}(\tau)]^{1/2} = [\det \mathbf{A}/\det \mathbf{Q}^{(y)}(\tau)]^{1/2} \; . \tag{41}$$

The matrix $\mathbf{M}^{(y)}(\tau)$ obtained by both approaches is the same. The advantage of the first approach is its high flexibility. Once the 4×4 ray propagator matrix $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$ is known, the matrix $\mathbf{M}^{(y)}(\tau)$ can be obtained from (36) for various specifications of $\mathbf{M}^{(y)}(\tau_0)$ with practically no additional work. Contrary to it, the second approach requires repeated solution of DRT for every new specification of $\mathbf{M}^{(y)}(\tau_0)$.

3.2 Paraxial approximation of the displacement vector

In this section, we discuss extension of the zero-order ray-theory expression (28) for the displacement vector, by including the paraxial travel times (39). Consequently, the resulting expression gives displacement vector not only along the ray Ω, but also in its paraxial vicinity. We obtain

$$\begin{aligned} \mathbf{u}^{par}(y_1, y_2, \tau) &= \mathbf{U}^{\Omega}(\tau)[\det(\mathbf{Q}_1^{(y)}(\tau, \tau_0) + \mathbf{Q}_2^{(y)}(\tau, \tau_0)\mathbf{M}^{(y)}(\tau_0))]^{-1/2} \\ &\times \ \exp[-\mathrm{i}\omega(t - T(\tau) - \tfrac{1}{2}\mathbf{y}^T\mathbf{M}^{(y)}(\tau)\mathbf{y})] \ . \end{aligned} \tag{42}$$

Here $\mathbf{y} \equiv (y_1, y_2)^T$ are wavefront orthonormal coordinates, $\mathbf{U}^{\Omega}(\tau)$ is the vectorial spreading-free amplitude, given by (29), and $\mathbf{M}^{(y)}(\tau)$ is the 2×2 matrix of second derivatives of the travel-time field with respect to wavefront orthonormal coordinates. If we know the 4×4 ray propagator matrix $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$ given by (32), the matrix $\mathbf{M}^{(y)}(\tau)$ can be determined from equation (36). The factor $[\det(\mathbf{Q}^{(y)}(\tau_0)/\det\mathbf{Q}^{(y)}(\tau))]^{1/2}$ may be expressed in terms of 2×2 matrices $\mathbf{Q}_1^{(y)}(\tau, \tau_0)$, $\mathbf{Q}_2^{(y)}(\tau, \tau_0)$ and $\mathbf{M}^{(y)}(\tau_0)$ using (37). If we solve the DRT system for specific initial conditions like in eq.(40), the matrix $\mathbf{M}^{(y)}(\tau)$ and the factor $[\det(\mathbf{Q}^{(y)}(\tau_0)/\det\mathbf{Q}^{(y)}(\tau))]^{1/2}$ must be determined from (41) after solving the DRT (B-20) with the initial conditions (40).

As shown in Section 2.4, expression $\mathbf{U}^{\Omega}(\tau)$ cannot be used for a point source situated at the initial point $\tau = \tau_0$. This does not cause any problem in our treatment, as we are interested in Gaussian beams, and in their real-valued version (called the paraxial ray approximation of the displacement vector). In this case, the point source is really not allowed at the initial point. Otherwise, however, paraxial ray approximation for the displacement vector for a point source at τ_0 can be simply derived by a limiting process $[\det\mathbf{Q}^{(y)}(\tau)]_{\tau\to\tau_0} \to 0$. This limiting process is, however, not permitted for Gaussian beams, and we do not discuss it here.

When the 4×4 ray propagator matrix $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$ is known, the computation of paraxial approximation (42) of the displacement vector requires the specification of an additional initial condition to the standard ray-theory initial conditions, namely the specification of the real-valued 2×2 matrix $\mathbf{M}^{(y)}(\tau_0)$ of second derivatives of the travel-time field at an initial point $\tau = \tau_0$ of the ray Ω. No initial conditions specified separately for $\mathbf{Q}^{(y)}(\tau_0)$ and $\mathbf{P}^{(y)}(\tau_0)$ are required. Fortunately, the 2×2 matrix $\mathbf{M}^{(y)}(\tau_0)$ has a simple physical meaning. It can be expressed in terms of the curvature matrix of the wavefront at τ_0. Remember that the 2×2 matrix $\mathbf{M}^{(y)}(\tau_0)$ must be specified symmetric and finite.

In equation (42), we used expansion of the travel time T from the central ray Ω to the paraxial vicinity of Ω. We, however, did not expand the amplitude term in the vicinity of Ω; it was left the same as on the ray Ω. The expansion of the amplitude term into the paraxial vicinity of Ω in layered, isotropic or anisotropic media would be considerably more complicated, and cannot be derived in the framework of standard DRT. It would require development of higher-order paraxial ray approximations. Such higher-order paraxial ray approximations, however, have not yet been investigated.

Another possibility how to improve the accuracy of the paraxial amplitudes is to trace auxiliary rays concentrated close to the central ray Ω and to take them into account in the computation of paraxial amplitudes. Actually, the most efficient and accurate way to do so consists in the computation of weighted summation of paraxial approximations of the displacement vectors (42), concentrated close to rays distributed densely in the region of interest. In the Cartesian coordinates, such method derived in a more sophisticated way, has been usually called the Chapman-Maslov method, see Chapman and Drummond (1982), Thomson and Chapman (1985), Chapman (2004).

To evaluate the paraxial approximation of the displacement vector $\mathbf{u}^{par}(R)$ at a paraxial point $R = R(y_I)$ using (42), we must know the wavefront orthonormal coordinates y_1, y_2 of the point R. The paraxial point R is, however, often specified in Cartesian coordinates. The determination of the wavefront orthonormal coordinates of the point R from its Cartesian coordinates is not a straightforward task. The problem can be fully removed if the 2×2 matrix $\mathbf{M}^{(y)}(\tau)$ is transformed to the 3×3 matrix $\mathbf{M}^{(x)}(\tau)$. This can be easily done with formula (38), described in Sec.2.6. We discuss the use of eq.(38) and its consequences, which hold for paraxial approximation of the displacement vector as well as for Gaussian beams, in detail in Secs.4.2 and 4.3.

4 Gaussian beams

Gaussian beam, concentrated close to the ray of any high-frequency seismic body wave, has a Gaussian amplitude distribution along any straight-line profile intersecting the ray. The amplitudes are frequency dependent. The expression for the Gaussian beam can be derived by a simple generalization of the paraxial approximation of the displacement vector (42), in which the 2×2 matrix $\mathbf{M}^{(y)}(\tau_0)$ of the second derivatives of the travel-time field is complex valued:

$$\mathbf{M}^{(y)}(\tau_0) = \mathrm{Re}\mathbf{M}^{(y)}(\tau_0) + \mathrm{i}\mathrm{Im}\mathbf{M}^{(y)}(\tau_0) \ . \tag{43}$$

In addition, we require that matrix $\mathbf{M}^{(y)}(\tau_0)$ satisfies the following Gaussian-beam existence conditions: a) $\mathbf{M}^{(y)}(\tau_0)$ is symmetric, b) $\mathbf{M}^{(y)}(\tau_0)$ is finite and c) $\mathrm{Im}\mathbf{M}^{(y)}(\tau_0)$ is positive definite. The properties of the DRT system and of the ray-propagator matrix then guarantee that the 2×2 matrix $\mathbf{M}^{(y)}(\tau)$ satisfies the conditions a) - c) at any point of the central ray. As $\mathrm{Im}\mathbf{M}^{(y)}(\tau)$ is positive definite along the whole central ray Ω, the amplitude profile is Gaussian at any point of ray Ω. Once a Gaussian beam, always a Gaussian beam!

Note that $\mathbf{M}^{(y)}(\tau_0)$ can also be chosen purely imaginary:

$$\mathbf{M}^{(y)}(\tau_0) = \mathrm{i}\mathrm{Im}\mathbf{M}^{(y)}(\tau_0) \ . \tag{44}$$

This choice is quite common in Gaussian beam migration in seismic exploration. It corresponds to the local plane wavefront initial conditions at τ_0, $\mathrm{Re}\mathbf{M}^{(y)}(\tau_0) = 0$.

Second important note. In the expression (38) for a 3×3 matrix $\mathbf{M}^{(x)}$ of the second derivatives of the travel time field with respect to global Cartesian coordinates, the only complex-valued quantity on the r.h.s. is $\mathbf{M}^{(y)}$, all other quntities are real-valued. Consequently, the properties of Gaussian beams are again controlled by 2×2 complex-valued 3×3 matrix $\mathbf{M}^{(y)}$. Equation (38) can be simply used to choose initial conditions for dynamic ray tracing using $\mathbf{M}^{(y)}(\tau_0)$, as all other expressions are known from ray tracing $(\mathbf{p}, \boldsymbol{\mathcal{U}}, \boldsymbol{\eta})$. The vectors $\mathbf{f}_1, \mathbf{f}_2$ can be simply determined from unit, mutually perpendicular vectors $\mathbf{e}_1, \mathbf{e}_2$, see (15). At the initial point of the ray, the unit mutually perpendicular vectors $\mathbf{e}_1, \mathbf{e}_2$ are tangential to the wavefront at $\tau = \tau_0$, but they may be arbitrarily rotated around the slowness vector.

Similarly as in Section 3, we have considered even in Section 4 mostly wavefront orthonormal coordinates y_1, y_2. For global Cartesian coordinates x_i and for the simplified global Cartesian coordinates, see Section 5.

4.1 Computation of complex-valued $\mathbf{M}^{(y)}(\tau)$

Similarly as the real-valued 2×2 matrix $\mathbf{M}^{(y)}(\tau)$ in Sec. 3.1, the complex-valued matrix $\mathbf{M}^{(y)}(\tau)$ can also be computed in two ways: a) using the ray propagator matrix $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$; b) using the solution of the DRT system with specific initial conditions.

4.1.1 Use of ray propagator matrix $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$

We consider a real-valued ray Ω, situated in an inhomogeneous anisotropic layered medium. We assume that the dynamic ray tracing in wavefront orthonormal coordinates y_1, y_2 has been performed along ray Ω and that the 4×4 ray propagator matrix $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$ has been determined. The ray propagator matrix $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$ and its 2×2 submatrices $\mathbf{Q}_1^{(y)}(\tau, \tau_0)$, $\mathbf{Q}_2^{(y)}(\tau, \tau_0)$, $\mathbf{P}_1^{(y)}(\tau, \tau_0)$ and $\mathbf{P}_2^{(y)}(\tau, \tau_0)$ are real-valued.

We now consider the complex-valued 2×2 matrix $\mathbf{M}^{(y)}(\tau_0)$, satisfying the Gaussian beam existence conditions. Equation (42) then yields a solution, which has usually been called the Gaussian beam:

$$\begin{aligned} \mathbf{u}^{beam}(y_1, y_2, \tau) &= \mathbf{U}^{\Omega}(\tau)(\det \mathbf{W})^{-1/2} \exp[-\tfrac{1}{2}\omega \mathbf{y}^T \mathrm{Im}\mathbf{M}^{(y)}(\tau)\mathbf{y}] \\ &\times \exp[-\mathrm{i}\omega(t - T(\tau) - \tfrac{1}{2}\mathbf{y}^T \mathrm{Re}\mathbf{M}^{(y)}(\tau)\mathbf{y})] \ . \end{aligned} \tag{45}$$

Here $\mathbf{y} = (y_1, y_2)^T$ and

$$\mathbf{W}(\tau, \tau_0) = \mathbf{Q}_1^{(y)}(\tau, \tau_0) + \mathbf{Q}_2^{(y)}(\tau, \tau_0)(\mathrm{Re}\mathbf{M}^{(y)}(\tau_0) + \mathrm{i}\mathrm{Im}\mathbf{M}^{(y)}(\tau_0)) \ . \tag{46}$$

The Gaussian beam existence conditions guarantee that $\mathbf{M}^{(y)}(\tau)$ is finite along the whole ray Ω. This guarantees that $\det \mathbf{W}(\tau, \tau_0)$ cannot become zero at any point of ray Ω, including point $\tau = \tau_0$. Consequently, matrix $\mathbf{W}(\tau, \tau_0)$ is regular for any τ. For $\tau = \tau_0$, we have $\mathbf{Q}_1^{(y)}(\tau_0, \tau_0) = \mathbf{I}$, $\mathbf{Q}_2^{(y)}(\tau_0, \tau_0) = \mathbf{0}$, so that

$$\mathbf{W}(\tau_0, \tau_0) = \mathbf{I} \ . \tag{47}$$

As $(\det \mathbf{W}(\tau,\tau_0))^{-1/2}$ is, in general, a complex-valued square root, we must be careful in choosing its argument. We can determine it in the following way: a) we put it zero for $\tau = \tau_0$, where $\mathbf{W}(\tau_0,\tau_0)$ is real valued; b) we require that it varies continuously along ray Ω.

Since there are no points along ray Ω, at which $\det \mathbf{W}(\tau,\tau_0)$ is zero, the use of Gaussian beams removes singularities at *caustic points*. This is a very important and useful property of Gaussian beams. The problem of caustics is one of the most serious problems in computing of ray synthetic seismic wave fields in inhomogeneous isotropic or anisotropic media. The method based on the summation of Gaussian beams removes this problem.

As the Gaussian beam is not singular at any point of the ray, it is also not singular at the initial point τ_0 of the ray. Consequently, the wave field generated by a point source at $\tau = \tau_0$ cannot be described by a single Gaussian beam. It can, however, be described approximately by the weighted sum of Gaussian beams. See the discussion at the end of Sec.2.4.

Let us consider again a real-valued ray Ω. Each Gaussian beam connected with this ray is specified by a 2×2 symmetric complex-valued matrix $\mathbf{M}^{(y)}(\tau_0)$ given at an arbitrary point of the ray, $\tau = \tau_0$. Consequently, we can construct a six-parametric system of Gaussian beams connected with each ray Ω. The three parameters $\mathrm{Re}M_{11}^{(y)}(\tau_0)$, $\mathrm{Re}M_{22}^{(y)}(\tau_0)$, and $\mathrm{Re}M_{12}^{(y)}(\tau_0)$, control the shape of the wavefront of the Gaussian beam at τ_0. The other three parameters, $\mathrm{Im}M_{11}^{(y)}(\tau_0)$, $\mathrm{Im}M_{22}^{(y)}(\tau_0)$ and $\mathrm{Im}M_{12}^{(y)}(\tau_0)$, control the width of the Gaussian beam at τ_0. The real-valued travel time T along ray Ω and the spreading-free vectorial amplitudes $\mathbf{U}^{\Omega}$ are the same for all the Gaussian beams connected with ray Ω.

Similarly as in the paraxial approximation of the displacement vector (42), the vectorial spreading-free amplitude $\mathbf{U}^{\Omega}(\tau)$ in the expression (45) for Gaussian beams is the same in the whole plane tangent to the wavefront at τ on Ω. The most efficient and accurate way of taking into account the paraxial changes of the amplitude consists in using the weighted summation of Gaussian beams.

4.1.2 Use of DRT with specific complex-valued initial conditions

If we do not wish to compute the ray propagator matrix $\mathbf{\Pi}^{(y)}(\tau,\tau_0)$, we can determine the complex-valued matrix $\mathbf{M}^{(y)}(\tau)$ and the complex-valued factor $[\det(\mathbf{Q}^{(y)}(\tau)/\det \mathbf{Q}^{(y)}(\tau_0))]^{1/2}$ by solving the DRT system with specific complex-valued initial conditions. These conditions can be given by relations analogous to (40):

$$\mathbf{Q}^{(y)}(\tau_0) = \mathbf{I} \; , \qquad \mathbf{P}^{(y)}(\tau_0) = \mathbf{M}_0^{(y)} \; . \tag{48}$$

The matrix $\mathbf{M}_0^{(y)} = \mathbf{M}^{(y)}(\tau_0)$ is now complex-valued. The 2×2 matrices $\mathrm{Re}\mathbf{M}_0^{(y)}$ and $\mathrm{Im}\mathbf{M}_0^{(y)}$ are real valued, symmetric and finite. Moreover, $\mathrm{Im}\mathbf{M}_0^{(y)}$ is positive definite. Quite often $\mathrm{Re}\mathbf{M}_0^{(y)} = 0$ is used, representing Gaussian beam with plane wavefront at $\tau = \tau_0$. The Gaussian beam is again given by (45), where the 2×2 complex-valued matrix $\mathbf{M}^{(y)}(\tau)$ is given by the relation $\mathbf{M}^{(y)}(\tau) = \mathbf{P}^{(y)}(\tau)(\mathbf{Q}^{(y)}(\tau))^{-1}$. The 2×2 matrices

$\mathbf{P}^{(y)}(\tau)$ and $\mathbf{Q}^{(y)}(\tau)$ are also complex valued, and are obtained as solutions of the DRT system with the initial conditions (48). Finally, the complex-valued 2×2 matrix $\mathbf{W}(\tau,\tau_0)$ is given by the relation $\mathbf{W}(\tau,\tau_0) = \mathbf{Q}^{(y)}(\tau)(\mathbf{Q}^{(y)}(\tau_0))^{-1}$.

For more details, see Appendix B.

4.2 Evaluation of a Gaussian beam at a specified paraxial point

Equation (45) with (46) can be simply used to calculate a Gaussian beam connected with ray Ω at any point R, situated in a paraxial vicinity of Ω in the plane tangent to the wavefront and intersecting Ω at a **known** point R_Ω. The position of point R_Ω on Ω is specified by the monotonic parameter τ and by $y_1 = y_2 = 0$. The wavefront orthonormal coordinates y_1, y_2 of point R are then determined in the plane tangent to the wavefront at R_Ω. This plane is specified by the known basis vectors $\mathbf{e}_1(\tau)$, $\mathbf{e}_2(\tau)$. If we wish, we can then determine the Cartesian coordinates of point R from the known Cartesian coordinates of R_Ω and from the known Cartesian components of $\mathbf{e}_1(\tau)$ and $\mathbf{e}_2(\tau)$.

A problem arises if point R is specified in Cartesian coordinates. This is the case, for example, of the method of summation of Gaussian beams, in which the observation point R is usually specified in Cartesian coordinates. We then face a considerably more complicated problem to determine point R_Ω on ray Ω, at which the plane tangent to the wavefront and containing the observation point R intersects Ω. Once point R_Ω is determined, the evaluation of Gaussian beam at R is easy.

Determination of R_Ω from known R in an inhomogeneous anisotropic medium is very cumbersome. It is cumbersome even in an isotropic inhomogeneous medium, when the plane tangent to the wavefront at R_Ω reduces to the plane perpendicular to Ω at R_Ω. Consequently, the solution of the problem in inhomogeneous isotropic media requires the numerical determination of the plane perpendicular to Ω, containing point R. For an inhomogeneous anisotropic medium, however, the problem is not purely geometrical, since the plane perpendicular to Ω must be replaced by the plane tangent to the wavefront at R_Ω.

Described cumbersome procedure could be avoided if the expression (45) for the Gaussian beam, and specifically the argument of the exponential function, were transformed to Cartesian coordinates. Actually, it is sufficient to transform the 2×2 matrix of second derivatives of the travel time with respect to wavefront orthonormal coordinates to the 3×3 matrix of second derivatives of the travel time with respect to Cartesian coordinates. For isotropic media, such a procedure was proposed by Klimeš (1984), see also Červený (2001; Sec.4.1.8) and Introduction of this paper. The procedure is simple and efficient, and fully removes the above-mentioned problems. Then, with the known Cartesian coordinates of point R_Ω, chosen arbitrarily on Ω, but as close as possible to arbitrarily situated point R, it is easy to evaluate the Gaussian beam at R. The same approach can be used in inhomogeneous anisotropic media. It is just sufficient to replace the terms containing the matrix $\mathbf{M}^{(y)}$ in (45) by $\mathbf{M}^{(x)}$, specified in Cartesian coordinates.

The transformation from $M^{(y)}_{IJ} = \partial^2 T/\partial q_I \partial q_J$, computed by DRT in ray-centred coordinates, to $M^{(x)}_{ij} = \partial^2 T/\partial x_i \partial x_j$, expressed in Cartesian coordinates, can be performed with formula (38) given in Sec.2.6.

Inserting (38) into (39), we obtain the quadratic expansion of the complex-valued paraxial travel-time field in Cartesian coordinates at an arbitrary point R, situated in the vicinity of point R_Ω, chosen arbitrarily on ray Ω:

$$\begin{aligned} T(R) &= T(R_\Omega) + (\mathbf{x}(R) - \mathbf{x}(R_\Omega))^T \mathbf{p}(R_\Omega) \\ &+ \tfrac{1}{2}(\mathbf{x}(R) - \mathbf{x}(R_\Omega))^T \mathbf{M}^{(x)}(R_\Omega)(\mathbf{x}(R) - \mathbf{x}(R_\Omega)) \; . \end{aligned} \tag{49}$$

The elements of the 3×3 matrix $\mathbf{M}^{(x)}(R_\Omega)$ are given by (38). The only complex-valued expression in (49) is the 2×2 matrix $\mathbf{M}^{(y)}(R_\Omega)$, which is included in $\hat{\mathbf{M}}^{(x)}(R_\Omega)$, see (38).

The quadratic expansion of the travel time can be used not only with respect to paraxial distances from ray Ω, but also with respect to the distances along ray Ω; it can be used in a "quadratic" vicinity of an arbitrarily chosen point R_Ω on the ray Ω. If the termination points of different rays are situated along a target surface, the contributions of individual Gaussian beams at the point R can be computed from values at these termination points. Point R is no longer required to be situated in planes tangent to the wavefronts at points R_Ω on each central ray Ω. This increases considerably the efficiency of procedures like the summation of Gaussian beams.

We now consider a Gaussian beam, connected with ray Ω and determine its contribution at the paraxial point R specified in Cartesian coordinates and situated in the vicinity of point R_Ω on Ω. Point R_Ω corresponds to the sampling parameter $\gamma_3 = \tau$ along Ω. The coordinates of point R_Ω can, of course, be expressed also in Cartesian coordinates. To some extent, the point R_Ω may be chosen arbitrarily on ray Ω, but it must be close to R. Then the contribution at point R of the Gaussian beam, concentrated to Ω, is as follows:

$$\begin{aligned} \mathbf{u}^{beam}(R) &= \mathbf{U}^\Omega(R_\Omega)(\det \mathbf{W})^{-1/2} \exp[-\omega \mathrm{Im}T(R)] \\ &\times \exp[-\mathrm{i}\omega(t - \mathrm{Re}T(R))] \; . \end{aligned} \tag{50}$$

Here $\mathbf{U}^\Omega$ and $\det \mathbf{W}$ have the same meaning as in equations (45) and (46). The complex-valued paraxial travel time $T(R)$ is given by (49) with (38).

For completeness, we present here the expressions for $\mathrm{Re}T(R)$ and $\mathrm{Im}T(R)$ in Cartesian coordinates:

$$\mathrm{Re}T(R) = T(R_\Omega) + \overline{p} + \frac{1}{2}\overline{\mathbf{f}}\, \mathrm{Re}\mathbf{M}^{(y)}(R_\Omega)\, \overline{\mathbf{f}}^T + \overline{p}\,\overline{\eta} + \frac{1}{2}(\boldsymbol{\mathcal{U}}^T \boldsymbol{\eta})\overline{p}^2 \; , \tag{51}$$

$$\mathrm{Im}T(R) = \frac{1}{2}\overline{\mathbf{f}}\, \mathrm{Im}\mathbf{M}^{(y)}(R_\Omega)\, \overline{\mathbf{f}}^T \; . \tag{52}$$

Here

$$\begin{aligned} \overline{p} &= (\mathbf{x}(R) - \mathbf{x}(R_\Omega))^T \mathbf{p}(R_\Omega) \; , \\ \overline{\eta} &= (\mathbf{x}(R) - \mathbf{x}(R_\Omega))^T \boldsymbol{\eta}(R_\Omega) \; , \\ \overline{\mathbf{f}} &= (\overline{f}_1, \overline{f}_2) = (\mathbf{x}(R) - \mathbf{x}(R_\Omega))^T (\mathbf{f}_1, \mathbf{f}_2) \; . \end{aligned} \tag{53}$$

Quantities $\overline{p}$ and $\overline{\eta}$ represent scalar products of vectors $\mathbf{p}(R_\Omega)$ and $\boldsymbol{\eta}(R_\Omega)$ with the "observation vector" $\mathbf{x}(R) - \mathbf{x}(R_\Omega)$. The 1×2 matrix $\overline{\mathbf{f}}$ with elements $\overline{f}_1$ and $\overline{f}_2$ contains the scalar products of vectors $\mathbf{f}_1$ and $\mathbf{f}_2$ with the observation vector $\mathbf{x}(R) - \mathbf{x}(R_\Omega)$. The 2×2 matrix $\mathbf{M}^{(y)}(R_\Omega)$, with elements $M_{IJ}^{(y)}(R_\Omega)$, is obtained by dynamic ray tracing in wavefront orthonormal coordinates.

In the described approach, we calculate the paraxial travel-time field (49) and the Gaussian beam (50) in Cartesian coordinates, although the dynamic ray tracing has been performed in ray-centred coordinates. We apply (38) only locally at point R_Ω on ray Ω, close to the observation point R.

4.3 Properties of Gaussian beams

The expression for the Gaussian beam, given by (50) with (51)–(53) is very general and flexible, as the position of the observation point R may be specified in Cartesian coordinates. Moreover, we can use any point R_Ω situated on ray Ω, close to R, as the reference point in (50). Similarly as in equation (45), we need to determine the 2×2 matrix $\mathbf{M}^{(y)}(R_\Omega)$ of the second derivatives of the complex-valued travel-time field with respect to wavefront orthonormal coordinates y_1, y_2 at point R_Ω. The 2×2 matrix $\mathbf{M}^{(y)}(R_\Omega)$ may be determined by solving the DRT system in wavefront orthonormal coordinates.

In fact, equations (45) with (46) are a special case of (50)–(53) and can be simply obtained from them. Consider the observation point R situated in the plane tangent to the wavefront at R_Ω. We can then express the position of point R in ray-centred coordinates using the relation

$$\mathbf{x}(R) - \mathbf{x}(R_\Omega) = y_1\mathbf{e}_1 + y_2\mathbf{e}_2 \; . \tag{54}$$

Using (53) for $\overline{\mathbf{f}}$, we obtain

$$\overline{\mathbf{f}} = (y_1\mathbf{e}_1 + y_2\mathbf{e}_2)^T(\mathbf{f}_1, \mathbf{f}_2) = (y_1, y_2) = \mathbf{y}^T \; . \tag{55}$$

From (55), we obtain the Gaussian beam exponential factor, which controls the amplitude profile of the beam, see (50) and (52):

$$\exp[-\tfrac{1}{2}\omega\overline{\mathbf{f}}\mathrm{Im}\mathbf{M}^{(y)}(R_\Omega)\overline{\mathbf{f}}^T] = \exp[-\tfrac{1}{2}\omega\mathbf{y}^T\mathrm{Im}\mathbf{M}^{(y)}(R_\Omega)\mathbf{y}] \; . \tag{56}$$

Now consider the equation (51) for $\mathrm{Re}T(R)$ in the plane tangent to the wavefront at R_Ω. As slowness vector $\mathbf{p}$ is perpendicular to this plane at R_Ω, we obtain $\overline{p} = 0$. Consequently, the three terms in (51) containing $\overline{p}$ vanish. Using eq.(55), term $\frac{1}{2}\overline{\mathbf{f}}\mathrm{Re}\mathbf{M}^{(y)}(R_\Omega)\overline{\mathbf{f}}^T$ can be transformed into $\frac{1}{2}\mathbf{y}^T\mathrm{Re}\mathbf{M}^{(y)}(R_\Omega)\mathbf{y}$. Thus, equations (45)–(46) for the Gaussian beam in the plane tangent to the wavefront at R_Ω, obtained in Section 4.1.1, follow, as a special case, from equations (50)–(53).

It is obvious that equations (50)–(53) for $\mathrm{Re}T(R)$ and $\mathrm{Im}T(R)$ in Cartesian coordinates are more general and flexible than equations (45)–(46) in ray-centred coordinates.

Similarly, the algorithm for the summation of Gaussian beams at an observation point R specified in Cartesian coordinates is considerably simpler if equations (50)–(53) are used. Equations (50)–(53) may also be conveniently used in studying properties of Gaussian beams in any section and along any profile. It is easy to show that the amplitude distribution of Gaussian beams is Gaussian along any straight-line profile intersecting the ray. The most convenient formulae for Gaussian beams are, however, obtained in the plane tangent to the wavefront at any point R_Ω of ray Ω. In this plane, the Gaussian beam is fully described by equations (45)–(46). As the wavefront orthonormal coordinates y_1, y_2 in this plane are actually 2D Cartesian coordinates, we can simply use $\exp[-\frac{1}{2}\omega\mathbf{y}^T\mathrm{Im}\mathbf{M}^{(y)}(R_\Omega)\mathbf{y}]$ from (45) to study the properties of Gaussian beams in this plane.

The amplitudes of a Gaussian beam in the plane tangent to the wavefront decrease exponentially with square of the distance from point R_Ω on ray Ω. The exponential decrease is frequency-dependent; it is faster for higher frequencies and slower for lower frequencies. Quadratic curve $\frac{1}{2}\omega\mathbf{y}^T(R)\mathrm{Im}\mathbf{M}^{(y)}(R_\Omega)\mathbf{y}(R) = 1$ in the plane tangent to the wavefront at R_Ω represents the spot ellipse for frequency ω. Along the spot ellipse, the amplitude of the Gaussian beam is constant.

The amplitude distribution of the Gaussian beam in the plane tangent to the wavefront at R_Ω is fully controlled by the 2×2 matrix $\mathrm{Im}\mathbf{M}^{(y)}(R_\Omega)$. As $\mathrm{Im}\mathbf{M}^{(y)}(R_\Omega)$ is symmetric and positive definite at any point R_Ω of ray Ω, it has two positive real-valued eigenvalues $M_1^I(R_\Omega)$ and $M_2^I(R_\Omega)$ at any point R_Ω. Instead of the eigenvalues $M_1^I(R_\Omega)$ and $M_2^I(R_\Omega)$ of the 2×2 matrix $\mathrm{Im}\,\mathbf{M}^{(y)}(R_\Omega)$, we can also use the quantities $L_1(R_\Omega)$ and $L_2(R_\Omega)$, given by the relation

$$L_{1,2}(R_\Omega) = [\pi M_{1,2}^I(R_\Omega)]^{-1/2} \; . \tag{57}$$

Quantities $L_1(R_\Omega)$ and $L_2(R_\Omega)$ represent the half-axes of the spot ellipse in the plane tangent to the wavefront at R_Ω for frequency $f = 1$Hz (i.e., $\omega = 2\pi$). We call them the half-widths of the Gaussian beam in the plane tangent to the wavefront at R_Ω. Let us note that spot ellipses can be simply constructed and studied in any plane intersecting ray Ω, including the plane perpendicular to Ω at R_Ω. The half-width of the Gaussian beam varies along ray Ω. We can determine these variations from the equations (57) and (36). Consequently, the Gaussian beams may be narrow in some regions of the ray, but broad in other regions.

The 2×2 matrix $\mathrm{Re}\mathbf{M}^{(y)}(R_\Omega)$ describes the geometric properties of the wavefront of the Gaussian beam. Since $\mathrm{Re}\mathbf{M}^{(y)}(R_\Omega)$ is always symmetrical, its eigenvalues $M_1^R(R_\Omega)$ and $M_2^R(R_\Omega)$ are always real. Instead of $\mathrm{Re}\mathbf{M}^{(y)}(R_\Omega)$, we can introduce the 2×2 matrix $\mathbf{K}(R_\Omega)$ of the curvature of the wavefront at R_Ω on ray Ω by relation

$$\mathbf{K}(R_\Omega) = \mathcal{C}(R_\Omega)\mathrm{Re}\mathbf{M}^{(y)}(R_\Omega) \; , \tag{58}$$

where $\mathcal{C}(R_\Omega)$ is the phase velocity (the velocity in the direction of the slowness vector $\mathbf{p}(R_\Omega)$). The eigenvalues of $\mathbf{K}(R_\Omega)$ then represent the principal curvatures of the wavefront of the Gaussian beam at point R_Ω on Ω.

The curvature of the wavefront of the Gaussian beam varies along ray Ω and may be determined using (58) and (36). For $\mathrm{Re}\mathbf{M}^{(y)}(R_\Omega) = \mathbf{0}$, the wavefront is locally planar at R_Ω.

5 Discussion and concluding remarks

In Sections 3 and 4, we studied the paraxial ray approximation and the Gaussian beams, expressed in wavefront orthonormal coordinates y_1, y_2 in a great detail. When we wish to consider global Cartesian coordinates, we have to make some minor changes.

The basic quantity we shall use is the 2×2 matrix $\mathbf{M}^{(y)}$, which is real-valued for paraxial ray approximation, and complex-valued for Gaussian beams. We shall try to express the initial conditions for the DRT system in global Cartesian coordinates in terms of this matrix.

In global Cartesian coordinates x_i, the DRT system is the simplest from the systems shown in this paper. Moreover, it does not require to solve the ordinary differential equation (14) for $\mathbf{e}_1$ and $\mathbf{e}_2$ along the ray. The disadvantage of the system is a large number of equations.

The initial conditions for the DRT system in global Cartesian coordinates can be again expressed in terms of matrix $\mathbf{M}^{(y)}(\tau_0)$. We merely use (38) and determine $\mathbf{M}^{(x)}(\tau_0)$ from $\mathbf{M}^{(y)}(\tau_0)$. The quantities $\mathbf{p}(\tau_0)$, $\boldsymbol{\eta}(\tau_0)$ and $\boldsymbol{\mathcal{U}}(\tau_0)$ are known from ray tracing, and $\mathbf{f}_1(\tau_0)$, $\mathbf{f}_2(\tau_0)$ are determined from $\mathbf{e}_1(\tau_0)$, $\mathbf{e}_2(\tau_0)$. The unit vectors $\mathbf{e}_1(\tau_0)$ and $\mathbf{e}_2(\tau_0)$ must be tangential to the wavefront and mutually perpendicular, but otherwise they can be chosen freely.

By DRT in global Cartesian coordinates, we determine the whole matrix $\mathbf{M}^{(x)}(\tau)$ along the ray by dynamic ray tracing. Consequently, the application of equation (38) at points $\tau \neq \tau_0$ is not required. If we, however, wish to determine $\mathbf{M}^{(y)}(\tau)$, we can use (38) and solve the differential equation (14) for $\mathbf{e}_I$, and to determine $\mathbf{f}_I$ from $\mathbf{e}_I$.

The advantage of the *simplified DRT in global Cartesian coordinates* is the application of the simplest DRT system. The next advantage is the lower number of equations in the system. The disadvantage with respect to complete DRT system in global Cartesian coordinates is that the ordinary differential equation for $\mathbf{e}_I$ must be solved to compute $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{f}_1$ and $\mathbf{f}_2$ along the ray.

The determination of initial conditions in the simplified DRT system in terms of $\mathbf{M}^{(y)}(\tau_0)$ is explained in Section 2.3.3, see (24). At any point of the ray, we can simply compute $Q^{(y)}_{IN}$ and $P^{(y)}_{IN}$ from $Q^{(x)}_{iN}$ and $P^{(x)}_{iN}$ using (25). Consequently, all derivations presented in Section 3 may be also applied to simplified DRT in global Cartesian coordinates.

The expressions derived for the Gaussian beams in inhomogeneous anisotropic layered media are valid for anisotropy of arbitrary symmetry, specified by upto 21 density-normalized elastic moduli a_{ijkl}, which are arbitrary functions of coordinates x_n. Of course, it is assumed that the general conditions of applicability of the ray method are satisfied to compute sufficiently accurate ray propagator matrices along the ray.

The basic quantity in the computation of Gaussian beams in inhomogeneous anisotropic layered media is the complex-valued 2×2 matrix $\mathbf{M}^{(y)}(\tau)$ of the second derivatives of the travel-time field with respect to wavefront orthonormal coordinates, with elements $\partial^2 T/\partial y_N \partial y_M$. Once the 2×2 matrix $\mathbf{M}^{(y)}(\tau)$ is known along the central ray Ω, the flexible and numerically efficient expressions (50)–(53) for the Gaussian beam can be used.

As the model of an inhomogeneous anisotropic layered medium under consideration is very general, the relevant procedures are rather complex. They may, however, be simplified in many special cases. Such simplifications lead to more efficient algorithms. The most significant role is played by the simplifications of the ray-tracing computations.

In the following, we list several such possible simplifications.

a) Considerable simplifications can be obtained for higher anisotropic symmetries, for example for the transversely isotropic (TI) or orthorhombic media.

b) Further simplification can be obtained for weakly anisotropic media. This is true particularly for P waves. For example, approach proposed by Zhu, Gray and Wang (2007) can be used in this case. For S waves, their coupling in weakly anisotropic media must be taken into account. In this case, Gaussian beams should be constructed along so-called common rays. It may be useful to use first-order ray tracing (Pšenčík and Farra, 2005; Farra and Pšenčík, 2010).

c) For media with spatially varying elements of higher anisotropic symmetry, the ray-tracing computations are quite cumbersome and even inaccurate. Iversen and Pšenčík (2007) proposed a procedure based on the evaluation of quantities important for ray tracing in a coordinate system connected with the symmetry elements of the considered medium. The procedure conserves the symmetry of the studied medium throughout the model and increases the efficiency (computer-time and memory savings) considerably.

d) Considerably more efficient algorithms for ray tracing and ray propagator matrix computations are obtained for inhomogeneous factorized anisotropic media, in which all density normalized elastic moduli vary spatially in the same way. The simplest inhomogeneous factorized anisotropic media are media, in which the gradient (e.g., the vertical gradient) of all density-normalized elastic moduli is the same in the whole region of interest. For more details, refer to Červený (1989), Shearer and Chapman (1989).

e) A simple and numerically very efficient algorithm would be obtained for a medium composed of homogeneous anisotropic layers (blocks), or composed of different inhomogeneous factorized anisotropic layers (blocks).

f) Instead of general 3-D configurations, considered in this text, we can sometimes consider 2-D configurations, which simplifies the procedures considerably. For inhomogeneous media of general anisotropy, however, the 2-D configurations do not play such an important role as for isotropic inhomogeneous media. This is due to the different direction of the ray-velocity vector, the slowness vector and the polarization vector, because of which the wave propagation in inhomogeneous anisotropic media is generally three-dimensional. Only in the planes of symmetry of transversely isotropic or orthorhombic media may a 2-D configuration play an important role.

Acknowledgements: The authors are greatly indebted to Luděk Klimeš for helpful suggestions and valuable discussions. This research has been supported by the Consortium Project "Seismic Waves in Complex 3-D Structures", by the Research Projects 205/08/0332 and P210/10/0736 of the Grant Agency of the Czech Republic, and by the Research Project MSM 0021620860 of the Ministry of Education of the Czech Republic.

Appendix A
Ray-tracing in inhomogeneous anisotropic media with interfaces

We consider the eikonal equation for inhomogeneous anisotropic media, given in the Hamiltonian form $\mathcal{H}(x_i, p_j) = 0$. Hamiltonian $\mathcal{H}(x_i, p_j)$ is defined by the relation

$$\mathcal{H}(x_i, p_j) = \tfrac{1}{2} G_m(x_i, p_j) \ , \qquad (A-1)$$

where G_m is one of the three eigenvalues of the Christoffel matrix $\Gamma_{ik} = a_{ijkl} p_j p_l$. We assume that the eigenvalue G_m differs from other two eigenvalues. The ray tracing system is given by the system of six nonlinear ordinary differential equations of the first order,

$$\frac{\mathrm{d}x_i}{\mathrm{d}\tau} = \frac{\partial \mathcal{H}}{\partial p_i} \ , \quad \frac{\mathrm{d}p_i}{\mathrm{d}\tau} = -\frac{\partial \mathcal{H}}{\partial x_i} \ . \qquad (A-2)$$

Here the parameter τ along the ray represents the travel time. The right-hand sides of the ray tracing system (A-2) represent components $\mathcal{U}_i = \partial \mathcal{H}/\partial p_i$ of the ray velocity vector $\boldsymbol{\mathcal{U}}$, and $\eta_i = -\partial \mathcal{H}/\partial x_i$ of vector $\boldsymbol{\eta}$. They can be determined from the equations

$$\frac{\partial \mathcal{H}}{\partial p_i} = \mathcal{U}_i = a_{ijkl} p_l g_j^{(m)} g_k^{(m)} \ , \quad -\frac{\partial \mathcal{H}}{\partial x_i} = \eta_i = -\frac{1}{2} \frac{\partial a_{jkln}}{\partial x_i} p_k p_n g_j^{(m)} g_l^{(m)} \ . \qquad (A-3)$$

(no summation over m). Vector $\mathbf{g}^{(m)}$ denotes the unit eigenvector of the Christoffel matrix corresponding to the eigenvalue G_m of the wave under consideration (P, S1, S2).

In equations (A-2), we can use alternative expressions for $\partial \mathcal{H}/\partial p_i$ and $-\partial \mathcal{H}/\partial x_i$, which do not contain eigenvector $\mathbf{g}^{(m)}$ explicitly. These expressions are, however, algebraically more complicated. They read

$$\frac{\partial \mathcal{H}}{\partial p_i} = \mathcal{U}_i = a_{ijkl} p_l D_{jk}/D_{ss} \ , \quad -\frac{\partial \mathcal{H}}{\partial x_i} = \eta_i = -\frac{1}{2} \frac{\partial a_{jkln}}{\partial x_i} p_k p_n D_{jl}/D_{ss} \ . \qquad (A-4)$$

Here

$$D_{ij} = \tfrac{1}{2} \epsilon_{ikl} \epsilon_{jrs} (\Gamma_{kr} - \delta_{kr})(\Gamma_{ls} - \delta_{ls}) \ . \qquad (A-5)$$

Symbol ϵ_{ijk} represents the Levi-Civita symbol ($\epsilon_{123} = \epsilon_{312} = \epsilon_{231} = 1$, $\epsilon_{132} = \epsilon_{213} = \epsilon_{321} = -1$, $\epsilon_{ijk} = 0$ otherwise).

Another alternative expression for $\partial \mathcal{H}/\partial x_i$ in ray tracing equations (A-2) is as follows. Taking into account that $G_m(x_i, p_j)$ in (A-1) is a homogeneous function of the second degree in p_j, we have $G_m(x_i, p_j) = p_k p_k G_m(x_i, N_j)$ and $G_m(x_i, N_j) = \mathcal{C}^2(x_i, N_j)$, where $\mathbf{N} = \mathbf{p}/|\mathbf{p}|$ and $p_k p_k = \mathcal{C}^{-2}(x_i, N_j)$, $\mathcal{C}(x_i, N_j)$ being the phase velocity in the direction of the vector $\mathbf{N}$. Then we can write

$$-\frac{\partial \mathcal{H}(x_n, p_j)}{\partial x_i} = -\frac{1}{2} \frac{\partial G_m(x_n, p_j)}{\partial x_i} = -\frac{1}{2} \mathcal{C}^{-2}(x_n, N_j) \frac{\partial \mathcal{C}^2(x_n, N_j)}{\partial x_i} \ . \qquad (A-6)$$

For $\partial\mathcal{H}/\partial p_i$ and $-\partial\mathcal{H}/\partial x_i$ we can thus write:

$$\frac{\partial\mathcal{H}}{\partial p_i} = \mathcal{U}_i(x_n, p_j) \ , \quad -\frac{\partial\mathcal{H}}{\partial x_i} = -\mathcal{C}^{-1}(x_n, N_j)\frac{\partial\mathcal{C}(x_n, N_j)}{\partial x_i} \ . \qquad (A-7)$$

The ray tracing (A-2) with (A-7) was proposed by Zhu, Gray and Wang (2005). The phase velocity $\mathcal{C}$ and the ray-velocity vector $\boldsymbol{\mathcal{U}}$ can be expressed exactly or approximately in terms of Thomsen's (1986) or weak-anisotropy (WA) parameters. In the case of approximate expressions and WA parameters, the ray tracing system (A-2) with (A-7) corresponds to the system proposed by Pšenčík and Farra (2005) for P-wave ray tracing.

The initial conditions for the ray tracing system (A-2) differ from the initial conditions for ray tracing in isotropic media. For a given direction $\mathbf{N}$ specifying the direction of the slowness vector at the initial point of the ray, we can determine phase velocities of all three waves, $\mathcal{C}_m$. The phase velocities are obtained as a solution of a cubic equation resulting from the condition of solvability of the Christoffel equation

$$\det(a_{ijkl}N_jN_l - \mathcal{C}\delta_{ik}) = 0 \ . \qquad (A-8)$$

Selecting the value of $\mathcal{C}_m$ corresponding to the studied wave, we can construct the slowness vector of the studied wave at the initial point of the ray, $\mathbf{p} = \mathbf{N}/\mathcal{C}_m$. The initial conditions for the ray-velocity vector $\boldsymbol{\mathcal{U}}$ need not be specified among initial conditions for (A-3).

Anisotropic ray tracing system (A-2), with (A-3) or (A-4) or (A-7), can be used quite universally for P waves, including P waves in heterogeneous isotropic and weakly anisotropic media. For S waves, however, the situation is more complicated. Ray tracing fails in the vicinity of S-wave singularities, where the two eigenvalues corresponding to S waves are equal or close to each other. Ray tracing for S waves may fail globally in very weakly anisotropic media, and fails fully in isotropic media. Remember that two eigenvalues of S waves are equal in isotropic media. These problems can be removed if the coupling ray theory for shear waves or its various modifications are used (Kravtsov, 1968; Coates and Chapman, 1990; Bulant and Klimeš, 2008; Farra and Pšenčík, 2010).

Let us now assume that a ray hits a curved structural interface. In the framework of the zero-order ray method, the reflection/transmission problem at the point of incidence of an arbitrary high-frequency wave at a curved interface Σ separating two inhomogeneous media is reduced to the problem of incidence of a plane wave at a plane interface separating two homogeneous media. Three reflected and three transmitted waves (P, S1, S2) are generated at the point of incidence; some of them may be inhomogeneous. The slowness vector of any of reflected or transmitted waves at the point of reflection/transmission is given by the relation

$$\tilde{\mathbf{p}} = \sigma\mathbf{n} + \mathbf{p}^{\Sigma} \ . \qquad (A-9)$$

Tilde indicates that the quantity corresponds to a generated wave. In (A-9), $\mathbf{n}$ is the unit normal to the interface Σ at the point of incidence. Symbol $\mathbf{p}^{\Sigma}$ denotes the tangential component to the interface of the slowness vector of the incident wave. Component $\mathbf{p}^{\Sigma}$ is the same for the incident and all generated waves. This equality is just another expression of the Snell law. The projection σ of slowness vector $\tilde{\mathbf{p}}$ to normal $\mathbf{n}$ is a root of the algebraic equation of the sixth degree:

$$\det[a_{ijkl}(\sigma n_k + p_k^{\Sigma})(\sigma n_l + p_l^{\Sigma}) - \delta_{ij}] = 0 \ . \qquad (A-10)$$

For reflected waves, we use the same elastic moduli a_{ijkl} as for incident waves. For transmitted waves, we use a_{ijkl} corresponding to the halfspace on the other side of the interface.

Equation (A-10) has six solutions for each halfspace. The physical solutions corresponding to the three reflected and three transmitted waves are selected from them according to the direction of the relevant ray-velocity vector $\tilde{\boldsymbol{\mathcal{U}}}$ (for real-valued roots) and according to the radiation conditions (for complex-valued roots). For more details, see Gajewski and Pšenčík (1987), Červený (2001, section 2.3.3).

The ray method was first proposed for the computation of high-frequency seismic wave fields in inhomogeneous anisotropic media by Babich (1961). Ray-tracing equations, namely (A-2) with (A-4), were first derived by Červený (1972). For more details on ray tracing in inhomogeneous anisotropic media, see Červený (2001, Chap.3.6). The computer program for ray tracing in inhomogeneous anisotropic layered structures based on the above-mentioned formulae is the computer package ANRAY (Gajewski and Pšenčík, 1987,1990). The package ANRAY is freely available on the web pages of the SW3D Consortium (http://sw3d.mff.cuni.cz/).

Appendix B
Dynamic ray tracing in inhomogeneous anisotropic media

Dynamic ray tracing consists in the solution of a system of linear ordinary differential equations of the first order along the ray Ω. The dynamic ray tracing system in various forms of Cartesian coordinates (general, local) was studied by several authors. For a detailed derivations and for references see Červený (2001). We shall discuss here three DRT systems, based on Cartesian coordinates: a) The DRT system in global Cartesian coordinates x_1, x_2, x_3. b) The DRT system in the wavefront orthonormal coordinates (local Cartesian coordinates) y_1, y_2. c) Simplified DRT system in global Cartesian coordinates x_1, x_2, x_3.

a) **DRT system in global Cartesian coordinates** x_i.

It consists of six equations for $Q_i^{(x)} = \partial x_i/\partial\gamma$ and $P_i^{(x)} = \partial p_i/\partial\gamma$, where γ is any of initial values x_{i0}, p_{i0}, or any ray parameter or a variable along the ray. The DRT system reads

$$\mathrm{d}Q_i^{(x)}/\mathrm{d}\tau = A_{ij}^{(x)}Q_j^{(x)} + B_{ij}^{(x)}P_j^{(x)} , \quad \mathrm{d}P_i^{(x)}/\mathrm{d}\tau = -C_{ij}^{(x)}Q_j^{(x)} - D_{ij}^{(x)}P_j^{(x)} , \qquad (B-1)$$

where

$$A_{ij}^{(x)} = \partial^2\mathcal{H}/\partial p_i\partial x_j , \quad B_{ij}^{(x)} = \partial^2\mathcal{H}/\partial p_i\partial p_j ,$$
$$C_{ij}^{(x)} = \partial^2\mathcal{H}/\partial x_i\partial x_j , \quad D_{ij}^{(x)} = \partial^2\mathcal{H}/\partial x_i\partial p_j , \qquad (B-2)$$

and where the Hamiltonian $\mathcal{H}$ is given by (7). The elements of 3×3 matrices $A_{ij}^{(x)}$, $B_{ij}^{(x)}$, $C_{ij}^{(x)}$ and $D_{ij}^{(x)}$ satisfy three symmetry relations:

$$B_{ij}^{(x)} = B_{ji}^{(x)} , \quad C_{ij}^{(x)} = C_{ji}^{(x)} , \quad D_{ij}^{(x)} = A_{ji}^{(x)} . \qquad (B-3)$$

The 1×3 vectors $\mathbf{Q}^{(x)}$ and $\mathbf{P}^{(x)}$ in (B-1) satisfy the following constrain relation:

$$\mathcal{U}_k Q_k^{(x)} - \eta_k P_k^{(x)} = 0 \ . \qquad (B-4)$$

It is common to write the dynamic ray tracing system (B-1) in a matrix form:

$$\mathrm{d}\mathbf{Q}^{(x)}/\mathrm{d}\tau = \mathbf{A}^{(x)}\mathbf{Q}^{(x)} + \mathbf{B}^{(x)}\mathbf{P}^{(x)} \ , \quad \mathrm{d}\mathbf{P}^{(x)}/\mathrm{d}\tau = -\mathbf{C}^{(x)}\mathbf{Q}^{(x)} - \mathbf{D}^{(x)}\mathbf{P}^{(x)} \ , \qquad (B-5)$$

where elements of the 3×3 matrices $\mathbf{Q}^{(x)}$ and $\mathbf{P}^{(x)}$ have the following form:

$$Q_{ij}^{(x)} = \partial x_i/\partial \gamma_j \ , \qquad P_{ij}^{(x)} = \partial p_i/\partial \gamma_j \ . \qquad (B-6)$$

Here $\gamma_i = x_{i0}$ or $\gamma_i = p_{i0}$ or γ_i are ray coordinates.

Consequently, the 6×6 system matrix $\mathbf{S}^{(x)}(\tau)$ of the DRT system in global Cartesian coordinates reads:

$$\mathbf{S}^{(x)}(\tau) = \begin{pmatrix} \mathbf{A}^{(x)} & \mathbf{B}^{(x)} \\ -\mathbf{C}^{(x)} & -\mathbf{D}^{(x)} \end{pmatrix} \ . \qquad (B-7)$$

Let us now multiply the DRT equations (B-5) from the right-hand sides by a constant finite non-degenerate matrix $\mathbf{A}^{-1}$. Then matrices

$$\overline{\mathbf{Q}}^{(x)} = \mathbf{Q}^{(x)}\mathbf{A}^{-1} \ , \quad \overline{\mathbf{P}}^{(x)} = \mathbf{P}^{(x)}\mathbf{A}^{-1} \qquad (B-8)$$

are also solutions of (B-5). This fact allows us to specify simply the initial conditions for (B-5). Taking

$$\mathbf{A} = \mathbf{Q}^{(x)}(\tau_0) \ , \qquad (B-9)$$

we obtain initial conditions for $\overline{\mathbf{Q}}^{(x)}$ and $\overline{\mathbf{P}}^{(x)}$ as follows

$$\overline{\mathbf{Q}}^{(x)}(\tau_0) = \mathbf{I} \ , \quad \overline{\mathbf{P}}^{(x)}(\tau_0) = \mathbf{P}^{(x)}(\tau_0)(\mathbf{Q}^{(x)}(\tau_0))^{-1} = \mathbf{M}^{(x)}(\tau_0) \ , \qquad (B-10)$$

see (13). Initial conditions (B-10) are considerably more suitable than initial conditions for $\mathbf{Q}^{(x)}$ and $\mathbf{P}^{(x)}$. They reduce to the specification of only the identity matrix and of the matrix of second derivatives of the travel-time field with respect to Cartesian coordinates x_i, $\partial^2 T/\partial x_i \partial x_j$. This matrix has a simple and clear physical meaning. For initial conditions (B-10), the solutions of the DRT system has the following form:

$$\overline{\mathbf{Q}}^{(x)}(\tau) = \mathbf{Q}^{(x)}(\tau)(\mathbf{Q}^{(x)}(\tau_0))^{-1} \ ,$$

$$\overline{\mathbf{P}}^{(x)}(\tau) = \mathbf{P}^{(x)}(\tau)(\mathbf{Q}^{(x)}(\tau_0))^{-1} = \mathbf{M}^{(x)}(\tau)\overline{\mathbf{Q}}^{(x)}(\tau) \ . \qquad (B-11)$$

Here we have used the relation $\mathbf{M}^{(x)}(\tau) = \mathbf{P}^{(x)}(\tau)(\mathbf{Q}^{(x)}(\tau))^{-1}$, see (13).

We can further modify initial conditions (B-10) using eq. (38) for the 3×3 matrix $\mathbf{M}^{(x)}(\tau)$. With the exception of the second derivatives of the travel time field in the plane tangential to the wavefront $\partial^2 T/\partial y_I \partial y_J$, all other quantities are known in (38). Thus, it is sufficient to consider the 2×2 matrix $\mathbf{M}^{(y)}(\tau_0)$ as an initial condition. This 2×2 symmetric matrix $\mathbf{M}^{(y)}$ is represented by only three elements, $M_{11}^{(y)}$, $M_{22}^{(y)}$ and $M_{12}^{(y)}$. This is well known from the theory of DRT in wavefront orthonormal coordinates, but we just showed that the same 2×2 initial matrix $\mathbf{M}^{(y)}$ may be used even when we consider global Cartesian coordinates.

Solving the dynamic ray tracing system (B-1) for $Q_i^{(x)}$ and $P_i^{(x)}$, with the initial conditions for $\mathbf{\Pi}^{(x)}(\tau, \tau_0)$ given by the 6×6 identity matrix $\mathbf{I}$, we can calculate the 6×6 ray propagator matrix $\mathbf{\Pi}^{(x)}(\tau, \tau_0)$ in global Cartesian coordinates. We use it to generalize dynamic ray tracing for media with curved structural interfaces. Let us consider a ray Ω of a wave reflected/transmitted at an interface Σ. We denote by τ^Σ the travel time corresponding to the point of incidence and $\tilde{\tau}^\Sigma$ to the point of reflection or transmission. Of course, $\tilde{\tau}^\Sigma = \tau^\Sigma$, but the slowness vectors $\mathbf{p}$, ray-velocity vectors $\boldsymbol{\mathcal{U}}$, etc., are different for τ^Σ and $\tilde{\tau}^\Sigma$. The ray propagator matrix $\mathbf{\Pi}^{(x)}(\tau, \tau_0)$ from the initial point τ_0 situated on the ray of the incident wave to the termination point situated on the ray of the reflected/transmitted wave is then given by the relation

$$\mathbf{\Pi}^{(x)}(\tau, \tau_0) = \mathbf{\Pi}^{(x)}(\tau, \tilde{\tau}^\Sigma)\mathbf{\Pi}^{(x)}(\tilde{\tau}^\Sigma, \tau^\Sigma)\mathbf{\Pi}^{(x)}(\tau^\Sigma, \tau_0) \qquad (B-12)$$

The 6×6 matrix $\mathbf{\Pi}^{(x)}(\tilde{\tau}^\Sigma, \tau^\Sigma)$ is called the interface propagator matrix. The ray propagator matrix in a layered medium is obtained by considering the interface propagator matrix at every point of incidence of the ray at a structural interface.

We consider a smooth structural interface Σ which separates two inhomogeneous anisotropic media with a smooth distribution of elastic moduli and density. We specify Σ by the relation $\mathbf{x} = \mathbf{g}(u_1, u_2)$, where u_1 and u_2 are Gaussian coordinates on the interface with origin at the point of incidence. As a special case of u_1, u_2, we can consider local Cartesian coordinates in a plane tangent to the interface Σ at the point of incidence τ^Σ of the ray Ω on Σ. The vectorial function $\mathbf{g}(u_1, u_2)$ has first- and second- order partial derivatives given by $\mathbf{g}_I = \partial \mathbf{g}/\partial u_I$ and $\mathbf{g}_{IJ} = \partial^2 \mathbf{g}/\partial u_I \partial u_j$. The vectors $\mathbf{g}_1$ and $\mathbf{g}_2$ are generally non-orthogonal and non-unit vectors tangent to the surface Σ and the vectors $\mathbf{g}_{IJ}$ are related to the curvature of Σ. In the vicinity of the point of incidence, interface Σ may be approximated, to the second order in u_1, u_2, by the relation

$$\mathbf{x}(u_K) = \mathbf{x}_0 + \mathbf{g}_I u_J + \tfrac{1}{2}\mathbf{g}_{IJ} u_I u_J \; . \qquad (B-13)$$

Unit normal to the interface $\mathbf{n}$ is defined as $\mathbf{n} = (\mathbf{g}_1 \times \mathbf{g}_2)/|\mathbf{g}_1 \times \mathbf{g}_2|$.

The 6×6 interface propagator matrix $\mathbf{\Pi}^{(x)}(\tilde{\tau}_\Sigma, \tau_\Sigma)$ in global Cartesian coordinates x_i in a particularly concise form has derived by Moser (2004). It reads

$$\mathbf{\Pi}^{(x)}(\tilde{\tau}^\Sigma, \tau^\Sigma) = \begin{pmatrix} \tilde{\mathbf{X}}\mathbf{X}^{-1} & \mathbf{0} \\ \tilde{\mathbf{X}}^{-1}\mathbf{R}\mathbf{X}^{-1} & \tilde{\mathbf{X}}^{-1}\mathbf{X}^T \end{pmatrix} , \qquad (B-14)$$

where $\mathbf{X}$, $\tilde{\mathbf{X}}$ and $\mathbf{R}$ are 3×3 matrices:

$$\mathbf{X} = (\mathbf{g}_1, \mathbf{g}_2, \boldsymbol{\mathcal{U}}) \; , \qquad \tilde{\mathbf{X}} = (\mathbf{g}_1, \mathbf{g}_2, \tilde{\boldsymbol{\mathcal{U}}}) \; , \qquad (B-15)$$

$$\mathbf{R} = \begin{pmatrix} -\mathbf{g}_{IJ}^T \Delta \mathbf{p} & \mathbf{g}_I^T \Delta \boldsymbol{\eta} \\ \mathbf{g}_J \Delta \boldsymbol{\eta} & \tilde{\boldsymbol{\eta}}^T \tilde{\boldsymbol{\mathcal{U}}} - \boldsymbol{\eta}^T \boldsymbol{\mathcal{U}} \end{pmatrix} . \qquad (B-16)$$

Quantities with Δ in (B-16) denote the differences between quantities related to generated and incident waves, $\Delta \mathbf{p} = \tilde{\mathbf{p}} - \mathbf{p}$, $\Delta \boldsymbol{\eta} = \tilde{\boldsymbol{\eta}} - \boldsymbol{\eta}$. All quantities in (B-14) can be evaluated as soon as the ray tracing across the interface has been computed. It was proved by Moser (2004) that the interface propagator matrix (B-14) is symplectic, that it preserves the eikonal constraint relation (B-4), and that $\det \Pi^{(x)}(\tilde{\tau}^{\Sigma}, \tau^{\Sigma}) = 1$ and satisfies chain rule (B-24).

b) **DRT system in wavefront orthonormal Cartesian coordinates** y_I

It consists of four equations for $Q^{(y)}_M = \partial y_M/\partial\gamma$ and $P^{(y)}_M = \partial p^{(y)}_M/\partial\gamma$, where γ is any ray parameter. The wavefront orthonormal coordinates y_1, y_2 are described in Section 2.3. The DRT system in wavefront orthonormal coordinates reads (Červený, 2001, Sec. 4.2.2).

$$\mathrm{d}Q_M/\mathrm{d}\tau = A^{(y)}_{MN}Q^{(y)}_N + B^{(y)}_{MN}P^{(y)}_N \ , \qquad \mathrm{d}P^{(y)}_M/\mathrm{d}\tau = -C^{(y)}_{MN}Q^{(y)}_N - D^{(y)}_{MN}P^{(y)}_N \ , \qquad (B-17)$$

where

$$\begin{aligned} A^{(y)}_{MN} &= H_{iM}H_{jN}[\partial^2\mathcal{H}/\partial p_i\partial x_j + \mathcal{U}_i\eta_j] \ , \\ B^{(y)}_{MN} &= H_{iM}H_{jN}[\partial^2\mathcal{H}/\partial p_i\partial p_j - \mathcal{U}_i\mathcal{U}_j] \ , \\ C^{(y)}_{MN} &= H_{iM}H_{jN}[\partial\mathcal{H}/\partial x_i\partial x_j - \eta_i\eta_j] \ , \\ D^{(y)}_{MN} &= H_{iM}H_{jN}[\partial^2\mathcal{H}/\partial x_i\partial p_j - \eta_i\mathcal{U}_j] \ . \end{aligned} \qquad (B-18)$$

Here H_{iN} are components of the transformation matrix from wavefront orthonormal coordinates y_N to global Cartesian coordinates x_i, see (17). They can be expressed in terms of basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$, $H_{iN} = e_{Ni}$, see (18). The expressions (B-18) satisfy the following symmetry relations

$$B^{(y)}_{MN} = B^{(y)}_{NM} \ , \quad C^{(y)}_{MN} = C^{(y)}_{NM} \ , \quad D^{(y)}_{MN} = A^{(y)}_{NM} \ . \qquad (B-19)$$

The components p_i, $\mathcal{U}_i$ and η_i are expressed in global Cartesian coordinates, and are known from ray tracing, $H_{iN} = e_{Ni}$ are computed using (14).

It is common to express the DRT system (B-17) in a matrix form,

$$\mathrm{d}\mathbf{Q}^{(y)}/\mathrm{d}\tau = \mathbf{A}^{(y)}\mathbf{Q}^{(y)} + \mathbf{B}^{(y)}\mathbf{P}^{(y)} \ , \quad \mathrm{d}\mathbf{P}^{(y)}/\mathrm{d}\tau = -\mathbf{C}^{(y)}\mathbf{Q}^{(y)} - \mathbf{D}^{(y)}\mathbf{P}^{(y)} \ , \qquad (B-20)$$

where $\mathbf{Q}^{(y)}$ and $\mathbf{P}^{(y)}$ are 2×2 matrices with elements,

$$Q^{(y)}_{IJ} = \partial y_I/\partial\gamma_J \ , \qquad P^{(y)}_{IJ} = \partial p^{(y)}_I/\partial\gamma_J \ . \qquad (B-21)$$

Here γ_J are ray parameters.

Consequently, the 4×4 system matrix $\mathbf{S}^{(y)}(\tau)$ in wavefront orthonormal coordinates reads:

$$\mathbf{S}^{(y)}(\tau) = \begin{pmatrix} \mathbf{A}^{(y)} & \mathbf{B}^{(y)} \\ -\mathbf{C}^{(y)} & -\mathbf{D}^{(y)} \end{pmatrix} \ . \qquad (B-22)$$

Similarly as for global Cartesian coordinates, we can choose initial conditions for (B-20) as follows:

$$\overline{\mathbf{Q}}^{(y)}(\tau_0) = \mathbf{I} \ , \qquad \overline{\mathbf{P}}^{(y)}(\tau_0) = \mathbf{M}^{(y)}(\tau_0) \ , \qquad (B-23)$$

where $\mathbf{I}$ is the 2×2 identity matrix. For initial conditions (B-23), the DRT system (B-19) has the solution:

$$\overline{\mathbf{Q}}^{(y)}(\tau) = \mathbf{Q}^{(y)}(\tau)(\mathbf{Q}^{(y)})^{-1}(\tau_0) \ ,$$

$$\overline{\mathbf{P}}^{(y)}(\tau) = \mathbf{P}^{(y)}(\tau)(\mathbf{Q}^{(y)})^{-1}(\tau_0) = \mathbf{M}^{(y)}(\tau)\overline{\mathbf{Q}}^{(y)}(\tau) \ . \qquad (B-24)$$

We now present the 4×4 interface propagator matrix $\mathbf{\Pi}^{(y)}(\tilde{\tau}^{\Sigma},\tau^{\Sigma})$ in wavefront orthonormal coordinates. The chain rule across a structural interface yields

$$\mathbf{\Pi}^{(y)}(\tau,\tau_0) = \mathbf{\Pi}^{(y)}(\tau,\tilde{\tau}^{\Sigma})\mathbf{\Pi}^{(y)}(\tilde{\tau}^{\Sigma},\tau^{\Sigma})\mathbf{\Pi}^{(y)}(\tau^{\Sigma},\tau_0) \ . \qquad (B-25)$$

Using the same notations related to the structural interface Σ as in global Cartesian coordinates, the 4×4 interface propagator matrix $\mathbf{\Pi}^{(y)}(\tilde{\tau}_{\Sigma},\tau_{\Sigma})$ is given by the relation

$$\mathbf{\Pi}^{(y)}(\tilde{\tau}^{\Sigma},\tau^{\Sigma}) = \begin{pmatrix} \tilde{\mathbf{K}}^T\mathbf{K}^{-T} & \mathbf{0} \\ \tilde{\mathbf{K}}^{-1}[\mathbf{E}-\tilde{\mathbf{E}}-(\sigma-\tilde{\sigma})\mathbf{D}]\mathbf{K}^{-T} & \tilde{\mathbf{K}}^{-1}\mathbf{K} \end{pmatrix} \ . \qquad (B-26)$$

For detailed derivation of (B-26), see Červený and Moser (2007, Sec.6). In (B-25), the symbols without tilde correspond to the point of incidence, $\tau=\tau^{\Sigma}$, the symbols with tilde to the point of reflection/transmission, $\tau=\tilde{\tau}^{\Sigma}$. We explain here only the symbols corresponding to the point of incidence τ^{Σ}, the explanation of symbols with tilde is analogous.

The 2×2 matrices $\mathbf{K}$ and $\mathbf{K}^{-1}$ are given by relations

$$\mathbf{K} = (\mathbf{g}_1,\mathbf{g}_2)^T(\mathbf{f}_1,\mathbf{f}_2), \qquad \mathbf{K}^{-1} = (\mathbf{e}_1,\mathbf{e}_2)^T(\mathbf{h}_1,\mathbf{h}_2) \qquad (B-27)$$

with

$$\mathbf{h}_1 = (\mathbf{g}_2\times\boldsymbol{\mathcal{U}})/\boldsymbol{\mathcal{U}}^T(\mathbf{g}_1\times\mathbf{g}_2), \qquad \mathbf{h}_2 = (\boldsymbol{\mathcal{U}}\times\mathbf{g}_1)/\boldsymbol{\mathcal{U}}^T(\mathbf{g}_1\times\mathbf{g}_2). \qquad (B-28)$$

Vectors $\mathbf{e}_I$ and $\mathbf{f}_J$ are defined in (14) and (15).

Elements of the 2×2 **inhomogeneity matrix E** read

$$E_{IJ} = (\mathbf{g}_I^T\mathbf{p})(\mathbf{e}_K^T\boldsymbol{\eta})(\mathbf{g}_J^T\mathbf{f}_K) + (\mathbf{g}_I^T\boldsymbol{\eta})(\mathbf{g}_J^T\mathbf{p}). \qquad (B-29)$$

The symbols $\mathbf{p}$ and $\boldsymbol{\eta}$ denote the slowness and eta vectors, respectively, which are determined from ray tracing at the point of incidence. In homogeneous media, $\mathbf{E}=\mathbf{0}$, since $\boldsymbol{\eta}=\mathbf{0}$. The elements of the 2×2 **curvature matrix D** are given by the relation

$$D_{IJ} = \mathbf{g}_{IJ}^T\mathbf{n}. \qquad (B-30)$$

For a plane interface, $\mathbf{D}=\mathbf{0}$. Finally, scalar σ in (B-26) is given by the relations $\sigma=\mathbf{n}^T\mathbf{p}$.

Note that $\mathbf{g}_I=\tilde{\mathbf{g}}_I$, but $\mathbf{e}_I\neq\tilde{\mathbf{e}}_I$, $\mathbf{f}_I\neq\tilde{\mathbf{f}}_I$ and $\mathbf{h}_I\neq\tilde{\mathbf{h}}_I$. The vectors $\tilde{\mathbf{e}}_I$ can be chosen as two arbitrary mutually perpendicular unit vectors in the plane perpendicular to $\tilde{\mathbf{p}}$. Once $\tilde{\mathbf{p}}$ and $\tilde{\boldsymbol{\mathcal{U}}}$ are known, the determination of $\tilde{\mathbf{f}}_I$ and $\tilde{\mathbf{h}}_I$ is straightforward.

It should be emphasized that the interface propagator matrix must be used even at interfaces of the second order (where gradients of parameters of the medium or of density vary discontinuously). Ignoring this may lead to inaccuracies and cause instability of the solution of the DRT system. Let us mention that interfaces of the second order are often introduced to the model artificially, during the approximation of the model, for example, when bi- or trilinear interpolation in a rectangular grid is used or if a piece-wise polynomial approximation using triangles (2D) or tetrahedrons (3D) is used. From these reasons, it is important to use splines or other approximations, which do not generate interfaces of the second order. The mentioned approximations may, on the other hand, generate false oscillations of the distribution of density-normalized elastic moduli. To avoid these effects, techniques developed for isotropic media and based on Sobolev scalar products and Lyapunov exponents (see Klimeš, 2000 for theory and Bulant, 2002 and Žáček, 2002 for numerical examples; see also Červený et al., 2007, Sec.6.2.1) should be extended to anisotropic media.

c) **Simplified DRT system in global Cartesian coordinates x_i**

In order to construct the 4×4 propagator matrix $\mathbf{\Pi}^{(y)}(\tau,\tau_0)$ in wavefront orthonormal coordinates, it is sufficient to solve the DRT system (B-5) for only six elements of matrix $\mathbf{Q}^{(x)}$ and six elements of matrix $\mathbf{P}^{(x)}$:

$$Q^{(x)}_{iN} = \partial x_i/\partial\gamma_N \ , \qquad P^{(x)}_{iN} = \partial p_i/\partial\gamma_N \ , \qquad (B-31)$$

where γ_1 and γ_2 are ray parameters. The elements $Q^{(x)}_{i3}$ and $P^{(x)}_{i3}$, $i = 1,2,3$, are not necessary to compute. The simplified DRT system in global Cartesian coordinates x_i reads:

$$\mathrm{d}Q^{(x)}_{iN}/\mathrm{d}\tau = A^{(x)}_{ij}Q^{(x)}_{jN} + B^{(x)}_{ij}P^{(x)}_{jN} \ , \qquad \mathrm{d}P^{(x)}_{iN}/\mathrm{d}\tau = -C^{(x)}_{ij}Q^{(x)}_{jN} - D^{(x)}_{ij}P^{(x)}_{jN} \ . \qquad (B-32)$$

Here the 3×3 matrices $A^{(x)}_{ij}$, $B^{(x)}_{ij}$, $C^{(x)}_{ij}$ and $D^{(x)}_{ij}$ are given by (B-2). The system (B-31) consists of twelve equations, not eighteen like (B-1).

Equations (24) and (25) give simple transformation rules from $Q^{(x)}_{iN}$, $P^{(x)}_{iN}$ to $Q^{(y)}_{JN}$ and $P^{(y)}_{JN}$, and back. Consequently, we can specify the initial conditions for the DRT system (B-31) in terms of initial conditions for $Q^{(y)}_{JN}$ and $P^{(y)}_{JN}$, which have clear physical meaning.

For a point source initial conditions, we have $Q^{(y)}_{JN}(\tau_0) = 0$ and $P^{(y)}_{JN}(\tau_0) = \delta_{JN}$. From this, we obtain, see (24):

$$Q^{(x)}_{iN}(\tau_0) = 0 \ , \qquad P^{(x)}_{iN}(\tau_0) = e_{Ni} - e_{Nj}p_i\mathcal{U}_j \ . \qquad (B-33)$$

Similarly, for plane wave initial conditions, we have $Q^{(y)}_{JN}(\tau_0) = \delta_{JN}$ and $P^{(y)}_{JN} = 0$. The initial conditions for $Q^{(x)}_{iN}(\tau_0)$ and $P^{(x)}_{iN}(\tau_0)$ are then as follows, see (24):

$$Q^{(x)}_{iN}(\tau_0) = e_{Ni} \ , \qquad P^{(x)}_{iN}(\tau_0) = e_{Nj}p_i\eta_j \ . \qquad (B-34)$$

Assume that we wish to compute the 4×4 propagator matrix $\Pi^{(y)}(\tau,\tau_0)$ corresponding to wavefront orthonormal coordinates using the DRT system (B-32). For this we need to solve (B-32) for a point source initial conditions (B-33) and for plane a wave initial conditions (B-34). Once this is done, we can transform $Q^{(x)}_{iN}(\tau)$ and $P^{(x)}_{iN}(\tau)$ into $Q^{(y)}_{JN}(\tau)$

and $P^{(y)}_{JN}(\tau)$ using (25) at an arbitrarily selected point τ on the ray Ω, and construct the required ray propagator matrix in wavefront orthonormal coordinates. Note that for this, we need to compute vectors $\mathbf{e}_I$ along the ray Ω.

Appendix C

Properties of ray propagator matrix $\mathbf{\Pi}(\tau, \tau_0)$

All the relations given here for the ray propagator matrix $\mathbf{\Pi}(\tau, \tau_0)$ are valid both for 4×4 propagator matrices $\mathbf{\Pi}^{(y)}(\tau, \tau_0)$ in wavefront orthonormal coordinates y_I, and for 6×6 propagator matrices $\hat{\mathbf{\Pi}}^{(x)}(\tau, \tau_0)$ in global Cartesian coordinates x_i. For this reason, we do not distinguish them here and express them without the superscripts (y) and (x).

The most important property of the ray propagator matrix is its symplecticity. This means that it satisfies the following matrix relation:

$$\mathbf{\Pi}^T(\tau, \tau_0)\mathbf{J}\mathbf{\Pi}(\tau, \tau_0) = \mathbf{J} \ . \qquad (C-1)$$

Here $\mathbf{J}$ is the matrix

$$\mathbf{J} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{pmatrix} \ . \qquad (C-2)$$

In (C-2), $\mathbf{I}$ is the identity matrix and $\mathbf{0}$ is the null matrix. The symplecticity of $\mathbf{\Pi}(\tau, \tau_0)$ has several important and useful consequences:

a) Matrix $\mathbf{\Pi}(\tau, \tau_0)$ satisfies Liouvillé's theorem: For $\det \mathbf{\Pi}(\tau_0, \tau_0) = 1$,

$$\det \mathbf{\Pi}(\tau, \tau_0) = 1 \ . \qquad (C-3)$$

Equation (C-3) is satisfied for any τ. Consequently, $\mathbf{\Pi}(\tau, \tau_0)$ is regular along the whole ray Ω.

b) Matrix $\mathbf{\Pi}(\tau, \tau_0)$ satisfies the chain rule:

$$\mathbf{\Pi}(\tau, \tau_0) = \mathbf{\Pi}(\tau, \tau_1)\mathbf{\Pi}(\tau_1, \tau_0) \ . \qquad (C-4)$$

The point corresponding to τ_1 may be any point of ray Ω, not necessarily between τ_0 and τ. The chain rule (C-4) can be extended to an arbitrary number of points τ_1, τ_2, τ_3, ... τ_n along the ray Ω.

c) The inverse propagator matrix $\mathbf{\Pi}^{-1}(\tau, \tau_0) = \mathbf{\Pi}(\tau_0, \tau)$ of $\mathbf{\Pi}(\tau, \tau_0)$ is always regular and is given by the relation

$$\mathbf{\Pi}^{-1}(\tau, \tau_0) = \begin{pmatrix} \mathbf{P}_2^T(\tau, \tau_0) & -\mathbf{Q}_2^T(\tau, \tau_0) \\ -\mathbf{P}_1^T(\tau, \tau_0) & \mathbf{Q}_1^T(\tau, \tau_0) \end{pmatrix} \ . \qquad (C-5)$$

References

Alkhalifah, T., 1995. Gaussian beam depth migration for anisotropic media. *Geophysics*, **60**, 1474 - 1484.

Bakker, P.M. , 1996. Theory of anisotropic dynamic ray tracing in ray-centered coordinates. *PAGEOPH*, **148**, 583 - 589.

Babich, V.M., 1961. Ray method of the computation of the intensity of wavefronts in elastic inhomogeneous anisotropic medium. In: *Problems of the dynamic theory of propagation of seismic waves* (in Russian), Vol. 5, ed. G.I. Petrashen, pp. 36–46. Leningrad: Leningrad Univ. Press. (Translation to English: *Geophys. J. Int.*, **118**, 379–383, 1994.)

Babich, V.M., 1968. Eigenfunctions, concentrated in the vicinity of a closed geodesics. In: *Mathematical problems of the theory of propagation of waves* (in Russian), Vol. 9, ed. V.M. Babich, pp. 15 - 63. Leningrad: Nauka.

Babich, V,M,, and Popov, M.M., 1981. Propagation of concentrated acoustical beams in three-dimensional inhomogeneous media (in Russian). *Akust. Zh.*, **27**, 828 - 835.

Bleistein, N., 2007. Mathematical modeling, migration and inversion with Gaussian beams. Lecture Notes. Center for Wave Phenomena, Colorado School of Mines, Golden, 118 p.

Bleistein, N., and Gray, S.H., 2010. Amplitude calculation for 3-D Gaussian beam migration using complex-valued traveltimes. Preprint.

Bulant, P., 1996. Two-point ray tracing in 3-D. *PAGEOPH*, **148**, 421 - 447.

Bulant, P., 2002. Sobolev scalar products in the construction of velocity models - application to model Hess and to SEG/EAGE Salt model. *PAGEOPH*, **159**, 1487 - 1506.

Bulant, P., and Klimeš, L., 2008. Numerical comparison of the isotropic-common-ray and anisotropic-common-ray approximations of the coupling ray theory. *Geophys. J. Int.*, **175**, 357–374.

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

Červený, V., 1985. Gaussian beam synthetic seismograms. *J. Geophys.*, **58**, 44 - 72.

Červený, V., 1989. Ray tracing in factorized anisotropic inhomogeneous media. *Geophys. J. Int.*, **99**, 91 - 100.

Červený, V., 2001. *Seismic ray theory.* Cambridge: Cambridge Univ. Press.

Červený, V., 2007. A note on dynamic ray tracing in ray-centered coordinates in anisotropic inhomogeneous media. *Stud. Geophys. Geod.*, **51**, 401 - 422.

Červený, V., and Klimeš, L., 2010. Transformation relations for second derivatives of travel time in anisotropic media. *Stud. Geophys. Geod.*, **54**, 257 - 267.

Červený, V., Klimeš, L., and Pšenčík, I., 2007. Seismic ray method: Recent developments. *Advances in Geophysics*, **48**, 1–126.

Červený, V., and Moser, T.J., 2007. Ray propagator matrices in three-dimensional anisotropic inhomogeneous layered structures. *Geophys. J. Int.*, **168**, 593–604.

Červený, V., Popov, M.M., and Pšenčík, I., 1982. Computation of wave fields in inhomogeneous media. Gaussian beam approach. *Geophys. J.R. astr. Soc.*, **88**, 43 - 79.

Červený, V., and Pšenčík, I., 1983a. Gaussian beams in two-dimensional elastic inhomogeneous media. *Geophys. J.R. astr. Soc.*, **72**, 417–433.

Červený, V., and Pšenčík, I., 1983b. Gaussian beams and paraxial ray approximation in three-dimensional elastic inhomogeneous media. *J. Geophys.*, **53**, 1–15.

Červený, V., and Pšenčík, I., 2009. Gaussian beams in inhomogeneous anisotropic layered structures. In *Seismic waves in complex 3-D structures*, Report 19, pp. 123–156. Charles University, Faculty of Mathematics and Physics, Dept. of Geophysics, Praha.

Červený, V., and Pšenčík, I., 2010. Gaussian beams in inhomogeneous anisotropic layered structures. *Geophys. J. Int.*, **180**, 798–812.

Coates, R.T., and Chapman, C.H., 1990. Quasi–shear wave coupling in weakly anisotropic 3-D media. *Geophys. J. Int.*, **103**, 301–320.

Chapman, C. H., 2004. *Fundamentals of seismic wave propagation.* Cambridge: Cambridge Univ. Press.

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

Farra,V., and Pšenčík,I., 2010. Coupled S waves in inhomogeneous weakly anisotropic media using first-order ray tracing. *Geophys.J.Int.*, **180**, 405-417.

Felsen, L.B., 1976. Complex-source point solutions of the field equations and their relations to the propagation and scattering of Gaussian beams. In *Ist.Naz.Alta Matem., Symp. Math.*, **18**, 39 - 56.

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.

George, T., Virieux,J., and Madariaga, R., 1987. Seismic wave synthesis by Gaussian beam summation: A comparison with finite differences. *Geophysics*, **52**, 1065 - 1073.

Gray, S., 2005. Gaussian beam migration of common shot records. *Geophysics*, **70**, S71-S77.

Gray, S.H., and Bleistein, N., 2009. True-amplitude Gaussian-beam migration. *Geophysics*, **74**, S11-S23.

Hanyga, A., 1982. Dynamic ray tracing in an anisotropic medium. *Tectonophysics*, **90**, 243 - 251.

Hanyga, A., 1986. Gaussian beams in anisotropic elastic media. *Geophys. J. R. astr.*

Soc., **85**, 473 - 503.

Hill, N., 1990. Gaussian beam migration. *Geophysics*, **55**, 1416 - 1428.

Hill, N.R., 2001. Prestack Gaussian beam depth migration. *Geophysics*, **66**, 1240 - 1250.

Iversen, E., and Pšenčík,I., 2007. Ray tracing for continuously rotated local coordinates belonging to a specified anisotropy *Stud. Geophys. Geod.*, **51**, 37–58.

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., 1984. Expansion of high-frequency time harmonic wavefield given on an initial surface into Gaussian beams. *Geophys. J.R. astr. Soc.*, **79**, 105 - 118.

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

Klimeš, L., 2000. Sobolev scalar products in the construction of velocity models. In: *Seismic Waves in Complex 3-D Structures*, Report 10, pp. 15-40, Charles Univ., Faculty of Mathematics and Physics, Dept. of Geophysics, Praha.

Klimeš, L., 2006. Ray-centered coordinate systems in anisotropic media. *Stud. Geoph. Geod.,*, **50**, 431 - 447.

Kravtsov, Yu.A., 1968. "Quasiisotropic" approximation to geometrical optics, *Dokl. AN SSSR*, **183**, No.1, 74-77 (in Russian).

Kravtsov, Yu.A., and Berczynski, P., 2007. Gaussian beams in inhomogeneous media; A review. *Stud. Geoph. Geod.,*, **51**, 1 - 36.

Leung, S., Qian, J., and Burridge, R., 2007. Eulerian Gaussian beams for high-frequency wave propagation. *Geophysics*, **72**, SM61 - SM76.

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

Moser, T.J., and Červený, V., 2007. Paraxial ray methods for anisotropic inhomogeneous media. *Geophys. Prospecting*, **55**, 21 - 37.

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

Popov, M.M., 2002. Ray theory and Gaussian beam method for geophysicists. Lecture Notes. Univ. of Bahia, Salvador, Brasil, 172 p.

Pšenčík, I., and Farra,V., 2005. First-order ray tracing for qP waves in inhomogeneous weakly anisotropic media. Geophysics, **70**, D65-D75.

Ralston, J., 1983. Gaussian beams and the propagation of singularities. In *Studies in partial differential equations*, ed. W. Littman, Math. Ass. of Am. Studies in

Mathematics, **23**, 206 - 248.

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.

Tanushev, N., 2008. Superpositions of higher-order Gaussian beams. *Commun. Math. Sci.*, **6**, 449 - 475.

Thomsen, L., 1986. Weak elastic anisotropy. *Geophysics*, **51**, 1954 - 1966.

Thomson, C. J., and Chapman, C. H., 1985. An inroduction to Maslov's asymptotic method. *Geophys. J. R. astr. Soc.*, **83**, 143 – 168.

Vinje, V., Roberts, G., and Taylor, R., 2008. Controlled beam migration: a versabile structure imaging tool. First Break, **26**, 109 - 113.

White, B., Norris, A., Bayliss, A., and Burridge, R., 1987. Some remarks on the Gaussian beam summation method. *Geophys. J.R. astr. Soc.*, **89**, 579 - 636.

Žáček, K., 2002. Smoothing the Marmousi model. *PAGEOPH*, **159**, 1507 - 1526.

Zhu, T., Gray, S.H., and Wang, D., 2005. Kinematic and dynamic ray tracing in anisotropic media: Theory and applications. In: *75th Annual Int. Meeting, SEG, Expanded Abstracts*, 96 - 99.

Zhu, T., Gray, S.H., and Wang, D., 2007. Prestack Gaussian-beam depth migration in anisotropic media. *Geophysics*, **72**, S133 - S138.