Geophysical Journal International

Geophys. J. Int. (2010) **180,** 798–812 doi: 10.1111/j.1365-246X.2009.04442.x

Gaussian beams in inhomogeneous anisotropic layered structures

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

Accepted 2009 October 31. Received 2009 October 24; in original form 2009 July 29

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 straightline 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 $\mathbf{M}$ of the second derivatives of the traveltime field with respect to ray-centred coordinates. Matrix $\mathbf{M}$ can be simply determined along the ray if the ray propagator matrix is known and if the value of $\mathbf{M}$ is specified at a selected point of the ray. The ray propagator matrix can be calculated along the ray by solving the dynamic ray tracing system twice: once for the real-valued initial plane-wave conditions and once for the real-valued initial point-source conditions. Alternatively, matrix $\mathbf{M}$ 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 (ray-centred, Cartesian, etc.). Here we use the ray-centred coordinate system, but propose a simple local transformation to Cartesian coordinates. This simplifies the computation of the Gaussian beams at the observation points situated in the vicinity of the central ray. The paper is self-contained and presents all the equations needed in computing the Gaussian beam. The proposed expressions for Gaussian beams are applicable to general 3-D inhomogeneous layered structures of arbitrary anisotropy (specified by up to 21 independent position-dependent elastic moduli). Possible simplifications are outlined.

Key words: 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 straightline 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, for example, 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 traveltime in the vicinity of the central ray. If we determine the traveltimes in the vicinity of the central ray by exact solution of the eikonal equation, we speak of 'strict Gaussian beams' (Červený *et al.* 2007, p. 76). In the vicinity of the central ray, called the paraxial vicinity, the complex-valued traveltime 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 traveltimes 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. In this paper, we call them simply 'Gaussian beams', without emphasizing the paraxial validity of their equations.

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 & Popov (1981), Popov (1982) and Červený *et al.* (1982). For the elastodynamic isotropic wave equation, see Červený & Pšenčík (1983a,b), Klimeš (1984), Červený (1985), George *et al.* (1987), White *et al.* (1987), Popov (2002), Bleistein (2007), Červený *et al.* (2007), Kravtsov & Berczynski (2007) and Leung *et al.* (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 algorithm and excellent results can be found in Hill (1990, 2001), Gray (2005), Vinje *et al.* (2008) and Gray & Bleistein (2009). 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) and Červený (2001, section 5.8). To seismic migration in anisotropic media, the Gaussian beam summation method was applied by Alkhalifah (1995) and Zhu *et al.* (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 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 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 Riccati equation for the 2×2 complex-valued matrix $\mathbf{M}$ of the second derivatives of the traveltime field with respect to 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 matrix $\mathbf{M}$ can be linearized by specifying $\mathbf{M} = \mathbf{P}\mathbf{Q}^{-1}$, where $\mathbf{P}$ and $\mathbf{Q}$ are again 2×2 matrices. Matrices $\mathbf{P}$ and $\mathbf{Q}$ 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 traveltime is computed not only along the central ray, but also in its 'paraxial vicinity'. The only difference is that matrices $\mathbf{Q}$, $\mathbf{P}$ and $\mathbf{M}$ 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 $\mathbf{Q}$, $\mathbf{P}$, $\mathbf{M}$ 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 DRT. The DRT system, and, consequently, the expressions for Gaussian beams, can be expressed in various coordinate systems (Cartesian, ray-centred, etc.). Here we use the ray-centred coordinate system, as it yields physically simple and understandable results. For anisotropic media, the DRT system in ray-centred coordinates is, of course, algebraically more complicated than in isotropic media, but the principles remain the same. The DRT system may be used for the computation of the matrix $\mathbf{M}(\tau)$ in two alternatives. (1) In the first alternative, 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 $\mathbf{M}(\tau)$ can be determined at any point of the central ray from the initial matrix $\mathbf{M}(\tau_0)$ by a single matrix multiplication. (2) In the second alternative, the DRT is solved directly, using specific complex-valued initial conditions containing $\mathbf{M}(\tau_0)$. 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 ray-centred coordinates at a paraxial (observation) point, it is useful to transform locally the 2×2 matrix $\mathbf{M}$ in ray-centred coordinates to the analogous 3×3 matrix $\mathbf{M}^{(x)}$ in Cartesian coordinates. We could obtain this matrix by using DRT in Cartesian coordinates as, for example, Červený (1972), Ralston (1983), Leung *et al.* (2007) and Tanushev (2008). However, here we concentrate on the DRT in ray-centred coordinates, and determine $\mathbf{M}^{(x)}$ only locally by transforming 2×2 matrix $\mathbf{M}$. 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, section 4.6). Since then, the procedure of Klimeš (1984) 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 approach of Klimeš (1984) 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 the Appendices. In the Appendices, however, we include all ray-theory pre-requisites 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 equations and DRT equations, and describe the computation of ray-theory amplitudes. In anisotropic media, both ray tracing and DRT 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 also explained. In Section 2, we also present an important transformation formula allowing conversion of matrix $\mathbf{M}$ in ray-centred coordinates into $\mathbf{M}^{(x)}$ in Cartesian coordinates. Section 3 is devoted to paraxial ray approximations. The expressions for paraxial traveltimes 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 discussion of possible simplifications of the procedure of computing Gaussian beams.

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 **a** and **b** reads $\mathbf{a}^T\mathbf{b}$, the dyadic reads $\mathbf{a}\mathbf{b}^T$. Whenever there may be reason for confusion, the dimensions of the matrices are explicitly indicated. 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, DRT 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 a heterogeneous 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 traveltime, $\mathbf{U}(x_i)$ 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_i:

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

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 **p**.

The Christoffel matrix (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 mth elementary wave. It follows from (3) that the eigenvalue G_m of this wave satisfies the relation

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

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

$$\mathcal{H}(x_i, p_j) = \frac{1}{2}[G_m(x_i, p_j) - 1] = 0. \tag{7}$$

The Hamiltonian $\mathcal{H}(x_i, p_j)$ is used in ray tracing, see Appendix A, from which the traveltime is usually computed.

The vectorial ray-theory amplitude **U** can be expressed in terms of the unit real-valued eigenvector $\mathbf{g}^{(m)}$ of the Christoffel matrix (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. Eq. (8) shows that eigenvector $\mathbf{g}^{(m)}$ specifies the polarization of the wave under consideration. For this reason, we call $\mathbf{g}^{(m)}(x_i)$ the polarization 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* 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 different ways. In inhomogeneous anisotropic media, it is most convenient to take $\gamma_3 = \tau$, where τ is the traveltime T along ray Ω of the wave under consideration. The ray tracing equations for inhomogeneous anisotropic media 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 *et al.* 2005, 2007) or weak-anisotropy (WA) parameters (see Pšenčík & 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 $\boldsymbol{\mathcal{U}}(\tau) = \mathrm{d}\mathbf{x}(\tau)/\mathrm{d}\tau$, 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 wave front, the vector $\boldsymbol{\eta}(\tau) = \mathrm{d}\mathbf{p}(\tau)/\mathrm{d}\tau$, 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 traveltime $T(\tau)$ 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 traveltimes (the complex-valued traveltimes 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 (DRT).

2.3 Dynamic ray tracing

DRT consists in a solution of a system of linear ordinary differential equations of the first order, which can be solved along the central ray Ω, together with the ray tracing system or along an *a priori* known ray Ω. The results of DRT can be used to compute the traveltime not only along the ray, but also in its vicinity. The DRT system can be expressed in various coordinate systems (ray-centred, Cartesian, etc.) In this paper, we consider the ray-centred coordinate system q_i, $i = 1, 2, 3$. For many other details on ray-centred coordinates and DRT in ray-centred coordinates see Klimeš (1994) and Červený (2001).

The basic property of the ray-centred coordinate system is that the selected central ray Ω represents its coordinate axis q_3. The q_1 and q_2 coordinate axes may be introduced in many ways. For an up-to-date review see Klimeš (2006). Mostly, they are introduced as mutually perpendicular straight lines situated in the plane tangent to the wave front, and intersecting at the central ray Ω. The transformation equation between Cartesian coordinates x_i and ray-centred coordinates q_j is then given by the relation

$$x_i(q_j) = x_i^{\Omega}(q_3) + H_{iM}(q_3)q_M, \tag{9}$$

where $i, j = 1, 2, 3$ and $M = 1, 2$. Central ray Ω is specified by $q_1 = q_2 = 0$, so that $\mathbf{x}(q_3) = \mathbf{x}^{\Omega}(q_3)$, where $\mathbf{x}^{\Omega}$ denotes a point on the ray Ω. Similarly as in kinematic ray tracing, we take coordinate q_3 to equal the traveltime τ along central ray Ω, $q_3 = \tau$. Let us emphasize that the choice $q_3 = \tau$ simplifies the computations considerably, as it leads to the following simple relations, valid at any point of the central ray Ω:

$$\partial T/\partial q_3 = 1, \quad \partial^2 T/\partial q_3 \partial q_i = 0. \tag{10}$$

Note also that $\partial T/\partial q_I = 0$ at any point of Ω. Simple relations (10) are not valid for any other parameter q_3 along Ω, for example, for the arclength s. It is important to emphasize this, since the arclength along Ω has been commonly used in ray-centred coordinates in heterogeneous isotropic media. In this aspect, our treatment differs from the common treatment in isotropic media.

The elements of the 3×3 transformation matrices $\mathbf{H}$ and $\overline{\mathbf{H}}$ from ray-centred to Cartesian coordinates and back, are defined as follows:

$$H_{im} = \partial x_i/\partial q_m, \quad \overline{H}_{im} = \partial q_i/\partial x_m. \tag{11}$$

As the ray-centred coordinate system is curvilinear, we have to distinguish two systems of basis vectors: the contravariant basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ (tangential to coordinate axes), and the covariant basis vectors $\mathbf{f}_1$, $\mathbf{f}_2$, $\mathbf{f}_3$ (perpendicular to coordinate surfaces). These basis vectors are defined as columns of 3×3 transformation matrices $\mathbf{H}$ and $\overline{\mathbf{H}}^T$:

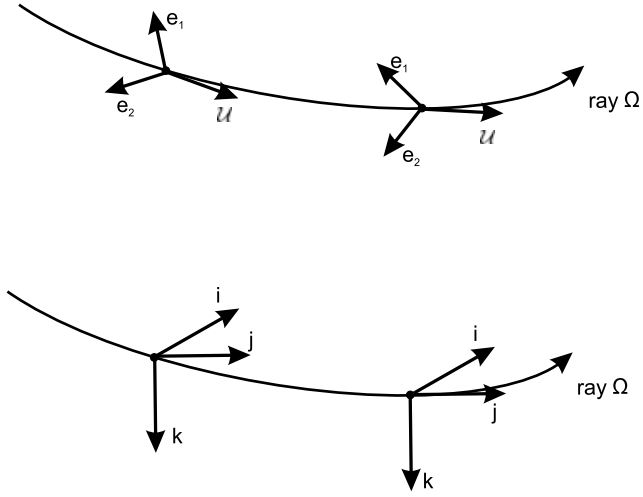

Figure 1. Top panel: Central ray Ω with basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ situated in the plane tangent to the wave front at two points of Ω. The vectors $\boldsymbol{\mathcal{U}}$ tangent to the ray Ω are ray-velocity vectors. Bottom panel: Central ray Ω with the basis vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ of the Cartesian coordinate system at the same points of Ω as above. The basis vectors $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ do not vary along the ray.

$$\mathbf{H} = (\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 = \boldsymbol{\mathcal{U}}), \quad \overline{\mathbf{H}}^T = (\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3 = \mathbf{p}). \tag{12}$$

The obvious relation $\overline{H}_{ik}H_{kj} = \delta_{ij}$ can be expressed in terms of $\mathbf{e}_i$ and $\mathbf{f}_i$ as follows:

$$\mathbf{f}_i^T \mathbf{e}_j = \delta_{ij}. \tag{13}$$

The basis vectors $\mathbf{e}_1, \mathbf{e}_2$ (tangent to the wave front) and $\mathbf{e}_3$ (tangent to the central ray Ω) are shown schematically in the upper part of Fig. 1. In the bottom part, Cartesian basis vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are shown. In contrast to the vectors $\mathbf{e}_i$, the latter vectors do not vary along the ray Ω. Eq. (13) yields vectorial relations $\mathbf{p}^T\boldsymbol{\mathcal{U}} = 1, \mathbf{p}^T\mathbf{e}_I = 0, \boldsymbol{\mathcal{U}}^T\mathbf{f}_I = 0$. Thus, vectors $\mathbf{e}_I$ are perpendicular to the slowness vector $\mathbf{p}$, and vectors $\mathbf{f}_I$ are perpendicular to the ray-velocity vector $\boldsymbol{\mathcal{U}}$, that is, to the ray. Eq. (13) also yields

$$\mathbf{f}_1 = \frac{\mathbf{e}_2 \times \boldsymbol{\mathcal{U}}}{\boldsymbol{\mathcal{U}}^T(\mathbf{e}_1 \times \mathbf{e}_2)} = \mathcal{C}^{-1}(\mathbf{e}_2 \times \boldsymbol{\mathcal{U}}),$$
$$\mathbf{f}_2 = \frac{\boldsymbol{\mathcal{U}} \times \mathbf{e}_1}{\boldsymbol{\mathcal{U}}^T(\mathbf{e}_1 \times \mathbf{e}_2)} = \mathcal{C}^{-1}(\boldsymbol{\mathcal{U}} \times \mathbf{e}_1). \tag{14}$$

The contravariant basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ can be specified by solving a simple ordinary differential equation of the first order along central ray Ω:

$$\mathrm{d}\mathbf{e}_I/\mathrm{d}\tau = -\left(\mathbf{e}_I^T\boldsymbol{\eta}\right)\mathbf{p}/(\mathbf{p}^T\mathbf{p}). \tag{15}$$

Here $\mathbf{p}$ and $\boldsymbol{\eta} = \mathrm{d}\mathbf{p}/\mathrm{d}\tau$ are known from kinematic ray tracing. At a selected point τ_0 of central ray Ω, we take the initial values for $\mathbf{e}_1$ and $\mathbf{e}_2$ such that they form a triplet of mutually perpendicular unit vectors with $\mathbf{N} = \mathcal{C}\mathbf{p}$. It is possible to prove that vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{N}$ are then right-handed, unit and mutually perpendicular along the whole central ray Ω. Consequently, it is sufficient to calculate only $\mathbf{e}_1(\tau)$ using (15), as $\mathbf{e}_2(\tau) = \mathbf{N}(\tau) \times \mathbf{e}_1(\tau)$. Moreover, the numerical determination of $\mathbf{e}_1(\tau)$ using (15) is sufficient to determine analytically all vectors $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{f}_1$, $\mathbf{f}_2$. Since $\mathbf{e}_3 = \boldsymbol{\mathcal{U}}$ and $\mathbf{f}_3 = \mathbf{p}$ are known from ray tracing, we can also determine the complete transformation matrices $\mathbf{H}$ and $\overline{\mathbf{H}}$, see (12). Note that vectors $\mathbf{f}_1$ and $\mathbf{f}_2$ are not necessarily unit and mutually perpendicular, but $\mathbf{e}_1$ and $\mathbf{e}_2$ are.

The DRT system in ray-centred coordinates can be used to compute the following partial derivatives along the central ray Ω:

$$Q_I = \partial q_I/\partial\gamma, \quad P_I = \partial p_I^{(q)}/\mathrm{d}\gamma. \tag{16}$$

Here γ is any of the ray parameters. Symbol $p_I^{(q)}$ denotes the Ith component of slowness vector $\mathbf{p}$ in the ray-centred coordinate system, $p_I^{(q)} = H_{kl}\, p_k$. Quantities Q_I and $P_I (I = 1, 2)$ specify the off-ray paraxial changes of q_I and $p_I^{(q)}$, caused by the changes of γ. Q_I and P_I may be used to calculate paraxial rays, paraxial traveltimes, the paraxial slowness vectors, etc. In 3-D models, the DRT system consists of four linear ordinary differential equations of the first order in Q_I and P_I. The actual form of this DRT system is presented in Appendix B. The transformation of the DRT system across a structural interface is realized by the so-called interface propagator matrix, see also Appendix B.

In an orthonomic system of rays, in which any ray is specified by two ray parameters γ_1, γ_2, it is useful to express the DRT system in ray-centred coordinates in matrix form. We introduce two 2×2 matrices $\mathbf{Q}$ and $\mathbf{P}$, with elements

$$Q_{IJ} = \partial q_I/\partial\gamma_J, \quad P_{IJ} = \partial p_I^{(q)}/\partial\gamma_J. \tag{17}$$

To compute the 2×2 matrices $\mathbf{Q}$ and $\mathbf{P}$, we need to solve the DRT system for Q_I and P_I, consisting of four equations twice, with different initial conditions corresponding to ray parameters γ_1 and γ_2. The coefficients in both DRT systems, however, remain the same. Both systems can be solved together. As the coefficients in both systems are evaluated only once, the solution for Q_{IJ} and P_{IJ} is not much slower than the solution for Q_I and P_I.

From the 2×2 matrices $\mathbf{Q}$ and $\mathbf{P}$, we can obtain the important 2×2 matrix $\mathbf{M}$ of the second derivatives of traveltime field T with respect to ray-centred coordinates q_1 and q_2,

$$M_{IJ} = \partial^2 T/\partial q_I \partial q_J. \tag{18}$$

It is easy to show that the symmetric matrix $\mathbf{M}$ with three independent elements M_{IJ} can be expressed in terms of $\mathbf{Q}$ and $\mathbf{P}$. As

$$M_{IJ} Q_{JK} = \frac{\partial^2 T}{\partial q_I \partial q_J}\frac{\partial q_J}{\partial \gamma_K} = \frac{\partial p_I^{(q)}}{\partial q_J}\frac{\partial q_J}{\partial \gamma_K} = \frac{\partial p_I^{(q)}}{\partial \gamma_K} = P_{IK}, \tag{19}$$

we obtain

$$\mathbf{M} = \mathbf{P}\mathbf{Q}^{-1}. \tag{20}$$

The 2×2 matrix $\mathbf{M}$ could also be computed directly, using the Riccati equation solved along the central ray Ω. The Riccati equation, however, is non-linear and is not so suitable for computations as the DRT equations. Moreover, the linear DRT system can be used to construct the 4×4 ray-propagator matrix, which plays an important role in paraxial computations and in the computation of Gaussian beams. The Riccati equation cannot be used for this purpose.

A brief note to the DRT system in Cartesian coordinates. In this paper, we consider the DRT system in ray-centred coordinates, but we do not want to claim that ray-centred coordinates are always preferable in comparison with Cartesian coordinates. In the DRT system in Cartesian coordinates, the 3×3 matrices $\mathbf{Q}^{(x)}$ and $\mathbf{P}^{(x)}$, with elements $Q_{ij}^{(x)}$ and $P_{ij}^{(x)}$, are computed along the ray. Thus the number of solved equations is higher than in ray-centred coordinates. The DRT system in Cartesian coordinates, however, has simpler right-hand sides, especially in anisotropic media. The DRT system in Cartesian coordinates for anisotropic inhomogeneous media was derived by Červený (1972) with the purpose to compute the geometrical spreading and ray amplitudes. The DRT system in Cartesian coordinates can be simplified. By a suitable choice of the initial conditions (point-source and plane-wave initial conditions) for DRT, it is sufficient to compute only the first two columns of the 3×3 matrices $\mathbf{Q}^{(x)}$ and $\mathbf{P}^{(x)}$. For more details see Červený (2001, section 4.2) and Červený & Pšenčík (2009). Let us note that such simplified Cartesian DRT system with point-source initial conditions is used in the program package ANRAY (Gajewski & Pšenčík, 1991; http://sw3d.mff.cuni.cz). The numerical efficiency of the DRT systems in different coordinates has not yet been systematically investigated.

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):

$$\mathbf{u}(\tau) = A(\tau)\mathbf{g}(\tau)\exp\{-\mathrm{i}\omega[t - T(\tau)]\}. \tag{21}$$

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 traveltime along ray Ω. The traveltime 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. We consider two points on the ray specified by τ and τ_0, and assume that $A(\tau_0)$ is known. Then the continuation formula reads

$$A(\tau) = \left[\frac{\rho(\tau_0)\mathcal{C}(\tau_0)}{\rho(\tau)\mathcal{C}(\tau)}\right]^{1/2}\left[\frac{\det\mathbf{Q}(\tau_0)}{\det\mathbf{Q}(\tau)}\right]^{1/2}\mathcal{R}^C A(\tau_0). \tag{22}$$

In eq. (22), 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 $\mathbf{Q}$ is the 2×2 matrix whose elements are defined in eq. (17). Matrix $\mathbf{Q}$ is often called the matrix of geometrical spreading, and $|\det\mathbf{Q}|^{1/2}$ the geometrical spreading. In the ray theory, the term $[\det\mathbf{Q}(\tau_0)/\det\mathbf{Q}(\tau)]^{1/2}$ 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 (22), which is convenient in the study of Gaussian beams. Continuation relation (22) is regular as long as $\det\mathbf{Q}(\tau)$ is non-zero. We also assume that $\det\mathbf{Q}(\tau_0) \neq 0$. 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, for example, 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 of the wave front 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. (22) would include an additional multiplication factor.

We emphasize the important assumption made in the derivation of the continuation relation (22), namely that $\det \mathbf{Q}(\tau_0) \neq 0$. This means that (22) cannot be used for individual Gaussain beams if a point source is situated at 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 $\det \mathbf{Q}(\tau_0) \neq 0$, the superposition represents approximately the wavefield generated by a point source see, for example, Hill (2001; Appendix B).

There is another important aspect related to the use of eq. (22) 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, however, expect decrease of the accuracy of the scalar ray-theory amplitude $A(\tau)$.

It is useful to express (21) with (22) in the following form:

$$\mathbf{u}(\tau) = \mathbf{U}^{\Omega}(\tau) \left[\frac{\det \mathbf{Q}(\tau_0)}{\det \mathbf{Q}(\tau)} \right]^{1/2} \exp\{-\mathrm{i}\omega[t - T(\tau)]\}, \tag{23}$$

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

$$\mathbf{U}^{\Omega}(\tau) = \left[\frac{\rho(\tau_0)\mathcal{C}(\tau_0)}{\rho(\tau)\mathcal{C}(\tau)} \right]^{1/2} \mathcal{R}^C A(\tau_0)\mathbf{g}(\tau). \tag{24}$$

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.

We introduce the 4×4 real-valued ray propagator matrix $\mathbf{\Pi}(\tau, \tau_0)$ as the fundamental matrix of the DRT system in ray-centred coordinates

$$\mathrm{d}\mathbf{\Pi}(\tau, \tau_0)/\mathrm{d}\tau = \mathbf{S}(\tau)\mathbf{\Pi}(\tau, \tau_0), \tag{25}$$

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

$$\mathbf{\Pi}(\tau_0, \tau_0) = \mathbf{I}. \tag{26}$$

Here $\mathbf{S}(\tau)$ is the 4×4 system matrix of the DRT system and $\mathbf{I}$ is the 4×4 identity matrix. Note that the fundamental matrix is a matrix composed of four linearly independent solutions of the DRT system. The linearly independent solutions are guaranteed by the initial conditions (26). The specific expressions, which form the system matrix $\mathbf{S}(\tau)$ of the DRT systems, are shown in Appendix B.

Let us make two important comments to eq. (25) with the initial condition (26). First, all the four 4×1 linearly independent solutions of (25) 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 $\mathbf{\Pi}(\tau, \tau_0)$ using eqs (25) with (26) consists in the determination of the 4×4 system matrix $\mathbf{S}(\tau)$ along the central ray. The system matrix $\mathbf{S}(\tau)$ is, however, the same for all four linearly independent solutions forming $\mathbf{\Pi}(\tau, \tau_0)$. Consequently, there is no great difference between seeking one of the linearly independent columns of $\mathbf{\Pi}(\tau, \tau_0)$ and the whole $\mathbf{\Pi}(\tau, \tau_0)$. Second, the solution of DRT system (25) for $\mathbf{\Pi}(\tau, \tau_0)$ is much faster than standard ray tracing. Except the second-order derivatives of the Hamiltonian $\mathcal{H}(x_i, p_j)$, only the terms already used to determine the right-hand sides of the ray tracing equations are used to solve (25). We remind the reader that the DRT system is linear while the ray tracing system is generally non-linear.

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

$$\mathbf{\Pi}(\tau, \tau_0) = \begin{bmatrix} \mathbf{Q}_1(\tau, \tau_0) & \mathbf{Q}_2(\tau, \tau_0) \\ \mathbf{P}_1(\tau, \tau_0) & \mathbf{P}_2(\tau, \tau_0) \end{bmatrix}. \tag{27}$$

Here $\mathbf{Q}_1(\tau, \tau_0)$, $\mathbf{Q}_2(\tau, \tau_0)$, $\mathbf{P}_1(\tau, \tau_0)$ and $\mathbf{P}_2(\tau, \tau_0)$ are 2×2 real-valued matrices. For the orthonomic system of rays, these matrices have a very simple physical meaning following from (26):

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

$$\mathbf{Q}_1(\tau_0, \tau_0) = \mathbf{I}, \quad \mathbf{P}_1(\tau_0, \tau_0) = \mathbf{0}. \tag{28}$$

In (28), $\mathbf{I}$ 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 wave front at $\tau = \tau_0$.

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

$$\mathbf{Q}_2(\tau_0, \tau_0) = \mathbf{0}, \quad \mathbf{P}_2(\tau_0, \tau_0) = \mathbf{I}. \tag{29}$$

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 is solved twice: once with the initial plane-wave conditions corresponding to (28) and once with the initial point-source conditions corresponding to (29). Alternatively, we can say that the DRT system should be solved four times, with the initial conditions specified by the four columns of the 4×4 identity matrix.

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

$$\begin{bmatrix} \mathbf{Q}(\tau) \\ \mathbf{P}(\tau) \end{bmatrix} = \mathbf{\Pi}(\tau, \tau_0) \begin{bmatrix} \mathbf{Q}(\tau_0) \\ \mathbf{P}(\tau_0) \end{bmatrix}. \tag{30}$$

The 4×4 ray propagator matrix $\mathbf{\Pi}(\tau, \tau_0)$ has 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 $\mathbf{\Pi}(\tau, \tau_0)$ can be used to construct a simple analytical relation for the continuation of the 2×2 matrix $\mathbf{M}$ of the second spatial derivatives of the traveltime field along the ray Ω. Using (20), (27) and (30), we obtain

$$\mathbf{M}(\tau) = \mathbf{P}(\tau)\mathbf{Q}^{-1}(\tau) = [\mathbf{P}_1(\tau, \tau_0) + \mathbf{P}_2(\tau, \tau_0)\mathbf{M}(\tau_0)] \times [\mathbf{Q}_1(\tau, \tau_0) + \mathbf{Q}_2(\tau, \tau_0)\mathbf{M}(\tau_0)]^{-1}. \quad (31)$$

This is a very important relation in the theory of Gaussian beams. Once we know the ray propagator matrix $\mathbf{\Pi}(\tau, \tau_0)$ along ray Ω, and matrix $\mathbf{M}(\tau_0)$ at the point of ray Ω with $\tau = \tau_0$, eq. (31) provides a simple way of determining $\mathbf{M}(\tau)$ for arbitrary τ along ray Ω. If we use the interface propagator matrix (B-6) in $\mathbf{\Pi}(\tau, \tau_0)$, see (C-6), eq. (31) 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 $\mathbf{\Pi}(\tau, \tau_0)$ is in the transformation of the factor $[\det \mathbf{Q}(\tau_0)/\det \mathbf{Q}(\tau)]^{1/2}$ appearing in the formula for ray-theory amplitudes, see eq. (22). We obtain

$$[\det \mathbf{Q}(\tau_0)/\det \mathbf{Q}(\tau)]^{1/2} = \{\det[\mathbf{Q}_1(\tau, \tau_0) + \mathbf{Q}_2(\tau, \tau_0)\mathbf{M}(\tau_0)]\}^{-1/2}. \quad (32)$$

Similarly as eq. (31), eq. (32) plays an important role in the computation of Gaussian beams. The great advantage of the expressions on the right hand side of eqs (31) and (32) is that they are expressed in terms of known submatrices $\mathbf{Q}_1$, $\mathbf{Q}_2$, $\mathbf{P}_1$ and $\mathbf{P}_2$, of the propagator matrix and of physically well understandable initial conditions, namely in terms of the matrix of the second spatial derivatives of the traveltime field $\mathbf{M}(\tau_0)$ in ray-centred coordinates. No other initial values are required.

2.6 Transformation of M(τ) to Cartesian coordinates

For evaluation of paraxial traveltimes, it is useful to transform the 2×2 matrix $\mathbf{M}(\tau)$ in ray-centred coordinates to 3×3 matrix $\mathbf{M}^{(x)}(\tau)$ in Cartesian coordinates. The appropriate relation was derived by Červený & Klimeš (2009) and reads

$$\mathbf{M}^{(x)} = \mathbf{f}\mathbf{M}\mathbf{f}^T + \mathbf{p}\boldsymbol{\eta}^T + \boldsymbol{\eta}\mathbf{p}^T - \mathbf{p}\mathbf{p}^T(\boldsymbol{\mathcal{U}}^T\boldsymbol{\eta}). \quad (33)$$

Here $\mathbf{M}^{(x)}$ denotes the 3×3 matrix of the second derivatives of the traveltime field with respect to Cartesian coordinates, with elements $M^{(x)}_{ij} = \partial^2\tau/\partial x_i \partial x_j$, $\mathbf{M}$ is the 2×2 matrix of the second derivatives of the traveltime field with respect to ray-centred coordinates, with elements $M_{IJ} = \partial^2\tau/\partial q_I \partial q_J$. The quantities $\mathbf{p}$, $\boldsymbol{\eta}$ and $\boldsymbol{\mathcal{U}}$ are slowness, eta and ray-velocity vectors, which are known from ray tracing and $\mathbf{f}$ is a 3×2 matrix, $\mathbf{f} = (\mathbf{f}_1, \mathbf{f}_2)$, where $\mathbf{f}_I$ are the covariant basis vectors of ray-centred coordinates, see eq. (14). Note that vectors $\mathbf{f}_I$ are also needed in DRT, see Appendix B. Thus the transformation equation (33) can be used for computation of the matrix $\mathbf{M}^{(x)}$ from $\mathbf{M}$ at any point of central ray Ω without any additional computations.

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 traveltime field T in the 'paraxial' vicinity of the ray Ω, not only on the ray Ω itself. We call such traveltimes the paraxial traveltimes. The knowledge of the real-valued paraxial traveltimes is a necessary pre-requisite of many other extensions of the ray method including Gaussian beams. For this reason, we first derive useful expressions for the real-valued paraxial traveltimes. After this, we use the expressions for paraxial traveltimes in equations for the relevant displacement vector and obtain the so-called paraxial approximation of the displacement vector.

3.1 Paraxial traveltimes

Paraxial traveltime is usually specified by Taylor expansion of traveltime T to quadratic terms at point R_Ω situated on ray Ω. In ray-centred coordinates q_1, q_2, q_3, ray Ω is the q_3-axis of the ray-centred coordinate system, with $q_3 = \tau$. The other two coordinates q_1, q_2 of the ray-centred coordinate system are zero along Ω, so that the ray-centred coordinates of R_Ω are $(0, 0, q_3 = \tau)$. The ray-centred coordinates $\mathbf{q} \equiv (q_1, q_2)^T$ are Cartesian coordinates in the plane tangent to the wave front at R_Ω, with the origin at R_Ω. Then the paraxial traveltime $T(R)$ at the point $R(q_1, q_2, q_3 = \tau)$ can be expressed in terms of $T(R_\Omega)$ as follows:

$$T(R) = T(R_\Omega) + \frac{1}{2}\mathbf{q}^T(R)\mathbf{M}(R_\Omega)\mathbf{q}(R). \quad (34)$$

The linear term with respect to $\mathbf{q}(R)$ is absent in (34), as the ray-centred components of the slowness vector, $p_1^{(q)}(R_\Omega)$ and $p_2^{(q)}(R_\Omega)$ at R_Ω on Ω are zero (the slowness vector is perpendicular to the wave front). Eq. (34) for the paraxial traveltime in ray-centred coordinates is very simple. It only requires knowledge of $T(R_\Omega)$ and of the 2×2 matrix $\mathbf{M}(R_\Omega) \equiv \mathbf{M}(\tau)$.

There are two approaches to compute the 2×2 matrix $\mathbf{M}(R_\Omega)$ of the second derivatives of the traveltime field at an arbitrary point R_Ω on the central ray Ω. In both approaches, we need to know the 2×2 matrix $\mathbf{M}(\tau_0)$ at a given 'initial' point on Ω, corresponding to τ_0. The matrix $\mathbf{M}(\tau_0)$ must be (1) symmetric and (2) finite. The first approach requires computation of the 4×4 ray propagator matrix $\mathbf{\Pi}(\tau, \tau_0)$ in ray-centred coordinates along the central ray Ω, see Section 2.5. Then we can obtain the 2×2 matrix $\mathbf{M}(\tau)$ using the simple relation (31). The other approach is to compute $\mathbf{M}(\tau)$ from $\mathbf{M}(\tau) = \mathbf{P}(\tau)\,\mathbf{Q}^{-1}(\tau)$, where $\mathbf{P}(\tau)$ and $\mathbf{Q}(\tau)$ are calculated directly from DRT system with specific initial conditions. We can choose the initial conditions $\mathbf{Q}(\tau_0)$ and $\mathbf{P}(\tau_0)$ in the DRT system (B-1) in the following way:

$$\mathbf{Q}(\tau_0) = \mathbf{A}, \qquad \mathbf{P}(\tau_0) = \mathbf{M}_0\mathbf{A}. \quad (35)$$

Here $\mathbf{A}$ is an arbitrary constant real-valued finite 2×2 matrix, for which $\det \mathbf{A} \neq 0$. The 2×2 matrix $\mathbf{M}_0$ is real-valued, finite and symmetric. Note that $\mathbf{M}(\tau_0) = \mathbf{P}(\tau_0)\,\mathbf{Q}^{-1}\,(\tau_0) = \mathbf{M}_0$. Solving the DRT system (B-1) with initial conditions (35), we can obtain $\mathbf{Q}(\tau)$ and $\mathbf{P}(\tau)$ at an arbitrary point of central ray Ω. From $\mathbf{Q}(\tau)$ and $\mathbf{P}(\tau)$ we then get

$$\mathbf{M}(\tau) = \mathbf{P}(\tau)\mathbf{Q}^{-1}(\tau),$$
$$[\det \mathbf{Q}(\tau_0)/\det \mathbf{Q}(\tau)]^{1/2} = [\det \mathbf{A}/\det \mathbf{Q}(\tau)]^{1/2}. \quad (36)$$

The matrix $\mathbf{M}(\tau)$ obtained by both approaches is the same. The advantage of the first approach is its high flexibility. Once the ray propagator matrix $\mathbf{\Pi}(\tau, \tau_0)$ is known, the matrix $\mathbf{M}(\tau)$ can be obtained from (31) for various specifications of $\mathbf{M}(\tau_0)$ with practically no additional work. Contrary to it, the second approach requires repeated solution of DRT for every new specification of $\mathbf{M}(\tau_0)$.

3.2 Paraxial approximation of the displacement vector

In this section, we discuss extension of the zero-order ray-theory expression for the displacement vector (23), by including the paraxial traveltimes (34). Consequently, the resulting expression gives displacement vector not only along the ray Ω, but also in its paraxial vicinity. We obtain

$$\mathbf{u}^{par}(q_1, q_2, \tau) = \mathbf{U}^{\Omega}(\tau)\{\det[\mathbf{Q}_1(\tau, \tau_0) + \mathbf{Q}_2(\tau, \tau_0)\mathbf{M}(\tau_0)]\}^{-1/2} \times \exp\left\{-\mathrm{i}\omega\left(t - T(\tau) - \frac{1}{2}\mathbf{q}^T\mathbf{M}(\tau)\mathbf{q}\right)\right\}. \quad (37)$$

Here $\mathbf{q} \equiv (q_1, q_2)^T$ are ray-centred coordinates, $\mathbf{U}^{\Omega}(\tau)$ is the spreading-free amplitude, given by (24), and $\mathbf{M}(\tau)$ is the 2×2 matrix of second derivatives of the traveltime field with respect to ray-centred coordinates. If we know the ray propagator matrix $\mathbf{\Pi}(\tau, \tau_0)$ given by (27), the matrix $\mathbf{M}(\tau)$ can be determined from equation (31). The factor $\{\det[\mathbf{Q}(\tau_0)/\det\mathbf{Q}(\tau)]\}^{1/2}$ may be expressed in terms of 2×2 matrices $\mathbf{Q}_1(\tau, \tau_0)$, $\mathbf{Q}_2(\tau, \tau_0)$ and $\mathbf{M}(\tau_0)$ using (32). Consequently, there is no need for any additional computations. If we solve the DRT system for specific initial conditions like in eq. (35), the matrix $\mathbf{M}(\tau)$ and the factor $\{\det[\mathbf{Q}(\tau_0)/\det\mathbf{Q}(\tau)]\}^{1/2}$ must be determined from (36).

When the ray propagator matrix $\mathbf{\Pi}(\tau, \tau_0)$ is known, the computation of paraxial approximation of the displacement vector (37) requires the specification of an additional initial condition to the standard ray-theory initial conditions, namely the specification of the 2×2 matrix $\mathbf{M}(\tau_0)$ of second derivatives of the traveltime field at an arbitrary point $\tau = \tau_0$ of the ray Ω. No initial conditions specified separately for $\mathbf{Q}(\tau_0)$ and $\mathbf{P}(\tau_0)$ are required. Fortunately, the 2×2 matrix $\mathbf{M}(\tau_0)$ has a simple physical meaning. It can be expressed in terms of the curvature matrix of the wave front at τ_0. Remember that the 2×2 matrix $\mathbf{M}(\tau_0)$ must be specified symmetric and finite.

In eq. (37), we used expansion of the traveltime T from the central ray Ω to the paraxial vicinity of Ω. We, however, did not expand the ray-theory amplitudes in the vicinity of Ω; they were left the same as on the ray Ω. The expansion of the ray-theory amplitudes 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 methods. Such higher-order paraxial methods, however, have not yet been investigated.

Another possibility how to improve the paraxial ray-theory 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 (37), 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 & Drummond (1982), Thomson & Chapman (1985) and Chapman (2004).

To evaluate the paraxial approximation of the displacement vector $\mathbf{u}^{par}(R)$ at a paraxial point $R = R(q_i)$ using (37), we must know the ray-centred coordinates $q_1, q_2, q_3 = \tau$ of the point R. The paraxial point R is, however, often specified in Cartesian coordinates. The determination of the ray-centred 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 $\mathbf{M}(\tau)$ is transformed to the 3×3 matrix $\mathbf{M}^{(x)}(\tau)$. This can be easily done with formula (33), described in Section 2.6. We discuss the use of eq. (33) and its consequences, which hold for paraxial approximation of the displacement vector as well as for Gaussian beams, in detail in Sections 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 straightline 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 (37), in which the 2×2 matrix $\mathbf{M}(\tau_0)$ of the second derivatives of the traveltime field is complex valued:

$$\mathbf{M}(\tau_0) = \mathrm{Re}\mathbf{M}(\tau_0) + \mathrm{i}\mathrm{Im}\mathbf{M}(\tau_0). \quad (38)$$

In addition, we require that matrix $\mathbf{M}(\tau_0)$ satisfies the following Gaussian-beam existence conditions: (i) $\mathbf{M}(\tau_0)$ is symmetric, (ii) $\mathbf{M}(\tau_0)$ is finite and (iii) Im $\mathbf{M}(\tau_0)$ is positive definite. The properties of the DRT system and of the ray-propagator matrix then guarantee that the 2×2 matrix $\mathbf{M}(\tau)$ satisfies the conditions (i)–(iii) at any point of the central ray. As Im $\mathbf{M}(\tau)$ 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 $\mathbf{M}(\tau_0)$ can also be chosen purely imaginary:

$$\mathbf{M}(\tau_0) = \mathrm{i}\mathrm{Im}\mathbf{M}(\tau_0). \quad (39)$$

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

4.1 Computation of complex-valued M(τ)

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

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

We consider a real-valued ray Ω, situated in an inhomogeneous anisotropic layered medium. We assume that the DRT in ray-centred coordinates q_1, q_2 has been performed along ray Ω and that the 4×4 ray propagator matrix $\mathbf{\Pi}(\tau, \tau_0)$ has been determined. The ray propagator matrix $\mathbf{\Pi}(\tau, \tau_0)$ and its 2×2 submatrices $\mathbf{Q}_1(\tau, \tau_0)$, $\mathbf{Q}_2(\tau, \tau_0)$, $\mathbf{P}_1(\tau, \tau_0)$ and $\mathbf{P}_2(\tau, \tau_0)$ are real-valued.

We now consider the complex-valued 2×2 matrix $\mathbf{M}(\tau_0)$, satisfying the Gaussian beam existence conditions. Eq. (37) then yields a solution, which has usually been called the Gaussian beam:

$$\mathbf{u}^{beam}(q_1, q_2, \tau) = \mathbf{U}^{\Omega}(\tau)(\det\mathbf{W})^{-1/2}\exp\left[-\frac{1}{2}\omega\mathbf{q}^T\mathrm{Im}\mathbf{M}(\tau)\mathbf{q}\right] \times \exp\left\{-\mathrm{i}\omega\left[t - T(\tau) - \frac{1}{2}\mathbf{q}^T\mathrm{Re}\mathbf{M}(\tau)\mathbf{q}\right]\right\}. \quad (40)$$

Here $\mathbf{q} = (q_1, q_2)^T$ and

$$\mathbf{W}(\tau, \tau_0) = \mathbf{Q}_1(\tau, \tau_0) + \mathbf{Q}_2(\tau, \tau_0)[\mathrm{Re}\mathbf{M}(\tau_0) + \mathrm{i}\mathrm{Im}\mathbf{M}(\tau_0)]. \quad (41)$$

The Gaussian beam existence conditions guarantee that $\mathbf{M}(\tau)$ is finite along the whole ray Ω. This guarantees that $\det \mathbf{Q}(\tau)$ 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 $\mathbf{Q}_1(\tau_0, \tau_0) = \mathbf{I}$, $\mathbf{Q}_2(\tau_0, \tau_0) = \mathbf{0}$, so that

$$\mathbf{W}(\tau_0, \tau_0) = \mathbf{I}. \tag{42}$$

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: (i) we put it zero for $\tau = \tau_0$; (ii) we require that it varies continuously along ray Ω.

There are no points along ray Ω, at which $\det \mathbf{W}(\tau, \tau_0)$ is zero. Thus, 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 thus removes this problem.

As the Gaussian beam cannot be singular at any point of the ray, it cannot be 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 Section 2.4.

Consider again a real-valued ray Ω. Each Gaussian beam connected with this ray is specified by a 2×2 symmetric complex-valued matrix $\mathbf{M}(\tau_0)$ given at an arbitrary point of the ray, $\tau = \tau_0$. Consequently, we can construct a six-parametric system of Gaussian beams connected with ray Ω. The three parameters $\mathrm{Re}\, M_{11}(\tau_0)$, $\mathrm{Re}\, M_{22}(\tau_0)$ and $\mathrm{Re}\, M_{12}(\tau_0)$, control the shape of the wave front of the Gaussian beam at τ_0. The other three parameters, $\mathrm{Im}\, M_{11}(\tau_0)$, $\mathrm{Im}\, M_{22}(\tau_0)$ and $\mathrm{Im}\, M_{12}(\tau_0)$, control the width of the Gaussian beam at τ_0. The real-valued traveltime along ray Ω and the spreading-free 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 (37), the ray-theory amplitudes in the expression (40) for Gaussian beams are the same in the whole plane tangent to the wave front at τ on Ω. The most efficient and accurate way of taking into account the paraxial changes of the ray-theory amplitudes 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 $\mathbf{\Pi}(\tau, \tau_0)$, we can determine the complex-valued matrix $\mathbf{M}(\tau)$ and the complex-valued factor $\{\det[\mathbf{Q}(\tau_0)/\det \mathbf{Q}(\tau)]\}^{1/2}$ by solving the DRT system with specific complex-valued initial conditions. These conditions can be given by relations analogous to (35):

$$\mathbf{Q}(\tau_0) = \mathbf{A}, \qquad \mathbf{P}(\tau_0) = \mathbf{M}_0\mathbf{A}. \tag{43}$$

Here, again $\mathbf{A}$ is an arbitrary constant real-valued finite 2×2 matrix, for which $\det \mathbf{A} \neq 0$. The matrix $\mathbf{M}_0$ is, however, complex-valued. The 2×2 matrices $\mathrm{Re}\, \mathbf{M}_0$ and $\mathrm{Im}\, \mathbf{M}_0$ are real valued, symmetric and finite. Moreover, $\mathrm{Im}\, \mathbf{M}_0$ is positive definite. Quite often $\mathrm{Re}\, \mathbf{M}_0 = \mathbf{0}$ is used. The Gaussian beam is again given by (40), where the 2×2 complex-valued matrix $\mathbf{M}(\tau)$ is given by the relation $\mathbf{M}(\tau) = \mathbf{P}(\tau)\mathbf{Q}^{-1}(\tau)$. The 2×2 matrices $\mathbf{P}(\tau)$ and $\mathbf{Q}(\tau)$ are also complex valued, and are obtained as solutions of the DRT system with the initial conditions (43). Finally, the complex-valued 2×2 matrix $\mathbf{W}(\tau, \tau_0)$ is given by the relation $\mathbf{W}(\tau, \tau_0) = \mathbf{Q}(\tau)\mathbf{Q}^{-1}(\tau_0)$.

4.2 Evaluation of a Gaussian beam at a specified paraxial point

Eqs (40) with (41) 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 wave front at a known point R_Ω on Ω. The position of point R_Ω on Ω is specified by the monotonic parameter τ and by $q_1 = q_2 = 0$. The ray-centred coordinates q_1, q_2 of point R are then determined in the plane tangent to the wave front 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 containing the observation point R is tangent to the wave front. 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 wave front 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 wave front at R_Ω.

Described cumbersome procedure could be avoided if the expression (40) 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 traveltime with respect to ray-centred coordinates to the 3×3 matrix of second derivatives of the traveltime with respect to Cartesian coordinates. For isotropic media, such a procedure was proposed by Klimeš (1984), see also Červený (2001; section 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, 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 substitute the matrix $\mathbf{M}$ in the quadratic expansion of the traveltime in (40) by $\mathbf{M}^{(x)}$, specified in Cartesian coordinates. The transformation from $M_{IJ} = \partial^2 T/\partial q_I \partial q_J$, 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 (33) given in Section 2.6.

Inserting (33) into (34), we obtain the quadratic expansion of the paraxial traveltime field in Cartesian coordinates at an arbitrary point R, situated in the vicinity of point R_Ω, chosen arbitrarily on ray Ω:

$$T(R) = T(R_\Omega) + [\mathbf{x}(R) - \mathbf{x}(R_\Omega)]^T \mathbf{p}(R_\Omega) + \frac{1}{2}[\mathbf{x}(R) - \mathbf{x}(R_\Omega)]^T \mathbf{M}^{(x)}(R_\Omega)[\mathbf{x}(R) - \mathbf{x}(R_\Omega)]. \tag{44}$$

The elements of the 3×3 matrix $\mathbf{M}^{(x)}(R_\Omega)$ are given by (33). The only complex-valued expression in (44) is the 2×2 matrix $\mathbf{M}(R_\Omega)$, which is included in $\mathbf{M}^{(x)}(R_\Omega)$, see (33).

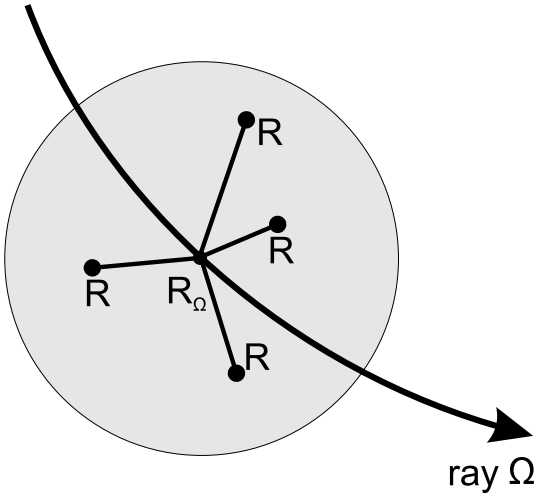

Figure 2. Section of the spherical vicinity of the point R_Ω, situated on the central ray Ω, in which quadratic expansion of the paraxial traveltime field (44) is valid.

The quadratic expansion of the traveltime 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 as shown in Fig. 2. 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 wave fronts 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:

$$\mathbf{u}^{beam}(R) = \mathbf{U}^{\Omega}(R_\Omega)(\det \mathbf{W})^{-1/2} \exp[-\omega \mathrm{Im} T(R)] \times \exp\{-\mathrm{i}\omega[t - \mathrm{Re} T(R)]\}. \quad (45)$$

Here $\mathbf{U}^{\Omega}$ and $\det \mathbf{W}$ have the same meaning as in equations (40) and (41). The complex-valued paraxial traveltime $T(R)$ is given by (44) with (33).

For completeness, we present here the expressions for Re $T(R)$ and Im $T(R)$:

$$\mathrm{Re} T(R) = T(R_\Omega) + \overline{p} + \frac{1}{2}\bar{\mathbf{f}}\,\mathrm{Re}\mathbf{M}(R_\Omega)\,\bar{\mathbf{f}}^T + \overline{p}\,\overline{\eta} + \frac{1}{2}(\boldsymbol{\mathcal{U}}^T\boldsymbol{\eta})\overline{p}^2, \quad (46)$$

$$\mathrm{Im} T(R) = \frac{1}{2}\bar{\mathbf{f}}\,\mathrm{Im}\mathbf{M}(R_\Omega)\,\bar{\mathbf{f}}^T. \quad (47)$$

Here

$$\overline{p} = [\mathbf{x}(R) - \mathbf{x}(R_\Omega)]^T \mathbf{p}(R_\Omega),$$
$$\overline{\eta} = [\mathbf{x}(R) - \mathbf{x}(R_\Omega)]^T \boldsymbol{\eta}(R_\Omega),$$
$$\bar{\mathbf{f}} = (\overline{f}_1, \overline{f}_2) = [\mathbf{x}(R) - \mathbf{x}(R_\Omega)]^T (\mathbf{f}_1, \mathbf{f}_2). \quad (48)$$

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 $\bar{\mathbf{f}}$ 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 $\mathbf{M}(R_\Omega)$, with elements $M_{IJ}(R_\Omega)$, is obtained by DRT in ray-centred coordinates.

In the described approach, we calculate the paraxial traveltime field (44) and the Gaussian beam (45) in Cartesian coordinates, although the DRT has been performed in ray-centred coordinates. We apply (33) 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 (45) with (46)–(48) 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 (45). Similarly as in eq. (40), we need to determine the 2×2 matrix $\mathbf{M}(R_\Omega)$ of the second derivatives of the complex-valued traveltime field with respect to ray-centred coordinates q_1, q_2 at point R_Ω. The 2×2 matrix $\mathbf{M}(R_\Omega)$ may be determined by solving the DRT system in ray-centred coordinates.

In fact, equations (40) with (41) are a special case of (45)–(48) and can be simply obtained from them. Consider the observation point R situated in the plane tangent to the wave front at R_Ω. We can then express the position of point R in ray-centred coordinates using the relation

$$\mathbf{x}(R) - \mathbf{x}(R_\Omega) = q_1\mathbf{e}_1 + q_2\mathbf{e}_2. \quad (49)$$

Using (48) for $\bar{\mathbf{f}}$, we obtain

$$\bar{\mathbf{f}} = (q_1\mathbf{e}_1 + q_2\mathbf{e}_2)^T(\mathbf{f}_1, \mathbf{f}_2) = (q_1, q_2) = \mathbf{q}^T. \quad (50)$$

From (50), we obtain the Gaussian beam exponential factor, which controls the amplitude profile of the beam, see (45) and (47):

$$\exp\left[-\frac{1}{2}\omega\bar{\mathbf{f}}\mathrm{Im}\mathbf{M}(R_\Omega)\bar{\mathbf{f}}^T\right] = \exp\left[-\frac{1}{2}\omega\mathbf{q}^T\mathrm{Im}\mathbf{M}(R_\Omega)\mathbf{q}\right]. \quad (51)$$

Now consider eq. (46) for Re $T(R)$ in the plane tangent to the wave front 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 (46) containing $\overline{p}$ vanish. Using eq. (50), term $\frac{1}{2}\bar{\mathbf{f}}\mathrm{Re}\mathbf{M}(R_\Omega)\bar{\mathbf{f}}^T$ can be transformed into $\frac{1}{2}\mathbf{q}^T\mathrm{Re}\mathbf{M}(R_\Omega)\mathbf{q}$. Thus, eqs (40)–(41) for the Gaussian beam in the plane tangent to the wave front at R_Ω, obtained in Section 4.1.1, follow also from eqs (45) to (48).

It is obvious that eqs (45)–(48) for Re $T(R)$ and Im $T(R)$ in Cartesian coordinates are more general and flexible than eqs (40)–(41) 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 eqs (45)–(48) are used. Eqs (45)–(48) 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 straightline profile intersecting the ray. The most convenient formulae for Gaussian beams are, however, obtained in the plane tangent to the wave front at any point R_Ω of ray Ω. In this plane, the Gaussian beam is fully described by eqs (40)–(41). As the ray-centred coordinates q_1, q_2 in this plane are actually 2-D Cartesian coordinates, we can simply use $\exp[-\frac{1}{2}\omega\mathbf{q}^T\mathrm{Im}\mathbf{M}(R_\Omega)\mathbf{q}]$ from (40) to study the properties of Gaussian beams in this plane.

The amplitudes of a Gaussian beam in the plane tangent to the wave front 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 $\frac{1}{2}\omega \mathbf{q}^T(R_\Omega)\mathrm{Im}\mathbf{M}(R_\Omega)\mathbf{q}(R_\Omega) = 1$ in the plane tangent to the wave front 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 wave front at R_Ω is fully controlled by the 2×2 matrix Im $\mathbf{M}(R_\Omega)$. As Im $\mathbf{M}(R_\Omega)$ 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 Im $\mathbf{M}(R_\Omega)$, we can also use the quantities $L_1(R_\Omega)$ and $L_2(R_\Omega)$, given by the relation

$$L_{1,2}(R_\Omega) = [\pi M_{1,2}^I(R_\Omega)]^{-1/2}. \tag{52}$$

Quantities $L_1(R_\Omega)$ and $L_2(R_\Omega)$ represent the half-axes of the spot ellipse in the plane tangent to the wave front at R_Ω for frequency $f = 1\mathrm{Hz}$ (i.e. $\omega = 2\pi$). We call them the half-widths of the Gaussian beam in the plane tangent to the wave front 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 2×2 matrix Re $\mathbf{M}(R_\Omega)$ describes the geometric properties of the wave front of the Gaussian beam. Since Re $\mathbf{M}(R_\Omega)$ is always symmetrical, its eigenvalues $M_1^R(R_\Omega)$ and $M_2^R(R_\Omega)$ are always real. Instead of Re $\mathbf{M}(R_\Omega)$, we can introduce the 2×2 matrix $\mathbf{K}(R_\Omega)$ of the curvature of the wave front at R_Ω on ray Ω by relation

$$\mathbf{K}(R_\Omega) = \mathcal{C}(R_\Omega)\mathrm{Re}\mathbf{M}(R_\Omega), \tag{53}$$

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 wave front of the Gaussian beam at point R_Ω on Ω.

The half-width of the Gaussian beam varies along ray Ω. We can determine these variations from the eqs (52) and (31). Consequently, the Gaussian beams may be narrow in some regions of the ray, but broad in other regions. Similarly, the curvature of the wave front of the Gaussian beam varies along ray Ω and may be determined using (53) and (31). For Re $\mathbf{M}(R_\Omega) = \mathbf{0}$, the wave front is locally planar at R_Ω.

5 CONCLUDING REMARKS

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 $a_{ijkl}(x_n)$, 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 $\mathbf{M}(\tau)$ of the second derivatives of the traveltime field with respect to ray-centred coordinates, with elements $\partial^2 T/\partial q_N \partial q_M$. This matrix can be computed in different ways. The most common is use of DRT in ray-centred coordinates q_1, q_2. Once the 2×2 matrix $\mathbf{M}(\tau)$ is known along the central ray Ω, the flexible and numerically efficient expressions (45)–(48) 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.

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

(ii) Further simplification can be obtained for weakly anisotropic media. This is true particularly for *P* waves. For example, approach proposed by Zhu *et al.* (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 could be constructed along so-called common rays. It may be useful to use first-order ray tracing (Pšenčík & Farra 2005; Farra & Pšenčík 2008).

(iii) For media with spatially varying elements of higher anisotropic symmetry, the ray tracing computations are quite cumbersome and even inaccurate. Iversen & 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.

(iv) 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) and Shearer & Chapman (1989).

(v) 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).

(vi) 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 generally 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 3-D. Only in the planes of symmetry of transversely isotropic and orthorhombic media such a situation may play an important role.

ACKNOWLEDGMENTS

The authors are greatly indebted to Luděk Klimeš for helpful suggestions and valuable discussions. Comments and recommendations of Sam Gray and an anonymous reviewer are highly appreciated. This research has been supported by the Consortium Project 'Seismic Waves in Complex 3-D Structures', by the Research Projects 205/08/0332 and 205/07/0032 of the Grant Agency of the Czech Republic, and by the Research Project MSM 0021620860 of the Ministry of Education of the Czech Republic.

REFERENCES

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

Babich, V.M., 1961. Ray method of the computation of the intensity of wave fronts in elastic inhomogeneous anisotropic medium, in *Problems of the Dynamic Theory of Propagation of Seismic Waves* (in Russian), Vol. 5, pp. 36–46, ed. Petrashen, G.I., 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, pp. 15–63, ed. Babich, V.M., Leningrad: Nauka.

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

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

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

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.

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. & Klimeš, L., 2009. Transformation relations for second derivatives of traveltime in anisotropic media, in *Seismic waves in complex 3-D structures*, Report 19, pp. 115–122. Charles University, Faculty of Mathematics and Physics, Dept. of Geophysics, Praha.

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

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

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

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

Červený, V., Klimeš, L. & Pšenčík, I., 2007. Seismic ray method: recent developments, *Adv. Geophys.,* **48,** 1–126.

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

Farra, V. & Pšenčík, I., 2008. First-order ray computations of coupled S waves in inhomogeneous weakly anisotropic media, *Geophys. J. Int.,* **173,** 979–989.

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

Gajewski, D. & 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. & Pšenčík, I., 1990. Vertical seismic profile synthetics by dynamic ray tracing in laterally varying layered anisotropic structures, *J. geophys. Res.,* **95,** 11 301–11 315.

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

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

Gray, S.H. & 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. & 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. &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. Geophys. Geod.,* **50,** 431–447.

Kravtsov, Yu.A., 1968. 'Quasiisotropic' approximation to geometrical optics, *Dokl. AN SSSR,* **183**(1), 74–77 (in Russian).

Kravtsov, Yu.A. & Berczynski, P., 2007. Gaussian beams in inhomogeneous media: a review, *Stud. Geophys. Geod.,* **51,** 1–36.

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

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

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

White, B., Norris, A., Bayliss, A. & 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. & Wang, D., 2005. Kinematic and dynamic ray tracing in anisotropic media: theory and applications, in *Proceedings of the 75th Annual Int. Meeting, SEG, Expanded Abstracts,* 96–99.

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

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

$$\mathcal{H}(x_i, p_j) = \frac{1}{2}[G_m(x_i, p_j) - 1], \tag{A1}$$

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

$$\frac{\mathrm{d}x_i}{\mathrm{d}\tau} = \frac{\partial \mathcal{H}}{\partial p_i}, \quad \frac{\mathrm{d}p_i}{\mathrm{d}\tau} = -\frac{\partial \mathcal{H}}{\partial x_i}. \tag{A2}$$

Here the parameter τ along the ray represents the traveltime. The right-hand sides of the ray tracing system (A2) 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

$$\begin{aligned} \frac{\partial \mathcal{H}}{\partial p_i} &= \mathcal{U}_i = a_{ijkl} p_l g_j^{(m)} g_k^{(m)}, \\ -\frac{\partial \mathcal{H}}{\partial x_i} &= \eta_i = -\frac{1}{2} \frac{\partial a_{jkln}}{\partial x_i} p_k p_n g_j^{(m)} g_l^{(m)}. \end{aligned} \tag{A3}$$

(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 or S2).

In eqs (A2), 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

$$\begin{aligned} \frac{\partial \mathcal{H}}{\partial p_i} &= \mathcal{U}_i = a_{ijkl} p_l D_{jk}/D_{ss}, \\ -\frac{\partial \mathcal{H}}{\partial x_i} &= \eta_i = -\frac{1}{2} \frac{\partial a_{jkln}}{\partial x_i} p_k p_n D_{jl}/D_{ss}. \end{aligned} \tag{A4}$$

Here

$$D_{ij} = \frac{1}{2} \epsilon_{ikl} \epsilon_{jrs} (\Gamma_{kr} - \delta_{kr})(\Gamma_{ls} - \delta_{ls}). \tag{A5}$$

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 expressions for $\partial \mathcal{H}/\partial x_i$ in ray tracing equations (A2) is as follows. Taking into account that $G_m(x_i, p_j)$ in (A1) 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

$$-\frac{\partial \mathcal{H}(x_n, p_j)}{\partial x_i} = -\frac{1}{2} \frac{\partial G_m(x_n, p_j)}{\partial x_i} = -\frac{1}{2} \mathcal{C}^{-2}(x_n, N_j) \frac{\partial \mathcal{C}^2(x_n, N_j)}{\partial x_i}. \tag{A6}$$

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

$$\frac{\partial \mathcal{H}}{\partial p_i} = \mathcal{U}_i(x_n, p_j), \quad -\frac{\partial \mathcal{H}}{\partial x_i} = -\mathcal{C}^{-1}(x_n, N_j) \frac{\partial \mathcal{C}(x_n, N_j)}{\partial x_i}. \tag{A7}$$

The ray tracing (A2) with (A7) was proposed by Zhu *et al.* (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 (A2) with (A7) corresponds to the system proposed by Pšenčík & Farra (2005) for *P*-wave ray tracing.

The initial conditions for the ray tracing system (A2) 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, $\mathcal{C}$. The phase velocities are obtained as a solution of a cubic equation resulting from the condition of solvability of the Christoffel equation

$$\det(a_{ijkl} N_j N_l - \mathcal{C} \delta_{ik}) = 0. \tag{A8}$$

Choosing the value of $\mathcal{C}$ corresponding to the studied wave, we can construct the slowness vector of the studied wave at the initial point of the ray, $\mathbf{p} = \mathbf{N}/\mathcal{C}$. The initial conditions for the ray-velocity vector $\boldsymbol{\mathcal{U}}$ need not be specified among initial conditions; they are obtained from initial conditions for $\mathbf{p}$.

Anisotropic ray tracing system (A2), with (A3) or (A4) or (A7), 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 & Chapman 1990; Farra & Pšenčík 2008).

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 (A2) with (A4), were first derived by Červený (1972). For more details on ray tracing in inhomogeneous anisotropic media, see Červený (2001, chapter 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 & Pšenčík 1987, 1990). The package ANRAY is freely available on the web pages of the SW3D Consortium (http://sw3d.mff.cuni.cz/).

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 in the vicinity of the point of incidence of an arbitrary high-frequency wave at a curved interface Σ 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 and S2) are generated at the point of incidence; some of them may be inhomogeneous. The initial slowness vector of any of reflected or transmitted waves is given by the relation

$$\tilde{\mathbf{p}} = \sigma \mathbf{n} + \mathbf{p}^{\Sigma}. \tag{A9}$$

Tilde indicates that the quantity corresponds to a generated wave. In (A9), $\mathbf{n}$ 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:

$$\det\left[a_{ijkl}\left(\sigma n_j + p_j^{\Sigma}\right)\left(\sigma n_l + p_l^{\Sigma}\right) - \delta_{ik}\right] = 0. \tag{A10}$$

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 half-space on the other side of the interface.

Eq. (A10) has six solutions for each half-space. 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 & Pšenčík (1987), Červený (2001, section 2.3.3).

APPENDIX B: DYNAMIC RAY TRACING IN INHOMOGENEOUS ANISOTROPIC MEDIA WITH INTERFACES

DRT consists in the solution of a system of linear ordinary differential equations of the first order along ray Ω. The system may be solved together with the ray tracing, or along an already known ray Ω. The DRT in ray-centred coordinates for inhomogeneous anisotropic media was studied by Hanyga (1982), Kendall *et al.* (1992), Klimeš (1994), Klimeš (2006), Bakker (1996) and Červený (2007). See also Červený (2001).

DRT system in ray-centred coordinates consists of four equations for $Q_I = \partial q_I/\partial\gamma$ and $P_I = \partial p_I^{(q)}/\partial\gamma$, with $I = 1, 2$. Symbol $p_I^{(q)}$ denotes the Ith ray-centred component of slowness vector. The DRT system reads

$$\mathrm{d}Q_I/\mathrm{d}\tau = A_{IJ}Q_J + B_{IJ}P_J, \qquad \mathrm{d}P_I/\mathrm{d}\tau = -C_{IJ}Q_J - D_{IJ}P_J. \tag{B1}$$

Here the 2×2 matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ are given by relations

$$\mathbf{A} = \mathbf{f}^T\hat{\mathbf{A}}^{(x)}\mathbf{e} + \mathbf{d}, \qquad \mathbf{B} = \mathbf{f}^T\hat{\mathbf{B}}^{(x)}\mathbf{f},$$
$$\mathbf{C} = \mathbf{e}^T(\hat{\mathbf{C}}^{(x)} - \boldsymbol{\eta}\boldsymbol{\eta}^T)\mathbf{e}, \quad \mathbf{D} = \mathbf{e}^T\hat{\mathbf{D}}^{(x)}\mathbf{f} + \mathbf{d}^T. \tag{B2}$$

The 3×3 matrices $\mathbf{A}^{(x)}$, $\mathbf{B}^{(x)}$, $\mathbf{C}^{(x)}$ and $\mathbf{D}^{(x)}$ are defined in terms of Cartesian derivatives of the Hamiltonian:

$$A_{ij}^{(x)} = \partial^2\mathcal{H}/\partial p_i\partial x_j, \quad B_{ij}^{(x)} = \partial^2\mathcal{H}/\partial p_i\partial p_j,$$
$$C_{ij}^{(x)} = \partial^2\mathcal{H}/\partial x_i\partial x_j, \quad D_{ij}^{(x)} = \partial^2\mathcal{H}/\partial x_i\partial p_j. \tag{B3}$$

Note that $D_{im}^{(x)} = A_{mi}^{(x)}$. The 3×2 matrix $\mathbf{e}$ is composed of two unit basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$, tangent to the wave front: $\mathbf{e} = (\mathbf{e}_1, \mathbf{e}_2)$. Using $\mathbf{e}_1$ and $\mathbf{e}_2$, we can also compute the 3×2 matrix $\mathbf{f} = (\mathbf{f}_1, \mathbf{f}_2)$ composed of two covariant basis vectors $\mathbf{f}_1$ and $\mathbf{f}_2$, perpendicular to the ray. Vectors $\mathbf{f}_1$ and $\mathbf{f}_2$ can be calculated from $\mathbf{e}_1$ and $\mathbf{e}_2$ using (14). Finally, $\mathbf{d} = \mathbf{f}^T\mathrm{d}\mathbf{e}/\mathrm{d}\tau$, where derivative $\mathrm{d}\mathbf{e}/\mathrm{d}\tau$ is given by (15).

The DRT system (B1) can also be written in matrix form:

$$\mathrm{d}\mathbf{Q}/\mathrm{d}\tau = \mathbf{A}\mathbf{Q} + \mathbf{B}\mathbf{P}, \quad \mathrm{d}\mathbf{P}/\mathrm{d}\tau = -\mathbf{C}\mathbf{Q} - \mathbf{D}\mathbf{P}. \tag{B4}$$

All matrices in (B4) are 2×2. The matrices $\mathbf{Q}$ and $\mathbf{P}$ are defined by relations $Q_{IJ} = \partial q_I/\partial\gamma_J$ and $P_{IJ} = \partial p_I^{(q)}/\partial\gamma_J$.

Numerically, the most time consuming procedure in DRT in ray-centred coordinates in anisotropic media is the computation of the 3×3 matrices $\mathbf{A}^{(x)} \equiv \mathbf{D}^{(x)T}$, $\mathbf{B}^{(x)}$ and $\mathbf{C}^{(x)}$, see (B3).

If a ray strikes an interface, the initial conditions for the DRT along central rays of reflected/transmitted waves at the relevant point of incidence must be introduced. They are realized in terms of the so-called interface propagator matrix $\mathbf{\Pi}(\tilde{\tau}^{\Sigma}, \tau^{\Sigma})$, see (C6) for its incorporation into the ray propagator matrix.

We consider again 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. As a special case of u_1, u_2 we can consider local Cartesian coordinates in a plane tangent to interface Σ at the point of incidence $\tau = \tau_\Sigma$ of central ray Ω on 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

$$\mathbf{x}(u_K) = \mathbf{x}_0 + \mathbf{g}_I u_I + \frac{1}{2}\mathbf{g}_{IJ}u_I u_J. \tag{B5}$$

Here $\mathbf{g}_I = \partial\mathbf{g}/\partial u_I$ are vectors tangent to Σ at $\mathbf{x} = \mathbf{x}_0$, and elements $\mathbf{g}_{IJ} = \partial^2\mathbf{g}/\partial u_I\partial u_J$ are related to the curvature of the interface Σ at $\mathbf{x} = \mathbf{x}_0$. In general, vectors $\mathbf{g}_I$ are neither unit nor mutually perpendicular. 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 4×4 interface propagator matrix $\mathbf{\Pi}(\tilde{\tau}_\Sigma, \tau_\Sigma)$ is given by the relation

$$\mathbf{\Pi}(\tilde{\tau}^{\Sigma}, \tau^{\Sigma}) = \begin{bmatrix} \tilde{\mathbf{K}}^T\mathbf{K}^{-T} & \mathbf{0} \\ \tilde{\mathbf{K}}^{-1}[\mathbf{E} - \tilde{\mathbf{E}} - (\sigma - \tilde{\sigma})\mathbf{D}]\mathbf{K}^{-T} & \tilde{\mathbf{K}}^{-1}\mathbf{K} \end{bmatrix}. \tag{B6}$$

For detailed derivation of (B6), see Červený & Moser (2007, section 6). In (B6), the symbols without a tilde correspond to the point of incidence, $\tau = \tau_\Sigma$, the symbols with a 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 a tilde is analogous.

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

$$\mathbf{K} = (\mathbf{g}_1, \mathbf{g}_2)^T(\mathbf{f}_1, \mathbf{f}_2), \qquad \mathbf{K}^{-1} = (\mathbf{e}_1, \mathbf{e}_2)^T(\mathbf{h}_1, \mathbf{h}_2) \tag{B7}$$

with

$$\mathbf{h}_1 = (\mathbf{g}_2 \times \boldsymbol{\mathcal{U}})/\boldsymbol{\mathcal{U}}^T(\mathbf{g}_1 \times \mathbf{g}_2), \qquad \mathbf{h}_2 = (\boldsymbol{\mathcal{U}} \times \mathbf{g}_1)/\boldsymbol{\mathcal{U}}^T(\mathbf{g}_1 \times \mathbf{g}_2). \tag{B8}$$

Vectors $\mathbf{e}_I$ and $\mathbf{f}_J$ are contravariant and covariant basis vectors of the ray-centred coordinate system, respectively, defined in Section 2.3.

Elements of the 2×2 inhomogeneity matrix $\mathbf{E}$ read

$$E_{IJ} = \left(\mathbf{g}_I^T\mathbf{p}\right)\left(\mathbf{e}_K^T\boldsymbol{\eta}\right)\left(\mathbf{g}_J^T\mathbf{f}_K\right) + \left(\mathbf{g}_I^T\boldsymbol{\eta}\right)\left(\mathbf{g}_J^T\mathbf{p}\right). \tag{B9}$$

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 $\mathbf{D}$ are given by the relation

$$D_{IJ} = \mathbf{g}_{IJ}^T\mathbf{n}. \tag{B10}$$

For a plane interface, $\mathbf{D} = \mathbf{0}$. Finally, scalars σ and $\tilde{\sigma}$ in (B6) are given by relations $\sigma = \mathbf{n}^T\mathbf{p}$ and $\tilde{\sigma} = \mathbf{n}^T\tilde{\mathbf{p}}$.

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 piecewise polynomial approximation using triangles (2-D) or tetrahedrons (3-D) 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; Žáček 2002 for numerical examples; see also Červený *et al.* 2007, section 6.2.1) should be extended to anisotropic media.

APPENDIX C: PROPERTIES OF RAY PROPAGATOR MATRIX $\mathbf{\Pi}(\tau, \tau_0)$

We consider a 4×4 ray propagator matrix $\mathbf{\Pi}(\tau, \tau_0)$ in ray-centred coordinates. We describe its important properties which can be conveniently used in the paraxial ray theory and in the theory of Gaussian beams.

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

$$\mathbf{\Pi}^T(\tau, \tau_0)\mathbf{J}\mathbf{\Pi}(\tau, \tau_0) = \mathbf{J}. \tag{C1}$$

Here $\mathbf{J}$ is the 4×4 matrix

$$\mathbf{J} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{pmatrix}. \tag{C2}$$

In (C2), $\mathbf{I}$ is the 2×2 identity matrix and $\mathbf{0}$ is the 2×2 null matrix. The symplecticity of $\mathbf{\Pi}(\tau, \tau_0)$ has several important and useful consequences.

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

$$\det \mathbf{\Pi}(\tau, \tau_0) = 1. \tag{C3}$$

Eq. (C3) is satisfied for any τ. Consequently, $\mathbf{\Pi}(\tau, \tau_0)$ is regular along the whole ray Ω.

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

$$\mathbf{\Pi}(\tau, \tau_0) = \mathbf{\Pi}(\tau, \tau_1)\mathbf{\Pi}(\tau_1, \tau_0). \tag{C4}$$

The point corresponding to τ_1 may be any point of ray Ω, not necessarily between τ_0 and τ. The chain rule (C4) can be extended to an arbitrary number of points $\tau_1, \tau_2, \tau_3, \ldots \tau_n$ along the ray Ω.

(3) 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

$$\mathbf{\Pi}^{-1}(\tau, \tau_0) = \begin{bmatrix} \mathbf{P}_2^T(\tau, \tau_0) & -\mathbf{Q}_2^T(\tau, \tau_0) \\ -\mathbf{P}_1^T(\tau, \tau_0) & \mathbf{Q}_1^T(\tau, \tau_0) \end{bmatrix}. \tag{C5}$$

(4) Matrix $\mathbf{\Pi}(\tau, \tau_0)$ can be constructed even if ray Ω is reflected/transmitted at a structural interface between τ_0 and τ. Let us introduce the traveltimes τ^Σ and $\tilde{\tau}^\Sigma$ corresponding to the point of incidence and to the point of reflection/transmission, respectively. 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 $\mathbf{\Pi}(\tau, \tau_0)$ then reads:

$$\mathbf{\Pi}(\tau, \tau_0) = \mathbf{\Pi}(\tau, \tilde{\tau}^\Sigma)\mathbf{\Pi}(\tilde{\tau}^\Sigma, \tau^\Sigma)\mathbf{\Pi}(\tau^\Sigma, \tau_0). \tag{C6}$$

The 2×2 matrix $\mathbf{\Pi}(\tilde{\tau}^\Sigma, \tau^\Sigma)$ is the 'interface propagator matrix', see (B6). Thus, 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.

APPENDIX D: BRIEF DESCRIPTION OF COMPUTATIONAL PROCEDURE OF A GAUSSIAN BEAM

D1 Model

First of all, a model of a layered medium with smoothly varying interfaces and smoothly varying density-normalized elastic moduli in layers between interfaces must be prepared. Interfaces and elastic parameters are usually specified at gridpoints of rectangular grids. Since computation of Gaussian beams requires solution of DRT equations, the second derivatives of the density-normalized elastic moduli and of the functions approximating interfaces must be continuous. Convenient are, for example, approximation techniques like cubic, bicubic and/or tricubic splines with tension. Remember, linear, bilinear or trilinear approximation of density-normalized elastic moduli introduces artificial interfaces of the second-order (interfaces, on which gradients of density-normalized elastic moduli vary discontinuously), and DRT then requires specification of boundary conditions.

D2 Ray-centred coordinates

Since we consider DRT in ray-centred coordinates in this paper, it is necessary to determine the contravariant basis vectors $\mathbf{e}_i$, specifying these coordinates, along the central ray. Their determination as well as the determination of covariant vector basis $\mathbf{f}_i$, is described in Section 2.3.

D3 Ray tracing and DRT

For tracing rays, eq. (A2) with (A3), (A4) or (A7) can be used. Specification of initial conditions at a point-source and after reflection/transmission is described in Appendix A. For DRT, eqs (B4) can be used. The initial conditions for the DRT system can be specified in two ways. If we wish to use the ray propagator matrix for the computation of Gaussian beams, the initial conditions are specified by the identity matrix, see Sections 2.5 and 4.1.1. The other option is to use the specific complex-valued initial conditions described in Section 4.1.2 They include the specification of the 2×2 matrix $\mathbf{M}(\tau_0)$ at an 'initial' point of the ray, τ_0. At the point of incidence at an interface, the DRT equations must be transformed according to the rules described in the Appendix B.

D4 Evaluation of a Gaussian beam

For this purpose, either eq. (40) or (45) can be used. In both, the most important role is played by the matrix of second derivatives of traveltimes with respect to ray-centred coordinates, $\mathbf{M}(\tau)$. If we use the ray propagator matrix, we can use (31) to determine $\mathbf{M}(\tau)$ from $\mathbf{M}(\tau_0)$ specified at τ_0. If we solve DRT with specific complex-valued initial conditions containing $\mathbf{M}_0$, see (43), we can determine $\mathbf{M}(\tau)$ using the relation $\mathbf{M}(\tau) = \mathbf{P}(\tau)\mathbf{Q}^{-1}(\tau)$. Eq. (40) is expressed in ray-centred coordinates. More convenient, especially for evaluation of Gaussian beams at observation points specified in Cartesian coordinates, is eq. (45), see detailed discussion in Section 4.2. All quantities appearing in eqs (40) and (45) can be evaluated from quantities obtained by ray tracing and dynamic ray tracing.