

Higher-order Hamilton–Jacobi perturbation theory for anisotropic heterogeneous media: dynamic ray tracing in Cartesian coordinates

Einar Iversen,¹ Bjørn Ursin,² Teemu Saksala,³ Joonas Ilmavirta⁴ and Maarten V. de Hoop³

¹Department of Earth Science, University of Bergen, P.O. Box 7803, N-5020 Bergen, Norway. E-mail: enar.iversen@uib.no

²Department of Geoscience and Petroleum, Norwegian University of Science and Technology (NTNU), S.P. Andersens vei 15A, NO-7491 Trondheim, Norway

³Department of Computational and Applied Mathematics, Rice University, 6100 Main MS-134, Houston, TX 77005-1892, USA

⁴Department of Mathematics and Statistics, University of Jyväskylä, P.O. Box 35 (MaD), FI-40014 University of Jyväskylä, Finland

Accepted 2018 December 19. Received 2018 December 7; in original form 2018 October 01

SUMMARY

With a Hamilton–Jacobi equation in Cartesian coordinates as a starting point, it is common to use a system of ordinary differential equations describing the continuation of first-order derivatives of phase-space perturbations along a reference ray. Such derivatives can be exploited for calculating geometrical spreading on the reference ray and for establishing a framework for second-order extrapolation of traveltimes to points outside the reference ray. The continuation of first-order derivatives of phase-space perturbations has historically been referred to as dynamic ray tracing. The reason for this is its importance in the process of calculating amplitudes along the reference ray. We extend the standard dynamic ray-tracing scheme to include higher-order derivatives of the phase-space perturbations. The main motivation is to extrapolate and interpolate amplitude and phase properties of high-frequency Green’s functions to nearby (paraxial) source and receiver locations. Principal amplitude coefficients, geometrical spreading factors, geometrical spreading matrices, ray propagator matrices, traveltimes, slowness vectors and curvature matrices are examples of quantities for which we enhance the computation potential. This, in turn, has immediate applications in modelling, mapping and imaging. Numerical tests for 3-D isotropic and anisotropic heterogeneous models yield clearly improved extrapolation results for the traveltimes and geometrical spreading. One important conclusion is that the extrapolation function for the geometrical spreading must be at least third order to be appropriate at large distances away from the reference ray.

Key words: Body waves; Computational seismology, Seismic anisotropy; Numerical approximations and analysis; Numerical modelling; Wave propagation.

1 INTRODUCTION

We consider a higher-order Hamilton–Jacobi perturbation theory for anisotropic heterogeneous media. This theory arises from the differentiation of the Hamilton system for ray tracing in phase space (Hamilton 1837). Specifically, we discuss the higher-order perturbations of a Hamiltonian flow with respect to its initial conditions in the phase space.

The resulting perturbation coefficients can be used for higher-order extrapolation or interpolation of important quantities related to the amplitude and phase of the high-frequency Green’s function: traveltimes, geometrical spreading, amplitude coefficients and polarization directions. The methodology has immediate applications in contexts where high-frequency Green’s functions are used extensively, for example, in modelling, mapping and imaging.

The leading-order perturbation yields the linearized or first-order Hamilton–Jacobi perturbation system, the integration of which is commonly used, for example, to construct the geometrical spreading. In solid earth geophysics this process is known as dynamic ray tracing. We focus on the integration of higher-order Hamilton–Jacobi perturbation equations—using point-source and local plane-wave initial conditions—in Cartesian phase-space coordinates.

Ray perturbation has been studied for decades commonly from a paraxial point of view (Červený 1972; Farra & Madariaga 1987; Červený *et al.* 1988; Bortfeld 1989; Gajewski & Pšenčík 1990; Hubral *et al.* 1992; Klimeš 1994; Červený 2001; Chapman 2004; Iversen 2004a; Moser & Červený 2007; Červený & Moser 2007; Iversen & Pšenčík 2008; Červený & Pšenčík 2010). We note that the leading-order perturbation of the traveltimes with respect to the source and receiver location requires only the ray propagator associated with the linearized

or first-order Hamilton–Jacobi system (Červený *et al.* 1984, 2012). Perturbation can not only be viewed as a local extrapolation but also be exploited for an interpolation with derivatives.

Throughout this paper any perturbation of the Hamiltonian flow will be assumed due to a perturbation of the initial conditions, or *ray parameters*, belonging to a given reference ray. This is in contrast to the ray perturbation arising because parameters of the (elastic) model are perturbed. In our case, the model is considered *fixed*.

Our main motivation is *extrapolating* not only *traveltime* but the full *ray propagator* away from the reference ray, that is, the geometry information for any neighbouring (paraxial) ray obtained by a perturbation in the phase space. For these paraxial rays the source point may be different from the source point of the reference ray. As a special case it is then possible to compute also *geometrical spreading* for general perturbations of the source and receiver location. In addition, the integration of the higher-order Hamilton–Jacobi perturbation equations opens for more accurate extrapolation of principal amplitude coefficients.

In fact, for higher-order derivatives of traveltime and amplitude only, under the assumption of a caustic-free two-parametric (orthonomic) system of rays, one does not need to consider the higher-order Hamilton–Jacobi perturbation equations (Červený 2001; Klimeš 2002a; Goldin & Duchkov 2003; Klimeš 2006a). Indeed, the higher-order derivatives of the traveltime can be obtained recursively by a number of closed-form integrations (quadratures) along a reference ray in an isotropic medium (e.g. Červený, 2001) and in an anisotropic medium (e.g. Klimeš 2002a). The procedure in Klimeš (2002a) included also the effects of perturbing the model. For isotropic media, Goldin & Duchkov (2003) took into account second-order spatial derivatives of amplitude, in an effort to make the recursive integration scheme applicable in the vicinity of caustics. A further development to the anisotropic media was proposed by Klimeš (2006a), comprising higher-order spatial derivatives of amplitude and higher-order model-perturbation derivatives.

The novel methodology introduced in this paper has additional qualities compared to the approaches proposed above:

(i) The methodology yields the possibility of extrapolating the *entire paraxial system*, accommodated by the ray propagator, to any paraxial ray. In this way, we can establish on a paraxial ray exactly the same information as is provided by conventional ‘complete’ dynamic ray tracing along the reference ray (Červený *et al.* 1988).

(ii) If one is primarily interested in computing the geometrical spreading, as is typically the case in seismic imaging, it adds an unnecessary complication to do this via a differentiation of the transport equation for amplitude. In our case, all computations are based on a Hamilton–Jacobi equation for stationary time (the Eikonal equation). This yields great advantages in many respects, for example, with respect to caustics—as the conventional dynamic ray-tracing equations and their higher-order extensions are all safely integrated.

(iii) Our methodology is fairly easy to implement, as we simply add extra sets of ordinary differential equations or closed-form integrals to those of the conventional approach. We remark, though, that for computation of paraxial traveltimes and amplitudes from a single source point in a caustic-free medium, the approach described in Klimeš (2002a, 2006a) is potentially faster, as the number of equations (quadratures) is smaller.

(iv) We provide in this paper numerical examples quantifying the errors involved in the higher-order extrapolation of the traveltime and geometrical spreading based on a single reference ray in an isotropic or anisotropic medium. It is clearly demonstrated that one achieves appropriate accuracy for geometrical spreading only if the extrapolation function is at least third order in the spatial coordinates.

The developed methodology has the following main applications:

(i) Fast computation of high-frequency elastic-wave Green’s functions corresponding to general paraxial rays, through (Hermite or spline) interpolation and extrapolation of amplitude and phase with derivatives. Our procedure holds in generally anisotropic media, leading to systems of equations describing the propagation of elastic waves, of principal type.

(ii) Fast generalized Radon transform inversion, where the amplitude and traveltime of the rays from the image point to the sources and receivers can be extrapolated from two reference rays (Beylkin & Burridge 1990; de Hoop *et al.* 1994; de Hoop & Bleistein 1997; Bleistein *et al.* 2001; Stolk & de Hoop 2002; Brandsberg-Dahl *et al.* 2003a,b; Sollid & Ursin 2003; Ursin 2004; Foss & Ursin 2004; Foss *et al.* 2004, 2005).

(iii) Extrapolation from the reference rays of map depth migration (Iversen & Gjøystdal 1996; Douma & de Hoop 2006), with the assumption that the scattering is from interfaces. Asymptotically, one only needs a narrow fan of rays illuminating a reflector.

(iv) True-amplitude time migration, that is, migration in image-ray coordinates and restricted-angle transform through extrapolation; here, the reference rays are image rays. This is considered a further development of earlier work on true-amplitude depth migration and time-to-depth mapping (Hubral 1977, 1983; Schleicher *et al.* 2007; Iversen *et al.* 2012; Tygel *et al.* 2012).

We make the observation that true-amplitude time migration formulated in this way explicitly shows that the relevant quantities can be obtained from the generalized Dix procedure for the reconstruction of a Riemannian metric in ray-centred coordinates or Fermi coordinates (Cameron *et al.* 2007; Iversen & Tygel 2008; de Hoop *et al.* 2014, 2015).

(v) Ray-based extended depth imaging through extrapolation (Stolk & de Hoop 2006; de Hoop *et al.* 2009). Here, the reference rays are the ones for map migration.

(vi) Map migration and depth imaging based on isochron rays (Iversen 2004b; Duchkov & de Hoop 2010).

(vii) Source–receiver continuation and characterization of the range of the single scattering operator. Here, extrapolation provides the local flow along characteristic strips (de Hoop & Uhlmann 2006).

(viii) Common-reflection-surface (CRS) processing techniques (e.g. Rabbel *et al.* 1991; Jäger *et al.* 2001). These techniques utilize that coherent local reflection events in the recorded data constitute a (hyper)surface, typically given in source–receiver coordinates or midpoint–offset coordinates (Ursin 1982). The CRS time surface is conventionally considered to be a second-order approximation. When higher-order coefficients are available from ray theory one could consider an extension also of the CRS techniques to higher orders.

The paper is organized as follows. First, we describe the basic concepts of the Hamilton–Jacobi theory, followed by a review of conventional dynamic ray tracing in Cartesian coordinates. Then, we introduce higher-order Hamilton–Jacobi phase-space perturbation equations and the constraint relations pertaining to them. We also specify initial conditions for the point-source and plane-wave situations. Assuming that the necessary phase-space perturbation data have been computed, we formulate approaches for the higher-order paraxial extrapolation of the traveltime and geometrical spreading. The Hamiltonian is mostly treated as a ‘black box’ with certain fundamental properties. As we see it, this widens the number of applications where the theory can be used. One section is however devoted to specific Hamiltonians. After the theory sections we show numerical examples for three related 3-D heterogeneous models. We also discuss briefly the connections to differential geometry. For an overview of the main mathematical symbols used in the paper, see Table 1.

2 HAMILTON–JACOBI EQUATION IN CARTESIAN PHASE SPACE

Consider a Cartesian coordinate system with the position vector $\mathbf{x} = (x_i)$ and slowness (momentum) vector $\mathbf{p} = (p_i)$. We form the phase space $\mathbf{w} = (w_r) = (x_i, p_i)$, where all six components vary freely. In the phase space (w_r) we further consider a reference ray Ω given as a function of the time τ , so that

$$w_r = \hat{w}_r(\tau). \quad (1)$$

Eq. (1) can be associated with a Hamilton–Jacobi equation for stationary time,

$$\mathcal{H}(\mathbf{w}) = \hat{\mathcal{H}}, \quad (2)$$

where the function $\mathcal{H}(\mathbf{w})$ is referred to as the Hamiltonian, and $\hat{\mathcal{H}}$ is a nonzero constant. The Hamilton–Jacobi eq. (2) is a nonlinear first-order partial differential equation for the time τ along Ω —in the context of wave propagation it is also often called the Eikonal equation.

One can interpret eq. (2) to represent a hypersurface (manifold) in phase space with five degrees of freedom. This hypersurface is typically not available as a specific, exact, function; rather, it will be known through derivatives evaluated up to a certain order with respect to phase-space coordinates at points on Ω .

We assume that \mathcal{H} is a homogeneous function of degree two in the slowness components, p_i . Then, Euler’s theorem for homogeneous functions yields

$$p_i \frac{\partial \mathcal{H}}{\partial p_i} = 2\mathcal{H}. \quad (3)$$

The specific formulation chosen for the Hamiltonian will determine what will be the independent variable along rays. For this variable to be the time τ , the Hamiltonian must satisfy

$$p_i \frac{\partial \mathcal{H}}{\partial p_i} = 1. \quad (4)$$

In view of eq. (3) the constant in eq. (2) is therefore $\hat{\mathcal{H}} = 1/2$.

The total temporal derivatives of position and momentum vectors can be computed using Hamilton’s equations,

$$\frac{dx_i}{d\tau} = \frac{\partial \mathcal{H}}{\partial p_i}, \quad \frac{dp_i}{d\tau} = -\frac{\partial \mathcal{H}}{\partial x_i}. \quad (5)$$

Integration of the ordinary differential equations (ODEs) in eq. (5) yields the solution functions $\hat{x}_i(\tau)$ and $\hat{p}_i(\tau)$ on Ω , as well as the time derivative of these functions,

$$v_i(\tau) = \frac{d\hat{x}_i}{d\tau}(\tau), \quad \eta_i(\tau) = \frac{d\hat{p}_i}{d\tau}(\tau). \quad (6)$$

We note that $\mathbf{v} = (v_i)$ signifies the ray-velocity (group-velocity) vector, while the time derivative of the slowness vector, $\boldsymbol{\eta} = (\eta_i)$, is referred to as just the eta vector. The fundamental requirement in eq. (4) and the first subequations of eqs (5) and (6) show that the slowness vector and ray-velocity vector must satisfy

$$p_i v_i = 1 \quad (7)$$

along the ray Ω .

Hamilton’s equations may alternatively be formulated compactly as

$$\frac{dw_r}{d\tau} = J_{rs} \frac{\partial \mathcal{H}}{\partial w_s}, \quad (8)$$

Table 1. Main mathematical symbols used in the paper. For multicomponent quantities the dimensions are specified.

Quantity	Dimension	Description
(x_1, x_2, x_3)	3	Cartesian coordinate system
$\mathbf{x} = (x_i)$	3	Position vector of the Cartesian coordinate system
$\mathbf{p} = (p_i)$	3	Slowness vector (momentum vector) of the Cartesian coordinate system
$\mathbf{w} = (w_r)$ $= (x_i, p_j)$	6	Phase-space vector of the Cartesian coordinate system
Ω		Reference ray
$\mathcal{H}(\mathbf{w})$		Hamiltonian
$\hat{\mathcal{H}}$		Constant value of the Hamiltonian
\mathcal{N}		Degree of the Hamiltonian
τ		Traveltime along the ray Ω
τ_0		Traveltime at the initial point of the ray Ω
c		Phase velocity
$\mathbf{v} = (v_i)$	3	Ray-velocity (group-velocity) vector
$\boldsymbol{\eta} = (\eta_i)$	3	Derivative of slowness vector \mathbf{p} with respect to traveltime τ
N_γ		Number of parameters specifying a perturbation of the initial phase-space location. Possible values are from 1 to 6.
(γ_a)	N_γ	Parameters specifying a perturbation of the initial phase-space location
$\mathbf{X} = \{X_{ra}\}$	$6 \times N_\gamma$	First-order derivatives of phase-space perturbations
$\mathbf{S} = \{S_{rs}\}$	6×6	ODE coefficients related to first-order derivatives of phase-space perturbations
$\mathbf{U} = \{U_{ij}\}$	3×3	Subset (submatrix) of the ODE coefficients $\{S_{rs}\}$
$\mathbf{V} = \{V_{ij}\}$	3×3	Subset (submatrix) of the ODE coefficients $\{S_{rs}\}$, the wave-propagation metric tensor
$\mathbf{W} = \{W_{ij}\}$	3×3	Subset (submatrix) of the ODE coefficients $\{S_{rs}\}$
$\delta\mathbf{w} = (\delta w_r)$	6	Perturbation of the phase-space vector
$\delta\mathbf{w}_0 = (\delta w_r)_0$	6	Perturbation of the phase-space vector at the initial point on Ω
$\mathbf{\Pi}(\tau, \tau_0)$ $= \{\Pi_{rs}(\tau, \tau_0)\}$	6×6	Ray propagator matrix
Π		Paraxial plane
$\mathcal{E} = \{\mathcal{E}_{iM}\}$ $= [\mathbf{e}_1, \mathbf{e}_2]$	3×2	Basis vectors in the plane Π
$\mathbf{H} = \{H_{ij}\}$ $= [\mathcal{E} \mathbf{v}]$	3×3	Transformation matrix related to the plane Π
$\mathcal{F} = \{\mathcal{F}_{iM}\}$	3×2	Submatrix of the 3×3 matrix \mathbf{H}^{-T} .
$\{\alpha_{ij}\}$	3×3	Projection operator with respect to the wave-propagation metric tensor
$\mathbf{M} = \{M_{ij}\}$	3×3	Second-order derivatives of traveltime with respect to Cartesian coordinates, on Ω
$\{M_{ijk}\}$	$3 \times 3 \times 3$	Third-order derivatives of traveltime with respect to Cartesian coordinates, on Ω
$\{M_{ijkl}\}$	$3 \times 3 \times 3 \times 3$	Fourth-order derivatives of traveltime with respect to Cartesian coordinates, on Ω
$\mathbf{Q} = \{Q_{ia}\}, \mathbf{P} = \{P_{ia}\}$	$3 \times N_\gamma$	First-order derivatives of phase-space perturbations, in Q – P notation
$\mathbf{Q} = \{Q_{iA}\}, \mathbf{P} = \{P_{iA}\}$	3×2	First-order derivatives of phase-space perturbations, in Q – P notation, for the case $N_\gamma = 2$
$\hat{\mathbf{Q}} = [\mathbf{Q} \mathbf{v}]$	3×3	Extension of 3×2 matrix \mathbf{Q} to size 3×3 , the geometrical spreading matrix
$\hat{\mathbf{P}} = [\mathbf{P} \boldsymbol{\eta}]$	3×3	Extension of 3×2 matrix \mathbf{P} to size 3×3
$\hat{\mathbf{Q}}^\dagger = \{Q_{ij}^\dagger\} = \hat{\mathbf{Q}}^{-1}$	3×3	Inverse geometrical spreading matrix
$\mathbf{Q}^\dagger = \{Q_{Ij}^\dagger\}$	2×3	Submatrix of the inverse geometrical spreading matrix
$\{X_{rab}\}$	$6 \times N_\gamma \times N_\gamma$	Second-order derivatives of phase-space perturbations
$\{Q_{iab}\}, \{P_{iab}\}$	$3 \times N_\gamma \times N_\gamma$	Second-order derivatives of phase-space perturbations, in Q – P notation
$\{S_{rst}\}$	$6 \times 6 \times 6$	Main ODE coefficients related to second-order derivatives of phase-space perturbations
$\{R_{rst}\}$	$6 \times 6 \times 6$	Additional ODE coefficients related to second-order derivatives of phase-space perturbations
$\{U_{ijk}\}$	$3 \times 3 \times 3$	Subset of the ODE coefficients $\{S_{rst}\}$
$\{V_{ijk}\}$	$3 \times 3 \times 3$	Subset of the ODE coefficients $\{S_{rst}\}$
$\{X_{rabc}\}$	$6 \times N_\gamma \times N_\gamma \times N_\gamma$	Third-order derivatives of phase-space perturbations
$\{Q_{iabc}\}, \{P_{iabc}\}$	$3 \times N_\gamma \times N_\gamma \times N_\gamma$	Third-order derivatives of phase-space perturbations, in Q – P notation

Table 1. Continued

Quantity	Dimension	Description
$\{S_{rstu}\}$	$6 \times 6 \times 6 \times 6$	Main ODE coefficients related to third-order derivatives of phase-space perturbations
$\{R_{rstu}\}$	$6 \times 6 \times 6 \times 6$	Additional ODE coefficients related to third-order derivatives of phase-space perturbations
$\mathbf{s} = (s_i)$	3	Source point
$\mathbf{r} = (r_i)$	3	Receiver point
$T(\mathbf{r}, \mathbf{s})$		Traveltime as a function of source–receiver coordinates
$\mathcal{L}(\mathbf{r}, \mathbf{s})$		Relative geometrical spreading as a function of source–receiver coordinates
$\{a_{ijkl}\}$	$3 \times 3 \times 3 \times 3$	Density-normalized elastic moduli
$\mathbf{\Gamma} = \{\Gamma_{ij}\}$	3×3	Christoffel matrix
G		Eigenvalue of the Christoffel matrix
P, Q, R		General invariants of the Christoffel matrix, for arbitrarily anisotropic media
P^{PSV}, R^{PSV}, G^{SH}		Particular invariants of the Christoffel matrix, for transversely isotropic media
G^P, G^{SV}		Particular invariants of the Christoffel matrix, for elliptically anisotropic media
$\mathbf{0}$	*	Zero multicomponent quantity. The dimensions follow from the context.
\mathbf{I}	3×3	Identity matrix
\mathbf{J}	6×6	Matrix for rearranging derivatives in Hamilton's equations

where J_{rs} are components of the 6×6 matrix

$$\mathbf{J} = \{J_{rs}\} = \begin{pmatrix} \{0_{ij}\} & \{\delta_{ij}\} \\ -\{\delta_{ij}\} & \{0_{ij}\} \end{pmatrix}. \quad (9)$$

The right-hand side of eq. (8) is the Hamiltonian vector field corresponding to the Hamiltonian \mathcal{H} .

Differentiation of eq. (3) with respect to p_i yields the important relation

$$\frac{\partial^2 \mathcal{H}}{\partial p_i \partial p_j} p_j = \frac{\partial \mathcal{H}}{\partial p_i}, \quad (10)$$

which holds for general locations in phase space. For other useful expressions involving derivatives of the Hamiltonian, see Appendix A.

On the ray Ω eq. (10) is recast to

$$V_{ij}(\tau) \hat{p}_j(\tau) = v_i(\tau), \quad (11)$$

where

$$V_{ij}(\tau) = \frac{\partial^2 \mathcal{H}}{\partial p_i \partial p_j}(\hat{\mathbf{w}}(\tau)). \quad (12)$$

In physics, the quantity $\{V_{ij}\}$ is often referred to as the wave-propagation metric tensor (Klimeš 2002b). In Riemannian geometry (e.g. Bao *et al.*, 2012), the hypersurface (2) is approximated using partial derivatives of slowness components p_k up to order two taken on the ray Ω . As a consequence, the second-order derivatives $\partial^2 \mathcal{H} / \partial p_i \partial p_j$ are considered invariant with respect to p_k .

3 CONVENTIONAL DYNAMIC RAY TRACING

As an introduction to higher-order Hamilton–Jacobi perturbation equations, we summarize the basics of conventional dynamic ray tracing.

3.1 Perturbations in phase space

Consider again a reference ray Ω with phase-space locations $(\hat{w}_r(\tau))$ consistent with eq. (2). A perturbed phase-space location is then generally expressed as

$$w_r = \hat{w}_r(\tau) + \delta w_r, \quad (13)$$

where all six perturbation components δw_r may vary freely. It is common to write the perturbed phase-space location as a vectorial function, with components $w_r = w_r(\boldsymbol{\gamma}, \tau)$. Here, the vector $\boldsymbol{\gamma} = (\gamma_a)$ has dimension N_γ and serves to parametrize a perturbation of the phase-space location corresponding to the initial point on Ω , for which $\tau = \tau_0$. The symbol $\hat{\boldsymbol{\gamma}}$ signifies no perturbation of this initial phase-space location. We require (i) that the variables γ_a are mutually independent, (ii) that none of them depend on the time, τ , and (iii) that none of them depend on the model of the medium. It follows that the dimension, N_γ , of the vector $\boldsymbol{\gamma}$ must have the maximum value $N_\gamma = 6$, that is, the dimension of the Cartesian phase space.

3.2 System of Hamilton–Jacobi perturbation equations

A system for *dynamic ray tracing* in Cartesian coordinates (x_i) can be derived by inserting eq. (13) on the left-hand side of eq. (8) followed by partial differentiation with respect to the variable γ_a . Since τ and γ_a are independent variables, the differentiations $d/d\tau$ and $\partial/\partial\gamma_a$ commute (Červený 2001; section 4.2.1). We obtain the system of ODEs

$$\frac{dX_{ra}}{d\tau}(\tau) = S_{rt}(\tau) X_{ta}(\tau); \quad \frac{d\mathbf{X}}{d\tau}(\tau) = \mathbf{S}(\tau) \mathbf{X}(\tau), \quad (14)$$

where $X_{ra}(\tau)$ can be equivalently defined by the partial derivatives

$$X_{ra}(\tau) = \frac{\partial(\delta w_r)}{\partial\gamma_a}(\hat{\boldsymbol{\gamma}}, \tau) = \frac{\partial w_r}{\partial\gamma_a}(\hat{\boldsymbol{\gamma}}, \tau), \quad (15)$$

and $S_{rt}(\tau)$ is formed by the second partial derivatives of the Hamiltonian,

$$S_{rt}(\tau) = J_{rs} \frac{\partial^2 \mathcal{H}}{\partial w_s \partial w_t}[\mathbf{w}(\hat{\boldsymbol{\gamma}}, \tau)]. \quad (16)$$

The quantity X_{ra} in eq. (15) is a first-order derivative of a perturbation in phase space related to a point with time τ on the reference ray Ω . For clarity of notation, we prefer mostly to write such derivatives as in the last expression of eq. (15), that is, without the perturbation (δ) symbol. It is emphasized that the derivative $\partial w_r/\partial\gamma_a$ belongs to a fixed value of the time τ .

Conventional dynamic ray tracing in Cartesian coordinates yields as output the $6 \times N_\gamma$ matrix $\mathbf{X}(\tau) = \{X_{ra}(\tau)\}$, with components of the form given in eq. (15). The matrix function $\mathbf{X}(\tau)$ is continued along the ray Ω by solving the system of ODEs in eq. (14) with suitable initial conditions. It is common to split the matrix $\{X_{ra}\}$ into $3 \times N_\gamma$ sub-matrices $\{Q_{ma}\}$ and $\{P_{ma}\}$, such that

$$Q_{ma}(\tau) = \frac{\partial x_m}{\partial\gamma_a}(\hat{\boldsymbol{\gamma}}, \tau), \quad P_{ma}(\tau) = \frac{\partial p_m}{\partial\gamma_a}(\hat{\boldsymbol{\gamma}}, \tau). \quad (17)$$

Eq. (14) can therefore be written equivalently as

$$\frac{d}{d\tau} \begin{pmatrix} \mathbf{Q}(\tau) \\ \mathbf{P}(\tau) \end{pmatrix} = \begin{pmatrix} \mathbf{W}^T(\tau) & \mathbf{V}(\tau) \\ -\mathbf{U}(\tau) & -\mathbf{W}(\tau) \end{pmatrix} \begin{pmatrix} \mathbf{Q}(\tau) \\ \mathbf{P}(\tau) \end{pmatrix}, \quad (18)$$

with

$$U_{ij}(\tau) = \frac{\partial^2 \mathcal{H}}{\partial x_i \partial x_j}[\mathbf{w}(\hat{\boldsymbol{\gamma}}, \tau)], \quad V_{ij}(\tau) = \frac{\partial^2 \mathcal{H}}{\partial p_i \partial p_j}[\mathbf{w}(\hat{\boldsymbol{\gamma}}, \tau)], \quad W_{ij}(\tau) = \frac{\partial^2 \mathcal{H}}{\partial x_i \partial p_j}[\mathbf{w}(\hat{\boldsymbol{\gamma}}, \tau)]. \quad (19)$$

Here, V_{ij} represents the wave-propagation metric tensor components introduced in eq. (12). Dynamic ray tracing, prescribed by the Hamilton–Jacobi perturbation eqs (18), may be performed simultaneously or subsequently with respect to ray tracing, prescribed by Hamilton’s eqs (5). In the latter case, the functions $\hat{\mathbf{x}}(\tau)$ and $\hat{\mathbf{p}}(\tau)$ will be known beforehand; the same is true for the functions $\mathbf{v}(\tau)$ and $\boldsymbol{\eta}(\tau)$ in eq. (6).

3.3 Ray propagator matrix

There are two common ways to find a solution of the system of ODEs in eq. (14) or (18) by means of integration. One is to integrate with right-hand sides of the differential equations exactly as specified in eq. (14) or (18), the other is to make use of a pre-calculated (known) first-order mapping between perturbed phase-space locations at the start and end point of the ray Ω . The coefficients of this mapping form the 6×6 ray propagator matrix in Cartesian coordinates. Below we introduce this matrix in a formal way.

A situation of particular interest arises if we choose the vector $\boldsymbol{\gamma}$ specifically as the 6-D phase-space perturbation at the initial point on Ω , for which $\tau = \tau_0$, that means,

$$\boldsymbol{\gamma}_r = (\delta w_r)_0 = w_r - \hat{w}_r(\tau_0). \quad (20)$$

Obviously, for this definition of $\boldsymbol{\gamma}$ we have $\hat{\boldsymbol{\gamma}} = \mathbf{0}$, the six-component zero vector. Using eq. (13), we establish the function

$$\delta w_r(\delta \mathbf{w}_0, \tau) = w_r(\delta \mathbf{w}_0, \tau) - \hat{w}_r(\tau), \quad (21)$$

where it is implicit that the freely varying (perturbation) vector $\delta \mathbf{w}_0$ belongs to the time τ_0 . The ray propagator matrix of size 6×6 in Cartesian coordinates can then be introduced as

$$\Pi_{ru}(\tau, \tau_0) = \frac{\partial(\delta w_r)}{\partial(\delta w_u)_0}(\delta \mathbf{w}_0 = \mathbf{0}, \tau). \quad (22)$$

The ray propagator matrix encapsulates the six fundamental solutions to the system (14) of ODEs. The Hamilton–Jacobi perturbation equations for the ray propagator matrix are given by

$$\frac{d\Pi_{ru}}{d\tau}(\tau, \tau_0) = S_{rt}(\tau) \Pi_{tu}(\tau, \tau_0); \quad \frac{d\Pi}{d\tau}(\tau, \tau_0) = \mathbf{S}(\tau) \Pi(\tau, \tau_0), \quad (23)$$

with the initial condition

$$\Pi_{ru}(\tau_0, \tau_0) = \delta_{ru}. \quad (24)$$

When the ray propagator matrix is known for the segment (τ, τ_0) of the ray Ω , any other dynamic ray-tracing solution on that segment can be computed using the linear combination of fundamental solutions,

$$X_{ra}(\tau) = \Pi_{rt}(\tau, \tau_0)X_{ta}(\tau_0); \quad \mathbf{X}(\tau) = \mathbf{\Pi}(\tau, \tau_0)\mathbf{X}(\tau_0). \quad (25)$$

In this way, the ray propagator matrix $\mathbf{\Pi}(\tau, \tau_0)$ represents the solution operator for the initial value problem (14).

The initial perturbation $\boldsymbol{\gamma} = \delta\mathbf{w}_0$ encompasses six degrees of freedom. A general perturbation $\delta\mathbf{w}_0$ can be considered to consist of (i) a *paraxial* contribution (four degrees of freedom), (ii) a *ray-tangent* contribution (one degree of freedom) and (iii) a *non-eikonal* contribution (one degree of freedom). The particular fundamental solutions resulting from these three types of initial conditions are often referred to as the paraxial, ray-tangent and non-eikonal solutions of dynamic ray tracing (Červený 2001).

3.4 Dynamic ray tracing specified by two paraxial ray parameters

Consider a situation with two parameters specifying the initial phase-space perturbation ($N_\gamma = 2$). For this particular situation we replace lowercase indices a and b in eq. (17) with corresponding uppercase indices A and B . We further assume that the parameters γ_A , $A = 1, 2$, have a purely paraxial nature, so that any initial phase-space perturbation is constrained not to have a ray-tangent or non-eikonal contribution. The parameters γ_A specify the initial conditions for paraxial rays, that means, rays in the vicinity of the reference ray Ω . We refer to γ_A as paraxial ray parameters or just *ray parameters*. For any reference ray or paraxial ray, the ray parameters are constant.

Together, the two quantities γ_A and the travelttime τ form a 3-D curvilinear *ray coordinate system*, $(\gamma_1, \gamma_2, \tau)$. The mapping from ray coordinates to Cartesian coordinates reads

$$x_i = x_i(\gamma_1, \gamma_2, \tau), \quad p_i = p_i(\gamma_1, \gamma_2, \tau). \quad (26)$$

We introduce 3×3 matrices $\widehat{\mathbf{Q}}$ and $\widehat{\mathbf{P}}$ for performing first-order transformation of position and momentum vectors from ray coordinates to Cartesian coordinates. The components of these matrices are

$$Q_{iA} = \frac{\partial x_i}{\partial \gamma_A}, \quad Q_{i3} = \frac{\partial x_i}{\partial \tau}, \quad (27)$$

$$P_{iA} = \frac{\partial p_i}{\partial \gamma_A}, \quad P_{i3} = \frac{\partial p_i}{\partial \tau}. \quad (28)$$

Matrix $\widehat{\mathbf{Q}}$ is the 3×3 geometrical spreading matrix for dynamic ray tracing in Cartesian coordinates.

For the inverse mapping operation, from Cartesian coordinates to ray coordinates, we introduce the *ray parameter function* $\gamma_A(\mathbf{x})$ and the *travelttime function* $\tau(\mathbf{x})$,

$$\gamma_A = \gamma_A(\mathbf{x}), \quad \tau = \tau(\mathbf{x}), \quad (29)$$

with first-order derivatives

$$\frac{\partial \gamma_A}{\partial x_i} = Q_{Ai}^\dagger, \quad \frac{\partial \tau}{\partial x_i} = p_i = Q_{3i}^\dagger. \quad (30)$$

The quantities Q_{ai}^\dagger in eq. (30) form the inverse of matrix $\widehat{\mathbf{Q}}$, such that

$$\widehat{\mathbf{Q}}^{-1} = \widehat{\mathbf{Q}}^\dagger. \quad (31)$$

More details on the first-order transformation between ray coordinates and Cartesian coordinates are given in Appendix B.

For the second derivatives of the travelttime function $\tau(\mathbf{x})$ we use the notation

$$\mathbf{M} = \{M_{ij}\} = \left\{ \frac{\partial^2 \tau}{\partial x_i \partial x_j} \right\}. \quad (32)$$

It is straightforward to show (see Appendix B) that matrix \mathbf{M} can be computed using

$$\mathbf{M} = \widehat{\mathbf{P}}\widehat{\mathbf{Q}}^{-1}. \quad (33)$$

4 HIGHER-ORDER HAMILTONIAN–JACOBI PERTURBATION EQUATIONS

In conventional dynamic ray tracing in Cartesian coordinates one continues along the ray Ω the first-order derivatives of a phase-space perturbation, given in eq. (15). That approach is extended here to include continuation of derivatives up to third order.

4.1 Continuation of second-order derivatives of phase-space perturbations

We formulate second-order derivatives of phase-space perturbations compactly as

$$X_{rab}(\tau) = \frac{\partial^2 w_r}{\partial \gamma_a \partial \gamma_b}(\hat{\boldsymbol{\gamma}}, \tau). \quad (34)$$

In Q – P notation we write them as

$$Q_{iab}(\tau) = \frac{\partial^2 x_i}{\partial \gamma_a \partial \gamma_b}(\hat{\boldsymbol{\gamma}}, \tau), \quad P_{iab}(\tau) = \frac{\partial^2 p_i}{\partial \gamma_a \partial \gamma_b}(\hat{\boldsymbol{\gamma}}, \tau). \quad (35)$$

We map the third-order derivatives of the Hamiltonian to a 3-D coefficient tensor,

$$S_{rtu}(\tau) = J_{rs} \frac{\partial^3 \mathcal{H}}{\partial w_s \partial w_t \partial w_u}[\mathbf{w}(\hat{\boldsymbol{\gamma}}, \tau)]. \quad (36)$$

Using the latter, we establish ODEs for continuation of second-order derivatives of perturbations in phase space,

$$\frac{dX_{rab}}{d\tau}(\tau) = S_{rt}(\tau)X_{iab}(\tau) + R_{rab}(\tau), \quad (37)$$

where

$$R_{rab}(\tau) = S_{rtu}(\tau)X_{ia}(\tau)X_{ub}(\tau). \quad (38)$$

The combination of ODEs given by eqs (14) and (37) can be integrated to yield the solution (34).

A different approach is to use an expression for the solution in terms of the initial condition, the ray propagator matrix and a closed-form integral along the ray Ω ,

$$X_{rab}(\tau) = \Pi_{rt}(\tau, \tau_0)X_{iab}(\tau_0) + \int_{\tau_0}^{\tau} \Pi_{rt}(\tau, \tau')R_{iab}(\tau')d\tau'. \quad (39)$$

This approach requires that the ray propagator matrix, $\Pi_{rt}(\tau, \tau_0)$, is a known (pre-computed) function.

It is convenient to reformulate eq. (39) such that the ray propagator matrix in the integrand corresponds to propagation from $\tau = \tau_0$ to $\tau = \tau'$. Using the chain property and the symplectic property of the ray propagator matrix we obtain

$$X_{rab}(\tau) = \Pi_{rt}(\tau, \tau_0) \left(X_{iab}(\tau_0) - \int_{\tau_0}^{\tau} J_{tv} \Pi_{sv}(\tau', \tau_0) J_{sq} R_{qab}(\tau') d\tau' \right). \quad (40)$$

4.2 Continuation of third-order derivatives of phase-space perturbations

We want to determine third-order derivatives of phase-space perturbations,

$$X_{rabc}(\tau) = \frac{\partial^3 w_r}{\partial \gamma_a \partial \gamma_b \partial \gamma_c}(\hat{\boldsymbol{\gamma}}, \tau), \quad (41)$$

or in Q – P notation,

$$Q_{iabc}(\tau) = \frac{\partial^3 x_i}{\partial \gamma_a \partial \gamma_b \partial \gamma_c}(\hat{\boldsymbol{\gamma}}, \tau), \quad P_{iabc}(\tau) = \frac{\partial^3 p_i}{\partial \gamma_a \partial \gamma_b \partial \gamma_c}(\hat{\boldsymbol{\gamma}}, \tau). \quad (42)$$

The ODEs for continuation of third-order derivatives of perturbations in phase space can be written,

$$\frac{dX_{rabc}}{d\tau}(\tau) = S_{rt}(\tau)X_{iabc}(\tau) + R_{rabc}(\tau), \quad (43)$$

where

$$S_{rtuv}(\tau) = J_{rs} \frac{\partial^4 \mathcal{H}}{\partial w_s \partial w_t \partial w_u \partial w_v}[\mathbf{w}(\hat{\boldsymbol{\gamma}}, \tau)], \quad (44)$$

$$R_{rabc}(\tau) = S_{rtu}(\tau)[X_{ia}(\tau)X_{ubc}(\tau) + X_{ib}(\tau)X_{uac}(\tau) + X_{uc}(\tau)X_{iab}(\tau)] + S_{rtuv}(\tau)X_{ia}(\tau)X_{ub}(\tau)X_{vc}(\tau). \quad (45)$$

The ODEs given by eqs (14), (37) and (43) can be integrated to yield the solution (41).

Alternatively we write the solution in terms of its initial condition, the ray propagator matrix and a closed-form integral along the ray Ω ,

$$X_{rabc}(\tau) = \Pi_{rt}(\tau, \tau_0)X_{iabc}(\tau_0) + \int_{\tau_0}^{\tau} \Pi_{rt}(\tau, \tau')R_{iabc}(\tau')d\tau'. \quad (46)$$

In this situation $\Pi_{rt}(\tau, \tau_0)$ and $X_{tab}(\tau)$ must be known along Ω . The chain and symplectic properties of the ray propagator matrix yield

$$X_{rabc}(\tau) = \Pi_{rt}(\tau, \tau_0) \left(X_{tabc}(\tau_0) - \int_{\tau_0}^{\tau} J_{tv} \Pi_{sv}(\tau', \tau_0) J_{sq} R_{qabc}(\tau') d\tau' \right). \quad (47)$$

Eq. (47) represents a third-order analogue of eq. (40).

5 CONSTRAINT RELATIONS

Derivatives of phase-space perturbations are in general not independent, as they will be constrained by the Hamilton–Jacobi eq. (2). To describe dependencies between first-order derivatives in conventional dynamic ray tracing Červený (2001) uses the notion *constraint relation*. When introducing derivatives of higher order it is necessary to consider also higher-order constraint relations.

5.1 Constraint relation for first-order derivatives of phase-space perturbations

The Hamiltonian is required to be constant along a trajectory in the phase space. As a consequence,

$$\frac{\partial \mathcal{H}}{\partial w_r} \frac{\partial w_r}{\partial \gamma_a} = 0, \quad \text{or equivalently,} \quad \frac{\partial \mathcal{H}}{\partial x_i} \frac{\partial x_i}{\partial \gamma_a} + \frac{\partial \mathcal{H}}{\partial p_i} \frac{\partial p_i}{\partial \gamma_a} = 0. \quad (48)$$

Thus, along Ω the following constraint applies,

$$v_i P_{ia} = \eta_i Q_{ia}. \quad (49)$$

Eq. (49) represents the constraint relation for first-order derivatives of phase-space perturbations in Cartesian coordinates (Červený 2001).

5.2 Constraint relations for second-order derivatives of phase-space perturbations

Differentiation of eq. (48) with respect to γ_b yields

$$\frac{\partial^2 \mathcal{H}}{\partial w_r \partial w_s} \frac{\partial w_r}{\partial \gamma_a} \frac{\partial w_s}{\partial \gamma_b} + \frac{\partial \mathcal{H}}{\partial w_r} \frac{\partial^2 w_r}{\partial \gamma_a \partial \gamma_b} = 0. \quad (50)$$

We use

$$\begin{aligned} \frac{\partial \mathcal{H}}{\partial w_r} &= -J_{rs} \dot{w}_s, \\ \frac{\partial^2 \mathcal{H}}{\partial w_r \partial w_s} &= -J_{rt} S_{ts}, \end{aligned}$$

so that eq. (50) becomes

$$-J_{rt} S_{ts} X_{ra} X_{sb} - J_{rs} \dot{w}_s X_{rab} = 0. \quad (51)$$

Applying the standard Hamilton–Jacobi perturbation eqs (14) then yields the second-order constraint relation,

$$-X_{rab} J_{rs} \dot{w}_s = X_{ra} J_{rs} \dot{X}_{sb}. \quad (52)$$

In Q – P notation the latter equation is restated

$$v_i P_{iab} = \eta_i Q_{iab} + Q_{ia} \dot{P}_{ib} - P_{ia} \dot{Q}_{ib}. \quad (53)$$

For a situation with two paraxial ray parameters γ_A , $A = 1, 2$, we may introduce in eq. (53) the 3×3 matrix \mathbf{M} of second derivatives of traveltime, defined in eq. (33), so that

$$v_i P_{iAB} = \eta_i Q_{iAB} + \dot{M}_{ij} Q_{iA} Q_{jB}. \quad (54)$$

5.3 Constraint relations for third-order derivatives of phase-space perturbations

We differentiate eq. (52) with respect to γ_c . Since the quantities γ_c are independent of the traveltime τ , the order of differentiation in γ_c and τ can be interchanged. We then obtain the third-order constraint relation,

$$-X_{rabc} J_{rs} \dot{w}_s = X_{rab} J_{rs} \dot{X}_{sc} + X_{rac} J_{rs} \dot{X}_{sb} + X_{ra} J_{rs} \dot{X}_{sbc}. \quad (55)$$

Introducing Q – P notation in eq. (55) yields the constraint relation

$$v_i P_{iabc} = \eta_i Q_{iabc} + Q_{iab} \dot{P}_{ic} - P_{iab} \dot{Q}_{ic} + Q_{iac} \dot{P}_{ib} - P_{iac} \dot{Q}_{ib} + Q_{ia} \dot{P}_{ibc} - P_{ia} \dot{Q}_{ibc}. \quad (56)$$

In the situation with two paraxial ray parameters, the lowercase indices a, b and c are replaced by their corresponding uppercase versions. The constraint relation (56) can then be rephrased in terms of second- and third-order derivatives of traveltime,

$$v_i P_{iABC} = \eta_i Q_{iABC} + \dot{M}_{ij} (Q_{iA} Q_{jBC} + Q_{iB} Q_{jAC} + Q_{iC} Q_{jAB}) + \dot{M}_{ijk} Q_{iA} Q_{jB} Q_{kC}. \quad (57)$$

6 INITIAL CONDITIONS

To be able to start the integration operations described above we need initial conditions.

At the initial point of the ray Ω , for which the time is $\tau = \tau_0$, we denote the position vector as $\mathbf{x} = \hat{\mathbf{x}}(\tau_0)$ and the slowness vector as $\mathbf{p} = \hat{\mathbf{p}}(\tau_0)$. It is necessary to specify the derivatives of the phase-space perturbations, $X_{ra}(\tau_0), X_{rab}(\tau_0), X_{rabc}(\tau_0), \dots$, up to the highest order under consideration in the system of Hamilton–Jacobi perturbation equations.

We limit our discussion of initial conditions to those prescribed by two paraxial ray parameters $\gamma_A, A = 1, 2$. In this respect, two cases are of particular interest: (i) the point-source situation where the initial wave front is degenerated and coincides with the source point, and (ii) the plane-wave situation where the initial wave front is a plane Π normal to the direction of the slowness vector $\hat{\mathbf{p}}(\tau_0)$.

To aid the setup of initial conditions, we introduce two linearly independent vectors \mathbf{e}_1 and \mathbf{e}_2 in the plane Π . Except for the requirements of linear independence and confinement to the plane Π , the orientation of \mathbf{e}_1 and \mathbf{e}_2 is arbitrary. The components of \mathbf{e}_1 and \mathbf{e}_2 form the 3×2 matrix $\mathcal{E} = \{\mathcal{E}_{iM}\}$. We establish a 3×3 matrix $\mathbf{H} = \{H_{ij}\}$ so that

$$\mathbf{H} = \begin{bmatrix} \mathcal{E} & \mathbf{v} \end{bmatrix}. \quad (58)$$

The inverse $\mathbf{H}^\dagger = \mathbf{H}^{-1}$ may then be expressed as

$$\mathbf{H}^{-1} = \begin{bmatrix} \mathcal{F}^T \\ \mathbf{p}^T \end{bmatrix}. \quad (59)$$

If \mathbf{e}_1 and \mathbf{e}_2 are chosen orthonormal, one will be able to compute \mathcal{F} from the relation

$$\mathcal{F}_{iM} = [\delta_{ij} - p_i v_j] \mathcal{E}_{jM} = \alpha_{ij} \mathcal{E}_{jM}, \quad (60)$$

where all quantities belong to $\tau = \tau_0$. The quantity α_{ij} represents a projection operator with respect to the wave-propagation metric tensor (Hanyga 1982; Klimeš 2006b). If \mathbf{e}_1 and \mathbf{e}_2 are not orthonormal, we compute \mathcal{F} from eq. (59), so that

$$\mathcal{F}_{iM} = H_{Mi}^\dagger. \quad (61)$$

6.1 Point source

For a point source, the two ray parameters γ_A will be parametrizing the slowness vectors of rays starting out from that point. Obviously, the location x_i of a point on the (degenerated) source wave front will be insensitive to any value of the parameters γ_A . As a consequence, all derivatives of position are zero at the source point,

$$Q_{iA} = \frac{\partial x_i}{\partial \gamma_A} = 0, \quad Q_{iAB} = \frac{\partial^2 x_i}{\partial \gamma_A \partial \gamma_B} = 0, \quad Q_{iABC} = \frac{\partial^3 x_i}{\partial \gamma_A \partial \gamma_B \partial \gamma_C} = 0, \quad \text{etc.} \quad (62)$$

We define here the two ray parameters γ_A at the source point by

$$\gamma_A = \mathcal{E}_{iA} [p_i - \hat{p}_i(\tau_0)]. \quad (63)$$

It is remarked that other definitions are possible, for example, one may let the parameters γ_A be Euler angles. The parameters γ_A in eq. (63) represent a projection of the slowness vector perturbation onto the coordinates corresponding to the vectors \mathbf{e}_1 and \mathbf{e}_2 . One important note in this context is that the three components of vector \mathbf{p} are constrained, as they have to satisfy the Hamilton–Jacobi equation.

From eq. (63) it follows that

$$\mathcal{E}_{iM} \frac{\partial p_i}{\partial \gamma_A}(\boldsymbol{\gamma}, \tau_0) = \delta_{MA}. \quad (64)$$

Moreover, the combination of constraint relation (49) and the first initial condition in eq. (62) gives

$$v_i(\tau_0) P_{iA}(\tau_0) = 0. \quad (65)$$

Then, solving the system of eqs (64) and (65) for $P_{iA}(\tau_0)$ yields at the initial point

$$P_{iA} = \mathcal{F}_{iA}, \quad (66)$$

where \mathcal{F}_{iA} is given by eq. (60) or (61).

To obtain $P_{iAB}(\tau_0)$, differentiate eq. (64) further,

$$\mathcal{E}_{iM} \frac{\partial^2 p_i}{\partial \gamma_A \partial \gamma_B}(\boldsymbol{\gamma}, \tau_0) = 0. \quad (67)$$

Also, we combine constraint relation (56) with initial conditions (62),

$$v_i(\tau_0) P_{iAB}(\tau_0) = -P_{iA}(\tau_0) \dot{Q}_{iB}(\tau_0). \quad (68)$$

Here, the time derivative on the right-hand side is given by the standard Hamilton–Jacobi perturbation eqs (18). We solve eqs (67) and (68) for $P_{iAB}(\tau_0)$, which yields

$$P_{iAB} = -p_i \mathcal{F}_{mA} \mathcal{F}_{nB} V_{mn}. \quad (69)$$

where all quantities belong to the initial point.

With values corresponding to a point source the constraint relation (56) becomes

$$v_j P_{jABC} = -P_{jA} \dot{Q}_{jBC} - P_{jAB} \dot{Q}_{jC} - P_{jAC} \dot{Q}_{jB}. \quad (70)$$

Applying eq. (69), we obtain

$$\dot{Q}_{jBC} = V_{jk} P_{kBC} + V_{jmn} P_{mB} P_{nC} = -V_{jk} p_k \mathcal{F}_{mB} \mathcal{F}_{nC} V_{mn} + V_{jmn} \mathcal{F}_{mB} \mathcal{F}_{nC} = -v_j \mathcal{F}_{mB} \mathcal{F}_{nC} V_{mn} + V_{jmn} \mathcal{F}_{mB} \mathcal{F}_{nC}, \quad (71)$$

where V_{jmn} is the tensor of third-order partial derivatives of the Hamiltonian with respect to slowness components, evaluated on the reference ray. Using eqs (66), (69) and (71) in (70) then yields

$$v_j P_{jABC} = -V_{jmn} \mathcal{F}_{jA} \mathcal{F}_{mB} \mathcal{F}_{nC}. \quad (72)$$

Since eq. (67) in addition implies that $\mathcal{E}_{iM} P_{iABC} = 0$, it follows that the initial condition for P_{iABC} is

$$P_{iABC} = -p_i V_{jmn} \mathcal{F}_{jA} \mathcal{F}_{mB} \mathcal{F}_{nC}. \quad (73)$$

If the Hamiltonian is a polynomial function of second degree in the slowness components, we have $V_{jmn} = 0$, and hence $P_{iABC} = 0$.

6.2 Plane-wave source

For a plane-wave source the two ray parameters γ_A may represent any pair of coordinates in the initial wave-front plane, Π . We choose here specifically these coordinates in the directions of the vectors \mathbf{e}_1 and \mathbf{e}_2 introduced earlier, such that $\partial \mathbf{x} / \partial \gamma_1 = \mathbf{e}_1$ and $\partial \mathbf{x} / \partial \gamma_2 = \mathbf{e}_2$ on Ω ; hence,

$$\frac{\partial x_i}{\partial \gamma_A}(\hat{\boldsymbol{\gamma}}, \tau_0) = \mathcal{E}_{iA}. \quad (74)$$

Since our choice of plane-wave ray parameters γ_A is connected with the matrix \mathcal{E} , the matrix \mathcal{F} in eq. (59) will also relate to these ray parameters. The connection is simply

$$\mathcal{F}_{iA} = \frac{\partial \gamma_A}{\partial x_i}, \quad (75)$$

where $\gamma_A = \gamma_A(\mathbf{x})$, see eq. (29), and where the derivative is taken at the source point on Ω . In view of eq. (75) we find it natural to define the plane-wave ray parameters by the linear expression

$$\gamma_A = \mathcal{F}_{iA} [x_i - \hat{x}_i(\tau_0)]. \quad (76)$$

The following relations must then be satisfied along the plane Π ,

$$\mathcal{F}_{iA} \frac{\partial x_i}{\partial \gamma_B}(\boldsymbol{\gamma}, \tau_0) = \delta_{AB}, \quad \mathcal{E}_{iM} p_i(\boldsymbol{\gamma}, \tau_0) = 0, \quad (77)$$

Eq. (74) readily yields the initial conditions

$$Q_{iA} = \mathcal{E}_{iA}, \quad Q_{iAB} = 0, \quad Q_{iABC} = 0. \quad (78)$$

Also, repeated differentiation of the last subequation in eq. (77) with respect to γ_A gives

$$\mathcal{E}_{iM} P_{iA} = 0, \quad \mathcal{E}_{iM} P_{iAB} = 0, \quad \mathcal{E}_{iM} P_{iABC} = 0. \quad (79)$$

To obtain P_{iA} we use eq. (79) and also invoke the constraint relation (49),

$$v_i P_{iA} = \eta_i Q_{iA}. \quad (80)$$

In combination, eqs (79) and (80) yield the initial condition

$$P_{iA} = p_i \eta_j \mathcal{E}_{jA}. \quad (81)$$

Proceeding to compute P_{iAB} , we take constraint relation (53) with $Q_{iA} = \mathcal{E}_{iA}$ and $Q_{mAB} = 0$, as prescribed by initial conditions (78),

$$v_m P_{mAB} = Q_{mA} \dot{P}_{mB} - P_{mA} \dot{Q}_{mB}. \quad (82)$$

Using the standard Hamilton–Jacobi perturbation equations then gives

$$v_m P_{mAB} = (-U_{ij} + 3\eta_i \eta_j) \mathcal{E}_{iA} \mathcal{E}_{jB}. \quad (83)$$

The combination of the middle sub-eq. (79) with (83) therefore yields

$$P_{mAB} = p_m [-U_{ij} + 3\eta_i \eta_j] \mathcal{E}_{iA} \mathcal{E}_{jB}. \quad (84)$$

For the initial plane wave front the constraint relation (56) can be formulated as (Appendix C)

$$\begin{aligned} v_m P_{mABC} &= Q_{mA} \dot{P}_{mBC} - P_{mA} \dot{Q}_{mBC} - P_{mAB} \dot{Q}_{mC} - P_{mAC} \dot{Q}_{mB} \\ &= (15\eta_i \eta_j \eta_k - 3\eta_i U_{jk} - 3\eta_j U_{ik} - 3\eta_k U_{ij} - U_{ijk}) \mathcal{E}_{iA} \mathcal{E}_{jB} \mathcal{E}_{kC}. \end{aligned} \quad (85)$$

Here, the tensor U_{ijk} represents the third-order derivatives of the Hamiltonian with respect to position components, evaluated on the reference ray. By combining the last sub-equation in (79) with (85) we obtain the initial condition

$$P_{mABC} = p_m (15\eta_i \eta_j \eta_k - 3\eta_i U_{jk} - 3\eta_j U_{ik} - 3\eta_k U_{ij} - U_{ijk}) \mathcal{E}_{iA} \mathcal{E}_{jB} \mathcal{E}_{kC}. \quad (86)$$

7 PARAXIAL EXTRAPOLATION

We describe extrapolation of geometrical spreading and traveltime away from the reference ray, Ω . The ray Ω is specified by paraxial ray parameters $\gamma_A = \gamma_{A0}$, $A = 1, 2$, and it includes a source point, $\mathbf{x} = \mathbf{s}_0$, and a reference receiver point, $\mathbf{x} = \mathbf{r}_0$. The traveltime at \mathbf{r}_0 is $\tau = \tau_0^R = \tau(\mathbf{r}_0)$.

7.1 Extrapolation of geometrical spreading

In the neighbourhood of the reference receiver point $\mathbf{x} = \mathbf{r}_0$ we consider the 3×3 geometrical spreading matrix $\widehat{\mathbf{Q}}$ to be a function of the receiver position, $\mathbf{x} = \mathbf{r}$, while the source position is kept fixed. A Taylor-series expansion of $\widehat{\mathbf{Q}}$ in $\Delta \mathbf{r} = \mathbf{r} - \mathbf{r}_0$ reads

$$Q_{ia}(\mathbf{r}, \mathbf{s}_0) = Q_{ia}(\mathbf{r}_0, \mathbf{s}_0) + \frac{\partial Q_{ia}}{\partial r_k}(\mathbf{r}_0, \mathbf{s}_0) \Delta r_k + \frac{1}{2} \frac{\partial^2 Q_{ia}}{\partial r_k \partial r_l}(\mathbf{r}_0, \mathbf{s}_0) \Delta r_k \Delta r_l + \frac{1}{6} \frac{\partial^3 Q_{ia}}{\partial r_k \partial r_l \partial r_m}(\mathbf{r}_0, \mathbf{s}_0) \Delta r_k \Delta r_l \Delta r_m + \dots \quad (87)$$

with the first three sets of derivatives given by

$$\frac{\partial Q_{ia}}{\partial r_k} = \frac{\partial^2 r_i}{\partial \gamma_a \partial \gamma_b} \frac{\partial \gamma_b}{\partial r_k} = Q_{iab} Q_{bk}^\dagger, \quad (88)$$

$$\frac{\partial^2 Q_{ia}}{\partial r_k \partial r_l} = \frac{\partial^3 r_i}{\partial \gamma_a \partial \gamma_b \partial \gamma_c} \frac{\partial \gamma_b}{\partial r_k} \frac{\partial \gamma_c}{\partial r_l} + \frac{\partial^2 r_i}{\partial \gamma_a \partial \gamma_b} \frac{\partial^2 \gamma_b}{\partial r_k \partial r_l} = Q_{iabc} Q_{bk}^\dagger Q_{cl}^\dagger + Q_{iab} Q_{bkl}^\dagger, \quad (89)$$

$$\begin{aligned} \frac{\partial^3 Q_{ia}}{\partial r_k \partial r_l \partial r_m} &= \frac{\partial^4 r_i}{\partial \gamma_a \partial \gamma_b \partial \gamma_c \partial \gamma_d} \frac{\partial \gamma_b}{\partial r_k} \frac{\partial \gamma_c}{\partial r_l} \frac{\partial \gamma_d}{\partial r_m} + \frac{\partial^3 r_i}{\partial \gamma_a \partial \gamma_b \partial \gamma_c} \left(\frac{\partial \gamma_b}{\partial r_k} \frac{\partial^2 \gamma_c}{\partial r_l \partial r_m} + \frac{\partial \gamma_b}{\partial r_l} \frac{\partial^2 \gamma_c}{\partial r_k \partial r_m} \right) + \frac{\partial^2 r_i}{\partial \gamma_a \partial \gamma_b} \frac{\partial^3 \gamma_b}{\partial r_k \partial r_l \partial r_m} \\ &= Q_{iabcd} Q_{bk}^\dagger Q_{cl}^\dagger Q_{dm}^\dagger + Q_{iabc} \left(Q_{bk}^\dagger Q_{clm}^\dagger + Q_{bl}^\dagger Q_{ckm}^\dagger \right) + Q_{iab} Q_{bklm}^\dagger. \end{aligned} \quad (90)$$

We see that a first-order expansion of matrix $\{Q_{ia}\}$ relies on second-order derivatives Q_{iab} , a second order expansion of $\{Q_{ia}\}$ relies on third-order derivatives Q_{iabc} , and so forth. The derivatives of the transformation from Cartesian coordinates to ray coordinates, that is, the quantities $Q_{bk}^\dagger, Q_{bkl}^\dagger, Q_{bklm}^\dagger, \dots$, can be obtained by repeated differentiation of eq. (B4).

It is assumed that the dynamic ray tracing is subject to point-source initial conditions at the point \mathbf{s}_0 . The relative geometrical spreading for a paraxial ray from \mathbf{s}_0 to \mathbf{r} can then be computed using

$$\mathcal{L}(\mathbf{r}, \mathbf{s}_0) = \left[\frac{1}{c(\mathbf{r})} \det\{Q_{ia}(\mathbf{r}, \mathbf{s}_0)\} \right]^{1/2}, \quad (91)$$

where $c(\mathbf{r})$ is the phase velocity of the paraxial ray evaluated at the position \mathbf{r} .

7.2 Extrapolation of traveltime

Consider the traveltime function $T(\mathbf{r}, \mathbf{s}_0) = \tau(\mathbf{r})$ corresponding to a fixed source point at $\mathbf{x} = \mathbf{s}_0$. We write a Taylor expansion of T in $\Delta \mathbf{r}$,

$$T(\mathbf{r}, \mathbf{s}_0) = \tau_0^R + p_k \Delta r_k + \frac{1}{2} M_{kl} \Delta r_k \Delta r_l + \frac{1}{6} M_{klm} \Delta r_k \Delta r_l \Delta r_m + \frac{1}{24} M_{klmn} \Delta r_k \Delta r_l \Delta r_m \Delta r_n + \dots, \quad (92)$$

where all the coefficients are evaluated at \mathbf{r}_0 .

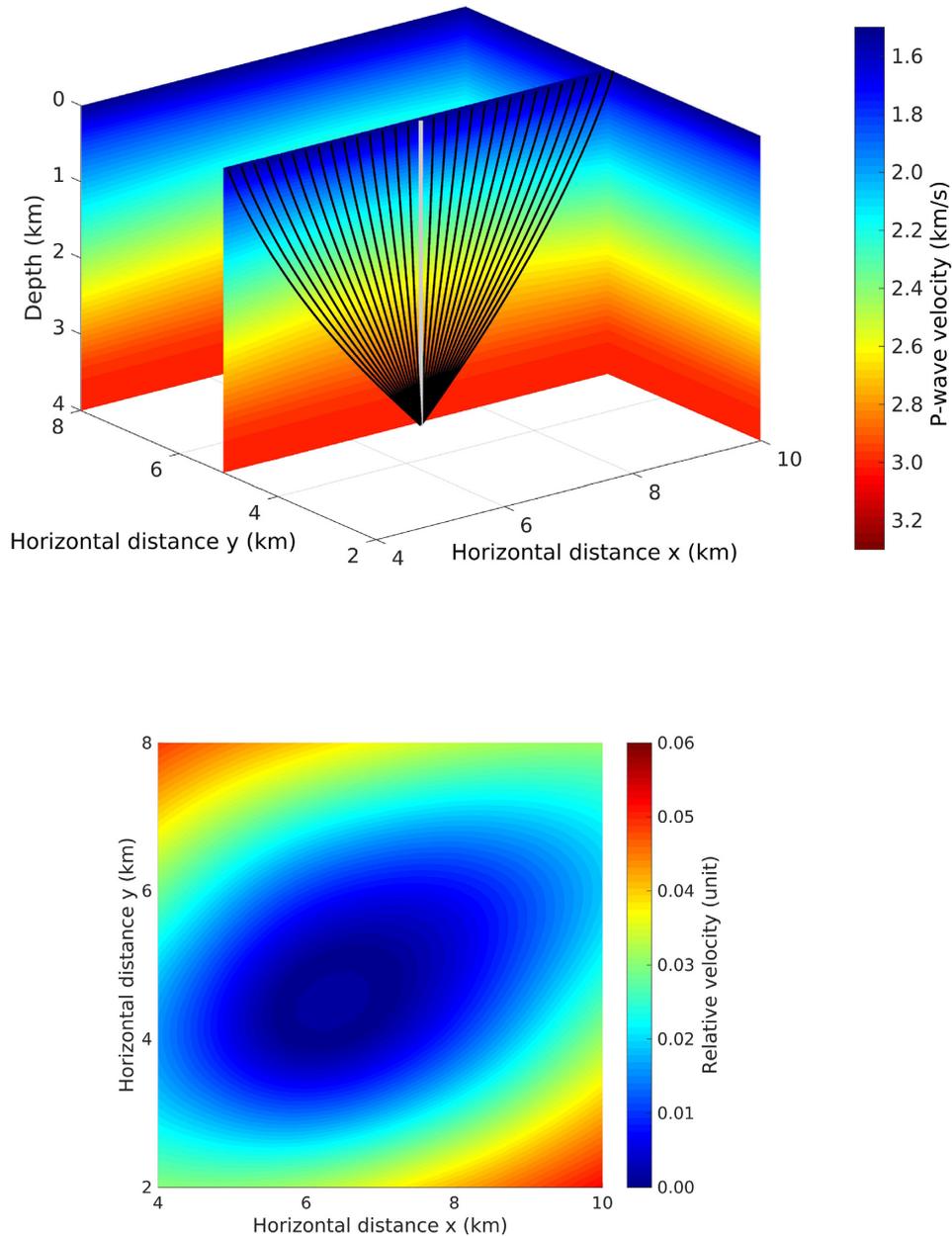


Figure 1. Model ISO and some of the rays used in numerical tests. Top: the vertical P -wave velocity is indicated by colour in vertical slices. Data for numerical comparisons are computed along rays (black) from a source point at depth 4 km. Coefficients for extrapolation of ray quantities are computed along a single reference ray (grey). Bottom: lateral differences in the vertical P -wave velocity, relative to the velocity on the reference ray, at depth 2 km.

The 3×3 matrix of second derivatives of traveltime, \mathbf{M} , can be computed using eq. (33). The latter is restated here as

$$M_{ij} Q_{ja} = P_{ia}, \tag{93}$$

where all indices run from 1 to 3. Differentiating eq. (93) twice with respect to the ray coordinates yields,

$$M_{ijk} Q_{ja} Q_{kb} = P_{iab} - M_{ij} Q_{jab}, \tag{94}$$

$$M_{ijkl} Q_{ja} Q_{kb} Q_{lc} = P_{iabc} - M_{ijk} (Q_{ja} Q_{kbc} + Q_{jb} Q_{kac} + Q_{jc} Q_{kab}) - M_{ij} Q_{jab}. \tag{95}$$

When the right-hand sides of eqs (93)–(95) have been evaluated, we obtain explicit expressions for the second-, third- and fourth-order derivatives of traveltime after multiplying by the relevant number of (inverse) matrices, $\{Q_{aj}^\dagger\}$.

To evaluate the right-hand side of (94) we need to know M_{ij} , which is pre-computed using (93). In addition, we need Q_{iAB} , P_{iAB} , \dot{Q}_{iA} , \dot{P}_{iA} , \dot{v}_i and $\dot{\eta}_i$. To evaluate the right-hand side of eq. (95) we will also need M_{ijk} , pre-computed using (94), as well as Q_{iABC} , P_{iABC} , \dot{Q}_{iAB} ,

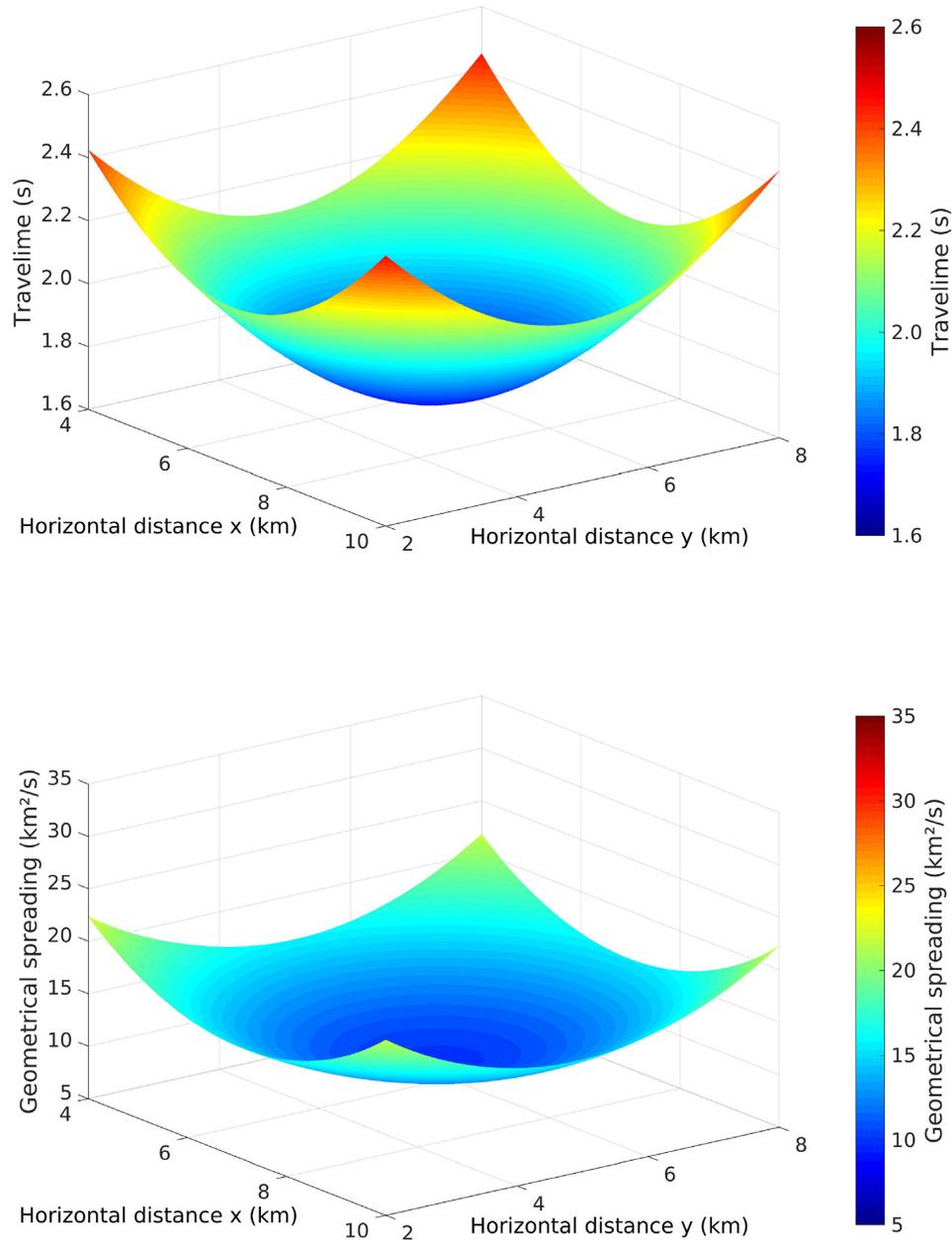


Figure 2. Model ISO: ray-tracing simulated P -wave traveltime (top) and geometrical spreading (bottom) in the plane at depth 0 km for a source point at depth 4 km.

\dot{P}_{iAB} , \ddot{Q}_{iA} , \dot{P}_{iA} , \ddot{v}_i and $\dot{\eta}_i$. The temporal derivatives \dot{Q}_{iA} , \dot{P}_{iA} , \dot{Q}_{iAB} and \dot{P}_{iAB} can be readily obtained from the relevant system of differential equations given in eq. (14) or (37). The derivatives \ddot{Q}_{iA} , \ddot{P}_{iA} , \ddot{v}_i , $\ddot{\eta}_i$, \ddot{v}_i and $\ddot{\eta}_i$ can be obtained after temporal differentiation of these equations.

In some situations it can be useful to do a Taylor expansion of squared traveltime rather than of the traveltime itself, as the expansion of $T^2(\mathbf{r}, \mathbf{s}_0)$ to second order in $\Delta \mathbf{r}$ is exact for waves from a point source in an isotropic homogeneous medium. For underlying theory and numerical examples, see Ursin (1982); Gjøystdal *et al.* (1984). Extrapolation of squared traveltime may be highly appropriate also for a transversely isotropic medium with a vertical axis of symmetry. For details on this matter, see Alkhalifah & Tsvankin (1995) and Tsvankin (2013).

8 HAMILTONIANS FOR P AND S WAVES IN ANISOTROPIC HETEROGENEOUS MEDIA

Up to this point, the theory has been described with the Hamiltonian appearing as a black box. In this section, we elaborate on specific Hamiltonians related to P and S waves in anisotropic heterogeneous media.

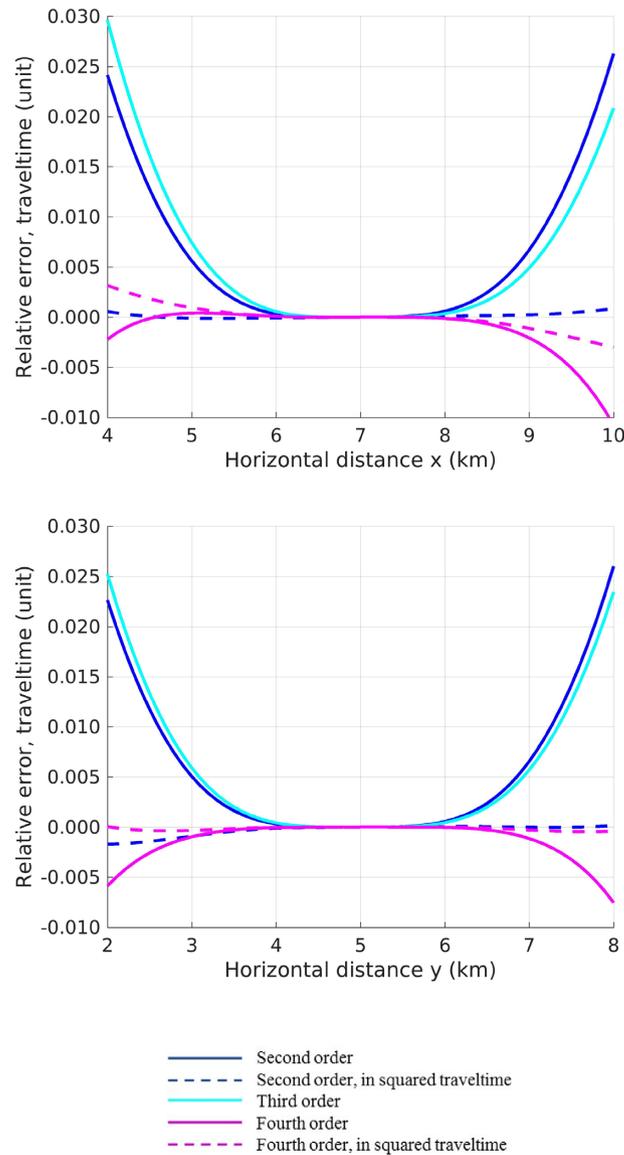


Figure 3. Model ISO: Relative error in traveltimes for different extrapolation approaches along the lines $y = 5$ km (top) and $x = 7$ km (bottom).

8.1 Arbitrary anisotropy

The Christoffel matrix Γ , as defined, for example, in Červený (2001), eq. (2.2.19), is of size 3×3 and has the components

$$\Gamma_{ik}(\mathbf{x}, \mathbf{p}) = a_{ijkl}(\mathbf{x}) p_j p_l. \tag{96}$$

Here, a_{ijkl} is the tensor of density-normalized elastic moduli. The Christoffel matrix has three real eigenvalues and three corresponding mutually orthogonal eigenvectors. One selected eigenvalue and its associated eigenvector are denoted, respectively, by the symbols G and \mathbf{g} . The eigenvalue G corresponds to an elementary P or S wave with polarization vector \mathbf{g} .

The eigenvalues of matrix Γ satisfy the characteristic equation

$$\det(\Gamma - G\mathbf{I}) = 0; \tag{97}$$

here, \mathbf{I} is the 3×3 identity matrix. Eq. (97) represents a third-order polynomial in G ,

$$G^3 - PG^2 + QG - R = 0, \tag{98}$$

where the quantities P , Q and R are invariants of Γ ,

$$P = \text{tr}\Gamma, \quad Q = \text{tr}(\text{cof}\Gamma), \quad R = \det \Gamma. \tag{99}$$

We note that P , Q and R are scalar functions in phase space, and they are homogeneous of degree two (P), four (Q) and six (R) in the slowness components, p_i . The function G is homogeneous of degree two in p_i .

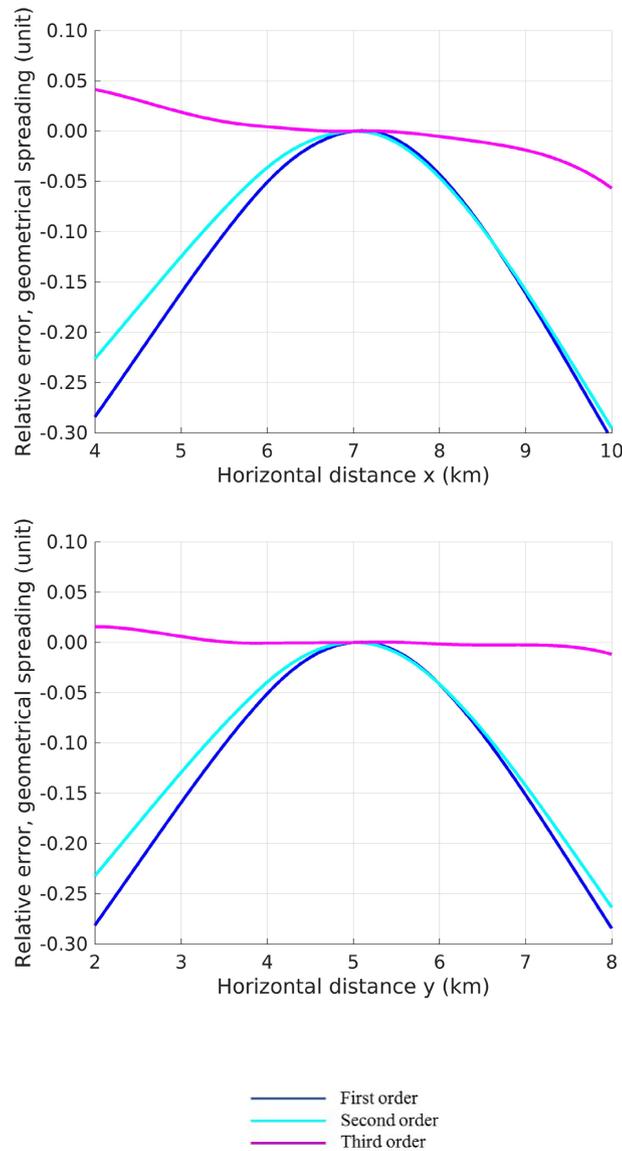


Figure 4. Model ISO: relative error in geometrical spreading for different extrapolation approaches along the lines $y = 5$ km (top) and $x = 7$ km (bottom).

For a given wave mode and a given point \mathbf{x} , we get an exact description of the relevant slowness-surface sheet if the eigenvalue function G is subject to the constraint $G(\mathbf{w}) = 1$. On the other hand, our Hamiltonian \mathcal{H} satisfies the Hamilton–Jacobi equation $\mathcal{H} = 1/2$. Hence, to ensure consistency with the slowness surface it is natural to express \mathcal{H} in terms of G , such that

$$\mathcal{H}(\mathbf{w}) = \frac{1}{2} G(\mathbf{w}). \tag{100}$$

In the case of arbitrary anisotropy, derivatives of \mathcal{H} are obtained by differentiation of eqs (98) and (100), followed by setting $G = 1$.

We remark that for some applications of ray perturbation theory (see, e.g. Klimeš 2002a; Červený & Klimeš 2009) it may be useful to redefine the Hamiltonian to

$$\mathcal{H}(\mathbf{w}) = \frac{1}{\mathcal{N}} [G(\mathbf{w})]^{\mathcal{N}/2}, \tag{101}$$

where \mathcal{N} is a nonzero scalar. Such a redefinition will only affect the non-eikonal solution to dynamic ray tracing; the other fundamental solutions are unaffected.

8.2 Partial factoring of the characteristic equation by polarization

For particular anisotropic symmetries, for example, transversely isotropic media, one can utilize a partial factoring of the characteristic eq. (98) by polarization. One of the elementary S waves is then SH polarized, meaning that the polarization vector is confined to the (locally) horizontal plane. The polarization vectors of the two other elementary waves, P and SV , form a (locally) vertical plane. The partial factoring

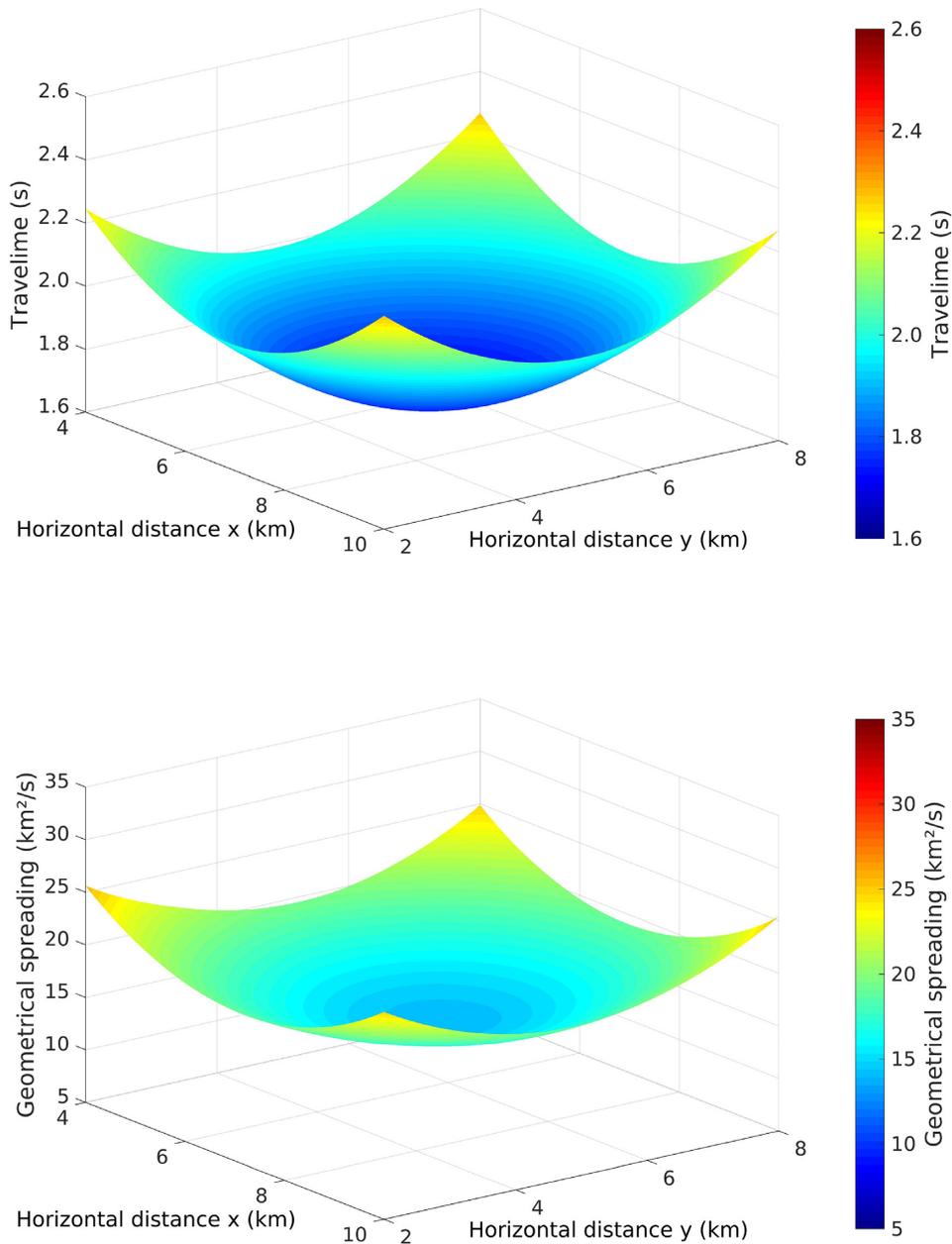


Figure 5. Model VEL: ray-tracing simulated P -wave traveltime (top) and geometrical spreading (bottom) in the plane at depth 0 km for a source point at depth 4 km.

is stated

$$(G^2 - P^{PSV}G + R^{PSV})(G - G^{SH}) = 0. \tag{102}$$

In this situation we get a specific equation for the Hamiltonian of the SH -polarized wave,

$$\mathcal{H} = \frac{1}{2}G^{SH}. \tag{103}$$

The eigenvalue function G^{SH} is homogeneous of degree two in the slowness components, and the slowness sheet for the SH wave is elliptical.

Eq. (102) further yields another, common, equation for the P - and SV -polarized waves,

$$G^2 - P^{PSV}G + R^{PSV} = 0. \tag{104}$$

The functions P^{PSV} and R^{PSV} are homogeneous of degree two and four in the slowness components.

For an SH wave, derivatives of \mathcal{H} in phase space are obtained by differentiating eq. (103). For a P or SV wave, we differentiate eqs (100) and (104), followed by setting $G = 1$.

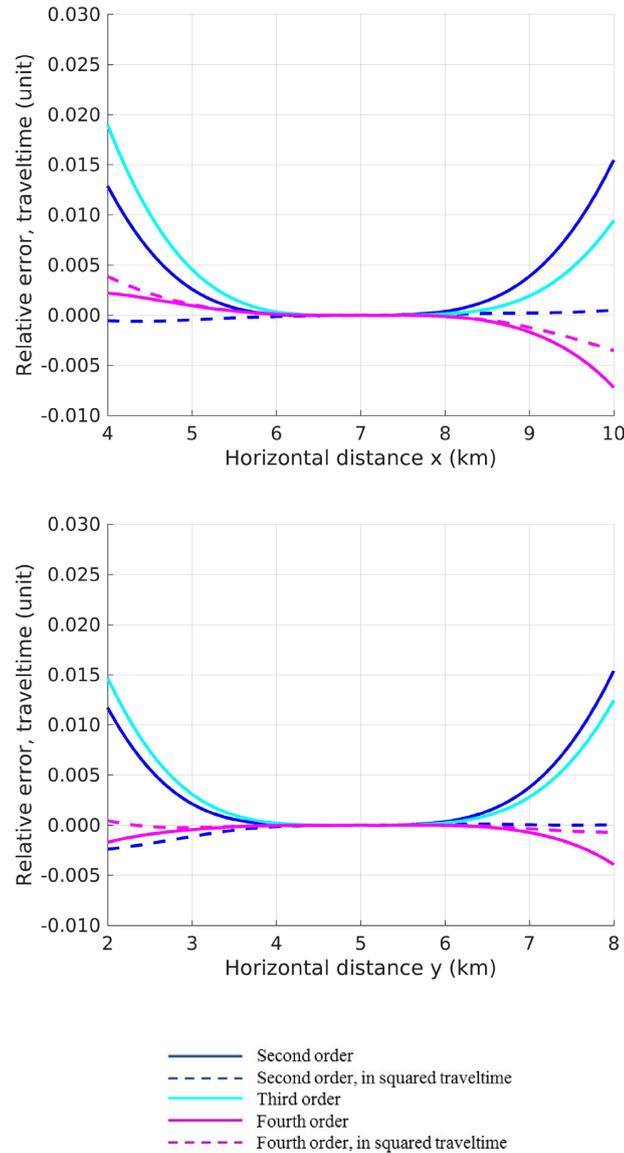


Figure 6. Model VEL: Relative error in traveltimes for different extrapolation approaches along the lines $y = 5$ km (top) and $x = 7$ km (bottom).

8.3 Full factoring of the characteristic equation by polarization

Consider a further factoring of eq. (102) so that

$$(G - G^P)(G - G^{SV})(G - G^{SH}) = 0. \tag{105}$$

This yields the following possibilities for the Hamiltonian,

$$\mathcal{H} = \frac{1}{2}G^P, \quad \mathcal{H} = \frac{1}{2}G^{SV}, \quad \mathcal{H} = \frac{1}{2}G^{SH}. \tag{106}$$

The eigenvalue functions $G^P(\mathbf{w})$, $G^{SV}(\mathbf{w})$ and $G^{SH}(\mathbf{w})$ are all homogeneous of second degree in the slowness components. The P - and SH -wave slowness sheets are elliptical; the SV -wave slowness sheet is spherical.

To obtain derivatives of the Hamiltonian \mathcal{H} , we differentiate the relevant eigenvalue function in eq. (106).

9 NUMERICAL EXAMPLES

We have performed numerical tests of the above described higher-order approaches to dynamic ray tracing in Cartesian coordinates, using three related 3-D heterogeneous models. All the simulation examples are for P waves and for a single source point.

We employ a point-source initialization of the dynamic ray tracing system. The two ray parameters that specify the initial conditions are horizontal components of the slowness vector at the source point. In this case the computed first- and higher-order derivatives of phase-space

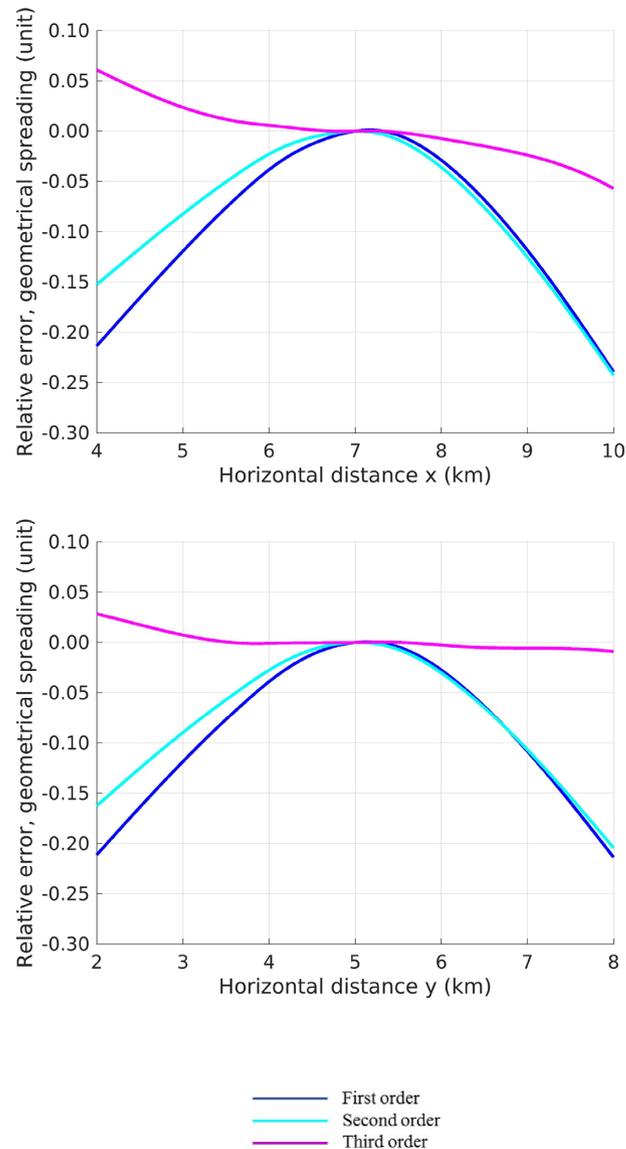


Figure 7. Model VEL: relative error in geometrical spreading for different extrapolation approaches along the lines $y = 5$ km (top) and $x = 7$ km (bottom).

perturbations represent a pure paraxial solution, which means that the choice of the value of the quantity \mathcal{N} in eq. (101) should, from a theoretical point of view, not influence the results. However, one cannot rule out that \mathcal{N} may have an effect in the presence of numerical errors, but this matter has not been subject to our study.

In this section the spatial coordinates of the models are referred to as (x, y, z) .

9.1 Model ISO

The first model, Model ISO, is adopted from Iversen & Tygel (2008)—it is isotropic and includes a gentle anticline structure (Fig. 1). However, in the implementation of higher-order dynamic ray tracing we use a quintic (fifth-order) B-spline representation (e.g. Farin *et al.* 2002) to ensure C^4 continuity of the volumetric medium parameter functions. As a consequence, the P -wave velocity field appears here in a somewhat smoother form than in Iversen & Tygel (2008). The ratio of S -wave to P -wave velocity is constant = 0.5. Data for numerical comparisons is obtained using conventional P -wave kinematic and dynamic ray tracing from a source point at depth $z = 4$ km to receivers in the plane at zero depth. Fig. 1 shows a subselection of rays (black) for receivers along the line $y = 5$ km. The ray (grey) arriving at the receiver location $(7, 5, 0)$ km is taken as a reference ray for higher-order dynamic ray tracing computations.

Fig. 2 shows the computed traveltimes data (top) and geometrical spreading data (bottom). Geometrical spreading was computed using eq. (91). As ray parameters γ_A in that equation we used the two horizontal components of the slowness vector at the source point.

Our results are shown as error curves for the extrapolated traveltimes (Fig. 3) and the extrapolated geometrical spreading (Fig. 4). The computed error curves belong to a line of constant $y (= 5$ km) and a line of constant $x (= 7$ km). We refer to the lateral distance between

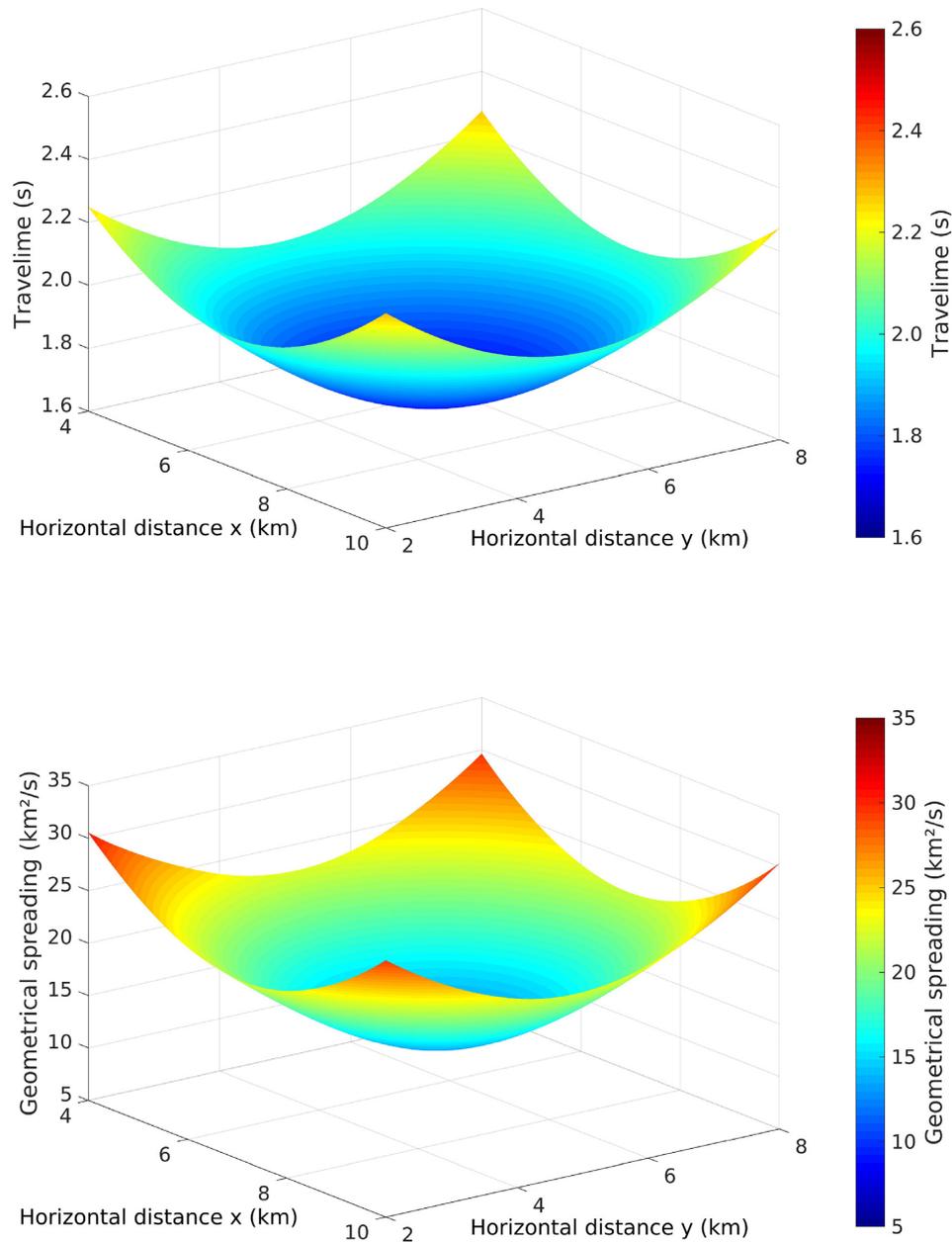


Figure 8. Model VTI: ray-tracing simulated P -wave traveltime (top) and geometrical spreading (bottom) in the plane at depth 0 km for a source point at depth 4 km.

a receiver and the reference ray as the *paraxial distance*. In the cross-sections $y = 5$ and $x = 7$ (km) the maximum paraxial distance is 3 km.

For traveltime extrapolation, we observe that second-order extrapolation of squared traveltime (dashed blue) yields a very good result. The maximum relative errors at 3 km paraxial distance are around 0.15 per cent. For heterogeneous media such a good result is not obvious (Gjøystdal *et al.* 1984). In the current test, however, the heterogeneities are moderate. We note that the results for fourth-order extrapolation of traveltime (magenta) and squared traveltime (dashed magenta) are also very good. The range of relative errors for the latter approach is 0–0.3 per cent.

Concerning extrapolation of geometrical spreading, a striking observation is that the extrapolation function needs to be at least third order in the spatial coordinates, in order to be appropriate at large paraxial distances. The relative errors obtained for the third-order approach (magenta) are below 1 per cent for paraxial distances 0–1.5 km and below 5 per cent for paraxial distances 1.5–3 km.

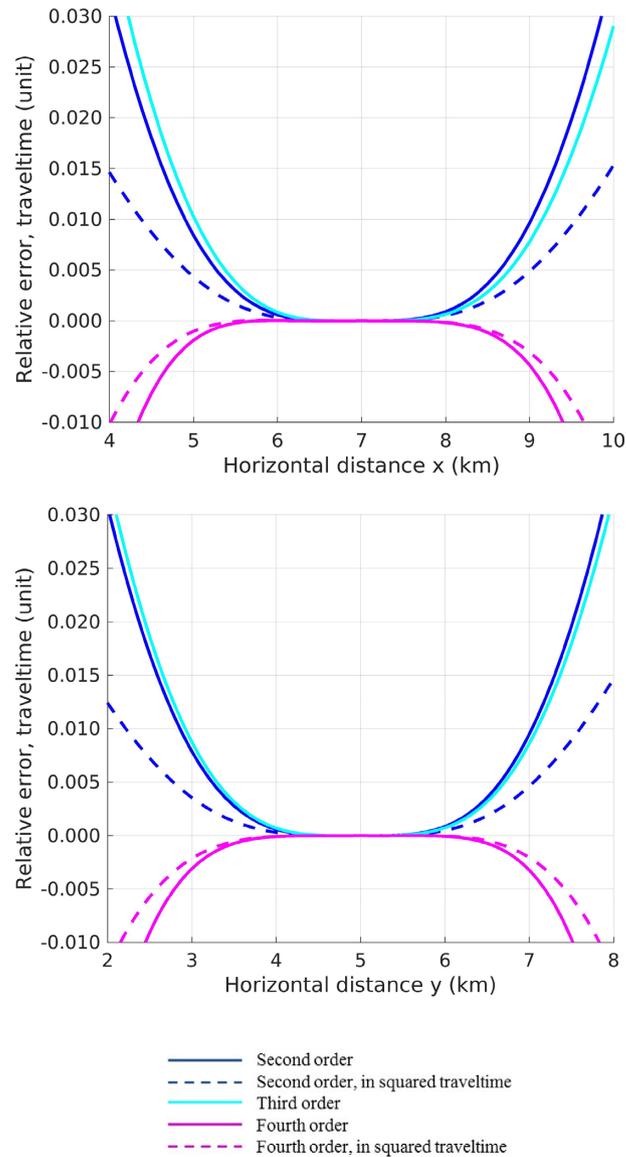


Figure 9. Model VTI: Relative error in traveltime for different extrapolation approaches along the lines $y = 5$ km (top) and $x = 7$ km (bottom).

9.2 Model VEL

The second model, Model VEL, differs from Model ISO only in one respect—we have introduced elliptic anisotropy related to a vertical axis of symmetry. The anisotropy was defined to be constant, with Thomsen's (1986) parameters specified as $\varepsilon = \delta = 0.2$. Since we consider P -wave simulation only, the value of Thomsen's parameter γ (not to be confused with ray parameters) does not affect our computations.

To get an impression of the effect of introducing strong elliptic anisotropy we can compare Figs 2 and 5. We note a decrease in traveltime at large lateral distances from the reference ray, and also a general increase in the values of geometrical spreading.

Figs 6 and 7 show the same type of error curves as was given for Model ISO (Figs 3 and 4). Extrapolation of squared traveltime to second order (dashed blue) yields also for Model VEL an excellent result. The reason is twofold—the heterogeneities are moderate and the anellipticity effect is zero. For a corresponding homogeneous model the fourth-order term of the extrapolation function for squared traveltime would have vanished completely (Alkhalifah & Tsvankin 1995). The results for fourth-order extrapolation of traveltime (magenta) and squared traveltime (dashed magenta) are also very good. We note that the results for second-order extrapolation of traveltime (blue) and third-order extrapolation of traveltime are significantly better than in the isotropic case. Third-order extrapolation of geometrical spreading (magenta) has the same level of relative accuracy as for the isotropic model, that is, less than 1 per cent / 5 per cent for paraxial distance ranges, respectively, 0–1.5 km/1.5–3 km.

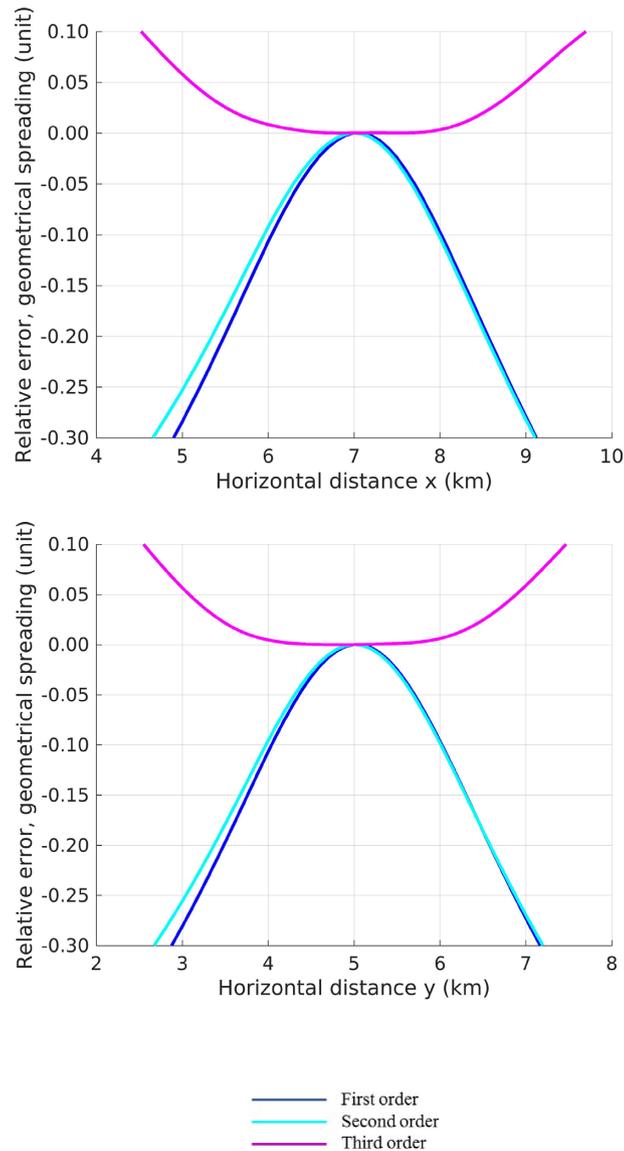


Figure 10. Model VTI: relative error in geometrical spreading for different extrapolation approaches along the lines $y = 5$ km (top) and $x = 7$ km (bottom).

9.3 Model VTI

In the third model, Model VTI, the anisotropy is transversely isotropic (hexagonal) relative to a vertical axis of symmetry. We have introduced anellipticity in the slowness surface by the parameter setting $\varepsilon = 0.3$, $\delta = 0.1$. The parameters are constant throughout the model, so the anellipticity is invariant.

By comparing Fig. 5 for Model VEL with Fig. 8, we see that the change in traveltime is quite minor, while geometrical spreading has substantial changes at large paraxial distances.

Figs 9 and 10 show curves of relative errors in traveltime and geometrical spreading. They are to be compared to the corresponding Figs 3 and 4 (for Model ISO) and Figs 6 and 7 (for Model VEL). We observe that it is now harder to get a good extrapolation result in the full range of paraxial distances (3 km). When using fourth-order extrapolation of squared traveltime, the errors are below 0.025 per cent within a paraxial distance of 2 km. For third-order extrapolation of geometrical spreading, errors are below 1 per cent within 1 km paraxial distance.

10 DISCUSSION AND CONCLUSIONS

We give a brief discussion on the connections to differential geometry, followed by our concluding remarks.

The theory derived in this paper has important connections to differential geometry—in particular to the Riemannian and Finslerian geometry systems (see, e.g. Shen 2001; Bao *et al.* 2012). Any Riemannian, Finslerian, or similar geometrical structure can be completely described by distances. In the propagation of elastic waves distance is naturally measured in traveltime: the distance between two points is the

shortest time it takes for a wave to go from one point to the other. By Fermat's principle elastic waves travel along geodesics (rays). Locally, the geodesics represent length-minimizing curves of the geometrical structure under consideration.

The elastic geometry of P waves in anisotropic media is Finsler geometry, as observed by Antonelli *et al.* (2003). If the anisotropy is elliptical, then the Finsler geometry simplifies to Riemannian geometry. For anisotropy of lower symmetry than elliptic, non-Riemannian geometric features are intrinsic properties of the medium. Leading-order variations from a reference ray can be adequately described by elliptic anisotropy, and similarly, a Finsler geometry has a canonical Riemannian metric along a reference geodesic. Higher-order variations require a Finsler treatment—elliptic approximations are insufficient.

To gain access to the powerful tools of differential geometry, it is reasonable to rephrase the study of elastic waves in a geometrical framework. Writing equations in terms of coordinate invariant geometrical quantities simplifies the structure significantly. Many extra terms in calculations in anisotropic heterogeneous media may be viewed as a symptom of computing in coordinates rather than invariantly. For example, the geometric counterpart of the Q – P quantities is a Jacobi field and its covariant derivative; for details on the geometrical aspects, see Paternain (2012). In the equation of motion of the geometrical variant, similar to eq. (18), the diagonal blocks are automatically zero and the top-right block is the identity. All geometrical information is packed into the curvature operator appearing as the lower left block. It is our intention to elaborate on these connections between differential geometry and seismology in future work.

A key motivation behind this paper is to contribute to better computational capabilities, with respect to accuracy as well as efficiency, for processes exploiting amplitude and/or phase properties of high-frequency Green's functions.

The starting point of the theory development is a Hamilton–Jacobi equation in Cartesian coordinates. Based on this fundamental equation, we first reviewed the standard approach to dynamic ray tracing, by which first-order derivatives of phase-space perturbations are continued along a reference ray. Thereafter, we developed a theoretical framework for computation of higher-order derivatives of the phase-space perturbations. Detailed results were exposed for the orders two and three in these derivatives.

We did numerical tests of higher-order dynamic ray tracing for three related 3-D heterogeneous models—one isotropic, one elliptically anisotropic and one transversely isotropic. The higher-order approach yields clear improvements of the paraxial extrapolation of traveltime and geometrical spreading, compared to results obtained using conventional dynamic ray tracing. One important observation is that the extrapolation function for geometrical spreading must be at least third order to be appropriate at large extrapolation distances.

For the special case of a point source in a caustic-free medium, the approach proposed by Klimeš (2002a, 2006a) is potentially faster than the methodology presented here, as the number of equations to be integrated are lower. Therefore, as a future issue, it would be interesting to see a comparative study of the two approaches that addresses computational efficiency, but in addition also accuracy and stability.

The presented methodology opens possibilities for further research and development in several directions. So far we have only considered continuous models—for completeness it should also be possible to use models with interfaces. Furthermore, when concerning S waves in anisotropic media one should take into account Hamiltonians that honour both elementary S waves simultaneously, such that problems with S -wave singularities are circumvented. Having available higher-order derivatives of the phase-space perturbations it will be possible to formulate also higher-order variants of Hamilton's two-point characteristic. Last, but not least, we intend to develop the methodology also for ray-centred coordinates, taking into account the connections to differential geometry outlined above.

ACKNOWLEDGEMENTS

MvdH gratefully acknowledges support from the Simons Foundation under the MATH + X program, the National Science Foundation under grant DMS-1815143 and the corporate members of the Geo-Mathematical Group at Rice University. TS gratefully acknowledges support from the Simons Foundation under the MATH + X program. JI has been supported by the Academy of Finland (decision 295853). BU has received support from the Research Council of Norway through the ROSE project. EI and BU are grateful for being invited to meetings at Rice University. In this work we have used academic software licenses from NORSAR and MATLAB.

REFERENCES

- Alkhalifah, T. & Tsvankin, I., 1995. Velocity analysis for transversely isotropic media, *Geophysics*, **60**(5), 1550–1566.
- Antonelli, P., Rutz, S. & Slawinski, M., 2003. A geometrical foundation for seismic ray theory based on modern Finsler geometry, in *Finsler and Lagrange Geometries*, pp. 17–54, eds Anastasiei, M. & Antonelli, P.L., Springer.
- Bao, D., Chern, S.-S. & Shen, Z., 2012. *An Introduction to Riemann-Finsler geometry*, Vol. 200, Springer Science & Business Media.
- Beylkin, G. & Burridge, R., 1990. Linearized inverse scattering problems in acoustics and elasticity, *Wave Motion*, **12**(1), 15–52.
- Bleistein, N., Cohen, J.K. & Stockwell, J.W., Jr, 2001. *Mathematics of Multidimensional Seismic Imaging, Migration, and Inversion*, Springer.
- Bortfeld, R., 1989. Geometrical ray theory: rays and traveltimes in seismic systems (second-order approximation of the traveltimes), *Geophysics*, **54**, 342–349.
- Brandsberg-Dahl, S., de Hoop, M.V. & Ursin, B., 2003a. Focusing in dip and AVA compensation on scattering-angle/azimuth common image gathers, *Geophysics*, **68**(1), 232–254.
- Brandsberg-Dahl, S., Ursin, B. & de Hoop, M.V., 2003b. Seismic velocity analysis in the scattering angle/azimuth domain, *Geophys. Prospect.*, **51**(4), 295–314.
- Cameron, M.K., Fomel, S. & Sethian, J.A., 2007. Seismic velocity estimation from time migration, *Inverse Probl.*, **23**, 1329–1369.
- Červený, V., 1972. Seismic rays and ray intensities in inhomogeneous anisotropic media, *Geophys. J. R. astr. Soc.*, **29**, 1–13.
- Červený, V., 2001. *Seismic Ray Theory*, Cambridge University Press.
- Červený, V. & Klimeš, L., 2009. Transformation relations for second derivatives of traveltime in anisotropic media, in *19th Annual Report of the Seismic Waves in Complex 3-D Structures (SW3D) Consortium*, pp. 115–122, Department of Geophysics, Faculty of Mathematics and Physics, Charles University.

- Červený, V. & Moser, T.J., 2007. Ray propagator matrices in three-dimensional anisotropic inhomogeneous layered media, *Geophys. J. Int.*, **168**, 593–604.
- Červený, V. & Pšenčík, I., 2010. Gaussian beams in inhomogeneous anisotropic layered structures, *Geophys. J. Int.*, **180**(2), 798–812.
- Červený, V., Klimeš, L. & Pšenčík, I., 1984. Paraxial ray approximations in the computation of seismic wavefields in inhomogeneous media, *Geophys. J. R. astr. Soc.*, **79**, 89–104.
- Červený, V., Klimeš, L. & Pšenčík, I., 1988. Complete seismic ray tracing in three-dimensional structures, in *Seismological Algorithms, Computational Methods and Computer Programs*, pp. 89–168, ed. Doornbos, D.J., Academic Press.
- Červený, V., Iversen, E. & Pšenčík, I., 2012. Two-point paraxial travel-times in an inhomogeneous anisotropic medium, *Geophys. J. Int.*, **189**(3), 1597–1610.
- Chapman, C.H., 2004. *Fundamentals of Seismic Wave Propagation*, Cambridge University Press.
- de Hoop, M.V. & Bleistein, N., 1997. Generalized Radon transform inversions for reflectivity in anisotropic elastic media, *Inverse Probl.*, **13**(3), 669–690.
- de Hoop, M.V. & Uhlmann, G., 2006. Characterization and ‘source-receiver’ continuation of seismic reflection data, *Commun. Math. Phys.*, **263**(1), 1–19.
- de Hoop, M.V., Burridge, R., Spencer, C. & Miller, D., 1994. Generalized radon transform/amplitude versus angle (GRT/AVA) migration/inversion in anisotropic media, *Proc. SPIE*, **2301**, 15–27.
- de Hoop, M.V., Smith, H., Uhlmann, G. & van der Hilst, R.D., 2009. Seismic imaging with the generalized Radon transform: a curvete transform perspective, *Inverse Probl.*, **25**(2), 025005.
- de Hoop, M.V., Holman, S.F., Iversen, E., Lassas, M. & Ursin, B., 2014. Reconstruction of a conformally euclidean metric from local boundary diffraction travel times, *SIAM J. Math. Anal.*, **46**(6), 3705–3726.
- de Hoop, M.V., Holman, S.F., Iversen, E., Lassas, M. & Ursin, B., 2015. Recovering the isometry type of a Riemannian manifold from local boundary diffraction travel times, *J. Math. Pure Appl.*, **103**(3), 830–848.
- Douma, H. & de Hoop, M.V., 2006. Explicit expressions for prestack map time migration in isotropic and VTI media and the applicability of map depth migration in heterogeneous anisotropic media, *Geophysics*, **71**(01), S13–S28.
- Duchkov, A.A. & de Hoop, M.V., 2010. Extended isochron rays in prestack depth (map) migration, *Geophysics*, **75**(4), S139–S150.
- Farin, G., Hoschek, J. & Kim, M.-S., 2002. *Handbook of Computer Aided Geometric Design*, Elsevier.
- Farra, V. & Madariaga, R., 1987. Seismic waveform modeling in heterogeneous media by ray perturbation theory, *J. geophys. Res.*, **92**, 2697–2712.
- Foss, S.-K. & Ursin, B., 2004. 2.5D modelling, inversion and angle migration in anisotropic media, *Geophys. Prospect.*, **52**, 65–84.
- Foss, S.-K., Ursin, B. & Sollid, A., 2004. A practical approach to PP seismic angle tomography, *Geophys. Prospect.*, **52**, 663–669.
- Foss, S.-K., de Hoop, M.V. & Ursin, B., 2005. Linearized 2.5-dimensional parameter imaging inversion in anisotropic elastic media, *Geophys. J. Int.*, **161**, 722–738.
- Gajewski, D. & Pšenčík, I., 1990. Vertical seismic profile synthetics by dynamic ray tracing in laterally varying layered anisotropic structures, *J. geophys. Res.*, **95**(B7), 11301–11315.
- Gjøystdal, H., Reinhardsen, J.E. & Ursin, B., 1984. Traveltime and wavefront curvature calculations in three-dimensional inhomogeneous layered media with curved interfaces, *Geophysics*, **49**, 1466–1494.
- Goldin, S.V. & Duchkov, A.A., 2003. Seismic wave field in the vicinity of caustics and higher-order travel time derivatives, *Stud. Geophys. Geod.*, **47**, 521–544.
- Hamilton, W.R., 1837. Third supplement to an essay on the theory of systems of rays, *Trans. R. Ir. Acad.*, **17**, 1–144, read January 23, 1832, and October 22, 1832.
- Hanyga, A., 1982. Dynamic ray tracing in an anisotropic medium, *Tectonophysics*, **90**, 243–251.
- Hubral, P., 1977. Time migration—some ray theoretical aspects, *Geophys. Prospect.*, **25**(04), 738–745.
- Hubral, P., 1983. Computing true amplitude reflections in a laterally inhomogeneous earth, *Geophysics*, **48**, 1051–1062.
- Hubral, P., Schleicher, J. & Tygel, M., 1992. Three-dimensional paraxial ray properties: part I. Basic relations, *J. Seism. Explor.*, **1**, 265–279.
- Iversen, E., 2004a. Reformulated kinematic and dynamic ray tracing systems for arbitrarily anisotropic media, *Stud. Geophys. Geod.*, **48**, 1–20.
- Iversen, E., 2004b. The isochron ray in seismic modeling and imaging, *Geophysics*, **69**(4), 1053–1070.
- Iversen, E. & Gjøystdal, H., 1996. Event-oriented velocity estimation based on prestack data in time or depth domain, *Geophys. Prospect.*, **44**(4), 643–686.
- Iversen, E. & Pšenčík, I., 2008. Ray tracing and inhomogeneous dynamic ray tracing for anisotropy specified in curvilinear coordinates, *Geophys. J. Int.*, **174**, 316–330.
- Iversen, E. & Tygel, M., 2008. Image-ray tracing for joint 3D seismic velocity estimation and time-to-depth conversion, *Geophysics*, **73**(3), P99–P114.
- Iversen, E., Tygel, M., Ursin, B. & de Hoop, M.V., 2012. Kinematic time migration and demigration of reflections in pre-stack seismic data, *Geophys. J. Int.*, **189**(3), 1635–1666.
- Jäger, R., Mann, J., Höcht, G. & Hubral, P., 2001. Common-reflection-surface stack: Image and attributes, *Geophysics*, **66**(1), 97–109.
- Klimeš, L., 1994. Transformations for dynamic ray tracing in anisotropic media, *Wave Motion*, **20**, 261–272.
- Klimeš, L., 2002a. Second-order and higher-order perturbations of travel time in isotropic and anisotropic media, *Stud. Geophys. Geod.*, **46**(2), 213–248.
- Klimeš, L., 2002b. Relation of the wave-propagation metric tensor to the curvatures of the slowness and ray-velocity surfaces, *Stud. Geophys. Geod.*, **46**(3), 589–597.
- Klimeš, L., 2006a. Spatial derivatives and perturbation derivatives of amplitude in isotropic and anisotropic media, *Stud. Geophys. Geod.*, **50**(3), 417–430.
- Klimeš, L., 2006b. Ray-centred coordinate systems in anisotropic media, *Stud. Geophys. Geod.*, **50**, 431–447.
- Moser, T.J. & Červený, V., 2007. Paraxial ray methods for anisotropic inhomogeneous media, *Geophys. Prospect.*, **55**, 21–37.
- Paternain, G.P., 2012. *Geodesic Flows*, Vol. **180**, Springer Science & Business Media.
- Rabbel, W., Bittner, R. & Gelchinsky, B., 1991. Seismic mapping of complex reflectors with the common-reflecting-element method (CRE method), *Phys. Earth planet. Inter.*, **67**(1), 200–210.
- Schleicher, J., Tygel, M. & Hubral, P., 2007. *Seismic True-Amplitude Imaging*, Society of Exploration Geophysicists.
- Shen, Z., 2001. *Lectures on Finsler Geometry*, World Scientific.
- Sollid, A. & Ursin, B., 2003. Scattering-angle migration of ocean-bottom seismic data in weakly anisotropic media, *Geophysics*, **68**(2), 641–655.
- Stolk, C.C. & de Hoop, M.V., 2002. Microlocal analysis of seismic inverse scattering in anisotropic elastic media, *Commun. Pure Appl. Math.*, **55**(3), 261–301.
- Stolk, C.C. & de Hoop, M.V., 2006. Seismic inverse scattering in the downward continuation approach, *Wave Motion*, **43**(7), 579–598.
- Thomsen, L., 1986. Weak elastic anisotropy, *Geophysics*, **51**, 1954–1966.
- Tsvankin, I., 2013. *Seismic Signatures and Analysis of Reflection Data in Anisotropic Media*, 3rd edn, Society of Exploration Geophysicists.
- Tygel, M., Ursin, B., Iversen, E. & de Hoop, M.V., 2012. Estimation of geological dip and curvature from time-migrated zero-offset reflections in heterogeneous anisotropic media, *Geophys. Prospect.*, **60**(2), 201–216.
- Ursin, B., 1982. Quadratic wavefront and travel time approximations in inhomogeneous layered media with curved interfaces, *Geophysics*, **47**, 1012–1021.
- Ursin, B., 2004. Parameter inversion and angle migration in anisotropic elastic media, *Geophysics*, **69**, 1125–1142.

APPENDIX A: RELATIONS FOR DERIVATIVES OF THE HAMILTONIAN

The considered Hamiltonian \mathcal{H} satisfies the relations

$$p_i \frac{\partial^2 \mathcal{H}}{\partial p_i \partial p_j} = \frac{\partial \mathcal{H}}{\partial p_j}, \quad p_i \frac{\partial^2 \mathcal{H}}{\partial p_i \partial x_j} = 2 \frac{\partial \mathcal{H}}{\partial x_j}. \quad (\text{A1})$$

These relations hold on the reference ray, Ω , as well as in its vicinity.

For the third-order derivatives of \mathcal{H} we have the following general relations,

$$p_i \frac{\partial^3 \mathcal{H}}{\partial p_i \partial x_j \partial x_k} = 2 \frac{\partial^2 \mathcal{H}}{\partial x_j \partial x_k}, \quad (\text{A2})$$

$$p_i \frac{\partial^3 \mathcal{H}}{\partial p_i \partial p_j \partial x_k} = \frac{\partial^2 \mathcal{H}}{\partial p_j \partial x_k}, \quad (\text{A3})$$

$$p_i p_j \frac{\partial^3 \mathcal{H}}{\partial p_i \partial p_j \partial x_k} = 2 \frac{\partial \mathcal{H}}{\partial x_k}, \quad (\text{A4})$$

It is important to note that \mathcal{H} may have nonzero derivatives of order three and higher in the slowness components. In particular, it follows from eq. (A1) that the third- and fourth-order derivatives must satisfy

$$p_i \frac{\partial^3 \mathcal{H}}{\partial p_i \partial p_j \partial p_k} = 0 \quad (\text{A5})$$

and

$$p_i \frac{\partial^4 \mathcal{H}}{\partial p_i \partial p_j \partial p_k \partial p_l} = - \frac{\partial^3 \mathcal{H}}{\partial p_j \partial p_k \partial p_l}. \quad (\text{A6})$$

APPENDIX B: TWO-PARAMETRIC SYSTEM OF RAYS

Consider a ray-parameter space (γ_A) with dimension $N_\gamma = 2$, and the associated ray-coordinate system, $(\gamma_1, \gamma_2, \tau)$. In a local region around a point on the reference ray we assume that a one-to-one mapping exists between the Cartesian coordinates (x_1, x_2, x_3) and the ray coordinates $(\gamma_1, \gamma_2, \tau)$.

B1 First-order transformation between Cartesian coordinates and ray coordinates

To the first order, the quantities on Ω describing the transformation from ray coordinates to Cartesian coordinates are

$$\frac{\partial x_i}{\partial \gamma_A} = Q_{iA}, \quad \frac{\partial x_i}{\partial \tau} = v_i = Q_{i3}. \quad (\text{B1})$$

For the inverse transformation we use the functions $\gamma_A = \gamma_A(\mathbf{x})$ and $\tau = \tau(\mathbf{x})$, with the first-order derivatives

$$\frac{\partial \gamma_A}{\partial x_i} = Q_{Ai}^\dagger, \quad \frac{\partial \tau}{\partial x_i} = p_i = Q_{3i}^\dagger. \quad (\text{B2})$$

The quantities in eqs (B1) and (B2) form the 3×3 matrices

$$\widehat{\mathbf{Q}} = \begin{bmatrix} \mathbf{Q} & \mathbf{v} \end{bmatrix}, \quad \widehat{\mathbf{Q}}^{-1} = \widehat{\mathbf{Q}}^\dagger = \begin{bmatrix} \mathbf{Q}^\dagger \\ \mathbf{p}^T \end{bmatrix}. \quad (\text{B3})$$

Based on the relations

$$Q_{ai}^\dagger Q_{ib} = \delta_{ab} \quad Q_{ia} Q_{aj}^\dagger = \delta_{ij}, \quad (\text{B4})$$

we can list some intrinsic properties of the transformation between ray coordinates and Cartesian coordinates on Ω ,

$$\begin{aligned} p_i Q_{iA} &= 0, \\ v_i Q_{Ai}^\dagger &= 0, \\ Q_{Ai}^\dagger Q_{iB} &= \delta_{AB}, \\ p_i v_i &= 1, \\ Q_{Ai}^\dagger Q_{jA} &= \alpha_{ij}, \end{aligned} \quad (\text{B5})$$

where α_{ij} is given by

$$\alpha_{ij} = \delta_{ij} - p_i v_j. \quad (\text{B6})$$

B2 Second-order derivatives of traveltime

Consider a specific traveltime function $\tau(\mathbf{x})$ which corresponds to two paraxial ray parameters $\gamma_A(\mathbf{x})$, $A = 1, 2$. The first- and second-order derivatives of τ on Ω are

$$\frac{\partial \tau}{\partial x_i} = p_i, \quad (B7)$$

$$\frac{\partial^2 \tau}{\partial x_i \partial x_j} = M_{ij}. \quad (B8)$$

The quantities M_{ij} are forming the 3×3 matrix \mathbf{M} .

Using the chain rule for differentiation, we then obtain

$$\frac{\partial^2 \tau}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_j} \{p_i [\gamma_A(\mathbf{x}), \tau(\mathbf{x})]\} = \frac{\partial p_i}{\partial \gamma_A} \frac{\partial \gamma_A}{\partial x_j} + \frac{\partial p_i}{\partial \tau} \frac{\partial \tau}{\partial x_j},$$

and consequently,

$$M_{ij} = P_{iA} Q_{Aj}^\dagger + \eta_i p_j, \quad \mathbf{M} = \mathbf{P} \mathbf{Q}^\dagger + \boldsymbol{\eta} \mathbf{p}^T. \quad (B9)$$

The last relation describes how the 3×3 matrix of second derivatives of traveltime, \mathbf{M} , can be obtained from the 3×2 matrices \mathbf{P} and \mathbf{Q} computed using the (standard) Hamilton–Jacobi perturbation equations. The 2×3 matrix \mathbf{Q}^\dagger is a submatrix of matrix $\widehat{\mathbf{Q}}^{-1}$ in eq. (B3).

APPENDIX C: CONSTRAINT RELATION FOR THIRD-ORDER DERIVATIVES OF PHASE-SPACE PERTURBATIONS: PLANE WAVE FRONT

Based on eq. (56), the constraint relation for third-order derivatives of phase-space perturbations, assuming two ray parameters, can be written

$$v_i P_{iABC} = \eta_i Q_{iABC} + Q_{iA} \dot{P}_{iBC} - P_{iA} \dot{Q}_{iBC} + Q_{iAB} \dot{P}_{iC} - P_{iAB} \dot{Q}_{iC} + Q_{iAC} \dot{P}_{iB} - P_{iAC} \dot{Q}_{iB}. \quad (C1)$$

In this appendix we derive a special version of eq. (C1), pertaining specifically to a plane wave front. On the way we use relations for the derivatives of the Hamiltonian given in Appendix A. It is convenient to use abbreviated forms for partial derivatives in these derivations, namely,

$$\partial_k^1 = \frac{\partial}{\partial x_k}, \quad \partial_k^2 = \frac{\partial}{\partial p_k}, \quad (C2)$$

and

$$\frac{\partial \mathcal{H}}{\partial x_k} = \mathcal{H}_{,k}^1, \quad \frac{\partial \mathcal{H}}{\partial p_k} = \mathcal{H}_{,k}^2, \quad \frac{\partial^2 \mathcal{H}}{\partial x_k \partial x_l} = \mathcal{H}_{,kl}^{11}, \quad \frac{\partial^2 \mathcal{H}}{\partial x_k \partial p_l} = \mathcal{H}_{,kl}^{12}, \quad \frac{\partial^2 \mathcal{H}}{\partial p_k \partial p_l} = \mathcal{H}_{,kl}^{22}, \quad (C3)$$

and so forth.

We find expressions for the various quantities on the right-hand side of eq. (C1). In particular, we can use the ODEs in eq. (37) to obtain

$$\dot{Q}_{iBC} = \mathcal{H}_{,mi}^{12} Q_{mBC} + \mathcal{H}_{,im}^{22} P_{mBC} + Q_{mB} [Q_{nC} \partial_n^1 + P_{nC} \partial_n^2] \mathcal{H}_{,mi}^{12} + P_{mB} [Q_{nC} \partial_n^1 + P_{nC} \partial_n^2] \mathcal{H}_{,im}^{22},$$

which yields

$$\dot{Q}_{iBC} = \mathcal{H}_{,mi}^{12} Q_{mBC} + \mathcal{H}_{,im}^{22} P_{mBC} + \mathcal{H}_{,mni}^{112} Q_{mB} Q_{nC} + \mathcal{H}_{,nmi}^{122} P_{mB} Q_{nC} + \mathcal{H}_{,mni}^{122} Q_{mB} P_{nC} + \mathcal{H}_{,mni}^{222} P_{mB} P_{nC}. \quad (C4)$$

Likewise,

$$\dot{P}_{iBC} = -\mathcal{H}_{,im}^{11} Q_{mBC} - \mathcal{H}_{,im}^{12} P_{mBC} + Q_{mB} [Q_{nC} \partial_n^1 + P_{nC} \partial_n^2] (-\mathcal{H}_{,im}^{11}) + P_{mB} [Q_{nC} \partial_n^1 + P_{nC} \partial_n^2] (-\mathcal{H}_{,im}^{12}),$$

so therefore,

$$\dot{P}_{iBC} = -\mathcal{H}_{,im}^{11} Q_{mBC} - \mathcal{H}_{,im}^{12} P_{mBC} - \mathcal{H}_{,imn}^{111} Q_{mB} Q_{nC} - \mathcal{H}_{,imn}^{112} P_{mB} Q_{nC} - \mathcal{H}_{,imn}^{112} Q_{mB} P_{nC} - \mathcal{H}_{,imn}^{122} P_{mB} P_{nC}. \quad (C5)$$

Using eqs (C4) and (C5) we obtain,

$$P_{iA} \dot{Q}_{iBC} = \mathcal{H}_{,mi}^{12} P_{iA} Q_{mBC} + \mathcal{H}_{,im}^{22} P_{iA} P_{mBC} + \mathcal{H}_{,mni}^{112} P_{iA} Q_{mB} Q_{nC} + \mathcal{H}_{,nmi}^{122} P_{iA} P_{mB} Q_{nC} + \mathcal{H}_{,mni}^{122} P_{iA} Q_{mB} P_{nC} + \mathcal{H}_{,mni}^{222} P_{iA} P_{mB} P_{nC}, \quad (C6)$$

$$Q_{iA} \dot{P}_{iBC} = -\mathcal{H}_{,im}^{11} Q_{iA} Q_{mBC} - \mathcal{H}_{,im}^{12} Q_{iA} P_{mBC} - \mathcal{H}_{,imn}^{111} Q_{iA} Q_{mB} Q_{nC} - \mathcal{H}_{,imn}^{112} Q_{iA} P_{mB} Q_{nC} - \mathcal{H}_{,imn}^{112} Q_{iA} Q_{mB} P_{nC} - \mathcal{H}_{,imn}^{122} Q_{iA} P_{mB} P_{nC}. \quad (C7)$$

C1 A special case

Assume that the quantities P_{iA} and P_{iAB} can be expressed in terms of the slowness components p_i as

$$P_{iA} = p_i E_A, \quad P_{iAB} = p_i F_{AB}. \quad (\text{C8})$$

Applying eq. (C8) in eqs (C6) and (C7) yields

$$P_{iA} \dot{Q}_{iBC} = \mathcal{H}_{,mi}^{,12} p_i Q_{mBC} E_A + \mathcal{H}_{,im}^{,22} p_i p_m E_A F_{BC} + \mathcal{H}_{,mni}^{,112} p_i Q_{mB} Q_{nC} E_A + \mathcal{H}_{,nmi}^{,122} p_i p_m Q_{nC} E_A E_B + \mathcal{H}_{,mni}^{,122} p_i p_n Q_{mB} E_A E_C, \quad (\text{C9})$$

$$Q_{iA} \dot{P}_{iBC} = -\mathcal{H}_{,im}^{,11} Q_{iA} Q_{mBC} - \mathcal{H}_{,im}^{,12} p_m Q_{iA} F_{BC} - \mathcal{H}_{,imn}^{,111} Q_{iA} Q_{mB} Q_{nC} - \mathcal{H}_{,imn}^{,112} p_m Q_{iA} Q_{nC} E_B - \mathcal{H}_{,imn}^{,112} p_n Q_{iA} Q_{mB} E_C - \mathcal{H}_{,imn}^{,122} p_m p_n Q_{iA} E_B E_C. \quad (\text{C10})$$

In eq. (C9) we have also applied the property (A5) of the Hamiltonian, which eliminates the term including the derivative $\mathcal{H}_{,imn}^{,222}$.

Using relations (A2)–(A4) in eqs (C9) and (C10) leads to

$$P_{iA} \dot{Q}_{iBC} = -2\eta_m Q_{mBC} E_A + E_A F_{BC} + 2\mathcal{H}_{,mn}^{,11} Q_{mB} Q_{nC} E_A - 2\eta_n Q_{nC} E_A E_B - 2\eta_m Q_{mB} E_A E_C, \quad (\text{C11})$$

$$Q_{iA} \dot{P}_{iBC} = -\mathcal{H}_{,im}^{,11} Q_{iA} Q_{mBC} + 2\eta_i Q_{iA} F_{BC} - \mathcal{H}_{,imn}^{,111} Q_{iA} Q_{mB} Q_{nC} - 2\mathcal{H}_{,in}^{,11} Q_{iA} Q_{nC} E_B - 2\mathcal{H}_{,im}^{,11} Q_{iA} Q_{mB} E_C + 2\eta_i Q_{iA} E_B E_C. \quad (\text{C12})$$

C2 Plane wave front

The next step is to apply in eqs (C11) and (C12) specifically the conditions for a plane wave front. This means to set

$$E_A = \eta_i \mathcal{E}_{iA}, \quad (\text{C13})$$

$$F_{AB} = (-U_{ij} + 3\eta_i \eta_j) \mathcal{E}_{iA} \mathcal{E}_{jB}, \quad (\text{C14})$$

$$Q_{iA} = \mathcal{E}_{iA}, \quad (\text{C15})$$

$$Q_{iAB} = 0. \quad (\text{C16})$$

This yields,

$$P_{iA} \dot{Q}_{iBC} = (-U_{jk} + 3\eta_j \eta_k) \eta_i \mathcal{E}_{iA} \mathcal{E}_{jB} \mathcal{E}_{kC} + 2\mathcal{H}_{,mn}^{,11} \mathcal{E}_{mB} \mathcal{E}_{nC} \eta_i \mathcal{E}_{iA} - 2\eta_n \mathcal{E}_{nC} \eta_i \mathcal{E}_{iA} \eta_j \mathcal{E}_{jB} - 2\eta_m \mathcal{E}_{mB} \eta_i \mathcal{E}_{iA} \eta_k \mathcal{E}_{kC} = (\eta_i U_{jk} - \eta_i \eta_j \eta_k) \mathcal{E}_{iA} \mathcal{E}_{jB} \mathcal{E}_{kC} \quad (\text{C17})$$

and

$$Q_{iA} \dot{P}_{iBC} = 2\eta_i \mathcal{E}_{iA} (-U_{jk} + 3\eta_j \eta_k) \mathcal{E}_{jB} \mathcal{E}_{kC} - \mathcal{H}_{,imn}^{,111} \mathcal{E}_{iA} \mathcal{E}_{mB} \mathcal{E}_{nC} - 2\mathcal{H}_{,in}^{,11} \mathcal{E}_{iA} \mathcal{E}_{nC} \eta_j \mathcal{E}_{jB} - 2\mathcal{H}_{,im}^{,11} \mathcal{E}_{iA} \mathcal{E}_{mB} \eta_k \mathcal{E}_{kC} + 2\eta_i \eta_j \eta_k \mathcal{E}_{iA} \mathcal{E}_{jB} \mathcal{E}_{kC} = (8\eta_i \eta_j \eta_k - 2\eta_i U_{jk} - 2\eta_j U_{ik} - 2\eta_k U_{ij} - U_{ijk}) \mathcal{E}_{iA} \mathcal{E}_{jB} \mathcal{E}_{kC}. \quad (\text{C18})$$

Taking the difference yields,

$$Q_{iA} \dot{P}_{iBC} - P_{iA} \dot{Q}_{iBC} = (9\eta_i \eta_j \eta_k - 3\eta_i U_{jk} - 2\eta_j U_{ik} - 2\eta_k U_{ij} - U_{ijk}) \mathcal{E}_{iA} \mathcal{E}_{jB} \mathcal{E}_{kC}. \quad (\text{C19})$$

In addition, we have

$$P_{iAB} \dot{Q}_{iC} = p_i F_{AB} (\mathcal{H}_{,mi}^{,12} Q_{mC} + \mathcal{H}_{,im}^{,22} P_{mC}) = (\mathcal{H}_{,mi}^{,12} p_i Q_{mC} + \mathcal{H}_{,im}^{,22} p_i p_m E_C) F_{AB} = (-2\eta_m Q_{mC} + E_C) F_{AB}. \quad (\text{C20})$$

Inserting as above yields

$$P_{iAB} \dot{Q}_{iC} = (-3\eta_i \eta_j \eta_k + \eta_k U_{ij}) \mathcal{E}_{iA} \mathcal{E}_{jB} \mathcal{E}_{kC}. \quad (\text{C21})$$

Finally, we apply the results (C18), (C19), and (C21) in eq. (C1), which yields the constraint relation for a plane wave front,

$$v_i P_{iABC} = Q_{iA} \dot{P}_{iBC} - P_{iA} \dot{Q}_{iBC} - P_{iAB} \dot{Q}_{iC} - P_{iAC} \dot{Q}_{iB} = (15\eta_i \eta_j \eta_k - 3\eta_i U_{jk} - 3\eta_j U_{ik} - 3\eta_k U_{ij} - U_{ijk}) \mathcal{E}_{iA} \mathcal{E}_{jB} \mathcal{E}_{kC}. \quad (\text{C22})$$