Geophysical Journal International

Geophys. J. Int. (2021) **226,** 893–927
Advance Access publication 2021 April 15
GJI Seismology

doi: 10.1093/gji/ggab151

Higher-order Hamilton–Jacobi perturbation theory for anisotropic heterogeneous media: transformation between Cartesian and ray-centred coordinates

Einar Iversen,[1] Bjørn Ursin,[2] Teemu Saksala,[3] Joonas Ilmavirta[4] and Maarten V. de Hoop[5]

[1] *Department of Earth Science, University of Bergen, P.O. Box* 7803*, N-*5020 *Bergen, Norway. E-mail: einar.iversen@uib.no*
[2] *Department of Electronic Systems, Norwegian University of Science and Technology (NTNU), O.S. Bragstadsplass* 2B*, NO-*7491 *Trondheim, Norway*
[3] *Department of Mathematics, Box* 8205*, North Carolina State University, Raleigh, NC* 27695-8205*, USA*
[4] *Unit of Computing Sciences, Tampere University, Kalevantie* 4*, FI-*33014*,* 33100 Tampere*, Finland*
[5] *Department of Computational and Applied Mathematics, Rice University,* 6100 *Main MS-*134*, Houston, TX* 77005-1892*, USA*

Accepted 2021 April 13. Received 2021 April 12; in original form 2020 December 22

SUMMARY

Within the field of seismic modelling in anisotropic media, dynamic ray tracing is a powerful technique for computation of amplitude and phase properties of the high-frequency Green's function. Dynamic ray tracing is based on solving a system of Hamilton–Jacobi perturbation equations, which may be expressed in different 3-D coordinate systems. We consider two particular coordinate systems; a Cartesian coordinate system with a fixed origin and a curvilinear ray-centred coordinate system associated with a reference ray. For each system we form the corresponding 6-D phase spaces, which encapsulate six degrees of freedom in the variation of position and momentum. The formulation of (conventional) dynamic ray tracing in ray-centred coordinates is based on specific knowledge of the first-order transformation between Cartesian and ray-centred phase-space perturbations. Such transformation can also be used for defining initial conditions for dynamic ray tracing in Cartesian coordinates and for obtaining the coefficients involved in two-point traveltime extrapolation. As a step towards extending dynamic ray tracing in ray-centred coordinates to higher orders we establish detailed information about the higher-order properties of the transformation between the Cartesian and ray-centred phase-space perturbations. By numerical examples, we (1) visualize the validity limits of the ray-centred coordinate system, (2) demonstrate the transformation of higher-order derivatives of traveltime from Cartesian to ray-centred coordinates and (3) address the stability of function value and derivatives of volumetric parameters in a higher-order representation of the subsurface model.

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

1 INTRODUCTION

For more than 40 years, dynamic ray tracing has been a powerful method to compute important amplitude and phase attributes of high-frequency Green's functions. Dynamic ray tracing can be expressed in Cartesian coordinates, in ray-centred coordinates, and in generally curvilinear coordinates (Červený 2001). The basic idea is to formulate a system of ordinary differential equations, Hamilton-Jacobi perturbation equations, by which one continues the first-order derivatives of perturbations in position/slowness as a function of traveltime, say, along a reference ray. These derivatives of phase-space perturbations constitute the basis for first-order extrapolation of position/slowness and second-order extrapolation of traveltime in the paraxial region, that is a close neighborhood of the reference ray where the traveltime is single-valued. The first derivatives of position in the dynamic ray tracing system yield the geometrical spreading—the primary contributor to the amplitude on the reference ray. In this paper, we focus on the *coordinates* used in dynamic ray tracing. We also address the consequences for the *model representation* when the transformation between coordinate systems is extended to higher orders. Our work fits in approximation theory, which makes it natural to consider 'higher order' in the context of *accuracy*.

Historically, dynamic ray tracing has often been done in Cartesian coordinates (e.g. Červený 1972, 2001; Farra & Madariaga 1987; Gajewski & Pšenčík 1990; Chapman 2004; Iversen 2004; Červený & Moser 2007; Červený & Pšenčík 2010; Klimeš 2013; Koren & Ravve 2021;

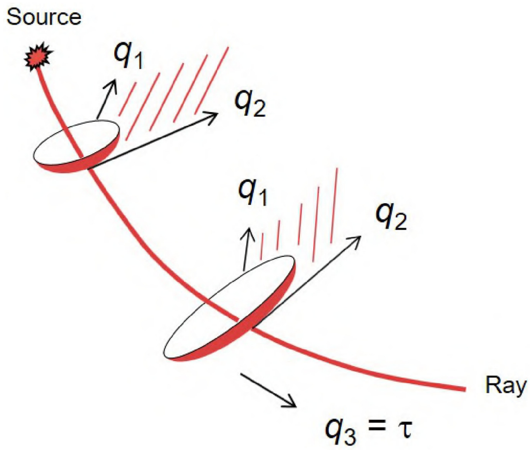

Figure 1. Ray-centred coordinates: A 2-D Cartesian coordinate frame (q_1, q_2) is continued along the selected reference ray. The coordinate q_3 is a monotonic variable along the ray—we take q_3 as the traveltime.

Ravve & Koren 2021) or in ray-centred coordinates (e.g. Popov & Pšenčík 1978; Hanyga 1982; Kendall *et al.* 1992; Červený 2001; Červený *et al.* 2007; Cameron *et al.* 2007; Iversen 2006; Iversen & Tygel 2008; Klimeš 1994, 2006a, 2012, 2019).

Ray-centred coordinates were introduced in seismology by Popov & Pšenčík (1978). Assuming an isotropic medium, a 2-D Cartesian coordinate frame is continued along the reference ray (Fig. 1). The dynamic ray tracing quantities, that is derivatives of the phase-space perturbations, are expressed relative to this moving frame. In this way, ray-centred coordinates interplay naturally with the wave propagation under study, the number of differential equations is limited to a minimum, the initial conditions become simple and intuitive, and numerical errors caused by redundant solutions are absent or reduced to a minimum. In differential geometry, ray-centred coordinates are a first-order approximation of *Fermi coordinates* (e.g. Chavel 2006), which are local coordinates associated with a *geodesic*—a generalized ray. As with Fermi coordinates, the basis vectors of ray-centred coordinates are continued along the reference ray. The basis vectors are grouped in two sets, referred to as the *contra-variant* and *co-variant* bases. Only one set needs to be integrated along the ray; the other set then follows from explicit expressions.

Recently, Iversen *et al.* (2019) extended dynamic ray tracing for anisotropic media to higher orders, using Cartesian coordinates. The main motivation was to attain more robust and accurate extrapolation or interpolation of amplitude and phase attributes of the high-frequency Green's function. Dynamic ray tracing in Cartesian and ray-centred coordinates have different strengths and weaknesses, and both formulations have proven very useful for applications in seismology and seismic exploration. In some modelling approaches, for example the wavefront construction method (Vinje *et al.* 1993), one can utilize the properties of both coordinate systems in a complimentary way. It is therefore valuable to extend also dynamic ray tracing in ray-centred coordinates to higher orders, and to develop a framework that allows to switch between the two formulations. The latter is provided in the current paper—a higher-order transformation of phase-space perturbations between Cartesian and ray-centred coordinates.

Also shown in the paper, the mentioned higher-order transformation can, in principle, be obtained without a higher-order representation of the model parameter functions. The reason is that the higher-order effects can be incorporated by interpolating along the reference ray the phase-space coordinates and their time derivatives. However, the derivatives of the phase-space perturbations considered for transformation may indeed depend explicitly on the higher-order derivatives of the model parameters. These perturbation quantities typically correspond to initial conditions or end results of dynamic ray tracing. Thus, to ensure consistency between transformation coefficients and the quantities to be transformed it is recommended that the same, higher-order, model representation is used both for computation of the transformation coefficients and the derivatives of the phase-space perturbations.

We utilize the properties of basic splines, or *B-splines* (de Boor 1972), to ensure a consistent higher-order transformation of phase-space perturbations between Cartesian and ray-centred coordinates. The term *spline* originated in the ship construction industry in England and dates back at least to 1752 (Farin *et al.* 2002). The spline was a mechanical device of wood, used for drawing smooth curves. The objective was to make the hull of the ship smooth. Later, in the 20th century, the science of curve fitting and computer-aided geometric design was driven by pioneers working in the car industry (e.g. Carl de Boor, Paul de Casteljau, Pierre Bézier). From the 1970s and onwards splines have been popular for use in computer graphics and geometric modelling in general (Bartles *et al.* 1987), and in solid earth geophysics in particular (e.g. Gjøystdal *et al.* 1985).

A cornerstone in the theory of B-splines is the *de Casteljau algorithm* (Nowak 2011), which is used for recursive determination of *Bernstein polynomials* (named after Sergei Natanovich Bernstein). With de Casteljau's algorithm one can, in principle, compute safely the value of a function representation and its derivatives to any order. In this paper we use *quintic* (fifth-degree) B-splines, which encapsulate C^4 continuity—it is then guaranteed that derivatives up to order four are continuous. In contrast, for conventional dynamic ray tracing and associated transformation between Cartesian and ray-centred coordinates, a cubic B-spline representation (continuity of type C^2) is sufficient.

As key results, we provide expressions for the relevant second-, third- and fourth-order coefficients in the transformation from ray-centred to Cartesian phase-space perturbations and vice versa. However, to avoid using long expressions for the higher-order coefficients of the inverse transformation, Cartesian to ray-centred, we also emphasize the option of computing these implicitly from the coefficients of the forward transformation, by means of the *symplectic property* (e.g. Červený 2001). The results obtained here are further used in a companion paper on higher-order dynamic ray tracing in ray-centred coordinates (Iversen *et al.* 2021).

The higher-order transformations are valuable in the following situations.

(i) When performing *dynamic ray tracing in ray-centred coordinates* the transformations are needed in the design of the coefficients of the system of Hamilton–Jacobi perturbation equations. These coefficients then belong to ray-centred coordinates but will typically depend on a set of property functions defined with respect to Cartesian coordinates. Moreover, at the final point the results will often need to be transformed to Cartesian coordinates, using the transformations presented in this paper.

(ii) *Dynamic ray tracing in Cartesian coordinates* yields the results directly in the Cartesian coordinate system. In some situations, however, it may be useful to do further analysis and computations in ray-centred coordinates.

(iii) *Initial conditions* for the Hamilton–Jacobi perturbation equations can be specified in Cartesian coordinates, ray-centred coordinates, or by other means. A transformation of such conditions between different coordinates may then be required, depending on the chosen coordinate system for the Hamilton–Jacobi perturbation equations. The initial conditions in ray-centred coordinates are often simple and intuitive, while this is generally not the case in Cartesian coordinates.

A general impact of the paper is that the presented transformation equations are of importance in paraxial ray methods. Furthermore, the provided higher-order transformation could be a stepping stone to establishing a transformation where the ray-centred coordinates are replaced by a more general curvilinear coordinate system. We expect the latter would be valuable in applications related to solid earth geophysics.

The paper is organized as follows. We first describe the notion of a 6-D phase space consisting of position and momentum, in ray-centred coordinates. Important in this context is the two sets of basis vectors and the continuation of these along the reference ray. Next, we derive explicit expressions for all relevant coefficients of the forward and inverse transformations, from the ray-centred to the Cartesian phase-space perturbations and vice versa. The resulting transformation coefficients are subsequently used to derive a framework for transformation of the higher-order derivatives of traveltime between ray-centred and Cartesian coordinates. Thereafter, we use the transformation coefficients to address the validity region of ray-centred coordinates. In a numerical examples section we show applications of the derived theory, for isotropic and vertical transversely isotropic (VTI) versions of the Marmousi model. In these examples, we (1) visualize the validity limits of the ray-centred coordinate system and (2) show some subtleties of the transformation of higher-order derivatives of traveltime from Cartesian to ray-centred coordinates. Later in the section we address the stability of the function value and the derivatives computed using a quintic B-spline representation.

Notes on the nomenclature—We use component and vector/matrix notations in parallel. Components of vectors, matrices and tensors are specified by lower- and uppercase subscript indices. The lowercase indices $a, b, c, \ldots, p, q$ run from 1 to 3, while corresponding uppercase indices $A, B, C, \ldots$ take the values 1, 2 only. In the remaining part of the alphabet the indices $r, s, t, \ldots$ run from 1 to 6. For equations in component notation we use Einstein's summation convention. Vectors are are written as lowercase bold symbols, $\mathbf{a}$, or in terms of components only, a_i. A vector $\mathbf{a}$ with N components is equivalently understood as an N-tuple $\mathbf{a} = (a_i) = (a_1, a_2, \ldots, a_N)$ or as an $N \times 1$ column matrix—the specific meaning follows from the context. Multicolumn matrices are written either in bold uppercase, $\mathbf{H}$, or in component notation, H_{ia}. To connect the two forms, we write $\mathbf{H} = \{H_{ia}\}$. The symbol † is used to signify components of an inverse matrix, for example $\mathbf{H}^{-1} = \{H^{\dagger}_{ai}\}$. In the context of derivatives of phase-space perturbations continued along the reference ray we use the perturbation symbol δ. Perturbations performed locally are signified by the symbol Δ. For overview of the mathematical symbols used in the paper, see Table 1.

2 PHASE-SPACE COORDINATES

We review and discuss the phase spaces arising from Cartesian and ray-centred coordinates.

2.1 Cartesian phase-space coordinates

For a fixed 3-D Cartesian coordinate system, $(x_i) = (x_1, x_2, x_3)$, we consider a position vector, $\mathbf{x} = (x_i)$, and a momentum vector, $\mathbf{p} = (p_i)$, with the measurement unit of inverse velocity. Because of the latter property, the vector $\mathbf{p}$ is commonly referred to as the *slowness vector* or briefly as just the *slowness*. The slowness vector components p_1, p_2 and p_3 can be combined with the position-vector components x_1, x_2 and x_3 to form the 6-D domain

$$(w_r) = (x_i, p_j) = (x_1, x_2, x_3, p_1, p_2, p_3); \quad r = 1, 2, \ldots, 6; \quad i, j = 1, 2, 3; \tag{1}$$

known as the *phase space* in Cartesian coordinates. In this domain all six coordinates vary freely.

The notion of six freely varying phase-space components is fundamental in the Hamiltonian formulation of ray theory. One can consider the phase space as a workspace in which we seek the ray solutions. The computed solutions form a subspace, a *hypersurface*, with five degrees of freedom. The hypersurface is defined by the Hamilton-Jacobi equation, introduced in the next subsection.

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
τ		Traveltime along the ray Ω
τ_0		Traveltime at the initial point of the ray Ω
c		Phase velocity
$\mathbf{c} = (c_i)$	3	Phase-velocity vector
$\mathbf{n} = (n_i)$	3	Normalized phase-velocity vector
$\mathbf{v} = (v_i)$	3	Ray-velocity (group-velocity) vector
$\boldsymbol{\eta} = (\eta_i)$	3	Time derivative of the slowness vector $\mathbf{p}$
$\boldsymbol{\alpha} = \{\alpha_{ij}\}$	3×3	Projection operator with respect to the wave-propagation metric tensor
(q_1, q_2, q_3)	3	Ray-centred coordinate system
$\mathbf{q} = (q_a)$	3	Position vector of the ray-centred coordinate system
$\mathbf{p}^{(q)} = (p_a^{(q)})$	3	Momentum vector of the ray-centred coordinate system
$\mathbf{w}^{(q)} = (w_r^{(q)})$ $= (q_a, p_b^{(q)})$	6	Phase-space vector of the ray-centred coordinate system
$(v_a^{(q)})$	3	Ray-velocity (group-velocity) vector, in ray-centred coordinates
$(\eta_a^{(q)})$	3	Time derivative of the momentum vector $\mathbf{p}^{(q)}$
$\boldsymbol{\mathcal{E}} = \{\mathcal{E}_{iA}\}$ $= [\mathbf{e}_1\ \mathbf{e}_2]$	3×2	Contra-variant (paraxial) basis of the ray-centred coordinate system
$\mathbf{H} = \{H_{ia}\}$ $= [\boldsymbol{\mathcal{E}}\ \mathbf{v}]$	3×3	Coefficients of coordinate transformation, ray-centred to Cartesian, first order
$\{H_{iab}\}$	$3 \times 3 \times 3$	Coefficients of coordinate transformation, ray-centred to Cartesian, second order
$\{H_{iabc}\}$	$3 \times 3 \times 3 \times 3$	Coefficients of coordinate transformation, ray-centred to Cartesian, third order
		...
$\boldsymbol{\mathcal{F}} = \{\mathcal{F}_{iA}\}$	3×2	Co-variant (paraxial) basis of the ray-centred coordinate system
$\mathbf{H}^{-1} = \{H_{ai}^{\dagger}\}$ $= [\boldsymbol{\mathcal{F}}\ \mathbf{p}]^T$	3×3	Coefficients of coordinate transformation, Cartesian to ray-centred, first order
$\{H_{aij}^{\dagger}\}$	$3 \times 3 \times 3$	Coefficients of coordinate transformation, Cartesian to ray-centred, second order
$\{H_{aijk}^{\dagger}\}$	$3 \times 3 \times 3 \times 3$	Coefficients of coordinate transformation, Cartesian to ray-centred, third order
		...
$\mathcal{T}(\mathbf{x}), \mathcal{T}(\mathbf{q})$		A general time function
$\tau(\mathbf{x}), \tau(\mathbf{q})$		A specific traveltime function
$\{M_{ij}\}$	3×3	Derivatives of traveltime, Cartesian coordinates, second order
$\{M_{ijk}\}$	$3 \times 3 \times 3$	Derivatives of traveltime, Cartesian coordinates, third order
$\{M_{ijkl}\}$	$3 \times 3 \times 3 \times 3$	Derivatives of traveltime, Cartesian coordinates, fourth order
		...
$\{\mathcal{M}_{ab}\}$	3×3	Derivatives of traveltime, ray-centred coordinates, second order
$\{\mathcal{M}_{abc}\}$	$3 \times 3 \times 3$	Derivatives of traveltime, ray-centred coordinates, third order
$\{\mathcal{M}_{abcd}\}$	$3 \times 3 \times 3 \times 3$	Derivatives of traveltime, ray-centred coordinates, fourth order
		...
$\{R_{ij}^{m}\}$	$3 \times 3 \times 3$	Operator in the computation of inverse-transform coefficients, second order
$\{R_{ijk}^{mn}\}$	$3 \times 3 \times 3 \times 3 \times 3$	Operator in the computation of inverse-transform coefficients, third order
		...
$\boldsymbol{\Lambda} = \{\Lambda_{xr}\}$	6×6	Coefficients of phase-space coordinate transformation, ray-centred to Cartesian, first order
$\boldsymbol{\Lambda}^{11} = \{\Lambda_{ia}^{11}\}$	3×3	Sub-matrix of $\boldsymbol{\Lambda}$
$\boldsymbol{\Lambda}^{21} = \{\Lambda_{ia}^{21}\}$	3×3	Sub-matrix of $\boldsymbol{\Lambda}$
$\boldsymbol{\Lambda}^{22} = \{\Lambda_{ia}^{22}\}$	3×3	Sub-matrix of $\boldsymbol{\Lambda}$
$\mathbf{J}$	6×6	Matrix for rearranging the sub-matrices of matrix $\boldsymbol{\Lambda}$
$\{\Lambda_{xrs}\}$	$6 \times 6 \times 6$	Coefficients of phase-space coordinate transformation, ray-centred to Cartesian, second order
$\{\Lambda_{xrst}\}$	$6 \times 6 \times 6 \times 6$	Coefficients of phase-space coordinate transformation, ray-centred to Cartesian, third order

Table 1. Continued

Quantity	Dimension	Description
		...
$\mathbf{\Lambda}^{-1} = \{\Lambda^{\dagger}_{rx}\}$	6×6	Coefficients of phase-space coordinate transformation, Cartesian to ray-centred, first order
$\{\Lambda^{\dagger}_{rxy}\}$	$6 \times 6 \times 6$	Coefficients of phase-space coordinate transformation, Cartesian to ray-centred, second order
$\{\Lambda^{\dagger}_{rxyz}\}$	$6 \times 6 \times 6 \times 6$	Coefficients of phase-space coordinate transformation, Cartesian to ray-centred, third order
		...
$\bar{\boldsymbol{\mathcal{E}}}$	3×2	An orthonormal version of the contra-variant basis $\boldsymbol{\mathcal{E}}$
$\boldsymbol{\mathcal{C}} = \{\mathcal{C}_{AB}\}$	2×2	Matrix that rotates and scales $\bar{\boldsymbol{\mathcal{E}}}$ into $\boldsymbol{\mathcal{E}}$
$\boldsymbol{\mathcal{B}} = \{\mathcal{B}_{AB}\}$	2×2	Matrix describing the deviation of the basis $\boldsymbol{\mathcal{E}}$ from orthonormality
$\boldsymbol{\mathcal{A}} = \{\mathcal{A}_{AB}\}$	2×2	Coefficient matrix, phase-velocity formulation for the derivative $\mathrm{d}\boldsymbol{\mathcal{E}}/\mathrm{d}\tau$
$\mathcal{A}$		Coefficient scalar, phase-velocity formulation for the derivative $\mathrm{d}\boldsymbol{\mathcal{E}}/\mathrm{d}\tau$, applying when $\boldsymbol{\mathcal{E}}$ is orthonormal
$\boldsymbol{\mathcal{K}} = \{\mathcal{K}_{AB}\}$	2×2	Coefficient matrix, ray-velocity formulation for the derivative $\mathrm{d}\boldsymbol{\mathcal{E}}/\mathrm{d}\tau$

2.2 Hamilton–Jacobi equation and Hamilton's equations for the reference ray

We consider a *Hamilton–Jacobi equation*

$$\mathcal{H}(\mathbf{x}, \mathbf{p}) = 1/2, \tag{2}$$

where $\mathcal{H}(\mathbf{x}, \mathbf{p})$ is the *Hamiltonian* (see, e.g. Červený 2001; Iversen *et al.* 2019). As indicated by the form of eq. (2), we have chosen the Hamiltonian as a homogeneous function of the second degree in the momentum vector components. This choice is practical and does not affect the main theoretical results derived in the paper. The reason is that the Hamiltonian is used here only for establishing a reference ray and an associated ray-centred coordinate system. Neither of these are affected by the chosen degree of the Hamiltonian.

The Hamiltonian $\mathcal{H}(\mathbf{x}, \mathbf{p})$ has six degrees of freedom. On the other hand, the Hamilton–Jacobi equation (2) imposes a constraint in the phase space, so that the solutions for position and momentum have (together) five degrees of freedom. As the position is unrestricted, it means that a momentum satisfying the Hamilton–Jacobi-equation can only have two degrees of freedom. This constraint on the momentum will in general include position-dependent medium properties.

Let τ be a generic independent time variable controlling the continuation of ray-field quantities. Moreover, consider a reference ray, signified as Ω. The position and momentum vectors along Ω are given by the function:

$$\begin{aligned} \mathbf{x} &= \hat{\mathbf{x}}(\tau), \qquad x_i = \hat{x}_i(\tau), \\ \mathbf{p} &= \hat{\mathbf{p}}(\tau), \qquad p_i = \hat{p}_i(\tau). \end{aligned} \tag{3}$$

Taking the time derivative of these two vector functions yields the *ray-velocity vector* (or group-velocity vector)

$$\mathbf{v} = \dot{\mathbf{x}} = \frac{\mathrm{d}\hat{\mathbf{x}}}{\mathrm{d}\tau}, \qquad v_i = \dot{x}_i = \frac{\mathrm{d}\hat{x}_i}{\mathrm{d}\tau}, \tag{4}$$

and the so-called 'eta-vector' (Červený 2001)

$$\boldsymbol{\eta} = \dot{\mathbf{p}} = \frac{\mathrm{d}\hat{\mathbf{p}}}{\mathrm{d}\tau}, \qquad \eta_i = \dot{p}_i = \frac{\mathrm{d}\hat{p}_i}{\mathrm{d}\tau}. \tag{5}$$

The functions $\hat{\mathbf{x}}(\tau)$ and $\hat{\mathbf{p}}(\tau)$ constitute the solution to a system of ordinary differential equations

$$\frac{\mathrm{d}x_i}{\mathrm{d}\tau} = \frac{\partial \mathcal{H}}{\partial p_i}; \qquad \frac{\mathrm{d}p_i}{\mathrm{d}\tau} = -\frac{\partial \mathcal{H}}{\partial x_i}, \tag{6}$$

referred to as *Hamilton's equations*. The solution to eq. (6) must in general be computed numerically, for example using the Runge–Kutta method. Along the ray Ω the fundamental ray-theory relation

$$p_i v_i = 1 \tag{7}$$

is always satisfied.

2.3 Phase space in ray-centred coordinates

The *ray-centred coordinate system* consist of 3-D curvilinear coordinates (q_1, q_2, q_3) related to the reference ray Ω, on which $q_1 = q_2 = 0$. The position vector in ray-centred coordinates is denoted $\mathbf{q} = (q_i)$. The third coordinate is curvilinear and changes monotonically along Ω. We choose q_3 as the traveltime; hence the points on Ω satisfy

$$q_I = 0, \qquad q_3 = \tau. \tag{8}$$

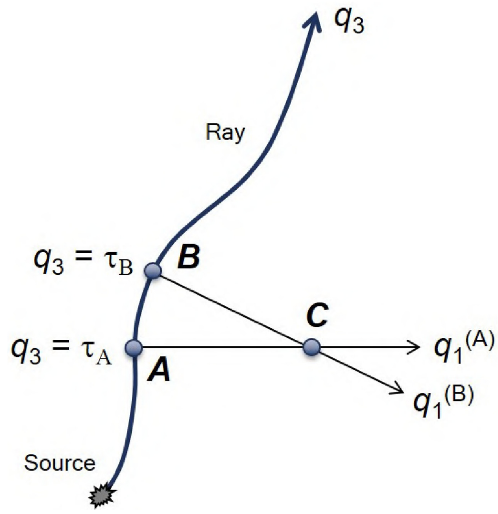

Figure 2. A 2-D illustration of a limitation of ray-centred coordinates. The q_1 coordinate lines corresponding to points *A* and *B* on the reference ray intersect in the point *C*. Therefore, in *C* the mapping between Cartesian and ray-centred coordinates is not one-to-one.

For a certain point on Ω the coordinate axes q_1 and q_2 are straight (but not necessarily perpendicular) lines. The q_1 and q_2 axes are situated in a moving plane, defined as a normal plane with respect to the slowness vector $\hat{\mathbf{p}}(\tau)$ on Ω.

The coordinate q_3 is *constant* in the specific q_1q_2 plane belonging a selected point on Ω. In this respect, the time variables q_3 and τ are fundamentally different. While τ signifies the time of a wavefront passing through an arbitrary point $\mathbf{x}$, this is *not* the case for q_3. Rather, q_3 is the *reference ray traveltime* resulting if we are able to construct a unique q_1q_2 plane through $\mathbf{x}$, such that the slowness vector $\hat{\mathbf{p}}(q_3)$ on Ω is normal to that plane. It follows that the ray-centred coordinates have a certain *region of validity* in the vicinity of Ω, arising from the fact that for a curved ray different q_1q_2 planes will intersect at some limiting transverse (paraxial) distance from Ω. Hence, for greater paraxial distances there will not be a one-to-one correspondence between ray-centred and Cartesian coordinates (see the illustration in Fig. 2). The width of the region of validity depends on the curvature of the ray trajectory.

In ray-centred coordinates we denote the momentum vector as $\mathbf{p}^{(q)} = (p_i^{(q)})$. The components of the ray-centred position and momentum vectors $\mathbf{q}$ and $\mathbf{p}^{(q)}$ form the phase-space coordinates

$$(w_r^{(q)}) = (q_i, p_j^{(q)}) = (q_1, q_2, q_3, p_1^{(q)}, p_2^{(q)}, p_3^{(q)}). \tag{9}$$

As with the phase space in Cartesian coordinates, see eq. (1), the entities in eq. (9) are independent. Moreover, in ray-centred coordinates eq. (2) has the counterpart $\mathcal{H}(\mathbf{q}, \mathbf{p}^{(q)}) = 1/2$. The latter yields five degrees of freedom.

2.4 Transformation from ray-centred to Cartesian coordinates and associated basis vectors

For any, general, point located close to the ray Ω, we assume a one-to-one correspondence between the coordinates (q_1, q_2, q_3) and (x_1, x_2, x_3). Violations of this assumption could occur—such cases are discussed in Section 6. Defining the one-to-one correspondence by the function $\mathbf{x}(q_1, q_2, q_3)$, the transformation from ray-centred and Cartesian coordinates is given as

$$\mathbf{x}(q_1, q_2, q_3 = \tau) = \hat{\mathbf{x}}(\tau) + \mathbf{e}_1(\tau)\, q_1 + \mathbf{e}_2(\tau)\, q_2, \tag{10}$$

where $\mathbf{e}_1(\tau)$ and $\mathbf{e}_2(\tau)$ are basis vectors corresponding, respectively, to the q_1 and q_2 coordinates. We assemble these vectors as columns in the 3×2 matrix

$$\boldsymbol{\mathcal{E}}(\tau) = \Big(\mathbf{e}_1(\tau)\ \mathbf{e}_2(\tau)\Big)\,. \tag{11}$$

The computation of matrix $\boldsymbol{\mathcal{E}}$ implies to solve additional differential equations along the ray Ω, see Appendix A. In the perspective of differential geometry, the ray-centred coordinates in eq. (10) result from a linearization of general coordinates associated with a geodesic—the *Fermi coordinates*.

In terms of the mapping (10) from ray-centred to Cartesian coordinates we can write the basis vectors $\mathbf{e}_1(\tau)$ and $\mathbf{e}_2(\tau)$ as

$$\mathbf{e}_1(\tau) = \frac{\partial \mathbf{x}}{\partial q_1}(0, 0, \tau)\,; \qquad \mathbf{e}_2(\tau) = \frac{\partial \mathbf{x}}{\partial q_2}(0, 0, \tau). \tag{12}$$

The basis vector belonging to the coordinate $q_3 = \tau$ is

$$\frac{\partial \mathbf{x}}{\partial q_3}(0, 0, \tau) = \frac{\mathrm{d}\hat{\mathbf{x}}}{\mathrm{d}\tau}(\tau) = \mathbf{v}(\tau), \tag{13}$$

that means, the ray-velocity vector introduced in eq. (4). The vector set ($\mathbf{e}_1(\tau)$, $\mathbf{e}_2(\tau)$, $\mathbf{v}(\tau)$) is referred to as the *contra-variant basis* (e.g. Klimeš 2006a; Červený & Moser 2007) and is represented as the 3×3 transformation matrix,

$$\mathbf{H}(\tau) = \{H_{ia}(\tau)\} = \left\{ \frac{\partial x_i}{\partial q_a}(0, 0, \tau) \right\} = \Big(\mathcal{E}(\tau)\ \mathbf{v}(\tau) \Big). \tag{14}$$

The terminology contra-variant and co-variant comes from differential geometry and corresponds, respectively, to vectors and covectors. A *vector* is strictly defined to describe a change of *position*. From this point of view, the quantity $\mathbf{v}$ in eq. (4) is therefore a vector. If the timescale in the direction of $\mathbf{v}$ is changed, say from unit seconds to milliseconds (scaling by a factor 1000), the corresponding scaling of $\mathbf{v}$ will be 1/1000 to compensate. The components of $\mathbf{v}$ therefore contra-varies with the change of basis, here exemplified by a change of the time variable. For a *co-vector* the situation is the opposite—it covaries with the change of basis. Co-vectors form the dual space to vectors in the sense of linear algebra.

Since vector $\hat{\mathbf{p}}(\tau)$ is normal to the $q_1 q_2$ plane, it follows that

$$\hat{p}_i(\tau) \frac{\partial x_i}{\partial q_A}(0, 0, \tau) = \hat{p}_i(\tau)\, \mathcal{E}_{iA}(\tau) = 0. \tag{15}$$

Moreover, eqs (7) and (13) yield

$$\hat{p}_i(\tau) \frac{\partial x_i}{\partial q_3}(0, 0, \tau) = \hat{p}_i(\tau)\, v_i(\tau) = 1. \tag{16}$$

The basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ can be chosen orthonormal, but more general options are available (see Appendix A). In the latter case it is of course important to ensure that $\mathbf{e}_1$ and $\mathbf{e}_2$ do not become co-linear, and that neither of them vanish. We remark that $\mathbf{e}_1$ and $\mathbf{e}_2$ do not depend on the curvature of wavefronts, so that the determinant of matrix $\mathbf{H}$ will not vanish as a result of caustics in the wavefield.

2.5 Momentum vector in ray-centred coordinates

Consider a general differentiable time function $\mathcal{T}$, expressed either in Cartesian coordinates or ray-centred coordinates, so that

$$\mathcal{T}(\mathbf{x}) = \mathcal{T}(\mathbf{q}). \tag{17}$$

We emphasize that $\mathcal{T}$ is *general*, which means it can be chosen arbitrarily, as long as eq. (17) is satisfied. Applying the chain rule for derivatives to $\mathcal{T}$, we obtain

$$\frac{\partial \mathcal{T}}{\partial x_i} = \frac{\partial \mathcal{T}}{\partial q_a}\frac{\partial q_a}{\partial x_i}, \qquad \frac{\partial \mathcal{T}}{\partial q_a} = \frac{\partial \mathcal{T}}{\partial x_i}\frac{\partial x_i}{\partial q_a}. \tag{18}$$

In the following, it is necessary to consider the momenta p_i and $p_a^{(q)}$ either as independent variables or as functions of position. To serve both these purposes, we take

$$p_i = \frac{\partial \mathcal{T}}{\partial x_i}, \qquad p_a^{(q)} = \frac{\partial \mathcal{T}}{\partial q_a}, \tag{19}$$

and we note that the momentum component

$$p_3^{(q)} = \frac{\partial \mathcal{T}}{\partial q_3} \tag{20}$$

is dimensionless.

For a given choice of function $\mathcal{T}$, the momentum components in eq. (19) are functions of position. However, since the time function $\mathcal{T}$ can be chosen arbitrarily, the time gradients may in general have any direction and any magnitude. As a consequence, the relations (18) can be restated

$$p_i = p_a^{(q)} \frac{\partial q_a}{\partial x_i}[\mathbf{x}(q_1, q_2, q_3)], \tag{21}$$

$$p_a^{(q)} = p_i \frac{\partial x_i}{\partial q_a}(q_1, q_2, q_3), \tag{22}$$

where the momenta on the right-hand sides may vary freely.

Now recall the essence of eq. (15), that the basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are both normal to the slowness vector, and eq. (16), that the dot product of the slowness vector and the ray-velocity vector is one. Using eqs (15) and (16) in eq. (22), it follows that the ray-centred momentum components belonging to Ω are constants,

$$p_A^{(q)} = 0, \qquad p_3^{(q)} = 1. \tag{23}$$

Time differentiation of eqs (8) and (23) yields, in ray-centred coordinates, the counterparts to the vector functions $\mathbf{v}$ and $\boldsymbol{\eta}$ in eqs (4) and (5). We obtain

$$\big(v_a^{(q)}(\tau)\big) = (0, 0, 1), \qquad \big(\eta_a^{(q)}(\tau)\big) = (0, 0, 0). \tag{24}$$

Moreover, the combination of eqs (21) and (23) yields

$$\hat{p}_i(\tau) = \frac{\partial q_3}{\partial x_i}[\mathbf{x}(0, 0, \tau)]. \tag{25}$$

We discuss eq. (25). For an arbitrary point $\mathbf{q} = (0, 0, \tau)$ on the reference ray Ω, the slowness vector is equal to the time gradient of an hypothetically 'exploding' plane wave. The wavefront source plane coincides with the q_1q_2-plane of the ray-centred coordinate system at the point under consideration. However, since the local time gradient of a plane wavefront equals the time gradient corresponding to any other wavefront propagating in the same direction, it is clear that eq. (25) can be generalized to $p_i = \partial\tau/\partial x_i$ along the ray Ω and along any other ray. This represents a fundamental property in ray theory— for a certain location on a ray, the slowness vector must equal the traveltime gradient at that location (see, e.g. Červený 2001). We note in particular that a traveltime function with small (paraxial) variation with respect to q_1 and q_2 will yield a value of the momentum component $p_3^{(q)} = \partial\tau/\partial q_3$ that differs only slightly from 1.

Consider a perturbed momentum vector $\mathbf{p}^{(q)} = \hat{\mathbf{p}}^{(q)}(\tau) + \delta\mathbf{p}^{(q)}$ along Ω, specified such that $p_1^{(q)}$ and $p_2^{(q)}$ are both set to zero, while $p_3^{(q)}$ is free to vary. In view of eq. (20) one can then interpret $p_3^{(q)}$ in terms of a virtual stretch of the time q_3 along Ω. To describe this kind of stretch effect, Burridge (Chapman 2004, p. 152) introduced the variable ϵ. It relates to the momentum component $p_3^{(q)}$ simply by

$$\epsilon = p_3^{(q)} - 1, \tag{26}$$

given that $p_1^{(q)} = p_2^{(q)} = 0$.

2.6 Transformation from Cartesian to ray-centred coordinates and associated basis vectors

The two sets of first-order derivatives involved in the transformations between ray-centred and Cartesian coordinates have to satisfy the relations

$$\frac{\partial q_a}{\partial x_i}\frac{\partial x_i}{\partial q_b} = \delta_{ab}, \qquad \frac{\partial x_i}{\partial q_a}\frac{\partial q_a}{\partial x_j} = \delta_{ij}. \tag{27}$$

In (27) the left-hand subequation has the implications

$$\frac{\partial q_3}{\partial x_i}\frac{\partial x_i}{\partial q_b} = \delta_{3b}, \tag{28}$$

$$\frac{\partial q_a}{\partial x_i}\frac{\partial x_i}{\partial q_3} = \delta_{a3}, \tag{29}$$

$$\frac{\partial q_A}{\partial x_i}\frac{\partial x_i}{\partial q_B} = \delta_{AB}, \tag{30}$$

while the right-hand subequation yields

$$\frac{\partial q_A}{\partial x_i}\frac{\partial x_j}{\partial q_A} = \alpha_{ij}. \tag{31}$$

Here, we have introduced the quantity

$$\alpha_{ij} = \delta_{ij} - \frac{\partial q_3}{\partial x_i}\frac{\partial x_j}{\partial q_3}. \tag{32}$$

Moreover, on the ray Ω we define the 3×2 matrix

$$\mathcal{F}(\tau) = \left\{\frac{\partial q_A}{\partial x_i}[\mathbf{x}(0, 0, \tau)]\right\}. \tag{33}$$

In view of eqs (25) and (33) the matrix $\mathbf{H}(\tau)$ in eq. (14) has the inverse

$$\mathbf{H}^{-1}(\tau) = \left\{H^{\dagger}(\tau)\right\} = \left\{\frac{\partial q_a}{\partial x_i}[\mathbf{x}(0, 0, \tau)]\right\} = \left(\mathcal{F}(\tau)\ \hat{\mathbf{p}}(\tau)\right)^T. \tag{34}$$

The forms of matrices $\mathbf{H}$ and $\mathbf{H}^{-1}$ then yield the relations

$$\mathbf{v}^T\mathcal{F} = \{0_{1A}\}; \quad v_i\mathcal{F}_{iA} = 0, \tag{35}$$

$$\mathcal{F}^T\mathcal{E} = \{\delta_{AB}\}; \quad \mathcal{F}_{iA}\mathcal{E}_{iB} = \delta_{AB}, \tag{36}$$

$$\mathcal{F}\mathcal{E}^T = \{\delta_{ij}\} - \mathbf{p}\mathbf{v}^T; \quad \mathcal{F}_{iA}\mathcal{E}_{jA} = \delta_{ij} - p_iv_j = \alpha_{ij}, \tag{37}$$

with all quantities evaluated on the reference ray Ω.

It is remarked that the quantity α_{ij} on Ω, on the right-hand side of eq. (37), represents a projection operator with respect to the wave-propagation metric (Hanyga 1982; Klimeš 2002, 2006a)

$$\alpha_{ij} = P_i^j = \delta_i^j - p_ih_3^j \qquad \text{(notation of Klimeš)}.$$

The columns of matrix $\mathcal{F}$ represent two vectors $\mathbf{f}_1$ and $\mathbf{f}_2$,

$$\mathcal{F} = \begin{pmatrix} \mathbf{f}_1 & \mathbf{f}_2 \end{pmatrix}, \tag{38}$$

and the vector set ($\mathbf{f}_1$, $\mathbf{f}_2$, $\mathbf{p}$) is the *co-variant basis* (e.g., Klimeš 2006a; Červený & Moser 2007). The vectors $\mathbf{f}_1$ and $\mathbf{f}_2$ are related to the contra-variant basis ($\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{v}$) by the operations (Klimeš 2006a)

$$\mathbf{f}_1 = \frac{\mathbf{e}_2 \times \mathbf{v}}{\mathbf{e}_1 \cdot (\mathbf{e}_2 \times \mathbf{v})}, \qquad \mathbf{f}_2 = \frac{\mathbf{v} \times \mathbf{e}_1}{\mathbf{e}_2 \cdot (\mathbf{v} \times \mathbf{e}_1)}. \tag{39}$$

The third co-variant and contra-variant basis vectors, $\mathbf{p}$ and $\mathbf{v}$, are both known, but we include for completeness their relations to the other set of basis vectors,

$$\mathbf{p} = \frac{\mathbf{e}_1 \times \mathbf{e}_2}{\mathbf{v} \cdot (\mathbf{e}_1 \times \mathbf{e}_2)}, \qquad \mathbf{v} = \frac{\mathbf{f}_1 \times \mathbf{f}_2}{\mathbf{p} \cdot (\mathbf{f}_1 \times \mathbf{f}_2)}. \tag{40}$$

On the reference ray, the slowness vector $\mathbf{p}$ is, by definition, orthogonal to the contra-variant basis vectors $\mathbf{e}_1$ and $\mathbf{e}_2$, see eq. (15) and the first sub-eq. (40). On the other hand, eq. (35) and the second sub-eq. (40) show that the ray-velocity vector $\mathbf{v}$ is orthogonal to the co-variant basis vectors $\mathbf{f}_1$ and $\mathbf{f}_2$.

2.7 Continuation of basis vectors along the reference ray

Various options exist for continuation of the matrices $\mathcal{E}$ and $\mathcal{F}$ along the reference ray Ω; details concerning such options can be found in Červený (2001), Klimeš (2006a) and Červený *et al.* (2007). In Appendix A, we describe simple ways of computing matrix $\mathcal{F}$ from matrix $\mathcal{E}$, and we also provide general expressions for the time derivatives of $\mathcal{E}$ and $\mathcal{F}$.

In the *standard option* for the bases of the ray-centred coordinate system the vectors $\mathbf{e}_A$, $A = 1, 2$, are orthonormal; in addition the two derivative vectors $\dot{\mathbf{e}}_A$ are both parallel to the slowness vector, $\mathbf{p}$. This results in an ordinary differential equation for the vector $\mathbf{e}_A$ along the reference ray,

$$\frac{\mathrm{d}\mathbf{e}_A}{\mathrm{d}\tau} = -c^2\, \mathbf{p}\, (\boldsymbol{\eta} \cdot \mathbf{e}_A), \qquad A = 1 \text{ or } 2, \tag{41}$$

where c is the (scalar) phase velocity. It is sufficient to include this differential equation only for, say, vector $\mathbf{e}_1$. Then vector $\mathbf{e}_2$ can be found from the cross product formula

$$\mathbf{e}_2 = \frac{\mathbf{p}}{\|\mathbf{p}\|} \times \mathbf{e}_1. \tag{42}$$

For a proof of eq. (41), see the derivation of eq. (A8) in Appendix A. The form (41) is attained when we set the 2×2 matrix $\boldsymbol{\mathcal{A}}$ zero in (A8).

At the initial point on Ω the direction of vector $\mathbf{e}_1$ can, in principle, be chosen arbitrarily in the normal plane to the slowness vector. It may, however, be practical to use a convention for the orientation of $\mathbf{e}_1$, for example, to let it comply with a mesh on an initial surface. The standard option is used in the Complete Ray Tracing (CRT) software package (Klimeš 2006b, section 6.2).

3 TRANSFORMATION OF PHASE-SPACE PERTURBATIONS: FROM RAY-CENTRED TO CARTESIAN COORDINATES

We consider the transformation of phase-space perturbations from ray-centred coordinates to Cartesian coordinates. Recall then that the phase space is represented by the two coordinate systems, respectively, in eqs (1) and (9). The transformation of the perturbations can be expressed as the Taylor series

$$\begin{aligned}\Delta w_x = {} & \frac{\partial w_x}{\partial w_r^{(q)}} \Delta w_r^{(q)} + \frac{1}{2} \frac{\partial^2 w_x}{\partial w_r^{(q)} \partial w_s^{(q)}} \Delta w_r^{(q)} \Delta w_s^{(q)} + \frac{1}{6} \frac{\partial^3 w_x}{\partial w_r^{(q)} \partial w_s^{(q)} \partial w_t^{(q)}} \Delta w_r^{(q)} \Delta w_s^{(q)} \Delta w_t^{(q)} \\ & + \frac{1}{24} \frac{\partial^4 w_x}{\partial w_r^{(q)} \partial w_s^{(q)} \partial w_t^{(q)} \partial w_u^{(q)}} \Delta w_r^{(q)} \Delta w_s^{(q)} \Delta w_t^{(q)} \Delta w_u^{(q)} + \dots,\end{aligned} \tag{43}$$

where we use Einstein's summation convention for the indices $r, s, t, u = 1, 2, \dots, 6$. The indices of the computed perturbation in the Cartesian phase-space coordinates are $x = 1, 2, \dots, 6$. The objective of this section is to establish specific expressions for the partial derivatives in eq. (43), which are all evaluated on the reference ray Ω and constitute the coefficients of the transformation.

In eq. (10) the position vector in Cartesian coordinates is expressed as a function of ray-centred coordinates. Using component notation, this equation is equivalently written

$$x_i(\mathbf{q}) = \hat{x}_i(q_3) + \mathcal{E}_{iA}(q_3) q_A. \tag{44}$$

It is noted that the position vector in the Cartesian coordinate system depends on the position vector in ray-centred coordinates in a one-to-one fashion. The position-vector function on the left-hand side of eq. (44) does not depend on the momentum vector in ray-centred coordinates. Moreover, using eq. (21) the momentum (slowness) vector in Cartesian coordinates is expressed as a function of ray-centred phase-space coordinates,

$$p_i(\mathbf{q}, \mathbf{p}^{(q)}) = \frac{\partial q_m}{\partial x_i}[\mathbf{x}(\mathbf{q})]\, p_m^{(q)}. \tag{45}$$

3.1 Partial derivatives of Cartesian position coordinates

Based on eq. (44) we obtain first- and second-order partial derivatives of Cartesian position coordinates x_i with respect to ray-centred position coordinates q_a,

$$\frac{\partial x_i}{\partial q_A} = \mathcal{E}_{iA}, \qquad \frac{\partial x_i}{\partial q_3} = \frac{\mathrm{d}\hat{x}_i}{\mathrm{d}\tau} + \frac{\mathrm{d}\mathcal{E}_{iA}}{\mathrm{d}\tau} q_A, \tag{46}$$

$$\frac{\partial^2 x_i}{\partial q_A \partial q_B} = 0, \qquad \frac{\partial^2 x_i}{\partial q_3^2} = \frac{\mathrm{d}^2\hat{x}_i}{\mathrm{d}\tau^2} + \frac{\mathrm{d}^2\mathcal{E}_{iA}}{\mathrm{d}\tau^2} q_A, \qquad \frac{\partial^2 x_i}{\partial q_3 \partial q_A} = \frac{\mathrm{d}\mathcal{E}_{iA}}{\mathrm{d}\tau}. \tag{47}$$

We observe that the extension to third- and fourth-order derivatives of x_i with respect to q_a is trivial. Moreover, to any order partial derivatives of x_i with respect to ray-centred momentum coordinates $p_a^{(q)}$ are zero.

Evaluation of the partial derivatives in eqs (46) and (47) on the reference ray, where $q_A = 0$, yields

$$\frac{\partial x_i}{\partial q_A} = \mathcal{E}_{iA}, \qquad \frac{\partial x_i}{\partial q_3} = \dot{x}_i = v_i, \tag{48}$$

$$\frac{\partial^2 x_i}{\partial q_A \partial q_B} = 0, \qquad \frac{\partial^2 x_i}{\partial q_3^2} = \ddot{x}_i = \dot{v}_i, \qquad \frac{\partial^2 x_i}{\partial q_3 \partial q_A} = \dot{\mathcal{E}}_{iA}. \tag{49}$$

Expressions for higher-order partial derivatives follow readily.

3.2 Partial derivatives of Cartesian momentum coordinates

We derive relations for the derivatives of the Cartesian momentum coordinates p_i with respect to the ray-centred phase-space coordinates q_a and $p_a^{(q)}$. The starting point is eq. (45), which yields

$$\frac{\partial p_i}{\partial q_a} = \frac{\partial x_j}{\partial q_a} \frac{\partial^2 q_b}{\partial x_i \partial x_j} p_b^{(q)}, \tag{50}$$

$$\frac{\partial p_i}{\partial p_a^{(q)}} = \frac{\partial q_a}{\partial x_i}. \tag{51}$$

Applying eq. (51) in eq. (45), we can conclude that for any corresponding location in the phase space, (w_r) or $(w_r^{(q)})$, we must have

$$p_i = \frac{\partial p_i}{\partial p_a^{(q)}} p_a^{(q)}. \tag{52}$$

Differentiating eq. (52) with respect to q_j yields the general relation,

$$\frac{\partial p_i}{\partial q_a} = \frac{\partial^2 p_i}{\partial q_a \partial p_b^{(q)}} p_b^{(q)}, \tag{53}$$

in which

$$\frac{\partial^2 p_i}{\partial q_a \partial p_b^{(q)}} = \frac{\partial}{\partial q_a}\left(\frac{\partial p_i}{\partial p_b^{(q)}}\right) = \frac{\partial}{\partial q_a}\left(\frac{\partial q_b}{\partial x_i}\right) = \frac{\partial x_j}{\partial q_a}\frac{\partial^2 q_b}{\partial x_i \partial x_j}. \tag{54}$$

In the last step we used the differential operator $\partial/\partial q_a = (\partial x_j/\partial q_a)\,\partial/\partial x_j + (\partial p_j/\partial q_a)\,\partial/\partial p_j$, where the second term has no effect.

3.2.1 *First-order partial derivatives of Cartesian momentum coordinates*

After some elaboration (Appendix B) we can restate eq. (50) as

$$\frac{\partial p_i}{\partial q_A} = \frac{\partial q_3}{\partial x_i}\frac{\partial p_j}{\partial q_3}\frac{\partial x_j}{\partial q_A}, \tag{55}$$

$$\frac{\partial p_i}{\partial q_3} = p_A^{(q)} \frac{\partial}{\partial q_3}\left(\frac{\partial q_A}{\partial x_i}\right) + p_3^{(q)} \frac{\partial}{q_3}\left(\frac{\partial q_3}{\partial x_i}\right). \tag{56}$$

Evaluation of eqs (51), (55) and (56) on the reference ray yields the results

$$\frac{\partial p_i}{\partial q_A} = p_i \eta_j \mathcal{E}_{jA}, \qquad \frac{\partial p_i}{\partial q_3} = \eta_i, \tag{57}$$

$$\frac{\partial p_i}{\partial p_A^{(q)}} = \mathcal{F}_{iA}, \qquad \frac{\partial p_i}{\partial p_3^{(q)}} = p_i. \tag{58}$$

3.2.2 Higher-order partial derivatives of Cartesian momentum coordinates

Derivations of general higher-order partial derivatives of Cartesian momentum coordinates p_i, evaluated on the reference ray Ω, are given in Appendix B. We summarize here the results.

The second-order derivatives of p_i on Ω are

$$\frac{\partial^2 p_i}{\partial q_A \partial q_B} = 2 p_i \eta_j \eta_k \mathcal{E}_{jA} \mathcal{E}_{kB}, \qquad \frac{\partial^2 p_i}{\partial q_3^2} = \dot{\eta}_i, \qquad \frac{\partial^2 p_i}{\partial q_3 \partial q_A} = \frac{\mathrm{d}}{\mathrm{d}\tau}\left(p_i \eta_j \mathcal{E}_{jA}\right), \tag{59}$$

$$\frac{\partial^2 p_i}{\partial p_a^{(q)} \partial p_b^{(q)}} = 0, \tag{60}$$

$$\frac{\partial^2 p_i}{\partial q_A \partial p_3^{(q)}} = p_i \eta_j \mathcal{E}_{jA}, \qquad \frac{\partial^2 p_i}{\partial q_3 \partial p_3^{(q)}} = \eta_i,$$

$$\frac{\partial^2 p_i}{\partial q_A \partial p_B^{(q)}} = -p_i \mathcal{K}_{AB}, \qquad \frac{\partial^2 p_i}{\partial q_3 \partial p_A^{(q)}} = \dot{\mathcal{F}}_{iA}, \tag{61}$$

where we have introduced a 2×2 matrix $\boldsymbol{\mathcal{K}} = \{\mathcal{K}_{AB}\}$ (see Klimeš 2006a, eq. 24), such that

$$\mathcal{K}_{AB} \equiv \dot{\mathcal{E}}_{iA} \mathcal{F}_{iB}. \tag{62}$$

For the third-order derivatives of p_i on Ω we obtain the expressions

$$\frac{\partial^3 p_i}{\partial q_A \partial q_B \partial q_C} = 6 p_i \eta_j \eta_k \eta_l \mathcal{E}_{jA} \mathcal{E}_{kB} \mathcal{E}_{lC}, \qquad \frac{\partial^3 p_i}{\partial q_3^3} = \ddot{\eta}_i,$$

$$\frac{\partial^3 p_i}{\partial q_3^2 \partial q_A} = \frac{\mathrm{d}^2}{\mathrm{d}\tau^2}\left(p_i \eta_j \mathcal{E}_{jA}\right), \qquad \frac{\partial^3 p_i}{\partial q_3 \partial q_A \partial q_B} = 2 \frac{\mathrm{d}}{\mathrm{d}\tau}\left(p_i \eta_j \eta_k \mathcal{E}_{jA} \mathcal{E}_{kB}\right), \tag{63}$$

$$\frac{\partial^3 p_i}{\partial q_A \partial q_B \partial p_C^{(q)}} = -p_i \eta_j \left(\mathcal{E}_{jA} \mathcal{K}_{BC} + \mathcal{E}_{jB} \mathcal{K}_{AC}\right),$$

$$\frac{\partial^3 p_i}{\partial q_3^2 \partial p_A^{(q)}} = \ddot{\mathcal{F}}_{iA}, \qquad \frac{\partial^3 p_i}{\partial q_3 \partial q_A \partial p_B^{(q)}} = -\frac{\mathrm{d}}{\mathrm{d}\tau}\left(p_i\ \mathcal{K}_{AB}\right),$$

$$\frac{\partial^3 p_i}{\partial q_A \partial q_B \partial p_3^{(q)}} = 2 p_i \eta_j \eta_k \mathcal{E}_{jA} \mathcal{E}_{kB}, \qquad \frac{\partial^3 p_i}{\partial q_3^2 \partial p_3^{(q)}} = \dot{\eta}_i, \qquad \frac{\partial^3 p_i}{\partial q_3 \partial q_A \partial p_3^{(q)}} = \frac{\mathrm{d}}{\mathrm{d}\tau}\left(p_i \eta_j \mathcal{E}_{jA}\right). \tag{64}$$

Some third-order derivatives are always zero, namely,

$$\frac{\partial^3 p_i}{\partial p_a^{(q)} \partial p_b^{(q)} \partial p_c^{(q)}} = 0, \qquad \frac{\partial^3 p_i}{\partial p_a^{(q)} \partial p_b^{(q)} \partial q_c} = 0, \tag{65}$$

Concerning fourth-order derivatives of p_i on Ω we only need to consider the following results,

$$\frac{\partial^4 p_i}{\partial q_A \partial q_B \partial q_C \partial q_D} = 24\ p_i\ \eta_j \eta_k \eta_l \eta_m\ \mathcal{E}_{jA} \mathcal{E}_{kB} \mathcal{E}_{lC} \mathcal{E}_{mD}, \tag{66}$$

$$\frac{\partial^4 p_i}{\partial q_A \partial q_B \partial q_C \partial p_D^{(q)}} = -2\ p_i\ \eta_j \eta_k \left(\mathcal{E}_{jA} \mathcal{E}_{kB} \mathcal{K}_{CD} + \mathcal{E}_{jA} \mathcal{E}_{kC} \mathcal{K}_{BD} + \mathcal{E}_{jB} \mathcal{E}_{kC} \mathcal{K}_{AD}\right). \tag{67}$$

The remaining fourth-order derivatives are either zero or can be obtained trivially from lower-order derivatives.

4 INVERSE TRANSFORMATION OF PHASE-SPACE PERTURBATIONS: FROM CARTESIAN TO RAY-CENTRED COORDINATES

We consider in this section the inverse transformation of phase-space perturbations, that means, in the direction from Cartesian coordinates to ray-centred coordinates. This transformation can be expressed as the Taylor series

$$\Delta w_r^{(q)} = \frac{\partial w_r^{(q)}}{\partial w_x} \Delta w_x + \frac{1}{2} \frac{\partial^2 w_r^{(q)}}{\partial w_x \partial w_y} \Delta w_x \Delta w_y + \frac{1}{6} \frac{\partial^3 w_r^{(q)}}{\partial w_x \partial w_y \partial w_z} \Delta w_x \Delta w_y \Delta w_z$$
$$+ \frac{1}{24} \frac{\partial^4 w_r^{(q)}}{\partial w_x \partial w_y \partial w_z \partial w_\phi} \Delta w_x \Delta w_y \Delta w_z \Delta w_\phi + \dots, \tag{68}$$

with $r, x, y, z, \phi = 1, 2, \ldots, 6$, and where the involved coefficients (partial derivatives) belongs to the reference ray Ω. We derive specific expressions for the derivatives of the ray-centred position and momentum coordinates up to, respectively, the second and the fourth order in the Cartesian phase-space coordinates.

As opposed to the situation in eq. (44) for the forward transformation of phase-space perturbations, from ray-centred to Cartesian coordinates, we do not have an explicit expression for the ray-centred space coordinates in terms of Cartesian coordinates. However, after multiplying each side of (44) by $\mathcal{F}_{iB}(q_3)$ and using (36), we obtain the following relationship between the position coordinates,

$$q_A = \mathcal{F}_{kA}(q_3)\,[x_k - \hat{x}_k(q_3)]\,. \tag{69}$$

The ray-centred momentum vector can be expressed as

$$p_a^{(q)}(\mathbf{x},\mathbf{p}) = p_i \frac{\partial p_a^{(q)}}{\partial p_i}(\mathbf{x},\mathbf{p}) = p_i \frac{\partial x_i}{\partial q_a}[\mathbf{q}(\mathbf{x})]. \tag{70}$$

4.1 Symplecticity

A 6×6 matrix $\mathbf{\Lambda}$ is called *symplectic* (e.g. Červený *et al.* 2007, eq. 58) if it satisfies the equation

$$\mathbf{\Lambda}^T \mathbf{J} \mathbf{\Lambda} = \mathbf{J}, \tag{71}$$

where $\mathbf{J}$ is the 6×6 matrix

$$\mathbf{J} = \{J_{rs}\} = \begin{pmatrix} \{0_{ij}\} & \{\delta_{ij}\} \\ -\{\delta_{ij}\} & \{0_{ij}\} \end{pmatrix}. \tag{72}$$

From eqs (71) and (72) it follows that the inverse matrix $\mathbf{\Lambda}^{-1}$ is given by

$$\mathbf{\Lambda}^{-1} = -\mathbf{J}\,\mathbf{\Lambda}^T\,\mathbf{J}. \tag{73}$$

Consider now the 6×6 matrix $\mathbf{\Lambda}$ defined as

$$\mathbf{\Lambda} \;=\; \left\{ \frac{\partial w_x}{\partial w_r^{(q)}} \right\} \;=\; \begin{pmatrix} \left\{ \frac{\partial x_i}{\partial q_a} \right\} & \{0_{ij}\} \\ \left\{ \frac{\partial p_i}{\partial q_a} \right\} & \left\{ \frac{\partial p_i}{\partial p_a^{(q)}} \right\} \end{pmatrix} \;=\; \begin{pmatrix} \mathbf{\Lambda}^{11} & \mathbf{0} \\ \mathbf{\Lambda}^{21} & \mathbf{\Lambda}^{22} \end{pmatrix}. \tag{74}$$

Inverting both sides of eq. (51), we obtain

$$\frac{\partial x_i}{\partial q_a} = \frac{\partial p_a^{(q)}}{\partial p_i}. \tag{75}$$

Another important property related to the matrices $\mathbf{\Lambda}$ and $\mathbf{\Lambda}^{-1}$ is

$$\frac{\partial p_i}{\partial q_a} = -\frac{\partial p_a^{(q)}}{\partial x_i}. \tag{76}$$

For a proof of eq. (76), see Appendix B.

The matrix $\mathbf{\Lambda}$ in eq. (74) holds for any phase-space location $\mathbf{w}^{(q)}$ complying with the validity of ray-centred coordinates, not only for phase-space locations corresponding to the reference ray. It is straightforward to show that $\mathbf{\Lambda}$ satisfies eq. (71). As a consequence, $\mathbf{\Lambda}$ is symplectic at any phase-space location $\mathbf{w}^{(q)}$ where the mapping between ray-centred and Cartesian coordinates is one-to-one. Wherever $\mathbf{\Lambda}$ is non-singular its determinant equals one.

Using eqs (51), (73), (75) and (76) we formulate the inverse $\mathbf{\Lambda}^{-1}$ as

$$\mathbf{\Lambda}^{-1} = \left\{ \frac{\partial w_r^{(q)}}{\partial w_x} \right\} = \begin{pmatrix} \left\{ \frac{\partial p_i}{\partial p_a^{(q)}} \right\}^T & \{0_{ij}\} \\ -\left\{ \frac{\partial p_i}{\partial q_a} \right\}^T & \left\{ \frac{\partial x_i}{\partial q_a} \right\}^T \end{pmatrix} = \begin{pmatrix} \mathbf{\Lambda}^{22\,T} & \mathbf{0} \\ -\mathbf{\Lambda}^{21\,T} & \mathbf{\Lambda}^{11\,T} \end{pmatrix}. \tag{77}$$

4.2 Transformation matrices on the reference ray

Eqs (74) and (77) yield general forms of the 6×6 matrices describing to first order the transformation between ray-centred and Cartesian phase-space coordinates. We now summarize, in matrix and component form, the equations for the matrices $\mathbf{\Lambda}$ and $\mathbf{\Lambda}^{-1} = \mathbf{\Lambda}^{\dagger}$ corresponding to a point on the reference ray Ω. These matrices are well established (see e.g. Klimeš 1994; Červený 2001),

$$\mathbf{\Lambda} = \begin{pmatrix} \boldsymbol{\mathcal{E}} & \mathbf{v} & \{0_{iA}\} & \{0_{i1}\} \\ \mathbf{p}\boldsymbol{\eta}^T\boldsymbol{\mathcal{E}} & \boldsymbol{\eta} & \boldsymbol{\mathcal{F}} & \mathbf{p} \end{pmatrix}, \qquad \mathbf{\Lambda}^{-1} = \begin{pmatrix} \boldsymbol{\mathcal{F}}^T & \{0_{iA}\}^T \\ \mathbf{p}^T & \{0_{i1}\}^T \\ -\boldsymbol{\mathcal{E}}^T\boldsymbol{\eta}\mathbf{p}^T & \boldsymbol{\mathcal{E}}^T \\ -\boldsymbol{\eta}^T & \mathbf{v}^T \end{pmatrix}, \tag{78}$$

$$\{\Lambda_{xr}\} = \begin{pmatrix} \{\mathcal{E}_{iA}\} & \{v_i\} & \{0_{iA}\} & \{0_{i1}\} \\ \{p_i\eta_j\mathcal{E}_{jA}\} & \{\eta_i\} & \{\mathcal{F}_{iA}\} & \{p_i\} \end{pmatrix}, \quad \{\Lambda^\dagger_{rx}\} = \begin{pmatrix} \{\mathcal{F}_{iA}\}^T & \{0_{iA}\}^T \\ \{p_i\}^T & \{0_{i1}\}^T \\ -\{p_i\eta_j\mathcal{E}_{jA}\}^T & \{\mathcal{E}_{iA}\}^T \\ -\{\eta_i\}^T & \{v_i\}^T \end{pmatrix}. \tag{79}$$

To obtain $\mathbf{\Lambda}$ we use eqs (48), (57) and (58); to obtain $\mathbf{\Lambda}^{-1}$ we use eq. (77).

4.3 Coefficients of the inverse transformation determined by means of symplecticity

Eq. (71) can be utilized as a simple, indirect, means of obtaining the coefficients of the inverse transformation (Cartesian to ray-centred), given that the coefficients of the forward transformation (ray-centred to Cartesian) are known.

We restate eq. (73) in component form,

$$\Lambda^\dagger_{ry} = -J_{rs}\, J_{xy}\, \Lambda_{xs}, \tag{80}$$

where all indices take the values from 1 to 6. Successive differentiation to higher orders in the Cartesian phase-space coordinates then yields

$$\Lambda^\dagger_{ryz} = -J_{rs}\, J_{xy}\, \Lambda^\dagger_{tz}\, \Lambda_{xst}, \tag{81}$$

$$\Lambda^\dagger_{ryz\phi} = -J_{rs}\, J_{xy}\, (\Lambda^\dagger_{tz\phi}\, \Lambda_{xst} + \Lambda^\dagger_{tz}\, \Lambda^\dagger_{u\phi}\, \Lambda_{xstu}), \tag{82}$$

$$\Lambda^\dagger_{ryz\phi\mathring{a}} = -J_{rs}\, J_{xy}\, \Big[\Lambda^\dagger_{tz\phi\mathring{a}}\, \Lambda_{xst} + (\Lambda^\dagger_{tz}\, \Lambda^\dagger_{u\phi\mathring{a}} + \Lambda^\dagger_{t\phi}\, \Lambda^\dagger_{uz\mathring{a}} + \Lambda^\dagger_{t\mathring{a}}\, \Lambda^\dagger_{uz\phi})\, \Lambda_{xstu} + \Lambda^\dagger_{tz}\, \Lambda^\dagger_{u\phi}\, \Lambda^\dagger_{v\mathring{a}}\, \Lambda_{xstuv}\Big], \tag{83}$$

and so forth. Here, the index series r, s, t, u, v and $x, y, z, \phi, \mathring{a}$ represent ray-centred and Cartesian phase-space coordinates, respectively. For each order of differentiation, we use as input the results obtained in the differentiations for lower orders.

Eqs (80)–(83) have the advantage that they can be easily implemented. There are, however, contexts where it is useful to know explicit expressions for the transformation coefficients in terms of quantities of the wavefield, for example the slowness vector and the ray-velocity vector. We focus on such explicit expressions in the remaining part of this section.

4.4 Partial derivatives of ray-centred position coordinates

We obtain partial derivatives of the ray-centred position coordinates, q_a. To any order, partial derivatives of q_a with respect to momentum coordinates p_i are zero. In the following, we therefore need to consider only the partial derivatives with respect to position coordinates x_i. We provide here explicit expressions for such derivatives up to order four. For a certain spatial order of the transformation coefficients, the computation take the form of a linear operator acting on the time derivative of the coefficients of one order less.

4.4.1 First-order partial derivatives

We differentiate eq. (69) with respect to the Cartesian position coordinates. This yields

$$\frac{\partial q_A}{\partial x_i} = \dot{\mathcal{F}}_{kA}\frac{\partial q_3}{\partial x_i}\,[x_k - \hat{x}_k(q_3)] + \mathcal{F}_{kA}\left(\delta_{ki} - \dot{\hat{x}}_k\frac{\partial q_3}{\partial x_i}\right). \tag{84}$$

Here, the derivative $\partial q_3/\partial x_i$ can be expressed implicitly by two operations: First, form the 3×3 inverse $\{\partial q_i/\partial x_j\}$ based on eq. (50); secondly, extract the three components $\partial q_3/\partial x_i$ using

$$\frac{\partial q_3}{\partial x_i} = \delta_{3j}\,\frac{\partial q_i}{\partial x_j}. \tag{85}$$

On the reference ray Ω, for which $q_A = 0$, we naturally have

$$\frac{\partial q_A}{\partial x_i} = \mathcal{F}_{iA}, \qquad \frac{\partial q_3}{\partial x_i} = p_i. \tag{86}$$

4.4.2 Second-order partial derivatives

A derivation for the general second-order partial derivatives of ray-centred coordinates q_a is contained in Appendix C. On the ray Ω such derivatives can be expressed in the compact form

$$\frac{\partial^2 q_a}{\partial x_i\,\partial x_j} \;\equiv\; H^\dagger_{aij} \;=\; R^m_{ij}\,\dot{H}^\dagger_{am}, \tag{87}$$

where $H^{\dagger}_{aij} = \partial^2 q_a/\partial x_i \partial x_j$, and R^{m}_{ij} is the operator

$$R^{m}_{ij} = p_i \alpha_{jm} + p_j \alpha_{im} + p_i p_j v_m. \tag{88}$$

The upper and lower indices on the R^{m}_{ij} symbol correspond, respectively, to the input and output data of the operator. From eqs (87) and (88) we can state specific expressions for the derivatives of q_A and q_3,

$$\frac{\partial^2 q_A}{\partial x_i \partial x_j} = p_i \dot{\mathcal{F}}_{jA} + p_j \dot{\mathcal{F}}_{iA} - p_i p_j v_m \dot{\mathcal{F}}_{mA}, \tag{89}$$

$$\frac{\partial^2 q_3}{\partial x_i \partial x_j} = p_i \eta_j + \eta_i p_j - p_i p_j v_m \eta_m. \tag{90}$$

We note that eq. (90) is consistent with Červený & Klimeš (2010, eq. 35) .

4.4.3 Higher-order partial derivatives

Derivations for the third- and fourth-order derivatives of ray-centred coordinates q_a with respect to Cartesian coordinates x_i are given in Appendices D and E. The third-order derivatives are expressed in the form

$$\frac{\partial^3 q_a}{\partial x_i \partial x_j \partial x_k} \;\equiv\; H^{\dagger}_{aijk} \;=\; R^{mn}_{ijk}\, \dot{H}^{\dagger}_{amn}, \tag{91}$$

with $H^{\dagger}_{aijk} = \partial^3 q_a/\partial x_i \partial x_j \partial x_k$ and the operator R^{mn}_{ijk} given by

$$R^{mn}_{ijk} = p_i \alpha_{jm} \alpha_{kn} + p_j \alpha_{km} \alpha_{in} + p_k \alpha_{im} \alpha_{jn} + (p_i p_j \alpha_{kn} + p_j p_k \alpha_{in} + p_k p_i \alpha_{jn}) v_m + p_i p_j p_k v_m v_n. \tag{92}$$

For the fourth-order derivatives we get the result

$$\frac{\partial^4 q_a}{\partial x_i \partial x_j \partial x_k \partial x_l} \;\equiv\; H^{\dagger}_{aijkl} \;=\; R^{mnq}_{ijkl}\, \dot{H}^{\dagger}_{amnq}, \tag{93}$$

where $H^{\dagger}_{aijkl} = \partial^4 q_a/\partial x_i \partial x_j \partial x_k \partial x_l$, and R^{mnq}_{ijkl} is the operator

$$\begin{aligned} R^{mnq}_{ijkl} \;=\; & p_i \alpha_{jm} \alpha_{kn} \alpha_{lq} + p_j \alpha_{im} \alpha_{kn} \alpha_{lq} + p_k \alpha_{im} \alpha_{jn} \alpha_{lq} + p_l \alpha_{im} \alpha_{jn} \alpha_{kq} \\ & + (p_i p_j \alpha_{kn} \alpha_{lq} + p_i p_k \alpha_{jn} \alpha_{lq} + p_i p_l \alpha_{jn} \alpha_{kq} + p_j p_k \alpha_{in} \alpha_{lq} + p_j p_l \alpha_{in} \alpha_{kq} + p_k p_l \alpha_{in} \alpha_{jq})\, v_m \\ & + (p_i p_j p_k \alpha_{lq} + p_i p_j p_l \alpha_{kq} + p_i p_k p_l \alpha_{jq} + p_j p_k p_l \alpha_{iq})\, v_m v_n + p_i p_j p_k p_l\, v_m v_n v_q. \end{aligned} \tag{94}$$

4.5 Partial derivatives of ray-centred momentum coordinates

We obtain explicit expressions for the partial derivatives of ray-centred momentum coordinates with respect to Cartesian phase-space coordinates. We limit our specifications to include first- and second-order derivatives, to avoid unnecessary complexity. Higher-order derivatives are also of importance, but we can compute them as described above, by utilizing the symplectic property of the transformation.

4.5.1 First-order partial derivatives

The general first-order derivatives $\partial p^{(q)}_a/\partial x_i$ and $\partial p^{(q)}_a/\partial p_i$ are easily obtained using the connections of these derivatives to the derivatives $\partial p_i/\partial q_a$ and $\partial x_i/\partial q_a$, see eqs (75) and (76). On the reference ray we then have

$$\frac{\partial p^{(q)}_A}{\partial x_i} = -p_i \eta_j \mathcal{E}_{jA}, \qquad \frac{\partial p^{(q)}_3}{\partial x_i} = -\eta_i, \tag{95}$$

$$\frac{\partial p^{(q)}_A}{\partial p_j} = \mathcal{E}_{jA}, \qquad \frac{\partial p^{(q)}_3}{\partial p_j} = v_j. \tag{96}$$

4.5.2 Second-order partial derivatives

In consideration of eq. (75) we note that the following second-order partial derivatives with respect to Cartesian momentum coordinates will be zero,

$$\frac{\partial^2 p^{(q)}_a}{\partial p_i \partial p_j} = 0. \tag{97}$$

The same will be true for any higher-order derivative with respect to one of the three momentum coordinates.

Using eqs (75) and (76) we find general second-order mixed derivatives with respect to x_i and p_j, as follows. First,

$$\frac{\partial^2 p_A^{(q)}}{\partial x_i \partial p_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial x_j}{\partial q_A}\right) = \frac{\partial q_b}{\partial x_i}\frac{\partial}{\partial q_b}\left(\frac{\partial x_j}{\partial q_A}\right),$$

hence

$$\frac{\partial^2 p_A^{(q)}}{\partial x_i \partial p_j} = \frac{\partial q_3}{\partial x_i}\frac{\partial}{\partial q_3}\left(\frac{\partial x_j}{\partial q_A}\right). \tag{98}$$

Secondly,

$$\frac{\partial^2 p_3^{(q)}}{\partial x_i \partial p_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial x_j}{\partial q_3}\right) = \frac{\partial q_b}{\partial x_i}\frac{\partial}{\partial q_b}\left(\frac{\partial x_j}{\partial q_3}\right),$$

which yields

$$\frac{\partial^2 p_3^{(q)}}{\partial x_i \partial p_j} = \frac{\partial q_b}{\partial x_i}\frac{\partial}{\partial q_3}\left(\frac{\partial x_j}{\partial q_b}\right). \tag{99}$$

Evaluation of eqs (98) and (99) on the reference ray gives the results

$$\frac{\partial^2 p_A^{(q)}}{\partial x_i \partial p_j} = p_i \dot{\mathcal{E}}_{jA}, \qquad \frac{\partial^2 p_3^{(q)}}{\partial x_i \partial p_j} = H^\dagger_{ai}\dot{H}_{ja}, \tag{100}$$

We use eqs (75) and (76) to obtain the remaining second-order derivatives of ray-centred momentum coordinates, evaluated on the ray Ω. This yields

$$\begin{aligned}
\frac{\partial^2 p_A^{(q)}}{\partial x_i \partial x_j} &= \frac{\partial^2 q_3}{\partial x_i \partial x_j} p_k \dot{\mathcal{E}}_{kA} + p_i p_j p_k \ddot{\mathcal{E}}_{kA},\\
\frac{\partial^2 p_3^{(q)}}{\partial x_i \partial x_j} &= \left(\frac{\partial^2 q_a}{\partial x_i \partial x_j}\dot{H}_{ma} + H^\dagger_{ai}H^\dagger_{bj}\frac{\mathrm{d}}{\mathrm{d}\tau}\frac{\partial^2 x_m}{\partial q_a \partial q_b}\right)p_m.
\end{aligned} \tag{101}$$

5 RELATING THE DERIVATIVES OF A TRAVELTIME FUNCTION IN RAY-CENTRED AND CARTESIAN COORDINATES

In this section we use transformation coefficients derived above to relate the derivatives of a traveltime function given in ray-centred and Cartesian coordinates. It is assumed that the traveltime function in the vicinity of a point on the reference ray Ω can be expressed equivalently in ray-centred and Cartesian coordinates, as $\tau(\mathbf{q})$ and $\tau(\mathbf{x})$, respectively. On Ω, the second- and higher-order derivatives of τ with respect to x_i are denoted $M_{ij} = \partial^2\tau/\partial x_i \partial x_j$, $M_{ijk} = \partial^3\tau/\partial x_i \partial x_j \partial x_k$, etc.—the corresponding derivatives with respect to q_a are denoted $\mathcal{M}_{ab} = \partial^2\tau/\partial q_a \partial q_b$, $\mathcal{M}_{abc} = \partial^3\tau/\partial q_a \partial q_b \partial q_c$, etc. For the coefficients of the forward/inverse transformations we use, respectively, the notations $H_{ia} = \partial x_i/\partial q_a$, $H_{iab} = \partial^2 x_i/\partial q_a \partial q_b$, $H_{iabc} = \partial^3 x_i/\partial q_a \partial q_b \partial q_c$, etc., and $H^\dagger_{ai} = \partial q_a/\partial x_i$, $H^\dagger_{aij} = \partial^2 q_a/\partial x_i \partial x_j$, $H^\dagger_{aijk} = \partial^3 q_a/\partial x_i \partial x_j \partial x_k$, etc.

5.1 Mapping the derivatives of traveltime from ray-centred to Cartesian coordinates

One advantage of using ray-centred coordinates is that many derivatives of traveltime are zero in these coordinates. As a consequence, it can be useful to write the mapping of derivatives of traveltime, from ray-centred to Cartesian coordinates, explicitly in terms of the components $\mathcal{F}_{iA}$ and p_i of the co-variant basis vectors. As these expressions are space demanding as we proceed to higher orders, they are given in Appendix F. Here, we provide mapping equations given in terms of the inverse transformation components $H^\dagger_{ai}$. These are attractive because of compactness, but their computation may not be equally efficient as those in Appendix F.

On the reference ray Ω, the mapping for the first four orders of the derivatives of traveltime, from ray-centred to Cartesian coordinates, can be written,

$$p_i = p_a^{(q)} H^\dagger_{ai}, \tag{102}$$

$$M_{ij} = p_a^{(q)} H^\dagger_{aij} + \mathcal{M}_{ab} H^\dagger_{ai} H^\dagger_{aj}, \tag{103}$$

$$M_{ijk} = p_a^{(q)} H^\dagger_{aijk} + \mathcal{M}_{ab}(H^\dagger_{ai}H^\dagger_{bjk} + H^\dagger_{aj}H^\dagger_{bik} + H^\dagger_{ak}H^\dagger_{bij}) + \mathcal{M}_{abc}H^\dagger_{ai}H^\dagger_{bj}H^\dagger_{ck}, \tag{104}$$

$$\begin{aligned}
M_{ijkl} = p_a^{(q)} H^\dagger_{aijkl} &+ \mathcal{M}_{ab}\,(H^\dagger_{ai}H^\dagger_{bjkl} + H^\dagger_{aj}H^\dagger_{bikl} + H^\dagger_{ak}H^\dagger_{bijl} + H^\dagger_{al}H^\dagger_{bijk} + H^\dagger_{aij}H^\dagger_{bkl} + H^\dagger_{aik}H^\dagger_{bjl} + H^\dagger_{ail}H^\dagger_{bjk})\\
&+ \mathcal{M}_{abc}\,(H^\dagger_{ai}H^\dagger_{bj}H^\dagger_{ckl} + H^\dagger_{ai}H^\dagger_{bk}H^\dagger_{cjl} + H^\dagger_{ai}H^\dagger_{bl}H^\dagger_{cjk} + H^\dagger_{aj}H^\dagger_{bk}H^\dagger_{cil} + H^\dagger_{aj}H^\dagger_{bl}H^\dagger_{cik} + H^\dagger_{ak}H^\dagger_{bl}H^\dagger_{cij})\\
&+ \mathcal{M}_{abcd}\,H^\dagger_{ai}H^\dagger_{bj}H^\dagger_{ck}H^\dagger_{dl}.
\end{aligned} \tag{105}$$

This setup can be easily extended to order five and higher, if needed. Note that the first term on the right-hand side of eq. (103) corresponds to eq. (90) and also to Červený & Klimeš (2010, eq. 35).

5.2 Mapping the derivatives of traveltime from Cartesian to ray-centred coordinates

Proceeding as above, the complete mapping of the first four derivatives of traveltime from Cartesian to ray-centred coordinates can be stated, after evaluation on Ω,

$$p_a^{(q)} = p_i H_{ia}, \tag{106}$$

$$\mathcal{M}_{ab} = p_i H_{iab} + M_{ij} H_{ia} H_{jb}, \tag{107}$$

$$\mathcal{M}_{abc} = p_i H_{iabc} + M_{ij}(H_{ia} H_{jbc} + H_{ib} H_{jac} + H_{ic} H_{jab}) + M_{ijk} H_{ia} H_{jb} H_{kc}, \tag{108}$$

$$\begin{aligned}\mathcal{M}_{abcd} = p_i H_{iabcd} &+ M_{ij}\,(H_{ia} H_{jbcd} + H_{ib} H_{jacd} + H_{ic} H_{jabd} + H_{id} H_{jabc} + H_{iab} H_{jcd} + H_{iac} H_{jbd} + H_{iad} H_{jbc}) \\ &+ M_{ijk}\,(H_{ia} H_{jb} H_{kcd} + H_{ia} H_{jc} H_{kbd} + H_{ia} H_{jd} H_{kbc} + H_{ib} H_{jc} H_{kad} + H_{ib} H_{jd} H_{kac} + H_{ic} H_{jd} H_{kab}) \\ &+ M_{ijkl}\, H_{ia} H_{jb} H_{kc} H_{ld}.\end{aligned} \tag{109}$$

5.3 Special case: Traveltime function corresponding to an initial plane wave

We comment on a special case, for which the mapping of the derivatives of traveltime becomes particularly simple, namely, the traveltime function arising as a result of an initial plane wave.

Iversen *et al.* (2019) describes in detail how one can define the initial condition for the dynamic ray tracing quantities in Cartesian coordinates, referred to as *derivatives of phase-space perturbations*, and also how to derive from these quantities the derivatives of traveltime up to order four: p_i, M_{ij}, M_{ijk} and M_{ijkl}. Essential in this setup is the use of *constraint relations* of the standard type (Červený 2001) and of higher order (Iversen *et al.* 2019). In particular, when using the plane-wave initial condition, one can from these constraint relations alone obtain the complete set of dynamic ray tracing quantities.

For the initial plane wave, the computation of second-order derivatives of traveltime, M_{ij}, at the given (initial) point requires only knowledge of the first-order derivatives of the model parameters. Computation of third- and fourth order derivatives, M_{ijk} and M_{ijkl}, requires computation of second- and third-order derivatives of the model parameters, respectively. Hence, the initial plane wave situation is special, as it also permits a computation of fifth-order derivatives of traveltime, M_{ijklm}, given that we provide the corresponding fourth-order derivatives of the model parameters. These derivatives are naturally included when using the quintic B-spline as a basic element of the model representation.

To establish the fifth-order derivatives M_{ijklm}, we therefore have available the required model input data, but there are two minor extra issues to be handled, relative to the description in Iversen *et al.* (2019): (1) the set of constraint relations must be extended with one additional order; (2) the set of equations for the derivatives of traveltime must be extended with one additional equation, thus involving M_{ijklm}. We do not write the specific equations here, as they would need a formal introduction of the derivatives of perturbations in phase space, which is outside the scope of the paper. However—both these derivations are straightforward to carry out, and they are needed here basically to support the numerical examples below.

In view of eqs (106)–(109) it is straightforward to transform the fifth-order derivatives in Cartesian coordinates, M_{ijklm}, to corresponding derivatives in ray-centred coordinates, $\mathcal{M}_{abcde}$. There is, however, one issue that needs attention, as the transformation formulas contains terms of the type $p_i H_{iab}$, $p_i H_{iabc}$, $p_i H_{iabcd}$, $p_i H_{iabcde}$, etc. The only apparently problematic components among the fifth-order transformation derivatives in H_{iabcde} are $H_{ia3333} = \mathrm{d}^4 H_{ia}/\mathrm{d}\tau^4$—all other fifth-order derivatives are known from straightforward time differentiation of the fourth-order derivatives H_{iabcd}. The term $p_i\, \mathrm{d}^4 H_{ia}/\mathrm{d}\tau^4$ may however be computed without specific knowledge of the derivatives $\mathrm{d}^4 H_{ia}/\mathrm{d}\tau^4$, as we can utilize that the constraint $p_i H_{ia} = \delta_{3a}$ must hold along Ω. Indeed, we may use this constraint at any level of the transform, in terms of the relations

$$p_i H_{ia3} = -\eta_i H_{ia}, \tag{110}$$

$$p_i H_{ia33} = -\frac{\mathrm{d}\eta_i}{\mathrm{d}\tau} H_{ia} - 2\eta_i \frac{\mathrm{d}v_i}{\mathrm{d}\tau}, \tag{111}$$

$$p_i H_{ia333} = -\frac{\mathrm{d}^2\eta_i}{\mathrm{d}\tau^2} H_{ia} - 3\frac{\mathrm{d}\eta_i}{\mathrm{d}\tau}\frac{\mathrm{d}H_{ia}}{\mathrm{d}\tau} - 3\eta_i \frac{\mathrm{d}^2 H_{ia}}{\mathrm{d}\tau^2}, \tag{112}$$

$$p_i H_{ia3333} = -\frac{\mathrm{d}^3\eta_i}{\mathrm{d}\tau^3} H_{ia} - 4\frac{\mathrm{d}^2\eta_i}{\mathrm{d}\tau^2}\frac{\mathrm{d}H_{ia}}{\mathrm{d}\tau} - 6\frac{\mathrm{d}\eta_i}{\mathrm{d}\tau}\frac{\mathrm{d}^2 H_{ia}}{\mathrm{d}\tau^2} - 4\eta_i \frac{\mathrm{d}^3 H_{ia}}{\mathrm{d}\tau^3}, \text{ etc.} \tag{113}$$

Here, the time derivative $\mathrm{d}^3\eta_i/\mathrm{d}\tau^3$ relies on derivatives of the model parameters up to order four, which we have available.

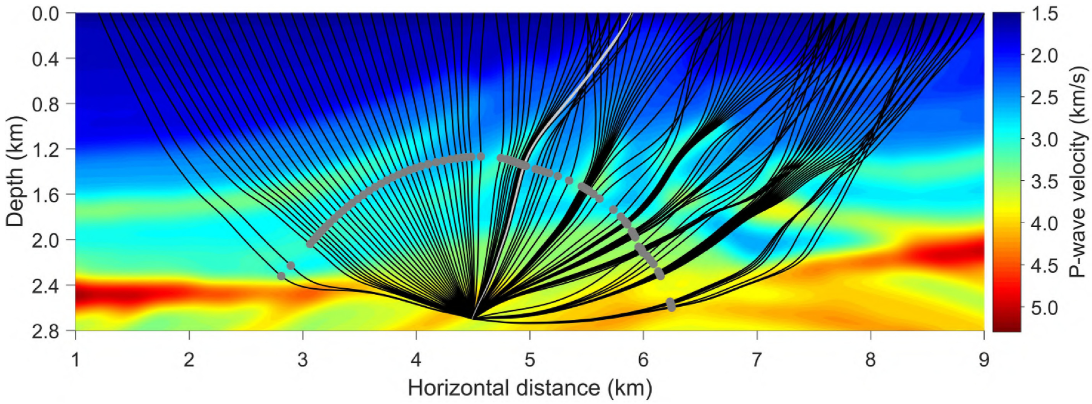

Figure 3. Marmousi isotropic model with source point ($x = 4.5$, $z = 2.7$) km. Rays have been computed to equidistant receivers located at zero depth. Grey dots show ray points corresponding to the given traveltime 0.44 s. Reference ray—light grey line; other rays—black lines.

6 ON THE VALIDITY OF RAY-CENTRED COORDINATES

We address the validity limits of ray-centred coordinates by computing the transformation coefficients $\Lambda^{11}_{ia}(q_1, q_2, q_3)$ at paraxial locations, not only on the reference ray Ω. Observe that on Ω we have $\Lambda^{11}_{ia}(0, 0, q_3) = H_{ia}(q_3)$.

In general, ray-centred coordinates are valid only in a certain region around the reference ray. In this region, the mapping between the three ray-centred coordinates and the three Cartesian coordinates must be one-to-one. A necessary condition to accomplish this is that the determinant of matrix $\mathbf{\Lambda}^{11}$ is strictly positive,

$$\det\mathbf{\Lambda}^{11} > 0, \tag{114}$$

which arises from the fact that the mapping collapses for $\det\mathbf{\Lambda}^{11} = 0$, while $\det\mathbf{\Lambda}^{11} > 0$ on Ω. We note that the condition (114) is not sufficient, as there are special cases where fulfillment of (114) does not correspond to a one-to-one mapping. The 6×6 matrix $\mathbf{\Lambda}$ is symplectic, so its determinant will be one as long as (114) is satisfied. All six phase-space coordinates of the Cartesian and ray-centred domains will then have a one-to-one relationship.

To enable a monitoring of criterion (114) we perform a paraxial expansion of the coefficients Λ^{11}_{ia} in the q_1q_2 plane. The expansion is linear in q_A,

$$\frac{\partial x_i}{\partial q_A}(q_1, q_2, q_3) \;=\; \mathcal{E}_{iA}(q_3), \qquad \frac{\partial x_i}{\partial q_3}(q_1, q_2, q_3) \;=\; v_i(q_3) + \dot{\mathcal{E}}_{iA}(q_3)\, q_A, \tag{115}$$

so the determinant to be checked in (114) is therefore

$$\det\mathbf{\Lambda}^{11}(q_1, q_2, q_3) \;=\; c(q_3)\,[1 - \eta_i(q_3)\,\mathcal{E}_{iA}(q_3)\, q_A]. \tag{116}$$

We observe that this determinant does not depend on the convention used for continuation of the basis vectors, although the second sub-equation of (115) indeed has this dependence. Eq. (116) can be used to monitor the behaviour of the coordinate transformation in a region around the reference ray.

7 NUMERICAL EXAMPLES

We illustrate by numerical examples some aspects of the above theory for the higher-order transformations between Cartesian and ray-centred coordinates. We wanted to use a quite drastically varying model for the experiments, to be able to address the limits of validity of the ray-centred coordinate system. Our choice was the well-known (isotropic, 2-D) Marmousi model. Before doing any tests, the P-wave velocity field of the model was smoothed in NORSAR, using a Hamming filter radius of 0.3 km. This was a necessity to make the model appropriate for ray tracing. We also made a VTI version of the smoothed model, by choosing Thomsen's (1986) parameters ϵ and δ as the constants $\epsilon = 0.3$, $\delta = 0.1$. The ratio of the vertical P- to S-wave velocity was set to 2 throughout the VTI model.

Moreover, since we are taking advantage of a model representation supporting quintic B-splines, which is not so common to use, we find it instructive to exemplify the function values and derivatives resulting when choosing different degrees of the B-spline representation.

In this section, we refer to the Cartesian coordinates x_1 and x_3 as x and z, respectively.

7.1 Rays and coordinate lines in the Marmousi isotropic model

In the Marmousi isotropic model (see Fig. 3), consider a source point at the location ($x = 4.5$, $z = 2.7$) km. Using two-point ray tracing, we compute rays from this source point to regularly spaced receivers in the acquisition surface, $z = 0$ km. Even though our model has

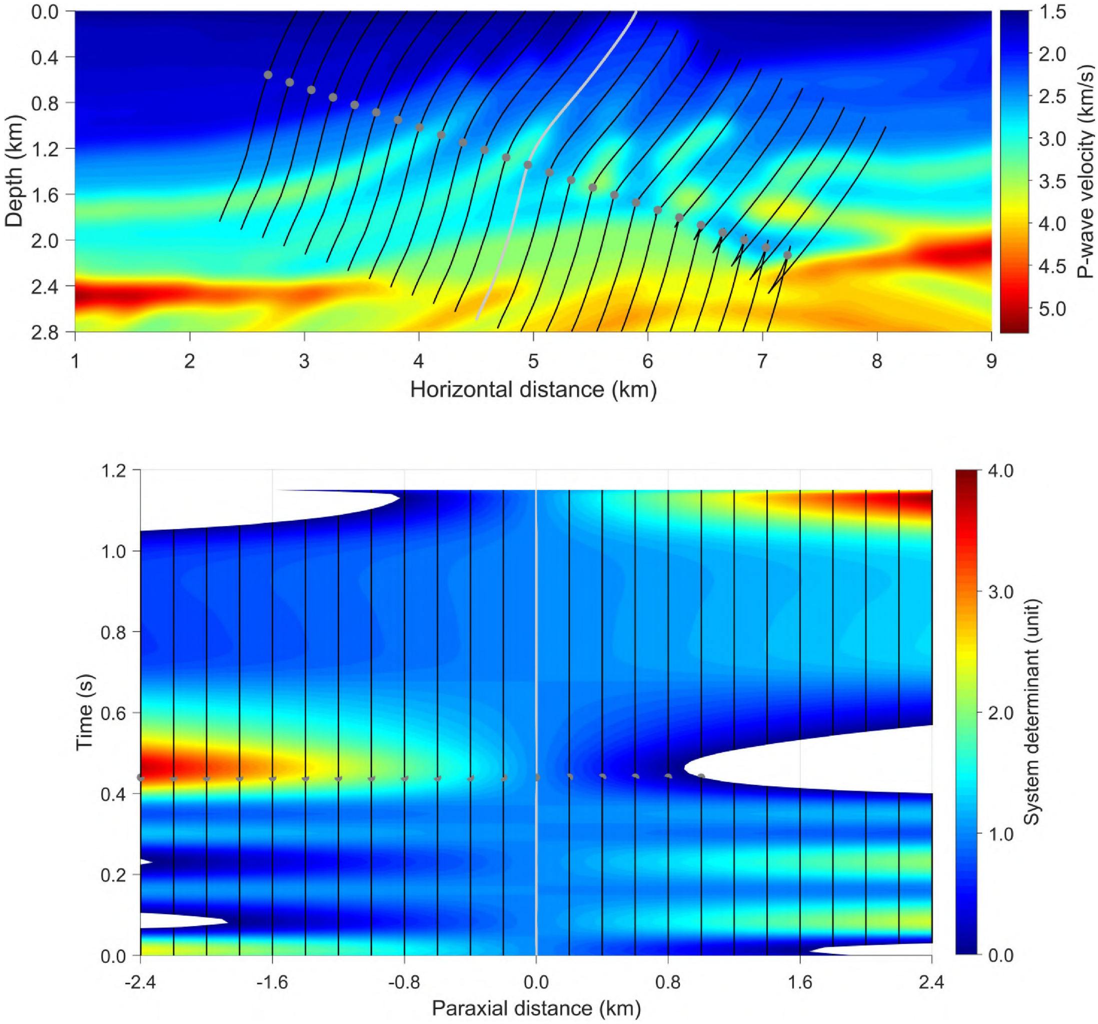

Figure 4. Top panel: Marmousi isotropic model and the reference ray (light grey) from Fig. 3. The black lines correspond to constant values of the q_1-coordinate, in the range $[-2.4, 2.4]$ km. Grey dots correspond to constant $q_3 = 0.44$ s. Bottom panel: The normalized determinant of the coordinate transformation matrix $\{\partial x_i/\partial q_j\}$ plotted in the ray-centred coordinates (q_1, q_3).

been smoothed, the ray field in Fig. 3 shows quite dramatic focusing/defocusing effects. We select one ray (light grey) as the reference ray. Moreover, in the part where this ray has strong curvature, we select a reference time, $\tau = 0.44$ s.

Next, in Fig. 4(top panel) we visualize the ray-centred coordinates that result for the given reference ray. The black lines in the figure are q_3-coordinate lines—they correspond to constant values of the q_1 coordinate. The grey dots refer to our selected reference time, $q_3 = \tau = 0.44$ s. We observe the evolvement of a triplication of the q_3 lines, starting around 1 km away from the reference ray. In Fig. 4(bottom panel) we have plotted the normalized determinant of the coordinate transformation matrix, given by eq. (116). In this context, 'normalized' means to divide by the phase velocity, $c(q_3)$.

In the white areas of the plot, the determinant is zero or negative, thus corresponding to a collapse of the ray-centred coordinate system. The mapping between Cartesian coordinates and ray-centred coordinates is then not one-to-one. This type of collapse is outlined by the point C in Fig. 2. The collapse is a result of the curvature of the ray path—it is not related to caustics of the ray field. We see that the formation of triplications in the coordinate lines (Fig. 4, top panel) is consistent with the system determinant becoming zero (Fig. 4, bottom panel).

7.2 Rays and coordinate lines in the Marmousi VTI model

We repeat the above exercise using the Marmousi VTI model (Fig. 5). It is then natural to choose a different reference ray, to ensure that we study a ray with significant curvature effects. Still, it seemed suitable to choose the same reference time, $\tau = 0.44$ s. As for the isotropic model, we obtain areas where the mapping between Cartesian coordinates and ray-centred coordinates is not one-to-one (Fig. 6, bottom panel). The deformation of the coordinate lines (Fig. 6, bottom panel) appears in a somewhat different form than for the isotropic case; the reason being that the coordinate lines not only depend on the curvature of the reference ray path—the angle between the slowness vector and the ray-velocity vector also plays a role.

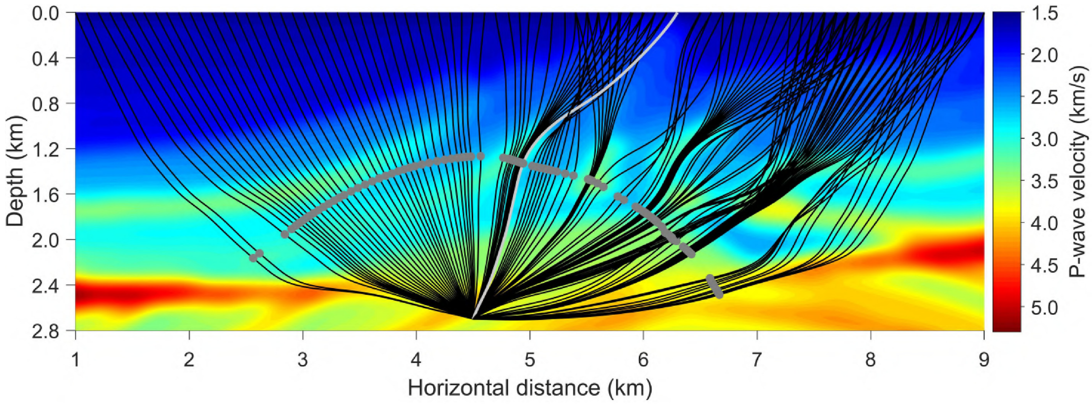

Figure 5. Marmousi VTI model with source point ($x = 4.5$, $z = 2.7$) km. Rays have been computed to equidistant receivers located at zero depth. Grey dots show ray points corresponding to the given traveltime 0.44 s. Reference ray—light grey line; other rays—black lines.

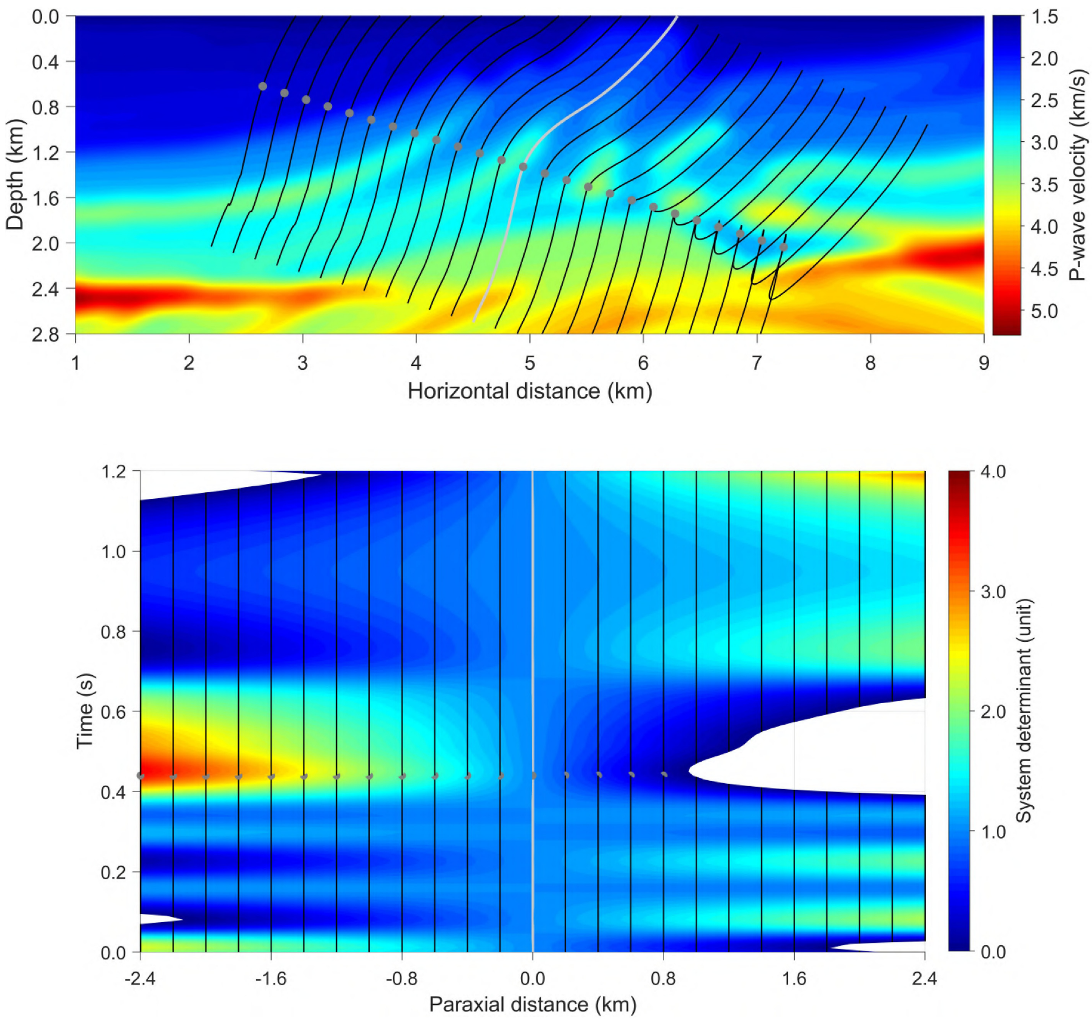

Figure 6. Top panel: Marmousi VTI model and the reference ray (light grey) from Fig. 5. The black lines correspond to constant values of the q_1-coordinate, in the range [−2.4, 2.4] km. Grey dots correspond to constant $q_3 = 0.44$ s. Bottom panel: The normalized determinant of the coordinate transformation matrix $\{\partial x_i/\partial q_j\}$ plotted in the ray-centred coordinates (q_1, q_3).

7.3 Traveltime expansions for an initial plane wave in the Marmousi VTI model

Based on our current reference point, that is the point on the reference ray that corresponds to the reference time, $q_3 = 0.44$ s, we now want to study the initiation of a plane wave. The initial slowness vector of the plane wave shall be equal to the slowness vector at the reference point, so that our reference ray for the initial-plane-wave ray field and the original common-source ray field will be the

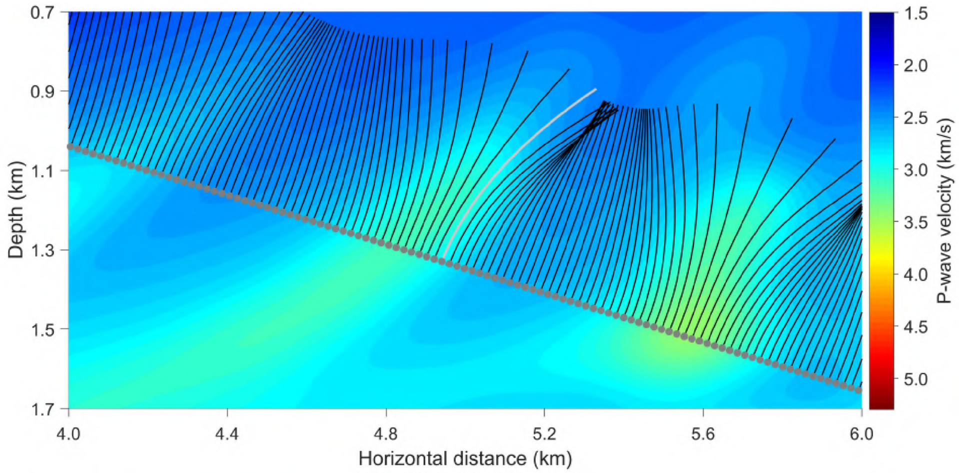

Figure 7. Initiation of a plane wave (grey dots) in the Marmousi VTI model, based on a point on the reference ray in Fig. 5 with traveltime 0.44 s from the original source point. All rays are traced a positive time 0.2 s away from the initial plane. Reference ray—light grey line; other rays—black lines.

same. In particular, we are interested in establishing higher-order approximations for the traveltime of the initial plane wave, given with respect to Cartesian coordinates as well as ray-centred coordinates. All computations are then done in the reference point, that is without using dynamic ray tracing. For the transformation of coefficients between the two coordinate systems we use the theory derived in this paper.

Fig. 7 shows a close-up of the Marmousi VTI model around the reference point. A number of rays (reference ray—light grey line; other rays—black lines) have been traced a positive time 0.2 s away from the initial plane (grey dots). Somewhat to the right of the reference ray, the formation of a caustic is visible.

Using the same close-up window as in Fig. 7, we have in Fig. 8 plotted four versions of the traveltime function arising from the plane-wave initial condition. The four versions corresponds to second-, third-, fourth- and fifth-order Taylor expansions, in Cartesian coordinates, around the reference point. Comparing the expansions of orders two and three, we see that the third-order expansion yields a somewhat better consistency with respect to the initial-plane-wave ray field. For the fourth-order expansion we observe a much better consistency on the right-hand side of the reference ray, where the caustic is located. Finally, the fifth-order expansion yields improved accuracy quite close to the reference ray but the consistency does not extend as far as for the fourth-order result.

In Fig. 9 we consider corresponding expansions of traveltime in ray-centred coordinates. The expansion of order two, which corresponds to the conventional dynamic ray tracing method, shows no lateral variation. Such variation appears with a third-order expansion of traveltime, but the degree of freedom is obviously too small to yield useful information at paraxial distances. For the fourth-order expansion a 'ridge' appears in the traveltime plot, to the right of the reference ray. This ridge indicates that paraxial rays experience a lower velocity than on the reference ray. In fact, a locally lower velocity is indeed the reason for the formation of the caustic to the right of the reference ray, as can be seen from Fig. 7. The fifth-order expansion in ray-centred coordinates provides one additional interesting detail: A local 'height' occurring at a paraxial distance of around 1 km. This local maximum is consistent with the previously mentioned collapse of the ray-centred coordinate system, see Fig. 5.

In summary, we find that Figs 8 and 9 illustrate well the increased information content inherent in the higher-order transformation of a local traveltime function between Cartesian and ray-centred coordinates.

7.4 Use of higher-order B-splines in the model representation

In this subsection we show an example of function values and derivatives that can result when different degrees are chosen for the B-spline function representing the discrete values of a given model parameter.

Fig. 10(top panel) shows a sequence of points (black circles), regularly spaced in the lateral direction, for which we want to generate a univariate function. In the literature of computer aided geometric design these input points are referred to synonymously as nodes, vertices, or control points. The first-degree spline (dashed black) is C^0 and corresponds to straightforward linear interpolation between these points. The function value is then continuous, while the first derivative is discontinuous. When choosing a cubic (magenta) and quintic (blue) B-spline function we observe the following: (1) The function does not traverse through the given input points; (2) The function gets progressively smoother as we increase the degree of the B-spline representation.

Consider the evaluation of a function value and eventual derivatives on a certain interval between two nodes. For a linear spline it is only these *two* nodes that determine the evaluation (by linear interpolation); to evaluate a cubic B-spline on the same interval we need the

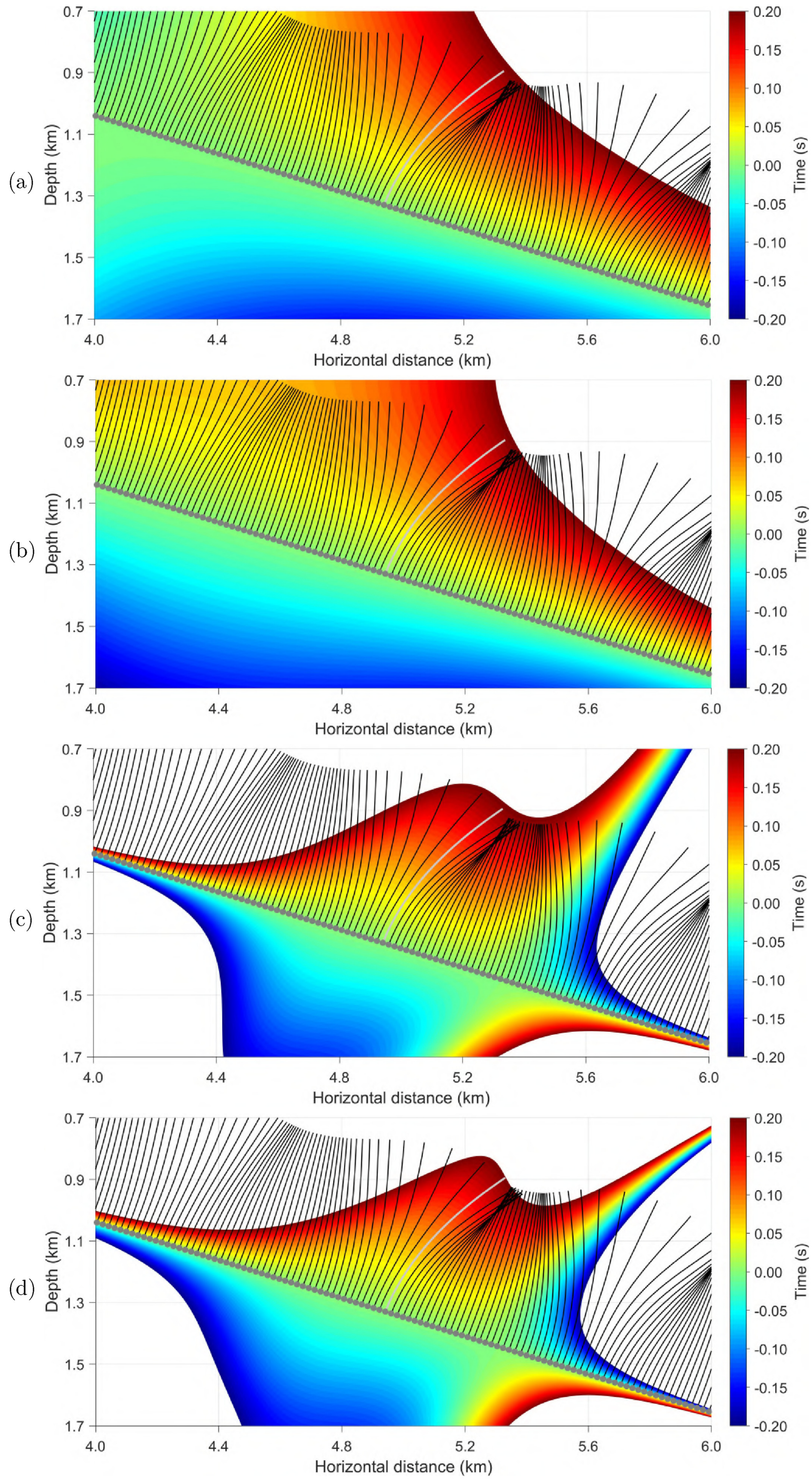

Figure 8. For an initial plane wave (grey dots) in the Marmousi VTI model: Traveltime extrapolation in Cartesian coordinates, to second (a), third (b), fourth (c) and fifth (d) order. Reference ray—light grey line; other rays—black lines.

two bounding nodes plus one node on each side, totally *four* nodes; evaluation of a quintic B-spline requires the two bounding nodes plus two nodes on each side, totally *six* nodes, and so forth.

In Fig. 10(bottom panel) the quintic 1-D B-spline representation in Fig. 10(top panel) has been restated as a volume function (in general a function of all three spatial coordinates, but here only shown in two dimensions). This is a simple example of a volume function well suited for seismic modelling. We note that cubic splines have been used in seismic modelling for a long time (Gjøystdal *et al.* 1985).

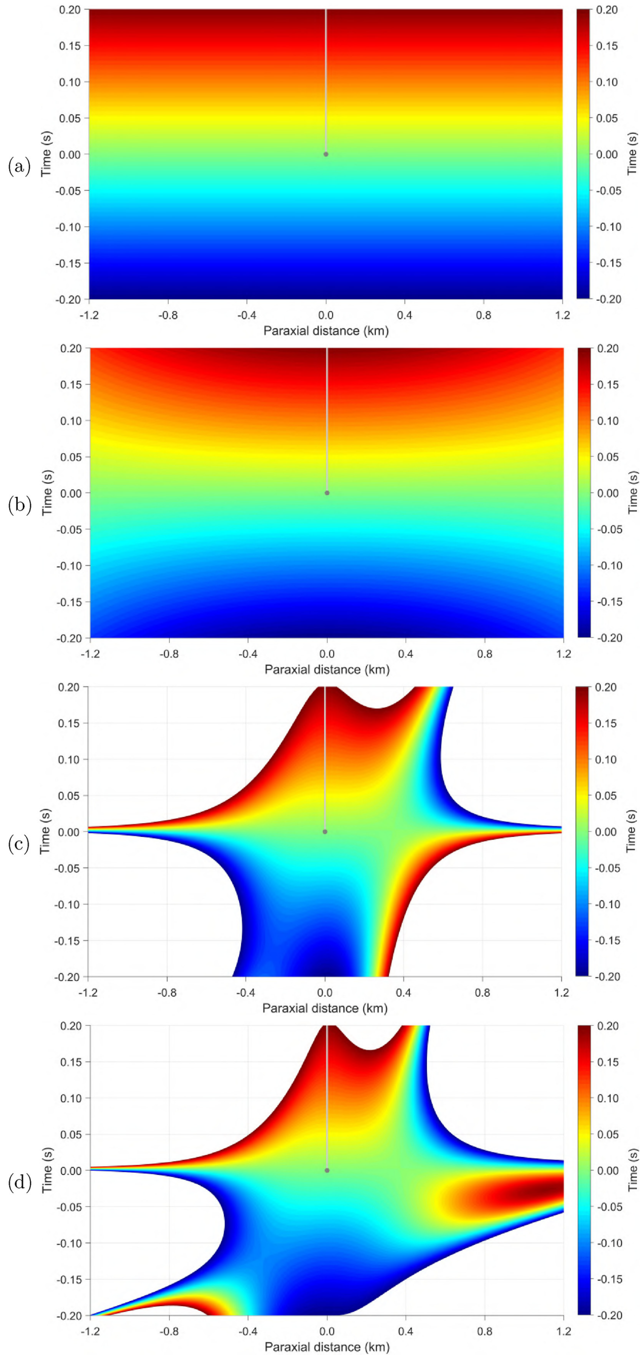

Figure 9. For an initial plane wave in the Marmousi VTI model: Traveltime extrapolation in ray-centred coordinates, to second (a), third (b), fourth (c) and fifth (d) order. Reference ray—light grey line; reference point—grey dot.

Fig. 11 shows graphs of first- (a) to fourth-order (d) derivatives evaluated using the cubic (magenta) and quintic (blue) B-spline functions shown in Fig. 10. The first-order derivatives of the cubic and quintic B-spline functions (Fig. 11a) are both continuous and smooth. We observe that the second-order derivative of the cubic B-spline function (Fig. 11b) is continuous but not smooth—the third-order derivative is discontinuous. For the quintic B-spline function, however, the second-order derivative (Fig. 11b) and the third-order derivative (Fig. 11c) are both continuous *and* smooth. Finally, the fourth-order derivative of the quintic B-spline function (Fig. 11d) is continuous but not smooth.

We find that Figs 10 and 11 illustrate well the stable behaviour of computed function values and derivatives of the quintic B-spline function, up to order three in the derivatives. For the fourth order the derivative is continuous, but in general not smooth. If more smoothness is needed, one can introduce a pre-processing step for the input (node) values. We further address the topic of higher-order B-spline functions in the Section 8.

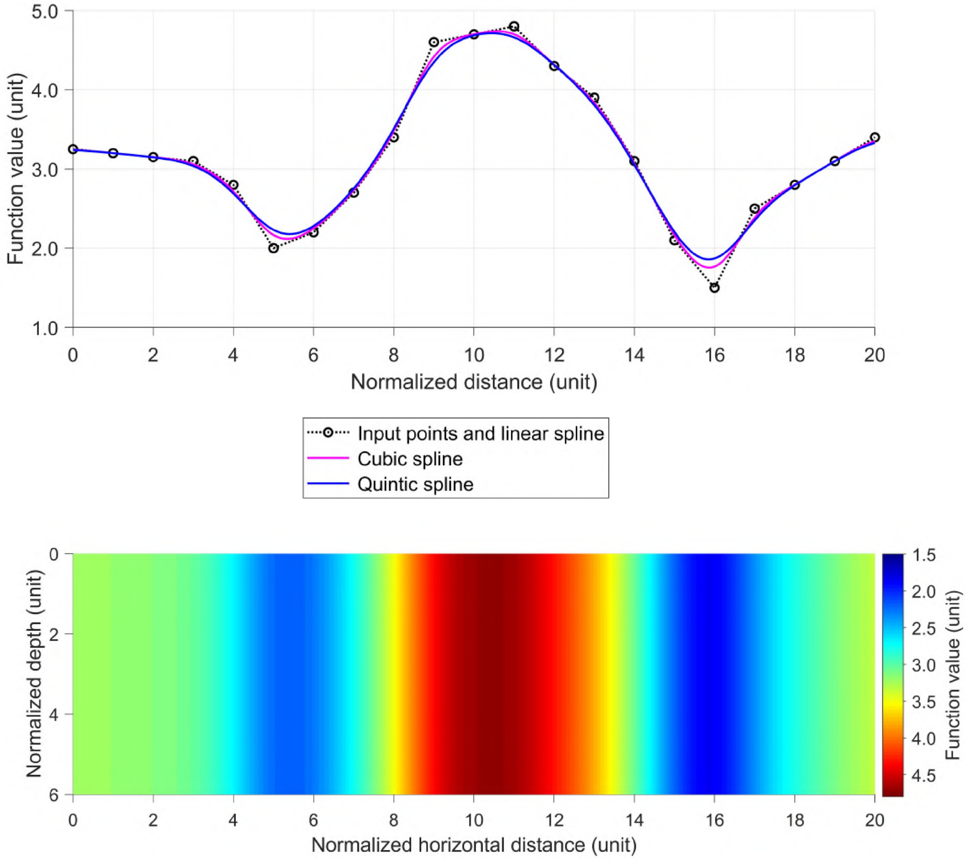

Figure 10. Top panel: B-spline representations of a sequence of node points (black circles) specified in one dimension only. First-degree (linear) B-spline representation—dotted black line; third-degree (cubic) B-spline representation—magenta line; fifth-degree (quintic) B-spline representation—blue line. Bottom panel: Using the quintic B-spline representation to form a volume function.

8 DISCUSSION

We discuss (1) the choice of model representation for computation of the higher-order transformation relating the phase-space perturbations in Cartesian and ray-centred coordinates and (2) the properties of B-splines in a higher-order context.

8.1 On the inclusion of higher-order derivatives of model parameters in the computation of transformation coefficients

The derived transformation coefficients do not depend explicitly on the derivatives of the model parameters (e.g. density-normalized elastic moduli). These derivatives are hidden in the ray-velocity vector, $\mathbf{v} = (v_i)$, and its time derivatives, and in the time derivatives of the eta vector, $\boldsymbol{\eta} = (\eta_i)$. In the tangent plane to the wavefront, first-order derivatives of the model parameters are sufficient to establish transformation coefficients of any order. Summarizing these last observations, it is clear that the extension of the (standard) first-order transformation approach to include higher-order transformation coefficients requires only a local, higher-order, extension of the Cartesian phase-space vector for the reference ray.

More specifically, a transformation to second order between Cartesian and ray-centred coordinates requires $\mathrm{d}v_i/\mathrm{d}\tau$ and $\mathrm{d}\eta_i/\mathrm{d}\tau$, while the third-order transformation needs $\mathrm{d}^2 v_i/\mathrm{d}\tau^2$ and $\mathrm{d}^2\eta_i/\mathrm{d}\tau^2$. We outline two different approaches to compute these time derivatives:

(i) By adapting a three times differentiable function to the computed discrete values of $[\mathbf{x}(\tau), \mathbf{p}(\tau)]$.
(ii) By using the first-order (conventional) and second-order dynamic ray tracing equations in Cartesian coordinates (Iversen *et al.* 2019), for which we need to know, respectively, the second- and third-order derivatives of the model parameters along the ray.

The first approach is simple, efficient and could typically be handled with a cubic B-spline representation of the model parameters. However, if the transformations are to be used in higher-order dynamic ray tracing, in which we utilize a quintic B-spline representation (Iversen *et al.* 2019), then one may risk inconsistencies resulting from the fact that the model representation is different in the computation of the transformation coefficients and in the specification/continuation of the dynamic ray tracing quantities. For an indication of the source of such potential inconsistencies, see Figs 10 and 11.

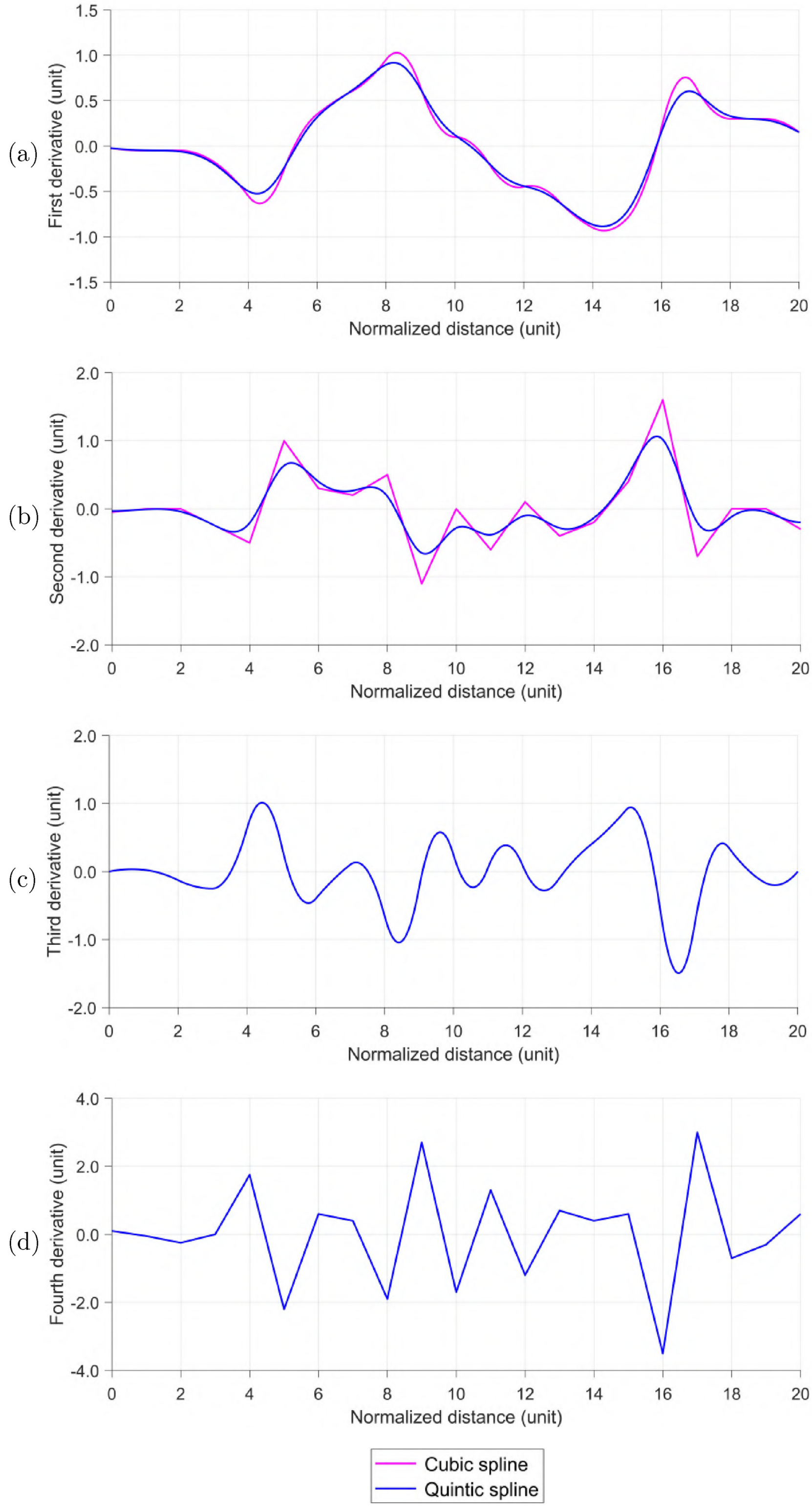

Figure 11. First-(a) to fourth-(d) order derivatives based on the cubic and quintic B-spline representations for the sequence of node points in Fig. 10.

If we want to specify a local wavefield, based on information in a single point on the reference ray, then we may need also derivatives of order two and higher in the coordinates q_1 and q_2. This is the case in the exploding plane wave situation, Figs 8 and 9. To evaluate the initial condition for the exploding plane wave, to all orders, we therefore use the quintic B-spline representation.

To be on the safe side, our recommendation is to use approach (ii) above for the computation of time derivatives along the ray, whenever the transformations are to be applied to the dynamic ray tracing quantities.

8.2 Aspects of the use of B-splines

There are certainly trade-offs involved in the choice of a model representation. B-splines have many good properties, but there are also effects that may be considered a disadvantage.

The B-spline function evaluation is local—this property makes it attractive with respect to efficiency, stability, and for model perturbation. Consider a certain model volumetric model parameter, and assume that the input points (nodes) correspond to a regular 3-D grid in the spatial coordinates. For a linear, cubic, and quintic B-spline the function evaluation inside a given grid cell requires, respectively, $2 \times 2 \times 2 = 8$, $4 \times 4 \times 4 = 64$ and $6 \times 6 \times 6 = 216$ model parameter values in the grid points of the current cell and its adjacent neighbors.

If we change the model parameter value in one of the grid points, then the function is affected only in a local region, the *region of influence*, surrounding that specific point. As with the function evaluation, the extent of the region depends on the B-spline degree—for linear, cubic, and quintic B-splines the region of influence consists of, respectively, 8, 64 and 216 grid cells.

We also discuss smoothness and stability. A B-spline curve does not traverse through the given input points, but is 'cutting the corners'. The function value is guaranteed to be within the *convex hull* formed by the input points (Farin *et al.* 2002). This is in contrast to an interpolating spline curve and a Bézier curve, for which the shape of the curve depends on *all* input points. It is noted the shape can be very sensitive to a perturbation of a single point, while a B-spline curve changes in a controlled way.

As is illustrated in Fig. 10(top panel), increasing the degree of the B-spline leads to increased smoothness. The second and the fourth derivative of, respectively, the cubic and the quintic B-spline are continuous, but not smooth (Fig. 11). To have a more smooth appearance of such derivatives a preprocessing of the input model parameter values is necessary. For the cubic spline this aspect has been known for years (e.g. Gjøystdal *et al.* 2002), but Fig. 11 (d) indicates that a pre-smoothing may be needed also for quintic B-splines.

When the latter point is handled, the quintic B-spline yields a stable computation of the value of a model parameter and its spatial derivatives up to order four. As a side effect, we note that the B-spline representation introduces a smoothing of the input parameter values, and this smoothing increases with the degree of the B-spline. The same is true for the computation time—it increases when more input points are needed in the evaluation of the function.

9 CONCLUSIONS

We have derived expressions for the second-, third-, and fourth-order coefficients of the transformation of phase-space perturbations from ray-centred coordinates to Cartesian coordinates and vice versa. The expressions depend on the (contra-variant and co-variant) sets of basis vectors related to ray-centred coordinates and on the time derivatives of these basis vectors. We provide a general formulation of the continuation of the contra-variant and co-variant basis along the reference ray.

With the transformation coefficients derived in this paper, the ground is prepared for introducing higher-order Hamilton–Jacobi perturbation equations (higher-order dynamic ray tracing) in ray-centred coordinates. For that purpose, we show and utilize that the 6×6 matrix of first-order derivatives relating the phase-space perturbations in ray-centred and Cartesian coordinates is symplectic. The symplecticity holds not only on the reference ray but at general locations in phase space complying with the validity region of ray-centred coordinates.

By numerical examples, we have used the transformation coefficients to visualize the validity region of ray-centred coordinates. A natural question is to what extent this validity region will be a problem in practice. For conventional dynamic ray tracing it seems unlikely that ray-centred coordinates may cause difficulties—on the other hand, as we have exemplified in this paper, the ability of the conventional method to describe paraxial variations of wavefield attributes in a relatively complex model is very limited. In total, the numerical examples show interesting effects for the traveltime extrapolations based on derivatives in a given point on the reference ray. The extrapolations to third, fourth and fifth order are progressively richer in information—we can see the effect of paraxial velocity changes, that potentially may lead to caustics. Moreover, the fifth-order result, in ray-centred coordinates, indicates that we can also see the imprint of the collapse of the coordinates.

We have given examples of the transformation of a traveltime function, for different extrapolation orders. We expect that equally significant results can be obtained for other wavefield properties, like geometrical spreading, amplitude coefficients, polarization vectors, and so forth. It would be interesting to see this further illuminated in future papers.

Furthermore, in view of our observations in this paper: what will happen first as we move paraxially away from the reference ray—a collapse of the wavefield (a caustic is encountered) or a collapse of the coordinate system? This is also a topic that requires more research.

The presented higher-order transformation theory for ray-centred coordinates is expected to be important in future developments of paraxial ray methods. Furthermore, the theory represents a good base for developing higher-order transformations related to more general curvilinear coordinate systems.

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 thanks the Research Council of Norway and the industry partners of the GAMES consortium at NTNU for financial support (grant No. 294404). EI and BU are grateful for being invited to meetings at Rice University. We have used academic software licenses from NORSAR and MathWorks. We are grateful for the feedback from the editorial board of *Geophysical Journal International*, represented by Louise Alexander, Jean Virieux and Michal Malinowski, and for the valuable technical reviews by Zvi Koren and Ludek Klimeš.

DATA AVAILABILITY STATEMENT

The input and output data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Bartles, R., Beatty, J. & Barsky, B., 1987. *An Introduction to the Splines in Computer Graphics*, Morean Kaufmann Publ.

Cameron, M.K., Fomel, S. & Sethian, J.A., 2007. Seismic velocity estimation from time migration, *Inverse Problems,* **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 Univ. Press.

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

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., 1979. Ray amplitudes of seismic body waves in laterally inhomogeneous media, *Geophys. J. R. astr. Soc.,* **57,** 91–106.

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., 2007. Seismic ray method: recent developments, *Adv. Geophys.,* **48,** 1–126.

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

Chavel, I., 2006. *Riemannian Geometry: A Modern Introduction,* Vol. **98,** Cambridge Univ. Press.

de Boor, C., 1972. On calculating with B-splines, *J. Approx. Theory,* **6**(1), 50–62.

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.

Gajewski, D. & Pšenčík, I., 1990. Vertical seismic profile synthetics by dynamic ray tracing in laterally varying layered anisotropic structures, *J. geophys. Res.,* **95**(B7), 11 301–11 315.

Gjøystdal, H., Reinhardsen, J. & Åstebøl, K., 1985. Computer representation of complex 3-D geological structures using a new 'solid modeling' technique, *Geophys. Prospect.,* **33**(8), 1195–1211.

Gjøystdal, H., Iversen, E., Laurain, R., Lecomte, I., Vinje, V. & Åstebøl, K., 2002. Review of ray theory applications in modelling and imaging of seismic data, *Stud. Geophys. Geod.,* **46**(2), 113–164.

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

Iversen, E., 2004. Reformulated kinematic and dynamic ray tracing systems for arbitrarily anisotropic media, *Stud. Geophys. Geod.,* **48,** 1–20.

Iversen, E., 2006. Amplitude, Fresnel zone, and NMO velocity for PP and SS normal-incidence reflections, *Geophysics,* **71**(02), W1–W14.

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., Ursin, B., Saksala, T., Ilmavirta, J. & de Hoop, M.V., 2019. Higher-order Hamilton-Jacobi perturbation theory for anisotropic heterogeneous media: dynamic ray tracing in Cartesian coordinates, *Geophys. J. Int.,* **216**(3), 2044–2070.

Iversen, E., Ursin, B., Saksala, T., Ilmavirta, J. & de Hoop, M.V., 2021. Higher-order Hamilton-Jacobi perturbation theory for anisotropic heterogeneous media: dynamic ray tracing in ray-centred coordinates, *Geophys. J. Int.,* doi:10.1093/gji/ggab152.

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

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

Klimeš, L., 2002. Second-order and higher-order perturbations of travel time in isotropic and anisotropic media, *Stud. Geophys. Geod.,* **46**(2), 213–248.

Klimeš, L., 2006a. Ray-centred coordinate systems in anisotropic media, *Stud. Geophys. Geod.,* **50,** 431–447.

Klimeš, L., 2006b. Common-ray tracing and dynamic ray tracing for S waves in a smooth elastic anisotropic medium, *Stud. Geophys. Geod.,* **50**(3), 449–461.

Klimeš, L., 2012. Zero-order ray-theory Green tensor in a heterogeneous anisotropic elastic medium, *Stud. Geophys. Geod.,* **56**(2), 373–382.

Klimeš, L., 2013. Relation between the propagator matrix of geodesic deviation and the second-order derivatives of the characteristic function, *J. Electromagn. Waves Appl.,* **27**(13), 1589–1601.

Klimeš, L., 2019. Calculation of the amplitudes of elastic waves in anisotropic media in Cartesian or ray-centred coordinates, *Stud. Geophys. Geod.,* **63**(2), 229–246.

Koren, Z. & Ravve, I., 2021. Eigenrays in 3D heterogeneous anisotropic media, part I: kinematics, *Geophys. Prospect.,* **69**(1), 3–27.

Nowak, G., 2011. A de Casteljau algorithm for q-Bernstein-Stancu polynomials, in *Abstract and Applied Analysis,* Vol. **2011,** Hindawi.

Popov, M.M. & Pšenčík, I., 1978. Computation of ray amplitudes in inhomogeneous media with curved interfaces, *Stud. Geophys. Geod.,* **22,** 248–258.

Ravve, I. & Koren, Z., 2021. Eigenrays in 3D heterogeneous anisotropic media, part II: dynamics, *Geophys. Prospect.,* **69**(1), 28–52.

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

Vinje, V., Iversen, E. & Gjøystdal, H., 1993. Traveltime and amplitude estimation using wavefront construction, *Geophysics,* **58**(8), 1157–1166.

APPENDIX A: CONTINUATION OF BASIS VECTORS ALONG THE REFERENCE RAY

We consider the continuation of the contra-variant and co-variant sets of basis vectors, $\mathcal{E}$ and $\mathcal{F}$, along the reference ray Ω. The process is typically organized by integrating the required differential equations for $\mathcal{E}$, i.e., we need $\mathrm{d}\mathcal{E}/\mathrm{d}\tau$ expressed in terms of $\mathcal{E}$ and quantities that are known on Ω. Once $\mathcal{E}$ and $\mathrm{d}\mathcal{E}/\mathrm{d}\tau$ are known for some location on Ω, then $\mathcal{F}$ and $\mathrm{d}\mathcal{F}/\mathrm{d}\tau$ can be computed by explicit formulas.

Alternatively, one may integrate a system of differential equations for $\mathcal{F}$ and from them obtain, by explicit expressions, the quantities $\mathcal{E}$ and $\mathrm{d}\mathcal{E}/\mathrm{d}\tau$.

A1 Continuation of the contra-variant basis

The generally non-orthonormal basis $\mathcal{E}$ can be related to an orthonormal basis $\bar{\mathcal{E}}$ by

$$\mathcal{E} = \bar{\mathcal{E}}\mathcal{C}, \tag{A1}$$

where $\mathcal{C}$ is a 2×2 matrix that accounts for rotation as well as a possible stretch within the plane orthogonal to the slowness vector $\hat{\mathbf{p}}(\tau)$. A rotation does not change the lengths of the two vectors in $\bar{\mathcal{E}}$, while a stretch may imply a change of relative length as well as of measurement unit. We use the symmetric matrix

$$\mathcal{B} = \{\mathbf{e}_A{}^T \mathbf{e}_B\} = \mathcal{E}^T \mathcal{E} \tag{A2}$$

to describe eventual deviation of the basis $\mathcal{E}$ from orthonormality, which is the case when $\mathcal{B} = \{\delta_{AB}\}$. As a consequence, eq. (A2) can also be expressed as

$$\mathcal{B} = \mathcal{C}^T \mathcal{C}. \tag{A3}$$

Let c and $\mathbf{n}$ be the phase velocity and the unit vector corresponding to the slowness vector $\mathbf{p}$, so that $\mathbf{p} = c^{-1}\mathbf{n}$. As the vectors in $\mathcal{E}$ and the unit vector $\mathbf{n}$ are linearly independent, any arbitrary vector, $\mathbf{s} = (s_i)$, may therefore be expressed as the linear combination

$$s_i = \mathcal{E}_{iB}\, a_B + n_i a_3, \tag{A4}$$

where a_1, a_2, and a_3 are coefficients. The corresponding expression for a matrix $\mathbf{S} = \{S_{iB}\}$ consisting of two arbitrary column vectors reads

$$S_{iC} = \mathcal{E}_{iB} \mathcal{A}_{BC} + n_i \mathcal{A}_{3C}. \tag{A5}$$

In the following, we elaborate on the nature of the coefficients $\mathcal{A}_{BC}$ and $\mathcal{A}_{3C}$ in the case where $S_{iC} = \mathrm{d}\mathcal{E}_{iC}/\mathrm{d}\tau$.

The basis vectors in $\mathcal{E}$ have to be orthogonal to $\mathbf{p}$, i.e., $\mathbf{p}^T\mathcal{E} = \{0_A\}$, implying that

$$\mathbf{p}^T \frac{\mathrm{d}\mathcal{E}}{\mathrm{d}\tau} = -\frac{\mathrm{d}\mathbf{p}^T}{\mathrm{d}\tau}\mathcal{E}, \tag{A6}$$

which can be recast as

$$\mathbf{n}^T \frac{\mathrm{d}\mathcal{E}}{\mathrm{d}\tau} = -c\frac{\mathrm{d}\mathbf{p}^T}{\mathrm{d}\tau}\mathcal{E}. \tag{A7}$$

It is clear that the coefficients $\{\mathcal{A}_{3C}\}$ are given by the right-hand side of eq. (A7), so we can write

$$\frac{\mathrm{d}\mathcal{E}}{\mathrm{d}\tau} = \mathcal{E}\mathcal{A} - \mathbf{c}\frac{\mathrm{d}\mathbf{p}^T}{\mathrm{d}\tau}\mathcal{E}. \tag{A8}$$

Here, $\mathbf{c} = c^2\mathbf{p}$ is the *phase-velocity vector*, and $\mathcal{A} = \{\mathcal{A}_{BC}\}$ is a 2×2 matrix, which we have not yet specified.

We elaborate on the specification of matrix $\mathcal{A}$. A first observation, based on eqs (A2) and (A7), is that $\mathcal{A}$ must satisfy

$$\mathcal{A} = \mathcal{B}^{-1}\mathcal{E}^T\frac{\mathrm{d}\mathcal{E}}{\mathrm{d}\tau}, \tag{A9}$$

Eqs (A2) and (A9) further yields

$$\frac{\mathrm{d}\mathcal{B}}{\mathrm{d}\tau} = \mathcal{A}^T\mathcal{B} + \mathcal{B}\mathcal{A}. \tag{A10}$$

Hence, if matrix $\mathcal{A}$ is zero then matrix $\mathcal{B}$ is constant along the ray Ω. In particular, if the basis $\mathcal{E}$ is orthonormal at the starting point, this property will be preserved along the whole ray.

The situation

$$\mathcal{A} = \{0_{AB}\}; \qquad \mathcal{B} = \{\delta_{AB}\}, \tag{A11}$$

has been referred to as the *standard option* for the bases of the ray-centred coordinate system (Červený 2001; Klimeš 2006a, section 5.4). It is used, for example, in the Complete Ray Tracing (CRT) software package (Klimeš 2006b, section 6.2) .

A2 Computation of the co-variant basis and its time derivative

The co-variant basis $\mathcal{F}$ can always be expressed in terms of the contra-variant basis $\mathcal{E}$ via

$$\mathcal{F} = \alpha\mathcal{E}\mathcal{B}^{-1}, \tag{A12}$$

with the components of $\boldsymbol{\alpha}$ given by eq. (37). In view of the definition of the 2×2 matrix $\boldsymbol{\mathcal{K}}$ in eq. (62), we can use eqs (A8) and (A12) to obtain

$$\boldsymbol{\mathcal{K}}^T = \boldsymbol{\mathcal{A}} + \boldsymbol{\mathcal{B}}^{-1}\boldsymbol{\mathcal{E}}^T\mathbf{v}\boldsymbol{\eta}^T\boldsymbol{\mathcal{E}}. \tag{A13}$$

Moreover, since $\boldsymbol{\mathcal{E}}^T\boldsymbol{\mathcal{F}} = \{\delta_{AB}\}$ and $\mathbf{v}^T\boldsymbol{\mathcal{F}} = \{0_A\}$, we have

$$\boldsymbol{\mathcal{E}}^T\frac{\mathrm{d}\boldsymbol{\mathcal{F}}}{\mathrm{d}\tau} = -\boldsymbol{\mathcal{K}}, \qquad \mathbf{v}^T\frac{\mathrm{d}\boldsymbol{\mathcal{F}}}{\mathrm{d}\tau} = -\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}\tau}\boldsymbol{\mathcal{F}}. \tag{A14}$$

Eq. (A14) then yields the time derivative of $\boldsymbol{\mathcal{F}}$ as

$$\frac{\mathrm{d}\boldsymbol{\mathcal{F}}}{\mathrm{d}\tau} = -\boldsymbol{\mathcal{F}}\boldsymbol{\mathcal{K}} - \mathbf{p}\frac{\mathrm{d}\mathbf{v}^T}{\mathrm{d}\tau}\boldsymbol{\mathcal{F}}. \tag{A15}$$

Note that eq. (A15) also follows from Klimeš (2006a, eqs 24 and 26).

A3 Equivalent formulations for the continuation of the contra-variant basis

By eq. (A8) we established a general formulation for the continuation of the contra-variant basis $\boldsymbol{\mathcal{E}}$ along the reference ray. Eq. (A8) is expressed in terms of the *phase-velocity vector*, $\mathbf{c}$.

With the help of eq. (A13) we can obtain a system of differential equations for $\boldsymbol{\mathcal{E}}$ that is fully equivalent to (A8),

$$\frac{\mathrm{d}\boldsymbol{\mathcal{E}}}{\mathrm{d}\tau} = \boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{K}}^T - \mathbf{v}\frac{\mathrm{d}\mathbf{p}^T}{\mathrm{d}\tau}\boldsymbol{\mathcal{E}}; \tag{A16}$$

see also Klimeš (2006a, eqs 24 and 25). In eq. (A16) it is now the *ray-velocity vector*, $\mathbf{v}$, that appears on the right-hand side. The phase-velocity vector is related to the ray-velocity vector by

$$\mathbf{c} = \left[\{\delta_{ij}\} - \boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{B}}^{-1}\boldsymbol{\mathcal{E}}^T\right]\mathbf{v}, \tag{A17}$$

or equivalently,

$$\mathbf{c} = \left[\{\delta_{ij}\} - \bar{\boldsymbol{\mathcal{E}}}\bar{\boldsymbol{\mathcal{E}}}^T\right]\mathbf{v}. \tag{A18}$$

A4 Continuation of the contra-variant basis— the orthonormality case

If $\mathbf{e}_1$ and $\mathbf{e}_2$ are orthogonal unit vectors, eqs (A2) and (A10) yields

$$\boldsymbol{\mathcal{A}} = -\boldsymbol{\mathcal{A}}^T = \mathcal{A}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \tag{A19}$$

where $\mathcal{A}$ is a scalar function of τ. It is common, but not mandatory, to define $\boldsymbol{\mathcal{E}}(\tau)$ strictly in terms of 1) the curvature and torsion of the ray trajectory and 2) the initial condition, $\boldsymbol{\mathcal{E}}(\tau_0)$; see, e.g., Červený & Pšenčík (1979). The time derivative vectors $\mathrm{d}\mathbf{e}_A/\mathrm{d}\tau$, $A = 1, 2$, will then both be parallel to $\mathbf{p}$, which yields

$$\mathcal{A}(\tau) = 0 \tag{A20}$$

along the ray. Eq. (A20) represents the standard option in eq. (A11) in the situation that $\boldsymbol{\mathcal{E}}$ is orthonormal. One can interpret $\mathcal{A}(\tau)$ as a function specifying an additional constraint on the rotation of $\boldsymbol{\mathcal{E}}$, i.e., a constraint not related to curvature and torsion of the ray trajectory.

Using eqs (A2) and (A12), it follows that for an orthonormal basis $\boldsymbol{\mathcal{E}}$ of any kind [$\mathcal{A}(\tau)$ may be nonzero in eq. (A19)] we have

$$\boldsymbol{\mathcal{F}} = \boldsymbol{\alpha}\boldsymbol{\mathcal{E}}. \tag{A21}$$

APPENDIX B: PARTIAL DERIVATIVES OF CARTESIAN MOMENTUM COORDINATES

B1 First-order partial derivatives of Cartesian momentum coordinates

We obtain first-order partial derivatives of p_i with respect to q_a.

Using eq. (51) in eq. (53) we find different forms for the derivative $\partial p_i/\partial q_3$,

$$\frac{\partial p_i}{\partial q_3} = p_a^{(q)}\frac{\partial^2 p_i}{\partial q_3 \partial p_a^{(q)}} = p_a^{(q)}\frac{\partial}{\partial q_3}\left(\frac{\partial p_i}{\partial p_a^{(q)}}\right) = p_a^{(q)}\frac{\partial}{\partial q_3}\left(\frac{\partial q_a}{\partial x_i}\right), \tag{B1}$$

where the last expression can also be stated

$$\frac{\partial p_i}{\partial q_3} = p_A^{(q)}\frac{\partial}{\partial q_3}\left(\frac{\partial q_A}{\partial x_i}\right) + p_3^{(q)}\frac{\partial}{q_3}\left(\frac{\partial q_3}{\partial x_i}\right). \tag{B2}$$

Furthermore, we differentiate the first sub-equation of (27) with respect to x_i, which yields

$$\frac{\partial^2 q_c}{\partial x_i \partial x_j}\frac{\partial x_j}{\partial q_a} + \frac{\partial q_c}{\partial x_j}\frac{\partial q_b}{\partial x_i}\frac{\partial^2 x_j}{\partial q_a \partial q_b} = 0, \tag{B3}$$

and therefore,

$$\frac{\partial x_j}{\partial q_a}\frac{\partial^2 q_c}{\partial x_i \partial x_j} = -\frac{\partial q_c}{\partial x_j}\frac{\partial q_b}{\partial x_i}\frac{\partial^2 x_j}{\partial q_a \partial q_b}. \tag{B4}$$

With the help of eq. (54) it is clear that

$$\frac{\partial^2 p_i}{\partial q_a \partial p_c^{(q)}} = -\frac{\partial q_c}{\partial x_j}\frac{\partial q_b}{\partial x_i}\frac{\partial^2 x_j}{\partial q_a \partial q_b}. \tag{B5}$$

Using eq. (53) further yields

$$\frac{\partial p_i}{\partial q_a} = -p_c^{(q)}\frac{\partial q_c}{\partial x_j}\frac{\partial q_b}{\partial x_i}\frac{\partial^2 x_j}{\partial q_a \partial q_b}. \tag{B6}$$

In order to find an expression for $\partial p_i/\partial q_A$ we use the property

$$\frac{\partial}{\partial q_3}\left(\frac{\partial q_c}{\partial x_j}\frac{\partial x_j}{\partial q_A}\right) = 0, \tag{B7}$$

which follows from eq. (27). We substitute A for a and elaborate on eq. (B6),

$$\begin{aligned}
\frac{\partial p_i}{\partial q_A} &= -p_c^{(q)}\frac{\partial q_c}{\partial x_j}\frac{\partial q_b}{\partial x_i}\frac{\partial^2 x_j}{\partial q_A \partial q_b}\delta_{l3} \\
&= -\frac{\partial q_3}{\partial x_i}p_m^{(q)}\frac{\partial q_m}{\partial x_j}\frac{\partial^2 x_j}{\partial q_3 \partial q_A} \\
&= \frac{\partial q_3}{\partial x_i}p_c^{(q)}\frac{\partial}{\partial q_3}\left(\frac{\partial q_c}{\partial x_j}\right)\frac{\partial x_j}{\partial q_A}.
\end{aligned}$$

Using also eq. (B1) we obtain the result

$$\frac{\partial p_i}{\partial q_A} = \frac{\partial q_3}{\partial x_i}\frac{\partial p_j}{\partial q_3}\frac{\partial x_j}{\partial q_A}. \tag{B8}$$

The derivatives $\partial p_i/\partial q_A$ and $\partial p_i/\partial q_3$ in eqs (B8) and (B2) correspond to general phase-space locations. As such they can be used to obtain higher-order derivatives. Evaluation on Ω yields

$$\frac{\partial p_i}{\partial q_A} = p_i \eta_j \mathcal{E}_{jA}, \qquad \frac{\partial p_i}{\partial q_3} = \eta_i. \tag{B9}$$

B2 A property of the symplectic transformation matrix

We derive an expression for the first-order derivatives $\partial p_a^{(q)}/\partial x_i$ of the transformation from Cartesian coordinates to ray-centred coordinates.

Differentiate eqs (70) and (75) with respect to the Cartesian position coordinate,

$$\frac{\partial p_a^{(q)}}{\partial x_i} = p_j \frac{\partial^2 p_a^{(q)}}{\partial x_i \partial p_j}, \qquad \frac{\partial^2 p_a^{(q)}}{\partial x_i \partial p_j} = \frac{\partial q_b}{\partial x_i}\frac{\partial^2 x_j}{\partial q_a \partial q_b},$$

and combine the latter expressions with eq. (B5). This yields,

$$\begin{aligned}
\frac{\partial p_a^{(q)}}{\partial x_i} &= p_j \frac{\partial q_b}{\partial x_i}\frac{\partial^2 x_j}{\partial q_a \partial q_b} = -p_j \frac{\partial x_j}{\partial q_c}\frac{\partial^2 p_i}{\partial q_a \partial p_c^{(q)}} \\
&= -p_j \frac{\partial p_c^{(q)}}{\partial p_j}\frac{\partial^2 p_i}{\partial q_a \partial p_c^{(q)}} \\
&= -p_c^{(q)}\frac{\partial^2 p_i}{\partial q_a \partial p_c^{(q)}} \\
&= -\frac{\partial}{\partial q_a}\left(p_c^{(q)}\frac{\partial p_i}{\partial p_c^{(q)}}\right),
\end{aligned}$$

and hence the important result

$$\frac{\partial p_a^{(q)}}{\partial x_i} = -\frac{\partial p_i}{\partial q_a}. \tag{B10}$$

B3 Higher-order partial derivatives of Cartesian momentum coordinates

We obtain general second- and third-order partial derivatives of Cartesian momentum coordinates.

The right-hand side of eq. (51) is not a function of $p_a^{(q)}$, so higher-order derivatives with respect to $p_a^{(q)}$ are therefore zero. Hence,

$$\frac{\partial^2 p_i}{\partial p_a^{(q)} \partial p_b^{(q)}} = 0, \tag{B11}$$

$$\frac{\partial^3 p_i}{\partial p_a^{(q)} \partial p_b^{(q)} \partial p_c^{(q)}} = 0, \qquad \frac{\partial^3 p_i}{\partial p_a^{(q)} \partial p_b^{(q)} \partial q_c} = 0. \tag{B12}$$

Moreover, the general result in eq. (B8) can equivalently be written

$$\frac{\partial p_i}{\partial q_A} = \frac{\partial p_i}{\partial p_3^{(q)}} \frac{\partial p_j}{\partial q_3} \frac{\partial x_j}{\partial q_A}. \tag{B13}$$

Differentiation of this equation with respect to $p_b^{(q)}$ yields

$$\frac{\partial^2 p_i}{\partial q_A \partial p_b^{(q)}} = \frac{\partial p_i}{\partial p_3^{(q)}} \frac{\partial^2 p_j}{\partial q_3 \partial p_b^{(q)}} \frac{\partial x_j}{\partial q_A}, \tag{B14}$$

hence,

$$\frac{\partial^2 p_i}{\partial q_A \partial p_b^{(q)}} = \frac{\partial q_3}{\partial x_i} \frac{\partial}{\partial q_3} \left(\frac{\partial q_b}{\partial x_j} \right) \frac{\partial x_j}{\partial q_A}. \tag{B15}$$

Moreover, one can write

$$\frac{\partial^2 p_i}{\partial q_3 \partial p_b^{(q)}} = \frac{\partial}{\partial q_3} \left(\frac{\partial q_b}{\partial x_i} \right). \tag{B16}$$

We evaluate eqs (B15) and (B16) on the reference ray, which gives

$$\begin{aligned} &\frac{\partial^2 p_i}{\partial q_A \partial p_B^{(q)}} = -p_i \mathcal{K}_{AB}, \qquad \frac{\partial^2 p_i}{\partial q_A \partial p_3^{(q)}} = p_i \eta_j \mathcal{E}_{jA}, \\ &\frac{\partial^2 p_i}{\partial q_3 \partial p_A^{(q)}} = \dot{\mathcal{F}}_{iA}, \qquad \frac{\partial^2 p_i}{\partial q_3 \partial p_3^{(q)}} = \eta_i. \end{aligned} \tag{B17}$$

For definition of the quantity $\mathcal{K}_{AB}$, see eq. (62).

From eqs (B9) and (B17) the results for higher-order partial derivatives of p_i with respect to q_3 follow readily,

$$\frac{\partial^2 p_i}{\partial q_3^2} = \dot{\eta}_i, \qquad \frac{\partial^2 p_i}{\partial q_3 \partial q_A} = \frac{\mathrm{d}}{\mathrm{d}\tau} \left(p_i \eta_j \mathcal{E}_{jA} \right), \tag{B18}$$

$$\frac{\partial^3 p_i}{\partial q_3^3} = \ddot{\eta}_i, \qquad \frac{\partial^3 p_i}{\partial q_3^2 \partial q_A} = \frac{\mathrm{d}^2}{\mathrm{d}\tau^2} \left(p_i \eta_j \mathcal{E}_{jA} \right), \qquad \frac{\partial^3 p_i}{\partial q_3^2 \partial p_A^{(q)}} = \ddot{\mathcal{F}}_{iA}, \tag{B19}$$

$$\frac{\partial^3 p_i}{\partial q_3 \partial q_A \partial p_B^{(q)}} = -\frac{\mathrm{d}}{\mathrm{d}\tau} \left(p_i \mathcal{K}_{AB} \right), \qquad \frac{\partial^3 p_i}{\partial q_3^2 \partial p_3^{(q)}} = \dot{\eta}_i, \qquad \frac{\partial^3 p_i}{\partial q_3 \partial q_A \partial p_3^{(q)}} = \frac{\mathrm{d}}{\mathrm{d}\tau} \left(p_i \eta_j \mathcal{E}_{jA} \right). \tag{B20}$$

We proceed to consider the momentum components p_i differentiated twice with respect to the ray-centred position coordinates, q_A. Using eq. (B13) as our starting point, we obtain

$$\begin{aligned} \frac{\partial^2 p_i}{\partial q_A \partial q_B} &= \left(\frac{\partial^2 p_i}{\partial p_3^{(q)} \partial q_B} \frac{\partial p_j}{\partial q_3} + \frac{\partial p_i}{\partial p_3^{(q)}} \frac{\partial^2 p_j}{\partial q_3 \partial q_B} \right) \frac{\partial x_j}{\partial q_A} \\ &= \frac{\partial q_3}{\partial x_i} \frac{\partial^2 p_k}{\partial q_3 \partial p_3^{(q)}} \frac{\partial x_k}{\partial q_B} \frac{\partial p_j}{\partial q_3} \frac{\partial x_j}{\partial q_A} + \frac{\partial p_i}{\partial p_3^{(q)}} \frac{\partial}{\partial q_3} \left(\frac{\partial q_3}{\partial x_j} \right) \frac{\partial p_k}{\partial q_3} \frac{\partial x_k}{\partial q_B} \frac{\partial x_j}{\partial q_A}, \end{aligned}$$

which can be restated

$$\frac{\partial^2 p_i}{\partial q_A \partial q_B} = \frac{\partial q_3}{\partial x_i} \left(\frac{\partial}{\partial q_3} \left(\frac{\partial q_3}{\partial x_k} \right) \frac{\partial p_j}{\partial q_3} + \frac{\partial}{\partial q_3} \left(\frac{\partial q_3}{\partial x_j} \right) \frac{\partial p_k}{\partial q_3} \right) \frac{\partial x_j}{\partial q_A} \frac{\partial x_k}{\partial q_B}. \tag{B21}$$

The evaluated expression on the reference ray is

$$\frac{\partial^2 p_i}{\partial q_A \partial q_B} = 2 p_i \eta_j \eta_k \mathcal{E}_{jA} \mathcal{E}_{kB}. \tag{B22}$$

As a consequence, we also get

$$\frac{\partial^3 p_i}{\partial q_3 \partial q_A \partial q_B} = 2 \frac{\mathrm{d}}{\mathrm{d}\tau} \left(p_i \eta_j \eta_k \mathcal{E}_{jA} \mathcal{E}_{kB} \right). \tag{B23}$$

Eq. (B21) is equivalently written

$$\frac{\partial^2 p_i}{\partial q_A \partial q_B} = \frac{\partial q_3}{\partial x_i}\left(\frac{\partial^2 p_k}{\partial q_3 \partial p_3^{(q)}}\frac{\partial p_j}{\partial q_3} + \frac{\partial^2 p_j}{\partial q_3 \partial p_3^{(q)}}\frac{\partial p_k}{\partial q_3}\right)\frac{\partial x_j}{\partial q_A}\frac{\partial x_k}{\partial q_B}. \tag{B24}$$

Differentiation with respect to $p_c^{(q)}$ then yields the third-order (mixed) derivative

$$\frac{\partial^3 p_i}{\partial q_A \partial q_B \partial p_c^{(q)}} = \frac{\partial q_3}{\partial x_i}\left(\frac{\partial^2 p_k}{\partial q_3 \partial p_3^{(q)}}\frac{\partial^2 p_j}{\partial q_3 \partial p_c^{(q)}} + \frac{\partial^2 p_j}{\partial q_3 \partial p_3^{(q)}}\frac{\partial^2 p_k}{\partial q_3 \partial p_c^{(q)}}\right)\frac{\partial x_j}{\partial q_A}\frac{\partial x_k}{\partial q_B}, \tag{B25}$$

which can also be stated

$$\frac{\partial^3 p_i}{\partial q_A \partial q_B \partial p_c^{(q)}} = -\frac{\partial q_3}{\partial x_i}\left(\frac{\partial^2 p_k}{\partial q_3 \partial p_3^{(q)}}\frac{\partial x_k}{\partial q_B}\frac{\partial}{\partial q_3}\left(\frac{\partial x_j}{\partial q_A}\right)\frac{\partial q_c}{\partial x_j} + \frac{\partial^2 p_j}{\partial q_3 \partial p_3^{(q)}}\frac{\partial x_j}{\partial q_A}\frac{\partial}{\partial q_3}\left(\frac{\partial x_k}{\partial q_B}\right)\frac{\partial q_c}{\partial x_k}\right). \tag{B26}$$

Evaluation on the reference ray gives

$$\frac{\partial^3 p_i}{\partial q_A \partial q_B \partial p_C^{(q)}} = -p_i \eta_j (\mathcal{E}_{jA}\mathcal{K}_{BC} + \mathcal{E}_{jB}\mathcal{K}_{AC}), \qquad \frac{\partial^3 p_i}{\partial q_A \partial q_B \partial p_3^{(q)}} = 2 p_i \eta_j \eta_k \mathcal{E}_{jA}\mathcal{E}_{kB}. \tag{B27}$$

To obtain the general third-order derivative $\partial p_i/\partial q_A \partial q_B \partial q_C$ we do further differentiation of eq. (B24). The derivation is straightforward and follows the same principles as above. On the reference ray we find,

$$\frac{\partial^3 p_i}{\partial q_A \partial q_B \partial q_C} = 6 p_i \eta_j \eta_k \eta_l \mathcal{E}_{jA}\mathcal{E}_{kB}\mathcal{E}_{lC}. \tag{B28}$$

Certain fourth-order derivatives of p_i have a role in the derivation of third-order Hamilton-Jacobi perturbation differential equations. It is straightforward, although space demanding, to develop expressions for these derivatives. Therefore, we state here only the evaluated results on the reference ray,

$$\frac{\partial^4 p_i}{\partial q_A \partial q_B \partial q_C \partial q_D} = 24\, p_i\, \eta_j \eta_k \eta_l \eta_m\, \mathcal{E}_{jA}\mathcal{E}_{kB}\mathcal{E}_{lC}\mathcal{E}_{mD}, \tag{B29}$$

$$\frac{\partial^4 p_i}{\partial q_A \partial q_B \partial q_C \partial p_D^{(q)}} = -2\, p_i\, \eta_j \eta_k\, (\mathcal{E}_{jA}\mathcal{E}_{kB}\mathcal{K}_{CD} + \mathcal{E}_{jA}\mathcal{E}_{kC}\mathcal{K}_{BD} + \mathcal{E}_{jB}\mathcal{E}_{kC}\mathcal{K}_{AD}). \tag{B30}$$

APPENDIX C: SECOND-ORDER PARTIAL DERIVATIVES OF RAY-CENTRED POSITION COORDINATES

We develop specific expressions for the second-order partial derivatives of ray-centred position coordinates, q_a.

Eq. (54) yields

$$\begin{aligned}\frac{\partial^2 q_a}{\partial x_i \partial x_j} &= \frac{\partial^2 p_i}{\partial q_c \partial p_a^{(q)}}\frac{\partial q_c}{\partial x_j} \\ &= \frac{\partial^2 p_i}{\partial q_C \partial p_a^{(q)}}\frac{\partial q_C}{\partial x_j} + \frac{\partial^2 p_i}{\partial q_3 \partial p_a^{(q)}}\frac{\partial q_3}{\partial x_j}.\end{aligned} \tag{C1}$$

Inserting from eqs (B15) and (B16) gives

$$\begin{aligned}\frac{\partial^2 q_a}{\partial x_i \partial x_j} &= \frac{\partial^2 p_i}{\partial q_c \partial p_a^{(q)}}\frac{\partial q_c}{\partial x_j} \\ &= \frac{\partial^2 p_i}{\partial q_C \partial p_a^{(q)}}\frac{\partial q_C}{\partial x_j} + \frac{\partial^2 p_i}{\partial q_3 \partial p_a^{(q)}}\frac{\partial q_3}{\partial x_j} \\ &= \frac{\partial q_3}{\partial x_i}\frac{\partial}{\partial q_3}\left(\frac{\partial q_a}{\partial x_m}\right)\frac{\partial x_m}{\partial q_C}\frac{\partial q_C}{\partial x_j} + \frac{\partial}{\partial q_3}\left(\frac{\partial q_a}{\partial x_i}\right)\frac{\partial q_3}{\partial x_j} \\ &= \frac{\partial q_3}{\partial x_i}\frac{\partial}{\partial q_3}\left(\frac{\partial q_a}{\partial x_m}\right)\left(\delta_{jm} - \frac{\partial q_3}{\partial x_j}\frac{\partial x_m}{\partial q_3}\right) + \frac{\partial}{\partial q_3}\left(\frac{\partial q_a}{\partial x_i}\right)\frac{\partial q_3}{\partial x_j} \\ &= \left(\frac{\partial q_3}{\partial x_i}\delta_{jm} + \frac{\partial q_3}{\partial x_j}\delta_{im} - \frac{\partial q_3}{\partial x_i}\frac{\partial q_3}{\partial x_j}\frac{\partial x_m}{\partial q_3}\right)\frac{\partial}{\partial q_3}\left(\frac{\partial q_a}{\partial x_m}\right),\end{aligned} \tag{C2}$$

where we utilized eq. (31). Eq. (C2) corresponds to a general location in phase space. We write this equation compactly as

$$\frac{\partial^2 q_a}{\partial x_i \partial x_j} = R_{ij}^m \frac{\partial}{\partial q_3}\left(\frac{\partial q_a}{\partial x_m}\right), \tag{C3}$$

where the operator R_{ij}^{m} has two equivalent forms—

$$R_{ij}^{m} = \frac{\partial q_3}{\partial x_i}\delta_{jm} + \frac{\partial q_3}{\partial x_j}\delta_{im} - \frac{\partial q_3}{\partial x_i}\frac{\partial q_3}{\partial x_j}\frac{\partial x_m}{\partial q_3} \tag{C4}$$

and

$$R_{ij}^{m} = \frac{\partial q_3}{\partial x_i}\alpha_{jm} + \frac{\partial q_3}{\partial x_j}\alpha_{im} + \frac{\partial q_3}{\partial x_i}\frac{\partial q_3}{\partial x_j}\frac{\partial x_m}{\partial q_3}. \tag{C5}$$

The quantity α_{jm} is defined in eq. (32).

Application of eqs (C3) and (C4) on the ray Ω yields

$$H_{aij}^{\dagger} = R_{ij}^{m}\,\dot{H}_{am}^{\dagger}, \tag{C6}$$

with the operator R_{ij}^{m} given equivalently by

$$R_{ij}^{m} = p_i\alpha_{jm} + p_j\alpha_{im} + p_i p_j v_m \tag{C7}$$

or

$$R_{ij}^{m} = p_i\delta_{jm} + p_j\delta_{im} - p_i p_j v_m. \tag{C8}$$

APPENDIX D: THIRD-ORDER PARTIAL DERIVATIVES OF RAY-CENTRED POSITION COORDINATES

We derive expressions for the third-order partial derivatives of the ray-centred position coordinates,

$$\frac{\partial^3 q_a}{\partial x_i \partial x_j \partial x_k}.$$

Multiplying this derivative expression with the first-order derivatives of the transformation from ray-centred to Cartesian coordinates leads to component quantities of the form

$$\begin{aligned} A_{cde}^{a} &\equiv \frac{\partial^3 q_a}{\partial x_i \partial x_j \partial x_k}\frac{\partial x_i}{\partial q_c}\frac{\partial x_j}{\partial q_d}\frac{\partial x_k}{\partial q_e} \\ &= \left(\frac{\partial^3 p_i}{\partial p_a^{(q)}\partial q_d \partial q_e} - \frac{\partial^2 q_a}{\partial x_i \partial x_j}\frac{\partial^2 x_j}{\partial q_d \partial q_e}\right)\frac{\partial x_i}{\partial q_c}. \end{aligned} \tag{D1}$$

It is straightforward to show that the components A_{CDE}^{a} are zero. Moreover, we utilize in the following that

$$\frac{\partial x_i}{\partial q_3}\frac{\partial^3 q_a}{\partial x_i \partial x_j \partial x_k} = \frac{\partial}{\partial q_3}\left(\frac{\partial^2 q_a}{\partial x_j \partial x_k}\right). \tag{D2}$$

For brevity, we introduce on the reference ray Ω a notation for the higher-order derivatives,

$$H_{aij}^{\dagger} = \frac{\partial^2 q_a}{\partial x_i \partial x_j}, \quad H_{aijk}^{\dagger} = \frac{\partial^3 q_a}{\partial x_i \partial x_j \partial x_k}, \text{ etc.,} \tag{D3}$$

so that eq. (D2) can be restated

$$v_i\,H_{aijk}^{\dagger} = \dot{H}_{ajk}^{\dagger} \tag{D4}$$

on Ω. The third-order derivatives of ray-centred position coordinates can now be written in terms of the co-variant basis as

$$\begin{aligned} \frac{\partial^3 q_a}{\partial x_i \partial x_j \partial x_k} = {} & A_{3DE}^{a}\left(p_i\mathcal{F}_{jD}\mathcal{F}_{kE} + p_i\mathcal{F}_{kD}\mathcal{F}_{jE} + p_j\mathcal{F}_{kD}\mathcal{F}_{iE}\right) \\ & + A_{33E}^{a}\left(p_i p_j\mathcal{F}_{kE} + p_i p_k\mathcal{F}_{jE} + p_j p_k\mathcal{F}_{iE}\right) \\ & + A_{333}^{a}\,p_i p_j p_k, \end{aligned} \tag{D5}$$

with the A-quantities given by

$$A_{3DE}^{a} = \dot{H}_{amn}^{\dagger}\mathcal{E}_{mD}\mathcal{E}_{nE}, \tag{D6}$$

$$A_{33E}^{a} = \dot{H}_{amn}^{\dagger}v_m\mathcal{E}_{nE}, \tag{D7}$$

$$A_{333}^{a} = \dot{H}_{amn}^{\dagger}v_m v_n. \tag{D8}$$

Applying eqs (D6) and (D8) in (D5) yields

$$
\begin{aligned}
H^{\dagger}_{aijk} = {} & \dot{H}^{\dagger}_{amn}\mathcal{E}_{mD}\mathcal{E}_{nE}\left(p_i\mathcal{F}_{jD}\mathcal{F}_{kE} + p_j\mathcal{F}_{iD}\mathcal{F}_{kE} + p_k\mathcal{F}_{iD}\mathcal{F}_{jE}\right) \\
& + \dot{H}^{\dagger}_{amn}v_m\mathcal{E}_{nE}\left(p_ip_j\mathcal{F}_{kE} + p_ip_k\mathcal{F}_{jE} + p_jp_k\mathcal{F}_{iE}\right) \\
& + \dot{H}^{\dagger}_{amn}v_mv_n\,p_ip_jp_k.
\end{aligned} \tag{D9}
$$

We can therefore write

$$
H^{\dagger}_{aijk} = R^{mn}_{ijk}\,\dot{H}^{\dagger}_{amn}, \tag{D10}
$$

where we can use eq. (37) and the symmetry of indices m and n to state the operator R^{mn}_{ijk} as

$$
\begin{aligned}
R^{mn}_{ijk} = {} & p_i\alpha_{jm}\alpha_{kn} + p_j\alpha_{km}\alpha_{in} + p_k\alpha_{im}\alpha_{jn} \\
& + \left(p_ip_j\alpha_{kn} + p_jp_k\alpha_{in} + p_kp_i\alpha_{jn}\right)v_m \\
& + p_ip_jp_kv_mv_n.
\end{aligned} \tag{D11}
$$

Further use of eq. (37) yields the equivalent form

$$
\begin{aligned}
R^{mn}_{ijk} = {} & p_i\delta_{jm}\delta_{kn} + p_j\delta_{km}\delta_{in} + p_k\delta_{im}\delta_{jn} \\
& - \left(p_ip_j\delta_{kn} + p_jp_k\delta_{in} + p_kp_i\delta_{jn}\right)v_m \\
& + p_ip_jp_kv_mv_n.
\end{aligned} \tag{D12}
$$

To evaluate the right-hand side of eq. (D10) we need the time derivative of the second-order coefficients, see eq. (C6). This time derivative is expressed in terms of the first and second time derivatives of the first-order coefficients,

$$
\dot{H}^{\dagger}_{aij} = \dot{R}^{m}_{ij}\dot{H}^{\dagger}_{am} + R^{m}_{ij}\ddot{H}^{\dagger}_{am}. \tag{D13}
$$

Here, we can obtain $\dot{R}^{m}_{ij}$ by straightforward differentiation of eq. (C8), which yields

$$
\dot{R}^{m}_{ij} = \eta_i\alpha_{jm} + \eta_j\alpha_{im} - p_ip_j\dot{v}_m. \tag{D14}
$$

APPENDIX E: FOURTH-ORDER PARTIAL DERIVATIVES OF RAY-CENTRED POSITION COORDINATES

In order to obtain the fourth-order derivatives of ray-centred position coordinates we introduce the component quantities

$$
\begin{aligned}
A^{a}_{cdef} & \equiv \frac{\partial^4 q_a}{\partial x_i\partial x_j\partial x_k\partial x_l}\frac{\partial x_i}{\partial q_c}\frac{\partial x_j}{\partial q_d}\frac{\partial x_k}{\partial q_e}\frac{\partial x_l}{\partial q_f} \\
& = \left[\frac{\partial^4 p_i}{\partial p^{(q)}_a\partial q_d\partial q_e\partial q_f} - \frac{\partial^3 q_a}{\partial x_i\partial x_j\partial x_k}\left(\frac{\partial^2 x_j}{\partial q_d\partial q_e}\frac{\partial x_k}{\partial q_f} + \frac{\partial^2 x_j}{\partial q_d\partial q_f}\frac{\partial x_k}{\partial q_e} + \frac{\partial^2 x_j}{\partial q_e\partial q_f}\frac{\partial x_k}{\partial q_d}\right) - \frac{\partial^2 q_a}{\partial x_i\partial x_j}\frac{\partial^3 x_j}{\partial q_d\partial q_e\partial q_f}\right]\frac{\partial x_i}{\partial q_c}.
\end{aligned} \tag{E1}
$$

Here, we note that the components A^{a}_{CDEF} are zero. We also emphasize the importance of the relation

$$
\frac{\partial x_i}{\partial q_3}\frac{\partial^4 q_a}{\partial x_i\partial x_j\partial x_k\partial x_l} = \frac{\partial}{\partial q_3}\left(\frac{\partial^3 q_a}{\partial x_j\partial x_k\partial x_l}\right), \tag{E2}
$$

which, on the reference ray, can be written in short-hand notation as

$$
v_i\,H^{\dagger}_{aijkl} = \dot{H}^{\dagger}_{ajkl}. \tag{E3}
$$

From eq. (E1) it follows that the fourth-order derivatives $H^{\dagger}_{aijkl}$, taken on the reference ray, can be expressed in terms of the co-variant basis as

$$
\begin{aligned}
H^{\dagger}_{aijkl} = {} & A^{a}_{3DEF}\left(p_i\mathcal{F}_{jD}\mathcal{F}_{kE}\mathcal{F}_{lF} + p_j\mathcal{F}_{iD}\mathcal{F}_{kE}\mathcal{F}_{lF} + p_k\mathcal{F}_{iD}\mathcal{F}_{jE}\mathcal{F}_{lF} + p_l\mathcal{F}_{iD}\mathcal{F}_{jE}\mathcal{F}_{kF}\right) \\
& + A^{a}_{33EF}\left(p_ip_j\mathcal{F}_{kE}\mathcal{F}_{lF} + p_ip_k\mathcal{F}_{jE}\mathcal{F}_{lF} + p_ip_l\mathcal{F}_{jE}\mathcal{F}_{kF} + p_jp_k\mathcal{F}_{iE}\mathcal{F}_{lF} + p_jp_l\mathcal{F}_{iE}\mathcal{F}_{kF} + p_kp_l\mathcal{F}_{iE}\mathcal{F}_{jF}\right) \\
& + A^{a}_{333F}\left(p_ip_jp_k\mathcal{F}_{lF} + p_ip_jp_l\mathcal{F}_{kF} + p_ip_kp_l\mathcal{F}_{jF} + p_jp_kp_l\mathcal{F}_{iF}\right) + A^{a}_{3333}\,p_ip_jp_kp_l,
\end{aligned} \tag{E4}
$$

where the A-components are

$$
A^{a}_{3DEF} = \dot{H}^{\dagger}_{amnq}\mathcal{E}_{mD}\mathcal{E}_{nE}\mathcal{E}_{qF}, \tag{E5}
$$

$$
A^{a}_{33EF} = \dot{H}^{\dagger}_{amnq}v_m\mathcal{E}_{nE}\mathcal{E}_{qF}, \tag{E6}
$$

$$
A^{a}_{333F} = \dot{H}^{\dagger}_{amnq}v_mv_n\mathcal{E}_{qF}, \tag{E7}
$$

$$A^{a}_{3333} = \dot{H}^{\dagger}_{amnq} v_m v_n v_q. \tag{E8}$$

Using eqs (E5)–(E8) in eq. (E4) yields the result

$$H^{\dagger}_{aijkl} = R^{mnq}_{ijkl}\, \dot{H}^{\dagger}_{amnq}, \tag{E9}$$

with the operator R^{mnq}_{ijkl} given by

$$\begin{aligned} R^{mnq}_{ijkl} &= p_i \alpha_{jm}\alpha_{kn}\alpha_{lq} + p_j \alpha_{im}\alpha_{kn}\alpha_{lq} + p_k \alpha_{im}\alpha_{jn}\alpha_{lq} + p_l \alpha_{im}\alpha_{jn}\alpha_{kq} \\ &\quad + (p_i p_j \alpha_{kn}\alpha_{lq} + p_i p_k \alpha_{jn}\alpha_{lq} + p_i p_l \alpha_{jn}\alpha_{kq} + p_j p_k \alpha_{in}\alpha_{lq} + p_j p_l \alpha_{in}\alpha_{kq} + p_k p_l \alpha_{in}\alpha_{jq})\, v_m \\ &\quad + (p_i p_j p_k \alpha_{lq} + p_i p_j p_l \alpha_{kq} + p_i p_k p_l \alpha_{jq} + p_j p_k p_l \alpha_{iq})\, v_m v_n + \ p_i p_j p_k p_l\, v_m v_n v_q. \end{aligned} \tag{E10}$$

To compute the right-hand side of eq. (E9) we need the time derivative of the third-order coefficients of the transformation from Cartesian to ray-centred coordinates, given by eq. (D10). The differentiation can be carried out as follows,

$$\dot{H}^{\dagger}_{aijk} = \dot{R}^{mn}_{ijk}\, \dot{H}^{\dagger}_{amn} + R^{mn}_{ijk}\, \ddot{H}^{\dagger}_{amn}, \tag{E11}$$

where

$$\frac{\mathrm{d}^2 H^{\dagger}_{amn}}{\mathrm{d}\tau^2} = \frac{\mathrm{d}^2 R^{q}_{mn}}{\mathrm{d}\tau^2}\frac{\mathrm{d}H^{\dagger}_{aq}}{\mathrm{d}\tau} + 2\frac{\mathrm{d}R^{q}_{mn}}{\mathrm{d}\tau}\frac{\mathrm{d}^2 H^{\dagger}_{aq}}{\mathrm{d}\tau^2} + R^{q}_{mn}\frac{\mathrm{d}^3 H^{\dagger}_{aq}}{\mathrm{d}\tau^3}. \tag{E12}$$

APPENDIX F: DETAILS OF THE MAPPING OF TRAVELTIME DERIVATIVES FROM RAY-CENTRED TO CARTESIAN COORDINATES

We elaborate on the details of the mapping of the derivatives of traveltime, from ray-centred to Cartesian coordinates. The objective is to write explicit expressions in terms of the co-variant basis vector components, $\mathcal{F}_{iA}$ and p_i, of the ray-centred coordinates. This can be useful, because many derivatives of traveltime in ray-centred coordinates are zero. It is clear, however, that when addressing derivatives of traveltime of orders four and higher, one comes to a point where it is more practical to write the mapping in terms of the inverse transformation matrix components, $H^{\dagger}_{ai}$.

Similarly to in the first sub-equation of (18), we use that the first derivatives of traveltime are related by

$$\frac{\partial \tau}{\partial x_i} = \frac{\partial \tau}{\partial q_a}\frac{\partial q_a}{\partial x_i}. \tag{F1}$$

Differentiation of eq. (F1) then yields

$$\frac{\partial^2 \tau}{\partial x_i \partial x_j} = \frac{\partial \tau}{\partial q_a}\frac{\partial^2 q_a}{\partial x_i \partial x_j} + \frac{\partial^2 \tau}{\partial q_a \partial q_b}\frac{\partial q_a}{\partial x_i}\frac{\partial q_b}{\partial x_j}. \tag{F2}$$

On Ω we use that

$$\frac{\partial \tau}{\partial q_e} = \delta_{3e}, \qquad \frac{\partial^2 \tau}{\partial q_e \partial q_f} = \frac{\partial^2 \tau}{\partial q_A \partial q_B}\delta_{Ae}\delta_{Bf}, \tag{F3}$$

Defining $M_{ij} \equiv \partial^2\tau/\partial x_i \partial x_j$ and $\mathcal{M}_{AB} \equiv \partial^2\tau/\partial q_A \partial q_B$, we therefore have

$$M_{ij} = H^{\dagger}_{3ij} + \mathcal{M}_{AB}\, \mathcal{F}_{iA}\mathcal{F}_{jB}. \tag{F4}$$

We further differentiate eq. (F2), to third order in the spatial coordinates,

$$\begin{aligned} \frac{\partial^3 \tau}{\partial x_i \partial x_j \partial x_k} &= \frac{\partial \tau}{\partial q_a}\frac{\partial^3 q_a}{\partial x_i \partial x_j \partial x_k} \\ &\quad + \frac{\partial^2 \tau}{\partial q_a \partial q_b}\left(\frac{\partial q_a}{\partial x_i}\frac{\partial^2 q_b}{\partial x_j \partial x_k} + \frac{\partial q_a}{\partial x_j}\frac{\partial^2 q_b}{\partial x_i \partial x_k} + \frac{\partial q_a}{\partial x_k}\frac{\partial^2 q_b}{\partial x_i \partial x_j}\right) \\ &\quad + \frac{\partial^3 \tau}{\partial q_a \partial q_b \partial q_c}\frac{\partial q_a}{\partial x_i}\frac{\partial q_b}{\partial x_j}\frac{\partial q_c}{\partial x_k}. \end{aligned} \tag{F5}$$

In the evaluation on Ω we utilize that

$$\begin{aligned} \frac{\partial^3 \tau}{\partial q_e \partial q_f \partial q_g} &= \frac{\partial^3 \tau}{\partial q_A \partial q_B \partial q_C}\delta_{Ae}\delta_{Bf}\delta_{Cg} \\ &\quad + \frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial^2 \tau}{\partial q_A \partial q_B}\right)(\delta_{3e}\delta_{Af}\delta_{Bg} + \delta_{3f}\delta_{Ae}\delta_{Bg} + \delta_{3g}\delta_{Ae}\delta_{Bf}), \end{aligned} \tag{F6}$$

which yields

$$\begin{aligned}
M_{ijk} = {} & H^{\dagger}_{3ijk} \\
& + \mathcal{M}_{AB}\left(\mathcal{F}_{iA}H^{\dagger}_{Bjk} + \mathcal{F}_{jA}H^{\dagger}_{Bik} + \mathcal{F}_{kA}H^{\dagger}_{Bij}\right) \\
& + \dot{\mathcal{M}}_{AB}\left(p_i\mathcal{F}_{jA}\mathcal{F}_{kB} + p_j\mathcal{F}_{iA}\mathcal{F}_{kB} + p_k\mathcal{F}_{iA}\mathcal{F}_{jB}\right) \\
& + \mathcal{M}_{ABC}\,\mathcal{F}_{iA}\mathcal{F}_{jB}\mathcal{F}_{kC}.
\end{aligned} \tag{F7}$$

In the same manner, we obtain an equation for transformation of the fourth-order derivatives of traveltime, from ray-centred to Cartesian coordinates. The evaluated result on Ω is

$$\begin{aligned}
M_{ijkl} = {} & H^{\dagger}_{3ijkl} + \mathcal{M}_{AB}\left(\mathcal{F}_{iA}H^{\dagger}_{Bjkl} + \mathcal{F}_{jA}H^{\dagger}_{Bikl} + \mathcal{F}_{kA}H^{\dagger}_{Bijl} + \mathcal{F}_{lA}H^{\dagger}_{Bijk} + H^{\dagger}_{Aij}H^{\dagger}_{Bkl} + H^{\dagger}_{Aik}H^{\dagger}_{Bjl} + H^{\dagger}_{Ail}H^{\dagger}_{Bjk}\right) \\
& + \dot{\mathcal{M}}_{AB}(\delta_{3e}\delta_{Af}\delta_{Bg} + \delta_{3f}\delta_{Ae}\delta_{Bg} + \delta_{3g}\delta_{Ae}\delta_{Bf}) \\
& \times\left(H^{\dagger}_{ei}H^{\dagger}_{fj}H^{\dagger}_{gkl} + H^{\dagger}_{ei}H^{\dagger}_{fk}H^{\dagger}_{gjl} + H^{\dagger}_{ei}H^{\dagger}_{fl}H^{\dagger}_{gjk} + H^{\dagger}_{ej}H^{\dagger}_{fk}H^{\dagger}_{gil} + H^{\dagger}_{ej}H^{\dagger}_{fl}H^{\dagger}_{gik} + H^{\dagger}_{ek}H^{\dagger}_{fl}H^{\dagger}_{gij}\right) \\
& + \mathcal{M}_{ABC}\left(\mathcal{F}_{iA}\mathcal{F}_{jB}H^{\dagger}_{Ckl} + \mathcal{F}_{iA}\mathcal{F}_{kB}H^{\dagger}_{Cjl} + \mathcal{F}_{iA}\mathcal{F}_{lB}H^{\dagger}_{Cjk} + \mathcal{F}_{jA}\mathcal{F}_{kB}H^{\dagger}_{Cil} + \mathcal{F}_{jA}\mathcal{F}_{lB}H^{\dagger}_{Cik} + \mathcal{F}_{lA}\mathcal{F}_{lB}H^{\dagger}_{Cij}\right) \\
& + \dot{\mathcal{M}}_{ABC}\left(p_i\mathcal{F}_{jA}\mathcal{F}_{kB}\mathcal{F}_{lC} + p_j\mathcal{F}_{iA}\mathcal{F}_{kB}\mathcal{F}_{lC} + p_k\mathcal{F}_{iA}\mathcal{F}_{jB}\mathcal{F}_{lC} + p_l\mathcal{F}_{iA}\mathcal{F}_{jB}\mathcal{F}_{kC}\right) \\
& + \ddot{\mathcal{M}}_{AB}\left(p_ip_j\mathcal{F}_{kA}\mathcal{F}_{lB} + p_ip_k\mathcal{F}_{jA}\mathcal{F}_{lB} + p_ip_l\mathcal{F}_{jA}\mathcal{F}_{kB} + p_jp_k\mathcal{F}_{iA}\mathcal{F}_{lB} + p_jp_l\mathcal{F}_{iA}\mathcal{F}_{kB} + p_kp_l\mathcal{F}_{iA}\mathcal{F}_{jB}\right) \\
& + \mathcal{M}_{ABCD}\,\mathcal{F}_{iA}\mathcal{F}_{jB}\mathcal{F}_{kC}\mathcal{F}_{lD}.
\end{aligned} \tag{F8}$$

It is straightforward, but quite laborious, to also write the mapping of fifth-order derivatives (or even higher derivatives) in terms of the co-variant basis. We leave this exercise to the interested reader.