Geophysical Prospecting
EAGE EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS
Geophysical Prospecting, 2021, **69**, 3–27 doi: 10.1111/1365-2478.13052

Eigenrays in 3D heterogeneous anisotropic media, Part I: Kinematics

Zvi Koren* and Igor Ravve*
Emerson, Houston, USA

Received March 2019, revision accepted November 2020

ABSTRACT

We present a new ray bending approach, referred to as the Eigenray method, for solving two-point boundary-value kinematic and dynamic ray tracing problems in 3D smooth heterogeneous general anisotropic elastic media. The proposed Eigenray method is aimed to provide reliable stationary ray path solutions and their dynamic characteristics, in cases where conventional initial-value ray shooting methods, followed by numerical convergence techniques, become challenging. The kinematic ray bending solution corresponds to the vanishing first traveltime variation, leading to a stationary path between two fixed endpoints (Fermat's principle), and is governed by the nonlinear second-order Euler–Lagrange equation. The solution is based on a finite-element approach, applying the weak formulation that reduces the Euler–Lagrange second-order ordinary differential equation to the first-order weighted-residual nonlinear algebraic equation set. For the kinematic finite-element problem, the degrees of freedom are discretized nodal locations and directions along the ray trajectory, where the values between the nodes are accurately and naturally defined with the Hermite polynomial interpolation. The target function to be minimized includes two essential penalty (constraint) terms, related to the distribution of the nodes along the path and to the normalization of the ray direction. We distinguish between two target functions triggered by the two possible types of stationary rays: a minimum traveltime and a saddle-point solution (due to caustics). The minimization process involves the computation of the global (all-node) traveltime gradient vector and the traveltime Hessian matrix. The traveltime Hessian is used for the minimization process, analysing the type of the stationary ray, and for computing the geometric spreading of the entire resolved stationary ray path. The latter, however, is not a replacement for the dynamic ray tracing solution, since it does not deliver the geometric spreading for intermediate points along the ray, nor the analysis of caustics. Finally, we demonstrate the efficiency and accuracy of the proposed method along three canonical examples.

Key words: Anisotropy, Rays, Computing aspects.

INTRODUCTION

Two-point ray tracing in general 3D heterogeneous anisotropic media is one of the cornerstones for simulating the propagation of the high-frequency components of seismic body waves between sources and receivers. It has been mainly performed using the ray shooting method followed by numerical convergence schemes to reach the destination points. In this method, a fan of rays is first traced from a given starting point to the acquisition surface, where groups of rays arriving near each target location (e.g. a receiver) with similar take-off angles (slowness vectors) are used for the convergence process. By covering a wide range of take-off

*Email: zvi.koren@emerson.com, igor.ravve@emerson.com

angles, multi-pathing stationary solutions can be found. For example, Bulant (1996, 2002) suggests an original ray shooting algorithm that makes it possible to find all two-point trajectories for general 3D layer-based smooth inhomogeneous isotropic media. However, in complex geological areas, characterized by heterogeneity and intrinsic anisotropic characteristics of the elastic properties of the rocks, the convergence to a given location can be highly sensitive to (infinitely) small changes in the take-off angles, resulting in shadow zones which the numerically traced rays can barely penetrate. The ray bending methods can be considered a complementary optimization approach to the ray shooting, where the proposed Eigenray method can be viewed as an extension of currently available ray bending solutions, in particular for general anisotropic elastic media. It primarily attempts to fill in the above-mentioned shadow zones using the following workflow: (a) constructing initial guess (non-stationary) trajectories by interpolating/extrapolating from nearby traced rays (e.g. provided by the ray shooting method), (b) updating the initial-guess trajectories until they satisfy Fermat's principle of stationary time: An iterative optimization procedure for finding the (nearest) stationary ray path between the two fixed endpoints using the finite-element approach, and (c) weighting (quantifying) the plausibility of the solution with a proposed complexity criterion based on the computed amplitude loss (normalized geometric spreading, discussed in both parts of this study). As mentioned, the Eigenray method is particularly attractive for areas that involve considerable isotropic/anisotropic velocity variations (e.g. between sediments and salt/basalt/carbonate rocks) or local velocity anomalies (e.g. gas clouds) smoothed along the transition zones. The method can also be extended for 'blocky' models with sharp velocity discontinuities across the surface interfaces, and we discuss this option as well. The case of simulating 'head waves' is an extreme example of Eigenrays providing controlled plausible solutions where conventional ray tracing methods become extremely challenging. Thomson (1989) suggested the use of a ray bending (correcting) solution to study head (grazing) waves, where the primary (incidence) and secondary (reflected/scattered) waves travel (graze) along a given reflector/refractor. We note that the term 'head waves' is used here to describe general phenomena of rays, mainly travelling laterally along/below transition zones or across/in-between local velocity anomalies.

Depending on the subsurface model representation, the geometry of the acquisition system and the problem to be solved, such as seismic forward modelling, migration and inversion, several different ray-based approaches for simulating the propagation of high-frequency body waves in inhomogeneous isotropic and anisotropic elastic (and also viscoelastic) media have been studied and implemented; each has its own advantages and disadvantages. These characteristics have been documented, for example, by Leidenfrost *et al.* (1999), where the authors study six different methods, in particular, finite-difference (FD) eikonal solvers, the graph method, wavefront construction (WFC) and a combined FD and Runge–Kutta method, for calculating seismic traveltimes from a point source to a regular subsurface grid. Additionally, Thurber and Kissling (2000) suggest a method for classifying different strategies for computing ray paths and travel times combining ray shooting, bending, perturbation and grid-based approaches. Other methods and strategies, or combinations of different methods, are also available and widely used.

WFC is an attractive method for simulating ray propagation in complex areas while attempting to avoid artificial 'shadow zones'. Vinje *et al.* (1993) published a pioneering study on the WFC method, along with later works by Lambaré *et al.* (1996), Lucio *et al.* (1996), Ettrich and Gajewski (1996), Gibson (1999), Gjøystdal *et al.* (2002) and Lai *et al.* (2009). In this method, a fan of dense rays is simultaneously propagated in time, where at each time step a new wavefront is constructed. During the propagation, the wavefront is normally expanded and hence, new ray segments, normal to the current wavefront, are added, to ensure reliable representation (sampling) of the wavefront. However, while crossing complex velocity regions, the wavefront begins to split into several branches (triplications) resulting in deformed (non-topological) shapes (e.g. cusps). The process of adding ray segments (normal to the non-topological wavefronts) becomes very challenging, and the accuracy of the constructed wavefront decreases. In extreme cases, a huge number of additional rays is required to fulfil the required accuracy, making this method expensive. Note that the WFC method does not naturally deliver the exact stationary ray paths, which is the main goal of the present work. If the ray paths are not explicitly needed, the method can be efficiently used to compute dynamic parameters as well. The WFC can be used along with the ray shooting and ray bending techniques, for example, to compute initial conditions for ray shooting and boundary conditions for ray bending.

Paraxial ray tracing methods have been intensively used to solve two-point ray tracing problems in smooth heterogeneous velocity media (e.g. Beydoun and Keho, 1987; Virieux *et al.*, 1988; Farra *et al.*, 1989; Gibson *et al.*, 1991; Farra, 1993; Strahilevitz *et al.*, 1998) and for the computation of dynamic properties (e.g. Popov and Pšenčík, 1978). In this tech-

nique, paraxial (nearby) rays are approximated (predicted) from given reference (central) rays using the first-order perturbation theory. These methods can also be used to interpolate traveltimes in the vicinity of the central rays (e.g. Bulant and Klimeš, 1999).

Gaussian beam summation methods (Popov, 1982; Červený, Popov and Pšenčík, 1982) or Maslov's methods (Chapman and Drummond, 1982; Thomson and Chapman, 1985; Huang *et al*., 1998) are alternative methods for the two-point ray tracing, with the advantage of overcoming the ray theory singularity problems related to the vanishing ray Jacobian while crossing caustics.

The approach suggested in this study belongs to the class of the ray bending optimization methods which have been extensively studied in the past, mainly for isotropic media or anisotropic media with high level of symmetry (e.g. transverse isotropy). An early ray tracing approach based on Fermat's principle has been suggested by Wesson (1971). Julian and Gubbins (1977) derived a boundary-value formulation of the ray tracing equations that can be solved iteratively (a ray bending approach), claiming that in some cases this method is more efficient than the ray shooting method. The proposed Eigenray method inherits several features of this early study; however, the method suggested by Julian and Gubbins (1977) is based on the FD formulation, applying globally the flow parameter for the whole interval between the source and the receiver, $0 \leq \xi \leq 1$, while the proposed Eigenray method is based on the finite-element formulation, applying a local flow variable $-1 \leq \xi \leq +1$ inside each finite element (which is generally more accurate). Furthermore, the proposed finite-element approach makes it possible to naturally insert different constraints (e.g. continuity/discontinuity) across the discretized nodes of the ray path, depending on the type of the geological layered model and the required type of solution; the implementation of these kinds of operations is nontrivial when using the FD approach. In order to solve the nonlinear algebraic set, these authors suggest the Newton method which requires computation of the traveltime Hessian matrix that can later be used to compute the geometric spreading. This is also the case with the proposed Eigenray method.

Smith *et al*. (1979) applied the ray bending technique for a full 3D velocity inversion. Pereyra *et al*. (1980) extended the ray bending technique to allow reflection and transmission through the interfaces. In later studies, Pereyra (1988, 1992) applied a combination of a fast-shooting algorithm and a multipoint boundary-value ray bending approach to obtain several source–receiver arrivals (multi-pathing), and to compute geometric spreading. The method was then extended for both modelling and inversion problems in complex 3D geological models, applying irregularly shaped blocks (Pereyra, 1996; Xu *et al*., 2014), with smoothly and slowly varying medium properties inside the blocks and sharp discontinuities at the interfaces. Thomson and Gubbins (1982) used equally spaced nodes along the horizontal x axis and applied a cubic spline interpolation to solve the boundary-value ray tracing problem with the ray bending method. They also computed geometric spreading in areas that involve velocity anomalies. Thomson (1983) solved the inverse problem for the velocity model parameters by applying the ray bending method to relate changes in the geometric spreading to the lateral velocity variations. Um and Thurber (1987) presented the ray trajectory as a set of points with a linear interpolation between them, where they adjusted the locations of these points to fit the kinematic ray tracing equations. Westwood and Vidmar (1987) applied the ray bending method to simulate the signals interacting with a layered ocean bottom. Waltham (1988) studied models consisting of constant velocity layers separated by curved interfaces and computed ray paths, whose traveltimes are stationary with respect to (wrt) changes in the ray/interface intersection points. Moser (1991) used this method to compute the traveltime between the source point and all points of a given network. Moser *et al*. (1992) improved the conventional ray bending approach by applying (a) gradient search methods and (b) interpolation by beta-splines between the nodes. Farra (1992) applied the Hamiltonian formulation with the propagator matrix to the ray bending approach. For isotropic media, Snieder and Spencer (1993) presented a unified approach to solve three (traditionally different) ray path update problems, namely, ray bending, ray velocity perturbation and paraxial ray methods. The ray bending method assumes a given velocity model and two fixed endpoints. Starting with an approximated non-stationary path between the endpoints, the aim is to find the nearest stationary ray. The ray perturbation and paraxial ray methods assume already a given stationary ray path, and the aim is to estimate a new stationary ray due to changes in the medium properties or the endpoint locations, respectively. Shashidhar and Anand (1995) solved the problem of 3D Eigenray tracing in an ocean channel. Grechka and McMechan (1996) developed a 3D two-point ray tracing technique based on Fermat's principle, suggesting a method that takes advantage of the global Chebyshev approximation of both the model and the curved rays, and makes it possible to find minima, maxima and saddle points of traveltime. Cores *et al*. (2000) assumed a piecewise-linear ray path and presented the problem of tracing rays under Fermat's principle in 2D and 3D heterogeneous isotropic media. They applied

biharmonic splines to model the reflector geometry and the velocity function, where they used the global spectral gradient technique for the optimization.

Ecoublet *et al.* (2002) suggested a 2D ray bending tomography (as an alternative to an initial-value ray shooting tomography) where the traveltime between the endpoints satisfies Fermat's principle. Bona and Slawinski (2003) demonstrated that Fermat's principle of stationary traveltime holds for general heterogeneous anisotropic media. Zhao *et al.* (2004) applied an irregular network 'shortest path' method for high-performance seismic ray tracing. This approach was later extended by Zhou and Greenhalgh (2005, 2006) for anisotropic media, in particular, for transverse isotropy (TI). Rawlinson *et al.* (2008) considered a variety of schemes for tracking the kinematics of seismic waves in heterogeneous 3D structures, including a ray shooting method, a ray bending method and a combined approach, with a linear interpolation between the nodes. Kumar *et al.* (2004) and Casasanta *et al.* (2008) developed a two-point ray bending algorithm for TI with vertical axis of symmetry (VTI) media for compressional waves, applying the group velocity approximation suggested by Byun *et al.* (1989). Wong (2010) extended the method for layered TI with tilted axis of symmetry (TTI) media, where the degrees of freedom (DoF) were intersections of the ray with the layer interfaces. Bona *et al.* (2009) developed a strategy for two-point ray tracing, using a stochastic simulated annealing global search method. Sripanich and Fomel (2014) presented an efficient algorithm for two-point ray tracing in layered media by means of the ray bending method, where the ray paths are discretized at the intersection of the rays with the structure's interfaces, applying the global traveltime Hessian and the Newton method to find a stationary ray path. Cao *et al.* (2017) suggested a fast-marching method to compute the traveltime along an expanding wavefront using Fermat's principle in transversely isotropic media with vertical and tilted axes of symmetry, where the ray (group) velocity was approximated from the moveout equation. Wu *et al.* (2019) applied the shortest-path ray tracing adhering to Fermat's principle, in order to suppress the noise and improve the quality of pre-stack seismic data (nonlinear optimal stacking). Hovem and Dong (2019) applied the Eigenray method to compute the trajectories, reflection points and incidence angles for a large number of rays in seawater with the same source and receiver, and multiple reflections from the sea floor and from the water surface. We note that in special seismic acquisition surveys, such as vertical seismic profiles (VSP) or well-to-well surveys, the ray bending approach can be particularly attractive.

One of the core parts of the ray bending procedure is computing the ray velocity magnitude from its direction, which, in turn, requires first to establish the slowness vector components; for compressional waves, the solution is unique. We follow the approach suggested by Musgrave (1954, 1970), Fedorov (1968), and Grechka (2017), where the slowness-related gradient of the vanishing Hamiltonian is parallel to the ray direction. A similar approach with a set of polynomial equations has been suggested by Vavryčuk (2006). Zhang and Zhou (2018) applied this method for TI media and extended it to separate compressional qP waves from qSV shear waves and to eliminate complex-valued roots, using corresponding relationships for the phase velocities of these wave modes. As a result, unique solutions were obtained for qP and SH waves, and one or three solutions for qSV waves in TI media. Conditions for qSV triplications have been formulated by Dellinger (1991) and further studied by Thomsen and Dellinger (2003), Schoenberg and Daley (2003), Vavryčuk (2003) and Roganov and Stovas (2010). Song and Every (2000), Stovas (2016), Xu and Stovas (2018) and Xu *et al.* (2021) studied triplications in orthorhombic media. Xu *et al.* (2020) explored triplications of converted waves in TTI media. Stovas *et al.* (2020) studied phase and ray velocity surfaces and triplications in elastic and acoustic models of different symmetry classes.

Computing the slowness vector from the ray direction vector has also been studied for general viscoelastic anisotropic media (Vavryčuk, 2007, 2008a), and the results were used for ray tracing in such media (Vavryčuk, 2008b, 2012).

Koren and Ravve (2018a) demonstrated the power of applying a ray bending solution (referred to as the Eigenray method) by using a nonlinear spectral element method (Lagrange interpolation-based elements) to efficiently find accurate stationary ray paths in complex geological areas, characterized by smooth heterogeneous isotropic media. In a following abstract (Koren and Ravve, 2018b), the theory was extended to general anisotropic media, using Hermite finite elements, which makes it possible to naturally impose continuity conditions at the ray nodes for both locations and directions of the ray velocity. This abstract and the following Eigenray studies are heavily based on our paper on the computation of spatial and directional, first and second, derivatives of the ray velocity (governing the corresponding derivatives of the traveltime) in 3D smooth heterogeneous general anisotropic elastic media (Ravve and Koren, 2019). Two later abstracts on the kinematic Eigenray (Koren and Ravve, 2020d) and dynamic Eigenray (Ravve and Koren, 2020e) briefly summarize the basic theoretical and im-

plementation concepts of the Eigenray method. Recently, we have published the detailed theoretical and implementation aspects of the Eigenray method as seven preprint papers in the arXiv website library of the Cornell University; three parts on the kinematics (Koren and Ravve, 2020a, 2020b, 2020c) and four parts on the dynamics (Ravve and Koren, 2020a, 2020b, 2020c, 2020d). The primary objective of this part of the paper is to provide the main results of the kinematic Eigenray method.

We propose an original arclength-related Lagrangian, and we derive the corresponding kinematic Euler–Lagrange nonlinear, second-order ordinary differential equations. We also provide the corresponding Hamiltonian, related to the proposed Lagrangian through the Legendre transformation. We then apply a finite-element approach to establish the stationary ray path, nearest to a specified initial-guess trajectory, for 3D smooth heterogeneous general anisotropic media. Finally, using the already computed global (all-node) traveltime Hessian of the stationary ray, we solve a particular (limited) 'dynamic' problem which does not require explicitly performing the dynamic ray tracing system. We propose an efficient method to condense the global traveltime Hessian into the source–receiver traveltime Hessian, which makes it possible to compute the geometric spreading of the entire ray path between the source and receiver. Computation of the geometrical spreading from the source–receiver traveltime Hessian is a well-known technique and we list the relevant published studies on this topic within the corresponding dedicated section. Note that in the context of this part of the study which deals with the kinematic ray characteristics (e.g. ray paths and traveltime), the geometric spreading is used to construct a ray complexity criterion which is proposed to quantify the reliability/plausibility of the resolved ray trajectories.

Finally, one cannot consider writing a theoretical study on seismic ray theory without explicitly mentioning the many excellent scientific books on this theme, among them, Sommerfeld (1964), Červený (2000), Bleistein *et al*. (2001), Chapman (2004), Schleicher *et al*. (2007), Tsvankin and Grechka (2011), Slawinski (2015) and Krebes (2019).

THE KINEMATIC EIGENRAY METHOD

The kinematic Eigenray tracing is based on the application of Fermat's principle to find the ray paths and their corresponding traveltimes between a given source–receiver pair in a known geological model parameterized with anisotropic elastic properties. According to Fermat's principle, the ray path between two fixed endpoints is the one that leads to a stationary time (normally, the least time, but may also be a saddle-point time, due to caustics). Considering anisotropic elastic media and fixed endpoint pairs (source–receiver), we obtain the stationary ray paths using the ray bending approach. Starting with an initial (non-stationary) trajectory discretized with a set of irregular sampled nodes, we solve the two-point kinematic ray tracing (KRT) problem using a finite-element approach, successively (iteratively) refining the locations and directions of the trajectory at the nodes. If a number of different stationary rays co-exist between the given endpoints (multi-pathing), a different initial trajectory is applied to each solution.

Overall, as a strategy for performing two-point ray tracing for conventional seismic acquisition surveys, we recommend starting with the ray shooting method, followed by numerical convergence techniques. In cases where the numerical convergence is difficult (the solution becomes very sensitive to fine changes in the take-off angles), the proposed Eigenray method can be used as an additional (complementary) attempt to obtain plausible solutions. In these cases, a (non-stationary) initial-guess solution between the endpoints is first predicted by interpolation/extrapolation from nearby rays.

We apply the Euler–Lagrange formula over an integrand of the traveltime functional (Lagrangian), valid for smooth heterogeneous general anisotropic elastic media. We propose a specific Lagrangian with a clear physical interpretation, which we find most convenient and efficient for our finite-element solution, in particular, in general anisotropic media.

For a given trajectory between two fixed endpoints (a stationary ray path or a non-stationary approximation path in its vicinity), the Lagrangian $L = L[\mathbf{x}(s),\ \dot{\mathbf{x}}(s)]$ used in this work is a function of the arclength-dependent ray path coordinates, $\mathbf{x}(s)$, and the arclength derivative of the location vector, $\dot{\mathbf{x}}(s) = \mathrm{d}\mathbf{x}/\mathrm{d}s$, along the trajectory. Parameter s is the arclength of the ray path. The arclength derivative of the location vector, $\dot{\mathbf{x}}(s)$, is noted by $\mathbf{r}(s) \equiv \dot{\mathbf{x}}(s)$, and it represents a vector tangent to the trajectory at any point along the ray, normalized to the unit length, $\mathbf{r} \cdot \mathbf{r} = 1$; hence it is the ray direction vector, or equivalently, the ray velocity direction vector. The Euler–Lagrange equation is then applied to the proposed Lagrangian, yielding a nonlinear, second-order ordinary differential equations for the ray locations and directions, in terms of the ray velocity, its gradients and Hessians, which is then solved by the finite-element approach.

The local gradients and Hessians of the traveltime require the corresponding derivatives of the ray velocity along the local segments of the path (finite elements). Due to anisotropy, we deal with two different types of ray velocity gradient vectors – spatial and directional – and three types of ray velocity

Hessian matrices: spatial, directional and mixed. The method for computing the ray velocity derivatives is a core component of this study. It has been recently published by Ravve and Koren (2019) for general anisotropic (triclinic) media, where the higher anisotropic symmetries were considered particular cases. The method establishes the spatial and directional gradients, $\nabla_{\mathbf{x}} v_{\text{ray}}$ and $\nabla_{\mathbf{r}} v_{\text{ray}}$, and the spatial, directional and mixed Hessians, $\nabla_{\mathbf{x}}\nabla_{\mathbf{x}} v_{\text{ray}}$, $\nabla_{\mathbf{r}}\nabla_{\mathbf{r}} v_{\text{ray}}$ and $\nabla_{\mathbf{x}}\nabla_{\mathbf{r}} v_{\text{ray}} = (\nabla_{\mathbf{r}}\nabla_{\mathbf{x}} v_{\text{ray}})^{\mathrm{T}}$. Although the theory described in this study includes all types of waves, we only focus on the implementation for compressional waves.

Comment: The method presented in this study can be applied to all types of wave modes in anisotropic elastic media. Note that the main challenge of shear waves in anisotropic media is the non-uniqueness of the ray (group) velocity surface, which, in turn, is related to the non-uniqueness of the slowness vector $\mathbf{p}$ given the ray velocity direction $\mathbf{r}$. Consider, for simplicity, a homogeneous triclinic medium with high degree of anisotropy. Given the source and receiver locations, $\mathbf{x}_S$ and $\mathbf{x}_R$, the shape of the ray path is a straight line, with the length of the path, $s = \sqrt{(\mathbf{x}_R - \mathbf{x}_S)\cdot(\mathbf{x}_R - \mathbf{x}_S)}$. To compute the traveltime, we need first to establish the slowness components, $t = (\mathbf{x}_R - \mathbf{x}_S)\cdot\mathbf{p} = s\,\mathbf{r}\cdot\mathbf{p} = s/v_{\text{ray}}$. However, there are up to 18 different shear-wave slowness vector solutions (Grechka, 2017) leading to corresponding different ray velocity magnitudes and traveltimes along the same straight trajectory. This, in turn, yields multi-valued Lagrangians, their gradients and Hessians. One needs to choose one of the multiple shear branches, and in a general case of the KRT for a shear wave, keeping the required chosen branch at each step. Note that simulating the amplitudes of shear waves with dynamic ray tracing is easier, since it is assumed that the kinematic problem has already been solved, and thus, at each node the proper shear branch has already been chosen. For compressional waves, the ray (group) velocity surface is convex, and the solution for the slowness vector components is unique.

THE ARCLENGTH-RELATED LAGRANGIAN

Consider a 3D smooth heterogeneous general anisotropic medium and a given initial-guess ray trajectory (non-stationary approximated ray path) between two fixed endpoints, S and R. Let $\mathbf{x}(s) = [\,x_1\ x_2\ x_3\,]$ be a point along the approximated ray path, and $\dot{\mathbf{x}}(s) \equiv \mathrm{d}\mathbf{x}(s)/\mathrm{d}s \equiv \mathbf{r}(s) = [\,r_1\ r_2\ r_3\,]$ the direction vector at this point, normalized to the unit length, $\mathbf{r}\cdot\mathbf{r} = 1$, where $\mathrm{d}s$ is an infinitesimal elementary arclength along the path. Note that $\mathbf{r}$ is also the direction of the ray velocity vector. Fermat's principle for a stationary traveltime path can be stated with the use of the Lagrangian $L(\mathbf{x}, \mathbf{r})$ that depends on the ray path coordinates $\mathbf{x}(s)$ and direction $\mathbf{r}(s)$ along the path,

$$t = \int_S^R L\,(\mathbf{x}, \mathbf{r})\,\mathrm{d}s \rightarrow \text{stationary}, \delta t\,(\mathbf{x}) = \int_S^R \delta L\,(\mathbf{x}, \mathbf{r})\mathrm{d}s = 0. \quad (1)$$

In this study, we propose the following arclength-related Lagrangian:

$$L\,(\mathbf{x}, \mathbf{r}) = \frac{\mathrm{d}t}{\mathrm{d}s} = \frac{\sqrt{\mathbf{r}\cdot\mathbf{r}}}{v_{\text{ray}}\,(\mathbf{x}, \mathbf{r})} = \frac{\sqrt{\mathbf{r}\cdot\mathbf{r}}}{v_{\text{ray}}\left[\tilde{\mathbf{C}}(\mathbf{x}), \mathbf{r}\right]}. \quad (2)$$

The proposed Lagrangian is a first-degree homogeneous function with respect to (wrt) the ray direction $\mathbf{r}$. The value of the square root $\sqrt{\mathbf{r}\cdot\mathbf{r}}$ in the numerator of the Lagrangian is 1, but we do not replace it by the constant value because it affects the partial derivatives needed for the Euler–Lagrange formulation. The ray velocity magnitude in the denominator, $v_{\text{ray}}[\tilde{\mathbf{C}}(\mathbf{x}), \mathbf{r}]$, depends implicitly on the location components $\mathbf{x}$, where $\tilde{\mathbf{C}}(\mathbf{x})$ is the medium density-normalized stiffness tensor (the tilde indicates the fourth-order tensor), and explicitly on the ray direction components (the latter is in particular important in anisotropic media). The stiffness tensor at each point is known and fixed: in this study, we do not update the model.

Note that the ray velocity magnitude is a zero-degree homogeneous function wrt any tangent vector $k\,\mathbf{r}$, which means that it depends only on the normalized direction vector $\mathbf{r}$ but not on its length k.

A comprehensive analysis of the proposed Lagrangian for anisotropic elastic media (equation (2)) is given by Koren and Ravve (2020a, 2020b). In particular, the latter cited paper provides analytical and numerical validations of the Lagrangian for different anisotropic symmetries and comparisons with alternative Lagrangians (e.g. Červený, 2002a, 2002b).

Notations for the derivatives wrt the flow parameter

A parameter with an upper dot and with no subscript means a derivative of the parameter wrt the arclength s (in the dynamic analysis – wrt the arclength of the central ray). A parameter with an upper dot and with a subscript means a derivative wrt the flow parameter indicated by that subscript. In particular, $\dot{\mathbf{x}} \equiv \dot{\mathbf{x}}_s$, $\dot{\mathbf{x}}_\tau$ and $\dot{\mathbf{x}}_\zeta$ represent the derivatives of the ray path location wrt the arclength s, current time τ and a generic flow parameter ζ, respectively. Additionally, in the finite-element implementation, we use the symbol 'prime' (instead of the upper dot) for derivatives wrt the internal flow parameter $-1 \le \xi \le +1$ within any individual element (be-

tween the nodes). In particular, $\mathbf{x}'$ and $\mathbf{r}'$ are, respectively, the derivatives of the location and direction wrt the internal parameter ξ.

EULER–LAGRANGE EQUATION

In this section, we obtain the nonlinear second-order Euler–Lagrange kinematic ray equation which is then solved with the finite-element approach.

For the spatial and directional gradients of the Lagrangian $L(\mathbf{x}, \mathbf{r})$, we apply the shorthand notations,

$$L_{\mathbf{x}} = \nabla_{\mathbf{x}} L = \frac{\partial L}{\partial \mathbf{x}} \quad \text{and} \quad L_{\mathbf{r}} = \nabla_{\mathbf{r}} L = \frac{\partial L}{\partial \mathbf{r}}. \tag{3}$$

With these notations, the Euler–Lagrange equation (manifesting the stationarity condition) reads (e.g. Gelfand and Fomin, 2000)

$$\frac{d}{\mathrm{d}s} L_{\mathbf{r}} = L_{\mathbf{x}}. \tag{4}$$

The spatial and directional partial derivatives (gradients) of the Lagrangian $L(\mathbf{x}, \mathbf{r})$ follow from equation (2),

$$L_{\mathbf{x}} = -\frac{\nabla_{\mathbf{x}} v_{\mathrm{ray}}}{v_{\mathrm{ray}}^2}\sqrt{\mathbf{r}\cdot\mathbf{r}} \text{ and } L_{\mathbf{r}} = \frac{\mathbf{r}}{v_{\mathrm{ray}}\sqrt{\mathbf{r}\cdot\mathbf{r}}} - \frac{\nabla_{\mathbf{r}} v_{\mathrm{ray}}}{v_{\mathrm{ray}}^2}\sqrt{\mathbf{r}\cdot\mathbf{r}}. \tag{5}$$

Introduction of equation (5) into equation (4) leads to the explicit form of the nonlinear, second-order, vector-form, Euler–Lagrange ordinary differential equations (ODE),

$$\frac{d}{\mathrm{d}s}\left(\frac{\mathbf{r}}{v_{\mathrm{ray}}\sqrt{\mathbf{r}\cdot\mathbf{r}}} - \frac{\nabla_{\mathbf{r}} v_{\mathrm{ray}}}{v_{\mathrm{ray}}^2}\sqrt{\mathbf{r}\cdot\mathbf{r}}\right) = -\frac{\nabla_{\mathbf{x}} v_{\mathrm{ray}}}{v_{\mathrm{ray}}^2}\sqrt{\mathbf{r}\cdot\mathbf{r}}. \tag{6}$$

We note that whenever there is no need to further differentiate with respect to the ray direction $\mathbf{r}$, equation (5) can be simplified to

$$L_{\mathbf{x}} = -\frac{\nabla_{\mathbf{x}} v_{\mathrm{ray}}}{v_{\mathrm{ray}}^2} \text{ and } L_{\mathbf{r}} = \frac{\mathbf{r}}{v_{\mathrm{ray}}} - \frac{\nabla_{\mathbf{r}} v_{\mathrm{ray}}}{v_{\mathrm{ray}}^2} \tag{7}$$

and the actual nonlinear, second-order, Euler–Lagrange ODE to be solved is given by

$$\frac{d}{\mathrm{d}s}\underbrace{\left(\frac{\mathbf{r}}{v_{\mathrm{ray}}} - \frac{\nabla_{\mathbf{r}} v_{\mathrm{ray}}}{v_{\mathrm{ray}}^2}\right)}_{L_{\mathbf{r}}=\mathbf{p}} = -\frac{\nabla_{\mathbf{x}} v_{\mathrm{ray}}}{v_{\mathrm{ray}}^2}. \tag{8}$$

The expression in the brackets constitutes the momentum equation, $L_{\mathbf{r}} = \mathbf{p}$ (where $\mathbf{p}$ is the slowness vector consisting of the tangent (to the ray) and normal components, as shown in Fig. 1). The directional gradient vector, $\nabla_{\mathbf{r}} v_{\mathrm{ray}}$, belongs to the plane that includes both vectors, $\mathbf{p}$ and $\mathbf{v}_{\mathrm{ray}}$, and is also normal to $\mathbf{v}_{\mathrm{ray}}$. It can be written as

$$\nabla_{\mathbf{r}} v_{\mathrm{ray}} = -\mathbf{v}_{\mathrm{ray}} \times \mathbf{p} \times \mathbf{v}_{\mathrm{ray}} = -v_{\mathrm{ray}}^2\, \mathbf{r} \times \mathbf{p} \times \mathbf{r}. \tag{9}$$

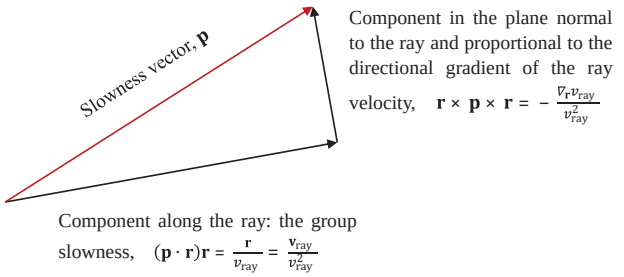

Figure 1 Slowness vector decomposition into components along the ray and in the normal plane.

Remark: Normally, for three arbitrary vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ a triple cross-product is non-associative $(\mathbf{a}\times\mathbf{b})\times\mathbf{c} \neq \mathbf{a}\times(\mathbf{b}\times\mathbf{c})$. However, in a particular case, $(\mathbf{a}\times\mathbf{b})\times\mathbf{a} = \mathbf{a}\times(\mathbf{b}\times\mathbf{a})$, it becomes associative. We therefore suggest omitting the brackets for the sake of symmetry, and write it just as $\mathbf{a}\times\mathbf{b}\times\mathbf{a}$. In our case, $\mathbf{r}\times\mathbf{p}\times\mathbf{r}$ yields a projection of vector $\mathbf{p}$ on a plane normal to the ray $\mathbf{r}$. Detailed explanations and derivation of equation (9) can be found in Appendix A of Koren and Ravve (2020a). Introducing equation (9) into the second equation of set 7 and taking into account that $v_{\mathrm{ray}}^{-1} = \mathbf{p}\cdot\mathbf{r}$, we see that the first term on the right-hand side is the projection of the slowness vector on the ray direction, while the second term is its projection on the plane normal to the ray. The sum of these terms results in the full slowness vector, i.e. $L_{\mathbf{r}} = (\mathbf{p}\cdot\mathbf{r})\mathbf{r} + \mathbf{r}\times\mathbf{p}\times\mathbf{r} = \mathbf{p}$.

We apply the weak formulation to equation (8) to be solved with the finite-element method (e.g. Segerlind, 1984; Hughes, 2000; Reddy, 2004; Zienkiewicz *et al.*, 2013; Bathe, 2014). The weak formulation effectively eliminates the second derivative of the position vector (or the first derivative of the ray direction), $\ddot{\mathbf{x}} = \dot{\mathbf{r}}$.

THE ARCLENGTH-RELATED HAMILTONIAN

In this section, we provide the arclength-related Hamiltonian, $H(s)$, that matches the proposed Lagrangian $L(s)$ through the Legendre transformation.

We consider the Christoffel equation, valid for general anisotropic media, as a *reference* (vanishing, unitless) ray tracing Hamiltonian:

$$H^{\bar{\tau}}(\mathbf{x}, \mathbf{p}) = \det\left[\mathbf{\Gamma}(\mathbf{x}, \mathbf{p}) - \mathbf{I}\right] = 0\,,\ \mathbf{\Gamma}(\mathbf{x}, \mathbf{p}) = \mathbf{p}\cdot\tilde{\mathbf{C}}(\mathbf{x})\cdot\mathbf{p}, \tag{10}$$

where $\mathbf{\Gamma}(\mathbf{x}, \mathbf{p})$ is the Christoffel matrix, $\mathbf{I}$ is the 3×3 identity matrix, and $\tilde{\mathbf{C}}(\mathbf{x})$ is the density-normalized fourth-order stiffness (elastic) tensor. The bar above the superscript index,

τ, indicates that the flow parameter of the Hamiltonian is a 'scaled time' $\bar{\tau}$ (rather than the actual time τ), and it has the units of time $[T]$. In Appendix C of Koren and Ravve (2020a) we relate it to the actual traveltime τ, using a unitless scaler:

$$\alpha_{\text{sc}}(\tau) = \frac{\mathrm{d}\bar{\tau}}{\mathrm{d}\tau} = \frac{\mathrm{d}\bar{\tau}/\mathrm{d}s}{\mathrm{d}\tau/\mathrm{d}s} = \frac{1}{\mathbf{p}\cdot H_{\mathbf{p}}^{\bar{\tau}}},$$

where

$$H_{\mathbf{p}}^{\bar{\tau}} \equiv \frac{\partial H^{\bar{\tau}}}{\partial \mathbf{p}}. \tag{11}$$

We then introduce the following arclength-related Hamiltonian, connected to the reference Hamiltonian:

$$H(\mathbf{x},\mathbf{p}) \equiv H^{s}(\mathbf{x},\mathbf{p}) = \frac{H^{\bar{\tau}}(\mathbf{x},\mathbf{p})}{\sqrt{H_{\mathbf{p}}^{\bar{\tau}}\cdot H_{\mathbf{p}}^{\bar{\tau}}}}. \tag{12}$$

With this Hamiltonian, the ray equation set for the kinematic ray tracing can be written in a simple form

$$\mathbf{r} \equiv \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s} = H_{\mathbf{p}}\,,\ \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}s} = -H_{\mathbf{x}}. \tag{13}$$

Note that for the proposed L and H, the Legendre transformation holds (e.g. Arnold, 1989; Slawinski, 2015; Červený, 2002a, 2002b)

$$L(\mathbf{x},\mathbf{r}) = \mathbf{p}\cdot\mathbf{r} - H(\mathbf{x},\mathbf{p})\,. \tag{14}$$

We re-emphasize that the arclength-related Hamiltonian is defined as that whose associated flow variable in the ray tracing equations is the arclength. In this study, we obtained the matching arclength-related Hamiltonian $H(\mathbf{x},\mathbf{p})$ from the reference Hamiltonian $H^{\bar{\tau}}(\mathbf{x},\mathbf{p})$, but it can also be obtained directly from our proposed arclength-related Lagrangian $L(\mathbf{x},\mathbf{r})$, using the Legendre transformation, provided the Christoffel equation is taken into account, as demonstrated by Koren and Ravve (2020b).

In the computational workflow of the Eigenray method, equation (2) for the Lagrangian is used rather than equation (14). The latter is given solely to demonstrate that our proposed Lagrangian and Hamiltonian (both arclength-related) match the Legendre transformation.

COMPUTING THE SLOWNESS VECTOR GIVEN THE RAY DIRECTION

The ray direction vector $\mathbf{r}$ and the slowness vector $\mathbf{p}$ are dependent ray characteristics and should match each other. We distinguish between the forward problem of finding $\mathbf{r}(\mathbf{x},\mathbf{p})$ and the inverse problem of finding $\mathbf{p}(\mathbf{x},\mathbf{r})$. The forward problem is simpler, but in the proposed Eigenray approach, the ray location $\mathbf{x}$ and its direction $\mathbf{r}$ are the primary (input) degrees of freedom, while the slowness vector $\mathbf{p}$ is a dependent parameter to be established, given the medium properties $\tilde{\mathbf{C}}(\mathbf{x})$ (the stiffness tensor) and the ray direction $\mathbf{r}$. Thus, we mostly deal with this kind of (nontrivial) inverse problem. The forward problem can be solved directly, applying the gradient of the arclength-related Hamiltonian with respect to (wrt) the slowness vector, $H_{\mathbf{p}} = \mathbf{r}$ (this is also one of the kinematic ray tracing equations). The inverse problem in this study is solved by applying the Hamiltonian-based approach, exploiting the collinearity of the Hamiltonian gradient (wrt the slowness vector) and the ray direction (e.g. Musgrave, 1954; Fedorov, 1968; Grechka, 2017), along with the condition for the vanishing Hamiltonian,

$$\underbrace{H_{\mathbf{p}}^{\bar{\tau}}(\mathbf{x},\mathbf{p})\times\mathbf{r} = 0}_{\substack{\text{pick two equations}\\ \text{from three}}}\,,\ \underbrace{H^{\bar{\tau}}(\mathbf{x},\mathbf{p}) = 0}_{\substack{\text{the third}\\ \text{equation}}}\,. \tag{15}$$

In this equation set, the reference Hamiltonian $H^{\bar{\tau}}(\mathbf{x},\mathbf{p})$ (defined in equation (10)) can be replaced by any other Hamiltonian (e.g. by the arclength-related Hamiltonian, $H \equiv H^{s}$); however, we consider the reference Hamiltonian $H^{\bar{\tau}}$ the simplest for this problem.

To solve set (15) for the slowness vector components, we apply the Newton method. In this paper, we mainly focus on compressional waves, where the solution (the components of the slowness vector) is unique, given the ray velocity direction. To obtain the initial guess, we may assume that the slowness direction does not differ much from the ray direction, $\mathbf{n}\approx\mathbf{r}$, and we find the compressional phase velocity magnitude, v_{phs}, from the Christoffel equation, where $\mathbf{p} = \mathbf{n}/v_{\text{phs}}$. A more robust initial guess is suggested by Pšenčík and Vavryčuk (2002), improved by Farra (2004) and later applied by Farra and Pšenčík (2013). Their method is valid for compressional waves only and is based on the weak-anisotropy assumption. Koren and Ravve (2020a) formulated it in a tensor form:

$$\mathbf{r}-\mathbf{n} \approx 2\frac{\mathbf{n}\times\mathbf{v}\times\mathbf{n}}{\mathbf{n}\cdot\mathbf{v}}\,,\ \text{where}\ \mathbf{v}\equiv\hat{\mathbf{\Gamma}}\mathbf{n}\,,\ \hat{\mathbf{\Gamma}} = \mathbf{n}\tilde{\mathbf{C}}\mathbf{n}, \tag{16}$$

where, due to the weak anisotropy, the phase direction $\mathbf{n}$ can be replaced by the ray direction $\mathbf{r}$ on the right-hand side of each equation of set (16). We use notation $\hat{\mathbf{\Gamma}}$ in equation (16), because $\mathbf{\Gamma}$ is reserved for $\mathbf{p}\tilde{\mathbf{C}}\mathbf{p}$. In cases of strong anisotropy (especially for low symmetry anisotropy), the compressional wave initial guess may converge to a final shear-wave solution. Special measures described by Ravve and Koren (2019) should be taken to avoid convergence to a wrong wave mode. The cited paper also provides a numerical example

for establishing the compressional slowness components in a triclinic medium with strong anisotropy. The technique for solving set (15) is explained in Ravve and Koren (2019) and Koren and Ravve (2020a).

Finally, after the slowness vector is found, the ray velocity magnitude is computed,

$$v_{\text{ray}} = \frac{1}{\mathbf{p}\cdot\mathbf{r}}. \tag{17}$$

SPATIAL AND DIRECTIONAL DERIVATIVES OF THE LAGRANGIAN

Using the proposed arclength-related Lagrangian (equation (2)), we derive and list the first and second derivatives of the proposed Lagrangian $L[\mathbf{x}(s), \mathbf{r}(s)]$ with respect to the locations and directions along the ray trajectory,

$$\begin{aligned}
L_{\mathbf{x}} &= -\frac{\nabla_{\mathbf{x}} v_{\text{ray}}\sqrt{\mathbf{r}\cdot\mathbf{r}}}{v_{\text{ray}}^2},\ L_{\mathbf{r}} = \frac{\mathbf{r}}{v_{\text{ray}}\sqrt{\mathbf{r}\cdot\mathbf{r}}} - \frac{\nabla_{\mathbf{r}} v_{\text{ray}}\sqrt{\mathbf{r}\cdot\mathbf{r}}}{v_{\text{ray}}^2},\\
L_{\mathbf{xx}} &= 2\frac{\nabla_{\mathbf{x}} v_{\text{ray}} \otimes \nabla_{\mathbf{x}} v_{\text{ray}}}{v_{\text{ray}}^3} - \frac{\nabla_{\mathbf{x}}\nabla_{\mathbf{x}} v_{\text{ray}}}{v_{\text{ray}}^2},\\
L_{\mathbf{xr}} = L_{\mathbf{rx}}^T &= -\frac{\nabla_{\mathbf{x}} v_{\text{ray}} \otimes \mathbf{r}}{v_{\text{ray}}^2} + 2\frac{\nabla_{\mathbf{x}} v_{\text{ray}} \otimes \nabla_{\mathbf{r}} v_{\text{ray}}}{v_{\text{ray}}^3} - \frac{\nabla_{\mathbf{x}}\nabla_{\mathbf{r}} v_{\text{ray}}}{v_{\text{ray}}^2},\\
L_{\mathbf{rr}} &= \frac{\mathbf{I} - \mathbf{r}\otimes\mathbf{r}}{v_{\text{ray}}} - \frac{\mathbf{r}\otimes\nabla_{\mathbf{r}} v_{\text{ray}} + \nabla_{\mathbf{r}} v_{\text{ray}}\otimes\mathbf{r}}{v_{\text{ray}}^2}\\
&\quad + 2\frac{\nabla_{\mathbf{r}} v_{\text{ray}} \otimes \nabla_{\mathbf{r}} v_{\text{ray}}}{v_{\text{ray}}^3} - \frac{\nabla_{\mathbf{r}}\nabla_{\mathbf{r}} v_{\text{ray}}}{v_{\text{ray}}^2},
\end{aligned} \tag{18}$$

where symbol $\otimes$ denotes a tensor product of two vectors resulting in a second-order tensor. The gradients of the Lagrangian $L_{\mathbf{x}}$, $L_{\mathbf{r}}$ are vectors of length 3, and the Hessians of the Lagrangian $L_{\mathbf{xx}}$, $L_{\mathbf{xr}}$, $L_{\mathbf{rx}}$, $L_{\mathbf{rr}}$ are square matrices (tensors) of dimension 3. Vectors $\nabla_{\mathbf{x}} v_{\text{ray}}$ and $\nabla_{\mathbf{r}} v_{\text{ray}}$ are the spatial and directional gradients of the ray velocity, respectively. Tensors $\nabla_{\mathbf{x}}\nabla_{\mathbf{x}} v_{\text{ray}}$ and $\nabla_{\mathbf{r}}\nabla_{\mathbf{r}} v_{\text{ray}}$ are, respectively, the spatial and directional Hessians of the ray velocity and $\nabla_{\mathbf{x}}\nabla_{\mathbf{r}} v_{\text{ray}}$ and $\nabla_{\mathbf{r}}\nabla_{\mathbf{x}} v_{\text{ray}}$ are the mixed Hessians. Ravve and Koren (2019) and Koren and Ravve (2020a, Appendix E) provide a computational workflow to establish these gradients and Hessians in smooth heterogeneous general anisotropic media. The first derivatives $L_{\mathbf{x}}$ and $L_{\mathbf{r}}$ define the local traveltime gradients, while the second derivatives $L_{\mathbf{xx}}$, $L_{\mathbf{xr}}$, $L_{\mathbf{rx}}$, $L_{\mathbf{rr}}$ define the local traveltime Hessians, used in both kinematic and dynamic analyses.

Alternatively, taking into account the momentum equation, $L_{\mathbf{r}} = \mathbf{p}$, the gradients and Hessians of the Lagrangian can be expressed through the corresponding derivatives of the slowness vector:

$$\begin{aligned}
&L_{\mathbf{r}} = \mathbf{p},\ L_{\mathbf{x}} = \mathbf{r}\cdot\nabla_{\mathbf{x}}\mathbf{p},\ L_{\mathbf{xx}} = \mathbf{r}\cdot\nabla_{\mathbf{x}}\nabla_{\mathbf{x}}\mathbf{p},\\
&L_{\mathbf{rx}} = \nabla_{\mathbf{x}}\mathbf{p},\ L_{\mathbf{xr}} = \nabla_{\mathbf{x}}^T\mathbf{p},\ L_{\mathbf{rr}} = \nabla_{\mathbf{r}}\mathbf{p}.
\end{aligned} \tag{19}$$

Remark: We note that for the kinematic problem, the Hessian matrices of the Lagrangian, $L_{\mathbf{xx}}$, $L_{\mathbf{rr}}$, $L_{\mathbf{xr}}$ and $L_{\mathbf{rx}} = L_{\mathbf{xr}}^T$, are used only within the method chosen for searching the stationary path (e.g. the Newton optimization method); other approaches (e.g. gradient-type methods) may require the Lagrangian gradients, $L_{\mathbf{x}}$ and $L_{\mathbf{r}}$ only, while these Hessians are the building stones in the calculation of the dynamic characteristics.

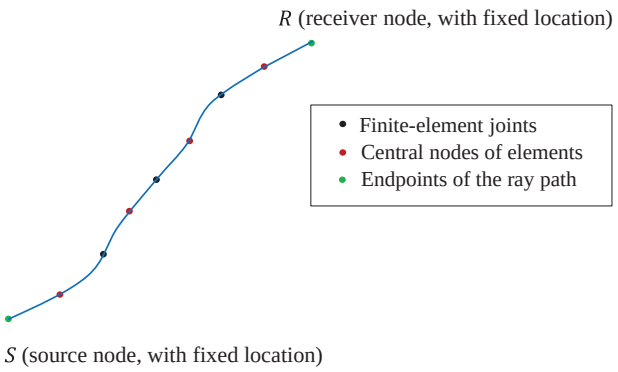

Figure 2 Discretization scheme of a ray path, with four three-node finite elements.

FINITE-ELEMENT DISCRETIZATION

In this study, we apply two- and three-node finite elements. A schematic ray path discretization is shown in Figure 2, where four three-node elements are used. The fixed endpoints of the ray path are shown in green, the finite-element joints in black and the internal nodes of the elements in red. (Obviously, a realistic scheme includes much more elements and nodes.) We define a segment of the path as an interval between two neighbour nodes. Two-node elements consist of a single segment, while three-node elements include two segments. The continuous ray trajectory $\mathbf{x}(s)$, where s is the arclength flow parameter (with an infinite number of degrees of freedom [DoF]) is represented (approximated) by a finite number of segments (intervals of finite elements) with nodes at their joints, $\mathbf{x}_i$. At the joint nodes connecting the neighbour elements, the locations, $\mathbf{x}(s)$, and ray directions $\mathbf{r}(s) \equiv \dot{\mathbf{x}}(s) \equiv d\mathbf{x}/ds$, related to the adjacent elements are normally continuous (except the directional discontinuities at the medium interfaces). At the internal nodes of the elements, the locations, ray directions and all higher derivatives of the locations with respect to

the arclength are continuous. The number of degrees of freedom (DoF) becomes finite: three location components $\mathbf{x}_i$ and three ray direction components $\mathbf{r}_i$, per node i. The three nodal direction DoF are dependent since a direction in 3D space is defined by two angles; we keep the ray direction normalized to the unit length, $\mathbf{r}_i \cdot \mathbf{r}_i = 1$, and we enforce this dependency with the soft normalization constraint.

The resolving nonlinear algebraic equation set of the finite-element analysis delivers a solution to the kinematic problem (locations and ray directions) at all nodes, and the Hermite interpolation makes it possible to compute the solution continuously for any point between the nodes.

HERMITE INTERPOLATION FUNCTIONS

Our proposed anisotropic Eigenray method is based on imposing continuities of the locations and ray velocity directions at the ray trajectory nodes. The ray direction Cartesian components are derivatives of the corresponding location components with respect to the arclength, and the ray trajectory along the intervals is presented by the Hermite interpolation polynomials. The Hermite polynomials provide the interpolation of a function accounting for its nodal values and the nodal values of its derivatives (e.g. Hildebrand, 1987; Burden and Faires, 2005). The Hermite finite elements naturally support these conditions. In 3D space, each node has three location components and three direction components, but since the direction is described by a tangent vector of unit length, the number of independent degrees of freedom (DoF) is five per node.

We follow conventional rules to construct the Hermite polynomials for two-node and three-node finite elements.

The internal unitless flow parameter

To keep the limits of integration fixed and identical for all finite elements, it is natural and convenient to define an internal unitless flow parameter, ξ, $-1 \le \xi \le 1$, inside the finite elements, to be used instead of the arclength s. For this we introduce a new, 'normalized' Lagrangian, $\hat{L}(\xi)$, and formulate the governing relationships in terms of the internal parameter ξ, where the transformation metric relates the arclength s to ξ.

Two-node Hermite element

An element with two end nodes has 12 DoF, only 10 of which are independent due to the abovementioned constraints. Let A and B be the 'left' and 'right' ends of a finite element, as shown in Figure 3. The traveltime increases from A to B, $t_a < t_b$. The

Figure 3 Scheme of two-node Hermite element.

internal unitless parameter ξ takes the values -1 and $+1$ at the element endpoints A and B, respectively.

Assume that a function $f(\xi)$ (that may be, for example, a scalar function or any Cartesian component of a vector) and its derivative, $f'(\xi) = \mathrm{d}f/\mathrm{d}\xi$, are specified at the endpoints,

$$\{f_a,\ f'_a,\ f_b,\ f'_b\}. \tag{20}$$

The interpolation function can be presented as

$$f(\xi) = f_a h_a(\xi) + f'_a d_a(\xi) + f_b h_b(\xi) + f'_b d_b(\xi). \tag{21}$$

Functions $h_a(\xi)$, $d_a(\xi)$, $h_b(\xi)$, $d_b(\xi)$ are all cubic polynomials, given by

$$\begin{aligned} h_a(\xi) &= +\tfrac{(1-\xi)^2(2+\xi)}{4}, d_a(\xi) = +\tfrac{(1-\xi)^2(1+\xi)}{4}, \\ h_b(\xi) &= +\tfrac{(1+\xi)^2(2-\xi)}{4}, d_b(\xi) = -\tfrac{(1+\xi)^2(1-\xi)}{4}. \end{aligned} \tag{22}$$

The shape functions are plotted in Figure 4.

Three-node Hermite element

A three-node Hermite element, with the nodes at the endpoints A, C, $\xi = \pm 1$, and an additional central node B, $\xi = 0$, is shown in Figure 5. It provides a better accuracy of the ray path for the same total number of trajectory nodes.

Given the nodal function values and nodal derivatives at the three nodes A, B, C,

$$\{f_a,\ f'_a,\ f_b,\ f'_b,\ f_c,\ f'_c\}. \tag{23}$$

The interpolation function reads

$$\begin{aligned} f(\xi) \ = \ & f_a h_a(\xi) + f_b h_b(\xi) + f_c h_c(\xi) + f'_a d_a(\xi) + f'_b d_b(\xi) \\ & + f'_c d_c(\xi), \end{aligned} \tag{24}$$

where the interpolation (shape) functions have properties similar to those of the two-node element. Functions $h_a(\xi)$, $d_a(\xi)$, $h_b(\xi)$, $d_b(\xi)$, $d_c(\xi)$ are fifth-degree polynomials, and $h_c(\xi)$ is a fourth-degree polynomial (as it has to be an even function), given by

$$\begin{aligned} h_a &= +\tfrac{1}{4}(1-\xi)^2\xi^2(4+3\xi),\ d_a = +\tfrac{1}{4}(1-\xi)^2\xi^2(1+\xi), \\ h_b &= +\tfrac{1}{4}(1+\xi)^2\xi^2(4-3\xi),\ d_b = -\tfrac{1}{4}(1+\xi)^2\xi^2(1-\xi), \\ h_c &= +\left(1-\xi^2\right)^2, d_c = +\xi\left(1-\xi^2\right)^2. \end{aligned} \tag{25}$$

The shape functions are plotted in Figure 6.

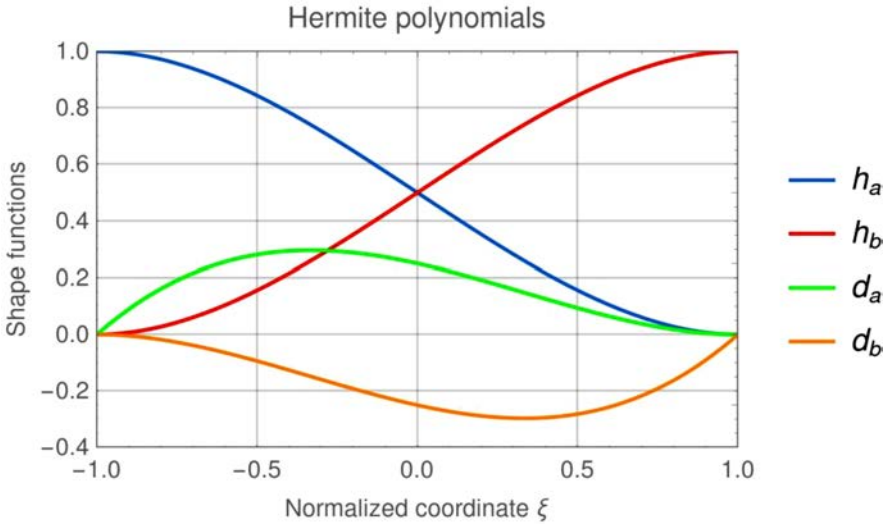

Figure 4 Shape functions for two-node Hermite interpolation.

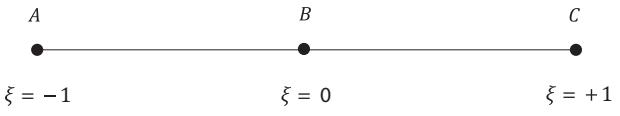

Figure 5 Scheme of three-node Hermite interpolation.

The location of a ray trajectory point between the nodes of a finite element is given by

$$\mathbf{x}(\xi) = \sum_{I=1}^{n} \mathbf{x}_I h_I(\xi) + \sum_{I=1}^{n} \mathbf{x}'_I d_I(\xi)\ ,\ \mathbf{x}' = \mathrm{d}\mathbf{x}/\mathrm{d}\xi, \tag{26}$$

where I is the index of the node and $n = 2,\ 3$ is the number of the nodes of the finite element. I stands for $a,\ b$ in two-node elements and for $a,\ b,\ c$ in three-node elements.

In this study, the interpolation functions are applied to the ray path points' locations. The quadrature formulae for Hermite interpolation functions are presented in Appendix B of Koren and Ravve (2020c).

DERIVATIVES WITH RESPECT TO THE INTERNAL FLOW PARAMETER

The total traveltime, which has to be stationary for the resolved path, is the sum of the local traveltimes, $t = \sum \Delta t_i$. The term 'local' means a characteristic within a single finite element. The quadrature (integration) formulae for computing the local traveltimes Δt_i are based on the Hermite interpolation and make it possible to compute the traveltime and its derivatives between the nodes of each element.

To keep the limits of integration fixed and identical for all finite elements, inside the finite elements, we introduce an internal unitless flow parameter, $\xi,\ \ -1 \le \xi \le 1$, to be used instead of the arclength s in the finite-element solver. Given the values of the nodal locations $\mathbf{x}$ and directions $\mathbf{r}_i = \dot{\mathbf{x}}_i = \mathrm{d}\mathbf{x}/\mathrm{d}s|_i$, we follow Yong and Cheng (2004) and minimize the pseudo strain energy within a single element in order to convert the nodal directions $\mathbf{r}_i$ into the nodal derivatives of the position with respect to (wrt) the internal parameter, $\mathbf{x}'_i = \mathrm{d}\mathbf{x}/\mathrm{d}\xi|_i$. For this we compute the nodal values of the metric, $s'_i = \mathrm{d}s/\mathrm{d}\xi|_i$. The details of the implementation can be found in Koren and Ravve (2020c, Appendix C). Note that the ray direction $\mathbf{r}$ is normally continuous at the finite-element joints (except special cases of sharp medium interfaces), while the derivative $\mathbf{x}'(\xi)$ need not to be continuous.

The ray direction wrt to the internal unitless flow parameter at any point reads

$$\mathbf{r}(\xi) = \dot{\mathbf{x}} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\xi}\frac{\mathrm{d}\xi}{\mathrm{d}s} = \frac{\mathbf{x}'(\xi)}{s'(\xi)} = \frac{\mathbf{x}'(\xi)}{\sqrt{\mathbf{x}'(\xi)\cdot\mathbf{x}'(\xi)}}. \tag{27}$$

We introduce the normalized Lagrangian, where the arclength flow parameter is replaced by the internal parameter, ξ. In particular, we relate the spatial/directional gradients and spatial/directional/mixed Hessians of the normalized Lagrangian $\hat{L}(\xi)$ (equation (29) below) to their corresponding gradients and Hessians of the arclength-related Lagrangian $L(s)$.

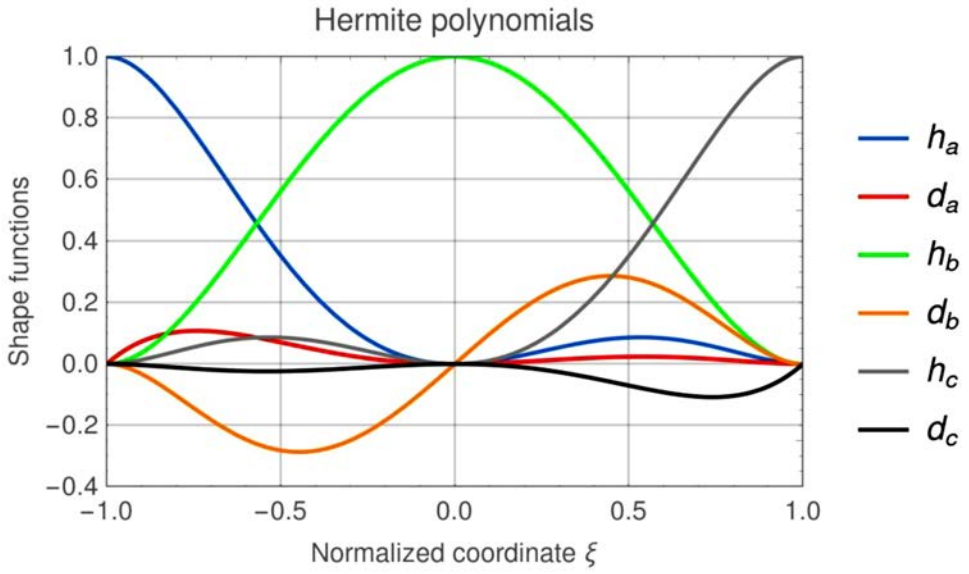

Figure 6 Shape functions for three-node Hermite interpolation.

Hence, for each element we compute the local traveltime

$$\Delta t = \int_{\xi=-1}^{\xi=+1} \frac{\sqrt{\mathbf{x}'(\xi)\cdot\mathbf{x}'(\xi)}}{v_{\text{ray}}\left[\mathbf{x}(\xi),\mathbf{x}'(\xi)\right]}\,\mathrm{d}\xi = \int_{\xi=-1}^{\xi=+1} \hat{L}(\xi)\,\mathrm{d}\xi. \tag{28}$$

The normalized Lagrangian $\hat{L}(\xi)$ is then the traveltime integrand wrt the unitless internal parameter ξ. Note that $\hat{L}(\xi)$ has units of time, $[T]$, while the arclength-related Lagrangian, $L(\mathbf{x},\mathbf{r}) = L(s)$, has units of slowness, $[T/L]$. The two Lagrangians are related with the metric, $\mathrm{d}s/\mathrm{d}\xi$,

$$\hat{L}(\mathbf{x},\mathbf{r})\,\mathrm{d}\xi = L(\mathbf{x},\mathbf{r})\,\mathrm{d}s = \mathrm{d}\tau \;\rightarrow\; \hat{L}(\mathbf{x},\mathbf{r}) = L(\mathbf{x},\mathbf{r})\,\frac{\mathrm{d}s}{\mathrm{d}\xi}. \tag{29}$$

The relations between the gradients and Hessians of the normalized and arclength-related Lagrangians are listed in Appendix D of Koren and Ravve (2020c). Next, we compute the local traveltime derivatives wrt the nodal positions and directions: the local gradient (vector of length $6n$) and the Hessian (matrix of dimensions $6n \times 6n$), where n is the number of nodes in a single element (i.e. either 2 or 3 in this study).

THE LOCAL ALGEBRAIC EQUATION SET

The weak formulation solution

We apply the weak formulation with the Galerkin (1915) method to equation (8), where both the interpolation and the test (weight) functions are the same Hermite polynomials. The residual of the ordinary differential equations (ODE) is orthogonal to each test function. 'Orthogonal' means that the integral of their product within the finite-element arclength vanishes. This procedure includes integration by parts and effectively reduces the second-order nonlinear ODE set to the first-order, nonlinear, local weighted-residual set,

$$\int_{s_{\text{ini}}}^{s_{\text{fin}}} \frac{\mathrm{d}}{\mathrm{d}s}\left(\frac{\mathbf{r}}{v_{\text{ray}}} - \frac{\nabla_{\mathbf{r}} v_{\text{ray}}}{v_{\text{ray}}^2}\right) w(s)\,\mathrm{d}s = -\int_{s_{\text{ini}}}^{s_{\text{fin}}} \frac{\nabla_{\mathbf{x}} v_{\text{ray}}}{v_{\text{ray}}^2}\, w(s)\,\mathrm{d}s, \tag{30}$$

where $h_I(\xi)$ and $d_I(\xi)$ are the weight (interpolation) functions, and s_{ini} and s_{fin} are the 'initial' and 'final' values of the arclength at the two external nodes of the finite element. It is suitable to switch from the arclength s to the normalized internal flow parameter $-1 \le \xi \le +1$. We then perform the integration by parts, and obtain the resulting equation with the boundary conditions at the finite element ends,

$$\begin{aligned}&(\mathbf{p}\,w)\big|_{\xi=-1}^{\xi=+1} - \int_{\xi=-1}^{\xi=+1}\left(\frac{\mathbf{r}}{v_{\text{ray}}} - \frac{\nabla_{\mathbf{r}} v_{\text{ray}}}{v_{\text{ray}}^2}\right) w'\,\mathrm{d}\xi \\ &= -\int_{\xi=-1}^{\xi=+1} \frac{\nabla_{\mathbf{x}} v_{\text{ray}}}{v_{\text{ray}}^2}\, w\, s'\mathrm{d}\xi,\end{aligned} \tag{31}$$

where prime means a derivative by ξ, and $s' = \mathrm{d}s/\mathrm{d}\xi = \sqrt{\mathbf{x}'(\xi)\cdot\mathbf{x}'(\xi)} \equiv l(\xi)/2$.

The boundary term in equation (31) is the slowness vector with the corresponding sign (a plus sign at the end node of the element pointing to the receiver, and a minus sign at the end node pointing to the source) that appears only for weight

functions that are 1 on the corresponding end of the interval. At the internal joints, the boundary terms cancel each other during the assembly, so we can omit them. At the source and receiver, the boundary terms remain after the assembly, but we remove the equations that include them and replace by our boundary conditions (given locations of the ray path endpoints). Thus, we can remove the slowness term for all cases, and equation (30) simplifies to

$$\int_{\xi=-1}^{\xi=+1}\left(\frac{\mathbf{r}\,w'}{v_{\text{ray}}}-\frac{\nabla_{\mathbf{r}}v_{\text{ray}}\,w'}{v_{\text{ray}}^2}-\frac{\nabla_{\mathbf{x}}v_{\text{ray}}}{v_{\text{ray}}^2}\,w\,s'\right)\,\mathrm{d}\xi=0. \tag{32}$$

Applying the spatial and directional shape functions, $w = h_I(\xi)$ and $w = d_I(\xi)$, respectively, and taking into account that, $\mathbf{r} = 2\mathbf{x}'(\xi)/l(\xi)$, we obtain

$$\begin{aligned}&\int_{\xi=-1}^{\xi=+1}\left[\frac{1}{2}l\,(\xi)\;h_I\,(\xi)\,L_{\mathbf{x}}+h'_I L_{\mathbf{r}}\right]\mathrm{d}\xi=0,\\&\int_{\xi=-1}^{\xi=+1}\left[\frac{1}{2}l\,(\xi)\;d_I\,(\xi)\,L_{\mathbf{x}}+d'_I L_{\mathbf{r}}\right]\mathrm{d}\xi=0,\end{aligned} \tag{33}$$

where $h_I(\xi)$ and $d_I(\xi)$ are the Hermite interpolation polynomial related to the location and direction components, respectively, of node I (numeration within a single element). Introducing the spatial and directional gradients of the Lagrangian (equation (7)) into equation (32), we obtain the algebraic equation set

$$\begin{aligned}&\int_{\xi=-1}^{\xi=+1}\left[\frac{2h'_I\,(\xi)\;\mathbf{x}'\,(\xi)}{v_{\text{ray}}\,l\,(\xi)}-\frac{h'_I\,(\xi)\;\nabla_{\mathbf{r}}v_{\text{ray}}}{v_{\text{ray}}^2}-\frac{l\,(\xi)\,h_I\,(\xi)\,\nabla_{\mathbf{x}}v_{\text{ray}}}{2v_{\text{ray}}^2}\right]\mathrm{d}\xi=0,\\&\int_{\xi=-1}^{\xi=+1}\left[\frac{2d'_I\,(\xi)\;\mathbf{x}'\,(\xi)}{v_{\text{ray}}\,l\,(\xi)}-\frac{d'_I\,(\xi)\;\nabla_{\mathbf{r}}v_{\text{ray}}}{v_{\text{ray}}^2}-\frac{l\,(\xi)\,d_I\,(\xi)\,\nabla_{\mathbf{x}}v_{\text{ray}}}{2v_{\text{ray}}^2}\right]\mathrm{d}\xi=0,\end{aligned} \tag{34}$$

where the spatial gradient of the ray velocity magnitude

$$\nabla_{\mathbf{x}}v_{\text{ray}}\,(\xi)=\nabla_{\mathbf{x}}v_{\text{ray}}\left[\mathbf{x}\,(\xi)\,,\mathbf{r}\,(\xi)\right]=\nabla_{\mathbf{x}}v_{\text{ray}}\left[\mathbf{x}\,(\xi)\,,\mathbf{x}'\,(\xi)\right], \tag{35}$$

and a similar equation holds for the directional gradient $\nabla_{\mathbf{r}}v_{\text{ray}}$.

Equation (34) represents the nonlinear kinematic algebraic set for a single finite element (i.e. before assembly). The set includes 12 or 18 scalar equations for two- and three-node elements, respectively. This set expresses the vanishing gradient of the local traveltime with respect to the degrees of freedom of the given element. This set does not account for the constraints yet; the contribution of the constraints into the target function is accounted separately. The derivatives of the left-hand sides of the functions in equation set (34) with respect to the nodal values of the ray path coordinates $\mathbf{x}_J$ and with respect to $\mathbf{x}'_J$ result in spatial, directional and mixed blocks IJ of the symmetric local Hessian matrix. The traveltime Hessians of the two- and three-node elements include 16 and 36 blocks, respectively, where each block is a 3×3 matrix. The local gradients include four and six blocks, respectively, where each block is a vector of length 3.

The weak formulation effectively reduces the second-order nonlinear ODE set to the first-order, nonlinear, algebraic local weighted-residual equation set. A detailed derivation is given by Koren and Ravve (2020c). The technical details for computing the local traveltime, its gradient vectors and Hessian matrices are presented in Appendices E and F of the cited work.

CONSTRAINT ON THE NODE DISTRIBUTION ALONG A STATIONARY PATH

While the stationary traveltime condition fully defines the ray path, it still allows some freedom for setting the distribution of the nodes along the ray. We therefore apply an additional constraint W_s on the segment lengths between the successive nodes, so that the nodes are located more densely along ray parts with high curvature. In these parts of the path, the ray velocity magnitude changes rapidly in the direction normal to the ray velocity vector. We construct constraints on the ratios between the arclengths connecting successive nodes. These lengths are inversely proportional to the average curvatures of the corresponding intervals,

$$W_s=\frac{w_s}{2}\sum_{i=1}^{N-1}\left(\frac{\Delta s_i}{\Lambda_i}-\frac{\Delta s_{i+1}}{\Lambda_{i+1}}\right)^2, \tag{36}$$

where w_s is a weighting factor of the constraint, N is the number of intervals (ray path segments), $N + 1$ is the total number of nodes enumerated from zero to N, and $N - 1$ is the number of internal nodes of the path. Δs_i is the arclength of segment i, and parameter Λ_i is the mean radius of curvature of the ray along this segment, normalized so that $\sum_{i=1}^{N}\Lambda_i = 1$; Λ_i is limited to a finite (large) value for straight segments.

In general, there are two ways to implement the constraints: hard constraints (e.g. applying the Lagrange multipliers method) and soft or relaxed constraints, by adding a penalty term to the target function to be minimized. The soft constraint method (used in this study) is simpler; it does not lead to additional unknown parameters and does not increase

the bandwidth of the resolving matrix, while still providing excellent accuracy. There is no need to keep the desired length ratio exact. The implementation details of the node distribution constraint are presented in Appendix H of the study by Koren and Ravve (2020c).

CONSTRAINT ON THE RAY DIRECTION NORMALIZATION

Throughout the workflow, the ray velocity direction in the governing equations is assumed normalized, where at each node, $\mathbf{r}\cdot\mathbf{r}=1$. However, the Newton iterative procedure (or any other iterative procedure, like the conjugate-gradient or anti-gradient descent, that does not explicitly require the second derivatives of the target function) provides a 'recommended' set of updated parameters at the end of each successive iteration, which can violate the normalization of the ray direction. For example, the direction at node i becomes $\mathbf{r}_i+\Delta\mathbf{r}_i$, where $|\mathbf{r}_i|=1$, but the length $|\mathbf{r}_i+\Delta\mathbf{r}_i|$ may essentially differ from 1. The remedy is to include within the target function an additional normalization penalty term W_r,

$$W_r=\frac{w_r}{2}\sum_{i=0}^{N}(\mathbf{r}_i\cdot\mathbf{r}_i-1)^2, \tag{37}$$

where w_r is a weighting factor of the constraint, and $N+1$ is the total number of nodes, including the endpoints of the path. The implementation details of the normalization constraint are presented in Appendix H of Koren and Ravve (2020c).

The normalization constraints might be, in principle, implemented with the Lagrange multipliers λ_i. This approach, however, increases the number of degrees of freedom per node from six to seven, which affects not only the size of the Hessian matrix, but also – more important – its band width.

In this study, both constraints are implemented as soft constraints. Our computational practice indicates that the approach with the two soft constraints works perfectly. At the stationary point, the values of the weighted constraint terms are of order 10^{-5} or 10^{-6} of the unconstrained target function.

OPTIMIZATION OF THE TARGET FUNCTION

As mentioned, in the case of a minimum traveltime, the target function T includes the traveltime and two weighted penalty terms. The target function and its gradient read,

$$T=t+W_s+W_r\rightarrow\min\,,\ \nabla_{\mathrm{d}}T=\nabla_{\mathrm{d}}t+\nabla_{\mathrm{d}}W_s+\nabla_{\mathrm{d}}W_r=0, \tag{38}$$

where t is the traveltime, symbol ∇_{d} means the gradient of a scalar function with respect to (wrt) all (location and direction) degrees of freedom (DoF), W_s is the node distribution penalty, and W_r is the ray direction normalization penalty. Theoretically, the stationary path may also deliver a maximum traveltime; however, in real problems this case is very rare (non-realistic) – it requires all eigenvalues of the traveltime Hessian matrix to be negative. On the other hand, the case of stationary traveltime with saddle points is very common, indicating caustic locations along the ray. In this case or in cases when the type of stationary point is unknown, we suggest that the target function consists of the traveltime gradient squared (instead of the traveltime) and the penalty terms, where the minimum of the target function corresponds to the vanishing (or negligibly small) norm of the gradient vector. For general or saddle-point cases, the target function and its gradient read

$$\begin{aligned} T&=\frac{\nabla_{\mathrm{d}}t\cdot\nabla_{\mathrm{d}}t}{2}+W_s+W_r\rightarrow\min\,,\\ \nabla_{\mathrm{d}}T&=\nabla_{\mathrm{d}}\nabla_{\mathrm{d}}t\cdot\nabla_{\mathrm{d}}t+\nabla_{\mathrm{d}}W_s+\nabla_{\mathrm{d}}W_r, \end{aligned} \tag{39}$$

where $\nabla_{\mathrm{d}}t$ and $\nabla_{\mathrm{d}}\nabla_{\mathrm{d}}t$ are respectively the gradient and Hessian of the traveltime wrt all DoF. We take into account that the location DoF have the units of distance, while the direction DoF are unitless. Note that for this general case of the stationary traveltime, in particular, for a saddle-point case when equation (39) is applied, the target function itself includes the traveltime gradient, and as a result, the gradient of the target function already includes the traveltime Hessian. Therefore, in these cases, we do not apply the Newton-type minimization methods that require the second derivative of the target function (and hence, higher order derivatives of the traveltime), and we only apply the gradient methods (e.g. the conjugate-gradient or anti-gradient descent). We further note that (a) in either case, whether the traveltime of the stationary path is the minimum or not, the target function is always minimized, and (b) for either of the minimization methods that we use (Newton-type or gradient), we need to compute both the gradient and the Hessian of the traveltime.

Convergence to a zero traveltime gradient squared (to the stationary ray path) may be sensitive to the initial-guess trajectory. If the initial guess is far from the stationary solution, the iterations may yield a non-vanishing minimum traveltime gradient squared.

For the traveltime minimum search, we use the Newton optimization method, where at each iteration of the optimization process, we update the nodal locations and directions of the ray and refine the trial trajectory until the stationary

condition is reached. The technical details of the Newton optimization procedure are explained in Appendix J of Koren and Ravve (2020c).

Since throughout the iterative process we often do not know in advance the type of the final stationary solution, and since the Eigenray method is normally applied in cases where conventional two-point kinematic ray tracing fails (these cases involve complex wave phenomena), the target function that includes the traveltime gradient squared, valid for both the minimum and the saddle-point solutions (equation (39)), is normally the default choice for the minimization. Numerical procedures for locating constrained saddle points have been studied by Zhang and Du (2012a, 2012b), Ren and Vanden-Eijnden (2013), Gao *et al*. (2015), Albareda *et al*. (2018), Li *et al*. (2019) and Li and Zhou (2019).

ASSEMBLY OF THE LOCAL TARGET FUNCTION GRADIENTS AND HESSIANS

The target function to be optimized includes the traveltime and two soft weighted constraints related to the distribution of nodes along the ray path and to the normalization of the directions of the ray velocity at the nodal points. Each element contributes to the traveltime, each joint node contributes to the distribution penalty, and each node contributes to the normalization penalty. The derivative of the sum equals to the sum of the derivatives, and this means that at the joints, where the local vectors and matrices overlap, the numbers are just added. Locations of the end nodes are known, and this constitutes the boundary conditions. The implementation details of the assembly procedure are presented in Appendix I of Koren and Ravve (2020c), where, in particular, a special case of assembly for discontinuous ray velocity direction through a sharp medium interface is explained.

TYPES OF STATIONARY PATHS AND EXISTENCE OF CAUSTICS

The first step of the analysis/computation of the dynamic characteristics along the derived stationary ray paths involves verifying whether the path delivers a minimum or a saddle-point (due to caustics) traveltime solution. This information can be obtained directly from the global traveltime Hessian matrix, by the analysis of all its eigenvalues (6×number of nodes). If all eigenvalues are positive, the stationary path is a minimum traveltime. We assume that the (hypothetic) case where all the eigenvalues are negative (a maximum traveltime solution) is not realistic for systems with multiple degrees of freedom. If one or more eigenvalues are negative, and the others are positive, it is a saddle-point solution, which is an indication for the existence of caustics: zeros of the (signed) cross-section area of the ray tube (ray Jacobian) along the ray. We note that this type of eigenvalue analysis of the global traveltime Hessian cannot provide the actual number of caustics, their locations and type: line or point, which are important for implementing the required phase correction of the asymptotic Green's function. This information requires explicit solution of the dynamic ray tracing vector equation.

GEOMETRIC SPREADING FROM GLOBAL TRAVELTIME HESSIAN

In many practical applications, only the geometric spreading between the source and receiver is needed, ignoring phase shifts due to caustics, where dynamic ray tracing is not a must. In this study, this type of geometric spreading is used for evaluating the reliability/plausibility of the resolved stationary ray paths using a special ray path complexity criterion described in the next section.

The geometric spreading is obtained from the mixed (off-diagonal) block of the source–receiver traveltime Hessian matrix. This symmetric endpoint matrix of dimension 6×6 includes the external degrees of freedom (DoF) of the source, S, and receiver, R, locations only. It consists of four blocks of dimension 3×3, where the upper and lower symmetric diagonal blocks are related to the source and receiver, respectively, while the two off-diagonal blocks, transposed to each other, are related to both the source and receiver. Goldin (1986) suggested a workflow to compute the geometric spreading using the 3×3 off-diagonal block, with its last line and column removed (related to the coordinate axis normal to the acquisition surface), so that the resulting 'mixed' traveltime Hessian matrix $\mathbf{M}_{RS}$ becomes 2×2.

Vanelle and Gajewski (2003, 2013) derived a relationship to compute the geometric spreading for general anisotropy, accounting for the transform between the Cartesian and ray-centred coordinates,

$$L_{GS} = \sqrt{\frac{\cos\theta_{\text{ray},S}}{\cos\beta_S}\frac{\cos\theta_{\text{ray},R}}{\cos\beta_R}\frac{1}{\det\mathbf{M}_{RS}}}, \tag{40}$$

where $\theta_{\text{ray},S}$ and $\theta_{\text{ray},R}$ are the ray (group) take-off angles at the endpoints of the path (i.e. angles between the ray velocities and the normal directions to the acquisition surfaces), β_S and β_R are angles between the phase and ray velocities, and subscripts S and R are related to the source and receiver, respectively.

The Eigenray method suggested in this study provides a natural way to compute the geometric spreading for the whole ray path between the source and receiver without explicitly performing dynamic ray tracing (DRT). Along with the ray path, arclength and traveltime, the Eigenray method provides the global traveltime Hessian matrix of the stationary path. The latter represents a large, narrow-band, square matrix that includes all DoF before imposing the boundary conditions. In this part of our study, we suggest a way to compress this matrix into a 6×6 endpoint traveltime Hessian, whose DoF are spatial locations of the source and receiver, $\mathbf{x}_S$ and $\mathbf{x}_R$.

The geometric spreading is obtained directly from the compressed Hessian matrix of the traveltime with respect to the endpoint locations. Each endpoint has three spatial DoF (coordinates); therefore, the endpoint traveltime Hessian is a symmetric matrix of dimension 6×6. Unlike the global traveltime Hessian that includes all nodal DoF and is subject to the boundary conditions, the endpoint traveltime Hessian is not necessarily positive definite even when the stationary path delivers the minimum traveltime.

The geometric spreading is computed in two stages:

- 'Condensing' the global traveltime Hessian matrix $\nabla_{\mathbf{d}}\nabla_{\mathbf{d}}\, t$ to the endpoint traveltime Hessian matrix $\widetilde{\nabla_{SR}\nabla_{SR}}t$, where the subscript $\mathbf{d}$ means all DoF (related to locations and directions), while S and R are related to the source and receiver locations, respectively. The dimension of Hessian $\widetilde{\nabla_{SR}\nabla_{SR}}t$ is 6×6 and we extract its 3×3 off-diagonal mixed block.
- Applying a conventional, well-known method (equation (40)) for computing the geometric spreading, given the extracted 3×3 mixed Hessian and other computed data.

We emphasize that when computing the Eigenray stationary path, the location components of the endpoints are fixed and not considered DoF. The DoF are the location and direction components of the internal nodes and also the direction components of the end nodes. On the other hand, when computing the geometric spreading for the whole ray, only the location components of the two end nodes (a total of six components) are considered independent (external) DoF. The locations of the internal nodes and the directions of the ray at all the nodes (including the end nodes) are internal, dependent DoF. Given the locations of the two end nodes, all these internal DoF are fully defined by the stationary ray. In the Eigenray approach, the traveltime depends on all DoF, $\mathbf{d}$, where $t = t(\mathbf{d}_E, \mathbf{d}_I)$, where $\mathbf{d}_E$ and $\mathbf{d}_I$ are the external and internal DoF, respectively. However, if the external DoF (endpoint locations) are fixed, then the internal DoF, $\mathbf{d}_I = \mathbf{d}_I(\mathbf{d}_E)$, can be explicitly computed, accounting for the vanishing traveltime gradient. Hence, the internal DoF can be excluded and the traveltime equation becomes only a function of the source and receiver locations,

$$t = t\left[\mathbf{d}_E, \mathbf{d}_I\left(\mathbf{d}_E\right)\right] = t\left(\mathbf{d}_E\right). \tag{41}$$

The corresponding 6×6 endpoint traveltime Hessian exists and reads

$$\widetilde{\nabla_{SR}\nabla_{SR}}\, t = \begin{bmatrix} \widetilde{\nabla_S\nabla_S}\, t & \widetilde{\nabla_S\nabla_R}\, t \\ \widetilde{\nabla_R\nabla_S}\, t & \widetilde{\nabla_R\nabla_R}\, t \end{bmatrix}. \tag{42}$$

The tilde symbol is needed to distinguish these blocks from the corresponding blocks of the global traveltime Hessian. Each of the four blocks in the matrix of equation (42) is 3×3. The implementation details for computing the endpoint traveltime Hessian matrix of equation (42) are explained in Ravve and Koren (2020a, Appendix A). The algorithm includes inversion of the 'stiffness' submatrix related to the internal DoF. However, this submatrix becomes invertible (i.e. its determinant does not vanish) only after implementing the constraints which are explained below. The physical reason for the vanishing determinant is that the internal nodes of a stationary paraxial ray can move along its path without actually affecting the path geometry (i.e. the paraxial ray path remains the same). Thus, the stationarity condition does not fully define the location of the nodes of the paraxial ray. To define the nodal locations of the paraxial ray, we assume that the shifts between the nodes of the central ray and paraxial rays are normal to the central ray. The left lower 3×3 block consisting of the mixed derivatives (i.e. those second derivatives where one coordinate belongs to the source and the other to the receiver) is needed to apply a conventional well-known approach for computing the geometric spreading, described in Appendix B of Ravve and Koren (2020a). As mentioned, the geometric spreading is a function of the compressed endpoint traveltime Hessian which, in turn, is a function of the global traveltime Hessian, related to the stationary ray path. The relative geometric spreading is insensitive to the swapping of the source and receiver. It does not matter whether it has been computed by applying the endpoint traveltime Hessian or by using the ray Jacobian (DRT).

RAY PATH COMPLEXITY CRITERION

In order to estimate the reliability/plausibility of the ray path obtained with the Eigenray kinematic ray tracing, we

introduce the normalized geometric spreading, L_{GS}/σ, and the *endpoint* propagation complexity criterion,

$$\hat{c}_r = \left(\frac{L_{GS}}{\sigma} - 1\right)^2. \tag{43}$$

A zero complexity corresponds to an isotropic medium with a constant (not necessarily vertical) velocity gradient, where the relative geometric spreading coincides with parameter sigma, defined as $\mathrm{d}\sigma = v_{\mathrm{ray}}\mathrm{d}s = v_{\mathrm{ray}}^2\mathrm{d}\tau$, i.e. $L_{GS} = \sigma$, and thus, the endpoint propagation complexity $\hat{c}_r$ vanishes. In Ravve and Koren (2021) of this study, we define the *weighted* propagation complexity $c_r \neq \hat{c}_r$ that accounts for the normalized relative geometric spreading not only at the receiver point, but along the whole stationary ray path.

NUMERICAL EXAMPLES

We present a number of synthetic examples for the proposed Eigenray solutions. In this part, we demonstrate the computational results for isotropic velocity fields. Two anisotropic numerical examples are included in Ravve and Koren (2021), where both kinematic (stationary ray) and dynamic (geometric spreading) characteristics of the ray paths have been computed. In this study, we consider direct waves between two endpoints (S and R). The transmission/reflection of the rays across interfaces are beyond the scope. In the following three examples, two-node Hermite elements were used.

Example 1. Eigenrays in high-velocity half-space under constant velocity layer ('head wave')

Consider a 1D velocity model whose vertical profile, gradient and second derivative are shown in Figure 7. The smoothed velocity model is depth dependent and described by

$$v = v_{\mathrm{o}} + \frac{\Delta v}{2}\left(1 + \tanh\frac{z - z_h}{\Delta z_h}\right), \tag{44}$$

where $z \equiv x_3$, $v_{\mathrm{o}} = 2\,\mathrm{km/s}$ is the velocity of the 'homogeneous' layer above the half-space, and $\Delta v = 2\,\mathrm{km/s}$ is the difference between the velocity of the half-space and that of the overlying layer; thus, the half-space velocity is $v_h = v_{\mathrm{o}} + \Delta v = 4\,\mathrm{km/s}$. Actually, neither the overlying layer nor the half-space is homogeneous, due to the transition zone. Parameter $z_h = 1.5\,\mathrm{km}$ is the mid-level of the vertical transition zone, and parameter $\Delta z_h = 0.2\,\mathrm{km}$ is the characteristic distance that shows the extent of the transition zone (which is approximately $5\,\Delta z_h$). The offset $h = 10$ km.

Figure 8(a) and 8(b) shows the Eigenray solution with 5 and 20 finite elements, respectively, and they are almost identical. The red line is the initial guess, presented by an elliptic arc, defined by three parameters: the bottom point depth and the endpoints' offset and take-off angle. The bottom point depth is 2 km, which is approximately the lower end of the transition zone. In Figure 8(a), lines of different colours show the segments of the path corresponding to different finite elements. In Figure 8(b), the stationary ray path is shown by a black line. Other initial paths, with overestimated maximum depth, lead to the same stationary solution. In the next example, we show that in the case of multi-arrivals, the final solution is sensitive to the initial guess.

Example 2. Eigenrays in a medium with low-velocity elliptic anomaly

Consider a constant background velocity with an elliptic anomaly region of a lower velocity, as shown in Figure 9(a). The coordinates of the ellipse centre are c_1, c_3, and the semi-axes of the ellipse are a_1, a_3. The velocity field is described by an analytic function:

$$v(x_1, x_3) = v_{\mathrm{o}} + \frac{\Delta v}{2}\left(\tanh A - 1\right), \tag{45}$$

where

$$A = \frac{1}{s_e}\left[\frac{(x_1 - c_1)^2}{a_1^2} + \frac{(x_3 - c_3)^2}{a_3^2} - 1\right]. \tag{46}$$

The background velocity outside the ellipse is indicated by v_{o}, and $v_{\mathrm{o}} - \Delta v$ is the anomalous low velocity inside the ellipse. Negative Δv leads to anomalous high velocity inside the ellipse. Parameter s_e is the smoothing scale: the smaller s_e is, the sharper the velocity change. For infinitesimal s_e, the velocity function becomes discontinuous. We accept the following parameters,

$$\begin{aligned} &v_{\mathrm{o}} = 5\,\mathrm{km/s}\,,\ \Delta v = 3\,\mathrm{km/s}\,,\ c_1 = 5\,\mathrm{km}\,,\ c_3 = 3\,\mathrm{km}\,,\\ &a_1 = 3\,\mathrm{km}\,,\ a_3 = 2\,\mathrm{km}\,,\ s_e = 0.2. \end{aligned} \tag{47}$$

The source is located at the subsurface point with zero horizontal coordinate and depth $d_{\mathrm{o}} = 6\,\mathrm{km}$, and the receiver is on the surface, with the one-way offset $h_s = 10\,\mathrm{km}$. The absolute value of the velocity gradient is shown in Figure 9(b). Three different initial guesses lead to three different solutions shown in Figure 10 by green, black and blue lines. The red ellipse is the contour of the velocity anomaly. The corresponding initial trajectories are shown by dashed lines of the same colours. The green and black lines are the 'shallow' and 'deep' solutions that bypass the anomaly from above and below, respectively. The blue line is the 'transmission' solution that penetrates into the anomalous region. The initial path for the

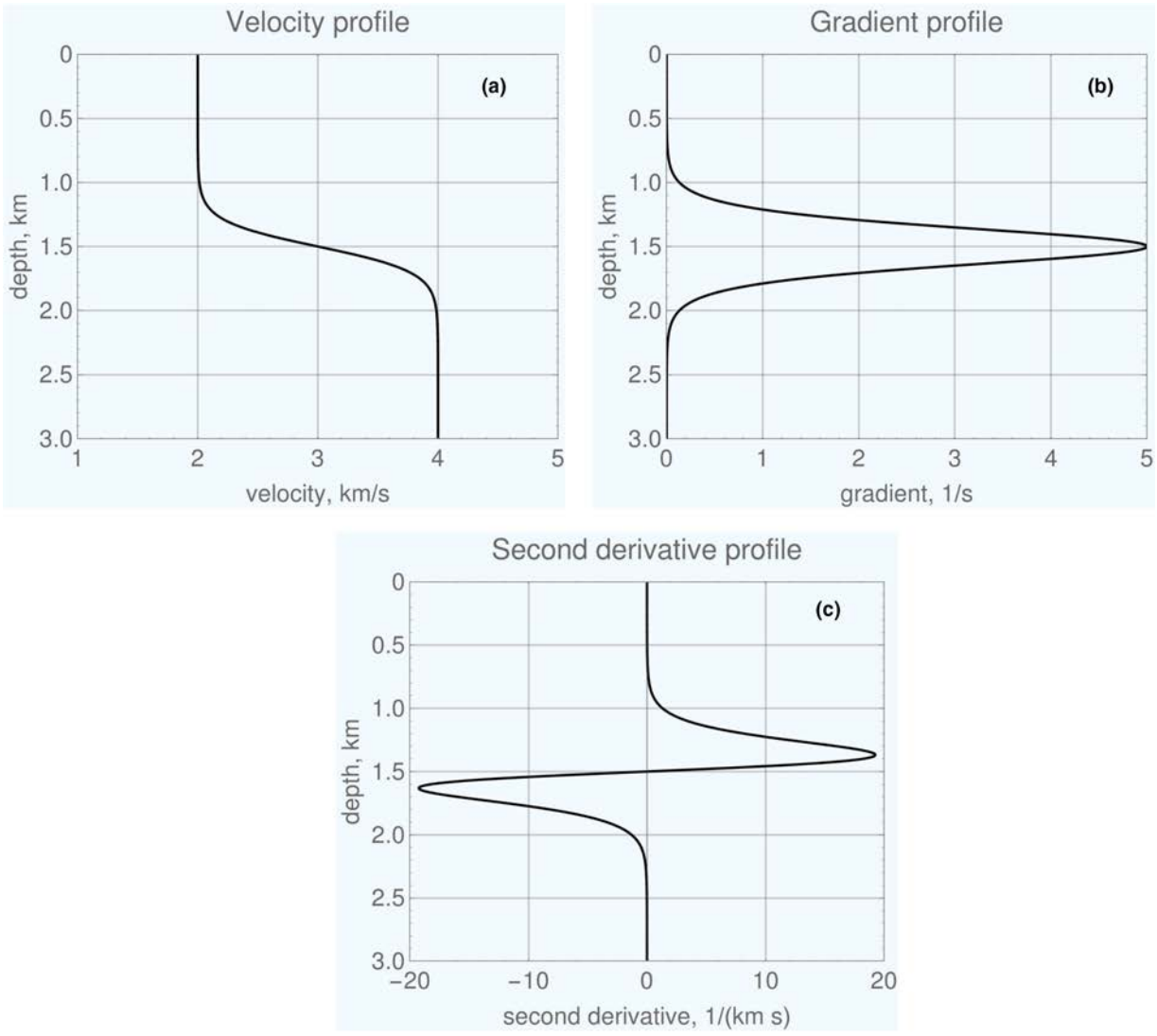

Figure 7 High-velocity half-space under a constant velocity layer: (a) vertical velocity profile, (b) vertical velocity gradient and (c) second vertical derivative of the velocity.

transmission solution is the straight line connecting the endpoints. The initial paths for the shallow and deep solutions are segments of a rectangle with rounded corners, whose parametric equation reads

$$x_1(\tau_u) = a_e|\cos\tau_u|^{1/m}\mathrm{sgn}(\cos\tau_u),$$
$$x_3(\tau_u) = b_e|\sin\tau_u|^{1/m}\mathrm{sgn}(\sin\tau_u), \tag{48}$$

where τ_u is a running (flow) parameter (in radians), a_e and b_e are 'semi-axes', and $m_u \geq 1$ is a real-number parameter ($m_u = 1$ leads to an ellipse; we applied $m_u = 5$). The initial ray direction components are

$$r_1(\tau_u) = -\frac{a_e}{d_u}|\cos\tau_u|^{1/m_u}\sin\tau_u|\sin\tau_u|,$$
$$r_3(\tau_u) = +\frac{b_e}{d_u}|\sin\tau_u|^{1/m_u}\cos\tau_u|\cos\tau_u|, \tag{49}$$
$$d_u(\tau_u) \equiv \sqrt{a_e^2\left(\cos^2\tau_u\right)^{1/m_u}\sin^4\tau_u + b_e^2\left(\sin^2\tau_u\right)^{1/m_u}\cos^4\tau_u}.$$

For the problem we solve, $a_e = 10$ km and $b_e = 6$ km. For the full rectangle, $0 \leq \tau_u \leq 2\pi$ (four quadrants). One can consider a single quadrant for a shallow/deep initial guess.

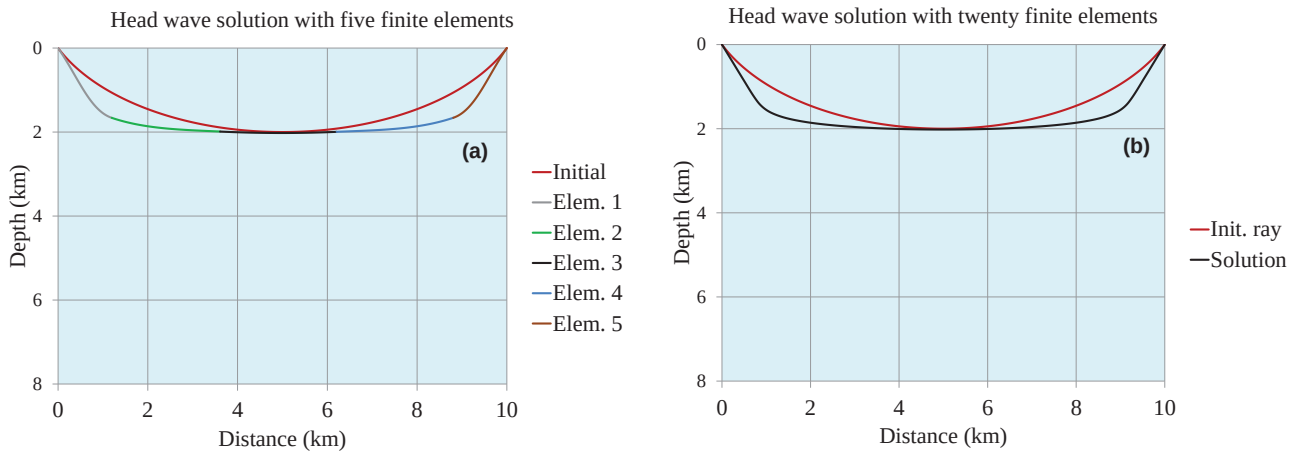

Figure 8 Head wave ray path in the medium with a high-velocity half-space: (a) Eigenray with five finite elements and (b) Eigenray with 20 finite elements.

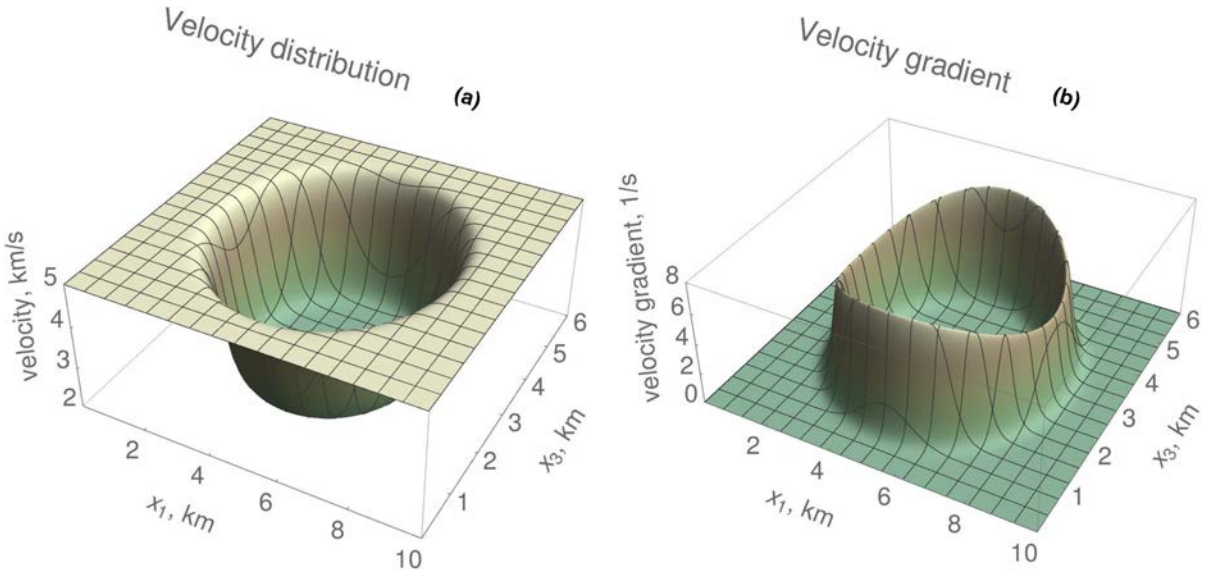

Figure 9 Medium with a low-velocity elliptic anomaly: (a) velocity distribution and (b) absolute value of the velocity gradient.

For the 'shallow' and 'deep' rays (black and green lines), the resulting (minimum) traveltime is $t = 2.61048$ s, and for the 'transmission' ray (blue line), the traveltime is $t = 3.71291$ s. 'Shallow' and 'deep' rays are symmetric solutions, where the ray bypasses the low-velocity inclusion and almost avoids penetration into the transition zone. These rays travel completely through the high-velocity background. For a 'transmission' ray, refraction occurs twice, upon entry to and exit from the low-velocity ellipse. Thus, we deal with a multi-arrival case, characterized by three local minima, two of which are also global.

Example 3: Eigenrays in velocity field with two elliptic anomalies

Consider a model that combines slow- and high-velocity elliptic anomalies and a deep high-velocity half-space. It can be analytically described by

$$v(x_1, x_3) = v_o - \frac{\Delta v}{2}\left(1 - \tanh A_a\right) + \frac{\Delta v}{2}\left(1 - \tanh A_b\right) + \frac{\Delta v_b}{2}\left(1 + \tanh A_c\right), \quad (50)$$

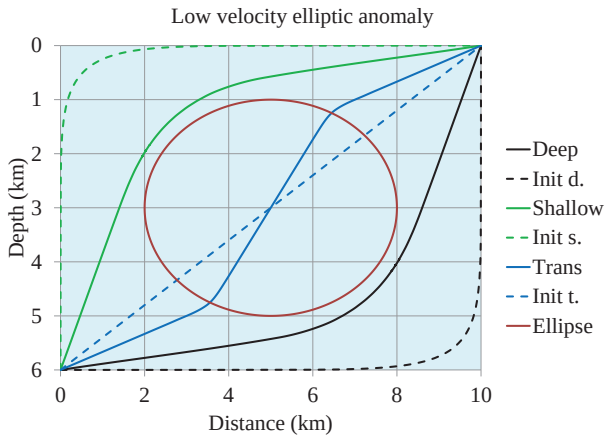

Figure 10 One-way path in a medium with low-velocity elliptic anomaly: Multi-arrival, corresponding to the shallow, deep and transmission solutions.

where

$$A_a = \frac{1}{s_e}\left[\frac{(x_1-c_a)^2}{a_1^2}+\frac{(x_3-c_3)^2}{a_3^2}-1\right],$$
$$A_b = \frac{1}{s_e}\left[\frac{(x_1-c_b)^2}{a_1^2}+\frac{(x_3-c_3)^2}{a_3^2}-1\right], \qquad A_c \equiv \frac{x_3-x_t}{\Delta x_o}. \tag{51}$$

Parameters c_a and c_b are horizontal coordinates of central points of elliptic anomalies, c_3 is their common vertical coordinate, x_t is the floor depth of the high-velocity half-space. Parameters s_e and Δx_o govern the extent of the transition zones,

$$v_o = 4\,\text{km/s}, \Delta v = 2\,\text{km/s}, \Delta v_b = 3\,\text{km/s}, c_a = 4\ \text{km},$$
$$c_b = 20\ \text{km}, c_3 = 2\ \text{km}, a_1 = 4\,\text{km}, a_3 = 1\,\text{km}, \tag{52}$$
$$x_t = 4\,\text{km}, \Delta x_o = 0.2\,\text{km}, s_e = 0.2,$$

and the offset $h = 22$ km. The velocity distribution and absolute value of the velocity gradient are shown in Figure 11(a, b), respectively. Figure 12 shows the Eigenray traveltime minimization results. Note that the scales in the horizontal and vertical directions are different (the horizontal space is much longer than shown). The grey and red ellipses show the contours of the low- and high-velocity anomalies, respectively. The mid-level of the transition zone between the layer with anomalies and high-velocity half-space is 4 km. The three dashed lines show initial paths, and the corresponding solid lines are stationary trajectories (local minima). All three solutions bypass the low-velocity anomaly and penetrate the high-velocity anomaly. Two initial guesses, shown by dashed black and dashed blue lines, represent elliptic arcs. The black line solution bypasses the low-velocity anomaly from the right. It has a shortest path, but not the least traveltime (not the global minimum), because this path does not reach the high-velocity half-space. The blue line solution bypasses the low-velocity anomaly from the left. Therefore, its path is longer, but it reaches the high-velocity half-space and thus arrives within a shorter time. The red line shows one more solution.

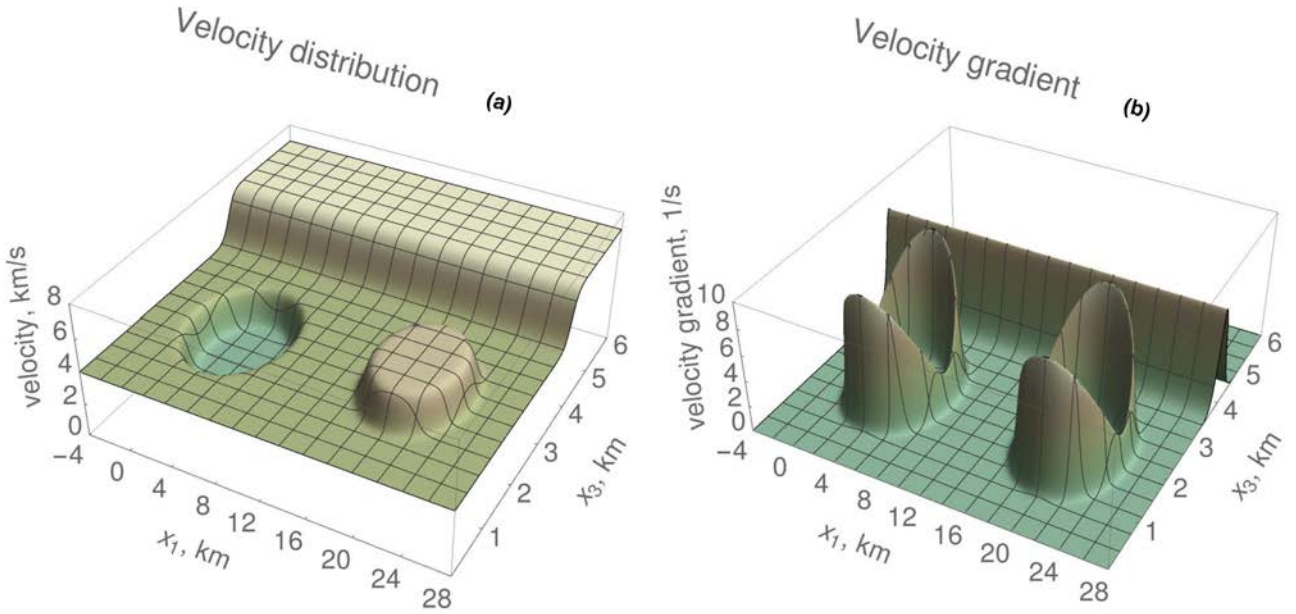

Figure 11 Medium with two elliptic anomalies and a high-velocity half-space: (a) velocity distribution and (b) absolute value of the velocity gradient.

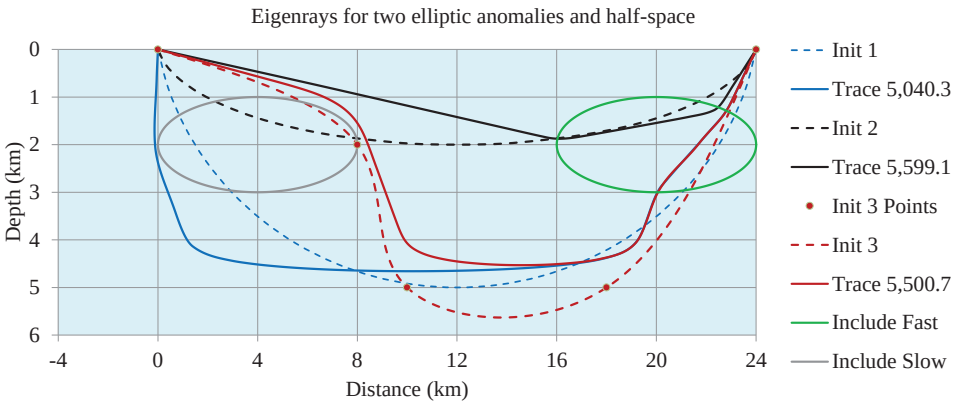

Figure 12 Eigenrays in a medium with two elliptic anomalies and a half-space.

Its initial guess was obtained with the control points shown by bold red points. There are five red points in total, but only three of them are internal/independent, the other two are endpoints of the trajectory (source and receiver). Using the five points, a standard cubic spline is created, shown by a dashed red line. The corresponding solution bypasses the low-velocity anomaly from the right and reaches the high-velocity half-space, but the portion of the path run in the half-space is shorter than that of the 'blue' solution; therefore, its traveltime is longer than that of the 'blue', but shorter than that of the 'black'. The real numbers in the legend of Figure 12 show the traveltimes of the three stationary paths in milliseconds.

DISCUSSION ON THE NUMBER OF FLOATING-POINT OPERATIONS

The computational time increases almost linearly with the number of nodes. The reason is that at each iteration we solve a linearized equation set with the global Hessian of the target function (which is the most time consumptive operation). The global Hessian is a narrow-band matrix, because the elements are connected only consequently at the joints. As mentioned by Koren and Ravve (2020c), due to the soft constraints, the band width becomes $n_b = 18$, for both two- and three-node finite elements, independent of the total number of elements. Note that the total number of degrees of freedom (DoF) is $m = 6(N+1)$, where $N+1$ is the number of the nodes (enumerated from zero to N). The global Hessian is a symmetric matrix. To solve the equation set, we apply Cholesky decomposition, where the number of floating-point operations is, $k = mn_b^2/3$. Since the band width of the matrix, n_b, is constant, the computational time of a single iteration depends almost linearly on the number of the DoF, m. The number of the optimization iterations, needed to reach a stationary point, may be assumed independent of (or only weakly dependent on) the number of the ray path nodes.

CONCLUSIONS

The kinematic Eigenray method, presented in this study, has been developed to solve two-point ray tracing problems in 3D smooth heterogeneous general anisotropic elastic media. We first establish the theoretical background of the variational (kinematic) Eigenray method, based on Fermat's principle, for obtaining stationary ray paths between two fixed endpoints. We propose an original arclength-related Lagrangian, depending on both the location and direction of the ray trajectory which allows efficient finite-element implementation. We also provide the corresponding Hamiltonian for the proposed Lagrangian; the two are related by the Legendre transformation. We then derive the second-order, Lagrangian-based, ordinary differential equation for the kinematic ray tracing. This equation is obtained in a convenient form suitable for the weak variational formulation, which is then solved using the finite-element approach with the Hermite interpolation.

Starting with an initial (non-stationary) discretized trajectory between two given endpoints, we construct a target function that includes two essential constraints. It is optimized (minimized) by computing corrections to the spatial locations and directions of the trajectory that yield a stationary traveltime solution. The first constraint is related to the element lengths governing the locations (distribution) of the nodes along the stationary ray, where a higher local curvature of the path leads to a denser grid. Without this constraint, the resolving equation set is under-defined. The second constraint is the normalization of the length of the ray (group) velocity direction at each node. Explicit expressions for the traveltime and its first and second derivatives allow the implementation

of the Newton method for the optimization, where we search for a minimum traveltime, and the target function includes explicitly the traveltime. In cases where the ray trajectory includes caustics, we search for a saddle-point stationary solution using the gradient method where the target function includes the traveltime gradient squared. The target function that includes the traveltime gradient squared is also used in cases where the type of the stationary solution is not known ahead.

Finally, we propose an original and efficient algorithm for computing the geometrical spreading of the entire stationary ray, applying the global (all-node) traveltime Hessian, already computed along the resolved stationary ray path. We first reduce (compress) the global traveltime Hessian to the endpoint (source–receiver) spatial traveltime Hessian, which makes it possible to apply the known technique using its off-diagonal mixed-derivative sub-block. The proposed method is very efficient; however, it does not explicitly deliver the important information about possible caustics along the ray, their location and their type (point or line). Moreover, it does not deliver the geometric spreading between the source and intermediate points along the ray path. Nevertheless, there may be still a class of problems in which the information only about the total geometric spreading is valuable. In this work, it is used for computing a ray path complexity criterion for evaluating the reliability/plausibility of the solution.

As a general strategy for ray tracing in complex subsurface geological media with complex wave phenomena and multi-arrivals (several stationary solutions between the two fixed endpoints), we suggest starting with the ray shooting method, and then use the proposed Eigenray method to obtain valid (acceptable) stationary rays between the source and the receivers located at the remaining shadow zones.

ACKNOWLEDGEMENTS

The authors are grateful to Emerson for the financial and technical support of this study and for the permission to publish its results. The gratitude is extended to Ivan Pšenčík, Einar Iversen, Michael Slawinski, Alexey Stovas, Vladimir Grechka, Vladimir Cheverda, Maxim Protasov, Yuriy Ivanov, Bing Zhou, Victor Pereyra and our colleague Beth Orshalimy, whose valuable remarks helped to improve the content and style of this paper.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to the article as no new data were created or analysed in this study.

ORCID


Igor Ravve https://orcid.org/0000-0001-6905-9236

REFERENCES

Albareda, G., Bofill, J., Moreira, I., Quapp, W. and Rubio-Martinez, J. (2018) Exploring potential energy surfaces with gentlest ascent dynamics in combination with the shrinking dimer method and Newtonian dynamics. *Theoretical Chemistry Accounts*, 137(6), article 73.

Arnold, V. (1989) *Mathematical Methods of Classical Mechanics*, 2nd edition. Berlin/Heidelberg/New York: Springer-Verlag.

Bathe, K. (2014) *Finite Element Procedures in Engineering Analysis*. Prentice-Hall Inc.

Beydoun, W. and Keho, T. (1987) The paraxial ray method. *Geophysics*, 52(12), 1639–1653.

Bleistein, N., Cohen, J. and Stockwell, J. (2001) *Mathematics of Multi-dimensional Seismic Imaging, Migration, and Inversion*. New York: Springer Publishing Co.

Bona, A. and Slawinski, M. (2003) Fermat's principle for seismic rays in elastic media. *Journal of Applied Geophysics*, 54(3–4), 445–451.

Bona, A., Slawinski, M. and Smith, P. (2009) Ray tracing by simulated annealing: bending method. *Geophysics*, 74(2), T25–T32.

Bulant, P. (1996) Two-point ray tracing in 3-D. *Pure and Applied Geophysics*, 148(3–4), 421–447.

Bulant, P. (2002) Sobolev scalar products in the construction of velocity models: application to model Hess and to SEG/EAGE salt model. *Pure and Applied Geophysics*, 159(7–8), 1487–1506.

Bulant, P. and Klimeš, L. (1999) Interpolation of ray-theory travel times within ray cells. *Geophysical Journal International*, 139(2), 273–282.

Burden, R. and Faires, D. (2005) *Numerical Analysis*. Thomson Higher Education, 130–137.

Byun, B., Corrigan, D. and Gaider, J. (1989) Anisotropic velocity analysis for lithology discrimination. *Geophysics*, 54(12), 1566–1574.

Cao, J., Hu, J. and Wang, H. (2017) Traveltime computation in TI media using Fermat's principle fast marching. EAGE 79th Conference and Technical Exhibition, Expanded Abstract, https://doi.org/10.3997/2214-4609.201700670.

Casasanta, L., Drufuca, G., Andreoletti, C. and Panizzardi, J. (2008) 3D anisotropic ray tracing by raypath optimization. SEG International Exposition and 78th Annual Meeting, Expanded Abstract, 2161–2165.

Červený, V. (2000) *Seismic Ray Theory*. Cambridge University Press.

Červený, V. (2002a) Fermat's variational principle for anisotropic inhomogeneous media. *Studia Geophysica et Geodaetica*, 46(3), 567–588.

Červený, V. (2002b) Fermat's variational principle for anisotropic inhomogeneous media, SW3D (seismic waves in complex 3D structures), report 11, 211–236, http://sw3d.mff.cuni.cz/papers/r11vc1.htm.

Červený, V., Popov, M. and Pšenčík, I. (1982) Computation of wave fields in inhomogeneous media – Gaussian beam approach. *Geophysical Journal International*, 70(1), 109–128.

Chapman, C. (2004) *Fundamentals of Seismic Wave Propagation*. Cambridge University Press.

Chapman, C. and Drummond, R. (1982) Body-wave seismograms in inhomogeneous media using Maslov asymptotic theory. Bulletin of Seismological Society of America, 72, S227–S317.

Cores, D., Fung, G. and Michelena, R. (2000) A fast and global two-point low-storage optimization technique for tracing rays in 2D and 3D isotropic media. *Journal of Applied Geophysics*, 45(4), 273–287.

Dellinger, J. (1991) *Anisotropic seismic wave propagation*: PhD Thesis, Stanford University, Stanford, CA (http://sepwww.stanford.edu/theses/sep69/).

Ecoublet, P., Singh, S., Chapman, C. and Jackson, G. (2002) Bent-ray traveltime tomography and migration without ray racing. *Geophysical Journal International*, 149(3), 633–645.

Ettrich, N. and Gajewski, D. (1996) Wave front construction in smooth media for prestack depth migration. *Pure and Applied Geophysics*, 148(3–4), 481–502.

Farra, V. (1992) Bending method revisited: a Hamiltonian approach. *Geophysical Journal International*, 109(1), 138–150.

Farra, V. (1993) Ray tracing in complex media. *Journal of Applied Geophysics*, 136(1–2), 55–73.

Farra, V. (2004) Improved first-order approximation of group velocities in weakly anisotropic media. *Studia Geophysica et Geodaetica*, 48, 199–213.

Farra, V. and Pšenčík, I. (2013) Moveout approximations for P- and SV waves in VTI media. *Geophysics*, 78(5), WC81–WC92.

Farra, V., Virieux, J. and Madariaga, R. (1989) Ray perturbation for interfaces. *Geophysical Journal International*, 99(2), 377–390.

Fedorov, F. (1968) *Theory of Elastic Waves in Crystals*. Plenum Press: New York.

Galerkin, B. (1915) On electrical circuits for the approximate solution of the Laplace equation. *Vestnik Inzheneroff*, 19, 897–908 (in Russian).

Gao, W., Leng, J. and Zhou, X. (2015) An iterative minimization formulation for saddle point search. *SIAM Journal on Numerical Analysis*, 53(4), 786–1805.

Gelfand, M. and Fomin, S. (2000) *Calculus of Variations*. Reprinted by Dover, originally published by Prentice Hall, New Jersey, 1963.

Gibson, R. (1999) *Ray tracing by wavefront construction in 3-D, anisotropic media: Eos transactions*, American Geophysical Union, 80, F696.

Gibson, R., Sena, A. and Toksöz, M. (1991) Paraxial ray tracing in 3d inhomogeneous, anisotropic media. *Geophysical Prospecting*, 39(4), 473–504.

Gjøystdal, H., Iversen, E., Lecomte, I., Vinje, V. and Åstebøl, K. (2002) Review of ray theory applications in modeling and imaging of seismic data. *Studia Geophysica et Geodaetica*, 46(2), 113–164.

Goldin, S. (1986) *Seismic Traveltime Inversion*. Tulsa: SEG. ISBN 978-0-93183-038-9.

Grechka, V. and McMechan, G. (1996) 3D two-point ray tracing for heterogeneous, weakly transversely isotropic media. *Geophysics*, 61(6), 1883–1894.

Grechka, V. (2017) Algebraic degree of a general group velocity surface. *Geophysics*, 82(4), WA45–Wa53.

Hildebrand, F. (1987) *Introduction to Numerical Analysis*. Dover.

Hovem, J. and Dong, H. (2019) Understanding ocean acoustics by Eigenray analysis. *Journal of Marine Science and Engineering*, 7(4), paper 118, 1–12. https://doi.org/10.3390/jmse7040118.

Huang, X., West, G. and Kendall, J. (1998) A Maslov–Kirchhoff seismogram method. *Geophysical Journal International*, 132(3), 595–602.

Hughes, T. (2000) *The Finite Element Method. Linear Static and Dynamic Finite Element Analysis*. Prentice-Hall Inc.

Julian, B. and Gubbins, D. (1977) Three-dimensional seismic ray tracing. *Journal of Geophysics*, 43, 95–113.

Koren, Z. and Ravve, I. (2018a) Eigenray tracing in 3D heterogeneous media. EAGE 80th Conference and Technical Exhibition, Expanded Abstract, https://doi.org/10.3997/2214-4609.201801325.

Koren, Z. and Ravve, I. (2018b) Eigenray Tracing in 3D heterogeneous anisotropic media using finite element method. 18th International Workshop on Seismic Anisotropy, Extended Abstracts.

Koren, Z. and Ravve, I. (2020a) Eigenrays in 3D heterogeneous anisotropic media. Part I – Kinematics, variational formulation: arXiv:2003.09406.

Koren, Z. and Ravve, I. (2020b) Eigenrays in 3D heterogeneous anisotropic media. Part II – Kinematics, validation of the Lagrangian. arXiv:2003. 09407.

Koren, Z. and Ravve, I. (2020c) Eigenrays in 3D heterogeneous anisotropic media. Part III – Kinematics, finite-element implementation. arXiv:2003.09408.

Koren, Z. and Ravve, I. (2020d) Eigenray in 3D heterogeneous general anisotropic media: kinematics. EAGE 82nd Conference and Technical Exhibition, Expanded Abstract.

Krebes, D. (2019) *Seismic Wave Theory*. Cambridge University Press.

Kumar, D., Sen, M. and Ferguson, R. (2004) Traveltime calculation and prestack depth migration in tilted transversely isotropic media. *Geophysics*, 69(1), 37–44.

Lai, H., Gibson, R. and Lee, K. (2009) Quasi-shear wave ray tracing by wavefront construction in 3-D, anisotropic media. *Journal of Applied Geophysics*, 69(2), 82–95.

Lambaré, G., Lucio, P. and Hanyga, A. (1996) Two-dimensional multivalued traveltime and amplitude maps by uniform sampling of a ray field. *Geophysical Journal International*, 125(2), 584–598.

Leidenfrost, A., Ettrich, N., Gajewski, D. and Kosloff, D. (1999) Comparison of six different methods for calculating travel times. *Geophysical Prospecting*, 47(3), 269–237.

Li, Z., Ji, B. and Zhou, J. (2019) A local minimax method using virtual geometric objects: Part I—For finding saddles. *SIAM Journal of Scientific Computing*, 78(1), 202–225.

Li, Z. and Zhou, J. ((2019) A local minimax method using virtual geometric objects: Part II—For finding equality constrained saddles. *SIAM Journal of Scientific Computing*, 78(1), 226–245.

Lucio, P., Lambaré, G. and Hanyga, A. (1996) 3D multidimensional travel time and amplitude maps. *Pure and Applied Geophysics*, 148(3–4), 449–479.

Moser, T.J. (1991) Shortest path calculation of seismic rays. *Geophysics*, 56(1), 59–67.

Moser, T.J., Nolet, G. and Snieder, R. (1992) Ray bending revisited. *Bulletin of the Seismological Society of America*, 82(1), 259–288.

Musgrave, M. (1954) On the propagation of elastic waves in aeolotropic media. I. General principles. *Proceedings of the Royal Society of London*, A226, 339–355, https://doi.org/10.1098/rspa.1954.0258.

Musgrave, M. (1970) *Crystal Acoustics: Holden-Day*, 2nd edition. Acoustical Society of America.

Pereyra, V. (1988) Numerical methods for inverse problems in three-dimensional geophysical modeling. *Applied Numerical Mathematics*, 4, 97–139.

Pereyra, V. (1992) Two-point ray tracing in general 3D media. *Geophysical Prospecting*, 40(3), 267–287.

Pereyra, V. (1996) Modeling, inversion and block nonlinear traveltime inversion in 3D. *Pure and Applied Geophysics*, 148(3–4), 345–386.

Pereyra, V., Lee, W. and Keller, H. (1980) Solving two-point seismic ray-tracing problems in heterogeneous medium. *Bulletin of the Seismological Society of America*, 70(1), 79–99.

Popov, M. (1982) A new method of computation of wave fields using Gaussian beams. *Wave Motion*, 4(1), 85–97.

Popov, M. and Pšenčík, I. (1978) Computation of ray amplitudes in inhomogeneous media with curved interfaces. *Studia Geophysica et Geodaetica*, 22, 248–258.

Pšenčík, I. and Vavryčuk, V. (2002) Approximate relation between the ray vector and wave normal directions in weakly anisotropic media. *Studia Geophysica et Geodaetica*, 46, 793–807.

Ravve, I. and Koren, Z. (2019) Directional derivatives of ray velocity in anisotropic elastic media. *Geophysical Journal International*, 216(2), 859–895.

Ravve, I. and Koren, Z. (2020a) Eigenrays in 3D heterogeneous anisotropic media: Part IV – Geometric spreading from traveltime Hessian. arXiv:2003.10242.

Ravve, I. and Koren, Z. (2020b) Eigenrays in 3D heterogeneous anisotropic media. Part V – Dynamics, variational formulation. arXiv:2003.10243.

Ravve, I. and Koren, Z. (2020c) Eigenrays in 3D heterogeneous anisotropic media. Part VI – Dynamics, Lagrangian vs. Hamiltonian approaches. arXiv:2003.10244.

Ravve, I. and Koren, Z. (2020d) Eigenrays in 3D heterogeneous anisotropic media. Part VII – Dynamics, finite-element implementation. arXiv:2003.11418.

Ravve, I. and Koren, Z. (2020e) Eigenray in 3D heterogeneous general anisotropic media: dynamics: EAGE 82nd Conference and Technical Exhibition, Expanded Abstract.

Ravve, I. and Koren, Z. (2021) Eigenrays in 3D heterogeneous anisotropic media, Part II: Dynamics. *Geophysical Prospecting*, 69, 28–52.

Rawlinson, N., Hauser, J. and Sambridge, M. (2008) Seismic ray tracing and wavefront tracking in laterally heterogeneous media. *Advances in Geophysics*, 49, 203–273.

Reddy, J. (2004) *An Introduction to the Finite Element Method*: McGraw-Hill Mechanical Engineering.

Ren, W. and Vanden-Eijnden, E. (2013) A climbing string method for saddle point search. *The Journal of Chemical Physics*, 138(13), 134105.

Roganov, Y. and Stovas, A. (2010) On shear-wave triplications in a multilayered transversely isotropic medium with vertical symmetry axis. *Geophysical Prospecting*, 58, 549–559.

Schleicher, J., Tygel, M. and Hubral, P. (2007) *Seismic True-Amplitude Imaging. Geophysical Developments no. 12*. Society of Exploration Geophysicists.

Segerlind, L. (1984) *Applied Finite Element Analysis*, 2nd edition. John Wiley & Sons.

Schoenberg, M. and Daley, T. (2003) qSV wavefront triplication in transversely isotropic material. SEG International Exposition and 73rd Annual Meeting, Expanded Abstract, 137–140.

Shashidhar, N. and Anand, G. (1995) Eigenray tracing in an ocean using Fermat's principle. *Journal of Sound and Vibration*, 186(2), 231–243.

Slawinski, M. (2015) *Waves and Rays in Elastic Continua*. World Scientific Publishing.

Smith, M., Julian, B., Engdahl, E., Gubbins, D. and Gross, R. (1979) Linearized inversion of traveltimes for three-dimensional earth structure, Abstract. *Eos Transactions of American Geophysical Union*, 59, 12.

Snieder, R. and Spencer, C. (1993) A unified approach to ray bending, ray perturbation and paraxial ray theories. *Geophysical Journal International*, 115, 456–470.

Sommerfeld, A. (1964) *Optics: Lectures on Theoretical Physics, 4*. Academic Press.

Song, L.P. and Every, A.G. (2000) Approximate formulae for acoustic wavegroup slownesses in weakly orthorhombic media. *Journal of Physics D: Applied Physics*, 33, L81–L85, https://doi.org/10.1088/0022-3727/33/17/101.

Sripanich, Y.S. and Fomel, S. (2014) Two-point seismic ray tracing in layered media using bending. SEG International Exposition and 84th Annual Meeting, Expanded Abstract, 453–457.

Stovas, A. (2016) Vertical on-axis triplications in orthorhombic media. *Journal of Geophysics and Engineering*, 13(6), 875–879.

Stovas, A., Roganov, Y. and Roganov, V. (2020) Geometrical characteristics of phase and group velocity surfaces in anisotropic media. *Geophysical Prospecting*, 68, Early view. https://doi.org/10.1111/1365-2478.13030.

Strahilevitz, R., Kosloff, D. and Koren, Z. (1998) Three-dimensional two-point ray tracing using paraxial rays in Cartesian coordinates. SEG International Exposition and 68th Annual Meeting, Expanded Abstract, 1887–1892.

Thomsen, L. and Dellinger, J. (2003) On shear-wave triplications in transversely isotropic media. *Journal of Applied Geophysics*, 54, 289–296.

Thomson, C. (1983) Ray-theoretical amplitude inversion for laterally varying velocity structure below NORSAR. *Geophysical Journal International*, 74(2), 525–558.

Thomson, C. (1989) Corrections for grazing rays to 2-D seismic modeling. *Geophysical Journal International*, 96(3), 415–446.

Thomson, C. and Chapman, C. (1985) An introduction to Maslov's asymptotic method. *Geophysical Journal International*, 83(1), 143–168.

Thomson, C. and Gubbins, D. (1982) Three-dimensional atmospheric modeling at NORSAR: linearity of the method and amplitude variations from the anomalies. *Geophysical Journal International*, 71(1), 1–36.

Thurber, C. and Kissling, E. (2000) *Advances in Travel-Time Calculations for 3-D Structures: "Advances in Seismic Event Location"*, 71–99. Kluwer Academic Publishers.

Tsvankin, I. and Grechka, V. (2011) *Seismology of Azimuthally Anisotropic Media and Seismic Fracture Characterization*. SEG.

Um, J. and Thurber, C. (1987) A fast algorithm for two-point seismic ray tracing. *Bulletin of the Seismological Society of America*, 77(3), 972–986.

Vanelle, C. and Gajewski, D. (2003) Determination of geometrical spreading from traveltimes. *Journal of Applied Geophysics*, 54(3), 391–400.

Vanelle, C. and Gajewski, D. (2013) True-amplitude Kirchhoff depth migration in anisotropic media: the traveltime-based approach. *Geophysics*, 78(5), WC33–WC39.

Vavryčuk, V. (2003) Generation of triplication on transversely isotropic media. *Physical Review B*, 68, 054107.

Vavryčuk, V. (2006) Calculation of the slowness vector from the ray vector in anisotropic media. *Proceedings of the Royal Society, Series A*, 462, 883–896.

Vavryčuk, V. (2007) Ray velocity and ray attenuation in homogeneous anisotropic viscoelastic media. *Geophysics*, 72(6), D119–D127.

Vavryčuk, V. (2008a) Velocity, attenuation, and quality factor in anisotropic viscoelastic media: a perturbation approach. *Geophysics*, 73(5), D63–D73.

Vavryčuk, V. (2008b) Real ray tracing in anisotropic viscoelastic media. *Geophysical Journal International*, 175, 617–626.

Vavryčuk, V. (2012) On numerically solving the complex eikonal equation using real ray-tracing methods: a comparison with the exact analytical solution. *Geophysics*(4), 77, T109–T116.

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

Virieux, J., Farra, V. and Madariaga, R. (1988) Ray tracing for earthquake location in laterally heterogeneous media. *Journal of Geophysical Research*, 93(B6), 6585–6599.

Waltham, D. (1988) Two-point ray tracing using Fermat's principle. *Geophysical Journal International*, 93(3), 575–582.

Wesson, R. (1971) Travel-time inversion for laterally inhomogeneous crustal velocity models. *Bulletin of Seismological Society of America*, 61(3), 729–746.

Westwood, E. and Vidmar, P. (1987) Eigenray finding and time series simulation in a layered-bottom ocean. *The Journal of the Acoustical Society of America*, 81(4), 912–924.

Wong, J. (2010) Fermat's principle and ray tracing in anisotropic layered media. *CREWES Research Report*, 22, 1–8.

Wu, C., Wang, H., Hu, J., Luo, F. and Xu, P. (2019) Nonlinear optimal stacking based on shortest path ray tracing for enhancing pre-stack seismic data. EAGE 80th Conference and Technical Exhibition, Expanded Abstract. https://doi.org/10.3997/2214-4609.201900847.

Xu, S. and Stovas, A. (2018) Triplications on traveltime surface for pure and converted wave modes in elastic orthorhombic media. *Geophysical Journal International*, 215(1), 677–694.

Xu, S., Stovas, A. and Mikada, H. (2020) Triplications for the converted wave in transversely isotropic media with a tilted symmetry axis. *Geophysical Prospecting*, 68, 1126–1138.

Xu, S., Stovas, A., Mikada, H. and Takekawa, J. (2021) On-axis triplications in elastic orthorhombic media. *Geophysical Journal International*, 224, 449–467.

Xu, T., Wu, C. and Wu, Z. (2014) Seismic traveltime inversion of 3D velocity model with triangulated interfaces. *Earthquake Science*, 27(2), 127–136.

Yong, J. and Cheng, F. (2004) Geometric Hermite curves with minimum strain energy. *Computer Aided Geometric Design*, 21, 281–301

Zhang, J. and Du, Q. (2012a) Shrinking dimer dynamics and its application for saddle point search. *SIAM Journal on Numerical Analysis*, 50(4), 1899–1921.

Zhang, J. and Du, Q. (2012b) Constrained shrinking dimer dynamics for saddle point search with constraints. *Journal of Computational Physics*, 231(14), 4745–4758.

Zhang, L. and Zhou, B. (2018) Calculation of slowness vectors from ray directions for qP-, qSV-, and qSH-waves in tilted transversely isotropic media. *Geophysics*, 83(4), C153–C160.

Zhao, A., Zhang, Z. and Teng, J. (2004) Minimum travel time tree algorithm for seismic ray tracing: improvements in efficiency. *Journal of Geophysics and Engineering*, 1, 245–251.

Zhou, B. and Greenhalgh, S. (2005) 'Shortest path' ray tracing for most general 2D/3D anisotropic media. *Journal of Geophysics and Engineering*, 2, 54–63.

Zhou, B. and Greenhalgh, S. (2006) Raypath and traveltime computations for 2D transversely isotropic media with dipping symmetry axes. *Exploration Geophysics*, 37, 150–159.

Zienkiewicz, O., Taylor, R. and Zhu, J. (2013) *The Finite Element Method, Its Basis and Fundamentals*. Elsevier Ltd.