Seismic ray method: Recent developments

Vlastislav Červený, Luděk Klimeš
Department of Geophysics, Faculty of Mathematics and Physics, Charles University, Ke Karlovu 3, 121 16 Praha 2, Czech Republic,
E-mails: vcerveny@seis.karlov.mff.cuni.cz, klimes@seis.karlov.mff.cuni.cz

Ivan Pšenčík
Geophysical Institute, Academy of Sciences of Czech Republic,
Boční II 1401, 141 31 Praha 4, Czech Republic, E-mail: ip@ig.cas.cz

Abstract

The seismic ray method has found broad applications in the numerical calculation of seismic wavefields in complex 3–D, isotropic and anisotropic, laterally varying layered structures and in the solution of forward and inverse problems of seismology and seismic exploration for oil. This article outlines the basic features of the seismic ray method, and reviews its possibilities and recent extensions. Considerable attention is devoted to ray tracing and dynamic ray tracing of S waves in anisotropic inhomogeneous media, to the coupling ray theory for S waves in such media, to the summation of Gaussian beams and packets, and to the selection of models suitable for ray tracing. The article also points to areas of the seismic ray method that are still open to further research.

Submitted for publication in book Wu, R-S. & Maupin, V. (eds.): *Advances in Wave Propagation in Heterogeneous Media.*
In: Seismic Waves in Complex 3-D Structures, Report 15 (Department of Geophysics, Charles University, Prague, 2005), pp. 057–172

Contents

1. Introduction

The seismic ray method is based on the asymptotic high–frequency solution of the elastodynamic equation. It consists in the decomposition of the wavefield into independent contributions called elementary waves, which propagate along rays. These elementary waves may represent various seismic body waves propagating in isotropic or anisotropic inhomogeneous, layered and block structures, such as direct, reflected, converted or multiply reflected/transmitted waves. The great advantage of the ray method is that individual elementary waves can be treated separately.

Elementary waves are specified by their travel times $\tau(x_i)$ and vectorial complex–valued amplitudes $\mathbf{U}(x_i)$, as functions of Cartesian coordinates x_i. In the zero–order approximation of the ray method, the displacement vector $\mathbf{u}(x_i,t)$ of an arbitrary time–harmonic elementary wave can be expressed in terms of $\tau(x_i)$ and $\mathbf{U}(x_i)$ as follows:

$$\mathbf{u}(x_i,t) = \mathbf{U}(x_i)\exp[-\mathrm{i}\omega(t-\tau(x_i))] \quad . \tag{1}$$

Here t is the time, and ω is the circular frequency. Both $\mathbf{U}(x_i)$ and $\tau(x_i)$ are assumed to be frequency independent.

Travel time $\tau(x_i)$ satisfies a nonlinear partial differential equation of the first order, called the eikonal equation. The explicit form of the eikonal equation is obtained directly from the elastodynamic equation. The eikonal equation is usually solved for $\tau(x_i)$ by the method of characteristics. The characteristics of the eikonal equation define the rays. Mathematically, the characteristics of the eikonal equations are represented by a system of ordinary differential equations of the first order, called the ray tracing system, see Section 2. Conventional numerical methods can be used to solve this system with a required accuracy, so that the computation of rays (as an initial value problem) is not difficult. In addition to the ray trajectory, we also obtain the travel times $\tau(x_i)$ and the Cartesian components $p_i = \partial\tau/\partial x_i$ of slowness vector $\mathbf{p}$ at each point of the ray. The procedure of solving the ray tracing system is called ray tracing, and the relevant travel times are called ray–theory travel times. At interfaces, the ray tracing system must be supplemented by Snell's law, see Section 8. For more details on initial–value ray tracing see Section 2.

A quantity which plays an important role in the seismic ray method is the polarization vector $\mathbf{g}$ of the wave under consideration, see Section 3. The polarization vector is very useful particularly in anisotropic inhomogeneous media, where it can be suitably applied even in ray tracing algorithms. It plays a fundamental role in amplitude studies both in isotropic and anisotropic inhomogeneous media.

In certain applications, the initial conditions for rays are not known. Instead, some other conditions are known at the initial and end points of the ray. For example, the ray is specified by the positions of its initial and end points. We then speak of a two–point ray and two–point ray tracing. In comparison with initial–value ray tracing, two–point ray tracing is considerably more complicated, particularly in 3–D layered/block models. An important role in two–point ray tracing and in some other applications is played by the so–called ray histories. For more details on two–point ray tracing and ray histories, see Section 4.

The vectorial complex–valued amplitudes $\mathbf{U}(x_i)$ can in most cases be expressed as the product of a unit real–valued polarization vector $\mathbf{g}$ and a scalar complex–valued amplitude $A(x_i)$. The scalar complex–valued amplitude $A(x_i)$ satisfies a linear partial differential equation, called the transport equation. The explicit form of the transport equation is obtained directly from the elastodynamic equation. The transport equation can be simply solved along rays. The transport equation then reduces to an ordinary differential equation of the first order, which can be solved analytically in terms of geometrical spreading. Geometrical spreading is mostly calculated along the ray using a procedure called dynamic ray tracing, see Section 5.2. Dynamic ray tracing is a very powerful procedure, which is based on the solution of a system of linear ordinary differential equations of the first order along the ray. As the system is linear, its solution is easy. Mostly, the system is expressed either in global Cartesian coordinates or in ray–centred coordinates, see Section 5.1. Using dynamic ray tracing, it is also possible to determine the DRT (dynamic ray tracing) propagator matrix along the ray, see Section 5.3. The DRT propagator matrix can be used not only to calculate geometrical spreading along the ray, but also in many other applications. For example, it can be used to determine approximately ray–theory travel times and other relevant quantities not only along a ray, but also in its "quadratic" vicinity. We then speak of the paraxial approximation of the ray method (shortly paraxial ray method), paraxial rays, paraxial travel times, etc., see Sections 5.4–5.6. To obtain the final equations for vectorial complex amplitudes, or for the Green function, we must also determine the amplitude contributions at the initial point of the ray (for example, the radiation pattern of the source), at the end point of the ray (for example, the conversion coefficients at the receiver), and at the points of reflection/transmission (R/T coefficients). See Section 6.

Once the travel time $\tau(x_i)$ and the vectorial complex–valued amplitude $\mathbf{U}(x_i)$ at a given spatial point x_i are known, the particle ground motion for an arbitrary high–frequency signal can be easily computed using the Fourier transform, see Section 7. Finally, considering receiver points distributed along some profile and selected elementary waves, we obtain the whole ray–theory seismogram section.

A simple procedure of solving approximately both direct and inverse kinematic problems of inhomogeneous, isotropic and anisotropic, media is based on the perturbation theory. See Section 9 for a description of basic principles of perturbation theory for travel times.

For finite frequencies, the ray method is only approximate. Its accuracy, however, is sufficient to solve many important 3–D wave propagation problems, which can hardly be treated by any other means. Various extensions of the ray method, which increase its accuracy and can be applied in situations where the standard ray method is not applicable, have been proposed. Many of them use ray fields as a framework. Several these extensions are described or reviewed to a greater or smaller extent in Sections 10–12. Section 10 is devoted to the coupling ray theory for S waves. This theory plays a fundamental role in the propagation of S waves in weakly anisotropic media, and in anisotropic media close to S–wave singularities. A detailed treatment of Gaussian beams and Gaussian packets can be found in Section 11. The summation of Gaussian beams and Gaussian packets overcomes the difficulties of the ray theory with caustics. Some other extensions of the ray method and some related methods are briefly described in Section 12.

The ray method can be applied to isotropic, anisotropic and weakly anisotropic 3–D structures, with arbitrary 3–D variations of elastic moduli and density, with curved interfaces, for arbitrary source–receiver configurations, and for very general types of seismic body waves. It is, however, limited to smooth models (smooth distribution of elastic moduli and density inside layers and blocks, smooth interfaces). For more details on the construction and smoothing of 3–D models for the application of ray methods, see Section 13.

Unless otherwise stated, we consider only perfectly elastic media. Moreover, we do not consider inhomogeneous waves, with complex–valued travel times and complex–valued rays. The waves in viscoelastic models are still the subject of research.

The seismic ray method has been described in great detail in several textbooks and papers. See, for example, Babich and Buldyrev (1972), Červený, Molotkov, & Pšenčík (1977), Hanyga, Lenartowicz & Pajchel (1984), Bleistein (1984), Virieux (1996), Dahlen and Tromp (1998), Bleistein, Cohen & Stockwell (2001), Červený (2001), Chapman (2002, 2004). Therefore, the well–known properties of the ray method are presented here without detailed derivation, just for the sake of completeness. This concerns mostly Sections 2–8, for which the book by Červený (2001) can be considered a general reference. More space is, however, devoted to several recently intensively studied, important concepts such as the coupling ray theory or seismic wave modelling with the use of Gaussian beams, see Sections 9–11 and 13.

Similarly, the paper does not contain a systematic and complete bibliography of the subject. This would exceed considerably the admissible extent of the paper. Exceptions are sections 10, 11 and 13, which contain more extensive lists of recent references. More references on the ray theory itself can be found in the books and papers mentioned in the above paragraph.

The ray method has also been used in other branches of physics, particularly in the electromagnetic theory. Although we consider the elastic waves described by the *elastodynamic equation*, most results are applicable to any non–dissipative hyperbolic second–order partial differential equation (wave equation), including the system of Maxwell equations. Most wave propagation problems discussed here can thus be applied with minor modifications in the electromagnetic theory. The basic principles of the summation of Gaussian beams and packets are independent of a particular selection of the wave equation. References relating to the electromagnetic ray theory, however, are given here only exceptionally. For general references, see the books by Babich & Buldyrev (1972), Felsen & Markuwitz (1973), Kravtsov & Orlov (1980).

To express the equations in this paper in concise form, we mostly use the componental notation for vectors and matrices, and exceptionally the matrix notation. In the componental notation, the upper–case indices $(A, B, C, ...)$ take the values of 1 and 2, the lower–case indices $(a, b, c, ...)$ the values 1, 2 and 3. In Section 9, the Greek lower–case subscripts $(\alpha, \beta, \gamma, ...)$ index the model parameters. In Section 11, the Greek lower–case subscripts $(\alpha, \beta, \gamma, ...)$ index the space–time coordinates, and take the values 1, 2, 3 and 4. The Einstein summation convention is used throughout the paper.

For more information on the seismic ray method and for additional references see also the WWW pages of the Consortium Project "Seismic Waves in Complex 3–D Structures" at "http://sw3d.mff.cuni.cz".

2. Seismic rays and travel times. Initial–value ray tracing

The eikonal equation is a nonlinear partial differential equation of the first order for travel time $\tau(x_i)$, and can be expressed in Hamiltonian form

$$\mathcal{H}(x_i, p_i) = 0 \quad , \tag{2}$$

where $\mathcal{H}$ is the Hamiltonian or Hamilton function, x_i are Cartesian coordinates and $p_i = \partial\tau/\partial x_i$ are Cartesian components of slowness vector $\mathbf{p}$. We consider Hamiltonians which are homogeneous functions of the second degree in p_i (with a possible additional constant). For example, for elementary waves propagating in an isotropic inhomogeneous medium, we use the Hamiltonian

$$\mathcal{H}(x_i, p_i) = \tfrac{1}{2}[V^2(x_i)p_i p_i - 1] \quad , \tag{3}$$

where $V(x_i)$ is the spatially variable velocity of P or S waves. In anisotropic inhomogeneous media, we use the Hamiltonian given by the relation

$$\mathcal{H}(x_i, p_i) = \tfrac{1}{2}[G_m(x_i, p_i) - 1] \quad . \tag{4}$$

We also use the averaged Hamiltonian

$$\mathcal{H}(x_i, p_i) = \tfrac{1}{4}[G_1(x_i, p_i) + G_2(x_i, p_i) - 2] \tag{5}$$

(Bakker, 2002) of both S–waves, which enables the *common S wave rays* to be traced. These rays are frequently used in the coupling ray theory, see Section 10. Here $G_m(x_i, p_i)$, $m = 1, 2, 3$, is an eigenvalue of the 3×3 matrix

$$\Gamma_{ik}(x_n, p_n) = a_{ijkl}p_j p_l \quad , \tag{6}$$

and $a_{ijkl}(x_n)$ is the tensor of density–normalized elastic moduli.

In the seismic ray method, the 3×3 matrix $\Gamma_{ik}(x_n, p_n)$ is usually called the Christoffel matrix, although it does not strictly correspond to the common definition of the Christoffel matrix, in which the slowness vector $\mathbf{p}$ in (6) is replaced by a real–valued unit vector $\mathbf{n}$, oriented in the direction of $\mathbf{p}$. Vector $\mathbf{n}$ is called the wavefront normal.

Consequently, the three eigenvalues G_m can be calculated by solving the appropriate algebraic equation of the third degree. We call the elementary waves, corresponding to eigenvalues G_1, G_2 and G_3, the S1, S2 and P waves, respectively. In isotropic media, eigenvalues G_1, G_2, G_3 can be expressed analytically:

$$G_1(x_i, p_i) = G_2(x_i, p_i) = \beta^2(x_i)p_n p_n \quad , \qquad G_3(x_i, p_i) = \alpha^2(x_i)p_n p_n \quad . \tag{7}$$

Here α and β are the velocities of P and S waves, propagating in isotropic inhomogeneous media. Thus, the Hamiltonian (3) for isotropic media is a special case of the Hamiltonian (4) for anisotropic media, in which the eigenvalues corresponding to S1 and S2 waves coincide.

It follows from (2)–(4) that the eikonal equations for individual elementary waves can also be expressed in the following form:

$$G_m(x_i, p_i) = 1 \quad , \tag{8}$$

both for isotropic and anisotropic media.

Once we know eigenvalue $G_m(x_i, p_i)$ for m specified, and if this eigenvalue is different from the two remaining eigenvalues, we can also uniquely determine the eigenvector

$\mathbf{g}^{(m)}$ of the Christoffel matrix (6). Eigenvector $\mathbf{g}^{(m)}$ then specifies the polarization vector of the elementary wave under consideration, see Section 3. If, however, eigenvalue G_m coincides with one of the two remaining eigenvalues, eigenvector $\mathbf{g}^{(m)}$ cannot be uniquely determined. We then speak of the degenerate case and degenerate Christoffel matrix. The two elementary waves are coupled, and the polarization of the resulting wave is elliptical.

In realistic cases, the P–wave eigenvalue G_3 is well separated from the S–wave eigenvalues G_1 and G_2. This means that the degenerate case does not exist for P waves. The algorithms proposed for P waves in anisotropic media can be used quite universally also in isotropic and weakly anisotropic media.

For S waves, the situation is considerably more complex. There are two different degenerate cases for S waves: (a) Eigenvalues G_1 and G_2 coincide *locally* along certain lines or at certain points in anisotropic media. We then speak of S–wave singularities. (b) Eigenvalues G_1 and G_2 coincide *globally* in isotropic media, see (7). In both these degenerate cases, the S1 and S2 waves are coupled, locally or globally, and propagate as a single wave. The algorithms for S waves propagating in anisotropic inhomogeneous media cannot be used for S waves propagating in isotropic inhomogeneous media. Consequently, there are two main ray theories for S waves: the first for the computation of S waves in anisotropic inhomogeneous media, based on equation (4), and called the anisotropic ray theory for S waves, and the second for the computation of S waves in isotropic inhomogeneous media, based on equation (3), and called the isotropic ray theory for S waves. In weakly anisotropic inhomogeneous media, none of these ray theories is valid for S waves, and a different approach, based on the coupling ray theory for S waves, must be used. For the coupling ray theory for S waves, see Section 10.

The eikonal equation can be solved by the method of characteristics, which represent rays. The parameter, specifying the points along the characteristics, depends on the selected Hamiltonian. For Hamiltonians which are homogeneous functions of the second degree in p_i (with a possible additional constant), the parameter has the physical meaning of travel time τ. Consequently, we consider consistently with equations (3), (4) and (5) travel time τ to be the parameter along the ray. The characteristic of the eikonal equation is then described by six ordinary differential equations in x_i and p_i,

$$\mathrm{d}x_i/\mathrm{d}\tau = \partial\mathcal{H}/\partial p_i \quad , \qquad \mathrm{d}p_i/\mathrm{d}\tau = -\partial\mathcal{H}/\partial x_i \quad . \tag{9}$$

Thus, travel time τ is a parameter along the characteristic curve, and need not be computed additionally by quadratures. The determination of the derivatives of the Hamiltonian in ray tracing system (9) may require the calculation of the eigenvectors of the Christoffel matrix. Whereas there are no problems with Hamiltonian (4) for a P wave and with Hamiltonians (3) or (5), the determination of Hamiltonian (4) for an S wave may be difficult. Refer to Section 3 for more details.

The system of six ordinary differential equations (9) represents the ray tracing system (also sometimes called the kinematic ray tracing system, to distinguish it from the dynamic ray tracing system). If the initial conditions $x_i = x_{i0}$ and $p_i = p_{i0}$ are specified for $\tau = \tau_0$ in such a way that $\mathcal{H}(x_{i0}, p_{i0}) = 0$, ray tracing system (9) can be solved by conventional numerical methods (predictor–corrector, for example). We then speak of initial–value ray tracing and of initial–value rays. Relation (2) must be satisfied along the whole ray and can be used to check and control the accuracy of ray tracing.

For isotropic inhomogeneous media, the ray tracing system reads, see (3) and (9),

$$\mathrm{d}x_i/\mathrm{d}\tau = V^2 p_i \quad , \qquad \mathrm{d}p_i/\mathrm{d}\tau = -\tfrac{1}{2} p_k p_k \partial V^2/\partial x_i \quad . \tag{10}$$

As $p_k p_k = 1/V^2$ along the ray, see (7) and (8), the R.H.S. of the second equation of (10) can also be replaced by $-\partial \mathrm{ln} V/\partial x_i$. Isotropic ray tracing system (10) can be used quite safely for both P and S waves, even if two rays with different initial conditions intersect (multipathing) or are tangent (caustics) at some points of the rays.

Ray tracing system (10) for isotropic inhomogeneous media has been used to compute rays in 2-D ray tracing computer packages SEIS, see Červený & Pšenčík (1984), and in 3–D ray tracing computer package CRT (complete ray tracing), see Červený, Klimeš & Pšenčík (1988).

For anisotropic inhomogeneous media, the ray tracing system can be expressed as follows:

$$\frac{\mathrm{d}x_i}{\mathrm{d}\tau} = a_{ijkl} p_l g_j^{(m)} g_k^{(m)} \quad , \qquad \frac{\mathrm{d}p_i}{\mathrm{d}\tau} = -\tfrac{1}{2}\, \frac{\partial a_{jkln}}{\partial x_i} p_k p_n g_j^{(m)} g_l^{(m)} \quad . \tag{11}$$

Here $g_i^{(m)}$ is the eigenvector of the Christoffel matrix corresponding to the considered wave. No summation over (m) is applied. Ray tracing system (11) has found useful applications in anisotropic inhomogeneous media.

Ray tracing system (11) can also be expressed in an alternative form. We define subdeterminants

$$D_{ij} = \tfrac{1}{2} \epsilon_{ikl} \epsilon_{jrs} (\Gamma_{kr} - \delta_{kr})(\Gamma_{ls} - \delta_{ls}) \tag{12}$$

of the Christoffel matrix. Here ϵ_{ijk} is the Levi–Civitta symbol ($\epsilon_{123} = \epsilon_{312} = \epsilon_{231} = 1$, $\epsilon_{321} = \epsilon_{213} = \epsilon_{132} = -1, \epsilon_{ijk} = 0$ otherwise). It is not difficult to prove that, for a selected elementary wave for which $G_m = 1$, the factor D_{ij}/D_{ss}, defined by (12), can be expressed for $D_{ss} \neq 0$ in terms of the eigenvector of the Christoffel matrix corresponding to the considered wave as follows:

$$D_{ij}/D_{ss} = g_i^{(m)} g_j^{(m)} \quad . \tag{13}$$

No summation over (m) is applied. Using this relation, the anisotropic ray tracing system (11) yields, for D_{ss} sufficiently different from zero, ray tracing system

$$\frac{\mathrm{d}x_i}{\mathrm{d}\tau} = a_{ijkl} p_l D_{jk}/D_{ss} \quad , \qquad \frac{\mathrm{d}p_i}{\mathrm{d}\tau} = -\tfrac{1}{2}\, \frac{\partial a_{jkln}}{\partial x_i} p_k p_n D_{jl}/D_{ss} \quad . \tag{14}$$

The mode of the wave we wish to compute is specified by the initial conditions x_{i0}, p_{i0}, which must satisfy equation (2) at the initial point, corresponding to the relevant wave. Anisotropic ray tracing system (14) is used for 3–D computations in the computer package ANRAY (Gajewski & Pšenčík, 1987).

Anisotropic ray tracing system (11) or (14) can be used quite universally for P waves, even for P waves in isotropic and weakly anisotropic inhomogeneous media. For S1 and S2 waves, however, the situation is more complicated. The anisotropic ray tracing systems do not work universally for these waves, but may fail in degenerate cases. Ray tracing system (14) may fail in the vicinity of S–wave singularities (including isotropic media), where $D_{ss} \approx 0$.

Let us first consider an elementary wave, corresponding to eigenvalue G_m of the Christoffel matrix (6), well separated from the other two eigenvalues. In this *non–degenerate case*, we may identify eigenvalue G_m by its unit value, or we may identify the corresponding eigenvector $\mathbf{g}^{(m)}$ by its continuity along the ray. Eigenvector $\mathbf{g}^{(m)}$ can then be inserted into ray tracing equations (11). In this case, ray tracing equations (11) may be replaced by ray tracing equations (14) without considerable loss of accuracy.

In the *vicinity of S–wave singularities*, eigenvalue G_m of the Christoffel matrix (6) cannot be safely identified by its unit value. We must thus identify eigenvector $\mathbf{g}^{(M)}$ by means of its continuity along the ray (Vavryčuk, 1999, sec. 5.1.1), if we are performing ray tracing in an anisotropic medium. In this case, we should use equations (11) for ray tracing and avoid equations (14).

If the segment of the ray in the vicinity of the S–wave singularity is sufficiently short, we may simply compare the eigenvectors in front of and beyond the vicinity of the S–wave singularity (Vavryčuk, 1999, sec. 5.1.2; Bulant & Klimeš, 2002, eqs. 18 and 37), and possibly interpolate the eigenvectors.

If the segment of the ray in the vicinity of the S–wave singularity (singular region) is too long, we must extend vectors $\mathbf{g}^{(M)}$ along the ray so that they do not rotate about vector $\mathbf{g}^{(3)}$,

$$\frac{\mathrm{d}g_i^{(1)}}{\mathrm{d}\tau} g_i^{(2)} = 0 \quad . \tag{15}$$

For this purpose, we may use, e.g., equation

$$\frac{\mathrm{d}g_i^{(M)}}{\mathrm{d}\tau} = \frac{g_i^{(3)}}{G_M - G_3}\, g_j^{(3)} \left(\frac{\partial \Gamma_{jk}}{\partial x_l} \mathcal{U}_l + \frac{\partial \Gamma_{jk}}{\partial p_l} \eta_l \right) g_k^{(M)} \quad , \tag{16}$$

where no summation over $M = 1, 2$ is applied. We then compare the eigenvectors determined in front of the singular region and continued through the singular region with the eigenvectors determined beyond the singular region. If the difference between the eigenvectors is not negligible, we should terminate tracing the ray. This may happen, e.g., in the vicinity of a kiss singularity (Rümpker & Thomson, 1994).

Solving the ray tracing system (10) or (11) or (14) for given initial conditions x_{i0} and p_{i0}, we obtain not only the ray trajectory (specified by $x_i(\tau)$), but also slowness vector $\mathbf{p}(\tau)$ (perpendicular to the wavefront) at any point of the ray. We also obtain two other important vectors $\boldsymbol{\mathcal{U}}$ and $\boldsymbol{\eta}$ at any point of the ray,

$$\boldsymbol{\mathcal{U}}(\tau) = \mathrm{d}\mathbf{x}/\mathrm{d}\tau \quad , \qquad \boldsymbol{\eta}(\tau) = \mathrm{d}\mathbf{p}/\mathrm{d}\tau \quad . \tag{17}$$

Explicit expressions for $\boldsymbol{\mathcal{U}}$ and $\boldsymbol{\eta}$ are given by the R.H.S.'s of (10), (11), or (14). Physically, $\boldsymbol{\mathcal{U}}$ is the ray velocity vector, tangent to the ray, with the magnitude equal to velocity $\mathcal{U}$ of propagation along the ray. This vector has the direction of the time–averaged energy flux. In the seismological literature, $\boldsymbol{\mathcal{U}}$ is also often called the group velocity vector. In isotropic media, $\boldsymbol{\mathcal{U}}$ and $\mathbf{p}$ are parallel, but not in anisotropic media. In other words, the time–averaged energy flux is not perpendicular to the wavefronts in anisotropic media. Both $\boldsymbol{\mathcal{U}}$ and $\mathbf{p}$ play an important role in the seismic ray method, particularly in dynamic ray tracing, and satisfy the well–known relation $\boldsymbol{\mathcal{U}} \cdot \mathbf{p} = 1$.

In applications, the whole system of rays is usually needed, not just one single ray. Mostly, a two–parameter system of rays, corresponding to a system of wavefronts of a selected elementary wave, propagating in the model under consideration (orthonomic system of rays) is considered. We also call this system the ray field. We denote the two parameters γ_1 and γ_2, and refer to them as the ray parameters. For example, for an elementary wave generated by a point source, the two ray parameters may be chosen as the take–off angles of initial slowness vectors at the source. Similarly, for a wave generated at an initial surface, the ray parameters can be chosen as curvilinear coordinates of initial points of rays along the initial surface. The properties of ray fields, expressed in terms of ray parameters, play a very important role in the seismic ray method and its applications. This applies to the concepts such as the ray tube, ray Jacobian, geometrical spreading, caustics, KMAH index, curvature of the wavefront, shadow zones, etc. The two ray parameters γ_1, γ_2 may be supplemented by a third parameter γ_3, which specifies uniquely the position of any point on the ray. We then speak of ray coordinates $\gamma_1, \gamma_2, \gamma_3$. See Červený (2001, sec. 3.10).

Similarly as isotropic ray tracing system (10), anisotropic ray tracing systems (11) and (14) can be used safely for any wave mode if the rays with different initial values intersect (multipathing) or pass through caustic points.

In media composed of layers or blocks, ray tracing must be supplemented by the formulae describing transformation of quantities, calculated along the ray, on boundaries of layers or blocks. Ray tracing must also be supplemented by the prescription specifying which of the generated waves should be considered after reflection/transmission. We refer to this prescription as the *code* of the elementary wave. Every elementary wave is thus specified by its code. This distinguishes it from other elementary waves.

As mentioned above, the ray method requires sufficiently smooth models. As soon as heterogeneities in the model exceed a certain limit, the rays may have a chaotic character, particularly for large travel times (long rays). For more details on chaotic rays, see Section 13.

3. Polarization vectors

There are two different sorts of polarization vectors, with two distinct applications:

(a) Polarization vectors are required in ray tracing equations (11). These vectors are discussed in Section 2, and may be called *ray–tracing polarization vectors.*

(b) Polarization vectors are used to determine the vectorial amplitude,

$$\mathbf{U} = A\,\mathbf{e} \quad , \tag{18}$$

where A is the complex–valued scalar amplitude and $\mathbf{e}$ is the real–valued unit polarization vector. These polarization vectors are discussed in this section, and should be distinguished from ray–tracing polarization vectors, because the vectorial amplitude need not have the direction of the ray–tracing polarization vector.

3.1. Non–degenerate case

Let us consider an elementary wave, corresponding to eigenvalue G_m of the Christoffel matrix (6), well separated from the other two eigenvalues. The complex–valued amplitude vector $\mathbf{U}$ of the wave under consideration, propagating in an isotropic or anisotropic inhomogeneous medium can then be expressed in the following form:

$$\mathbf{U} = A\,\mathbf{g}^{(m)} \quad , \tag{19}$$

where A is the complex–valued scalar amplitude and $\mathbf{e} = \mathbf{g}^{(m)}$ is the real–valued unit polarization vector. The polarization vector of the elementary wave under consideration can be simply computed as the eigenvector of the Christoffel matrix Γ_{ik} given by (6), corresponding to eigenvalue G_m. At any point of the ray, at which the slowness vector is known and at which the Christoffel matrix is not degenerate, polarization vectors can be determined uniquely by conventional methods. The three polarization vectors $\mathbf{g}^{(1)}, \mathbf{g}^{(2)}$ and $\mathbf{g}^{(3)}$, corresponding to one wave normal, are mutually orthogonal. If the ray under consideration (P, S1 or S2) corresponds to the selected eigenvalue G_m, the polarization vector of the wave propagating along the ray is $\mathbf{g}^{(m)}$.

Equation (19) for the amplitude vector $\mathbf{U}$ can be used only if the eigenvalue G_m, corresponding to the elementary wave under consideration, is different from the others. This applies to P waves propagating in both isotropic and anisotropic media, and to non–degenerate S waves in anisotropic media. In anisotropic media, we speak about the *anisotropic ray theory.* As we can easily see from (19), in this case the elementary wave under consideration is linearly polarized.

3.2. Polarization in the vicinity of S–wave singularities or in isotropic media

For S waves propagating in isotropic media, weakly anisotropic media, and for S waves propagating close to S–wave singularities in anisotropic inhomogeneous media, the complex–valued amplitude vector $\mathbf{U}$ must be considered in the form

$$\mathbf{U} = B\,\mathbf{g}^{(1)} + C\,\mathbf{g}^{(2)} \quad . \tag{20}$$

Here $\mathbf{g}^{(1)}$ and $\mathbf{g}^{(2)}$ are eigenvectors of the Christoffel matrix (6), corresponding to eigenvalues G_1 and G_2, and B and C are two complex–valued scalar quantities. As we can see from (20), the elementary wave under consideration is elliptically polarized if B and C are "out of phase". This applies globally to S waves in isotropic and weakly anisotropic inhomogeneous media, and locally to S waves propagating in anisotropic inhomogeneous media, close to S–wave singularities. In inhomogeneous weakly anisotropic media or in the vicinity of a S–wave singularity, B and C are frequency dependent. They can be determined from two coupled transport equations. More details are given in Section 10.

For S waves propagating in isotropic inhomogeneous media, the two transport equations for the complex–valued scalar amplitudes B and C in (20) are generally coupled. They can be decoupled only for a special choice of variations of vectors $\mathbf{g}^{(1)}$ and $\mathbf{g}^{(2)}$ along the ray. The decoupling condition for isotropic media can be expressed in the form of an ordinary differential equation of the first order, which must be satisfied along the ray:

$$\mathrm{d}\mathbf{g}^{(M)}/\mathrm{d}\tau = (\mathbf{g}^{(M)} \cdot \boldsymbol{\eta})\mathbf{p}/(\mathbf{p} \cdot \mathbf{p}) \quad , \qquad M = 1, 2 \quad . \tag{21}$$

This equation is a special case of equation (16). Here $\boldsymbol{\eta}$ is given by (17) and is known from ray tracing. In the terminology of Riemannian geometry, we say that vector $\mathbf{g}^{(M)}$ satisfying (21) is transported parallely along the ray. If $\mathbf{g}^{(1)}$ and $\mathbf{g}^{(2)}$ are chosen to be mutually perpendicular unit vectors perpendicular to $\mathbf{p}$ at the initial point of the ray, equation (21) guarantees that they have the same property along the whole ray. Consequently, $\mathbf{g}^{(1)}$ and $\mathbf{g}^{(2)}$ are always perpendicular to the ray and do not rotate around the ray as the wave progresses. As $\mathbf{g}^{(1)}$, $\mathbf{g}^{(2)}$ and $\mathbf{p}$ are always orthogonal, it is sufficient to compute just $\mathbf{g}^{(1)}$ (or $\mathbf{g}^{(2)}$) from (21), the second vector $\mathbf{g}^{(2)}$ (or $\mathbf{g}^{(1)}$) can be simply determined using the orthogonality condition.

As the transport equation is decoupled for $\mathbf{g}^{(1)}$ and $\mathbf{g}^{(2)}$ satisfying (21), $\mathbf{g}^{(1)}$ and $\mathbf{g}^{(2)}$ represent the polarization vectors of S waves along the whole ray in an isotropic medium, once they represent them at the initial point of the ray. For this reason, we call vectors $\mathbf{g}^{(1)}$ and $\mathbf{g}^{(2)}$, satisfying (21), the polarization vectors of S waves.

At interfaces between isotropic media, the polarization vectors of reflected or transmitted S waves can be chosen as two arbitrary, mutually perpendicular unit vectors, perpendicular to the ray of a generated wave. The corresponding R/T coefficients, however, must be adjusted to this choice.

Polarization vectors have many important physical applications. They can be used to display the polarization diagrams, which clearly show the vectorial character of the wavefield and play an important role in the determination of the mode of the wave under consideration (P, S1, S2). They have been traditionally used in the investigation of S–wave splitting, etc.

There is one important difference in the behaviour of polarization vectors of S waves in isotropic and anisotropic media. Whereas the polarization vectors of an S wave in an isotropic inhomogeneous medium do not rotate in the plane perpendicular to the ray, in an anisotropic inhomogeneous media they can rotate in the plane perpendicular to the third eigenvector of the Christoffel matrix, as the wave progresses. As the direction of this eigenvector is usually close to the direction of the ray, we can also roughly say that the polarization vectors of S waves rotate around the ray in anisotropic inhomogeneous media. The rotation may be fast, particularly close to S–wave singularities.

4. Ray histories, two–point ray tracing, wavefront tracing

The introduction of ray histories and the division of the ray–parameter domain into regions of equal ray histories brought about a considerable simplification and generalization of the 2–D two–point ray tracing algorithm and enabled a highly accurate and reliable 3–D two–point ray tracing algorithm (Bulant 1996) to be designed. The concept of ray histories has facilitated the introduction of calculating travel times and other quantities on dense rectangular grids of points by interpolation within ray cells (Bulant & Klimeš 1999), which is of crucial importance in prestack seismic migrations and nonlinear determination of seismic hypocentres. The relevant algorithms are simultaneously used in very successful wavefront tracing methods, and allow diffracted edge waves to be calculated. Note that during recent years, ray histories have also become very important in 3–D computer graphics and animation, where, however, the rays are straight, which considerably simplifies the determination of possible boundaries between ray histories compared to seismology.

The definition of ray histories allows the ray–parameter domain to be decomposed into regions of equal ray histories, and these regions to be triangulated. This procedure is, in general, called here *controlled initial–value ray tracing.* Controlled initial–value ray tracing has many applications. It is the basis of boundary–value (two–point) ray tracing, it is applied in wavefront tracing, it serves as the preprocessor for interpolation within ray cells, and it is of principal importance in the asymptotic summation of Gaussian beams or in the asymptotic summation of Gaussian packets.

4.1. Model and ray histories

The travel time of an elementary wave is multivalued with respect to spatial coordinates. It thus cannot globally be defined as the function of spatial coordinates x_i. It has to be parametrized by ray coordinates γ_a, composed of two ray parameters γ_1, γ_2 and of parameter γ_3 along the rays. The spatial distribution of travel time can be described by the projection of ray coordinates onto spatial coordinates and travel time,

$$\gamma_a \longrightarrow x_i(\gamma_a), \tau(\gamma_a) \quad . \tag{22}$$

We may then say that travel time τ is a "continuous function" if both coordinates x_i and travel time τ are continuous functions of the ray coordinates, and that the n–th derivative of travel time τ is a "continuous function" if both the n–th derivatives of coordinates x_i and the n–th derivative of travel time τ are continuous functions of the ray coordinates. We may define the continuity of other quantities calculated along the rays analogously.

If the n–th derivatives of the travel time are to be continuous, the n–th derivatives of the density–normalized elastic moduli (i.e. of the propagation velocity in isotropic media) must also be continuous. Similarly for other quantities calculated along the rays. For example, if the amplitudes are to be continuous, the second derivatives of the density–normalized elastic moduli must also be continuous.

For the calculation of the n–th derivatives of travel time, a *smooth model* (smooth velocity model, smooth macro model) thus means the continuous n–th derivatives of the density–normalized elastic moduli. The common minimum requirement on a smooth model for the zero–order ray theory are the continuous second derivatives of the density–normalized elastic moduli.

The density–normalized elastic moduli may be discontinuous only along "*interfaces*". An interface of the $(n+1)$–th order means that the n–th derivatives of the density–normalized elastic moduli are discontinuous. Interfaces thus divide the model into *blocks* (e.g., a layer or a salt dome). The variation of the density–normalized elastic moduli within each block should be smooth in the same sense as in the smooth model (Gjøystdal, Reinhardsen & Åstebøl, 1985; Červený, Klimeš & Pšenčík, 1988). A model composed of these blocks, which are separated by interfaces, may be called the *block model* (block velocity model, block macro model).

The interfaces are composed of one or several smooth surfaces or of parts thereof. The surfaces forming the interfaces may be defined implicitly, as the zero isosurfaces of given functions. If we wish the n–th derivatives of travel time to be continuous within the block beyond the interface, we require the $(n-m)$–th derivatives of the function describing the interface of the $(m+1)$–th order to be continuous.

Unfortunately, in models with interfaces, not all rays pass through an equal sequence of blocks and interfaces. The continuity or smoothness of the travel time or of other quantities calculated along the rays is then violated. We thus have to introduce, for each elementary wave (specified by the code of the elementary wave), the *ray histories*. Rays of a single ray history pass through an equal sequence of blocks and interfaces. Each sequence of blocks and interfaces encountered during ray tracing thus defines the corresponding ray history. The ray histories may be indexed by integers, and may thus be understood to be the integer functions of ray coordinates (similarly as, e.g., the KMAH index defined in Section 6.3). This integer function of ray coordinates is called the *history function*.

If we perform ray tracing from a given source to a given reference surface (e.g., two–point ray tracing), we are usually interested in the complete ray history from the source to the reference surface. In this case, the ray history depends on two ray parameters γ_A. If the spatial distribution of the results of ray tracing is required (e.g., interpolation within ray cells, or wavefront tracing), the ray history depends on three ray coordinates γ_a, and is determined along the ray gradually, starting from the source. The history function dependent on three ray coordinates γ_a is discontinuous with ray parameters, whose images $x_i(\gamma_a)$ are the points of intersection of rays with interfaces.

Travel time is a smooth function for a constant ray history. At the boundary between ray histories, travel time or its derivatives may be discontinuous. Analogously for the spatial coordinates $x_i(\gamma_a)$ of rays and for many other quantities calculated along the rays. Geometrical spreading may even approach infinity at the boundary between ray histories. Whereas travel time is continuous at the caustics, some other quantities are not (e.g., the KMAH index, amplitude, second derivatives of travel time). If we are interested in the quantities discontinuous at the caustics, we may optionally wish to include also the KMAH index into the definition of ray histories. In this case, the ray history including the KMAH index is determined both by the ray history without the KMAH index and by the KMAH index. The KMAH index is defined in Section 6.3, and dynamic ray tracing (Section 5.2) has to be applied to determine it.

Since the continuity and smoothness is violated at the boundaries between ray histories, neither interpolation, nor paraxial ray approximation, nor perturbation expansion can be applied across these boundaries. That is why the boundaries between ray histories should be carefully determined, and why methods such as interpolation, paraxial ray approximation or perturbation expansion can be applied just within a single ray history. The determination of the boundaries between ray histories is thus the basic and most involved step in applying such methods. Interpolation, paraxial ray approximation or perturbation expansion on their own are usually just relatively easy and routine subsequent steps of the application.

4.2. Two–point ray tracing

We briefly present here the algorithm for determining all two–point rays of a given elementary wave by means of the shooting method, proposed by Bulant (1996). The algorithm is designed for general 3–D models composed of heterogeneous blocks separated by curved interfaces. The algorithm is independent of the type and shape of the source (point source, curved linear source, curved surface source) and of its parametrization.

Solving the *two–point ray tracing* problem means finding all rays of a given elementary wave, which take off from the source and pass through the receiver. Note that we have used the term *two–point ray* in place of the term *boundary–value ray*, which would be more appropriate because the *source* is general and is not limited to a point source. The *receiver* is a given point, situated at the *reference surface*. This receiver need not exactly coincide with the actual receiver, it may represent the projection of the actual receiver onto the reference surface. The actual receivers should be situated in the vicinity of the reference surface. The receivers at the reference surface are parametrized by two reference coordinates ξ_K along the reference surface. The reference coordinates may be general functions of model coordinates x_i, $\xi_K = \xi_K(x_i)$.

Rays taking off from the source are parametrized by ray parameters γ_A. The given 2–D set of all ray parameters (γ_1, γ_2) is the *ray–parameter domain*. To simplify the description of the ray–parameter domain, it is often assumed that the ray–parameter domain is the image of the rectangular *normalized ray–parameter domain* in the given mapping $\gamma_A = \gamma_A(\eta_B)$, projecting *normalized ray parameters* η_B onto ray parameters γ_A. The normalized ray–parameter domain is usually selected as a unit square.

The history function for two–point ray tracing is the integer function of two ray parameters, and describes the complete ray history from the source to the reference surface. The history function for two–point ray tracing should include the KMAH index in order to catch all two–point rays on a triplication of travel time.

Initial–value ray tracing assigns the integer value of the history function to every ray. Initial–value ray tracing also assigns unique coordinates ξ_K of incidence to the rays incident at the reference surface. These rays are referred to as *successful rays*.

For two–point ray tracing, the ray–parameter domain should be divided into regions of the same value of the history function. If defining the ray–parameter domain as the image of the normalized ray–parameter domain, this division is performed in the normalized ray–parameter domain.

The demarcation belts between the different ray histories, corresponding to the numerical uncertainty of the boundaries between the ray histories, must be kept reasonably narrow, because all two–point rays situated inside the demarcation belts will be lost. Once the ray–parameter domain is decomposed, with the given accuracy, into regions of the same value of the history function, the two–point rays may be found iteratively within the individual regions corresponding to successful rays.

The ray–parameter domain is divided into regions of equal value of the history function by means of triangulating the ray–parameter domain. The triangulation is performed iteratively.

First, the ray–parameter domain is covered by *basic triangles*. The size of the basic triangles is measured with respect to the *ray–parameter metric tensor*. The ray–parameter metric tensor may specify the metric in the ray–parameter domain (e.g., the metric on the unit sphere for a point source). This ray–parameter metric tensor must be defined for all ray parameters, including rays, which do not leave the source. This ray–parameter metric tensor should also be reasonably smooth over all ray histories. The basic triangles are roughly equilateral with respect to the ray–parameter metric.

A *homogeneous triangle* is a triangle whose apexes are formed by rays with equal value of the history function. All "inhomogeneous" triangles have to be divided into homogeneous triangles.

In the first step of the division, all *boundary rays* have to be found on the sides of the inhomogeneous triangle. The boundary rays are pairs of rays with very slightly different ray parameters, but with different ray histories. Their maximum distance in the ray–parameter domain is given a priori. The pairs of boundary rays serve to separate the regions of equal value of the history function. The boundary rays are sought using the method of halving intervals, which requires the auxiliary rays, which are not further used in the triangulation, to be traced, see Figure 1. During this step, narrow strips of new ray histories can often be found along the sides of the basic triangles. In this way, several pairs of boundary rays may be found along one side of the triangle.

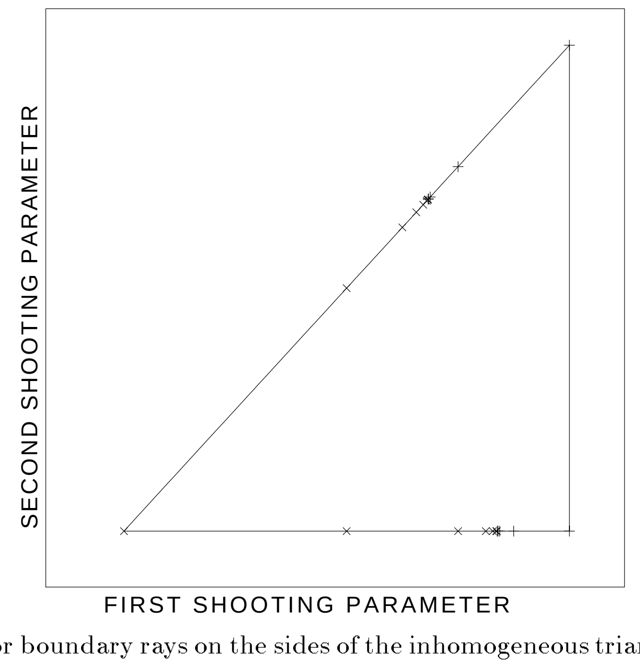

Figure 1. Searching for boundary rays on the sides of the inhomogeneous triangle by halving intervals. Rays are symbol–coded according to their history. The three vertices of the triangle are the basic rays. Note that the distance between the two rays in each pair of boundary rays is very small.

In the second step, the boundary between one ray history and the other ray histories is traced, see Figure 2. For the description of tracing the boundary, refer to Bulant (1996, sec. 3.3). In this way, we obtain homogeneous polygons, which should then be divided into homogeneous triangles.

In the third step, the homogeneous polygons should be divided into homogeneous triangles. The division of homogeneous polygons into triangles, not too different from equilateral, requires tracing several new rays, situated both on the long sides of the homogeneous polygons and inside the homogeneous polygons, see Figure 3.

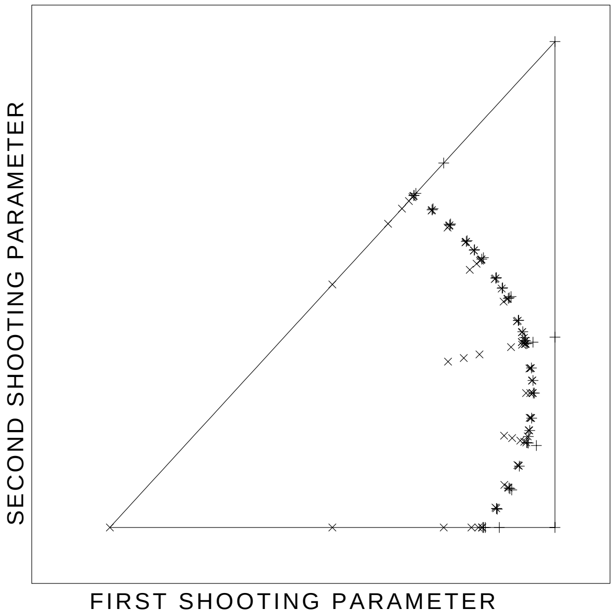

Figure 2. The boundary between the two ray histories is demarcated by pairs of boundary rays. All the boundary and basic rays of the same ray history create a homogeneous polygon. Here are two homogeneous polygons, which will be divided into homogeneous triangles in the next step.

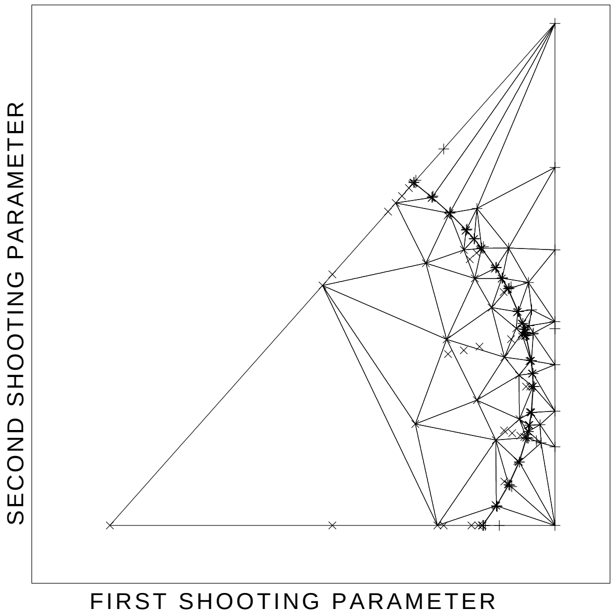

Figure 3. Both homogeneous polygons are covered by homogeneous triangles. Each homogeneous triangle is formed by three rays of equal ray history. Several new rays were traced in order to create homogeneous triangles "not too different" from equilateral.

If a ray of a different ray history is identified in a homogeneous triangle, the ray together with the vertices of the triangle is used to create three new inhomogeneous triangles. The algorithm then reverts to the division of these three triangles into homogeneous triangles.

We would appreciate if the ray–parameter metric tensor could also take into account geometrical spreading or another quantification of the complexity of the system of rays. Unfortunately, this might create considerable discontinuities of the ray–parameter metric tensor at the boundaries between ray histories. That is why we also define the *secondary ray–parameter metric tensor*, which is smooth only within the individual regions of equal value of the history function. The secondary ray–parameter metric tensor is used to measure the size of homogeneous triangles. If a homogeneous triangle is large with respect to the secondary ray–parameter metric tensor, it is divided into smaller triangles. In two–point ray tracing, the secondary ray–parameter metric tensor may take into account, e.g., geometrical spreading at the reference surface, see Figure 4.

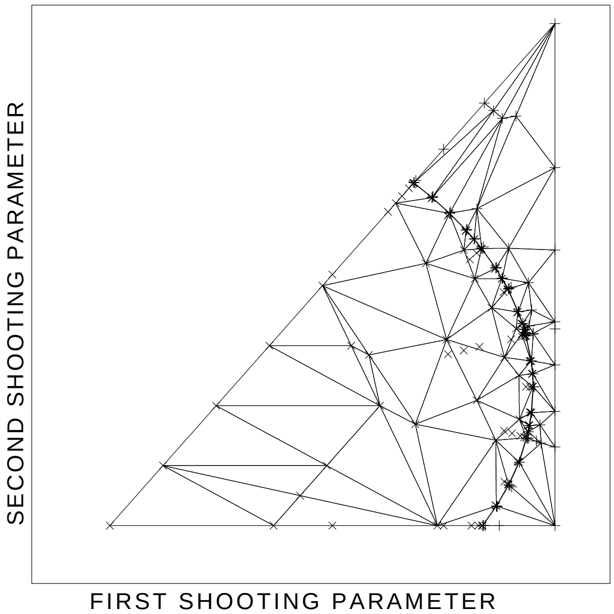

Figure 4. Again the same basic triangle as in Figure 3. Now all the homogeneous triangles corresponding to successful rays, which are too large on the reference surface, are divided into smaller ones. This means that the secondary ray–parameter metric tensor was defined so that it measures the distances between the points of intersection of rays with the reference surface.

For the coverage of the normalized ray–parameter domain by rays of various ray histories, traced in the model depicted in Figure 5, refer to Figure 6. The arrangement of the individual homogeneous subdomains is quite complicated. Some regions of equal ray histories are disconnected, because the rays reflect from disconnected parts of the lower interface. The small separation of neighbouring two–point rays indicates large geometrical spreading.

The two–point rays of the wave diffracted at the edge of a lenticular inclusion, traced using the described algorithm, are displayed in Figure 7.

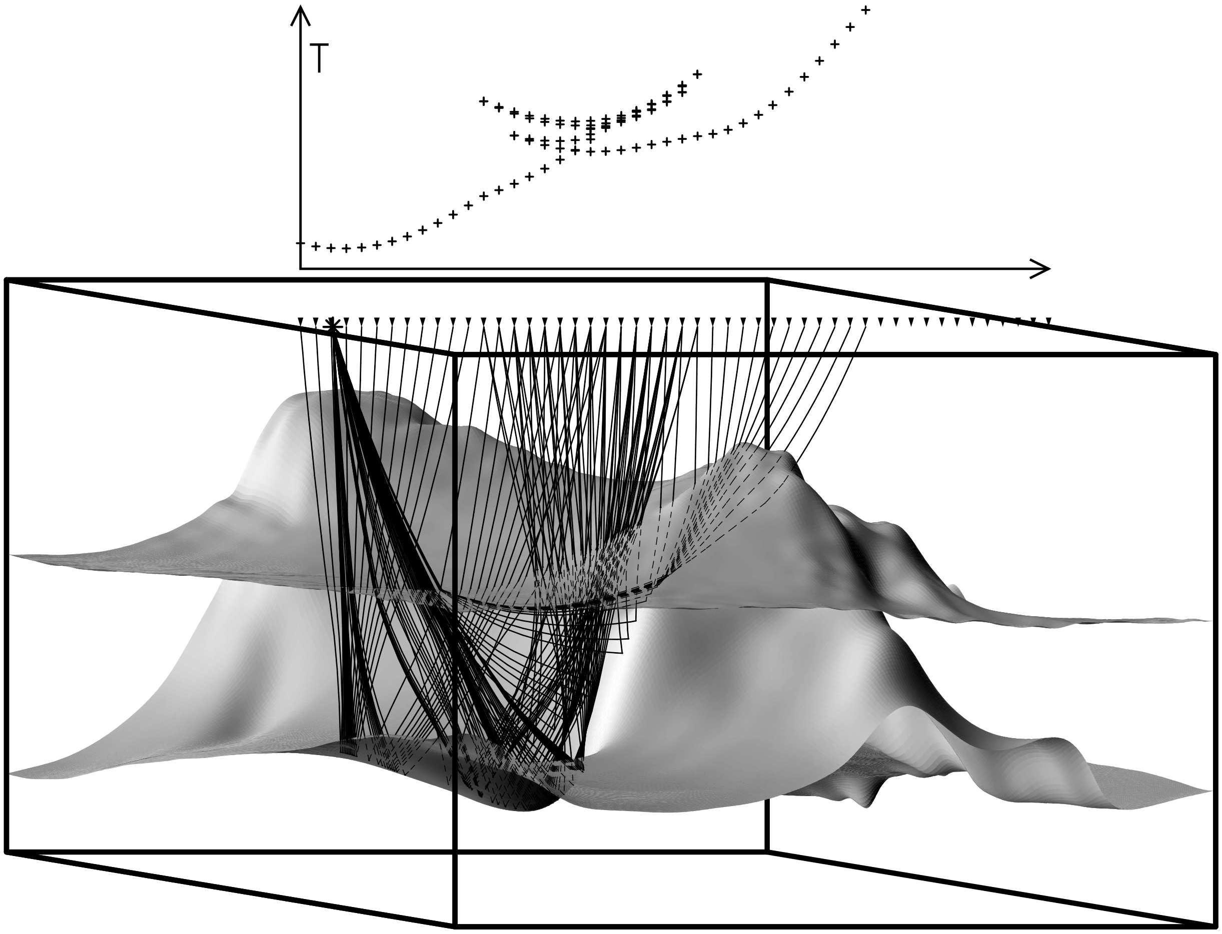

Figure 5. The three–layer model used for two–point ray tracing (Bulant, 1999), with two–point rays reflected from the lower interface. The asterisk represents the point source, the small triangles are the receivers. The time curve above the model demonstrates the corresponding two–point travel times. Only each fourth receiver and each fourth two–point ray indicated in Figure 6 has been plotted.

4.3. Controlled initial–value ray tracing

Controlled initial–value ray tracing is nearly equal to two–point ray tracing.

Controlled initial–value ray tracing for the asymptotic summation of Gaussian beams differs from two–point ray tracing by no receivers specified along the reference surface.

In controlled initial–value ray tracing for the asymptotic summation of Gaussian packets or for the interpolation within ray cells, no reference surface is usually specified. If no reference surface is specified, the history function describes the complete ray history from the source to the endpoint of the ray.

The history function for controlled initial–value ray tracing usually does not include the KMAH index, because the asymptotic summation of Gaussian beams, the asymptotic summation of Gaussian packets, and usually also the interpolation within ray cells do not require sorting of rays according to the KMAH index.

In controlled initial–value ray tracing for the interpolation within ray cells, the secondary ray–parameter metric tensor may take into account geometrical spreading along the whole ray.

In controlled initial–value ray tracing for the asymptotic summation of Gaussian beams, the secondary ray–parameter metric tensor should control the discretization

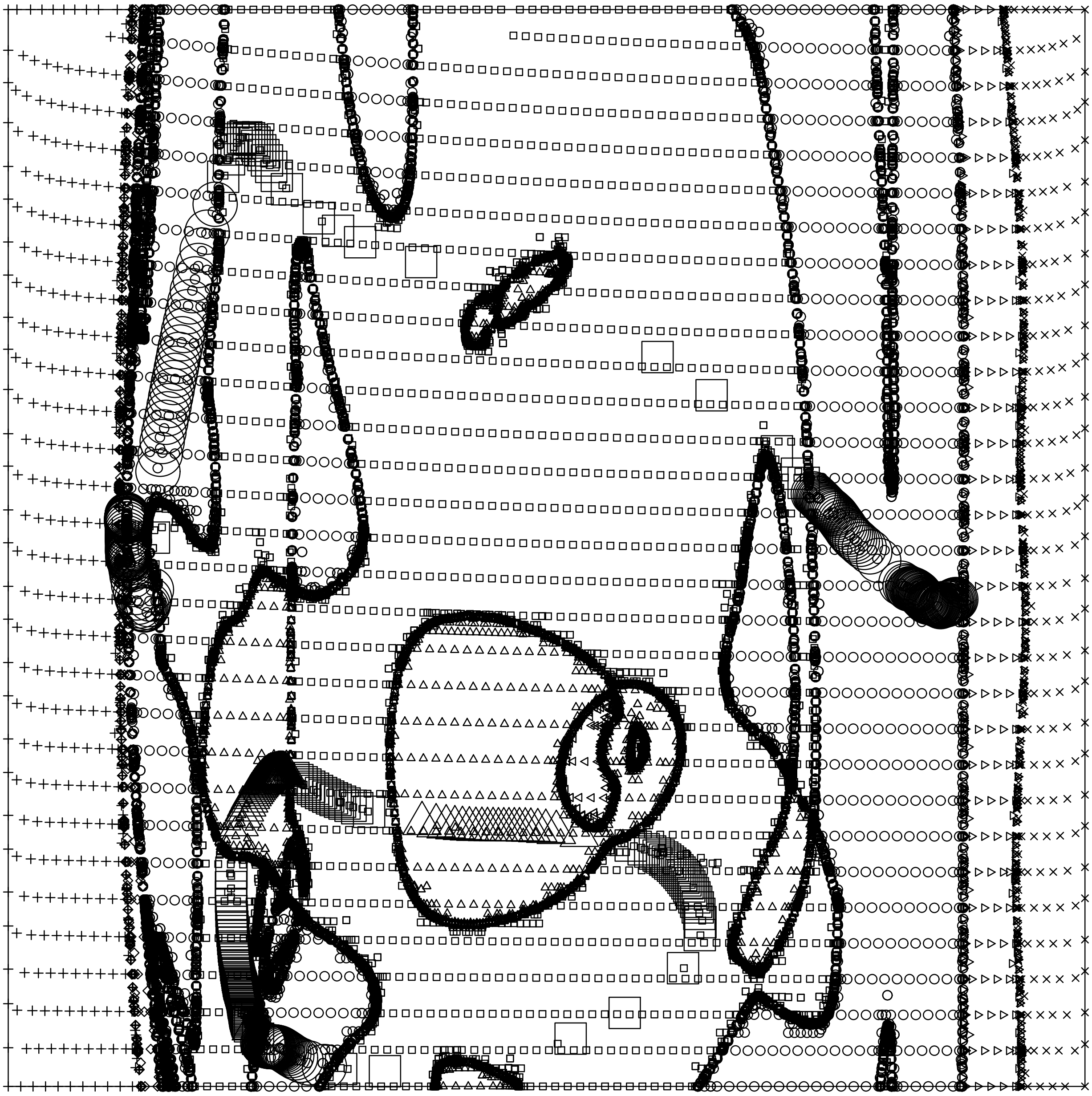

Figure 6. The coverage of the ray–parameter domain by rays shot during the computation. The rays are symbol–coded according to the ray history. Two–point rays are plotted as larger symbols. The basic triangles are roughly equilateral with respect to the ray–parameter metric rather than with respect to the Euclidean metric on this page.

error, see equation (147). The secondary ray–parameter metric tensor then corresponds to the shape of Gaussian beams optimized for the propagation from the source to the reference surface.

In controlled initial–value ray tracing for the asymptotic summation of Gaussian packets, the secondary ray–parameter metric tensor should also be used to control the discretization error. In this case, the secondary ray–parameter metric tensor may take into account metric tensors (147) corresponding to the shapes of Gaussian packets optimized for the propagation from the source to various points along the ray.

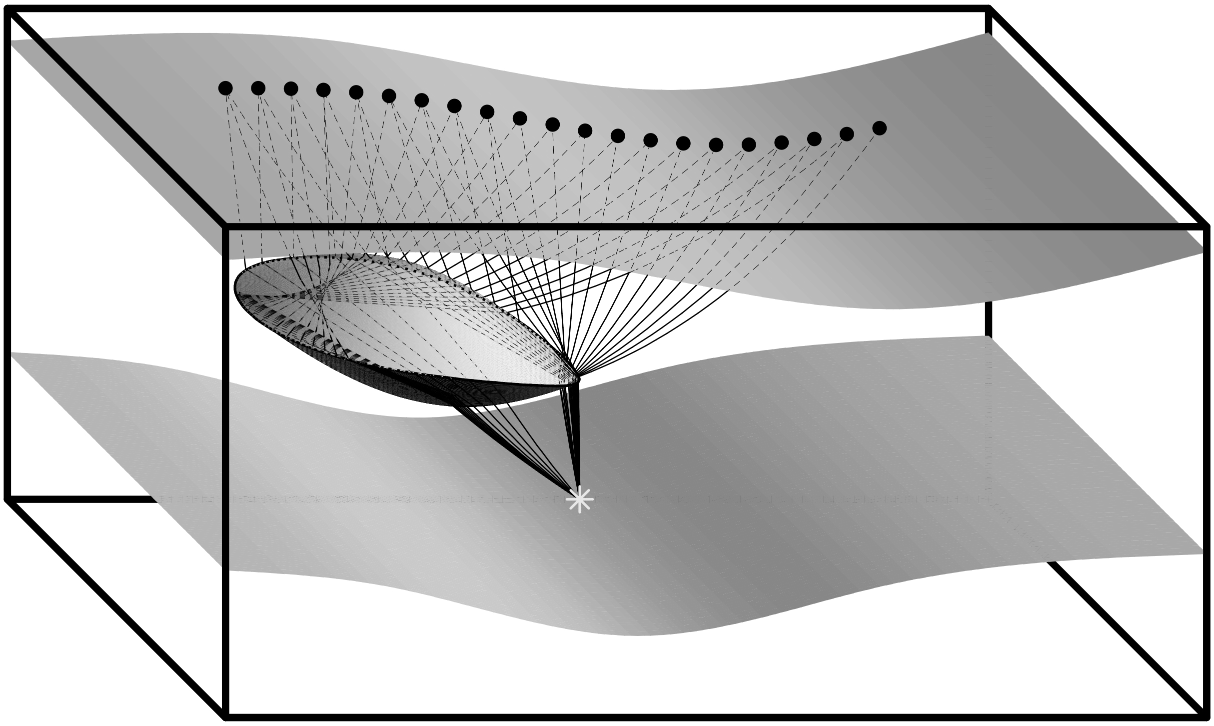

Figure 7. Two–point rays diffracted at the edge of the lenticular inclusion.

4.4. Wavefront tracing

In *wavefront tracing* (Vinje, Iversen & Gjøystdal, 1993a; Vinje et al., 1993b, 1996a, 1996b; Lambaré, Lucio & Hanyga, 1996; Gjøystdal et al., 2002), a system of wavefronts with a given step in travel time is considered for each elementary wave. The rays are then traced from the source to the first wavefront, from the first wavefront to the second wavefront, and so on. The ray histories for tracing to individual wavefronts then differ. The ray history for tracing to a wavefront corresponds to the propagation from the source to this wavefront.

The triangulation of the ray–parameter domain for tracing to the first wavefront may be carried out analogously as in two–point ray tracing or in controlled initial–value ray tracing. When tracing from the first wavefront to the second wavefront, the homogeneous triangles may be divided. There are two different reasons for this division:
(a) Rays of an equal ray history on the first wavefront are of different ray histories on the second wavefront. A homogeneous triangle then becomes inhomogeneous, and has to be divided.
(b) The evolution of the secondary ray–parameter metric tensor makes the homogeneous triangle too large on the second wavefront. The homogeneous triangle then has to be divided into smaller homogeneous triangles (Vinje, 1997).

When tracing from the first wavefront to the second wavefront, new rays required for the division of the triangles are not shot from the source. They are shot from the first wavefront instead. The initial values for the new rays are determined by the interpolation within homogeneous triangles corresponding to the first wavefront. The interpolation of the initial coordinates should be performed along the spatial boundary of the ray cell rather than along the wavefront, if the interpolation within ray cells is to be applied, see Section 4.5. Analogously for tracing to subsequent wavefronts.

Wavefront tracing is especially suitable for interpolating within ray cells. It has a great advantage for this purpose over controlled initial–value ray tracing. In controlled initial–value ray tracing, the ray history corresponds to the whole ray. The narrow demarcation belts between different ray histories then create demarcation volumes which always extend from the source, even if the ray histories begin to differ later on. These demarcation volumes then form gaps inside the volume corresponding to equal ray history. This problem does not occur in wavefront tracing.

Another, but less important advantage of wavefront tracing is the smaller total length of traced rays than in controlled initial–value ray tracing. In particular, the auxiliary rays required during the determination of boundaries between ray histories are traced only between two neighbouring wavefronts. The difference in the total length of useful rays (vertices of homogeneous triangles) is less pronounced, and is partly reduced by more expensive numerical integration. Because of short ray elements between wavefronts, wavefront tracing requires the application of Runge–Kutta methods, which are roughly twice more expensive than predictor–corrector methods recommended for two–point ray tracing and controlled initial–value ray tracing (Červený, Klimeš & Pšenčík, 1988, sec. 5.8). Let us emphasize that the most important factor, influencing the efficiency of both controlled initial–value ray tracing and wavefront tracing by an order of magnitude, is the determination of the boundaries between ray histories. The second important factor is a good selection of the secondary ray–parameter metric tensor. The total length of traced rays is not too important.

Wavefront tracing is thus an excellent method for calculating travel times and other quantities. The reliability and efficiency of wavefront tracing strongly depends, especially in block models with interfaces, on detailed numerical algorithms for triangulation of the ray–parameter domains corresponding to the individual wavefronts.

4.5. Interpolation within ray cells

Each homogeneous triangle in the ray–parameter domain, created during *controlled initial–value ray tracing*, generates a ray tube limited by the corresponding three rays. Each traced ray is recorded as a set of points of two types. First, there are the points on the ray stored with a given travel–time step; we shall refer to them as *time points*. Second, there are the points of interaction of the ray with interfaces and other surfaces (reflection, transmission or termination); we shall refer to them as *interaction points*. A ray tube is thus represented as a set of the above points on the three rays, which form the ray tube, see Figure 8.

For interpolation, the ray tube is decomposed into *ray cells*. *Regular ray cells* are defined by six points on the rays. Both the bottom and the top of the regular ray cell are formed by triangles defined by three points on the rays. These are usually the time points of the same travel–time level. *Degenerate ray cells* are formed by five or four points. They occur mainly at a point source, where the bottom of a ray cell is formed by a single point, in front of and beyond interfaces, where one or two points of the top of a ray cell coincide with the corresponding point(s) of the bottom, and in front of the end surface where the ray tube terminates.

We start the decomposition at the source, taking the first points on the rays as the bottom of the first ray cell. The bottom is formed by a single point for point sources, by a line segment for straight–line sources, or by a triangle for curved–line sources and

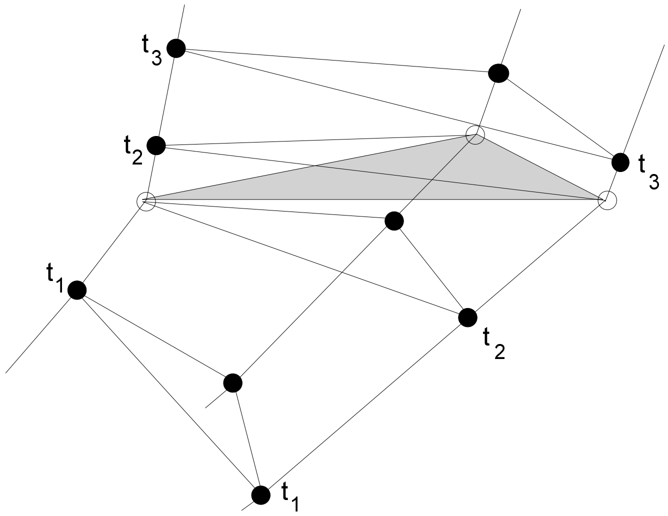

Figure 8. Decomposition of a ray tube into ray cells. Three rays, forming a ray tube, cross an interface. Bullets represent the points at three travel–time levels t_1, t_2 and t_3 ("time points"). Circles are the points of intersection of the rays with the interface ("interaction points"). Starting from the bottom of the figure, we can see one regular ray cell formed by six points, then two degenerate cells formed by five and four points, and again a regular ray cell. The bottom of the first ray cell and the top of the fourth cell are formed by triangles which approximate wavefronts at travel-time levels t_1 and t_3. The shaded triangle, which is the top of the second cell and the bottom of the third cell, approximates the interface.

surface sources. We then proceed along the rays to the second points on the rays, taking the second points simultaneously as the top of the first ray cell and as the bottom of the second ray cell (if all the second points are time points). We continue this procedure along the rays and create regular ray cells, until we reach the first interaction point on anyone of the rays. Here we postpone proceeding along this ray, and we proceed only along the remaining two rays (or one ray), until we reach the interaction points on all the three rays. We may thus create one or more degenerate ray cells, until the three interaction points form the top of the last created ray cell. If they lie at the end surface, the decomposition of the ray tube into ray cells is completed. If they are located at an interface, we continue the process of creating ray cells beyond this interface. We may again have to postpone proceeding along some ray(s) and we may create some degenerate ray cell(s), until we get time points of equal travel time at the top of the recently created ray cell. We can then continue with creating regular ray cells with equal travel time at their bottoms and tops. See Figure 8.

Ray cells in *wavefront tracing* are analogous to ray cells created during controlled initial–value ray tracing. The only difference is that the ray tube can be split into several ray tubes in wavefront tracing. This splitting should always be realized along the planar top triangle of the last cell before splitting in order to assure perfect contact of the ray cells.

In decomposing ray tubes into ray cells, we must ensure, that the sides of all the ray cells are equal for neighbouring ray tubes. Otherwise, some gridpoints of a receiver grid might be considered to be located within none, or within both of the neighbouring ray cells.

By determining a ray cell we understand the decision, whether an arbitrary point lies within the ray cell, or not. Assume a ray cell defined by six points on three rays:

points $\mathbf{B}_i$ at the bottom, and points $\mathbf{C}_i$ at the top of the ray cell, see Figure 9. We introduce a local coordinate system parametrized by parameters w_1, w_2, w_B. The coordinates of arbitrary point $\mathbf{X}$ are then given as

$$\mathbf{X} = \mathbf{A}_3 + w_1(\mathbf{A}_1 - \mathbf{A}_3) + w_2(\mathbf{A}_2 - \mathbf{A}_3) \quad , \tag{23}$$

where

$$\mathbf{A}_i = \mathbf{C}_i + w_\mathrm{B}(\mathbf{B}_i - \mathbf{C}_i) \quad , \quad i = 1, 2, 3 \quad . \tag{24}$$

Parameters w_1, w_2 are the barycentric coordinates in triangle $\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3$. A degenerate ray cell is treated as a special case of a regular ray cell with coinciding vertices. Equation (23) is the parametric description of planar triangle $\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3$. Equations (24) are parametric descriptions of straight-line segments $\mathbf{B}_1\mathbf{C}_1$, $\mathbf{B}_2\mathbf{C}_2$ and $\mathbf{B}_3\mathbf{C}_3$, approximating the curved ray segments. Equations (23) and (24) may also be expressed as

$$\mathbf{X} = w_\mathrm{B} \sum_{i=1}^{3} w_i \mathbf{B}_i + w_\mathrm{C} \sum_{i=1}^{3} w_i \mathbf{C}_i \tag{25}$$

with $w_\mathrm{C} = 1 - w_\mathrm{B}$, $w_3 = 1 - w_1 - w_2$.

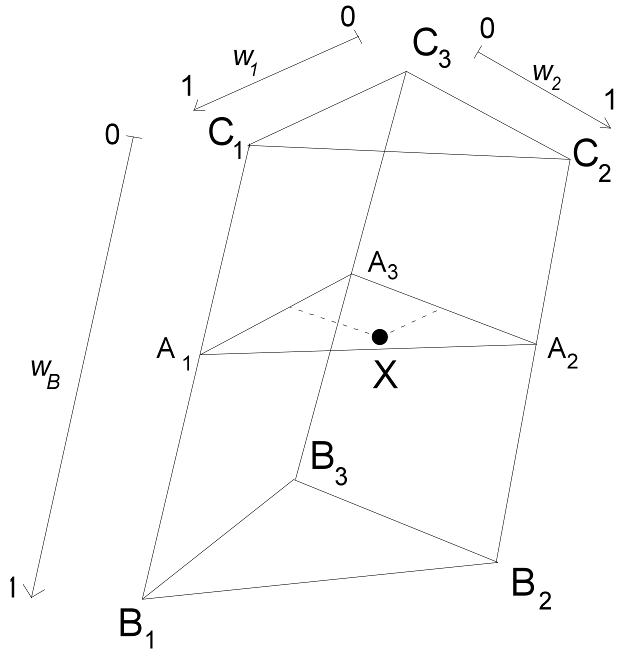

Figure 9. A ray cell is defined by six points $\mathbf{B}_1$, $\mathbf{B}_2$, $\mathbf{B}_3$, $\mathbf{C}_1$, $\mathbf{C}_2$, $\mathbf{C}_3$, situated on three rays. Point $\mathbf{X}$ is examined, whether it is located within the ray cell. Its position may be expressed using the local coordinate system parametrized by parameters w_1, w_2, w_B. If all of the quantities w_1, w_2, w_B, $1 - w_1 - w_2$ are in interval $\langle 0, 1 \rangle$, point $\mathbf{X}$ lies within the ray cell.

Point $\mathbf{X}$ lies within the ray cell if all values w_1, w_2, w_3, w_B are within interval $\langle 0, 1 \rangle$. Thus, for each receiver, which might be located within the selected ray cell, we compute its local coordinates w_1, w_2, w_B and quantities w_3, w_C. If all of them display values from interval $\langle 0, 1 \rangle$, point $\mathbf{X}$ is within the ray cell, and the values of all required quantities will be determined by interpolation within the ray cell.

Local coordinate w_B is the root of the cubic equation (Bulant & Klimeš, 1999, eq. 6). For each root w_B from interval $\langle 0, 1 \rangle$, local coordinates w_1 and w_2 are the unique solutions of the linear equation (Bulant & Klimeš, 1999, eq. 17). If the ray cell touches a caustic, two or three (in 3–D) different sets of values w_1, w_2, w_3, $w_{\rm B}$ may be found within interval $\langle 0,1\rangle$ for the same point $\mathbf{X}$, see Figure 10. Different coordinates w_1, w_2, $w_{\rm B}$ of the same point $\mathbf{X}$ correspond to different ray take–off parameters γ_1, γ_2, and the interpolation should be performed for each set of coordinates w_1, w_2, $w_{\rm B}$.

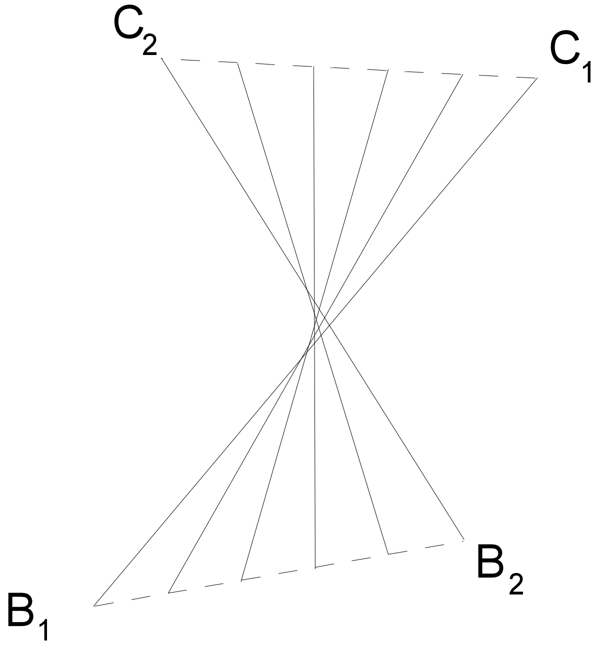

Figure 10. A 2–D ray cell touches a caustic. The lines approximate rays, the dashed lines represent the bottom and the top of the cell. Each point of a small area in the middle of the cell lies on two rays. For such points we get two (two or three in 3–D) different sets of local coordinates, and we then obtain two (two or three in 3–D) different values of interpolated quantities.

In each ray cell, travel time and other quantities are interpolated as functions of local coordinates w_1, w_2, $w_{\rm B}$. Since both the travel time and its gradient are known at all vertices of a ray cell, the travel time can be approximated by bicubic interpolation (Bulant & Klimeš, 1999, eqs. 25–27 and 35), see Figure 11a. The application of the second derivatives of travel time during interpolation should be avoided because they may be infinite. Other quantities (e.g., amplitudes) may be approximated by bilinear interpolation (Bulant & Klimeš, 1999, eqs. 18 and 19). Refer to Figures 11a–d for the comparison between bicubic and bilinear interpolation of travel time.

The bicubic interpolation fits the values and the first spatial partial derivatives at the six vertices of the ray cell. The functional values at each cell face depend on the functional values and derivatives at the vertices of the face only. The interpolated values are thus continuous across the cell faces. The gradient of the interpolated function is continuous at the triangular cell faces and along the cell edges, but may be discontinuous at the tetragonal cell faces separating ray tubes. The bicubic interpolation fits linear and quadratic functions exactly.

The described algorithm incorporates both the decision, whether a receiver lies in the ray cell and the interpolation to the receiver. It is applicable to all ray cells formed by six, five or four points. All the quantities computed along rays may be interpolated, using the values of the quantities only at the vertices of the corresponding ray cell.

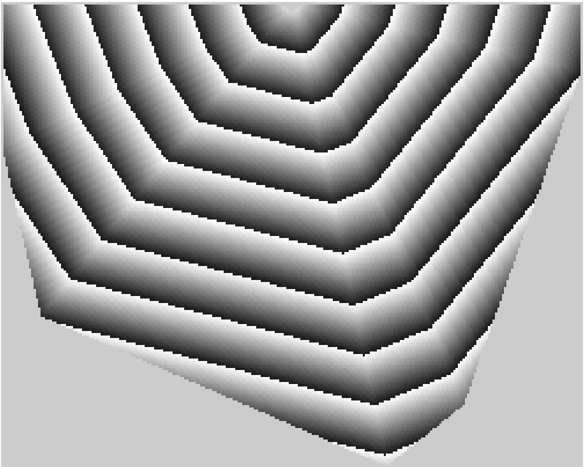

Figure 11a. Travel times in a homogeneous model interpolated bilinearly within wide 3–D ray tubes. One grey–scale cycle from white to dark grey corresponds to 0.3 s. The solid light–grey region is not covered by ray cells.

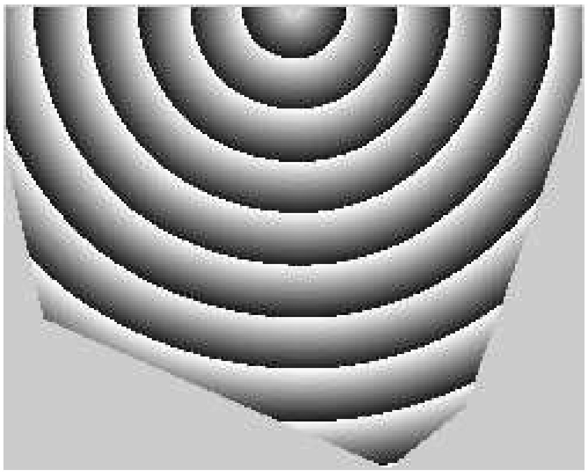

Figure 11b. Travel times in a homogeneous model calculated within the same wide 3–D ray tubes using the bicubic interpolation. The maximum travel time is about 2.7 s. The solid light–grey region is not covered by ray cells.

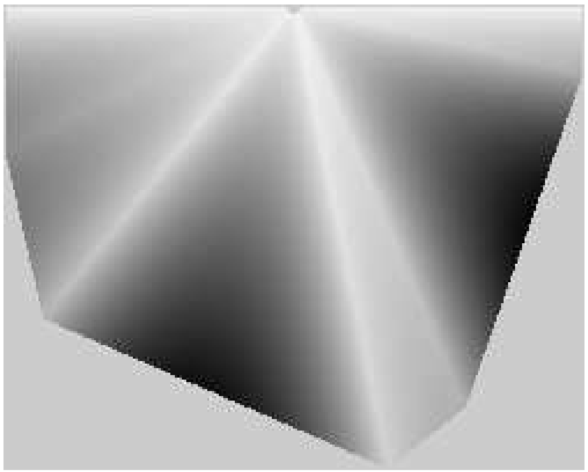

Figure 11c. Differences of the bilinearly interpolated travel times from the exact solution. The maximum difference is about 0.308 s. The grey scale is the same as in Figure 11a. The solid light–grey region is not covered by ray cells.

Figure 11d. Differences of the bicubically interpolated travel times from the exact solution. The maximum difference is about 0.017 s. The grey scale is the same as in Figure 11a. The solid light–grey region is not covered by ray cells.

5. Dynamic ray tracing. Paraxial ray methods

Standard ray tracing can be used to compute the ray as a space trajectory, and the travel time τ, slowness vector $\mathbf{p}$, ray–velocity vector $\boldsymbol{\mathcal{U}}$, polarization vector $\mathbf{g}$ and the vector $\boldsymbol{\eta}$ of the wave under consideration at any point of the ray. These quantities, however, are not known outside the traced ray (in its vicinity). To find them outside the traced ray, it would be necessary to compute new rays. In addition, geometrical spreading and ray amplitudes cannot be computed by single ray tracing, without considering vicinity rays.

These disadvantages can be reduced by an additional procedure, called dynamic ray tracing, or also paraxial ray tracing. Ray tracing, supplemented by dynamic ray tracing (DRT), is quite sufficient to determine geometrical spreading and ray amplitudes along a single ray. Moreover, DRT can be used to determine the DRT propagator matrix, which can be applied to compute all kinematic quantities not only along a single ray, but also in some vicinity of the ray. We call this vicinity the paraxial vicinity of the ray, and speak of paraxial rays, paraxial travel times, paraxial slowness vectors, etc. The DRT propagator matrix can also be used to solve various boundary–value ray tracing problems.

Dynamic ray tracing can be performed in various coordinate systems: It is particularly suitable to use the ray–centred coordinate system, in which one coordinate axis coincides with the ray. The ray–centred coordinate system is described in Section 5.1, and the dynamic ray tracing systems in Cartesian and ray–centred coordinates in Section 5.2. In Section 5.3, the computation of DRT propagator matrices is discussed, and in Sections 5.4 and 5.5, the application of DRT propagator matrices in the computation of higher–order partial derivatives of the travel time field is described. Finally, Section 5.6 contains a list of some other applications of DRT propagator matrices.

5.1. Ray–centred coordinates

Let us consider an arbitrarily selected ray Ω and call it the central ray. Along it, let us introduce the ray–centred coordinate system q_1, q_2, q_3, in which the central ray Ω is the q_3–coordinate axis. Coordinate q_3 may be an arbitrary monotonic variable along the ray. It is most natural to take $q_3 = \tau$, similarly as in ray tracing, but not necessarily. We introduce q_3 by the relation

$$q_3 = q_{30} + \int_{\tau_0}^{\tau} w(\tau)\mathrm{d}\tau \quad , \tag{26}$$

where the integration is taken along ray Ω. Consequently, q_3 equals τ for $w(\tau) = 1$, q_3 equals the arclength for $w(\tau) = \mathcal{U}(\tau)$, etc. Ray–centred coordinates q_1 and q_2 are introduced in a plane tangent to the wavefront, intersecting the central ray Ω at the point specified by q_3. At points along ray Ω, coordinates q_1 and q_2 are zero. The mutual relation between the global Cartesian coordinates x_i and the ray–centred coordinates q_k may be expressed as follows:

$$x_i(q_n) = x_i^{\Omega}(q_3) + H_{iN}(q_3)q_N \quad . \tag{27}$$

Here $x_i = x_i^{\Omega}(q_3)$ represents the equation of the central ray (along which $q_1 = q_2 = 0$). The 3×3 transformation matrices from ray–centred to Cartesian coordinates and back are denoted by H_{im} and $\bar{H}_{mi}$, and satisfy relations

$$H_{im}\bar{H}_{mj} = \delta_{ij} \quad , \qquad \bar{H}_{mi}H_{in} = \delta_{mn} \quad . \tag{28}$$

The elements H_{i3} and $\bar{H}_{3i}$ are given by simple explicit relations:

$$H_{i3} = w^{-1}\mathcal{U}_i \quad , \qquad \bar{H}_{3i} = wp_i \quad . \tag{29}$$

In anisotropic inhomogeneous media, the ray–centred coordinate system is non–orthogonal. Vectors H_{i1}, H_{i2} and H_{i3} represent contravariant basis vectors (with $H_{i3} = w^{-1}\mathcal{U}_i$ tangent to the ray, and H_{i1}, H_{i2} tangent to the wavefront), and vectors $\bar{H}_{1i}$, $\bar{H}_{2i}$ and $\bar{H}_{3i}$ covariant basis vectors (with $\bar{H}_{3i} = wp_i$, perpendicular to the wavefront, and $\bar{H}_{1i}$, $\bar{H}_{2i}$ perpendicular to the ray). As vectors $\boldsymbol{\mathcal{U}}$ and $\mathbf{p}$ are known from ray tracing at any point of ray Ω, the covariant basis vectors can be simply recalculated into contravariant, and vice versa. Thus, it is quite sufficient to calculate vectors H_{1i} and H_{2i} (or $\bar{H}_{i1}$ and $\bar{H}_{i2}$); the complete system of covariant and contravariant basis vectors can then be simply obtained at any point of the central ray Ω. The choice whether we compute the covariant or contravariant basis vectors depends merely on the numerical efficiency of the relevant algorithms.

An efficient numerical algorithm has been proposed to compute the contravariant basis vectors H_{i1} and H_{i2}. They may be determined using ordinary differential equations

$$\mathrm{d}H_{iM}/\mathrm{d}q_3 = -(H_{lM} \cdot \eta_l)p_i/(p_k p_k) \quad , \qquad M = 1,2 \quad . \tag{30}$$

Vectors $\boldsymbol{\eta}$ and $\mathbf{p}$ are known from ray tracing, see (17) for $\boldsymbol{\eta}$. The ordinary differential equation (30) does not depend on w, and may be used for any monotonic coordinate q_3.

The contravariant basis vectors H_{i1} and H_{i2}, calculated from (30), have certain very interesting and important properties. Consider an initial point $q_3 = q_{30}$ of ray Ω, and two mutually perpendicular unit vectors $H_{i1}(q_{30})$, $H_{i2}(q_{30})$, perpendicular also to slowness vector $\mathbf{p}$ at this point. The ordinary differential equation (30) then guarantees that H_{i1} and H_{i2} satisfy the same properties at any point of the ray. We do not need to calculate both vectors H_{i1} and H_{i2}; it is sufficient to calculate only one of them. Let us denote $\mathbf{H}_1 = [H_{11}, H_{21}, H_{31}]^{\mathrm{T}}$, $\bar{\mathbf{H}}_1 = [\bar{H}_{11}, \bar{H}_{12}, \bar{H}_{13}]^{\mathrm{T}}$, and similarly $\mathbf{H}_2$ and $\bar{\mathbf{H}}_2$. Once $\mathbf{H}_1$ is known, $\mathbf{H}_2$, $\bar{\mathbf{H}}_1$ and $\bar{\mathbf{H}}_2$ can be calculated using the relations:

$$\mathbf{H}_2 = \frac{\mathbf{p} \times \mathbf{H}_1}{|\mathbf{p} \times \mathbf{H}_1|} \quad , \qquad \bar{\mathbf{H}}_1 = \frac{\mathbf{H}_2 \times \mathbf{H}_3}{\mathbf{H}_3 \cdot (\mathbf{H}_1 \times \mathbf{H}_2)} \quad , \qquad \bar{\mathbf{H}}_2 = \frac{\mathbf{H}_3 \times \mathbf{H}_1}{\mathbf{H}_3 \cdot (\mathbf{H}_1 \times \mathbf{H}_2)} \quad . \tag{31}$$

In isotropic inhomogeneous media, the contravariant and covariant basis vectors coincide, because we have chosen vectors H_{i1} and H_{i2} unit and mutually perpendicular. It is very interesting to see that the basis vectors of the ray–centred coordinate systems also represent the polarization vectors of S waves propagating in isotropic inhomogeneous media. Compare equations (21) and (30), which coincide in this case. For anisotropic inhomogeneous media, however, the basis vectors are different from the polarization vectors.

The numerical solution of (30) is easy, fast and stable, and may be used both for isotropic and anisotropic inhomogeneous media. Once the contravariant basis vectors of the ray–centred coordinate system $H_{i1}(\tau)$, $H_{i2}(\tau)$ and $H_{i3}(\tau)$ are known along the ray, the Cartesian and ray–centred coordinates of any point in the vicinity of the central ray Ω can be simply mutually recalculated.

There are many other possibilities of introducing the ray–centred basis vectors and relevant ray–centred coordinates than (30). For example, we could use an equation analogous to (30) directly for the covariant basis vectors $\bar{H}_{i1}$ and $\bar{H}_{i2}$, perpendicular to the ray. However, Equation (30) with (31) yields probably the simplest algorithm to calculate the basis vectors along the ray. The advantages are that both $\mathbf{p}$ and $\boldsymbol{\eta} = \mathrm{d}\mathbf{p}/\mathrm{d}\tau$ are known from ray tracing and that the same system (30) may be used both for anisotropic and isotropic inhomogeneous media.

5.2. Dynamic ray tracing

Dynamic ray tracing (DRT) consists in solving a system of linear ordinary differential equations of the first order along the ray (DRT system). It yields derivatives of coordinates of a ray and of the corresponding slowness vectors with respect to, e.g., ray parameters. Although the DRT system is very simple and can be integrated without significant numerical effort (as opposed to ray tracing), it extends the possibilities of ray tracing considerably. DRT was originally used to determine geometrical spreading, needed to calculate the amplitudes along the ray. Recently, however, the number of applications of DRT have increased so much that DRT should be a part of any ray tracing computer package.

The DRT system can be expressed in many forms and in various coordinate systems. The most common coordinate systems, in which the DRT system is expressed, are general Cartesian coordinate systems and some kinds of ray–centred coordinates, connected with the ray. Both systems are alternative. The number of equations in the DRT system in ray–centred coordinates is lower than that in Cartesian coordinates, but the DRT system in ray–centred coordinates additionally requires also the basis vectors of the ray–centred coordinate system to be computed, see (30).

In global Cartesian coordinates, DRT is designed to compute six quantities Q_n and P_n $(n = 1, 2, 3)$ along the ray,

$$Q_n = \partial x_n/\partial\gamma \quad , \qquad P_n = \partial p_n/\partial\gamma \quad , \tag{32}$$

where γ is a chosen ray parameter or some initial parameter of the ray (x_{i0}, p_{i0}), or some other parameter influencing the ray. The DRT system for Q_n and P_n is then simply obtained by taking the partial derivatives of (9) with respect to γ:

$$\mathrm{d}Q_n/\mathrm{d}\tau = A_{nm}Q_m + B_{nm}P_m \ , \quad \mathrm{d}P_n/\mathrm{d}\tau = -C_{nm}Q_m - D_{nm}P_m \ . \tag{33}$$

Here

$$\begin{aligned} A_{nm} &= \partial^2\mathcal{H}/\partial p_n \partial x_m \quad , \qquad & B_{nm} &= \partial^2\mathcal{H}/\partial p_n \partial p_m \quad , \\ C_{nm} &= \partial^2\mathcal{H}/\partial x_n \partial x_m \quad , \qquad & D_{nm} &= \partial^2\mathcal{H}/\partial x_n \partial p_m \quad , \end{aligned} \tag{34}$$

and $\mathcal{H}$ denotes the Hamiltonian, which is a homogeneous function of the second degree in p_i (with a possible additional constant).

Thus, DRT system (33) consists of six linear ordinary differential equations. DRT system (33) can be used both for isotropic and anisotropic inhomogeneous media, only the Hamiltonian $\mathcal{H}$ should be specified correspondingly. The general form of Hamiltonian $\mathcal{H}$ for anisotropic inhomogeneous media is given by (4), and for isotropic inhomogeneous media by (3). DRT system (33) does not require the basis vectors of the ray–centred coordinate system to be computed. We do not present detailed equations for A_{nm}, B_{nm}, C_{nm} and D_{nm} here, they can be found in the ray theory literature, see, e. g., Červený (2001, sec. 4.14.1).

Let us briefly discuss the physical meaning of the computed quantities Q_n and P_n, given by (32). Quantities Q_n and P_n show how point x_n and slowness vector p_n at the point on the ray change, when parameter γ is changed. Thus, using the results of DRT, we can approximately compute various ray–theory quantities (rays, travel times, slowness vectors, etc.) not only along the ray, but also in the "paraxial" vicinity of the ray. We then speak of paraxial rays, paraxial travel times, paraxial slowness vectors, etc.

The DRT system can also be expressed in the ray–centred coordinates, connected with the basis vectors described in the preceding section. We denote these coordinates by q_i. Coordinates q_1 and q_2 correspond to the contravariant basis vectors $\mathbf{H}_1$ and $\mathbf{H}_2$, and measure the distance from the ray along the plane tangent to the wavefront. The ray is then specified by relations $q_1 = q_2 = 0$.

Klimeš (1994) proved that the six equations of the DRT system (33) can be reduced to four equations in ray–centred coordinates q_1, q_2. We denote

$$Q_N^{(q)} = \partial q_N / \partial \gamma \quad , \qquad P_N^{(q)} = \partial(\partial \tau / \partial q_N)/\partial \gamma \quad , \tag{35}$$

$N = 1, 2$. The DRT system for $Q_N^{(q)}$ and $P_N^{(q)}$, $N = 1, 2,$, is then given by four equations:

$$\mathrm{d}Q_N^{(q)}/\mathrm{d}\tau = A_{NM}^{(q)} Q_M^{(q)} + B_{NM}^{(q)} P_M^{(q)} \ , \quad \mathrm{d}P_N^{(q)}/\mathrm{d}\tau = -C_{NM}^{(q)} Q_M^{(q)} - D_{NM}^{(q)} P_M^{(q)} \ . \tag{36}$$

Here

$$\begin{aligned} A_{NM}^{(q)} &= \bar{H}_{Ni} H_{jM} A_{ij} - d_{NM} \ , \qquad & B_{NM}^{(q)} &= \bar{H}_{Ni} \bar{H}_{Mj} B_{ij} \ , \\ C_{NM}^{(q)} &= H_{iN} H_{jM} (C_{ij} - \eta_i \eta_j) \ , \qquad & D_{NM}^{(q)} &= H_{iN} \bar{H}_{Mj} D_{ij} - d_{MN} \ , \end{aligned} \tag{37}$$

and

$$d_{NM} = \bar{H}_{Ni}\, \mathrm{d}H_{iM}/\mathrm{d}\tau = -(\bar{H}_{Ni} p_i)(H_{lM} \eta_l)/(p_k p_k) \quad . \tag{38}$$

The 3×3 matrices $A_{ij}, B_{ij}, C_{ij}, D_{ij}$ are given by (34), η_i by (17), H_{iM} (and $\bar{H}_{Mi}$) are the Cartesian components of the contravariant (covariant) basis vectors. They are calculated using (30) and (31).

Mostly, DRT system (36) is applied to orthonomic systems of rays, specified by two ray parameters γ_1, γ_2. We introduce two 2×2 matrices $\mathbf{Q}^{(q)}$ and $\mathbf{P}^{(q)}$, given by their components,

$$Q_{NK}^{(q)} = \partial q_N / \partial \gamma_K \quad , \qquad P_{NK}^{(q)} = \partial(\partial \tau / \partial q_N)/\partial \gamma_K \quad . \tag{39}$$

DRT system (36) is then suitably expressed in matrix form:

$$\mathrm{d}\mathbf{Q}^{(q)}/\mathrm{d}\tau = \mathbf{A}^{(q)}\mathbf{Q}^{(q)} + \mathbf{B}^{(q)}\mathbf{P}^{(q)} \ , \quad \mathrm{d}\mathbf{P}^{(q)}/\mathrm{d}\tau = -\mathbf{C}^{(q)}\mathbf{Q}^{(q)} - \mathbf{D}^{(q)}\mathbf{P}^{(q)} \ . \tag{40}$$

The 2×2 matrices play an important role in the computation of geometrical spreading and in paraxial ray methods. Quantities $\det(\mathbf{Q}^{(q)})$ or $\det(\mathbf{P}^{(q)})$ may vanish at certain points of the ray. The point at which $\det(\mathbf{Q}^{(q)}) = 0$ is called the caustic point, and the point at which $\det(\mathbf{P}^{(q)}) = 0$ is known as the pseudocaustic (or telescopic) point.

DRT is universal for P waves propagating in arbitrary anisotropic inhomogeneous media. The anisotropic DRT system for P waves reduces to the isotropic DRT system for P waves, when anisotropy vanishes. This is not the case of S waves.

For isotropic inhomogeneous media, the dynamic ray tracing system in ray–centred coordinates is considerably simpler. Equation (10) yields the DRT system in the following matrix form:

$$\mathrm{d}\mathbf{Q}^{(q)}/\mathrm{d}\tau = V^2 \mathbf{P}^{(q)} \quad , \qquad \mathrm{d}\mathbf{P}^{(q)}/\mathrm{d}\tau = -V^{-1}\mathbf{V}^{(q)}\mathbf{Q}^{(q)} \quad . \tag{41}$$

Here velocity V equals α or β, and $\mathbf{V}^{(q)}$ is a 2×2 matrix with elements $V^{(q)}_{,IJ} = (\partial^2 V/\partial q_I \partial q_J)_{q_1=q_2=0}$. This DRT system was introduced to seismology by Popov & Pšenčík (1978a,b). It is used to compute geometrical spreading along 2–D rays in isotropic inhomogeneous layered media in computer package SEIS, see Červený & Pšenčík (1984), and along 3–D rays in computer package CRT, see Červený, Klimeš & Pšenčík (1988). In 3–D *anisotropic* inhomogeneous layered media, DRT system (33) in Cartesian coordinates is used to compute geometrical spreading and to perform two–point ray tracing in computer package ANRAY by Gajewski & Pšenčík (1987).

5.3. DRT propagator matrix

The dynamic ray tracing system consists of linear ordinary differential equations of the first order. Consequently, it allows the fundamental matrix, consisting of linearly independent solutions of the DRT system, to be introduced. This can be done by specifying the fundamental matrix by the identity matrix at an arbitrarily selected point $\tau=\tau_0$ of the ray. We call this fundamental matrix the *DRT propagator matrix*, and denote it $\mathbf{\Pi}(\tau,\tau_0)$. Clearly, the DRT propagator matrix $\mathbf{\Pi}(\tau,\tau_0)$ depends on variable τ and on the initial time τ_0. Once a routine for the solution of the DRT system is available, the computation of the DRT propagator matrix is easy: we merely perform dynamic ray tracing along the relevant ray, starting from τ_0, with the initial conditions specified by the identity matrix.

In this section, we consider only the 4×4 DRT propagator matrix $\mathbf{\Pi}(\tau,\tau_0)$ for 3–D models, corresponding to DRT system (36) in ray–centred coordinates. The DRT propagator matrix $\mathbf{\Pi}(\tau,\tau_0)$ has several very interesting and important properties:

(a) If the DRT propagator matrix $\mathbf{\Pi}(\tau,\tau_0)$ is known, the solution of the DRT system can be easily found at point τ of the ray for arbitrary initial conditions at τ_0:

$$\begin{pmatrix} \mathbf{Q}^{(q)}(\tau) \\ \mathbf{P}^{(q)}(\tau) \end{pmatrix} = \mathbf{\Pi}(\tau,\tau_0) \begin{pmatrix} \mathbf{Q}^{(q)}(\tau_0) \\ \mathbf{P}^{(q)}(\tau_0) \end{pmatrix} \quad . \tag{42}$$

(b) The DRT propagator matrix $\mathbf{\Pi}(\tau,\tau_0)$ is symplectic,

$$\mathbf{\Pi}^{\mathrm{T}}(\tau,\tau_0)\,\mathbf{J}\,\mathbf{\Pi}(\tau,\tau_0) = \mathbf{J} \quad , \qquad \text{where} \qquad \mathbf{J} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{pmatrix} \quad . \tag{43}$$

Here $\mathbf{I}$ and $\mathbf{0}$ are 2×2 identity and null matrices. It follows from (43) that $\det \mathbf{\Pi}(\tau,\tau_0) = 1$, so that $\mathbf{\Pi}(\tau,\tau_0)$ is regular along the whole ray.

(c) The DRT propagator matrix $\mathbf{\Pi}(\tau,\tau_0)$ satisfies the chain rule,

$$\mathbf{\Pi}(\tau,\tau_0) = \mathbf{\Pi}(\tau,\tau_1)\,\mathbf{\Pi}(\tau_1,\tau_0) \quad , \tag{44}$$

where τ_1 specifies an arbitrary point on the ray. Equation (44) also implies

$$\mathbf{\Pi}^{-1}(\tau_1,\tau_0) = \mathbf{\Pi}(\tau_0,\tau_1) \quad , \tag{45}$$

for points τ_0 and τ_1 arbitrarily situated on the ray.

(d) The DRT propagator matrix can also be defined for rays interacting with interfaces (reflections, transmissions). We merely introduce the *interface propagator matrix* $\mathbf{\Pi}(\tau^{\mathrm{r}},\tau^{\mathrm{inc}})$, where τ^{inc} corresponds to the point of incidence, and τ^{r} to the corresponding point of reflection/transmission (at the same interface), $\tau^{\mathrm{r}}=\tau^{\mathrm{inc}}$. For the point of incidence τ^{inc} situated between τ_0 and τ, we then modify (44) as follows:

$$\mathbf{\Pi}(\tau,\tau_0) = \mathbf{\Pi}(\tau,\tau^{\mathrm{r}})\,\mathbf{\Pi}(\tau^{\mathrm{r}},\tau^{\mathrm{inc}})\,\mathbf{\Pi}(\tau^{\mathrm{inc}},\tau_0) \quad . \tag{46}$$

Thus, it is only necessary to insert the interface propagator matrix $\mathbf{\Pi}(\tau^{\mathrm{r}}, \tau^{\mathrm{inc}})$ into the proper place of the chain. For more details on interface propagator matrices refer to Farra & Le Bégat (1995), and Moser (2004). Of course, (46) may be further chained.

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

$$\mathbf{\Pi}(\tau, \tau_0) = \begin{pmatrix} \mathbf{Q}_1(\tau, \tau_0) & \mathbf{Q}_2(\tau, \tau_0) \\ \mathbf{P}_1(\tau, \tau_0) & \mathbf{P}_2(\tau, \tau_0) \end{pmatrix} \quad . \tag{47}$$

The 2×2 submatrix $\mathbf{Q}_2(\tau, \tau_0)$ can be used to determine relative geometrical spreading $\mathcal{L}(\tau, \tau_0)$,

$$\mathcal{L}(\tau, \tau_0) = |\det \mathbf{Q}_2(\tau, \tau_0)| \quad . \tag{48}$$

It can also be used to determine the complete phase shift of the Green function due to caustics $T^{\mathrm{G}}(\tau, \tau_0)$, which is closely related to the number of zeros of det $\mathbf{Q}_2(\tau, \tau_0)$ along the ray from τ_0 to τ. Both $\mathcal{L}(\tau, \tau_0)$ and $T^{\mathrm{G}}(\tau, \tau_0)$ play a fundamental role in the determination of complex–valued vectorial amplitudes of the Green function along the ray, see Section 6.

Note that matrix $\mathbf{Q}_2(\tau, \tau_0)$ in (48) corresponds to the DRT system (36), with the basis vectors H_{i1} and H_{i2} calculated by (30) and (31). Consequently, relative geometrical spreading $\mathcal{L}(\tau, \tau_0)$ is related to the surface element cut out of the *wavefront* by the ray tube, not out of the plane perpendicular to the ray. Relative geometrical spreading $\bar{\mathcal{L}}(\tau, \tau_0)$, related to the surface element cut out of the plane perpendicular to the ray is simply related to $\mathcal{L}(\tau, \tau_0)$ by formula

$$[\mathcal{C}(R)\mathcal{C}(S)]^{1/2}\mathcal{L}(R, S) = [\mathcal{U}(R)\mathcal{U}(S)]^{1/2}\bar{\mathcal{L}}(R, S) \quad . \tag{49}$$

Here $\mathcal{C}$ is the phase velocity and $\mathcal{U}$ the ray velocity.

5.4. Second–order spatial derivatives of travel time

The first–order partial travel–time derivatives with respect to Cartesian coordinates x_i (components of slowness vector $\mathbf{p}$) are known along the ray from ray tracing, see (10) and (14). To compute the second–order and higher–order derivatives, DRT is needed.

The second–order partial derivatives of the travel–time field are obtained directly as a by–product of DRT. If we denote by $\mathbf{M}$ the 2×2 matrix of the second derivatives of the travel time field with respect to ray–centred coordinates q_1, q_2, we can put

$$\mathbf{M} = \mathbf{P}^{(q)}(\mathbf{Q}^{(q)})^{-1} \quad . \tag{50}$$

The 2×2 matrices $\mathbf{P}^{(q)}$ and $\mathbf{Q}^{(q)}$ are known from DRT system (40). Analogously to (50), we can also introduce the 3×3 matrix $\mathbf{N} = \mathbf{P}\mathbf{Q}^{-1}$ of the second derivatives of the travel–time field with respect to Cartesian coordinates x_i and the 3×3 matrices $\mathbf{P}$ and $\mathbf{Q}$.

Actually, the DRT system can also be expressed directly in terms of $\mathbf{M}$. It then transforms to a nonlinear matrix Riccati equation for $\mathbf{M}$. In isotropic media, the Riccati DRT system for the 2×2 matrix $\mathbf{M}$ reads

$$\mathrm{d}\mathbf{M}/\mathrm{d}\tau = -V^2\mathbf{M}\mathbf{M} - V^{-1}\mathbf{V}^{(q)} \quad . \tag{51}$$

Here V is the velocity and $\mathbf{V}^{(q)}$ the 2×2 matrix introduced below equation (41). The disadvantage of the Riccati DRT system is that it is nonlinear, whereas the DRT system (40) is linear. The Riccati DRT system is, however, broadly used in the theory of Gaussian beams, see Section 11.1.

5.5. Third–order and higher–order spatial derivatives of travel time

Once DRT has been performed, the third–order and higher–order partial derivatives of travel time can be computed by simple numerical quadratures along rays. The general form of the equations for the third–order and higher–order partial derivatives of travel time was suggested by Babich, Buldyrev & Molotkov (1985). The explicit equations for the third–order and higher–order partial derivatives of travel time in Cartesian coordinates in an isotropic medium with a homogeneous Hamiltonian of the second degree were presented by Červený (2001, sec. 4.7.6). The explicit equations for the third–order and higher–order partial derivatives of travel time in both isotropic and anisotropic media were derived by Klimeš (2002a). The equations are derived for a general Hamiltonian. The derivatives of travel time are expressed in the form of integrals along the ray, with the appropriate initial values. The integrands are composed of partial derivatives of travel time of orders lower than the one being computed, and of the phase–space partial derivatives of the Hamiltonian. The derivatives of the Hamiltonian contain partial derivatives of density–normalized elastic moduli up to the order corresponding to that of the calculated derivatives of travel time.

Examples of possible application of the third–order and higher–order partial derivatives of travel time are:

(a) Estimation of the *accuracy of travel–time interpolation within ray cells*, and the control of the accuracy by optimizing the size of ray cells in wavefront tracing and in controlled initial–value ray tracing followed by travel–time interpolation.

(b) Estimation and control of the *accuracy of the paraxial ray approximation.*

(c) Estimation of the *accuracy of the paraxial approximation of Gaussian beams.*

(d) Estimation of the *accuracy of the paraxial expansions of Fresnel edge waves* approximating the diffractions.

(e) Application to the *local and uniform asymptotic approximations* in heterogeneous media. They may be useful in deriving general high–frequency asymptotic approximations of the Maslov and Gaussian–beam integrals with respect to the ray parameters, such as the Airy approximations (third–order Taylor expansion of travel time with respect to ray parameters) and the Pearcey approximations (fourth–order Taylor expansion of travel time with respect to ray parameters).

(f) Calculating or approximating the *higher–order terms in the ray series.*

(g) Determination of the *dependence of the smoothness of travel time on the smoothness of the model.* For example, continuous and finite n–th velocity derivatives ($n \geq 2$) imply continuous and finite n–th derivatives of travel time off caustics.

(h) The third– and higher–order spatial derivatives of travel time are required when calculating the *third– and higher–order derivatives of travel time with respect to the perturbations of model parameters.*

Note that the equations for the partial derivatives of the vectorial amplitude may also be derived (e.g., Klimeš 2003a).

5.6. Applications of DRT propagator matrix. Paraxial ray methods

The DRT propagator matrix has recently found many applications in the ray method, both in isotropic and anisotropic inhomogeneous layered media. A list of several of them follows: 1) Paraxial ray theory. Computation of paraxial rays, paraxial travel times, and paraxial slowness vectors in the vicinity of the ray. (Červený, 2001, sec. 4.6). 2) Computation of the curvature of the wavefront along the ray (see Section 5.4). 3) Computation of the third–order and higher–order spatial derivatives of travel time (see Section 5.5). 4) Computation of the second–order and higher–order perturbations of travel time (see Section 9.2). 5) Computation of the perturbations or spatial derivatives of amplitude (Klimeš, 2003a). 6) Computation of paraxial Fresnel volumes (Červený & Soares, 1992). 7) Investigation of chaotic rays and evaluation of Lyapunov coefficients (see Section 13). 8) Construction of surface–to–surface propagator matrices (Hubral, Schleicher, & Tygel, 1992a, 1992b; Schleicher, Tygel & Hubral, 1993; Sun, 2004). 9) Two–point ray tracing. 10) Various boundary–value problems for a ray. 11) Investigation of reciprocity relations along a ray (see Sections 6.2 and 6.3). 12) Determination of geometrical spreading from travel time measurements (Červený, 2001). 13) Factorization of geometrical spreading. Fresnel matrices. 14) Computation of Gaussian beams and Gaussian packets (see Section 11). 15) Maslov method (see Section 11.4.6). 16) Numerous applications in ray perturbation theory (Farra, 1999). 17) Applications in migration methods (Tygel, Schleicher & Hubral, 1994; Sun, 2004). 18) Kirchhoff–Helmholtz integrals (Tygel, Schleicher & Hubral, 1994; Chapman, 2004). For more details and many references see Červený (2001, chap. 4).

6. Ray–theory amplitudes. Ray–theory Green function

Once the ray tracing and dynamic ray tracing have been performed, and polarization vectors and geometrical spreading have been determined along the ray, the computation of vectorial complex–valued amplitudes $\mathbf{U}(x_i)$ of an arbitrary elementary wave is not difficult. It is, however, still necessary to take into account the conditions at the initial point of the ray, at the end point of the ray, and at points of contact of the ray with interfaces (or with the Earth's surface). These conditions are known theoretically for most situations of practical interest, and many of them are included in the available ray program packages. For this reason, we shall be brief here. (a) The initial point of the ray often represents a point source. The most important point sources are single–force and moment–tensor point sources, including the explosive point source. The source may be situated in a smooth medium, at the surface of the Earth, or at an interface. Ray–theory equations are known even for a line source (arbitrarily curvilinear) and for a surface source (arbitrarily curved). (b) The most common conditions for the receiver point are expressed in terms of conversion coefficients. The conversion coefficients may correspond to the receiver situated at the Earth's surface, at an interface, or in a smooth medium. (c) The algorithms for evaluating the (complex–valued) reflection/transmission coefficients at interfaces and at the Earth's surface are also well known for anisotropic media. Displacement R/T coefficients, or normalized R/T coefficients (normalized with respect to the energy flux across the interface) are mostly, used. The normalized R/T coefficients have the very useful property of reciprocity.

6.1. Ray–theory Green function

A very important role in various applications in seismology and seismic exploration for oil is played by the elastodynamic Green function. The elastodynamic Green function $G_{in}(R,t,S,t_0)$ represents the i–th Cartesian component of the displacement vector at location R and time t due to a single–force point source situated at location S, oriented along the n–th Cartesian axis, with the time dependence $\delta(t-t_0)$. Quite analogously, we can introduce the elementary ray–theory elastodynamic Green function. The only differences are that the elementary ray–theory elastodynamic Green function does not correspond to the complete wavefield, but to the wavefield of an arbitrarily selected elementary wave, and that it is not exact, but only represents a zero–order ray approximation. The ray method can be used to obtain a surprisingly simple expression of the elementary ray–theory elastodynamic Green function for any selected *elementary wave* (e.g., any multiply reflected/transmitted wave), propagating in an arbitrary 3–D anisotropic or isotropic inhomogeneous layered structure along any *ray* connecting the point source S and the receiver at R. In the frequency domain, the expression reads:

$$G_{in}(R,t;S,t_0) = \frac{g_n(S)g_i(R)\exp[\mathrm{i}T^{\mathrm{G}}(R,S)+\mathrm{i}\omega\tau(R,S)]}{4\pi[\varrho(S)\varrho(R)\mathcal{C}(S)\mathcal{C}(R)]^{1/2}\mathcal{L}(R,S)}\mathcal{R}^C \quad . \tag{52}$$

Here g_i (resp. g_n) is the i–th (resp. the n–th) Cartesian component of the relevant polarization vector $\mathbf{g}$, $\mathcal{C}$ is the phase velocity, and ϱ the density. Further, $\tau(R,S)$ is the travel time from S to R, $T^{\mathrm{G}}(R,S)$ the complete phase shift due to caustics along the ray from S to R (see Section 6.3), $\mathcal{L}(R,S)$ relative geometrical spreading along the ray from S to R, see (48), and $\mathcal{R}^C$ the product of all normalized displacement R/T coefficients at all points of contact of the ray with interfaces or with the Earth's surface, between S and R.

Alternatively, we can use (49) and replace the factor $[\mathcal{C}(S)\mathcal{C}(R)]^{1/2}\mathcal{L}(R,S)$ by the factor $[\mathcal{U}(S)\mathcal{U}(R)]^{1/2}\bar{\mathcal{L}}(R,S)$ in (52).

Similar expressions for the elementary ray–theory elastodynamic Green function are also valid in isotropic structures. Actually, they are formally a little more complicated than (52), due to the degeneracy of S waves in isotropic media. It is suitable to express the relevant expression in matrix form. The expressions are valid for source and receiver, situated at arbitrary locations, including interfaces and the Earth's surface.

Before we compute the ray–theory Green function (52), we have to perform ray tracing and dynamic ray tracing along the ray from S to R. By ray tracing, we determine $\tau(R,S)$, and by dynamic ray tracing $\mathcal{L}(R,S)$ and $T^{\mathrm{G}}(R,S)$. Equation (52) is valid for points S and R, situated in the smooth parts of the model. For points S and R situated on an interface and/or the Earth's structure, polarization vectors $g_n(S)$ and $g_i(R)$ must be substituted by more complicated expressions.

The elementary ray–theory elastodynamic Green function may be used in similar applications as the exact elastodynamic Green function. Of course, it may be useful to consider several elementary ray–theory Green functions, corresponding to most energetic elementary waves propagating along different rays from point S to point R (direct P, S1 and S2 waves, various primary reflected waves, etc.), and thus form the ray–theory Green function.

6.2. Reciprocity of the ray–theory Green function

It is possible to prove that the elementary ray–theory elastodynamic Green function is reciprocal in the following sense:

$$G_{in}(R,t;S,t_0) = G_{ni}(S,t;R,t_0) \quad . \tag{53}$$

Here $G_{ni}(S,t;R,t_0)$ corresponds to the backward propagation, from R to S. The reciprocity relation (53) is valid for any multiply reflected (possibly converted) elementary wave propagating in a 3–D anisotropic laterally varying layered structure with any number of reflection/transmission points. Similarly, reciprocity (53) remains valid for the wavefield, composed of any number of elementary waves propagating from S to R along different ray trajectories. Reciprocity relation (53) follows simply from (52), if we take into account that travel time $\tau(R,S)$, relative geometrical spreading $\mathcal{L}(R,S)$, the complete phase shift due to caustics $T^{\mathrm{G}}(R,S)$ and the product of normalized R/T displacement coefficients $\mathcal{R}^C$ are reciprocal. The reciprocity of travel time $\tau(R,S)$ is obvious. The reciprocity of $\mathcal{L}$ follows from the symplecticity of the DRT propagator matrix, and the reciprocity of $\mathcal{R}^C$ was proved by Chapman (1994). It should be noted that the product of the conventional displacement R/T coefficients is not reciprocal. For the reciprocity of the phase shift due to caustics see Section 6.3. See also Kendall, Guest & Thomson (1992).

Note that reciprocity relation (53) is valid only for the single force point source. It is not valid for other types of point sources, e.g., for the explosive source (centre of dilatation). The reciprocity for a dilatational point source would be valid if pressure is recorded at the receiver, instead of the displacement vector components.

6.3. Phase shift of the Green function due to caustics. KMAH index

In anisotropic inhomogeneous media, the complete phase shift due to caustics $T^{\mathrm{G}}(R,S)$ is given by the relation,

$$T^{\mathrm{G}}(R,S) = T^{\mathrm{C}}(R,S) + \tfrac{1}{2}\pi\sigma_0(S) = -\tfrac{1}{2}\pi(k(R,S) - \sigma_0(S)) \quad . \tag{54}$$

Here $\sigma_0(S)$ corresponds to the phase shift of the point source, and $T^{\mathrm{C}}(R,S)$ is the phase shift due to caustics outside the source, between S and R. Phase shift $\sigma_0(S)$ depends on the form of the ray–velocity surface at source S and on the direction of the ray at S. For rays corresponding to the forward branch of the ray–velocity surface, $\sigma_0 = 0$ (similarly as in isotropic media). For rays corresponding to the reverse branch of the ray–velocity surface, $\sigma_0 = 1$ or even $\sigma_0 = 2$ (Pšenčík & Teles, 1996). The quantity $k(R,S)$, called the KMAH index, is an integer, changing only when the ray under consideration passes through a caustic point.

In isotropic media, the slowness surface is always convex, so that $\sigma_0(S) = 0$. Moreover, the KMAH index $k(R,S)$ is always positive. The increment of the KMAH index equals unity at a line caustic, and two at a point caustic. The caustic points, and consequently $k(R,S)$, can be found as a by–product of DRT.

We can also introduce the KMAH index for anisotropic media $k^{\mathrm{G}}(R,S)$ by the relation

$$k^{\mathrm{G}}(R,S) = k(R,S) - \sigma_0(S) \quad . \tag{55}$$

In this section, we consider the KMAH index in the form of (55). It is possible to prove that the KMAH index (55) is reciprocal in the following sense

$$k^{\mathrm{G}}(R,S) = k^{\mathrm{G}}(S,R) \quad , \tag{56}$$

Here $k^{\mathrm{G}}(R,S)$ corresponds to the forward ray, and $k^{\mathrm{G}}(S,R)$ to the reverse ray. For more details refer to Klimeš (1997), Chapman (2004, p. 176).

Unlike in isotropic media, where the increment of the KMAH index is always positive, the increment of the KMAH index of S waves in anisotropic media may be either positive, or negative, depending on the convexity or non–convexity of the slowness surface. Lewis (1965) derived a general phase–shift rule for a point caustic, expressed in terms of the signature of the matrix of the second derivatives of travel time. Garmany (2000) expressed a general phase–shift rule for a line (simple) caustic in terms of the second derivatives of the eigenvalue of the Christoffel matrix with respect to the slowness vector. Bakker (1998) derived the equations for the phase shift corresponding to a general wavefield, due to both line (simple) and point caustics in 3–D anisotropic media. Bakker's rules are expressed in terms of the second derivatives of the eigenvalue of the Christoffel matrix with respect to the slowness vector.

Here, a simple example is given to illustrate simultaneously the positive and negative phase shifts in anisotropic media, and the reciprocity (56) of the phase shift of the Green function.

6.4. Example of phase shifts

Assume a 2–D homogeneous anisotropic medium with a triplication on the S–wave ray–velocity surface. Select a ray corresponding to the reverse branch of the ray–velocity surface as the axial ray. Assume that the homogeneous anisotropic medium is separated from the homogeneous isotropic medium by a planar interface perpendicular to the axial ray. For the sake of simplicity, we may assume that the (phase and ray) velocities in the direction of the axial ray are similar (or equal) in both media.

The slowness vectors of the paraxial rays, propagating in the anisotropic medium to one side of the axial ray, point to the other side of the axial ray, see Figure 12. On entering the isotropic medium, the paraxial rays are refracted to the direction of the slowness vector and cross the axial ray at the caustic.

Consider 3 points on the axial ray: Point A in the anisotropic medium, point B situated in the isotropic medium between the interface and the caustic corresponding to the point source at A, and point C situated in the isotropic medium beyond the caustic corresponding to the point source at A. The paraxial rays from all 3 points are depicted in Figure 12. The paraxial rays corresponding to the same paraxial slowness vector are parallel in each of the halfspaces.

Accumulation of the phase shift along the four rays between points A, B and C is shown in Figures 13, 14, 15 and 16.

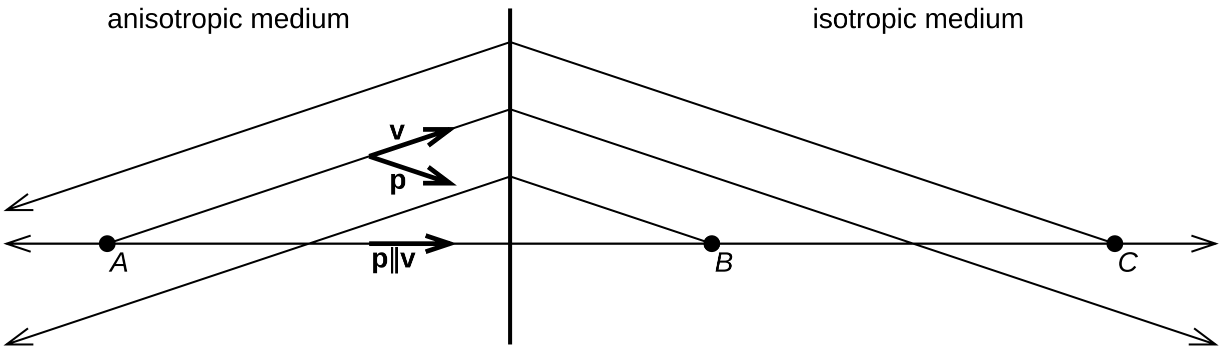

Figure 12. The axial ray and paraxial rays from points A, B, C in the discussed example. In the anisotropic halfspace, paraxial ray–velocity vector $\mathbf{v}$ and slowness vector $\mathbf{p}$ are deflected in opposite directions.

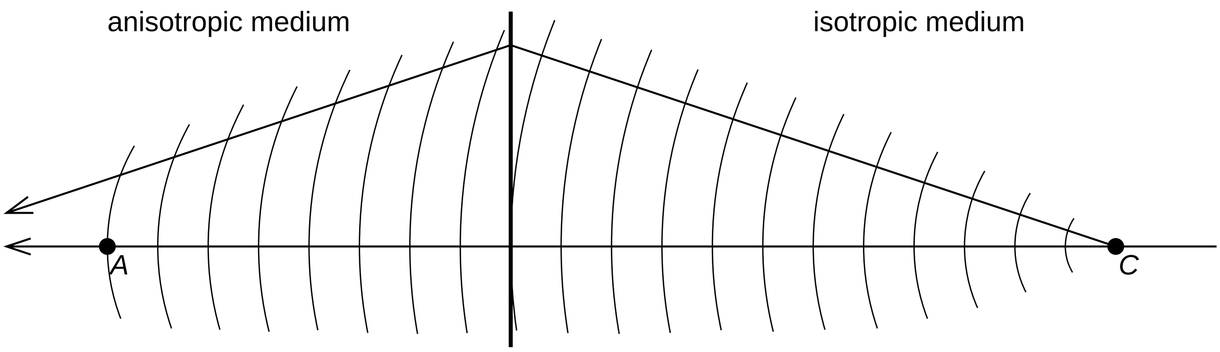

Figure 13. $C \rightarrow A$: There is no caustic along the ray from point C to point A. As the KMAH index is defined zero at the point source in the isotropic medium, the resulting KMAH index at point A is $\mathbf{0}$.

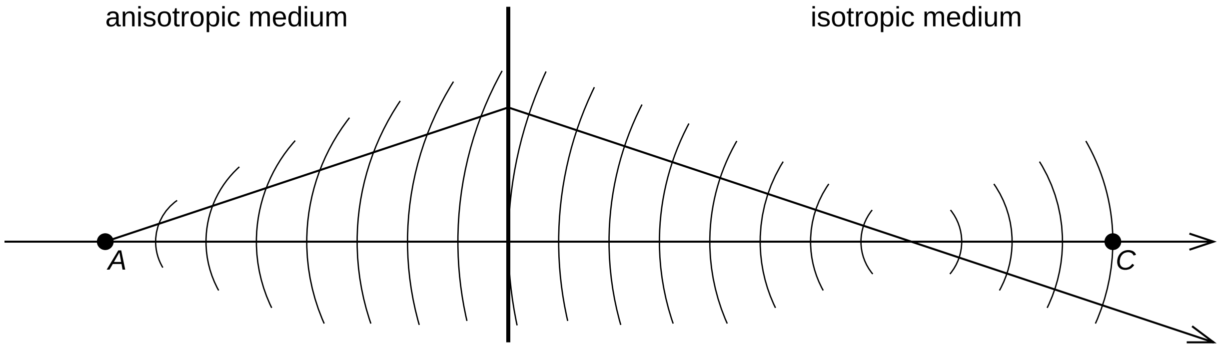

Figure 14. $A \to C$: The initial KMAH index at point A on the reverse wavefront branch equals -1, see (55) with $\sigma_0 = 1$. There is a single caustic along the ray from point A to point C, located in the isotropic medium, causing the KMAH index to increase by $+1$. The resulting KMAH index at point C is then 0.

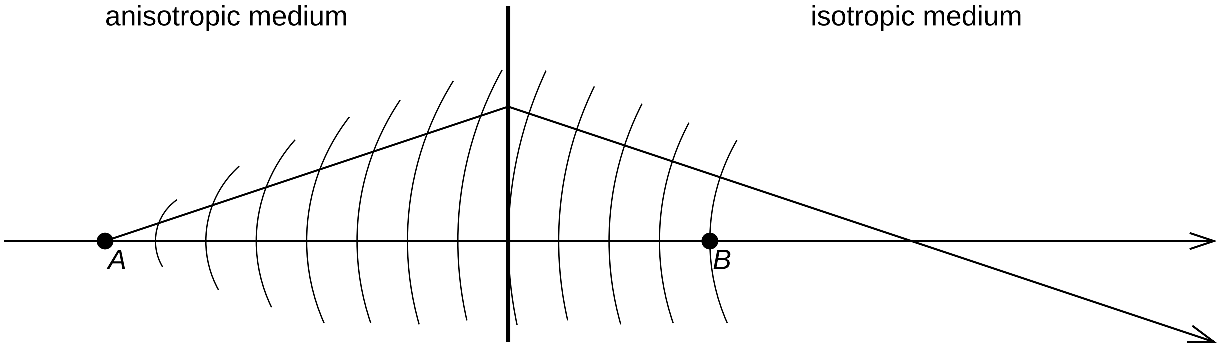

Figure 15. $A \to B$: The initial KMAH index at point A is again -1, but there is no caustic along the ray from point A to point B. The resulting KMAH index at point B is now -1.

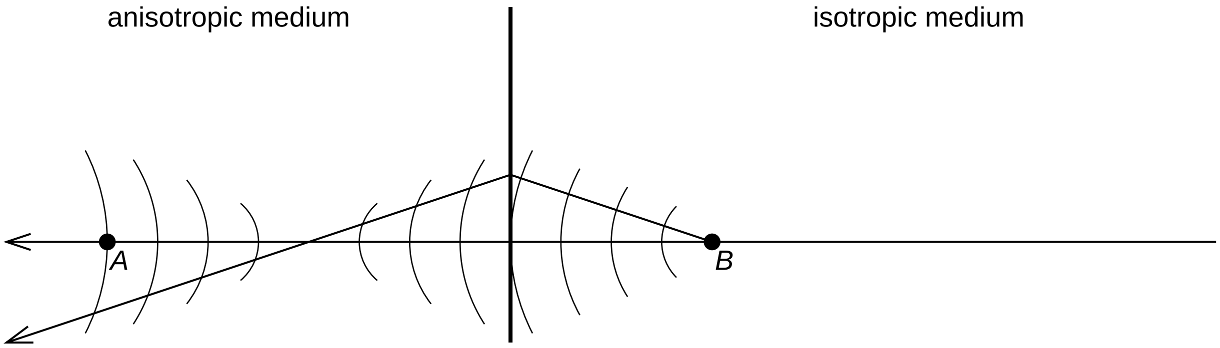

Figure 16. $B \to A$: The initial KMAH index at point B is 0, as for $C \to A$. There is a single caustic along the ray from point B to point A, located in the anisotropic medium on the reverse branch of the ray–velocity surface, causing the KMAH index to decrease by 1. The resulting KMAH index at point A is then -1.

7. Ray–theory seismograms

Let us consider an elementary wave, described in the frequency domain by equation (1), propagating in a perfectly elastic, inhomogeneous, anisotropic or isotropic, layered structure. Using the Fourier transform, we easily obtain the real–valued, time–domain alternative of (1), called the elementary ray–theory seismogram, or elementary ray–theory signal:

$$u_i(x_n,t) = \mathrm{Re}[U_i(x_n)F(t-\tau(x_i))] \ , \quad i=1,2,3 \quad . \tag{57}$$

Here $F(\zeta)$ is the analytical signal, given by relation

$$F(\zeta) = x(\zeta) + \mathrm{i}g(\zeta) \quad . \tag{58}$$

The real–valued function $x(\zeta)$ may be arbitrary, and may represent, for example, the source–time function. For simplicity, we shall refer to it as the input signal. The real–valued function $g(\zeta)$ is the Hilbert transform of $x(\zeta)$. Once the input signal $x(\zeta)$ is known, the analytical signal $F(\zeta)$ can be determined. Two simple examples of the analytical signals are as follows: (a) For $x(\zeta) = \delta(\zeta)$, where $\delta(\zeta)$ is the Dirac delta function, $F(\zeta) = \delta(\zeta) - \mathrm{i}/\pi\zeta$. (b) For $x(\zeta) = \cos(\zeta)$, the analytical signal is $F(\zeta) = \exp[-\mathrm{i}\zeta]$.

We call $x(\zeta)$ the causal input signal if $x(\zeta) = 0$ for $\zeta < 0$. It is well–known that the Hilbert transform $g(\zeta)$ of a causal input signal $x(\zeta)$ is not causal. Thus, the elementary ray–theory signal $u_i(x_n,t)$ *is in principle non–causal*, even when the input signal $x(\zeta)$ is causal.

The situation is simple if $U_i(x_n)$ is real–valued. Equation (57) then yields

$$u_i(x_n,t) = U_i(x_n)\, x(t-\tau(x_i)) \quad , \tag{59}$$

and the form of the elementary ray–theory signal $u_i(x_n,t)$ remains preserved along the whole ray. Further, $u_i(x_n,t)$ is causal along the whole ray if $x(\zeta)$ is causal.

The situation is quite different for complex–valued $U_i(x_n)$. Note that $U_i(x_n)$, real–valued at the initial point of the ray, may become complex–valued in other parts of the ray. This is mainly due to the caustic points along the ray, and due to the postcritical reflections at interfaces. Thus, without considering complex–valued $U_i(x_n)$, the ray method would be incomplete.

Let us now briefly discuss certain important properties of elementary ray–theory signals (57), valid both for $U_i(x_n)$ real–valued and complex–valued and for non–dissipative media.

1) Analytical signal $F(\zeta)$ is preserved along the whole ray.

2) The normalized envelope of the elementary ray–theory signal (57) is preserved along the whole ray. The normalization is taken with respect to the maximum value of the envelope. Thus, the normalized envelope is the same at the initial point of the ray, at interfaces, at caustic points, and at the receiver. It is always equal to the normalized envelope of the analytical signal.

3) Although the elementary ray–theory signal is not, in principle, causal, we may say that it is *effectively causal*, particularly for smooth, absolutely integrable input signals $x(\zeta)$. The explanation of effective causality is as follows: Assume that the ray–theory travel time corresponds to the maximum of the normalized envelope, and denote it at a given point of the ray by $t_{\max}$. The effective first arrival travel time $t_{\min}$ may then be taken at a time, for which the value of the normalized envelope is negligibly

small, for example 0.001. Property 2 then guarantees that $t_{\max} - t_{\min}$ is preserved along the whole ray, including the interfaces, caustic points, etc.

It might seem useful to consider very narrow causal input signals $x(\zeta)$ in the ray method, for example signals close to $x(\zeta) = \delta(\zeta)$. For such signals, however, the envelope is broad (due to the broad Hilbert transform), and $t_{\max} - t_{\min}$ is large. Thus, the effect of non–causality is particularly strong in this case. A very suitable option is to use smooth, non–causal, high–frequency input signal $x(\zeta)$, for example the Gabor signal. Fortunately, the three–parametric Gabor signal can easily simulate most signals $x(\zeta)$ known from seismological and seismic exploration records. The effects of changing the phase shifts of $U_i(x_n)$ on the ray–theory signals $u_i(x_n, t)$ are then expressed only in the changes of $u_i(x, t)$ under its fixed envelope.

8. Effects of interfaces

Reflection or transmission of a wave at an interface of the first order (across which parameters of the medium change abruptly) leads to abrupt changes of most quantities calculated along a ray. Only the ray itself and travel time remain continuous. Other quantities like the direction of the ray, the direction of the corresponding slowness vector, values of all the quantities calculated in dynamic ray tracing and, consequently, geometrical spreading and the ray amplitude change discontinuously. Let us briefly describe the transformation of the individual quantities.

Due to the continuity of the travel times of incident and generated waves along the interface, specifically due to the continuity of their first–order derivatives, the slowness vectors of the incident and generated waves must have the same projection into the plane tangent to the interface at the point of incidence. This equality of the tangential projections of the slowness vectors is but the expression of Snell's law. It holds in isotropic as well as anisotropic media. The determination of the slowness vector of a generated wave thus reduces to the determination of the component of the slowness vector into the normal n_i to the interface. The slowness vector $p_i^{(m)}$ of a generated wave can thus be expressed as

$$p_i^{(m)} = [p_i - n_i(n_k p_k)] + \xi^{(m)} n_i \quad . \tag{60}$$

Here p_i is the slowness vector of the incident wave, n_i is the unit normal to the interface and $\xi^{(m)}$ is the sought component of the slowness vector into normal n_i. The term in the square brackets is the tangential vectorial component of the slowness vector, which is the same for incident and generated waves. The slowness vector of any generated wave propagating in an anisotropic medium must satisfy the condition of solvability of the Christoffel equation

$$\det(a_{ijkl} p_j p_l - \delta_{ik}) = 0 \quad . \tag{61}$$

Here a_{ijkl} again denotes the tensor of density–normalized elastic moduli of the medium in which the generated wave propagates. By inserting (60) into (61) we arrive at the sixth–order polynomial equation for $\xi^{(m)}$. Since the coefficients of the polynomial equation are real–valued, its complex roots appear in complex conjugate pairs. Real roots correspond to subcritical incidence, complex roots indicate postcritical incidence of the incident wave. One equation (61) with parameters a_{ijkl} specifying the medium, in which incident and reflected waves propagate, holds for the three reflected waves; another equation (61) with parameters a_{ijkl} specifying the medium, in which transmitted waves propagate, holds for the three transmitted waves. Since each equation has six roots, only three of them have physical meaning and the remaining three must be eliminated. Of the real–valued roots only those are acceptable for which the corresponding ray–velocity vectors (parallel to the time–averaged energy flux vectors) point into the medium, in which the generated wave propagates. Of the complex conjugate roots only those are acceptable, which correspond to inhomogeneous waves (waves with complex–valued slowness vector, whose real and imaginary parts make a non–zero angle) decreasing exponentially with distance from the interface. The limiting case, in which real–valued roots become complex–valued, is the case of critical incidence, in which the ray–velocity vector is parallel to the interface. Note that in anisotropic media, the slowness vector of the wave, generated by a critically incident wave, is generally not parallel to the interface. Since the ray–velocity vector and the slowness vector have generally different directions in anisotropic media, it may happen that the slowness and ray–velocity vectors of a generated wave point to opposite sides of the interface.

If the generated wave propagates in an isotropic medium, condition (61) simplifies. The slowness vector must satisfy the corresponding eikonal equation, from which its component into the normal to the interface can be determined. In this case, two roots are obtained for a single wave and one of them must be eliminated. The conditions of elimination are the same as in the case of an anisotropic medium. Since in an isotropic medium the ray–velocity vector and the slowness vector are parallel, the ray–velocity vector can be substituted by the slowness vector in the elimination process.

Thus, to determine slowness vectors of generated waves in an anisotropic medium, a polynomial equation of the sixth order must be solved numerically. It yields the normal components of slowness vectors of all waves generated on the relevant side of the interface. In isotropic media, simple analytic formulae exist for the determination of the slowness vector of a given generated wave alone.

Equation (60) clearly indicates that the slowness vectors of generated waves are situated in the plane of incidence defined by the slowness vector of the incident wave and the normal to the interface. In anisotropic media, the ray–velocity vectors may point out of the plane of incidence. In isotropic media, the ray–velocity vectors of all generated waves are situated in the plane of incidence.

The transformation relations at an interface for the solutions of the dynamic ray tracing system also follow from the continuity of the travel times of generated and incident waves. In this case, second–order travel–time derivatives are used, which means that the transformation is affected by the first–order derivatives of the parameters of the medium and by the curvature of the interface at the point of incidence. Due to this dependence, the transformation relations for dynamic ray tracing must be applied even at interfaces of the second order, i.e. interfaces, at which gradients of elastic parameters change abruptly while the parameters themselves are continuous.

The ray amplitudes of generated waves at the point of incidence are obtained from the ray amplitude of the incident wave by multiplying it by the corresponding reflection or transmission coefficients. In the zero–order ray approximation, the coefficients correspond to the plane–wave reflection/transmission coefficients. Mostly displacement reflection/transmission coefficients are used. Normalized coefficients, which have the convenient property of reciprocity, see Section 6.2, can be used instead of them. The coefficients depend on the parameters of the medium, the density and the angle of incidence. In addition to the coefficients, which are applied at the points of reflection or transmission at interfaces, the conversion coefficients should be used at sources and receivers when these are situated at interfaces.

9. Ray–theory perturbations

Perturbation methods play an important role in ray methods. They can be used for fast but approximate solution of forward problems in complicated models. Perturbation methods play an equal or even more important role in inverse problems.

There is a great variety of perturbation methods in use, most of them being based on the following assumptions. We assume that a model, in which we wish to study wave propagation, differs only little from another model called the *background* or *reference model.* The solution in the perturbed model can then be sought in the form of a power series in the deviations of the perturbed and reference models. If only the first term of the series is considered, which is often the case, we speak of the *first–order perturbation expansion.* The perturbed and reference models may both be isotropic or anisotropic, or the reference model may be isotropic and the perturbed model anisotropic. Various quantities may be sought in this way.

Most often the quantity sought is the travel time. The corresponding first–order formula for the isotropic reference and perturbed models has been known for a long time, see Aki & Richards (1981). For the anisotropic perturbed model, the formulae for the first–order travel–time perturbations can be found in Červený (1982), Červený & Jech (1982), Hanyga (1982), Jech & Pšenčík (1989), Chapman & Pratt (1992). In these approaches a reference ray is traced in the reference isotropic or anisotropic medium, and the first–order travel–time correction is sought by integration along the reference ray. Even higher–order travel–time perturbations can be calculated. Farra (1999) proposed a procedure for calculating the second–order travel–time perturbations. Her procedure is based on the knowledge of the first–order perturbations of a reference ray. Recently, developing ideas of Babich, Buldyrev & Molotkov (1985), Klimeš (2002a) proposed a procedure for calculating higher–order perturbations (or derivatives) of the travel time along a reference ray in the reference medium. Refer to Section 9.2 for more details.

The perturbation theory can be used not only to determine the travel time, but the whole wavefield. Farra & Madariaga (1987) proposed such an approach for an isotropic reference and perturbed medium. The approach consists in computating perturbed rays and related dynamic ray tracing by integrating along a reference ray. The approach was later extended also to anisotropic media, see, e.g., Farra (1989), Nowack & Pšenčík (1991). In all the above–mentioned approaches, it is necessary to trace rays in the reference medium, and then to calculate the perturbations of the sought quantities along these rays. Recently, Farra & Pšenčík (2004) and Farra (2004) proposed an alternative approach of calculating rays in inhomogeneous weakly anisotropic media, in which the deviation of anisotropy from isotropy is considered to be of the first order. In ray tracing system (9), they substituted the exact Hamiltonian by its first-order approximation. In this way, their ray tracing yields "first–order rays" and corresponding first–order travel times without calculating the reference rays. If it becomes necessary to increase the accuracy of the calculated travel times, second–order travel–time perturbation derivatives can be easily calculated by integrating along the first–order rays. In weakly anisotropic media, this ray tracing represents a natural generalization of ray tracing for isotropic media. Indeed, for vanishing anisotropy, the first–order ray tracing reduces to the exact "isotropic" ray tracing. The first–order ray tracing equations are, as in isotropic media, different for P and S waves. In each case, the ray tracing equations depend on 15 weak anisotropy parameters. Weak anisotropy parameters represent a generalization of linearized Thomsen's (1986) parameters for anisotropy of arbitrary symmetry, see, e.g., Farra & Pšenčík (2003). There are 9 parameters, which are common for P and S waves, and 6 parameters, which are related only to P waves, and 6 parameters, which are related only to S waves. During numerical tests using a configuration and models typical for seismic exploration, it was found that the relative errors of the travel times including the second–order correction were well under 0.05% for anisotropy of about 8%, and they did not exceed 0.3% for anisotropy of about 20%.

A possibility of how to calculate the whole wavefield using the perturbation theory is the *generalized Born scattering*, see Coates & Chapman (1991), Chapman (2004). This approach is based on the first–order Born approximation, in which the exact Green function in the reference medium is substituted by the ray–theory Green function. The errors caused by this substitution are introduced into the scattering integral so that scattering occurs not only from perturbations of the medium, but also from the errors.

In this contribution, we concentrate on the perturbations of travel time only. Similar formulae as for travel time may also be derived for other ray–theory quantities.

9.1. Model parameters

We assume that the Hamiltonian is a function of *phase–space coordinates*, composed of spatial coordinates x_i and slowness–vector components p_j, and of any number of *model parameters* f_κ,

$$\mathcal{H} = \mathcal{H}(x_i, p_j, f_\kappa) \quad . \tag{62}$$

The travel time and other quantities calculated in the model using this Hamiltonian are then functions of *spatial coordinates* x_i and of model parameters f_κ,

$$\tau = \tau(x_i, f_\kappa) \quad . \tag{63}$$

In this section, we index the model parameters by lower–case Greek subscripts. Model parameters are the parameters to be perturbed. They may have various meanings. For example, model parameters f_κ may parametrize the model. Parameters f_κ may be the B–spline coefficients of the functions describing the model. However, the applicability of model parameters is more general. For example, there are two S waves with the two respective Hamiltonians corresponding to the anisotropic ray theory in anisotropic media. Instead of selecting one of these two Hamiltonians, we may consider a one–parametric set of Hamiltonians, parametrized by a parameter f_μ, such that we obtain the Hamiltonian corresponding to one S wave for $f_\mu = -1$, the Hamiltonian corresponding to the other S wave for $f_\mu = +1$, and the averaged Hamiltonian of both S waves corresponding to the anisotropic common ray tracing (Bakker, 2002) for $f_\mu = 0$. Another example: Instead of using Hamiltonian $\mathcal{H} = \mathcal{H}(x_i, p_j, f_\kappa)$ and perturbation expansion from $f_\kappa = f_\kappa^1$ to $f_\kappa = f_\kappa^2$, we may define a new one–parametric Hamiltonian $\tilde{\mathcal{H}} = \tilde{\mathcal{H}}(x_i, p_j, f)$,

$$\tilde{\mathcal{H}}(x_i, p_j, f) = \mathcal{H}(x_i, p_j, f_\kappa^1) + [\mathcal{H}(x_i, p_j, f_\kappa^2) - \mathcal{H}(x_i, p_j, f_\kappa^1)]\, f \tag{64}$$

and apply the perturbation expansion from $f = 0$ to $f = 1$.

We shall refer to the partial derivatives with respect to model parameters f_κ as the *perturbation derivatives.*

We denote the partial derivatives with respect to spatial coordinates x_i by lower–case Roman subscripts following a comma, and the *perturbation derivatives by lower–case Greek subscripts* following a comma. In phase space, we denote the partial derivatives with respect to components p_j of the slowness vector by lower–case Roman superscripts following a comma. For instance,

$$\mathcal{H}^{,ab\ldots f}_{,ij\ldots n\alpha\ldots\nu} = \frac{\partial}{\partial f_\alpha} \cdots \frac{\partial}{\partial f_\nu} \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} \cdots \frac{\partial}{\partial x_n} \frac{\partial}{\partial p_a} \frac{\partial}{\partial p_b} \cdots \frac{\partial}{\partial p_f} \mathcal{H} \tag{65}$$

denotes the perturbation derivatives of the partial *phase–space derivatives* of the Hamiltonian. The perturbation derivatives are taken at fixed phase–space coordinates x_i, p_j, and the partial phase–space derivatives are calculated at fixed model parameters f_κ. Similarly,

$$\tau_{,ij\ldots n\alpha\ldots\nu} = \frac{\partial}{\partial f_\alpha} \cdots \frac{\partial}{\partial f_\nu} \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} \cdots \frac{\partial}{\partial x_n} \tau \tag{66}$$

denotes the perturbation derivatives of the *spatial derivatives* of travel time τ. The perturbation derivatives are taken at fixed spatial coordinates x_i, and the partial spatial derivatives are calculated at fixed model parameters f_κ. Analogously for other quantities, e.g., amplitude $A = A(x_i, f_\kappa)$.

9.2. Perturbation of travel time

The first–order perturbation derivatives of travel time are very important in (a) linearized inversion of travel times; (b) approximate solution of the complex–valued eikonal equation in the vicinity of real–valued rays; (c) common–ray approximations of the coupling ray theory. In all these methods, the error of the first–order perturbation expansion should be estimated and controlled. The error of the first–order perturbation expansion can be approximated by the neglected second–order term. It is thus of principal importance in all the above applications to estimate the second–order perturbation derivatives of travel time. A brief note on the application of the second–order perturbation derivatives of travel time to the linearized inversion of travel times is given in Section 9.3. For estimating the errors of common–ray approximations of the coupling ray theory using the second–order perturbation derivatives of travel time, refer to Klimeš & Bulant (2004).

The perturbation derivatives of the travel time of all orders can be calculated by simple numerical quadratures along the unperturbed ray. The perturbation derivatives of the spatial derivatives of travel time can also be calculated by analogous numerical quadratures along the unperturbed ray. As regards the general equations for the perturbation derivatives of all orders, refer to Klimeš (2002a). Here we only present the equations for the first–order and second–order perturbation derivatives of travel time, expressed in terms of a general Hamiltonian.

The individual perturbation derivatives of travel time depend on the form of the Hamiltonian. For different Hamiltonians, we obtain different perturbation expansions of travel time. Some of these perturbation expansions may be more accurate for some applications than other perturbation expansions (Klimeš, 2002a).

9.2.1. First–order perturbation derivatives of travel time

Equation

$$\tau_{,\alpha}(\gamma_3) = \tau_{,\alpha}(\gamma_3^0) + \int_{\gamma_3^0}^{\gamma_3} \mathrm{d}\gamma_3 \; (-\mathcal{H}_{,\alpha}) \tag{67}$$

for the first–order perturbation derivatives of travel time is well–known (e.g., Červený, 2001, eq. 3.9.6). The integration variable γ_3 along the ray is determined by the form of the Hamiltonian. For the Hamiltonians considered in this paper, the integration variable is the travel time, $\gamma_3 = \tau$.

9.2.2. First–order perturbation derivatives of the travel–time gradient

The first–order perturbation derivatives of the travel–time gradient may be calculated using equation

$$\tau_{,i\alpha} = T_{a\alpha} \; Q^{-1}_{ai} \quad , \tag{68}$$

where Q^{-1}_{ai} are the elements of the matrix inverse to matrix $Q_{ia} = \partial x_i / \partial \gamma_a$ of geometrical spreading in Cartesian coordinates. The perturbation derivatives of the covariant spatial derivatives $T_{a\alpha}$ of travel time with respect to ray coordinates can be obtained by quadrature

$$T_{a\alpha}(\gamma_3) = T_{a\alpha}(\gamma_3^0) + \int_{\gamma_3^0}^{\gamma_3} \mathrm{d}\gamma_3 \; (-\mathcal{H}_{,i\alpha} \; Q_{ia} - \mathcal{H}^{;i}_{,\alpha} \; P_{ia}) \quad . \tag{69}$$

The matrix $P_{ia} = \partial p_i / \partial \gamma_a$ of the transformation from ray coordinates to the components of the slowness vector is calculated, together with the matrix Q_{ia} of geometrical spreading, by dynamic ray tracing.

9.2.3. Second–order perturbation derivatives of travel time

The second–order perturbation derivatives of travel time can be calculated by quadrature

$$\tau_{,\alpha\beta}(\gamma_3) = \tau_{,\alpha\beta}(\gamma_3^0) + \int_{\gamma_3^0}^{\gamma_3} \mathrm{d}\gamma_3 \; K_{\alpha\beta} \quad , \tag{70}$$

where the integration kernel is given by

$$K_{\alpha\beta} = -\mathcal{H}_{,\alpha\beta} - T_{a\alpha} \; Q^{-1}_{ar} \mathcal{H}^{;r}_{,\beta} - T_{a\beta} \; Q^{-1}_{ar} \mathcal{H}^{;r}_{,\alpha} - T_{a\alpha} \; Q^{-1}_{ar} T_{b\beta} \; Q^{-1}_{bs} \mathcal{H}^{;rs} \quad . \tag{71}$$

Quadrature (70) is applicable only to ray segments with the regular matrix Q_{ia} of geometrical spreading. Along ray segments with singular matrix Q_{ia} but regular matrix P_{ia}, equation (70) may be replaced by equation

$$\tau_{,\alpha\beta}(\gamma_3) = \tau_{,\alpha\beta}(\gamma_3^0) + \left[T_{a\alpha} \; P^{-1}_{ar} T_{b\beta} \; Q^{-1}_{br} \right]_{\gamma_3^0}^{\gamma_3} + \int_{\gamma_3^0}^{\gamma_3} \mathrm{d}\gamma_3 \; K^{\mathrm{caust}}_{\alpha\beta} \quad , \tag{72}$$

where the integration kernel is given by

$$K^{\mathrm{caust}}_{\alpha\beta} = -\mathcal{H}_{,\alpha\beta} + T_{a\alpha} \; P^{-1}_{ar} \mathcal{H}_{,r\beta} + T_{a\beta} \; P^{-1}_{ar} \mathcal{H}_{,r\alpha} - T_{a\alpha} \; P^{-1}_{ar} T_{b\beta} \; P^{-1}_{bs} \mathcal{H}_{,rs} \quad . \tag{73}$$

9.3. Optimizing model updates during linearized inversion of travel times

The size of the individual iterations in the linearized inversion of travel times should be controlled by means of the second–order perturbation derivatives of travel time. This control can be achieved by including the square of the relevant Sobolev norm of the model update into the objective function for the linearized inversion of travel times. The Sobolev norm of two functions is the square root of the Sobolev scalar product of these two functions. The Sobolev scalar product of two functions is a linear combination of the L_2 Lebesgue scalar products of the zero–order, first–order, second–order or higher–order derivatives of the functions, see, e.g., Tarantola (1987).

Since the second–order perturbation derivatives of travel time are proportional to the square of the first–order perturbation derivatives of the slowness or velocity gradient (Klimeš, 2002a), the error of a single iteration of the linearized inversion of travel times is roughly proportional to the square of the Sobolev norm composed of the first derivatives of the slowness or velocity update. We should thus include the square of the Sobolev norm composed of the first derivatives of the slowness or velocity update into the objective function (Klimeš, 2002c).

The weighting factor of this Sobolev norm in the objective function may iteratively be adjusted. For this purpose, we only need to calculate the homogeneous second–order perturbation derivative of travel time in the direction of the model update, and the mixed second–order perturbation derivative of travel time in the direction of the model update and in the direction of the derivative of the model update with respect to the weighting factor. The calculation of these second–order perturbation derivatives of travel time costs 4 numerical quadratures along each two–point ray. The results are then used to correct the estimate of the weighting factor and to calculate a new model update, new second–order perturbation derivatives of travel time, and a new weighting factor, until reaching the nearly–optimum size of the model update for the rapid convergence of the travel–time inversion.

10. Coupling ray theory for S waves

There are two different high–frequency asymptotic ray theories: the *isotropic ray theory* assuming equal velocities of both S waves and the *anisotropic ray theory* assuming both S waves strictly decoupled. In the isotropic ray theory, the S–wave polarization vectors do not rotate about the ray, whereas in the anisotropic ray theory they may rotate rapidly about the ray, see Figure 17. Thomson, Kendall & Guest (1992) demonstrated analytically that the high–frequency asymptotic error of the anisotropic ray theory is inversely proportional to the second or higher root of the frequency if a ray passes through the point of equal S–wave eigenvalues of the Christoffel matrix.

In weakly anisotropic models, at moderate frequencies, the S–wave polarization tends to remain unrotated about the ray, but is partly attracted by the rotation of the eigenvectors of the Christoffel matrix. The intensity of the attraction increases with frequency. This behaviour of the S–wave polarization is described by the *coupling ray theory* proposed by Coates & Chapman (1990) and (in the quasi–isotropic approximation) by Pšenčík (1998). The coupling ray theory is applicable at all degrees of anisotropy, from isotropic models to considerably anisotropic ones. The frequency–dependent coupling ray theory is the generalization of both the zero–order isotropic and anisotropic ray theories and provides continuous transition between them. The numerical algorithm for calculating the frequency–dependent complex–valued S–wave polarization vectors of the coupling ray theory has been designed by Bulant & Klimeš (2002).

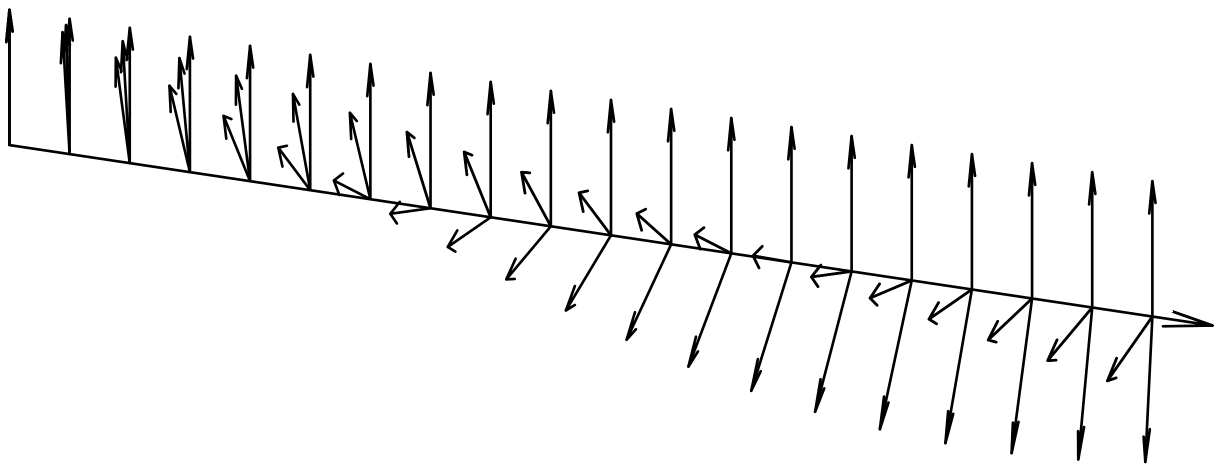

Figure 17. S–wave polarization along a ray: Isotropic–ray–theory polarization points upwards. Anisotropic–ray–theory polarization points downwards on the right–hand side. Coupling–ray–theory polarization points between the isotropic–ray–theory and anisotropic–ray–theory polarizations.

The results of the isotropic ray theory, anisotropic ray theory and coupling ray theory have been compared with the exact solution in the "simplified twisted crystal" model designed by Vavryčuk (1999), analytically by Klimeš (2004a) and numerically by Bulant et al. (2004). The quasi–isotropic approximation of the coupling ray theory has been numerically compared with the isotropic ray theory, anisotropic ray theory and reflectivity method in a more realistic 1–D model by Pšenčík & Dellinger (2001).

10.1. Coupling equation for S waves

Assume a *reference ray* in phase space, parametrized by reference travel time τ, with reference slowness vectors $p_i(\tau)$ known at all its points $x_j(\tau)$. Using the reference slowness vector, we can calculate the reference Christoffel matrix

$$\Gamma_{jk}(\tau) = p_i(\tau)\, a_{ijkl}(x_m(\tau))\, p_l(\tau) \tag{74}$$

and its eigenvectors $g_{i1}(\tau)$, $g_{i2}(\tau)$, $g_{i3}(\tau)$ along the reference ray, see (6). Assume that eigenvectors $g_{i1}(\tau)$ and $g_{i2}(\tau)$ correspond to S waves and that they vary continuously along the reference ray. Continuity is not required in regions where the corresponding two eigenvalues are equal. Let us denote by $\tau_1(\tau)$ and $\tau_2(\tau)$ the travel times corresponding to eigenvectors $g_{i1}(\tau)$ and $g_{i2}(\tau)$, respectively. They may be approximated by quadratures along the unperturbed reference ray,

$$\frac{\mathrm{d}\tau_1}{\mathrm{d}\tau} = [\Gamma_{jk} g_{j1} g_{k1}]^{-\frac{1}{2}} \quad , \qquad \frac{\mathrm{d}\tau_2}{\mathrm{d}\tau} = [\Gamma_{jk} g_{j2} g_{k2}]^{-\frac{1}{2}} \tag{75}$$

(Klimeš, 2003a, eqs. 43 and 65). The coupling–ray–theory solution u_i of the elastodynamic equations may then be expressed, for S waves, as the linear combination

$$u_i = A\, g_{iM}\, a_M \exp(\mathrm{i}\omega\overline{\tau}) \tag{76}$$

of the anisotropic–ray–theory solutions, where

$$\overline{\tau}(\tau) = \tfrac{1}{2}[\tau_1(\tau) + \tau_2(\tau)] \tag{77}$$

is the average travel time, and $A = A(\tau)$ is the complex–valued scalar amplitude in the high–frequency approximation (Červený, 1972), corresponding to the system of reference rays. The coupling–ray–theory (Coates & Chapman, 1990) equation for complex–valued amplitude factors $a_M = a_M(\tau)$ reads (Bulant & Klimeš, 2002, eq. 9)

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \left[\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\frac{\mathrm{d}\varphi}{\mathrm{d}\tau} + \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}\frac{\omega}{2}\frac{\mathrm{d}(\tau_1-\tau_2)}{\mathrm{d}\tau}\right]\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad , \tag{78}$$

where

$$\frac{\mathrm{d}\varphi}{\mathrm{d}\tau} = \frac{\mathrm{d}g_{k1}}{\mathrm{d}\tau}\, g_{k2} = -g_{k1}\frac{\mathrm{d}g_{k2}}{\mathrm{d}\tau} \tag{79}$$

is the angular velocity of the eigenvector rotation.

10.2. Coupling–ray–theory S–wave propagator matrix

Propagator matrix $\mathbf{\Pi}^g$ of equation (78), defined as

$$\Pi^g_{MN}(\tau,\tau_0) = \frac{\partial\, a_M(\tau)}{\partial\, a_N(\tau_0)} \quad , \tag{80}$$

is a complex–valued 2×2 matrix satisfying equation (Bulant & Klimeš, 2002, eq. 11)

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{\Pi}^g = \left[\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\frac{\mathrm{d}\varphi}{\mathrm{d}\tau} + \begin{pmatrix} \mathrm{i} & 0 \\ 0 & -\mathrm{i} \end{pmatrix}\frac{\mathrm{d}\psi}{\mathrm{d}\tau}\right]\mathbf{\Pi}^g \quad , \tag{81}$$

directly following from equation (78). Here

$$\psi(\tau) = \tfrac{1}{2}\omega\, [\tau_1(\tau) - \tau_2(\tau)] \quad . \tag{82}$$

Note that propagator matrix $\mathbf{\Pi}^g$ is symplectic and unitary.

It is difficult to integrate equation (81) by the Runge–Kutta or another numerical method that requires derivative $\mathrm{d}\varphi/\mathrm{d}\tau$ along the reference ray to be calculated, because this derivative is undefined in the singular regions, in which the two eigenvalues of Christoffel matrix (74) are equal. The method of calculating propagator matrix $\mathbf{\Pi}^g$, suitable for equation (81), has been proposed by Červený (2001) and Bulant & Klimeš (2002), with emphasis on numerical implementation. The method does not require the calculation of the angular velocity $\mathrm{d}\varphi/\mathrm{d}\tau$ of the rotation of the eigenvectors of the Christoffel matrix along the reference ray and does not require $\mathrm{d}\varphi/\mathrm{d}\tau$ to be smooth or finite along the reference ray.

The proposed method of solving equation (81) takes advantage of the chain rule. Since $\mathbf{\Pi}^g$ is a propagator matrix satisfying the chain rule, it may be numerically calculated as the product of propagator matrices $\mathbf{\Pi}^g$ corresponding to reasonably small segments of the reference ray (Červený, 2001). Frequency–dependent propagator matrices along the individual small ray segments may be approximated by the method of mean coefficients (Červený, 2001). The accuracy of the proposed algorithm of numerical integration of the coupling equation has been estimated by Bulant & Klimeš (2002). The estimate enables the integration step to be controlled, so that the relative error in the wavefield amplitudes due to the integration is kept below a given limit, which is of principal importance for numerical applications.

10.3. Quasi–isotropic approximations of the coupling ray theory

The coupling ray theory by Coates & Chapman (1990) is applicable at all degrees of anisotropy, but it is often modified by various *quasi–isotropic approximations*. There are many frequently used quasi–isotropic approximations of the coupling ray theory (Klimeš & Bulant, 2004), which impair the accuracy of the coupling ray theory both with increasing frequency and increasing degree of anisotropy. For example, the reference ray may be calculated in different ways (Bakker, 2002; Klimeš & Bulant, 2004), the Christoffel matrix may be approximated by its quasi–isotropic projections onto the plane perpendicular to the reference ray and onto the line tangent to the reference ray (Pšenčík, 1998), travel times corresponding to the anisotropic ray theory may be approximated in several ways, e.g., by the first–order quasi–isotropic perturbation with respect to the density–normalized elastic moduli (Pšenčík, 1998), etc. Most of these quasi–isotropic approximations can be avoided with minimum effort (Bulant & Klimeš, 2002; 2004), except for the *common ray approximation* for S waves.

The effects of the quasi–isotropic approximation of the Christoffel matrix, the quasi–isotropic projection of the Green function and the quasi–isotropic perturbation of travel times have been demonstrated by numerical comparison with the exact solution for the one–way plane–wave propagator matrix in a simple 1–D anisotropic "oblique twisted crystal" model by Bulant & Klimeš (2004). The effect of the quasi–isotropic projection of the Green function has been numerically demonstrated by Bulant & Klimeš (2002) on the example of seismograms calculated by the coupling ray theory in a more realistic 1–D anisotropic model.

When using any of these quasi–isotropic approximations of the coupling ray theory, the errors due to the applied quasi–isotropic approximation should be calculated and checked.

10.3.1. Selection of the reference ray

The isotropic ray theory is always the limiting case of the coupling ray theory for decreasing anisotropy at a fixed frequency. On the other hand, the high–frequency limit of the coupling ray theory at a fixed anisotropy depends on the choice of the reference ray, and even on the choice of the *system* of reference rays, because the amplitudes are determined by the paraxial reference rays.

From the point of view of the high–frequency asymptotic validity, the frequency–independent reference ray is best represented by the *anisotropic–ray–theory reference ray*, provided that we choose the initial condition for the polarization vector in the coupling equation given by the eigenvector of the Christoffel matrix corresponding to the reference ray. The anisotropic–ray–theory travel time corresponding to the selected polarization is then exact, and only the difference between the two anisotropic–ray–theory S–wave travel times is approximate. The coupling ray theory may then also be used at high frequencies because the approximate travel–time difference only influences the coupling due to low–frequency scattering. The coupling ray theory then correctly converges to the anisotropic ray theory for high frequencies. For other choices of reference rays, the high–frequency limit of the coupling ray theory at a fixed anisotropy is incorrect, although the differences may be small at the finite frequencies under consideration. Note that the anisotropic–ray–theory reference ray can be traced only if the eigenvectors of the Christoffel matrix vary continuously along the whole ray (Vavryčuk, 2001).

In the *common ray approximation*, only one reference ray is traced for both anisotropic–ray-theory S waves, and both S–wave anisotropic–ray-theory travel times are approximated by the perturbation expansion from the common reference ray. The common ray approximation thus considerably simplifies the coding of the coupling ray theory and numerical calculations, but may introduce errors in travel times due to the perturbation. These travel–time errors can deteriorate the coupling–ray–theory solution at high frequencies. The travel–time errors due to the common ray approximations can be calculated using the equations proposed by Klimeš & Bulant (2004).

In the *anisotropic common ray approximation*, the common reference ray is traced using the averaged Hamiltonian of both anisotropic–ray-theory S waves (Bakker, 2002; Klimeš, 2003b). This is probably the best common ray approximation (Klimeš & Bulant, 2004).

In the less accurate *isotropic common ray approximation*, the reference ray is traced in the reference isotropic model. The reference isotropic model may be selected in different ways, yielding quasi–isotropic approximations of differing accuracies.

In the common ray approximations, the S–wave travel times are usually approximated by the first–order perturbation expansion from the common reference ray. The errors of S–wave travel times may then be approximated by second–order terms in the perturbation expansion. A method of estimating the errors due to the isotropic common ray approximation and the anisotropic common ray approximation has been proposed and numerically demonstrated by Klimeš & Bulant (2004). The method is based on the equations for the second–order perturbations of travel time derived by Klimeš (2002a).

The accuracy of the anisotropic common ray approximation can be studied along isotropic common rays, without tracing the anisotropic common rays. If the error of the isotropic common ray approximation exceeds an acceptable limit, we can immediately decide whether the anisotropic common ray approximation (Bakker, 2002; Klimeš, 2003b) would be sufficiently accurate, or whether the anisotropic–ray–theory rays should be traced as reference rays for the coupling ray theory. The numerical results by Klimeš & Bulant (2004) suggest that the anisotropic common ray approximation by Bakker (2002) is worth coding and applying. Klimeš & Bulant (2004) studied the effects of the anisotropic common ray approximation and the less accurate isotropic common ray approximation, and compared them with the effects of the quasi–isotropic approximation of the Christoffel matrix, of the quasi–isotropic projection of the Green function, and of the quasi–isotropic perturbation of travel times.

10.3.2. Quasi–isotropic projection of the Green function

The coupling–ray–theory solution (76) may be approximated by its projection

$$\tilde{u}_i = h_{iM} h_{mM} u_m \tag{83}$$

onto the reference S–wave polarization plane, given by two orthonormal reference polarization vectors h_{k1}, h_{k2}. This approximation may simplify the modification of existing isotropic ray tracing codes for the coupling ray theory, and was used by Bulant & Klimeš (2002) in numerical calculation. The error of this approximation is obvious and simple to calculate.

10.3.3. Quasi–isotropic approximation of the Christoffel matrix

The Christoffel matrix may be approximated by its projections onto the reference S–wave polarization plane and onto the reference P–wave polarization line. Denote the polarization vectors of the isotropic ray theory, or the reference polarization vectors in general, by h_{k1}, h_{k2} and h_{k3}. If the Christoffel matrix is approximated by its projections onto plane h_{j1}, h_{k2} and onto vector h_{l3}, namely,

$$\begin{aligned}\tilde{\Gamma}_{jk} &= h_{jM} h_{mM} \Gamma_{mn} h_{nN} h_{kN} + h_{j3} h_{m3} \Gamma_{mn} h_{n3} h_{k3} \\ &= \Gamma_{jk} - (h_{jM} h_{k3} + h_{j3} h_{kM}) h_{mM} \Gamma_{mn} h_{n3} \quad ,\end{aligned} \tag{84}$$

then eigenvectors g_{k1} and g_{k2} become located in plane h_{j1}, h_{k2} as in the zero-order quasi–isotropic approximation of Pšenčík (1998). This approximation includes the quasi–isotropic projection of the Green function.

The quasi–isotropic approximation of the Christoffel matrix usually generates much greater error than the other quasi–isotropic approximations. The alteration of the coupling–ray–theory seismograms due to the quasi–isotropic approximation of the Christoffel matrix has been demonstrated by Bulant & Klimeš (2002).

10.3.4. Quasi–isotropic perturbation of travel times

The anisotropic–ray–theory travel times used in the coupling ray theory should be calculated by the numerical quadrature of equations (75) along the reference ray. In the quasi–isotropic perturbation of travel times, the anisotropic–ray–theory travel times are calculated from the reference travel time by the first–order perturbation with respect to the density–normalized elastic moduli,

$$\frac{\mathrm{d}\tau_1}{\mathrm{d}\tau} \approx (\Gamma^0_{jk} g_{j1} g_{k1})^{-\frac{1}{2}} - \frac{1}{2} (\Gamma_{jk} - \Gamma^0_{jk}) g_{j1} g_{k1} (\Gamma^0_{jk} g_{j1} g_{k1})^{-\frac{3}{2}} \quad . \tag{85}$$

If the isotropic common ray approximation and the quasi–isotropic approximation of the Christoffel matrix are also applied, then

$$\Gamma^0_{jk} g_{j1} g_{k1} = 1 \quad , \tag{86}$$

and equation (85) becomes

$$\frac{\mathrm{d}\tau_1}{\mathrm{d}\tau} \approx \frac{3}{2} - \frac{1}{2}\Gamma_{jk} g_{j1} g_{k1} \quad , \tag{87}$$

as in the quasi–isotropic approximation of Pšenčík (1998). Analogously, we obtain $\frac{\mathrm{d}\tau_2}{\mathrm{d}\tau}$. The quasi–isotropic perturbation of travel times leads to an erroneous time shift in coupling–ray–theory seismograms, but has a negligible impact on the amplitudes.

10.4. Numerical examples

The vertically heterogeneous 1–D anisotropic model QI was used by Pšenčík & Dellinger (2001, model WA rotated by 45°) for comparison of the coupling–ray–theory synthetic seismograms with the reflectivity method. The density–normalized elastic moduli a_{ijkl} in both model QI and in the reference isotropic model QI0 are interpolated linearly with depth. For the discussion and description of this model refer to Pšenčík & Dellinger (2001).

The differences of the elastic moduli of model QIH from the elastic moduli of the reference isotropic model QI0 are exactly twice smaller than the differences of model QI. The differences of the elastic moduli of model QI2 from the elastic moduli of the reference isotropic model QI0 are exactly twice larger than the differences of model QI. The differences of the elastic moduli of model QI4 from the elastic moduli of the reference isotropic model QI0 are exactly 4 times larger than the differences of model QI. For a description of models QI0, QI, QI2 and QI4 refer to Klimeš & Bulant (2004).

The synthetic seismograms were calculated at 29 receivers located in a vertical well drilled 1 km from a vertical force. The source time function is the Gabor signal with reference frequency 50 Hz, band–pass filtered by a cosine filter.

Isotropic–ray–theory, coupling–ray–theory and anisotropic–ray–theory seismograms in model QIH are compared in Figures 18, 19 and 20. Coupling–ray–theory seismograms in all 4 anisotropic models QIH, QI, QI2 and QI4 are compared in Figure 21.

Figure 22 shows the comparison of the synthetic seismograms computed by the quasi–isotropic approximation of the coupling ray theory and by the Chebyshev spectral method (Tessmer, 1995) in model QI. Slight discrepancies (quasi–isotropic synthetics are "faster") appearing at shallower receivers have not been explained yet. According to Klimes & Bulant (2004), the quasi–isotropic approximation should yield sufficiently accurate results for model QI.

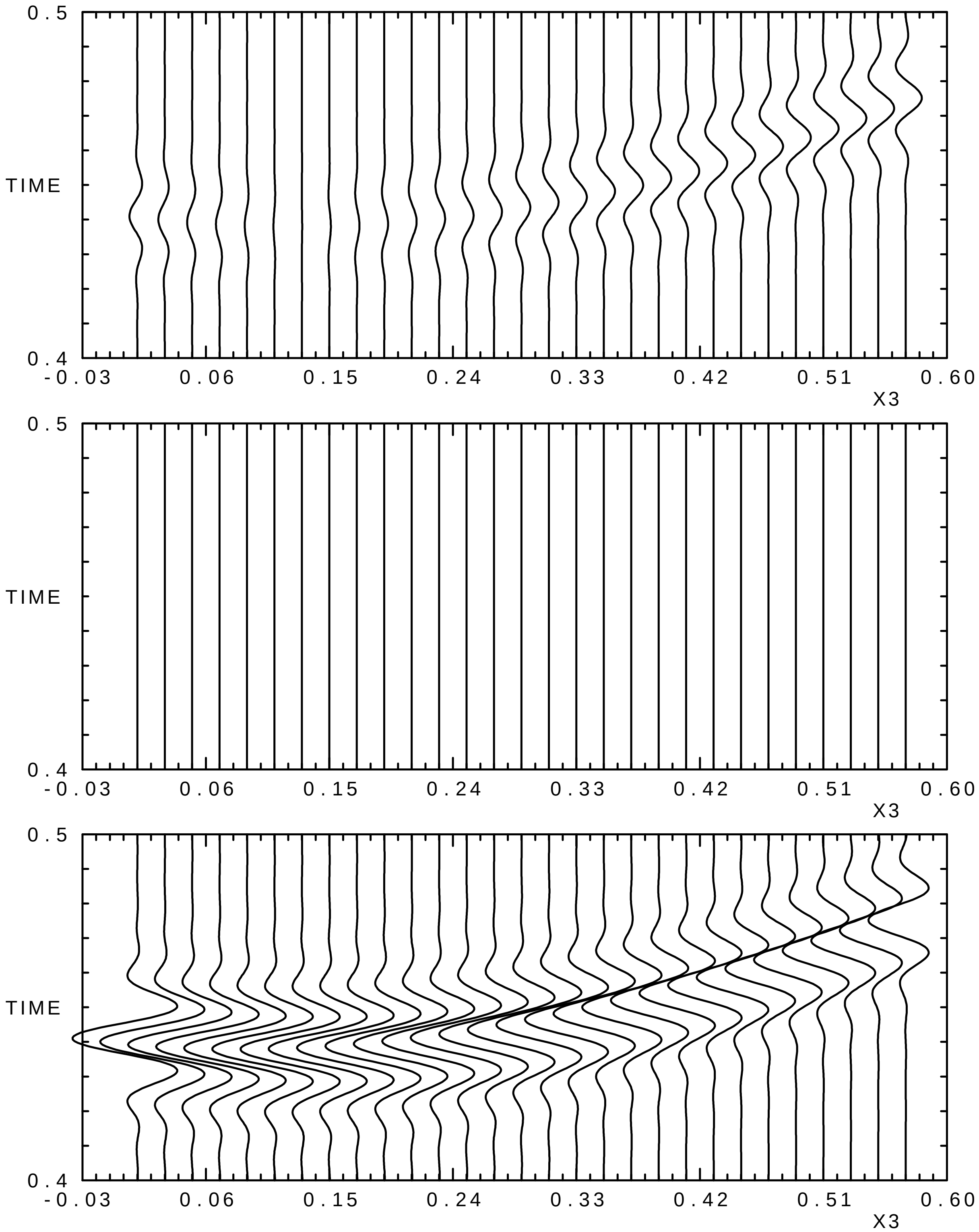

Figure 18. Isotropic–ray–theory seismograms in model QIH (actually calculated in isotropic model QI0). From top to bottom: the first (radial) component, the second (transverse) component, the third (vertical) component.

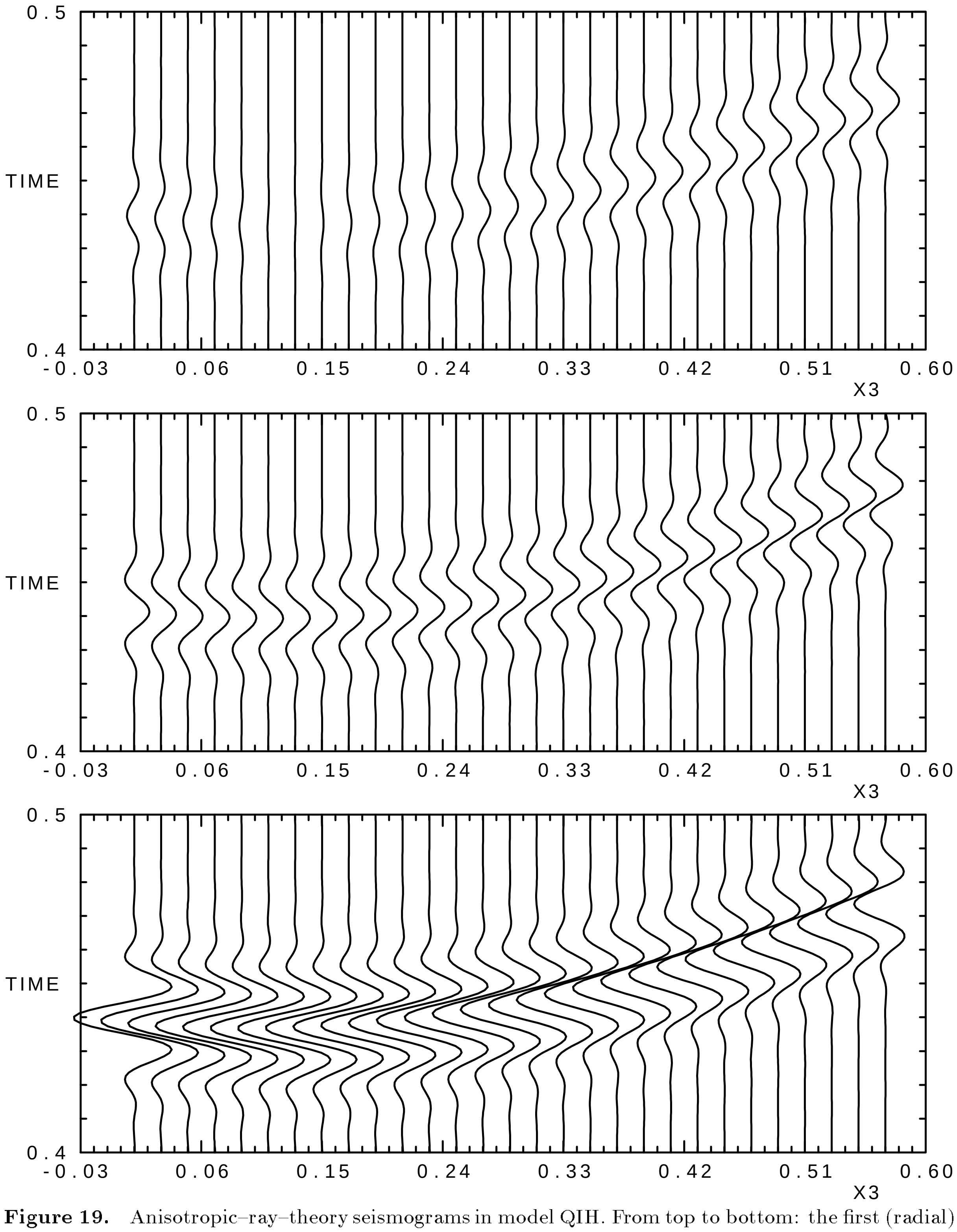

Figure 19. Anisotropic–ray–theory seismograms in model QIH. From top to bottom: the first (radial) component, the second (transverse) component, the third (vertical) component.

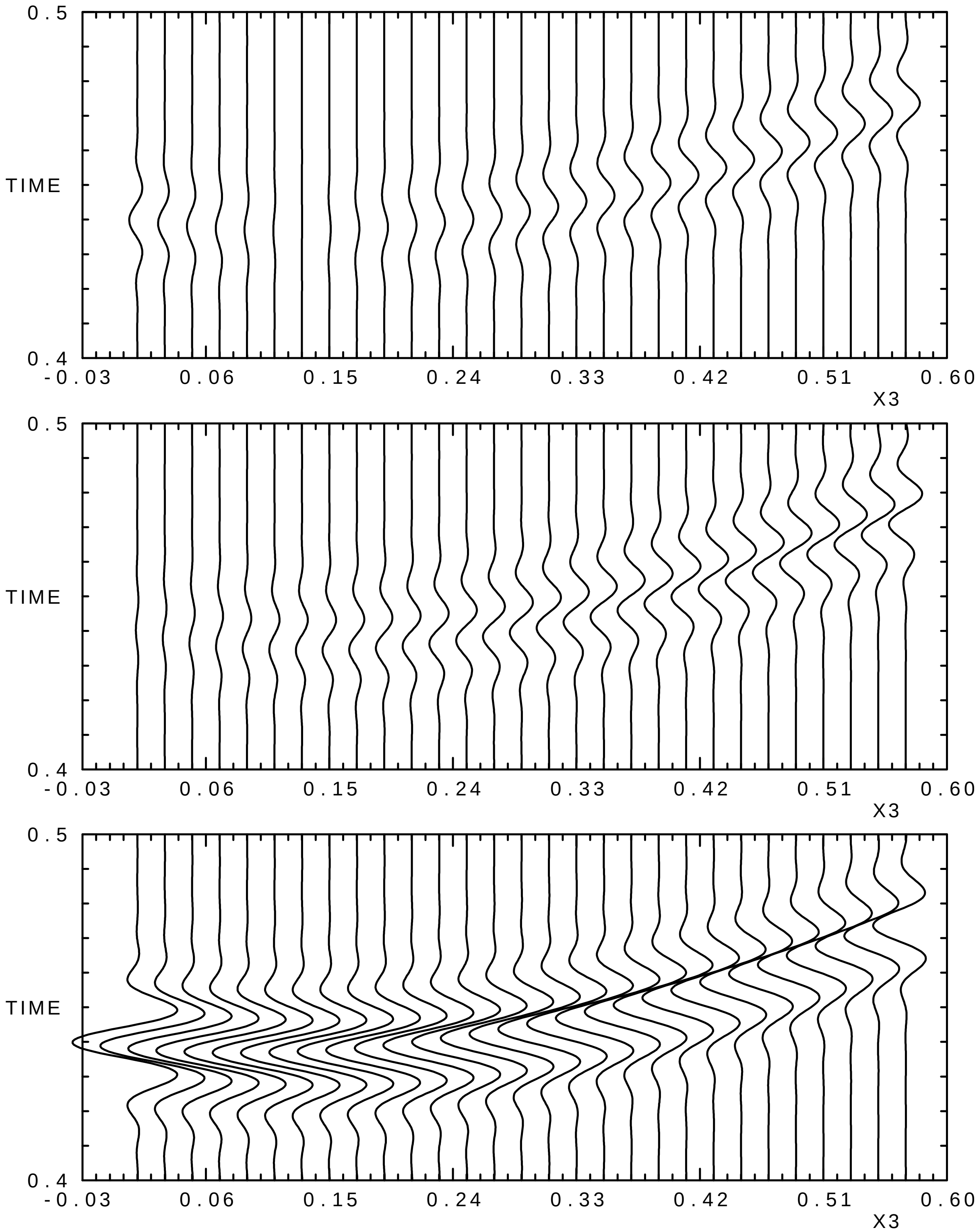

Figure 20. Coupling–ray–theory seismograms in model QIH. From top to bottom: the first (radial) component, the second (transverse) component, the third (vertical) component.

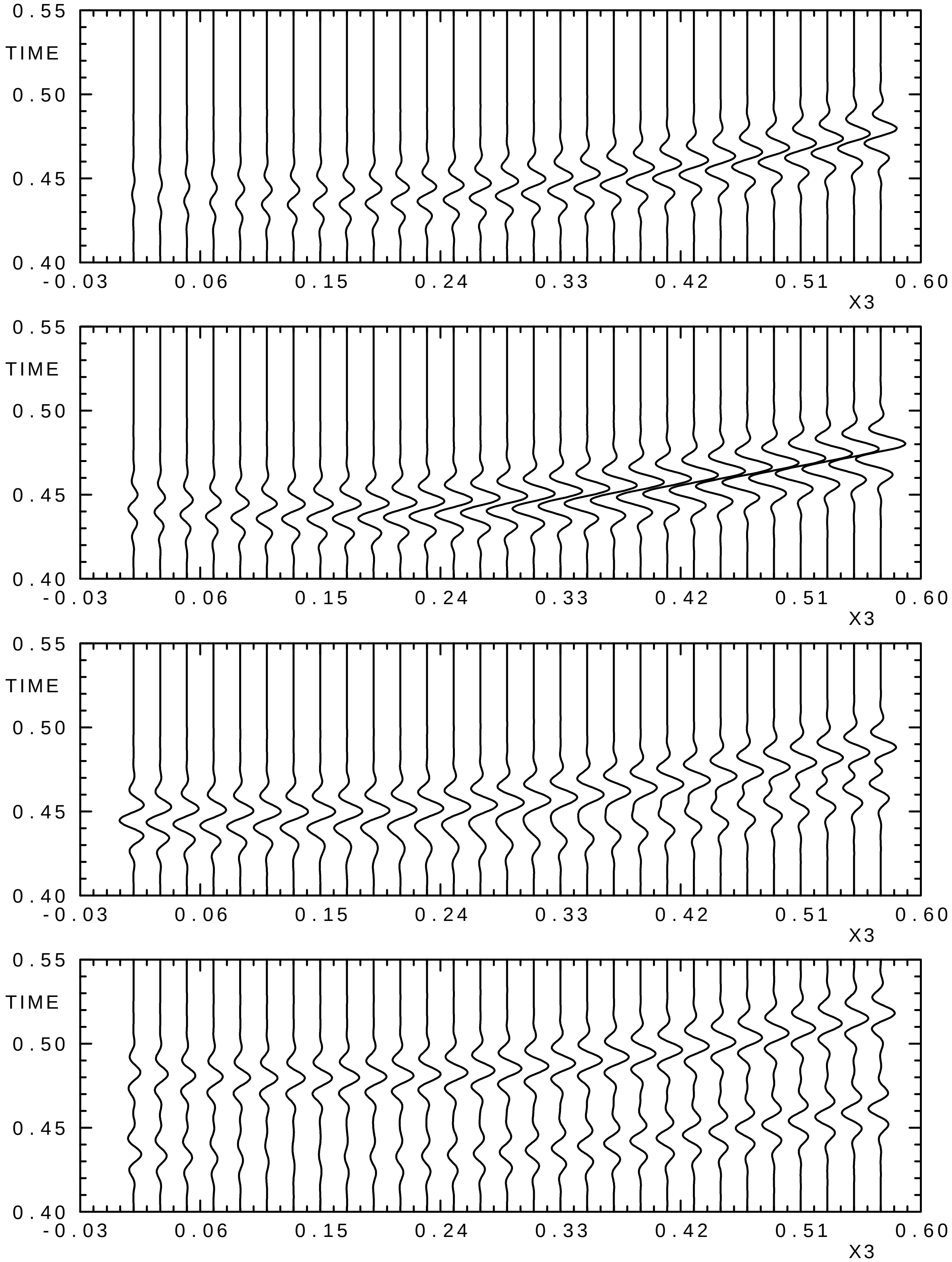

Figure 21. Coupling–ray–theory seismograms in models QIH, QI, QI2 and QI4 (from top to bottom). Only the second (transverse) component is shown. This component vanishes in the reference isotropic model QI0. See the clear increment of amplitudes in the two upper models, and the clear development of S–wave splitting in the two bottom models.

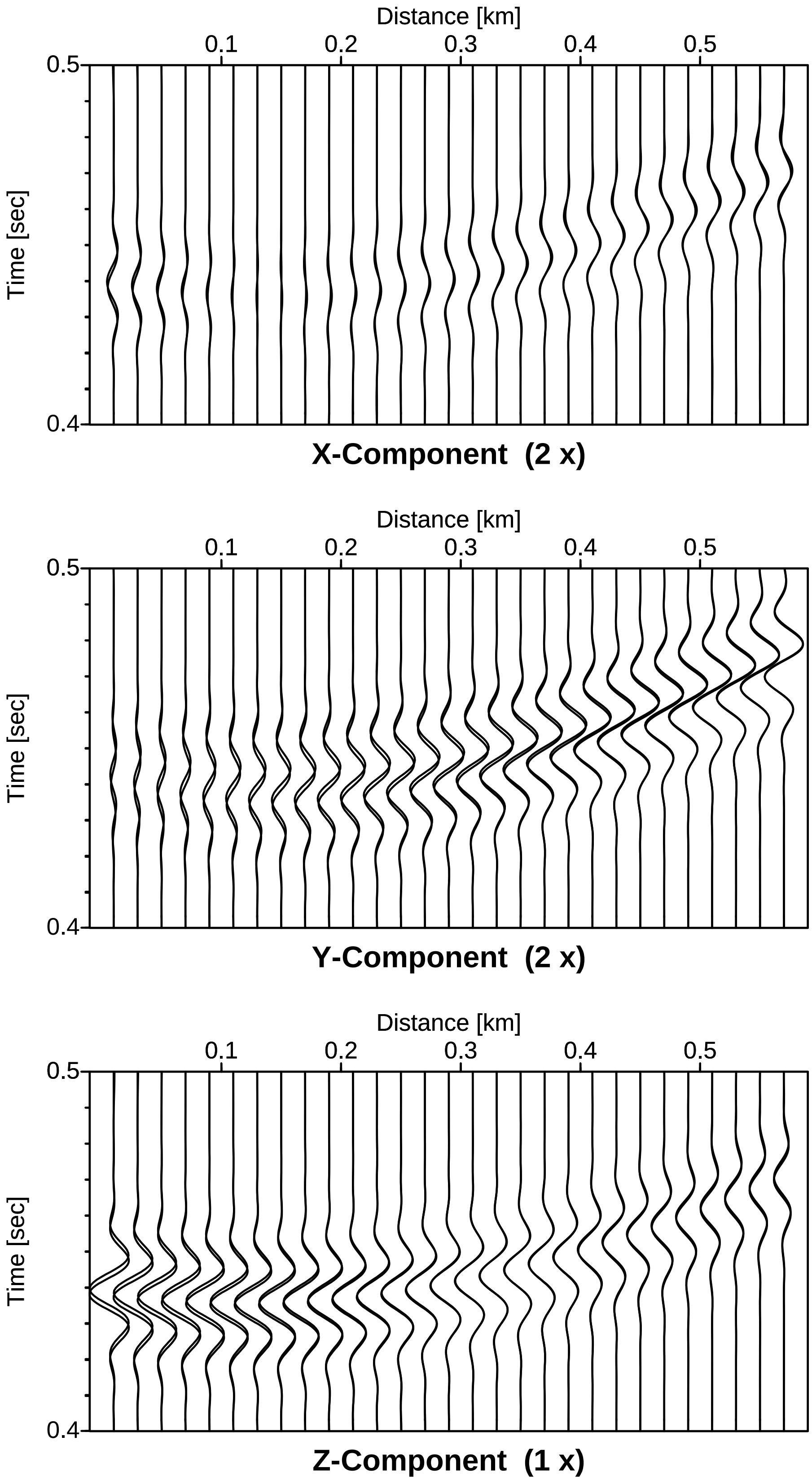

Figure 22. Comparison of the coupling–ray–theory seismograms with the Chebyshev spectral method in model QI. From top to bottom: the first (radial) component (twice enlarged), the second (transverse) component (twice enlarged), the third (vertical) component. For better viewing, the first two components are twice enlarged with respect to the third one.

11. Summation of Gaussian beams and packets

Gaussian beams and packets may serve as building blocks of a wavefield. The summation of Gaussian beams (Červený, Popov & Pšenčík, 1982; Popov, 1982; Červený, 1985) and packets overcomes the problems of the standard ray theory with caustics. The summation of Gaussian beams and packets is considerably comprehensive and flexible. It may be formulated in many ways, including the Maslov method and its various generalizations as special cases. The form of the summation depends primarily on the specification of the wavefield, also comprising the system of Gaussian packets scattered from Gabor functions forming medium perturbations.

We shall now concentrate on the representation of non–directional wavefields without pronounced diffractions rather than on the decomposition of directional beams or of diffracted waves. We thus omit the higher–order Gaussian beams (Bessel–Gaussian beams, Hermite–Gaussian beams, Laguerre–Gaussian beams) and diffracted Gaussian beams.

A Gaussian beam or packet is expressed in terms of a complex–valued vectorial amplitude and complex–valued travel time (phase function, eikonal). In this Section 11, we prefer the term "phase function" specifically for Gaussian packets, otherwise we shall use the term "travel time". The travel time depends on the wave equation through the eikonal equation, whereas the amplitude depends on the wave equation through the transport equation.

In Section 11, we consider only isotropic media. Asymptotic summations of Gaussian beams or packets are performed in the frequency domain, because the frequency domain is the natural representation in time–independent media.

11.1. Gaussian beams

A Gaussian beam is a high–frequency asymptotic time–harmonic solution of the elastodynamic equation, with an approximately Gaussian profile perpendicularly to the central ray. Refer to Figure 24 for an illustration. A Gaussian beam represents a bundle of complex–valued rays concentrated in the vicinity of a real–valued central ray. A paraxial Gaussian beam is a Gaussian beam approximated by the paraxial ray approximation in the vicinity of the central ray, or in the vicinity of a selected reference point. The paraxial ray approximation consists in the second–order Taylor expansion of travel time at a point on the central ray and in a constant amplitude.

The evolution of the shape of a Gaussian beam along the central ray is determined by the respective ordinary differential equations. The travel time and its first derivatives along the real–valued central ray are determined, together with the central ray, by ray tracing and are real–valued. The complex–valued second derivatives of travel time, which control the shape of the Gaussian beam, are determined by dynamic ray tracing, see Section 5. Dynamic ray tracing is also required for determining the complex–valued amplitude of the Gaussian beam.

Infinitely broad paraxial Gaussian beams (with real–valued second derivatives of travel time) correspond to the paraxial approximation of time–harmonic ray–theory wavefields. If used in a superposition, they must differ from the paraxial approximation of the decomposed ray–theory wavefield. Ribbon Gaussian beams are paraxial Gaussian beams infinitely broad only in a single *singular direction* perpendicular to the central ray. If used in a superposition, they correspond to the paraxial approximation of the decomposed ray–theory wavefield in the singular direction. If the ribbon Gaussian beams used in a superposition are infinitely broad in all directions, they coincide with the paraxial approximation of the decomposed ray–theory wavefield only in the singular direction, but must differ from it in other directions perpendicular to the central ray. Note that infinitely broad Gaussian beams correspond to the standard Maslov method. However, these infinitely broad Gaussian beams are not recommended for the frequency–domain superposition integrals.

11.1.1. Gaussian beam

A Gaussian beam is an approximate time–harmonic high–frequency asymptotic solution of the elastodynamic equation (or other hyperbolic wave equation), and may be expressed in the form of

$$\mathbf{u}(x_i,\omega) = \mathbf{U}\exp(\mathrm{i}\omega\tau) \quad , \tag{88}$$

where both the complex–valued vectorial amplitude $\mathbf{U} = \mathbf{U}(x_i)$ and the complex–valued travel time $\tau = \tau(x_i)$ are smooth functions of spatial coordinates x_i.

The complex–valued travel time satisfies the eikonal equation

$$V^2\tau_{,i}\tau_{,i} = 1 \quad . \tag{89}$$

Introducing slowness vector

$$p_i = \tau_{,i} \tag{90}$$

and Hamiltonian (3), the eikonal equation may be solved by the method of characteristics. The corresponding Hamilton ray tracing equations for frequency independent propagation velocity $V = V(x_k)$ are

$$\frac{\mathrm{d}y_i}{\mathrm{d}\gamma_3} = [V(y_k)]^2 p_i \quad , \qquad \frac{\mathrm{d}p_i}{\mathrm{d}\gamma_3} = -[V(y_n)]^{-1}V_{,i}(y_k) \quad , \tag{91}$$

see (10). Here we have denoted the coordinates by y_i instead of x_i in order to be able to distinguish the points of the ray from general coordinates x_i. We have also denoted the travel time along the ray γ_3 in order to distinguish it from the complex–valued travel time $\tau = \tau(x_i)$ in equation (88).

Since the Hamiltonian is real–valued along the real–valued rays, the imaginary part of the travel time is constant. We shall refer to the ray, along which the imaginary part of the travel time is minimum, as the central ray. We shall assume that the value of the imaginary part of the travel time along the central ray is zero,

$$\mathrm{Im}\{\tau[y_i(\gamma_3)]\} = 0 \quad . \tag{92}$$

Since (92) corresponds to the minimum of the imaginary part of τ, inequality

$$\mathrm{Im}[\tau(x_i)] \geq 0 \tag{93}$$

is the consequence of option (92). The Gaussian beam is thus concentrated close to its central ray.

11.1.2. Paraxial Gaussian beam

We shall refer to the approximation of a Gaussian beam, obtained by substituting the exact travel time τ by its second–order Taylor expansion, as the paraxial approximation of the Gaussian beam or, briefly, as the *paraxial Gaussian beam.* We shall refer to Gaussian beam (88) with the exact travel time τ, satisfying eikonal equation (89) as the *strict Gaussian beam.*

The paraxial approximation of a strict Gaussian beam centred at point $y_i(\gamma_3)$, is

$$\mathbf{u}(x_k,\omega) = \mathbf{U}\ \exp\{\mathrm{i}\omega[p_i^{\mathrm{R}} r_i + \tfrac{1}{2} N_{ij} r_i r_j]\} \quad , \tag{94}$$

where

$$r_i = x_i - y_i(\gamma_3) \tag{95}$$

is the difference of the spatial coordinates from reference point y_i. p_i^{R} is the ray–theory slowness vector at the reference point, and

$$N_{ij} = \tau_{,ij} \tag{96}$$

are the second derivatives of the travel time at the reference point with respect to Cartesian coordinates. Quantities $\mathbf{U}$, p_i and N_{ij} on the right–hand side of equation (94) are taken at the reference point $y_i(\gamma_3)$ of the Gaussian beam.

The complex–valued second partial derivatives N_{ij} of the travel time with respect to Cartesian coordinates x_i at reference point y_i determine the shape of the Gaussian beam. The imaginary part of the 3×3 matrix N_{kl} has rank 2 and should be chosen positive–semidefinite. The evolution of the complex–valued second derivatives N_{ij} of the travel time along the spatial central ray is determined by the quantities calculated by dynamic ray tracing. Dynamic ray tracing is also required for determining the complex–valued amplitude of the Gaussian beam.

The difference between the strict Gaussian beam and its paraxial approximation is of the asymptotic order $\omega^{-\frac{1}{2}}$. A single paraxial Gaussian beam is then a worse solution to the equations of motion than the strict Gaussian beam by the asymptotic order $\omega^{\frac{1}{2}}$. However, the differences of order $\omega^{-\frac{1}{2}}$ between the paraxial and strict Gaussian beams are odd functions of the coordinates and so vanish in the integration over the space in the asymptotic summation of Gaussian beams. The error of the 2–parametric (1–parametric in 2–D models) time–harmonic superposition of paraxial Gaussian beams is thus of the asymptotic order ω^{-1}, i.e. of the same order as the error of the superposition of the strict Gaussian beams.

As the travel time, the amplitude can also be replaced by its Taylor expansion. The restriction of the amplitude up to the n–th order introduces an error of the asymptotic order $\omega^{-\frac{n+1}{2}}$. Since the error of the second–order expansion of the travel time is of order $\omega^{-\frac{1}{2}}$, the amplitude can be replaced by the constant in the paraxial approximation.

11.1.3. Equations for the second derivatives of travel time in Cartesian coordinates

Differentiating eikonal equation (89) with respect to spatial coordinates and considering (96), we obtain relation

$$V^2 p_k N_{ki} = -V^{-1} V_{,i} \quad . \tag{97}$$

Differentiating eikonal equation (89) twice with respect to spatial coordinates, and inserting it together with ray tracing equations (91) into equation $\mathrm{d}\tau_{,jk}/\mathrm{d}\gamma_3 = \tau_{,ijk}\,\mathrm{d}y_k/\mathrm{d}\gamma_3$, we arrive at the Riccati equation

$$\frac{\mathrm{d}N_{jk}}{\mathrm{d}\gamma_3} = -V^2 N_{ij} N_{ik} - V^{-1} V_{,jk} - V^{-2} V_{,j} V_{,k} - 2V V_{,j} N_{ik} p_i - 2V V_{,k} N_{ij} p_i \quad . \tag{98}$$

Inserting (97) into (98), we obtain the Riccati equation for the 3×3 matrix N_{ij},

$$\frac{\mathrm{d}N_{jk}}{\mathrm{d}\gamma_3} = -V^2 N_{ij} N_{ik} - V^{-1} V_{,jk} + 3V^{-2} V_{,j} V_{,k} \quad . \tag{99}$$

Equation (99) describes the evolution of the shape of a Gaussian beam in general Cartesian coordinates. Note that equation (99) can also be transformed into the system of two linear ordinary differential equations for two 3×3 matrices.

11.1.4. Equations for the second derivatives of travel time in ray–centred coordinates

The second–order covariant derivatives M_{mn} of the travel time in ray–centred coordinates are related to the partial derivatives N_{mn} in Cartesian coordinates through equation

$$N_{jk} = H_{jm} H_{kn} M_{mn} \quad . \tag{100}$$

Riccati equation (99) for matrix N_{ij} is equivalent to the Riccati equation for 3×3 matrix M_{mn}. Here H_{im} is the i–th component of m–th basis vector of the ray–centred coordinate system, see Section 5.1.

The derivatives of the basis vectors of the ray–centred coordinate system along the ray satisfy relations, see (30),

$$\frac{\mathrm{d}H_{jM}}{\mathrm{d}\gamma_3} H_{jN} = 0 \ , \quad \frac{\mathrm{d}H_{j3}}{\mathrm{d}\gamma_3} H_{jN} = -V^{(q)}_N \ , \quad \frac{\mathrm{d}H_{jM}}{\mathrm{d}\gamma_3} H_{j3} = V^{(q)}_M \ , \quad \frac{\mathrm{d}H_{j3}}{\mathrm{d}\gamma_3} H_{j3} = 0 \ , \tag{101}$$

where

$$V^{(q)}_m = V_{,i} H_{im} \tag{102}$$

is the velocity gradient in ray–centred coordinates.

Equation (97), transformed using (100) into ray–centred coordinates, reads

$$M_{m3} = M_{3M} = -V^{-2} V^{(q)}_m \quad . \tag{103}$$

Equation (99), transformed using (100) into ray–centred coordinates, reads

$$\frac{\mathrm{d}M_{mn}}{\mathrm{d}\gamma_3} = -M_{jn} \frac{\mathrm{d}H_{ij}}{\mathrm{d}\gamma_3} H_{im} - M_{mk} \frac{\mathrm{d}H_{ik}}{\mathrm{d}\gamma_3} H_{in} - V^2 M_{km} M_{kn} - V^{-1} V^{(q)}_{mn} + 3V^{-2} V^{(q)}_m V^{(q)}_n \ , \tag{104}$$

where

$$V^{(q)}_{mn} = V_{,ij} H_{im} H_{jn} \tag{105}$$

are the covariant velocity derivatives in ray–centred coordinates. Inserting (101) for the derivatives of the basis vectors, equation (104) may be expanded into the equation for the 2×2 submatrix M_{MN} and equations for matrix elements M_{M3} and M_{33}. The equation for the 2×2 submatrix M_{MN}, with (103) inserted for M_{M3}, reads

$$\frac{\mathrm{d}M_{MN}}{\mathrm{d}\gamma_3} = -V^2 M_{MK} M_{KN} - V^{-1} V^{(q)}_{MN} \quad . \tag{106}$$

The equations for matrix elements M_{M3} and M_{33} need not be used because their solution is given by equation (103).

The covariant velocity derivatives $V^{(q)}_i$ and $V^{(q)}_{ij}$ in ray–centred coordinates are defined by equations (102) and (105).

11.1.5. Solving the equation for a Gaussian beam

To express equation (106) in matrix notation, we introduce matrices

$$\mathbf{M} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \quad , \qquad \mathbf{V}^{(q)} = \begin{pmatrix} V^{(q)}_{11} & V^{(q)}_{12} \\ V^{(q)}_{21} & V^{(q)}_{22} \end{pmatrix} \quad . \tag{107}$$

Equation (106) in matrix notation has form (51), and its solution may be expressed as

$$\mathbf{M} = \mathbf{M}^{\mathrm{T}} = \mathbf{P}^{(q)} (\mathbf{Q}^{(q)})^{-1} \quad , \tag{108}$$

see (50). Complex–valued matrices $\mathbf{Q}^{(q)}$ and $\mathbf{P}^{(q)}$ satisfy dynamic ray tracing equations (41). Introducing the DRT propagator matrix (47), matrices $\mathbf{Q}^{(q)}$ and $\mathbf{P}^{(q)}$ may be expressed in terms of their initial values $\mathbf{Q}^0$ and $\mathbf{P}^0$ as

$$\mathbf{Q}^{(q)} = \mathbf{Q}_1\, \mathbf{Q}^{(q)0} + \mathbf{Q}_2\, \mathbf{P}^{(q)0} \quad , \qquad \mathbf{P}^{(q)} = \mathbf{P}_1\, \mathbf{Q}^{(q)0} + \mathbf{P}_2\, \mathbf{P}^{(q)0} \quad , \tag{109}$$

with

$$\mathbf{P}^{(q)0} = \mathbf{M}^0\, \mathbf{Q}^{(q)0} \quad . \tag{110}$$

Matrix (108) may then be expressed as

$$\mathbf{M} = (\mathbf{P}_1 + \mathbf{P}_2\, \mathbf{M}^0)\, (\mathbf{Q}_1 + \mathbf{Q}_2\, \mathbf{M}^0)^{-1} \quad . \tag{111}$$

11.1.6. Amplitude of a Gaussian beam

The vectorial amplitude $\mathbf{U}$ of the Gaussian beam may be expressed in terms of the complex–valued scalar amplitude A and unit polarization vector $\mathbf{e}$,

$$\mathbf{U} = A\, \mathbf{e} \quad . \tag{112}$$

The unit polarization vector $\mathbf{e}$ is identical to the analogous ray–theory polarization vector.

The complex–valued scalar amplitude evolves along the central ray according to equation

$$A = A^0 \sqrt{\frac{V^0 \varrho^0}{V \varrho \det(\mathbf{Q}_1 + \mathbf{Q}_2 \mathbf{M}^0)}} \quad . \tag{113}$$

Here velocity V and density ϱ correspond to the elastodynamic equation, and should be replaced by the respective material parameters for other wave equations.

11.2. Gaussian packets

Gaussian packets, also called (space–time) Gaussian beams (Ralston, 1983), quasi-photons (Babich & Ulin, 1981) or coherent states (Combescure, Ralston & Robert, 1999), are high–frequency asymptotic space–time solutions of the elastodynamic equation. A Gaussian packet has an approximately Gaussian profile in all spatial directions and in time. A Gaussian packet is concentrated to a real–valued space–time ray, as a Gaussian beam to a spatial ray. In a stationary medium, a Gaussian packet propagates along its real–valued spatial central ray. A paraxial Gaussian packet is a Gaussian packet approximated by the paraxial ray approximation in the vicinity of its central point. The time–dependent central point is the spatial position of the maximum of the envelope of the Gaussian packet. The central point of a Gaussian packet moves along the spatial central ray according to the ray tracing equations. The shape of a Gaussian packet is determined by the second derivatives of the phase function.

The equations for the evolution of a Gaussian packet were derived by Babich & Ulin (1981) and Ralston (1983). The phase function and its first derivatives along the real–valued central ray are determined by ray tracing and are real–valued. The evolution of the complex–valued second derivatives of the phase function, which control the shape of the Gaussian packet, along the central ray is determined by the quantities calculated by dynamic ray tracing with additional quadratures along the central ray. Dynamic ray tracing is also required for determining the complex–valued amplitude of the Gaussian packet.

The evolution equations are required to propagate Gaussian packets constituting a general time–domain wavefield (Žáček, 2003a, 2003b), and to optimize the shape of Gaussian packets (Žáček, 2001a, 2001b). The evolution equations for the shape of Gaussian packets are unnecessary for the asymptotic decomposition of the time–harmonic wavefield specified in terms of amplitude and phase (Klimeš, 1984b, 1989b), but may be applied to the optimization of the shape of Gaussian packets used in this asymptotic decomposition.

Infinitely long paraxial Gaussian packets correspond to the paraxial approximations of Gaussian beams. Infinitely broad paraxial Gaussian packets correspond to the paraxial approximations of space–time ray–theory wavefields.

11.2.1. Gaussian packet and the space-time eikonal equation

A Gaussian packet is an approximate space–time high–frequency asymptotic solution of the corresponding hyperbolic wave equation, and may be expressed in the form of

$$\mathbf{u}(x_\kappa) = \mathbf{U}\exp(\mathrm{i}\omega\tau) \quad , \tag{114}$$

where both the complex–valued vectorial amplitude $\mathbf{U} = \mathbf{U}(x_\alpha)$ and the complex–valued phase function $\tau = \tau(x_\alpha)$ are smooth functions of the four space–time coordinates x_α. Hereinafter, the lower–case Greek subscripts ${}_{\alpha,\beta,\ldots} = {}_{1,2,3,4}$ will correspond to four space–time coordinates: spatial coordinates x_i with ${}_i = {}_{1,2,3}$, and time x_4.

The complex–valued phase function satisfies the space–time eikonal equation

$$V^2\tau_{,i}\tau_{,i} - \tau_{,4}\tau_{,4} = 0 \quad . \tag{115}$$

Introducing space–time slowness vector

$$N_\alpha = \tau_{,\alpha} \tag{116}$$

and Hamiltonian $\mathcal{H} = \mathcal{H}(x_\alpha, N_\alpha) = V^2(x_k)N_iN_i - N_4N_4$, the eikonal equation may be solved by the method of characteristics. The corresponding Hamilton space–time ray tracing equations for temporally independent propagation velocity $V = V(x_k)$ are

$$\frac{\mathrm{d}y_i}{\mathrm{d}\gamma_3} = [V(y_k)]^2 N_i \ , \quad \frac{\mathrm{d}y_4}{\mathrm{d}\gamma_3} = -N_4 \ , \quad \frac{\mathrm{d}N_i}{\mathrm{d}\gamma_3} = -[V(y_n)]^{-1} V_{,i}(y_k) \ , \quad \frac{\mathrm{d}N_4}{\mathrm{d}\gamma_3} = 0 \ . \quad (117)$$

Here again we have denoted the space–time coordinates by y_α instead of x_α in order to be able to distinguish the points of the space–time ray from general space–time coordinates x_α.

The imaginary part of phase function τ must have, at a fixed time $x_4 = y_4(\gamma_3)$, a strict minimum over space at some point $x_i = y_i(\gamma_3)$. The imaginary part of the phase function can be approximated in the vicinity of its minimum y_i by the quadratic form of spatial distances $[x_i - y_i(\gamma_3)]$. For sufficiently high frequencies, the amplitude profile of the Gaussian packet is thus *nearly* Gaussian in space and concentrated close to point $y_i(\gamma_3)$. We shall refer to point $y_i(\gamma_3)$ as the central point of the Gaussian packet.

Since the Hamiltonian is zero along the space–time rays, the complex–valued phase function is constant. In this way, the minimum of the imaginary part of τ, which yields the maximum of the high–frequency Gaussian packet, is identical with the space–time ray passing through point $y_i(\gamma_3^0)$ at time $y_4(\gamma_3^0)$ (Norris, White & Schreiffer, 1987). We shall assume that the value of the phase function at this space–time ray is zero,

$$\tau[y_\alpha(\gamma_3)] = 0 \quad . \tag{118}$$

Since (118) corresponds to the minimum of the imaginary part of τ, inequality

$$\mathrm{Im}[\tau(x_\alpha)] \geq 0 \tag{119}$$

is the consequence of option (118).

The Gaussian packet is thus concentrated close to the space–time ray as the Gaussian beam is concentrated close to the spatial ray, and is also called a (space–time) Gaussian beam (Ralston 1983). Thus, in a time–independent perfectly elastic and isotropic model, the Gaussian packet moves along the spatial ray with the ray velocity. The Gaussian packet remains approximately Gaussian in space at all times $x_4 = y_4(\gamma_3)$. At any time, the principal part of the energy of the Gaussian packet is effectively concentrated in some vicinity of the central point $y_i(\gamma_3)$ of the Gaussian packet. This vicinity is called the *effective region* of the Gaussian packet. At a specified time, the effective region of a Gaussian packet is limited in space and, for a sufficiently high frequency, it is concentrated close to its central point. In space–time, however, the effective region of a Gaussian packet is not limited.

We select

$$N_4 = -1 \quad . \tag{120}$$

Derivative $\mathrm{d}/\mathrm{d}\gamma_3$ along the space–time ray then coincides with derivative $\mathrm{d}/\mathrm{d}y_4$ with respect to time y_4, and space–time ray tracing equations (117) reduce to the spatial ray tracing equations (10).

11.2.2. Paraxial Gaussian packet

We shall refer to the approximation of a Gaussian packet, obtained by substituting the exact phase function τ by its second–order Taylor expansion, as the paraxial approximation of the Gaussian packet or, briefly, as the *paraxial Gaussian packet.* The paraxial approximation of the Gaussian packet in space is an analogy of the paraxial approximation of the Gaussian beam in the plane perpendicular to the ray. We shall refer to the Gaussian packet (114) with the exact phase function τ satisfying eikonal equation (115) as the *strict Gaussian packet.*

The space–time paraxial approximation of a Gaussian packet centred at point $y_\alpha(\gamma_3)$, is

$$\mathbf{u}(x_\kappa) = \mathbf{U}\ \exp\{\mathrm{i}\omega[N_\alpha r_\alpha + \tfrac{1}{2}N_{\alpha\beta}r_\alpha r_\beta]\} \quad , \tag{121}$$

where

$$r_i = x_i - y_i(\gamma_3) \quad , \tag{122}$$

$$r_4 = x_4 - y_4(\gamma_3) = x_4 - \gamma_3 \tag{123}$$

is the difference of the space–time coordinates from the central point y_α. N_α defined by (116) is the space–time slowness vector of the Gaussian packet, and

$$N_{\alpha\beta} = \tau_{,\alpha\beta} \tag{124}$$

are the second derivatives of the phase function at the central point. Quantities $\mathbf{U}$, N_α and $N_{\alpha\beta}$ on the right–hand side of equation (121) are taken at the central point $y_\alpha(\gamma_3)$ of the Gaussian packet.

Since we have selected $N_4 = -1$, see (120),

$$N_i = p_i^{\mathrm{R}} \tag{125}$$

is the ray–theory slowness vector at the central point. The first derivatives N_α of the phase function at the central point of the Gaussian packet are thus determined by standard ray tracing and are real–valued.

The complex–valued second partial derivatives $N_{\alpha\beta}$ of the phase function with respect to Cartesian coordinates x_i at the central point y_α determine the shape of the Gaussian packet. The imaginary part of the 3×3 spatial submatrix N_{kl} of the 4×4 matrix $N_{\alpha\beta}$ should be chosen so that it is positive–definite. The imaginary part of the whole 4×4 matrix $N_{\alpha\beta}$ is singular. The evolution of the complex–valued second derivatives $N_{\alpha\beta}$ of the phase function along the spatial central ray is determined by quantities calculated by dynamic ray tracing. Dynamic ray tracing is also required for determining the complex–valued amplitude of the Gaussian packet.

The spatial effective region of the Gaussian packet may be sufficiently small for sufficiently high frequencies. Then, at any time $x_4 = y_4(\gamma_3)$, the phase function τ of the strict Gaussian packet can be replaced, in the spatial effective region, by its Taylor expansion up to the second order with respect to the spatial coordinates,

$$\mathbf{u}(x_\kappa) = \mathbf{U}\ \exp\{\mathrm{i}\omega[N_i r_i + \tfrac{1}{2}N_{ij}r_i r_j]\} \quad , \tag{126}$$

where N_{ij} is the spatial part of the 4×4 matrix (124) of the second space–time derivatives of the phase function.

The difference between the strict Gaussian packet and its spatial paraxial approximation is of the asymptotic order $\omega^{-\frac{1}{2}}$. A single paraxial Gaussian packet is then a worse solution to the equations of motion than the strict Gaussian packet by the asymptotic order $\omega^{\frac{1}{2}}$. However, the differences of order $\omega^{-\frac{1}{2}}$ between the paraxial and strict Gaussian packets are odd functions of the coordinates, which therefore vanish in the integration over the space in the asymptotic summation of Gaussian packets. The error of the 3–parametric (2–parametric in 2–D models) time–harmonic superposition of paraxial Gaussian packets is thus of the asymptotic order ω^{-1}, i.e. of the same order as the error of the superposition of the strict Gaussian packets.

As the phase function, the amplitude can also be replaced by its Taylor expansion. The restriction of the amplitude up to the n–th order introduces an error of the asymptotic order $\omega^{-\frac{n+1}{2}}$. Since the error of the second–order expansion of the phase function is of the order $\omega^{-\frac{1}{2}}$, the amplitude can be replaced by the constant in the paraxial approximation.

11.2.3. Equations for the second derivatives of phase function in Cartesian coordinates

The Riccati equation for the 4×4 space–time matrix (124) may be obtained by differentiating eikonal equation (115) twice with respect to the space–time coordinates x_α, see Ralston (1983, eq. 2.4). The corresponding linear dynamic ray tracing system may be obtained by differentiating the ray tracing system with respect to the space–time ray parameters, see Ralston (1983, eq. 2.6). The Riccati equation for the 4×4 space–time matrix (124) may be reduced to the Riccati equation for its 3×3 spatial submatrix N_{ij}, see Kachalov (1984, eq. 9), Babich, Buldyrev & Molotkov (1985, eq. 2.15), Norris, White & Schreiffer (1987, eq. 2.27a). For the corresponding linear dynamic ray tracing system refer to Kachalov (1984, eq. 12), Babich, Buldyrev & Molotkov (1985, eq. 2.18), Norris, White & Schreiffer (1987, eq. 2.25ab). Here we shall directly derive the Riccati equation for the 3×3 spatial submatrix of matrix (124) and then switch to ray–centred coordinates in order to decouple it.

Differentiating eikonal equation (115) with respect to space–time coordinates and considering (120), we obtain relation

$$N_{4\alpha} = -V^{-1}V_{,\alpha} - V^2 N_i N_{i\alpha} \tag{127}$$

for calculating N_{4i} from N_{ij}, and for calculating N_{44} from N_{4i}.

Differentiating eikonal equation (115) twice with respect to spatial coordinates, and inserting it together with space–time ray tracing equations (117) and with equation (127) into equation $\mathrm{d}\tau_{,jk}/\mathrm{d}\gamma_3 = \tau_{,ij\alpha}\mathrm{d}y_\alpha/\mathrm{d}\gamma_3$, we arrive at the Riccati equation

$$\frac{\mathrm{d}N_{jk}}{\mathrm{d}\gamma_3} = -V^2 N_{ij}N_{ik} - V^{-1}V_{,jk} - VV_{,j}N_{ik}N_i - VV_{,k}N_{ij}N_i + V^4 N_i N_{ij} N_{lk} N_l \tag{128}$$

for symmetric 3×3 matrix N_{ij}. Equations (127) and (128) describe the evolution of the shape of a Gaussian packet in general Cartesian coordinates.

11.2.4. Equations for the second derivatives of phase function in ray–centred coordinates

Covariant derivatives M_{mn} of the phase function in ray–centred coordinates are related to the partial derivatives N_{mn} in Cartesian coordinates through equations

$$N_{jk} = H_{jm} H_{kn} M_{mn} \quad , \tag{129}$$

$$N_{j4} = H_{jm} M_{m4} \quad , \tag{130}$$

$$N_{44} = M_{44} \quad . \tag{131}$$

Here H_{im} is the i–th component of m–th basis vector of the ray–centred coordinate system, see Section 5.1.

Riccati equation (128) for matrix N_{ij} is equivalent to the Riccati equation for the 3×3 matrix M_{mn}, see Norris, White & Schreiffer (1987, eq. 4.8d), with the corresponding linear dynamic ray tracing system (Norris, White & Schreiffer, 1987, eq. 4.8ab). An analogous Riccati equation for the 3×3 matrix of the second partial derivatives in ray–centred coordinates was derived by Babich & Ulin (1981, eqs. 2.7–2.9). The evolution equations for the 2×2 submatrix M_{MN}, corresponding to the plane perpendicular to the ray, are decoupled, i.e. they do not depend upon the outer elements M_{m3}. These equations are identical with the equations for Gaussian beams and can be treated in the same way. The equations for M_{m3} can be solved analytically (Babich and Ulin, 1981). Here we shall stick to the equations for the 3×3 matrix of the second covariant derivatives M_{mn} in ray–centred coordinates.

Equation (127) yields

$$M_{i4} = -V M_{i3} - V^{-1} V^{(q)}_{i} \tag{132}$$

and

$$M_{44} = V^2 M_{33} + V^{(q)}_{3} \quad , \tag{133}$$

where velocity gradient $V^{(q)}_{m}$ in ray–centred coordinates is defined by equation (102).

Equation (128), transformed using (129) into ray–centred coordinates, reads

$$\begin{aligned} \frac{\mathrm{d}M_{mn}}{\mathrm{d}\gamma_3} = & - M_{jn} \frac{\mathrm{d}H_{ij}}{\mathrm{d}\gamma_3} H_{im} - M_{mk} \frac{\mathrm{d}H_{ik}}{\mathrm{d}\gamma_3} H_{in} - V^2 M_{Km} M_{Kn} \\ & - V^{-1} V^{(q)}_{mn} - V^{(q)}_{m} M_{3n} - V^{(q)}_{n} M_{3m} \quad , \end{aligned} \tag{134}$$

where covariant velocity derivatives $V^{(q)}_{mn}$ in ray–centred coordinates are defined by equation (105).

Inserting (101) for the derivatives of the basis vectors, equation (134) may be expanded into three equations

$$\frac{\mathrm{d}M_{MN}}{\mathrm{d}\gamma_3} = -V^2 M_{MK} M_{KN} - V^{-1} V^{(q)}_{MN} \quad , \tag{135}$$

$$\frac{\mathrm{d}M_{M3}}{\mathrm{d}\gamma_3} = -M_{Mi} V^{(q)}_{i} - V^2 M_{MI} M_{I3} - V^{-1} V^{(q)}_{M3} \quad , \tag{136}$$

$$\frac{\mathrm{d}M_{33}}{\mathrm{d}\gamma_3} = -2 M_{i3} V^{(q)}_{i} - V^2 M_{3I} M_{I3} - V^{-1} V^{(q)}_{33} \quad . \tag{137}$$

Equations (135) to (137) for the second covariant derivatives of the phase function in ray–centred coordinates are equivalent to the equations by Babich & Ulin (1981, eqs. 2.7–2.9) for the second partial derivatives of the phase function in ray–centred coordinates.

Riccati equation (135) is *identical with the equation for the corresponding Gaussian beam* and its solution is given by equation (111). Equations (136) and (137) extend the Gaussian–beam solution to a Gaussian–packet solution.

11.2.5. Solving the equations for a Gaussian packet

Denoting

$$\mathbf{M}_3 = \begin{pmatrix} M_{13} \\ M_{23} \end{pmatrix} \quad , \qquad \mathbf{M}_4 = \begin{pmatrix} M_{14} \\ M_{24} \end{pmatrix} \quad , \qquad \mathbf{v}^{(q)} = \begin{pmatrix} V_1^{(q)} \\ V_2^{(q)} \end{pmatrix} \quad , \tag{138}$$

we may express the solutions of equations (132), (133), (136) and (137) as

$$\mathbf{M}_3 = -V^{-1}\mathbf{M}_4 - V^{-2}\mathbf{v}^{(q)} \quad , \tag{139}$$

$$M_{33} = V^{-2}\left[M_{44} - V_3^{(q)}\right] \quad , \tag{140}$$

$$M_{34} = -V^{-1}M_{44} \quad , \tag{141}$$

$$\mathbf{M}_4 = [(\mathbf{Q}_1 + \mathbf{Q}_2\mathbf{M}^0)^{-1}]^{\mathrm{T}}\mathbf{M}_4^0 \quad , \tag{142}$$

$$M_{44} = M_{44}^0 - \mathbf{M}_4^{0\,\mathrm{T}}(\mathbf{Q}_1 + \mathbf{Q}_2\mathbf{M}^0)^{-1}\mathbf{Q}_2\,\mathbf{M}_4^0 \tag{143}$$

(Klimeš, 2004b). Equation (142) is convenient for calculating vector $\mathbf{M}_4$ along the central spatial ray. To calculate M_{44}, we may use equation (143). Quantities $\mathbf{M}_3$, M_{33} and M_{34} can then be obtained through simple relations (139) to (141).

Non–vanishing vector (142) implies that the Gaussian packet is asymmetric with respect to the central ray. The evolution of the shape of a Gaussian packet along the central ray is illustrated in Figure 23.

11.2.6. Amplitude of a Gaussian packet

The equations for the vectorial amplitude of the Gaussian packet are identical with equations (112) and (113) for the corresponding Gaussian beam.

11.3. Optimization of the shape of Gaussian beams or packets

Gaussian beams and packets are high–frequency approximate solutions of the elastodynamic equation. The accuracy of these approximate solutions depends on their shape, especially on their width. The “optimum” shape depends on the distance from the source (Klimeš, 1989a).

The optimum shape of Gaussian beams depends on the error of the Gaussian–beam solution of the elastodynamic equation, which is, unfortunately, unknown for general media. However, in considerably heterogeneous media, the optimum Gaussian beams are very close to the narrowest Gaussian beams.

The root–mean–square width of each individual Gaussian beam between the source and receiver can be minimized using the algorithm by Klimeš (1989a), see Figure 24.

The same algorithm can be used to minimize the width of a Gaussian packet, measured perpendicularly to the central ray. The optimum Gaussian packet is *symmetric* with respect to its central ray, see Figure 23. The error of the Gaussian–packet solution of the elastodynamic equation is smaller for longer Gaussian packets, but the

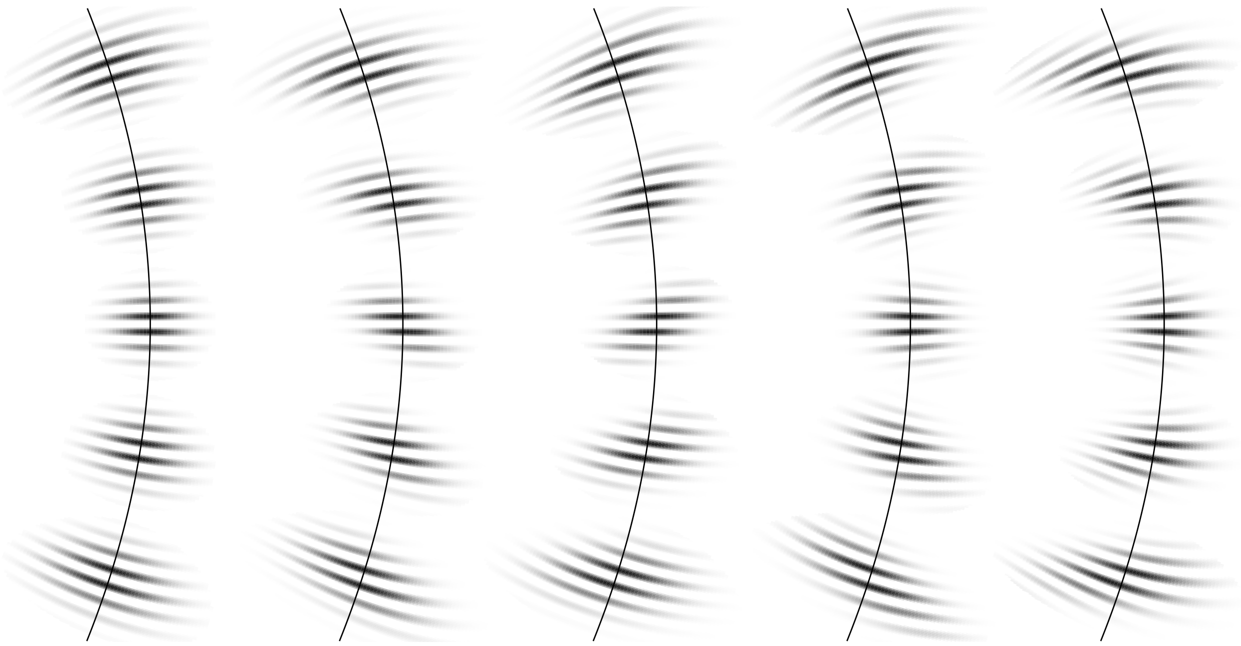

Figure 23. Symmetric and asymmetric Gaussian packets propagating along a spatial ray. All five packets correspond to Gaussian beams of equal shapes. The first Gaussian packet is axially symmetric with respect to its central ray. The second and third packets have envelopes oblique with respect to their central rays. The frequencies of the second and third packets at the waists of the corresponding Gaussian beams are symmetric with respect to the central rays. The fourth and fifth packets have different frequencies to the left and to the right of the central rays. We may call them "dispersive". The envelopes of the fourth and fifth packets at the waists of the corresponding Gaussian beams are symmetric with respect to the central rays. During the evolution of Gaussian packets, the dispersion of the wavefronts generates obliquity of the envelopes and vice versa. The asymmetric Gaussian packets are thus both oblique and dispersive in general.

length of Gaussian packets is limited by the accuracy of their paraxial approximation. This problem with long Gaussian packets may be overcome by decomposing each long Gaussian packet into the summation of shorter ones, similarly as composing a Gaussian beam of Gaussian packets in the asymptotic summation of Gaussian packets. Each long Gaussian packet is then numerically calculated as a combination of the paraxial approximations of a Gaussian packet from different reference points (Žáček, 2004).

To preserve the accuracy of the approximate decomposition of a wavefield into Gaussian beams or packets (Section 11.5), the optimization of the shape of the individual beams or packets should be supplemented with smoothing the dependence of the shape on the summation parameters (Žáček, 2001a, 2001b).

A possibility to optimize clearly the shape of Gaussian beams or packets in the asymptotic superposition integral (Sections 11.4.1 and 11.4.2) is probably the main difference of the summation of Gaussian beams and packets from the integral representations formally derived using the Maslov method, coherent–state method, or their various generalizations, notwithstanding formally equal integral representations.

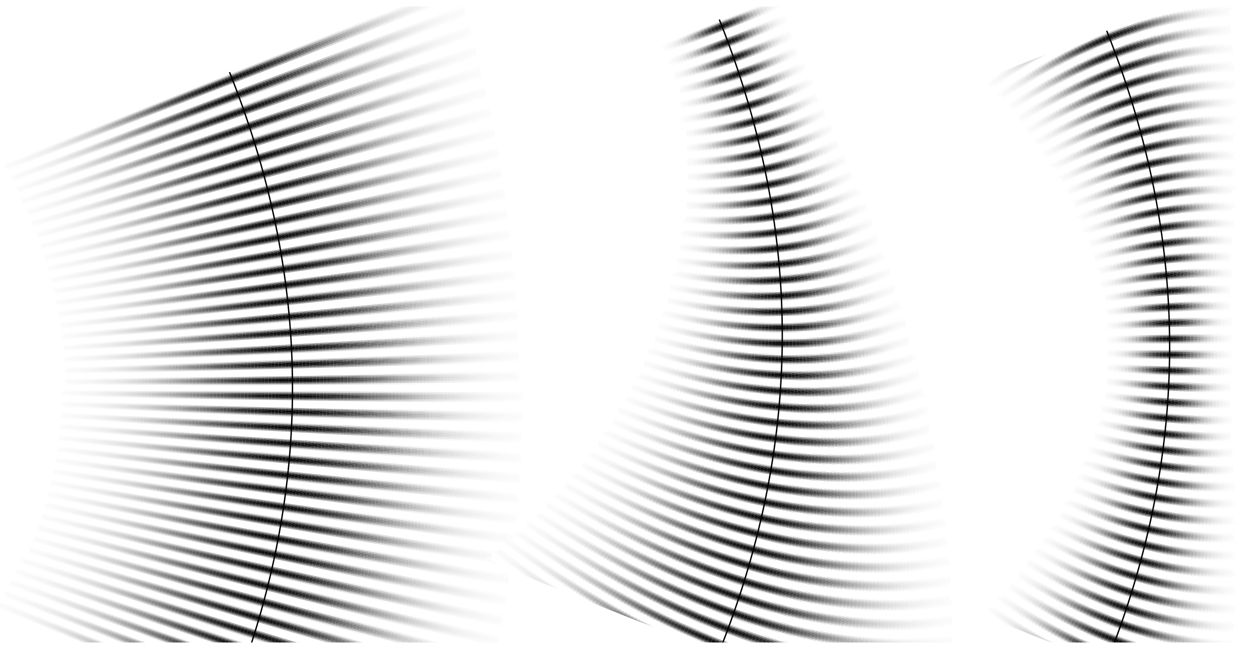

Figure 24. Optimization of the shape of Gaussian beams. The left Gaussian beam is too wide. The middle Gaussian beam has been selected too narrow at its top end, which results in its extensive spreading downwards. The narrowest Gaussian beam with respect to the mean square width is displayed on the right.

11.4. Asymptotic summation of Gaussian beams and packets

11.4.1. Asymptotic decomposition into Gaussian beams

A general time–harmonic wavefield given along a surface, decomposed into individual elementary waves and expressed in terms of the vectorial amplitude and travel time, can be asymptotically expressed as a two–parametric integral superposition of Gaussian beams (Klimeš, 1984a).

The two–parametric integral superposition of Gaussian beams, corresponding to ray–theory wavefield

$$\mathbf{u}^{\mathrm{R}} = \mathbf{U}^{\mathrm{R}} \exp(\mathrm{i}\omega\tau^{\mathrm{R}}) \quad , \tag{144}$$

may be expressed as (Klimeš, 1984a, eq. 77)

$$\mathbf{u}(x_i,\omega) = \frac{\omega}{2\pi} \iint \mathrm{d}\gamma_1 \mathrm{d}\gamma_2 \; \mathbf{U}^{\mathrm{R}} \; |\det(\mathbf{Q}^{(q)\mathrm{R}})| \; \sqrt[2]{\det}[\mathrm{i}\,(\mathbf{M}^{\mathrm{R}} - \mathbf{M})] \times \exp\{\mathrm{i}\omega[\tau^{\mathrm{R}} + p_k^{\mathrm{R}}(x_k - y_k) + \tfrac{1}{2}(x_k - y_k)N_{kl}(x_l - y_l)]\} \quad , \tag{145}$$

where the function $\sqrt[2]{\det}$ of a symmetric complex–valued matrix is defined by the following two conditions: (a) $\sqrt[2]{\det}(\mathbf{B})$ is the product of the second roots of the eigenvalues of matrix $\mathbf{B}$, (b) the roots with the greatest real parts are selected. All the quantities on the right–hand side (except x_i) are taken at the reference point $y_i = y_i(\gamma_1,\gamma_2)$ selected for the paraxial approximation of a Gaussian beam concentrated to the ray with ray parameters γ_1,γ_2. The reference points are usually determined as the points of intersection of rays with a given reference surface, see Figure 25. Receivers should be situated in the vicinity of the reference surface.

Here τ^{R} and p_k^{R} are the ray–theory travel time and ray–theory slowness vector, taken at point y_i. Similarly $\mathbf{U}^{\mathrm{R}}$ is the complex–valued vectorial ray–theory amplitude

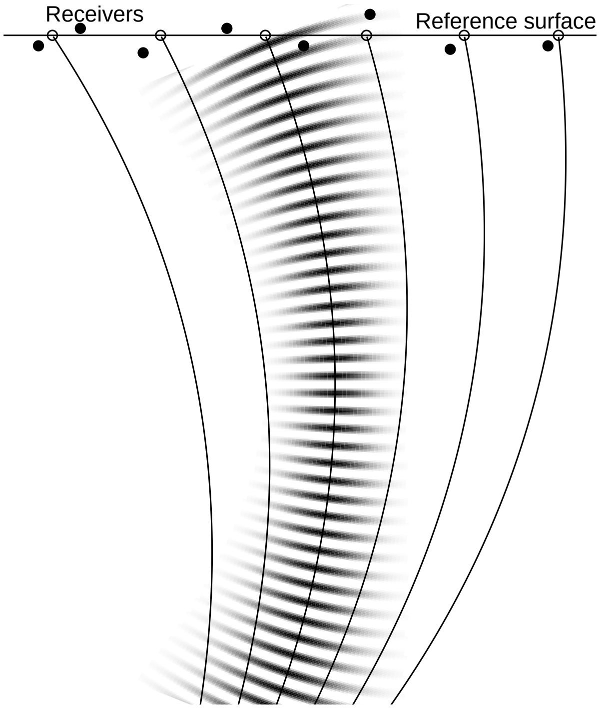

Figure 25. Asymptotic summation of Gaussian beams. Receivers are situated in the vicinity of the reference surface. The points of intersection of rays with the reference surface are selected as the reference points for the paraxial approximation of Gaussian beams.

at point y_i. Matrix $\mathbf{M}^{\mathrm{R}}$ is the 2×2 matrix of the second derivatives of the ray–theory travel time with respect to ray–centred coordinates q_1, q_2 perpendicular to the central ray. Matrix $\mathbf{M}$ is the 2×2 matrix of the second derivatives of the complex–valued travel time of the Gaussian beam with respect to ray–centred coordinates q_1, q_2. Matrix $\mathbf{M}$ controls the shape of Gaussian beams and should be selected with the positive–definite imaginary part. Matrix N_{kl} is the 3×3 matrix of the second derivatives of the complex–valued travel time of the Gaussian beam with respect to Cartesian coordinates x_i, and is completely determined by the 2×2 matrix $\mathbf{M}$ (Klimeš, 1984a, eqs. 77 and 78). The transformation Jacobian from ray parameters γ_1, γ_1 to ray–centred coordinates q_1, q_2 is $|\det(\mathbf{Q}^{(q)\mathrm{R}})|$, where $\mathbf{Q}^{(q)\mathrm{R}}$ is the 2×2 matrix of geometrical spreading in ray–centred coordinates, see (39).

The asymptotic integral superposition (145) is independent of the selection of the surface for decomposition of the ray–theory wavefield into Gaussian beams (Klimeš, 1984a). That is why the ray–theory wavefield may asymptotically be decomposed into Gaussian beams locally, even in its singular regions. The accuracy of the integral superposition of Gaussian beams thus depends on the optimization of the shape of beams only, not on the surface for decomposition. The variation of the optimum shape of beams with distance from the source is quite different from the evolution of the individual Gaussian beams along the same central ray. In other words, we decompose the wavefield into different Gaussian beams for different propagation distances.

11.4.2. Asymptotic decomposition into Gaussian packets

A time–harmonic Gaussian beam may be expressed as a one–parametric integral superposition of space–time Gaussian packets. A general time–harmonic wavefield, specified in terms of the amplitude and travel time, can be asymptotically expressed as a three–parametric integral superposition of space–time Gaussian packets (Klimeš, 1984b, 1989b).

The three–parametric integral superposition of Gaussian packets corresponding to the ray–theory wavefield (144) may be expressed as (Klimeš, 1984b, eq. 51; 1989b, eq. 31)

$$\mathbf{u}(x_i,\omega) = \left(\frac{\omega}{2\pi}\right)^{3/2} \iiint \mathrm{d}\gamma_1 \mathrm{d}\gamma_2 \mathrm{d}\gamma_3 \; \mathbf{U}^{\mathrm{R}} \; V \; |\det(\mathbf{Q}^{(q)\mathrm{R}})| \; \sqrt[2]{\det}[\mathrm{i}\,(\mathbf{N}^{\mathrm{R}} - \mathbf{N})]$$
$$\times \exp\{\mathrm{i}\omega[\tau^{\mathrm{R}} + p_k^{\mathrm{R}}(x_k - y_k) + \tfrac{1}{2}(x_k - y_k)N_{kl}(x_l - y_l)]\} \quad (146)$$

where all the quantities on the right–hand side (except x_k) are taken at the central point $y_i = y_i(\gamma_1, \gamma_2, \gamma_3)$ of a Gaussian packet. Points $y_i = y_i(\gamma_1, \gamma_2, \gamma_3)$ are the points of spatial rays corresponding to ray–theory wavefield $\mathbf{u}^{\mathrm{R}}$. The quadrature is performed over ray parameters γ_1, γ_2 and over travel time γ_3 along the rays, see Figures 26 and 27.

Here τ^{R}, p_k^{R} and N_{kl}^{R} are the ray–theory travel time, ray–theory slowness vector and the matrix of second derivatives of the ray–theory travel time in Cartesian coordinates x_i, taken at point y_i. Similarly $\mathbf{U}^{\mathrm{R}}$ is the matrix of the complex–valued vectorial ray–theory amplitude at point y_i. The transformation Jacobian from ray parameters γ_l to Cartesian coordinates x_i is $V \; |\det(\mathbf{Q}^{(q)\mathrm{R}})|$, where V is the wave–propagation velocity and $\mathbf{Q}^{(q)\mathrm{R}}$ is the 2×2 matrix of geometrical spreading in ray–centred coordinates (Červený, 2001).

Matrix N_{kl} is the 3×3 matrix of second derivatives of the complex–valued phase function of the Gaussian packet with respect to Cartesian coordinates x_i. Matrix N_{kl} controls the shape of Gaussian packets and should be selected with the positive–definite imaginary part.

The properties of the integral superposition of space–time Gaussian packets follow from the properties of the integral superposition of Gaussian beams.

11.4.3. Discretization error

In numerical algorithms, the integral superposition (145) of Gaussian beams is discretized into the summation over rays. These rays are calculated by controlled initial–value ray tracing. We thus have to divide the ray–parameter domain into elements corresponding to individual rays used in the summation. We first divide the ray–parameter domain into areas corresponding to the individual ray histories. We then divide each area into elements corresponding to the individual rays. We choose each element as the Voronoi polygon of the corresponding ray, i.e. as the set of points, which are closer to the corresponding ray than to other traced rays of equal ray history. Since the discretization error is approximately controlled by contravariant ray–parameter metric tensor

$$\mathbf{G}_{(\gamma)} = (\mathbf{Q}^{(q)\mathrm{R}})^{\mathrm{T}} \, (\mathbf{M} - \mathbf{M}^{\mathrm{R}})^{*} \; [\mathrm{Im}(\mathbf{M})]^{-1} \; (\mathbf{M} - \mathbf{M}^{\mathrm{R}}) \; \mathbf{Q}^{(q)\mathrm{R}} \qquad (147)$$

(Klimeš 1986), the Voronoi polygons are defined with respect to this metric tensor. Here the 2×2 matrices $\mathbf{Q}^{(q)\mathrm{R}}$, $\mathbf{M}^{\mathrm{R}}$ and $\mathbf{M}$ are expressed in ray–centred coordinates. Matrix $\mathbf{M}$ controls the shape of the Gaussian beam concentrated at the ray, and should be determined by optimization of the shape of the Gaussian beam. Matrix $\mathbf{Q}^{(q)\mathrm{R}}$ of geometrical spreading and matrix $\mathbf{M}^{\mathrm{R}}$ of the second derivatives of travel time describe the orthonomic system of rays, at which the Gaussian beams are concentrated. Contravariant ray–parameter metric tensor (147) varies smoothly within each ray history, but is often discontinuous at the boundary between different ray histories. Since we

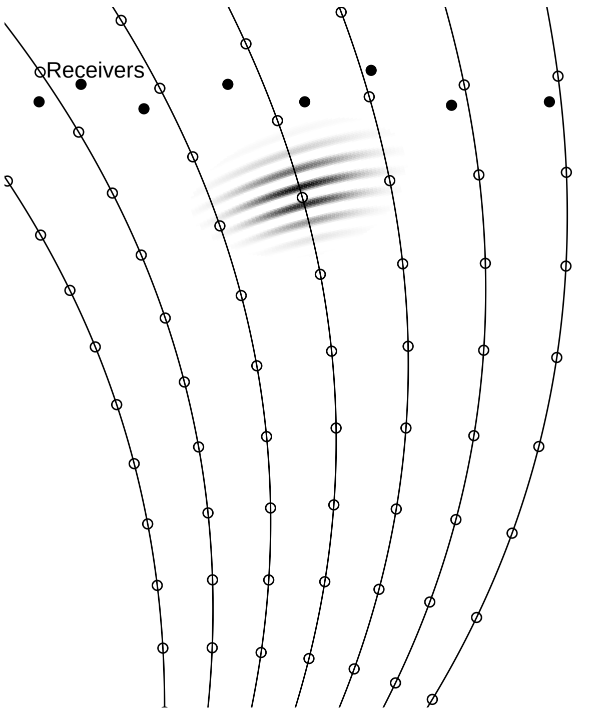

Figure 26. Asymptotic summation of Gaussian packets. The central points for the paraxial approximation of Gaussian packets are regularly distributed along rays. The density of rays and of central points depends on the shape of the Gaussian packets. Receivers may be situated arbitrarily, no reference surfaces are required.

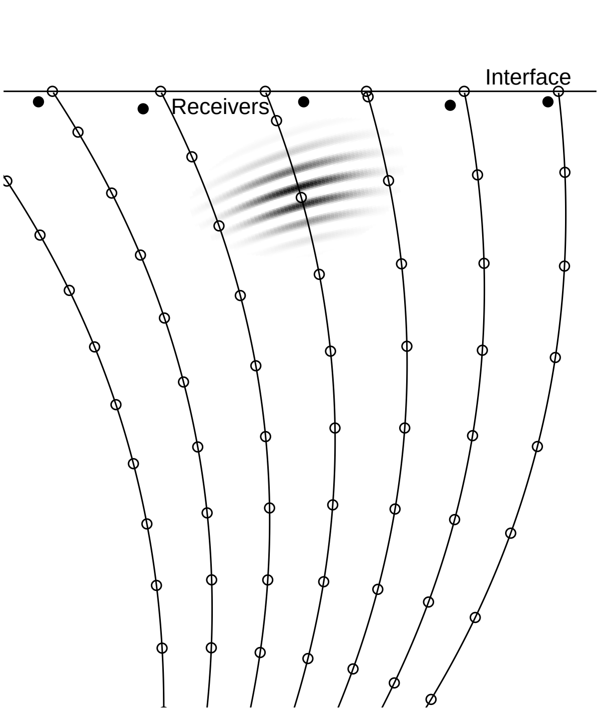

Figure 27. Summation of Gaussian packets at an interface. Paraxial approximation cannot be applied across the interface. The points of intersection of rays with the interface serve as the reference points for the paraxial approximation of Gaussian packets centred beyond the interface. Receivers may be situated arbitrarily.

know the values of the metric tensor only at the ray–parameter points corresponding to the calculated rays, we can only determine the elements as an approximation to the Voronoi polygons.

The error due to the discretization depends on frequency and on the shape of the Gaussian beams and can be controlled by the selection of the maximum distance between rays with respect to ray–parameter metric (147). This maximum distance between rays influences the sizes and shapes of the Voronoi polygons. The size and shape of a Voronoi polygon can be quantified by its moment with respect to the corresponding ray–parameter point, which represents the central ray of a Gaussian beam. Unfortunately, the exact equations for the discretization into an irregular system of central rays of Gaussian beams are not known yet.

The discretization error applicable to a system of central rays regular with respect to metric (147), with smooth variation of the weights of Gaussian beams in the decomposition, was derived by Klimeš (1986), for 2–D also by Hill (1990). This "regular" error represents the lower estimate of the discretization error. Since the upper estimate based on the discretization of the integral over a single Voronoi polygon into a single Gaussian beam (Klimeš, 1985) is too conservative, the "regular" error is now used as the "optimistic" approximation of the discretization error also for an irregular system of central rays of Gaussian beams. The assumption of smooth variation of the weights of Gaussian beams corresponds to the conditions for the integral superposition of Gaussian beams. This smooth variation includes wavefields generated by various point sources, or wavefields with smooth variation of the amplitude and travel time along an initial surface.

The error due to the discretization of the integral superposition of Gaussian packets can be controlled analogously to the error due to the discretization of the integral superposition of Gaussian beams. Refer to Klimeš (1989b) for more details.

It is even more difficult to estimate the discretization error for the decomposition of a general 3–D wavefield into Gaussian packets, described in Section 11.5. The upper estimate of the discretization error for the decomposition of a general 1–D wavefield into the regular two–parametric system of Gabor functions was derived by Daubechies (1991, eq. 3.2.4).

11.4.4. Linear canonical transforms

Various superposition integrals similar or equivalent to the asymptotic summation of Gaussian beams or to the asymptotic summation of Gaussian packets have been derived by many authors using the Maslov method, its various generalizations, or coherent–state transforms. The generalizations of the Maslov method are based on the linear canonical transform or on its special cases such as the fractional Fourier transform. Before discussing these integral transform methods, we briefly recall the definition of the linear canonical transform and mention their relation to the coherent–state transform.

Spatial coordinate x_i in phase space corresponds to operator $\hat{x}_i$ of the multiplication by x_i. The momentum coordinate p_i^x in phase space corresponds to the differential operator

$$\hat{p}_i = \frac{1}{\mathrm{i}\omega}\frac{\partial}{\partial x_i} \quad . \tag{148}$$

Let us consider a linear coordinate transform in phase space,

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{p}^x \end{pmatrix} = \boldsymbol{\mathcal{M}} \begin{pmatrix} \mathbf{x}' \\ \mathbf{p}^{x'} \end{pmatrix} \quad , \tag{149}$$

where the complex–valued 6×6 matrix $\boldsymbol{\mathcal{M}}$ may be composed of four 3×3 matrices,

$$\boldsymbol{\mathcal{M}} = \begin{pmatrix} \tilde{\mathbf{Q}} & \mathbf{Q} \\ \tilde{\mathbf{P}} & \mathbf{P} \end{pmatrix} \quad . \tag{150}$$

We shall see later that matrix $\boldsymbol{\mathcal{M}}$ should be symplectic. The linear coordinate transform corresponds to the linear transform

$$\begin{pmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{p}}^x \end{pmatrix} = \boldsymbol{\mathcal{M}} \begin{pmatrix} \hat{\mathbf{x}}' \\ \hat{\mathbf{p}}^{x'} \end{pmatrix} \tag{151}$$

of the operators. The *linear canonical transform* is the linear transform

$$f_{\mathcal{M}}(\mathbf{x}) = \int \mathrm{d}^3\mathbf{x}' \; C_{\mathcal{M}}(\mathbf{x},\mathbf{x}')\, f(\mathbf{x}') \tag{152}$$

of functions, with integration kernel of the form

$$C_{\mathcal{M}}(\mathbf{x},\mathbf{x}') = c_{\mathcal{M}} \exp\left[\frac{\mathrm{i}\omega}{2} \begin{pmatrix} \mathbf{x} \\ \mathbf{x}' \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} \mathbf{N}_{\mathcal{M}}^{xx} & \mathbf{N}_{\mathcal{M}}^{xx'} \\ \mathbf{N}_{\mathcal{M}}^{x'x} & \mathbf{N}_{\mathcal{M}}^{x'x'} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{x}' \end{pmatrix} \right] \quad , \tag{153}$$

which corresponds to transform (151) of operators. Applying both sides of operator transform (151) with (148) and (150) to the right–hand side of equation (152) and assuming that the integrand vanishes at the boundary of the integration volume, we can prove that matrix (150) must be *symplectic* and that

$$\mathbf{N}^{x'x}_{\mathcal{M}} = (\mathbf{N}^{xx'}_{\mathcal{M}})^{\mathrm{T}} = -\mathbf{Q}^{-1} \quad , \tag{154}$$

$$\mathbf{N}^{x'x'}_{\mathcal{M}} = \mathbf{Q}^{-1}\tilde{\mathbf{Q}} \quad , \tag{155}$$

$$\mathbf{N}^{xx}_{\mathcal{M}} = \mathbf{P}\mathbf{Q}^{-1} \quad . \tag{156}$$

The relation of symplectic matrix (150) and integration kernel (153) resembles the relation of the paraxial ray propagator matrix and the paraxial two–point eikonal (Arnaud, 1972, eq. 11.1; Červený, Klimeš & Pšenčík, 1984, eq. 22).

The inverse transforms to (149) and (151) are

$$\begin{pmatrix} \mathbf{x}' \\ \mathbf{p}^{x'} \end{pmatrix} = \mathcal{M}^{-1} \begin{pmatrix} \mathbf{x} \\ \mathbf{p}^{x} \end{pmatrix} \tag{157}$$

and

$$\begin{pmatrix} \hat{\mathbf{x}}' \\ \hat{\mathbf{p}}^{x'} \end{pmatrix} = \mathcal{M}^{-1} \begin{pmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{p}}^{x} \end{pmatrix} \quad , \tag{158}$$

where the inverse matrix to symplectic matrix $\mathcal{M}$ is

$$\mathcal{M}^{-1} = \begin{pmatrix} \mathbf{P}^{\mathrm{T}} & -\mathbf{Q}^{\mathrm{T}} \\ -\tilde{\mathbf{P}}^{\mathrm{T}} & \tilde{\mathbf{Q}}^{\mathrm{T}} \end{pmatrix} \quad . \tag{159}$$

The inverse transform to (152) is

$$f_{\mathcal{M}^{-1}}(\mathbf{x}') = \int \mathrm{d}^3\mathbf{x}\; C_{\mathcal{M}^{-1}}(\mathbf{x}',\mathbf{x})\, f(\mathbf{x}) \quad . \tag{160}$$

Integration kernel $C_{\mathcal{M}^{-1}}(\mathbf{x}',\mathbf{x})$ is obtained from symplectic matrix $\mathcal{M}^{-1}$ by equations analogous to (153) and (154) to (156).

Constants $c_{\mathcal{M}}$ and $c_{\mathcal{M}^{-1}}$ must be selected so that the composition of transforms (152) and (160) forms the identity transform,

$$c_{\mathcal{M}} = c_{\mathcal{M}^{-1}} = \left[\sqrt[2]{\det}\left(\frac{2\pi}{\omega}\mathbf{Q}\right)\right]^{-1} \quad . \tag{161}$$

Here the matrix function $\sqrt[2]{\det}$ is defined by the following two conditions: (a) $\sqrt[2]{\det}(\mathbf{A})$ is the product of the fourth roots of the eigenvalues of matrix $\mathbf{A}^{\mathrm{T}}\mathbf{A}$, (b) the roots with the greatest real parts are selected. For more details refer to Wolf (1974, 1979) or Ozaktas, Zalevsky & Kutay (2001).

Interesting special cases of 3–D linear canonical transform (152) are 3–D Fourier transform

$$\mathcal{M} = \begin{pmatrix} \mathbf{0} & -\mathbf{I} \\ \mathbf{I} & \mathbf{0} \end{pmatrix} \quad , \tag{162}$$

and 3–D separable fractional Fourier transform (Condon, 1937; Ozaktas, Zalevsky & Kutay, 2001)

$$\mathcal{M} = \begin{pmatrix} \cos(\alpha_1) & 0 & 0 & -\sin(\alpha_1) & 0 & 0 \\ 0 & \cos(\alpha_2) & 0 & 0 & -\sin(\alpha_2) & 0 \\ 0 & 0 & \cos(\alpha_3) & 0 & 0 & -\sin(\alpha_3) \\ \sin(\alpha_1) & 0 & 0 & \cos(\alpha_1) & 0 & 0 \\ 0 & \sin(\alpha_2) & 0 & 0 & \cos(\alpha_2) & 0 \\ 0 & 0 & \sin(\alpha_3) & 0 & 0 & \cos(\alpha_3) \end{pmatrix} \quad . \tag{163}$$

For real–valued, imaginary–valued or complex–valued α_i, (163) defines real–ordered, imaginary–ordered or complex–ordered fractional Fourier transforms, respectively.

11.4.5. Coherent state transforms

The paraxial approximation of a Gaussian packet (coherent state) in the vicinity of its central point $\mathbf{y}$ is performed in local Cartesian coordinates

$$\mathbf{r} = \mathbf{x} - \mathbf{y} \quad . \tag{164}$$

It is thus also reasonable to consider linear canonical transforms (152) and (160) with respect to local Cartesian coordinates $\mathbf{r}$. For a given function f, we define new function

$$\phi(\mathbf{r},\mathbf{y}) = f(\mathbf{r}+\mathbf{y}) \quad . \tag{165}$$

The local inverse linear canonical transform

$$\psi(\mathbf{r}',\mathbf{y}) = \phi_{\mathcal{M}^{-1}}(\mathbf{r}',\mathbf{y}) \tag{166}$$

of function $\phi(\mathbf{r},\mathbf{y})$ with respect to the first argument is called the *coherent state transform* of function $f(\mathbf{x})$ (Klauder, 1987, eq. 3; Foster & Huang, 1991, eq. 39; Thomson, 2001, eq. 2.4). Function $\phi(\mathbf{r},\mathbf{y})$ may be obtained from the coherent state transform $\psi(\mathbf{r}',\mathbf{y})$ by local linear canonical transform

$$\phi(\mathbf{r},\mathbf{y}) = \psi_{\mathcal{M}}(\mathbf{r},\mathbf{y}) \quad . \tag{167}$$

Function $f(\mathbf{x})$ may be restored from $\phi(\mathbf{r},\mathbf{y})$ in many ways, yielding various linear transforms of operators $\hat{\mathbf{x}}$ and $\hat{\mathbf{p}}^x$ to operators $\hat{\mathbf{r}}'$, $\hat{\mathbf{p}}^{r'}$, $\hat{\mathbf{y}}'$ and $\hat{\mathbf{p}}^{y'}$. Functional transform

$$f(\mathbf{x}) = \phi(\mathbf{0},\mathbf{x}) \tag{168}$$

(Klauder, 1987, eq. 4; Foster & Huang, 1991, eq. 40; Thomson, 2001, eq. 2.6) yields operator transform

$$\begin{pmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{p}}^x \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{y}} \\ \hat{\mathbf{p}}^y \end{pmatrix} \quad , \tag{169}$$

whereas functional transform

$$f(\mathbf{x}) = \phi(\mathbf{y},\mathbf{x}-\mathbf{y}) \tag{170}$$

yields operator transform

$$\begin{pmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{p}}^x \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{y}}+\hat{\mathbf{r}} \\ \hat{\mathbf{p}}^y \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{y}}+\tilde{\mathbf{Q}}\hat{\mathbf{r}}'+\mathbf{Q}\hat{\mathbf{p}}^{r'} \\ \hat{\mathbf{p}}^y \end{pmatrix} \tag{171}$$

actually used by Foster & Huang (1991, eq. 46) and Thomson (2001, eqs. 4.1 and 4.2). Functional transform

$$f(\mathbf{x}) = \phi(\mathbf{x}-\mathbf{y},\mathbf{y}) \quad , \tag{172}$$

naturally corresponding to equations (164) and (165), yields operator transform

$$\begin{pmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{p}}^x \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{y}}+\hat{\mathbf{r}} \\ \hat{\mathbf{p}}^r \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{y}} \\ \mathbf{0} \end{pmatrix} + \mathcal{M}\begin{pmatrix} \hat{\mathbf{r}}' \\ \hat{\mathbf{p}}^{r'} \end{pmatrix} \quad . \tag{173}$$

Functional transform

$$f(\mathbf{x}) = \phi(\mathbf{x},\mathbf{0}) \tag{174}$$

reduces the coherent state transform to the linear canonical transform.

11.4.6. Maslov methods

The standard ray theory is derived and expressed in the "coordinate representation", i.e. with respect to the spatial coordinates. In order to obtain various special cases of the summation of Gaussian beams and packets, the high–frequency approximation may be developed with respect to 3 appreciably general "representation coordinates" $\mathbf{x}'$ chosen in 6–D phase space, see (157), and then transformed to the coordinate representation.

The original Maslov method (Maslov, 1965; Chapman & Drummond, 1982) consists in weighted combination of the standard ray–theory approximation with the Maslov methods of the first, second and third order. The Maslov method of the first order corresponds to one spatial coordinate replaced by the respective momentum coordinate. The Maslov method of the second or third order corresponds to two or three spatial coordinates replaced by the respective momentum coordinates, see (162). We obtain the generalized eikonal and transport equations by the high–frequency approximation in this representation, and solve them. The travel time in this representation is the Legendre transform of the ray–theory travel time with respect to 1, 2 or 3 coordinates. The high–frequency asymptotic approximation of the wavefield is then transformed to the coordinate representation by the 1–D, 2–D or 3–D Fourier transform, respectively. This Fourier transform forms the superposition integral. The Maslov method of the first order represents the one–parametric superposition of infinitely broad "ribbon" Gaussian beams with second–order derivatives of travel time vanishing along the summation lines. The Maslov method of the second order represents the two–parametric superposition of infinitely broad Gaussian beams with second–order derivatives of travel time vanishing along the summation surfaces. The Maslov method of the third order represents the superposition of infinitely broad Gaussian packets with vanishing second–order spatial derivatives of the phase function.

Alonso & Forbes (1998) selected each representation coordinate x'_i as a real–valued linear combination of a spatial coordinate x_i and the corresponding momentum coordinate p^x_i, see (163). They then solved the generalized eikonal and transport equations obtained by the high–frequency approximation in this representation. The travel time in this representation is the separable real–ordered fractional Legendre transform of the ray–theory travel time (Alonso & Forbes 1995). The high–frequency asymptotic approximation of the wavefield is then transformed to the coordinate representation by the 1–D, 2–D or 3–D separable real–ordered fractional Fourier transform (Condon, 1937; Ozaktas, Zalevsky & Kutay, 2001). This real–ordered fractional Fourier transform forms the superposition integral. The resulting approximation represents the superposition of infinitely broad Gaussian beams (1–D, 2–D) or packets (3–D). Equivalent results have been achieved by application of the Maslov method in local curvilinear coordinates or with respect to the "reference travel time" (Kendall & Thomson, 1993). This approximation may artificially be supplemented with Gaussian windowing through the imaginary–ordered fractional Fourier transform, sometimes also called the Gaussian–windowed Fourier transform (Alonso & Forbes, 1998; Forbes & Alonso, 1998; Kravtsov & Orlov, 1999). Analogous Gaussian windowing may be introduced using the coherent state transform (Foster & Huang, 1991).

The Maslov method yields general superpositions of Gaussian beams or packets if the representation coordinates $\mathbf{x}'$ are sufficiently general complex–valued linear combinations of phase–space coordinates $\mathbf{x}$ and $\mathbf{p}^x$ (Klimeš, 1984b, eq. 24). We then solve the generalized eikonal and transport equations obtained by the high–frequency approximation in the new representation. The travel time in the new representation may be obtained by the generalized Legendre transform of the ray–theory travel time (Klimeš, 1984b, eq. 39). The high–frequency asymptotic approximation of the wavefield is then transformed to the coordinate representation by the 3–D complex linear canonical transform (Klimeš, 1984b, eq. 27). Analogous superpositions may also be obtained by means of the coherent state transform (Thomson, 2001).

To obtain a general superposition of Gaussian packets by the Maslov method, we transform the elastodynamic equation to phase–space coordinates $\mathbf{x}'$, $\mathbf{p}^{x'}$, given by general complex–valued symplectic transform (157), by inserting expressions (151) for operators $\hat{\mathbf{x}}$, $\hat{\mathbf{p}}^x$. The high–frequency approximation then yields the generalized eikonal and transport equations. The solutions of these equations may be expressed in terms of the solutions of the standard eikonal and transport equations. The obtained approximate high–frequency solution can then be transformed back to coordinates $\mathbf{x}'$ by 3–D complex linear canonical transform (152), which is a generalization of both the real–ordered fractional Fourier transform and imaginary–ordered fractional Fourier transform (Gaussian–windowed Fourier transform). The resulting approximate high–frequency solution is identical to general 3–parametric superposition (146) of Gaussian packets, including general 2–parametric superposition (145) of Gaussian beams as a special case. Matrix $\mathbf{N}$ controlling the shape of Gaussian packets in (146) is identical to matrix (156),

$$\mathbf{N} \equiv \mathbf{N}^{xx}_{\mathcal{M}} \quad . \tag{175}$$

The approximate high–frequency solution (146) derived in this way is thus independent of the choice of $\mathbf{Q}$ in symplectic matrix (150), provided that we set $\mathbf{P} = \mathbf{NQ}$. Solution (146) is also independent of matrices $\tilde{\mathbf{Q}}$ and $\tilde{\mathbf{P}}$ forming with $\mathbf{Q}$ and $\mathbf{P}$ symplectic matrix $\boldsymbol{\mathcal{M}}$. Klimeš (1984b) thus performed the above described derivation with matrix (150) selected in the special form

$$\boldsymbol{\mathcal{M}} = \begin{pmatrix} \mathbf{0} & -\mathbf{I} \\ \mathbf{I} & -\mathbf{N} \end{pmatrix} \quad , \tag{176}$$

which corresponds to the form

$$\boldsymbol{\mathcal{M}}^{-1} = \begin{pmatrix} -\mathbf{N} & \mathbf{I} \\ -\mathbf{I} & \mathbf{0} \end{pmatrix} \tag{177}$$

of matrix (159).

Note that Foster & Huang (1991) and Thomson (2001) selected matrix (159) for the coherent state transform (166) in the separable form

$$\boldsymbol{\mathcal{M}}^{-1} = \begin{pmatrix} \mathrm{i}\Omega & 0 & 0 & 1 & 0 & 0 \\ 0 & \mathrm{i}\Omega & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad . \tag{178}$$

Operators $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}^r$ are then transformed to operators $\hat{\mathbf{r}}'$ and $\hat{\mathbf{p}}^{r'}$ using equation (151) with matrix

$$\boldsymbol{\mathcal{M}} = \begin{pmatrix} 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & \mathrm{i}\Omega & 0 & 0 \\ 0 & 1 & 0 & 0 & \mathrm{i}\Omega & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad . \tag{179}$$

Compared with equations (175) to (177), where matrix $\mathbf{N}$ has the positive–definite imaginary part, Foster & Huang (1991) and Thomson (2001) selected the analogous 2×2 matrix with the negative–definite imaginary part.

Unfortunately, the mathematical formalism of Maslov methods obscures the evolution of Gaussian beams or Gaussian packets along rays and makes the optimization of their shapes more difficult.

11.5. Decomposition of a general wavefield into Gaussian packets or beams

Assume a time–dependent wavefield specified along a given surface, and call it the "time section". The trace of a Gaussian packet in the time section is approximately a Gabor function. The widths of the envelopes of Gabor functions are inversely proportional to the square root of frequency, not constant as for the Gabor transforms (discrete, integral) nor inversely proportional to frequency as for the wavelet transforms. Moreover, the shape of the packets has to be optimized to some extent with respect to the elastodynamic equation, and is thus often dependent on time and on the coordinates and wavenumbers along the surface. The Gabor functions corresponding to optimized Gaussian packets have envelopes considerably dependent on frequency and on the direction of propagation, and moderately dependent on the position and time. This makes the decomposition of a general wavefield into Gaussian packets intricate. Žáček (2003a, 2003b) generalized the integral Gabor transform towards the approximate integral expansion of a time section into Gabor functions of varying shape (refer to Figure 28 for an illustration). The approximate expansion into Gaussian packets has the form of a linear integral transform. The expansion is exact if the envelopes of Gaussian packets depend on frequency and on the angle of incidence only. The system of Gaussian packets is four–parametric in 2–D and six–parametric in 3–D.

The decomposition of a general wavefield into Gaussian packets of given, optimized envelopes is crucial for Gaussian–packet true–amplitude prestack depth migrations. It may also enable to develop general hybrid methods, combining the Gaussian packet summation method with finite differences, finite elements, and other highly accurate methods, which can be applied only to small parts of large models at high frequencies.

A time–dependent wavefield specified along a given surface can also be Fourier transformed into the frequency domain, where it consists of individual time–harmonic wavefields. Each time–harmonic wavefield may then be decomposed using the oversampled Gabor transform into the four–parametric (in 3–D) system of Gabor functions, which are the traces of Gaussian beams along the surface (Hill, 1990, 2001; Shlivinski et al., 2004).

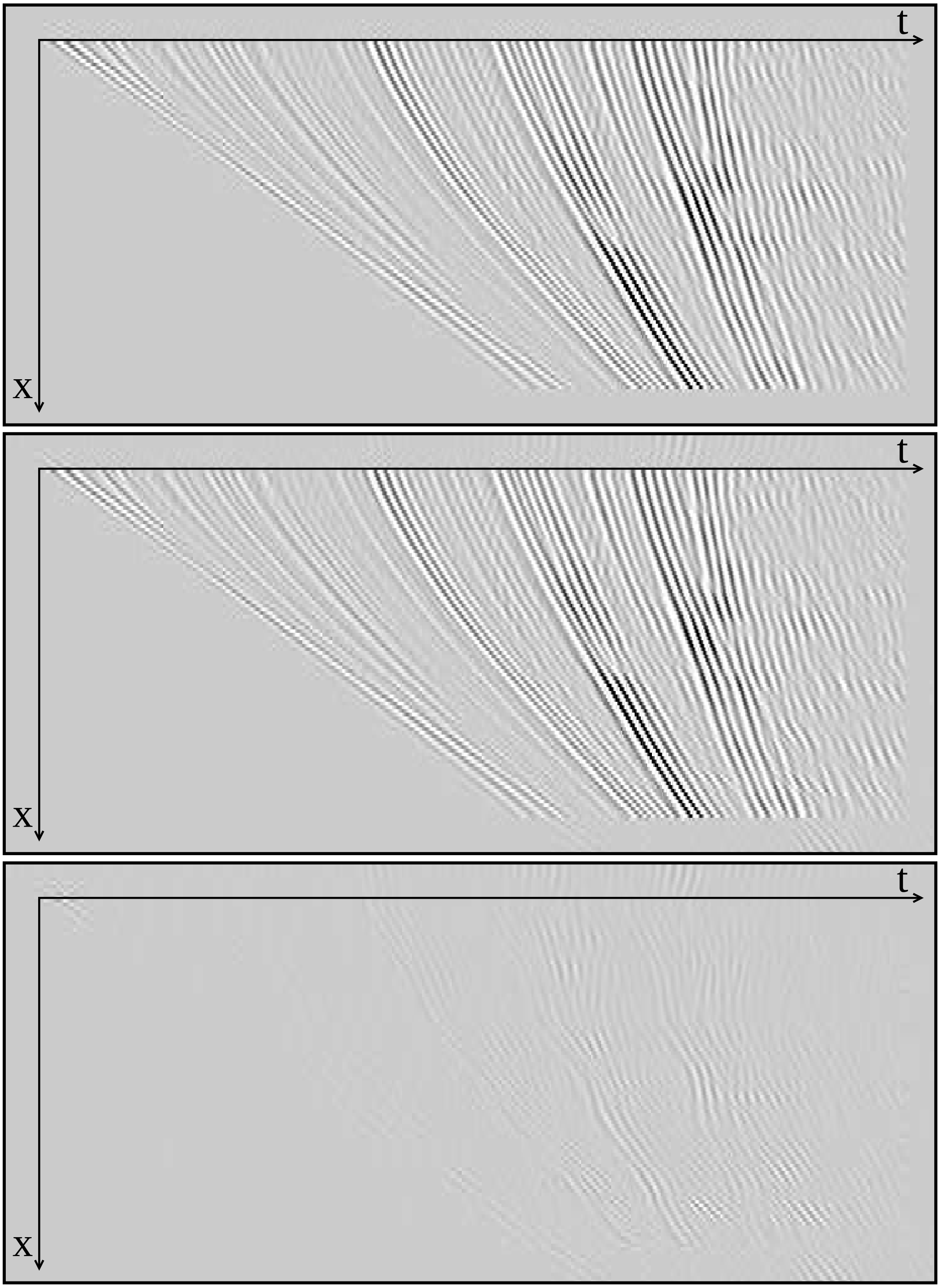

Figure 28. Decomposition of a general wavefield into Gaussian packets. **Top:** Recorded wavefield. **Middle:** Wavefield composed of the Gabor functions corresponding to optimized Gaussian packets (six–parametric summation in 3–D). **Bottom:** Difference between the original wavefield and the wavefield composed of the Gabor functions.

11.6. Sensitivity of waves to heterogeneities

We decompose perturbations of the coefficients of the elastodynamic equation (e.g., elastic moduli and density in the elastodynamic equation) into Gabor functions

$$g(\mathbf{x}) = \exp[\,\mathrm{i}\mathbf{k}^{\mathrm{T}}(\mathbf{x}-\mathbf{y}) - \tfrac{1}{2}(\mathbf{x}-\mathbf{y})^{\mathrm{T}}\mathbf{K}(\mathbf{x}-\mathbf{y})] \tag{180}$$

centred at various spatial positions $\mathbf{y}$ and having various structural wavenumber vectors $\mathbf{k}$. We consider a short–duration incident wavefield with a smooth broad–band frequency spectrum. The wavefield scattered by the perturbations is then composed of waves scattered by the individual Gabor functions. The scattered waves are estimated using the first–order Born approximation with paraxial ray approximation. Each Gabor function usually generates only a few narrow–band space–time Gaussian packets propagating in specific directions (Žáček & Klimeš, 2003). The only exceptions are broad–band forward scattering and broad–band narrow–angle scattering from the lowest structural wavenumbers, and rather rare broad–band critical scattering with mode conversion. Each scattered Gaussian packet is sensitive to just a single linear combination of the coefficients of the elastodynamic equation. This information about the Gabor function is lost if the scattered Gaussian packet does not fall into the aperture covered by the receivers and into the legible frequency band.

11.7. Migrations

A "prestack depth migration" is a simple back–projection of a wavefield, roughly approximating the inversion of wide–angle scattering. It often includes even additional rough approximations. The back–propagated wavefield is compared with the incident wavefield, forming an "image" (convolutional transform) of the gradient of a particular linear combination of the coefficients of the elastodynamic equation.

11.7.1. Gaussian packet migrations

The recorded wavefield (time section) is decomposed into Gaussian packets as described in Section 11.5. The individual Gaussian packets are back–propagated and compared with the incident wavefield (Žáček, 2004). The image of each back–propagated Gaussian packet is approximately formed by one or a few Gabor functions. We thus obtain not only the image of small–scale structural heterogeneities, but also the relation between the time section and the heterogeneities.

The algorithm of the true–amplitude common–source prestack depth migrations based on Gaussian packets consists of the following basic steps: optimization of the model for ray tracing and for Gaussian packets, calculation of travel times and other ray–theory quantities from the source to the dense rectangular grid of points covering the target zone, optimization of the envelopes of Gaussian packets travelling from various parts of the target zone to the receivers, decomposition of the time section into Gaussian packets, and the backprojection of the individual Gaussian packets from the time section onto the migrated image of the target zone.

11.7.2. Gaussian beam migrations

The recorded wavefield, Fourier transformed into the frequency domain, may be decomposed into Gaussian beams as described in Section 11.5. The individual Gaussian beams are back–propagated and compared with the incident wavefield (Hill, 1990, 2001). In the frequency domain, the image of each back–propagated Gaussian beam extends along its whole central ray.

To improve the numerical efficiency, this procedure of Gaussian beam migration can be transformed from the frequency domain to the time domain. Refer to Hill (1990, 2001) for more details.

12. Other extensions of the ray method and related methods

12.1. Higher–order ray approximations

The zero–order ray approximation (1) of the displacement vector can be generalized in the form of an asymptotic ray series in inverse powers of frequency ω. In the frequency domain, the ray–series solution reads:

$$\mathbf{u}(x_i,t) = \exp[-\mathrm{i}\omega(t-\tau(x_i))] \sum_{n=0}^{\infty} (\mathrm{i}\omega)^{-n} \mathbf{U}^{(n)}(x_i) \quad . \tag{181}$$

Here $\mathbf{U}^{(n)}(x_i)$ are vectorial amplitude coefficients of the ray series. The zero–order term ($n=0$) yields the zero–order approximation, and the other terms ($n \geq 1$) are called higher–order approximations. The procedures of determining the higher–order amplitude coefficients ($U^{(1)}, U^{(2)}, ...$) are described in Červený (2001, sections 5.6 and 5.7), where many other references can also be found.

In most cases it is sufficient to consider the zero–order approximation. In regular regions, the higher–order terms are usually small and thus it is not necessary to compute them. In singular regions, the zero–order approximation fails to describe the wavefield properly and the higher–order approximation cannot fix this. Moreover, considering higher–order terms in inhomogeneous isotropic or anisotropic media is generally complicated, because the terms of lower order and their spatial derivatives have to be known. There are, however, certain exceptions.

Such exceptions, in which higher–order approximations may play an important role are situations, in which the zero–order amplitude or some of its component (if we are dealing with a vectorial wavefield) are zero. This occurs, for example, in the vicinity of the Brewster angle, for which R/T coefficients are zero. Another example are nodal lines of seismic sources, in whose directions the radiation function is zero. Higher–order approximations are also important in the study of the problem of R/T at interfaces of orders higher than first (at an interface of the first order, parameters of the medium change discontinuously). A zero–order wave incident at an interface of the second order generates a zero–order transmitted wave but first–order reflected wave, for whose evaluation, the first–order approximation of the ray theory is necessary. To a certain extent, the situation with the zero amplitude of the zero–order wave also concerns head waves, which are generated in postcritical regions by zero–order waves propagating separately from the incident wave, for more details refer to Červený (2001, sec. 5.6) and to references therein.

A special case, in which higher–order terms of the ray series play an important role are the so–called additional components. We concentrate on the first–order additional components of the ray series.

The higher–order amplitude coefficients of the vectorial ray series can be separated into *principal terms*, whose directions coincide with the direction of the zero–order term, and *additional terms*, whose directions are perpendicular to the direction of the zero–order term. Whereas the first–order principal term represents only a certain correction of the zero–order term, the first–order additional terms represent leading terms in the direction perpendicular to the zero–order term. We can thus speak of the vectorial leading term of the ray series, see Eisner & Pšenčík, (1996), or the two–term solution, see Fradkin & Kiselev (1997), where many additional references can be found. See also Goldin (1989). In contrast to the principal terms, which are obtained by integrating along the corresponding rays, the additional terms are local quantities and can be evaluated relatively easily in arbitrary varying media. Spatial derivatives of the parameters of the medium and of the zero–order amplitude term are necessary for their evaluation. The evaluation of the derivatives of the parameters of the medium is straightforward. For the evaluation of the amplitude term, the procedure described by Eisner & Pšenčík (1996), based on the substitution of the derivatives by the differences of values calculated on neighbouring rays, can be used.

Eisner and Pšenčík (1996) showed that the use of the first–order additional terms in the expression for the ray Green function in a homogeneous isotropic medium improves the ray Green function much more than the addition of any other higher–order term (note that the ray series corresponding to the exact Green function in a homogeneous isotropic medium consists of three terms). This indicates that the first–order additional terms should be used routinely. Their most obvious effects are deviations of the polarizations of waves from the directions specified by their zero–order ray approximation. The polarization of a P wave may be close to transverse and elliptical, the polarization of an S wave, on the contrary, can be close to longitudinal. The first–order additional components can also be considered in anisotropic media. In weakly anisotropic media, they reduce to their "isotropic" versions, see Pšenčík (1998).

In homogeneous media, it is possible, with the use of symbolic manipulation software, to generate and evaluate formulae for higher–order terms theoretically without limitation. Vavryčuk (1999) used this approach to remove the failure of the zero–order ray approximation in the proper description of the problem of coupling of the two S waves propagating vertically in a weakly transversely isotropic medium with the axis of symmetry rotating in the horizontal plane. Unfortunately, in inhomogeneous media this procedure cannot be used, and we have to resort to the coupling ray theory, see Section 10.

12.2. Direct computation of first–arrival travel times

Recently, considerable attention has been devoted to the direct computation of travel times $\tau(x_i)$ on 2–D and 3–D grids without using ray tracing. These methods, however, do not yield the ray–theory travel times, but the first–arrival travel times. The first–arrival travel times are not related to the propagation of energy and should strictly be distinguished from the ray–theory travel times related to energetic arrivals. In the interpretational procedures of seismic exploration and seismology (tomography, migration), the energetic arrivals should mostly be considered, first–arrival travel times playing a smaller role. For this reason, we shall discuss the methods for the computation of first–arrival travel times only very briefly. Two basic methods in the computation of first–arrival travel times are network shortest–path ray tracing, and "finite–difference" travel time tracing. Refer to Figures 29–33 for simple numerical examples. The network shortest–path ray tracing is based on the the theory of graphs. The trajectory corresponding to the minimum time is usually called the shortest path, where "shortest" means the minimum travel time. For more details and other references see Moser (1991), Klimeš & Kvasnička (1994). The "finite–difference" methods do not usually consist in the direct finite–difference solutions of the eikonal equation (remember that the eikonal equation is nonlinear), but only in certain finite–difference approximations for the first–arrival travel–time continuation. Many such "finite–difference" approximations

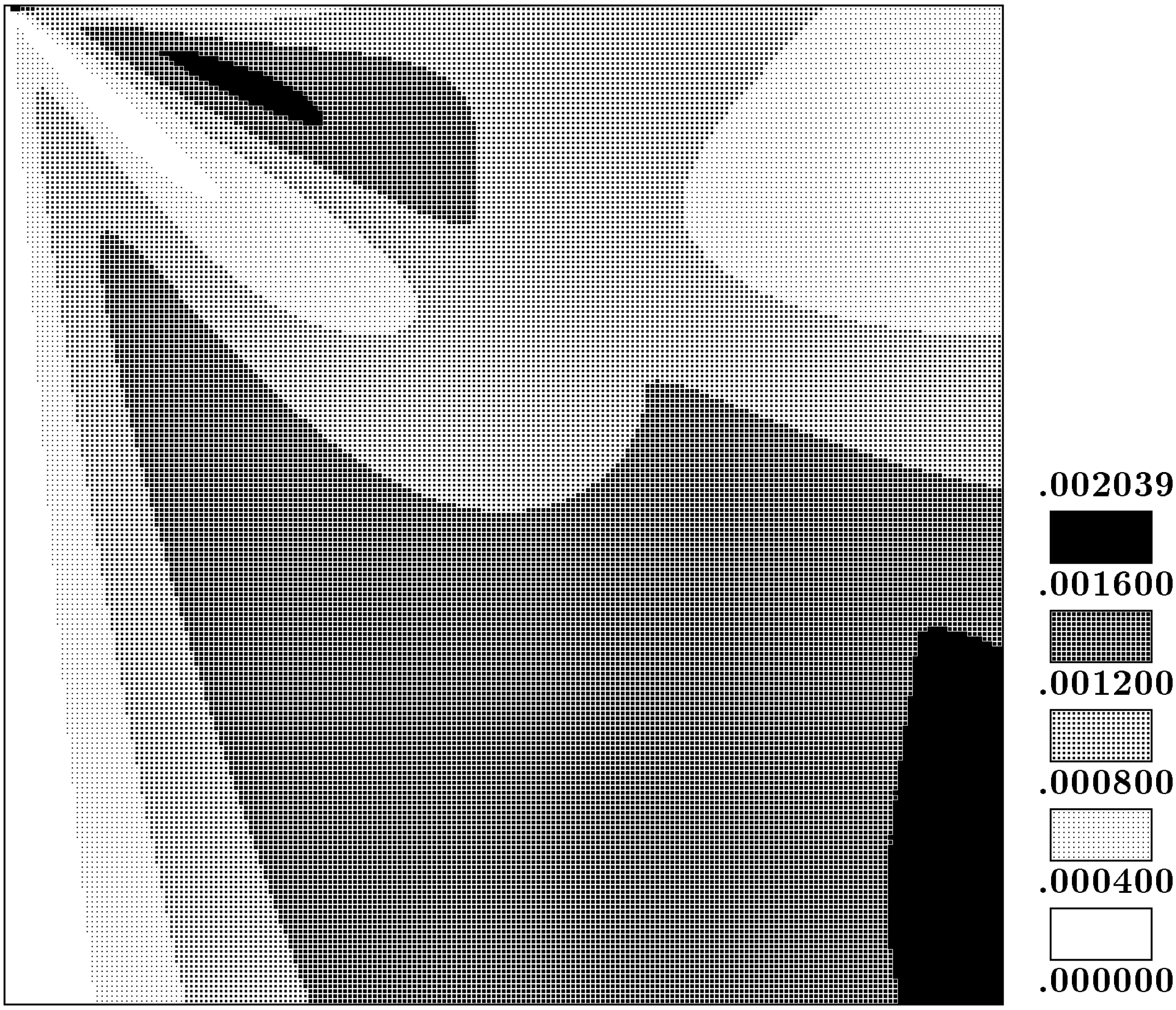

Figure 29. Constant vertical velocity gradient model, velocity contrast 1:3. Travel–time errors of the first–order grid travel–time tracing method proposed by Podvin & Lecomte (1991). The absolute errors of the first–arrival travel time are scaled in seconds.

have been proposed. See Vidale (1990), Klimeš (1996), and for many other references Červený (2001, p. 187). An extension of the "finite–difference" method, which also takes into account later arrivals, was proposed by Abgrall & Benamou (1999).

It should be emphasized that the ray–theory travel times, related to energy transport, are safely obtained by shooting methods with standard ray tracing, or by the wavefront construction method, see Section 4.

12.3. Ray method with complex eikonal

In the standard zero–order approximation of the ray method, travel time (eikonal) τ and slowness vector $\mathbf{p}$ are real–valued. The ray method can, however, be extended to complex–valued τ and $\mathbf{p}$. This generalization is then called *the ray method with complex eikonal* and the rays in it are complex–valued. Note that the elementary waves along complex rays with $\mathrm{Re}\mathbf{p}$ and $\mathrm{Im}\mathbf{p}$ parallel are usually called *homogeneous* and with $\mathrm{Re}\mathbf{p}$ and $\mathrm{Im}\mathbf{p}$ non–parallel are called *inhomogeneous*.

Homogeneous and inhomogeneous waves play an important role in many wave propagation problems including perfectly elastic media. As an example of inhomogeneous

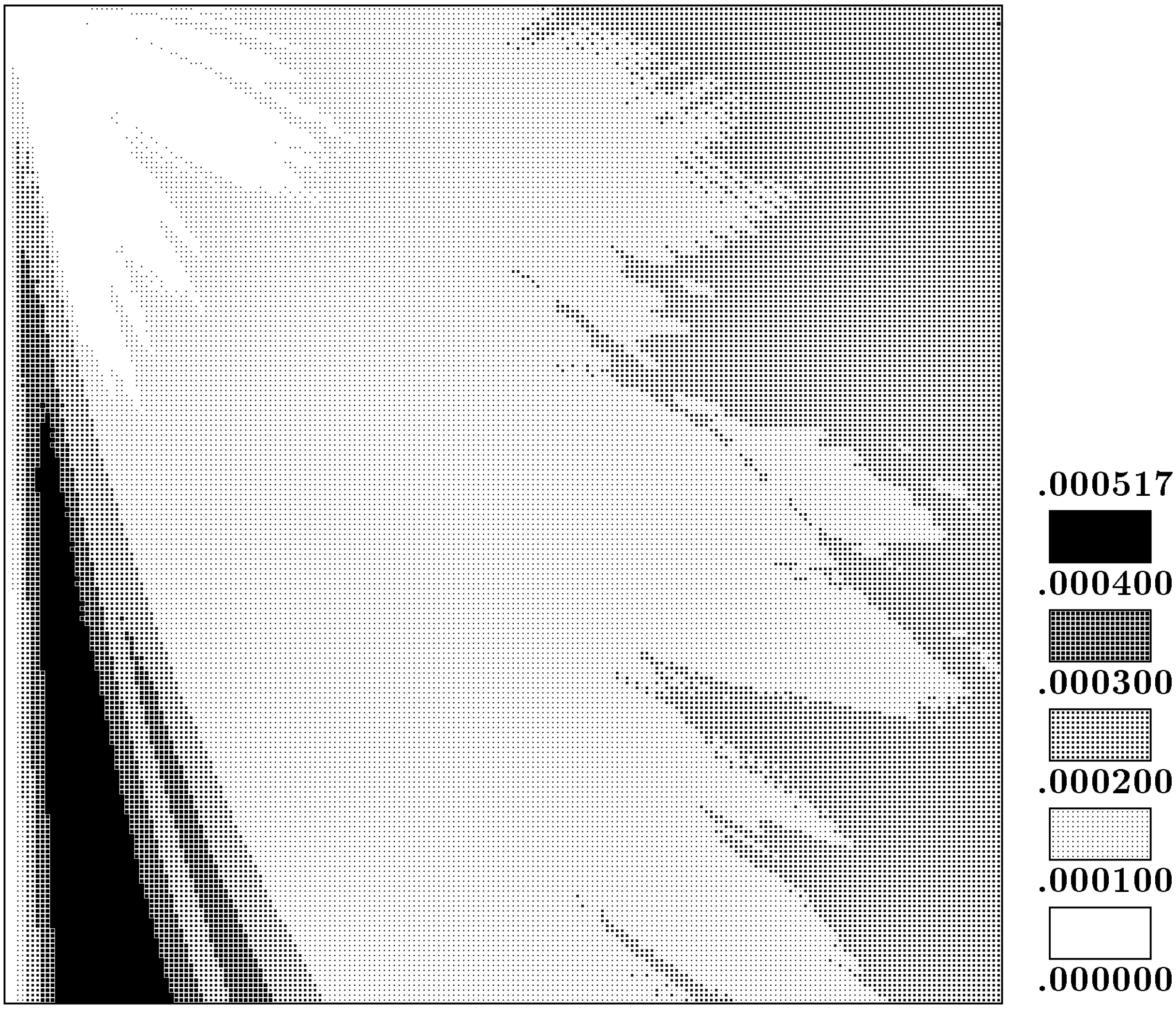

Figure 30. Constant vertical velocity gradient model, velocity contrast 1:3. Travel–time errors of the network shortest–path ray tracing according to Klimeš & Kvasnička (1994). The absolute errors of the travel time are scaled in seconds.

waves in perfectly elastic media let us mention the waves penetrating the caustic shadow, postcritically transmitted waves, and the inhomogeneous waves generated by a point source. The literature devoted to theoretical problems of inhomogeneous waves propagating in perfectly elastic media is broad, see for example recent publications by Chapman et al. (1999) or Kravtsov, Forbes & Asatryan (1999), where many additional references can be found. However, as far as we know, the calculation of complex rays has not yet been implemented into ray tracing program packages for perfectly elastic inhomogeneous media. Only the amplitudes of the complex rays at the point of their generation at an interface are considered when postcritically reflected waves are calculated.

In *dissipative media*, the eikonal τ is intrinsically complex–valued, and the absorption is always connected with velocity dispersion (causal absorption). Consequently, the rays are also complex–valued, see Thomson (1997). In the seismic ray theory the study of complex rays in dissipative media has not been completed yet. In most cases, it is assumed that dissipation is weak and it is taken into account only as a first–order perturbation of a perfectly elastic medium, see Figure 34. The corresponding waves

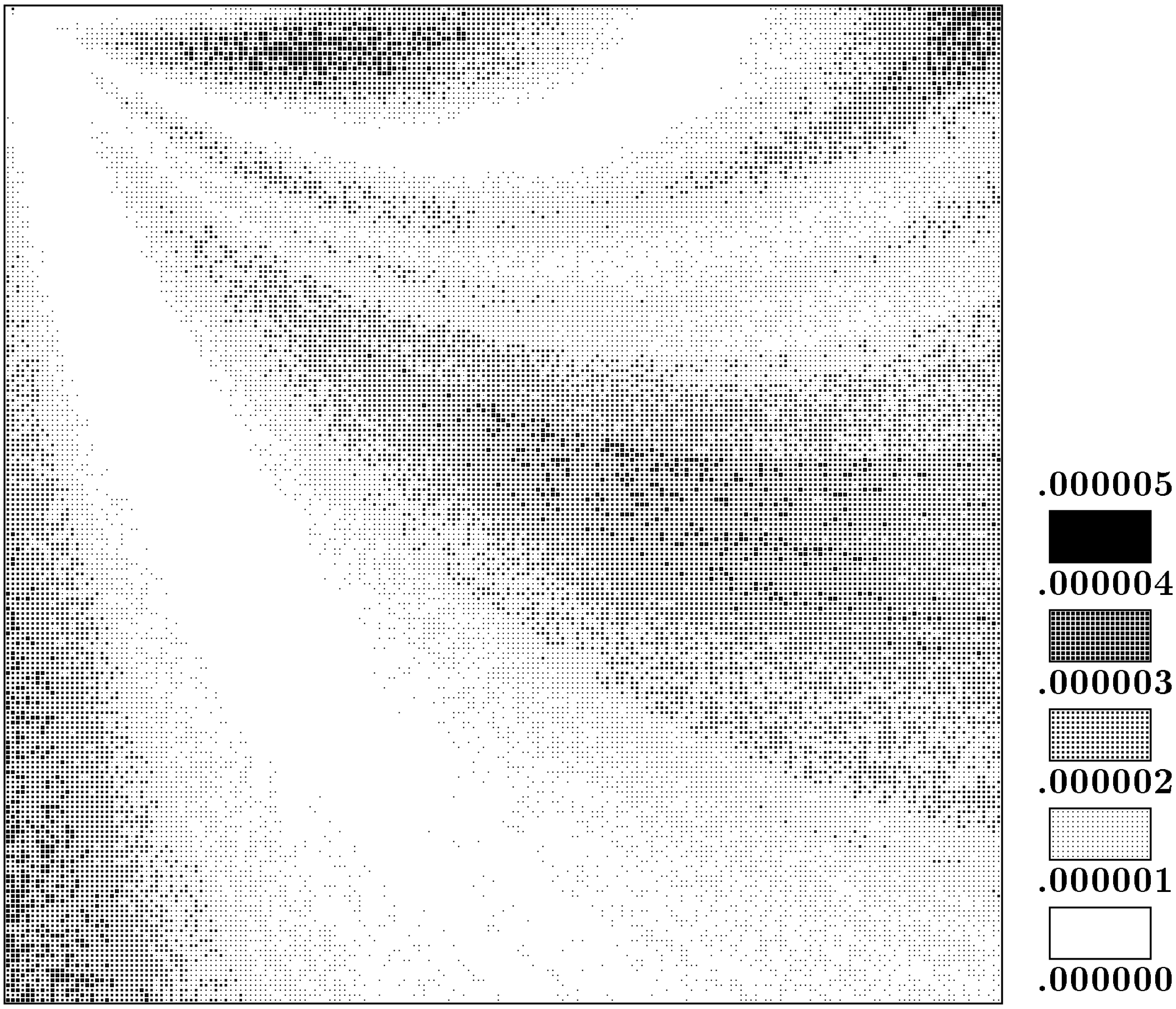

Figure 31. Constant vertical velocity gradient model, velocity contrast 1:3. Travel–time errors of the second–order grid travel–time tracing according to Klimeš (1996). The absolute errors of the travel time are scaled in seconds.

then have the character of homogeneous waves. Refer, for example, to Moczo, Bard & Pšenčík (1987) for isotropic dissipative media, and to Gajewski & Pšenčík (1992) for anisotropic dissipative media. Such a simple approximation is at present implemented in most ray–based program packages. As the dissipation within the Earth's interior is usually weak, the first–order perturbation method yields sufficiently accurate amplitude decay due to absorption and frequency–dependent time shift due to the relevant dispersion.

12.4. Hybrid methods

In hybrid methods, the ray method is locally combined with some other method. There are several such existing combinations. We shall now discuss briefly three of them: (a) the hybrid ray–finite–difference method; (b) the hybrid ray–mode method; (c) the hybrid ray–matrix method (sometimes called the hybrid ray–reflectivity method).

In the hybrid ray–finite–difference method, the wavefield is computed in a globally smooth model, with a more complicated structure localized in a small part of the model. The ray method is used in the smooth part of the model, finite differences in the small

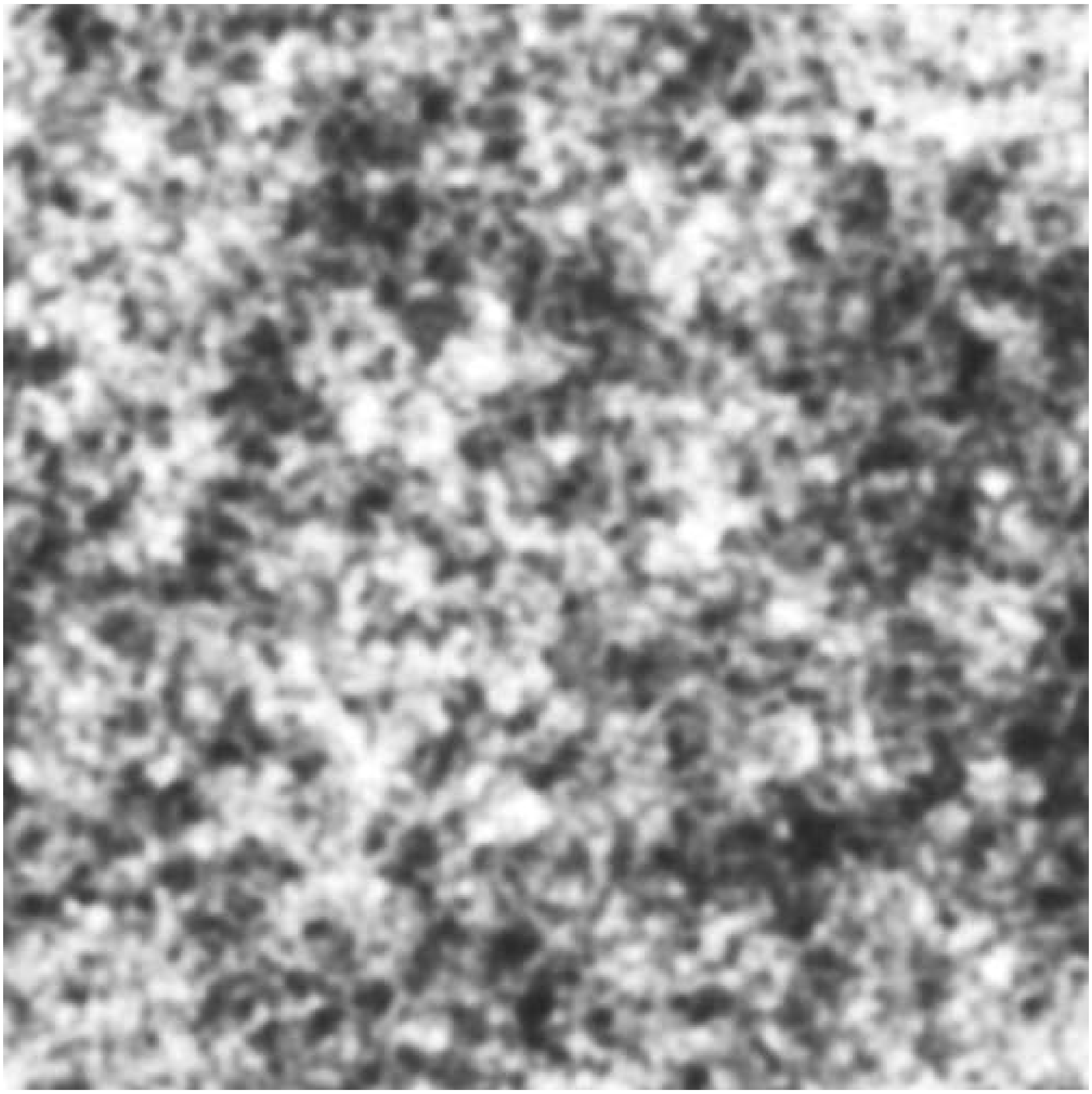

Figure 32. A 2–D realization of a random velocity distribution.

complicated region. The basic problem is to match the solutions of both methods at the boundary of the complicated region. See, e.g., Lecomte (1996), Gjøystdal et al. (2002), Opršal et al. (2002).

In the hybrid ray–mode method, the modal and ray approaches are combined. The aim is to simulate the complete wavefield by several ray contributions and by several modes. See, e.g., Kapoor & Felsen (1995), or Zhao & Dahlen (1996).

The hybrid ray–matrix method has been used in smooth models containing transition layers thin with respect to the prevailing wavelength. The ray method is again used in smooth parts of the model, the matrix method is used to compute R/T coefficients of thin layers, see Figure 35. Červený (1989) applied this method to the Earth's crust model with the MOHO discontinuity approximated by a thin layer with varying internal structure (transition layer, laminated layer).

The method was also applied to the study of tunnelling of seismic waves through

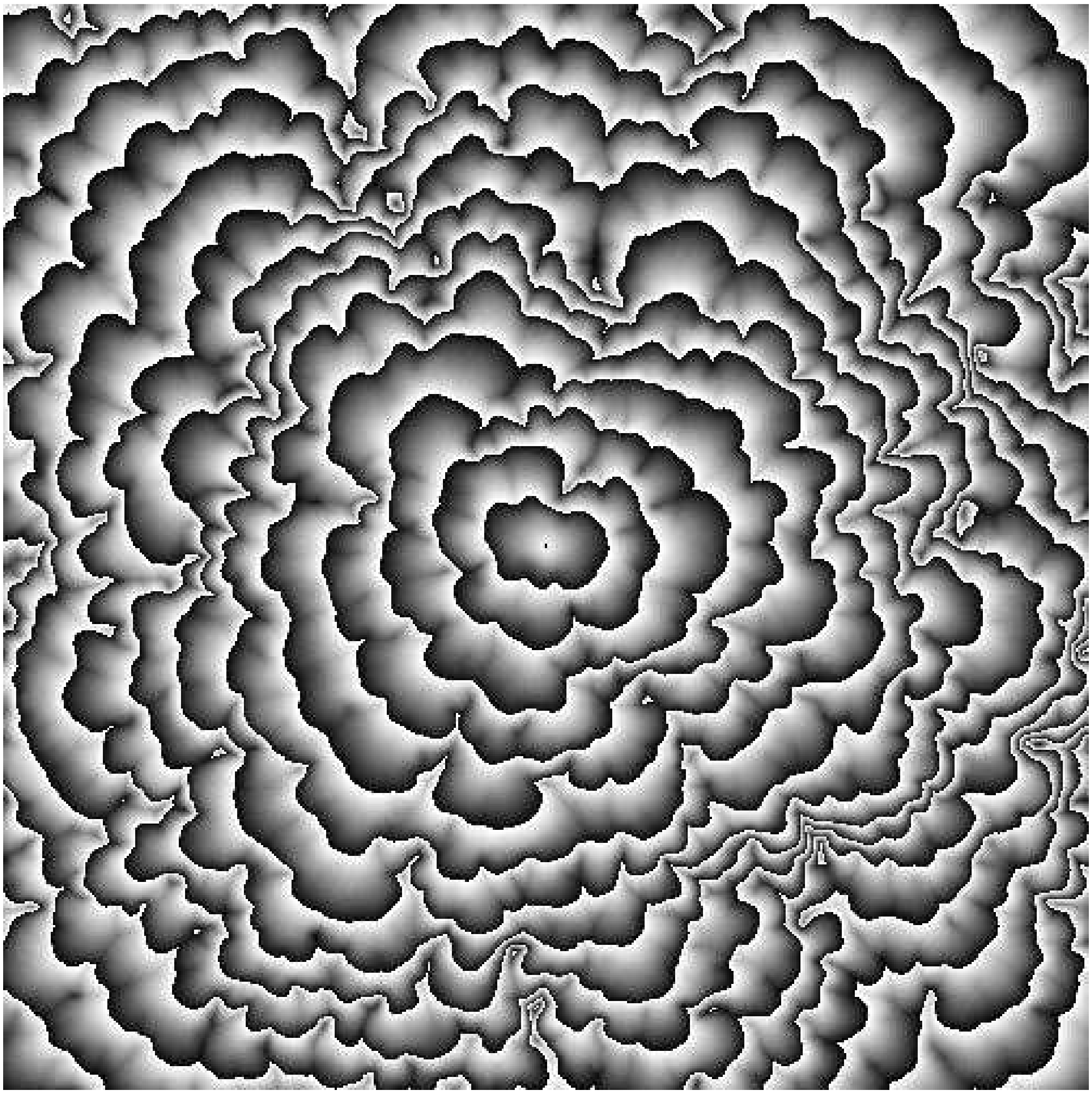

Figure 33. First–arrival travel times calculated in the 2–D model shown in Figure 32 using grid travel–time tracing according to Klimeš (1996).

a thin high–velocity layer (Červený & Aranha, 1992), and to the study of the effects of a near–surface sedimentary layer on the wavefield measured on the Earth's surface (Červený & Andrade, 1992). The method is also a part of the CRT program package.

12.5. Several other extensions of the ray method

There are many other useful extensions of the ray method. We briefly mention several of them and supplement each of them by at least one reference, where more details and additional references can be found.

(a) Space–time ray method (Babich, Buldyrev & Molotkov, 1985); (b) Surface–wave ray tracing (Martin & Thomson, 1997; Thomson, 1997; Dahlen & Tromp, 1998; Červený, 2001, sec. 3.12); (c) Asymptotic diffraction theory including geometric theory of diffraction (Hanyga, 1995; Hanyga et al. 2001); (d) Ray methods in directional

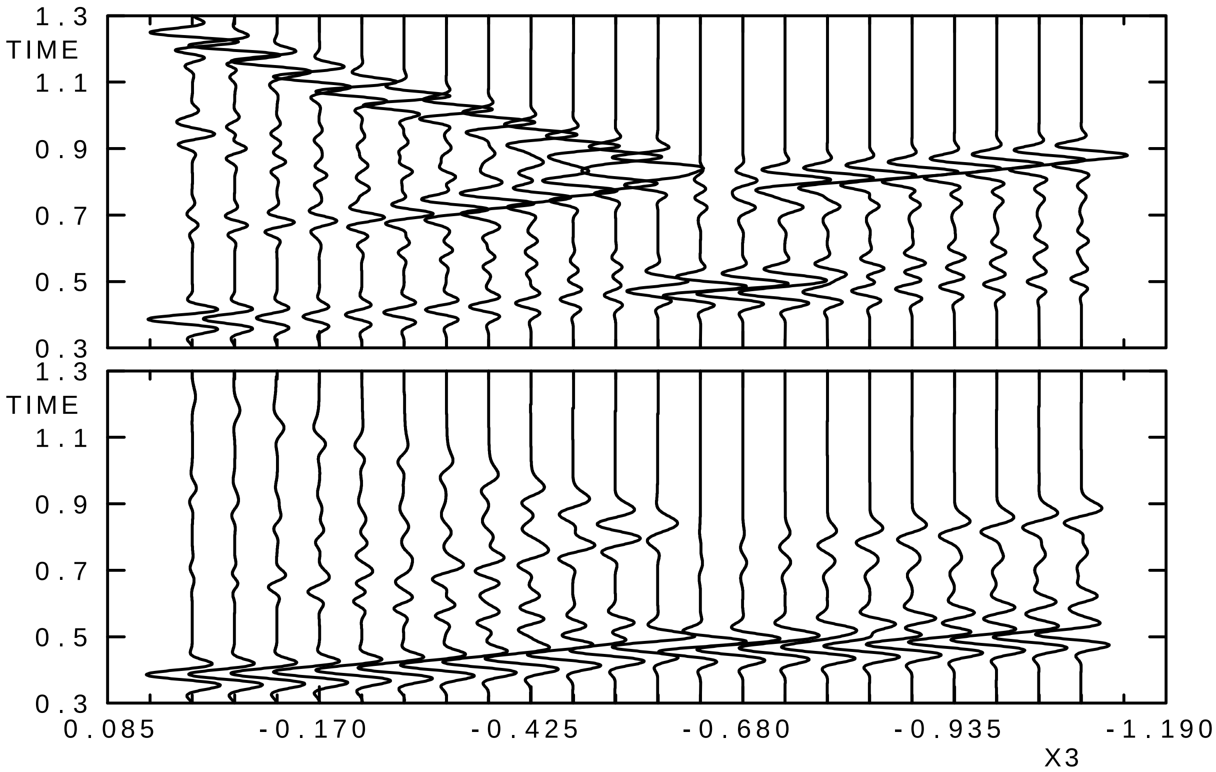

Figure 34. Trace–normalized ray–theory seismograms without attenuation (top) and with causal attenuation (bottom). Complex–valued travel times are approximated using the ray perturbation method.

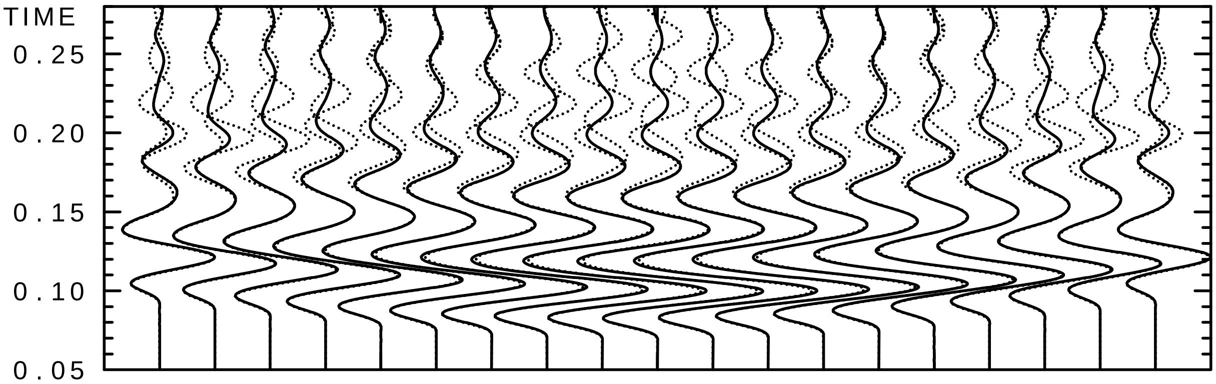

Figure 35. Vertical component of synthetic seismograms at the top of a sedimentary layer. **Solid line:** elastic seismograms calculated by the ray–matrix method are quite accurate in this case. The standard ray method would yield only the onset of the wave–train. **Two dotted lines:** finite–difference elastic seismograms are deteriorated by reflections due to imperfect non–reflecting boundary conditions.

(one way) propagation (Thomson, 1999); (e) Ray method in Born scattering of seismic waves (Ursin & Tygel, 1997; Chapman, 2004); (f) Kirchhoff–Helmholtz integrals (Tygel, Schleicher & Hubral, 1994, Schleicher et al., 2001; Chapman, 2004).

13. Ray chaos, Lyapunov exponents, models suitable for ray tracing

In complex models, the behaviour of rays becomes chaotic and geometrical spreading, the number of arrivals and the density of caustic surfaces exponentially increase with travel time (Keers, Dahlen & Nolet, 1997). The exponential increment may be quantitatively described in terms of the Lyapunov exponents. The average twisting of ray tubes in phase space, which determines the average frequency of caustic points along rays, may be quantitatively described in terms of the rotation numbers. Note that the average Lyapunov exponent and the average rotation number are two different characteristics of ray chaos.

Lyapunov exponents enable to determine quantitative criteria on models, suitable for ray tracing, in terms of Sobolev scalar products. These criteria enable to construct the optimum models of geological structures for ray tracing by inversion of seismic data.

13.1. Lyapunov exponents and rotation numbers

If heterogeneities in a model (velocity model, macro model) exceed a certain degree, average geometrical spreading exponentially increases with length of the rays and, in consequence, the average number of travel times (i.e. the average number of rays intersecting at the same point) exponentially increases with distance from the source. This behaviour of rays strictly limits the possibility of calculating two–point rays and travel times in overly complex models because:

(a) Geometrical spreading may be so large that two–point rays cannot be found within the numerical accuracy. Similarly, the ray tubes cannot be sufficiently narrow for travel–time interpolation.

(b) The number of two–point travel times at each point is so large that all travel times cannot be calculated within reasonable computational time and costs.

(c) The number of two–point travel times at each point is so large that they can hardly be useful for any application, independently of the applicability of the ray theory, which is not considered here.

It is thus of principal interest to quantify the exponential divergence of neighbouring rays with respect to the complexity of the model, and to formulate explicit criteria enabling models suitable for ray tracing to be constructed.

The exponential divergence of rays is quantified by the Lyapunov exponents (Lyapunov, 1949; Oseledec, 1968). The average frequency of caustic points along rays is quantified by the rotation numbers (Johnson, 1986). The Lyapunov exponents and the rotation numbers are two different characteristics of ray chaos.

Wolfson & Tappert (2000, eq. 36) estimated the Lyapunov exponent for 2–D ray tracing in random models with infinitely weak heterogeneities. Their method, designed for underwater acoustics, is thus not applicable to strong heterogeneities present in the solid Earth. Here we present the method for the quantitative estimation of the Lyapunov exponent, mean Lyapunov exponent, rotation number and mean rotation number for rays in 2–D models without interfaces according to Klimeš (2002b). The presented method is designed especially for deterministic models (velocity models, macro models) with strong heterogeneities. The method is thus not applicable to weak heterogeneities. It should be further generalized to 3–D models and to models with interfaces.

We define the Lyapunov exponents for general ray tracing, and the rotation number for 2–D ray tracing. We then introduce and estimate the *average Lyapunov exponent*, describing the average spreading of ray tubes and average number of travel times, in smooth 2–D models without interfaces. The equations allow the average exponential divergence of rays and exponential growth of the number of travel–time branches in the model to be estimated prior to ray tracing. The test of the equations for estimating the average Lyapunov exponents in a given smooth 2–D model without interfaces is illustrated on a numerical example.

13.1.1. Lyapunov exponents for ray tracing

Lyapunov exponents may be defined in several ways (Lyapunov, 1949; Oseledec, 1968; Katok, 1980). Some definitions rely on an unspecified norm in phase space, which may be chosen arbitrarily. Although the phase–space norm does not affect the values of the Lyapunov exponents defined asymptotically for infinitely long rays, it may considerably affect the estimated values of the Lyapunov exponents along finite rays in finite models.

The estimates of the Lyapunov exponents based on the *characteristic values* of the DRT propagator matrix are not affected by the free parameters corresponding to the norm in phase space. On the other hand, the characteristic values oscillate along rays which makes the estimation of the Lyapunov exponents difficult. The oscillations of the characteristic values are caused by the rotation of ray tubes in phase space.

Let us denote by μ_1, μ_2, ..., μ_{2N} the *characteristic values* of the $2N \times 2N$ (in N–dimensional space) DRT propagator matrix $\mathbf{\Pi}$ in Cartesian coordinates, i.e. the solutions of the characteristic equation

$$\det[\mathbf{\Pi}(\tau,\tau_0) - \mu(\tau,\tau_0)\mathbf{I}] = 0 \quad , \tag{182}$$

sorted according to their absolute values,

$$|\mu_1| \geq |\mu_2| \geq ... \geq |\mu_{2N}| \quad . \tag{183}$$

The complex–valued characteristic values, with the same absolute value, are assumed to be sorted according to their argument $-\pi < \arg\mu \leq \pi$.

The *positive Lyapunov exponents* along a ray parametrized by travel time τ may be defined as

$$\lambda_k = \limsup_{\tau\to+\infty} \frac{\ln|\mu_k(\tau,\tau_0)|}{\tau-\tau_0} \quad , \qquad k = 1,2,...,N-1 \quad . \tag{184}$$

The Lyapunov exponents are thus defined here with respect to travel time. Note that the Lyapunov exponents may also be defined with respect to another monotonic parameter along the ray, which may or may not differ from the parameter determined by the form of the Hamiltonian.

Because of the symplectic property of the DRT propagator matrix, its inverse $\mathbf{\Pi}^{-1}$ has the same set of characteristic values as $\mathbf{\Pi}$. That is why the characteristic values of all Hamiltonian systems form reciprocal pairs $\mu_1\mu_{2N} = 1$, $\mu_2\mu_{2N-1} = 1$, ... $\mu_N\mu_{N+1} = 1$. Each positive Lyapunov exponent of ray tracing (as of other Hamiltonian systems) is thus accompanied by a negative Lyapunov exponent of the same absolute value. It is thus sufficient to study the positive Lyapunov exponents for ray tracing.

Two Lyapunov exponents λ_N and λ_{N+1} of ray tracing are always zero. The remaining $N-1$ pairs of positive and negative Lyapunov exponents correspond to the

$(2N-2)\times(2N-2)$ (in N–dimensional space) DRT propagator matrix $\mathbf{\Pi}^{(q)}$ in ray–centred coordinates. It is thus often convenient to replace the DRT propagator matrix in Cartesian coordinates in equation (182) by the DRT propagator matrix in ray–centred coordinates.

13.1.2. Rotation number for 2–D ray tracing

In isotropic 2–D models, the dynamic ray tracing equation (41) in ray–centred coordinates simplifies to equation

$$\frac{\mathrm{d}\mathbf{\Pi}^{(q)}}{\mathrm{d}\tau} = \begin{pmatrix} 0 & B^{(q)} \\ -C^{(q)} & 0 \end{pmatrix} \mathbf{\Pi}^{(q)} \quad , \tag{185}$$

where the second phase–space derivatives of the Hamiltonian corresponding to the in–plane ray–centred coordinate q are

$$C^{(q)} = V^{-1}\frac{\partial^2 V}{\partial q \partial q} \tag{186}$$

and

$$B^{(q)} = V^2 \quad . \tag{187}$$

The solutions of dynamic ray tracing equation (185) rotate in phase space. Denote the angle of rotation of solution of equation (185) between travel times τ_0 and τ by $\vartheta(\tau,\tau_0)$. The exact value of angle $\vartheta(\tau,\tau_0)$ depends on the phase–space metric and is not unique. Fortunately, each increment of $\vartheta(\tau,\tau_0)$ by an integer multiple of π can uniquely be defined, at least in isotropic media. The *rotation number* φ is then defined as

$$\varphi = \lim_{\tau\to+\infty} \frac{\vartheta(\tau,\tau_0)}{\tau-\tau_0} \quad . \tag{188}$$

The number of travel times and the number of caustics are not related in a simple way, because the number of travel times remains unchanged when the ray touches a caustic, whereas the number of travel times may be either increased by 2 or decreased by 2 when the ray crosses a caustic. The average Lyapunov exponent, characterizing the exponential increment of the number of travel times in dependence on the length of the rays, and the rotation number, characterizing the average frequency of caustic points along rays, are thus two different characteristics of ray chaos. The caustic point is understood to be the point where the ray touches the caustic.

13.1.3. Approximation of the positive Lyapunov exponent in 2–D models without interfaces

We introduce parameters

$$\Lambda = \int \sqrt{\mathrm{neg}(C^{(q)}B^{(q)})}\,\mathrm{d}\tau = \int\sqrt{\mathrm{neg}\left(V\frac{\partial^2 V}{\partial q\partial q}\right)}\,\mathrm{d}\tau \tag{189}$$

and

$$\Phi = \int \sqrt{\mathrm{pos}(C^{(q)}B^{(q)})}\,\mathrm{d}\tau = \int\sqrt{\mathrm{pos}\left(V\frac{\partial^2 V}{\partial q\partial q}\right)}\,\mathrm{d}\tau \tag{190}$$

calculated along the ray. Here

$$\mathrm{pos}(x) = \max(x,0) \quad , \qquad \mathrm{neg}(x) = -\min(x,0) \tag{191}$$

denote the positive and negative parts of real number x, respectively. Parameter Λ accounts for the exponential spreading of the ray tube in phase space, and parameter Φ describes how the ray tube is twisted in phase space. Parameter Λ increases in the *defocusing zones*, where the second velocity derivative perpendicularly to the ray is negative. Parameter Φ increases in the *focusing zones*, where the second velocity derivative perpendicularly to the ray is positive. Integral (189) is thus the sum of integrals $\Delta\Lambda_n$ over individual ray segments corresponding to defocusing zones, and integral (190) is the sum of integrals $\Delta\Phi_n$ over the individual ray segments corresponding to focusing zones.

We approximate the positive Lyapunov exponent (184) by the limit

$$\lambda_1 \approx \lim_{\tau\to+\infty} \lambda(\tau,\tau_0) \tag{192}$$

of function

$$\lambda(\tau,\tau_0) = \frac{L(\tau,\tau_0)}{\tau-\tau_0} \quad . \tag{193}$$

The estimate of function L in equation (193) is (Klimeš, 2002b)

$$\begin{aligned} L(\tau,\tau_0) = \sum_{n=1}^{N} \Delta\Lambda_n + \sum_{n=1}^{N-1} \ln \left|\cos\!\left[\min\!\left(\Delta\Phi_n, \frac{\pi}{3}\right)\right]\right| \\ + \frac{1}{2}\ln\left\{1-\sin\!\left[2\min\!\left(\Delta\Phi_0,\frac{\pi}{6}\right)\right]\sin\!\left[2\min\!\left(\Delta\Phi_N,\frac{\pi}{6}\right)\right]\right\} \ , \end{aligned} \tag{194}$$

where $\Delta\Lambda_n$, $n=1,2,...,N$, correspond to the defocusing ray segments ($C^{(q)}B^{(q)}<0$) and $\Delta\Phi_n$, $n=0,1,2,...,N$, correspond to the focusing ray segments ($C^{(q)}B^{(q)}>0$), $\Delta\Phi_0$ and $\Delta\Phi_N$ may become zero if the ray starts or terminates in a defocusing region. Terms $\Delta\Lambda_n$ account for exponential deformation of phase space due to defocusing. After phase–space rotation of ray tubes in a focusing region, the exponential deformation does not continue in the previous deformation, and the deformations may partly cancel out. This is taken into account by the terms with $\Delta\Phi_n$ in (194). These terms are deterministic for small rotations $\Delta\Phi_n$ (the first argument of the minimum functions in (194)), and statistically averaged for greater rotations $\Delta\Phi_n$ (the second argument of the minimum functions). A rough upper estimate $L_{\max}(\tau,\tau_0)=\sum_{n=1}^{N}\Delta\Lambda_n$ of (194) may be used especially for very strong and large heterogeneities $\Lambda_n \gg \frac{\pi}{3}$.

Functions (193) and (194) are defined and have reasonable values also for rays of finite lengths. Quantity $\lambda(\tau,\tau_0)$ is thus our estimate of the Lyapunov exponent for a finite ray.

13.1.4. Approximation of the rotation number in 2–D models without interfaces

We analogously approximate rotation number (188) by the limit

$$\varphi \approx \lim_{\tau\to+\infty} \varphi(\tau,\tau_0) \tag{195}$$

of function

$$\varphi(\tau,\tau_0) = \frac{\Phi(\tau,\tau_0)}{\tau-\tau_0} \quad , \tag{196}$$

where

$$\Phi(\tau,\tau_0) = \sum_{n=0}^{N} \Delta\Phi_n \quad . \tag{197}$$

This estimate is in accordance with the example of White, Nair & Bayliss (1988) but has not been further numerically tested yet.

13.1.5. Lyapunov exponent and rotation number for a system of finite rays

Equation (194) allows the average Lyapunov exponent

$$\overline{\lambda} = \frac{\sum\limits_{\mathrm{ray}} L_{\mathrm{ray}}(\tau_{\mathrm{ray}}, \tau_{\mathrm{ray}\,0})}{\sum\limits_{\mathrm{ray}} [\tau_{\mathrm{ray}} - \tau_{\mathrm{ray}\,0}]} \tag{198}$$

over all rays to be introduced. This average Lyapunov exponent expresses the global properties of the model. However, for a single source, it is still dependent on the source geometry and position with respect to the model boundaries. The average Lyapunov exponent over various sources should describe the global properties of the model. The average Lyapunov exponent over various sources may depend on the geometry of model boundaries, but only if the statistical properties of the model are anisotropic.

Analogously, the average rotation number $\overline{\varphi}$ for the system of finite rays may be approximated by

$$\overline{\varphi} \approx \frac{\sum\limits_{\mathrm{ray}} \Phi_{\mathrm{ray}}(\tau_{\mathrm{ray}}, \tau_{\mathrm{ray}\,0})}{\sum\limits_{\mathrm{ray}} [\tau_{\mathrm{ray}} - \tau_{\mathrm{ray}\,0}]} \quad , \tag{199}$$

where parameter Φ along the individual rays is defined by (197).

The average distance between two consecutive caustic points along a ray, measured in terms of travel time τ, is

$$\overline{\tau}_{\mathrm{caustic}} = \frac{\pi}{\overline{\varphi}} \quad . \tag{200}$$

At this distance, the average number of travel times increases by the factor

$$\overline{N}_{\mathrm{caustic}} = \exp(\frac{\overline{\lambda}}{\overline{\varphi}}\pi) \quad . \tag{201}$$

We can conjecture from equations (194), (198) and (199) that

$$0 \leq \overline{\lambda} < \overline{\varphi} \quad , \tag{202}$$

depending on the kind of heterogeneities in the model. In consequence, the average increment $\overline{N}_{\mathrm{caustic}}$ of travel times between two consecutive caustic points depends on the kind of heterogeneities in the model, varying within interval

$$1 \leq \overline{N}_{\mathrm{caustic}} < \exp\pi \simeq 23.141 \quad . \tag{203}$$

13.1.6. Average Lyapunov exponent for the model and average rotation number for the model

Our estimate of the Lyapunov exponent of a single finite ray depends on the position and direction of the ray. Let us now average the Lyapunov exponent over the whole model volume.

We cover the model by a dense system of parallel straight lines. For each direction of the lines, we calculate the average Lyapunov exponent (198) using equation (194), similarly as for a system of rays, and call it the *directional Lyapunov exponent*. We then average the calculated directional Lyapunov exponent over the directions, applying a selected directional weighting function. For instance, the directional weighting function may represent the shape of a model box, with the origin at the centre of the box or at the mean position of the intended point sources. The directional weighting function may also correspond to the probability of the ray directions estimated for ray tracing in the model. For example, if we are going to trace nearly vertical rays only and wish to estimate the average Lyapunov exponent for these rays, the directional weighting function should reflect the probability of the ray directions.

13.1.7. Numerical example

The 2–D model designed by Jean–David Benamou is formed by the homogeneous background of velocity $1.0\,\mathrm{km\,s^{-1}}$, perturbed by a stretched bi–sine egg–box of amplitude $0.2\,\mathrm{km\,s^{-1}}$, see Figure 36. The horizontal dimension of the model box is 3 km, vertical 6 km. There are focusing low–velocity regions close to the four corners. The average travel times in seconds closely correspond to the average lengths in kilometres in this model.

Some rays, shot from the point source situated at the bottom of the model box, 1.55 km from the left–hand corner, are shown in Figure 37. Figure 38 shows the rays shot from the point source shifted to 1.85 km from the left–hand corner. Although the model is highly regular and periodic, the ray paths are quite irregular. If we divide the model into the smallest equal cells, rays enter and leave the individual cells at and in quite different positions and directions, 0% of rays being periodic. The rays traced in this model can thus serve as an example of the chaotic behaviour of rays in general heterogeneous 2–D models, keeping in mind that the statistical properties of this model are strongly anisotropic.

Figure 39 displays the numbers of travel times corresponding to the point source situated at the bottom of the model box, 1.55 km from the left–hand corner. The travel times are calculated in a grid of 121×241 points covering the model box by means of interpolation within the ray cells (Bulant, 1999; Bulant & Klimeš, 1999). The maximum number of travel times found is 49. Figure 40 shows the numbers of travel times for the point source shifted to 1.85 km from the left–hand corner. The maximum number of travel times found is 59.

Ninety directions with an angular increment of 2 degrees have been chosen to estimate the average Lyapunov exponent for the model. For each direction, the model has been covered by 45 equally spaced straight lines. The directional Lyapunov exponent according to equations (198) and (194) has been numerically calculated along the straight lines, with a step corresponding to 45 steps along the longest line for the direction. Since the statistical properties of the model are strongly anisotropic, the

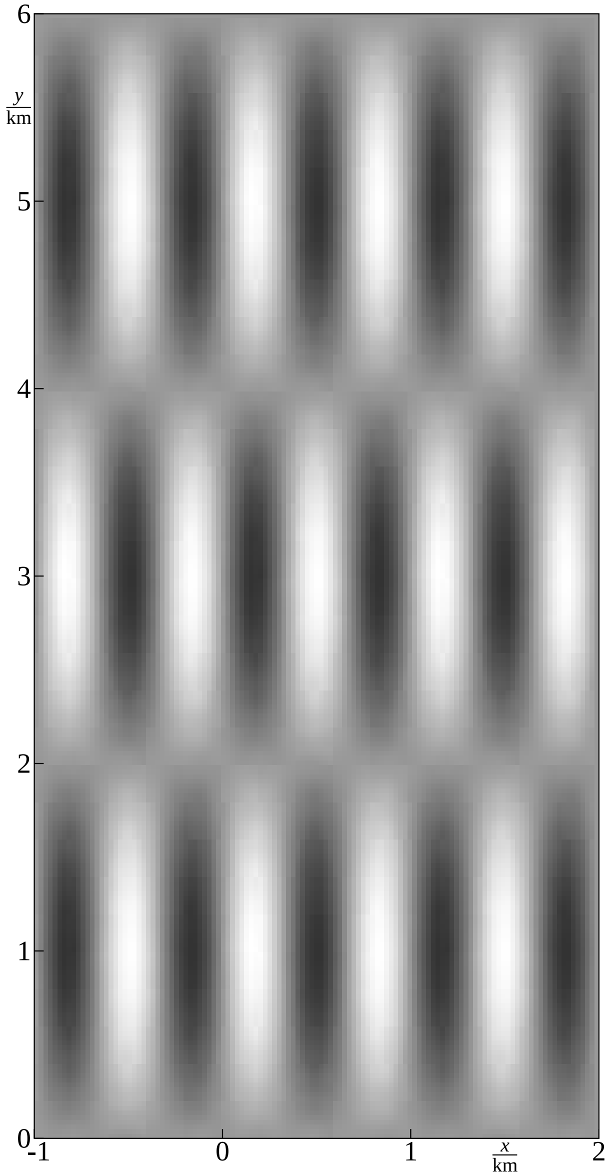

Figure 36. The model is formed by a homogeneous background of velocity $1.0\,\mathrm{km\,s^{-1}}$, perturbed by a stretched bi–sine egg–box of amplitude $0.2\,\mathrm{km\,s^{-1}}$. The dark zones situated close to the four corners are focusing low–velocity regions.

directional Lyapunov exponents vary between $0.170\,\mathrm{s}^{-1}$ and $1.019\,\mathrm{s}^{-1}$. The selected directional weighting function corresponds to the model box with the origin at the centre of the bottom edge. This directional weighting function is suitable for the point sources situated at the bottom or at the top of the model box. The average Lyapunov exponent for the model, calculated with this directional weighting function, is

$$\overline{\lambda}_{\mathrm{model}} = 0.698\,\mathrm{s}^{-1} \quad . \tag{204}$$

The average Lyapunov exponent does not noticeably vary with the horizontal translation of the origin of the directional weighting function within the middle third of the horizontal model dimension.

Figure 41 displays the natural logarithms of the average and maximum numbers of travel times along the individual horizontal grid lines of Figures 39 and 40. The

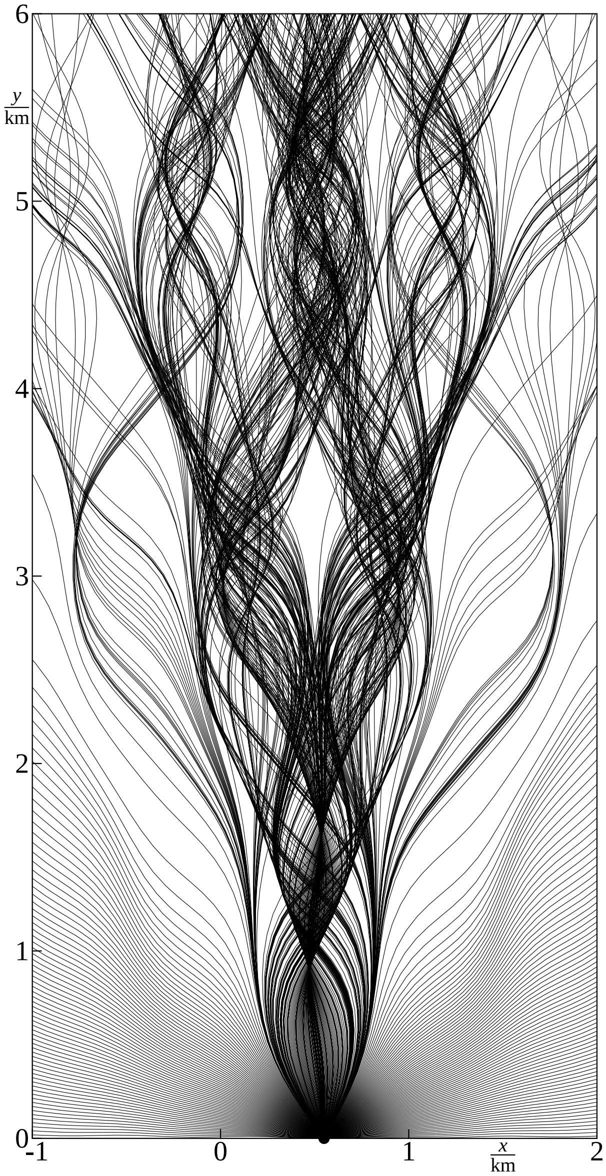

Figure 37. Rays in the model of Figure 36, corresponding to the point source situated at the bottom of the model box, 1.55 km from the left–hand corner.

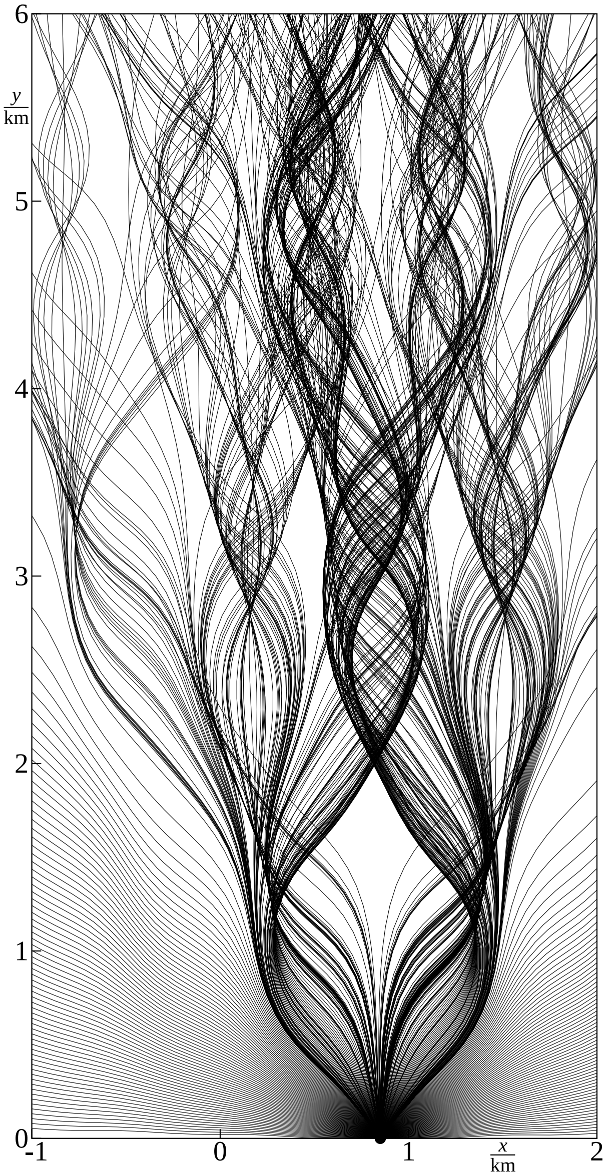

Figure 38. Rays in the model of Figure 36, corresponding to the point source situated at the bottom of the model box, 1.85 km from the left–hand corner.

horizontal axis is the distance of the grid line from the bottom of the model box in kilometres, and serves as a rough approximation of the travel time in seconds. The slope of the straight solid lines is given by the average Lyapunov exponent (204) for the model.

The numerical example demonstrates the good correspondence between the average logarithms of the numbers of ray–theory travel times and the estimate of the average Lyapunov exponent for the model.

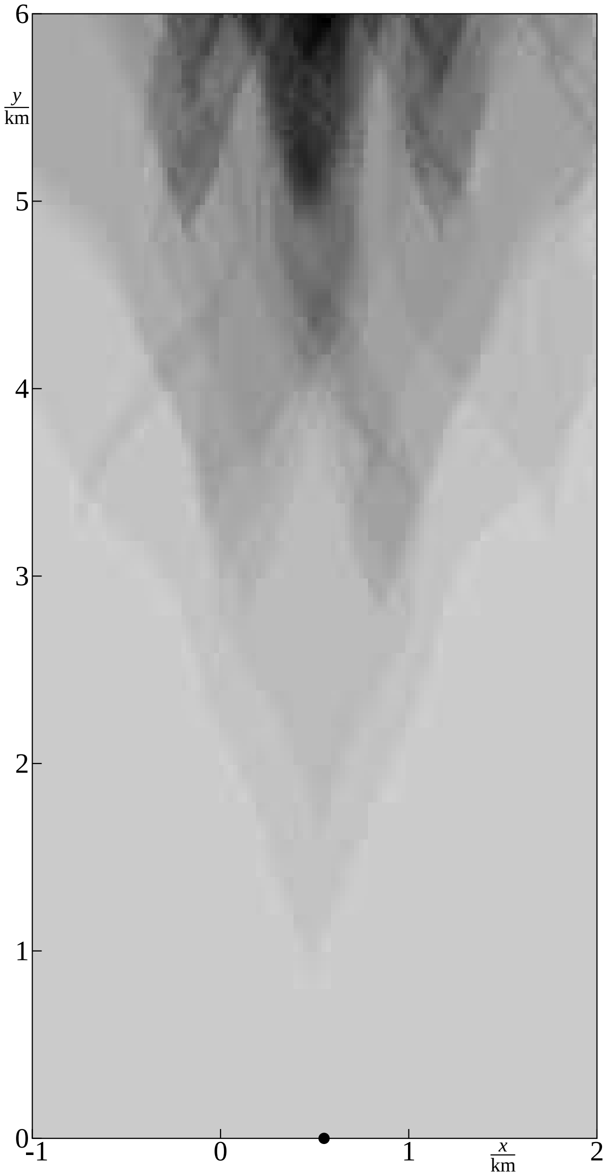

Figure 39. The numbers of travel times corresponding to the point source situated at the bottom of the model box, 1.55 km from the left–hand corner. The maximum number of travel times found is 49.

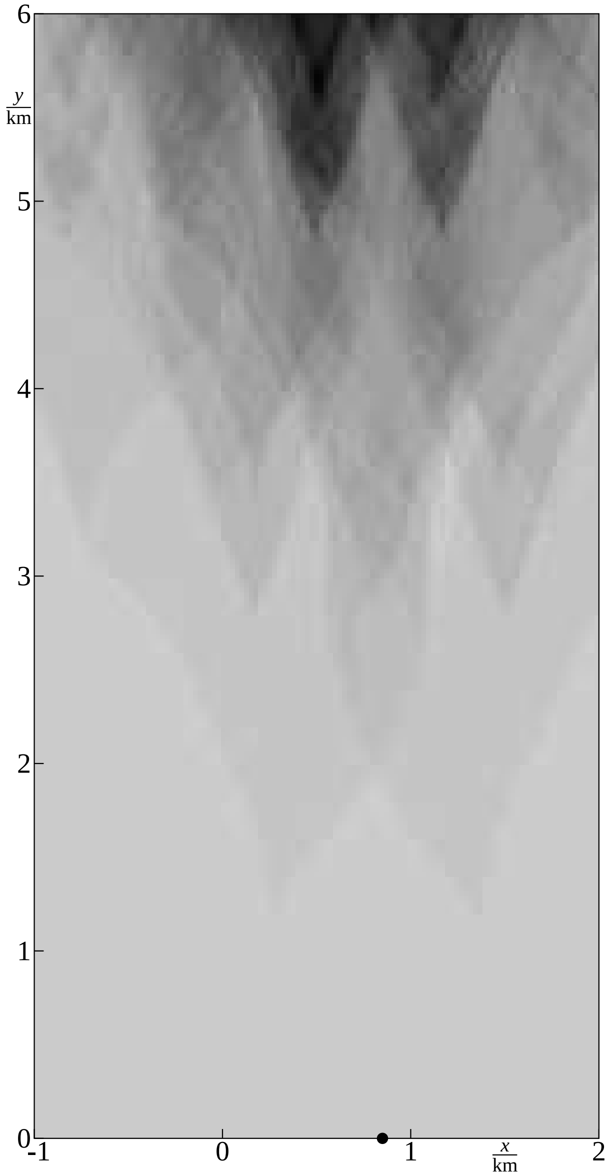

Figure 40. The numbers of travel times corresponding to the point source situated at the bottom of the model box, 1.85 km from the left–hand corner. The maximum number of travel times found is 59.

13.2. Models suitable for ray tracing

Since the average Lyapunov exponent for the model may be approximated in terms of the Sobolev norm composed of the second velocity derivatives (Klimeš, 2000), the average Lyapunov exponent and the Sobolev norm may be used in the construction of models optimized for the calculation of ray–theory Green functions (Bulant, 2002; Žáček, 2002).

The Sobolev scalar product of two functions is a linear combination of the L_2 Lebesgue scalar products of the zero, first, second or higher derivatives of the functions. There are two main reasons for including the Sobolev norm in the objective function during the model smoothing (Žáček, 2002) or during the inversion of seismic data (Klimeš, 2002c). Firstly, the Sobolev norm regularizes ill–conditioned inversions,

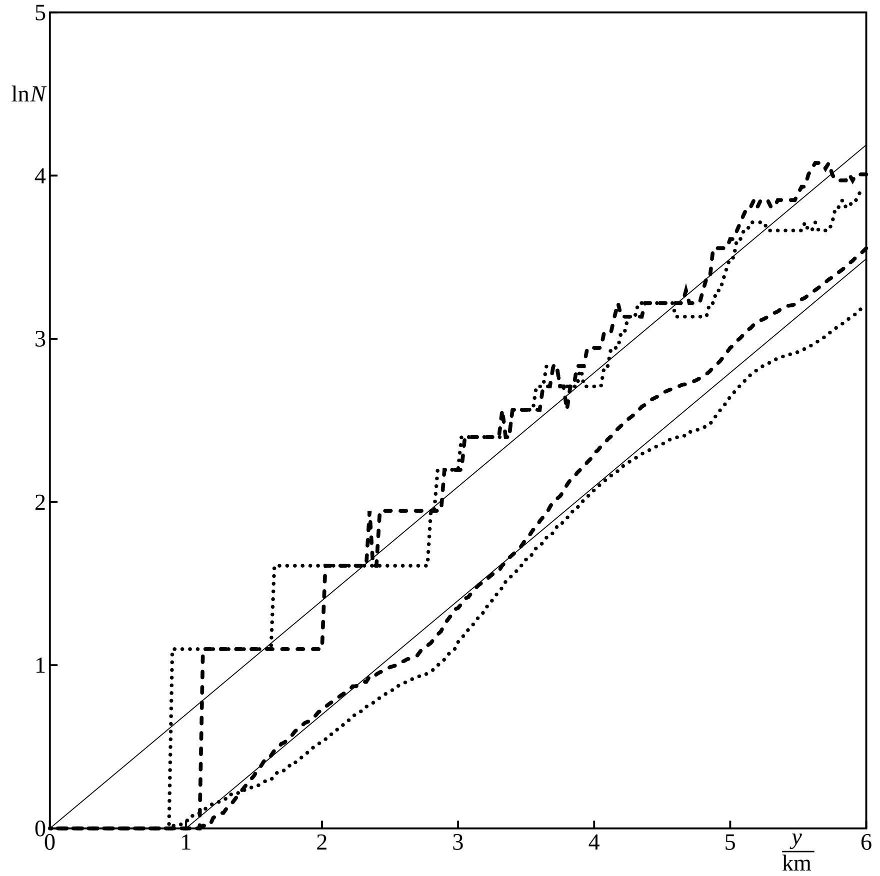

Figure 41. The natural logarithms of the average and maximum numbers of travel times along the individual horizontal grid lines of Figure 39 [*bold dotted lines*] and Figure 40 [*bold dashed lines*]. The horizontal axis represents the distance of the grid line from the bottom of the model box in kilometres, and serves as a rough approximation of the travel time in seconds. The slope of the *thin solid lines* is given by the average Lyapunov exponent (204) for the model.

which enables us to control the behaviour of the model in regions not illuminated by the data. A typical example of such application is the fitting of interfaces for which the data are available only in some parts of the model (Bulant, 2002). Secondly, the minimization of the corresponding Sobolev norm during the inversion of the given data allows us to construct a model optimal for the selected computational method, i.e. for ray tracing. Note that we still have no quantitative criteria of applicability and accuracy of ray methods or their extensions. Instead, we use the criterion of the numerical efficiency of ray tracing, based on the numbers of arrivals calculated in the model.

13.2.1. Application of Sobolev scalar products to smoothing models

The Lyapunov exponent describes the complexity of ray and travel–time fields with increasing travel time. We should thus establish the maximum average Lyapunov exponent for the model. This maximum Lyapunov exponent reflects the smoothness of the model required by future applications of the model.

Since the Lyapunov exponent depends on the second velocity derivatives, we should minimize the *second velocity derivatives* during the model smoothing or during the inversion of seismic data. The minimization of the second velocity derivatives can be achieved by including the square of the relevant Sobolev norm of the model into the objective function. The problem now is the choice of the weighting factor of the Sobolev norm in the objective function.

We may roughly estimate the maximum Sobolev norm of the model corresponding to the maximum Lyapunov exponent (Klimeš, 2000). Starting with zero or minimum (with respect to the stability of inversion) weighting factor, we may iteratively adjust the weighting factor so as to obtain the model with the Sobolev norm roughly equal to the estimated maximum Sobolev norm.

For this model, we calculate the average Lyapunov exponent. If the average Lyapunov exponent for the model does not accord with the maximum Lyapunov exponent, we should adjust our estimation of the maximum Sobolev norm and of the weighting factor in the objective function. If the average Lyapunov exponent for the model accords with the maximum Lyapunov exponent, we should perform ray tracing and other tests of suitability of the model for our applications. Based on these tests, we may possibly wish to adjust the maximum Lyapunov exponent, and consequently the maximum Sobolev norm and the weighting factor in the objective function.

Note that the selection of the Sobolev norm may reflect the application of the model. For example, if the model is designed for tracing only nearly vertical rays, vertical smoothing of the second velocity derivatives should be slighter than horizontal smoothing, according to the estimated deviation of rays from the vertical direction. The resulting model, optimized for nearly vertical rays, may then be much more accurate than the model smoothed for all directions of rays, including refracted ones.

14. Conclusions. Areas open to further research

In this contribution, we have concentrated on the forward problems of the zero–order approximation of the ray method only, particularly on the problem of the numerical calculation of seismic wavefields in 2–D and 3–D laterally varying, isotropic, anisotropic or weakly anisotropic, perfectly elastic layered and block structures. We have not touched the applications in inverse problems (tomography, migration, etc.); our conclusions apply to these problems only indirectly. We have explained the methods and algorithms designed to perform ray tracing and dynamic ray tracing, which provide travel times, polarization vectors, DRT propagator matrices, vectorial ray amplitudes, ray–theory Green functions, and ray–theory seismograms. We have also discussed certain extensions of the zero–order approximation of the ray method, particularly the coupling ray theory of S waves, and the methods of summation of Gaussian beams and Gaussian packets. Suitable algorithms have been explained and numerical examples of computations have been presented.

Unfortunately, even in smooth models with interfaces, the proposed algorithms are not yet fully universal. Some open problems remain not only in the computation of amplitudes, but even in ray tracing and travel time computations.

Let us start with *ray tracing and travel time computations.* The theoretical formulation of ray tracing is now closer and closer to completeness. This is particularly true of P waves. Indeed, ray tracing and travel time computations of P waves are quite safe for isotropic, anisotropic and weakly anisotropic layered models. For S waves, however, the situation is not as simple. Although great progress has been made recently in the coupling ray theory of S waves, see Section 10, routine algorithms for S waves propagating in weakly anisotropic laterally varying media *with interfaces* do not yet exist. Moreover, the algorithms based on common ray tracing should still be carefully tested. The algorithms of anisotropic ray tracing should be tested carefully in the vicinity of S–wave singularities. We are, however, convinced that the problems of safe and stable algorithms for ray tracing and travel–time computations of S waves in weakly anisotropic laterally varying models with interfaces will be completed soon.

The next problem of ray tracing which has not yet been completely solved consists in ray tracing of *inhomogeneous waves* propagating in complex structures. At present, only a little on ray tracing of inhomogeneous waves is known, even if a perfectly elastic structure is considered, although these waves may theoretically be rather strong in certain cases. This applies to postcritically transmitted waves, to waves tunnelling through a thin high–velocity layer, to waves penetrating into a caustic shadow, etc. Whereas the inhomogeneous waves are exceptional in perfectly elastic media, they are common in dissipative media. There are two principal approaches for ray tracing and travel–time computations of inhomogeneous waves in dissipative media: (a) Complex ray tracing, in 12–dimensional phase space. There are still open problems in complex ray tracing, both theoretical and algorithmical. In particular, we still do not know how to introduce complex–valued interfaces into a complex model. (b) The perturbation methods for weakly dissipative media, starting from non–dissipative models. As the dissipation is mostly weak in the Earth's interior, the first–order perturbation methods will usually be sufficient. They could also be applied to viscoelastic, poroelastic, poroviscoelastic and Biot media. It would, however, be convenient to have the possibility of exploiting also complex ray tracing, at least for testing purposes. Of course, in this case, ray tracing will be frequency dependent (or space–time dependent). The problem of rays in dissipative media is very interesting, at least from the theoretical point of view. Note that dissipative properties may play an important role in many regions of the Earth's interior.

Next let us discuss the open problems in the computation of *ray amplitudes, ray–theory seismograms, etc.* For simplicity, we shall speak only of ray amplitudes. In contrast to ray tracing, the main problem consists in the accuracy of the computed ray amplitudes, not in their computation. Moreover, the ray amplitudes are more sensitive to details of the structure than the travel times. Once rays have been computed, the computation of ray amplitudes along the ray is usually simple. It consists in dynamic ray tracing along the ray and in the subsequent computation of geometrical spreading and of the phase shift due to caustics. Once geometrical spreading and the phase shift due to caustics are known, the computation of ray amplitudes is easy and straightforward. It is only necessary to compute the R/T coefficients at points of contact of the ray with interfaces, the radiation function of the source and to include the receiving conditions at the receiver. For isotropic, anisotropic and weakly anisotropic media, most of these problems have been theoretically solved. The only problem which has not yet been fully solved consists in the coupling ray theory across interfaces. Nevertheless, we think that suitable algorithms for this case will be designed soon.

For dissipative media (isotropic, anisotropic or weakly anisotropic), the problem of evaluation of ray amplitudes is practically open. Similarly as in the case of ray tracing, we think that the most suitable approach would be based on perturbations from non–dissipative to weakly dissipative media.

The accuracy of ray amplitudes may be poor, particularly in singular regions. Such singular regions may be rather extensive and may even overlap.

For smooth isotropic media without interfaces, the situation is simpler, as only one type of singularity exists in this case — the caustic. The caustics themselves may, however, attain various forms, and various extensions of the ray method must be used to compute the amplitudes. In smooth anisotropic media, additional complications are caused by S–wave singularities, i. e. directions, in which two S–waves propagate with the same phase velocities. In models with interfaces, other singularities often appear (edge points, vertex points, critical points, points of tangency of the ray with the interface, etc.). The ray tracing procedures for diffracted waves are usually no more complicated than for zero–order waves. For example, for edge and vertex waves, we only need to specify properly the initial conditions for the rays, but the ray tracing itself remains the same as for zero–order elementary waves, see Figure 7. The amplitudes of diffracted waves, however, must be calculated using the equations of the geometric theory of diffraction. For certain types of diffracted waves, however, the ray tracing procedures must be modified. This applies, for example, to rays grazing along smooth interfaces. Such rays have been discussed mostly in some special analytic cases, but general procedures are not yet available (as far as we know).

The computation of amplitudes in singular regions may be performed by specific methods, designed for specific types of singularities. The disadvantage is that the algorithms are different for different diffracted waves. Moreover, singular regions often overlap. We would prefer to use some extensions of the ray method, applicable uniformly for any singular region, or at least for most of them. If such extensions were available, the ray computations of amplitudes would simplify considerably. Several such extensions of the ray method have been proposed (summation of Gaussian beams, summation of Gaussian packets, the Maslov method, the method of coherent states, parabolic equation methods, etc). We use mainly the methods based on the summation of Gaussian beams and packets for this purpose. A lot of work on Gaussian beams and packets in laterally varying isotropic layered structures has been done recently, see Section 11. All this work, however, has been mostly done for isotropic media; it should also be extended to anisotropic and weakly anisotropic media. We expect that the summation of Gaussian beams or packets will be one of most important topics in the seismic ray method for several years. We also believe that Gaussian beams and Gaussian packets will find numerous applications in seismology and seismic exploration.

Acknowledgements

Petr Bulant prepared many figures used in this paper.

The research has been supported by the Grant Agency of the Czech Republic under Contracts 205/01/0927, 205/01/D097, 205/04/1104 and 205/05/2182, by the Grant Agency of the Charles University under Contract 375/2004/B-GEO/MFF, by the Grant Agency of the Academy of Sciences of the Czech Republic under Contract A3012309, by the Ministry of Education of the Czech Republic within Research Project MSM113200004, and by the members of the consortium "Seismic Waves in Complex 3–D Structures" (see "http://sw3d.mff.cuni.cz").

References

Abgrall, R. & Benamou, J.-D. (1999): Big ray–tracing and eikonal solver on unstructured grids: Application to the computation of a multivalued traveltime field in the Marmousi model. *Geophysics*, **64**, 230–239.

Aki, K. & Richards, P. (1980): *Quantitative Seismology (2 vols).* W.H. Freeman & Co, San Francisco.

Alonso, M.A. & Forbes, G.W. (1995): Fractional Legendre transformation. *J. Phys. A*, **28**, 5509–5527.

Alonso, M.A. & Forbes, G.W. (1998): Asymptotic estimation of the optical wave propagator. I. Derivation of a new method. *J. Opt. Soc. Am. A*, **15**, 1329–1340.

Arnaud, J.A. (1972): Modes in helical gas lenses. *Applied Optics*, **11**, 2514–2521.

Babich, V.M. & Buldyrev, V.S. (1972): *Asymptotic Methods in Problems of Diffraction of Short Waves* (in Russian). Nauka, Moscow, Translated to English by Springer, Berlin 1991, under the title *Short-wavelength diffraction theory.*.

Babich, V.M., Buldyrev, V.S. & Molotkov, I.A. (1985): *Space–time Ray Method. Linear and Non–linear Waves* (in Russian). Leningrad Univ. Press, Leningrad.

Babich, V.M. & Ulin, V.V. (1981): Complex space–time ray method and "quasiphotons" (in Russian). In: Babich, V.M. (ed.): *Mathematical problems of the theory of propagation of waves, Vol. 12*, pp. 5–12, Nauka, Leningrad, English transl.: *J. Sov. Math.*, **24**(1984), 269–273.

Bakker, P.M. (1998): Phase shift at caustics along rays in anisotropic media. *Geophys. J. int.*, **134**, 515–518.

Bakker, P.M. (2002): Coupled anisotropic shear wave raytracing in situations where associated slowness sheets are almost tangent. *Pure appl. Geophys.*, **159**, 1403–1417.

Bleistein, N. (1984): *Mathematical methods for wave phenomena.* Academic Press, New York.

Bleistein, N., Cohen, J.K. & Stockwell Jr., J.W. (2001): *Mathematics of multidimensional seismic imaging, migration, and inversion.* Springer, Berlin.

Bulant, P. (1996): Two–point ray tracing in 3–D. *Pure appl. Geophys.*, **148**, 421–447.

Bulant, P. (1999): Two–point ray–tracing and controlled initial–value ray–tracing in 3–D heterogeneous block structures. *J. seism. Explor.*, **8**, 57–75.

Bulant, P. (2002): Sobolev scalar products in the construction of velocity models — application to model Hess and to SEG/EAGE Salt Model. *Pure appl. Geophys.*, **159**, 1487–1506.

Bulant, P. & Klimeš, L. (1999): Interpolation of ray theory traveltimes within ray cells. *Geophys. J. int.*, **139**, 273–282.

Bulant, P. & Klimeš, L. (2002): Numerical algorithm of the coupling ray theory in weakly anisotropic media. *Pure appl. Geophys.*, **159**, 1419–1435.

Bulant, P. & Klimeš, L. (2004): Comparison of quasi–isotropic approximations of the coupling ray theory with the exact solution in the 1–D anisotropic "oblique twisted crystal" model. *Stud. geophys. geod.*, **48**, 97–116.

Bulant, P., Klimeš, L., Pšenčík, I. & Vavryčuk, V. (2004): Comparison of ray methods with the exact solution in the 1–D anisotropic "simplified twisted crystal" model. *Stud. geophys. geod.*, **48**, 675–688.

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

Červený, V. (1982): Direct and inverse kinematic problems for inhomogeneous anisotropic media - Linearization approach. *Contrib. Geophys. Inst. Slov. Acad. Sci.*, **13**, 127–133.

Červený, V. (1985): Gaussian beam synthetic seismograms. *J. Geophys.*, **58**, 44–72.

Červený, V. (1989): Synthetic body wave seismograms for laterally varying media containing thin transmission layers. *Geophys. J. int.*, **99**, 331–349.

Červený, V. (2001): *Seismic Ray Theory.* Cambridge Univ. Press, Cambridge.

Červený, V. & Andrade F.C.M. (1992): Influence of a near–surface structure on seismic wave fields recorded at the Earth's surface. *J. seism. Explor.*, **1**, 107–116.

Červený, V. & Aranha P.R.A. (1992): Tuneling of seismic body waves through thin high–velocity layers in complex structures. *Stud. geophys. geod.*, **36**, 115–138.

Červený, V. & Jech, J. (1982): Linearized solutions of kinematic problems of seismic body waves in inhomogeneous slightly anisotropic media. *J. Geophys.*, **51**, 96–104.

Červený, V., Klimeš, L. & Pšenčík, I. (1988): Complete seismic–ray tracing in three–dimensional structures. In: Doornbos, D.J. (ed.): *Seismological Algorithms*, pp. 89–168, Academic Press, New York.

Červený, V., Molotkov, I.A. & Pšenčík, I. (1977): *Ray Method in Seismology.* Charles Univ., Praha.

Červený, V., Popov, M.M. & Pšenčík, I. (1982): Computation of wave fields in inhomogeneous media — Gaussian beam approach. *Geophys. J. R. astr. Soc.*, **70**, 109–128.

Červený, V. & Pšenčík, I. (1984): SEIS83 – Numerical modelling of seismic wave fields in 2–D laterally varying layered structures by the ray method. In: Engdahl,E.R. (ed.): *Documentation of Earthquake Algorithms, Report SE-35*, pp. 36–40, World Data Center A for Solid Earth Geophysics, Boulder.

Červený, V. & Soares, J.E.P. (1992): Fresnel volume ray tracing. *Geophysics*, **57**, 902–915.

Chapman, C.H. (1994): Reflection/transmission coefficient reciprocities in anisotropic media. *Geophys. J. int.*, **116**, 498–501.

Chapman, C.H. (2002): Seismic ray theory and finite frequency extensions. In: Lee, W.H.K., Kanamori, H. & Jennings, P.C. (eds.): *International Handbook of Earthquake and Engineering Seismology, Part A*, pp. 103–123, Academic Press, New York.

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

Chapman, C.H. & Drummond, R. (1982): Body–wave seismograms in inhomogeneous media using Maslov asymptotic theory. *Bull. Seismol. Soc. Am.*, **72**, S277–S317.

Chapman, C.H. & Pratt, R.G. (1992): Traveltime tomography in anisotropic media — I. Theory. *Geophys. J. int.*, **109**, 1–19.

Chapman, S.J., Lawry, J.M.H., Ockendon, J.R. & Tew, R.H. (1999): On the theory of complex rays. *SIAM Rev.*, **41**, 417–509.

Coates, R.T. & Chapman, C.H. (1990): Quasi–shear wave coupling in weakly anisotropic 3–D media. *Geophys. J. int.*, **103**, 301–320.

Coates, R.T. & Chapman, C.H. (1991): Generalized Born scattering of elastic waves in 3–D media. *Geophys. J. int.*, **107**, 231–263.

Combescure, M., Ralston, J. & Robert, D. (1999): A proof of the Gutzwiller semiclassical trace formula using coherent state decomposition. *Commun. math. Phys.*, **202**, 463–480.

Condon, E.U. (1937): Immersion of the Fourier transform in a continuous group of functional transformations. *Proc. Nat. Acad. Sci.*, **23**, 158–164.

Dahlen, F.A. & Tromp,J. (1998): *Theoretical Global Seismology.* Princeton Univ. Press, Princeton.

Daubechies, I. (1991): *Ten Lectures on Wavelets.* Society for Industrial and Applied Mathematics, Pennsylvania.

Eisner, L. & Pšenčík, I. (1996): Computation of additional components of the first-order ray approximation in isotropic media. *Pure appl. Geophys.*, **148**, 227–253.

Farra, V. (1989): Ray perturbation theory for heterogeneous hexagonal anisotropic medium. *Geophys. J. int.*, **99**, 723–737.

Farra, V. (1999): Computation of second–order traveltime perturbation by Hamiltonian ray theory. *Geophys. J. int.*, **136**, 205–217.

Farra, V. (2004): First–order ray tracing for qS waves in inhomogeneous weakly anisotropic media. *Geophys. J. int.*, submitted.

Farra, V & Le Bégat, S. (1995): Sensitivity of qP–wave traveltimes and polarization vectors to heterogeneity, anisotropy, and interfaces. *Geophys. J. int.*, **121**, 371–384.

Farra, V. & Madariaga, R. (1987): Seismic waveform modeling in heterogeneous media by ray perturbation theory. *J. geophys. Res.*, **92B**, 2697–2712.

Farra, V. & Pšenčík, I. (2003): Properties of the zero-, first- and higher-order approximations of attributes of elastic waves in weakly anisotropic media. *J. Acoust. Soc. Am.*, **114**, 1366–1378.

Farra, V. & Pšenčík, I. (2004): First–order ray tracing for qP waves in inhomogeneous weakly anisotropic media. In: *Seismic Waves in Complex 3–D Structures, Report 14*, pp. 87–97, Dep. Geophys., Charles Univ., Prague, online at "http://sw3d.mff.cuni.cz".

Felsen, L.B. & Markuvitz, N. (1973): *Radiation and scattering of waves.* Prentice Hall, Englewood Cliffs.

Forbes, G.W. & Alonso, M.A. (1998): Asymptotic estimation of the optical wave propagator. II. Relative validity. *J. Opt. Soc. Am. A*, **15**, 1341–1354.

Foster, D.J. & Huang, J-I. (1991): Global asymptotic solutions of the wave equation. *Geophys. J. int.*, **105**, 163–171.

Fradkin, L.Yu. & Kiselev A.P. (1997): The two-component representation of time-harmonic elastic body waves in the high- and intermediate-frequency regimes. *J. Acoust. Soc. Am.*, **101**, 52–65.

Gajewski, D. & Pšenčík, I. (1987): Computation of high–frequency seismic wavefields in 3–D laterally inhomogeneous anisotropic media. *Geophys. J. R. astr. Soc.*, **91**, 383–411.

Gajewski, D. & Pšenčík, I. (1992): Vector wavefields for weakly attenuating anisotropic media by the ray method. *Geophysics*, **57**, 27–38.

Garmany, J. (2000): Phase shifts at caustics in anisotropic media. In: *Anisotropy 2000: Fractures, Converted Waves and Case Studies*, pp. 419–425, Soc. Explor. Geophysicists, Tulsa.

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

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

Goldin, S.V. (1989): Physical analysis of the additional components of seismic waves in the first approximation of ray series. *Geology and Geophysics*, **30**, 128–132.

Hanyga, A. (1982): The kinematic inverse problem for weakly laterally inhomogeneous anisotropic media. *Tectonophysics*, **90**, 253–262.

Hanyga, A. (1995): Asymptotic edge-and-vertex diffraction theory. *Geophys. J. int.*, **123**, 227–290.

Hanyga, A., Druzhinin, A.B., Dzhafarov, A.D. & Frøyland, L. (2001): A Hamiltonian approach to asymptotic seismic reflection and diffraction modeling. *Geophys. Prospecting*, **49**, 213–227.

Hanyga, A., Lenartowicz, E. & Pajchel, J. (1984): *Seismic Wave Propagation in the Earth.* Elsevier, Amsterdam.

Hill, N.R. (1990): Gaussian beam migration. *Geophysics*, **55**, 1416–1428.

Hill, N.R. (2001): Prestack Gaussian–beam depth migration. *Geophysics*, **66**, 1240–1250.

Hubral, P., Schleicher, J. & Tygel, M. (1992a): Three-dimensional paraxial ray properties: Part I. Basic relations. *J. seism. Explor.*, **1**, 265–279.

Hubral, P., Schleicher, J. & Tygel, M. (1992b): Three-dimensional paraxial ray properties: Part II. Applications. *J. seism. Explor.*, **1**, 347–362.

Jech, J. & Pšenčík, I. (1989): First–order perturbation method for anisotropic media. *Geophys. J. int.*, **99**, 369–376.

Johnson, R.A. (1986): An example concerning the geometrical significance of the rotation number — integrated density of states. In: Arnold, L. & Wihstutz, V. (eds.): *Lyapunov Exponents*, pp. 216–226, Springer, Berlin.

Kachalov, A.P.A coordinate system for describing the "quasiphoton" (in Russian). In: Babich, V.M. (ed.): *Mathematical problems of the theory of propagation of waves, Vol. 14*, pp. 73–76, Nauka, Leningrad, English transl.: *J. Sov. Math.*, **32**(1986), 151–153.

Kapoor, T.K. & Felsen, L.B. (1995): Hybrid ray-mode analysis of acoustic scattering from a finite, fluid loaded plate. *Wave Motion*, **22**, 109–131.

Katok, S.R. (1980): The estimation from above for the topological entropy of a diffeomorphism. In: Netecki, Z. & Robinson, C. (eds.): *Global Theory of Dynamical Systems*, pp. 258–264, Lecture Notes in Mathematics, Vol. 819, Springer, Berlin.

Keers, H., Dahlen, F.A. & Nolet, G. (1997): Chaotic ray behaviour in regional seismology. *Geophys. J. int.*, **131**, 361–380.

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

Kendall, J-M. & Thomson, C.J. (1993): Maslov ray summation, pseudo–caustics, Lagrangian equivalence and transient seismic waveforms. *Geophys. J. int.*, **113**, 186–214.

Klauder, J.R. (1987): Semiclassical quantization of classically chaotic systems. *Phys. Rev. Lett.*, **59**, 748–750.

Klimeš, L. (1984a): Expansion of a high–frequency time–harmonic wavefield given on an initial surface into Gaussian beams. *Geophys. J. R. astr. Soc.*, **79**, 105–118.

Klimeš, L. (1984b): The relation between Gaussian beams and Maslov asymptotic theory. *Stud. geophys. geod.*, **28**, 237–247.

Klimeš, L. (1985): *Computation of seismic wavefields in 3–D media by the Gaussian beam method. Program package SW84.* Research Report No. 68 for Geofyzika Brno, Institute of Geophysics, Charles University, Prague.

Klimeš, L. (1986): Discretization error for the superposition of Gaussian beams. *Geophys. J. R. astr. Soc.*, **86**, 531–551.

Klimeš, L. (1989a): Optimization of the shape of Gaussian beams of a fixed length. *Stud. geophys. geod.*, **33**, 146–163.

Klimeš, L. (1989b): Gaussian packets in the computation of seismic wavefields. *Geophys. J. int.*, **99**, 421–433.

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

Klimeš, L. (1996): Grid travel–time tracing: second–order method for the first arrivals in smooth media. *Pure appl. Geophys.*, **148**, 539–563.

Klimeš, L. (1997): Phase shift of the Green function due to caustics in anisotropic media. In: *Seismic Waves in Complex 3–D Structures, Report 6*, pp. 167–173, Dep. Geophys., Charles Univ., Prague, online at "http://sw3d.mff.cuni.cz".

Klimeš, L. (2000): Sobolev scalar products in the construction of velocity models. In: *Seismic Waves in Complex 3–D Structures, Report 10*, pp. 15–40, Dep. Geophys., Charles Univ., Prague, online at "http://sw3d.mff.cuni.cz".

Klimeš, L. (2002a): Second–order and higher–order perturbations of travel time in isotropic and anisotropic media. *Stud. geophys. geod.*, **46**, 213–248.

Klimeš, L. (2002b): Lyapunov exponents for 2–D ray tracing without interfaces. *Pure appl. Geophys.*, **159**, 1465–1485.

Klimeš, L. (2002c): Application of the medium covariance functions to travel–time tomography. *Pure appl. Geophys.*, **159**, 1791–1810.

Klimeš, L. (2003a): Perturbations and spatial derivatives of amplitude in isotropic and anisotropic media. In: *Seismic Waves in Complex 3–D Structures, Report 13*, pp. 107–118, Dep. Geophys., Charles Univ., Prague, online at "http://sw3d.mff .cuni.cz".

Klimeš, L. (2003b): Common ray tracing and dynamic ray tracing for S waves in a smooth elastic anisotropic medium. In: *Seismic Waves in Complex 3–D Structures, Report 13*, pp. 119–141, Dep. Geophys., Charles Univ., Prague, online at "http://sw3d.mff.cuni.cz".

Klimeš, L. (2004a): Analytical one–way plane–wave solution in the 1–D anisotropic "simplified twisted crystal" model. *Stud. geophys. geod.*, **48**, 75–96.

Klimeš, L. (2004b): Gaussian packets in smooth isotropic media. In: *Seismic Waves in Complex 3–D Structures, Report 14*, pp. 43–54, Dep. Geophys., Charles Univ., Prague, online at "http://sw3d.mff.cuni.cz".

Klimeš, L. & Bulant, P. (2004): Errors due to the common ray approximations of the coupling ray theory. *Stud. geophys. geod.*, **48**, 117–142.

Klimeš, L. & Kvasnička, M. (1994): 3–D network ray tracing. *Geophys. J. int.*, **116**, 726–738.

Kravtsov, Yu.A., Forbes, G.W. & Asatryan, A.A. (1999): Theory and applications of complex rays. In: Wolf, E. (ed.): *Progress in Optics, Vol.39*, pp. 1–62, Elsevier, Amsterdam.

Kravtsov, Yu.A. & Orlov, Yu.I. (1980): *Geometrical Optics of Inhomogeneous Media* (in Russian). Nauka, Moscow, Translation to English by Springer, Berlin, 1990.

Kravtsov, Yu.A. & Orlov, Yu.I. (1999): *Caustics, Catastrophes and Wave Fields.* Springer, Heidelberg.

Lambaré, G., Lucio, P.S. & Hanyga, A. (1996): Two–dimensional multivalued traveltime and amplitude maps by uniform sampling of ray field. *Geophys. J. int.*, **125**, 584–598.

Lecomte, I. (1996): Hybrid modeling with ray tracing and finite difference. In: *Expanded Abstracts of 66th SEG Annual Meeting*, pp. 699–702, Soc. Explor. Geophysicists, Tulsa.

Lewis, R.M. (1965): Asymptotic theory of wave–propagation. *Arch. ration. Mech. Anal.*, **20**, 191–250.

Lyapunov, A.M. (1949): Problème Général de la Stabilité du Mouvement. In: *Annals of Mathematical Studies, Vol. 17*, pp. Princeton Univ. Press, Princeton.

Martin, B.E. & Thomson, C.J. (1997): Modelling surface waves in anisotropic structures. II. Examples. *Phys. Earth Planet. Int.*, **103**, 253–279.

Maslov, V.P. (1965): *Theory of perturbations and asymptotic methods* (in Russian). Izd. MGU, Moscow.

Moczo, P., Bard, P.Y. & Pšenčík, I. (1987): Seismic response of 2–D absorbing structure by the ray method. *J. Geophys.*, **62**, 38–49.

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

Moser, T-J. (2004): Review of the anisotropic interface ray propagator: Symplecticity, eigenvalues, invariants and applications. *Studia geophys. geod.*, **48**, 47–73.

Norris, A., White, B.S. & Schreiffer, J.R. (1987): Gaussian wave packets in inhomogeneous media with curved interfaces. *Proc. R. Soc. London*, **A 412**, 93–123.

Nowack, R.L. & Pšenčík, I. (1991): Travel–time and ray path perturbation from isotropic to anisotropic media. *Geophys. J. int.*, **106**, 1–10.

Opršal, I., Brokešová, J., Fäh, D. & Girardini, D. (2002): 3D hybrid ray–FD and DWN–FD seismic modeling for simple models containing complex local structures. *Stud. geophys. geod.*, **46**, 711–730.

Oseledec, V.I. (1968): A multiplicative ergodic theorem: Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc.*, **19**, 197–231.

Ozaktas, H.M., Zalevsky, Z. & Kutay, M.A. (2001): *The Fractional Fourier Transform with Applications in Optics and Signal Processing.* John Wiley & Sons, Chichester.

Podvin, P. & Lecomte, I. (1991): Finite difference computation of traveltimes in very contrasted velocity models: a massively parallel approach and its associated tools. *Geophys. J. int.*, **105**, 271–284.

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

Popov, M.M. & Pšenčík, I. (1978a): Ray amplitudes in inhomogeneous media with curved interfaces. *Travaux Instit. Géophys. Acad. Tchécosl. Sci. No. 454, Geofys. Sborník*, **24**, 111–129, Academia, Praha.

Popov, M.M. & Pšenčík, I. (1978b): Computation of ray amplitudes in inhomogeneous media with curved interfaces. *Stud. geophys. geod.*, **22**, 248–258.

Pšenčík, I. (1998): Green's functions for inhomogeneous weakly anisotropic media. *Geophys. J. int.*, **135**, 279–288.

Pšenčík, I. & Dellinger, J. (2001): Quasi–shear waves in inhomogeneous weakly anisotropic media by the quasi–isotropic approach: A model study. *Geophysics*, **66**, 308–319.

Pšenčík, I. & Teles, T.N. (1996): Point–source radiation in inhomogeneous anisotropic structures. *Pure appl. Geophys.*, **148**, 591–623.

Ralston, J. (1983): Gaussian beams and the propagation of singularities. In: Littman W. (ed.): *Studies in Partial Differential Equations, MAA Studies in Mathematics, Vol.23*, pp. 206–248, The Math. Assoc. Amer., Washington, DC.

Rümpker, G. & Thomson, C.J. (1994): Seismic–waveform effects of conical points in gradually varying anisotropic media. *Geophys. J. int.*, **118**, 759–780.

Schleicher, J., Tygel, M. & Hubral, P. (1993): 3–D true–amplitude finite–offset migration. *Geophysics*, **58**, 1112–1126.

Schleicher, J., Tygel, M., Ursin, B. & Bleistein, N. (2001): The Kirchhoff-Helmholtz integral for anisotropic elastic media. *Wave Motion*, **34**, 353–364.

Shlivinski, A., Heyman, E., Boag, A., & Letrou, C. (2004): Phase–space beam summation formulations for ultra wideband (UWB) radiation. In: *Proceedings of URSI International Symposium on Electromagnetic Theory (Pisa)*, pp. 936–938, Edizioni PLUS, Università di Pisa.

Sun, J. (2004): True-amplitude weight functions in 3D limited aperture migration revisited. *Geophysics*, **60**, 1025–1036.

Tarantola, A. (1987): *Inverse Problem Theory.* Elsevier, Amsterdam.

Tessmer, E. (1995): 3–D seismic modelling of general material anisotropy in the presence of the free surface by a Chebyshev spectral method. *Geophy. J. int.*, **121**, 557–575.

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

Thomson, C.J. (1997): Complex rays and wave packets for decaying signals in inhomogeneous, anisotropic and anelastic media. *Stud. geophys. geod.*, **41**, 345–381.

Thomson, C.J. (1999): The 'gap' between seismic ray theory and 'full' wavefield extrapolation. *Geophys. J. int.*, **137**, 364–380.

Thomson, C.J. (2001): Seismic coherent states and ray geometrical spreading. *Geophys. J. int.*, **144**, 320–342.

Thomson, C.J., Kendall, J-M. & Guest, W.S. (1992): Geometrical theory of shear–wave splitting: corrections to ray theory for interference in isotropic/anisotropic transitions. *Geophys. J. int.*, **108**, 339–363.

Tygel, M., Schleicher, J. & Hubral, P. (1994): Kirchhoff-Helmholtz theory in modelling and migration. *J. seism. Explor.*, **3**, 203–214.

Ursin, B. & Tygel, M. (1997): Reciprocal volume and surface scattering integrals for anisotropic elastic media. *Wave Motion*, **26**, 31–42.

Vavryčuk, V. (1999): Applicability of higher–order ray theory for S wave propagation in inhomogeneous weakly anisotropic elastic media. *J. geophys. Res.*, **104B**, 28829–28840.

Vavryčuk, V. (2001): Ray tracing in anisotropic media with singularities. *Geophys. J. int.*, **145**, 265–276.

Vidale, J.E. (1990): Finite–difference calculation of traveltimes in three dimensions. *Geophysics*, **55**, 521–526.

Vinje, V. (1997): A new interpolation criterion for controlling accuracy in wavefront construction. In: *Expanded Abstracts of 67th SEG Annual Meeting (Dallas)*, pp. 1723–1726, Soc. Explor. Geophysicists, Tulsa.

Vinje, V., Iversen, E., Åstebøl, K. & Gjøystdal, H. (1996a): Estimation of multivalued arrivals in 3D models using wavefront construction — Part I. *Geophys. Prospecting*, **44**, 819–842.

Vinje, V., Iversen, E., Åstebøl, K. & Gjøystdal, H. (1996b): Part II: Tracing and interpolation. *Geophys. Prospecting*, **44**, 843–858.
Vinje, V., Iversen, E. & Gjøystdal, H. (1993a): Traveltime and amplitude estimation using wavefront construction. *Geophysics*, **58**, 1157–1166.
Vinje, V., Iversen, E., Gjøystdal, H. & Åstebøl, K. (1993b): Estimation of multivalued arrivals in 3D models using wavefront construction. In: *Extended Abstracts of 55th SEG Annual Meeting (Stavanger)*, pp. B019, Eur. Assoc. Explor. Geophysicists, Zeist.
Virieux, J. (1996): Seismic ray tracing. In: Boschi, E., Ekström, G. & Morelli, A (eds.): *Seismic Modelling of the Earth Structure*, pp. 223–304, Instituto Nazionale di Geofisica, Roma.
White, B.S., Nair, B. & Bayliss, A. (1988): Random rays and seismic amplitude anomalies. *Geophysics*, **53**, 903–907.
Wolf, K.B. (1974): Canonical Transforms. I. Complex linear transforms. *J. math. Phys.*, **15**, 1295–1301.
Wolf, K.B. (1979): *Integral Transforms in Science and Engineering.* Plenum Press, New York.
Wolfson, M.A. & Tappert, F.D. (2000): Study of horizontal multipaths and ray chaos due to ocean mezoscale structure. *J. Acoust. Soc. Am.*, **107**, 154–162.
Žáček, K. (2001a): Optimization of the shape of Gaussian beams. In: *Seismic Waves in Complex 3–D Structures, Report 11*, pp. 181–201, Dep. Geophys., Charles Univ., Prague, online at "http://sw3d.mff.cuni.cz".
Žáček, K. (2001b): Optimization of the shape of Gaussian beams. In: *Expanded Abstracts of 71st SEG Annual Meeting (San Antonio)*, pp. 2128–2131, Soc. Explor. Geophysicists, Tulsa.
Žáček, K. (2002): Smoothing the Marmousi model. *Pure appl. Geophys.*, **159**, 1507–1526.
Žáček, K. (2003a): Decomposition of the wavefield into optimized Gaussian packets. In: *Seismic Waves in Complex 3–D Structures, Report 13*, pp. 35–43, Dep. Geophys., Charles Univ., Prague, online at "http://sw3d.mff.cuni.cz".
Žáček, K. (2003b): Decomposition of the wave field into optimized Gaussian packets. In: *Expanded Abstracts of 73rd SEG Annual Meeting (Dallas)*, pp. 1869–1872, Soc. Explor. Geophysicists, Tulsa.
Žáček, K. (2004): Gaussian–packet prestack depth migration. In: *Seismic Waves in Complex 3–D Structures, Report 14*, pp. 17–25, Dep. Geophys., Charles Univ., Prague, online at "http://sw3d.mff.cuni.cz".
Žáček, K. & Klimeš, L. (2003): Sensitivity of seismic waves to the structure. In: *Expanded Abstracts of 73rd SEG Annual Meeting (Dallas)*, pp. 1857–1860, Soc. Explor. Geophysicists, Tulsa.
Zhao, L. & Dahlen, F.A. (1996): Mode-sum to ray-sum transformation in a spherical and an aspherical earth. *Geophys. J. int.*, **126**, 389–412.