Fermat's variational principle for anisotropic inhomogeneous media

Vlastislav Červený

Department of Geophysics, Faculty of Mathematics and Physics,
Charles University, Ke Karlovu 3, 121 16 Praha 2, Czech Republic.
E-mail vcerveny@seis.karlov.mff.cuni.cz

Summary

In seismology, Fermat's variational principle has mostly been used in parameteric form. It is valid for any parameter u used to specify the position of points on curves. The relevant Lagrangian $\mathcal{L}(x^k, \mathrm{d}x^k/\mathrm{d}u)$, where $x^k, k = 1, 2, 3$, are general curvilinear coordinates, is then a homogeneous function of the first degree in $\mathrm{d}x^k/\mathrm{d}u$. It is shown that the Legendre transform cannot be applied to this Lagrangian to derive the relevant Hamiltonian $\mathcal{H}(x^k, p_k)$ and Hamiltonian ray equations. The reason is that the Hessian determinant of the transformation vanishes identically if the Lagrangian is a homogeneous function of the first degree. The Lagrangians must be modified so that the Hessian determinant is different from zero. Two such modifications are proposed in this article. In the first modification, the selected parameter u along the curves is chosen to correspond to travel time τ, and the modified Lagrangian $\mathcal{L}^M(x^k, \mathrm{d}x^k/\mathrm{d}\tau)$ is introduced by the relation $\mathcal{L}^M(x^k, \mathrm{d}x^k/\mathrm{d}\tau) = \frac{1}{2}[\mathcal{L}(x^k, \mathrm{d}x^k/\mathrm{d}\tau)]^2$. The modified Lagrangian $\mathcal{L}^M(x^k, \mathrm{d}x^k/\mathrm{d}\tau)$ yields the same Euler-Lagrange equations as the standard parameteric Lagrangian $\mathcal{L}(x^k, \mathrm{d}x^k/\mathrm{d}u)$, but represents a homogeneous function of the second order in $\mathrm{d}x^k/\mathrm{d}\tau$ (not of the first order). Consequently, the relevant Hessian determinant does not vanish identically. In the second modification, one of the coordinates x^k, e.g., the coordinate x^3, is chosen to represent parameter u. Here the relevant Lagrangian $\mathcal{L}^R(x^k, \mathrm{d}x^1/\mathrm{d}x^3, \mathrm{d}x^2/\mathrm{d}x^3)$ is referred to as the reduced Lagrangian. Again, the Hessian determinant does not identically vanish in this case. In both cases, the Legendre transform can be used to compute the Hamiltonian from the Lagrangian, and vice versa, and the Hamiltonian canonical equations can be derived from the Euler-Lagrange equations. The relations between modified Hamiltonians and Lagrangians are discussed in detail. It is shown that the standard form of the Hamiltonian, derived from the elastodynamic equation and representing the eikonal equation, which has been broadly used in the seismic ray method, corresponds to the modified Lagrangian $\mathcal{L}^M(x^k, \mathrm{d}x^k/\mathrm{d}\tau)$, not to the standard parameteric Lagrangian $\mathcal{L}(x^k, \mathrm{d}x^k/\mathrm{d}u)$. It is also shown that the relations $\mathcal{L}^M(x^k, \mathrm{d}x^k/\mathrm{d}\tau) = \frac{1}{2}$ and $\mathcal{H}^M(x^k, p_k) = \frac{1}{2}$ are valid along the whole ray and that they represent the group velocity surface and the slowness surface, respectively. All procedures and derived equations are valid for general anisotropic inhomogeneous media, and for general curvilinear coordinates x^i. To make certain procedures and equations more transparent and objective, the simpler cases of isotropic and ellipsoidally anisotropic media are briefly discussed as special cases.

Seismic Waves in Complex 3-D Structures, Report 11,
Department of Geophysics, Charles University, Praha 2001.

1 Introduction

The problem of Fermat's variational principle in an anisotropic inhomogeneous medium is mostly of academic interest only. Practically everything that follows from Fermat's variational principle can also be derived by applying high-frequency asymptotic methods to the elastodynamic equation, and from the consequent eikonal equation. The author wants to emphasize that he does not expect any practical consequences to follow from the presented treatment. The eikonal equation, following from the anisotropic elastodynamic equation using the asymptotic methods, reads,

$$G(x^k, p_k) = 1 \ , \tag{1}$$

where $G(x^k, p_k)$ is the eigenvalue of the Christoffel matrix $\Gamma_{ik} = a_{ik}^{jl} p_j p_l$. Here p_i are covariant components of the slowness vector, x^i the general curvilinear coordinates, and a_{ik}^{jl} the density normalized elastic moduli. The eikonal equation (1) corresponds to any of the three elementary waves (qP, qS1 or qS2) which may propagate in anisotropic inhomogeneous medium. Eikonal equation (1) may be used to write the Hamiltonian ray tracing equations:

$$\frac{\mathrm{d}x^i}{\mathrm{d}\tau} = \frac{1}{2}\frac{\partial G}{\partial p_i} \ , \quad \frac{\mathrm{d}p_i}{\mathrm{d}\tau} = -\frac{1}{2}\frac{\partial G}{\partial x^i} \ , \tag{2}$$

see Červený (2001a). Here τ denotes a parameter along the ray, corresponding to the travel time. It is not difficult to prove that $G(x^i, p_i)$ is a homogeneous function of the second degree in p_i. By a homogeneous function $f(p_i)$ of the k-th degree in p_i, we understand a function $f(p_i)$ which satisfies the relation

$$f(ap_i) = a^k f(p_i) \ , \tag{3}$$

for any nonvanishing constant a. The homogeneous function $f(p_i)$ of the k-th degree satisfies the Euler's theorem

$$p_j \partial f(p_i)/\partial p_j = k f(p_i) \ . \tag{4}$$

The ray tracing equations (2) take the standard form of Hamiltonian canonical equations

$$\frac{\mathrm{d}x^i}{\mathrm{d}\tau} = \frac{\partial \mathcal{H}}{\partial p_i} \ , \quad \frac{\mathrm{d}p_i}{\mathrm{d}\tau} = -\frac{\partial \mathcal{H}}{\partial x^i} \ , \tag{5}$$

if we introduce Hamiltonian $\mathcal{H}(x^k, p_k)$ by the relation

$$\mathcal{H}(x^k, p_k) = \tfrac{1}{2} G(x^k, p_k) \ . \tag{6}$$

Hamiltonian $\mathcal{H}(x^k, p_k)$ is again a homogeneous function of the second degree in p_i, similarly as $G(x^k, p_k)$. The relation

$$\mathcal{H}(x^k, p_k) = \tfrac{1}{2} \tag{7}$$

is satisfied along the whole ray. It is well known that relations (1) and (7) represent alternatively the slowness surface in a phase space specified by coordinates p_i.

Note that the same Hamiltonian ray equations (5) are obtained also if we use Hamiltonian $\mathcal{H}(x^k, p_k) = \frac{1}{2}(G(x^k, p_k) - 1)$. Then $\mathcal{H}(x^k, p_k)$ equals zero along the whole ray. This form is also common in the seismological literature, see Červený (2001a, Eq.(3.6.3)). Here, however, we shall use the Hamiltonian given by (6) consistently.

In this paper, we wish to derive general equations (1)–(2), (5)–(7) using Fermat's variational principle. We would also like to find the Lagrangian corresponding to Hamiltonian (6), and to establish general rules of computing Lagrangians from Hamiltonians, and vice versa. The Lagrangians are introduced here in terms of the wave propagation metric tensor, explained in Section 2, see (9). The wave propagation metric tensor is very closely related to the group velocity surface, see Section 3.2.

The Legendre transform will be used to compute the Lagrangians from Hamiltonians, and vice versa. It is shown that the Legendre transform cannot be used if the standard form of Fermat's variational principle in parameteric form is used, as the Hessian determinant of the transform vanishes in this case. For this reason, the Lagrangian must be modified so that the Hessian determinant is different from zero. Two ways of doing this are proposed in Sections 3 and 4. The modified Lagrangian $\mathcal{L}^M(x^k, \mathrm{d}x^k/\mathrm{d}\tau)$ is introduced in Section 3, and the reduced Lagrangian $\mathcal{L}^R(x^k, \mathrm{d}x_1/\mathrm{d}x_3,\ \mathrm{d}x_2/\mathrm{d}x_3)$ is introduced in Section 4. In both cases, the Hessian determinant is not identically zero, and the Legendre transform may be used. Particularly important is the modified Lagrangian $\mathcal{L}^M(x^k, \mathrm{d}x^k/\mathrm{d}\tau)$, as it corresponds to the Hamiltonian given by (6).

We consider general curvilinear coordinates x^i, $i = 1, 2, 3$. The Einstein summation convention is used and applies to repetitive subscripts and superscripts. Only in some examples in subsections 3.3, 3.4, 4.3 and 4.4, related to isotropic and ellipsoidally anisotropic media, we also use standard Cartesian coordinates. In these cases, all indices are written as subscripts, and the Einstein summation convention applies to repetitive lower indices.

2 Fermat's variational principle

Fermat's principle, in its original version, states that the ray connecting two points, S and R, is represented by a curve along which the least time for a signal to travel from S to R is required. We shall speak of Fermat's minimum-time principle. Fermat's minimum-time principle has been used recently in network ray tracing and yields the rays corresponding to first arrival travel times. See Moser (1991), Klimeš and Kvasnička (1994), and other references given there.

In this paper, we shall consider the variational form of Fermat's principle, also sometimes called Fermat's variational principle. Fermat's variational principle also yields the rays corresponding to later arrival travel times. More details on Fermat's variational principle for anisotropic inhomogeneous media can be found in Babich (1961), Kline and Kay (1965), Hanyga, Lenartowicz and Pajchel (1984), Chapman and Pratt (1992), Epstein and Śniatycki (1992), Hanyga (1996), Stephen (1996) Pšenčík (1997).

We introduce Fermat's functional $\mathcal{I}(l)$ as the travel time of the signal from S to R along curve l. Fermat's functional, defined in this way, depends, of course, on the curve l, connecting S and R. Then Fermat's variational principle reads: The signal propagates from point S to R along a curve $\bar{l}$ which renders Fermat's functional stationary.

We now have to specify what we understand under the travel time $\mathcal{I}(l)$ of a signal along curve l. In general, $\mathcal{I}(l)$ is given by the relation

$$\mathcal{I}(l) = \int_S^R \mathrm{d}\tau \ , \tag{8}$$

where $\mathrm{d}\tau$ is the elementary time increment measured along the elementary arclength increment $\mathrm{d}l$ of curve l. To specify $\mathrm{d}\tau$, we shall introduce the metric tensor g_{ij} of a specially chosen Finsler space (Asanov, 1985). Actually, the introduction of this metric tensor does not play a decisive role here. It would be possible to avoid it. However, it makes our treatment more transparent. The square of the distance $\mathrm{d}s^2$ between two adjacent points in a Finsler space is given by the relation $\mathrm{d}s^2 = g_{ij}\mathrm{d}x^i\mathrm{d}x^j$. The distances may have various physical meanings. We shall consider the Finsler space in which distance s is measured by travel time τ, so that

$$\mathrm{d}\tau^2 = g_{ij}\mathrm{d}x^i\mathrm{d}x^j \; . \tag{9}$$

We shall refer to the relevant metric tensor g_{ij} as the *Finslerian metric tensor*, or the *wave propagation metric tensor*. Fermat's functional (8) then reads

$$\mathcal{I}(l) = \int_S^R [g_{ij}x^{i\prime}x^{j\prime}]^{1/2}\mathrm{d}u \; , \qquad x^{i\prime} = \mathrm{d}x^i/\mathrm{d}u \; . \tag{10}$$

Here u is an arbitrary parameter which specifies the position of a point on curve l.

Several notes on the Finslerian metric tensor g_{ij}. It is a function both of position x^k and of the direction of vector $x^{k\prime}$, tangential to curve l at x^k,

$$g_{ij} = g_{ij}(x^k, x^{k\prime}) \; . \tag{11}$$

This is a great difference with respect to the Riemanian metric tensor $g_{ij}(x^k)$, and reflects its directional dependence of velocities in anisotropic media. The Finslerian metric tensor $g_{ij}(x^k, x^{k\prime})$ reduces to the Riemanian metric tensor $g_{ij}(x^k)$, if g_{ij} is independent of direction. The so-called *Cartan torsion tensor* C_{ijl} can been introduced,

$$C_{ijl}(x^k, x^{k\prime}) = \tfrac{1}{2}\frac{\partial g_{ij}(x^k, x^{k\prime})}{\partial x^{l\prime}} \; . \tag{12}$$

In Riemanian geometry, the Cartan torsion tensor C_{ijl} vanishes identically, $C_{ijl}(x^k, x^{k\prime}) = 0$. In Finslerian geometry, however, it is generally different from zero. For more details, see Asanov (1985).

We shall assume that the Finslerian metric tensor $g_{ij}(x^k, x^{k\prime})$ satisfies the following properties:

1) $g_{ij}(x^k, x^{k\prime})$ is symmetric,

$$g_{ij}(x^k, x^{k\prime}) = g_{ji}(x^k, x^{k\prime}) \; . \tag{13}$$

2) $g_{ij}(x^k, x^{k\prime})$ is a homogeneous function of degree zero in $x^{k\prime}$, so that

$$g_{ij}(x^k, ax^{k\prime}) = a^0 g_{ij}(x^k, x^{k\prime}) = g_{ij}(x^k, x^{k\prime}) \; , \tag{14}$$

see (3). Thus, $g_{ij}(x^k, x^{k\prime})$ is not changed if the vector with the contravariant components $x^{k\prime}$ is replaced by a vector with a different length, but the same direction. This implies that the Finslerian metric tensor $g_{ij}(x^k, x^{k\prime})$ is a function of the position and the direction. The Euler's theorem (4) also yields

$$x^{n\prime}\partial g_{ij}(x^k, x^{k\prime})/\partial x^{n\prime} = 0 \; . \tag{15}$$

3) $g_{ij}(x^k, x^{k\prime})$ is nondegenerate,

$$\det(g_{ij}(x^k, x^{k\prime})) \neq 0 \ . \tag{16}$$

4) $g_{ij}(x^k, x^{k\prime})$ satisfies the following relation,

$$\frac{\partial g_{ij}(x^k, x^{k\prime})}{\partial x^{n\prime}} = \frac{\partial g_{in}(x^k, x^{k\prime})}{\partial x^{j\prime}} \ . \tag{17}$$

For a detailed discussion of (17) see Asanov (1985, p. 266).

In mathematical works, devoted to Finsler geometry, it has often been also assumed that g_{ij} is positive definite. Such an assumption would have serious consequences for our studies of anisotropic elastic media, and would exclude many important anisotropic media from our treatment. Fortunately, the Finsler geometry can be developed even for indefinite g_{ij} (indefinite Finsler spaces). See Asanov (1985, p. 21), and other references given therein (for example, Beem 1970, Beem and Kishta 1974).

Actually, even the Riemanian metric tensor may be used in some simple cases of anisotropic media, like ellipsoidally anisotropic media, but not in a general case. In Section 3.2, we shall explain how the Finslerian metric tensor can be determined from travel time measurements.

As the simplest example of wave propagation metric tensor g_{ij}, let us consider an isotropic medium and Cartesian coordinates x^i. In this case g_{ij} does not depend on $x^{k\prime}$,

$$g_{ij} = V^{-2}(x^k)\delta_{ij} \ , \qquad g_{ij}\mathrm{d}x^i\mathrm{d}x^j = V^{-2}(x^k)[(x^{1\prime})^2 + (x^{2\prime})^2 + (x^{3\prime})^2]\mathrm{d}\tau^2 \ . \tag{18}$$

Here V is the velocity of propagation of the wave under consideration. Fermat's functional (10) is then given by the familiar form,

$$\mathcal{I}(l) = \int_S^R V^{-1}(x^k)[(x^{1\prime})^2 + (x^{2\prime})^2 + (x^{3\prime})^2]^{1/2}\mathrm{d}u \ , \quad x^{i\prime} = \mathrm{d}x^i/\mathrm{d}u \ . \tag{19}$$

We shall now use the notation

$$\mathcal{L}(x^k, x^{k\prime}) = [g_{ij}(x^k, x^{k\prime})x^{i\prime}x^{j\prime}]^{1/2} \ , \qquad x^{i\prime} = \mathrm{d}x^i/\mathrm{d}u \ , \tag{20}$$

and call $\mathcal{L}(x^k, x^{k\prime})$ the Lagrangian. Fermat's functional $\mathcal{I}(l)$, given by (10), then reads

$$\mathcal{I}(l) = \int_S^R \mathcal{L}(x^k, x^{k\prime})\mathrm{d}u \ . \tag{21}$$

The statement that functional $\mathcal{I}(l)$ is stationary along curve $\bar{l}$ means that the first variation of the functional vanishes along $\bar{l}$:

$$\delta\mathcal{I}(l) = \delta\int_S^R \mathrm{d}\tau = \delta\int_{u(S)}^{u(R)} \mathcal{L}(x^k, x^{k\prime})\mathrm{d}u = 0 \ , \quad x^{k\prime} = \mathrm{d}x^k/\mathrm{d}u \ . \tag{22}$$

Here $u(S)$ and $u(R)$ are the values of the parameter u corresponding to fixed points S and R, respectively. Curve $\bar{l}$, for which Fermat's functional is stationary, is called the extremal curve, or the *extremal of Fermat's functional*. The ray may then be identified as the extremal of Fermat's functional. It also represents the geodesic curve connecting points S and R in a Finsler space specified by the wave propagation metric tensor g_{ij} (11). Consequently, the ray in an anisotropic inhomogeneous medium is also represented by a Finslerian geodesic curve.

It is well known that the extremal of Fermat's functional satisfies the Euler-Lagrange equations:

$$\frac{\mathrm{d}}{\mathrm{d}u}\Big(\frac{\partial\mathcal{L}}{\partial x^{i\prime}}\Big) - \frac{\partial\mathcal{L}}{\partial x^i} = 0 \ , \qquad i = 1,2,3 \ . \tag{23}$$

The Euler-Lagrange equations (23) represent the ray equations in Lagrangian form, also called Lagrangian ray equations.

The Lagrangian ray equations (23) consist of three ordinary differential equations of the second order. It would be useful to transform (23) into a system of six ordinary differential equations of the first order. The well-known procedure of doing this is based on the so-called dual Legendre transform. However, some restrictions are imposed on this transform. For this reason, we shall briefly recapitulate it here. For more details, see Babich and Buldyrev (1972), Brdička and Hladík (1987), Goldstein (1980), Kline and Kay (1965), etc.

We introduce "momenta" p_i by the relation

$$p_i = \partial\mathcal{L}/\partial x^{i\prime} \ , \qquad i = 1,2,3 \ . \tag{24}$$

The most important step in the dual Legendre transform consists in the determination of $x^{i\prime}$ $(i = 1,2,3)$ from known p_i $(i = 1,2,3)$, using (24). This is possible only if the relevant Hessian determinant is different from zero,

$$\det\Big(\frac{\partial^2\mathcal{L}}{\partial x^{i\prime}\partial x^{j\prime}}\Big) \neq 0 \ . \tag{25}$$

If (25) is satisfied, we can solve (24) for $x^{i\prime}$ and determine $x^{i\prime} = f^i(x^k, p_k)$. We can then determine the Hamiltonian $\mathcal{H}(x^k, p_k)$ by relation

$$\mathcal{H}(x^k, p_k) = [x^{i\prime}p_i - \mathcal{L}(x^k, x^{k\prime})]_{x^{i\prime}=f^i(x^k,p_k)} \ . \tag{26}$$

Differential $\mathrm{d}\mathcal{L}$ can be expressed in two alternative ways:

$$\mathrm{d}\mathcal{L} \ = \ \frac{\partial\mathcal{L}}{\partial x^i}\mathrm{d}x^i + \frac{\partial\mathcal{L}}{\partial x^{i\prime}}\mathrm{d}x^{i\prime} = -\frac{\partial\mathcal{H}}{\partial x^i}\mathrm{d}x^i - \frac{\partial\mathcal{H}}{\partial p_i}\mathrm{d}p_i + x^{i\prime}\mathrm{d}p_i + p_i\mathrm{d}x^{i\prime} \ . \tag{27}$$

In the second relation of (27), we have used $\mathcal{L} = x^{i\prime}p_i - \mathcal{H}(x^k, p_k)$. Then (27) yields two important relations

$$\frac{\partial\mathcal{L}}{\partial x^i} = -\frac{\partial\mathcal{H}}{\partial x^i} \ , \quad x^{i\prime} = \frac{\partial\mathcal{H}}{\partial p_i} \ . \tag{28}$$

Equations (23), (24) and (28) then yield the final Hamiltonian canonical system of six ordinary differential equations of the first order, alternative to the Euler-Lagrange equations (23),

$$\frac{\mathrm{d}x^i}{\mathrm{d}u} = \frac{\partial\mathcal{H}}{\partial p_i} \ , \qquad \frac{\mathrm{d}p_i}{\mathrm{d}u} = -\frac{\partial\mathcal{H}}{\partial x^i} \ . \tag{29}$$

The system of equations (29) represents the ray tracing system in Hamiltonian form, or the Hamiltonian ray equations.

Thus, if Hessian (25) is different from zero, we can construct Hamiltonian $\mathcal{H}(x^k, p_k)$ from Lagrangian $\mathcal{L}(x^k, x^{k\prime})$, and vice versa. Condition (25), however, often causes serious difficulties. We shall now explain the main problem, connected with Lagrangians represented by homogeneous functions of the first degree in $x^{i\prime}$.

In seismology, Fermat's principle has mostly been used in parameteric form (22). It is required that the Euler-Lagrange equations (23) hold for any parameter u. For a parameter σ, different from u, (22) reads

$$\delta\mathcal{I}(l) = \delta \int_S^R \mathcal{L}(x^k, x^{k\prime}) \mathrm{d}\sigma = 0 \ , \quad x^{k\prime} = \mathrm{d}x^k/\mathrm{d}\sigma \ .$$

Substituting $\mathrm{d}u = (\mathrm{d}u/\mathrm{d}\sigma)\mathrm{d}\sigma$ into (22) yields

$$\delta\mathcal{I}(l) = \delta \int_S^R \mathcal{L}(x^k, a^{-1}x^{k\prime}) a \mathrm{d}\sigma \ , \qquad a = \mathrm{d}u/\mathrm{d}\sigma \ , \qquad x^{k\prime} = \mathrm{d}x^k/\mathrm{d}\sigma \ .$$

Both expressions for $\delta\mathcal{I}(l)$ are the same if $a\mathcal{L}(x^k, a^{-1}x^{k\prime}) = \mathcal{L}(x^k, x^{k\prime})$, that is, if $\mathcal{L}(x^k, x^{k\prime})$ is a homogeneous function of the first degree in $x^{i\prime}$. Thus (22) and (23) may be used for arbitrarily chosen parameter u only if $\mathcal{L}(x^k, x^{k\prime})$ is a homogeneous function of the first degree in $x^{k\prime}$.

It is, however, possible to prove that the Hessian determinant (25) vanishes identically if $\mathcal{L}(x^k, x^{k\prime})$ is a homogeneous function of the first degree in $x^{i\prime}$. Function $\partial\mathcal{L}/\partial x^{j\prime}$ is then a homogeneous function of the zeroth degree in $x^{j\prime}$, and the Euler theorem yields

$$x^{i\prime} \frac{\partial^2 \mathcal{L}(x^k, x^{k\prime})}{\partial x^{i\prime} \partial x^{j\prime}} = x^{i\prime} \frac{\partial}{\partial x^{i\prime}} \left(\frac{\partial \mathcal{L}(x^k, x^{k\prime})}{\partial x^{j\prime}} \right) = 0 \ .$$

Thus, for any fixed j, vector $\partial^2\mathcal{L}/\partial x^{i\prime}\partial x^{j\prime}$ is perpendicular to vector $x^{i\prime}$. This shows that all the three column vectors in the Hessian determinant are situated in a plane. Consequently, they are linearly dependent and the Hessian determinant vanishes.

The final result is that the dual Legendre transformation cannot be used to compute the Hamiltonian from the Lagrangian, if Lagrangian $\mathcal{L}(x^k, x^{k\prime})$ is a homogeneous function of the first degree in $x^{i\prime}$. Consequently, the Hamiltonian ray tracing equations cannot be obtained from the Lagrangian ray tracing equations by the application of the standard dual Legendre transform.

We shall show a simple consequence of the Lagragian being a homogeneous function of the first degree in $x^{i\prime}$. In this case, we obtain $x^{i\prime}p_i = x^{i\prime}(\partial\mathcal{L}/\partial x^{i\prime}) = \mathcal{L}(x^k, x^{k\prime})$, so that

$$x^{i\prime}p_i - \mathcal{L}(x^k, x^{k\prime}) = \mathcal{L}(x^k, x^{k\prime}) - \mathcal{L}(x^k, x^{k\prime}) = 0 \ .$$

Equation (26) then yields

$$\mathcal{H}(x^k, p_k) = 0 \ . \tag{30}$$

Thus, we have obtained the interesting result that $\mathcal{H}(x^k, p_k) = 0$ for Lagrangians represented by homogeneous function of the first degree in $x^{i\prime}$. We are, however, not able to derive the actual form of $\mathcal{H}(x^k, p_k)$.

Nevertheless, it is possible to modify the Lagrangian into a new form, for which the Euler-Lagrange equations are still satisfied, but the Hessian determinant does not identically vanish. In this contribution, we shall describe two such possibilities.

1) We can use the variable $u = \tau$ along the curve, where τ represents the travel time along the curve. We can then introduce the modified Lagrangian $\mathcal{L}^M(x^k, \dot{x}^k)$ by the relation

$$\mathcal{L}^M(x^k, \dot{x}^k) = \tfrac{1}{2}[\mathcal{L}^2(x^k, \dot{x}^k)]^2 \ , \qquad \dot{x}^i = \mathrm{d}x^i/\mathrm{d}\tau \ .$$

We emphasize that $\dot{x}^i$ is used to denote $\mathrm{d}x^i/\mathrm{d}\tau$ to distinguish it from $x^{i\prime} = \mathrm{d}x^i/\mathrm{d}u$, valid for any parameter u. For a detailed treatment, see Section 3.

2) As the variable u along the ray, we can choose one of the coordinates x^i, for example, coordinate x^3. We shall refer to the relevant Lagrangian as the reduced Lagrangian, and denote it by $\mathcal{L}^R$. See Section 4.

In both these cases, the Hessian determinant is not identically zero, and the dual Legendre transform can be used. The Hamiltonian can be calculated from the Lagrangian, and the ray tracing system in Lagrangian form can be transformed to the Hamiltonian ray tracing system, and vice versa.

The statement that the Hessian is not identically equal zero does not mean that the Hessian cannot vanish for some specific directions. Actually, such situations may appear when the Hamiltonian (and/or the Lagrangian) do not represent globally convex surfaces. More details will be given later. We shall also discuss some simpler situations later, in which all transformations may be performed analytically, and show what the relation is of the Lagrangians under consideration to the group velocity surfaces.

3 Modified Lagrangians and Hamiltonians

We shall consider travel time τ to be a variable along the curves connecting S and R. Fermat's variational principle (22) then reads

$$\delta \int_S^R \mathcal{L}(x^k, \dot{x}^k)\mathrm{d}\tau = 0 \ , \qquad \dot{x}^k = \mathrm{d}x^k/\mathrm{d}\tau \ , \tag{31}$$

and the relevant Euler-Lagrange equations are

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial \mathcal{L}}{\partial \dot{x}^i}\right) - \frac{\partial \mathcal{L}}{\partial x^i} = 0 \ , \qquad i = 1, 2, 3 \ . \tag{32}$$

Lagrangian $\mathcal{L}(x^k, \dot{x}^k)$ is given by the relation

$$\mathcal{L}(x^k, \dot{x}^k) = [g_{ij}(x^k, \dot{x}^k)\dot{x}^i\dot{x}^j]^{1/2} \ . \tag{33}$$

It immediately follows from (9) that

$$\mathcal{L}(x^k, \dot{x}^k) = 1 \tag{34}$$

along the ray. As the Finslerian metric tensor $g_{ij}(x^k, \dot{x}^k)$ is a homogeneous function of degree zero in $\dot{x}^k$, Lagrangian $\mathcal{L}(x^k, \dot{x}^k)$ is a homogeneous function of the first degree in $\dot{x}^i$, so that the relevant Hessian determinant $\det(\partial^2\mathcal{L}/\partial\dot{x}^i\partial\dot{x}^j)$ vanishes. Consequently, the dual Legendre transform cannot be used to derive Hamiltonian canonical equations.

3.1 Modified Lagrangian $\mathcal{L}^M$. Derivation of Hamiltonian canonical equations

We now introduce the modified Lagrangian $\mathcal{L}^M(x^k, \dot{x}^k)$ by the relation

$$\mathcal{L}^M(x^k, \dot{x}^k) = \tfrac{1}{2}[\mathcal{L}(x^k, \dot{x}^k)]^2 = \tfrac{1}{2} g_{ij}(x^k, \dot{x}^k)\dot{x}^i \dot{x}^j \ . \tag{35}$$

Consequently, the modified Lagrangian $\mathcal{L}^M(x^k, \dot{x}^k)$ is a homogeneous function of the second degree in $\dot{x}^k$. It follows from (34) that

$$\mathcal{L}^M(x^k, \dot{x}^k) = \tfrac{1}{2} \ . \tag{36}$$

Equations (35) and (17) also immediately yield an important expression for g_{ij} in terms of $\mathcal{L}^M$,

$$g_{ij}(x^k, \dot{x}^k) = \frac{\partial^2 \mathcal{L}^M(x^k, \dot{x}^k)}{\partial \dot{x}^i \partial \dot{x}^j} \ , \tag{37}$$

and the relation of g_{ij} to the Hessian determinant:

$$\det\left(\frac{\partial^2 \mathcal{L}^M}{\partial \dot{x}^i \partial \dot{x}^j}\right) = \det(g_{ij}) \ . \tag{38}$$

We shall now evaluate $\frac{\mathrm{d}}{\mathrm{d}\tau}(\partial\mathcal{L}^M/\partial\dot{x}^i) - \partial\mathcal{L}^M/\partial x^i$. We obtain

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial\mathcal{L}^M}{\partial\dot{x}^i}\right) - \frac{\partial\mathcal{L}^M}{\partial x^i} = 2\left[\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\mathcal{L}\frac{\partial\mathcal{L}}{\partial\dot{x}^i}\right) - \mathcal{L}\,\frac{\partial\mathcal{L}}{\partial x^i}\right] = 2\left[\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial\mathcal{L}}{\partial\dot{x}^i}\right) - \frac{\partial\mathcal{L}}{\partial x^i}\right] = 0 \ ,$$

as $\mathcal{L}(x^k, \dot{x}^k) = 1$ along the ray, see (34) and (32). Thus, the modified Lagrangian $\mathcal{L}^M(x^k, \dot{x}^k)$ satisfies the Euler-Lagrange equations,

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\partial\mathcal{L}^M}{\partial\dot{x}^i}\right) - \frac{\partial\mathcal{L}^M}{\partial x^i} = 0 \ , \tag{39}$$

corresponding to the functional

$$\mathcal{I}(l) = 2\int_S^R \mathcal{L}^M(x^k, \dot{x}^k)\mathrm{d}\tau = \int_S^R g_{ij}(x^k, \dot{x}^k)\dot{x}^i\dot{x}^j \mathrm{d}\tau \ . \tag{40}$$

We have introduced the factor 2 before the integral, as we wish to keep the same variable τ in (40) as in (31), corresponding to the travel time, see (36). This factor does not influence Fermat's variational principle and the consequent Euler-Lagrange equations at all. Consequently, variational principle

$$\delta\int_S^R 2\mathcal{L}^M(x^k, \dot{x}^k)\mathrm{d}\tau = \delta\int_S^R g_{ij}(x^k, \dot{x}^k)\dot{x}^i\dot{x}^j \mathrm{d}\tau = 0 \tag{41}$$

yields the same extremals as Fermat's variational principle (31).

The modified Lagrangian $\mathcal{L}^M(x^k, \dot{x}^k)$ is not a homogeneous function of the first degree in $\dot{x}^i$ as $\mathcal{L}(x^k, \dot{x}^k)$, but a homogeneous function of the second degree in $\dot{x}^i$. In this case, the Hessian determinant (38) does not vanish identically, and the Legendre dual transform may be used to construct the Hamiltonian. See also (16).

We introduce momenta p_i and Hamiltonian $\mathcal{H}^M(x^k, p_k)$ by relations

$$p_i = \frac{\partial\mathcal{L}^M}{\partial\dot{x}^i} \ , \quad \mathcal{H}^M(x^k, p_k) = [\dot{x}^i p_i - \mathcal{L}^M(x^k, \dot{x}^k)]_{\dot{x}^i = f^i(x^k, p_k)} \ . \tag{42}$$

As the Hessian determinant for $\mathcal{L}^M$ does not vanish, the first equation of (42) can be inverted to yield $\dot{x}^i = f^i(x^k, p_k)$. In the following, we shall refer to $\mathcal{H}^M(x^k, p_k)$ as the *modified Hamiltonian.* Using also (28),

$$\frac{\partial \mathcal{L}^M}{\partial x^i} = -\frac{\partial \mathcal{H}^M}{\partial x^i} \ , \quad \dot{x}^i = \frac{\partial \mathcal{H}^M}{\partial p_i} \ , \tag{43}$$

we arrive at the Hamiltonian ray equations in the following form,

$$\frac{\mathrm{d}x^i}{\mathrm{d}\tau} = \frac{\partial \mathcal{H}^M}{\partial p_i} \ , \quad \frac{\mathrm{d}p_i}{\mathrm{d}\tau} = -\frac{\partial \mathcal{H}^M}{\partial x^i} \ . \tag{44}$$

This is the final result. Later on, we shall show that the modified Hamiltonian $\mathcal{H}^M(x^k, p_k)$, obtained in this way, corresponds to the Hamiltonian defined by (6). See Section 3.2.

3.2 Discussion of the modified Lagrangian $\mathcal{L}^M$ and relevant modified Hamiltonian $\mathcal{H}^M$

We shall now discuss the modified Lagrangian $\mathcal{L}^M(x^k, x^{k\prime})$ and the modified Hamiltonian $\mathcal{H}^M(x^k, p_k)$ in a greater detail. We shall also add several notes to clarify the meaning and the properties of certain quantities introduced above. We shall try to use the terminology common in seismic ray theory.

1) Momenta p_i represent the covariant components of the slowness vector. It should be noted that p_i computed from the modified Lagrangian $\mathcal{L}^M$ using (42) are exactly the same as the p_i computed from the parameteric Lagrangian $\mathcal{L}$ using (24) for $u = \tau$, $p_i = \partial \mathcal{L}^M / \partial \dot{x}^i = \mathcal{L} \partial \mathcal{L} / \partial \dot{x}^i = \partial \mathcal{L} / \partial x^i$. Similarly, quantities $\dot{x}^i = \mathrm{d}x^i / \mathrm{d}\tau$ represent the contravariant components of group velocity vector $\vec{\mathcal{U}}$,

$$\dot{x}^i = \mathrm{d}x^i / \mathrm{d}\tau = \mathcal{U}^i \ . \tag{45}$$

Multiplying the first equation of (42) by $\dot{x}^i$ and using Euler's theorem, we obtain

$$\dot{x}^i p_i = \dot{x}^i \frac{\partial \mathcal{L}^M}{\partial \dot{x}^i} = 2\mathcal{L}^M = 1 \ .$$

This yields the well-known relation connecting p_i and $\mathcal{U}^i$,

$$p_i \mathcal{U}^i = 1 \ . \tag{46}$$

2) The modified Lagrangian $\mathcal{L}^M(x^k, \dot{x}^k)$ can be expressed in terms of group velocity $\mathcal{U}$. If we use (36) and Euler's theorem (4), we obtain

$$\mathcal{L}^M(x^k, \mathcal{U}^k) = \mathcal{U}^2 \mathcal{L}^M(x^k, \mathcal{U}^k / \mathcal{U}) = \tfrac{1}{2} \ .$$

We now take into account that $\mathcal{U}^k / \mathcal{U} = \mathrm{d}x^k / \mathrm{d}s$, where s is the arclength along the ray. This yields

$$\mathcal{U}^2 = \frac{1}{2\mathcal{L}^M(x^k, \mathrm{d}x^k/\mathrm{d}s)} = \frac{1}{\mathcal{L}^2(x^k, \mathrm{d}x^k/\mathrm{d}s)} \ , \quad \mathcal{U} = \frac{1}{\mathcal{L}(x^k, \mathrm{d}x^k/\mathrm{d}s)} \ . \tag{47}$$

3) We shall now show that surface $\mathcal{L}^M(x^k,\mathcal{U}^k) = \frac{1}{2}$ in the $\mathcal{U}^i$-space represents the group velocity surface for a fixed x^k. We define the group velocity surface as the locus of the end points of group velocity vectors $\vec{\mathcal{U}}$, constructed in all possible directions at point x^i. Consequently, the group velocity surface represents a polar graph of group velocity $\mathcal{U}$ as a function of two take-off angles that specify the direction of unit vector $\vec{t} = \vec{\mathcal{U}}/\mathcal{U}$. We use the relation

$$\mathcal{U}^2 \mathcal{L}^M(x^k, t^k) = \tfrac{1}{2} \ , \tag{48}$$

where $t^k = \mathrm{d}x^k/\mathrm{d}s$ are components of the unit vector tangent to the ray. We further use the polar representation for $\mathcal{U}^k$, $\mathcal{U}^k = \rho t^k$. Then $\mathcal{L}^M(x^k, t^k) = \rho^{-2}\mathcal{L}^M(x^k,\mathcal{U}^k)$, and

$$(\mathcal{U}/\rho)^2 \mathcal{L}^M(x^k,\mathcal{U}^k) = \tfrac{1}{2} \ . \tag{49}$$

This shows that $\mathcal{U} = \rho$ along the surface $\mathcal{L}^M(x^k,\mathcal{U}^k) = \frac{1}{2}$. Consequently, the surface $\mathcal{L}^M(x^k,\mathcal{U}^k) = \frac{1}{2}$ in the $\mathcal{U}^i$-space represents the group velocity surface for a fixed x^k.

In the same way, it can be simply proved that an alternative expression for the group velocity surface is $\mathcal{L}(x^k,\mathcal{U}^k) = 1$. Thus, we can alternatively express the group velocity surface by any of the two equations

$$\mathcal{L}^M(x^k,\mathcal{U}^k) = \tfrac{1}{2} \ , \qquad \mathcal{L}(x^k,\mathcal{U}^k) = 1 \ . \tag{50}$$

4) Similar argumentation may be used to prove that equation $\mathcal{H}^M(x^k, p_k) = \frac{1}{2}$ represents the slowness surface in the p_i-phase space for a fixed x^i. By the slowness surface we understand the locus of the end points of slowness vector $\vec{p}$ constructed in all directions at point x^i . Often, the eigenvalues $G(x^i, p_i)$ of the Christoffel matrix are used in the seismic ray theory to specify the slowness surface. It is related to the modified Hamiltonian as follows:

$$\mathcal{H}^M(x^k, p_k) = \tfrac{1}{2} G(x^k, p_k) \ . \tag{51}$$

Both $\mathcal{H}^M(x^k, p_k)$ and $G(x^k, p_k)$ are homogeneous functions of the second degree in p_i, and the slowness surface is given by two alternative equations:

$$\mathcal{H}^M(x^k, p_k) = \tfrac{1}{2} \ , \qquad G(x^k, p_k) = 1 \ . \tag{52}$$

Thus, function $G(x^k, p_k)$ represents the eigenvalue of the Christoffel matrix Γ_{ij}, see (1) and (6). However, to obtain the actual form of slowness surface $G(x^k, p_k) = 1$ from the known group velocity surface $\mathcal{L}^M(x^k, \dot{x}^k) = \frac{1}{2}$, it would be necessary to solve equation $p_i = \partial\mathcal{L}^M/\partial\dot{x}^i$ for $\dot{x}^i = f^i(x^k, p_k)$, and insert it into (42). Also a graphical construction can be used, as both these surfaces are polarly reciprocal (see Helbig, 1994).

5) We shall now introduce the contravariant components of wave propagation metric tensor $\overline{g}^{ij}$ in the x^k, p_i domain. To distinguish it from the expressions for the metric tensor in the $x^k, \dot{x}^i$ domain, we shall use a bar above the letter. We define $\overline{g}^{ij}$ as

$$\overline{g}^{mn}(x^k, p_k) = \tfrac{1}{2} \, \frac{\partial^2 G}{\partial p_m \partial p_n} = \frac{\partial^2 \mathcal{H}^M}{\partial p_m \partial p_n} \ . \tag{53}$$

Using the first equation of the Hamilton canonical equations (44) and (51), we obtain

$$\dot{x}^i = \mathcal{U}^i = \tfrac{1}{2} \, \frac{\partial G}{\partial p_i} = \tfrac{1}{2} p_m \, \frac{\partial^2 G}{\partial p_m \partial p_i} = p_m \overline{g}^{mi}(x^k, p_k) \ . \tag{54}$$

The contravariant components of metric tensor $\overline{g}^{mi}$ are related to its covariant components $\overline{g}_{ij}$ as

$$\overline{g}_{ik}\overline{g}^{ij} = \delta^j_k \ , \tag{55}$$

where δ^j_k is the mixed co- and contravariant Kronecker delta symbol. It equals 1 for $j = k$, and 0 for $j \neq k$. Equation (54) then yields

$$p_m = \overline{g}_{mk}(x^i, p_i)\dot{x}^k \ . \tag{56}$$

We shall now use the relation

$$p_i p_m \frac{\partial^2 \mathcal{H}^M}{\partial p_i \partial p_m} = p_i p_m \overline{g}^{im}(x^k, p_k) = \overline{g}_{ik}\dot{x}^k \overline{g}_{ml}\dot{x}^l \overline{g}^{im} = \dot{x}^m \dot{x}^l \overline{g}_{ml}(x^k, p_k) \ ,$$

and obtain

$$p_i p_m \overline{g}^{im}(x^k, p_k) = \dot{x}^m \dot{x}^l \overline{g}_{ml}(x^k, p_k) = 1 \ . \tag{57}$$

Relation (57) is satisfied along the ray. Thus, $\overline{g}_{ml}(x^k, p_k)$ is the p_i-domain alternative of the wave propagation metric tensor $g_{ml}(x^k, \dot{x}^k)$, given in (9).

A similar expression for the contravariant components of the wave propagation metric tensor was derived earlier by Hanyga, Lenartowicz and Pajchel (1984), and by Klimeš (1994). Hanyga, Lenartowicz and Pajchel (1984) even refer to an older reference Rund (1966). Wave propagation metric tensor $\overline{g}^{mn}$ is also closely related to the metric tensor of the general theory of relativity.

6) The first equation of (42) can also be expressed as follows, see (37),

$$p_i = \frac{\partial \mathcal{L}^M}{\partial \dot{x}^i} = \dot{x}^j \frac{\partial^2 \mathcal{L}^M}{\partial \dot{x}^i \partial \dot{x}^j} = g_{ij}(x^k, \dot{x}^k)\dot{x}^j \ , \tag{58}$$

$\mathcal{L}^M(x^k, \dot{x}^k)$ being a homogeneous function of the second degree in $\dot{x}^i$. This also yields

$$\dot{x}^i = g^{ij}(x^k, \dot{x}^k)p_j \ . \tag{59}$$

Here $g^{ij}(x^k, \dot{x}^k)$ are the contravariant components of the wave propagation metric tensor, and can be determined by inverting $g_{ij}(x^k, \dot{x}^k)$.

We now take into account the transformation relations between the covariant y_i and contravariant y^i components of vector $\vec{y}$, in coordinates specified by the Finslerian metric tensor $g_{ij}(x_k, \dot{x}_k)$,

$$y_i = g_{ij}y^j \ , \qquad y^i = g^{ij}y_j \ .$$

Equations (58) and (59) then indicate that the covariant components p_i of the slowness vector and the contravariant components $\dot{x}^i$ of the group velocity vector represent the components of the same vector in the Finslerian space.

7) The comparison of Equation (54) with (59) and of (56) with (58) shows that

$$\overline{g}_{mn} = g_{mn} \ , \qquad \overline{g}^{mn} = g^{mn} \ . \tag{60}$$

The only difference is that $\overline{g}_{mn}$ and $\overline{g}^{mn}$ are expressed in terms of p_i, but g_{mn} and g^{mn} in terms of $\dot{x}^i$. Relation (55) can then be expressed in the following form: $g_{ij}\overline{g}^{in} = \delta^n_j$. This yields

$$\frac{\partial^2 \mathcal{L}^M(x^k, \dot{x}^k)}{\partial \dot{x}^i \partial \dot{x}^j} \, \frac{\partial^2 \mathcal{H}^M(x^k, p_k)}{\partial p_i \partial p_n} = \delta^n_i \ . \tag{61}$$

Using (33), (35) and (57), we can also express Lagrangians $\mathcal{L}(x^k, \dot{x}^k)$ and $\mathcal{L}^M(x^k, \dot{x}^k)$ in the p_i-domain as follows:

$$\mathcal{L}(x^k, \dot{x}^k) = [\overline{g}^{im}(x^k, p_k)p_i p_m]^{1/2} \ , \quad \mathcal{L}^M(x^k, \dot{x}^k) = \tfrac{1}{2}\overline{g}^{im}(x^k, p_k)p_i p_m \ . \tag{62}$$

As $\overline{g}^{im}$ is given by (53), we further obtain, in the p_i domain,

$$\begin{aligned} \mathcal{L}(x^k, \dot{x}^k) &= [2\mathcal{H}^M(x^k, p_k)]^{1/2} = [G(x^k, p_k)]^{1/2} \ , \\ \mathcal{L}^M(x^k, \dot{x}^k) &= \mathcal{H}^M(x^k, p_k) = \tfrac{1}{2}G(x^k, p_k) \ . \end{aligned} \tag{63}$$

This clearly reflects the relations between the Lagrangians in the $\dot{x}^i$-domain and the Hamiltonians in the p_i-domain.

8) Using (61), we obtain

$$\det(\partial^2 \mathcal{L}^M / \partial \dot{x}^i \partial \dot{x}^j) = 1/\det(\partial^2 \mathcal{H}^M / \partial p_i \partial p_j) \ . \tag{64}$$

Thus, the anomalous behaviour of the modified Hamiltonian $\mathcal{H}^M$ for some direction specified by slowness vector p_i corresponds to the anomalous behaviour of the modified Lagrangian $\mathcal{L}^M$, for the direction of the relevant group velocity vector. Well-known examples are the parabolic points on the slowness surface (where $\det(\partial^2 \mathcal{H}^M / \partial p_i \partial p_j) = 0$) and the corresponding cuspoidal points of the group velocity surface (where $\det(\partial^2 \mathcal{L}^M / \partial \dot{x}^i \partial \dot{x}^j) = \infty$).

9) Lagrangian $\mathcal{L}^M$ may be multivalued for certain directions of the group velocity vector, even if only one selected wave is considered. The exception is the qP wave. The slowness surface of the qP wave is always completely convex, and the relevant group velocity surface is always completely single-valued and convex. The slowness surfaces of qS waves, however, may be hyperbolic or concave for certain directions of the slowness vector. The explanation of this fact is well-known. Equation $\mathcal{U}^i = \partial \mathcal{H} / \partial p_i$ may yield the same direction of $\vec{\mathcal{U}}$ for several different points along the slowness surface, if the slowness surface is not fully concave. This yields the multivaluedness of the modified Lagrangian, the multivaluedness of the group velocity surface, and the multivaluedness of wave propagation metric tensor $g_{ij}(x^k, \dot{x}^k)$. In this case, typical loops appear on the group velocity surface.

10) Let us return to metric tensor $g_{ij}(x^k, \dot{x}^k)$. It is given by relation (37). Thus, metric tensor g_{ij} is related to the curvature of the group velocity surface and may be calculated from it. It is well known that the group velocity surface may be determined from travel-time measurements. This suggests that even wave propagation metric tensor $g_{ij}(x^k, \dot{x}^k)$ may be determined from the travel-time measurements. We only need to determine the second derivatives (37) of the group velocity surface. As the simplest example, let us consider an isotropic medium, and Cartesian coordinates x^i. The group velocity surface is then a sphere, given by the relation $\mathcal{L}^M = \frac{1}{2}V^{-2}(x^k)[(\dot{x}^1)^2 + (\dot{x}^2)^2 + (\dot{x}^3)^2] = \frac{1}{2}$. Equation (37) then yields $g_{ij} = V^{-2}(x^k)\delta_{ij}$. In this case, the metric tensor g_{ij} does not depend on $\dot{x}^k$.

11) We shall now write equations for travel times $T(R, S)$ along the ray connecting S and R. Using (21), (40) and (63), we obtain several alternative expressions in the x_i'- and p_i- domains:

$$T(R, S) = \int_S^R \mathcal{L}(x^k, \dot{x}^k) \mathrm{d}\tau = \int_S^R [G(x^k, p_k)]^{1/2} \mathrm{d}\tau \ . \tag{65}$$

Here $\mathcal{L} = [g_{mn}(x^k, \dot{x}^k)\dot{x}^m\dot{x}^n]^{1/2}$, $G = g^{mn}(x^k, p_k)p_m p_n$. See also Pšenčík (1997). Alternatively,

$$T(R,S) = 2\int_S^R \mathcal{L}^M(x^k, \dot{x}^k)\mathrm{d}\tau = 2\int_S^R \mathcal{H}^M(x^k, p_k)\mathrm{d}\tau = \int_S^R G(x^k, p_k)\mathrm{d}\tau \ , \qquad (66)$$

where $\mathcal{L}^M = \frac{1}{2}\mathcal{L}^2 = \frac{1}{2}g_{mn}(x^k, \dot{x}^k)\dot{x}^m\dot{x}^n$, $\mathcal{H}^M = \frac{1}{2}G = \frac{1}{2}g^{mn}(x^k, p_k)p_m p_n$. The first relation of (65) is very useful, as it can be used for any parameter u along the ray, see (21),

$$T(R,S) = \int_S^R \mathcal{L}(x^k, \mathrm{d}x^k/\mathrm{d}u)\mathrm{d}u \ . \qquad (67)$$

We remind the reader that the Hessian determinant corresponding to Lagrangian $\mathcal{L}$ vanishes, as $\mathcal{L}$ is a homogeneous function of the first degree in $\mathrm{d}x^k/\mathrm{d}u$. Consequently, the Legendre transform cannot be used to determine the associated Hamiltonian directly.

If we use the parameter u representing arclength s along the ray, we obtain from (67),

$$T(R,S) = \int_S^R \frac{\mathrm{d}s}{\mathcal{U}(x^k, \mathrm{d}x^k/\mathrm{d}s)} \ , \qquad (68)$$

where $\mathcal{U}$ is the group velocity, see (47). This relation has been often used in seismological literature and corresponds fully to simple physical considerations. It should be emphasized that the integrations in (65)–(68) are performed along a ray.

3.3 Isotropic medium

The ray tracing equations derived in Sections 3.1 and 3.2, and all other relations of these sections, are valid for general anisotropic inhomogeneous media and for arbitrary curvilinear coordinates. They may be simplified in many ways. If we consider an isotropic medium, the metric tensor g_{ij} does not depend on $\dot{x}^i$. In addition, we can also use orthogonal coordinates x^i, or even Cartesian coordinates. To make certain steps in the derivations and some equations in these sections more transparent and objective, we shall first apply the general procedure to isotropic inhomogeneous media, and to Cartesian coordinates. In such a case, all mathematical operations can be simply performed analytically. In the second part of this section, we shall return to general curvilinear coordinates.

We shall denote the Cartesian coordinates using standard lower-case indices and assume the summation over the same lower-case indices: $x^1 = x_1,\ x^2 = x_2,\ x^3 = x_3$. The covariant components of the wave propagation metric tensor are then given by relation (18),

$$g_{ij} = V^{-2}(x_k)\delta_{ij} \ . \qquad (69)$$

Lagrangian $\mathcal{L}(x_k, \dot{x}_k^{'})$ reads, see (20),

$$\mathcal{L}(x_k, x_k^{'}) = V^{-1}(x_k)(x_i^{'}x_i^{'})^{1/2} \ , \qquad x_i^{'} = \mathrm{d}x_i/\mathrm{d}u \ . \qquad (70)$$

The relevant Fermat's variational principle is as follows:

$$\delta\int_S^R V^{-1}(x_k)(x_i^{'}x_i^{'})^{1/2}\mathrm{d}u = 0 \ , \qquad x_i^{'} = \mathrm{d}x_i/\mathrm{d}u \ . \qquad (71)$$

and the Euler-Lagrange equations read

$$\frac{\mathrm{d}}{\mathrm{d}u}(V^{-1}(x_k)D^{-1}x_i^{'}) = D\frac{\partial}{\partial x_i}\Big(\frac{1}{V(x_k)}\Big) \ , \quad D = (x_k^{'}x_k^{'})^{1/2} \ . \qquad (72)$$

Note that $D = V$ for $u = \tau$, and the Euler-Lagrange equations read:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\Big(\frac{1}{V^2(x_k)}\dot{x}_i\Big) - V(x_k)\frac{\partial}{\partial x_i}\Big(\frac{1}{V(x_k)}\Big) = 0 \ , \quad \dot{x}_i = \mathrm{d}x_i/\mathrm{d}\tau \ . \tag{73}$$

We shall now calculate Hessian determinant (25). We obtain

$$\partial^2 \mathcal{L}/\partial x_i^{'}\partial x_j^{'} = (\delta_{ij} D^2 - x_i^{'} x_j^{'})/V D^3 \ ,$$

so that

$$\det(\partial^2 \mathcal{L}/\partial x_i^{'}\partial x_j^{'}) = V^{-3} D^{-9} \begin{pmatrix} x_2^{'2} + x_3^{'2} & x_1^{'} x_2^{'} & x_1^{'} x_3^{'} \\ x_1^{'} x_2^{'} & x_1^{'2} + x_3^{'2} & x_2^{'} x_3^{'} \\ x_1^{'} x_3^{'} & x_2^{'} x_3^{'} & x_1^{'2} + x_2^{'2} \end{pmatrix} = 0 \ . \tag{74}$$

As the Hessian determinant vanishes, the dual Legendre transform cannot be used to determine Hamiltonian $\mathcal{H}(x_i, p_i)$ from Lagrangian (70) and to write the canonical Hamiltonian equations.

For this reason, we shall introduce the modified Lagrangian. We use $u = \tau$, where τ is the travel time along the curves, and use (35)

$$\mathcal{L}^M(x_k, \dot{x}_k) = \tfrac{1}{2} V^{-2}(x_k)(\dot{x}_i \dot{x}_i) \ , \quad \dot{x}_i = \mathrm{d}x_i/\mathrm{d}\tau \ . \tag{75}$$

Modified Fermat's variational principle (41) then reads

$$\delta \int_S^R V^{-2}(x_k)(\dot{x}_i\dot{x}_i)\mathrm{d}\tau = 0 \ . \tag{76}$$

The corresponding Euler-Lagrange equations are

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\Big(\frac{2}{V^2(x_k)}\dot{x}_i\Big) - (\dot{x}_i\dot{x}_i)\frac{\partial}{\partial x_i}\Big(\frac{1}{V^2(x_k)}\Big) = 0 \ . \tag{77}$$

If we take into account $\dot{x}_i\dot{x}_i = V^2$, we can see that (77) is fully equivalent to (73). Thus, Fermat's variational principles (71) and (76) yield the same extremals, even though Fermat's variational principle in (76) is not constructed in the standard way of (71).

It is simple to determine the Hessian determinant for Lagrangian $\mathcal{L}^M$. From (75), we obtain

$$\frac{\partial^2 \mathcal{L}^M(x_k, \dot{x}_k)}{\partial \dot{x}_i \partial \dot{x}_j} = V^{-2}(x_k)\delta_{ij} \ ,$$

so that

$$\det\Big(\frac{\partial^2 \mathcal{L}^M(x_k, \dot{x}_k)}{\partial \dot{x}_i \partial \dot{x}_j}\Big) = V^{-6} \neq 0 \ . \tag{78}$$

Thus, the dual Legendre transform can be used to compute the relevant modified Hamiltonian and find the Hamiltonian canonical equations. Before we do this, we only note that the modified Lagrangian (75) is constant along the whole extremal and equals $\frac{1}{2}$,

$$\mathcal{L}^M(x_k, \dot{x}_k) = \tfrac{1}{2} \ , \tag{79}$$

see (36). This also immediately follows from (75), if we take into account that $\dot{x}_i\dot{x}_i = V^2$.

Using (42), we obtain

$$p_i = \frac{\partial \mathcal{L}^M}{\partial \dot{x}_i} = V^{-2}\dot{x}_i \ . \tag{80}$$

This can be simply inverted to yield

$$\dot{x}_i = V^2 p_i \ . \tag{81}$$

The modified Hamiltonian $\mathcal{H}^M(x_k, p_k)$ is then given by relation (42)

$$\mathcal{H}^M(x_k, p_k) = [\dot{x}_i p_i - \mathcal{L}^M(x_k, \dot{x}_k)]_{\dot{x}_i = V^2 p_i} = V^2 p_i p_i - \tfrac{1}{2} V^2 p_i p_i \ .$$

Thus,

$$\mathcal{H}^M(x_k, p_k) = \tfrac{1}{2} V^2(x_k) p_i p_i \ . \tag{82}$$

The modified Hamiltonian is constant along the whole extremal,

$$\mathcal{H}^M(x_k, p_k) = \tfrac{1}{2} \ ,$$

see (52). As we know, $V^2(x_k)p_i p_i$ is eigenvalue $G(x_k, p_k)$ of the Christoffel matrix for isotropic media, $G(x_k, p_k) = V^2(x_k)p_i p_i$. Consequently,

$$G(x_k, p_k) = V^2 p_i p_i = 1 \ . \tag{83}$$

along the whole extremal. This represents the eikonal equation for isotropic media.

The Hamiltonian ray tracing system (44) for isotropic media in Cartesian coordinates then reads

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

Using (83), we obtain $p_k p_k = 1/V^2$, so that

$$\frac{\mathrm{d}x_i}{\mathrm{d}\tau} = V^2 p_i \ , \quad \frac{\mathrm{d}p_i}{\mathrm{d}\tau} = -\frac{\partial \ln V}{\partial x_i} \ . \tag{84}$$

This corresponds fully to Červený (2001a, Equation (3.1.19)).

It is possible to see from (75) and (79) that group velocity $\mathcal{U} = |\vec{\mathcal{U}}| = (\dot{x}_i \dot{x}_i)^{1/2}$ equals V and does not depend on the direction of propagation. The group velocity surface is spherical, with radius V. Similarly, (83) shows that $|\vec{p}| = (p_i p_i)^{1/2} = 1/V$ and that it does not depend on the direction. The slowness surface is spherical, with radius $1/V$. Equation (81) shows that group velocity vector $\vec{\mathcal{U}}$ and slowness vector $\vec{p}$ are parallel in isotropic media, and $\vec{\mathcal{U}} \cdot \vec{p} = 1$.

We shall now briefly return to curvilinear coordinates. In isotropic media, the relations between modified Lagrangians and modified Hamiltonians are simple, even in general curvilinear coordinates.

We shall introduce the metric tensor $G_{ij}(x^k)$ in standard Riemannian space, in which the distance is measured by length. In isotropic media, the relation between metric tensor G_{ij} and wave propagation metric tensor g_{ij}, introduced by (9), is as follows:

$$g_{ij}(x^k, \dot{x}^k) = V^{-2}(x^k) G_{ij}(x^k) \ . \tag{85}$$

All relations derived in Section 3.1 simplify in this case. The modified Lagrangian reads, see (35):

$$\mathcal{L}^M(x^k,\dot{x}^k) = \tfrac{1}{2}V^{-2}(x^k)G_{ij}(x^k)\dot{x}^i\dot{x}^j \ . \tag{86}$$

The relevant Euler-Lagrange equations (39) are as follows:

$$\frac{\mathrm{d}}{\mathrm{d}\tau}(V^{-2}G_{ij}\dot{x}^j) - \tfrac{1}{2}\dot{x}^k\dot{x}^l\frac{\partial}{\partial x^i}(V^{-2}G_{kl}) = 0 \ . \tag{87}$$

For p_i, we can use (42) to obtain

$$p_i = \frac{\partial\mathcal{L}^M}{\partial\dot{x}^i} = V^{-2}(x^k)G_{ij}(x^k)\dot{x}^j \ . \tag{88}$$

This can be simply inverted for $\dot{x}^j$. We take into account that $G_{ij}G^{ik} = \delta^k_j$, where G^{ik} are the contravariant components of the metric tensor. Then

$$\dot{x}^j = V^2(x^k)G^{ji}(x^k)p_i \ . \tag{89}$$

The modified Lagrangian $\mathcal{L}^M$ can then be expressed in terms of p_k, see (86) and (89):

$$\mathcal{L}^M = \tfrac{1}{2}V^2p_jp_kG^{jk} \ . \tag{90}$$

The reduced Hamiltonian $\mathcal{H}^M(x^k,p_k)$ is obtained from (42), (89) and (90),

$$\mathcal{H}^M(x^k,p_k) = \tfrac{1}{2}V^2(x^k)G^{ij}(x^k)p_ip_j \ , \tag{91}$$

see also (63). We remind the reader that $\mathcal{L}^M(x^k,\dot{x}^k) = \frac{1}{2}$ and $\mathcal{H}^M(x^k,p_k) = \frac{1}{2}$ along the whole ray, see (36) and (52). This yields, along the whole ray,

$$G_{ij}(x^k)\dot{x}^i\dot{x}^j = V^2(x^k) \ , \quad G^{ij}(x^k)p_ip_j = V^{-2}(x^k) \ . \tag{92}$$

Finally, the Hamiltonian ray tracing system is as follows:

$$\begin{aligned}\frac{\mathrm{d}x^i}{\mathrm{d}\tau} &= \frac{\partial\mathcal{H}^M}{\partial p_i} = V^2G^{ik}p_k \ , \\ \frac{\mathrm{d}p_i}{\mathrm{d}\tau} &= -\frac{\partial\mathcal{H}^M}{\partial x^i} = -\tfrac{1}{2}G^{jk}p_jp_k\frac{\partial V^2}{\partial x^i} - \tfrac{1}{2}V^2p_jp_k\frac{\partial G^{jk}}{\partial x^i} \ .\end{aligned}$$

As $G^{jk}p_jp_k\partial V^2/\partial x^i = V^{-2}(x^k)\partial V^2/\partial x^i = 2\partial\mathrm{ln}V/\partial x^i$, we obtain

$$\frac{\mathrm{d}x^i}{\mathrm{d}\tau} = V^2G^{ik}p_k \ , \quad \frac{\mathrm{d}p_i}{\mathrm{d}\tau} = -\frac{\partial\mathrm{ln}V}{\partial x^i} - \tfrac{1}{2}V^2p_jp_k\frac{\partial G^{jk}}{\partial x^i} \ . \tag{93}$$

This corresponds fully to the Hamiltonian ray equations (3.5.54) in Červený (2001a) for $n = 0$, if we take into account that

$$\lim_{n\to 0}\left(n^{-1}\,\frac{\partial(1/V)^n}{\partial x^i}\right) = -\frac{\partial\mathrm{ln}V}{\partial x^i} \ .$$

Note that the modified Lagrangian for isotropic media and curvilinear coordinates was also used by Babich and Buldyrev (1972). The authors apply the Legendre transformation to it to derive the relevant modified Hamiltonian (even though they do not use the term "modified").

3.4 Ellipsoidal anisotropy. Cartesian coordinates

In the next analytical example, we shall consider the ellipsoidal anisotropy, and Cartesian coordinates. Ellipsoidal anisotropy has been traditionally introduced in the p_i-domain, see Burridge, Chadwick and Norris (1993), Mensch and Farra (1999), Ettrich, Gajewski and Kashtan (2001). For this reason, we shall choose an approach opposite to that in Section 3.3: We shall start from a known modified Hamiltonian, and derive the relevant modified Lagrangian. The Hamiltonian ray tracing equations are obtained automatically in this case, as the modified Hamiltonian is pressumably known. Using the dual Legendre transform, we shall derive the Lagrangian and the Lagrangian ray tracing system. In this case, all computations can be performed analytically.

We assume that the modified Hamiltonian $\mathcal{H}^M(x_k, p_k)$ is given by the relation

$$\mathcal{H}^M(x_k, p_k) = \tfrac{1}{2}G(x_k, p_k) \ , \quad G(x_k, p_k) = A^2(x_k)p_1^2 + B^2(x_k)p_2^2 + C^2(x_k)p_3^2 \ . \tag{94}$$

As we immediately see, slowness surface $\mathcal{H}^M(x_k, p_k) = \frac{1}{2}$ is an ellipsoid,

$$A^2(x_k)p_1^2 + B^2(x_k)p_2^2 + C^2(x_k)p_3^2 = 1 \ , \tag{95}$$

with half-axes $1/A$, $1/B$ and $1/C$. Actually, it would be more general to consider the generally oriented ellipsoid $G(x_k, p_k) = g^{ij}(x_k)p_ip_j$, and even the following derivations would be more concise due to Einstein summation convention. The diagonalized ellipsoidal form (95) is, however, more transparent. We denote $p_i = |\vec{p}|N_i$, where N_i is the unit vector along $\vec{p}$ (perpendicular to the wavefront). This yields

$$|\vec{p}| = [A^2(x_k)N_1^2 + B^2(x_k)N_2^2 + C^2(x_k)N_3^2]^{-1/2} \ . \tag{96}$$

As phase velocity $\mathcal{C}$ equals $1/|\vec{p}|$,

$$\mathcal{C} = [A^2(x_k)N_1^2 + B^2(x_k)N_2^2 + C^2(x_k)N_3^2]^{1/2} \ . \tag{97}$$

Hamiltonian ray tracing system (44) is easily obtained from (94),

$$\begin{aligned}
\frac{\mathrm{d}x_1}{\mathrm{d}\tau} &= A^2p_1 \ , \quad & \frac{\mathrm{d}p_1}{\mathrm{d}\tau} &= -\tfrac{1}{2}\Big(\frac{\partial A^2}{\partial x_1}p_1^2 + \frac{\partial B^2}{\partial x_1}p_2^2 + \frac{\partial C^2}{\partial x_1}p_3^2\Big) \ , \\
\frac{\mathrm{d}x_2}{\mathrm{d}\tau} &= B^2p_2 \ , \quad & \frac{\mathrm{d}p_2}{\mathrm{d}\tau} &= -\tfrac{1}{2}\Big(\frac{\partial A^2}{\partial x_2}p_1^2 + \frac{\partial B^2}{\partial x_2}p_2^2 + \frac{\partial C^2}{\partial x_2}p_3^2\Big) \ , \\
\frac{\mathrm{d}x_3}{\mathrm{d}\tau} &= C^2p_3 \ , \quad & \frac{\mathrm{d}p_3}{\mathrm{d}\tau} &= -\tfrac{1}{2}\Big(\frac{\partial A^2}{\partial x_3}p_1^2 + \frac{\partial B^2}{\partial x_3}p_2^2 + \frac{\partial C^2}{\partial x_3}p_3^2\Big) \ .
\end{aligned} \tag{98}$$

Thus, group velocity components $\mathcal{U}_i = \dot{x}_i = \partial\mathcal{H}/\partial p_i$ are given by relations

$$\mathcal{U}_1 = A^2p_1 \ , \quad \mathcal{U}_2 = B^2p_2 \ , \quad \mathcal{U}_3 = C^2p_3 \ . \tag{99}$$

The inversion for p_i is simple

$$p_1 = A^{-2}\mathcal{U}_1 \ , \quad p_2 = B^{-2}\mathcal{U}_2 \ , \quad p_3 = C^{-2}\mathcal{U}_3 \ . \tag{100}$$

Equation (42) can be used to compute the modified Lagrangian from the modified Hamiltonian,

$$\mathcal{L}^M(x_k, \dot{x}_k) = \dot{x}_ip_i - \mathcal{H}^M(x_k, p_k) \ ,$$

where p_i are expressed in terms of $\mathcal{U}_i = \dot{x}_i$ using (100). Using also (94), we obtain

$$\begin{aligned}\mathcal{L}^M(x_k, \dot{x}_k) &= \dot{x}_1 p_1 + \dot{x}_2 p_2 + \dot{x}_3 p_3 - \tfrac{1}{2}(A^2 p_1^2 + B^2 p_2^2 + C^2 p_3^2) \\ &= \dot{x}_1^2 A^{-2} + \dot{x}_2^2 B^{-2} + \dot{x}_3^2 C^{-2} - \tfrac{1}{2}(\dot{x}_1^2 A^{-2} + \dot{x}_2^2 B^{-2} + \dot{x}_3^2 C^{-2}) \ .\end{aligned}$$

This yields,

$$\mathcal{L}^M(x_k, \dot{x}_k) = \tfrac{1}{2}(A^{-2}\dot{x}_1^2 + B^{-2}\dot{x}_2^2 + C^{-2}\dot{x}_3^2) \ . \qquad (101)$$

Thus, group velocity surface $\mathcal{L}^M(x_k, \mathcal{U}_k) = \frac{1}{2}$ is again ellipsoidal

$$A^{-2}\mathcal{U}_1^2 + B^{-2}\mathcal{U}_2^2 + C^{-2}\mathcal{U}_3^2 = 1 \ . \qquad (102)$$

The half-axes of the group velocity ellipsoid are A, B, and C. They are reciprocal to the half-axes of the slowness ellipsoid.

Using $\mathcal{U}_i = \mathcal{U} t_i$, where t_i is the unit vector along the ray, $t_i = \mathrm{d}x_i/\mathrm{d}s$, we obtain from (102)

$$\mathcal{U}^2(A^{-2}t_1^2 + B^{-2}t_2^2 + C^{-2}t_3^2) = 1 \ .$$

This yields

$$\mathcal{U} = (A^{-2}t_1^2 + B^{-2}t_2^2 + C^{-2}t_3^2)^{-1/2} \ . \qquad (103)$$

The Euler-Lagrange equations (39) corresponding to modified Lagrangian $\mathcal{L}^M(x_i, \dot{x}_i)$ given by (101) read:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\tau}(A^{-2}\dot{x}_1) - \tfrac{1}{2}\Big(\frac{\partial A^{-2}}{\partial x_1}\dot{x}_1^2 + \frac{\partial B^{-2}}{\partial x_1}\dot{x}_2^2 + \frac{\partial C^{-2}}{\partial x_1}\dot{x}_3^2\Big) &= 0 \ , \\ \frac{\mathrm{d}}{\mathrm{d}\tau}(B^{-2}\dot{x}_2) - \tfrac{1}{2}\Big(\frac{\partial A^{-2}}{\partial x_2}\dot{x}_1^2 + \frac{\partial B^{-2}}{\partial x_2}\dot{x}_2^2 + \frac{\partial C^{-2}}{\partial x_2}\dot{x}_3^2\Big) &= 0 \ , \\ \frac{\mathrm{d}}{\mathrm{d}\tau}(C^{-2}\dot{x}_3) - \tfrac{1}{2}\Big(\frac{\partial A^{-2}}{\partial x_3}\dot{x}_1^2 + \frac{\partial B^{-2}}{\partial x_3}\dot{x}_2^2 + \frac{\partial C^{-2}}{\partial x_3}\dot{x}_3^2\Big) &= 0 \ .\end{aligned} \qquad (104)$$

We can see that Lagrangian ray tracing system (104) also follows directly from Hamiltonian ray tracing system (98), if we use (100).

Note that modified Lagrangian (101) can also be obtained from modified Hamiltonian (94) in a more straightforward way, using the wave propagation metric tensor. Using (53), we obtain from (94) for $g^{ij} = \overline{g}^{ij}$,

$$g^{ij} = \begin{pmatrix} A^2 & 0 & 0 \\ 0 & B^2 & 0 \\ 0 & 0 & C^2 \end{pmatrix} \ .$$

This yields, see (55),

$$g_{ij} = \begin{pmatrix} A^{-2} & 0 & 0 \\ 0 & B^{-2} & 0 \\ 0 & 0 & C^{-2} \end{pmatrix} \ .$$

The modified Lagrangian is then given by (35) and yields (101).

4 Reduced Lagrangians and Hamiltonians

The next possibility of overcoming the difficulties connected with the vanishing Hessian determinant consists in introducing reduced Lagrangians and Hamiltonians. They correspond to the special choice of parameter u, namely $u = x^1$, $u = x^2$ or $u = x^3$.

Consider first the general parameteric Fermat's variational principle (22),

$$\delta \int_S^R \mathcal{L}(x^k, x^{k\prime})\mathrm{d}u = 0 \ , \qquad x^{k\prime} = \mathrm{d}x^k/\mathrm{d}u \ . \tag{105}$$

This principle can be used for any parameter u, including $u = x^1$, $u = x^2$ or $u = x^3$. The Euler-Lagrange equations corresponding to (105) are given by (23). It was shown in Section 2 that, in this case, Lagrangian $\mathcal{L}(x^k, x^{k\prime})$ is a homogeneous function of the first degree in $x^{i\prime}$ and that the corresponding Hamiltonian $\mathcal{H}(x^k, p_k)$ vanishes, $\mathcal{H}(x^k, p_k) = 0$, see (30). We also have

$$p_i = \frac{\partial \mathcal{L}}{\partial x^{i\prime}} \ , \quad x^{i\prime} = \frac{\partial \mathcal{H}}{\partial p_i} \ , \quad \mathcal{H}(x^k, p_k) = x^{i\prime} p_i - \mathcal{L}(x^k, x^{k\prime}) \ . \tag{106}$$

4.1 Legendre transform for reduced Lagrangians and Hamiltonians

Assume now that $u = x^3$. Then $\mathcal{L}(x^k, x^{k\prime}) = \mathcal{L}^R(x^k, x^{K\prime})$, where $x^{K\prime} = \mathrm{d}x^K/\mathrm{d}x^3$ and $K = 1, 2$. Lagrangian $\mathcal{L}^R$ is no longer a homogeneous function of the first degree in $x^{K\prime}$. The Euler-Lagrange equations for the Lagrangian $\mathcal{L}^R(x^k, x^{K\prime})$ are as follows:

$$\frac{\mathrm{d}}{\mathrm{d}x^3}\left(\frac{\partial \mathcal{L}^R}{\partial x^{I\prime}}\right) - \frac{\partial \mathcal{L}^R}{\partial x^I} = 0 \ , \qquad I = 1, 2 \ . \tag{107}$$

Thus, (107) consists of two ordinary differential equations of the second order, not of three as (23).

Using, (106), we also obtain $x^{3\prime} = \partial\mathcal{H}/\partial p_3 = 1$, as $x^{3\prime} = \mathrm{d}x^3/\mathrm{d}x^3$. This yields,

$$\mathcal{H}(x^k, p_k) = p_3 + \mathcal{H}^R(x^k, p_K) = 0 \ . \tag{108}$$

We shall call $\mathcal{H}^R(x^k, p_K)$ the *reduced Hamiltonian*. Using (108), p_3 is obtained by the equation

$$p_3 = -\mathcal{H}^R(x^k, p_K) \ . \tag{109}$$

This equation represents an explicit equation of the slowness surface. The reduced Hamiltonian is obtained from equation $\mathcal{H}(x_k, p_k) = \frac{1}{2}$ by solving it for p_3. In seismology, this is a well-known procedure, even in anisotropic media, as it is used in the reflection/transmission problem and in one-way propagation. See Červený (2001a).

As $p_i = \partial T/\partial x^i$, Equation (109) represents a partial differential equation of the first order for $T = T(x^i)$,

$$\partial T/\partial x^3 = -\mathcal{H}^R(x^k, \mathrm{d}T/\partial x^K) \ , \quad K = 1, 2 \ , \quad k = 1, 2, 3 \ .$$

This equation is usually called the evalution Hamilton-Jacobi equation. Actually, the evalution Hamilton-Jacobi equation is obtained from the stationary Hamilton-Jacobi equation $\mathcal{H}(x^k, p_k) = \frac{1}{2}$ as its one-way alternative; specifying $u = x^3$ and solving it for p_3.

Equations (106) simplify for $u = x_3$ as follows:

$$p_I = \frac{\partial \mathcal{L}^R}{\partial x^{I\prime}} \ , \quad x^{I\prime} = \frac{\partial \mathcal{H}^R}{\partial p_I} \ , \quad \mathcal{H}^R(x^k, p_K) = x^{I\prime} p_I - \mathcal{L}^R(x^k, x^{K\prime}) \ . \tag{110}$$

Thus, the problem reduces from three dimensions to two dimensions.

As the reduced Lagrangian $\mathcal{L}^R(x^k, x^{K\prime})$ is no longer a homogeneous function of the first degree in $x^{I\prime}$, the Hessian determinant is not identically zero, and the dual Legendre transform (110) can be applied to compute $\mathcal{H}^R(x^k, p_K)$. The final Hamiltonian ray tracing system consists of four ordinary differential equations of the first order,

$$\frac{\mathrm{d}x^I}{\mathrm{d}x^3} = \frac{\partial \mathcal{H}^R}{\partial p_I} \ , \qquad \frac{\mathrm{d}p_I}{\mathrm{d}x^3} = -\frac{\partial \mathcal{H}^R}{\partial x^I} \ . \tag{111}$$

To complete the solution, (109) for p_3 must be added.

The most difficult step in computing the reduced Hamiltonian $\mathcal{H}^R(x^k, p_K)$ from a known reduced Lagrangian $\mathcal{L}^R(x^k, x^{K\prime})$ consists in the solution of equations $p_I = \partial \mathcal{L}^R / \partial x^{I\prime}$ $(I = 1, 2)$ for $x^{I\prime} = x^{I\prime}(x^k, p_K)$. The same applies to the solution of $x^{I\prime} = \partial \mathcal{H}^R / \partial p_I$ for p_1, p_2, if we wish to determine the reduced Lagrangian $\mathcal{L}^R(x^k, x^{K\prime})$ from the known reduced Hamiltonian.

The disadvantage of the approach based on the reduced Lagrangian and Hamiltonian is that it is applicable only to one-way propagation, along or against the x^3 axis. It fails at turning points, where the rays change their direction with respect to the x_3-axis. For one-way propagation without turning points, however, it is fully alternative to the approach based on the modified Lagrangian and Hamiltonian, only the computational schemes are different. See the example for ellipsoidal anisotropy in Section 4.4.

In general, the ray tracing equations based on the modified Hamiltonian (or on modified Lagrangian) are preferable, as they do not pose any difficulties with the turning points of the ray.

4.2 Reduced Hamiltonian ray tracing in general anisotropic media

The reduced Hamiltonian ray tracing system (111) is applicable to general anisotropic inhomogeneous media and to arbitrary curvilinear coordinates x^I, $I = 1, 2$. It consists of four equations only. It is, however, not simple to calculate $\partial p_3 / \partial p_I$ and $\partial p_3 / \partial x^I$. Actually, p_3 is a solution of an algebraic equation of the sixth order.

It is, however, known from the seismic ray theory that p_3 is an eigenvalue of some 6×6 matrix $\mathbf{A}$. See Červený (2001a), Equations (5.4.124) and (5.4.135), and also Červený (2001b), where other references can be found. If we denote the eigenvalue of $\mathbf{A}$ by σ and the relevant 6×1 eigenvector by $\mathbf{W}$, the eigenvalue system reads

$$\mathbf{A}\mathbf{W} = \mathbf{W}\sigma \ . \tag{112}$$

It should be noted that the 6×6 matrix $\mathbf{A}$ is not symmetrical, and, consequently, the 6×1 eigenvectors $\mathbf{W}$ are not orthogonal. Matrix $\mathbf{A}$, however, can be symmetrized by multiplying it by the 6×6 matrix $\mathbf{I}_1$ from the left, where $\mathbf{I}_1$ is given by the relation

$$\mathbf{I}_1 = \begin{pmatrix} \hat{\mathbf{0}} & \hat{\mathbf{I}} \\ \hat{\mathbf{I}} & \hat{\mathbf{0}} \end{pmatrix} . \qquad (113)$$

Here $\hat{\mathbf{0}}$ and $\hat{\mathbf{I}}$ are 3×3 null and identity matrices, respectively. As we can see, $\mathbf{I}_1$ is a 6×6 analogue of the well-known 2×2 Pauli matrix.

We shall now express eigenvalue σ in term of $\mathbf{A}$ and $\mathbf{W}$. Multiplying (112) from the left by $\mathbf{W}^T\mathbf{I}_1$, we obtain

$$\sigma = \mathbf{W}^T\mathbf{I}_1\mathbf{A}\mathbf{W}/\mathbf{W}^T\mathbf{I}_1\mathbf{W} . \qquad (114)$$

Denoting the partial derivative of σ with respect to any quantity by a prime, we can simply determine σ' from (114), in terms of $\mathbf{A}'$. We obtain, see Červený (2001b),

$$\sigma' = \mathbf{W}^T\mathbf{I}_1\mathbf{A}'\mathbf{W}/\mathbf{W}^T\mathbf{I}_1\mathbf{W} . \qquad (115)$$

Thus, the Hamiltonian ray tracing system (111) reads

$$\frac{\mathrm{d}x^I}{\mathrm{d}x^3} = -\mathbf{W}^T\mathbf{I}_1\frac{\partial \mathbf{A}}{\partial p_I}\mathbf{W}/\mathbf{W}^T\mathbf{I}_1\mathbf{W} , \qquad \frac{\mathrm{d}p_I}{\mathrm{d}x^3} = \mathbf{W}^T\mathbf{I}_1\frac{\partial \mathbf{A}}{\partial x^I}\mathbf{W}/\mathbf{W}^T\mathbf{I}_1\mathbf{W} . \qquad (116)$$

For more details and detailed expressions for $\mathbf{A}$ and $\mathbf{W}$ refer to Červený (2001b) in this volume. Some simple analytical examples are also discussed there.

4.3 Isotropic media, Cartesian coordinates

We shall again consider an isotropic medium and Cartesian coordinates x_1, x_2, x_3, as a simple example of the general procedure based on the reduced Lagrangian. The standard Lagrangian, in parameteric form, then reads

$$\mathcal{L}(x_k, x_k') = V^{-1}(x_k)[(x_1')^2 + (x_2')^2 + (x_3')^2]^{1/2} , \quad x_i' = \mathrm{d}x_i/\mathrm{d}u .$$

If we choose the parameter along curve $u = x_3$, we obtain the reduced Lagrangian $\mathcal{L}^R(x_k, x_K')$,

$$\mathcal{L}^R(x_k, x_K') = V^{-1}(x_k)[(x_1')^2 + (x_2')^2 + 1]^{1/2} , \quad x_K' = \mathrm{d}x_K/\mathrm{d}x_3 . \qquad (117)$$

Here $K = 1, 2$. Fermat's variational principle then reads

$$\delta \int_S^R V^{-1}(x_k)[(x_1')^2 + (x_2')^2 + 1]^{1/2}\mathrm{d}x_3 = 0 . \qquad (118)$$

The relevant Euler-Lagrange equations are as follows:

$$\frac{\mathrm{d}}{\mathrm{d}x_3}\left\{\frac{x_I'}{V(x_i)[1 + (x_1')^2 + (x_2')^2]^{1/2}}\right\} - [1 + (x_1')^2 + (x_2')^2]^{1/2}\frac{\partial}{\partial x_I}(\frac{1}{V}) = 0 . \qquad (119)$$

To compute p_I, we use (110):

$$p_I = \partial\mathcal{L}^R/\partial x_I' = x_I'/V[1 + (x_1')^2 + (x_2')^2]^{1/2} , \quad I = 1, 2 . \qquad (120)$$

The two equations (120) can be solved for x'_I to yield

$$x'_I = p_I V [1 - V^2(p_1^2 + p_2^2)]^{-1/2} \ . \tag{121}$$

Using (121), the reduced Lagrangian $\mathcal{L}^R(x_i, x'_I)$ can be expressed in terms of p_I, see (117):

$$\mathcal{L}^R = V^{-1}[1 - V^2(p_1^2 + p_2^2)]^{-1/2} \ . \tag{122}$$

The final equation for the reduced Hamiltonian $\mathcal{H}^R(x_i, p_i)$ is given by (110), where we insert (121) and (122):

$$\mathcal{H}^R(x_k, p_K) = -V^{-1}(x_k)[1 - V^2(p_1^2 + p_2^2)]^{1/2} \ . \tag{123}$$

The relevant Hamiltonian ray tracing system (111) then reads

$$\frac{\mathrm{d}x_I}{\mathrm{d}x_3} = \frac{V p_I}{[1 - V^2(p_1^2 + p_2^2)]^{1/2}} \ , \quad \frac{\mathrm{d}p_I}{\mathrm{d}x_3} = \frac{1}{[1 - V^2(p_1^2 + p_2^2)]^{1/2}} \frac{\partial}{\partial x_I}\left(\frac{1}{V}\right) \ . \tag{124}$$

To complete the solution, we must add (109):

$$p_3 = -\mathcal{H}^R(x_k, p_K) = V^{-1}(x_k)[1 - V^2(p_1^2 + p_2^2)]^{1/2} \ . \tag{125}$$

The choice of sign in (125) corresponds to the propagation in the positive direction of the x_3 axis.

The reduced Hamiltonian derived here from the reduced Lagrangian using the Fermat's variational principle corresponds exactly to the reduced Hamiltonian derived from the eikonal equation $V^2 p_i p_i = 1$ (which further follows from the elastodynamic equation). Compare Equation (123) given here with the expression for the reduced Hamiltonian used in Červený (2001a), Eq.(3.1.28). One sees that, even the ray tracing systems, given here by (124), and the ray tracing systems (3.1.29) in Červený (2001a), are the same.

The well-known disadvantage of ray tracing system (124) consists in the fact that parameter x_3 along the ray is not necessarily monotonic along the ray. The ray tracing system (124) fails at ray turning points, at which the ray is perpendicular to the x_3 axis (the minimum with respect to x_3, for example).

4.4 Ellipsoidal anisotropy, Cartesian coordinates

We shall start from Fermat's variational principle in parameteric form (22), and from the homogeneous ellipsoidal Lagrangian:

$$\mathcal{L}(x_k, x'_k) = [(x'_1)^2 A^{-2} + (x'_2)^2 B^{-2} + (x'_3)^2 C^{-2}]^{1/2} \ , \quad x'_i = \mathrm{d}x_i/\mathrm{d}u \ . \tag{126}$$

Here A, B, C are smooth functions of Cartesian coordinates x_i. As Lagrangian (126) is a homogeneous function of the first degree in x'_i, the Hessian determinant vanishes, and the dual Legendre transform cannot be used to derive the corresponding Hamiltonian. We can, however, introduce parameter $u = x_3$ and reduced Lagrangian $\mathcal{L}^R(x_i, x'_I)$:

$$\mathcal{L}^R(x_k, x'_K) = [(x'_1)^2 A^{-2} + (x'_2)^2 B^{-2} + C^{-2}]^{1/2} \ . \tag{127}$$

Fermat's variational principle then reads

$$\delta \int_S^R [(x'_1)^2 A^{-2} + (x'_2)^2 B^{-2} + C^{-2}]^{1/2} \mathrm{d}x_3 = 0 \ ,$$

and the relevant Euler-Lagrange equations are as follows:

$$\frac{\mathrm{d}}{\mathrm{d}x_3}\left(\frac{\partial D}{x_I^{'}}\right) - \frac{\partial D}{\partial x_I} = 0 \ , \qquad D = [(x_1^{'})^2 A^{-2} + (x_2^{'})^2 B^{-2} + C^{-2}]^{1/2} \ . \tag{128}$$

To compute p_I, we use (110),

$$p_1 = x_1^{'} A^{-2} D^{-1} \ , \qquad p_2 = x_2^{'} B^{-2} D^{-1} \ , \tag{129}$$

which can be solved for $x_I^{'}$ to yield

$$x_1^{'} = \frac{C^{-1} A^2 p_1}{[1 - A^2 p_1^2 - B^2 p_2^2]^{1/2}} \ , \quad x_2^{'} = \frac{C^{-1} B^2 p_2}{[1 - A^2 p_1^2 - B^2 p_2^2]^{1/2}} \ . \tag{130}$$

We can now express the reduced Lagrangian in terms of p_I

$$\mathcal{L}^R = C^{-1} [1 - A^2 p_1^2 - B^2 p_2^2]^{-1/2} \ . \tag{131}$$

The corresponding reduced Hamiltonian $\mathcal{H}^R(x_i, p_i)$ is given by (110), where we insert (130) and (131)

$$\mathcal{H}^R(x_k, p_K) = x_I^{'} p_I - \mathcal{L}(x_k, x_K^{'}) = -\left[\frac{1}{C^2} - \frac{A^2}{C^2} p_1^2 - \frac{B^2}{C^2} p_2^2\right]^{1/2} \ . \tag{132}$$

The relevant Hamiltonian ray tracing equations follow from (132), using (111)

$$\begin{aligned}
\frac{\mathrm{d}x_1}{\mathrm{d}x_3} &= \frac{A^2 p_1}{C[1 - A^2 p_1^2 - B^2 p_2^2]^{1/2}} \ , \\
\frac{\mathrm{d}x_2}{\mathrm{d}x_3} &= \frac{B^2 p_2}{C[1 - A^2 p_1^2 - B^2 p_2^2]^{1/2}} \ , \\
\frac{\mathrm{d}p_1}{\mathrm{d}x_3} &= -\frac{C}{2[1 - A^2 p_1^2 - B^2 p_2^2]^{1/2}} \left[\frac{\partial}{\partial x_1}\left(\frac{1}{C^2}\right) - p_1^2 \frac{\partial}{\partial x_1}\left(\frac{A}{C}\right)^2 - p_2^2 \frac{\partial}{\partial x_1}\left(\frac{B}{C}\right)^2\right] \ , \\
\frac{\mathrm{d}p_2}{\mathrm{d}x_3} &= -\frac{C}{2[1 - A^2 p_1^2 - B^2 p_2^2]^{1/2}} \left[\frac{\partial}{\partial x_2}\left(\frac{1}{C^2}\right) - p_1^2 \frac{\partial}{\partial x_2}\left(\frac{A}{C}\right)^2 - p_2^2 \frac{\partial}{\partial x_2}\left(\frac{B}{C}\right)^2\right] \ .
\end{aligned} \tag{133}$$

This system must be supplemented by Equation (109),

$$p_3 = -\mathcal{H}^R(x_i, p_I) = \left[\frac{1}{C^2} - \frac{A^2}{C^2} p_1^2 - \frac{B^2}{C^2} p_2^2\right]^{1/2} \ . \tag{134}$$

We shall now compare the results obtained here with those obtained using modified Lagrangian $\mathcal{L}^M$ and modified Hamiltonian $\mathcal{H}^M$. We consider only rays without turning points with respect to x_3. Reduced Lagrangian $\mathcal{L}^R(x_k, x_K^{'})$ is given by (127)

$$\mathcal{L}^R(x_k, x_K^{'}) = [(x_1^{'})^2 A^{-2} + (x_2^{'})^2 B^{-2} + C^{-2}]^{1/2} \ , \quad x_I^{'} = \mathrm{d}x_I / \mathrm{d}x_3 \ .$$

As $x_I^{'} = \dot{x}_I / \dot{x}_3$, where $\dot{x}_i = \mathrm{d}x_i / \mathrm{d}\tau$, we obtain

$$\mathcal{L}^R(x_k, \dot{x}_k) = [(\dot{x}_1)^2 A^{-2} + (\dot{x}_2)^2 B^{-2} + (\dot{x}_3)^2 C^{-2}]^{1/2} / \dot{x}_3 \ .$$

Consequently, the relevant modified Lagrangian $\mathcal{L}^M(x_k, \dot{x}_k)$, corresponding to the parameter $u = \tau$ along the curves, is

$$\mathcal{L}^M(x_k, \dot{x}_K) = \tfrac{1}{2}[(\dot{x}_1)^2 A^{-2} + (\dot{x}_2)^2 B^{-2} + (\dot{x}_3)^2 C^{-2}] = \tfrac{1}{2} \ ,$$

see (35) and (101). The relevant modified Hamiltonian $\mathcal{H}^M(x_k, p_k)$ is given by (94)

$$\mathcal{H}^M(x_k, p_k) = \tfrac{1}{2}(A^2 p_1^2 + B^2 p_2^2 + C^2 p_3^2) = \tfrac{1}{2} \ .$$

Solving this for p_3 yields, for the forward propagation,

$$p_3 = C^{-1}[1 - A^2 p_1^2 - B^2 p_2^2]^{1/2} \ .$$

This corresponds strictly to (134).

Even Hamiltonian ray tracing equations (133) can be simply obtained from ray tracing equations (98). We merely divide (98) by $\mathrm{d}x_3/\mathrm{d}\tau = C^2 p_3 = C[1 - A^2 p_1^2 - B^2 p_2^2]^{1/2}$, and insert $p_3^2 = C^{-2} - (A/C)^2 p_1^2 - (B/C)^2 p_2^2$, see (134).

Acknowledgements

The research has been supported by the Consortium Project "Seismic Waves in Complex 3-D Structures".

References

Asanov, G.S. (1985). *Finsler geometry, relativity and gauge theories.* Dordrecht: Reidel.

Babich, V.M. (1961). Ray method of the computation of the intensity of wave fronts in elastic inhomogeneous anisotropic medium. In *Problems of the dynamic theory of propagation of seismic waves* (in Russian), vol. 5, ed. G.I. Petrashen, pp. 36–46. Leningrad: Leningrad Univ Press. (Translation to English: *Geophys. J. Int.* **118**, 379–83, 1994.)

Babich, V.M., and Buldyrev, V.S. (1972). *Asymptotic methods in problems of diffraction of short waves* (in Russian). Moscow: Nauka. (Translated to English by Springer, 1991, under the title *Short-wavelength diffraction theory.*)

Beem, J.K. (1970). Indefinite Finsler spaces and timelike spaces. *Canad. J. Math.* **22**, 1035–9.

Beem,, J.K. and Kishta, M.A. (1974). On generalized indefinite Finsler spaces. *Indiana Univ. Math. J.* **23**, 845–53.

Brdička, M., and Hladík, A. (1987). *Theoretical mechanics* (in Czech). Praha: Academia.

Burridge, R., Chadwick, P., and Norris, A.N. (1993). Fundamental elastodynamic solutions for anisotropic media with ellipsoidal slowness surfaces. *Proc. Roy. Soc. London* **A440**, 655–81.

Červený, V. (2001a). *Seismic ray theory.* New York: Cambridge Univ. Press.

Červený, V. (2001b). Reduced Hamiltonian in ray tracing systems for anisotropic inhomogeneous media. In this report.

Chapman, C.H. and Pratt, R.G. (1992). Traveltime tomography in anisotropic media - I. Theory. *Geophys. J. Int.* **109**, 1–19.

Epstein, M. and Śniatycki, J. (1992). Fermat's principle in elastodynamics. *J. Elasticity* **27**, 45–56.

Ettrich, N., Gajewski, D., and Kashtan, B. (2001). Reference ellipsoids for anisotropic media. *Geophysical Prospecting* **49**, 321–334.

Goldstein, H. (1980). *Classical mechanics.* Reading, MA: Addison Wesley.

Hanyga, A. (1996). Fermat's principle in anisotropic elasticity. Presented at the 7th IWSA, Miami, 1996.

Hanyga, A., Lenartowicz, E., and Pajchel, J. (1984). *Seismic wave propagation in the Earth.* Amsterdam: Elsevier.

Helbig, K. (1994). *Foundations of anisotropy for exploration seismics.* Oxford: Pergamon.

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

Klimeš, L., and Kvasnička, M. (1994). 3-D network ray tracing. *Geophys. J. Int.* **116**, 726–38.

Kline, M. and Kay, I.W. (1965). *Electromagnetic theory and geometrical optics.* New York: Interscience.

Mensch, T., and Farra, V. (1999). Computation of qP-wave rays, traveltimes and slowness vectors in orthorhombic media. *Geophys. J. Int.* **138**, 244–56.

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

Pšenčík, I. (1997). Fermat's principle in inhomogeneous anisotropic media. In *Seismic Waves in Complex 3-D Structures*, *Report 6*, pp. 175–7. Prague: Charles Univ., Faculty of Mathematics and Physics, Depart. of Geophysics.

Rund, H. (1966). *The Hamilton-Jacobi theory in the calculus of variations.* London: Van Nostrand. (Revised edition: Huntington, N.Y.: Krieger Publ., 1973.)

Stephen, R. (1996). Fermat's principle for anisotropic elastic media. Presented at the 7th IWSA, Miami, 1996.

Thomson, C.J. (1999). The 'gap' between seismic ray theory and 'full' wavefield extrapolation. *Geophys. J. Int.* **137**, 364–80.