KINEMATIC HYPOCENTRE DETERMINATION USING THE PARAXIAL RAY APPROXIMATION

LUDĚK KLIMEŠ*

Department of Geophysics, Charles University, Prague

Summary: *The robust nonlinear approach by Tarantola and Valette, consisting in direct evaluation of the "probability" density function, is supplemented with the paraxial ray approximation of the travel time. A sufficiently dense 2-parametric system of rays from each receiver is evaluated only once for all hypocentre determinations. The interpolation formulae for the travel times apply to all travel-time branches. Their derivation is based on the summation of Gaussian packets. The proposed algorithm for determining the hypocentre is able to find all of its possible locations.*

1. INTRODUCTION

This paper is devoted to the determination of a seismic hypocentre in general 3-D laterally inhomogeneous media. The robust nonlinear approach consisting in direct evaluation of the "probability" density function has been proposed by Tarantola and Valette [1]. In this paper, Tarantola and Valette's method is supplemented with the paraxial ray approximation of the travel time, based on a system of rays traced once for all hypocentre determinations. On the other hand, Moser et al. [2] have combined Tarantola and Valette's method with the network shortest path calculation of travel times. Their shortest path method is more reliable than the ray tracing, but it deals just with the first arrivals.

For simple structures, of course, there are more efficient methods than those evaluating whole "probability" density functions, but they are applicable only to special classes of models (such as homogeneous models, 1-D models, etc.) Such efficient but special methods are not studied in this paper.

It is assumed that the seismic model of the medium is known, with a specified accuracy. At present, in many regions, the quantity and quality of seismic recordings used for hypocentre determination barely suffices to determine a hypocentre accurately, and often there is no significant additional information to improve the model.

The input data for the kinematic hypocentre determination are the arrival times of selected groups of elementary waves at the receivers. We understand a group of elementary waves to include the elementary waves whose arrivals cannot be mutually distinguished by a seismologist determining the travel times. For instance, a group of elementary waves may consist of waves refracted in different layers but having close travel times. In such a case we may not be able to identify an observed wave with a layer. In other words, we can determine the group of waves to which the arrival corresponds, but we cannot say to which wave of the group the arrival corresponds. The arrival times cannot be determined exactly and, moreover, they are often not unique. That is why they should be expressed in terms of the (probability) density functions rather than as numbers. Thus, we assume that each group of elementary waves arriving at

* Address: Ke Karlovu 3, 121 16 Praha 2, Czech Republic

a receiver is specified in terms of a density function describing the probability of various arrival times. Without loss of generality, we may assume that each receiver is related to just one group of elementary waves. Several groups of waves arriving at the same receiver may be interpreted as groups arriving at various receivers with the same coordinates. Hereafter we shall assume that there are several receivers, just one group of elementary waves arriving at each receiver. For the sake of simplicity, we shall not consider the cross-correlation between the density functions of arrival times of the individual groups of waves at the individual receivers.

Kinematic hypocentre determination is an inverse problem. In solving this inverse problem we shall follow the formulation of Tarantola and Valette [1] which should be familiar to the reader. The applied probability formulation allows all possible hypocentre locations, which are in conformity with the input arrival times, to be determined. This approach utilizes no linearization, and is very effective if the solution of the forward problem (i.e. determination of the travel time) at the gridpoints of a small grid of hypocentral coordinates does not imply a considerable increase of computer time with respect to the solution of the forward problem at a single point. A small grid is understood here to have at least

$$2^n e \tag{1}$$

points, where n is the number of parameters of the inverse problem. For hypocentre determination, the number of hypocentral parameters is $n = 4$ if all 3 spatial coordinates and the hypocentral time are being determined.

Kinematic hypocentre determination requires the solution of the two-point ray tracing problem, which is the most expensive part of the numerical procedures. To find a ray connecting a point with a receiver by means of the shooting method, it is usually necessary to repeat the tracing of a ray several times. Moreover, it is necessary to find all the rays connecting the point with the receiver, and this must be done for all receivers, and repeated for each new position of the point. Simultaneously, the computed rays must be sorted into individual branches of elementary waves. An elementary wave is one of the waves, of which the total wavefield is theoretically composed. It is usually defined by a specification of the behaviour of its rays. A branch of an elementary wave is its part, which has at the most one arrival at each spatial point. For instance, if there is a triplication of the travel-time curve, the elementary wave must be decomposed into three branches. The sorting of rays into individual branches of elementary waves can hardly be realized, while it is the necessary condition for the application of the paraxial ray approximation [3]. Multiply repeated two-point ray tracing, solved by means of a shooting method, becomes impractical in complex laterally inhomogeneous structures. Instead, it is more suitable to shoot a sufficiently dense system of rays from each receiver, and to interpolate in the net of the computed points. The interpolation formulae may be derived with the help of the Gaussian packet summation method. Besides avoiding the two-point ray tracing and the sorting of rays into individual branches of elementary waves, this algorithm has two great advantages: (a) it allows separation of ray tracing from the actual hypocentre determination, (b) it facilitates application of the above-mentioned probability approach to the nonlinear kinematic hypocentre determination.

To evaluate travel times at any point of a model, interpolation including smoothing, derived from the summation of Gaussian packets [4,5], is applied. A basic knowledge of the paraxial ray approximation and the Gaussian beam method may be useful to the reader. These methods were described in detail by Červený [6] and an ample list of further references is given by Červený [7]. Some notes on the dynamic ray tracing were concisely outlined by Červený et al. [8]. An algorithm for ray tracing and for the evaluation along rays of the quantities necessary for

hypocentre determination has been proposed by Červený et al. [9]. The output files of this "complete ray tracing", containing the quantities computed along rays from the individual receivers, become the input files for the kinematic hypocentre determination algorithm.

For the sake of simplicity, in this paper we consider only smooth models without interfaces. However, the version of the algorithm, proposed for such models, works even in models with structural interfaces, but with a small defect: If a hypocentre is situated close to a structural interface, the density function of the hypocentre location is spread over a larger region than that corresponding to the input data. In other words, not all of the information contained in the input data is utilized in the vicinity of a structural interface. Here, the vicinity has the dimensions of the effective regions of Gaussian packets used.

Both the component and matrix notations will be used. In the case of component notation, the capital-letter indices take values $K,L,\ldots = 1,2$; the lower-case indices take values $k,l,\ldots = 1,2,3$. The Einstein summation convention is applied to both repetitive subscripts and superscripts. Matrices are denoted by boldface letters. In order to distinguish between 2-dimensional and 3-dimensional vectors and matrices, the latter are denoted with a circumflex above the letter. If the same boldface letter is used for 3×3 and 2×2 matrices with and without the circumflex, then the matrix without circumflex ($\mathbf{M}$) is the 2×2 upper left-hand minor of the 3×3 matrix with a circumflex ($\hat{\mathbf{M}}$). Thus M_{JK} are components of $\mathbf{M}$, and M_{jk} those of $\hat{\mathbf{M}}$. The dagger (e.g. $\hat{\mathbf{M}}^{+}$) denotes the Hermitian adjoint (transpose for real-valued matrices).

2. THEORY

2.1 Theoretical travel times

The concept of the travel time is based on the asymptotic ray theory and is not precisely defined for the general wavefield generated by a point source. In this paper, we use the following definition:

Assume that the Fourier transform of the complex-valued Green function can be, at least approximately, expressed in the form of

$$\hat{G}(\hat{\mathbf{x}}, \hat{\mathbf{r}}, \omega) = \sum_{\text{BEW}} \hat{A}(\hat{\mathbf{x}}, \hat{\mathbf{r}}) \exp\{ i\omega\tau(\hat{\mathbf{x}}, \hat{\mathbf{r}})\} \quad , \tag{2}$$

where the summation is over all branches BEW of all elementary waves, see Červený et al. [10], equations 18 and 19. Quantities $\tau(\hat{\mathbf{x}}, \hat{\mathbf{r}})$, corresponding to individual branches of elementary waves, are referred to as the *theoretical travel times* of these branches of elementary waves. Let us now assume that the Green function (2) may be separated into groups of elementary waves. Each such group may again be expressed in the form of (2), where the summation is only over elementary wave branches BEW of the wave group we are interested in.

The *theoretical density function* of the travel time t from $\hat{\mathbf{r}}$ to $\hat{\mathbf{x}}$ of the chosen group of elementary waves is then assumed to be of the form

$$\vartheta(\hat{\mathbf{x}}, \hat{\mathbf{r}}, t) = \sum_{\text{BEW}} w(\hat{\mathbf{A}}, \tau)\, \delta(t - \tau) \quad , \tag{3}$$

where the summation is over all branches BEW of all elementary waves of the considered group. Here $w(\hat{\mathbf{A}}, \tau)$ are the weighting functions. Since travel times $\tau = \tau(\hat{\mathbf{x}}, \hat{\mathbf{r}})$ are not

computed exactly (especially due to the inaccurate model of the medium), equation (3) has to be replaced by a modified form

$$\vartheta(\hat{\mathbf{x}},\hat{\mathbf{r}},t) = \sum_{\text{BEW}} w(\hat{\mathbf{A}},\tau)\,(2\pi)^{-1/2}\,(\delta\tau)^{-1} \exp\left[-\frac{(t-\tau)^2}{2\,(\delta\tau)^2}\right] \quad . \tag{4}$$

Here the probability distributions of the travel-time errors are assumed Gaussian, with standard deviations $\delta\tau = \delta\tau(\hat{\mathbf{x}},\hat{\mathbf{r}})$.

Standard deviations of theoretical travel times. Assume that the model is expressed in terms of an expansion of slowness $u(\hat{\mathbf{x}}) = [v(\hat{\mathbf{x}})]^{-1}$ into a finite number of localized functions $B_\beta(\hat{\mathbf{x}})$, e.g. B-splines,

$$u(\hat{\mathbf{x}}) = \sum_\beta u_\beta B_\beta(\hat{\mathbf{x}}) \quad . \tag{5}$$

Here u_β are the coefficients of this expansion. The travel time may then be expressed as

$$\tau = \int_0^s u(\hat{\mathbf{x}})\,\mathrm{d}s = \sum_\beta u_\beta \int_0^s B_\beta(\hat{\mathbf{x}})\,\mathrm{d}s \quad , \tag{6}$$

where s is the length along the ray. The standard deviation of the theoretical travel time is then

$$\delta\tau = \sqrt{\sum_\alpha \sum_\beta \frac{\partial\tau}{\partial u_\alpha} \left(\delta^2 u\right)_{\alpha\beta} \frac{\partial\tau}{\partial u_\beta}} \tag{7}$$

where

$$\frac{\partial\tau}{\partial u_\alpha} = \int_0^s B_\alpha(\hat{\mathbf{x}})\,\mathrm{d}s \quad , \tag{8}$$

and $\left(\delta^2 u\right)_{\alpha\beta}$ is the covariance matrix of expansion coefficients u_α of the slowness. For instance, if the covariance matrix is assumed in the form of

$$\left(\delta^2 u\right)_{\alpha\beta} = (\delta u)^2\,\delta_{\alpha\beta} \quad , \tag{9}$$

where $\delta_{\alpha\beta}$ is an identity matrix, and standard deviation δu is the same for all expansion coefficients u_β, equation (7) simplifies to read

$$\delta\tau = \delta u \sqrt{\sum_\beta \left(\frac{\partial\tau}{\partial u_\beta}\right)^2} = \delta u \sqrt{\sum_\beta \left[\int_0^s B_\beta(\hat{\mathbf{x}})\,\mathrm{d}s\right]^2} \tag{10}$$

Assuming the basis functions $B_\beta(\hat{\mathbf{x}})$ normalized in such a way that their sum $\sum_\beta B_\beta(\hat{\mathbf{x}})$ is a unit function in the whole model, equation (10) yields

$$\delta\tau \le \delta u \sqrt{\left[\sum_\beta \int_0^s B_\beta(\hat{\mathbf{x}})\,\mathrm{d}s\right]^2} = \delta u\, s \quad , \tag{11}$$

where s is the length of the ray. Simultaneously,

$$\delta\tau \le \delta u \sqrt{b \sum_\beta \int_0^s B_\beta(\hat{\mathbf{x}})\,\mathrm{d}s} = \delta u \sqrt{b\, s} \quad , \tag{12}$$

where b is the maximum possible value of any integral $\int_0^s B_\beta(\hat{\mathbf{x}})\,\mathrm{d}s$. In the case of a B-spline interpolation in a regular 3-D grid with a constant grid step of d, $b \approx (4/9)\,d$. Estimation (11) is stronger than (12) for $s \le b$, and vice versa. In many applications, a simple linear bound

$$\delta\tau \le \Delta\,\tau \quad , \tag{13}$$

to the travel-time standard deviation may suffice. Here the maximum relative error Δ of the travel time may equal either the maximum relative error of the slowness, or the maximum relative error of the velocity in the model.

Weighting functions. Theoretical weighting functions $w(\hat{\mathbf{A}},\tau)$ in (4) may be chosen somewhat arbitrarily. The selection of the form of these weighting functions should be left to seismologists who pick the arrival times, because it should reflect the nature of the input data (arrival times). The arrival times may correspond to the most pronounced phases, or may correspond to the first observed impulse. For instance, some very simple choices are

$$w(\hat{\mathbf{A}},\tau) = 1 \quad , \tag{14a}$$

$$w(\hat{\mathbf{A}},\tau) = \left[\sqrt{\mathrm{tr}\left(\hat{\mathbf{A}}^+\hat{\mathbf{A}}\right)}\right]^a \quad , \tag{14b}$$

$$w(\hat{\mathbf{A}},\tau) = \left[\sqrt{\mathrm{tr}\left(\hat{\mathbf{A}}^+\hat{\mathbf{A}}\right)}\right]^a \exp(-\tau/T) \quad , \tag{14c}$$

and others like these. The simplest option (14a) is assumed to be the most common choice. Weighting function (14b), where $\sqrt{\mathrm{tr}\left(\hat{\mathbf{A}}^+\hat{\mathbf{A}}\right)}$ is the magnitude of the complex-valued vectorial amplitude $\hat{\mathbf{A}}$ and a is a given exponent, may be selected if the arrivals of small amplitudes are not likely to be picked out in the seismic recordings. Weighting function (14c), where T is the specified constant (in time units), prefers first arrivals.

2.2 Summation of Gaussian packets

The sum of Gaussian packets over the space, corresponding to the Green function (2) may be expressed as, [5],

$$\hat{\mathbf{G}}(\hat{\mathbf{x}},\hat{\mathbf{r}},\omega) = \left(\frac{\omega}{2\pi}\right)^{3/2} \sum_{\mathrm{EW}} \sum_{\mathrm{GP}} \Delta^3\gamma\, v \left|\det\left(Q^{\mathrm{R}}_{KL}\right)\right| \sqrt{\det\left[\mathrm{i}\left(\hat{\mathbf{N}}^{\mathrm{R}} - \hat{\mathbf{N}}\right)\right]}\, \hat{\mathbf{A}} \times$$

$$\times \exp\left\{\mathrm{i}\omega\left[\tau_0 + p_k\left(x_k - y_k\right) + \tfrac{1}{2}\left(x_k - y_k\right) N_{kl}\left(x_l - y_l\right)\right]\right\} \quad , \qquad (15)$$

where all the quantities on the right-hand side (except $\hat{\mathbf{x}}$) are taken at the central point $\hat{\mathbf{y}}$ of a Gaussian packet GP of an elementary wave EW. The summation $\Sigma_{\mathrm{EW}}\,\Sigma_{\mathrm{GP}}$ in (15) over Gaussian packets GP of the elementary waves EW we are interested in, may alternatively be expressed as the summation $\Sigma_{\mathrm{BEW}}\,\Sigma_{\mathrm{GPB}}$ over Gaussian packets GPB of the individual branches of these elementary waves and then over the branches BEW, similarly as in (2). The summed packets are concentrated at discrete points $\hat{\mathbf{y}}$ along rays traced from the point source at $\hat{\mathbf{r}}$.

Here τ_0, p_k and N^{R}_{kl} are the travel time, slowness vector and the matrix of second derivatives of the travel time in the x_i-coordinates, taken at point $\hat{\mathbf{y}}$ and corresponding to the point source at $\hat{\mathbf{r}}$. Similarly $\hat{\mathbf{A}} = \hat{\mathbf{A}}(\hat{\mathbf{y}},\hat{\mathbf{r}})$ is the matrix of the complex-valued vectorial ray amplitudes at point $\hat{\mathbf{y}}$, corresponding to the three possible independent orientations of a single force at point source $\hat{\mathbf{r}}$. The transformation Jacobian from ray parameters γ_l to x_k-coordinates is $v\left|\det\left(Q^{\mathrm{R}}_{KL}\right)\right|$.

As in [11] and [5], we assume that the take-off ray parameters γ_1, γ_2 of the traced rays form a locally regular grid with approximately constant grid steps $\Delta\gamma_1$, $\Delta\gamma_2$. Points $\hat{\mathbf{y}}$ along rays (central points of the packets) are assumed to be stored with constant step $\Delta\gamma_3$ in parameter γ_3 along the ray. Volume $\Delta^3\gamma$ of the ray-parameter space, corresponding to a Gaussian packet, is determined by steps $\Delta\gamma_1$, $\Delta\gamma_2$, $\Delta\gamma_3$ in ray parameters when storing the results of the ray tracing into a file. For a locally regular grid we have

$$\Delta^3\gamma = \Delta\gamma_1\,\Delta\gamma_2\,\Delta\gamma_3 \quad . \qquad (16)$$

Independent variable γ_3 along the ray is assumed to be the travel time. In another case (e.g. for γ_3 being the length along the ray), the product of wave-propagation velocity v with the absolute value of the determinant of the 2×2 geometrical spreading matrix

$$Q^{\mathrm{R}}_{JK} = \partial q_J / \partial \gamma_K \qquad (17)$$

should be replaced by the determinant of the 3×3 matrix

$$Q^{\mathrm{R}}_{jk} = \partial q_j / \partial \gamma_k \quad . \qquad (18)$$

For a more detailed description of the quanties in (15) refer to [5].

The 3×3 complex-valued matrix $\hat{\mathbf{N}}$ with a symmetric real part and a positive-definite symmetric imaginary part may be chosen arbitrarily. For the purpose of the kinematic hypocentre determination it is convenient to choose the packets as narrow as possible for two reasons: (a) We have defined the theoretical travel times (see Section 2.1) in terms of the asymptotic ray theory, and the ray-theory limit of the superposition (15) of Gaussian packets is reached for infinitely narrow Gaussian packets. (b) The computation time (unlike the access time to the disk files) of the hypocentral determination is proportional to the volume of the effective regions of Gaussian packets. That is why we shall choose the packets as narrow as possible.

For the given steps $\Delta\gamma_i$, discretization error δ_{MAX} increases with increasing matrix

$$\hat{\mathbf{Z}} = \frac{\omega}{2} \left(\hat{\mathbf{X}}^{\mathrm{R}}\right)^{+} \left(\hat{\mathbf{N}} - \hat{\mathbf{N}}^{\mathrm{R}}\right)^{+} \left(\operatorname{Im} \hat{\mathbf{N}}\right)^{-1} \left(\hat{\mathbf{N}} - \hat{\mathbf{N}}^{\mathrm{R}}\right) \hat{\mathbf{X}}^{\mathrm{R}} \quad , \tag{19}$$

see [5], eq. (42). Here the transformation matrix

$$X^{\mathrm{R}}_{ik} = \partial x_i / \partial \gamma_k = \left(\partial x_i / \partial q_j\right) Q^{\mathrm{R}}_{jk} \tag{20}$$

from ray coordinates to Cartesian coordinates is closely connected with the ray geometrical spreading matrix. Simultaneously, for narrow packets, matrix $\hat{\mathbf{Z}}$ increases with increasing imaginary part $\operatorname{Im} \hat{\mathbf{N}}$ of matrix $\hat{\mathbf{N}}$, i.e. with decreasing width of the packets. Thus, for given $\hat{\mathbf{Z}}$, the narrowest possible Gaussian packets are obtained for

$$\frac{\mathrm{i}\omega}{2}\left(\hat{\mathbf{N}}^{\mathrm{R}} - \hat{\mathbf{N}}\right) = \left[\left(\hat{\mathbf{X}}^{\mathrm{R}}\right)^{+}\right]^{-1} \hat{\mathbf{Z}} \left[\hat{\mathbf{X}}^{\mathrm{R}}\right]^{-1} . \tag{21}$$

For the given steps $\Delta\gamma_i$ and specified limit δ_{MAX} of the maximum discretization error, the largest possible $\hat{\mathbf{Z}}$ is

$$\hat{\mathbf{Z}} = \kappa^2 \begin{bmatrix} (\Delta\gamma_1)^2 & 0 & 0 \\ 0 & (\Delta\gamma_2)^2 & 0 \\ 0 & 0 & (\Delta\gamma_3)^2 \end{bmatrix}^{-1} \quad , \tag{22}$$

see [5], eqs. (40a-d) and (41). Finally, the narrowest possible Gaussian packets are obtained for

$$\frac{\mathrm{i}\omega}{2}\left(\hat{\mathbf{N}}^{\mathrm{R}} - \hat{\mathbf{N}}\right) = \kappa^2 \left[\hat{\mathbf{X}}^{\mathrm{R}} \begin{bmatrix} (\Delta\gamma_1)^2 & 0 & 0 \\ 0 & (\Delta\gamma_2)^2 & 0 \\ 0 & 0 & (\Delta\gamma_3)^2 \end{bmatrix} \left(\hat{\mathbf{X}}^{\mathrm{R}}\right)^{+}\right]^{-1} \quad , \tag{23}$$

see [5], with

$$\kappa^2 = \pi^2 / \ln\left(6/\delta_{\mathrm{MAX}}\right) \quad , \tag{24}$$

where δ_{MAX} is the specified limit of the maximum discretization error, see [11]. Note that δ_{MAX} has effectively no influence on the travel times. In the kinematic hypocentre determination, it only controls the small spurious oscillations of the amplitude of the density function. Thus, there is no reason to choose κ^2 smaller than 2 ($\delta_{MAX} = 4\%$) which is the option preferred by the author. Values up to $\kappa^2 = 3{\cdot}1$ ($\delta_{MAX} = 25\%$) or $\kappa^2 = \pi$ ($\delta_{MAX} = 26\%$) are recommended.

2.3 Interpolation of travel times

In order to use the definition of theoretical travel times from Section 2.1, the superposition of Gaussian packets (15) must take the form of (2). For this purpose, let us assume that the quadratic expansion

$$\tau(\hat{\mathbf{x}},\hat{\mathbf{r}}) = \tau_0 + p_k(x_k - y_k) + \tfrac{1}{2}(x_k - y_k)\, N^{\mathrm{R}}_{kl}\,(x_l - y_l) \quad , \tag{25}$$

of the travel time in the vicinity of point $\hat{\mathbf{y}}$ [3,12] is nearly independent (within the corresponding standard deviation) of the position of point $\hat{\mathbf{y}}$, if point $\hat{\mathbf{y}}$ lies in the effective vicinity of $\hat{\mathbf{x}}$ and if it lies on the same branch of an elementary wave. This condition can be satisfied for fairly narrow Gaussian packets, see (23). Note that quantities τ_0, p_k and N^{R}_{kl} in (25) are taken at point $\hat{\mathbf{y}}$, and have already appeared in (15).

If we replace the theoretical travel times $\tau(\hat{\mathbf{x}},\hat{\mathbf{r}})$ in (2) by their approximations (25), we find the superposition of Gaussian packets (15), written with $\Sigma_{\mathrm{BEW}}\,\Sigma_{\mathrm{GPB}}$ instead of $\Sigma_{\mathrm{EW}}\,\Sigma_{\mathrm{GP}}$, equivalent to (2), with

$$\hat{\mathbf{A}}(\hat{\mathbf{x}},\mathbf{r}) = \sum_{\mathrm{GPB}} (2\pi)^{-3/2}\, \Delta^3 \gamma\, v \left|\det\left(Q^{\mathrm{R}}_{KL}\right)\right| \sqrt{\det\left[\,\mathrm{i}\,\omega\left(\hat{\mathbf{N}}^{\mathrm{R}} - \hat{\mathbf{N}}\right)\right]}\; \hat{\mathbf{A}}\,(\hat{\mathbf{y}},\hat{\mathbf{r}}) \times$$

$$\times \exp\left\{\,\mathrm{i}\,\omega\,\tfrac{1}{2}(x_k - y_k)\left(N_{kl} - N^{\mathrm{R}}_{kl}\right)(x_l - y_l)\right\} \quad . \tag{26}$$

Here the summation is over all Gaussian packets GPB of a branch BEW of an elementary wave. Equation (26) may be interpreted as the interpolation with smoothing of function $\hat{\mathbf{A}}(\hat{\mathbf{x}},\hat{\mathbf{r}})$,

$$\hat{\mathbf{A}}(\hat{\mathbf{x}},\hat{\mathbf{r}}) = \sum_{\mathrm{GPB}} P(\hat{\mathbf{x}},\hat{\mathbf{y}})\, \hat{\mathbf{A}}(\hat{\mathbf{y}},\hat{\mathbf{r}}) \quad , \tag{27}$$

by means of the interpolation functions

$$P(\hat{\mathbf{x}},\hat{\mathbf{y}}) = (2\pi)^{-3/2} \Delta^3 \gamma\, v \left|\det\left(Q^{\mathrm{R}}_{KL}\right)\right| \sqrt{8 \det \hat{\mathbf{Y}}}\, \exp\left\{-(x_k - y_k)\, Y_{kl}\,(x_l - y_l)\right\} \quad . \tag{28}$$

Here we have introduced, for the sake of conciseness, matrix

$$\hat{\mathbf{Y}} = \frac{\mathrm{i}\,\omega}{2}\left(\hat{\mathbf{N}}^{\mathrm{R}} - \hat{\mathbf{N}}\right) \quad . \tag{29}$$

For possible choices of $\hat{\mathbf{Y}}$, refer to (23).

The sum $\Sigma_{GPB} P(\hat{\mathbf{x}},\hat{\mathbf{y}})$ of the interpolation functions is approximately a unit function in all regions covered by the corresponding branch of the elementary wave, and decreases to zero in the shadow zones of the branch. The deviation of the sum from the unit function is controlled by κ^2, see (24), and should not exceed δ_{MAX} in the regions regular from the point of view of the ray theory. In irregular regions, the deviation may increase in order, but the functions should still approximately maintain their interpolation properties, see Fig. 1a. The widths of the peripheries of shadow zones, where the sum $\Sigma_{GPB} P(\hat{\mathbf{x}},\hat{\mathbf{y}})$ of the interpolation functions smoothly increases from 0 to 1 instead of being zero, is of the same order as the distance between central points $\hat{\mathbf{y}}$ of the Gaussian packets, see Fig. 1a. Similar behaviour also applies to the interpolation (27) of ray amplitudes, see Fig. 1b, and to the interpolation of any other quantity smoothly varying inside the illuminated region of the branch of the elementary wave. Consequently, interpolation functions (28) may be used not only to interpolate ray amplitudes, see (27), but also to interpolate any other quantity related to the branch of the elementary wave and sufficiently smooth within the effective regions of the Gaussian packets, if the same accuracy as for the amplitudes (up to tens per cent) is acceptable.

That is why we can also use interpolation functions (28) to interpolate weighting functions $w(\hat{\mathbf{A}},\tau)$. In other words, assuming that functions $\hat{\mathbf{A}}(\hat{\mathbf{x}},\hat{\mathbf{r}})$ and $w(\hat{\mathbf{A}},\tau)$ are sufficiently smooth within the effective regions of Gaussian packets, the interpolation with smoothing of function $\hat{\mathbf{A}}(\hat{\mathbf{x}},\hat{\mathbf{r}})$ may be replaced by the interpolation with smoothing of function $w(\hat{\mathbf{A}}(\hat{\mathbf{x}},\hat{\mathbf{r}}),\tau)$,

$$w(\hat{\mathbf{A}}(\hat{\mathbf{x}},\hat{\mathbf{r}}),\tau) = \sum_{GPB} P(\hat{\mathbf{x}},\hat{\mathbf{y}})\, w(\hat{\mathbf{A}}(\hat{\mathbf{y}},\hat{\mathbf{r}}),\tau) \quad . \tag{30}$$

Then, substituting (30) into (4), we arrive at

$$\vartheta(\hat{\mathbf{x}},\hat{\mathbf{r}},t) = \sum_{BEW}\sum_{GPB} P(\hat{\mathbf{x}},\hat{\mathbf{y}})\, w(\hat{\mathbf{A}},\tau)\,(2\pi)^{-1/2}(\delta\tau)^{-1}\exp\left[-\frac{(t-\tau)^2}{2(\delta\tau)^2}\right] \quad . \tag{31}$$

Now, we may rearrange summation $\Sigma_{BEW}\Sigma_{GPB}$ back to $\Sigma_{EW}\Sigma_{GP}$ to get the final expression for the theoretical density function of the travel time,

$$\vartheta(\hat{\mathbf{x}},\hat{\mathbf{r}},t) = \sum_{EW}\sum_{GP} (2\pi)^{-2} A^3 \gamma\, v \left|\det\left(Q^{R}_{KL}\right)\right| \sqrt{8\det\hat{\mathbf{Y}}}\; w(\hat{\mathbf{A}},\tau)\,(\delta\tau)^{-1} \times$$

$$\times \exp\left\{-(x_k - y_k)\, Y_{kl}\,(x_l - y_l)\right\} \exp\left[-\frac{(t-\tau)^2}{2(\delta\tau)^2}\right] \quad . \tag{32}$$

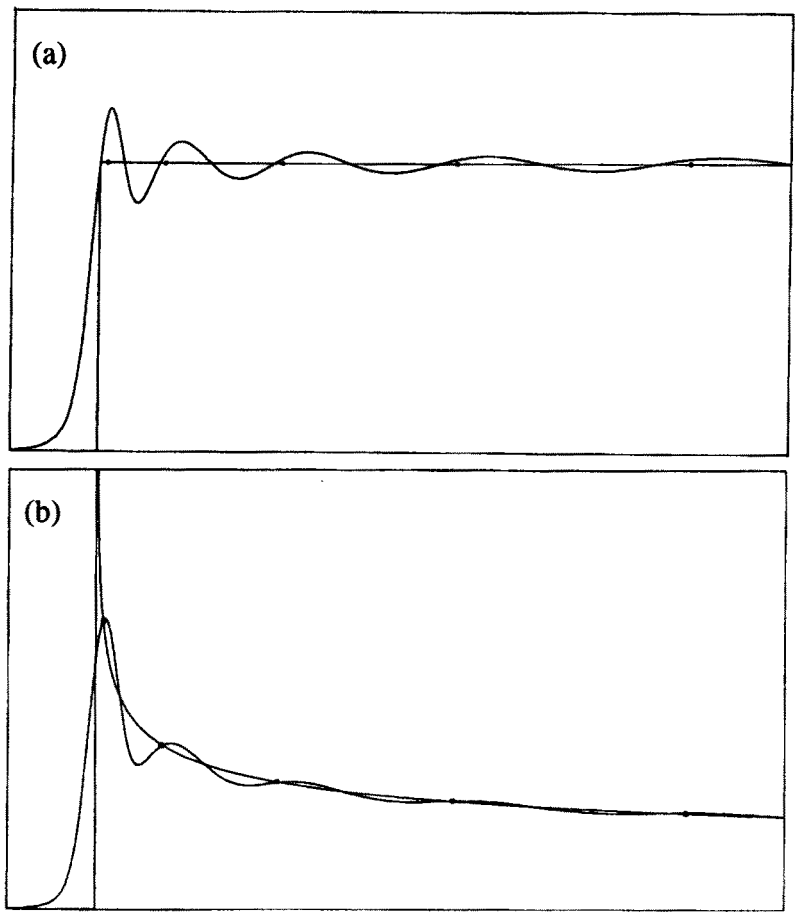

Fig. 1. Example of interpolation in the vicinity of a caustic. (a) The sum $\sum_{GPB} P(\hat{x},\hat{y})$ of interpolation functions compared with the characteristic function of the illuminated region. The characteristic function is zero in the shadow zone left of the caustic, and unity in the illuminated region. The dots in the graph of the characteristic function denote central points $\hat{y}$ of Gaussian interpolation functions $P(\hat{x},\hat{y})$. The distribution of the central points corresponds to a constant step in the ray take-off parameter, the caustic ray being situated in the middle of two successive axial rays. The width of the packets is given by $\kappa^2 = 2$, see (24). Assuming that the packets are very narrow, the horizontal scale is fully determined by the density of the central points. (b) Interpolated ray amplitude (27) compared with the ray amplitude curve. The dots in the graph of the ray amplitude curve denote the central points of the Gaussian interpolation functions. In other words, the dots are the discrete values being interpolated. The horizontal scale is determined by the central points.

2.4 Kinematic hypocentre determination

In this section we closely follow the formulation of Tarantola and Valette [1]. The inverse problem of hypocentre determination has 4 parameters and N data: 3 hypocentral coordinates $\hat{\mathbf{x}}$ and hypocentral time x_4, and N arrival times $T_1,T_2,...,T_N$ of several groups of elementary waves at several receivers. As mentioned in the Introduction, we may assume, without loss of generality, that each arrival time T_ν corresponds to one receiver. In this way, two receivers may have the same coordinates while recording two different groups of elementary waves.

Null information. Assume that the *null information density function* is not dependent on arrival times,

$$\mu = \mu(\hat{\mathbf{x}},x_4) \quad . \tag{33}$$

With no information, all values of arrival times are equally probable. For instance, in Cartesian coordinates, no information corresponds to constant density function $\mu(\hat{\mathbf{x}},x_4)$, e.g. $\mu(\hat{\mathbf{x}},x_4) = 1$.

A priori information. For hypocentre determination, no a priori information is probably available. Thus, the *a priori density function* of the hypocentral parameters is probably equal to the null density function,

$$\rho_0(\hat{\mathbf{x}},x_4) = \mu(\hat{\mathbf{x}},x_4) \quad . \tag{34}$$

However, any prior information on the hypocentral parameters may be included.

Input data. Input data, i.e. arrival times $T_1,T_2,...,T_N$, are described in terms of the *data density function* $\rho_D(T_1,T_2,...,T_N)$. Let us assume that the arrival times at the individual receivers are mutually independent, i.e. that the data density function may be expressed in the form of

$$\rho_D(T_1,T_2,...,T_N) = \rho_1(T_1)\,\rho_2(T_2)\,...\,\rho_N(T_N) \quad . \tag{35}$$

The whole prior density function describing both the data and the prior information on the parameters is then

$$\rho(\hat{\mathbf{x}},x_4,T_1,T_2,...,T_N) = \rho_0(\hat{\mathbf{x}},x_4)\,\rho_D(T_1,T_2,...,T_N) \quad . \tag{36}$$

The *partial data density functions* $\rho_\nu(T_\nu)$, for $\nu = 1,2,...,N$, are general functions. From the point of view of describing them by a small number of parameters, it seems convenient to specify them in the form of

$$\rho_\nu(T_\nu) = \sum_{\mu=1}^{M_\nu} W_{\mu\nu}\,(2\pi)^{-1/2}\left(\Delta T_{\mu\nu}\right)^{-1} \exp\left[-\frac{\left(T_\nu - T_{\mu\nu}\right)^2}{2\left(\Delta T_{\mu\nu}\right)^2}\right] \quad , \tag{37}$$

by several (i.e. M_ν) triplets $T_{\mu\nu}$, $\Delta T_{\mu\nu}$, $W_{\mu\nu}$ of real numbers. Here $T_{\mu\nu}$, with $\mu = 1,2,...,M_\nu$, are possible mean values of arrival times at the ν-th receiver, picked from seismic records. The case of $M_\nu > 1$ corresponds to a situation when a seismologist

hesitates to choose between two or more possibilities. Quantities $\Delta T_{\mu\nu}$ are standard deviations of the picked arrival times $T_{\mu\nu}$, and $W_{\mu\nu}$ are the corresponding weighting factors, which are important for $M_\nu > 1$. It is convenient to choose, at each receiver ν, the maximum weighting factor $W_\nu^{\mathrm{MAX}} = \max\left(W_{\mu\nu},\ \mu = 1,2,...,M_\nu\right)$ equal to 1.

Theoretical relations. Theoretical relations are described in terms of the *theoretical density function*

$$\theta(\hat{\mathbf{x}},x_4,T_1,T_2,...,T_N) = \mu(\hat{\mathbf{x}},x_4)\,\theta_1(\hat{\mathbf{x}},x_4,T_1)\,\theta_2(\hat{\mathbf{x}},x_4,T_2)\,...\,\theta_N(\hat{\mathbf{x}},x_4,T_N)\ , \tag{38}$$

where

$$\theta_\nu(\hat{\mathbf{x}},x_4,T_\nu) = \vartheta(\hat{\mathbf{x}},\hat{\mathbf{r}}_\nu,T_\nu - x_4)\ . \tag{39}$$

The relative probability of $\vartheta(\hat{\mathbf{x}},\hat{\mathbf{r}}_\nu,\tau)$ of travel time τ from receiver $\hat{\mathbf{r}}_\nu$ to point $\hat{\mathbf{x}}$ is given by (32).

A posteriori state of information. The *posterior density function* is given by, [1],

$$\sigma(\hat{\mathbf{x}},x_4,T_1,T_2,...,T_N) = \rho(\hat{\mathbf{x}},x_4,T_1,T_2,...,T_N)\,\theta(\hat{\mathbf{x}},x_4,T_1,T_2,...,T_N)\,/\mu(\hat{\mathbf{x}},x_4)\ , \tag{40}$$

which, after the substitution of (36), (35) and (38), becomes

$$\sigma(\hat{\mathbf{x}},x_4,T_1,T_2,...,T_N) = \rho_0(\hat{\mathbf{x}},x_4)\,\rho_1(T_1)\,\rho_2(T_2)\,...\,\rho_N(T_N)\,\times$$

$$\times\,\theta_1(\hat{\mathbf{x}},x_4,T_1)\,\theta_2(\hat{\mathbf{x}},x_4,T_2)\,...\,\theta_N(\hat{\mathbf{x}},x_4,T_N) \tag{41}$$

A posteriori marginal density function. We are not interested in the whole posterior density function (40), but just in the posterior *marginal density function*

$$\sigma_{\mathrm{P}}(\hat{\mathbf{x}},x_4) = \int_{-\infty}^{+\infty} \mathrm{d}T_1 \int_{-\infty}^{+\infty} \mathrm{d}T_2 ... \int_{-\infty}^{+\infty} \mathrm{d}T_N\ \sigma(\hat{\mathbf{x}},x_4,T_1,T_2,...,T_N)\ . \tag{42}$$

Due to (41), it can be expressed as product

$$\sigma_{\mathrm{P}}(\hat{\mathbf{x}},x_4) = \rho_0(\hat{\mathbf{x}},x_4)\,\sigma_1(\hat{\mathbf{x}},x_4)\,\sigma_2(\hat{\mathbf{x}},x_4)\,...\,\sigma_N(\hat{\mathbf{x}},x_4)\ , \tag{43}$$

where

$$\sigma_\nu(\hat{\mathbf{x}},x_4) = \int_{-\infty}^{+\infty} \mathrm{d}T_\nu\ \rho_\nu(T_\nu)\,\theta_\nu(\hat{\mathbf{x}},x_4,T_\nu) = \int_{-\infty}^{+\infty} \mathrm{d}T_\nu\ \rho_\nu(T_\nu)\,\vartheta(\hat{\mathbf{x}},\hat{\mathbf{r}}_\nu,T_\nu - x_4) \tag{44}$$

are posterior *partial marginal density functions.* After insertion of (32) and (37), these integrals may be computed analytically to yield

$$\sigma_v(\hat{\mathbf{x}}, x_4) = \sum_{\text{EW}} \sum_{\text{GP}} (2\pi)^{-3/2} \Delta^3 \gamma v \left| \det\left(Q_{KL}^{\text{R}}\right) \right| \sqrt{8 \det \hat{\mathbf{Y}}}\, w(\hat{\mathbf{A}}, \tau) \times$$

$$\times \exp\left\{-\left(x_k - y_k\right) Y_{kl} \left(x_l - y_l\right)\right\} \sum_{\mu=1}^{M_v} W_{\mu v} (2\pi)^{-1/2} \left[(\delta\tau)^2 + \left(\Delta T_{\mu v}\right)^2\right]^{-1/2} \times$$

$$\times \exp\left[-\frac{\left(T_{\mu v} - x_4 - \tau\right)^2}{2\left[(\delta\tau)^2 + \left(\Delta T_{\mu v}\right)^2\right]}\right] , \tag{45}$$

where τ is given by (25). For the purposes of the kinematic hypocentre determination algorithm, it is probably convenient to substitute the posterior partial marginal density functions (45) by the posterior *partial modified density functions*

$$\sigma_v(\hat{\mathbf{x}}, x_4) = \sum_{\text{EW}} \sum_{\text{GP}} (2\pi)^{-3/2} \Delta^3 \gamma v \left| \det\left(Q_{KL}^{\text{R}}\right) \right| \sqrt{8 \det \hat{\mathbf{Y}}}\, w(\hat{\mathbf{A}}, \tau) \times$$

$$\times \exp\left\{-\left(x_k - y_k\right) Y_{kl} \left(x_l - y_l\right)\right\} \sum_{\mu=1}^{M_v} W_{\mu v} \exp\left[-\frac{\left(T_{\mu v} - x_4 - \tau\right)^2}{2\left[(\delta\tau)^2 + \left(\Delta T_{\mu v}\right)^2\right]}\right] \tag{46}$$

This modification is also used by Moser et al. [2]. If the input data and their deviations are mutually consistent, the maximum of the posterior *modified density function* (43) with (46) should approximately correspond to the product of weighting factors $w(\hat{\mathbf{A}}, \tau)$ and $W_{\mu v}$. Especially if unit option (14a) is selected for $w(\hat{\mathbf{A}}, \tau)$, the global maximum of (43) with (46) should be roughly the product $W_1^{\text{MAX}} W_2^{\text{MAX}} \ldots W_N^{\text{MAX}}$. Here all the maximum weights W_v^{MAX} have been advised above to be chosen equal to 1, see (37). This property may be used to check the consitency of the input data. If the maximum of the posterior modified density function (43) with (46) is considerably lower than that estimated from the weighting factors, there is probably no hypocentre corresponding to the given model and travel-time data, in the specified volume. Equations (43) and (46) are the basic relations for the proposed kinematic hypocentre determination algorithm.

3. ALGORITHM

The nonlinear kinematic hypocentre determination algorithm may be divided into three basic steps: (1) Complete ray tracing; (2) Preparation of disk files for travel-time interpolation; (3) Actual kinematic hypocentre determination. In the computer implementation of the algorithm, these three steps will probably correspond to three individual basic programs, connected by means of data files.

3.1 Complete ray tracing

In a given model, a sufficiently dense system of rays corresponding to each group of elementary waves recorded at the receiver, is computed for each receiver position. A possible ray tracing algorithm has been described by Červený et al. [9]. Quantities such as the coordinates, travel time, slowness vector, basis vectors of the ray-centred coordinate system, paraxial ray propagator matrix and amplitudes, are stored, with a specified step along all rays, in a file.

3.2 Disk files for travel-time interpolation

The quantities stored at points along rays in the preceding step are used to evaluate the following 20 quantities at these points: 3 coordinates x_k, travel time τ_0, 3 components p_k of the slowness vector, 6 components of the 3×3 real symmetric matrix $\frac{1}{2}\hat{\mathbf{N}}^{\mathrm{R}}$, 6 components of the 3×3 positive-definite real symmetric matrix $\hat{\mathbf{Y}}$ (see (29)), and the weighting factor

$$(2\pi)^{-3/2} \Delta^3 \gamma \, v \left| \det \left(Q_{KL}^{\mathrm{R}} \right) \right| \sqrt{8 \det \hat{\mathbf{Y}}} \, w(\hat{\mathbf{A}}, \tau) \quad , \tag{47}$$

see (46) and (14a–c). For option (23), the weighting factor (47) reads

$$(\pi)^{-3/2} \kappa^3 \Delta^3 \gamma (\Delta\gamma_1 \, \Delta\gamma_2 \, \Delta\gamma_3)^{-1} \, w(\hat{\mathbf{A}}, \tau) \quad . \tag{48}$$

Matrix $\hat{\mathbf{Y}}$ may also be chosen complex-valued and described by means of 12 reals, but the author presently sees no reason to do so. The 21st quantity $\delta\tau$ has to be stored only if it does not depend on the above 20 quantities, see (5)–(13). For each receiver, these quantities are stored in the receiver file. From now on each receiver is fully specified by the corresponding filename. Note that the choice of matrix $\hat{\mathbf{Y}}$, see (29), is free, but option (23) seems to be the most convenient for our ray-theory interpretation of the travel time.

3.3 Kinematic hypocentre determination

Input data. Assuming that the receiver file described in the previous section is available for each receiver, we need the following four data sets for each hypocentre determination:

1. Coefficients of the equation for the standard deviation $\delta\tau$ of the theoretical travel time, see Section 2.1. These data specify the accuracy of the model and, therefore, are probably the same for all hypocentre determinations. For instance, they may consist of coefficients δu and b of equations (11) and (12), or of coefficient Δ of equation (13).
2. The space region and the time interval in which we are looking for a hypocentre. This region in 4-D space of 3 hypocentral coordinates x_k and hypocentral time x_4 may be specified in terms of coordinate limits

$$x_1^{MIN}, x_1^{MAX}, x_2^{MIN}, x_2^{MAX}, x_3^{MIN}, x_3^{MAX}, x_4^{MIN}, x_4^{MAX} \quad . \tag{49}$$

3. The division of the 4-D region into uniform rectangular blocks, specified by numbers

$$N_1, N_2, N_3, N_4 \tag{50}$$

of blocks in the directions of the individual coordinate axes. These data may be generated automatically by a program in order to achieve maximum efficiency of the algorithm.

4. For each receiver ν, the possible values $T_{\mu\nu}$ of the arrival time, their standard deviations $\Delta T_{\mu\nu}$ and weights $W_{\mu\nu}$, for $\mu = 1,2,...,M_\nu$ (see (37)).

Flattening the marginal density function. Marginal density function (43) is to be evaluated at gridpoints situated in the centres of the individual rectangular blocks. We denote the grid intervals by d_κ, for $\kappa = 1,2,3,4$,

$$d_\kappa = \left(x_\kappa^{MAX} - x_\kappa^{MIN}\right)/N_\kappa \quad . \tag{51}$$

The partial modified density functions (46) are roughly concentrated close to some 3-D hypersurfaces in the 4-D hypocentral time-space. The standard halfwidths of these functions measured in the direction of time axis x_4 are

$$\Delta_4 = \sqrt{(\delta\tau)^2 + (\Delta T_{\mu\nu})^2} \quad . \tag{52}$$

The standard halfwidths measured in the directions of space axes x_k are

$$\Delta_k = \Delta_4 / p_k \quad . \tag{53}$$

In order not to lose the marginal density function in gaps between the gridpoints, relations

$$\Delta_k \geq d_\kappa/2 \ , \quad \kappa = 1,2,3,4 \tag{54}$$

should roughly hold true. For this reason we replace the standard deviation $\sqrt{(\delta\tau)^2 + (\Delta T_{\mu\nu})^2}$ in (46) by the value of

$$\max\left(\sqrt{(\delta\tau)^2 + (\Delta T_{\mu\nu})^2},\ d_1 p_1/2,\ d_2 p_2/2,\ d_3 p_3/2,\ d_4/2\right) \ . \tag{55}$$

Functions (46) are then spread sufficiently to be able to identify them from the discrete values at the gridpoints.

A *single iteration.* The grid values of marginal density function (43) are stored in one real array, while partial modified density functions (46) are accumulated in a second real array. The dimensions of these arrays must at least equal to the number of gridpoints $N1, N2, N3, N4$. Initially, prior density function $\rho_0(\hat{\mathbf{x}}, x_4)$ is stored in the first array. After finishing the accumulation of each partial modified density function (46) in the second array, the first array is updated by multiplication with the second array. Each partial

modified density function (46) in the second array is formed by the summation $\Sigma_{EW}\,\Sigma_{GP}$ over all points $\hat{y}$ stored in the corresponding receiver file. Note that the individual contributions to function (46) are effectively nonzero only at the gridpoints in the vicinity of points $(x_k, x_4) = (y_k, T_{\mu\nu} - \tau)$. Moreover, we need not evaluate function (46) in the second array at gridpoints which already have zero values in the first array.

Complete hypocentre determination. If the grid is sufficiently dense, the support (region of nonzero values) of marginal density function (43) is smaller than the previous region (49). Limits (49) are then updated to close this support tightly. The numbers (50) of grid lines may also be updated. The new grid must be denser than the previous one. The next iteration is then performed for a new, more moderate flattening (55) of the marginal density function. These iterations are terminated when the grid is so dense that inequalities (54) are satisfied. At that time flattenings (55) have no effect, and the digitization of marginal density function (43) is sufficiently dense. If we are interested only in hypocentral coordinates, marginal density function (43) may be summed over hypocentral time x_4. If the hypocentre location problem has several solutions, i.e. if marginal density function (43) is concentrated in several separated regions, we run into difficulties. Only one of the regions is then selected and covered by the grid, the other regions being held aside. After this branching, the algorithm continues for the selected region. When marginal density function (43) is determined in the selected region, the iterations of the other branches take place.

Acknowledgements: The preliminary, Czech version [13] of this paper was written during the author's employment in the Institute of Geology and Geotechnics of the Czechoslovak Academy of Sciences. The English version of this paper was written during the author's fellowship in the laboratories of Energy, Mines and Resources, Canada. The author is indebted to Dr. Don White and Dr. Jan Šílený for their careful reading of the manuscript.

Received: 26. 1. 1993
Revised: 17. 12. 1993

References

[1] A. Tarantola, B. Valette: Inverse problems = quest for information. J. Geophys., **50** (1982), 159–170.

[2] T-J. Moser, T. Van Eck, G. Nolet: Hypocenter determination in strongly heterogeneous Earth models using the shortest path method. J. geophys. Res., **97B** (1992), 6563–6572.

[3] V. Červený, L. Klimeš, I. Pšenčík: Paraxial ray approximations in the computation of seismic wavefields in inhomogeneous media. Geophys. J. R. astr. Soc., **79** (1984), 89–104.

[4] L. Klimeš: The relation between Gaussian beams and Maslov asymptotic theory. Studia geoph. et geod., **28** (1984), 237–247.

[5] L. Klimeš: Gaussian packets in the computation of seismic wavefields. Geophys. J. int., **99** (1989), 421–433.

[6] V. Červený: The application of ray tracing to the numerical modeling of seismic wavefields in complex structures. In: G. Dohr (Ed.), Seismic Shear Waves, Part A: Theory, (K. Helbig, S. Treitel (Eds.), Handbook of Geophysical Exploration, Section I: Seismic Exploration, Vol. 15A), Geophysical Press, London, 1985, pp.1–124.

[7] V. Červený: Gaussian beam synthetic seismograms. J. Geophys., **58** (1985), 44–72.

[8] V. Červený, L. Klimeš, I. Pšenčík: Applications of the dynamic ray tracing. Phys. Earth planet. Interiors, **51** (1988), 25–35.

[9] V. Červený, L. Klimeš, I. Pšenčík: Complete seismic-ray tracing in three-dimensional structures. In: D. J. Doornbos (Ed.), Seismological Algorithms, Academic Press, New York 1988, pp. 89–168.

[10] V. Červený, J. Pleinerová, L. Klimeš, I. Pšenčík: High-frequency radiation from earthquake sources in laterally varying layered structures. Geophys. J. R. astr. Soc., **88** (1987), 43–79.

[11] L. Klimeš: Discretization error for the superposition of Gaussian beams. Geophys. J. R. astr. Soc., **86** (1986), 531–551.

[12] L. Klimeš: Expansion of a high-frequency time-harmonic wavefield given on an initial surface into Gaussian beams. Geophys. J. R. astr. Soc., **79** (1984), 105–118.

[13] L. Klimeš: Kinematic location of a seismic hypocentre. Acta montana, No. 75, Instit. Geol. Geotechn. Czechosl. Acad. Sci., Praha 1987, pp.51–64 (in Czech).