OPTIMIZATION OF THE SHAPE OF GAUSSIAN BEAMS

K. ŽÁČEK

Department of Geophysics, Faculty of Mathematics and Physics, Charles University, Ke Karlovu 3, 121 16 Praha 2, Czech Republic (zacek@karel.troja.mff.cuni.cz, karelzacek@yahoo.com)

Received: September 27, 2005; Revised: February 12, 2006; Accepted: March 23, 2006

ABSTRACT

The applicability and accuracy of the Gaussian beam method depend on the proper choice of the shape of beams. Gaussian beams become inaccurate solutions of the elastodynamic equation if the velocity field changes considerably within the beam width. We present a procedure of determining the optimum initial shape of Gaussian beams based on minimizing the average squared widths of Gaussian beams and smoothing the distribution of the optimum parameters of Gaussian beams on the Hamiltonian hypersurface in the phase-space. The original method of smoothing represents an essential part of the algorithm, which is designed particularly for the optimization of the shape of Gaussian beams for Gaussian beam or packet migrations.

Key words: Gaussian beam, Hamiltonian hypersurface, optimization, smoothing, Sobolev norm

1. INTRODUCTION

Gaussian beams (GBs) represent high-frequency asymptotic time-harmonic solutions of the elastodynamic equation, which are concentrated close to rays (e.g., *Babich and Popov, 1989*; *Červený et al., 1982*; *Červený and Pšenčík, 1983*; *Popov, 1982*). The distribution of the amplitude of the principal component of the displacement in the profile perpendicular to the ray is Gaussian (bell-shaped). The great advantage of the Gaussian beam method is that GBs are regular along the whole ray, even at caustics.

However, the applicability and accuracy of the Gaussian beam method depend on the proper choice of the shape of beams. It is necessary to keep GBs narrow in relation to the velocity changes in the model, because GBs become inaccurate solutions of the elastodynamic equation if the velocity field changes considerably within the beam width. Unfortunately, if GBs are too narrow, they quickly increase in width as they propagate. Thus, we can use neither too narrow nor too wide beams as the initial choice of the shape of GBs. Furthermore, in a complex structure, we cannot judge solely from the final width of the beam whether the beam is or is not a reasonably accurate solution of the elastodynamic equation. The beam must be sufficiently narrow along the whole part of the ray path between the source and the point under consideration.

All these requirements force us to use a more sophisticated theory in the choice of the initial shape of GBs than just a wild guess or some kind of empirical rules. *Klimeš (1989)* proposed a procedure, which is followed in this paper, based on minimizing the integral of a certain expression along a fixed part of the beam's central ray. This approach allows us to minimize not only the width of GBs, but also the quadratic variations of the complex-valued phase along an arbitrary surface, along a structural interface, or along a wavefront tangent plane.

In a complex structure, for various positions of the initial point of the beam's central ray (e.g., source), for various take-off angles of the beam's central ray, and for various travel times, the optimum initial parameters of GBs can vary considerably. This can bring about serious problems in the decomposition of the wave field into Gaussian beams or packets (*Žáček, 2003, 2006*). The decomposition of the wave field into beams or packets, whose shape depends on the coordinate of the intersection of the central ray with the profile and on the travel time, is strictly mathematically speaking incorrect. In such a case, we can achieve satisfactory results only for a sufficiently smooth distribution of the initial parameters. Hence, we should be able to simultaneously optimize and smooth the distribution of the initial parameters of GBs.

2. SPECIFICATION OF SOME USED QUANTITIES

In the case of the component notation, the upper-case indices take the values $I, J, \ldots = 1,2$ and the lower-case indices take the values $i, j, \ldots = 1,2,3$. The Einstein summation over the pairs of identical indices is used. The matrices are denoted by boldface letters (e.g., **A**) or by means of their components (e.g., A_{ij}).

We denote by A_j the amplitude and by θ the complex-valued phase of a frequency-domain Gaussian beam

$$g_j = A_j \exp(\mathrm{i}\omega\theta), \tag{1}$$

where i is the imaginary unit and ω is the circular frequency. In ray-centred coordinates q_j (e.g., *Popov and Pšenčík, 1978*; *Červený, 2001*), where q_3 is an independent variable along the ray and q_j are Cartesian coordinates in the phase-front tangent plane, the quadratic Taylor expansion (paraxial approximation) of the phase has the form of

$$\theta(q_j) = \tau(q_3) + \frac{1}{2} q_K M_{KL}(q_3) q_L, \tag{2}$$

where τ is the travel time along the central ray and **M** is the second differential of the phase along the plane tangent to the phase-front.

Matrix **M** consists of a symmetric real part **R** and of a positive-definite symmetric imaginary part **Y**,

$$\mathbf{M} = \mathbf{R} + \mathrm{i}\mathbf{Y}, \tag{3}$$

where **Y** determines the Gaussian beam width and **R** defines the curvature of the phasefront. Matrix **M** may be also expressed as

$$\mathbf{M} = \mathbf{P}\mathbf{Q}^{-1} , \tag{4}$$

where

$$\begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix} = \begin{pmatrix} \mathbf{1} \\ \mathbf{M} \end{pmatrix} \mathbf{Q} \tag{5}$$

is the solution of the dynamic ray-tracing system (e.g., *Červený, 2001*)

$$\frac{\mathrm{d}}{\mathrm{d}\tau} \begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix} = \begin{pmatrix} \mathbf{0} & v^2 \mathbf{1} \\ -v^{-1}\mathbf{V} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix} , \tag{6}$$

v is the propagation velocity, **V** is the second differential of the propagation velocity along the phasefront tangent plane, **0** and **1** being zero and identity 2 × 2 matrices.

Any solution of the dynamic ray-tracing system with the initial conditions

$$\begin{pmatrix} \mathbf{Q}\left(q_3^{(0)} \right) \\ \mathbf{P}\left(q_3^{(0)} \right) \end{pmatrix} = \begin{pmatrix} \mathbf{Q}_0 \\ \mathbf{P}_0 \end{pmatrix} = \begin{pmatrix} \mathbf{1} \\ \mathbf{M}_0 \end{pmatrix} \mathbf{Q}_0 \tag{7}$$

may be expressed as

$$\begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix} = \Pi \begin{pmatrix} \mathbf{Q}_0 \\ \mathbf{P}_0 \end{pmatrix} , \tag{8}$$

where the paraxial-ray propagator matrix

$$\Pi\left(q_3, q_3^{(0)} \right) = \begin{pmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \\ \mathbf{P}_1 & \mathbf{P}_2 \end{pmatrix} \tag{9}$$

is the fundamental 4 × 4 matrix of the solutions of the dynamic ray-tracing system. $\mathbf{Q}_1$ and $\mathbf{P}_1$ are solutions of the dynamic ray-tracing system for the normalized plane wavefront initial conditions

$$\begin{pmatrix} \mathbf{Q}_1 \\ \mathbf{P}_1 \end{pmatrix} = \begin{pmatrix} \mathbf{1} \\ \mathbf{0} \end{pmatrix} \tag{10}$$

and $\mathbf{Q}_2$ and $\mathbf{P}_2$ are solutions of the dynamic ray-tracing system for the normalized point source initial conditions

$$\begin{pmatrix} \mathbf{Q}_2 \\ \mathbf{P}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{1} \end{pmatrix} . \tag{11}$$

3. MINIMIZATION OF THE OBJECTIVE FUNCTION

Klimeš (1989) proposed a procedure for determining the shape of Gaussian beams so that they minimize the integral of certain expression along a fixed part of the beam's central ray. He introduced an objective function, which roughly expresses the dependency of the error of the GB solution on the shape of the GB. The general form of the minimized objective function is

$$T(\mathbf{G}) = \int_{q_3^{(1)}}^{q_3^{(2)}} \mathrm{tr}\left\{\mathbf{G}(q_3)\,\mathrm{Re}\left[\mathbf{\Psi}(q_3)\right]\right\} \mathrm{d}q_3 \,, \tag{12}$$

where

$$\mathrm{Re}\left[\mathbf{\Psi}(q_3)\right] = \begin{pmatrix} \left[\mathbf{Y}(q_3)\right]^{-1} & \left[\mathbf{Y}(q_3)\right]^{-1}\mathbf{R}(q_3) \\ \mathbf{R}(q_3)\left[\mathbf{Y}(q_3)\right]^{-1} & \mathbf{Y}(q_3)+\mathbf{R}(q_3)\left[\mathbf{Y}(q_3)\right]^{-1}\mathbf{R}(q_3) \end{pmatrix} \tag{13}$$

and $\mathbf{G}$ is the weighting 4×4 matrix. We control the physical quantity to be minimized by the choice of the form of the matrix $\mathbf{G}$.

In order to minimize the mean square of the width of GBs, we choose the matrix $\mathbf{G}$ in the form of

$$\mathbf{G} = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}. \tag{14}$$

Hence, the objective function reads

$$T(\mathbf{G}) = \int_{q_3^{(1)}}^{q_3^{(2)}} \mathrm{tr}\left\{\left[\mathbf{Y}(q_3)\right]^{-1}\right\} \mathrm{d}q_3 \,. \tag{15}$$

The objective function may be also written as

$$T(\mathbf{G}) = \mathrm{tr}\left\{\mathbf{B}(\mathbf{G})\,\mathrm{Re}(\mathbf{\Psi}_0)\right\}, \tag{16}$$

where

$$\mathbf{B}(\mathbf{G}) = \int_{q_3^{(1)}}^{q_3^{(2)}} \mathbf{\Pi}^{\dagger}\left(q_3, q_3^{(0)}\right)\mathbf{G}(q_3)\,\mathbf{\Pi}\left(q_3, q_3^{(0)}\right) \mathrm{d}q_3 \,, \tag{17}$$

the dagger ($\dagger$) is used to denote transpose, and $\mathrm{Re}(\mathbf{\Psi}_0)$ is given by

$$\mathrm{Re}(\mathbf{\Psi}) = \mathbf{\Pi}\,\mathrm{Re}(\mathbf{\Psi}_0)\,\mathbf{\Pi}^{\dagger}, \tag{18}$$

$$\mathrm{Re}\left(\boldsymbol{\Psi}_0\right) = \begin{pmatrix} \mathbf{Y}_0^{-1} & \mathbf{Y}_0^{-1}\mathbf{R}_0 \\ \mathbf{R}_0\mathbf{Y}_0^{-1} & \mathbf{Y}_0 + \mathbf{R}_0\mathbf{Y}_0^{-1}\mathbf{R}_0 \end{pmatrix} . \tag{19}$$

Let us now decompose the real positive-definite symmetric 4×4 matrix $\mathbf{B}$ into 2×2 submatrices

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix} , \tag{20}$$

which yields

$$T = \mathrm{tr}\left\{\mathbf{B}_{11}\mathbf{Y}_0^{-1} + \mathbf{B}_{12}\mathbf{R}_0\mathbf{Y}_0^{-1} + \mathbf{B}_{21}\mathbf{Y}_0^{-1}\mathbf{R}_0 + \mathbf{B}_{22}\left(\mathbf{R}_0\mathbf{Y}_0^{-1}\mathbf{R}_0 + \mathbf{Y}_0\right)\right\} . \tag{21}$$

This objective function has just one local extreme which is simultaneously the global minimum.

Differentiating the objective function (21) with respect to the real symmetric matrix $\mathbf{R}_0$ and putting the result to equal zero, we obtain

$$\mathbf{Y}_0^{-1}\mathbf{B}_{12} + \mathbf{B}_{21}\mathbf{Y}_0^{-1} + \mathbf{Y}_0^{-1}\mathbf{R}_0\mathbf{B}_{22} + \mathbf{B}_{22}\mathbf{R}_0\mathbf{Y}_0^{-1} = 0 . \tag{22}$$

Finally, we obtain

$$\mathbf{R}_0 = \mathbf{X} - \left(\mathbf{Y}_0\mathbf{B}_{22}\right)\left(\mathbf{X} - \mathbf{X}^{\dagger}\right)\left[\mathrm{tr}\left(\mathbf{Y}_0\mathbf{B}_{22}\right)\right]^{-1} , \tag{23}$$

where

$$\mathbf{X} = -\mathbf{B}_{12}\mathbf{B}_{22}^{-1} . \tag{24}$$

Differentiating the objective function (16) with respect to the real symmetric matrix $\mathbf{Y}_0$ and putting the result to equal zero, we obtain

$$-\mathbf{Y}_0^{-1}\left[\mathbf{B}_{11} + \mathbf{B}_{12}\mathbf{R}_0 + \mathbf{R}_0\mathbf{B}_{21} + \mathbf{R}_0\mathbf{B}_{22}\mathbf{R}_0\right]\mathbf{Y}_0^{-1} + \mathbf{B}_{22} = 0 . \tag{25}$$

Finally, we arrive at

$$\mathbf{Y}_0 = \mathbf{B}_{22}^{-1/2}\mathbf{S}\mathbf{B}_{22}^{-1/2}\left\{1 + \det\left(\mathbf{X} - \mathbf{X}^{\dagger}\right)\det\left(\mathbf{B}_{22}\right)\left[\mathrm{tr}\left(\mathbf{S}\right)\right]^{-2}\right\}^{1/2} , \tag{26}$$

where

$$\mathbf{S} = \left(\mathbf{B}_{22}^{1/2}\mathbf{C}_{11}\mathbf{B}_{22}^{1/2}\right)^{1/2} \tag{27}$$

and

$$\mathbf{C}_{11} = \mathbf{B}_{11} - \mathbf{B}_{12}\mathbf{B}_{22}^{-1}\mathbf{B}_{21} . \tag{28}$$

The matrix **B**, which is the result of the integration of the ordinary differential equations

$$\frac{\mathrm{d}\mathbf{B}}{\mathrm{d}q_3} = \mathbf{\Pi}^{\dagger}\mathbf{G}\mathbf{\Pi} , \tag{29}$$

is ill-conditioned. This may be overcome by using the symmetric matrix

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{B}_{11} - \mathbf{B}_{12}\mathbf{B}_{22}^{-1}\mathbf{B}_{21} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix} . \tag{30}$$

The differential equations for the matrix **C** for optimization read

$$\frac{\mathrm{d}\mathbf{C}_{11}}{\mathrm{d}q_3} = \left(\mathbf{\Pi}_1 - \mathbf{\Pi}_2\mathbf{C}_{22}^{-1}\mathbf{C}_{21}\right)^{\dagger} \mathbf{G} \left(\mathbf{\Pi}_1 - \mathbf{\Pi}_2\mathbf{C}_{22}^{-1}\mathbf{C}_{21}\right) ,$$

$$\frac{\mathrm{d}\mathbf{C}_{12}}{\mathrm{d}q_3} = \mathbf{\Pi}_1^{\dagger}\mathbf{G}\mathbf{\Pi}_2 , \tag{31}$$

$$\frac{\mathrm{d}\mathbf{C}_{22}}{\mathrm{d}q_3} = \mathbf{\Pi}_2^{\dagger}\mathbf{G}\mathbf{\Pi}_2 ,$$

where

$$\mathbf{\Pi}_1 = \begin{pmatrix} \mathbf{Q}_1 \\ \mathbf{P}_1 \end{pmatrix} \tag{32}$$

and

$$\mathbf{\Pi}_2 = \begin{pmatrix} \mathbf{Q}_2 \\ \mathbf{P}_2 \end{pmatrix} \tag{33}$$

are 4×2 submatrices of the paraxial-ray propagator matrix (9).

4. TRANSFORMATION OF THE MATRIX C

We define the Cartesian components of the slowness vector

$$p_i^{(z)} = \frac{\partial \tau}{\partial z_i} , \tag{35}$$

where (z_1, z_2, z_3) is a local Cartesian coordinate system with its origin at the initial point of the ray and basis vectors $\boldsymbol{i}_1^{(z)}$, $\boldsymbol{i}_2^{(z)}$ and $\boldsymbol{i}_3^{(z)}$. We choose the unit vector $\boldsymbol{i}_3^{(z)}$ to coincide with the unit vector normal to the initial surface at the initial point of the ray. The vectors $\boldsymbol{i}_1^{(z)}$ and $\boldsymbol{i}_2^{(z)}$ are then obviously situated in the plane tangent to the initial surface.

The unitary transformation matrix is defined by

$$H_{ij} = \frac{\partial z_i}{\partial q_j} = \frac{\partial q_j}{\partial z_i} , \tag{36}$$

the columns of which constitute the local vector basis of the ray-centred coordinate system expressed in local Cartesian coordinates z_i . We shall also denote the components of the velocity gradient in the ray-centred coordinate system on the central ray

$$V_i = \left(\frac{\partial v}{\partial q_i} \right)_{q_I = 0} = \left(\frac{\partial z_j}{\partial q_i} \frac{\partial v}{\partial z_j} \right)_{q_I = 0} . \tag{37}$$

The second-order Taylor expansion of the time field along the initial surface is given by the relation (*Klimeš, 1984*)

$$\theta^{\Sigma} = \tau + p_I^{(z)} z_I + \frac{1}{2} z_I z_J M_{IJ}^{\Sigma} . \tag{38}$$

The matrix $\mathbf{M}^{\Sigma}$ is defined by

$$M_{IJ}^{\Sigma} = H_{IK} H_{JL} M_{KL} \left(q_3^{(0)} \right) + p_3^{(z)} D_{IJ} + E_{IJ} , \tag{39}$$

where

$$E_{IJ} = -H_{I3} H_{JK} V_K v^{-2} - H_{J3} H_{IK} V_K v^{-2} - H_{I3} H_{J3} V_3 v^{-2} \tag{40}$$

and **D** is the matrix of curvature of the initial surface.

Let us now introduce the matrix **F**,

$$\mathbf{F} = \begin{pmatrix} \mathbf{H}^{\dagger} & \mathbf{0} \\ -\mathbf{H}^{-1} \left(p_3^{(z)} \mathbf{D} + \mathbf{E} \right) & \mathbf{H}^{-1} \end{pmatrix} , \tag{41}$$

where **H** is the 2 × 2 upper left minor of the 3 × 3 matrix (36). Then,

$$\begin{pmatrix} \mathbf{1} \\ \mathbf{M}_0 \end{pmatrix} = \mathbf{F} \begin{pmatrix} \mathbf{1} \\ \mathbf{M}_0^{\Sigma} \end{pmatrix} \left(\mathbf{H}^{\dagger} \right)^{-1} . \tag{42}$$

Inserting Eq.(42) into Eq.(19), we obtain

$$\mathrm{Re}\left(\Psi_0 \right) = \mathbf{F} \, \mathrm{Re}\left(\Psi_0^{\Sigma} \right) \mathbf{F}^{\dagger} , \tag{43}$$

where $\mathrm{Re}\left(\boldsymbol{\Psi}_0^{\Sigma}\right)$ is defined by equation analogous to Eq.(19), with $\mathbf{M}_0$ replaced by $\mathbf{M}_0^{\Sigma}$. Considering Eq.(43), the objective function can be expressed in the form analogous to Eq.(16), with $\mathrm{Re}\left(\boldsymbol{\Psi}_0\right)$ replaced by $\mathrm{Re}\left(\boldsymbol{\Psi}_0^{\Sigma}\right)$ and $\mathbf{B}(\mathbf{G})$ replaced by

$$\mathbf{B}^{\Sigma} = \mathbf{F}^{\dagger}\mathbf{B}\mathbf{F} . \tag{44}$$

From Eqs.(30), (41), and (44), we can finally derive

$$\mathbf{C}_{11}^{\Sigma} = \mathbf{H}\mathbf{C}_{11}\mathbf{H}^{\dagger} , \tag{45}$$

$$\mathbf{C}_{21}^{\Sigma} = \left(\mathbf{H}^{\dagger}\right)^{-1}\left[\mathbf{C}_{21}\mathbf{H}^{\dagger} - \mathbf{C}_{22}\mathbf{H}^{-1}\left(p_3^{(z)}\mathbf{D} + \mathbf{E}\right)\right], \tag{46}$$

$$\mathbf{C}_{12}^{\Sigma} = \left(\mathbf{C}_{21}^{\Sigma}\right)^{\dagger} , \tag{47}$$

and

$$\mathbf{C}_{22}^{\Sigma} = \left(\mathbf{H}^{\dagger}\right)^{-1}\mathbf{C}_{22}\mathbf{H}^{-1} . \tag{48}$$

5. 2-D CASE WITH A FLAT INITIAL SURFACE

In 2-D, the submatrices of the matrix $\mathbf{C}$ may be written as

$$\mathbf{C}_{11} = \begin{pmatrix} C_{11} & 0 \\ 0 & C_{11}^{\perp} \end{pmatrix}, \tag{49}$$

$$\mathbf{C}_{22} = \begin{pmatrix} C_{22} & 0 \\ 0 & C_{22}^{\perp} \end{pmatrix}, \tag{50}$$

and

$$\mathbf{C}_{12} = \mathbf{C}_{21} = \begin{pmatrix} C_{12} & 0 \\ 0 & C_{12}^{\perp} \end{pmatrix}. \tag{51}$$

In other words, we have three independent in-plane parameters, C_{11}, C_{12}, and C_{22}. Parameters $C_{11}^{\perp}$, $C_{12}^{\perp}$, and $C_{22}^{\perp}$ describe the optimum initial parameters of GBs perpendicularly to the plane of calculation. In the case of a flat initial surface, matrix $\mathbf{D}$ is given by

$$\mathbf{D} = \mathbf{0} . \tag{52}$$

Hence, we can write that

$$C_{11}^{\Sigma} = \left(v p_3^{(z)} \right)^2 C_{11} , \tag{53}$$

$$C_{22}^{\Sigma} = \left(v p_3^{(z)} \right)^{-2} C_{22} , \tag{54}$$

and

$$C_{12}^{\Sigma} = C_{12} - \left(v p_3^{(z)} \right)^{-2} E C_{22} , \tag{55}$$

where $E = E_{11}$ is given by Eq.(40),

$$E = -2 p_1^{(z)} p_3^{(z)} \left(\frac{\partial v}{\partial q_1} \right)_{q_{1,3}=0} - \left(p_1^{(z)} \right)^2 \left(\frac{\partial v}{\partial q_3} \right)_{q_{1,3}=0} . \tag{56}$$

Finally, we present the initial parameters of the shape of the Gaussian beams projected on the initial surface,

$$M_0^{\Sigma} = R_0^{\Sigma} + \mathrm{i} Y_0^{\Sigma} , \tag{57}$$

where

$$R_0^{\Sigma} = -C_{12}^{\Sigma} \left(C_{22}^{\Sigma} \right)^{-1} \tag{58}$$

and

$$Y_0^{\Sigma} = \left[C_{11}^{\Sigma} \left(C_{22}^{\Sigma} \right)^{-1} \right]^{1/2} . \tag{59}$$

6. SMOOTHING THE DISTRIBUTION OF R_0^{Σ} AND Y_0^{Σ}

The rays may be defined as the characteristic curves of the eikonal equation (e.g., *Červený, 2001*). In smoothly inhomogeneous isotropic media, the eikonal equation reads

$$p_i p_i = v^{-2} \left(\boldsymbol{x} \right) , \tag{60}$$

where

$$p_i = \frac{\partial \tau}{\partial x_i} , \tag{61}$$

x_i being the general Cartesian coordinates. In general, we shall write the eikonal equation as

$$H(\boldsymbol{x}, \boldsymbol{p}) = 0 , \tag{62}$$

where the Hamiltonian function $H(\boldsymbol{x}, \boldsymbol{p})$ may be specified in various ways.

We consider x_i and p_i to be independent coordinates in a six-dimensional phase-space (four-dimensional phase-space in 2-D). The eikonal equation then defines a Hamiltonian hypersurface (five-dimensional in 3-D and three-dimensional in 2-D) in the phase-space.

In 2-D, let us define a new coordinate system y_i in the three-dimensional Hamiltonian hypersurface. Coordinate y_1 corresponds to an independent variable along the ray, y_2 corresponds to take-off angle of the ray, and y_3 corresponds to the position of the initial point of the ray along the initial line. We shall call it the phase-space ray coordinate system.

Let us remind that the parameters R_0^{Σ} and Y_0^{Σ} depend upon an independent variable along the ray, take-off angles of the beam's central ray, and the position of the initial point of the ray,

$$R_0^{\Sigma} = R_0^{\Sigma}(\boldsymbol{y}) \tag{63}$$

and

$$Y_0^{\Sigma} = Y_0^{\Sigma}(\boldsymbol{y}) . \tag{64}$$

For numerical purposes, we have to choose certain discretization of the Hamiltonian hypersurface. This may be done by choosing a sufficiently dense grid in the Hamiltonian hypersurface, the grid points of which are used for storing the necessary physical quantities, and also for storing the optimum initial parameters of GBs. In determining the optimum parameters of GBs for migration, the grid points form a regular rectangular grid in coordinates

$$w_1 = t , \; w_2 = y_2 , \; w_3 = y_3 , \tag{65}$$

where t is the time in the common shot gather. Coordinates w_i are suitable for optimization of the shape of GBs for Gaussian beam or packet migrations. The inverted data are transformed from coordinates y_i to coordinates w_i using a coordinate transform

$$w_1 = \tau^S\left(x_i(y_1)\right) + y_1 , \; w_2 = y_2 , \; w_3 = y_3 , \tag{66}$$

where $\tau^S(x_i)$ is the travel time from the source corresponding to the time section to point x_i. Multivalued travel time $\tau^S(x_i)$ then creates multiple data points.

Since we need to have the distribution of the initial parameters of GBs sufficiently smooth in the decomposition of the wave field into GBs or packets, we should be able to smooth it. Let us denote the smoothed initial parameters of GBs by R_0^M and Y_0^M .

In obtaining a smoother distribution of the optimum initial parameters of GBs, we minimize the squares of the relevant Sobolev norms of the parameters of GBs together with the mean squares of the widths of the corresponding GBs. The objective function to be minimized reads

$$O = \frac{T}{q_3^{(2)} - q_3^{(1)}} + \left\| R_0^M \right\|^2 + \left\| Y_0^M \right\|^2 , \tag{67}$$

where T is defined by Eq.(15) and $\|\bullet\|$ is the appropriate Sobolev norm. The Sobolev scalar product is a linear combination of the L2 Lebesgue scalar products of the zero, first, second or higher partial derivatives (*Tarantola, 1987*).

Expressing T in the form of

$$T = C_{11}^{\Sigma} \left(Y_0^M \right)^{-1} + \left[R_0^M + C_{12}^{\Sigma} \left(C_{22}^{\Sigma} \right)^{-1} \right]^2 C_{22}^{\Sigma} \left(Y_0^M \right)^{-1} + C_{22}^{\Sigma} Y_0^M , \tag{68}$$

we see that the objective function (67) is minimized by R_0^M minimizing the objective function

$$O_R = \left\| R_0^M \right\|^2 + \left| \frac{R_0^M - R_0^D}{\sigma_R} \right|_{\mathrm{L2}}^2 , \tag{69}$$

where $|\bullet|_{\mathrm{L2}}$ is the standard L2 Lebesgue norm, R_0^D is given by

$$R_0^D = R_0^{\Sigma} , \tag{70}$$

and the standard deviations σ_R are defined as

$$\sigma_R = \sqrt{\frac{Y_0^M}{C_{22}^{\Sigma}}} . \tag{71}$$

During the iterative linearized smoothing, σ_R is calculated using Y_0^M from the previous iteration, with the initial estimate corresponding to value

$$\sigma_R = \sqrt{\frac{Y_0^{\Sigma}}{C_{22}^{\Sigma}}} . \tag{72}$$

Eq.(68) may be rearranged to read

$$T = 2 C_{22}^{\Sigma} Y_0^D + C_{22}^{\Sigma} \left(Y_0^M \right)^{-1} \left(Y_0^M - Y_0^D \right)^2 , \tag{73}$$

where

$$Y_0^D = \sqrt{C_{11}^{\Sigma} \left(C_{22}^{\Sigma} \right)^{-1} + \left(R_0^M - R_0^D \right)^2} . \tag{74}$$

We see that the objective function (67) is minimized by Y_0^M minimizing the objective function

$$O_Y = \left\| Y_0^M \right\|^2 + \left| \frac{Y_0^M - Y_0^D}{\sigma_Y} \right|_{\mathrm{L2}}^2 , \tag{75}$$

where the standard deviations σ_Y are given by

$$\sigma_Y = \sqrt{\frac{Y_0^M}{C_{22}^{\Sigma}}} . \tag{76}$$

During the iterative linearized smoothing, σ_Y is calculated using Y_0^M from the previous iteration, with the initial estimate corresponding to value

$$\sigma_Y = \sqrt{\frac{Y_0^D}{C_{22}^{\Sigma}}} . \tag{77}$$

In obtaining the parameters R_0^M , we minimize the objective function O_R defined by formula

$$\begin{aligned} O_R = & \sum_{GRID} \left(\frac{R_0^D \left(\boldsymbol{w}^{GRID} \right) - R_0^M \left(\boldsymbol{w}^{GRID} \right)}{\sqrt{N} \sigma_R^{GRID}} \right)^2 \\ & + \left[\int d^3 w \right]^{-1} \int b_{ij} \left(\frac{\partial R_0^M \left(\boldsymbol{w} \right)}{\partial w_i} \right) \left(\frac{\partial R_0^M \left(\boldsymbol{w} \right)}{\partial w_j} \right) d^3 w , \end{aligned} \tag{78}$$

where $\boldsymbol{w} = \left(w_1, w_2, w_3 \right)$ and b_{ij} are the weighting coefficients of the Sobolev scalar product. Superscript *GRID* takes values *GRID* = 1, 2, …, *N*, where *N* is the number of grid points of the original data grid. As we have not had any prior information about the optimum smoothness of the distribution of the initial parameters of GBs, we have used here only the first derivatives in the Sobolev norms in constructing the objective function.

We can express R_0^M as a linear combination of tricubic B-splines B_α

$$R_0^M \left(\boldsymbol{w} \right) = B_\alpha \left(\boldsymbol{w} \right) R_\alpha^B , \tag{79}$$

where R_α^B are the values of the smoothed initial parameters of GBs at grid points of the B-spline grid, which is a sub-grid of the original data grid. Subscript α takes values α = 1, 2, …, *P*, where *P* is the number of B-splines describing the smoothed distribution of the optimum initial parameters of GBs.

Eq.(78) now reads

$$O_R = \sum_{GRID} \left(\frac{R_0^D \left(\boldsymbol{w}^{GRID} \right) - B_\alpha \left(\boldsymbol{w}^{GRID} \right) R_\alpha^B}{\sqrt{N} \sigma_R^{GRID}} \right)^2 + R_\alpha^B D_{\alpha\beta} R_\alpha^B , \tag{80}$$

where

$$D_{\alpha\beta} = \left[\int d^3 w \right]^{-1} \int b_{ij} \left(\frac{\partial B_\alpha (\boldsymbol{w})}{\partial w_i} \right) \left(\frac{\partial B_\beta (\boldsymbol{w})}{\partial w_j} \right) d^3 w \, . \tag{81}$$

Since we do not know the coefficients b_{ij} which lead to the optimum distribution of the initial parameters of GBs, the problem is not linear. Thus, parameters R_α^B cannot be determined analytically. Since we do not want to solve the non-linear inverse problem numerically, we need to "linearize" Eq.(81). The linearization of Eq.(81) yields

$$D_{\alpha\beta} = s_R^2 D'_{\alpha\beta} \, , \tag{82}$$

$$D'_{\alpha\beta} = \left[\int d^3 w \right]^{-1} \int b'_{ij} \left(\frac{\partial B_\alpha (\boldsymbol{w})}{\partial w_i} \right) \left(\frac{\partial B_\beta (\boldsymbol{w})}{\partial w_j} \right) d^3 w \, , \tag{83}$$

where s_R is a free parameter and b'_{ij} are fixed coefficients of the Sobolev scalar product.

Coefficients b'_{ij} may be constructed as a completely symmetric tensor. The 3×3 matrix $\mathbf{b}'$ is then defined by

$$b'_{ij} = \frac{\langle e_i e_j \rangle}{d} \, , \tag{84}$$

where $\mathbf{e}$ is a unit vector, $\langle \cdots \rangle$ indicates averaging over all possible directions of a unit vector, $d = 1$ in 1-D, $d = 2$ in 2-D, and $d = 3$ in 3-D. The average of a unit vector over all directions can be calculated analytically. Generally in d-D, we may put

$$b'_{ij} = \frac{\delta_{ij}}{d} \, , \tag{85}$$

where δ_{ij} is the Kronecker symbol. In 3-D, the desired matrix $\mathbf{b}'$ may be expressed as

$$\mathbf{b}' = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{pmatrix} . \tag{86}$$

We can now rewrite Eq.(58) to read

$$O_R = \left(\mathbf{R}_0^D - \mathbf{B}\mathbf{R}^B \right)^\dagger \mathbf{C}_R^{-1} \left(\mathbf{R}_0^D - \mathbf{B}\mathbf{R}^B \right) + s_R^2 \left(\mathbf{R}^B \right)^\dagger \mathbf{D}'\mathbf{R}^B \, , \tag{87}$$

where $\mathbf{R}_0^D$ is defined as $\left(R_0^D \right)_i = R_0^D (w_i)$, $\mathbf{B}$ is defined as $B_{i\alpha} = B_\alpha (w_i)$, $\mathbf{D}'$ is a $P \times P$ matrix given by Eq.(83), and $\mathbf{C}_R$ is a $N \times N$ diagonal matrix, composed of $N\left(\sigma_R^{GRID} \right)^2$, see Eq.(78).

Differentiating the objective function (87) with respect to the vector $\mathbf{R}^B$ and putting the result to equal zero, we obtain

$$\mathbf{B}^{\dagger}\mathbf{C}_R^{-1}\left(\mathbf{B}\mathbf{R}^B - \mathbf{R}_0^D\right) + s_R^2 \mathbf{D}'\mathbf{R}^B = 0 . \tag{88}$$

The resulting vector $\mathbf{R}^B$ is

$$\mathbf{R}^B = \left[\mathbf{B}^{\dagger}\mathbf{C}_R^{-1}\mathbf{B} + s_R^2 \mathbf{D}'\right]^{-1} \mathbf{B}^{\dagger}\mathbf{C}_R^{-1}\mathbf{R}_0^D . \tag{89}$$

By analogy, following the procedure from Eq.(78) to Eq.(89), the resulting vector $\mathbf{Y}^B$ is

$$\mathbf{Y}^B = \left[\mathbf{B}^{\dagger}\mathbf{C}_Y^{-1}\mathbf{B} + s_Y^2 \mathbf{D}'\right]^{-1} \mathbf{B}^{\dagger}\mathbf{C}_Y^{-1}\mathbf{Y}_0^D , \tag{90}$$

where $\mathbf{C}_Y$ is a $N \times N$ diagonal matrix, composed of $N\left(\sigma_Y^{GRID}\right)^2$.

7. ALGORITHM

First of all, we need to compute a sufficiently dense set of rays and store several important quantities along the rays. In dependency upon the independent variable y_1 along the ray (e.g., travel time), the take-off angle y_2 of the beam's central ray, and the position y_3 of the initial point of the ray (e.g., source), we have to store the paraxial-ray propagator matrix, see Eq.(9), the Cartesian components of the slowness vector, see Eq.(35), and the transformation matrix, the columns of which constitute the local vector basis of the ray-centred coordinates expressed in Cartesian coordinates, see Eq.(36).

Then, we solve the ordinary differential equation (31) by numerical integration along the ray. The results of the numerical integration describe the optimum initial profiles of GBs in a plane perpendicular to the central ray of GBs. We have to transform the results of the numerical integration to the analogous quantities describing the optimum shape of GBs along the initial surface, see Eqs.(45–48).

Finally, we can calculate the optimum initial parameters of GBs R_0^{Σ} and Y_0^{Σ}, which are discretized in dependency upon coordinates w_i, see Eqs.(65) and (66). Thus, each initial parameter of GBs is defined on the 3-D parameter grid. We also need to store the values of C_{22}^{Σ} for smoothing.

If necessary, we smooth the distribution of initial parameters of GBs on the Hamiltonian hypersurface in the phase-space. In order to obtain a smoother distribution, we minimize the squares of the relevant Sobolev norms of the parameters of GBs together with the mean squares of the widths of the corresponding GBs, see Eq.(67).

8. NUMERICAL EXAMPLES

We have decided to use the Marmousi model (*Versteeg and Grau, 1991*) as the velocity model. Since the original Marmousi model is too complex for ray-based methods, we have used the smoothed Marmousi model (*Žáček, 2002*) in the computations, see

Fig. 1. The dimensions of the model are 9200 m (length) by 3000 m (depth). The grided values of velocity vary from 1520 ms^{-1} to 4550 ms^{-1}.

We have prepared two groups of sets of the initial parameters of GBs: (a) very little smoothed sets (with various numbers of iterations), where we have used $s_R = 5 \times 10^9\ \mathrm{m}^3\mathrm{s}^{-1}$ and $s_Y = 5 \times 10^{10}\ \mathrm{m}^3\mathrm{s}^{-1}$ in the smoothing, see Eqs.(82), (89), and (90), and (b) sets smoothed to a constant value (with various numbers of iterations), where we have used $s_R = 1 \times 10^{13}\ \mathrm{m}^3\mathrm{s}^{-1}$ and $s_Y = 1 \times 10^{13}\ \mathrm{m}^3\mathrm{s}^{-1}$ in the smoothing.

Standard halfwidth a of a Gaussian beam of crossection

$$\exp\left(-\frac{q_1^2}{2a^2}\right), \tag{91}$$

multiplied by the square root of ($2\pi f$), f being the frequency, has been interpolated between the rays and displayed,

$$W = a\sqrt{2\pi f}\ , \tag{92}$$

which can be also expressed as

$$W = \sqrt{2\pi f\, Y^{-1}}\ . \tag{93}$$

The color coded quantity W is displayed at the respective points along the central rays of GBs. Note that the displayed quantity W describes the halfwidth of GBs at the endpoints of the ray segments, not the root mean square of the halfwidth along the ray segments, although the latter quantity is minimized. The yellow colour corresponds to the GB halfwidth of 0 m for all frequencies. The halfwidth increases from yellow through green and blue towards red. The red colour corresponds to the GB halfwidths of 1010 m and more for the frequency of 35 Hz, and of 1890 metres and more for the frequency of 10 Hz.

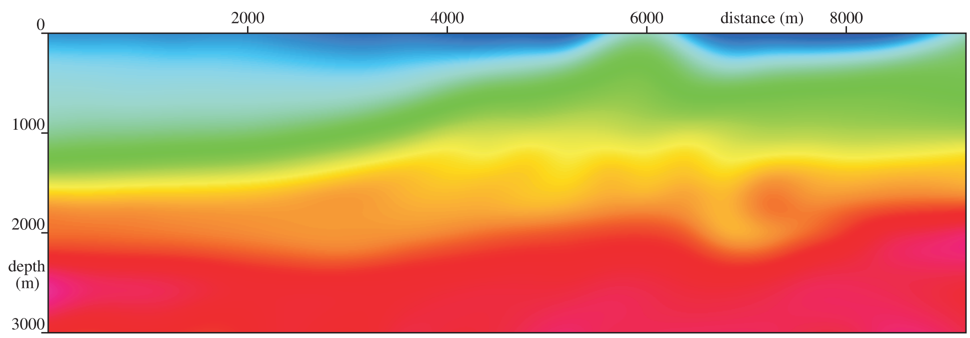

Fig. 1. Smoothed Marmousi model.

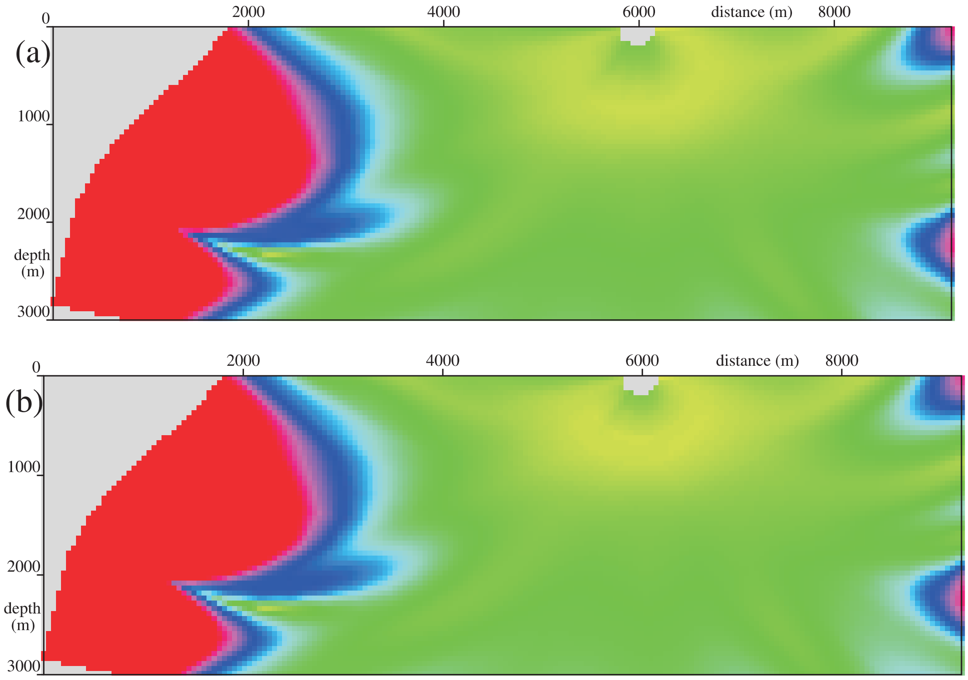

Fig. 2. The halfwidths of GBs for the position of the source of 5975 m. The green colour corresponds to the GB halfwidth of 202 and 378 m for the frequencies of 35 Hz and 10 Hz, respectively. The red colour corresponds to the GB halfwidth of 1010 m and more, and of 1890 m and more for the frequencies of 35 Hz and 10 Hz, respectively. The coefficients used for smoothing the optimum initial parameters of GBs are $s_R = 5 \times 10^9\ \mathrm{m^3 s^{-1}}$ and $s_Y = 5 \times 10^{10}\ \mathrm{m^3 s^{-1}}$; **(a)** one iteration, and **(b)** two iterations.

In Fig. 2, the optimum initial parameters of GBs are very little smoothed with one (Fig. 2a) and two (Fig. 2b) iterations. As you can see, there is no significant difference between the figures. For these initial parameters of GBs, we obtain the best achievable results in terms of the average GB width along a fixed part of the ray. Unfortunately, it would be impossible to decompose the wave field accurately into Gaussian beams or packets corresponding to such rapidly varying initial parameters.

In Fig. 3, the optimum initial parameters of GBs are smoothed to a constant value with one (Fig. 3a) and two (Fig. 3b) iterations. We can clearly see that the initial parameters of GBs smoothed with two iterations give much better results. We do not show the GB widths for the initial parameters smoothed with three or more iterations, because they do not change anymore. Such result indicates that the iterative method quickly converges even for an extensive smoothing. For the initial parameters of GBs smoothed to a constant value, we are able to control the error of decomposition of the wave field into Gaussian beams or packets (*Klimeš, 1986*; *Žáček, 2003, 2006*).

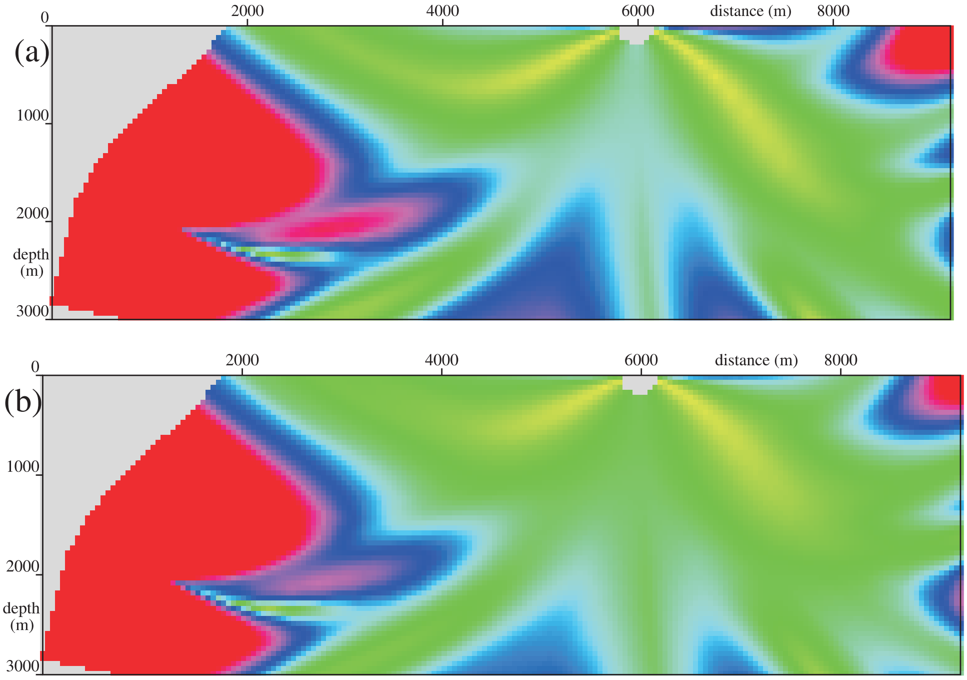

Fig. 3. The same as in Fig. 2, but for $s_R = 1 \times 10^{13}$ m^{3}s^{-1} and $s_Y = 1 \times 10^{13}$ m^{3}s^{-1}.

9. CONCLUSIONS

A proper choice of the initial shape of Gaussian beams represents a crucial step of any method based on Gaussian beams. For imaging purposes, Gaussian beams should stay narrow with respect to the velocity changes in the structure. Although we can easily minimize the average width of a single beam, such minimization alone is not sufficient when dealing with the whole wave field. In a complex structure, the optimum initial parameters of Gaussian beams may vary considerably in dependency upon the position of the initial point of the beam's central ray, the take-off angle of the beam's central ray, and the travel time. This would cause great problems in the decomposition of the wave field into optimized Gaussian beams or packets. Therefore, it is essential to smooth the distribution of the initial parameters of Gaussian beams on the Hamiltonian hypersurface in the phase-space.

Unfortunately, such smoothing is responsible for a faster spreading of beams, which may lead, in a complex structure, to an inaccurate solution of the elastodynamic equation. As a result, we have to deal with a trade-off between the accuracy of the decomposition of the wave field and the accuracy of the Gaussian beam solution of the elastodynamic equation.

Acknowledgements: The author is greatly indebted to L. Klimeš for guidance throughout the work on this topic. The research has been supported by the Grant Agency of the Czech Republic under Contract 205/04/1104, by the Grant Agency of the Charles University under Contract 375/2004/B-GEO/MFF, and by the members of the consortium "Seismic Waves in Complex 3-D Structures" (see http://sw3d.mff.cuni.cz).

References

Babich V.M. and Popov M.M., 1989. Gaussian summation method (review). *Radiophysics and Quantum Electronics*, **32**, 1063−1081.

Červený V., Popov M.M. and Pšenčík I., 1982, Computation of wave fields in inhomogeneous media - Gaussian beam approach. *Geophys. J. R. Astr. Soc.*, **70**, 109−128.

Červený V. and Pšenčík I., 1983. Gaussian beams and paraxial ray approximation in three-dimensional elastic inhomogeneous media. *J. Geophys.*, **53**, 1−15.

Červený V., 2001. *Seismic Ray Theory*. Cambridge Univ. Press, Cambridge, U.K.

Klimeš L., 1984. Expansion of a high-frequency time-harmonic wavefield given on an initial surface into Gaussian beams. *Geophys. J. R. Astr. Soc.*, **79**, 105−118.

Klimeš L., 1986. Discretization error for the superposition of Gaussian beams. *Geophys. J. R. Astr. Soc.*, **86**, 531−551.

Klimeš L., 1989. Optimization of the shape of Gaussian beams of a fixed length. *Stud. Geophys. Geod.*, **33**, 146−163.

Popov M.M., 1982. A new method of computation of wavefields using Gaussian beams. *Wave Motion*, **4**, 85−97.

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

Tarantola A., 1987. *Inverse Problem Theory*. Elsevier, Amsterdam, The Netherlands.

Versteeg R.J. and Grau G. (Eds.), 1991. *The Marmousi experience*. Proc. EAGE workshop on Practical Aspects of Seismic Data Inversion (Copenhagen, 1990), Eur. Assoc. Expl. Geophysicists, Zeist.

Žáček K., 2002. Smoothing the Marmousi model. *Pure Appl. Geophys.*, **159**, 1507−1526.

Žáček K., 2003. Decomposition of the wave field into optimized Gaussian packets. *Soc. Expl. Geophys. Expanded Abstracts*, **22**, 1869−1872.

Žáček K., 2006. Decomposition of the wavefield into optimized Gaussian packets. *Stud. Geophys. Geod.*, **50**, 367−380.