Optimization of the Shape of Gaussian Beams

Karel Žáček

Department of Geophysics, Charles University, Ke Karlovu 3, 121 16 Praha 2, Czech Republic, E-mail: zacek@karel.troja.mff.cuni.cz

Summary

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 determinig 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.

This procedure can increase the applicability of the Gaussian beam or packet method, especially in complex structures. Moreover, it can make the use of the Gaussian beam or packet method more comfortable. The presented procedure is suitable for the optimization of the shape of Gaussian beams for Gaussian beam or packet migrations.

Key words

Asymptotic ray theory, Gaussian beams, 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., Červený et al., 1982; Červený & Pšenčík, 1983). 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 ray path.

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.

Seismic Waves in Complex 3-D Structures, Report 11 (Department of Geophysics, Charles University, Prague, 2000), pp. 181–201

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 GBs or packets. 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 $I, J, ... = 1, 2$ and the lower–case indices take the values $i, j, ... = 1, 2, 3$. The Einstein summation over the pairs of identical indices is used. The matrices are denoted by boldface letters (e.g., $\mathbf{A}$) 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

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

where i is the imaginary unit and ω is the circular frequency. In *ray–centred coordinates* q_j (eg., Popov & 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

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

where τ is the travel time along the central ray and $\mathbf{M}$ is the second differential of the phase along the plane tangent to the phase–front.

Matrix $\mathbf{M}$ consists of a symmetric real part $\mathbf{R}$ and of a positive–definite symmetric imaginary part $\mathbf{Y}$,

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

It may be also expressed as

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

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* (eg., Červený, 2001)

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

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

Any solution of the *dynamic ray–tracing system* with the initial conditions

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

may be expressed as

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

where the *paraxial–ray propagator matrix*

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

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

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

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. The general form of the minimized objective function is

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

where

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

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

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

Hence, the objective function reads

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

The objective function may be also written as

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

where

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

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

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

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

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

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

which yields

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

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

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

Finally, we obtain

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

where

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

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

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

Finally, we arrive at

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

where

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

and

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

The matrix $\mathbf{B}$, which is the result of the integration of the ordinary differential equations

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

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

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

The differential equations for the matrix $\mathbf{C}$ for optimization read

$$\begin{aligned} \frac{\mathrm{d}\mathbf{C}_{11}}{\mathrm{d}q_3} &= (\mathbf{\Pi}_1 - \mathbf{\Pi}_2\mathbf{C}_{22}^{-1}\mathbf{C}_{11})^{\dagger}\mathbf{G}(\mathbf{\Pi}_1 - \mathbf{\Pi}_2\mathbf{C}_{22}^{-1}\mathbf{C}_{11}) \quad , \\ \frac{\mathrm{d}\mathbf{C}_{12}}{\mathrm{d}q_3} &= \mathbf{\Pi}_1^{\dagger}\mathbf{G}\mathbf{\Pi}_2 \quad , \\ \frac{\mathrm{d}\mathbf{C}_{22}}{\mathrm{d}q_3} &= \mathbf{\Pi}_2^{\dagger}\mathbf{G}\mathbf{\Pi}_2 \quad , \end{aligned} \tag{31}$$

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

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

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 $\mathbf{i}_1^{(z)}$, $\mathbf{i}_2^{(z)}$ and $\mathbf{i}_3^{(z)}$. We choose the unit vector $\mathbf{i}_3^{(z)}$ to coincide with the unit vector normal to the initial surface at the initial point of the ray. The vectors $\mathbf{i}_1^{(z)}$ and $\mathbf{i}_2^{(z)}$ are then obviously situated in the plane tangent to the initial surface.

The unitary transformation matrix is defined by

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

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

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

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

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

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

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

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 $\mathbf{D}$ is the matrix of curvature of the initial surface.

Let us now introduce the matrix $\mathbf{F}$,

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

where $\mathbf{H}$ is the 2×2 upper left minor of the 3×3 matrix (36). Then

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

Inserting (42) into (19), we obtain

$$\mathrm{Re}\,(\mathbf{\Psi}_0) = \mathbf{F}\mathrm{Re}\left(\mathbf{\Psi}_0^{\Sigma}\right)\mathbf{F}^{\dagger} \quad , \tag{43}$$

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

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

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

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

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

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

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

5 2–D case with a flat initial surface

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

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

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

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

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

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

Hence, we can write that

$$C_{11}^{\Sigma} = (vp_3^{(z)})^2 C_{11} \quad , \tag{53}$$

$$C_{22}^{\Sigma} = (vp_3^{(z)})^{-2} C_{22} \tag{54}$$

and

$$C_{12}^{\Sigma} = C_{12} - (vp_3^{(z)})^{-2} E C_{22} \quad , \tag{55}$$

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

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

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

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

where

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

and

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

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* (eg., Červený, 2001). In smoothly inhomogeneous isotropic media, the *eikonal equation* reads

$$p_i p_i = v^{-2}(\mathbf{x}) \quad , \tag{60}$$

where

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

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

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

where the Hamiltonian function $H(\mathbf{x},\mathbf{p})$ 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,

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

and

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

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

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

where t is the time in the common–source time section. 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

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

where $\tau^{\mathrm{S}}(x_i)$ is the travel time from the source corresponding to the time section to point x_i. Multivalued travel time $\tau^{\mathrm{S}}(x_i)$ 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

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

where T is defined by (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

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

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

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

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

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

and the standard deviations σ_{R} are defined as

$$\sigma_{\mathrm{R}} = \sqrt{\frac{Y_0^{\mathrm{M}}}{C_{22}^{\Sigma}}} \quad . \tag{71}$$

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

$$\sigma_{\mathrm{R}} = \sqrt{\frac{Y_0^{\Sigma}}{C_{22}^{\Sigma}}} \quad . \tag{72}$$

Equation (68) may be rearranged to read

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

where

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

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

$$O_{\mathrm{Y}} = \|Y_0^{\mathrm{M}}\|^2 + \left| \frac{Y_0^{\mathrm{M}} - Y_0^{\mathrm{D}}}{\sigma_{\mathrm{Y}}} \right|_{\mathrm{L2}}^2 \quad , \tag{75}$$

where the standard deviations σ_{Y} are given by

$$\sigma_{\mathrm{Y}} = \sqrt{\frac{Y_0^{\mathrm{M}}}{C_{22}^{\Sigma}}} \quad . \tag{76}$$

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

$$\sigma_{\mathrm{Y}} = \sqrt{\frac{Y_0^{\mathrm{D}}}{C_{22}^{\Sigma}}} \quad . \tag{77}$$

In obtaining the parameters R_0^{M}, we minimize the objective function O_{R} defined by formula

$$O_{\mathrm{R}} = \sum_{\mathrm{GRID}} \left(\frac{R_0^{\mathrm{D}}(\mathbf{w}^{\mathrm{GRID}}) - R_0^{\mathrm{M}}(\mathbf{w}^{\mathrm{GRID}})}{\sqrt{N}\sigma_{\mathrm{R}}^{\mathrm{GRID}}} \right)^2 +$$

$$+ \left[\int \mathrm{d}^3 w \right]^{-1} \int b_{ij} \left(\frac{\partial R_0^{\mathrm{M}}(\mathbf{w})}{\partial w_i} \right) \left(\frac{\partial R_0^{\mathrm{M}}(\mathbf{w})}{\partial w_j} \right) \mathrm{d}^3 w \quad , \tag{78}$$

where $\mathbf{w} = (w_1, w_2, w_3)$ 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

$$R_0^{\mathrm{M}}(\mathbf{w}) = B_\alpha(\mathbf{w}) R_\alpha^{\mathrm{B}} \quad , \tag{79}$$

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 $\alpha = 1, 2, .., P$, where P is the number of B–splines describing the smoothed distribution of the optimum initial parameters of GBs.

Equation (78) now reads

$$O_{\mathrm{R}} = \sum_{\mathrm{GRID}} \left(\frac{R_0^{\mathrm{D}}(\mathbf{w}^{\mathrm{GRID}}) - B_\alpha(\mathbf{w}^{\mathrm{GRID}}) R_\alpha^{\mathrm{B}}}{\sqrt{N}\sigma_{\mathrm{R}}^{\mathrm{GRID}}} \right)^2 + R_\alpha^{\mathrm{B}} D_{\alpha\beta} R_\beta^{\mathrm{B}} \quad , \tag{80}$$

where

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

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" formula (81). The linearization of (81) yields

$$D_{\alpha\beta} = s_{\mathrm{R}}^2 D'_{\alpha\beta} \quad , \tag{82}$$

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

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

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

where $\mathbf{e}$ is a unit vector, $\langle ... \rangle$ 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

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

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

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

We can now rewrite equation (58) to read

$$O_{\mathrm{R}} = \left(\mathbf{R}_0^{\mathrm{D}} - \mathbf{B}\mathbf{R}^{\mathrm{B}}\right)^{\dagger} \mathbf{C}_{\mathrm{R}}^{-1} \left(\mathbf{R}_0^{\mathrm{D}} - \mathbf{B}\mathbf{R}^{\mathrm{B}}\right) + s_{\mathrm{R}}^2 (\mathbf{R}^{\mathrm{B}})^{\dagger} \mathbf{D}^{'} \mathbf{R}^{\mathrm{B}} \quad , \tag{87}$$

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

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

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

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

$$\mathbf{R}^{\mathrm{B}} = [\mathbf{B}^{\dagger}\mathbf{C}_{\mathrm{R}}^{-1}\mathbf{B} + s_{\mathrm{R}}^2 \mathbf{D}^{'}]^{-1} \mathbf{B}^{\dagger}\mathbf{C}_{\mathrm{R}}^{-1}\mathbf{R}_0^{\mathrm{D}} \quad . \tag{89}$$

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

$$\mathbf{Y}^{\mathrm{B}} = [\mathbf{B}^{\dagger}\mathbf{C}_{\mathrm{Y}}^{-1}\mathbf{B} + s_{\mathrm{Y}}^2 \mathbf{D}^{'}]^{-1} \mathbf{B}^{\dagger}\mathbf{C}_{\mathrm{Y}}^{-1}\mathbf{Y}_0^{\mathrm{D}} \quad , \tag{90}$$

where $\mathbf{C}_{\mathrm{Y}}$ is a $N \times N$ diagonal matrix, composed of $N(\sigma_{\mathrm{Y}}^{GRID})^2$.

7 Algorithm

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

(b) We solve the ordinary differential equations (31) by numerical integration along the ray.

(c) 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 equations (45), (46), (47) and (48).

(d) 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 (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.

(e) If necessary, we smooth the distribution of the optimum parameters of GBs. In obtaining 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 equation (67). The presented procedure also allows us to smooth iteratively.

8 Numerical examples

We have decided to use the Marmousi model (Versteeg & 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 Figure 1. The dimensions of the model are 9200 metres (length) by 3000 metres (depth). The grided values of velocity vary from $1520\mathrm{ms}^{-1}$ to $4550\mathrm{ms}^{-1}$.

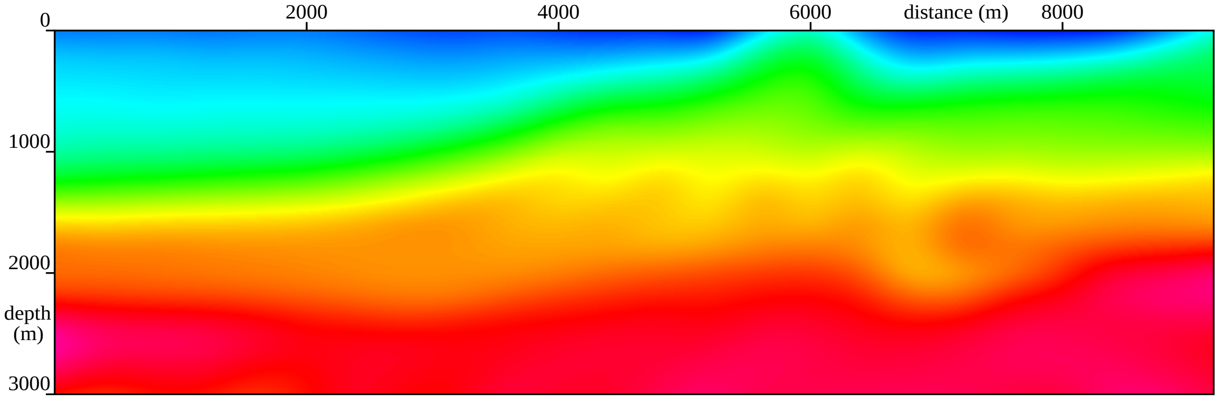

Figure 1. The smoothed Marmousi model.

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 $s_{\mathrm{R}} = 5 \times 10^9 \mathrm{m}^3\mathrm{s}^{-1}$ and $s_{\mathrm{Y}} = 5 \times 10^{10} \mathrm{m}^3\mathrm{s}^{-1}$ in the smoothing, see equations (82), (89) and (90), and
(b) sets smoothed to a constant value (with various numbers of iterations), where we have used $s_{\mathrm{R}} = 1 \times 10^{13} \mathrm{m}^3\mathrm{s}^{-1}$ and $s_{\mathrm{Y}} = 1 \times 10^{13} \mathrm{m}^3\mathrm{s}^{-1}$ in the smoothing.

In Tables 1a, b and 2a, b, we show the relative root–mean–square (RMS) differences between the parameters R_0^Σ and Y_0^Σ and the parameters R_0^{M} and Y_0^{M} obtained by the smoothing with one, two and three iterations. Note that the differences between the second and the third iterations are not very pronounced. For more iterations, there is no difference at all.

Although the relative RMS differences, see Tables 1a, b, close to 100% may look as too great, this is the least possible smoothing due to the numerical problems. The differences in Tables 2a and 2b considering the initial parameters of GBs smoothed to a constant value are, naturally, even worse. Especially for the position of the source of 5975 metres, where the relative RMS difference between R_0^{M} and R_0^Σ is up to 40000%. This is caused by a broad range of the values of the parameter R_0^Σ and clearly corresponds to a fact, that the spreading of GBs is the most pronounced for this position of the source.

$s_{\mathrm{Y}} = 5 \times 10^{10}\mathrm{m}^3\mathrm{s}^{-1}$	**425 m**	**3200 m**	**5975 m**	**8775 m**
relative RMS diff. (1$^{\mathrm{st}}$ it.)	65	37	42	69
relative RMS diff. (2$^{\mathrm{nd}}$ it.)	150	34	48	160
relative RMS diff. (3$^{\mathrm{rd}}$ it.)	160	34	52	190

Table 1a. The relative RMS difference between the smoothed parmeters Y_0^{M} and the original parameters Y_0^{Σ} (in [%]). Columns correspond to various positions of the source.

$s_{\mathrm{R}} = 5 \times 10^{9}\mathrm{m}^3\mathrm{s}^{-1}$	**425 m**	**3200 m**	**5975 m**	**8775 m**
relative RMS diff. (1$^{\mathrm{st}}$ it.)	84	84	140	110
relative RMS diff. (2$^{\mathrm{nd}}$ it.)	75	84	120	94
relative RMS diff. (3$^{\mathrm{rd}}$ it.)	83	83	120	100

Table 1b. The relative RMS difference between the smoothed parmeters R_0^{M} and the original parameters R_0^{Σ} (in [%]). Columns correspond to various positions of the source.

$s_{\mathrm{Y}} = 1 \times 10^{13}\mathrm{m}^3\mathrm{s}^{-1}$	**425 m**	**3200 m**	**5975 m**	**8775 m**
relative RMS diff. (1$^{\mathrm{st}}$ it.)	95	110	85	97
relative RMS diff. (2$^{\mathrm{nd}}$ it.)	230	200	130	130
relative RMS diff. (3$^{\mathrm{rd}}$ it.)	230	200	130	130

Table 2a. The relative RMS difference between the smoothed parmeters Y_0^{M} and the original parameters Y_0^{Σ} (in [%]). Columns correspond to various positions of the source.

$s_{\mathrm{R}} = 1 \times 10^{13}\mathrm{m}^3\mathrm{s}^{-1}$	**425 m**	**3200 m**	**5975 m**	**8775 m**
relative RMS diff. (1$^{\mathrm{st}}$ it.)	350	530	40000	130
relative RMS diff. (2$^{\mathrm{nd}}$ it.)	1900	1900	38000	160
relative RMS diff. (3$^{\mathrm{rd}}$ it.)	1900	1900	39000	160

Table 2b. The relative RMS difference between the smoothed parmeters R_0^{M} and the original parameters R_0^{Σ} (in [%]). Columns correspond to various positions of the source.

Note that the greatest jump in the relative RMS differences is between smoothing to a constant value with one iteration and with two iterations.

Standard halfwidth a of a Gaussian beam of crossection

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

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

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

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 metres for all frequencies. The halfwidth increases from yellow through green and blue towards red. The red colour corresponds to the GB halfwidths

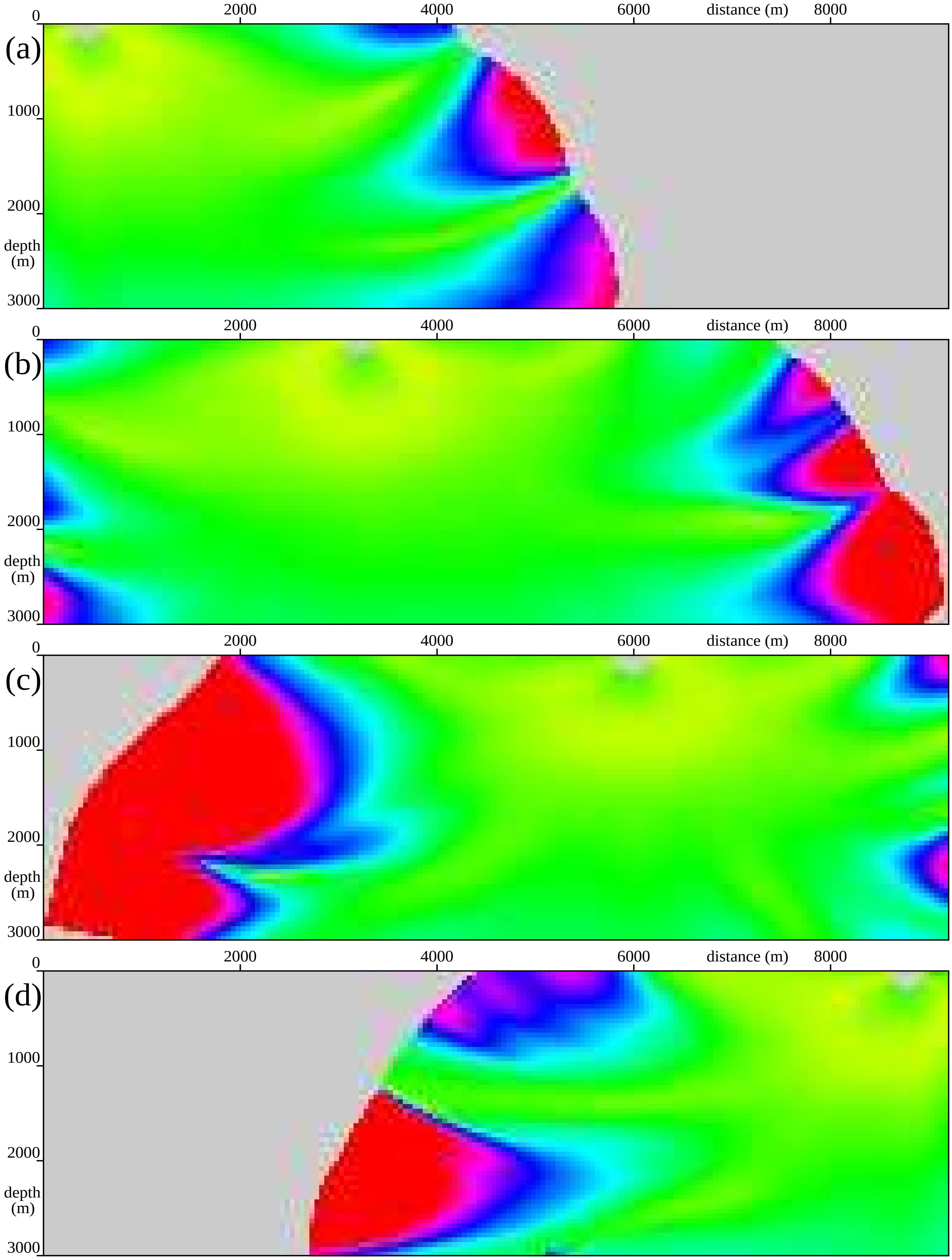

Figure 2. The halfwidths of GBs for the position of the source of (a) 425m, (b) 3200m, (c) 5975m and (d) 8775m. The *green* colour corresponds to the GB halfwidth of 202 and 378 metres for the frequencies of 35 Hz and 10 Hz, respectively. The *red* colour corresponds to the GB halfwidth of 1010 and more, and of 1890 metres and more for the frequencies of 35 Hz and 10 Hz, respectively. The coefficients used for smoothing the optimum initial parameters of GBs are $s_{\mathrm{R}} = 5 \times 10^{9}\mathrm{m}^3\mathrm{s}^{-1}$ and $s_{\mathrm{Y}} = 5 \times 10^{10}\mathrm{m}^3\mathrm{s}^{-1}$, one iteration.

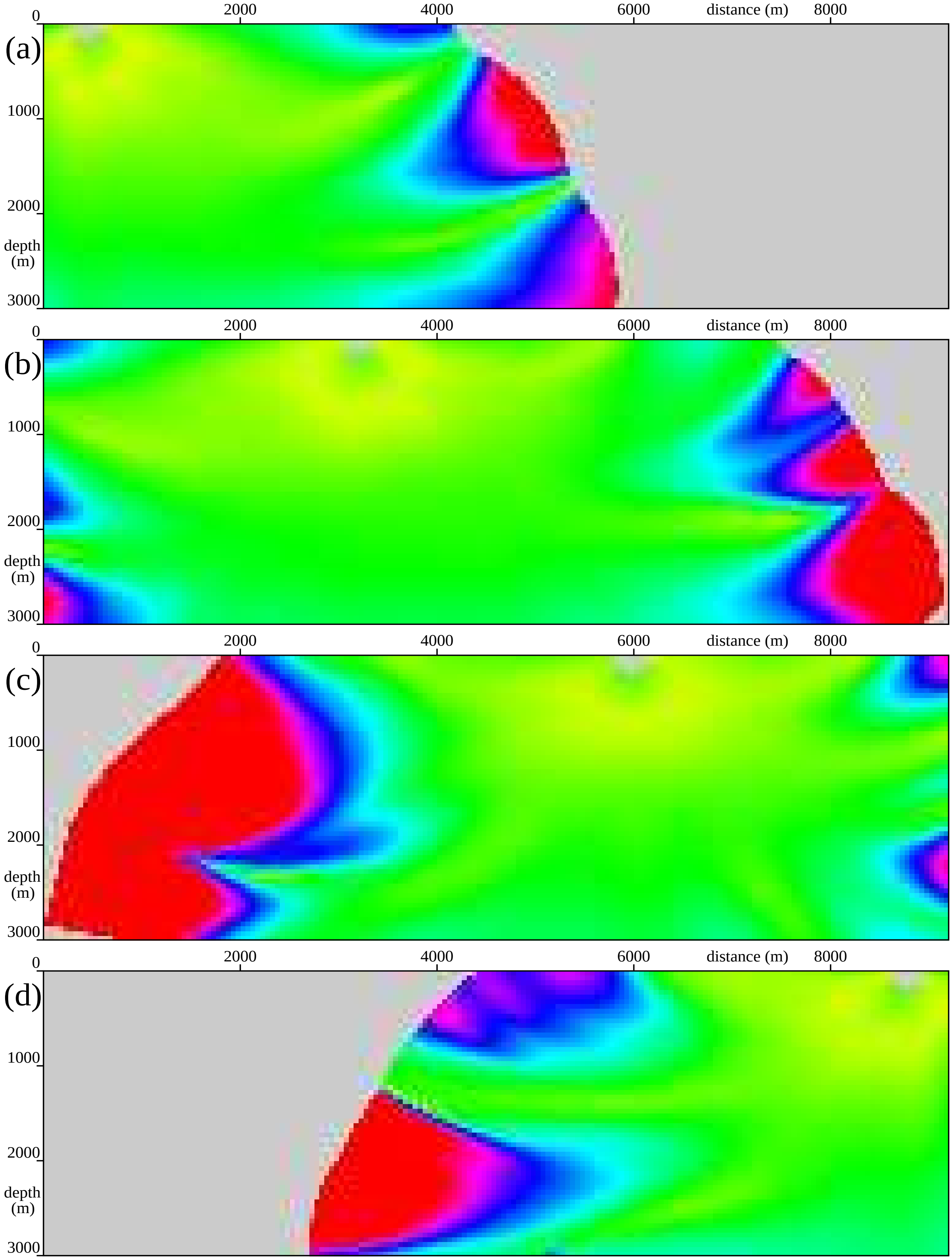

Figure 3. The halfwidths of GBs for the position of the source of (a) 425m, (b) 3200m, (c) 5975m and (d) 8775m. The *green* colour corresponds to the GB halfwidth of 202 and 378 metres for the frequencies of 35 Hz and 10 Hz, respectively. The *red* colour corresponds to the GB halfwidth of 1010 and more, and of 1890 metres and more for the frequencies of 35 Hz and 10 Hz, respectively. The coefficients used for smoothing the optimum initial parameters of GBs are $s_{\mathrm{R}} = 5 \times 10^{9}\mathrm{m}^{3}\mathrm{s}^{-1}$ and $s_{\mathrm{Y}} = 5 \times 10^{10}\mathrm{m}^{3}\mathrm{s}^{-1}$, two iterations.

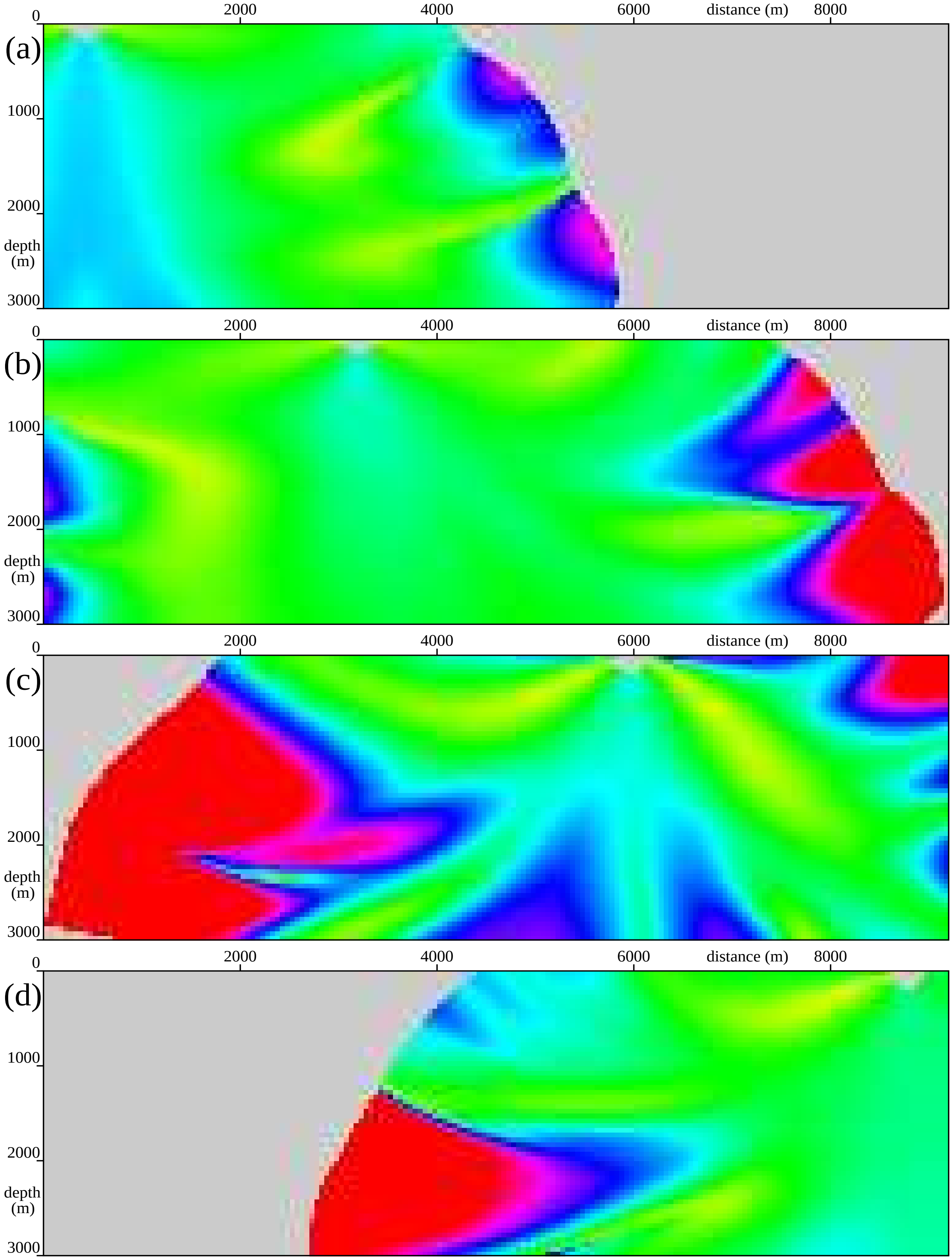

Figure 4. The halfwidths of GBs for the position of the source of (a) 425m, (b) 3200m, (c) 5975m and (d) 8775m. The *green* colour corresponds to the GB halfwidth of 202 and 378 metres for the frequencies of 35 Hz and 10 Hz, respectively. The *red* colour corresponds to the GB halfwidth of 1010 and more, and of 1890 metres and more for the frequencies of 35 Hz and 10 Hz, respectively. The coefficients used for smoothing the optimum initial parameters of GBs are $s_{\mathrm{R}} = 1 \times 10^{13}\mathrm{m}^3\mathrm{s}^{-1}$ and $s_{\mathrm{Y}} = 1 \times 10^{13}\mathrm{m}^3\mathrm{s}^{-1}$, one iteration.

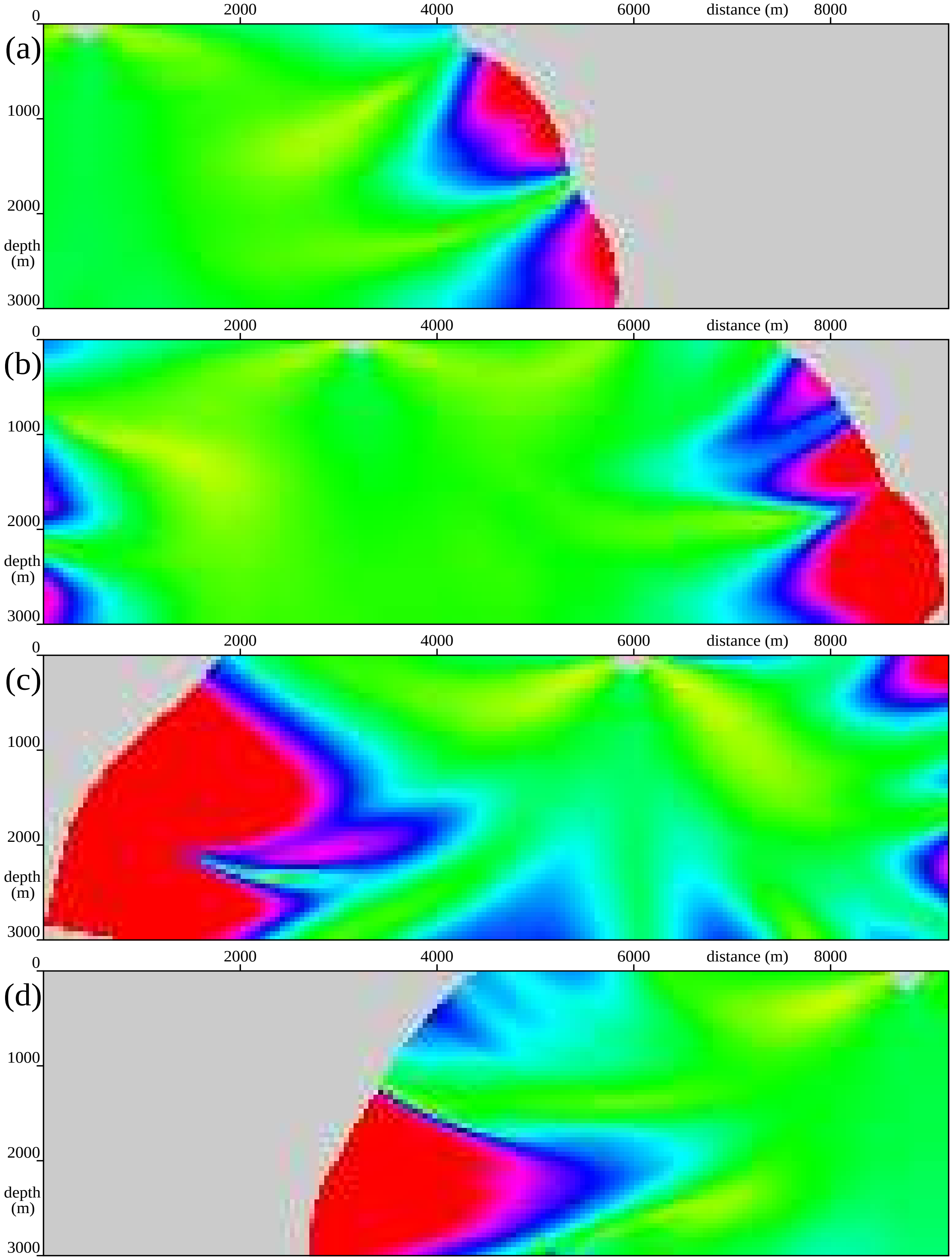

Figure 5. The halfwidths of GBs for the position of the source of (a) 425m, (b) 3200m, (c) 5975m and (d) 8775m. The *green* colour corresponds to the GB halfwidth of 202 and 378 metres for the frequencies of 35 Hz and 10 Hz, respectively. The *red* colour corresponds to the GB halfwidth of 1010 and more, and of 1890 metres and more for the frequencies of 35 Hz and 10 Hz, respectively. The coefficients used for smoothing the optimum initial parameters of GBs are $s_R = 1 \times 10^{13}\mathrm{m}^3\mathrm{s}^{-1}$ and $s_Y = 1 \times 10^{13}\mathrm{m}^3\mathrm{s}^{-1}$, two iterations.

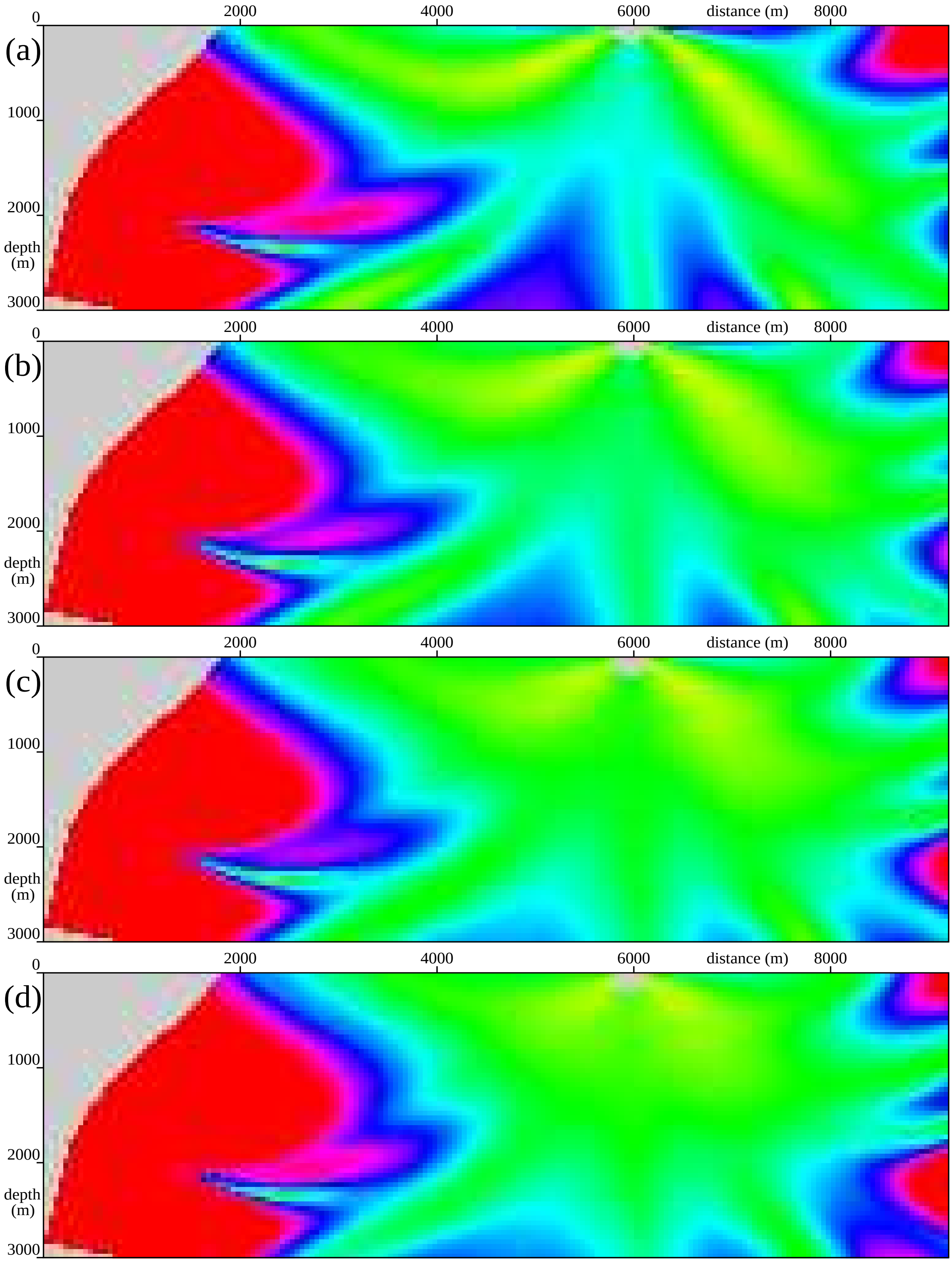

Figure 6. The halfwidths of GBs for the value of the initial parameter $R_0 = -0.26\times10^{-6}\mathrm{sm}^{-2}$ and for the value of the initial parameter Y_0 of (a) $0.29\times10^{-7}\mathrm{sm}^{-2}$, (b) $0.59\times10^{-7}\mathrm{sm}^{-2}$, (c) $0.12\times10^{-6}\mathrm{sm}^{-2}$ and (d) $0.24\times10^{-6}\mathrm{sm}^{-2}$. The *green* colour corresponds to the GB halfwidth of 202 and 378 metres for the frequencies of 35 Hz and 10 Hz, respectively. The *red* colour corresponds to the GB halfwidth of 1010 and more, and of 1890 metres and more for the frequencies of 35 Hz and 10 Hz, respectively.

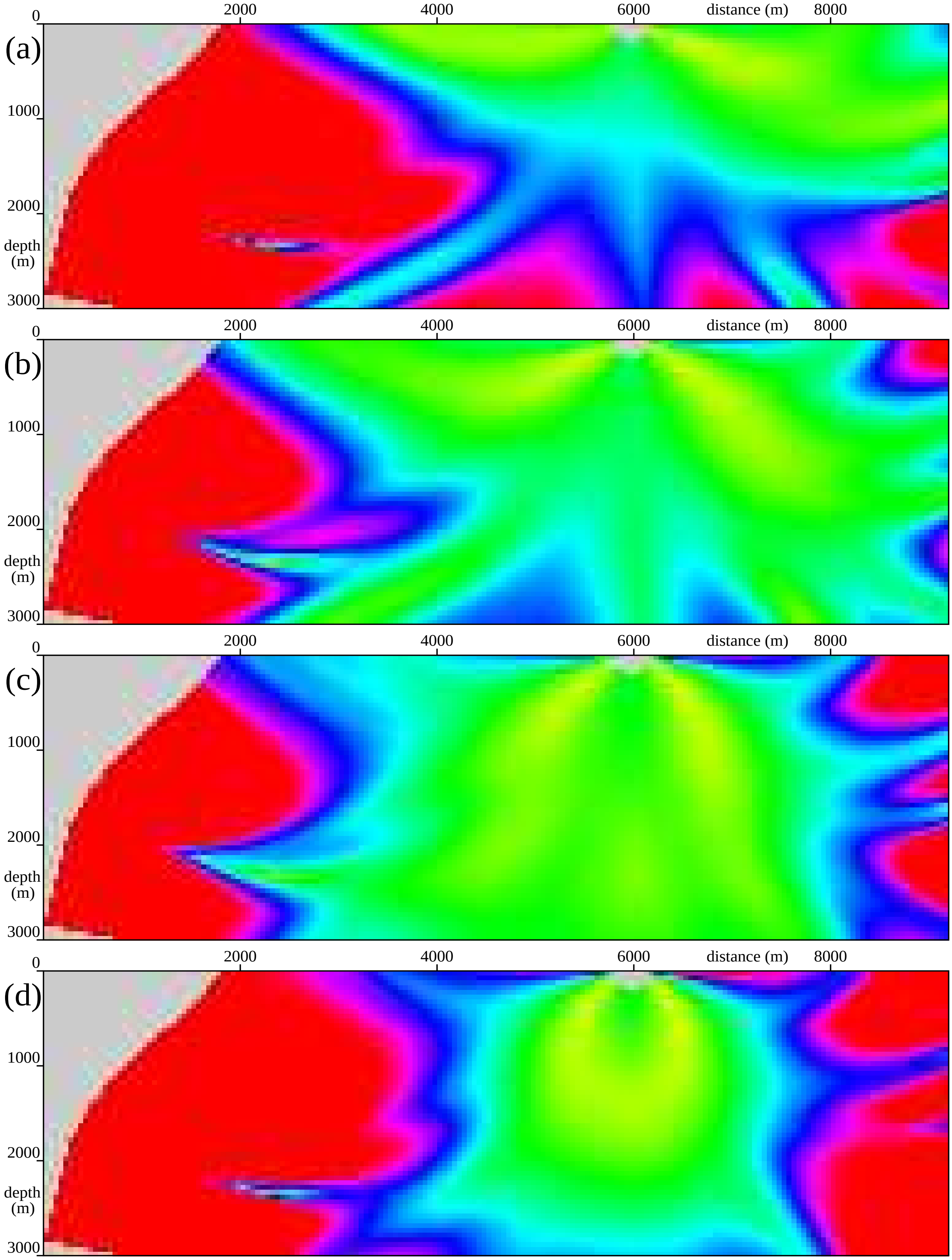

Figure 7. The halfwidths of GBs for the value of the initial parameter $Y_0 = 0.59 \times 10^{-7}\mathrm{sm}^{-2}$ and for the value of the initial parameter R_0 of (a) $-0.13 \times 10^{-6}\mathrm{sm}^{-2}$, (b) $-0.26 \times 10^{-6}\mathrm{sm}^{-2}$, (c) $-0.39 \times 10^{-6}\mathrm{sm}^{-2}$ and (d) $-0.52 \times 10^{-6}\mathrm{sm}^{-2}$. The *green* colour corresponds to the GB halfwidth of 202 and 378 metres for the frequencies of 35 Hz and 10 Hz, respectively. The *red* colour corresponds to the GB halfwidth of 1010 and more, and of 1890 metres and more for the frequencies of 35 Hz and 10 Hz, respectively.

of 1010 metres and more for the frequency of 35Hz, and of 1890 metres and more for the frequency of 10Hz.

In Figures 2 and 3, the optimum inital parameters of GBs are very little smoothed (just for numerical purposes) with one (Figure 2) and two (Figure 3) iterations. As you can see, there is no significant difference between the figures. This also corresponds to Table 1, where is no considerable jump in the relative RMS differences of the inital parameters of GBs.

In Figures 4 and 5, the optimum inital parameters of GBs are smoothed to a constant value with one (Figure 4) and two (Figure 5) 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, beacause they do not change anymore.

The halfwidths of GBs for the value of the initial parameter $R_0 = -0.26\times10^{-6}\mathrm{sm}^{-2}$ and for the value of the initial parameter Y_0 of $0.29 \times 10^{-7}\mathrm{sm}^{-2}$ (which corresponds to the optimum initial parameters of GBs smoothed to a constant value with one iteration, see Figure 4c), $0.59 \times 10^{-7}\mathrm{sm}^{-2}$ (which corresponds to the optimum initial parameters of GBs smoothed to a constant value with two iteration, see Figure 5c), $0.12\times10^{-6}\mathrm{sm}^{-2}$ and $0.24 \times 10^{-6}\mathrm{sm}^{-2}$ are in Figure 6.

The halfwidths of GBs for the value of the initial parameter $Y_0 = 0.59 \times 10^{-7}\mathrm{sm}^{-2}$ and for the value of the initial parameter R_0 of $-0.13 \times 10^{-6}\mathrm{sm}^{-2}$, $-0.26 \times 10^{-6}\mathrm{sm}^{-2}$ (which corresponds to the optimum initial parameters of GBs smoothed to a constant value with two iteration, see Figure 5c), $-0.39 \times 10^{-6}\mathrm{sm}^{-2}$ and $-0.52 \times 10^{-6}\mathrm{sm}^{-2}$ are shown in Figure 7.

As we can see, the constant optimum initial parameters of GBs obtained by the presented procedure give the best results in terms of GB width. Although a slightly different choice of the initial parameters can be also good (eg., see Figure 6c), we should not forget, that we have achieved the optimum parameters automatically. This is a great advantage of this method.

Currently, we cannot say anything meaningful about the influence of the optimization of the shape of GBs on the Gaussian beam or packet migrations. Naturally, we believe it will improve the accuracy of these methods.

9 Conclusions

The presented procedure of determining the optimum initial shape of GBs can improve the applicability and accuracy of the Gaussian beam or packet method. We obtain the optimum parameters almost automatically. Only the smoothness of the distribution of the optimum parameters of GBs, which plays the key role in the decomposition of the wave field into Gaussian beams or packets, have to be chosen by the user.

Naturally, it is not necessary to use this procedure in simple structures. But the more complex is the model, the more important is to optimize the initial parameters of GBs. Furthemore, especially in simple structures, we can smooth the optimum initial parameters of GBs to make them constant on the whole *Hamiltonian hypersurface*. It is more comfortable than trying to obtain the constant optimum initial parameters of GBs just by chance.

Let us remind that our goal was to find the optimum initial parameters of GBs in terms of GB width and that we have not tested the influence of the choice of the initial parameters on the wave field. This is a theme for further study.

Acknowledgements

The author wishes to thank Luděk Klimeš for his kind guidance throughout the work on this paper. This research has been supported by the Grant Agency of the Czech Republic under Contract 205/01/0927, by the Ministry of Education of the Czech Republic within Research Project J13/98 113200004, by the Grant Agency of the Charles University under Contract 237/2001/B–GEO/MFF and by the members of the consortium "Seismic Waves in Complex 3–D Structures" (see "http:/seis.karlov.mff.cuni .cz/consort/main.htm").

References

Červený, V., Popov, M.M. & Pšenčík, I. (1982): Computation of wave fields in inhomogeneous media — Gaussian beam approach. Geophys. J. R. astr. Soc., **70**, 109–128.

Červený, V. & Pšenčík, I. (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, in press.

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. (1989): Optimization of the shape of Gaussian beams of a fixed length. Stud. geophys. geod., **33**, 146–163.

Popov, M.M. & 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.

Versteeg, R.J. & Grau, G. (eds.) (1991): The Marmousi experience. Proc. EAGE workshop on Practical Aspects of Seismic Data Inversion (Copenhagen, 1990), Eur. Assoc. Explor. Geophysicists, Zeist.

Žáček, K. (2002): Smoothing the Marmousi model. Pageoph, **159**, in press.