Department of Geophysics
Faculty of Mathematics and Physics
Charles University in Prague

Gaussian Packet Prestack Depth Migration

Doctoral Thesis

Karel Žáček

Supervisor: RNDr. Luděk Klimeš, DrSc.

Prague
2005

I am greatly indebted to Luděk Klimeš for a kind guidance throughout my work on this thesis and for everything he taught me.

I will never forget the time I spent with Ross Hill — thanks for the opportunity, inspiration and all those magic trips.

This thesis would have never been completed without the invaluable technical support of Václav Bucha and Petr Bulant.

I also wish to express my thanks to Professor Vlastislav Červený, Johana Brokešová and Ivan Pšenčík for their comments and advice.

This research has been supported by the Grant Agency of the Charles University under Contracts 375/2004/B-GEO/MFF and 237/2001/B–GEO/MFF, by the Grant Agency of the Czech Republic under Contracts 205/04/1104 and 205/01/0927, by the Ministry of Education of the Czech Republic within Research Project J13/98 113200004, and by the members of the consortium "Seismic Waves in Complex 3-D Structures".

My heart will always stay with Hanka, Susana and Alex — sincere thanks for everything.

Prohlašuji, že jsem svou doktorskou práci napsal samostatně a výhradně s použitím citovaných pramenů. Souhlasím se zapůjčováním práce.

V Praze dne 21.6.2005 Karel Žáček

Contents

Decomposition of the wave field into optimized Gaussian packets

Gaussian packet prestack depth migration

Preface

This doctoral thesis consists of an expanded abstract and four separate papers. Each paper is devoted to one step of the Gussian packet prestack depth migration algorithm:
a) "Smoothing the Marmousi model", which introduces a method of smoothing the velocity model by minimizing the relevant Sobolev norm of slowness.
b) "Optimization of the shape of Gaussian beams", which explains how to decrease the spreading of Gaussian beams and packets by optimizing their initial shape.
c) "Decomposition of the wave field into optimized Gaussian packets", which offers a procedure for determining the initial amplitudes of back-propagated Gaussian packets.
d) "Gaussian packet prestack depth migration", which describes the back-propagation of the wave field using Gaussian packets and application of the imaging functional, and which summarizes the whole Gussian packet migration algorithm.

We present all four papers in their original form as they were published in the journal *Pure and applied geophysics* and in progress reports of the consortium *Seismic Waves in Complex 3-D Structures*.

The results of the project have been presented at annual meetings of the Society of Exploration Geophysicists (SEG) and of the European Association of Geoscientists & Engineers (EAGE). The author of this thesis received the "Award of Merit — Best Student Paper Presented at 2001 SEG Annual Meeting" for the contribution "Optimization of the shape of Gaussian beams", "Award of the Minister of Education" (2001), which is awarded annually to five outstanding university students in the Czech Republic, and the 1^{st} position in the student category of the "Babuška Prize" (2001) for the papers "Smoothing the Marmousi model" and "Optimization of the shape of Gaussian beams".

Abstract

1 Introduction

The ultimate goal of seismic processing is to produce a correct image of a geological structure. Currently, prestack depth migration represents the most accurate (even though the most expensive) tool for imaging complex subsurface stuctures. Our project follows the pioneering work of N. Ross Hill in the field of prestack depth migrations based on Gaussian beams (Hill 1990, 2001).

Instead of using Gaussian beams, we have shifted our attention to Gaussian packets (e.g., Babich & Ulin, 1981; Ralston, 1983; Klimeš, 1989a, 2004). Like Gaussian beams, Gaussian packets represent high-frequency asymptotic solutions of the elastodynamic equation. But while Gaussian beams are concentrated close to the *central ray* of the beam, Gaussian packets are concentrated close to the *central point* of the packet. The main advantage of the Gaussian packet migration over the methods based on Gaussian beams is a direct relation between the regions in the recorded wave field and corresponding localized regions in the migrated section.

Naturally, Gaussian packets spread as they propagate through the structure. It is necessary to keep Gaussian packets narrow in relation to the velocity changes in the model, because Gaussian packets become inaccurate solutions of the elastodynamic equation if the velocity field changes considerably within the packet width. This spreading depends on the complexity of the velocity model and on the initial shape of Gaussian packets. Therefore, before proceeding to the migration, we need to prepare a suitable velocity model and choose the appropriate initial shape of Gaussian packets.

The Gaussian packet migration algorithm consists of four basic steps:
a) Preparation of a suitable smooth velocity model (Žáček, 2002).
b) Optimization of the shape of Gaussian packets (Klimeš, 1989b; Žáček, 2001a, 2001b).
c) Decomposition of the recorded wave field into optimized Gaussian packets (Žáček, 2003, 2005a).
d) Back-propagation of the recorded wave field using Gaussian packets and application of the imaging functional (Žáček 2004, 2005b).

The word *optimized* implies that the shape of Gaussian packets, in the plane perpendicular to the central ray of the packet, depends not only on the frequency, but also on the coordinate of the intersection of the central ray of a Gaussian packet with the profile, on its arrival time, and on the component of the slowness vector along the profile.

2 Velocity model

A velocity model should prevent Gaussian packets from excessive spreading and approximately preserve original travel times. The question of the smoothness of the velocity model is closely related to the problem of finding the limits of applicability of the ray theory, which remains unsolved and open for further research.

In a complex model, the geometrical spreading and number of arrivals exponentially increase with increasing travel time. The exponential increment is controlled by the Lyapunov exponent (e.g., Klimeš, 2002). Consequently, the Lyapunov exponent determines the horizon, where the ray behavior becomes chaotic. Since the Lyapunov exponent depends on the second spatial derivatives of velocity or slowness, the second

derivatives should be minimized. Therefore, we smooth slowness by minimizing the relevant Sobolev norm composed of the second derivatives of slowness (Žáček, 2002).

3 Optimization of the shape of Gaussian packets

The applicability and accuracy of the Gaussian packet method depend on the proper choice of the initial shape of packets. Unfortunately, narrow Gaussian packets quickly increase in width as they propagate. Thus, we can use neither too narrow nor too wide packets as the initial choice of the shape of Gaussian packets. Furthermore, in a complex structure, we cannot judge solely from the final width of the packet whether the packet is or is not a reasonably accurate solution of the elastodynamic equation. The packet must be sufficiently narrow along the whole ray path.

The initial shape of Gaussian packets is determined by the matrix of complex-valued second space-time derivatives of the phase function. To prevent Gaussian packets from excessive spreading, we optimize their initial shape by minimizing their width along the whole ray path. Since we wish to control the spreading in the plane perpendicular to the central ray of the packet, it is sufficient to optimize the initial shape of corresponding Gaussian beams (Klimeš, 1989b; Žáček, 2001b). We can restrict our attention to Gaussian beams, because the Ricatti equation for a Gaussian packet may be decoupled into three equations, where one equation represents the Ricatti equation for a Gaussian beam, and the other two equations extend the Gaussian beam solution to a Gaussian packet solution.

In 2-D, the shape of a Gaussian beam is characterized by a single complex-valued parameter, see Žáček (2001b, 2005a, 2005b). The imaginary part of this parameter determines the Gaussian beam width and the real part defines the curvature of the phase-front at the reference point.

The parameter, which describes the shape of a corresponding Gaussian beam along the initial surface, may depend on the coordinate of the intersection of the central ray of a Gaussian packet with the profile, on its arrival time, and on the component of the slowness vector along the profile. In other words, this parameter is situated on a Hamiltonian hypersurface in the phase-space (Žáček, 2001b). In a complex structure, the optimum initial parameters may vary in orders of magnitude. This would cause great problems in the decomposition of the wave field into optimized Gaussian packets. Therefore, we have developed a procedure, which allows us to smooth iteratively the distribution of the initial parameter on the Hamiltonian hypersurface (Žáček, 2001b).

To simplify further calculations, we can smooth the distribution of the initial parameter to a constant value and obtain *uniform* Gaussian packets. Naturally, the smoothing brings about more rapid spreading of Gaussian packets.

4 Decomposition of the wave field into Gaussian packets

The recorded wave field, which consists of seismic traces that have different receiver locations, but that were generated by the same shot, is termed a *common-shot gather*. In order to determine the initial amplitudes of Gaussian packets, we need to decompose the common-shot gather into optimized Gaussian packets (Žáček, 2005a). Each Gaussian packet arriving at the receivers is represented by a Gabor function, whose shape depends on the shape of the Gaussian packet. Using the coherent-state transform, we decompose the common-shot gather into these Gabor functions. The complex valued amplitude of

the Gaussian packet then equals the amplitude of the corresponding Gabor function. The amplitudes depend on the coordinate of the intersection of the central ray of a Gaussian packet with the profile, its arrival time, the component of the slowness vector along the profile, and the frequency. Let us point out that the intersection of the central ray of a Gaussian packet with the profile does not generally coincide with any receiver.

5 Migration

The Gaussian packet migration operates in the common-shot domain. The incident wave field may be calculated in various ways. We benefit from a method of interpolation within ray cells proposed by Bulant & Klimeš (1999), which requires only a sufficiently dense set of rays from the shot location to be traced. By interpolation, we obtain the multi-valued amplitudes of the incident wave field corresponding to multi-valued travel times.

We determine the scattered wave field using the Gaussian packet method. Gaussian packets propagate along their central rays. We move along these central rays, determine the amplitude and the second space-time derivatives of the phase function of the back-propagated Gaussian packet, and apply the imaging functional, which yields the localized image of a single Gabor function from the common-shot gather (Žáček, 2005b). Such image represents the basic building block of the Gaussian packet migration. Then, we superimpose the images of all Gabor functions corresponding to a single common-shot gather, and obtain the prestack migrated image of this common-shot gather. Finally, we can stack all of the images to produce the migrated image of the whole data set.

6 Target-oriented imaging

Each Gabor function from the common-shot gather generates its localized image in the depth section. In this way, we obtain a one-to-one relation between the Gabor functions from the common-shot gather and their localized images from the depth section. This relation was discussed in greater detail by Žáček & Klimeš (2003).

Let us say that we are interested in a particular area in the subsurface structure. We call this area a *target zone*. Unlike Gaussian beams, Gaussian packets are, at any given time, concentrated close to the central point of the packet. Thus, we can consider only those packets, which fall into the target zone.

Moreover, we can pick out Gaussian packets contributing to the target zone, multiply their amplitudes by corresponding weighting factors, and use them in the recomposition of the common-shot gather. In such a case, we obtain the wave field scattered specifically from the target zone.

7 Numerical examples

We tested our method on the Marmousi data set (Versteeg & Grau, 1991). Following the procedure explained in Section 4, we decompose the common-shot gather, see Figure 1a, into individual Gabor functions, see Figure 1b. The complex-valued amplitude of a Gaussian packet equals the amplitude of the corresponding Gabor function. Then, we calculate the localized image of a single Gabor function from the common-shot gather, see Figure 1c. Let us point out that Figures 1b and 1c demonstrate the unique one-to-one relation between the Gabor function from the common-shot gather and its localized

image from the migrated section. Finally, we superimpose all relevant images and obtain the prestack migrated image of the common-shot gather, see Figure 1d.

The Marmousi model represents a very complex structure, see Figure 2a. The dimensions of the model are 9200 metres (length) by 3000 metres (depth). The values of velocity, which are defined on a grid of cells of 4×4 metres, vary from 1500 ms^{-1} to 5500 ms^{-1}. Since the original model is not suitable for any ray method, we minimized the relevant Sobolev norm of slowness as described in Section 2 and obtained the smooth velocity model, see Figure 2b. The model is characterized by its *average Lyapunov exponent* λ (Klimeš, 2002; Žáček, 2002), which controls the average geometrical spreading of the ray field. The average Lyapunov exponent for this model reads $\lambda = 0.89$ s^{-1}.

The stacked migrated section of the Marmousi data set is displayed in Figure 2c. In order to speed the calculation up, we used a sparser grid of cells of 20×20 metres. Although we can recognize several features of the Marmousi model, we failed in reconstructing the bottom part of the model. We believe three main causes of such result consist in (a) the use of *uniform* Gaussian packets, (b) an oversimplified stacking procedure and (c) a considerably smoothed velocity model, see Versteeg (1993). Nevertheless, we are still intensively testing our method, and therefore the numerical example shown in Figure 2c should be considered as very preliminary.

In the future, we would like to refine the method of stacking in order to improve the quality of stacked migrated sections. Moreover, we should consider the application of Gaussian packets with varying initial parameters. This would make the Gaussian packet prestack depth migration more expensive, but also more suitable for complex velocity models.

8 Conclusions

The Gaussian packet prestack depth migration represents a new and promising imaging method operating in the common-shot domain. It can handle multi-valued travel times and allows a target-oriented approach.

Although the common-shot Gaussian packet migration cannot match the efficiency of the common-offset Gaussian beam migration algorithm proposed by Hill (2001), it is more general and provides a one-to-one relation between the Gabor functions from the common-shot gather and their localized images from the depth section. We hope that this unique feature could help us in understanding the true meaning of the migrated section.

Ultimately, we would like to inspire and encourage other researchers and possibly attract their attention to the Gaussian packet domain. We believe that it could show even greater potential than the well-known Gaussian beam domain.

References

Babich, V. M. & Ulin, V. V. (1981): Complex space time ray method and 'quasiphotons'. Zap. Nauchn. Sem. Leningr. Otd. Math. Inst., **117**, 5–12 (in Russian); English transl.: J.Sov. Math., **24**, 269–273, 1984.

Bulant, P. & Klimeš, L. (1999): Interpolation of ray-theory travel times within ray cells. Geophys. J. Int., **139**, 273–282.

Hill, N. R. (1990): Gaussian beam migration. Geophysics, **55**, 1416–1428.

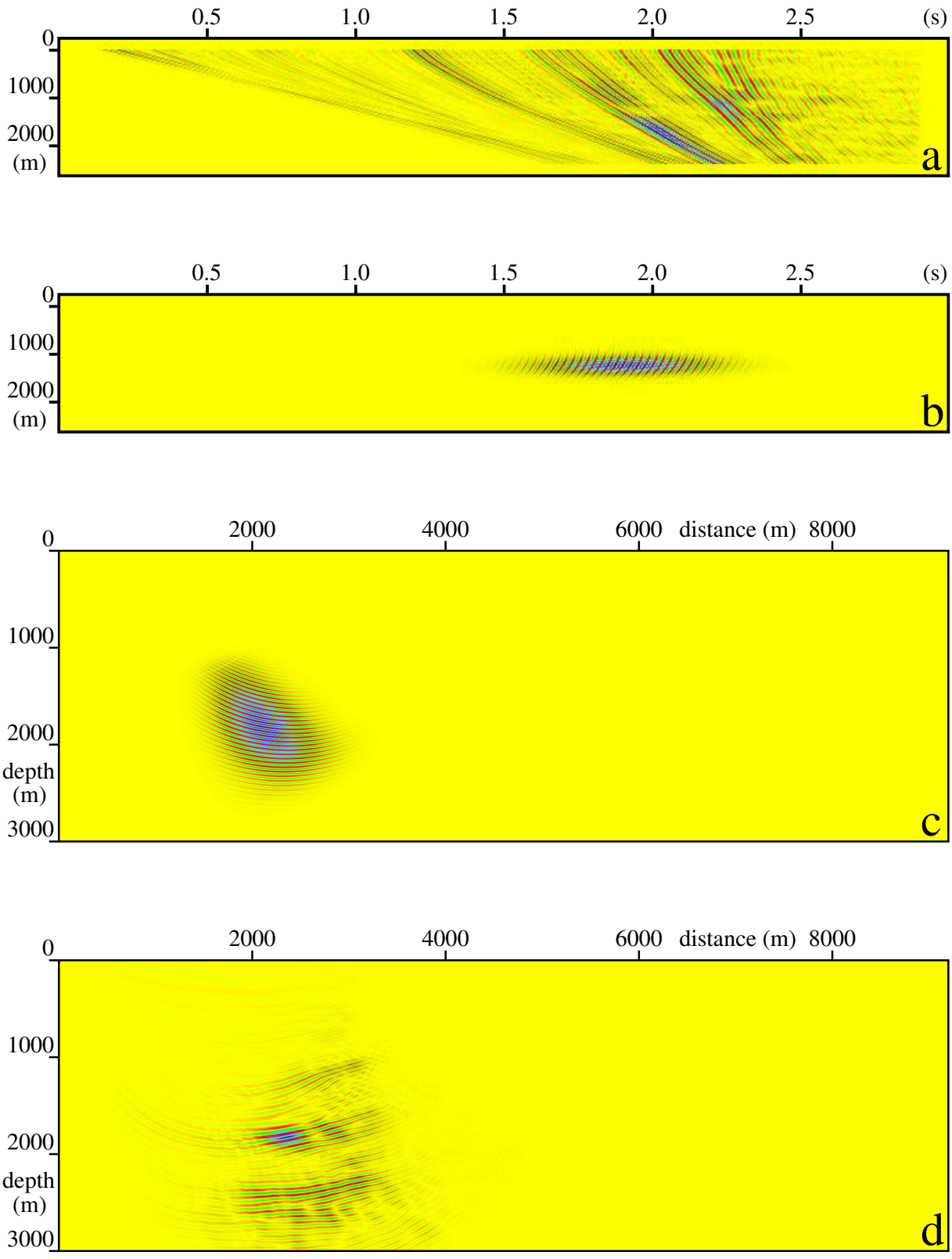

Figure 1. We decompose the common-shot gather of the Marmousi data set (a) into individual Gabor functions (b). Each Gabor function generates its localized image in the depth section (c). Superimposing all relevant images, we obtain the migrated image of the common-shot gather (d).

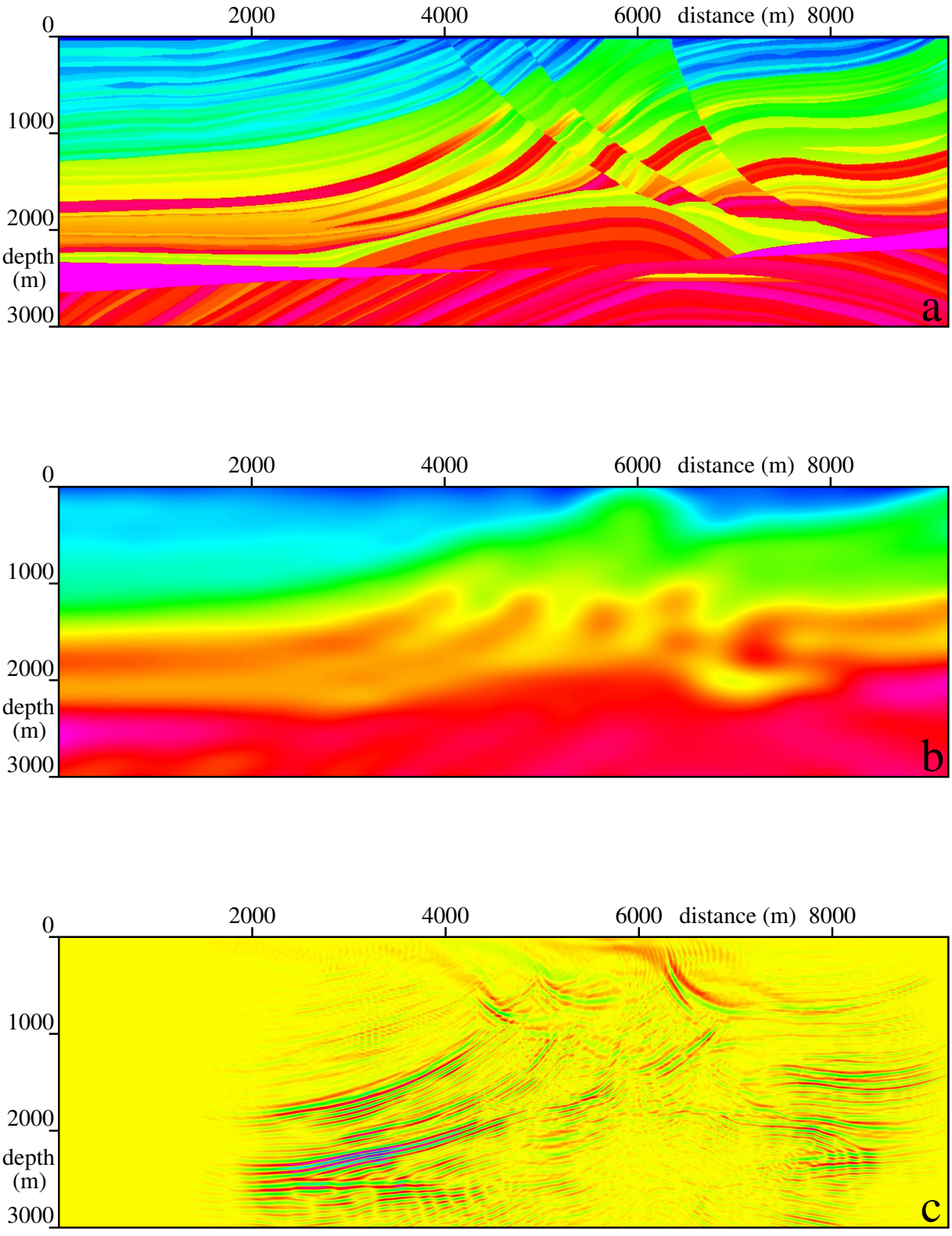

Figure 2. (a) The original Marmousi model, (b) the moothed velocity model with the average Lyapunov exponent $\lambda = 0.89\ s^{-1}$, and (c) the stacked migrated section of the Marmousi data set. We used a sparse grid of cells of 20×20 metres and applied no post-processing.

Hill, N. R. (2001): Prestack Gaussian-beam depth migration. Geophysics, **66**, 1240–1250.

Klimeš, L. (1989a): Gaussian packets in the computation of seismic wavefields. Geophys. J. Int., **99**, 421–433.

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

Klimeš, L. (2002): Lyapunov exponents for 2-D ray tracing without interfaces. Pure and Appl. Geophys., **159**, 1465–1485.

Klimeš, L. (2004): Gaussian packets in smooth isotropic media. In: Seismic Waves in Complex 3-D Structures, Report 14, Charles University, Prague, 43–54.

Ralston, J. (1983): Gaussian beams and the propagation of singularities. MAA Studies in Mathematicc., **23**, 206–248.

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.

Versteeg, R. J. (1993): Sensitivity of prestack depth migration to the velocity model. Geophysics, **58**, 873–882.

Žáček, K. (2001a): Optimization of the shape of Gaussian beams. 71st Ann. Int. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 2128–2131.

Žáček, K. (2001b): Optimization of the shape of Gaussian beams. In: Seismic Waves in Complex 3-D Structures, Report 11, Charles University, Prague, 181–201.

Žáček, K. (2002): Smoothing the Marmousi model. Pure and Appl. Geophys., **159**, 1507–1526.

Žáček, K. (2003): Decomposition of the wave field into optimized Gaussian packets. 73st Ann. Int. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1869–1872.

Žáček, K. & Klimeš, L. (2003): Sensitivity of seismic waves to the structure. 73rd Ann. Int. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1857–1860.

Žáček, K. (2004): Gaussian packet pre-stack depth migration. 74th Ann. Int. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 957–960.

Žáček, K. (2005a): Decomposition of the wave field into optimized Gaussian packets. In: Seismic Waves in Complex 3-D Structures, Report 15, Charles University, Prague, 17–27.

Žáček, K. (2005b): Gaussian packet prestack depth migration. In: Seismic Waves in Complex 3-D Structures, Report 15, Charles University, Prague, 29–48.

Pure appl. geophys. 159 (2002) 1507–1526
0033–4553/02/081507–20 $ 1.50 + 0.20/0

Pure and Applied Geophysics

Smoothing the Marmousi Model

KAREL ŽÁČEK[1]

Summary — The only way to make an excessively complex velocity model suitable for application of ray-based methods, such as the Gaussian beam or Gaussian packet methods, is to smooth it. We have smoothed the Marmousi model by choosing a coarser grid and by minimizing the second spatial derivatives of the slowness. This was done by minimizing the relevant Sobolev norm of slowness. We show that minimizing the relevant Sobolev norm of slowness is a suitable technique for preparing the optimum models for asymptotic ray theory methods. However, the price we pay for a model suitable for ray tracing is an increase of the difference between the smoothed and original model. Similarly, the estimated error in the travel time also increases due to the difference between the models. In smoothing the Marmousi model, we have found the estimated error of travel times at the verge of acceptability. Due to the low frequencies in the wavefield of the original Marmousi data set, we have found the Gaussian beams and Gaussian packets at the verge of applicability even in models sufficiently smoothed for ray tracing.

Key words: Velocity model, smoothing, asymptotic ray theory, Gaussian beams, Lyapunov exponent, Sobolev norm.

1. Introduction

The computation of rays is extremely sensitive to the smoothness of the model. In rough models, the behaviour of rays becomes chaotic and geometrical spreading and the number of arrivals increase with travel time rapidly (e.g., SMITH *et al.*, 1992; ABDULLAEV, 1993; TAPPERT and TANG, 1996; WITTE *et al.*, 1996; KEERS *et al.*, 1997). Moreover, a large number of two-point rays to each receiver makes calculation of two-point travel times slow and expensive. Often, two-point rays cannot be found within the numerical accuracy.

We need a reasonably smooth velocity model for a depth migration technique-based on Gaussian packets. In the Gaussian packet method (e.g., KLIMEŠ, 1989), we do not need to find two-point rays, however a sufficiently dense set of rays must be calculated. Thus, the desired model should be suitable for ray tracing. Since we wish to keep the width of the Gaussian packets sufficiently small, the width of the

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

Gaussian packet also depending on frequency, the model should be sufficiently smooth for the frequencies under consideration.

Various methods of smoothing the velocity model have been developed and published. The authors have been interested in determining the most suitable physical quantity to be smoothed (MÜLLER and SHAPIRO, 2000; GOLD *et al*., 2000), in finding a way to smooth the velocity model (GRUBB and WALDEN, 1995; VERSTEEG, 1991; BRAC and NGUYEN, 1990), or in studying the effects of smoothing on the wavefield (VERSTEEG, 1991, 1993).

An optimum way to smooth a complex velocity model for ray-based methods, which is presented in this paper, is to minimize the appropriate Sobolev norm of the velocity or slowness. We show that minimizing the Sobolev norm may be used for efficiently controlling the behaviour of rays in complex structures.

2. The Marmousi Model

Since we wish to use a "realistic" 2-D velocity model, we have decided to smooth the Marmousi model (VERSTEEG and GRAU, 1991; VERSTEEG, 1991, 1993). The Marmousi model, based on a real geological structure, is very complex, see Figure 1. The dimensions of the model are 9200 metres (length) by 3000 metres (depth). Values of velocity, which correspond to P waves, are defined at each gridpoint of the grid of cells of 4×4 metres. The grided values of velocity vary from 1500 ms^{-1} to 5500 ms^{-1}.

In the Marmousi model, the synthetic seismograms were computed by the finite difference method (VERSTEEG and Grau, 1991). We wish to use these seismograms as the "real data" for the migration. The length of the seismograms is 2.9 seconds with a sampling interval of 4 milliseconds. A trapezoidal frequency filter determined by frequencies of 0 Hz, 10 Hz, 35 Hz and 55 Hz has been applied to the data by the developers of the Marmousi data set.

3. Basic Ideas about the Desired Model

The desired smoothed model must fulfil two main and, unfortunately, contradictive requirements:

(a) to be in "good agreement" with the original Marmousi model, and

(b) to be "sufficiently smooth" for ray tracing and Gaussian packet computations.

Under the term "good agreement," we understand a slight difference between the smoothed and original model, expressed in terms of the standard L2 Lebesgue norm.

The meaning of the term "sufficiently smooth" is more complicated. In a complex model, the geometrical spreading and number of arrivals exponentially increase with increasing travel time. The exponential increment is controlled by the Lyapunov exponent (LYAPUNOV, 1949; MCCAULEY, 1993; ADDISON, 1997; KLIMEŠ, 2001). Since

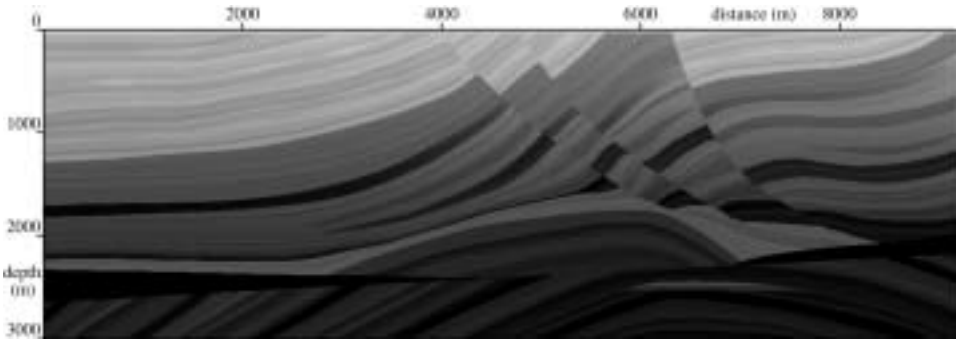

Figure 1
The Marmousi model.

the Lyapunov exponent depends on the second spatial derivatives of the velocity or slowness, the second derivatives should be minimized.

By minimizing the square of the Sobolev norm of slowness we may minimize the corresponding partial derivatives. 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).

The vague terms "good agreement" and "sufficiently smooth" cannot be easily quantified before a detailed study of the behaviour of rays and Gaussian packets in smoothed models is made.

The original Marmousi velocity model consists of discrete values of velocity at grid points of a regular, dense grid. In obtaining a smoothed model we

(a) choose a coarser data grid (which is a subgrid of the original grid) to reduce the amount of data to fit,
(b) arithmetically average the densely sampled slowness of the Marmousi model over cells centred at the grid points of the coarse data grid,
(c) choose a coarse B-spline model grid (which is a subgrid of the coarse data grid) and
(d) fit the averaged slowness values by the smoothed model.

We need to interpolate the discrete values of slowness on a coarse model grid for ray tracing. We have chosen bicubic B-splines as the interpolating functions, benefiting from the continuity of the second derivatives.

We summarize all types of grids being used in this paper. The first one is the original grid of the Marmousi model. The second is the coarser data grid constructed from the Marmousi model as explained above, which is used to fit the smoothed model. The third is the B-spline grid of the smoothed model.

4. Inversion

In order to find optimum parameters of the smoothed model, we minimize the objective function S defined by formula

$$S = \sum_{\mathrm{GRID}} \left(\frac{u^{\mathrm{D}}(\mathbf{x}^{\mathrm{GRID}}) - u^{\mathrm{M}}(\mathbf{x}^{\mathrm{GRID}})}{\sigma^{\mathrm{GRID}}} \right)^2 + \left[\int \mathrm{d}^2 x \right]^{-1} \int b_{ijkl} \left(\frac{\partial^2 u^{\mathrm{M}}(\mathbf{x})}{\partial x_i \partial x_j} \right) \left(\frac{\partial^2 u^{\mathrm{M}}(\mathbf{x})}{\partial x_k \partial x_l} \right) \mathrm{d}^2 x \ , \tag{1}$$

where u^{D} is the value of slowness in the data grid, u^{M} is the value of slowness in the model being sought, $\mathbf{x} = (x_1, x_2)$, σ^{GRID} are the weighting parameters of the grid points, b_{ijkl} are the weighting coefficients of the Sobolev scalar product. Superscript GRID takes values $\mathrm{GRID} = 1, 2, \ldots, N$, where N is the number of grid points of the coarser data grid with averaged values of slowness mentioned above. Subscripts take values $i, j, k, l = 1, 2$ in a 2-D model. Einstein summation over the pairs of identical indices is used. Integration is performed over the whole model.

We can express u^{M} as a linear combination of bicubic B-splines $B_\alpha(\mathbf{x})$

$$u^{\mathrm{M}}(\mathbf{x}) = B_\alpha(\mathbf{x}) u_\alpha \ , \tag{2}$$

where u_α are the model parameters (values of slowness at grid points of the B-spline grid). Subscript α takes values $\alpha = 1, 2, \ldots, P$, where P is the number of model parameters. Consequently, P is the number of B-splines describing the smoothed model.

Equation (1) now reads

$$S = \sum_{\mathrm{GRID}} \left(\frac{u^{\mathrm{D}}(\mathbf{x}^{\mathrm{GRID}}) - B_\alpha(\mathbf{x}^{\mathrm{GRID}}) u_\alpha}{\sigma^{\mathrm{GRID}}} \right)^2 + u_\alpha D_{\alpha\beta} u_\beta \ , \tag{3}$$

where

$$D_{\alpha\beta} = \left[\int \mathrm{d}^2 x \right]^{-1} \int b_{ijkl} \left(\frac{\partial^2 B_\alpha(\mathbf{x})}{\partial x_i \partial x_j} \right) \left(\frac{\partial^2 B_\beta(\mathbf{x})}{\partial x_k \partial x_l} \right) \mathrm{d}^2 x \ . \tag{4}$$

Since we do not know the coefficients b_{ijkl} which lead to the optimum model, the problem is not linear. Thus, parameters u_α cannot be determined analytically. Since we do not want to solve the nonlinear inverse problem numerically, we need to "linearize" formula (4). The linearization of (4) yields

$$D_{\alpha\beta} = s^2 D'_{\alpha\beta} \ , \tag{5}$$

$$D'_{\alpha\beta} = \left[\int \mathrm{d}^2 x \right]^{-1} \int b'_{ijkl} \left(\frac{\partial^2 B_\alpha(\mathbf{x})}{\partial x_i \partial x_j} \right) \left(\frac{\partial^2 B_\beta(\mathbf{x})}{\partial x_k \partial x_l} \right) \mathrm{d}^2 x \ , \tag{6}$$

where s is a free parameter and b'_{ijkl} are fixed coefficients of the Sobolev scalar product. The choice of coefficients b'_{ijkl} will be discussed in Section 6.

We can now rewrite equation (3) to read

$$S = [\mathbf{u}^{\mathrm{D}} - \mathbf{B}\mathbf{u}]^{\mathrm{T}} \mathbf{C}^{-1} [\mathbf{u}^{\mathrm{D}} - \mathbf{B}\mathbf{u}] + s^2 \mathbf{u}^{\mathrm{T}} \mathbf{D}' \mathbf{u} \ , \tag{7}$$

where $\mathbf{B}$ is defined as $B_{i\alpha} = B_\alpha(x_i)$, $\mathbf{D}'$ is a $P \times P$ matrix given by formula (6), $\mathbf{C}$ is a $N \times N$ diagonal matrix, composed of $(\sigma^{\mathrm{GRID}})^2$, see equation (1). N is the number of grid points.

The condition for the minimum of the objective function is

$$\frac{\partial S}{\partial u_\alpha} = 0 \ , \tag{8}$$

which yields

$$\mathbf{B}^{\mathrm{T}}\mathbf{C}^{-1}[\mathbf{B}\mathbf{u} - \mathbf{u}^{\mathrm{D}}] + s^2\mathbf{D}'\mathbf{u} = 0 \ . \tag{9}$$

The resulting vector of the model parameters is

$$\mathbf{u} = [\mathbf{B}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{B} + s^2\mathbf{D}']^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{u}^{\mathrm{D}} \ . \tag{10}$$

5. Criteria of Acceptability

In a complex 2-D model, the width of ray tube Q increases with increasing travel time τ approximately according to the asymptotic formula

$$Q \propto \mathrm{e}^{\lambda\tau} \ , \tag{11}$$

where λ is the Lyapunov exponent corresponding to the ray (LYAPUNOV, 1949; OSELEDEC, 1968; KATOK, 1980).

The number of arrivals at each point of the model is an important indication as to whether the behaviour of rays is regular or chaotic. We wish the number of arrivals not to exceed, let us say, 10. In a finite model, the number of arrivals v is proportional to the widths of the ray tubes. This is caused by the overlaping of the ray tubes, see Figure 5. As we wish to smooth the model for migration, the sum of travel times from source τ_{S} and receiver τ_{R} to a point of the model should be substituted for travel time τ in equation (11). Hence,

$$v \propto \mathrm{e}^{\lambda(\tau_{\mathrm{S}}+\tau_{\mathrm{R}})} \ . \tag{12}$$

Consequently, the number of arrivals v may be expressed as the product of the numbers of arrivals from source v_{S} and receiver v_{R}

$$v = v_{\mathrm{S}}v_{\mathrm{R}} \ . \tag{13}$$

For $\tau = \tau_{\mathrm{S}} + \tau_{\mathrm{R}} = 0$, we obtain $\mathrm{e}^{\lambda\tau} = 1$. Since this corresponds to the number of arrivals in the nearest vicinity of the source (or of the receiver for the migration), we can alter equation (12) to read

$$v \approx \mathrm{e}^{\lambda(\tau_{\mathrm{S}}+\tau_{\mathrm{R}})} \ . \tag{14}$$

We want to work with the "average Lyapunov exponent" $\hat{\lambda}$. The "average Lyapunov exponent" $\hat{\lambda}$ is the Lyapunov exponent averaged over a large set of rays (KLIMEŠ, 1999). The value of the "average Lyapunov exponent" may be one of the criteria of the smoothness of the model. Hence, we wish $e^{\hat{\lambda}\tau_{\max}}$ not to exceed 10, $\tau_{\max}$ being the maximum sum of travel times from the source and receiver to a point of the model. Since the sum of travel times from the source and receiver cannot exceed the length of the seismogram, $\tau_{\max} = 2.9$ s may be used for estimating the optimum value of $\hat{\lambda}$. Thus, for the number of arrivals not exceeding 10, we obtain the optimum value of $\hat{\lambda}$ close to 0.8 s^{-1}.

The width of Gaussian packets should be kept small. For very wide packets the obtained wavefield would not be the correct solution of the equations being solved. Accordingly, the desired migrated section would be wrong. The maximum halfwidth should probably not be greater than the B-spline interval.

We mention that the width of the Gaussian beams or packets depends not only on the smoothness of the model, but also on the frequencies under consideration. From this point of view the model is not complex for Gaussian beams or packets by itself, but in relation to the frequency.

The relative root-mean-square (RMS) difference of slowness between the original and the smoothed model may be the criterion of "good agreement." The relative RMS difference of slowness corresponds approximately to the relative error of the travel time. This is an asymptotic relation valid for short rays. The relative error of the travel time may be smaller for longer rays.

By the term "error of travel time" we understand the difference between the real travel time in the original structure and the computed travel time in the smoothed model. Although we cannot determine the real travel time, we can estimate the error caused by the difference between the original and the smoothed model.

6. Choice of the Coefficients and of the Density of the Grids

We need to specify coefficients b'_{ijkl} and s, the matrix $\mathbf{C}$ and the density of the grids before the computation.

As we have no prior information, we choose $\sigma^{\mathrm{GRID}} = \sqrt{N}$, where N is the number of values to be fitted. This makes the value of objective function S approximately independent of the number of gridpoints.

Coefficients b'_{ijkl} may be constructed as a completely symmetric tensor (e.g., BULANT, 2001). The 4×4 matrix $\mathbf{b}'$ is then defined by

$$b'_{ijkl} = \frac{d(d+2)}{3} \langle e_i e_j e_k e_l \rangle \ , \tag{15}$$

where $\mathbf{e}$ is a unit vector, $\langle \ldots \rangle$ indicates averaging over all directions of the unit vector, $d = 1$ in 1-D, $d = 2$ in 2-D and $d = 3$ in 3-D. We have introduced the formal scaling

coefficient $1/3d(d+2)$ in order to make the respective coefficients b'_{ijkl} equal in 1-D, 2-D and 3-D.

The average of the unit vector over all directions can be calculated analytically. Generally in d-D for $d = 1$, 2 or 3, we may put

$$b'_{ijkl} = \frac{\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}}{3} \quad , \tag{16}$$

where δ_{ij} is the Kronecker symbol. In 2-D,

$$b'_{1111} = b'_{2222} = 1,$$
$$b'_{1122} = b'_{1212} = b'_{1221} = b'_{2112} = b'_{2121} = b'_{2211} = \tfrac{1}{3} \text{ and}$$
$$b'_{1112} = b'_{1121} = b'_{1211} = b'_{1222} = b'_{2111} = b'_{2122} = b'_{2212} = b'_{2221} = 0 \ .$$

We must keep in mind that this is only our special choice of coefficients b'_{ijkl}, and that there are various other ways of constructing matrix $\mathbf{b}'$. For example, we can increase coefficient b'_{1111} and decrease coefficient b'_{2222} (or *vice versa*) and use this new matrix of coefficients for anisotropic smoothing.

The original grid of the Marmousi model consists of cells of 4×4 metres, which yields $751 \times 2301 = 1,728,051$ grid points. Three B-spline grids of cells of (a) 100×400 metres, (b) 200×230 metres and (c) 200×400 metres are studied. These grids consist of $P = 744$, 656 and 384 grid points, respectively.

Three data grids of cells of (a) 20×80 metres, (b) 40×40 metres and (c) 40×80 metres are used in the inversion. The values at the grid points are calculated by averaging the values of slowness in the Marmousi model, as described in Section 3. These grids consist of $N = 17516$, 17556 and 8816 grid points, respectively.

Finally, we need to choose the values of parameter s. We choose the initial value of parameter s for the linearized inversion as

$$s_{\text{init}} \approx \frac{|u - u_0|}{\sigma} \frac{1}{\|u\|_{\text{init}}} \quad , \tag{17}$$

where $\|u\|_{\text{init}}$ is the initial value of the Sobolev norm of slowness, $|u - u_0|$ is the standard slowness deviation of the model, and σ is the given slowness deviation. We have made this rough estimate assuming that the first term on the right-hand side of equation (1) does not exceed dramatically the second term, or *vice versa*. According to equation (A.10), we can estimate the maximum value of the Sobolev norm of slowness as $\|u\|_{\text{init}} \approx \sqrt{8/3}(2\Lambda_{\text{init}})^2 u_{\text{A}}^3$, where u_{A} is the average slowness in the model, and Λ_{init} is the initial value of the "average Lyapunov exponent" without compensation for the focusing low-velocity zones, see equation (A.1). The standard slowness deviation of the model may be estimated by $|u - u_0| \approx \varepsilon u_{\text{A}}$, ε being the relative travel-time error. The given slowness deviation is determined by $\sigma = \sigma^{\text{GRID}} N^{-1/2}$. Hence, we can alter equation (17) as

$$s_{\text{init}} \approx \sqrt{\frac{3}{8}}\, \varepsilon (2u_{\text{A}}\Lambda_{\text{init}})^{-2} \sqrt{N} \left(\sigma^{\text{GRID}}\right)^{-1} \quad . \tag{18}$$

Note that parameter s is proportional to $\hat{\lambda}^{-2}$, see equation (18). Thus, the n-fold decrement of $\hat{\lambda}$ requires an n^4-fold increment of the square of the Sobolev norm in the objective function in equation (1). Consequently, the decrement of $\hat{\lambda}$ increases the difference between the new and the original model. From this point of view $\hat{\lambda}$ should not be too small. We should keep $\hat{\lambda}$ close to the optimum value estimated above.

For $\varepsilon = 0.1$, $(u_A)^{-1} = 3000$ ms^{-1} and $\Lambda_{init} = 1.3$ s^{-1}, see equation (A.12), we obtain $s_{init} = 81529$. We study the values of s of (1) 0 m^2 (without the Sobolev norm included in the inversion), (2) $\sqrt{3/8} \cdot 10,000$ m$^2 \approx 6124$ m^2, (3) $\sqrt{3/8} \cdot 25,000$ m$^2 \approx 15,309$ m^2, (4) $\sqrt{3/8} \cdot 50,000$ m$^2 \approx 30,619$ m^2, (5) $\sqrt{3/8} \cdot 100,000$ m$^2 \approx 61,237$ m^2 and (6) $\sqrt{3/8} \cdot 225,000$ m$^2 \approx 137,784$ m^2.

7. Numerical Examples

We have calculated the smoothed models, the corresponding values of the relative RMS difference of slowness between the smoothed and original model, the angular dependence of the Lyapunov exponents, the values of the "average Lyapunov exponents," rays, numbers of arrivals and the halfwidths of Gaussian beams.

Figure 2 displays the models without the minimized Sobolev norm, smoothed just by the use of the coarse B-spline grid.

Figure 3 presents the smoothed models with the grid of cells of 200 × 400 metres and with the minimized Sobolev norm. The respective figures for the models with the grid of cells of 100 × 400 or 200 × 230 metres look similar and are not shown. At first glance we can see that the models with the values of parameter $s = 61{,}237$ m^2 and $s = 137{,}784$ m^2 do not show features of the original Marmousi model.

Figure 4 displays the angular dependence of the Lyapunov exponents and the values of the "average Lyapunov exponents" for smoothed models with the B-spline grid of cells of 200 × 400 metres. We can see that the model without the minimized Sobolev norm of slowness seems to be too rough, whereas the "average Lyapunov exponents" of the models with the minimized Sobolev norm are close to or less than our initial assumption of the optimum value. Unfortunately, the strong angular dependence of the Lyapunov exponents (and consequently the excessive maximum value of the Lyapunov exponent) indicates that even models with $\hat{\lambda}$ close to 1 (as the model with $s = 6124$ m^2) may still be too rough.

The synthetic seismograms, used for the migration in the Marmousi model, are computed with the length of 2.9 seconds. The streamer composed of hydrophone groups has been used for data acquisition for each shot. The farthest hydrophone was located 2575 metres from the watergun. Since the maximum travel time cannot exceed the length of the seismogram, let us assume a fixed travel time of 2.9 seconds. For an almost horizontal ray with the endpoint in the farthest hydrophone and for the velocity of 1500 ms^{-1} we obtain the farthest possible reflection point, at 4350 metres from the watergun and 900 metres from the farthest hydrophone. As 900

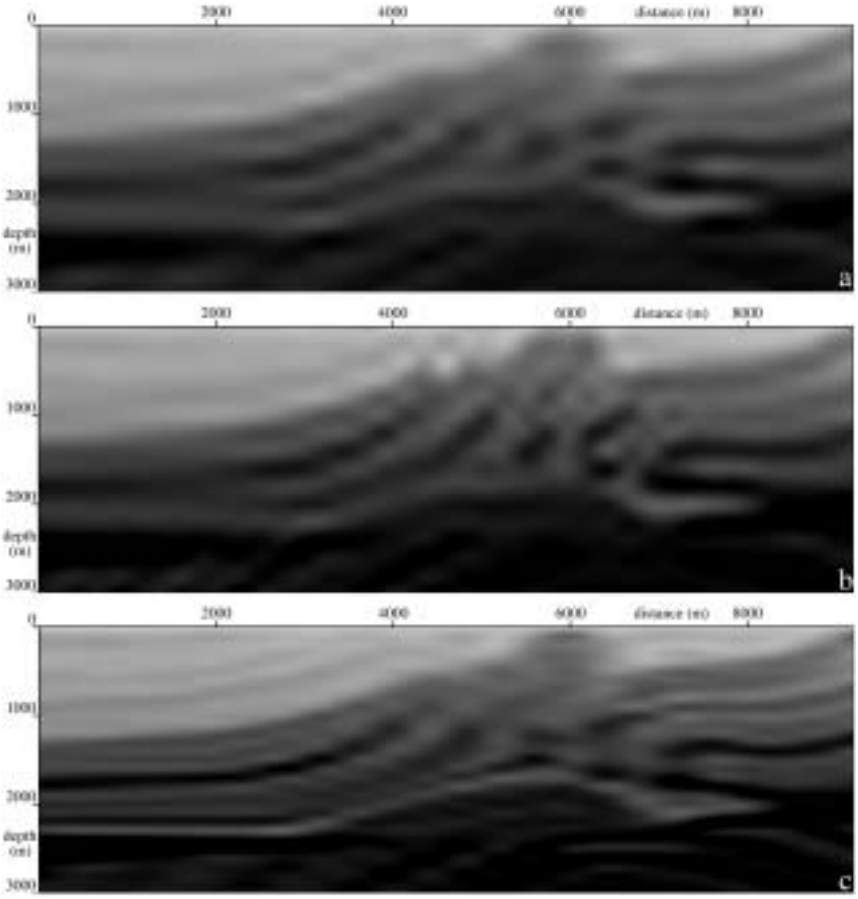

Figure 2
The smoothed models without minimized Sobolev norm of slowness ($s = 0$ m^2) for the B-spline grids of cells of (a) 200×400 metres, (b) 200×230 metres and (c) 100×400 metres.

metres corresponds to 0.6 seconds, we may estimate the maximum useful value of the travel time as 2.3 seconds. The rays have been calculated for this value of the maximum travel time.

Figure 5 displays rays computed in the smoothed models with a constant step in the take-off angles. We can see the dependence of the behaviour of rays on parameter s. We moved the source along the whole profile and tested ray tracing. The behaviour of rays was always of the same kind as in these illustrative figures. Models with $s = 0$ m^2 and $s = 6124$ m^2 seem to be unsuitable for ray methods due to the density of caustics. In Figure 5a we can see rays trapped in the low velocity channels.

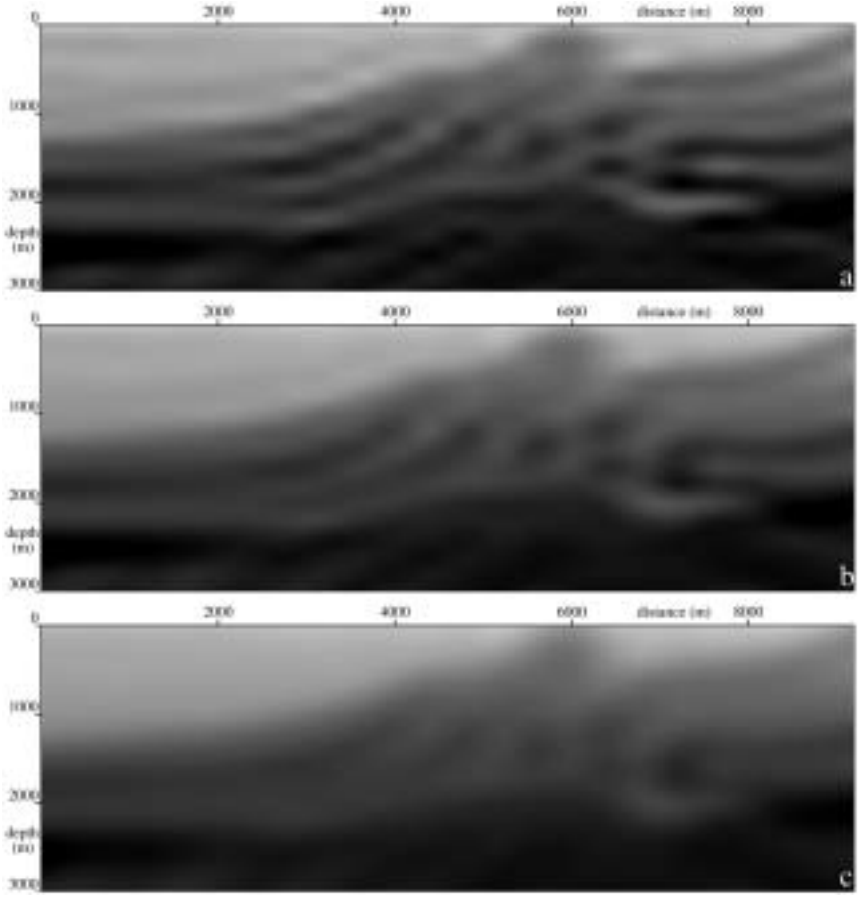

Figure 3
The smoothed models for the B-spline grid of cells of 200 × 400 metres and for values of parameter s of (a) 0 m^2, (b) 6124 m^2, (c) 15,309 m^2, (d) 30,619 m^2, (e) 61,237 m^2 and (f) 137,784 m^2.

The maximum number of arrivals for the models with the B-spline grid of cells of 200 × 400 metres is

(a) 19 for the value of $s = 0$ m^2,
(b) 18 for the value of $s = 6124$ m^2,
(c) 7 for the value of $s = 15{,}309$ m^2,
(d) 5 for the value of $s = 30{,}619$ m^2,
(e) 3 for the value of $s = 61{,}237$ m^2 and
(f) 2 for the value of $s = 137{,}784$ m^2.

In the model without interfaces, the number of arrivals should be odd. Even numbers for $s = 6124$ m^2 and $s = 137,784$ m^2 may be explained by the influence of the

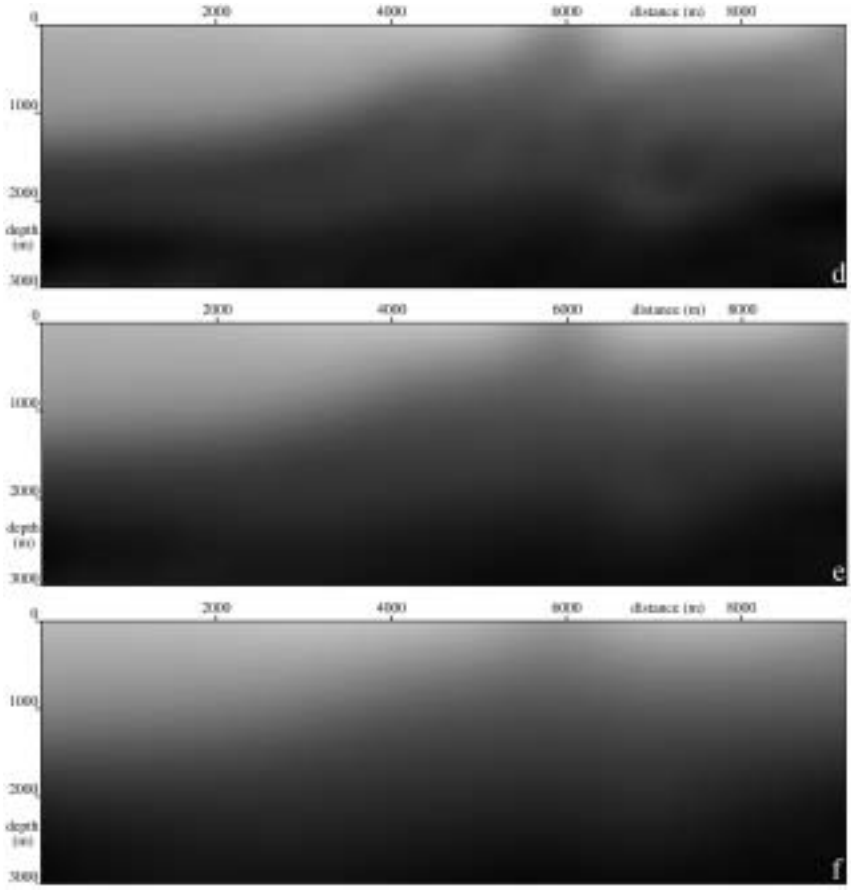

Figure 3d–f

borders of the model. Due to the requirements established above, the models with $s = 0$ m^2 and $s = 6124$ m^2 are probably not suitable for ray-based methods.

The relative RMS difference of slowness between all calculated models and the original Marmousi model is in Table 1. We can see that the price for a model suitable for ray tracing is a considerable increment of the relative RMS difference of slowness between the smoothed and original model, representing here the geological structure. This resulted from the complexity of the original Marmousi model. If the value of parameter s is larger, the relative RMS difference is the same for all the studied B-spline grids. Hence, the model with the B-spline grid of cells of 200×400 metres (with only 384 B-spline grid points) is probably the best choice.

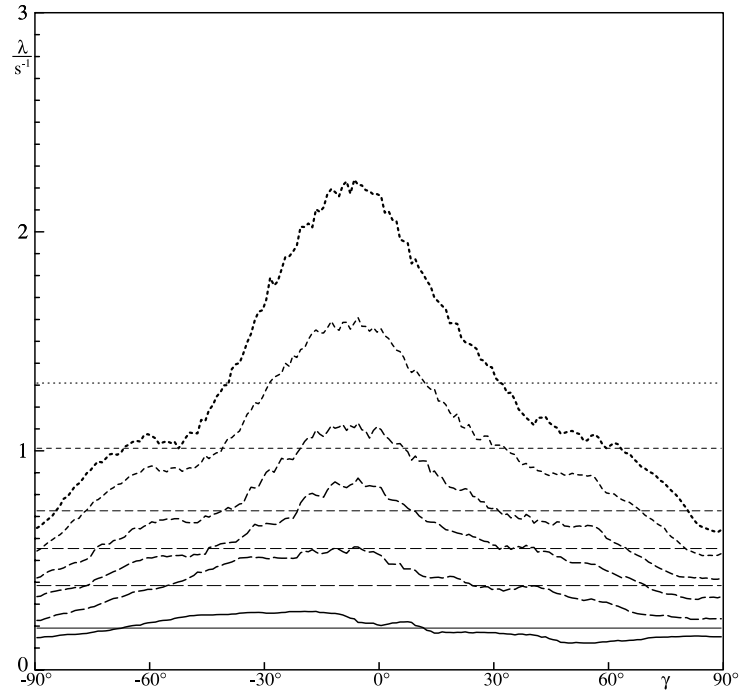

Figure 4
The angular dependence of the Lyapunov exponents for the models with the B-spline grid of cells of 200×400 metres. The left and right border corresponds to the vertical ray, the middle corresponds to the horizontal ray. The values of parameter s are (1) 0 m^2 for the dotted lines, (2) 6124 m^2, (3) 15,309 m^2, (4) 30,619 m^2 and (5) 61,237 m^2 for the dashed lines and (6) 137,784 m^2 for the solid lines. The thin horizontal lines correspond to the "average Lyapunov exponents," averaged over angles with a uniform weight.

In general, we believe it is useless to work with an overly dense B-spline grid, because the smoothed models with various densities of the model grid converge with increasing weight of the Sobolev norm.

Let us form a short summary of what we have accomplished. We have prepared the smoothed models. We have calculated the values of corresponding "directional" and "average Lyapunov exponents" and the values of the relative RMS difference between the smoothed and the original Marmousi model. Finally, we have studied the behaviour of rays in the smoothed models. That is, we already know how to smooth the Marmousi model for the computation of rays and travel times.

8. Effects of Smoothing on Gaussian Beams

Our primary objective was to prepare a suitable velocity model for Gaussian packet migration. Since the width of the Gaussian beam is equal to the width of the

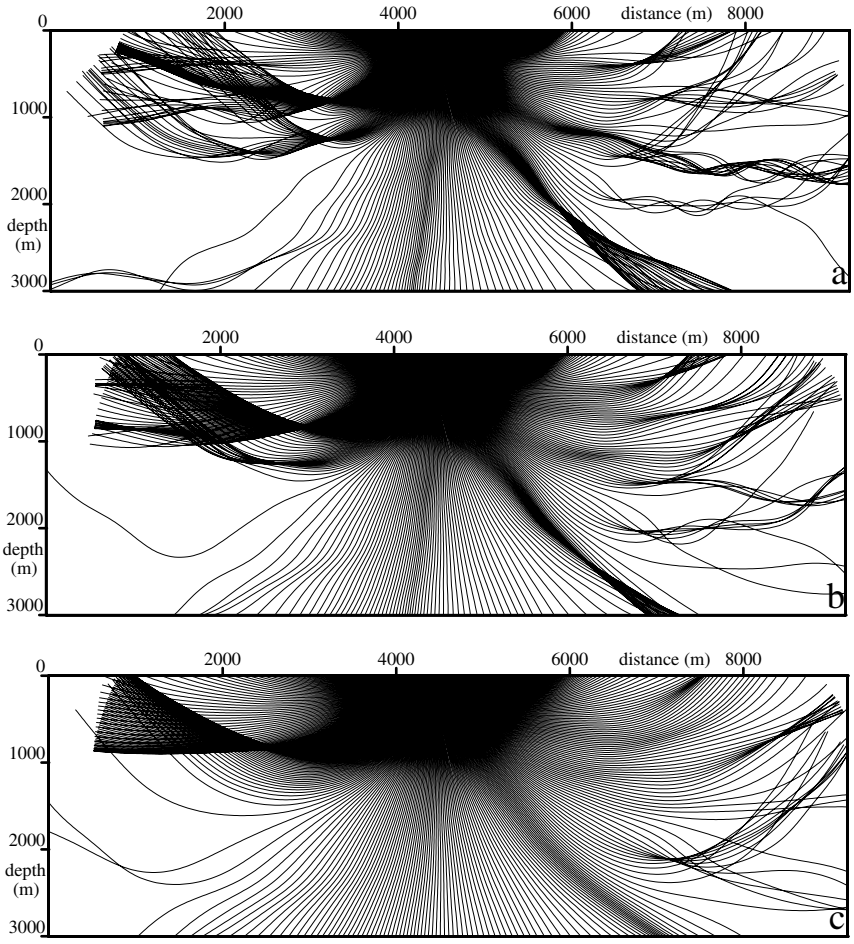

Figure 5
Rays in models with the B-spline grid of cells of 200×400 metres and with values of parameter s of (a) 0 m^2, (b) 6124 m^2, (c) 15,309 m^2, (d) 30,619 m^2, (e) 61,237 m^2, and (f) 137,784 m^2. The maximum travel time is of 2.3 seconds.

corresponding symmetric Gaussian packets, and the computation of the beams is easier, we study the width of the Gaussian beams. In 2-D, the profile of the Gaussian beam in a cross section orthogonal to the ray is controlled by the factor

$$\exp\left(\mathrm{i}\pi f M q^2\right) \ , \tag{19}$$

where i is the imaginary unit, f is frequency, q is the ray-centred coordinate orthogonal to the ray, and M is the second derivative of the complex-valued travel time. The quadratic term in the Taylor expansion of the complex-valued travel time of the Gaussian beam thus reads

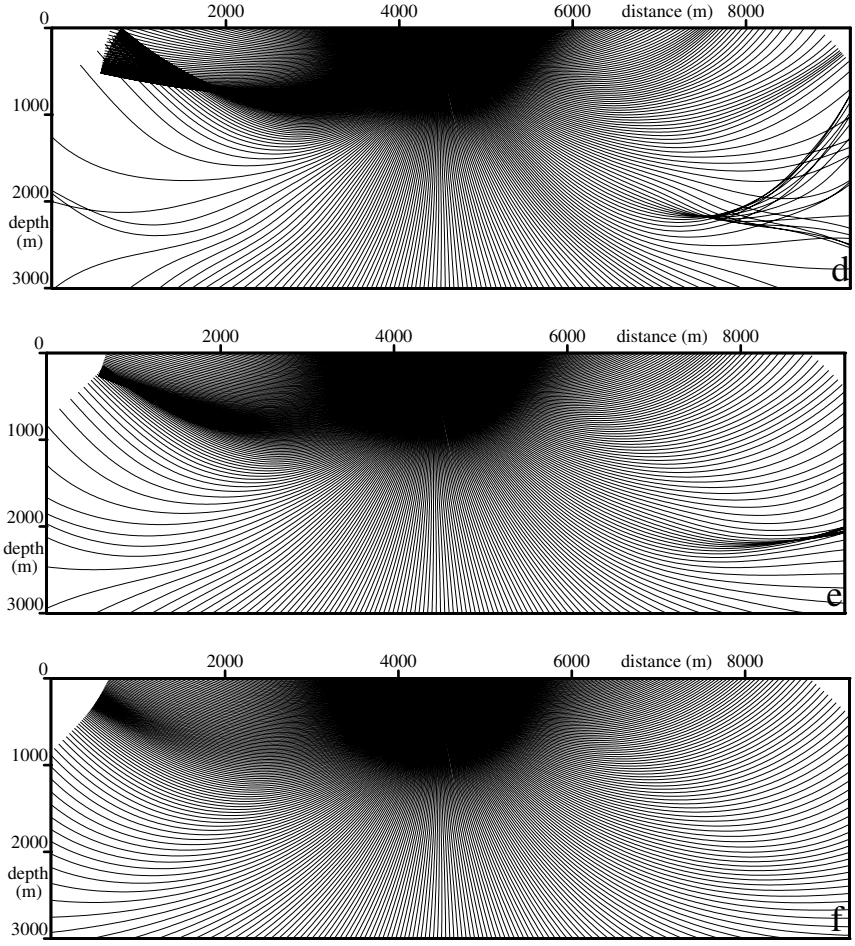

Figure 5d–f

$$\tfrac{1}{2} Mq^2 \ . \tag{20}$$

The quadratic term in the Taylor expansion of the complex-valued travel time of the Gaussian beam along the surface is

$$\tfrac{1}{2}(G^{\mathrm{R}} + \mathrm{i}G^{\mathrm{I}})(x - x_0)^2 \ , \tag{21}$$

where $(x - x_0)$ is the distance from the initial point of the central ray of the Gaussian beam to the respective point on the surface and G^{R} and G^{I} are real-valued parameters determining M uniquely (KLIMEŠ, 1984).

We have calculated the standard halfwidths of Gaussian beams in various smoothed models for various initial values of parameters G^{R} and G^{I}. Standard halfwidth a of a Gaussian beam of cross section

$$\exp\left(-\frac{q^2}{2a^2}\right) \ , \tag{22}$$

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

$$W = a\sqrt{2\pi f} \ . \tag{23}$$

The halfwidths of Gaussian beams calculated for the models with the B-spline grid of cells of 200×400 metres and with $s = 15{,}309$ m^2, $s = 30{,}619$ m^2, $s = 61{,}237$ m^2 and $s = 137{,}784$ m^2 are shown in Figure 6. These halfwidths have been calculated for the initial values of parameters $G^{\mathrm{R}} = 0$ and $G^{\mathrm{I}} = 0.250 \times 10^{-6}$. The models with lower values of s were excluded.

The gray-scale coded quantity W is displayed at the respective points along the central rays of the beams. The white colour corresponds to the Gaussian beam halfwidth of 0 metres for all frequencies. The black colour corresponds to the Gaussian beam halfwidths of 1010 metres and more for the frequency of 35 Hz, and of 1890 metres and more for the frequency of 10 Hz. Therefore, the black coloured regions of Figure 6 indicate too wide Gaussian beams for the frequencies under consideration.

We can see that the model with $s = 15{,}309$ m^2 is not suitable for Gaussian beams or packets. Especially if the position of the source is close to the middle of the profile, the Gaussian beams become wider too quickly. On the other hand, the models with $s = 61{,}237$ m^2 and $s = 137{,}784$ m^2 seem to be acceptable. Unfortunately, these models are smoothed to an extent which may jeopardize the migration. We hope that we will be able to use the model with $s = 30{,}619$ m^2 in the migration. We have studied the behaviour of Gaussian beams for various initial parameters G^{R} and G^{I}. We have realized that different initial values of these parameters are suitable for different positions of the source, or of the receiver in the migration. Consequently, we will try to develop a method to optimize the shapes of Gaussian beams or packets dependent on the position of the source, or of the receiver in the migration. This would allow the use of models not so smoothed.

We mention that even in models with a sufficiently small number of arrivals, the widths of Gaussian beams are at the verge of acceptability. This is due to the low frequencies under consideration.

Let us summarize that models with parameter s equal to or greater than $15,309$ m^2 seem to be suitable for ray tracing with the travel time of 2.3 seconds, see Figures 5 and 6. From this point of view, these models are sufficiently smooth. However, the low frequencies under consideration make the use of the Gaussian beam or packet method almost impossible. We probably should improve the applicability of the Gaussian packet method by using shapes of Gaussian packets optimized in dependence on the position of the source, or of the receiver in the migration.

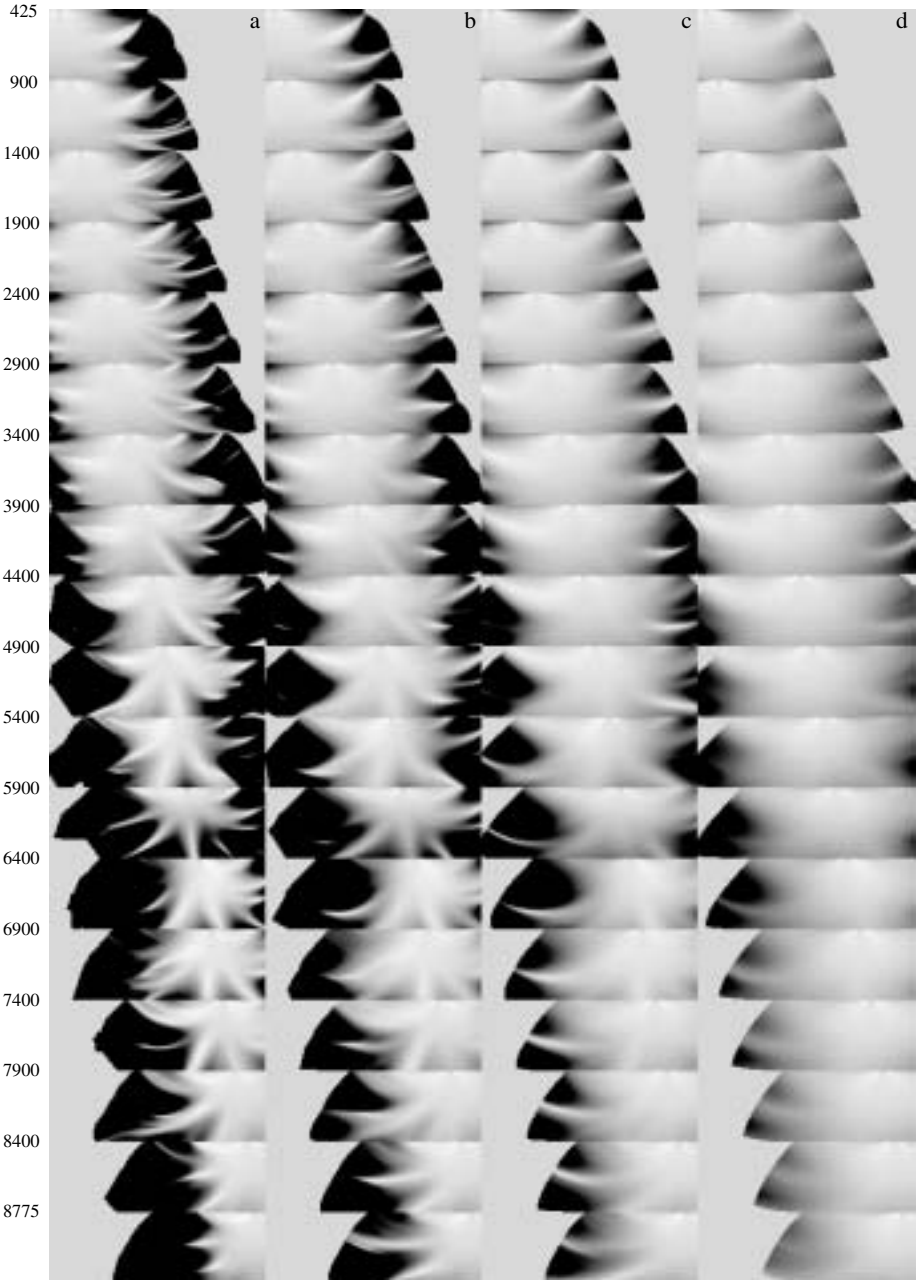

Figure 6
The halfwidths of Gaussian beams in smoothed models with the B-spline grid of cells of 200 × 400 metres. Columns correspond from left to right to the values of parameter s of (a) 15,309 m^2, (b) 30,619 m^2, (c) 61,237 m^2 and (d) 137,784 m^2. Rows correspond to various positions of the source (in metres).

Table 1

The relative RMS difference between the smoothed and the Marmousi model. Columns correspond to various B-spline grids, rows correspond to various values of parameter s

	100–400 m	200–230 m	200–400 m
0 m^2	8.3%	10.8%	11.1%
6,124 m^2	11.4%	11.9%	12.0%
15,309 m^2	13.2%	13.2%	13.3%
30,619 m^2	14.3%	14.3%	14.3%
61,237 m^2	15.1%	15.1%	15.1%
137,784 m^2	15.8%	15.8%	15.8%

9. Conclusions

The minimization of the relevant Sobolev norm of slowness is a powerful tool for preparing the optimum models for the asymptotic ray theory methods. As we have illustrated in numerical examples, it can be used for smoothing very complex models. However, the difference of slowness between the smoothed and the original model then increases rapidly. Also, the error of the travel time then increases.

We must keep in mind that there exists a natural relation between the complexity of the original model and the resulting difference between the sufficiently smoothed model and the original model. The more complex the original model, the more change it requires. Thus, the decision rests with the user, whether or not the model is too complex for smoothing. The required maximum error of travel time is then a key argument.

We have also demonstrated that even in models sufficiently smoothed for ray tracing, the Gaussian beams may still be too wide for the frequencies under consideration. In preparing a model for Gaussian beams or packets, we cannot judge solely from the number of arrivals and values of the ''average Lyapunov exponents'', whether the model is convenient. The widths of Gaussian beams or packets in relation to the frequency should be studied as well.

Acknowledgements

The author wishes to thank Luděk Klimeš for his kind guidance throughout the work on this paper, Professor V. Červený and Johana Brokešová for valuable comments and advice, and also Andreas Ehinger for providing the Marmousi model and data set. This research has been supported by the Ministry of Education of the Czech Republic within Research Project J13/98 113200004 and by the members of the consortium ''Seismic Waves in Complex 3-D Structures'' (see ''http:/seis.karlov.mff.cuni.cz/consort/main.htm'').

Appendix A

In evaluating a meaningful initial value of parameter s (see Section 6), we need to find some approximative relation between the Sobolev norm and the Lyapunov exponent.

According to KLIMEŠ (1999), the "average Lyapunov exponent" $\hat{\lambda}$ may be approximated by

$$\hat{\lambda} \approx \Lambda + \Delta\Phi \ , \tag{A.1}$$

where $\Delta\Phi$ is the decompensation for the low-velocity focusing zones. In 2-D, Λ is defined as

$$\Lambda = \left[\int v^{-1}\,\mathrm{d}^2x\right]^{-1} \int \sqrt{\mathrm{neg}(v_{,ij}\,e_i\,e_j)v^{-1}}\,\mathrm{d}^2x \ , \tag{A.2}$$

where $\mathrm{neg}(f) = 1/2(f - |f|)$ is the negative part of f, v is the velocity, $v_{,ij}$ is the second velocity derivative and $\mathbf{e}$ is a unit vector perpendicular to the ray.

We assume that the model is so smooth that the number of velocity oscillations, K_{osc}, along rays of length corresponding to $\tau_{\max}$ is small,

$$K_{\mathrm{osc}} = \frac{\tau_{\max}}{\tau_{\mathrm{osc}}} \ , \tag{A.3}$$

where τ_{osc} is the average wavelength of the velocity oscillations in the smoothed model, expressed in travel-time units. This assumption allows for the approximation

$$\Delta\Phi \approx -\frac{\ln 2}{\tau_{\mathrm{osc}}} = -\frac{K_{\mathrm{osc}} \ln 2}{\tau_{\max}} \ . \tag{A.4}$$

Let us now perform several approximations to express Λ in terms of the Sobolev norm of slowness u in the model without interfaces,

$$\Lambda \approx \left[\int u\,\mathrm{d}^2x\right]^{-1} \int \sqrt{\mathrm{pos}(u_{,ij}\,e_i\,e_j)u^{-1}}\,\mathrm{d}^2x \ , \tag{A.5}$$

where $\mathrm{pos}(f) = 1/2(f + |f|)$ is the positive part of f,

$$\Lambda \approx \frac{1}{2}\left[\int u^{\frac{3}{2}}\,\mathrm{d}^2x\right]^{-1} \int \sqrt{|u_{,ij}\,e_i\,e_j|}\,\mathrm{d}^2x \ , \tag{A.6}$$

and

$$\Lambda \approx \frac{1}{2}u_{\mathrm{A}}^{-\frac{3}{2}}\left\{\left[\int \mathrm{d}^2x\right]^{-1} \int (u_{,ij}\,e_i\,e_j)^2\,\mathrm{d}^2x\right\}^{\frac{1}{4}} \ , \tag{A.7}$$

where

$$u_{\mathrm{A}} = \left\{ \left[\int \mathrm{d}^2x \right]^{-1} \int u^{\frac{3}{2}} \mathrm{d}^2x \right\}^{\frac{2}{3}} \quad . \tag{A.8}$$

Finally, we arrive at

$$\Lambda \approx \frac{1}{2}\, u_{\mathrm{A}}^{-\frac{3}{2}} \left(\frac{3}{8}\right)^{\frac{1}{4}} \sqrt{\|u\|} \quad , \tag{A.9}$$

where $\|u\|$ is the Sobolev norm of slowness given by matrix $\mathbf{b}'$, see equation (17). This approximation may also be expressed as

$$\|u\| \approx \sqrt{\frac{8}{3}}\, u_{\mathrm{A}}^3 (2\Lambda)^2 \quad . \tag{A.10}$$

As we need to find some initial value of parameter s, we should estimate the respective value of the Sobolev norm $\|u\|_{\mathrm{init}}$. Since we have already derived an approximative relation between $\|u\|$ and Λ, see equation (A.10), we need to find the value of Λ_{init}. We have decided to keep the number of arrivals less than 10, see Section 5. With a view to equations (14), (A.1) and (A.4),

$$\Lambda \leq \frac{\ln 10 + K_{\mathrm{osc}} \ln 2}{\tau_{\max}} \quad . \tag{A.11}$$

Since we assume at least one shift of $-\ln 2$ for the source and one for the receiver, we assume $K_{\mathrm{osc}} = 2$. For $\tau_{\max} = 2.9$ s, we can put

$$\Lambda_{\mathrm{init}} \approx 1.3\,\mathrm{s}^{-1} \quad . \tag{A.12}$$

References

Abdullaev, S.S., *Chaos and Dynamics of Rays in Waveguide Media* (Gordon and Breach, Amsterdam 1993).

Addison, P.S., *Fractals and Chaos: An Illustrated Course* (IOP Publishing Ltd, Bristol and Philadelphia 1997).

Brac, J. and Nguyen, L.L., *Modeling geological objects with splines*. In *PSI 1990 Annual Report* (Institut Francais du Petrole, Rueil Malmaison, France 1990).

Bulant, P. (2001), *Sobolev Scalar Products in the Construction of Velocity Models — Application to Model Hess and to SEG/EAGE Salt Model*, Pure appl. geophys. *159*, 1487–1506.

Gold, N., Shapiro, S.A., Bojinsky, S., and Müller, T.M. (2000), *An Approach to Upscaling for Seismic Waves in Statistically Isotropic Heterogenous Elastic Media*, Geophysics *65*, 1837–1850.

Grubb, H.J. and Walden, A.T. (1995), *Smoothing Seismically Derived Velocities*, Geophys. Prosp. *43*, 1061–1082.

Keers, H., Dahlen, F.A., and Nolet, G. (1997), *Chaotic Ray Behaviour in Regional Seismology*, Geophys. J. Int. *131*, 361–380.

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), *Gaussian Packets in the Computation of Seismic Wavefields*, Geophys. J. Int. *99*, 421–433.
KLIMEŠ, L., *Lyapunov exponents for 2-D ray tracing without interfaces*. In *Seismic Waves in Complex 3-D Structures, Report 8* (Dep. Geophys., Charles Univ., Prague 1999) pp. 83–96.
KLIMEŠ, L. (2001), *Lyapunov Exponents for 2-D Ray Tracing without Interfaces*, Pure appl. geophys. *159*, 1465–1485.
KATOK, S.R., *The estimation from above for the topological entropy of a diffeomorphism*. In *Global Theory of Dynamical Systems. Lecture Notes in Mathematics, vol. 819* (eds. Netecki, Z. and Robinson, C.) (Springer, Berlin, Heidelberg, New York 1980) pp. 258–264.
LYAPUNOV, A.M., *Problème Général de la Stabilité du Mouvement, Annals of Mathematical Studies, vol. 17* (Princeton University Press 1949).
MCCAULEY, J.L., *Chaos, Dynamics and Fractals: An Algorithmic Approach to Deterministic Chaos* (Cambridge University Press, Cambridge 1993).
MÜLLER, T.M. and SHAPIRO, S.A. (2000), *Most Probable Seismic Pulses in Single Realizations of Two- and Three-dimensional Random Media*, Geophys. J. Int. *144*, 83–95.
OSELEDEC, V.I. (1968), *A Multiplicative Ergodic Theorem: Lyapunov Characteristic Numbers for Dynamical Systems*, Trans. Moscow Math. Soc. *19*, 179–210 in Russian, 197–231 English translation.
SMITH, K.B., BROWN, M.G., and TAPPERT, F.D. (1992), *Ray Chaos in Underwater Acoustics*, J. Acoust. Soc. Am. *91*, 1939–1949.
TAPPERT, F.D. and TANG, X. (1996), *Ray Chaos and Eigenrays*, J. Acoust. Soc. Am. *99*, 185–195.
TARANTOLA, A., *Inverse Problem Theory* (Elsevier, Amsterdam 1987).
VERSTEEG, R.J. (1993), *Sensitivity of Prestack Depth Migration to the Velocity Model*, Geophysics *58*, 873–882.
VERSTEEG, R.J., *Analysis of the Problem of the Velocity Model Determination for Seismic Imaging* (Ph.D. Thesis, University of Paris VII 1991).
VERSTEEG, R.J. and GRAU, G. (eds.), *The Marmousi Experience* (Eur. Assoc. Explor. Geophysicists, Zeist 1991).
WITTE, O., ROTH, M., and MÜLLER, G. (1996), *Ray Tracing in Random Media*, Geophys. J. Int. *124*, 159–169.

(Received October 18, 2000, revised April 9, 2001, accepted May 15, 2001)

Optimization of the shape of Gaussian beams

Karel Žáček
Department of Geophysics, Faculty of Mathematics and Physics, 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.

In: Seismic Waves in Complex 3-D Structures, Report 11 (Department of Geophysics, Charles University, Prague, 2001), 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

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

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

$$\frac{\mathrm{d}\mathbf{C}_{22}}{\mathrm{d}q_3} = \mathbf{\Pi}_2^\dagger \mathbf{G} \mathbf{\Pi}_2 \quad , \tag{31}$$

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),

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

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

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

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

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

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

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

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 $N(\sigma_Y^{GRID})^2$.

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 $4550\mathrm{ms}^{-1}$.

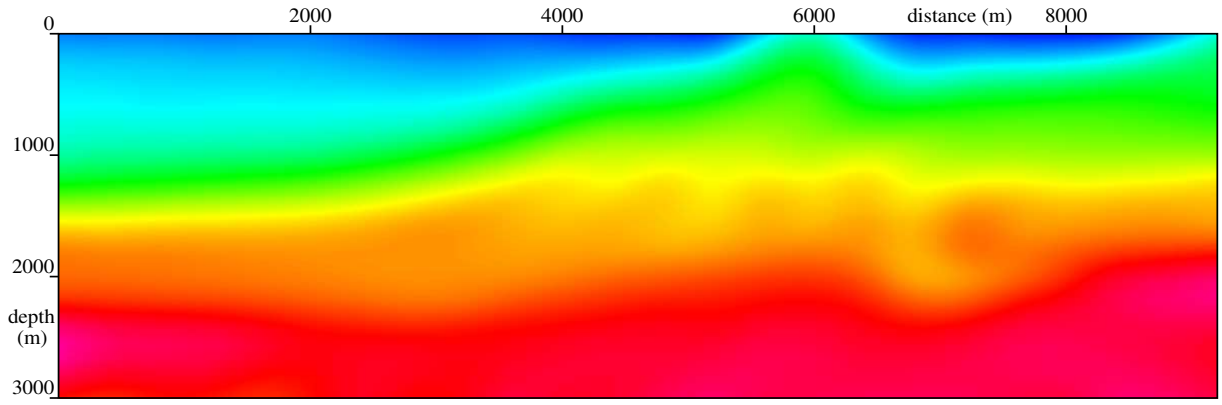

Figure 1. The smoothed Marmousi model.

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.

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

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

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

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

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 $R_0 = -0.26 \times 10^{-6}\mathrm{sm}^{-2}$ 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), $0.12 \times 10^{-6}\mathrm{sm}^{-2}$ 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.

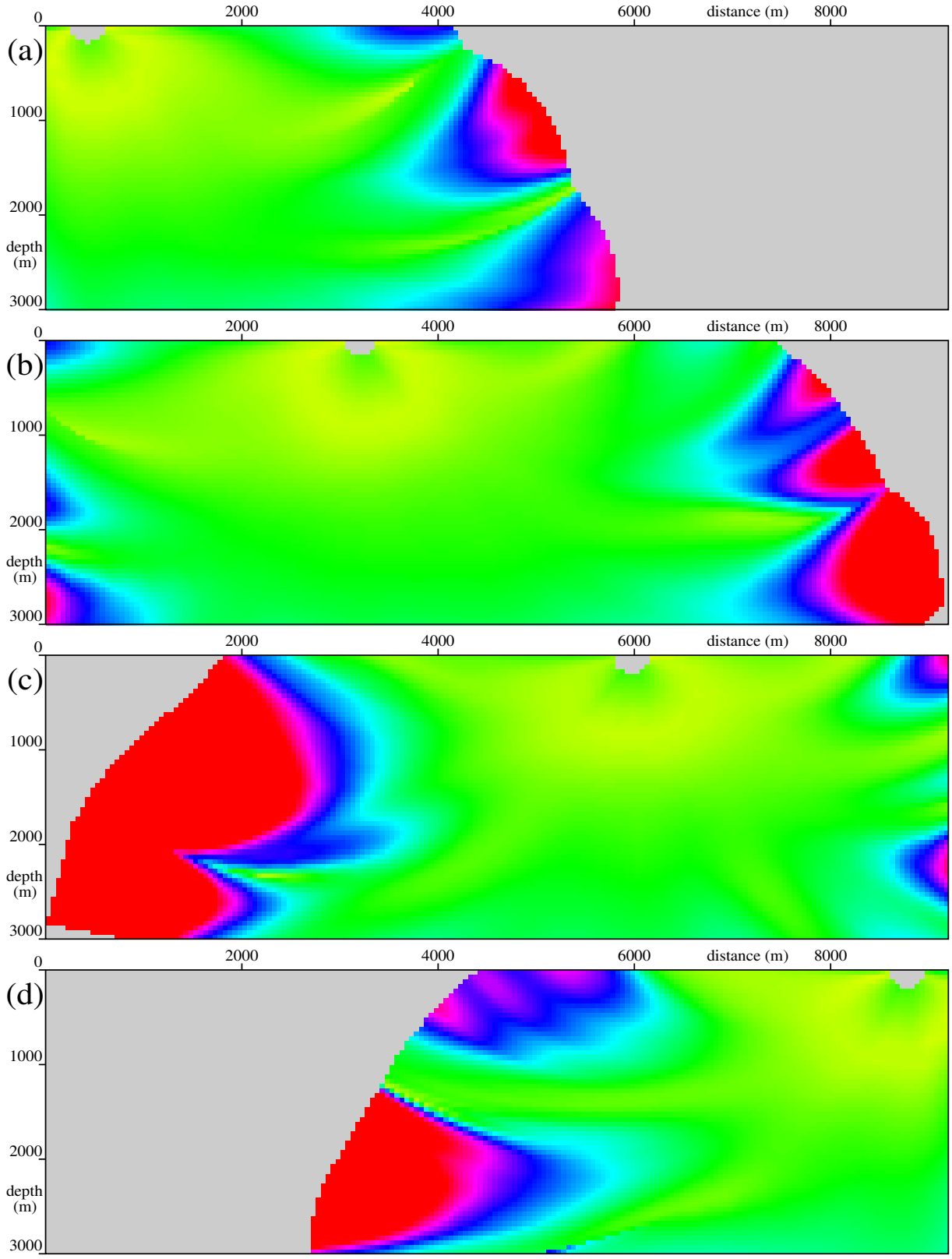

Figure 2. The halfwidths of GBs for the position of the source of (a) 425m, (b) 3200m, (c) 5975m and (d) 8775m. The *green* colour corresponds to the GB halfwidth of 202 and 378 metres for the frequencies of 35 Hz and 10 Hz, respectively. The *red* colour corresponds to the GB halfwidth of 1010 and more, 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.

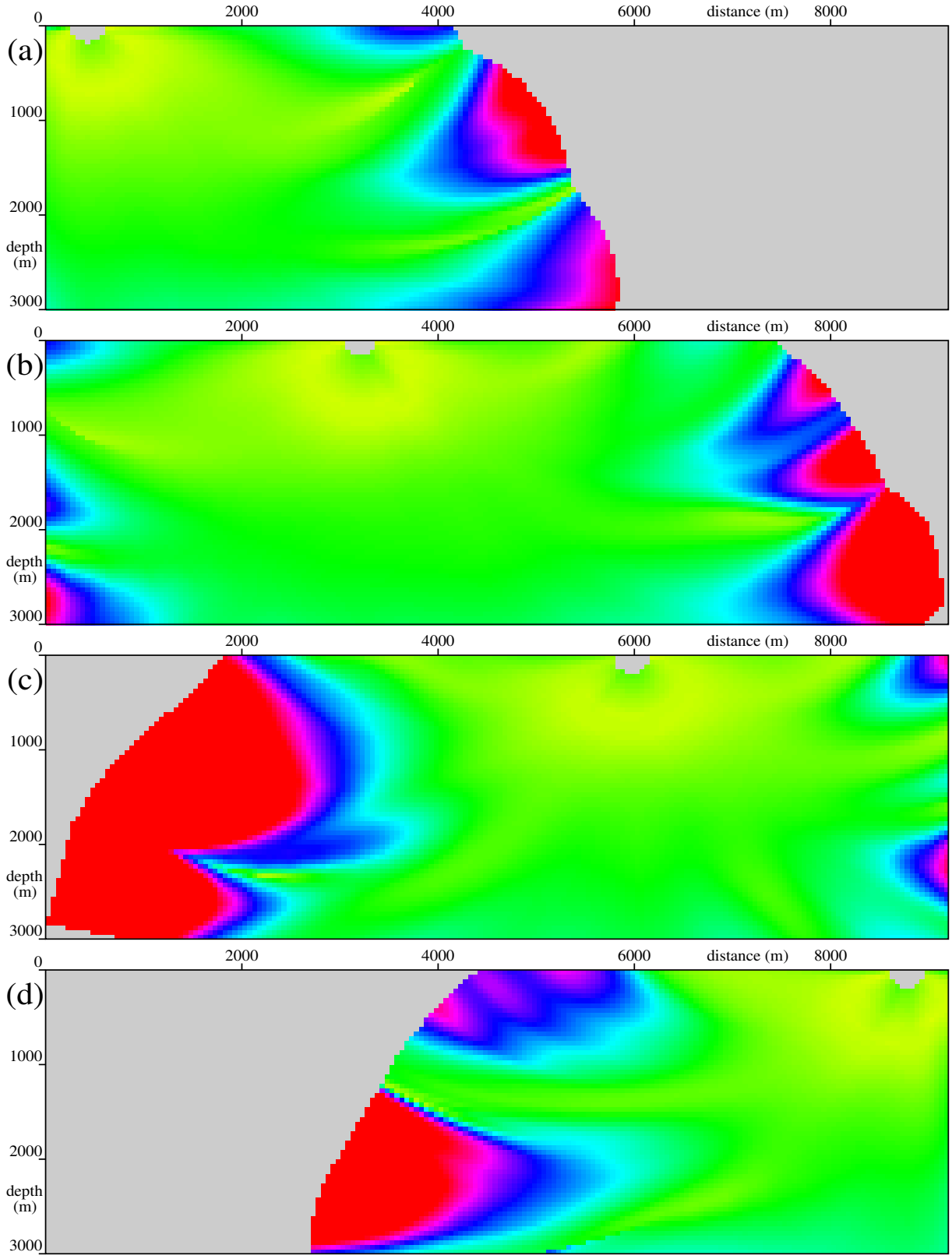

Figure 3. The halfwidths of GBs for the position of the source of (a) 425m, (b) 3200m, (c) 5975m and (d) 8775m. The *green* colour corresponds to the GB halfwidth of 202 and 378 metres for the frequencies of 35 Hz and 10 Hz, respectively. The *red* colour corresponds to the GB halfwidth of 1010 and more, 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.

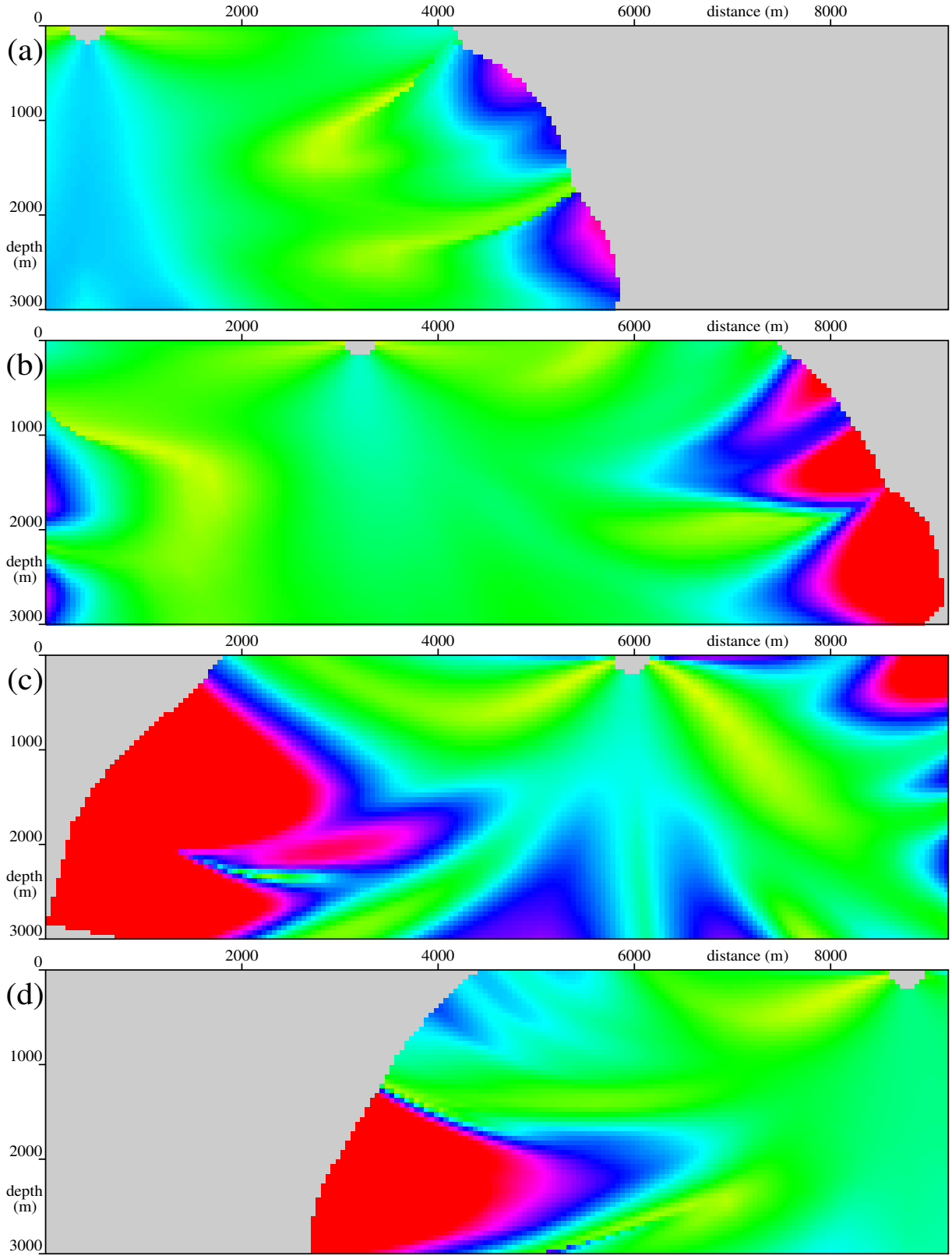

Figure 4. The halfwidths of GBs for the position of the source of (a) 425m, (b) 3200m, (c) 5975m and (d) 8775m. The *green* colour corresponds to the GB halfwidth of 202 and 378 metres for the frequencies of 35 Hz and 10 Hz, respectively. The *red* colour corresponds to the GB halfwidth of 1010 and more, 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.

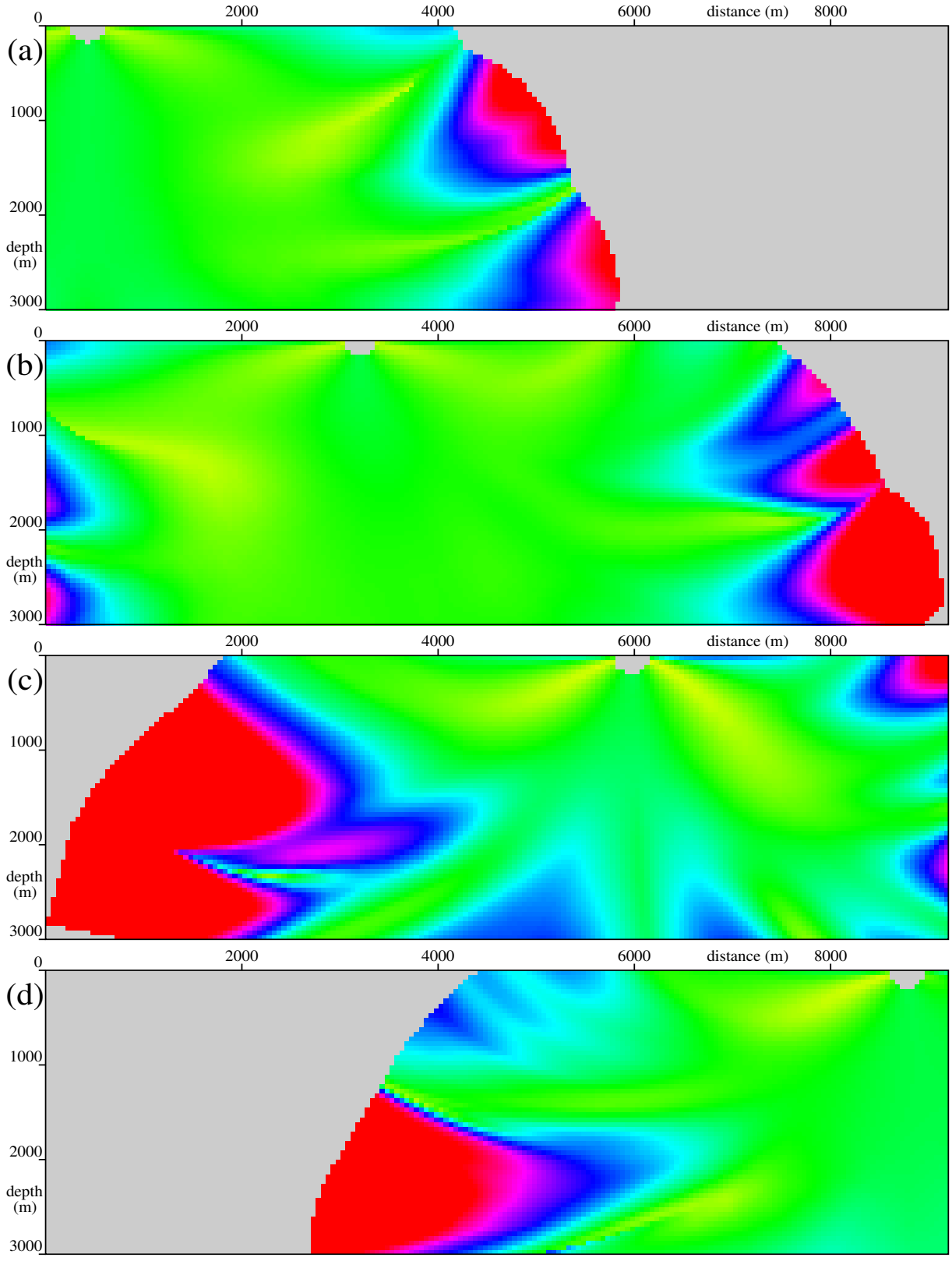

Figure 5. The halfwidths of GBs for the position of the source of (a) 425m, (b) 3200m, (c) 5975m and (d) 8775m. The *green* colour corresponds to the GB halfwidth of 202 and 378 metres for the frequencies of 35 Hz and 10 Hz, respectively. The *red* colour corresponds to the GB halfwidth of 1010 and more,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

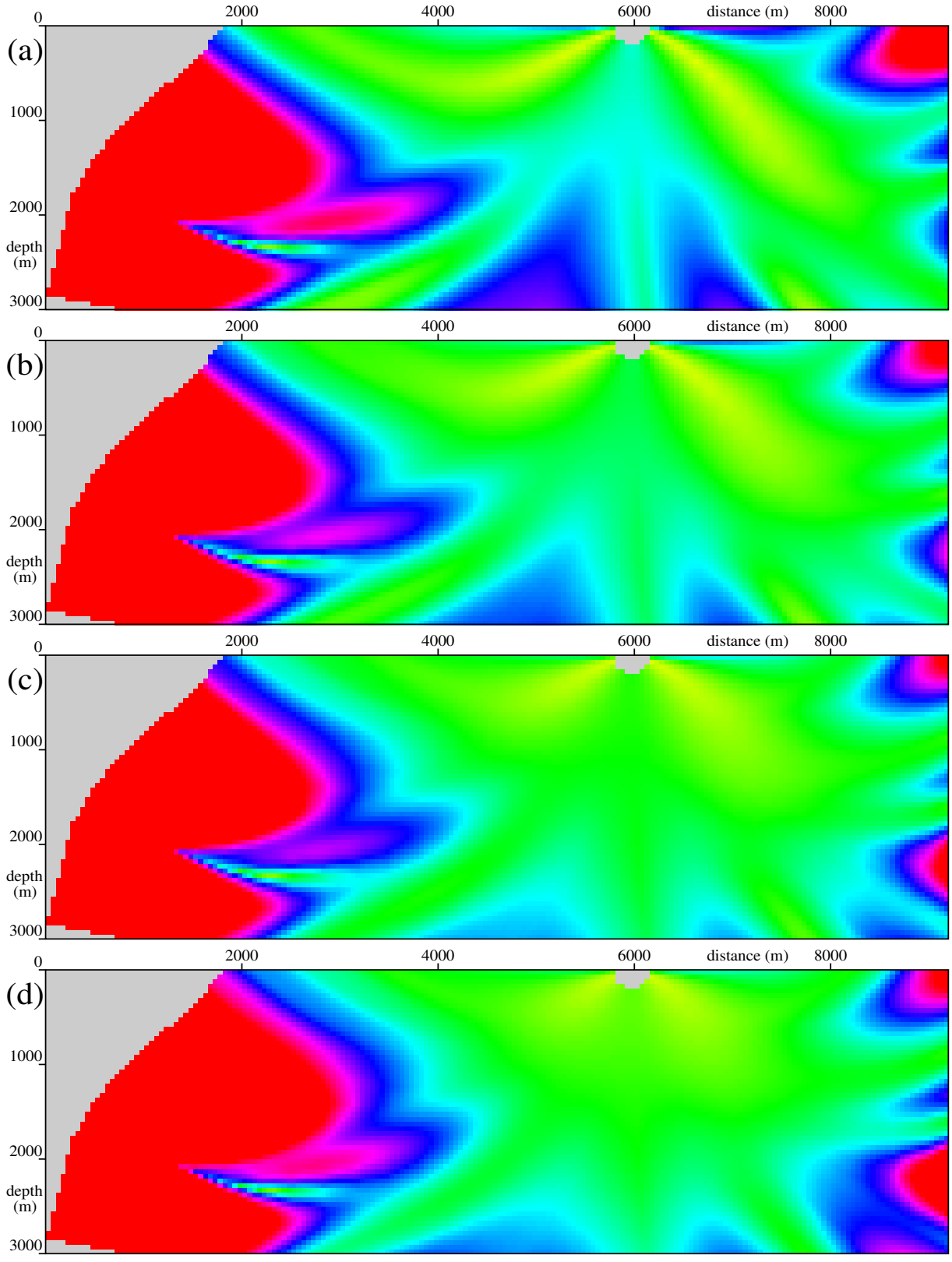

Figure 6. The halfwidths of GBs for the value of the initial parameter $R_0 = -0.26 \times 10^{-6} \mathrm{sm}^{-2}$ and for the value of the initial parameter Y_0 of (a) $0.29 \times 10^{-7} \mathrm{sm}^{-2}$, (b) $0.59 \times 10^{-7} \mathrm{sm}^{-2}$, (c) $0.12 \times 10^{-6} \mathrm{sm}^{-2}$ and (d) $0.24 \times 10^{-6} \mathrm{sm}^{-2}$. 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.

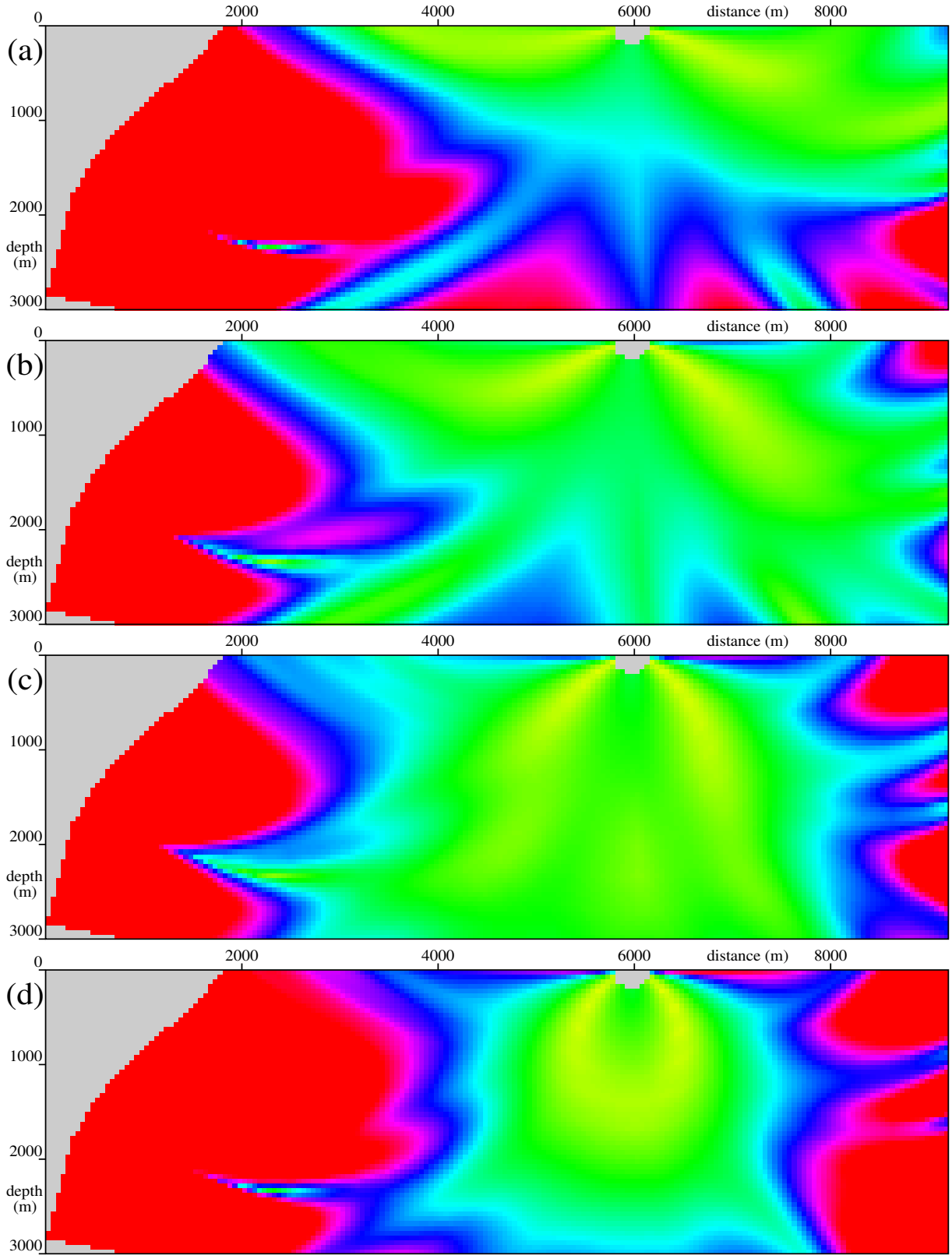

Figure 7. The halfwidths of GBs for the value of the initial parameter $Y_0 = 0.59 \times 10^{-7}\mathrm{sm}^{-2}$ and for the value of the initial parameter R_0 of (a) $-0.13 \times 10^{-6}\mathrm{sm}^{-2}$, (b) $-0.26 \times 10^{-6}\mathrm{sm}^{-2}$, (c) $-0.39 \times 10^{-6}\mathrm{sm}^{-2}$ and (d) $-0.52 \times 10^{-6}\mathrm{sm}^{-2}$. 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.

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.

Decomposition of the wave field into optimized Gaussian packets

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

Summary

The decomposition of the wave field into optimized Gaussian packets is of key importance in the Gaussian packet migration. We present the theoretical basics and a set of numerical examples of our method.

Key words

Gaussian packet, Gabor function, coherent-state transform, prestack depth migration, common-shot gather.

1 Introduction

Our long-term project is to explore the properties of a depth migration method based on Gaussian packets. Gaussian packets are waves whose envelopes at any given time are nearly Gaussian in space. They represent high-frequency asymptotic solutions of the elastodynamic equation, which are concentrated close to the central point of the packet (e.g., Babich & Ulin, 1981; Ralston, 1983; Klimeš, 1989a, 2004).

The main advantage of the Gaussian packet migration over the methods based on Gaussian beams is a direct relation between the regions in the common-shot gather and corresponding localized regions in the migrated section.

Like Gaussian beams, Gaussian packets become inaccurate solutions of the elastodynamic equation if they spread excessively as they propagate. This spreading depends on the complexity of the velocity model and on the initial shape of Gaussian packets. Thus, we should use sufficiently smooth velocity model and optimize the initial shape of the packets.

The Gaussian packet migration algorithm consists of four basic steps:
a) Preparation of a suitable smooth velocity model (Žáček, 2002).
b) Optimization of the shape of Gaussian packets (Klimeš, 1989b; Žáček, 2001a, 2001b).
c) Decomposition of the wave field into optimized Gaussian packets, which was already presented by Žáček (2003), and which we thoroughly describe in this paper. The word *optimized* implies that the shape of Gaussian packets, in the plane perpendicular to the central ray of the packet, depends not only on the frequency, but also on the coordinate of the intersection of the central ray of a Gaussian packet with the profile, on its arrival time, and on the component of the slowness vector along the profile.
d) Back-propagation of the wave field using Gaussian packets and application of the imaging functional (Žáček, 2004, 2005).

The algorithm of the decomposition is designed here especially for numerical testing on the 2-D Marmousi data set (Versteeg & Grau, 1991). That is why only 2-D case is dealt with in this paper, although the generalization to 3-D is straightforward.

In: Seismic Waves in Complex 3-D Structures, Report 15 (Department of Geophysics, Charles University, Prague, 2005), pp.17–27

2 Decomposition

We would like to approximate the wave field (common-shot gather) $f(x,t)$ in the form of

$$\widetilde{f}(x',t') = \int \mathrm{d}t_\mathrm{R} \int \mathrm{d}\omega \, \exp\left[-\mathrm{i}\omega(t'-t_\mathrm{R})\right] \int \mathrm{d}x_\mathrm{R} \int \mathrm{d}p \, \exp\left[\mathrm{i}\omega p(x'-x_\mathrm{R})\right]$$

$$\times \, W(x'-x_\mathrm{R}, t'-t_\mathrm{R}, x_\mathrm{R}, p, t_\mathrm{R}, \omega) \; F(x_\mathrm{R}, p, t_\mathrm{R}, \omega) \quad , \tag{1}$$

where x_R is the coordinate of the intersection of the central ray of a Gaussian packet with the profile, t_R denotes corresponding arrival time of a Gaussian packet, p is the component of the slowness vector a Gaussian packet along the profile, and ω is the positive circular frequency of a Gaussian packet. The complex-valued function W determines the shape of the envelope of a Gaussian packet, and the complex-valued function $F(x_\mathrm{R}, p, t_\mathrm{R}, \omega)$ specifies the amplitude of a Gaussian packet. Hereafter, the limits of integration with respect to x_R, t_R, x, t and p are $-\infty$ and $+\infty$, and the limits of integration with respect to ω are 0 and $+\infty$.

Let us express the amplitude of a Gaussian packet in the form of an integral transform similar to the forward 2-D *coherent-state transform*,

$$F(x_\mathrm{R}, p, t_\mathrm{R}, \omega) = \frac{\omega}{2\pi^2} \int \mathrm{d}x \, \exp\left[-\mathrm{i}\omega p(x-x_\mathrm{R})\right] \int \mathrm{d}t \, \exp\left[\mathrm{i}\omega(t-t_\mathrm{R})\right]$$

$$\times \, w(x_\mathrm{R}-x, t_\mathrm{R}-t, x_\mathrm{R}, p, t_\mathrm{R}, \omega) \; f(x,t) \quad , \tag{2}$$

where w denotes a complex-valued analyzing function. Then, equation (1) can be expressed as

$$\widetilde{f}(x',t') = \int \mathrm{d}t \int \mathrm{d}\omega \, \exp\left[\mathrm{i}\omega(t-t')\right] \int \mathrm{d}x \int \mathrm{d}p \, \exp\left[-\mathrm{i}\omega p(x-x')\right]$$

$$\times \, \frac{\omega}{2\pi^2} \; f(x,t) \; I(x'-x, t'-t, p, \omega) \quad , \tag{3}$$

where

$$I(x'-x, t'-t, p, \omega)$$

$$= \int \mathrm{d}t_\mathrm{R} \int \mathrm{d}x_\mathrm{R} \, W(x'-x_\mathrm{R}, t'-t_\mathrm{R}, x_\mathrm{R}, p, t_\mathrm{R}, \omega) \; w(x_\mathrm{R}-x, t_\mathrm{R}-t, x_\mathrm{R}, p, t_\mathrm{R}, \omega) \quad . \tag{4}$$

Considering the convolution (4) independent of p, integration with respect to p in (3) yields a Dirac distribution $\delta(x'-x)$,

$$\widetilde{f}(x',t') = \int \mathrm{d}t \int \mathrm{d}\omega \, \exp\left[\mathrm{i}\omega(t-t')\right] \int \mathrm{d}x \, \delta(x'-x) \, \frac{1}{\pi} \, f(x,t) \, I(x'-x, t'-t, 0, \omega) \quad , \tag{5}$$

which can be integrated with respect to x,

$$\widetilde{f}(x',t') = \int \mathrm{d}t \int \mathrm{d}\omega \, \exp\left[\mathrm{i}\omega(t-t')\right] \; \frac{1}{\pi} \; f(x',t) \; I(0, t'-t, 0, \omega) \quad . \tag{6}$$

Considering $I(0, t'-t, 0, \omega)$ independent of ω, integration with respect to ω in (6) yields a complex-valued analytic Dirac distribution $\delta(t'-t) - \frac{\mathrm{i}}{\pi(t'-t)}$. Moreover, let us consider real-valued $I(0, t'-t, 0, \omega)$, which leads to

$$\mathrm{Re}\left(\widetilde{f}(x',t')\right) = \int \mathrm{d}t \; \delta(t'-t) \; f(x',t) \; I(0, t'-t, 0, \omega_0) \quad , \tag{7}$$

which can be integrated with respect to t,

$$\mathrm{Re}\left(\widetilde{f}(x',t')\right) = f(x',t')\ I(0,0,0,\omega_0) \quad . \tag{8}$$

For

$$I(0,0,0,\omega_0) = 1 \quad , \tag{9}$$

the transformation would be exact,

$$\mathrm{Re}\left(\widetilde{f}(x,t)\right) = f(x,t) \quad . \tag{10}$$

The envelope of a trace of a Gaussian packet along the profile may be expressed as

$$W(x'-x_\mathrm{R}, t'-t_\mathrm{R}, x_\mathrm{R}, p, t_\mathrm{R}, \omega) = \exp\left[\mathrm{i}\omega\frac{1}{2}(\mathbf{y}'-\mathbf{y}_\mathrm{R})^\mathrm{T}\mathbf{K}(\mathbf{y}'-\mathbf{y}_\mathrm{R})\right] \quad , \tag{11}$$

where

$$\mathbf{y}' = \begin{pmatrix} x' \\ t' \end{pmatrix} \quad , \tag{12}$$

$$\mathbf{y}_\mathrm{R} = \begin{pmatrix} x_\mathrm{R} \\ t_\mathrm{R} \end{pmatrix} \tag{13}$$

and

$$\mathbf{K} = N\ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + N_{44}\ \begin{pmatrix} p^2 & -p \\ -p & 1 \end{pmatrix} \quad . \tag{14}$$

Please note that parameters N and N_{44} correspond to N^0 and N^0_{44}, respectively, in Žáček (2005, eq. 65).

Let us write the analyzing function w in an analogous form,

$$w(x'-x_\mathrm{R}, t'-t_\mathrm{R}, x_\mathrm{R}, p, t_\mathrm{R}, \omega) = \widetilde{a}\ \exp\left[\mathrm{i}\omega\frac{1}{2}(\mathbf{y}-\mathbf{y}_\mathrm{R})^\mathrm{T}\widetilde{\mathbf{K}}(\mathbf{y}-\mathbf{y}_\mathrm{R})\right] \quad , \tag{15}$$

where

$$\mathbf{y} = \begin{pmatrix} x \\ t \end{pmatrix} \tag{16}$$

and

$$\widetilde{\mathbf{K}} = \widetilde{N}\ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \widetilde{N}_{44}\ \begin{pmatrix} \widetilde{p}^2 & -\widetilde{p} \\ -\widetilde{p} & 1 \end{pmatrix} \quad . \tag{17}$$

The meaning and choice of parameters $\widetilde{a}$, $\widetilde{p}$, $\widetilde{N}$ and $\widetilde{N}_{44}$ is discussed below.

Neglecting the dependence of parameters N, N_{44}, $\widetilde{a}$, $\widetilde{p}$, $\widetilde{N}$ and $\widetilde{N}_{44}$ on x_R and t_R, we may calculate the convolution (4),

$$I(x'-x, t'-t, p, \omega)$$
$$= \widetilde{a}\int \mathrm{d}t_\mathrm{R} \int \mathrm{d}x_\mathrm{R}\ \exp\left[\mathrm{i}\omega\frac{1}{2}\left((\mathbf{y}'-\mathbf{y}_\mathrm{R})^\mathrm{T}\mathbf{K}(\mathbf{y}'-\mathbf{y}_\mathrm{R}) + (\mathbf{y}-\mathbf{y}_\mathrm{R})^\mathrm{T}\widetilde{\mathbf{K}}(\mathbf{y}-\mathbf{y}_\mathrm{R})\right)\right] \quad , \tag{18}$$

which leads to

$$I(x'-x,t'-t,p,\omega)$$

$$= \frac{2\pi\widetilde{a}}{\omega\sqrt{\det[-\mathrm{i}(\mathbf{K}+\widetilde{\mathbf{K}})]}} \exp\left[\mathrm{i}\omega\frac{1}{2}\left((\mathbf{y}'-\mathbf{y})^{\mathrm{T}}(\mathbf{K}^{-1}+\widetilde{\mathbf{K}}^{-1})^{-1}(\mathbf{y}'-\mathbf{y})\right)\right] \quad . \quad (19)$$

Inverse matrices to (14) and (17) read

$$\mathbf{K}^{-1} = \frac{1}{N_{44}}\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix} + \frac{1}{N}\begin{pmatrix}1 & p\\ p & p^2\end{pmatrix} \quad (20)$$

and

$$\widetilde{\mathbf{K}}^{-1} = \frac{1}{\widetilde{N}_{44}}\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix} + \frac{1}{\widetilde{N}}\begin{pmatrix}1 & \widetilde{p}\\ \widetilde{p} & \widetilde{p}^2\end{pmatrix} \quad . \quad (21)$$

In order to proceed from (3) to (5) and (6), we require $I(x'-x,t'-t,p,\omega)$ to be independent of p. Therefore, the sum $\mathbf{K}^{-1}+\widetilde{\mathbf{K}}^{-1}$ of matrices (20) and (21) in (19) should be independent of p. To proceed from (6) to (7), we need $I(0,t'-t,0,\omega)$ to be independent of ω. This means that the rightmost diagonal component of matrix $(\mathbf{K}^{-1}+\widetilde{\mathbf{K}}^{-1})^{-1}$ in (19) should be proportional to ω^{-1}.

To meet these requirements, we choose the sum of matrices (20) and (21) in the form of

$$\mathbf{K}^{-1}+\widetilde{\mathbf{K}}^{-1} = \begin{pmatrix}\frac{2}{N_0} & 0\\ 0 & \frac{\omega}{K_0}+\frac{\omega}{\widetilde{K}_0}\end{pmatrix} \quad , \quad (22)$$

where N_0, K_0 and $\widetilde{K}_0$ are constants. We introduce two generally different constants K_0 and $\widetilde{K}_0$, because we would like to choose $N_{44}=N_{44}(K_0)$ and $\widetilde{N}_{44}=\widetilde{N}_{44}(\widetilde{K}_0)$. Inserting expressions (20) and (21) into equation (22), we obtain three conditions,

$$\frac{1}{N}+\frac{1}{\widetilde{N}} = \frac{2}{N_0} \quad , \quad (23)$$

$$\frac{p}{N}+\frac{\widetilde{p}}{\widetilde{N}} = 0 \quad (24)$$

and

$$\frac{p^2}{N}+\frac{1}{N_{44}}+\frac{\widetilde{p}^2}{\widetilde{N}}+\frac{1}{\widetilde{N}_{44}} = \frac{\omega}{K_0}+\frac{\omega}{\widetilde{K}_0} \quad , \quad (25)$$

where N_0, K_0 and $\widetilde{K}_0$ are free complex-valued constants that must be selected.

In 2-D, the initial shape of a Gaussian packet is defined by two complex-valued parameters — N and K_0. The imaginary part of N determines the Gaussian packet width along the profile and the real part of N defines the curvature of the phase-front. We choose parameter K_0, which describes the shape of a Gaussian packet along the central ray, and which is connected with parameter N_{44} by equation (25), pure imaginary and uniform for all Gaussian packets.

In order to minimize the spreading of Gaussian packets, parameter N should be determined by the optimization of the shape of corresponding Gaussian beams (Klimeš, 1989b; Žáček, 2001a, 2001b). Consequently, we should consider *optimized* Gaussian packets,

$$N = N(x_{\mathrm{R}},t_{\mathrm{R}},p) \quad . \quad (26)$$

In this place, we must point out that the transformation is exact,

$$\widetilde{f}(x,t) = f(x,t) \quad , \tag{27}$$

only for N independent of x_{R} and t_{R}, see equations (18) and (19). Nevertheless, for a reasonably smooth distribution of N with respect to x_{R} and t_{R}, we could achieve satisfactory results. Therefore, we have developed a procedure, which allows us to smooth iteratively the distribution of parameter N on a *Hamiltonian hypersurface* in the phase-space (Žáček, 2001a, 2001b).

Using equations (23) and (24), we may write

$$\widetilde{N} = \frac{NN_0}{2N - N_0} \tag{28}$$

and

$$\widetilde{p} = -\frac{\widetilde{N}}{N}p \quad . \tag{29}$$

Please note that parameter p is real-valued, but $\widetilde{p}$ is generally complex-valued.

Equation (25) may be rearranged in several ways to determine parameters N_{44} and $\widetilde{N}_{44}$. For instance, we can get two symmetrical expresions,

$$\widetilde{N}_{44} = \frac{\widetilde{N}\widetilde{K}_0}{\omega\widetilde{N} - \widetilde{p}^2\widetilde{K}_0} \tag{30}$$

and

$$N_{44} = \frac{NK_0}{\omega N - p^2 K_0} \quad . \tag{31}$$

Then, pure imaginary constants $\widetilde{K}_0$ and K_0 must obey

$$\mathrm{Im}(\widetilde{K}_0) < \mathrm{Im}\left(\frac{\omega\widetilde{N}}{\widetilde{p}^2}\right) \tag{32}$$

and

$$\mathrm{Im}(K_0) < \mathrm{Im}\left(\frac{\omega N}{p^2}\right) \quad . \tag{33}$$

Another approach would allow the use of the fast Fourier transform (FFT) in the decomposition (but not in the re-composition) of the wave field,

$$\widetilde{N}_{44} = \frac{\widetilde{K}_0}{\omega} \quad , \tag{34}$$

$$N_{44} = \left(\frac{\omega}{K_0} - \frac{p^2}{N} - \frac{\widetilde{p}^2}{\widetilde{N}}\right)^{-1} \tag{35}$$

and

$$\mathrm{Im}(K_0) < \mathrm{Im}\left(\frac{\omega\,(2N - N_0)}{p^2}\right) \quad . \tag{36}$$

Equation (9) requires

$$\widetilde{a} = \frac{\omega}{2\pi}\sqrt{\det[-\mathrm{i}(\mathbf{K} + \widetilde{\mathbf{K}})]} \quad , \tag{37}$$

where matrix $\mathbf{K}$ is given by (14) and matrix $\widetilde{\mathbf{K}}$ by (17). Equation (37) with (14) and (17) can be converted using equations (29) and (25) into

$$\widetilde{a} = \frac{\omega}{2\pi}\sqrt{-\mathrm{i}(N+\widetilde{N})}\sqrt{-\mathrm{i}N_{44}\widetilde{N}_{44}\left(\frac{\omega}{K_0}+\frac{\omega}{\widetilde{K}_0}\right)} \quad , \tag{38}$$

where the square roots should be taken with positive real parts.

3 Discretization

A proper choice of the discretization of integral (1) is not only of key importance in the decomposition of the wave field into Gaussian packets, but it also affects the accuracy and efficiency of the Gaussian packet migration.

Let us consider a 2-D integral

$$\int \mathrm{d}\xi_1 \int \mathrm{d}\xi_2 \, \exp\left(\mathrm{i}\,\xi_J D_{JK} \xi_K\right) \tag{39}$$

discretized on a regular rectangular grid with intervals $\Delta\xi_1$ and $\Delta\xi_2$. Sufficient conditions for the discretization were derived by Klimeš (1986),

$$\Delta\xi_1 \le \kappa\sqrt{-[\mathrm{Im}(\mathbf{D}^{-1})]_{11}} \tag{40}$$

and

$$\Delta\xi_2 \le \kappa\sqrt{-[\mathrm{Im}(\mathbf{D}^{-1})]_{22}} \quad . \tag{41}$$

Matrix $\mathbf{D}$ must satisfy

$$\left|[\mathrm{Im}(\mathbf{D}^{-1})]_{12}\right| \le \frac{1}{2}\sqrt{[\mathrm{Im}(\mathbf{D}^{-1})]_{11}[\mathrm{Im}(\mathbf{D}^{-1})]_{22}} \quad . \tag{42}$$

The error of discretization depends on the properties of the wave field. For a simple wave field with linear amplitudes and quadratic arrival times, the relative maximum discretization error is about

$$\delta_{\mathrm{MAX}} = 4\,\exp(-\pi^2\kappa^{-2}) \quad . \tag{43}$$

Then, the maximum relative error for $\kappa^2 = \frac{\pi}{2}$ would be about 0.7 %. Unfortunately, the error may raise rapidly for a more complex wave field. In the worst case, the maximum relative error could reach up to 21 % for $\kappa^2 = \frac{\pi}{2}$, see Table 3.3 in Daubechies (1992). Nevertheless, we believe that a reasonable value of κ for a realistic wave field would be close to $\kappa^2 = \frac{\pi}{2}$.

The error of discretization of integral (18) with respect to x_{R} and t_{R} is controlled by matrix

$$\mathbf{D} = \frac{1}{2}\omega(\mathbf{K} + \widetilde{\mathbf{K}}) \quad . \tag{44}$$

According to conditions (40) and (41), maximum step in x_{R} for replacement of integral (18) by its discretized counterpart is

$$\Delta x_{\mathrm{R}} = \kappa\sqrt{-\frac{2}{\omega}\left[\mathrm{Im}\left((\mathbf{K}+\widetilde{\mathbf{K}})^{-1}\right)\right]_{11}} \quad , \tag{45}$$

maximum step in t_{R} is

$$\Delta t_{\mathrm{R}} = \kappa\sqrt{-\frac{2}{\omega}\left[\mathrm{Im}\left((\mathbf{K}+\widetilde{\mathbf{K}})^{-1}\right)\right]_{22}} \quad , \tag{46}$$

and we require

$$\left|\left[\mathrm{Im}\left((\mathbf{K}+\widetilde{\mathbf{K}})^{-1}\right)\right]_{12}\right| \leq \frac{1}{2}\sqrt{\left[\mathrm{Im}\left((\mathbf{K}+\widetilde{\mathbf{K}})^{-1}\right)\right]_{11}\left[\mathrm{Im}\left((\mathbf{K}+\widetilde{\mathbf{K}})^{-1}\right)\right]_{22}} \quad . \tag{47}$$

Equation (3) can be understood as a convolution of functions

$$J(x,t,p,\omega) = \frac{\omega}{2\pi^2} I(x,t,p,\omega)\ \exp(\mathrm{i}\omega p x)\ \exp(-\mathrm{i}\omega t) \tag{48}$$

and $f(x,t)$ with respect to x and t, integrated over p and ω. Fourier transform $\widetilde{f}(k',\omega')$ of $\widetilde{f}(x',t')$ is then the product of Fourier transforms $J(k',\omega',p,\omega)$ and $f(k',\omega')$, integrated over p and ω. Inserting equation (19) into (48) and using the Fourier transform, we obtain $J(k',\omega',p,\omega)$ in the form similar to (39) with $\xi_1 = k'$ and $\xi_2 = \omega'$. The error of discretization of the integral of $J(k',\omega',p,\omega)$ with respect to $k = \omega p$ and ω is then controlled by matrix

$$\mathbf{D} = -\frac{1}{2}\omega^{-1}(\mathbf{K}^{-1} + \widetilde{\mathbf{K}}^{-1}) \quad . \tag{49}$$

According to conditions (40) and (41), maximum step in p is

$$\Delta p = \kappa\sqrt{\frac{2}{\omega}\left[\mathrm{Im}\left((\mathbf{K}^{-1}+\widetilde{\mathbf{K}}^{-1})^{-1}\right)\right]_{11}} \quad , \tag{50}$$

maximum step in ω is

$$\Delta\omega = \kappa\sqrt{2\omega\left[\mathrm{Im}\left((\mathbf{K}^{-1}+\widetilde{\mathbf{K}}^{-1})^{-1}\right)\right]_{22}} \quad , \tag{51}$$

and we require

$$\left|\left[\mathrm{Im}\left((\mathbf{K}^{-1}+\widetilde{\mathbf{K}}^{-1})^{-1}\right)\right]_{12}\right| \leq \frac{1}{2}\sqrt{\left[\mathrm{Im}\left((\mathbf{K}^{-1}+\widetilde{\mathbf{K}}^{-1})^{-1}\right)\right]_{11}\left[\mathrm{Im}\left((\mathbf{K}^{-1}+\widetilde{\mathbf{K}}^{-1})^{-1}\right)\right]_{22}} \quad . \tag{52}$$

Using conditions (23), (24) and (25), we can simplify expressions (50) and (51) into

$$\Delta p = \kappa\sqrt{\frac{\mathrm{Im}(N_0)}{\omega}} \tag{53}$$

and

$$\Delta\omega = \kappa\sqrt{2\ \mathrm{Im}\left(\frac{K_0\widetilde{K}_0}{K_0+\widetilde{K}_0}\right)} \quad . \tag{54}$$

Presented discretization conditions are valid for a general smooth distribution of parameter N on a *Hamiltonian hypersurface* in the phase-space, see relation (26). However, to simplify the whole migration procedure, it might be useful to smooth the optimized distribution of parameter N to a constant value as described by Žáček (2001a, 2001b). Considering *uniform* Gaussian packets with $N = \widetilde{N} = N_0$, $K_0 = \widetilde{K}_0$, $\widetilde{N}_{44}$ given by relation (30) and N_{44} determined by (31), maximum discretization steps according to (45), (46), (53) and (54) read

$$\Delta x_{\mathrm{R}} = \frac{\kappa}{\omega}\sqrt{-\mathrm{Im}\left[(\omega N_0 - p^2 K_0)\, N_0^{-2}\right]} \quad , \tag{55}$$

$$\Delta t_{\mathrm{R}} = \kappa \sqrt{-\mathrm{Im}\left[\left(\omega N_0 - p^2 K_0\right) N_0^{-1} K_0^{-1}\right]} \quad , \tag{56}$$

$$\Delta p = \kappa \sqrt{\frac{\mathrm{Im}\left(N_0\right)}{\omega}} \tag{57}$$

and

$$\Delta \omega = \kappa \sqrt{\mathrm{Im}\left(K_0\right)} \quad . \tag{58}$$

Naturally, *uniform* Gaussian packets bring about more rapid spreading, which could jeopardize the accuracy of the migration.

4 Numerical examples

We tested our method on a common-shot gather belonging to the Marmousi data set (Versteeg & Grau, 1991), see Figure 1. We decomposed the gather $f(x,t)$ using equation (2) into individual Gabor functions, see Figure 2. The complex-valued amplitude of a Gaussian packet equals the amplitude of the corresponding Gabor function.

In order to examine the accuracy of our method, we composed the approximated common-shot gathers $\widetilde{f}(x,t)$ according to equation (1), see Figure 3. The differences between the original gather and the approximated gathers are shown in Figure 4. We calculated all the numerical examples with $\kappa^2 = \frac{\pi}{2}$.

In the the first example, we present the decomposition of the wave field into *uniform* Gaussian packets with $\mathrm{Re}(N_0) = -0.25 \times 10^{-6}$ sm^{-2} and $\mathrm{Im}(N_0) = 0.25 \times 10^{-6}$ sm^{-2}. See Figure 3a for the corresponding approximated common-shot gather. Figure 4a indicates that the transformation is almost exact.

The next example demostrates the decomposition of the wave field into Gaussian packets with varying parameter N. In this case, $\mathrm{Re}(N) = -0.25 \times 10^{-6}$ sm^{-2} and $\mathrm{Im}(N)$ ranges from 0.125×10^{-6} sm^{-2} to 0.5×10^{-6} sm^{-2}. The approximated gather is in Figure 3b. According to Figure 4b, the differences between the original and the approximated wave field can be spotted not only at the edges of the gather. This is caused by the dependency of parameter N on x_{R} and t_{R}.

The last example is the decomposition of the wave field with $\mathrm{Re}(N) = -0.25 \times 10^{-6}$ sm^{-2} and $\mathrm{Im}(N)$ varying from 0.75×10^{-7} sm^{-2} to 0.1×10^{-5} sm^{-2}. The approximated gather is displayed in Figure 3b. Although we can note several differences between the original and the approximated gather in Figure 4c, the result seems to be satisfactory.

The only serious problem of our method consists in the computational cost. Unfortunately, we cannot use 2-D FFT. The mathematical formulation of our task allows only a series of 1-D FFTs in the decomposition, but not in the re-composition of the wave field. In order to make the decomposition of the wave field more efficient, we are constantly trying to refine the algorithm and speed up the computations.

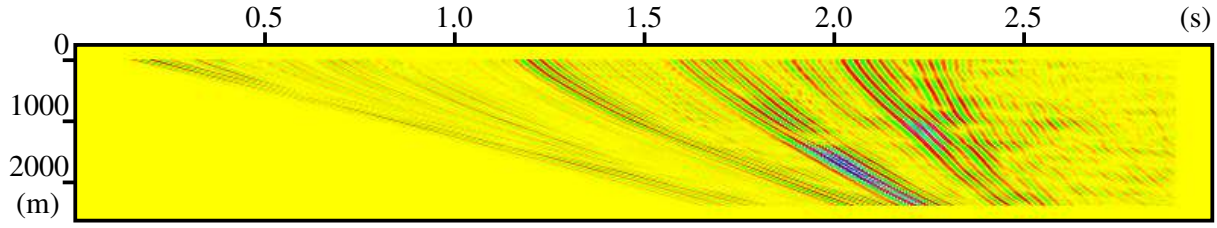

Figure 1. Original common-shot gather $f(x,t)$ of the Marmousi data set (Versteeg & Grau, 1991).

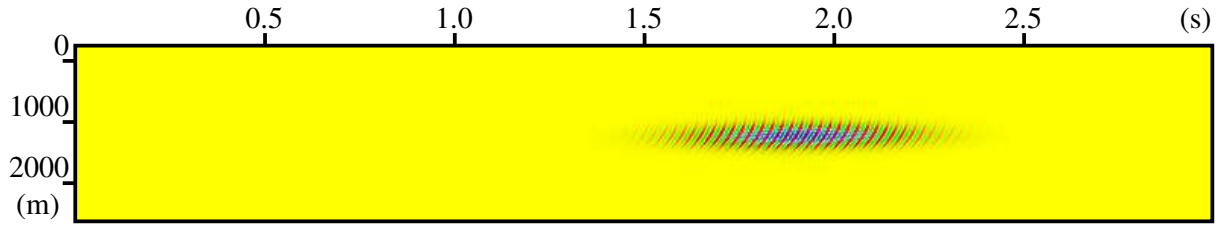

Figure 2. Gabor function.

5 Conclusions

The presented procedure of the decomposition of the wave field into Gaussian packets is fully functional. It might be suitable not only for Gaussian packet migration, but also for many different applications of the Gaussian packet method.

The choice of the discretization of the integral transform plays a crucial role in the decomposition of the wave field. It affects not only the accuracy, but also the efficiency of the method. We are able to control the discretization error of the decomposition into *uniform* Gaussian packets.

Decomposition of the wave field into Gaussian packets, whose shape depends on the coordinate of the intersection of the central ray of a Gaussian packet with the profile and on its arrival time, is strictly mathematically speaking incorrect. In order to achieve satisfactory results even in such a case, we smooth the distribution of the parameter determining the shape of a packet in the plane perpendicular to its central ray (Žáček, 2001a, 2001b).

The main disadvantage of our approach seems to be the computational cost of the decomposition. Nevertheless, once we obtain the amplitudes of Gaussian packets, we can calculate many migrated sections for various reasonably similar velocity models.

Acknowledgements

The author is greatly indebted to Luděk 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").

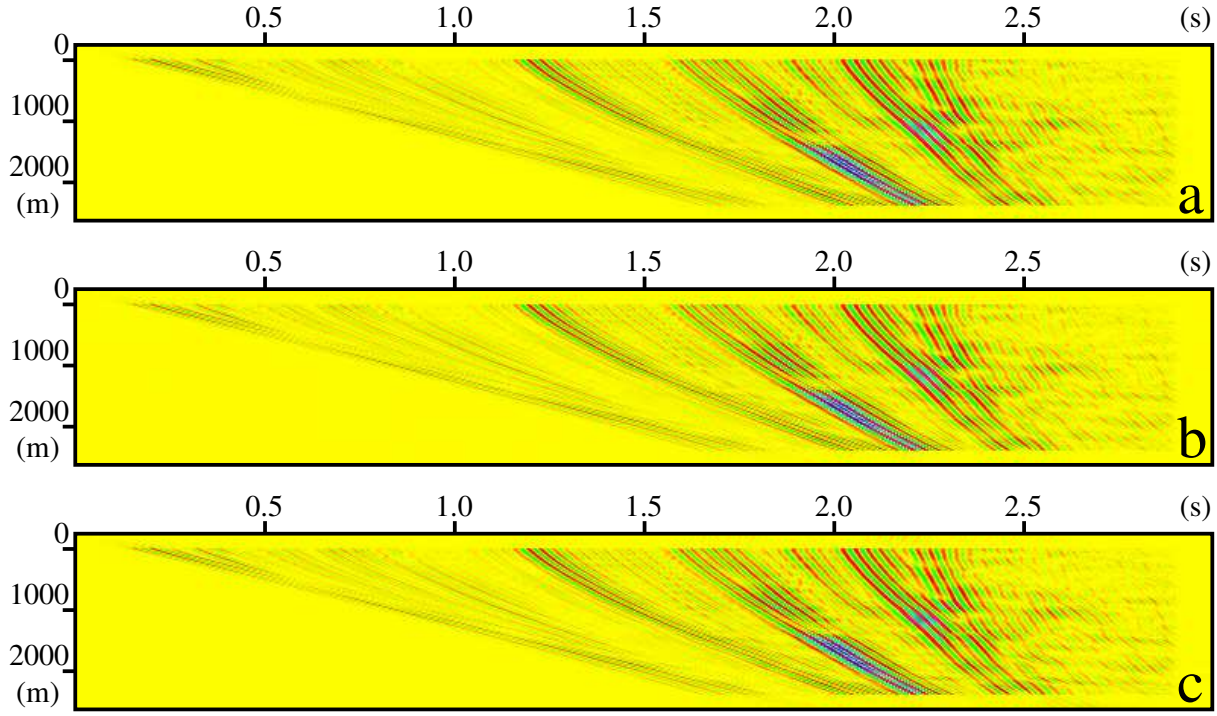

Figure 3. Approximated common-shot gathers $\widetilde{f}(x,t)$ for (a) *uniform* Gaussian packets with $\mathrm{Re}(N_0) = -0.25 \times 10^{-6}$ sm^{-2} and $\mathrm{Im}(N_0) = 0.25 \times 10^{-6}$ sm^{-2}; (b) Gaussian packets with $\mathrm{Re}(N) = -0.25 \times 10^{-6}$ sm^{-2}, and $\mathrm{Im}(N)$ ranging from 0.125×10^{-6} sm^{-2} to 0.5×10^{-6} sm^{-2}; (c) Gaussian packets with $\mathrm{Re}(N) = -0.25 \times 10^{-6}$ sm^{-2}, and $\mathrm{Im}(N)$ ranging from 0.75×10^{-7} sm^{-2} to 0.1×10^{-5} sm^{-2}.

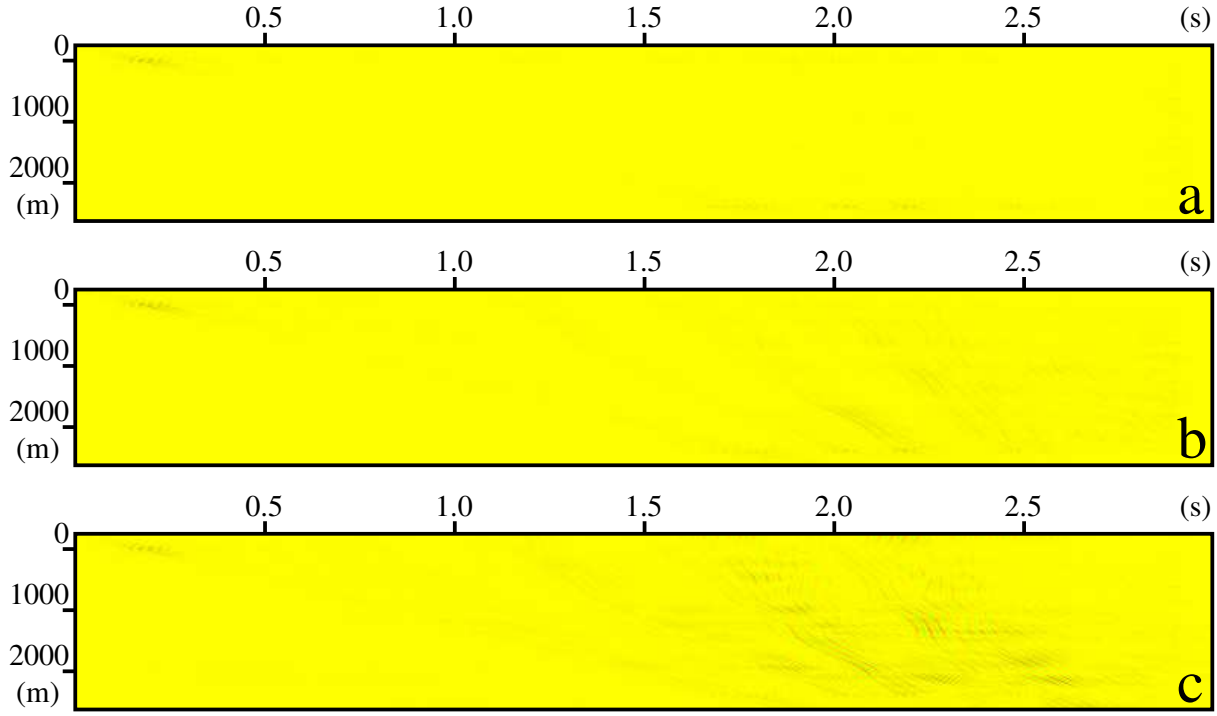

Figure 4. Differences between the original common-shot gather $f(x,t)$ and the approximated common-shot gathers $\widetilde{f}(x,t)$ for (a) *uniform* Gaussian packets with $\mathrm{Re}(N_0) = -0.25 \times 10^{-6}$ sm^{-2} and $\mathrm{Im}(N_0) = 0.25 \times 10^{-6}$ sm^{-2}; (b) Gaussian packets with $\mathrm{Re}(N) = -0.25 \times 10^{-6}$ sm^{-2}, and $\mathrm{Im}(N)$ ranging from 0.125×10^{-6} sm^{-2} to 0.5×10^{-6} sm^{-2}; (c) Gaussian packets with $\mathrm{Re}(N) = -0.25 \times 10^{-6}$ sm^{-2}, and $\mathrm{Im}(N)$ ranging from 0.75×10^{-7} sm^{-2} to 0.1×10^{-5} sm^{-2}.

References

Babich, V. M. & Ulin, V. V. (1981): Complex space time ray method and 'quasiphotons'. Zap. Nauchn. Sem. Leningr. Otd. Math. Inst., **117**, 5–12 (in Russian); English transl.: J.Sov. Math., **24**, 269–273, 1984.

Daubechies, I. (1992): Ten lectures on wavelets. CBMS-NSF regional conference series in applied mathematics, **61**, SIAM, Philadelphia.

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

Klimeš, L. (1989a): Gaussian packets in the computation of seismic wavefields. Geophys. J. Int., **99**, 421–433.

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

Klimeš, L. (2004): Gaussian packets in smooth isotropic media. In: Seismic Waves in Complex 3-D Structures, Report 14, Charles University, Prague, 43–54.

Ralston, J. (1983): Gaussian beams and the propagation of singularities. MAA Studies in Mathematicc., **23**, 206–248.

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. (2001a): Optimization of the shape of Gaussian beams. 71st Ann. Int. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 2128–2131.

Žáček, K. (2001b): Optimization of the shape of Gaussian beams. In: Seismic Waves in Complex 3-D Structures, Report 11, Charles University, Prague, 181–201.

Žáček, K. (2002): Smoothing the Marmousi model. Pure and Appl. Geophys., **159**, 1507–1526.

Žáček, K. (2003): Decomposition of the wave field into optimized Gaussian packets. 73st Ann. Int. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1869–1872.

Žáček, K. (2004): Gaussian packet pre-stack depth migration. 74th Ann. Int. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 957–960.

Žáček, K. (2005): Gaussian packet prestack depth migration. In: Seismic Waves in Complex 3-D Structures, Report 15, Charles University, Prague, 29–48.

Gaussian packet prestack depth migration

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

Summary

We present a complete algorithm and the first complete 2-D acoustic numerical examples of a brand new imaging method operating in the common-shot domain — the Gaussian packet prestack depth migration.

The main advantage of our method over the methods based on Gaussian beams is a direct relation between the regions in the common-shot gather and corresponding regions in the migrated section. Thus, the Gaussian packet prestack depth migration is especially suitable for a target-oriented imaging.

Key words

Gaussian packet, Gabor function, prestack depth migration, target-oriented imaging, multi-valued travel time, common-shot gather.

1 Introduction

Over the past years, several imaging techniques based on the seismic ray theory have been developed. Particularly simple and popular are the methods using the Kirchhoff integral. Another successful class of ray theory migrations employs Gaussian beams. An especially elegant and effective migration algorithm using Gaussian beams has been proposed by Hill (1990, 2001).

In search for an alternative to Gaussian beams, we have shifted our attention to Gaussian packets. Gaussian packets are waves whose envelopes at any given time are nearly Gaussian in space. They represent high-frequency asymptotic solutions of the elastodynamic equation, which are concentrated close to the central point of the packet (e.g., Babich & Ulin, 1981; Ralston, 1983; Klimeš, 1989a, 2004).

Like Gaussian beams, Gaussian packets become inaccurate solutions of the elastodynamic equation if they spread excessively as they propagate. This spreading depends on the complexity of the velocity model and on the initial shape of Gaussian packets. Thus, we should use a sufficiently smooth velocity model and optimize the initial shape of the packets.

The Gaussian packet migration algorithm consists of four basic steps:
a) Preparation of a suitable smooth velocity model (Žáček, 2002).
b) Optimization of the shape of Gaussian packets (Klimeš, 1989b; Žáček, 2001a, 2001b).
c) Decomposition of the wave field into optimized Gaussian packets (Žáček, 2003, 2005).
d) Back-propagation of the wave field using Gaussian packets and application of the imaging functional, which was already presented by Žáček (2004), and which we thoroughly describe in this paper.

The word *optimized* implies that the shape of Gaussian packets, in the plane perpendicular to the central ray of the packet, depends not only on the frequency, but

In: Seismic Waves in Complex 3-D Structures, Report 15 (Department of Geophysics, Charles University, Prague, 2005), pp.29–48

also on the coordinate of the intersection of the central ray of a Gaussian packet with the profile, on its arrival time, and on the component of the slowness vector along the profile.

The migration algorithm is designed here especially for numerical testing in the 2-D acoustic Marmousi model (Versteeg & Grau, 1991). That is why only scalar waves are dealt with in the imaging part of this paper, although the generalization of the Gaussian packet migration to the elastic waves is straightforward.

2 Paraxial Gaussian packet

In the case of the component notation, the upper-case indices take the values $I, J, ... = 1, 2$; the lower-case indices take the values $i, j, ... = 1, 2, 3$; the Greek indices take the values $\alpha, \beta, ... = 1, 2, 3, 4$. The Einstein summation over the pairs of identical indices is used. The asterisk (e.g., u^*) denotes the complex conjugate.

Let us write the space-time paraxial approximation of a Gaussian packet centred at point $y_\alpha(t)$ in the form of

$$\mathbf{U} = \mathbf{A} \exp\left[\mathrm{i}\omega(N_\alpha r_\alpha + \tfrac{1}{2} N_{\alpha\beta} r_\alpha r_\beta)\right] \quad , \tag{1}$$

where

$$r_i = x_i - y_i(t) \quad , \tag{2}$$

$$r_4 = x_4 - y_4(t) = x_4 - t \quad , \tag{3}$$

$$N_{\alpha\beta} = \tau_{,\alpha\beta} \tag{4}$$

and

$$N_\alpha = \tau_{,\alpha} \quad . \tag{5}$$

Here, τ denotes the complex-valued phase function. The complex-valued vectorial amplitude $\mathbf{A}$, the complex-valued second space-time derivatives of the phase function $N_{\alpha\beta}$ and the space-time slowness vector of the Gaussian packet N_α are taken at the central point $y_\alpha(t)$ of the packet. We choose

$$N_4 = -1 \quad . \tag{6}$$

As it is much easier to deal with Gaussian packets in the *ray-centred coordinate system* (e.g., Červený, 2001), we constitute the transform relations between the Cartesian and ray-centred coordinates. In our notation, h_{im} represents the i^{th} component of m^{th} orthonormal basis vector of the ray-centred coordinate system. Basis vectors h_{i1} and h_{i2} are perpendicular to the central ray; basis vector h_{i3} is tangent to the central ray.

The complex-valued second covariant derivatives $M_{\alpha\beta}$ of the phase function in ray-centred coordinates are related to the second space-time derivatives $N_{\alpha\beta}$ of the phase function in Cartesian coordinates as

$$M_{mn} = N_{jk} h_{jm} h_{kn} \quad , \tag{7}$$

$$N_{jk} = h_{jm} h_{kn} M_{mn} \quad , \tag{8}$$

$$M_{m4} = M_{4m} = N_{j4} h_{jm} \quad , \tag{9}$$

$$N_{j4} = N_{4j} = h_{jm} M_{m4} \tag{10}$$

and

$$M_{44} = N_{44} \quad . \tag{11}$$

Differentiating the eikonal equation

$$v^2\tau_{,i}\tau_{,i} - \tau_{,4}\tau_{,4} = 0 \tag{12}$$

with respect to space-time coordinates, and using relation (6), we arrive at

$$M_{i4} = -vM_{i3} - v^{-1}V_i \tag{13}$$

and

$$M_{44} = v^2 M_{33} + V_3 \quad , \tag{14}$$

where

$$V_m = v_{,i}h_{im} \tag{15}$$

is the velocity gradient in ray-centred coordinates.

Differentiating twice the eikonal equation (12) with respect to space-time coordinates, we obtain the Ricatti equation for the 4×4 space-time matrix (4), see Ralston (1983). Using relations (7) to (11), let us transform the Ricatti equation into the ray-centred coordinate system. Inserting equations (13) and (14), we obtain the Ricatti equation for the 3×3 spatial submatrix M_{mn} of the 4×4 space-time matrix $M_{\alpha\beta}$,

$$\frac{\mathrm{d}M_{mn}}{\mathrm{d}t} = -M_{jn}\frac{\mathrm{d}h_{ij}}{\mathrm{d}t}h_{im} - M_{mk}\frac{\mathrm{d}h_{ik}}{\mathrm{d}t}h_{in} - v^2 M_{Km}M_{Kn} - v^{-1}V_{mn} - V_m M_{3n} - V_n M_{3m} \quad , \tag{16}$$

where

$$V_{mn} = v_{,ij}h_{im}h_{jn} \tag{17}$$

are the covariant velocity derivatives in ray-centred coordinates. The derivatives of the basis vectors of the ray-centred coordinate system along the ray satisfy relations

$$\frac{\mathrm{d}h_{jM}}{\mathrm{d}t}h_{jN} = 0 \ , \quad \frac{\mathrm{d}h_{j3}}{\mathrm{d}t}h_{jN} = -V_N \ , \quad \frac{\mathrm{d}h_{jM}}{\mathrm{d}t}h_{j3} = V_M \ , \quad \frac{\mathrm{d}h_{j3}}{\mathrm{d}t}h_{j3} = 0 \quad . \tag{18}$$

Inserting relations (18), equation (16) may be decoupled into three equations,

$$\frac{\mathrm{d}M_{MN}}{\mathrm{d}t} = -v^2 M_{MK}M_{KN} - v^{-1}V_{MN} \quad , \tag{19}$$

$$\frac{\mathrm{d}M_{M3}}{\mathrm{d}t} = -M_{Mi}V_i - v^2 M_{MI}M_{I3} - v^{-1}V_{M3} \tag{20}$$

and

$$\frac{\mathrm{d}M_{33}}{\mathrm{d}t} = -2M_{i3}V_i - v^2 M_{3I}M_{I3} - v^{-1}V_{33} \quad . \tag{21}$$

While equation (19) is the Ricatti equation for a Gaussian beam, equations (20) and (21) extend the Gaussian beam solution to a Gaussian packet solution.

To express equations (19), (20) and (21) in the vector and matrix notation, we introduce following matrices and vectors,

$$\mathbf{M} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \quad , \tag{22}$$

$$\mathbf{M}_3 = \begin{pmatrix} M_{13} \\ M_{23} \end{pmatrix} \quad , \tag{23}$$

$$\mathbf{M}_4 = \begin{pmatrix} M_{14} \\ M_{24} \end{pmatrix} \quad , \tag{24}$$

$$\mathbf{W} = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \tag{25}$$

and

$$\mathbf{V} = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \quad . \tag{26}$$

The Ricatti equation for a Gaussian beam (19) in the matrix notation reads

$$\frac{\mathrm{d}\mathbf{M}}{\mathrm{d}t} = -v^2\mathbf{M}\mathbf{M} - v^{-1}\mathbf{W} \quad . \tag{27}$$

Its solution may be expressed as

$$\mathbf{M} = \mathbf{M}^{\mathrm{T}} = \mathbf{P}\mathbf{Q}^{-1} \quad , \tag{28}$$

where matrices $\mathbf{Q}$ and $\mathbf{P}$ satisfy the dynamic ray tracing equations (e.g., Červený, 2001)

$$\frac{\mathrm{d}\mathbf{Q}}{\mathrm{d}t} = v^2\mathbf{P} \quad , \qquad \frac{\mathrm{d}\mathbf{P}}{\mathrm{d}t} = -v^{-1}\mathbf{W}\mathbf{Q} \quad . \tag{29}$$

We introduce the *paraxial-ray propagator matrix*

$$\mathbf{\Pi} = \begin{pmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \\ \mathbf{P}_1 & \mathbf{P}_2 \end{pmatrix} \quad , \tag{30}$$

which solves equations (29) with unit initial conditions

$$\begin{pmatrix} \mathbf{Q}_1^0 & \mathbf{Q}_2^0 \\ \mathbf{P}_1^0 & \mathbf{P}_2^0 \end{pmatrix} = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \quad , \tag{31}$$

where $\mathbf{1}$ denotes a unit 2×2 matrix. Matrices $\mathbf{Q}$ and $\mathbf{P}$ may be determined from their initial values $\mathbf{Q}^0$ and $\mathbf{P}^0$ using the paraxial-ray propagator matrix as

$$\begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix} = \mathbf{\Pi} \begin{pmatrix} \mathbf{Q}^0 \\ \mathbf{P}^0 \end{pmatrix} \quad . \tag{32}$$

Equations (20) and (21), which extend the Gaussian beam solution of the Ricatti equation to a Gaussian packet solution, may be rewritten in the vector and matrix notation as follows,

$$\frac{\mathrm{d}(v\mathbf{M}_3 + v^{-1}\mathbf{V})}{\mathrm{d}t} = -v^2\mathbf{M}(v\mathbf{M}_3 + v^{-1}\mathbf{V}) \tag{33}$$

and

$$\frac{\mathrm{d}M_{33}}{\mathrm{d}t} = -2V_3M_{33} - v^2(\mathbf{M}_3 + v^{-2}\mathbf{V})^{\mathrm{T}}(\mathbf{M}_3 + v^{-2}\mathbf{V}) + v^{-2}\mathbf{V}^{\mathrm{T}}\mathbf{V} - v^{-1}V_{33} \quad . \tag{34}$$

Solutions of equations (33) and (34) read

$$\mathbf{M}_3 = v^{-1}\left[(\mathbf{Q}^{\mathrm{T}})^{-1}\mathbf{C} - v^{-1}\mathbf{V}\right] \tag{35}$$

and

$$M_{33} = v^{-2}\left\{-\frac{\mathrm{i}}{2}\mathbf{C}^{\mathrm{T}}\mathbf{Q}^{-1}\left[\mathrm{Im}(\mathbf{M})\right]^{-1}(\mathbf{Q}^{\mathrm{T}})^{-1}\mathbf{C} + C - V_3\right\} \quad , \tag{36}$$

where $\mathbf{C}$ is a constant vector determined by the initial conditions and C is an integration constant. Equations for elements $\mathbf{M}_3$, $\mathbf{M}_4$, M_{33}, M_{34} and M_{44}, which supplement the 2×2 Gaussian beam matrix M_{MN} to the 4×4 Gaussian packet matrix $M_{\alpha\beta}$, consist of two explicit evolution equations,

$$\mathbf{M}_4 = -(\mathbf{Q}^{\mathrm{T}})^{-1}\mathbf{C} \tag{37}$$

(Babich & Ulin, 1981),

$$M_{44} = -\frac{\mathrm{i}}{2}\mathbf{C}^{\mathrm{T}}\mathbf{Q}^{-1}\left[\mathrm{Im}(\mathbf{M})\right]^{-1}(\mathbf{Q}^{\mathrm{T}})^{-1}\mathbf{C} + C = -\frac{\mathrm{i}}{2}\mathbf{M}_4^{\mathrm{T}}\left[\mathrm{Im}(\mathbf{M})\right]^{-1}\mathbf{M}_4 + C \quad , \tag{38}$$

and three simple relations,

$$\mathbf{M}_3 = -v^{-1}\mathbf{M}_4 - v^{-2}\mathbf{V} \quad , \tag{39}$$

$$M_{33} = v^{-2}\left(M_{44} - V_3\right) \tag{40}$$

and

$$M_{34} = -v^{-1}M_{44} \quad . \tag{41}$$

Following equations (37) and (38), we determine constants $\mathbf{C}$ and C from the initial values $\mathbf{M}_4^0$ and M_{44}^0 of elements $\mathbf{M}_4$ and M_{44},

$$\mathbf{C} = -\mathbf{Q}^{0\,\mathrm{T}}\mathbf{M}_4^0 \tag{42}$$

and

$$C = M_{44}^0 + \frac{\mathrm{i}}{2}\mathbf{M}_4^{0\,\mathrm{T}}\left[\mathrm{Im}(\mathbf{M}^0)\right]^{-1}\mathbf{M}_4^0 \quad . \tag{43}$$

The complex-valued vectorial amplitude of a Gaussian packet $\mathbf{A}$, see equation (1), may be expressed as

$$\mathbf{A} = A\mathbf{e} \quad , \tag{44}$$

where A is the complex-valued scalar amplitude; $\mathbf{e}$ denotes the unit polarization vector of a Gaussian packet, which is identical to the Gaussian beam polarization vector and to the ray-theory polarization vector. The complex-valued scalar amplitude evolves along the central ray of a Gaussian packet according to equation (Babich & Ulin, 1981)

$$A = A^0\sqrt{\frac{v^0\rho^0\det(\mathbf{Q}^0)}{v\rho\,\det(\mathbf{Q})}} \quad , \tag{45}$$

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

3 Imaging

We choose the imaging functional for scalar waves in the form of

$$M(u,U) = \int \mathrm{d}\omega\ \frac{\widehat{F}(\omega)\widehat{u}^*(\omega)\widehat{U}(\omega)}{\widehat{u}^*(\omega)\widehat{u}(\omega)}\left[\int \mathrm{d}\omega\ \widehat{F}(\omega)\right]^{-1} \quad , \tag{46}$$

where $\widehat{u}$ is the temporal Fourier transform of the incident wavefield u and $\widehat{U}$ is the temporal Fourier transform of the back-propagated scattered wavefield U. Let us take the spectral filter $\widehat{F}$ in the form of the Fourier transform $\widehat{s}$ of the source time function s, possibly shifted by time $-t_{\mathrm{S}}$ or phase $-\varphi_{\mathrm{S}}$,

$$\widehat{F}(\omega) = \widehat{s}(\omega)\exp(\mathrm{i}\varphi_{\mathrm{S}} - \mathrm{i}\omega t_{\mathrm{S}}) \quad . \tag{47}$$

We assume the incident wavefield $\widehat{u}(\omega)$ approximated by several ray-theory elementary seismograms $\widehat{u}_\Omega(\omega)$ corresponding to individual branches Ω of possibly multi-valued travel time,

$$\widehat{u}(\omega) = \sum_\Omega \widehat{u}_\Omega(\omega) \quad , \tag{48}$$

with

$$\widehat{u}_\Omega(\omega) = \widehat{s}(\omega)\widehat{a}_\Omega \exp(\mathrm{i}\omega\tau_\Omega) \quad , \tag{49}$$

where a_Ω is the amplitude corresponding to travel time τ_Ω. To avoid singularities and to simplify the calculation, we apply approximation

$$\widehat{u}^*(\omega)\widehat{u}(\omega) \approx \widehat{s}^*(\omega)\widehat{s}(\omega) \sum_\Omega \widehat{a}^*_\Omega(\omega)\widehat{a}_\Omega(\omega) \quad . \tag{50}$$

The back-propagated scattered wavefield may be composed of individual Gaussian packets U_R,

$$\widehat{U}(\omega) = \sum_\mathrm{R} \widehat{U}_\mathrm{R}(\omega) \quad . \tag{51}$$

Then, the imaging functional takes the form of

$$M(u,U) = \sum_\mathrm{R} M(u,U_\mathrm{R}) \quad , \tag{52}$$

where

$$M(u,U_\mathrm{R}) \approx \int \mathrm{d}\omega \; \frac{\exp(\mathrm{i}\varphi_\mathrm{S} - \mathrm{i}\omega t_\mathrm{S}) \sum_\Omega \widehat{a}^*_\Omega(\omega) \exp(-\mathrm{i}\omega\tau_\Omega)\, \widehat{U}_\mathrm{R}(\omega)}{\sum_\Omega \widehat{a}^*_\Omega(\omega)\widehat{a}_\Omega(\omega)} \left[\int \mathrm{d}\omega \; \widehat{F}(\omega)\right]^{-1} . \tag{53}$$

In the time domain, equation (53) reads

$$M(u,U_\mathrm{R}) \approx \frac{\sum_\Omega \exp(\mathrm{i}\varphi_\mathrm{S})\, a^*_\Omega\, U_\mathrm{R}(t_\mathrm{S} + \tau_\Omega)}{F(0)\ \sum_\Omega a^*_\Omega a_\Omega} \quad . \tag{54}$$

Although Gaussian packets are moving from the target zone towards the corresponding receiver, we expect that the central rays of the packets are calculated from the receiver towards the target zone.

Travel time T from the receiver is then used as the parameter along the central ray, whereas $y_i(T)$ and $y_4(T) = t_\mathrm{R} - T$ are spatial coordinates and the travel time corresponding to the centre of the Gaussian packet moving towards the receiver, respectively.

Because of the opposite propagation, the slowness vector of the packet N_i is minus the slowness vector P_i of the ray coming from the receiver,

$$N_i = -P_i \quad . \tag{55}$$

We choose

$$N_4 = -1 \quad . \tag{56}$$

Paraxial approximation of a Gaussian packet propagating towards the corresponding receiver to reach it at time $x_4 = t_\mathrm{R}$, centred at point $y_\alpha(T)$, is

$$U_\mathrm{RT}(x_4) = A_\mathrm{RT} \exp\left[\mathrm{i}\omega(N_\alpha r_\alpha + \tfrac{1}{2} N_{\alpha\beta} r_\alpha r_\beta)\right] \quad , \tag{57}$$

where

$$r_i = x_i - y_i(T) \tag{58}$$

and

$$r_4 = x_4 - y_4(T) = x_4 - (t_\mathrm{R} - T) \quad . \tag{59}$$

Matrix $N_{\alpha\beta}$ corresponds to a Gaussian packet moving towards a receiver, although it is, like the travel time y_4 and slowness vector N_α, expressed in terms of the quantities calculated along the ray emanated from the receiver.

Paraxial approximation (57) is accurate in the vicinity of point $y_\alpha(T)$, but becomes inaccurate at large space-time distances from $y_\alpha(T)$. Assuming discrete points $y_i(T)$ with step ΔT along the central ray, we thus limit each paraxial approximation by weighting function

$$W_{\mathrm{T}} = \sqrt{\pi}\,\kappa^{-1}\,\exp\left[-\kappa^2 (N_i r_i)^2 \Delta T^{-2}\right] \quad , \tag{60}$$

where parameter κ controls the shape of the Gaussian function and, consequently, the error of the weighting. Sum of these weighting functions is approximately 1 in the vicinity of the central ray and thus

$$U_{\mathrm{R}}(t_{\mathrm{S}} + \tau_\Omega) = \sum_{\mathrm{T}} W_{\mathrm{T}}\, U_{\mathrm{RT}}(t_{\mathrm{S}} + \tau_\Omega) \quad . \tag{61}$$

According to equations (54) and (61), the image of a single Gaussian packet reads

$$M(u, U_{\mathrm{R}}) = \sum_{\mathrm{T}} M(u, W_{\mathrm{T}} U_{\mathrm{RT}}) \quad , \tag{62}$$

where the image of a single paraxial approximation is

$$M(u, W_{\mathrm{T}} U_{\mathrm{RT}}) = \frac{\sqrt{\pi} \sum_\Omega a^*_\Omega\, A_{\mathrm{RT}}\, \exp\left[\mathrm{i}\varphi_{\mathrm{S}} + \mathrm{i}\omega(N_\alpha r_\alpha + \frac{1}{2} N_{\alpha\beta} r_\alpha r_\beta) - \kappa^2 (N_i r_i)^2 \Delta T^{-2}\right]}{F(0)\ \kappa\ \sum_\Omega a^*_\Omega a_\Omega} \ , \tag{63}$$

with

$$r_4 = t_{\mathrm{S}} + \tau_\Omega(x_i) + T - t_{\mathrm{R}} \quad . \tag{64}$$

Inserting equation (62) into (52), the imaging functional takes the form of

$$M(u, U) = \sum_{\mathrm{R}} \sum_{\mathrm{T}} M(u, W_{\mathrm{T}} U_{\mathrm{RT}}) \quad . \tag{65}$$

4 Algorithm

Like other types of wave packets, Gaussian packets spread as they propagate through the structure. It is necessary to keep Gaussian packets narrow in relation to the velocity changes in the model, because Gaussian packets become inaccurate solutions of the elastodynamic equation if the velocity field changes considerably within the packet width. This spreading depends on the complexity of the velocity model and on the initial shape of Gaussian packets. Therefore, before proceeding to the migration, we need to prepare a suitable velocity model and choose the appropriate initial shape of Gaussian packets.

Since the Gaussian packet prestack depth migration operates in the wave field domain, the initial amplitudes of Gaussian packets are obtained by decomposing the common-shot gathers into Gabor functions.

In the process of imaging, we compare the incident wave field u and the back-propagated scattered wave field U. The back-propagated wave field consists of individual Gaussian packets. We choose the imaging functional in the form of expression (46), although other varieties of imaging functionals could be admissible as well.

4.1 Velocity model

A velocity model should prevent Gaussian packets from excessive spreading and approximately preserve original travel times. The question of the smoothness of the velocity model is closely related to the problem of finding the limits of applicability of the ray theory, which remains unsolved and open for further research.

In a complex model, the geometrical spreading and number of arrivals exponentially increase with increasing travel time. The exponential increment is controlled by the Lyapunov exponent (e.g., Klimeš, 2002). Consequently, the Lyapunov exponent determines the horizon, where the ray behavior becomes chaotic. Since the Lyapunov exponent depends on the second spatial derivatives of velocity or slowness, the second derivatives should be minimized. Therefore, we smooth slowness by minimizing the relevant Sobolev norm composed of the second derivatives of slowness (Žáček, 2002).

4.2 Optimization of the shape of Gaussian packets

The applicability and accuracy of the Gaussian packet method depend on the proper choice of the initial shape of packets. Unfortunately, narrow Gaussian packets quickly increase in width as they propagate. Thus, we can use neither too narrow nor too wide packets as the initial choice of the shape of Gaussian packets. Furthermore, in a complex structure, we cannot judge solely from the final width of the packet whether the packet is or is not a reasonably accurate solution of the elastodynamic equation. The packet must be sufficiently narrow along the whole ray path.

The initial shape of Gaussian packets is determined by the complex-valued second space-time derivatives of the phase function $N^0_{\alpha\beta}$, see equation (1). To prevent Gaussian packets from excessive spreading, we optimize their initial shape by minimizing their width along the whole ray path. Since we wish to control the spreading in the plane perpendicular to the central ray of the packet, see equation (27), it is sufficient to optimize the initial shape of corresponding Gaussian beams (Klimeš, 1989b; Žáček, 2001a, 2001b). We can restrict our attention to Gaussian beams, because the Ricatti equation for a Gaussian packet (16) may be decoupled into three equations, where equation (19) is the Ricatti equation for a Gaussian beam, and equations (20) and (21) extend the Gaussian beam solution to a Gaussian packet solution.

The shape of a Gaussian beam is characterized by a 2×2 complex-valued matrix in 3-D and by a single complex number in 2-D. Let us denote by N^0 the initial complex-valued parameter in 2-D, which was defined as parameter N by Žáček (2005). The imaginary part of N^0 determines the Gaussian beam width and the real part of N^0 defines the curvature of the phase-front.

Parameter N^0, which describes the shape of a corresponding Gaussian beam along the initial surface, may depend on the coordinate of the intersection of the central ray of a Gaussian packet with the profile x_{R}, on its arrival time t_{R}, and on the component of the slowness vector along the profile p. In other words, parameter N^0 is situated on a *Hamiltonian hypersurface* in the phase-space (Žáček, 2001a, 2001b). In a complex structure, the optimum initial parameters $N^0(x_{\mathrm{R}}, t_{\mathrm{R}}, p)$ may vary in orders of magnitude. This would cause great problems in the decomposition of the wave field into optimized Gaussian packets. Therefore, we have developed a procedure, which allows us to smooth iteratively the distribution of parameter N^0 on the Hamiltonian hypersurface (Žáček, 2001a, 2001b).

4.3 Decomposition of the wave field into Gaussian packets

In order to determine the initial amplitudes of Gaussian packets A^0, see equation (45), we need to decompose the common-shot gather into optimized Gaussian packets (Žáček, 2003, 2005). Each Gaussian packet arriving at the receivers is represented by a Gabor function, whose shape depends on the shape of the Gaussian packet. Using the *coherent-state transform*, we decompose the common-shot gather into these Gabor functions. The complex valued amplitude of the back-propagated Gaussian packet A^0 then equals the amplitude of the corresponding Gabor function. The amplitudes depend on the coordinate of the intersection of the central ray of a Gaussian packet with the profile x_{R}, its arrival time t_{R}, the component of the slowness vector along the profile p, and the frequency ω. Let us point out that the intersection of the central ray of a Gaussian packet with the profile x_{R} does not generally coincide with any receiver.

In 2-D, the initial shape of a Gaussian packet is defined by two free complex-valued parameters — N^0 and K^0. Parameter N^0 should be determined by the optimization of the shape of corresponding Gaussian beams. We choose parameter K^0, which was introduced as parameter K_0 by Žáček (2005), and which describes the shape of a Gaussian packet along the central ray, pure imaginary and uniform for all Gaussian packets. Please note that parameter K^0 is used in the choice of N_{44}, see Žáček (2005, eq. 25) and equation (66) below.

Let us mention two problems closely connected with optimized Gaussian packets. The decomposition is exact only for N^0 independent on x_{R} and t_{R}. However, we may achieve satisfactory results for a reasonably smooth distribution of parameter N^0 on the Hamiltonian hypersurface (Žáček, 2003, 2005).

Unfortunately, the non-uniform shapes of Gaussian packets also affect the amount of data we have to work with, because the maximum steps in x_{R} and p depend on parameter N^0. In 2-D, for the imaginary part of N^0 varying from $[\mathrm{Im}(N^0)]_{\mathrm{MIN}}$ to $[\mathrm{Im}(N^0)]_{\mathrm{MAX}}$, the number of grid points in our 4-dimensional phase-space is approximately proportional to $\{[\mathrm{Im}(N^0)]_{\mathrm{MAX}}/[\mathrm{Im}(N^0)]_{\mathrm{MIN}}\}^{1/2}$. This could seriously jeopardize the efficiency of our method.

We overcome these difficulties by smoothing the distribution of parameter N^0 on the Hamiltonian hypersurface. To simplify the calculation, we can smooth N^0 to a constant value and obtain uniform Gaussian packets. Naturally, the smoothing brings about more rapid spreading of Gaussian packets.

4.4 Migration

We begin the main process of migration by determining the incident wave field u and the back-propagated scattered wave field U. Then, we construct the migrated image, which is given by expressions (63) and (65).

The incident wave field u may be calculated in various ways. We benefit from a method of interpolation within ray cells proposed by Bulant & Klimeš (1999), which requires only a sufficiently dense set of rays from the shot location to be traced. By interpolation, we obtain the multi-valued amplitudes of the incident wave field a_Ω corresponding to travel times τ_Ω.

We determine the scattered wave field U using the Gaussian packet method. Gaussian packets propagate along their central rays, which have to be shot with the same steps in x_{R} and p as in the decomposition of the wave field. We move along these central

rays with step ΔT, determine matrix $N_{\alpha\beta}$ and the amplitude of a Gaussian packet A_{RT}, and calculate expression (63).

In 2-D, we choose basis vector h_{i1} perpendicular to the plane of calculation and, consequently, basis vector h_{i2} parallel to the plane of calculation. We select the Cartesian coordinate system with vector $(0,1,0)$ tangent to the initial surface in the plane of calculation and vector $(0,0,1)$ perpendicular to the initial surface. Then, matrix $M^0_{\alpha\beta}$ reads

$$\mathbf{M}^0 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{N^0}{h_{22}^2} + 2\frac{V_2}{v^2}\frac{h_{23}}{h_{22}} + \frac{V_3}{v^2}\left(\frac{h_{23}}{h_{22}}\right)^2 & -\frac{V_2}{v^2} & 0 \\ 0 & -\frac{V_2}{v^2} & \frac{N^0_{44}-V_3}{v^2} & -\frac{N^0_{44}}{v} \\ 0 & 0 & -\frac{N^0_{44}}{v} & N^0_{44} \end{pmatrix} \quad , \tag{65}$$

see equations (8), (39) with $\mathbf{M}_4 = \mathbf{0}$, (40) and (41). Please note that parameter N^0 corresponds to N^0_{22}, see equation (8). According to Žáček (2005, eq. 31), we choose

$$N^0_{44}(x_{\mathrm{R}}, t_{\mathrm{R}}, p, \omega) = \frac{N^0(x_{\mathrm{R}}, t_{\mathrm{R}}, p) K^0}{\omega N^0(x_{\mathrm{R}}, t_{\mathrm{R}}, p) - p^2 K^0} \quad . \tag{66}$$

In order to calculate expressions (28), (37), (38), (39), (40) and (41), which yield matrix $M_{\alpha\beta}$ corresponding to its initial choice $M^0_{\alpha\beta}$, we need to determine matrices $\mathbf{Q}$ and $\mathbf{P}$. Since the central ray of a Gaussian packet is calculated from the surface towards the target zone, whereas the packet propagates from the target zone towards the surface, we have to change the sign of the slowness vector, see equation (55), and modify equation (32) into

$$\begin{pmatrix} \mathbf{Q} \\ \mathbf{P} \end{pmatrix} = \begin{pmatrix} \mathbf{Q}_1 & -\mathbf{Q}_2 \\ -\mathbf{P}_1 & \mathbf{P}_2 \end{pmatrix} \begin{pmatrix} \mathbf{Q}^0 \\ \mathbf{P}^0 \end{pmatrix} \quad , \tag{67}$$

where

$$\mathbf{Q}^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{68}$$

and

$$\mathbf{P}^0 = \begin{pmatrix} 0 & 0 \\ 0 & M^0_{22} \end{pmatrix} \quad . \tag{69}$$

Using relations (8), (10) and (11), we transform matrix $M_{\alpha\beta}$ into matrix $N_{\alpha\beta}$, which describes the shape of a Gaussian packet in Cartesian coordinates.

The amplitude of a Gaussian packet evolves along the central ray according to equation (45). Following our choice of matrix $\mathbf{Q}^0$, see relation (68), we can write

$$A_{\mathrm{RT}} = A^0(x_{\mathrm{R}}, t_{\mathrm{R}}, p, \omega) \sqrt{\frac{v^0 \rho^0}{v \rho}} \frac{1}{\sqrt{\det(\mathbf{Q})}} \quad , \tag{70}$$

where the initial amplitude A^0 comes from the decomposition of the wave field. In the case of forward propagation, the complex-valued square root depends on the number of caustic points along the central ray as

$$\frac{1}{\sqrt{\det(\mathbf{Q})}} = \sqrt{\frac{\mathrm{i}\ \mathrm{sgn}\left[\det(\mathbf{Q}_2)\right]}{\det(\mathbf{Q})}}\ \exp\left[-\mathrm{i}\frac{\pi}{4}(2k+1)\right] \quad , \tag{71}$$

where the square root is taken with the positive real part. Index k denotes the *KMAH* index, which is defined as the sum of the caustic points along the ray (e.g., Červený,

2001). As the Gaussian packet propagates from the target zone towards the surface, whereas the initial value of matrix $\mathbf{Q}$ is defined on the surface, we have to change the sign of the *KMAH* index. Thus,

$$A_{\mathrm{RT}} = A^0(x_{\mathrm{R}}, t_{\mathrm{R}}, p, \omega)\sqrt{\frac{v^0\rho^0}{v\rho}}\sqrt{\frac{\mathrm{i}\ \mathrm{sgn}\left[\det(\mathbf{Q}_2)\right]}{\det(\mathbf{Q})}}\ \exp\left[\mathrm{i}\frac{\pi}{4}(2k-1)\right] \quad . \tag{72}$$

As soon as we determine the multi-valued amplitudes of the incident wave field a_Ω corresponding to travel times τ_Ω, and the amplitude A_{RT} and the second space-time derivatives of the phase function $N_{\alpha\beta}$ of the back-propagated Gaussian packet, we calculate expression (62), which yields the localized image of a single Gabor function from the common-shot gather. Then, we superimpose the images of all Gabor functions corresponding to a single common-shot gather, see expression (65), and obtain the prestack migrated image of this common-shot gather. Finally, we can stack all of the images to produce the migrated image of the whole data set.

5 Target-oriented imaging

Each Gabor function from the common-shot gather generates its localized image in the depth section. In this way, we obtain a one-to-one relation between the Gabor functions from the common-shot gather and their localized images from the depth section. This relation was discussed in greater detail by Žáček & Klimeš (2003).

Let us say that we are interested in a particular area in the subsurface structure. We call this area a *target zone*. Unlike Gaussian beams, Gaussian packets are, at any given time, concentrated close to the central point of the packet. Thus, we can consider only those packets, which fall into the target zone.

Moreover, we can pick out Gaussian packets contributing to the target zone, multiply their amplitudes by corresponding weighting factors, and use them in the recomposition of the common-shot gather. In such a case, we obtain the wave field scattered specifically from the target zone.

6 Numerical examples

We tested our method on the Marmousi data set (Versteeg & Grau, 1991). The Marmousi model represents a very complex structure, see Figure 1. The dimensions of the model are 9200 metres (length) by 3000 metres (depth). The values of velocity, which are defined on a grid of cells of 4×4 metres, vary from 1500 ms^{-1} to 5500 ms^{-1}. Since the original model is not suitable for any ray method, we minimized the relevant Sobolev norm of slowness as described in Section 4.1, and obtained three different smooth velocity models, see Figure 2. Each model is characterized by its *average Lyapunov exponent* λ (Klimeš, 2002; Žáček, 2002), which controls the average geometrical spreading of the ray field. The average Lyapunov exponent reads $\lambda = 0.52\ \mathrm{s}^{-1}$ for the model in Figure 2a, $\lambda = 0.68\ \mathrm{s}^{-1}$ for the model in Figure 2b, and $\lambda = 0.89\ \mathrm{s}^{-1}$ for the model in Figure 2c.

The Gaussian packet migration operates in the common-shot domain. Following the procedure explained in Section 4.3, we decompose the common-shot gather, see Figure 3a, into individual Gabor functions, see Figure 3b. The complex-valued amplitude of a Gaussian packet equals the amplitude of the corresponding Gabor function. Using expression (62), we calculate the localized image of a single Gabor function from the

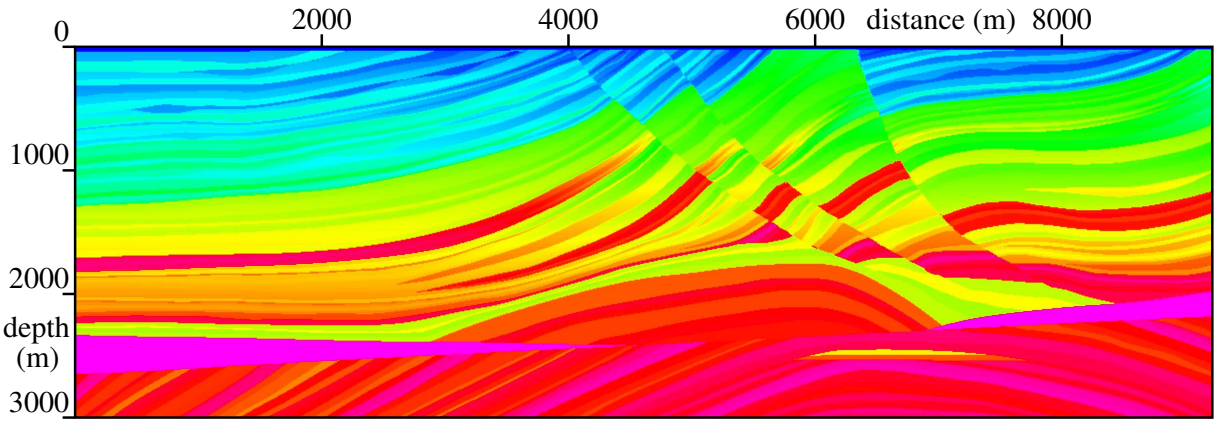

Figure 1. Marmousi model (Versteeg & Grau, 1991).

common-shot gather, see Figure 3c. Let us point out that Figures 3b and 3c demonstrate the unique one-to-one relation between the Gabor function from the common-shot gather and its localized image from the migrated section. Finally, we superimpose all relevant images according to relation (65), and obtain the prestack migrated image of the common-shot gather, see Figure 3d.

In Figure 4, we present the migrated images of individual Gabor functions for three different arrival times of a Gaussian packet t_{R}, but for the same coordinate of the intersection of the central ray of a Gaussian packet with the profile x_{R}, the same component of the slowness vector along the profile p, and the same frequency ω.

In Figure 5, we show the migrated images of individual Gabor functions for three different frequencies ω, but for the same coordinate of the intersection of the central ray of a Gaussian packet with the profile x_{R}, the same arrival time t_{R}, and the same component of the slowness vector along the profile p.

Summation of the migrated images of Gabor functions of a fixed shape over all arrival times t_{R} yields the migrated image of a corresponding Gaussian beam. In Figure 6, we demonstrate such summation for three different frequencies ω, but for the same coordinate of the intersection of the central ray with the profile x_{R} and the same component of the slowness vector along the profile p. Please note that the colour coded amplitudes in Figures 3, 4, 5 and 6 are normalized and do not represent the true amplitudes.

Since we wish to minimize the spreading, we optimize the initial shape of Gaussian packets according to the procedure suggested in Section 4.2. To simplify the tests, we have decided to smooth the resulting distribution of parameter N^0 on the Hamiltonian hypersurface in the phase-space to a constant value. Such smoothing is sensitive to the maximum travel time we take into account. Moreover, the smoothed uniform parameters of Gaussian packets vary for different common-shot gathers. As not all the common-shot gathers are equally important for the optimization, we cannot simply average corresponding parameters N^0 with uniform weights. Therefore, the choice of the most suitable parameter N^0 for the whole data set depends not only on the strict mathematical procedure suggested by Žáček (2001a, 2001b), but also on our experience. We select $\mathrm{Re}(N^0) = -0.25 \times 10^{-6}$ sm^{-2} and $\mathrm{Im}(N^0) = 0.25 \times 10^{-6}$ sm^{-2}. Parameter K^0, which describes the shape of a Gaussian packet along the central ray, is considered pure imaginary, $\mathrm{Im}(K^0) = 8.8$ s^{-2}.

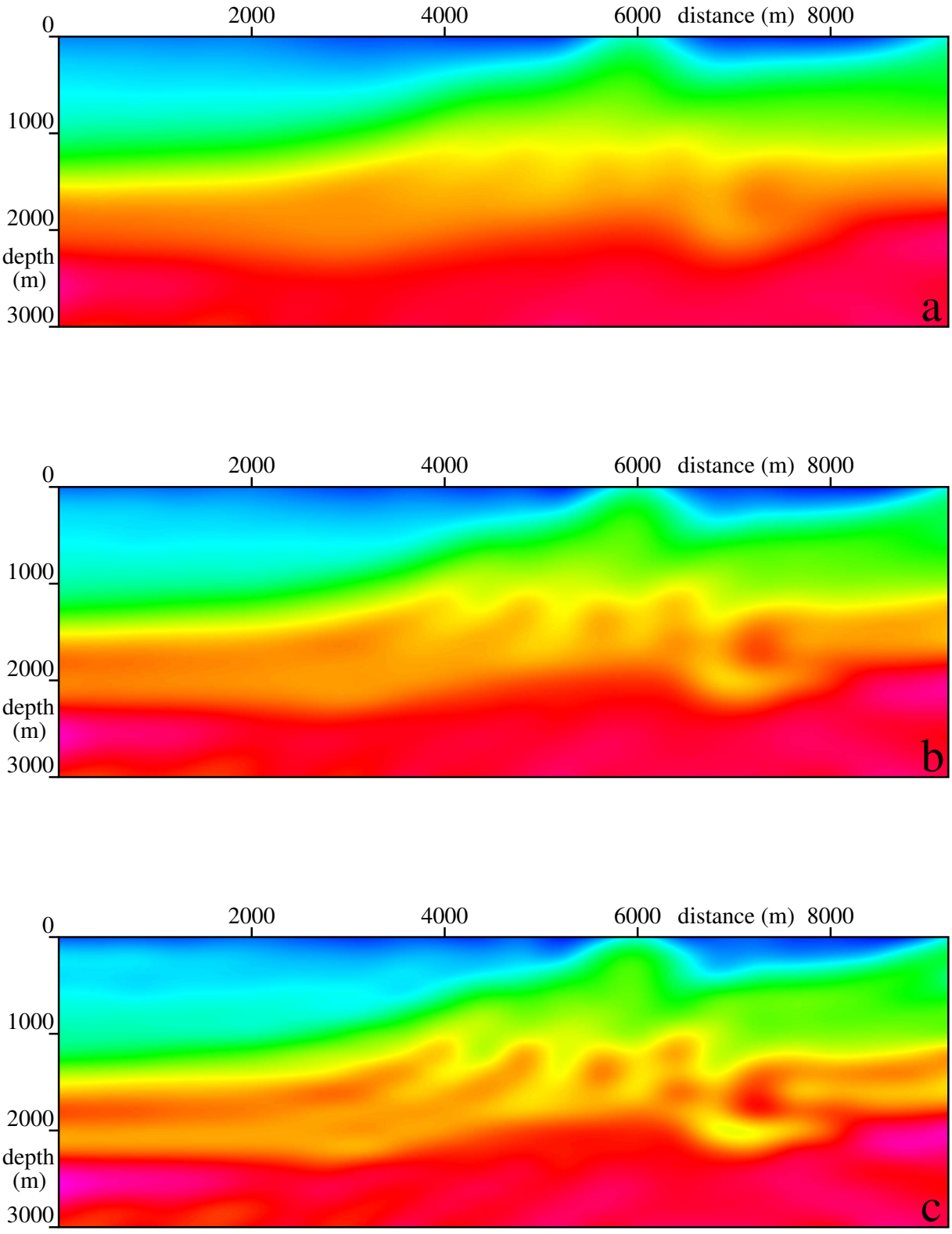

Figure 2. Smoothed models with the average Lyapunov exponents λ of (a) 0.52 s^{-1}, (b) 0.68 s^{-1} and (c) 0.89 s^{-1}. The average Lyapunov exponent controls the average geometrical spreading of the ray field.

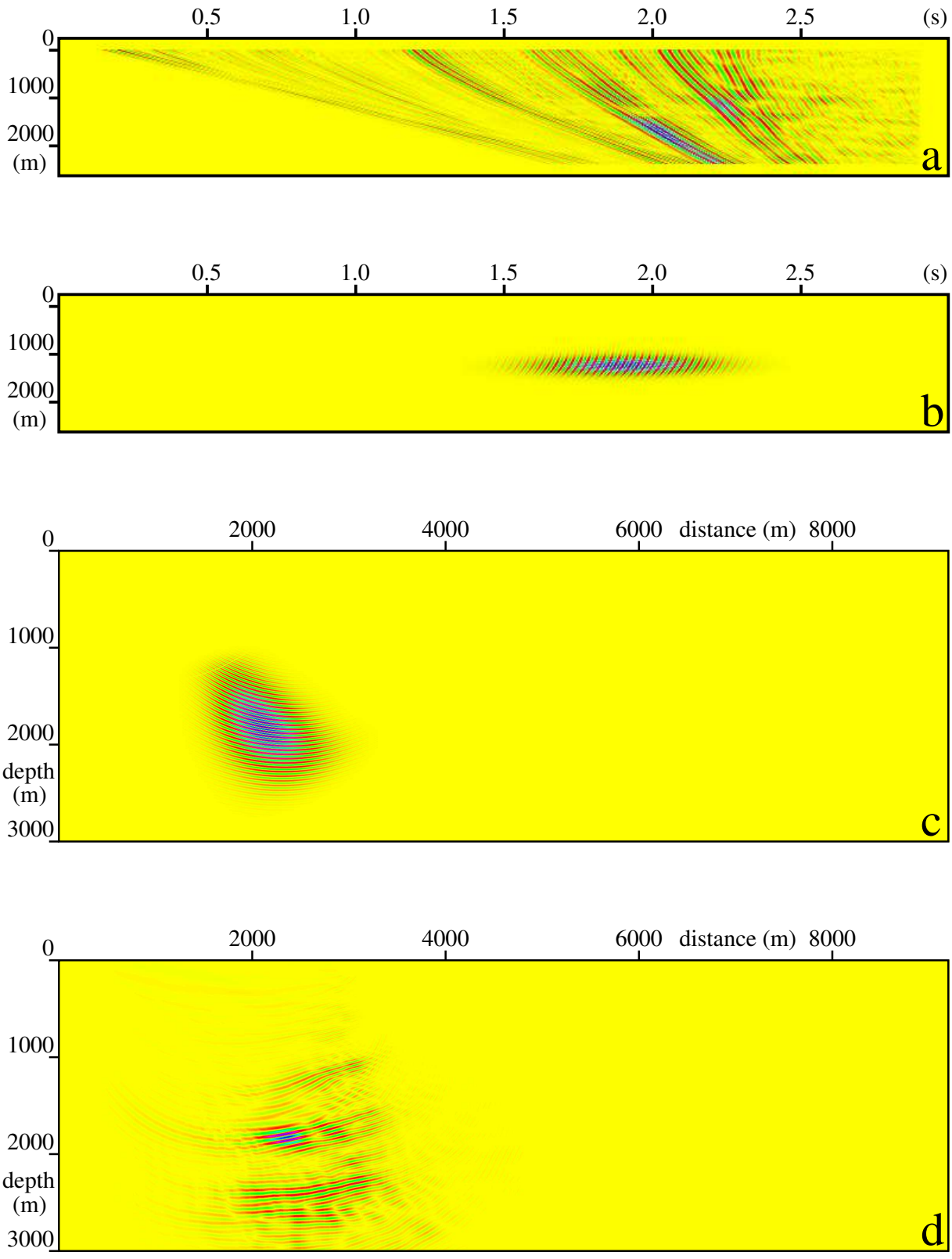

Figure 3. We decompose the common-shot gather of the Marmousi data set (a) into individual Gabor functions (b). Each Gabor function generates its localized image in the depth section (c). Superimposing all relevant images, we obtain the migrated image of the common-shot gather (d).

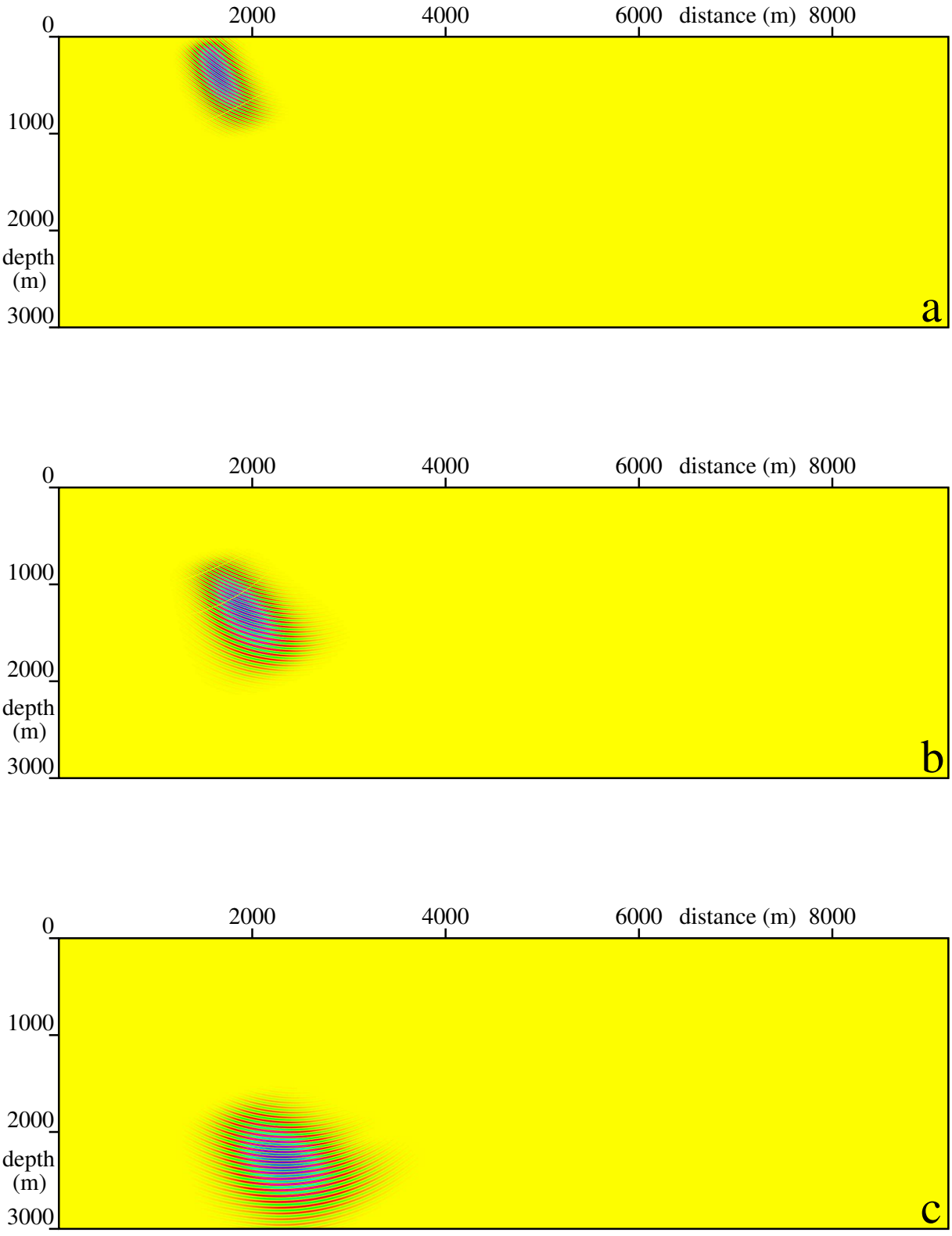

Figure 4. Migrated images of individual Gabor functions for three different arrival times t_{R}, but for the same coordinate of the intersection of the central ray of a Gaussian packet with the profile x_{R}, the same component of the slowness vector along the profile p, and the same frequency ω.

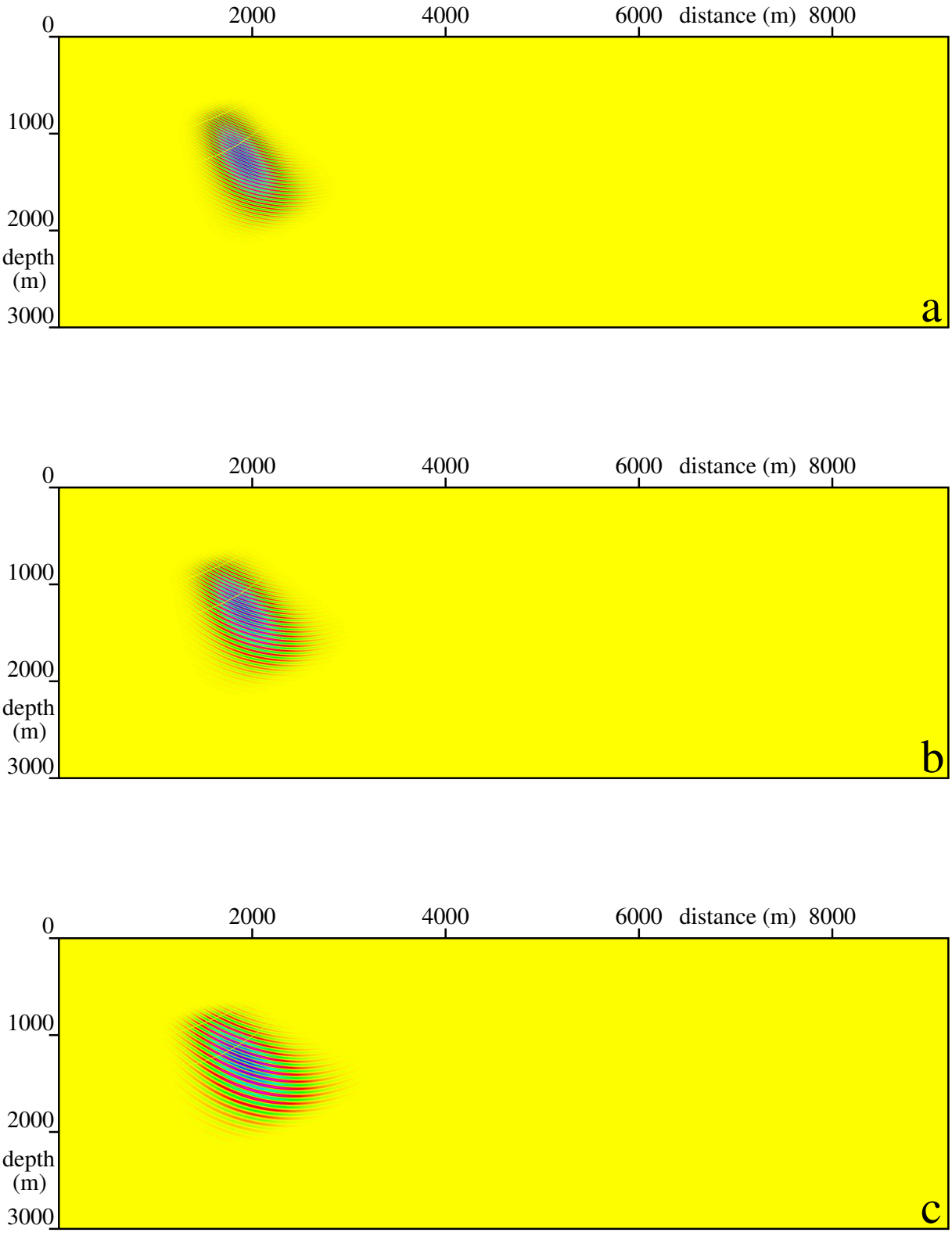

Figure 5. Migrated images of individual Gabor functions for three different frequencies ω, but for the same coordinate of the intersection of the central ray of a Gaussian packet with the profile x_{R}, the same arrival time t_{R}, and the same component of the slowness vector along the profile p.

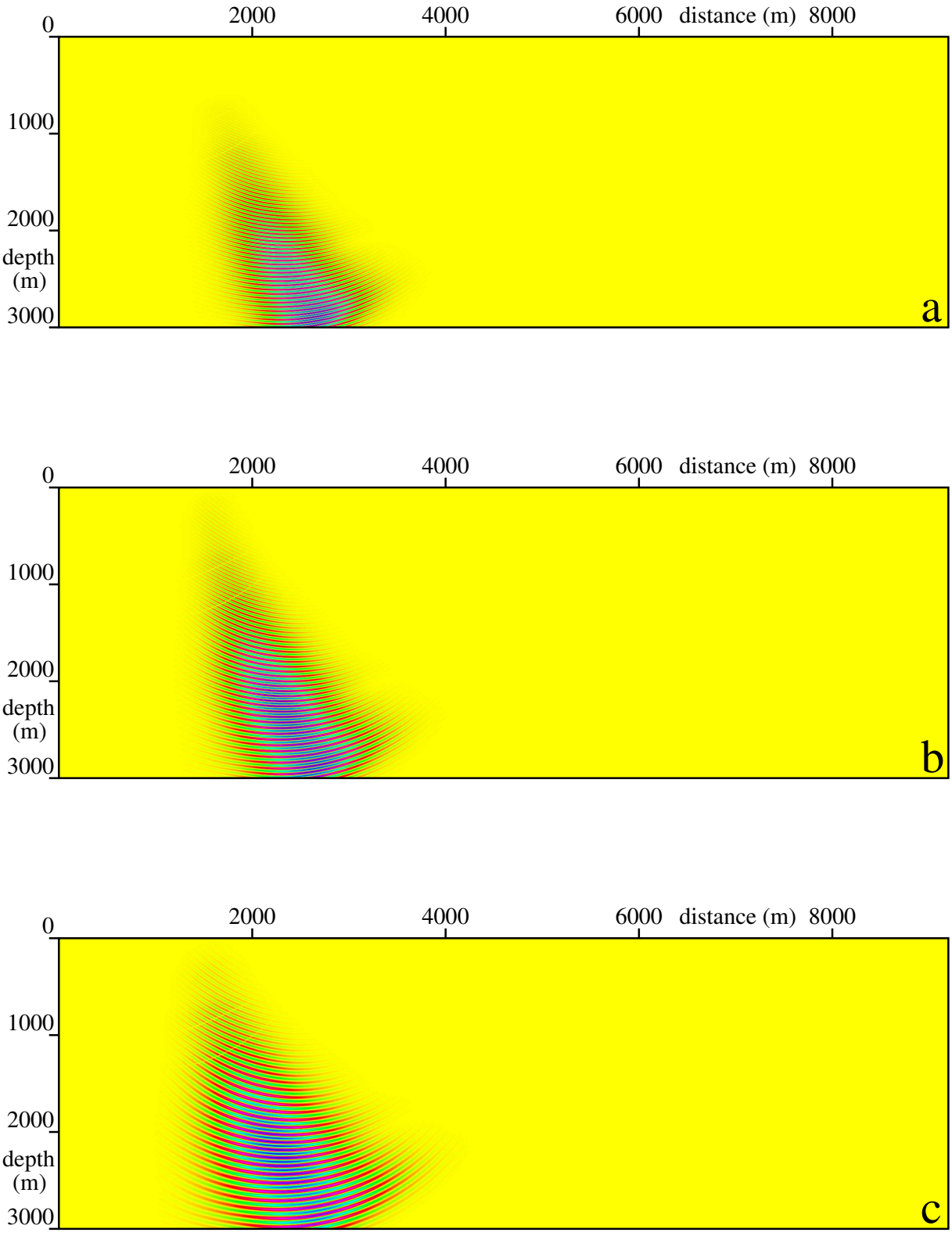

Figure 6. Migrated images of Gabor functions summed over all arrival times t_{R} for three different frequencies ω, but for the same coordinate of the intersection of the central ray of a Gaussian packet with the profile x_{R} and the same component of the slowness vector along the profile p. These images correspond to migrated images of Gaussian beams.

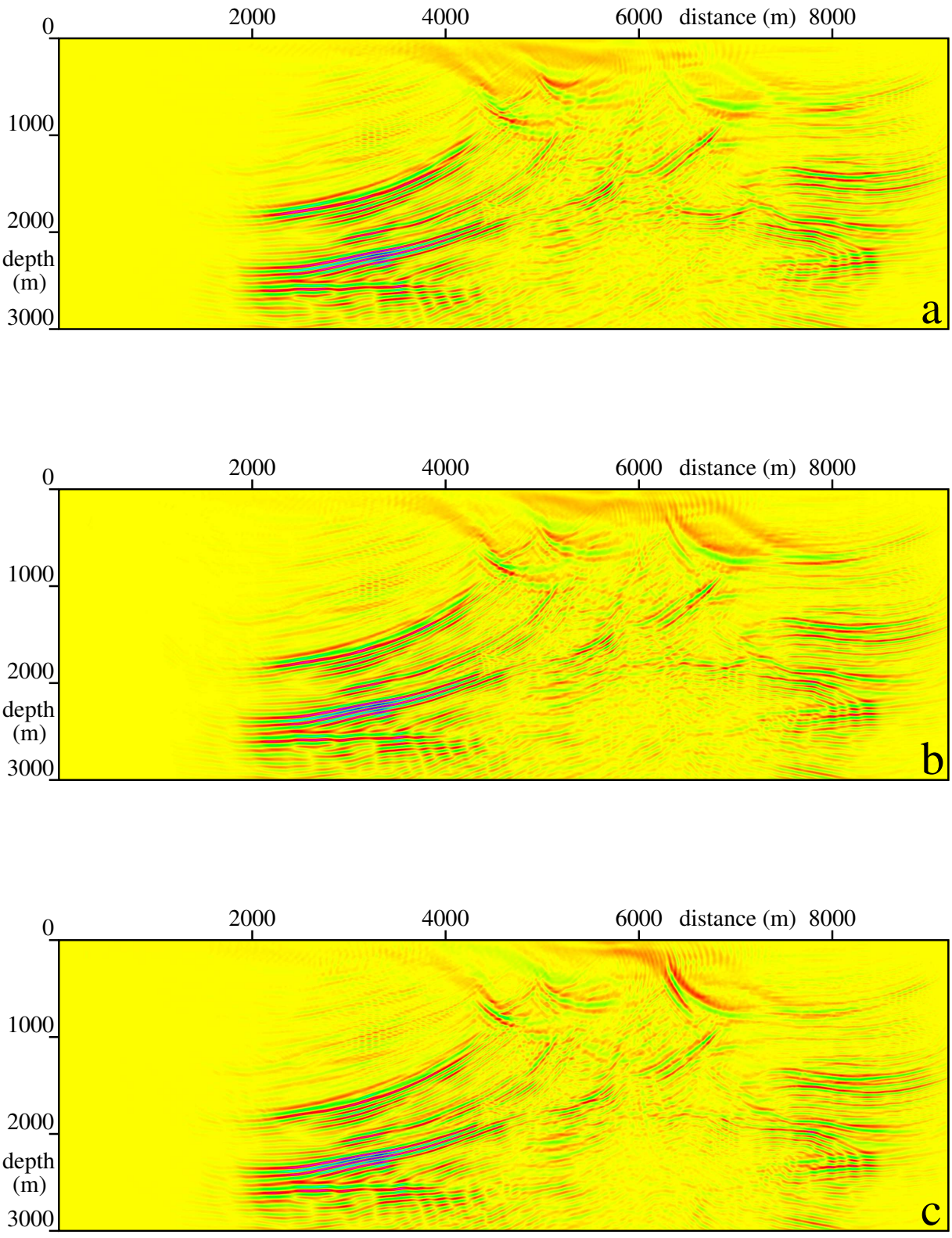

Figure 7. Stacked migrated sections of the Marmousi data set calculated in the velocity models with the average Lyapunov exponents λ of (a) 0.52 s^{-1}, see Figure 2a, (b) 0.68 s^{-1}, see Figure 2b, and (c) 0.89 s^{-1}, see Figure 2c. We used a sparse grid of cells of 20×20 metres and applied no post-processing.

A proper choice of the discretization of the phase-space is of key importance in the decomposition of the wave field into Gaussian packets. Using the method proposed by Žáček (2003, 2005), we obtain $\Delta x_{\mathrm{R}} = 82.1$ m, $\Delta t_{\mathrm{R}} = 0.272$ s, $\Delta p = 0.404 \times 10^{-4}$ sm^{-1} and $\Delta\omega = 3.70$ s^{-1}. We decomposed the original Marmousi data and did not apply any pre-processing.

The stacked migrated sections of the Marmousi data set are displayed in Figure 7. They correspond to three different smooth models, see Figure 2, and represent the raw output of our method. In order to speed the calculation up, we used a sparser grid of cells of 20×20 metres.

Although we can recognize several features of the Marmousi model, we failed in reconstructing the bottom part of the model. We believe three main causes of such result consist in the use of (a) *uniform* Gaussian packets, (b) an oversimplified stacking procedure, and (c) considerably smoothed velocity models, see Versteeg (1993). Nevertheless, we are still intensively testing our method, and therefore the numerical examples shown in Figure 7 should be considered as very preliminary.

In the future, we would like to refine the method of stacking in order to improve the quality of stacked migrated sections. Moreover, we should consider the application of Gaussian packets with varying parameter N^0. This would make the Gaussian packet prestack depth migration more expensive, but also more suitable for complex velocity models.

7 Conclusions

The Gaussian packet prestack depth migration represents a new and promising imaging method operating in the common-shot domain. It can handle multi-valued travel times and allows a target-oriented approach.

Although the common-shot Gaussian packet migration cannot match the efficiency of the common-offset Gaussian beam migration algorithm proposed by Hill (2001), it is more general and provides a one-to-one relation between the Gabor functions from the common-shot gather and their localized images from the depth section. We hope that this unique feature could help us in understanding the true meaning of the migrated section.

Ultimately, we would like to inspire and encourage other researchers and possibly attract their attention to the Gaussian packet domain. We believe that it could show even greater potential than the well-known Gaussian beam domain.

Acknowledgements

The author is greatly indebted to Luděk 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. & Ulin, V. V. (1981): Complex space time ray method and 'quasiphotons'. Zap. Nauchn. Sem. Leningr. Otd. Math. Inst., **117**, 5–12 (in Russian); English transl.: J.Sov. Math., **24**, 269–273, 1984.

Bulant, P. & Klimeš, L. (1999): Interpolation of ray-theory travel times within ray cells. Geophys. J. Int., **139**, 273–282.

Červený, V. (2001): Seismic Ray Theory. Cambridge Univ. Press, Cambridge.

Hill, N. R. (1990): Gaussian beam migration. Geophysics, **55**, 1416–1428.

Hill, N. R. (2001): Prestack Gaussian-beam depth migration. Geophysics, **66**, 1240–1250.

Klimeš, L. (1989a): Gaussian packets in the computation of seismic wavefields. Geophys. J. Int., **99**, 421–433.

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

Klimeš, L. (2002): Lyapunov exponents for 2-D ray tracing without interfaces. Pure and Appl. Geophys., **159**, 1465–1485.

Klimeš, L. (2004): Gaussian packets in smooth isotropic media. In: Seismic Waves in Complex 3-D Structures, Report 14, Charles University, Prague, 43–54.

Ralston, J. (1983): Gaussian beams and the propagation of singularities. MAA Studies in Mathematicc., **23**, 206–248.

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.

Versteeg, R. J. (1993): Sensitivity of prestack depth migration to the velocity model. Geophysics, **58**, 873–882.

Žáček, K. (2001a): Optimization of the shape of Gaussian beams. 71st Ann. Int. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 2128–2131.

Žáček, K. (2001b): Optimization of the shape of Gaussian beams. In: Seismic Waves in Complex 3-D Structures, Report 11, Charles University, Prague, 181–201.

Žáček, K. (2002): Smoothing the Marmousi model. Pure and Appl. Geophys., **159**, 1507–1526.

Žáček, K. (2003): Decomposition of the wave field into optimized Gaussian packets. 73st Ann. Int. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1869–1872.

Žáček, K. & Klimeš, L. (2003): Sensitivity of seismic waves to the structure. 73rd Ann. Int. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1857–1860.

Žáček, K. (2004): Gaussian packet pre-stack depth migration. 74th Ann. Int. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 957–960.

Žáček, K. (2005): Decomposition of the wave field into optimized Gaussian packets. In: Seismic Waves in Complex 3-D Structures, Report 15, Charles University, Prague, 17–27.