Smoothing the Marmousi model

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

Summary

The 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 & 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 Gaussian packet method (e.g., Klimeš, 1989), we do not need to find two-point rays, but 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 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 already developed and published. The authors have been interested in determining the most suitable physical

Seismic Waves in Complex 3-D Structures, Report 11 (Department of Geophysics, Charles University, Prague, 2000), pp. 161–179; to appear in PAGEOPH, vol. 159, 2002

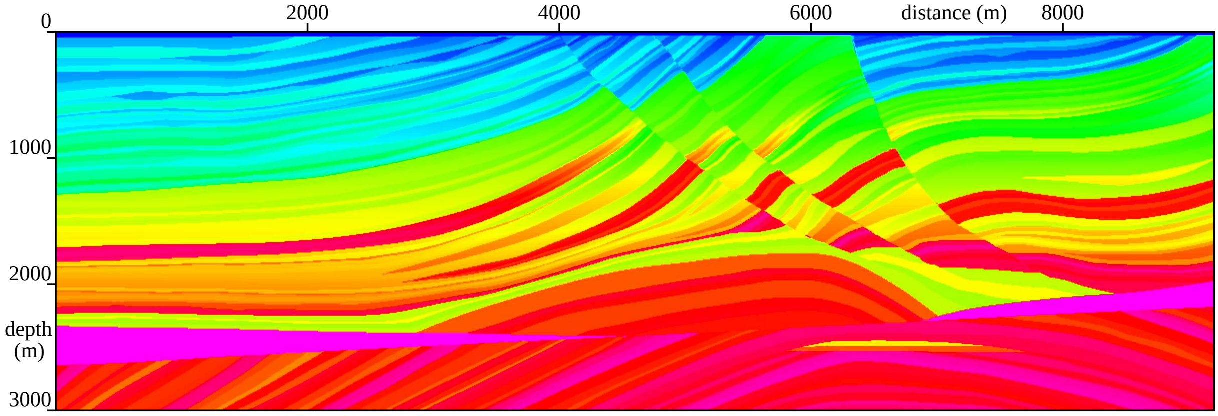


Figure 1. The Marmousi model.

quantity to be smoothed (Müller & Shapiro, 2000; Gold et. al., 2000), in finding a way to smooth the velocity model (Grubb & Walden, 1995; Versteeg, 1991; Brac & 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 & 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 $4{\times}4$ 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 & 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 has to 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 small 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, 2002). Since 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 sub-grid 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 sub-grid 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.

Let us 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

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

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 weighing parameters of the grid points, b_{ijkl} 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 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 $B_\alpha(\mathbf{x})$

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

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

Equation (1) now reads

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

where

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

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 non-linear inverse problem numerically, we need to "linearize" formula (4). The linearization of (4) yields

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

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

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

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

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 $(\sigma^{\text{GRID}})^2$, see equation (1). N is the number of grid points.

The condition for the minimum of the objective function is

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

which yields

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

The resulting vector of the model parameters is

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

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

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

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, 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 ν 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,

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

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

$$\nu = \nu_S \nu_R \quad . \tag{13}$$

For $\tau = \tau_{\mathrm{S}} + \tau_{\mathrm{R}} = 0$, we obtain $e^{\lambda\tau} = 1$. 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

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

We want to work with the "average Lyapunov exponent" $\widehat{\lambda}$. The "average Lyapunov exponent" $\widehat{\lambda}$ 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 $e^{\widehat{\lambda}\tau_{\max}}$ 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, $\tau_{\max} = 2.9\mathrm{s}$ may be used for estimating the optimum value of $\widehat{\lambda}$. Thus, for number of arrivals not exceeding 10, we obtain the optimum value of $\widehat{\lambda}$ close to $0.8\mathrm{s}^{-1}$.

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. Thus, the desired migrated section would be wrong. The maximum halfwidth should probably not be greater than the B-spline interval.

Let us 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 do not have any 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 (Bulant, 2002). The 4×4 matrix $\mathbf{b}'$ is then defined by

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

where $\mathbf{e}$ is a unit vector, $\langle ... \rangle$ 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 $\frac{1}{3}d(d+2)$ 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,
$b'_{1111} = b'_{2222} = 1$,
$b'_{1122} = b'_{1212} = b'_{1221} = b'_{2112} = b'_{2121} = b'_{2211} = \frac{1}{3}$ 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 $751 \times 2301 = 1728051$ 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 linarized inversion as

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

where $\|u\|_{\mathrm{init}}$ 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 $\sigma = \sigma^{\text{GRID}} N^{-\frac{1}{2}}$. Hence, we can alter equation (17) as

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

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

For $\varepsilon = 0.1$, $(u_{\text{A}})^{-1} = 3000$ ms^{-1} and $\Lambda_{\text{init}} = 1.3$ s^{-1}, see equation (A−12), we obtain $s_{\text{init}} = 81529$. We study the values of s of (1) 0 m^2 (without the Sobolev norm included in the inversion), (2) $\sqrt{3/8}\cdot 10000\text{m}^2 \approx 6124\text{m}^2$, (3) $\sqrt{3/8}\cdot 25000\text{m}^2 \approx 15309\text{m}^2$, (4) $\sqrt{3/8}\cdot 50000\text{m}^2 \approx 30619\text{m}^2$, (5) $\sqrt{3/8}\cdot 100000\text{m}^2 \approx 61237\text{m}^2$ and (6) $\sqrt{3/8}\cdot 225000\text{m}^2 \approx 137784\text{m}^2$.

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

Figure 3 shows 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 $s = 61237\text{m}^2$ and $s = 137784\text{m}^2$ 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 $\widehat{\lambda}$ close to 1 (as the model with $s = 6124\text{m}^2$) 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

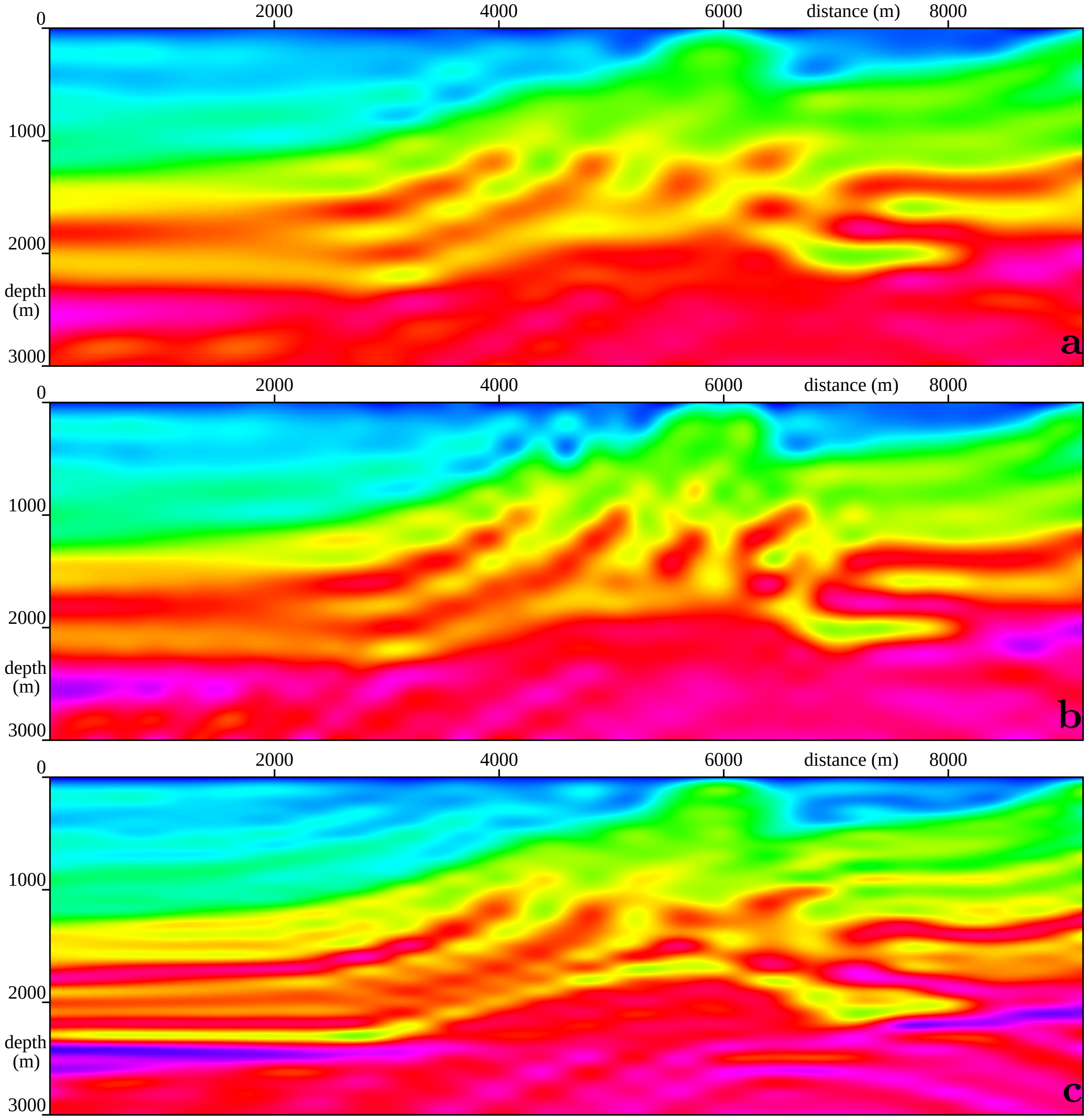
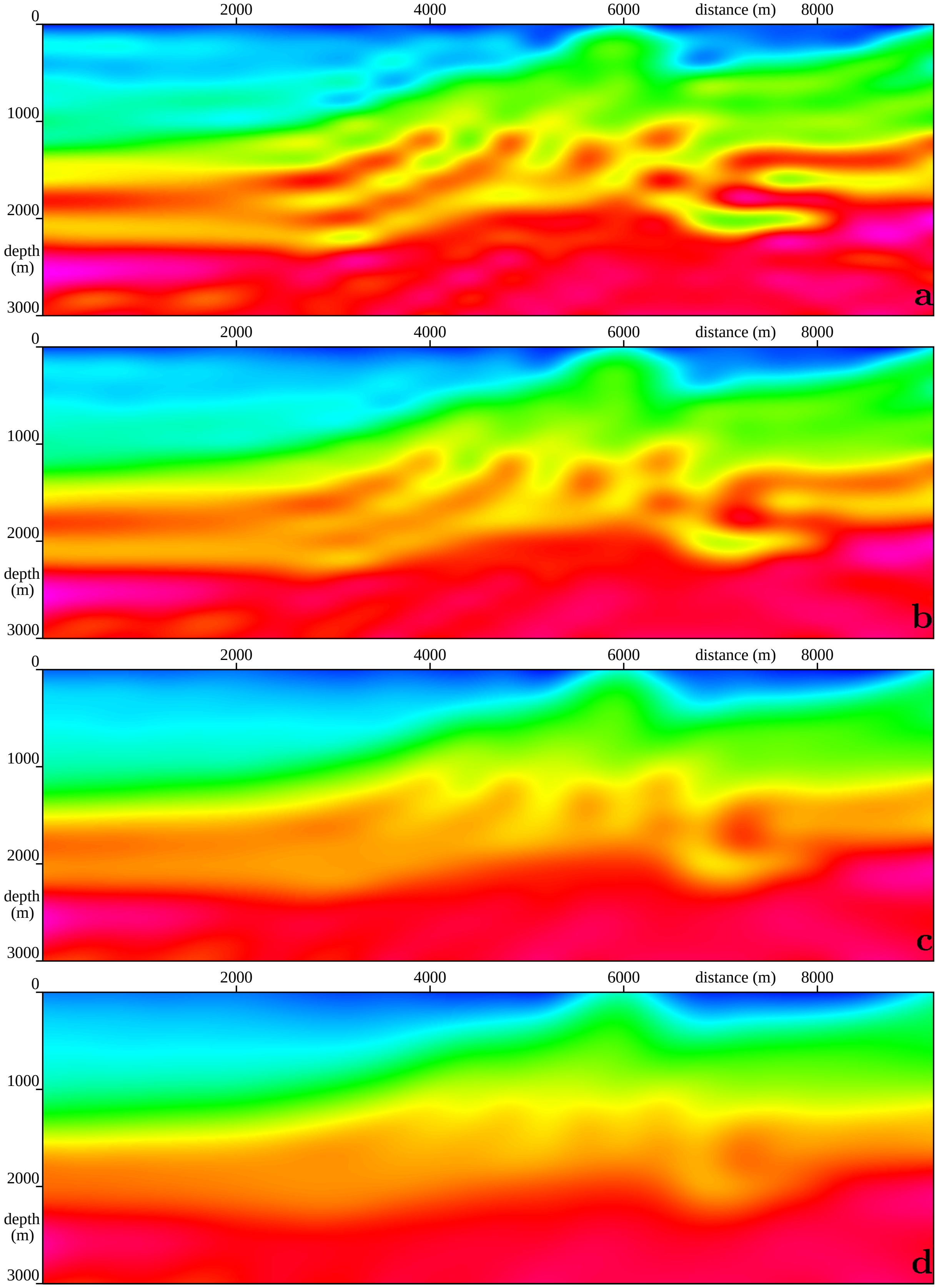


Figure 2. The smoothed models without minimized Sobolev norm of slowness ($s = 0\text{m}^2$) for the B-spline grids of cells of (a) 200×400 metres, (b) 200×230 metres and (c) 100×400 metres.

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 1500ms^{-1} we obtain the farthest possible reflection point at 4350 metres from the watergun and 900 metres from the farthest hydrophone. As 900 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 $s = 0\text{m}^2$ and $s = 6124\text{m}^2$ do not seem to be suitable for ray methods due to the density of


0
2000
4000
6000
distance (m)
8000
1000
2000
depth
(m)
3000
a
0
2000
4000
6000
distance (m)
8000
1000
2000
depth
(m)
3000
b
0
2000
4000
6000
distance (m)
8000
1000
2000
depth
(m)
3000
c
0
2000
4000
6000
distance (m)
8000
1000
2000
depth
(m)
3000
d

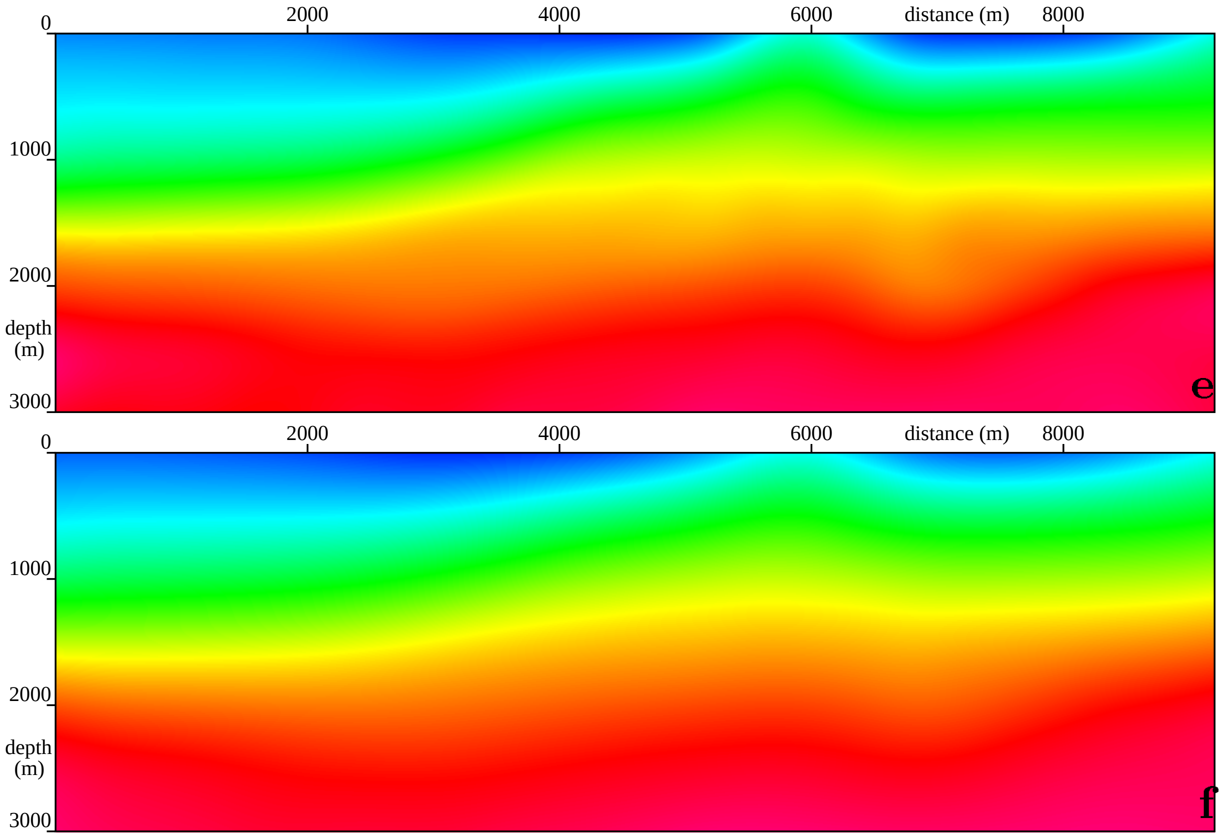


Figure 3. The smoothed models for the B-spline grid of cells of 200×400 metres and for values of parameter s of (a) 0m^2, (b) 6124m^2, (c) 15309m^2, (d) 30619m^2, (e) 61237m^2 and (f) 137784m^2.

caustics. In Figure 5a, we can see rays trapped in the low velocity channels.

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 $s = 0\text{m}^2$,
(b) 18 for the value of $s = 6124\text{m}^2$,
(c) 7 for the value of $s = 15309\text{m}^2$,
(d) 5 for the value of $s = 30619\text{m}^2$,
(e) 3 for the value of $s = 61237\text{m}^2$ and
(f) 2 for the value of $s = 137784\text{m}^2$.
In the model without interfaces, the number of arrivals should be odd. Even numbers for $s = 6124\text{m}^2$ and $s = 137784\text{m}^2$ may be explained by the influence of the borders of the model. Due to the requirements established above, the models with $s = 0\text{m}^2$ and $s = 6124\text{m}^2$ are probably not suitable for ray-based methods.

The relative RMS difference of slowness between all calculated models and the original Marmousi model are 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 is caused by 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.

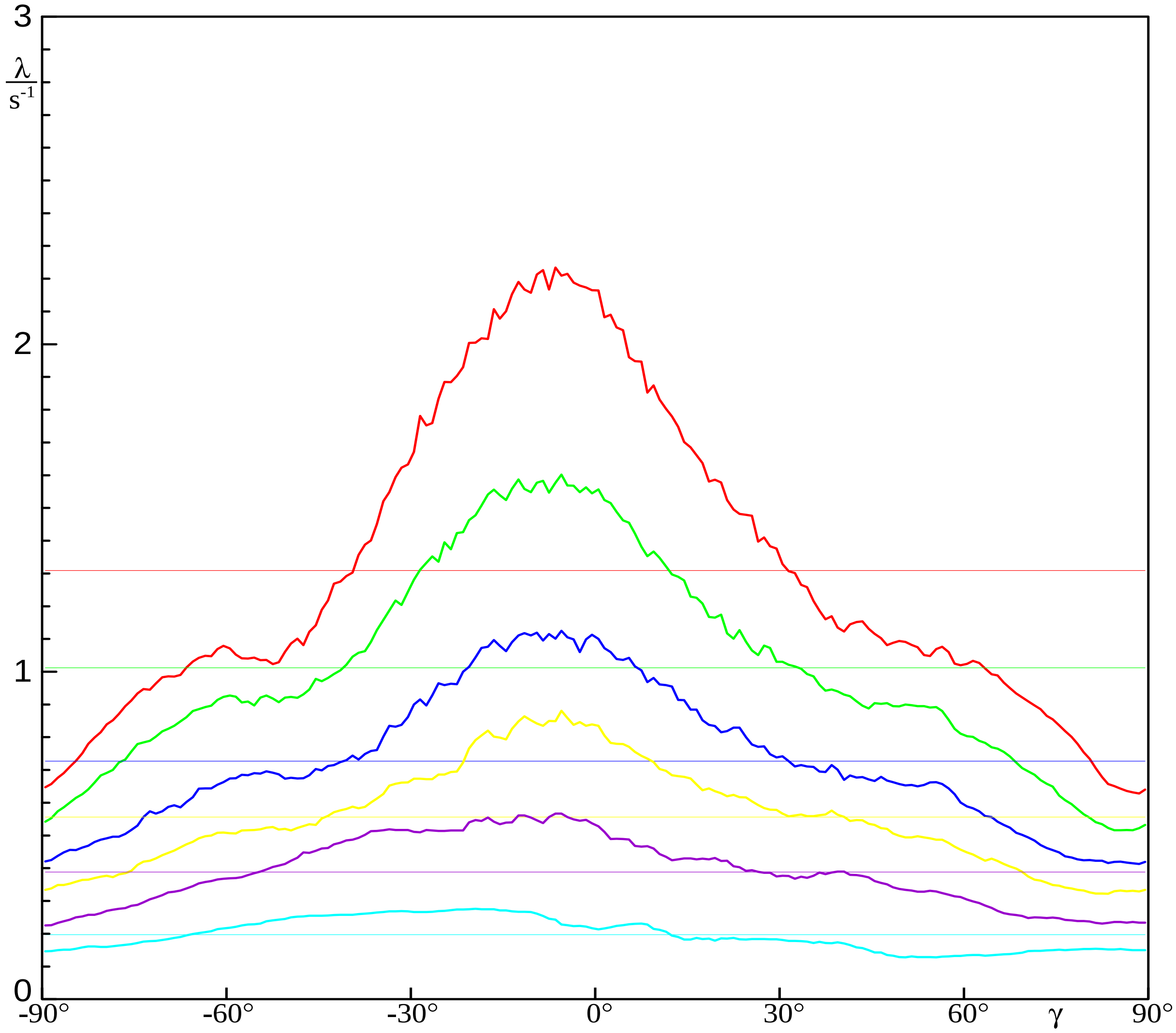
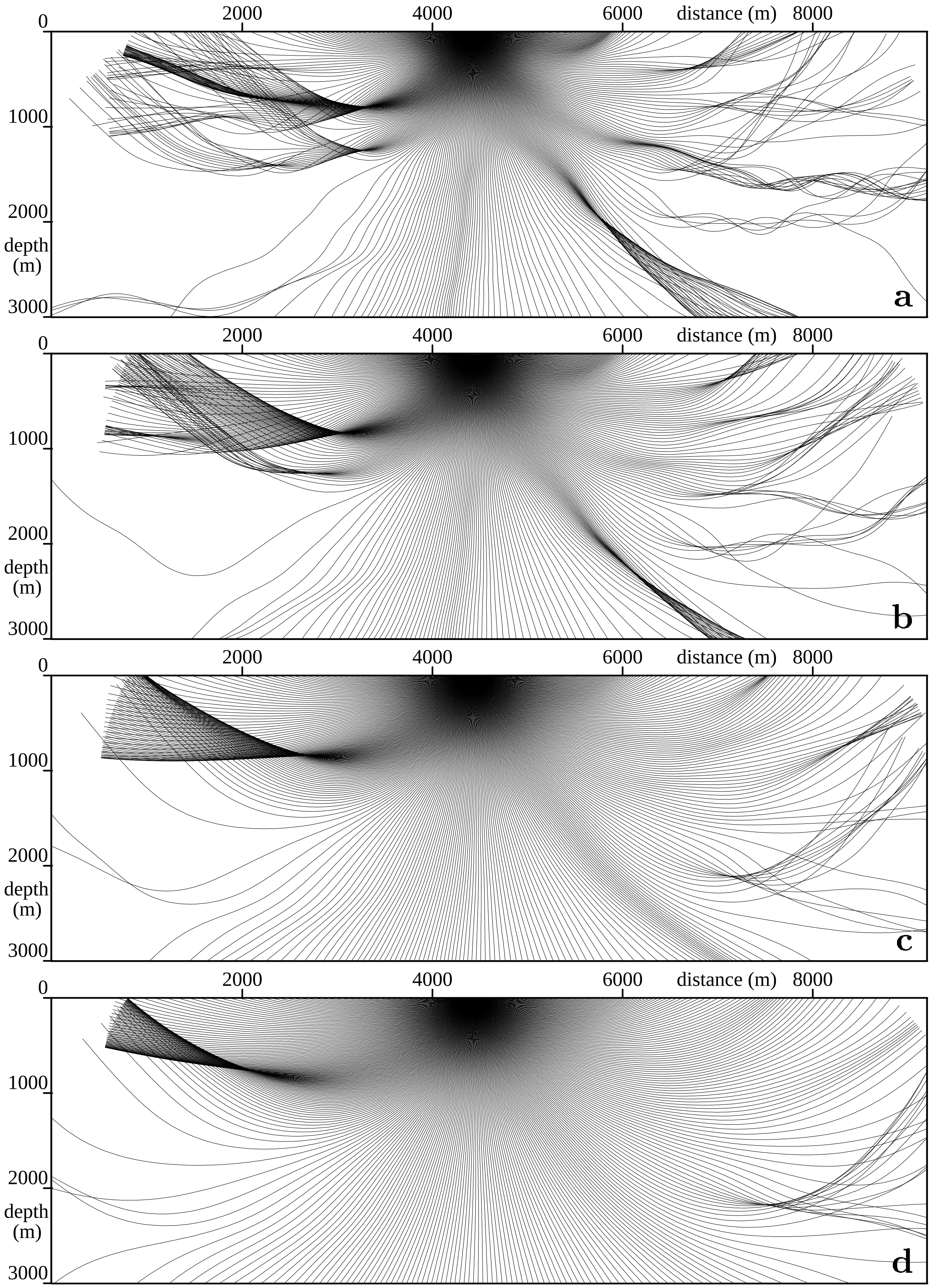


Figure 4. The angular dependence of the Lyapunov exponents for the models with the B-spline grid of cells of 200×400 metres. The left and right border corresponds to the vertical ray, the middle corresponds to the horizontal ray. The values of parameter s are (1) $0\mathrm{m}^2$ for *red*, (2) $6124\mathrm{m}^2$ for *green*, (3) $15309\mathrm{m}^2$ for *blue*, (4) $30619\mathrm{m}^2$ for *yellow*, (5) $61237\mathrm{m}^2$ for *magenta* and (6) $137784\mathrm{m}^2$ for *cyan*. 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 a too 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 make a short summary of what we have already done. 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. In other words, we already know, how to smooth the Marmousi model for the computation of rays and travel times.


0
2000
4000
6000
distance (m)
8000
1000
2000
depth
(m)
3000
a
b
c
d

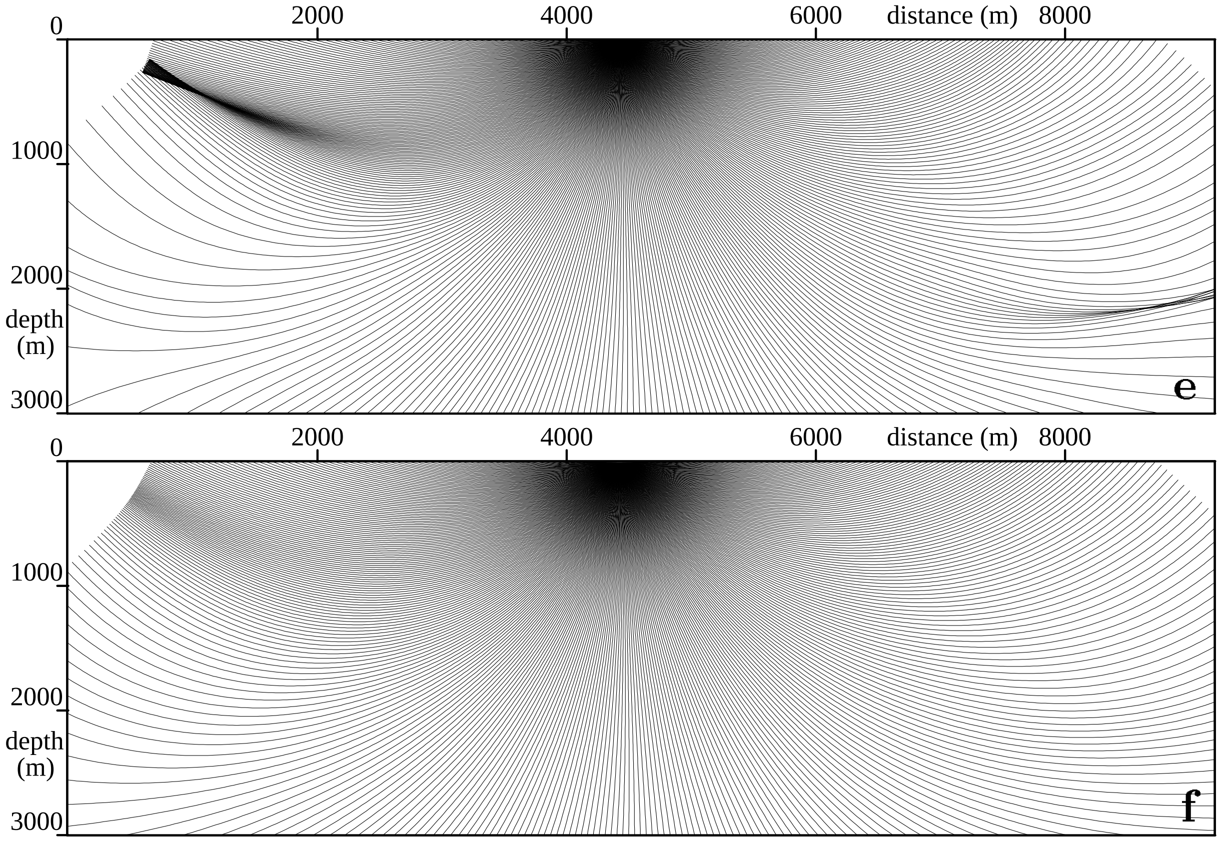


Figure 5. Rays in models with the B-spline grid of cells of 200×400 metres and with values of parameter s of (a) $0\mathrm{m}^2$, (b) $6124\mathrm{m}^2$, (c) $15309\mathrm{m}^2$, (d) $30619\mathrm{m}^2$, (e) $61237\mathrm{m}^2$ and (f) $137784\mathrm{m}^2$. The maximum travel time is of 2.3 seconds.

	100 − 400m	**200 − 230m**	**200 − 400m**
0m²	8.3%	10.8%	11.1%
6124m²	11.4%	11.9%	12.0%
15309m²	13.2%	13.2%	13.3%
30619m²	14.3%	14.3%	14.3%
61237m²	15.1%	15.1%	15.1%
137784m²	15.8%	15.8%	15.8%

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.

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 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

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

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

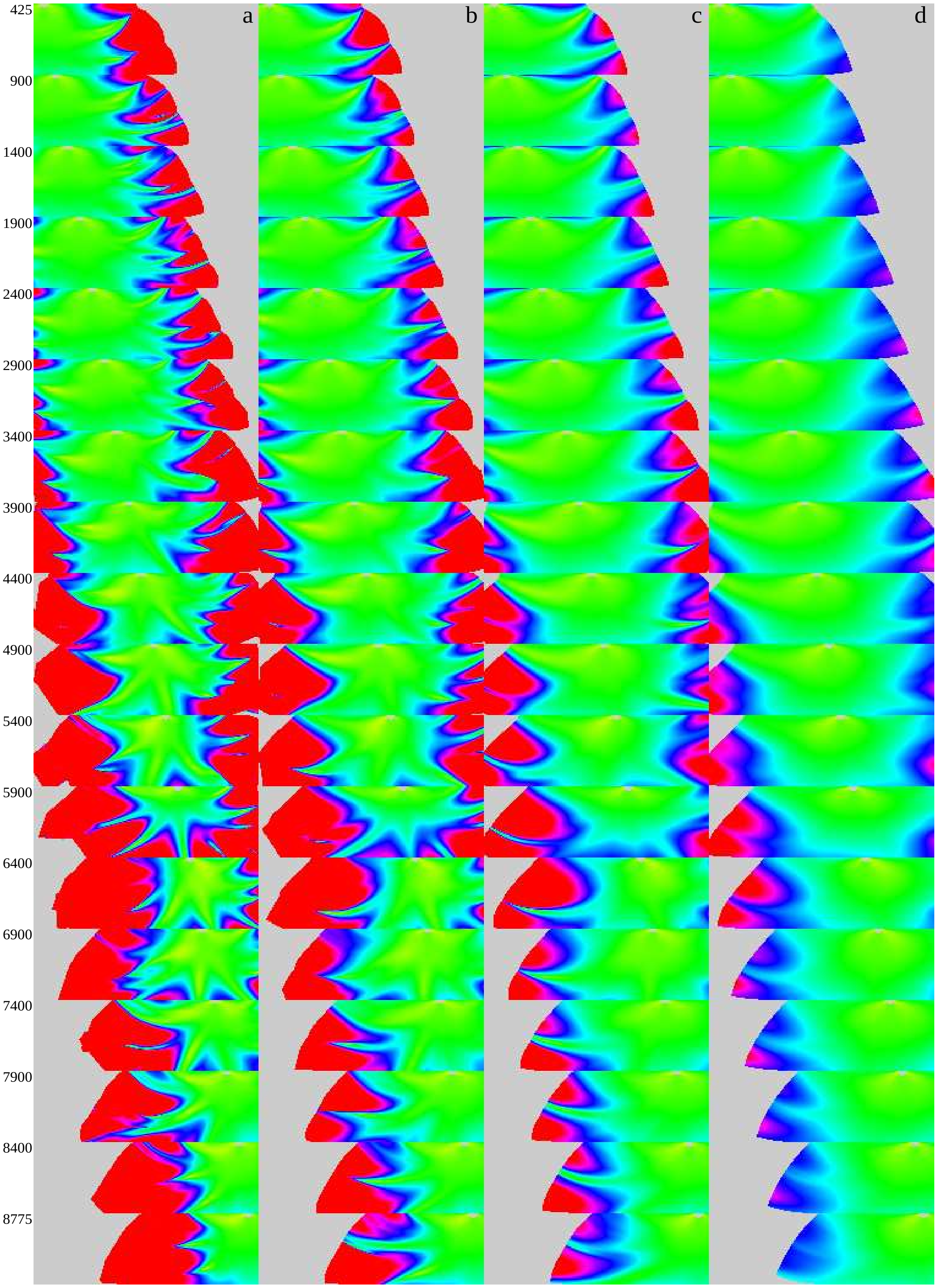


Figure 6. The halfwidths of Gaussian beams in smoothed models with the B-spline grid of cells of 200×400 metres. Columns correspond from left to right to the values of parameter s of (a) $15309\mathrm{m}^2$, (b) $30619\mathrm{m}^2$, (c) $61237\mathrm{m}^2$ and (d) $137784\mathrm{m}^2$. Rows correspond to various positions of the source (in metres).

quadratic term in the Taylor expansion of the complex-valued travel time of the Gaussian beam thus reads

$$\frac{1}{2}Mq^2 \quad . \tag{20}$$

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

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

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 crossection

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

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

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

The halfwidths of Gaussian beams calculated for the models with the B-spline grid of cells of 200×400 metres and with $s = 15309\mathrm{m}^2$, $s = 30619\mathrm{m}^2$, $s = 61237\mathrm{m}^2$ and $s = 137784\mathrm{m}^2$ 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 colour coded quantity W is displayed at the respective points along the central rays of the beams. The yellow colour corresponds to the Gaussian beam halfwidth of 0 metres for all frequencies. The green colour corresponds to the Gaussian beam halfwidth of 202 and 378 metres for the frequencies of 35 Hz and 10 Hz, respectively. The red colour corresponds to the Gaussian beam halfwidths of 1010 metres and more for the frequency of 35Hz, and of 1890 metres and more for the frequency of 10Hz. Thus, the red coloured regions of Figure 6 indicate too wide Gaussian beams for the frequencies under consideration

We can see that the model with $s = 15309\mathrm{m}^2$ 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 $s = 61237\mathrm{m}^2$ and $s = 137784\mathrm{m}^2$ 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 $s = 30619\mathrm{m}^2$ 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. In future, we will try to develop a method to optimize the shapes of Gaussian beams or packets in dependence on the position of the source, or of the receiver in the migration. This would allow the use of models not so smoothed.

Let us 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 caused by the low frequencies under consideration.

Let us summarize that models with parameter s equal to or greater than $15309\mathrm{m}^2$ seem to be suitable for ray tracing with the travel time of 2.3 seconds, see Figure 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 should probably 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.

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 is up to 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 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

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

Addison, P.S. (1997): Fractals and Chaos: an illustrated course. IOP Publishing Ltd, Bristol and Philadelphia.

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

Bulant, P. (2002): Sobolev scalar products in the construction of velocity models — application to model Hess and to SEG/EAGE Salt Model. Pageoph, **159**, in press.

Gold, N., Shapiro, S.A., Bojinsky, S. & Müller, T.M. (2000): An approach to upscaling for seismic waves in statistically isotropic heterogenous elastic media. Geophysics, **65**, 1837–1850.

Grubb, H.J. & Walden, A.T. (1995): Smoothing seismically derived velocities. Geophys. Prosp., **43**, 1061–1082.

Keers, H., Dahlen, F.A. & Nolet,G. (1997): Chaotic ray behaviour in regional seismology. Geophys. J. Int., **131**, pp. 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. (1999): Lyapunov exponents for 2-D ray tracing without interfaces. In: Seismic Waves in Complex 3-D Structures, Report 8, 83–96, Dep. Geophys., Charles Univ., Prague.

Klimeš, L. (2002): Lyapunov exponents for 2-D ray tracing without interfaces. Pageoph, **159**, in press.

Katok, S.R. (1980): The estimation from above for the topological entropy of a diffeomorphism. In: Netecki, Z. & Robinson, C. (eds.): Global Theory of Dynamical Systems. Lecture Notes in Mathematics, Vol. 819, Springer, Berlin, Heidelberg, New York, pp. 258–264.

Lyapunov, A.M. (1949): Problème Général de la Stabilité du Mouvement. Annals of Mathematical Studies, Vol. 17, Princeton Univ. Press.

McCauley, J.L. (1993): Chaos, Dynamics and Fractals: an algorithmic approach to deterministic chaos. Cambridge University Press, Cambridge.

Müller, T.M. & 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 in English translation.

Smith, K.B., Brown, M.G. & Tappert, F.D. (1992): Ray chaos in underwater acoustics. J. Acoust. Soc. Am., **91**, 1939–1949.

Tappert, F.D. & Tang, X. (1996): Ray chaos and eigenrays. J. Acoust. Soc. Am., **99**, 185–195.

Tarantola, A. (1987): Inverse Problem Theory. Elsevier, Amsterdam.

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

Versteeg, R.J. (1991): Analysis of the problem of the velocity model determination for seismic imaging. PhD thesis, University of Paris VII.

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.

Witte, O., Roth, M. & Müller, G. (1996): Ray tracing in random media. Geophys. J. Int., **124**, 159–169.

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" $\widehat{\lambda}$ may be approximated by

$$\widehat{\lambda} \approx \Lambda + \Delta\Phi \quad , \tag{A–1}$$

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

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

where $\mathrm{neg}(f) = \frac{1}{2}(f - |f|)$ 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,

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

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

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

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

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

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

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

and

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

where

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

Finally, we arrive at

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

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

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

As we need to find some initial value of parameter s, we should estimate the respective value of the Sobolev norm $\|u\|_{\text{init}}$. 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),

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

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

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