Three-dimensional two-point paraxial ray tracing problem in the presence of caustics

Wilson M. Figueiró and Raúl I. Madariaga
CPGG-UFBA - DTAO-ENS

Abstract

We propose an approach that allows us to continue, in a fast and accurate way, the iterative process defined by the paraxial ray tracing method when a caustic is found in the observation plane by the reference ray considering a three-dimensional velocity model. Seismic tomography methods depends, strongly, on the quality (quickness and accuracy) of ray tracing methods and, one of them, the called paraxial, is very useful and efficient for tomographic aims; but, the presence of caustics can damage, in a dramatic way, the offering of rays by this iterative method. In order to overcome this problem, we choose a three-dimensional velocity model that permits us to get analytical formulas of the paraxial ray tracing elements and, in addition, be used as a good approximation for more complicated media in a vicinity of the reference ray. The behavior of a property related to the matrix paraxial operator, such as condition number is shown graphically by a three-dimensional surface. A geometrical strategy (that can be applied for more complicated models by means of a sequence of local approximations using models that admit analytical treatments) is built in order to avoid the effect of the presence of caustic points when the reference ray arrives in one of them. Then, the desired convergence to the receiver point is obtained, almost always, in only one iteration.

INTRODUCTION

Seismic tomography requires efficient seismic ray tracing. One of the most common ways of solving the two point ray tracing problem is the so-called paraxial ray method (Popov and Pšenčík, 1978; Popov, 1982; Červený, Popov and Pšenčík, 1982). Any relatively complicated geological model has a number of caustic surfaces where paraxial ray tracing does not work well when the arrival point of the reference ray is close to one of them. The purpose of this work is to present an approach that allows us to solve this important problem.
Paraxial ray tracing is a shooting method consisting in an iterative process that, at each iteration computes a perturbation of the initial slowness vector at the source of the current (reference or central) ray using a first-order expansion perturbation of the ray-tracing equations. This positional perturbation, in the two-point ray-tracing problem, is the difference between the arrival point of the central ray and the receiver. The obtained slowness vector perturbation allows us to trace a new ray, that will be considered a new reference ray in the iterative process. If an analytical approach is desired, the mathematical aspects of this procedure can become complicated for more realistic models, but it is possible to find a family of seismic velocity models for which all calculations can be done analytically. For a numerical study of more complicated models, a geometrical strategy to avoid caustics effects can be applied too.
With a three-dimensional seismic velocity model, that describes a linear variation of the square slowness with position, it is possible to make experiments that show a very good performance of the paraxial ray tracing method when the rays arrival points are far from caustics during the iterative process. But, instabilities in the paraxial matrix operator arise when caustics or quasi-caustics are found by the ray in the observation surface. In this case, it is not possible to continue searching the receiver.
The approach developed here consists in a geometrical scheme that makes an extension of the central ray and generates news equations using an external point for the observation plane.
When the determinant of the paraxial operator is close to zero the approach developed can be used, in order to arrive to the receiver point. This is done in a single iteration. The same result is observed even when the determinant is identically equal to zero.

PARAXIAL RAY TRACING METHOD

A seismic source point $\mathbf{S}$ in a $3D$ isotropic velocity model M defines a travel-time field $T(\mathbf{x})$, $\mathbf{x} = (x_1, x_2, x_3)$, on it. By means of the $3D$ wave equation, it is possible to get the eikonal equation:

$$\sum_{i=1}^{3} \left(\frac{\partial T}{\partial x_i}\right)^2 = \frac{1}{v^2(\mathbf{x})}, \tag{1}$$

where $v(\mathbf{x})$ is the wave velocity function. The gradient $\vec{\nabla} T$ is perpendicular to the wave fronts ($T(\mathbf{x})$ = constant); this gradient is called slowness vector and represented by $\mathbf{p} = (p_1, p_2, p_3)$. The rays are admitted to be tangent to $\mathbf{p}$ at every point of M. Then,

$$\mathbf{p}(\mathbf{x}(\ell)) = \frac{1}{v(\mathbf{x}(\ell))}.\frac{d\mathbf{x}(\ell)}{d\ell}, \tag{2}$$

where ℓ is the arc length measured along the ray and $\mathbf{x}(\ell)$ is its corresponding point on the ray. Consequently, $||\mathbf{p}(\ell)|| = u(\ell)$, where $u = 1/v$ is the slowness. As suggested by Burridge (1976) we can define the Hamiltonian

$$H(\mathbf{x}, \mathbf{p}) = \frac{1}{2}[||\mathbf{p}||^2 - u^2(\mathbf{x})] \tag{3}$$

which is equal to zero along ray.
Let us consider a seismic ray that originates in $\mathbf{S}$ and is described by the pair $\mathbf{y}_o(\tau) = (\mathbf{x}_o(\tau), \mathbf{p}_o(\tau))$ in a $3D$ velocity model, where $\mathbf{p}_o(\tau) = (p_{o1}(\tau), p_{o2}(\tau), p_{o3}(\tau))$ represents the slowness vector tangent to this ray in its point

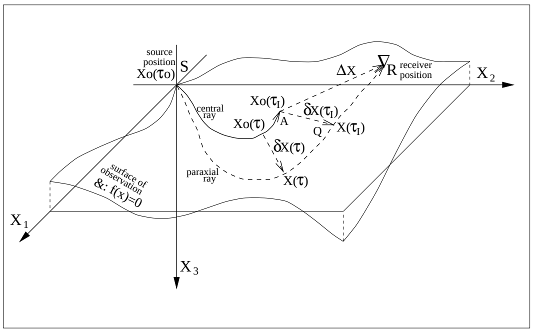

Figure 1: A three-dimensional schematic view of the problem.

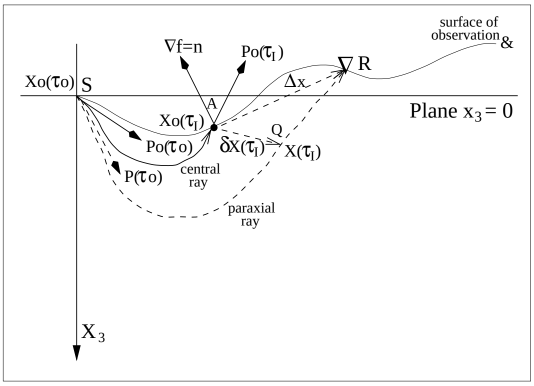

Figure 2: A two-dimensional view of some elements of the paraxial ray tracing scheme.

$\mathbf{x}_o(\tau) = (x_{o1}(\tau), x_{o2}(\tau), x_{o3}(\tau))$ and τ is a ray parameter defined by $\int_0^\ell v(\ell)d\ell$ (Thomson and Chapman, 1985). We want to compute perturbations $\delta\mathbf{x}(\tau) = (\delta x_1(\tau), \delta x_2(\tau), \delta x_3(\tau))$ and $\delta\mathbf{p}(\tau) = (\delta p_1(\tau), \delta p_2(\tau), \delta p_3(\tau))$ of the central ray $\mathbf{y}_o(\tau)$. The perturbed ray $\mathbf{y}(\tau) = (\mathbf{x}(\tau), \mathbf{p}(\tau))$ will be called a paraxial ray; and is given by:

$$\begin{cases} \mathbf{x}(\tau) = \mathbf{x}_o(\tau) + \delta\mathbf{x}(\tau) \\ \mathbf{p}(\tau) = \mathbf{p}_o(\tau) + \delta\mathbf{p}(\tau). \end{cases} \quad (4)$$

The ray tracing equations (Červený, 1987) are given by:

$$\begin{cases} \frac{d\mathbf{x}}{d\tau} = \vec{\nabla}_p H = \mathbf{p} \\ \frac{d\mathbf{p}}{d\tau} = -\vec{\nabla}_x H = \frac{1}{2}\vec{\nabla}_x(u^2(\mathbf{x})), \end{cases} \quad (5)$$

where $\vec{\nabla}_x$ and $\vec{\nabla}_p$ are the gradients with respect to the vectors $\mathbf{x}$ and $\mathbf{p}$, respectively.

The first order Taylor expansion of $\vec{\nabla}_x H$ around $\mathbf{x}_o$ combined with systems (4) and (5) produces the paraxial ray tracing equations:

$$\begin{bmatrix} \frac{d\delta\mathbf{x}}{d\tau} \\ \frac{d\delta\mathbf{p}}{d\tau} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{U} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \delta\mathbf{x} \\ \delta\mathbf{p} \end{bmatrix}, \quad (6)$$

where $\mathbf{I}$ is the identity matrix of order 3 and $U_{ij} = \frac{1}{2}\frac{\partial^2 u^2(\mathbf{x}_o)}{\partial x_i \partial x_j}$; with $i, j \in \{1, 2, 3\}$. Calling $\mathbf{\Lambda}$ the 6×6 matrix of equation (6), we have a more synthetic expression:

$$\frac{d\delta\mathbf{y}(\tau)}{d\tau} = \mathbf{\Lambda}\delta\mathbf{y}(\tau). \quad (7)$$

The formal solution of (7) (Aki & Richards, 1980) is:

$$\delta\mathbf{y}(\tau) = \mathbf{P}(\tau, \tau_o).\delta\mathbf{y}(\tau_o), \quad (8)$$

where τ_o is the initial value of τ and

$$\mathbf{P}(\tau, \tau_o) = \mathbf{I} + \sum_{j=1}^{\infty} \frac{1}{j!}\left(\int_{\tau_o}^{\tau} \mathbf{\Lambda}(\tau_1)d\tau_1\right)^j = \exp\left[\int_{\tau_o}^{\tau} \mathbf{\Lambda}(\tau_1)d\tau_1\right] \quad (9)$$

is known as the propagator matrix. Denoting $\int_{\tau_o}^{\tau} \mathbf{U}(\tau_1)d\tau_1$ by $\mathbf{\Upsilon}(\tau, \tau_o) = \mathbf{\Upsilon}$, we can write:

$$\int_{\tau_o}^{\tau} \mathbf{\Lambda}(\tau_1)d\tau_1 = \begin{bmatrix} \mathbf{0} & (\tau - \tau_o)\mathbf{I} \\ \mathbf{\Upsilon}(\tau, \tau_o) & \mathbf{0} \end{bmatrix}. \quad (10)$$

Then, $\mathbf{P}$ can be written in the matrix form:

$$\mathbf{P}(\tau, \tau_o) = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \\ \mathbf{P}_1 & \mathbf{P}_2 \end{bmatrix}, \quad (11)$$

where

$$\mathbf{Q}_2 = (\tau - \tau_o)\mathbf{I} + \sum_{j=1}^{\infty} \frac{1}{(2j+1)!}(\tau - \tau_o)^{j+1}\mathbf{\Upsilon}^j. \quad (12)$$

Figure 1 shows the $3D$ geometry of the problem, where the central ray, $\mathbf{x}_o$, arrives at the surface of observation, & ($f(\mathbf{x}) = 0$), at $\tau = \tau_I$ at the point $\mathbf{x}_o(\tau_I)$. In general, the paraxial ray, $\mathbf{x}$, for this value of $\tau = \tau_I$ will not belong to &. It is necessary to establish a relation between $\delta\mathbf{y}(\tau_I) = (\delta\mathbf{x}(\tau_I), \delta\mathbf{p}(\tau_I))$ and $\Delta\mathbf{y} = (\Delta\mathbf{x}, \Delta\mathbf{p})$, where $\Delta\mathbf{x} = \mathbf{R} - \mathbf{x}_o(\tau_I) = (\Delta x_1, \Delta x_2, \Delta x_3)$ and $\Delta\mathbf{p} = \mathbf{p}(\tau_R) - \mathbf{p}_o(\tau_I) = (\Delta p_1, \Delta p_2, \Delta p_3)$.

Regarding figure 2, we observe that the scalar products $< \mathbf{n}|\Delta\mathbf{x} >$ and $< \mathbf{n}|\mathbf{p}_o(\tau_I) >$ are known and $\mathbf{p}_o(\tau_I) = (p_{o1_I}, p_{o2_I}, p_{o3_I})$ is tangent to the central ray at $\mathbf{x}_o(\tau_I)$. Similarly, $\mathbf{p}_o(\tau_o) = (p_{o1_o}, p_{o2_o}, p_{o3_o})$ and $\mathbf{p}(\tau_o) = (p_{1o}, p_{2o}, p_{3o})$ are tangent at $\mathbf{S}$ to the central and paraxial rays, respectively. Calling $q_{2_{ij}}$ each entry of $\mathbf{Q}_2$, we have:

$$\begin{bmatrix} \Delta x_1 \\ \Delta x_2 \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \begin{bmatrix} \delta p_1(\tau_o) \\ \delta p_2(\tau_o) \end{bmatrix}, \quad (13)$$

where

$$\alpha_{ij} = q_{2_{ij}} - \frac{p_{oi_I}}{p_{o3_I}}.q_{2_{3j}} - \frac{p_{oj_o}}{p_{o3_o}}(q_{2_{i3}} - \frac{p_{oi_I}}{p_{o3_I}}.q_{2_{33}}). \quad (14)$$

The problem now consists in solving (13) to find $\delta\mathbf{p}(\tau_o)$ that produces a new $\mathbf{p}(\tau_o)$ and, consequently, a new central ray for the iterative process.

Calling $\mathbf{A}$ the 2×2 matrix operator of the system (13), our problem is to study the behavior of $\mathbf{A}$ when the ray arrives at a caustic point located in & and, in this case, to find alternatives in order to continue the iterative procedure in a stable way.

STUDY OF A PARTICULAR VELOCITY MODEL

Let us consider the squared slowness function defined as follows:

$$u^2(\mathbf{x}) = \frac{1}{v^2(\mathbf{x})} = a + bx_1 + cx_2 + dx_3, \quad (15)$$

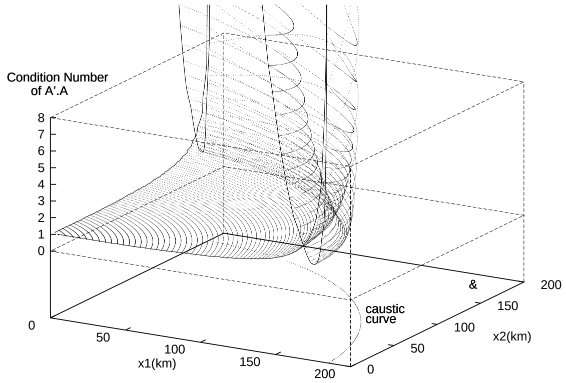

Figure 3: Representation of the condition number (CN) of $\mathbf{A}^T\mathbf{A}$ as a function of $\mathbf{p}_o(0)$, with spherical coordinates θ and φ varying in the interval $(0, \frac{\pi}{2})$.

where a, b, c and d are real numbers. For this model, it is easy to see that the matrix $\mathbf{\Upsilon}$ is equal to zero.
For model (15), when $\tau_I \neq 0$, the following statements are equivalent:

- $i.$ $(x_{o1}(\tau_I), x_{o2}(\tau_I), 0)$ is a caustic,
- $ii.$ $<\mathbf{p}_o(0)|\mathbf{p}_o(\tau_I)>= 0$,
- $iii.$ $det(\mathbf{A}) = 0$,
- $iv.$ minimum eigenvalue of $\mathbf{A}^T\mathbf{A}$ (λ_{min}) equal to zero,
- $v.$ $CN(\mathbf{A}^T\mathbf{A}) = +\infty$,
- $vi.$ $E_1 = \dfrac{det(\mathbf{A})}{tr^2(\mathbf{A})} = 0$ and
- $vii.$ $E_2 = \dfrac{det(\mathbf{A})}{tr(\mathbf{A}^T\mathbf{A})} = 0.$

Where $\mathbf{A}$ is calculated at τ_I and CN is the condition number.
To perform computations, model (15) is taken with $a = 6.25 \times 10^{-2}$ s^{2}km^{-2}, $b = -5.0 \times 10^{-5}$ s^2 km^{-1}, $c = -6.0 \times 10^{-5}$ s^2 km^{-1} and $d = -6.2 \times 10^{-4}$ s^{2}km^{-1}.
This model is certainly simple, but not completely unrealistic. Down to 40 km it gives a good approximation to acceptable seismic velocity distribution without discontinuities in the Earth.
The ray field is built in a continuous way as a function of the take off angles θ and φ. Let us define a curve γ contained in a spherical surface of center $\mathbf{S}$ and radius $\sqrt{a}$ such that the caustic curve in $\&$ is its image by the transformation that for each $\mathbf{p}_o(0)$, gives the exit point $(x_1, x_2, 0)$ of the ray with initial slowness $\mathbf{p}_o(0)$. For a fixed θ, φ_c is the φ for which $\mathbf{p}_o(0)$ is on γ.
Figure 3 shows the ratio $\frac{\lambda_{max}}{\lambda_{min}}$, the condition number $CN(\mathbf{A}^T\mathbf{A})$. For φ near $\frac{\pi}{2}$ ($\mathbf{p}_o(0)$ almost horizontal) CN tends to $+\infty$. As φ decreases to φ_c, CN decreases to 1, but, near and before φ reaches φ_c, CN grows to $+\infty$ on the caustic. For $\varphi < \varphi_c$, there is a strong fall of CN from $+\infty$ to almost 1, that value is kept until φ be close to 0. CN is adimensional.
Graphics of CN are useful to show where the system (13) is unstable (ill-conditioned, CN assumes high values). When such regions of instabilities are detected, it is necessary to have methods to overcame difficulties to solve (13) with a singular or quasi-singular matrix $\mathbf{A}$

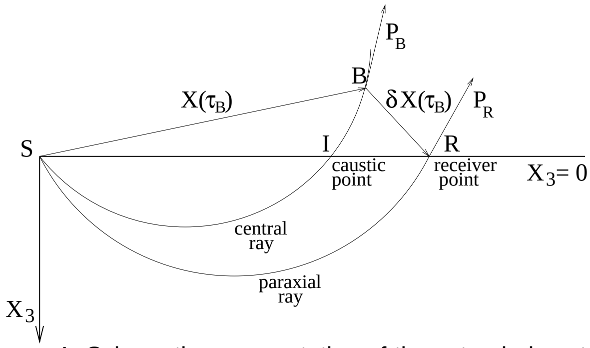

Figure 4: Schematic representation of the extended central ray method.

A METHOD TO AVOID THE CAUSTIC ARRIVAL POINTS

Considering figure 4, where the central ray arrives at the caustic point I, we observe that:

$$\mathbf{R} = \mathbf{x}(\tau_B) + \delta\mathbf{x}(\tau_B), \tag{16}$$

where B is a point situated on the extended central ray and τ_B is τ at this point.
Combining equation

$$\mathbf{p}_R = \mathbf{p}_B + \delta\mathbf{p}(\tau_B), \tag{17}$$

where $\mathbf{p}_R$ and $\mathbf{p}_B$ are the slowness vector at $\mathbf{R}$ and B, respectively, with the eikonal calculated at $\mathbf{R}$, produces:

$$\langle \mathbf{p}_B|\mathbf{p}_B\rangle + 2\langle \mathbf{p}_B|\delta\mathbf{p}(\tau_B)\rangle + \langle \delta\mathbf{p}(\tau_B)|\delta\mathbf{p}(\tau_B)\rangle = \frac{1}{v^2(\mathbf{R})}. \tag{18}$$

Solving the system (6) for model (15), we get:

$$\delta\mathbf{x}(\tau) = (k_1, k_2, k_3)\tau \tag{19}$$

and

$$\delta\mathbf{p}(\tau) = (k_1, k_2, k_3). \tag{20}$$

Then, we get the following system:

$$\left\{ \begin{array}{l} (b/4)\tau_B^2 + (p_{o_{1o}} + k_1)\tau_B = x_{1_R} \\ (c/4)\tau_B^2 + (p_{o_{2o}} + k_2)\tau_B = x_{2_R} \\ (d/4)\tau_B^2 + (p_{o_{3o}} + k_3)\tau_B = 0 \\ ((b/2)\tau_B + p_{o_{1o}} + k_1)^2 + \\ + ((c/2)\tau_B + p_{o_{2o}} + k_2)^2 + \\ + ((d/2)\tau_B + p_{o_{3o}} + k_3)^2 = \frac{1}{v^2(\mathbf{R})}, \end{array} \right. \tag{21}$$

where k_1, k_2, k_3 and τ_B are unknown real numbers. The perturbation $\delta\mathbf{p}(0)$ is given by (20), using the solution of (21). The basic idea of this method is to avoid the caustic region of the ray field at $\&$, moving out of the model to a region with a lower density of rays, by an extension of the central ray. Thus we move the point on the central ray away from the caustic in order to avoid instabilities of $\mathbf{A}$. From now on, it is called of extended central ray method.

NUMERICAL RESULTS

For the model studied here, we can see that when we choose an initial $\mathbf{p}_o(0)$ that produces a ray arriving near the caustic curve ($det(\mathbf{A}) \to 0$), the first iterations of the iterative process is characterized by alternating points near the caustic curve.

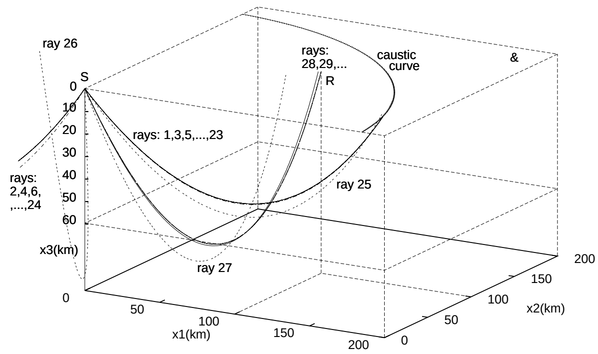

Figure 5: Successive rays computed during the iterative process of the paraxial ray method.

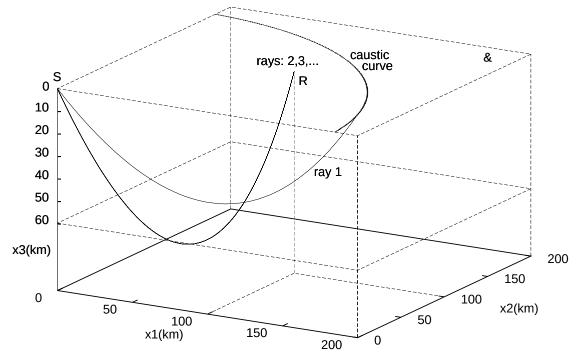

Figure 6: Ray 1 arrives at a quasi-caustic point, in which $det(\mathbf{A}) = 0.55 \times 10^{-9}$ km^{4}s^{-2}. Using the extended central ray tracing method the convergence is obtained in only one iteration.

This experiment, until the 25th iteration is marked by an alternation of arrival points situated very close to the quasi-caustic points $(179.3, 33.8, 0.0)$ and $(-217.5, -41.1, 0.0)$. Figure 5 shows the rays during the process of convergence. Depending on the degree of the proximity of the caustic curve arrival point, the oscillation shown can have a very long duration. This can produce a very poor performance of the paraxial method and even its divergence.
Keeping the same $\mathbf{S}$, $\mathbf{R}$ and $\mathbf{p}_o(0)$ of the last experiment and using the extended central ray method to escape of caustic arrival points, the receiver $\mathbf{R}$ is found in only one iteration as shown in the figure 6.

CONCLUSIONS

The extended central ray method is a way to escape from caustic points of central (current) rays during an iterative process defined by the paraxial ray tracing method. It transfers the work to a point of the central ray where the ray field is less dense, by an extension of the central ray.
The two experiments with this method show that it is able to solve in a fast and accurate way the two point ray tracing problem near caustics where the paraxial method does not work well. The performance of the method developed is very good. Even with the central ray arriving exactly at a caustic point, the receiver point is found in a single iteration.
It is shown that some properties related to the paraxial ray tracing operator (such as: determinant, minimum eigenvalue and condition number) can characterize the vicinity of a caustic on the surface of observation.
The model studied admits an analytical solution that may be not directly applied to treat more complicated models, but it can be useful. As an example, for a large amount of models, it is possible to find approximations, using the model studied, in the vicinity of the central ray, to calculate a slowness vector perturbation, to trace a new central ray in the original model and to repeat the same sequence of procedures until to reach the desired results. In addition, we believe that the results obtained for the particular $3D$ model studied here can be verified for more complicated models using the strategy proposed by the extended central ray method considering other kinds of peculiarities presented by different models.
The method can find applications in all procedures of three-dimensional seismic inversion that use the paraxial ray tracing method to trace seismic rays in a seismic velocity model with caustic points in the surface of observation. Anyway, the extended central ray method, applied to the model (15), can be the base to treat more complicated models and to suggest alternative strategies for ray tracing near caustics.

ACKNOWLEDGMENTS

This work was supported by the CAPES with a post-doctoral scholarship. “Bolsista da CAPES - Brasília-Brasil”. Thanks to the DTAO and “Laboratoire de Géologie” of ENS (Paris-France) and “Departamento de Geologia e Geofísica Aplicada” and CPGG of UFBA (Brazil) by the offer of means and opportunity to accomplish this work. A very special thanks to Maria Fátima de Freitas.

REFERENCES

Aki, K. & Richards, P. G., 1980, Quantitative Seismology: Theory and method, W. H. Freeman and Co., San Francisco.

Burridge, R., 1976, Some Mathematical Topics in Seismology, Courant Institute of Mathematical Sciences, New York University, New York.

Červený, V., 1987, Ray methods for three-dimensional seismic modeling, The Norwegian Institute of Technology.

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.

Popov, M. M., 1982, A new method of computation of wave fields using Gaussian beams, Wave Motion, 4, 85-97.

Popov, M. M. & Pšenčík, I., 1978, Computation of ray amplitudes in inhomogeneous media with curved interfaces, Studia geophys. geod., 22, 248-258.

Thomson, C. J. & Chapman, C. H., 1985, An introduction to Maslov’s asymptotic method, Geophys. J. R. Astron. Soc., 61, 729-746.