GEOPHYSICS, VOL. 64, NO. 1 (JANUARY-FEBRUARY 1999); P. 162–181, 23 FIGS., 1 TABLE.

Fast 2-D ray+Born migration/inversion in complex media

Philippe Thierry*, Stéphane Operto‡, and Gilles Lambaré*

ABSTRACT

In this paper, we evaluate the capacity of a fast 2-D ray+Born migration/inversion algorithm to recover the true amplitude of the model parameters in 2-D complex media. The method is based on a quasi-Newtonian linearized inversion of the scattered wavefield. Asymptotic Green's functions are computed in a smooth reference model with a dynamic ray tracing based on the wavefront construction method. The model is described by velocity perturbations associated with diffractor points. Both the first traveltime and the strongest arrivals can be inverted. The algorithm is implemented with several numerical approximations such as interpolations and aperture limitation around common midpoints to speed the algorithm. Both theoritical and numerical aspects of the algorithm are assessed with three synthetic and real data examples including the 2-D Marmousi example. Comparison between logs extracted from the exact Marmousi perturbation model and the computed images shows that the amplitude of the velocity perturbations are recovered accurately in the regions of the model where the ray field is single valued. In the presence of caustics, neither the first traveltime nor the most energetic arrival inversion allow for a full recovery of the amplitudes although the latter improves the results. We conclude that all the arrivals associated with multipathing through transmission caustics must be taken into account if the true amplitude of the perturbations is to be found. Only 22 minutes of CPU time is required to migrate the full 2-D Marmousi data set on a Sun SPARC 20 workstation. The amplitude loss induced by the numerical approximations on the first traveltime and the most energetic migrated images are evaluated quantitatively and do not exceed 8% of the energy of the image computed without numerical approximation. Computational evaluation shows that extension to a 3-D ray+Born migration/inversion algorithm is realistic.

INTRODUCTION

During the last decades, seismic imaging methods have made great progress. In pioneer works, Claerbout (1970, 1976) proposed an imaging principle and the concept of wave equation migration. Other approaches, such as the Kirchhoff integral method (French, 1975; Schneider, 1978) or the frequency-wavenumber method (Stolt, 1978), emerged rapidly. All these methods have been recast in the general framework of linearized seismic inverse theory (Tarantola, 1987), and we can now refer to migration/inversion methods.

Constant improvements in numerical schemes and in computer power already allow routine applications in 2-D prestack linearized inversion. The actual challenge is rather in the extension to three dimensions. Theoretically, extension to three-dimensions does not exhibit any major difficulty even if it generally leads to excessive computing time or memory requirements. At the present time, only kinematic Kirchhoff 3-D prestack depth migration codes are used in the industry. Moreover, the use of 3-D migration is motivated strongly by 3-D complex model studies. In this context, the limitation to first-arrival traveltime computed by some Eikonal solver (Podvin and Lecomte, 1991) in kinematic migration may provide poor quality results (Geoltrain and Brac, 1993). Thus, alternative approaches allowing both accurate complex media imaging and CPU efficiency must be investigated (Audebert et al., 1997).

With this prospect, the goal here is to assess the 3-D extension of the 2-D ray+Born migration/inversion method with application to complex media.

CPU efficiency of migration methods depends strongly on the computation of Green's functions. In two dimensions, various approaches have been proposed using finite-difference solvers of the wave equation (Kolb et al., 1986; Crase et al.,

---

Published on Geophysics Online, October 8, 1998. Manuscript received by the Editor April 14, 1997; revised manuscript received February 2, 1998.

*Ecoles des Mines de Paris, ARMINES-GEOPHY, 35 rue Saint Honore, 77305 Fontainebleau Cedex, France. E-mail: philippe.thierry@geophy.ensmp.fr, gilles.lambare@geophy.ensmp.fr.

‡Formerly Ecole des Mines de Paris, ARMINES-GEOPHY, 35 rue Saint Honore, 77305 Fontainebleau Cedex, France; currently Observatoire Océanique de Villefranche sur Mer, UMR Géosciences Azur CNRS, BP48, 06235 Villefranche sur Mer Cedex, France. E-mail: operto@obs-vlfr.fr.

1990; Pica et al., 1990) and of the paraxial one-way equation (Claerbout, 1976), and asymptotic methods such as WKBJ (Ikelle et al., 1988; Cao et al., 1990) or ray theory (Beydoun and Mendes, 1989; Lambaré et al., 1992). All these methods may be extended conceptually to three dimensions, but the computing time and memory requirements for 3-D applications may be prohibitive (Aminzadeh et al., 1994).

Among these seismic modeling methods, ray theory appears to offer a good compromise between precision and computational efficiency. Ray theory has been used for a long time in seismology (Červený et al., 1977; Aki and Richards, 1980), but the development of ray-tracing codes especially dedicated to migration in complex velocity models is rather recent (Lambaré et al., 1992, 1996; Vinje et al., 1993b; Rekdal and Biondi, 1994; Lucio et al., 1996).

A significant breakthrough was obtained with the finite-difference computation of minimum traveltime (Vidale, 1988; Podvin and Lecomte, 1991). An extremely fast and efficient algorithm for kinematic Kirchhoff migration can be built on Noble et al. (1996), but the results are poor when imaging complex structures (Geoltrain and Brac, 1993).

A solution for overcoming this limitation is to appeal to dynamic ray tracing, enabling the computation of both traveltime and amplitude. Several ray-tracing codes dedicated to migration have been proposed (Lambaré et al., 1992; Vinje et al., 1993a).

Compared to other methods, the wavefront construction method (Vinje et al., 1993a) may be more easily extended to three dimensions (Chilcoat and Hildebrand, 1995; Vinje et al., 1996; Lucio et al., 1996). Lambaré et al. (1996) and Lucio et al. (1996) improved this method by introducing a criterion for ray density ensuring a uniform precision for the ray-field sampling. These algorithms provide interpolated maps of traveltime, amplitude, and angles of rays for all the arrivals in 2-D and 3-D smooth velocity fields.

In addition to CPU efficiency, ray theory has additional advantages. The decomposition of Green's functions in terms of traveltime, amplitude, and source signature allows a low storage space of Green's functions and great simplifications in the migration algorithm.

Approaches for properly taking into account the amplitude information are based on a reformulation of migration in the frame of asymptotic linearized inversion. Ray+Born and ray+Kirchhoff forward approximations constitute the bases of such approaches. In these methods, the model is split into a diffractor or reflector distribution and a velocity macro model where ray tracing is performed. The velocity field is supposed to be known, and only the diffractor/reflector distribution is to be determined.

The first studies were devoted to the particular case of homogeneous velocity fields where the relation between perturbation of data and perturbation of model parameter can be rigorously diagonalized in time-space Fourier domain (Cohen and Bleistein, 1979).

Generalization of this diagonalization to more general velocity fields is not straighforward and requires additional approximations. Extensions to the case of a depth-dependent velocity field were first proposed by Clayton and Stolt (1981), Cohen and Hagin (1985), and Bleistein and Gray (1985), but the significant breakthrough was performed by Beylkin (1985). He recognized in the ray+Born operator a generalized Radon transform (Deans, 1983) and proposed an asymptotic inverse operator in terms of the inverse generalized Radon transform. In fact, since the Radon transform can be expressed in terms of the Fourier transform (Deans, 1983; Chapman, 1985), the Beylkin approach can be seen as an approximative extension to general velocity fields of the rigorous diagonalization of the ray+Born operator.

The work of Beylkin formed the basis of numerous developments. Kirchhoff migration was reformulated (Bleistein, 1987; Schleicher et al., 1993; Lumley, 1993) giving quantitative distributions of reflection coefficients and specular reflection angles. Applications to 2.5-D (Bleistein, 1987) and 3-D migration/inversion (Cohen et al., 1986; Schleicher et al., 1993) were proposed, and multiparameter linearized inversion was also developed (Beylkin and Burridge, 1990).

In parallel to all these studies, a specific effort was devoted to recasting ray-based migration in the general frame of the inverse problem theory (Tarantola, 1987). Solving the inverse problem consists of minimizing a given cost function quantifying the misfit between the observed data and calculated data. The works of Ikelle (Ikelle et al., 1986, 1988) and Beydoun and Mendes (1989) are expressions of such efforts.

With the classical ℓ^2 norm as a cost function, the solution of the linearized inverse problem can be expressed explicitly as a combination of gradient and Hessian terms (Tarantola, 1987). Except in the case of a depth-dependent velocity model (Ikelle et al., 1988), the computation of the solution requires the inversion of a huge Hessian matrix, which is essentially band diagonal. However, the approximation by its diagonal (Beydoun and Mendes, 1989; Plessix, 1996) is a rather poor approximation, and an iterative gradient minimization was proposed.

At this stage, the question was: how can one take advantage of Beylkin's approximations within the general frame of inverse theory? This step is not straightforward and was achieved by Jin et al. (1992). They proposed a weighted ℓ^2 cost function which provides asymptotically a diagonal Hessian matrix. Since approximations were introduced, an iterative (so-called quasi-Newtonian) minimization was still required but with a reduced number of iterations. The explicit form of the solution allowed theoretical studies of resolution (Jin et al., 1992), as well as discussion of conditioning in the case of multiparameter acoustic and elastic linearized inversion (Jin et al., 1992; Forgues and Lambaré, 1997). Applications to the 3-D case have already been proposed (Thierry et al., 1996; Operto et al., 1997; Sevink and Herman, 1997), and recently the approach was also extended to linearized inversion for seismic attenuation (Ribodetti et al., 1995).

Our main objective in this paper is to show, using a 2-D case study (Marmousi example), that the use of a migration/inversion method that accounts properly for the amplitude of the data improved the quality of the migration in case of complex media while remaining sufficiently fast to be extended to three dimensions.

In the first part of this paper, we recall Jin's approach of seismic 2-D and 2.5-D ray+Born inversion that was implemented in our algorithm. We define our ray+Born migration/inversion (also called in the literature "preserved amplitude prestack migration") as the result of the first iteration of a single parameter ray+Born inversion. We verify on a canonical example the ability of our approach to recover quantitatively the diffracting model. Then, we address the problem of the numerical implementation. We briefly describe the wavefront construction method used for computing traveltime, amplitude, and other ray-related parameters throughout target, and the implementation of the numerical approximations to speed the algorithm.

Finally, the algorithm is tuned on a structurally simple real data set and then applied to the 2-D complex synthetic Marmousi model and data set. The capacity of the method to recover the true amplitude of the perturbations is finally analysed by comparing the exact perturbation model and the migrated sections.

THEORETICAL ASPECTS

2-D ray+Born migration/inversion

Ray+Born migration/inversion can be viewed as a preserved amplitude prestack depth migration based on ray theory, the Born approximation, and linearized inversion (namely, a prestack depth migration which accounts for the amplitude of the data to recover the true amplitude of the model paremeters).

The Born approximation is a linear approximation around a reference model of the relation connecting the traces to the model, here described by the square slowness field. This approximation is relevant when the perturbation of the model requires only small wavelength perturbations. The Born approximation involves the computation of Green's functions in the reference model. We get the ray+Born approximation when such Green's functions are computed by ray theory. Details about the derivation of the 2-D and 2.5-D ray+Born approximation of the scalar wave equation, which we use throughout the paper, are given in Appendix A.

The 2-D and 2.5-D ray+Born summations [equations (A-7) and (A-9)] can be discretized. When all ray fields are single valued, we obtain in the frequency domain (ω denotes the angular frequency) the expression

$$\delta\mathcal{G}_{cal}(\mathbf{r},\omega;\mathbf{s}) \approx \sum_{\mathbf{x}} \mathcal{B}(\mathbf{r},\omega,\mathbf{s},\mathbf{x})\,\delta m(\mathbf{x}), \tag{1}$$

where $\mathbf{r}$ and $\mathbf{s}$ denote, respectively, the receiver and shot positions, $\mathbf{x}$ the position of the discretized model perturbations, $\delta\mathcal{G}_{cal}$ the perturbation of the Green's function, δm the square slowness perturbation, and $\mathcal{B}$ the kernel of the discretized ray+Born operator, with

$$\mathcal{B}(\mathbf{r},\omega,\mathbf{s},\mathbf{x}) = \Delta\mathbf{x}\,\mathcal{A}(\mathbf{r},\mathbf{x},\mathbf{s})\,\mathcal{K}(\omega)\,e^{i\omega\mathcal{T}(\mathbf{r},\mathbf{x},\mathbf{s})}. \tag{2}$$

$\mathcal{A}$ and $\mathcal{T}$ denote the product of the amplitudes and the sum of the traveltimes associated with the rays $\mathbf{s}\rightarrow\mathbf{x}$ and $\mathbf{x}\rightarrow\mathbf{r}$, $\mathcal{K}$ the ray+Born signature, and $\Delta\mathbf{x}$ the surface of the elementary pixel in the model for the discretized ray+Born summation (see Appendix A for more details).

For a given data set and model, the linear relation (1) can be inverted within the general frame of inverse problem theory (Tarantola, 1987). Consider the weighted ℓ^2 cost function

$$\mathcal{C}[\delta m] = \frac{1}{2}\sum_{\mathbf{s}}\sum_{\mathbf{r}}\sum_{\Omega} \mathcal{Q}|\delta\mathcal{G}_{obs}-\delta\mathcal{G}_{cal}|^2, \tag{3}$$

where $\delta\mathcal{G}_{obs}$ is the observed data (data have to be deconvolved from the source signature and water-bottom reverberations), $\mathbf{s}$ and $\mathbf{r}$ denote the shot and receiver numbers, and Ω the discretized frequencies. $\mathcal{Q}$ is a weighting function equivalent to a covariance matrix in the data space, as introduced in the inverse problem theory (Tarantola, 1987). We assume the data space $(\mathbf{s},\mathbf{r},\omega)$ to have one dimension more than the model space $(\mathbf{x})$ (this is the classical situation in standard—multichannel—seismic surveys), and we consider both positive and negative values of the angular frequency ω.

The expression of the solution δm minimizing the cost function is well known:

$$\delta m = \left(\mathcal{B}^\dagger\mathcal{Q}\mathcal{B}\right)^{-1}\mathcal{B}^\dagger\mathcal{Q}\,\delta\mathcal{G}_{obs}, \tag{4}$$

where † denotes a transposed conjugated matrix. The vector, $-\mathcal{B}^\dagger\mathcal{Q}\,\delta\mathcal{G}_{obs}$, is the gradient of the cost function and the square symetric matrix, $\mathcal{B}^\dagger\mathcal{Q}\mathcal{B}$, is the Hessian

$$\mathcal{B}^\dagger\mathcal{Q}\mathcal{B}(\mathbf{x},\mathbf{x}_0) = \sum_{\mathbf{s}}\sum_{\mathbf{r}}\sum_{\Omega}\mathcal{B}^\dagger(\mathbf{r},\omega,\mathbf{s},\mathbf{x})\;\mathcal{Q}\;\mathcal{B}(\mathbf{r},\omega,\mathbf{s},\mathbf{x}_0) \tag{5}$$

$$= \sum_{\mathbf{s}}\sum_{\mathbf{r}}\sum_{\Omega}\mathcal{D}(\mathbf{r},\mathbf{x},\mathbf{x}_0,\mathbf{s},\omega)\;e^{-i\omega\Delta\mathcal{T}(\mathbf{r},\mathbf{x},\mathbf{x}_0,\mathbf{s})}, \tag{6}$$

where

$$\begin{cases} \mathcal{D}(\mathbf{r},\mathbf{x},\mathbf{x}_0,\mathbf{s},\omega) = (\Delta\mathbf{x})^2|\,\mathcal{K}(\omega)|^2\,\mathcal{A}(\mathbf{r},\mathbf{x},\mathbf{s})\,\mathcal{A}(\mathbf{r},\mathbf{x}_0,\mathbf{s}) \\ \Delta\mathcal{T}(\mathbf{r},\mathbf{x},\mathbf{x}_0,\mathbf{s}) = \mathcal{T}(\mathbf{r},\mathbf{x},\mathbf{s}) - \mathcal{T}(\mathbf{r},\mathbf{x}_0,\mathbf{s}) \end{cases}. \tag{7}$$

In the context of seismic inversion, it is generally impossible to invert numerically this huge matrix. One has to resort to iterative gradient minimization, and any improvement in the estimation of the inverse of the Hessian should improve convergence (Tarantola, 1984; Beydoun and Mendes, 1989; Plessix, 1996).

Notice that the summation (6) involves a kernel expressed in terms of amplitude $\mathcal{D}$ and phase $\Delta\mathcal{T}$. It appears analogous to the integral expression of the Dirac function in two dimensions,

$$\delta(\mathbf{x}-\mathbf{x}_0) = \frac{1}{(2\pi)^2}\iint_{\mathbb{R}^2} d\mathbf{k}\;e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}_0)}. \tag{8}$$

To improve the minimization algorithm and to take advantage of explicit expressions of the Hessian, Jin et al. (1992) took advantage of the analogy between Hessian and Dirac expressions [equations (6) and (8)] and proposed to choose $\mathcal{Q}$ in such a way that the Hessian matrix is approximately diagonalized.

Following Jin et al. (1992), we choose $\mathcal{Q}$ as a local weight depending on the imaged point $\mathbf{x}_0$ in the model

$$\mathcal{Q}(\mathbf{s},\mathbf{r},\omega,\mathbf{x}_0) = \frac{\Delta\mathbf{s}\,\Delta\mathbf{r}\,\Delta\omega}{(2\pi)^2\,\mathcal{D}(\mathbf{r},\mathbf{x}_0,\mathbf{x}_0,\mathbf{s},\omega)}\left|\frac{\partial(\mathbf{k},\Theta)}{\partial(\mathbf{s},\mathbf{r},\omega)}\right|, \tag{9}$$

where Θ is the angle between vector $\mathbf{p}_r$ and $\mathbf{p}_s$ (Figure 1), $\mathbf{k} = \omega(\mathbf{p}_r+\mathbf{p}_s)$, $\Delta\mathbf{s}$ and $\Delta\mathbf{r}$ denote the shot and receiver step along the line, and $\Delta\omega$ is the sampling step for the angular frequency.

For the Hessian operator we obtain (see Appendix B for the demonstration)

$$\mathcal{B}^{\dagger}\mathcal{Q}\mathcal{B}(\mathbf{x},\mathbf{x}_0) \approx \int d\Theta\,\frac{1}{(2\pi)^2}\iint d\mathbf{k}\; e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}_0)}$$
$$\approx [\Theta]_{\min}^{\max}\,\delta(\mathbf{x}-\mathbf{x}_0). \tag{10}$$

Here, $[\Theta]_{\min}^{\max}$ denotes the gap between the maximum and minimum value of Θ, obtained by considering the set of rays $\mathbf{s}\to\mathbf{x}_0$ and $\mathbf{r}\to\mathbf{x}_0$ for all the traces. The Hessian matrix has now been approximately diagonalized and can thus be inverted easily.

With the weighting function $\mathcal{Q}$, expression (4) of the solution becomes

$$\delta m(\mathbf{x}_0) \approx \frac{1}{[\Theta]_{\min}^{\max}}\sum_{\mathbf{s}}\sum_{\mathbf{r}}\mathcal{E}(\mathbf{r},\mathbf{x}_0,\mathbf{s})\frac{1}{2\pi}$$
$$\times\int_{-\infty}^{+\infty} d\omega\,\mathcal{Z}(\omega)\,\delta\mathcal{G}_{obs}(\mathbf{r},\omega,\mathbf{s})\,e^{-i\omega\mathcal{T}(\mathbf{r},\mathbf{x}_0,\mathbf{s})} \tag{11}$$

$$\approx \frac{1}{[\Theta]_{\min}^{\max}}\sum_{\mathbf{s}}\sum_{\mathbf{r}}\mathcal{E}(\mathbf{r},\mathbf{x}_0,\mathbf{s})$$
$$\times\,[\mathcal{Z}(t)*\delta\mathcal{G}_{obs}(\mathbf{r},t,\mathbf{s})](\mathcal{T}(\mathbf{r},\mathbf{x}_0,\mathbf{s})), \tag{12}$$

where $*$ denotes the time convolution and

$$\begin{cases}\mathcal{E}(\mathbf{r},\mathbf{x}_0,\mathbf{s}) = \dfrac{1}{2\pi}\,\dfrac{|\mathbf{q}|^2\Delta\mathbf{s}\,\Delta\mathbf{r}}{\mathcal{A}(\mathbf{r},\mathbf{x}_0,\mathbf{s})}\left|\dfrac{\partial\,(\Phi_{\mathbf{s}})}{\partial(\mathbf{s})}\right|\left|\dfrac{\partial(\Phi_{\mathbf{r}})}{\partial(\mathbf{r})}\right|, \\[2ex] \mathcal{Z}(\omega) = |\omega|(\mathcal{K}(\omega))^{-1}\end{cases} \tag{13}$$

where $|\mathbf{q}|$ denotes the norm of vector $\mathbf{q}$, and $\Phi_{\mathbf{s}}$ and $\Phi_{\mathbf{r}}$ are, respectively, the angles associated with the slowness vectors of the rays $\mathbf{s}\to\mathbf{x}$ and $\mathbf{r}\to\mathbf{x}$ in $\mathbf{x}$ (Figure 1). Expressions of $\mathcal{A}$ and $\mathcal{K}$ in two dimensions and two-and-a-half dimensions are given by equations (A-8) and (A-10). For $\mathcal{Z}$, we get

$$\mathcal{Z}^{2\text{-}D}(t) = \mathcal{H}[\delta(t)] \quad \mathcal{Z}^{2.5\text{-}D}(t) = \mathcal{H}\left[\frac{H(t)}{\sqrt{t\pi}}\right] = \frac{H(-t)}{\sqrt{-t\pi}}, \tag{14}$$

where $\mathcal{H}$ denotes the Hilbert transformation and H the Heaviside function. To deduce expression (13), we used (see Appendix C for a proof)

$$\left|\frac{\partial(\mathbf{k},\Theta)}{\partial(\mathbf{s},\mathbf{r},\omega)}\right| = |\omega||\mathbf{q}|^2\left|\frac{\partial(\Phi_{\mathbf{s}})}{\partial(\mathbf{s})}\right|\left|\frac{\partial(\Phi_{\mathbf{r}})}{\partial(\mathbf{r})}\right|. \tag{15}$$

The computation of the Jacobians $|\partial(\Phi_{\mathbf{s}})/\partial(\mathbf{s})|$ and $|\partial(\Phi_{\mathbf{r}})/\partial(\mathbf{r})|$ can be done by paraxial ray tracing (Farra and Madariaga, 1987).

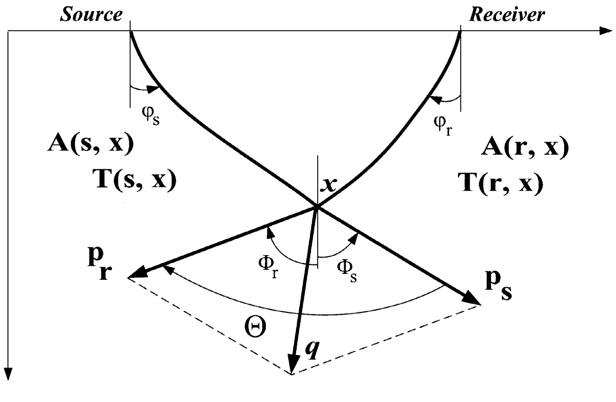


FIG. 1. Ray+Born approximation. Vector $\mathbf{q}$ is defined as the sum of the slowness vectors $\mathbf{p}_r = \nabla_{x_0}T(\mathbf{x}_0,\mathbf{r})$ and $\mathbf{p}_s = \nabla_{x_0}T(\mathbf{x}_0,\mathbf{s})$. Angle Θ is the angle between vector $\mathbf{p}_r$ and $\mathbf{p}_s$. Angles $\Phi_{\mathbf{s}}$ and $\Phi_{\mathbf{r}}$ are, respectively, the angles associated with the slowness vectors $\mathbf{p}_s$ and $\mathbf{p}_r$. Angles $\phi_{\mathbf{s}}$ and $\phi_{\mathbf{r}}$ are the takeoff angles at the source and receiver, respectively.

Although $\mathcal{Q}$ was introduced merely to obtain mathematical simplifications, it retains a clear physical meaning: $\mathcal{Q}$ corrects for amplitude decay, differential two-way isochron spreading, and Born's operator signature. The basic philosophy in Jin's approach can be summarized as: when no physical reason favors a particular cost function in an inverse problem, use the one that leads to the simplest solution. In the original works of Jin et al. (1992), Lambaré et al. (1992), Forgues and Lambaré (1997), or Sevink and Herman (1997), the numerical ray+Born schemes were iterative with residual computations. In the present study, the final images are the first iteration results, but we will see with the following canonical example that the limitation to a single iteration does not dramatically affect the resolution of the method.

The main differences between the ray+Born migration/inversion (12) and a Kirchhoff-preserved amplitude migration (Bleistein, 1987; Beydoun and Jin, 1994) are the weights and the filters applied to the data and, thus, the nature of the output. Whereas Kirchhoff-preserved amplitude migration returns (for each common-shot or common-offset gather) a migrated image of the specular reflectivity of continuous interfaces, the ray+Born migration/inversion returns a single migrated image of the perturbation of model parameters.

A canonical example

The model consists of a homogeneous medium (constant velocity of 3500 m/s) within which is buried a 100-m-thick homogeneous layer at 1.5-km depth (constant velocity of 3700 m/s). The synthetic data set was computed by ray theory taking into account the 3-D propagation (Figure 2). It consists of 120 shots

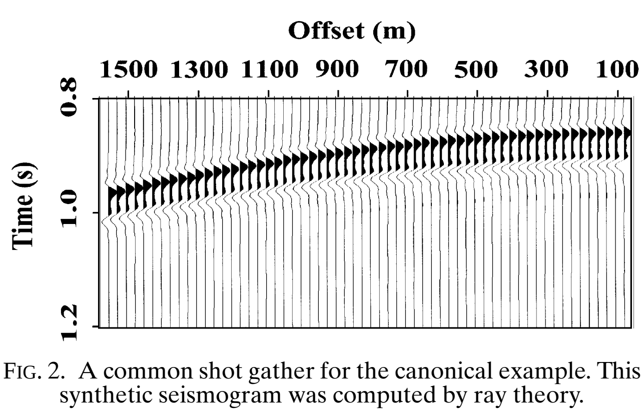


FIG. 2. A common shot gather for the canonical example. This synthetic seismogram was computed by ray theory.

spaced 25 m apart with 60 receivers per shot spaced 50 m apart. The first offset was 50 m, the time sampling was 2 ms, and the source signature was a zero-phase signal with a trapezoidal spectrum [0, 10, 35, 55] Hz.

A smooth representation of the true model (Figure 3a) was used for the migration. As suggested by Versteeg (1991), a lowpass Gaussian filter, $[(1/r\sqrt{\pi})\exp(-x^2/r^2)]$ combined with a coarse B-spline parameterization was used to smooth the exact model (the same parameters will be used later for the Marmousi example). The radius r of the Gaussian filter and the spacing between spline knots were both 75 m.

After migration, we summed the velocity macro model c_0 and the perturbation model δm inferred from the migration and compared the resulting model with the true model (Figure 3b). The amplitude and the shape of the 100-m-thick layer buried in the homogeneous medium are quite well recovered with a single iteration. Note that the shape of the recovered model (velocity macro model + perturbation model) exhibits some ripples on the top and bottom of the layers interpreted as a Gibbs phenomenon. Second, the average amplitude of the recovered model ($c_0 + \delta m$) slightly exceeds the amplitude of the true model in the layer. This suggests an overlap of the spatial bandwidth of the velocity macro model and the perturbation model. The use of a smoother velocity macro model would have lead to a closer match of the recovered and the exact model.

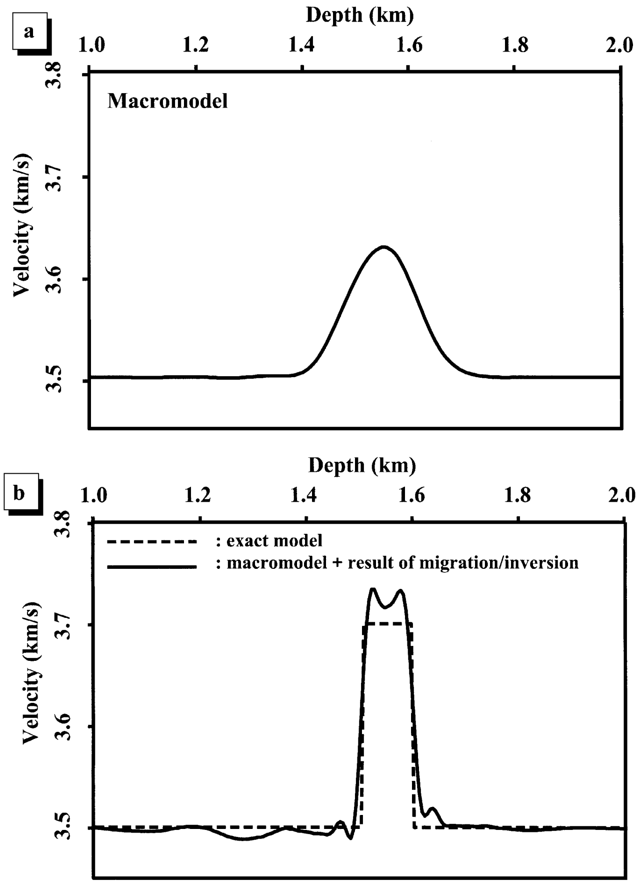


FIG. 3. The 2.5-D ray+Born migrated profile for the canonical example. (a) The velocity macro model is a smooth representation of the true model. It was lowpass filtered to 150 m (see Marmousi example) and projected into a cubic cardinal B-splines basis with knots spaced by 75 m apart. (b) Comparison between the exact model and the model resulting from the sum of the background model and the inversion result.

IMPLEMENTATION

Ray tracing for imaging

Numerical implementation of the ray+Born migration/inversion expression (12) requires computation of $\mathcal{E}$ and $\mathcal{T}$ by ray tracing in the reference model $c_0(\mathbf{x})$. The difficulties encountered when building such an algorithm mainly arise from the singularities associated with caustics and multiple arrivals, but also from the necessity to control the ray-field sampling.

Recently proposed by Vinje et al. (1993a), the wavefront construction method offers an interesting basis. The ray field associated with the 2-D Green's functions is divided into elementary quadrangular cells defined by adjacent rays and successive wavefronts. The sampled wavefronts are propagated along a set of rays by constant traveltime increments. When the "distance" between two adjacent rays violates some ray density criterion, a new ray is added on the previous wavefront by interpolation and propagated along with the old rays. Once the cells have been constructed, the desired parameters are interpolated on the target from the values at the vertices of the cells.

The ray density criterion proposed by Vinje et al. (1993a) used the metric distance between the rays. This criterion leads to a drastic undersampling in caustic regions. To remedy such limitations, Lambaré et al. (1996) (in 2-D) and Lucio et al. (1996) (in 3-D) recast the wavefront construction method in the general frame of the Hamiltonian formulation of ray theory and paraxial ray theory (Farra and Madariaga, 1987). Their approach avoids the problems associated with caustics and multiple arrivals. Indeed, their ray density criterion ensures a "uniform" precision for the ray-field sampling along the sampled wavefronts, even in the case of multiple arrivals or caustics. Their algorithm appears robust, precise, and CPU efficient. Compared to those developed by Podvin and Lecomte (1991), for estimating the first-arrival traveltime with an eikonal solver, the computing time is ten times longer for a test over a complex velocity field, but could provide nine ray parameters for all the arrivals (Lambaré et al., 1996).

In our imaging code, we used the wavefront algorithm developed by Lambaré et al. (1996). Six parameters are required in two-and-a-half dimensions: the traveltime T, the amplitude A, the parameter τ [see equation (A-10)], the slowness vector $\mathbf{p}$, and the takeoff angle ϕ (Figure 1).

The wavefront construction method has already been used intensively for 2-D migration (Moser and Pajchel, 1995; Thierry and Lambaré, 1995; Ettrich and Gajewski, 1996) and such applications will certainly develop in the coming years especially in the 3-D case (Thierry et al., 1996; Operto et al., 1997; Tura et al., 1997). Interfaces can be incorporated into the wavefront construction method (Vinje et al., 1993a, 1995; Moser and Pajchel, 1997), although this strategy yields a very significant increase in computational cost and algorithmic complexity.

For CPU efficiency, we decided to stick to smooth reference velocity fields. This is still a highly debated question, but we believe that, in many cases of practical interest, the introduction of interfaces in the reference model ("blocky" model), while inducing a huge computational overload, will not improve the imaging results significantly (Versteeg, 1991; Clar et al., 1996; Mispel and Hanitzsch, 1996). On the contrary, refining the quality of the smooth model itself and accounting in detail for the real acquisition geometry can definitely yield appreciable improvements.

Practical aspects of the migration/inversion

The CPU efficiency as well as the applicability to reasonably complex models was an unavoidable prerequisite for the extension to three dimensions. In this context, we will now focus on simple numerical approximations that we introduced to speed the code.

Target interpolation of the migration operator.—The kernel of the migration operator has a classical high-frequency asymptotic form [equation (12)] with a "rapidly" oscillating waveform $\mathcal{Z} * \delta\mathcal{G}_{obs}$ shifted by "slowly" varying time $\mathcal{T}$ and modulated by a "slowly" varying amplitude $\mathcal{E}$. The calculus of $\mathcal{T}$, Θ, and especially $\mathcal{E}$ can be an expensive operation in terms of computing time if it is done directly at the final sampling in the target. If these quantities have slow variations over our target zone, it seems reasonable to calculate them over a coarse regular target grid and to interpolate them linearly all over the fine target grid (Figure 4b).

The main advantages are a significant decrease of the required memory space and a slight decrease of the computing time required to compute these coarsely sampled Green's functions. Typically, the target interpolation of the migration operator divides the total computing time by a factor of three.

Linear interpolation has the advantage of CPU efficiency, and it must be clear that higher order interpolations would improve the accuracy while increasing the computing time.

Surface Interpolation of Green's functions.—At the surface, the source or receiver positions are densely sampled. In the same way as we interpolate the migration operator, we can interpolate at the surface the maps obtained by ray tracing (Figure 4a).

Linear interpolation has the great advantage of CPU efficiency, but there are many ways of interpolating the maps for the ray shooting position. Interpolation essentially depends on the map sampling and parameterization (for example, Pica, 1997, proposed a reparameterization of traveltime maps in celerity).

The map sampling may be expressed in terms of regular positions in the target or in terms of regular offsets with respect to the ray shooting positions at the surface. This second solution is interesting because the accuracy of the surface interpolation depends only on the lateral variations of the velocity macromodel (it was used, e.g., by Thierry et al., 1996; Operto et al., 1997). However, it must be clear that at the final step, maps will have to be sampled regularly in terms of regular and common positions in the target. Consequently, any other choice for the sampling will require this final interpolation in the target.

In case of multiple arrivals, linear interpolations of maps appear inappropriate. If one wants to keep interpolations to maintain CPU efficiency (from our point of view, it is unavoidable in three dimensions), the strategy must be reevaluated. A first possibility should consist of performing the interpolation for each branch of the ray field. This was already exploited in two dimensions for the interpolation in the target (Lambaré et al., 1992). The extension to interpolation at the surface and moreover to three dimensions appears difficult, because it imposes the precise determination of the boundaries of branches that can be problematic in three dimensions (Bulant, 1996). In fact, a rigorous approach should involve the interpolation of the ray field (the cells in the case of wavefront construction) in $(\mathbf{x}, \mathbf{p})$ space, where $\mathbf{p} = \nabla T$ denotes the slowness vector (Lambaré et al., 1996). In this space, the ray fields are never folded. To our knowledge, such approaches have not been developed and applied to real data. From a very practical point of view, with this paper we test the linear interpolation of maps at a constant position in the target and for, respectively, the first and the strongest arrivals in a complex model with multiple arrivals (Marmousi).

Dynamic storage of Green's functions.—The classical 2-D marine acquisition pattern exhibits a strong redundancy of receiver positions for successive shots. In former 2-D codes

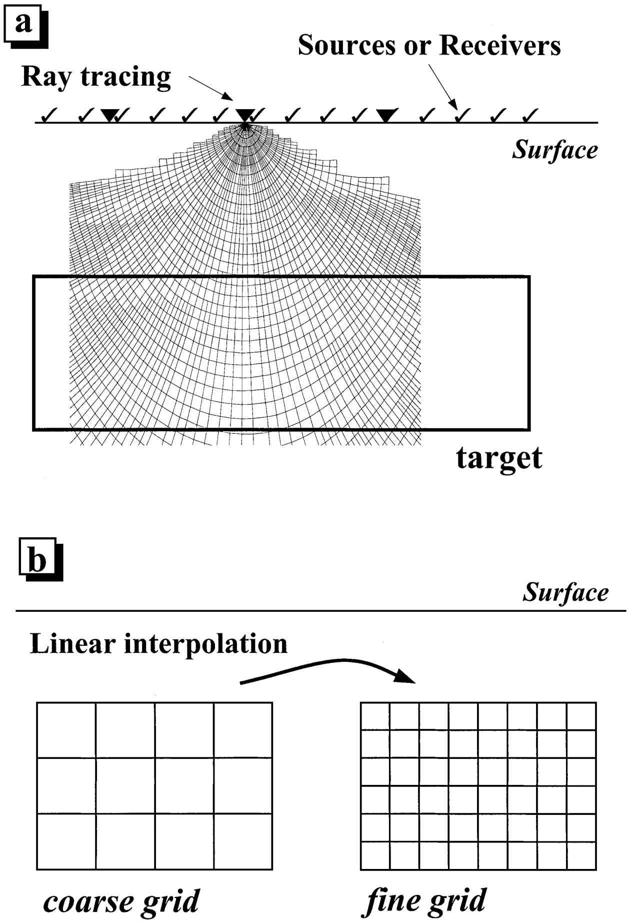


FIG. 4. (a) Surface interpolation of Green's functions. Rays are shot from the surface into the target zone for a limited number of shooting positions. Maps are then linearly interpolated for intermediate shooting positions. (b) Target interpolation of the preserved amplitude migration operator. The amplitude $\mathcal{E}$ and time T of the migration operator is calculated over a coarse grid in the target and then linearly interpolated over a fine target grid.

(Lambaré et al., 1992; Forgues and Lambaré, 1997), Green's functions were stored on disk because of small available core memory, and input/output represented a significant part of the imaging process. Present memory size on standard workstations allows the preservation of information on relevant Green's functions for successive shots using a specific dynamic storage. When migrating a shot, the maps associated with all receivers of the streamer are stored in memory. When the next shot is migrated, only maps corresponding to the new receiver positions are computed. They are incorporated into the memory storage, replacing the maps of the overtaken receiver positions (Figure 5b). This simple improvement is particularly interesting in terms of CPU efficiency for the 2-D migration even if it cannot be extented simply to three dimensions because of the lateral feathering in the 3-D marine acquisitions (Thierry et al., 1996).

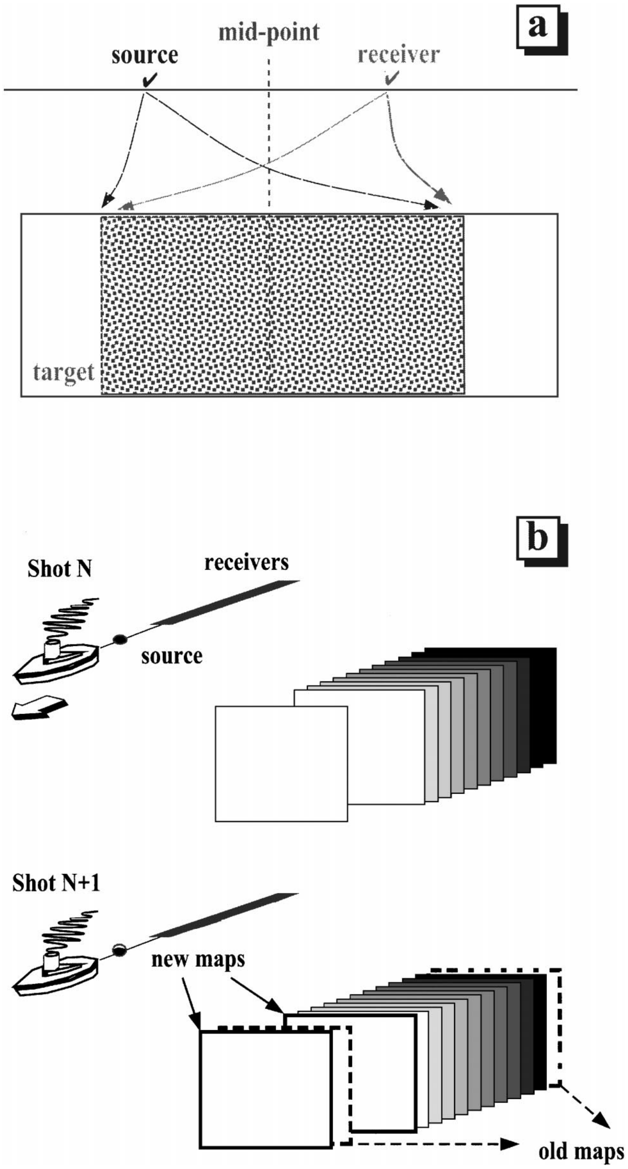


FIG. 5. (a) Limitation of the contribution zone of each trace. Margins are added around the CMP. Typically the total size in the x-direction is about half the length of the streamer. (b) Dynamic storage of Green's functions. For each shot, maps associated with all receivers of the streamer are stored in memory. For the next shot, only maps corresponding to new receiver positions are calculated. They are incorporated into storage, replacing maps of overtaken receiver positions.

Limited aperture.—In the case of reasonable lateral variations in the model, the contribution of each trace to the final migrated image can be limited to an area around the common midpoint (CMP) position (Figure 5a). The aperture can be evaluated heuristically in view of the complexity of the model, the frequency content of the source and the offset range.

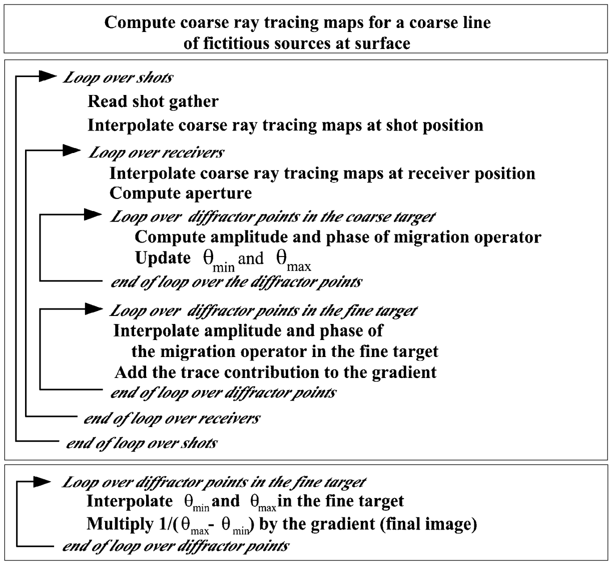
Rather than having theoretical discussions on the relevance and tuning of such approximations, we shall discuss them with the following validation tests. The implementation of the algorithm is summarized on a flow chart (Figure 6).

VALIDATION TESTS

A real marine seismic line: Tuning of the migration/inversion code

This first example concerns a structurally simple model: a marine profile provided by Norsk Hydro. It consists of 398 shots spaced 37.5 m apart with 120 receivers per shot spaced 25 m apart. The time sampling is 4 ms for a cutoff frequency of 67 Hz, and each trace contains about 1000 samples for a total size of about 200 Mbytes. Preprocessing, performed by Norsk Hydro, consisted of source signature deconvolution and water bottom reverberations removal. Figure 7 shows the data for shot 200.

Sixteen rms velocity logs were provided by Norsk Hydro. They were first interpolated in time using smooth cubic B-spline knots every 80 ms, and then converted into interval velocity profiles with the Dix formula. Finally, the velocity macromodel was defined by cubic B-spline profiles with 300-m steps in the x- and z-directions and with weights simply defined by undersampling the finely sampled smooth interval velocity profiles. Moreover, before the final migration, we improved



```
Compute coarse ray tracing maps for a coarse line
of fictitious sources at surface

Loop over shots
    Read shot gather
    Interpolate coarse ray tracing maps at shot position
    Loop over receivers
        Interpolate coarse ray tracing maps at receiver position
        Compute aperture
        Loop over diffractor points in the coarse target
            Compute amplitude and phase of migration operator
            Update $\theta_{min}$ and $\theta_{max}$
        end of loop over the diffractor points
        Loop over diffractor points in the fine target
            Interpolate amplitude and phase of
                the migration operator in the fine target
            Add the trace contribution to the gradient
        end of loop over diffractor points
    end of loop over receivers
end of loop over shots

Loop over diffractor points in the fine target
    Interpolate $\theta_{min}$ and $\theta_{max}$ in the fine target
    Multiply $1/(\theta_{max} - \theta_{min})$ by the gradient (final image)
end of loop over diffractor points
```

FIG. 6. Flow chart of the ray+Born migration/inversion algorithm. The main partitions of the algorithm are framed. The main loops are indicated by italic characters.

the velocity field slightly by trial-and-error flattening of common reflection point (CRP) sections (Figure 8). The velocity macro model does not exhibit strong lateral variations, and no caustic arose when ray tracing (Figure 9). It is a convenient model for applying our interpolation strategies. The following tunable parameters are used for the wavefront construction: $\Delta T = 40$ ms, $dX_{\max} = 2$ m, and $dP_{\max} = 2 \times 10^{-6}$ m/s (Lambaré et al., 1996).

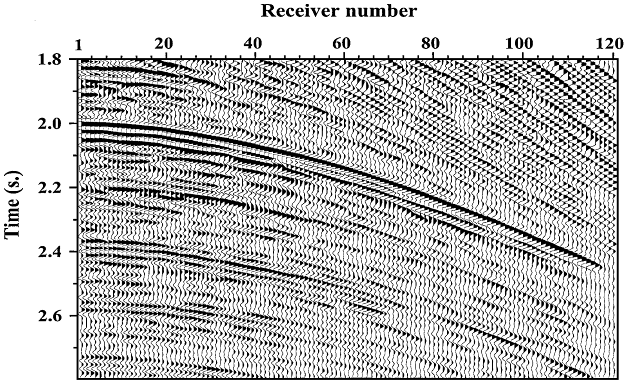


FIG. 7. Common shot gather for shot 200. The time sampling is 4 ms for a highest frequency of 67 Hz. Water bottom reverberations were removed, and a spiking deconvolution was applied.

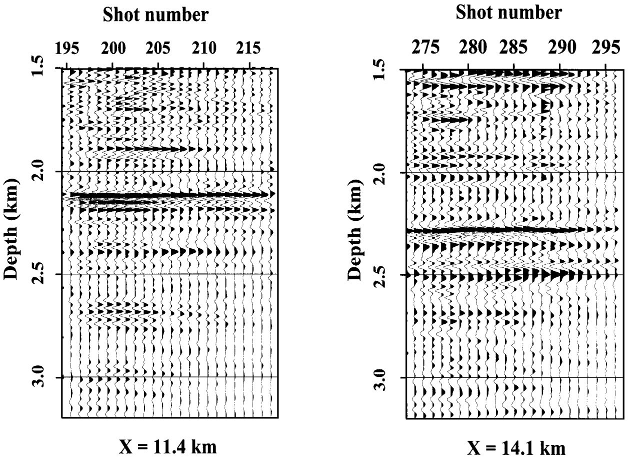


FIG. 8. Validation of the velocity macro model: CRP panels at $x = 11.4$ km and $x = 14.1$ km.

We applied our 2.5-D ray+Born migration/inversion scheme to the whole data set in the target zone of 7000 × 810 m (560 × 136 points with sampling steps 12.5 × 6 m). Tests were done for evaluating the tunable parameters in the migration code (sizes of the coarse surface and target grids, and size of the contribution zone). The aperture was fixed to 1000 m on each side of the CMP, and we checked the steps of the coarse surface and target grids considering the misfit with respect to a reference image obtained without approximations. Four sets of values of the grid steps were tested (Table 1). The size of the interpolation steps increased with a benefit in computing time and memory requirement and, conversely, with a decreasing image accuracy [see the final image for the reference computation without interpolation (Figure 10a), the result for set (3) (Table 1; Figure 10b), and the residual section (Figure 10c)]. No gain was applied to the sections that were plotted with the same clip (about one third of the maximum value of Figure 10a).

The migrated sections represent the perturbation of velocity given by the first-order relation $\delta c = -\delta m\, c_0^3/2$ rather than squared slowness δm. Some spurious events can be seen on the section (Figure 10a) for $x \in [8, 9]$ km. They correspond to spurious plane waves in some data. Also noticeable is that the water-bottom multiple has not been eliminated sufficiently by preprocessing. Note, however, that the purpose of the test was not to check the absolute quality of imaging but simply to check the influence of our approximations on the image.

Table 1 summarizes various indicators of the validations tests. We see that computing time is divided by four while the memory requirement is divided by 30. On the other hand, the energy of the misfit grew up to 6.3%. We define the "energy" of an image as

$$E = \sum_{i,j} (\delta c_{ij})^2. \tag{16}$$

We used linear interpolations at the surface and in the target. The interpolation of the traveltimes is important since it may result in traveltime shifts, which can be seen as velocity macro

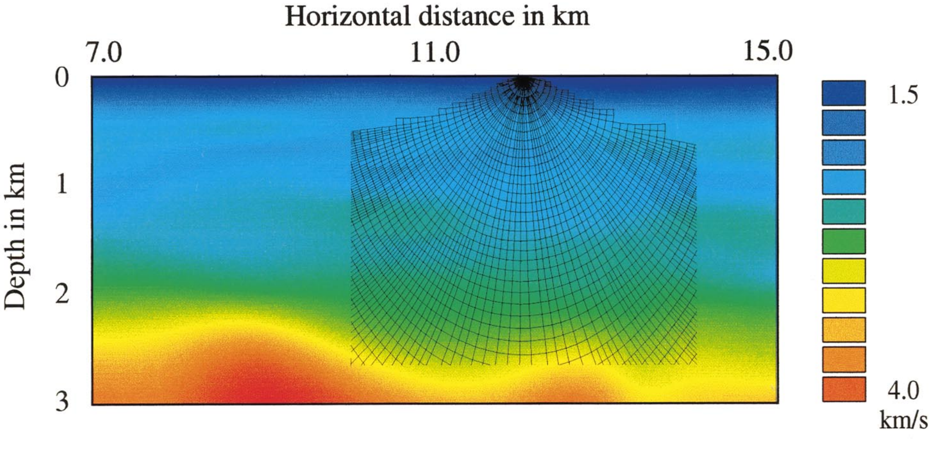


FIG. 9. Smooth velocity macro model. This velocity field is described by cubic cardinal B-splines (knot points are 300 m apart in both the x- and z-directions) and was derived from standard velocity analysis. A ray-field sampling by wavefront construction is superimposed.

Table 1. Computing performance on a Sun SPARC 20 workstation for various values of the tunable parameters of ray+Born migration/inversion and for the real marine line. The contribution zone of each trace is fixed to a 1000-m margin on both sides of the CMP. The reference set consists of a migration/inversion performed without CPU efficient interpolations, and sets (1)–(4) correspond to increasing interpolations steps at the surface and in the target. Energy of the image is defined by expression (16).

	Reference	Set (1)	Set (2)	Set (3)	Set (4)
Coarse sampling at surface (m)	25	50	75	100	150
Coarse sampling (targets) (x/z in m)	12.5/6	48/48	72/72	96/96	150/150
Number of ray tracings	716	362	242	182	122
Computing time for ray tracing (s)	923	330	218	160	99
Computing time for data reading (s)	120	120	120	120	120
Total computing time (s)	1800	812	660	555	540
Total memory requirement (Mbytes)	60	22	10	6	3
E_{misfit}/E_{ref} (%)	0.0	1.0	1.9	3.3	6.3

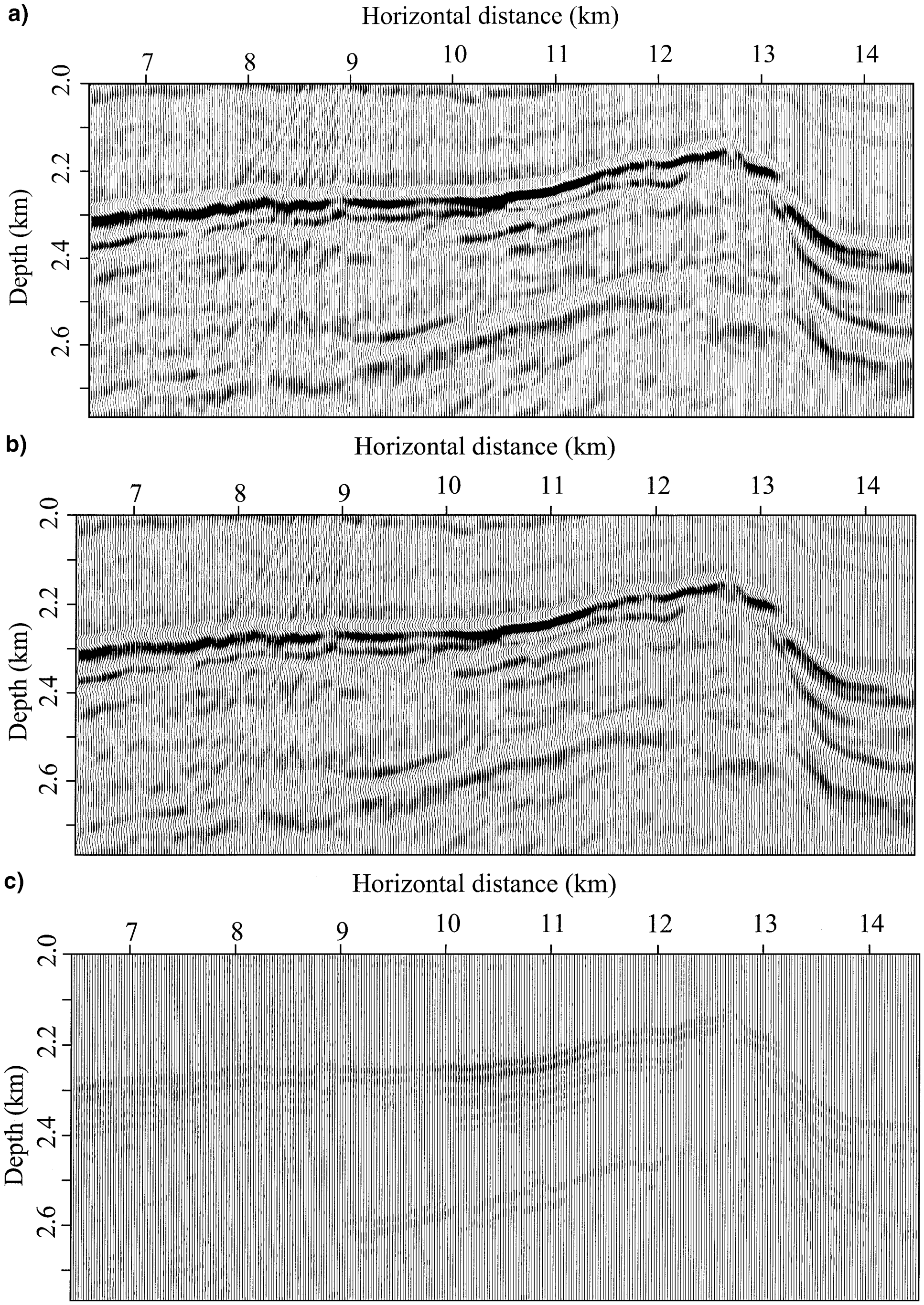


FIG. 10. Sections of the velocity perturbation obtained by 2.5-D ray+Born migration/inversion on a real marine line (see Table 1): (a) without interpolations, (b) with interpolations [parameter set (3), see Table 1], (c) difference of the sections. The same clip was used for the three sections and was about a third of the maximum amplitude in the image (a).

model variations. However, we must admit that our confidence in the velocity model as well as in the preserved amplitude quality of the data make us think that a 5% precision in the energy of the image should be seen as a satisfactory result. From these considerations, we adopted parameter set (3) as a reasonable compromise between computing cost and accuracy.

As a final test for checking the influence of the size of the contribution zone of traces around the CMP, we increased the margin from 1000 to 1500 m for parameter set (3). The energy of the difference section was 3.21%.

Marmousi data set: Application to a highly complex data set

In the previous application, our algorithm was tuned in the case of a real data set with a velocity macro model with gentle lateral variations. It is now tested in the context of a highly complex model. For ray-based migration, the concept of complex models is related to the existence of triplications when tracing rays in the velocity macro model.

Our ray+Born migration/inversion formula was derived considering a single arrival. Even if recent theoretical works (ten Kroode et al., 1994) has shown that the formula could be generalized to the case of multivalued ray fields, our numerical scheme is definitely based on the use of a single arrival. In this example, for complex media we evaluate both the efficiency of ray+Born migration/inversion using the first or the strongest arrivals and the influence of the numerical approximations (mainly, the interpolation scheme) on the amplitude of the perturbations.

We used the synthetic Marmousi model and data set (Versteeg and Grau, 1991). The model consists of a 2-D velocity/density grid of dimensions 751 × 2301 with a mesh spacing of 4 m (9200 m long and 3000 m deep) (Figure 11a shows the velocity grid). The synthetic data were computed by finite differences of the acoustic equation for a standard marine acquisition. There are 240 shots and 96 receivers per shot. The spacing beetween the shots and between the receivers is 25 m, and the first offset is 200 m. The shots are at 8 m depth and the receivers at 12 m. The time sampling is 4 ms.

Standard preserved amplitude preprocessing was applied to the data. First, a deterministic zero-phase deconvolution was applied. The far-field signature was estimated from the source given in Bourgeois et al. (1991) and convolved by the ghost signature. The data were muted, and finally a predictive deconvolution was applied to remove the water-bottom multiples (minimum prediction lag 40 ms; maximum prediction lag 190 ms; percent of prewhitening 10; window for autocorrelation [0.0, 2.9] s).

For the velocity macro model, following Versteeg (1991), we used a smooth representation (Figure 11b) of the true velocity model (Figure 11a). Smoothing was applied to the slowness field combining a Gaussian filter and a coarse B-spline parameterization. We filtered out the structures of spatial dimensions lower than 150 m with a Gaussian filter given by $(1/r\sqrt{\pi})\exp(-x^2/r^2)$ with $r = 76$ m. Finally, the smooth model was projected onto a cubic cardinal B-spline basis with nodes 76 m apart. The parameters for wavefront construction were $\Delta T = 10.0$ ms, $dX_{\max} = 2.0$ m, and $dP_{\max} = 2 \times 10^{-6}$ m/s. The depth of the ray shooting positions was fixed at 10 m (i.e., a mean value between the shots and receivers depths). Multiple arrivals arise when ray tracing (Figure 11c). Upgoing rays were eliminated systematically.

2-D ray+Born migration/inversion with the first arrival.— The Marmousi model is an acoustic model parameterized both by velocity and density. Our inverse formula required a slight adaptation. It is well known that, at near offsets, reflection/diffraction is essentially sensitive to the impedance contrast, $I = \rho c$. In this case, the ray+Born approximation [equation (A-7)] must be modified by replacing the 2-D amplitude of the Green function, $A^{2D}(\mathbf{x}, \mathbf{s})$ [equation (A-4)], by the amplitude

$$A^{2D}_{acous.}(\mathbf{x}, \mathbf{s}) = \sqrt{\rho_0(\mathbf{s})\rho_0(\mathbf{x})}\,A^{2D}(\mathbf{x}, \mathbf{s}), \tag{17}$$

and the perturbation of model, $\delta m(\mathbf{x})$, by

$$\delta m(\mathbf{x}) = \frac{-2}{\rho_0^2(\mathbf{x})c_0^3(\mathbf{x})}\,\delta I(\mathbf{x}). \tag{18}$$

The inversion formula for the perturbation of impedance can be derived from formula (12) simply by taking into account equations (17) and (18).

For the Marmousi data set, the 2-D migration formula was used. The Hilbert transform was first applied to the data set, and only the first arrival was stored while ray tracing. For migration, we took parameter set (3) of Table 1: the coarse sampling for ray tracing at the surface was 100 m, as well as the x and z coarse samplings in the target; the summation aperture was restricted to 2000 m around the CMP. The wider summation zone (in the former example, we took only 1000 m) was motivated by the complexity of the Marmousi model. The migrated profile ranges from 2000 to 9000 m in the x-direction and from 100 to 2800 m in the z-direction with fine sampling steps of 25 m and 6.25 m, respectively (281 × 449 points). The total computing time was 22 minutes on a Sun SPARC 20 workstation with 7.5 minutes for the 90 ray tracings. To favor the amplitude display, we first plotted the migrated section (reconstructed $\delta I/I_0$ field) with a color scale and without any gain correction (Figure 12).

The migrated section can be compared with the exact relative impedance perturbation section (Figure 13) plotted with the same color scale. To take the limited bandwidth of the source into account, the exact depth sections were converted into time sections, band-pass filtered with a trapezoidal filter ([5, 10, 35, 55] Hz), and converted back into depth sections. Qualitatively, the distribution of amplitudes of our image fits quite well with that of the exact perturbation model except in the complex deep zone (the hydrocarbon trap). Figure 14 shows the migrated section with a wiggle display and a low clip (0.05) (for comparison, the color scale used on Figures 12 and 13 for plotting the relative impedance perturbation $\delta I/I_0$ ranged from −0.48 to 0.48) to enhance the impedance contrast in the deep part of the model. Within the zone of the hydrocarbon trap, the amplitudes are strongly underestimated, but the shape of the trap is slightly visible among the noise.

A closer look at the amplitude and waveform match is provided by the direct comparison of three 1-D sections at positions 3700, 6200 (through the hydrocarbon trap), and 8000 m (see Figure 15) extracted from the first-arrival migrated section (Figure 12) and the exact perturbation model (Figure 13). It confirms the good match between the exact and computed sections except within the zone of the hydrocarbon trap. Concerning the resolution of first-arrival migration, our results are in full agreement with former studies on the Marmousi data set (Geoltrain and Brac, 1993; Ettrich and Gajewski, 1996). As stated in Ettrich and Gajewski (1996), the amplitude information is not sufficient for improving the structural resolution of the first-arrival migration in the complex area.

Sensitivity of first-arrival migration/inversion to the interpolations.—The previous migrated sections were generated using our CPU-efficient interpolations (surface and target interpolations). To evaluate the effects of these approximations on the image, we generate a first arrival migrated section without interpolation (Figure 16). The differences between the first arrival migrated sections with and without interpolation are distributed uniformly all over the section (Figure 17). The energy of the residual section is 4.17% of the energy of the migrated section obtained without interpolation. Given the interpolation-related parameters, the sensitivity of image to the interpolations depends on both the spectral content of data and the complexity of the model. With respect to the former example of the real marine seismic line, the cut-off frequency is lower in the Marmousi data set (55 Hz compared to 67 Hz), but the Marmousi model is much more complex, and the interpolations finally result in slightly stronger degradation of the image (which however remains rather reasonable).

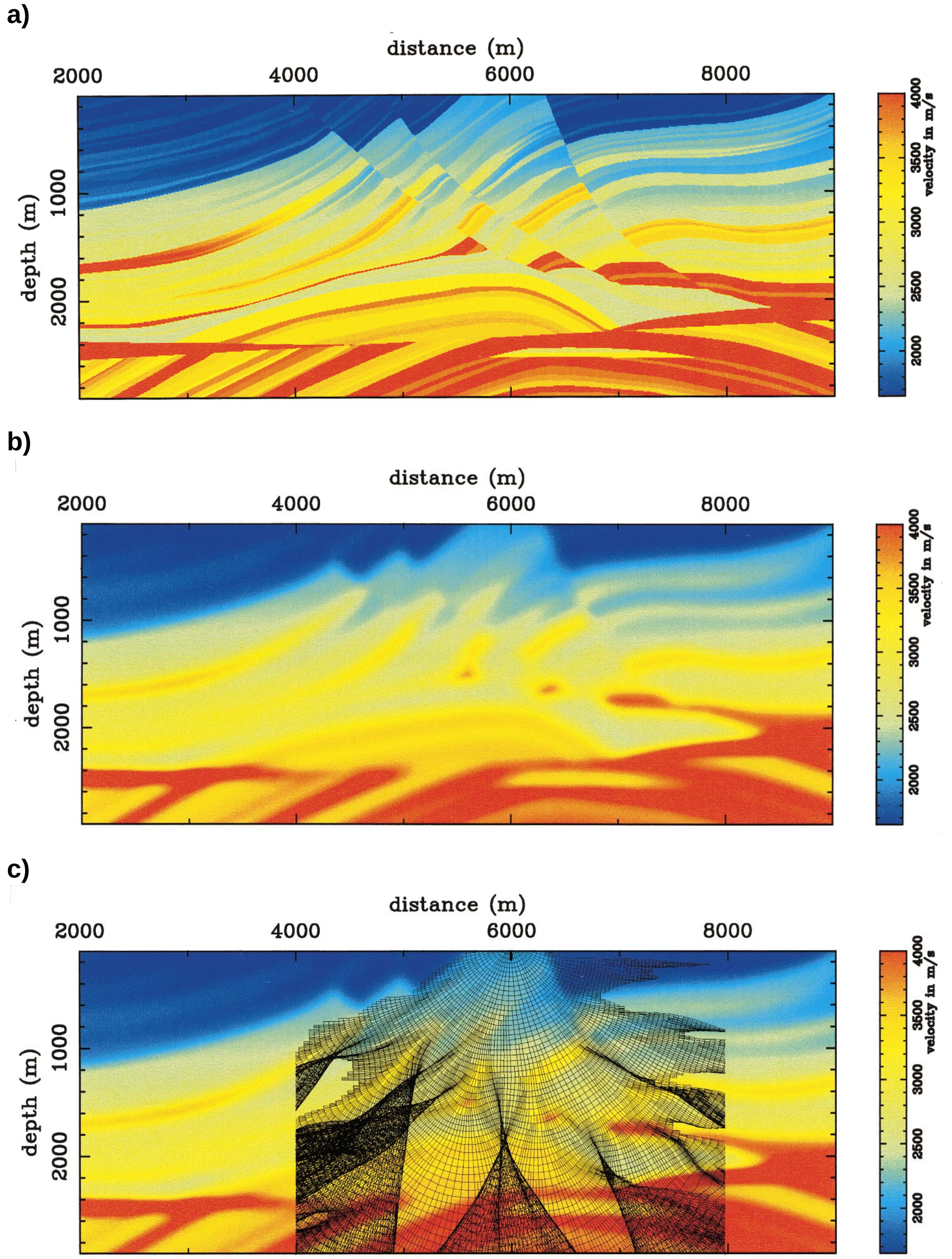


FIG. 11. (a) The velocity grid defined in the Marmousi experience. (b) The smooth velocity macro model. (c) A computed ray field superimposed to the smooth velocity macro model. Note the presence of multivalued arrivals at short offsets in the region of the hydrocarbon trap ($x \in [6.0, 7.5]$ km and $z \in [2.4, 2.6]$ km).

2-D ray+Born migration/inversion with the strongest arrival.—Several authors have mentioned that secondary strong-amplitude arrivals should be used in the migration (Geoltrain and Brac, 1993; Ettrich and Gajewski, 1996; Nichols, 1996; Audebert et al., 1997). To account for these later arrivals, we now investigate the extension of our algorithms to the strongest arrival, which is known to provide much more irregular maps for amplitude, traveltime, etc. (Lambaré et al., 1996).

The use of the strongest arrival only implies minor adaptations in the algorithm. The main modification is associated with the requirement of the KMAH index [equation (A-2)]. The real signal $\delta\mathcal{G}_{obs}$ in equation (11) is replaced by the analytical signal

$$\mathrm{Real}[(\delta G_{obs} + i\mathcal{H}(\delta G_{obs}))(i)^{\alpha}], \qquad (19)$$

where $\mathcal{H}$ denotes the Hilbert transform, and α the sum of the KMAH index for the rays connecting the source and the

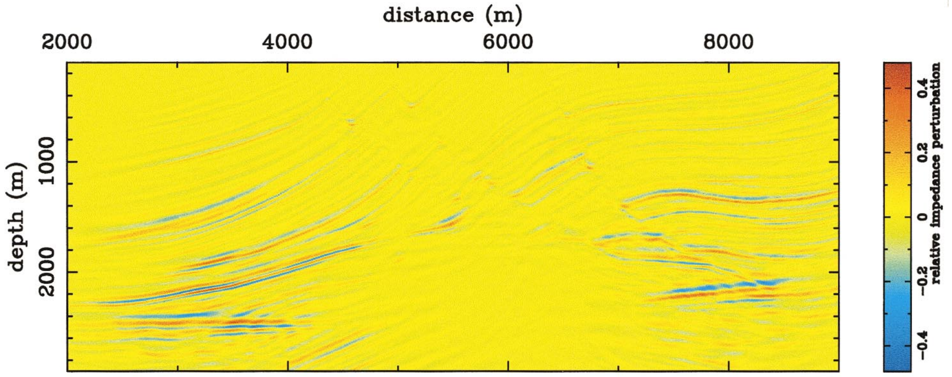


FIG. 12. First arrival migrated section of the Marmousi data set. The relative impedance perturbation $\delta I/I_0$ is plotted with the same color scale as that used for the exact model (Figure 13).

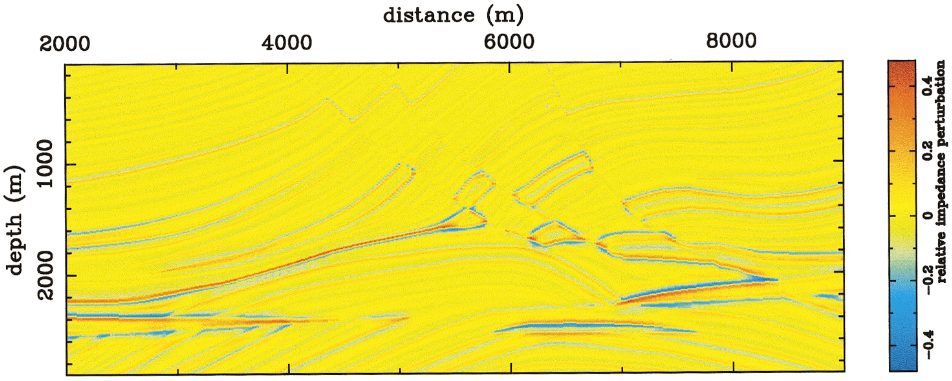


FIG. 13. Exact relative impedance perturbation $\delta I/I_0$ profile. To fit with the wavelengths of the source signature, the exact depth sections were converted into time sections, band-pass filtered with a trapezoidal filter ([5, 10, 35, 55] Hz), and converted back into depth sections.

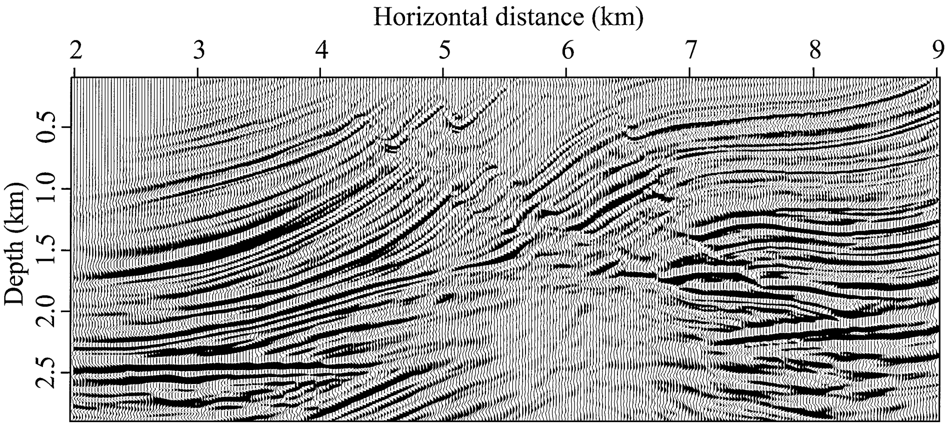


FIG. 14. First arrival migrated section of the Marmousi data set (relative impedance perturbation $\delta I/I_0$). Note that the top of the hydrocarbon trap is slightly visible among the noise.

receiver to the diffracting point. The computing time is equivalent to the one obtained with the first-arrival migration/inversion. The image obtained with the strongest arrival is shown on Figures 18 and 19 with the same color scale and the same clip the ones used for the first arrival (Figures 12 and 14, respectively).

At a quick glance, the strongest arrival migrated section still exhibits strongly underestimated amplitudes within the hydrocarbon trap zone when compared to the exact section (Figure 13). Several reasons could be proposed for such a failure. The limitation to a single arrival (even if it is the strongest!) is not sufficient in this complex model for focusing the whole

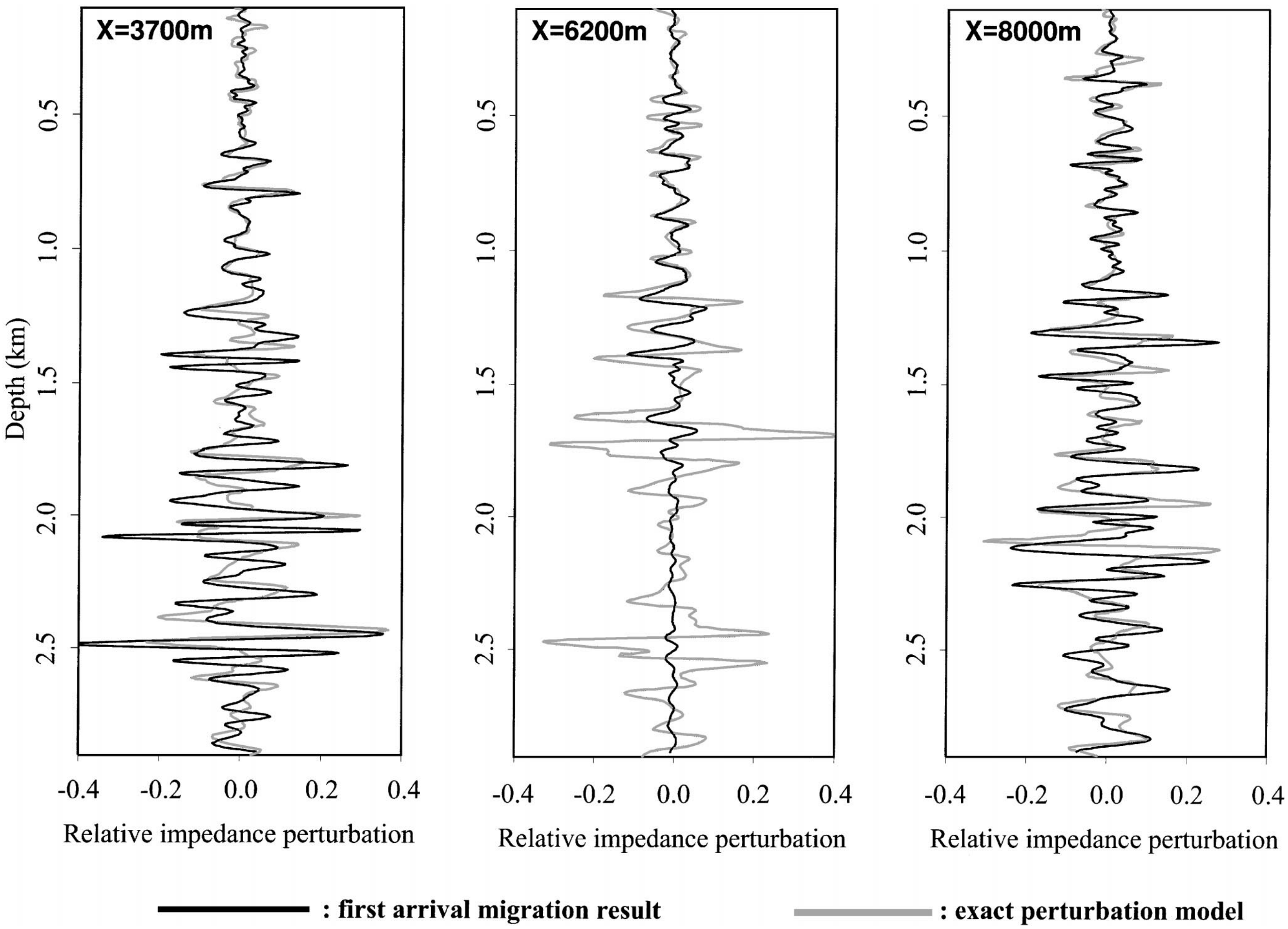


FIG. 15. Comparison of three $\delta I/I_0$ 1-D sections extracted from the exact section (Figure 13) and from the first arrival migrated section (Figure 12).

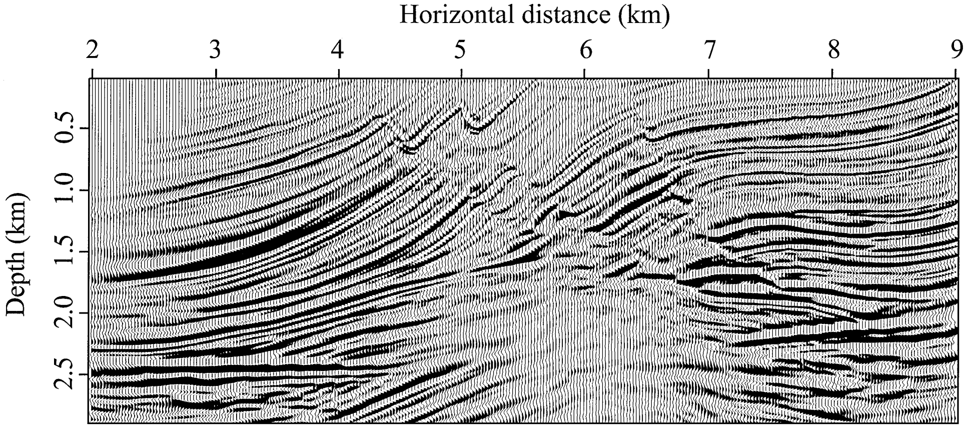


FIG. 16. First arrival migrated section of Marmousi data set (relative impedance perturbation, $\delta I/I_0$). With respect to Figure 14, no CPU-efficient interpolations were used in the target or at the surface. The clip was fixed to 0.05 (about a tenth of the maximum value).

energy of the data. Moreover, the singularities predicted by asymptotic ray theory in the presence of caustics affect the quality of the image. The effects of these singularities on the image are identified clearly by a bright spot at $x = 5000$ m and $z = 1000$ m on the strongest arrival migrated section (Figure 19).

However, we must also recognize that when compared to the first arrival migrated section (Figure 12), there is an obvious improvement of the image within the zone of the hydrocarbon trap. This improvement is more evident on the wiggle display shown on Figure 19 and plotted with the same clip as the one on Figure 14. This result confirms those obtained by Ettrich and Gajewski (1996), who concluded that there is an advantage in migrating with the strongest arrival with respect to the first arrival (see on Figure 20 the residual section between the

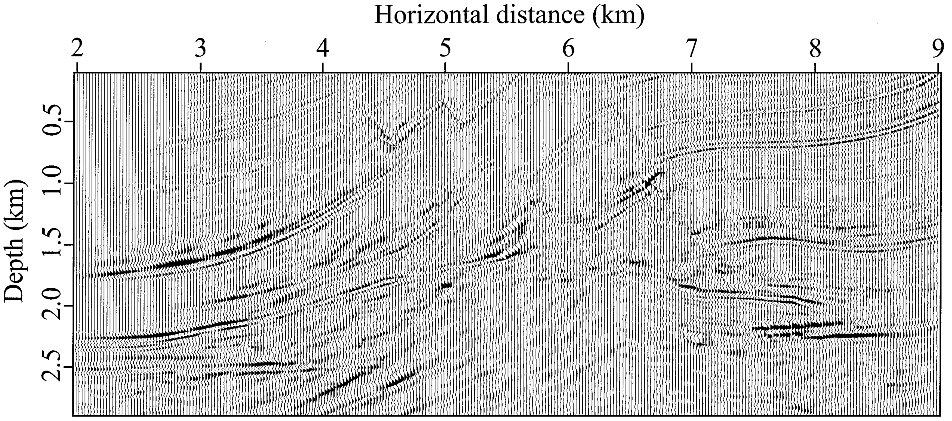


FIG. 17. Residual section between the first arrival migrated sections of Marmousi data set computed with and without the CPU-efficient interpolations (Figures 14 and 16, respectively). The clip was fixed to 0.05 (about a tenth of the maximum value).

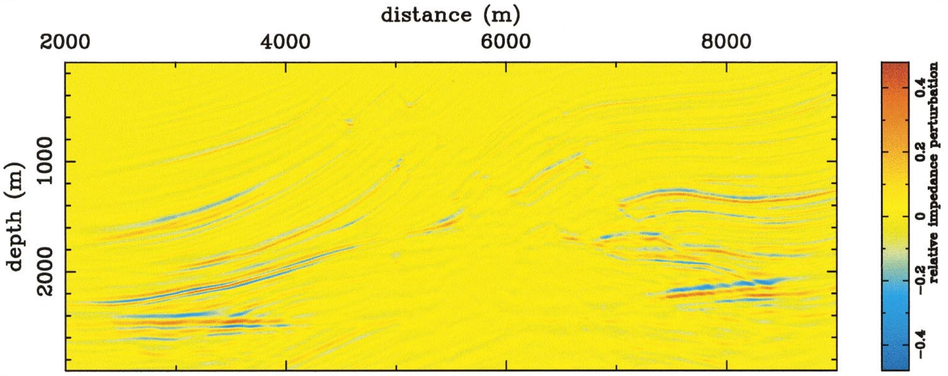


FIG. 18. Strongest arrival migrated section of Marmousi data set (relative impedance perturbation $\delta I/I_0$). Compare with the exact section on Figure 13 and with the first arrival migrated section on Figure 12.

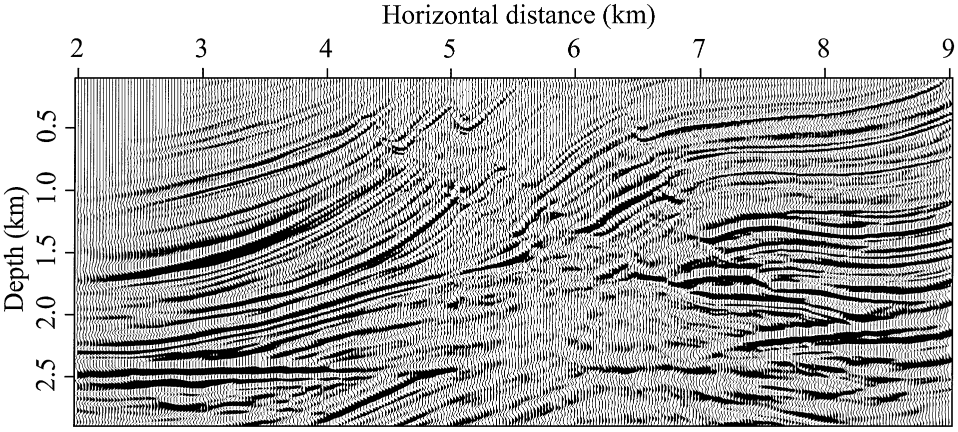


FIG. 19. Strongest arrival migrated section of Marmousi data set (relative impedance perturbation $\delta I/I_0$). The hydrocarbon trap appears clearly when compared to the first arrival migrated section (Figure 14).

strongest and the first arrival sections plotted with the same 0.05 clip, about a tenth of the maximum value). As expected, the images of Figures 12 and 18 essentially differ in the deep part of the image. Indeed, caustics develop in the shallow part of the model only at significant offsets and thus do not contribute significantly to the image.

The improvement of quantitative imaging near the hydrocarbon trap is still more obvious when comparing (Figure 21) the deepest part of the 1-D sections at 6200 m extracted from the true model (Figure 13), the first arrival (Figure 12), and the strongest arrival sections (Figure 18). The amplitudes are higher on the strongest arrival section than on the first arrival section. This suggests that the strongest arrival covers the specular reflections better (i.e., the reflections which focus more energy on the migrated section) than the first arrival.

Sensitivity of the strongest arrival migration/inversion to the interpolations.—The strongest arrival migrated section computed without interpolation is shown on Figure 22. In fact, the interpolations do not affect the image dramatically (compare Figures 19 and 22), but we see that the residuals are higher than those obtained when migrating the first arrival (compare Figures 17 and 23). It is simply due to the strongest discontinuities of the ray parameters maps (traveltime, amplitude, etc.), when using the strongest arrival (Lambaré et al., 1996). On such maps, linear interpolations have more severe effects than on the smoother maps of first arrival and thus may be very bad approximations. In compensation, interpolations allow us to pragmatically smooth the singularities predicted by ray theory at caustics.

Despite these limitations, we must recognize that using the strongest arrival with interpolations yields better results than the first arrival.

COMPUTING TIME

We can now compare our computing time with those obtained with other prestack depth migration codes. Keep in mind that on a Sun SPARC 20 worstation, 10 minutes of computing time was required to migrate the structurally simple real data set (398 shots with 120 groups of receivers and a grid of

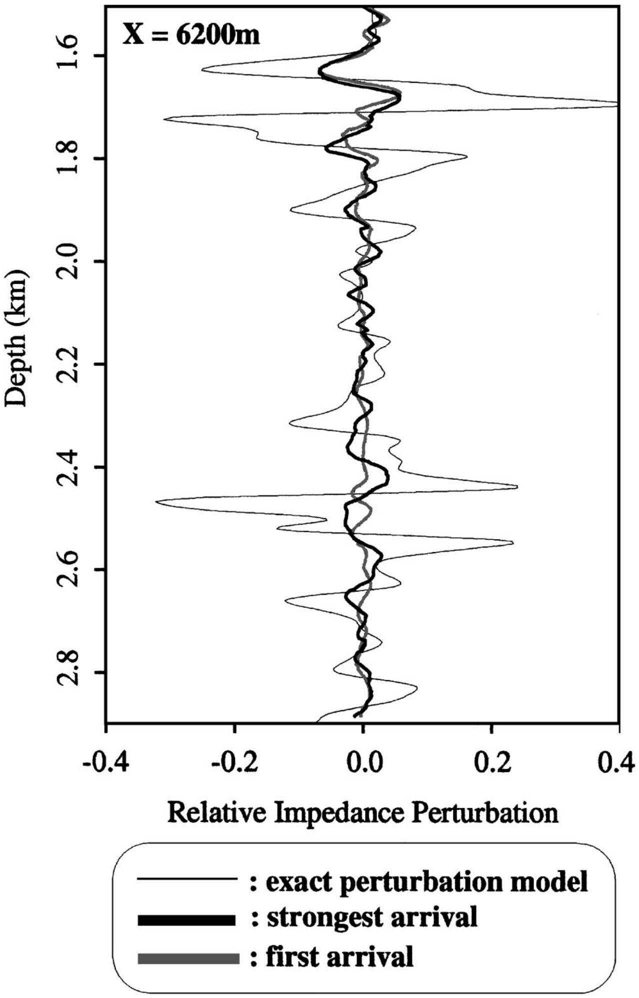


FIG. 21. Comparison among 1-D depth sections ($x = 200$ m) of relative impedance perturbation extracted from exact perturbation Marmousi model (Figure 13), from the strongest arrival (Figure 18), and from the first arrival migrated sections (Figure 12).

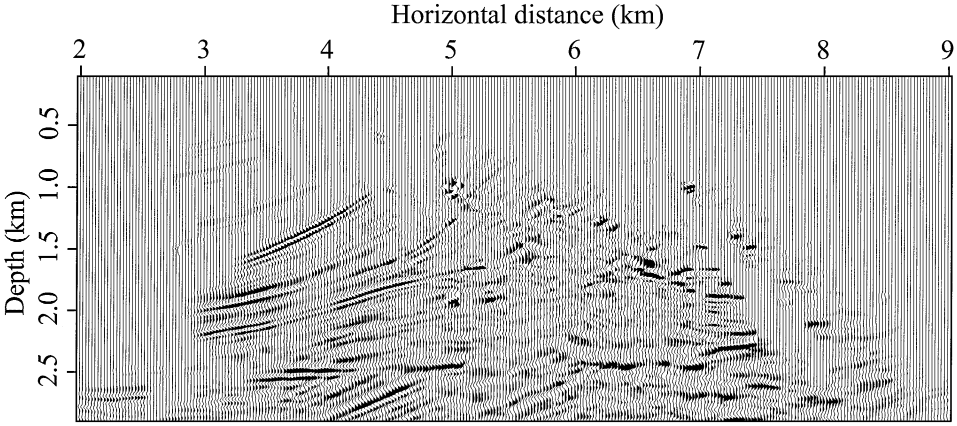


FIG. 20. Residual section between the strongest arrival (Figure 19) and the first arrival migrated sections (Figure 14).

560 × 136), whereas 22 minutes was required to migrate the whole complex Marmousi profile (240 shots with 96 receivers and a grid of 449 × 281 points).

From the literature, we list here some other computing times for migrating the Marmousi data set. Ettrich and Gajewsky (1996), reported 4 hours on a CONVEX C210 using a preserved amplitude Kirchhoff migration code using either the first or the strongest arrivals. Results were interesting and demonstrated the higher quality imaging achieved using the strongest arrival. However, no precise quantitative evaluation of the result with the exact model was proposed.

Using a common-offset wave-equation migration, Ehinger et al. (1996) reported a computing time of 40 minutes and a memory requirement of 35 Mwords on a Fujitsu VP2400 computer. The image was excellent even in the hydrocarbon trap, and wave-equation migration is considered by many as the favorite method for imaging complex structures. Unfortunately, the numerical extension of the method to three dimensions remains difficult with current technology, whereas 3-D preserved amplitude ray+Born and Kirchhoff migration codes have already been applied on real data sets (Thierry et al., 1996, 1999; Operto et al., 1997; Tura et al., 1997). Moreover, wave-equation migration is not a migration/inversion approach and, consequently, it does not provide a reliable quantitative estimation of the reflectivity or of the model perturbation. On the contrary, preserved amplitude ray+Born or Kirchhoff migration can be used for efficient amplitude variation with offset migration.

CONCLUSION

We showed that real sized 2-D and 2.5-D ray+Born migration/inversion can be done on standard workstations in a relatively short computing time (about 20 minutes), equivalent to the computing time obtained for standard industrial 2-D prestack depth migrations on supercomputers (Ehinger et al., 1996). The 2-D migration/inversion code appears robust, and the traditional instabilities of ray tracing codes are now overcome with the introduction of the wavefront construction method (Vinje et al., 1993a) associated with the uniform ray density criterion (Lambaré et al., 1996).

The power of the method in terms of quantitative imaging has also been demonstrated. Preserved amplitude migrated

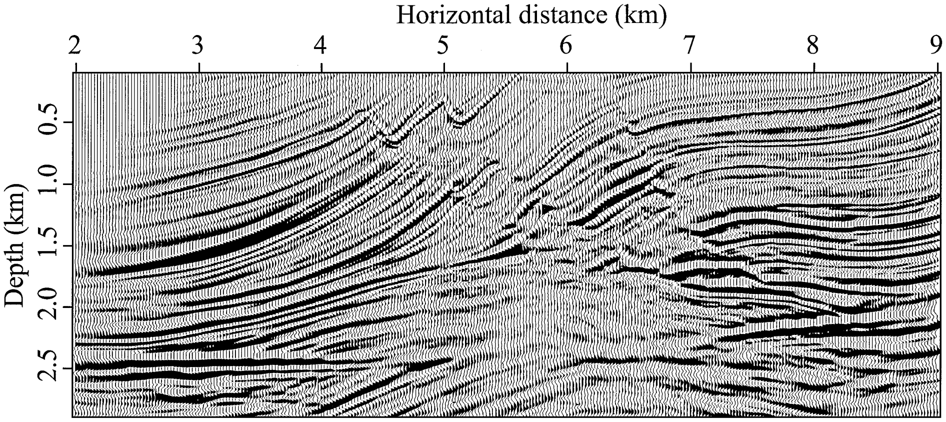


FIG. 22. Strongest arrival migrated section of Marmousi data set computed without CPU-efficient interpolations. The clip was fixed to 0.05 (about a tenth of the maximum value).

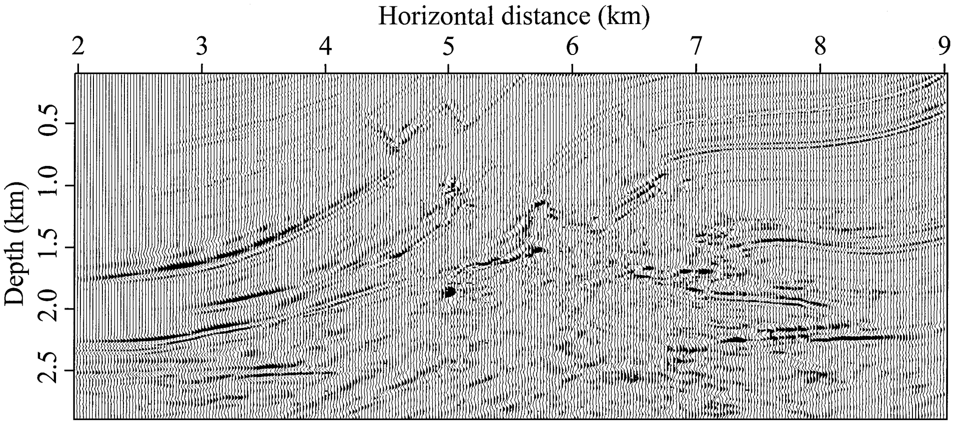


FIG. 23. Residual section between the strongest arrival sections computed without and with CPU-efficient interpolations (Figures 22 and 19, respectively). The clip was fixed to 0.05 (about a tenth of the maximum value of the migrated sections.). The energy [see expression (16)] of the residual section is 8.77% of the energy of the section obtained without interpolations (Figure 22).

sections were computed and plotted without any gain correction, and comparisons with exact models, when available, revealed quite good agreement. In the case of a complex model, taking into account the strongest arrival yields the best images. However, the quality and the amplitude of the migrated section remain poorer in zones of triplications. Using all the arrivals should improve the result, but it is not clear if we could avoid the introduction of globally valid asymptotic Green's functions as Gaussian beams or Maslov summations to overcome the artifacts of standard ray theory at caustics.

CPU efficiency is very promising for improving velocity analysis by methods such as CRP focusing gathers (Jin and Madariaga, 1993, 1994), which is a very expensive task. Concerning the extension to three dimensions, our 2-D code has fullfilled its role as testing platform because a fast 3-D ray+Born migration/inversion code has already been developed and applied to real data on a workstation (Thierry et al., 1996, 1999; Operto et al., 1997).

ACKNOWLEDGMENTS

This work was partly funded by the European Commission and Norwegian Research Council in the framework of the JOULE project "3-D Asymptotic Seismic Imaging." We thank Jan Pajchel from Norsk Hydro for providing us with the preprocessed data set and for fruitful discussions and collaboration. We thank Mark Noble, Pascal Podvin, and Henri Calandra for helpful discussions.

REFERENCES

Aki, K., and Richards, P., 1980, Quantitative seismology: Theory and Methods: W. H. Freeman & Co.

Aminzadeh, F., Burehard, N., Nicoletis, L., Rocca, F., and Wyatt, K., 1994, SEG/EAEG 3-D modeling project: First report: The Leading Edge, **13**, 110–111.

Audebert, F., Nichols, D., Rekdal, T., Biondi, B., Lumley, D., and Urdaneta, H., 1997, Imaging complex geologic structure with single-arrival Kirchhoff prestack depth migration: Geophysics, **62**, 1533–1543.

Beydoun, W. B., and Jin, S., 1994, Born or Kirchhoff migration/inversion: What is the earth's point of view?, *in* Hassanzadeh, S., Ed., Mathematical methods in geophysical imaging II: SPIE, 82–87.

Beydoun, W. B., and Mendes, M., 1989, Elastic ray-Born ℓ^2-migration/inversion: Geophys. J., **97**, 151–160.

Beylkin, G., 1985, Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform: J. Math. Phys., **26**, 99–108.

Beylkin, G., and Burridge, R., 1990, Linearized inverse scattering problems in acoustics and elasticity: Wave Motion, **12**, 15–52.

Bleistein, N., 1987, On the imaging of reflectors in the earth: Geophysics, **52**, 931–942.

Bleistein, N., and Gray, S., 1985, An extension of the Born inversion method to a depth dependent reference profile: Geophys. Prosp., **33**, 999–1022.

Bourgeois, A., Bourget, M., Lailly, P., Poulet, P., Ricarte, P., and Versteeg, R., 1991, Marmousi, model and data, *in* The Marmousi experience: Eur. Assn. Expl. Geophys., 5–16.

Bulant, P., 1996, Two-point ray tracing in 3-D: Pageoph, **148**, 421–447.

Cao, D., Beydoun, W. B., Singh, S., and Tarantola, A., 1990, A simultaneous inversion for background velocity and impedance maps: Geophysics, **55**, 458–469.

Červený, V., Molotkov, I. A., and Psencik, I., 1977, Ray theory in seismology: Charles University Press.

Chapman, C. H., 1985, Ray theory and its extensions: WKBJ and Maslov seismogram: J. Geophys., **58**, 27–43.

Chilcoat, S. R., and Hildebrand, S. T., 1995, Wavefront construction in 3-D: 65rd Ann. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1255–1257.

Claerbout, J. F., 1970, Coarse grid calculations of wave in inhomogeneous media with application to delineation of complicated seismic structure: Geophysics, **35**, 407–418.

——— 1976, Fundamentals of geophysical data processing: McGraw-Hill Book Co.

Clar, S., Ettrich, N., and Ruhl, T., 1996, Can we image complex structures using smoothed macro-velocity models?: 66rd Ann. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 543–546.

Clayton, R. W., and Stolt, R., 1981, A Born-WKBJ inversion method for acoustic reflection data: Geophysics, **46**, 1558–1565.

Cohen, J. K., and Bleistein, N., 1979, Velocity inversion procedure for acoustic waves: Geophysics, **44**, 1077–1087.

Cohen, J. K., and Hagin, F., 1985, Velocity inversion using a stratified reference: Geophysics, **50**, 1689–1700.

Cohen, J. K., Hagin, F., and Bleistein, N., 1986, Three-dimensional Born inversion with an arbitrary reference: Geophysics, **51**, 1552–1558.

Crase, E., Pica, A., Noble, M., McDonald, J., and Tarantola, A., 1990, Robust elastic non-linear waveform inversion: Application to real data: Geophysics, **55**, 527–538.

Deans, S. R., 1983, The Radon transform and some of its applications: John Wiley & Sons Inc.

Ehinger, A., Lailly, P., and Marfurt, K., 1996, Green's function implementation of common-offset wave equation migration: Geophysics, **61**, 1813–1821.

Ettrich, N., and Gajewski, D., 1996, Wavefront construction in smooth media for prestack depth migration: Pageoph, **148**, 481–502.

Farra, V., and Madariaga, R., 1987, Seismic waveform modeling in heterogeneous media by ray perturbation theory: J. Geophys. Res., **92**, 2697–2712.

Forgues, E., and Lambaré, G., 1997, Parameterization study for acoustic and elastic ray+Born inversion: J. Seis. Expl., **6**, 253–278.

French, W. S., 1975, Computer migration of oblique seismic reflection profiles: Geophysics, **40**, 961–980.

Geoltrain, J., and Brac, S., 1993, Can we image complex structures with first-arrival traveltime?: Geophysics, **58**, 564–575.

Ikelle, L., Diet, J. P., and Tarantola, A., 1986, Linearized inversion of multioffset seismic reflection data in the ω-k domain: Geophysics, **51**, 1266–1276.

——— 1988, Linearized inversion of multioffset seismic reflection data in the ω-k domain: Depth-dependent reference medium: Geophysics, **53**, 50–64.

Jin, S., and Madariaga, R., 1993, Background velocity inversion with a genetic algorithm: Geophys. Res. Lett., **20**, 93–96.

——— 1994, Nonlinear velocity inversion by a two-step Monte Carlo method: Geophysics, **59**, 577–590.

Jin, S., Madariaga, R., Virieux, J., and Lambaré, G., 1992, Two-dimensional asymptotic iterative elastic inversion: Geophys. J. Internat., **108**, 575–588.

Kolb, P., Collino, F., and Lailly, P., 1986, Prestack inversion of 1-D medium: Inst. Electr. Electron. Eng., Extended Abstracts, **74**, 498–508.

Lambaré, G., Lucio, P. S., and Hanyga, A., 1996, Two-dimensional multivalued traveltime and amplitude maps by uniform sampling of ray field: Geophys. J. Internat., **125**, 584–598.

Lambaré, G., Virieux, J., Madariaga, R., and Jin, S., 1992, Iterative asymptotic inversion of seismic profiles in the acoustic approximation: Geophysics, **57**, 1138–1154.

Lucio, P. S., Lambaré, G., and Hanyga, A., 1996, 3-D multivalued travel time and amplitude maps: Pageoph, **148**, 449–479.

Lumley, D. E., 1993, Angle-dependent reflectivity estimation: 63st Ann. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 746–749.

Mispel, J., and Hanitzsch, C., 1996, The use of layered or smoothed velocity models for prestack Kirchhoff depth migration: 66th Ann. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 519–521.

Moser, T., and Pajchel, J., 1995, Inversion in a nonsmooth background: 57th Ann. Mtg., Eur. Ass. Expl. Geophys. Extended Abstracts. E036.

——— 1997, Recursive seismic ray modeling: application in inversion and VSP: Geophys. Prosp., **45**, 885–908.

Nichols, D., 1996, Maximum energy traveltimes calculated in the seismic frequency band: Geophysics, **61**, 253–263.

Noble, M., Marsset, B., Missiaen, T., and Versteeg, W., 1996, Near surface 2D and 3D data processing—Beyond stack: 58th Ann. Mtg. Eur., Ass. Expl. Geophys., Expanded Abstracts, M033.

Operto, S., Lambaré, G., Podvin, P., and Thierry, P., 1997, 3D preserved amplitude prestack imaging of the Overthrust model: 59th Ann. Mtg., Eur. Assoc. Geoscientists Engin., Expanded Abstracts. A043.

Pica, A., 1997, Fast and accurate finite-difference solutions of the 3-D eikonal equation parameterized in celerity: 67th Ann. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1774–1777.

Pica, A., Diet, J. P., and Tarantola, A., 1990, Nonlinear inversion of seismic reflection data in laterally invariant medium: Geophysics, **55**, 284–292.

Plessix, R. E., 1996, Détermination de la vitesse pour l'interprétation de données sismiques très haute résolution à l'échelle géotechnique: Ph.D. thesis, Université Paris IX Dauphine.

Podvin, P., and Lecomte, I., 1991, Finite difference computation of traveltimes in very contrasted velocity model: A massively parallel approach and its associated tools: Geophys. J. Internat., **105**, 271–284.

Rekdal, T., and Biondi, B., 1994, Paraxial ray tracing in the marmousi model: 64th Ann. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1405–1409.

Ribodetti, A., Virieux, J., and Durand, S., 1995, Asymptotic theory for viscoacoustic seismic imaging: 65th Ann. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 631–634.

Schneider, W. A., 1978, Integral formulation for migration in two and three dimensions: Geophysics, **43**, 49–76.

Schleicher, J., Tygel, M., and Hubral, P., 1993, 3-D true-amplitude finite-offset migration: Geophysics, **58**, 1112–1126.

Sevink, A., and Herman, G., 1997, Three-dimensional inverse scattering versus imaging methods.: J. Seis. Expl., **6**, 227–238.

Stolt, R. H., 1978, Migration by fourier transform: Geophysics, **43**, 23–48.

Tarantola, A., 1984, Linearized inversion of seismic reflection data: Geophys. Prosp., **32**, 998–1015.

——— 1987, Inverse problem theory: Methods for data fitting and model parameter estimation: Elsevier.

ten Kroode, A. P. E., Smith, D. J., and Verdel, A. R., 1994, Linearized inverse scattering in the presence of caustics, *in* Hassanzadeh, S., Ed., Mathematical methods in geophysical imaging II: SPIE, 28–42.

Thierry, P., and Lambaré, G., 1995, 2.5-D true amplitude migration on a workstation: 65th Ann. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 156–159.

Thierry, P., Lambaré, G., Podvin, P., and Noble, M., 1996, 3D prestack preserved amplitude migration: application to real data: 66th Ann. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 555–558.

——— 1999, 3-D preserved amplitude prestack depth migration on a workstation: Geophysics, **64**, 222–229.

Tura, A., Hanitzsch, C., and Calandra, H., 1997, 3-D AVO migration/inversion of field data: J. Seis. Expl., **6**, 117–125.

Versteeg, R. J., 1991, Analyse du problème de la détermination du modèle de vitesse pour l'imagerie sismique: Ph.D. thesis, Université Paris VII.

Versteeg, R. J., and Grau, G., Eds., 1991, The Marmousi experience: Proceedings of the 1990 EAEG workshop on Practical Aspects of Seismic Data Inversion: Eur. Ass. Expl. Geophys.

Vidale, D., 1988, Finite-difference calculation of travel time: Bull. Seis. Soc. Am., **78**, 2062–2076.

Vinje, V., Iversen, E., Astebøl, K., and Gjøystdal, H., 1996, Estimation of multivalued arrivals in 3D models using wavefront construction—Part I: Geophys. Prosp., **44**, 819–842.

Vinje, V., Iversen, E., and Gjøystdal, H., 1993a, Traveltime and amplitude estimation using wavefront construction: Geophysics, **58**, 1157–1166.

——— 1993b, Estimation of multivalued arrivals in 3-D models using wavefront construction: 55th Ann. Mtg., Eur. Ass. Expl. Geophys., Extended Abstracts, B019.

Vinje, V., Iversen, E., Gjøystdal, H., and Astebøl, K., 1995, 3-D ray modeling by wavefront construction in open models: 65th Ann. Mtg., Soc. Expl. Geophys., Extended Abstracts, 1266–1268.

APPENDIX A

RAY+BORN APPROXIMATION

In this section, we develop a solution of the linearized forward problem using asymptotic Green's functions and the Born approximation.

Consider the scalar wave equation. The Green's function $\mathcal{G}_0(\mathbf{x}, t; \mathbf{s})$ ($\mathbf{s}$ denotes source position, $\mathbf{x}$ receiver position, and t the time) is solution of

$$\left(\frac{1}{c_0^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right)\mathcal{G}_0(\mathbf{x}, t; \mathbf{s}) = \delta\,(\mathbf{x}-\mathbf{s})\delta(t), \tag{A-1}$$

where $c_0(\mathbf{x})$ is the wave propagation velocity. It satisfies a reciprocity property $\mathcal{G}_0(\mathbf{x}, t; \mathbf{s}) = \mathcal{G}_0(\mathbf{s}, t; \mathbf{x})$.

Ray theory gives a high-frequency asymptotic approximation of the Green's function (Červený et al., 1977). In the frequency domain,

$$\mathcal{G}_0(\mathbf{x}, \omega; \mathbf{s}) = S(\omega) \sum_{n=1}^{N(\mathbf{x},\mathbf{s})} A_n(\mathbf{x}, \mathbf{s})\, e^{i\omega T_n(\mathbf{x},\mathbf{s}) - i\frac{\pi}{2}\,\mathrm{sign}(\omega)\,\alpha_n(\mathbf{x},\mathbf{s})}, \tag{A-2}$$

where n denotes the number of the ray branch, N the total number of ray branches, A the amplitude, T the traveltime, S the Green's function dimensional signature, ω the angular frequency, and α the KMAH index (Chapman, 1985). Our conventions for the Fourier transform are

$$f(\omega) = \int_{-\infty}^{+\infty} dt\; f(t)\, e^{i\omega t} \quad \text{and} \quad f(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega\; f(\omega)\, e^{-i\omega t}. \tag{A-3}$$

In the 2-D case, S and A are expressed as

$$S_{2\text{-}D}(\omega) = \frac{1}{\sqrt{-i\omega}} \quad \text{and} \quad A^{2\text{-}D}(\mathbf{x}, \mathbf{s}) = \sqrt{\frac{c_0(\mathbf{x})}{8\pi\,\mathcal{J}_{2\text{-}D}(\mathbf{x}, \mathbf{s})}}, \tag{A-4}$$

where $\mathcal{J}_{2\text{-}D}(\mathbf{r}, \mathbf{s})$ denotes the 2-D geometrical spreading associated with the 2-D asymptotic Green's functions

$$\mathcal{J}_{2\text{-}D}(\mathbf{x}, \mathbf{s}) = \frac{\partial \mathbb{L}(\mathbf{x})}{\partial \phi(\mathbf{s})}, \tag{A-5}$$

where $\partial\mathbb{L}(\mathbf{r})$ is the length of the elementary orthogonal cross-section of the ray tube in two dimensions, and $\partial\phi(\mathbf{s})$ the associated initial elementary angular aperture. All the parameters involved in the asymptotic Green's function (A-2) may be estimated along rays by the integration of ordinary differential equations.

The Born approximation is a linear approximation of the relation connecting the data $\mathcal{G}$ to the model $c(\mathbf{x})$. Consider a perturbation $\delta m(\mathbf{x})$ of the reference squared slowness model, $1/c_0^2(\mathbf{x})$, and the associated perturbation of data, $\delta\mathcal{G}$. The first-order Born approximation yields

$$\delta\mathcal{G}(\mathbf{r}, \omega; \mathbf{s}) = \omega^2 \int d\mathbf{x}\, \delta m(\mathbf{x})\, \mathcal{G}_0(\mathbf{r}, \omega; \mathbf{x})\, \mathcal{G}_0(\mathbf{x}, \omega; \mathbf{s}). \tag{A-6}$$

Introducing the asymptotic Green's functions (A-2) in the Born approximation (A-6) provides the ray+Born approximation

$$\delta\mathcal{G}(\mathbf{r}, \omega; \mathbf{s}) = \mathcal{K}(\omega) \int d\mathbf{x}\, \delta m(\mathbf{x}) \times \sum_{n=1}^{N}\sum_{\ell=1}^{L} \mathcal{A}_{n\ell}(\mathbf{r}, \mathbf{x}, \mathbf{s})\, e^{i\omega \mathcal{T}_{n\ell}(\mathbf{r},\mathbf{x},\mathbf{s})}, \tag{A-7}$$

where

$$\begin{cases} \mathcal{A}_{n\ell}^{2\text{-}D}(\mathbf{r}, \mathbf{x}, \mathbf{s}) = A_n(\mathbf{r}, \mathbf{x})A_\ell(\mathbf{x}, \mathbf{s}) \\ \mathcal{T}_{n\ell}(\mathbf{r}, \mathbf{x}, \mathbf{s}) = T_n(\mathbf{r}, \mathbf{x}) + T_\ell(\mathbf{x}, \mathbf{s}) \\ \qquad\qquad - \dfrac{\pi}{2}\dfrac{\mathrm{sign}\,(\omega)}{\omega}(\alpha_n(\mathbf{r}, \mathbf{x}) + \alpha_\ell(\mathbf{x}, \mathbf{s})) \\ \mathcal{K}^{2\text{-}D}(\omega) = i\omega \end{cases}. \tag{A-8}$$

$\mathcal{K}$, $\mathcal{A}$, and $\mathcal{T}$ denote the signature, the amplitude, and the phase of the Born operator associated with the ray branches n and ℓ of the Green's function of the source and the receiver, respectively.

The 2.5-D case (3-D in plane propagation for a crossline-invariant 3-D model) can be seen as a variation of the 2-D formula. It can be obtained from the 3-D ray+Born expression by approximating the integral along the invariant direction y with the stationary phase approximation (Bleistein, 1987),

$$\begin{aligned} \delta\mathcal{G}^{2.5\text{-}D}(\mathbf{r},\omega;\mathbf{s})\big|_{(y=0)} &= \mathcal{K}^{2.5\text{-}D}(\omega) \\ &\times \int_{(y=0)} d\mathbf{x}\,\delta m(\mathbf{x}) \sum_{n=1}^{N(\mathbf{r},\mathbf{x})} \sum_{\ell=1}^{L(\mathbf{x},\mathbf{s})} \mathcal{A}_{n\ell}^{2.5\text{-}D}(\mathbf{r},\mathbf{x},\mathbf{s})\, e^{i\omega\mathcal{T}_{n\ell}(\mathbf{r},\mathbf{x},\mathbf{s})}, \end{aligned} \tag{A-9}$$

with

$$\begin{cases} \mathcal{A}_{n\ell}^{2.5\text{-}D}(\mathbf{r},\mathbf{x},\mathbf{s}) = \dfrac{\mathcal{A}_{n\ell}^{2\text{-}D}(\mathbf{r},\mathbf{x},\mathbf{s})}{\sqrt{2\pi(\tau_n(\mathbf{r},\mathbf{x})+\tau_l(\mathbf{x},\mathbf{s}))}} \\ \mathcal{K}^{2.5\text{-}D}(\omega) = \dfrac{1}{\sqrt{-i\omega}}\omega^2 \end{cases}, \tag{A-10}$$

where τ is the parameter estimated along the ray by integrating $d\tau = dT\; c_0^2(\mathbf{x})$.

APPENDIX B

ASYMPTOTIC DIAGONALIZATION OF HESSIAN

Consider the expression of the Hessian matrix [equation (6)]. Let us first introduce Beylkin's approximations. In the context of high-frequency asymptotic approximations, only stationary phase contributions are important. The leading terms of the high-frequency asymptotic approximation of the Hessian only depend on the neighborhood of the fixed point $\mathbf{x}_0$ (ten Kroode et al., 1994). Asymptotically, the following approximations can be done,

$$\begin{cases} \mathcal{D}(\mathbf{r},\mathbf{x},\mathbf{x}_0,\mathbf{s},\omega) \approx \mathcal{D}(\mathbf{r},\mathbf{x}_0,\mathbf{x}_0,\mathbf{s},\omega) \\ \omega\Delta\mathcal{T}(\mathbf{r},\mathbf{x},\mathbf{x}_0,\mathbf{s},\omega) \approx \omega\nabla_{\mathbf{x}_0}\mathcal{T}(\mathbf{r},\,\mathbf{x}_0,\mathbf{s})\cdot(\mathbf{x}-\mathbf{x}_0), \\ \qquad\qquad = \mathbf{k}\cdot(\mathbf{x}-\mathbf{x}_0) \end{cases} \tag{B-1}$$

where we define the vectors $\mathbf{k}$ and $\mathbf{q}$,

$$\mathbf{k} = \omega\mathbf{q} \quad \text{and} \quad \mathbf{q} = \nabla_{\mathbf{x}_0}\mathcal{T}(\mathbf{r},\mathbf{x}_0,\mathbf{s}) = \mathbf{p_r} + \mathbf{p_s}, \tag{B-2}$$

where $\mathbf{p_r} = \nabla_{\mathbf{x}_0}T(\mathbf{x}_0,\mathbf{r})$ and $\mathbf{p_s} = \nabla_{\mathbf{x}_0}T(\mathbf{x}_0,\mathbf{s})$ are the slowness vectors in $\mathbf{x}_0$ for the rays $[\mathbf{r}\to\mathbf{x}_0]$ and $[\mathbf{s}\to\mathbf{x}_0]$ (see Figure 1).

We then transform the discrete summation in the Hessian (6) into an integral and change the parameters of integration from $\mathbf{s}$, $\mathbf{r}$ and ω to Θ and $\mathbf{k}$, where Θ is the angle between vectors $\mathbf{p_r}$ and $\mathbf{p_s}$ (Figure 1),

$$\mathcal{B}^{\dagger}\mathcal{Q}\mathcal{B}(\mathbf{x},\mathbf{x}_0)$$

$$= \sum_{\mathbf{s}}\sum_{\mathbf{r}}\sum_{\Omega} \mathcal{Q}\mathcal{D}(\mathbf{r},\mathbf{x},\mathbf{x}_0,\mathbf{s},\omega)\,e^{-i\omega\Delta\mathcal{T}(\mathbf{r},\mathbf{x},\mathbf{x}_0,\mathbf{s})} \tag{B-3}$$

$$\approx \int d\mathbf{s}\int d\mathbf{r}\int d\omega \frac{\mathcal{Q}\,\mathcal{D}(\mathbf{r},\mathbf{x}_0,\mathbf{x}_0,\mathbf{s},\omega)}{\Delta\mathbf{s}\,\Delta\mathbf{r}\,\Delta\omega}\,e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}_0)} \tag{B-4}$$

$$\approx \int d\Theta \frac{1}{(2\pi)^2}\iint d\mathbf{k}\,\mathcal{P}(\mathbf{s},\mathbf{r},\omega,\mathbf{x}_0)\,e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}_0)}, \tag{B-5}$$

where $\Delta\mathbf{s}$ and $\Delta\mathbf{r}$ denote the shot and receiver step along the line, $\Delta\omega$ is the sampling step for the angular frequency, and

$$\begin{aligned} \mathcal{P}(\mathbf{s},\mathbf{r},\omega,\mathbf{x}_0) &= \mathcal{Q}\frac{\mathcal{D}(\mathbf{r},\mathbf{x}_0,\mathbf{x}_0,\mathbf{s},\omega)(2\pi)^2}{\Delta\mathbf{s}\,\Delta\mathbf{r}\,\Delta\omega} \\ &\times \left|\frac{\partial(\mathbf{s},\mathbf{r},\omega)}{\partial(\mathbf{k},\Theta)}\right|. \end{aligned} \tag{B-6}$$

Expression (B-5) must be compared to equation (8) and, following Jin et al. (1992), we choose $\mathcal{Q}$ as a local weight depending on the imaged point $\mathbf{x}_0$ in the model in such a way that $\mathcal{P}(\mathbf{s},\mathbf{r},\omega,\mathbf{x}_0) = 1$, which imposes that

$$\mathcal{Q}(\mathbf{s},\mathbf{r},\omega,\mathbf{x}_0) = \frac{\Delta\mathbf{s}\,\Delta\mathbf{r}\,\Delta\omega}{(2\pi)^2\mathcal{D}(\mathbf{r},\mathbf{x}_0,\mathbf{x}_0,\mathbf{s},\omega)}\left|\frac{\partial(\mathbf{k},\Theta)}{\partial(\mathbf{s},\mathbf{r},\omega)}\right|. \tag{B-7}$$

Introducing expression (B-7) in equation (B-5) gives for the Hessian

$$\begin{aligned} \mathcal{B}^{\dagger}\mathcal{Q}\mathcal{B}(\mathbf{x},\mathbf{x}_0) &\approx \int d\Theta\,\frac{1}{(2\pi)^2}\iint d\mathbf{k}\,e^{-i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}_0)} \\ &\approx [\Theta]_{\min}^{\max}\,\delta(\mathbf{x}-\mathbf{x}_0). \end{aligned} \tag{B-8}$$

Here, $[\Theta]_{\min}^{\max}$ denotes the gap between the maximum and minimum value of Θ obtained by considering the set of rays $\mathbf{s}\to\mathbf{x}_0$ and $\mathbf{r}\to\mathbf{x}_0$ for all the traces. The Hessian matrix has now been approximately diagonalized and thus can be inverted easily.

APPENDIX C

DERIVATION OF THE JACOBIAN

In this appendix, we prove equation (15):

$$\left|\frac{\partial(\mathbf{k},\Theta)}{\partial(\mathbf{s},\mathbf{r},\omega)}\right| = |\omega||\mathbf{q}|^2\left|\frac{\partial(\Phi_\mathbf{s})}{\partial(\mathbf{s})}\right|\left|\frac{\partial(\Phi_\mathbf{r})}{\partial(\mathbf{r})}\right|. \qquad \text{(C-1)}$$

Let us remember that $\mathbf{k} = \omega\mathbf{q} = \omega(\mathbf{p}_s + \mathbf{p}_r)$ and that $\Phi_\mathbf{s}$ and $\Phi_\mathbf{r}$ are, respectively, the angles associated with the slowness vectors $\mathbf{p}_s$ and $\mathbf{p}_r$ (Figure 1). This result can be obtained using a polar representation for vector $\mathbf{k}$, $\mathbf{k} \to (|\mathbf{k}|, \Phi)$, where $\Phi = (\Phi_\mathbf{s} + \Phi_\mathbf{r})/2$. We get

$$\left|\frac{\partial(\mathbf{k},\Theta)}{\partial(\mathbf{s},\mathbf{r},\omega)}\right| = \left|\frac{\partial(\mathbf{k},\Theta)}{\partial(|\mathbf{k}|,\Phi,\Theta)}\right|\left|\frac{\partial(|\mathbf{k}|,\Phi,\Theta)}{\partial(\mathbf{s},\mathbf{r},\omega)}\right| \qquad \text{(C-2)}$$

$$= |\mathbf{k}|\left|\frac{\partial(|\mathbf{k}|,\Phi,\Theta)}{\partial(\mathbf{s},\mathbf{r},\omega)}\right| \qquad \text{(C-3)}$$

$$= |\omega||\mathbf{q}|\left|\frac{\partial(|\mathbf{k}|)}{\partial(\omega)}\right|\left|\frac{\partial(\Phi,\Theta)}{\partial(\mathbf{s},\mathbf{r})}\right| \qquad \text{(C-4)}$$

$$= |\omega||\mathbf{q}|^2\left|\frac{\partial(\Phi,\Theta)}{\partial(\mathbf{s},\mathbf{r})}\right|. \qquad \text{(C-5)}$$

Finally, we use

$$\begin{cases} \Phi = (\Phi_\mathbf{s} + \Phi_\mathbf{r})/2 \\ \Theta = \Phi_\mathbf{s} - \Phi_\mathbf{r} \end{cases} \qquad \text{(C-6)}$$

to get

$$\left|\frac{\partial(\Phi,\Theta)}{\partial(\mathbf{s},\mathbf{r})}\right| = \left|\frac{\partial(\Phi,\Theta)}{\partial(\Phi_\mathbf{s},\Phi_\mathbf{r})}\right|\left|\frac{\partial(\Phi_\mathbf{s},\Phi_\mathbf{r})}{\partial(\mathbf{s},\mathbf{r})}\right| \qquad \text{(C-7)}$$

$$= \left|\frac{\partial(\Phi_\mathbf{s})}{\partial(\mathbf{s})}\right|\left|\frac{\partial(\Phi_\mathbf{r})}{\partial(\mathbf{r})}\right| \qquad \text{(C-8)}$$

and to prove equation (15).