

## Two-dimensional asymptotic iterative elastic inversion

Side Jin,<sup>1</sup> Raul Madariaga,<sup>1</sup> Jean Virieux<sup>2</sup> and Gilles Lambaré<sup>1</sup>

<sup>1</sup>Laboratoire de Sismologie, Université Paris 7, Institut de Physique du Globe de Paris, F-75252 Paris 05, France

<sup>2</sup>Institut de Géodynamique, Université de Nice—Sophia Antipolis, F-06034 Nice, France

Accepted 1991 August 19. Received 1991 August 19; in original form 1991 January 31

### SUMMARY

An asymptotic linearized iterative elastic inversion method is proposed to invert 2-D Earth parameters from multicomponent data and is tested numerically. The forward problem is solved by a combination of the Born approximation and ray theoretical methods. We express the perturbed seismogram in terms of perturbations of  $P$ - and  $S$ -wave impedances and density. The inversion method is based on generalized least squares. We introduce a special form of the  $\ell^2$  norm with a weighting function that corrects for geometrical spreading and obliquity of the reflectors. The Hessian for this norm could be estimated in a closed form that is asymptotically valid at high frequencies. We propose a quasi-Newtonian iterative method for the solution of the inverse problem. The first iteration of this inversion method resembles the operator proposed by Beylkin (1985) and Beylkin & Burridge (1990) for the asymptotic inversion of seismic data. Our method is more general than theirs because it can handle arbitrary discrete distributions of sources and receivers. Elastic inversion is generally ill-posed because the problem is overdetermined but undersampled. We study the resolution of the asymptotic inversion method for general sets of sources and receivers. We show that simultaneous inversion for both  $P$ - and  $S$ -wave impedance is generally ill-conditioned if data for a single scattering mode are available. In particular, it seems that only one parameter can be reliably resolved from marine data. Simultaneous inversion for a finite set of parameters can be resolved only for multicomponent elastic data containing both  $P$ -wave and  $S$ -wave information. Inversion tests using synthetic data calculated by finite-differences demonstrates that it is possible to invert simultaneously for  $P$  and  $S$  impedances.

**Key words:** inversion, seismic reflection.

### INTRODUCTION

The inverse problem for finding the Earth's structure from seismic reflection data is currently treated in two apparently different ways. Firstly, from the point of view optimization theory, inversion attempts to find an Earth model that minimizes the difference between synthetic seismograms computed for this Earth model and the observed field data. This is the most frequently used approach (see, e.g. Tarantola 1984; Gauthier, Virieux & Tarantola 1986; Ikelle, Diet & Tarantola 1988; Mora 1987, 1988; Cao *et al.* 1990; Crase *et al.* 1990). Different definitions of the distance between observed data and synthetics lead to different optimization criterions. The advantage of this method is that it can deal with all sorts of source–receiver configurations, but it is computationally very expensive. Both finite differences (Crase *et al.* 1990) and ray-theoretical methods

(Beydoun & Mendez 1989) have been used to solve the forward problem. Often, linearization is required by the huge amount of data and parameters.

The other approach to seismic inversion is to find an approximate one-to-one relationship between the observed data and the Earth model parameters (Bleistein & Cohen 1979; Cohen & Bleistein 1979; Clayton & Stolt 1981; Beylkin 1985; Miller, Oristaglio & Beylkin 1987; Beylkin & Burridge 1990). The relation between the data and model parameters is usually approximated by ray theory so that, at the end, the data and the model can be related by a linear integral equation that can be readily inverted. These methods are computationally very attractive but they can only deal with particular data sets; in particular, it is not clear how to deal with data sets which are both redundant and incomplete, the most common situation in practice.

Both methods share stability problems: for the first

method, numerical inversion of huge matrices is always a difficult exercise, while the exact kernel inversion for the second method might lead to unphysical results. Regularization procedures exist and incorporate *a priori* information into the inversion. When linearization is applied as assumed in this paper, results may depend on the reference model from which perturbations are computed. As already observed by Gauthier *et al.* (1986), inadequate reference models lead to other minima of the cost function. It is enough to misalign reflections over half a wavelength to meet this problem for both acoustic and elastic inversions.

In this paper we present a method of linear elastic inversion that attempts to exploit the relative advantages of each of these two approaches. The inversion algorithm is derived from a generalized least-squares criterion, but we use the high-frequency approximation for the estimation of an approximate Hessian. We solve the forward problem by a combination of ray theory and Born approximation. The principal advantage of ray-Born modelling is computational efficiency. We solve the inverse problem by a least-squares iterative procedure using a quasi-Newton method. We demonstrate furthermore that the first iteration of our method is closely related to the spectral method developed by Beylkin (1985), Bleistein (1987) and Beylkin & Burridge (1990). Thus, the apparent incompatibility between the two approaches that some authors have put forward (Tarantola 1986) concerns more the philosophy of inversion than the practical algorithms proposed.

Several authors have applied elastic inversion methods to marine data. For instance, Cao *et al.* (1990) and Crase *et al.* (1990) tried to retrieve several independent parameters from a single *P* arrival; their results show that the simultaneous inversion of *P*- and *S*-wave impedances from single *P*-wave observations is very difficult in practice. The *P* and *S* impedances are strongly coupled in their inversions. For this reason, we carefully studied the resolution of multiparameter inversion and demonstrate that the inverse problem for more than one parameter is very ill-conditioned for single component data in the usual range of offsets. Thus, the full separation of two or more parameters from marine data is practically very difficult; this result was obtained by Santosa & Symes (1988) for a layered fluid. We give two examples of numerical tests to illustrate the usefulness of our method. We show that if we could use *P*- and *S*-wave arrivals simultaneously, we are able to retrieve two elastic parameters from the inversion of land data. Inadequate geometry of data acquisition, variations in diffraction coefficients between *P*-*P*, *P*-*S*, *S*-*P* are enough to allow separation between elastic parameters.

## FORWARD PROBLEM

We study the propagation of elastic waves in a medium where the material properties differ slightly from those of a reference medium with continuous smooth distribution of elastic impedance and density. The properties of the reference medium are supposed to be known, previously determined by a velocity analysis or other equivalent method. We assume that a reflection seismogram is generated by the waves scattered by small heterogeneities in the velocity and density distributions in the Earth.

The forward problem consists of finding the Green

function which is the solution of the elastodynamic equation for a point source located at  $\mathbf{s}$  (Aki & Richards 1980):

$$\rho \partial_{tt}^2 G_{ij} - \partial_{x_m} (c_{imprq} \partial_{x_q} G_{pj}) = \delta_{ij} \delta(t) \delta(\mathbf{x} - \mathbf{s}). \quad (1)$$

The solution,  $G_{ij}(\mathbf{x}, t; \mathbf{s}, 0)$  is the *i*th component of displacement at the point  $\mathbf{x}$  due to a point force in the *j*-direction applied at the source  $\mathbf{s}$ .  $\rho$  is the density and the  $c_{jmrpq}$  are the elastic parameters of the medium which is considered to be isotropic:

$$c_{ijpq} = \lambda \delta_{ij} \delta_{pq} + \mu (\delta_{ip} \delta_{jq} + \delta_{iq} \delta_{jp}). \quad (2)$$

$\lambda$  and  $\mu$  are the Lamé constants of the medium, and  $\delta_{jm}$  is Kröneckers symbol. The solution of problem (1) is only possible by numerical methods, for instance finite differences or finite elements. This is the approach adopted by Gauthier *et al.* (1986), Kolb, Collino & Lailly (1986) and several other authors. Such an approach is extremely slow and costly because it requires solving the wave propagation problem with many points per wavelengths in order to avoid the classical dispersion problems in finite differences (see, e.g. Virieux 1986). In this paper finite differences will be used to generate synthetics but inversion will be based on the ray theoretical solution of equation (1).

We assume further that the medium can be separated in two parts: firstly, a smooth reference medium on which the Green function may be computed by ray theory; and, secondly, a small-amplitude short spatial wavelength perturbation of the elastic parameters:

$$\rho = \rho_0 + \delta\rho, \quad \lambda = \lambda_0 + \delta\lambda, \quad \mu = \mu_0 + \delta\mu \quad (3)$$

where  $\rho_0$ ,  $\lambda_0$ ,  $\mu_0$  are the known smooth reference parameters and  $\delta\rho$ ,  $\delta\lambda$ ,  $\delta\mu$  the corresponding small perturbations. The effect of these small perturbations will be calculated by Born's approximation.

For a given source-receiver pair ( $\mathbf{s}$ ,  $\mathbf{r}$ ), we postulate that the Green tensor  $\mathbf{G}^0$  for the smooth reference medium can be calculated by the ray theoretical approximation. In this approximation, *P*- and *S*-waves propagate independently and  $\mathbf{G}^0$  can be expressed as

$$G_{ij}^0(\mathbf{x}, \mathbf{s}, t) = G_{ij}^p(\mathbf{x}, \mathbf{s}, t) + G_{ij}^s(\mathbf{x}, \mathbf{s}, t) \quad (4)$$

where each term on the right of (4) can be written as the following 2-D Green functions for a line source

$$G_{ij}^n(\mathbf{x}, \mathbf{s}, t) = A_{ij}^n(\mathbf{x}, \mathbf{s}) \frac{H[t - \tau^n(\mathbf{x}, \mathbf{s})]}{\sqrt{t - \tau^n(\mathbf{x}, \mathbf{s})}} \quad (5)$$

where *n* stands for the *P*-wave or the *S*-wave.  $A^n$  and  $\tau^n$  are the ray-theoretical amplitude and traveltimes for the wave of type *n*, which will be calculated by ray tracing using techniques developed, for example, by Lambaré *et al.* (1991).

The full solution of (1) can also be written as the sum of two terms, the Green function in the reference medium plus a small perturbation  $\mathbf{u}$  due to the scattering from the perturbations of elastic parameters and density  $\delta\lambda$ ,  $\delta\mu$ ,  $\delta\rho$ :

$$G_{ij} = G_{ij}^0 + u_{ij}. \quad (6)$$

The reflection seismograms calculated by perturbation theory contain all the waves singly diffracted by the heterogeneities in the medium. Including multiply diffracted waves is not possible within ray theory because the

computational effort is too big. If multiples were important it would be necessary to use numerical solutions by finite differences or finite elements.

Within first-order perturbation theory, waves scattered from heterogeneities separate into four different type of waves (Knopoff 1956):  $P$ - $P$  and  $P$ - $S$  diffraction for a  $P$ -wave incident on a heterogeneity; and  $S$ - $P$  and  $S$ - $S$  for  $S$  incident on the scatterer. As shown in Appendix A (see also Wu & Toksöz 1989) the complete reflection seismogram is the sum of five terms:

$$u_{ij}(\mathbf{r}, \mathbf{s}, t) = \sum_{\ell=1}^5 \gamma_i \epsilon(\mathbf{r}) \mathcal{R}_j \epsilon(\mathbf{s}) u_{\ell}(\mathbf{r}, \mathbf{s}, t), \quad (7)$$

where  $u_{\ell}$  is a component of the two-way Green function with  $\ell$  as a subscript for each type of scattered wave ( $1 = P$ - $P$ ,  $2 = P$ - $S$ ,  $3 = S$ - $P$ ,  $4 = SV$ - $SV$ ,  $5 = SH$ - $SH$ ).  $\gamma_{\ell}(\mathbf{r})$  is the polarization vector of the wave at the observer, longitudinal for  $P$ -waves, and transverse for  $S$ -waves.  $\mathcal{R}(\mathbf{s})$  is the radiation pattern for of  $P$ - and  $S$ -waves by the point source and it can be estimated by standard seismological methods (see, e.g. Aki & Richards 1980). The radiation pattern at the source and the polarization at the receiver have been written out explicitly because they have a key contribution for the multiparameter inversion. To first order with respect to the perturbation parameters, we find

$$u_{\ell}(\mathbf{r}, \mathbf{s}, t) = \frac{\partial}{\partial t} \int_M A_{\ell}(\mathbf{r}, \mathbf{s}, \mathbf{x}) \delta[t - \tau_{\ell}(\mathbf{r}, \mathbf{s}, \mathbf{x})] \times \sum_{i=1}^m W_{\ell i}(\mathbf{r}, \mathbf{s}, \mathbf{x}) f_i(\mathbf{x}) d\mathbf{x}, \quad (8)$$

where  $\mathbf{f}(\mathbf{x}) = [\delta\lambda(\mathbf{x}), \delta\mu(\mathbf{x}), \delta\rho(\mathbf{x})]$  is the vector of perturbation parameters. Referring to Fig. 1, (8) has a very simple geometrical interpretation: the integrand represents the diffraction by the heterogeneity located at point  $\mathbf{x}$ . A ray propagates from the source to the heterogeneity where it is diffracted and a new diffracted ray joins the scatterer to the receiver at  $\mathbf{r}$ .  $A_{\ell}$  is the product of amplitude coefficients along the trajectory from source to scatterer and back to the receiver, while  $\tau_{\ell}$  is the two-way traveltime from source to scatterer and back to the receiver.  $\mathbf{W}(\mathbf{r}, \mathbf{s}, \mathbf{x})$  is the scattering matrix which, for a given position of the scatterer, depends only on the scattering angle  $\theta$  between the incident and the

scattered ray as shown in Fig. 1. For  $P$ -to- $P$  reflection ( $\ell = 1$ ),  $A_{\ell} = A^P(\mathbf{r}, \mathbf{x})A^P(\mathbf{x}, \mathbf{s})$  and  $\tau_{\ell} = \tau^P(\mathbf{r}, \mathbf{x}) + \tau^P(\mathbf{x}, \mathbf{s})$ . For other conversion modes, similar definitions of  $A_{\ell}$  and  $\tau_{\ell}$  are deduced.

In the integral representation (8) of the scattering field, we use the Lamé constants and density for describing the Earth model. As shown by various authors (see, e.g. Tarantola 1986; Beydoun & Mendez 1989),  $P$ -wave impedance,  $S$ -wave impedance and density are better parameters for inversion than the elastic parameters. To first order, the relation between the perturbations of the Lamé parameters and the perturbations of  $P$ -wave and  $S$ -wave impedance are

$$\lambda = (IP^2 - 2IS^2)/\rho, \quad \mu = IS^2/\rho, \quad (9)$$

$$\delta\lambda = 2\alpha_0\delta IP - 4\beta_0\delta IS - (\alpha_0^2 - 2\beta_0^2)\delta\rho, \quad (10)$$

$$\delta\mu = 2\beta_0\delta IS - \beta_0^2\delta\rho.$$

Using these relations in the definition of the different terms of the scattering matrix, previous authors obtained for

$$\mathbf{f}(\mathbf{x}) = (\delta IP, \delta IS, \delta\rho) \quad (11)$$

the following diffraction coefficients:

$\ell = 1$ :  $P$ - $P$  diffraction:

$$W_{11} = 2\alpha_0^{-1}, \quad W_{12} = -4\kappa/\alpha_0 \sin^2(\theta_{pp}/2), \\ W_{13} = 2[\cos(\theta_{pp}/2) + \kappa^2 \sin^2(\theta_{pp})], \quad (12a)$$

$\ell = 2, \ell = 3$ :  $P$ - $S$  or  $S$ - $P$  diffraction:

$$W_{21} = 0, \quad W_{22} = 2\alpha_0^{-1} \sin(2\theta_{ps}), \\ W_{23} = [\sin\theta_{ps} - \kappa \sin(2\theta_{ps})], \quad (12b)$$

$\ell = 4$ :  $SV$ - $SV$  diffraction:

$$W_{41} = 0, \quad W_{42} = 2\beta_0^{-1} \cos(2\theta_{ss}), \\ W_{43} = \cos\theta_{ss} - \cos(2\theta_{ss}), \quad (12c)$$

$\ell = 5$ :  $SH$ - $SH$  diffraction:

$$W_{51} = 0, \quad W_{52} = 2\beta_0^{-1} \cos\theta_{ss}, \quad W_{53} = 1 - \cos\theta_{ss}, \quad (12)$$

where  $\kappa = \beta_0/\alpha_0$  and  $\theta$  is the scattering angle with obvious subscript notation associated with the mode of diffraction.

This completes our brief exposition of the direct problem, more details can be found in Appendix A and in Wu & Toksöz (1989) or Beylkin & Burridge (1990).

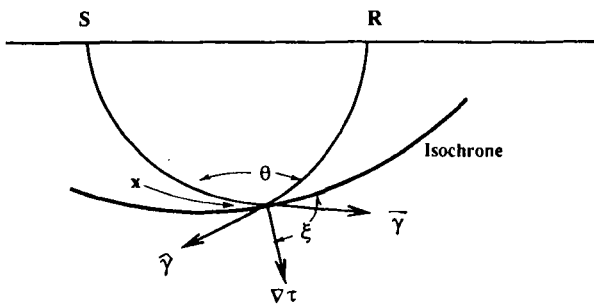
## INVERSION THEORY

### General approach to elastic inversion

From the solution of the linearized forward problem, we know that the scattered field is determined by a linear combination of Earth parameters. We can rewrite the general representation (8) for each scattering component in the Fourier domain in the following form:

$$u_{\ell}(\mathbf{r}, \mathbf{s}, \omega) = \int_M \sum_{k=1}^m A_{\ell}(\mathbf{r}, \mathbf{s}, \omega) i\omega e^{i\omega\tau_{\ell}(\mathbf{r}, \mathbf{s}, \mathbf{x})} \times W_{\ell k}(\mathbf{r}, \mathbf{s}, \mathbf{x}) f_k(\mathbf{x}) d\mathbf{x} \quad (13)$$

where  $\ell$  denotes the scattering mode ( $P$ - $P$ ,  $P$ - $S$ ,  $S$ - $P$  and  $SV$ - $SV$ ,  $SH$ - $SH$ ). Generally, if the data are generated by a



**Figure 1.** Geometry of the single scattering forward problem. The source and receiver locations are parametrized by the position  $\mathbf{s}$  and  $\mathbf{r}$ , respectively, on the recording surface. A ray from the source is scattered by the impedance heterogeneity located at point  $\mathbf{x}$ . At the scatterer we define two angles:  $\theta$  the aperture angle between the incident and diffracted rays and  $\xi$  the angle of the local normal  $\mathbf{p}$  to the isochrone of the two-way traveltime with respect to the horizon.

single vertical force source, we can use the  $P$ - $P$  and  $P$ - $S$  wave information from the data. Because (13) is a linear expression giving  $u_\ell$  from  $f_k$ , we can combine scattering components using (7) in order to obtain vertical and horizontal recorded displacements. Presenting one scattering mode simplifies notation. If only one displacement or pressure component is available, all that can be used is the  $P$ - $P$  scattering component.  $\mathbf{r}$  and  $\mathbf{s}$  are the positions of the receiver and source, respectively.  $m$  is the number of parameters in the linearized Earth model  $\mathbf{f}$  defined in (11). In principle, we may attempt to retrieve three parameters and define  $(f_1, f_2, f_3) = (\delta IP, \delta IS, \delta \rho)$ , but in practice  $\delta \rho$  can not be inverted from scattering data.

In order to properly pose the inverse problem we define model and data spaces, and the operators between these two spaces. The model space  $\mathcal{M}$  is the space of all possible perturbations of the impedance and density model; it will be parametrized by position  $\mathbf{x}$  and the  $k$ th component of the model vector  $\mathbf{f}$ . In a compact, we denote

$$\mathcal{M}:(k, \mathbf{x}). \quad (14)$$

The data space  $\mathcal{D}$  consists in all the seismograms recorded on the surface. This space will be parametrized by the indices of the source and receiver positions  $s$  and  $r$ , respectively; and by the frequency  $\omega$ . For simplicity of the analytic developments that we need, we considered that the data have been Fourier transformed to the frequency domain, but in fact at the end we will return to the time domain. All actual computations are carried out in the time domain. Thus seismograms are defined in the mixed data space:

$$\mathcal{D}:(\mathbf{r}, \mathbf{s}, \ell, \omega). \quad (15)$$

We rewrite the forward problem in the compact operator form

$$\mathbf{u} = \mathbf{G}\mathbf{f}, \quad (16)$$

where  $\mathbf{G}:\mathcal{D}\rightarrow\mathcal{M}$  is the two-way Green function, the integral operator on the right-hand side of (13). The solution of the linearized inverse problem consists in finding the inverse of operator  $\mathbf{G}$ .

### Inversion by a least-squares method

We seek an approximate solution to the inversion of (13) by an optimization method, introducing a cost function that measures the misfit between observed and calculated seismograms. Several choices for the cost function are possible; Crase *et al.* (1990) examined some of them. Here we will adopt the least-squares norm  $\mathcal{L}^2$  of the difference between observed and predicted seismograms as the least-squares criterion. We will demonstrate that this criterion leads to a quasi-Newtonian iterative solution method whose first iteration is closely related to the approximate inverse proposed by Beylkin & Burridge (1990) for the elastic problem.

In order to define the cost function we introduce the following explicit definition of inner product in data space:

$$\langle \mathbf{u} | \mathbf{v} \rangle_{\mathcal{D}} = \sum_{s,r,\ell} \int_{\Omega} d\omega u_\ell^*(\mathbf{r}, \mathbf{s}, \omega) Q_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x}_0, \omega) v_\ell(\mathbf{r}, \mathbf{s}, \omega) \quad (17)$$

where  $\mathbf{u}$  and  $\mathbf{v}$  are two sets of seismograms.  $^+$  denotes

complex conjugate. The sum in (17) extends over the data space defined by (15).  $Q$  is a diagonal covariance matrix with elements:

$$Q_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x}_0, \omega) = \frac{|\mathbf{p}_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x}_0)|^2}{4\pi^2 \omega A_\ell^2(\mathbf{r}, \mathbf{s}, \mathbf{x}_0)} \quad (18)$$

where  $\mathbf{p}_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x}_0, \omega) = \nabla \tau_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x}_0, \omega)$ , is the gradient of the two-way traveltime  $\tau_\ell$ .

The particular form of the covariance matrix  $Q_\ell$  is introduced in order to correct the amplitudes for geometrical spreading, and for the spectral contents of the Green function in two dimensions. Let us remark that the covariance function depends on  $\mathbf{x}_0$ , the coordinate of the point at which the model will be inverted. This choice is not classical, in most of the work on the optimization approach to inversion the covariance matrix is also diagonal, but it is independent of the coordinates of the point  $\mathbf{x}_0$  where the model is being estimated.  $Q$  is in fact a pre-conditioning applied to the gradient of the cost function. Other choices for a locally modulated covariance function may be interesting, but we prefer this one because it leads to an algorithm that follows closely that of Beylkin & Burridge (1990).

The covariance matrix  $Q$  given by (18) upgrades weak late arrivals. Instabilities might arise when noise exists in the late part of the seismogram or when the reference medium leads to strong defocusing in an area of the seismogram where energetic signals are observed. Summing over the data acquisition reduces considerably these incoherent instabilities. If not enough, *a priori* information as the maximum amplitude for the perturbation in the impedance parameters stabilizes the procedure. However, in practical situations, amplitudes of signal decay with time in a rather coherent way and defocusing effects are not so strong that  $A_\ell$  term goes to zero and creates instabilities (Lambaré *et al.* 1991). These instabilities have nothing to do with instabilities of exact inversion, where noise amplification are performed at each step of the layered stripping, for example, in order to fit exactly the data at the current depth.

For the inverse problem we also need a definition of inner product between any two functions  $\mathbf{f}(\mathbf{x})$  and  $\phi(\mathbf{x})$  in model space  $\mathcal{M}$ :

$$\langle \mathbf{f} | \phi \rangle_{\mathcal{M}} = \int_{\mathcal{M}} \sum_k f_k^*(\mathbf{x}) \phi_k(\mathbf{x}) d\mathbf{x}^2. \quad (19)$$

From the definition of inner product in data space (17), we obtain the  $\mathcal{L}^2$  norm, or cost function  $S(\mathbf{f}) = 1/2 \langle \mathbf{u}^{\text{obs}} - \mathbf{G}\mathbf{f} | \mathbf{u}^{\text{obs}} - \mathbf{G}\mathbf{f} \rangle_{\mathcal{D}}$  where  $\mathbf{u}^{\text{obs}}$  are the observed data and  $\mathbf{G}\mathbf{f}$  are the synthetic seismograms estimated by (13). With this cost function we formulate the inversion problem as:

$$\text{find } \mathbf{f} : \min_{\mathcal{M}} \{ S(\mathbf{f}) = 1/2 \langle \mathbf{u}^{\text{obs}} - \mathbf{G}\mathbf{f} | \mathbf{u}^{\text{obs}} - \mathbf{G}\mathbf{f} \rangle_{\mathcal{D}} \}. \quad (20)$$

Minimization of the cost function leads to the classical 'system' of normal equations:

$$\mathbf{G}^+ \mathbf{G} \mathbf{f} = \mathbf{G}^+ \mathbf{u}^{\text{obs}} \quad (21)$$

for all  $\mathbf{x}$  where  $\mathbf{G}^+$  is the adjoint function to the two-way Green function  $\mathbf{G}$ . This adjoint operator is defined by the classical relationship  $\langle \mathbf{u} | \mathbf{G}\mathbf{f} \rangle_{\mathcal{D}} = \langle \mathbf{G}^+ \mathbf{u} | \mathbf{f} \rangle_{\mathcal{M}}$ . From (13), we find the following explicit expression for the kernel of the

transposed operator  $\mathbf{G}^\dagger$

$$\mathbf{G}^\dagger(\mathbf{r}, \mathbf{s}, \mathbf{x}) = W_{k\ell}(\mathbf{r}, \mathbf{s}, \mathbf{x}) A_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x}) Q_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x}_0, \omega) \times (-i\omega) e^{-i\omega\tau_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x})}. \quad (22)$$

The formal solution of the inverse problem (21) is

$$\mathbf{f}(\mathbf{x}_0) = \mathbf{H}^{-1}(\mathbf{x}_0, \mathbf{x}) \gamma^0(\mathbf{x}) \quad (23)$$

where

$$\gamma^0 = \mathbf{G}^\dagger \mathbf{u}^{\text{obs}} \quad (24)$$

is the gradient of  $S(\mathbf{f})$  at  $\mathbf{f} = 0$ . Explicitly, the  $k$ th component of  $\gamma$  is

$$\gamma_k^0(\mathbf{x}) = \sum_{s,r,\ell} \frac{1}{4\pi^2} \int_{\Omega} d\omega \frac{A_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x})}{A_\ell^2(\mathbf{r}, \mathbf{s}, \mathbf{x}_0)} \times (-i |\mathbf{p}_\ell|^2) e^{-i\omega\tau_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x})} W_{k\ell}(\mathbf{r}, \mathbf{s}, \mathbf{x}) u_\ell^{\text{obs}}(\mathbf{r}, \mathbf{s}, \omega). \quad (25)$$

For computational simplicity, we return to time  $t$  and obtain:

$$\gamma_k^0(\mathbf{x}) = \frac{1}{\pi} \sum_{s,r,\ell} \frac{A_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x})}{A_\ell^2(\mathbf{r}, \mathbf{s}, \mathbf{x}_0)} |\mathbf{p}_\ell|^2 \mathcal{H}\{u_\ell^{\text{obs}}[\mathbf{s}, \mathbf{r}, t = \tau(\mathbf{r}, \mathbf{s}, \mathbf{x})]\} \quad (26)$$

where  $\mathcal{H}(\mathbf{u}^{\text{obs}})$  is the Hilbert transform of  $\mathbf{u}^{\text{obs}}$ . The Hilbert transform comes from the 2-D geometry of the problem. The presence of the argument  $t = \tau$  in (26) shows that this expression is in fact a ray theoretical migration with correction of the amplitudes by the scattering matrix and the inverse of the geometrical spreading.

The operator  $\mathbf{H}^{-1}$  in (23) is the formal inverse of

$$\mathbf{H} = \mathbf{G}^\dagger \mathbf{G} = \frac{\partial^2 S}{\partial \mathbf{f}^2}. \quad (27)$$

In optimization theory  $\mathbf{H}$ , the set of second-order functional derivatives of  $S$  with respect to  $\mathbf{f}$ , is called Hessian operator of  $S(\mathbf{f})$ . Unfortunately inverting the Hessian  $\mathbf{H}$  is seldom possible for continuous problems like (21). The usual approach in inverse theory is to replace the continuous problem by a discrete version of it, so that the problem reduces to the finding the inverse of a huge normal matrix. This is not possible in the present case, because of the sheer size of the normal matrix that we would have to invert. The usual method for solving large inverse problems of this kind is to use an iterative quasi-Newton method, where the inverse of the Hessian is approximately evaluated. Several choices are possible, for instance, conjugate gradient, preconditioned gradients, etc. (see, e.g. Tarantola 1987 for a review). If the inverse problem is ill-posed, however, iterative techniques are very delicate to use, because they can easily converge to numerical minima, they can oscillate and most of all, they can be very sensitive to numerical precision. Fortunately the inverse problem (21) is not so ill-posed when a single parameter like the  $P$ -impedance is being sought from reflection data gathered from a set of sources and receivers that cover well the studied area. When two parameters (like  $IP$  and  $IS$ ) are being inverted, the problem is more likely to be ill-posed as we will show later.

### Approximation to the Hessian

For an iterative technique to converge rapidly to the global minimum, it is absolutely necessary to get a good

approximate inverse of the Hessian. In the following, we will show that the high-frequency methods provide an excellent approximation to the Hessian, and that this approximation, derived from Beylkin's (1985) work, can be used with confidence in the inversion of seismic profiles. Let us start from the explicit expression for (27) in the ray theoretical approximation:

$$H_{ij}(\mathbf{x}, \mathbf{x}_0) = \sum_{s,r,\ell} \frac{1}{4\pi^2} \int_{\Omega} d\omega \omega |\mathbf{p}_\ell|^2 W_{i\ell}(\theta_{r,s}, \mathbf{x}_0) W_{j\ell}(\theta_{r,s}, \mathbf{x}) \times \frac{A_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x})}{A_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x}_0)} e^{-i\omega[\tau_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x}_0) - \tau_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x})]}. \quad (28)$$

The term under the integral over frequency is an oscillatory integral. At high frequencies this integral can be estimated by standard asymptotic methods (see, e.g. Copson 1967). The contribution of the oscillatory term to the integral is close to zero everywhere except where the phase of the exponential is close to zero. If the background velocity is sufficiently smooth, this occurs only when  $\mathbf{x}$  is close to  $\mathbf{x}_0$ . For more complex backgrounds, with caustics or focal points, the exponent may be close to zero at other points. We consider here only the simple case. Let us consider the integral for  $\mathbf{x}$  in the vicinity of point  $\mathbf{x}_0$ . We expand the phase and the amplitude terms in Taylor series around this point, keeping zeroth- and first-order terms for the phase function and only zeroth-order terms for the amplitudes:

$$\begin{aligned} \tau_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x}) &\sim \tau_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x}_0) + \mathbf{p}_\ell \cdot (\mathbf{x} - \mathbf{x}_0), \\ A_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x}) &\sim A_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x}_0), \quad W_{i\ell}(\mathbf{r}, \mathbf{s}, \mathbf{x}) \sim W_{i\ell}(\mathbf{r}, \mathbf{s}, \mathbf{x}_0) \end{aligned} \quad (29)$$

where  $\mathbf{p}_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x}_0) = \nabla \tau_\ell(\mathbf{r}, \mathbf{s}, \mathbf{x}_0)$  has a simple geometrical interpretation shown in Fig. 1. Since  $\tau$  is the sum of the traveltimes from the source  $\mathbf{s}$  to the scatterer  $\mathbf{x}_0$  and from the scatterer to the receiver  $\mathbf{r}$ , the curves  $\tau = \text{constant}$ , define for fixed  $\mathbf{s}$  and  $\mathbf{r}$ , a curve in the 2-D space of diffracting points  $\mathbf{x}$  that we call an isochrone.  $\mathbf{p} = \nabla \tau$  is the normal to the isochrone passing through the point  $\mathbf{x}_0$ . Using (29), the expression (28) for the Hessian  $\mathbf{H}$  simplifies to

$$H_{ij}(\mathbf{x}, \mathbf{x}_0) \sim \sum_{s,r,\ell} W_{i\ell}(\theta_{r,s}, \mathbf{x}_0) W_{j\ell}(\theta_{r,s}, \mathbf{x}) \frac{1}{4\pi^2} \times \int_{\Omega} d\omega \omega |\mathbf{p}_\ell|^2 e^{i\omega \mathbf{p}_\ell \cdot (\mathbf{x} - \mathbf{x}_0)} \quad (30)$$

which, as shown in Appendix B, can be easily integrated to:

$$H_{ij}(\mathbf{x}, \mathbf{x}_0) \sim \sum_{r,s,\ell} W_{i\ell}(\theta_{r,s}, \mathbf{x}_0) W_{j\ell}(\theta_{r,s}, \mathbf{x}) \delta(\mathbf{x} - \mathbf{x}_0) \quad (31)$$

where the sum extends over all the angles  $\theta_{r,s}$  available at the point  $\mathbf{x}_0$ .

The approximate Hessian (31) was obtained under the assumption that high-frequency waves dominate the operator  $G$ . While this is not exact, it is a very good approximation for the usual conditions of reflection seismology. In fact, the frequency contents of reflection seismograms is controlled by the source and the frequency pass-band of seismic instruments. As discussed by Miller *et al.* (1987), it is very likely that the conditions for the validity of approximation (31) are closely satisfied in practice. There is however an additional approximation in (31), namely that incident and diffracted rays at a heterogeneity are well

distributed in azimuth at the points where the material properties are being inverted. This added condition is difficult to evaluate in practice so that the relation (31) should be considered as an approximation to the Hessian which will be used to set up an iterative inversion method.

### Iterative quasi-Newton inversion method

Since we can not get an exact expression for the Hessian, we can use the approximation (30) in order to set up an iterative solution method. Let  $\mathbf{h}$  denote the approximation (31) to the Hessian; the quasi-Newton solution of the inverse problem (20) is

$$\mathbf{f}^{n+1} = \mathbf{f}^n - \mathbf{h}^{-1} \boldsymbol{\gamma}^n \quad (32)$$

where  $\boldsymbol{\gamma}^n$  is the gradient of  $S(\mathbf{f})$  calculated around the value of  $\mathbf{f}$  at the  $n$ th iteration:

$$\boldsymbol{\gamma}^n = \mathbf{G}^\dagger (\mathbf{u}^{\text{obs}} - \mathbf{G}\mathbf{f}^n). \quad (33)$$

Using the expression for the approximate Hessian obtained in (31), we get the iterative method:

$$\mathbf{f}^0(\mathbf{x}) = 0, \quad \mathbf{f}^1(\mathbf{x}) = \mathbf{G}^{-g} \mathbf{u}, \quad (34)$$

$$\mathbf{f}^{n+1}(\mathbf{x}) = \mathbf{f}^n(\mathbf{x}) + \mathbf{G}^{-g} (\mathbf{u}^{\text{obs}} - \mathbf{G}\mathbf{f}^n),$$

where  $\mathbf{G}^{-g} = (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{G}^\dagger$  is an approximate generalized inverse operator. Because the forward problem was linearized around the reference low-wavenumber model, the inverse operator does not change with the iterations. Therefore, after  $n$  iterations, the iterative method yields the following approximation:

$$\mathbf{f}_n = \left[ \mathbf{I} + \sum_{i=1}^{n-1} (\mathbf{I} - \mathbf{G}^{-g} \mathbf{G})^i \right] \mathbf{G}^{-g} \mathbf{u}^{\text{obs}} \quad (35)$$

that converges in the limit to

$$\lim_{n \rightarrow \infty} \mathbf{f}_n = (\mathbf{G}^{-g} \mathbf{G})^{-1} \mathbf{G}^{-g} \mathbf{u}^{\text{obs}}. \quad (36)$$

Thus the iterative method corrects for the error made in the estimation by the Hessian by the approximation (31). Let us remark that the first iteration of (36) yields an approximation that is very similar to the high-frequency inversion method developed by Beylkin (1985) and Beylkin & Burridge (1990).

### ANALYSIS OF RESOLUTION OF THE INVERSION METHOD

There are two problems that affect the solution of the inverse problem: first, the spatial resolution and, second, the separation of different components of impedance and density. In order to study the spatial resolution of our inversion algorithm, we consider the simple problem of a single scattering component:  $PP$  scattering ( $\ell = 1$ ), and a single parameter for inversion:  $P$ -wave impedance  $IP$  ( $k = 1$ ). In this case, the first iteration of the inverse problem (34) is explicitly

$$IP(\mathbf{x}) = \frac{1}{N} \sum_{r,s} \frac{1}{4\pi^2} \int_{\Omega} d\omega \frac{-i\mathbf{p}^2}{A_1(\mathbf{r}, \mathbf{s}, \mathbf{x})} \times e^{-i\omega\tau_1(\mathbf{r}, \mathbf{s}, \mathbf{x})} W_{11}(\mathbf{r}, \mathbf{s}, \mathbf{x}) u_1^{\text{obs}}(\mathbf{r}, \mathbf{s}, \omega) \quad (37)$$

where  $N = \sum_{r,s} W_{11}^2(\mathbf{r}, \mathbf{s}, \mathbf{x})$  and  $p_1 = |\mathbf{p}_1|$ .

Let  $\mathbf{x}_0$  be a fixed reference point inside the model and explore the solution of the inverse problem (37) around this point. Using again the relations (29) we can rewrite (37) in the form

$$IP(\mathbf{x}) = \frac{1}{N} \sum_{r,s} \frac{1}{4\pi^2} \int_{\Omega} d\omega \frac{-i\mathbf{p}^2}{A_1(\mathbf{r}, \mathbf{s}, \mathbf{x})} e^{-i\omega\tau_1(\mathbf{r}, \mathbf{s}, \mathbf{x}_0)} \times W_{11}(\mathbf{r}, \mathbf{s}, \mathbf{x}_0) u_1^{\text{obs}}(\mathbf{r}, \mathbf{s}, \omega) e^{-i\boldsymbol{\kappa} \cdot \Delta \mathbf{x}} \quad (38)$$

where  $\Delta \mathbf{x} = \mathbf{x} - \mathbf{x}_0$  and, as shown in Appendix B,

$$\boldsymbol{\kappa} = \omega \nabla \tau_1 = \omega \mathbf{p}_1, \quad \kappa = |\boldsymbol{\kappa}| = \omega p_1, \quad d\boldsymbol{\kappa} = d\omega p_1 \quad (39)$$

where  $p = |\mathbf{p}|$ . Following Beylkin & Burridge (1990), we use some geometrical properties of the diffraction problem discussed in Appendix B, in order to rearrange the sum in (38) as a sum over a set of discrete values of the scattering angle  $\theta$  and the angle  $\xi$  of the normal to the isochrone. These angles are defined in Fig. 1. Rewriting (38) in this fashion, and changing variables from  $\omega$  to  $\kappa$ , we get

$$IP(\mathbf{x}) = \frac{1}{4\pi^2} \sum_{\xi} \int_{\kappa} \kappa d\kappa IP(\kappa, \xi) e^{-i\boldsymbol{\kappa} \cdot \Delta \mathbf{x}} \quad (40)$$

where  $IP(\kappa, \xi)$  is given by

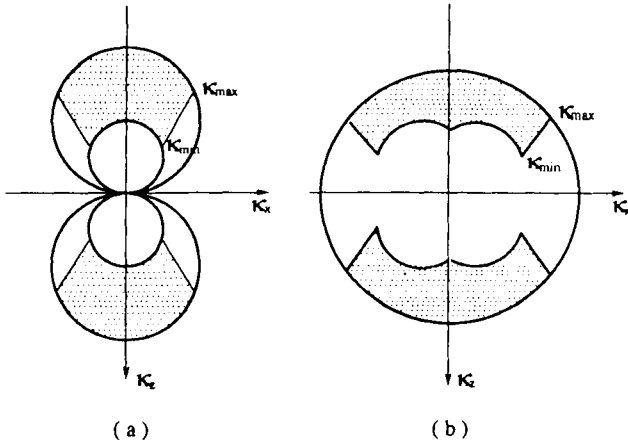
$$IP(\kappa, \xi) \sim \frac{1}{N} \sum_{\theta} W_{11}(\theta) \frac{-i\mathbf{p}}{\kappa A_1(\mathbf{r}, \mathbf{s}, \mathbf{x}_0)} \times e^{-i\boldsymbol{\kappa}/p_1 \tau_1(\mathbf{r}, \mathbf{s}, \mathbf{x}_0)} u_1^{\text{obs}}(\theta, \xi, \kappa/p_1). \quad (41)$$

Therefore, the inverted image  $IP(\mathbf{x})$  corresponds to a discretized, band-limited version of the exact Fourier transform of  $IP(\mathbf{x})$  in cylindrical coordinates  $\kappa, \xi$ :

$$IP(\mathbf{x}) = \frac{1}{4\pi^2} \int_0^{2\pi} d\xi \int_{-\infty}^{+\infty} \kappa d\kappa IP(\kappa, \xi) e^{-i\boldsymbol{\kappa} \cdot \Delta \mathbf{x}}. \quad (42)$$

As can be observed from Fig. 1, to each source–receiver pair  $(r, s)$ , there corresponds a unique couple of angles  $(\theta, \xi)$ . Examining (39) and (40) we observe that  $\theta$  and  $\xi$  play very different roles in the inversion. Each term in the sum over  $\xi$  in the approximate Fourier series (39) is calculated summing over all available angles  $\theta$ . The procedure for calculating  $IP$  is not very different from the usual data processing in exploration geophysics. The observed seismograms are migrated to the point of inversion  $\mathbf{x}_0$ , and sorted by scattering angle  $\theta$  and the local orientation of the isochrone  $\xi$ . The migrated records are stacked summing over the scattering angle  $\theta$ . Once the stack over  $\theta$  is done for every angle  $\xi$ , the local impedance  $IP(\mathbf{x})$  is calculated by the Fourier sum (40).

If we could obtain  $IP(\kappa, \xi)$  for all  $\kappa$  in the interval  $[-\infty, +\infty]$ , and for all angles  $\xi$  is the range 0 to  $2\pi$ , we could perfectly reconstruct the impedance field  $IP$  in the vicinity of  $\mathbf{x}_0$ . Unfortunately this is impossible for sources and receivers located on the surface of the Earth. Examining Fig. 1, it is clear that  $\xi$  may vary over a limited set of angles that depend on the distribution of sources and the aperture of the receiver set. Thus, angles  $\xi$  close to the vertical will be better resolved than shallow angles close to 0 or  $\pi$ . Furthermore, since the source signal is limited to a finite frequency domain  $\omega_{\min}$  to  $\omega_{\max}$ , say, the resolution in wavenumber domain  $\kappa$  will also be limited. Fig. 2(a) shows the spectral coverage that may be obtained for a single shot at a point situated at depth directly under the shot point. The angular spread  $\xi_{\min} - \xi_{\max}$  is controlled by the aperture



**Figure 2.** The spectral coverages of (a) a single shot data, and of (b) multisource data. We assume that receivers are evenly distributed along the surface in both cases.  $\kappa$  is the 2-D wavenumber vector in the space of parameters. The range of well-resolved wavenumber  $\kappa_{\min}$  to  $\kappa_{\max}$  depends on the frequency bandwidth of the used source. The angular sector resolved by the data is determined by the source and receiver spread.

of the receiver array. The larger the aperture, the wider the angular domain in  $(\kappa, \xi)$  space that may be resolved. Fig. 2(a) reflects the well-known fact that while vertical resolution may be good, the horizontal resolution is limited by the receiver spread. In Fig. 2(b), a similar estimation of spectral coverage is schematically presented for the case where a set of evenly distributed sources is available. The improvement in the horizontal resolution obtained in this case is obvious, but the reconstructed image is still a vertically high-passed and horizontally low-passed version of the actual impedance model.

### CONDITIONING OF THE SIMULTANEOUS INVERSION FOR $IP$ AND $IS$

In principle the iterative method proposed in (34) would allow for the simultaneous inversion of perturbations of density,  $P$ -wave and  $S$ -wave impedances. Each iteration would consist of the solution of the relatively simple linear system

$$(\mathbf{W}^T \mathbf{W}) \mathbf{f}_n = \mathbf{G}^+(\mathbf{u}^{\text{obs}} - \mathbf{G} \mathbf{f}). \quad (43)$$

This is a system of 3-by-3 equations that has to be solved at each point of the medium where we want to evaluate the perturbation vector  $\mathbf{f}$ . Unfortunately, the solution of this linear problem is very difficult because it is in general severely ill-conditioned. It is well known that density perturbations may not be separated from  $P$ -wave impedance because the corresponding scattering parameter in the matrix  $\mathbf{W}$  is small. A similar difficulty may affect the simultaneous estimation of both  $P$  and  $S$  impedances. Crase *et al.* (1990) did numerical tests to simulate the marine recording environment. Their results show that the two impedances can not be well separated from the marine synthetic data.

Here we will demonstrate from numerical analysis that the linear system (42) is ill-conditioned if scattering data of only a single mode are available. We illustrate this problem

by considering  $P$ - $P$  diffraction (marine data corresponds to this case). The conditioning of the 3-by-3 matrix  $\mathbf{A} = \mathbf{W}^T \mathbf{W}$  can be described by its condition number

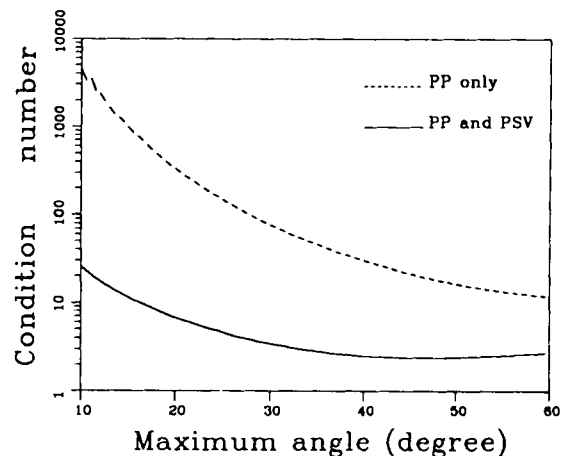
$$\text{cond}(\mathbf{A}) = \|\mathbf{A}\| \|\mathbf{A}^{-1}\| \quad (44)$$

where  $\|\mathbf{A}\|$  defines a certain norm for the matrix. For the euclidian norm the condition number is

$$\text{cond}(\mathbf{A}) = \frac{\lambda_{\max}}{\lambda_{\min}} \quad (45)$$

where  $\lambda_{\max}$  and  $\lambda_{\min}$  are the largest and smallest singular values of the matrix. Let us consider the condition number for the simultaneous inversion of the impedances  $IP$  and  $IS$  from reflection data of the single scattering mode  $P$ - $P$  ( $\ell = 1$ ). In Fig. 3, we show with dashed lines the logarithm of the condition number for this inversion as a function to the maximum scattering angle  $\theta_{\max}$ . For these computations we assumed a  $P$ - to  $S$ -wave velocity ratio  $\alpha/\beta \sim 1.732$ , and  $\theta_{\min} = 0$ . Let us underline that the matrix  $\mathbf{W}^T \mathbf{W}$  has a rank of 3.

A matrix is well conditioned if the condition number is not too far from 1. When the condition number is large, small errors in the data may cause large fluctuations of the solution. A realistic value of  $\theta_{\max}$  for the usual surface seismic configuration is around  $45^\circ$  for the inversion at shallow depths and  $25^\circ$  for deeper parameters. For shallow depths the condition number is not too great (about 30) and we may in principle invert for the two impedances. Unfortunately, in practice, it is difficult to separate the two parameters because of another obstacle: strong, complicated water-bottom multiples that appear at wide reflection angles (see Crase *et al.* 1990). At depth where  $\theta_{\max}$  is less than  $25^\circ$ , the condition number is larger than 100 and the separation of  $IS$  from  $IP$  becomes increasingly difficult. The reason for this difficulty in the inversion is that the elements of the scattering matrix  $\mathbf{W}$  given by (12a) show that  $P$ - $P$  diffraction for angles  $\theta < 35^\circ$  is basically controlled by the perturbation of  $P$  impedance (see e.g. Wu & Toksöz 1987). Therefore, from marine data, only  $P$ -wave impedance can be reliably inverted.  $S$  impedance and density perturbations are poorly controlled by the data and the inversion results will be suspect. Previous work by Santosa & Symes (1988)



**Figure 3.** Conditioning of simultaneous inversion for  $P$ - and  $S$ -wave impedances from single  $P$ - $P$  data (dashed line) and from  $P$ - $P$  and  $P$ - $SV$  data (solid line).

on the 2-by-2 scattering matrix of a layered fluid has led to similar conclusions.

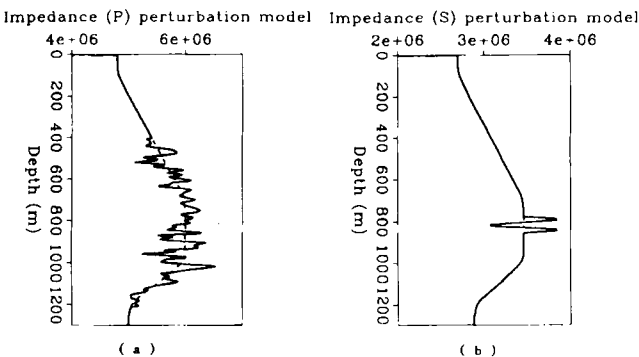
If three-component land data were available, we would be able to resolve the two impedances by using *P*-wave and *S*-wave arrivals. In Fig. 3 we show with a solid line the conditioning number for simultaneous inversion of *IP* and *IS* from *P*-*P* and *P*-*SV* reflection data. This illustrates that simultaneous inversion of *P* and *S* impedance is well conditioned (condition number is about 2.5) for  $\theta_{\max} > 20^\circ$ . Notice that the condition number is practically invariant if  $\theta_{\max}$  is larger than  $20^\circ$ . Now, we investigate practical performance of the iterative asymptotic inversion in order to verify numerically the conditioning analysis.

**NUMERICAL TESTS**

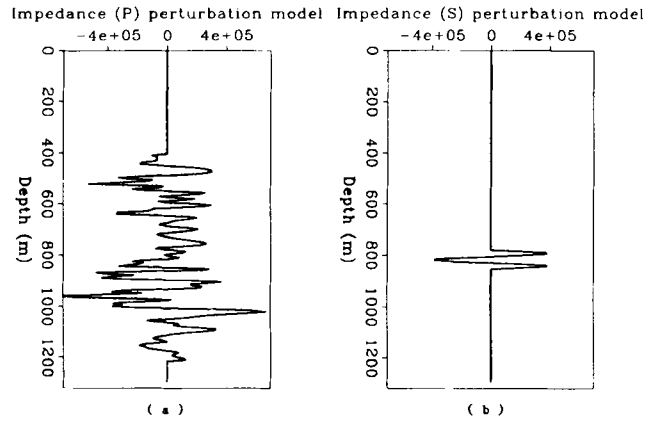
We test the inversion method that we propose on two synthetic examples. We used the finite difference method proposed by Virieux (1986) to generate the synthetics that will be used as ‘observed’ data for inversion. Inversion is done using the ray theoretical methods presented in this paper. The problem of ray tracing and the application of our method to real data will be taken up in Lambaré *et al.* (1991).

**Example 1**

We consider as the first example a single common shot gather of land data. The impedance models are shown in Fig. 4. Because we will only attempt a linearized inversion, we define a reference or ‘background’ velocity model (dashed lines in Fig. 4) that is a smoothed version of the full heterogeneous model (solid lines). The perturbation model is the difference between the perturbed model and the background model and is shown separately in Fig. 5. The *P*-wave impedance perturbation (Fig. 5a) was obtained from non-linear acoustic inversion of real data and checked against a sonic log by Pica, Diet & Tarantola (1990). We have no information about the distribution of *S*-wave impedance, but in order to test the possibility of separating *P*-wave impedance from *S*-wave impedance perturbations, we assumed a very simple, albeit unrealistic, sinusoidal perturbation of *S*-wave impedance (Fig. 5b).



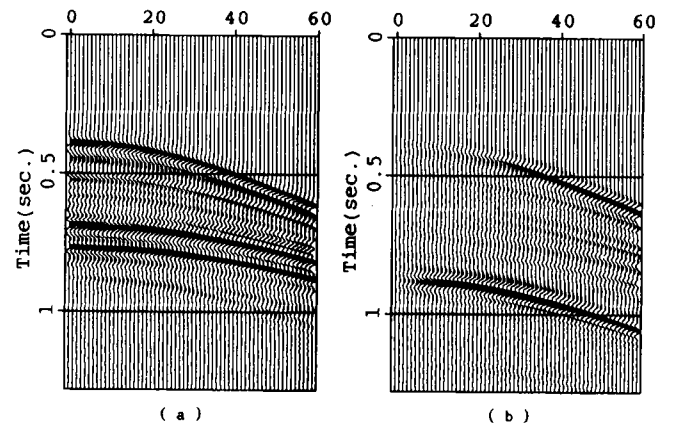
**Figure 4.** The models used to generate finite difference data: (a) the *P*-wave impedance model, and (b) the *S*-wave impedance model. The models are decomposed into a smooth reference model (dashed lines) and a perturbation (Fig. 5).



**Figure 5.** The perturbation model: (a) the *P*-wave impedance perturbation model and (b) the *S*-wave impedance perturbation model. The *P*-wave impedance perturbation is taken from the result of a non-linear acoustic inversion (Pica *et al.* 1990) and the *S*-wave model is a very simple localized perturbation of *S* impedance.

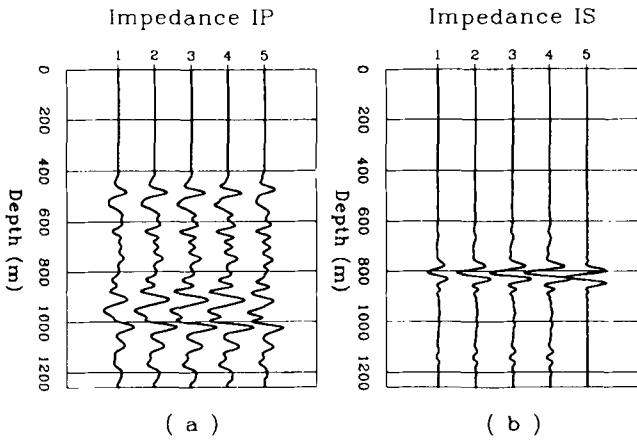
Figure 6 shows the synthetics generated by a finite-difference method with absorbing boundary conditions. The direct wave has been muted from these profiles. In the numerical computation, a grid with a spacing of 5 m was adopted both horizontally and vertically. We used a Dirac-like explosive source whose frequency bandwidth is 1 to 36 Hz. The sampling interval was 0.8 ms. Synthetic seismograms were calculated for a set of 60 geophones evenly distributed on the surface with a spacing of 20 m. The first geophone is located at the source and the last one has an offset of 1180 m from the source. Because we assumed a purely compressional source, only *P*-*P* and *P*-*SV* diffracted waves appear in the synthetics. The finite-difference synthetics shown in Fig. 6 were used as ‘observed’ data for inversion. *P*-*P* and *P*-*SV* waves were inverted simultaneously to obtain the variations of *P*-wave and *S*-wave impedances.

The vertical profiles for the four iterations of inversion are shown in Fig. 7. In order to compare the results of the iterative inversion with the original model, we present a



**Figure 6.** Muted synthetics calculated by finite differences for the model of Fig. 4: (a) vertical component and (b) horizontal component. The first offset is zero, and the last is 1180 m from the source. Because we simulate an explosive source, only *P*-*P* and *P*-*SV* diffracted waves dominate these synthetics.





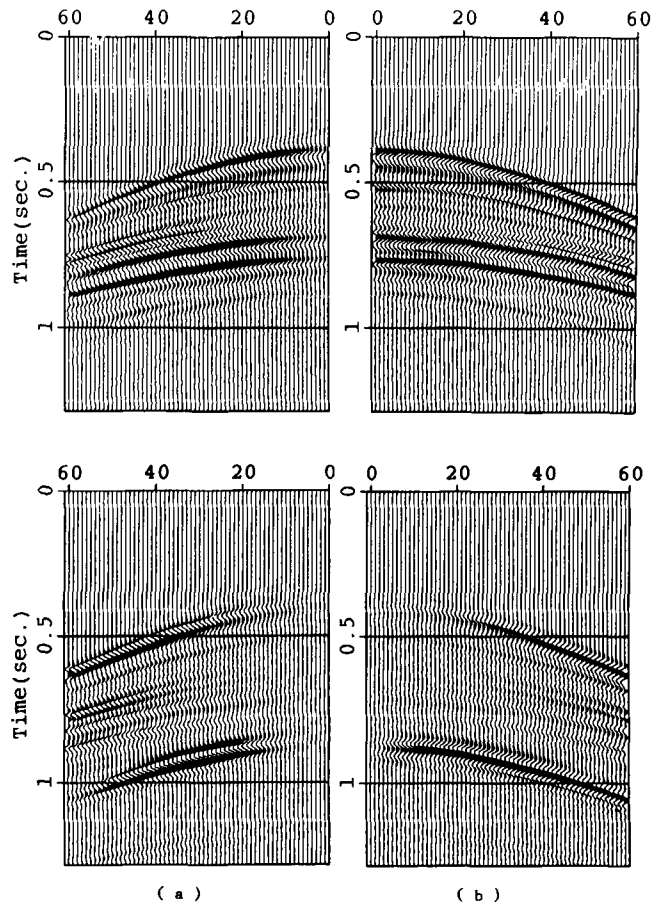
**Figure 7.** Comparison of the results of four iterations. To facilitate the comparison, a band-limited version of the original model is presented at the right (the fifth profile). (a) Comparison of *P*-wave impedance, (b) comparison of *S*-wave impedance. We can see that the *P*-wave impedance and the *S*-wave impedance are well separated. Note that the results of the third iteration are in excellent agreement with the true model.

band-limited version of the true model as the fifth profile in Fig. 7. Because the true model contains certain frequencies that cannot be inverted because of the finite frequency band of the source, we do not compare our results directly with the true model. The band-limited true model was computed filtering out frequency components above the Nyquist frequency. As can be observed from Fig. 7, the first iteration recovers the major characteristics of the original model. The main effect of higher order iterations is to improve the resolution. By iteration, we can retrieve lower and higher spatial wave frequencies which were poorly resolved in the first iteration. The results show that it was possible to distinguish variations of *P* impedance from those of *S* impedance; we have obtained similar results for other different numerical experiments. We are in agreement with the previously done conditioning analysis. Let us note that the inversion converges in only three iterations. It is, therefore, a very efficient inversion method thanks to the appropriate estimation of the Hessian and the use of the weighting function *Q* in the definition of the inner product in data space.

In Fig. 8 we show a comparison between the original data and synthetics calculated using the result of the fourth iteration in the inversion. We observe that the synthetics for the fourth iteration agree very well with the finite-difference data. In Fig. 9, we present the residuals of the fourth iteration of the inversion for the two components. The reduction of variance of the misfit function reaches about 78 per cent. The residuals contain some of the multiple reflections that our method cannot explain, but they are clearly very small, well within the numerical noise of the finite difference and ray methods. The computation of one iteration required 40 min of CPU time on a Convex C1 computer for this example.

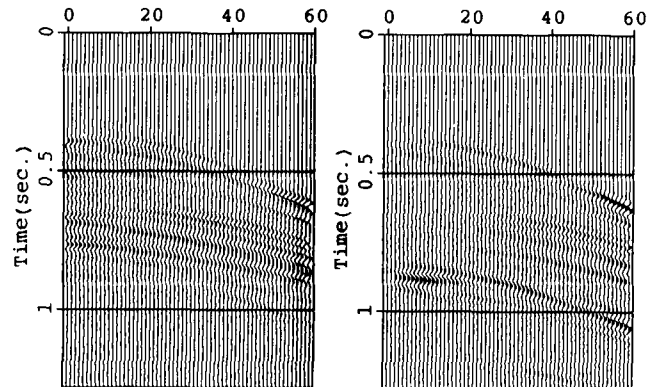
**Example 2**

The second example simulates a more complex geological structure. The *P*-wave impedance perturbation is shown in

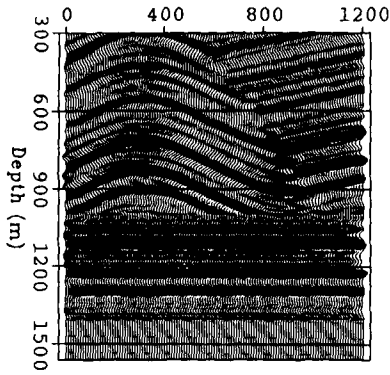


**Figure 8.** Real 'data' (left) and synthetics (right) after four iterations. The top panel is the vertical component and the bottom panel is the horizontal component.

Fig. 10. The *S*-wave impedance is the same as that of the *P* impedance with a factor of  $1/\sqrt{3}$ . The reference models are homogeneous: the *P*-wave velocity is  $3200 \text{ m s}^{-1}$ , *S*-wave velocity  $2000 \text{ m s}^{-1}$  and the density  $2000 \text{ kg m}^{-3}$ . With these models, we generated finite difference synthetics for 12 shot points. The grid spacing and time sampling intervals are the same as the previous example, 5 m and 0.8 ms. We used



**Figure 9.** The residuals between original 'data' and synthetics calculated from the model after four iterations. The reduction of the variance in the data reaches about 78 per cent. The left panel is the vertical component, while the right panel is the horizontal component.

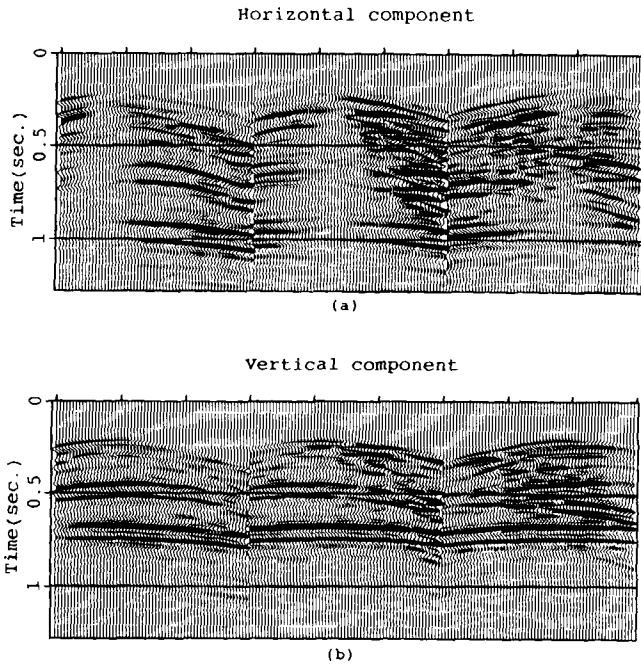


**Figure 10.** 2-D perturbation model for the second example: *P*-wave impedance. The *P*-wave impedance simulates a fold and a fault. *S*-wave impedance is proportional to *P* impedance.

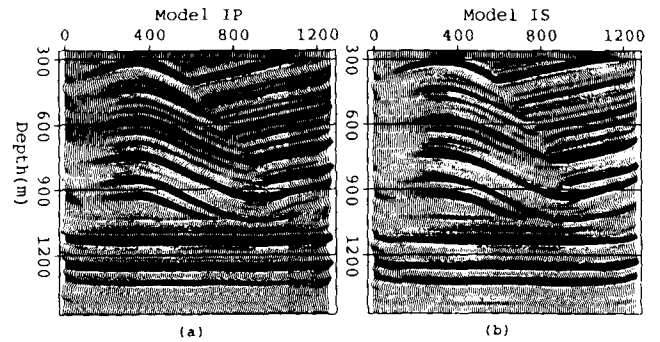
1.28 s of synthetic data for the inversion. To illustrate the synthetics, the second, sixth and tenth shot are shown in Fig. 11.

Since we use only surface data, the horizontal resolution is poorer than the vertical one. Thus we use the grid of 10 m horizontally and 5 m vertically for the inversion. Fig. 12 shows the results of the third iteration: *P*-wave impedance (a) and *S*-wave impedance (b). The parameters of the deeper part are underestimated because of relatively little coverage of data. By comparing (a) and (b) in Fig. 12, we see that our method separates *P* impedance perturbation from *S* impedance perturbation very efficiently: the different angular pattern of *P*- and *S*-wave energies allows this separation.

The synthetics and residuals for the third iteration are



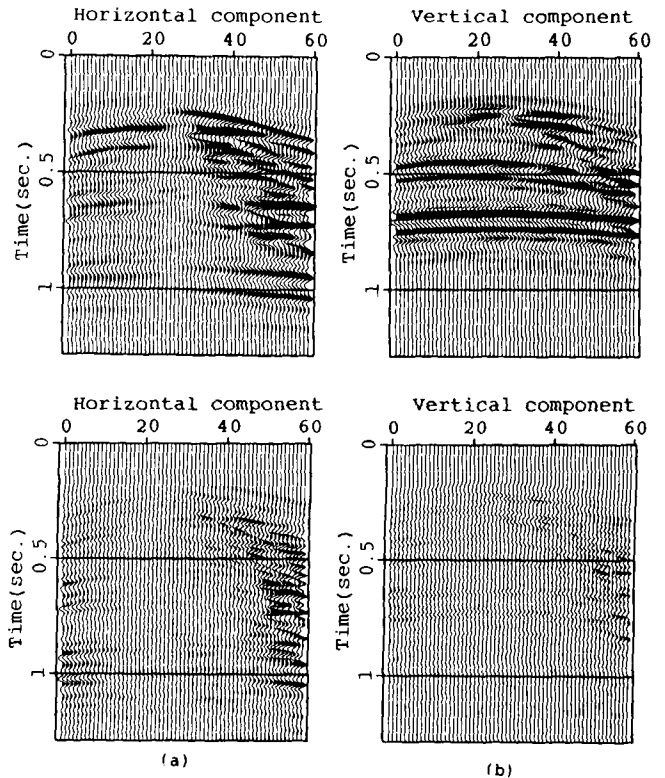
**Figure 11.** Synthetics seismograms generated by finite differences for the model shown in Fig. 10: (a) horizontal component and (b) vertical component. Synthetics calculated for three out of 10 shots are shown here.



**Figure 12.** Results obtained at the end of the third iteration: (a) *P*-wave impedance and (b) *S*-wave impedance. Impedances are well separated, but the image quality is poor. This is because we have a good multiplicity in subsurface coverage, but not enough range of offsets.

shown in Fig. 13. The variance of the residuals of the third iteration represent 60 per cent of the variance of the 'observed' data. We needed 5 hr of CPU time to compute one iteration of the inversion of the complete set of two-component seismograms for the 12 shots on a Convex C1 computer.

As a partial conclusion, we observe that our numerical simulation allow to recover both *IP* and *IS* impedances for the reference medium we have selected. For other reference media, recovered impedances might be worse but the geometry of acquisition still allows the separation between



**Figure 13.** Synthetics (top) and residuals (bottom) after three iterations. The reduction of variance in the data has reached about 60 per cent. The panels (a) are the vertical component, while panels (b) are the horizontal components.

parameters, a result we have analysed theoretically in a previous section.

## DISCUSSION

Numerical tests illustrate that our method is efficient for the inversion of land data. As usual in ray theory, the medium must satisfy certain validity conditions for the high-frequency methods to be applicable. These conditions are essentially that (1) the wavelength of the source be short compared to the dominant wavelength of the velocity model and (2) that the sources and receivers be far away from the scattering volume. In addition to these limitations that are inherent to ray methods, we also require that the Born approximation be valid, which means that (3) the dominant wavelength of the source signal be larger than the characteristic length of the scatterers. These three conditions determine the range of the frequency in which the ray method is applicable.

The use of the ray method improves the stability of the inversion by the coherence in time required by the ray seismogram. Iterations over new reference media would have increased the independence from the initial media. We have not attempted this step because it is not obvious to deduce a new smooth reference medium from the spatial high-frequency picture we recovered.

The two examples that we presented were chosen in order to illustrate that  $P$  and  $S$  impedances can be simultaneously inverted only if data sets containing both  $P$ - $P$  and  $P$ - $S$  diffractions are available. This may be the case if three-component land reflection data were available, or in the processing of three-component VSP data (Beydoun & Mendez 1989). We also showed that it is possible to perform iterative asymptotic inversion when the reference velocity field is known accurately. In the numerical tests, we chose the models as close to a real Earth subsurface as possible.

Although the inversion procedure works well with synthetic data, it would be unrealistic to expect such good results when applying this technique to real data. The magnitude of the perturbations may not be correct because:

(1) We need to have a precise knowledge of the source time function of the shots if we want to deconvolve the inverted profile from the effect of the source. In practice, however, source strength and phase are not very well known and, furthermore, change from source to source.

(2) Results depend also on the recording geometry, since the coverage of the data varies according to seismic configuration. Incomplete coverage obviously degrades the image. As we showed from an analysis of resolution of a single iteration of our method, Earth parameters are estimated by a sort of discretized, band-limited Fourier integral operator. Beylkin & Burridge (1990) obtained a similar result assuming that sources and receivers were uniformly distributed on the surface. In practice, however, the range of offsets ( $\theta$  in Fig. 1) and diffraction angles ( $\xi$  in the same figure) is limited. In this case, the impedance at depth can be only partially reconstructed.

(3) We did not take into account ground roll, refraction and multiples which were all muted from our synthetic 'data'. In reality, their effects cannot be suppressed completely.

(4) The Earth is certainly anisotropic and the velocity distribution is actually 3-D, whereas our inversion method assumes a 2-D isotropic Earth. It would not be difficult to relax the theoretical restriction to two dimensions. The computational cost would be enormous however even in present-day parallel computers. A problem that will certainly appear in 3-D will be that of computational and algorithmic complexity that we can hardly evaluate at present.

(5) Finally, the initial reference medium might blur the reconstructed impedance pictures in a coherent way whatever is the data acquisition geometry. Noise might also be a source of imaging degradation which can be partially solved by the data acquisition geometry.

In spite of these caveats, due mainly to the fact that we can only approach the real complexity of the Earth step by step, the method proposed here gives a good kinematic image of the Earth's interior that can be explicitly checked by comparison of observed and synthetic seismograms. We can compute explicitly the condition number of the Hessian and in this fashion estimate the resolving power. Alternative elastic inversion methods have been proposed by Pica *et al.* (1990), Crase *et al.* (1990) and others, but their method is based on finite-difference simulation of wave propagation which is extremely expensive. The method we presented here is more complicated analytically but is certainly much faster if ray tracing can be done efficiently and accurately. Actually, the volume of computation for the backprojection step is very small compared with the CPU time for direct modelling. Once the forward problem is solved all ray theoretical quantities can be stored and reused in the backpropagation step. Thus, the number of propagations to be simulated numerically in each iteration equals the number of shots. In finite difference inversion, or in the ray theoretical method of Beydoun & Mendez (1989), backpropagation is done also numerically at a cost similar to the solution of the forward problem.

An important question that remains is whether ray theory is sufficiently accurate for the synthesis of reflection seismograms. The answer to this question depends on how good the smooth reference model is. If, as in the examples shown here, the smooth model gives the correct arrival times for the single reflected phases, ray theory is at least as accurate as finite differences. The main source of inaccuracies in ray theory is ray tracing over long distances. This is not the case in usual reflection seismology, as is clearly shown by the fact that most of the usual data processing is done with extremely simple 1-D reference models. We verified this assertion by comparing synthetics generated for the same model by finite differences and by ray theory. As long as the frequency content of the source is close to that used in exploration geophysics, seismograms are practically indistinguishable. In fact, we do not think that limitations come from ray theory. It is much more likely that the real problem will come from the Born approximation on which all inversion methods, linear and non-linear, are based. Seismograms calculated by first-order perturbation lack the contribution of multiple scattering, local ground amplification and so on, which may be important under certain circumstances. Some tests of the validity of ray theory will be the subject of further work.

## CONCLUSIONS

We proposed an asymptotic linearized inversion method for 2-D seismic reflection data. Two fundamental assumptions were made: first, the velocity model can be separated into a smooth reference velocity model and, second, a perturbation of impedance that we try to invert. The second assumption is that ray theory is applicable for the computation of diffracted waves from the perturbations of impedance. In contrast with several recently proposed methods for asymptotic inversion (Beylkin 1985; Beylkin & Burridge 1990), our method was derived from the well-known methods of generalized inverse theory. We introduced an efficient approximation to the Hessian that leads to a rapidly convergent quasi-Newtonian iterative inversion method. Beydoun & Mendez (1989) also approached elastic inversion by ray theory in order to estimate the Hessian. We found that a single step of our inversion procedure as proposed in (34) yields a band-limited version of the perturbation field. Resolution in the spectral domain is controlled by both the frequency content of the source signal and the angular spread of the source-receiver distribution. The first iteration of our method is similar to Beylkin & Burridge (1990) method, but we do not depend on the availability of continuous distribution of sources and/or receivers on the surface.

Several synthetic examples are treated to show that the inversion method is robust and converges rapidly towards the solution thanks to a good estimation of the Hessian. We show theoretically that the simultaneous inversion of  $P$  and  $S$  impedances from data of a single scattering mode, like  $P$ - $P$  scattering, is ill posed if only a limited range of source-receiver offsets is available. Thus, from single wave type data (marine data, for example), only one parameter can be resolved reliably. From theoretical analysis, multiparameter inversion can be achieved if the data sets recording several components of ground motion were available. As a numerical verification, two simple cases of synthetic 'data' containing both  $P$ - $P$  and  $P$ - $SV$  diffractions were used to invert simultaneously for  $P$  and  $S$  impedances. These tests show that linear elastic inversion works on realistically sized problems with very good resolution of differences between  $P$  and  $S$  impedances. From our analytical study and our numerical experiences, we can state that recovering multiparameters is possible as long as we are in the linear domain. Of course, the way how to be in this domain or, in other words, how to find a good reference medium has not been addressed in this paper and requires further investigation. An application of our method to real marine data will be presented in a separate paper (Lambaré *et al.* 1991).

## ACKNOWLEDGMENTS

This work is a contribution of Groupement Scientifique Sismique, sponsored by Secteur TOAE of Centre National de la Recherche scientifique, Elf Aquitaine Company, and Institut National du Pétrole. Precise comments by Roel Snieder improved the manuscript. Many thanks to two unknown reviewers for their helpful comments. Publication IPG Paris No. 1185.

## REFERENCES

- Aki, K. & Richards, P., 1980. *Quantitative Seismology: Theory and Methods*, Freeman, San Francisco.
- Bernard, P. & Madariaga, R., 1984. A new asymptotic method for the modeling of near-field accelerograms, *Bull. seism. Soc. Am.*, **74**, 5399-5570.
- Beydoun, W. & Mendez, M., 1989. Elastic ray-Born  $L^2$  migration/inversion, *Geophys. J. Int.*, **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. & 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. & Cohen, J. K., 1979. Direct inversion procedure for Claerbout's equation, *Geophysics*, **44**, 1034-1040.
- Cao, D., Beydoun, W. B., Singh, S. C. & Tarantola, A., 1990. A simultaneous inversion for background velocity and impedance maps, *Geophysics*, **55**, 458-469.
- Cohen, J. K. & Bleistein, N., 1979. Velocity inversion procedure for acoustic waves, *Geophysics*, **44**, 1077-1087.
- Copson, E. T., 1967. *Asymptotic Expansions*, Cambridge University Press, Cambridge, UK.
- Clayton, R. W. & Stolt, R. H., 1981. A Born-WKB inversion method for acoustic reflection data, *Geophysics*, **46**, 1559-1567.
- Cruse, E., Pica, A., Noble, M., McDonald, J. & Tarantola, A., 1990. Robust elastic nonlinear waveform inversion: Application to real data, *Geophysics*, **55**, 527-538.
- Gauthier, O., Virieux, J. & Tarantola, A., 1986. Two-dimensional nonlinear inversion of seismic waveforms: numerical tests, *Geophysics*, **51**, 1387-1403.
- Ikelle, L., Diet, P. & Tarantola, A., 1988. Linearized inversion of multi-offset seismic reflection data in the  $\omega - k$  domain. Depth dependent reference medium, *Geophysics*, **53**, 50-64.
- Kolb, P., Collino, F. & Lailly, P., 1986. Pre-stack inversion of 1-D medium, *Proc. IEEE*, **74**, 498-508.
- Knopoff, L., 1956. Diffraction of elastic waves, *J. acoust. Soc. Am.*, **28**, 217-229.
- Lambaré, G., Virieux, J., Madariaga, R. & Jin, S., 1991. Iterative asymptotic inversion of seismic profiles in the acoustic approximation, *Geophysics*, in press.
- Miller, D., Oristaglio, M. L. & Beylkin, G., 1987. A new slant on seismic imaging: Migration and integral geometry, *Geophysics*, **52**, 943-964.
- Mora, P., 1987. Nonlinear two-dimensional elastic inversion of multioffset seismic data, *Geophysics*, **52**, 1211-1228.
- Mora, P., 1988. Elastic wavefield inversion of reflection and transmission data, *Geophysics*, **53**, 750-759.
- Pica, A., Diet, J. P. & Tarantola, A., 1990. Nonlinear inversion of seismic reflection data in a laterally invariant medium, *Geophysics*, **55**, 284-292.
- Santosa, F. & Symes, W. W., 1988. High-frequency perturbation analysis of the surface point-source response of a layered fluid, *J. Comp. Phys.*, **74**, 318-381.
- Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, **49**, 1259-1266.
- Tarantola, A., 1986. A strategy for nonlinear inversion of seismic reflection data, *Geophysics*, **51**, 1893-1903.
- Tarantola, A., 1987. *Inverse Problem Theory*, Elsevier, Amsterdam.
- Virieux, J., 1986.  $P$   $SV$ -wave propagation in heterogeneous media, velocity-stress finite-difference method, *Geophysics*, **51**, 889-899.
- Wu, R. S. & Toksöz, M. N., 1987. Diffraction tomography and multisource holography applied to seismic imaging, *Geophysics*, **52**, 11-25.

## APPENDIX A

Steps for the definition of scattering matrix  $W$ 

In (6) we separated the solution of the forward problem into a known field  $G_{ji}^0$ —the Green function for the reference medium—and the scattered field  $u_{ij}$  that we want to calculate. Within the first Born approximation, the scattered field is determined from the linearized equation of motion by

$$\rho_0 \partial_{tt}^2 u_{ij} - \partial_{x_m} (c_{impq}^0 \partial_{x_q} u_{pj}) = -\delta\rho \partial_{tt}^2 G_{ij}^0 + \partial_{x_m} (\delta c_{impq} \partial_{x_q} G_{pj}^0). \quad (\text{A1})$$

Noting that the Green function for this problem is that of the unperturbed problem  $G^0$ , the solution of (A1) for a source located at  $\mathbf{s}$  and a receiver at  $\mathbf{r}$  is

$$u_{ij}(\mathbf{r}, \mathbf{s}, t) = - \int_M dx^2 \{ \delta\rho \partial_{tt}^2 G_{in}^0(\mathbf{r}, \mathbf{x}, t) * G_{nj}^0(\mathbf{x}, \mathbf{s}, t) - G_{in}^0(\mathbf{r}, \mathbf{x}, t) * \partial_{x_m} [\delta c_{nmpq} \partial_{x_q} G_{pj}^0(\mathbf{x}, \mathbf{s}, t)] \} \quad (\text{A2})$$

where  $M$  is the heterogeneous region, and  $*$  denotes convolution in time.  $G(\mathbf{x}, \mathbf{s}, t)$  is the downgoing Green function from the source to the scattering point  $\mathbf{x}$ , while  $G(\mathbf{r}, \mathbf{x}, t)$  is the upgoing Green function from the scattering point  $\mathbf{x}$  to the observation point at  $\mathbf{s}$ . Integrating the second term on the right by parts and noticing that the boundary term vanishes, we arrive at

$$u_{ij}(\mathbf{r}, \mathbf{s}, t) = - \int_M dx^2 [ \delta\rho \partial_{tt}^2 G_{in}^0(\mathbf{r}, \mathbf{x}, t) * G_{nj}^0(\mathbf{x}, \mathbf{s}, t) - \partial_{x_m} G_{in}^0(\mathbf{r}, \mathbf{x}, t) * \delta c_{nmpq} \partial_{x_q} G_{pj}^0(\mathbf{x}, \mathbf{s}, t) ]. \quad (\text{A3})$$

This is the exact expression for the first Born approximation for elastic wave scattering from a general heterogeneity  $\delta\rho(\mathbf{x})$  and  $\delta c_{ijkl}(\mathbf{x})$ .

We can now introduce the ray theoretical approximation for the calculation of the Green functions  $G^0$  in (A3). Within this approximation, the Green's functions can be separated into  $P$ - and  $S$ -waves as shown in (3). Each component of the Green function,  $n = P$  or  $n = S$ , is then given by

$$G_{ij}^n(\mathbf{x}, \mathbf{s}, t) = A_{ij}^n(\mathbf{x}, \mathbf{s}) \frac{H(t - \tau^n)}{\sqrt{t - \tau^n}} \quad (\text{A4})$$

where the tensor amplitudes  $A_{ij}$  are given by

$$A_{ij}^n(\mathbf{x}, \mathbf{s}) = \gamma_i^n(\mathbf{x}) \mathcal{R}_j^n(\mathbf{s}) A^n(\mathbf{x}, \mathbf{s}) \quad (\text{A5})$$

with the scalar amplitudes

$$A^P(\mathbf{x}, \mathbf{s}) = \frac{1}{2\alpha(\mathbf{x}) \sqrt{2\pi\rho(\mathbf{x})\rho(\mathbf{s})\alpha(\mathbf{s})} J^P(\mathbf{x}, \mathbf{s})}, \quad (\text{A6})$$

$$A^S(\mathbf{x}, \mathbf{s}) = \frac{1}{2\beta(\mathbf{x}) \sqrt{2\pi\rho(\mathbf{x})\rho(\mathbf{s})\beta(\mathbf{s})} J^S(\mathbf{x}, \mathbf{s})},$$

where  $\gamma^n$  is the unit polarization vector calculated at the receiver  $\mathbf{x}$ ; for  $P$ -waves, it is tangent to the ray and, for  $S$ -waves, it is perpendicular to it.  $\mathcal{R}^n$  is the radiation pattern, and  $J^n(\mathbf{x}, \mathbf{s})$  is the geometrical spreading from a source point at  $\mathbf{s}$  to an observer at  $\mathbf{x}$ .  $\alpha$  and  $\beta$  are the  $P$ - and  $S$ -wave velocity, respectively. The traveltime functions  $\tau^P$  and  $\tau^S$  satisfy the eikonal equations

$$(\nabla\tau^P)^2 = 1/\alpha^2, \quad (\nabla\tau^S)^2 = 1/\beta^2, \quad (\text{A7})$$

respectively. The traveltime  $\tau$ , polarization, radiation pattern and geometrical spreading are calculated by ray tracing in the reference medium.

Inserting the ray theoretical approximations to the Green functions in the scattering equations (A2), we obtain the scattered field which consists of five terms

$$u_{ij}(\mathbf{r}, \mathbf{s}, t) = \sum_{\ell=1}^5 U_{ij}^{\ell}(\mathbf{r}, \mathbf{s}, t). \quad (\text{A8})$$

Each term corresponds to one of five possible scattering modes:

$$\begin{aligned} \ell = 1, & \quad P-P, & \ell_1 = P, & \ell_2 = P, \\ \ell = 2, & \quad P-S, & \ell_1 = P, & \ell_2 = S, \\ \ell = 3, & \quad S-P, & \ell_1 = S, & \ell_2 = P, \\ \ell = 4, & \quad S-S, & \ell_1 = S, & \ell_2 = S, \\ \ell = 5, & \quad SH-SH, & \ell_1 = SH, & \ell_2 = SH, \end{aligned} \quad (\text{A9})$$

where  $\ell_1$  indicates the type of incoming wave and  $\ell_2$  the outgoing wave on the scattering point. Each scattering mode can be expressed in the very simple form:

$$U_{ij}^{\ell}(\mathbf{r}, \mathbf{s}, t) = \gamma_i^{\ell_2}(\mathbf{r}) \mathcal{R}_j^{\ell_1}(\mathbf{s}) u_{\ell}(\mathbf{r}, \mathbf{s}, t), \quad (\text{A10})$$

with

$$\begin{aligned} u_{\ell}(\mathbf{r}, \mathbf{s}, t) = & - \partial_i \int_M [ \delta\rho \delta_{pn} + \delta c_{nmpq} \partial_{x_q} \tau^{\ell_2}(\mathbf{x}, \mathbf{s}) \partial_{x_m} \tau^{\ell_1}(\mathbf{r}, \mathbf{x}) ] \\ & \times A^{\ell_2}(\mathbf{x}, \mathbf{s}) \gamma_p^{\ell_1}(\mathbf{x}) \mathcal{R}_n^{\ell_2}(\mathbf{x}) A^{\ell_1}(\mathbf{r}, \mathbf{x}) \\ & \times \delta [ t - \tau^{\ell_1}(\mathbf{r}, \mathbf{x}) - \tau^{\ell_2}(\mathbf{x}, \mathbf{s}) ] dx^2 \end{aligned} \quad (\text{A11})$$

where the appropriate value for the indices  $\ell$ ,  $\ell_1$  and  $\ell_2$  have to be taken from the scattering modes defined in (A9).

The space derivatives of the traveltime function  $\partial_{x_q} \tau^n$  are calculated at the scattering point  $\mathbf{x}$ . They are in fact the components of the slowness vector  $\mathbf{p}^n = \nabla\tau^n$ . Introducing the angles  $\theta_{\ell}$  between the slowness vectors of the incident and scattered rays at  $\mathbf{x}$ , we can rewrite equations (A11) with the variables  $\theta$ ,  $\delta p$ ,  $\delta IP$  and  $\delta IS$ , and obtain expression (12) for the scattering matrix  $W$  used in a concise notation of (A11) which is equation (8).

## APPENDIX B

## An efficient ray theoretical estimation of the Hessian

So far we have parametrized the receivers and sources by their positions  $\mathbf{r}$  and  $\mathbf{s}$  along the recording surface. It is possible however to reorganize the receivers and sources as a function of the angles that the incident and scattered ray make with the  $x$ -axis at the scattering point. Following Fig. 1,  $\theta$  is the scattering angle already found in the definition of the scattering matrix  $W$  and  $\xi$  is the angle that the local normal to the curve  $\tau_{\ell} = \text{constant}$  makes with the horizontal axis. For any given  $\mathbf{s}$  and  $\mathbf{r}$ , and any scattering mode,  $\theta$  and  $\xi$  are uniquely defined. Therefore, if rays in the reference medium do not have caustics or other ray singularities, the mapping  $(r, s) \rightarrow (\theta, \xi)$  is one-to-one.

Let us notice that the gradient of the two-way traveltime  $\mathbf{p}$  is also the normal to the curves  $\tau_{\ell} = \text{constant}$ . These curves, called isochrones, correspond to points on the medium which have equal two-way traveltime between a fixed source and a given receiver. This simple geometrical interpretation

of traveltime gradients has found use in several domains (see e.g. Bernard & Madariaga 1984; Miller *et al.* 1987). Consider first, the approximate Hessian given by equation (34):

$$H_{kj}(\mathbf{x}, \mathbf{x}_0) \sim \sum_{s,r,\ell} W_{k\ell}(\theta_{r,s}, \mathbf{x}_0) W_{j\ell}(\theta_{r,s}, \mathbf{x}) \times \frac{1}{4\pi^2} \int_{\Omega} d\omega \omega |\mathbf{p}_{\ell}|^2 e^{i\omega \mathbf{p}_{\ell} \cdot (\mathbf{x} - \mathbf{x}_0)}$$

is transformed by reorganizing the sum with the new angles into

$$H_{kj}(\mathbf{x}, \mathbf{x}_0) \sim \sum_{\theta,\ell} W_{k\ell}(r, \mathbf{x}_0) W_{j\ell}(r, \mathbf{x}) \times \frac{1}{4\pi^2} \int_{\Omega} d\omega \omega |\mathbf{p}_{\ell}|^2 e^{i\omega \mathbf{p}_{\ell} \cdot (\mathbf{x} - \mathbf{x}_0)}. \tag{B1}$$

Using the following transformation of variables  $\kappa = \omega |\mathbf{p}_{\ell}|$  we get

$$H_{kj}(\mathbf{x}, \mathbf{x}_0) \approx \sum_{\theta,\ell} W_{k\ell}(r, \mathbf{x}_0) W_{j\ell}(r, \mathbf{x}) \frac{1}{4\pi^2} \int_{\mathcal{K}} d\kappa \kappa e^{i\kappa \cdot (\mathbf{x} - \mathbf{x}_0)} \tag{B2}$$

but

$$\frac{1}{4\pi^2} \sum_{\xi} \int_{\mathcal{K}} d\kappa \kappa e^{i\kappa \cdot (\mathbf{x} - \mathbf{x}_0)} = \delta(\mathbf{x} - \mathbf{x}_0) \tag{B3}$$

is a Fourier series defining an approximation to the delta function around  $\mathbf{x}_0$ . (B3) is exact only if the angular domain  $\xi$  and the wavenumber domain  $\mathcal{K}$  cover the entire 2-D Fourier transform domain of  $\kappa$ .