LINEARIZED INVERSION OF SEISMIC REFLECTION DATA*

A. TARANTOLA**

ABSTRACT

TARAN Otla, A. 1984, Linearized Inversion of Seismic Reflection Data, Geophysical Prospecting, 32, 998-1015.

This is the first of a series of papers giving the solution of the inverse problem in seismic exploration. The acoustic approximation is used together with the assumption that the velocity field has the form \( c_1 c_2 + \delta c_2 \), with \( |\delta c_2| \ll c_2 \). The forward problem is then linearized (thus neglecting multiple reflected waves) and the inverse problem of estimating \( \delta c_2 \) is set up. Its rigorous solution can be obtained using an iterative algorithm, each step consisting of a classical Kirchhoff migration (hyperbola summation) plus a classical forward modeling step (circle summation).

1. INTRODUCTION

There are three different approaches to the problem of transforming seismic reflection data from the "image space" to the "object space," namely:

1. The "migration" approach, based either on an imaging principle (Claerbout 1971) or on an exploding reflector concept (Loewenthal, Lu, Roberson and Sherwood 1976). Examples of migration methods are the Kirchhoff (hyperbola summation) method (French 1974, Schneider 1978), the Fourier-domain methods (Stolt 1978, Gazdag 1978), or the finite-difference methods (Claerbout 1971).

2. The "direct inversion" approach, in which the data set is reduced to a set with the same number of degrees of freedom as the unknown model, so that the resolution of the forward problem can be considered a bijective transformation which, applied to the model, gives the data set. The inverse problem is then solved by looking for the inverse transformation in exactly the same way as the Fourier of Laplace transforms are inverted (see, for instance, Cohen and Bleistein 1979, Raz 1981). This approach has a major handicap: it can handle neither redundant data, nor overdetermined. In addition, weighting of the data or a priori information on the model cannot easily be introduced.

* Received May 1983, revised version accepted for publication April 1984.
** Equipe de Tomographie Géophysique, Laboratoire de Sismologie (CNRS, LA 193), Institut de Physique du Globe, Université P. et M. Curie, 4 Place Jussieu, 75230 Paris Cedex 05, France.

2. THE FORWARD PROBLEM

The starting point of our theory is the equation for acoustic waves

\[
\frac{1}{K(r)} \frac{\partial^2}{\partial t^2} \text{div} \left( \frac{1}{\rho(r)} \text{grad} \right) U(r, t) = s(r, t),
\]

where \( K(r) \) represents the bulk modulus, \( \rho(r) \) the density, \( s(r, t) \) the source of acoustic waves, and \( U(r, t) \) the pressure variation. Although this equation holds only in media (i.e., in media with vanishing shear modulus) it is customary in the industrial community to use it for an approximate description of the propagation of seismic (i.e., elastic) waves. As our aim in this paper is to discover the links between migration and inversion, we accept this approximation without further discussion.

Results from borehole data indicate that density variations are not the main source of reflected waves (Flood 1981). If we neglect the derivatives of the density in the acoustic equation we arrive at

\[
\left( \frac{1}{C(r)} \frac{\partial^2}{\partial t^2} - V^2 \right) U(r, t) = s(r, t),
\]

where \( C(r) = \sqrt{[K(r)/\rho(r)]} \) is the velocity of the acoustic waves.

We introduce the Green's function \( g(t; t', r') \) of the problem by the definition

\[
\left( \frac{1}{C(r)} \frac{\partial^2}{\partial t^2} - V^2 \right) g(t; t', r') = \delta(r - r') \delta(t - t'),
\]

where \( \delta \) is the Dirac function. It is well known that if \( g(t; t'; r') \), for fixed \( r' \) and \( t' \), satisfies the initial and boundary conditions, then the solution of (1) can be obtained by

\[
U(r, t) = \int dr' \int dt' g(t; t'; r', t) s(r', t').
\]
We first consider an homogeneous medium:
\[ C(r) = c = \text{const.} \]

If we denote by \( u(r, t) \) the corresponding pressure field, (1) reduces to
\[ \left( \frac{1}{c^2 \frac{\partial^2}{\partial t^2} - \nabla^2} \right) u(r, t) = s(r, t). \]  
(5)

If we don’t worry for the moment about boundary conditions, we can take the free space Green’s function for the homogeneous medium (see, for example, Morse and Feshbach 1953):
\[ g(r, t; r', t') = \frac{1}{4\pi} \frac{1}{\|r - r'\|} \delta \left( t - t' - \frac{\|r - r'\|}{c} \right). \]  
(4)

Using (2) we have
\[ u(r, t) = \frac{1}{4\pi} \int dr' \frac{1}{\|r - r'\|} s \left( r', t - \frac{\|r - r'\|}{c} \right). \]  
(5)

We assume now that the waves are generated by an isotropic point source with a time function \( S(t) \) located at \( r = r_s \):
\[ s(r, t) = 8\pi^2 c^2 \delta(r - r_s) S(t), \]  
(6)

where the factor \( 8\pi^2 c^2 \) is introduced for simplification of subsequent formulae. Denoting by \( u(r, t; r_s) \) the wavefield corresponding to this source location, we easily obtain
\[ u(r, t; r_s) = 2\pi c^2 \frac{1}{\|r - r_s\|} S \left( t - \frac{\|r - r_s\|}{c} \right). \]  
(7)

We turn now to the problem of computing the wavefield \( U(r, t; r_s) \) corresponding to a perturbed velocity distribution
\[ C(r) = c + \delta c(r), \]
where we explicitly assume that the perturbation is small, i.e.,
\[ |\delta c(r)| \ll c. \]

We define \( \delta u(r, t; r_s) \) as the difference between the wavefield corresponding to the perturbed medium and the one corresponding to the homogeneous reference medium (with the same source):
\[ U(r, t; r_s) = u(r, t; r_s) + \delta u(r, t; r_s). \]

The field \( u(r, t; r_s) \) corresponds to the direct wave, while the field \( \delta u(r, t; r_s) \) corresponds to the scattered field. All through the rest of the paper we will use the term “reflected” as synonymous with “scattered.” We have
\[ \left( \frac{1}{(c + \delta c)^2 \frac{\partial^2}{\partial t^2} - \nabla^2} \right) [u(r, t; r_s) + \delta u(r, t; r_s)] = 8\pi^2 c^3 \delta(r - r_s) S(t). \]

Using
\[ \frac{1}{(c + \delta c)^2} - \frac{2\delta c}{c^3} + \alpha(\delta c)^2 \]
we readily obtain
\[ \left( \frac{1}{c^2 \frac{\partial^2}{\partial t^2} - \nabla^2} \right) \delta u(r, t; r_s) = \Delta s(r, t; r_s), \]  
(8)

where
\[ \Delta s(r, t; r_s) = \frac{2\delta c(r)}{c^3} \frac{\partial^2 u}{\partial t^2} (r, t; r_s) + \alpha(\delta c, \delta u)^2. \]  
(9)

We see that the reflected field \( \delta u \) generated by a perturbation \( \delta c \) in velocity can be interpreted as a secondary field propagating in the unperturbed medium (8) and due to secondary sources excited by the primary field (9).

As \( \delta c \) is assumed to be small, the term \( \alpha(\delta c, \delta u)^2 \) can be neglected, and it must be emphasized that this approximation contains, in particular, the Born approximation: only first-order scattering is considered, so that higher order effects (as multiple reflections) are neglected.

As the solution of (3) was obtained using (5), the solution of (8) is
\[ \delta u(r, t; r_s) = \frac{1}{4\pi} \int dr' \frac{1}{\|r - r'\|} \Delta s \left( r', t - \frac{\|r - r'\|}{c}; r_s \right). \]

i.e.,
\[ \delta u(r, t; r_s) = \frac{1}{2\pi c^3} \int dr' \frac{1}{\|r - r'\|} \frac{\partial^2 u}{\partial t^2} \left( r', t - \frac{\|r - r'\|}{c}; r_s \right) \delta c(r'). \]

Using (7) we have
\[ \frac{\partial^2 u}{\partial t^2} (r, t; r_s) = 2\pi c^3 \frac{1}{\|r - r_s\|} S \left( t - \frac{\|r - r_s\|}{c} \right). \]

This gives
\[ \delta s(r, t; r_s) = \int dr' \frac{1}{\|r - r'\| \cdot \|r_s - r'\|} S \left( \frac{\|r - r'\| + \|r_s - r'\|}{c} \right) \delta c(r'), \]
which can finally be written
\[ \delta u(r, t; r_s) = \left( \int dr' \frac{1}{\|r - r'\| \cdot \|r_s - r'\|} \delta \left( t - \frac{\|r - r'\| + \|r_s - r'\|}{c} \right) \delta c(r') \right) \ast S(t), \]  
(10)

where the symbol \( \ast \) denotes time convolution. For coincident source and receiver (zero offset) we have \( r = r_s \), and (10) simplifies to
\[ \delta u(r, t) = \left( \int dr' \frac{1}{\|r - r'\|} \delta \left( t - \frac{\|r - r'\|}{c/2} \right) \delta c(r') \right) \ast S(t), \]  
(11)

where we write \( \delta u(r, t) \) for \( \delta u(r, t; r_s) \).
These equations have a clear physical meaning. Equation (11) for instance means that the signal arriving at time \( t \) at point \( r \) comes from a sphere of radius \( ct/2 \) centered at point \( r \). For the finite-offset case (10) the signal comes from the surface of an ellipsoid with foci at points \( r \) and \( r_s \). Replacing \( \delta c(r') \) in (10) by a Dirac function we can obtain the effect of a single scatterer (a concentrated perturbation of velocity). We see in particular that the diffracted field is isotropic and that the source's time function has been differentiated twice. The term \( \| r - r' \|^{-1} \) represents the geometric decay of amplitude of the diffracted field itself, and the term \( \| r_s - r' \|^{-1} \) represents the decay of amplitude of the primary field.

2-D forward problem

Most seismic reflection data are still collected along a (nearly) straight line. It is clear that no serious attempt can be made for the interpretation of such data sets unless we can make the hypothesis that there are no strong lateral variations in the geology (i.e., in the direction perpendicular to the profile). More precisely, we assume that the geometric model is invariant against translation parallel to the \( y \)-axis (taking the \( x \)-axis as the profile line): 

\[
\delta c(x, y, z) = \delta c(x, y = 0, z) = \delta c(x, z).
\]

This assumption is often accompanied by the assumption that the primary field is generated by line sources perpendicular to the profile, because this leads to a wavefield \( \delta u \) which is also independent of \( y \) (thus reducing the forward problem to a “false” 2-D problem). We do not need this hypothesis: our sources are point sources (with 3-D spreading) illuminating a 2-D medium.

If \( \delta c \) is independent of \( y \), the sum over the \( y \)-coordinate can be performed analytically (see appendix A) yielding

\[
\delta u(x, t; x_0) = \frac{4}{c} \left( \int dx' \int dz' \frac{H(t - t_0)}{t \sqrt{(z' - z)^2 - t_1^2}} \delta c(x', z') \right) \ast S(t),
\]

where \( H(t) \) is the Heaviside function and where

\[
t_0 = \frac{d_1 + d_2}{c},
\]

\[
t_1 = \frac{2(d_1^2 + d_2^2)}{c^2} \frac{(d_1^2 - d_2^2)^2}{c^4 t_1^2}
\]

and

\[
d_1 = \sqrt{[(x - x)^2 + z'^2]},
\]

\[
d_2 = \sqrt{[(x - x)^2 + z'^2]}.
\]

For coincident source and receiver (zero offset), this equation simplifies to

\[
\delta u(x, t) = \frac{4}{c} \left( \int dx' \int dz' \frac{H(t - t_0)}{t \sqrt{(z' - z)^2 - t_1^2}} \delta c(x', z') \right) \ast S(t),
\]

where

\[
t_0 = \frac{\sqrt{(x - x)^2 + z'^2}}{c/2}.
\]

Replacing \( \delta c(x', z') \) in (12) by a Dirac function we obtain the response at the point \( x \) of a line scatterer excited by a point source at \( x_s \).

Let us point out that (13) is completely equivalent to (6) of Cohen and Beinstein (1979) (although (13) looks simpler).

3. The Inverse Problem

Equations (10) to (13) allow the computation of the reflected field at any receiver position (and for any source position) if the velocity perturbation \( \delta c \) is given everywhere in the space. In attacking the inverse problem we must take into account the fact that the source and receiver positions are inherently discrete (and finite in number). The time variable is also discretized in present-day recording systems, but this discretization is less essential to the theory; and throughout this paper we will use the notations corresponding to a continuous time variable, for simplifying the notation. The observations take the form

\[
\delta u(r_i; t; r_j),
\]

where \( r_i \) and \( r_j \) represent the positions of the receivers and sources, respectively, and where the subscript \( o \) stands for "observed." Through the rest of the paper a latin subscript stands for discrete receiver position while a greek subscript stands for discrete source position.

A quite general way for describing estimated errors in the data set is by introducing the covariances

\[
C_{\delta u}(r_i; t; r_j; r_k),
\]

which can take a very general form. For simplicity, we will later assume the form

\[
C_{\delta u}(r_i; t; r_j; r_k; r_m) = \sigma^2 \sigma_\delta \delta_{ij} \delta_{km},
\]

which means that we assume that errors are uncorrelated and that the estimated error for the \( i \)th trace is \( \sigma_\delta \).

In the interest of a compact notation, let us denote by \( \delta u \) any generic data vector; by \( \delta u_0 \), the observed vector, and by \( C_0 \) the covariance operator with kernel \( C_0(r_i; r_j; r_k; r_m) \). If the time variable is discrete (as it is assumed in (14)) then this kernel is a matrix (more precisely, a diagonal six-dimensional array).

The question now is the following: can we infer from the discrete data set \( \delta u_0 \) a realistic model \( \delta c \) for the earth? The response can be affirmative only if we have some a priori information about admissible earth models. A quite powerful way to specify the a priori information is by giving an a priori model \( c_0(r) \) and a covariance function \( C_c(r; r') \), which describes the confidence we have in the a priori model \( c_0(r) \).
use an iterative algorithm of the form
\[
\delta c_{k+1} = \delta c_k + \gamma_k \Delta c_k,
\]
where \(\Delta c_k\) is an arbitrary, suitably chosen, direction in the model space and where the scalar \(\gamma_k\) is computed for minimizing \(S\) along that direction. A simple calculation gives the optimum value for \(\gamma_k\):
\[
\gamma_k = \frac{\Delta c_k^* \gamma_c}{(G \Delta c_k)^* C_{u}^{-1} (G \Delta c_k) + \Delta c_k^* C_{c}^{-1} \Delta c_k},
\]
where \(\gamma_c\) denotes the value of \(\gamma\) at the point \(\delta c_k\).

There are many ways for choosing adequate directions \(\Delta c_k\), as for instance the conjugate directions (Fletcher and Reeves 1964):
\[
\Delta c_k = \gamma_k + \frac{\gamma_k^* \gamma_k}{\gamma_k^* \gamma_k - 1} \Delta c_{k-1},
\]
or the direction of steepest descent:
\[
\Delta c_k = \gamma_k,
\]
or the modified gradient direction:
\[
\Delta c_k = W \gamma_k,
\]
where \(W\) is a suitably chosen, positive definite operator which accelerates the convergence, or simplifies the computations (or both).

If, for instance, we take
\[
W = C_c
\]
we obtain the algorithm
\[
\delta c_{k+1} = \delta c_k + \gamma_k \Delta c_k,
\]
\[
\Delta c_k = C_c G^* C_{u}^{-1} (\delta u_k - G \delta c_k) - \delta c_k,
\]
\[
\gamma_k = \frac{\Delta c_k^* \gamma_c}{(G \Delta c_k)^* C_{u}^{-1} (G \Delta c_k) + \Delta c_k^* C_{c}^{-1} \Delta c_k},
\]
The algorithm converges because at each step the value of \(S\) diminishes; we can be sure that it converges to the right solution, because at the limit we have \(\Delta c_k = 0\), which is equivalent to \(\gamma_c = 0\), i.e., the gradient of \(S\) vanishes (and we know that the solution is unique).

A discussion of the number of iterations needed to obtain a good approximate of the solution will be given after the physical interpretation of the formulae.

Our first iterate is
\[
\delta c_1 = \alpha_s C_c G^* C_{u}^{-1} \delta u_k.
\]
Let us interpret this result starting with the simple case of a zero-offset data set.
With (11) the solution of the forward problem can be written as
\[ \delta u(r_1, t) = \int \frac{1}{\| r_1 - r \|^2} \mathcal{S} \left( t - \frac{\| r_1 - r \|}{c/2} \right) \delta c(r). \]
The formal solution of the forward problem was written in (16) as
\[ \delta u = G \delta c, \]
or, in explicit form,
\[ \delta u(r_1, t) = \int \mathcal{D} G(r_1, t | r) \delta c(r). \]
(30)
where \( G(r_1, t | r) \) represents, by definition, the kernel of \( G \):
\[ G(r_1, t | r) = \frac{1}{\| r_1 - r \|^2} \mathcal{S} \left( t - \frac{\| r_1 - r \|}{c/2} \right). \]
(31)
The observed values are denoted by \( \delta u_o(r_1, t) \). Let us assume here that for this zero-offset case the estimated errors take the simple form (14), i.e.,
\[ C_u(r_1, t | r_j, t') = \sigma_i^2 \delta_{ij} \delta_{tt}. \]
With
\[ \delta u_o = C_u^{-1} \delta u, \]
it is easy to see that we have
\[ \delta u_o(r_1, t) = \frac{1}{\sigma_i^2} \delta u(r_1, t), \]
(32)
which means that the application of \( C_u^{-1} \) to the observed reflected field \( \delta u_o \) has the effect of weighting the traces by the inverse of their estimated error.
The next step for understanding (29) is the interpretation of the action of \( G^* \) or \( \delta u'_o \). Let us denote by \( \delta c' \) the corresponding result:
\[ \delta c' = G^* \delta u'_o. \]
If we denote by \( G^*(r_1 | r, t) \) the kernel of \( G^* \), this equation is written, explicitly,
\[ \delta c'(r) = \sum \int \mathcal{D} G^*(r_1 | r, t) \delta u'_o(r_1, t). \]
From the definition of the adjoint of a linear operator, we have
\[ G^*(r_1 | r, t) = G(r_1, t | r) \]
and, using (31) we have
\[ \delta c'(r) = \sum \frac{1}{\| r_1 - r \|^2} \int \mathcal{D} \mathcal{S} \left( t - \frac{\| r_1 - r \|}{c/2} \right) \delta u'_o(r_1, t). \]

Denoting the integral by \( I \) we successively have
\[ I = \int \mathcal{D} \left( \delta \left( t - \frac{\| r_1 - r \|}{c/2} \right) \right) \delta u'_o(r_1, t) \]
\[ = \int \mathcal{D} \delta \left( t - \frac{\| r_1 - r \|}{c/2} \right) (\delta u'_o(r_1, t) * \mathcal{S}(-t)), \]
where we have used the property
\[ \int \mathcal{D} f(t) * g(t) h(t) = \int \mathcal{D} f(t) (g(t) * h(-t)). \]
Defining
\[ \delta u_o(r_1, t) = \delta u'_o(r_1, r) * \mathcal{S}(-t) \]
(33)
we finally obtain
\[ \delta c'(r) = \sum \frac{1}{\| r_1 - r \|^2} \delta u_o(r_1, t = \frac{\| r_1 - r \|}{c/2}). \]
(34)
We see thus that \( \delta c'(r) \) (i.e., the action of \( G^* \) over the corrected data set \( \delta u'_o(r_1, t) \)) is obtained through two actions: we must first cross-corrrelate the corrected data set with the source function \( \mathcal{S}(t) \) (eq. (33)); then the value of \( \delta c' \) at a point \( r \) is obtained by performing a sum over all the traces along the hyperboloid in the \( (t, r_1) \) space
\[ t = \frac{\| r_1 - r \|}{c/2}, \]
(eq. (34), i.e., by performing a classical "migration by diffraction stack" (Kirchhoff sum) with a multiplicative factor
\[ \frac{1}{\| r_1 - r \|^2}. \]
This factor is not identical to the one directly obtained using the Kirchhoff formula (Schneider 1978), because the Kirchhoff method merely performs a downward continuation of the observed field, while we here perform an inversion for obtaining a velocity model.
It is surprising to discover that the action of \( G^* \) corresponds to two classical treatments of seismic reflection data sets (i.e., cross-correlation with the source and migration). We can remark that our theory requires a cross-correlation with the source function, rather than a deconvolution. As cross-correlation is a much more stable operation than deconvolution, this gives some confidence on the stability of the inverse solution defined by our least-squares theory. On the other hand, it is not yet clear whether the convergence of the algorithm cannot be ameliorated if it is applied to a previously deconvolved data set.
We emphasize that the interpretation of \( G^* \) as a "cross-correlation plus migration" is possible only because we think of \( G \) as an operator, working over a conveniently defined space, rather than as a matrix acting over a discretized space.
This is a quite general conclusion in any inverse problem theory: discretization, if any, must be reserved for the final computations, and not for developing the formulæ.

The last step for the implementation of (29) is the application of the operator \( C_c \) to \( \delta c' \). This gives the final result

\[
\delta c_c(r) = \gamma \int dr' C_c(r, r') \delta c'(r').
\] (35)

As covariance operators are in general smoothing operators (as is the case for our example (15)), we see that \( \delta c_c \) is merely a filtered version of \( \delta c' \) (\( \gamma \) is a scalar).

Putting together (32) to (35), we list all the operations needed for the computation of the first iterate of the solution of the inverse problem:

**General formula:**
\[
\delta c_c = \gamma C_c G^* C_a^{-1} \delta u_n;
\]

**Weighting of the data:**
\[
\delta u_n(r_i, t) = \frac{\delta u_n^c(r_i, t)}{\sigma_i^2};
\]

**Cross-correlation with the source function:**
\[
\delta u_0(r_i, t) = \delta u_n^c(r_i, t) \ast \tilde{S}(-t);
\] (36)

**Migration:**
\[
\delta c'(r) = \sum_i \frac{1}{\|r_i - r\|^2} \delta u_0(r_i, t = \frac{\|r_i - r\|}{c/2});
\]

**Smoothing:**
\[
\delta c_c(r) = \gamma \int dr' C_c(r, r') \delta c'(r').
\]

As all steps correspond essentially to a sophisticated migration, we have thus arrived at the fundamental conclusion that migration of seismic reflection data only corresponds to a first iterate of a rigorous solution of the inverse problem.

Returning to the general iterative sequence (25), we see that for the computation of \( \gamma \), as well as for subsequent iterations, we need, in addition to the operations already described for the first iteration, the computation of
\[
\delta u_n = G \delta c_n.
\]

\( \delta u_n(r_i, t) \) clearly represents the computed data for the \( n \)th model. We have from (11)
\[
\delta u_n(r_i, t) = \left( \int dr \frac{1}{\|r_i - r\|^2} \delta \left( t - \frac{\|r_i - r\|}{c/2} \right) \delta c_n^c(r) \right) \ast \tilde{S}(t).
\]

Let us assume that \( r_i \) is at the surface of a half-space. For given \( r_i \) we can use spherical coordinates centered at \( r_i \). We have then
\[
\delta u_n(r_i, t) = \left[ \int_0^\pi d\phi \int_0^\pi d\theta \int_0^{\infty} d\rho \rho^2 \sin \theta \frac{1}{\rho^2} \delta \left( t - \frac{\rho}{c/2} \right) \delta c_n^c(r) \right] \ast \tilde{S}(t)
\]
\[= \frac{1}{2c} \left[ \int_0^\pi d\phi \int_0^\pi d\theta \sin \theta \delta c_n^c(r) \right] \ast \tilde{S}(t) \] (37)

with
\[
2\rho = 2\|r_i - r\| = ct,
\]

which corresponds to the fact that under homogeneous conditions, the signal arriving at the point \( r_i \) at the time \( t \) is obtained by integration of the model along the surface of a half-sphere, centered at \( r_i \), and of radius \( ct/2 \) (so that the time for going and coming back is \( t \)).

We see thus that the rigorous solution of the linearized inverse problem can be reduced to an iterative series of "migration plus forward modeling".

Let us remark here that, although each solution \( \delta c_n \) is defined everywhere in space, it has to be evaluated over a grid dense enough to compute the sum (37) with sufficient accuracy.

Equations (36)–(37) have been discussed for zero-offset data for a 3-D problem. It is easy to find the results corresponding to the other three situations considered in section 2, namely, finite-offset data for a 3-D problem, and zero- and finite-offset data, respectively, for a 2-D problem. The inverse algorithm (29) remains rigorously the same with the only difference that the operator \( G \) in (31) has to be replaced by the operator corresponding to the suitable forward problem ((10), (12), or (13), respectively).

Let us give these results without demonstration.

For finite-offset data for a 3-D problem we obtain
\[
\delta c_c = \gamma C_c G^* C_a^{-1} \delta u_n,
\]
\[
\delta u_n^c(r_i, t; r_j) = \frac{\delta u_n^c(r_i, t; r_j)}{\sigma_i^2},
\]
\[
\delta u_0(r_i, t; r_j) = \delta u_n^c(r_i, t; r_j) \ast \tilde{S}(-t),
\]
\[
\delta c'(r) = \sum_i \frac{1}{\|r_i - r\|^2} \delta \left( t - \frac{\|r_i - r\|}{c/2} \right) \delta u_0^c(r_i, t = \frac{\|r_i - r\| + \|r_j - r\|}{c/2}; r_j),
\]
\[
\delta c_c(r) = \gamma \int dr' C_c(r, r') \delta c'(r').
\] (38)
where \( r_i \) and \( r_a \) are, respectively, receiver and source location, and where \( \sigma_{ia} \) is the estimated error for the \( ia \)th trace. When \( \| r_i - r_a \| \to 0 \) and there is only one receiver position per source, (38) reduces to (36).

For subsequent iterations we also have to solve the forward problem (10):

\[
\delta u_a(r_i, t; r_a) = \int dr \frac{1}{\| r_i - r \| \| r_a - r \|} \left( t - \frac{\| r_i - r \| + \| r_a - r \|}{c} \right) \delta c_a(r) \ast S(t).
\]

It can be demonstrated that, for \( r_i \) and \( r_a \), on the surface of a half-space, this equation reduces to

\[
\delta u_a(r_i, t; r_a) = \frac{1}{2c} \int_0^\pi d\phi \int_0^\pi d\chi \sin \chi \int dx \delta c_a(x, z) \ast S(t),
\]

where the angle \( \chi \) is defined in Fig. 1, and where the point \( r \) lies on the ellipsoid

\[
\| r_i - r_a \| + \| r_a - r \| = ct.
\]

![Fig. 1. Geometry of equal traveltime for zero offset (left) and finite offset (right).](image)

We remark that (39) is the same as expression (37) for zero offset.

For two-dimensional geological structures we obtain

\[
\delta c_1 = \alpha_e C_e G^* C_r^{-1} \delta u_a,
\]

\[
\delta u_a(r_i, t) = \frac{\delta u_a(r_i, t)}{\sigma^2_{ia}},
\]

\[
\delta u_a(r_i, t) = \delta u_a(r_i, t) \ast S(-t),
\]

\[
\delta \tilde{u}_a(r_i, t) = \frac{1}{c/2} \int_{-t}^{t} dt \delta \tilde{u}_a(r_i, t) \ast S(t),
\]

\[
\delta c'(r) = 2 \sum_i \delta \tilde{u}_a(r_i, t = \frac{\| r - r_i \|}{c/2}),
\]

\[
\delta c_2(r) = \alpha_e \int dr' C_{r}(r, r') \delta c'(r'),
\]

where

\[
r_i \equiv (x_i, 0, 0),
\]

\[
r \equiv (x, 0, z).
\]

It is interesting to note that these equations are very similar to the set (36) corresponding to the 3-D problem. The important difference is the introduction of \( \delta u_a \). As \( \delta c' \) is obtained from \( \delta u_a \), in a way similar to that used for obtaining \( \delta c' \) from \( \delta u_a \) in (36), we can conclude that \( \delta u_a \) represents a data set corrected for the effect of the (infinite) lateral extent of the structures.

The solution of the forward problem was given by (13):

\[
\delta u_a(r_i, t) = \frac{4}{c} \int dx \int dr \frac{H(t - t_0)}{\sqrt{(t^2 - t_0^2)}} \delta c_a(x, z) \ast S(t).
\]

For \( r_i \) on the surface of a half-space, this can be very efficiently calculated by computing both for each \( x \), the following "equivalent radial model":

\[
\delta \tilde{u}_a(x_i, t) = \int_0^\pi d\theta \delta c_a(x_i, z),
\]

where \( \theta \) is a polar angle centered at \( x_i \), and where

\[
(x - x_i)^2 + z^2 = c^2 t^2,
\]

and then by computing

\[
\delta u_a(x_i, t) = \frac{2}{c} \int_{-t}^{t} dt \frac{\delta \tilde{u}_a(x_i, t)}{\sqrt{(t^2 - t_0^2)}} \ast S(t).
\]

The expressions for finite-offset data for a 2-D problem are:

\[
\delta c_1 = \alpha_e C_e G^* C_r^{-1} \delta u_a;
\]

\[
\delta u_a(r_i, t; r_a) = \frac{\delta u_a(r_i, t; r_a)}{\sigma^2_{ia}};
\]

\[
\delta \tilde{u}_a(r_i, t; r_a) = \delta \tilde{u}_a(r_i, t; r_a) \ast S(t);
\]

\[
\delta c'(r) = \frac{4}{c} \sum_i \sum_a \int dt \frac{H(t - t_0)}{\sqrt{(t^2 - t_0^2)}} \delta \tilde{u}_a(r_i, t; r_a);
\]

\[
\delta c_2(r) = \alpha_e \int dr' C_{r}(r, r') \delta c'(r'),
\]

where

\[
r_i \equiv (x_i, 0, 0),
\]

\[
r_a \equiv (x_a, 0, 0),
\]

\[
r \equiv (x, 0, z).
\]

The forward problem is solved by using (12). It is easy to see that the integral is extended over the interior of an ellipsoid with foci at \( r_i \) and \( r_a \).

Equations (40) and (43) only define the model \( \delta c_1 \) on the plane \( (x, 0, z) \), but \( \delta c_1 \) is, by definition, invariant under translation parallel to the \( y \)-axis. The reader can verify that the model thus obtained is identical to the one obtained using the 3-D formulae (36) and (38) if an infinite correlation length \( L_y \) in the \( y \)-direction is a priori assumed.
4. Discussion

Free surface

We have not taken into account the fact that the medium has a free surface. The effect of this surface can be treated approximately by adding to the actual source its virtual, inverted image with respect to the free surface. If we assume that the actual source is close enough to the surface, then we can approximate the double source by a single source with a more complicated time function. If we reinterpret then our function $S(t)$ as corresponding to that equivalent source, all the previous results remain approximately valid.

Applicability of the theory to actual data

The success of the crude "exploring reflector model" for the migration of seismic reflection data is due in part to the internal compatibility between this model and the traditional data processing prior to migration. In particular, amplitude equalization along traces and between traces with different offset plays a major role in the production of a CDP stack. Equating a CDP stack to a zero-offset profile is not compatible with the use of our theory, so a special effort has to be made for preserving the true amplitudes of seismograms if we want inversion to be better than migration. In particular, this means that inversion will only give interesting results when applied to unstacked seismic reflection data, unless the forward problem is modified and includes the description of the effects that CDP-stacking has on the data sets.

Number of iterations

The solution of our linearized forward problem was written $\delta u = G \delta c$. We have shown that each iteration of the gradient algorithm for solving the inverse problem consists essentially of the application of $G^*$ to the current value of the residuals. We have also shown that $G^*$ corresponds to a migration. Although gradient methods are known to converge very slowly in general, experience shows that a single migration gives models which resemble the true model. Therefore we expect to gain valuable information with only a few iterations. This should be even more so with highly redundant (i.e., unstacked) data.

Reference medium

The hypothesis of a homogeneous reference medium (with constant velocity) is somewhat difficult to accept. A proper generalization of these results to the inhomogeneous case is in preparation. For the moment, some simple tricks may be suggested. For instance, if we assume that the velocity of the medium is essentially depth-dependent, we can simply replace the expressions $||r_i - r||$ and $||r_s - r||$ appearing in (30)-(32) by the lengths of the corresponding ray paths between the points $r_i$ and $r$, and the point $r$. This approximation is probably good enough unless the medium produces strong focusing effects. If the medium is not only vertically but also laterally heterogeneous, ray-tracing should be used.

Static corrections

The solution can be forced to have some horizontal continuity (by choosing $L_i \gg L_s$). The predicted data at each iterative step will then also have this horizontal continuity. If at each iterative step we compare each predicted trace with the corresponding observed trace, it is easy to deduce the best static correction to be applied to each observed trace for optimizing the fit. We guess that, at the limit, the solution will be similar to that of a more general inverse problem including the static correction in its formulation.

5. Conclusion

The main conclusion of the paper is that the rigorous solution of the linearized seismic reflection inverse problem can be obtained using the classical methods of migration (plus forward modeling). The solution incorporates weighting of the data (by its estimated errors), and a priori information on the model.

Acknowledgments

We thank P. Bernard, A. Cisternas, G. Jobert, P. Lalil, and A. Nercissian for helpful comments and suggestions. This work has been supported partly by the RCP 264 (Etude Interdisciplinaire des Problemes Inverses) and by the ATP Calcul Vectoriel 1982.

Appendix

We demonstrate equation (12). We have

$$
\delta u(x_i, 0, z, t; x_s, 0, z_s) = \int dx \int d\gamma (x, z) \int dy \frac{\delta t - (y(d_x^2 + y^2) + (d_y^2 + y^2)c)}{(d_x^2 + y^2)^{1/2}(d_y^2 + y^2)^{1/2}} \ast S(t),
$$

where

$$
d_i = \sqrt{(x_i - x)^2 + (z - z_i)^2} \quad d_s = \sqrt{(x_s - x)^2 + (z_s - z)^2}.
$$

Using

$$
\delta(\psi(y)) = \sum_k \frac{\delta(y - y_k)}{|\psi'(y_k)|}
$$

we have

$$
\int dx \int d\gamma (x, z) \int dy \frac{\delta t - (y(d_x^2 + y^2) + (d_y^2 + y^2)c)}{(d_x^2 + y^2)^{1/2}(d_y^2 + y^2)^{1/2}} \ast S(t),
$$

where

$$
d_i = \sqrt{(x_i - x)^2 + (z - z_i)^2} \quad d_s = \sqrt{(x_s - x)^2 + (z_s - z)^2}.
$$

Using

$$
\delta(\psi(y)) = \sum_k \frac{\delta(y - y_k)}{|\psi'(y_k)|}
$$
for $\psi(y_k) = 0$, we obtain

$$
\delta\left( t - \frac{\sqrt{d_1^2 + y^2}}{c} + \frac{\sqrt{d_2^2 + y^2}}{c} \right) = H\left( t - \frac{d_1 + d_2}{c} \right) \sum_{k=1,2} \frac{\delta(y - y_k)}{t \left| y_k \right| \sqrt{\left( \frac{d_k}{c} \right)^2 + \frac{y_k^2}{c^2}}},
$$

where

$$
y_k = \pm \frac{1}{2}c\left( t^2 - \frac{2(d_1 + d_2)}{c^2} + \frac{(d_1^2 - d_2^2)^{1/2}}{c_4 t^2} \right).
$$

This gives directly (12) with

$$
t_0 = \frac{d_1 + d_2}{c},
$$

$$
t_i^2 = \frac{2(d_1^2 + d_2^2) - (d_1^2 - d_2^2)^2}{c_4 t^2}.
$$

**References**


Claerbout, J.F. 1971, Toward a unified theory of reflector mapping, Geophysics 36, 467–481.


Morse, P.M. and Feshbach, H. 1953, Methods of Theoretical Physics, McGraw-Hill, New York.