Geophysical Journal (1988) **95**, 449–462

Simultaneous determination of source wavelet and velocity profile using impulsive point-source reflections from a layered fluid*

K. Bube,* P. Lailly,† P. Sacks,‡ F. Santosa§ and W. W. Symes¶

* *Department of Mathematics, University of Washington, Seattle WA* 98195 *USA*
† *Institut Français du Petrol, B.P.* 311, 92506 *Rueil-Malmaison Cedex, France*
‡ *Department of Mathematics, Iowa State University, Ames, IA* 50011, *USA*
§ *Department of Mathematical Sciences, University of Delaware, Newark, DE* 19716, *USA*
¶ *Department of Mathematical Sciences, Rice University, Houston, TX* 77251, *USA*

Accepted 1988 April 1. Received 1988 February 23; marginal form 1987 July 17

SUMMARY

The determination of source signature is a major calibration problem in reflection seismology. This 'deconvolution' problem is conventionally approached by way of statistical methods, by direct measurement, or by the location of a clean reflection in an otherwise quiet part of a reflection section. We show that a quasi-impulsive, isotropic point source may be recovered simultaneously with the velocity profile from reflection data over a layered fluid, in linear (perturbation) approximation. Our approach is completely deterministic, and does not depend on the presence of an isolated reflection in a quiet part of the section, as we illustrate with a numerical example. After describing the algorithm and a numerical implementation, we give a complete mathematical treatment, which shows that our estimates of source wavelet and velocity profile are stable in a certain sense. Because of this stability property we conjecture that our approach to simultaneous estimation of source and medium parameters actually applies to a much broader class of models than that treated here.

Key words: waveform inversion, source signature estimation

1 INTRODUCTION

A major calibration problem in reflection seismology is that of source-signature identification. The seismic section depends not just on the mechanical structure of the Earth, the elucidation of which is the aim of reflection seismology, but also on the time dependence and spatial distribution of the seismic energy source. The separation of these two factors (Earth structure, energy source) is often regarded as a deconvolution problem ('source-signature deconvolution') and is commonly attacked with statistical tools (predictive deconvolution, ARMA modeling, maximum likelihood estimation), which aim simultaneously to suppress other sorts of seismic 'noise' (notably multiple reflections). These widely-used methods have been criticized as being based on unwarranted assumptions about both source and Earth structure (Ziolkowski 1984). An alternate method of source calibration is actual measurement of the direct wave, or use of a strong reflection in an otherwise relatively quiet part of the section.

In this paper we investigate the possibility that the time-dependence of an isotropic point-source might be determined directly (and deterministically) from the section, even when no isolated reflection event can be located. In effect, we ask whether the source parameters can be determined simultaneously with the earth model. We show that, under some restrictions, the answer is 'yes'.

This possibility of codetermination of source and velocity arises from the different dependence of their data influence on offset: that is, the effects of changing the source wavelet, respectively the velocity structure, *move out and scale differently*. These effects may be separated algebraically; there results an equation linking the data section with the velocity, sampled at several different, but related, depths. The wavelet has been eliminated from this relation altogether. Once the velocity has been determined, the wavelet may be extracted easily. On the other hand, functional-delay equations, such as the one we derive here for the velocity, are not commonplace in the mathematics of seismology, and it is not entirely obvious that such equations can be solved in a reasonable sense. Accordingly we give a complete theoretical treatment, ensuring that solutions can be found computationally as well. We implement an algorithm based on our theoretical developments, and show by numerical illustration that wavelet and velocity may indeed be separately determined even when their influence is coexistent in time.

To derive this result, we have imposed some hypotheses on both the earth model and the energy source. Those

* This research was supported by the Office of Naval Research under contracts N00014-83-K0051 and N00014-85-K0725, by National Science Foundation grants DMS-84-03148 and 86-03164 and by AFOSR 84-0252. Technical Report 87-17, June 1987.

concerning the earth model are:

(M1) that the Earth structure varies only with depth (layered medium);

(M2) that the Earth is a linearly elastic fluid, with only the sound velocity varying with position (depth);

(M3) that the variation of sound velocity with depth is somewhat smooth.

Concerning the source, we have assumed that

(S1) the spatial distribution of the source is point-like, to adequate approximation;

(S2) the source radiation pattern is isotropic, so that the source is described by a single function of time (wavelet);

(S3) the source wavelet is quasi-impulsive, i.e. differs from the Dirac delta function by a square-integrable function.

None of these assumptions (with the possible exception of S1) is valid for an accurate model of the typical seismic reflection experiment: The Earth is non-layered and inhomogeneous on interesting length scales and supports shear (as well as compressional) waves. Also the seismic source often has an anisotropic radiation pattern, and (most important) is bandlimited (not close to delta).

It seems clear that most of these assumptions could be relaxed, at least to some extent. The most important and involved step is the introduction of non-impulsive sources. It is plausible on the basis of velocity analysis (and is carefully justified on theoretical grounds in Santosa & Symes 1986) that bandlimited point-source data (with known source) determine a layered velocity structure, provided that the target structure is sufficiently rich in reflectors (thus is sufficiently *non-smooth*). We expect to be able to combine these bandlimited inversion ideas with those presented in this paper to codetermine bandlimited sources as well. We will further discuss the extent to which the restrictions M1–M3 and S1–S3 might be relaxed in the concluding section.

We note that Canadas & Kolb (1986) give numerical evidence that the velocity and source wavelet (and, to some extent, the density) of a layered fluid may be recovered from simulated seismic reflection data. They do not, however, give a theoretical justification for their procedure. Ramm (1987) has given a theoretical result for quasi-impulsive sources, which differs from ours in that it makes use of the low-frequency (rather than high-frequency) asymptotics of the response, hence is intrinsically restricted to the quasi-impulsive case.

The book by Lavrent'ev, Reznitskaya & Yakhno (1986) has recently come to our attention (January 1988). In Section 2.2 the authors consider a problem similar to that studied here. Their work is similar to ours in that they also find that the recovery of first-order perturbations in medium and source parameters hinges on the solution of a multiplicative-delay equation like (3.2) below. It differs from ours in that they do not use the plane-wave decomposition, and restrict their attention to perturbations about a homogeneous background model.

The model described above is quantified in the following boundary value problem, connecting

$\rho(x)$ —fluid density

$\lambda(x)$ —bulk modulus

$u(x, t)$—particle displacement field

$f(t)$ —source wavelet:

$$\rho(x)\frac{\partial^2 u}{\partial t^2}(x, t) = \nabla(\lambda(x)\nabla\cdot u(x, t))$$

in $\{(x, t): x_3 > 0\}$

$$\lambda(x)\nabla\cdot u(x, t) = -f(t)\delta(x_1, x_2)$$

for $x_3 = 0$

$$u(x, t) \equiv 0, \qquad t \ll 0.$$

We write $z = x_3$, and note that (M1) is stated $\rho = \rho(z)$, $\lambda = \lambda(z)$. We assume that $\lambda \equiv \text{const.}$, so that $c(z) = \lambda^{1/2}\rho(z)^{-1/2}$ parameterizes the medium (M2).

We remark that this is not a significant restriction: the problem of simultaneous determination of ρ and λ from known (and impulsive) source data is well understood (see e.g. Santosa & Symes 1987, and other work cited there). Extension of our present results to this more general case is routine. We have chosen to present the result for $\lambda \equiv$ *const.* merely to avoid obscuring the novelty of the present paper with irrelevant complexities.

The hypothesis (S3) is that

$$f(t) = \delta(t) + f_1(t)$$

with f_1 square-integrable. Causality is expressed by the requirement that $f_1(t) = 0$, $t < 0$ and we also assume that $f_1(t) \equiv 0$, $t \geq t_{\max}$. Consequently, f defines an invertible convolution operator (on $L^2[0, T]$ for any $T > 0$).

We idealize the seismogram as the surface trace of the vertical displacement field $u = (u_x, u_y, u_z)$ viewed as a function of the velocity profile c and the wavelet f:

$$s[c, f] = u_z(z = 0).$$

We shall use the commonplace notation of $\delta(\cdot)$ to denote perturbations of $(\cdot)$; the distinction from the Dirac delta function will be clear from the context in which it is used. Our main result concerns the *perturbational seismogram* $\delta s[c, f]$, which is the result of first-order perturbation of the velocity profile and source wavelet:

$$\delta s[c, f] = \delta u_z(z = 0),$$

where the perturbation field δu satisfies

$$\rho\frac{\partial^2 \delta u}{\partial t^2} = \nabla\nabla\cdot\delta u - \delta\rho\frac{\partial^2 u}{\partial t^2}$$

$$\nabla\cdot\delta u = -\delta f(t)\delta(x_1, x_2)$$

$$\delta u \equiv 0 \qquad t \ll 0$$

and $\delta\rho = -2\delta c\cdot c^{-3}$.

Our main result may be stated:

The velocity profile and source wavelet perturbations δc, δf are uniquely determined by the perturbational point-source seismogram δs, under hypotheses M1–M3 and S1–S3.

Note that we do *not* assume the presence of an isolated reflector. Thus the influences of δf and δc are coexistent in time. Nonetheless they can be recovered separately. In our view, it is this feature which makes the present result interesting despite the severe restrictions imposed on our model.

This result concerns the perturbational ('variable-background Born approximation') problem. The technique is constructive, and estimates δc, δf in terms of δs. It seems clear that regularity results similar to those explained in Symes (1986b, c) could be combined with the conclusions presented here to yield unique determination of f and c from s. Such matters will be discussed elsewhere.

We begin the study of this problem by transforming the point-source seismogram to a plane-wave (p-tau) section, and determining its structure, in Section 2. We use this structural information in Section 3 to derive a constructive procedure for separate determination of source and velocity perturbations. We suggest numerical techniques for the solution of least-squares formulations of the inverse problems and present the results of numerical experiments based on one of the possible implementations, in Section 4. These experiments were performed by Paul Sacks at Iowa State University. Section 5 contains the proofs of the main mathematical result and of several technical lemmas needed in Section 3. We end the paper with a discussion of possibilities for relaxing the restrictions, and extending the scope, of our results (Section 6).

2 STRUCTURE OF THE PERTURBATIONAL SEISMOGRAM

Our main tool in this paper is the *plane-wave decomposition*: since the coefficients are independent of time and of the horizontal coordinates, the Radon integral·

$$\iint dx\,dy\,u(x, y, z, t + p \cdot x)$$

produces for each p a field satisfying a system of partial differential equations in z, t. In fact, we are only interested in a finite depth interval $0 \le z \le z_{\max}$, so without loss of generality we assume

$$0 < c_* \le c(z) \le c^*, \qquad 0 \le z$$

for suitable constants c_*, c^*. In Santosa & Symes (1985, Appendix) it is shown that for sufficiently small p, the plane-wave component of normal displacement

$$U(z, \tau, p) = \iint dx\,dy u_z(x, y, z, \tau + px)$$

solves the boundary value problem

$$(\rho - \lambda p^2)\frac{\partial^2 U}{\partial \tau^2} - \lambda\frac{\partial^2 U}{\partial z^2} = 0$$

$$-\lambda\frac{\partial U}{\partial z}(0, \tau, p) = f(\tau) \tag{2.1}$$

$$U \equiv 0, \qquad \tau \ll 0.$$

Recall that the wave velocity is given by $c(z) = \lambda^{1/2}\rho^{-1/2}(z)$. After normalization, we may assume for the rest of the paper that $\lambda \equiv 1$, so that the mechanics are described by c alone.

The equation (2.1) is hyperbolic only so long as $c\,|p| < 1$ (*precritical slowness*). We define

$$z_{\max}(p) = \sup\{z : c(z')p < 1 \text{ for } 0 \le z' < z\}$$

and consider (2.1) and related equations only in the slab $[(z, t) : 0 \le z < z_{\max}(p)]$ for each p.

We note that the Radon transform integral given above must in general be modified by the introduction of a mute or cutoff for large but still precritical p. For a suitable choice of mute, the components still satisfy the plane-wave equation (2.1) up to an error which may be controlled by the techniques presented here: see Santosa & Symes (1988).

The plane-wave seismogram (p-tau section) S is the surface ($z = 0$) time history of the plane wave normal displacement fields

$$S[c, f](\tau, p) = U(0, \tau, p).$$

The topic of our paper is the 'inverse' problem:

Given a measurement of G of plane-wave reflection data, find a velocity-source pair (c, f) for which

$$S[c, f](\tau, p) = U(0, \tau, p) = G(\tau, p).$$

This formulation supposes implicitly that the (τ, p)-pairs at which the data G is given are precritical for the sought velocity profile c, which presumption is itself a nonlinear constraint on c.

Now it is natural to think that any feasible measurement G would be contaminated by error, hence might not fit any model exactly. Accordingly, it is popular to replace the 'exact inverse' problem above by a best-fit formulation. A common choice of error measure is the mean square, which leads to the least-squares inverse problem:

Find c, f to minimize

$$\iint d\Omega(\tau, p)\,|S[c, f](\tau, p) - G(\tau, p)|^2,$$

where $d\Omega$ is a measure (i.e. possibly nonuniform weight, continuous covariance matrix).

In this paper we shall limit our weights to those of the form

$$d\Omega(\tau, p) = d\tau\,d\mu(p).$$

Thus time points are uniformly weighted in each trace ($p = const.$), but we allow the traces to be weighted according to the measure $d\mu(p)$. For example, if $d\mu(p) = dp$, then all traces are uniformly weighted (over the domain of integration), whereas if $d\mu(p)$ is a finite sum of point masses, then the error measure above is concentrated on the corresponding finite set of traces.

The problem stated above is nonlinear. To begin our study in this paper we shall assume that the velocity is a sum of a (smooth, slowly varying) background velocity c and a (possibly rapidly varying) perturbation δc. We make a similar assumption concerning the source, i.e. that it is the sum of a background wavelet $f = \delta + f_1$, f_1 square-integrable, and a perturbation δf. We shall study the corresponding perturbation δS in the seismogram, which depends linearly on δc, δf.

Formally, to first order, the result of perturbing the velocity profile $c \leftarrow c + \delta c$ and the source wavelet $f \leftarrow f + \delta f$ is to perturb the plane-wave field U by a field δU satisfying

$$\left(\frac{1}{c^2} - p^2\right)\frac{\partial^2 \delta U}{\partial \tau^2} - \frac{\partial^2 \delta U}{\partial z^2} = \frac{2\delta c}{c^3}\frac{\partial^2 U}{\partial \tau^2}$$

$$-\frac{\partial \delta U}{\partial z}(0, \tau, p) = \delta f(\tau) \tag{2.2}$$

$$\delta U \equiv 0, \qquad \tau < 0.$$

The perturbational seismogram δS is the surface value of

δU:

$$\delta S(\tau, p) = \delta U(0, \tau, p).$$

Presumably $S[c + \delta c, \ f + \delta f] \approx S[c, f] + \delta S$. Conditions under which this is true, i.e. δS is actually the derivative of S. are discussed in Symes (1986a).

To understand the structure of the perturbational seismogram δS, we now introduce asymptotic approximations for the various fields. There will result an expansion of δS, which is of crucial importance to what follows.

As mentioned in the introduction, we consider in this paper only deconvolvable source wavelet f: that is, we assume that the convolution operator with kernel f has a bounded inverse on $L^2[0, T]$ for any $T > 0$. It is sufficient (but not necessary) that the Fourier transform $\hat{f}(\omega)$ have a uniformly bounded reciprocal. (Of course, we make no such restrictions on δf.) Since (with the obvious notation)

$$S_f = f^* S_\delta,$$

we could in principle apply the convolution inverse of f to the seismogram, and to the perturbational seismogram. This amounts to replacing δf by $(f^*)^{-1}\delta f$, and f by δ, in all of the above equations.

The field U is now the plane-wave *impulse-response*. As is well-known, this field possesses a *progressing wave expansion*: see Courant & Hilbert (1962, chapter 6) for generalities and Symes (1981, section 2) for computation of the present example. We obtain

$$U(z, \tau, p) = v(0, p)^{1/2} v(z, p)^{1/2} H(\tau - T(z, p)) + R(z, \tau, p), \tag{2.3}$$

where

$$v(z, p) = \frac{c(z)}{\sqrt{1 - c^2(z)p^2}}$$

is the plane-wave vertical wave velocity,

$$T(z, p) = \int_0^z \frac{dz'}{v(z', p)}$$

is the vertical travel time for the plane-wave at slowness p, H is the Heaviside unit step function, and R is a continuous remainder term. This expansion and others like it depend for their validity on the existence of sufficiently many derivatives of c, which we assume. See the concluding section for a little discussion of this point.

To derive the promised expansion of δS, first consider the case $\delta f \equiv 0$. Then a Green's identity argument, detailed in Santosa & Symes (1987, section 5) results in the expansion, expressed in terms of $\gamma := \delta \log c = \delta c / c$

$$\delta S|_{\delta f \equiv 0}(\tau, p) = v(0, p)\{(1 - c^2p^2)^{-1}\gamma\}(Z(\tau, p)) + \mathbb{K}_c\gamma(\tau, p), \tag{2.4}$$

where $c_0 = c(0)$ and $\mathbb{K}_c$ is an operator of Volterra type:

$$\mathbb{K}_c\gamma(\tau, p) = \int_0^{Z(\tau,p)} dz k_c(p, \tau, z)\gamma(z)$$

with continuous kernel k_c, and $Z(\tau, p)$ is the inverse of the two-way travel time:

$$z \equiv Z(2T(z, p), p).$$

On the other hand, a perturbation in $f(=\delta)$ gives the solution of the inhomogeneous problem

$$\left[\left(\frac{1}{c^2} - p^2\right)\frac{\partial^2}{\partial \tau^2} - \frac{\partial^2}{\partial z^2}\right]\delta U = 0$$

$$\left.\frac{\partial \delta U}{\partial z}\right|_{z=0} = -\delta f$$

$$\delta U \equiv 0, \qquad \tau < 0$$

which is in turn the convolution of δf with the solution of the same problem, but with δf replaced by (Dirac) δ. We recognize that the latter problem is identical to (2.1) with $f = -\delta$, so

$$\delta S|_{\delta c \equiv 0} = \delta f^* U|_{z=0}.$$

Using the progressing wave expansion we see that this is proportional to the integrated wavelet

$$\phi(\tau) = \int_0^\tau \delta f$$

up to an error given by a Volterra operator on ϕ. Introducing the velocity trace

$$V(\tau, p) = \frac{\partial U}{\partial \tau}(0, \tau, p)$$

we can re-write the above relation as

$$\delta S|_{\delta c \equiv 0} = V^* \phi \tag{2.5}$$

and $V(\tau, p) = v(0, p)\delta(\tau) + (\textit{continuous error})$ is an invertible (on L^2) τ-convolution kernel, known from the reference seismogram.

Combining (2.5) and (2.4) we get

$$\delta S(\tau, p) = V^*\phi(\tau, p) - v(0, p)\{(1 - c^2p^2)^{-1}\gamma\}(Z(\tau, p)) + \mathbb{K}\gamma(\tau, p). \tag{2.6}$$

Observe that V^* has a simple convolution inverse

$$V_1(\tau, p) = v(0, p)^{-1}\delta(\tau) + k_1(\tau, p),$$

where k_1 is continuous and causal for each p. We shall use this fact in the algorithm in Section 4.

3 SEPARATION OF WAVELET AND VELOCITY PERTURBATIONS

It is evident from (2.6) that the p-dependence ('moveout' and scaling) of δf and δc, as they appear in δS, are different. In this section, we draw a consequence: that both δf and δc may be determined from δS.

We do not address the *quality* of this determination here, or even the precise conditions under which it holds. We want the reader to be aware of the limitations of such formal analysis: it gives no insight whatsoever into the effectiveness of any algorithms developed to exploit the theoretical observation. Such insight can only be gained by a rigorous (i.e. correct and complete) analysis, which we give in Section 5.

In keeping with the formal approach of this section, we will drop all 'Volterra' terms [i.e. $\mathbb{K}$ in (2.6)] as negligible. The sense in which this is actually true is explained in Section 5.

Thus (2.6) becomes

$$\delta S \approx v(0, p)(\phi(\tau) - [(1 - c^2p^2)^{-1}\gamma](Z(\tau, p)).$$

[Here we have used the remark at the end of Section 2, writing $V(\tau, p) \approx v(0, p)\delta(z)$.]

Thus, except for scale, ϕ appears without any p-dependence at all. Accordingly, form the auxiliary data set

$$\begin{aligned}\bar{G}(\tau, p_1, p_2) &= v(0, p_1)^{-1}\delta S(\tau, p_1) - v(0, p_2)^{-1}\delta S(\tau, p_2)\\ &= [(1 - c^2p_2^2)^{-1}\gamma](Z(\tau, p_2))\\ &\quad - [(1 - c^2p_1^2)^{-1}\gamma](Z(\tau, p_1)).\end{aligned}$$

Suppose that $p_1 > p_2$. Then $v(z, p_1) > v(z, p_2)$ for all z, so the p_1-wave reaches any given depth earlier than the p_2-wave. Accordingly,

$$\alpha(z, p_1, p_2) := Z(2T(z, p_1), p_2) < z. \tag{3.1}$$

The transformed auxiliary data set

$$\tilde{G}(z, p_1, p_2) = (1 - c^2(z)p_1^2)\bar{G}(2T(z, p_1), p_1, p_2)$$

is the left-hand side in the functional equation for γ:

$$\tilde{G}(z, p_1, p_2) = \gamma(z) + \beta(z, p_1, p_2)\gamma(\alpha(z, p_1, p_2)), \tag{3.2}$$

where

$$\beta(z, p_1, p_2) = (1 - c^2(z)p_1)(1 - c^2(\alpha(z, p_1, p_2))^2p_2^2)^{-1}.$$

Because of the 'delay' inequality (3.1), the equation (3.2) gives $\gamma(z)$, for any z, as a combination of the data $\tilde{G}$ and the value of γ at some *shallower* $z'(<z)$. Thus it would seem reasonable that γ would be entirely determined if we know $\gamma(z)$ for some very shallow near surface layer $0 \le z \le z_*$, for then we could work our way downward, using (3.2) recursively, to obtain γ at any depth. Thus some restriction on the behaviour of γ near $z = 0$ would allow us to determine γ.

While this is true, it is unnecessary to restrict γ near the surface, because of the properties of the coefficients α and β. This result is somewhat technical, and will be given in detail in Section 5.

In sum, the functional equation (3.2) determines $\gamma(z)$ (without any restrictions). Then $\gamma(z)$ may be substituted in (2.6), which may then be solved for $\phi(\tau)$ (for any p). Thus both $\gamma(z)$ and $\phi(\tau)$ are determined.

In the next section, we consider the algorithmic implications of this observation.

4 NUMERICAL SOLUTION OF THE INVERSE PROBLEM

In this section we show how the considerations of Section 3 yield an iterative algorithm for solution of the least-squares inverse problem in the (τ, p) domain, and discuss its implementation for a special case.

Thus we attempt to minimize, for given (τ, p) data G,

$$J(c, f) := \|S(c, f) - G\|_\mu^2 = \int d\mu(p) \int_0^{T_{\max}(p)} d\tau\, |S - G|^2 \tag{4.1}$$

in the notation of Section 3. In fact, we assume here that $T_{\max}(p) \equiv T$ and that the domain of integration consists of precritical (τ, p) pairs for all velocity profiles to be constructed during the iteration. This assumption *a priori* restricts the velocity profiles and allows us to use only a minimal set of 'safely precritical' data. Obviously relaxation of these restrictions would be desirable. For some discussion see Santosa & Symes (1987, sections 2 and 7).

From (2.6) we can write

$$DS[c, \delta](\delta c, \delta f) = \mathbb{E}(\phi, \gamma) + \mathbb{K}(\phi, \gamma), \tag{4.2}$$

where

$$DS[c, \delta](\delta f, \delta c) = \delta S$$

is the directional derivative of the seismogram S,

$$\mathbb{E}(\phi, \gamma)(\tau, p) = v(0, p)\{\phi(\tau) + [(1 - c^2p^2)^{-1}\gamma](Z(p, \tau))\}$$

and $\mathbb{K}$ is a Volterra-type integral operator with continuous kernel.

We shall suppose that the wavelet is a slowly-varying perturbation of δ:

$$f(\tau) = \delta(z) + f_1(z),$$

where $f_1(z)$ is square-integrable, as mentioned in the introduction. Thus the integrated wavelet ϕ is continuous for $\tau \ge 0$, and satisfies

$$\lim_{\tau \to 0^+} \phi(\tau) = 1.$$

Then, since

$$DS[c, f](\delta c, \delta f) = f^* DS[c, \delta](\delta c, (f^*)^{-1}\delta f)$$

in fact $DS[c, f](\delta c, \delta f)$ is given by the same expression (4.2), where $\mathbb{K}$ is modified to include convolution by the square-integrable part of f.

The Gauss–Newton method for minimizing $J(c, f)$ updates a current estimate (c_c, f_c) by solving a linear least-squares problem:

$$c_+ = c_c + \delta c, \qquad f_+ = f_c + \delta f,$$

where $(\delta c, \delta f)$ minimizes

$$\|DS[c_c, f_c](\delta c, \delta f) - (G - S[c_c, f_c])\|_\mu^2. \tag{4.3}$$

The minimizer of (4.3) satisfies the *normal equations*

$$\begin{aligned}&DS[c_c, f_c]^* DS[c_c, f_c](\delta c, \delta f)\\ &\quad = DS[c_c, f_c]^*(G - S[c_c, f])\end{aligned} \tag{4.4}$$

where DS^* is the *adjoint* of DS [regarded as a linear map acting on (ϕ, γ) from $L^2[0, T] \times L^2[0, Z] \to L^2\{[0, T] \times \operatorname{supp} d\mu, d\tau\, d\mu(p)\}$].

Since $\mathbb{K}$ in (4.2) is Volterra, it represents a small perturbation if Z, T are small enough. This suggests replacing DS and DS^* in (4.4) with the simpler operators $\mathbb{E}$ and $\mathbb{E}^*$, with the latter given by

$$\mathbb{E}^*\Phi = (\mathbb{E}_\phi^*\Phi, \mathbb{E}_\gamma^*\Phi)$$

$$\mathbb{E}_\phi^*\Phi(\tau) = \int d\mu(p) v(0, p)\Phi(\tau, p)$$

$$\mathbb{E}_\gamma^*\Phi(z) = 2\int d\mu(p) \frac{v(0, p)\Phi(2T(z, p), p)}{v(z, p)(1 - c^2(z)p^2)}.$$

It follows from the heuristics of Section 3 and the estimates in Section 5 that $\mathbb{E}^*\mathbb{E}$ is invertible—in fact positive definite—and $\mathbb{E}$ and $\mathbb{E}^*$ are easy to evaluate. Thus the modified normal equations

$$\mathbb{E}^*\mathbb{E}(\phi, \gamma) = \mathbb{E}^*(G - S[c_c, f_e]) \tag{4.5}$$

are uniquely solvable, and an iterative method like conjugate gradients should be quite efficient.

In rough outline, the resulting algorithm is: repeat until convergence

(1) compute the seismogram $S[c, f]$;
(2) solve the modified normal equations (4.5) for ϕ, γ:
(3) update $c \leftarrow ce^{\gamma}$, $f \leftarrow f + \phi'$; go to (1)

Essentially the same approach (replace DS by $\mathbb{E}$) is used with success by Sacks & Santosa (1987) in recovering c alone (they consider the 'consistent' case and solve a functional equation but the idea is very similar).

Since the replacement $DS \leftarrow \mathbb{E}$ involves only a small change for τ and z small, we would expect this algorithm to behave like the Gauss–Newton iteration near $\tau = z = 0$. The estimates of Section 5, which show that DS is bounded below, would suffice to show that the Gauss–Newton algorithm is convergent if S were differentiable. Unfortunately, differentiability of S requires more constraints on c than we have included here—essentially, c must be relatively smooth. The reasons for this, and a convergence-inducing regularization of the least-squares problem, are discussed in Symes (1986c). Granted that this regularization has been implemented, either by restricting the class of admissible c's or by adding a penalty term to J, Gauss–Newton will converge locally and thus so will the modified algorithm given above.

We will now consider a special case of this procedure. More general examples will be explored elsewhere.

We take (for $p_1 > p_2 \geq 0$)

$$d\mu(p) = \delta(p - p_1) + \delta(p - p_2)$$

which amounts to choosing for our data exactly two plane-wave traces. Now we should have 'just enough' data to solve the problem: two time series to determine two time series. In fact, a minor modification of the discussion of Section 3 leads to the conclusion that, for appropriate z and τ intervals, the mapping

$$\begin{pmatrix} \phi \\ \gamma \end{pmatrix} \mapsto \begin{pmatrix} \mathbb{E}(\phi, \gamma)(p_1, \cdot) \\ \mathbb{E}(\phi, \gamma)(p_2, \cdot) \end{pmatrix}$$

is invertible (as is the analogous map with $\mathbb{E}$ replaced by $DS[c, f]$). It follows that $\mathbb{E}^*$ is invertible as well, so the modified normal equations (4.5) are (in this special case) equivalent to

$$\mathbb{E}(\phi, \gamma) = G - S[c, f] \tag{4.6}$$

which is an approximation to

$$DS[c, f](\delta c, \delta f) = G - S[c, f]. \tag{4.7a}$$

Now (4.7) defines the Newton update for the functional equation

$$S[c, f] = G \tag{4.7b}$$

so we would expect iterative use of (4.6) to yield an approximate solution of this functional equation, rather than merely a least-squares solution. This is precisely the analogue of the algorithm presented in Sacks & Santosa (1987).

The prescription for solving (4.6) is actually given in Sections 3 and 5; in outline, repeat until convergence the following steps:

(i) Create the auxiliary data set

$$\bar{G}(\tau) = V_1^*(G - S[c, f])(\tau, p_1) - V_1^*(G - S[c, f])(\tau, p_2), \tag{4.8a}$$

where V_1 is the convolution inverse of the (current) velocity trace;

(ii) Create the auxiliary data set

$$\tilde{G}(z) = (1 - c(z)^2 p_1^2)\bar{G}(2T(z, p_1)).$$

(ii) Solve for γ the functional equation

$$\tilde{G}(z) = \gamma(z) + \beta(z)\gamma(\alpha(z)) =: (\mathbb{I} + \mathbb{T})\gamma(z), \tag{4.8b}$$

where

$$\alpha(z) = Z(2T(z, p_1), p_2) \tag{4.9}$$

$$\beta(z) = \frac{1 - c(z)^2 p_1^2}{1 - c^2(\alpha(z))p_2^2}$$

using the Neumann series

$$(\mathbb{I} + \mathbb{T})^{-1} \simeq \sum_{n=0}^{\infty} (-\mathbb{T})^n \tag{4.10}$$

the convergence of which is guaranteed by the properties of α and β and by Lemma 1 of Section 5.

(iv) Compute ϕ from the relation

$$V_1^*\{(G - S[c, f])(\tau, p_1) - v(0, p)[(1 - c^2 p_1^2)^{-1}\gamma](Z(\tau, p_1))\} = \phi(\tau).$$

(v) Update c and f:

$$c \leftarrow ce^{\gamma}$$

$$f \leftarrow f + \phi'.$$

Remark. These steps may be simplified further by noticing that $V_1(\tau, p) = v(0, p)^{-1}\delta(\tau) + k_1(\tau, p)$, as noted earlier, so that the convolution V_1^* may be replaced by multiplication with $v(0, p)^{-1}$ at the cost of another ('Volterra') error of the type we are already ignoring. This was also done in Section 3.

We have implemented a single step of this procedure. We compute the plane-wave seismogram $S[c, f]$ using a standard leapfrog finite-difference method, applied to the first-order system equivalent to the wave equation, in which the velocity is one of the dependent variables. These and other functions were represented by grid-functions on suitable space-time grids.

To obtain the unknown $\delta c(z)$ we must compute the function $\tilde{G}(z)$, defined in (4.8a, b), then evaluate the series (4.10). The compositions with travel-time, and with the depth change-of-coordinates (4.9), are computed using piecewise linear interpolation, the accuracy of which is compatible with that of the difference scheme. We replaced convolution with V by its discrete version:

$$(V^*\phi)(\tau_j) \sim \sum_i \Delta\tau V_{j-i}\phi_i$$

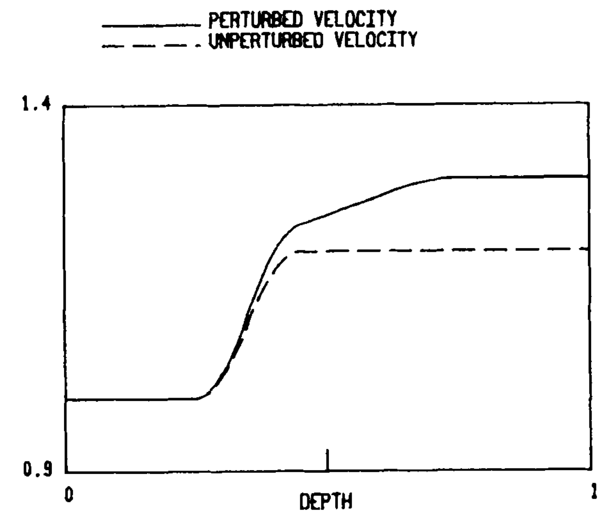

Figure 1. Reference and perturbed velocity profiles for first set of experiments (Figs 1–5).

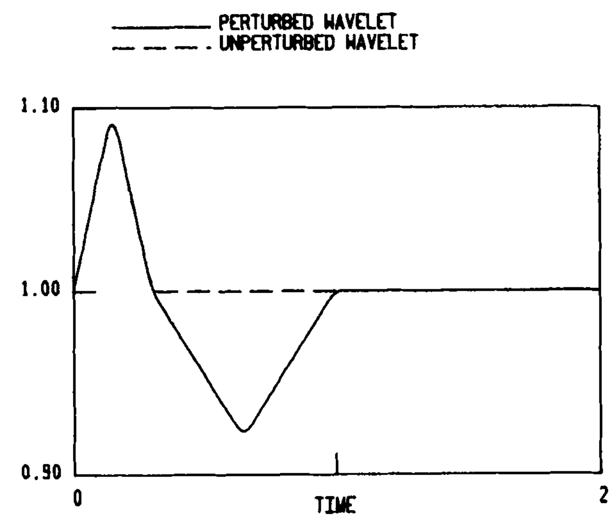

Figure 2. Reference and perturbed wavelets for first set of experiments.

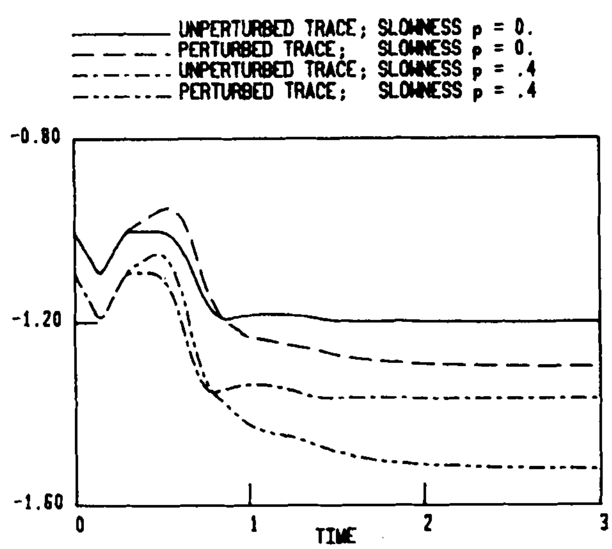

Figure 3. Reference and perturbed traces for $p = 0.0$, 0.4.

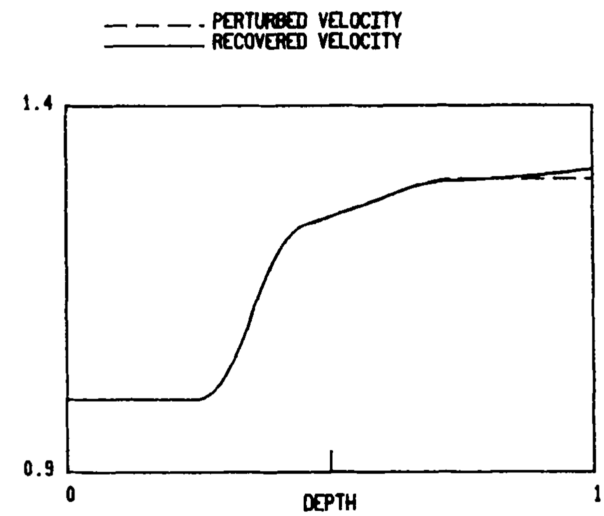

Figure 4. Comparison of exact and recovered velocity profiles from data of Fig. 3.

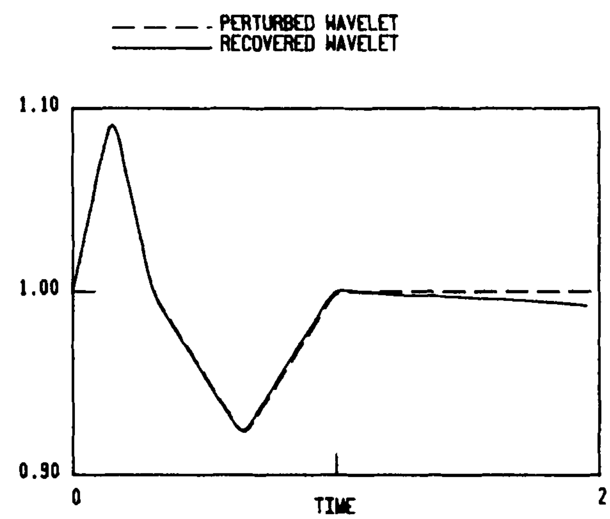

Figure 5. Comparison of exact and recovered wavelets from data of Fig. 3.

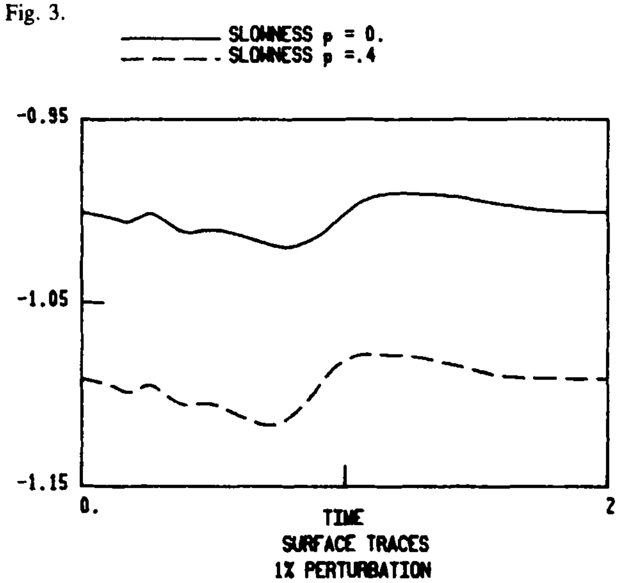

Figure 6. Perturbed traces, $p = 0.0$, 0.4; 1 per cent rms perturbation in $c(z)$, $f(t)$. Reference traces same as in Fig. 3.

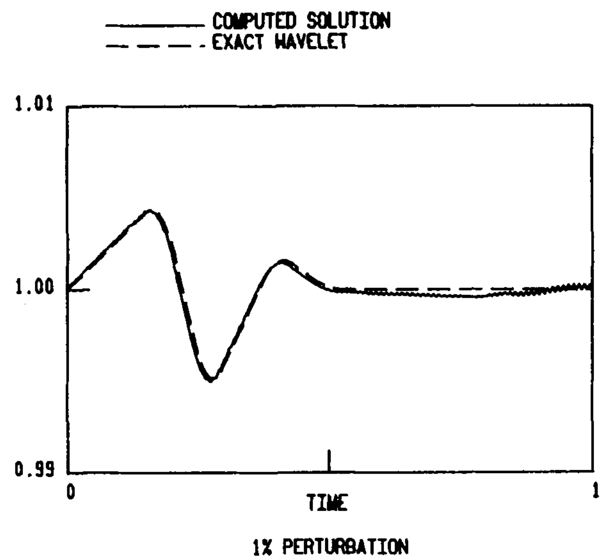

Figure 7. Comparison of exact and recovered wavelets from data of Fig. 6.

and deconvolution by V (i.e. convolution with V_1) with the solution of this triangular system by back-substitution.

A synthetic example of this procedure is presented in Figs 1–5. Figure 1 shows the reference velocity c and the velocity to be recovered $c + \delta c$, and Fig. 2 shows the corresponding reference and perturbed wavelets. For the two slownesses we chose $p_2 = 0$ and $p_1 = 0.4$. In Fig. 3 we show $S[c, f](\tau, p_i)$, $i = 1, 2$ and $G(\tau, p_i) = S[c + \delta c, f + \delta f](\tau, p_i)$, $i = 1, 2$.

The velocity perturbation δc is computed as discussed above; it was necessary to use three terms in the series (4.10). The resulting estimate of $c + \delta c$ is displayed in Fig. 4, along with the exact velocity profile. Finally, the exact and recovered wavelets are shown in Fig. 5.

We emphasize that this example illustrates our contention that the source wavelet (perturbation) may be recovered

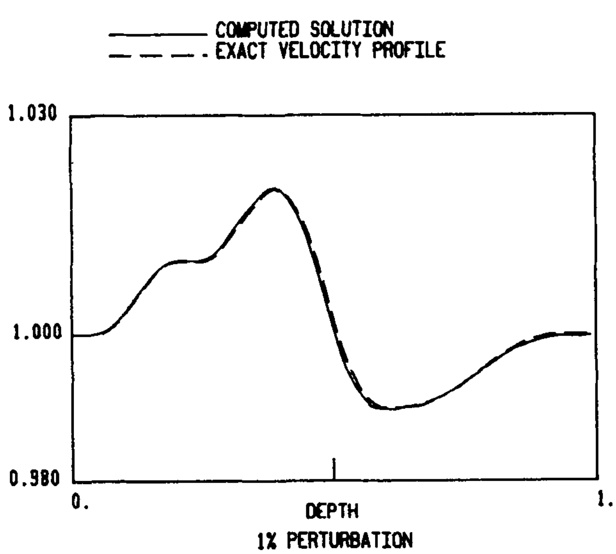

Figure 8. Comparison of exact and recovered velocity profiles from data of Fig. 6.

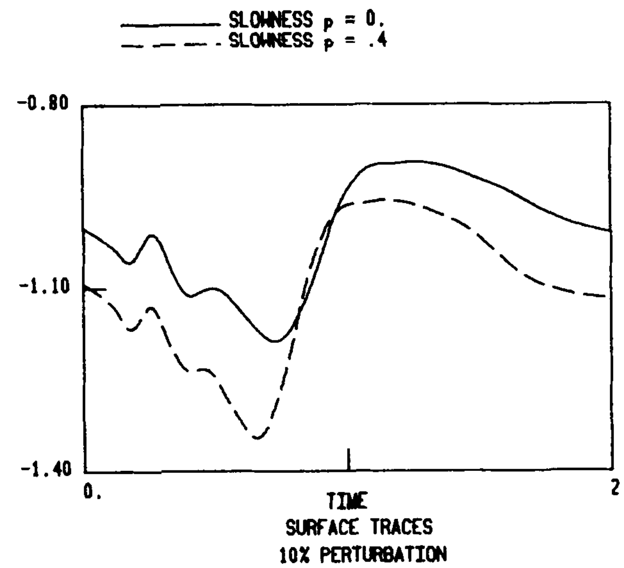

Figure 9. Perturbed traces, $p = 0.0$, 0.4; 10 per cent rms perturbation in $c(z)$, $f(t)$. Reference traces same as in Fig. 3.

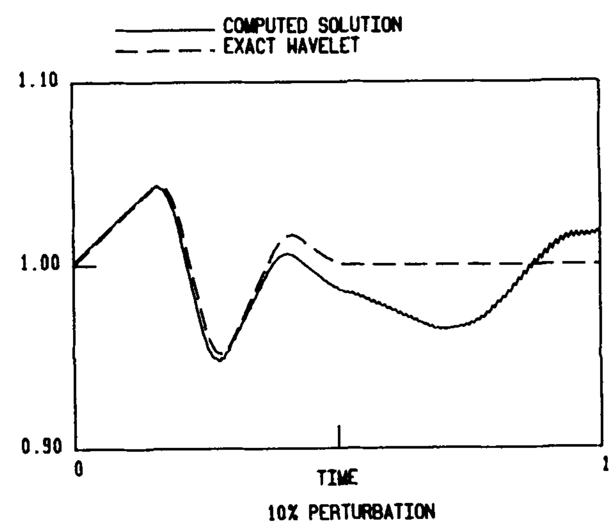

Figure 10. Comparison of exact and recovered wavelets from data of Fig. 9.

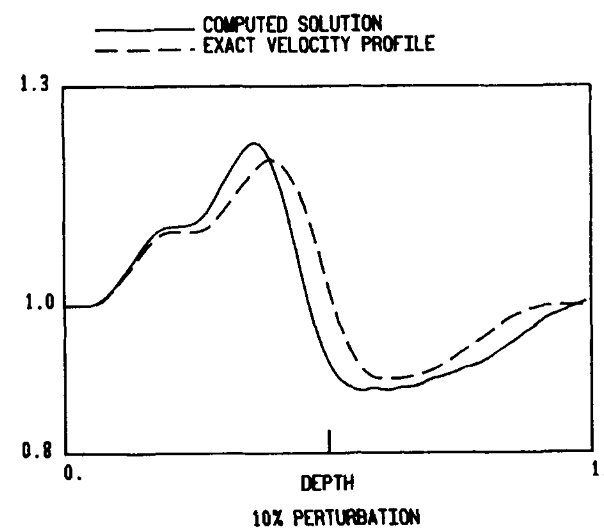

Figure 11. Comparison of exact and recovered velocity profiles from data of Fig. 9.

even without the presence of a strong clean reflection in an otherwise quiet part of the section. In this example, the perturbations δc and δf yield trace perturbations δs which are completely time-coincident.

Note that in the interesting special case that $c(z) \equiv 1$, $f(t) = \delta(t)$, the above procedure simplifies substantially, since $S[c, f]$ can then be computed exactly.

Since this procedure approximates one step of Newton's method for the functional equation (4.7b), we would expect the accuracy to degrade as δc becomes larger. To illustrate this degradation, we show in Fig. 6 the perturbed surface traces $S[c + \delta c, f + \delta f]$ for $p_2 = 0$, $p_1 = 0.4$ and for δc and δf representing 1 per cent of the 'energy' (L^2-norm) of c and f. The recovered wavelets and velocity are compared with the targets in Figs 7 and 8. Figures 9, 10, and 11 repeat this comparison for 10 per cent perturbations δc and δf.

5 ESTIMATION OF THE WAVELET AND VELOCITY PERTURBATIONS

Our aim in this section is to show that the differential section δS dominates the perturbations $\gamma = \delta \log c$ and $\phi = \int \delta f$:

$$\|\phi\|^2 + \|\gamma\|^2 \le C \, \|\delta S\|^2,$$

where the vertical bars denote appropriate L^2-norms. Such estimates are necessary to ensure stable linear estimation of δc and δf, and also form the heart of the construction of nonlinear least-squares estimators for c and f.

Our derivation will be constructive, and will justify a computational technique which we explored in the last section, as well as the heuristics of Section 3.

Recall that

$$V_1(\tau, p) = \frac{1}{v(0, p)} \delta(\tau) + k_1(\tau, p)$$

is the convolutional inverse of $V(\tau, p)$, with k_1 continuous and causal. From the structure of V_1, and the expansion for δS in (2.6), the auxiliary data (Section 3) must have the form

$$\begin{aligned}\bar{G}(\tau, p_1, p_2) = {} & (1 - c^2 p^1)^{-1} \gamma(Z(\tau, p_1)) \\ & - (1 - c^2 p_2)^{-1} \gamma(Z(\tau, p_2)) + \mathbb{K}_1 \gamma(\tau, p_1, p_2),\end{aligned} \tag{5.1}$$

where $\mathbb{K}_1$ is an integral operator with piecewise continuous kernel $k_1(p_1, p_2, \tau, z)$ supported in $\{0 \le z \le \max[Z(\tau, p_1), Z(\tau, p_2)]\}$. This last term was dropped in Section 3.

We have already indicated that when $p_1 > p_2$,

$$\alpha(z, p_1, p_2) := Z(2T(z, p_1), p_2) < z.$$

We will now quantify this inequality.

The quantity $\alpha(z, p_1, p_2)$ is crucial to the argument which follows: it gives the depth reached by a wave travelling at the slower wave speed $v(\cdot, p_2)$ in the two-way time taken by a wave travelling as the faster wave speed $v(\cdot, p_1)$ to reach depth z, thus it represents the 'spatial delay' caused by replacing the (faster) $v(\cdot, p_1)$ by the (slower) $v(\cdot, p_2)$. We shall assume that we consider only pairs (τ, p) for which $z = Z(\tau, p)$ satisfies

$$\frac{1}{c^2(z')} - p^2 \ge \Delta^2 \quad \text{for} \quad 0 \le z' \le z \tag{5.2}$$

for some fixed $\Delta > 0$. Denote by $\bar{z}_{\max}(p)$ the largest z for which (5.2) holds.

To take into account the possibly finite duration of the p-τ traces $[0 \le \tau \le T_{\max}(p)]$, set

$$z_{\max}(p) = \min(\bar{z}_{\max}(p), Z(T_{\max}(p), p)).$$

[Note that $\bar{z}_{\max}(p) = \infty$, for instance, for $p = 0$, if c is bounded on $0 \le z < \infty$ and $T_{\max}(0) = \infty$.] Correspondingly, set

$$\tau_{\max}(p) = 2T(z_{\max}(p), p).$$

Then for $0 \le z < z_{\max}(p)$ a little algebra yields

$$T(z, p_2) \ge \left(1 + \frac{c_*^2}{2}(p_1^2 - p_2^2)\right) T(z, p_1),$$

whence

$$\begin{aligned} z - Z(2T(z, p_1), p_2) &\ge c_*[T(z, p_2) - T(z, p_1)] \\ &\ge \frac{c_*^3}{2}(p_1^2 - p_2^2) T(z, p_1) \\ &\ge \frac{c_*^3 \Delta}{2}(p_1^2 - p_2^2) z. \end{aligned}$$

Thus

$$\alpha(z, p_1, p_2) \le \alpha^*(p_1, p_2)z \tag{5.3}$$

with $\alpha^* = 1 - \frac{1}{4}c_*^3 \Delta(p_1^2 - p_2^2) < 1$. Similarly , it is easy to see that $\alpha' \ge \alpha_* > 0$ for some α_* independent of z, p.

Recall the definition of $\bar{G}(z, p_1, p_2)$ and its representation (3.2):

$$\begin{aligned}\bar{G}(z, p_1, p_2) &:= (1 - c^2(z)p_1^2)\bar{G}(2T(z, p_1), p_1, p_2).\\ &= \gamma(z) + \beta(z, p_1, p_2)\gamma(\alpha(z, p_1, p_2)) + \mathbb{K}_2\gamma(z, p_1, p_2),\end{aligned} \tag{5.4}$$

where $\mathbb{K}_2$ is another integral operator with piecewise continuous kernel $k_2(p_1, p_2, z, z')$ supported in $\{0 \le z' \le z\}$, and

$$\beta(z, p_1, p_2) = (1 - c^2(z)p_1^2)(1 - c^2(\alpha(z, p_1, p_2))p_2^2)^{-1}.$$

[Thus (5.4) is the precise statement of which (3.2) is an approximation.] Note that

$$\beta(0, p_1, p_2) = \frac{1 - c^2(0)p_1^2}{1 - c^2(0)p_2^2} < 1. \tag{5.5}$$

Now it follows from (5.3), (5.4), (5.5) and Lemma 3 below that

$$\begin{aligned}\|\gamma\|_{L^2[0, z_{\max}(p_1)]} &\leqq C_1 \|\bar{G}(\cdot, p_1, p_2)\|_{L^2[0, z_{\max}(p_1)]}\\ &\le C_2 \|\bar{G}(\cdot, p_1, p_2)\|_{L^2[0, \tau_{\max}(p_1)]}\\ &\le C_3\{\|\delta S(\cdot, p_1)\|_{L^2[0, \tau_{\max}(p_1)]}\\ &\quad + \|\delta S(\cdot, p_2)\|_{L^2[0, \tau_{\max}(p_2)]}\}.\end{aligned} \tag{5.6}$$

From the foregoing calculations and the statement of Lemma 3, the constants C_1, C_2, C_3 evidently depend on c_*, Δ, and $p_1^2 - p_2^2 > 0$.

To express the dependence of δc and δf on the entire precritical p-tau section, choose a positive measure $d\mu$ and set

$$\|\delta S\|_\mu^2 = \int_0^{p_{\max}} d\mu(p) \int_0^{\tau_{\max}(p)} d\tau\, |\delta S(\tau, p)|^2.$$

Thus $\| \ \|_\mu^2$ is a weighted mean-square error measure. For example, if we choose $d\mu(p) = dp$ (Lebesque measure), the all p-tau traces are weighted uniformly, whereas choosing $d\mu(p)$ to be a linear combination of point masses has the effect of selecting a discrete set of slownesses.

Now suppose that for some $\varepsilon > 0$,

$$\{(p_1, p_2) : |p_1 - p_2| \ge \varepsilon\} \cap (\operatorname{supp} d\mu \times \operatorname{supp} d\mu) \ne 0.$$

Let $p_{\min} = \inf\{|p| : p, p + \varepsilon \in \operatorname{supp} \mu\}$. Then from (5.6) follows immediately the estimate

$$\begin{aligned}&\|\gamma\|_{L^2[0, z_{\max}(p_{\min})]}^2\\ &\le C_4 \int_{|p_1 - p_2| > \varepsilon} d\mu(p_1) \int d\mu_2(p_2) \int_0^{\max[\tau_{\max}(p_1), \tau_{\max}(p_2)]} d\tau\, |G(\tau, p_1, p_2)|^2\\ &\le C_5 \|\delta S\|_\mu^2,\end{aligned} \tag{5.7}$$

where C_4 and C_5 depend on G and μ as well.

Perhaps a refined analysis of these constants is possible, in the spirit of Santosa & Symes (1988).

Finally, with γ known, ϕ can be extracted from any of the components via (2.6), whence estimates for ϕ follow immediately.

To complete this section we give estimates needed above, which establish the invertibility on L^2 of operators of the form

$$\mathbb{I} + \mathbb{T} + \mathbb{K},$$

where $\mathbb{T}$ is multiplicative-delay:

$$\mathbb{T}u(z) = \beta(z)u(\alpha(z)),$$

with $\alpha(z) \le \alpha^* < 1$, $0 \le z \le 1$ and $\mathbb{K}$ is an integral operator of Volterra type

$$\mathbb{K}u(z) = \int_0^z d\zeta k(z, \zeta)u(\zeta).$$

For convenience we restrict our attention to the scalar case. Analogous results for vector-valued functions follow from trivial modifications of the proofs. We begin with the case $k \equiv 0$.

Lemma 1. Suppose that α, $\beta \in C^1[0, 1]$ satisfy

(1) $\alpha(0) = 0$,

(2) for some α_*, α^* with $0<\alpha_*$, α^* and $\alpha^*<1$,

$$\left.\begin{aligned}\alpha'(z)\geq\alpha_*\\ \alpha(z)\leq\alpha^* z\end{aligned}\right\}\quad 0\leq z\leq 1.$$

(3) $\beta(0)[\alpha'(0)]^{-1/p}<1$.

Then for any $\mu<1$, the operator $\mathbb{T}: L^p[0,1]\to L^p[0,1]$ defined by

$$\mathbb{T}u(z)=\beta(z)u(\alpha(z))$$

satisfies

$$\|\mathbb{T}^N\|_p\leq\mu$$

for some N depending on $\|\alpha\|_{C^1[0,1]}$, $\|\beta\|_{C^1[0,1]}$, α_*, α^*, $\beta(0)(\alpha'(0))^{-1/p}$, and μ.

Proof. Define sequences α_n, β_n in $C^1[0,1]$ by

$$\left.\begin{aligned}\alpha_0(z)&=z\\ \alpha_1(z)&=\alpha(z)\\ \alpha_n(z)&=\alpha(\alpha_{n-1}(z))\end{aligned}\right\}\quad\text{for}\quad 0\leq z\leq 1,\qquad n>1$$

$$\alpha_{-1}(z)=\alpha^{-1}(z)\quad\text{for}\quad 0\leq z\leq\alpha(1)$$

$$\alpha_{-n}(z)=\alpha_{-n+1}(\alpha_{-1}(z))\quad\text{for}\quad 0\leq z\leq\alpha_n(1),\qquad n>1$$

$$\left.\begin{aligned}\beta_1(z)&=\beta(z)\\ \beta_n(z)&=\beta(z)\beta_{n-1}(\alpha(z))\\ &=\prod_{k=0}^{n-1}\beta(\alpha_k(z))\end{aligned}\right\}\quad\text{for}\quad 0\leq z\leq 1,\qquad n>1.$$

Then

$$\mathbb{T}^n u(z)=\beta_n(z)u(\alpha_n(z)).$$

Now that

$$\alpha_n(z)\leq\alpha^*\alpha_{n-1}(z)\leq\cdots\leq(\alpha^*)^n z,\qquad 0\leq z\leq 1$$

so that $a_n(1)\leq(\alpha^*)^n\to 0$ as $n\to\infty$. Choose δ so that for $0\leq z\leq\delta$,

$$\beta(z)(\alpha'(z))^{-1/p}\leq\tfrac{1}{2}[1+\beta(0)(\alpha'(0))^{-1/p}]:=\lambda<1$$

and select n so that $(\alpha^*)^n\leq\delta$.

Note that

$$\begin{aligned}\alpha_n'&=\alpha'\circ\alpha_{n-1}\cdot\alpha_{n-1}'\\ &=\cdots=\prod_{k=0}^{n-1}\alpha'\circ\alpha_k.\end{aligned}$$

Note also that if $z\in[0,\alpha_k(1)]$, then

$$\alpha_j\circ\alpha_{-k}(z)=\alpha_{j-k}(z),\qquad j=0,1,2,\ldots$$

Thus

$$\begin{aligned}\|T^N u\|_p^p&=\int_0^1 dz\,|\beta_N u\circ\alpha_N|^p\\ &=\int_0^{\alpha_N(1)}\frac{|\beta_N\circ\alpha_{-N}|^p}{\alpha_N'\circ\alpha_{-N}}|u|^p\\ &=\int_0^{\alpha_N(1)}\left[\prod_{k=0}^{N-1}\frac{|\beta\circ\alpha_{k-N}|^p}{\alpha'\circ\alpha_{k-N}}\right]|u|^p.\end{aligned}\tag{5.8}$$

Set

$$C:=\|\beta(\alpha')^{1/p}\|_\infty^n.$$

Then

$$\prod_{k=0}^{n-1}\frac{|\beta\circ\alpha_{k-N}|^p}{\alpha'\circ\alpha_{k-N}}\leq C^p\tag{5.9}$$

on $[0, \alpha_N(1)]$. Also note that, if $z \in [0, \alpha_N(1)]$ and $k \geq n$, $\alpha_{k-N}(z) = \alpha_{k-N}[\alpha_N(w))] = \alpha_k(w) \leq \delta$ for some $w \in [0, 1]$. Thus

$$\frac{|\beta \circ \alpha_{k-N}|^p}{\alpha' \circ \alpha_{k-N}}(z) \leq \lambda^p \quad \text{for} \quad k \geq n, \qquad 0 \leq z <= \alpha_N(1). \tag{5.10}$$

Taken together, (5.8), (5.9), and (5.10) imply that

$$\|\mathbb{T}^N\|_p^p \leq C^p \lambda^{(N-n)p} \|u\|_p^p$$

for $N \geq n$. If

$$N \geq n + \frac{\log C - \log \mu}{(-\log \lambda)}$$

then

$$\|\mathbb{T}^N u\|_p \leq \mu \|u\|_p$$

as required. Clearly C, λ and n depend on the stated quantities.

We now consider operators of 'Volterra' form, that is,

$$\mathbb{K}u(z) = \int_0^z d\zeta k(z, \zeta) u(\zeta).$$

As is well known, for continuous kernels k, such operators have spectral radius zero as operators on L^p. See e.g. Widom (1969, pp. 9–11).

Actually, slightly less than continuity is required. We consider operators K defined by kernels k, possibly matrix-valued, for which the function

$$\tilde{k}(z) = k(\cdot, z)$$

lies in $C^0([0, 1], L^2[0, 1])$ with the property that $\operatorname{supp} k(z) \subset [z, 1]$, $0 \leq z < 1$.

Note that the Volterra property is equivalent to the statement that

$$\operatorname{supp} u \subset [z_1, 1] \Rightarrow \operatorname{supp} \mathbb{K}u \subset [z_1, 1].$$

Lemma 2. Suppose that $k \in C^0([0, 1]; L^2[0, 1])$ satisfies $\operatorname{supp} \tilde{k}(z) \subset [z, 1]$ and define for $u \in L^2[0, 1]$

$$\mathbb{K}u = \int_0^1 dz \tilde{k}(z) u(z) \in L^2[0, 1].$$

Then for any positive $\mu < 1$ there exists a positive integer N depending on

$$\operatorname{supp}_{0 \leq z \leq 1} \|\tilde{k}(z)\|_{L^2[0,1]}$$

and on μ, for which

$$\|\mathbb{K}^N\|_2 < 1.$$

Proof. [This is a small modification of the standard argument. See e.g. Widom (1969).] Define the iterated kernels $\tilde{k}_n$ by

$$\tilde{k}_1 = \tilde{k}$$

$$\tilde{k}_n(z_1) = \int dz \tilde{k}_{n-1}(z) k(z_1, z) \qquad n = 2, 3, 4 \cdots$$

One easily verifies that $\tilde{k}_n \in C^0([0, 1]; L^2[0, 1])$. Because of the observation preceding the statement of the lemma, $\operatorname{supp} \tilde{k}_n(z) \subset [z, 1]$ as well.

Note that

$$\begin{aligned}
\|\tilde{k}_2(z_1)\| &= \int dz \tilde{k}_1(z) k(z, z_1)\| \\
&= \left\| \int_{z_1}^1 dz \tilde{k}(z) k(z, z_1) \right\| \\
&\leq \int_{z_1}^1 dz \, \|\tilde{k}(z)| \, |k(z, z_1)| \\
&\leq C \int_{z_1}^1 dz \, |k(z, z_1)| \\
&\leq C^2 (1 - z_1)^{1/2},
\end{aligned}$$

where we have written $C = \sup_{0 \le z \le 1} \|\tilde{k}(z)\|$. Assume that

$$\|\tilde{k}_j(z)\| \le \frac{C^n(1-z)^{n/2}}{(n!)^{1/2}}, \qquad j = 1, 2, \ldots, n.$$

In general,

$$\begin{aligned}\|\tilde{k}_{n+1}(z_1)\| &\le \int_{z_1}^{1} dz\, \|\tilde{k}_n(z)\|\, |k(z, z_1)| \\ &\le \frac{C^{n+1}}{(n!)^{1/2}} \int_{z_1}^{1} dz(1-z)^{n/2}\, |k(z, z_1)| \\ &\le \frac{C^{n+1}}{(n!)^{1/2}} \left(\int_{z_1}^{1} dz(1-z)^n\right)^{1/2} \|\tilde{k}(z_1)\| \\ &= \frac{C^{n+2}}{((n+1)!)^{1/2}}(1-z_1)^{n+1/2}.\end{aligned}$$

Thus $\|\mathbb{K}^n\| < 1$ for large n, as desired. Q.E.D.

Lemma 3. Suppose that k is a Volterra kernel as described in Lemma 2, and $\mathbb{T}$ is a multiplicative delay operator defined by scalars α, β obeying the conditions of Lemma 1. Then

$$\mathbb{I} + \mathbb{T} + \mathbb{K}$$

is invertible: in particular, for some $\varepsilon > 0$ depending on the quantities described in Lemmas 1 and 2,

$$\|(\mathbb{I} + \mathbb{T} + \mathbb{K})u\| \ge \varepsilon\, \|u\|$$

for $u \in L^2[0, 1]$.

Proof. The convergence of the Neumann series of $\mathbb{I} + \mathbb{T}$ in operator norm is a consequence of Lemma 1. Thus $(\mathbb{I} + \mathbb{T})^{-1}\mathbb{K}$ is a bounded operator on $L^2[0, 1]$, given by a kernel

$$\tilde{k}_1(z) = (\mathbb{I} + \mathbb{T})^{-1}\tilde{k}_0(z).$$

The Volterra property of $\tilde{k}_1$ follows from the delay property ($\alpha < 1$) of $\mathbb{T}$, and an estimate for k_1 in $C^0([0, 1]; L^2[0, 1])$ follows drectly from Lemma 1 as well. Thus

$$\begin{aligned}\|(\mathbb{I} + \mathbb{T} + \mathbb{K})u\| &\geqq \|(\mathbb{I} + \mathbb{T})(\mathbb{I} + (\mathbb{I} + \mathbb{T})^{-1}\mathbb{K})u)\| \\ &\ge \varepsilon_1 \|(\mathbb{I} + (\mathbb{I} + \mathbb{T})^{-1}\mathbb{K})u\| \\ &\ge \varepsilon\, \|u\|\end{aligned}$$

by the application of Lemma 2. Q.E.D.

6 DISCUSSION

In this section we discuss briefly the consequences of weakening the hypotheses M1–M3, S1–S3 stated in Section 1, which underlay our arguments but which are unacceptably restrictive for practical application.

The extension of the perturbational results to nonlayered background media is straightforward, to some extent, so long as the background is smooth. In some sense the literature on migration is concerned exactly with this extension: for some recent, mathematically correct results for the fixed f case see Beylkin (1985) and Rakesh (1986). Results analogous to those presented here should be achievable in this context. The (full) nonlinear nonlayered problem is much more difficult; for some limited results see Sacks & Symes (1985), Symes (1986a).

The isotropic elasticity model is the general choice for the 'premium' model level; see e.g. Tarantola (1984). Some results on the determination of a layered elastic medium from various point-source data sets may be found in Clarke (1984), Yagle & Levy (1986), Carazzone (1986), and Sacks & Symes (1987). On the other hand, at least transverse, and possibly more general, anisotropy is evident in reflection data in some locales (Thomsen 1986); Coen & Meadows (1985) give some results concerning anisotropic elastic inversion. We see no difficulty in extending the results reported here to these more general layered models.

The smoothless question is quite interesting and is presently inadequately understood. Symes (1986c) gives a review of the extent to which smoothness constraints can be relaxed for a single-plane-wave, layered medium problem, and Sacks & Symes (1985) gives some idea of the difficulties which arise in the study of several-dimensional problems in the presence of limited smoothness. An analytic approach to the nonlinear problem requires that this issue be addressed, as perturbations and backgrounds must then be regarded as having the same degree of smoothness. On the other hand, non-smoothness may have very favourable consequences, as will be discussed below in relation to band-extrapolation, and must certainly be regarded as a feature of the actual parameter distributions in earth materials.

As mentioned in the introduction, the point-source assumption seems adequate for reflection seismology on physical grounds. On the other hand, common land and marine energy sources are strongly anisotropic. To some extent this feature might be modelled by a multipole source term, and the moment tensor components recovered by a generalization of our technique. This seems an important matter for further work.

Perhaps the most unrealistic assumption of our work is S3 (quasi-impulsive nature of sources): it is simply violently wrong. Typical reflection seismic sources have significant energy content in the range 4–60 Hz, at best. Thus a reasonable isotropic point-source model should involve a very non-impulsive wavelet.

In the last several years, considerable numerical evidence has emerged to support the contention that bandlimited point-source data does determine layered velocity profiles by least-squares data fitting: see for example McAulay (1985) and Kolb, Collino & Lailly (1986). Santosa & Symes (1986) give a comprehensive analysis of this problem, and show that success of the least-squares approach is crucially dependent on

(i) sufficient (precritical) aperture and,
(ii) sufficient reflector density, i.e. sufficient *lack of smoothness*.

Thus the nonsmooth distribution of parameters may have as a positive consequence the feasibility of seismic inversion.

We believe that the ideas of this paper and those of Santosa & Symes (1986) might be synthesized to obtain codetermination of velocity profiles and bandlimited point-sources, under appropriate conditions. In support of this conjecture we cite the numerical experiments of Canadas & Kolb (1986) who solved this problem numerically *via* least-squares data fitting.

ACKNOWLEDGMENTS

The initial work was carried out at Cornell University under the auspices of the SRO-III project on inverse problems sponsored by ONR. We are grateful to the directors of that project, Professors Y.-H. Pao and L. Payne, for their encouragement and support. We would also like to thank the editor (Dr J. Stevens) and the referees (Dr J. Carazzone and Dr T. Clarke) for suggestions which led to substantial improvements in the exposition.

REFERENCES

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

Canadas, G. & Kolb, P. 1986. Least-squares inversion of prestack data: simultaneous identification of density and velocity of stratified media, *Presented at 56th Ann. Mtng Soc. Exploration Geophysicists*, Houston, TX.

Carazzone, J. J., 1986. Inversion of P-SV seismic data, *Geophysics*, **51,** 1056–1068.

Clarke, T. J., 1984. Full reconstruction of a layered elastic medium from *P-SV* slant-stack data, *Geophys. J. R. astr. Soc.*, **78,** 775–793.

Coen, S. & Meadows, M., 1985. Exact and approximate inversion of plane-layered isotropic and anisotropic elastic media by layer stripping, *Abstract presented at 55th Annual Meeting of Society of Exploration Geophysicists*, Washington, DC.

Courant, R. & Hilbert, D., 1962. *Methods of Mathematical Physics*, vol. II, Wiley-Interscience, New York.

Kolb, P., Collino, F. & Lailly, P., 1986. Prestack inversion of a 1-D medium, *Proc. IEEE*, **74** (4).

Lavrent'ev, M. M., Reznitskaya, K. G. & Yakhno, V. G., 1986. *One-dimensional Inverse Problems of Mathematical Physics*, Translations, Series, 2, vol. 130, American Mathematical Society, Providence, RI.

McAulay, A., 1985. Prestack inversion with plane-layer point source modeling, *Geophysics*, **50,** 77–89.

Rakesh, 1986. A coefficient determination problem for the wave equation, *PhD thesis*, Department of Mathematics, Cornell University.

Sacks, P. & Santosa, F. 1987. A simple computational scheme for determining the sound speed of an acoustic medium from the surface impulse response, *SIAM J. Sci. stat. Comp.* **3,** (4), 501–520.

Sacks, P. & Symes, W., 1985. Uniqueness and continuous dependence for a multidimensional hyperbolic inverse problem, *Comm. in P.D.E.* **10,** 635–676.

Sacks, P. & Symes, W., 1987. Recovery of the elastic parameters of a layered half-space, *Geophys. J. R. Astr. Soc.* **88,** 593–620.

Santosa, F. & Symes, W. 1985. Determination of layered acoustic media *via* multiple impedance profile inversions from plane wave data, *Geophys. J. R. astr. Soc.*, **81,** 175–195.

Santosa, F. & Symes, W., 1986. An analysis of least squares velocity inversion, preprint; to appear as *SEG* Monograph (1988).

Santosa, F. & Symes, W. W., 1988. High-frequency perturbational analysis for the point-source response of a layered acoustic medium. *J. comp. Phys.*, **74,** (22), 318–381.

Symes, W., 1981. Stable solution of the inverse reflection problem for a smoothly stratified elastic medium, *SIAM J. math. Anal.*, **12,** 421–453.

Symes, W., 1985. Stability and instability results for inverse problems in several-dimensional wave propagation, *Proc. of 7th Internat. Conf. on Computing Methods in Applied Science and Engineering*, INRIA.

Symes, W., 1986a. Linearization stability for an inverse problem in several-dimensional wave propagation, *SIAM J. math. Anal.*, **17,** 132–151.

Symes, W., 1986b. On the relation between coefficient and boundary values for solutions of Webster's horn equation, *SIAM J. math. Anal.*, **17,** 1400–1420.

Symes, W., 1986c. Stability considerations for the velocity inversion problem, in *Proc. SEG/SIAM/SPE Conference on Seismic Exploration and Reservoir Modeling*, ed. Fitzgibbon, W. SIAM, Philadelphia.

Tarantola, A., 1984. The seismic reflection inverse problem, in *Inverse Problems of Acoustic and Elastic Waves*, eds Santosa, F., Pao, Y.-H., Symes, W. W. & Holland, C. H., SIAM, Philadelphia.

Thomsen, L., 1986. Reflection seismology in azimuthally anisotropic media, *Presented at 56th Annual Meeting, Society of Exploration Geophysicists*, Houston, TX.

Widom, H., 1969. *Lectures on Integral Equations*, van Nostrand, New York.

Yagle, A. & Levy, B., 1985. A layer-stripping solution of the inverse problem for a one-dimensional elastic medium, *Geophysics* **50**.

Ziolkowski, A., 1984. *Deconvolution*, International Human Resources Development Corporation, Boston, MA.