Inverting Waveforms for Velocities in the Presence of Caustics

William W. Symes
The Rice Inversion Project
Department of Computational and Applied Mathematics
Rice University

June 9, 1994

Abstract

Transmission caustics cause velocity estimation techniques based based on straightforward analysis of image volume moveout to fail. For example the *differential semblance* objective function (a quantitative measure of image volume moveout) is smooth and has only kinematically correct critical points when the incident wavefield is free of caustics, but loses these properties when caustics are present. The reasons for this phenomenon are closely related to those identified by Geoltrain and Brac for the inadequacy of Kirchhoff migration based on first arrival times for complex velocity structure. *Controlled illumination* of the subsurface, introduced by Rietvold and Berkhout to improve imaging given an adequate macromodel, also suggests a new approach to image volume moveout analysis for macromodel estimation. The transmission inversion (crosswell tomography) problem provides a simple framework within which to illustrate the controlled illumination reformulation of differential semblance and its effectiveness.

1 Introduction

In the last few years the author has introduced a method for waveform inversion of both macro- and micro-model which offers the advantages of both traveltime tomography (convergence from inaccurate initial guess) and conventional (least squares) inversion (estimates of short scale amplitudes, avoidance of time picking). This technique is called differential semblance optimization, and is based on a modification of the least squares objective function. The characteristic features of the method are:

- the micromodel (short scale components) is regarded as depending *a priori* on the shot (or experiment) parameter, thus forming an analogue or special case of the image volume of prestack migration;

- residual moveout in the inverted micromodel data volume is penalized by the mean square of nearest-neighbor differences - that is, by a *differential* measure of semblance;
- the differential semblance measure and the least squares error are combined into an objective function, which is minimized in two stages: an inner optimization for micromodel produces a function of the macromodel only; then this macromodel objective is minimized.

Regarding the last point, note that the micromodel is implicitly computed each time the macromodel is adjusted: that is, the optimizations are *nested*, rather than sequential as is envisaged in eg. Berkhout [1].

Considerable mathematical and numerical evidence exists that this technique is actually capable of estimating macro- and micromodel components accurately, albeit at considerable computational cost, directly from waveform data (Symes [19], [18], [22], [20], [21], Symes and Kern [26], [27]). Moreover implementations based on various physical modeling assumptions, from constant density acoustics to viscoelasticity, have been used to process field seismic data with reasonably satisfactory results (Symes and Carazzone [24], Minkoff and Symes [13]). So far as I know this is the *only* approach demonstrated to be capable of estimating both macro- and micromodel directly from waveform data sets using efficient gradient-based optimization algorithms. Related algorithms have been described by Bunks [2], Clement and Chavent [3], Jervis et al. [12], and Ehinger [4].

However differential semblance optimization ("DSO"), as originally formulated, will fail (lose its good numerical properties leading to global convergence) in the presence of complex ray geometry including multiple arrivals. Song and Symes [17] have shown an explicit example of this failure in the context of DSO for transmission velocity inversion (crosswell tomography). I will recapitulate a modification of this example below. In essence, when the current estimate of velocity is substantially wrong, the various branches of time present in the data appear as several residual moveouts in the inverted data volume, none of which is dominant. Fitting one of these branches implies that the other branches are not fit. Only when the velocity is very nearly correct are the amplitudes well enough reconstructed that one branch of moveout in the inverted data volume becomes clearly dominant. I will document this effect below, both analytically and numerically, in a simple example. The result: the DSO objective function takes on the appearance of the output least squares objective. Some part of the DSO data misfit term always saturated, as is generally the case with the OLS objective.

The problem is not peculiar to DSO: it will afflict all velocity analysis methods, including traveltime inversion, that use multiple arrival times either explicitly or implicitly. Since waveform inversion implicitly takes all arrival times into account, any velocity estimation technique based on waveform data must confront this difficulty.

Two recent papers, Geoltrain and Brac [6] and Rietvold and Berkhout [15] shed considerable light on the influence of trasmission caustics and multiarrivals. Both address principally the effect of transmission caustics on imaging, assuming that a correct macromodel is available. However the conclusions reached in these papers have important implications for macromodel estimation, which I will attempt to develop here.

The paper of Geoltrain and Brac [6] points out that first arrival traveltimes in complex media are often not energetically important, so that Kirchhoff migration algorithms using only first arrivals will give incorrect images. In effect the bulk of the energy, which comes into the data on later arrivals, is mapped into the image according to the first arrival. The effect on the image is similar to that of an incorrectly high migration velocity, i.e. apparent overmigration. The paper illustrates this effect in a completely convincing way, using the Marmousi data set and velocity model (Versteeg and Grau [28]). Of course precisely the same effect would obtain were an inversion based on first arrival times only used, which differs from a migration only in amplitude adjustments. The authors point out (p. 573) that such overmigration would "fool" a velocity analysis: even the exact velocity profile would appear too slow. This is because the amplitudes of the secondary arrivals are eliminated entirely. These secondary arrival amplitudes can be actually far larger than the first arrival amplitudes. The authors convincingly argue this proposition from geometrical optics considerations. They suggest as one possible ("deluxe") remedy using a complete WKBJ Green's function to formulate the migration, with all arrivals present with their correct amplitudes. As I shall show below this will indeed work, but only if the macromodel is correct. Otherwise, the migration - or even inversion - partitions energy incorrectly, assigning large amplitudes to events which should be very small when migrated with the correct macromodel. This mispartition of energy "fools" velocity analysis in exactly the way suggested by Geoltrain and Brac for first arrival migration: it introduces apparently important but fallacious residual moveout into the image volume.

Rietvold and Berkhout [15] address the question: with what form of data is an image best produced in complex velocity structure? They suggest a method of *controlled illumination* which produces *areal shot records*, synthetic data sets corresponding to highly distorted distributed ("areal")sources. These complex sources produce simple, single arrival incident waveforms in a specified target zone, with which imaging is easily and efficiently accomplished. The paper points out that the complex sources necessary to accomplish controlled illunination are easily computed, given the macromodel. Since the simple incident waveform can only be produced locally at depth, controlled illumination is presented as a technique for *target oriented* imaging. The salient points are also illustrated with the Marmousi model.

In the following pages I will combine the observations made in the papers of Geoltrain and Brac [6] and Rietvold and Berkhout [15] with the idea of DSO to suggest a velocity analysis technique for complex structure. Geoltrain and Brac [6] show that imaging must take into account all arrivals. I extend this lesson to understand the influence of velocity errors in the presence of multiple arrivals: the natural partition of energy, suggested by Geoltrain and Brac to correctly image events, is completely degraded by relatively small velocity errors. Thus the presence of multiple arrivals effectively disables velocity analysis, by DSO or any other means. Rietvold and Berkhout [15] describe a straightforward method to eliminate multiple arrivals from the incident front in specified depth zones. I give below an alternative derivation of areal sources for controlled illumination, based strictly on the wave equation and its propagation of singularity properties.

The simplest problem for which transmission caustics complicate velocity estimation is the transmission inversion or crosswell waveform tomography problem. Song and Symes [17] introduced and analyzed DSO inversion for the crosswell problem, assuming the absence of caustics in point source data. In this paper I introduce an areal source reformulation of differential semblance optimization for transmission inversion, and suggest its use to determine velocities when transmission caustics are present. I present analytical arguments and numerical examples to support the contention that areal source DSO retains the good properties possessed by the original formulation in the non-caustic case.

The reflection inversion problem (surface seismology) presents additional complications, which would obscure the basic message of this paper. A form of areal source DSO suitable for reflection seismology appears in Symes [23].

2 The DSO formulation of Crosswell Tomography

2.1 A simple model

Regard the earth as an acoustic medium occupying all of $\mathbf{R}^3$. That is, ignore the boundary conditions at the earth's surface, elastic wave effects, attenuation, anisotropy, and all of the other features which make actual seismic data interesting. Nonetheless the key difficulties of waveform based velocity estimation are present in this highly simplified model.

The "wells" are two vertical line segments. The source well is located at $x = x_s, y = y_s$, the receiver well at $x = x_r, y = y_r$. It suffices to consider a single receiver located at depth z_r, and an interval of source positions $z_{s,\min} < z < z_{s,\max}$. The theory developed below can accomodate an array of receivers at the price of slightly more complex notation; the conclusions remain the same.

The pressure field $P(z_s; x, y, z, t)$ satisfies the constant density acoustic wave equation

$$\left(\frac{1}{v^2}\frac{\partial^2 P}{\partial t^2} - \nabla^2 P\right)(z_s; x, y, z, t) = F(z_s; z, t)\delta(x - x_s)\delta(y - y_s)$$

$$P \equiv 0,\ t << 0$$

The line source density $F(z_s; z, t)$ depends parametrically on the position variable z_s. The choice of z_s is somewhat arbitrary; it is natural for reproducible isotropic point sources, for which

$$F(z_s; z, t) = f(t)\delta(z - z_s)$$

This is the *point source model*, to be assumed until further notice. The notation emphasizes that I will eventually consider more general line sources.

The velocity field v is subject to two major assumptions from the outset, both of which are natural for crosswell inversion theory (and assumed without comment in much work on the topic):

1. v is smooth on the wavelength scale, in particular produces no reflected waves;

2. v is constant in the crossline direction.

Of course neither assumption is true in practice. Recent work by Harris et al. [9] and Zhou et al. [29] shows that considerable information can be extracted from reflections in crosswell data. However in that work reflections are separated from direct arrivals and migrated to produce high resolution images. The emphasis in this paper is on interpreting direct arrivals to produce a tomogram of the sort obtained in the cited works by traveltime inversion - but from the waveform data directly.

The second assumption is highly questionable (eg. Mufti [14]) but underlies much of crosswell tomography - it is even implicit in the use of "tomogram" and related words.

The signal S measured at the receiver location (x_r, y_r, z_r) depends on the velocity field and the source density:

$$S[v]f(z_s,t) = P(z_s; x_r, y_r, z_r, t)$$

The notation emphasizes the linear dependence of S on f. Of course the dependence of S on v is quite nonlinear.

In view of the modeling assumptions made so far, the transmission inversion ("crosswell waveform tomography") problem is: given a crosswell seismogram $S_{\rm data}$, find model parameters v and f so that $S[v]f \simeq S_{\rm data}$.

2.2 Output least squares

An obvious formulation is *output least squares ("OLS")*: find v, f to minimize

$$J_{\rm OLS}[v,f] \equiv \int\int dz_s dt |S[v]f - S_{\rm data}|^2$$

The OLS objective function $J_{\rm OLS}$ is highly nonconvex in v, with many local minima. Therefore it is unsuitable for rapid optimization by Newton's method or similar algorithms. The reader may consult for instance Gauthier et al. [5], Santosa and Symes [16], Herman [10], and Symes and Carazzone [25] for graphical evidence supporting this contention for reflection inversion, and Song and Symes [17] for similar evidence for transmission inversion.

Note that the source waveform figures quadratically in $J_{\rm OLS}$. Thus it is relatively easy to compute the *reduced output least squares objective*

$$J^{\rm red}_{\rm OLS}[v] \equiv \min_f J_{\rm OLS}[v,f]$$

The minimum on the right hand side is taken over the set of square integrable source waveforms.

The presence of local minima, and the very narrow domain of attraction of the global minimizer, are not peculiar to the constant velocity model subspace, but characteristic of the output least squares method for waveform inverse problems.

If the sources and receiver are connected by unique rays of geometric optics, the behaviour of the output least squares objective is particularly easy to understand. In that case the source-receiver traveltime $\tau(z_s)$ is well defined. Suppose that the data $S_{\rm data}$ are noise-free,

$$S_{\rm data} = S[v^*]f^*$$

and that the trial velocity v and target velocity v^* both produce unique source-receiver rays. Geometric optics gives to excellent approximation

$$\begin{aligned} S[v]f(z_s,t) &\cong a(z_s)f(t-\tau(z_s)) \\ S_{\rm data}(z_s,t) &\cong a^*(z_s)f^*(t-\tau^*(z_s)) \end{aligned}$$

$$\begin{aligned} J_{\rm OLS}[v,f] &= \int\int dz_s dt\, |a(z_s)f(t-\tau(z_s)) - a^*(z_s)f^*(t-\tau^*(z_s))|^2 \\ &= \int\int dz_s dt\, a(z_s)^2 \left| f(t) - \frac{a^*(z_s)}{a(z_s)} f^*(t-\tau^*(z_s)+\tau(z_s))\right|^2 \end{aligned}$$

Evidently this quantity is minimized over f by

$$f[v](t) = \left(\int dz_s a^2(z_s)\right)^{-1}\left(\int dz_s a(z_s)a^*(z_s)f^*(t-\tau^*(z_s)+\tau(z_s))\right)$$

This gives $f[v^*] = f^*$, but when $\tau \neq \tau^*$ the target wavelet is averaged over a time interval. When $\tau - \tau^*$ varies by a wavelength or more over the source array, destructive interference occurs and $f[v] \approx 0$. Thus $J^{\rm red}_{\rm OLS}[v] \approx \|S_{\rm data}\|^2$ when v is sufficiently in error to cause the traveltime error to change by more than a wavelength. This *saturation* phenomenon effectively disables output least squares as an inversion principle for velocities.

2.3 Differential semblance

The *differential semblance* modification of output least squares begins with the observation that a *single trace* (for a single source location) is quite easy to fit, with arbitrary velocity, simply by chosing the source waveform appropriately. The point source model assumes that the source is an isotropic radiator independent of location, thus forcing one to attempt to fit all traces in the seismogram simultaneously. This task is impossible to carry out with any degree of success unless the velocity is nearly correct, thus nearly matching the data moveout with that of the simulated data.

Therefore allow the source waveform part of the model to depend explicitly on the source position parameter z_s. That is, $f = f(z_s, t)$. Then fitting the data becomes easy, but the coherence implicit in the initial model has been lost. The point source model assumes a reproducible, position independent source. Thus the position dependent source representation just introduced is *infeasible*. The differential semblance principle reintroduces feasibility in a controlled way, through a penalty term. Amongst all possible penalties for infeasibility, the *differential* choice gives the smoothest response to velocity changes, since nearby traces in the inverted source model are similar even when the velocity is incorrect. These considerations lead to the following objective functional:

$$J_\sigma[v,f] \equiv J_{\rm OLS}\left[v, \frac{\partial f}{\partial t}\right] + \sigma^2 \int\int dz_s dt \left|\frac{\partial f}{\partial z_s}\right|^2$$

where the first term on the right is reinterpreted in the obvious way for position dependent sources. The use of the t-derivative of f is a technical device to give the two terms equal weight in the frequency domain. Thus the source waveform f appearing in the DSO objective function is actually the primitive of the source waveform in the original model description.

We will carry this normalization throughout the paper. Therefore for convenience we redefine the seismogram operator as

$$S[v]f(z_s,t) = \frac{\partial}{\partial t}P(z_s; x_r, y_r, z_r, t)$$

Just as for OLS, the DSO objective function is quadratic in the source waveform f, which means that the reduced objective

$$J^{\rm red}_\sigma[v] \equiv \min_f J_\sigma[v,f]$$

is relatively easy to compute, using one of a number of iterative methods. The minimum here is taken over the set of (position dependent) source waveforms whose first derivatives are square integrable. Song and Symes [17] shows that the reduced DSO objective is smooth and has only kinematically correct minimizers over large subsets of velocity. This conclusion holds for moderate values of the differential semblance weight σ, *provided* that a unique ray connects sources and receiver (i.e. no transmission caustics). In that case, the two-step minimization of J_σ *via* gradient based optimization methods succeeds, i.e. produces velocity estimates fitting well the arrival times in the data. Note that gradient based optimization of the reduced OLS function typically does not succeed in this sense, unless the initial velocity produces traveltimes which fit the data well already. As one would expect from the definition, the reduced OLS function is the limit as $\sigma \to \infty$ of the reduced DSO function.

To illustrate these points consider the example in Figure 1. The data set displayed in Figure 1b is synthesized from the geometry depicted in Figure 1a. The velocity is constant except for a low velocity anomaly consisting of two depressions in which velocity is lowered by 5%. The target source time function is a Ricker wavelet. Figure 1c graphs the dependence of the reduced DSO objective for various convex combinations of the target velocity field and the ambient constant, with the differential semblance weight $\sigma = 1.0$. Evidently the DSO objective (or at least these slices of it) tend to be convex. Extensive theoretical and numerical evidence suggests that the DSO reduced objective function tends to be more convex for lower σ, whereas its behaviour incresingly resembles that of the OLS reduced objective function as $\sigma \to \infty$ (see eg. Gockenbach [7]).

Figure 1d illustrates the flattening, i.e. decreasing dependence on z_s, of the estimated source as the velocity moves from the ambient constant towards the target weight. These displays show the *residual moveout* of the inverted source waveforms as functions of source position and time. The information thus displayed is analogous to that contained in NMO or migration velocity analysis panels (coherency panels, common image gather, iso-x gathers, ...). Any significant dependence of the phases of the events in the panel on the source offset z_s is indicative of velocity error; velocity corrections should flatten the events. Since inversion is being used, rather than migration or normal moveout correction, the amplitudes are "flattened" as well, so that at the correct velocity the traces in the panel are almost identical in amplitude as well as in phase.

Song and Symes [17] and Gockenbach [7] explained these phenomena. The reader may consult these references for a detailed mathematical account of the following material.

The key to analysis of the DSO objective function is its $\sigma = 0$ limit, the behaviour of which is easy to anlayse, and which determines the behaviour of J_σ^{red} for small σ. Define the scaled limit

$$J_0[v] = \lim_{\sigma\to 0}\sigma^{-2} J_\sigma^{\mathrm{red}}[v]$$
$$= \int\int dz_s dt \left| \frac{\partial f_0[v]}{\partial z_s} \right|^2$$

where $f_0[v](z_s, t)$ is the (unique) solution of

$$S[v]f = S_{\mathrm{data}}$$

This prescription only makes sense if a solution to the above equation exists, which happens when the data are free of noise, i.e. for some $v^*(x, z), f^*(t)$

$$S_{\mathrm{data}} = S[v^*]f^*$$

To see what this means, recall that for the moment source/receiver rays have been assumed unique, so that the arrival time $\tau(z_s)$ of the ray from the source at z_s to the receiver at z_r in the velocity field v is well-defined. Then the geometric optics approximation gives

$$f_0[v](z_s,t) \cong \frac{a^*(z_s)}{a(z_s)} f^*(t - \tau(z_s) + \tau^*(z_s))$$

where τ^* is the traveltime corresponding to the velocity field v^*. Assuming that f^* is much more oscillatory than the amplitudes a and a^*,

$$J_0[v] \simeq \text{const} \left(\int dt \left|\frac{\partial f^*}{\partial t}\right|^2\right) \left(\int dz_s \left|\frac{\partial}{\partial z_s}(\tau - \tau^*)\right|^2\right)$$

Thus $J_0[v] \simeq 0$ when v and v^* give traveltimes differing essentially by a constant across the source location array. In fact $J_0[v]$ is proportional to the mean square error in the z-derivative of the traveltime, up to the approximation implied by the "$\simeq$". It's relatively easy to show that this measure of traveltime error has only stationary points in v which are kinematically correct, i.e. at which the traveltime is in error at most by a constant. Therefore the same is true of J_0. It follows that *the global optima of J_0 can be found by local methods, as there are no spurious (kinematically incorrect) stationary points.*

Since J_0 is proportional to the mean square error in the traveltime z-derivative, rather than the mean square error in the traveltime itself, minimization of J_0 determines the traveltime only up to a constant. This is easy to understand: if one takes $f(t) = f^*(t + t_0)$, then a compensating change in the velocity so that $\tau = \tau^* - t_0$ leaves the value of J_0 unchanged. That is, this ambiguity expresses the the freedom to choose the source time origin implicit in the formulation of the enlarged model space. The constant in the traveltime error vanishes if for instance it is known *a priori* that v and v^* share a ray. This would occur for instance if the difference between v and v^* were localized, so that the rays for extreme sources are the same. Alternatively for example a constraint on some moment of the source could play the same role, by limiting the translational degree of freedom enjoyed by the source in this formulation. A more realistic geometry with multiple receivers will also eliminate this ambiguity.

Note that the source figures in this expression only through its mean square (recall that f has been reinterpreted to be the primitive of the source waveform). In particular, the general shape of $J_0[v]$ is *independent of the source bandwidth*, in very strong contrast to the output least squares objective.

The additional error indicated by the use of "$\simeq$" is proportional to the ratio of the dominant data wavelength to the length scale of the velocity fluctuations allowed in the model. If the permitted velocity models v are subject to a suitable prior smoothness constraint, as is often appropriate in tomography, then the traveltime error (except for the constant noted above) is provably dominated by a measure of source wavelength. In particular this component of the error vanishes as the source bandwidth becomes infinite: in this "ray theoretic" limit $J_0[v]$ becomes literally an objective function for traveltime tomography.

In any case minimization of $J_0[v]$ for noise free data produces a kinematically correct velocity v, matching the traveltimes in the data (up to an additive constant). Moreover the minimization can be carried out by local, as opposed to global, optimisation methods. As one would expect, the same is true for $J_\sigma^{\rm red}$, provided that σ is "small". For data with suffieciently small noise level e, minimization of $J_\sigma^{\rm red}$ with appropriate positive σ gives kinematically correct v up to an error proportional to e and to a dominant wavelength. This solution can be continued as $\sigma \rightarrow \infty$, and in that limit becomes the OLS solution. See the cited references for details and examples.

In sum, so long as transmission caustics do not appear, DSO enables effective waveform inversion for velocities *via* local optimization methods, i.e. "tomography without picking".

2.4 Transmission caustics cause nonconvexity

Unfortunately the pretty picture painted in the preceding section is no longer correct when transmission caustics are present. For example, Figure 2a presents a model similar in geometry to that of Figure 1 but with a more severe low velocity anomaly. This velocity field produces a triplication in the vicinity of the receiver for a range of source positions, as can be seen in the data (Figure 2b). Figure 2c shows the reduced DSO objective, graphed versus anomaly strength. Evidently the DSO objective for this example has the same character as the OLS objective; in particular the value does not drop to near its minimum until the anomaly strength is nearly correct.

Once again it is possible to understand the behaviour of J_σ by exploring the $\sigma \rightarrow 0$ limit, using geometric optics. Away from the caustic locus geometric optics gives the approximation

$$S[v]f(z_s,t) \simeq \sum_{i=1}^{R} a_i(z_s) M_i \frac{\partial}{\partial t} f(z_s, t-\tau_i(z_s))$$

where the τ_i are the various ($R > 1$) arrival times, the a_i are the corresponding amplitudes, and the operators M_i are either the identity or the Hilbert transform of f depending on the number of caustics encountered by the ray ending in the local wavefront normal. For the first arrival ($i = 1$) always $M_i = I$, so we can write

$$S[v] = a_1 T_{-\tau_1}(I + \sum_{i=2}^{R} b_i M_i T_{\tau_1-\tau_i})\frac{\partial}{\partial t}$$

$$\equiv a_1 T_{-\tau_1}(I+K)\frac{\partial}{\partial t}$$

where $b_i = a_i/a_1$ and T is the translation operator defined by

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

Assume as before that

$$S_{\rm data} = S[v^*]f^* = a_1^* T_{-\tau_1^*}(I+K^*)\frac{\partial}{\partial t} f^*$$

and in particular that v and v^* are so close that their traveltime fields have the same number of branches. Then

$$f_0[v](z_s,t) = (I+K)^{-1}\frac{a_1^*}{a_1} T_{-\tau_1^*+\tau_1}(I+K^*)f^*$$

$$= \frac{a_1^*}{a_1} T_{-\tau_1^*+\tau_1} f^*$$

$$+\left(\sum_{i=2}^{R} b_i^* M_i T_{\tau_1^*-\tau_i^*} - \sum_{i=2}^{R} b_i M_i T_{\tau_1-\tau_i}\right) \frac{a_1^*}{a_1} T_{-\tau_1^*+\tau_1} f^* + \ldots$$

Although $-\tau_1^* + \tau_1$ and its z_s derivative get small as the trial velocity v approaches the target velocity v^*, the same is not true of the interbranch phase differences $\tau_1^* - \tau_i^*$ and $\tau_1 - \tau_i$. Thus the differential semblance of f_0 involves *differences of shifted wavelets* just as does the OLS objective. As with OLS, the differential semblance will saturate away from the correct velocity. field. Also like OLS, and in contrast to its behaviour when only a single arrival is present in the data, the objective function $J_0[v]$ saturates more quickly with higher source frequency content.

When the velocity is correct, Geoltrain and Brac [6] suggest that the amplitudes of the "crosstalk" events in the solution will be small, and so will not confuse the identification of reflectors. Figure 2d, rightmost panel (panel 6) illustrates this contention. However when the velocity is incorrect by even a small amount, the distribution of energy between the various events in the inverted micromodel volume becomes much more even. In Figure 2d, panel 5, the strength of the anomaly is only 20% below its target strength. However, note the very steeply dipping event between 200 and 500 ms. The error in the time difference $\tau_1^* - \tau_2^*$ between the first arrival and more energetic later arrival *in the data* causes this event, as suggested by the formula above for f_0. Its position changes very little as the velocity is altered, because the time shift responsible for its position is a *data*, not model, property. Only when the model becomes so accurate that $\tau_1 - \tau_2 \cong \tau_1^* - \tau_2^*$ does this "crosstalk" event begin to disappear. Because the moveout of such events is very large, they cause a large change in the DSO objective function when they appear. Note that in this example, $\sigma = 1 > 0$, and the normal equations are solved; the right hand sides of these involve the adjoint translation operator $T_\delta^T = T_{-\delta}$, and so the solution exhibits energetic precursors.

The displays in Figure 3 are analogous to coherency or common image panels in migration velocity analysis for reflection data. Multiple arrivals with significant energy would also "fool" these into showing large events with substantial moveout, disproportionate to the actual velocity error.

3 Definition of Areal Sources

Evidently a reformulation of the DSO objective is required which no longer involves multiple arrival times: these must be removed somehow from the incident wavefield. Of the several possibilities for accomplishing this goal, the most promising seems the *controlled illumination* method of Rietvold and Berkhout [15]. This section gives a wave equation based account of this idea, in the crosswell context.

The basic insight is appealingly simple: for any velocity field, it is possible to combine point sources in the source well to produce a simple wavefront with a *single phase* near the receiver. The *areal sources* so produced also give simple wavefronts for velocities near that used in their design. Therefore at least locally in velocity space, i.e. for small changes in velocity, the use of the areal source removes the difficulty explained above.

I carry out the construction in two space dimensions, and give a couple of remarks at the end about the three dimensional case. It will produce a source which yields the waveform

$$\frac{1}{4\pi r}f\left(t-\frac{r}{v_0}\right)$$

on the line $\{(x,z):x=x_r\}$ in the vicinity of the receiver point (x_r,z_r).
Here $r=\sqrt{(x_s-x_r)^2+(z-z_s)^2}$ and v_0 is some convenient reference velocity, eg. $v_0=v(x_r,z_r)$. That is, the areal source for parameter z_s produces exactly the wavefield that would have been produced by a point source in a constant velocity medium.

This is not the only possible "target" source: Rietvold and Berkhout [15] for example produce plane waves, rather than spherical waves, near the target. Plane waves have some advantages for reflection inversion. Spherical waves seem natural in the present context, as they give exactly the point source as the areal source when the velocity is in fact constant and equal to v_0.

Let $\phi(z,t)$ be a smooth cutoff function satisfying

$$\phi(z_r,t_r)=1,\ t_r=\frac{1}{v_0}\sqrt{(x_r-x_s)^2+(z_r-z_s)^2}$$

and vanishing a few wavelengths away from (z_r,t_r). Denote by $P_f[v](z_s,\cdot,\cdot,\cdot)$ the solution of the reverse time initial boundary value problem

$$\frac{1}{v^2}\frac{\partial^2 P}{\partial t^2}-\nabla^2 P=0,\ x<x_r$$

$$P(x_r,z,t)=\phi(z,t)\frac{1}{4\pi r}f\left(t-\frac{r}{v_0}\right),\ r=\sqrt{(x_r-x_s)^2+(z-z_s)^2}$$

$$P\equiv 0,\ t>>0$$

Several steps in the following derivation depend on the assumption that the field P, which is a piece of an outgoing spherical wave at $x=x_r$, consists of downgoing waves. Mathematically, the propagation of singularities theorem of Hörmander [11], justifies these steps, i.e. shows that the errors committed in these constructions are of lower order in frequency (smoother) than the terms retained. I reserve complete account of this reasoning for another place.

Assume that the velocity field v has been extended evenly about x_s, that is, $v(x,z)=v(2x_s-x,z)$. This has no effect on waves propagating between the wells. Define the mirror image field $\tilde{P}$ by

$$\begin{aligned}\tilde{P}(x,z,t) &= P(x,z,t)\ \text{if} x>x_s\\ &= P(2x_s-x,z,t)\ \text{if}\ x<x_s\end{aligned}$$

Evidently $\tilde{P}$ is not likely to be smooth enough to satisfy the wave equation in the usual sense, so the wave equation satisfied by it must be interpreted in the sense of distributions. For a smooth test function ψ,

$$\int\int\int dx dz dt \tilde{P}\left(\frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2}-\nabla^2\psi\right)$$

$$= \int \int dzdt \int_{-\infty}^{0} dx P(2x_s - x, z, t) \left(\frac{1}{v^2} \frac{\partial^2 \psi}{\partial t^2} - \nabla^2 \psi \right) (x, z, t)$$

$$+ \int \int dzdt \int_{0}^{\infty} dx P(x, z, t) \left(\frac{1}{v^2} \frac{\partial^2 \psi}{\partial t^2} - \nabla^2 \psi \right) (x, z, t)$$

$$= \int \int dzdt \int_{-\infty}^{0} dx \psi(x, z, t) \left(\frac{1}{v^2} \frac{\partial^2 P}{\partial t^2} - \nabla^2 P \right) (2x_s - x, z, t)$$

$$+ \int \int dzdt \int_{0}^{\infty} dx \psi(x, z, t) \left(\frac{1}{v^2} \frac{\partial^2 P}{\partial t^2} - \nabla^2 P \right) (x, z, t)$$

$$+2 \int \int dzdt \frac{\partial P}{\partial x} \psi(x_s, z, t)$$

whence it follows that in the sense of distributions,

$$\frac{1}{v^2} \frac{\partial^2 \tilde{P}}{\partial t^2} - \nabla^2 \tilde{P} = 2 \frac{\partial P}{\partial x} \delta(x - x_s)$$

Since I have assumed that P consists of downgoing waves, there exists a time $T_{\min}$ (which can be estimated by ray tracing) for which

$$P = \tilde{P} \simeq 0, \; t < T_{\min}$$

(actually this means that P and $\tilde{P}$ are *smoother* than the data for $t < T_{\min}$).

So far in the discussion f has been an arbitrary source waveform. If $f = \delta$, then the construction produces a kernel which can be used to compute areal sources by convolution with any source waveform.

To summarize the reasoning explained so far in an algorithm for computation of areal sources:

1. extend the velocity field evenly about $x = x_s$;
2. solve the reverse time initial boundary value problem given above with $f = \delta$ to find $P_\delta[v]$;
3. compute the trace (impulsive areal source)

$$G[v](z_s, z, t) = 2 \frac{\partial P_\delta[v]}{\partial x}(z_s, x_s, z, t);$$

4. set $F[v](z_s, z, \cdot) = G[v](z_s, z, \cdot) * f(\cdot)$ (areal source for the waveform f);

Then the solution of

$$\frac{1}{v^2} \left(\frac{\partial^2 U}{\partial t^2} - \nabla^2 U \right) (x, z, t) = F[v](z_s, z, t) \delta(x - x_s)$$

$$U \equiv 0, \; t < T_{\min}$$

produces a field U satisfying

$$U(x_r, z, t) = \frac{1}{4\pi r} f\left(t - \frac{r}{v_0}\right), \; r = \sqrt{(x_r - x_s)^2 + (z - z_s)^2}$$

for z near z_r.

Note that this problem satisfied by $F[v] = G[v] * f$ is an *inverse source problem*, i.e. a source is to be computed to create a field satisfying prescribed boundary values on a distant hyperplane.

Define the *areal source forward map* by

$$S^a[v]f(z_s, t) = P(z_s; x_r, z_r, t)$$

where P solves the initial value problem

$$\left(\frac{1}{v^2}\frac{\partial^2 P}{\partial t^2} - \nabla^2 P\right)(z_s; x, y, z, t) = (f * G[v](z_s; z, \cdot))(t)\delta(x - x_s)$$

$$P \equiv 0, \; t << 0$$

(i.e. P is the field radiated by the areal source $f * G[v]$).

Note that by definition,

$$S^a[v]f(z_s, t) = \frac{1}{4\pi r} f\left(z_s; t - \frac{r}{v_0}\right).$$

where r is as before the Euclidean distance from the source point (x_s, z_s) to the receiver point (x_r, z_r). That is, *the areal source forward map is independent of the velocity v.* Therefore from now on this map will be denoted simply S^a.

Now comes a most important point: *because the areal source is a linear combination of point sources in the source well, the data seismogram must be reformulated in the same way as the forward map!* That is, the appropriate data for the areal source inverse problem is

$$S^a_{\text{data}}[v](z_s, t) = \int\int dz dt' G[v](z_s; z, t - t') S_{\text{data}}(z, t')$$

The discussion to come depends crucially on a close relation between the areal source construction and the Radon transform. Indeed, suppose that the field $H(z, t)$ solves the inverse source problem

$$\left(\frac{1}{v^2}\frac{\partial^2 W}{\partial t^2} - \nabla^2 W\right)(x, z, t) = H(z, t)\delta(x - x_r)$$

$$W \equiv 0, \; t << 0$$

$$W(x_s, z, t) = S_{\text{data}}(z, t)$$

Then

$$S^a_{\text{data}}(z_s, t_r) = \int\int dz dt G[v](z_s; z, t_r - t) S_{\text{data}}(z, t)$$

$$= \int\int\int dx dz dt G[v](z_s; z, t_r - t)\delta(x - x_s)W(x, z, t)$$

The first two factors in the integrand are the right-hand side of the above boundary value problem for P with the choice $f = \delta$, so this integral is

$$= \int\int\int \left(\frac{1}{v^2}\frac{\partial^2 P}{\partial t^2} - \nabla^2 P\right)(z_s, x, z, t_r - t)W(x, z, t)$$

$$= \int\int\int dx dz dt P(z_s; x, z, t_r - t)\delta(x - x_r)H(z, t)$$

$$= \frac{-1}{4\pi}\int\int dz dt r^{-1} H(z, t)\delta\left(t_r - t - \frac{r}{v_0}\right)$$

where as before r is the source-receiver distance. Finally note that this last integral gives the solution of *constant coefficient* wave equation with velocity v_0 for the source $H(z, t)\delta(x - x_r)$, restricted to the plane $\{x = x_s\}$.

That is, the areal source seismogram is obtained by solving the inverse source problem for H with the point source seismogram as data, then performing the hyperbolic Radon transform, i.e. solving the forward source problem with constant velocity. As with the inverse source problem defining $G[v]$, to good approximation H is the Neumann data for the solution of the backwards Dirichlet problem with data $S_{\rm data}$ on $x = x_s$. Thus the areal source seismogram $S_{\rm data}[v]$ is the product of an inverse modeling step with the velocity v applied to the point source data, followed by a forward modeling step with constant velocity v_0.

To illustrate the construction of areal sources, I have carried out the construction described above for the model of Figure 2a. The impulsive areal source ($G(z_s, z, t)$, fixed z_s) appears as Figure 3a. Since I used finite difference approximation to carry out these calculations, the results show considerable numerical noise. The areal source seismogram for a 15 Hz Ricker time dependence appears as Figure 3b, and is essentially a perfect representation of the constant velocity target seismogram.

This concludes our account of the Rietvold and Berkhout [15] method of controlled illumination for the 2D crosswell configuration. It is natural to wonder whether an extension to fully 3D wave propagation is possible. As the source well is not a hypersurface in 3D, sources confined to the well cannot control amplitudes along rays in all directions. This observation appears to preclude elimination of multiarrivals due to out of plane propagation, at least through the sort of distributed source construction discussed here. Of course volumetric control of velocity from crosswell data is problematic in any case.

4 The DSO objective function defined by an Areal Source

Recall that

$$S[v]f = S_{\rm data}$$

if and only if

$$S^a f = S^a_{\rm data}[v]$$

Here the source waveform f depends on the "source coordinate" z_s as in the earlier discussion of the differential semblance principle, although of course the source is now distributed and z_s is merely a parameter. It follows from the assumed isotropic, position independent nature of the source that a *feasible* source waveform should be position independent, just as before. These considerations lead to the definition of the *areal source differential semblance objective function*:

$$J^a_\sigma[v,f] \equiv \int\int dz_s dt \left\{ |S^a f - S^a_{\rm data}[v]|^2 + \sigma^2 \left|\frac{\partial f}{\partial z_s}\right|^2 \right\}$$

Note that if $S_{\rm data} = S[v^*]f^*$ for a coherent (z-independent) source f^*, then $J^a_\sigma[v^*,f^*] = 0$.

J^a_σ is quadratic in f, which can therefore be eliminated to give

$$J^{a,{\rm red}}_\sigma[v] \equiv \min_f J^a_\sigma[v,f]$$

The scaled $\sigma \rightarrow 0$ limit is

$$J^a_0[v] \equiv \lim_{\sigma\rightarrow 0}\sigma^{-2} J^{a,{\rm red}}_\sigma[v] = \left|\frac{\partial f^a_0[v]}{\partial z_s}\right|^2$$

where

$$S^a f^a_0[v] = S^a_{\rm data}[v]$$

As before this limit only makes sense for noise-free data, i.e.

$$S_{\rm data} = S[v^*]f^*$$

with f^* independent of z_s.

As in the single arrival case, the behaviour of the DSO objective function for areal sources J^a_σ behaves for noise-free data and for small σ like its $\sigma \rightarrow 0$ limit J^a_0. In view of the expression for the point source seismogram, it is natural to expect that (except at caustic points)

$$S^a_{\rm data}[v](z_s,t) \simeq \sum_{i=1}^R a^{*,a}_i(z_s) M_i \frac{\partial f^*}{\partial t} f(z_s, t - \tau^{*,a}_i(z_s))$$

whence

$$f^a_0[v] \simeq -\sum_{i=1}^R \frac{a^a_i(z_s)}{4\pi r} M_i f\left(z_s, t - \tau^{*,a}_i(z_s) + \frac{r}{v_0}\right)$$

Thus the entire issue hinges on the behaviour of the phases $\tau^{*,a}_i$. Note that the sum above is *finite*, in contrast to the expression for $f_0[v]$ in the case of multiple arrivals.

Note also that only the difference(s) of the (possibly multiple) data phases $\tau^{*,a}_i$ and the predicted phase r/v_0 appear in the expression for $f^a_0[v]$. The interbranch phase differences appearing in the expression given above for $f_0[v]$ (with transmission caustics in the data) are absent.

When $v = v^*$ (the correct velocity), the number R of phases collapses to one, and

$$S^a_{\rm data}[v^*] = -\frac{1}{4\pi r}\frac{\partial f}{\partial t}\left(t - \frac{r}{v_0}\right)$$

In fact the phases are *stable*, that is depend smoothly on v. In particular for v near v^* the phase is uniqe ($R = 1$) and then

$$S^a_{\rm data}[v] \simeq a^{*,a}[v]\frac{\partial f}{\partial t}(t - \tau^{*,a}[v])$$

The sense of "near" is quite strong, involving several spatial derivatives. When v is sufficiently far from v^* the phase becomes multivalued (i.e. $R > 1$), but is still stable in a suitable sense (see below).

The gradient of J^a_0 admits a similar analysis to that of the caustic-free case above, provided that v is near enough v^* that the phase is single valued. From the definition of S^a one has immediately that

$$f^a_0[v] \simeq -\frac{a^{*,a}[v]}{4\pi r} f^*\left(t - \tau^{*,a}[v] + \frac{r}{v_0}\right)$$

whence

$$J^a_0[v] \simeq \int dt \left|\frac{\partial f^*}{\partial t}\right|^2 \int dz_s \left|\frac{\partial}{\partial z_s}\left(\tau^{*,a}[v] - \frac{r}{v_0}\right)\right|^2$$

This expression is very similar to that derived above for $J_0[v]$ in the caustic-free case. Note in particular that the target source wavelet f^* appears only through the mean square of its t-derivative, so that the shape of J^a_0 is independent of the frequency content of the wavelet. In this respect J^a_0 is similar to the function J_0 in the single arrival case, and differs from J_0 when multiple arrivals are present.

The gradient of $J^a_0[v]$ for v close to v^* is also free of spurious zeroes, i.e. all stationary points v satisfy

$$\frac{\partial}{\partial z_s}\left(\tau^a[v] - \frac{r}{v_0}\right) \simeq 0$$

The proof of this result relies on the single valued nature of the phase τ^a, so does not extend immediately to the larger set of v far away from v^*. I conjecture that the result remains true in a large set in which τ^a becomes multivalued.

In support of this conjecture, and to illustrate the improved nature of the differential semblance objective function defined by an areal source, consider once again the data of Figure 2. Figure 3c graphs the values of $J^{a,{\rm red}}_\sigma$ along the line segment between the constant background velocity and the target velocity with low velocity anomalies. While not convex the behaviour of $J^{a,{\rm red}}_\sigma$ is much more suggestive of good response to local optimization methods than are the curves in Figure 2c. Figure 3d is analogous to Figure 2d, and shows the source as function of source depth parameter. Note that the limbs of source estimate systematically weaken and flatten, in contrast to the behaviour of the souce estimates in Figure 2d.

To conclude this section, I present a sketch of the proof of stability of the multiarrival phases τ^a_i. I believe that this sketch can serve as the outline of a full mathematical treatment. Recall that the areal source data is obtained by solving the inverse source problem with the point source response as data and with trial velocity v for a distribution on $x = x_s$, then by solving the forward (source) problem with constant velocity to obtain $S^a_{\rm data}[v]$. Corresponding wavefronts are constructed by the method of characteristics. Thus a point in the wavefront of $S^a_{\rm data}$ is the result of forward ray tracing from the source to the receiver in v^*, backwards ray tracing from receiver to $x = x_s$ in v, and then forward ray tracing in constant velocity v_0 back to the receiver plane $x = x_r$. Near

$x = x_s$ typically several points in this wavefront cover a given spatial location, unless v is close to v^*. That is the surface built in this way is singular, unless v is close to v^*. On the other hand for $v = v^*$ the surface is spherical. Since the rays depend smoothly on the velocity (thanks to the fundamental theorem on ordinary differential equations), so does the surface constructed from them. This establishes the existence of the smooth phase mentioned above for v near v^*. For v further away, the (singular) phase surface is better represented by its normal conoid (Lagrangian, see eg. Guillemin and Sternberg [8], Ch. 2). This conoid is also smoothly dependent on the velocity, which establishes the smooth dependence of the multiple phases globally in v, in a suitable sense.

Acknowlegement

This work was partially supported by the Texas Geophysical Parallel Computation Project, the Schlumberger Foundation, and The Rice Inversion Project. TRIP Sponsors for 1994 are Advance Geophysical, Amoco Production Co., Conoco Inc., Cray Research Inc., Exxon Production Research Co., Interactive Network Technologies, Mobil Research and Development Corp., and Texaco Inc.

References

[1] A.J. BERKHOUT. Seismic inversion in terms of pre-stack migration and multiple elimination. *Proc. IEEE*, 74:415–427, 1986.

[2] C. BUNKS. Paraxial wave equation inversion with geometric constraints. Abstract B012. In *Proceedings*. European Assoc. Explor. Geophys., 1992.

[3] FRANÇOIS CLÉMENT and GUY CHAVENT. Waveform inversion through MBTT formulation. In E. Kleinman, T. Angell, D. Colton, F. Santosa, and I. Stakgold, editors, *Mathematical and Numerical Aspects of Wave Propagation*. SIAM, Philadelphia, 1993.

[4] A. EHINGER. How to pick traveltime in complex structures? In *54th Meeting*. European Association of Exploration Geophysicists, 1992. Abstract.

[5] O. GAUTHIER, A. TARANTOLA, and J. VIRIEUX. Two-dimensional nonlinear inversion of seismic waveforms. *Geophysics*, 51:1387–1403, 1986.

[6] S. GEOLTRAIN and J. BRAC. Can we image complex structures with first-arrival traveltime? *Geophysics*, 58:564–575, 1993.

[7] MARK GOCKENBACH. *An abstract analysis of differential semblance optimization*. PhD thesis, Computational and Applied Mathematics Department, Rice University, Houston, Texas, U.S.A., 1994.

[8] V. GUILLEMIN and S. STERNBERG. *Geometric Asymptotics*. American Mathematical Society, Providence, 1979.

[9] J. HARRIS. Practical algorithms for crosswell seismic imaging. In S. E. Zarantonello, editor, *Mathematical Methods in Geophysical Imaging*. Society of Photooptical Instrumentation Engineers, Bellingham, 1993.

[10] G. HERMAN. Generalization of traveltime inversion. *Geophysics*, 57:9–14, 1992.

[11] L. HORMANDER. Fourier integral operators. *Acta Mathematica*, 127:79–123, 1971.

[12] M. JERVIS, M.K. SEN, and P.L. STOFFA. Optimization methods for 2-D pre-stack migration velocity estimation. In *Proc. 63rd Annual International Meeting*, pages 965–968, Washington, D.C., 1993. Society of Exploration Geophysicists.

[13] S.E. MINKOFF and W.W. SYMES. Viscoelastic modeling and inversion of a marine data set. Rice inversion project technical report, Department of Computational and Applied Mathematics, Rice University, Houston, TX, 1994.

[14] I. MUFTI. Tomographic seismic imaging of reservoirs in three-dimensional environments. In *Proc. 62st Annual International Meeting*, pages 79–82, Houston, 1992. Society of Exploration Geophysicists. Expanded abstract.

[15] W. RIETVOLD and A.J. BERKHOUT. Depth migration combined with controlled illumination. In *Proc. Soc. Explor. Geophys. Annual Mtg.*, pages 931–934. Society of Exploration Geophysicists, 1992.

[16] F. SANTOSA and W.W. SYMES. *An Analysis of Least-Squares Velocity Inversion*, volume 4 of *Geophysical Monographs*. Soc. of Expl. Geophys., Tulsa, 1989.

[17] H. SONG and W.W. SYMES. A differential semblance formulation of crosswell tomography. Technical Report TR92-22, Department of Computational and Applied Mathematics, Rice University, Houston, TX, 1993.

[18] W.W. SYMES. A differential semblance algorithm for the inverse problem of reflection seismology. *Computers and Math. with Appl.*, 22(4/5):147–178, 1991.

[19] W.W. SYMES. Layered velocity inversion: a model problem from reflection seismology. *SIAM J. Math. Anal.*, 22:680–716, 1991.

[20] W.W. SYMES. The plane-wave detection problem. *Inverse Problems*, in press, 1992.

[21] W.W. SYMES. In S. Hassanzadegh, editor, *Proceedings, Symposium on Geophysical Image Processing*. SPIE, 1993.

[22] W.W. SYMES. A differential semblance criterion for inversion of multioffset seismic reflection data. *J. Geophys. Res.*, 98:2061–2073, 1993.

[23] W.W. SYMES. Velocity estimation from reflected waves in the presence of transmission caustics. In S. Zarantonello, editor, *Mathematical Methods in Geophysical Imaging II*. SPIE, Bellingham, 1994.

[24] W.W. SYMES and J. CARAZZONE. Velocity inversion by differential semblance optimization. *Geophysics*, 56(5):654–663, 1991.

[25] W.W. SYMES and J.J. CARAZZONE. Velocity inversion by coherency optimization. In J.B. Bednar, editor, *Geophysical Inversion*, pages 59–89. Society for Industrial and Applied Mathematics, Philadelphia, 1992.

[26] W.W. SYMES and M. KERN. Geophysical prospecting. in press.

[27] W.W. SYMES and R. VERSTEEG. Velocity model determination using differential semblance optimization. In *Proc. 63rd Annual International Meeting*, pages 696–699, Washington, D.C., 1993. Society of Exploration Geophysicists. Expanded abstract.

[28] R. VERSTEEG and G. GRAU. Practical aspects of inversion: The Marmousi experience. Proceedings of the EAEG, The Hague, 1991.

[29] Y. ZHOU, C. CAI W. LUO and G. SCHUSTER. Hybrid wave equation traveltime and waveform inversion of crosswell seismic data. In S. E. Zarantonello, editor, *Mathematical Methods in Geophysical Imaging*. Society of Photooptical Instrumentation Engineers, Bellingham, 1993.

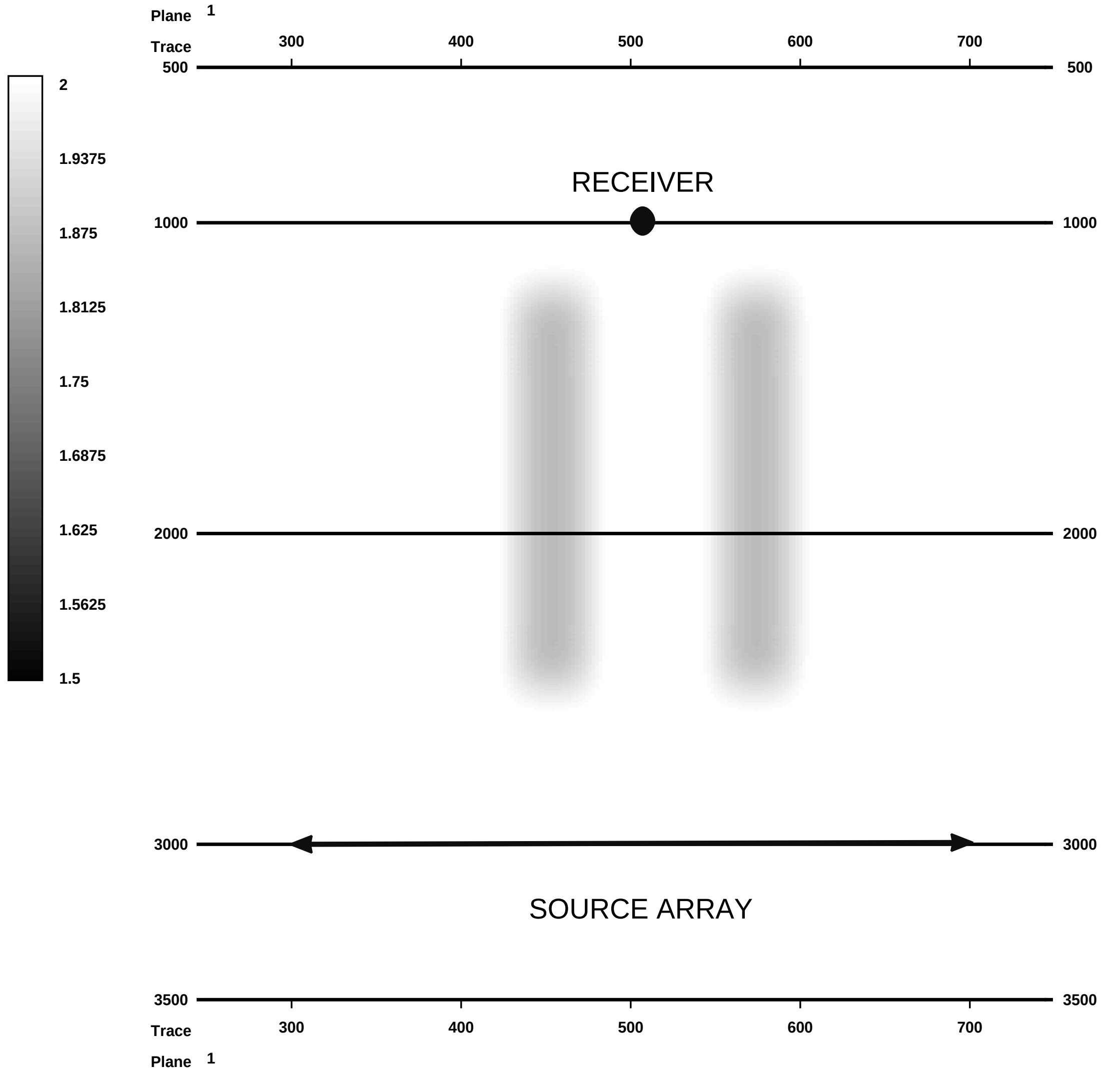

FIGURE 1A: Velocity model with two low velocity anomalies. Ambient level: 2 km/s. Minimum in center of anomalies: 1.9 km/s. Trace spacing in plot: 8 m. Depth scale is in meters. Sources form a line at depth 3 km, from 2.4 to 5.6 km offset. Receiver is at depth 1 km, offset 4 km.

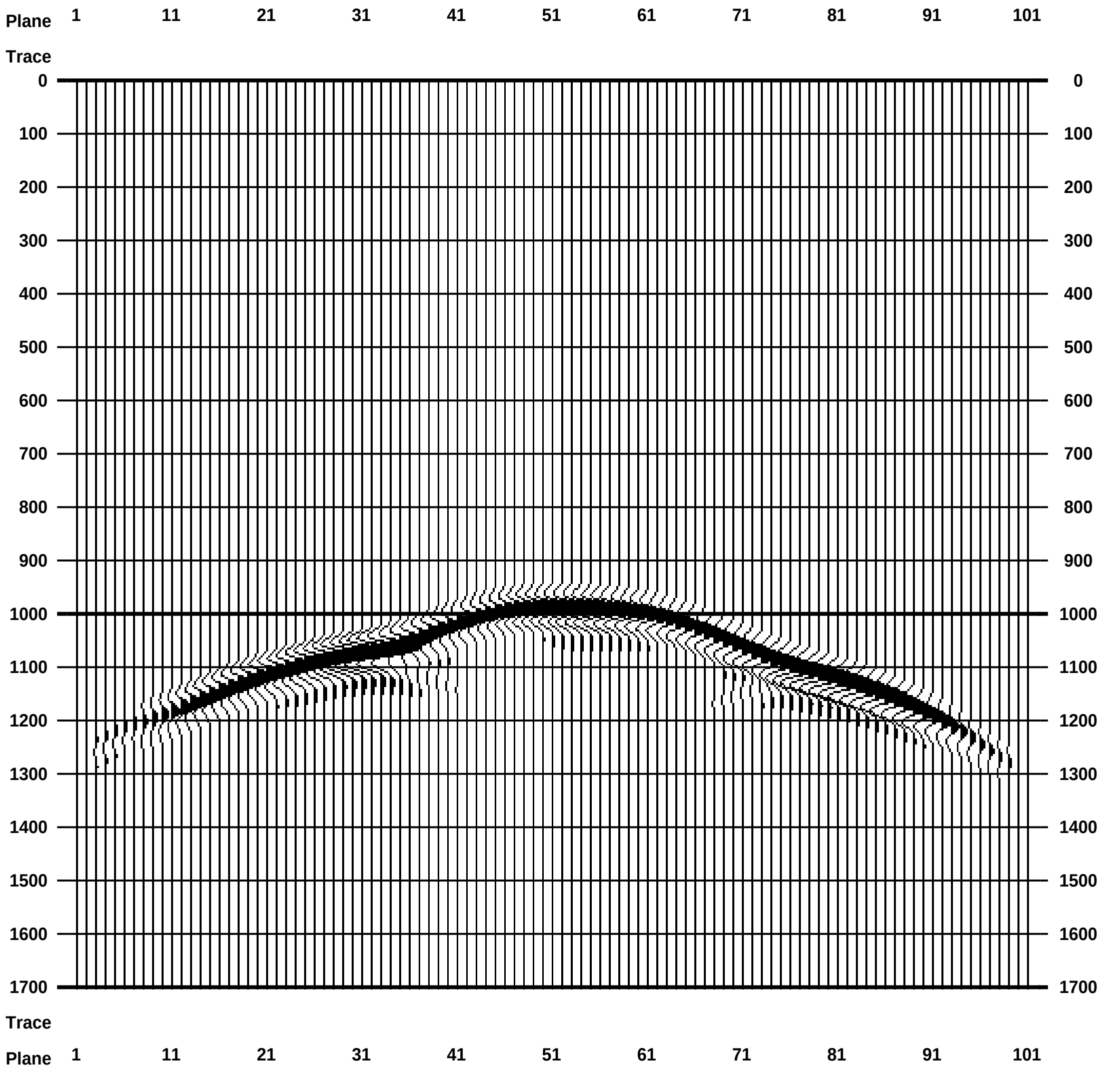

FIGURE 1B: Transmission seismogram for configuration of Figure 1a. Receiver spacing is 32 m; offset at left is 2400 m, at right 5600 m. Vertical axis marked in milliseconds.

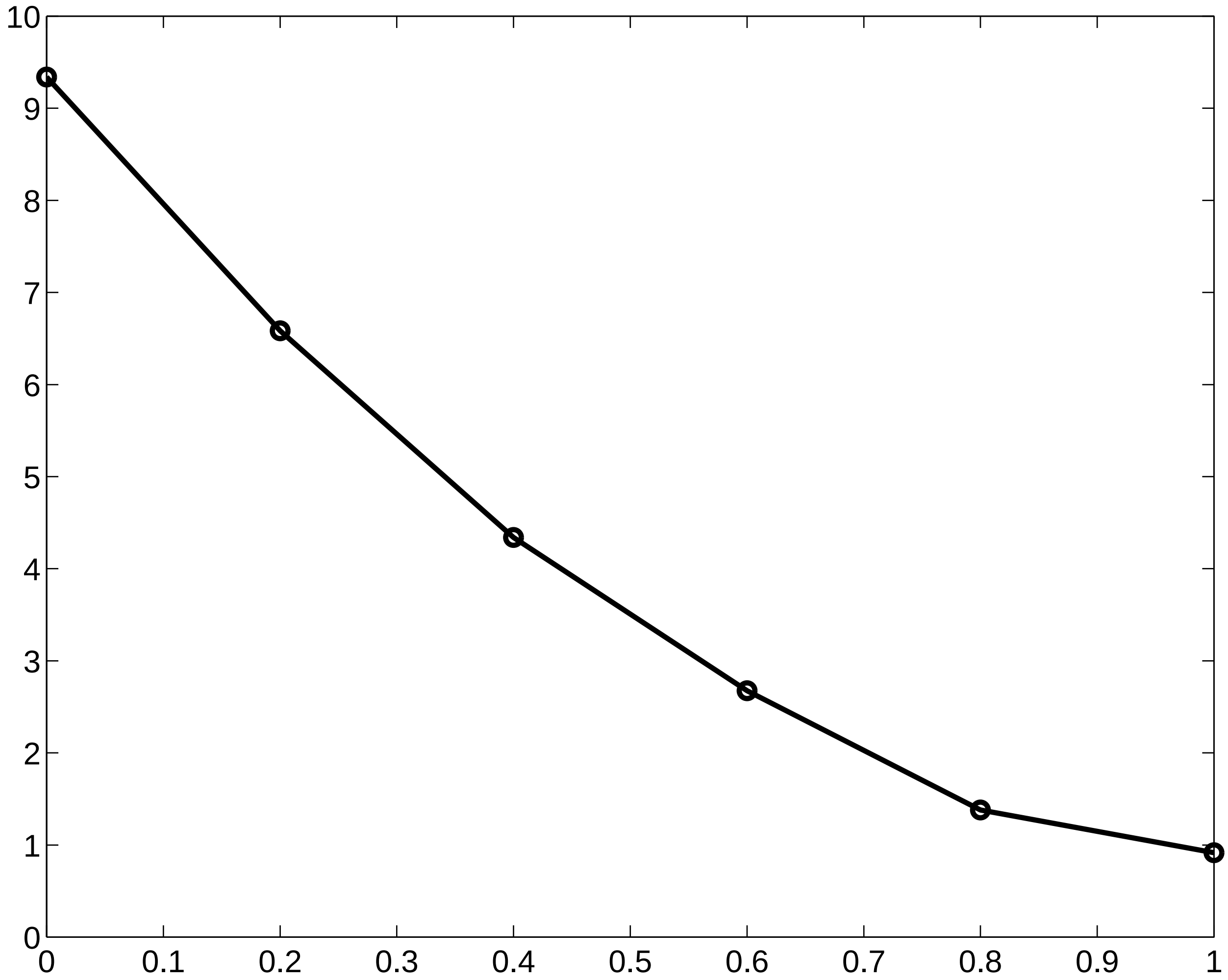

FIGURE 1C: Plot of differential semblance along line segment $(1-h)v_0 + hv^*$, for $h = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0$. Here $v_0 \equiv 2$km/s, and v^* is the velocity of Figure 1a.

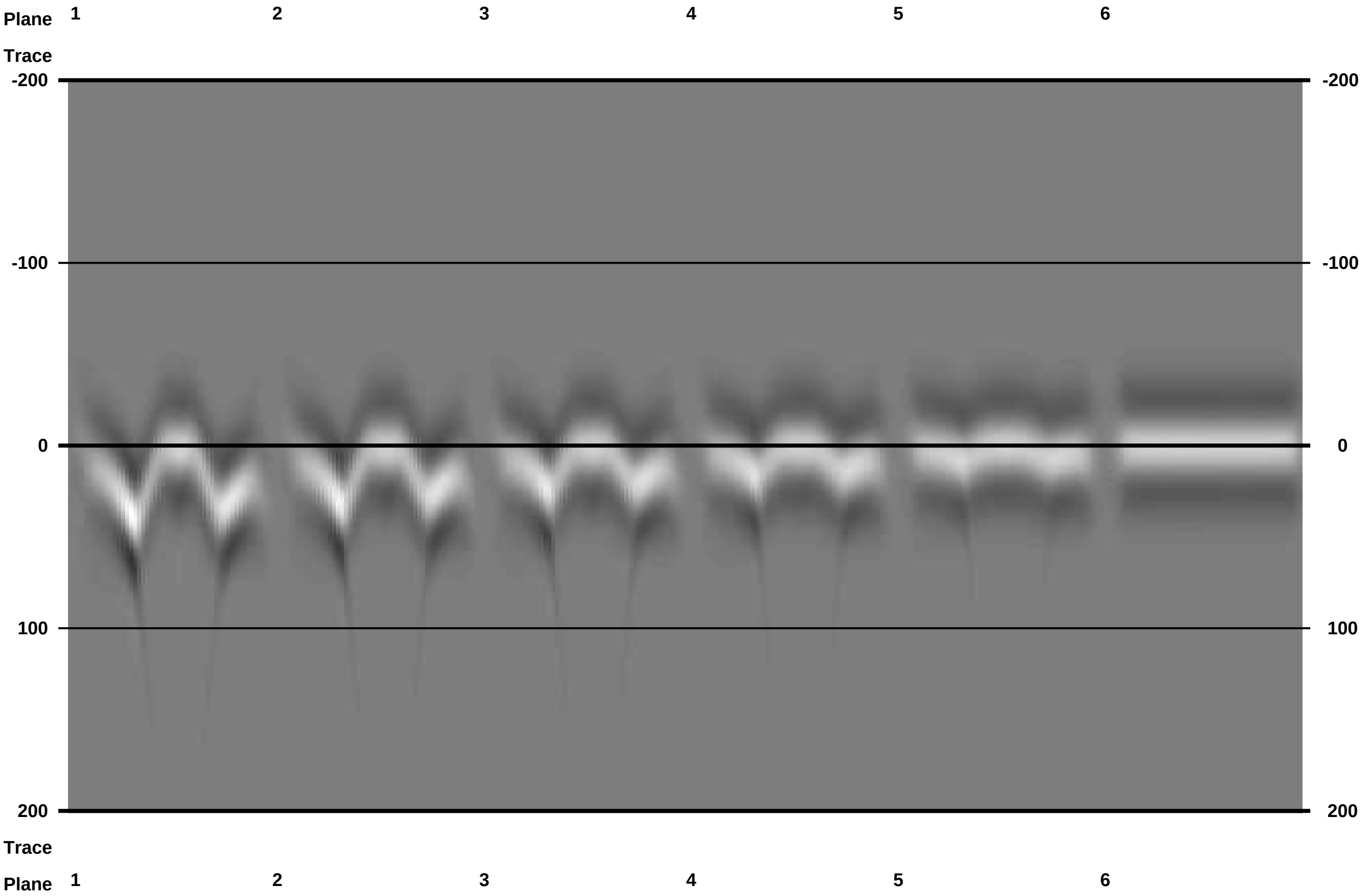

FIGURE 1D: A coherency display of the inverted sources for velocities on the line segment $(1 - h)v_0 + hv^*$, for $h = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0$. Here $v_0 \equiv 2$km/s, and v^* is the velocity of Figure 1a. The vertical axis is marked in milliseconds. Each plane (1 - 6) represents one value of h.

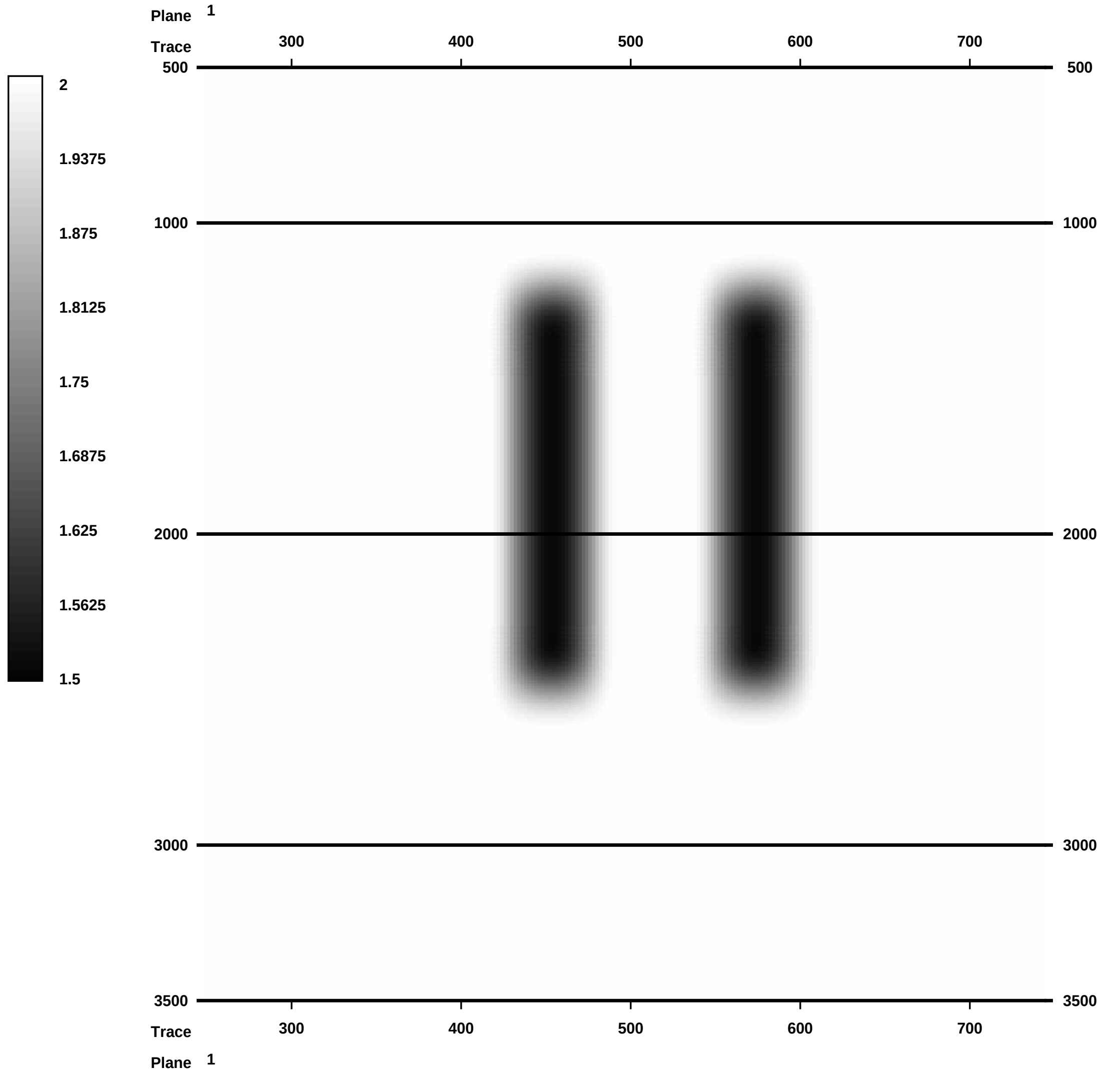

FIGURE 2A: Velocity model with two low velocity anomalies, deeper than the example of Figure 1. Ambient level: 2 km/s. Minimum in center of anomalies: 1.5 km/s. Geometry and scales same as in Figure 1a.

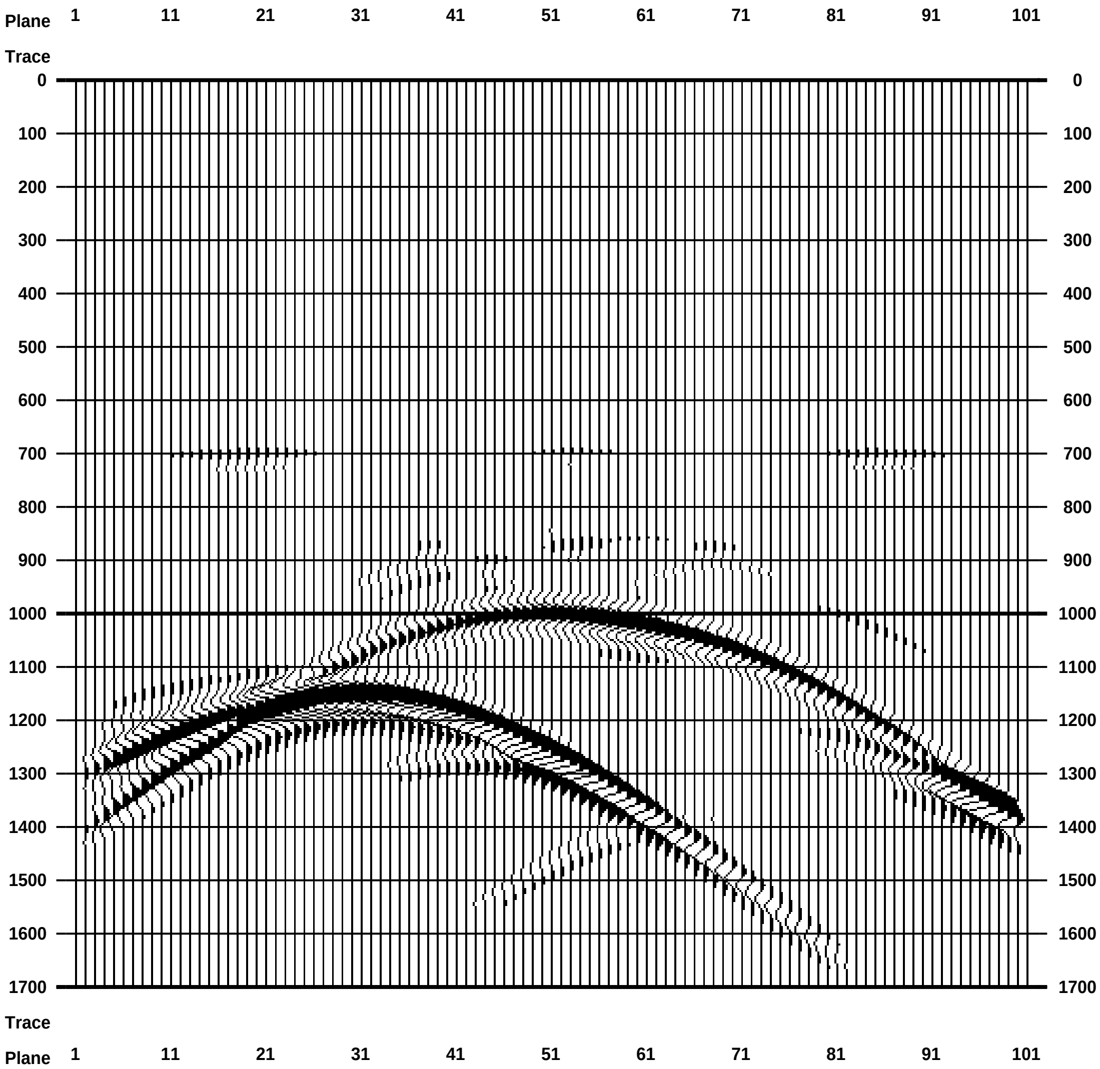

FIGURE 2B: Transmission seismogram for configuration of Figure 2a. Note energetic late arrivals.

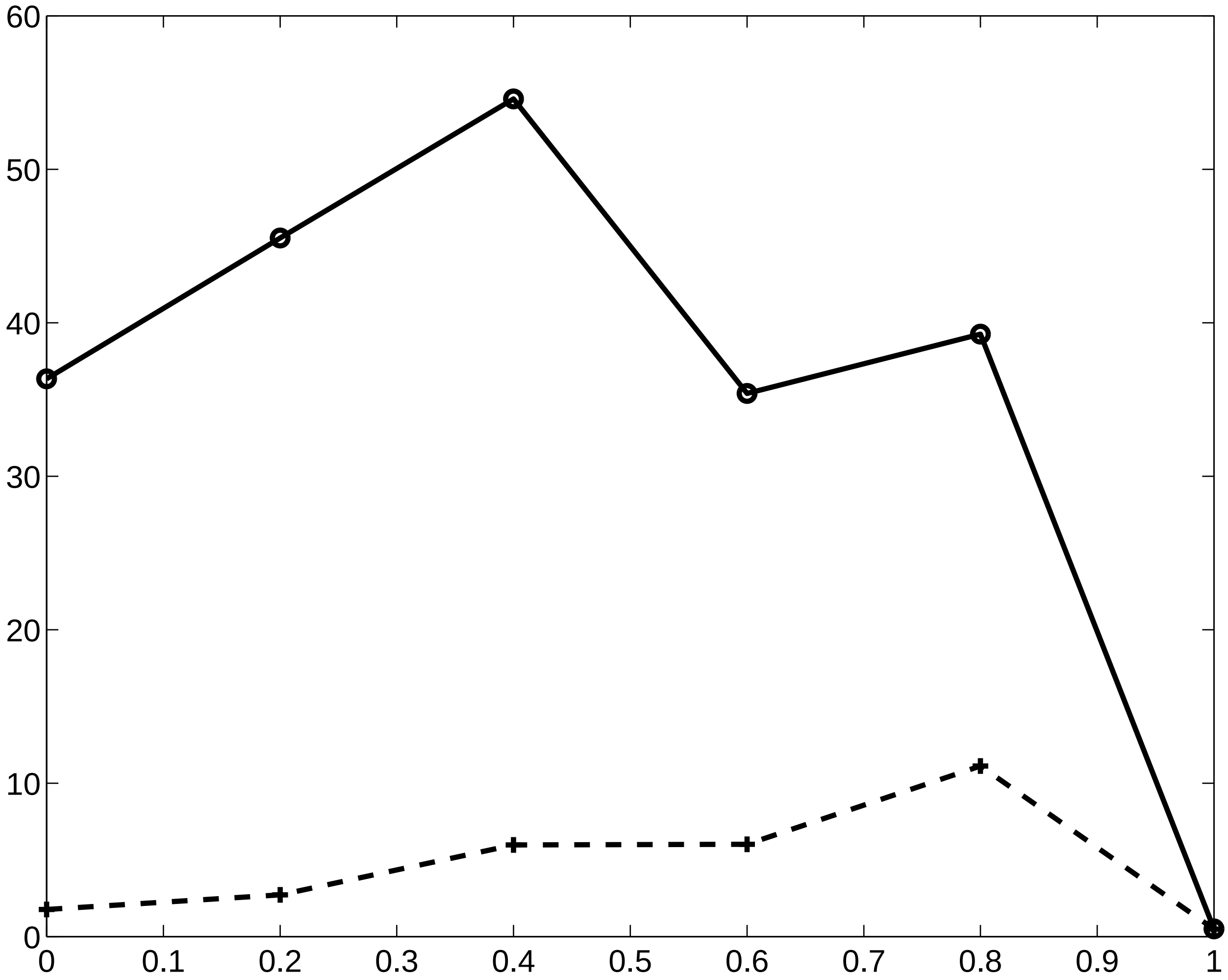

Figure 2c: Plot of differential semblance along line segment $(1-h)v_0 + hv^*$, for $h = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0$. Here $v_0 \equiv 2$km/s, and v^* is the velocity of Figure 2a. For the solid line $\sigma = 1.0$, for the dashed line $\sigma = 0.1$.

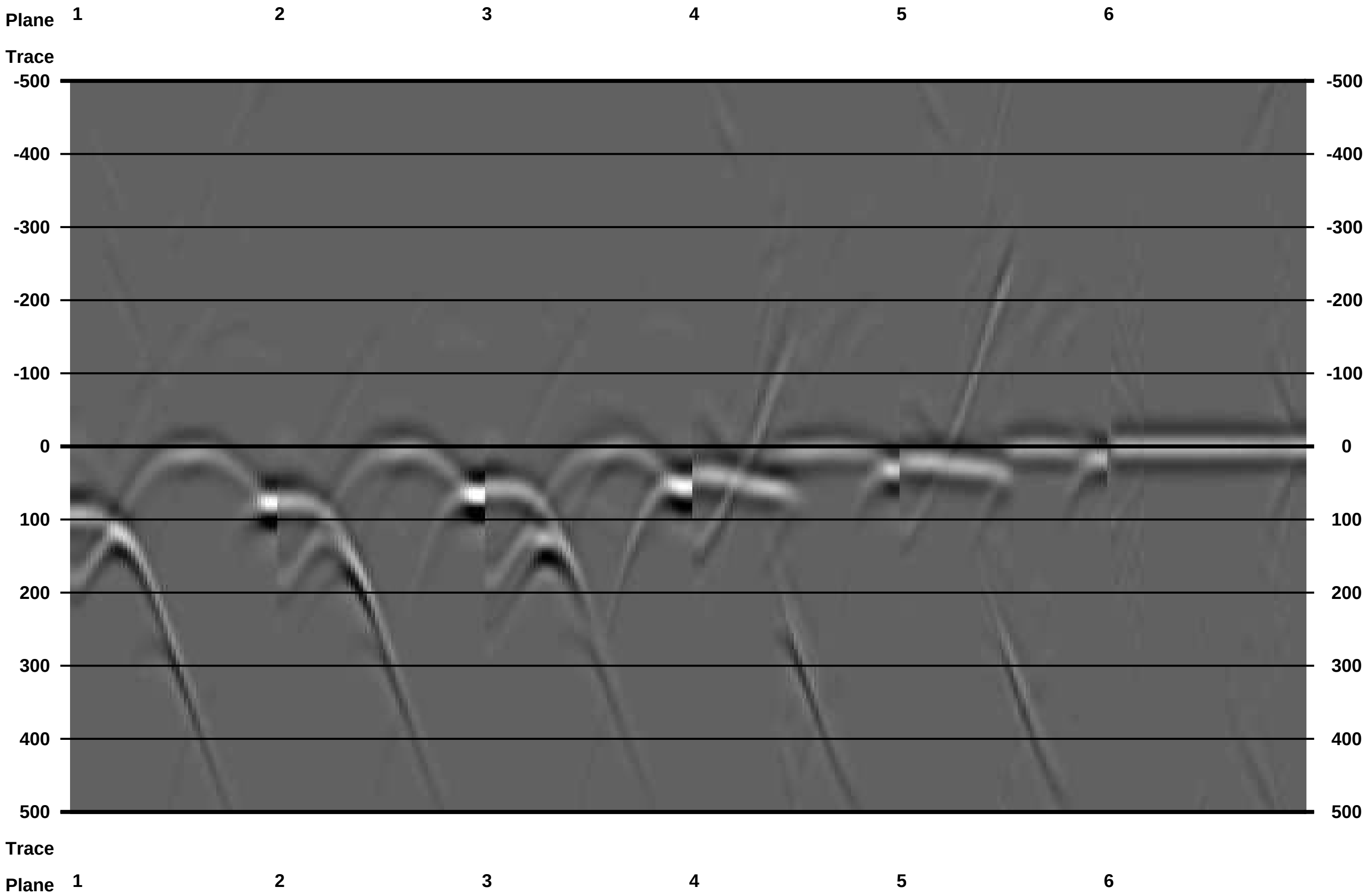

FIGURE 2D: A coherency display of the inverted sources for velocities on the line segment $(1-h)v_0 + hv^*$, for $h = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0$. Here $v_0 \equiv 2$km/s, and v^* is the velocity of Figure 2a. Note the persistance of the steeply dipping event between 200 and 500 ms for panels 1 - 5.

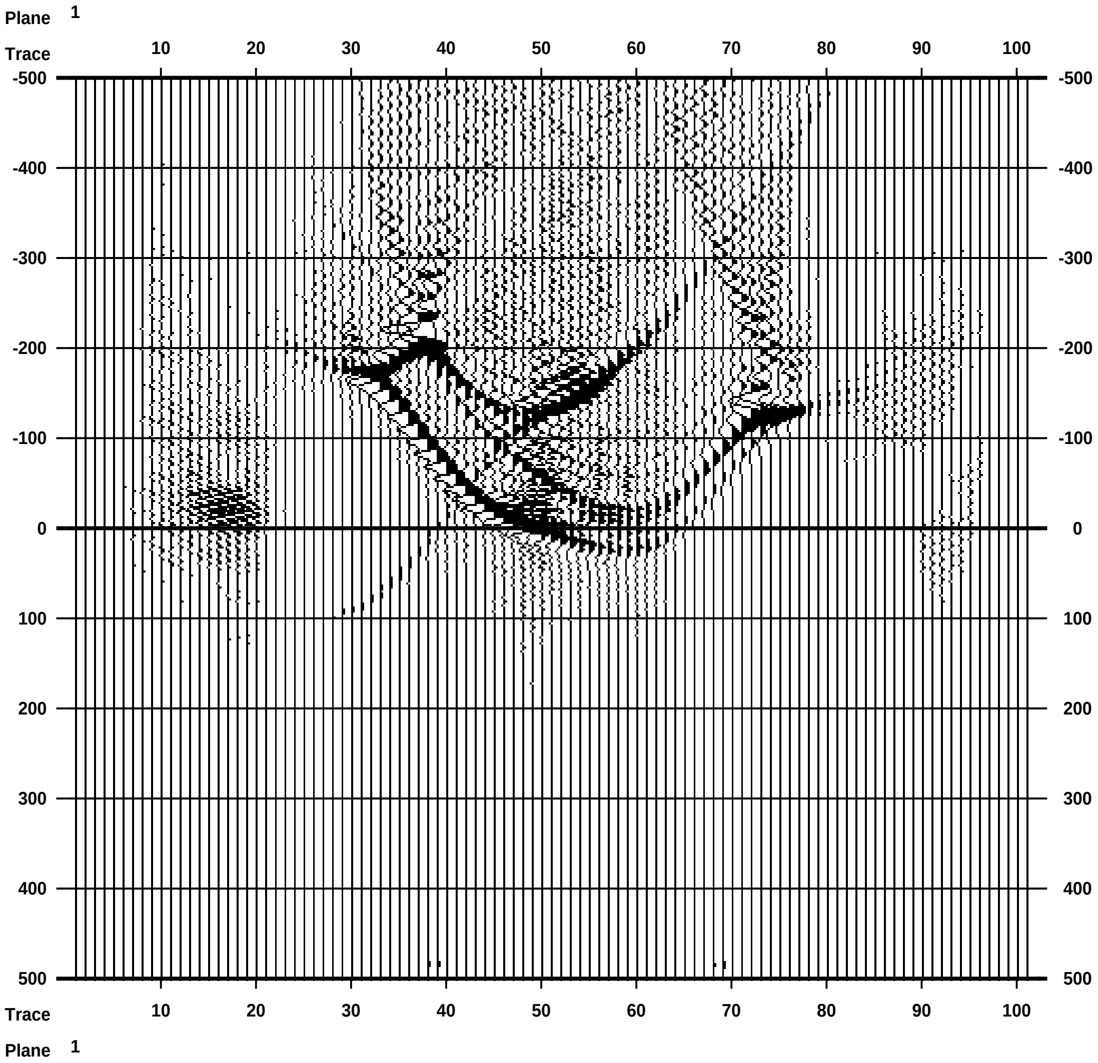

Figure 3a: Areal source impulsive kernel $G(z_s; z, t)$ for model of Figure 2a, $z_s = 4000$ m.

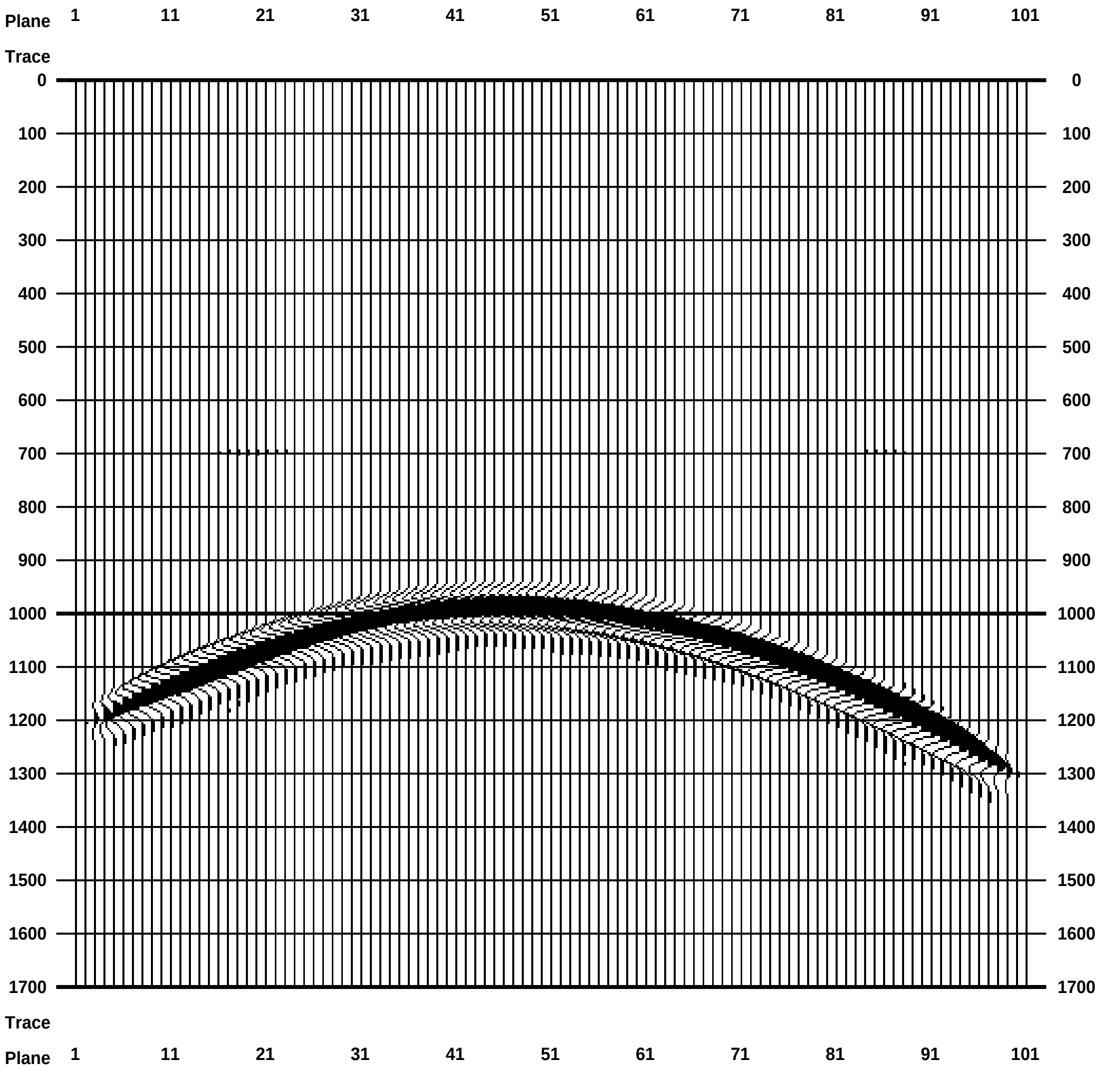

FIGURE 3B: Transmission seismogram for configuration of Figure 2a, areal source as in Figure 3a.

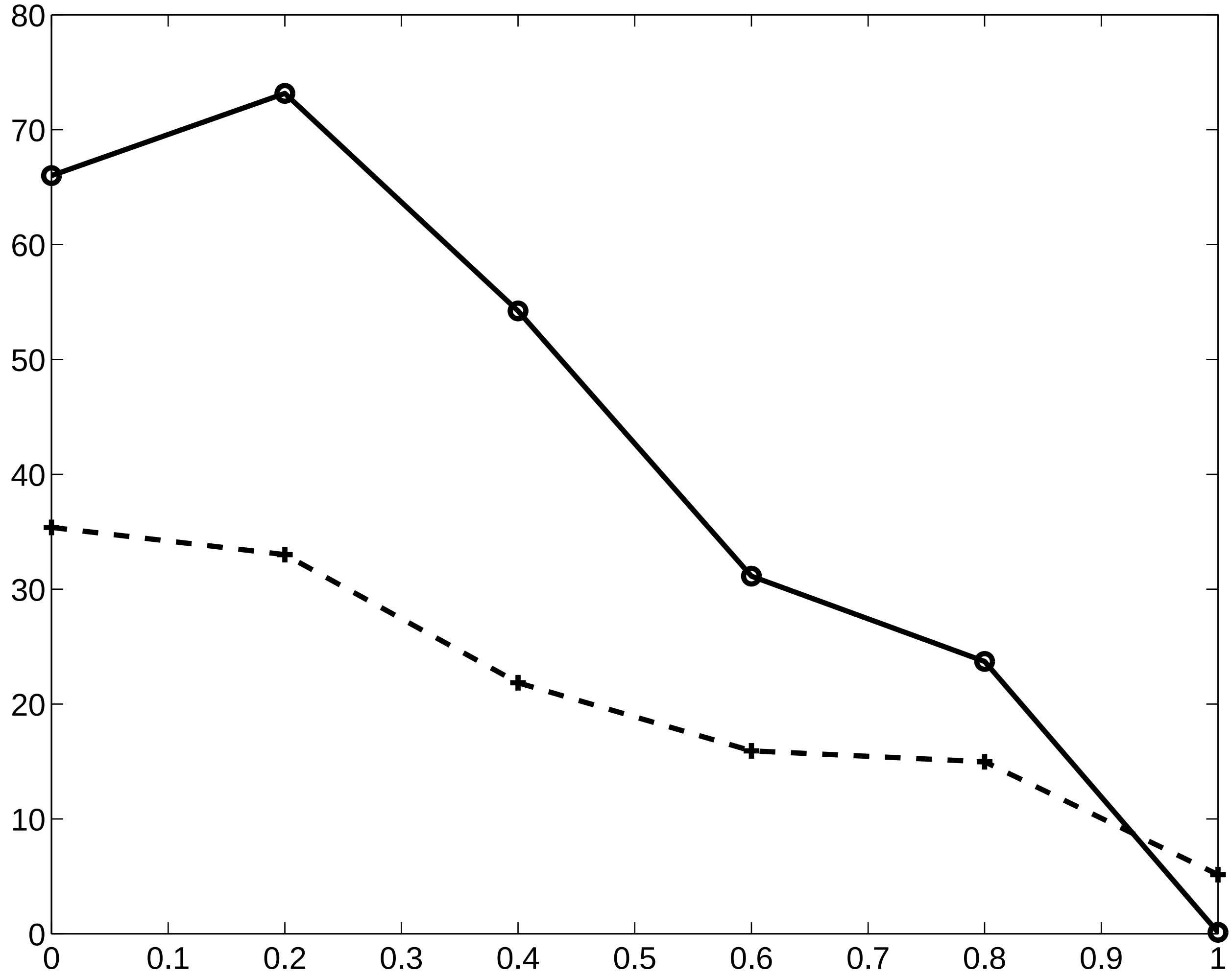

FIGURE 3C: Plot of areal source differential semblance along line segment $(1-h)v_0 + hv^*$, for $h = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0$. Here $v_0 \equiv 2$km/s, and v^* is the velocity of Figure 2a. For the solid line $\sigma = 1.0$, for the dashed line $\sigma = 0.1$. The $\sigma = 0.1$ curve has been scaled by a factor of 20 to render the curves visually comparable.

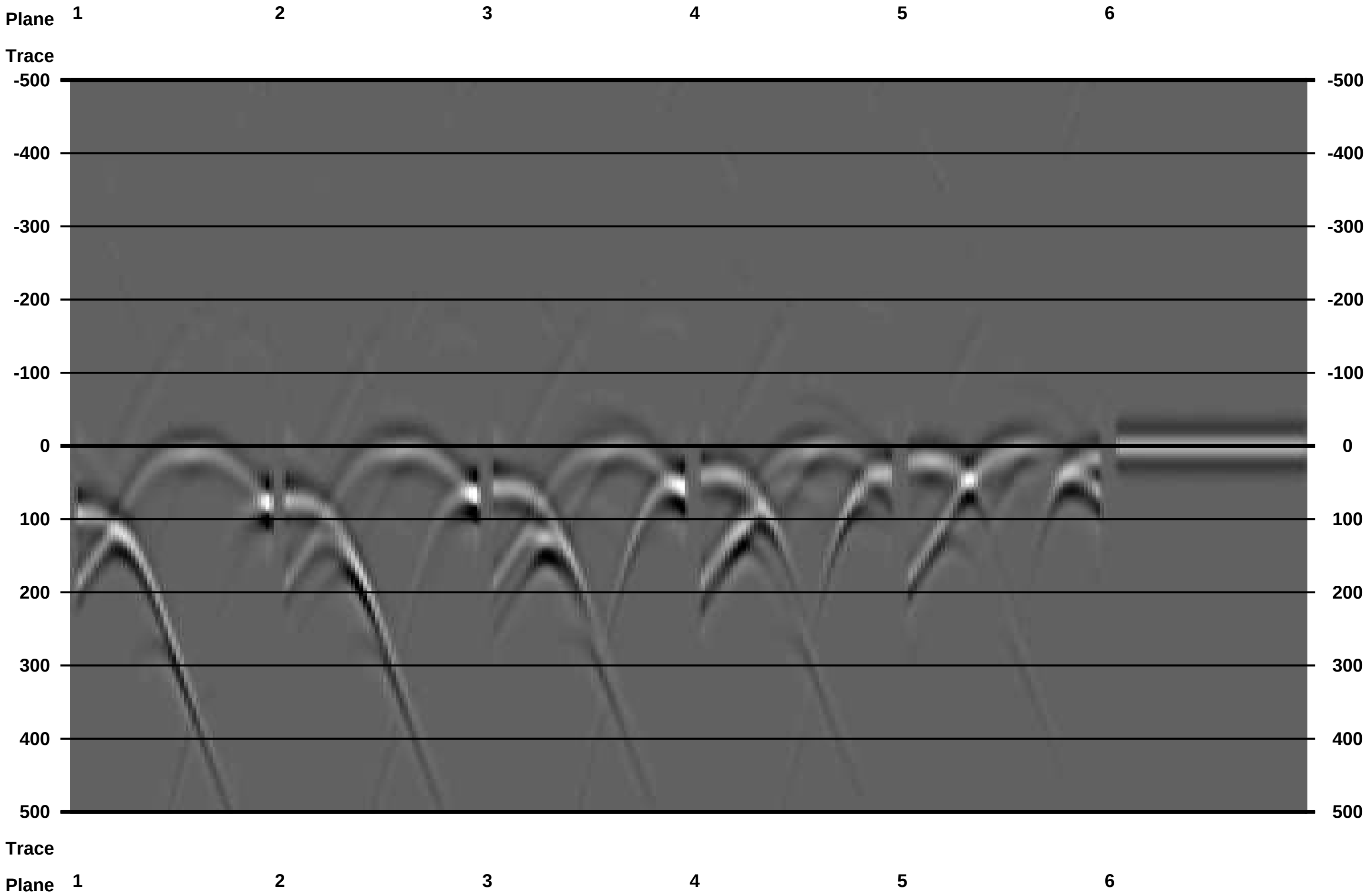

FIGURE 3D: A coherency display of the inverted sources for velocities on the line segment $(1-h)v_0 + hv^*$, for $h = 0.0, 0.2, 0.4, 0.6, 0.8, 1.0$, using areal source differential semblance. Here $v_0 \equiv 2\text{km/s}$, and v^* is the velocity of Figure 2a. Note that dipping events steadily lose strength and flatten from panel to panel.