Appendix A

Fine points of CDR picking

Methods of picking CDR parameters are based on empirical observations more than upon theory. I have developed one such method; it has been tested on a single set of seismic data. This data set consisted of marine seismic observations provided courtesy of British Petroleum. Indispensable pre-processing of the data was performed by Paul Fowler at Stanford, and I thank him for allowing me to use the results of his work.

A.1 Slant stacks and picking

In the CDR method, adjacent traces are slant stacked and then picked. Soviet researchers have done extensive work on the problem of automatic picking of slant stacks (Rapoport, 1977). I have used some of their results, but for the most part I have developed my own approaches. The Soviet methods were developed to work optimally in an environment where the computers have less internal memory and less disk storage.

A.1.1 Slant stack interpolation

A set of traces $u_i(t)$, consisting of $2n_{\text{bh}}+1$ traces, is slant stacked according to the equation

$$r_j(t) = \sum_{i=-n_{\text{bh}}}^{n_{\text{bh}}} u_i(t + ij\,\Delta p\,\Delta x), \tag{A.1}$$

where i and j are the subscripts representing trace location and ray parameter, respectively, and Δx and Δp are the respective intervals of trace location and ray parameter. See Chapter 2 for more details.

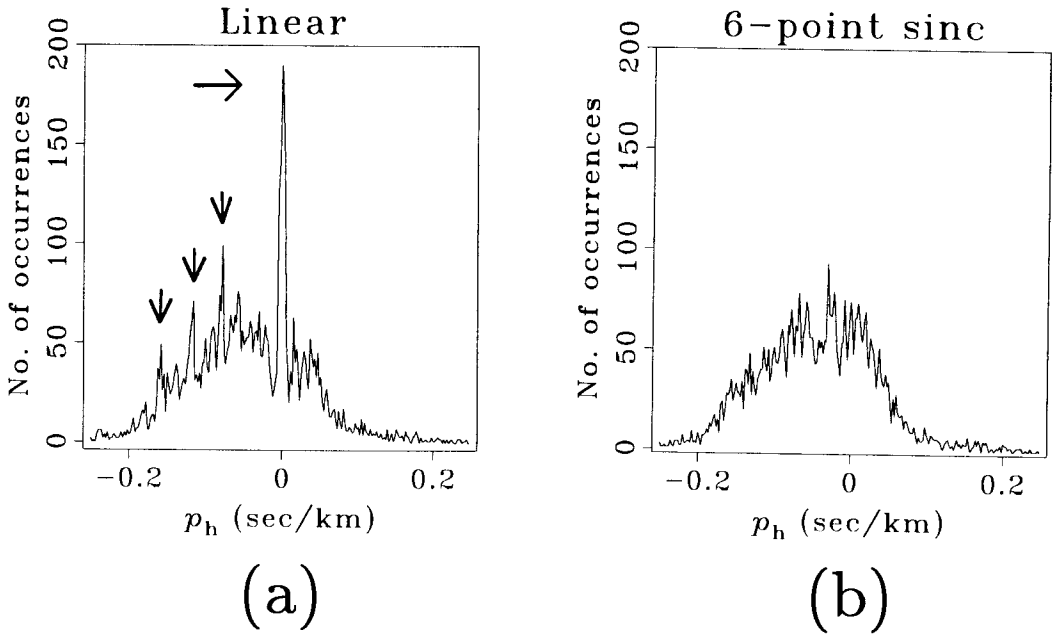

Figure A.1: Biases caused by linear interpolation. Histograms of ray-parameter picks (p_h), made after slant stacking the British Petroleum marine data set, are shown. Diagram (a) shows the result of using a linear interpolator in the slant stacking. Diagram (b) shows the result of using a six-point truncated-sinc interpolator. The presence in Diagram (a) of regularly spaced peaks (labeled by arrows) suggests that linear interpolation biases the picks towards certain values. The binning interval is .002 sec/km.

Equation (A.1) is based on the assumption that the traces are continuous functions of time. Real seismic data, however, is discretely sampled at intervals of Δt. A discrete series u_{il}, where i is the trace number and l is the time sample, can be converted to a continuous function $u_i(t)$ according to the formula

$$u_i(t) = \sum_l u_{il} f(t - l\,\Delta t), \tag{A.2}$$

where $f(t)$ is an interpolation function and Δt is the discretization in time. A typical interpolation function is the linear interpolator:

$$f(t) = \begin{cases} 1 - |t|/\Delta t & \text{if } -\Delta t \le t \le \Delta t, \\ 0 & \text{otherwise.} \end{cases} \tag{A.3}$$

To include interpolation, equation (A.1) is rewritten as

$$r_{jk} = \sum_{i=-n_{\rm bh}}^{n_{\rm bh}} \sum_l f(k\,\Delta t + ij\,\Delta p\,\Delta x - l\,\Delta t), \tag{A.4}$$

where k represents the time sample, and $f(t)$ is the interpolation function.

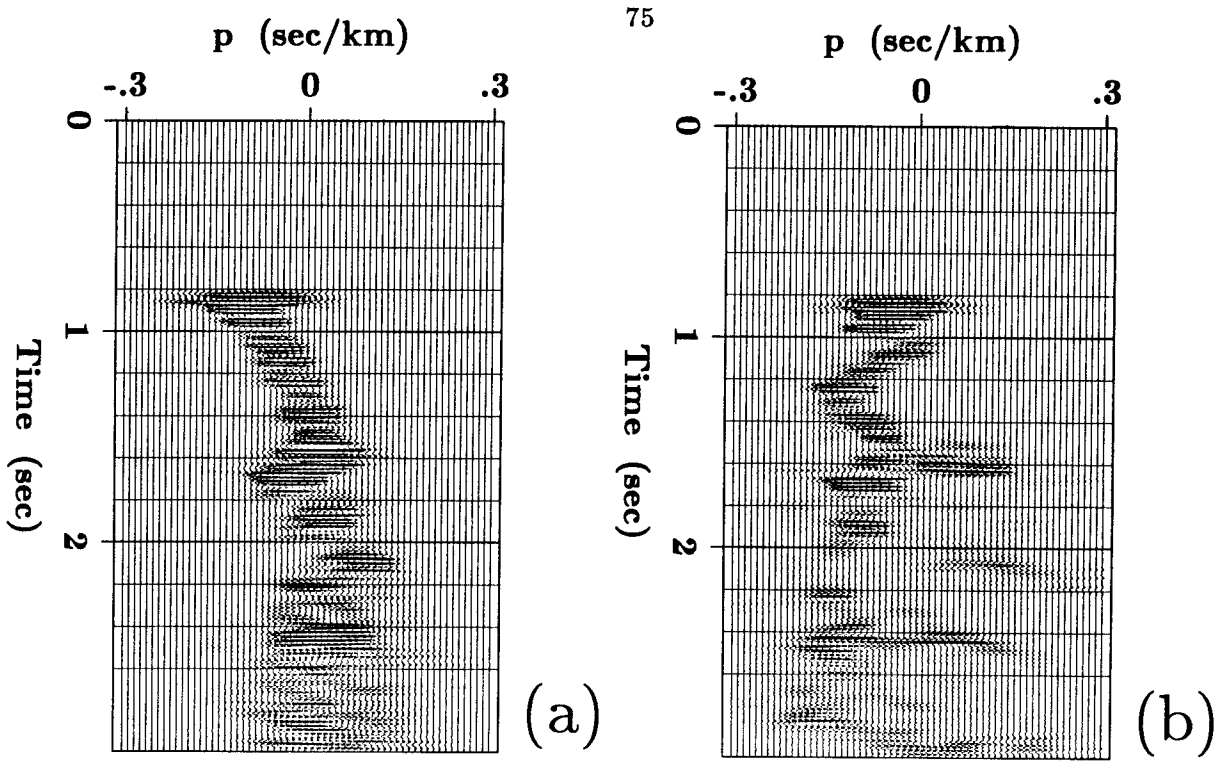

Figure A.2: Slant-stack sections. Diagrams (a) and (b) show the result of slant stacking over reciprocal gathers of field data. Diagram (a) represents a slant stack over different values of offset; Diagram (b) represents a slant stack over different values of midpoint.

Tests have shown that if linear interpolation is used, biases arise when the ray parameter p is picked. This effect is seen in Figure A.1a, which shows a histogram of picked ray parameters (the picking algorithm is described below). Note the peaks at regular intervals. Field data would not be expected to have such peaks; they are therefore artifacts. Decreasing the ray-parameter interval Δp does not remove these artifacts. A six-point truncated-sinc interpolation function effectively removes them, as is seen in Figure A.1b. A four-point interpolation function would probably be adequate.

A.1.2 Automatic picking

Once the slant-stack sections have been generated, they must be picked. To illustrate the picking process, I shall use a field-data example. Figure A.2 shows two slant-stack sections from a subset of the British Petroleum marine data. Figure A.2a shows the result of slant stacking over different values of offset, and Figure A.2b shows the result of slant stacking over different values of midpoint.

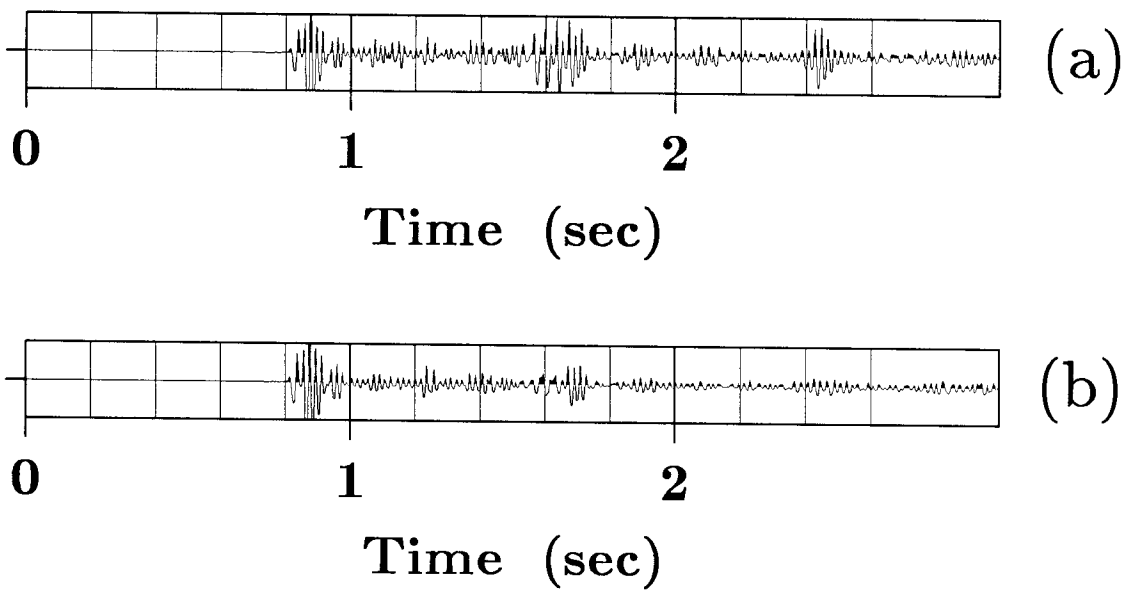

Figure A.3: Maximum traces. The maximum trace values at each time, from the panels in Figures A.2a and A.2b respectively, are shown.

Maximum traces

The picking process begins with finding the maximum amplitude at each time in each slant-stack section. This is equivalent to finding the maximum value along every horizontal line in Figures A.2a and A.2b. The two "maximum traces", from Figures A.2a and A.2b respectively, are shown in Figures A.3a and A.3b; they are designated $s_s(t)$ and $s_g(t)$.

Correlation

The two maximum traces, $s_s(t)$ and $s_g(t)$, can be combined to give a correlated-maximum trace. Such a trace should have large amplitudes wherever $s_s(t)$ and $s_g(t)$ are correlated, and low amplitudes wherever they are not.

One measure of the correlation is the quantity $s_s(t) - s_g(t)$. When this quantity is near zero, the traces are correlated; otherwise they are not (this quantity is not a good measure if the traces differ only by a multiplicative constant). I use the following equation for $q(t)$,

the correlated-maximum trace:

$$q(t) = \text{amax}(s_s(t),\, s_g(t)) \exp\left(-\frac{1}{\sigma_q^2} \left[\frac{s_s(t) - s_g(t)}{\frac{1}{2}(s_s(t) + s_g(t))} \right]^2 \right), \tag{A.5}$$

where the value of σ_q is selected by the user (usually on a trial-and-error basis); amax(x, y) is a function that returns x if $|x| > |y|$, and returns y otherwise.

Picking and interpolation

The correlated-maximum trace, $q(t)$, is the trace that is first picked. The maximum value $q_{\max}$, and corresponding traveltime $t_{\max}$, are found on $q(t)$. Then, the slant-stack sections are examined to find the respective points of maximum amplitude at time $t_{\max}$. The positions of these maxima give the picked values of p on the two sections.

Recall that the slant-stack sections are discretized in p and t. Once the discrete maximum-amplitude p and t values are found, quadratic interpolations in p and t are performed to find the true position of the points of maximum amplitude.

Near-time windowing

It is undesirable to pick peaks that are too close together. After a peak has been picked at $t_{\max}$, I therefore window the surrounding region on the correlated-maximum trace:

$$q(t) = \begin{cases} q'(t)\, |t_{\max} - t| / w_h & \text{if } -w_h \le t_{\max} - t \le w_h, \\ q'(t) & \text{otherwise;} \end{cases} \tag{A.6}$$

where $q'(t)$ is the unwindowed correlated-maximum trace, and w_h, a user-selected value, is the half-width of the triangular window.

Bad picks

The picking process described here does a good job of finding peaks (maxima) on the slant-stack sections. These sections, however, may contain non-valid maxima. The main cause of non-valid maxima, other than a high noise level, is spatial aliasing; spatial aliasing arises when the distance between adjacent shots or receivers is greater than some fraction of the wavelength of the recorded wave. Semblance weighting helps reduce the magnitude of spatially aliased peaks, but inevitably, some of these peaks are picked anyway. These bad picks (or rather, good picks on bad peaks) can often be eliminated on the basis of CDR

velocity, amplitude, and dip. This topic is covered more fully in section 2.4.2 (page 33) and in section A.2.2.

A.2 A marine data set

The picking procedure was tested on a marine data set provided by British Petroleum, with vital pre-processing work performed by Paul Fowler. A near-offset section of this data is shown in Figure 2.10 (page 32). There were 300 shot points, stationed 16.6 meters apart. Each shot was recorded on a 36-receiver cable, with geophone arrays at 33.3 meter intervals. Each trace contained 750 samples, with a sampling interval of .004 seconds. A time-dependent gain was applied to each trace (the gain function equaled $t^{1.5}$); a normal-moveout correction was applied corresponding to a constant velocity of 1.6 km/sec (on page 14 of section 2.2.1 is a discussion of the reasons for applying such a correction).

A.2.1 CDR and picking parameters

There are a number of parameters that need to be specified in order to pick successfully a particular set of seismic data. These parameters might be divided into three categories: trace-selection parameters, slant-stack parameters, and picking parameters. Their values in this thesis were chosen by trial-and-error methods.

Trace selection

A useful concept is that of the "summation base", defined as the set of traces used in any particular slant stack. In the CDR method, slant stacking is performed over a pair of summation bases, and the reciprocal parameters are picked. For marine data, the most logical approach is to pick p_y and p_h (these two quantities were defined in section 2.2.1, page 12). As shown in Figure 2.2 (page 13), the pair of summation bases used to pick p_h and p_y form an X-shaped figure.

I used a summation base consisting of 13 adjacent traces. Figure 2.4a shows the centers of the summation bases in a subset of the marine data (86 shot profiles out of 300). Figure A.4b shows some of the summation bases themselves, as well as the centers shown in the previous figure. Figure A.4c shows all the summation bases. These diagrams give an idea of the sampling density used.

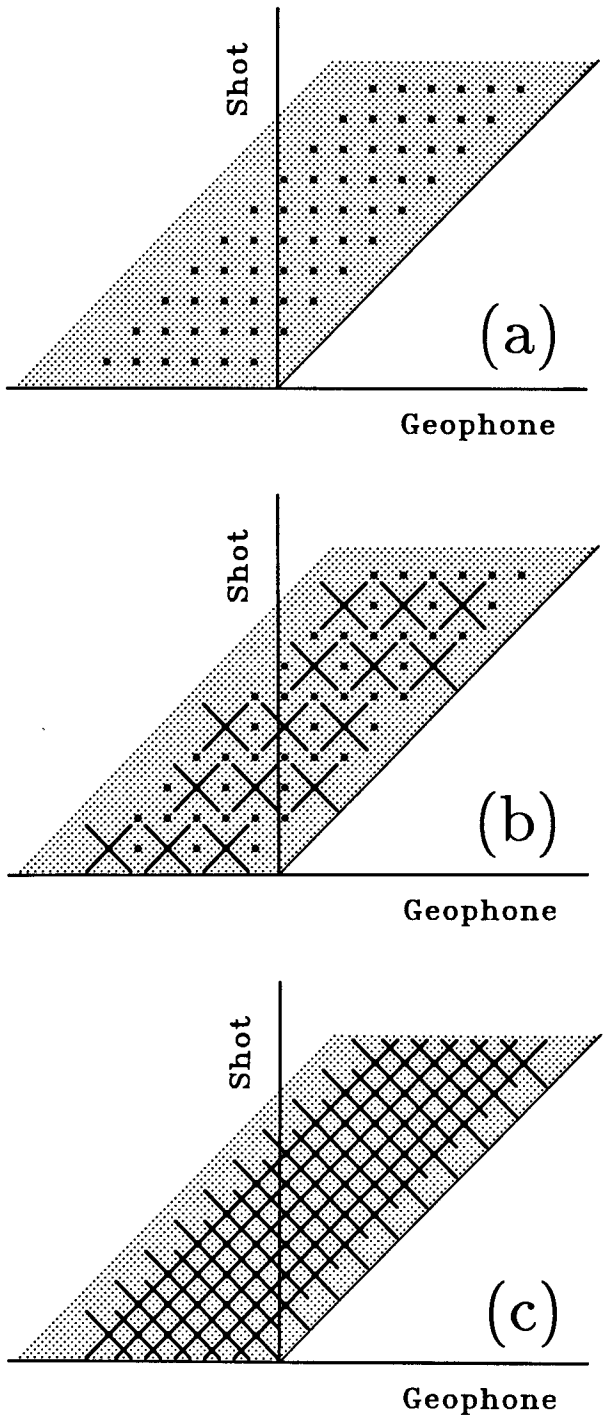

Figure A.4: Summation bases. Diagram (a) is a stacking chart of a subsection of the marine field-data example, with the locations of the centers of the summation bases denoted by large dots. The small dots represent traces. Diagram (b) is the same as diagram (a), but with some of the summation bases represented by heavy lines. Diagram (c) shows all of the summation bases.

Slant-stack parameters

Once a gather of traces has been selected, it is slant stacked over a range of p values. I chose the range by looking at slant stacks, and deciding which p values contained most of the geologically significant reflectors without including too many aliased events from the strong water-bottom reflections. I allowed p to vary from -0.8 to 0.8 sec/km. I also selected a value for Δp, the distance between adjacent values of p. One can use theoretical approaches in order to choose Δp. For instance, one could make use of formulas that relate the average wavelength of the signal and the width of the summation base to the optimum value of Δp. I used a value, .01 sec/km, much smaller than the optimum. As a result, the slant stack was, in a theoretical sense, oversampled in p. I felt this oversampling was necessary, since in subsequent processing I effectively used nearest-neighbor interpolation between adjacent p traces. As mentioned in section 2.2.1 (page 12), I weighted the slant stack by a smoothed semblance-based weighting function. The smoothing is carried out in the time domain; a boxcar smoothing function is used. The length of this boxcar function is another user-selected parameter. I used a boxcar function that was 21 points long.

Picking parameters

Several user-selected parameters play a role during the picking of events on the slant stack. Two of these parameters are discussed in section A.1.2. In picking the marine data I set σ_q (equation (A.5)) to .8, and w_h (equation (A.6)) to .05 seconds. Another user-selected parameter is the percentile value; this value determines which peaks on each $q(t)$ trace are considered interesting. For the marine data, only peaks larger than the 95th-percentile amplitude of $q(t)$ were picked.

A.2.2 Filtering parameters

After the reciprocal parameters were picked, they were filtered to remove multiple reflections and spurious events. Figure 2.7 (page 29) shows the picked data before filtering (some low-amplitude events have not been plotted); there were 6,507 sets of reciprocal parameters. A filtering program was used that performs v_{CDR}-based time migration (section 2.3.3, page 28) and eliminates picks on the basis of their migrated position and other criteria. First, all picks were filtered out that had less than a certain amplitude (about 25% of the picks). All picks were eliminated that had shots or geophones outside the

region of the velocity analysis, as were all picks whose migrated dip bars lay deeper than 2.5 seconds. I drew a line just above the first water-bottom multiple; all picks below that line were eliminated that had a value of v_{CDR} less than 1.65 km/sec. I drew a line just above the second water-bottom multiple; all picks below that line were eliminated that had a value of v_{CDR} less than 1.9 km/sec. Finally, some spurious picks in a certain region were eliminated that had dips greater than .5 sec/km on the migrated time section. These filtering operations would be difficult to perform on conventional seismic data, but were easily carried out on the picked data. After the filtering 3,244 sets of reciprocal parameters remained. Figure 2.12 (page 33) shows the data after filtering.

A.2.3 Computation time

Considering the size of the data set (10,800 traces, with 750 samples per trace), the slant-stacking and picking operations did not require an inordinate amount of computer time: only 30 minutes of CPU time on a Convex C-1 computer. I did not determine how much of this time was spent performing the slant stacks, and how much was spent doing the picking. The filtering was less expensive. Four different filtering programs were used, and each took about 2 minutes of CPU time. Most of the time in each program was spent performing v_{CDR} time migration. If the four programs had been consolidated, this migration would have been performed only once. I did not see any need for such a consolidation, however, given the small amount of CPU time that would be saved.

Appendix B

The line search

An important element of the Gauss-Newton optimization method is the line search. At each Gauss-Newton iteration step, a descent distance and direction are found by solving a linearized least-squares problem. Since the objective function is non-linear, the descent distance is only an estimate. It is necessary to try several descent distances along the descent direction, in order to find the one that best minimizes the objective function. The *line search* (Luenberger, 1984) is the term for the procedure used to find this distance.

The line search can be expressed in mathematical terms. Given a velocity model $\mathbf{v}$, a descent direction $\Delta\mathbf{v}$ is found by following the procedures outlined in equations (3.8), (3.9), and (3.11). The task, then, is to find the value of α, a scalar, that minimizes $\Phi(\mathbf{v}^{(k)}+\alpha\,\Delta\mathbf{v})$, where Φ is the objective function. Luenberger (1984) gives several techniques for determining α; I have chosen to develop my own variant, suitable for the idiosyncrasies of the CDR tomographic inversion method. Recall from equation (3.5) that Φ depends on $\mathbf{x}_{\text{err}}$, which is in turn non-linearly dependent on the velocity model and the picked parameters. The only way to determine Φ for a given velocity model is by ray tracing. Every time Φ is evaluated, therefore, rays must be traced for each set of reciprocal parameter. Such ray tracing is expensive, and should be done as seldom as possible.

I minimize the number of ray-tracing steps by evaluating Φ for three values of α. A fourth value, at $\alpha = 0.$, is already known. I fit a cubic curve of Φ as a function of α through these four points, and find the point on this curve where the minimum value of Φ should lie. This point gives me a new trial value of α. I evaluate Φ at this new point, fit a cubic curve through the new point and its nearest neighbors, and find a new minimum. This process continues until the new minimum value of Φ is greater than .99 times the

previous minimum.

The curve-fitting procedure can be described mathematically. If Φ were exactly a quadratic function of the velocity model, then α would equal 1.0. I therefore look for values of α in this neighborhood. Four initial trial values of α are used: 0.0, 0.4, 0.8, and 1.2. The corresponding values of Φ are found through ray tracing. These values are used to find the coefficients a, b, c, and d in the equation

$$\Phi = a\alpha^3 + b\alpha^2 + c\alpha + d. \tag{B.1}$$

It can be shown that the minimum value of Φ in equation (B.1) lies at

$$\alpha = \frac{-c}{b + \sqrt{b^2 - 3ac}}. \tag{B.2}$$

This new value of α is used to refine the position of the minimum. The line search continues until two successive values of Φ differ by less than 1%.

Appendix C

Ray tracing in a gridded velocity field

In my tomographic inversion process the velocity field is represented by a grid. The grid can be either laterally homogeneous (for layered media), or laterally heterogeneous. I have developed ray-tracing methods for both of these cases.

C.1 Ray tracing in a laterally homogeneous medium

Ray tracing in a laterally homogeneous medium is relatively easy. If p is the ray parameter and z is depth, then

$$\sin\theta(z) = p\,v(z), \tag{C.1}$$

where $\theta(z)$ is the angle, measured from the vertical, of the ray at depth z, and $v(z)$ is the velocity of the medium. The ray parameter p is constant in a horizontally layered medium. If a particular layer has a thickness Δz, then the ray will travel within that layer a horizontal distance (defined as Δx) of

$$\Delta x = \Delta z \tan\theta(z). \tag{C.2}$$

Similarly, travel time (Δt) through that layer is

$$\Delta t = \frac{\Delta z}{v(z)\cos\theta(z)}. \tag{C.3}$$

It is assumed that velocity v (and thus θ as well) remains constant over the distance Δz. These three equations, plus some well-known trig identities ($\cos\theta = \sqrt{1-\sin^2\theta}$,

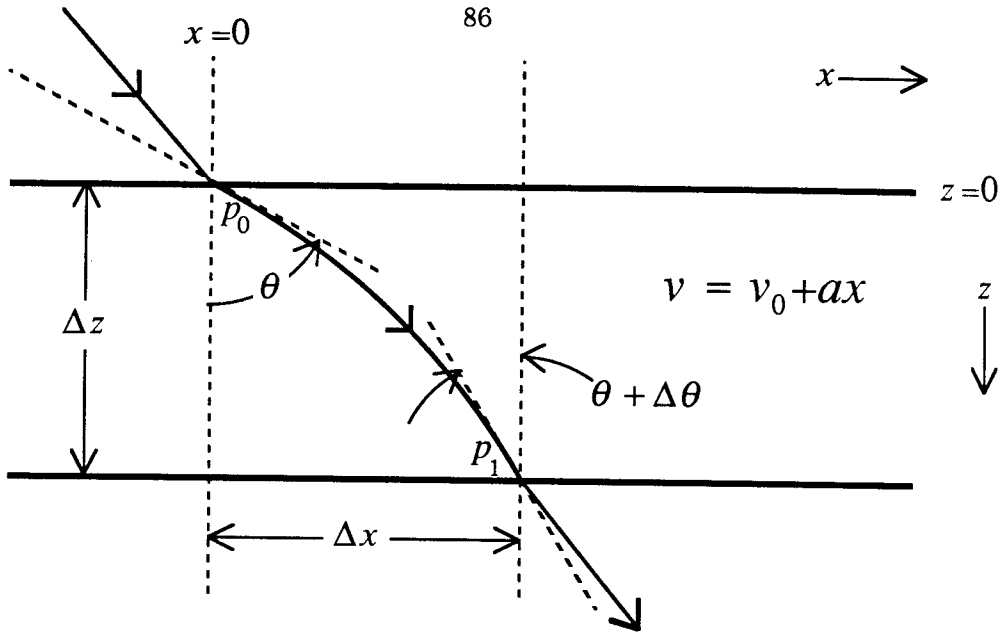

Figure C.1: Ray-tracing notation. This figure shows the notation used in the text. The figure is drawn so that the values of all parameters except $\Delta\theta$ are positive.

for instance) are enough to find x_{err}, given a particular layered velocity model and the parameters x_{s}, x_{g}, p_{s}, p_{g}, and t.

C.2 Ray tracing in a laterally heterogeneous medium

When velocity varies only vertically, so that $v = v(z)$, it is possible to discretize the velocity model in terms of layers. When velocity varies horizontally as well, however, it is necessary to discretize the model into boxes. I allow a horizontal velocity gradient to exist within each box. If v_i is the velocity associated with a particular box, a_i is the value of the associated velocity gradient, and x_i is the horizontal position of the center of the box, then

$$v(x,z) = v_i + a_i(x - x_i) \tag{C.4}$$

within the box. The values of a_i are chosen to make $v(x,z)$ a continuous (but not continuously differentiable) function in x, even at the boundaries between horizontally adjacent boxes. Note, however, that $v(x,z)$ will not be a continuous function in z, just as it was not a continuous function in z in the horizontally layered case.

Tracing rays is, of course, more complicated in the presence of a horizontal gradient. Suppose a ray, with ray parameter p_0, crosses into a new box at $x = 0$, $z = 0$ (see Figure C.1). Let the velocity at that point be v_0, let the velocity within the box be

described by the formula $v_0 + ax$, and define $k \equiv -a/v_0$. Then

$$\sin\theta = v_0 p_0, \tag{C.5}$$

where θ is the angle of the ray from the vertical immediately *after* it has entered the box. Suppose the ray emerges from the box at the point $x = \Delta x$, $z = \Delta z$, and that just before it emerges, it has an angle from the vertical of $\theta + \Delta\theta$. The following two formulas adapted from Bishop et al. (1985) then hold:

$$\cos(\theta + \Delta\theta) = (1 - k\,\Delta x)\cos\theta, \tag{C.6}$$

and

$$\sin(\theta + \Delta\theta) = \sin\theta + k\,\Delta z\cos\theta. \tag{C.7}$$

Once $\sin\theta$ has been determined from equation (C.5), $\cos\theta$ may be obtained from a well-known trig identity, and Δz is known (it is the thickness of the layer, if the ray travels all the way from the top of the layer to the bottom). Then, $\sin(\theta + \Delta\theta)$ may be obtained from equation (C.7), $\cos(\theta + \Delta\theta)$ obtained by a trig identity, and equation (C.6) solved to yield

$$\Delta x = \frac{1}{k}\left(1 - \frac{\cos(\theta + \Delta\theta)}{\cos\theta}\right). \tag{C.8}$$

Note that when k is close to zero, equation (C.2) should be used in place of this equation. If Δx is known, and Δz is the unknown quantity—this could happen if the ray intersects the vertical boundary between two horizontally adjacent boxes—equations (C.6) and (C.7) can be solved to yield a similar formula for Δz as a function of Δx. I will not give that formula here.

Another important parameter that must be determined is p_1, the ray parameter at the point of emergence. It is easily shown that $\sin(\theta + \Delta\theta) = (v_0 + a\,\Delta x)p_1$, and this equation, with equation (C.6) and the definition of k, can be solved to yield

$$p_1 = \frac{\cos\theta}{v_0}\tan(\theta + \Delta\theta). \tag{C.9}$$

Since velocity discontinuities occur only with changes in z (recall that the gradients a_i are chosen so as to make this statement true), p_1 remains unchanged as the ray crosses the interface into the next box. It can therefore be used as the input ray parameter, p_0, when the ray is traced through this next box.

It is worth noting that nowhere, so far, has it been necessary to evaluate a transcendental function. That is, the ray tracing can be carried out without the computer ever needing

to evaluate a sine, cosine, arcsine, or arccosine function. The only function (besides the four basic arithmetic functions) that needs to be called by the ray-tracing program is the square root function (to solve such identities as $\sin^2\theta + \cos^2\theta = 1$). The calculation of the traveltime, however, will prove to be an exception.

The traveltime, Δt, is the amount of time it take the ray to travel through a particular box. This quantity is determined from the integral

$$\Delta t = \int_s \frac{ds}{v(s)}, \tag{C.10}$$

where s represents the path that the ray follows within the box. This integral can be transformed algebraically into the explicit formula

$$\Delta t = \frac{1}{kv} \log\left(\frac{\cos\theta\,(1+\sin(\theta+\Delta\theta))}{\cos(\theta+\Delta\theta)\,(1+\sin\theta)}\right). \tag{C.11}$$

The ray-tracing program must therefore evaluate a logarithm. When k is close to zero, it is preferable to use equation (C.3) in place of equation (C.11).

Appendix D

Determining the Fréchet matrix

The Fréchet matrix is relatively easy to determine when velocity is a function of depth only. It is more difficult to determine when velocity varies laterally as well as vertically.

D.1 The Fréchet matrix when $v = v(z)$

Recall from equation (3.8) that $\underset{\sim}{\mathbf{A}}^k$, the Fréchet matrix at the kth iteration, is defined according to the formula

$$A_{ij}^k \equiv \left(\frac{\partial x_{\mathrm{err}(j)}}{\partial v_i} \right)_{\mathbf{v}=\mathbf{v}^k}. \tag{D.1}$$

Further definitions are necessary: Δz_i is the vertical distance that the jth ray travels in the ith layer (here, as in later definitions, the j subscript has been omitted for the sake of clarity); Δx_i is the horizontal distance that this ray travels in the ith layer; and Δt_i is the travel time of this ray within the layer. Then, according to the chain rule,

$$\frac{\partial x_{\mathrm{err}(j)}}{\partial v_i} = \frac{\partial x_{\mathrm{err}(j)}}{\partial \Delta x_i} \cdot \frac{d\Delta x_i}{dv_i} + \frac{\partial x_{\mathrm{err}(j)}}{\partial \Delta t_i} \cdot \frac{d\Delta t_i}{dv_i}. \tag{D.2}$$

The problem is now reduced to one of finding the value of each term.

Before continuing, there is a point that should be clarified. Recall that Δx_i was defined to be the horizontal distance traveled by the jth ray in the ith layer. This definition does not take into account the fact that two rays are traced simultaneously to find $x_{\mathrm{err}(j)}$, and that $x_{\mathrm{err}(j)}$ is thus dependent on both of these rays. That is,

$$x_{\mathrm{err}} = x_{\mathrm{ge}} - x_{\mathrm{se}}, \tag{D.3}$$

where x_{se} is the horizontal position of the endpoint of the jth down-going (shot) ray and x_{ge} is the horizontal position of the endpoint of the jth up-going (geophone) ray. (The endpoint of a ray is defined to be the point where the ray intersects the line $z = z_e$, with z_e defined as the depth where the combined computed travel times of the two rays equals the measured travel time.) A velocity change in a given layer will thus produce changes in both rays. It is most convenient, however, to consider the effect of the change on only one ray at a time, and later sum the results into A_{ij}. This is the approach that will be used here.

The first term in equation (D.2) is easily determined. Since v is a function only of z, x_{se} (or x_{ge}) varies exactly the same amount as Δx_i varies, independently of the velocity structure of the model. As a consequence of this result and of equation (D.3), and depending on which ray is under consideration,

$$\frac{\partial x_{\mathrm{err}(j)}}{\partial \Delta x_i} = -1 \quad \text{(down-going)}; \qquad \text{otherwise,} \quad \frac{\partial x_{\mathrm{err}(j)}}{\partial \Delta x_i} = 1 \quad \text{(up-going)}. \tag{D.4}$$

The second term in equation (D.2) can be determined by differentiating equation (C.2). The result is:

$$\frac{d\Delta x_i}{dv_i} = \frac{p\,\Delta z_i}{\cos^3 \theta_i}, \tag{D.5}$$

where θ_i is the angle of the ray in layer i as measured from the vertical.

The third term in equation (D.2) is somewhat more complicated to derive. To begin with,

$$\frac{\partial x_{\mathrm{err}(j)}}{\partial \Delta t_i} = \frac{dx_{\mathrm{err}}}{dz_e} \cdot \frac{dz_e}{d\Delta t_i}, \tag{D.6}$$

where z_e is the depth of the endpoints of the jth pair of rays. From equation (D.3),

$$\frac{dx_{\mathrm{err}}}{dz_e} = \frac{dx_{ge}}{dz_e} - \frac{dx_{se}}{dz_e}, \tag{D.7}$$

which through simple trigonometry becomes

$$\frac{dx_{\mathrm{err}}}{dz_e} = \tan\theta_{ge} - \tan\theta_{se}, \tag{D.8}$$

where θ_{ge} and θ_{se} are the angles at depth z_e of the up-going and down-going rays respectively. These angles can be derived from p_s and p_g through the rule that $vp = \sin\theta$.

The second part of equation (D.6) must be approached indirectly. First, note that even though travel time Δt_i within a layer may vary, the overall travel time t must remain constant. Thus, as Δt_i increases, $t - \Delta t_i$ decreases. As a result,

$$\frac{dz_e}{d\Delta t_i} = -\frac{dz_e}{dt}. \tag{D.9}$$

Next, it is useful to note that

$$\frac{dz_e}{dt} = \frac{1}{dt/dz_e}. \tag{D.10}$$

Then

$$\frac{dt}{dz_e} = \frac{dt_s}{dz_e} + \frac{dt_g}{dz_e}, \tag{D.11}$$

where t_s and t_g are the total travel times of the down-going (shot) ray and the up-going (geophone) ray, respectively. An application of the chain rule yields

$$\frac{dt}{dz_e} = \frac{dt_s}{dl_s}\frac{dl_s}{dz_e} + \frac{dt_g}{dl_g}\frac{dl_g}{dz_e}, \tag{D.12}$$

where l_s and l_g are the total path lengths of the down-going (shot) ray and the up-going (geophone) ray, respectively. Through the application of some trigonometry, and thanks to the relationship $v = dl/dt$,

$$\frac{dt}{dz_e} = \frac{1}{v_{se}}\frac{1}{\cos\theta_{se}} + \frac{1}{v_{ge}}\frac{1}{\cos\theta_{ge}}, \tag{D.13}$$

where $v_{se} \equiv v(x_{se}, z_e)$, and $v_{ge} \equiv v(x_{ge}, z_e)$.

Now it is possible to combine equations (D.6), (D.8), (D.9), (D.10), and (D.13) and write:

$$\frac{\partial x_{err(j)}}{\partial \Delta t_i} = -v_e \frac{\sin\theta_{ge}\cos\theta_{se} - \sin\theta_{se}\cos\theta_{ge}}{\cos\theta_{se} + \cos\theta_{ge}}, \tag{D.14}$$

where $v_e \equiv v(z_e)$.

The fourth term in equation (D.2) can be found by differentiating equation (C.3), with the result:

$$\frac{d\Delta t_i}{dv_i} = -\frac{\Delta t_i}{v_i} + \frac{p\,\Delta z_i \sin\theta_i}{v_i \cos^3\theta_i}. \tag{D.15}$$

The final result, found by substituting equations (D.4), (D.5), (D.14) and (D.15) into equations (D.1) and (D.2), is:

$$A_{ij}^k = \pm\frac{p\,\Delta z_i}{\cos^3\theta_i} - v_e \frac{\sin\theta_{ge}\cos\theta_{se} - \sin\theta_{se}\cos\theta_{ge}}{\cos\theta_{se} + \cos\theta_{ge}}\left(-\frac{\Delta t_i}{v_i} + \frac{p\,\Delta z_i \sin\theta_i}{v_i \cos^3\theta_i}\right). \tag{D.16}$$

The $\pm$ sign is positive when the ray under consideration is the up-going (geophone) ray; otherwise it is negative.

D.2 The Fréchet matrix when $v = v(x, z)$

When the medium is both laterally and vertically heterogeneous, calculation of the Fréchet matrix is more complicated, but not unduly so.

Several different approaches to this calculation can be taken. The simplest approach is to perturb each grid point in the velocity model in turn, and see how all the x_{err} values vary. This simple approach is much too expensive, since the number of rays that would have to be traced equals the number of grid points in the model times the number of sets of picked parameters. In the approach shown here, fewer perturbations are made, and the results of these perturbations are propagated by means of a transfer matrix.

To begin, an equation analogous to equation (D.2) is written:

$$\frac{\partial x_{\text{err}}}{\partial v_i} = \frac{\partial x_{\text{err}}}{\partial x_{\text{e}}} \cdot \frac{\partial x_{\text{e}}}{\partial v_i} + \frac{\partial x_{\text{err}}}{\partial t} \cdot \frac{\partial t}{\partial v_i}. \tag{D.17}$$

This equation refers to the jth ray, but the subscript j has been dropped for the sake of clarity. Here v_i is defined to be the velocity in the ith box of the model, x_{e} is the horizontal position of the endpoint of the jth ray, and t is the travel time of the jth ray from the surface to the depth z_{e}. Strictly speaking, v_i is the velocity at the ith grid point, but it is more convenient to speak in terms of boxes. As in the previous section, the endpoint of the ray is defined to be the point where the ray intersects the line $z = z_{\text{e}}$ (recall that z_{e} is the depth where the computed travel time equals the measured travel time). Likewise, attention is focused on only one of the two rays associated with the jth set of picked parameters.

Equation (D.17) gives the overall formula for determining the Fréchet matrix; now it is necessary to find the values of the individual terms of this formula.

The first term in equation (D.17) is determined by a formula similar to equation (D.4):

$$\frac{\partial x_{\text{err}}}{\partial x_{\text{e}}} = -1 \quad \text{(down-going)}; \qquad \text{otherwise,} \quad \frac{\partial x_{\text{err}}}{\partial x_{\text{e}}} = 1 \quad \text{(up-going)}. \tag{D.18}$$

The third term in equation (D.17) is analogous to equation (D.14), but without the assumption that $v_{se} = v_{ge}$:

$$\frac{\partial x_{\text{err}}}{\partial t} = -v_{se} v_{ge} \frac{\sin\theta_{ge}\cos\theta_{se} - \sin\theta_{se}\cos\theta_{ge}}{v_{se}\cos\theta_{se} + v_{ge}\cos\theta_{ge}}, \tag{D.19}$$

with the variables on the right-hand side defined as in the previous section.

The second and fourth terms of equation (D.17) are not found so easily. First, some new notation is defined, as illustrated in Figure D.1. Up until now, a specific velocity box

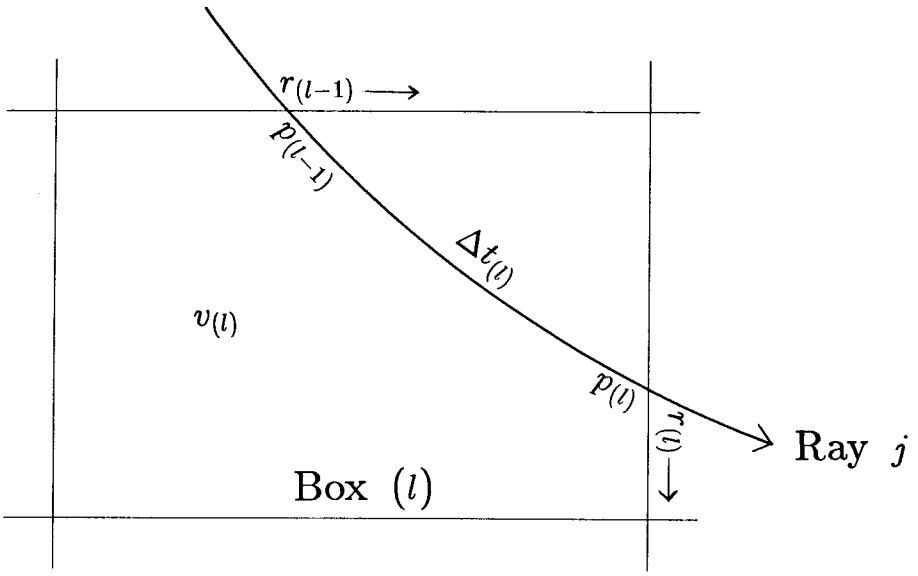

Figure D.1: Definition of terms. Ray j travels through box (l), which is the lth box along its path. It enters the box at spatial coordinate $r_{(l-1)}$ (measured parallel to the boundary where the ray enters), where it has ray parameter $p_{(l-1)}$. It leaves the box at $r_{(l)}$ (measured parallel to the boundary where the ray exits), with ray parameter $p_{(l)}$. The velocity within the box is $v_{(l)}$, and the ray takes an amount of time $\Delta t_{(l)}$ to travel through the box.

has been denoted by the subscript i, meaning that it is the ith box in the velocity model. It is useful at this point to introduce a different subscript, (l), which denotes the lth box through which ray j passes (the parentheses are meant to distinguish this new type of subscript from the previous type). For example, ray j starts at the surface, so the first box it passes through has velocity $v_{(1)}$, the second box it goes through has velocity $v_{(2)}$, and so on.

A ray entering the lth box has ray parameter $p_{(l-1)}$, and as it leaves the box it has ray parameter $p_{(l)}$. The position of the ray as it enters the box is denoted by $r_{(l-1)}$, which is a spatial coordinate measured along a certain axis. This axis is set to be parallel to the boundary that the ray crosses as it enters the box. The origin of this coordinate system is unspecified, since I will be concerned only with changes in $r_{(l-1)}$, not with its absolute value. Similarly, the position of the ray as it exits the box is denoted by $r_{(l)}$, which is measured along an axis parallel to the boundary that the ray crosses as it exits. This may seem like an odd way to measure the spatial coordinates of the ray, but it will make subsequent work easier.

The lth box has, as noted previously, a velocity $v_{(l)}$, and the ray takes an amount of time $\Delta t_{(l)}$ to travel through this box.

The box containing the endpoint of the ray may be denoted by L; by definition $r_{(L)}$ is measured along an axis parallel to the line $z = z_e$.

Now that these terms have been defined, it is possible to specify a transfer matrix $\underset{\sim}{\mathbf{T}}_{(l)}$:

$$\underset{\sim}{\mathbf{T}}_{(l)} \equiv \begin{pmatrix} \dfrac{\partial r_{(l)}}{\partial r_{(l-1)}} & \dfrac{\partial r_{(l)}}{\partial p_{(l-1)}} & 0 \\ \dfrac{\partial p_{(l)}}{\partial r_{(l-1)}} & \dfrac{\partial p_{(l)}}{\partial p_{(l-1)}} & 0 \\ \dfrac{\partial \Delta t_{(l)}}{\partial r_{(l-1)}} & \dfrac{\partial \Delta t_{(l)}}{\partial p_{(l-1)}} & 1 \end{pmatrix}. \tag{D.20}$$

The values of the terms in $\underset{\sim}{\mathbf{T}}_{(l)}$ may be determined analytically, but I prefer a finite-difference method: perturb $r_{(l-1)}$ and $p_{(l-1)}$, and see how $r_{(l)}$, $p_{(l)}$, and $\Delta t_{(l)}$ vary in response.

The next formula is essential, and I give it without proof:

$$\begin{pmatrix} \dfrac{\partial r_{(L)}}{\partial v_{(l)}} \\ \dfrac{\partial p_{(L)}}{\partial v_{(l)}} \\ \dfrac{\partial t}{\partial v_{(l)}} \end{pmatrix} = \underset{\sim}{\mathbf{T}}_{(L)} \cdot \underset{\sim}{\mathbf{T}}_{(L-1)} \cdots \underset{\sim}{\mathbf{T}}_{(l+2)} \cdot \underset{\sim}{\mathbf{T}}_{(l+1)} \cdot \begin{pmatrix} \dfrac{\partial r_{(l)}}{\partial v_{(l)}} \\ \dfrac{\partial p_{(l)}}{\partial v_{(l)}} \\ \dfrac{\partial \Delta t_{(l)}}{\partial v_{(l)}} \end{pmatrix}. \tag{D.21}$$

Be sure to distinguish between L and l in this formula. The quantities in the far right-hand vector ($\partial r_{(l)}/\partial v_{(l)}$, etc.) can be determined analytically, but again I prefer to find them through a finite-differencing method, perturbing $v_{(l)}$ and seeing the response.

Equation (D.21) gives the second and fourth terms of equation (D.17) directly, if the subscripts i and (l) refer to the same box (that is, if the ith box in the velocity model is also the lth box that ray j passes through). It is only necessary to make the identifications

$$\frac{\partial x_e}{\partial v_i} = \frac{\partial r_{(L)}}{\partial v_{(l)}} \quad \text{and} \quad \frac{\partial t}{\partial v_i} = \frac{\partial t}{\partial v_{(l)}}. \tag{D.22}$$

Bibliography

Bandurin, S.I., Bogoliubskii, A.D., Kopilevich, E.A., Sorin, V.Ia., and Ianovskii, A.K., 1982, A program for constructing seismic depth sections from CDP data on a medium-class ES computer: Exploration Geophysics, **95**, 31–39 (in Russian: Бандурин С.И., Боголюбский А.Д., Копилевич Е.А., Сорин В.Я., Яновский А.К., 1982, Математическое обеспечение построения глубинных сейсмических разрезов по данным МОГТ на ЭВМ ЕС среднего класса: Разведочная геофизика, вып. 95, с. 31–39).

Biondi, B., 1987, Interval velocity estimation from beam-stacked data: Stanford Exploration Project report SEP–51, 13–27.

Bishop, T.N., Bube, K.P., Cutler, R.T., Langan, R.T., Love, P.L., Resnick, J.R., Shuey, R.T., Spindler, D.A., and Wyld, H.W., 1985, Tomographic determination of velocity and depth in laterally varying media: Geophysics, **50**, 903–923.

Červený, V., and Pšenčík, I., 1984a, Gaussian beams in elastic 2-D laterally varying layered structures: Geophys. J. Roy. Astr. Soc., **78**, 65–91

Červený, V., and Pšenčík, I., 1984b, Program SEIS83: Numerical modelling of seismic wave fields in 2-D laterally varying layered structures by the ray method (Fortran program).

Cutler, R.T., Langan, R.T., Love, P.L., and Bishop, T.N., 1985, A tomographic solution to the travel-time problem in general inverse seismology: Advances in Geophysical Data Processing, **2**, 199–221.

Etgen, J., 1987, Velocity analysis by iterative prestack depth migration: a proposal: Stanford Exploration Project report SEP–51, 107–117.

Fowler, P., 1984, Velocity independent imaging of seismic reflectors: 54th Annual International SEG Meeting, Atlanta, Georgia.

Fowler, P., 1986, Migration velocity analysis by optimization: linear theory: 56th Annual International SEG Meeting, Houston, Texas.

Gardner, L.W., 1949, Seismograph determination of salt-dome boundary using well detector deep on dome flank: Geophysics, **14**, 29–38.

Gill, P.E., Murray, W., and Wright, M.H., 1981, Practical optimization: Academic Press Inc.

Glogovskii, V.M., Grinshpun, A.V., Meshbei, V.I., and Tseitlin, M.I., 1977, Solution of the inverse kinematic problem of reflection seismology in a layered medium, using reciprocal points: Applied Geophysics, **87**, 40–46 (in Russian: Глоговский В.М., Гринштун А.В., Мешбей В.И., Цейтлин М.И., 1977, Решение обратной кинематической задачи сейсморазведки в слоистой среде с использованием взаимных точек: Прикладная геофизика, вып. 87, с. 40–46).

Glogovskii, V.M., Meshbei, V.I., and Tseitlin, M.I., 1979, An algorithm for determining parameters of layered media using the reciprocal points of reflection seismic traveltime curves: Exploration Geophysics, **86**, 30–42 (in Russian: Глоговский В.М., Мешбей В.И., Цейтлин М.И., 1979, Алгоритм определения параметров слоистой среды по взаимным точкам годографов отраженных волн: Разведочная геофизика, вып. 86, с. 30–42).

Glogovskii, V., Meshbei, V., Tseitlin, M., and Langman, S., 1982, Kinematic-dynamic transformation of seismic traces for the determination of the velocity and depth structure of a medium, *in* Collection of reports of the second scientific seminar of COMECON countries on oil geophysics: Moscow, vol. 1, 327–331 (in Russian: Глоговский В., Мешбей В., Цейтлин М., Лангман С., 1982, Кинематико-динамическое преобразование сейсмической записи для определения скоростного и глубинного строения среды: Сборник докладов второго научного семинара стран-членов СЭВ по нефтяной геофизике, Москва, т. 1, с. 327–331).

Goldin, S.V., 1986, Seismic traveltime inversion: Society of Exploration Geophysicists.

Gray, W.C., and Golden, J.E., 1983, Velocity determination in a complex earth: 53rd Annual International SEG Meeting, Las Vegas, Nevada.

Harlan, W., and Burridge, R., 1983, A tomographic velocity inversion for unstacked data: Stanford Exploration Project report SEP–37, 1–7.

Hermont, A.J., 1979, Letter to the Editor, re: Seismic controllable directional reception as practiced in the U.S.S.R.: Geophysics, **44**, 1601–1602.

Hubral, P., and Krey, T., 1980, Interval velocities from seismic reflection time measurements: Society of Exploration Geophysicists.

Inderwiesen, P., 1986, A ray theoretical, direct inversion method for determining velocity-depth structures from seismic data: 56th Annual International SEG Meeting, Houston, Texas.

Kong, S.M., Phinney, R.A., and Chowdhury, K.R., 1985, A nonlinear signal detector for enhancement of noisy seismic record sections: Geophysics, **50**, 539–550.

Kopilevich, E.A., Sharapova, E.S., Bogoliubskii, A.D., Brodskii, A.Ia., Parasyna, V.S., and Nediliuk, L.P., 1986, The effectiveness of implementing a programming-methodological complex for the automatic interpretation of reflection seismic data on an ES computer: Exploration Geophysics, **104**, 45–52 (in Russian: Копилевич Е.А., Шарапова Е.С., Боголюбский А.Д., Бродский А.Я., Парасына В.С., Недилюк Л.П., 1986, Эффективность внедрения программно-методического комплекса автоматизированной интерпретации данных МОГТ на ЕС ЭВМ: Разведочная геофизика, вып. 104, с. 45–52).

Kostov, C., and Biondi, B., 1987, Improved resolution of slant stacks using beam stacks: Stanford Exploration Project report SEP–51, 343–350.

Kozlov, E.A., Mushin, I.A., and Tsyplakova, N.M., 1975, Construction of dynamic depth sections using the principles of the method of controlled directional reception: Applied Geophysics, **79**, 3–17 (in Russian: Козлов Е.А., Мушин И.А., Цыплакова Н.М., 1975, Построение динамических глубинных разрезов с использованием принципов метода регулируемого направленного приема: Прикладная геофизика, вып. 79, с. 3–17).

Luenberger, D.G., 1984, Linear and nonlinear programming (2nd edition): Addison-Wesley Publishing Company Inc.

Menke, W., 1984, Geophysical data analysis: Discrete inverse theory: Academic Press, Inc.

Nevinnii, A.V., and Urupov, A.K., 1976, The determination of interval velocities in media with curvilinear boundaries: Applied Geophysics, **83**, 3–20 (in Russian: Невинный А.В., Урупов А.К., 1976, Определение пластовых скоростей в средах с криволинейными границами: Прикладная геофизика, вып. 83, с. 3–20.).

Paige, C.C., and Saunders, M.A., 1982, LSQR: an algorithm for sparse linear equations and sparse least squares: ACM Transactions on Mathematical Software, **8**, 43-71.

Pan, N.D., and Gardner, G.H.F., 1986, The p, q algorithm for seismic data analysis: 56th Annual International SEG Meeting, Houston, Texas.

Phinney, R.A., and Jurdy, D.M., 1979, Seismic imaging of deep crust: Geophysics, **44**, 1637–1660.

Puzyrev, N.N., 1979, Time fields of reflected waves and the method of effective parameters: Nauka (in Russian: Пузырев Н.Н., 1979, Временные поля отраженных волн и метод эффективных параметров: Наука).

Rapoport, M.B., 1977, Determination of wave parameters through the CDR summation of seismic traces, *in* Riabinkin, L.A., Ed., Digital processing of reflection seismic data: Transactions of the Gubkin Institute of Petrochemical and Gas Production (Moscow), **120**, 17–22 (in Russian: Рапопорт М.Б., 1977, Определение параметров волн при суммировании сейсмических записей по МРНП: в сб.: Цифровая обработка данных сейсморазведки: Труды Московского института нефтехимической и газовой промышленности им. Губкина, вып. 120., с. 17–22).

Reshef, M., and Kosloff, D., 1984, Three common-shot migration methods: 54th Annual International SEG Meeting, Atlanta, Georgia.

Riabinkin, L.A., Napalkov, Iu.V., Znamenskii, V.V., Voskresenskii, Iu.N., and Rapoport, M.B., 1962, Theory and practice of the CDR seismic method: Transactions of the Gubkin Institute of Petrochemical and Gas Production (Moscow), **39** (in Russian: Рябинкин Л.А., Напалков Ю.В., Знаменский В.В., Воскресенский Ю.Н., Рапопорт М.Б., 1962, Теория и практика сейсмического метода РНП: Труды Московского института нефтехимической и газовой промышленности им. Губкина, вып. 39).

Rieber, F., 1936, A new reflection system with controlled directional sensitivity: Geophysics, **1**, 97–106.

Schultz, P.S., and Claerbout, J.F., 1978, Velocity estimation and downward continuation by wavefront synthesis: Geophysics, **43**, 691–714.

Stoffa, P.L., Buhl, P., Diebold, J.B., and Wenzel, F., 1981, Direct mapping of seismic data to the domain of intercept time and ray parameter—a plane wave decomposition: Geophysics, **46**, 255–267.

Stork, C., and Clayton, R., 1986, Analysis of the resolution between ambiguous velocity and reflector position for traveltime tomography: 56th Annual International SEG Meeting, Houston, Texas.

Tarantola, A., and Valette, B., 1982, Generalized nonlinear inverse problems solved using the least squares criterion: Reviews of Geophysics and Space Physics, **20**, 219–232.

Urupov, A.K., and Levin, A.N., 1985, The determination and interpretation of velocities in reflection seismology: Nedra (in Russian: Урупов А.К., Лёвин А.Н., 1985, Определение и интерпретация скоростей в методе отраженных волн: Недра).

Vasil'ev, S.A., and Urupov, A.K., 1978, New possibilities for studying the propagation velocity of seismic waves and the structure of the medium by means of observations at reciprocal points: Applied Geophysics, **92**, 3–16 (in Russian: Васильев С.А., Урупов А.К., 1978, Новые возможности изучения скорости распространения сейсмических волн и строения среды по наблюдениям во взаимных точках: Прикладная геофизика, вып. 92, с. 3–16).

Zavalishin, B.R., 1975, On the dimensions of the portion of a boundary which forms a reflected wave: Applied Geophysics, **77**, 67–74 (in Russian: Завалишин Б.Р., 1975, О размерах участка границы формирующего отраженную волну: Прикладная геофизика, вып. 77, с. 67–74).

Zavalishin, B.R., 1981, Perfection of methods for constructing seismic images using controlled direct reception: Soviet Geology and Geophysics, **22**, no. 10, 98–104 (English translation of: Завалишин, Б.Р., 1981, Совершенствование приемов построения сейсмических изображений по методу РНП: Геология и геофизика, № 10, с. 114–122).

Zavalishin, B.R., Voskresenskii, Iu.N., and Konopliantsev, M.A., 1982, Interpretational selection of waves for the faster construction of CDR dynamic depth sections: Applied Geophysics, **102**, 37–48 (in Russian: Завалишин Б.Р., Воскресенский Ю.Н., Коноплянцев М.А., 1982, Интерпретационый отбор волн при ускоренном построении динамических глубинных разрезов МРНП: Прикладная геофизика, вып. 102, с. 37–48).