0.1 (title page)

Imaging the Earth's Interior

Chapter 3. Offset

Jon F. Claerbout

3.0 Offset, Another Dimension

Earlier chapters have assumed that the shot and the geophone are located in the same place. The reality is that there is often as much as a 3-km horizontal separation between them. The 3-km offset is comparable to the depth of many petroleum reservoirs.

Offset is another *dimension* in the analysis of data. At the time of writing, this dimension is often represented in field operations by about 48 channels. No one seems to believe, however, that 48 channels is enough. Recording systems with as many as 1024 channels are coming into use.

The offset dimension adds three important aspects to reflection seismology. First, it enables us to routinely measure the velocity of seismic waves in rocks. This velocity has been assumed to be known in the previous chapters of this book. Second, it gives us data redundancy: it gives independent measurements of quantities that should be identical. Superposition of the measurements (stacking) offers the potential for signal enhancement by destructive interference of noise. Third (a disadvantage), since the offset is nonzero, procedures for migration take on another element of complexity. By the end of this chapter we will be trying to deal with three confusing subjects at the same time — namely, dip, offset, and lateral velocity variation.

Theoretically it seems that offset should offer us the possibility of identifying rocks by observing the reflection coefficient as a function of angle, both for p waves and for p-to-s converted waves. The reality seems to be that neither measurement can be made reliably, if at all. See Section 1.4 for a fuller discussion of this subject, which would be an interesting one for research, with a large potential for practical rewards. See also Ostrander (1982) and Tatham and Stoffa (1976). The reasons for the difficulty in measurement, and the resolution of the difficulty, are, however, not the goal of this book. This goal is instead to enable us to deal effectively with that which is routinely observable.

Stacking Diagrams

First, define the midpoint y between the shot and geophone, and define h to be half the horizontal offset between the shot and

geophone:

$$y = \frac{g + s}{2} \tag{1a}$$

$$h = \frac{g - s}{2} \tag{1b}$$

The reason for using ***half*** the offset in the mathematical equations is to simplify and symmetrize many later equations. The definition of offset is made with $g - s$ rather than with $s - g$ so that positive offset means waves moving in the positive x direction. In the marine case, this means the ship is presumed to sail negatively along the x-axis. In reality the ship may go either way, and shot points may either increase or decrease as the survey proceeds. In some situations you may clarify matters by setting the field observer's shot-point numbers to negative values.

Data is defined experimentally in the space of (s, g). Equation (1) represents a change of coordinates to the space of (y, h). Midpoint-offset coordinates are especially useful for interpretation and data processing. Since the data is also a function of the travel time t, the full dataset lies in a volume. Because it is so difficult to make a satisfactory display of such a volume, what is customarily done is to display slices. The names of slices vary slightly from one company to the next. The following names seem to be well known and clearly understood:

$(y, h=0, t)$	zero-offset section
$(y, h=h_{\text{min}}, t)$	near-trace section
$(y, h=const, t)$	constant-offset section
$(y, h=h_{\text{max}}, t)$	far-trace section
$(y=const, h, t)$	common-midpoint gather
$(s=const, g, t)$	common-shot gather or profile
$(s, g=const, t)$	common-geophone gather

These slices may be found in figure 1.

A common-depth-point (CDP) gather is defined by the industry and by common usage to be the same thing as a common-midpoint (CMP) gather. But in this book there will be a pedantic distinction. A CDP gather is defined to be a CMP gather with its time axis stretched according to some velocity model, say,

$$(y=const, h, \sqrt{t^2 - 4h^2/v^2}) \quad \text{common-depth-point gather}$$

This offset dependent stretching makes the time axis of the gather become more like a *depth* axis, thus providing the D in CDP. This

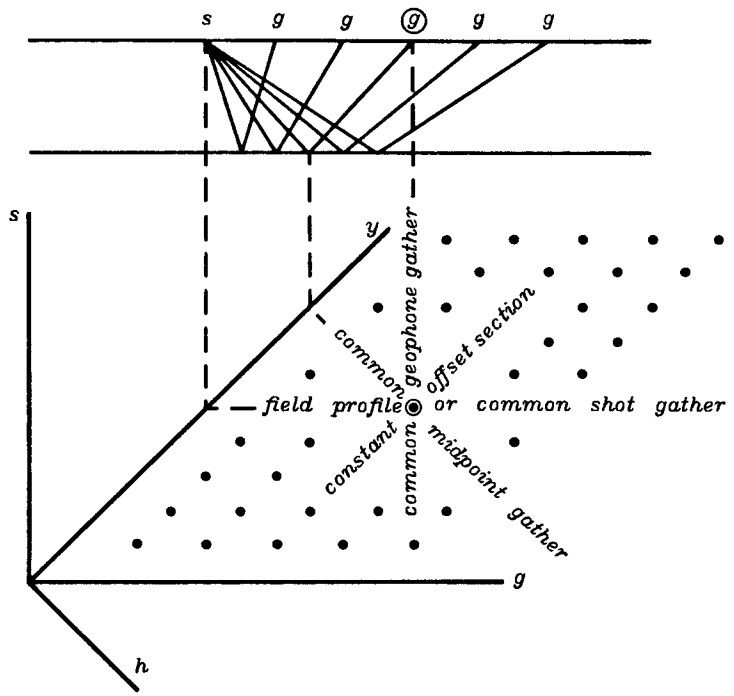


FIG. 1. Top depicts a field recording of marine seismograms from a shot at s to geophones labeled g. There is a horizontal reflecting layer to aid interpretation. The lower diagram is called a stacking diagram. (It is ***not*** a perspective drawing.) Each dot in this plane depicts a possible seismogram. Think of time running out from the plane. The center geophone above (circled) records the seismogram (circled) that may be found in various geophysical displays. Labels in the diagram below give common names for the displays.

stretching of the time axis is called normal moveout correction (NMO). Notice that as the velocity goes to infinity, the amount of stretching goes to zero.

In routine industrial practice the data is rarely displayed as a function of offset. Instead, each CDP gather is summed over offset. The resulting sum is a single trace. Such a trace can be constructed at each midpoint. The collection of such traces, a function of midpoint and time, is called a CDP stack. Roughly speaking, a CDP stack is like a zero-offset section, but it has a less noisy

appearance.

The construction of a CDP stack requires that a numerical choice be made for the moveout correction velocity. This choice is called the stacking velocity. Naturally, it is usually your best guess of the earth's actual velocity. Indeed a popular means of estimating the earth's velocity is to try several stacking velocities and see which CDP stack is strongest and least noisy.

What is "Poor Quality" Data?

Vast regions of the world have good petroleum potential but are hard to explore because of the difficulty of obtaining good quality reflection seismic data. The reasons are often unknown. What is "poor quality" data? From an experimentalist view, almost all seismic data is good in the sense that it is repeatable. The real problem is that the data makes no sense.

Take as an earth model a random arrangement of point reflectors. Its migrated zero-offset section should look random too. Given the repeatability that is experienced in data collection, data with a random appearance implies a random jumble of reflectors. With only zero-offset data little else can be deduced. But with the full range of offsets at our disposal, a more thoughtful analysis can be tried. This chapter provides some of the required techniques.

An interesting model of the earth is a random jumble of point scatterers in a constant-velocity medium. The data would be a random function of time and a random function of the horizontal location of the shot-geophone midpoint. But after suitable processing, for each midpoint, the data should be a perfectly hyperbolic function of shot-geophone offset. This would determine the earth velocity exactly, even if the random scatterers were distributed in three dimensions, and the survey were only along a surface line.

This particular model could fail to explain the "poor quality" data. In that case other models could be tried. The effects of random velocity variations in the near surface or the effects of multiple reflections could be analyzed. Noise in seismology can usually be regarded as a failure of analysis rather than as something polluting the data. It is the offset dimension that gives us the redundancy we need to try to figure out what is really happening.

Texture of Horizontal Bedding, Marine Data

Gravity is a strong force for the stratification of rocks, and in many places in the world rocks are laid down in horizontal beds. Yet even in the most ideal environment the bedding is not mirror smooth; it has some *texture*. We begin the study of offset with synthetic data that mimics the most ideal environment. Such an environment is almost certainly marine, where sedimentary deposition can be slow and uniform. The wave velocity will be taken to be constant, and all rays will reflect as from horizontally lying mirrors. Mathematically, *texture* is introduced by allowing the reflection coefficients of the beds to be laterally variable. The lateral variation is presumed to be a random function, though not necessarily with a white spectrum. Let us examine the appearance of the resulting field data.

Randomness is introduced into the earth with a random function of midpoint y and depth z. This randomness is impressed on some geological "layer cake" function of depth z. For every point in (y,z)-space, a hyperbola of the appropriate random amplitude must be superposed in the space of offset h and travel time t.

What does the final data space look like? This question has little meaning until we decide how the three-dimensional data volume will be presented to the eye. Let us view the data much as it is recorded in the field. For each shot point we see a frame in which the vertical axis is the travel time and the horizontal axis is the distance from the ship down the towed geophone cable. The next shot point gives us another frame. Repetition gives us a movie. And what does the movie show?

A single frame shows hyperbolas with imposed texture. The movie shows the texture moving along each hyperbola to increasing offsets. (I find that no sequence of still pictures can give the impression that the movie gives.) Really the ship is moving; the texture of the earth is remaining stationary under it. This is truly what most marine data looks like, and the computer program of figure 2 does a good job of simulating it. Comparing synthetic data to real marine-data movies, I am impressed by the large amount of random lateral variation required in the synthetics to achieve resemblance to field data. The randomness seems too great to represent lithologic variation. Apparently it is the result of something not modeled. Perhaps it results from our incomplete understanding of the mechanism of reflection from the quasi-random earth. Or perhaps it is an effect of the partial focusing of waves

```
# Synthetic marine data tape movie generation
integer kbyte,it,nt,ih,nh,is,ns,iz,nz,it0,iy
real p(512),b(512),refl(25,16),z(25),geol(25),random
open(3,file="plot",status='new',access='direct',form='unformatted',recl=1)
nt = 512;      nh = 48;       ns = 10;       nz = 25;kbyte = 1
do iz=1,nz                              # Reflector depth
      z(iz) =  nt*random()              # random() is on the interval (0.,1.)
do iz=1,nz                              # Reflector strength with depth.
      geol(iz) = 2.*random()-1.
do is = 1,ns                            # Give texture to the Geology
      do iz = 1,nz
            refl(iz,is) = (1.+random())*geol(iz)
do it = 1,nt                                    # Prepare a wavelet
      b(it) = exp(-it*.08)*sin(.5*it-.5)
do is = ns,1,-1 {                               # Shots. Run backwards.
      do ih = 1,nh {                            # down cable h = (g-s)/2
            iy = (is-1)+(ih-1)                  # y = midpoint
            iy = 1 + (iy-ns*(iy/ns))            # periodic with midpoint
            do it = 1,nt
                  p(it) = 0.
            do iz = 1,nz {                      # Add in a hyperbola for each layer
                  it0 = sqrt( z(iz)**2 + 100.*(ih-1)**2 )
                  do it = 1,nt-it0 {            # Add in the wavelet
                        p(it+it0) = p(it+it0) + refl(iz,iy)*b(it)
                        }
                  }
            write(3,rec=kbyte) (p(it),it=1,nt);kbyte = kbyte+nt*4
            }
      }
stop;       end
```

FIG. 2. Computer program to make synthetic field tapes in an ideal marine environment.

sometime after they reflect from minor topographic irregularities. A full explanation awaits more research.

Texture of Land Data: Near-Surface Problems

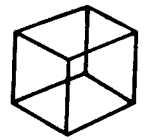
Reflection seismic data recorded on land frequently displays randomness because of the irregularity of the soil layer. Often it is so disruptive that the seismic energy sources are deeply buried (at much cost). The geophones are too many for burial. For most land reflection data, the texture caused by these near-surface irregularities exceeds the texture resulting from the reflecting layers.

To clarify our thinking, an ideal mathematical model will be proposed. Let the reflecting layers be flat with no texture. Let the geophones suffer random time delays of several time points. Time delays of this type are called *statics*. Let the shots have random strengths. For this movie, let the data frames be *common-midpoint gathers*, that is, let each frame show data in (h,t)-space at a fixed midpoint y. Successive frames will show successive midpoints. The study of figure 1 should convince you that the travel-time irregularities associated with the geophones should move leftward, while the amplitude irregularities associated with the shots should move rightward. In real life, both amplitude and time anomalies are associated with both shots and geophones.

EXERCISES

1. Note that figure 1 is drawn for a shot interval Δs equal to half the geophone interval Δg. Redraw figure 1 for $\Delta s = \Delta g$. Common-midpoint gathers now come in two types. Suggest two possible definitions for "near-offset section."
2. Make the synthetic land data movie. Observing this movie you will note the perceptual problem of being able to see the leftward motion along with the rightward motion.

   Your mind will often see only one, blocking out the other, similar to the way you perceive a 3-D cube, from a 2-D projection of its edges. Define recursive dip filters to pass and reject the various textures of shot, geophone, and midpoint.



3.1 Absorption and a Little Focusing

Sometimes the earth strata lie horizontally with little irregularity. There we may hope to ignore the effects of migration. Seismic rays should fit a simple model with large reflection angles occurring at wide offsets. Such data should be ideal for the measurement of reflection coefficient as a function of angle, or for the measurement of the earth acoustic absorptivity $1/Q$. In his doctoral dissertation, Einar Kjartansson reported such a study. The results were so instructive that the study will be thoroughly reviewed here. I don't know to what extent the Grand Isle gas field typifies the rest of the earth, but I do know it is an excellent place to begin learning about the meaning of shot-geophone offset.

The Grand Isle Gas Field: A Classic Bright Spot

The dataset Kjartansson studied was a seismic line across the Grand Isle gas field, off the shore of Louisiana, and was supplied by the Gulf Oil Company. The data contain several classic "bright spots" (strong reflections) on some rather flat undisturbed bedding. Of interest are the lateral variations in amplitude on reflections at a time depth of about 2.3 seconds. (See figure 3.) It is widely believed that such bright spots arise from shallow gas-bearing sands.

Theory predicts that reflection coefficient should be a function of angle. For an anomalous physical situation like gas-saturated sands, the function should be distinctive. Evidence should be found on common-midpoint gathers like those shown in figure 1. Looking at any one of these gathers you will note that the reflection strength versus offset seems to be a smooth, sensibly behaved function, apparently quite measurable. Using layered media theory, however, it was determined that only the most improbably bizarre medium could exhibit such strong variation of reflection coefficient with angle, particularly at small angles of incidence. (The reflection angle of the energy arriving at wide offset at time 2.5 seconds is not a large angle. Assuming constant velocity, $\arccos(2.3/2.6) = 28°$.) Compounding the puzzle, each common-midpoint gather shows a ***different*** smooth, sensibly behaved, measurable function. Furthermore, these midpoints are near one another, ten shot points spanning a horizontal distance of 820 feet.

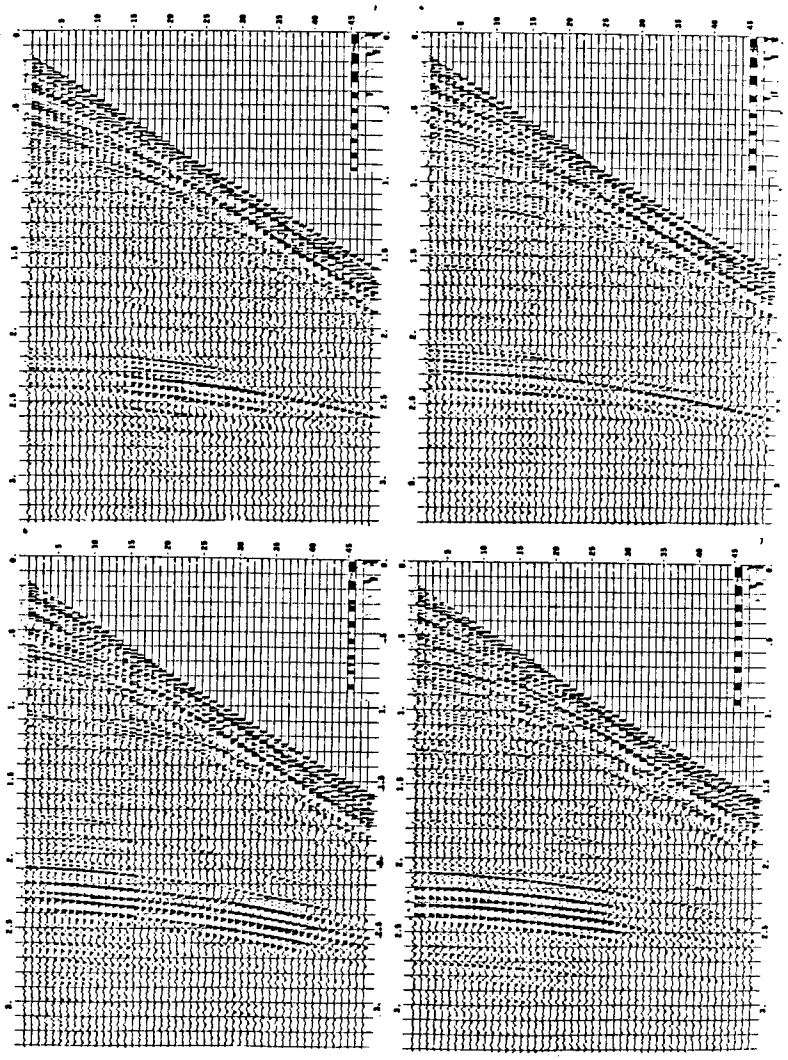
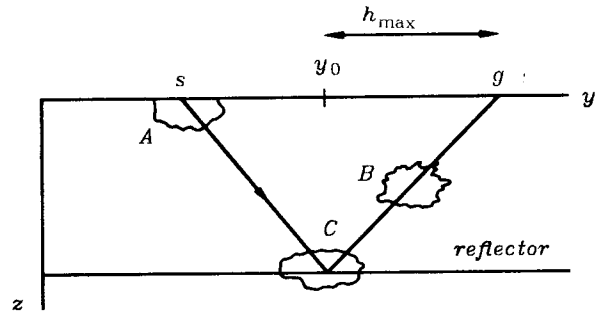


FIG. 1. Top left is shot point 210; top right is shot point 220. No processing has been applied to the data except for a display gain proportional to time. Bottom shows shot points 305 and 315. (Kjartansson, Gulf)

Kjartansson's Model for Lateral Variation in Amplitude

The Grand Isle data is incomprehensible in terms of the model based on layered media theory. Kjartansson proposed an alternative model. Figure 2 illustrates a geometry in which rays travel in straight lines from any source to a flat horizontal reflector, and thence to the receivers. The only complications are "pods" of some material that is presumed to disturb seismic rays in some anomalous way. Initially you may imagine that the pods absorb wave energy. (In the end it will be unclear whether the disturbance results from energy focusing or absorbing.)



The model above produces the disturbed data space sketched below.

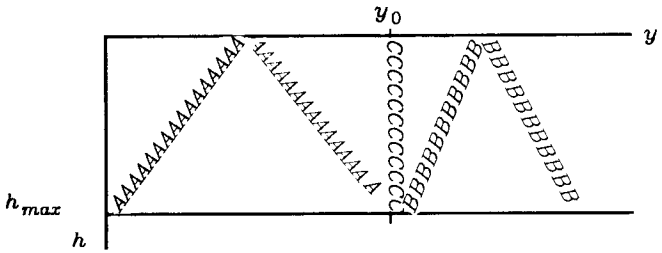


FIG. 2. Kjartansson's model. Anomalous material in pods A, B, and C may be detected by its effect on reflections from a deeper layer.

Pod A is near the surface. The seismic survey is affected by it twice — once when the pod is traversed by the shot and once when it is traversed by the geophone. Pod C is near the reflector and encompasses a small area of it. Pod C is seen at all offsets h but only at one midpoint, y_0. The raypath depicted on the top of figure 2 is one that is affected by all pods. It is at midpoint y_0 and at the widest offset h_{max}. Find the raypath on the lower diagram in figure 2.

Pod B is part way between A and C. The slope of affected points in the (y,h)-plane is part way between the slope of A and the slope of C.

Figure 3 shows a common-offset section across the gas field. The offset shown is the fifth trace from the near offset, 1070 feet from the shot point. Don't be tricked into thinking the water was deep. The first break at about .33 seconds is wide-angle propagation.

The power in each seismogram was computed in the interval from 1.5 to 3 seconds. The logarithm of the power is plotted in figure 4a as a function of midpoint and offset. Notice streaks of energy slicing across the (y,h)-plane at about a 45° angle. The strongest streak crosses at exactly 45° degrees through the near offset at shot point 170. This is a missing shot, as is clearly visible in figure 3. Next, think about the gas sand described as pod C in the model. Any gas-sand effect in the data should show up as a streak across all offsets at the midpoint of the gas sand — that is, horizontally across the page. I don't see such streaks in figure 4a. Careful study of the figure shows that the rest of the many clearly visible streaks cut the plane at an angle noticeably *less* than ±45°. The explanation for the angle of the streaks in the figure is that they are like pod B. They are part way between the surface and the reflector. The angle determines the depth. Being closer to 45° than to 0°, the pods are closer to the surface than to the reflector.

Figure 4b shows timing information in the same form that figure 4a shows amplitude. A CDP stack was computed, and each field seismogram was compared to it. A residual time shift for each trace was determined and plotted in figure 4b. The timing residuals on one of the common-midpoint gathers is shown in figure 5.

The results resemble the amplitudes, except that the results become noisy when the amplitude is low or where a "leg jump" has confounded the measurement. Figure 4b clearly shows that the disturbing influence on timing occurs at the same depth as that which disturbs amplitudes.

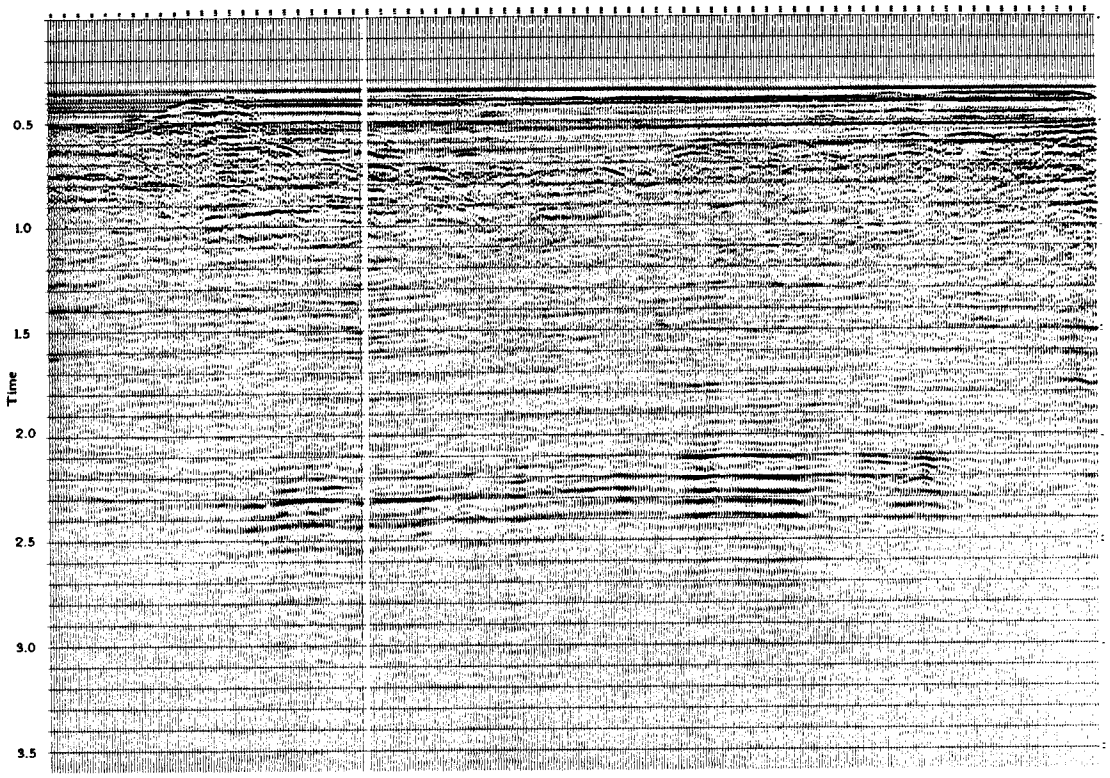


FIG. 3. A constant-offset section across the Grand Isle gas field. The offset shown is the fifth from the near trace. (Kjartansson, Gulf)

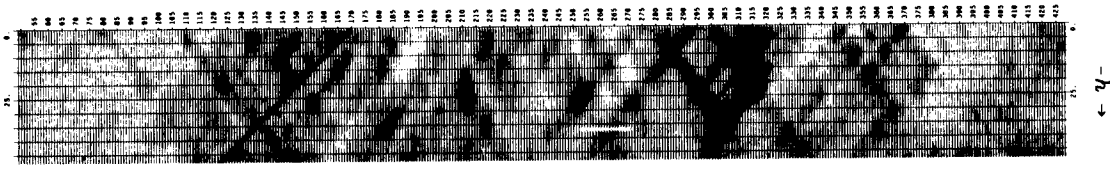


FIG 4a. amplitude (h,y)

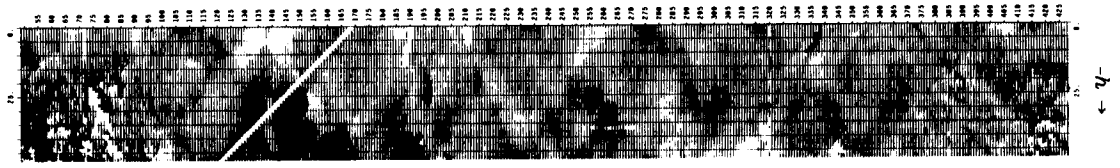


FIG 4b. timing (h,y)

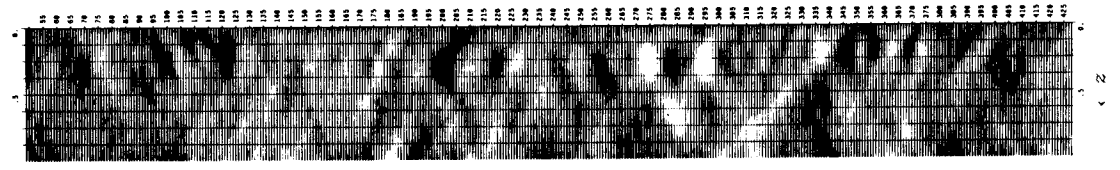


FIG 4c. amplitude (z,y)

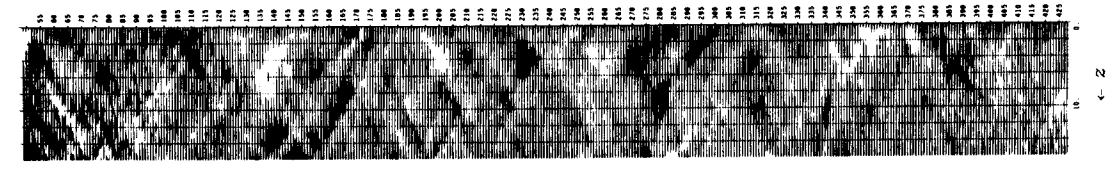


FIG 4d. timing (z,y)

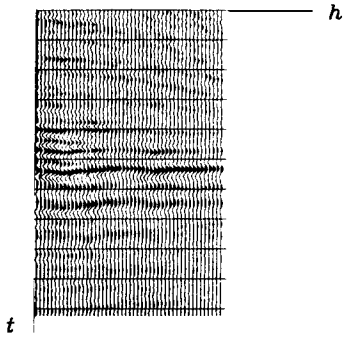


FIG. 5. Midpoint gather 220 (same as in figure 1b) after moveout. Shown is a one-second window centered at 2.3 seconds, time shifted according to an NMO for an event at 2.3 seconds, using a velocity of 7000 feet/sec. (Kjartansson)

The process of ***inverse slant stack***, briefly described in Section 1.6, enables us to determine the depth distribution of the pods. This distribution is displayed in figures 4c and 4d.

Rotten Alligators

The sediments carried by the Mississippi River are dropped at the delta. There are sand bars, point bars, old river bows now silted in, a crow's foot of sandy distributary channels, and between channels, swampy flood plains are filled with decaying organic material. The landscape is clearly laterally variable, and eventually it will all sink of its own weight, aided by growth faults and the weight of later sedimentation. After it is buried and out of sight the lateral variations will remain as pods that will be observable by the seismologists of the future. These seismologists may see something like figure 6.

Focusing or Absorption?

Highly absorptive rocks usually have low velocity. Behind a low velocity pod, waves should be weakened by absorption. They should also be strengthened by focusing. Which effect dominates? How does the phenomenon depend on spatial wavelength? A full reconstruction of the physical model remains to be done. Maybe

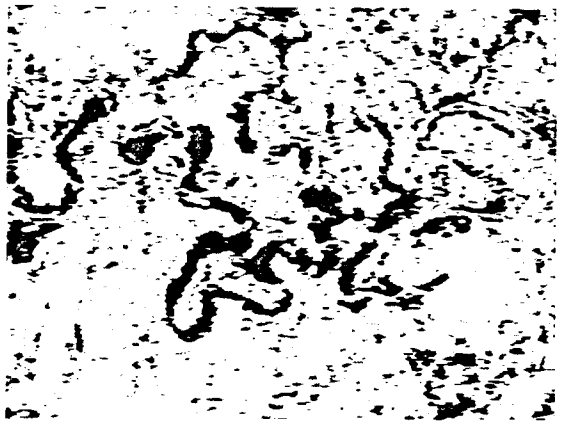


FIG. 6. Map view of seismic data at 196 milliseconds travel time in the Gulf of Thailand. The data shows ancient river meanders now submerged. (Geophysical Services Inc.)

you can figure it out knowing that black on figure 4c denotes low amplitude or high absorption, and black on figure 4d denotes low velocities.

EXERCISE

1. Consider waves converted from pressure P waves to shear S waves. Assume an S-wave speed of about half the P-wave speed. What would figure 2 look like for these waves?

3.2 Introduction to Dip

The study of seismic travel time dependence upon source/receiver offset begins by calculating the travel times for rays in some ideal environments.

Sections and Gathers for Planar Reflectors

The simplest environment for reflection data is a single horizontal reflection interface, which is shown in figure 1. As expected, the zero-offset section mimics the earth model. The common-midpoint gather is a hyperbola whose asymptotes are straight lines with slopes of the velocity v_1. The most basic data processing is called *common-depth-point stack* or CDP stack. In it, all the traces on the common-midpoint (CMP) gather are time shifted into alignment and then added together. The result mimics a zero-offset trace. The collection of all such traces is called the *CDP-stacked section*. In practice the CDP-stacked section is often interpreted and migrated as though it were a zero-offset section. In this chapter we will learn to avoid this popular, oversimplified assumption.

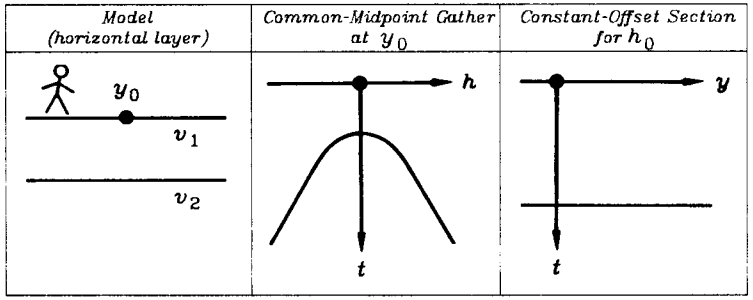


FIG. 1. Simplest earth model.

The next simplest environment is to have a planar reflector that is oriented vertically rather than horizontally. This is not typical, but is included here because the effect of earth dip is more comprehensible in an extreme case. Now the wave propagation is along the air-earth interface. To avoid confusion the reflector may be inclined at a slight angle from the vertical, as in figure 2.

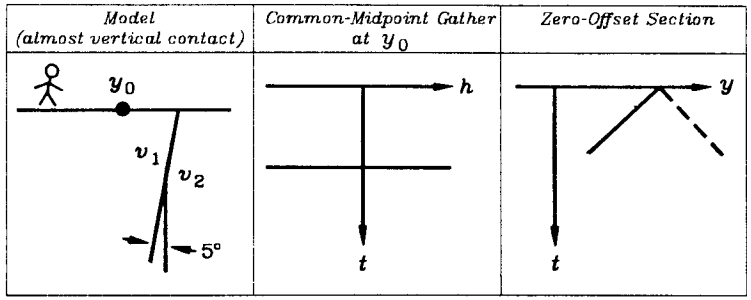


FIG. 2. Near-vertical reflector, a gather, and a section.

Figure 2 shows that the travel time does not change as the offset changes. It may seem paradoxical that the travel time does not increase as the shot and geophone get further apart. The key to the paradox is that midpoint is held constant, not shotpoint. As offset increases, the shot gets further from the reflector while the geophone gets closer. Time lost on one path is gained on the other.

A planar reflector may have any dip between horizontal and vertical. Then the common-midpoint gather lies between the common-midpoint gather of figure 1 and that of figure 2. The zero-offset section in figure 2 is a straight line, which turns out to be the asymptote of a family of hyperbolas. The slope of the asymptote is the velocity v_1.

The Dipping Bed

While the travel-time curves resulting from a dipping bed are simple, they are not simple to derive. Before the derivation, the result will be stated: for a bed dipping at angle α from the horizontal, the travel-time curve is

$$t^2 v^2 = 4(y-y_0)^2 \sin^2\alpha + 4h^2 \cos^2\alpha \qquad (1)$$

For $\alpha = 45°$, equation (1) is the familiar Pythagoras cone, it is just like $t^2 = z^2 + x^2$. For other values of α, the equation is still a cone, but a less familiar one because of the stretched axes.

For a common-midpoint gather at $y = y_1$ in (h,t)-space, equation (1) looks like $t^2 = t_0^2 + 4h^2/v^2_{apparent}$. Thus the common-midpoint gather contains an ***exact*** hyperbola, regardless of the earth dip angle α. The effect of dip is to change the asymptote of the hyperbola, thus changing the apparent velocity. The result has great significance in applied work and is known as Levin's dip correction (1971):

$$v_{apparent} = \frac{v_{earth}}{\cos(\alpha)} \qquad (2)$$

In summary, dip increases the stacking velocity.

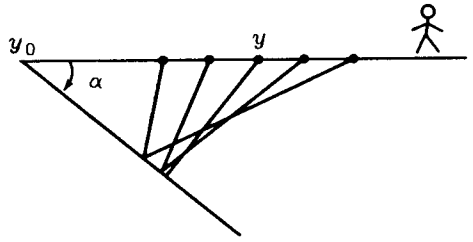


FIG. 3. Rays from a common-midpoint gather.

Figure 3 depicts some rays from a common-midpoint gather. Notice that each ray strikes the dipping bed at a different place. So a common-***midpoint*** gather is not a common-***depth-point*** gather. To realize why the reflection point moves on the reflector, recall the basic geometrical fact that an angle bisector in a triangle generally doesn't bisect the opposite side. The reflection point moves ***up*** dip with increasing offset.

Finally, equation (1) will be proved. Figure 4 shows the basic geometry along with an "image" source on another reflector of twice the dip. For convenience, the bed intercepts the surface at $y_0 = 0$. The length of the line $s'g$ in figure 3 is determined by the trigonometric Law of Cosines to be

$$t^2v^2 = s^2 + g^2 - 2sg\cos 2\alpha$$
$$t^2v^2 = (y-h)^2 + (y+h)^2 - 2(y-h)(y+h)\cos 2\alpha$$
$$t^2v^2 = 2(y^2+h^2) - 2(y^2-h^2)(\cos^2\alpha - \sin^2\alpha)$$

$$t^2v^2 = 4y^2\sin^2\alpha + 4h^2\cos^2\alpha$$

which is equation (1).

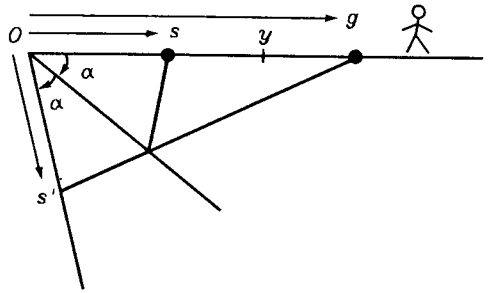


FIG. 4. Travel time from image source at s' to g may be expressed by the law of cosines.

Another facet of equation (1) is that it describes the constant-offset section. Surprisingly, the travel time of a planar bed becomes curved — it too becomes hyperbolic.

The Point Response

Another simple geometry is a reflecting point within the earth. A wave incident on the point from any direction reflects waves in all directions. This geometry is particularly important because any model is a superposition of such point scatterers. Figure 5 shows an example. The curves in figure 5 include flat spots for the same reasons that some of the curves in figures 1 and 2 were straight lines.

The point-scatterer geometry for a point located at (x,z) is shown in figure 6.

The equation for travel time t is the sum of the two travel paths

$$tv = \sqrt{z^2 + (s-x)^2} + \sqrt{z^2 + (g-x)^2} \qquad (3)$$

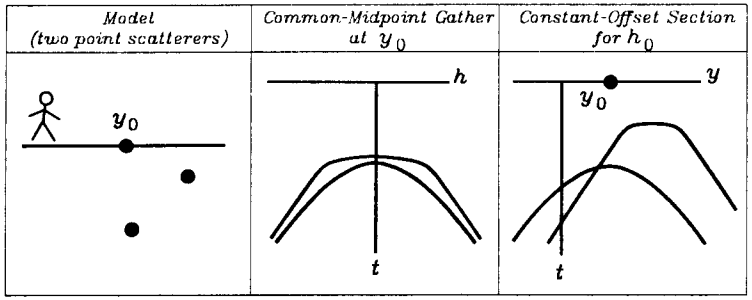


FIG. 5. Response of two point scatterers. Note the flat spots.

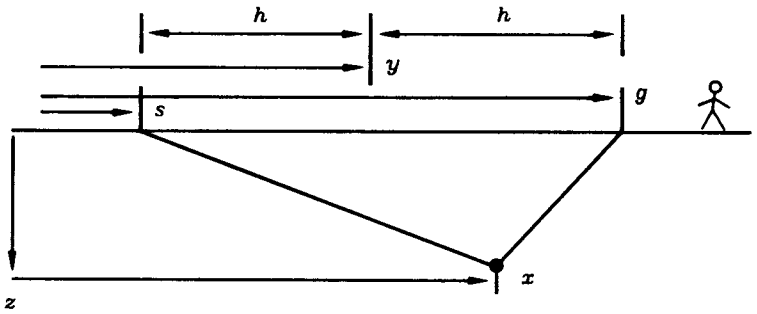


FIG. 6. Geometry of a point scatterer.

Cheops' Pyramid

Because of the importance of the point-scatterer model, we will go to considerable lengths to visualize the functional dependence among t, z, x, s, and g in equation (3). This picture is more difficult — by one dimension — than is the conic section of the exploding-reflector geometry.

To begin with, suppose that the first square root in (3) is constant because everything in it is held constant. This leaves the familiar hyperbola in (g,t)-space, except that a constant has been added to the time. Suppose instead that the other square root is constant. This likewise leaves a hyperbola in (s,t)-space. In (s,g)-space, travel time is a function of s plus a function of g. I think of this as one coat hanger, which is parallel to the s-axis, being hung from another coat hanger, which is parallel to the g-axis.

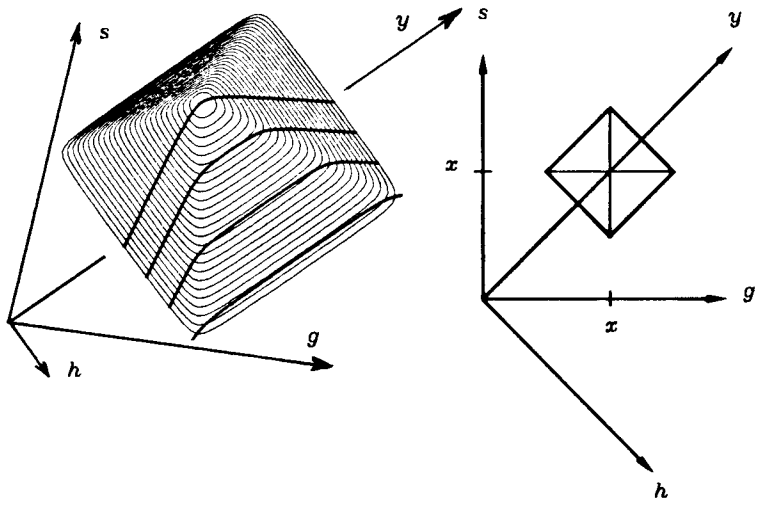


FIG. 7a,b. Left is a picture of the travel-time mountain of equation (3) for fixed x and z. The darkened lines are constant-offset sections. Right is a cross section through the mountain for large t (or small z). (Ottolini)

A view of the travel-time mountain on the (s,g)-plane or the (y,h)-plane is shown in figure 7a. Notice that a cut through the mountain at large t is a square, the corners of which have been smoothed. A constant value of t is the square contoured in (s,g)-space, as in figure 7b. Algebraically, the squareness becomes evident for a point reflector near the surface, say, $z \to 0$. Then (3) becomes

$$v\,t \;=\; |s - x| \;+\; |g - x| \qquad (4)$$

The center of the square is located at $(s,g)=(x,x)$. Taking travel time t to increase downward from the horizontal plane of (s,g)-space, the square contour is like a horizontal slice through the

Egyptian pyramid of Cheops. To walk around the pyramid at a constant altitude is to walk around a square. Alternately, the altitude change of a traverse over g at constant s is simply a constant plus an absolute-value function, as is a traverse of s at constant g.

More interesting and less obvious are the curves on common-midpoint gathers and constant-offset sections. Recall the definition that the midpoint between the shot and geophone is y. Also recall that h is half the horizontal offset from the shot to the geophone.

$$y = \frac{g + s}{2} \tag{5a}$$

$$h = \frac{g - s}{2} \tag{5b}$$

A traverse of y at constant h is shown in figure 7. At the highest elevation on the traverse, you are walking along a flat horizontal step like the flat-topped hyperboloids of figure 5. Some erosion to smooth the top and edges of the pyramid gives a model for nonzero reflector depth.

For rays that are near the vertical, the travel-time curves are far from the hyperbola asymptotes. Then the square roots in (3) may be expanded in Taylor series, giving a parabola of revolution. This describes the eroded peak of the pyramid.

Random Point Scatterers

Figure 8 shows a synthetic constant-offset section (COS) taken from an earth model containing about fifty randomly placed point scatterers. Late arrival times appear hyperbolic. Earlier arrivals have flattened tops. Nothing arrives before a ray going horizontally from the shot to the geophone.

Figure 9 shows a synthetic common-shot profile (CSP) from the same earth model of random point scatterers. Each scatterer produces a hyperbolic arrival. The hyperbolas are not symmetric around zero offset; their locations are random. They must, however, all lie under the lines $|g-s| = vt$. Hyperbolas with sharp tops can be found at late times as well as early times. However, the sharp tops, which are from shallow scatters near the geophone, must lie near the lines $|g-s| = vt$.

Figure 10a shows a synthetic common-midpoint gather (CMP) from an earth model containing about fifty randomly placed point

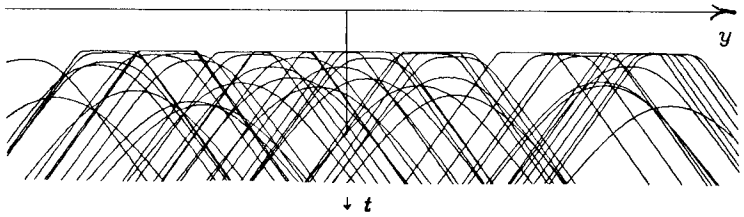


FIG. 8. Constant-offset section over random point scatterers.

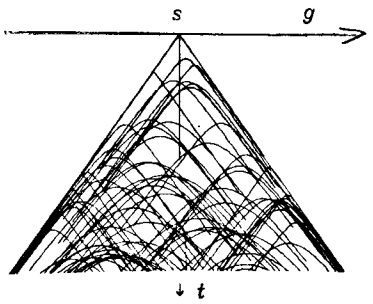


FIG. 9. Common-shot profile over random point scatterers.

scatterers. Because this is a common-*midpoint* gather, the curves are symmetric through zero offset, and the negative offsets of field data are never plotted. Some of the arrivals have flattened tops, which indicate scatterers that are not directly under the midpoint.

Normal-moveout (NMO) correction is a stretching of the data to try to flatten the hyperbolas. This correction assumes flat beds, but it also works for point scatterers that are directly under the midpoint. Figure 10b shows what happens when normal-moveout correction is applied on the random scatterer model. Some reflectors are flattened; others are "overcorrected."

Forward and Backward Scattering: Larner's Streaks

At some locations, near-surface waves overwhelm the deep reflections of geologic interest. Compounding our difficulty, the near-surface waves are usually irregular because the earth is comparatively more irregular at its surface than deeper down. On land, these interfering waves are called ground roll. At sea, they

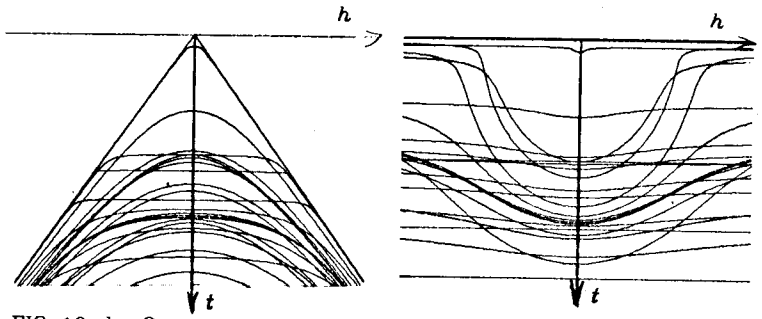


FIG. 10a,b. Common-midpoint gather on earth of randomly located point scatterers (left). The same gathers after NMO correction (right).

are called water waves (not to be confused with surface waves on water).

A model for such near-surface noises is suggested by the vertical reflecting wall in figure 2. In this model the waves remain close to the surface. Randomly placed vertical walls could result in a zero-offset section that resembles the field data of figure 11. Another less extreme model for the surface noises is the flat-topped curves in the random point-scatterer model.

In the random point-reflector model the velocity was a constant. In real life the earth velocity is generally slower for the near-surface waves and faster for the deep reflections. This sets the stage for some unexpected noise amplification.

CDP stacking enhances events with the stacking velocity, and discriminates against events with other velocities. Thus you might expect that stacking at deeper, higher velocity would discriminate against low-velocity, near-surface events. Near-surface noises, however, are not reflections from horizontal layers; they are more like reflections from vertical walls. Looking back to figure 2, you will notice that the common-midpoint gather appears to have been already moveout corrected although no moveout correction was done. Because the dip of the bed is 90°, equation (2) gives a stacking velocity equal to infinity. This means that no moveout correction is required. In practice the technique of CDP stacking does not have perfect velocity-resolving power. In the presence of irregularities, its velocity-resolving power may not be very good at all. So it is not surprising that stacking at deep-sediment, high velocities can enhance surface noises. This phenomenon was clearly

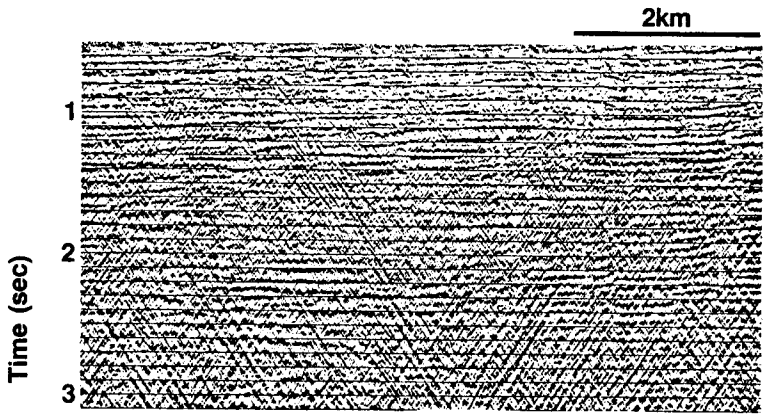


FIG. 11. CDP stack with water noise from the Shelikof Strait, Alaska. (from Larner)

described and explained by Larner et al, 1981, and figure 11 was given to me by him.

Velocity of Sideswipe

Shallow-water noise can come from waves scattering from a sunken ship or from the side of an island or iceberg several kilometers to the side of the survey line. Think of boulders strewn all over a shallow sea floor, not only along the path of the ship, but also off to the sides. Since the velocity of water is constant, the travel-time curves for reflections from the boulders will nicely match the random point-scatterer model. Because of the long wavelengths of seismic waves, our sending and receiving equipment does not enable us to distinguish waves going up and down from those going sideways.

Imagine one of these shallow scatterers several kilometers to the side of the ship. More precisely, let the scatterer be on the earth's surface, perpendicular to the midpoint of the line connecting the shot point to the geophone. A common-midpoint gather for this scatterer is a perfect hyperbola, as from the deep reflector contributions on figure 9. Since it is a water-velocity hyperbola, this scatterered noise should be nicely suppressed by CDP stacking with the higher, sediment velocity. So the "streaking" scatterers mentioned earlier are not sidescatter. The "streaking" scatterers

are those along the survey line, not those perpendicular to it.

The Migration Ellipse

Rather than regarding the reflection point (y_0,z) as being fixed, another insight into equation (3) is to regard the offset h and the total travel time t as fixed constants. Then the locus of possible reflectors turns out to describe an ellipse in the plane of $(y-y_0,z)$. The reason it is an ellipse follows from the geometric definition of an ellipse. To draw an ellipse, place a nail or tack into s on figure 6 and another into g. Connect the tacks by a string that is exactly long enough to go through (y_0,z). An ellipse going through (y_0,z) may be constructed by sliding a pencil along the string, keeping the string tight. The string keeps the total distance tv constant.

Recall that one method for the migration of zero-offset sections is to take every data value in (y,t)-space and use it to superpose an appropriate semicircle in (y,z)-space. For nonzero offset the circle should be generalized to an ellipse.

It is not easy to show that equation (3) can be cast in the standard mathematical form of an ellipse, namely, a stretched circle. But the result is a simple one, and an important one for later analysis, so here we go. Equation (3) in (y,h)-space is

$$tv = \sqrt{z^2+(y-y_0-h)^2}+\sqrt{z^2+(y-y_0+h)^2} \quad (6)$$

To help reduce algebraic verbosity, define a new y equal to the old one shifted by y_0. Also make the definitions

$$tv_{rock} = 2d = 2tv_{half} \quad (7a)$$

$$a = z^2+(y+h)^2 \quad (7b)$$

$$b = z^2+(y-h)^2 \quad (7c)$$

$$a-b = 4yh \quad (7d)$$

With these definitions, (6) becomes

$$2d = \sqrt{a}+\sqrt{b} \quad (8)$$

Square.

$$4d^2-(a+b) = 2\sqrt{a}\sqrt{b} \quad (9)$$

Square again.

$$16d^4-8d^2(a+b)+(a+b)^2 = 4ab \quad (10a)$$

$$16d^4-8d^2(a+b)+(a-b)^2 = 0 \quad (10b)$$

Introduce definitions of a and b.

$$16d^4-8d^2[2z^2+2y^2+2h^2]+16y^2h^2 = 0 \quad (11)$$

Bring y and z to the right.

$$d^4-d^2h^2 = d^2(z^2+y^2)-y^2h^2 \quad (12a)$$

$$d^2(d^2-h^2) = d^2z^2+(d^2-h^2)y^2 \quad (12b)$$

$$d^2 = \frac{z^2}{1-\frac{h^2}{d^2}}+y^2 \quad (12c)$$

Finally, recalling all earlier definitions,

$$t^2v_{half}^2 = \frac{z^2}{1-\frac{h^2}{t^2v_{half}^2}}+(y-y_0)^2 \quad (13)$$

Fixing t, equation (13) is the equation for a circle with a stretched z-axis. Our algebra has confirmed that the "string and tack" definition of an ellipse matches the "stretched circle" definition. An ellipse in model space is the earth model given the observation of a delta function on a constant-offset section. See figure 12.

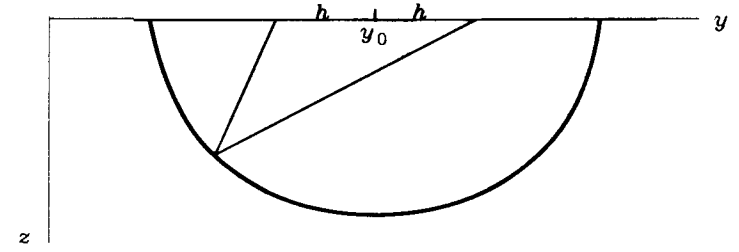


FIG. 12. Migration ellipse.

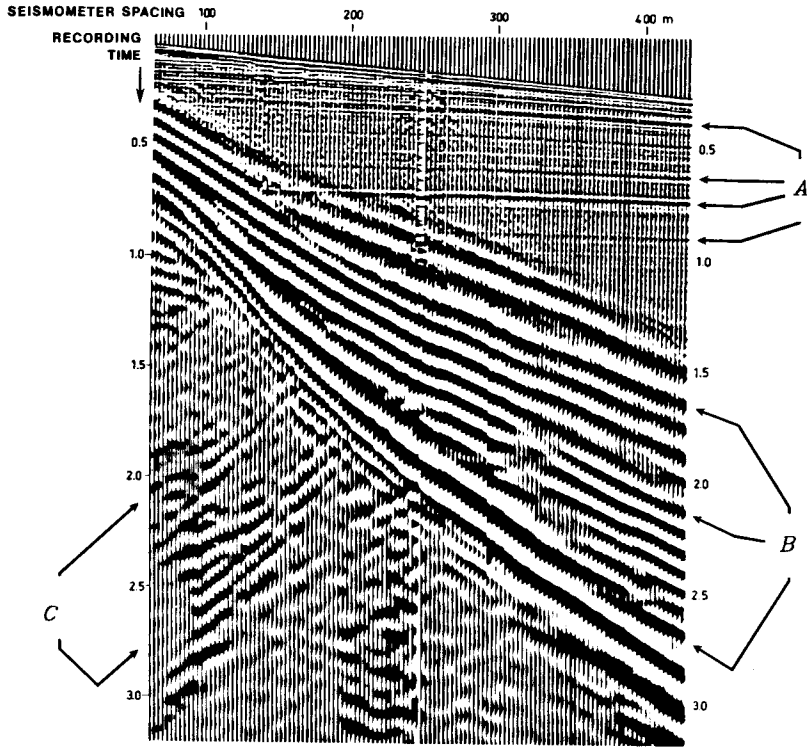


FIG. 13. Undocumented data from a recruitment brochure. This data may be assumed to be of textbook quality. The speed of sound in water is about 1500 m/sec. Identify the events at A, B, and C. Is this a common-shotpoint gather or a common-midpoint gather? (Shell Oil Company)

3.3 Experiment Sinking with the Double-Square-Root Equation

Exploding-reflector imaging will be replaced by a broader imaging concept, *experiment sinking*. A new equation called the double-square-root (DSR) equation will be developed to implement experiment-sinking imaging. The function of the DSR equation is to downward continue an entire seismic survey, not just the geophones but also the shots. After deriving the DSR equation, the remainder of this chapter will be devoted to explaining migration, stacking, migration before stack, velocity analysis, and corrections for lateral velocity variations in terms of the DSR equation.

Peek ahead at equation (9) and you will see an equation with two square roots. One represents the cosine of the wave *arrival* angle. The other represents the *takeoff* angle at the shot. One cosine is expressed in terms of k_g, the Fourier component along the geophone axis of the data volume in (s,g,t)-space. The other cosine, with k_s, is the Fourier component along the shot axis.

Our field seismograms lie in the (s,g)-plane. To move onto the (y,h)-plane inhabited by seismic interpreters requires only a simple rotation. The data could be Fourier transformed with respect to y and h, for example. Then downward continuation would proceed with equation (17) instead of equation (9).

The Experiment-Sinking Concept

The exploding-reflector concept has great utility because it enables us to associate the seismic waves observed at zero offset in many experiments (say 1000 shot points) with the wave of a single thought experiment, the exploding-reflector experiment. The exploding-reflector analogy has a few tolerable limitations connected with lateral velocity variations and multiple reflections, and one major limitation: it gives us no clue as to how to migrate data recorded at nonzero offset. A broader imaging concept is needed.

Start from field data where a survey line has been run along the x-axis. Assume there have been an infinite number of experiments, a single experiment consisting of placing a point source or shot at s on the x-axis and recording echoes with geophones at each possible location g on the x-axis. So the observed data is an upcoming wave that is a two-dimensional function of s and g, say $P(s,g)$. (Important practical questions about the actual spacing and extent of shots and geophones will be deferred until Chapter 4.)

Previous chapters have shown how to downward continue the *upcoming* wave. Downward continuation of the upcoming wave is really the same thing as downward continuation of the geophones. It is irrelevant for the continuation procedures where the wave originates. It could begin from an exploding reflector, or it could begin at the surface, go down, and then be reflected back upward.

To apply the imaging concept of experiment sinking, it is necessary to downward continue the sources as well as the geophones. The main idea here comes from the principle of reciprocity, which says that the same seismogram should be recorded if the locations of the shot and geophone are swapped.

(One reason for the validity of reciprocity is that no matter how complicated a geometrical arrangement, the speed of sound along a ray is the same in either direction. A more thorough consideration of seismic reciprocity is found in FGDP. There has been little material published on field experience with reciprocity, but some workshop abstracts may be found in *Geophysics*, vol. 46, no. 4. The abstracts of Rocca, Savelli, Michon, and Norotte particularly address the issue of validity under field conditions. No one seems to seriously question the use of reciprocity in the present application.)

We know how to downward continue geophones. Since reciprocity permits interchanging geophones with shots, we really know how to downward continue shots too.

Shots and geophones may be downward continued to different levels, and they may be at different levels during the process, but for the final result they are only required to be at the same level. That is, taking z_s to be the depth of the shots and z_g to be the depth of the geophones, the downward-continued survey will be required at all levels $z = z_s = z_g$.

The image of a reflector at (x,z) is defined to be the strength and polarity of the echo seen by the closest possible source-geophone pair. Taking the mathematical limit, this closest pair is a source and geophone located together on the reflector. The travel time for the echo is zero. This experiment-sinking concept of imaging is summarized by

$$Image(x,z) \;=\; Wave(s=x,\, g=x,\, z,\, t=0) \qquad (1)$$

For good quality data, that is, data that fits the assumptions of the downward-continuation method, energy should migrate to zero offset at zero travel time. Study of the energy that doesn't do so should enable improvement of the model. Model improvement usually amounts to improving the spatial distribution of velocity.

Review of the Paraxial Wave Equation

In Chapter 1 an equation was derived for paraxial waves. The assumption of a *single* plane wave means that the arrival time of the wave is given by a single-valued $t(x,z)$. On a plane of constant z, such as the earth's surface, Snell's parameter p is measurable. It is

$$\left.\frac{\partial t}{\partial x}\right]_z \;=\; \frac{dt}{dx} \;=\; \frac{\sin\theta}{v} \;=\; p \qquad (2a)$$

In a borehole there is the constraint that measurements must be made at a constant x, where the relevant measurement from an *upcoming* wave would be

$$\left.\frac{\partial t}{\partial z}\right]_x \;=\; \frac{dt}{dz} \;=\; \frac{\cos\theta}{v} \;=\; -\left[\frac{1}{v^2} - \left(\frac{dt}{dx}\right)^2\right]^{1/2} \qquad (2b)$$

Recall the time-shifting partial-differential equation and its solution U as some arbitrary functional form f:

$$\frac{\partial U}{\partial z} \;=\; -\frac{dt}{dz}\frac{\partial U}{\partial t} \qquad (3a)$$

$$U \;=\; f\left(t - \int_0^z \frac{dt}{dz}\,dz\right) \qquad (3b)$$

The partial derivatives in equation (3a) are taken to be at constant x, just as is equation (2b). After inserting we have

$$\frac{\partial U}{\partial z} \;=\; \left[\frac{1}{v^2} - \left(\frac{dt}{dx}\right)^2\right]^{1/2}\frac{\partial U}{\partial t} \qquad (4a)$$

Fourier transforming the wave field over (x,t), we replace $\partial/\partial t$ by $-i\omega$. Likewise, for the traveling wave of the Fourier kernel $\exp(-i\omega t + ik_x x)$, constant phase means that $dt/dx = k_x/\omega$. With this, (4a) becomes

$$\frac{\partial U}{\partial z} \;=\; -i\omega\left(\frac{1}{v^2} - \frac{k_x^2}{\omega^2}\right)^{1/2} U \qquad (4b)$$

The solutions to (4b) agree with those to the scalar wave equation unless v is a function of z, in which case the scalar wave equation has both upcoming and downgoing solutions, whereas (4b) has

only upcoming solutions. Chapter 2 taught is how to go into the lateral space domain by replacing ik_x by $\partial/\partial x$. The resulting equation is useful for superpositions of many local plane waves and for lateral velocity variations $v(x)$.

The DSR Equation in Shot-Geophone Space

Let the geophones descend a distance dz_g into the earth. The change of the travel time of the observed upcoming wave will be

$$\frac{dt}{dz_g} = -\left[\frac{1}{v^2} - \left(\frac{dt}{dg}\right)^2\right]^{1/2} \tag{5a}$$

Suppose the shots had been let off at depth dz_s instead of at $z=0$. Likewise then

$$\frac{dt}{dz_s} = -\left[\frac{1}{v^2} - \left(\frac{dt}{ds}\right)^2\right]^{1/2} \tag{5b}$$

Both (5a) and (5b) require minus signs because the travel time decreases as both geophones and shots move down.

Simultaneously downward project both the shots and geophones by an identical amount $dz = dz_g = dz_s$. The travel-time change is the sum of (5a) and (5b), namely,

$$dt = \frac{dt}{dz_g}dz_g + \frac{dt}{dz_s}dz_s = \left(\frac{dt}{dz_g} + \frac{dt}{dz_s}\right)dz \tag{6}$$

or

$$\frac{dt}{dz} = -\left\{\left[\frac{1}{v^2} - \left(\frac{dt}{dg}\right)^2\right]^{1/2} + \left[\frac{1}{v^2} - \left(\frac{dt}{ds}\right)^2\right]^{1/2}\right\} \tag{7}$$

This expression for dt/dz may be substituted into equation (3a):

$$\frac{\partial U}{\partial z} = +\left\{\left[\frac{1}{v^2} - \left(\frac{dt}{dg}\right)^2\right]^{1/2} + \left[\frac{1}{v^2} - \left(\frac{dt}{ds}\right)^2\right]^{1/2}\right\}\frac{\partial U}{\partial t} \tag{8}$$

Three-dimensional Fourier transformation converts upcoming wave data $u(t,s,g)$ to $U(\omega,k_s,k_g)$. Expressing equation (8) in Fourier space gives

$$\frac{\partial U}{\partial z} = -i\omega\left\{\left[\frac{1}{v^2} - \left(\frac{k_g}{\omega}\right)^2\right]^{1/2} + \left[\frac{1}{v^2} - \left(\frac{k_s}{\omega}\right)^2\right]^{1/2}\right\}U \tag{9}$$

Recall the origin of the two square roots in equation (9). One is the cosine of the arrival angle at the geophones divided by the velocity at the geophones. The other is the cosine of the takeoff angle at the shots divided by the velocity at the shots. With the wisdom of previous chapters we know how to go into the lateral space domain by replacing ik_g by $\partial/\partial g$ and ik_s by $\partial/\partial s$. To incorporate lateral velocity variation $v(x)$, the velocity at the shot location must be distinguished from the velocity at the geophone location. Thus,

$$\frac{\partial U}{\partial z} = \left\{\left[\left(\frac{-i\omega}{v(g)}\right)^2 - \frac{\partial^2}{\partial g^2}\right]^{1/2} + \left[\left(\frac{-i\omega}{v(s)}\right)^2 - \frac{\partial^2}{\partial s^2}\right]^{1/2}\right\}U \tag{10}$$

Equation (10) is known as the double-square-root (DSR) equation in shot-geophone space. It might be more descriptive to call it the experiment-sinking equation since it pushes geophones and shots downward together. Recalling the section on splitting and full separation we realize that the two square-root operators are commutative [$v(s)$ commutes with $\partial/\partial g$], so it is completely equivalent to downward continue shots and geophones separately or together. This equation will produce waves for the rays that are found on zero-offset sections but are absent from the exploding-reflector model.

The DSR Equation in Midpoint-Offset Space

The trouble with having the experiment-sinking equation in shot-geophone space is that it offers little geometrical insight. By converting to midpoint-offset space we hope to identify the familiar zero-offset migration part along with corrections for offset.

The transformation between (g,s) recording parameters and (y,h) interpretation parameters is

$$y = \frac{g+s}{2} \tag{11a}$$

$$h = \frac{g-s}{2} \tag{11b}$$

Travel time t may be parameterized in (g,s)-space or (y,h)-space. Differential relations for this conversion are given by the chain rule for derivatives:

$$\frac{dt}{dg} = \frac{dt}{dy}\frac{dy}{dg} + \frac{dt}{dh}\frac{dh}{dg} = \frac{1}{2}\left(\frac{dt}{dy} + \frac{dt}{dh}\right) \tag{12a}$$

$$\frac{dt}{ds} = \frac{dt}{dy}\frac{dy}{ds} + \frac{dt}{dh}\frac{dh}{ds} = \frac{1}{2}\left[\frac{dt}{dy} - \frac{dt}{dh}\right] \quad (12b)$$

Having seen how stepouts transform from shot-geophone space to midpoint-offset space, let us next see that spatial frequencies transform in much the same way. Clearly, data could be transformed from (s,g)-space to (y,h)-space with (11) and then Fourier transformed to (k_y,k_h)-space. The question is then, what form would the double-square-root equation (9) take in terms of the spatial frequencies (k_y,k_h)? Define the seismic data field in either coordinate system as

$$U(s,g) = U'(y,h) \quad (13)$$

This introduces a new mathematical function U' with the same physical meaning as U but, like a computer subroutine or function call, with a different subscript look-up procedure for (y,h) than for (s,g). Applying the chain rule for partial differentiation to (13) gives

$$\frac{\partial U}{\partial s} = \frac{\partial y}{\partial s}\frac{\partial U'}{\partial y} + \frac{\partial h}{\partial s}\frac{\partial U'}{\partial h} \quad (14a)$$

$$\frac{\partial U}{\partial g} = \frac{\partial y}{\partial g}\frac{\partial U'}{\partial y} + \frac{\partial h}{\partial g}\frac{\partial U'}{\partial h} \quad (14b)$$

and utilizing (11) gives

$$\frac{\partial U}{\partial s} = \frac{1}{2}\left[\frac{\partial U'}{\partial y} - \frac{\partial U'}{\partial h}\right] \quad (15a)$$

$$\frac{\partial U}{\partial g} = \frac{1}{2}\left[\frac{\partial U'}{\partial y} + \frac{\partial U'}{\partial h}\right] \quad (15b)$$

In Fourier transform space where $\partial/\partial x$ transforms to ik_x, equation (15), when i and $U=U'$ are cancelled, becomes

$$k_s = \frac{1}{2}(k_y - k_h) \quad (16a)$$

$$k_g = \frac{1}{2}(k_y + k_h) \quad (16b)$$

Equation (16) is a Fourier representation of (15). Substituting (16) into (9) achieves the main purpose of this section, which is to get the double-square-root migration equation into midpoint-offset coordinates:

$$\frac{d}{dz}U = -i\frac{\omega}{v}\left\{\left[1-\left[\frac{vk_y+vk_h}{2\omega}\right]^2\right]^{1/2} + \left[1-\left[\frac{vk_y-vk_h}{2\omega}\right]^2\right]^{1/2}\right\}U \quad (17)$$

Equation (17) is the takeoff point for many kinds of common-midpoint seismogram analyses. Some convenient definitions that simplify its appearance are

$$G = \frac{v\,k_g}{\omega} \quad (18a)$$

$$S = \frac{v\,k_s}{\omega} \quad (18b)$$

$$Y = \frac{v\,k_y}{2\,\omega} \quad (18c)$$

$$H = \frac{v\,k_h}{2\,\omega} \quad (18d)$$

Chapter 1 showed that the quantity $v\,k_x/\omega$ can be interpreted as the angle of a wave. Thus the new definitions S and G are the sines of the takeoff angle and of the arrival angle of a ray. When these sines are at their limits of ± 1 they refer to the steepest possible slopes in (s,t)- or (g,t)-space. Likewise, Y may be interpreted as the dip of the data as seen on a seismic section. The quantity H refers to stepout observed on a common-midpoint gather. With these definitions (17) becomes slightly less cluttered:

$$\boxed{\frac{d}{dz}U = -\frac{i\,\omega}{v}\left[\sqrt{1-(Y+H)^2} + \sqrt{1-(Y-H)^2}\right]U} \quad (19)$$

Most present-day before-stack migration procedures involve an interpretation of equation (19). Further analysis of it will explain the limitations of conventional processing procedures as well as

suggest improvements in the procedures.

EXERCISE

1. Adapt equation (17) to allow for a difference in velocity between the shot and the geophone.

3.4 The Meaning of the DSR Equation

The double-square-root equation controls most nonstatistical aspects of seismic data processing for petroleum prospecting. This equation, which was derived in the previous section, is not easy to understand because it is an operator in a four-dimensional space, namely, (z,s,g,t). We will approach it through various applications, each of which is like a picture in a space of lower dimension. In this section lateral velocity variation will be neglected (things are bad enough already!). Begin with

$$\frac{dU}{dz} = \frac{-i\omega}{v}\left(\sqrt{1-G^2} + \sqrt{1-S^2}\right) U \tag{1a}$$

$$\frac{dU}{dz} = \frac{-i\omega}{v}\left(\sqrt{1-(Y+H)^2} + \sqrt{1-(Y-H)^2}\right) U \tag{1b}$$

Zero-Offset Migration (H = 0)

One way to reduce the dimensionality of (1) is simply to set $H=0$. Then the two square roots become the same, so that they can be combined to give the familiar paraxial equation:

$$\frac{dU}{dz} = -i\omega\frac{2}{v}\sqrt{1-\frac{v^2k_y^2}{4\omega^2}}\, U \tag{2}$$

In both places in equation (2) where the rock velocity occurs, the rock velocity is divided by 2. Recall that the rock velocity needed to be halved in order for field data to correspond to the exploding-reflector model. So whatever we did by setting $H=0$, it had the effect of making the experiment-sinking concept functionally equivalent to the exploding-reflector concept.

Zero-Dip Stacking (Y = 0)

When dealing with the offset h it is common to assume that the earth is horizontally layered so that experimental results will be independent of the midpoint y. With such an earth the Fourier transform of all data over y will vanish except for $k_y = 0$, or, in other words, for $Y = 0$. The two square roots in (1) again become identical and the resulting equation is once more the paraxial equation:

$$\frac{dU}{dz} = -i\omega\frac{2}{v}\sqrt{1-\frac{v^2k_h^2}{4\omega^2}}\, U \tag{3}$$

Using this equation to downward continue hyperboloids from the earth's surface, we find the hyperboloids shrinking with depth, until the correct depth where best focus occurs is reached. This is shown in figure 1.

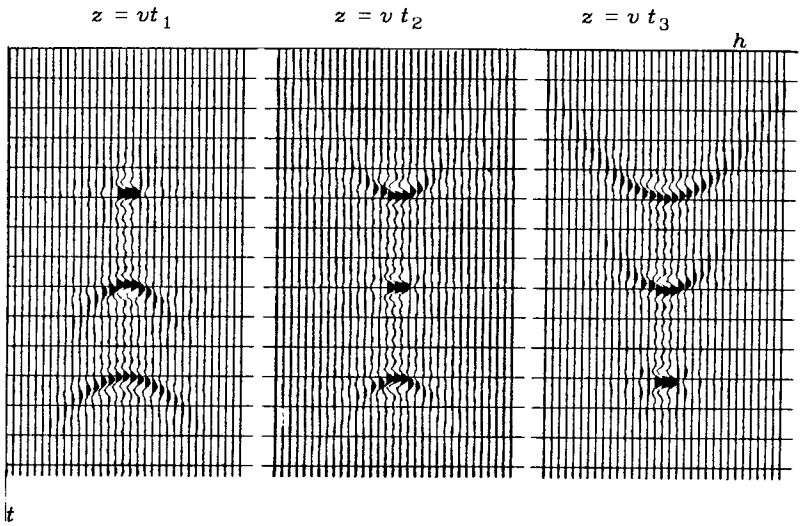


FIG. 1. With an earth model of three layers, the common-midpoint gathers are three hyperboloids. Successive frames show downward continuation to successive depths where best focus occurs. (Gonzalez)

The waves focus best at zero offset. The focus represents a downward-continued experiment, in which the downward continuation has gone just to a reflector. The reflection is strongest at zero travel time for a coincident source-receiver pair just above the reflector. Extracting the zero-offset value at $t = 0$ and abandoning the other offsets is a way of eliminating noise. (Actually it is a way of ***defining*** noise.) Roughly it amounts to the same thing as the conventional procedure of summation along a hyperbolic trajectory on the original data. Naturally the summation can be expected to be best when the velocity used for downward continuation comes closest to the velocity of the earth. Later offset space will be used to *determine* velocity.

Conventional Processing — the Separable Approximation

The *DSR* operator is formally defined as the parenthesized operator in equation (1b):

$$DSR(Y,H) \;=\; \sqrt{1-(Y-H)^2} + \sqrt{1-(Y+H)^2} \qquad (4)$$

There is a serious problem with this operator: it is ***not separable*** into a sum of an offset operator and a midpoint operator. ***Nonseparable*** means that a Taylor series for (4) contains terms like Y^2H^2. Such terms cannot be expressed as a function of Y plus a function of H. Nonseparability is a data-processing disaster. It implies that migration and stacking must be done simultaneously, not sequentially. The only way to recover pure separability would be to return to the space of S and G. (That is a drastic alternative, far from conventional processing. We will return to it later.)

Let us review the general issue of separability. The obvious way to get a separable approximation of the operator $\sqrt{1-X^2-Y^2}$ is to form a Taylor series expansion, and then drop all the cross terms. A more clever approximation is $\sqrt{1-X^2}+\sqrt{1-Y^2}-1$, which fits all Y exactly when $X=0$ and all X exactly when $Y=0$. Applying this idea to the *DSR* operator gives

$$SEP(Y,H) \;=\; 2 + [DSR(Y,0) - 2] + [DSR(0,H) - 2] \qquad (5a)$$

$$SEP(Y,H) \;=\; 2\,[1 + (\sqrt{1-Y^2} - 1) + (\sqrt{1-H^2} - 1)] \qquad (5b)$$

Notice that at $H = 0$ (5) becomes equal to the *DSR* operator. At $Y=0$ (5) also becomes equal to the *DSR* operator. Only when both H and Y are nonzero does *SEP* depart from *DSR*.

The splitting of (5) into a sum of three operators offers an advantage like the one offered by the 2-D Fourier kernel $exp(ik_y y + ik_h h)$, which has a phase that is the sum of two parts. It means that Fourier integrals may have either y or h nested on the inside. So downward continuation with *SEP* could be done in (k_h,k_y)-space as implied by (1b), or we could choose to Fourier transform to (h,k_y), (k_h,y), or (y,h) by appropriate nesting operations.

It is convenient to give familiar names to the three terms in (5b). The first is associated with time-to-depth conversion, the second with normal moveout, and the third with migration:

$$SEP(Y,H) \;=\; TD + MIG(Y) + NMO(H) \qquad (5c)$$

The approximation (5) can be interpreted as "standard processing." The first stage in standard processing is NMO correction. In (5) the *NMO* operator downward continues all offsets at the earth's surface, to all offsets at depth. Selecting zero offset is no more than abandoning all other offsets. Like stacking over offset, selecting zero offset reduces the amount of data under consideration.

Ordinarily the abandoned offsets are not migrated. (Alternately, a clever procedure for changing stacking velocities after migration involves migrating several offsets near zero offset.)

Since all terms in the *SEP* operator are interchangeable, it would be foolish and wasteful to use it to migrate all offsets before stack. The result of doing so should be identical to after-stack migration.

Various Meanings of H = 0

Recall the various forms of the stepout operator:

Forms of stepout operator $2H/v$		
ray trace	*Fourier*	*PDE*
$\frac{dt}{dh}$	$\frac{k_h}{\omega}$	$\partial_h^t = \int_{-\infty}^{t} dt \, \frac{\partial}{\partial h}$

Reciprocity suggests that travel time is a symmetrical function of offset; thus dt/dh vanishes at $h = 0$. In that sense it seems appropriate to apply equation (2) to zero-offset sections. More precisely, the ray-trace expression dt/dh strictly applies only when a single plane wave is present. Spherical wavefronts are made from the superposition of plane waves. Then the Fourier interpretation of H is slightly different and more appropriate. To set $\omega = 0$ would be to select a zero frequency component, a simple integral of a seismic trace. To set $k_h = 0$ would be to select a zero spatial-frequency component, that is, an integration over offset. Conventional stacking may be defined as integration (or summation) over offset along a hyperbolic trajectory. Simply setting $k_h = 0$ is selecting a hyperbolic trajectory that is flat, namely, the hyperbola of infinite velocity. Such an integration will receive its major contribution from the top of the data hyperboloid, where the data events come tangent to the horizontal line of integration. (For some historical reason, such a data summation is often called *vertical stack.)* Of the total contribution to the integral, most comes from a zone near the top, before the stepout equals a half-wavelength. The width of this zone, which is called a Fresnel zone, is the major factor contributing to the integral. See figure 2. The main differences between a zero-offset section and a vertical stack are the amplitude and a small phase shift. In practical cases they are unlikely to migrate in a significantly different way. It might be nice if we could find an equation to downward continue data that is stacked at velocities other than infinite velocity.

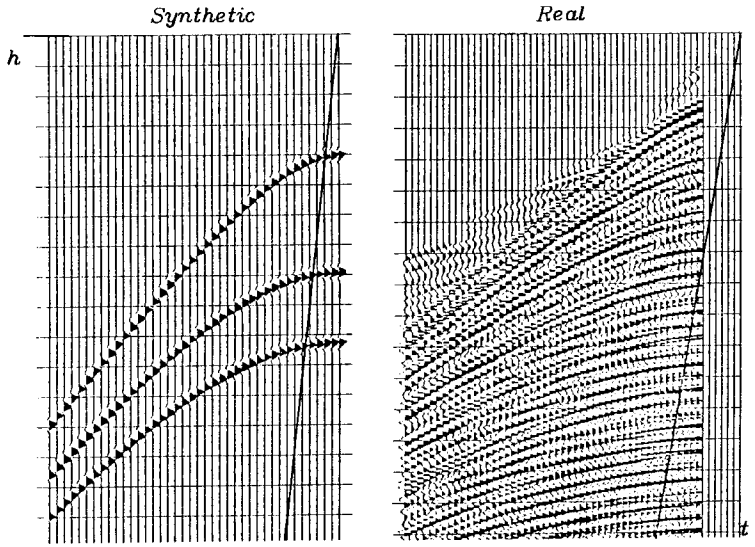


FIG. 2. The Fresnel zone for a vertical stack is defined to be the horizontal extent on the offset h axis within which the time shift of seismograms is about a half-wavelength. To be mathematically precise, it is necessary to specify a frequency. For practical purposes it is usually enough to look at zero crossings, or just to remember that typically $\Delta t/t \approx 2h^2/t^2v^2 \approx 1/100$, or $\cos 8° = .99$. (Gonzalez)

The partial-differential-equation point of view of setting $H = 0$ is identical with the Fourier view when the velocity is a constant function of the horizontal coordinate; but otherwise the PDE viewpoint is a slightly more general one. To be specific, but not cluttered, express equation (1) in 15-degree, retarded, space-domain form. Thus,

$$\left[\frac{\partial}{\partial z} + \frac{v}{-i\omega 8}\left[\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial h^2}\right]\right] U = 0 \qquad (6)$$

Integrate this equation over offset h. The integral commutes with the differential operators. Recall that the integral of a derivative is the difference between the function evaluated at the upper limit and the function evaluated at the lower limit. Thus,

$$\left\{\frac{\partial}{\partial z} + \frac{v}{-i\omega 8}\frac{\partial^2}{\partial y^2}\right\}\left\{\int U\, dh\right\} + \left.\frac{\partial U}{\partial h}\right|_{h=-\infty}^{h=+\infty} = 0 \qquad (7a)$$

The wave should vanish at infinite offset and so should its horizontal offset derivative. Thus the last term in (7a) should vanish. So, setting $H = 0$ means

$$(\textit{Paraxial operator})\ (\textit{vertical stack}) = 0 \qquad (7b)$$

A problem in the development of (7b) was that, twice, it was assumed that velocity is independent of offset: first, when the thin-lens term was omitted from (6), and second, when the offset integration operator was interchanged with multiplication by velocity. If the velocity depends on the horizontal x-axis, then it certainly depends on both midpoint and offset. In conclusion: If velocity changes slowly across a Fresnel zone, then setting $H = 0$ provides a valid equation for downward continuation of vertically stacked data.

Clayton's Cosine Corrections

A tendency exists to associate the sine of the earth dip angle with Y and the sine of the shot-geophone offset angle with H. While this is roughly valid, there is an important correction. Consider the dipping bed shown in figure 3.

The dip angle of the reflector is α, and the offset is expressed as the offset angle β. Clayton showed, and it will be verified, that

$$Y = \sin\alpha \cos\beta \qquad (8a)$$

$$H = \sin\beta \cos\alpha \qquad (8b)$$

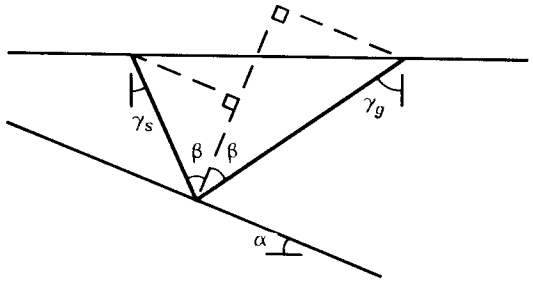


FIG. 3. Geometry of a dipping bed. Note that the line bisecting the angle 2β does not pass through the midpoint between g and s. (Clayton)

For small positive or negative angles the cosines can be ignored, and it is then correct to associate the sine of the earth dip angle with Y and the sine of the offset angle with H. At moderate angles the cosine correction is required. At angles exceeding 45° the sensitivities reverse and conventional wisdom is exactly opposite to the truth. The reader should be wary of informal discussions that simply associate Y with dip and H with velocity. "Larner's streaks" in Section 3.2 were an example of mixing the effects of dip and offset. Indeed, at steep dips the usual procedure of using H to determine velocity should be changed somehow to use Y.

Next, (8) will be proven. The source takeoff angle is γ_s, and the incident receiver angle is γ_g. First, relate γ_s and γ_g to α and β. Adding up the angles of the smaller constructed triangle gives

$$\left(\frac{\pi}{2} - \gamma_s - \alpha\right) + \beta + \frac{\pi}{2} = \pi$$

$$\gamma_s = \beta - \alpha \qquad (9a)$$

Adding up the angles around the larger triangle gives

$$\gamma_g = \beta + \alpha \qquad (9b)$$

To associate the angles at depth, α and β, with the stepouts dt/ds and dt/dg at the earth's surface requires taking care with the signs, noting that travel time increases as the geophone moves right and decreases as the shot moves right. Recall from Section 3.3 equations (16) and (18), the definitions of apparent

angles Y and H,

$$Y - H = S = \frac{v\, k_s}{\omega} = v\,\frac{dt}{ds} = -\sin\gamma_s = \sin(\alpha - \beta)$$

$$Y + H = G = \frac{v\, k_g}{\omega} = v\,\frac{dt}{dg} = +\sin\gamma_g = \sin(\alpha + \beta)$$

Adding and subtracting this pair of equations and using the angle sum formula from trigonometry gives Clayton's cosine corrections (8):

$$Y = \frac{1}{2}\sin(\alpha+\beta) + \frac{1}{2}\sin(\alpha-\beta) = \sin\alpha\,\cos\beta$$

$$H = \frac{1}{2}\sin(\alpha+\beta) - \frac{1}{2}\sin(\alpha-\beta) = \sin\beta\,\cos\alpha$$

Snell-Wave Stacks and CMP Slant Stacks

Setting the takeoff angle S to zero also reduces the double-square-root equation to a single-square-root equation. The meaning of $S = 0$ is that $k_s = 0$ or equivalently that the data should undergo a summation (without time shifting) over shot s. Such a summation simulates a downgoing plane wave. The imaging principle behind the summation would be to look at the upcoming wave at the arrival time of the downgoing wave. Historically, this is the idea behind my first finite-difference, wave-equation migrations of CDP stacks.

A Snell wave is a generalization of a downgoing plane wave at nonvertical incidence. The shots are not fired simultaneously, but sequentially at an inverse rate of $dt/ds = S/v$. This could be simulated with field data by summing across the (t,s)-plane along a line of slope dt/ds. Setting S to be some constant, say $S = v\ dt/ds$, also reduces the double-square-root equation to a paraxial wave equation, just the equation needed to downward continue the downgoing Snell wave experiment. Snell waves could be constructed for various $p = dt/ds$ values. Each could be migrated and imaged, and the images stacked over p. These ideas have been around longer than the DSR equation, yet they have gained no popularity. What could be the reason?

A problem with Snell wave simulation is that the wave field is usually sampled at coarse intervals along a geophone cable, which itself never seems to extend as far as the waves propagate. Crafty techniques to interpolate and extrapolate the data are frustrated by the fact that on a common-geophone gather, the top of the hyperbola need not be at zero offset. For dipping beds the earliest arrival is often off the end of the cable. So the data processing depends strongly on the missing data.

These difficulties provide an ecological niche for the common-midpoint slant stack, namely, $H = p\,v$. At common *midpoint* the hyperbolas go through zero offset with zero slope. The data are thus more amenable to the interpolation and extrapolation required for integration over a slanted line. Setting $H = p\,v$ yields

$$k_z = -\frac{i\,\omega}{v}\left[\sqrt{1-(Y+pv)^2} + \sqrt{1-(Y-pv)^2}\right] \qquad (10)$$

This has not reduced the DSR equation to a paraxial wave equation, but it has reduced the problem to a form manageable with the available techniques, such as the Stolt or phase-shift methods. Details of this approach can be found in the dissertation of Richard Ottolini.

Why Not Downward Continue in (S,G)-Space?

If the velocity were known and the only task were to migrate, then there would be no fundamental reason why the downward continuation could not be done in (S,G)-space. But the velocity really isn't well known. The sensitivity of migration to velocity error increases rapidly with angle, and angle accuracy is the presumed advantage of (S,G)-space. Furthermore, the finite extent of the recording cable and the tendency to spatial aliasing create the same problems with (S,G)-space migration as are experienced with Snell stacks. I see no fundamental reason why (S,G)-space migration should be any better than CMP slant stacks, and the aliasing and truncation situations seem likely to be worse. Less ambitious and more practical approaches to the wide-angle migration problem are found later in this chapter.

On the other hand, lateral velocity variation (if known) could demand that migration be done in (s,g)-space.

Still another reason to enter shot-geophone space would be that the shots were far from one another. Then the data would be aliased in both midpoint space and offset space. Geophones could still be downward continued, however. Downward continuation of shots would amount to calculating travel times, followed by time shifting traces. The arrival of high-density geophone arrays and three-dimensional interest may regenerate these old ideas.

3.5 Stacking and Velocity Analysis

Hyperbolic stacking over offset may be the most important computer process in the prospecting industry. It is more important than migration because it reduces the data base from a volume in (s,g,t)-space to a plane in (y,t)-space. At the present time few people who interpret seismic data have computerized seismic data movies, so most interpreters must have their data stacked before they can even look at it. Migration merely converts one plane to another plane. Furthermore, migration has the disadvantage that it sometimes compounds the mess made by near-surface lateral velocity variation and multiple reflections. Stacking can compound the mess too, but in bad areas nothing can be seen until the data is stacked. In addition to its other drawing points, stacking has as a by-product the determination of rock velocity.

Historically, stacking has been done using ray methods, and it is still being done almost exclusively in this way. Migration, on the other hand, is more often done using wave-equation methods, that is to say, by Fourier or finite-difference methods. Both migration and stacking are hyperbola-recognition processes. The advantages of wave-equation methods in migration have been many. Shouldn't these advantages apply equally to stacking? It would seem so, but current industrial practice does not bear this out. The reasons are not yet clear. Except for the central role of its subject matter, this section really belongs to a research monograph with the facetious title "Theory That Should Work Out Soon." The wave-equation stacking and velocity-determination methods described in this Section are indeed ingenious ones. Perhaps they have not yet been satisfactorily tested, or perhaps they are just imperfectly assembled. The reader can guess, and time will tell.

One possible reason why much of this theory is not in routine industrial use is that the issue of stacking to remove redundancy may be more appropriately a statistical problem than a physical one. To allow for this contingency I have included a bit on "wave-equation moveout," a way of deferring statistical analysis until after downward continuation. Another possibility is that the problems of missing data off the ends of the recording cable and spatial aliasing within the cable may be more flexibly attacked by ray methods than by wave-equation methods. For this contingency I have included a brief subsection on data restoration. Whatever the case, the data-manipulation procedures in this chapter should be helpful.

Normal Moveout (NMO)

To do a conventional velocity analysis and stack, you need some skill at interpolation, and you need a table of computed travel times as functions of offset and depth for some velocity model. Wide offset traces should have their time axes stretched so that the arrival of an event at wide offset occurs at the same time as the event at zero offset. This stretching is called normal-moveout correction (NMO), a technique that was first discussed in Section 3.0. After, in some cases, intermediate processing, such as balancing amplitudes and spectra, the data is averaged over offset. Presumably, the more closely the earth velocity matches the velocity in the travel-time table, the better (bigger) will be the result.

In principle, NMO converts common-midpoint gathers, one of which, say, is denoted by $P(h,t)$, to an earth model, say,

$$Q(h,z) \quad = \quad earth(z) \times const(h) \tag{1}$$

Actually, $Q(h,z)$ doesn't turn out to be a constant function of h, but that is the goal.

The NMO procedure can be regarded as a simple copying. Conceptually, it is easy to think of copying every point of the (h,t)-plane to its appropriate place in the (h,z)-plane. Such a copying process could be denoted as

$$Q[h,z(h,t)] \quad = \quad P(h,t) \tag{2}$$

Care must be taken to avoid leaving holes in the (h,z)-plane. It is better to scan every point in the output (h,z)-plane and find its source in the (h,t)-plane. With a table $t(h,z)$, data can be moveout corrected by the copying operation

$$Q(h,z) \quad = \quad P[h,t(h,z)] \tag{3}$$

Using the terminology of this book, the input $P(h,t)$ to the moveout correction is called a CMP gather, and the output Q is called a CDP gather. Mathematically, the creation of Q from P is a ***linear*** mapping because the original data Q could be decomposed into arbitrary parts, the parts could be separately NMOed, and the outputs would superpose as they should.

In practice, the first step in generating the travel-time tables is to change the depth-variable z to a vertical travel-time-variable τ. So the required table is $t(h,\tau)$. To get the output data for location (h,τ) you take the input data at location (h,t). The most straightforward and reliable way to produce this table seems to be to march down in steps of z, really τ, and trace rays. That is, for

various fixed values of Snell's parameter p, you compute $t(p,\tau)$ and $h(p,\tau)$ from $v(\tau)$ by integrating

$$\frac{dt}{d\tau} = \frac{dz}{d\tau}\frac{dt}{dz} = v\,\frac{1}{v\cos\theta} = \frac{1}{\sqrt{1-p^2v^2}} \qquad (4)$$

$$\frac{dh}{d\tau} = \frac{dz}{d\tau}\frac{dh}{dz} = v\tan\theta = \frac{p\,v^2}{\sqrt{1-p^2v^2}} \qquad (5)$$

Given $t(p,\tau)$ and $h(p,\tau)$, iteration and interpolation are required to eliminate p and find $t(h,\tau)$. It sounds awkward — and it is — because at wide angles there can be multiple arrivals. But once the job is done you can save the table and reuse it many times.

Mutes

An important part of conventional processing is the definition of a *mute*. A mute is a weighting function used to suppress some undesirable portions of the data. Weights and mutes have a substantial effect on the quality of a stack. Therefore it is not surprising that in practice, they are the subject of much theorizing and experimentation.

Often the mute is a one-dimensional function of $r = h/t$. Reasons can be given to mute data at both large and small values of r. At large values of r there are wide-angle problems which it might be nice to avoid, such as refractions. At small values of r is found the energy that remains in the vicinity of the shot, such as falling dirt or water or slow ground roll. Even if noise problems are disregarded, two reasons may be given to justify the common practice of scaling data upward in proportion to offset. First, think of the true three-dimensional problem. Stacking should focus waves spreading three-dimensionally from the point source. (Migration corrects for the two-dimensional focusing of "bent bedding.") So in stacking, you really wish to integrate over a hyperbola of revolution. The second justification for scaling data upward in proportion to offset is that there is less velocity information near zero offset, because there is necessarily little moveout.

Linearity Allows Postponing Statistical Estimation Until Last.

The linearity of wave-equation data processing allows us to decompose a dataset into parts, process the parts separately, then recombine them. The result is the same as if they were never separated.

For example, suppose a CMP gather is divided into two parts, say, inner traces A and outer traces B. Let $(A,0)$ denote a CMP gather where the outer traces have been replaced by zeros. Likewise, $(0,B)$ could be another gather where the inner traces have been replaced by zeros. We could downward continue $(A,0)$ and then also downward continue $(0,B)$. After downward continuation, $(A,0)$ and $(0,B)$ could be added. Alternately, we could pause, do some thinking about statistics, and then choose to combine them with some weighting function. Figure 1 shows a dataset of three traces decomposed into three datasets, one for each trace. Semicircles depict the separate downward continuation of each trace. Each semicircle goes through zero offset, giving the appropriately stretched, NMOed trace.

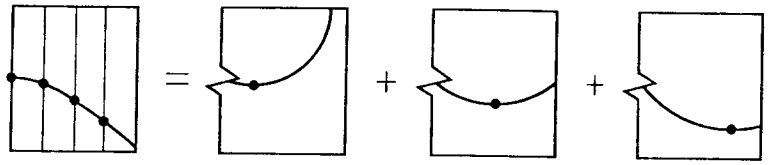


FIG. 1. A three-trace CMP gather decomposed by traces, each focused (migrated) into a semicircle going through zero offset at the apex of the hyperbola.

The idea of using a weighting function is a drastic departure from our previous style of analysis. It represents a disturbing recognition that we have been neglecting something important in all scientific analysis, namely, statistics ! What are the ingredients that go into the choice of a weighting function? They are many. Signal and noise variances play a role. Some channels may be noisy or absent. When final display is contemplated, it is necessary to consider human perception and the need to compress the dynamic range, so that small values can be perceived. Dynamic-range compression must be considered not only in the obvious (h,t)-space, but also in frequency space, dip space, or any other space in which the wave fields may get too far out of balance. (One such "other space" that I'd like to be able to define is (v,t)-space, where v is velocity.)

There are many ways to decompose a dataset. The choice depends on your statistical model and your willingness to repeat the processing many times. Perhaps the parts of the data gather should be decomposed not by their h values but by their values of $r = h/t$. Clearly, there is a lot to think about.

Lateral Interpolation and Extrapolation of a CMP Gather

Practical problems dealing with common-midpoint gathers arise because of an insufficient number of traces. *Truncation* problems are those that arise because the geophone cable has a fixed length, not as long as the distance over which seismic energy propagates. Figure 2 shows why cable truncations are a problem for conventional, ray-trace, stacking methods as well as for wave-equation methods. *Aliasing* problems are those that arise because shots and geophones are not close enough together. Spatial aliasing of data on the offset axis seems to be a more serious problem for wave-equation methods than it is for ray-trace methods. The reason is that normal-moveout correction reduces the spatial frequencies. *Gaps* in the data, resulting from practical problems with the geophones, cable, and access to the terrain, are also frequently a snag.

Here these problems will all be attacked together with a systematic approach to estimating missing traces. The technique to be described is the simplest member of a more general family of *missing data* estimation procedures currently being developed at the Stanford Exploration Project.

First do normal-moveout correction, that is, stretch the time axis to flatten hyperbolas. The initial question is what velocity to use for the normal-moveout correction. For trace interpolation the appropriate moveout velocity turns out to be that of the *dominating energy on the gather*. On a given dataset this velocity could be primary velocity at some times and multiple velocity at other times. The reason for such a nonphysical velocity is this: The strong events must be handled well, in order to save the weak ones. Truncations of weak events can be ignored as a "second-order" problem. The practical problem is usually to suppress strong water-velocity events in the presence of weak sedimentary reflections, particularly at high frequencies. In principle, we might be seeking weak $P-SV$ waves in the presence of strong $P-P$ waves.

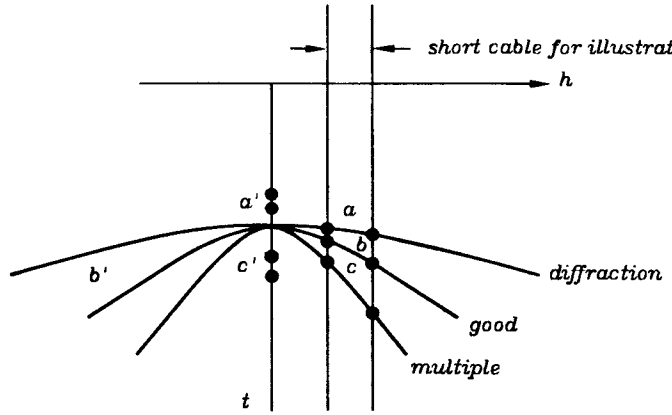


FIG. 2. Normal moveout at the earth velocity brings the cable truncations on good events to a good place, causing no problems. The cable truncations of diffractions and multiples, however, move to a' and c', where they could be objectionable. Such corruption could make folly of sophisticated time-series analysis of the waveform found on a CDP stack.

After NMO, the residual energy should have little dip, except of course where missing data, now replaced by zeros, forces the existing data to be broad-banded in spatial frequency. In order to improve our view of this badly behaved energy, the data is passed through a "badpass" filter, such as the high-pass recursive dip filter

$$\frac{\frac{k^2}{-i\omega}}{\alpha + \frac{k^2}{-i\omega}} \tag{6}$$

discussed in Section 2.5. Notice that this filter greatly weakens the energy with small k, namely, the energy that was properly moveout corrected. On the other hand, near the missing traces, notice that the spectrum should be broad-band with k and that such energy passes through the filter with almost unit gain.

The output from the "badpass" filter is now ready to be subtracted from the data. The subtraction is done selectively. Where recorded data exists, nothing is subtracted.

The first iteration is performed. Next the whole process is repeated, and iterated. Convergence is finally achieved when

nothing comes out of the badpass filter at the locations where data was not recorded. An example of this process can be found in figure 3.

FIG. 3. Space reserved for future figure by Jeff Thorson.

The above procedure is limited because it ignores the possibility that several velocities may be simultaneously present on a dataset. To really do a good job of extending such a dataset may require a parsimonious model concept and a velocity spectral concept such as the one developed later in Section 5.2, which deals with linear moveout.

The above procedures have also ignored the possibility of dip in the midpoint direction. This is taken up in Section. 3.6.

The (z,t)-Plane Method

In the 15° continuation equation $U_{zt} = -\frac{1}{2}v U_{hh}$, scaling the depth z is indistinguishable from scaling the velocity. Thus, downward continuation with the wrong velocity is like downward continuation to the wrong depth. Stephen M. Doherty was the first to use this idea in a velocity estimation scheme. See figure 4.

The idea is to downward continue with a preliminary velocity model and to display the zero-offset trace, a function of t', at all travel-time depths τ. If the maximum amplitude occurs at $t' = \tau$, then your preliminary model is good. If the maximum is shifted, then you have some analysis to do before you can say what velocity should be used on the next iteration.

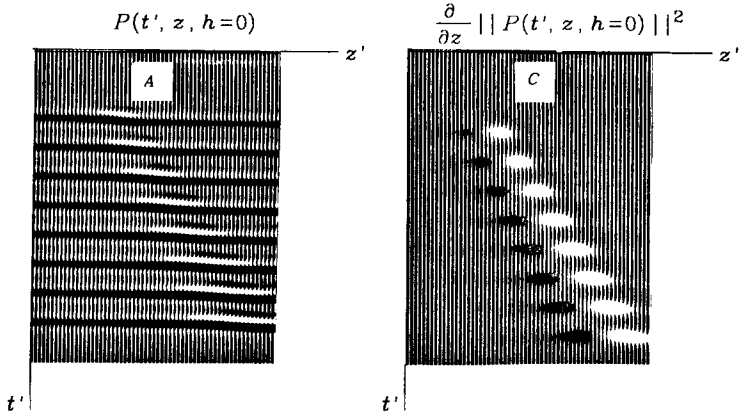


FIG. 4. Two displays of the (z,t)-plane at zero offset. The earth model is eight uniformly spaced reflectors under a water layer [a family of hyperboloids in (h,t) at $z=0$]. The left display is the zero-offset trace. The amplitude maximum at the focus is not visually striking, but the phase shift is apparent. The right display is the z-derivative of the envelope of the zero-offset trace. A linear alignment along $z=vt$ is more apparent. (Doherty)

Splitting a Gather into High- and Low-Velocity Components

A process will be defined that can partition a CMP gather, ***both reflections and refractions***, into one part with RMS velocity greater than that of some given model $\bar{v}(z)$ and another part with velocity less than $\bar{v}(z)$.

After such a partitioning, the low-velocity noise could be abandoned. Or the earth velocity could be found through iteration, by making the usual assumption that the velocity spectrum has a peak at earth velocity. As will be seen later, various data interpolation, lateral extrapolation, and other statistical procedures are also made possible by the linearity and invertibility of the partitioning of the data by velocity.

The procedure is simple. Begin with a common-midpoint gather, zero the negative offsets, and then downward continue according to the velocity model $\bar{v}(z)$. The components of the data with velocity less than $\bar{v}(z)$ will overmigrate through zero offset to negative offsets. The components of the data with velocity

greater than $v(z)$ will undermigrate. They will move toward zero offset but they will not go through. So the low-velocity part is at negative offset and the high-velocity part is at positive offset. If you wish, the process can then be reversed to bring the two parts back to the space of the original data.

Obviously, the process of multiplying data by a step function may create some undesirable diffractions, but then, you wouldn't expect to find an infinitely sharp velocity cutoff filter. Clearly, the false diffractions could be reduced by using a ramp instead of a step. An alternative to zeroing negative h would be to go into (k_h,ω)-space and zero the two quadrants of sign disagreement between k_h and ω.

This partitioning method unfortunately does not, by itself, provide a velocity spectrum. Energy away from $h=0$ is unfocused and not obviously related to velocity. The need for a velocity spectrum motivates the development of other processes.

Reflected Refractions on Sections

It is common for an interpreter looking at a stacked section to identify a reflected refraction. This is just a hyperbolic asymptote seen in (y,t)-space. This event provides a quick, easy, and accurate velocity estimate, namely, $v = 2\,dy/dt$. From a processing point of view, such a velocity measurement is unexpected, because automatic processing extracts all velocity information in offset space, a space which many interpreters prefer to leave inside the computer. Of course for a reflected refraction to be identified a special geological circumstance must be present: a scatterer strong enough to have its hyperbolic asymptote visible. Furthermore, the point scatterer must also be strong enough to get through the typical suppression effects of shot and geophone patterns and CDP stacking. The most highly suppressed events, water velocity and ground roll, are just those whose velocities are most often apparent on stacked sections. (Recall that the reason for this is that, seen from the side, a point scatterer is like a dipping bed. It has a higher stacking velocity.) Some strong reflected refraction energy was present on the common-shot profile shown in Section 3.2.

Processing seems to ignore or discriminate against the reflected refraction, yet the reflected refraction is often seen and used. There must be an explanation. Perhaps there is also a latent opportunity. From a theoretical point of view, Clayton's cosines showed that at wide angles the velocity and dip sensitivity of midpoint and offset are exchanged. At late times another important theoretical aspect presents itself. The aperture of a cable length can be much less than the width of a migration hyperbola. So, although it is easy to find an asymptote in midpoint space, there is little time shift at the end of the cable in offset space.

What processing could take advantage of lateral reflectivity and could enhance, instead of suppress, our ability to determine velocity in this way? Start by stacking at a high velocity. Then use the idea that at any depth z, the power spectrum of the data $U(\omega,k_y)$ should have a cutoff at the evanescent stepout $p(z) = k_y/\omega = 1/v(z)$. This would show up in a plot the power spectrum U^*U, or better yet the dip spectrum, as a function of depth. Perhaps it would be still better to visually inspect the seismic section itself after filtering in dip space about the expected velocity.

The wave-extrapolation equation is an all-pass filter, so why does the power spectrum change with depth? It changes because at any depth z it is necessary to exclude all the seismic data before $t=0$. This data should be zeroed before computing the dip spectrum. The procedure is depicted in figure 5.

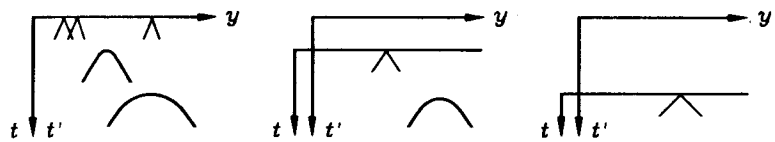


FIG. 5. The dip-spectrum method of velocity determination. To find the velocity at any depth, seek the steepest dip on the section at that depth. On the left, at the earth's surface, you see the surface ground roll. In frames B and C the slowest events are the asymptotes of successively faster hyperbolas.

To my knowledge this method has never been tried. I believe it is worth some serious testing. Even in the most layered of geological regions there are always faults and irregularities to illuminate the full available spectrum. Difficulty is unlikely to come from weak signals. More probably, the potential for difficulty lies in the near-surface irregularities.

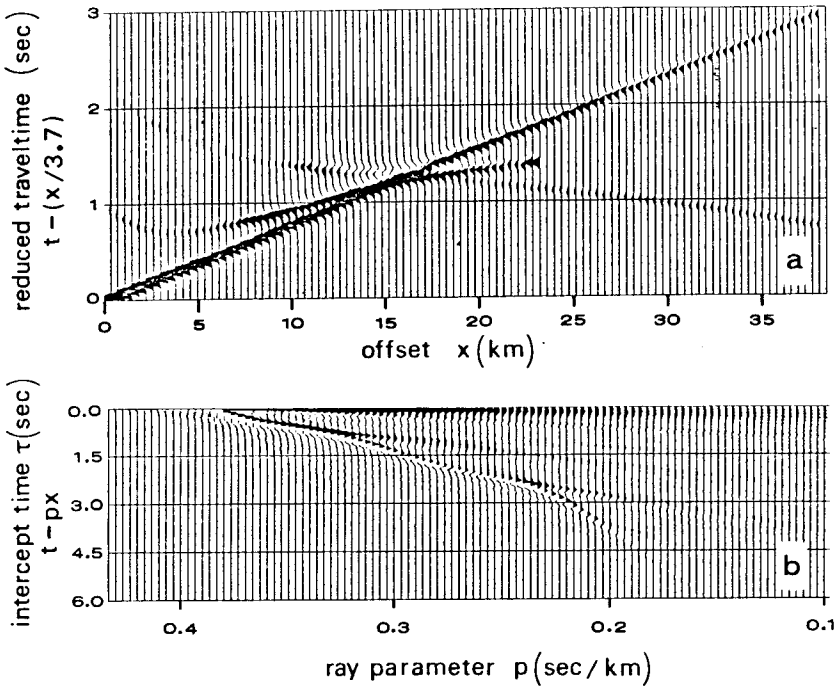


FIG. 6. The upper figure (a) contains a synthetic refraction profile (plotted with linear moveout). The data is transformed by slant stack to the lower half of the figure (b). The result of downward continuation of this slant-stacked wave field (b) is shown in figure 7. (Clayton & McMechan)

Clayton and McMechan's Method for Refractions on Gathers

The same process for getting velocity from reflected refractions on sections could be used on common-midpoint gathers. On gathers, however, you must also take into account the fact that downward continuation focuses energy on zero offset. The focus is not a featureless point. Taking original data to consist of a refraction only, with no reflection, the focus is a concentrated patch of energy oriented with the same stepout dt/dh as the original unfocused refraction. Summing through the focus at all possible orientations (slant stack) transforms the data — say, $U(h,\tau)$ — to dip space — say, $U(p,\tau)$. The velocity of the earth at travel-time depth τ is found by seeking maxima on $p(z) = 1/v(z)$.

Clayton and McMechan invented and developed the velocity determination procedure described here. They actually did the downward continuation and the slant stack in opposite orders, but I don't believe it makes any difference. Figures 6 and 7 show one of their examples.

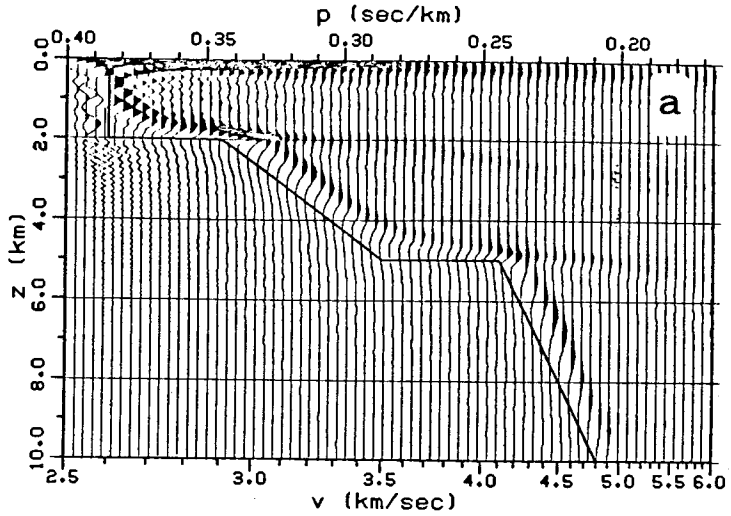


FIG. 7. The result of downward continuation of the slant-stacked wave field in figure 6b with the correct velocity-depth function (the solid line). (Clayton & McMechan)

Because it is limited to refractions, the method of Clayton and McMechan may seem to be rather remote from routine industrial velocity analysis. But it has an important feature. It is a completely linear and invertible function of the data. The inverse operation, which can be used for making synthetic data, is just an inverse slant stack followed by upward continuation. Besides its ability to make synthetic data, the method offers processing advantages that stem from its ability to partition data by velocity

and then return to the space of the original data.

EXERCISES

1. Assume that the data $P(y,h,t)$ is constant with respect to midpoint y. Given a common-midpoint gather $P(h,t,z=0)$, define a Stolt-type integral transformation from $P(h,t,z=0)$ to $P(h=0,t,z)$ based on the double-square-root equation:

$$\frac{\partial}{\partial z}P \quad = \quad -i\,\frac{\omega}{v}\left[\sqrt{1-(Y+H)^2}+\sqrt{1-(Y-H)^2}\right]P$$

   As with Stolt migration, your answer should be expressed as a 2-D inverse Fourier transform.

2. Start with a CDP gather $u(h,t)$ defined (by reciprocity) at both positive and negative values of h. Describe the effect of the following operations: Fourier transform to $U(k_h,\omega)$; multiply by $1 + sgn(\omega)\,sgn(k_h)$; transform back to (h,t)-space.

3.6 Migration with Velocity Estimation

We often face the three complications dip, offset, and unknown velocity at the same time. The double-square-root equation provides an attractive avenue when the velocity is known, but when it isn't, we are left with velocity-estimation procedures, such as that in the last section, ***which assume no dip***. In this section a means will be developed of estimating velocity in the presence of dip.

Dip Moveout — Sherwood's Devilish

Recall (Section 3.2) Levin's expression for the travel time of the reflection from a bed dipping at angle α from the horizontal:

$$t^2v^2 \quad = \quad 4(y-y_0)^2\sin^2\alpha + 4h^2\cos^2\alpha \qquad (1)$$

In (h,t)-space this curve is a hyperbola. Scaling the velocity by $\cos\alpha$ makes the travel-time curve identical to the travel-time curve of the dip-free case. This is the conventional approach to stacking and velocity analysis. It is often satisfactory. Sometimes it is unsatisfactory because the dip angle is not a single-valued function of space. For example, near a fault plane there will be diffractions. They are a superposition of all dips, each usually being weaker than the reflections. Many dips are present in the same place.

A ***devilishly*** clever insight of J.W.C. Sherwood (1976) showed that it is possible to perform appropriate dip-dependent time shifts before the usual NMO and velocity analysis. He called his process *Devilish*, an acronym for "dipping-event velocity inequalities licked." The process is less cryptically described as ***prestack partial migration***, but it has finally come to be called simply ***dip moveout*** (DMO).

The idea is to replace $\cos^2\alpha$ in (1) by $1-\sin^2\alpha$ and then interpret $\sin^2\alpha \approx Y^2$ using the familiar representations of $Y = v\,k_y/2\omega$. Refer to Clayton's cosine corrections in Section 3.4 for the more precise result $Y = \sin\alpha\cos\beta$. The approximation $\cos\beta = 1$ limits us to small offsets.

Consider the processing implied for the special case of constant velocity and modest angles. Let t_0 be defined by $(y-y_0)\sin\alpha$. Then expand the square root in (1):

$$t \quad \approx \quad t_0 + \frac{2h^2}{t_0v^2}\cos^2\alpha \qquad (2)$$

$$t \approx t_0 + \frac{2h^2}{t_0 v^2}(1 - \sin^2\alpha)$$

$$t \approx t_0 + \frac{2h^2}{t_0 v^2} - \frac{2h^2}{t_0 v^2}\sin^2\alpha$$

$$t \approx t_0 + \frac{2h^2}{t_0 v^2} - \frac{2h^2}{t_0 v^2} Y^2$$

$$t \approx t_0 + \frac{2h^2}{t_0 v^2} - \frac{h^2}{2t_0}\frac{k_y^2}{\omega^2}$$

$$t \approx t_0 + \Delta t_{NMO} - \Delta t_{DMO} \qquad (3)$$

Notice that Δt_{DMO} is velocity independent.

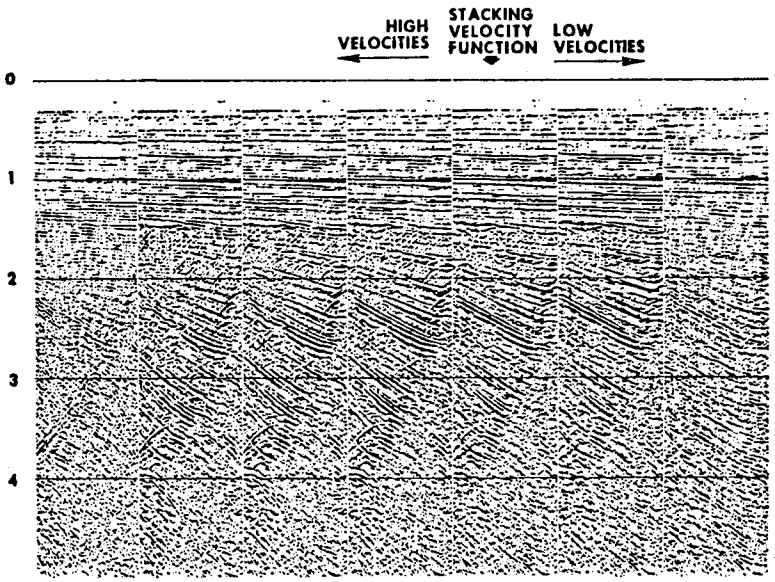


FIG. 1. Conventional stacks with varying velocity. (Sherwood, Digicon)

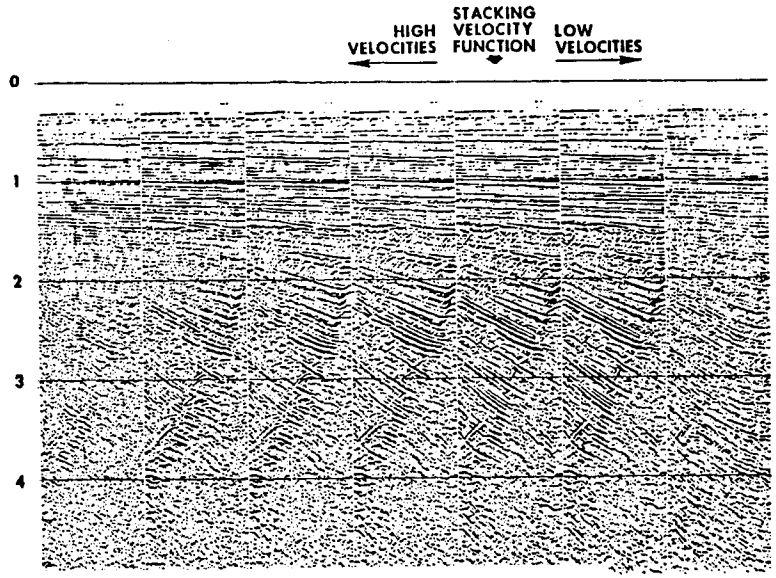


FIG. 2. *Devilish* stacks with varying velocity. (Sherwood, Digicon)

Figure 1 contains a panel from a stacked section. The panel is shown several times; each time the stacking velocity is different. It should be noted that at the low velocities, the horizontal events dominate, whereas at the high velocities, the steeply dipping events dominate. After the *Devilish* correction was applied, the data was restacked as before. Figure 2 shows that the stacking velocity no longer depends on the dip. This means that after *Devilish*, the velocity may be determined without regard to dip. In other words, events with all dips contribute to the same consistent velocity rather than each dipping event predicting a different velocity. So the *Devilish* process should provide better velocities for data with conflicting dips. And we can expect a better final stack as well.

It is a difficult task to carry through a precise analysis that incorporates a depth-variable velocity and a wide range of angles. So we may hope that this is possible, and so we may get an idea of what a DMO operator might look like, we next look at Rocca's smile construction.

Rocca's Smear Operator

Fabio Rocca developed a clear conceptual model for Sherwood's dip corrections. Figure 3 illustrates Rocca's concept of a prestack partial-migration operator. Imagine a constant-offset section $P(t,y,h=h_0)$ containing an impulse function at some particular (t_0,y_0). The earth model implied by this data is a reflector shaped like an ellipse, with the shot point at one focus and the receiver at the other. Starting from this earth model, a zero-offset section is made by forward modeling — that is, each point on the ellipse is expanded into a hyperbola. Combining the two operations — constant-offset migration and zero-offset diffraction — gives the Rocca operator.

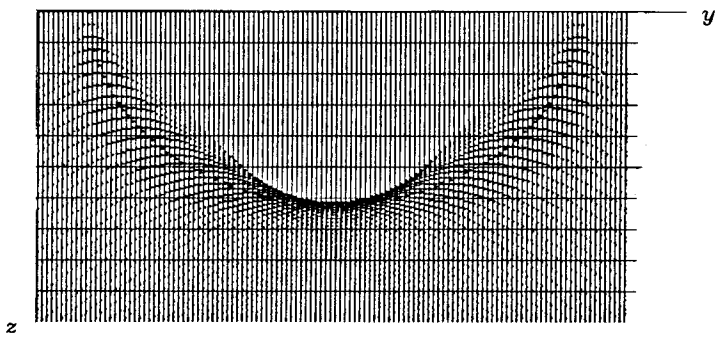


FIG. 3. Rocca's prestack partial-migration operator is a superposition of hyperbolas, each with its top on an ellipse. Convolving (over midpoint) Rocca's operator onto a constant-offset section converts it to a zero-offset section. (Gonzalez)

The Rocca operator is the curve of osculation in figure 3, the smile-shaped curve where the hyperbolas reinforce one another. If the hyperbolas in figure 3 had been placed everywhere on the ellipse instead of at isolated points, then the osculation curve would be the only thing visible (and you wouldn't be able to see where it came from).

The analytic expression for the travel time on the Rocca smile turns out to be the end of a narrow ellipse, shown in figure 4. The equation for the narrow ellipse is the velocity-independent curve

$$1 = \frac{(y - y_0)^2}{h^2} + \frac{t^2}{t_0^2} \qquad (4)$$

The Rocca operator is not completely velocity-independent, however, because the curve should be cut off at $dt/dy = 2/v$.

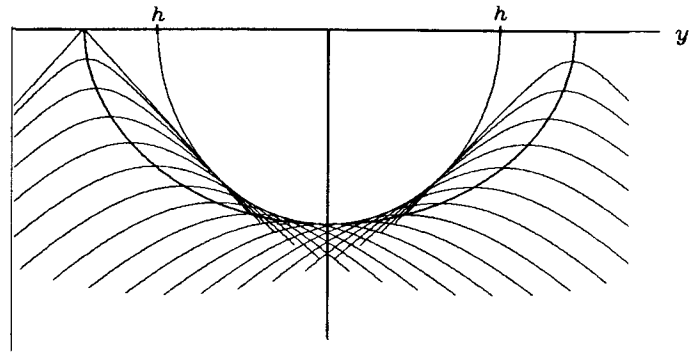


FIG. 4. Rocca's smile. (Ronen)

The Rocca operator transforms a constant-offset section into a zero-offset section. This transformation achieves two objectives: first, it does normal-moveout correction; second, it does Sherwood's dip corrections. The operator of figure 3 is convolved across the midpoint axis of the constant-offset section, giving as output a zero-offset section at just one time, say, t_0. For each t_0 a different Rocca operator must be designed. The outputs for all t_0 values must be superposed. Figure 5 shows a superposition of several Rocca smiles for several values of t_0.

From figure 5 we see that the energy in the operator is concentrating near the bottom. In the limiting case that the ellipse tends to a circle, that is to say, where h/vt_0 is small, the energy does all go to the bottom. When all the energy is concentrated near a single point, the Rocca operator becomes a delta function.

This operator is particularly attractive from a practical point of view. Instead of using a big, wide ellipse and doing the big job of migrating each constant-offset section, only the narrow, little Rocca operator is needed. After compensating each offset to zero

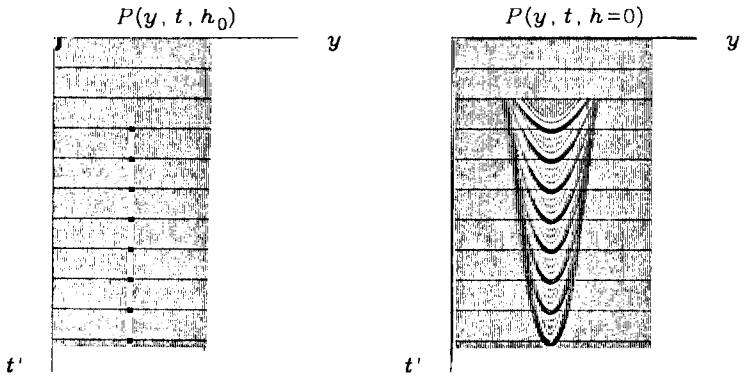


FIG. 5. The prestack partial-migration operator. A model (left) of impulses at various depths at some fixed offset, say, $P(y,t,h_0)$, and the dip moveout filter response $P(y,t,0)$ (right). (Yilmaz)

offset, velocity is determined by the normal-moveout residual; then data is stacked and migrated.

The *narrowness* of the Rocca ellipsoid is an advantage in two senses. Practically, it means that not many midpoints need to be brought into the computer main memory before velocity estimation and stacking are done. (The low-amplitude tails would be ignored.) More fundamentally, since the operator is so compact, it does not do a lot to the data. This is important because the operation is done at an early stage, before the velocity is well known. So it may be satisfactory to choose the velocity for the Rocca operator as a constant, regional value, say, 2.5 km/sec.

More desirable than the exact expression for the travel-time curve would be a Fourier representation for the operator itself. That would help us implement it. For this we turn to Ottolini's radial traces.

Ottolini's Radial Traces

Ordinarily we regard a common-midpoint gather as a collection of seismic traces, that is, a collection of time functions, each one for some particular offset h. But this (h,t) data space could be represented in a different coordinate system. A system with some nice attributes is the radial-trace system introduced by Turhan Taner. In this system the traces are not taken at constant h, but at constant angle. The idea is illustrated in figure 6.

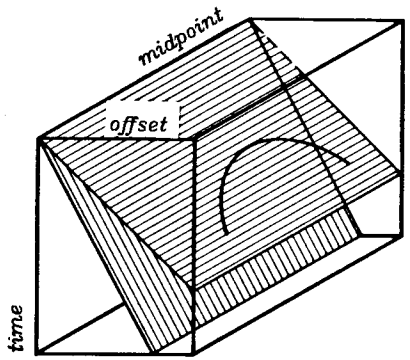


FIG. 6. Inside the data volume of a reflection seismic traverse are planes called *radial-trace sections*. A point scatterer inside the earth puts a hyperbola on a radial-trace section.

Besides having some theoretical advantages, which will become apparent, this system also has some practical advantages, notably: (1) the traces neatly fill the space where data are nonzero; (2) the traces are close together at early times where wavelengths are short, and wider apart where wavelengths are long; and (3) the energy on a given trace tends to represent wave propagation at a fixed angle. The last characteristic is especially important with multiple reflections. But for our purposes the best attribute of radial traces is still another one.

Richard Ottolini noticed that a point scatterer in the earth appears on a radial-trace section as an *exact* hyperbola, not a flat-topped hyperboloid. The travel-time curve for a point scatterer, Cheops' pyramid, can be written as a "string length" equation, or a stretched-circle equation (Section 3.2). Making the definition

$$\sin\psi = \frac{2h}{vt} \tag{5}$$

and substituting into Section 3.2, equation (12) yields

$$vt = 2\left[\frac{z^2}{\cos^2\psi} + (y - y_0)^2\right]^{1/2} \tag{6}$$

Scaling the z-axis by $\cos\psi$ gives the circle and hyperbola case all over again! The hidden hyperbola is shown as a three-dimensional sketch in figure 7.

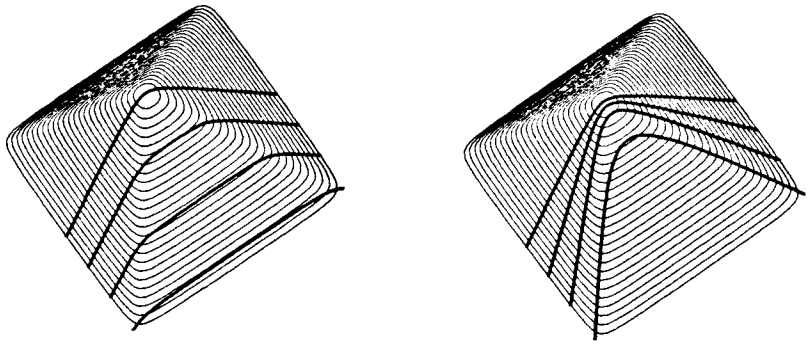


FIG. 7. An unexpected hyperbola in Cheops' pyramid is the diffraction hyperbola on a radial-trace section. (Ottolini)

This means that on a radial-trace section the diffraction response is a simple hyperbola. From a data-processing point of view the simple hyperbola is much easier to handle than the flat-topped hyperboloid that is seen at constant h. It is easy to convert a data hyperbola of one ψ-value to a data hyperbola of another: just rescale the z-axis.

According to Fourier transform theory, scaling z by a $\cos\psi$ divisor will scale k_z by a $\cos\psi$ multiplier. So the downward extrapolation is given by

$$k_z = -\frac{\omega}{v}\frac{1}{\cos\psi}\left[1-\left[\frac{v\,k_y}{2\,\omega}\right]^2\right]^{1/2} \qquad (7a)$$

The offset information is buried in the angle parameter ψ. There are no derivatives or wavenumbers in the offset or ψ direction. The data seem to be happy to remain at a fixed ψ-value.

Define parameters a and b by writing (7a) as

$$k_z = -\frac{\omega}{v}\,a\,b \qquad (7b)$$

This can be interpreted as

$$a\,b = 1 + (a-1) + (b-1) + (a-1)(b-1)$$

$$a\,b = TD + NMO + Mig + Dev \qquad (8)$$

This result is delightfully simple. The dip moveout operator *Dev* is a migration $(b-1)$ with a scaling factor $(a-1)$ that depends on the radial plane being migrated.

The generalization to stratified media may not be easy, but on the other hand it may not be necessary. There are other definitions for *NMO* and *Mig* in media that vary in depth (and even vary laterally). The practical view is that *Dev* is a modest correction and thus need not itself be represented with great accuracy — constant velocity should be adequate. The processing sequence envisioned is as follows:

1. Convert offset space to radial traces.
2. Apply *Dev* using any cheap algorithm with v = 2.5 km/sec.
3. Revert from radial traces to offset space.
4. Use your favorite conventional velocity analysis.
5. Stack.
6. Migrate.

The main disadvantage of this approach to velocity estimation seems to be the unfamiliarity of radial traces. Most people would prefer an operator to operate on constant-offset sections. Also, the treatment of $v(z)$ seems crude.

Anti-Alias Characteristic of Dip Moveout

You might think that if (y,h,t)-space is sampled along the y-axis at a sample interval Δy, then any final migrated section $P(y,z)$ would have a spatial resolution no better than Δy. This is not the case.

The basic principle at work here has been known since the time of Shannon. If a time function and its derivative are sampled at a time interval $2T$, they can both be fully reconstructed provided that the original bandwidth of the signal is lower than $1/2T$. More generally, if a signal is filtered with m independent filters, and these m signals are sampled at an interval mT, then the signal can be recovered.

Here is how this concept applies to seismic data. The basic signal is the earth model. The various filtered versions of it are the constant-offset sections. Further details can be found in a paper by Bolondi, Loinger, and Rocca (1982), who first pointed out the anti-alias properties of dip moveout.

Yilmaz's Deviation

Now that the basic geometrical concepts of prestack partial migration have been illustrated by the approaches of Sherwood, Rocca, and Ottolini, we can try to establish the prestack partial-migration operator using the wave equation. Such an approach, if successful, should lead to appropriate treatment of amplitude, phase, wavelet and all other aspects of wave equations. A natural starting point is the double-square-root equation:

$$\frac{\partial P}{\partial z} = -i\,\frac{\omega}{v}(DSR)\,P \tag{9}$$

$$DSR(Y,H) = \sqrt{1-(Y-H)^2} + \sqrt{1-(Y+H)^2} \tag{10}$$

It has already been noted (Section 3.4) that the DSR operator is ***not separable*** into a sum of an offset operator and a midpoint operator. ***Non-separable*** means that a Taylor series for (10) contains terms like Y^2H^2. Such terms cannot be expressed as a function of Y plus a function of H. Nonseparability is a data-processing disaster. It implies that migration and stacking must be done simultaneously, not sequentially. The only way to recover pure separability is to return to the space of S and G, but that would retreat from the practical problem, which can be solved using only a modest correction.

Recall from Section 3.4 the separable approximation to (10):

$$Sep(Y,H) = 2 + [DSR(Y,0) - 2] + [DSR(0,H) - 2] \tag{11a}$$

$$= 2\,[1 + (\sqrt{1-Y^2} - 1) + (\sqrt{1-H^2} - 1)] \tag{11b}$$

$$= TD + Mig(Y) + NMO(H) \tag{11c}$$

The separable approximation describes "standard processing."

To devise a prestack partial-migration operator it seems necessary to cope with the physical effects found in the Y^2H^2 term, but to do so in such a way that H is not interpreted as an operator. Recall that the stepout operator H is interpreted in Fourier space by k_h/ω and in physical space by ∂_h^t. Both of these forms imply mixing information from different offsets. The ray interpretation avoids this. Define $H_0 = \frac{1}{2}v\,dt/dh$. Then H is interpreted as a scalar function of the coordinates h and t rather than as an operator. So $Y^2H_0^2$ will be the partial-migration operator.

To find the value of H, examine the simplest geometry

$$(2h)^2 + (2z)^2 = v^2t^2 \tag{12}$$

Differentiating with respect to h at constant z gives

$$\frac{dt}{dh} = 4\,\frac{h}{t\,v^2} \tag{13a}$$

$$H_0 = \frac{v}{2}\,\frac{dt}{dh} = \frac{2h}{v\,t} \tag{13b}$$

This means that the stepout dt/dh on a CMP gather can be predicted by knowing only velocity and position (h,t).

Equation (13b) is readily generalized to stratified media. If the velocity is stratified, then H is also a function of z. There is no need, however, to be extremely careful when computing only a small correction. When Yilmaz applied these ideas in his dissertation he found it satisfactory to use an RMS velocity, although he also incorporated a regional dip correction

$$H_0 = \frac{2h}{v_{RMS}\,t}\cos 45° \tag{13c}$$

Doing this may seem crude, but it is much better than standard processing, which sets the term to zero.

We can do better than the $Y^2H_0^2$ term. Recall the reasoning that led to defining Sep as an approximation to DSR. That procedure led to an accurate answer for $H_0 = 0$. Now try this:

$$Dev(Y) = DSR(Y,H_0) - Sep(Y,H_0) \tag{14}$$

Equation (14) defines a deviation operator. We might even say equation (14) is a ***devilishly*** clever way of representing the difference between the DSR equation and its separable approximation in the form of a migration operator that depends on the scalars h and t.

To obtain a more familiar result, make all the substitutions and also some approximations. Use square-root expansions that allow H_0 to be large (≈ 1), but keep terms that are second order in Y:

$$Dev(Y) \approx \left[1 - \frac{1}{(1-H_0^2)^{3/2}}\right] Y^2 \approx -\frac{3}{2}H_0^2\,Y^2 \tag{15a}$$

Incorporating the missing $-i\omega/v$ from the *DSR* equation, introducing travel-time depth $\tau = 2z/v$, and going to the time domain gives

$$\frac{\partial^2 P}{\partial \tau \, \partial t} = -\frac{3\cos 45^\circ}{4} \frac{h^2}{t^2} \frac{\partial^2}{\partial y^2} P \qquad (15b)$$

Notice again the velocity-independence of the result. Yilmaz implemented (15a), the center term, and performed the partial migrations of impulses, getting what should be Rocca's operator without the *NMO*. The result was shown in figure 5. Then he tested the operator on some synthetic constant-offset sections, as shown in figure 8. The results shown in figures 5 and 8, while imperfect, are reasonable. Their greatest difficulty lies in the vicinity of the flat top. This is the pole in equation (15a). There both numerical and theoretical approximations could break down.

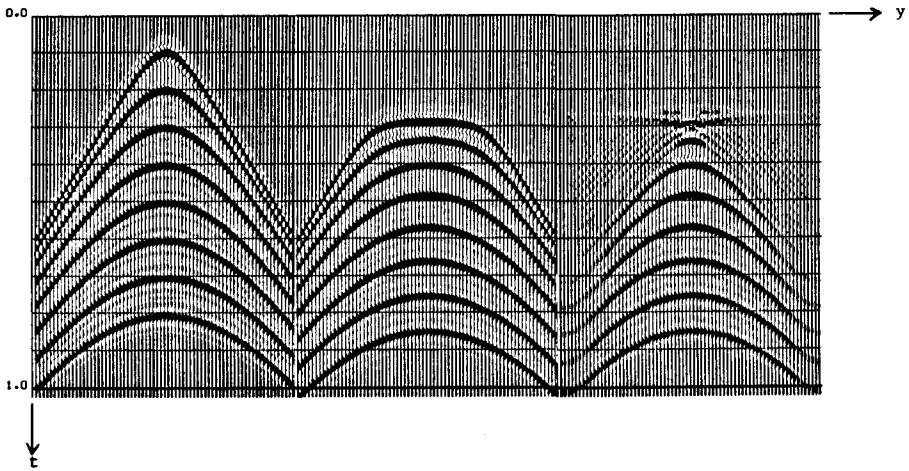


FIG. 8. Zero-offset section over point scatterers (left). Constant-offset section over same scatterers (center). Attempt of *Dev* operator to convert constant offset to zero offset. (Yilmaz)

Even if the numerical work were done perfectly, the result should not be perfect, for two reasons. The first reason is the cos 45° patch. The second and more fundamental reason is that it was assumed that dt/dh could be expressed as a function of the coordinate system (h,t) instead of being attached to the events themselves.

The Rocca smear operator has correct timing for all dips, but the Yilmaz deviation operator is valid over a limited range of dip. The strength of the Yilmaz operator is that it is rooted in the DSR. Thus it should have an appropriate amplitude and wavelet, and it should be a logical approach to depth-variable velocity. Too bad it breaks down at large angles.

Hale's Constant Offset DMO

Dave Hale (dissertation, in preparation) has recently found an exact Fourier representation of Rocca's smile and thus a method for constant-offset sections that is not so radical as Ottolini's radial-trace method. When the velocity is constant, Hale's method is exact for all angles.

3.7 Lateral Velocity Variation in Bigger Doses

To the interpreting *geologist*, lateral velocity variation produces a strange distortion in the seismic section. And the distortion is worse than it looks. The *geophysicist* is faced with the challenge of trying to deal with lateral velocity variation in a quantitative manner. First, how can reliable estimates of the amount of lateral velocity variation be arrived at? Then, do we dare use these estimates for reprocessing data?

Our studies of *dip* and *offset* have resulted in straightforward procedures to handle them, even when they are simultaneously present. Unfortunately, increasing lateral velocity variation leads to increasing confusion. It is confusion we must try to overcome. Strong lateral velocity variation overlies the largest oil field in North America, Prudhoe Bay. Luckily, we have many idealized examples that are easy to understand. Any "ultimate" theory would have to explain these examples as limiting cases.

Let us review. The double-square-root equation presumably works if the square roots are expanded and if we accept the usual limitation of accuracy with angle. Our problem with the DSR is that it merely tells us how to migrate and stack *once the velocity is known*. Kjartansson's method of determining the distribution $v(x,z)$ assumes straight rays, no dip, and a single, planar reflector. On the other hand, stacking along with prestack partial migration allows any scattering geometry but enables determination of $v(z)$ only under the presumption that there is no lateral variation of velocity. Clearly, there are many gaps. We begin with comprehensible, special cases but ultimately sink into a sea of confusion.

Replacement Velocity: Freezing the Water

Sometimes you are lucky and you know the velocity. Maybe you know it because you are dealing with synthetic data. Maybe you know it because you have already drilled 300 shallow holes. Or maybe you can make a good estimate because you have a profile of water depth and you are willing to guess at the sediment velocity. Often the velocity problem is really a near-surface problem. Perhaps you have been dragging your seismic streamer over the occasional limestone reefs in the Red Sea.

Assuming that you know the velocity and that the lateral variations are near the surface, then you should think about the idea of a *replacement velocity*. For example, suppose you could freeze the water in the Red Sea, just until it is hard enough that the ice velocity and the velocity of the limestone reefs are equal. That would remove the unnecessary complexity of the reflections from deep targets. Of course you can't freeze the Red Sea, but you can reprocess the data to try to mimic what would be recorded if you could.

First, downward continue the data to some datum beneath the lateral variations. Then upward continue it back to the surface through the homogeneous replacement medium.

While in principle the DSR could be used for this job, in practice using it would be expensive and impractical. The best approach is to study the two operations —going down, then going up —in combination. Since the two operations are largely in opposition to each other, whatever is done to the data should be just a function of the difference. For example, the equation

$$\frac{\partial P}{\partial z} = i\omega\left[\frac{1}{v(s)} + \frac{1}{v(g)} - \frac{2}{v_{avg}}\right]P \qquad (1)$$

combines the downward continuation with the upward continuation and makes little change to the wavefield P when the velocities are nearly the same.

I haven't tried wider angle representations, but it seems like a good idea to try them, particularly for the offset angle. This is clearly a task that would be well suited to the talent found in geophysical contracting companies. Figure 1 illustrates a process of this type, called *REVEAL* by Digicon, Inc., who have chosen not to *reveal* the details of their implementation.

Lateral Shift of the Hyperbola Top

Shown in figure 2 is a point scatterer below a 30° wedge with low velocity on the high side and high velocity beneath. This is a simple prototype for many lateral-velocity-variation problems. Surface arrival times will be roughly hyperbolic with some distortion because of the velocity jump at the wedge. The minimum travel time (hyperboloid top) has been displaced from its usual location directly above the point scatterer. Notice and remember that

1. At minimum time, the ray emerges going straight up.
2. Minimum time is on the high velocity side of the point scatterer.

The travel-time curve is roughly hyperbolic, but the asymptote on the left side gives the velocity of the medium on the left side and

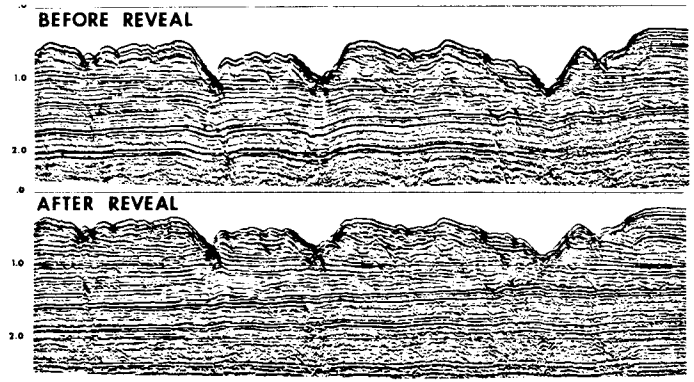


FIG. 1. Example of processing with a replacement velocity. Observe that deeper bedding is now flatter and more continuous. (Digicon, Inc., *Geophysics*, 1977, v. 42, n. 3, p. A76)

the asymptote on the right gives the velocity on the right. We can also deduce that

3. Displacement of the hyperboloid from the scatterer increases with offset.

Let $T(x)$ denote the travel time from the point scatterer to the surface point x. The travel time for a constant-offset section is then $t(y) = T(y+h) + T(y-h)$. To find the earliest arrival, set $dt/dy = 0$. This proves that the slope at a on figure 2 is the negative of the slope at b. This, with the asymmetry of the travel-time curve, establishes point 3.

Lateral velocity variation causes hyperbolas to lose their symmetry. Computationally, it is the lens term that tilts hyperbolas, causing their tops to move laterally.

Phantom Diffractor

A second example of lateral velocity variation is figure 3, also taken from Kjartansson's dissertation. The physical model shown on the inset in figure 3 is three constant velocity wedges separated by broken line segments representing reflectors. The bottom edge of the model also represents a reflector. The wave field in figure 3 was made using the exploding-reflector calculation, which Kjartansson regarded as a reasonable approximation to a zero-offset

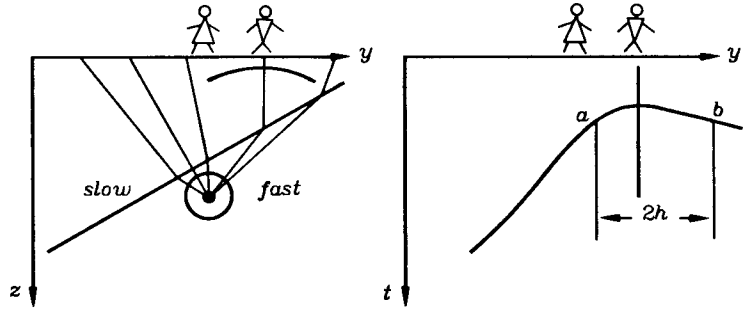


FIG. 2. Rays emerging from a point scatterer beneath a velocity wedge (left). Travel time curve (right). Slope at a is the negative of that at b.

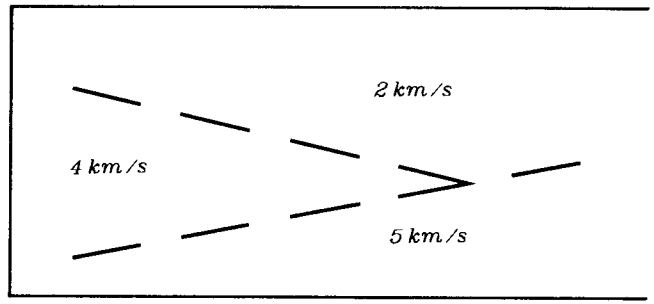
section. Notice that under the tip of the 4 km/sec wedge is a small diffraction on the bottom horizontal reflector. Because such a diffraction has nothing to do with the flat reflector on which it is seen, it is termed a "phantom" diffraction. Phantom diffractions are not easy to recognize, but they do occur. In reality, the "bright spots" in Section 3.1 were probably phantom diffractions. It has been reported that phantom diffractions provide a means of prospecting for small, high-velocity, carbonate reefs.

Wavefront Healing

Another example of ray bending, reproduced from FGDP, is shown in figure 4. The first frame on the left shows a plane wave just after it has been distorted into a wavy shape by the thin-lens term. After this the thin-lens term vanishes. Later frames show the effect of increasing amounts of diffraction.

Fault-Plane Reflection

Across a single vertical fault in the earth the velocity will be a simple step function of the horizontal coordinate. Rays traveling across such a fault suffer in amplitude because of reflection and transmission coefficients, depending on the angle. Since near-vertical rays are common, only small velocity contrasts are required to generate strong internal reflections. By this reasoning, steep faults should be more distorted, and hence more



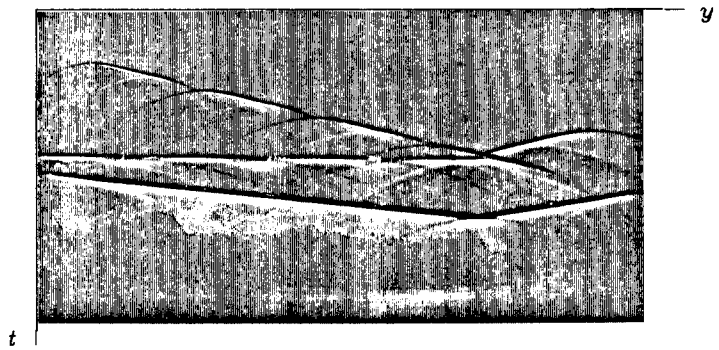


FIG. 3. The model in the upper panel was taken from Western Geophysical's *Depth Migration* brochure. The model is not physical because of the segmenting of the interface; however, the segments make it a good case for the study of lateral shifts. The synthetic data is in the lower panel (from Kjartansson). The phantom diffraction is on the latest arrival just below the tip of the wedge.

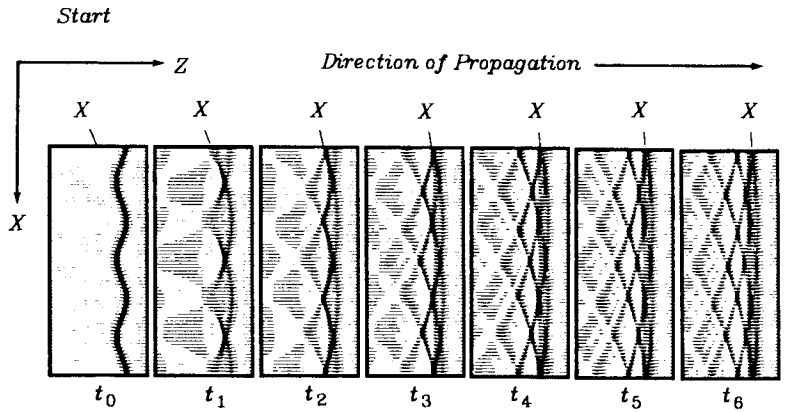


FIG. 4. The first frame on the left shows a plane wave just after it has been distorted into a wavy shape by the thin lens term. After this the thin lens term vanishes. Subsequent frames show the effect of increasing amounts of diffraction. Notice the lengthening of the wave packet and the ***healing*** of the first arrival. (FGDP, p. 213, figure 10-22)

recognizable, on small-offset sections than on wide-offset sections or stacks.

This phenomenon is somewhat more confusing when seen in (x,t)-space. Figure 5 was computed by Kjartansson and used in a quiz. Study this figure and answer the questions in the caption. Here is a hint: A reflected ray beyond critical angle undergoes a phase shift. This will turn a pulse into a doublet that might easily be mistaken for two rays.

Figure 5 exhibits a geometry in which the exploding-reflector model fails to produce all the rays seen on a zero-offset section. The exploding-reflector model produces two types of rays: the ray that goes directly to the surface, and the ray that reflects from the fault plane before going to the surface. A zero-offset section has three rays: the two rays just mentioned, but moving at double travel time, once up, once down; and in addition the ray not present in figure 5, which hits the fault plane going one way but not the other way.

There is a simple way to make constant-offset sections in laterally variable media when the reflector is just a point. The

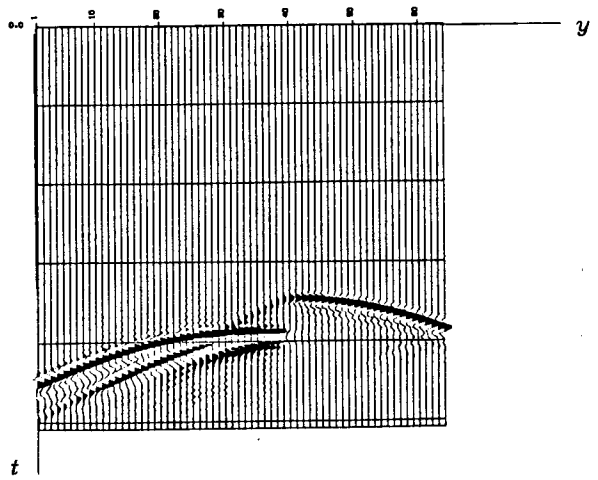


FIG. 5. Synthetic data from an exploding-reflector calculation for an earth model containing a point scatterer and a velocity jump $v_1/v_2 = 1.2$ across a vertical contact. (Kjartansson)

(a) Is the point scatterer in the slow or the fast medium?
(b) Identify four arrivals and diagram their raypaths.
(c) Identify and explain two kinds of computational artifacts.
(d) Find an evanescent wave.
(e) Find phase shift on a beyond-critical-angle reflection.
(f) A zero-offset section should have a ray not shown above. Where?

exploding-reflector seismogram recorded at $x=s$ is simply time convolved with the one recorded at $x=g$. Convolution causes the travel times to add. Even the non-exploding-reflector rays are generated. Too bad this technique doesn't work for reflector models that are more complicated than a simple reflecting point.

Misuse of $v(x)$ for Depth Migration

The program that generated figure 5 could be run in reverse to do a migration. All the energy from all the interesting rays would march back to the impulsive source. Would this be an effective migration program in a field environment? It is unlikely that it would. Even assuming that all the theoretical weaknesses could be eliminated, the process is far too sensitive to quantitative knowledge of the lateral velocity jump. It is the quantitative value that determines the reflection coefficient and ultimately the correct recombination of all the wavefronts back to a pulse. To see how an incorrect value can result in further error, imagine using the hyperbola-summation migration method. Applied to this geometry this method implies weighted summation over all the raypaths in the figure. The incorrect value would put erroneous amplitudes on various branches. An erroneous location for the fault would likewise mislocate several branches.

The lesson to be learned from this example is clear. Unnecessary bumps in the velocity function can create imaginary fault plane reflections. Consistent with known information, a presumed migration velocity should be as smooth as possible in the lateral direction. Unskilled and uninformed staff at a processing center remote from the decision making should not have the freedom to introduce rapid lateral changes in the velocity model.

First-Order Effects, the Lens Term

Now let us be specific about what is meant by the lens term in the present context of before-stack migration in the presence of lateral velocity variation. Specializing the DSR equation to 15° angles gives

$$\frac{\partial U}{\partial z} = -\left\{\left[\frac{i\omega}{v(s)^2} + \frac{v(s)^2}{2i\omega}\frac{\partial^2}{\partial s^2}\right] + \left[\frac{i\omega}{v(g)^2} + \frac{v(g)^2}{2i\omega}\frac{\partial^2}{\partial g^2}\right]\right\} U \qquad (2)$$

Rearranging the terms to group by behavior gives

$$\frac{\partial U}{\partial z} = -\left[\frac{i\omega}{v(s)^2} + \frac{i\omega}{v(g)^2}\right] U - \left[\frac{v(s)^2}{2i\omega}\frac{\partial^2}{\partial s^2} + \frac{v(g)^2}{2i\omega}\frac{\partial^2}{\partial g^2}\right] U \qquad (3a)$$

$$\frac{\partial U}{\partial z} = \textit{lens term} + \textit{diffraction term} \qquad (3b)$$

So you see the familiar type of lens term, but it has two parts, one for shifting at the shot, and one for shifting at the geophone.

The Migrated Time Section: An Industry Kludge

One of the things you notice first when you encounter lateral velocity variation is that CMP stacking behaves strangely. The moveout may become nonhyperbolic or the stacking velocity may become nonphysical. Quantitatively the effect is more severe than the effect of cosine correction for dip. We will take the expedient approach, which is to assert that a stack is merely an estimate of the zero-offset section. Thus migration should proceed as it would for a zero-offset section. Later we will return to the stacking problem.

My discussion on ***travel-time depth*** (in Section 1.3) explained the industrial practice of avoiding reference to depth on a migrated section. This was done by the use of a travel-time depth τ defined by

$$\frac{d\tau}{dz} = \frac{2}{v_{rock}} = \frac{1}{v_{half}} \tag{4}$$

The purpose of the transformation is to reduce the velocity sensitivity of the data display. The sensitivity is reduced because a flat horizontal reflector may be migrated with ***any*** velocity without its position on a migrated time section being affected.

The effect of velocity error increases with dip, as can be seen by noting that the downward-continuation operation

$$\exp\left\{ i\,\frac{\omega}{v}\left[1-\left[\frac{v\,k_y}{\omega}\right]^2\right]^{1/2}\Delta z\right\} \tag{5}$$

is converted by (4) to

$$\exp\left\{ i\omega\,\Delta\tau\left[1-\left[\frac{v\,k_y}{\omega}\right]^2\right]^{1/2}\right\} \tag{6}$$

In (6), the velocity v multiplies the stepout k_y/ω, so an error in one is like an error in the other. The two possible errors are conveniently lumped together and called an angle error. In practice, it is often valid to say that 15° migration requires little velocity accuracy while 45° migration demands great velocity accuracy.

This conventional wisdom becomes doubtful in the presence of lateral velocity variation. There seems to be no satisfactory, automatic way of dealing with the data over the quantitative range of typical geological parameters. The downward-continuation equations clarify the difficulty.

Even in the presence of lateral velocity variation, the square root in (5) is interpretable in a "local plane wave" sense, and it may be used to define a reasonable equation for downward continuation. The exploding-reflector model means that local spatial wavelengths are controlled by

$$k_z(y,z) = \frac{\omega}{v(y,z)}\left[1-\left[\frac{v(y,z)\,k_y(y,z)}{\omega}\right]^2\right]^{1/2} \tag{7}$$

The difficulty arises in converting (7) to travel-time depth using (4). The issue is whether to let the velocity in (4) be laterally variable. This will be tried, not because it is justifiable, but because it leads to a definition of the ***migrated time section***, an important industrial product. The definition is given by any reasonable implementation of

$$k_\tau(y,\tau) = \omega\left[1-\left[\frac{v(y,\tau)\,k_y(y,\tau)}{\omega}\right]^2\right]^{1/2} \tag{8}$$

This definition contains a serious pitfall, which only shows up when the velocity is laterally variable. The (y,z) coordinate system is an orthogonal coordinate system, but the (y,τ) system is not [unless $v(y)=const$]. So equation (7), which says that $\cos\theta = \sqrt{1-\sin^2\theta}$, is not correctly interpreted by (8).

In summary, the situation is this: In a production environment a great deal of data gets processed before anyone has a clear idea of how much lateral velocity variation is present. So the lens term is omitted. The results are OK if the lens term happens to commute with the diffraction term. The terms do commute when the lateral velocity variation is slow enough. Otherwise, you have a nasty interpretation problem.

Effect on Stacking Velocity (Lynn)

The dissertation of Walter S. Lynn and the associated article in *Geophysics* show how PDEs can be written to describe the influence of lateral velocity variation on stacking velocity and migration.

You are on the island of Princes and Pawns. Princes always tell the truth, and pawns always tell lies. You cannot tell them apart from appearance. You come across two people, A and B. A says: "Either I am a pawn or B is a prince."

What are A and B?

##########

You are on the island of Princes and Pawns. Princes always tell the truth, and pawns always tell lies. You cannot tell them apart from appearance. You meet three natives: A, B, and C. A and B make the following statements:

A: All of us are pawns.

B: Exactly on of us is a prince.

What are A, B, and C?

##########

You are on the island of Princes and Pawns. Princes always tell the truth, and pawns always tell lies. You cannot tell them apart from appearance. You come across two people, A and B. A says:

"I am a pawn, but B isn't."

What are A and B?

##########

You are on the island of Princes and Pawns. Princes always tell the truth, and pawns always tell lies. You cannot tell them apart from appearance. You meet three natives: A, B, and C. A and B make the following statements:

A: "B is a pawn."

B: A and C are of the same type.

What is C?

##########

You are on the island of Princes and Pawns. Princes always tell the truth, and pawns always tell lies. You cannot tell them apart from appearance. You meet three natives: A, B, and C. A makes the following statement:

"B and C are of the same type."

If you asked C, "Are A and B of the same type?", what would C say?

##########

I visited the island of Princes and Pawns. Princes always tell the truth, and pawns always tell lies. I came across two of the natives resting under a tree. I asked one of them, "Is either of you a prince?" He responded, and I knew the what each one was. What are the two natives?

##########