6.2 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.

6.2.1 Sections and gathers for planar reflectors

The simplest environment for reflection data is a single horizontal reflection interface, which is shown in Figure 6.12. As expected, the zero-offset section mimics the earth

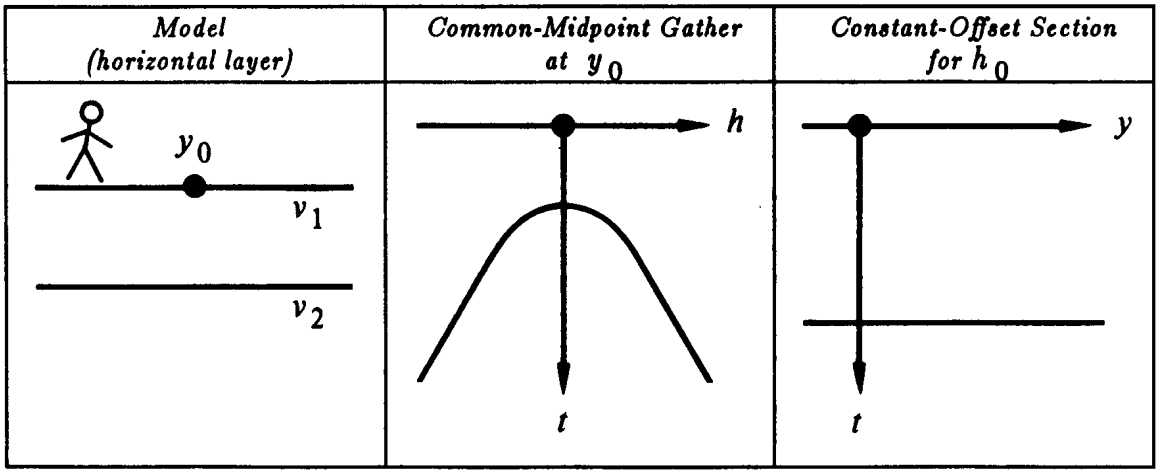

Figure 6.12: Simplest earth model. ofs-simple [NR]

model. The common-midpoint gather is a hyperbola whose asymptotes are straight lines with slopes of the inverse 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.

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 6.13.

Figure 6.13 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.

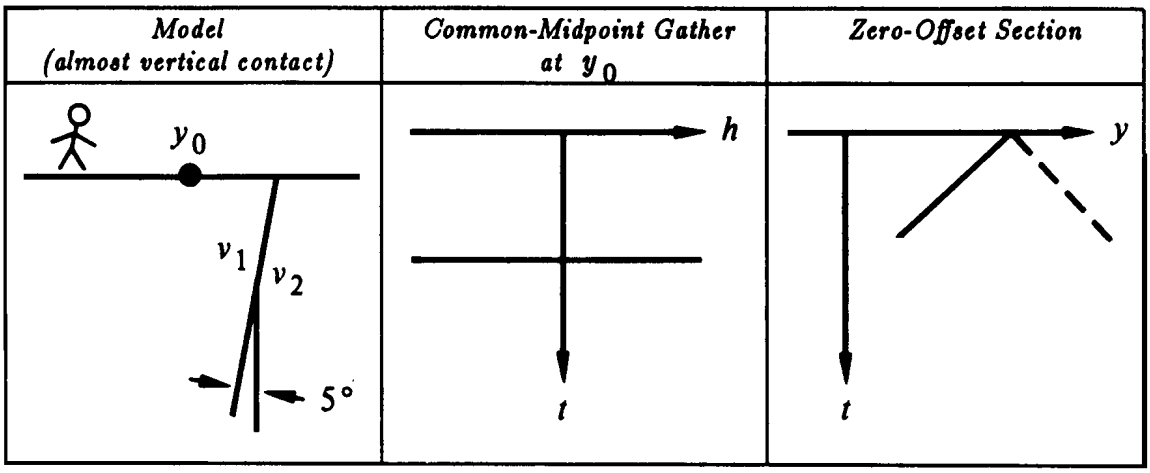

Figure 6.13: Near-vertical reflector, a gather, and a section. ofs-vertlay [NR]

A planar reflector may have any dip between horizontal and vertical. Then the common-midpoint gather lies between the common-midpoint gather of Figure 6.12 and that of Figure 6.13. The zero-offset section in Figure 6.13 is a straight line, which turns out to be the asymptote of a family of hyperbolas. The slope of the asymptote is the inverse of the velocity v_1.

6.2.2 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

$$t^2\, v^2 \quad = \quad 4\,(y - y_0)^2\, \sin^2\alpha \;+\; 4\,h^2\, \cos^2\alpha \tag{6.3}$$

For $\alpha = 45°$, equation (6.3) 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 (6.3) looks like $t^2 = t_0^2 + 4h^2/v^2_{\text{apparent}}$. 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]:

$$v_{\text{apparent}} \quad = \quad \frac{v_{\text{earth}}}{\cos(\alpha)} \tag{6.4}$$

(See also Slotnick [1959]). In summary, dip increases the stacking velocity.

Figure 6.14 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

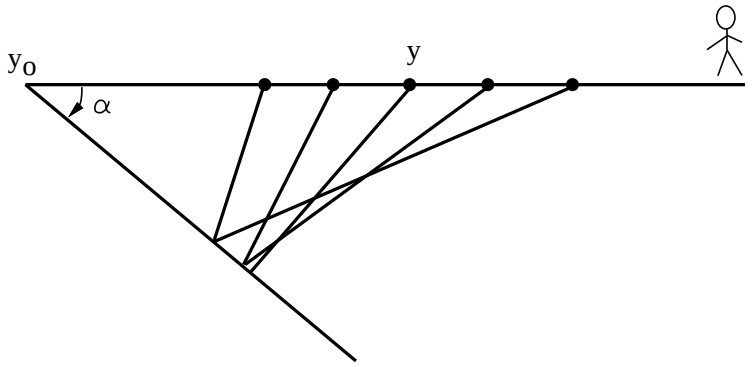

Figure 6.14: Rays from a common-midpoint gather. ofs-dipray [NR]

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 (6.3) will be proved. Figure 6.15 shows the basic geometry along with an "image" source on another reflector of twice the dip. For convenience,

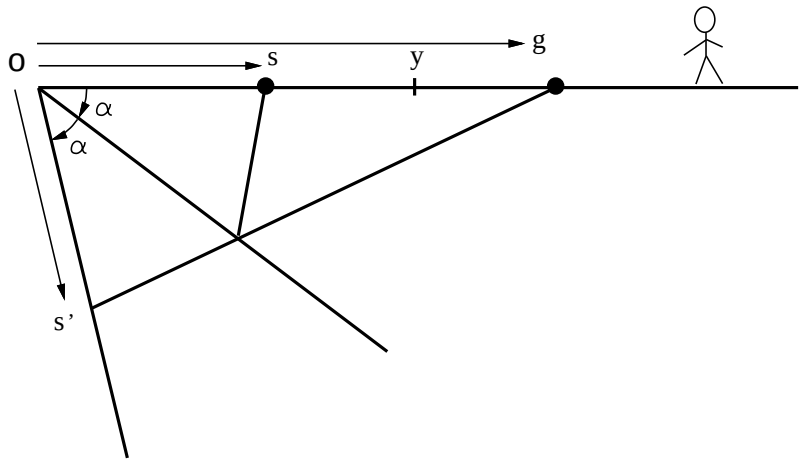

Figure 6.15: Travel time from image source at s' to g may be expressed by the law of cosines. ofs-lawcos [NR]

the bed intercepts the surface at $y_0 = 0$. The length of the line $s'g$ in Figure 6.15 is determined by the trigonometric Law of Cosines to be

$$\begin{aligned} t^2 v^2 &= s^2 + g^2 - 2\,s\,g\,\cos 2\alpha \\ t^2 v^2 &= (y-h)^2 + (y+h)^2 - 2\,(y-h)(y+h)\,\cos 2\alpha \\ t^2 v^2 &= 2\,(y^2+h^2) - 2\,(y^2-h^2)\,(\cos^2\alpha - \sin^2\alpha) \\ t^2 v^2 &= 4\,y^2\,\sin^2\alpha + 4\,h^2\,\cos^2\alpha \end{aligned}$$

which is equation (6.3).

Another facet of equation (6.3) is that it describes the constant-offset section. Surprisingly, the travel time of a dipping planar bed becomes curved at nonzero offset—it too becomes hyperbolic.

6.2.3 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 6.16 shows an example. The curves in Figure 6.16 include flat spots for the same reasons

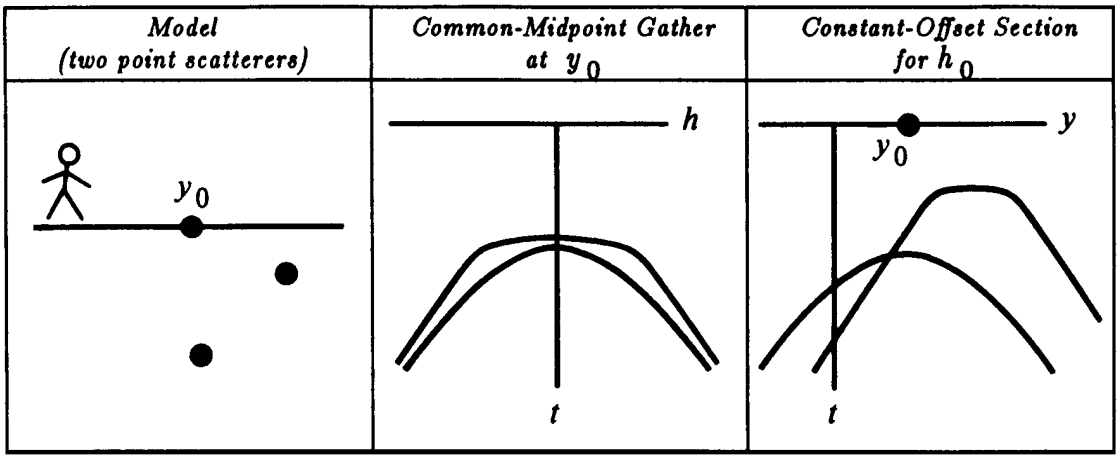

Figure 6.16: Response of two point scatterers. Note the flat spots. ofs-twopoint [NR]

that some of the curves in figures6.12 and 6.13 were straight lines.

The point-scatterer geometry for a point located at (x, z) is shown in Figure 6.17.

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

$$t\,v \quad = \quad \sqrt{z^2 + (s - x)^2} + \sqrt{z^2 + (g - x)^2} \tag{6.5}$$

6.2.4 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 (6.5). 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 (6.5) 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

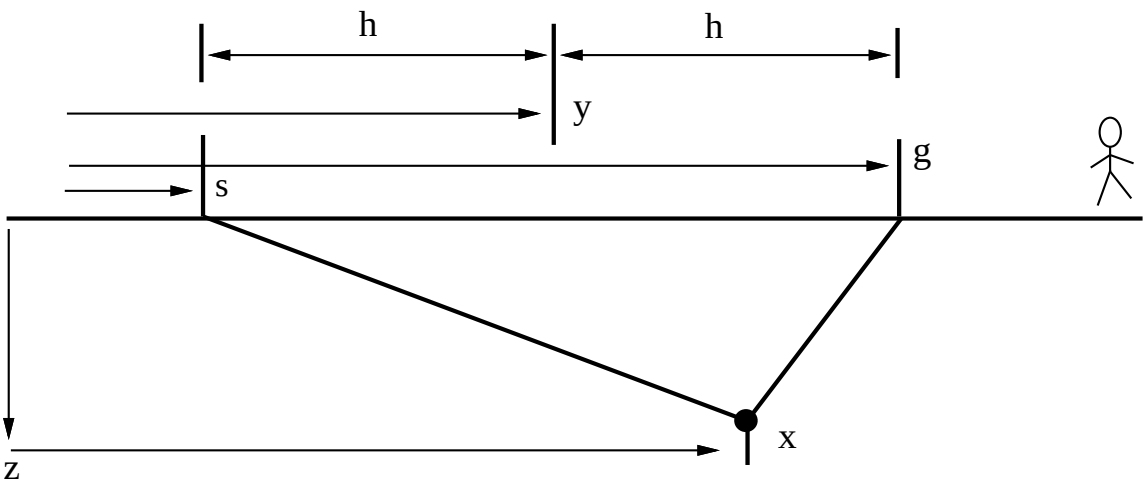

Figure 6.17: Geometry of a point scatterer. ofs-pgeometry [NR]

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.

A view of the travel-time pyramid on the (s,g)-plane or the (y,h)-plane is shown in Figure 6.18a. Notice that a cut through the pyramid at large t is a square, the corners of which have been smoothed. At very large t, a constant value of t is the square contoured in (s,g)-space, as in Figure 6.18b. Algebraically, the squareness becomes evident for a point reflector near the surface, say, $z \rightarrow 0$. Then (6.5) becomes

$$v\,t \quad = \quad |s \;-\; x| \;\; + \;\; |g \;-\; x| \tag{6.6}$$

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.

$$y \quad = \quad \frac{g \;+\; s}{2} \tag{6.7}$$

$$h \quad = \quad \frac{g \;-\; s}{2} \tag{6.8}$$

A traverse of y at constant h is shown in Figure 6.18. At the highest elevation on the traverse, you are walking along a flat horizontal step like the flat-topped hyperboloids

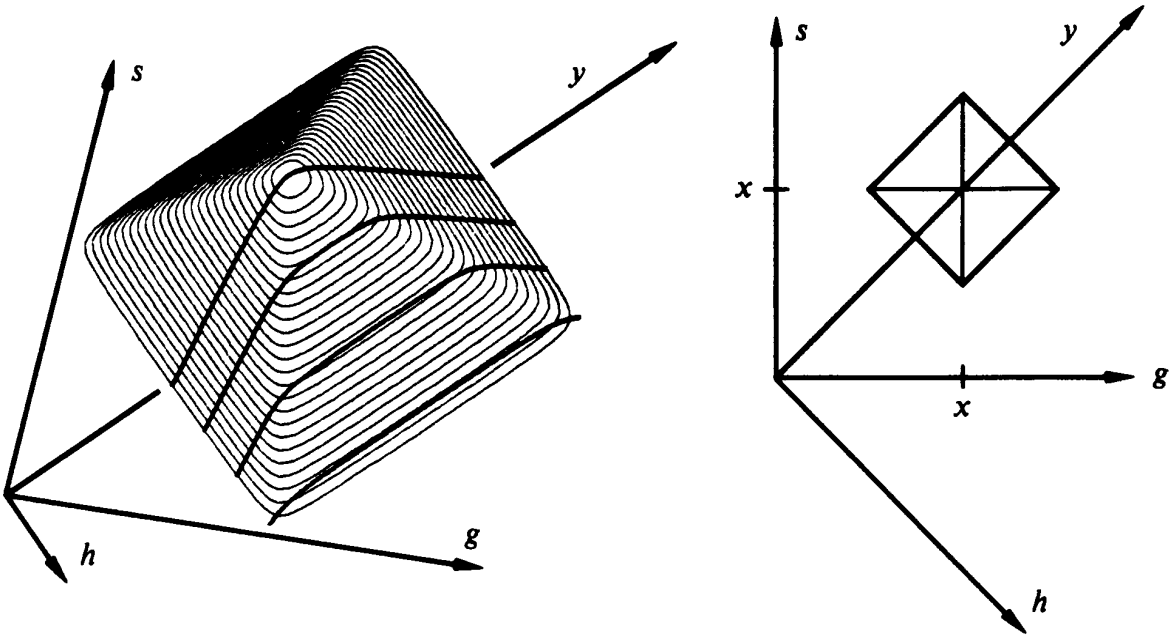

Figure 6.18: Left is a picture of the travel-time pyramid of equation ?? for fixed x and z. The darkened lines are constant-offset sections. Right is a cross section through the pyramid for large t (or small z). (Ottolini) ofs-cheop [NR]

of Figure 6.16. 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 (6.5) may be expanded in Taylor series, giving a parabola of revolution. This describes the eroded peak of the pyramid.

6.2.5 Random point scatterers

Figure 6.19 shows a synthetic constant-offset section (COS) taken from an earth model containing about fifty randomly placed point scatterers. Late arrival times

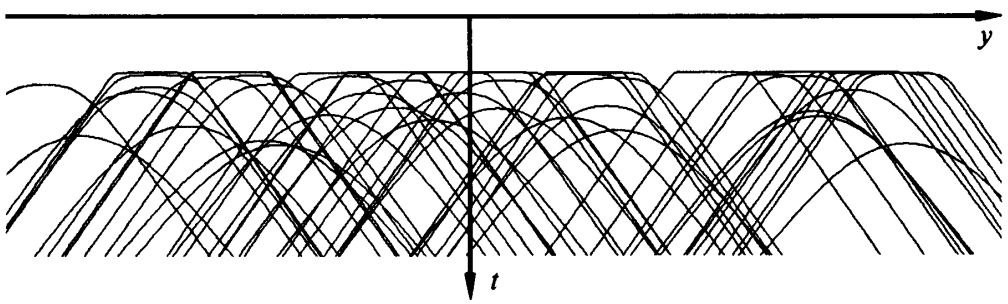

Figure 6.19: Constant-offset section over random point scatterers. ofs-randcos [NR]

appear hyperbolic. Earlier arrivals have flattened tops. The earliest possible arrival corresponds to a ray going horizontally from the shot to the geophone.

Figure 6.20 shows a synthetic common-shot profile (CSP) from the same earth model of random point scatterers. Each scatterer produces a hyperbolic arrival.

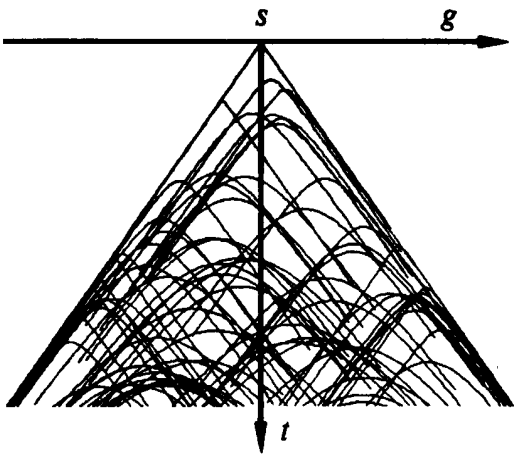

Figure 6.20: Common-shot profile over random point scatterers. ofs-randcsp [NR]

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 scatterers near the geophone, must lie near the lines $|g - s| = vt$.

The left side of Figure 6.21 shows a synthetic common-midpoint gather (CMP) from an earth model containing about fifty randomly placed point scatterers.

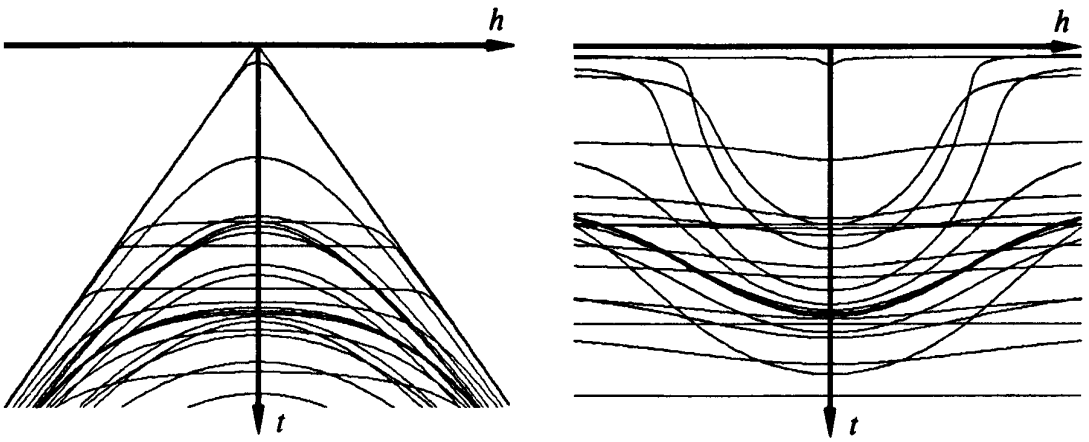

Figure 6.21: Common-midpoint gather on earth of randomly located point scatterers (left). The same gathers after NMO correction (right). ofs-randcmp [NR]

Because this is a common-*midpoint* gather, the curves are symmetric through zero offset. (The negative offsets of field data are hardly ever 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. The right side of Figure 6.21 shows what happens when normal-moveout correction is applied on the random scatterer model. Some reflectors are flattened; others are "overcorrected."

6.2.6 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 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 6.13. 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 6.22. Another less extreme model for the surface noises is the flat-topped curves in the random point-scatterer model.

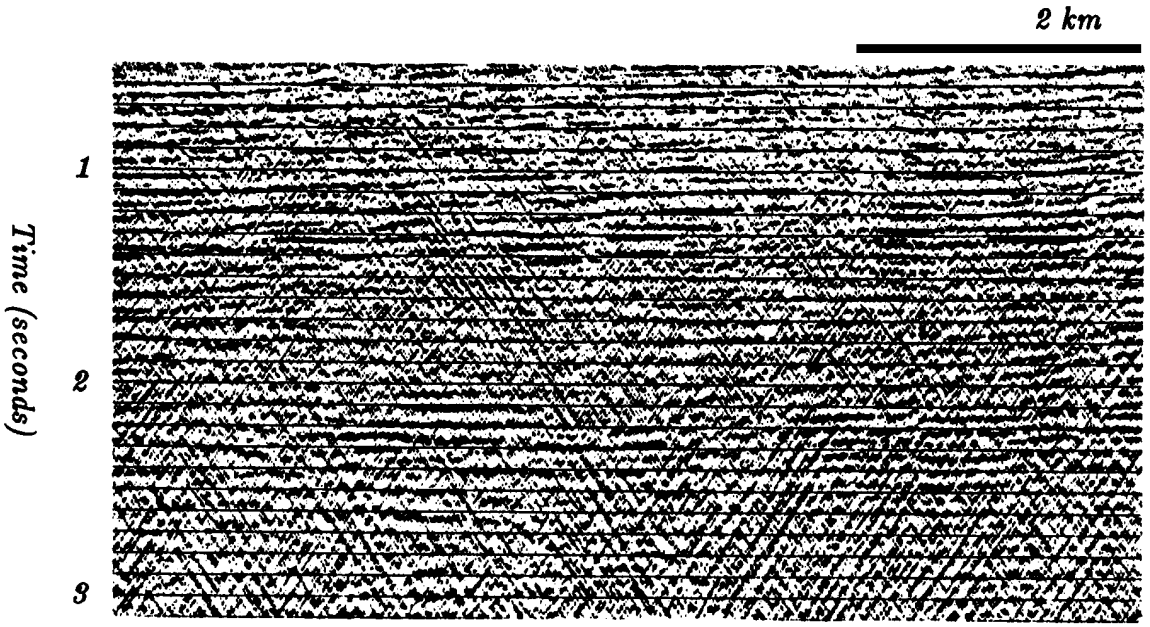

Figure 6.22: CDP stack with water noise from the Shelikof Strait, Alaska. (by permission from *Geophysics,* Larner et al. [1983]) ofs-shelikof [NR]

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 or steeply dipping layers. But equation (6.4) shows that dip increases the apparent velocity. So it is not surprising that stacking at deep-sediment, high velocities can enhance surface noises. Occurrence of this problem in practice was nicely explained and illustrated by Larner et al. [1983]

6.2.7 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. (Reality is more impressive. Wind blown ice flows drag themselves along the bottom making huge scars in it.) The travel-time curves for reflections from the boulders nicely matches 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 6.20. 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 in Figure 6.22 are not sidescatter.

The "streaking" scatterers are those neither along the survey line nor those perpendicular to it. The "in-line" scatterers have infinite velocity (on the flat spot). The scatterers perpendicular to the survey line have the water velocity. Scatterers at other angles have inbetween velocities and they stack strongly at the sediment velocity and account for the streaks.

6.2.8 The migration ellipse

Another insight into equation (6.5) 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 6.17 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 migrating 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 (Figure 3.6).

It is not easy to show that equation (6.5) 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 (6.5) in (y,h)-space is

$$t\,v \quad = \quad \sqrt{z^2 \,+\, (y \,-\, y_0 \,-\, h)^2} \,+\, \sqrt{z^2 \,+\, (y \,-\, y_0 \,+\, h)^2} \tag{6.9}$$

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

$$t\,v_{\text{rock}} \quad = \quad 2\,d \quad = \quad 2\,t\,v_{\text{half}} \tag{6.10}$$

$$a \quad = \quad z^2 \,+\, (y \,+\, h)^2 \tag{6.11}$$

$$b \quad = \quad z^2 \,+\, (y \,-\, h)^2 \tag{6.12}$$

$$a \,-\, b \quad = \quad 4\,y\,h \tag{6.13}$$

With these definitions, (6.9) becomes

$$2\,d \quad = \quad \sqrt{a} \,+\, \sqrt{b} \tag{6.14}$$

Square to get a new equation with only one square root.

$$4\,d^2 \,-\, (a \,+\, b) \quad = \quad 2\,\sqrt{a}\sqrt{b} \tag{6.15}$$

Square again to eliminate the square root.

$$16\,d^4 \,-\, 8\,d^2(a \,+\, b) \,+\, (a \,+\, b)^2 \quad = \quad 4\,a\,b \tag{6.16}$$

$$16\,d^4 \,-\, 8\,d^2(a \,+\, b) \,+\, (a \,-\, b)^2 \quad = \quad 0 \tag{6.17}$$

Introduce definitions of a and b.

$$16\,d^4 \,-\, 8\,d^2\,[2\,z^2 \,+\, 2\,y^2 \,+\, 2\,h^2] \,+\, 16\,y^2\,h^2 \quad = \quad 0 \tag{6.18}$$

Bring y and z to the right.

$$d^4 \,-\, d^2\,h^2 \quad = \quad d^2\,(z^2 \,+\, y^2) \,-\, y^2\,h^2 \tag{6.19}$$

$$d^2\,(d^2 \,-\, h^2) \quad = \quad d^2\,z^2 \,+\, (d^2 \,-\, h^2)\,y^2 \tag{6.20}$$

$$d^2 \quad = \quad \frac{z^2}{1 \,-\, \frac{h^2}{d^2}} \,+\, y^2 \tag{6.21}$$

Finally, recalling all earlier definitions,

$$t^2\,v_{\text{half}}^2 \quad = \quad \frac{z^2}{1 \,-\, \frac{h^2}{t^2\,v_{\text{half}}^2}} \,+\, (y \,-\, y_0)^2 \tag{6.22}$$

Fixing t, equation (6.22) 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 an impulse on a constant-offset section.

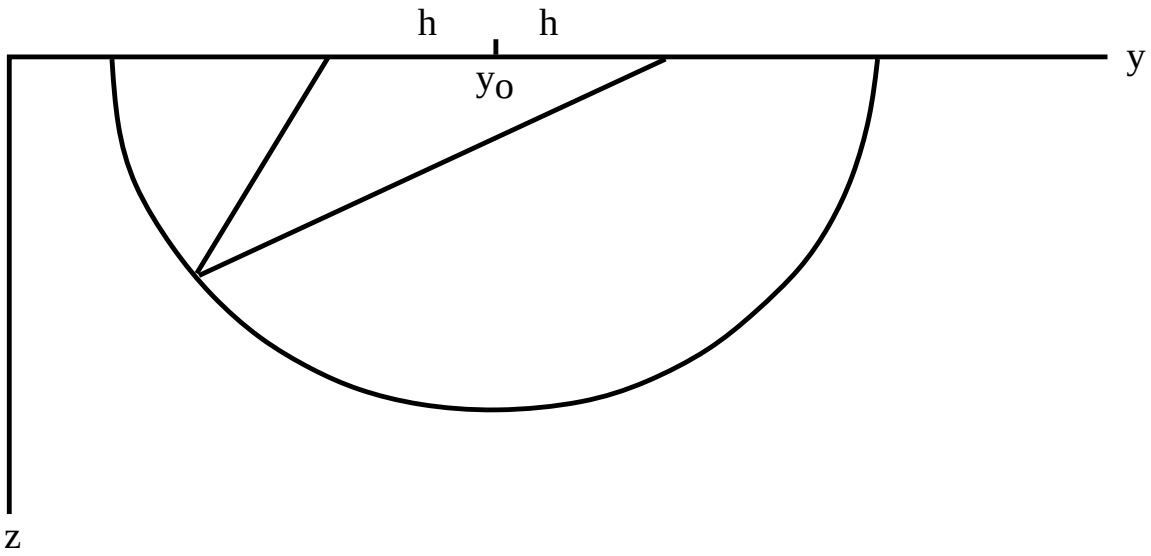

Figure 6.23: Migration ellipse. ofs-ellipse [NR]

6.3 SURVEY SINKING WITH THE DOUBLE-SQUARE-ROOT EQUATION

Exploding-reflector imaging will be replaced by a broader imaging concept, *survey sinking.* A new equation called the double-square-root (DSR) equation will be developed to implement survey-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 (6.35) 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 (6.48) instead of equation (6.35).

The DSR equation depends upon the reciprocity principle which we will review first.

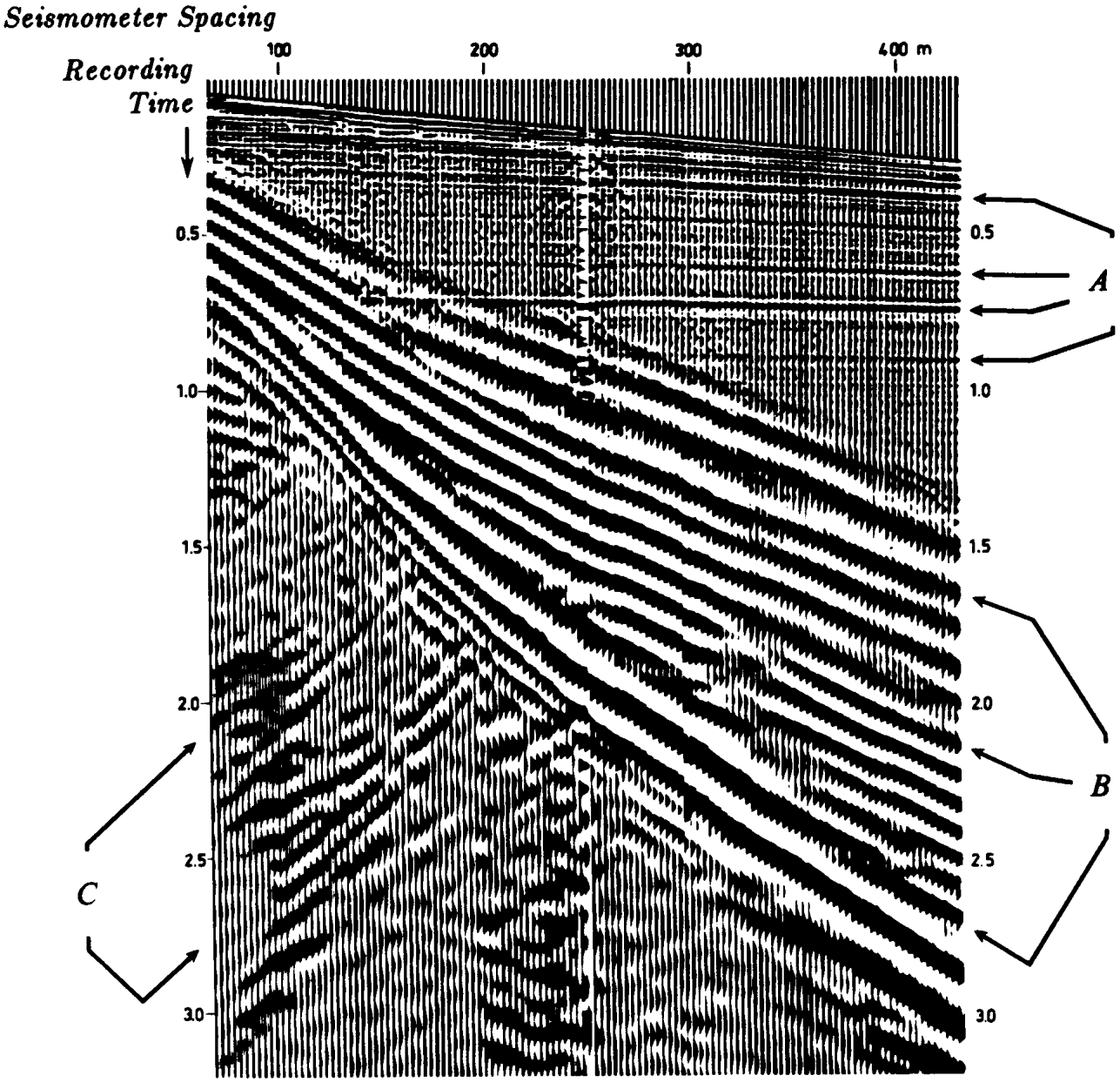

Figure 6.24: Undocumented data from a recruitment brochure. This data may be assumed to be of textbook quality. The speed of sound in water is about 1500 m/sec. Identify the events at A, B, and C. Is this a common-shotpoint gather or a common-midpoint gather? (Shell Oil Company) ofs-shell [NR]

6.3.1 Seismic reciprocity in principle and in practice

The principle of reciprocity says that the same seismogram should be recorded if the locations of the source and geophone are exchanged. A physical 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.

Mathematically, the reciprocity principle arises because the physical equations of elasticity are self adjoint. The meaning of the term *self adjoint* is illustrated in FGDP where it is shown that discretized acoustic equations yield a *symmetric* matrix even where density and compressibility are space variable. The inverse of any such symmetric matrix is another symmetric matrix called the impulse-response matrix. Elements across the matrix diagonal are equal to one another. Each element of any pair is a response to an impulsive source. The opposite element of the pair refers to the interchanged source and receiver.

A tricky thing about the reciprocity principle is the way antenna patterns must be handled. For example, a single vertical geophone has a natural antenna pattern. It cannot see horizontally propagating pressure waves nor vertically propagating shear waves. For reciprocity to be applicable, antenna patterns must not be interchanged when source and receiver are interchanged. The antenna pattern must be regarded as attached to the medium.

I searched our data library for split-spread land data that would illustrate reciprocity under field conditions. The constant-offset section in Figure 6.25 was recorded by vertical vibrators into vertical geophones. The survey was not intended to be a test of reciprocity, so there likely was a slight lateral offset of the source line from the receiver line. Likewise the sender and receiver arrays (clusters) may have a slightly different geometry. The earth dips in Figure 6.25 happen to be quite small although lateral velocity variation is known to be a problem in this area.

In Figure 6.26, three seismograms were plotted on top of their reciprocals. Pairs were chosen at near offset, at mid range, and at far offset. You can see that reciprocal seismograms usually have the same polarity, and often have nearly equal amplitudes. (The figure shown is the best of three such figures I prepared).

Each constant time slice in Figure 6.27 shows the reciprocity of many seismogram pairs. Midpoint runs horizontally over the same range as in Figure 6.25. Offset is vertical. Data is not recorded near the vibrators leaving a gap in the middle. To minimize irrelevant variations, moveout correction was done before making the time slices. (There is a missing source that shows up on the left side of the figure). A movie of panels like Figure 6.27 shows that the bilateral symmetry you see in the individual panels is characteristic of all times. Notice however that there is a significant departure from reciprocity on the one-second time slice around midpoint 120.

In the laboratory, reciprocity can be established to within the accuracy of measurement. This can be excellent. (See White's example in FGDP). In the field, the validity of reciprocity will be dependent on the degree that the required con-

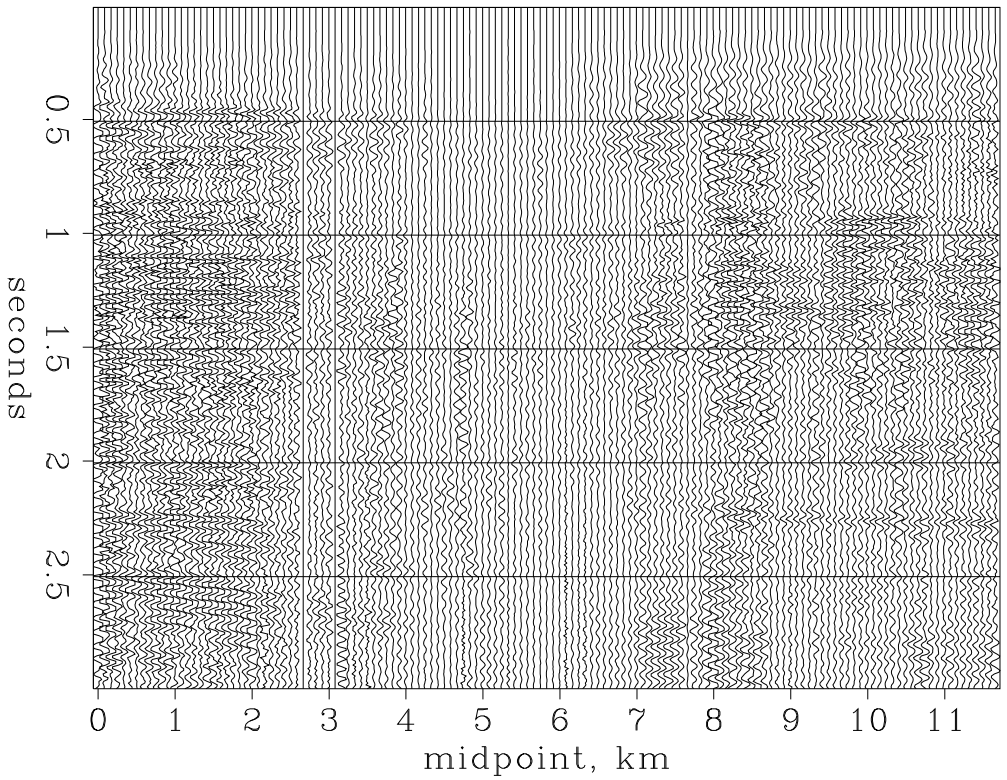

Figure 6.25: Constant-offset section from the Central Valley of California. (Chevron) ofs-toldi [ER]

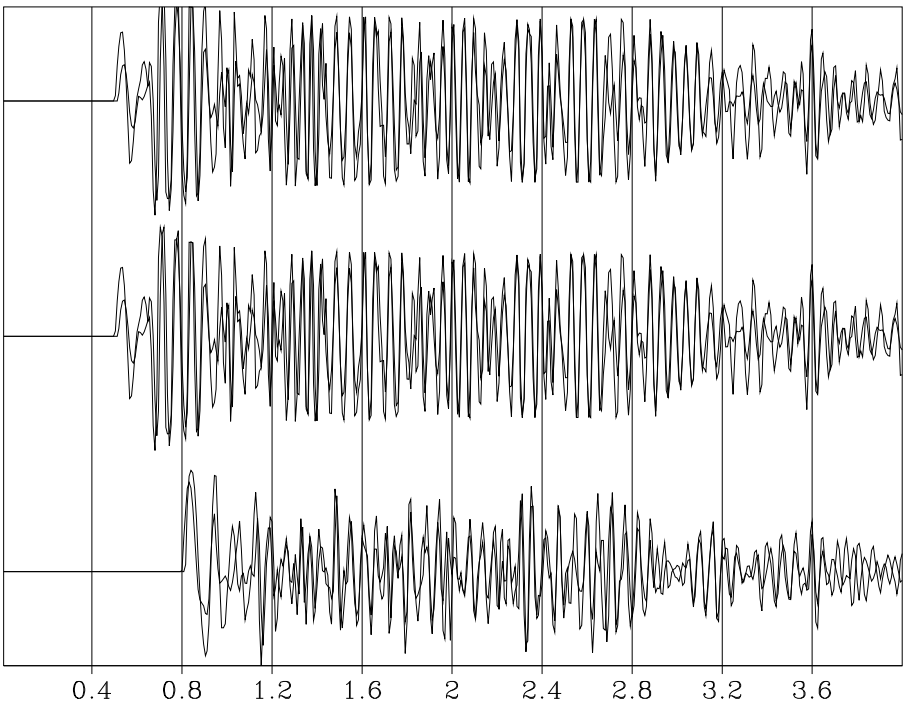

Figure 6.26: Overlain reciprocal seismograms. ofs-reciptrace [ER]

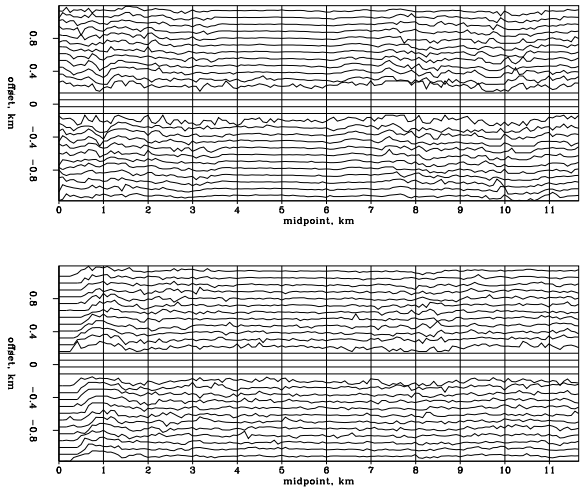

Figure 6.27: Constant time slices after NMO at 1 second and 2.5 seconds. [ER] ofs-recipslice

ditions are fulfilled. A marine air gun should be reciprocal to a hydrophone. A land-surface weight-drop source should be reciprocal to a vertical geophone. But a buried explosive shot need not be reciprocal to a surface vertical geophone because the radiation patterns are different and the positions are slightly different. Fenati and Rocca [1984] studied reciprocity under varying field conditions. They reported that small positioning errors in the placement of sources and receivers can easily create discrepancies larger than the apparent reciprocity discrepancy. They also reported that theoretically reciprocal experiments may actually be less reciprocal than presumably nonreciprocal experiments.

Geometrical complexity within the earth does not diminish the applicability of the principle of linearity. Likewise, geometrical complexity does not reduce the applicability of reciprocity. Reciprocity does not apply to sound waves in the presence of wind. Sound goes slower upwind than downwind. But this effect of wind is much less than the mundane irregularities of field work. Just the weakening of echoes with time leaves noises that are not reciprocal. Henceforth we will presume that reciprocity is generally applicable to the analysis of reflection seismic data.

6.3.2 The survey-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 has 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 $P(s, g, t)$.

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 survey sinking, it is necessary to downward continue the sources as well as the geophones. We already 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 survey-sinking concept of imaging is summarized by

$$\text{Image}(x, z) \quad = \quad \text{Wave}(s = x, g = x, z, t = 0) \tag{6.23}$$

For good quality data, i.e. 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.

6.3.3 Review of the paraxial wave equation

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

$$\frac{\partial t}{\partial x} \quad = \quad \frac{\sin \theta}{v} \quad = \quad p \tag{6.24}$$

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

$$\frac{\partial t}{\partial z} \quad = \quad -\frac{\cos \theta}{v} \quad = \quad -\sqrt{\frac{1}{v^2} - \left(\frac{\partial t}{\partial x}\right)^2} \tag{6.25}$$

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

$$\frac{\partial U}{\partial z} = -\frac{\partial t}{\partial z}\frac{\partial U}{\partial t} \tag{6.26}$$

$$U = f\left(t - \int_0^z \frac{\partial t}{\partial z}\, dz\right) \tag{6.27}$$

The partial derivatives in equation (6.26) are taken to be at constant x, just as is equation (6.25). After inserting (6.25) into (6.26) we have

$$\frac{\partial U}{\partial z} = \sqrt{\frac{1}{v^2} - \left(\frac{\partial t}{\partial x}\right)^2}\ \frac{\partial U}{\partial t} \tag{6.28}$$

Fourier transforming the wavefield 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 $\partial t/\partial x = k_x/\omega$. With this, (6.28) becomes

$$\frac{\partial U}{\partial z} = -i\omega \sqrt{\frac{1}{v^2} - \frac{k_x^2}{\omega^2}}\ U \tag{6.29}$$

The solutions to (6.29) 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 (6.29) has only upcoming solutions. Chapter 4 taught us 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 $v(x)$.

6.3.4 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

$$\frac{\partial t}{\partial z_g} = -\sqrt{\frac{1}{v^2} - \left(\frac{\partial t}{\partial g}\right)^2} \tag{6.30}$$

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

$$\frac{\partial t}{\partial z_s} = -\sqrt{\frac{1}{v^2} - \left(\frac{\partial t}{\partial s}\right)^2} \tag{6.31}$$

Both (6.30) and (6.31) require minus signs because the travel time decreases as either geophones or shots move down.

Simultaneously downward project both the shots and geophones by an identical vertical amount $dz = dz_g = dz_s$. The travel-time change is the sum of (6.30) and (6.31), namely,

$$dt \quad = \quad \frac{\partial t}{\partial z_g}\, dz_g \;+\; \frac{\partial t}{\partial z_s}\, dz_s \quad = \quad \left(\frac{\partial t}{\partial z_g} \;+\; \frac{\partial t}{\partial z_s} \right)\, dz \tag{6.32}$$

or

$$\frac{\partial t}{\partial z} \quad = \quad - \left(\sqrt{\frac{1}{v^2} \;-\; \left(\frac{\partial t}{\partial g} \right)^2} \;+\; \sqrt{\frac{1}{v^2} \;-\; \left(\frac{\partial t}{\partial s} \right)^2} \right) \tag{6.33}$$

This expression for $\partial t/\partial z$ may be substituted into equation (6.26):

$$\frac{\partial U}{\partial z} \quad = \quad + \left(\sqrt{\frac{1}{v^2} \;-\; \left(\frac{\partial t}{\partial g} \right)^2} \;+\; \sqrt{\frac{1}{v^2} \;-\; \left(\frac{\partial t}{\partial s} \right)^2} \right) \frac{\partial U}{\partial t} \tag{6.34}$$

Three-dimensional Fourier transformation converts upcoming wave data $u(t, s, g)$ to $U(\omega, k_s, k_g)$. Expressing equation (6.34) in Fourier space gives

$$\frac{\partial U}{\partial z} \quad = \quad - i\omega \left[\sqrt{\frac{1}{v^2} \;-\; \left(\frac{k_g}{\omega} \right)^2} \;+\; \sqrt{\frac{1}{v^2} \;-\; \left(\frac{k_s}{\omega} \right)^2} \right] U \tag{6.35}$$

Recall the origin of the two square roots in equation (6.35). 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,

$$\frac{\partial U}{\partial z} \quad = \quad \left[\sqrt{\left(\frac{-i\omega}{v}(g) \right)^2 - \frac{\partial^2}{\partial g^2}} \;+\; \sqrt{\left(\frac{-i\omega}{v}(s) \right)^2 - \frac{\partial^2}{\partial s^2}} \right] U \tag{6.36}$$

Equation (6.36) is known as the double-square-root (DSR) equation in shot-geophone space. It might be more descriptive to call it the survey-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.

6.3.5 The DSR equation in midpoint-offset space

By converting the DSR equation to midpoint-offset space we will be able 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

$$y = \frac{g + s}{2} \tag{6.37}$$

$$h = \frac{g - s}{2} \tag{6.38}$$

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:

$$\frac{\partial t}{\partial g} = \frac{\partial t}{\partial y}\frac{\partial y}{\partial g} + \frac{\partial t}{\partial h}\frac{\partial h}{\partial g} = \frac{1}{2}\left(\frac{\partial t}{\partial y} + \frac{\partial t}{\partial h}\right) \tag{6.39}$$

$$\frac{\partial t}{\partial s} = \frac{\partial t}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial t}{\partial h}\frac{\partial h}{\partial s} = \frac{1}{2}\left(\frac{\partial t}{\partial y} - \frac{\partial t}{\partial h}\right) \tag{6.40}$$

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 (6.37) and (6.38) and then Fourier transformed to (k_y, k_h)-space. The question is then, what form would the double-square-root equation (6.35) take in terms of the spatial frequencies (k_y, k_h)? Define the seismic data field in either coordinate system as

$$U(s, g) = U'(y, h) \tag{6.41}$$

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 (6.41) gives

$$\frac{\partial U}{\partial s} = \frac{\partial y}{\partial s}\frac{\partial U'}{\partial y} + \frac{\partial h}{\partial s}\frac{\partial U'}{\partial h} \tag{6.42}$$

$$\frac{\partial U}{\partial g} = \frac{\partial y}{\partial g}\frac{\partial U'}{\partial y} + \frac{\partial h}{\partial g}\frac{\partial U'}{\partial h} \tag{6.43}$$

and utilizing (6.37) and (6.38) gives

$$\frac{\partial U}{\partial s} = \frac{1}{2}\left(\frac{\partial U'}{\partial y} - \frac{\partial U'}{\partial h}\right) \tag{6.44}$$

$$\frac{\partial U}{\partial g} = \frac{1}{2}\left(\frac{\partial U'}{\partial y} + \frac{\partial U'}{\partial h}\right) \tag{6.45}$$

In Fourier transform space where $\partial/\partial x$ transforms to ik_x, equations (6.44) and (6.45), when i and $U = U'$ are cancelled, become

$$k_s = \frac{1}{2}(k_y - k_h) \tag{6.46}$$

$$k_g = \frac{1}{2}(k_y + k_h) \tag{6.47}$$

Equations (6.46) and (6.47) are Fourier representations of (6.44) and (6.45). Substituting (6.46) and (6.47) into (6.35) achieves the main purpose of this section, which is to get the double-square-root migration equation into midpoint-offset coordinates:

$$\frac{\partial}{\partial z} U = -i\frac{\omega}{v}\left[\sqrt{1-\left(\frac{vk_y + vk_h}{2\,\omega}\right)^2} + \sqrt{1-\left(\frac{vk_y - vk_h}{2\,\omega}\right)^2}\right] U \tag{6.48}$$

Equation (6.48) is the takeoff point for many kinds of common-midpoint seismogram analyses. Some convenient definitions that simplify its appearance are

$$G = \frac{v\ k_g}{\omega} \tag{6.49}$$

$$S = \frac{v\ k_s}{\omega} \tag{6.50}$$

$$Y = \frac{v\ k_y}{2\ \omega} \tag{6.51}$$

$$H = \frac{v\ k_h}{2\ \omega} \tag{6.52}$$

Chapter 3 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 (6.48) becomes slightly less cluttered:

$$\boxed{\frac{\partial}{\partial z}U = -\frac{i\omega}{v}\left(\sqrt{1-(Y+H)^2} + \sqrt{1-(Y-H)^2}\right)U} \tag{6.53}$$

Most present-day before-stack migration procedures can be interpreted through equation (6.53). Further analysis of it will explain the limitations of conventional processing procedures as well as suggest improvements in the procedures.

EXERCISES:

1 Adapt equation (6.48) to allow for a difference in velocity between the shot and the geophone.

2 Adapt equation (6.48) to allow for downgoing pressure waves and upcoming shear waves.

6.4 THE MEANING OF THE DSR EQUATION

The double-square-root equation contains 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

$$\frac{dU}{dz} = \frac{-i\omega}{v} \left(\sqrt{1 - G^2} + \sqrt{1 - S^2} \right) U \tag{6.54}$$

$$\frac{dU}{dz} = \frac{-i\omega}{v} \left(\sqrt{1 - (Y+H)^2} + \sqrt{1 - (Y-H)^2} \right) U \tag{6.55}$$

6.4.1 Zero-offset migration (H = 0)

One way to reduce the dimensionality of (6.55) 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:

$$\frac{dU}{dz} = -i\omega \frac{2}{v} \sqrt{1 - \frac{v^2 k_y^2}{4\omega^2}} U \tag{6.56}$$

In both places in equation (6.56) 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 $H = 0$, gave us the same migration equation we used in chapter 3. Setting $H = 0$ had the effect of making the survey-sinking concept functionally equivalent to the exploding-reflector concept.

6.4.2 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 (6.54) and (6.55) again become identical, and the resulting equation is once more the paraxial equation:

$$\frac{dU}{dz} = -i\omega \frac{2}{v} \sqrt{1 - \frac{v^2 k_h^2}{4\omega^2}} U \tag{6.57}$$

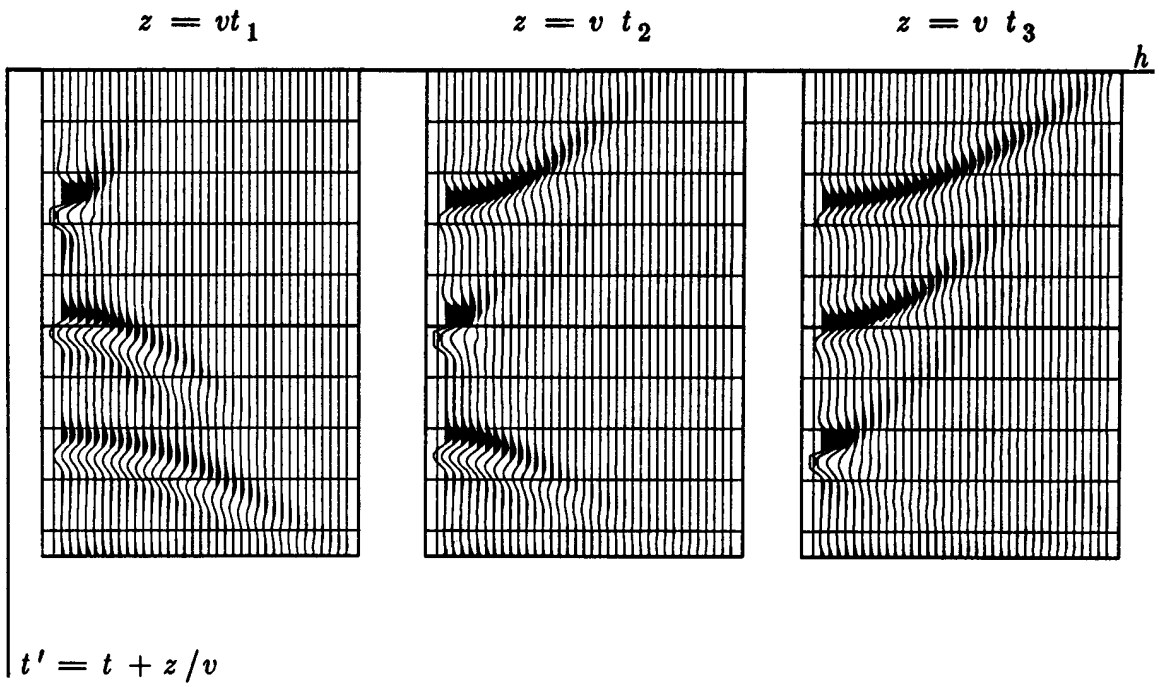

Figure 6.28: With an earth model of three layers, the common-midpoint gathers are three hyperboloids. Successive frames show downward continuation to successive depths where best focus occurs. ofs-dc2 [NR]

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 6.28.

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.

6.4.3 Conventional processing—separable approximation

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

$$\text{DSR}(Y,H) \quad = \quad \sqrt{1 \; - \; (Y-H)^2} \; + \; \sqrt{1 \; - \; (Y+H)^2} \tag{6.58}$$

In Fourier space, downward continuation is done with the operator $\exp(i\omega v^{-1}\,\text{DSR}\,z)$.

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 (6.58) contains terms like $Y^2 H^2$. 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 (though not the same equation) to the DSR operator gives

$$\text{SEP}(Y, H) \;=\; 2 \;+\; [\text{DSR}(Y, 0) \;-\; 2] \;+\; [\text{DSR}(0, H) \;-\; 2] \tag{6.59}$$

$$\text{SEP}(Y, H) \;=\; 2\,[1 \;+\; (\sqrt{1 - Y^2} \;-\; 1) \;+\; (\sqrt{1 - H^2} \;-\; 1)] \tag{6.60}$$

Notice that at $H = 0$ (6.59) and (6.60) become equal to the DSR operator. At $Y = 0$ (6.59) and (6.60) also become equal to the DSR operator. Only when both H and Y are nonzero does SEP depart from DSR.

The splitting of (6.59) and (6.60) 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 (6.55), 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 (6.60). The first is associated with time-to-depth conversion, the second with migration, and the third with normal moveout.

$$\text{SEP}(Y, H) \quad = \quad \text{TD} \;+\; \text{MIG}(Y) \;+\; \text{NMO}(H) \tag{6.61}$$

The approximation (6.59) and (6.60) can be interpreted as "standard processing." The first stage in standard processing is NMO correction. In (6.59) and (6.60) 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 seem wasteful to use it to migrate all offsets before stack. The result of doing so should be identical to after-stack migration.

6.4.4 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 (6.56) 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. The Fresnel zone has been extracted from a field profile in Figure 6.29.

The definition of the Fresnel zone involves a frequency. For practical purposes we may just look at zero crossings. Examining Figure 6.29 near one second we see a variety of frequencies. In the interval between $t = 1.0$ and $t = 1.1$ I see about two wavelengths of low frequencies and about 5 wavelengths of high frequencies. The highest frequencies are the main concern, because they define the limit of seismic resolution. The higher frequency has about 100 half wavelengths between time zero and a time of one second. As a rough generality, this observed value of 100 applies to all travel times. That is, at any travel time, the highest frequency that has meaningful spatial correlations is often observed to have a half period of about 1/100 of the total travel time. We may say that the quality factor Q of the earth's sedimentary crust is often about 100. So the angle that we are typically thinking about is $\cos 8^\circ = .99$.

Theoretically, 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

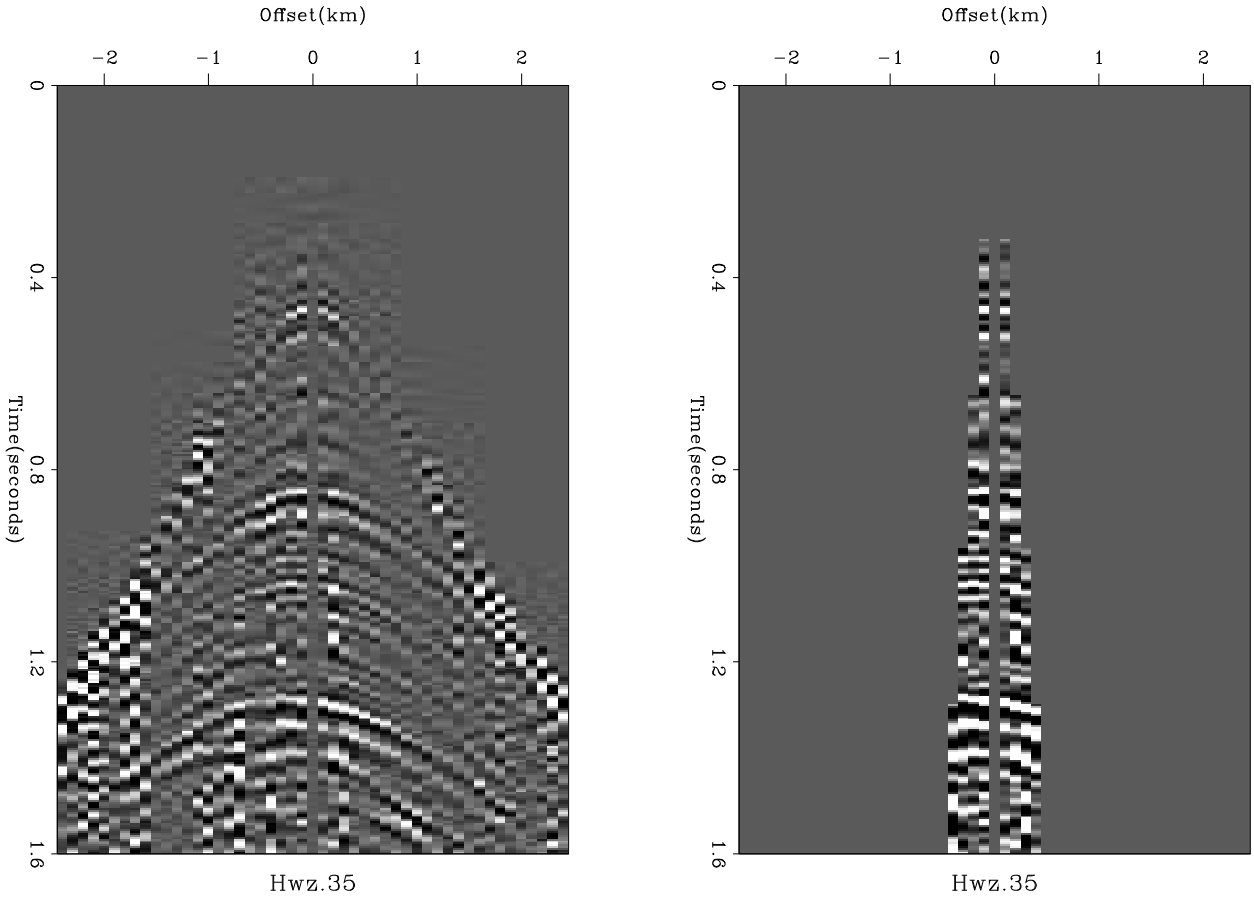

Figure 6.29: (left) A land profile from Denmark (Western Geophysical) with the Fresnel zone extracted and redisplayed (right). ofs-denmark [ER]

to migrate in a significantly different way. It would be nice if we could find an equation to downward continue data that is stacked at velocities other than infinite velocity.

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, equations (6.54) and(6.55) can be expressed in 15° retarded, space-domain form. Thus,

$$\left[\frac{\partial}{\partial z} + \frac{v}{-i\omega\, 8} \left(\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial h^2} \right) \right] U' = 0 \tag{6.62}$$

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,

$$\left(\frac{\partial}{\partial z} + \frac{v}{-i\omega\, 8} \frac{\partial^2}{\partial y}^2 \right) \left(\int U\, dh \right) + \frac{v}{-i\omega\, 8} \frac{\partial U}{\partial h} \Bigg|_{\substack{h = +\infty \\ h = -\infty}} = 0 \tag{6.63}$$

The wave should vanish at infinite offset and so should its horizontal offset derivative. Thus the last term in (6.63) should vanish. So, setting $H = 0$ has the meaning

$$\text{(Paraxial operator (vertical stack} = 0 \tag{6.64}$$

A problem in the development of (6.64) was that, twice, it was assumed that velocity is independent of offset: first, when the thin-lens term was omitted from (6.62), 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.

6.4.5 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 6.30.

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

$$Y = \sin\alpha \cos\beta \tag{6.65}$$

$$H = \sin\beta \cos\alpha \tag{6.66}$$

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

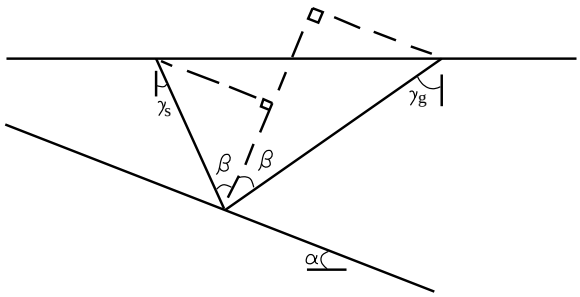

Figure 6.30: Geometry of a dipping bed. The line bisecting the angle 2β does not pass through the midpoint between g and s. (Clayton) ofs-clay [NR]

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" 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, (6.65) and (6.66) 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

$$\begin{aligned} (\frac{\pi}{2} - \gamma_s - \alpha) + \beta + \frac{\pi}{2} &= \pi \\ \gamma_s &= \beta - \alpha \end{aligned} \tag{6.67}$$

Adding up the angles around the larger triangle gives

$$\gamma_g = \beta + \alpha \tag{6.68}$$

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 equations (6.46), (6.47), (6.49), (6.50), (6.51) and (6.52), the definitions of apparent angles Y and H,

$$Y - H = S = \frac{v\,k_s}{\omega} = v\,\frac{dt}{ds} = -\sin\gamma_s = \sin(\alpha - \beta) \tag{6.69}$$

$$Y + H = G = \frac{v\,k_g}{\omega} = v\,\frac{dt}{dg} = +\sin\gamma_g = \sin(\alpha + \beta) \tag{6.70}$$

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

$$Y = \frac{1}{2}\sin(\alpha + \beta) + \frac{1}{2}\sin(\alpha - \beta) = \sin\alpha\cos\beta \tag{6.71}$$

$$H = \frac{1}{2}\sin(\alpha + \beta) - \frac{1}{2}\sin(\alpha - \beta) = \sin\beta\cos\alpha \tag{6.72}$$

6.4.6 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. S could also be set equal a constant, to simulate a downgoing Snell wave.

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 wavefield 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, $H = pv$. (A fuller explanation of slant stack comes later.) 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 $H = pv$ yields

$$k_z \quad = \quad -\frac{\omega}{v} \left[\sqrt{1 - (Y + pv)^2} + \sqrt{1 - (Y - pv)^2} \right] \tag{6.73}$$

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 [1982].

6.4.7 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.

Chapter 7

Velocity and dip moveout

7.1 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 gives as a byproduct estimates 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. So the latter part of this section really belongs to a research monograph with the facetious title "Theory That Should Work Out Soon." More advanced ideas of velocity estimation are later. Wave-equation stacking and velocity-determination methods are ingenious. 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.

7.1.1 Normal moveout (NMO)

Normal moveout correction (NMO) is a stretching of the time axis to make all seismograms look like zero-offset seismograms. In its simplest form, NMO is based on the Pythagorean relation $t_{NMO}^2 = t^2 - x^2/v^2$. In a constant velocity earth, the NMO correction would take the asymptote of the hyperbola family and move it up to $t = 0$. This abandons anything on the time axis before the first arrival, and stretches the remainder of the seismogram. The stretching is most severe near the first arrival, and diminishes at later times. In the NMO example in Figure 7.1 you will notice the low frequencies caused by the stretch.

NMO correction may be done to common-shot field profiles or to CMP gathers. NMO applied to a field profile makes it resemble a small portion of a zero-offset section. Then geologic structure is prominently exhibited. NMO on a CMP gather is the principal means of determining the earth's velocity-depth function. This is because CMP gathers are insensitive to earth dip.

Mathematically, the NMO transformation is a *linear* operation. It may seem paradoxical that a non-uniform axis-stretching operation is a linear operation, but axis stretching does satisfy the mathematical conditions of linearity. Do not confuse the widespread linearity condition with the less common condition of time invariance. Linearity requires only that for any decomposition of the original data P into parts (say P_1 and P_2) the sum of the NMOed parts is equal the NMO of the sum. Examples of decompositions include: (1) separation into early times and late times, (2) separation into even and odd time points, (3) separation into high frequencies and low frequencies, and (4) separation into big signal values and small ones.

To envision NMO as a linear operator, think of a seismogram as a vector. The NMO operator resembles a diagonal matrix, but the matrix contains *interpolation filters* along its diagonal, and the interpolation filters are shifted off from the diagonal to create the desired time delay.

7.1.2 Conventional velocity analysis

A conventional velocity analysis uses a collection of trial velocities. Each trial velocity is taken to be a constant function of depth and is used to moveout correct the data. Figure 7.2 (left) exhibits the CMP gather of Figure 7.1 (left) after moveout correction by a constant velocity. Notice that the events in the middle of the gather are nearly flattened, whereas the early events are undercorrected and later events are overcorrected. This is typical because the amount of moveout correction varies inversely with velocity (by Pythagoras), and the earth's velocity normally increases with depth. A measure of the goodness of fit of the NMO velocity to the earth velocity is found by

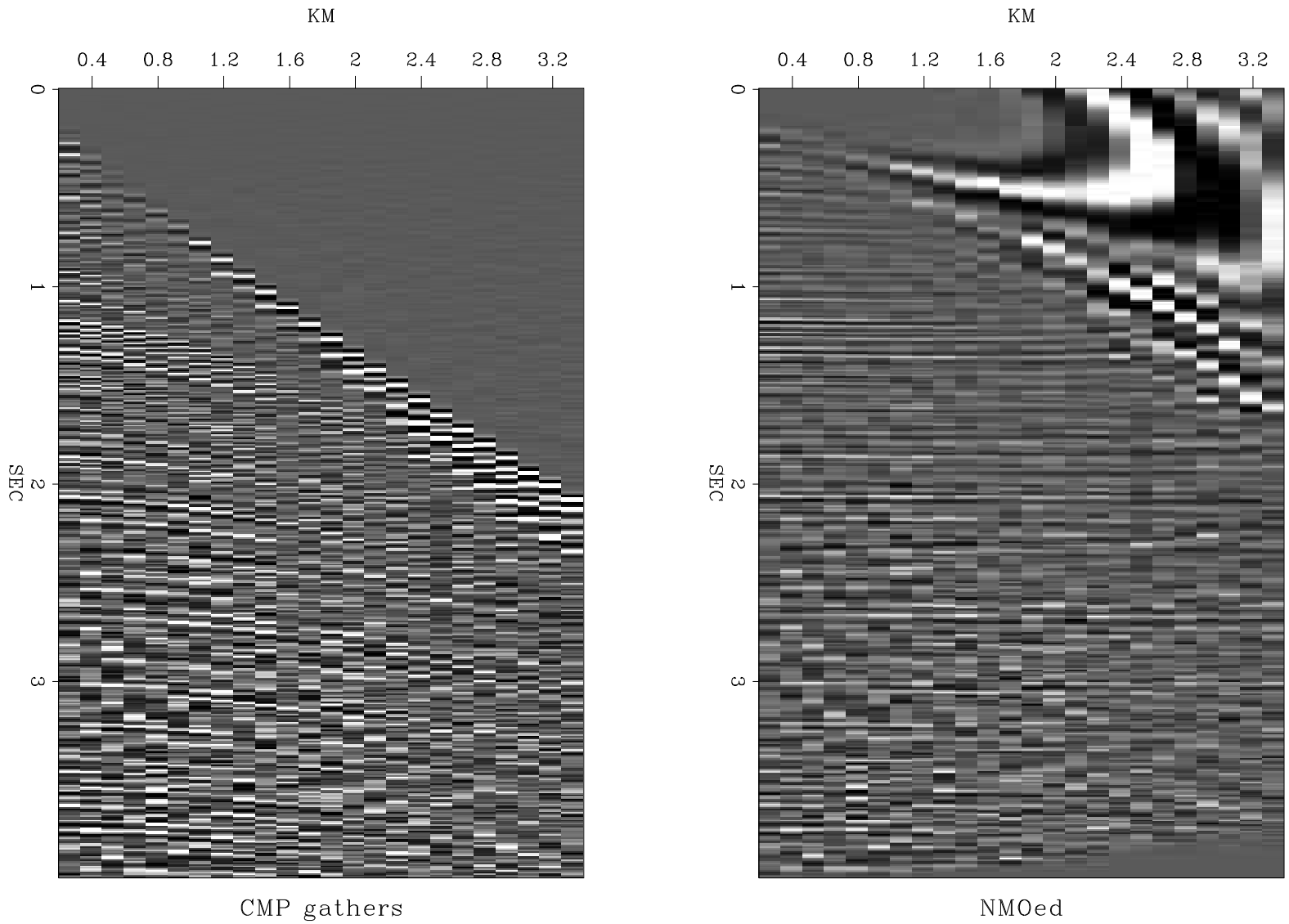

Figure 7.1: CMP gather (Western Geophysical) from the Gulf coast shown at the left was NMO corrected and displayed at the right. vdmo-cmpnmo [ER]

summing the CDP gather over offset. Presumably, the better the velocities match, the better (bigger) will be the sum. The process is repeated for many velocities. The amplitude of the sum, contoured as a function of time and velocity, is shown in Figure 7.2 (right).

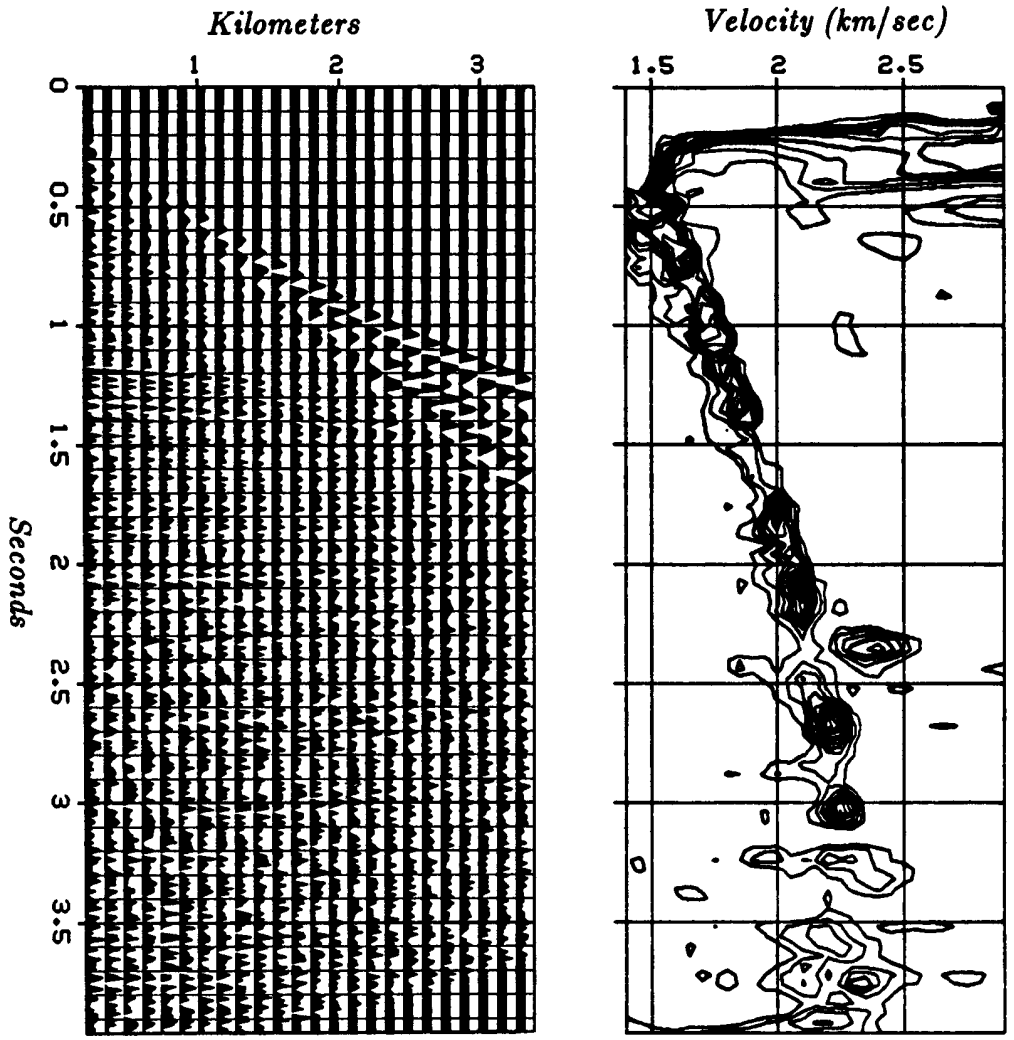

Figure 7.2: NMO at constant velocity with velocity analysis. (Hale) vdmo-cmphale [NR]

In practice additional steps may be taken before summing. The traces may be balanced (scaled to be equal) in their powers and in their spectra. Likewise the amplitude of the sum may be normalized and smoothed. (See Taner and Koehler [1969]). Also the data may be *edited* and *weighted* as explained in the next subsection.

The velocity giving the best stack is an average of the earth's velocity above the reflector. The precise definition of this average is deferred till later.

7.1.3 Mutes and weights

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. Figure 7.3 shows an example of a muted field profile. Weights and mutes have a substantial

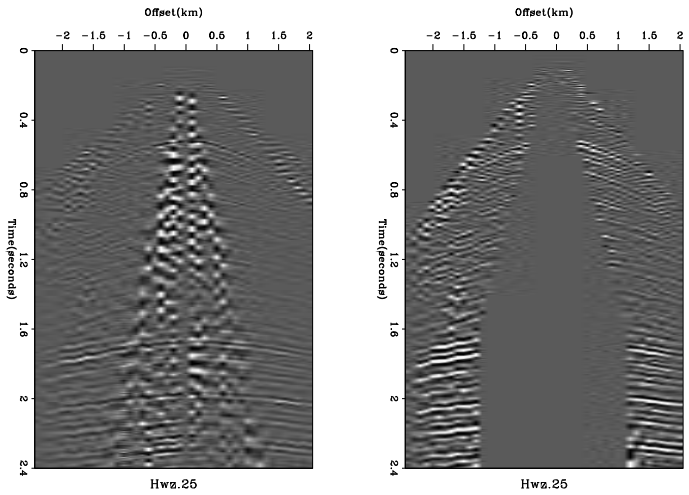

Figure 7.3: Left is a land profile from Alberta (Western Geophysical). On the right it is muted to remove ground roll (at center) and head waves (the first arrivals). vdmo-mute [ER]

effect on the quality of a stack. So 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 small values of r, energy is found that remains near the shot, such as falling dirt or water or slow ground roll.

At large values of r, there are problems with the first arrival. Here the NMO stretch is largest and most sensitive to the presumed velocity. The first arrival is often called a *head wave* or *refraction*. Experimentally, a head wave is a wave whose travel time appears to be a linear function of distance. Theoretically, a head wave is readily defined for layered media. The head wave has a ray that propagates horizontally along a layer boundary. In practice, a head wave may be weaker or stronger than the reflections. A strong head wave may be explained by the fact that reflected waves spread in three dimensions, while head waves spread in only two dimensions.

Muting may be regarded as weighting by zero. More general weights may be chosen to produce the most favorable CDP stack. A sophisticated analysis would certainly include noise and truncation. Let us do a simplified analysis. It leads to the most basic weighting function.

Ordinarily we integrate over offset along a hyperbola. Instead, think of the *three*-dimensional problem. You really wish to integrate over a hyperbola of revolution. Assume that the hyperboloid is radially symmetric. Weighting the integrand by h

allows the usual line integral to simulate integration over the hyperboloid of revolution. A second justification for scaling data by offset h before stacking is that there is less velocity information near zero offset, where there is little moveout, and more velocity information at wider offset where $\Delta t/\Delta h$ is larger.

7.1.4 NMO equations

The earth's velocity typically ranges over a factor of two or more within the depth range of a given data set. Thus the Pythagorean analysis needs reexamination. In practice, depth variable velocity is often handled by inserting a time variable velocity into the Pythagorean relation. (The classic reference, Taner and Koehler [1969], includes many helpful details). This approximation is much used, although it is not difficult to compute the correct nonhyperbolic moveout. Let us see how the velocity function $v(z)$ is mathematically related to the NMO. Ignoring dip, NMO converts common-midpoint gathers, one of which, say, is denoted by $P(h,t)$, to an earth model, say,

$$Q(h,z) \quad = \quad \text{earth}(z) \ \times \ \text{const}(h) \tag{7.1}$$

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

$$Q[h, z(h,t)] \quad = \quad P(h,t) \tag{7.2}$$

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

$$Q(h,z) \quad = \quad P[h, t(h,z)] \tag{7.3}$$

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.

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 the following equations over τ:

$$\frac{dt}{d\tau} \quad = \quad \frac{dz}{d\tau}\,\frac{dt}{dz} \quad = \quad v\,\frac{1}{v\ \cos\ \theta} \quad = \quad \frac{1}{\sqrt{1\ -\ p^2\,v(\tau)^2}} \tag{7.4}$$

$$\frac{dh}{d\tau} \quad = \quad \frac{dz}{d\tau}\,\frac{dh}{dz} \quad = \quad v\ \tan\ \theta \quad = \quad \frac{p\,v(\tau)^2}{\sqrt{1\ -\ p^2\,v(\tau)^2}} \tag{7.5}$$

(In equations (7.4) and (7.5) dt/dz and dh/dz are based on rays, not wavefronts). 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 usually are head waves arriving in the middle of the reflections. But once the job is done you can save the table and reuse it many times. The multibranching of the travel time curves at wide offset motivates a wave-equation based velocity analysis. The greatest velocity sensitivity occurs just where the classic hyperbolic assumption and the single-arrival assumption break down.

7.1.5 Linearity allows postponing statistical estimation

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 zeroes. Likewise, $(0,B)$ could be another copy of the gather where the *inner* traces have been replaced by zeroes. We could downward continue $(A,0)$ and separately 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 7.4 shows a dataset of three traces decomposed into three datasets, one for each trace.

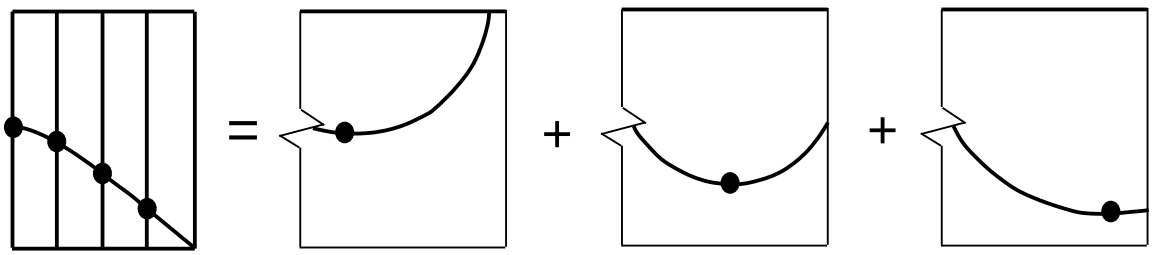

Figure 7.4: A three-trace CMP gather decomposed by traces. At the left, impulses on the data are interpolated, depicting a hyperbola. At the right, data points are expanded into migration semicircles, each of which goes through zero offset at the apex of the hyperbola. vdmo-decomp [NR]

Semicircles depict the separate downward continuation of each trace. Each semicircle goes through zero offset, giving the appropriately stretched, NMOed trace.

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 wavefields may get too far out of balance.

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.

7.1.6 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 that is not as long as the distance over which seismic energy propagates. Figure 7.5 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

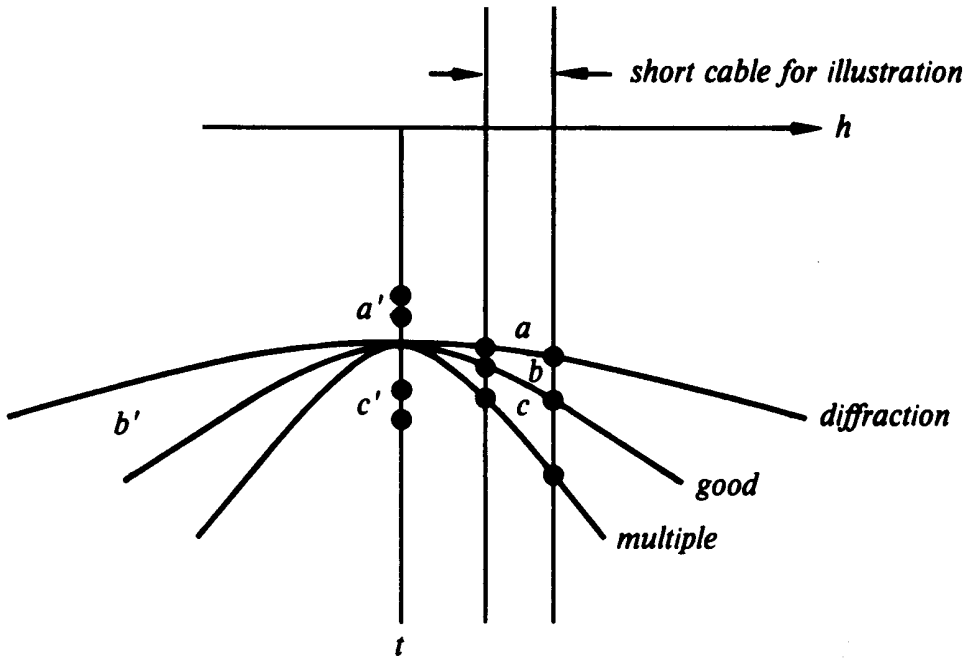

Figure 7.5: Normal moveout at the earth velocity brings the cable truncations on good events to a good place, causing no problems. The cable truncations of diffractions and multiples, however, move to a' and c', where they could be objectionable. Such corruption could make folly of sophisticated time-series analysis of the waveform found on a CDP stack. vdmo-nmotrunc [NR]

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.

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

$$\frac{\frac{k^2}{-i\omega}}{\alpha + \frac{k^2}{-i\omega}} \tag{7.6}$$

Notice that this filter greatly weakens the energy with small k, that is, 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. This completes the first iteration. Next the steps are 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 7.6.

The above procedure has ignored the possibility of dip in the midpoint direction.

This procedure is also 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 and a velocity spectral concept such as the ones developed later.

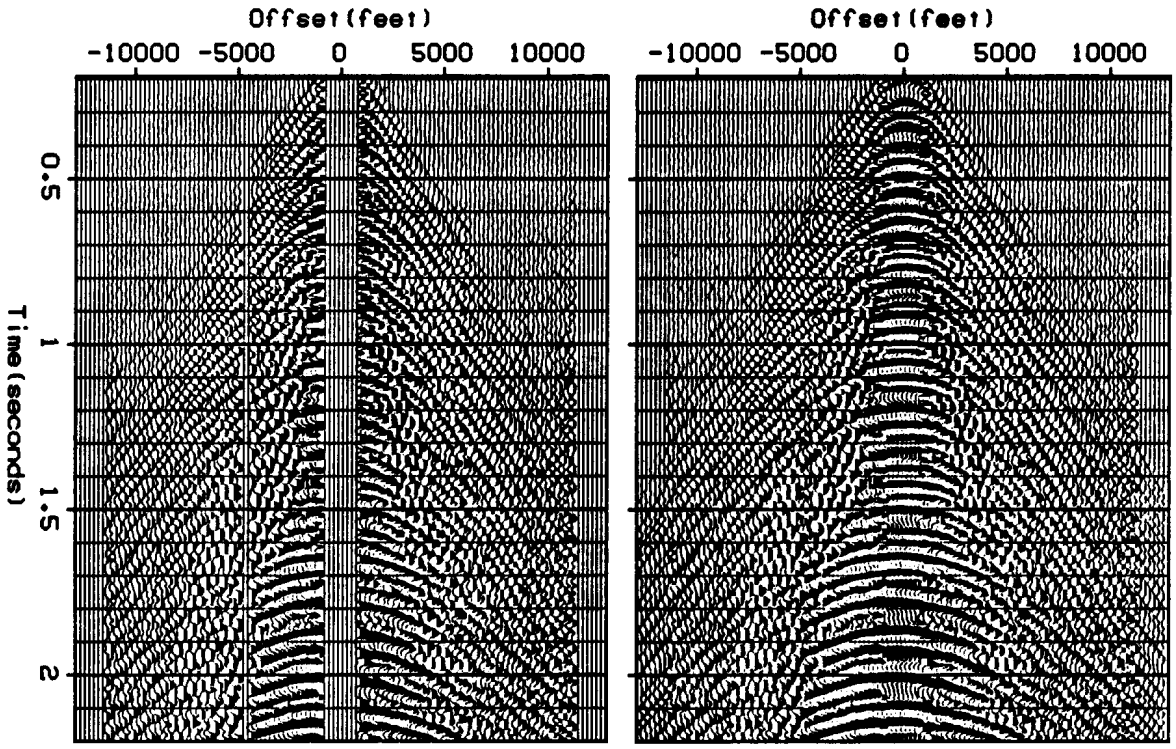

Figure 7.6: Field profile from Alaska with missing channels on the left (Western Geophysical), restored by iterative spatial filtering on the right. (Harlan) vdmo-missfil [NR]

7.1.7 In and out of velocity space

Summing a common-midpoint gather on a hyperbolic trajectory over offset yields a stack called a *constant-velocity* or a *CV* stack. A velocity space may be defined as a family of CV stacks, one stack for each of many velocities. CV stacking is a transformation from *offset* space to *velocity* space. CV stacking creates a (t, v)-space velocity panel from a (t, h)-space common-midpoint gather. Conventional industrial velocity estimation amounts to CV stacking supplemented by squaring and normalizing. Linear transformations such as CV stack are generally *invertible,* but the transformation to velocity space is of very high dimension. Forty-eight channels and 1000 time points make the transformation 48,000-dimensional. With present computer technology, matrices this large cannot be inverted by algebraic means. However, there are some excellent approximate means of inversion.

For unitary matrices, the transpose matrix equals the inverse matrix. In wave-propagation theory, a *transpose* operator is often a good approximation to an *inverse* operator. Thorson [1984] pointed out that the transpose operation to CV stacking is just about the same thing as CV stacking itself. To do the operation transposed to CV stacking, begin with a velocity panel, that is, a panel in (t, v)-space. To create some given offset h, each trace in the (t, v)-panel must be first compressed to undo the original NMO stretch. That is, events must be pushed from the zero-offset time that they have in the (t, v)-panel to the time appropriate for the given h. Then stack the (t, v)-panel over v to produce the seismogram for the given h. Repeat the process for all desired values of h. The program for transpose CV stack is like the program for CV stack itself, except that the stretch formula is changed to a compensating compression.

The inversion of a CV stack is analogous to inversion of slant stack or Radon transformation. That is, the CV stack is almost its own inverse, but you need to change a sign, and at the end, a filtering operation, like rho filtering, is also needed to touch up the spectrum, thereby finishing the job. It is the rho filtering that distinguishes *inverse* CV stack from the *transpose* of CV stack.

The word *transpose* refers to *matrix transpose.* It is difficult to visualize why the word transpose is appropriate in this case because we are discussing data spaces that are two-dimensional and operators that are four-dimensional. But if you will map these two- and four-dimensional objects to familiar one- and two-dimensional objects by a transformation then you will see that the word *transpose* is entirely appropriate. The rho-type filtering required for CV stacks is slightly more complicated than ordinary rho filtering—refer to Thorson [1984].

Figure 7.7 shows an example of Thorson's velocity space inversion. Panel D is the original common-midpoint gather. Next is panel LU, the approximate reconstruction of D from velocity space. The hyperbolic events are reconstructed much better than the random noise. The random noise was not reconstructed so well because the range of velocities in the CV stack was limited between water velocity and 3.5 km/sec. The next two panels (U) and (LTD) are theoretically related by the "rho"

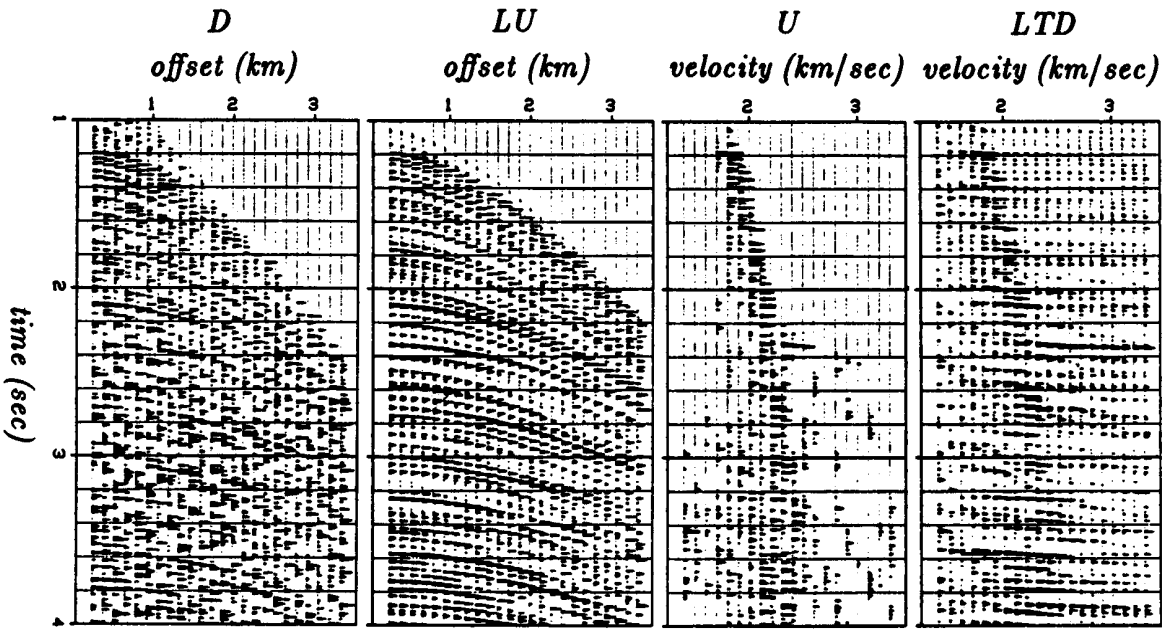

Figure 7.7: Panel D at the left is a CMP gather from the Gulf of Mexico (Western Geophysical). The second panel (LU) is reconstructed data obtained from the third panel (U) by inverse NMO and stack. The last panel (LTD) is a CV stack of the first panel. (Thorson) vdmo-thor [NR]

filtering. LTD is the CV stack of D. LU is the transpose CV stack of U.

It is worth noting that there is a substantial amount of work in computing a velocity panel. A stack must be computed for each velocity. Velocity discrimination by wave-equation methods will be described next. The wave-equation methods are generally cheaper, though not fully comparable in effect.

7.1.8 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 [1975] used this idea in a velocity-estimation scheme—see Figure 7.8.

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.

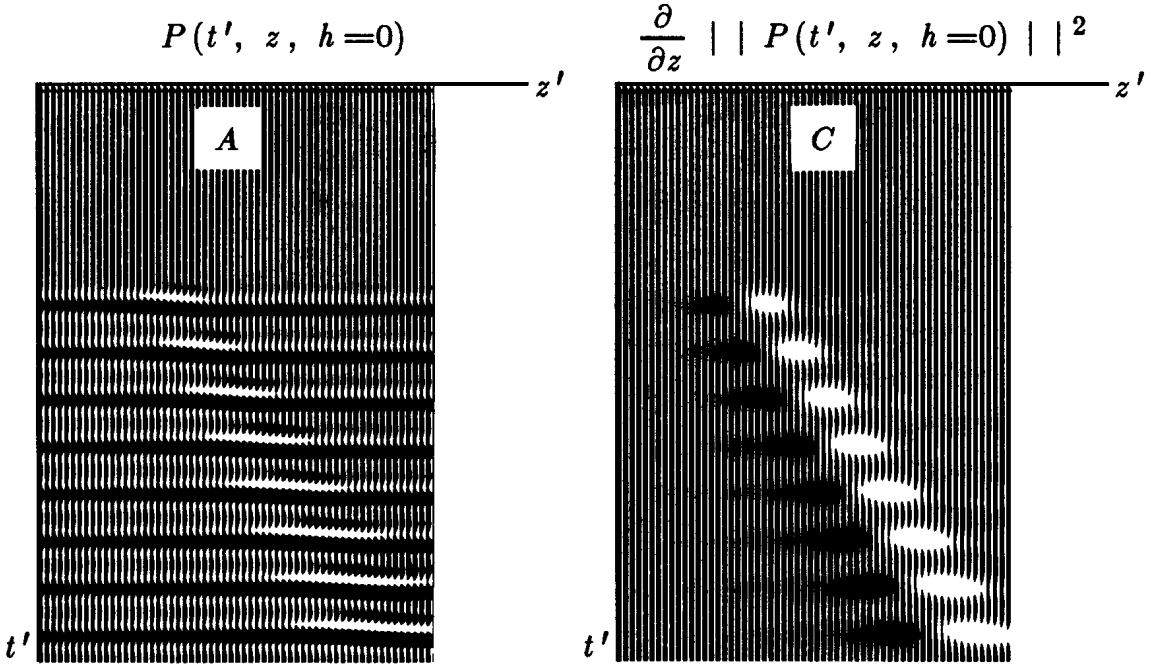

Figure 7.8: Two displays of the (z,t)-plane at zero offset. The earth model is eight uniformly spaced reflectors under a water layer (a family of hyperboloids in (h,t) at $z=0$). The left display is the zero-offset trace. The amplitude maximum at the focus is not visually striking, but the phase shift 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) vdmo-doherty [NR]

7.1.9 Decomposing a gather by velocity

A process will be defined that can partition a CMP gather, *both reflections and head waves,* 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 $\bar{v}(z)$ 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.

7.1.10 Reflected head waves on sections

It is common for an interpreter looking at a stacked section to identify a reflected head wave. Experimentally it is just a hyperbolic asymptote seen in (y, t)-space. Theoretically, it is a ray that moves away from a source along a horizontal interface until it encounters an irregularity, a fault perhaps, from which it reflects and returns toward the source. Reflected head waves are sometimes called reflected refractions. This event provides an easy 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 head wave to be identified, a special geological circumstance must be present—a scatterer strong enough to have its hyperbolic asymptote visible. 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 Larner's

streaks).

Velocity estimates made from reflections are averages of all the layers above the reflection point. To get depth resolution, it is necessary to subtract velocity estimates of different depth levels. Because of the subtraction, accuracy is lost. So with reflected waves, there is naturally a trade-off between accuracy and depth resolution. On the other hand, velocity estimates from head waves naturally have a high resolution in depth.

Processing seems to ignore or discriminate against the backscattered head wave, yet it 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 factor becomes significant: 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 of the power spectrum $U^* U$, or better yet of 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 7.9.

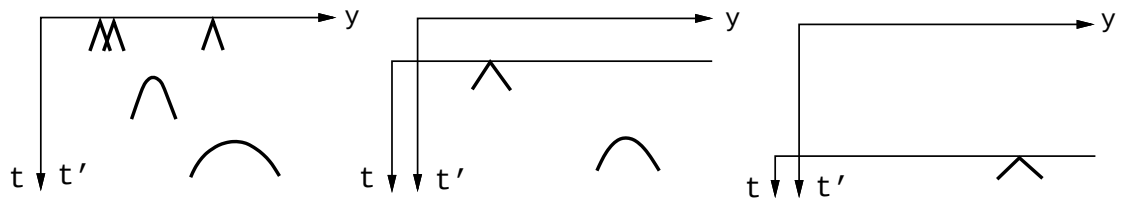

Figure 7.9: The dip-spectrum method of velocity determination. To find the velocity at any depth, seek the steepest dip on the section at that depth. On the left, at the earth's surface, you see the surface ground roll. In frames B and C the slowest events are the asymptotes of successively faster hyperbolas. vdmo-specz [NR]

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 near-surface velocity irregularity.

EXERCISES:

1 Assume that the data $P(y, h, t)$ is constant with 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 $1 \, + \, \mathrm{sgn}(\omega)\,\mathrm{sgn}(k_h)$; transform back to (h, t)-space.

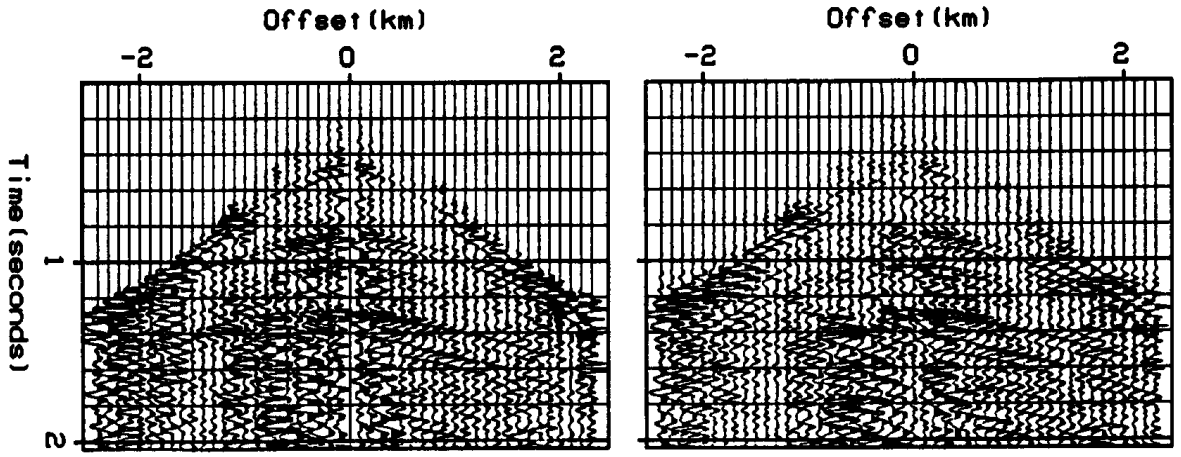

Figure 7.10: What is this? vdmo-exer [NR]

7.2 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 previous section, *which assume no dip.* In this section a means will be developed of estimating velocity in the presence of dip.

7.2.1 Dip moveout—Sherwood's devilish

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

$$t^2 \, v^2 \quad = \quad 4 \, (y - y_0)^2 \, \sin^2 \alpha \, + \, 4 \, h^2 \, \cos^2 \alpha \qquad (7.7)$$

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. They blur the velocity estimate and the stack.

In principle, migration before stack—some kind of implementation of the full DSR equation—solves this general problem. But where do we get the velocity to use in the migration equations? Although migration is somewhat insensitive to velocity when only small angles are involved, migration becomes sensitive to velocity when wide angles are involved.

The migration process should be thought of as being interwoven with the velocity estimation process. J.W.C. Sherwood [1976] showed how the two processes, migration and velocity estimation, should be interwoven. The moveout correction should be considered in two parts, one depending on offset, the NMO, and the other depending on dip. This latter process was conceptually new. Sherwood described the process as a kind of filtering, but he did not provide implementation details. He called his process *Devilish,* an acronym for "dipping-event velocity inequalities licked." The process was later described more functionally by Yilmaz as *prestack partial migration,* but the process has finally come to be called simply *dip moveout* (DMO). We will first see Sherwood's results, then Rocca's conceptual model of the DMO process, and finally two conceptually distinct, quantitative specifications of the process.

Figure 7.11 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 7.12 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.

7.2.2 Rocca's smear operator

Fabio Rocca developed a clear conceptual model for Sherwood's dip corrections. Figure 7.13 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—

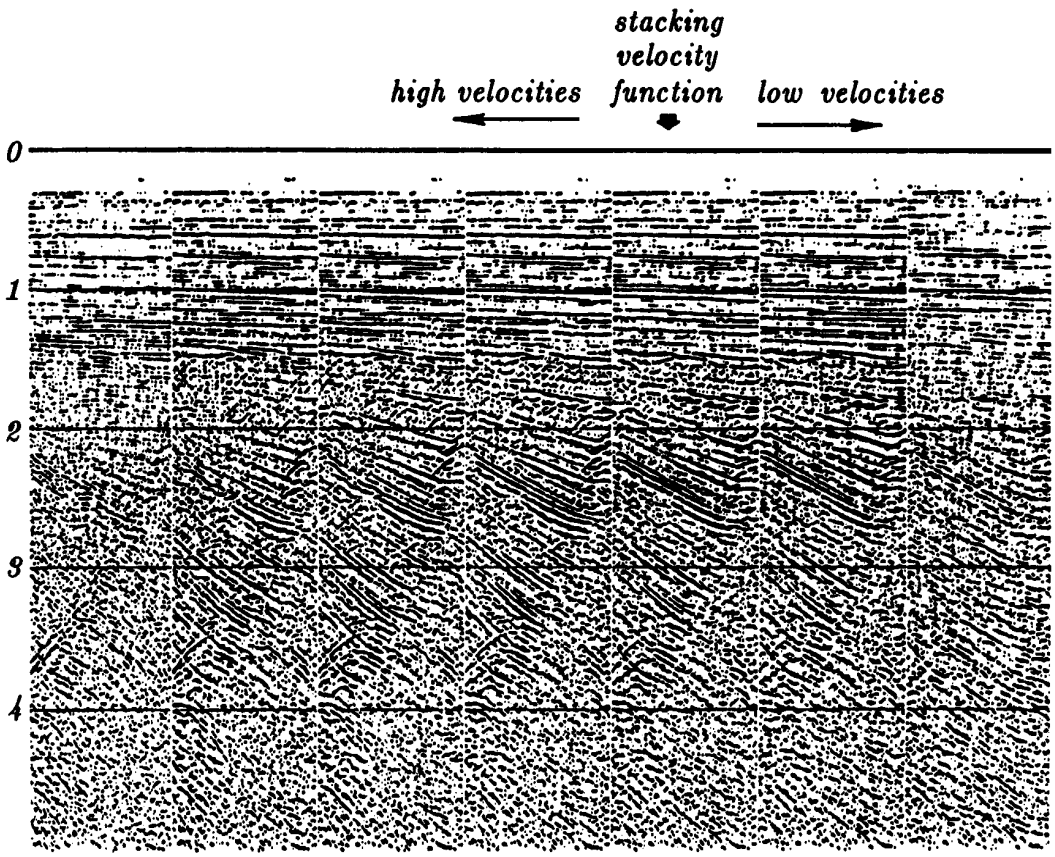

Figure 7.11: Conventional stacks with varying velocity. (distributed by Digicon, Inc.) vdmo-digicon [NR]

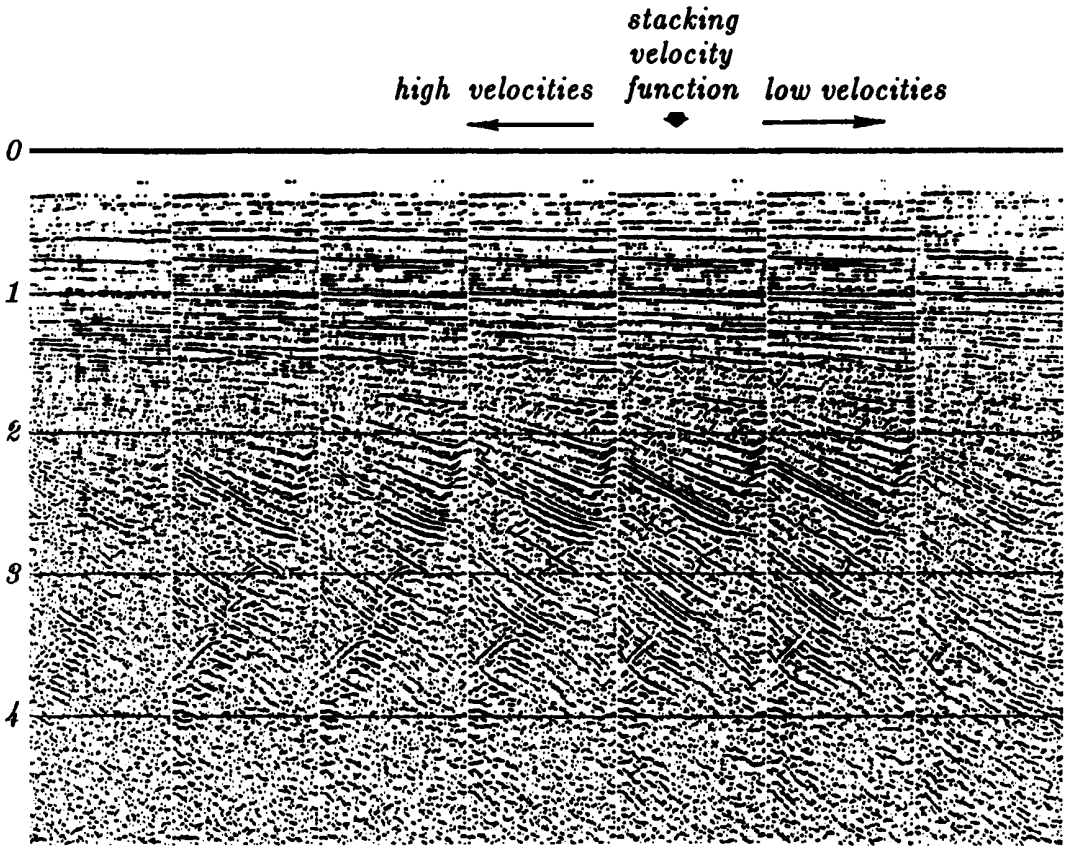

Figure 7.12: *Devilish* stacks with varying velocity. (distributed by Digicon, Inc.) vdmo-devlish [NR]

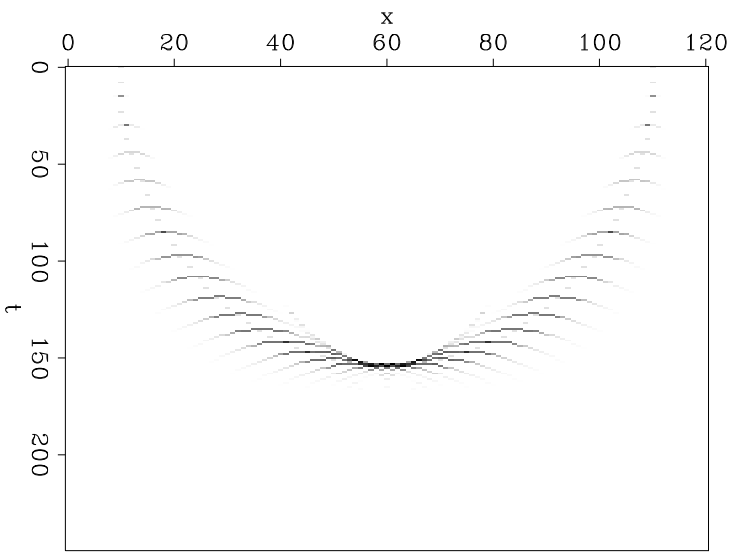

Figure 7.13: Rocca's prestack partial-migration operator is a superposition of hyperbolas, each with its top on an ellipse. Convolving (over midpoint) Rocca's operator onto a constant-offset section converts the constant-offset section into a zero-offset section. (Gonzalez) vdmo-rocca [ER]

constant-offset migration and zero-offset diffraction—gives the Rocca operator.

The Rocca operator is the curve of osculation in Figure 7.13, i.e., the smile-shaped curve where the hyperbolas reinforce one another. If the hyperbolas in Figure 7.13 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 is the end of a narrow ellipse, shown in Figure 7.14. We will omit the derivation of the equation

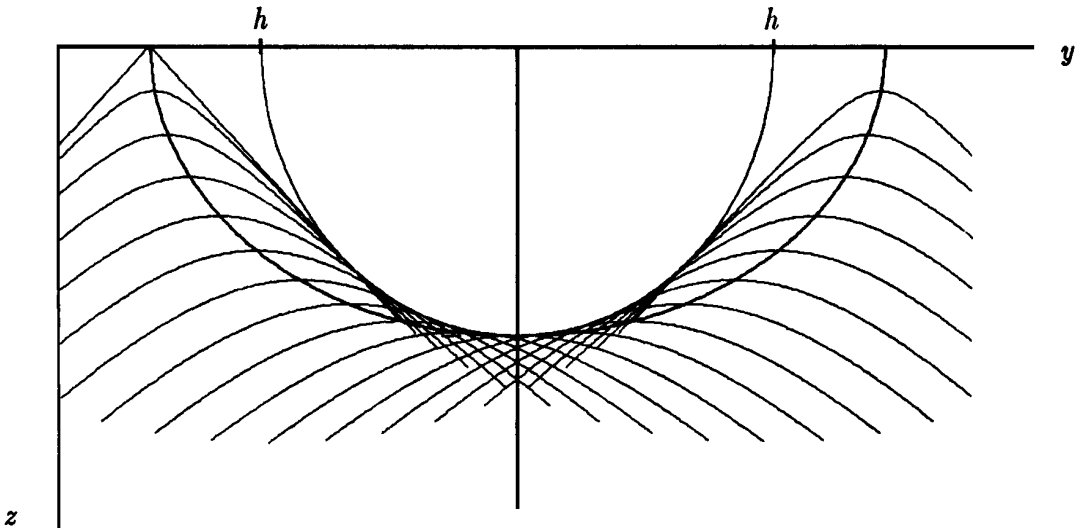

Figure 7.14: Rocca's smile. (Ronen) vdmo-rocca2 [NR]

for this curve which turns out to be

$$1 \quad = \quad \frac{(y \ - \ y_0)^2}{h^2} \quad + \quad \frac{t^2}{t_0^2} \tag{7.8}$$

The Rocca operator appears to be velocity independent, but it is not completely so because the curve cuts off at $dt/dy = 2/v$.

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 7.13 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 7.15 shows a superposition of several Rocca smiles for several values of t_0.

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. From Figure 7.15 we see that the energy in the dip moveout operator concentrates narrowly near the bottom. In the limiting case that h/vt_0 is small, the energy all goes to the bottom. When all