4.0 OFFSET, ANOTHER DIMENSION

Earlier chapters have assumed the circumstances required by the exploding reflector concept, namely that the shot and the geophone are located in the same place. The reality is that there is often as much as a 3-km horizontal separation between them. The 3-km offset is roughly comparable to the depth of the petroleum reservoirs that are the targets of search and analysis.

Offset is another dimension in the analysis of data. At the time of writing, this dimension is commonly represented in field operations by about 48 channels. But everyone seems to believe that this is not enough channels. Recording systems with as many as 1024 channels are coming into use.

The offset dimension adds three important aspects to reflection seismology. First, it enables us to routinely measure the velocity of seismic waves in rocks. This velocity has been assumed known in the previous chapters on migration. Second, it gives us data redundancy. We now have independent measurements of what should be the same thing. Superposition of the measurements (stacking) offers the potential for signal enhancement by destructive interference of noise. Third (a disadvantage), the fact that the offset is not zero means that procedures for migration take on another element of complexity. By the end of the chapter we will be trying to deal with three confusing subjects at the same time -- namely, dip, offset, and lateral velocity variation.

Theoretically it seems that offset should bring us the possibility of identifying rocks by observing the reflection coefficient as a function of angle, both for p waves and for p-to-s- converted waves. The reality seems to be that neither measurement can be made reliably, if at all. This is a very interesting subject for research. See for example *Geophysics* (1976) vol 41, no 5. It has a large potential for practical rewards. But the reasons for the frustration, and the resolution of the difficulty, are not the goals of this book, which are instead to enable us to deal effectively with that which is routinely observable.

What is "poor-quality data"?

There are vast regions of the world with good petroleum potential that are difficult to explore because of the difficulty of obtaining good-quality reflection seismic data. The reasons are often unknown. What is "poor-quality" data? From an experimentalist view almost all seismic data are good in the sense they are repeatable. The real problem is that the data make no sense.

A migrated zero-offset section would look random for data recorded at a location where the earth consisted of a random arrangement of point reflectors. Given the repeatability that we experience in data collection, are we to assume that a random appearance of data implies a random jumble of reflectors? With only zero-offset data we could assume little else. With the full range of offsets at our disposal, we can attempt a more thoughtful analysis. This chapter provides some of the required techniques. For example, suppose the earth really were a random jumble of point scatterers in a constant-velocity medium. The data would be a random function of time and a random function of the horizontal location of the shot-geophone midpoint, but after suitable processing, at each midpoint the data should be a perfectly hyperbolic function of shot-geophone offset. This would determine the earth's velocity exactly, even if the random scatterers were distributed in three dimensions, and the survey was only along a surface line.

This model could fail to explain the "poor-quality" data, then other models could be tried. The effects of random velocity variations in the near surface or the effects of multiple reflections could be analyzed. Noise in seismology can usually be regarded as a failure of analysis rather than as something polluting the data. Basically, it is the offset dimension which gives us the redundancy needed to try to figure out what is happening.

The Experiment Sinking Concept

The exploding-reflector concept has great utility in that it enables us to associate the seismic waves observed at zero offset from many experiments (say 1000 shot points) with the wave of a single thought experiment, the exploding-reflector experiment. The analogy has a few tolerable limitations with lateral velocity variations and multiple reflections. But the main limitation of the exploding-reflector concept is that it gives us no clue at all as to how we can interpret data recorded at nonzero offset. We need a broader imaging concept.

Start from the field data. We take this to be a survey line run along the x-axis. Mathematically, we presume to have the results of infinitely many experiments, a single experiment being to have a point source or shot at s on the x-axis and to record echos with geophones, at each possible location g on the x-axis. So the observed data is an upcoming wave which is a two-dimensional function of s and g, say $p(s,g)$. (Important practical questions about the actual spacing and extent of shots and geophones will be deferred until much later.)

Previous chapters have shown how to downward continue the upcoming wave. The procedures are quite independent of whether the source of the disturbance is an exploding reflector or a surface point source. (In this chapter, and in all previous chapters, we neglect multiple reflections. This will be seen to mean that the geophone sees upcoming waves, but not downgoing waves.)

To apply the imaging concept of experiment sinking we will need to learn how to downward continue the sources as well as the geophones. Actual techniques to do this will be developed later, but the main idea comes from the principle of reciprocity. Reciprocity says that the same seismogram should be recorded if the shot and geophone have their locations swapped. (A reason for the validity of reciprocity is that no matter how complicated a geometrical arrangement, the speed of sound is the same in either direction along a ray. A more thorough consideration of seismic reciprocity is found in FGDP.) We know how to downward continue geophones, and with reciprocity we can then interchange them with shots, so we really know how to downward continue shots.

We could downward continue shots and geophones to different levels, and they may be at different levels during the process, but for the final results we only require shots and geophones at the same level. That is, taking z_s to be the depth of the shots and z_g that of geophones, then 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 is a source and geophone located together on the reflector. The traveltime for the echo is zero. This experiment-sinking concept of imaging is summarized by

$$Image(x=g,z) \quad = \quad Wave(s=g,g,z,t=0) \tag{1}$$

4.1 CHEOP'S PYRAMID

We begin our study of offset by calculating some traveltimes for rays in some idealized environments. First we formally define the midpoint y between the shot and the geophone. We also define h to be half the horizontal offset between the shot and geophone. Thus

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

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

The definition of offset is taken to be $g - s$ rather than $s - g$ so that positive offset means waves moving in the positive x-direction. In the marine case this means the ship will sail negatively along the x-axis. The reason for using only half the total offset in the mathematical equations is to simplify and symmetrize many subsequent equations.

Sections and Gathers for Planar Reflectors

Data are defined experimentally in the space of (s,g). Equation (1) is a change of coordinates to the space of (y,h). Midpoint-offset coordinates are especially useful for interpretation and data processing. Since the data are also a function of traveltime t, the full dataset lies in a volume. It is impossible to find a satisfactory display of such a volume, so what is customarily done is to display slices. The names of slices vary from one company to the next. We will use the following names, which seem to be well known and clearly understood.

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

The simplest environment for reflection data is a single horizontal reflection interface, as shown in figure 1.

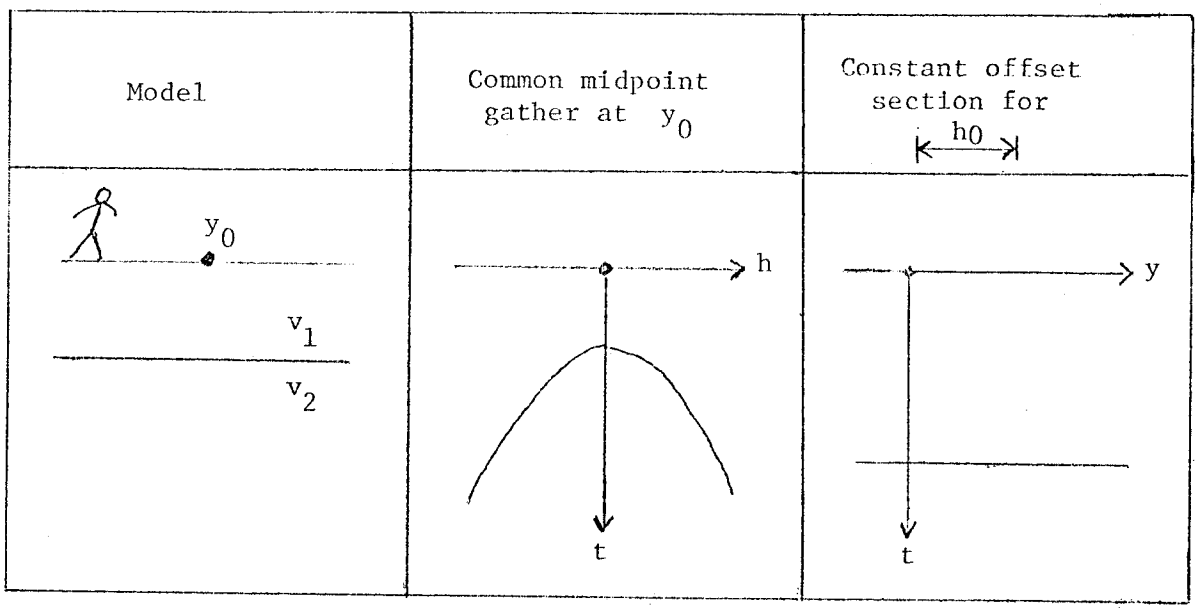

FIG. 1. A horizontal reflector and a resulting common-midpoint gather (center) and constant-offset section (right).

The asymptotes of the hyperbola on the common-midpoint gather are straight lines whose slope is the velocity v_1.

The next simple environment is to have a planar reflector which is oriented vertically rather than horizontally. In this case the wave propagation is along the air-earth interface. To avoid any confusion this might create we could also incline the reflector at a slight angle from vertical, as shown in figure 2.

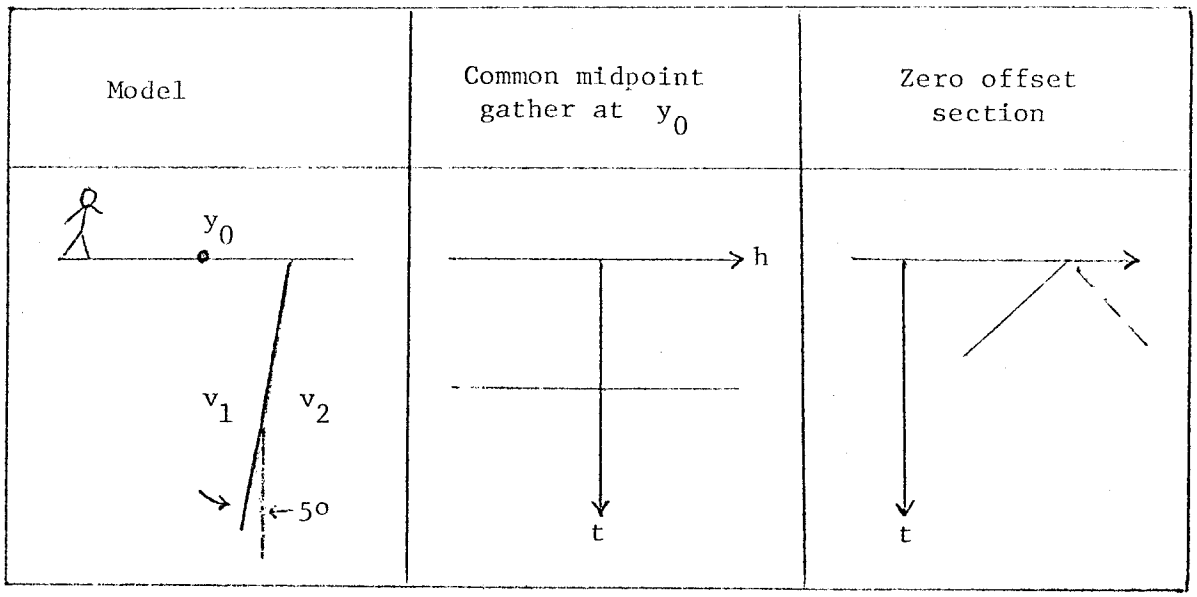

FIG. 2. Near-vertical reflector, a gather and a section.

In figure 2 the traveltime does not change as the offset changes. This is because the midpoint is held constant. As offset increases, the shot gets further from the reflector while the geophone gets closer in such a way as to keep the total time constant. In reality a planar reflector can have any dip between horizontal and vertical. Then the common-midpoint gathers lie between that of figure 1 and that of figure 2. The zero-offset section in figure 2 is a straight line. It turns out to be the asymptote of a family of hyperbolas. The slope of the asymptote is the velocity v_1.

The Point Response

The next simple geometry to consider 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 reflection data can always be considered to be a superposition of such point scatterers. Figure 3 shows an example. The curves in figure 3 show flat spots for the same reasons that some of the curves in figures 1 and 2 were straight lines.

The point scatterer geometry for a point located at (x,z) is shown in figure 4.

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

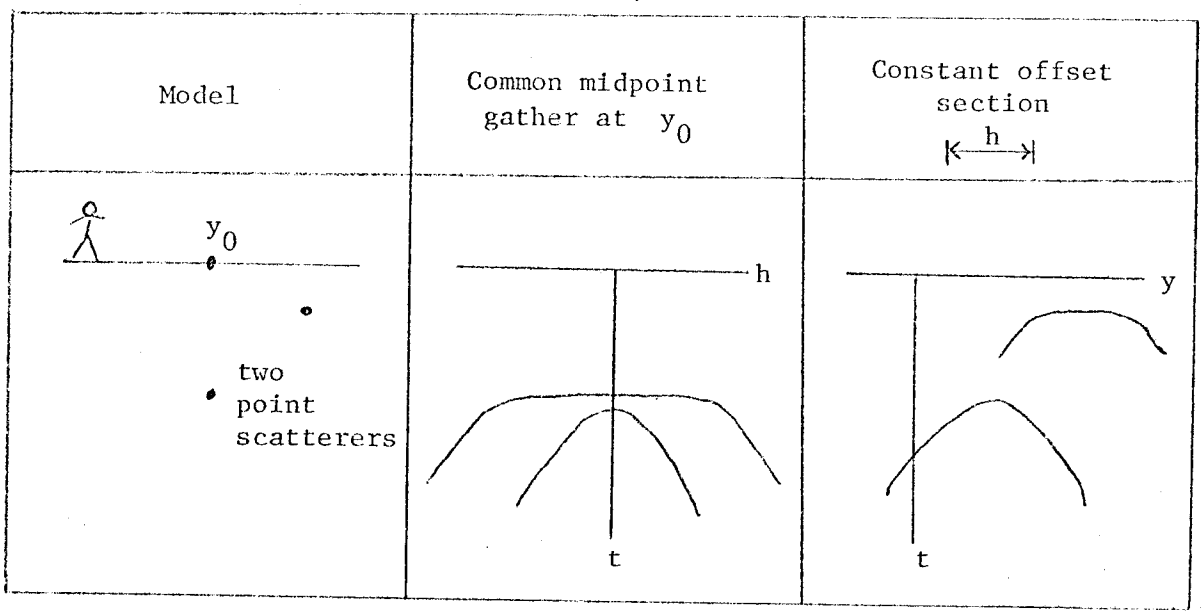

FIG. 3. Response of two point scatterers. Note flat spots.

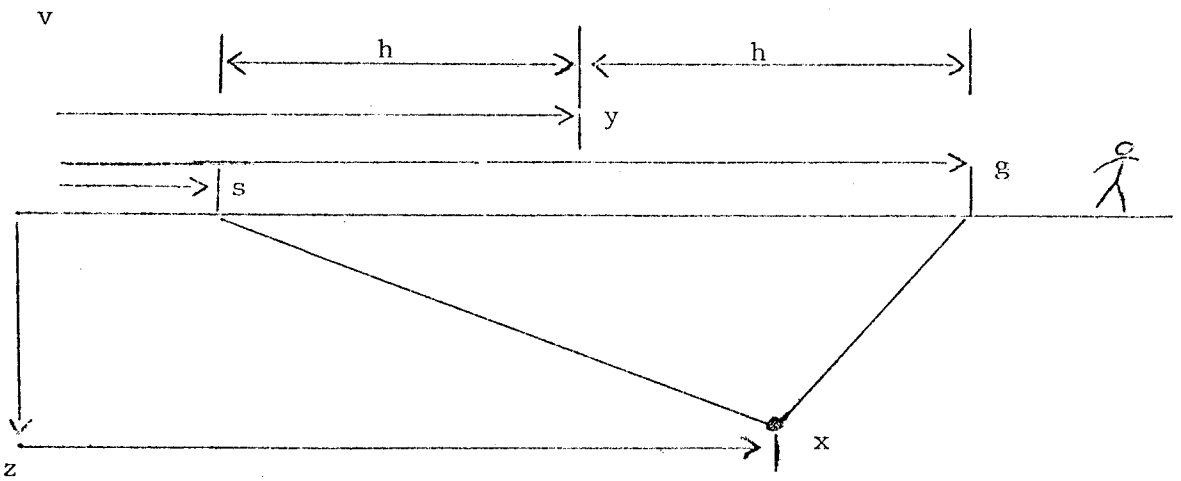

FIG. 4. Geometry of a point scatterer.

$$tv = \left[z^2 + (s - x)^2\right]^{1/2} + \left[z^2 + (g - x)^2\right]^{1/2} \qquad (2)$$

Cheop's 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 (2). The picture is one dimension more difficult than the conic section that relates to the exploding-reflector geometry.

To begin with, suppose the first square root in (2) is constant because everything in it is held constant. Then we have the familiar hyperbola in (g,t) space, the only difference being that there is an additive constant to the time. Taking instead the other square root to be constant, likewise we get the familiar hyperbola again in (s,t) space. Trying to think in (s,g) space, we think of traveltime being a function of s plus a function of g. I think of this as one coat hanger parallel to the s axis being hung from another coat hanger, which is parallel to the g axis.

The traveltime mountain in the (s,g) plane or the (y,h) plane has been plotted in figure 5. What doesn't show very clearly is that a cut through the mountain at a large constant t is a square whose corners have been smoothed. To see this asymptotic behavior, consider a point reflector very near the surface, say $z \to 0$. Then (2) becomes

$$t = |g - x| + |s - x| \qquad (3)$$

A constant value of t is the square contoured in (s,g)-space in figure 6.

The center of the square is located at $(s,g)=(x,x)$. Taking traveltime t to be 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. 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 is a traverse of h at constant y or a traverse of y at constant h. At the highest elevation on the traverse, you are walking along a flat horizontal step like the flat-topped hyperboloids of figure 3. To imagine a nonzero reflector depth you need some erosion to smooth the top and edges of the pyramid.

For rays which are near to the vertical, the traveltime curves are far from the hyperbola asymptotes. In this case the square roots in (2) may be expanded in Taylor series giving a parabola of revolution. This describes the eroded peak of the pyramid.

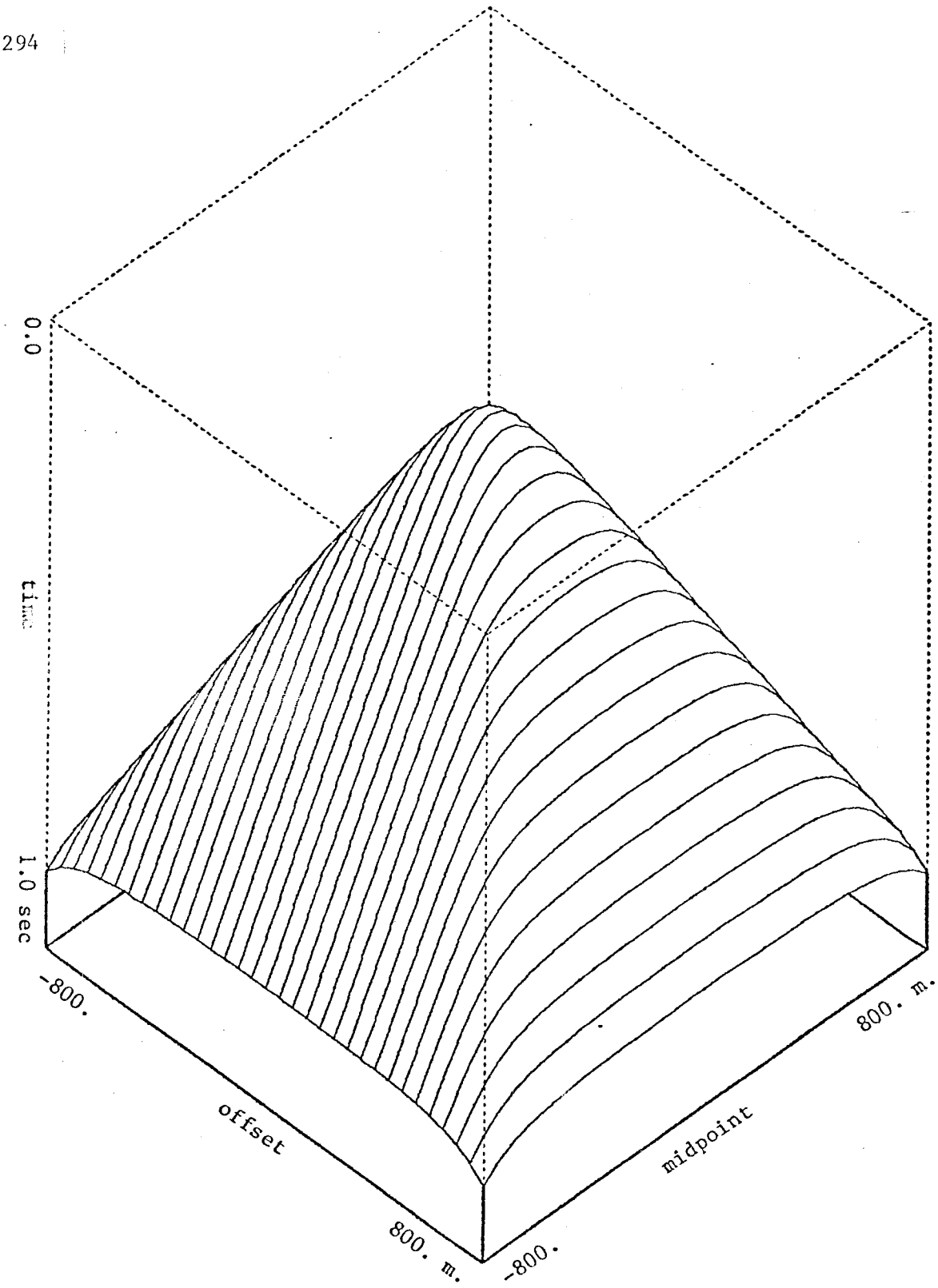

FIG. 5. (Ottolini) Equation (2) for fixed x and z.

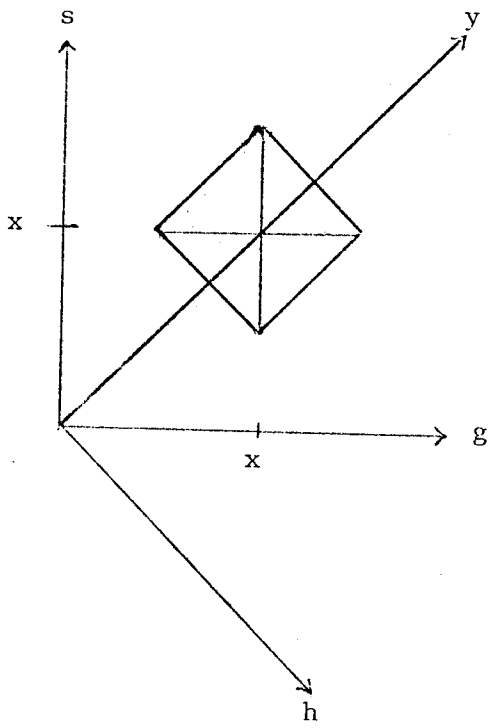

FIG. 6. A square in (h,y)-space is a contour of constant time.

The Migration Ellipse

Rather than regarding the reflection point (x,z) as being fixed, another insight into equation (2) is to regard offset h and the total traveltime t as fixed constants. Then the locus of point reflectors turns out to describe an ellipse in the plane of $(y-x,z)$. The reason it is an ellipse follows from the geometric definition. To make an ellipse, place a nail or tack into s on figure 4 and another into g. Connect the tacks by a string which is just long enough to go through x. An ellipse going through x may be constructed by moving a pencil along the string, keeping the string tight. The string keeps the total distance tv equal to a constant. It is left for the exercises to show that equation (2) can be cast in the standard mathematical form of an ellipse, namely a stretched circle.

Recall that one of the methods for the migration of zero-offset sections is to take every data value in (y,t)-space and use it to superpose an appropriate semicircle in (y,z)-space. For nonzero offset the circle should be generalized to an ellipse. What is unknown at this stage is the appropriate amplitude and phase distribution along the ellipse. How can the answer be expressed in terms of wave-extrapolation equations? After many false starts this question was answered and the answer is found in the next

section.

Exercises

1. The field data shown in figure 7 may be assumed to be of textbook quality. The speed of sound in water is about 1500 m/sec.
   a. Is this a common-shotpoint gather or a common-midpoint gather?
   b. Identify the events at A, B, and C.
2. Express equation (2) in the familiar mathematical form for an ellipse, namely a stretched circle. Hint: first square it, then move the remaining square root to one side of the equation and square again, removing all square roots.

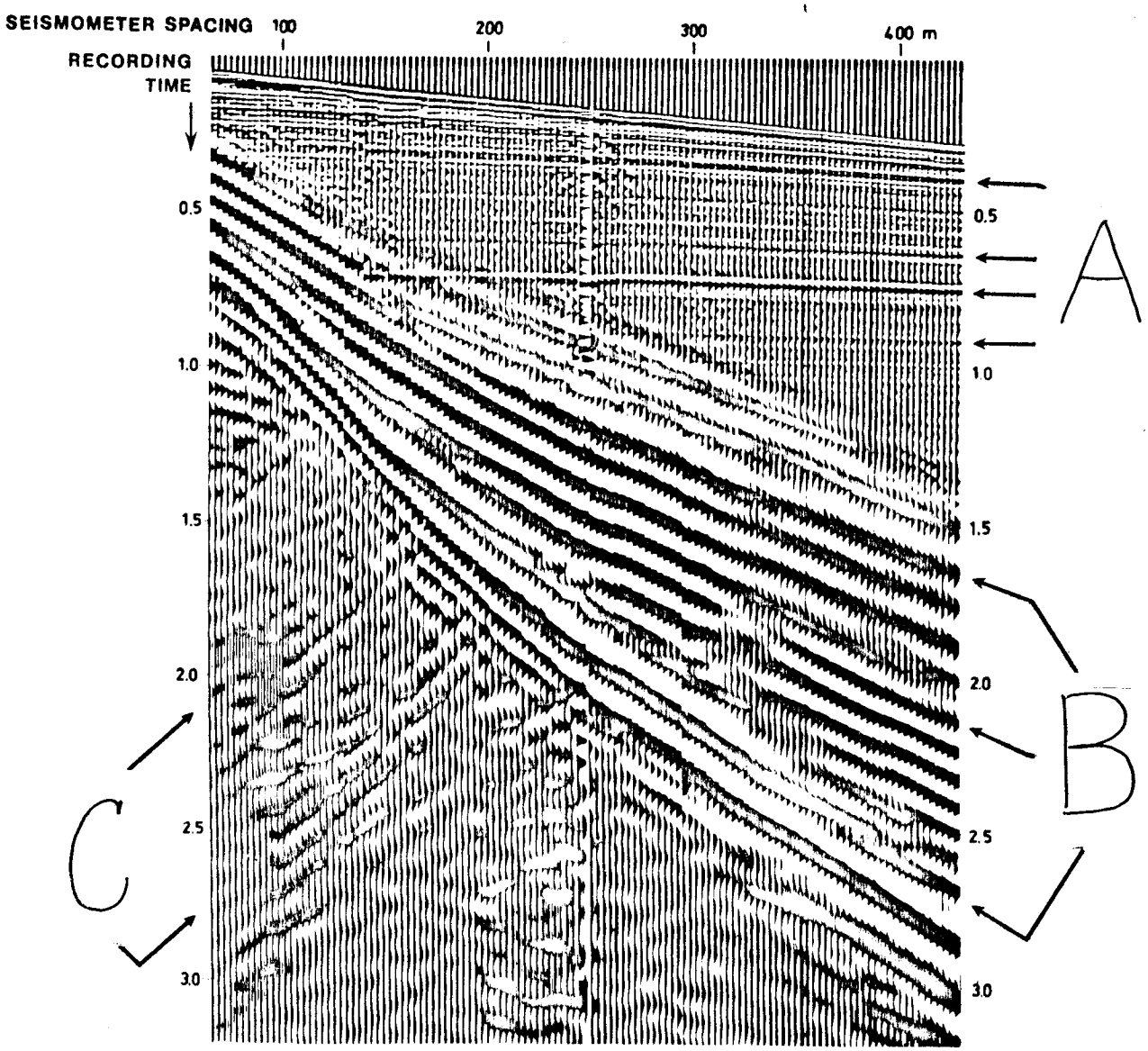

FIG. 7. Field data sample from an Exxon Corporation university student recruitment brochure. Answer the questions in exercise 1.

4.2 DERIVATION OF THE DOUBLE-SQUARE-ROOT EQUATION

The function of the DSR equation is to downward continue an entire seismic survey, not just the geophones but also the shots. One of the square roots in the DSR equation is for the cosine of the angle of waves arriving at the geophone. The other square root is for the angle at the shot. The remainder of the chapter explains migration, stacking, migration before stack, velocity analysis, and corrections for lateral velocity variations in terms of the DSR equation.

Review of the Single-Square-Root Equation

In Chapter 1 we derived the single-square-root equation. The assumption of a *single* plane wave means that its arrival time is given by a single-valued $t(x,z)$. On a plane of constant z, such as the earth's surface, we can measure Snell's parameter p which is

$$\left.\frac{\partial t}{\partial x}\right]_z = \frac{dt}{dx} = \frac{\sin\vartheta}{v} = p \tag{1a}$$

In a borehole we would be constrained to stay at a constant x where the relevant measurement from an *upcoming* wave would be

$$\left.\frac{\partial t}{\partial z}\right]_x = \frac{dt}{dz} = \frac{\cos\vartheta}{v} = -\left[\frac{1}{v^2} - \left[\frac{dt}{dx}\right]^2\right]^{1/2} \tag{1b}$$

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

$$\frac{\partial U}{\partial z} = -\frac{dt}{dz}\,\frac{\partial U}{\partial t} \tag{2a}$$

$$U = f\left[t - \int_0^z \frac{dt}{dz}\,dz\right] \tag{2b}$$

The partial derivatives in equation (2a) are taken to be at constant x, just as is equation (1b). Inserting, we have

$$\frac{\partial U}{\partial z} = \left[\frac{1}{v^2} - \left[\frac{dt}{dx}\right]^2\right]^{1/2} \frac{\partial U}{\partial t} \tag{3}$$

Fourier transforming the wave field over (x,t) we replace $\partial/\partial t$ by $-i\omega$. Likewise, for the traveling wave of the Fourier kernal $\exp(-i\omega t + ik_x x)$, we find that $dt/dx = k_x/\omega$. With this, (3) becomes

$$\frac{\partial U}{\partial z} = -i\omega\left[\frac{1}{v^2} - \left[\frac{k_x}{\omega}\right]^2\right]^{1/2} U \tag{4}$$

With the wisdom of Chapters 2 and 3 we know how to go into the lateral space domain by replacing ik_x by $\partial/\partial x$. Then the equation is valid for superpositions of many local plane waves and for lateral velocity variations $v(x)$. The solutions to (4) agree with those of the scalar wave equation unless v is a function of z, in which case the scalar wave equation has both up- and downgoing solutions, whereas (4) has only upgoing solutions.

The DSR Equation in Shot-Geophone Space

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

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

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

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

We also need a minus sign here because the traveltime in the experiment must decrease as the shots are pushed downward.

Now suppose we simultaneously downward project both the shots and geophones by an identical amount $dz = dz_g = dz_s$. The traveltime change is the sum of (5a) and (5b), namely

$$dt = \frac{dt}{dz_g} dz_g + \frac{dt}{dz_s} dz_s = \left[\frac{dt}{dz_g} + \frac{dt}{dz_s}\right] dz \tag{6a}$$

or

$$\frac{dt}{dz} = -\left\{\left[\frac{1}{v^2} - \left[\frac{dt}{dg}\right]^2\right]^{1/2} + \left[\frac{1}{v^2} - \left[\frac{dt}{ds}\right]^2\right]^{1/2}\right\} \tag{6b}$$

This expression for dt/dz may be substituted into the time-shifting partial differential equation which operates on the upcoming wave field $U(x,z,t)$, namely equation (2a):

$$\frac{\partial U}{\partial z} = -\frac{dt}{dz}\frac{\partial U}{\partial t} \tag{7}$$

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

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

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

With the wisdom of chapters 2 and 3 we know how to go into the lateral space domain by replacing ik_g by $\partial/\partial g$ and ik_s by $\partial/\partial s$. Then we may wish to incorporate the effects of lateral velocity variation $v(x)$. Recall the origin of the two square roots in equation (9). One is the cosine of the arrival angle at the geophones divided by the velocity at the geophones. The other is the cosine of the take-off angle at the shots divided by the velocity at the shots. Making clear our intention to permit the velocity to differ from the shot location to the geophone location we have

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

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

The DSR Equation in Midpoint-Offset Space

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

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

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

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

Traveltime t may be parameterized in (g,s)-space or (y,h)-space. Differential relations for converting are given by the chain rule for derivatives, namely

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

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

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

$$U(s,g) = U'(y,h) \tag{13}$$

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

$$\frac{\partial U}{\partial s} = \frac{\partial y}{\partial s}\frac{\partial U'}{\partial y} + \frac{\partial h}{\partial s}\frac{\partial U'}{\partial h} \tag{14a}$$

$$\frac{\partial U}{\partial g} = \frac{\partial y}{\partial g}\frac{\partial U'}{\partial y} + \frac{\partial h}{\partial g}\frac{\partial U'}{\partial h} \tag{14b}$$

and utilizing (11), we get

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

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

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

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

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

Equation (16) is a Fourier representation of (12). Substituting (16) into (9) we achieve the main purpose of this section, to get the double-square-root migration equation in midpoint-offset coordinates

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

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

$$C = \frac{vk_g}{\omega} \tag{18a}$$

$$S = \frac{vk_s}{\omega} \tag{18b}$$

$$Y = \frac{vk_y}{2\omega} \tag{18c}$$

$$H = \frac{vk_h}{2\omega} \tag{18d}$$

As noted earlier, the definitions of S and C are the sines of the takeoff angle and 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. Similarly, if $H=0$, then Y is bounded by ± 1 [since $Y+(H=0)=C=\pm 1$]. Thus, the quantities S,C,Y,H are all sine-like and refer to angles from vertical in the spaces of shot, geophone, midpoint and offset. With these definitions (17) becomes slightly less cluttered:

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

At this stage I would like to make the claim that the DSR equation forms the basis for all correct migration, stacking, and velocity analysis procedures and that further analysis of it will explain the limitations of standard processing procedures as well as suggest improvements to the standard procedures. Before making the claim and illustrating it by examples, there will be a short diversion to reconfirm that when the velocity is constant, the DSR equation predicts the same traveltimes as Cheop's pyramid.

Stationary Phase to Reconfirm Cheop's Pyramid

In constant velocity media, downward continuation with the DSR equation is expressed in the Fourier domain as

$$U(z) = U(0) \exp\left[-i\left[\frac{\omega^2}{v^2} - k_g^2\right]^{1/2} z\right] \exp\left[-i\left[\frac{\omega^2}{v^2} - k_s^2\right]^{1/2} z\right] \tag{20}$$

In the domain of (ω,k_g,k_s) the downward-continuation operator is the product of two exponential functions. Thus, in the domain of (t,g,s) it will be a convolution of two functions. What are those functions? Each function is the familiar, conic-section, Huygens hyperbola. The first exponential operator is a function of ω and k_g, but it is a constant function of k_s. This means it is a hyperbola in (g,t) space times a delta function at $s=0$. If you can imagine two hyperbolic coat hangers, one hanging from the other, then you may have the correct picture. The sum of the two times represents the travel path from the reflecting point to the shot and from the reflection point to the receiver.

As an exercise in the use of the stationary phase method, we may now see how it can be used to produce the migration ellipse. To inverse Fourier transform the operator of (20) or its representation in midpoint-offset space, we multiply by the Fourier kernal and then integrate. The total phase angle of the exponential for the midpoint-offset case is

$$\Phi = \left\{\left[\frac{\omega^2}{v^2} - \frac{1}{4}\left[k_y + k_h\right]^2\right]^{1/2} + \left[\frac{\omega^2}{v^2} - \frac{1}{4}(k_y - k_h)^2\right]^{1/2}\right\} z + k_h h + k_y y - \omega t \tag{21}$$

At large values of $(z,h,y,$ and $t)$ the integrand is a rapidly oscillating function of (ω,k_h,k_y). Only at the flat spots on the phase is the integral expected to amount to much. These flat spots may be found by the three conditions

$$\frac{\partial\Phi}{\partial\omega} = 0\,; \quad \frac{\partial\Phi}{\partial k_h} = 0\,; \quad \frac{\partial\Phi}{\partial k_y} = 0 \tag{22}$$

Performing the differentiations you find expressions containing not just (z,h, and y) but also (ω,k_h,k_y). Persistent algebraic efforts (made easier if you have a copy of Rob Clayton's personal notes) enables you to use (21) and two equations from (22) to eliminate all the frequencies. Having persevered you are left with Cheop's equation.

4.3 THE MEANING OF THE DOUBLE-SQUARE-ROOT EQUATION

The double-square-root (DSR) equation controls the bulk of the non-statistical 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 lower dimensional space. In this section we will neglect lateral velocity variation (things are bad enough already!) and start from

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

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

Zero-Offset Migration (H=0)

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

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

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

Zero-Dip Stacking (Y=0)

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

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

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

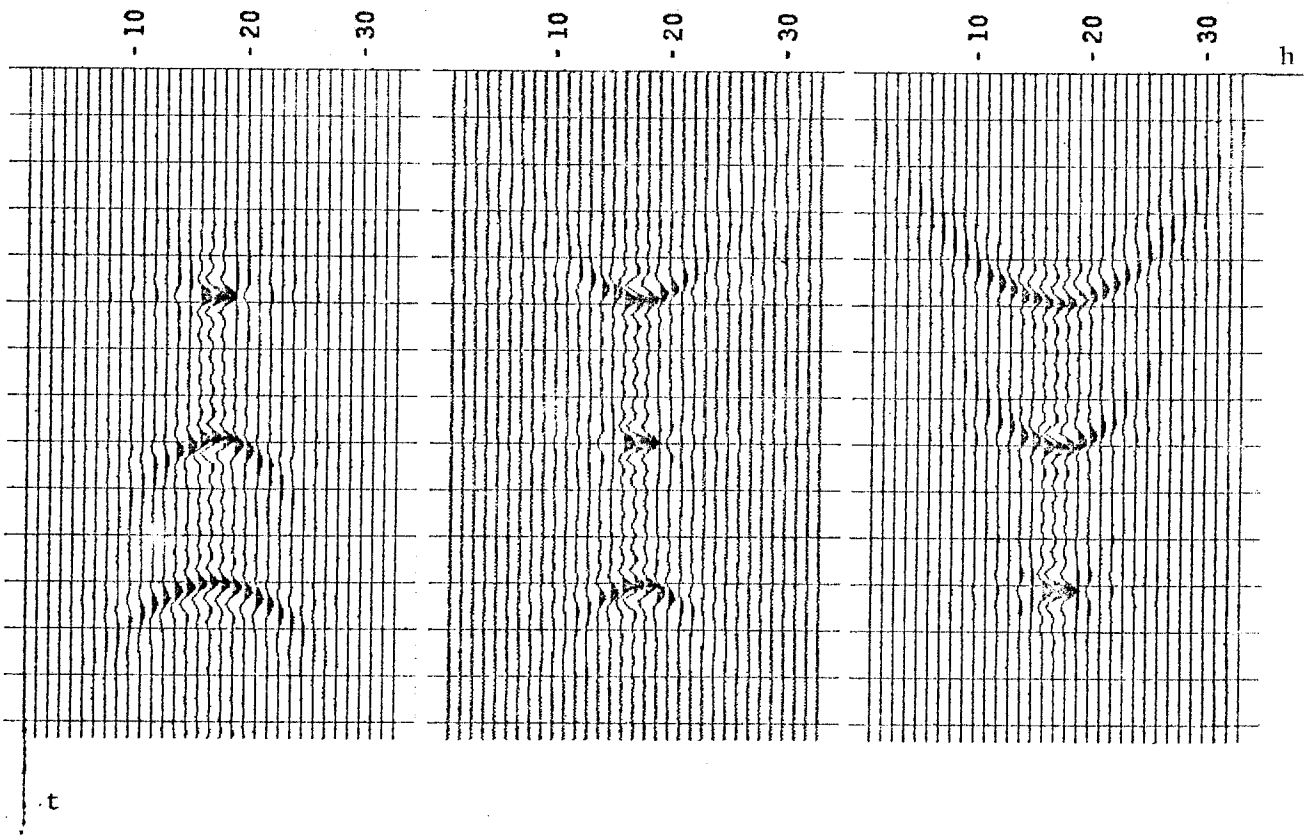

FIG. 1. (Gonzalez) With an earth model of three layers, the common-midpoint gathers are three hyperboloids. Successive frames show downward continuation to successive depths where best focus occurs.

Because of the symmetry of surface data, and all subsequent subsurface data, this best focus occurs at zero offset. For a zero-offset, source-receiver pair, the largest signal strength must be seen in the limit of zero traveltime, when the pair is just above the reflector. Extracting the zero-offset value at $t=0$ and abandoning the other offsets is a way of eliminating noise. Roughly, it amounts to the same thing as the conventional procedure of summation along a hyperbolic trajectory on the original data. More specifically, it may be the zero lag in (h,t)-space of the crosscorrelation between the field data and the Huygens secondary source hyperbola. Naturally the correlation can be expected to be best when the velocity used for downward continuation comes closest to the velocity of the earth. Later we will see how offset space is used to *determine* velocity.

Various Meanings of "H = 0"

Recall the various forms of the stepout operator:

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

For any single ray, reciprocity implies that t is a symmetrical function of offset; thus dt/dh vanishes at $h=0$. In that sense it seems appropriate to apply equation (2) to zero-offset sections. To be more precise, however, 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 more appropriate and it is slightly different. To set $\omega = 0$ would be to select a zero frequency component, a simple integral of a seismic trace. Setting $k_h = 0$ is selecting 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 which is very flat, namely the hyperbola of infinite velocity. Such an integration will have its major contribution at the top of the data hyperboloid, where the data events come tangent to the horizontal line of integration. (For some historical reason, this kind of 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, called a Fresnel zone, is the major factor contributing to the integral. See figure 2. The main differences between a zero-offset section and a vertical stack are the amplitude and a small phase shift. In practical cases they are unlikely to migrate in a significantly different way. It might be nice if we could find an equation to downward continue data that are stacked at velocities other than infinite velocity.

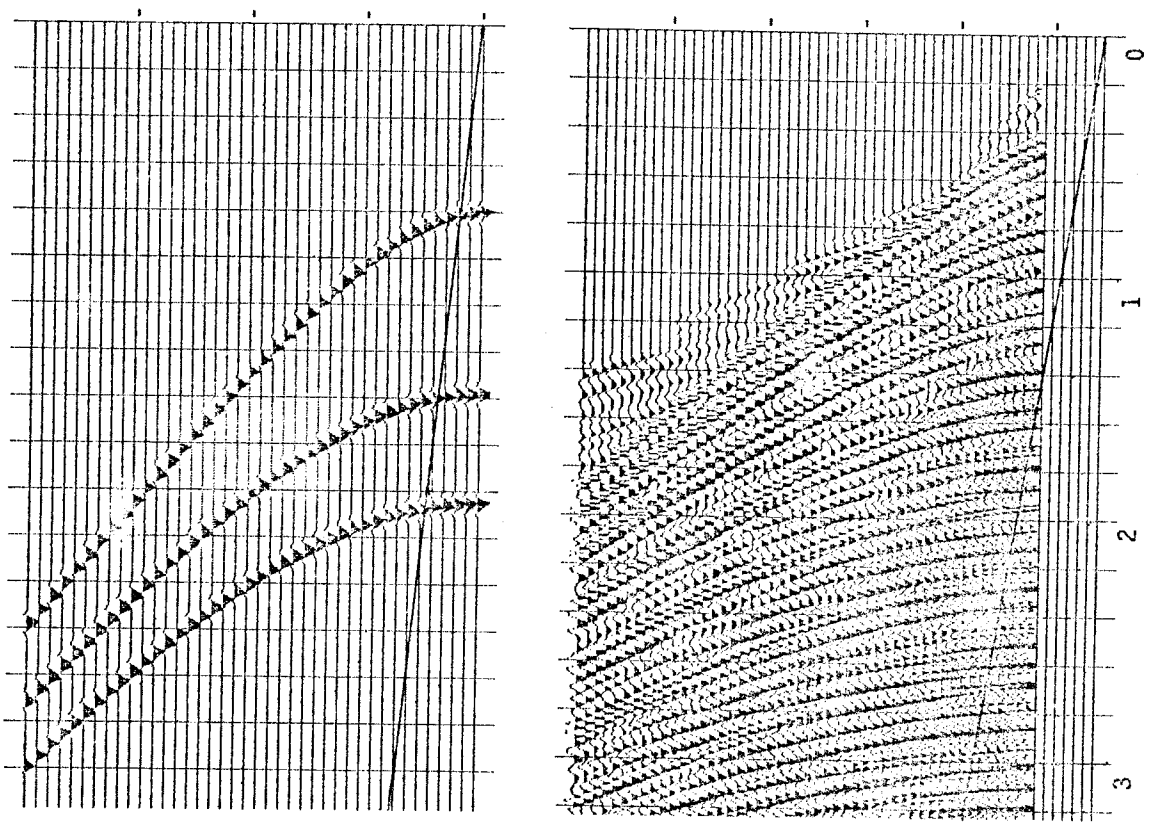

FIG. 2. (Gonzalez) The Fresnel zone for a vertical stack is defined to be the horizontal extent on the offset h-axis within which the time shift of seismograms is about a half-wavelength. To be mathematically precise, it is necessary to specify a frequency. For practical purposes it is usually sufficient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

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 former viewpoint is often a more powerful one. To be specific, without being cluttered, express equation (1) in 15-degree, retarded, space-domain form. We get

$$\left[\frac{\partial}{\partial z} + \frac{v}{-i\omega 8}\left[\frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial h^2}\right]\right] U = 0 \qquad (4)$$

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 and lower limit. We get

$$\left[\frac{\partial}{\partial z} + \frac{v}{-i\omega 8}\frac{\partial^2}{\partial y^2}\right]\left[\int U\, dh\right] + \left.\frac{\partial U}{\partial h}\right|_{h=-\infty}^{h=+\infty} = 0 \qquad (5a)$$

The wave should vanish at infinite offset and so should its horizontal offset derivative. So we should be able to neglect the last term in (4) and assert that setting $H=0$ means just that.

$$(Single\ SQRT\ operator)\ (vertical\ stack) = 0 \qquad (5b)$$

A problem in developing (5b) is that, twice, we have assumed that velocity was independent of offset, first when we omitted the thin lens term from (4), and second when we commuted the offset-integration operator with the operation of multiplication of velocity. If the velocity depends on the horizontal x-axis, then it certainly depends on both midpoint and offset. One final conclusion is this: If velocity changes slowly across a Fresnel zone, then setting $H = 0$ provides a valid equation for downward continuation of vertically stacked data.

Clayton's Cosine Corrections

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

The dip angle of the reflector is α and the offset is expressed as the offset angle β. Clayton showed, and we will verify that

$$Y = \sin\alpha \cos\beta \qquad (6a)$$

$$H = \sin\beta \cos\alpha \qquad (6b)$$

Thus it will be seen that for small positive or negative angles it is correct to associate the sine of the earth dip angle with Y and the sine of the offset angle with H. At moderate angles we see a cosine correction in each case. At angles exceeding 45 degrees (which are rarely considered) the sensitivities reverse. The reader should be wary of informal discussions which simply associate Y with dip and H with velocity. Perhaps at steep dips the usual procedure of using H to determine velocity should be changed somehow to use Y. Let us return to the derivation of (6).

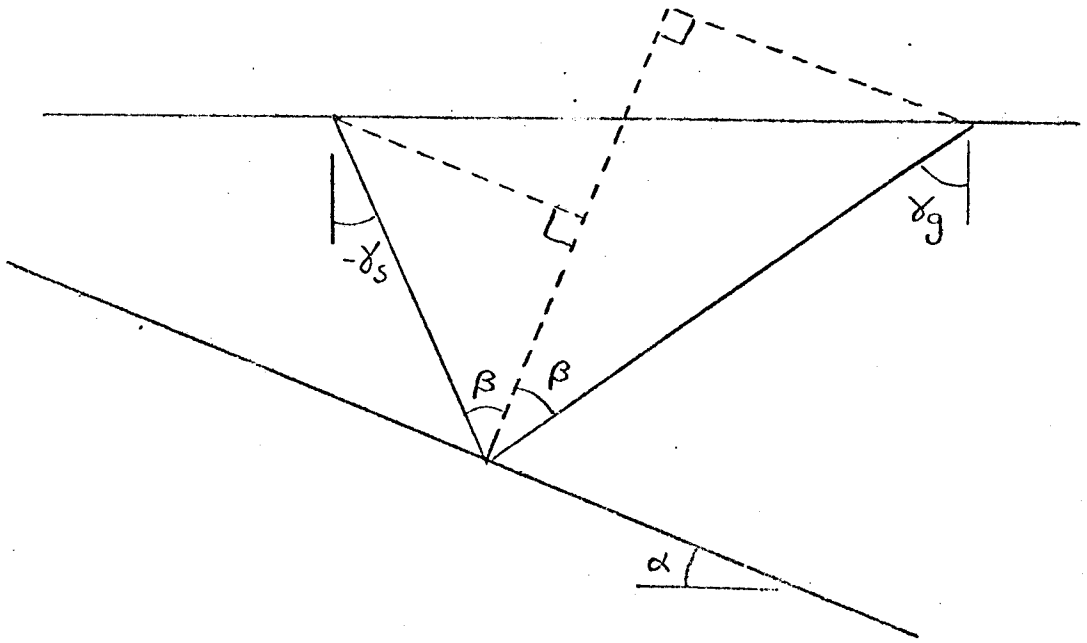

FIG. 3. (Clayton) Geometry of dipping bed. Note that the line bisecting the angle 2β does not pass through the midpoint between g and s.

The source takeoff angle is γ_s and the incident receiver angle is γ_g. First we start by relating γ_s and γ_g to α and β. Adding up the angles of the small constructed triangle we have

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

or

$$\gamma_s = \beta - \alpha$$

Adding up the angles around the larger triangle we have

$$\gamma_g = \beta + \alpha$$

By Snell's law we know that $S = \sin(\gamma_s)$. From the previous section [equations (16) and (18)] we see that $S = Y - H$. Hence,

$$Y - H = S = \sin\gamma_s = \sin(\beta - \alpha)$$

and similarly,

$$Y + H = G = \sin\gamma_g = \sin(\beta + \alpha)$$

Solving for Y and H in this pair of equations we have

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

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

Adding and subtracting, we can use the angle sum formula from trigonometry to get Clayton's cosine corrections.

Snell-Wave Stacks and CMP Slant Stacks

Now that we have considered the meaning of setting to zero the dip angle Y and the offset angle H we should perhaps consider the more obvious possibility of setting the takeoff angle S to zero. This also reduces the double-square-root equation to a single square root. The meaning of $S=0$ is that $k_s = 0$ or that the data should undergo a summation (without time shifting) over shot s. Such a summation simulates a downgoing plane wave. Conceptually this is a nice idea. The imaging principle would be to look at the upcoming wave at the arrival time for the downgoing wave.

In practice it would be a problem that the Fresnel zone is not very wide and the energy outside the zone is ignored. This energy at wider offsets could be included in the analysis by generating a Snell wave with some non-zero Snell parameter p. Such a wave is generated by time shifting the shotpoints before superposition. In other words, this is linear moveout over shots, at constant geophone, followed by stack. This is depicted in figure 4.

Snell waves could be constructed for various p values. Each can be migrated and imaged, and the images stacked over p. The geometry of this is more confusing than you might expect.

Another problem with Snell-wave simulation is that the data are usually known at rather 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 dataset are frustrated by the fact that on a common-geophone gather, the earliest arrival need not be at zero offset, and for steeply dipping beds the earliest arrival is often off the end of the cable.

All this provides an ecological nitch for the common-midpoint slant stack, namely $H=pv$. It has the advantage that the hyperbolas go through zero offset with zero slope. So the data are more amenable to interpolation and extrapolation.

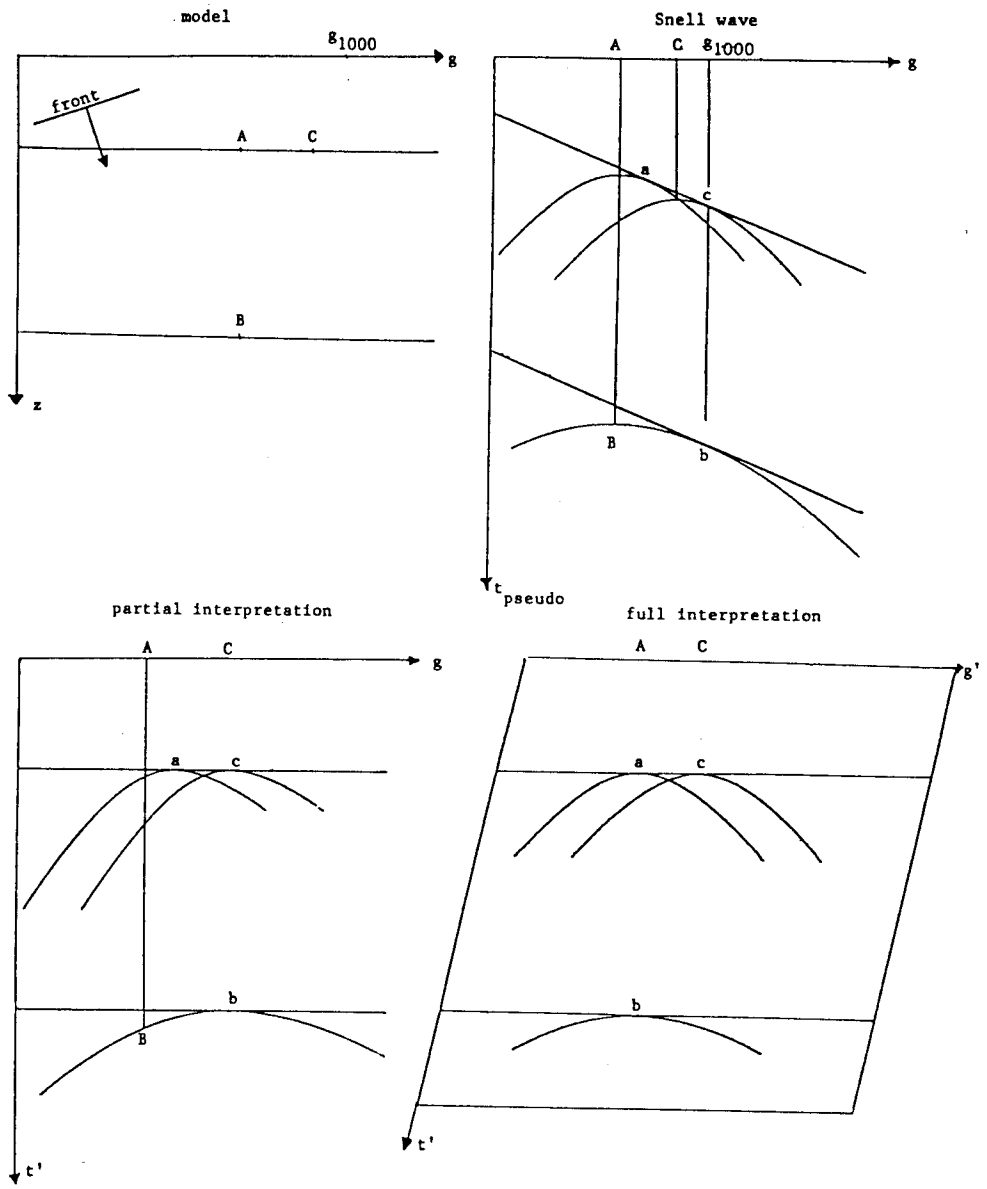

FIG. 4. (Gonzalez) Top left is three point scatterers on two reflectors. Top right is the expected Snell wave. Bottom left is the Snell wave after linear moveout. Bottom right is after tansform to full interpretation coordinate. At last a, b, and c are located where A, B, and C began.

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. The harsh reality is that the velocity is not 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 kind of 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. Considerably less ambitious and more practical approaches to the problem are found later in the chapter.

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

4.4 STACKING AND VELOCITY ANALYSIS

Hyperbolic stacking over offset may be the most important process in the prospecting industry. It is far more important than migration of stacks. The reason why stacking is so much more important than migration is that it reduces the data base from a volume in (s,g,t)-space to a plane in (y,t)-space. Only after such a reduction can the data be plotted and perceived by human beings. (A useful byproduct of stacking is velocity determination.) Migration merely maps 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 you can hardly look at your data until you stack them.

Historically, stacking has been done by ray methods, and it is still being done almost exclusively by ray methods. Migration, on the other hand, is more commonly done by wave-equation methods, that is to say, by Fourier or finite-difference methods. The advantages of wave-equation methods in migration have been many. Don't these advantages apply equally to stacking? Maybe, or maybe not.

This section describes a number of ingenious wave-equation stacking and velocity-determination methods. Perhaps they have not been satisfactorily tested, or perhaps they are fundamentally flawed. The reader can guess, and time will tell. One possible basic flaw 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. Another possibility for a basic flaw is that perhaps the issue of stacking to remove redundancy is more appropriately a statistical problem than a physical problem. This must certainly be the case in some land areas where the complexity of the near surface geology far exceeds the redundancy in our data. Even so, the data manipulation procedures in this chapter should be helpful.

Conventional Methods

To do a conventional velocity analysis and stack, you basically need some skill at interpolation and you need a table of computed traveltimes as a function of offset and depth for some velocity model. Wide offset traces have their time axes stretched to make the arrival of an event at wide offset come out to be at the same time as the event at zero offset. After possible intermediate processing, such as balancing amplitudes and spectra, the data are averaged over offset. Presumably, the more closely the earth velocity matches the velocity in the traveltime table, the better (bigger) will be the result.

The first step in generating the traveltime tables is to change the depth-variable z to a vertical traveltime-variable τ. Then we create a table $t(h,\tau)$. To find the output data for location (h,τ) you look at the input data at location (h,t). The most straight forward and reliable way to get this table seems to be to march down in steps of z, actually τ, and trace rays. That is, for various fixed values of Snell's parameter p, you compute $t(p,\tau)$ and $h(p,\tau)$ from $v(\tau)$ by integrating

$$\frac{dt}{d\tau} = \frac{dz}{d\tau}\frac{dt}{dz} = v\,\frac{1}{v\cos\vartheta} = \left[1 - p^2v^2\right]^{-1/2} \tag{1}$$

$$\frac{dh}{d\tau} = \frac{dz}{d\tau}\frac{dh}{dz} = v\tan\vartheta = \frac{pv^2}{(1-p^2v^2)^{1/2}} \tag{2}$$

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

The subject of dip corrections is taken up in a later section; however, it can be approximately stated that the effect of dip is to increase the stacking velocity by the cosine of the dip.

The (z,t) Plane Method

In the 15-degree continuation equation $U_{zt} = -1/2vU_{hh}$ it may be noted that a scaling in depth z is indistinguishable from a scaling in velocity. This indicates the practical fact that downward continuation with the wrong velocity is very similar to downward continuation to the wrong depth. Stephen M. Doherty was the first to utilize this idea for a velocity estimation scheme. See figure 1.

The idea is to downward continue with a preliminary velocity model and display the zero offset trace, a function of t', at all traveltime depths τ. If the maximum

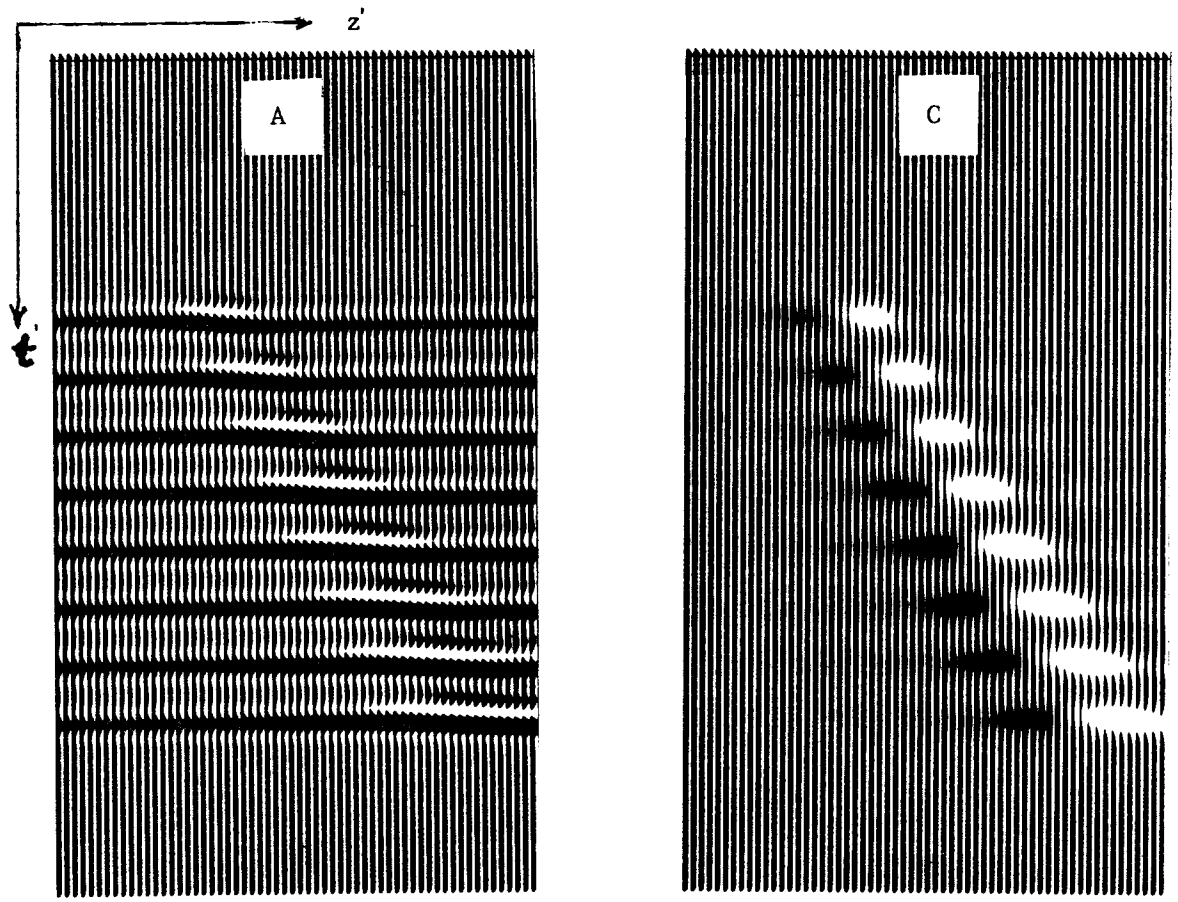

FIG. 1. (Doherty) Two displays of the (z,t)-plane at zero offset. The earth model is 8 uniformly spaced reflectors under a water layer (a family of hyperboloids in (h,t) at $z=0$). The left display is the zero offset trace. The amplitude maximum at the focus is not visually striking, but the phase shift 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.

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.

Reflected Refractions on Sections

It is fairly common for an interpreter looking at a stacked section to identify a reflected refraction. It is just a hyperbolic asymptote seen in (y,t)-space. This event provides a quick, easy, and accurate velocity estimate, namely $v=2dy/dt$. From a processing point of view, such a velocity measurement is rather unexpected, because automatic processing extracts all velocity information in offset space, a space which many interpreters prefer to leave inside the computer. Of course rather special geological circumstances must be present: a point scatterer strong enough to have its hyperbolic asymptote visible. Furthermore, the point scatterer has to be strong enough to get through the typical suppression effects of shot and geophone patterns and CDP stacking. It is rather remarkable that the most highly suppressed events, water velocity and ground roll, are just those whose velocities are most commonly apparent on stacked sections. We saw some very strong reflected refraction energy on the common-shot profile shown in the section on Cheop's Pyramid.

All processing seems to ignore or discriminate against the reflected refraction, yet we see it and use it. There must be an explanation and perhaps there is a latent opportunity. From a theoretical point of view we already noted that at wide angles the velocity and dip sensitivity of midpoint and offset interchange their normal roles. At late times there is another important theoretical consideration. The aperture of a cable length can be much less than the width of a migration hyperbola. So, although we can easily get to an asymptote in midpoint space, we can see very little time shift at the end of the cable in offset space.

What kind of processing could take advantage of lateral reflectivity and enhance, instead of suppress, our ability to determine velocity in this way? Start by stacking at a high velocity. Then we can utilize the basic idea that at any depth z, the power spectrum of the data $U(\omega,k_y)$ should have a cutoff at the evanescent stepout $p(z) = k_y/\omega = 1/v(z)$. We could plot the power spectrum U^*U, or better yet the dip spectrum, as a function of depth. Perhaps is 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 $t=0$. Such data should be zeroed before computing the dip spectrum. The procedure is depicted in figure 2.

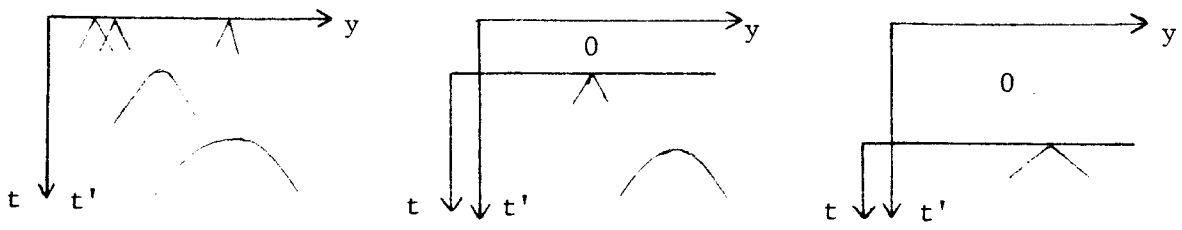

FIG. 2. The dip-spectrum method of velocity determination. To find the velocity at any depth you seek the steepest dip on the section at that depth. On the left, at the earth's surface, you see the surface ground roll. In frames B and C the slowest events are the asymptotes of successively faster hyperbolas.

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

Clayton and McMechan's Method for Refractions on Gathers

The same process for getting velocity from reflected refractions on sections could be used on a common-midpoint gather. But on a gather there is the further interesting feature that downward continuation focuses all the energy upon zero offset. A focus is not a featureless point. Taking original data to consist of a refraction only, with no reflection, then the focus is a concentrated patch of energy oriented with the same stepout dt/dh as the original unfocused refraction. Summing through the focus at all possible orientations (slant stack), we transform the data, say $U(h,\tau)$, to dip space, say $U(p,\tau)$. The velocity of the earth at traveltime depth τ is found by seeking maxima on $p(z)=1/v(z)$.

Clayton and McMechan invented and developed the velocity determination procedure described here. Actually, they did the downward continuation and the slant stack in opposite orders, but I don't believe it makes any difference. Figures 3 and 4 show one of their examples.

Because it is limited to refractions, the method of Clayton and McMechan may seem to be rather remote from routine industrial velocity analysis. But it has an

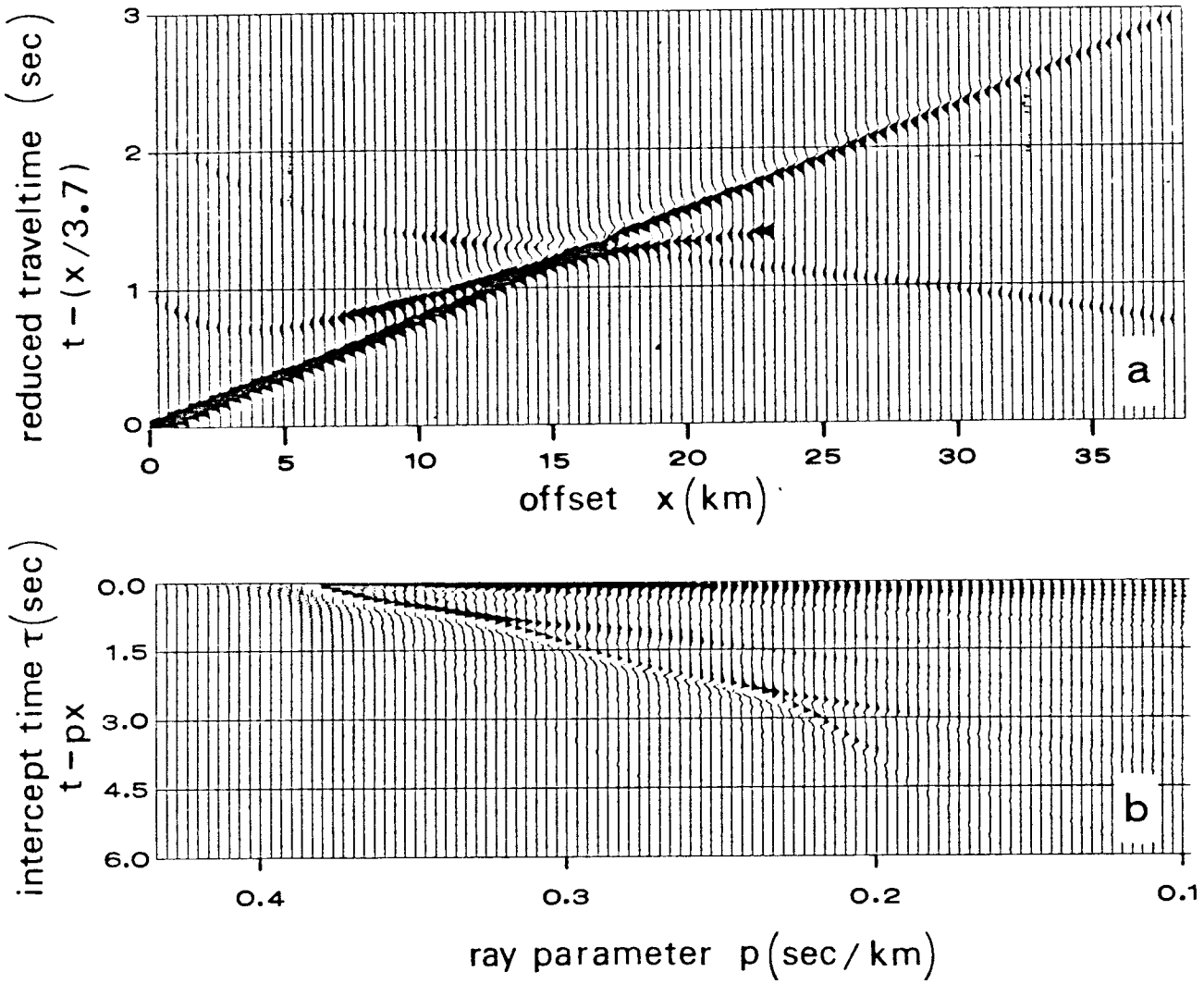

FIG. 3. (Clayton & McMechan) The first transformation. The upper (a) half of this figure contains a synthetic refraction profile plotted in reduced time format. The reduction velocity is 3.7 km/s. These data are transformed by slant stacking into the plane-wave decomposition shown in the lower (b) half of the figure. This transformation is the first half of the process of inversion of the data wave field. The result of migrating this slant-stacked wave field (b) is shown in figure 4.

important feature. It is a completely linear and invertible function of the data. The inverse operation, which can be used for making synthetic data, is just an inverse slant stack followed by upward continuation. Besides the ability to make synthetic data, important processing advantages stem, as we will see, from the ability to partition data by velocity, then return to the space of the original data.

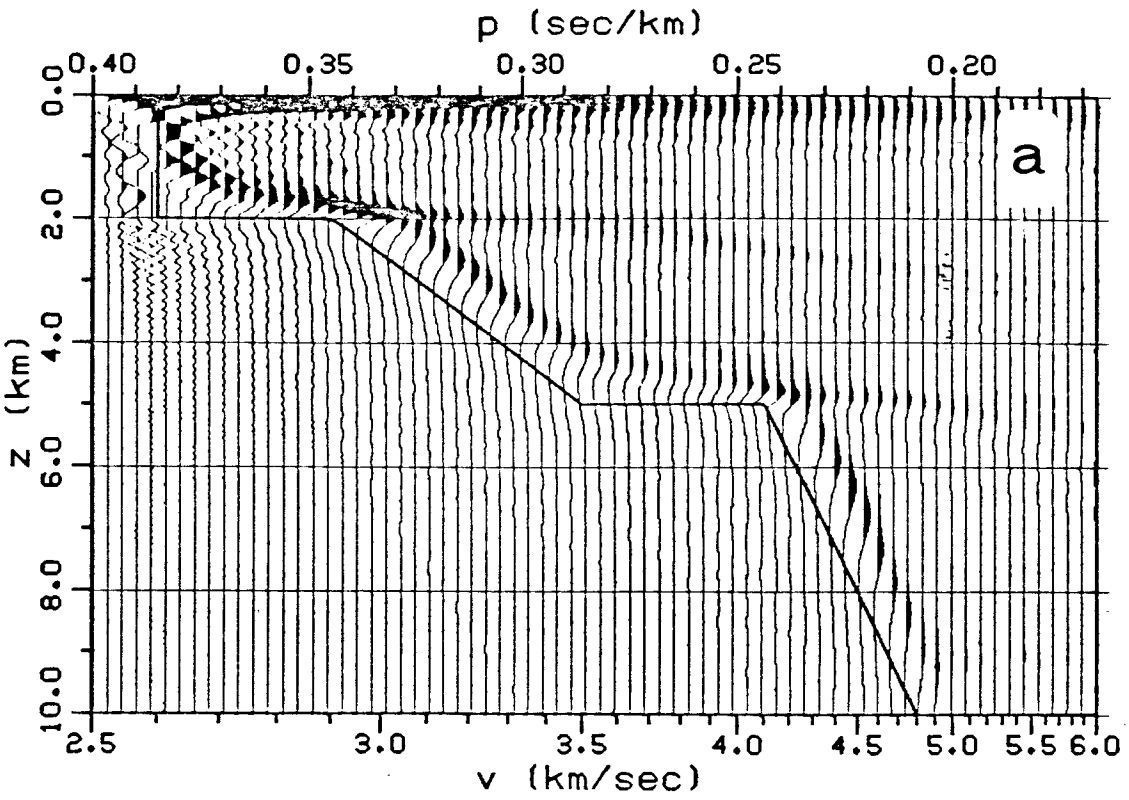

FIG. 4. (Clayton & McMechan) This figure contains the result of migration of the slant-stacked wave field in figure 3b with the correct velocity-depth function (the solid line).

Splitting a Gather into High- and Low-Velocity Components

The pie-slice process shares a nice feature in common with the Clayton and McMechan process. Both enable a dataset to be split into dip components greater than some velocity and dip components less than that velocity. Then the low-velocity part can be abandoned as noise. A difficulty with both methods is that they are troubled by the reflected energy. Near the top of any slow-velocity-noise hyperbola there is little stepout; hence this reflected energy is not distinguishable on the basis of stepout alone from high-velocity signals.

Let us define a process which can partition a CMP gather, *both reflections and refractions*, into a part with RMS velocity greater than that of some given model $\overline{v}(z)$ and another part with velocity less than $\overline{v}(z)$.

After such a partitioning is made, we could abandon the low-velocity noise. We could also find the earth velocity through iteration, by making the usual assumption that the velocity spectrum has a peak at earth velocity. As we will later see, 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 $\overline{v}(z)$. As a result, the components of the data with velocity less than $\overline{v}(z)$ will over-migrate through zero offset to negative offsets. The components of data with velocity greater than $\overline{v}(z)$ will under-migrate. 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. Then, if you wish, the process can 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 some kind of 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.

Unfortunately, this partitioning method 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 development of other processes.

Lateral Interpolation and Extrapolation of a CMP Gather

Figure 5 shows why cable truncations are a problem for conventional, ray-trace, stacking methods as well as for wave-equation methods. Spatial aliasing of data on the offset axis may not be quite the problem for ray-trace methods that it is for wave-equation methods. The reason is that the spatial-aliasing problem is really one of bandwidth, rather than maximum frequency, and the ray-trace methods seem somewhat better prepared to handle it. Here we will attack both problems, namely the task of using *a priori* knowledge of limits of reasonable velocities in order to fill in the *tag* data (tag = *t*runcated, *a*liased, and *g*apped) to achieve both better stacks and better velocity spectra within the range of reasonable values. These problems are more seriously being attacked by current SEP research projects.

We plan iterative use of the partitioning method described previously. Figure 6 delimits various regions for discussion. We define region A to be a place where no velocity is so low as to spatially alias the raw data. Region B has important high-velocity events as well as some aliased low velocity events. Region C has overmigrated low-

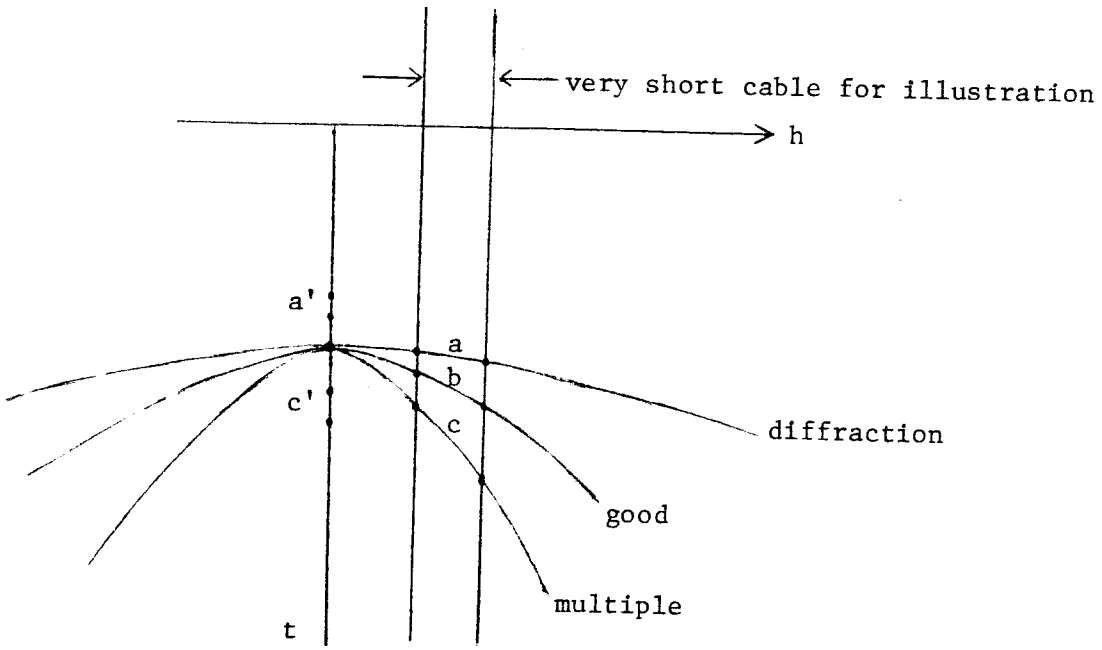

FIG. 5. Normal moveout at the earth velocity brings the cable truncations on good events to a good place, causing no problems. But the cable truncations of diffractions and multiples move to a' and c', where they will be objectionable if they are of sufficient size.

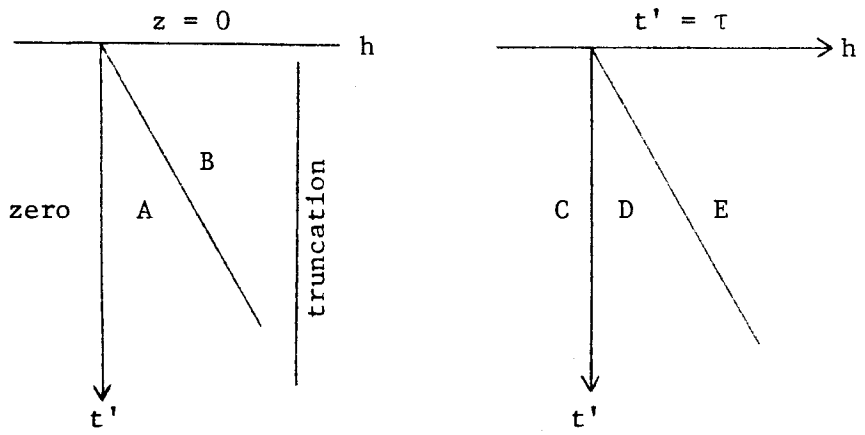

FIG. 6. Regions of space on one side of a surface gather (left) and a focused gather (right).

velocity energy. The ultimate stack lies on the border between C and D. Region D has model velocities, and higher velocities due to dip. Region E contains various kinds of junk, which we can suppress to help fill in the *tag* data. Notice that region E contains truncations from the far end of the cable, and possibly truncations from the near end. On a more subtle level, recall the effect of wave-equation migration on aliased data. Aliased data may have a stepout of 2π per trace so it can appear to have little or no stepout. Spatial aliasing of low velocities makes them high velocities. Such information, which exists in B will not move, so it will be found in E.

The iterative procedure is as follows:

1. Downward continue.
2. Zap region E.
3. Upward continue.
4. Restore data where known.

It seems clear that the iterative procedure will interpolate traces, enabling wave-equation methods to work as well as, or better than, ray-trace methods. It also seems clear that data will be extrapolated off the end of the cable. It is not totally clear that the extrapolated data will improve the stack along the CD boundary. Perhaps it will, and perhaps it will not.

To really do a good job of extending a dataset may require a parsimonious model concept and a better data-partitioning procedure, similar to that described in the next subsection.

The Linear Moveout Method

The linear-moveout method began as a graphical method of finding the velocity of reflected events. It was intended for use by interpreters on raw field data (SEP-11, p. 41-43; SEP-14, p. 13-15; SEP-15, p. 61-66.) Then a focusing procedure was added (SEP-15, p. 81-86) and tested by Gonzales (SEP-16, p. 181-204; SEP 20, p. 49-56). The present status of the method is unclear, but it does have the attributes of linearity, invertibility, and parsimony. "Parsimony" in this case means the following: start with ideal data, which has a number of events with various velocities and traveltimes. The data may contain multiples, so that more than one velocity could occur at once. On transformation into (h,τ)-space the events should be found in various places, according to their velocity, and all should be moderately well focused. Then the principle of minimum entropy can be used to extend and fill the dataset.

Exact Graphical Method for Interval Velocity Measurement [1]

Consider a point source. The wavefront after a time t is a circle of radius vt and is given by

$$v^2t^2 = x^2 + z^2 \tag{3,4}$$

Letting f denote the lateral source- receiver offset and z_s denote the depth to an image source under a horizontal plane layer we have

$$v^2t^2 = f^2 + (z - z_s)^2 \tag{5}$$

We make our measurements at the earth's surface where $z=0$. Differentiating (5) with respect to t we obtain

$$v^2 2t = 2f \frac{df}{dt}$$

$$v^2 = \frac{f}{t}\frac{df}{dt} = \frac{f}{pt} \tag{6}$$

Figure 7 shows that the three parameters required by (6) to compute the material velocity are readily measured on a common midpoint gather.

Of course, we can measure some kind of velocity by means of Equation (6) even if the earth does not have the assumed constant velocity. The question then becomes, what does the measurement mean? In the case of a stratified medium $v(z)$ we can quickly establish the answer to be the familiar RMS, or root-mean-square velocity. To do so, first note that the bit of energy arriving at the point of tangency has throughout its entire trip into the earth been propogating with a constant Snell's parameter p. The best way to specify velocity in a stratified earth is to give it as some function $v(z)$. Another way is to pick a Snell's parameter p and start descending into the earth on a ray with this p. As the ray goes into the earth from the surface $z=0$ at $t=0$, the ray would be moving with a speed of, say, $v'(p,t)$. It is an elementary exercise to compute $v'(p,t)$ from $v(z)$ and vice versa. So, when convenient, we may refer to the velocity as some function $v'(p,t)$. The horizontal distance f which a ray will travel in time t is given by the time integral of the horizontal component of velocity, namely

$$f = \int_0^t v'(p,t) \sin\vartheta \, dt \tag{7}$$

[1] SEP 20, p 61.

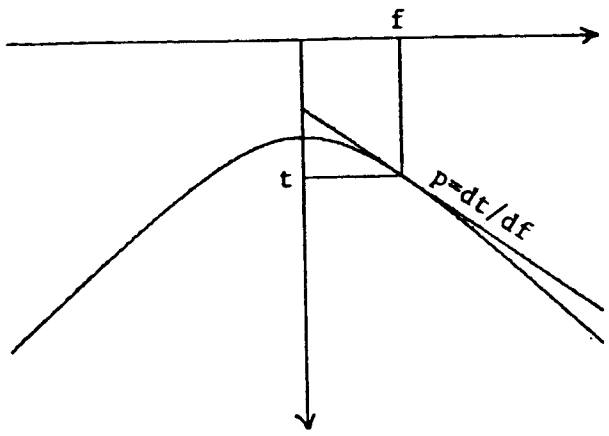

FIG. 7. (Gonzalez) A straight line, drawn tangent to hyperbolic observations. The slope p of the line is arbitrary and it may be chosen so that the tangency occurs at a place of good signal- to-noise ratio.

Replacing $\sin\vartheta$ by pv and taking the constant p out of the integral yields

$$f = p \int_0^t v^2 dt \tag{8}$$

Inserting (8) into (6) we get

$$v^2_{measured} = \frac{f}{pt} = \frac{1}{t} \int_0^t v^2 dt \tag{9}$$

which justifies the assertion that

$$v_{measured} = v_{root\text{-}mean\text{-}square} = v_{RMS} \tag{10}$$

Equation (9) is exact. It does not involve a "small offset" assumption or a "straight ray" assumption.

Next let us consider the so-called *interval* velocity. Figure 8 shows hyperboloidal arrivals from two flat layers where a straight line of slope p has been constructed to have the same slope p. Then the tangencies are measured to have locations (f_1,t_1) and (f_2,t_2). From (8) and (2), using the subscript i to denote the i-th tangency (f_i,t_i), we have

$$f_i \frac{df}{dt} = \int_0^{t_i} v^2 dt \tag{11}$$

Assume that the velocity between successive events is a constant $v_{interval}$ and subtract (11) with $i + 1$ from (11) with i to get

$$(f_{i+1} - f_i)\frac{df}{dt} = (t_{i+1} - t_i)v^2_{interval} \tag{12}$$

Solving for the interval velocity,

$$v^2_{interval} = \frac{f_{i+1} - f_i}{t_{i+1} - t_i}\frac{df}{dt} \tag{13}$$

So the velocity of the material between the i-th and the $i+1$-st reflectors can be measured directly by the square root of the product of the two slopes in (13), which are the dashed and solid straight lines in figure 8. The advantage of manually placing straight lines on the data, over automated analysis, is that you can graphically visualize the sensitivity of the measurement to noise, and you can select the best offsets on the data at which to make the measurement. When doing this routinely one quickly discovers that the major part of the effort is in accurately constructing two lines which are tangent to the events. When this happens, it is convenient to replot the data with linear moveout $t' = t - pf$. After replotting, the sloped lines have become horizontal so that any of the many timing lines can be used. Locating tangencies is now a question of finding the tops of convex events. This is depicted in figure 9. In terms of the time t', equation (13) becomes

$$v^2_{interval} = \frac{1}{\frac{\Delta t}{\Delta f}}\frac{1}{p} = \frac{1}{\frac{\Delta t'}{\Delta f} + p}\frac{1}{p} \tag{14}$$

Finally, the advantages of the manual technique of interval velocity determination presented here, compared with the automated hyperbola scan technique of current industry practice, are:

1) We have made no analytical approximations which deteriorate with angle.
2) We select that portion of the data (by selecting the p value) where the data quality is best for the task at hand.
3) Although it is not shown here, it turns out that migration techniques are available to pre-process the data to remove dip effects.

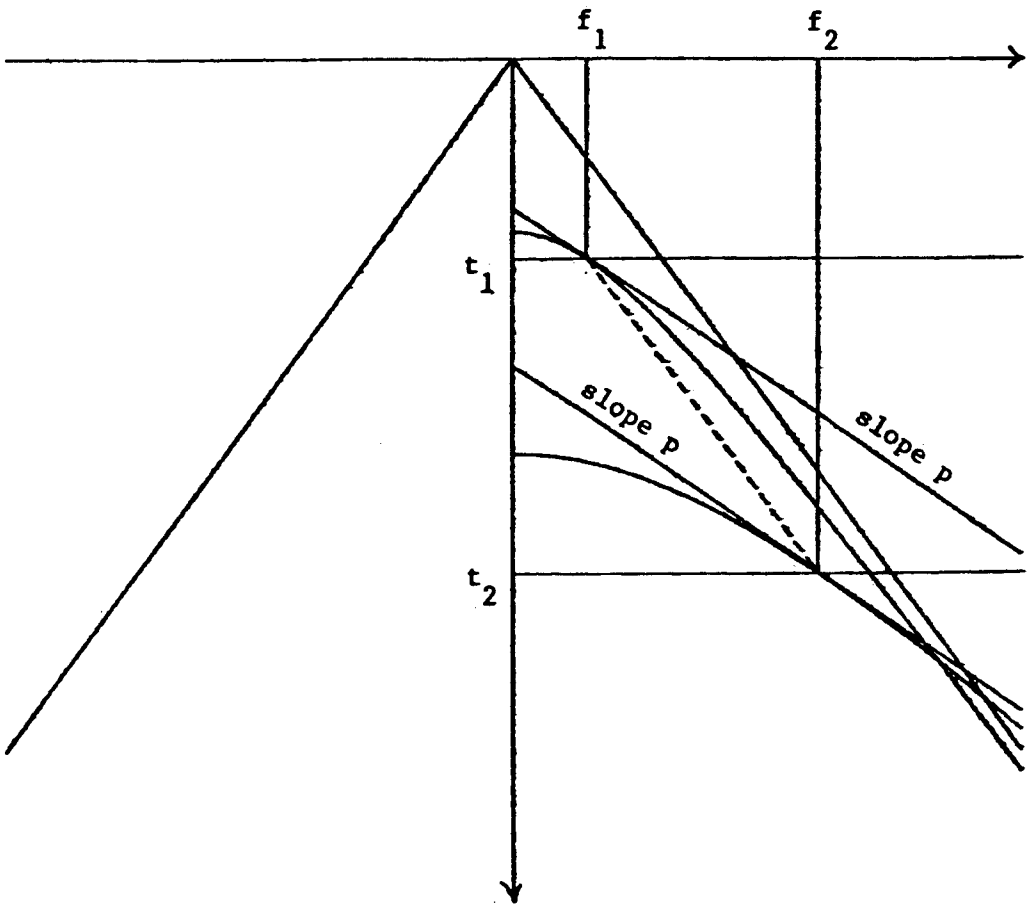

FIG. 8. (Gonzalez) Construction of two parallel lines on a common midpoint gather tangent to reflections from two plane layers.

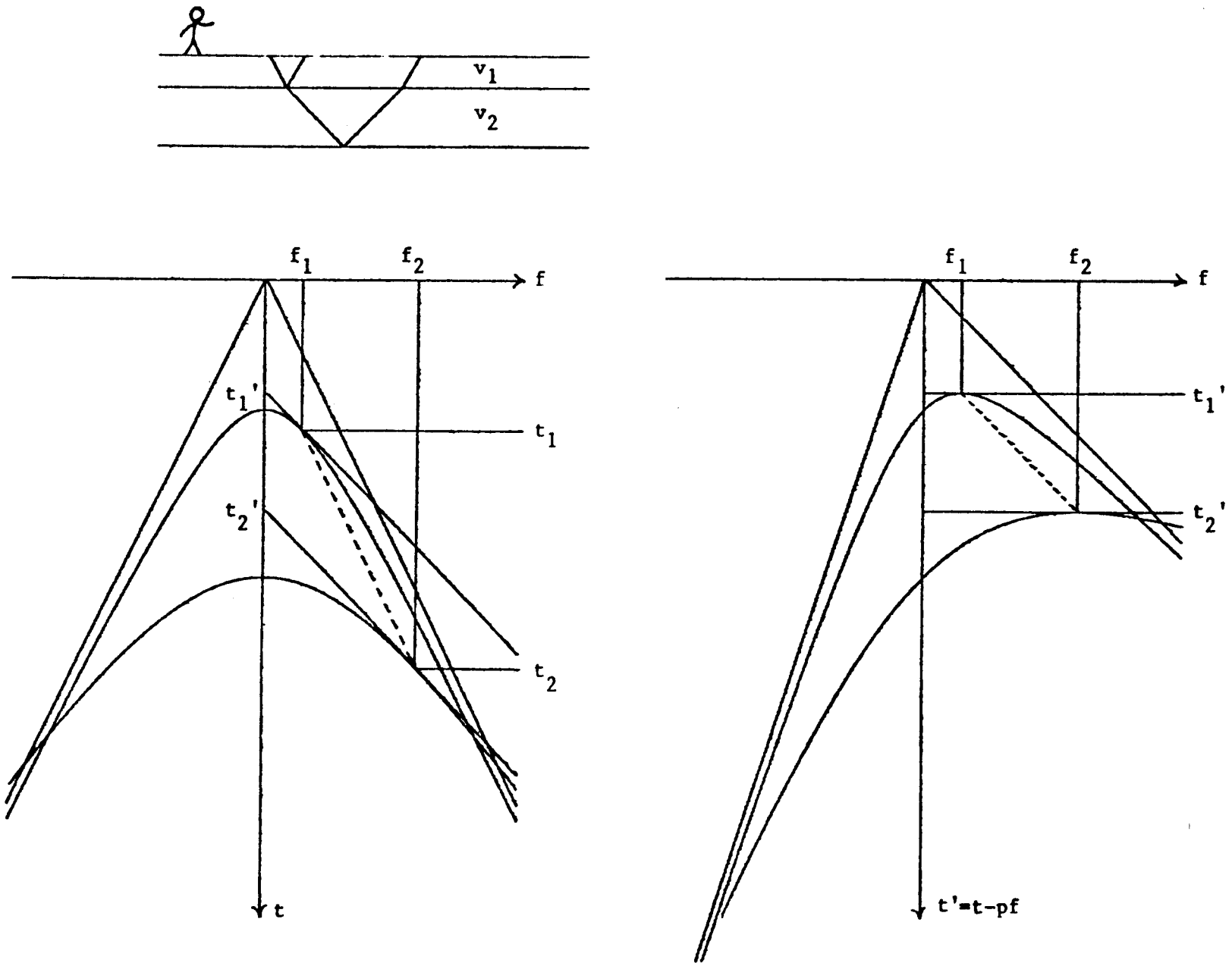

FIG. 9. (Gonzalez) Linear moveout converts the task of identifying tangencies to constructed parallel lines, to the task of locating tops of convex events.

4.5 DIP AND OFFSET TOGETHER

Before data can be summed over offset, it is necessary to apply the time shifting, described in the previous section, called *NMO* (Normal Moveout Correction). The calculation of the time shift is ordinarily based on the zero-dip assumption. Where the earth dip does not vanish it is customary to make a correction. This correction is usually regarded as an adjustment to the velocity. The data are not stacked at the earth velocity, but at a higher velocity that is equal to the earth velocity divided by the cosine of the dip. (See Clayton's Cosine Corrections.) The stacking velocity is *higher* than the earth velocity because the hyperboloid has a *flattened* top (see Cheop's Pyramid). Clearly there is quite a lot of mismatch between a flat-topped hyperboloid and a higher-velocity hyperbola. But great accuracy is never demanded when a small correction is being calculated. So in practice the *stacking velocity* is defined by the adjustment of v in the hyperbola equation, which gives the largest value of the sum.

More serious than the mismatch is the fact that, when stacking a common-midpoint gather, the amount of dip is usually not known. In fact, the data could simultaneously contain events with various dips, so that no time shift could correct all the events simultaneously.

This is the kind of problem which is solved, in principle, by migrating with the full *DSR* equation. But there are several reasons why the straightforward approach is not a realistic approach to the problem. The main practical considerations are these: The *DSR* equation does the wide angles correctly. But at wide angles the sensitivity to velocity becomes very great. In the past we have determined the velocity by the low dip assumption. Using such velocities we may expect to lose th accuracy advantage of the *DSR*. It is hard enough to run a single *DSR* job. The prospect of iteration with various velocities seems overwhelming.

Sherwood's Devilish

J.W.C. Sherwood devised a practical solution to the problem. Figure 1 shows a panel from a stacked section. The panel is shown several times, each for a different stacking velocity. It should be noticed that at the low velocities, the near horizontally dipping events predominate, whereas at the high velocities, the steeply dipping events

predominate. What Sherwood did was to find some way of correcting the raw data for dip. He called the process *Devilish*, meaning "Dipping-Event Velocity Inequalities Licked." After the *Devilish* correction, he restacked the data as before. Figure 2 shows that the stacking velocity no longer depends on the dip.

To get some idea of how the *Devilish* process could improve stacks and sections, Sherwood and his colleagues Judson, Lin, and Schultz, presented some stacked sections and migrated sections with and without the process. A stacked and migrated section without *Devilish* is shown in figure 3. With *Devilish*, it is shown in figure 4. The improvement is striking, particularly the enhancement of the fault plane reflection. It is a pitfall, however, to assume that the result is better because the fault-plane reflection is stronger. Fault-plane enhancement can always be achieved by increasing the stacking velocity. It is not very clear what were the stacking velocities in figure 3, nor as it turns out, is this comparison the most significant issue.

Of course, a good stack is nice but even more important was figure 2, which showed that the pre-stack dip corrections made it possible to stack without regard to dip. This means that it is possible to determine velocity without regard to dip. In other words, events with all dips should help 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.

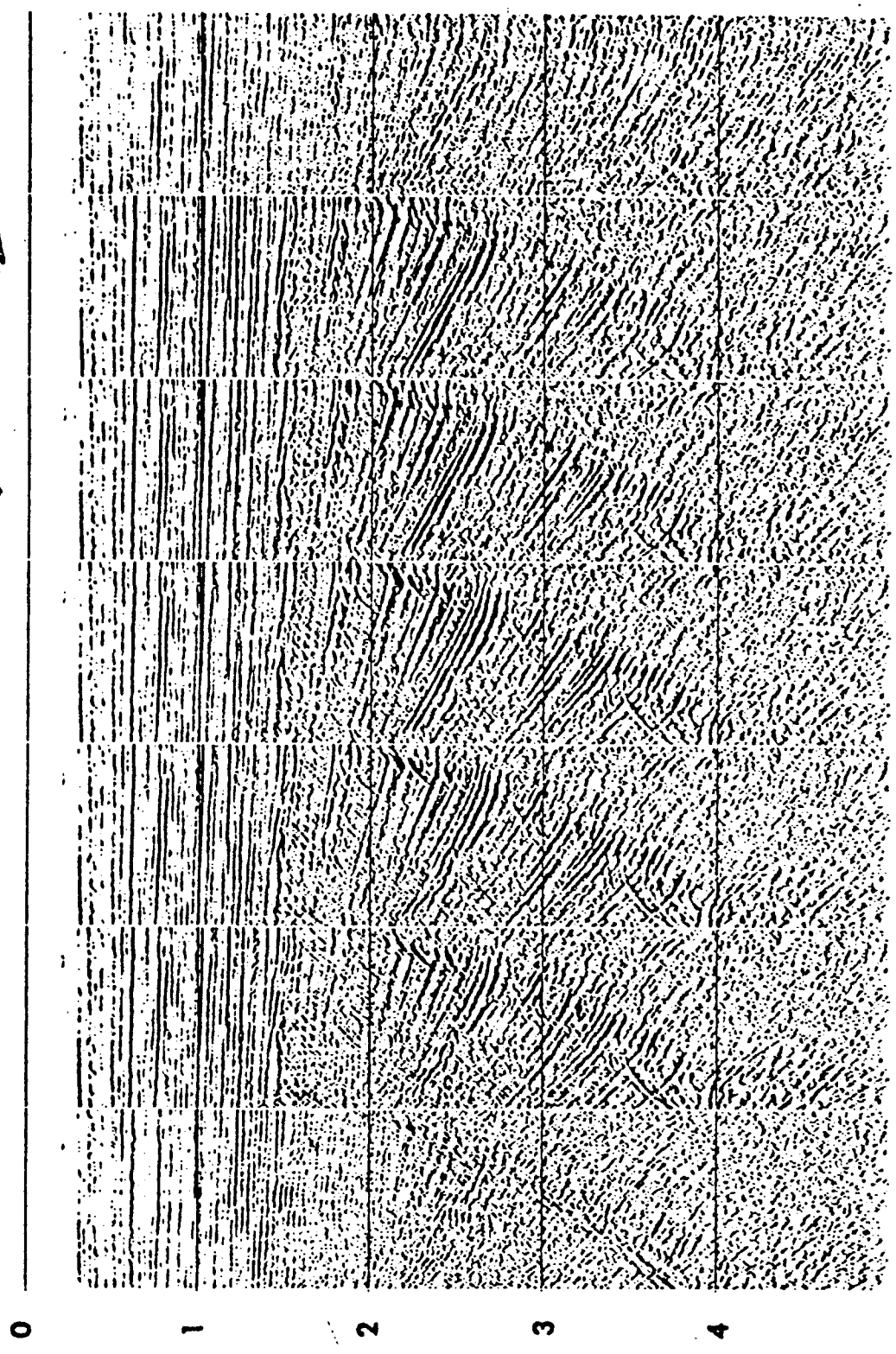

FIG. 1. (Sherwood) Conventional stacks with varying velocity.

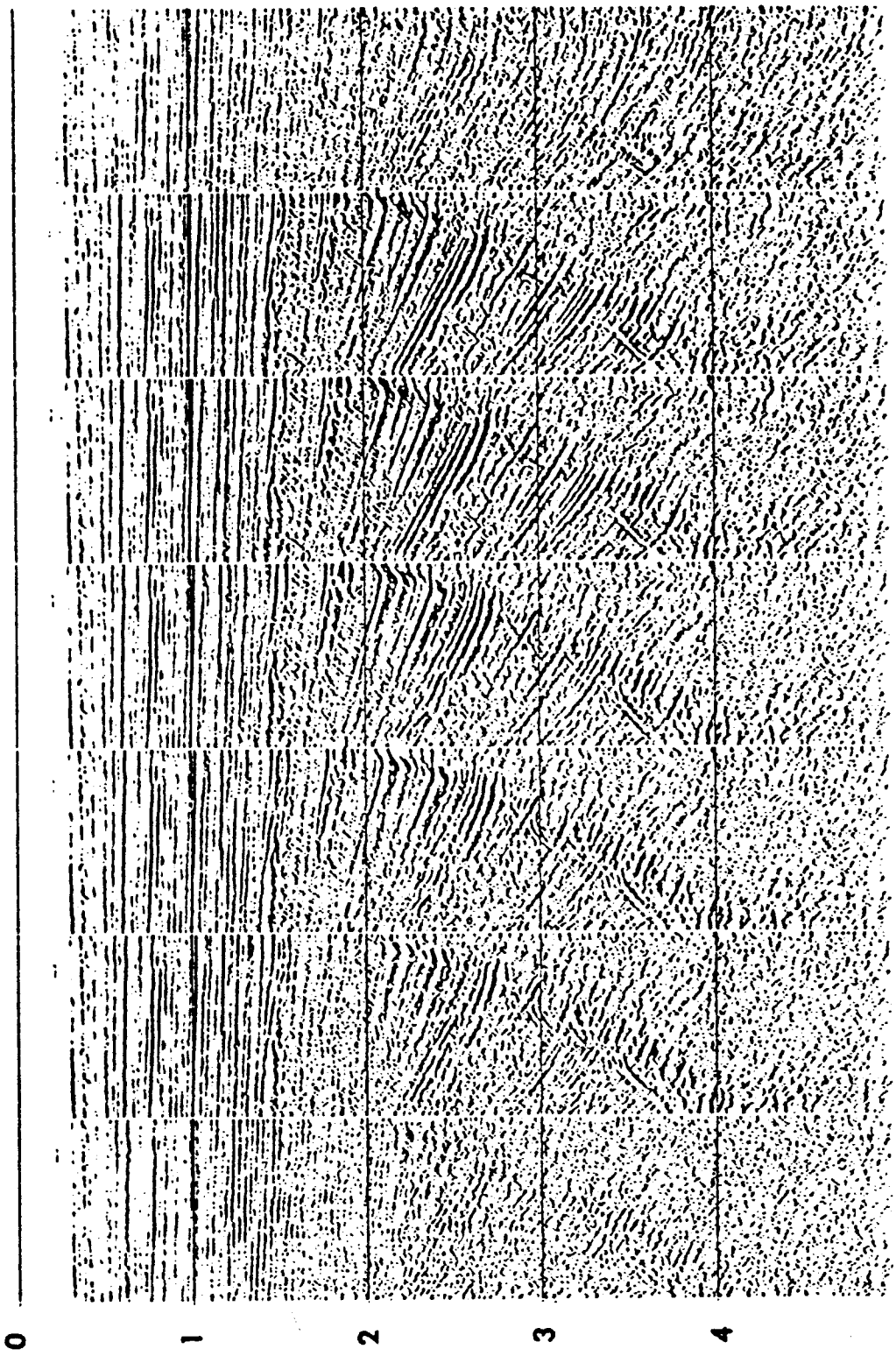

FIG. 2. (Sherwood) *Devilish* stacks with varying velocity.

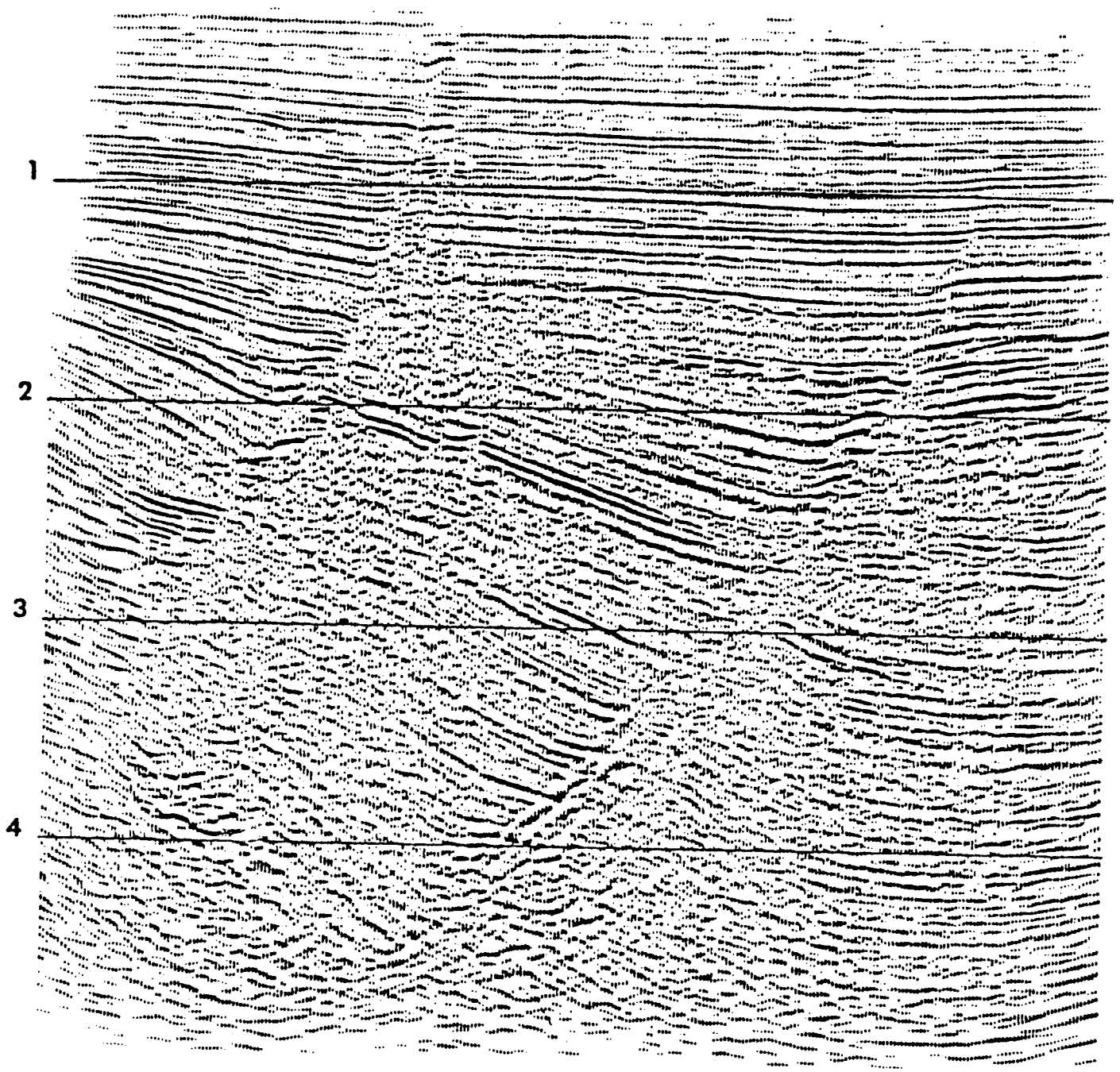

FIG. 3. (Sherwood) *CDP* sum followed by migration.

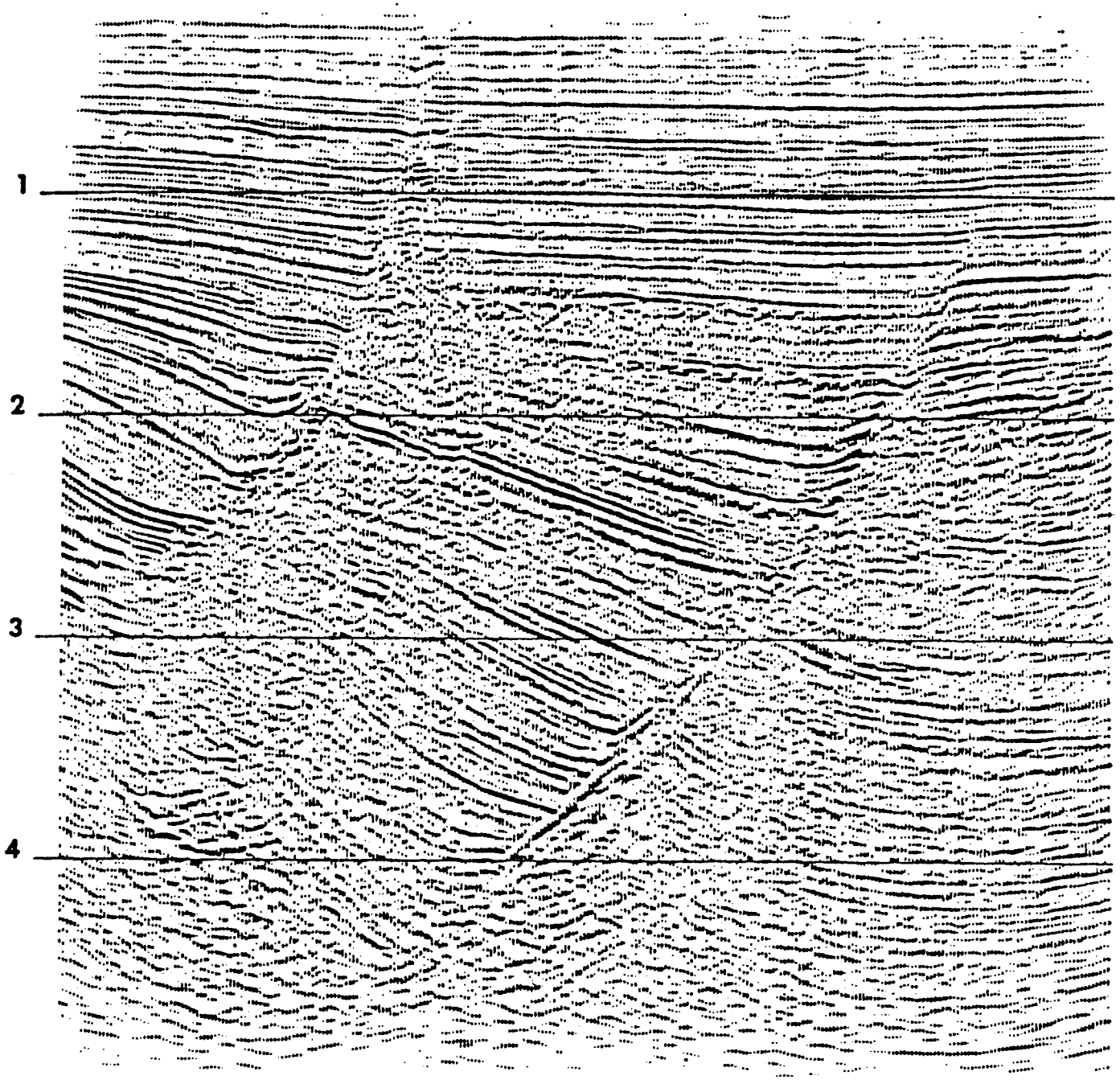

FIG. 4. (Sherwood) The *Devilish* process (pre-stack partial migration) followed by *CDP* sum, followed by migration.

Rocca's Partial Migration

Fabio Rocca developed a clear conceptual model for Sherwood's dip corrections. Figure 5 illustrates Rocca's concept of a pre-stack partial-migration operator. Imagine a constant-offset section $P(t,y,h=h_0)$, on which the data P is an impulse function at some particular (t_0,y_0). The earth model implied by this data would be a reflector shaped like an ellipse, with the shot point at one focus and the receiver at the other. Starting from this earth model we can create a zero offset section by forward modeling, that is, each point on the ellipse can be expanded into an hyperbola. Combining the two operations -- constant offset migration and zero offset diffraction -- we get the Rocca operator.

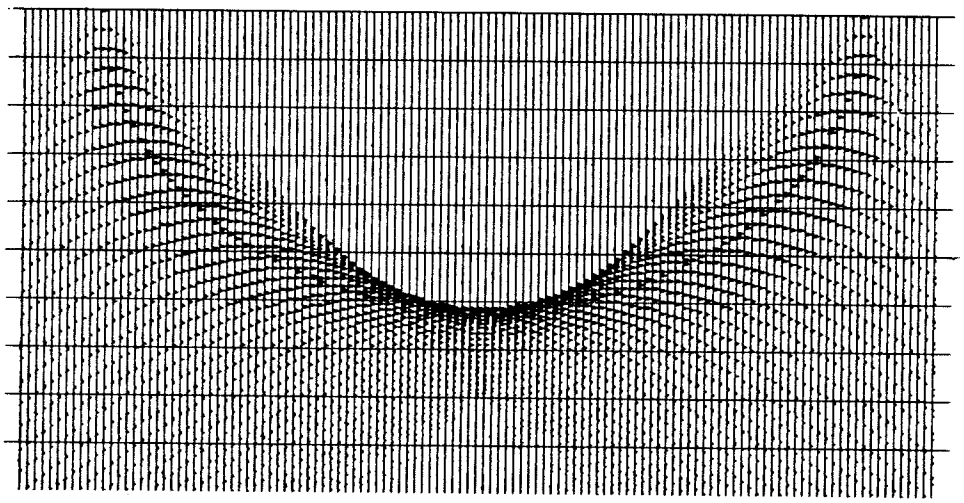

FIG. 5. (Gonzales) Rocca's pre-stack partial-migration operator is a superposition of hyperbolas, each with its top on an ellipse. Convolving (over midpoint) Rocca's operator onto a constant offset section converts it to a zero offset section.

The Rocca operator is the curve of osculation on figure 5. It is the curve where the hyperbolas reinforce one another. If the hyperbolas on figure 5 had been placed everywhere on the ellipse instead of at isolated points, then the osculation curve would be the only thing visable (and you wouldn't be able to see where it came from).

The energy in the operator is seen to be concentrating near the bottom. In fact the energy will all go to the bottom in the limiting case that the ellipse tends to a circle, that is to say, the limit where h/vt_0 is small (h = half offset, v = velocity, t_0 = travel time to the point on the constant offset section). When the energy is all

concentrated near a single place, the Rocca operator becomes a delta function.

The Rocca operator can transform a constant offset section into a zero offset section. To do this, the operator of figure 5 must be convolved across the midpoint axis of the constant offset section. This must be done for each t_0 in the input constant offset section and the results for all t_0 values must be superposed.

The Rocca operator achieves two objectives. First, it does normal moveout correction. Second, it does Sherwood's dip corrections.

From a practical point of view this operator is very attractive. Instead of implementing the big job of migrating each constant-offset section with a big, wide ellipse, we use only the narrow, little Rocca operator. After compensating each offset to zero offset, we can determine velocity by the normal-moveout residual, then stack, and finally migrate.

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

Yilmaz's Deviation

With the basic geometrical concepts of pre-stack partial migration having been illustrated by the approaches of Sherwood and Rocca, we can now set ourselves to the task of trying to find a formulation of the pre-stack, partial-migration operator which is related to the wave equation. Such an approach, if successful, should lead to more accurate treatment of amplitude, phase, and all other aspects which are included in wave equations. The natural starting point seems to be the double-square-root (DSR) equation

$$\frac{\partial P}{\partial z} = -i \frac{\omega}{v} (DSR)\, P \tag{1}$$

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

There is a serious problem at the very outset. The DSR operator is *not separable* into a sum of an offset operator and a midpoint operator. *Non- separable* means that a

Taylor series for (2) contains terms like Y^2H^2. Such terms cannot be expressed as a function of Y plus a function of H. This lack of separability is a data-processing disaster. It implies that migration and stacking must be done simultaneously, not sequentially. The only way to recover pure separability is to return to the space of S and G, but that would not at all be the approach of Sherwood and Rocca and would retreat from the belief that we seek only a modest correction.

To better understand separability and how it relates to data processing, we next make a separable approximation to (2).

$$Sep(Y,H) = 2 + [DSR(Y,0) - 2] + [DSR(0,H) - 2] \tag{3a}$$

$$= 2\,[1 + (\sqrt{1 - Y^2} - 1) + \sqrt{1 - H^2} - 1)] \tag{3b}$$

$$= TD + Mig(Y) + NMO(H) \tag{3c}$$

The result will be interpreted as "standard processing." Notice that at $H = 0$ (3) becomes equal to the DSR operator, and at $Y=0$ it becomes equal to the DSR operator. Only when both are nonzero does Sep depart from DSR. [To understand how the Sep approximation was first guessed, it helps to note that $DSR(0,0) = 2$.]

The fact that (3) has split into the sum of three operators offers an advantage like the one offered by the 2-D Fourier kernal $exp(ik_y y + ik_h h)$, which has a phase that is the sum of two parts. It means that we may nest Fourier integration with either y or h on the inside. So downward continuation with Sep could be done in (k_h,k_y)-space as implied by (1), or we could choose to Fourier transform to (h,k_y), (k_h,y), or (y,h) by appropriate nesting operations.

How is stacking to be regarded as an aspect of the NMO operator? It isn't. The NMO operator downward continues all offsets to all other offsets. Stacking is the selection of zero offset after downward continuation. Selecting zero offset is no more than abandoning all other offsets. Ordinarily the abandoned offsets would not be migrated. Alternately, a clever procedure for changing stacking velocities after migration would involve migrating several offsets near zero offset.

Since all terms in the Sep operator are interchangeable, it would be foolish and wasteful to use it to migrate all offsets prior to stack. The result should be identical to after-stack migration.

An ellipse can migrate a constant-offset section. How can the constant offset section be migrated with the DSR equation? This is a perplexing question for which I have never had a fully satisfactory answer. If the downward continuation is to be done at constant offset h, then stacking will simply be abandoning this offset. If the offset

h should decrease during the downward continuation, then another coordinate system seems to be required, one in which h depends on z as well as on $g-s$.

To devise a pre-stack partial-migration operator it seems to be necessary to cope with the kind of physical effect found in the Y^2H^2 term, but to do it in such a way as not to interpret H as an operator. Recall that the stepout operator H is interpreted in Fourier space by k_h/ω and in physical space by ∂_h^t. Both of these forms imply mixing information from different offsets. We need the ray interpretation. Let us define $2H_0 = v\ dt/dh$. Then H can be interpreted as a number, dt/dh, rather than an operator. So $Y^2H_0^2$ will be the partial-migration operator. But what should the number be?

Consider the simplest geometry in which

$$(2h)^2 + (2z)^2 = v^2t^2 \tag{4}$$

Differentiating with respect to h at constant z, we get

$$\frac{dt}{dh} = \frac{4h}{tv^2} \tag{5a}$$

$$H_0 = \frac{v}{2}\frac{dt}{dh} = \frac{2h}{vt} \tag{5b}$$

This says that the stepout dt/dh on a data gather will be predicted by knowing only velocity and position (h,t). The assumption of constant velocity could be improved by ray tracing. But there is no need to be extremely careful when computing only a small correction. When Yilmaz applied these ideas (SEP-18) he found it satisfactory to use an RMS velocity, but he also incorporated a regional dip correction

$$H_0 = \frac{2h}{v_{RMS}t}\cos 45^\circ \tag{5c}$$

This seems crude, but it is much better than standard processing, which sets the term to zero.

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

$$Dev(Y) = DSR(Y,H_0) - Sep(Y,H_0) \tag{6}$$

This defines a deviation operator. We might even call it a *devilishly* clever way of representing the difference between the DSR equation and its separable approximation in the form of a migration operator that depends on the scalars h and t.

All this may now seem rather abstract, so let us make all the substitutions and also some approximations in order to have a more familiar result. Using square-root expansions that allow H_0 to be large ≈1) we get to second order in Y

$$Dev(Y) \approx \left[1 - \frac{1}{(1-H_0^2)^{3/2}}\right] Y^2 \approx -\frac{3}{2} H_0^2 Y^2 \tag{7a}$$

Incorporating the missing $-i\omega/v$ from the DSR equation, and going to the time domain we get

$$\frac{\partial^2 P}{\partial z \, \partial t} = -\frac{3 \cos 45^\circ}{2v} \left[\frac{h}{t}\right]^2 \frac{\partial^2}{\partial y^2} P \tag{7b}$$

Yilmaz implemented (7a), the center term, and performed the partial migrations of impulses, giving what should be Rocca's operator without the NMO. The result is shown in figure 6. Then he tested the operator on some synthetic constant-offset sections, as shown in figure 7. These results, while imperfect, can be considered to be quite reasonable. The greatest difficulty seems to be in the vicinity of the flat top. This is the pole in equation (7a). There we can expect both numerical and theoretical considerations to break down.

Even if the numerical work were done perfectly, we should not expect a perfect result, for two reasons. First was the cos 45° patch. More fundamentally, we assumed that dt/dh could be expressed as a function of the coordinate system (h,t) rather than being attached to the events themselves. The Rocca operator seems to have correct timing for all dips, but the deviation operator defined here seems to have a limited range of validity in dip.

Ottolini's Chance

Ordinarily we regard a common-midpoint gather as a collection of seismic traces, that is, a collection of time functions, each at constant offset h. But this (h,t) data space could be seen with a different coordinate system. One system with some particularly nice attributes is the radial-trace system introduced by Turhan Taner. In this system the traces are not taken at constant h, but at constant angle. The idea is depicted in figure 8. Besides some nice theoretical attributes, which will become apparent, this system also has some nice practical features. Consider these: (1) The traces neatly fill the space where data are nonzero. (2) The traces are close together at early times where wavelengths are short, and wider apart where they are long. (3)

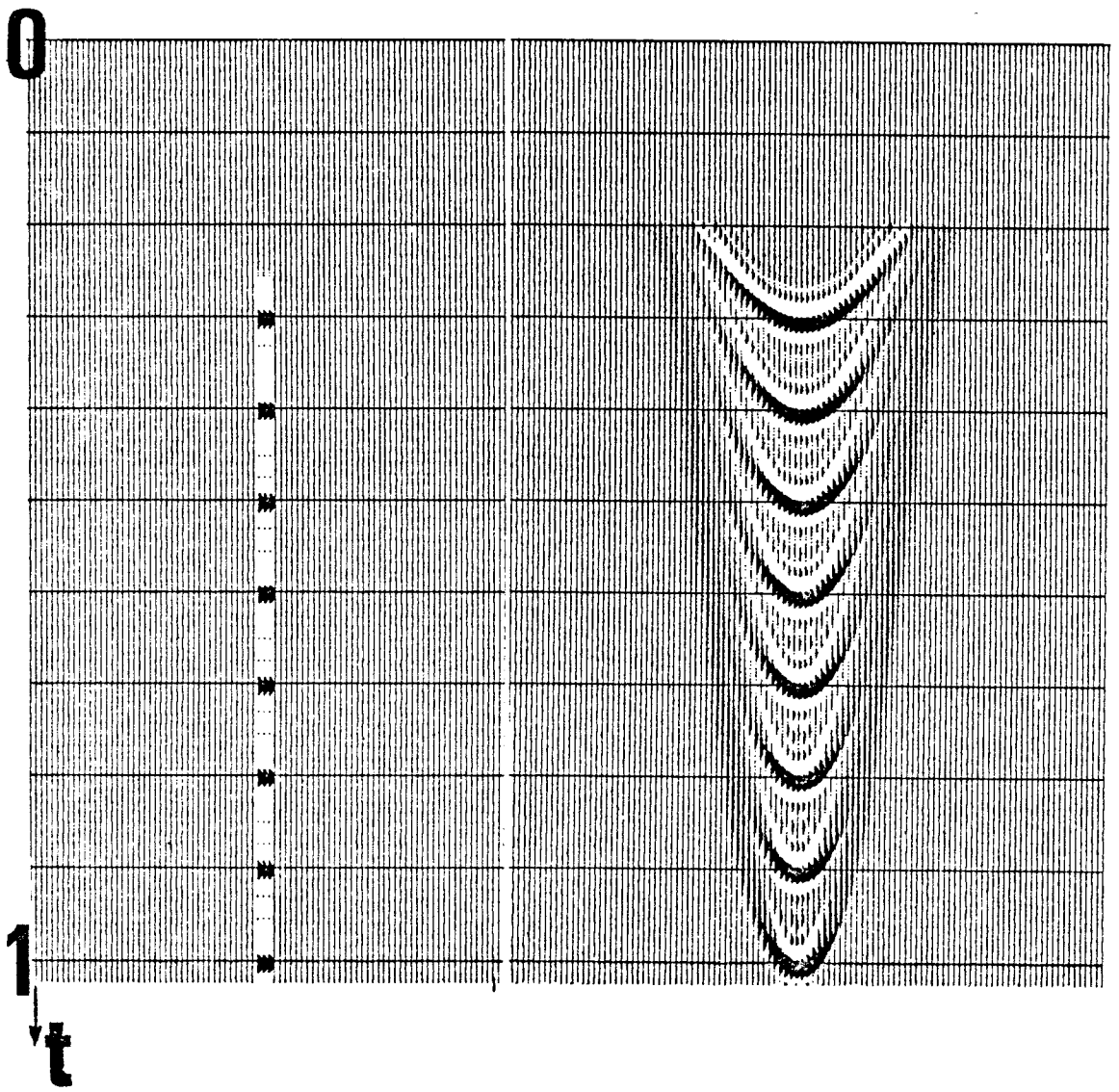

FIG. 6. (Yilmaz) Rocca's pre-stack partial-migration operator. A model (left) of impulses at various depths at some fixed offset, that is $P(y,t,h_0)$, and the response $P(y,t,0)$ (right).

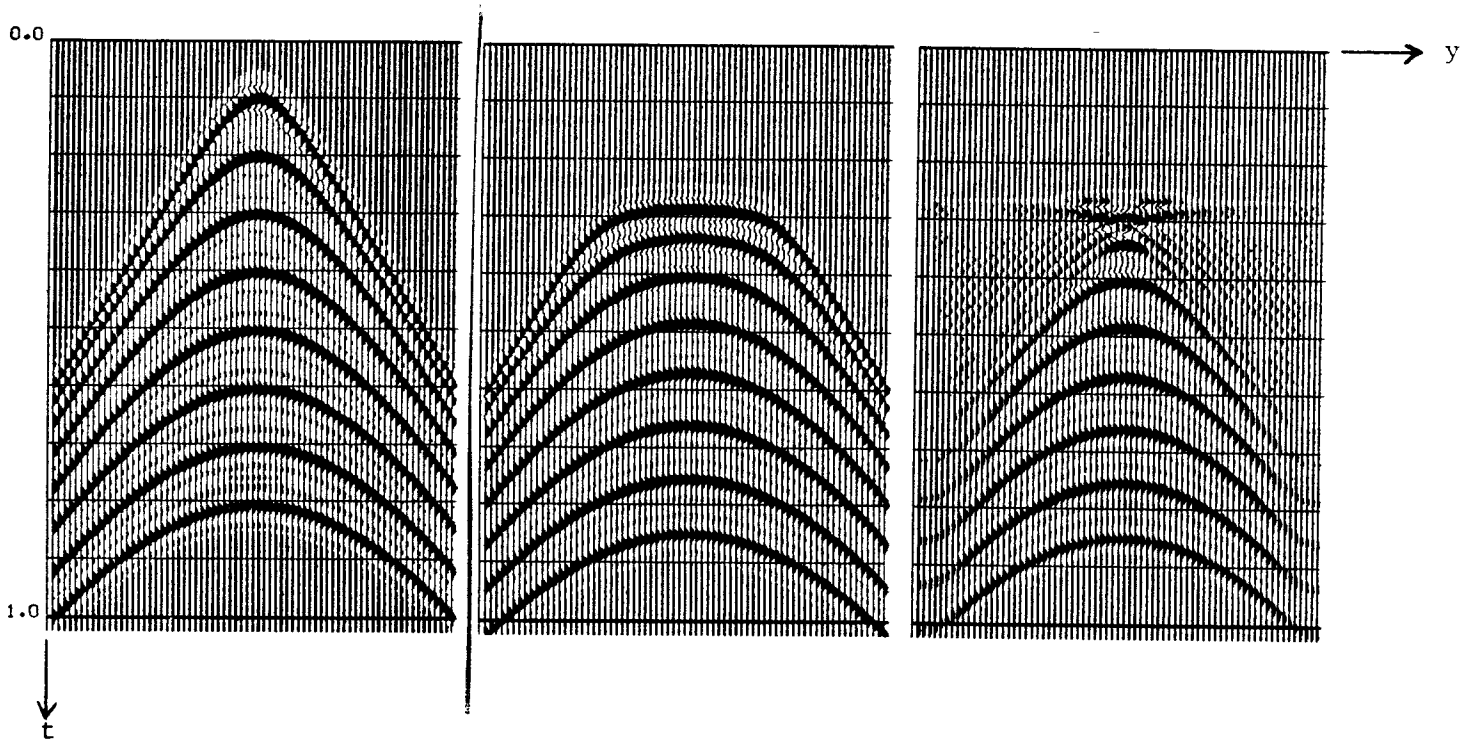

FIG. 7. (Yilmaz) Zero-offset section over point scatterers (left). Constant-offset section over same scatterers (center). Attempt of *Dev* operator to convert constant offset to zero offset.

The energy on a given trace represents wave propagation at a fixed angle. The last point is especially important with multiple reflections. Obviously for the velocity-stratified case $v=v(z)$ the radial traces would be defined for fixed Snell parameter p instead of fixed h/t.

For our present purposes the best attribute of radial traces is yet to come. Richard Ottolini noticed that the substitution $2h = vt \sin\vartheta$ into the Cheop's pyramid equation

$$vt = \sqrt{z^2 + (y+h)^2} + \sqrt{z^2 + (y-h)^2} \tag{8}$$

along with some heavy algebraic work [square (8), move the cross term to one side, square again, solve] leads to the simpler form

$$vt = 2\left[\frac{z^2}{\cos^2\vartheta} + y^2\right]^{1/2} \tag{9}$$

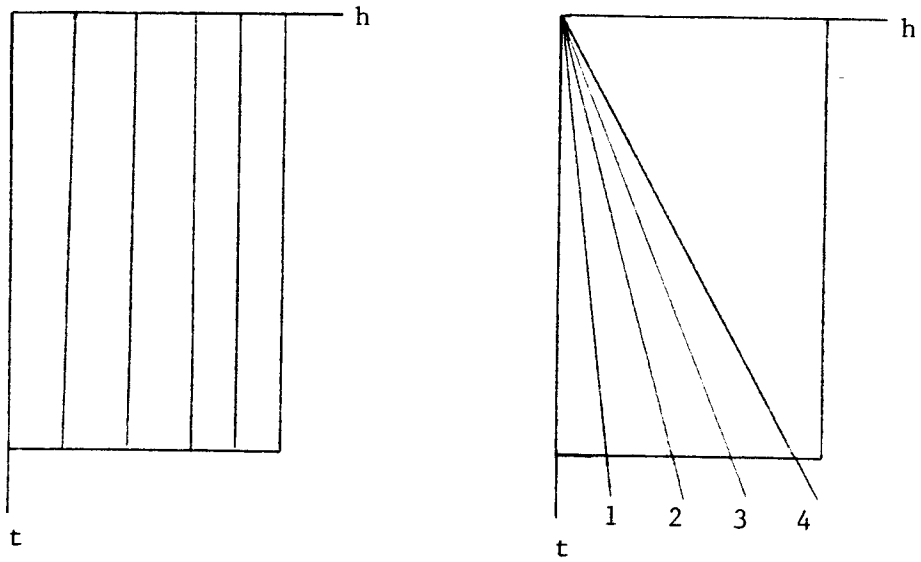

FIG. 8. The usual seismic trace is $P(h,t)$ for some fixed h. The radial trace is $P(h,t)$ for some $2h/t = \sin\vartheta$.

Scaling the z-axis by $\cos\vartheta$ we have the circle and hyperbola situation all over again! The hidden hyperbola is shown as a three-dimensional sketch on figure 9.

This means that on a section in (y,t)-space which is generated for a fixed p, the diffraction response is a simple hyperbola. From a data processing point of view it is much easier to handle than the flat-topped hyperbola that is seen at constant h. It is easy to convert a data hyperbola of one p-value to that of another; all we need to do is rescale the z-axis.

From Fourier transform theory we know that scaling z by a $\cos\vartheta$ divisor will scale k_z by a $\cos\vartheta$ multiplier. So the downward extrapolation is determined by

$$k_z = -\frac{\omega}{v}\cos\vartheta\left[1-\left[\frac{vk_y}{2\omega}\right]^2\right]^{1/2} \tag{10a}$$

All of the offset information is buried in the Snell parameter p in the $\cos\vartheta = \sqrt{1-p^2v^2}$. There are no derivatives or wavenumbers in the offset direction. The data seem to be happy to remain at a fixed p-value.

Define parameters a and b by writing (10a) as

$$k_z = -\frac{\omega}{v}\,a\,b \tag{10b}$$

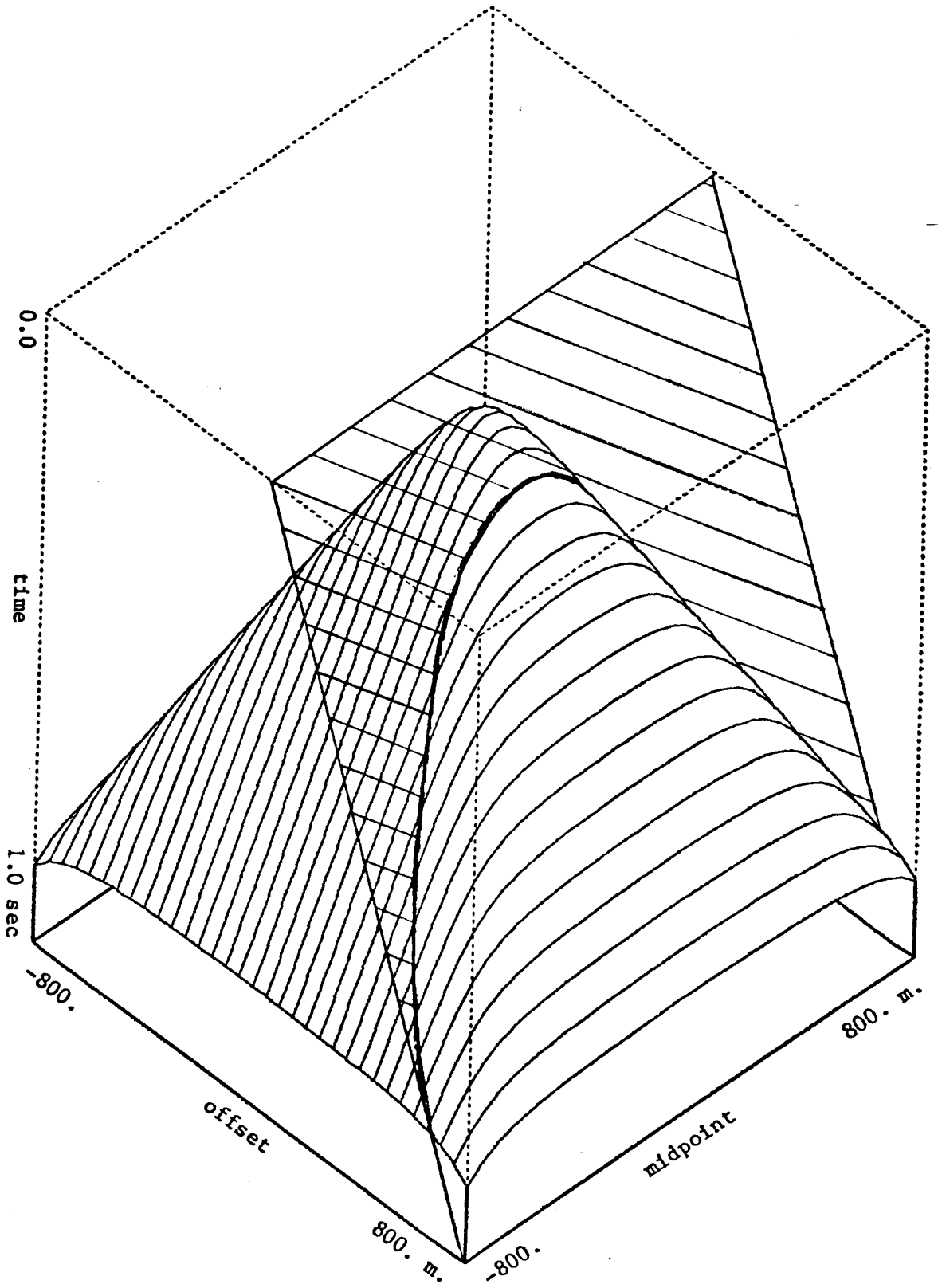

FIG. 9. (Ottolini) An unexpected hyperbola in Cheop's pyramid is the diffraction hyperbola on a radial trace section.

We have the interpretation

$$ab = 1 + (a-1) + (b-1) + (a-1)(b-1)$$

$$ab = TD + NMO + Mig + Dev \tag{11}$$

This result is delightfully simple. The generalization to stratified media may not be easy, but that is probably unnecessary. We have lots of other definitions for *NMO* and *Mig* in media which vary in depth (and even laterally). The practical view is that *Dev* is a modest correction and thus need not itself be represented with great accuracy, constant velocity should be adequate. To review, the processing sequence envisioned is

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