1.1 EXPLODING REFLECTORS

The most basic reflection seismic prospecting equipment is a source for impulsive sound waves, a geophone (something like a microphone), and a multichannel waveform display system. A survey line is defined along the earth's surface. It could be the path for a ship, in which case the receiver is called a hydrophone. About every 25 meters or so the source is activated and the echoes are recorded nearby. The sound receiver will have almost no directional tuning capability, owing to the fact that the frequencies which have deep-earth penetrating ability are those with wavelengths longer than the ship. Consequently, echoes can arrive from several directions at the same time. It is the joint task of geophysicists and geologists to interpret the results. Geophysicists assume the quantitative, physical and statistical tasks. Their main goals, and the goal to which this book is mainly directed, is to make good pictures of the earth's interior from the echoes.

A Powerful Analogy

Figure 1 depicts two wave-propagation situations which are apparently quite different. The first is our situation with field recording. The second is a thought experiment in which all of the reflectors in the earth suddenly explode. Waves from the hypothetical explosion propagate up to the earth's surface where they are observed by a hypothetical string of geophones along the earth's surface. Even if the earth had exploding reflectors, we would have difficulty recording the waves because of the need for so many geophones. It is surely much easier to tow one geophone past a thousand locations than to operate a one-thousand-channel recording system.

Notice in the figure that the raypaths in the field recording situation seem to be the same as those in the exploding reflector situation. It is a great conceptual advantage to imagine that the two wave fields, the observed and the hypothetical, are indeed the same. If they are the same, then we can ignore the many thousands of experiments which have actually been done and think only of the one hypothetical experiment. The one major, obvious difference between the two situations is that in the field geometry waves must first go down and then return upward along the same path, whereas in the hypothetical experiment they just go up. This difference could be

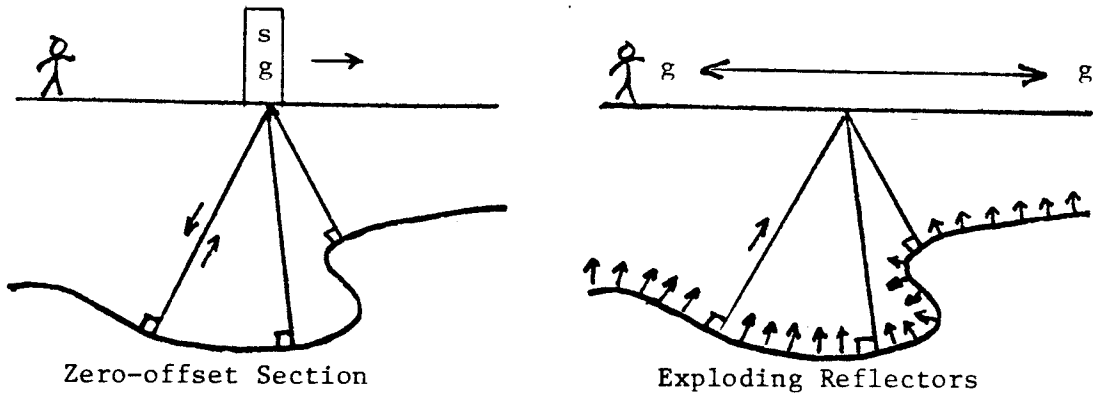


FIG. 1. The field geometry of echoes collected with a source-receiver pair at all places on the earth's surface (left) and the "exploding reflectors" conceptual model (right).

accounted for in either of two ways. We could take the traveltime in field experiments and divide by two. In practice, the data of the field experiments (two-way time) is analyzed assuming the sound velocity to be half its true value.

Huygens Secondary Point Source

Waves on the ocean have wavelengths comparable to those of waves in seismic prospecting (15-500 meters), but they are conveniently different in that they move slowly enough to be easily observed. Imagine a long harbor barrier parallel to the beach with a small entrance in the barrier for the passage of ships. This is depicted in figure 2. A plane wave incident on the barrier from the open ocean will send a wave through the gap in the barrier. It is an observed fact that in the harbor the wavefront becomes a circle with the gap as its center. The difference between this beam of water waves and a light beam through a window is in the ratio of wavelength to hole size.

A Cartesian coordinate system has been superimposed upon the ocean surface with x going along the beach and z measuring the distance from shore. To draw the analogy to reflection seismology we must say that we are confined to the beach (the earth's surface) where we can make only measurements of wave height as a function of x and t. From this data we can make inferences about the existence of a gap in the barrier out in the (x,z)-plane. Figure 3a depicts the arrival time at the beach of a wave from the ocean. The earliest arrivals occur nearest the gap. What mathematical expression determines the shape of the arrival curve seen in the (x,t)-plane?

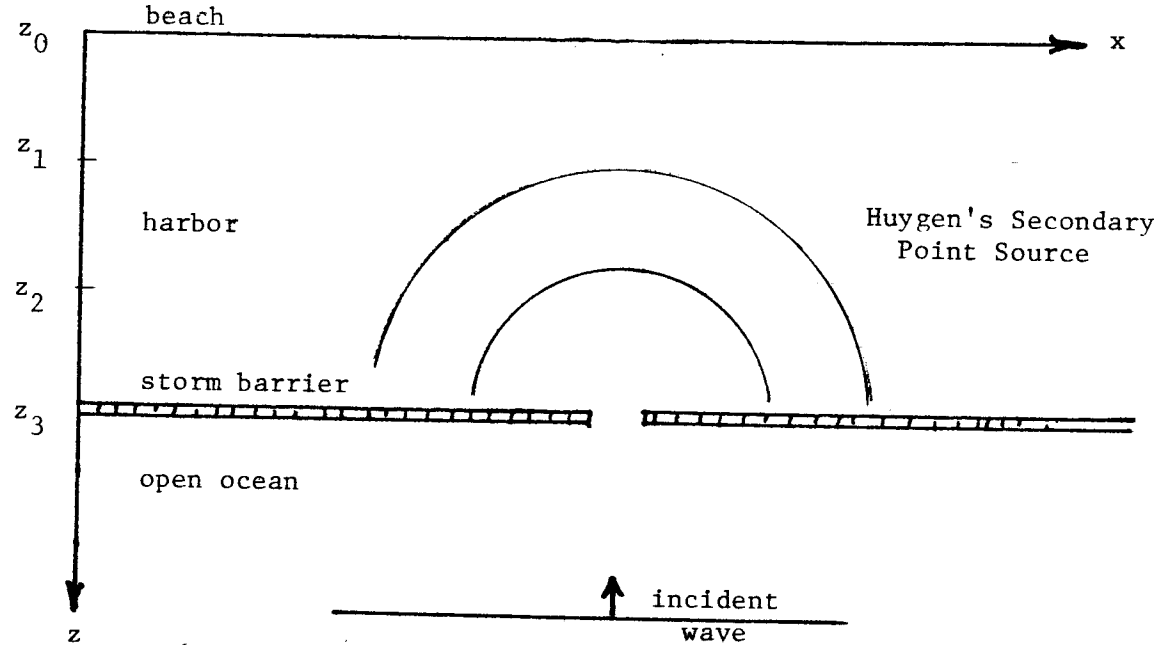


FIG. 2. Waves going through a gap in a barrier have semi-circular wavefronts (provided that the wavelength is long compared to the gap size).

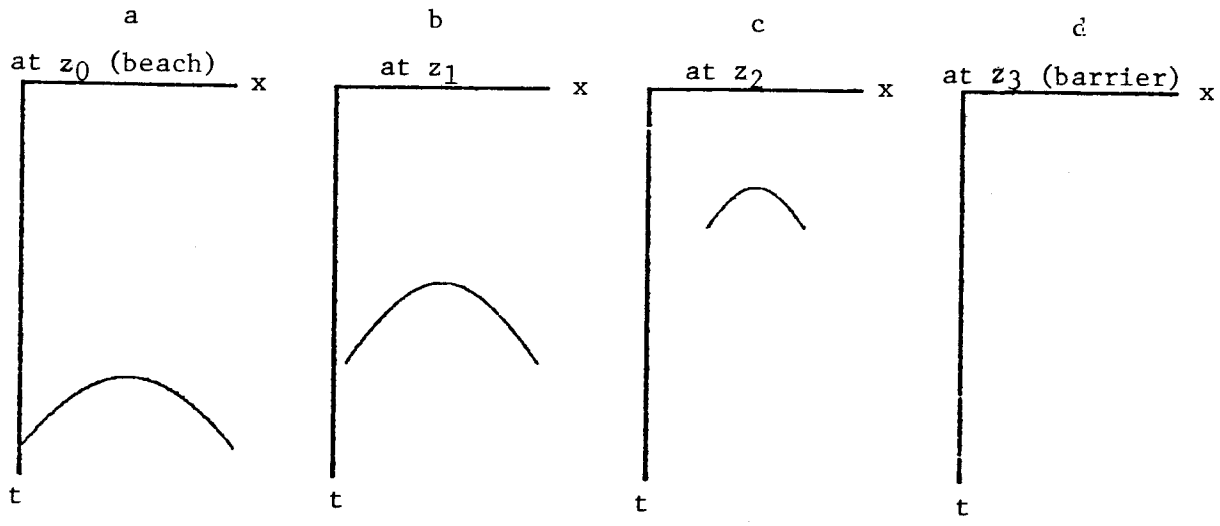


FIG. 3. The left frame shows the hyperbolic wave arrival time seen at the beach. Frames to the right show arrivals at increasing distances out in the water. (The x-axis is compressed from figure 2.)

The waves of interest are expanding circles. An equation for a circle expanding with velocity v about a point (x_3,z_3) is

$$(x-x_3)^2 + (z-z_3)^2 = v^2t^2 \qquad (1)$$

Considering t to be a constant, i.e. taking a snapshot, (1) is the equation of a circle. Considering z to be a constant, (1) is an equation in the (x,t)-plane for a hyperbola. Considered in the (t,x,z)-volume, (1) is the equation of a cone. Slices at various values of t show circles of various sizes. Slices of various values of z show various hyperbolas. Figure 3 shows four hyperbolas. The first is our observation on the beach $z_0 = 0$. The second is a hypothetical set of observations at some distance z_1 out in the water. The third, at z_2, is an even greater distance from the beach. The fourth, z_3, is nearly all the way out to the barrier where the hyperbola has degenerated to a point. All these hyperbolas are from a family of hyperbolas, each with the same asymptote. The asymptote refers to a wave which turns nearly 90 degrees at the gap and is found moving nearly parallel to the shore at the speed dx/dt of a water wave. [For this water wave analogy we presume (incorrectly) that the speed of water waves is a constant independent of water depth.]

Linearity is a property of all low-amplitude waves (not those foamy, breaking waves you see near the shore). This means that if we have two gaps in the harbor barrier we will have two semi-circular wavefronts. Where the circles cross, the wave heights combine by simple linear addition. It is interesting to think of a barrier with very many holes such as that shown in figure 4. The many semi-circles and hyperbolas combine, tending to give the wave which would have been seen if there were no barrier. Indeed, in the limiting case where the barrier disappears, being nothing but one gap alongside another, the semi-circles and the hyperbolas should all combine to make only the incident plane wave. All those waves at non-vertical angles must somehow combine with one another to extinguish all evidence of anything but the plane wave. If the original incident wave was a positive pulse, then the Huygens secondary source must consist of both positive and negative polarities in order to enable the destructive interference of all but the plane wave. So the Huygens waveform has a phase shift. Eventually we will find mathematical expressions for the Huygens secondary source. Another property we will discover, well known to boaters, is that the Huygens semi-circle has its largest amplitude pointing straight towards shore. The amplitude drops to zero for waves moving parallel to the beach. In optics this amplitude dropoff with angle is called the *obliquity factor*.

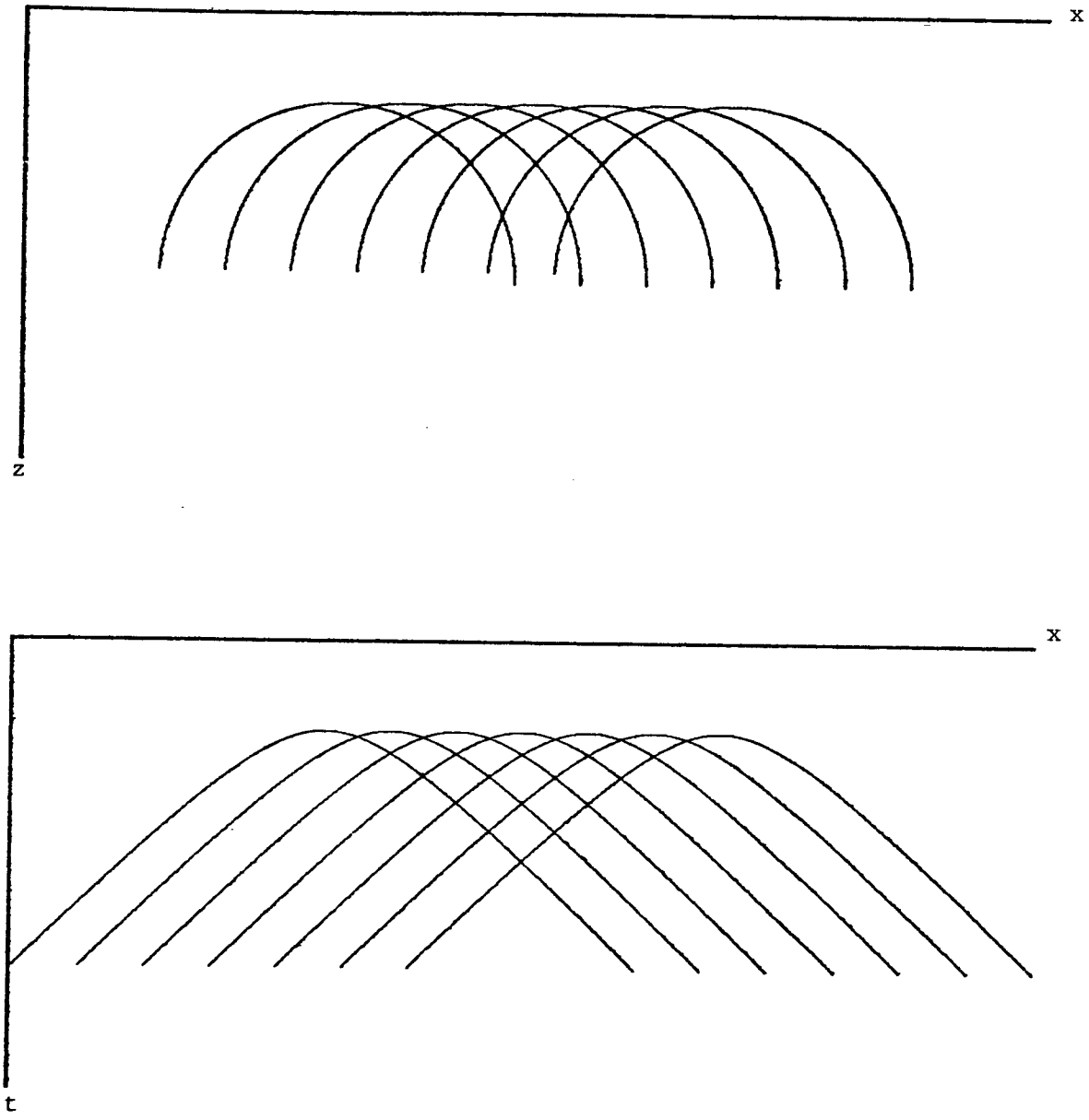


FIG. 4. (Gonzalez) Top shows a superposition of many Huygens semi-circular wavefronts to create a nearly planar wave. Bottom shows the superposition of the hyperboloids.

Migration Defined

Looking in the dictionary at the word "run" you find many definitions. They are related, but they are distinct. The word "migration" in geophysical prospecting likewise has about four related but distinctive meanings. The simplest is like the meaning of the word "move." When an object at some location in the (x,z)-plane is found at a different location at a later time t, then we say it *moves*. Analogously when a wave arrival (often called "an event") at some location in the (x,t)-space of geophysical observations is found at a different position for a different survey line at a greater depth z, then we say it *migrates*.

To see this more clearly we imagine the four frames of figure 3 being taken from a movie. During the movie, the depth z changes beginning from the beach (earth's surface) going out to the storm barrier. The frames are superimposed in figure 5a. Mainly what happens in the movie is that the event migrates upward toward $t=0$. To remove this dominating effect of vertical translation we make another superimposition, keeping the hyperbola tops all in the same place. This is done by replacing the time t-axis by a so-called *retarded* time axis $t'=t+z/v$, as shown in figure 5b. Our second, more precise, definition of *migration* is the motion of an event in (x,t')-space as z changes. Having removed the vertical shift, we are seeing mainly a shape change.

It is of interest to see how the shape actually changes. Think of a pebble thrown into the water and the ensuing circular wave. At the end of any ray from the center to the circle is a wavefront whose slope is given by some $dx/dz = \tan\vartheta$. This angle is constant as the circle grows with t. Likewise, in (x,t)-space, the wavefront, called an event, has a slope $dt/dx = \sin\vartheta/v$ which remains constant as z increases. Figure 5b was drawn so that the hyperbolas end at $\vartheta = 45$ degrees. These endpoints migrate along a straight line in the (x,t')-plane toward the center, which they hit at depth z_3.

In this case the exploding reflector is like a short line segment across the barrier gap. At depth z_3 all the energy in the (x,t')-space of migrated data is located in the position of the gap. In other words, it is focused. The third definition of migration is that it is the process which somehow pushes observational data -- wave height as a function of x and t -- from the beach to the barrier.

To go farther we need a more general example than the storm barrier example. The barrier example is confined to making Huygens sources only at some particular z, and we need sources at other depths as well. Then, given a wave extrapolation process to move data to increasing z values, we can construct our exploding reflector images with

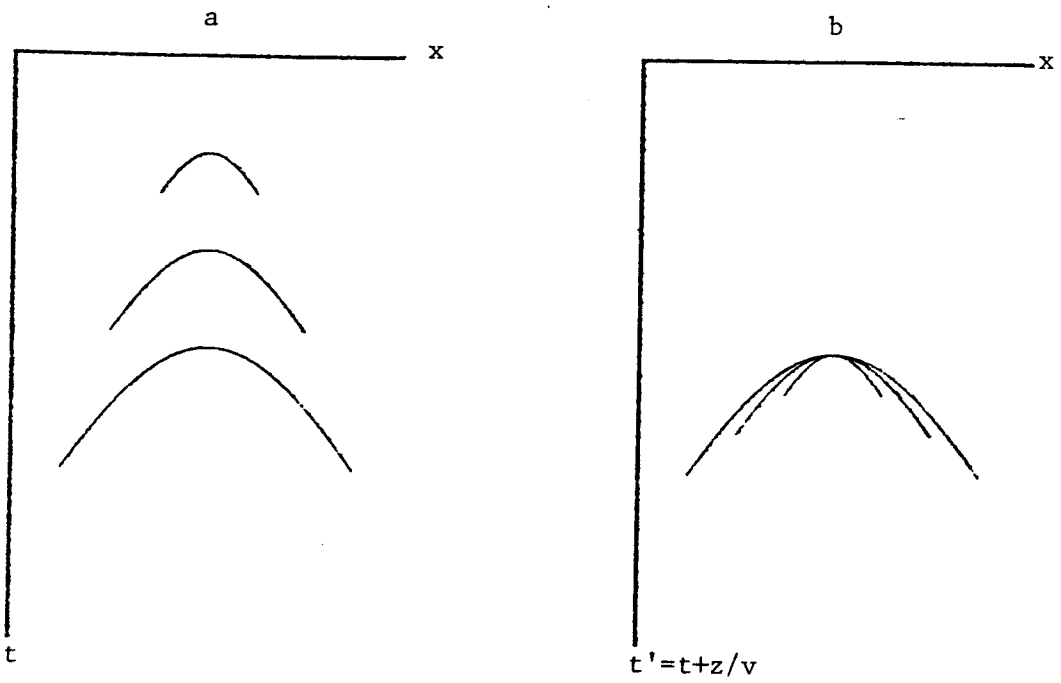
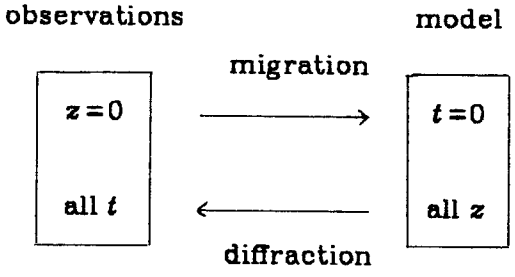


FIG. 5. (Gonzalez) Left shows a superposition of the hyperbolas of figure 3. At the right the superposition incorporates a shift, called retardation $t'=t+z/v$, to keep the hyperbola tops together.

$$Image\,(x,z) \;=\; Wave\,(t=0,x,z) \tag{2}$$

Our fourth definition of migration also incorporates the definition of "diffraction" as the opposite of migration.



Diffraction is sometimes regarded as the natural process which creates and enlarges hyperboloids. *Migration* is the computer process which does the reverse.

Another aspect of the use of the word "migration" arises in Chapter 4 where the horizontal coordinate can be either midpoint y or shot to geophone offset h. Hyperboloids can be downward continued in both the (y,t) and the (h,t) plane. In the

(y,t) plane this is called *migration* or *imaging* and in the (h,t) plane it is called *focussing* or *velocity analysis*

An Impulse in the Data

We have seen that Huygens diffraction takes an isolated pulse function (delta function) in (x,z)-space and makes it into a hyperbola in (x,t)-space at $z=0$. The converse is to start from a delta function in (x,t)-space at $z=0$. This converse refers to a seismic survey in which you record no echoes except at one particular location and there you record only one echo. What earth model is consistent with such observations? As shown in figure 6 this earth must contain a spherical mirror whose center is at the anomalous recording position.

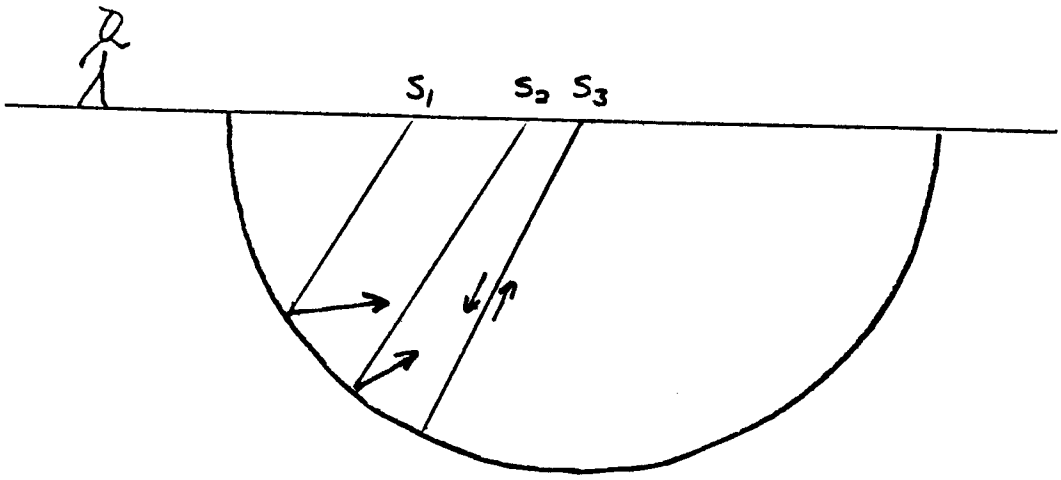


FIG. 6. When the seismic source S is at the exact center of a semi-circular mirror, then, and only then, will an echo return to the geophone at the source. This semi-circular reflector is the logical consequence of a dataset where one echo is found at only one place on the earth.

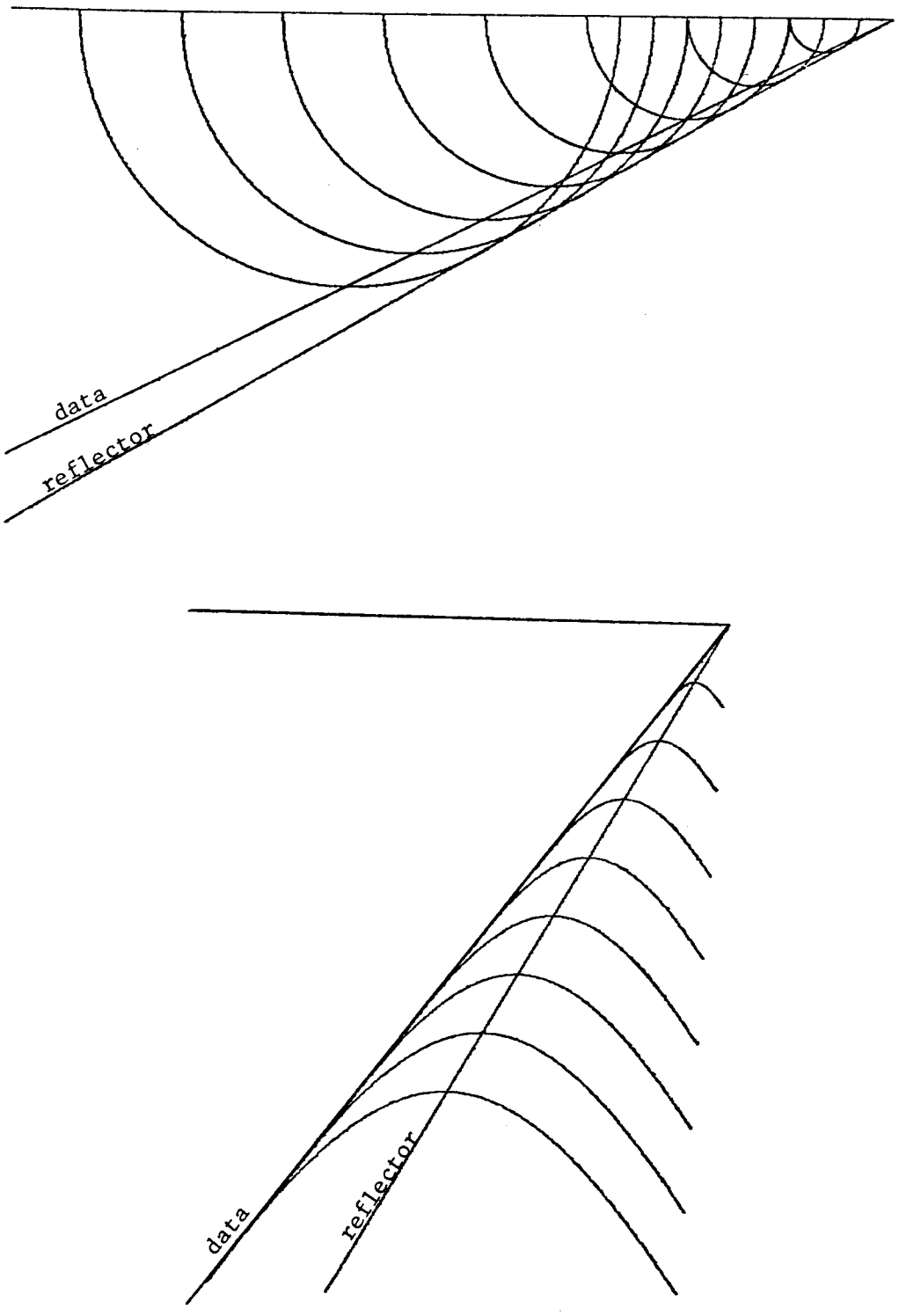


FIG. 7. Top shows the circular mirrors in (x,z)-space for a 60-degree dipping line of point sources. Bottom shows the hyperboloidal wavefronts seen at the earth's surface.

Migration Steepens Reflectors

It is true that flanks of hyperbolas migrate without change of slope. But a hyperbola is a special kind of event which comes from a single source at a single depth. Superposing point sources from different depths into a dipping planar reflector we find that migration steepens the reflections. This could be suspected by consideration of the limiting case, a vertical wall. Its reflections, the asymptotes of a hyperbola, have a non-vertical steepness. To see this in a less extreme case, see figure 7, where a dipping bed, of dip of about 60 degrees, is made from a series of points in a line.

Limitations of the Exploding Reflector Concept

The exploding reflector concept is a most powerful and fortunate analogy. For people who spend their time working entirely on data interpretation rather than processing, the exploding reflector concept is more than a vital crutch. It's the only means of transportation! For those of us who work on data processing, the exploding reflector concept has a very serious shortcoming. No one has yet figured out how to extend the concept to apply to data recorded at nonzero offset. Furthermore, most data is recorded at rather large offsets. In a modern marine prospecting survey, the recording cable (a cable containing not one but many hundreds of hydrophones) is typically 2-3 kilometers long. Drilling may be about 3 kilometers deep. So in practice the angles are big. Therein lie both new problems and new opportunities, none of which we will consider until Chapter 4.

Furthermore, even at zero offset, the exploding reflector concept is not quantitatively correct. Later efforts, mainly in Chapter 4 will elucidate, to some degree, the region of validity. For the moment let us just note two obvious failings. First, figure 8 shows a ray which is not predicted by the exploding-reflector model, but which will be present in a zero-offset section. Notice that lateral velocity variation is required for this situation to exist.

Second, consider the situation with multiple reflections. For a flat seafloor with a two-way traveltime t_1, multiple reflections are predicted at times $2t_1$, $3t_1$, $4t_1$, etc. In the exploding-reflector geometry the first multiple has first a path from reflector to surface, then from surface to reflector, then from reflector to surface, for a total time $3t_1$. Subsequent multiples occur at times $5t_1$, $7t_1$, etc. Clearly there is no relationship between the multiple reflections generated on the zero-offset section and those of the exploding-reflector model. This explains why Chapter 5 of this book, which has to do with modeling and suppressing multiple reflections, completely abandons the zero-

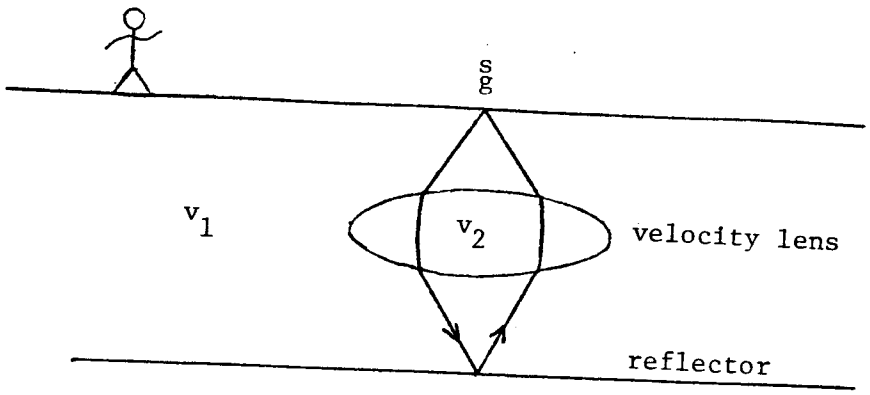


FIG. 8. A ray, not predicted by the exploding reflector model, which would nevertheless be found on a zero-offset section.

offset approach.

1.2 WAVE EXTRAPOLATION AS A 2-D FILTER

In Fourier analysis we are familiar with the idea that an impulse function (delta function) can be constructed by superposition of sinusoids (or complex exponentials). In the study of time series this construction is used for the *impulse response* of a filter. In the study of functions of space, it is used to make a physical point source.

Taking time and space together, Fourier components can be interpreted as monochromatic waves. Physical optics (and with it reflection seismology) becomes an extension to filter theory. In this section we learn the mathematical form, in Fourier space, of the Huygen's secondary source. It is a two-dimensional (2-D) filter for spatial extrapolation of wave fields.

Rays and Fronts

Figure 1 depicts a ray moving down into the earth at an angle ϑ from the vertical. Perpendicular to the ray is a wavefront. By elementary geometry the angle between the wavefront and the earth's surface is also ϑ. The ray increases its length at a speed v. The speed which is observable on the earth's surface is the intercept of the wavefront with the earth's surface. This speed, namely $v/\sin\vartheta$, is faster than v. Likewise, the speed of the intercept of the wavefront and the vertical axis is $v/\cos\vartheta$. A mathematical expression for a straight line, like that shown to be the wavefront in figure 1 is

$$z = z_0 - x \tan\vartheta \tag{1}$$

In this expression z_0 is the intercept between the wavefront and the vertical axis. To make the intercept move downward, we replace it by the appropriate velocity times time

$$z = \frac{vt}{\cos\vartheta} - x \tan\vartheta \tag{2}$$

Solving for time we get

$$t(x,z) = \frac{z}{v}\cos\vartheta + \frac{x}{v}\sin\vartheta \tag{3}$$

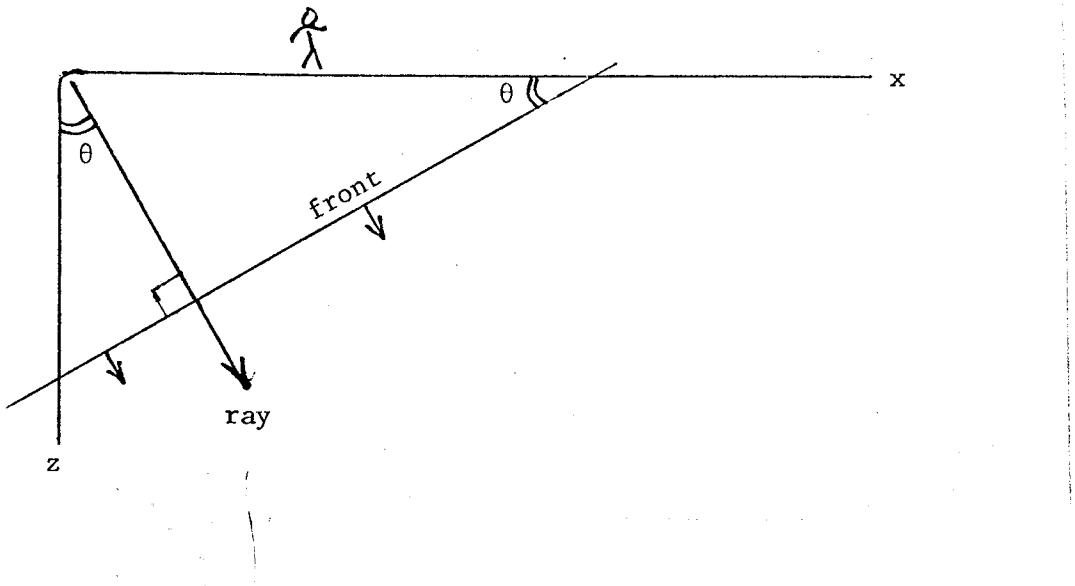


FIG. 1. Downgoing ray and wavefront.

Equation (3) tells us what time the wavefront will pass any particular location (x,z). The expression for an arbitrary shifted waveform is $f(t - t_0)$. Using (3) to define the time shift t_0 we have an expression for a wave field which is some waveform moving on a ray.

$$moving\ wave\ field\ =\ f\left[t - \frac{x}{v}\sin\vartheta - \frac{z}{v}\cos\vartheta\right] \tag{4}$$

Waves in Fourier Space

Arbitrary functions can be made from the superposition of sinusoids. Sinusoids and complex exponentials commonly occur. One reason they occur is because they are the solutions to linear partial differential equations (PDE's) with constant coefficients. The PDE's arise because most laws of physics are expressible as PDE's.

Specializing the moving wave field in equation (4) to be a cosine function of frequency ω, and using the fact that cosine is even, i.e. $\cos\vartheta = \cos -\vartheta$, we have

$$cosine\ on\ a\ ray\ =\ \cos\left[\omega\left[\frac{x}{v}\sin\vartheta + \frac{z}{v}\cos\vartheta - t\right]\right] \tag{5}$$

Using Fourier integrals on time functions we encounter the *Fourier kernel* $\exp(-i\omega t)$. To use Fourier integrals on the space-axis x we need to define the spatial angular

frequency. Since we will ultimately encounter quite a few different space axes (three for shot, three for geophone, also the midpoint and offset), we will adopt the convention of using a subscript on the letter k to denote the axis being Fourier transformed. So k_x is the angular spatial frequency on the x-axis and $\exp(ik_x x)$ is its Fourier kernel. For each axis and Fourier kernel there is the question of the choice of the sign of i. The sign choice is discussed later in more detail, but essentially we will choose the sign convention of most physics books, namely, to agree with equation (5), which is a wave moving in the positive direction along the space axes. Thus the Fourier kernel for (t,x,z)-space will be taken to be

$$\textit{Fourier kernel} \;=\; \exp[i(k_x x + k_z z - \omega t)] \tag{6}$$

Now for the whistles, bells, and trumpets. Comparing (5) and (6) we learn how to relate physical angles to velocity and Fourier components. These relations should be memorized!

Angles and Fourier Components	
$\sin\vartheta = \dfrac{vk_x}{\omega}$	$\cos\vartheta = \dfrac{vk_z}{\omega}$

(7)

Equally important is what comes next. We may insert the angle definitions into the familiar relation $\sin^2\vartheta + \cos^2\vartheta = 1$. This gives a most important relationship, known by the impressive name as the *dispersion relation of the scalar wave equation.*

$$\boxed{k_x^2 + k_z^2 \;=\; \frac{\omega^2}{v^2}} \tag{8}$$

We'll encounter *dispersion relations* and the *scalar wave equation* later. The reason why (8) is so important is that enables us to make the distinction between an arbitrary function and an apparently chaotic function which actually is a wave field. Take any function $p(t,x,z)$. Fourier transform it to $P(\omega,k_x,k_z)$. Look in the (ω,k_x,k_z)-volume for any non-vanishing values of P. You will have a wave field if and only if all non-vanishing P have coordinates which satisfy (8). Even better, in practice we often know the (x,t) dependence at $z=0$, but we do not know the z-dependence. Then we find the z-dependence by the assumption that we have a wavefield, so the z-dependence is implied from (8).

Two-Dimensional Fourier Transform

Before going any further, let us review some basic facts about two-dimensional Fourier transformation. A two-dimensional function is represented in a computer as numerical values in a matrix. A one-dimensional Fourier transform in a computer is an operation on a vector. A two-dimensional Fourier transform may be accomplished by a sequence of one-dimensional Fourier transforms. You may first transform each column vector of the matrix and then transform each row vector of the matrix. Alternately you may first do the rows and later do the columns.

We can diagram the calculation as follows:

$$\begin{array}{ccc} p(t,x) & ----- & P(\omega,x) \\ | & & | \\ P(t,k_x) & ----- & P(\omega,k_x) \end{array}$$

A notational problem on the diagram is that we cannot maintain the usual convention of using a lower case letter for the domain of physical space and an upper case letter for the Fourier domain, because the convention cannot include the mixed objects $P(t,k_x)$ and $P(\omega,x)$. Rather than invent some new notation it seems best to let the reader use the context to cope with this notational problem. The arguments of the function must help serve as the name of the function.

Altogether, the two-dimensional Fourier transform of a collection of seismograms involves only twice as much arithmetic as the one-dimensional Fourier transform of each seismogram. This is lucky. Let us write a few equations to establish that the asserted procedure does indeed do a two-dimensional Fourier transform. First of all we express the idea that any function of x and t may be expressed as a superposition of sinusoidal functions

$$p(t,x) = \int\int e^{-i\omega t + ik_x x} P(\omega,k_x)\, d\omega\, dk_x \tag{9}$$

The kernel in this *inverse* Fourier transform takes the form of a wave moving in the plus x-direction. Likewise, in the *forward* Fourier transform, the sign of both exponentials changes, preserving the fact that the kernel is a wave moving positively. The scale factor and the infinite limits are omitted as a matter of convenience. (The limits and scale both differ from the discrete computation, so why bother?) With (9) we are doing the inverse transform. Now let us nest the double integration in a form which indicates that the temporal transforms are done first (inside):

$$p(t,x) = \int e^{ik_x x} \left[\int e^{-i\omega t} P(\omega,k_x)\, d\omega \right] dk_x = \int e^{ik_x x} P(t,k_x)\, dk_x$$

The quantity in brackets indicates temporal Fourier transforms being done for each and every k_x. Alternately, we could do the nesting with the k_x-integral on the inside. That would imply rows first instead of columns (or vice versa). It is the separability of $\exp(-i\omega t + ik_x x)$ into a product of exponentials which makes the computation this easy and cheap.

The Input-Output Relation

Let us return to the dispersion relation (8)

$$k_x^2 + k_z^2 = \frac{\omega^2}{v^2}$$

In applications where time evolves it is convenient to solve (8) for $\omega(k_x,k_z)$. In extrapolation applications it is convenient to solve for $k_z(\omega,k_x)$. Consider first an evolution situation. Inspect the integral below.

$$p(x,z,t) = \int \int \left[P(k_x,k_z,t=0)\, e^{-i\omega(k_x,k_z)t} \right] e^{ik_x x + ik_z z}\, dk_x\, dk_z \qquad (10)$$

At $t=0$ it is just a double inverse Fourier transform which represents initial conditions in the (x,z)-plane. Taking $P(k_x,k_z,0)$ to be constant would be a point source at $(x,z) = (0,0)$. The time-dependence in (10) has been chosen [by selecting $\omega(k_x,k_z)$] to ensure that $p(x,z,t)$ is a wave field which fits the initial conditions at $t=0$.

Next consider the wave-extrapolation situation implied below:

$$p(x,z,t) = \int \int \left[P(k_x,z=0,\omega)\, e^{ik_z(\omega,k_x)z} \right] e^{-i\omega t + ik_x x}\, d\omega\, dk_x \qquad (11)$$

At $z=0$ it is just a double inverse Fourier transform which could represent geophysical observations in the (t,x)-plane at the earth's surface. The depth-dependence has been chosen [by selecting $k_z(\omega,k_x)$] to ensure that $p(x,z,t)$ is a wave field that matches the surface observations at $z=0$.

So far we have not mentioned that $\omega(k_x,k_z)$ and $k_z(\omega,k_x)$ are square-root functions and consequently that there is a choice of signs. Initial conditions will determine what combination of the two solutions are desired. In the extrapolation case we have

$$k_z = \pm \left[\frac{\omega^2}{v^2} - k_x^2 \right]^{1/2} \qquad (12a)$$

$$= \pm \frac{\omega}{v}\left[1 - \frac{v^2 k_x^2}{\omega^2}\right]^{1/2} \tag{12b}$$

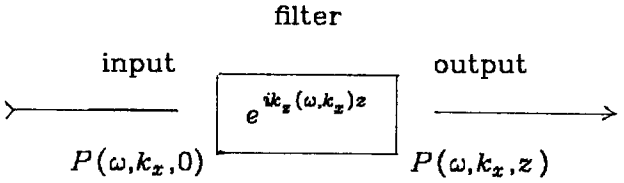
$$= \pm \frac{\omega}{v} \cos \vartheta \tag{12c}$$

Choice of the plus means that $\exp(-i\omega t + ik_z z)$ is a downgoing wave. The minus sign makes it upcoming (the usual case).

The quantity in brackets in (11) may be evaluated at any value of z. Given its value at one value of z, say $z=0$, we can determine its value at another. That is,

$$P(\omega,k_x,z) = P(\omega,k_x,0)\, e^{ik_z(\omega,k_x)z} \tag{13}$$

This is a product relationship in both the ω-domain and the k_x-domain. Hence it can be regarded as a convolutional filter in t and x. In terms of engineering flow diagrams with inputs and outputs, equation (13) may be thought of as



What does this filter look like in the time and space domain? Physically, it is the Huygen's secondary wave source which was described in terms of ocean waves entering a gap through a storm barrier. Adding up the response of multiple gaps in the barrier would be convolution over x. Superposing many incident ocean waves would be convolution over t. Mathematically, the exact inverse 2-D transform of the filter is a difficult task, well beyond the level of our present efforts. As a practical matter the 2-D transforms are rather easy in a computer. Some slices of the conic section are found in FGDP on pages 199-200.

Exercise

1. Let $P(k_x,k_z)$ in (10) be a constant signifying a point source at the origin in (x,z)-space. Let t be very large, meaning that the phase = $\varphi = [-\omega(k_x,k_z) + k_x\,(x/t) + k_z\,(z/t)]t$ in the integration is rapidly alternating with changes in k_x and k_z. Assume the maximum contribution to the integral comes when the phase is stationary. That is, where $\partial\varphi/\partial k_x$ and $\partial\varphi/\partial k_z$ both vanish. Where is the energy in (x,z,t)-space?

1.3 FOUR WIDE-ANGLE MIGRATION METHODS

The four methods of migration of reflection seismic data that are described here may all be found in modern production environments. As a group they are all strong in their ability to handle wide-angle rays. As a group they are all weak in their ability to deal with lateral velocity variation.

Traveltime Depth

Conceptually, the output of a migration program is a picture in the (x,z)-plane. In practice the vertical axis is almost never depth z: it is the *vertical traveltime* τ. In a constant-velocity earth the time and the depth are related by a simple scale factor. The meaning of the scale factor is that the (x,τ)-plane has a vertical exaggeration compared to the (x,z)-plane. In reconnaissance work, the vertical is often exaggerated by about a factor of five. By the time prospects have been narrowed to the point where a drill site is being selected, the vertical exaggeration factor in use is likely to be about unity (no exaggeration).

The traveltime depth τ is usually defined to include both the time for the wave going down and for the wave coming up. The factor of 2 thus introduced quickly disappears into the rock velocity. Recall that zero-offset data sections are generally interpreted in terms of exploding-reflector wave fields. To make the correspondence, the rock velocity is halved for the wave analysis:

$$\tau = \frac{2z}{v_{true}} = \frac{z}{v_{half}} \tag{1}$$

The first task in interpretation of seismic data is to figure out the approximate numerical value of the vertical exaggeration. It is doubtful that it will be printed on the data header for the simple reason that it is not exactly known. This is because the seismic velocity is not exactly known. Furthermore, the velocity usually *increases* with depth, which means that the vertical exaggeration *decreases* with depth. For velocity-stratified media, we may write the time-to-depth conversion formula

$$\tau(z) = \int_0^z \frac{dz}{v(z)} \quad \text{or} \quad \frac{d\tau}{dz} = \frac{1}{v} \tag{2}$$

Hyperbola Summation and Semicircle Superposition Methods

These methods are the most comprehensible of all known methods. Conceptually, at least, they seem to predate the use of computers. Computer implementations of these methods seem to predate the exploding-reflector concept, and they certainly predate the idea of downward extrapolating a wave field with $exp(ik_z z)$ followed by imaging at $t=0$.

First of all, recall the equation for a conic section, a circle in (x,z)-space or a hyperbola in (x,t)-space. With traveltime depth τ, we get

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

$$\frac{x^2}{v^2} + \tau^2 = t^2 \tag{3b}$$

Figure 1 illustrates the circle-superposition method. Taking the data field to contain a few impulse functions, then the output should be a superposition of the appropriate semicircles. Each semicircle denotes the spherical reflector earth model, which would be implied by a dataset with a single pulse. Taking the data field to be a thousand seismograms of a thousand points each, then the output is a superposition of a million semicircles. Since a seismogram has both positive and negative polarities, about half the semicircles will be superposed with negative polarities. The resulting superposition could look like almost anything. Indeed, the semicircles might mutually destroy one another almost everywhere except at one isolated impulse in (x,τ)-space. Should this happen you might rightly suspect that the input data section in (x,t)-space is a Huygens secondary source, namely energy concentrated along a hyperbola. This leads us to the *hyperbola summation* method.

The *hyperbola summation* method of migration is depicted in figure 2. The idea is to create one point in (x,τ)-space at a time, unlike the semicircle method, where each point in (x,τ)-space is built up bit by bit as the one million semicircles are stacked together. To create one fixed point in the output (x,τ)-space, a hyperbola, equation (3b), is imagined set down with its top upon the corresponding position of (x,t)-space. All data values touching the hyperbola are added together to produce a value for the output at the appropriate place in (x,τ)-space. In the same way, all other locations in (x,τ)-space are filled.

The opposite of data processing, building models from data, is constructing synthetic data from models. By means of a slight modification, the above two processing programs can be converted to modeling programs. Instead of *hyperbola summation* or

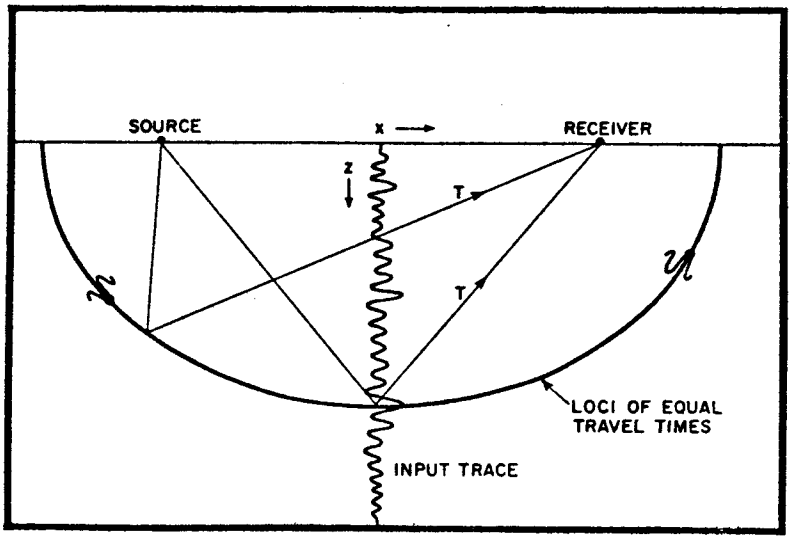


FIG. 1. [from Schneider, W. A., 1971, Developments in seismic data processing and analysis (1968-1970): Geophysics, v. 36, no. 6, p. 1043-1073] The ------- process may be described in numerous ways; however, two very simple and equally valid representations are indicated in figures 1 and 2. Shown here is a representation of the process in terms of what happens to a single input trace plotted in depth (time may also be used) midway between its source and receiver. Each amplitude value of this trace is mapped into the subsurface along a curve representing the loci of points for which the traveltime from source to reflection point to receiver is constant. If the velocity is constant, these curves are ellipses with source and receiver as foci. The picture produced by this operation is simply a wavefront chart modulated by the trace amplitude information. This clearly is not a useful image in itself, but when the map is composited with similar maps from neighboring traces (and common-depth-point traces of different offsets), useful subsurface images are produced by virtue of constructive and destructive interference between wavefronts in the classical Huygens sense. For example, wavefronts from neighboring traces will all intersect on a diffraction source, adding constructively to produce an image of the diffractor in the form of a high-amplitude blob whose (z,x) resolution is controlled by the pulse bandwidth and the horizontal aperture of the array of neighboring traces composited. For a reflecting surface, on the other hand, wavefronts from adjacent traces are tangent to the surface and produce an image of the reflector by constructive interference of overlapping portions of adjacent wavefronts. In subsurface regions devoid of reflecting and scattering bodies, the wavefronts tend to cancel by random addition.

semicircle superposition, one does *hyperbola superposition* or *semicircle summation*. You might wonder whether the processing programs really are the inverse to the modeling programs. You might also wonder whether the two different methods of modeling or processing are equivalent. If they differ, which is better? Clearly some facts which have been glossed over are (1) the angle-dependence of amplitude (obliquity function) of the Huygens waveform, (2) spherical spreading of energy, and (3) the phase shift on the Huygens waveform. Actually, results are reasonably good even

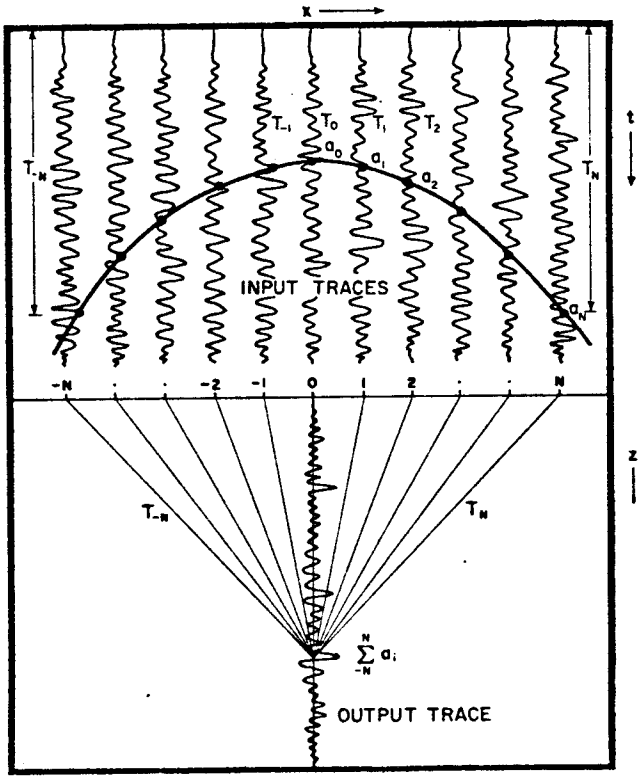


FIG. 2. (from Schneider, 1971 - see figure 1 caption for complete reference) A second description of the ------- process is provided here. The process is represented in terms of how an output trace is developed from an ensemble of input traces, shown as CDP-stacked traces in the upper half of the figure. The output in the lower half reflects how each amplitude value at (x,z) is obtained by summing input amplitudes along the traveltime curve shown. This curve defines a diffraction hyperbola, and if a diffraction source existed in the subsurface at the output point shown, a large amplitude would result. The process also works for reflectors since we may regard a reflector as a continuum of diffracting elements whose individual images merge to produce a smooth continuous boundary.

without these. Since proper migration is an all-pass filter, the inverse to the filter should be simply its time reverse. So crosscorrelating the data (a superposition of hyperbolas) with another hyperbola should result in a filter that is a pretty good inverse to the filter which is a delta function on a hyperbola.

As later methods of migration were developed, the deficiencies of the earlier methods became more clearly understood, and were largely correctable by careful implementation. One advantage of the later methods was that they implemented true all-pass filters. Such migrations preserve the general appearance of the data. This suggests restoration of high frequencies, which tend to be destroyed by hyperbolic integrations. Work by Trorey, Schneider, Hilterman, and possibly others with the Kirchoff diffraction integral suggested quantitative means of bringing hyperbola methods into agreement with other methods, at least for constant velocity. Common terminology nowadays is to refer to any hyperbola or semicircular method as a Kirchoff

method, although, strictly speaking, the Kirchoff integral applies only in the constant-velocity case.

There seems to be no automatic method for migrating data which is spatially aliased (a common problem). The hyperbola-sum-type methods run the risk that the migration operator itself can become spatially aliased. This is a situation to be avoided by means of careful implementation. The first thing to realize is that you should be *integrating* along a hyperbolic trajectory. A summation incorporating only one point per trace is a poor approximation. It is better to incorporate more points, as depicted in figure 3.

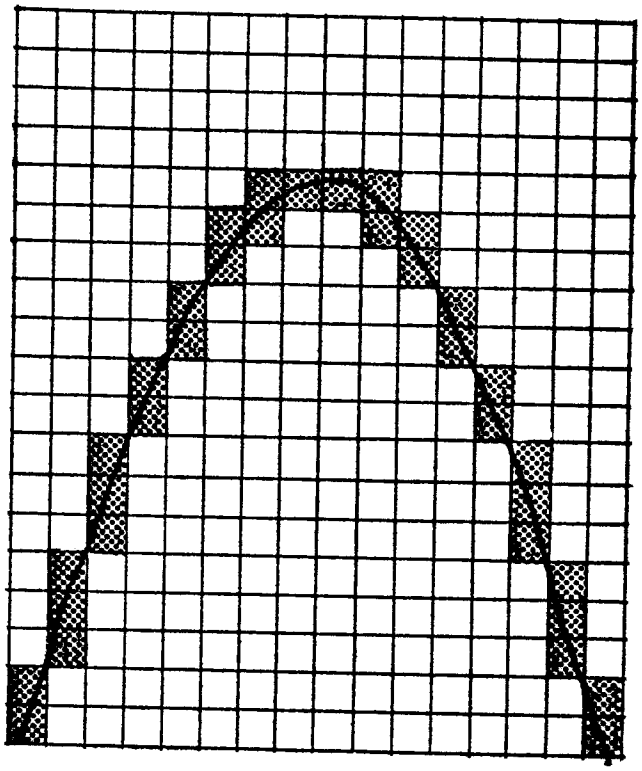


FIG. 3. For a low velocity hyperbola, integration will require more than one point per channel.

The likelihood of getting an aliased operator increases where the hyperbola is steep-sloped. In production examples an aliased operator often stands out on the seafloor reflection where -- although it may be perfectly flat -- it acquires a noisy precursor from the steep-flanked hyperbola of the water path. Figure 4 shows an example.

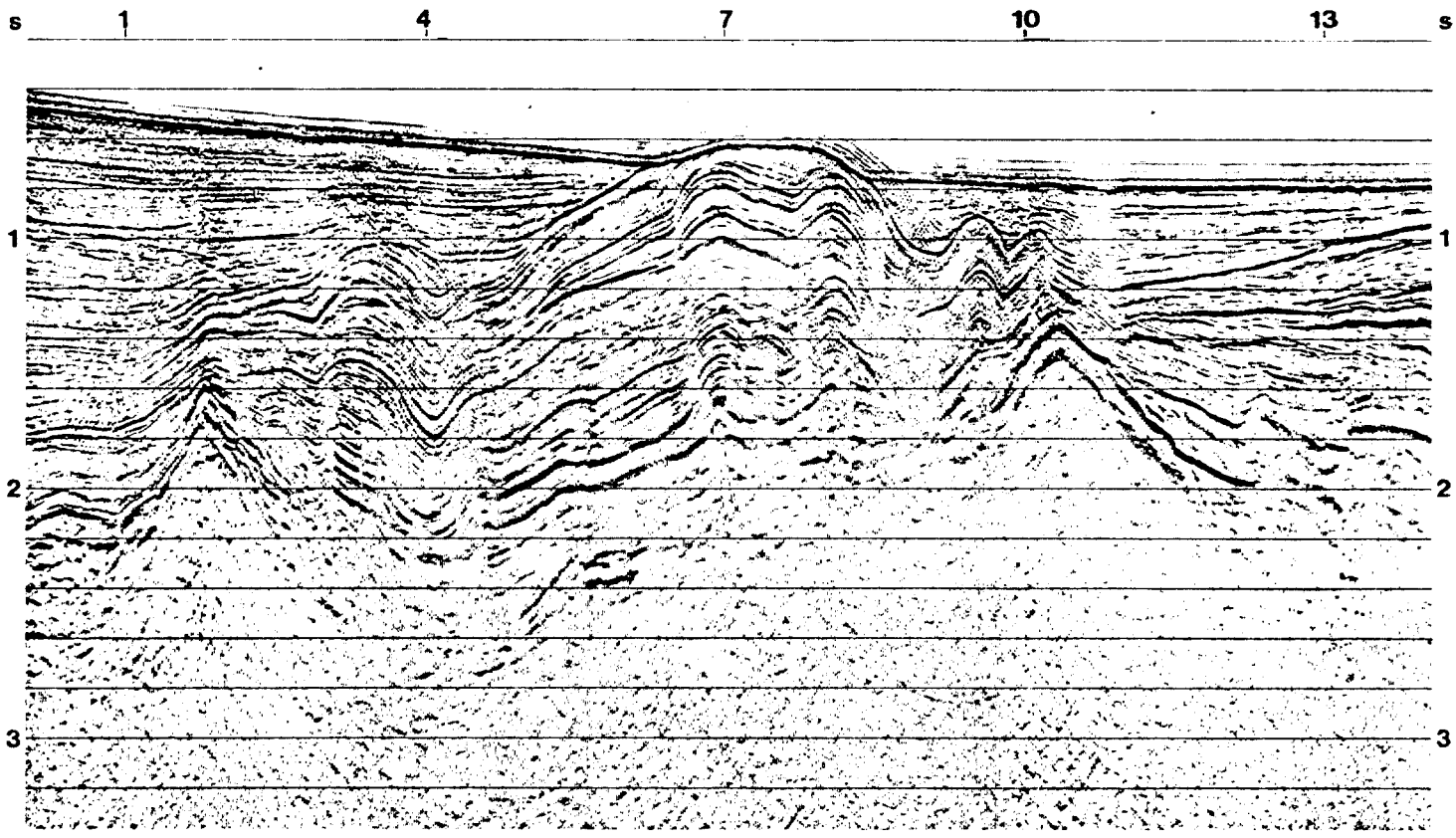


FIG. 4. (from "Wave equation migration: two approaches," 1975, a Western Geophysical brochure) Kirchoff-summation migration. A complex of interfering diffractions has been resolved into a geologically plausible section showing tight folding and faulting. No muting of traces above the water bottom has been used in these examples, so that the high level of "migration noise" generated by the Kirchoff-summation method can be seen.

The Phase-Shift Method

The phase-shift method proceeds straightforwardly by extrapolating downward with $\exp(ik_z z)$ and subsequently evaluating the wave field at $t=0$ (that is, when the reflectors explode). Of all the wide-angle methods it most easily incorporates depth variation in velocity. Even the phase angle and obliquity function are correctly included, automatically. Unlike Kirchoff methods, there is no danger of aliasing the operator. This method is also quite comprehensible. Once the exploding-reflector concept and wave-extrapolation concept became known, a number of workers independently developed phase-shift migration programs (among them Gazdag, Stoffa, who else?).

To start with, you need to do a two-dimensional Fourier transform (2D-FT) of your dataset. Some practical details about 2D-FT are described in a later section. Then you

push the transformed data values, all in the (ω,k_x)-plane, downward to a depth Δz by means of multiplication by

$$e^{ik_z \Delta z} = \exp\left\{ -i \frac{\omega}{v}\left[1 - \left[\frac{vk_x}{\omega}\right]^2\right]^{1/2} \Delta z \right\} \tag{4}$$

Ordinarily the time-sample interval $\Delta\tau$ for the output migrated section will be chosen equal to the time-sample rate of the input data (often 4 ms). Thus, choosing the depth $\Delta z = v\Delta\tau$, the downward-extrapolation operator for a single time unit is

$$\exp\left\{ -i\ \omega\ \Delta\tau\left[1 - \left[\frac{vk_x}{\omega}\right]^2\right]^{1/2} \right\} \tag{5}$$

One hardly ever knows the velocity very precisely, so although the velocity may be increasing fairly steadily with depth, it is often approximated as constant in about 20 layers, rather than slowly changing at each of the thousand or so points on a seismogram. The advantage of velocity being constant in layers is one of economy. Once the square root and the sines and cosines in (5) have been computed, then the complex multiplier (5) can be re-used for all 20 layers.

Next is the task of imaging. At each depth we imagine an inverse Fourier transform followed by selection of the value at $t=0$. Luckily, only the Fourier transform at one point, $t=0$, is needed, so that is all that need be computed. It is especially easy since the value at $t=0$ is merely a summation of each ω frequency component. Finally, inverse Fourier transform k_x to x. The overall migration process may be summarized as follows.

$P(\omega,k_x) = FT[p(t,x)]$
For $\tau = \Delta\tau, 2\Delta\tau, \ldots,$ end of time axis on seismogram
    For all k_x:
        For all ω
            $P(\omega,k_x) = P(\omega,k_x) \exp[-i\ \omega\ \Delta\tau \cos \vartheta(\omega,k)]$
        End loop on ω
    $Image(k_x,\tau) = \sum_{\omega} P$
    End loop on k_x
    $Image(x,\tau) = FT[Image(k_x,\tau)]$
End loop on τ

The Stolt Method

On most computers the Stolt method of migration is the fastest method, by a considerable margin. For many applications, this will be the most important attribute to consider. In a constant-velocity earth the Huygens wave source is treated exactly correctly. Like the other methods, this migration method can be reversed and made into a modeling program. One drawback, a matter of principle, is that the method does not handle depth variation in velocity. Despite the fact that migration effects tend to be in proportion to velocity squared, this drawback has been rumored to be largely offset in practice by the existence of an approximate correction by an axis-stretching procedure. A practical drawback is the periodicity of all the Fourier transforms. In principle this is no problem at all, being solvable by a sufficient surrounding of the data by zeros. A single line sketch of the Stolt method is this:

$$P(x,t) \rightarrow P(k_x,\omega) \rightarrow P\left[k_x, k_z = \left[\frac{\omega^2}{v^2} - k_x^2\right]^{1/2}\right] \rightarrow P(x,z)$$

To see why this works, begin with the input-output relation for downward extrapolation of wave fields:

$$P(\omega,k_x,z) = e^{ik_z z}\, P(\omega,k_x,z=0) \tag{6}$$

Perform a two-dimensional inverse Fourier transform:

$$p(t,x,z) = \int\int e^{ik_x x - i\omega t + ik_z z}\, P(\omega,k_x,0)\, d\omega\, dk_x$$

Apply the idea that the image at (x,z) is the exploding reflector wave at time $t=0$:

$$Image(x,z) = \int\int e^{ik_x x}\, e^{ik_z(\omega,k_x)z}\, P(\omega,k_x,0)\, d\omega\, dk_x \tag{7}$$

Equation (7) states the answer we want, but it is in a very unattractive form. The computational effort implied by (7) is that a two dimensional integration must be done for each and every z-level. The Stolt procedure will be to convert the three dimensional calculation implied by (7) to a single two dimensional fourier transform.

So far we have done nothing to specify that we have an *upgoing* wave instead of a downgoing wave. If ω were always positive, then $+k_z$ would always refer to a downgoing wave and $-k_z$ to an upgoing wave. We need negative frequencies ω as well as positive frequencies in order to describe waves that have real values (not complex). The direction of the wave is defined by the relationship of z and t required to keep the phase constant in the expression $\exp(-i\omega t + ik_z z)$. So the proper description for a

downgoing wave is that the signs of ω and k_z must agree. For an upgoing wave it is the reverse. With this clarification we prepare to change the integration variable in (7) from ω to k_z.

$$\omega = -sgn(k_z)\, v \sqrt{|k_x^2 + k_z^2|} \tag{8a}$$

$$\frac{d\omega}{dk_z} = -sgn(k_z)\, v \frac{k_z}{+\sqrt{k_x^2 + k_z^2}} \tag{8b}$$

$$\frac{d\omega}{dk_z} = \frac{-v\,|k_z|}{+\sqrt{k_x^2 + k_z^2}} \tag{8c}$$

Now we will introduce (8) into (7) including also a minus sign so that the integration on k_z may be taken from minus infinity to plus infinity as was the integration on ω.

$$Image(x,z) = \int\int e^{ik_z z + ik_x x} \left\{ P[\omega(k_x,k_z),k_x,0] \frac{v\,|k_z|}{\sqrt{k_x^2 + k_z^2}} \right\} dk_z\, dk_x \tag{9}$$

Equation (9) states the final result as a two-dimensional inverse Fourier transform. The Stolt migration method is a direct implementation of (9). The steps of the algorithm are:

1. Double Fourier transform field data from $p(t,x,0)$ to $P(\omega,k_x,0)$.
2. Reinterpolate P onto a new mesh so that it is a function of k_x and k_z. Multiply P by the scale factor (which has the interpretation $\cos\vartheta$).
3. Inverse Fourier transform to (x,z)-space.

The major implementation difficulty with the Stolt algorithm is the interpolation. Recall that a delta function late on the time axis is a rapidly oscillating function of frequency. Consider the difficulty of interpolating functions oscillating near the Nyquist frequency. An example by Walter S. Lynn in figure 5 shows that using linear interpolation causes the second half of the time axis to become useless, giving impulse responses which are semicircles downward instead of upward. This means we need an extraordinary amount of zero padding on the time axis. To keep memory costs reasonable, the algorithm can be reorganized.

We need storage space for a long vector, say $u(t)$, (about four times as much as for a typical seismogram). If you have an array processor, that is where this vector belongs. The Stolt algorithm becomes:

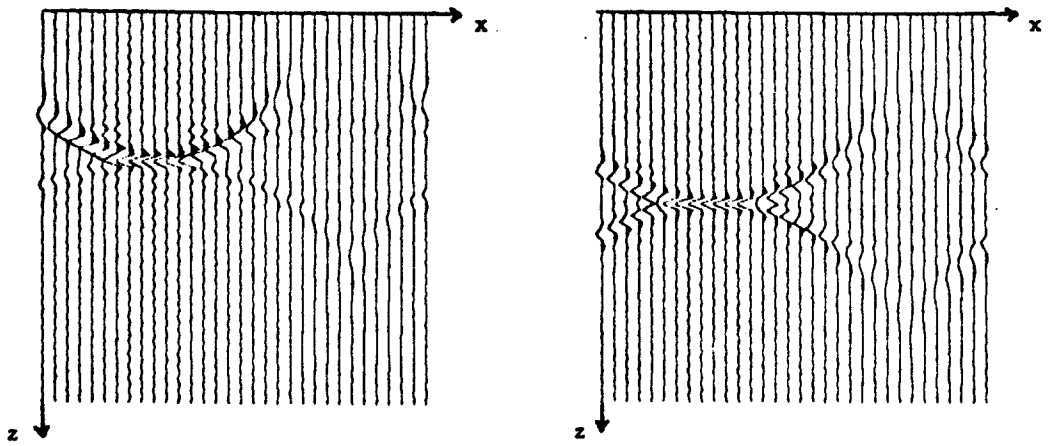


FIG. 5. (Lynn) Impulse response of Stolt method using linear interpolation. The downward-arcing circle is a processing artifact of the linear interpolation in the frequency domain. It worsens rapidly as the delay of the impulse approaches half the length of the time axis. This figure is based on the complex to complex FT, so the maximum frequency is 2π. With the real to complex FT the maximum frequency is π, so hopefully the problem arises twice as late on the time (or depth) axis.

$P(k_x,t) = FT[p(x,t)]$
For all k_x:
$u(t) = P(k_x,t)$
Pad out remaining length of u with zeros.
$U(\omega) = FT[u(t)]$
$$U(k_z) = U[-\sqrt{k_x^2 + k_z^2}] \frac{-v\,|k_z|}{\sqrt{k_x^2 + k_z^2}}$$
$u(z) = FT[U(k_z)]$
$P(k_x,z) = u(z)$
End loop on k_x.
$p(x,z) = FT[P(k_x,z)]$

Even this improved algorithm is not trouble-free. The periodicity in x still requires padding with lots of zeros on x.

Subjective Comparison and Evaluation of Methods

The three basic methods of migration in this section are compared subjectively in table 1.

	Hyperbola Sum Semicircle Sup.	Phase Shift	Stolt
Speed	slow	average	very fast
Memory organization	awkward	good	good
$v(z)$	easily	easily	approximately by stretching
wide angle?	Beware of data alias and operator alias.	Beware of data alias.	Beware of data alias.
Correct phase and obliquity?	possible with some effort for const v	easily for any $v(z)$	yes, for const v
wraparound noise?	no	worst on x, but reasonable	(x,z,t) a problem
$v(x)$	Production programs have serious pitfalls.	no known production program	no known research program

TABLE 1. Subjective comparison of three wide-angle migration methods.

Finally, the perspective of later chapters allows some remarks on the overall quality of the wide-angle methods as a group. Their greatest weakness is their near inability to deal with lateral velocity variation. Their greatest strength, the wide-angle capability, is lessened in value by the weakness of other links in the data collection and processing chain. Namely:

1. Shot-to-geophone offset angles are commonly large but ignored. A CDP stack is not a zero-offset section.

2. Why process to the very wide angles seen in the survey line when even tiny angles perpendicular to the line are being ignored?

3. Data is often insufficiently densely sampled to represent steeply dipping data without aliasing.

4. Accuracy in knowledge of velocity is seldom sufficient to justify processing to wide angles. Recall that the output is a migrated *time* section. Arbitrary velocity error makes no difference when processing horizontal bedding. Velocity error sensitivity increases with angle up to 90 degrees, where the accuracy needed to avoid destructive interference is in the ratio of half a seismic wavelength divided by the traveltime. Commonly this is about 1%, much less than the common accuracy of velocity knowledge. At 45 degrees it is 1.4%.

Exercises

1. Define the computer program for modeling with the phase shift method -- that is, create the surface data $P(x,z=0,t)$ from some exploding reflector distribution $P(x,z,t=0)$.

2. Define the computer program for the inverse to the Stolt algorithm - that is, create synthetic data from a given model.

1.4 THE PHYSICAL BASIS

Previous sections have considered the *geometrical* aspects of wave propagation and how they relate to seismic imaging. Now we consider how the *physical* aspects relate to imaging. The propagation medium has a mass density and compressibility. The waves have a material acceleration vector and a pressure gradient. Static deformation, ground roll, shear, rigidity, dissipation, sedimentary deposition -- how are these related to our image construction?

Derivation of the Acoustic Wave Equation

The acoustic wave equation describes sound waves in a liquid or gas. Another more complicated set of equations describes elastic waves in solids. We begin with the acoustic case. Newton's law of momentum conservation says that a small volume within the gas will accelerate if there is an applied force. The force arises from pressure differences at opposite sides of the small volume. Define

ρ = mass per unit volume of the fluid
u = velocity flow of fluid in the x-direction
w = velocity flow of fluid in the z-direction
P = pressure in the fluid

Newton's law says

$$mass \times acceleration = force = -\, pressure\ gradient$$

$$\rho \frac{\partial u}{\partial t} = -\frac{\partial P}{\partial x} \tag{1a}$$

$$\rho \frac{\partial w}{\partial t} = -\frac{\partial P}{\partial z} \tag{1b}$$

The second physical process to consider is the possibility of energy storage by compression and volume change. If the velocity vector u at $x + \Delta x$ exceeds that at x then the flow is said to be diverging. In other words, the small volume between x and $x + \Delta x$ is expanding. This expansion must lead to a pressure drop. The amount of the pressure drop is in proportion to a property of the fluid called its

incompressibility K. In one dimension the equation is

$$\textit{pressure decrease} \;=\; (\textit{incompressibility}) \times (\textit{divergence of velocity})$$

$$-\frac{\partial P}{\partial t} \;=\; K\,\frac{\partial u}{\partial x} \tag{2a}$$

In two dimensions it is

$$-\frac{\partial P}{\partial t} \;=\; K\left[\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z}\right] \tag{2b}$$

To make the one-dimensional wave equation from (1a) and (2a), first divide (1a) by ρ and take its x-derivative:

$$\frac{\partial}{\partial x}\,\frac{\partial}{\partial t}\,u \;=\; -\,\frac{\partial}{\partial x}\,\frac{1}{\rho}\,\frac{\partial P}{\partial x} \tag{3}$$

Second, the time-derivative of (2) will be taken. In solid-earth sciences we are fortunate that the material does not change during the course of the experiments. This means that K is a constant function of time

$$\frac{\partial^2 P}{\partial t^2} \;=\; -\,K\,\frac{\partial}{\partial t}\,\frac{\partial}{\partial x}\,u \tag{4}$$

Inserting (3) into (4), the one-dimensional scalar wave equation appears

$$\frac{\partial^2 P}{\partial t^2} \;=\; K\,\frac{\partial}{\partial x}\,\frac{1}{\rho}\,\frac{\partial P}{\partial x} \tag{5a}$$

In two space dimensions, the scalar wave equation is

$$\frac{\partial^2 P}{\partial t^2} \;=\; K\left[\frac{\partial}{\partial x}\,\frac{1}{\rho}\,\frac{\partial}{\partial x} + \frac{\partial}{\partial z}\,\frac{1}{\rho}\,\frac{\partial}{\partial z}\right]P \tag{5b}$$

Expand derivatives.

$$\left[\frac{\partial^2}{\partial t^2} - \frac{K}{\rho}\left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}\right]\right]P \;=\; -\,\frac{K}{\rho^2}\left[\frac{\partial \rho}{\partial x}\,\frac{\partial P}{\partial x} + \frac{\partial \rho}{\partial z}\,\frac{\partial P}{\partial z}\right] \tag{6}$$

Let us compare the wave equation (6) with our earlier assertion that a wave field could be represented by a complex exponential, namely

$$P \;=\; \exp(-i\omega t + ik_x x + ik_z z) \tag{7}$$

Inserting (7) as a trial solution into (6) we can cancel the complex exponential, getting

$$-\omega^2 + \frac{K}{\rho}(k_x^2 + k_z^2) = -\frac{K}{\rho^2}\left[\frac{\partial \rho}{\partial x} ik_x + \frac{\partial \rho}{\partial z} ik_z\right] \tag{8}$$

In the substitution of (7) into (6) it was assumed that the three frequencies (ω, k_x, k_z) were independent of space, but this will not be consistent with (8) unless we have the material properties ρ and K independent of space. This is an example of the general fact that Fourier methods fail on equations with space-variable coefficients. We will return at length to the issue of space-variable material properties. Taking the material properties to be constant, equation (8) becomes the *dispersion relation of the two-dimensional scalar wave equation*

$$\frac{\omega^2}{K/\rho} = k_x^2 + k_z^2 \tag{9}$$

Earlier an equation like (9) was developed considering only the geometrical behavior of waves. In that development we had wave velocity squared where K/ρ stands in equation (9). Thus there is the the association

$$v^2 = \frac{K}{\rho} \tag{10}$$

Reflections and the High Frequency Limit

It is well known that the contact between two different materials can cause acoustic reflections. A material contact is defined to be a place where either K or ρ changes by a spatial step function. In one dimension either $\partial K/\partial x$ or $\partial \rho/\partial x$ or both would be infinite at a point. It is well known that either can cause a reflection. So it is perhaps a little surprising that the density derivative is explicitly found in (6) but the incompressibility derivative is not explicitly there. This means that if we drop the density gradients in (6) we will not lose all possible reflections. Dropping the terms would slightly simplify further analysis. Since constant density is a reasonable enough physical situation to consider, the terms are often dropped.

There are also some well-known mathematical circumstances under which the first-order terms may be ignored. Suppose the two media in contact gradually blend into one another so that $\partial \rho/\partial x$ is less than infinity. Suppose we are interested only in high frequencies, that is, large values of ω, k_x, k_z. As the frequencies tend to infinity in (8), the second-order terms get larger much faster than the gradient terms which are first order. In that mathematical limit the gradient terms may be neglected.

Of course, practical situations may arise for which these terms need to be included. It is usually not difficult to do so. I believe the terms are generally neglected for the same reason we often write equations in two dimensions instead of three. The extension is usually possible but it is rarely required.

Evanescence and Ground Roll

Completing the physical derivation of the dispersion relation

$$k_x^2 + k_z^2 = \frac{\omega^2}{v^2} \tag{11}$$

we can now have a new respect for it. It carries more meaning than could have been anticipated on the basis of the earlier geometrical derivation. We originally thought of it merely in terms of $\sin^2\vartheta + \cos^2\vartheta = 1$ where $\sin\vartheta = vk_x/\omega$. Originally we attached no meaning to $\sin\vartheta$ exceeding unity, in other words, to vk_x exceeding ω. Now we can. In fact there was a hidden ambiguity in two of the migration methods we considered. Since data could be an arbitrary function in the (t,x)-plane, then its Fourier transform could be an arbitrary function in the (ω,k_x)-plane. So we certainly could find energy with an angle sine greater than one. What should we do with it?

In the most extreme case, $\omega = 0$, k_x is real, and $k_z = \pm ik_x$. This means that the depth-dependence of the physical solution is a growing or a damped exponential. In the elastic-wave situation it describes the deformation of ground under a parked airplane. Only if the airplane can move faster than the speed of sound in the earth will there be a wave radiated into the earth. Moving at a subsonic speed the deformation is said to be *quasi- static*.

Perhaps a better physical description is to imagine a sinusoidally corrugated sheet. One sometimes sees such metallic sheet used for roofs or garage doorways. The wavelength of the corrugation fixes k_x. Moving such a sheet past your ear at velocity V you would hear a frequency of oscillation equal to Vk_x, regardless of whether V is larger or smaller than the speed of sound in air. But the sound you hear would get weaker exponentially with distance from the sheet unless the plate moved very fast, $V > v$, in which case the moving sheet would be radiating sound to great distances.

What should a migration program do with such low-velocity energy? Theoretically it should be exponentially damped. Quantitatively the damping is so rapid that the offending region of (ω,k_x)-space may as well be replaced by zeros. You might not expect to find a lot of energy in your data at these low velocities. Actually, with land

data, the low-velocity soil layer often creates conditions under which there can be a lot of energy at these low velocities, so much that this is a major problem. This energy is called "ground roll." It could be defined as any energy moving with a velocity less than that of the velocity of the rocks at the depths of interest. Such energy is unwanted noise since its exponential decay effectively prevents it from being influenced by deep objects of interest.

Shear Waves and Lithology

In earthquake seismology and in laboratory measurement there are two clearly observed velocities. The faster velocity is a pressure wave (p-wave) and the slower velocity (by about a factor of two) is a shear wave (s-wave). Theory, field data, and laboratory measurement are in excellent agreement.

I have been watching for shear waves in the reflection seismic data that are recorded for petroleum prospecting. I have never found them. They should show up routinely in velocity surveys. In stacked sections they should appear to be similar to multiple reflections. Shear waves should have good diagnostic value in exploration. But the likelihood of seeing shear waves in conventional data seems to be so remote that most interpreters have given up trying.

To compound this puzzle, reflection data are usually of better quality than either earthquake data or laboratory data.

Shear waves can be seen in reflection data when special transverse generating and recording equipment is set up for the task. The puzzle is why p-to-s conversions are not routinely observed with the standard operating arrangement. Theory predicts that every p-wave which hits an interface at an angle -- the usual case at nonzero offset -- should generate an s-wave which is nicely separated from the p arrival. So why don't we see converted waves in conventional data? Four reasons may be offered:

1. In marine data there would have to be a conversion to p for the water path.
2. In land data the soil seems to be very absorptive of shear.
3. Explosive sources tend to generate more p than s.
4. Vertical component recorders tend to see p better than s.

I find the above four reasons are not convincing because they do not scale amplitudes by very large values. We record a wide dynamic range in a wide variety of environments. We frequently display data with automatic gain control (AGC). Weakened amplitude appears to be insufficient cause for the failure of observation.

As described theoretically, experimentally, and in observational earthquake seismology, pressure to shear wave conversion is associated with planar contact between two dissimilar materials. With earthquakes, the earth's surface and the interface between the liquid core are solid mantle are two reliable producers of converted waves. Other boundaries are more problematic in their effect. In sedimentary geology there seems to be an abundance of flat layers. But are they really flat enough in their shear characteristics?

Strange as it may seem, even for p-waves there is not good agreement about the true nature of seismic reflections. Physicists tend to think of the reflections as being caused by the interface between rock types, as a sand-to-shale contact. Many geologists, particularly a group known as seismic stratigraphers, have a different concept. They have studied thousands of miles of reflection data along with well logs. They believe a reflection marks a constant geological time horizon. They assert that a long, continuous reflector could represent terrigenous deposition on one end and marine deposition on the other end with a variety of rock types going from one extreme to the other.

Looking at sedimentary rocks you can gain an appreciation for both points of view. Generally speaking, most reservoir rocks are sandstones and most sands are deposited near the mouth of a river where the water velocity is no longer sufficient to move them. The sands are never laid down in flat layers. They deposit along the terminus of the sand bars found at the river mouth. Commonly they deposit along a slope of 25 degrees or so, as indicated in figure 1.

The shales (more fine-grained material, like dirt) deposit in deeper water and tend to be somewhat more layered. Specific locations of sand deposition change with the passing of storms and seasons leaving a wood-grain-like appearance in the rock.

The delta itself is a complicated, ever-changing arrangement of channels and bars. Throughout time the delta moves up and down the coastline. At any one time it seems to be moving seaward as the deposits are made, but subsequent settling, compression, or raising of sea level can cause it to move landward.

Sand is important because its porosity and permeability enable oil to accumulate in traps. Shale is important because it contains the products of former life on earth, and their hydrocarbons. These escape to the sands, but often not to the earth's surface, due to covering impermeable shales. The acoustic properties of sands and shales often overlap, though there is a slight tendency for shales to have a lower velocity. We geophysicists on the surface see with seismic wavelengths ($\approx$ 30 meters), the final interbedded three-dimensional mixture of sands and shales.

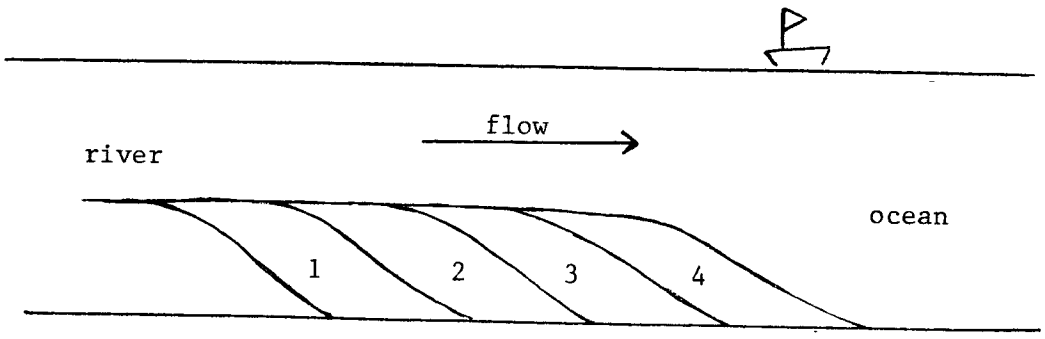


FIG. 1. Sands (petroleum reservoir rocks) deposit on fairly steep slopes where rivers run into the ocean.

Although there is not a great deal of transversely generated shear-wave data available, some indications are that such data give a rather different picture of the earth, one which is difficult to relate to the p-wave picture.

Could it be that the nature of depositional rocks is such that $p-s$ conversions do not occur? Or perhaps they do not occur in the organized way that $p-$reflections do, thus contributing only to the chaos?

The failure of simple elastic equations to describe the observed absence of $p-s$ coupling is not the first failure of Newtonian theory. Even more remarkable and well-documented is the situation with the seismic dissipation parameter Q. It is observed to be nearly a frequency-independent constant. Newtonian viscosity predicts inverse frequency dependence. Many simple models fail by predicting a resonance absorption. The heterogeneity of the rock, at all scales, seems to be an essential attribute to a successful theory (such as that in the section on impedance).

Finally, departure of theory from experiment is not cause for embarrassment: it is indicative of opportunities for discovery. I should not assert that conversions do not occur. Perhaps they do, and with more careful analysis we will observe them. They may provide the best prospecting tool yet!

Philosophy of Inverse Problems

Physical processes are often simulated with computers in much the way that they occur in nature. The machine memory is used as a map of physical space, and time evolves in the calculation as it does in the simulated world. A nice thing about solving problems this way is that there is never any question about the uniqueness of the solution. Errors of initial data and model discretization do not tend to have a catastrophic effect. Exploration geophysicists, however, rarely solve problems of this type. Instead of having (x,z)-space in the computer memory and letting t evolve, we usually have (x,t)-space in memory and we are extrapolating in depth z. This is basically our business, taking information (data) at the earth's surface and attempting to extrapolate to information at depth. Stable time evolution in nature provides no "existence proof" that what we want to do is reasonable, stable, or even possible. When it isn't, we must redefine our goals.

Commonly the time-evolution problems are called *forward problems* and the depth-extrapolation problems are called *inverse problems*. In a *forward problem*, such as with acoustic waves, it is clear what you need and what you can get. You need the density, $\rho(x,z)$ and incompressibility $K(x,z)$ and you need to know the initial source of disturbance. You can get the wave field everywhere at later times but you usually only want it at the earth's surface for comparison to some data. In the *inverse problem* you have the waves seen at the surface, the source specification, and you would like to determine the material properties $\rho(x,z)$ and $K(x,z)$. What has been learned from experience is that the observations cannot give reasonable estimates of images or maps of ρ and K.

Luckily, it has been discovered that certain functions of ρ and K can be reliably determined and mapped. Speaking in the sense of production data processing, not research, it can be said that the density ρ cannot be mapped at all so it may as well be set equal to a constant or to some laboratory-determined function of velocity. What we can determine is the velocity, namely $v^2 = K/\rho$. Even this is not determined in the unambiguous sense we might like. We see velocity through two different, virtually non-overlapping windows. The most valuable window is the one on high spatial frequencies. Reflections may be imagined as coming from cracks within rocks of constant velocity. The imagined cracks are small delta functions added to the constant-velocity field. The cracks give the velocity field a high spatial-frequency content which is sensed through the typical 10 to 100 Hz spectral window of good-quality field data. In reality we probably do not see cracks. We see the sedimentary configurations described earlier.

The other window on the velocity function will be described in more detail in Chapter 4. It involves study of traveltime with respect to variation of shot-to-geophone offset. With this second window we consider ourselves to be fortunate when we can discern sixteen independent velocity measurements on a 4-second reflection time axis. So this window goes from zero to about 2 Hz.

In routine processing and in this book the two views of the velocity are treated as independent entities. The function in the big window will be called the reflectivity, $c(x,z)$, and the function in the small window will be called (confusingly) the velocity v, usually $v(z)$.

In mathematics the solution to an inverse problem has come to mean the "determination" of material properties from wave fields. Often this is achieved by means of a "convergent sequence." We are considerably less precise (more inclusive) about what we mean by "determination." In Chapters 1-3 of this book we "determine" reflectors by the exploding-reflection concept. In Chapter 4 we incorporate shot-to-geophone offset, and we "determine" reflectivity $c(x,z)$ and velocity $v(z)$ by means of a buried-experiment concept. In Chapter 5 we suppress multiple reflections and find "true" amplitudes of reflections by an imaging concept which seeks to have the upcoming wave vanish before the onset of the downgoing wave. Other imaging concepts seem likely to result from future processing schemes. It might be possible to show that some of our "determinations" coincide with those of mathematicians, but that is not our goal.

1.5 THE SINGLE-SQUARE-ROOT EQUATION

The function of the *single- square- root equation* is to extrapolate waves down into the earth. This equation seems to be the basis for all valid migration methods. It is a partial differential equation of transcendental type. It matches the ordinary scalar wave equation in some respects but departs from it in others.

Snell Waves

It is natural to begin studies of waves by using equations describing plane waves in media of constant velocity. However, in reflection seismic surveys the velocity contrast between shallowest and deepest reflectors ordinarily exceeds a factor of two. So in the analysis of field data, depth variation of velocity is almost always included. Rays bend and wavefronts curve. A *Snell wave* will be defined to be the natural extension of the idea of a *plane wave* for a medium with depth-variable velocity $v(z)$. For a stratified medium $v=v(z)$ we will define the *single- square- root equation*, that is, we will define its form in stratified media $v=v(z)$ to be a differential equation with a Snell wave as its solution. The diagram in figure 1 illustrates the geometry of a downgoing Snell wave.

From the diagram we see the differentials

$$\frac{dt}{dx} = \frac{\sin \vartheta}{v} \tag{1a}$$

$$\frac{dt}{dz} = \frac{\cos \vartheta}{v} \tag{1b}$$

The two equations define two (inverse) speeds. First is a horizontal speed, measured along the earth's surface, called the *horizontal phase velocity*. Second is a vertical speed, measurable in a borehole, called the *vertical phase velocity*. Notice that both these speeds *exceed* the velocity v of wave propagation in the medium. Reference to wave *fronts* gives speeds larger than v and reference to *rays* gives speeds smaller.

A Snell wave could be generated by a source function moving horizontally along the earth's surface at a speed dx/dt. By symmetry (the medium is invariant under lateral shift), the horizontal speed dx/dt seen at any other depth, say z_0, must equal

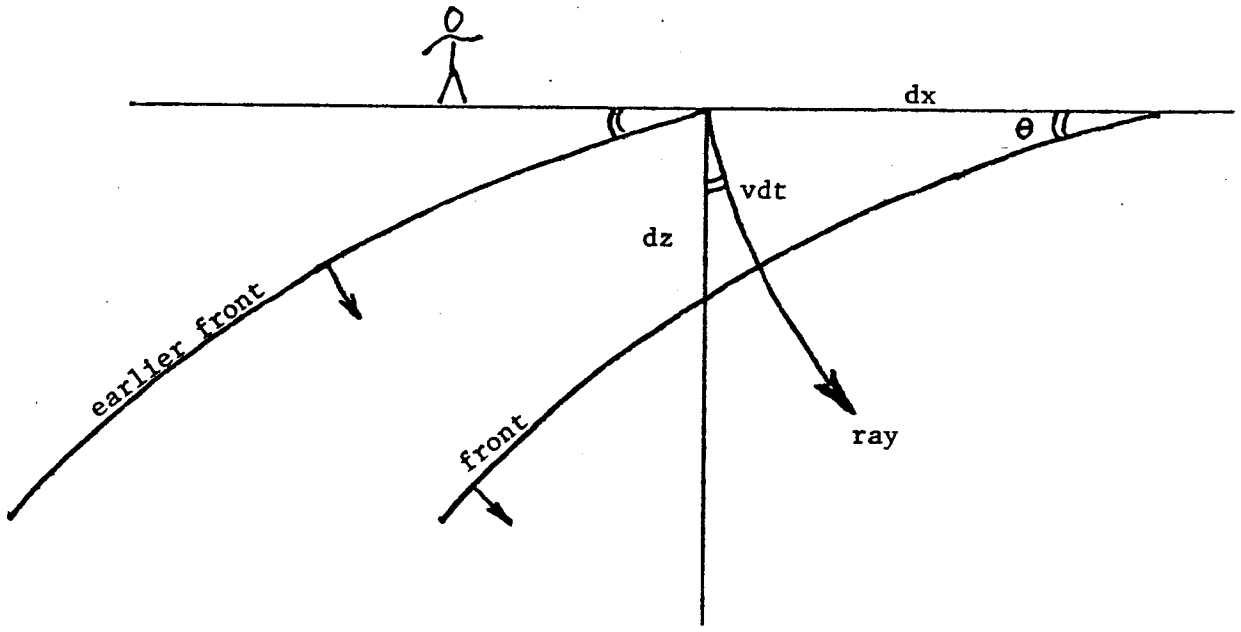


FIG. 1. Downgoing fronts and rays in stratified medium $v(z)$

that seen at the surface $z=0$. This depth-invariant parameter, or rather its inverse dt/dx, is known as Snell's parameter. It is noteworthy that Snell's parameter, denoted p, is directly observable at the surface, whereas neither v nor ϑ are directly observable. Since p is not only observable, but constant in depth, it is customary to use it to eliminate ϑ from equation (1):

$$\frac{dt}{dx} = \frac{\sin\vartheta}{v} = p \tag{2a}$$

$$\frac{dt}{dz} = \frac{\cos\vartheta}{v} = \left[\frac{1}{v(z)^2} - p^2\right]^{1/2} \tag{2b}$$

With plane waves we speak of their angle of propagation. With Snell waves we speak instead of their Snell parameter p.

Taking the Snell wave to go through the origin at time zero, an expression for the arrival time of the Snell wave at any other location is given by

$$t(x,z) = \frac{\sin\vartheta}{v}\,x + \int_0^z \frac{\cos\vartheta}{v}\,dz \tag{3a}$$

$$= px + \int_0^z \left[\frac{1}{v(z)^2} - p^2\right]^{1/2} dz \tag{3b}$$

The validity of (3b) is readily checked by computing $\partial t/\partial x$ and $\partial t/\partial z$, then comparing with (2).

An arbitrary waveform $f(t)$ may be carried by the Snell wave. Using (3) to define a delay time to the location (x,z), the delayed wave $f(t-t_0)$ is

$$\textit{Snell wave field} = f\left\{ t - px - \int_0^z \left[\frac{1}{v(z)^2} - p^2 \right]^{1/2} dz \right\} \tag{4}$$

Time-Shifting Equations

An important task is to predict the wave field inside the earth given the waveform at the surface. For a downgoing plane wave this can be done by the time-shifting partial differential equation

$$\frac{\partial P}{\partial z} = -\frac{1}{v}\frac{\partial P}{\partial t} \tag{5}$$

as may be readily verified by substituting the trial solutions

$$P = f\left[t - \frac{z}{v} \right] \quad \textit{for constant } v \tag{6}$$

or

$$P = f\left[t - \int_0^z \frac{dz}{v(z)} \right] \quad \textit{for } v(z) \tag{7}$$

Heeding some important restrictions, this also works for non-vertically incident waves with the partial differential equation

$$\frac{\partial P}{\partial z} = -\frac{dt}{dz}\frac{\partial P}{\partial t} \tag{8}$$

which has the solution

$$P = f\left(t - px - \int_0^z \frac{dt}{dz}\, dz \right) \tag{9}$$

In interpreting (8) and (9) we recall that dz/dt is the apparent velocity in a borehole. The partial derivative of wave field P with respect to depth z is taken at constant x, i.e., the wave is extrapolated down the borehole. The idea that downward extrapolation can be achieved by merely time-shifting clearly holds only for situations in which a

single Snell wave is present, that is, the *same* arbitrary time function must be seen at all locations.

Substitution from (1) also enables us to rewrite (8) in the various forms

$$\frac{\partial P}{\partial z} = -\frac{\cos\vartheta}{v}\frac{\partial P}{\partial t} \tag{10a}$$

$$\frac{\partial P}{\partial z} = -\left[\frac{1}{v(z)^2} - p^2\right]^{1/2}\frac{\partial P}{\partial t} \tag{10b}$$

$$\frac{\partial P}{\partial z} = -\left[\frac{1}{v(z)^2} - \left[\frac{dt}{dx}\right]^2\right]^{1/2}\frac{\partial P}{\partial t} \tag{10c}$$

Since $dt/dx=p$ can be measured along the surface of the earth, it seems that equation (10c), along with an assumed velocity $v(z)$ and some observed data $P(x,t)$, would enable us to determine $\partial P/\partial z$, which is the necessary first step of downward continuation. But we must not forget that we are dealing by assumption with a single Snell wave and not a superposition of several Snell waves. Superposition of different waveforms on different Snell paths will cause different time functions to be seen at different places. Then a mere time-shift will not achieve downward continuation. Luckily, a complicated wave field that is variable from place to place may be decomposed, by mathematical techniques not yet discussed, into many Snell waves, each of which can be downward extrapolated with the differential equation (10) or its solution (9). One such decomposition technique is Fourier analysis.

Fourier analyzing the function $f(x,t,z=0)$ seen on the earth's surface, we will need the Fourier kernel $exp(-i\omega t + ik_x x)$. Moving on the earth's surface at an inverse speed of $dt/dx=k_x/\omega$, the phase of the Fourier kernel, hence the kernel itself, remains constant. Only those sinusoidal components which move at the same speed as the Snell wave can have a non-zero correlation with it. So if the disturbance is a single Snell wave then all Fourier components vanish except for those which satisfy $p=k_x/\omega$. You should memorize these basic relations:

$$\boxed{\frac{dt}{dx} = \frac{\sin\vartheta}{v} = p = \frac{k_x}{\omega}} \tag{11}$$

In seismology it is very likely that the appearance of any square-root function comes

from use of (11) to make a cosine.

Utilization of this Fourier-domain interpretation of Snell's parameter p enables us to write the square-root equation (10) in an even more useful form. But first we must re-express the square root equation in the Fourier domain, which we do by replacing the $\partial/\partial t$ operator in (10) by $-i\omega$. The result is

$$\frac{\partial P}{\partial z} = + \frac{i\omega}{v(z)} \left\{ 1 - \left[\frac{v(z) k_x}{\omega} \right]^2 \right\}^{1/2} P \tag{12}$$

At present it is equivalent to specify either the differential equation (12) or its solution (9) with f being the complex exponential, namely

$$P = exp\, -i\omega \left\{ t - px - \int_0^z \left[\frac{1}{v(z)^2} - \frac{k_x^2}{\omega^2} \right]^{1/2} dz \right\} \tag{13}$$

Later, when we consider lateral velocity variation $v(x)$, the solution (13) becomes wrong, whereas the differential equation (12) indicates a useful computational algorithm. But before we go on to lateral velocity gradients we need to consider more carefully the situation with vertical velocity gradients.

Velocity Gradients

Inserting the Snell wave-field expression into the scalar wave equation, one discovers that our definition of a Snell wave does not satisfy the scalar wave equation. The discrepancy arises only in the presence of velocity gradients. In other words, if there is a shallow constant velocity v_1 and a deep constant velocity v_2, the equation is satisfied everywhere except where v_1 changes to v_2. Solutions to the scalar wave equation must show amplitude changes across an interface, because of transmission coefficients. Our definition of a Snell wave is a wave of constant amplitude with depth. This raises the question of whether our definition is wrong and whether we could better use the scalar wave equation, rather than the single square root equation, to downward extrapolate wave fields. Several compelling reasons will later be given as to why the scalar wave equation should *not* be used for wave-field extrapolation. Actually, the amplitudes can be correctly incorporated in the single- square-root equation, as indicated by the exercises. The reason why it is customary not to do so is probably the same reason that density gradients are commonly ignored. They add to the clutter in writing equations while their contribution to better results, namely more correct amplitudes and possible small phase shifts, has marginal utility. Indeed, if they are

included then other deeper questions should also be considered, such as the question of why we use the acoustic equation rather than various other forms of scalar elastic equations.

Wave Equation + Exploding Reflectors $\neq$ Migration

The scalar wave equation will not be advocated for wave extrapolation for three reasons: (1) initial conditions, (2) unstable solutions, and (3) multiple reflections.

First, the scalar wave equation has a second depth z derivative. This means that two boundary conditions are required on the z-axis. Since we observe at $z=0$, it seems natural that these should be knowledge of P and dP/dz at $z=0$. But we don't observe dP/dz. Even supposing this difficulty could somehow be overcome, what about the second difficulty?

Theoretically, the scalar wave equation contains evanescent waves, waves with low apparent surface velocity, which will be either exponentially decaying or exponentially growing with depth. When obtaining numerical solutions, it is likely to be difficult to reject the growing solutions. Even if we are somehow able to achieve stability, the third difficulty remains.

Where dissimilar media are in contact, that is, when the velocity is space-variable, then there are reflected waves as well as transmitted waves at the interface. The multiple reflections in the exploding-reflector model differ from those on zero-offset sections, so the multiples generated will be incorrect for our application. Furthermore, the multiples are likely to be seriously damaging because they will be generated at high- amplitude signals and will migrate to positions where they interfere with low-amplitude signals. Errors in transmission coefficients are much less serious because transmission errors on high-amplitude events remain with the high-amplitude events as small scaling errors.

Exercises

1. Devise a mathematical expression for a plane wave which is a delta function of time with propagation angle of 15 degrees from the vertical z-axis in the plus z direction. Express the result in the domain of

(a) (t,x,z)

(b) (ω,x,z)

(c) (ω,k_x,z)

(d) (ω,p,z)

2. Find an amplitude function $A(z)$ which, when multiplied by f in equation (9), causes it to be a solution to the scalar wave equation in stratified media $v(z)$.

3. Modify the single-square-root equation so that the result of exercise 1 is the solution. Hint:

$$\frac{\partial P}{\partial z} \quad \textit{can be changed to} \quad \frac{1}{B(z)} \frac{\partial}{\partial z} B(z)P(z)$$

1.6 MASTERY OF 2-D FOURIER TECHNIQUES

Here is a collection of helpful tips for those who will be involved in implementations of migration methods.

Signs and Scales in Fourier Transforms[1]

We have been doing Fourier transform on t, x, and z coordinates. On each coordinate there is a choice for the sign convention of Fourier transform. Electrical engineers have chosen one convention and physicists another. Both have good reasons for their choices. Our circumstances more closely resemble those of physicists, so we will use their convention, which for the *inverse* Fourier transformation is

$$p(t,x,z) = \int\int\int e^{-i\omega t + ik_x x + ik_z z}\, P(\omega,k_x,k_z)\, d\omega\, dk_x\, dk_z \qquad (1)$$

The limits on the integrations and the scale factor differ in the continuous case from the discrete case. We rarely do the transforms analytically in either case. The extra notation required for limits and scales usually clutters rather than clarifies a discussion. So we omit them altogether except when they play a useful role in the discussion.

The sign convention is more important. Because there are so many space axes (later, midpoint and offset space axes are introduced and transformed as well), it is worthwhile to establish a good sign convention. The approach of "changing the signs around until it works" can be perplexed by the number of possible permutations of the number of signs to be changed. There are good reasons for the sign conventions chosen by physicists, and once the reasons are known, it is easy to remember the conventions.

First of all, waves should, by convention, move in the positive direction on the space axis. This is especially evident on work for which the space axis is a radius. Atoms, like geophysical sources, always radiate from a point to infinity, not the other way around. So we always choose waves moving positively on any space axis. In equation (1) this means that the sign on the spatial frequencies must be opposite to the sign on the temporal frequency.

---

[1] Adapted from SEP 16, p 361-2

We could still change the sign of all three frequencies. But there is a reason not to do so. There are more space axes than time axes. We'll see the fewest number of minus signs and the fewest sign changes if we take the spatial gradient $\partial/\partial x$, $\partial/\partial z$, etc. to be associated with the positive k-vector, i.e. with ik_x, ik_z, etc. Of course, we are left associating the time derivative with $-i\omega$.

This sign convention leaves us inconsistent with the practice of electrical engineers who rarely work with space axes and naturally enough have chosen to associate $\partial/\partial t$ with $+i\omega$. The only good reason I know to adopt the engineering choice is that we compute with an array processor built and microcoded by engineers who naturally use their own sign convention. It doesn't make any difference on the programs which transform complex-valued, time functions to complex-valued, frequency functions, because then the sign convention is under user control. But it does make a difference on the program which converts real time functions to complex frequency functions. With the Stolt algorithm it is common to transform space first. Then the engineering convention is an advantage. The way to live in both worlds is to imagine that the frequencies produced by the program do not range from 0 to $+\pi$ as the description says, but that they range from 0 to $-\pi$. You can always take the complex conjugate of the transform which will swap the sign of the ω-axis.

How to Transpose a Big Matrix[2]

It is a lucky thing that we can easily transpose very large matrices. This is what makes wave-equation seismic data processing reasonable on a mini-computer. The transpose algorithm is simple but tricky. I shall begin, therefore, by describing a card trick. I have in my hands a deck of cards from which I have removed the nines, tens, and face cards. Let a, b, c, and d denote hearts, spades, clubs, and diamonds. Also, I have arranged these cards in the order (let ace be denoted by a one):

$$1a \quad 1b \quad 1c \quad 1d \quad 2a \quad 2b \quad 2c \quad 2d \quad 3a \quad \cdots \quad 8d$$

Now I deal the cards face up alternately, one onto pile A and one onto pile B. You see

$$\textit{Pile } A\text{:} \quad 1a \quad 1c \quad 2a \quad 2c \quad 3a \quad 3c \quad \cdots \quad 8a \quad 8c$$

$$\textit{Pile } B\text{:} \quad 1b \quad 1d \quad 2b \quad 2d \quad 3b \quad 3d \quad \cdots \quad 8b \quad 8d$$

Next I place pile A on top of (in front of) pile B, and again deal the cards out alternately

[2] Adapted from SEP 11, p 211-2

into pile A' and pile B'. You see

$$\textit{Pile } A': \quad 1a \quad 2a \quad 3a \quad \cdots \quad 8a \quad 1b \quad 2b \quad \cdots \quad 8b$$

$$\textit{Pile } B': \quad 1c \quad 2c \quad 3c \quad \cdots \quad 8c \quad 1d \quad 2d \quad \cdots \quad 8d$$

Now I place pile A' on top of B'. We started with all the aces together, the twos together, etc. Now we have all the hearts together, the spades together, etc. So you see that in just two deals of the cards, I have transposed the deck. We never spread the cards out all over the table because we have never had randomly to access the deck. We just made sequential passes over it. In principle, this algorithm transposes a matrix requiring four magnetic tapes but almost no core memory.

Now let us try the inverse transpose. You see that it takes me *three* deals of the cards rather than the *two* deals it took for the original transpose. This is because the deck has $2^2 = 4$ suits and $2^3 = 8$ numbers. Actually, there is another algorithm which will allow me to do the inverse tranpose in only two passes rather than three. You just do everything backwards. Start with piles A' and B'. Then create pile A by alternately selecting cards one from pile A' and one from pile B'. Likewise construct pile B. Then do it one more time. This algorithm is the "merge" algorithm.

So we see that the matrix transpose of a matrix of size $2^n \times 2^m$ can be done by the lesser of n or m passes over the data.

A variety of generalizations are possible. With 4 card piles we could work out techniques for matrices of dimension 4^n. This would decrease the number of passes but increase the required number of tape drives. Likewise it turns out that arbitrary order may be factored into primes, etc. But this takes us too far afield.

If you wish to minimize the number of passes over the data, you turn out to maximize the number of tapes. In reality you probably won't be using real tapes when you are transposing. But you are likely to be simulating tape operations on a large disk memory. Then the number of "tapes" you choose to use will be controlled by the ratio of the speed of random transfers compared to the speed of sequential transfers.

Rocca's 2-D F.T. Without Transposing[3]

The most direct method of two dimensional Fourier transformation in a computer is the repetitive application of a one dimensional Fourier transform method. The

[3] Adapted from SEP 15, p 247-50

easiest part is the "fast" direction, that is, if the data matrix is stored by columns -- as in the Fortran language -- then the column transforms are a trivial exercise in repetitive use of a one dimensional program. Conceptually the easiest way to handle the transformation over the row direction is to transpose the matrix, transform each column, and transpose back. Fabio Rocca suggested another means of Fourier transformation over the row index which is not much more difficult conceptually but which seems to be a more practical approach. The essential idea of Rocca's method is this:

The data matrix can be regarded as a row vector whose entries are columns. Taking the "fast" index to range down the column, the columns may be transformed by one-dimensional transforms either before or after the row operations are done. To do the row operations you modify an ordinary one-dimensional Fourier transform program by replacing each scalar add or multiply operation by the same operation upon every element in the corresponding column.

As an illustration of the idea, a row Fourier transformation was based on the one dimensional Fourier transformation program found in FGDP. The appropriate generalization is listed in table 1. For real to complex Fourier transforms, you should beware of the assumption that real and imaginary parts are stored contiguously. This assumption is true for the column index, but not the row index.

As a practical matter, there are two important special cases. The easiest case is when the computer memory is large enough to hold the data matrix. Some machines have a virtual memory operating system. The program of Table 1 should be satisfactory in such an environment even if the data matrix exceeds the core memory. This is because the inner loops all run down the column index.

The less easy case is when the data matrix exceeds the memory size. Then the matrix is stored on disk and two columns at a time are read into memory for processing. A pseudo-Fortran program of Robert W. Clayton is found in SEP-15, pages 247-250. Some comments of his on this subject follow:

"...the inner loop (or twiddle step) of the one-dimensional FT subroutine FORK (FGDP, p. 12) is:

```
        do 50 i=m, nx, istep
        ctemp= cw*cx(i+n)
        cx(i+n)= cx(i)-ctemp
50      cx(i)= cx(i)+ctemp
```

In the two-dimensional FFT, this step would appear as (in pseudo-Fortran):

```
      subroutine rowcc(n1,n2,cx,sign2,scale)
c     Try Rocca's row Fourier Transform.
c     sign2 should be +1. or -1. It is the sign of i.
      complex cx(n1,n2),cmplx,cw,cdel
      do 05 i1=1,n1
      do 05 i2=1,n2
05    cx(i1,i2)=cx(i1,i2)*scale
      j=1
      do 30 i=1,n2
      if(i.gt.j) go to 10
      call twid1(n1,cx(1,i),cx(1,j))
10    m=n2/2
20    if(j.le.m) go to 30
      j=j-m
      m=m/2
      if(m.ge.1) go to 20
30    j=j+m
      lstep=1
40    istep=2*lstep
      cw=1.
      arg=sign2*3.14159265/lstep
      cdel=cmplx(cos(arg),sin(arg))
      do 60 m=1,lstep
      do 50 i=m,n2,istep
50    call twid2(n1,cw,cx(1,i),cx(1,i+lstep))
60    cw=cw*cdel
      lstep=istep
      if(lstep.lt.n2) go to 40
      return
      end

      subroutine twid1(n,cx,cy)
      complex cx(n),cy(n),ct
      do 10 i=1,n
      ct=cx(i)
      cx(i)=cy(i)
10    cy(i)=ct
      return
      end

      subroutine twid2(n,cw,cx,cy)
c     If you feel like optimizing, this is the place.
      complex cx(n),cy(n),ctemp,cw
      do 10 i=1,n
      ctemp=cw*cy(i)
      cy(i)=cx(i)-ctemp
10    cx(i)=cx(i)+ctemp
      return
      end
```

TABLE 1. Program for Fourier transformation over row index.

```
      do 50 i=m, nx, istep
      call read("read (i)-th column of matrix into vector a")
      call read("read (i+n)-th column of matrix into vector  b")
      do 45 k=1, ny
      ctemp= cw*b(k)
      b(k)= a(k)-ctemp
45    a(k)= a(k)+ctemp
      call write("write vector  a  into  (i-th) column of matrix")
50    call write("write vector  b  into  (i+n)-th column of matrix")
```

Here "read" and "write" are input/output routines to move data in and out of core.

Some further considerations for production programs follow: The row FFT's require $\log_2 N$ passes over the matrix, where N is the number of columns in the matrix. If four or more column vectors can be held in core at one time, then the I/O operations can be considerably reduced by "unfolding" the innermost loop. That is, instead of basing the twiddle step on a Fourier transform of length 2, it could be based on lengths 4, 8, or 16, etc. This would eliminate the I/O operations of the intermediate steps.

The method can be adapted for use with an array processor. The twiddle step which contains all the floating point operations could be micro-coded for the array processor. This would leave only the I/O operations for the host computer."

Inverse Slant Stack[4]

The processes of *slant stack* and *inverse slant stack* find some application in seismology, though not as much as in medical imaging. The process has 2-D Fourier interpretation. Here we will just touch on it, not master it. Slant stack is defined by

$$g(t,p) = \int f(t+ph,h)\, dh \tag{2}$$

The inverse may be represented in the Fourier domain (Thorson, SEP-14, p. 81-85) by

$$f(t,h) = \int \{ \int [G(\omega,p) e^{-i\omega ph} \, |\omega| \,] \, e^{i\omega t} \, d\omega \} \, dp \tag{3}$$

The product of three functions in the ω-domain is a convolution in the time domain. The three functions are the slant stack g, a time shift by ph, and an $|\omega|$ filter. Letting g' denote the filtered slant stack, the inversion becomes

$$f(t,h) = \int g'(t - ph)\, dp \tag{4}$$

In practice, the result is not as clean as 2-D FT, and other techniques, such as optimization, may be used. (See Kjartansson, SEP-20, p. 12-24.)

---

[4] Adapted from SEP 14, p 81-85.