1.1 Exploding Reflectors

The basic equipment for reflection seismic prospecting 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 the source is activated, and the echoes are recorded nearby. The sound source and receiver have almost no directional tuning capability because the frequencies that penetrate the earth have 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 shows two wave-propagation situations. The first is realistic field sounding. The second is a thought experiment in which 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.

Notice in the figure that the raypaths in the field-recording case seem to be the same as those in the exploding-reflector case. It is a great conceptual advantage to imagine that the two wavefields, the observed and the hypothetical, are indeed the same. If they are the same, then the many thousands of experiments that have really been done can be ignored, and attention can be focused on the one hypothetical experiment. One obvious difference between the two cases 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. Travel time in field experiments could be divided by two. In practice, the data of the field experiments (two-way time) is analyzed

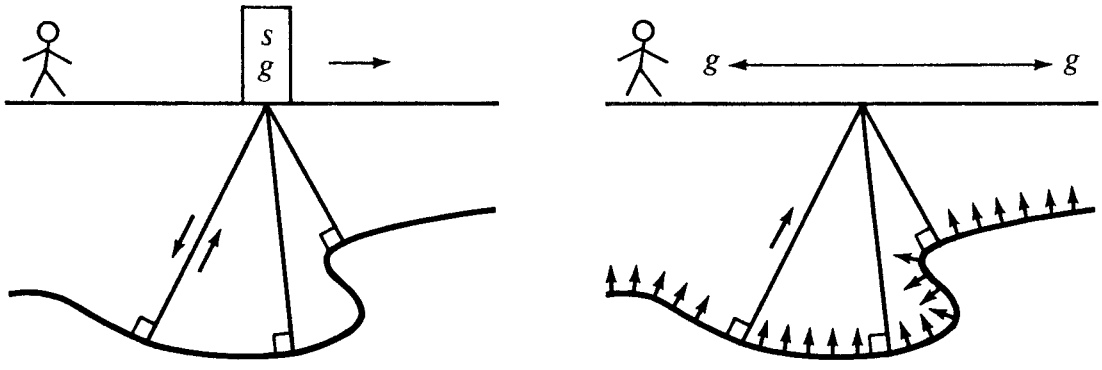


FIG. 1.1-1. Echoes collected with a source-receiver pair moved to all points on the earth's surface (left) and the "exploding-reflectors" conceptual model (right).

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 ocean waves move slowly enough to be seen. Imagine a long harbor barrier parallel to the beach with a small entrance in the barrier for the passage of ships. This is shown 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 the wavefront in the harbor 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.

Linearity is a property of all low-amplitude waves (not those foamy, breaking waves near the shore). This means that two gaps in the harbor barrier make two semicircular wavefronts. Where the circles cross, the wave heights combine by simple linear addition. It is interesting to think of a barrier with many holes. In the limiting case of very many holes, the barrier disappears, being nothing but one gap alongside another. Semicircular wavefronts combine to make only the incident plane wave. Hyperbolas do the same. Figure 3 shows hyperbolas increasing in density from left to right. All those waves at nonvertical angles must somehow combine with one another to extinguish all evidence of anything but the plane wave.

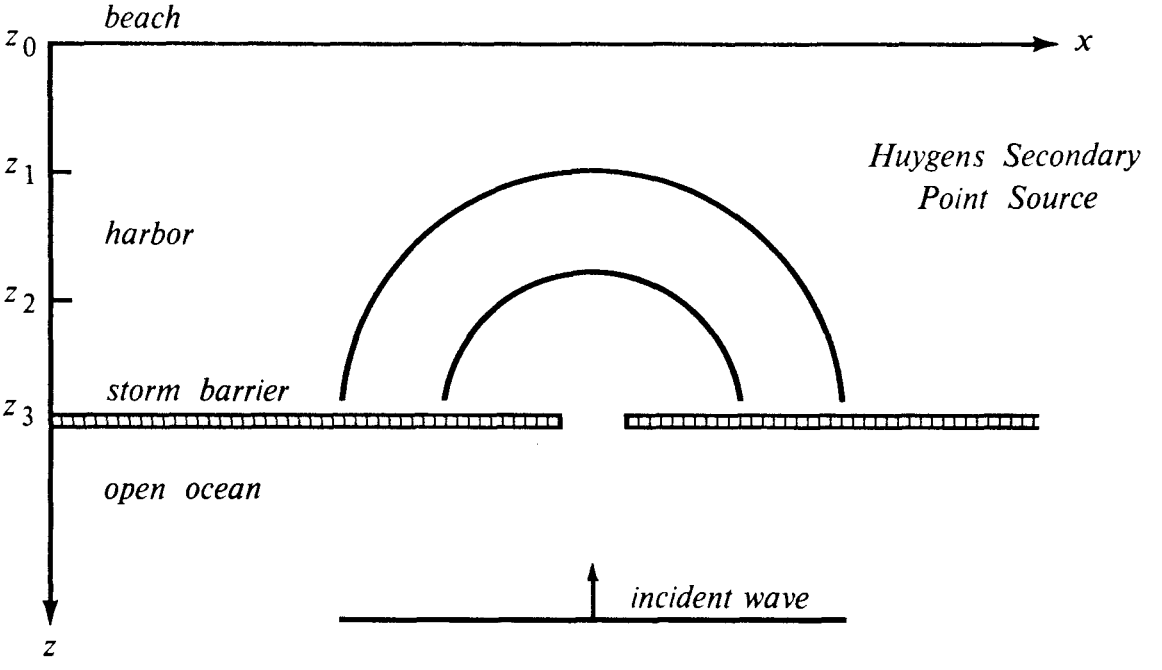


FIG. 1.1-2. Waves going through a gap in a barrier have semicircular wavefronts (if the wavelength is long compared to the gap size).

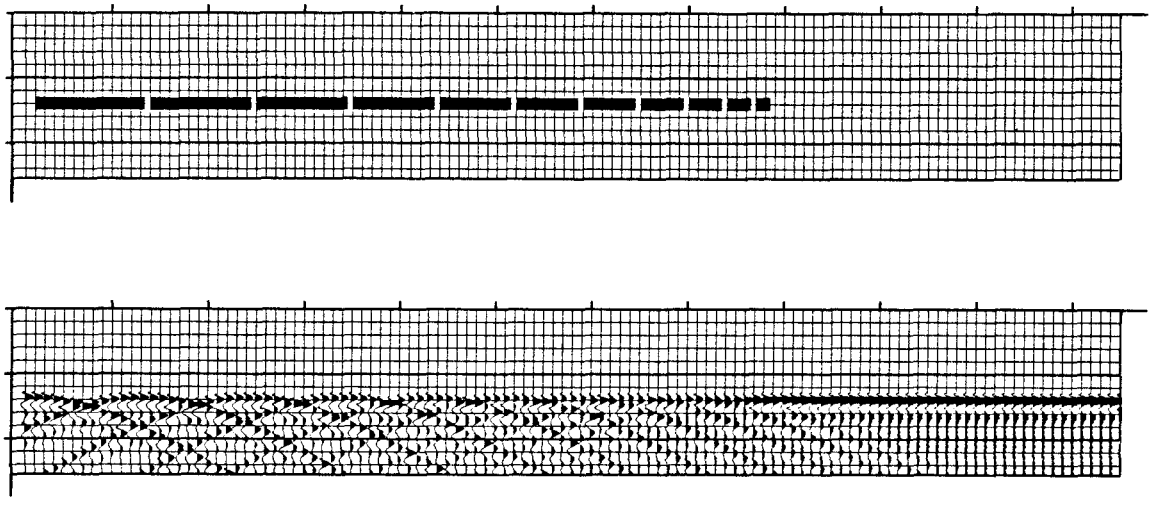


FIG. 1.1-3. A barrier with many holes (top). Waves, (x, t)-space, seen beyond the barrier (bottom).

A Cartesian coordinate system has been superimposed on the ocean surface with x going along the beach and z measuring the distance from shore. For the analogy with reflection seismology, people are confined to the beach (the earth's surface) where they make measurements of wave height as a function of x and t. From this data they can make inferences about the existence of gaps in the barrier out in the (x, z)-plane. Figure 4a shows the arrival time at the beach of a wave from the ocean through a gap. The earliest arrival occurs nearest the gap. What mathematical expression determines the shape of the arrival curve seen in the (x, t)-plane?

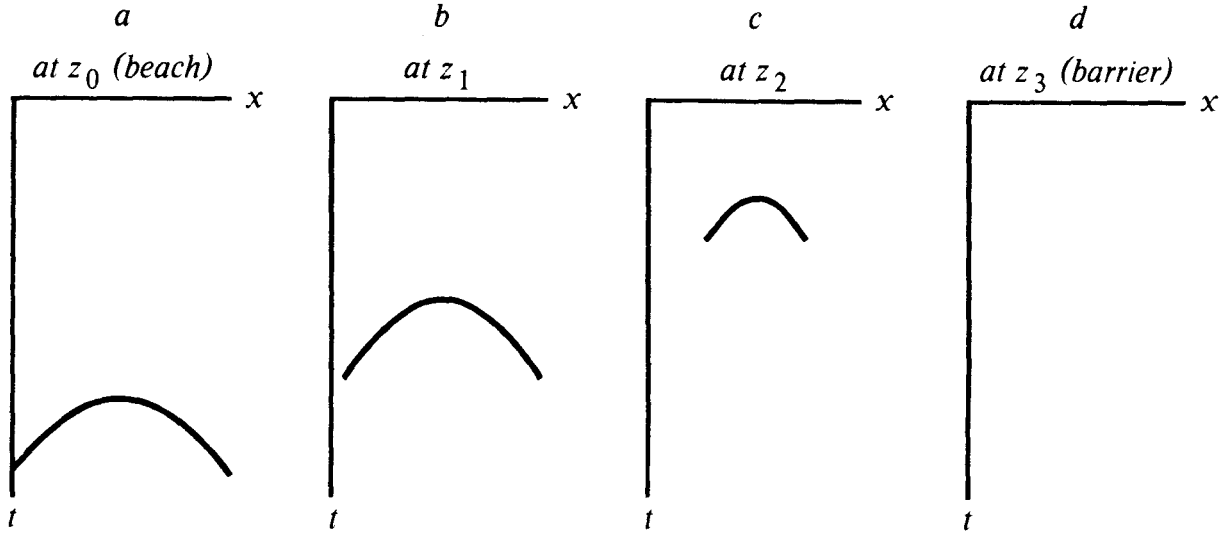


FIG. 1.1-4. The left frame shows the hyperbolic wave arrival time seen at the beach. Frames to the right show arrivals at increasing distances out in the water. (The x-axis is compressed from figure 2). (Gonzalez)

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

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

Considering t to be a constant, i.e. taking a snapshot, equation (1) is that of a circle. Considering z to be a constant, it is an equation in the (x, t)-plane for a hyperbola. Considered in the (t, x, z)-volume, equation (1) is that of a cone. Slices at various values of t show circles of various sizes. Slices of various values of z show various hyperbolas. Figure 4 shows four hyperbolas. The first is the observation made at the beach $z_0 = 0$. The second is a

hypothetical set of observations at some distance z_1 out in the water. The third set of observations is at z_2, an even greater distance from the beach. The fourth set of observations is at z_3, 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 that turns nearly 90° 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 it is presumed —incorrectly— that the speed of water waves is a constant independent of water depth).

If the original incident wave was a positive pulse, then the Huygens secondary source must consist of both positive and negative polarities to enable the destructive interference of all but the plane wave. So the Huygens waveform has a phase shift. In the next section, mathematical expressions will be found for the Huygens secondary source. Another phenomenon, well known to boaters, is that the largest amplitude of the Huygens semicircle is in the direction pointing straight towards shore. The amplitude drops to zero for waves moving parallel to the shore. In optics this amplitude dropoff with angle is called the *obliquity factor*.

Migration Defined

A dictionary gives many definitions for the word *run*. They are related, but they are distinct. The word *migration* in geophysical prospecting likewise has about four related but distinct 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 imagine the four frames of figure 4 being taken from a movie. During the movie, the depth z changes beginning at the beach (the earth's surface) and 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 make another superposition, keeping the hyperbola tops all in the same place. Mathematically, the time t axis is replaced by a so-called *retarded* time axis $t'=t+z/v$, shown in figure 5b. The second, more precise definition of *migration* is the motion of an event in (x, t')-space as z changes. After removing the vertical shift, the residual motion is mainly a shape change. By this definition hyperbola tops, or horizontal layers, don't migrate.

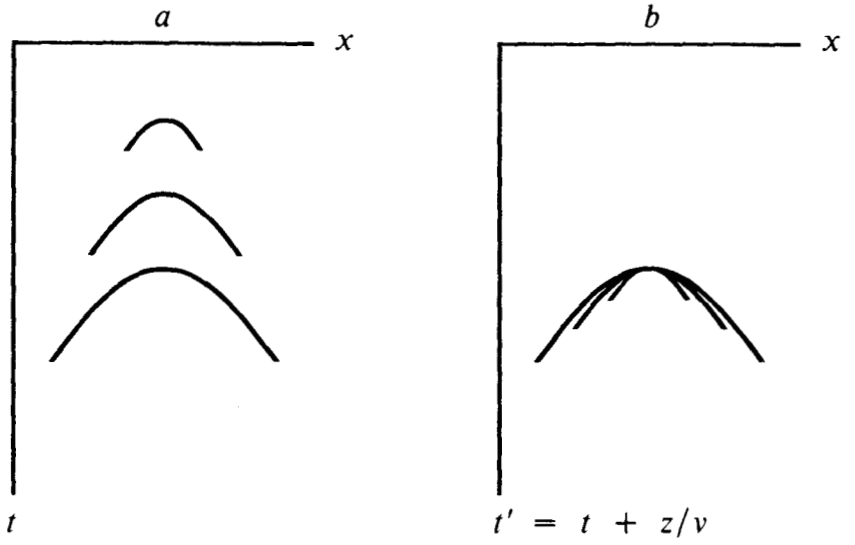


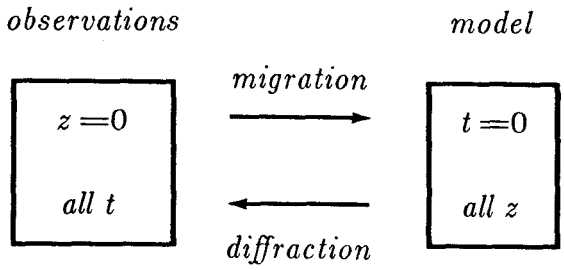
FIG. 1.1-5. Left shows a superposition of the hyperbolas of figure 4. At the right the superposition incorporates a shift, called retardation $t' = t + z/v$, to keep the hyperbola tops together. (Gonzalez)

The hyperbolas in figure 5 really extend to infinity, but the drawing cuts each one off at a time equal $\sqrt{2}$ times its earliest arrival. Thus the hyperbolas shown depict only rays moving within 45° of the vertical. It is good to remember this, that the ratio of first arrival time on a hyperbola to any other arrival time gives the cosine of the angle of propagation. The cutoff on each hyperbola is a ray at 45°. Notice that the end points of the hyperbolas on the drawing can be connected by a straight line. Also, the slope at the end of each hyperbola is the same. For any wavefront, the angle of the wave is $\tan \theta = dx/dz$ in physical space. For any seismic event, the slope $v\ dt/dx$ is $\sin \theta$, as you can see by considering a wavefront intercepting the earth's surface at angle θ. So, energy moving on a straight line in physical (x, z)-space migrates along a straight line in data (x, t)-space. As z increases, the energy of all angles comes together to a focus. The focus is the exploding reflector. It is the gap in the barrier. This third definition of migration is that it is the process that somehow pushes observational data — wave height as a function of x and t — from the beach to the barrier. The third definition stresses not so much the motion itself, but the transformation from the beginning point to the ending point.

To go further, a more general example is needed than the storm barrier example. The barrier example is confined to making Huygens sources only at some particular z. Sources are needed at other depths as well. Then, given a wave-extrapolation process to move data to increasing z values, exploding-reflector images are constructed with

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

The fourth definition of migration also incorporates the definition of *diffraction* as the opposite of migration.



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

Another aspect of the use of the word *migration* arises in Chapter 3, where the horizontal coordinate can be either shot-to-geophone midpoint y, or 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 *focusing* or *velocity analysis.*

An Impulse in the Data

The 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 no echoes are recorded except at one particular location, and at that location only one echo is recorded. 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.

It is unlikely that the processes of nature have created many spherical mirrors inside the earth. But when we look at processed geophysical data, we often see spherical mirrors. Obviously, such input data contains impulses that are not consistent with the wave-propagation theory being explained here.

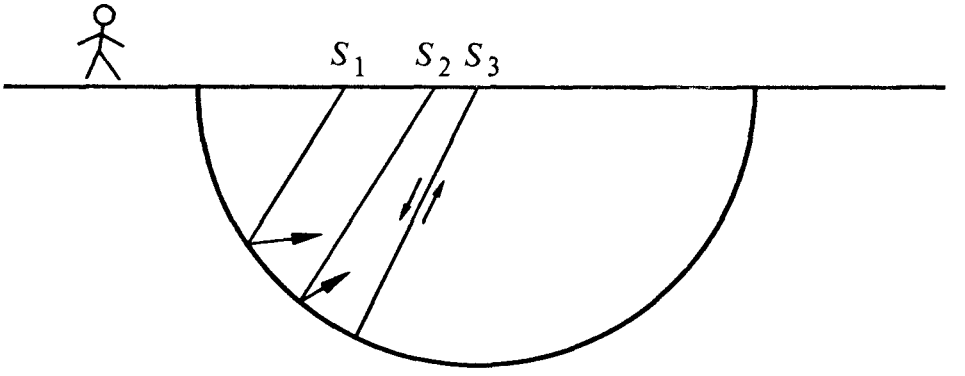


FIG. 1.1-6. When the seismic source S is at the exact center of a semicircular mirror, then, and only then, will an echo return to the geophone at the source. This semicircular reflector is the logical consequence of a dataset where one echo is found at only one place on the earth.

This illustrates why petroleum prospectors study reflection seismic data processing, even though they personally plan to write no processing programs. The raw data is too complex to comprehend. The processed data gives an earth model, but its reliability is difficult to know. You may never plan to build an automobile, but when you drive alone far out into the desert, you should know as much as you can about automobiles.

Hand Migration

Given a seismic event at (x_0, t_0) with a slope $p = dt/dx$, let us determine its position (x_m, t_m) after migration. Consider a planar wavefront at angle θ to the earth's surface traveling a distance dx in a time dt. Assuming a velocity v we have the wave angle in terms of measurable quantities.

$$\sin\theta = \frac{v\ dt}{dx} = p\ v \tag{3}$$

The vertical travel path is less than the angled path by

$$t_m = t_0 \cos\theta = t_0 \sqrt{1 - p^2 v^2} \tag{4a}$$

A travel time t_0 and a horizontal component of velocity $v \sin\theta$ gives the lateral location after migration:

$$x_m = x_0 - t_0\, v \sin\theta = x_0 - t_0\, p\ v^2 \tag{4b}$$

Consideration of a hyperbola migrating towards its apex shows why (4b) contains a minus sign. Equations (4a) and (4b) are the basic equations for

manual migration of reflection seismic data. They tell you where the point migrates, but they do not tell you how the slope p will change.

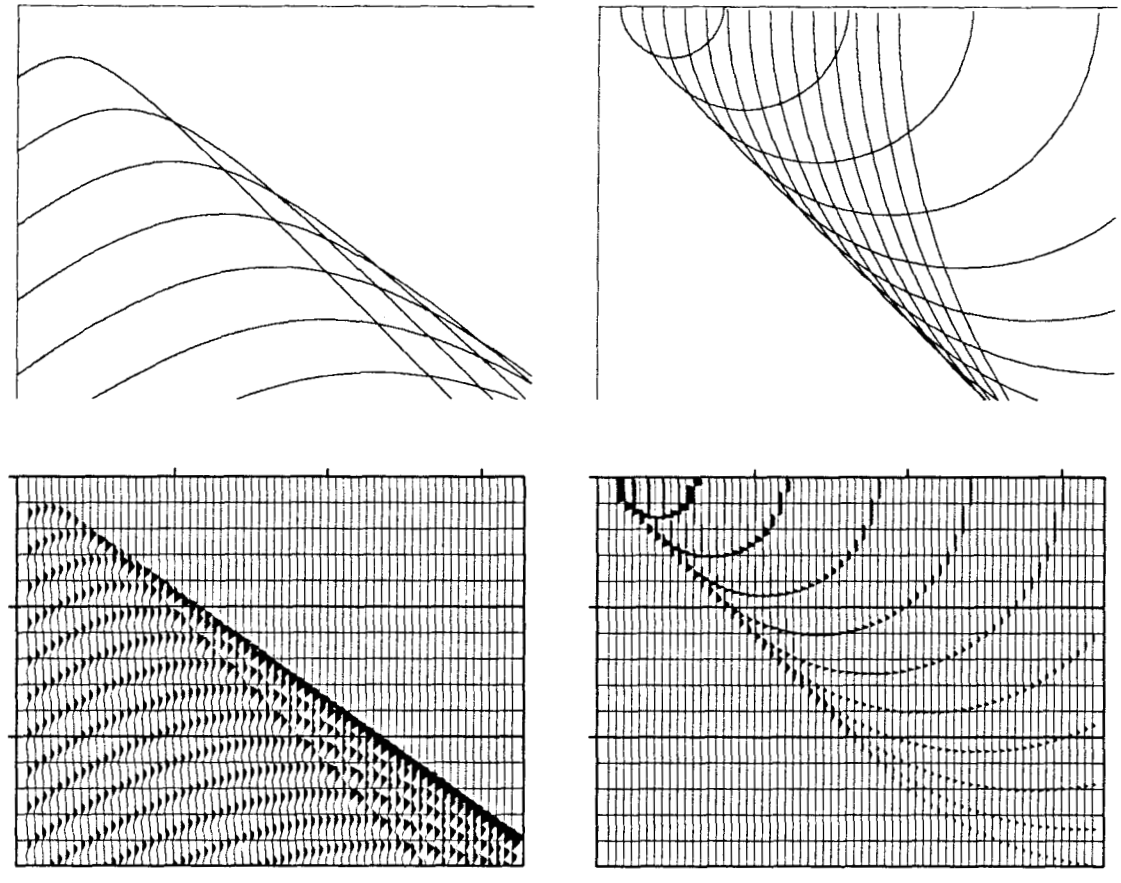


FIG. 1.1-7. Left is a superposition of many hyperbolas. The top of each hyperbola lies along a straight line. That line is like a reflector, but instead of using a continuous line, it is a sequence of points. Constructive interference gives an apparent reflection off to the side.

Right shows a superposition of semicircles. The bottom of each semicircle lies along a line that could be the line of an observed plane wave. Instead the plane wave is broken into point arrivals, each being interpreted as coming from a semicircular mirror. Adding the mirrors yields a more steeply dipping reflector.

Reflector Steepening

Consider a vertical wall, a limiting case of a dipping bed. Its reflections, the asymptotes of a hyperbola, have a nonvertical steepness. This establishes that migration increases the apparent steepness of dipping beds. I use the

words *apparent steepness* because it is the slope as seen in the (x, t)-plane that has steepened. Migration really produces its output in z. but z/v is often overlain on t to create a *migrated time section.* When we say a hyperbola migrates to its apex, we are of course thinking of the migrated time section. Let us determine the steepening as a function of angle.

Consider a point $(x_{0+}, t_{0+}) = x_0 + \Delta, t_0 + p\,\Delta$ neighboring the original point (x_0, t_0). By equation (4), this neighbor migrates to

$$t_{m+} = (t_0 + p\ \Delta)\sqrt{1 - p^2 v^2} \tag{5a}$$

$$x_{m+} = x_0 + \Delta - (t_0 + p\ \Delta)\, p\ v^2 \tag{5b}$$

Now we compute the stepout p_m of the migrated event

$$p_m = \frac{dt_{m+}}{dx_{m+}} = \frac{dt_{m+}/d\,\Delta}{dx_{m+}/d\,\Delta}$$

$$p_m = \frac{p\ \sqrt{1 - p^2 v^2}}{1 - p^2 v^2} = \frac{p}{\sqrt{1 - p^2 v^2}} = \frac{\tan\theta}{v} \tag{6}$$

So slopes on migrated time sections, like slopes in Cartesian space, imply tangents of angles while slopes on unmigrated time sections imply sines.

It may seem paradoxical that dipping beds change slope on migration whereas flanks of hyperbolas do not change slope during downward continuation. One reason is that migration is downward continuation *plus* imaging (selecting $t = 0$). Another reason is that a hyperbola is a special event that comes from a single source at a single depth whereas a dipping bed is a superposition of point sources from different depths. Figure 7 shows how points making up a line reflector diffract to a line reflection, and how points making up a line reflection migrate to a line reflector.

Limitations of the Exploding-Reflector Concept

The exploding-reflector concept is a powerful and fortunate analogy. For people who spend their time working entirely on data interpretation rather than on processing, the exploding-reflector concept is more than a vital crutch. It's the only means of transportation! But for those of us who work on data processing, the exploding-reflector concept has a 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, there is not one hydrophone, but hundreds, which are strung out in a cable towed behind the ship. The recording cable 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 will be considered until Chapter 3.

Furthermore, even at zero offset, the exploding-reflector concept is not quantitatively correct. For the moment, note three obvious failings: figure 8 shows rays that are not predicted by the exploding-reflector model. These rays will be present in a zero-offset section. Lateral velocity variation is required for this situation to exist.

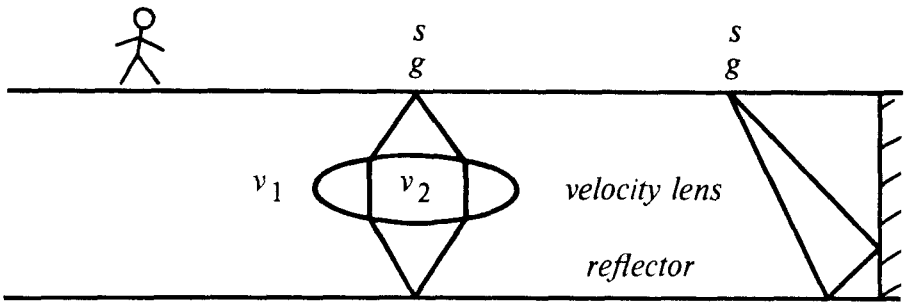


FIG. 1.1-8. Two rays, not predicted by the exploding-reflector model, that would nevertheless be found on a zero-offset section.

Second, the exploding-reflector concept fails with multiple reflections. For a flat sea floor with a two-way travel time t_1, multiple reflections are predicted at times $2t_1$, $3t_1$, $4t_1$, etc. In the exploding-reflector geometry the first multiple goes 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 the multiple reflections generated on the zero-offset section differ from those of the exploding-reflector model.

The third failing of the exploding-reflector model is where we are able to see waves bounced from both sides of an interface. The exploding-reflector model predicts the waves emitted by both sides have the same polarity. The physics of reflection coefficients says reflections from opposite sides have opposite polarities.

Plate Tectonics Example

Plate tectonic theory says the ocean floors are made of thin plates that are formed at volcanic ridges near the middle of the oceans. These plates move toward trenches in the deepest part of the ocean where they plunge back down into the earth. The best evidence for the theory is the lack of old

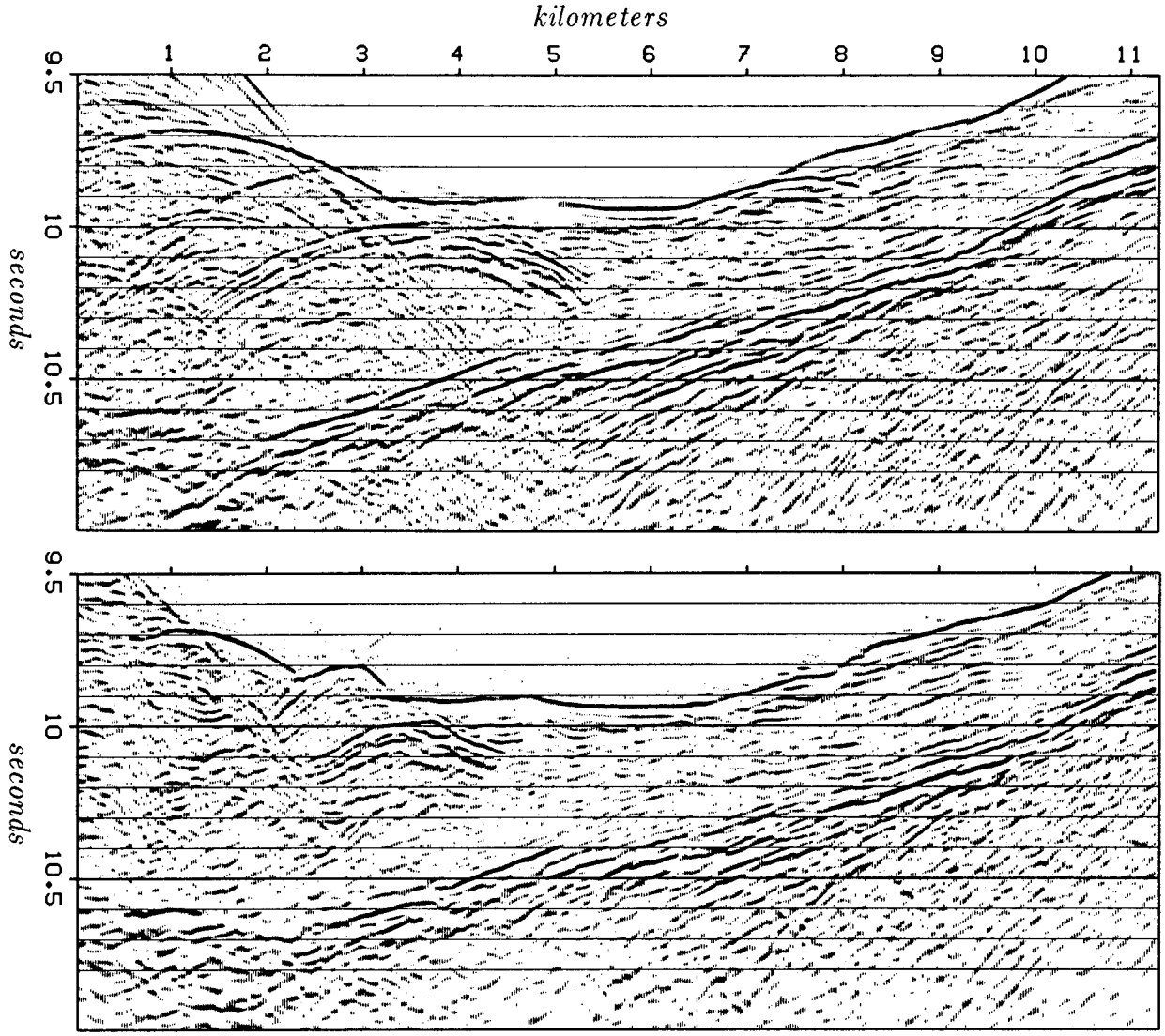


FIG. 1.1-9. Top is 11 kilometers of reflection data from a survey line across the Japan trench (Tokyo University Oceanographic Research Institute). Bottom shows the result of migration processing. (Ottolini)

rocks on the floors of the earth's oceans. Generally, continents are older rocks jostled by the younger moving oceanic plates. The formation of plates by mid-ocean ridge volcanism is readily observed in a variety of ways. Whether the plates really do plunge at the trenches is not so clear observationally. The evidence comes from earthquake locations and from reflection seismology. Figure 9 shows some reflection data from the Japan trench. Two reflections dominate, the sea floor reflection and a deeper layer dipping down to the left. This latter is presumably the top part of a plate that is beginning its descent into the earth. We can examine it for evidence of bending downward, such as

tension fractures near the surface. (The topmost layer is soft recent marine sediment loosely attached to the plate).

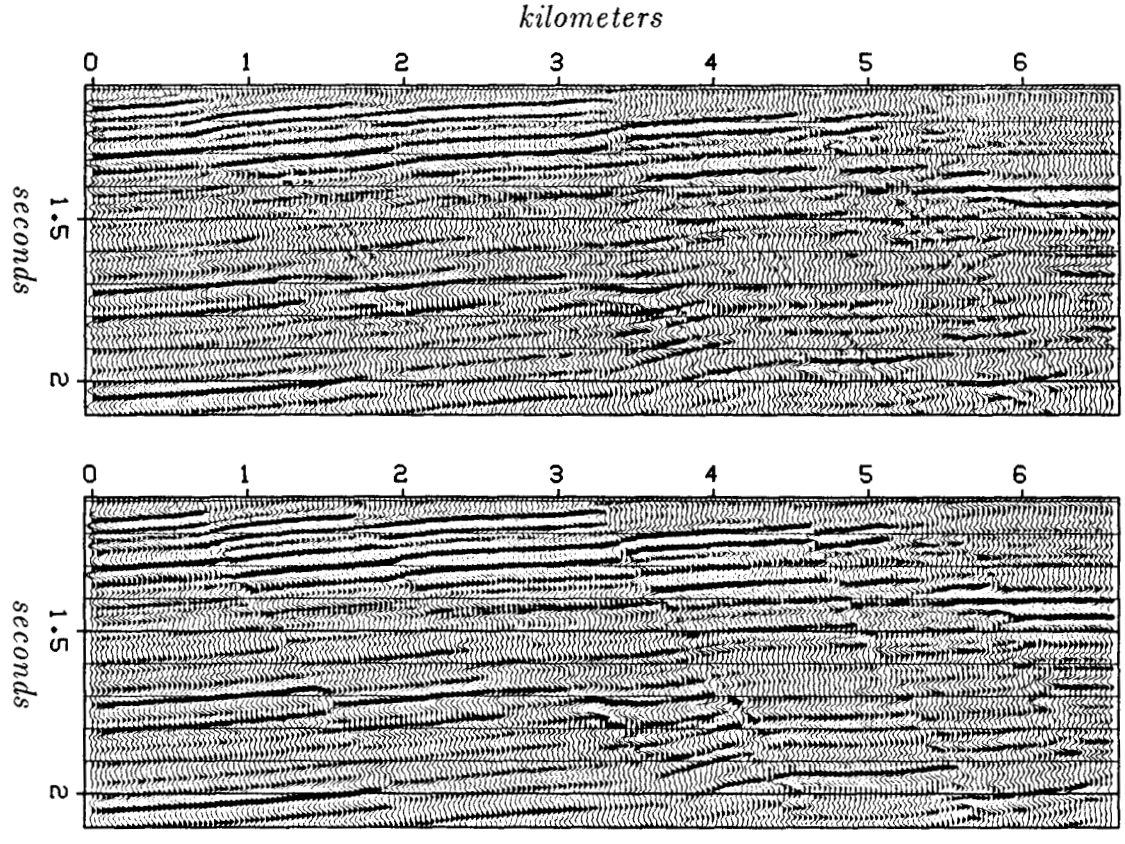


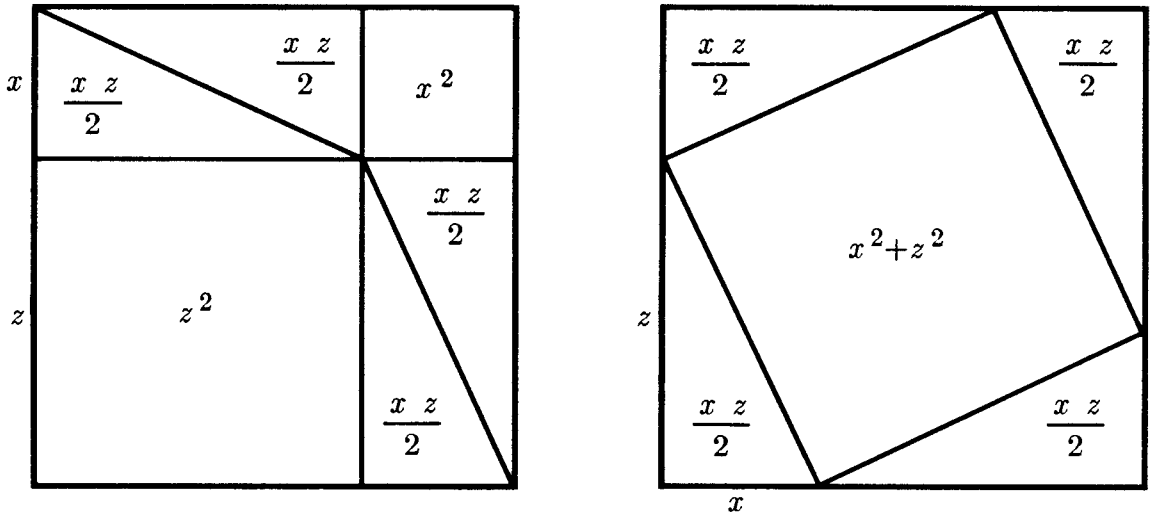
FIG. 1.1-10. Top is 6.5 kilometers of reflection data from a survey line offshore from the Texas coast of the Gulf of Mexico. Bottom shows the result of migration processing. (Rothman).

Notice that the top of the plot is not zero time. The time axis runs from 9.5 to 11.0 seconds. Before 9.5 sec there are no echoes — we are waiting for the waves to go between the ship and the ocean floor. Hyperbolic reflections around kilometers 1-3 are collapsed by migration to form interesting "blocky" shapes. Look at the sea floor topography near kilometer 8 and the difference between migrated and unmigrated data sections. After migration, the sea floor diffraction hyperbolas move away from the plate echo (kilometer 4). Fractures (especially the one at 6.2 km) are more sharply defined. Finally, if the plate bends downward, it is not apparent from the data given. The bending question really requires a more detailed analysis of lateral variation in seismic velocity.

For an example with petroleum interest, see figure 10, data from offshore Texas. Sediments are dropped where coastal rivers enter the Gulf of Mexico. The added weight causes slumping along steep faults. After a permeable sandstone layer has been identified by drilling, its reflection can be extrapolated up dip to the nearest fault on data like figure 10. The fault is likely to break the continuity of the permeability trapping the upward flowing hydrocarbons. A sandstone at this depth can have a porosity of 25%. Assume a seismic velocity of 2.2 km/sec. Deduce the scale between physical volume and the data in figure 10. Comparing the value of a volume of oil to the size of that same volume on figure 10, you can see the importance of good images.

EXERCISES

1. Prove the Pythagorean theorem, that is, the length of the hypotenuse $v\ t$ of a right triangle is determined by $x^2 + z^2 = v^2 t^2$. Hint:



2. Compute propagation angles for the hyperbola flanks in figure 9.
3. Using the result of exercise 2, deduce the plunge angle of the plate.
4. How deep is the Japan trench (water velocity is 1.5 km/sec)?
5. On the Gulf Coast data, which direction is offshore? Why?

1.2 Wave Extrapolation as a 2-D Filter

One of the main ideas in Fourier analysis is that an impulse function (a delta function) can be constructed by the 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 plane 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 Huygens secondary source. It is a two-dimensional (2-D) filter for spatial extrapolation of wavefields.

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 that is observable on the earth's surface is the intercept of the wavefront with the earth's surface. This speed, namely $v/\sin\theta$, is faster than v. Likewise, the speed of the intercept of the wavefront and the vertical axis is $v/\cos\theta$. A mathematical expression for a straight line, like that shown to be the wavefront in figure 1, is

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

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

$$z = v\,\frac{t}{\cos\theta} - x \tan\theta \tag{2}$$

Solving for time gives

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

Equation (3) tells the time that the wavefront will pass any particular location (x,z). The expression for a shifted waveform of arbitrary shape is $f(t - t_0)$. Using (3) to define the time shift t_0 gives an expression for a

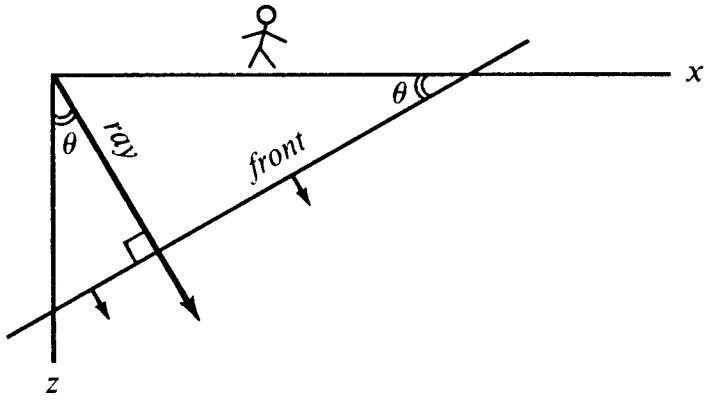


FIG. 1.2-1. Downgoing ray and wavefront.

wavefield that is some waveform moving on a ray.

$$\textit{moving wavefield} \;=\; f\left(\; t \;-\; \frac{x}{v} \sin\theta - \frac{z}{v} \cos\theta \right) \tag{4}$$

Waves in Fourier Space

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

Using Fourier integrals on time functions we encounter the *Fourier kernel* $\exp(-i\omega t)$. Specializing the arbitrary function in equation (4) to be the real part of the function $\exp[-i\omega(t-t_0)]$ gives

$$\textit{moving cosine wave} \;=\; \cos\left[\; \omega \left(\; \frac{x}{v} \sin\theta + \frac{z}{v} \cos\theta - t \; \right) \; \right] \tag{5}$$

To use Fourier integrals on the space-axis x the spatial angular frequency must be defined. Since we will ultimately encounter many space axes (three for shot, three for geophone, also the midpoint and offset), the convention will be to use 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 sign of i. The sign convention used here is the one used in most physics books, namely, the one that agrees with equation (5). Reasons for the choice are given in Section 1.6. With this convention, a wave moves

in the *positive* direction along the space axes. Thus the Fourier kernel for (x, z, t)-space will be taken to be

Fourier kernel =

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

Now for the whistles, bells, and trumpets. Equating (5) to the real part of (6), physical angles and velocity are related to Fourier components. These relations should be memorized!

Angles and Fourier Components	
$\sin\theta = \dfrac{v\ k_x}{\omega}$	$\cos\theta = \dfrac{v\ k_z}{\omega}$

(7)

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

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

We'll encounter *dispersion relations* and the *scalar wave equation* later. The importance of (8) is that it enables us to make the distinction between an arbitrary function and a chaotic function that actually is a wavefield. 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 nonvanishing values of P. You will have a wavefield if and only if all nonvanishing P have coordinates that satisfy (8). Even better, in practice the (x, t)-dependence at $z = 0$ is usually known, but the z-dependence is not. Then the z-dependence is found by assuming P is a wavefield, so the z-dependence is inferred from (8).

Migration Improves Horizontal Resolution

In principle, migration converts hyperbolas to points. In practice, hyperbolas don't collapse to a point, they collapse to a *focus.* A focus has measurable dimensions. Migration is said to be "good" because it increases spatial resolution. It squeezes a large hyperbola down to a tiny focus. To quantitatively describe the improvement of migration, the size of the hyperbola and the size of the focus must be defined. Figure 2 shows various ways of measuring the size of a hyperbola.

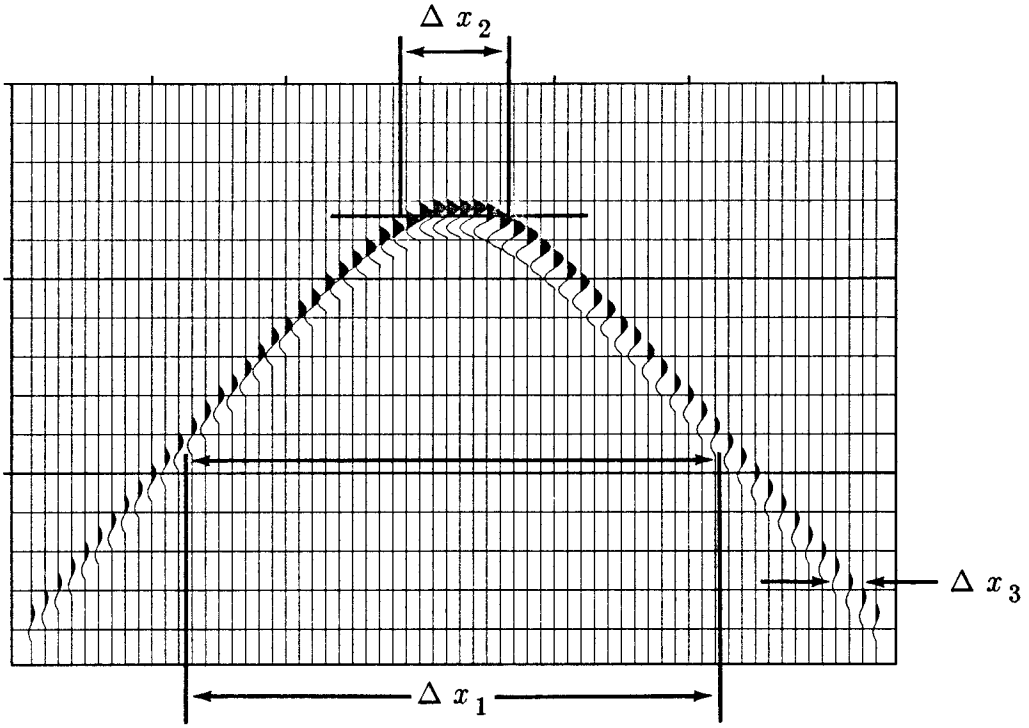


FIG. 1.2-2. Measurements of width parameters of a hyperbola.

The hyperbola carries an impulsive arrival. So the ω-bandwidth of the hyperbola is roughly given by the zero crossings on the time axis of the main energy burst. I'll mention 50 Hz as a typical value, though you could encounter values four times higher or four times lower. Knowledge of a seismic velocity determines depth resolution. I'll suggest 3 km/sec, though once again you could encounter velocities four times greater or four times less. These values imply a seismic wavelength of $v/f = 60$ meters. But the effective seismic wavelength is half the actual wavelength. The half comes from halving the velocity v in exploding reflector calculations, or equivalently, from realizing that the seismic wavelength is divided equally into upgoing and downgoing parts. Resolving power is customarily defined as about half the effective wavelength or about 15 meters. (Whether seismic resolution should be *half* the effective wavelength or a smaller fraction is an issue that involves signal-to-noise considerations outside our present study).

The lateral resolution requires estimates of hyperbola width and focus width. Figure 2 shows three hyperbola widths. The widest, Δx_1, includes about three-quarters of the energy in the hyperbola. Next is the width Δx_2, called the *Fresnel Zone*. It is measured across the hyperbola at the time when the first arrival has just changed polarity. Third is the smallest measurable width, found far out on a flank. This width, Δx_3, is the shortest

horizontal wavelength to be found. Resolution is the study of the size of error, and it is not especially useful to be precise about the error in the error. The main idea is that $\Delta x_1 > \Delta x_2 > \Delta x_3$. The bandwidth of the spatial k_x spectrum is roughly $1/\Delta x_3$. How small a focus can migration make? It will be limited by the available bandwidth in the k_x spectrum. The size of the focus will be about the same as Δx_3.

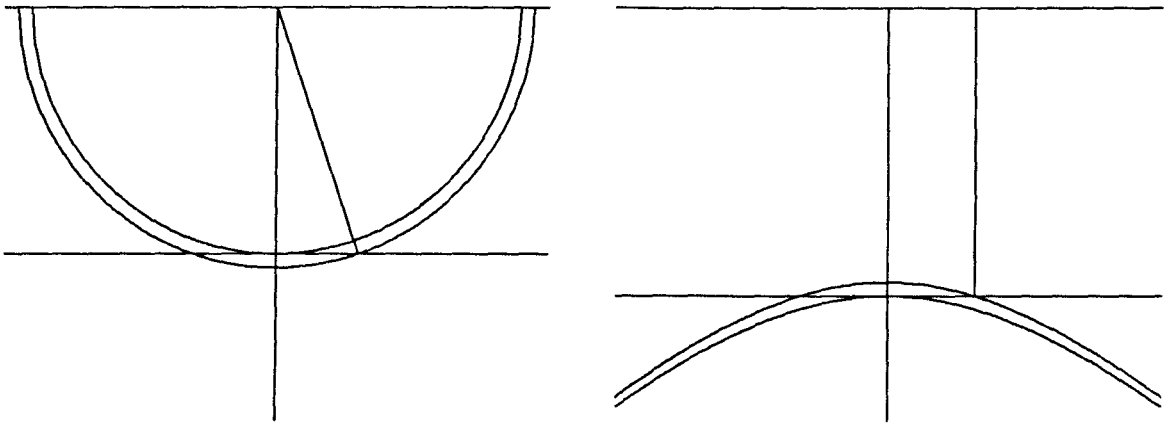


FIG. 1.2-3. Fresnel zone in (x, z)-space (left) and in (x, t)-space (right).

Figure 3 shows the geometry of the Fresnel zone concept. A Fresnel zone is an intercept of a spherical wave with a plane. The intercept is defined when the spherical wave penetrates the plane to a depth of a half wavelength. What is the meaning of the Fresnel width Δx_2? Imagine yourself in Berlin. There is a wall there. You may not go near it. Imagine a hole in the wall. You are shouting to a friend on the opposite side. How does the loudness of the sound depend on the size of the hole ΔX? It is not obvious, but it is well known, both theoretically and experimentally, that holes larger than the Fresnel zone cause little attenuation, but smaller holes restrict the sound in proportion to their size.

Wave propagation is a convolutional filter that smears information from a region Δx_2 along a reflector (or Δx_1 in the subsurface) to a point on the surface. Migration, the reverse of wave propagation, is the deconvolution operation. The final amount of lateral resolution is limited by the spatial bandwidth of the data.

Migration may be called for even where reflectors show no dip. When a well site is to be chosen within an accuracy of less than Δx_2 then the

interpreter is looking at subtle changes in amplitude or waveform along the reflector. Migration causes these amplitude and waveform variations to change and to move horizontally along the reflector. The distance moved is about equal the Fresnel zone.

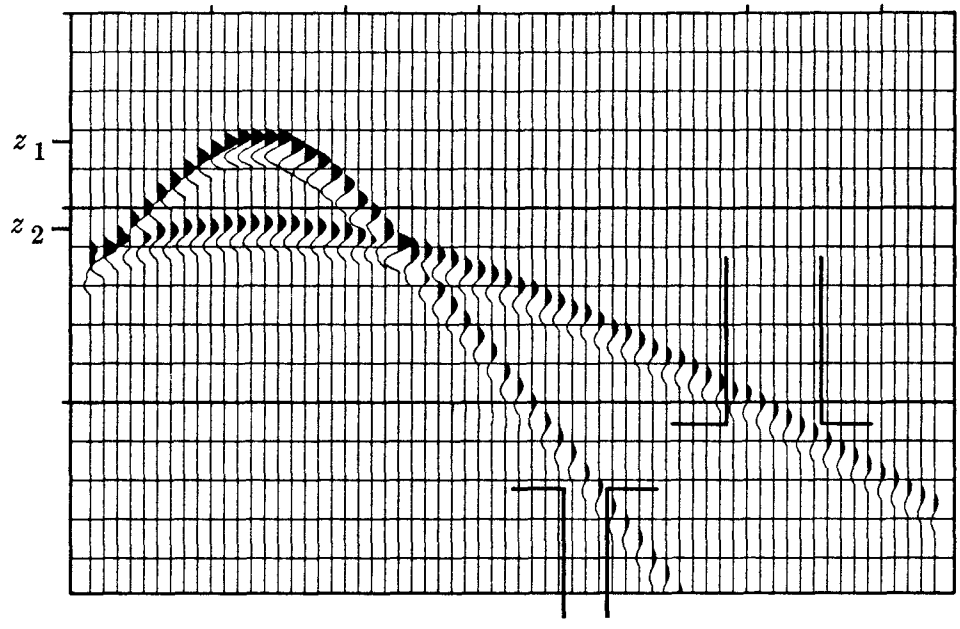


FIG. 1.2-4. Hyperboloids for an earth of velocity increasing with depth. Observable lateral wavelengths get longer with increasing depth. Thus lateral resolving power decreases with depth.

A basic fact of seismology is the resolution limitation caused by the increase with depth of the seismic velocity. What happens is that as the waves get deeper into the earth, their spatial wavelengths get longer because of the increasing velocity. The case of vertical resolution is simply this: longer wavelengths, less resolution. The case of horizontal resolution is similar, but the horizontal wavelength is directly measurable at the earth's surface. Figure 4 demonstrates this. Hyperboloids from shallow and deep scatterers are shown. Shallow hyperbolas have early tops and steep asymptotes. Deep scatterers have late tops and less steep asymptotes. The less steep asymptotes have longer horizontal wavelength. Horizontal wavelengths measured at the surface are unchanged at depth, even though velocity increases with depth. (This implication of Snell's law is shown in Section 1.5). Thus, lateral spatial resolution gets worse with depth. Compounding the above reason for decreasing resolution is the loss of high-frequency energy at late travel time.

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 computed 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. This is diagramed as follows:

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

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 that 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 name the function.

An example of these transformations on typical deep-ocean data is shown in figure 5.

In the deep ocean, sediments are fine-grained and deposit slowly in flat, regular, horizontal beds. The lack of permeable rocks like sandstone severely reduces the potential for petroleum production from the deep ocean. The fine-grained shales overlay irregular, igneous, basement rocks. In the plot of $P(t,k_x)$ the lateral continuity of the sediments is shown by the strong spectrum at low k_x. The igneous rocks show a k_x spectrum extending to such large k_x that the deep data may be somewhat spatially aliased (sampled too coarsely). The plot of $P(\omega,x)$ shows that the data contains no low-frequency energy. At large ω the energy is not dropping off as fast as one might like, which indicates temporal frequency aliasing. This aliasing is also apparent in the plot of $p(t,x)$ in the steplike appearance of the sea-floor

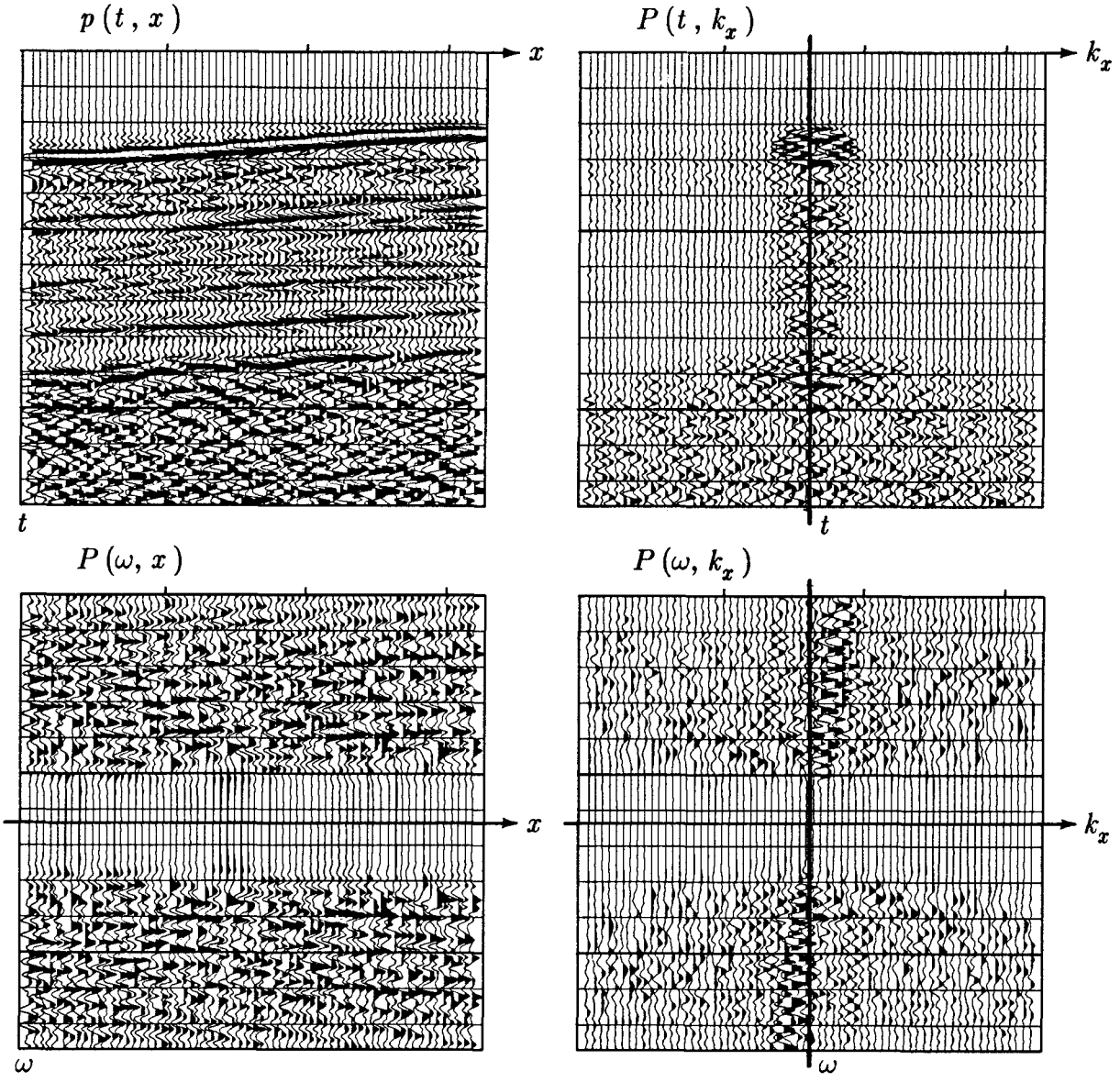


FIG. 1.2-5. A deep-marine dataset $p(t, x)$ from Alaska (U.S. Geological Survey) and the real part of various Fourier transforms of it. Because of the long travel time through the water, the time axis does not begin at $t=0$.

arrival. The dip of the sea floor shows up in (ω, k_x)-space as the energy crossing the origin at an angle.

Altogether, the two-dimensional Fourier transform of a collection of seismograms involves only twice as much computation as the one-dimensional Fourier transform of each seismogram. This is lucky. Let us write some equations to establish that the asserted procedure does indeed do a two-dimensional Fourier transform. Say first that any function of x and t

may be expressed as a superposition of sinusoidal functions:

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

(Sign convention used in Fourier transformation is explained in Section 1.6). The kernel in this *inverse* Fourier transform has the form of a wave moving in the plus x direction. Likewise, in the *forward* Fourier transform, the signs of both exponentials change, 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 sampled-time computation, so why bother?) The double integration can be nested to show that the temporal transforms are done first (inside):

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

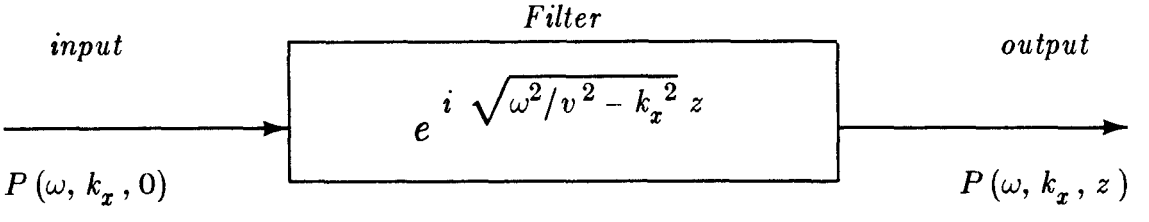
The quantity in brackets is a Fourier transform over ω done for each and every k_x. Alternately, the nesting could be done with the k_x-integral on the inside. That would imply rows first instead of columns (or vice versa). It is the separability of $\exp(-i\omega t + i k_x x)$ into a product of exponentials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

The Input-Output Relation

At the heart of the migration process is the operation of downward continuing data. Given the input data on the plane of the earth's surface $z=0$, we must manufacture the data that could be recorded at depth z. This is most easily done in the Fourier domain. The method will be seen to be simply multiplication by a complex exponential, namely,

$$P(\omega, k_x, z) = P(\omega, k_x, 0)\, e^{i k_z(\omega, k_x) z} \tag{10}$$

Since the operation is a multiplication in the Fourier domain, it may be described as an engineering diagram.



Downward continuation is a product relationship in both the ω-domain and the k_x-domain. What does the filter look like in the time and space

domain? It turns out like a cone, that is, it is roughly an impulse function of $x^2 + z^2 - v^2 t^2$. More precisely, it is the Huygens secondary wave source that was exemplified by 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.

Now let us see why the downward continuation filter has the mathematical form stated. Every point in the (ω, k_x)-plane refers to a sinusoidal plane wave. The variation with depth will also be sinusoidal, namely $\exp(ik_z z)$. The value of k_z for the plane wave is found simply by solving equation (8):

$$k_z = \pm \left(\frac{\omega^2}{v^2} - k_x^2 \right)^{1/2} \tag{11a}$$

$$= \pm \frac{\omega}{v} \left(1 - \frac{v^2 k_x^2}{\omega^2} \right)^{1/2} \tag{11b}$$

$$= \pm \frac{\omega}{v} \cos \theta \tag{11c}$$

Choice of the *plus* sign means that $\exp(-i\omega t + i k_z z)$ is a *down*going wave (because the phase will stay constant if z increases as t increases). Choice of a *minus* sign makes the wave *up*coming. The exploding-reflector concept requires upcoming waves, so we nearly always use the minus sign, whether we are migrating or modeling.

The input-output filter, being of the form $e^{i\phi}$, appears to be a phase-shifting filter with *no amplitude scaling*. This bodes well for our plans to deconvolve. It means that signal-to-noise power considerations will be much less relevant for migration than for ordinary filtering.

EXERCISES

1. Suppose that you are able to observe some shear waves at ordinary seismic frequencies. Is the spatial resolution better, equal, or worse than usual? Why?
2. Scan this book for hyperbolic arrivals on field data and measure the Fresnel zone width. Where zero offset recordings are not made, a valid approximation is to measure Δx_2 along a *tilted* line.
3. Explain the horizontal "layering" in figure 1.2-5 in the plot of $P(\omega, x)$. What determines the "layer" separation? What determines the "layer"

slope?

4. Evolution of a wavefield with time is described by

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

Let $P(k_x, k_z, 0)$ be constant, signifying a point source at the origin in (x, z)-space. Let t be very large, meaning that phase $= \phi =$ $[-\omega(k_x, k_z) + k_x (x/t) + k_z (z/t)]t$ in the integration is rapidly alternating with changes in k_x and k_z. Assume that the only significant contribution to the integral comes when the phase is stationary, that is, where $\partial\phi/\partial k_x$ and $\partial\phi/\partial k_z$ both vanish. Where is the energy in (x, z, t)-space?

5. Downward continuation of a wave is expressed by

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

Let $P(k_x, 0, \omega)$ be constant, signifying a point source at the origin in (x, t)-space. 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 are all found in modern production environments. As a group they handle wide-angle rays easily. As a group they are used less successfully to deal with lateral velocity variation.

Travel-Time 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 travel time* 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 sufficiently narrowed for a drill site to

be selected, the vertical exaggeration factor in use is likely to be about unity (no exaggeration).

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

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

The first task in interpretation of seismic data is to figure out the approximate numerical value of the vertical exaggeration. It probably won't be printed on the data header because the seismic velocity is not really known. Furthermore, the velocity usually *increases* with depth, which means that the vertical exaggeration *decreases* with depth. For velocity-stratified media, the time-to-depth conversion formula is

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

Hyperbola-Summation and Semicircle-Superposition Methods

The methods of hyperbola summation and semicircle superposition are the most comprehensible of all known methods.

Recall the equation for a conic section, that is, a circle in $(x\,,\,z)$-space or a hyperbola in $(x\,,\,t)$-space. Converting to travel-time depth τ

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

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

Figure 1 illustrates the *semicircle-superposition* method. (Both the figure and its caption are from Schneider's classic paper [1971]). Taking the data field to contain a few impulse functions, the output should be a superposition of the appropriate semicircles. Each semicircle denotes the spherical-reflector earth model that would be implied by a dataset with a single pulse. Taking the data field to be one thousand seismograms of one thousand points each, then the output is a superposition of one 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 $(x\,,\,\tau)$-

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.

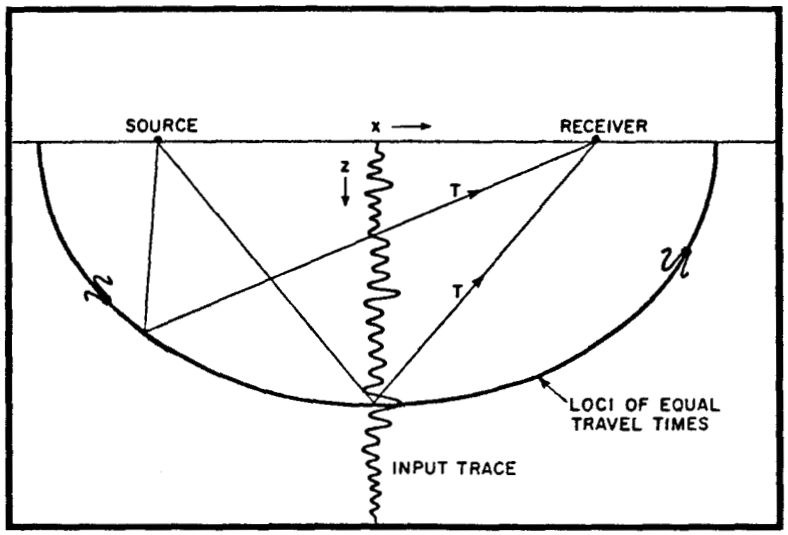


FIG. 1.3-1. The process may be described in numerous ways. Two very simple and equally valid representations are indicated in figures 1 and 2. Shown here is a representation of the process in terms of what happens to a single input trace plotted in depth (time may also be used) midway between its source and receiver. Each amplitude value of this trace is mapped into the subsurface along a curve representing the loci of points for which the travel time 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 as 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. (from Schneider, W. A., 1971 [by permission])

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 in 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, imagine a hyperbola, equation (3b), set down with its top on 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. We can wonder whether the hyperbola-migration method is better or worse than or equivalent to the semicircle method.

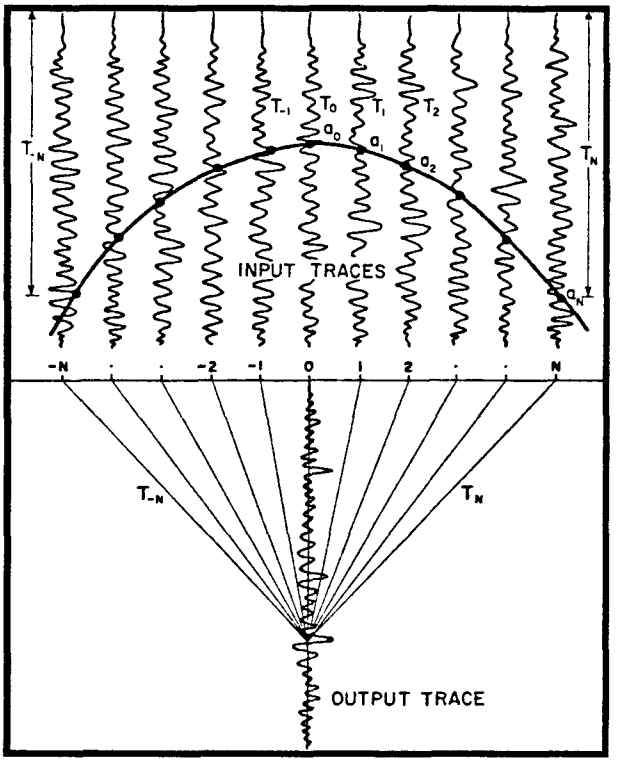


FIG. 1.3-2. A second description of the process is provided here. The process is represented in terms of how an output trace is developed from an ensemble of input traces, shown as CDP-stacked traces in the upper half of the figure. The output in the lower half reflects how each amplitude value at (x, z) is obtained by summing input amplitudes along the travel-time curve shown. This curve defines a diffraction hyperbola. If a diffraction source existed in the subsurface at the output point shown, then a large amplitude would result. The process also works for reflectors since a reflector may be regarded as a continuum of diffracting elements whose individual images merge to produce a smooth continuous boundary. (also from Schneider, 1971)

The opposite of data processing or building models from data is constructing synthetic data from models. With a slight change, the above two processing programs can be converted to modeling programs. Instead of *hyperbola summation* or *semicircle superposition,* you do *hyperbola superposition* or *semicircle summation.* We can also wonder whether the processing programs really are inverse to the modeling programs. Some factors that need to be considered are (1) the angle-dependence of amplitude (the obliquity function) of the Huygens waveform, (2) spherical spreading of energy, and (3) the phase-shift on the Huygens waveform. It turns out that results are reasonably good even when these complicating factors are ignored.

As other methods of migration were developed, the deficiencies of the earlier methods were more clearly understood and found to be largely correctable by careful implementation. One advantage of the later methods is that they implement 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 with the Kirchhoff diffraction integral by Trorey [1970] and Hilterman [1970] led to forward modeling programs. Eventually (Schneider [1977]) this work 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 Kirchhoff method, although, strictly speaking, the Kirchhoff integral applies only in the constant-velocity case.

Spatial Aliasing

Spatial aliasing means insufficient sampling of the data along the space axis. This difficulty is so universal, that all migration methods must consider it.

Data should be sampled at more than two points per wavelength. Otherwise the wave arrival direction becomes ambiguous. Figure 3 shows synthetic data that is sampled with insufficient density along the x-axis. You can see that the problem becomes more acute at high frequencies and steep dips.

There is no generally-accepted, automatic method for migrating spatially aliased data. In such cases, human beings may do better than machines, because of their skill in recognizing true slopes. When the data is adequately sampled, however, computer migration based on the wave equation gives better results than manual methods. Contemporary surveys are usually adequately sampled along the line of the survey, but there is often difficulty in the perpendicular direction.

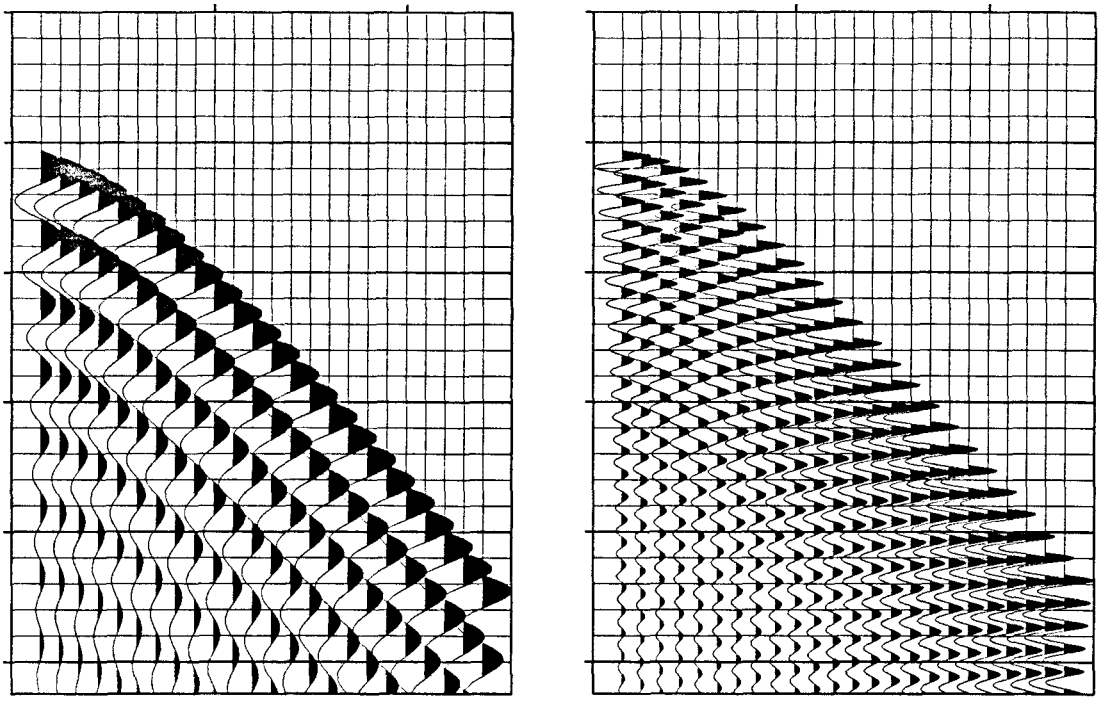


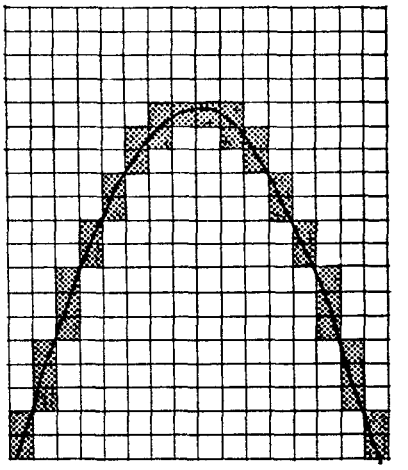
FIG. 1.3-3. Insufficient spatial sampling of synthetic data. To better perceive the ambiguity of arrival angle, view the figures at a grazing angle from the side.

The hyperbola-sum-type methods run the risk of the migration operator itself becoming spatially aliased. This should be avoided by 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 4. The likelihood of getting an aliased operator increases where the hyperbola is steeply sloped. In production examples an aliased operator often stands out above the sea-floor reflection, where — although the sea floor may be flat — it acquires a noisy precursor due to the steeply flanked hyperbola crossing the sea floor.

The Phase-Shift Method (Gazdag)

The phase-shift method proceeds straightforwardly by extrapolating downward with $\exp(ik_z z)$ and subsequently evaluating the wavefield at $t=0$ (the reflectors explode at $t=0$). 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 Kirchhoff methods, with this method there is no danger of aliasing the operator.

FIG. 1.3-4. For a low-velocity hyperbola, integration will require more than one point per channel.



The phase-shift method begins with a two-dimensional Fourier transform (2D-FT) of the dataset. (Some practical details about 2D-FT are described in Section 1.7). Then the transformed data values, all in the (ω, k_x)-plane, are downward continued to a depth Δz by multiplying by

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

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

$$C = \exp\left\{ -i\ \omega\ \Delta\tau \left[1 - \left(\frac{v\ k_x}{\omega} \right)^2 \right]^{1/2} \right\} \tag{5}$$

Data will be multiplied many times by C, thereby downward continuing it by many steps of $\Delta\tau$.

Next is the task of imaging. At each depth an inverse Fourier transform is followed by selection of its value at $t=0$. (Reflectors explode at $t=0$). Luckily, only the Fourier transform at one point, $t=0$, is needed, so that is all that need be computed. The computation is especially easy since the value at $t=0$ is merely a summation of each ω frequency component. (This may be seen by substituting $t=0$ into the inverse Fourier integral). Finally, inverse Fourier transform k_x to x. The migration process, computing the image from the upcoming wave u, may be summarized as follows:

$U(\omega, k_x) = FT[\, u(t, x)]$
For $\tau = \Delta\tau,\ 2\Delta\tau,\ \cdots,$ end of time axis on seismogram {
  For all k_x {
    $Image(k_x, \tau) = 0.$
    For all ω {
      $C = \exp(-i\ \omega\ \Delta\tau \sqrt{1 - v^2\ k_x^2/\omega^2}\ )$
      $U(\omega, k_x) = U(\omega, k_x) * C$
      $Image(k_x, \tau) = Image(k_x, \tau) + U(\omega, k_x)$
      }
    }
  $image(x, \tau) = FT[Image(k_x, \tau)]$
  }

Inverse migration (modeling) proceeds in much the same way. Beginning from an upcoming wave that is zero at great depth, the wave is marched upward in steps by multiplication with $\exp(i\ k_z\ \Delta\ z)$. As each level in the earth is passed, exploding reflectors from that level are added into the upcoming wave. The program for modeling the upcoming wave u is

$Image(k_x, z) = FT[\, image(x, z)]$
For all ω and all k_x
  $U(\omega, k_x) = 0.$
For all ω {
For all k_x {
For $z = z_{\max},\ z_{\max} - \Delta z,\ z_{\max} - 2\Delta z,\ \cdots, 0$ {
  $C = \exp(+i\ \Delta z\ \omega \sqrt{v^{-2} - k_x^2/\omega^2}\ )$
  $U(\omega, k_x) = U(\omega, k_x) * C + Image(k_x, z)$
  } } }
$u(t, x) = FT[\, U(\omega, k_x)]$

The positive sign in the complex exponential is a combination of two negatives, the *up*coming wave and the *up*ward extrapolation. The three loops on ω, k_x, and z are interchangeable. When the velocity v is a constant function of depth the program can be speeded by moving the computation of the complex exponential C out of the inner loop on z.

The velocity is hardly ever known precisely, so although it may be increasing steadily with depth, it is often approximated as constant in layers

instead of slowly changing at each of the thousand or so time points on a seismogram. The advantage of this approximation is economy. Once the square root and the sines and cosines in (5) have been computed, the complex multiplier (5) can be reused many times. With a 4-millisecond sample rate and a layer 200 milliseconds thick, the complex multiplier gets used 50 times before it is abandoned.

The Stolt Method

On most computers the Stolt method of migration is the fastest one — by a wide margin. For many applications, this will be its most important attribute. For a constant-velocity earth it incorporates the Huygens wave source 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 Stolt method does not handle depth variation in velocity. This drawback is largely offset in practice by an approximate correction that uses an axis-stretching procedure (Section 4.5). A practical problem is the periodicity of all the Fourier transforms. In principle this is no problem at all, since it can be solved by adequately surrounding the data by zeroes.

A single line sketch of the Stolt method is this:

$$p(x, t) \rightarrow P(k_x, \omega) \rightarrow P'(k_x, k_z = \sqrt{\omega^2/v^2 - k_x^2}) \rightarrow P(x, z)$$

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

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

Perform a two-dimensional inverse Fourier transform:

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

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

$$Image(x, z) = \iint e^{i k_x x} e^{i k_z(\omega, k_x) z} P(\omega, k_x, 0)\, d\omega\, dk_x \tag{7}$$

Equation (7) gives the final image, but it is in an unattractive form, since it implies that a two-dimensional integration must be done for each and every z-level. The Stolt procedure converts the three-dimensional calculation thus implied by (7) to a single two-dimensional Fourier transform.

So far nothing has been done to specify an *upcoming* wave instead of a downgoing wave. The direction of the wave is defined by the relationship of z and t that is required to keep the phase constant in the expression $\exp(-i\omega t + ik_z z)$. If ω were always positive, then $+k_z$ would always refer

to a downgoing wave and $-k_z$ to an upcoming wave. Negative frequencies ω as well as positive frequencies are needed to describe waves that have real (not complex) values. So the proper description for a downgoing wave is that the signs of ω and k_z must be the same. The proper description for an upcoming wave is the reverse. With this clarification the integration variable in (7) will be changed from ω to k_z.

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

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

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

Put (8) into (7), and include also a minus sign so that the integration on k_z goes from minus infinity to plus infinity as was the integration on ω.

$$Image(x, z) = \tag{9}$$

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

Equation (9) states the 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 data from $p(t, x, 0)$ to $P(\omega, k_x, 0)$.
2. Interpolate 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 $\cos\theta$).
3. Inverse Fourier transform to (x, z)-space.

Samples of Stolt migration of impulses are shown in figure 5. You can see the expected semicircular smiles. You can also see a semicircular frown hanging from the bottom of each semicircle. The worst frown is on the deepest spike. The semicircular mirrors have centers not only at the earth's surface $z=0$ but also at the bottom of the model $z=z_{max}$. It is known that these frowns can be suppressed by interpolating more carefully. (Interpolation is the way you convert from a uniform mesh in ω, to a uniform mesh in k_z). Interpolate with say a *sinc* function instead of a linear interpolator. (See Section 4.5). A simpler alternative is to stay away from the bottom of

the model, i.e. pad with lots of zeroes.

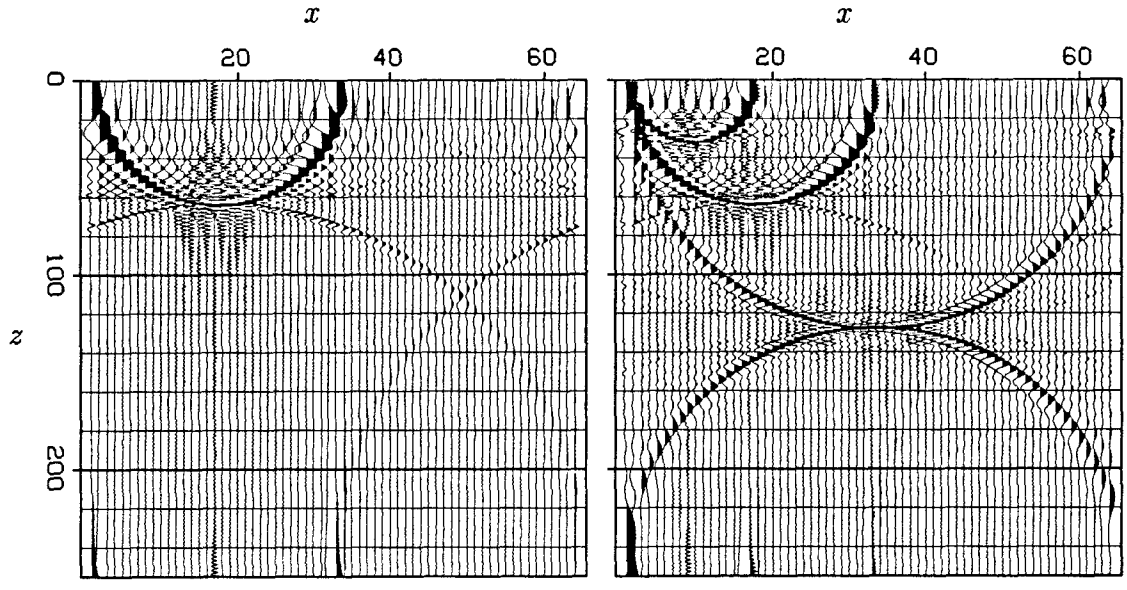


FIG. 1.3-5. Response of Stolt method to data with impulses. Semicircles are seen, along with computation artifacts.

It seems that an extraordinary amount of zero padding is required on the time axis. To keep memory requirements reasonable, the algorithm can be reorganized as described in an exercise. Naturally, the periodicity in x also requires padding the x-axis with zeroes.

Hyperbola Summation Refined into the Kirchhoff Method

Schneider [1978] states the analytic representation for the Huygens secondary wavelet

$$FT^{-1}(e^{i k_z z}) = \frac{1}{\pi} \frac{\partial}{\partial z_0} \frac{\text{step}(t - r/v)}{\sqrt{t^2 - r^2/v^2}} \tag{10}$$

where r is the distance $\sqrt{x^2 + (z - z_0)^2}$ between the (exploding reflector) source and the receiver. The function (10) contains a pole and the derivative of a step function. Because of the infinities it really cannot be graphed. But from the mathematical form you immediately recognize that the disturbance concentrates on the expected cone. The derivative of the step function gives a positive impulsive arrival on the cone. The derivative of the inverse square root gives the impulse a tail of negative polarity decaying with a $-3/2$ power. The cosine obliquity arises because the derivative is a z derivative and not

an r derivative.

Equation (10) states the *two*-dimensional Huygens wavelet, not the *three*-dimensional wavelet (which differs in some minor aspects). Although waves from point sources are mainly spherical, the focusing of bent layers is mainly a two-dimensional focusing, i.e., bent layers are more like cylinders than spheres.

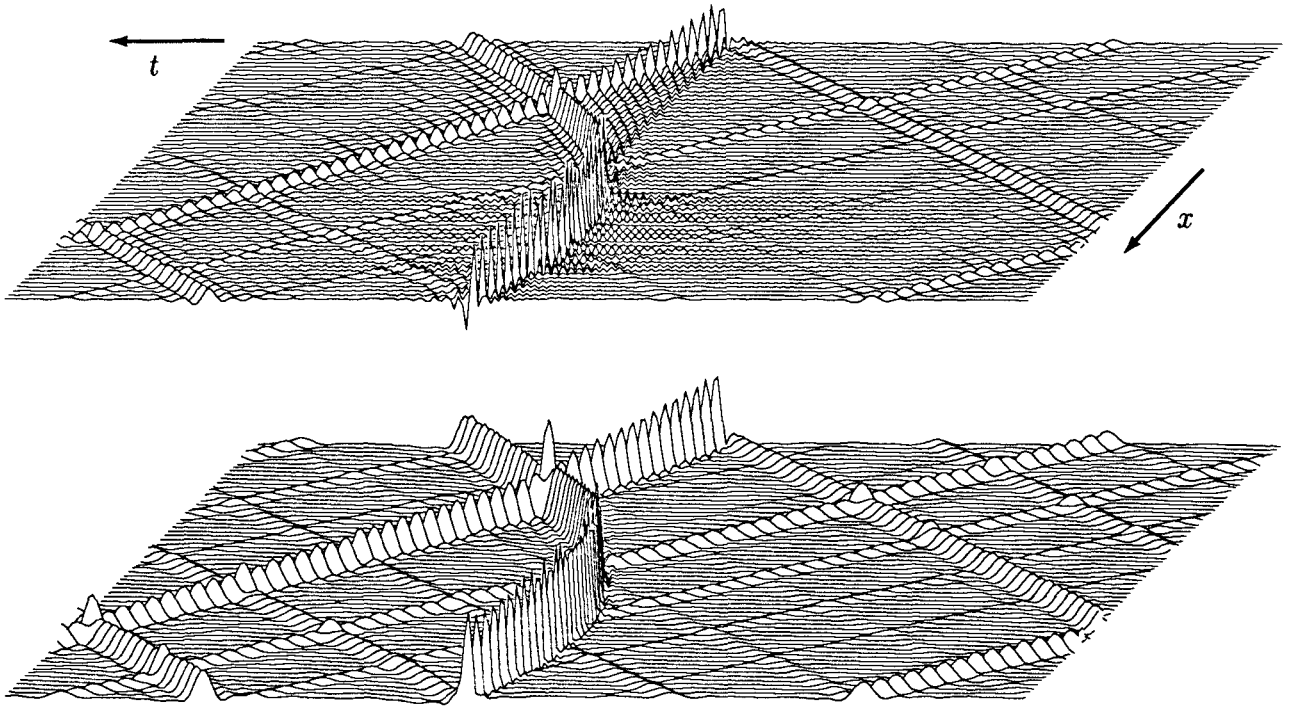


FIG. 1.3-6. The Huygens wavelet (top) and a smoothed time integral (bottom).

You might wonder why anyone would prefer approximations, given the exact inverse transform (10). The difficulty of graphing (10) shows up in practice as a difficulty in convolving it with data. That is why early Kirchhoff migrations were generally recognizable by precursor noise above a flat sea floor. Chapters 2 and 4 are largely devoted to extensions of (10) that are valid with variable velocity and that are better representations on a data mesh.

In the Fourier domain, the Huygens secondary source function is simple and smooth. It is a straightforward matter to evaluate the function on a rectangular mesh and inverse transform with the programs in Section 1.7. Figure 6 shows the result on a 256 $\times$ 64 point mesh. (In practice the mesh would be about 1024 $\times$ 256 or more, but the coarser mesh used here provides

a plot of suitable detail). Because of the difficulty in plotting functions that resemble an impulsive doublet, a second plot of the time integral, (with gentle band limiting) is displayed in the lower part of figure 6.

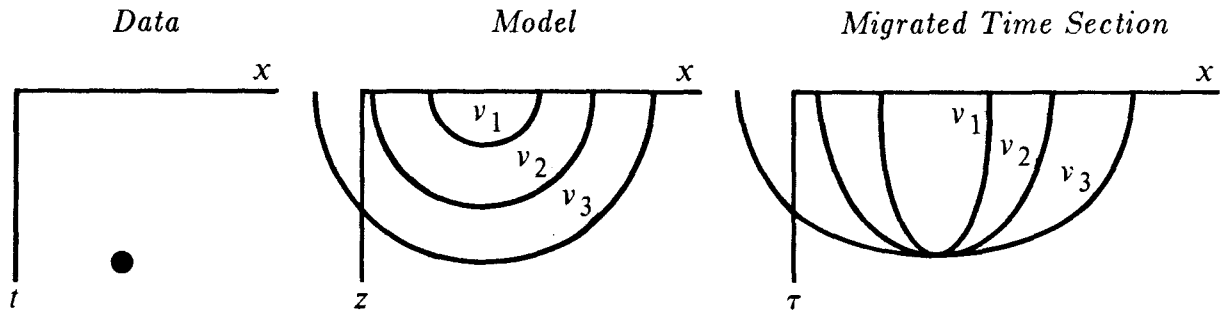


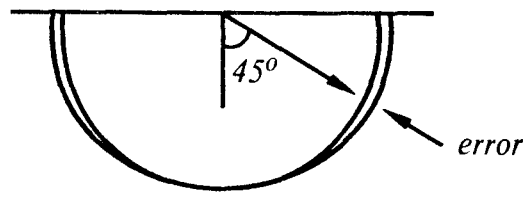
FIG. 1.3-7. Velocity error sensitivity increases with angle up to 90° . Migration of a data impulse as a function of velocity. Three possible choices of constant velocity are shown superposed on one plane.

Sensitivity of Migration to Velocity Error

Figure 7 shows how the migration impulse response depends on velocity. Recall that migrated data is ordinarily displayed as a *time* section. Arbitrary velocity error makes no difference for horizontal bedding.

Different people have different accuracy criteria. A reasonable criterion is that the positioning error of the energy in the semicircles should be less than a half-wavelength. For the energy moving horizontally, the positioning error is simply related to the dominant period ΔT and the travel time T. The ratio $T/\Delta T$ is rarely observed to exceed 100. This 100 seems to be a fundamental observational parameter of reflection seismology in sedimentary rock. (Theoretically, it might be related to the "Q" of sedimentary rock or it may relate to generation of chaotic internal multiple reflections. Larger values than 100 occur when (1) much of the path is in water or (2) at time depths greater than about 4 seconds). Figure 8 compares two nearby migration velocities. The separation of the curves increases with angle. For the separation to be less than a wavelength, for 90° dip the velocity error must be less than one part in 100. For 45° migration velocity error could be larger by $\sqrt{2}$.

FIG. 1.3-8. Timing error of the wrong velocity increases with angle.



Velocities are rarely known this accurately. So we may question the value of migration at wide angles.

Subjective Comparison and Evaluation of Methods

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

	Hyperbola Sum or Semicircle Superpose	Phase Shift	Stolt
$v(z)$	ray tracing	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	easy for any $v(z)$	for const v
wraparound noise?	no	on x, see Section 4.5 to alleviate on t	on x, see Section 4.5 to alleviate on (t, z)
$v(x)$	Production programs have serious pitfalls.	approximately by iteration and interpolation	no known program
side boundaries and irregular spacing	excellent	poor	poor
Speed	slow	average	very fast
memory organization	awkward	good	good

TABLE 1.3-1. Subjective comparison of three wide-angle migration methods.

The perspective of later chapters makes it possible to remark on the quality of the wide-angle methods as a group, and it is useful to do that now. Their greatest weakness is their difficulties with lateral velocity variation. Their greatest strength, wide-angle capability, is reduced by the weakness of other links in the data collection and processing chain, namely:

1. Shot-to-geophone offset angles are often 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 not sampled densely enough to represent steeply dipping data without aliasing.
4. Accuracy in the knowledge of velocity is seldom enough to justify processing to wide angles.
5. Noise eventually overpowers all echoes and this also implies an angle cutoff. For example, imagine oil reservoirs at a time depth of two seconds, where data recording stopped at four seconds. The implied angle cutoff is at 60° .

EXERCISES

1. The wave modeling program sketch assumes that the exploding reflectors are impulse functions of time. Modify the program sketch for wave modeling to include a source waveform $s(t)$.
2. The migration program sketch allows the velocity to vary with depth. However the program could be speeded considerably when the velocity is a constant function of depth. Show how this could be done.
3. Define the program sketch for the inverse to the Stolt algorithm — that is, create synthetic data from a given model.
4. The Stolt algorithm can be reorganized to reduce the memory requirement of zero padding the time axis. First Fourier transform x to k_x. Then select, from the (t, k_x)-plane of data, vectors of constant k_x. Each vector can be moved into the space of a long vector, then zero padded and interpolated. Sketch the implied program.
5. Given seismic data that is cut off at four seconds, what is the deepest travel time depth from which 80° dips can be observed?

1.4 The Physical Basis

Previous sections have considered the *geometrical* aspects of wave propagation and how they relate to seismic imaging. Now we will 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 image construction?

Because of the complexity of sedimentary rocks, there is not universal agreement on an appropriate mathematical description. To help you understand the degree to which theory can be used as a guide, I will point to some inconsistencies between theory and current industrial practice.

The Clastic Section

Generally speaking, most petroleum reservoir rocks are sandstones. Sandstones are most often made by the sands that are deposited near the mouths of rivers where the water velocity is no longer sufficient to move them. The sands deposit along the terminus of the sand bars found at the river mouth, often along a slope of 25° or so, as depicted in figure 1. Although the sands are not laid down in flat layers, the process may build a horizontal layer.

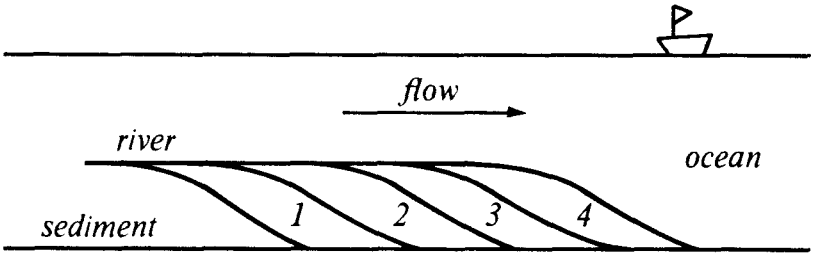


FIG. 1.4-1. Sands (petroleum reservoir rocks) deposit on fairly steep slopes where rivers run into the ocean.

Clays are more fine-grained materials (dirt) that are carried out to deeper water before they settle to make shale. Shale deposits tend to be layered somewhat more horizontally than sandstones. Specific locations of sand deposition change with the passing of storms and seasons, leaving a wood-grain-like appearance in the rock.

The river delta itself is a complicated, ever-changing arrangement of channels and bars, constantly moving along the coastline. At any one time the delta seems to be moving seaward as deposits are left, but later settling, compression, or elevation of sea level can cause it to move landward.

Sand is important because its porosity enables oil to accumulate and its permeability enables the oil to move to a well. Shale is important because it contains the products of former life on earth, and their hydrocarbons. These escape to nearby sands, but often not to the earth's surface, because of 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 than sands. Geophysicists on the surface see with seismic wavelengths ($\approx$ 30 meters) the final interbedded three-dimensional mixture of sands and shales.

Mixtures of sands and shales are called *clastic* rocks. The word *clast* means *break*. Clastic rocks are made from broken bits of crystalline rock. Most sedimentary rocks are clastic rocks. Most oil is found in clastic rocks. But much oil is also found in association with carbonates such as limestone. Carbonates are formed in shallow marine environments by marine organisms. Many carbonates (and clastics) contain oil that cannot be extracted because of lack of permeability. Permeability in carbonate rocks arises through several obscure processes. The seismologist knows carbonates as rocks with greater velocity than clastic rocks. Typically, a carbonate has a 20-50% greater velocity than a nearby clastic rock. Clastic rock sometimes contains limestone, in which case it is called *marl*.

Chrono-Stratigraphy

Strange as it may seem, there is not universal agreement about the exact nature of seismic reflections. Physicists tend to think of the reflections as caused by the interface between rock types, as a sand-to-shale contact. The problem with this view is that sands and shales interlace in complex ways, both larger and smaller than the seismic wavelength. Many geologists, particularly a group known as *seismic stratigraphers,* have a different concept. (See *Seismic Stratigraphy — Applications to Hydrocarbon Exploration,* memoir 26 of the American Association of Petroleum Geologists). 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 in between. Data interpretation based on this assumption is called *chrono-stratigraphy.* The view of the seismic stratigrapher seems reasonable enough for areas that are wholly clastic, but when carbonates and other rocks are present, the physicist's view seems more appropriate. For further details, the book of Sheriff [1980] is recommended.

Nonobservation of Converted Shear Waves

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 is a shear wave (S-wave). The shear wave can be polarized with ground motion in a horizontal plane (SH) or in a vertical plane (SV). Theory, field data, and laboratory measurement are in agreement. Successful experimental work with S-waves in the prospecting environment was done by Cherry and Waters [1968] and Erickson, Miller and Waters [1968].

It is remarkable that more than 99% of industrial petroleum prospecting ignores the existence of shear waves. Mathematically the earth is treated as if it were a liquid or a gas. The experimental work with shear waves used special equipment to generate and record vibration perpendicular to the survey line, i.e. SH-waves. The picture of the earth given by these transverse waves is often impaired by the soil layers, but sometimes the SH-wave picture is clear and consistent. Surprisingly, even good SH-wave data is often difficult to relate to the P-wave picture. These experimental studies show that the shear waves typically travel about half the speed of the pressure waves except in the soil layer, where the shear wave speed is often much slower and more variable. Observed shear waves usually have lower frequency than pressure waves. A shear wave with half the frequency and half the velocity of a pressure wave has just the same wavelength and hence the same resolving power as the pressure wave. Indeed, experimental work shows that shear waves do offer us about the same spatial resolution as pressure waves. Most land seismic data shows only the vertical component of motion, and all marine seismic data records the pressure. So in the conventional recording geometry, ideally we should never see SH-waves. More precisely, SH-waves should be small, arising only from the earth's departure from simple stratification.

The puzzling aspect of shear waves in reflection seismology is the failure of petroleum prospectors using the standard operating arrangement to routinely observe P-to-S conversions. Theory predicts that P-waves hitting an interface at an angle should be partially converted to SV-waves.

Furthermore, for the 30°-60° angle reflections that are routinely encountered, these converted waves should have a size comparable to the P-wave.

The routine geometry of recording and processing discriminates somewhat against converted shear waves. But it discriminates against multiple reflections too (in much the same way) and we see multiples all the time. Furthermore the signature of converted waves should resemble that of multiples, but be distinctly different. Converted waves should show up routinely in velocity surveys (Chapter 3). Figure 2 shows a zero-offset section containing some multiple reflections. The multiple reflection is recognizable as a replica of earlier topography. Converted waves would replicate the topography but the time scaling would be in the ratio of about 3/2 instead of exactly 4/2. With a sufficiently complex topography, as in figure 2, the probability is low that the converted wave would be mistaken for another primary reflection.

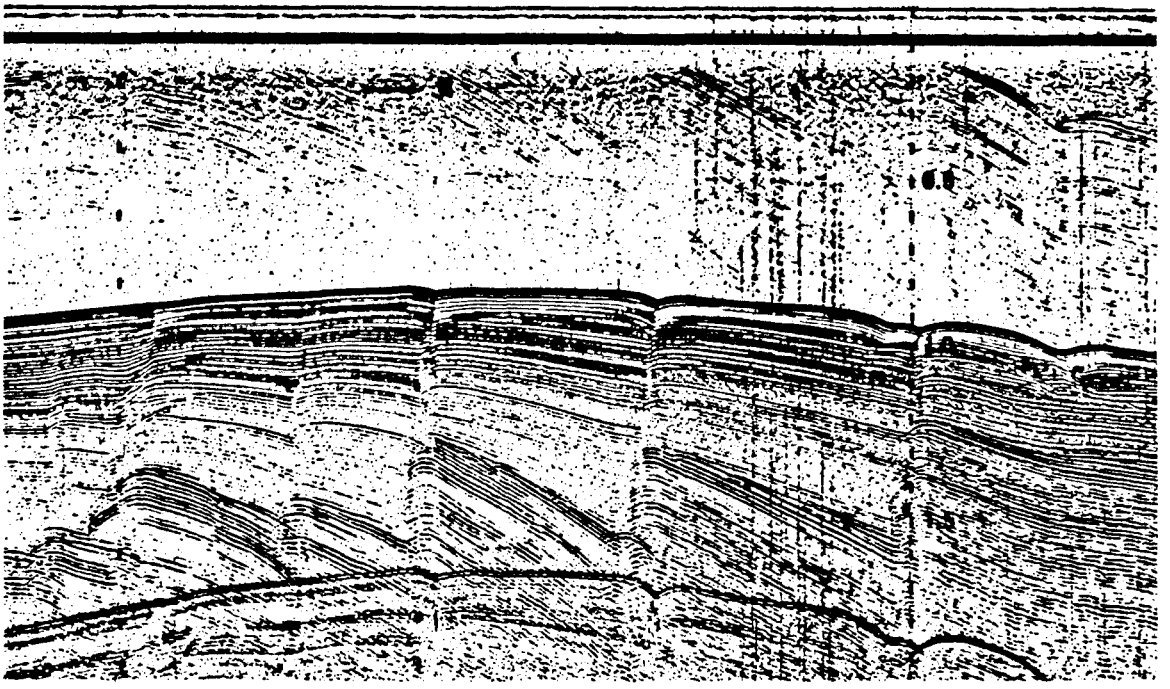


FIG. 1.4-2. A zero-offset section from east Africa with multiple reflections. (Teledyne)

Converted waves should have good diagnostic value in exploration. But the likelihood of seeing converted shear waves in conventional data seems to be so remote that most interpreters have given up looking. Why aren't converted waves seen in conventional data? Some reasons can be offered:

1. In marine data there would have to be an additional conversion to P for the water path.
2. In land data the soil is especially absorptive of shear.
3. Vertical component recorders tend to see P better than S. This is especially so because rays bend toward the vertical in the near surface.

Of all the reasons why converted shear waves should be weaker than pressure waves, none is overwhelming. A wide range of amplitudes are recorded in a wide variety of environments. Data are often displayed with automatic gain control (AGC). Weakened amplitude appears to be insufficient cause for the failure of observation. We should keep looking. Converted waves are certainly more prevalent than our recognition of them. (I have never identified a converted shear wave on conventional recordings).

So although converted shear waves might some day play a significant role in reflection seismology, we now return to the mainstream — how to deal effectively with that which is routinely observable.

Reliability of Reverberation Modeling

The seismological literature contains an abundance of theory to describe seismic waves in layered media. A significant aspect of applied seismology is the general neglect of intrabed reverberation. When a wave reflects from an interface, the strength of the reflected wave is a small fraction, typically less than 10%, of the strength of the incident wave. This reflected wave is the one that is mainly dealt with in this book. However, the reflected wave itself reflects again and again, ad infinitum. For short path geometries, there can be very many of these rays. The question is whether these reverberations can ever amount to enough to make considering them worthwhile. The answer seems to be that although reverberation may be significant, seismologists are rarely able to improve interpretation of reflection survey data with the more complicated theory that is required to incorporate reverberation. A few more details are in Section 5.5.

The situation is somewhat improved when well logs are available, but even then there are serious difficulties. The best possible lateral resolving power, say about 20-50 meters, is obtained after migration. The well log, however, is not a 20-50 meter lateral average of the earth. Next time you see

a highway cut through sedimentary rocks, think of the difference between a point and a lateral average over 20-50 meters. In practice, people smooth the well log (vertical smoothing). Too little smoothing gives too much reverberation. Too much smoothing gives no reverberation. The amount of vertical smoothing is an empirically determined parameter, and results are significantly sensitive to it. Vertical averages of the well log may or may not be a satisfactory approximation to the required horizontal average.

Failure of Newtonian Viscosity

Also remarkable is the failure of basic textbook seismology to explain the observed frequency-dependence of the dissipation parameter Q. The simplest theoretical approach to dissipation is to add a strain-rate term to Hooke's stress-strain law. This predicts stronger relative dissipation of high frequencies than of low frequencies. Experimentally, relative dissipation is observed to be roughly constant over many decades of frequency. Other simple Newtonian theories yield polynomial ratios in $-i\omega$ for the stress/strain ratio. These theories contain scale lengths and characteristic frequencies. They do not predict constant Q. The heterogeneity of the rock at all scales seems to be an essential attribute of a successful theory (Section 4.6).

Philosophy of Inverse Problems

Physical processes are often simulated with computers in much the same way 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 extrapolate in depth z. This is 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 our extrapolation goals are reasonable, stable, or even possible.

The time-evolution problems are often called *forward problems* and the depth-extrapolation problems *inverse problems.* In a *forward problem,* such as one with acoustic waves, it is clear what you need and what you can get. You need the density $\rho(x, z)$ and the incompressibility $K(x, z)$, and you need to know the initial source of disturbance. You can get the wavefield 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 routine observations do not give reasonable estimates of images or maps of ρ and K.

What You Can Get from Reflection Seismology

Luckily, it has been discovered that certain functions of ρ and K can be reliably determined and mapped. The velocity v and the acoustic impedance R are given by the equations

$$v = \sqrt{K/\rho} \tag{1a}$$

$$R = \sqrt{K\rho} \tag{1b}$$

Mathematically, it is a simple job to back-solve equation (1), which gives

$$K = v\,R \tag{2a}$$

$$\rho = \frac{R}{v} \tag{2b}$$

In practice, the solution (2) has little value because the two parameters v and R are seen through nonoverlapping spectral windows. The acoustic impedance R is seen through the typical 10 to 100 Hz spectral window of good quality reflection data. Since the low frequency part of the spectrum is missing, it is common to say that it is not the impedance which is seen, but the gradient $\textit{reflectivity} = c(x, z) = \nabla \log(R)$.

The velocity v is seen through a much smaller window. Observation of it involves studying travel time as shot-to-geophone offset varies and will be described in Chapter 3. With this second window it is hard to discern sixteen independent velocity measurements on a 4-second reflection time axis. So this window goes from zero to about 2 Hz, as depicted in figure 3.

Note that there is an information gap from 2-10 Hz. Even presuming that *rock physics* can supply us with a relationship between ρ and K, the gap seriously interferes with the ability of a seismologist to predict a well log before the well is drilled. What seismologists can do somewhat reliably is predict a filtered log.

The observational situation described above has led reflection seismologists to a specialized use of the word *velocity*. To a reflection seismologist, *velocity* means the low spatial frequency part of "real velocity." The high-frequency part of the "real velocity" isn't called *velocity*: it is called *reflectivity*. Density is generally disregarded as being almost unmeasurable by surface reflection seismology.

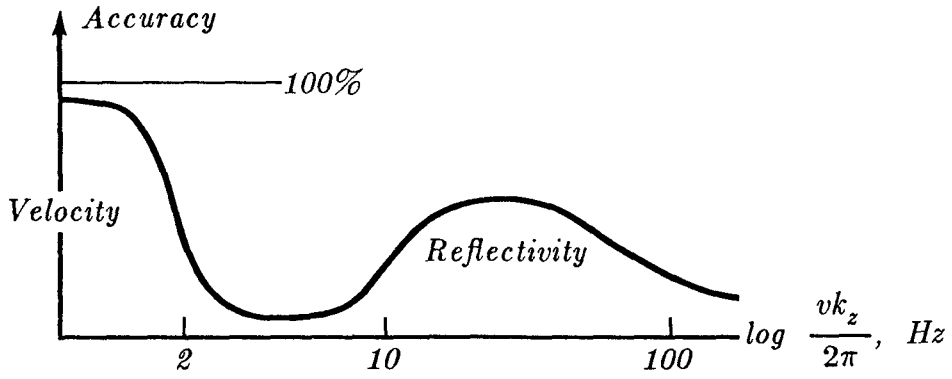


FIG. 1.4-3. Reliability of information obtained from surface seismic measurements.

Mathematical Inverse Problems

In mathematics the solution to an inverse problem has come to mean the "determination" of material properties from wavefields. Often this is achieved with a "convergent sequence." Geophysicists are less precise (or more inclusive) about what they mean by "determination." In Chapters 1-2 of this book reflectors are "determined" by the exploding-reflection concept. In Chapter 3 shot-to-geophone offset is incorporated, and reflectivity $c(x, z)$ and velocity $v(z)$ are "determined" with a buried-experiment concept. In Chapter 5 the concept is developed of suppressing multiple reflections and finding the "true" amplitudes of reflections by having 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 such coincidence is not our goal.

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. Begin with the acoustic case. 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 of momentum conservation says that a small volume within a gas will accelerate if there is an applied force. The force arises from pressure differences at opposite sides of the small volume. Newton's law says

$$\textit{mass} \times \textit{acceleration} \;=\; \textit{force} \;=\; -\ \textit{pressure gradient}$$

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

$$\rho \frac{\partial w}{\partial t} \;=\; -\frac{\partial P}{\partial z} \tag{3b}$$

The second physical process is 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

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

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

In two dimensions it is

$$-\frac{\partial P}{\partial t} \;=\; K \left(\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} \right) \tag{4b}$$

To arrive at the one-dimensional wave equation from (3a) and (4a), first divide (3a) by ρ and take its x-derivative:

$$\frac{\partial}{\partial x} \frac{\partial}{\partial t}\, u \;=\; -\frac{\partial}{\partial x} \frac{1}{\rho} \frac{\partial P}{\partial x} \tag{5}$$

Second, take the time-derivative of (4). In the solid-earth sciences we are fortunate that the material in question does not change during our experiments. This means that K is a constant function of time:

$$\frac{\partial^2 P}{\partial t^2} \;=\; -K \frac{\partial}{\partial t} \frac{\partial}{\partial x}\, u \tag{6}$$

Inserting (5) into (6), the one-dimensional scalar wave equation appears.

$$\frac{\partial^2 P}{\partial t^2} \;=\; K \frac{\partial}{\partial x} \frac{1}{\rho} \frac{\partial P}{\partial x} \tag{7a}$$

In two space dimensions, the exact, acoustic scalar wave equation is

$$\frac{\partial^2 P}{\partial t^2} \;=\; K \left(\frac{\partial}{\partial x} \frac{1}{\rho} \frac{\partial}{\partial x} + \frac{\partial}{\partial z} \frac{1}{\rho} \frac{\partial}{\partial z} \right) P \tag{7b}$$

You will often see the scalar wave equation in a simplified form, in which it is assumed that ρ is not a function of x and z. Two reasons are often given for this approximation. First, observations are generally unable to determine density, so density may as well be taken as constant. Second, we will soon

see that Fourier methods of solution do not work for space variable coefficients. Before examining the validity of this approximation, its consequences will be examined. It immediately reduces (7b) to the usual form of the scalar wave equation:

$$\boxed{\frac{\partial^2 P}{\partial t^2} = \frac{K}{\rho}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}\right) P} \tag{8}$$

To see that this equation is a restatement of the geometrical concepts of previous sections, insert the trial solution

$$P = \exp(-i\omega t + i k_x x + i k_z z) \tag{9}$$

What is obtained is the *dispersion relation of the two-dimensional scalar wave equation:*

$$\frac{\omega^2}{K/\rho} = k_x^2 + k_z^2 \tag{10}$$

Earlier (Section 1.2, equation(8)) an equation like (10) was developed by considering only the geometrical behavior of waves. In that development the wave velocity squared was found where K/ρ stands in equation (10). Thus physics and geometry are reconciled by the association

$$v^2 = \frac{K}{\rho} \tag{11}$$

Last, let us see why Fourier methods fail when the velocity is space variable. Assume that ω, k_x, and k_z are constant functions of space. Substitute (9) into (8) and you get the contradiction that ω, k_x, and k_z must be space variable if the velocity is space variable. Try again assuming space variability, and the resulting equation is still a differential equation, not an algebraic equation like (10).

Evanescence and Ground Roll

Completing the physical derivation of the dispersion relation,

$$k_x^2 + k_z^2 = \frac{\omega^2}{v^2} \tag{12}$$

we can now have a new respect for it. It carries more meaning than could have been anticipated by the earlier geometrical derivation. The dispersion relation was originally regarded merely as $\sin^2\theta + \cos^2\theta = 1$ where $\sin\theta = v\, k_x/\omega$. There was no meaning in $\sin\theta$ exceeding unity, in other

words, in $v\ k_x$ exceeding ω. Now there is. There was a hidden ambiguity in two of the previous migration methods. Since data could be an arbitrary function in the (t, x)-plane, its Fourier transform could be an arbitrary function in the (ω, k_x)-plane. In practice then, there is always energy with an angle sine greater than one. This is depicted in figure 4. What should be done with this energy?

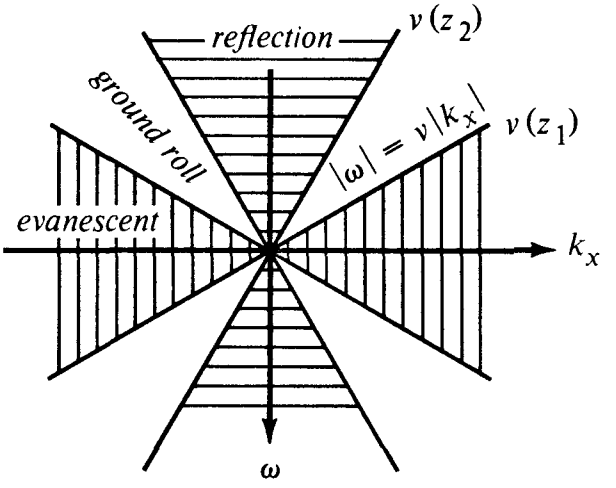


FIG. 1.4-4. The triangle(s) of reflection energy $|\omega| > v(z)\ |k_x|$ become narrower with velocity, hence with depth. Ground roll is energy that is propagating at the surface, but evanescent at depth.

When $v\ k_x$ exceeds ω, the familiar downward-extrapolation expression is better rewritten as

$$e^{\pm i \sqrt{\omega^2/v^2 - k_x^2}\, z} \quad = \quad e^{\pm \sqrt{k_x^2 - \omega^2/v^2}\, z} \tag{13}$$

This says that the depth-dependence of the physical solution is a growing or a damped exponential. These solutions are termed *evanescent* waves. In the most extreme case, $\omega = 0$, k_x is real, and $k_z = \pm i\ k_x$. For elastic waves, that would be the deformation of the ground under a parked airplane. Only if the airplane can move faster than the speed of sound in the earth will a wave be radiated into the earth. If the airplane moves at a subsonic speed the deformation is said to be *quasi-static*.

Perhaps a better physical description is a thought experiment with a sinusoidally corrugated sheet. Such metallic sheet is sometimes 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 $V\ k_x$, 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 it moved very fast, $V > v$, in which case the moving sheet would be radiating sound to great distances. This is why supersonic airplanes use so much fuel.

What should a migration program do with energy that moves slower than the sound speed? Theoretically, such energy should be exponentially damped in the direction going away from the source. The damping in the offending region of (ω, k_x)-space is, quantitatively, extremely rapid. Thus, simple exploding-reflector theory predicts that there should be almost no energy in the data at these low velocities.

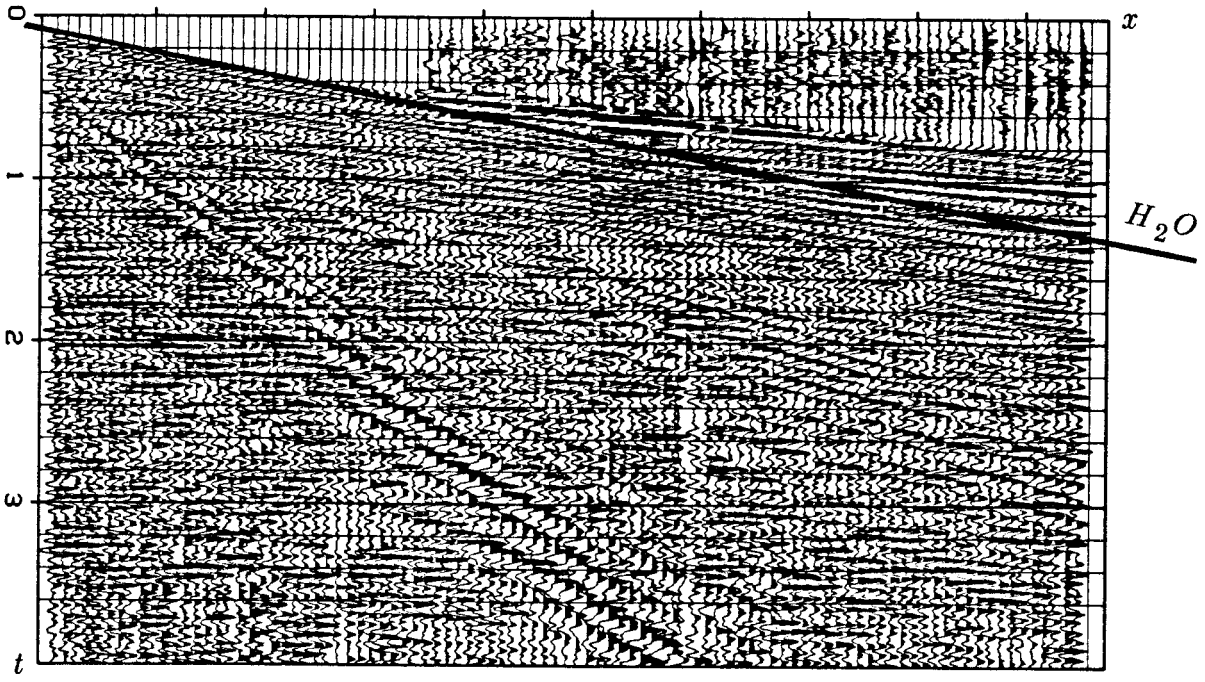


FIG. 1.4-5. Florida shallow marine profile, exhibiting ground roll with frequency dispersion. (Conoco, Yedlin)

The reality is that, instead of tiny amounts of energy in the evanescent region of (ω, k_x)-space, there is often a great deal. This is another breakdown of the exploding-reflector concept. The problem is worst with land data. Waves that are evanescent in deep, fast rocks of interest can be propagating in the low-velocity soil layer. This energy is called *ground roll.* Figure 5 shows an example. Like the surface of the earth, which varies greatly from place to place, the immediate subsurface which controls the ground roll varies substantially. So although figure 5 is a nice example, no example can really be typical. This data is not a zero-offset section. The shot is on the left, and the traces to the right are from geophones at increasing distances from the shot. The straight line drawn onto the data defines a slope equal to the water velocity. Steeper events are ground roll. In this figure there are two types of ground roll, one at about half of water velocity, and a stronger one at about a quarter of water velocity. The later and stronger one shows an interesting feature known as *frequency dispersion.* Viewing the data from the side, you should be able to notice that the high frequencies arrive before the low frequencies.

Ground roll is unwanted noise since its exponential decay effectively prevents it from being influenced by deep objects of interest. In practice, energy in the offending region of (ω, k_x)-space should be attenuated. A mathematical description is to say that the composite mapping from model space to data space and back to model space again is not an identity transformation but an idempotent transformation.

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, and we know that either can cause a reflection. So it is perhaps a little surprising that while the density derivative is explicitly found in (7b), the incompressibility derivative is not. This means that dropping the density gradients in (7b) will not eliminate all possible reflections. Dropping the terms will slightly simplify further analysis, however, and since constant density is a reasonable case, the terms are often dropped.

There are also some well-known mathematical circumstances under which the first-order terms may be ignored. Fix your attention on a wave going in any particular direction. Then ω, k_x, and k_z have some prescribed ratio. In the limiting case that frequency goes to infinity, the P_{tt}, P_{xx}, and P_{zz} terms in (8) all tend to the *second* power of infinity. Suppose two media

gradually blend into one another so that $\partial\rho/\partial x$ is less than infinity. The terms neglected in going from (7b) to (8) are of the form $\rho_x P_x$ and $\rho_z P_z$. As frequency tends to infinity, these terms only tend to the *first* power of infinity. Thus, in that limit they can be neglected.

These terms are usually included in theoretical seismology where the goal is to calculate synthetic seismograms. But where the goal is to create earth models from seismic field data — as in this book — these terms are generally neglected. Earth imaging is more difficult than calculating synthetic seismograms. But often the reason for neglecting the terms is simply to reduce the clutter. These terms may be neglected for the same reason that equations are often written in two dimensions instead of three: the extension is usually possible but not often required. Further, these terms are often ignored to facilitate Fourier solution techniques. Practical situations might arise for which these terms need to be included. With the finite-difference method (Section 2.2), they are not difficult to include. But any effort to include them in data processing should also take into account other factors of similar significance, such as the assumption that the acoustic equation approximates the elastic world.

EXERCISES

1. Soil is typically saturated with water below a certain depth, which is known as the *water table*. Experience with hammer-seismograph systems shows that seismic velocity typically jumps up to water velocity ($V_{H_2O} = 1500\ m/s$) at the water table. Say that in a certain location, the ground roll is observed to be greater than the reflected waves, so a decision has been made to bury geophones. The troublesome ground roll is observed to travel at six-tenths the speed of a water wave. How deep must the geophones be buried below the water table to attenuate the ground roll by a factor of ten? Assume the data contains all frequencies from 10 to 100 Hz. (Hints: $\log_e 10 \approx 2$, $2\pi \approx 6$, etc.)

2. Consider the function

$$P(z,t) = P_0 \frac{1}{\sqrt{Y(z)}} e^{i\omega t - i\int_0^z \omega \sqrt{\frac{\rho(\xi)}{K(\xi)}}\, d\xi} \qquad \text{(1.4E1)}$$

where

$$P_0 = \text{constant}$$

$$Y \equiv \frac{1}{\sqrt{\rho(z)\,K(z)}}$$

as a trial solution for the one-dimensional wave equation:

$$\left(\frac{\partial^2}{\partial t^2} - \frac{K(z)}{\rho(z)} \frac{\partial^2}{\partial z^2} \right) P = - \frac{K(z)}{\rho(z)^2} \frac{\partial \rho}{\partial z} \frac{\partial P}{\partial z} \qquad (1.4E2)$$

Substitute the trial solution (E1) into the wave equation (E2). Discuss the trade-off between changes in material properties and the validity of your solution for different wavelengths.

1.5 The Paraxial Wave Equation

The scalar wave equation, unlike Fourier equations, allows arbitrary spatial variation in density and velocity. Because of this you might expect that it would be used directly in the manufacture of migrated sections. But it is used little for migration, and we will first review why this is so. Then we will meet the *paraxial wave equation,* which is the basis for most production migration.

Philosophically, the paraxial wave equation is an intermediary between the simple concepts of rays and plane waves and deeper concepts embodied in the wave equation. (The paraxial wave equation is also called the *single-square-root* equation. In Chapter 2, a specialization of it is called the *parabolic* wave equation). The derivation of the parabolic wave equation does not proceed from simple concepts of classical physics. Its development is more circuitous, like the Schroedinger equation of quantum physics. You must study it for a while to see why it is needed. When I introduced the parabolic wave equation to seismic calculations in 1970, it met with considerable suspicion. Fortunately for you, years of experience have enabled me to do a better job of explaining it, and fortunately for me, its dominance of the industrial scene will give you the interest to persevere.

The paraxial equation will be introduced by means of Fourier methods. Fourier methods are incompatible with space-variable coefficients. Since we want to incorporate spatial variations in velocity, this limitation is ultimately to be avoided, so after getting the paraxial equation in the Fourier domain, ik_z is replaced by $\partial/\partial z$, and ik_x is replaced by $\partial/\partial x$. Then, being in the

space domain, the velocity can be space-variable. The result is a partial differential equation often solved by the finite-differencing method. This procedure turns out to be valid, but new students of migration understandably regard it with misgiving. Thus, the final part of this section is a derivation of the paraxial wave equation which makes no use of Fourier methods.

Why the Scalar Wave Equation is Rarely Used for Migration

Life would be simpler if migration could be done with the scalar wave equation instead of the paraxial equation. Indeed, migration can be done with the scalar wave equation, and there are some potential advantages (Kosloff and Baysal [1983]). But more than 99% of current industrial migration is done with the paraxial equation.

The main problem with the scalar wave equation is that it will generate unwanted internal multiple reflections. The exploding-reflector concept cannot deal with multiple reflections. Primary reflections can be modeled with only upcoming waves, but multiple reflections involve both up and downgoing paths. The multiple reflections observed in real life are completely different from those predicted by the exploding-reflector concept. For the sea-floor multiple reflection, a sea-floor two-way travel-time depth of t_0 yields sea-floor multiple reflections at times $2t_0$, $3t_0$, $4t_0$, $\cdots$. In the exploding-reflector conceptual model, a sea-floor one-way travel-time depth of t_0 yields sea-floor multiple reflections at times $3t_0$, $5t_0$, $7t_0$, $\cdots$. In building a telescope, microscope, or camera, the designer takes care to suppress backward reflected light because it creates background noise on the image. Likewise, in building a migration program we do not want to have energy moving around that does not contribute to the focused image. The scalar wave equation with space-variable coefficients will generate such energy. This unwanted energy is especially troublesome if it is coherent and migrates to a time when primaries are weak. It is annoying, as the reflection of a bright window seen on a television screen is annoying. So if you were trying to migrate with the scalar wave equation, you would make the velocity as smooth as possible.

Another difficulty of imaging with the scalar wave equation arises with evanescent waves. These are the waves that are exponentially growing or decaying with depth. Nature extrapolates waves forward in time, but we are extrapolating them in depth. Growing exponentials can have tiny sources, even numerical round-off, and because they grow rapidly, some means must be found to suppress them.

A third difficulty of imaging with the scalar wave equation derives from initial conditions. The scalar wave equation has a second depth z-derivative. This means that two boundary conditions are required on the z-axis. Since

data is recorded at $z=0$, it seems natural that these boundary conditions should be knowledge of P and $\partial P/\partial z$ at $z=0$. But $\partial P/\partial z$ isn't recorded.

Luckily, in building an imaging device that operates wholly within a computer, we have ideal materials to work with, i.e., reflectionless lenses. Instead of the scalar wave equation of the real world we have the paraxial wave equation.

Fourier Derivation of the Paraxial Wave Equation

Start from the dispersion relation of the scalar wave equation:

$$k_x^2 + k_z^2 = \frac{\omega^2}{v^2} \tag{1}$$

Take a square root.

$$k_z = \pm \sqrt{\frac{\omega^2}{v^2} - k_x^2} \tag{2}$$

The simple act of selecting the minus sign in (2) includes the upcoming waves and eliminates the downgoing waves. Equation (1) is the three-dimensional Fourier transform of the scalar wave equation. Inverse transforming (2) will give us an equation for upcoming (or downgoing) waves only, without the other. Inverse Fourier transformation over a dimension is just a matter of selecting one or more of the following substitutions:

$$\frac{\partial}{\partial t} = -i\,\omega \tag{3a}$$

$$\frac{\partial}{\partial x} = i\,k_x \tag{3b}$$

$$\frac{\partial}{\partial z} = i\,k_z \tag{3c}$$

After inverse transformation over z there is a differential equation in z in which the velocity may be taken to be z-variable. Likewise for x. Any equation resulting from any of the substitutions of (3) into (2) is called a paraxial equation. Chapter 2 of this book goes into great detail about the meaning of these equations. Before beginning this interpretation the paraxial wave equation will be derived without the use of Fourier transformation. Besides giving a clear path to the basic migration equation, this derivation also gives a better understanding of what the equation really does, and how it differs from the scalar wave equation.

Snell Waves

It is natural to begin studies of waves with equations that describe plane waves in a medium of constant velocity. However, in reflection seismic surveys the velocity contrast between shallowest and deepest reflectors ordinarily exceeds a factor of two. Thus depth variation of velocity is almost always included in the analysis of field data. Seismological theory needs to consider waves that are just like plane waves except that they bend to accommodate the velocity stratification $v(z)$. Figure 1 shows this in an idealized geometry: waves radiated from the horizontal flight of a supersonic airplane.

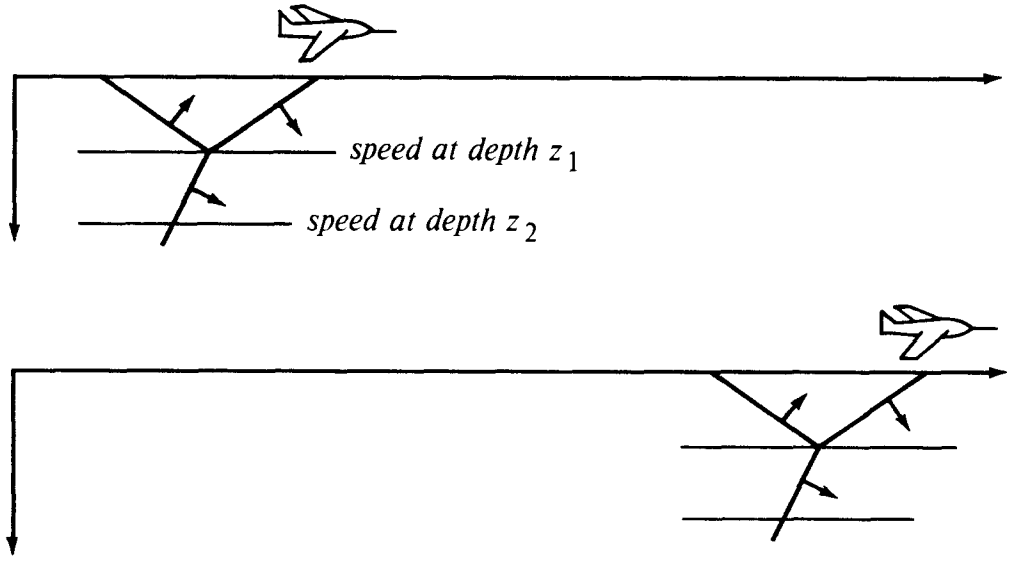


FIG. 1.5-1. Fast airplane radiating a sound wave into the earth. From the figure you can deduce that $\partial t / \partial x$ is the same at depth z_1 as it is at depth z_2. This leads (in isotropic media) to Snell's law.

The airplane flies horizontally at a constant speed. It goes from $x = -\infty$ to $x = +\infty$. Imagine an earth of horizontal plane layers. In this model there is nothing to distinguish any point on the x-axis from any other point on the x-axis. But the seismic velocity varies from layer to layer. There may be reflections, head waves, shear waves, and multiple reflections. Whatever the picture is, it moves along with the airplane. A picture of the wavefronts near the airplane moves along with the airplane. The top of the picture and the bottom of the picture both move laterally at the same speed even if the earth velocity increases with depth. If the top and bottom didn't go at the same speed, the picture would become distorted, contradicting the presumed symmetry of translation. This horizontal speed, or rather its

inverse $\partial t / \partial x$, has several names. In practical work it is called the *stepout*. In theoretical work it is called the *ray parameter*. It is very important to note that $\partial t / \partial x$ does not change with depth, even though the seismic velocity does change with depth. In a constant-velocity medium, the angle of a wave does not change with depth. In a stratified medium, $\partial t / \partial x$ does not change with depth.

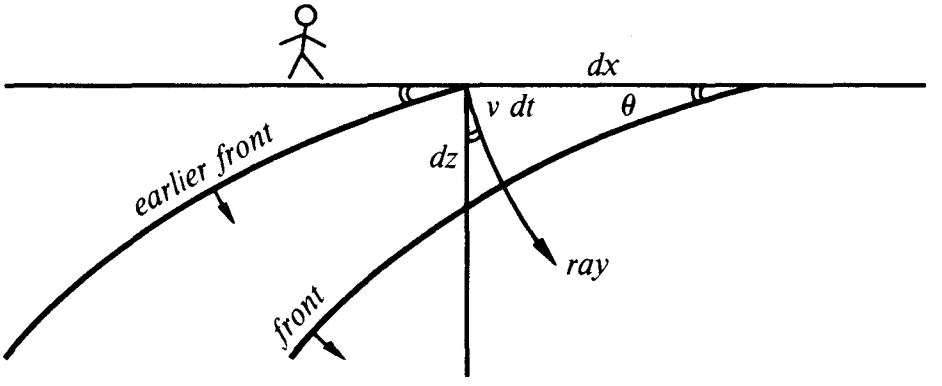


FIG. 1.5-2. Downgoing fronts and rays in stratified medium $v(z)$. The wavefronts are horizontal translations of one another.

Figure 2 illustrates the differential geometry of the wave. The diagram shows that

$$\frac{\partial t}{\partial x} = \frac{\sin \theta}{v} \tag{4a}$$

$$\frac{\partial t}{\partial z} = \frac{\cos \theta}{v} \tag{4b}$$

These two equations define two (inverse) speeds. The first is a horizontal speed, measured along the earth's surface, called the *horizontal phase velocity*. The 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. Projection of wave *fronts* onto coordinate axes gives speeds larger than v, whereas projection of *rays* onto coordinate axes gives speeds smaller than v. The inverse of the phase velocities is called the *stepout* or the *slowness*.

Snell's law relates the angle of a wave in one layer with the angle in another. The constancy of equation (4a) in depth is really just the statement of Snell's law. Indeed, we have just derived Snell's law. All waves in seismology propagate in a velocity-stratified medium. So they cannot be called plane

waves. But we need a name for waves that are near to plane waves. A *Snell wave* will be defined to be the generalization of a plane wave to a stratified medium $v(z)$. A plane wave that happens to enter a medium of depth-variable velocity $v(z)$ gets changed into a Snell wave. While a plane wave has an angle of propagation, a Snell wave has instead a *Snell parameter* $p = \partial t / \partial x$.

It is noteworthy that Snell's parameter $p = \partial t / \partial x$ is directly observable at the surface, whereas neither v nor θ is directly observable. Since $p = \partial t / \partial x$ is not only observable, but constant in depth, it is customary to use it to eliminate θ from equation (4):

$$\frac{\partial t}{\partial x} = \frac{\sin\theta}{v} = p \tag{5a}$$

$$\frac{\partial t}{\partial z} = \frac{\cos\theta}{v} = \left(\frac{1}{v(z)^2} - p^2 \right)^{1/2} \tag{5b}$$

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

$$t(x,z) = \frac{\sin\theta}{v}\, x + \int_0^z \frac{\cos\theta}{v}\, dz \tag{6a}$$

$$t(x,z) = p\, x + \int_0^z \left(\frac{1}{v(z)^2} - p^2 \right)^{1/2} dz \tag{6b}$$

The validity of (6b) is readily checked by computing $\partial t / \partial x$ and $\partial t / \partial z$, then comparing with (5).

An arbitrary waveform $f(t)$ may be carried by the Snell wave. Use (6) to define a delay time t_0 for a delayed wave $f[t - t_0(x,z)]$ at the location (x,z).

$$SnellWave(t,x,z) = f\left\{ t - p\, x - \int_0^z \left(\frac{1}{v(z)^2} - p^2 \right)^{1/2} dz \right\} \tag{7}$$

Time-Shifting Equations

An important task is to predict the wavefield 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

$$\frac{\partial P(t,z)}{\partial z} = -\frac{1}{v} \frac{\partial P(t,z)}{\partial t} \tag{8}$$

as may be readily verified by substituting the trial solutions

$$P(t,z) = f\left(t - \frac{z}{v}\right) \qquad \text{for constant } v \tag{9}$$

or

$$P(t,z) = f\left(t - \int_0^z \frac{dz}{v(z)}\right) \qquad \textit{for } v(z) \tag{10}$$

This also works for nonvertically incident waves with the partial differential equation

$$\frac{\partial P(t,x,z)}{\partial z} = -\frac{\partial t}{\partial z}\,\frac{\partial P(t,x,z)}{\partial t} \tag{11}$$

which has the solution

$$P(t,x,z) = f\left(t - p\,x - \int_0^z \frac{\partial t}{\partial z}\,dz\right) \tag{12}$$

In interpreting (11) and (12) recall that $1/(\partial t/\partial z)$ is the apparent velocity in a borehole. The partial derivative of wavefield $P(t,x,z)$ 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 holds only when a *single* Snell wave is present; that is, the *same* arbitrary time function must be seen at all locations.

Substitution from (5) also enables us to rewrite (11) in the various forms

$$\frac{\partial P(t,x,z)}{\partial z} = -\frac{\cos\theta}{v}\,\frac{\partial P(t,x,z)}{\partial t} \tag{13a}$$

$$\frac{\partial P(t,x,z)}{\partial z} = -\left(\frac{1}{v(z)^2} - p^2\right)^{1/2}\,\frac{\partial P(t,x,z)}{\partial t} \tag{13b}$$

$$\frac{\partial P(t,x,z)}{\partial z} = -\left[\frac{1}{v(z)^2} - \left(\frac{\partial t}{\partial x}\right)^2\right]^{1/2}\,\frac{\partial P(t,x,z)}{\partial t} \tag{13c}$$

Equation (13) is a paraxial wave equation. Since $\partial t/\partial x = p$ can be measured along the surface of the earth, it seems that equation (13c), along with an assumed velocity $v(z)$ and some observed data $P(t,x,z=0)$, would enable us to determine $\partial P/\partial z$, which is the necessary first step of downward continuation. But the presumption was that there was only a single Snell wave and not a superposition of several Snell waves. Superposition of different waveforms on different Snell paths would cause different time functions to be seen at different places. Then a mere time shift would not achieve downward continuation. Luckily, a complicated wavefield that is variable

from place to place may be decomposed into many Snell waves, each of which can be downward extrapolated with the differential equation (13) or its solution (12). One such decomposition technique is *Fourier analysis.*

Fourier Decomposition

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

$$\boxed{\frac{\partial t}{\partial x} \;=\; \frac{\sin\theta}{v} \;=\; p \;=\; \frac{k_x}{\omega}} \tag{14}$$

In theoretical seismology a square-root function often appears as a result of using (14) to make a cosine.

Utilization of this Fourier domain interpretation of Snell's parameter p enables us to write the square-root equation (13) in an even more useful form. But first the square-root equation must be reexpressed in the Fourier domain. This is done by replacing the $\partial/\partial t$ operator in (13) by $-i\,\omega$. The result is

$$\frac{\partial P\,(\omega,\,k_x\,,\,z\,)}{\partial z} \;=\; +\,i\,\omega \left[\; \frac{1}{v\,(z\,)^2} \;-\; \frac{k_x^{\,2}}{\omega^2} \;\right]^{1/2} P\,(\omega,\,k_x\,,\,z\,) \tag{15}$$

At present it is equivalent to specify either the differential equation (15) or its solution (12) with f as the complex exponential:

$$P\,(\omega,\,k_x\,,\,z\,)\,e^{-i\,\omega t} \;=\; \exp\left\{\; i\,\omega \int_0^z \left[\; \frac{1}{v\,(z\,)^2} \;-\; \frac{k_x^{\,2}}{\omega^2} \;\right]^{1/2} dz \;\right\} \tag{16}$$

Later, when we consider lateral velocity variation $v\,(x\,)$, the solution (16) becomes wrong, whereas the differential equation (13c) is a valid description of any *local* plane-wave behavior. But before going to lateral velocity gradients we should look more carefully at vertical velocity gradients.

Velocity Gradients

Inserting the Snell wavefield expression into the scalar wave equation, we discover 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. The paraxial wave equation could be modified to incorporate a transmission coefficient effect. The reason it rarely is modified may be the same reason that density gradients are often ignored. They add clutter to equations while their contribution to better results — namely, more correct amplitudes and possible tiny phase shifts — has marginal utility. Indeed, if they are included, then other deeper questions should also be included, such as the question, why use the acoustic equation instead of various other forms of scalar elastic equations?

Even if the paraxial wave equation were modified to incorporate a transmission coefficient effect, its solution would still fail to satisfy the scalar wave equation because of the absence of the reflected wave. But that is just fine, because it is the paraxial equation, with its reflection-free lenses, that is desired for data processing.

EXERCISES

1. Devise a mathematical expression for a plane wave that is an impulse function of time with a propagation angle of 15° 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 (12), yields an approximate solution to the scalar wave equation for stratified media $v(z)$. For $p = 0$, the solution should reduce to the solution of Exercise 2 in Section 1.4.

1.6 Mastery of 2-D Fourier Techniques

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

Signs and Scales in Fourier Transforms

In Fourier transforming t-, x-, and z-coordinates, a sign convention must be chosen for each coordinate. Electrical engineers have chosen one convention and physicists another. While both have good reasons for their choices, our circumstances more closely resemble the circumstances of physicists, so their convention will be used. For the *inverse* Fourier transformation this is

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

For the *forward* Fourier transform, the space variables carry a negative sign and time carries a positive sign. The limits on the integrations and the scale factor in the continuous case differ from the discrete case. We rarely do the transforms analytically in either case. Since the extra notation required for limits and scales usually clutters rather than clarifies a discussion, they will be omitted altogether except when they play a useful role.

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. Someone using the approach of "changing the signs around until it works" is likely to be perplexed by the number of possible permutations. There are good reasons for the sign conventions chosen by physicists, and once the reasons are known, it is easy to remember the conventions.

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 our convention will be always to 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. This statement applies to both the forward and the inverse transform.

This leaves the choice of whether to use the positive sign for the time axis or the space axes. There are many space axes but only one time axis.

There will be the fewest number of minus signs and the fewest sign changes if the spatial gradient $\partial/\partial x$, $\partial/\partial z$, etc. is chosen to be associated with the positive k-vector, i.e., with ik_x, ik_z, etc. Of course, this leaves the time derivative with $-i\omega$.

This sign convention brings our practice into conflict 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 have of course used their own sign convention. It doesn't matter for the programs that transform complex-valued time functions to complex-valued frequency functions, because then the sign convention is under the user's control. But it does make a difference with the program that converts real time functions to complex frequency functions. 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 from 0 to $-\pi$. Alternately, you could always take the complex conjugate of the transform, which would swap the sign of the ω-axis. With the Stolt algorithm it is common to transform space first. Then the array processor convention turns out to have our notation.

How to Transpose a Big Matrix

It is lucky that very large matrices can easily be transposed. This is what makes wave-equation seismic data processing reasonable on a small minicomputer. By *very large* matrix, I mean one that is too big to fit in a computer's random access memory (RAM). If two copies of the data fit in the RAM, then transposition is simply the copy operation $T(i, j) = M(j, i)$.

The transpose algorithm for very large matrices 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 following order (let ace be denoted by 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 onto pile A' and pile B'. You see

Pile A': $1a$ $2a$ $3a$ $\cdots$ $8a$ $1b$ $2b$ $\cdots$ $8b$

Pile B': $1c$ $2c$ $3c$ $\cdots$ $8c$ $1d$ $2d$ $\cdots$ $8d$

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

Now I will try the inverse transpose. Note 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 transpose in only two passes rather than three. I just do everything backwards. I start with piles A' and B'. Then I create pile A by alternately selecting cards one from pile A' and one from pile B'. Likewise I create pile B. Then I repeat this procedure. The first algorithm is called the *sort* algorithm, and the second is called the *merge* algorithm. With these two algorithms, 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 four card piles, techniques could be developed 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 can be factored into primes, etc. But this takes us too far afield.

Minimizing the number of passes over the data turns out to maximize the number of tapes. In reality you won't be using real tapes when you are transposing. Instead you will 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 to the speed of sequential transfers.

Rocca's 2-D Fourier Transform without Transposing

The most direct method of two-dimensional Fourier transformation in a computer is the repetitive application of a one-dimensional Fourier transform method. The 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 the repetitive use of a one-dimensional program. Now for the rows. If the matrix fits in the RAM, then everything is

easy: one row at a time can be copied into a vector; the vector can be Fourier transformed and then copied back into the matrix row. The more typical case is that the data doesn't fit in the RAM but does fit in the "virtual" memory. This means that the programmer could write $T(i, j) = M(j, i)$ but the program would run prohibitively slowly because an entire page of virtual memory would be fetched from disk just to find a single number.

Conceptually, an easy way to handle the transformation over the row direction is to transpose the matrix, transform each column, and transpose back. Fabio Rocca suggested a quicker and easier means of Fourier transformation over the row index. The basic Fourier transform program has certain overhead calculations, such as computing or fetching sines and cosines. Ordinarily, these overhead calculations are repeated each time a Fourier transformation is performed. With Rocca's method the overhead calculations are done just once, and all the rows get Fourier transformed. So it is even quicker than the straightforward approach. The method follows.

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, just modify an ordinary one-dimensional Fourier transform program by replacing each scalar add or multiply operation by the same operation on every element in the corresponding column.

The order in which data is accessed makes Rocca's row algorithm efficient in a virtual memory environment. Before the days of virtual memory, we implemented the Rocca row algorithm with reads and writes around the inner loops.

To illustrate Rocca's method, a row Fourier transformation program was written based on the one-dimensional Fourier transformation program found in FGDP. It is included in the next section. That program transforms complex time functions to complex frequency functions. If you should decide to write a real-to-complex Fourier transform, you should beware of the assumption that real and imaginary parts are stored contiguously. This assumption is true for the column index, but not for the row index.

1.7 Sample Programs

The programs in this section generated many of the examples given in this book. They were written for clarity and brevity, and they are excellent for experimental work. Good production programs will be faster (by factors of from 1.01 to about 4). Speed can be gained by taking advantage of various special circumstances. For example, data is real, but these expository programs assume it to be complex.

RATional FORtran = Ratfor

Bare bones Fortran is our most universal computer language. But it is hardly appropriate for expository discussion of algorithms. The ideal expository language is *Ratfor*. Ratfor is *Rat*ional *For*tran, namely, Fortran without the blemishes. Ratfor programs (including the Ratfor preprocessor) are readily converted to Fortran by means of the Ratfor preprocessor. Since the preprocessor is publicly available Ratfor is practically as universal as Fortran.†

You won't really need the preprocessor or any precise definitions if you already know Fortran or almost any other computer language, because then the Ratfor language will be easy to understand. Statements on a line may be separated by ";". Statements may be grouped together with { }. Do loops don't require statement numbers because { } defines the range. Given that "if ( )" is true, the statements in the following { } are done. "Else { }" does what you would expect it to. Indentation is used for readability. Choose your own style. I have overcondensed. Anything following # is a comment. You may omit the braces { } when they contain only one statement. "Break" will cause premature termination of the enclosing { }. "Break 2" escapes from {{ }}. "While ( ) { }" repeats the statements in { } while the condition ( ) is true. "Repeat { } until ( )" is a loop that tests at the bottom. A looping statement more general than "do" is "for(initialize; condition; reinitialize) { }". "Next" causes skipping to the end of any loop and a retrial of the test condition. The Fortran relational operators .gt., .ge., .ne., etc. may be written >, >=, !=, etc. The logical operators .and. and .or. may be written & and |. Anything that doesn't make sense to the Ratfor preprocessor, such as

† Kernighan, B.W. and Plauger, P.J., 1976, *Software Tools:* Addison-Wesley Publishing Company.

Fortran input-output, is passed through without change.

Two-Dimensional Fourier Transformation

Two-dimensional Fourier transforms are based on the one-dimensional Fourier transform. An extremely rapid way to compute the one-dimensional Fourier transform exists and is called the *Cooley-Tukey algorithm* or the *Fast Fourier Transform.* Unfortunately it bears little resemblance to the Fourier integral. This method is so fast and effective that you will hardly ever see the transform being done in the obvious way. All functions are taken to be periodic, so physically transient functions must be regarded as functions of very long period. Usually there is the further restriction that the period must be exactly 2^N points long, where N is an integer. To understand this program, you should look at FGDP or any number of electrical engineering books. To write and use two-dimensional Fourier transform programs, it is only necessary to know the one-dimensional definition of inputs and outputs. Figure 1 shows that humans like to have $t=0$ in the middle of the time axis and $\omega=0$ in the middle of the frequency axis, whereas the standard one-dimensional Fourier transform programs place $t=0$ and $\omega=0$ at one end of a vector.

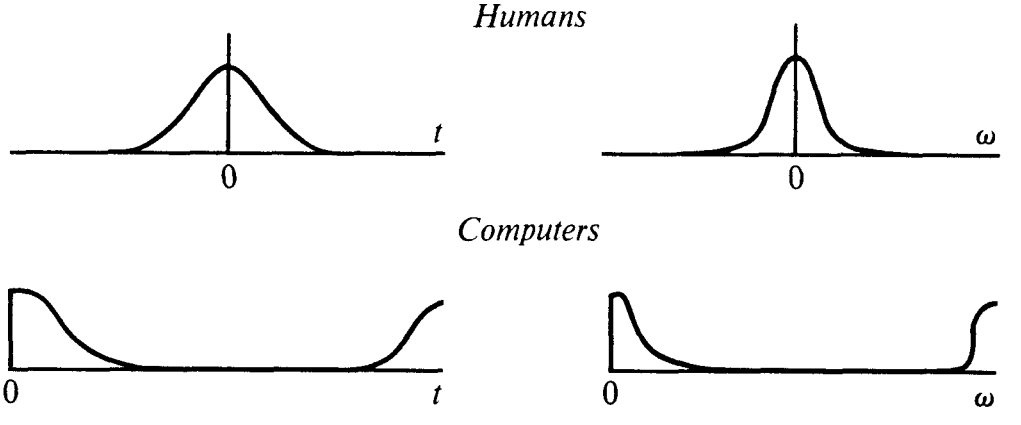


FIG. 1.7-1. Computer storage arrangement in one-dimensional Fourier transform programs.

Take the one-dimensional Fourier transform of an eight-point time function. The zero frequency is output in the first vector element. The Nyquist frequency π, which is the highest frequency representable on a mesh, namely the function $+1, -1, +1, -1, \cdots$, is in the fifth element of the eight point function, after which follow the negative frequencies. The smallest nonzero

negative frequency is in the eighth vector element. If there were a ninth element, it would by periodicity be equal to the first element. It is a common beginner's error to find that the output of a migration is not real. The imaginary part should be about 10^{-6} of the real part, as expected for single precision arithmetic. A much larger imaginary part, proportional to $1/N$ where N is the vector length, indicates a programming error.

Below is the test program for the two-dimensional program. The "write" statement is local Fortran, not Ratfor. The function being transformed is something of a low-frequency function on the time axis, and very much a low-frequency function on the space axis.

```
# Test case for two-dimensional Fourier Transformation
integer it,nt,ix,nx;        complex cp(64,64), cwork(64)
open(4,file='plotfile',status='new',access='direct',form='unformatted',recl=1)
nx = 64;        nt = 64;
do it=1,nt
        do ix=1,nx
                cp(it,ix)=0.
cp(16,3)=1.; cp(16,4)=4.; cp(16,5)=6.; cp(16,6)=4.; cp(16,7)=1.
cp(17,3)=1.; cp(17,4)=4.; cp(17,5)=6.; cp(17,6)=4.; cp(17,7)=1.
call ft2d(nt,nx,cp,+1.,+1.,cwork)
write(4,rec=1) ((real(cp(it,ix)),it=1,nt),ix=1,nx)
stop;    end
```

The most basic two-dimensional Fourier transform is shown below.

```
# 2D Fourier transform by using 1D program
subroutine ft2d (n1,n2,cp,sign1,sign2,cwork)
complex  cp(n1,n2),cwork(n2)
integer  n1,n2
real     sign1,sign2
do i2 = 1,n2                        # transform over the fast dimension
       call fork (n1,cp(1,i2),sign1)        # one-dimensional Fourier transform
do i1 = 1,n1 {                              # transform over the slow dimension
       do i2 = 1,n2
              cwork(i2) = cp(i1,i2)
       call fork (n2,cwork,sign2)  # one-dimensional Fourier transform
       do i2 = 1,n2
              cp(i1,i2) = cwork(i2)
       }
return;        end
```

Finally we have the one-dimensional fast Fourier transform program. This one is the Ratfor version of Fortran "fork" found in FGDP on p.12. As usual, lx is a power of 2, the output cx(1) is the zero frequency, cx(lx/2+1) is the so-called Nyquist frequency, and cx(lx) is the smallest negative frequency. The algorithm is short, but tricky, and you should not expect the program to

be readable unless you consult other references.

```
# 1D fast Fourier transform
subroutine fork(lx,cx,signi)
complex cx(lx),carg,cexp,cw,ct
j = 1; k = 1;        sc = sqrt(1./lx)
do i = 1,lx {
        if (i<=j) { ct=cx(j)*sc;    cx(j)=cx(i)*sc;    cx(i)=ct }
        m = lx/2
        while (j>m) { j=j-m;    m=m/2;    if (m<1) break }
        j = j+m
        }
repeat {
        istep = 2*k
        do m = 1,k {
                carg = (0.,1.)*(3.14159265*signi*(m-1))/k;        cw = cexp(carg)
                do i = m,lx,istep
                        { ct=cw*cx(i+k);    cx(i+k)=cx(i)-ct;    cx(i)=cx(i)+ct }
                }
        k = istep
        } until(k>=lx)
return;         end
```

Fourier transforms have both real and imaginary parts. Sometimes both are displayed. Often the imaginary part is ignored. This is because most of our time functions vanish before $t=0$. Thus, their Fourier transforms must satisfy certain conditions, namely, real and imaginary parts must be related by Hilbert transform. Locally, one often looks like cosine, the other like sine. So, seeing the real part, it is often easy to imagine the imaginary part. Figure 2 shows the output of the test program.

Stolt Migration

The Stolt migration program shown next uses linear interpolation to convert the ω-axis to the k_z-axis. The scaling by $dk_z/d\omega$ has little effect, so it was omitted to shorten the program. (Something needed to be saved for the exercises). The test case is to make semicircles from impulses.

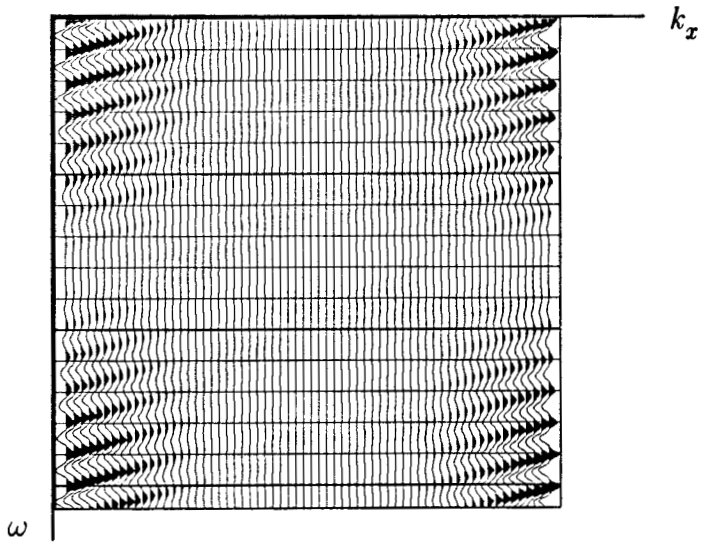


FIG. 1.7-2. Output of two-dimensional Fourier transformation test program.

```
# Test case for Stolt migration.
integer it,nt,ix,nx;  real vdtodx;  complex cp(256,64)
open(4,file='plotfile',status='new',access='direct',form='unformatted',recl=1)
nx = 64;     nt = 256;     vdtodx = 1./4.              # vdtodx = v dt / dx
do it=1,nt
        do ix=1,nx
                cp(it,ix)=0.
cp(32,9)=1.;     cp(64,17)=1.;     cp(128,33)=1.
call stolt(nt,nx,cp,vdtodx)
write(4,rec=1) ((real(cp(it,ix)),it=1,nt),ix=1,nx)
stop;  end
```

```
# Stolt migration subroutine without cosine weight.
subroutine stolt(nt,nx,cp,vdtodx)
integer ikx,nx,nt,nth,iktau,iom
real om,vkx,wl,wh,aktau,pi,pionth,vdtodx
complex cp(nt,nx),cbf(1025)
pi = 3.14159265;    nth=nt/2;    pionth = pi/nth;
call ft2d(nt,nx,cp,1.,-1.,cbf)
do ikx = 1,nx {
        vkx = (ikx-1)*2.*pi*vdtodx/nx
        if ( ikx > nx/2 ) vkx = 2.*pi*vdtodx-vkx        # negative k_x
        cbf(1) = 0.;            cbf(nt+1)=0.            # cbf = working buffer
        do iom = 1,nt
                cbf(iom) = cp(iom,ikx)          # Omit weighting
        cp(1,ikx)=0.                            # Ignore zero freq
        do iktau = 2,nth+1 {                            # Stretch
                aktau = (iktau-1.01)*pionth
                om = sqrt(aktau*aktau+vkx*vkx);         iom = 1+om/pionth
                if(iom<nth) {
                        wl = iom-om/pionth;     wh = 1.-wl
                        cp(iktau,ikx)     = wl*cbf(iom)    +wh*cbf(iom+1)
                        cp(nt-iktau+2,ikx) = wl*cbf(nt-iom+2)+wh*cbf(nt-iom+1)
                        }
                else
                        cp(iktau,ikx) = 0.
                }
        }
call ft2d(nt,nx,cp,-1.,1.,cbf)
return;         end
```

The output of this test program was shown in Section 1.3. To better illustrate the periodic nature of the solution all but one semicircle was removed and the result plotted with a nonlinear gain. Four identical plots appear side-by-side in figure 3.

Rocca's Row Fourier Transform

Rocca's Fourier transform over rows is somewhat faster than the rudimentary program because the basic overhead is done *once,* while *every* row gets Fourier transformed. But the main advantage of the Rocca method over the rudimentary method is that the data need not be transposed, and the program runs efficiently even in a paged environment.

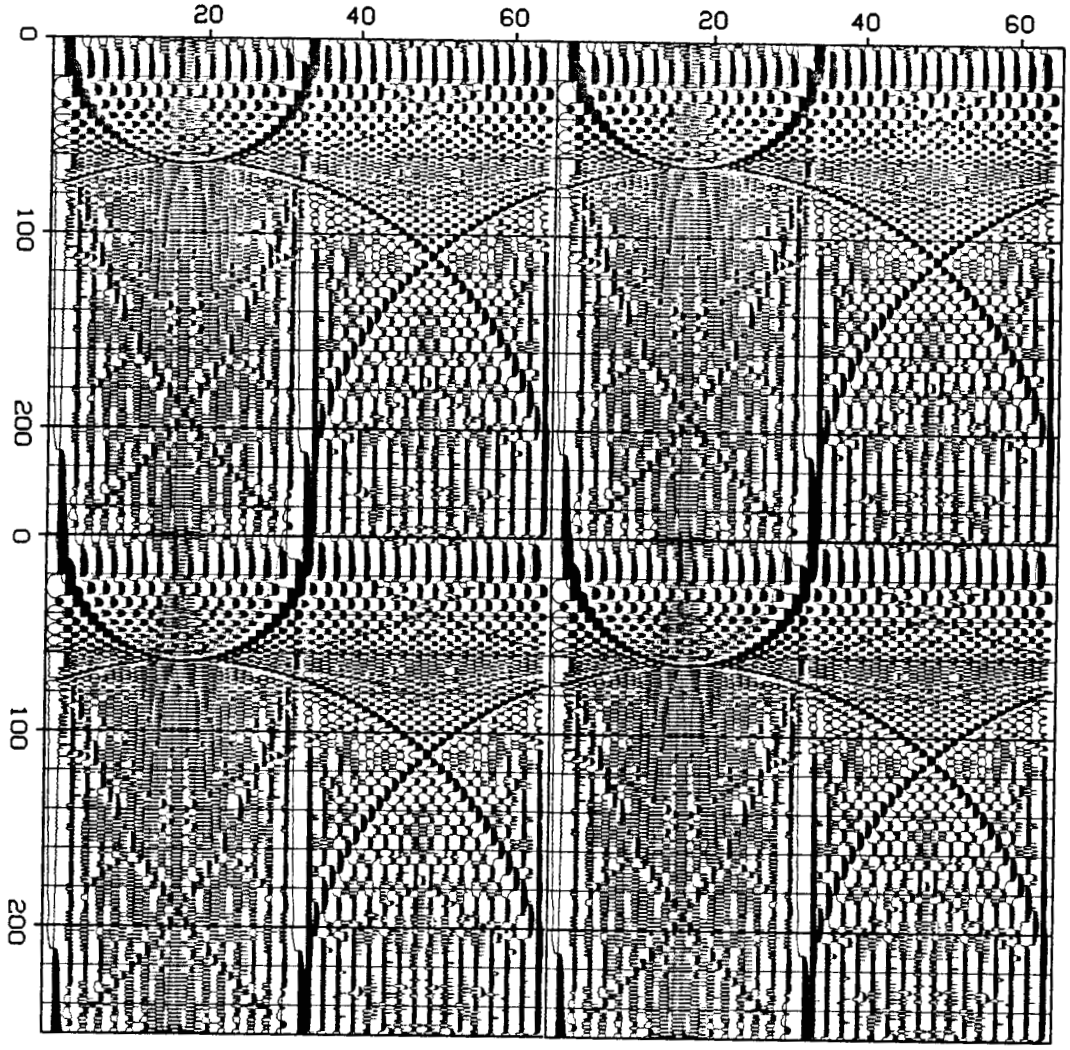


FIG. 1.7-3. Periodicity of the output of the Stolt migration program.

```
# Try Rocca's row Fourier transform.
#       sign2 should be +1. or -1. it is the sign of i.
subroutine rowcc(n1,n2,cx,sign2,scale)
complex cx(n1,n2),cmplx,cw,cdel
do i1 = 1,n1
        do i2 = 1,n2
                cx(i1,i2) = cx(i1,i2)*scale
j = 1
do i = 1,n2 {
        if (i<=j)    call twid1(n1,cx(1,i),cx(1,j))
        m = n2/2
        while (j>m) { j = j-m; m = m/2; if (m<1) break }
        j = j+m }
lstep = 1
repeat {
        istep = 2*lstep;      cw = 1.
        arg = sign2*3.14159265/lstep;  cdel = cmplx(cos(arg),sin(arg))
```

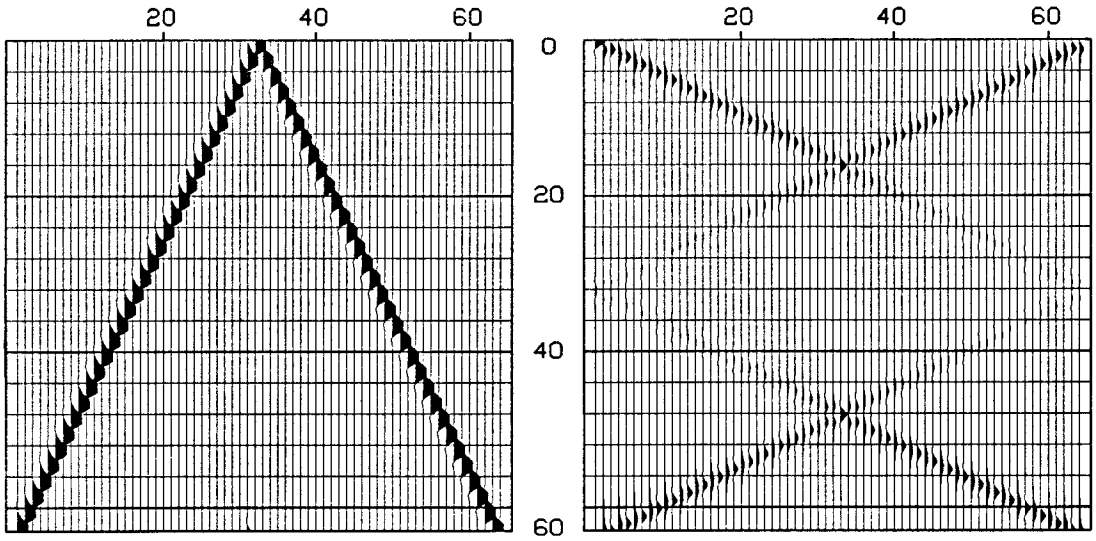
```
        do m = 1,lstep {
                do i = m,n2,istep
                        call twid2(n1,cw,cx(1,i),cx(1,i+lstep))
                cw = cw*cdel
                }
        lstep = istep
        } until(lstep>=n2)
return;         end

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

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

EXERCISES

1. Most time functions are real. Their imaginary part is zero. Show that this means that $F(\omega, k)$ can be determined from $F(-\omega, -k)$.
2. Verify by using your computer and plotter that figure 2 is produced by the program given.
3. The real part of the FT plotted in the previous exercise is somewhat difficult to interpret because of the awkward placement of the negative frequencies and wavenumbers. Modify the program so that $F(\omega, k)$ has its origin at the center (33,33) of the plotted grid. Hint: a simple modification of $f(t, x)$ before Fourier transforming is sufficient; recall the "shift theorem." Write $f(t, x)$ and the new, more easily interpreted $F(\omega, k)$. Label axes. (Hale)
4. A point explosion on the earth's surface at time $t=0$ and location $x=32$ provides synthetic observations in the (t, x)-plane shown on the left. On the right is the magnitude of the two-dimensional Fourier transform, (ω, k_x)-plane. The origin is in the upper left corner of each plot. What would these plots look like on an earth of half the velocity? (Toldi)



5. Insert the appropriate cosine obliquity function into the Stolt migration program. Test, and verify little difference but some angle-dependent scaling.
6. Write a program for diffraction by the Stolt method. That is, given point scatterers inside the earth, generate the appropriate hyperbolas.
7. If you include the inverse cosine weighting function in a Stolt diffraction program, beware of the pole at the evanescent edge. Is it better to stretch before weighting or after? Why?
8. Interpolation error in the Stolt program may be reduced by reducing the speed of oscillation of $P(\omega)$ with ω. To do this note that $p(t)$ vanishes for negative t. So multiply $P(\omega)$ by $e^{-i\omega\tau}$ before interpolation, and then divide it out after. What is an appropriate value of the constant τ to use in the program?