Chapter 3

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. They are all based on Fourier transforms.

3.1 PLANE-WAVE SUPERPOSITION

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.

3.1.1 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 1.1 shows various ways of measuring the size 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

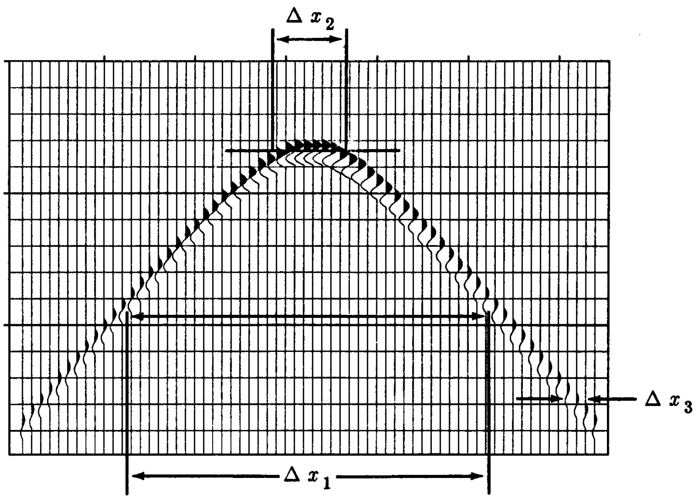

Figure 3.1: Measurements of width parameters of a hyperbola. omk-fwidth [NR]

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 1.1 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.

Figure 1.2 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 was a wall there. You could 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

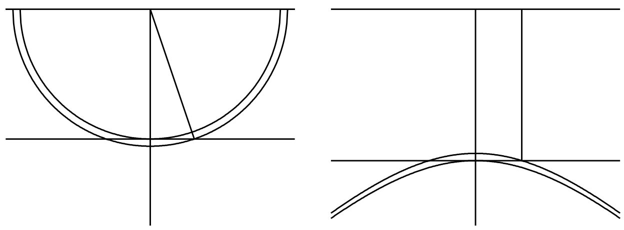

Figure 3.2: Fresnel zone in (x,z)-space (left) and in (x,t)-space (right). omk-fresnel [ER]

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). 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.

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 1.3 demonstrates this. Hyperboloids from

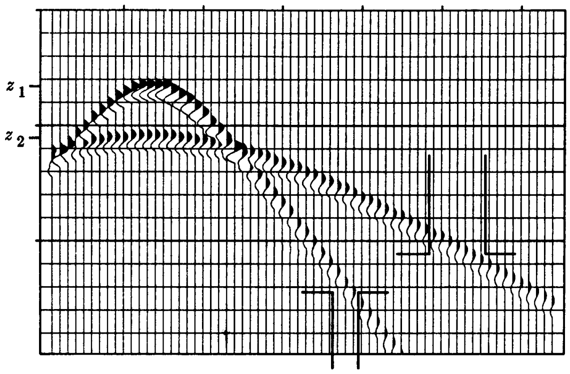

Figure 3.3: Hyperboloids for an earth of velocity increasing with depth. Observable lateral wavelengths get longer with increasing depth. Thus lateral resolving power decreases with depth. omk-hyp2 [NR]

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. 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.

3.1.2 Rays and fronts

Figure 1.4 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

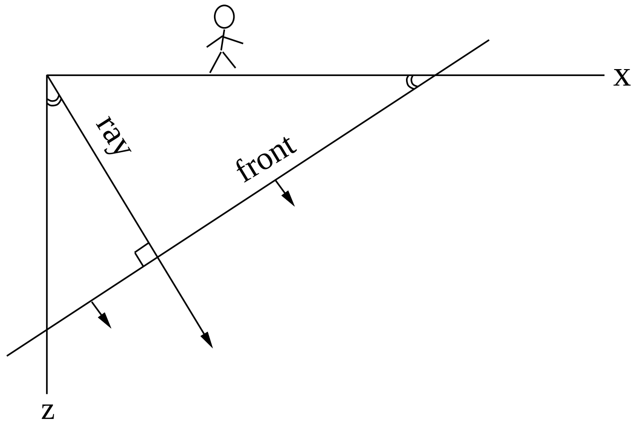

Figure 3.4: Downgoing ray and wavefront. omk-front [NR]

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 would be

$$z \quad = \quad z_0 \ - \ x \ \tan\theta \tag{3.1}$$

like that shown to be the wavefront in Figure 1.4.

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:

$$z \quad = \quad v\,\frac{t}{\cos\theta} \ - \ x \ \tan\theta \tag{3.2}$$

Solving for time gives

$$t(x,z) \quad = \quad \frac{z}{v}\ \cos\theta \ + \ \frac{x}{v}\ \sin\theta \tag{3.3}$$

Equation (1.3) 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 (1.3) to define the time shift t_0 gives an expression for a wavefield that is some waveform moving on a ray.

$$\text{moving wavefield} \quad = \quad f\left(t \ - \ \frac{x}{v}\ \sin\theta \ - \ \frac{z}{v}\ \cos\theta \right) \tag{3.4}$$

3.1.3 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 (1.4) to be the real part of the function $\exp[-i\omega(t-t_0)]$ gives

$$\text{moving cosine wave} \quad = \quad \cos\left[\,\omega\left(\frac{x}{v}\ \sin\theta\ +\ \frac{z}{v}\ \cos\theta\ -\ t\right)\right] \tag{3.5}$$

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 (1.5). 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

$$\text{Fourier kernel} \quad = \quad e^{i\,k_x x}\ e^{i\,k_z z}\ e^{-\,i\omega t} \quad = \quad \exp[i(k_x x\ +\ k_z z\ -\ \omega t)] \tag{3.6}$$

Now for the whistles, bells, and trumpets. Equating (1.5) to the real part of (1.6), physical angles and velocity are related to Fourier components. These relations should be memorized!

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

(3.7)

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, seen earlier as the *dispersion relation of the scalar wave equation.*

$$k_x^2\ +\ k_z^2 \quad = \quad \frac{\omega^2}{v^2} \tag{3.8}$$

The importance of (1.8) 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 (1.8). 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 (1.8).

3.1.4 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. Mathematically shifting on the z-axis is like shifting on the time t-axis. To shift on time by an amount t_0 we multiply by $\exp(i\omega t_0)$ and to shift on depth by an amount z_0 we multiply by $\exp(ik_z z_0)$. It is not quite this trivial however: the data field is not evidently a function of k_z because we do not have data given on a z-axis so we cannot simply Fourier transform it to k_z. Instead we have

$$P(\omega, k_x, z) \quad = \quad P(\omega, k_x, 0) \; e^{i\,k_z(\omega,k_x)z} \tag{3.9}$$

Since the operation is a multiplication in the Fourier domain, it may be described as the engineering diagram in Figure 1.5.

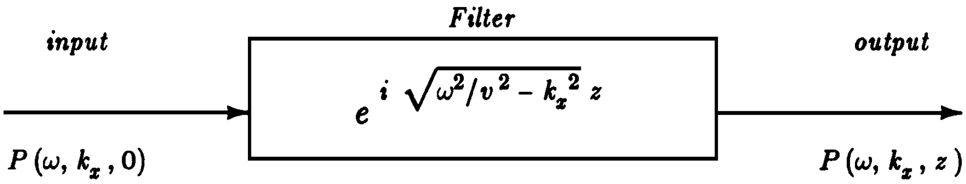

Figure 3.5: $\boxed{\text{omk-inout}}$ [NR]

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 $x^2 + z^2 - v^2t^2$. 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 (1.8):

$$k_z \quad = \quad \pm\sqrt{\frac{\omega^2}{v^2} \; - \; k_x^2} \tag{3.10}$$

$$\quad = \quad \pm\,\frac{\omega}{v}\,\sqrt{1 \; - \; \frac{v^2\,k_x^2}{\omega^2}} \tag{3.11}$$

$$\quad = \quad \pm\,\frac{\omega}{v}\;\cos\theta \tag{3.12}$$

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.

Mathematically we can imagine an arbitrary function of t, x, and z. Its three-dimensional Fourier transform fills a volume in the space of ω, k_x, and k_z. This arbitrary function is not a wavefield unless it vanishes everywhere except the cone-shaped surface $k_x^2+k_z^2 = \omega^2/v^2$ given by the wave equation. Thus equation (1.9) makes sense even though we cannot Fourier transform data over the z-axis because we don't observe data along the z-axis. It makes sense because if we know the mathematical function is a wavefield and if we know its Fourier transform over ω and k_x, then we know its Fourier transform at any k_z and we can "delay" it in z (downward shift) by multiplying by $\exp(ik_z z)$.

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 Evolution of a wavefield with time is described by

$$p(x,z,t) \;=\; \int\int \left[P(k_x,k_z,t=0)\, e^{-i\omega(k_x,k_z)t} \right] \; e^{ik_x x + ik_z z} dk_x \, dk_z \qquad (3.13)$$

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?

4 Downward continuation of a wave is expressed by

$$p(x,z,t) \;=\; \int\int \left[P(k_x,z=0,\omega)\; e^{ik_z(\omega,k_x)z} \right] \; e^{-i\omega t + ik_x x}\; d\omega \, dk_x \qquad (3.14)$$

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?

3.1.5 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:

$$\tau \quad = \quad \frac{2\,z}{v_{\text{true}}} \quad = \quad \frac{z}{v_{\text{half}}} \tag{3.15}$$

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

$$\tau(z) \quad = \quad \int_0^z \frac{dz}{v(z)} \qquad \text{or} \qquad \frac{d\tau}{dz} \quad = \quad \frac{1}{v} \tag{3.16}$$

3.2 SUMMATION AND SUPERPOSITION OF CONICS

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

$$x^2 \;+\; z^2 \quad = \quad v^2\, t^2 \tag{3.17}$$

$$\frac{x^2}{v^2} \;+\; \tau^2 \quad = \quad t^2 \tag{3.18}$$

Figure 1.6, which is from Schneider's classic paper [1971], illustrates the *semicircle-superposition* method. 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

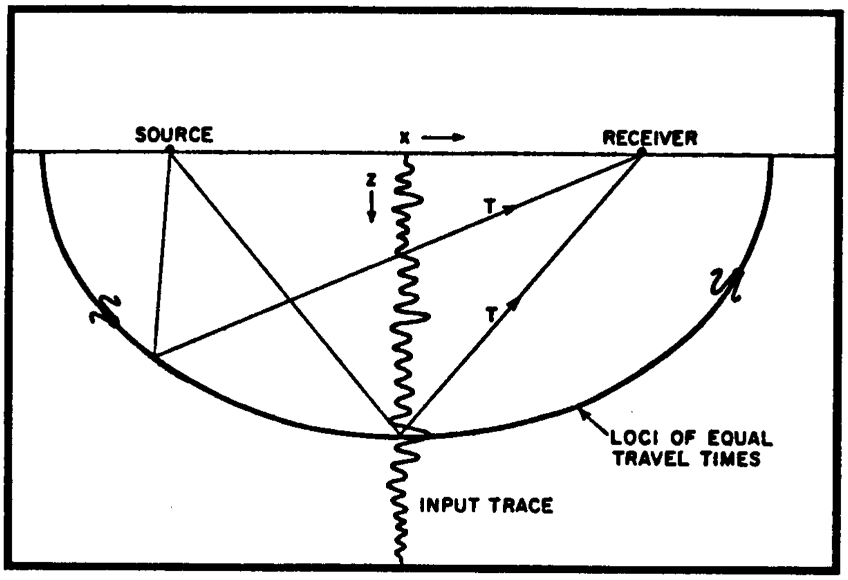

Figure 3.6: The process may be described in numerous ways. Two very simple and equally valid representations are indicated in Figures 1.6 and 1.7. Shown here is a representation of the process in terms of what happens to a single input trace plotted in depth (time may also be used) midway between its 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. (figure and caption from Schneider, W. A., 1971 [by permission]) omk-schneider [NR]

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, τ)-space. Should this happen you might rightly suspect that the input data section in (x, t)-space is a Huygens secondary source, namely, energy concentrated along a hyperbola. This leads us to the *hyperbola-summation* method.

The *hyperbola-summation* method of migration is depicted in Figure 1.7. The

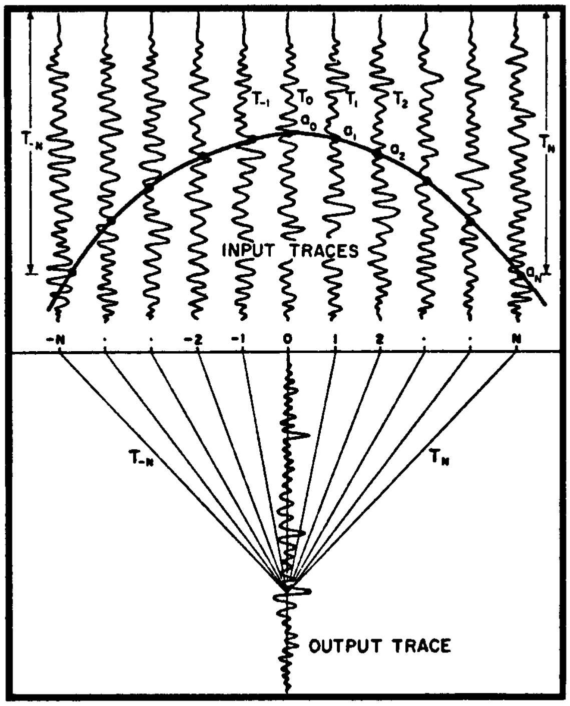

Figure 3.7: A second description of the process is provided here. The process is represented in terms of how an output trace is developed from an ensemble of input traces, shown as CDP-stacked traces in the upper half of the figure. The output in the lower half reflects how each amplitude value at (x, 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. (figure and caption from Schneider, 1971) omk-schneider2 [NR]

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 (1.18), 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.

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.

3.2.1 Hyperbola summation refined into the Kirchhoff method

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

$$\mathrm{FT}^{-1}(e^{i\,k_z\,z}) \quad = \quad \frac{1}{\pi}\,\frac{\partial}{\partial z_0}\,\frac{\mathrm{step}(t\;-\;r/v)}{\sqrt{t^2\;-\;r^2/v^2}} \tag{3.19}$$

where r is the distance $\sqrt{x^2\;+\;(z-z_0)^2}$ between the (exploding reflector) source and the receiver. The function (1.19) 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 (1.19) 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.

You might wonder why anyone would prefer approximations, given the exact inverse transform (1.19). The difficulty of graphing (1.19) 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 4, 8, and 9, are largely devoted to extensions of (1.19) 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 `ft1axis()` on page 32 and `ft2axis()` on page 33. Figure 1.8 shows the result on a 256×64 point mesh. (In practice the mesh would be

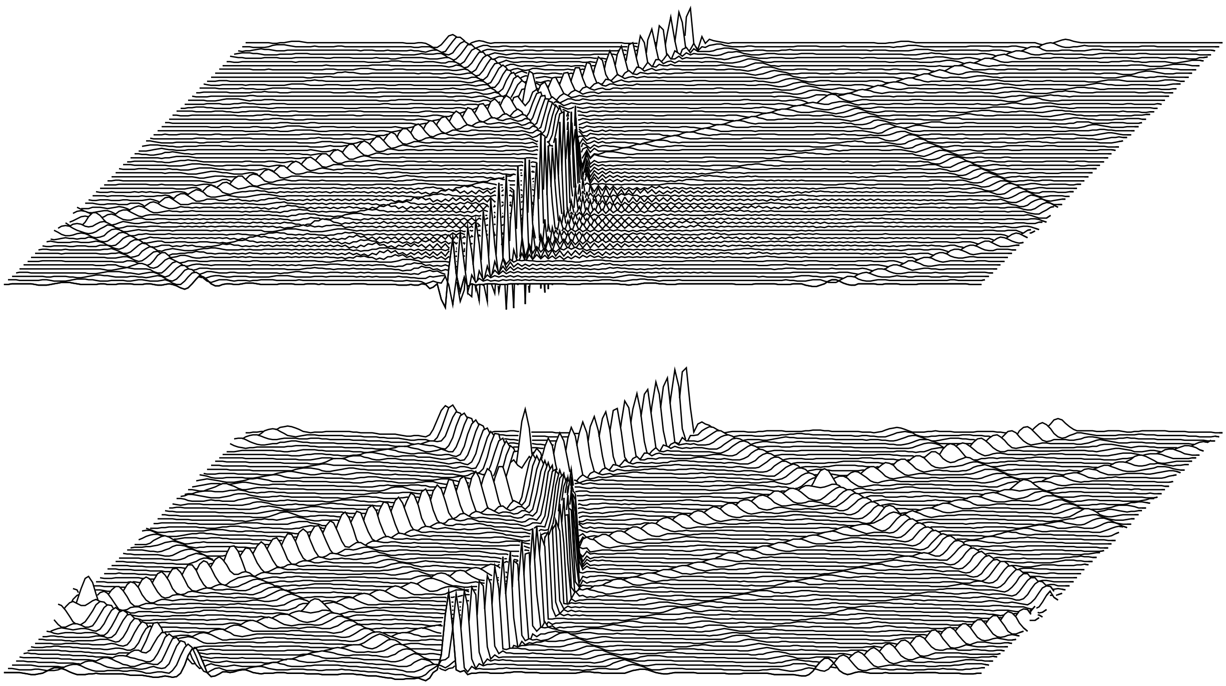

Figure 3.8: The Huygens wavelet (top) and a smoothed time integral (bottom). omk-huygens [ER]

about 1024×256 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 1.8.

3.3 THE PHASE-SHIFT METHOD

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.

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

$$e^{i\,k_z \Delta z} \quad = \quad \exp\left\{ \, -\, i\, \frac{\omega}{v} \, \sqrt{1 \; - \; \left(\frac{v\,k_x}{\omega} \right)^2} \; \Delta z \right\} \tag{3.20}$$

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

$$C \quad = \quad \exp\left\{ \; -\, i\,\omega \; \Delta\tau \, \sqrt{1 \; - \; \left(\frac{v\,k_x}{\omega} \right)^2} \; \right\} \tag{3.21}$$

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 in the following pseudo code:

```
U(ω, k_x, τ = 0) = FT[u(t, x)]
For τ = Δτ, 2Δτ, ..., end of time axis on seismogram
    For all k_x
        For all ω
            C = exp(−iωΔτ sqrt(1 − v² k_x² / ω²)
            U(ω, k_x, τ) = U(ω, k_x, τ − Δτ) * C
    For all k_x
        Image(k_x, τ) = 0.
        For all ω
            Image(k_x, τ) = Image(k_x, τ) + U(ω, k_x, τ)
    image(x, τ) = FT[Image(k_x, τ)]
```

This pseudo code Fourier transforms a wavefield observed at the earth's surface $\tau = 0$, and then it marches that wavefield down into the earth ($\tau > 0$) filling up a three-dimensional function, $U(\omega, k_x, \tau)$. Then it selects $t = 0$, the time of the exploding reflectors by summing over all frequencies ω. (Mathematically, this resembles finding the signal at $\omega = 0$ by summing over all t).

Turning from pseudocode to real code, an important practical reality is that computer memories are not big enough for the three-dimensional function $U(\omega, k_x, \tau)$. But it is easy to intertwine the downward continuation with the summation over ω so a three-dimensional function need not be kept in memory. This is done in the real code in subroutine `phasemig()`.

```
subroutine phasemig( u, nt, nx, dt, dx, image, ntau, dtau, v)
integer nt,nw,nx,ntau, iw,ikx,itau
real    dt,dx, w,w0,dw, kx,kx0,dkx,dtau, kzkz, v, sig1,sig2,pi, signum
complex u(nt,nx), image(ntau,nx), c
                        pi = 3.14159265;        sig1 =+1.;  sig2 =-1.
call ft1axis( 0, sig1, nt, nx, u)
call ft2axis( 0, sig2, nt, nx, u)

w0  = -pi/dt;   dw = 2.*pi/(nt*dt);     nw = nt
kx0 = -pi/dx;   dkx= 2.*pi/(nx*dx)

call zero( ntau*nx, image)
do itau = 1, ntau {
do ikx  = 2, nx    {        kx = kx0 + (ikx-1) * dkx
do iw = 2, nw      {        w  = w0  + (iw-1)  * dw
            kzkz = w*w - v*v * kx*kx
            if (kzkz >0) { c = cexp(cmplx(0., -signum(w)*dtau*sqrt(kzkz))) }
            else         { c = 0.                                          }
            u(iw,ikx) = u(iw,ikx) * c
            image(itau,ikx) = image(itau,ikx) + u(iw,ikx)
            }}}
call ft2axis( 1, sig2, ntau, nx, image)
return; end
```

An aspect of the computation that was hidden in the pseudo code that you can see in the real code is that we must also handle negative frequencies. When doing so, we use the signum function $\mathrm{sgn}(\omega) = \omega/|\omega|$ to choose the sign of k_z so that the sign of k_z is always opposite that of ω.

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. Pseudo code for modeling the upcoming wave u is

$$
\begin{array}{l}
\text{Image}(k_x, z) = FT[\text{image}(x, z)] \\
\text{For all } \omega \text{ and all } k_x \\
\qquad U(\omega, k_x) = 0. \\
\text{For all } \omega \ \{ \\
\text{For all } k_x \ \{ \\
\text{For } z = z_{\max}, z_{\max} - \triangle z, z_{\max} - 2\triangle z, \ldots, 0 \ \{ \\
\qquad C = \exp(+i\triangle z\omega\sqrt{v^{-2} - {k_x}^2/\omega^2} \\
\qquad U(\omega, k_x) = U(\omega, k_x) * C \\
\qquad U(\omega, k_x) = U(\omega, k_x) + \text{Image}(k_x, z) \\
\qquad \} \ \} \ \} \\
u(t, x) = FT[U(\omega, k_x)]
\end{array}
$$

Some real code for this job is in subroutine `phasemod()`.

```
subroutine phasemod( image, nz, nx, dz, dx, u, nt, dt, v)
integer nt,nw,nx,nz, iw,ikx,iz
real    dt,dx,dz, w,w0,dw, kx,kx0,dkx, kzkz, v, sig1,sig2,pi, signum
complex u(nt,nx), image(nz,nx), c
                pi = 3.14159265;        sig1=+1.;       sig2=-1.
call ft2axis( 0, sig2, nz, nx, image)

w0  = -pi/dt;   dw = 2.*pi/(nt*dt);     nw = nt
kx0 = -pi/dx;   dkx= 2.*pi/(nx*dx)
call zero( nw*nx, u)
do iw  = 2, nw     {  w  = w0  + (iw-1)  * dw
do ikx = 2, nx     {  kx = kx0 + (ikx-1) * dkx
do iz  = nz, 1, -1 {
            kzkz = w*w - (v*v) * kx*kx
            if (kzkz >0) { c = cexp(cmplx(0., signum(w)*dz*sqrt(kzkz))) }
            else         { c = 0.                                        }
            u(iw,ikx) = u(iw,ikx) * c
            u(iw,ikx) = u(iw,ikx)   +  image(iz,ikx)
            }}}
call ft1axis( 1, sig1, nt, nx, u)
call ft2axis( 1, sig2, nt, nx, u)
return; end
```

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.

3.4 THE STOLT METHOD

On most computers the Stolt [1978] 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. 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:

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

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

$$P(\omega, k_x, z) \quad = \quad e^{i\,k_z z}\; P(\omega, k_x, z=0) \tag{3.22}$$

Perform a two-dimensional inverse Fourier transform:

$$p(t,x,z) \quad = \quad \int\!\!\int \; e^{i\,k_x x - i\omega t + i\,k_z z}\; P(\omega, k_x, 0)\; d\omega\; dk_x \tag{3.23}$$

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

$$\text{Image}(x,z) \quad = \quad \int\!\!\int \; e^{i\,k_x x}\; e^{i\,k_z(\omega,k_x)\,z}\; P(\omega, k_x, 0)\; d\omega\; dk_x \tag{3.24}$$

Equation (1.24) 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 (1.24) 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 (1.24) will be changed from ω to k_z.

$$\omega \quad = \quad -\operatorname{sgn}(k_z)\; v\sqrt{k_x^2 \;+\; k_z^2} \tag{3.25}$$

$$\frac{d\omega}{dk_z} \quad = \quad -\operatorname{sgn}(k_z)\; v\; \frac{k_z}{\sqrt{k_x^2\; +\; k_z^2}} \tag{3.26}$$

$$\frac{d\omega}{dk_z} \quad = \quad \frac{-\,v|k_z|}{\sqrt{k_x^2\; +\; k_z^2}} \tag{3.27}$$

Put (1.25), (1.26) and (1.27) into (1.24), 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.

$$\mathrm{Image}(x,z) \quad = \quad \int\!\!\int e^{ik_z z + ik_x x} \left\{ P[\omega(k_x,k_z),k_x,0]\, \frac{v|k_z|}{\sqrt{k_x^2+k_z^2}} \right\} dk_z\, dk_x \tag{3.28}$$

Equation (1.28) states the result as a two-dimensional inverse Fourier transform. The Stolt migration method is a direct implementation of (1.28). The steps of the algorithm are

- Double Fourier transform data from $p(t,x,0)$ to $P(\omega,k_x,0)$.
- 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 θ).
- Inverse Fourier transform to (x,z)-space.

Samples of Stolt migration of impulses are shown in Figure 1.9. You can see the

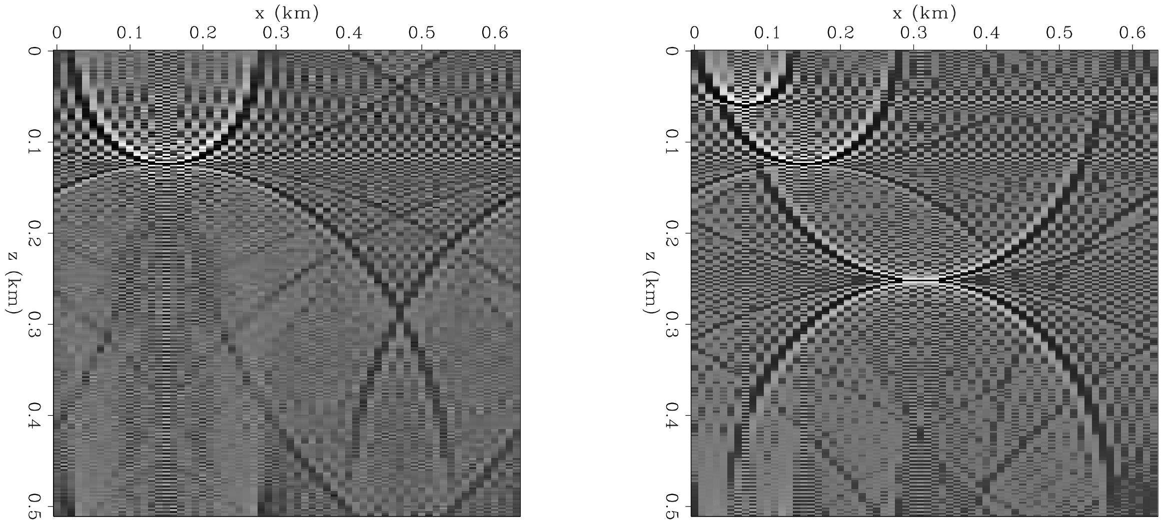

Figure 3.9: Response of Stolt method to data with impulses. Semicircles are seen, along with computation artifacts. omk-stolt [ER]

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. A simpler alternative is to stay away from the bottom of the model, i.e. pad with lots of zeroes.

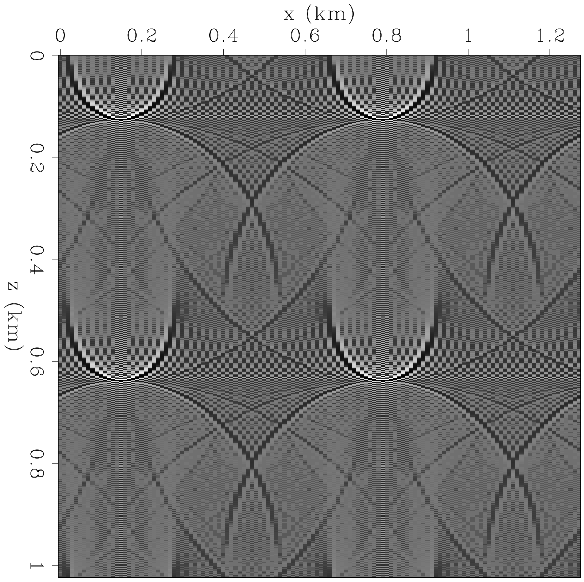

Figure 3.10: Periodicity of the output of the Stolt migration program. omk-stolt4 [ER]

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.

Below is a subroutine for Stolt migration. Its input is assumed to be data after 2-D Fourier transformation to (ω, k_x)-space, and its output is the image in to (k_z, k_x)-space, so it needs a final transformation to (z, x)-space.

```
subroutine stoltmig( nt,nz,nx, dt,dz,dx, vel, cftdata, cftimage)
integer it,iz,ix, nt,nz,nx, iw,ikz,ikx, nw,nkz,nkx
real    dt,dz,dx, w0,kz0,kx0, dw,dkz,dkx, pi, vel,vhalf, w,kz,kx, signum
complex ckzkx, cftdata(nt,nx), cftimage(nz,nx)
pi = 3.1415926
nw  = nt;       w0  = -pi/dt;   dw  = 2.*pi/(nt*dt);
nkz = nz;       kz0 = -pi/dz;   dkz = 2.*pi/(nz*dz);
nkx = nx;       kx0 = -pi/dx;   dkx = 2.*pi/(nx*dx);
do ikz= 1, nkz {        cftimage( ikz,   1) = (0.,0.) } # zero the Nyquist.
do ikx= 1, nkx {        cftimage(   1, ikx) = (0.,0.) } # zero the Nyquist.
vhalf = 0.5 * vel
do ikx = 2, nkx {       kx = kx0 + dkx * (ikx-1)
do ikz = 2, nkz {       kz = kz0 + dkz * (ikz-1)
        w = - signum( kz) * vhalf * sqrt( kx*kx + kz*kz)
        call cinterp1( w, nw,w0,dw, cftdata(1,ikx), ckzkx)
        if( w != 0.)
                cftimage(ikz,ikx) = ckzkx * abs( kz / w)
        else
```

```
                cftimage(ikz,ikx) = 0.
        }}
return;         end
```

`stoltmig()` uses the linear interpolation routine `cinterp1()` on the current page and the signum function $\operatorname{sgn}(x) = x/|x|$ done by routine `signum()` on page 19. In `stoltmig()` I have omitted the theoretically required $\cos\theta$ multiplication. This angle-dependent scaling can be left as an exercise. The effect of such an angle-dependent multiplier is to scale the signal amplitude along the hyperbola according to the local slope of the hyperbola.

```
# Nearest neighbor interpolation would do this:    cbx = cb( 1.5 + (x-x0)/dx)
# Linear interpolation with the same definitions of variables is:
#
subroutine cinterp1( x, nx,x0,dx, cb, cbx)
integer ix,ixc,         nx
real    x, xc, x0, dx, fraction
complex cb(nx), cbx
xc      = (x-x0) / dx
ixc     =  xc
fraction = xc - ixc
ix      = 1 + ixc
if( ix < 1 )
                cbx = cb( 1)
else if( ix+1 > nx)
                cbx = cb( nx)
else
                cbx = (1.-fraction) * cb(ix)  +  fraction * cb(ix+1)
return; end

real function signum(x)
real x
        if      (x>0)  { signum =  1 }
        else if (x<0)  { signum = -1 }
        else           { signum =  0 }
return;end
```

3.5 INFLUENCES ON MIGRATION QUALITY

Several things can degrade of the quality of a migrated image. Some, such as spatial aliasing and velocity uncertainty affect all migration methods. Other quality differences stem from the migration method.

3.5.1 Spatial aliasing concept

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 1.11 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.

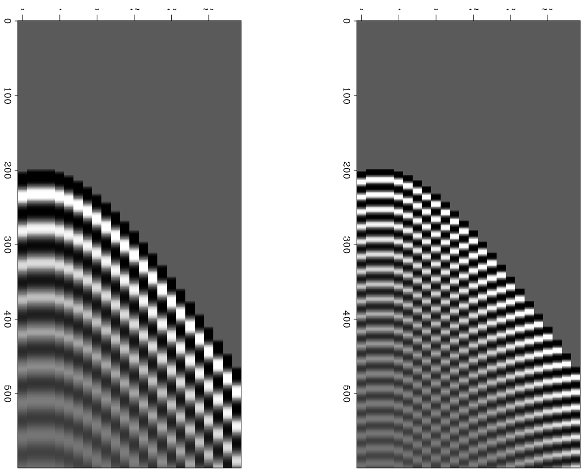

Figure 3.11: Insufficient spatial sampling of synthetic data. To better perceive the ambiguity of arrival angle, view the figures at a grazing angle from the side. omk-alias [ER]

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.

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 1.12. 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.

3.5.2 Sensitivity of migration to velocity error

Figure 1.13 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

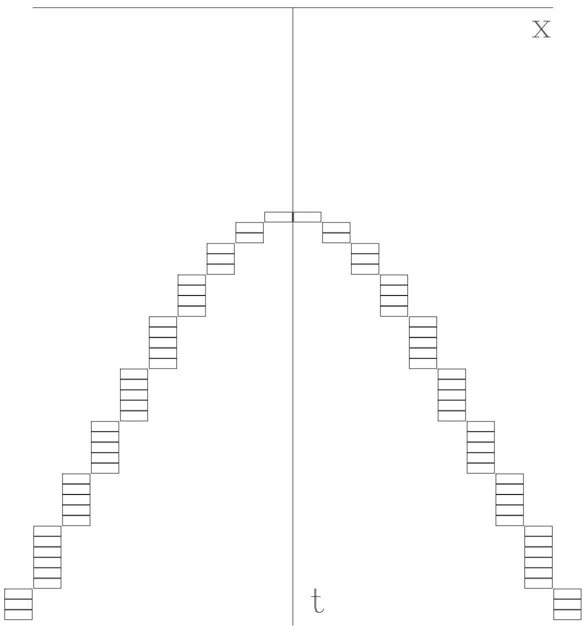

Figure 3.12: For a low-velocity hyperbola, integration will require more than one point per channel. omk-lineint [ER]

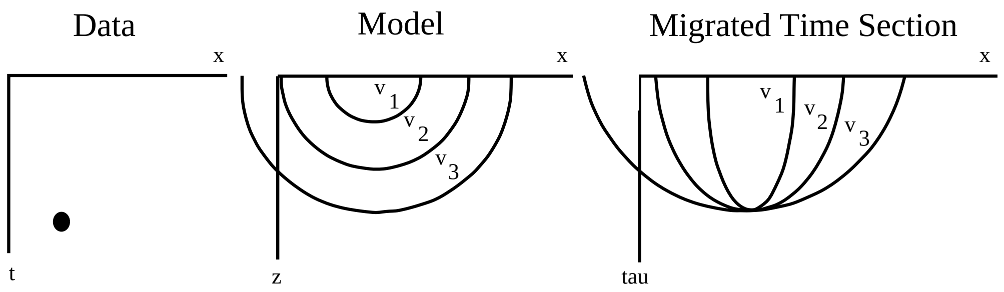

Figure 3.13: Velocity error sensitivity increases with angle up to 90°. Migration of a data impulse as a function of velocity. Three possible choices of constant velocity are shown superposed on one plane. omk-sensitive [NR]

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 1.14 compares two nearby migration velocities. The separation of the curves increases with angle. For the separation to

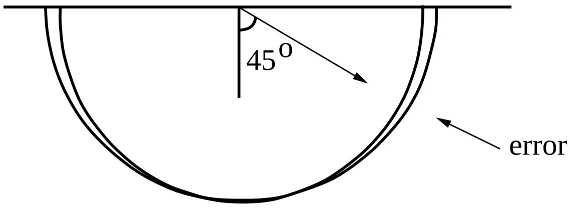

Figure 3.14: Timing error of the wrong velocity increases with angle. omk-senser [NR]

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 $\sqrt{2}$.

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

3.5.3 Subjective comparison and evaluation of methods

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

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:

- Shot-to-geophone offset angles are often large but ignored. A CDP stack is not a zero-offset section.
- Why process to the very wide angles seen in the survey line when even tiny angles perpendicular to the line are being ignored?
- Data is often not sampled densely enough to represent steeply dipping data without aliasing.
- Accuracy in the knowledge of velocity is seldom enough to justify processing to wide angles.
- Noise eventually overpowers all echoes, say at time $t_{\max}$. Given a depth of interest, the maximum angle is $\cos(z/vt)$. 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°.

	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	on x
$v(x)$	Production programs have serious pitfalls.	approximately by iteration and interpolation	no known program
side boundaries and irregular spaceing	excellent	poor	poor
Speed	slow	average	very fast

Table 3.1: Subjective comparison of three wide-angle migration methods.

3.5.4 Evanescence and ground roll

Completing the physical derivation of the dispersion relation,

$$k_x^2 \ + \ k_z^2 \quad = \quad \frac{\omega^2}{v^2} \tag{3.29}$$

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 $\sin\theta = vk_x/\omega$. 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 1.15. What should be done with this energy?

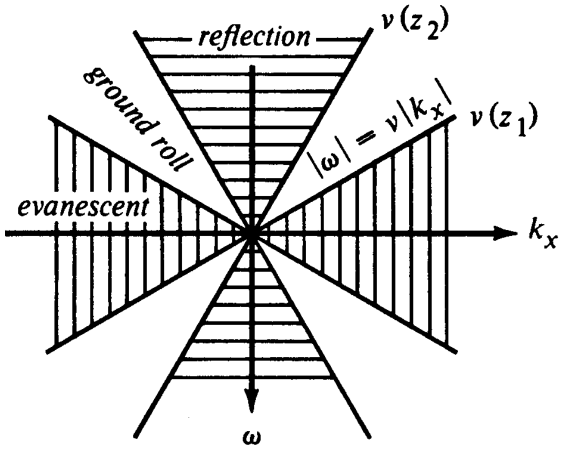

Figure 3.15: The triangle(s) of reflection energy $|\omega| > v(z)\,|k_x|$ become narrower with velocity, hence with depth. Ground roll is energy that is propagating at the surface, but evanescent at depth. omk-evtheory [NR]

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

$$e^{\pm i\sqrt{\omega^2/v^2 - k_x^2}\,z} \quad = \quad e^{\pm\sqrt{k_x^2 - \omega^2/v^2}\,z} \tag{3.30}$$

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 $k_z = \pm ik_x$. 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.

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 1.16 shows an example. This data is not a zero-

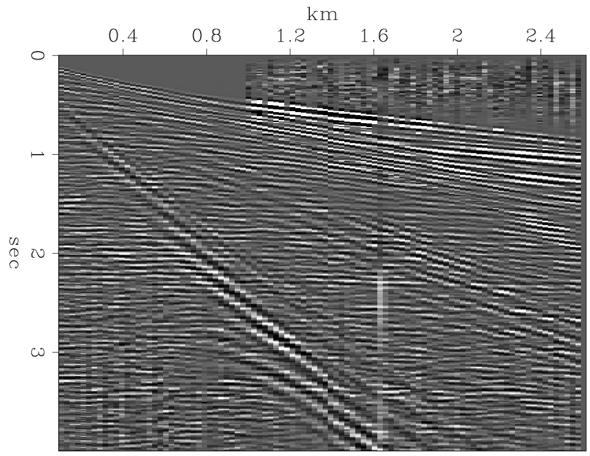

Figure 3.16: Florida shallow marine profile, exhibiting ground roll with frequency dispersion. (Conoco, Yedlin) omk-evdata [ER]

offset section. The shot is on the left, and the traces to the right are from geophones at increasing distances from the shot. First draw a line on the figure going through origin and at a slope of the water velocity, 1.5 km/sec. Steeper slopes are slower surface waves. Geological events have lesser slopes. 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 1.16 is a nice example, no example can really be typical. 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.

EXERCISES:

1 Why does `phasemig()` on page 14 use a negative sign on k_z whereas `phasemod()` on page 15 uses a positive sign even though both are dealing with *upcoming* waves?

2 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 $s(t)$.

3 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.

4 Define the program sketch for the inverse to the Stolt algorithm—that is, create synthetic data from a given model.

5 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.

6 Given seismic data that is cut off at four seconds, what is the deepest travel time depth from which 80° dips can be observed?

3.6 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, or a *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.

3.6.1 Why the 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 (Hemon [1978], 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 $2t_0,\ 3t_0,\ 4t_0, ...$ In the exploding-reflector conceptual model, a sea-floor one-way travel-time depth of t_0 yields sea-floor multiple reflections at times $3t_0,\ 5t_0,\ 7t_0, ...$ 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.

3.6.2 Fourier derivation of the paraxial wave equation

Start from the dispersion relation of the scalar wave equation:

$$k_x^2 \;+\; k_z^2 \;=\; \frac{\omega^2}{v^2} \tag{3.31}$$

Take a square root.

$$k_z \;=\; \pm\sqrt{\frac{\omega^2}{v^2} \;-\; k_x^2} \tag{3.32}$$

The simple act of selecting the minus sign in (1.32) includes the upcoming waves and eliminates the downgoing waves. Equation (1.31) is the three-dimensional Fourier transform of the scalar wave equation. Inverse transforming (1.32) 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:

$$\frac{\partial}{\partial t} \;=\; -\,i\,\omega \tag{3.33}$$

$$\frac{\partial}{\partial x} \;=\; i\,k_x \tag{3.34}$$

$$\frac{\partial}{\partial z} \;=\; i\,k_z \tag{3.35}$$

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 (1.33), (1.34) and (1.35) into (1.32) is called a paraxial equation. Here we will go 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.

3.6.3 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 1.17 shows this in an idealized geometry: waves radiated from the horizontal flight of a supersonic airplane. The airplane passes location x at time $t_0(x)$ flying horizontally

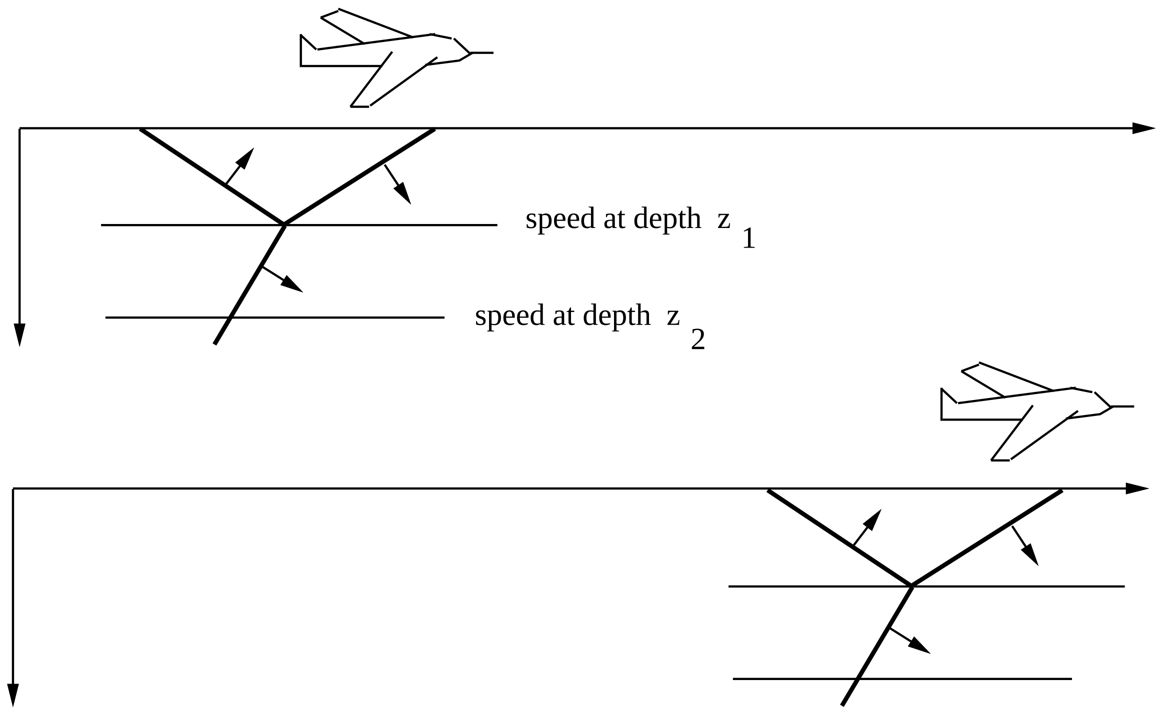

Figure 3.17: Fast airplane radiating a sound wave into the earth. From the figure you can deduce that the horizontal speed of the wavefront is the same at depth z_1 as it is at depth z_2. This leads (in isotropic media) to Snell's law. omk-airplane [NR]

at a constant speed. 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 $\partial t_0/\partial x$, 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 $\partial t_0/\partial x$ 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, $\partial t_0/\partial x$ does not change with depth.

Figure 1.18 illustrates the differential geometry of the wave. The diagram shows

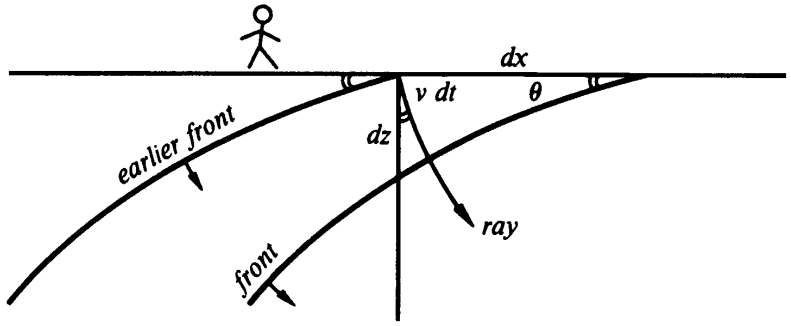

Figure 3.18: Downgoing fronts and rays in stratified medium $v(z)$. The wavefronts are horizontal translations of one another. omk-frontz [NR]

that

$$\frac{\partial t_0}{\partial x} = \frac{\sin\theta}{v} \tag{3.36}$$

$$\frac{\partial t_0}{\partial z} = \frac{\cos\theta}{v} \tag{3.37}$$

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 (1.36) 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* $p = \partial t_0/\partial x$.

It is noteworthy that Snell's parameter $p = \partial t_0/\partial x$ is directly observable at the surface, whereas neither v nor θ is directly observable. Since $p = \partial t_0/\partial x$ is not only observable, but constant in depth, it is customary to use it to eliminate θ from equations (1.36) and (1.37):

$$\frac{\partial t_0}{\partial x} = \frac{\sin\theta}{v} = p \tag{3.38}$$

$$\frac{\partial t_0}{\partial z} = \frac{\cos\theta}{v} = \sqrt{\frac{1}{v(z)^2} - p^2} \tag{3.39}$$

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

$$t(x,z) \quad = \quad \frac{\sin\theta}{v}\, x \;+\; \int_0^z \frac{\cos\theta}{v}\, dz \tag{3.40}$$

$$t(x,z) \quad = \quad p\,x \;+\; \int_0^z \sqrt{\frac{1}{v(z)^2} \;-\; p^2}\;\, dz \tag{3.41}$$

The validity of (1.41) is readily checked by computing $\partial t_0/\partial x$ and $\partial t_0/\partial z$, then comparing with (1.38) and (1.39).

An arbitrary waveform $f(t)$ may be carried by the Snell wave. Use (1.40) and (1.41) to *define* the time t_0 for a delayed wave $f[t - t_0(x,z)]$ at the location (x,z).

$$\mathrm{SnellWave}(t,x,z) \quad = \quad f\left(t \;-\; p\,x \;-\; \int_0^z \sqrt{\frac{1}{v(z)^2} \;-\; p^2}\;\, dz \right) \tag{3.42}$$

3.6.4 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

$$\frac{\partial P(t,z)}{\partial z} \quad = \quad -\,\frac{1}{v}\,\frac{\partial P(t,z)}{\partial t} \tag{3.43}$$

as may be readily verified by substituting either of the trial solutions

$$P(t,z) \;=\; f\left(t \;-\; \frac{z}{v} \right) \qquad \text{for constant } v \tag{3.44}$$

$$P(t,z) \;=\; f\left(t \;-\; \int_0^z \frac{dz}{v(z)} \right) \qquad \text{for } v(z) \tag{3.45}$$

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

$$\frac{\partial P(t,x,z)}{\partial z} \quad = \quad -\,\frac{\partial t_0}{\partial z}\,\frac{\partial P(t,x,z)}{\partial t} \tag{3.46}$$

which has the solution

$$P(t,x,z) \quad = \quad f(t \;-\; p\,x \;-\; \int_0^z \frac{\partial t_0}{\partial z}\, dz) \tag{3.47}$$

In interpreting (1.46) and (1.47) recall that $1/(\partial t_0/\partial z$ 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 (1.38) and (1.39) also enables us to rewrite (1.46) in the various forms

$$\frac{\partial P(t,x,z)}{\partial z} \quad = \quad -\frac{\cos\theta}{v}\frac{\partial P(t,x,z)}{\partial t} \tag{3.48}$$

$$\frac{\partial P(t,x,z)}{\partial z} \quad = \quad -\sqrt{\frac{1}{v(z)^2} - p^2}\;\frac{\partial P(t,x,z)}{\partial t} \tag{3.49}$$

$$\frac{\partial P(t,x,z)}{\partial z} \quad = \quad -\sqrt{\frac{1}{v(z)^2} - \left(\frac{\partial t_0}{\partial x}\right)^2}\;\frac{\partial P(t,x,z)}{\partial t} \tag{3.50}$$

Equations (1.48), (1.49) and (1.50) are paraxial wave equations. Since $\partial t_0/\partial x = p$ can be measured along the surface of the earth, it seems that equation (1.50), 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 equations (1.48), (1.49) and (1.50) or their solution (1.47). One such decomposition technique is *Fourier analysis.*

3.6.5 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 $\partial t_0/\partial x \;=\; k_x/\omega$, 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:

$$\boxed{\frac{\partial t_0}{\partial x} \quad = \quad \frac{\sin\theta}{v} \quad = \quad p \quad = \quad \frac{k_x}{\omega}} \tag{3.51}$$

In theoretical seismology a square-root function often appears as a result of using (1.51) to make a cosine.

Utilization of this Fourier domain interpretation of Snell's parameter p enables us to write the square-root equations (1.48), (1.49) and (1.50) 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 (1.48), (1.49) and (1.50) by $-i\omega$. The result is

$$\frac{\partial P(\omega, k_x, z)}{\partial z} \quad = \quad + \, i\omega \sqrt{\frac{1}{v(z)^2} \quad - \quad \frac{k_x^2}{\omega^2}} \; P(\omega, k_x, z) \tag{3.52}$$

At present it is equivalent to specify either the differential equation (1.52) or its solution (1.47) with f as the complex exponential:

$$P(\omega, k_x, z) \quad = \quad \exp \left(i\omega \int_0^z \sqrt{\frac{1}{v(z)^2} \quad - \quad \frac{k_x^2}{\omega^2}} \;\; dz \right) \tag{3.53}$$

Later, when we consider lateral velocity variation $v(x)$, the solution (1.53) becomes wrong, whereas the differential equation (1.50) 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.

3.6.6 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 (1.47), yields an approximate solution to the scalar wave equation for stratified media $v(z)$.

Index