3.0 MIGRATION, DEPENDENCE ON VELOCITY

Chapter 1 described the main ideas in migration and wave-field continuation. It also presented Fourier techniques for use when the velocity is laterally constant. Chapter 2 developed wave-equation techniques for use in Fourier domains or time and space domains. The space domain is mandatory when the velocity is laterally variable. In this chapter we will see what actually happens when you migrate. Migration processing depends on an assumed velocity model. It is particularly sensitive to lateral variation in velocity. Even in stratified velocity, migration has both physical and cosmetic effects on the data. We will identify what can go wrong: velocity error, dispersion, side-boundary reflection, and instability. We will see how to quantify and suppress the effects of errors.

Migration in (x,z,ω)-Space

Recall the phase-shift method from Chapter 1. The idea was to iterate on a two-step process. The first step was to downward continue with $exp\,[ik_z(\omega,k_x)\Delta z\,]$. The second step was to sum over all frequencies. Summation has the effect of evaluating the wave at $t=0$. There is an optional, additional step, which is to subtract the wave at $t=0$ from the dataset. If you don't do this optional step, it probably won't matter because you probably won't look at the data before $t=0$. However, in practice, the sampling of the ω-axis forces periodicity in t, which later shows up in z. So sometimes the optional step is worthwhile.

With lateral-velocity variation, things proceed in a very similar way. The major difference is that the database is represented in the space of (ω,x) rather than (ω,k_x). The downward-continuation step is now done in two stages. The first stage is the thin-lens, phase-shifting part. The second stage is a Crank-Nicolson, tridiagonal, diffraction stage.

Sensitivity of Migration to Velocity Error

Figure 1 shows how the migration impulse response depends on velocity.

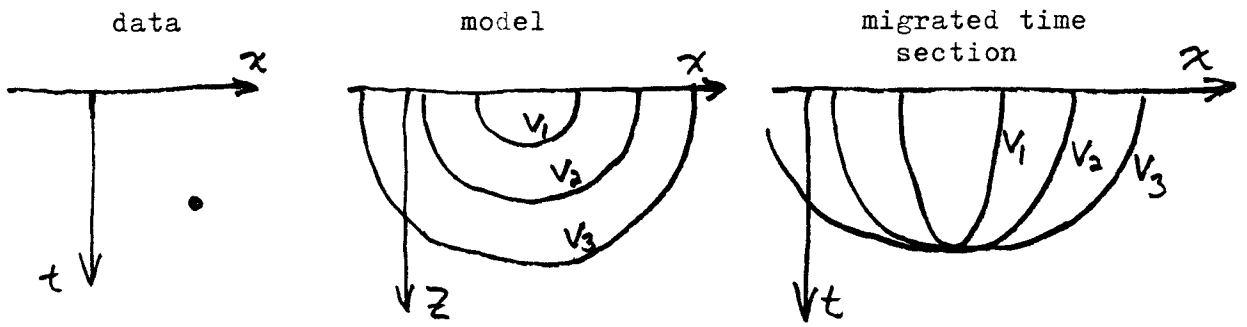

FIG. 1. Migration of a data impulse as a function of velocity.

Significant timing error will be assumed to be about a half-wavelength. Observationally the ratio of traveltime to wavelength is usually a hundred or less. (Notable exceptions are when much of the path is in water or else at time depths greater than about four seconds.) Consequently the velocity accuracy required for 90-degree migration is about 1%. For 45-degree migration velocity error could be larger by the square root of 2. This is illustrated in Figure 2.

Velocities are rarely known this accurately. What is the utility of processing in which erroneous time shifts exceed a half-wavelength?

Migrated Time Sections and Lateral Velocity Variation

The discussion on *traveltime depth* (in the section, "Four Wide-Angle Migration Methods") explained the industrial practice of avoiding reference to depth on a migrated section. This was done by use of a traveltime depth τ defined by

$$\frac{d\tau}{dz} = \frac{2}{v_{rock}} = \frac{1}{v_{half}} \tag{1}$$

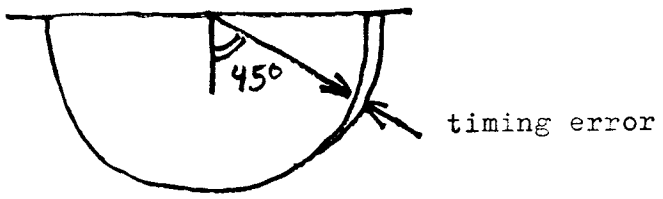

FIG. 2. Timing error of the wrong velocity increases with angle.

The purpose of the transformation is to reduce the velocity sensitivity of the data display. Sensitivity is reduced because a flat horizontal reflector may be migrated with *any* velocity without affecting its position on a migrated time section. The effect of velocity error increases with dip as can be observed by noting that the downward-continuation operation

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

is converted by (1) to

$$\exp\left\{ i\omega\, \Delta\tau \left[1 - \left[\frac{vk_x}{\omega}\right]^2\right]^{1/2} \right\} \tag{3}$$

Now we notice that the velocity v multiplies the stepout k_x/ω, so an error in one is like an error in the other. They are conveniently lumped together and called an angle error. In practice, it is often valid to say that 15° migration requires little velocity accuracy while 45° migration demands a lot.

All of this "conventional wisdom" becomes doubtful in the presence of lateral velocity variation. There seems to be no satisfactory, automatic way of dealing with the data over the quantitative range of geological parameters which commonly occurs. But the downward-continuation equations clarify the nature of the difficulty.

Even in the presence of lateral velocity variation, the square root in (2) is interpretable in a "local plane wave" sense, and it may be used to define a reasonable equation for downward continuation. That is, we have some local spatial wavelengths given by

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

The difficulty arises when we try to convert to traveltime depth with (1). The issue is whether to let the velocity in (1) be laterally variable. This will be tried, not because it is justifiable, but because it leads to a definition of *migrated time section*, an important industrial product. The definition is given by any reasonable implementation of

$$k_\tau(x,\tau) = \omega \left[1 - \left[\frac{v(x,\tau)\, k_x(x,\tau)}{\omega} \right]^2 \right]^{1/2} \tag{5}$$

The definition of *migrated time section* contains a serious pitfall which only shows up when the velocity is laterally variable. The (x,z) coordinate system is an orthogonal coordinate system but the (x,τ) system will not be orthogonal [unless $v(x)=const$], so equation (4), which basically says that $\cos\vartheta = (1 - \sin^2\vartheta)^{1/2}$, is not correctly interpreted by (5).

To avoid this pitfall, the velocity for time-to-depth conversion in (1) can be kept independent of x, call it $\overline{v}(z)$. It is like a retardation velocity. Since $\overline{v}(z)$ does not depend on x the conversion from z to τ is a simple scaling given by (1), which means the extrapolation equation is transformed by

$$\frac{\partial}{\partial \tau} = \frac{dz}{d\tau} \frac{\partial}{\partial z} = \overline{v} \frac{\partial}{\partial z} \tag{6a}$$

$$k_\tau = \overline{v}\, k_z \tag{6b}$$

instead of the complicated chain rule expression which would be required if $\overline{v}$ depended on x.

In principle, $\overline{v}(z)$ could be any function of z, but in practice it is convenient to take it to be some horizontal average of $v(x,z)$, say $\overline{v}(z) = \ll v(x,z) \gg$. This definition ensures that (6b) produces a migrated time section where that is possible, by virtue of the velocities being equal. Where the velocities are unequal, the difference of $1/v$ and $1/\overline{v}$ occurs in a lens-type term which becomes an essential part of the downward continuation. In this chapter we will see some of the effects of the lens term.

You can't time shift in the time domain.

You might wish to do the migration in (x,t)-space, as described at the end of the previous chapter. Then the thin-lens stage would be implemented by time shifting instead of multiplying by exp $\{i\omega [v(x,z)^{-1} - \overline{v}(z)^{-1}]\Delta z\}$. Time shifting is a delightfully easy operation when data is to be shifted by an integral amount of sample units. But repetitive time shifting by a fractional number of digital units is a nightmare. Multipoint interpolation operators are required. Even then, pulses tend to disperse.

3.2 PHYSICAL AND COSMETIC ASPECTS OF THE 45-DEGREE EQUATION

Dip filtering and gain control are two processes whose purpose seems to be largely cosmetic, that is, the changes they make to the data are planned to improve its appearance. Criteria invoked to choose quantitative parameters of such processes are often vague and relate to human experience or visual perception. Objective criteria as signal and noise dip spectra are rarely used in a quantitative way. But the importance of cosmetic processes is not to be underestimated. On many occasions a comparison of processing techniques (for choice of contractor?) has been frustrated by accidental change in cosmetic parameters. Nor are cosmetic processes totally outside the world of wave-equation analysis. Indeed they can be made even more effective when built into a process rather than added on at the end.

Dip filtering often comes under suspicion as being a deceptive violator of the purity of real data. Indeed it can be misused to create events at will. But dip filtering does occur naturally and it also occurs as a by-product of various other processes. Our purpose here is to see how it can be used to advantage in a perfectly justifiable way.

False Semi-Circles in Migrated Data

A commonly missed opportunity is the failure to make effective use of dip filtering to suppress multiples. Without going into a detailed exposition of the theory and properties of focused multiple reflections, it can be stated that multiples are unlike primaries in one important respect. Their strength may change rapidly in the horizontal direction. They need not be spread out into broad diffraction hyperbolas as primaries must. This difference arises because they often spend much time focusing themselves in the irregular, near-surface areas. Common evidence for this behavior is contained in the appearance of wide-angle migrated sections. Such sections often show semi-circular arcs coming all the way up to the surface. Such arcs are obviously not primary reflections. They can be multiples or unexplained impulsive noise. In either case they can be partially suppressed without touching primaries.

Zapping Multiples in Dip Space

Think of the migration of CDP stack as downward continuation in (ω, k_x, z)-space. Ordinarily, velocity increases with depth. As the downward continuation proceeds the velocity cutoff at the evanescent point bites out more and more area from the (ω, k_x)-space. Energy beyond this cutoff does not fit the primary wave propagation model, and it should be suppressed as soon as it is encountered. Such noise suppression can lead to a considerable drop in total power at late times.

Mixed Appearance of Dip-Filtered Data

An objection often raised against dip filtering is that it can give data a *mixed* appearance. By *mixed* is meant that adjacent channels appear to have been averaged and that they are no longer independent. This is true and it is inevitable at late times. It is inevitable because the horizontal resolving power of reflection data decreases with time. There are two reasons for decreasing lateral resolution. First, dissipation causes high frequencies to disappear. Second, even at constant frequency, horizontal wavelengths must increase as rays reach the higher velocities found at greater depths. It is unrealistic to ignore this fundamental limitation and imagine that channels should always have an appearance of independence. If a mixed appearance is to be avoided for display purposes then I advocate removing the low-velocity, coherent, signal-generated noise and replacing it by low-velocity, incoherent, Gaussian, random noise.

Accentuating Faults

It often happens that the location of oil is controlled by faulting. But the dominating effect of stratified reflectors may overwhelm the weak diffraction evidence of faulting. This situation calls for a cosmetic process which weakens zero and small dips, accentuates dips in the range of 10 to 60 degrees and then represses the very wide angles and evanescent energy. As with frequency filtering, sharp cutoffs are not desirable because of the implied long (and in space, wide) impulse response. It turns out that adjustable parameters that help achieve these goals are already available in the 45-degree equation.

Decomposition of the 45 Degree Equation into Effects

There are various means of entering viscosity into wave propagation theory. A well known means is to introduce a complex velocity into the ω^2/v^2 term of the scalar wave equation. This is much like introducing a complex ω. It may be recalled from Fourier transform theory that multiplication of a time function by a decaying exponential $\exp(-\alpha t)$ is the equivalent of replacing $-i\omega$ by $-i\omega+\alpha$ in the transform domain. In a later section on impedance functions it is shown that replacing $-i\omega/v$ by $(-i\omega/v)^\gamma$ describes the so-called "constant Q" absorption, which accurately matches laboratory measurements.

Performing two iterations of the Muir square root expansion we get an expression like

$$iVk_z^{(45)} = -i\omega_0 + \frac{X^2}{-i\omega_1 2 + \dfrac{X^2}{-i\omega_2 2}} \tag{1}$$

In this expression X^2 denotes $V^2k_x^2$ (or the positive definite matrix $(V\partial_x)(V\partial_x)^T = -V\partial_{xx}V$). Previously, in expressions like (1) we have always written simply $-i\omega$, never expecting to want to make the distinction between ω_0, ω_1, or ω_2. Indeed, we usually want each ω to have the same real part. However, by introducing different imaginary parts we can introduce a viscosity which is angle dependent.

For example, we could choose each $-i\omega_j$ in (1) to be the constant Q impedance function $(-i\omega)^\gamma$. The implied migration equation would then back out presumed frequency dissipation in the rocks. But that would lead to a ridiculous enhancement of high frequencies. A better idea would be to keep ω_0 non-viscous but choose $-i\omega_1 = -i\omega_2 = (-i\omega)^\gamma$. With this idea there would be no attempt to back out the Q of the vertical path but there would be compensation of non-zero offsets to the zero-offset.

The choice of different real parts for the $-i\omega_j$ functions creates an amplification or attenuation which depends on dip. We could select the real part of $-i\omega_1$ for the for-mentioned compensation of offset for Q. Then we could use the real part of $-i\omega_2$ for the purpose of suppressing evanescent energy. It would be simple if we could choose the real part of $-i\omega_2$ so as to attenuate all dips above (i.e. velocities below) the medium's cutoff. What happens is almost as good but not quite as simple. Larry Morley discovered (SEP 16, p 109-119) that the absorption turns out to be extremely strong at 3/4 of the medium velocity but, unfortunately, not so strong at lower velocities. (The 45-degree dispersion relation curve hits zero at $Vk/\omega = 4/3$.) The width and depth of

the absorption is somewhat controllable without messing up the migration, but massive very low velocity noise is better eliminated by some other means. In summary,

Main effect of terms in the 45 degree equation			
term	Real part	Imaginary Part	
	(influences traveltime)	cosmetic	physical
ω_0	time/depth conversion	t.v. filter	absorption
ω_1	migration/stacking vel.	fault enhancement	Q-offset compensation
ω_2	migration/stacking vel.	steep dip suppress	evanescent junk

Gain Control Does Dip Filtering Too

If data is exponentially gained upward before migration, then hyperbola flanks are boosted in strength before migration moves them in with the hyperbola top. This is certainly a dip-enhancement feature. Let us consider this more specifically. Take the $Z-transform$ of a time function a_t.

$$A(Z) = a_0 + a_1 Z + a_2 Z^2 + \cdots$$

We now define the exponentially gained time function by

$$\uparrow A(Z) = a_0 + a_1 e^{\alpha} Z + a_2 e^{2\alpha} Z^2 + \cdots$$

The symbol $\uparrow$ is indicative of the exponential gain. Mathematically $\uparrow$ means that the Z is replaced by $e^{\alpha}Z$. Consider a polynomial multiplication or convolution of time functions

$$C(Z) = A(Z)\,B(Z)$$

Obviously,

$$\uparrow C = (\uparrow A)\,(\uparrow B)$$

This means we can do exponential gain either before or after convolution.

Think of the downward-continuation operator $\exp(+ik_z z)$ for some fixed z and some fixed k_x. It is a function of ω which may be expressed in the time domain as a filter a_t. But the hyperbola flanks move *upward* on migration. So the filter is

anticausal which we can indicate by

$$A(Z) = a_0 + a_1 \frac{1}{Z} + a_2 \frac{1}{Z^2} + \cdots$$

Exponentially boosting the coefficients of positive powers of Z is associated with diminishing negative powers. So $\uparrow A$ has a weakened precursor; it tends to attenuate flanks rather than moving them; and it is said to be viscous.

A purely physical point of view dictates that cosmetic functions like gain control and dip filtering should be done after processing, say $\uparrow(AB)$. But this is equivalent to $(\uparrow A)(\uparrow B)$, which means using a viscous operator on exponentially gained data. In practice it is common to forget the viscosity and create $A(\uparrow B)$. I like such cosmetics in the sense that I think dipping events carry more information than flat ones. But going beyond 45-degree dip, attenuation seems to be preferable. The decomposition of the 45-degree equation into the three main parts gives much flexibility for reaching toward these goals.

Rejection by Incoherence or Rejection by Filtering?

We should avoid the pitfall of judging a supposed non-cosmetic process by a cosmetic effect. I once got caught. The process was migration before stack. The feature deemed desirable was the relative strength of the steepest clear event on the record, a fault plane reflection. But even gain control can affect dip spectra! I hoped the process was working by correctly eliminating some of the rejection of steep dips by CDP stack. Perhaps it was, but how could I know whether this was really happening or whether the process had an accidental ability to enhance dips by spatial filtering?

The Substitution Operator

The upward $\uparrow$ operator has been defined to mean the substitution $Z \rightarrow Ze^{\alpha}$. The main property of this operator is that if $C=AB$ then $\uparrow C = (\uparrow A)(\uparrow B)$. This property would be shared by any algebraic substitution for Z, not just the one for exponential gain. Another relatively trivial substitution may be used to achieve time-axis stretching or compression. For example replacing Z by Z^2 stretches the time axis by two. Another substitution which has a considerably deeper meaning than either of the previous two is the substitution of the constant Q dissipation operator $(-i\omega)^{\gamma}$. In summary,

Substitutions for Z-transform variable Z [all preserve $C(Z)=A(Z)B(Z)$]	
Exponential Growth	$Z \rightarrow Ze^{\alpha}$ $(i\omega \rightarrow i\omega + \alpha)$
Time expansion ($\alpha > 1$)	$Z \rightarrow Z^{\alpha}$
(Inverse) Constant Q dissipation	$-i\omega \rightarrow (-i\omega)^{\gamma}$

3.7 ABSORBING SIDES

Computer memory cells are often used to model points in a volume containing propagating waves. Regrettably, the number of cells is necessarily finite, though we often wish to model an infinite volume. Waves in the computer reflect back from the boundaries of the finite computer memory although we would usually prefer that the waves had gone away to infinity. To avoid the need for infinite computer capacity it is natural to try to develop absorptive side boundary conditions. In this section we will develop some highly absorptive boundary conditions, not as absorptive as an infinitude of appended zeros but obviously much cheaper.

Truncation at the Ends of the Cable and at the Ends of the Survey

In exploration we have two kinds of horizontal truncation problem. The first of these is at the end of the geophone cable. This problem mainly affects common-midpoint stacking. The second is at the geographical boundaries of the survey. This problem mainly affects migration. In each case the ideal solution is *not* an infinite amount of zero padding. The ideal solution is some kind of a horizontal extrapolation of the dataset. Two important ingredients to such an extrapolation are a noise model, and a wave-propagation signal model. Such an ideal extrapolation is rarely, if ever, approached in practice. Usually we settle for zero padding and possibly some tapering. For stacking, the zero padding has no cost. The migration situation seems to be analogous to stacking in that hyperboloids are collapsed to points. Actually the two situations are different because of the data itself. With migration we can have reflectors dipping either downward or upward toward the end of the section. When the energy dips downward, then migration will tend to move the energy back into the section. This is also the case with stacking in that energy moves from the far end of the cable back toward zero offset. It is in the other situation, the one which occurs only with migration, which is the troublesome one. In this case energy will tend to move off the mesh on which the data is defined. Again, philosophically, the best solution is some kind of an extrapolation of the dataset. In practice, the zero-slope side condition is perfect when dips vanish, and it is tolerable when they are small. Difficulty occurs when upward dipping energy is seen at the edges of the survey. On migration it tends to

falsely reflect back into the dataset. Such spurious reflections are usually objectionable.

The problem may be reduced by appending zeros to the sides of the dataset, thus providing the dipping energy with a place to go. This is helpful but it does not solve the whole problem for two reasons. The deeper and more basic reason is that some kind of an edge diffraction will still be produced. A secondary reason is that the zero padding cannot continue for an infinite distance. A diminishing value-for-cost requires that the zeros terminate somwhere beyond the end of the survey. At the termination there is a reasonable chance of finding remaining energy which we would prefer to absorb rather than reflect. It is this problem which Enqguist solved. The main practical benefit is the economic one of reduction in the number of padded zeros. To understand Engquist's solution we begin with a different problem, which is conceptually easier.

Engquist Boundaries for the Scalar Wave Equation

The simplest boundary condition is that a function should vanish on the boundary. A wave incident onto such a boundary reflects with a change in polarity (so that the incident wave plus the reflected wave will vanish on the boundary). The next-to-simplest boundary condition is the zero-slope boundary condition. It is also a perfect reflector, but the reflection coefficient is +1 instead of -1. Two points at the edge of the differencing mesh are required to represent the zero-slope boundary. The most general boundary condition usually considered is a linear combination of function and slope. This is also a two-point boundary condition. It so happens that our extrapolation equations contain only a single depth derivative so that on the z-axis they are a two-point condition. Observing this, Björn Engquist recognized a new application for extrapolation equations. Many researchers in other disciplines are interested in forward modeling, that is, evolving forward in time with an equation like the scalar wave equation, say $P_{xx} + P_{zz} = P_{tt}/v^2$. These people severely suffer the consequences of limited memory. Engquist's idea was that they should use our extrapolation equation as their boundary conditions. Suppose they desire an infinite absorbing volume surrounding a box in the (x,z)-plane. Then they need a boundary condition going all the way around the box. They could use our downgoing wave equation on the bottom of the box and our upcoming wave equation on the top edge. The sides could be handled analogously with an interchange of x and z. This idea was thoroughly tested and confirmed by Robert Clayton. An example of one of his comparisons is in figure 1.

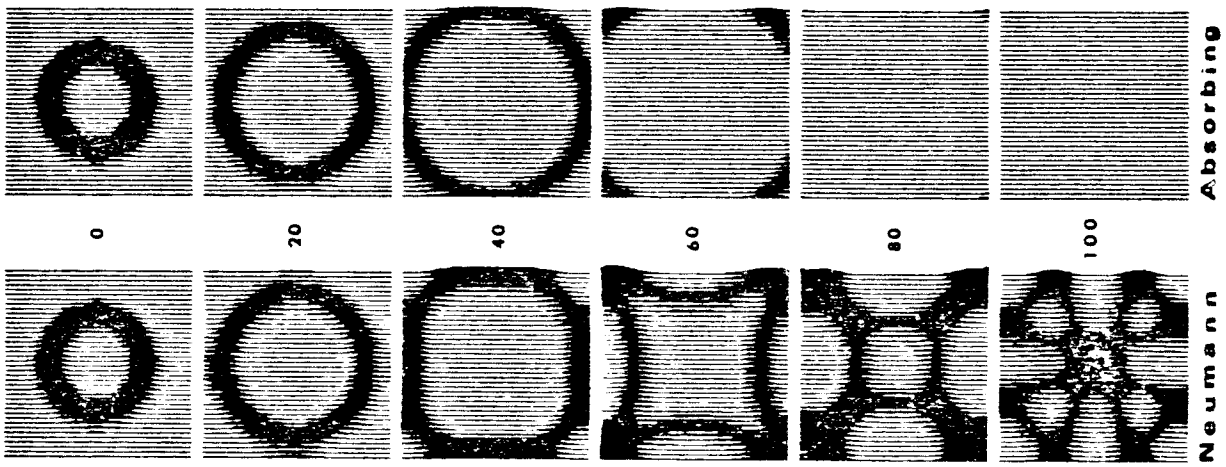

FIG. 1. (Clayton) Expanding circular wavefront in a box with absorbing sides (top) and with zero-slope sides (bottom).

Engquist Side Conditions for the Extrapolation Equations

In data processing we use the extrapolation equation in the interior of the region under study. This is unlike the forward modeling in which the full scalar wave equation is used in the interior and an extrapolation equation can be used on the boundary. The scalar wave equation has a circular dispersion relation whereas the extrapolation equation ideally has a semi-circular one. Reasoning by analogy, Engquist speculated that a quarter-circular dispersion relation might be some sort of ideal side boundary for wave-extrapolation problems. To be more specific and immediately applicable he proposed that the quarter circle be approximated by a straight line. This is depicted in figure 2.

The advantage of the straight-line dispersion relation is that in the space domain it represents a very simple, first-order, differential equation. A first-order equation has first derivatives which can be expressed over just two data points. Thus it can be used as a conventional, two-point, side-boundary condition. The right-side equation on figure 2 defines the boundary dispersion relation D.

$$0 = \frac{vk_z}{\omega} - 1 + const\ \frac{k_x}{\omega} = D(\omega,k_x,k_z) \tag{1}$$

In (t,x,z)-space it is

$$0 = (v\ \frac{\partial}{\partial z} + \frac{\partial}{\partial t} + const\ \frac{\partial}{\partial x})P \tag{2}$$

In retarded time, $\partial/\partial z$ may be eliminated with the interior equation.

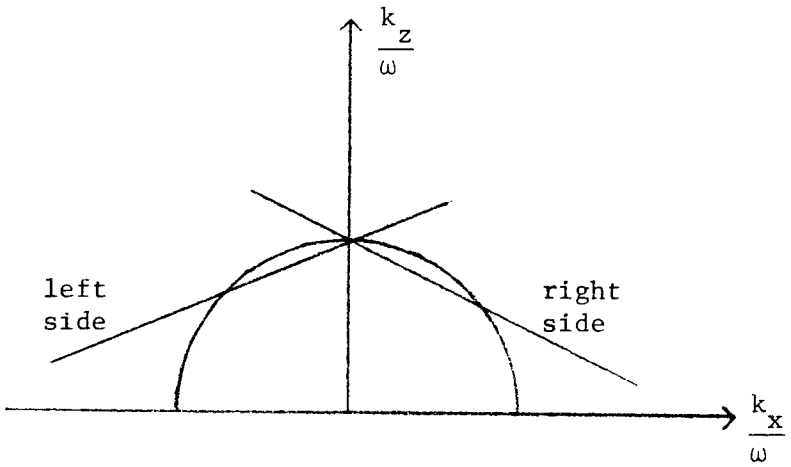

FIG. 2. Dispersion relation of simple absorbing side conditions.

A mathematical point of view which is non-physical is to imagine some peculiar physics which prescribe that the physical equation which applies in some region is just that which has the dispersion relation of the absorbing side condition. Aside this fictitious region is another in which the extrapolation equation applies. At the point of contact the solutions must match. It may come as no great surprise that the smallest boundary reflections occur where the two dispersion relations are a good match to each other. So the slope of the straight line is picked to form a good fit over the range of angles of interest. A nice example of side-boundary absorption for the diffraction equation is shown in figure 3, which is a reproduction of a result of Clayton in SEP-10, p 24.

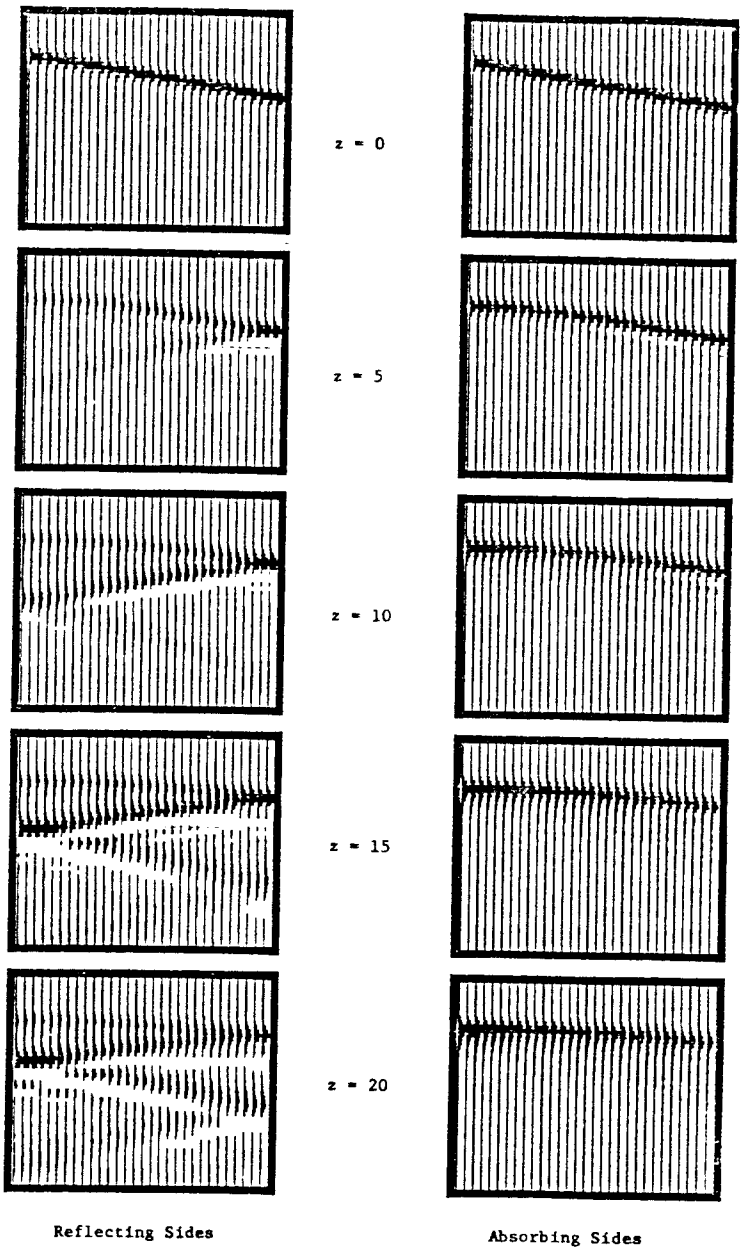

FIG. 3. (Clayton) A comparison of zero-slope side conditions versus absorbing sides. In this figure it is the diffraction equation, not the migration equation, which is used. The first arriving energy is drifting rightward and being absorbed at the right boundary. No energy enters at the left boundary. So we see a weakening diffraction on the left. On the right the amplitudes appear unchanged because each trace is rescaled to unit amplitude for display.

Size of the Reflection Coefficient

Let us look at some of the details of the reflection coefficient calculation. A mathematical expression for a unit amplitude monochromatic plane-wave incident on the side boundary superposed with a reflected wave of magnitude c is given by

$$P(x,z) = e^{-i\omega t + ik_z z}\left[e^{+ik_x x} + c\, e^{-ik_x x}\right] \tag{3}$$

In equation (3) we understand ω and k_x to be arbitrary and k_z to be determined from ω and k_x by the dispersion relation of the *interior* region, i.e. a semi-circle approximation. Assuming this interior solution to be applicable at the side boundary we may insert (3) into the differential equation (2) which represents the side boundary. Doing so we find $\partial/\partial x$ converted to $+ik_x$ on the incident wave, $\partial/\partial x$ converted to $-ik_x$ on the reflected wave, and $\partial/\partial z$ converted to ik_z. Thus the first term in (3) produces the dispersion relation $D(\omega,k_x,k_z)$, times the amplitude P. The second term produces the reflection coefficient c times $D(\omega,-k_x,k_z)$, times P.

$$c = \frac{-D(\omega,k_x,k_z)}{D(\omega,-k_x,k_z)} \tag{4}$$

The case of zero reflection arises when the numerical value of k_z selected by the interior equation at (ω,k_x) happens also to satisfy exactly the dispersion relation D of the side boundary condition. That explains why we try to match the quarter circle as well as we can. You might suspect that the straight line dispersion relation corresponds to the most general form of a side boundary condition, which is expressible on just two end points. Actually a more general expression with an extra adjustable parameter b_3 which fits even better is

$$D(\omega,k_x,k_z) = \left[1 - b_3 \frac{vk_x}{\omega}\right]\frac{vk_z}{\omega} - \left[b_1 - b_2 \frac{vk_x}{\omega}\right]$$

This choice was considered in more detail by Claerbout and Clayton in SEP-10. In fact figure 3 was computed with this side condition.

Stability

Claerbout and Clayton established absolute stability of absorbing side boundaries for the 15-degree equation including the discretization of the x-axis. Unfortunately an air-tight analysis of stability seems to be outside the framework of the Muir impedance rules. Consequently I don't believe the stability has been established for

the 45-degree equation. I'd like to speculate that a stability analysis may eventually fit within the "causal branch cut" analysis in SEP-20, but that is pure speculation. It has worked in examples, but that doesn't prove it always will.