Chapter 9

The craft of wavefield extrapolation

9.1 ABSORBING SIDES

Computer memory cells are often used to model points in a volume that contains propagating waves. Though we often wish to model an infinite volume, the number of computer cells is, regrettably, finite. Waves in the computer reflect back from the boundaries of the finite computer memory when we would prefer that the waves had gone away to infinity. To avoid the need for infinite computer capacity this section develops the theory of absorptive side boundary conditions.

There are two kinds of side boundary difficulties. First is where we are at the end of our observations. Second is where we somehow decided to limit the extent of our calculations. These boundaries might be the same. But to avoid confusion, let us presume that the data is of more limited extent than the computer memory. So alongside the data, which comprises the initial conditions for the calculation, is a region which will be called the *data padding* .

9.1.1 Data padding

The crudest assumption is that additional zero-valued data may be presumed for the padded area. To avoid an edge diffraction artifact the data must merge smoothly with the padding. So zero padding is a good assumption only if the data is already small around the side boundary. When data is zero padded, it is debatable whether or not it should be tapered (gradually scaled to zero) to match up smoothly with the zero padding. I prefer to avoid tapering the data. That amounts to falsifying it. Instead I prefer to pad the data not with zeroes, but with something that looks more like the data. A simple way is to replicate the last trace, scaling it downward with distance from the boundary. This works best when the stepout of the data matches the stepout of the extension. Any theory for optimum data padding has two important ingredients: a noise model and a signal model. An ideal data extrapolation is rarely, if ever, available in practice. My other book *Earth soundings analysis: processing versus inversion* addresses more directly the question of extending gathers.

9.1.2 Truncation at cable ends and at survey ends

In exploration there are two kinds of horizontal truncation problems. The first, which is at the end of the geophone cable, affects mainly common-midpoint stacking. The second is at the geographical boundaries of the survey and affects mainly migration. In both migration and stacking, hyperboloids are collapsed to points. But the processes differ because of the data itself. With stacking, it is predictable that energy will dip downward toward the far-offset cable truncation. With migration, reflectors can dip either downward or upward at the ends of the section. The downward-dipping case is better behaved. There seismic events move smoothly from the boundary to the interior.

The troublesome case occurs with migration when the seismic events dip upward at the edges of the survey. Then downward continuation moves seismic energy toward the boundary. On arriving at the boundary, it bounces back with opposite dip, and interferes with energy still moving toward the boundary. The problem may be reduced by appending space to the sides of the dataset, thus providing the dipping energy with a place to go. (You have previously decided what initial data padding to put in this space).

9.1.3 Engquist boundaries for the scalar wave equation

The simplest "textbook" 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 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 value and slope. This is also a two-point boundary condition. It so happens that our extrapolation equations (chapter 3) 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 our 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 suffer severely the consequences of limited memory. Engquist's idea was that they should use our extrapolation equations for their boundary conditions. (This idea led to his winning the SIAM prize). Suppose they desire an infinite absorbing volume surrounding a box in the (x, z)-plane. Then they need a boundary condition that goes all the way around the box. They could use the downgoing wave equation on the bottom of the box and the 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. For an example of one of his comparisons, see Figure 1.1.

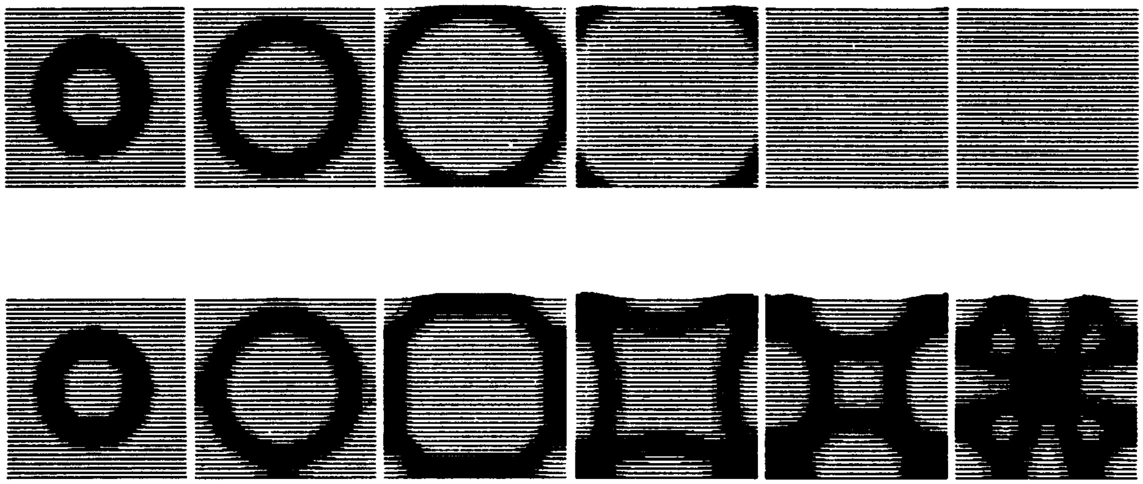

Figure 9.1: Expanding circular wavefront in a box with absorbing sides (top) and zero-slope sides (bottom). (Clayton) crft-box [NR]

9.1.4 Engquist sides for the extrapolation equation

In data processing an extrapolation equation is used in the interior of the region under study. This is unlike 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 has, ideally, a semicircular one. Reasoning by analogy, Engquist speculated that a quarter-circular dispersion relation might be an ideal side boundary for wave-extrapolation problems. To make his idea more specific and immediately applicable, he proposed that the quarter-circle be approximated by a straight line. This is shown in Figure 1.2.

The advantage of the straight-line dispersion relation is that in the space domain it is represented by a simple, first-order, differential equation. A first-order equation has first derivatives that can be expressed over just two data points, and thus it can be used as a conventional, two-point, side-boundary condition. The right-side equation in Figure 1.2 defines the boundary dispersion relation D:

$$0 \quad = \quad \frac{vk_z}{\omega} - 1 + \text{const}\ \frac{k_x}{\omega} \quad = \quad D(\omega, k_x, k_z) \tag{9.1}$$

In (t, x, z)-space this equation is

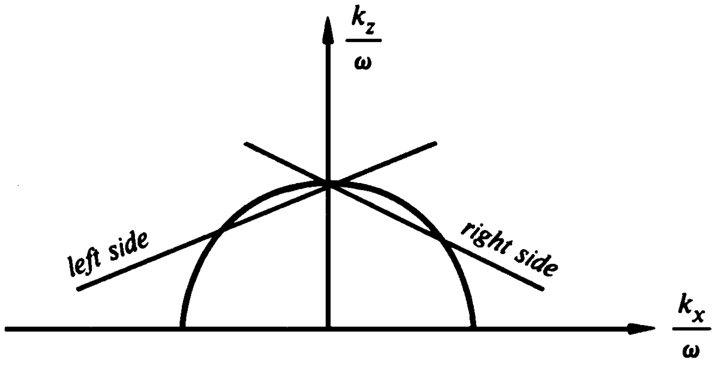

Figure 9.2: Dispersion relation of simple absorbing side conditions. crft-disperside [NR]

$$0 \quad = \quad (v\,\frac{\partial}{\partial z} \;+\; \frac{\partial}{\partial t} \;+\; \text{const}\;\frac{\partial}{\partial x}\,)\;P \tag{9.2}$$

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

For a mathematical, nonphysical point of view, imagine some peculiar physics which prescribe that the physical equation that applies in some region is just the equation which has the dispersion relation of the absorbing side condition. Beside this fictitious region imagine another in which the usual extrapolation equation applies. At the point of contact between the regions the solutions would match. It may come as no great surprise that the smallest boundary reflections would occur where the two dispersion relations were a good match to each other. So the slope of the straight line would be selected to form a good fit over the range of angles of interest. An example of side-boundary absorption during migration is shown in Figure 1.3.

9.1.5 Size of the reflection coefficient

Let us look at some of the details of the reflection coefficient calculation. A unit amplitude, monochromatic plane wave incident on the side boundary generates a reflected wave of magnitude c. The mathematical representation is:

$$P(x,z) \quad = \quad e^{-\,i\,\omega t + i\,k_z z}\,\left(\,e^{+\,i\,k_x x} \;+\; c\;e^{-\,i\,k_x x}\right) \tag{9.3}$$

In equation (1.3) ω and k_x are arbitrary, and k_z is determined from ω and k_x using the dispersion relation of the *interior* region, i.e., a semicircle approximation. Assuming this interior solution is applicable at the side boundary, you insert equation (1.3) into the differential equation (1.2), which represents the side boundary. As a result, $\partial/\partial x$ is converted to $+\,i\,k_x$ on the incident wave, and $\partial/\partial x$ is converted to $-\,i\,k_x$ on the reflected wave. Also, $\partial/\partial z$ is converted to ik_z. Thus the first term in (1.3) produces

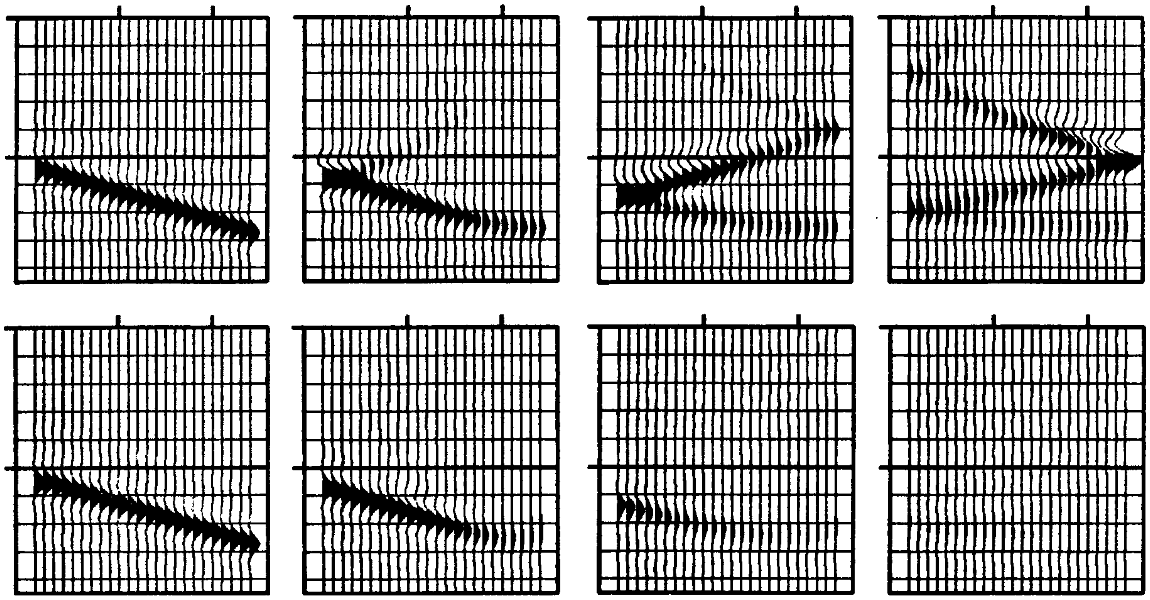

Figure 9.3: Downward continuation with zero slope side boundaries (top), and absorbing side boundaries (bottom). (Toldi) crft-sideabsorb [NR]

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. So (1.2) with (1.3) inserted becomes:

$$c \quad = \quad \frac{-\,D(\omega, k_x, k_z)}{D(\omega, -\,k_x, k_z)} \tag{9.4}$$

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. This explains why we try to match the quarter-circle as closely as possible. The straight-line dispersion relation does *not* correspond to the most general form of a side boundary condition, which is expressible on just two end points. A more general expression with adjustable parameters b_1, b_2, and b_3, which fits even better, is

$$D(\omega, k_x, k_z) \quad = \quad \left(1 \; - \; b_3 \, \frac{v k_x}{\omega}\right) \frac{v k_z}{\omega} \; - \; \left(b_1 \; - \; b_2 \, \frac{v k_x}{\omega}\right) \tag{9.5}$$

The absolute stability of straight-line absorbing side boundaries for the 15° equation can be established, including the discretization of the x-axis. Unfortunately, an airtight analysis of stability seems to be outside the framework of the Muir impedance rules. As a consequence, I don't believe that stability has been established for the 45° equation.

9.2 TUNING UP FOURIER MIGRATIONS

First we will see how to migrate dips greater than 90°. Then we will attack the main two disadvantages of Fourier migrations, namely, their periodicity and their poor tolerance of space-variable velocity.

9.2.1 Dips greater than 90 degrees

Migration of dips greater than 90° requires careful handling of evanescent energy. As this is being written, most migration-by-depth-extrapolation programs ignore or set to zero the energy that turns evanescent. The proper thing to do with energy becoming evanescent at depth z is to save it for a second pass *upward* . The upward pass begins from the bottom of the section with a zero downgoing wave. As the downgoing wave is extrapolated upward, the saved evanescent energy is reintroduced. As usual, the images are withdrawn from the wave at time $t = 0$.

To illustrate the concept, a program will be sketched that makes two images, first the usual image of the top side of the reflector, and second the image of the under side. The images may be viewed separately or summed.

The program makes the simplifying restriction on the velocity that $dv/dz \geq 0$. Because of this assumption, evanescent energy can be stored "in place" and ignored until the return pass. It is worth noting that the second pass is cheaper than the first pass because the region in which evanescence never occurred, $|k| < |\omega|/v(\tau_{\max})$, need not be processed.

first pass of conventional phase-shift migration.
$P(\omega, k_x) \ = \mathrm{FT}[u(t, x)]$
For $\tau = \triangle\tau,\ 2\triangle\tau,\ \ldots,\ \tau_{\max}$ {
    For all k_x {
        $\mathrm{Uimage}(k_x, \tau) \ = \ 0.$
        For all $\omega >\mid k \mid \ u(\tau)$ {
            $C \ = \ \exp(-i\omega\triangle\tau \sqrt{1 - v(\tau)^2 k_x^2/\omega^2}$
            $P(\omega, k_x) \ = \ P(\omega, k_x) \ * \ C$
            $\mathrm{Uimage}(k_x, \tau) \ = \mathrm{Uimage}(k_x, \tau) \ + \ P(\omega, k_x)$
            }
        }
    $\mathrm{uimage}(x, \tau) \ = \mathrm{FT}[\mathrm{Uimage}(k_x, \tau)]$
    }

```
# Second pass for underside image.
For τ = τmax, τmax − △τ, τmax − 2△τ, ..., 0 {
    For all kx {
        Dimage(kx, τ) = 0.
        For all ω =| k | v(τ) to ω = | k | v(τmax {
            # The wave changes direction but so does △τ
            C = exp(−iω△τ sqrt(1 − v(τ)^2 kx^2/ω^2)
            P(ω, kx) = P(ω, kx) * C
            Dimage(kx, τ) =Dimage(kx, τ) + P(ω, kx)
            }
        }
    dimage(x, τ) = FT[Dimage(kx, τ)]
    }
```

9.2.2 Stopping phase-shift migration wraparound (S. Levin)

Figure 1.4 shows a family of hyperbolas. Notice that these hyperbolas do not extend

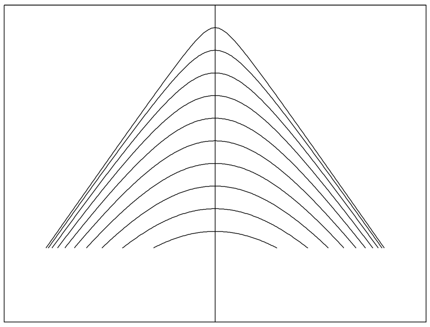

Figure 9.4: Hyperbolas truncated at a particular time. crft-hyptrunc [ER]

to infinite time but they truncate at a cut-off time t_c. A Fourier method will be described to create such time-truncated data. The method leads to a phase-shift migration program without wraparound artifacts.

When the fast Fourier transformation algorithm first came into use people noticed that it could be used for filtering. Transient filtering could be done exactly in the periodic Fourier domain if signals and filters were surrounded by enough zero padding. The same concept applies with migration. If field data and migration hyperbolas are surrounded by enough zeroes in the time- and space-domain then migration can be done in the Fourier domain with *no* wraparound. The trick is to see how the truncated hyperbolas in Figure 1.4 can be constructed in the Fourier domain.

To have truncations at time t_c, special point sources must be used. The deeper the source, the narrower must be its angular aperture. Take a hyperbola with first arrival at time t_0 to be truncated at some time t_c. The propagation angle θ of energy at the cutoff is given by cos $\theta = t_0/t_c$. So exploding reflectors have their k_x-spectrum truncated at $\sin\theta = vk_x\omega$. A 90° aperture implies echoes with an infinite time delay. Here is a sketch of the program.

```
# Modeling with time truncation at t_c
Model(k_x, z)  = FT[model(x, z)]
For all ω and all k_x
    U(ω, k_x)  =  0.
For z  =  z_max,  z_max − △z,  . . . , 0 {
    For all ω {
        For all | k_x |<| ω | /v {
            if (z  < v t_c) {
                sine =  sqrt(1 − z^2/v^2 t_c^2)
                if( | v k_x | < | ω | sine  )
                    aperture =  1.
                else
                    aperture =  0.
                }
            else
                aperture =  0.
            U(ω, k_x) = U(ω, k_x) e^{−iΔzω sqrt(v^{−2} − k_x^2/ω^2)} + aperture×Model(k_x, z)
            }
        }
    }
u(t, x)  = FT2D[U(ω, k_x)]
```

The above modeling program may be converted to a migration program by running the depth z loop down instead of up and by multiplying the downward continued data by the aperture function. The modifications to the program not only improve the quality of the migration, but the calculation is faster.

9.2.3 Controlling Stolt migration wraparound

As with the phase-shift algorithm, the key to reducing the computational artifacts of the Stolt algorithm is to suppress the time-domain wraparound. We will see that this amounts to accurate frequency-domain interpolation.

First, consider an impulse function at time t_0. Its Fourier transform is $\exp(-i\omega t_0)$. If t_0 is large, then the Fourier transform is a rapidly oscillating function of ω. Rapidly oscillatory functions are always difficult to interpolate. It is better to shift backward the time function, thereby smoothing its frequency function, then to interpolate the frequency function, and finally to undo the shift. Given seismic data on the time interval $0 < t < T$, the frequency function will be smoother if the data is shifted to an interval $-T/2 < t < T/2$. So the first proposed improvement to the Stolt migration program is to multiply in the frequency domain by $\exp(i\omega T/2)$, then interpolate, and finally multiply by $\exp(-i\omega T/2)$.

Linear interpolation is almost the easiest form of interpolation. On the other hand, Fourier transform theory suggests interpolation with the *sinc* function (by definition $\operatorname{sinc} u = (\sin u)/u$). The *sinc* function of frequency, when brought back into the time domain is a rectangle function of time. Take this rectangle function to be nonzero on the interval $-T/2 < t < T/2$. Recall that the fast (inverse) Fourier transform algorithm sums at uniform intervals in the frequency domain. This implicitly assumes zero between sample points, which in turn assumes that the time-domain function is periodic outside the given time interval. Now take the rectangle function of time to be the multiplier in the time domain that converts the periodic time function to the observed transient one. This multiplication in the time domain is equivalent to a convolution in the frequency domain with the appropriate *sinc* function. Convolution of the continuous *sinc* function with the given discrete-interval frequency function is really interpolation. Unfortunately, the *sinc* function extends infinitely down the frequency axis. Worse yet, it decays slowly. So some approximation or truncation of the *sinc* is used. Bill Harlan showed that tapered *sinc* functions achieve satisfactory accuracy more cheaply than zero padding. It seems, however, that the best approach is both to zero pad and to use some *sinc*-like interpolation. A definitive study of interpolation is that of Rosenbaum and Boudreaux [1981].

9.2.4 Stolt stretch

The great strength and the great weakness of the Stolt migration method is that it uses Fourier transformation over depth. This is a strength because it makes his method much faster than all other methods. And it is a weakness because it requires a velocity that is a constant function of depth. The earth velocity typically ranges over a factor of two within the seismic section, and the effect of velocity on migration tends to go as its square. To ameliorate this difficulty, Stolt suggested stretching the time axis to make the data look more like it had come from a constant-velocity earth. Stolt proposed the stretching function

$$\tau(t) \quad = \quad \sqrt{\frac{2}{v_0^2} \int_0^t \; t \; v_{\mathrm{RMS}}^2(t) \; dt} \tag{9.6}$$

where

$$v_{\mathrm{RMS}}^2(t) \quad = \quad \frac{1}{t} \int_0^t \; v^2(t) \; dt \tag{9.7}$$

At late times, which are associated with high velocities, Stolt's stretch implies that τ grows faster than t. The τ-axis is uniformly sampled to allow the fast Fourier transform. Thus at late time the samples are increasingly dense on the t-axis. This is the opposite of what earth Q and the sampling theorem suggest, but most people consider this a fair price.

The most straightforward derivation of (1.6) and (1.7) is based on the idea of matching the curvature of ideal hyperbola tops to the curvature on the stretched data. The equation of an ideal hyperbola in (x, τ)-space is

$$v_0^2 \, \tau^2 \quad = \quad x^2 \ + \ z^2 \tag{9.8}$$

Simple differentiation shows that the curvature at the hyperbola top is

$$\left. \frac{d^2 \tau}{dx^2} \right|_{x=0} \quad = \quad \frac{1}{\tau \ v_0^2} \tag{9.9}$$

It can be shown that in a stratified medium, equation (1.9) applies, except that the velocity must be replaced by the RMS velocity:

$$\left. \frac{d^2 t}{dx^2} \right|_{x=0} \quad = \quad \frac{1}{t \ v_{\mathrm{RMS}}^2} \tag{9.10}$$

We seek a stretched time $\tau(t)$. We would like to match the curves $t(x)$ and $\tau(x)$ for all x. But that would overdetermine the problem. Instead we could just match the derivatives at the hyperbola top, i.e., the second derivative of $\tau[t(x)]$ with respect to x at $x = 0$. With the substitutions (1.9) and (1.10), this would give an expression for $\tau \, d\tau/dt$ which, after integrating and taking its square root, yields (1.6) and (1.7).

A different derivation of the stretch gives a more accurate result at steeper angles. Instead of matching hyperbola curvature at the top, we go some distance out on the flank and match the slope and value. It is the flanks of the hyperbola that actually migrate, not the tops, so this result is more accurate. Algebraically the derivation is also easier, because only *first* derivatives are needed. Differentiating equation (1.8) with respect to x for a reflector at any depth z_j gives

$$\frac{d\tau}{dx} \quad = \quad \frac{x}{\tau \ v_0^2} \tag{9.11}$$

There is an analogous expression in a stratified medium. To obtain it, solve $x = \int v \sin\theta \, dt = p \int v^2 dt$ for $p = dt/dx$:

$$\frac{dt}{dx} \quad = \quad \frac{x}{\int_0^t \ v^2(p,t) \ dt} \tag{9.12}$$

Expressions (1.11) and (1.12) play the same role as (1.9) and (1.10), but (1.11) and (1.12) are valid everywhere, not just at the hyperbola top. Differentiating $\tau(t)$ gives

$$\frac{d\tau}{dx} \quad = \quad \frac{d\tau}{dt} \, \frac{dt}{dx} \tag{9.13}$$

Inserting (1.11) and (1.12) into (1.13) gives

$$\frac{x}{\tau \ v_0^2} \quad = \quad \frac{d\tau}{dt} \, \frac{x}{\int_0^t \ v^2(p,t) \ dt} \tag{9.14}$$

$$\tau \ d\tau \quad = \quad \left[\ \frac{1}{v_0^2} \int_0^t \ v^2(p,t') \ dt' \ \right] \ dt \tag{9.15}$$

Integrating (1.15) gives $\tau^2/2$ on the left. Then, taking the square root gives (1.6) but with a new definition for RMS velocity:

$$v_{\mathrm{RMS}}^2(t) \quad = \quad \frac{1}{t} \int_0^t \ v^2(p,t) \ dt \tag{9.16}$$

The thing that is new is the presence of the Snell parameter p. In a stratified medium characterized by some velocity, say, $v'(z)$, the velocity $v(p,t)$ is defined for the tip of the ray that left the surface at an angle with a stepout p. In practice, what value of p should be used? The best procedure is to look at the data and measure the $p = dt/dx$ of those events that you wish to migrate well. A default value is $p = 2(\sin 30°)/(2.5$ km/sec$) = .4$ millisec/meter. The factor of 2 is from the exploding-reflector model.

9.2.5 Gazdag's v(x) method

The phase-shift method of migration is attractive because it allows for arbitrary depth variation in velocity and arbitrary angles of propagation up to 90°. Unfortunately, lateral variation in velocity is not permitted because of the Fourier transformation over the x-axis. To alleviate this difficulty, Gazdag and Sguazzero [1984] proposed an interpolation method. Recall from chapter 3 that the phase-shift method 2-D Fourier transforms the data $p(x,t)$ to $P(k_x,\omega)$. Then $P(k_x,\omega)$ is downward continued in steps of depth by multiplication with $\exp[ik_z(\omega,k_x)\,\Delta z]$. Gazdag proposed several reference velocities, say, v_1, v_2, v_3, and v_4. He downward continued one depth step with each of the velocities, obtaining several reference copies of the downward-continued data, say, P_1, P_2, P_3, and P_4. Then he inverse Fourier transformed each of the P_j over k_x to $p_j(x,\omega)$. At each x, he interpolated the reference waves of nearest velocity to get a final value, say, $p(x,\omega)$ which he retransformed to $P(k_x,\omega)$ ready for another step. This appears to be an inefficient method since it duplicates the usual migration computation for each velocity. Surprisingly, the method seems to be successful, perhaps because of the peculiar nature of computation using an array processor.

EXERCISES:

1 To obtain a sharp cutoff in time t_c requires a broad bandwidth in the spectral domain. Given that Figure 1.4 is expressed on a 1000×1000 mesh, deduce the uncertainty in the cutoff t_c.

2 The phase-shift method tends to produce a migration that is periodic with z because of the periodicity of the Fourier transform over t. Ordinarily, this is not troublesome because we do not look at large z. The upcoming wave at great depth should be zero before $t = 0$. Kjartansson pointed out that periodicity in

z could be avoided if the wave at $t = 0$ is subtracted from the wavefield before the computation descends further. Thus, information could never get to negative time and "wrap around." Indicate how the program should be changed.

9.3 IMPEDANCE

Classical physics gives much attention to energy conservation and dissipation. Engineering filter theory gives much attention to causality—that there can be no response before the excitation. In geophysics we often need to ensure both causality and energy loss. We need to incorporate both, not only in theoretical derivations, but also in computations, and sometimes in computations that are discretized in time. There is a special class of mathematical functions called *impedance* functions that describe causal, linear disturbances in physical objects that dissipate energy.

Nature evolves forward in time. Naturally, impedance functions play a fundamental role in any modeling calculation where time evolves from past to future. Besides their use in physical modeling, impedances also find use in the depth extrapolation of waves. We geophysicists take data on the earth's surface and extrapolate downward to get information at depth. It is not the same as nature's extrapolation in time. In principle we don't require impedance functions to extrapolate in depth. But depth extrapolations made without impedance functions could exhibit growing oscillations, much like a physical system receiving energy from an external source. In fact, "straightforward" implementations of physical equations often exhibit unstable extrapolations. By formulating our extrapolation problems with impedance functions, we ensure stability. Of all the virtues a computational algorithm can have—stability, accuracy, clarity, generality, speed, modularity, etc.—the most important seems to be stability.

In this section we examine the theory of impedance functions, their precise definition, their computation in the world of discretized time, and the rules for combining simple impedances to get more complicated ones. We will also examine other special functions, the *minimum-phase* filter and the *reflectance* filter in their relation to the impedance filter. Wide-angle wave extrapolation and migration in the time domain will be formulated with impedance functions. Rocks are unlike "pure" substances because they contain irregularities at all scales. A particularly simple impedance function will be found that mimics the dissipation of energy in rocks, unlike the classical equations of Newtonian viscoelasticity.

9.3.1 Beware of infinity!

To prove that one equals zero take an infinite series, for example, $1, -1, +1, -1, +1,$ ... and group the terms in two different ways, and add them in this way:

$$(1-1)\ +\ (1-1)\ +\ (1-1)\ +\ \cdots \qquad = \qquad 1\ +\ (-1+1)\ +\ (-1+1)\ +\ \cdots$$

$$\begin{array}{rcl} 0 \; + \; 0 \; + \; 0 \; + \cdots & = & 1 + \; 0 \; + \; 0 \; + \; \cdots \\ 0 & = & 1 \end{array}$$

Of course this does not prove that one equals zero: it proves that care must be taken with infinite series. Next, take another infinite series in which the terms may be regrouped into any order without fear of paradoxical results. Let a pie be divided into halves. Let one of the halves be divided in two, giving two quarters. Then let one of the two quarters be divided into two eighths. Continue likewise. The infinite series is 1/2, 1/4, 1/8, 1/16, ... No matter how the pieces are rearranged, they should all fit back into the pie plate and exactly fill it.

The danger of infinite series is not that they have an infinite number of terms but that they may sum to infinity. Safety is assured if the sum of the absolute values of the terms is finite. Such a series is called *absolutely convergent.*

9.3.2 Z - transform

The Z-transform of an arbitrary, time-discretized function x_t is defined by

$$X(Z) \quad = \quad \cdots x_{-2}\,Z^{-2} \;+\; x_{-1}\,Z^{-1} \;+\; x_0 \;+\; x_1\,Z \;+\; x_2\,Z^2 \;+\; \cdots \tag{9.17}$$

Give Z the physical interpretation of time delay by one time unit. Then Z^2 delays two time units. Expressions like $X(Z)\,U(Z)$ and $X(Z)\,U(1/Z)$ are useful because they imply convolution and cross-correlation of the time-domain coefficients. (See FGDP).

Going on to consider numerical values for the delay operator Z, we discover that it is useful to ask whether $X(Z)$ is finite or infinite. Numerical values of Z that are of particular interest are $Z = +1$, $Z = -1$, and all those complex values of Z which are unit magnitude, say, $|Z| = 1$ or

$$Z \quad = \quad e^{i\,\omega\,\Delta t} \tag{9.18}$$

where ω is the real Fourier transform variable. Taking ω to be real means that Z is on the unit circle. Then the Z-transform is a discrete Fourier transform. Our attention can be restricted to time functions with a finite amount of energy by demanding that $U(Z)$ be finite for all values of Z on the unit circle $|Z| = 1$. Filter functions are always restricted to have finite energy.

The most straightforward way to say that a filter is causal is to say that its time domain coefficients vanish before zero lag, that is $u_t = 0$ for $t < 0$. Another way to say it is to say that $U(Z)$ is finite for $Z = 0$. At $Z = 0$ the Z-transform would be infinite if the coefficients u_{-1}, u_{-2}, etc. were not zero. For a causal function, each term in $|U(Z)|$ will be smaller if Z is taken inside the disk $|Z| < 1$ rather than on it. Thus convergence at $Z = 0$ and on the circle $|Z| = 1$ implies convergence everywhere inside the unit disk. So boundedness combined with causality means convergence in the unit disk. Convergence at $Z = 0$ but not on the circle $|Z| = 1$ would refer to a causal function with infinite energy, a case of no practical interest. What kind of function converges on the circle, at $Z = \infty$, but not at $Z = 0$? What function converges at all three places, $Z = 0$, $Z = \infty$, and $|Z| = 1$?

The filter $1/(1 - 2Z)$ can be expanded into powers of Z in (at least) two different ways. These are

$$\begin{aligned} \frac{1}{1 - 2Z} &= 1 + 2Z + 4Z^2 + 8Z^3 + \cdots \\ &= -\frac{1}{2Z}\frac{1}{1 - \frac{1}{2Z}} = -\frac{1}{2Z}\left[1 + \frac{1}{2Z} + \frac{1}{4Z^2} + \cdots\right] \end{aligned} \tag{9.19}$$

Which of these two infinite series converges depends on the numerical value of Z. For $|Z| = 1$ the first series diverges, but the second converges. So the only acceptable filter is *anticausal.* Is a series expansion unique? It is if it converges. Complex-variable theory proves this.

Let b_t denote a filter. Then a_t is its inverse filter if the convolution of a_t with b_t is a delta function. In the Fourier domain, we would say that filters are inverse to one another if their Fourier transforms are inverse to one another. Z-transforms can be used to define the inverse filter, say, $A(Z) = 1/B(Z)$. Whether the filter $A(Z)$ is causal depends on whether it is finite everywhere inside the unit disk, or really on whether $B(Z)$ vanishes *anywhere* inside the disk. For example, $B(Z) = 1 - 2Z$ vanishes at $Z = 1/2$. There $A(Z) = 1/B(Z)$ must be infinite, that is to say, the series $A(Z)$ must be nonconvergent at $Z = 1/2$. Thus—as we have just seen—a_t is noncausal. A most interesting case, called *minimum phase,* occurs when both a filter $B(Z)$ and its inverse are causal. In summary:

causal	$\mid B(Z) \mid < \infty$ for $\mid Z \mid \le 1$
causal inverse	$\mid 1/B(Z) \mid < \infty$ for $\mid Z \mid \le 1$
minimum phase	both above conditions

9.3.3 Review of impedance filters

Use Z-transform notation to define a filter $R(Z)$, its input $X(Z)$, and its output $Y(Z)$. Then

$$Y(Z) = R(Z)\, X(Z) \tag{9.20}$$

The filter $R(Z)$ is said to be *causal* if the series representation of $R(Z)$ has no negative powers of Z. In other words, y_t is determined from present and past values of x_t.

Additionally, the filter $R(Z)$ is *minimum phase* if $1/R(Z)$ has no negative powers of Z. This means that x_t can be determined from present and past values of y_t by straightforward polynomial division in

$$X(Z) \quad = \quad \frac{Y(Z)}{R}(Z) \tag{9.21}$$

Given that $R(Z)$ is already minimum phase, it can in addition be an impedance function if positive energy or work is represented by

$$
\begin{aligned}
0 \ \leq \ \text{work} \quad &= \quad \sum_t \ \text{force} \ \times \ \text{velocity} \ = \ \sum_{\text{t}} \ \text{voltage} \ \times \ \text{current} && (9.22)\\
&= \quad \frac{1}{2} \sum_t \ (\bar{x}_t y_t \ + \ \bar{y}_t x_t) && (9.23)\\
&= \quad \text{coef of } Z^0 \text{ of } \left[\ \bar{X} \ \left(\frac{1}{Z}\right) \ Y(Z) \ + \ \bar{Y} \ \left(\frac{1}{Z}\right) \ X(Z)\right] && (9.24)\\
&= \quad \frac{1}{2\pi} \int_0^{2\pi} \ \Re \ (\bar{X} \ Y) \ d\omega && (9.25)\\
&= \quad \int \ \Re \ (\bar{X} \ R \ X) \ d\omega \quad = \quad \int \ \bar{X} \ X \ \Re \ (R) \ d\omega && (9.26)
\end{aligned}
$$

Since $\bar{X}X$ could be an impulse function located at any ω, it follows that $\Re\left[R(\omega)\right] \geq 0$ for all real ω. In summary:

Definition of an Impedance	
causality	$r_t \ = \ 0$ for $t \ < 0$ i.e. $\mid R(Z) \mid < \ \infty$ for $\mid Z \mid \leq 1$
causal inverse	$\mid 1/R(Z) \mid < \ \infty$ for $\mid Z \mid \leq 1$
dissipates energy	$2\Re R(\omega) \ = \ R(Z) + \ \overline{R}(1/Z) \geq 0$ real ω

Adding an impedance to its Fourier conjugate gives a real positive function (the imaginary part of which is zero) like a power spectrum, say,

$$
\begin{aligned}
&\left(r_0 \ + \ r_1 Z \ + \ r_2 Z^2 \ + \ \cdots\right) \ + \\
&\qquad \left(\bar{r}_0 \ + \ \bar{r}_1 \frac{1}{Z} \ + \ \bar{r}_2 \frac{1}{Z^2} \ + \ \cdots\right) \quad \geq \quad 0 \quad \text{for real } \omega && (9.27)\\
&\qquad\qquad R(Z) \ + \ \bar{R}\left(\frac{1}{Z}\right) \quad \geq \quad 0 \quad \text{for real } \omega && (9.28)
\end{aligned}
$$

which is the basis for the remarkable fact that every impedance time function is one side of an autocorrelation function.

Impedances also arise in economic theory when X and Y are price and sales volume. I suppose that there the positivity of the impedance means that in the game of buying and selling you are bound to lose!

9.3.4 Causal integration

Begin with a function in discretized time p_t. The Fourier transform with the substitution $Z = \exp(i\omega\Delta t)$ is the Z-transform

$$P(Z) \quad = \quad \cdots p_{-2}\,Z^{-2} \;+\; p_{-1}\,Z^{-1} \;+\; p_0 \;+\; p_1\,Z \;+\; p_2\,Z^2 \;+\; \cdots \tag{9.29}$$

Define $-i\hat{\omega}$ (which will turn out to be an approximation to $-i\omega$) by

$$\frac{1}{-i\hat{\omega}\,\Delta t} \quad = \quad \frac{1}{2}\,\frac{1\;+\;Z}{1\;-\;Z} \tag{9.30}$$

Define another time function q_t with Z-transform $Q(Z)$ by applying the operator to $P(Z)$:

$$Q(Z) \quad = \quad \frac{1}{2}\,\frac{1\;+\;Z}{1\;-\;Z}\;P(Z) \tag{9.31}$$

Multiply both sides by $(1 - Z)$:

$$(1\;-\;Z)\;Q(Z) \quad = \quad \frac{1}{2}\;(1\;+\;Z)\;P(Z) \tag{9.32}$$

Equate the coefficient of Z^t on each side:

$$q_t \;-\; q_{t-1} \quad = \quad \frac{p_t\;+\;p_{t-1}}{2} \tag{9.33}$$

Taking p_t to be an impulse function, we see that q_t turns out to be a step function, that is,

$$p \quad = \quad \cdots\; 0,0,0,0,0,1,0,0,0,0,0,0,\cdots \tag{9.34}$$

$$q \quad = \quad \cdots 0,0,0,0,0,\frac{1}{2},1,1,1,1,1,1,\cdots \tag{9.35}$$

So q_t is the discrete domain representation of the integral of p_t from minus infinity to time t. It is the same as a Crank-Nicolson-style numerical integration of the differential equation $dQ/dt = P$. The operator $(1+Z)/(1-Z)$ is called the bilinear transform. The accuracy of the approximation to differentiation can be seen by multiplication on top and bottom by $Z^{-.5}$ and substitution of $Z = e^{i\omega\Delta t}$.

$$-\,i\,\frac{\hat{\omega}\,\Delta t}{2} \quad = \quad \frac{1\;-\;Z}{1\;+\;Z} \quad = \quad \frac{Z^{-.5}\;-\;Z^{+.5}}{Z^{-.5}\;+\;Z^{+.5}} \quad = \tag{9.36}$$

$$-\,i\,\frac{\hat{\omega}\,\Delta t}{2} \quad = \quad -\,i\,\frac{\sin\,\omega\,\Delta t/2}{\cos\,\omega\,\Delta t/2} \quad = \quad -\,i\,\tan\,\frac{\omega\,\Delta t}{2} \tag{9.37}$$

The integration operator has a pole at $Z = 1$, which is exactly on the unit circle. This raises the possibility of the paradox of infinity. In other words, there are other noncausal expansions too. For example, taking $1/(-i\omega)$ to be an imaginary, antisymmetric function of ω implies a real, antisymmetric time function, namely, $\mathrm{sgn}(t) = t/|t|$, which is not usually regarded as the integration operator. To avoid any ambiguity, we introduce here a small positive number ϵ and define $\rho = 1 - \epsilon$. The integration operator becomes

$$I = \frac{1}{2}\frac{1 + \rho Z}{1 - \rho Z} \tag{9.38}$$

$$= \frac{1}{2}(1 + \rho Z)\left[1 + \rho Z + (\rho Z)^2 + (\rho Z)^3 + \cdots\right] \tag{9.39}$$

$$= \frac{1}{2} + \rho Z + (\rho Z)^2 + (\rho Z)^3 + \cdots \tag{9.40}$$

Because ρ is slightly less than one, this series converges for any value of Z on the unit circle. If ϵ had been slightly negative instead of positive the expansion would have come out in negative instead of positive powers of Z.

Now the big news is that the causal integration operator is an example of an impedance function. The operator is clearly causal with a causal inverse. Let us check in the frequency domain that the real part is positive. Rationalizing the denominator gives

$$I = \frac{1}{2}\frac{(1 + \rho Z)}{(1 - \rho Z)}\frac{(1 - \rho/Z)}{(1 - \rho/Z)} = \frac{(1 - \rho^2) + \rho(Z - 1/Z)}{\text{positive}} \tag{9.41}$$

$$= \frac{(1 - \rho^2) + 2\,i\,\rho\sin\omega\Delta t}{\text{positive}} \tag{9.42}$$

Again, it is the choice of a positive ϵ that has caused $1 - \rho^2$, and hence the real part to be positive for all ω, as shown in Figure 1.5.

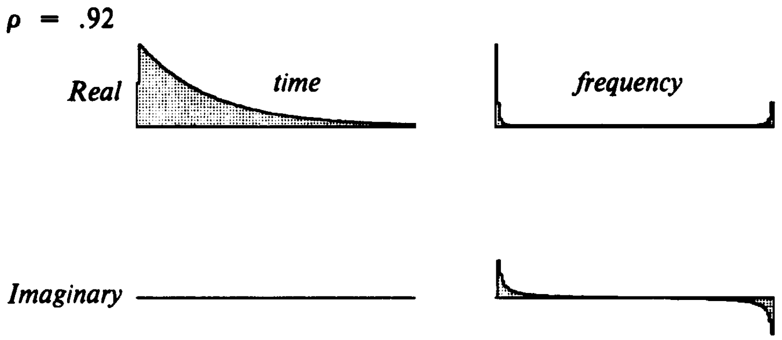

Figure 9.5: The causal integration operator I. The frequency axis is represented by a discrete Fourier transform over 256 points. Zero time and zero frequency are on the left end of their respective axes. crft-cintegral [NR]

As multiplication by $-i\omega$ in the frequency domain is associated with differentiation d/dt in the time domain, so is division by $-i\omega$ associated with integration. People usually associate the asymmetric operator $(1, -1)$ with differentiation. But notice that the inverse to the causal integration operator, namely,

$$I^{-1} \quad = \quad 2\,\frac{1\ -\ \rho Z}{1\ +\ \rho Z} \tag{9.43}$$

$$I^{-1} \quad = \quad 2\ -\ 4\rho Z\ +\ 4(\rho Z)^2\ -\ 4(\rho Z)^3\ +\ \cdots \tag{9.44}$$

also represents differentiation, although it is completely causal and not at all asymmetric. In linear systems analysis this representation of discrete differentiation is often the preferred one. The construction of higher-order, stable differential equations is subject to certain rules, to be covered, for combining impedances.

Occasionally it is necessary to have a *negative* real part for the differentiation operator. This can be achieved by taking ϵ to be negative, which means taking $\rho > 1$, and doing the infinite series expansion in powers of Z^{-1}, that is, anticausally instead of causally. In either the anticausal or the causal case the imaginary part will still be $-i\omega$, but the real part will have the opposite sign.

9.3.5 Muir's rules for combining impedances

For every physical system that conserves or dissipates energy there is an impedance function. Impedance functions are *special* combinations of differential operators and positive-valued physical constants. We will see just what combinations are allowed.

To ensure stable computations, it is important to be able to ensure that a supposed impedance function really is an impedance function. A difficulty in applied geophysics is this: Although you might require results only over a limited range of frequencies, and you might make approximations that are reasonable within that range, if the calculated impedance becomes negative outside the applicable range (it often happens near the Nyquist frequency), then the impedance filter will yield a numerically divergent output. So even though the impedance is almost correct, it is not usable.

Francis Muir provided three rules for combining simple impedances to get more complicated ones. [1] These rules are especially useful because we can start from the discrete-time forms of the differentiation and integration operators. Let R' denote a new impedance function generated from known impedance functions R, R_1, and R_2. These three ways of combining impedance are

1. Multiplication by positive scalar a $\qquad R' \ =\ a\,R$

2. Inversion $\qquad R' \ =\ \dfrac{1}{R}$

3. Addition $\qquad R' \ =\ R_1 + R_2$

[1]Personal communication with Francis Muir.

These rules do not include multiplication. Multiplication is not allowed because squaring, for example, doubles the phase angle, and thus may destroy the positivity of the real part. Since these rules do not include multiplication, but only scaling, summation, and inversion, the impedance functions that occur in nature will often be represented mathematically as *continued fractions.*

The first two of Muir's rules are so obvious we will not prove them. The third rule deserves more careful attention. To prove any rule, we need to show three things about R', namely, it is causal, it is PR (the Fourier transform has a positive real part), and it has an inverse. This last part is the hard part with Muir's third rule, namely, that the sum of two impedances has a causal inverse. Proof of this fact will take about two pages, and introduce several additional concepts.

9.3.6 Impedance defined from reflectance

The size of the class of filters called *impedances* will be seen to be large, because impedances are derived by transformation from an easily specified family of filters called *reflectances,* say, c_t and its Fourier transform $C(\omega)$. To be a reflectance, the time function must be strictly causal, and the frequency function must be strictly less than unity. By *strictly causal* it is meant that the time function vanishes both at zero time and before. For example, take $-1 \ < \ \rho \ < \ +1$ and the reflectance c_t to be an impulse of size ρ after a time Δt. The Fourier transform is

$$C \quad = \quad \rho Z \quad = \quad \rho \, e^{i\omega\Delta t} \tag{9.45}$$

Obviously, the product of two reflectances is another reflectance.

An impedance has been defined to be a causal filter with a causal inverse and a Fourier transform whose real part is positive. It will be shown that from any reflectance C the expression

$$R \quad = \quad \frac{1 \ - \ C}{1 \ + \ C} \tag{9.46}$$

generates an impedance. There are three things to show: R is causal, has a causal inverse, and is PR. First because of the assumption that C has a magnitude strictly less than unity, $C\bar{C} \ < \ 1$, the denominator expands to a convergent $1 + C + C^2 + \cdots$. Second, the inverse of R is found by simply changing the sign of C. Third, multiply top and bottom by the complex conjugate:

$$\Re\, R \quad = \quad \Re\, \frac{(1 \ - \ C)\,(1 \ + \ \bar{C})}{\text{positive}} \quad \overset{?}{\geq} \quad 0 \tag{9.47}$$

$$\Re\, R \quad = \quad \Re\, \frac{(1 \ - \ \bar{C}) \ \ + \ \ \text{imaginary}}{\text{positive}} \quad \geq \quad 0 \tag{9.48}$$

which shows that R has a real part that is positive.

The expression for $R(C)$ is easily back-solved for $C(R)$, but the converse theorem, that every R generates a reflectance, is harder to show. Nevertheless, it will be proved, along with a deeper theorem. A filter that is both causal and PR is said to be CPR. The deeper theorem is that *every CPR has an inverse and hence is an impedance.* This will be proved by showing that every CPR—say, $\hat{R}$, —can be used to construct a reflectance $\hat{C}$, which, since it is a reflectance, implies that the CPR $\hat{R}$ is an impedance R. Backsolving gives

$$\hat{C} \quad = \quad \frac{1 \ - \ \hat{R}}{1 \ + \ \hat{R}} \tag{9.49}$$

Proof requires that two things be shown—first, that the magnitude of $\hat{C}$ is less than unity. To show this, form the magnitude of the denominator and subtract the magnitude of the numerator. The result is four times the real part of $\hat{R}$, which is positive. Second, $\hat{C}$ must be proved causal. This is harder. $(1 \ + \ \hat{R})^{-1}$ can be expanded into a sum of positive powers of $\hat{R}$ and hence of positive powers of the delay operator. But the convergence of the series is not assured, because nothing requires $\hat{R}$ to be less than unity.

To prove that $\hat{C}$ is causal, we will take advantage of rule 1, namely, that an impedance can be scaled by any real positive number that you like, and it will still be an impedance. Consider a function that is similar to $\hat{C}$.

$$B \quad = \quad \frac{1 \ - \ \epsilon \hat{R}}{1 \ + \ \epsilon \hat{R}} \tag{9.50}$$

Choose ϵ small enough that for all ω, $\epsilon|\hat{R}| < 1$. This ensures a convergent expansion for the denominator in positive powers of $\hat{R}$ and hence Z. The expansion contains only positive powers in the delay operator. Thus B is a reflectance, and its corresponding impedance is $\hat{\epsilon}$. But an impedance can always be scaled by a positive number. Taking the number to be $1/\epsilon$ shows that $\hat{R}$ is an impedance. This completes the proof that every CPR is an impedance.

So impedances arise more easily than you might think. It is not necessary to have a reflectance C to insert into the relation $R = (1-C)/(1+C)$. We only need to have a CPR.

9.3.7 Functional analysis

We will establish the following theorems about exponentials, logarithms, and powers of Fourier transforms of filters:

- The exponential of a causal filter is causal.
- The exponential of a causal filter is a minimum-phase filter.
- The logarithm of a minimum-phase filter is causal.

- The Fourier domain representation of a minimum-phase filter is a curve that does not enclose the origin of the complex plane.
- Any power of a minimum-phase filter is minimum phase.
- Any real fractional power $-1 \leq \rho \leq 1$ of an impedance function is an impedance function.

To establish theorem 1, define the Z-transform of an arbitrary causal function

$$U(Z) \quad = \quad u_0 \ + \ u_1 \ Z \ + \ u_2 \ Z^2 \ + \ \cdots \tag{9.51}$$

and substitute it into the familiar power series for the exponential function:

$$B(Z) \quad = \quad e^U \quad = \quad 1 \ + \ U \ + \ \frac{U^2}{2!} \ + \ \frac{U^3}{3!} \ + \cdots \qquad (|U| \ < \ \infty) \tag{9.52}$$

No negative powers of Z can be found in the right side of (1.52), so $B(Z)$ will have no negative powers of Z. Also, the factorials in the denominator assure us that (1.52) always converges, so b_t is always causal.

To establish theorem 2, that the exponential is not just causal but also minimum phase, consider

$$B_+ \quad = \quad e^{+U} \tag{9.53}$$
$$B_- \quad = \quad e^{-U} \tag{9.54}$$

Clearly both B_+ and B_- are causal, and they are inverses of one another. A minimum-phase filter is defined to be causal with a causal inverse. So B_+ and B_- are minimum phase.

Now we will establish the converse of theorem 2—namely, theorem 3—which states that the logarithm of a minimum-phase filter is causal. Take the logarithm of (1.52) and form the Z-derivative:

$$U \quad = \quad \ln \ B \tag{9.55}$$
$$\frac{dU}{dZ} \quad = \quad u_1 \ + \ 2u_2 Z \ + \ 3u_3 Z^2 \ + \ \cdots \tag{9.56}$$
$$\frac{dU}{dZ} \quad = \quad \frac{1}{B} \, \frac{dB}{dZ} \tag{9.57}$$

Since B was assumed to be minimum phase, both $1/B$ and dB/dZ on the right of (1.57) are causal. Since the product of two causals is causal, dU/dZ is causal. But dU/dZ cannot be causal unless U is causal. That proves theorem 3, disregarding the remote danger that B might converge while dB/dZ diverges.

Theorem 4 refers to the Fourier domain representation of the minimum-phase filter. In the complex plane, the filter gives parametric equations for a curve, say $[x(\omega), y(\omega)] = [\Re B(Z), \, \Im B(Z)]$. The phase angle $\phi(\omega)$ is defined by the arctangent of the ratio y/x. For example, the causal, non-minimum-phase filter $U(Z) = Z = e^{i\omega}$ gives the parametric equations $x = \cos\omega$ and $y = \sin\omega$ which define a circle surrounding the origin. Notice that the phase of $Z = e^{i\omega}$ is $\phi(\omega) = \omega$, which is a monotonically increasing function of ω. In the minimum phase case, $\phi(\omega = 0) = \phi(\omega = 2\pi)$. In the non-minimum-phase case, the curve loops the origin, so $\phi(\omega = 0) = \phi(\omega = 2\pi) + 2\pi$. Theorem 3 allows us to say that a general formula for minimum-phase filters is

$$\begin{aligned} B \quad = \quad e^{U(Z)} \quad &= \quad \exp\left(\sum_{k=0}^{N} U_k \cos k\omega \; + \; i \sum_{k=0}^{N} U_k \sin k\omega \right) && (9.58) \\ &= \quad \exp[r(\omega) \; + \; i\,\phi(\omega)] && (9.59) \end{aligned}$$

The phase $\phi(\omega)$, being a sum of periodic functions, is itself a periodic function of ω, which means that in the plane of $(\Re B,\ \Im B)$ the curve representing $B(\omega)$ does not enclose the origin.

On to theorem 5, which says that any power of a minimum-phase function is minimum phase. Consider

$$B^r \quad = \quad \left(e^{\ln\,B}\right)^r \quad = \quad e^{r\ \ln\ B} \tag{9.60}$$

Since B is assumed to be minimum phase, by theorem 3, $\ln B$ will be causal. Scaling by a real or complex constant r does not change causality. Exponentiating shows, by theorem 2, that B^r is minimum phase.

Finally the proof of theorem 6, that an impedance function can be raised to any real fractional power $-1 \leq \rho \leq +1$ and the result will still be an impedance function. An impedance function is defined to be a minimum-phase function with the additional property that the real part of its Fourier transform is positive. This means that the phase angle ϕ lies in the range $-\pi/2 < \phi < +\pi/2$. Raising the impedance function to the ρ power will compress the range to $-\pi\rho/2 < \phi < \pi\rho/2$. This will keep the real part of the impedance function positive. Theorem 5 states that any power of a minimum-phase function is causal, which is more than we need to be certain that a *fractional* real power of an impedance function will be causal.

9.3.8 Wide-angle wave extrapolation

Let $s = -i\hat{\omega}$ denote the causal, positive, discrete representation of the differentiation operator, say,

$$s \quad = \quad -i\,\hat{\omega} \quad = \quad \frac{2}{\Delta t}\,\frac{1 \; - \; \rho Z}{1 \; + \; \rho Z} \tag{9.61}$$

Figure 1.6 compares hyperbolas constructed with ω to those constructed with $\hat{\omega}$. You see a pleasing drop in wraparound noise. It seems to work better than the ϵ in chapter 8. As we will see, the introduction of complex-valued $\hat{\omega}$ leads to a more natural handling of the square root at the evanescent transition.

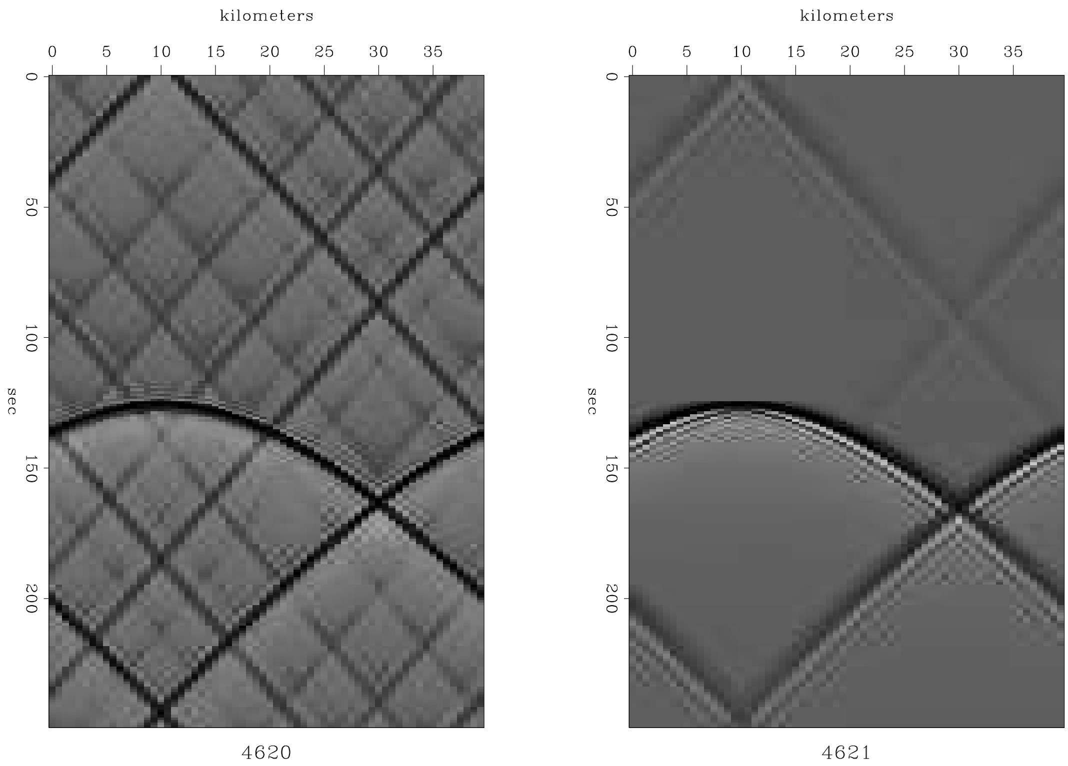

Figure 9.6: Hyperbolas with real frequency (left) and complex frequency (right). (Plotting uses square root gain described in chapter 8). crft-cxfreq [ER]

Consider the following recursion starting from $R_0 = s$:

$$R_{n+1} \quad = \quad s \ + \ \frac{X^2}{s \ + \ R_n} \tag{9.62}$$

This recursion produces continued fractions. Francis Muir introduced it as a means of developing wide-angle square-root approximations for migration (chapter 3). and he developed his three rules to show that every R_n is an impedance function. To see why every R_n is an impedance function, first note that the denominator $s + R_n$ is, for $n = 0$, the sum of two impedance functions. Then its inverse is an impedance function, and multiplication by the real positive constant X^2 and addition of another s both preserve the properties of impedance functions. Recursively we see that all the R_n are impedances.

As N becomes large this recursion either converges or it does not. Supposing that it does, we can see what it will converge to by setting $R_{n+1} = R_n = R_\infty = R$. Thus,

$$R \quad = \quad s \ + \ \frac{X^2}{s \ + \ R} \tag{9.63}$$

$$R\ (s \ + \ R) \quad = \quad s(s \ + \ R) \ + \ X^2 \tag{9.64}$$

$$R^2 \quad = \quad s^2 \ + \ X^2 \tag{9.65}$$

$$R \quad = \quad \sqrt{s^2 \ + \ X^2} \tag{9.66}$$

In wave-extrapolation problems X^2 is $v^2\ k_x^2$, where v is the wave velocity and k_x is the horizontal spatial frequency, namely, the Fourier dual to the horizontal x-axis. Performing these substitutions we have

$$R \quad = \quad \pm\sqrt{-\hat{\omega}^2 \ + \ v^2\, k_x^2} \tag{9.67}$$

So R is like $\pm ik_z v$. Remember that R_0, the first approximation to R, is $-i\hat{\omega}$. So downgoing waves are

$$D(x,z,t) \quad = \quad D(x,0,t)\ \ e^{ik_x x}\ e^{-\,R\,z/v}\ e^{-i\omega t} \tag{9.68}$$

To switch from downgoing to upcoming waves, we could either change the sign in front of R or we take the complex conjugate of R. The difference is what you want to do with the real part—do you want the wave to grow or not? Consider the dissipation of waves in the exploding reflector model. They damp as they propagate from the explosion to the surface. This means that as we migrate them, they should be exponentially growing. But we don't really want that. We really want to assure that they are not growing, perhaps we even want them decaying as we extrapolate them back. So for migration we downward continue monochromatic waves with

$$U(x,z,t) \quad = \quad U(x,0,t)\ \ e^{ik_x x}\ e^{-\bar{R}\,z/v}\ e^{-i\omega t} \tag{9.69}$$

although the real behavior of a wave from an exploding reflector wave would be

$$U(x,z,t) \quad = \quad U(x,0,t)\ \ e^{ik_x x}\ e^{+\,R\,z/v}\ e^{-i\omega t} \tag{9.70}$$

To examine the phase of the complex quantity R, set $v = 1$ obtaining

$$R \quad = \quad \sqrt{(-\,i\hat{\omega})^2 \ + \ k_x^2} \tag{9.71}$$

First note that $(-\,i\hat{\omega})$ is causal because of its Z-transform representation. By squaring the Z-transform we see that $(-\,i\hat{\omega})^2$ is also causal. In the time domain, k_x^2 is a delta function at the time origin. Thus R^2 given by (1.71) is causal. Figure 1.7 shows how the phase of (1.71) is constructed from its constituents. To illustrate the behavior of $-i\hat{\omega}$ from zero to infinity, I include both an artist's conception and the function on itself, overlain at various magnifications. The function $-i\hat{\omega}$ is a periodic with ω and its real and imaginary parts plot to a closed curve. To show the rate of change of the function, I sampled ω at 2° intervals. From great distance the function is a circle. Close up it looks like a line parallel to the imaginary axis.

R^2 is causal and from Figure 1.7 we can see that it has a "branch cut" property. That is, the phase of R has the positive real property. Theorem 5 forces R to be causal and minimum phase. That, with the phase defined by Figure 1.7, proves that R, given by (1.71), is an impedance function.

9.3.9 Fractional integration and constant Q

By equation (1.61) and theorem 6, fractional powers of integration and differentiation are also impedance functions. Kjartansson [1979] has advocated the fractional power as a stress-strain law for rocks. See also Madden [1976]. Classical studies in rock mechanics begin with a stress-strain law such as

$$\text{stress} \quad = \quad \text{stiffness} \quad \times \quad \text{strain} \quad + \quad \text{viscosity} \quad \times \quad \text{strain-rate}$$

which in the transform domain is

$$\text{stress} \quad = \quad [(-\,\mathrm{i}\omega)^0 \quad \times \quad \text{stiffness} \quad + \quad (-\,\mathrm{i}\omega)^1 \quad \times \quad \text{viscosity}] \,\times\, \text{strain} \tag{9.72}$$

Experimentally, the viscoelastic law (1.72) does a poor job of describing real rocks. Let us try another mathematical form that is like (1.72) in its limiting behavior at high and low viscosity:

$$\begin{aligned} \text{stress} \quad &= \quad \text{const} \ \times \ (-\,\mathrm{i}\omega)^{\epsilon} \ \times \ \text{strain} && (9.73)\\ &= \quad \text{const} \ \times \ (-\,\mathrm{i}\omega)^{\epsilon-1} \ \times \ \text{strain-rate} && (9.74) \end{aligned}$$

Here ϵ close to zero gives elastic behavior and ϵ close to one gives viscous behavior. The fact that $(-\,i\omega)^{\epsilon-1}$ is an impedance function meshes nicely with the concepts that (1.17) stress may be determined from strain history and strain may be determined from stress history, and (1.18) stress times strain-rate is dissipated power. Kjartansson [1979] points out that $(-\,i\omega)^{\gamma}$ exhibits the mathematical property called *constant Q,* so that as a stress/strain law for fitting experimental data on rocks, it is far

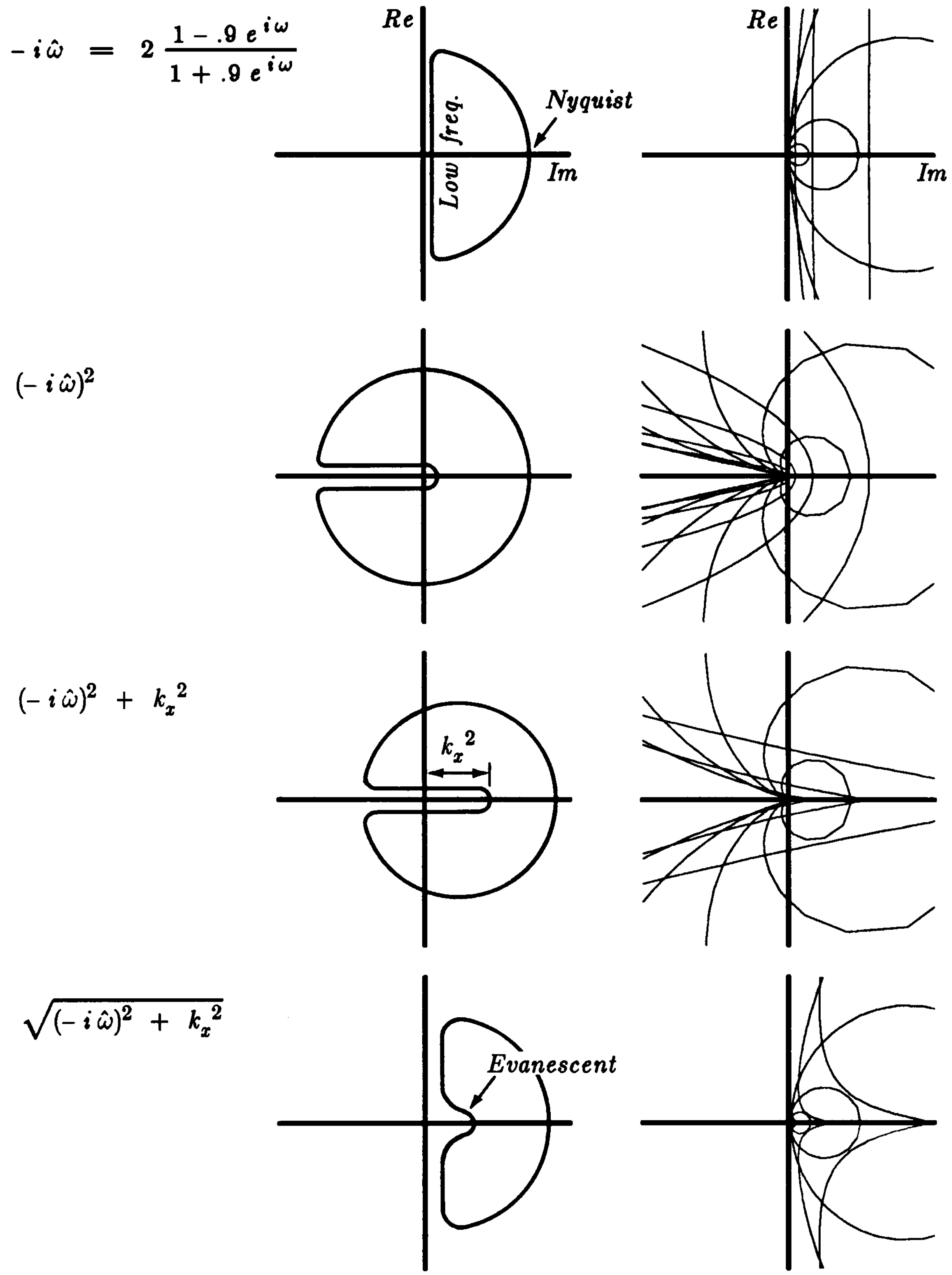

Figure 9.7: Complex plane diagram of constituents of the extrapolation operator R given by equation (1.71). The center column shows an artist's conception. The right column shows the function at several magnifications simultaneously. crft-kzplane [NR]

superior to (1.72). To see the constant Q property more clearly, express $(-i\omega)^\gamma$ in real and imaginary parts:

$$
\begin{aligned}
(-i\omega)^\gamma &= |\omega|^\gamma \, [-i \,\mathrm{sgn}(\omega)]^\gamma && (9.75)\\
&= |\omega|^\gamma \, [e^{-i\pi \; \mathrm{sgn}(\omega)/2}]^\gamma && (9.76)\\
&= |\omega|^\gamma \left\{ \cos\left[\frac{\pi\,\gamma}{2}\,\mathrm{sgn}(\omega)\right] - i \sin\left[\frac{\pi\,\gamma}{2}\,\mathrm{sgn}(\omega)\right] \right\} && (9.77)\\
&= |\omega|^\gamma \left[\cos\left(\frac{\pi\,\gamma}{2}\right) - i\,\mathrm{sgn}(\omega)\,\sin\left(\frac{\pi\,\gamma}{2}\right) \right] && (9.78)
\end{aligned}
$$

The constant Q property follows from the constant ratio between the real and imaginary parts of this function. Q itself is defined by

$$\frac{1}{Q} = \tan \pi\epsilon \approx \pi\epsilon \qquad (9.79)$$

A pulse with a Q of about 10 is shown in Figure 1.8.

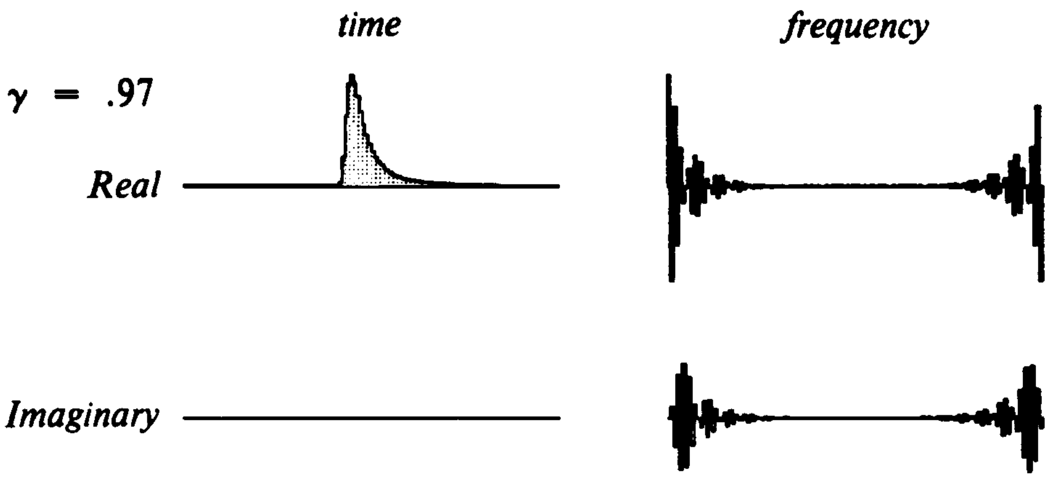

Figure 9.8: The constant Q pulse given by $e^{-(-i\omega)^{.97} t_0}$. The frequency axis is represented by a discrete Fourier transform over 256 points. Zero time and zero frequency are on the left end of their respective axes. crft-Qq [NR]

EXERCISES:

1 Take $\epsilon < 0$ and expand the integration operator for negative powers of Z. Explain the sign difference.

2 Let $\alpha > 0$ be a real, positive scaling constant, and let C be a reflectance function. Without using Muir's rules, prove that C' is a reflectance, where

$$\frac{1 - C'}{1 + C'} = \alpha \frac{1 - C}{1 + C}$$

Note that you have proven Muir's first rule. Muir's third rule can also be proven in an analogous way, but with much more algebraic detail.

3 The word *isomorphism* means not only that any impedance R_1, R_2, R' can be mapped into a reflectance C_1, C_2, C' , but also that Muir's three rules will be mapped into three rules for combining reflectances.

a. What are these three rules?

b. Although $C' = C_1 C_2$ does not turn out to be one of the three rules, it is obviously true. Either show that it is a consequence of the three rules or conclude that it is an independent rule that can be mapped back into the domain of the impedances to constitute a fourth rule.

4 Show that the log of the discrete causal integration operator, $\ln[(1+Z)/(1-Z)]$, is one side of the discrete Hilbert transform. Show that the reflected pulse from a boundary between two media with the same velocity but slightly different Q is one side of the Hilbert transform.

5 Consider the fourth-order Taylor expansion for square root in an extrapolation equation

$$\frac{dP}{dz} \quad = \quad i\omega \left[1 \ - \ \frac{1}{2} \left(\frac{vk}{\omega} \right)^2 \ - \ \frac{1}{8} \left(\frac{vk}{\omega} \right)^4 \right] \ P$$

a. Will this equation be stable for the complex frequency $-i\omega = -i\omega_0 + \epsilon$? Why?

b. Consider causal and anticausal time-domain calculations with the equation. Which, if any, is stable?

6 Consider a material velocity that may depend on the frequency ω and on the horizontal x-coordinate as well. Suppose that, luckily, the velocity can be expressed in the factored form $v(x, \omega) = v_1(x) \ v_2(\omega)$. Obtain a stable 45° wave-extrapolation equation. Hints: try

$$s \quad = \quad - \frac{i\omega}{v_2}$$

$$X^2 \quad = \quad \text{positive eigenvalue of } \ (v_1 \partial_x)(v_1 \partial_x)^T$$

7 Is the Levinson Recursion described in FGDP related to the rules in this section? If so, how? Hint: see Jones and Thron [1980].

8 Show the converse to theorem 4, namely, that if the phase curve of a causal function does not enclose the origin, then the inverse is causal.

9.4 THE BULLETPROOFING OF MUIR AND GODFREY

Stable extrapolation can be assured by preserving certain symmetries. It will be shown that stability is assured in both the differential equation

$$\frac{d\mathbf{q}}{dz} \quad = \quad - \, \mathbf{R} \ \mathbf{q} \tag{9.80}$$

and its Crank-Nicolson approximation

$$\frac{\mathbf{q}_{n+1} \;-\; \mathbf{q}_n}{\Delta z} \quad = \quad -\,\frac{\mathbf{R}}{2}\,(\mathbf{q}_{n+1} \;+\; \mathbf{q}_n) \tag{9.81}$$

provided that $\mathbf{R} \;+\; \mathbf{R}^*$ is a positive definite (actually, semidefinite) matrix. When stability was studied in the previous section the operator $\mathbf{R}$ was a scalar Z-transform. Because Z-transforms were used, the mathematics of that section was particularly suitable for time domain migrations. Because $\mathbf{R}$ was a scalar, the mathematics of that section was particularly suitable when data has been Fourier transformed over x. Here we will focus on the *matrix* character of $\mathbf{R}$. Thus we are concerning ourselves with migration in the x-domain. Our purpose in doing this theoretical work is to gain the ability to write a "bulletproof" program for migrating seismic data in the presence of lateral velocity variation. As an example, the familiar 45° extrapolation equation will be put in the bulletproof form. This section, combined with the previous one, gives a general theory for stable migration in (t, x)-space.

9.4.1 Stability of the differential equation

Let $\mathbf{q}^*$ denote the Hermitian conjugate of $\mathbf{q}$. For equation (1.80) to be stable the energy $\mathbf{q}^*\,\mathbf{q}$ must be either constant or decaying during depth extrapolation.

$$\begin{aligned} \frac{d}{dz}\,(\mathbf{q}^*\,\mathbf{q}) \quad &\leq \quad 0 \\ \mathbf{q}^*\,\mathbf{q}_z \;+\; \mathbf{q}_z^*\,\mathbf{q} \quad &\leq \quad 0 \end{aligned} \tag{9.82}$$

Substituting equation (1.80) into equation (1.82) gives

$$\begin{aligned} \mathbf{q}^*\,\mathbf{R}\,\mathbf{q} \;+\; \mathbf{q}^*\mathbf{R}^*\,\mathbf{q} \quad &\geq \quad 0 \\ \mathbf{q}^*\,(\mathbf{R} \;+\; \mathbf{R}^*)\mathbf{q} \quad &\geq \quad 0 \end{aligned} \tag{9.83}$$

Equation (1.83) shows that $\mathbf{R} \;+\; \mathbf{R}^*$ must be positive semidefinite for the differential equation to be stable.

9.4.2 Stability of the difference equation

The stability of the difference equation can be shown in a similar way, but with some extra clutter. First observe the identity

$$(\mathbf{a}^*\,\mathbf{a} \;-\; \mathbf{b}^*\,\mathbf{b}) \quad \equiv \quad \frac{1}{2}\,[(\mathbf{a} \;+\; \mathbf{b})^*\,(\mathbf{a} \;-\; \mathbf{b}) \;+\; (\mathbf{a} \;-\; \mathbf{b})^*\,(\mathbf{a} \;+\; \mathbf{b})] \tag{9.84}$$

Letting $\mathbf{a} \;=\; \mathbf{q}_{n+1}$ and $\mathbf{b} \;=\; \mathbf{q}_n$, equation (1.84) becomes

$$(\mathbf{q}_{n+1}^*\mathbf{q}_{n+1} \;-\; \mathbf{q}_n^*\mathbf{q}_n) \quad =$$

$$= \quad \frac{1}{2}\left[(\mathbf{q}_{n+1} \ + \ \mathbf{q}_n)^* (\mathbf{q}_{n+1} \ - \ \mathbf{q}_n) \ + \ (\mathbf{q}_{n+1} \ - \ \mathbf{q}_n)^* (\mathbf{q}_{n+1} \ + \ \mathbf{q}_n)\right] \tag{9.85}$$

Now, replace the $(\mathbf{q}_{n+1} \ - \ \mathbf{q}_n)$ terms by equation (1.81):

$$= \quad -\frac{\Delta z}{4}\left[\ (\mathbf{q}_{n+1} + \mathbf{q}_n)^* \mathbf{R}\,(\mathbf{q}_{n+1} + \mathbf{q}_n) + (\mathbf{q}_{n+1} + \mathbf{q}_n)^* \mathbf{R}^*\,(\mathbf{q}_{n+1} + \mathbf{q}_n)\right]$$

$$= \quad -\frac{\Delta z}{4}\left[(\mathbf{q}_{n+1} \ + \ \mathbf{q}_n)^* (\mathbf{R} \ + \ \mathbf{R}^*)\,(\mathbf{q}_{n+1} \ + \ \mathbf{q}_n)\right] \tag{9.86}$$

This equation establishes the result: If the matrix $\mathbf{R} \ + \ \mathbf{R}^*$ is positive definite, then $\mathbf{q}^*_{n+1}\mathbf{q}_{n+1}$ is less than $\mathbf{q}^*_n\mathbf{q}_n$.

9.4.3 Application to 45 degree wavefield extrapolation

The scalar wave equation for the extrapolation of a downgoing wavefield is

$$\frac{d\mathbf{q}}{dz} \quad = \quad i\,k_z\mathbf{q} \quad = \quad -\,\mathbf{R}\,\mathbf{q} \tag{9.87}$$

where the $\mathbf{R}$ operator takes the usual form

$$\mathbf{R} \quad = \quad -\,i\,k_z \quad = \quad \frac{-\,i\omega}{v}\sqrt{1 \ - \ \frac{v^2\,k_x^2}{\omega^2}} \tag{9.88}$$

Our plan is to approximate the square root by the usual continued fraction expansion and then identify $i\,k_x$ with ∂_x to obtain a space-domain equation. The main effort we must make stems from our refusal to make the usual assumption that $v(x,z)$ is independent of x. Since $\partial_x v q$ differs from $v\partial_x q$, the space representation does not seem to be unique, and we may wonder how the variable q relates to physical wave variables like pressure and displacement. Since (1.88) is purely imaginary, the depth-invariance of the quadratic $\mathbf{q}^*\,\mathbf{q}$ can be interpreted as the downward energy flux across the datum at depth z. Our main effort will be to assure that $\mathbf{q}^*\,\mathbf{q}$ does indeed remain depth-invariant when $v(x,z) \ \ != \ \ const.$ The task of determining the relation between the energy flux variable $\mathbf{q}$ and the physical variables will be left to the reader.

First $v^2k_x^2$ must be represented in the space domain. Thinking of the x-derivative operator $\partial/\partial x \ = \ \partial_x$ as a large bidiagonal matrix with $(1,-1)/\Delta x$ along the diagonal and $\mathbf{V}\,(x)$ as a diagonal matrix, we are attracted to expressions like $(\mathbf{V}\,\partial_x)^{\mathbf{T}}(\mathbf{V}\,\partial_x)$ or $(\mathbf{V}\,\partial_x)(\mathbf{V}\,\partial_x)^{\mathbf{T}}$ because they are symmetric, positive, semidefinite matrices. In simplest form, such numerical representations are tridiagonal matrices that can be abbreviated as either

$$\mathbf{T} \quad = \quad (\mathbf{V}\,\partial_x)\,(\mathbf{V}\,\partial_x)^{\mathbf{T}} \tag{9.89}$$

or

$$\mathbf{T} \quad = \quad (\mathbf{V}\,\partial_x)^{\mathbf{T}}\,(\mathbf{V}\,\partial_x) \tag{9.90}$$

At a later time accuracy or some other consideration could determine the choice in (1.89) and (1.90). Even other expressions could be used, provided they are real, symmetric, and positive definite.

In the previous section the constant velocity, 45° expansion of (1.88) was shown to be

$$\mathbf{R} \quad = \quad \frac{1}{v} \; \frac{v^2 k_x^2}{-\,i\omega\, 2 \;+\; \frac{v^2 k_x^2}{-\,i\omega\, 2}} \tag{9.91}$$

This scalar $\mathbf{R}$ always has a positive real part because $-\,i\omega$ is always represented in an impedance form, and the whole expression is built up satisfying Muir's rules for combining impedance functions. In going to the x-domain notice that $(ik_x)^2 \;=\; -\partial_{xx}$ and $(\partial_x)^{\mathbf{T}} \;=\; -\,\partial_x$. So the positive scalar $v^2 k_x^2$ corresponds to the positive eigenvalues of (1.89) and (1.90).

The expression of the bulletproof, square-root operator $\mathbf{R}$ in the space domain will now be given as

$$\mathbf{M} \quad = \quad \frac{\mathbf{T}}{-\,i\omega\, 2\,\mathbf{I} \;+\; \frac{\mathbf{T}}{-\,i\omega\, 2}} \tag{9.92}$$

$$\mathbf{R} \quad = \quad \mathbf{V}^{\,-\frac{1}{2}}\,\mathbf{M}\,\mathbf{V}^{\,-\frac{1}{2}} \tag{9.93}$$

Use of the division sign in (1.92) and (1.93) is justifiable because the matrix $\mathbf{T}$ commutes with the identity matrix $\mathbf{I}$. (A hazard in this work is that $\mathbf{T}$ does not commute with the diagonal matrix $\mathbf{V}$). The matrix $\mathbf{M}$ has the properties required of $\mathbf{R}$ since a basic matrix theorem says that the eigenvalues of a polynomial of a real symmetric matrix are the polynomials of the eigenvalues. In other words, replacing $\mathbf{T}$ in (1.92) and (1.93) by one of its eigenvalues produces a complex $\mathbf{M}$ whose real part is positive, so that $\mathbf{M}^* \;+\; \mathbf{M}$ is positive as required. What is needed is to show that the following matrix is positive definite:

$$\mathbf{R} \;+\; \mathbf{R}^* \quad = \quad \mathbf{V}^{\,-\frac{1}{2}}\,(\mathbf{M} \;+\; \mathbf{M}^*)\,\mathbf{V}^{\,-\frac{1}{2}} \tag{9.94}$$

A matrix $\mathbf{A}$ is positive definite if for arbitrary d, the scalar $d^*\,\mathbf{A}\,d$ is positive. The diagonal matrix $\mathbf{V}^{\,-\frac{1}{2}}$ can certainly be absorbed into d and d will still be arbitrary, so the proof is complete.

In programming it is a nuisance to put $\mathbf{V}^{\,-\frac{1}{2}}$ on each side of the matrix $\mathbf{M}$. Actually you can put $\mathbf{V}^{\,-1}$ on either side. In general, some other quadratic form such as $\mathbf{q}^*\,\mathbf{U}\,\mathbf{q}$ where $\mathbf{U}$ is strictly positive definite will be decreasing if $\mathbf{R}^*\,\mathbf{U} \;+\; \mathbf{U}\,\mathbf{R}$ is positive definite.