Fundamentals of Geophysical Data Processing
(first 9 of 11 original chapters)

Jon F. Clærbout
Cecil and Ida Green Professor of Geophysics
Stanford University

http://sepwww.stanford.edu/sep/prof
email: jon@sep.stanford.edu

Contents

Preface

Fundamentals of Geophysical Data Processing was first published in 1976 by McGraw Hill and later republished by Blackwell in 1985. I was disappointed that it went out of print again so I am now making it available free on the WWW. This electronic reincarnation lacks the final two chapters of the original book which I dropped because they are mostly better covered in my later books. A few errors are corrected in the electronic version, but I fear new errors were introduced in the retyping so if you have a copy of the original paper book, keep it!

This book contains basic and interesting material which because it does not play a big role in industrial practice was neglected in my later books. I enjoy teaching these basics from time to time and I use especially chapters 3,4,7,8,9. This is classic material about Toeplitz matrices and layered media. Material from chapters 1,2,5,6 was improved and included in my third book, Earth Soundings Analysis, Processing vrs. Inversion, now in print with Blackwells. This is material about Z-transforms and least squares fitting.

Many people have told me that this book, which was my first book, is the most creative of all my books. I agree. In defense of my more recent books, I would add that they benefitted greatly from the developement of word processors and from my teaching experience. So here you have it. I hope you enjoy it, and I hope you are getting it free.

Chapter 1

Transforms

The first step in data analysis is to learn how to represent and manipulate waveforms in a digital computer. Time and space are ordinarily regarded as continuous, but for purposes of computer analysis we must discretize them. This discretizing is also called digitizing or sampling. Discretizing continuous functions may at first be regarded as an evil that is necessary only because our data are not always known analytic functions. However, after gaining some experience with sampled functions, one realizes that many mathematical concepts are easier with sampled time than with continuous time. For example, in this chapter, the concept of the Z transform is introduced and is shown to be equivalent to the Fourier transform. The Z transform is readily understood on a basis of elementary algebra, whereas the Fourier transform requires substantial experience in calculus.

1.1 SAMPLED DATA AND Z TRANSFORMS

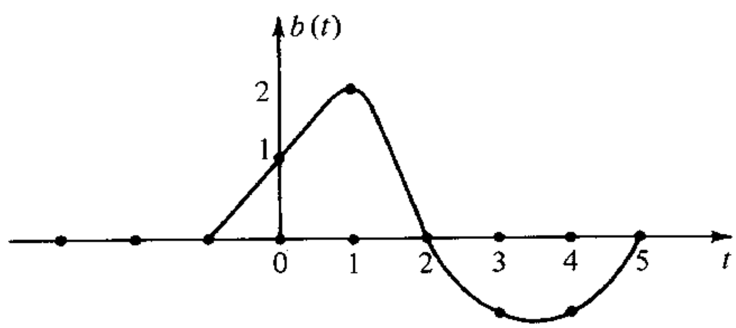


Figure 1.1: A continuous time function sampled at uniform time intervals. c1-1-1 [NR]

Consider the time function in Figure 1.1. To analyze such an observed time function in a computer it is necessary to approximate it in some way by a list of numbers. The usual way to do this is to evaluate or observe $b(t)$ at a uniform spacing of points in time. For this example, such a digital approximation to the continuous function could be denoted by the vector

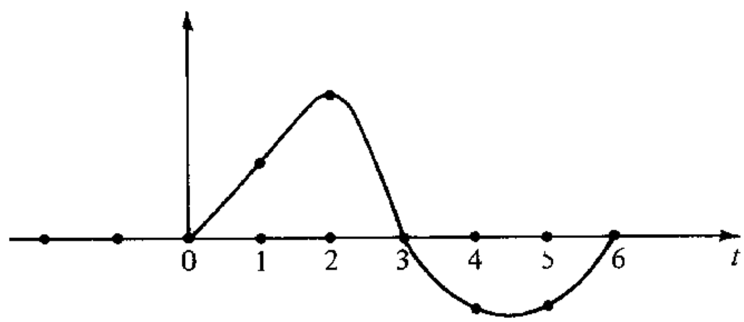
$$b_t \quad = \quad (\ldots\ 0,\ 0,\ 1,\ 2,\ 0,\ -1,\ -1,\ 0,\ 0,\ \ldots) \tag{1.1}$$

Of course if time points were taken more closely together we would have a more accurate approximation. Besides a vector, a function can be represented as a polynomial where the *coefficients* of the polynomial represent the values of $b(t)$ at successive time points. In this example, we have

$$B(Z) \quad = \quad 1 + 2Z + 0Z^2 - Z^3 - Z^4 \tag{1.2}$$

This polynomial is called a Z transform. What is the meaning of Z in this polynomial? The meaning is not that Z should take on some numerical value; the meaning of Z is that it is the unit delay operator. For example the coefficients of $ZB(Z) = Z + 2Z^2 - Z^4 - Z^5$ are plotted in Figure 1.2. It is the same waveform as in Figure 1.1, but it has been delayed.

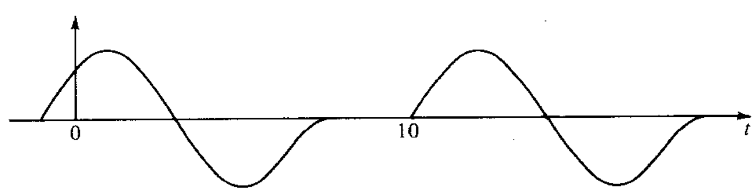
Figure 1.2: Coefficients of $ZB(Z)$ are shifted version of the coefficients of $B(Z)$ $\boxed{\text{c1-1-2}}$ [NR]



We see that the time function b_t is delayed n time units when $B(Z)$ is multiplied by Z^n. The delay operator Z is very important in analyzing waves simply because waves take a certain amount of time to get from place to place.

Another value of the delay operator is that it may be used to build up more complicated time functions from simpler ones. Suppose $b(t)$ represents the acoustic pressure function or the seismogram observed after a distant explosion. Then $b(t)$ is called the *impulse response*. If another explosion occurs at $t = 10$ time units after the first, we expect the pressure function $y(t)$ depicted in Figure 1.3.

Figure 1.3: Response to two explosions. $\boxed{\text{c1-1-3}}$ [NR]



In terms of Z transforms this would be expressed as $Y(Z) = B(Z) + Z^{10}B(Z)$. If the first explosion were followed by an implosion of half strength, we would have

$B(Z) - \frac{1}{2}Z^{10}B(Z)$. If pulses overlap one another in time [as would be the case if $B(Z)$ was of degree greater than 10], the waveforms would just add together in the region of overlap. The supposition that they just add together without any interaction is called the *linearity assumption*. This linearity assumption is very often true in practical cases. In seismology we find that—although the earth is a very heterogeneous conglomerations of rocks of different shapes and types—when seismic waves (of usual amplitude) travel through the earth, they do not interfere with one another. They satisfy linear superposition. The plague of nonlinearity arises from large amplitude disturbances. Nonlinearity does not arise from geometrical complications.

Now suppose there was an explosion at $t = 0$, a half-strength implosion at $t = 1$, and another, quarter-strength explosion at $t = 3$. This sequence of events determines a "source" time series, $x_t = (1, -\frac{1}{2}, 0, \frac{1}{4})$. The Z transform of the source is $X(Z) = 1 - \frac{1}{2}Z + \frac{1}{4}Z^3$. The observed y_t for this sequence of explosions and implosions through the seismometer has a Z transform $Y(Z)$ given by

$$
\begin{aligned}
Y(Z) &= B(Z) - \frac{Z}{2}B(Z) + \frac{Z^3}{4}B(Z) \\
&= \left(1 - \frac{Z}{2} + \frac{Z^3}{4}\right) B(Z) \\
&= X(Z)\,B(Z) \qquad (1.3)
\end{aligned}
$$

The last equation illustrates the underlying basis of linear-system theory that the output $Y(Z)$ can be expressed as the input $X(Z)$ times the impulse response $B(Z)$.

There are many examples of linear systems. A wide class of electronic circuits is comprised of linear systems. Complicated linear systems are formed by taking the output of one system and plugging it into the input of another. Suppose we have two linear systems characterized by $B(Z)$ and $C(Z)$, respectively. Then the question arises whether the two combined systems of Figure 1.4 are equivalent. The use of Z transforms makes it obvious that these two systems are equivalent since products of polynomials commute, i.e.,

$$
\begin{aligned}
Y_1(Z) &= [X(Z)B(Z)]\,C(Z) = XBC \\
Y_2(Z) &= [X(Z)C(Z)]\,B(Z) = XCB = XBC \qquad (1.4)
\end{aligned}
$$

Consider a system with an impulse response $B(Z) = 2 - Z - Z^2$. This polynomial can be factored into $2 - Z - Z^2 = (2+Z)(1-Z)$, and so we have the three equivalent systems in Figure 1.5. Since any polynomial can be factored, any impulse response can be simulated by a cascade of two-term filters (impulse responses whose Z transforms are linear in Z).

What do we actually do in a computer when we multiply two Z transforms together? The filter $2 + Z$ would be represented in a computer by the storage in memory of the coefficients $(2, 1)$. Likewise, for $1 - Z$ the numbers $(1, -1)$ are stored.

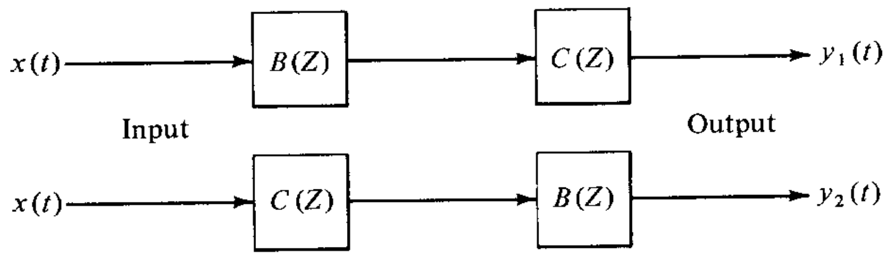


Figure 1.4: Two equivalent filtering systems. c1-1-4 [NR]

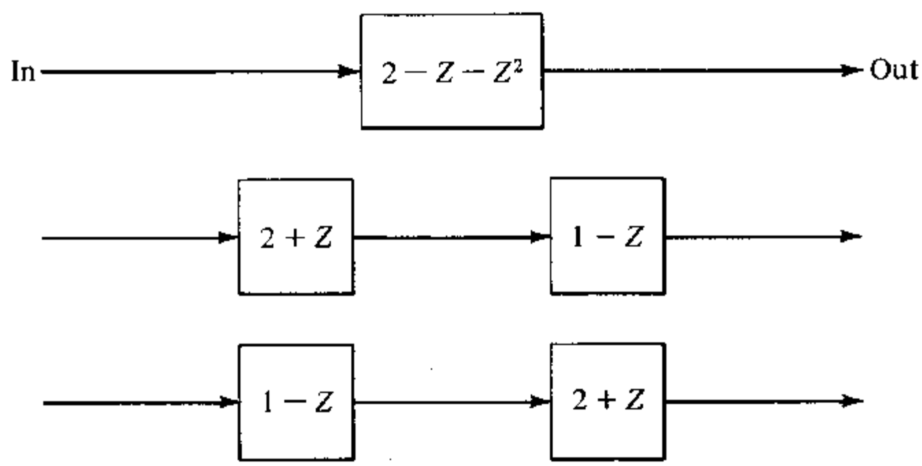


Figure 1.5: Three equivalent filtering systems. c1-1-5 [NR]

The polynomial multiplication program should take these inputs and produce the sequence $(2, -1, -1)$. Let us see how the computation proceeds in a general case, say

$$X(Z)\,B(Z) \quad = \quad Y(Z) \tag{1.5}$$

$$(x_0 + x_1 Z + x_2 Z + \cdots)\,(b_0 + b_1 Z + b_2 Z) \quad = \quad (y_0 + y_1 Z + y_2 Z^2 + \cdots) \tag{1.6}$$

Identifying coefficients of successive powers of Z, we get

$$\begin{aligned} y_0 &= x_0 b_0 \\ y_1 &= x_1 b_0 + x_0 b_1 \\ y_2 &= x_2 b_0 + x_1 b_1 + x_0 b_2 \\ y_3 &= x_3 b_0 + x_2 b_1 + x_1 b_2 \\ y_4 &= x_4 b_0 + x_3 b_1 + x_2 b_2 \end{aligned} \tag{1.7}$$

$$y_k = \sum_{i=0}^{2} x_{k-i} b_i \tag{1.8}$$

Equation (1.8) is called a *convolution equation.* Thus, we may say that the product of two polynomials is another polynomial whose coefficients are found by convolution. A simple Fortran computer program which does convolution, including end effects on both ends, is this:

```
DIMENSION X(LX), B(LB), Y(LY)
        DO 10 IY=1,LY
10      Y(IY) = 0.
        DO 20 IX=1,NX
        DO 20 IB=1,NB
        IY = IX+IB-1
20      Y(IY) = Y(IY) + X(IX)*B(IB)
```

The reader should notice that $X(Z)$ and $Y(Z)$ need not strictly be polynomials; they may contain both positive and negative powers of Z; that is

$$X(Z) \quad = \quad \cdots \frac{x_{-2}}{Z^2} + \frac{x_{-1}}{Z} + x_0 + x_1 Z + \cdots \tag{1.9}$$

$$Y(Z) \quad = \quad \cdots \frac{y_{-2}}{Z^2} + \frac{y_{-1}}{Z} + y_0 + y_1 Z + \cdots \tag{1.10}$$

The effect of using negative powers of Z in $X(Z)$ and $Y(Z)$ is merely to indicate that data are defined before $t = 0$. The effect of using negative powers of Z in the filter is quite different. Inspection of (1.8) shows that the output y_k which occurs at time k is a linear combination of current and previous inputs; that is, $(x_i,\ i \leq k)$. If the filter $B(Z)$ had included a term like b_{-1}/Z, then the output y_k at time k would be a linear combination of current and previous inputs and x_{k+1}, an input which really has not arrived at time k. Such a filter is called a *nonrealizable* filter because it could not operate in the real world where nothing can respond now to an excitation which has not yet occurred. However, nonrealizable filters are occasionally useful in computer simulations where all of the data are prerecorded.

EXERCISES:

1 Let $B(Z) = 1 + Z + Z^2 + Z^3 + Z^4$. Graph the coefficients of $B(Z)$ as a function of the powers of Z. Graph the coefficients of $[B(Z)]^2$.

2 If $x_t = \cos\omega_0 t$, where t takes on integral values, $b_t = (b_0, b_1)$, and $Y(Z) = X(Z)B(Z)$, what are A and B in $y_t = A\cos\omega_0 t + B\sin\omega_0 t$?

3 Deduce that, if $x_t = \cos\omega_0 t$ and $b_t = (b_0,\, b_1,\, \ldots,\, b_n)$, then y_t always takes the form $A\cos\omega_0 t + B\sin\omega_0 t$.

1.2 Z-TRANSFORM TO FOURIER TRANSFORM

We have defined the Z transform as

$$B(Z) = \sum_t b_t Z^t \tag{1.11}$$

If we make the substitution $Z = e^{i\omega}$ we have a "Fourier sum"

$$B(Z) \quad = \quad B(e^{i\omega}) \quad = \quad \sum_t b_t e^{i\omega t} \tag{1.12}$$

This is like a Fourier integral, and we could do a limiting operation to make it into an integral. Another point of view is that the Fourier integral

$$B(\omega) \quad = \quad \int_{-\infty}^{+\infty} b(t)\, e^{i\omega t}\, dt \tag{1.13}$$

reduces to the sum (1.12) when $b(t)$ is not a continuous function of time but is defined as

$$b(t) \quad = \quad \sum_k b_k\, \delta(t-k) \tag{1.14}$$

where δ is an impulse function.

In the last section we saw that to multiply two polynomials the coefficients must be convolved. The same process in Fourier transform language is that a product in the frequency domain corresponds to a convolution in the time domain.

Although one thinks of a Fourier transform as an integral which may be difficult or impossible to do, the Z transform is always easy, in fact trivial. To do a Z transform one merely attaches powers of Z to successive data points. When one has $B(Z)$ one can refer to it either as a time function or a frequency function, depending on whether one graphs the polynomial coefficients or if one evaluates and graphs $B(Z = e^{i\omega})$ for various frequencies ω. Notice that as ω goes from zero to 2π, $Z = e^{i\omega} = \cos\omega + i\sin\omega$ migrates once around the unit circle in the counterclockwise direction.

If taking a Z transform amounts to attaching powers of Z to successive points of a time function, then the inverse Z transform must be merely identifying coefficients of various powers of Z with different points in time. How can this simple "identification of coefficients" be the same as the apparently more complicated operation of inverse Fourier integrals? The inverse Fourier integral is

$$b(t) \quad = \quad \frac{1}{2\pi}\int_{-\infty}^{+\infty} B(\omega)\, e^{-i\omega t}\, d\omega \tag{1.15}$$

First notice that the integration of Z^n about the unit circle or $e^{in\omega}$ over $-\pi \leq \omega < +\pi$ gives zero unless $n = 0$ because cosine and sine are oscillatory; that is,

$$\begin{aligned} \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{in\omega}\, d\omega \quad &= \quad \frac{1}{2\pi}\int_{-\pi}^{\pi} (\cos n\omega + i\sin n\omega)\, d\omega \\ &= \quad \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{if } n = \text{non-zero integer} \end{cases} \end{aligned} \tag{1.16}$$

In terms of our discretized time functions, the inverse Fourier integral (1.15) is

$$b_t \quad = \quad \frac{1}{2\pi}\int_{-\pi}^{\pi} (\cdots + b_{-1}e^{-i\omega} + b_0 + b_1 e^{+i\omega} + \cdots)\, e^{-i\omega t}\, d\omega \tag{1.17}$$

Of all the terms in the integrand (1.17) we see by (1.16) that only the term with b_t will contribute to the integral; all the rest oscillate and cancel. In other words, it is only the coefficient of Z to the zero power which contributes to the integral, reducing (1.17) to

$$b_t \quad = \quad \frac{1}{2\pi} b_t \int_{-\pi}^{+\pi} d\omega \quad = \quad b_t \tag{1.18}$$

This shows how inverse Fourier transformation is just like identifying coefficients of powers of Z.

In this book and many others, it is common to assume that the time span between data samples $\Delta t = 1$ is unity. To adapt given equations to other values of Δt, one only need replace ω by $\omega\Delta t$; that is,

$$\omega_{\text{book}} \quad = \quad \omega_{\text{book}}\, \Delta t_{\text{book}} \quad = \quad \omega_{\text{true}}\, \Delta t_{\text{true}} \tag{1.19}$$

With Z transforms we have the spectrum given on a range of 2π for ω_{book}. In the limit Δt_{true} goes to zero, ω_{true} has the same infinite limits as the Fourier integral.

When a continuous function is approximated by a sampled function, it is necessary to take the sample spacing Δt_{true} small enough. The basic result of elementary texts is that, if there is no appreciable energy in a Fourier transform for frequencies higher than some frequency $\omega_{\max}$, then there is no appreciable loss of information if

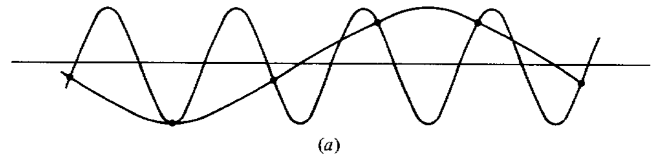


Figure 1.6: Subsampled sinusoid. c1-1-7a [NR]

the sampled at least two points per wavelength. Figure 1.6 shows how insufficient sampling of a sine wave often causes it to appear as a sine wave of lower frequency.

If a high-frequency sinusoid is sampled insufficiently often, it becomes indistinguishable from a lower-frequency sinusoid. For this reason $\omega_{\max} = \pi/\Delta t$ is said to be the folding frequency, as higher frequencies are folded down to look like lower frequencies.

In practice, quasi-sinusoidal waves are always sampled more frequently than twice per wavelength.

Next we wish to examine odd/even symmetries to see how they are affected in Fourier transformation. The even part e_t of a time function b_t is defined as

$$e_t \quad = \quad \frac{b_t - b_{-t}}{2} \tag{1.20}$$

The odd part is

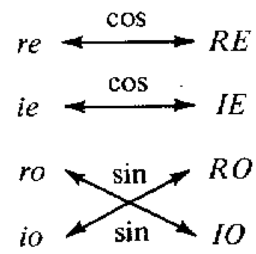
$$o_t \quad = \quad \frac{b_t - b_{-t}}{2} \tag{1.21}$$

A function is the sum of its even and odd parts. By adding (1-2-10) and (1-2-11), we get

$$b_t \quad = \quad e_t + o_t \tag{1.22}$$

Consider a simple, real, even time function such as $(b_{-1}, b_0, b_1) = (1, 0, 1)$. Its transform $Z + 1/Z = 2\cos\omega$ is an even function of ω since $\cos\omega = \cos(-\omega)$. Consider the real, odd time function $(b_{-1}, b_0, b_1) = (-1, 0, 1)$. Its transform $Z - 1/Z = 2(\sin\omega)/i$ is imaginary and odd, since $\sin\omega = -\sin(-\omega)$. Likewise, the transform of the imaginary even function $(i, 0, i)$ is the imaginary even function $i\cos\omega$ and the transform of the imaginary odd function $(-i, 0, i)$ is real and odd. Let r and i refer to real and imaginary, e and o refer to even and odd, and lower-case and upper-case refer to time and frequency functions. A summary of the symmetries of Fourier transformation is shown in Figure 1.7.

More elaborate time functions can be made up by adding together the two point functions we have considered. Since sums of even functions are even, and so on, the table of Figure 1.7 applies to all time functions. Note that an arbitrary time function takes the form $b_t = (\mathrm{re} + \mathrm{ro})_t + i(\mathrm{ie} + \mathrm{io})_t$. On transformation of b_t, each of the four individual parts transforms according to the table



(b)

Figure 1.7: Mnemonic table illustrating how even/odd and real/imaginary properties are affected by Fourier transformation. $\boxed{\text{c1-1-7b}}$ [NR]

EXERCISES:

1 Normally a function is specified entirely in the time domain or entirely in the frequency domain. When one is known, the other is determined by transformation. Now let us give half the information in the time domain by specifying that $b_t = 0$ for $t < 0$, and half in the frequency domain by giving the real part RE + RO in the frequency domain. How can you determine the rest of the function?

1.3 THE FAST FOURIER TRANSFORM

When we write the expression

$$B(Z) \quad = \quad b_0 + b_1 Z + b_2 Z^2 + \cdots \tag{1.23}$$

we have both a time function and its Fourier transform. If we plot the coefficients $(b_0, b_1, \ldots)$, we plot the time function. If we evaluate and plot (1.23) at numerous real ω, then we have plotted the transform. (Note that for real ω, Z is of unit magnitude; i.e., on the unit circle.) Since ω is a continuous variable and everything in a computer is finite, how do we select a finite number of values ω_k for plotting? The usual choice is to take evenly spaced frequencies. The lowest frequency can be zero. [Note $Z(\omega = 0) = e^{io} = 1$.] A frequency as high as $\omega = 2\pi$ [Note $Z(\omega = 2\pi) = e^{i2\pi} = 1$ also] need not be considered, since (1.23) gives the same value for it as for zero frequency. Choosing uniformly spaced frequencies between these limits we have

$$\omega_k \quad = \quad \frac{(0, 1, 2, \ldots, M-1)\, 2\pi}{M} \tag{1.24}$$

where M is some integer. Now let us abbreviate $B(Z(\omega_k))$ as B_k.

For the special case of an N-point time function where $N = 4$, (1.23) may be expressed by the matrix multiplication

$$\begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ B_3 \end{bmatrix} \quad = \quad \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & W & W^2 & W^3 \\ 1 & W^2 & W^4 & W^6 \\ 1 & W^3 & W^6 & W^9 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} \tag{1.25}$$

where

$$W \quad = \quad e^{2\pi i/N} \tag{1.26}$$

It is not essential to choose $N = M$ as we have done in (1.25), but it is a convenience. There is no loss of generality because one may always append zeros to a time function before inserting it into (1.25). A convenience of the choice $N = M$ is that the matrix in (1.25) will then be square and there will be an exact inverse. In fact, the inverse to (1.25) may be easily shown to be

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad = \quad 1/N \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1/W & 1/W^2 & 1/W^3 \\ 1 & 1/W^2 & 1/W^4 & 1/W^6 \\ 1 & 1/W^3 & 1/W^6 & 1/W^9 \end{bmatrix} \begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ B_3 \end{bmatrix} \tag{1.27}$$

Since $1/W$ is the complex conjugate of W, the matrices of (1.25) and (1.27) are just complex conjugates of one another. In fact, one observes no fundamental mathematical difference between time functions and frequency functions. This "duality" would be even more complete if we had used a scale factor of $N^{-1/2}$ in each of (1.25) and (1.27) rather than 1 in (1.25) and N^{-1} in (1.27). Note also that time functions and frequency functions could be interchanged in the mnemonic table describing symmetries. In fact, our earlier observation that the product of two frequency functions amounts to a convolution of two time functions corresponds to the convolution of the corresponding two frequency functions. We will not "provide" this duality as it is standard fare in both mathematics and systems theory books. However we will occasionally call upon the reader to realize that in any theorem the meanings of "time" and "frequency" may be interchanged.

In making a plot of the transform B_k for $(k = 0, 1, \ldots, M-1)$, the frequency axis ranges as $0 \le \omega_k < 2\pi$. It is often more natural to display the interval $-\pi \le \omega < \pi$. Since the transform is periodic with period 2π, values of B_k on the interval $\pi \le \omega < 2\pi$ may simply be moved to the interval $-\pi \le \omega < 0$ for display.

Thus, for $N = 8$ one might plot successively

$$B_4 \;\; B_5 \;\; B_6 \;\; B_7 \;\; B_0 \;\; B_1 \;\; B_2 \;\; B_3 \tag{1.28}$$

corresponding to values of ω equal to

$$-\pi, -\frac{3\pi}{4}, -\frac{\pi}{2}, -\frac{\pi}{4}, 0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4} \tag{1.29}$$

One advantage of this display interval is that for continuous time series which are sampled sufficiently densely in time the transform values B_k get small on both ends. If the time series is real, the real part of B_k has even symmetry about B_0; the imaginary part has odd symmetry about B_0. Then, one need not bother to display half the values. Choice of an odd value of N would enable us to put $\omega = 0$ exactly in the middle of the interval, but the reader will soon see why we stick to an even number of data points.

The matrix times vector operation in (1.25) requires N^2 multiplications and additions. The rest of this section describes a trick method, called the fast Fourier transform, of accomplishing the matrix multiplication in $N \log_2 N$ multiplications and additions. Since, for example, $\log_2 1024$ is 10, this is a tremendous saving in effort.

A basic building block in the fast Fourier transform is called doubling. Given a series $(x_0, x_1, \ldots, x_{N-1})$ and its sampled Fourier transform $(X_0, X_1, \ldots, X_{N-1})$ and another series $(y_0, y_1, \ldots, y_{N-1})$, one finds the transform of the interlaced double-length series

$$z_t \quad = \quad (x_0, y_0, x_1, y_1, \ldots, x_{N-1}, y_{N-1}) \tag{1.30}$$

The process of doubling is used many times during the process of computing a fast Fourier transform. As the word *doubling* might suggest, it will be convenient to suppose that N is an integer formed by raising 2 to some integer power. Suppose $N = 8 = 2^3$. We begin by dividing our eight-point series of one point each. The Fourier transform of each of the one-point series is just the point. Next, we use doubling four times to get the transforms of the four different two point series (x_0, x_4), (x_1, x_5), (x_2, x_6), and (x_3, x_7). We use doubling twice more to get the transforms of the two different four point series (x_0, x_2, x_4, x_6) and (x_1, x_3, x_5, x_7). Finally, we use doubling once more to get the transform of the original eight-point series $(x_0, x_1, x_2, \ldots, x_7)$.

It remains to look into the details of the doubling process.

Let

$$V \quad = \quad e^{i2\pi/2N} = W^{1/2} \tag{1.31}$$

$$V^N \quad = \quad e^{i\pi} = -1 \tag{1.32}$$

The transforms of two N-point series are by definition

$$X_k \quad = \quad \sum_{j=0}^{N-1} x_j V^{2jk} \quad (k = 0, 1, \ldots, N-1) \tag{1.33}$$

$$Y_k \quad = \quad \sum_{j=0}^{N-1} y_j V^{2jk} \quad (k = 0, 1, \ldots, N-1) \tag{1.34}$$

The transform of the interlaced series $z_j = (x_0, y_0, x_1, y_1, \ldots, x_{N-1}, y_{N-1})$ is by definition

$$Z_k = \sum_{l=0}^{2N-1} z_l V^{lk} \quad (k = 0, 1, \ldots, 2N-1) \tag{1.35}$$

To make Z_k from X_k and Y_k we require two separate formulas: one for $k = 0$, $1, \ldots, N-1$, and the other for $k = N, N+1, \ldots, 2N-1$.

First

$$Z_k = \sum_{l=0}^{2N-1} z_l V^{lk} \quad (k = 0, 1, \ldots, N-1)$$

We split the sum into two parts, noting that x_j multiplies even powers of V and y_j multiplies odd powers.

$$
\begin{aligned}
Z_k &= \sum_{j=0}^{N-1} x_j V^{2jk} + V^k \sum_{j=0}^{N=1} y_j V^{2jk} && (1.36)\\
&= X_k + V^k Y_k && (1.37)
\end{aligned}
$$

We obtain the last half of the Z_k by

$$
\begin{aligned}
Z_k &= \sum_{l=0}^{2N-1} z_l V^{lk} \quad (k = N, N+1, \ldots, 2N-1) && (1.38)\\
&= \sum_{l=0}^{2N-1} z_l V^{l(m+N)} \quad (k - N = m = 0, 1, \ldots, N-1) && (1.39)\\
&= \sum_{l=0}^{2N-1} z_l V^{lm} (V^N)^l && (1.40)\\
&= \sum_{l=0}^{2N-1} z_l V^{lm} (-1)^l
\end{aligned}
$$

1.4 Phase delay and group delay

This material was revised and included in my third book, ESA:PVI.

1.5 Correlation and spectra

This material was revised and included in my third book, ESA:PVI.

1.6 Hilbert transform

This material was revised and included in my third book, ESA:PVI.

```
      SUBROUTINE FORK(LX,CX,SIGNI)
C FAST FOURIER                                           2/15/69
C                             LX
C  CX(K) = SQRT(1/LX) SUM (CX(J)*EXP(2*PI*SIGNI*I*(J-1)*(K-1)/LX))
C                             J=1                FOR K=1,2,...,(LX=2**INTEGER)
      COMPLEX CX(LX),CARG,CEXP,CW,CTEMP
      J=1
      SC=SQRT(1./LX)
      DO 30 I=1,LX
      IF(I.GT.J) GO TO 10
      CTEMP=CX(J)*SC
      CX(J)=CX(I)*SC
      CX(I)=CTEMP
 10   M=LX/2
 20   IF(J.LE.M) GO TO 30
      J=J-M
      M=M/2
      IF(M.GE.1) GO TO 20
 30   J=J+M
      L=1
 40   ISTEP=2*L
      DO 50 M=1,L
      CARG=(0.,1.)*(3.14159265*SIGNI*(M-1))/L
      CW=CEXP(CARG)
      DO 50 I=M,LX,ISTEP
      CTEMP=CW*CX(I+L)
      CX(I+L)=CX(I)-CTEMP
 50   CX(I)=CX(I)+CTEMP
      L=ISTEP
      IF(L.LT.LX) GO TO 40
      RETURN
      END
```

Figure 1.8: A program to do fast Fourier transform. Modified from Brenner. Calling this program twice returns the original data. SIGNI should be +1. on one call and −1. on the other. LX must be a power of 2. c1-1-8 [NR]

Chapter 2

One-sided functions

All physical systems share the property that they do not respond before they are excited. Thus the impulse response of any physical system is a one-sided time function (it vanishes before $t = 0$). In system theory such a filter function is called realizable. In wave propagation this property is associated with causality in that no wave may begin to arrive before it is transmitted. The lag-time point $t = 0$ plays a peculiar and an important role. For this reason, many subtle matters will be much more clearly understood with sampled time than with continuous time. When a filter responds at and after lag time $t = 0$, we will say the filter is realizable or causal. The word *causal* is appropriate in physics where stress may cause (practically) instantaneous strain and vise versa, but one should revert to the more precise words *realizable* or *one-sided* when using filter theory to describe economic or social systems where simultaneity is quite different from cause and effect.

2.1 INVERSE FILTERS

To understand causal filters better, we now take up the task of undoing what a causal filter has done. Consider the output y_t of a filter b_t is known but the input x_t is unknown. See Figure 2.1.

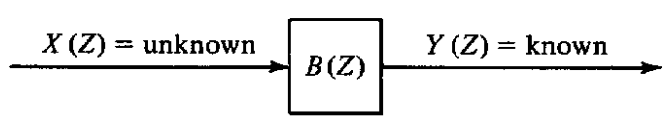


Figure 2.1: Sometimes the input to a filter is unknown. c2-2-1 [NR]

This is the problem that one always has with a transducer/recorder system. For example, the output of a seismometer is a wiggly line on a piece of paper from which the seismologist may wish to determine the displacement, velocity, or acceleration of the ground. To undo the filtering operation of the filter $B(Z)$, we will try to find another filter $A(Z)$ as indicated in Figure 2.2.

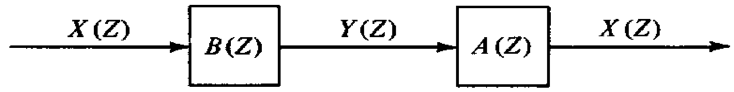


Figure 2.2: The filter $A(Z)$ is inverse to the filter $B(Z)$. c2-2-2 [NR]

To solve for the coefficients of the filter $A(Z)$, we merely identify coefficients of powers of Z in $B(Z)A(Z) = 1$. For $B(Z)$, a three-term filter, this is

$$(a_0 + a_1 Z + a_2 Z^2 + a_3 Z^3 + \cdots)\,(b_0 + b_1 Z + b_2 Z^2) \quad = \quad 1 \tag{2.1}$$

The coefficients of $Z^0, Z^1, Z^2, \cdots$ in (2.1) are

$$\begin{aligned}
a_0 b_0 &= 1 && (2.2)\\
a_1 b_0 + a_0 b_1 &= 0 && (2.3)\\
a_2 b_0 + a_1 b_1 + a_0 b_2 &= 0 && (2.4)\\
a_3 b_0 + a_2 b_1 + a_1 b_2 &= 0 && (2.5)\\
a_4 b_0 + a_3 b_1 + a_2 b_2 &= 0 && (2.6)\\
\cdots\cdots\cdots\cdots\cdots\cdots & && \\
a_k b_0 + a_{k-1} b_1 + a_{k-2} b_2 &= 0 && (2.7)
\end{aligned}$$

From (2.2) one may get a_0 from b_0. From (2.3) one may get a_1 from a_0 and the b_k. From (2.4) one may get a_2 from a_1, a_0, and the b_k. Likewise, in the general case a_k may be found from a_{k-1}, a_{k-2}, and the b_k. Specifically, from (2.7) the a_k may be determined recursively by

$$a_k \quad = \quad \frac{-\sum_{i=1}^{2} a_{k-i} b_i}{b_0} \tag{2.8}$$

Consider the example where $B(Z) = 1 - Z/2$; then, by equations like (2.2) to (2.7), by the binomial theorem, by polynomial division, or by Taylor's power series formula we obtain

$$A(Z) \quad = \quad \frac{1}{1 - Z/2} \quad = \quad 1 + \frac{Z}{2} + \frac{Z^2}{4} + \frac{Z^3}{8} + \cdots \tag{2.9}$$

We see that there are an infinite number of filter coefficients but that they drop off rapidly in size so that approximation in a computer presents no problem. The situation is not so rosy with the filter $B(Z) = 1 - 2Z$. Here we obtain

$$A(Z) \quad = \quad \frac{1}{1 - 2Z} \quad = \quad 1 + 2Z + 4Z^2 + 8Z^3 + 16Z^4 + 32Z^5 + \cdots \tag{2.10}$$

The coefficients of the series increase without bound. The outputs of the filter $A(Z)$ depend infinitely strongly on inputs of the infinitely distant past. [Recall that the present output of $A(Z)$ is a_0 times the present input x_1 plus a_1 times the previous input x_{t-1}, etc., so a_n represents memory on n time units earlier.] The implication of this is that some filters $B(Z)$ will not have useful finite approximate inverses $A(Z)$ determined from (2.2) to (2.8). We now seek ways to identify the good filters from the bad ones. With a two-pulse filter, the criterion is merely that the first pulse in $B(Z)$ be larger than the second. A more mathematical description of the state of affairs results from solving for the roots of $B(Z)$, that is, find the values of Z_0 for which $B(Z_0) = 0$. For example $1 - Z/2$ we find $Z_0 = 2$. For the example $1 - 2Z$, we find $Z_0 = \frac{1}{2}$. The general case for wavelets with complex coefficients is that, if the solution value Z_0 of $B(Z_0) = 0$ lies inside the unit circle in the complex plane, then $1/B(Z)$ will have coefficients which blow up; and if the root lies outside the unit circle, then the inverse $1/B(Z)$ will be bounded.

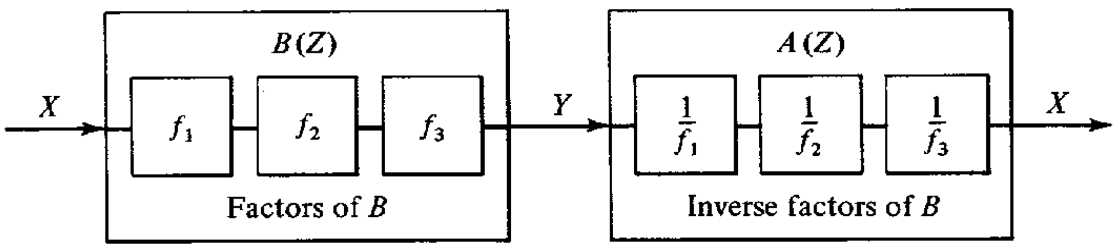


Figure 2.3: Factoring the polynomial $B(Z)$ breaks the filter into many two-term filters. Each one should have a bounded inverse. c2-2-3 [NR]

Recalling earlier discussion that a polynomial $B(Z)$ of degree N may be factored into N subsystems and that the ordering of subsystems is unimportant (see Figure 2.3), we suspect that if any of the N roots of $B(Z)$ lies inside the unit circle we may have difficulty with $A(Z)$. Actual proof of this suspicion relies on a theorem from complex-variable theory about absolutely convergent series. The theorem is that the product of absolutely convergent series is convergent, and conversely the product of any convergent series with a divergent series is divergent. Another proof may be based upon the fact that a power series for $1/B(Z)$ converges in a circle about the origin with a radius from the origin out to he first pole [the zero of $B(Z)$ of smallest magnitude]. Convergence of $A(Z)$ on the unit circle means, in terms of filters, that the coefficients of $A(Z)$ are decreasing. Thus, if all the zeros of $B(Z)$ are outside the unit circle, we will get a convergent filter from (2.8).

Can anything at all be done if there is one root or more inside the circle? An answer is suggested by the example

$$\frac{1}{1-2Z} \quad = \quad \frac{1}{2Z}\frac{1}{1-1/2Z} \quad = \quad -\frac{1}{2Z}\left[1+\frac{1}{2Z}+\frac{1}{(2Z)^2}+\cdots\right] \tag{2.11}$$

Equation (2.11) is a series expansion in $1/Z$, that is, a Taylor series about infinity. It converges from $Z = \infty$ all the way in to a circle of radius $1/2$. This means that the inverse converges on the unit circle where it must, if the coefficients are to be bounded. In terms of filters it means that the inverse filter must be one of those filters which responds to future inputs and hence is not physically realizable but may be used in computer simulation.

In the general case, then, one must factor $B(Z)$ into two parts: $B(Z) = B_{out}(Z)B_{in}(Z)$ where B_{out} contains roots outside the unit circle and B_{in} contains the roots inside. Then the inverse of B_{out} is expressed as a Taylor series about the origin and the inverse of B_{in} is expressed as a Taylor series about infinity. The final expression for $1/B(Z)$ is called a Laurent expansion for $1/B(Z)$, and it converges on a ring surrounding the unit circle. Cases with zeros exactly on the unit circle present special problems. Sometimes you can argue yourself out of the difficulty but at other times roots on or even near the circle may mean that a certain computing scheme won't work out well in practice.

Finally, let us consider a mechanical interpretation. The stress (pressure) in a material may be represented by x_t, and the strain (volume charge) may be represented by y_t. The following two statements are equivalent; that is, in some situations they are both true, and in other situations they are both false:

STATEMENT A The stress in a material may be expressed as a linear combination of present and past strains. Likewise, the strain may be deduced from present and past stresses.

STATEMENT B The filter which relates stress to strain and vice versa has all poles and zeros outside the unit circle.

EXERCISES:

1 Find the filter which is inverse to $(2 - 5Z + 2Z^2)$. You may just drop higher-order powers of Z, but an exact expression for the coefficients of any power of Z is preferred. (Partial fractions is a useful, though not necessary, technique.) Sketch the impulse response.

2 Show that multiplication by $(1 - Z)$ in discretized time is analogous to time differentiation in continuous time. Show that dividing by $(1 - Z)$ is analogous to integration. What are the limits on the integral?

3 Describe a general method for determining $A(Z)$ and $B(Z)$ from a Taylor series of $B(Z)/A(Z) = C_0 + C_1Z + C_2Z^2 + \cdots + C_\infty Z^\infty$ where $B(Z)$ and $A(Z)$ are polynomials of unknown degree n and m, respectively. Work out the case $C(Z) = \frac{1}{2} - \frac{3}{4}Z - \frac{3}{8}Z^2 - \frac{3}{16}Z^3 - \frac{3}{32}Z^4 - \cdots$. Don't try this problem unless you are quite familiar with determinants. [HINT: Identify coefficients of $B(Z) = A(Z)C(Z)$.]

2.2 MINIMUM PHASE

In Sec. 2-1 we learned that knowledge of convergence of the Taylor series of $1/B(Z)$ on $|Z| = 1$ is equivalent to knowledge that $B(Z)$ has no roots inside the unit circle.

Now we will see that these conditions are also equivalent to a certain behavior of the phase of $B(Z)$ on the unit circle.

Let us consider the phase shift of the two-term filter

$$\begin{aligned} B &= 1 - \frac{Z}{Z_0} \quad (Z_0 = \rho e^{i\omega_o}) \\ &= 1 - \rho^{-1} e^{i(\omega-\omega_o)} \\ &= 1 - \rho^{-1}\cos(\omega-\omega_o) - i\rho^{-1}\sin(\omega-\omega_o) \end{aligned}$$

By definition, phase is the arctangent of the ratio of the imaginary part to the real part.

A graph of phase as a function of frequency looks radically different for $\rho < 1$ than for $\rho > 1$. See Figure 2.4 for the case $\rho > 1$.

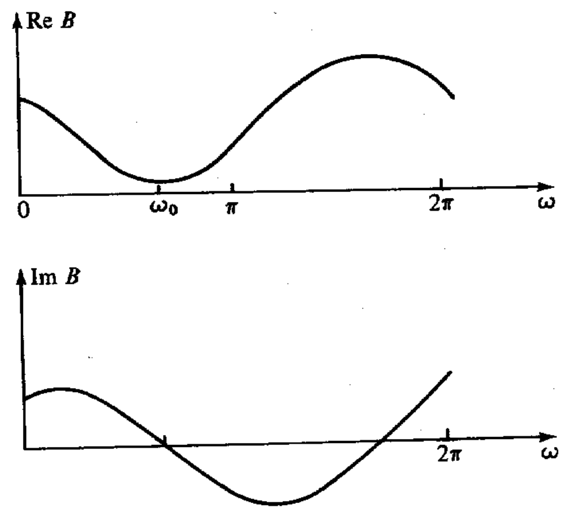


Figure 2.4: Real and imaginary parts of the Z transform $1 - Z/(1.25e^{i2\pi/3})$. $\boxed{\text{c2-2-4}}$ [NR]

The phase is the arctangent of ImB/ReB. The easiest way to keep track of the phase is in the complex B plane. This is shown in Figure 2.5.

Thus phase as a function of frequency is shown in Figure 2.6. Notice that the phase ϕ at $\omega = 0$ is the same as the phase at $\omega = 2\pi$. This follows because the real and imaginary parts are periodic with 2π. The situation will be different when there is a zero inside the unit circle; that is, $\rho < 1$. The real and imaginary parts are shown in Figure 2.7 and the complex plane in Figure 2.8.

The phase ϕ increases by 2π as ω goes from zero to 2π because the circular path surrounds the origin. The phase curve is shown in Figure 2.9. The case $\rho > 1$ where $\phi(\omega) = \phi(\omega + 2\pi)$ has come to be called *minimum phase* or *minimum delay.*

Now we are ready to consider a complicated filter like

$$B(Z) = \frac{(Z-c_1)(Z-c_2)\cdots}{(Z-a_1)(Z-a_2)\cdots} \tag{2.12}$$

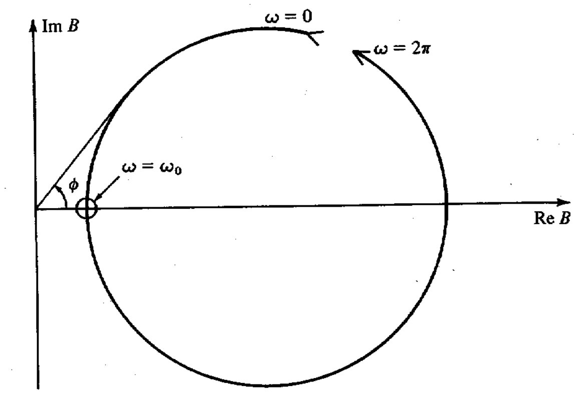


Figure 2.5: Phase of the two-term filter of Figure 2.4. c2-2-5 [NR]

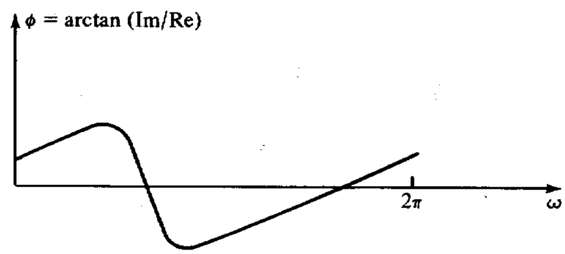


Figure 2.6: The phase of a two-term minimum-phase filter. c2-2-6 [NR]

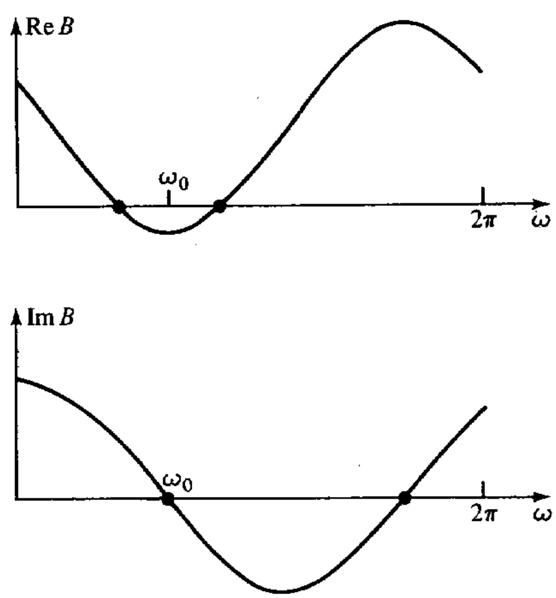


Figure 2.7: Real and imaginary parts of the two-term nonminimum-phase filter, $1 - 1.25Ze^{-i2\pi/3}$. c2-2-7 [NR]

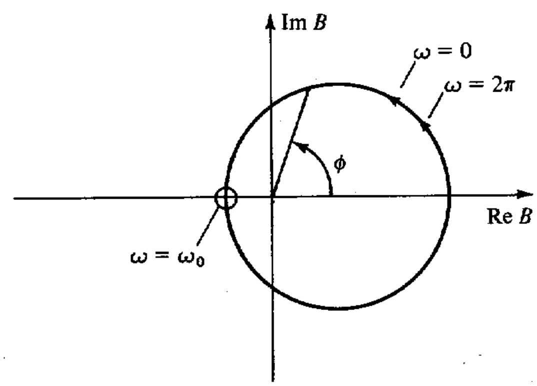


Figure 2.8: Phase in complex plane. c2-2-8 [NR]

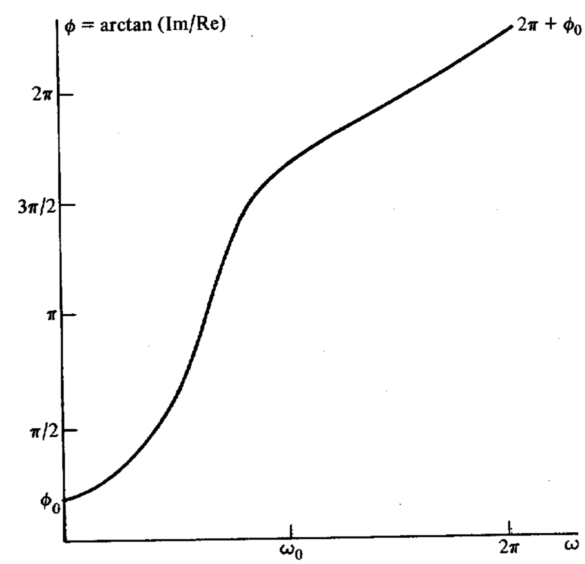


Figure 2.9: The phase of a two-term nonminimum-phase filter. c2-2-9 [NR]

By the rules of complex-number multiplication the phase of $B(Z)$ is the sum of the phases in the numerator minus the sum of the phases in the denominator. Since we are discussing realizable filters the denominator factors must all be minimum phase, and so the denominator phase curve is a sum of curves like Figure 2.6. The numerator factors may or may not be minimum phase. Thus the numerator phase curve is a sum of curves like either Figure 2.6 or Figure 2.9. If any factors at all are like Figure 2.9, then the total phase will resemble Figure 2.9 in that the phase at $\omega = 2\pi$ will be greater than the phase at $\omega = 0$. Then the filter will be nonminimum phase.

2.3 FILTERS IN PARALLEL

We have seen that in a cascade of filters the filter polynomials are multiplied together. One might conceive of adding two polynomials $A(Z)$ and $G(Z)$ when they correspond to filters which operate in parallel. See Figure 2.10.

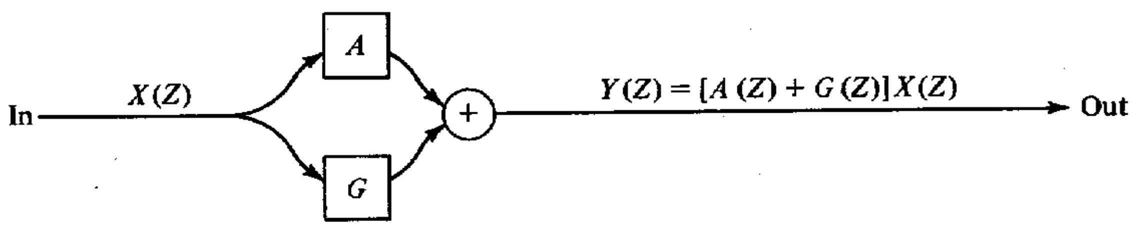


Figure 2.10: Filters operating in parallel. c2-2-10 [NR]

When filters operate in parallel their Z transforms add together. We have seen that a cascade of filters is minimum phase if, and only if, each element of the product is minimum phase. Now we will see a sufficient (but not necessary) condition that the sum $A(Z) + G(Z)$ be minimum phase. First of all, let us assume that $A(Z)$ is minimum phase. Then we may write

$$A(Z) + G(Z) \quad = \quad A(Z)\left[1 + \frac{G(Z)}{A(Z)}\right] \tag{2.13}$$

The question whether $A(Z) + G(Z)$ is minimum phase is now reduced to determining whether $A(Z)$ and $1 + G(Z)/A(Z)$ are both minimum phase. We have assumed that $A(Z)$ is minimum phase. Before we ask whether $1 + G(Z)/A(Z)$ is minimum phase we need to be sure that it is causal. Since $1/A(Z)$ is expandable in positive powers of Z only, then $G(Z)/A(Z)$ is also causal. We will next see that a sufficient condition for $1 + G(Z)/A(Z)$ to be minimum phase is that the spectrum of A exceeds that of G at all frequencies. In other words, for any real ω, $| A | > | G |$. Thus, if we plot the curve of $G(Z)/A(Z)$ in the complex plane, for real $0 \leq \omega \leq 2\pi$ it lies everywhere inside the unit circle. Now if we add unity —getting $1 + G(Z)/A(Z)$, the curve will always have a positive real part. See Figure 2.11.

Since the curve cannot enclose the origin, the phase must be that of a minimum-phase function. In words, "You can add garbage to a minimum-phase wavelet if you

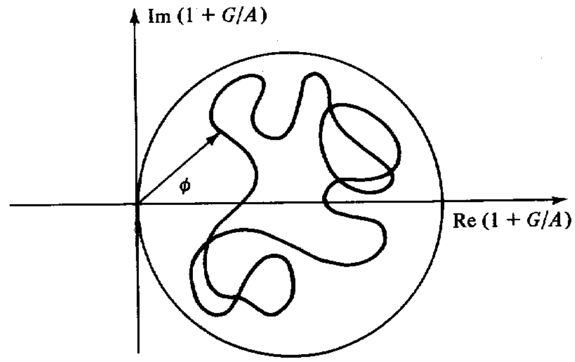


Figure 2.11: Phase of a positive real function lies between $\pm\pi/2$. c2-2-11 [NR]

do not add too much." This somewhat abstract theorem has an immediate physical consequence. Suppose a wave characterized by a minimum phase $A(Z)$ is emitted from a source and detected at a receiver some time later. At a still later time an echo bounces off a nearby object and is also detected at the receiver. The receiver sees the signal $Y(Z) = A(Z) + Z^n \alpha A(Z)$ where n measures the delay from the first arrival to the echo and α represents the amplitude attenuation of the echo. To see that $Y(Z)$ is minimum phase, we note that the magnitude of Z^n is unity and that the reflection coefficient α must be less than unity (to avoid perpetual motion) so that $Z^n \alpha A(Z)$ takes the role of $G(Z)$. Thus a minimum-phase wave along with its echo is minimum phase. We will later consider wave propagation situations with echoes of the echoes ad infinitum.

EXERCISES:

1 Find two nonminimum-phase wavelets whose sum is minimum phase.

2 Let $A(Z)$ be a minimum-phase polynomial of degree N. Let $A'(Z) = Z^N \bar{A}(1/Z)$. Locate in the complex Z plane the roots of $A'(Z)$. $A'(Z)$ is called *maximum phase.* [HINT: Work the simple case $A(Z) = a_0 + a_1 Z$ first.]

3 Suppose $A(Z)$ is maximum phase and that the degree of $G(Z)$ is less than or equal to the degree of $A(Z)$. Assume $\mid A \mid > \mid G \mid$. Show that $A(Z) + G(Z)$ is maximum phase.

4 Let $A(Z)$ be minimum phase. Where are the roots of $A(Z) + cZ^N \bar{A}(1/Z)$ in the three cases $\mid c \mid < 1, \mid c \mid > 1, \mid c \mid = 1$? (HINT: The roots of a polynomial are continuous functions of the polynomial coefficients.)

2.4 POSITIVE REAL FUNCTIONS

Two similar types of functions called *admittance functions* $Y(Z)$ and *impedance functions* $I(Z)$ occur in many physical problems. In electronics, they are ratios of current to voltage and of voltage to current; in acoustics, impedance is the ratio of pressure to velocity. When the appropriate electrical network or acoustical region contains no sources of energy, then these ratios have the positive real property. To see this in a mechanical example, we may imagine applying a known force $F(Z)$ and observing the resulting velocity $V(Z)$. In filter theory, it is like considering that $F(Z)$ is input to a filter $Y(Z)$ giving output $V(Z)$. We have

$$V(Z) \quad = \quad Y(Z)F(Z) \tag{2.14}$$

The filter $Y(Z)$ is obviously causal. Since we believe we can do it the other way around, that is, prescribe the velocity and observe the force, there must exist a convergent causal $I(Z)$ such that

$$F(Z) \quad = \quad I(Z)V(Z) \tag{2.15}$$

Since Y and I are inverses of one another and since they are both presumed bounded and causal, then they both must be minimum phase.

First, before we consider any physics, note that if the complex number $a+ib$ has a positive real part a, then the real part of $(a+ib)^{-1}$ namely $a/(a^2+b^2)$ is also positive. Taking $a+ib$ to represent a value of $Y(Z)$ or $I(Z)$ on the unit circle, we see the obvious fact that if either Y or I has the positive real property, then the other does, too.

Power dissipated is the product of force times velocity, that is

$$\text{Power} \quad = \quad \cdots + f_0 v_0 + f_1 v_1 + f_2 v_2 + \cdots \tag{2.16}$$

This may be expressed in terms of Z transforms as

$$\begin{aligned}
\text{Power} \quad &= \quad \frac{1}{2} \text{ coeff of } Z^0 \text{ of } V\left(\frac{1}{Z}\right) F(Z) + F\left(\frac{1}{Z}\right) V(Z) \\
&= \quad \frac{1}{2}\frac{1}{2\pi} \int_{-\pi}^{+\pi} \left[V\left(\frac{1}{Z}\right) F(Z) + F\left(\frac{1}{Z}\right) V(Z) \right] d\omega
\end{aligned} \tag{2.17}$$

Using (2.14) to eliminate $V(Z)$ we get

$$\text{Power} \quad = \quad \frac{1}{2}\frac{1}{2\pi} \int_{-\pi}^{+\pi} F\left(\frac{1}{Z}\right) \left[Y\left(\frac{1}{Z}\right) + Y(Z) \right] F(Z)\, d\omega \tag{2.18}$$

We note that $Y(Z)+Y(1/Z)$ looks superficially like a spectrum because the coefficient of Z^k equals that of Z^{-k}, which shows the symmetry of an autocorrelation function. Defining

$$R(Z) \quad = \quad Y(Z) + Y\left(\frac{1}{Z}\right) \tag{2.19}$$

(2.17) becomes

$$\text{Power} \quad = \quad \frac{1}{2}\frac{1}{2\pi} \int_{-\pi}^{+\pi} R(Z) F\left(\frac{1}{Z}\right) F(Z)\, d\omega \tag{2.20}$$

The integrand is the product of the arbitrary positive input force spectrum and $R(Z)$. If the power dissipation is expected to be positive at all frequencies (for all $\bar{F}F$), then obviously $R(Z)$ must be positive at all frequencies; thus R is indeed a spectrum. Since we have now discovered that $Y(Z)$ and $Y(1/Z)$ must be positive for all frequencies, we have discovered that $Y(Z)$ is not an arbitrary minimum-phase filter. The real part of both $Y(Z)$ and $Y(1/Z)$ is

$$\mathrm{Re}[Y(Z)] \quad = \quad \mathrm{Re}\left[Y\left(\frac{1}{Z}\right)\right] \quad = \quad y_o + y_1 \cos\,\omega + y_2 \cos\,2\omega + \cdots \tag{2.21}$$

Since the real part of the sum must be positive, then obviously the real part of each of the equal parts be positive.

Now if the material or mechanism being studied is passive (contains no energy sources) then we must have positive dissipation over a time gate from minus infinity up to any time t. Let us find an expression for dissipation in such a time gate. For simplicity take both the force and velocity vanishing before $t = 0$. Let the end of the time gate include the point $t = 2$ but not $t = 3$.

Define

$$f'_t \quad = \quad \begin{cases} f_t & t \leq 2 \\ 0 & t > 2 \end{cases} \tag{2.22}$$

To find the work done over all time we may integrate (2.20) over all frequencies. To find the work done in the selected gate we may replace F by F' and integrate over all frequencies, namely

$$W_2 \quad = \quad \frac{1}{2}\,\frac{1}{2\pi}\int_{-\pi}^{+\pi} F'\left(\frac{1}{Z}\right) R(Z) F'(Z)\, d\omega \tag{2.23}$$

As we have seen, this integral merely selects the coefficient of Z^0 of the integrand. Let us work this out. First, collect coefficients of powers of Z in $R(Z)F'(Z)$. We have

$$\begin{aligned} Z^0 &: \quad r_o f'_0 + r_{-1} f'_1 + r_{-2} f'_2 \\ Z^1 &: \quad r_1 f'_0 + r_0 f'_1 + r_{-1} f'_2 \\ Z^2 &: \quad r_2 f'_0 + r_1 f'_1 + r_0 f'_2 \end{aligned}$$

To obtain the coefficient of Z^0 in $F'(1/Z)[R(Z)F'(Z)]$ we must multiply the top row above by f'_0, the second row by f'_1 and the third row by f'_2. The result can be arranged in a very orderly fashion by

$$\begin{aligned} W_2 &= \frac{1}{2}[f_0\ f_1\ f_2] \begin{bmatrix} r_0 & r_{-1} & r_{-2} \\ r_1 & r_0 & r_{-1} \\ r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \end{bmatrix} \\ &= \frac{1}{2}[f_0\ f_1\ f_2] \begin{bmatrix} 2y_0 & y_1 & y_2 \\ y_1 & 2y_0 & y_1 \\ y_2 & y_1 & 2y_0 \end{bmatrix} \begin{bmatrix} f_0 \\ f_1 \\ f_2 \end{bmatrix} \end{aligned} \tag{2.24}$$

Not only must the 3 x 3 quadratic form (2.24) be positive (i.e., $W_2 \geq 0$ for arbitrary f_t) but all $t \times t$ similar quadratic forms W_t must be positive.

In conclusion, the positive real property in the frequency domain means that $Y(Z) + Y(1/Z)$ is positive for any real ω and the positive real property in the time domain means that all $t \times t$ matrices like that of (2.24) are positive definite. Figure 2.12 summarizes the function types which we have considered.

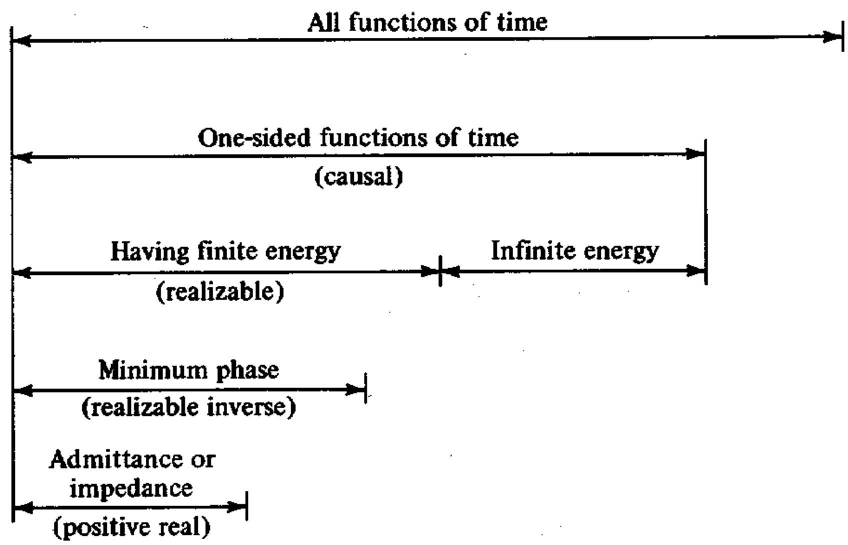


Figure 2.12: Important classes of time functions. c2-2-12 [NR]

EXERCISES:

1 In mechanics we have force and velocity of a free unit mass related by $dv/dt = f$ or $v = \int_{-\infty}^{t} f dt$. Compute the power dissipated as a function of frequency if integration is approximated by convolution with $(.5, 1., 1., 1., ...)$. [HINT: Expand $(1 + Z)/2(1 - Z)$ in positive powers of Z.]

2 Construct an example of a simple function which is minimum phase but not positive real.

2.5 NARROW-BAND FILTERS

Filters are often used to modify the spectrum of given data. With input $X(Z)$, filters $B(Z)$, and output $Y(Z)$ we have $Y(Z) = B(Z)X(Z)$ and the Fourier conjugate $\bar{Y}(1/Z) = \bar{B}(1/Z)\bar{X}(1/Z)$. Multiplying these two relations together we get

$$\overline{Y}Y = (\bar{B}B)(\bar{X}X) \tag{2.25}$$

which says that the spectrum of the input times the spectrum of the filter equals the spectrum of the output. Filters are often characterized by the shape of their spectra. Some examples are shown in Figure (2.13).

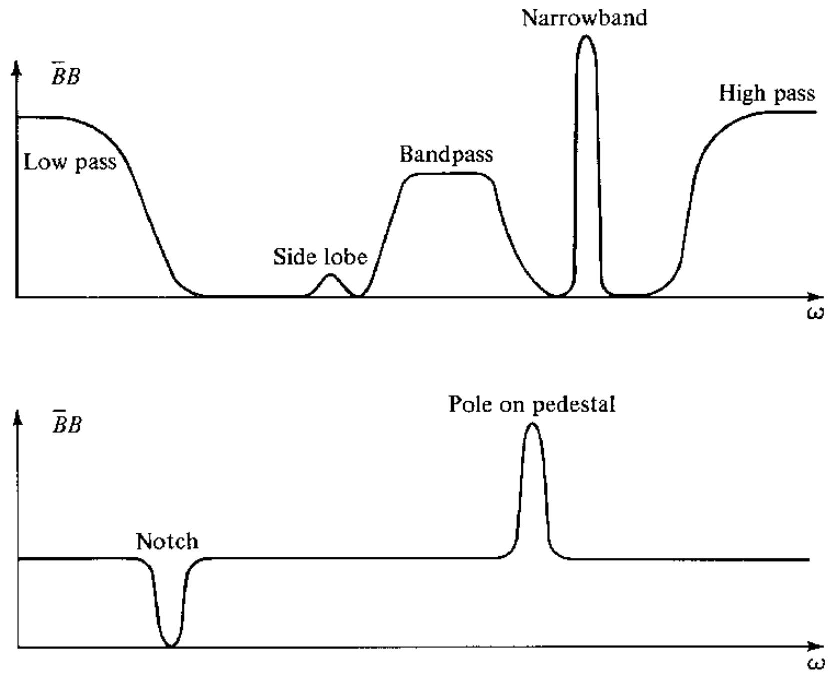


Figure 2.13: Spectra of various filters. c2-2-13 [NR]

We will have frequent occasion to deal with sinusoidal time functions. A simple way to represent a sinusoid by Z transforms is

$$\frac{1}{1 - Ze^{i\omega_0}} \quad = \quad 1 + Ze^{i\omega_0} + Z^2 e^{i2\omega_0} + \cdots \tag{2.26}$$

The time function associated with this Z transform is $e^{i\omega_0 t}$, but it is "turned on" at $t = 0$. Actually, the left-hand side of (2.26) contains a pole exactly on the unit circle, so that the series sits on the borderline between convergence and divergence. This can cause paradoxical situations [you could expand (2.26) so that the sinusoid turns off at $t = 0$] which we will avoid by pushing the pole from the unit circle to a small distance ε outside the unit circle. Let $Z_0 = (1 + \varepsilon)e^{i\omega_0}$. Then define

$$\begin{aligned} B(Z) \quad &= \quad \frac{1}{A(Z)} \quad = \quad \frac{1}{1 - Z/Z_0} \\ &= \quad 1 + \frac{Z}{Z_0} + \left(\frac{Z}{Z_0}\right)^2 + \cdots \end{aligned} \tag{2.27}$$

The time function corresponding to $B(Z)$ is zero before $t = 0$ and is $e^{-i\omega_0 t}/(1 + \varepsilon)^t$ after $t = 0$. It is a sinusoidal function which decreases gradually with time according to $(1 + \varepsilon)^{-t}$. The coefficients are shown in Figure 2.14.

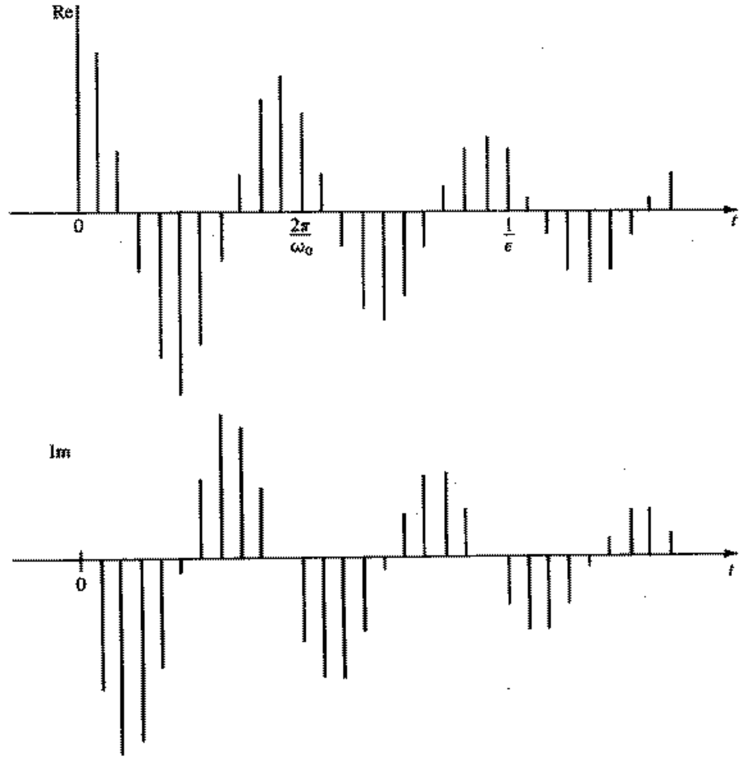


Figure 2.14: The time function associated with a simple pole just outside the unit circle at $Z_0 = 1.1e^{i\pi/5}$. c2-2-14 [NR]

It is intuitively obvious, although we will prove it later, that convolution with the coefficients of (2.27), which are sketched in Figure 2.14, is a narrow-banded filtering operation. If the pole is chosen very close to the unit circle, the filter bandpass becomes narrower and the coefficients of $B(Z)$ drop off more and more slowly. To actually perform the convolution it is necessary to truncate, that is, to drop powers of Z beyond a certain practical limit. It turns out that there is a very much cheaper method of narrow-band filtering than convolution with the coefficients of $B(Z)$. This method is polynomial division by $A(Z)$. We have for the output $Y(Z)$

$$Y(Z) = B(Z)X(Z) \tag{2.28}$$

$$Y(Z) = \frac{X(Z)}{A(Z)} \tag{2.29}$$

Multiply both sides of (2.29) by $A(Z)$

$$Y(Z)A(Z) = X(Z) \tag{2.30}$$

For definiteness, let us suppose the x_t and y_t vanish before $t = 0$. Now identify coefficients of successive powers of Z. We get

$$y_0 a_0 = x_0$$

$$\begin{aligned} y_1 a_0 + y_0 a_1 &= x_1 \\ y_2 a_0 + y_1 a_1 + y_0 a_2 &= x_2 \\ y_3 a_0 + y_2 a_1 + y_1 a_2 + y_0 a_3 &= x_3 \\ y_4 a_0 + y_3 a_1 + y_2 a_2 + y_1 a_3 &= x_4 \\ &\cdots \end{aligned} \tag{2.31}$$

A general equation is

$$y_k a_0 + \sum_{i=1}^{\infty} y_{k-i} a_i = x_k \tag{2.32}$$

Solving for y_k we get

$$y_k = \frac{x_k - \sum_{i=1}^{\infty} y_{k-i} a_i}{a_0} \tag{2.33}$$

Equation (2.33) may be used to solve for y_k once $y_{k-1}, y_{k-2}, \cdots$ are known. Thus the solution is recursive, and it will not diverge if the a_i are coefficients of a minimum-phase polynomial. In practice the infinite limit on the sum is truncated whenever you run out of coefficients of either $A(Z)$ or $Y(Z)$. For the example we have been considering, $B(Z) = 1/A(Z) = 1/(1 - Z/Z_0)$, there will be only one term in the sum. Filtering in this way is called *feedback filtering*, and for narrowband filtering it will be vastly more economical than filtering by convolution, since there are much fewer coefficients in $A(Z)$ than $B(Z) = 1/A(Z)$. Finally, let us examine the spectrum of $B(Z)$. We have

$$\begin{aligned} A(Z) &= 1 - \frac{Z}{Z_0} \\ &= 1 - \frac{e^{i\omega}}{(1+\varepsilon)e^{i\omega_0}} \\ &= 1 - \frac{e^{i(\omega-\omega_0)}}{(1+\varepsilon)} \end{aligned}$$

and

$$\bar{A}\left(\frac{1}{Z}\right) = 1 - \frac{e^{-i(\omega-\omega_0)}}{1+\varepsilon}$$

so

$$\begin{aligned} \bar{A}\left(\frac{1}{Z}\right) A(Z) &= \left(1 - \frac{e^{-i(\omega-\omega_0)}}{1+\varepsilon}\right)\left(1 - \frac{e^{i(\omega-\omega_0)}}{1+\varepsilon}\right) \\ &= 1 + \frac{1}{(1+\varepsilon)^2} - \frac{1}{1+\varepsilon}(e^{-i(\omega-\omega_0)} + e^{i(\omega-\omega_0)}) \\ &= 1 + \frac{1}{(1+\varepsilon)^2} - \frac{2\cos(\omega-\omega_0)}{1+\varepsilon} \end{aligned}$$

$$
\begin{aligned}
&= 1+\frac{1}{(1+\varepsilon)^2}-\frac{2}{1+\varepsilon}+\frac{2}{1+\varepsilon}[1-\cos(\omega-\omega_0)] \\
&= \left(1-\frac{1}{1+\varepsilon}\right)^2+\frac{4}{1+\varepsilon}\sin^2\frac{\omega-\omega_0}{2} \\
\bar{B}\left(\frac{1}{Z}\right)B(Z) &= \frac{(1+\varepsilon)^2}{\varepsilon^2+4(1+\varepsilon)\sin^2\left(\frac{\omega-\omega_0}{2}\right)} \qquad (2.34)
\end{aligned}
$$

To a good approximation this function may be thought of as $1/[\varepsilon^2+(\omega-\omega_0)^2]$. A plot of (2.34) is shown in Figure 2.15.

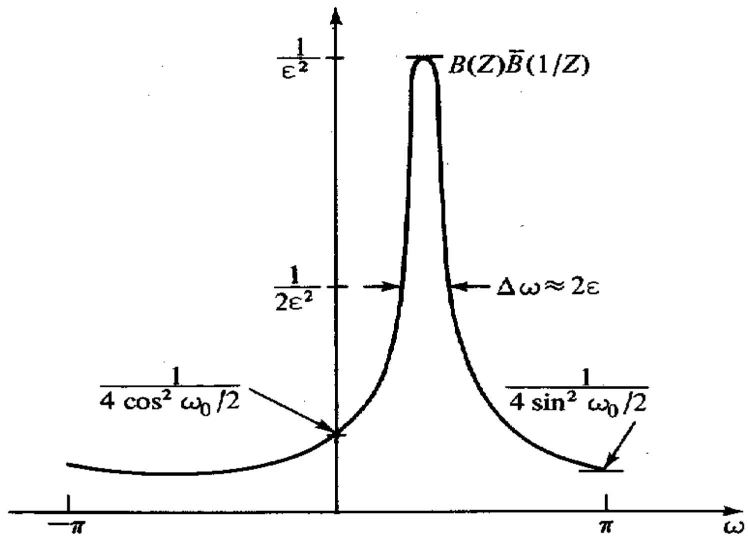


Figure 2.15: Spectrum associated with a single pole at $Z_0 = (1+\varepsilon)e^{i\omega_0}$. $\boxed{\text{c2-2-15}}$ [NR]

Now it should be apparant why this is called a narrowband filter. It amplifies a vary narrow band of frequencies and attenuates all others. The frequency window of this filter is said to be $\Delta\omega \approx 2\varepsilon$ in width. The time window is $\Delta t = 1/\varepsilon$, the damping time constant of the dampend sinusoid b_t.

One practical disadvantage of the filter under discussion is that although its input may be a real time series its output will be a complex time series. For many applications a filter with real coefficients may be preferred.

One approach is to follow the filter $[1, e^{i\omega_0}/(1+\varepsilon)]$ by the time-domain, complex conjugate filter $[1, e^{-i\omega_0}/(1+\varepsilon)]$. The composit time-domain operator is now $[1, (2\cos\omega_0)/(1+\varepsilon), 1/(1+\varepsilon)^2]$ which is real. [Note that the complex conjugate in the frequency domain is $\bar{B}(1/Z)$ but in the time domain it is $\bar{B}(Z) = \bar{b}_0 + \bar{b}_1 Z + \cdots$]. The composite filter may be denoted by $B(Z)\bar{B}(Z)$. The spectrum of this filter is $[B(Z)\bar{B}(1/Z)][\bar{B}(Z)B(1/Z)]$. One may quickly verify that the spectrum of $\bar{B}(Z)$ is like that of $B(Z)$, but the peak is at $-\omega_0$ instead of $+\omega_0$. Thus, the composite spectrum is the product of Figure 2.15 with itself reversed along the frequency axis. This is shown in Figure 2.16.

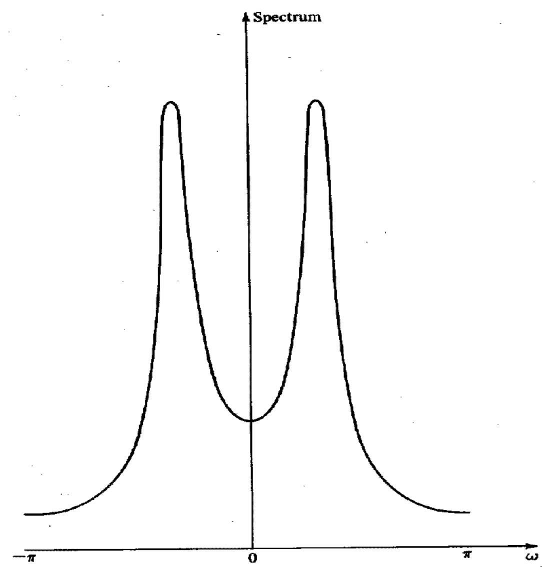


Figure 2.16: Spectrum of a two-pole filter where one pole is like Figure 2.15 and the other is at the conjugate position. $\boxed{\text{c2-2-16}}$ [NR]

EXERCISES:

1 A simple feedback operation is $y_t = (1-\varepsilon)y_{t-1} + x_t$. This operation is called leaky integration. Give a closed form expression for the output y_t if x_t is an impulse. What is the decay time τ of your solution (the time it takes for y_t to drop to $e^{-1}y_0$)? For small ε, say $=0.1$, $.001$, or 0.0001, what is τ?

2 How far from the unit circle are the poles of $1/(1-.1Z+.9Z^2)$? What is the decay time of the filter and its resonant frequency?

3 Find a three-term real feedback filter to pass 59-61 Hz on data which are sampled at 500 points/sec. Where are the poles? What is the decay time of the filter?

2.6 ALL-PASS FILTERS

In this section we consider filters with constant unit spectra, that is, $B(Z)\bar{B}(1/Z) = 1$. In other words, in the frequency domain $B(Z)$ takes the form $e^{i\phi(\omega)}$ where ϕ is real and is called the *phase shift*. Clearly $B\bar{B} = 1$ for all real ϕ. It is an easy matter to construct a filter with any desired phase shift; one merely Fourier transforms $e^{i\phi(\omega)}$ into the time domain. If $\phi(\omega)$ is arbitrary, the resulting time function is likely to be two-sided. Since we are interested in physical processes which are causal, we may wonder what class of functions $\phi(\omega)$ corresponds to one-sided time functions. The easiest way to proceed is to begin with a simple case of a single-pole, single-zero all-pass filter. Then more elaborate all-pass filters can be made up by cascading these simple filters. Consider the filter

$$P(Z) \quad = \quad \frac{Z - 1/(\bar{Z})_0}{1 - Z/Z_0} \tag{2.35}$$

Note that this is a simple case of functions of the form $Z^N \bar{A}(1/Z)/A(Z)$, where $A(Z)$ is a polynomial of degree N or less. Now observe that the spectrum of the filter p_t is indeed a frequency-independent constant. The spectrum is

$$\bar{P}\left(\frac{1}{Z}\right) P(Z) \quad = \quad \frac{1/Z - 1/Z_0}{1 - 1/(Z\bar{Z}_0)} \frac{Z - 1/\bar{Z}_0}{Z - 1/Z_0} \tag{2.36}$$

Multiply top and bottom on the left by Z. We now have

$$\bar{P}\left(\frac{1}{Z}\right) P(Z) \quad = \quad \frac{1 - Z/Z_0}{Z - 1/\bar{Z}_0} \frac{Z - 1/\bar{Z}_0}{1 - Z/Z_0} \quad = \quad 1 \tag{2.37}$$

It is easy to show that $\bar{P}(1/Z)P(Z) = 1$ for the general form $P(Z) = Z^N \bar{A}(1/Z)/A(Z)$. If Z_0 is chosen outside the unit circle, then the denominator of (2.35) can be expanded in positive powers of Z and the expansion in convergent on the unit circle. This means that causality is equivalent to Z_0 outside the unit circle. Setting the numerator of $P(Z)$ equal to zero, we discover that the zero $Z = 1/\bar{Z}_0$ is then inside the unit circle. The situation is depicted in Figure 2.17. To see that the pole and zero are on the same radius line, express Z_0 in polar form $r_0 e^{i\phi_0}$.

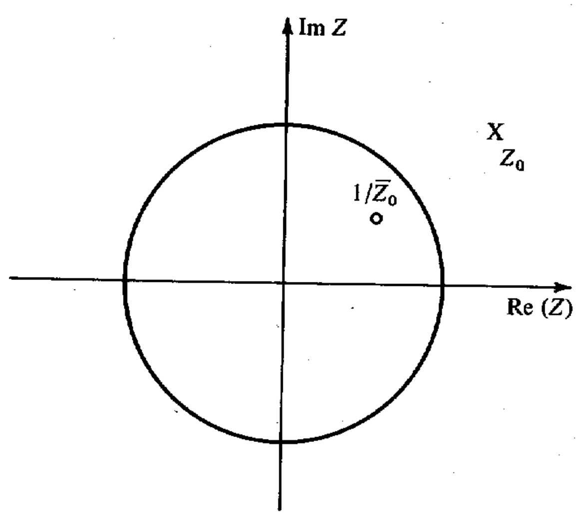


Figure 2.17: The pole of the all-phase filter lies outside the unit circle and the zero is inside. They lie on the same radius line. c2-2-17 [NR]

From Section 2.2 (on minimum phase) we see that the numerator of P is not minimum phase and its phase is augmented by 2π as ω goes from 0 to 2π. Thus the average delay $d\phi/d\omega$ is positive. Not only is the average positive but, in fact, the group delay turns out to be positive in every frequency. To see this, first note that

$$Z \quad = \quad e^{i\omega}$$

$$\begin{aligned} \frac{dZ}{d\omega} &= ie^{i\omega} = iZ \\ \frac{d}{d\omega} &= \frac{dZ}{d\omega}\frac{d}{dZ} = iZ\frac{d}{dZ} \end{aligned} \tag{2.38}$$

The phase of the all-pass filter (or any complex number) may be written as

$$\phi = \ln P(Z) \tag{2.39}$$

Since $|P| = 1$ the real part of the log vanishes; and so, for the all-pass filter (only) we may specialize (2.39) to

$$\begin{aligned} \phi &= \frac{1}{i}\ln P(Z) = \frac{1}{i}\ln\frac{Z - 1/\bar{Z}_0}{1 - Z/Z_0} \\ &= \frac{1}{i}\left[\ln\left(Z - \frac{1}{\bar{Z}_0}\right) - \ln\left(1 - \frac{Z}{Z_0}\right)\right] \end{aligned} \tag{2.40}$$

Using (2.38) the group delay is now found to be

$$\begin{aligned} \tau_g &= \frac{d\phi}{d\omega} = iZ\frac{d\phi}{dZ} = Z\left(\frac{1}{Z - 1/\bar{Z}_0} + \frac{1/Z_0}{1 - Z/Z_0}\right) \\ &= \frac{1}{1 - 1/\bar{Z}_0 Z} + \frac{Z/Z_0}{1 - Z/Z_0} \\ &= \frac{1 - Z/Z_0 + (1 - 1/\bar{Z}_0 Z)(Z/Z_0)}{(1 - 1/\bar{Z}_0 Z)(1 - Z/Z_0)} = \frac{1 - 1/Z_0\bar{Z}_0}{(1 - 1/\bar{Z}_0 Z)(1 - Z/Z_0)} \end{aligned} \tag{2.41}$$

The numerator of (2.41) is a positive real number (since $|Z_0| > 1$), and the denominator is of the form $\bar{A}(1/Z)A(Z)$, which is a spectrum and also positive. Thus we have shown that the group delay of this causal all-pass filter is always positive.

Now if we take a filter and follow it will an all-pass filter, the phases add and the group delay of the composite filter must necessarily be greater than the group delay of original filter. By the same reasoning the minimum-phase filter must have less group delay than any other filter with the same spectrum.

In summary, a single-pole, single-zero all-pass filter passes all frequency components with constant gain and a phase shift which may be adjusted by the placement of a pole. Taking Z_0 near the unit circle causes most of the phase shift to be concentrated near the frequency where the pole is located. Taking the pole further away causes the delay to be spread over more frequencies. Complicated phase shifts or group delays may be built up by cascading several single-pole filters.

EXERCISES:

1 An example of an all-pass filter is the time function $p_t = (\frac{1}{2}, -\frac{3}{4}, -\frac{3}{8}, -\frac{3}{16}, \cdots)$. Calculate a few lags of its autocorrelation by summing some infinite series.

2 Sketch the amplitude, phase, and group delay of the all-pass filter $(1-\bar{Z}_0 Z)/(Z_0 - Z)$ where $Z_0 = (1+\varepsilon)e^{i\omega_0}$ and ε is small. Indicate important parameters on the curve.

3 Show that the coefficients of an all-pass, phase-shifting filter made by cascading $(1-\bar{Z}_0 Z)/(Z_0 - Z)$ with $(1 - Z_0 Z)/(\bar{Z}_0 - Z)$ are real.

4 A continuous time function is the impulse response of a continuous-time, all-pass filter. Describe the function in both time domain and frequency domain. Interchange the words *time* and *frequency* in your description of the function. What is a physical example of such a function? What happens to the statement: "The group delay of an all-pass filter is positive."?

5 A graph of the group delay $\tau_g(\omega)$ in equation (2.41) shows τ_g to be positive for all ω. What is the area under τ_g in the range $0 < \omega < 2\pi$. (HINT: This is a trick question you can solve in your head.)

2.7 NOTCH FILTER AND POLE ON PEDESTAL

In some applications it is desired to reject a very narrow frequency band leaving the rest of the spectrum little changed. The most common example is 60-Hz noise from power lines. Such a filter can easily be made with a slight variation on the all-pass filter. In the all-pass filter the pole and zero have an equal (logarithmic) relative distance from the unit circle. All we need to do is to put the zero closer to the circle. In fact, there is no reason why we should not put the zero right on the circle. Then the frequency at which the zero is located is exactly canceled from the spectrum of input data. If the undesired frequency need not be completely rejected, then the zero can be left just inside or outside the circle. As the zero is moved farther away from the circle, the notch becomes less deep until finally the zero is farther from the circle than the pole and the notch has become a hump. The resulting filter which will be called *pole on pedestal* is in many respects like the narrowband filter discussed earlier. Some of these filters are illustrated in Figures 2.18 and 2.19. The difference between the pole-on-pedestal and the narrowband filters is in the asymptotic behavior away from ω_0. The former is flat, while the latter continues to decay with increasing $| \omega - \omega_0 |$. This makes the pole on pedestal more convenient for creating complicated filter shapes by cascades of single-pole filters.

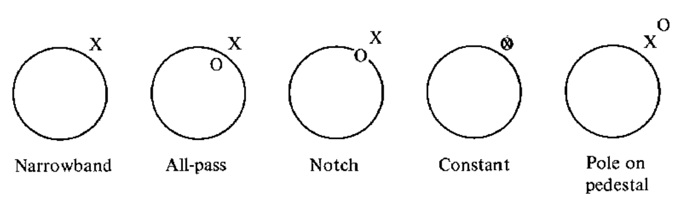


Figure 2.18: Pole and zero locations for some simple filters. Circles are unit circles in the Z plane. Poles are marked by X and *zeros* by 0. $\boxed{\text{c2-2-18}}$ [NR]

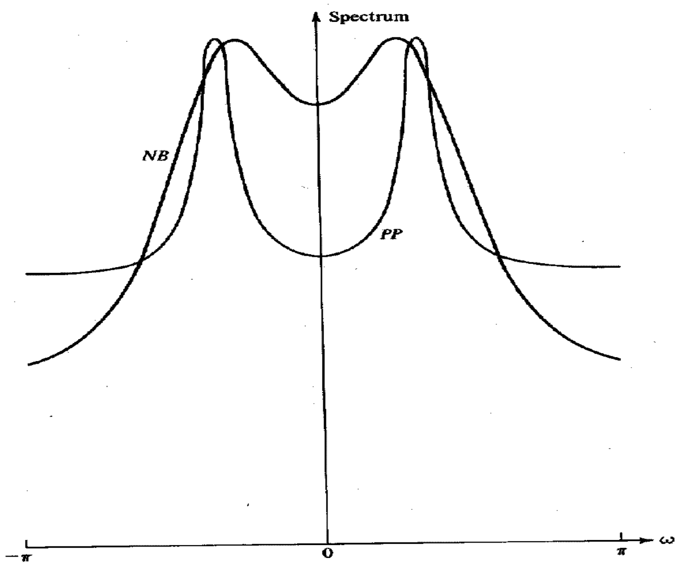


Figure 2.19: Amplitude vs. frequency for narrowband filter (NB) and pole-on-pedestal filter (PP). Each has one pole at $Z_0 = 1.2e^{i\pi/3}$. A second pole at $Z_0 = 1.2e^{-i\pi/3}$ enables the filters to be real in the time domain. c2-2-19 [NR]

Narrowband filters and sharp cutoff filters should be used with caution. An ever-present penalty for such filters is that they do not decay rapidly in time. Although this may not present problems in some applications, it will do so in others. Obviously, if the data collection duration is shorter or comparable to the impulse response of the narrowband filter, then the transient effects of starting up the experiment will not have time to die out. Likewise, the notch should not be too narrow in a 60-Hz rejection filter. Even a bandpass filter (easier to implement with fast Fourier transform than with a few poles) has a certain decay rate in the time domain which may be too slow for some experiments. In radar and in reflection seismology the importance of a signal is not related to its strength. Late-arriving echoes may be very weak, but they contain information not found in earlier echoes. If too sharp a frequency characteristic is used, then filter resonance from early strong arrivals may not have decayed sufficiently by the time that the weak late echoes arrive.

EXERCISES:

1 Consider a symmetric (nonrealizable) filter which passes all frequencies less than ω_0 with unit gain. Frequencies above ω_0 are completely attenuated. What is the rate of decay of amplitude with time for this filter?

2 Waves spreading from a point source decay in energy as the area on a sphere. The amplitude decays as the square root of energy. This implies a certain decay in time. The time-decay rate is the same if the waves reflect from planar interfaces. To what power of time t do the signal amplitudes decay? For waves backscattered to the source from point reflectors, energy decays as distance to the minus fourth power. What is the associated decay with time?

3 Discuss the use of the filter of Exercise 1 on the data of Exercise 2.

4 Design a single-pole, single-zero notch filter to reject 59 to 61 Hz on data which are sampled at 500 points per second.

2.8 THE BILINEAR TRANSFORM

Z transforms and Fourier transforms are related by the relations $Z = e^{i\omega}$ and $i\omega = \ln Z$. A problem with these relations is that simple ratios of polynomials in Z do not translate to ratios of polynomials in ω and vice versa. The approximation

$$-\, i\hat{\omega} \quad = \quad 2\frac{1-Z}{1+Z} \tag{2.42}$$

is easily solved for Z as

$$Z \quad = \quad \frac{1+\hat{\omega}/2}{1-\hat{\omega}/2} \tag{2.43}$$

These approximations are often useful. They are truncations of the exact power series expansions

$$-\, i\omega \quad = \quad -\ln\, e^{i\omega} \quad = \quad -\ln Z \quad = \quad 2\left[\frac{1-Z}{1+Z} + \frac{1}{3}\frac{(1-Z)^3}{(1+Z)^3} + \frac{1}{5}\cdots\right] \tag{2.44}$$

and

$$Z \quad = \quad e^{i\omega} \quad = \quad \frac{e^{i\omega/2}}{e^{-i\omega/2}} \quad = \quad \frac{1+i\omega/2+(i\omega/2)^2/2!+\cdots}{1-i\omega/2+(i\omega/2)^2/2!+\cdots} \tag{2.45}$$

For a Z transform $B(Z)$ to be minimum phase, any root Z_0 of $0 = B(Z_0)$ should be outside the unit circle. Since $Z_0 = \exp\{i[\mathrm{Re}(\omega_0) + i\,\mathrm{Im}(\omega_0)]\}$ and $\mid Z_0 \mid = e^{-\mathrm{Im}(\omega_0)}$, it means that for a minimum phase $\mathrm{Im}\,(\omega_0)$ should be negative. (In other words, ω_0 is in the lower half-plane.) Thus it may be said that $Z = e^{i\omega}$ maps the exterior of the unit circle to the lower half-plane. By inspection of Figures 2.20 and 2.21, it is found that the bilinear approximation (2.42) or (2.43) also maps the exterior of the unit circle into the lower half-plane.

Thus, although the bilinear approximation is an approximation, it turns out to exactly preserve the minimum-phase property. This is very fortunate because if a stable differential equation is converted to a difference equation via (2.42), the resulting difference equation will be stable. (Many cases may be found where the approximation of a time derivative by multiplication with $1 - Z$ would convert a stable differential equation into an unstable difference equation.)

A handy way to remember (2.42) is that $-i\omega$ corresponds to time differentiation of a Fourier transform and $(1 - Z)$ is the first differencing operator. The $(1 + Z)$ in the denominator gets things "centered" at $Z^{1/2}$

To see that the bilinear approximation is a low-frequency approximation, multiply top and bottom of (2.42) by $Z^{-1/2}$

Figure 2.20: Some typical points in the Z-plane, the ω-plane, and the $\hat{\omega}$-plane. c2-2-20 [NR]

	Z	$\omega = 2\pi n - i \ln Z$	$\hat{\omega} = 2i \frac{1-Z}{1+Z}$
A	1	$2\pi n + 0$	0
B	i	$2\pi n + \pi/2$	2
C	-1	$2\pi n + \pi$	$\pm\infty$
D	$-i$	$2\pi n - \pi/2$	-2
E	$\frac{1}{2}$	$2\pi n + .693i$	$i\frac{2}{3}$
F	2	$2\pi n - .693i$	$-i\frac{2}{3}$

$$
\begin{aligned}
-i\hat{\omega} &= 2\,\frac{Z^{-1/2} - Z^{1/2}}{Z^{-1/2} + Z^{1/2}} \\
&= -2i\,\frac{\sin\,\omega/2}{\cos\,\omega/2} \\
\hat{\omega} &= 2\,\tan\,\omega/2
\end{aligned}
\tag{2.46}
$$

Equation (2.46) implicitly refers to a sampling rate of one sample per second. Taking an arbitrary sampling rate Δt, the approximation (2.46) becomes

$$
\omega\,\Delta t \quad \approx \quad 2\,\tan\,\omega\Delta t/2 \tag{2.47}
$$

This approximation is plotted in Figure 2.22. Clearly, the error can be made as small as one wishes merely by sampling often enough; that is, taking Δt small enough.

From Figure 2.22 we see that the error will be only a few percent if we choose Δt small enough so that $\omega_{\max}\Delta t \leq 1$. Readers familiar with the folding theorem will recall that it gives the less severe restraint $\omega_{\max}\Delta t < \pi$. Clearly, the folding theorem is too generous for applications involving the bilinear transform.

Now, by way of example, let us take up the case of a pole $1/-i\omega$ at zero frequency. This is integration. For reasons which will presently be clear, we will consider the slightly different pole

$$
P \quad = \quad \frac{1}{-i\omega + \varepsilon} \tag{2.48}
$$

where ε is small. Inserting the bilinear transform, we get

$$
\begin{aligned}
P &= \frac{1}{2[(1-Z)/(1+Z)] + \varepsilon} \quad = \quad \frac{0.5(1+Z)}{1 - Z + \varepsilon[(1+Z)/2]} \\
&= \frac{0.5(1+Z)}{(1+\varepsilon/2) - Z(1-\varepsilon/2)}
\end{aligned}
\tag{2.49}
$$

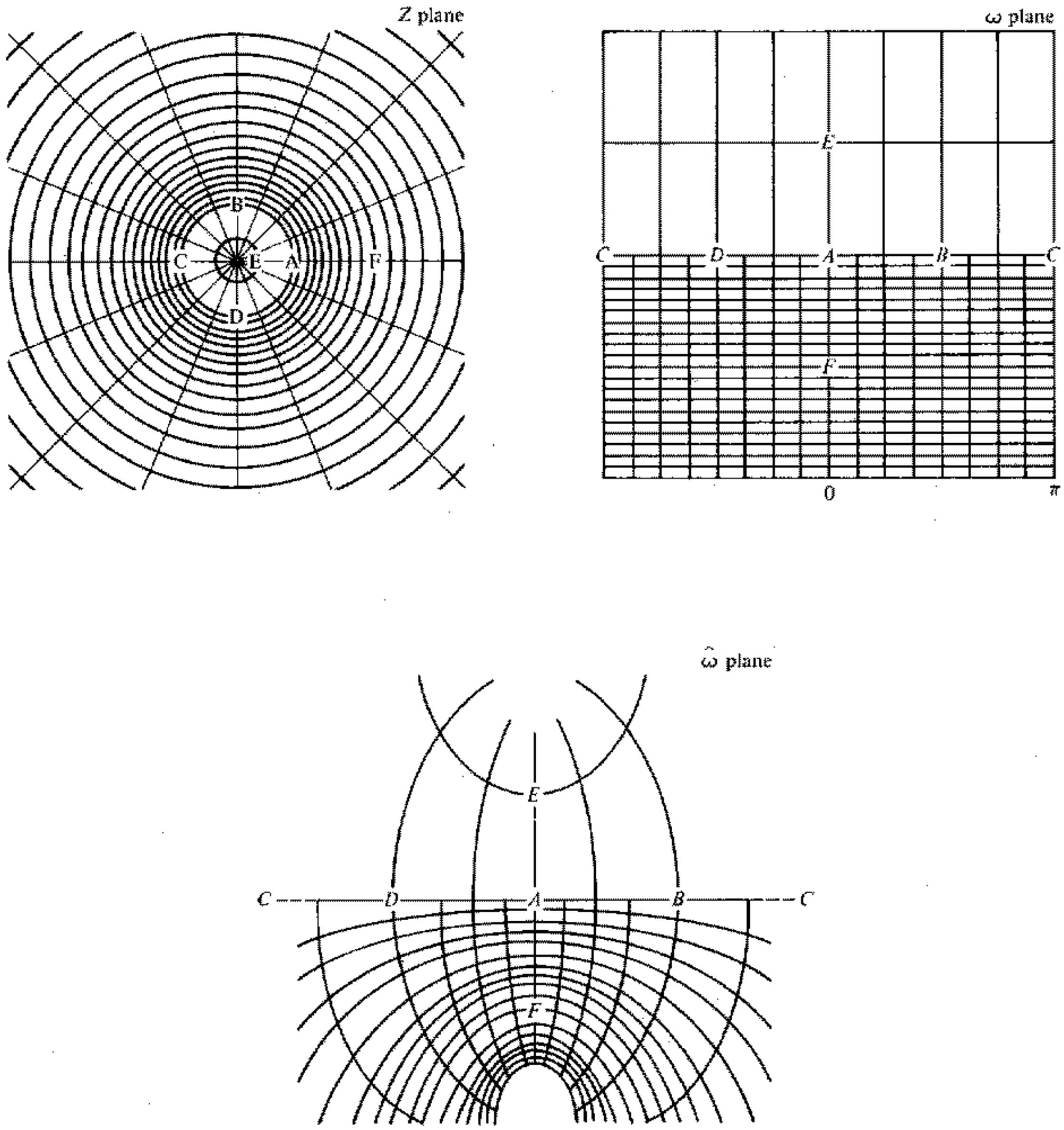


Figure 2.21: The points of Figure 2.20 displayed in the Z plane, the ω plane, and the $\hat{\omega}$-plane. $\boxed{\text{c2-2-21}}$ [NR]

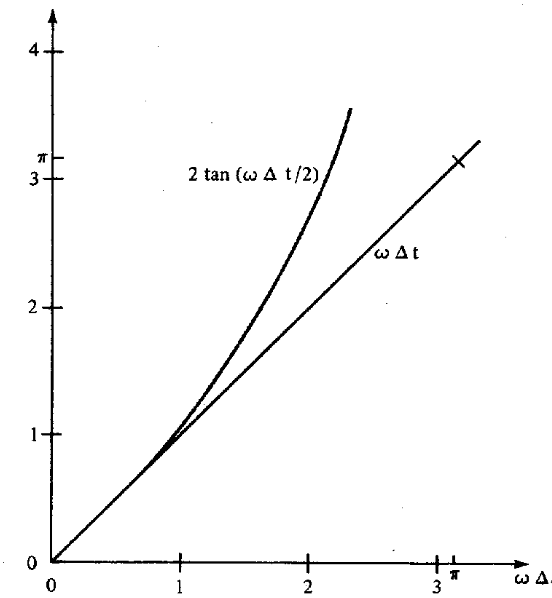


Figure 2.22: The accuracy of the bilinear transformation approximation. c2-2-22 [NR]

By inspection of (2.49) we see that the time-domain function is real, and as ε goes to zero it takes the form (.5, 1, 1, 1, ...). (Taking ε positive forces the step to go out into positive time, whereas ε negative would cause the step to rise at negative time.) The properties of this function are summarized in Figure 2.23.

EXERCISES:

1 In the solution to diffusion problems, the factor $F(\omega) = 1/(-i\omega)^{1/2}$ often arises as a multiplier. To see the equivalent convolution operation, find a causal, sampled-time representation f_t of $F(\omega)$ by identification of powers of Z in

$$(f_0 + f_1 Z + f_2 Z^2 + \cdots)^2 \quad = \quad 1/(-i\omega) \quad \simeq \quad \frac{1}{2}(1+Z)/(1-Z) \tag{2.50}$$

Solve numerically for f_0 through f_7.

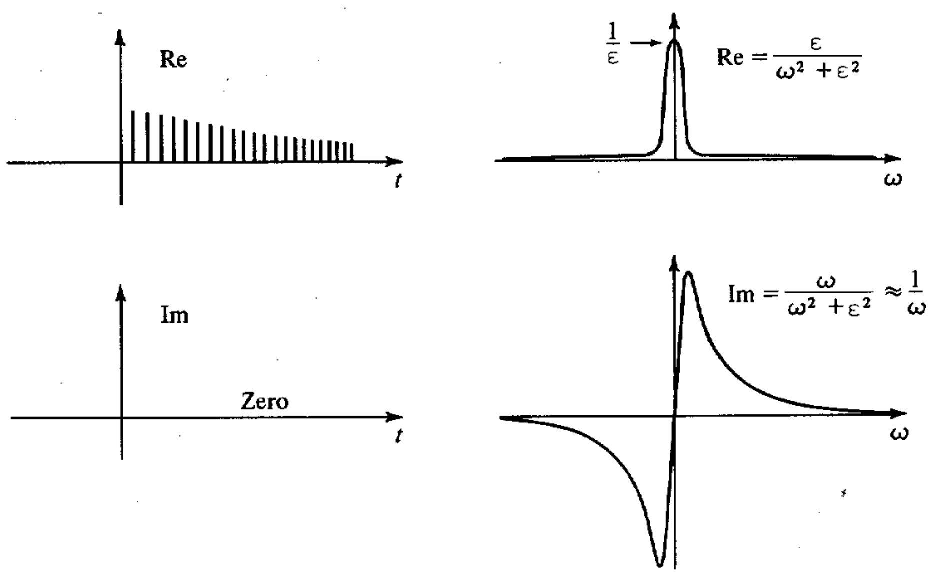


Figure 2.23: Properties of the integration operator. $\boxed{\text{c2-2-23}}$ [NR]

Chapter 3

Spectral factorization

As we will see, there is an infinite number of time functions with any given spectrum. Spectral factorization is a method of finding the one time function which is also minimum phase. The minimum-phase function has many uses. It, and it alone, may be used for feedback filtering. It will arise frequently in wave propagation problems of later chapters. It arises in the theory of prediction and regulation for the given spectrum. We will further see that it has its energy squeezed up as close as possible to $t = 0$. It determines the minimum amount of dispersion in viscous wave propagation which is implied by causality. It finds application in two-dimensional potential theory where a vector field magnitude is observed and the components are to be inferred.

This chapter contains four computationally distinct methods of computing the minimum-phase wavelet from a given spectrum. Being distinct, they offer separate insights into the meaning of spectral factorization and minimum phase.

3.1 ROOT METHOD

The time function (2, 1) has the same spectrum as the time function (1, 2). The autocorrelation is (2, 5, 2). We may utilize this observation to explore the multiplicity of all time functions with the same autocorrelation and spectrum. It would seem that the time reverse of any function would have the same autocorrelation as the function. Actually, certain applications will involve complex time series; therefore we should make the more precise statement that any wavelet and its complex-conjugate time-reverse share the same autocorrelation and spectrum. Let us verify this for simple two-point time functions. The spectrum of (b_0, b_1) is

$$\begin{aligned} \bar{B}\left(\frac{1}{Z}\right) B(Z) &= \left(\bar{b}_0 + \frac{\bar{b}_1}{Z}\right)(b_0 + b_1 Z) \\ &= \frac{\bar{b}_1 b_0}{Z} + (\bar{b}_0 b_0 + \bar{b}_1 b_1) + \bar{b}_0 b_1 Z \end{aligned} \tag{3.1}$$

The conjugate-reversed time function $(\bar{b}_1, \bar{b}_0)$ with Z transform $B_r(Z) = \bar{b}_1 + \bar{b}_0 Z$ has a spectrum

$$\begin{aligned} \bar{B}_r\left(\frac{1}{Z}\right) B_r(Z) &= \left(b_1 + \frac{b_0}{Z}\right)(\bar{b}_1 + \bar{b}_0 Z) \\ &= \frac{b_0\bar{b}_1}{Z} + (b_0\bar{b}_0 + b_1\bar{b}_1) + b_1\bar{b}_0 Z \end{aligned} \tag{3.2}$$

We see that the spectrum (3.1) is indeed identical to (3.2). Now we wish to extend the idea to time functions with three and more points. Full generality may be observed for three-point time functions, say $B(Z) = b_0 + b_1 Z + b_2 Z^2$. First, we call upon the fundamental theorem of algebra (which states that a polynomial of degree n has exactly n roots) to write $B(Z)$ in factored form

$$B(Z) \quad = \quad b_2\,(Z_1 - Z)(Z_2 - Z) \tag{3.3}$$

Its spectrum is

$$R(Z) \quad = \quad \bar{B}\left(\frac{1}{Z}\right) B(Z) \quad = \quad \bar{b}_2 b_2 \left(\bar{Z}_1 - \frac{1}{Z}\right)(Z_1 - Z)\left(\bar{Z}_2 - \frac{1}{Z}\right)(Z_2 - Z) \tag{3.4}$$

Now, what can we do to change the wavelet (3.3) which will leave its spectrum (3.4) unchanged? Clearly, b_2 may be multiplied by any complex number of unit magnitude. What is left of (3.4) can be broken up into a product of factors of form $(\bar{Z}_i - 1/Z)(Z_i - Z)$. But such a factor is just like (3.3). The time function of $(Z_i - Z)$ is $(Z_i, -1)$, and its complex-conjugate time-reverse is $(-1 + \bar{Z}_i Z)$. In a generalization of (3.3) there could be N factors $[(Z_i - Z),\quad i = 1, 2, \ldots, N]$. Any combination of them could be reversed. Hence there are 2^N different wavelets which may be formed by reversals, and all of the wavelets have the same spectrum. Let us look off the unit circle in the complex plane. The factor $(Z_i - Z)$ means that Z_i is a root of both $B(Z)$ and $R(Z)$. If we replace $(Z_i - Z)$ by $(-1 + \bar{Z}_i Z)$ in $B(Z)$, we have removed a root at Z_i from $B(Z)$ and replaced it by another at $Z = 1/\bar{Z}_i$. The roots of $R(Z)$ have not changed a bit because there were originally roots at both Z_i and $1/\bar{Z}_i$ and the reversal has merely switched them around. Summarizing the situation in the complex plane, $B(Z)$ has roots Z_i which occur anywhere, $R(Z)$ must have all the roots Z_i and, in addition, the roots $1/\bar{Z}_i$. Replacing some particular root Z_i by $1/\bar{Z}_i$ changes $B(Z)$ but not $R(Z)$. The operation of replacing a root at Z_i by one at $1/\bar{Z}_i$ may be written as

$$B'(Z) \quad = \quad \frac{Z - 1/\bar{Z}_i}{1 - Z/Z_i} B(Z) \tag{3.5}$$

The multiplying factor is none other than the all-pass filter considered in an earlier chapter. With that in mind, it is obvious that $B'(Z)$ has the same spectrum as $B(Z)$. In fact, there is really no reason for Z_i to be a root of $B(Z)$. If Z_i is a root of $B(Z)$, then $B'(Z)$ will be a polynomial; otherwise it will be an infinite series.

Now let us discuss the calculation of $B(Z)$ from a given $R(Z)$. First, the roots of $R(Z)$ are by definition the solutions to $R(Z) = 0$. If we multiply $R(Z)$ by Z^N (where $R(Z)$ has been given up to degree N), then $Z^N R(Z)$ is a polynomial and the solutions Z_i to $Z^N R(Z) = 0$ will be the same as the solutions of $R(Z) = 0$. Finding all roots of a polynomial is a standard though difficult task. Assuming this to have been done we may then check to see if the roots come in the pairs Z_i and $1/\bar{Z}_i$. If they do not, the $R(Z)$ was not really a spectrum. If they do, then for every zero inside the unit circle, we must have one outside. Refer to Figure 3.1.

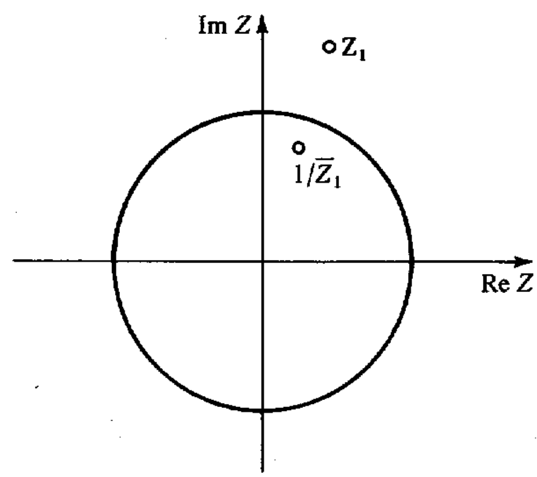


Figure 3.1: Roots of $\bar{B}(1/Z)\,B(Z)$. c3-3-1 [NR]

Thus, if we decide to make $B(Z)$ be a minimum-phase wavelet with the spectrum $R(Z)$, we collect all of the roots outside the unit circle. Then we create $B(Z)$ with

$$B(Z) \quad = \quad b_N\,(Z - Z_1)\,(Z - Z_2) \cdots (Z - Z_N) \tag{3.6}$$

This then summarizes the calculation of a minimum-phase wavelet from a given spectrum. When N is large, it is computationally very awkward compared to methods yet to be discussed. The value of the root method is that it shows certain basic principles.

1. Every spectrum has a minimum-phase wavelet which is unique within a complex scale factor of unit magnitude.

2. There are infinitely many time functions with any given spectrum.

3. Not all functions are possible autocorrelation functions.

The root method of spectral factorization was apparently developed by economists in the 1920s and 1930s. A number of early references may be found in Wold's book, *Stationary Time Series*.

EXERCISES:

1 How can you find the scale factor b_N in (3.6)?

2 Compute the autocorrelation of each of the four wavelets $(4, 0, -1)$, $(2, 3, -2)$, $(-2, 3, 2)$, $(1, 0, -4)$.

3 A power spectrum is observed to fit the form $P(\omega) = 38 + 10\cos\omega - 12\cos 2\omega$. What are some wavelets with this spectrum? Which is minimum phase? (HINT: $\cos 2\omega = 2\cos^2\omega - 1; 2\cos\omega = Z + 1/Z$; use quadratic formula.)

4 Show that if a wavelet $b_t = (b_0, b_1, \ldots, b_n)$ is real, the roots of the spectrum R come in the quadruplets Z_0, $1/Z_0$, $\bar{Z}_0$, and $1/\bar{Z}_0$. Look into the case of roots exactly on the unit circle and on the real axis. What is the minimum multiplicity of such roots?

3.2 ROBINSON'S ENERGY DELAY THEOREM

We will now show that a minimum-phase wavelet has less energy delay than any other one-side wavelet with the same spectrum. More precisely, we will show that the energy summed from zero to any time t for the minimum-phase wavelet is greater than or equal to that of any other wavelet with the same spectrum. Refer to Figure 3-2.

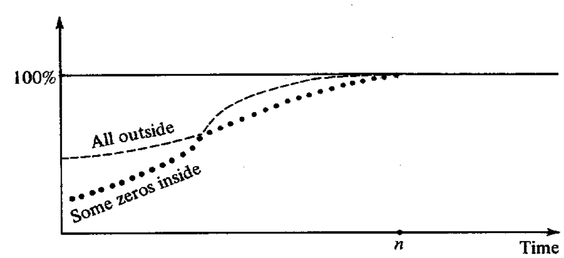


Figure 3.2: Percent of total energy in a filter between time 0 and time t. $\boxed{\text{c3-3-2}}$ [NR]

We will compare two wavelets P_{in} and P_{out} which are identical except for one zero, which is outside the unit circled for P_{out} and inside for P_{in}. We may write this as

$$\begin{aligned} P_{\text{out}}(Z) &= (b + sZ)\,P(Z) \\ P_{\text{in}}(Z) &= (s + bZ)\,P(Z) \end{aligned}$$

where b is bigger than s and P is arbitrary but of degree n. Next we tabulate the terms in question.

t	P_{out}	P_{in}	$P_{\text{out}}^2 - P_{\text{in}}^2$	$\sum_{k=0}^{t}(P_{\text{out}}^2 - P_{\text{in}}^2)$
0	bp_0	sp_0	$(b^2 - s^2)\,{p_0}^2$	$(b^2 - s^2)\,{p_0}^2$
1	$bp_1 + sp_0$	$sp_1 + bp_0$	$(b^2 - s^2)\,({p_1}^2 - {p_0}^2)$	$(b^2 - s^2)\,{p_1}^2$
$\vdots$	$\vdots$			
k	$bp_k + sp_{k-1}$	$sp_k + bp_{k-1}$	$(b^2 - s^2)\,({p_k}^2 - p_{k-1}^2)$	$(b^2 - s^2)\,{p_k}^2$
$\vdots$	$\vdots$			
$n+1$	sp_n	bp_n	$(b^2 - s^2)(-{p_n}^2)$	0

The difference, which is given in the right-hand column, is clearly always positive.

To prove that the minimum-phase wavelet delays energy the least, the preceding argument is repeated with each of the roots until they are all outside the unit circle.

EXERCISES:

1 Do the foregoing minimum-energy-delay proof for complex-valued b, s, and P. [CAUTION: Does $P_{\text{in}} = (s + bZ)\,P$ or $P_{\text{in}} = (\bar{s} + \bar{b}Z)P$?]

3.3 THE TOEPLITZ METHOD

The Toeplitz method of spectral factorization is based on special properties of Toeplitz matrices In this chapter we introduce the Toeplitz matrix to perform spectral factorization. In later chapters we will refer back several times to the algebra described here. When one desires to predict a time series, one can do this with a so-called *prediction filter.* This filter is found as the solution to Toeplitz simultaneous equations. Norman Levinson, in his explanatory appendix of Norbert Wiener's *Time Series,* first introduced the Toeplitz matrix to engineers; however, it had been widely known and used previously in the field of econometrics. It is only natural that it should appear first in economics because there the data are observed at discrete time points, whereas in engineering the idea of discretized time was rather artificial until the advent of digital computers. The need for prediction in economics is obvious. In seismology, it is not the prediction itself but the error in prediction which is of interest. Reflection seismograms are used in petroleum exploration. Ideally, the situation is like radar where the delay time is in direct proportion to physical distance. This is the case for the so-called *primary* reflections. A serious practical complication arises in shallow seas where large acoustic waves bounce back and forth between the sea surface and the sea floor. These are called *multiple* reflections. A mechanism for separation of the primary waves from the multiple reflections is provided by prediction. A multiple reflection is predictable from earlier echoes, but a primary reflection is not predictable from earlier echoes. Thus, the useful information is carried in the part of the seismogram which is *not* predictable. An oil company computer devoted to interpreting seismic exploration data typically solves about 100,000 sets of Toeplitz simultaneous equations in a day.

Another important application of the algebra associated with Toeplitz matrices is in high-resolution spectral analysis. This is where a power spectrum is to be estimated from a sample of data which is short (in time or space). The conventional statistical and engineering knowledge in this subject is based on assumptions which are frequently inappropriate in geophysics. The situation was fully recognized by John P. Burg who utilized some of the special properties of Toeplitz matrices to develop his maximum-entropy spectral estimation procedure described in a later chapter.

Another place where Toeplitz matrices play a key role is in the mathematical physics which describes layered materials. Geophysicists often model the earth by a stack of plane layers or by concentric spherical shells where each shell or layer is homogeneous. Surprisingly enough, many mathematical physics books do not mention Toeplitz matrices. This is because they are preoccupied with *forward* problems; that is, they wish to calculate the waves (or potentials) observed in a known configuration of materials. In geophysics, we are interested in both *forward* problems and in *inverse* problems where we observe waves on the surface of the earth and we wish to deduce material configurations inside the earth. A later chapter contains a description of how Toeplitz matrices play a central role in such inverse problems.

We start with a time function x_t which may or may not be minimum phase. Its spectrum is computed by $R(Z) = \bar{X}(1/Z)\,X(Z)$. As we saw in the preceding sections, given $R(Z)$ alone there is no way of knowing whether it was computed from a minimum-phase function or a nonminimum-phase function. We may suppose that there exists a minimum phase $B(Z)$ of the given spectrum, that is, $R(Z) = \bar{B}(1/Z)\,B(Z)$. Since $B(Z)$ is by hypothesis minimum phase, it has an inverse $A(Z) = 1/B(Z)$. We can solve for the inverse $A(Z)$ in the following way:

$$R(Z) = \bar{B}\left(\frac{1}{Z}\right) B(Z) = \frac{\bar{B}(1/Z)}{A(Z)} \tag{3.7}$$

$$R(Z)\,A(Z) = \bar{B}\left(\frac{1}{Z}\right) = \bar{b}_0 + \frac{\bar{b}_1}{Z} + \cdots \tag{3.8}$$

To solve for $A(Z)$, we identify coefficients of powers of Z. For the case where, for example, $A(Z)$ is the quadratic $a_0 + a_1 Z + a_2 Z^2$, the coefficient of Z^0 in (3.8) is

$$r_0 a_0 + r_{-1} a_1 + r_{-2} a_2 = \bar{b}_0 \tag{3.9}$$

The coefficient of Z^1 is

$$r_1 a_0 + r_0 a_1 + r_{-1} a_2 = 0 \tag{3.10}$$

and the coefficient of Z^2 is

$$r_2 a_0 + r_1 a_1 + r_0 a_2 = 0 \tag{3.11}$$

Bring these together we have the simultaneous equations

$$\begin{bmatrix} r_0 & r_{-1} & r_{-2} \\ r_1 & r_0 & r_{-1} \\ r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \bar{b}_0 \\ 0 \\ 0 \end{bmatrix} \tag{3.12}$$

It should be clear how to generalize this to a set of simultaneous equations of arbitrary size. The main diagonal of the matrix contains r_0 in every position. The diagonal just below the main one contains r_1 everywhere. Likewise, the whole matrix is filled. Such a matrix is called a Toeplitz matrix. Let us define $a'_k = a_k/a_0$. Recall by the polynomial division algorithm that $\bar{b}_0 = 1/\bar{a}_0$. Define a positive number $v = 1/a_0\bar{a}_0$. Now, dividing the vector on each side of (3.12) by a_0, we get the most popular form of the equations

$$\begin{bmatrix} r_0 & r_{-1} & r_{-2} \\ r_1 & r_0 & r_{-1} \\ r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} 1 \\ a'_1 \\ a'_2 \end{bmatrix} = \begin{bmatrix} v \\ 0 \\ 0 \end{bmatrix} \tag{3.13}$$

This gives three equations for the three unknowns a'_1, a'_2, and v. To put (3.13) in a form where standard simultaneous equations programs could be used one would divide the vectors on both sides by v. After solving the equations, we get a_0 by noting that it has magnitude $1/\sqrt{v}$ and its phase is arbitrary, as with the root method of spectral factorization.

At this point, a pessimist might interject that the polynomial $A(Z) = a_0 + a_1 Z + a_2 Z^2$ determined from solving the set of simultaneous equations might not turn out to be minimum phase, so that we could not necessarily compute $B(Z)$ by $B(Z) = 1/A(Z)$. The pessimist might argue that the difficulty would be especially likely to occur if the size of the set (3.13) was not taken to be large enough. Actually experimentalists have known for a long time that the pessimists were wrong. A proof can now be performed rather easily, along with a description of a computer algorithm which may be used to solve (3.13).

The standard computer algorithms for solving simultaneous equations require time proportional to n^3 and computer memory proportional to n^2. The Levinson computer algorithm for Toeplitz matrices requires time proportional to n^2 and memory proportional to n. First notice that the Toeplitz matrix contains many identical elements. Levinson utilized this special Toeplitz symmetry to develop his fast method.

The method proceeds by the approach called recursion. That is, given the solution to the $k \times k$ set of equations, we show how to calculate the solution to the $(k+1) \times (k+1)$ set. One must first get the solution for $k = 1$; then one repeatedly (recursively) applies a set of formulas increasing k by one at each stage. We will show how the recursion works for real-time functions ($r_k = r_{-k}$) going from the 3×3 set of equations to the 4×4 set, and leave it to the reader to work out the general case.

Given the 3×3 simultaneous equations and their solution a_1

$$\begin{bmatrix} r_0 & r_1 & r_2 \\ r_1 & r_0 & r_1 \\ r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} v \\ 0 \\ 0 \end{bmatrix} \tag{3.14}$$

then the following construction defines a quantity e given r_3 (or r_3 given e)

$$\begin{bmatrix} r_0 & r_1 & r_2 & r_3 \\ r_1 & r_0 & r_1 & r_2 \\ r_2 & r_1 & r_0 & r_1 \\ r_3 & r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ 0 \end{bmatrix} = \begin{bmatrix} v \\ 0 \\ 0 \\ e \end{bmatrix} \tag{3.15}$$

The first three rows in (3.15) are the same as (3.14); the last row is the new definition of e. The Levinson recursion shows how to calculate the solution a' to th 4×4 simultaneous equations which is like (3.14) but larger in size.

$$\begin{bmatrix} r_0 & r_1 & r_2 & r_3 \\ r_1 & r_0 & r_1 & r_2 \\ r_2 & r_1 & r_0 & r_1 \\ r_3 & r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} 1 \\ a'_1 \\ a'_2 \\ a'_3 \end{bmatrix} = \begin{bmatrix} v' \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{3.16}$$

The important trick is that from (3.15) one can write a "reversed" system of equations. (If you have trouble with the matrix manipulation, merely write out (3.16) as simultaneous equations, then reverse the order of the unknowns, and then reverse the order of the equations.)

$$\begin{bmatrix} r_0 & r_1 & r_2 & r_3 \\ r_1 & r_0 & r_1 & r_2 \\ r_2 & r_1 & r_0 & r_1 \\ r_3 & r_2 & r_1 & r_0 \end{bmatrix} \begin{bmatrix} 0 \\ a_2 \\ a_1 \\ 1 \end{bmatrix} = \begin{bmatrix} e \\ 0 \\ 0 \\ v \end{bmatrix} \tag{3.17}$$

The Levinson recursion consists of subtracting a yet unknown portion c_3 of (3.17) from (3.15) so as to get the result (3.16). That is

$$\begin{bmatrix} r_0 & r_1 & r_2 & r_3 \\ r_1 & r_0 & r_1 & r_2 \\ r_2 & r_1 & r_0 & r_1 \\ r_3 & r_2 & r_1 & r_0 \end{bmatrix} \left\{ \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ 0 \end{bmatrix} - c_3 \begin{bmatrix} 0 \\ a_2 \\ a_1 \\ 1 \end{bmatrix} \right\} = \left\{ \begin{bmatrix} v \\ 0 \\ 0 \\ e \end{bmatrix} - c_3 \begin{bmatrix} e \\ 0 \\ 0 \\ v \end{bmatrix} \right\} \tag{3.18}$$

To make the right-hand side of (3.18) look like the right-hand side of (3-3-8), we have to get the bottom element to vanish, so we must choose $c_3 = e/v$. This implies that $v' = v - c_3 e = v - e^2/v = v[1 - (e/v)^2]$. Thus, the solution to the 4×4 system is derived from the 3×3 by

$$e \quad \leftarrow \quad \sum_{i=0}^{2} a_i r_{3-i} \tag{3.19}$$

$$\begin{bmatrix} 1 \\ a'_1 \\ a'_2 \\ a'_3 \end{bmatrix} \quad \leftarrow \quad \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ 0 \end{bmatrix} - \frac{e}{v} \begin{bmatrix} 0 \\ a_2 \\ a_1 \\ 1 \end{bmatrix} \tag{3.20}$$

$$v' \quad \leftarrow \quad v\left[1 - (e/v)^2\right] \tag{3.21}$$

We have shown how to calculate the solution of the 4×4 Toeplitz equations from the solution of the 3×3 Toeplitz equations. The Levinson recursion consists of doing this type of step, starting from 1×1 and working up to $n \times n$.

```
      COMPLEX R,A,C,E,BOT,CONJG
      C(1)=-1.; R(1)=1.; A(1)=1.; V(1)=1.
200   DO 220 J=2,N
      A(J)=0.
      E=0.
      DO 210 I=2,J
210   E=E+R(I)*A(J-I+1)
      C(J)=E/V(J-1)
      V(J)=V(J-1)-E*CONJG(C(J))
      JH=(J+1)/2
      DO 220 I=1,JH
      BOT=A(J-I+1)-C(J)*CONJG(A(I))
      A(I)=A(I)-C(J)*CONJG(A(J-I+1))
220   A(J-I+1)=BOT
```

Figure 3.3: A computer program to do the Levinson recursion. It is assumed that the input r_k have been normalized by division by r_0. The complex arithmetic is optional. c3-3-3 [NR]

Let us reexamine the calculation to see why $A(Z)$ turns out to be minimum phase. First, we notice that $v = 1/\bar{a}_0 a_0$ and $v' = 1/(\bar{a}'_0 a'_0)$ are always positive. Then from (3.21) we see that $-1 < e/v < +1$. (The fact that $c = e/v$ is bounded by unity will later be shown to correspond to the fact that reflection coefficients for waves are so bounded.) Next, (3.20) may be written in polynomial form as

$$A'(Z) \quad = \quad A(Z) - (e/v)\, Z^3\, A(1/Z) \tag{3.22}$$

We know that Z^3 has unit magnitude on the unit circle. Likewise (for real time series), the spectrum of $A(Z)$ equals that of $A(1/Z)$. Thus (by the theorem of adding garbage to a minimum-phase wavelet) if $A(Z)$ is minimum phase, the $A'(Z)$ will also be minimum phase. In summary, the following three statements are equivalent:

1. $R(Z)$ is of the form $\bar{X}(\frac{1}{Z})X(Z)$.
2. $|c_k| < 1$.
3. $A(Z)$ is minimum phase.

If any one of the above three is false, then they are all false. A program for the calculation of a_k and c_k from r_k is given in Figure 3.3. In Chapter 8, on wave propagation in layers, programs are given to compute r_k from a_k or c_k.

EXERCISES:

1 The top row of a 4×4 Toeplitz set of simultaneous equations like (3.16) is $(1, \frac{1}{4}, \frac{1}{16}, \frac{1}{4})$. What is the solution a_k?

2 How must the Levinson recursion be altered if time functions are complex? Specifically, where do complex conjugates occur in (3.19), (3.8), and (3.21)?

3 Let $A_m(Z)$ denote a polynomial whose coefficients are the solution to an $m \times m$ set of Toeplitz equations. Show that if $B_k(Z) = Z^k A_k(Z^{-1})$ then

$$v_n \delta_{nm} \quad = \quad \frac{1}{2\pi} \int_0^{2\pi} R(Z)\, B_m(Z) Z^{-n}\, d\omega \qquad n \leq m$$

which means that the polynomial $B_m(Z)$ is orthogonal to polynomial Z^n over the unit circle under the positive weighting function R. Utilizing this result, state why B_m is orthogonal to B_n, that is

$$v_n \delta_{nm} \quad = \quad \frac{1}{2\pi} \int_0^{2\pi} R(Z)\, B_m(Z)\, \bar{B}_n \left(\frac{1}{Z} \right) d\omega$$

(HINT: First consider $n \leq m$, then all n.)

Toeplitz matrices are found in the mathematical literature under the topic of polynomials orthogonal on the unit circle. The author especially recommends Atkinson's book.

3.4 WHITTLE'S EXP-LOG METHOD

In this method of spectral factorization we substitute power series into other power series. Thus, like the root method, it is good for learning but not good for computing. We start with some given autocorrelation r_t where

$$R(Z) \quad = \quad \cdots + r_{-1} Z^{-1} + r_0 + r_1 Z + r_2 Z^2 \cdots$$

If $|R| > 2$ on the unit circle then a scale factor should be divided out. Insert this power series into the power series for logarithms.

$$\begin{aligned} U(Z) &= \ln\, R(Z) \\ &= (R-1) - \frac{(R-1)^2}{2} + \frac{(R-1)^3}{3} - \cdots \qquad 0 < R \leq 2 \\ &= \cdots + u_{-1} Z^{-1} + u_0 + u_1 Z + u_2 Z^2 + \cdots \end{aligned}$$

Of course, in practice this would be a lot of effort, but it could be done in a systematic fashion with a computer program. Now define U_t^+ by dropping negative powers of Z from $U(Z)$

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

Insert this into the power series for the exponential

$$B(Z) \quad = \quad e^{U^+(Z)} \quad = \quad 1 + U^+ + \frac{(U^+)^2}{2!} + \frac{(U^+)^3}{3!} + \cdots$$