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

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

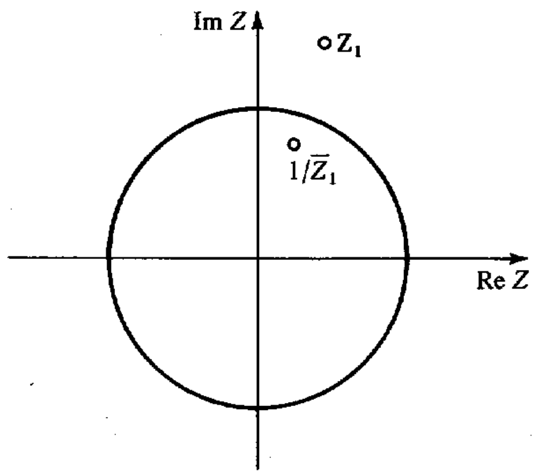
Now, what can we do to change the wavelet (3.3) which will leave its spectrum (3.4) unchanged? Clearly, b_2 may be multiplied by any complex number of unit magnitude. What is left of (3.4) can be broken up into a product of factors of form $(\bar{Z}_i - 1/Z)(Z_i - Z)$. But such a factor is just like (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 $\bar{B}(1/Z)\, B(Z)$. $\boxed{\text{c3-3-1}}$ [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.

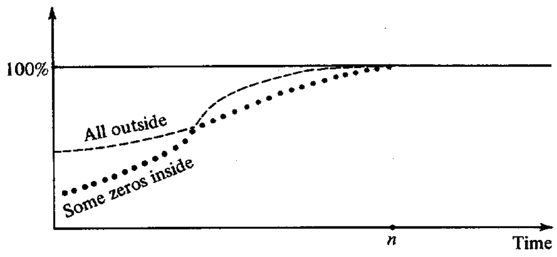


Figure 3.2: Percent of total energy in a filter between time 0 and time t. $\boxed{\text{c3-3-2}}$ [NR]

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:

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

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

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)

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

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.

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

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.)

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

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

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

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

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

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

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

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. $\boxed{\text{c3-3-3}}$ [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

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

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.

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

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

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

The desired minimum-phase wavelet is $B(Z)$; its spectrum is $R(Z)$. To see why this is so, consider the following identities.

$$\begin{aligned}
R(Z) &= e^{\ln R(Z)} \\
&= \exp\left(\frac{u_0}{2} + \sum_{-\infty}^{-1} u_k Z^k + \frac{u^0}{2} + \sum_{+1}^{\infty} u_k Z^k\right) \\
&= \exp\left(\frac{u_0}{2} + \sum_{-\infty}^{-1} u_k Z^k\right) \exp\left(\frac{u^0}{2} + \sum_{1}^{\infty} u_k Z^k\right) \\
&= \exp\left[\overline{U}^+\left(\frac{1}{Z}\right)\right] \exp[U^+(Z)] \\
&= B\left(\frac{1}{Z}\right) B(Z)
\end{aligned}$$

Thus we have factored $R(Z)$ into the desired conjugate parts. Finally, let us see why $B(Z) = e^{U^+(Z)}$ is minimum phase. All we need to know about $U^+(Z)$ is that it is finite and does not contain powers of Z^{-1}. There are two proofs.

The first proof goes by observing that the imaginary part of $U(Z)$ on the unit circle is the phase angle of $B(Z)$. To be minimum phase, the phase of $B(Z)$ must not be augmented by multiples of 2π as Z goes around the unit circle. For minimum phase, the phase should be periodic with period 2π. The phase $u_1 \sin\omega + u_2 \sin 2\omega + \cdots$ obviously satisfies this condition. The second proof is more abstract. In the second proof we note that the only way for $B(Z_0) = e^{U^+(Z_0)}$ to be zero for some Z_0 would be if $U^+(Z_0)$ were equal to $-\infty$; in other words, if U^+ were nonconvergent. This is impossible inside the unit circle because we took the log series for $U(Z)$ to be absolutely convergent on the unit circle; this means $U^+(Z)$ is convergent (finite) inside the circle. Since $B(Z)$ cannot have zeros inside the unit circle, it must be minimum phase.

EXERCISES:

1 How can you get $B(Z)$ if $|R| > 2$ on the unit circle?

2 Suppose $U^+(Z) = aZ$. For values of $a = -1, 0, 1, 100$, and i sketch the time function associated with $B(Z)$. Also sketch their spectra. Find one which resembles a gaussian.

3 Seismograms may be *spectrally balanced* by dividing the F.T. of each seismogram by its magnitude and multiplying by the average magnitude. This amounts to convolution with a symmetrical filter. How can spectral balancing be defined with causal filters?

4 Examine the coefficients of the Z derivative of $\ln B(Z) = U(Z)$ to find a recurrence for u_k given b_k and for b_k given u_k.

3.5 THE KOLMOGOROFF METHOD

If in a computer we have the coefficients $(x_0, x_1, x_2, \ldots)$ of a polynomial $X(Z)$, we say we are "working in the time domain." If we evaluate the polynomial $X(Z)$ at a number of positions on the unit circle, we have numbers, say $X(e^{i\omega_0}), X(e^{i\omega_1}), \ldots, X(e^{i\omega_n})$, which we call a frequency-domain representation of the polynomial $X(Z)$. We have seen that the fast Fourier transform is a very cheap way of going from the time domain to the frequency domain and back. This makes it very worthwhile to look at Whittle's factorization method in the frequency domain. Furthermore, we will understand spectral factorization from yet another point of view.

We may begin with a time function or Z transform $X(Z) = x_0 + x_1 Z + \cdots$. Let us denote by X_k the transform of the time function, that is, $X(Z)$ evaluated at numerous $(k = 0, 1, \ldots, n)$ places on the unit circle. Consider the identities

$$\begin{aligned} R_k &= \bar{X}_k X_k \\ &= e^{\ln(R_k)} = e^{U_k} \end{aligned}$$

Now we add and subtract a still arbitrary function Φ_k to the exponential

$$\begin{aligned} R_k &= \exp\left[\frac{1}{2}(U_k - i\Phi_k)\right] \exp\left[\frac{1}{2}(U_k + i\Phi_k)\right] \\ &= \bar{B}_k B_k \end{aligned}$$

Now the big question is what Φ_k should be used to guarantee that B_k transforms to a minimum-phase, one-sided time function? By looking at Whittle's method, we note that the only significant properties of $U^+(Z)$ are that it is finite and that the time function u_t vanishes before $t = 0$. Thus we expect that Φ_k should be chosen so that when $U_K + i\Phi_k$ is transformed into the time domain the resulting time function u_t^+ should vanish for negative time. This may be done as depicted in Figure 3.4.

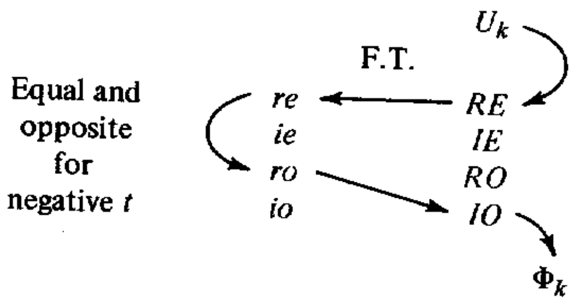


Figure 3.4: Determination of the phase function. $\boxed{\text{c3-3-4}}$ [NR]

To see how easy it really is to get the imaginary odd part IO, we fetch the integration filter from Sec. 2-8 (on bilinear transformation) and display it in Figure 3.5.

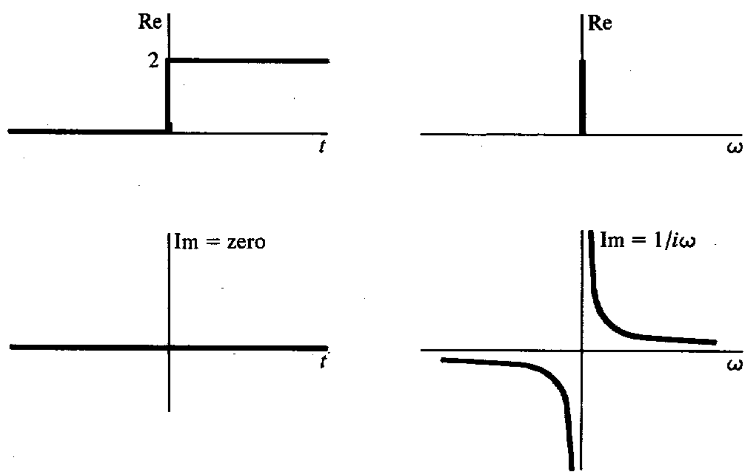


Figure 3.5: The transform pair used in the Hilbert transform. c3-3-5 [NR]

To get Φ_k we take U_k into the time domain, getting u_t. Then we multiply by the real step function of time in Figure 3.5, obtaining $u_t^+ = u_0, u_1, \ldots,$. This implies that in the frequency domain U_k has been convolved with $\delta_{k=0} + i \times$ (90° phase shift filter). Thus, Φ has been generated.

Let us reconsider the operation of dropping all of the negative powers of Z in $U(Z)$ as we did in the previous section to get $U^+(Z)$. For simplicity, consider the case r_t real; then u_t is real.

$$\begin{aligned} U(Z) &= \cdots + u_1 Z^{-1} + u_0 + u_1 Z + u_2 Z^2 + \cdots \\ U &= u_0(\cos 0) + 2u_1 \cos\omega + 2u_2 \cos 2\omega + \cdots \end{aligned}$$

Now let us make up a new function Φ by replacing cosine by sine in the foregoing expression

$$\Phi = 2u_i \sin\omega + 2u_2 \sin 2\omega + \cdots$$

We now see that combining U with $i\Phi$ we get U^+.

$$\begin{aligned} \frac{1}{2}(U + i\Phi) &= \frac{1}{2}u_0 + u_1 Z + u_2 Z^2 + \cdots \\ &= U^+(Z) \end{aligned}$$

Notice that the operation of changing cos t to sin t would be called 90° phase shift filtering. Here we have changed cos ω to sin ω with the result that $U^+(Z)$ has only positive coefficients of Z.

The Kolmogoroff method of spectral factorization is very fast in a computer because fast Fourier transforms may be used. Its principle disadvantage is that summation around the unit circle is always slightly different than integration about the circle. When the spectrum is simple but poles are very close to the unit circle, then the Toeplitz method may prove more satisfactory. A simple program to do spectral factorization is given in Figure 3.6.

EXERCISES:

1 Insert the additional arrows in Figure 3.4 which are required when dealing with complex time functions.

2 What is the meaning of *minimum-phase waveform* if the roles of time domain and frequency domain are interchanged?

3 Show how to do the inverse Hilbert transform, given ϕ find u. What is the interpretation of the fact that you cannot get u_0?

4 Consider a model of a portion of the earth where x is the north coordinate, $+z$ represents altitude above the earth, and magnetic bodies are distributed in the earth so as to create no magnetic field component in the east-west direction. One may show that the magnetic field h above the earth is represented by

$$\begin{bmatrix} h_x(x,z) \\ h_z(x,z) \end{bmatrix} = \int_{-\infty}^{+\infty} F(k) \begin{bmatrix} -ik \\ |k| \end{bmatrix} e^{ikx-|k|z}\, dk$$

Here $F(k)$ is some spatial frequency spectrum.

(a) By using Fourier transforms, how does one compute $h_x(x,0)$ from $h_z(x,0)$ and vice versa?

(b) Given $h_z(x,0)$, how does one compute $h_z(x,z)$?

(c) Notice that at $z=0$

$$w(x) = h_z(x) + ih_x(x) = \int_{-\infty}^{+\infty} e^{ikx} F(k)\,(|k|+k)\, dk$$

and that $F(k)(|k|+k)$ is a one-sided function of k. With a total field magnetometer one observes

$$h_x^2(x) + h_z^2(x) = w(x)\bar{w}(x)$$

What can you say about getting $F(k)$ from this?

(d) How unique are $h_x(x)$ and $h_z(x)$ if $w(x)\bar{w}(x)$ is given?

```
      LX=4                                 LX MUST BE A POWER OF 2.
      WRITE(6,FMT) (CX(I),I=1,LX)          PRINT THE TEST EXAMPLE.
      CALL FORK(LX,CX,1.)                  FOURIER TRANSFORM.
      DO 10 I=1,LX
10    CX(I)=.5*CLOG(CX(I)*CONJG(CX(I)))    FORM THE LOG OF THE SPECTRUM
      CALL FORK(LX,CX,-1.)                 BEGIN THE HILBERT TRANSFORM
      K=LX/2                               LX MUST BE EVEN
      DO 20 J=2,K                          LEAVE  T=0  ALONE
      CX(J)=CX(J)+CX(J)                    DOUBLE VALUES AT POSITIVE TIME
20    CX(LX+2-J)=0.                        ZERO VALUES AT NEGATIVE TIME
      CALL FORK(LX,CX,+1.)                 END HILBERT TRANSFORM
      DO 30 I=1,LX
30    CX(I)=CEXP(CX(I))                    EXPONENTIATE
      CALL FORK(LX,CX,-1.)                 INVERSE FOURIER TRANSFORM
      WRITE(6,FMT)(CX(I),I=1,LX)           PRINT MINIMUM PHASE WAVELET.

IN:  ( 0.0, 0.0),   ( 0.0, 8.0),   (16.0, 0.0),   ( 0.0, 0.0),

OUT: (15.6,-0.0),   ( 0.0,-8.7),   (-0.4,-0.0),   ( 0.0,-0.7),
```

Figure 3.6: A program to do spectral factorization by means of fast Fourier transform and Hilbert transform. Complex arithmetic is mandatory. Results are approximate since integration around the unit circle has been approximated by summation over four points. c3-3-6 [NR]

3.6 CAUSALITY AND WAVE PROPAGATION

The principle of causality, i.e., no response before a stimulus, places certain restraints upon the nature of wave propagation. Causality often seems to be violated because of

1. Approximations in a theory which are made to simplify it.
2. Fitting inappropriate curves to experimental observations.
3. Approximations made in converting differential equations to difference equations.

Violations of causality often result from seemingly inconsequential approximations at spectral frequencies outside the range of practical interest. The seemingly inconsequential approximations often take on significance because of the very weak convergence in the Hilbert transform. With regard to (1) and (2), the violation of causality may turn out to be a low price to pay compared with the cost of a more precise analysis. With regard to (3) or computation in general, the violation of causality usually has disastrous consequences. Even though difference equations may have proper behavior at frequencies of interest, exponential growth at other frequencies is almost always so severe as to completely obliterate the desired solution.

Most wave-propagation theories are worked out in the frequency domain in homogeneous materials. If they are further done in cartesian geometry, the end result is that propagation along the x axis is given by

$$Y(x) \quad = \quad Y(0)\, e^{ikx}. \tag{3.23}$$

What the wave theory provides is an expression for k in terms of frequency and all the physical parameters such as velocity, viscosity, etc. We will now take an "after-the-derivation" look at $k(\omega)$ and see if it satisfies causality. In other words, (3.23) is like the filter situation depicted in Figure 3.7. $Y(0)$ is the filter input, $Y(x)$ is like the output, and e^{ikx} is a filter which models the effect of propagation. The question is whether the filters are one-sided. This depends upon the detailed form of $k(\omega)$.

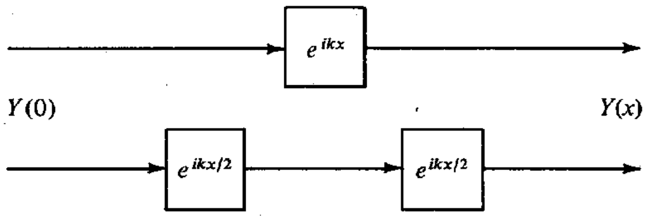


Figure 3.7: Modeling wave propagation with filters. $\boxed{\text{c3-3-7}}$ [NR]

First, consider the simple case when $k(\omega)$ is real for all real ω. Then e^{ikx} is an all-pass filter. It is realizable if and only if $k(\omega)x$ is a monotonically increasing function of ω.

Next, suppose $k(\omega)$ is complex. This means that attenuation will occur. Clearly energy conservation requires that the real part of ikx be negative.

Define $P(Z)$ to be a realizable all-pass filter and $B(Z)$ to be a realizable minimum-phase wavelet. Any realizable function can be represented in the form $B(Z)P(Z)$. We will require e^{ikx} to be so represented.

$$B(Z)P(Z) \quad = \quad e^{ikx} \quad = \quad \exp[i\,(k_r + ik_r)x] \tag{3.24}$$

Taking logs and using ϕ to denote phase, we have

$$\ln|B| + i\phi_B + i\phi_P \quad = \quad -k_i x + ik_r x \tag{3.25}$$

Splitting into real and imaginary parts we obtain

$$\begin{aligned} \ln|B| &= -k_i x \\ \phi_P &= k_r x - \phi_B \end{aligned}$$

First of all, from the attenuation $-k_i x = \ln|B|$ we can by Hilbert transform compute ϕ_B. When it is subtracted from $k_r x$ we are left with the phase shift ϕ_P of the all-pass filter. Recall from Sec. 2-6 (on all-pass filters) that if and only if ϕ_P is monotonically increasing, we have a causal all-pass filter. In conclusion, the test of causality for a function $k(\omega)$ lies in computing $\phi_P(\omega)$ and seeing whether it is monotonically increasing for all frequencies.

A great deal of confusion results from attenuation laws which are thought to be of the form $k_i \simeq \omega^2$ (gaussian attenuation) or $k_i \simeq |\omega|$ (sometimes called *constant* Q). The reason is that these functions are not integrable so they do not have convergent Hilbert transforms; that is, the ϕ_B turns out to be infinite. These functions correspond to time pulses like the gaussian time function e^{-t^2} which do not have a time before which they are zero. Such functions cannot strictly be considered to be associated with any causal process; however, they may give excellent practical approximations. In reality, physical dissipation is generally associated with a relaxation phenomenon having a finite relaxation time. For frequency values below the relaxation value, the attenuation may appear to be increasing indefinitely with frequency, but, in fact, the attenuation decreases above the characteristic relaxation frequency. The attenuation also provides a phase velocity aberration because of the addition of ϕ_B to ϕ_P but this is generally considered to be too small to measure.

Now let us take up an example of computer modeling of wave propagation. The simplest example is propagation without dissipation. Then $k = \omega/v$ where v is the velocity. Rather than attempt to construct a filter which will carry waves a long distance we will only attempt to carry them a distance Δx. The filter may be used N times to propagate a distance $x = N\,\Delta x$. We will use the bilinear truncation of the power series for exponential

$$e^u \quad \approx \quad \frac{1+\frac{u}{2}}{1-\frac{u}{2}} \tag{3.26}$$

and the bilinear representation for $i\omega\,\Delta t$, namely

$$i\omega\,\Delta t \quad \approx \quad 2\left(\frac{Z-1}{Z+1}\right) \tag{3.27}$$

With these and $k = \omega/v$ and the definition $a = \Delta x/v\,\Delta t$ we obtain

$$\begin{aligned} e^{ik\,\Delta x} \quad \approx \quad e^{i\hat{k}\,\Delta x} &= \frac{1+a\left(\frac{Z-1}{Z+1}\right)}{1-a\left(\frac{Z-1}{Z+1}\right)} \\ &= \frac{(Z+1)+a(Z-1)}{(Z+1)-a(Z-1)} \\ &= \frac{(1-a)+(1+a)\,Z}{(1+a)+(1-a)\,Z} \end{aligned} \tag{3.28}$$

That this is the general form of an all-pass filter may be seen by taking $Z_0 = -(1+a)/(1-a)$. Then (3.28) becomes

$$e^{ik\,\Delta x} \quad = \quad \frac{-1/Z_0 + Z}{1 - Z/Z_0} \tag{3.29}$$

The denominator will be minimum phase and the all-pass filter will be realizable if $Z_0 > 1$ or $\Delta x / v\, \Delta t > 0$. This means that present and past values of the inputs will determine the output if the wave is being projected in the $+x$ direction. If it is desired to project the wave backwards in space (Δx negative), then Δt must be taken negative, so that present and future input values determine the present output.

EXERCISES:

1 What is the velocity error as a function of ω for the filter of (3.28) for values of $a = .1$, 1., and 2. HINT: Let $\hat{k} = \omega / \hat{v}$ so $v/\hat{v} = \hat{k} v / \omega$. Then note that

$$\begin{aligned} i\hat{k}\,\Delta x &= \ln e^{i\hat{k}\Delta x} \quad = \quad \ln Z + \ln\left(\frac{-1}{Z_0 Z} + 1\right) - \ln\left(1 - \frac{Z}{Z_0}\right) \\ &= i\omega\,\Delta t - 2\,\mathrm{Im}\ln\left(1 - \frac{Z}{Z_0}\right) \end{aligned}$$