Elec4621:
Advanced Digital Signal Processing
Chapter 10: Subband Transforms and the Discrete Wavelet Transform

Dr. D. S. Taubman

May 16, 2011

1 Tree-Structured Subband Transforms

We have already seen two-channel filter banks used to represent a signal, $x[n]$, in terms of two subbands, $y_0[n]$ and $y_1[n]$, corresponding to the low and high frequency components of the original signal. The combined sampling rate of the two subbands is identical to that of the original signal. The representation can be alias free by appropriate choice of the filters in the filter bank and it is even possible to achieve perfect reconstruction. We will briefly consider the design of perfect reconstruction systems later. For the moment, however, we will simply assume the existence of appropriate designs, and we will use only perfect reconstruction filters.

It is obvious that our two-channel filter bank can be expanded to a larger number of channels by recursive application, in the manner suggested by Figure 1.In the figure, each subband is split into two further subbands, using the same subband transform. The invertibility of the complete transform follows trivially from the fact that the individual two-channel building blocks are invertible (perfect reconstruction). The figure also shows the pass band structure of this tree-structured transform. Note carefully the order of the subband pass bands. This is a consequence of the frequency reversal which occurs when high-pass subbands are sub-sampled.

An alternative tree-structured subband decomposition is that shown in Figure 2. In this case, only the low-pass channel is recursively expanded into further subbands, and the subbands have logarithmically spaced pass bands and bandwidths. The lowest frequency subbands have very narrow

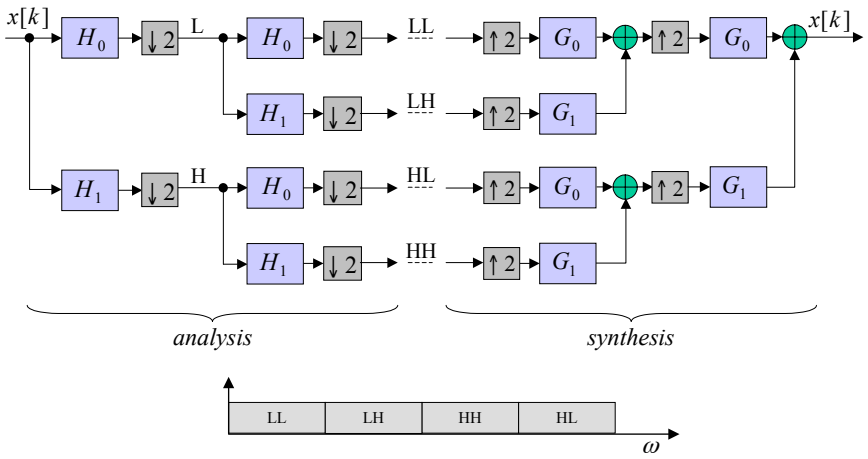

Figure 1: *Tree-structured subband transform with uniformly spaced pass-bands.*

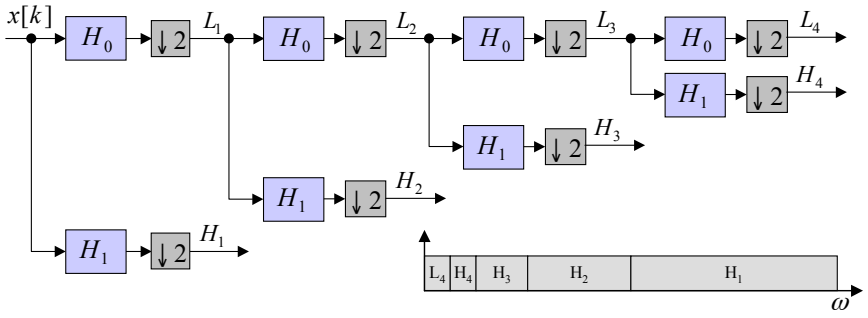

Figure 2: *Tree-structured subband transform with logarithmically spaced pass bands.*

bandwidths and correspondingly reduced sample rates, while the highest frequency subband occupies half the bandwidth of the original signal and has half the sampling rate.

The structure in Figure 2 is interesting for a variety of reasons. Firstly, notice that the synthesis system reconstructs the signal incrementally, with the intermediate results corresponding to successively higher resolution versions of the signal. That is, each of the intermediate signals, L_d, is a low-pass filtered and sub-sampled version of the original signal with successively higher bandwidths and sampling densities. If we stop the reconstruction at some point, we get a usable representation of the original signal at a correspondingly reduced resolution. This is useful in compression applications, since if we compress each of the subbands separately then it is not hard to see how we can discard some parts of the bit-stream to obtain lower resolution representations of the signal, without having to decompress and recompress everything.

In view of the above property, tree-structured subband transforms of the form shown in Figure 2 are known as multi-resolution transforms. This same structure is also commonly known as a Discrete Wavelet Transform (DWT). While we do not have time here to carefully investigate the subtle differences between wavelet transforms and generic tree-structured subband transforms, we point out that only certain choices for the subband filters lead to what we can strictly call a wavelet transform, while all discrete wavelet transforms are tree-structured subband transforms. Wavelet transforms provide a representation of the original signal, $x[n]$, in terms of a base resolution (or scale) component L_D, together with a progression of detail components, H_{D-1}, H_{D-2}, ..., H_1, which can be combined with the base resolution (subband synthesis) to recover versions of the signal at successively higher resolutions (or successively more detailed scales).

Logarithmically spaced subbands are also useful for representing audio waveforms, since the "hair cells" on the Basilar membrane in the human ear (the cochlear) have approximately logarithmically spaced frequency response characteristics. Most phenomena of interest for auditory perception, such as sensitivity to different frequencies, auditory masking and so on, can be reasonably modeled using a filter bank system which produces subbands with logarithmically spaced pass bands. Below about 500 Hz, the logarithmic spacing is best replaced with uniformly spaced subbands, suggesting an appropriate combination of the logarithmic and uniform decomposition structures shown in Figures 1 and 2. Such transforms form an excellent framework for efficient processing and compression of auditory signals, while preserving all of the original signal information (they can be inverted if nec-

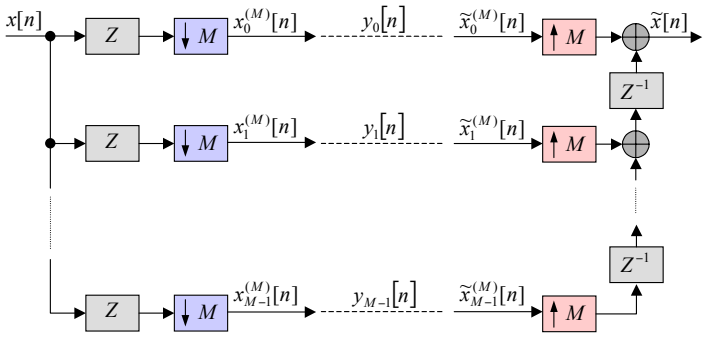

Figure 3: *Trivial subband filter bank and its inverse, where the subbands are simply the polyphase components of the input (and output) sequence.*

essary).

2 Polyphase Filter Bank Representation

Recall from the discussion of rational sampling rate changes in the previous chapter of your lecture notes, that we can write the input signal $x[n]$ in terms of its polyphase components,

$$x_p^{(M)}[n] = x[Mn + p]$$

where M is the number of channels in the filter bank. Similarly, we can write the output sequence produced by the subband synthesis system, $\tilde{x}[n]$, in terms of its polyphase components,

$$\tilde{x}_p^{(M)}[n] = \tilde{x}[Mn + p]$$

The polyphase components have the same sampling rate as the subbands and so in the simplest case we can identify the subbands with the polyphase components, i.e.

$$y_m[n] = x_m^{(M)}[n] = \tilde{x}_m^{(M)}[n]$$

and obtain a trivial perfect reconstruction filter bank illustrated in Figure 3. Of course, the subbands in this trivial system have no band pass properties and are riddled with aliasing.

Now if a non-trivial set of filters, $h_m[n]$, is involved, we can write down the subband analysis operation in terms of polyphase components by using

the efficient implementation derived in the last set of lecture notes for rational sampling rate changes (in this case, the decimation factor, $M_d = M$, while the interpolation factor, $M_u = 1$). We find that

$$\begin{aligned} y_m[n] &= \sum_k x[k] \cdot h_m[Mn-k] \\ &= \sum_{p=0}^{M-1} \sum_k x_p^{(M)}[k] \cdot h_m[Mn-(Mk+p)] \\ &= \sum_{p=0}^{M-1} \sum_k x_p^{(M)}[k]\, h_{m,p}[n-k] \end{aligned}$$

where

$$h_{m,p}[n] \triangleq h_m[Mn-p]$$

Thus, each subband sequence, $y_m[n]$, can be written as a sum of the filtered polyphase input components, where the polyphase filters, $h_{m,p}[n]$, are obtained by appropriately sub-sampling the original subband analysis filters, $h_m[n]$. The relationship can be conveniently expressed in the Z-transform domain as

$$Y_m(z) = \sum_{p=0}^{M-1} X_m^{(M)}(z) H_{m,p}(z)$$

and the entire analysis system can then be written in matrix form as

$$\begin{aligned} \begin{pmatrix} Y_0(z) \\ Y_1(z) \\ \vdots \\ Y_{M-1}(z) \end{pmatrix} &= \begin{pmatrix} H_{0,0}(z) & H_{0,1}(z) & \cdots & H_{0,M-1}(z) \\ H_{1,0}(z) & H_{1,1}(z) & \cdots & H_{1,M-1}(z) \\ \vdots & \vdots & \ddots & \vdots \\ H_{M-1,0}(z) & H_{M-1,1}(z) & \cdots & H_{M-1,M-1}(z) \end{pmatrix} \cdot \begin{pmatrix} X_0^{(M)}(z) \\ X_1^{(M)}(z) \\ \vdots \\ X_{M-1}^{(M)}(z) \end{pmatrix} \\ &= \mathbf{H}(z) \cdot \begin{pmatrix} X_0^{(M)}(z) \\ X_1^{(M)}(z) \\ \vdots \\ X_{M-1}^{(M)}(z) \end{pmatrix} \end{aligned}$$

where $\mathbf{H}(z)$ is known as the polyphase analysis matrix. The polyphase analysis matrix may be understood as an M-input/M-output LTI system.

We may proceed in an analogous manner for the synthesis system. In particular, appealing again to the efficient implementation of downsampling/upsampling systems in the previous set of lecture notes (here, we use

$M_u = M$ and $M_d = 1$), we see that the p^{th} polyphase component of the output sequence, $\tilde{x}_p^{(M)}[n]$, can be written as

$$\begin{aligned}\tilde{x}_p^{(M)}[n] &= \sum_{m=0}^{M-1} \sum_k y_m[k]\, g_m[(Mn+p) - Mk] \\ &= \sum_{m=0}^{M-1} \sum_k y_m[k]\, g_{p,m}[n-k]\end{aligned}$$

where

$$g_{p,m}[n] \triangleq g_m[Mn+p]$$

Thus, each output polyphase component, $\tilde{x}_p^{(M)}[n]$, may be written as a sum of the filtered sub-sequences, where the relevant filters are obtained by appropriately sub-sampling the synthesis filters, $g_m[n]$. Again, we may write this in the Z-transform domain as

$$\tilde{X}_p^{(M)}(z) = \sum_{m=0}^{M-1} Y_m(z) G_{p,m}(z)$$

and in matrix notation as

$$\begin{aligned}\begin{pmatrix} \tilde{X}_0^{(M)}(z) \\ \tilde{X}_1^{(M)}(z) \\ \vdots \\ \tilde{X}_{M-1}^{(M)}(z) \end{pmatrix} &= \begin{pmatrix} G_{0,0}(z) & G_{0,1}(z) & \cdots & G_{0,M-1}(z) \\ G_{1,0}(z) & G_{1,1}(z) & \cdots & G_{1,M-1}(z) \\ \vdots & \vdots & \ddots & \vdots \\ G_{M-1,0}(z) & G_{M-1,1}(z) & \cdots & G_{M-1,M-1}(z) \end{pmatrix} \cdot \begin{pmatrix} Y_0(z) \\ Y_1(z) \\ \vdots \\ Y_{M-1}(z) \end{pmatrix} \\ &= \mathbf{G}(z) \cdot \begin{pmatrix} Y_0(z) \\ Y_1(z) \\ \vdots \\ Y_{M-1}(z) \end{pmatrix}\end{aligned}$$

where $\mathbf{G}(Z)$ is known as the polyphase synthesis matrix. The polyphase synthesis matrix may also be understood as an M-input/M-output LTI system.

Note that the m^{th} row of the polyphase analysis matrix, $\mathbf{H}(Z)$, contains the filters derived from the m^{th} analysis filter, $h_m[n]$, whereas the m^{th} column of the polyphase synthesis matrix, $\mathbf{G}(Z)$, contains the filters derived from the m^{th} synthesis filter, $g_m[n]$. Moreover, the polyphase components of the analysis filters are obtained using slightly different definitions to the

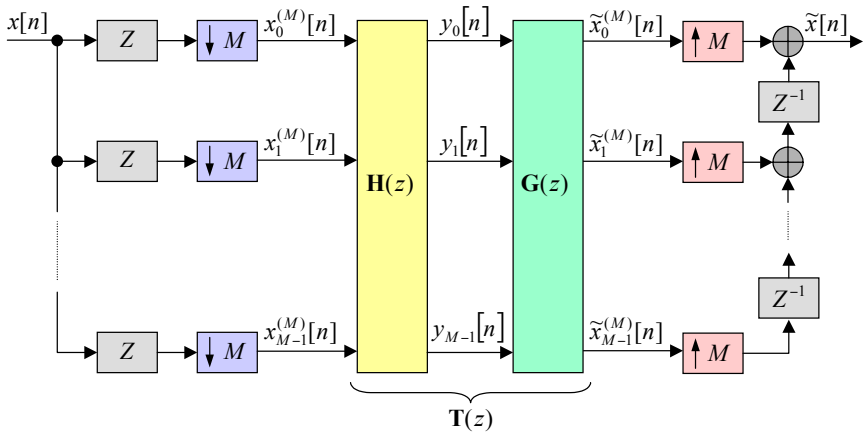

Figure 4: *Polyphase analysis, synthesis and transfer matrices.*

polyphase components of the synthesis filters. Rather than trying to remember these, it is often easiest to derive the equations from scratch each time, as done above.

Evidently, we can now put the analysis and synthesis subband systems together and we find that the input and output polyphase components are related according to

$$\begin{pmatrix} \tilde{X}_0^{(M)}(z) \\ \tilde{X}_1^{(M)}(z) \\ \vdots \\ \tilde{X}_{M-1}^{(M)}(z) \end{pmatrix} = \mathbf{T}(z) \cdot \begin{pmatrix} X_0^{(M)}(z) \\ X_1^{(M)}(z) \\ \vdots \\ X_{M-1}^{(M)}(z) \end{pmatrix}$$

where $\mathbf{T}(z) = \mathbf{G}(z) \cdot \mathbf{H}(z)$ is the known as the polyphase transfer matrix. This system is illustrated in Figure 4.By comparison with the trivial case in which we associated the subbands with the polyphase components (see Figure 3), it is obvious that the condition for perfect reconstruction is $\mathbf{T}(z) = \mathbf{I}$, or equivalently,

$$\mathbf{G}(z) = (\mathbf{H}(z))^{-1}$$

Let us consider the case of an $M = 2$ channel filter bank with polyphase analysis matrix,

$$\mathbf{H}(z) = \begin{pmatrix} H_{0,0}(z) & H_{0,1}(z) \\ H_{1,0}(z) & H_{1,1}(z) \end{pmatrix}$$

We can deduce from the above relationship that the synthesis matrix required for perfect reconstruction is given by

$$\begin{aligned}\mathbf{G}(z) &= \frac{1}{\det(\mathbf{H}(z))}\begin{pmatrix} H_{1,1}(z) & -H_{0,1}(z) \\ -H_{1,0}(z) & H_{0,0}(z)\end{pmatrix} \\ &= \frac{\begin{pmatrix} H_{1,1}(z) & -H_{0,1}(z) \\ -H_{1,0}(z) & H_{0,0}(z)\end{pmatrix}}{H_{0,0}(z)H_{1,1}(z) - H_{0,1}(z)H_{1,0}(z)}\end{aligned} \tag{1}$$

In general, then, if the analysis system consists of FIR filters, the synthesis system required for perfect reconstruction will consist of IIR filters with rational transfer functions. If the analysis and synthesis systems are both to involve FIR filters, the determinant polynomial,

$$\det(\mathbf{H}(z)) = H_{0,0}(z)H_{1,1}(z) - H_{0,1}(z)H_{1,0}(z)$$

must be of the form

$$\det(\mathbf{H}(z)) = \alpha z^d \tag{2}$$

for some integer, d, and scale factor, α.

To conclude this section, we will draw an important connection with our discussion of subband transforms at the end of the last chapter of your lecture notes. Given that perfect reconstruction FIR analysis and synthesis filter banks must satisfy equation (2), we may use equation (1) to deduce that

$$\begin{pmatrix} G_{0,0}(z) & G_{0,1}(z) \\ G_{1,0}(z) & G_{1,1}(z)\end{pmatrix} = \alpha^{-1} z^{-d} \begin{pmatrix} H_{1,1}(z) & -H_{0,1}(z) \\ -H_{1,0}(z) & H_{0,0}(z)\end{pmatrix}$$

Substituting the original filter coefficients for the polyphase components, we get

$$\begin{pmatrix} g_0[2n] & g_1[2n] \\ g_0[2n+1] & g_1[2n+1]\end{pmatrix} = \alpha^{-1} \begin{pmatrix} h_1[2(n-d)-1] & -h_0[2(n-d)-1] \\ -h_1[2(n-d)] & h_0[2(n-d)]\end{pmatrix}$$

which we may rewrite as

$$\begin{aligned} g_0[n] &= \alpha^{-1}(-1)^n h_1[n-1-2d] \\ g_1[n] &= -\alpha^{-1}(-1)^n h_0[n-1-2d]\end{aligned}$$

or, in the Z-domain,

$$\begin{pmatrix} G_0(z) \\ G_1(z)\end{pmatrix} = \alpha^{-1} z^{-(2d+1)} \begin{pmatrix} H_1(-z) \\ -H_0(-z)\end{pmatrix} \tag{3}$$

This is exactly the QMF selection which we made at the end of the previous chapter, as a means of ensuring alias free reconstruction, with $\kappa = 2d+1$.

Example 1 *Consider a two-channel filter bank in which the low-pass analysis filter, $h_0[n]$, has the following transfer function*

$$H_0(z) = 1 + \frac{1}{2}z^{-1} + \frac{1}{4}z^{-2}$$

Find a two tap high-pass analysis filter of the form

$$H_1(z) = 1 + az^{-1}$$

such that the perfect reconstruction synthesis filter bank involves only FIR filters. Find all of the relevant filters.

We begin by splitting $H_0(z)$ and $H_1(z)$ into their polyphase components and writing down the analysis matrix, $\mathbf{H}(z)$. We get

$$\mathbf{H}(z) = \begin{pmatrix} 1 + \frac{1}{4}z^{-1} & \frac{1}{2}z^{-1} \\ 1 & az^{-1} \end{pmatrix}$$

Now for $\det(\mathbf{H}(z))$ to be a monomial, we require

$$\alpha z^d = az^{-1} + \frac{1}{4}az^{-2} - \frac{1}{2}z^{-1}$$

which is solved by setting

$$a = \frac{1}{2}, \text{ in which case } \alpha z^d = \frac{1}{8}z^{-2}$$

Hence we get

$$G(z) = \frac{1}{\alpha z^d}\begin{pmatrix} az^{-1} & -\frac{1}{2}z^{-1} \\ -1 & 1 + \frac{1}{4}z^{-1} \end{pmatrix} = \begin{pmatrix} 4z & -4z \\ -8z^2 & 8z^2 + 2z \end{pmatrix}$$

so that

$$\begin{pmatrix} G_0(z) \\ G_1(z) \end{pmatrix} = \begin{pmatrix} 4z^2 - 8z^3 \\ -4z^2 + 8z^3 + 2z \end{pmatrix}$$

3 An Introductory Design Technique

The above example suggests that it is not all that difficult to find collections of filters, $H_0(z)$, $H_1(z)$, $G_0(z)$ and $G_1(z)$, all FIR, such that the complete subband transform system has the perfect reconstruction property. However, it is quite a different matter to design "good" low- and high-pass subband filters. In the example, we started with a reasonable low-pass filter, $H_0(z)$,

and designed the other three filters, but none of them turned out to be good low- or high-pass filters.

In this section, we describe a simple technique for designing good subband filters. This technique was originally proposed by Smith and Barnwell in 1984 as the first systematic technique for designing QMF filters with the perfect reconstruction property. The technique is based upon the particular QMF selections described at the end of the last chapter, which we reproduce here as

$$\begin{pmatrix} G_0(z) \\ G_1(z) \end{pmatrix} = z^{-1} \begin{pmatrix} H_1(-z) \\ -H_0(-z) \end{pmatrix}$$

and

$$G_0(z) = H_0\left(z^{-1}\right)$$

With this section, we found that G_0 and G_1 are both mirror images of H_0 and H_1, that all filters depend upon the selection of the single filter, $H_0(z)$, and that this filter has only to satisfy the power complementary property,

$$\left|\hat{h}_0(\omega)\right|^2 + \left|\hat{h}_0(\omega-\pi)\right|^2 = 2$$

or, equivalently,

$$H_0(z) H_0\left(z^{-1}\right) + H_0(-z) H_0\left(-z^{-1}\right) = 2$$

Letting $W(z) = H_0(z) H_0\left(z^{-1}\right)$, we see that $W(z)$ must be a zero-phase filter satisfying

$$\begin{aligned} W(z) + W(-z) &= \hat{w}(\omega) + \hat{w}(\omega-\pi) \\ &= \hat{w}(\omega) + \hat{w}(\pi-\omega) = 2 \end{aligned}$$

Smith and Barnwell noted that such a filter, $W(z)$, can be designed using the Parks-McClellan algorithm. In fact, so long as the ripple weighting factor, $\hat{\rho}(\omega)$ is symmetric about $\omega = \pi/2$, the Parks-McClellan algorithm will always design a filter whose pass-band ripples exactly cancel its stop-band ripples, so that

$$\hat{w}(\omega) + \hat{w}(\pi-\omega) = \hat{d}(\omega) + \hat{d}(\pi-\omega)$$

where $\hat{d}(\omega)$ is the desired frequency response supplied to the Parks-McClellan design procedure.

The final design procedure proposed by Smith and Barnwell is as follows:

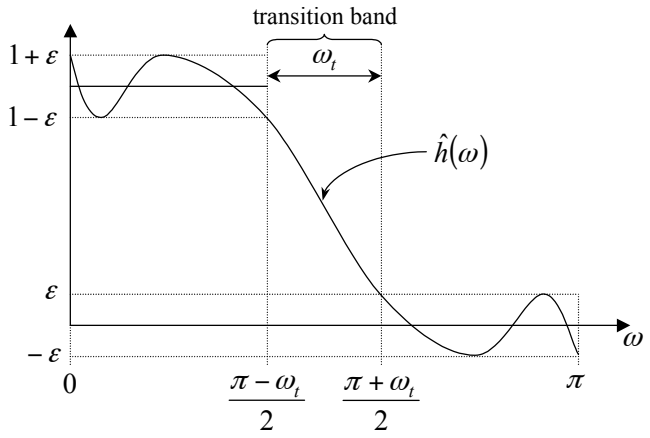

Figure 5: *Equiripple design obtained using the Parks-McClellan algorithm.*

1. Select a transition bandwidth, ω_t and set the ripple weighting function to
$$\hat{\rho}(\omega) = \begin{cases} 0 & \text{if } \frac{\pi}{2} - \frac{\omega_t}{2} < |\omega| < \frac{\pi}{2} + \frac{\omega_t}{2} \\ 1 & \text{otherwise} \end{cases}$$

2. Set the desired frequency response to that of an ideal half-band filter, i.e.,
$$\hat{d}(\omega) = \begin{cases} 1 & \text{if } |\omega| < \frac{\pi}{2} \\ 0 & \text{otherwise} \end{cases}$$

3. Design a zero phase filter, $H(z)$, with even order, $2N$, using the Parks-McClellan algorithm, with the above weighting function and desired frequency response. The FIR filter's impulse response, $h[n]$, has region of support $-N \leq n \leq N$ and the magnitude response has ripples of equal magnitude, ε, in both the pass and stop bands, such that $\hat{h}(\omega) + \hat{h}(\pi - \omega) = 1$. This is illustrated in Figure 5.

4. Since $\hat{w}(\omega) = \left|\hat{h}_0(\omega)\right|^2$, we require it to be strictly positive, but $\hat{h}(\omega)$ has ripples in the stop band which make both positive and negative transitions relative to the desired response of 0. To correct this response, set
$$\hat{w}(\omega) = \frac{2}{1+2\varepsilon}\left(\hat{h}(\omega) + \varepsilon\right)$$

or, equivalently,

$$w[n] = \frac{2}{1+2\varepsilon}\left(h[n] + \varepsilon \cdot \delta[n]\right)$$

Evidently, this filter satisfies the required properties

$$\begin{aligned} \hat{w}(\omega) &\geq 0, \quad \forall \omega \\ \hat{w}(\omega) + \hat{w}(\pi - \omega) &= 2, \quad \forall \omega \end{aligned}$$

5. Finally, use spectral factorization to find the N^{th} order filter, $H_0(z)$, such that

$$W(z) = H_0(z) H_0(-z)$$

The Smith Barnwell design procedure, described above, has one major drawback. Recall from our discussion of filter design techniques that a good low-pass filter design generally requires us to select a weighting function which is much larger in the stop band than in the pass band, so that ripples are suppressed more in the stop band. The Smith Barnwell procedure does not allow us to do this. Worse still, ripples in the stop band of the square-root filter, $\hat{h}_0(\omega)$, will be much larger in the stop band than in the pass band. For this reason, good low-pass filters generally require us to select very high orders, N.