Chapter 9

M-Channel Filter Banks

9.1 Freedom versus Structure

An M-channel filter bank has very considerable freedom for $M > 2$. That is true when we require perfect reconstruction. It is still true when we also require orthogonality. Those properties can be expressed directly as conditions on the $M \times M$ analysis polyphase matrix $\boldsymbol{H}_p(z)$:

1. Perfect reconstruction requires $\boldsymbol{H}_p(z)$ to be invertible for all z.

2. The synthesis filters are FIR when the determinant of $\boldsymbol{H}_p(z)$ is a delay z^{-l}.

3. The filter bank is orthogonal when $\boldsymbol{H}_p^T(z^{-1})\boldsymbol{H}_p(z) = \boldsymbol{I}$.

In the orthogonal case, each synthesis filter F_k is the transpose (or flip, or time-reversal) of the analysis filter H_k. Then delays make F_k causal.

How to design a good M-band filter bank? For $M = 2$, orthogonality is already restrictive (and effectively prevents linear phase). The PR condition is less strict; linear phase can be achieved. In both cases the product filters $F_0(z)H_0(z)$ and $F_1(z)H_1(z)$ are halfband. Furthermore, one comes from the other by alternating signs. All four filters are created by factoring *one halfband filter*.

For $M > 2$, the number of freedoms grows faster than the number of restrictions. The choice of one M-th band filter (or one lowpass filter $H_0(z)$) does not determine the other choices. This is attractive at first — we can have linear phase with orthogonality. But decisions are still needed. We can make those decisions late or early! The range of designs can be left very wide (late decision). Or we can restrict the filter bank to have a structure that we know is desirable (early decision, simplifying the design). This chapter studies both possibilities. In all cases we absolutely want *fast implementation*.

In practice, fast algorithms come by cascading simple functions. At the lowest level they are based on 2×2 butterflies. At a higher level three structures are quick:

Rotations and delays or ***DFT banks*** or ***DCT banks***.

All constant orthogonal matrices are products of $M(M-1)/2$ plane rotations. All paraunitary matrices are products of rotations and delays. (Householder would use reflections and delays — this factorization is an important theorem.) Under quantization and roundoff, an orthogonal butterfly remains orthogonal. The difficulty will be the large number of rotation angles, approximately $NM(M-1)/2$ for filters of lengths N. An optimal design may have excellent properties but it will not be easy to find.

This chapter will pursue all three structures. We develop the polyphase approach in Section 9.2 and the time domain approach (to LOT and GenLOT) in Section 9.3. Those sections extend to M channels the earlier ideas from two channels. Then Section 9.4 focuses on cosine modulation — not seen for two channels but now important.

It may assist the reader if we now briefly highlight DFT and DCT filter banks. These are *modulated* filter banks because all M filters come from frequency shifts of one prototype filter. It is usual to refer to DCT filter banks as *cosine-modulated filter banks*. In many applications, the DFT loses and the DCT wins.

Block Transforms

The simplest filter banks use only the M-point DFT or the M-point DCT. The signal is split into blocks of length M. These blocks are *separately* transformed. There is no overlapping or interaction between blocks, and no filter design is involved. The polyphase matrix can be the Fourier DFT matrix or the DCT matrix (this is the JPEG standard). In a block transform, $\boldsymbol{H}_p$ is just a constant matrix — and there is no smoothing between blocks.

The analysis half can be drawn with the modulators first or the decimators first. Figure 9.1 shows the direct form of the block DFT and the more efficient polyphase form ($W_M = e^{-j2\pi/M}$). Section 9.2 discusses the block generation step in detail, because a serial to parallel S/P converter is extremely important — as a way to start the filter bank.

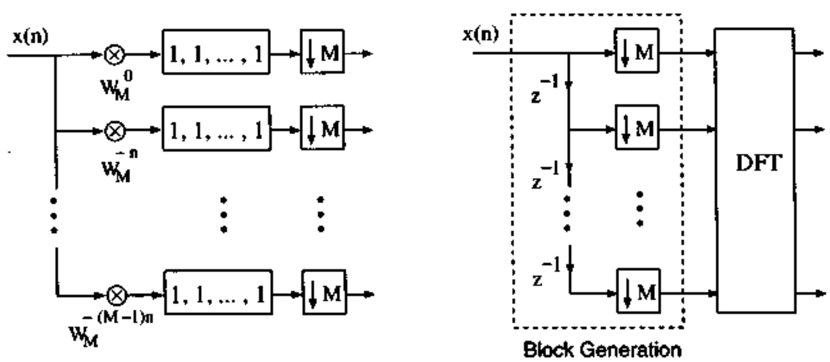


Figure 9.1: The DFT block transform: M samples at a time. Using DCT gives JPEG.

DFT Filter Banks

Now we include filters. The two-channel DFT filter bank appeared early in Chapter 4 (not with that name). The highpass frequency response shifted $H_0(e^{j\omega})$ by π. Consequently, $\boldsymbol{h}_1(n)$ was the ***alternating sign*** version of $\boldsymbol{h}_0(n)$. The reader will remember that $H_1(z) = H_0(-z)$ did not go well with perfect reconstruction. For two channels, the best arrangement is to alternate signs between H_0 and F_1 and between H_1 and F_0. This is not DFT.

For an M-channel DFT bank, the frequency response from H_k is $H_0(e^{j\omega})$ shifted by $\frac{2\pi k}{M}$. The corresponding z-transform and time-domain relations are

$$H_k(z) = H_0(ze^{-j2\pi k/M}) \quad \text{and} \quad \boldsymbol{h}_k(n) = \boldsymbol{h}_0(n)e^{j2\pi k/M}.$$

The synthesis filters are modulated in the same way. Note that the coefficients become *complex* for $M > 2$. This is a disadvantage. The great advantage is the simplicity of design and the speed of implementation, when the whole analysis bank is based on one filter H_0 (and the DFT).

Figure 9.2 shows the direct form and polyphase form of the DFT analysis bank. It differs from Figure 9.1 only by including filters — which come before $(\downarrow M)$ in the direct form and after $(\downarrow M)$ in the polyphase form. We are free to use the DFT or the IDFT in analysis, and reverse this choice in synthesis. The transfer functions $E_k(z)$ are the polyphase components of $H_0(z)$.

Figure 9.2 reduces to Figure 9.1 when $\boldsymbol{h}_0 = (1, 1, \ldots, 1)$. Then the product of DFT and IDFT gives perfect reconstruction trivially, a block at a time. Section 9.2 derives the PR condition for a DFT bank that involves filters. The requirements on those filters are quite restrictive. DFT banks are generally superseded by filter banks based on the DCT, if reconstruction is desired.

We saw the same for continuous time in Chapter 8. *Cosine modulation is the best.*

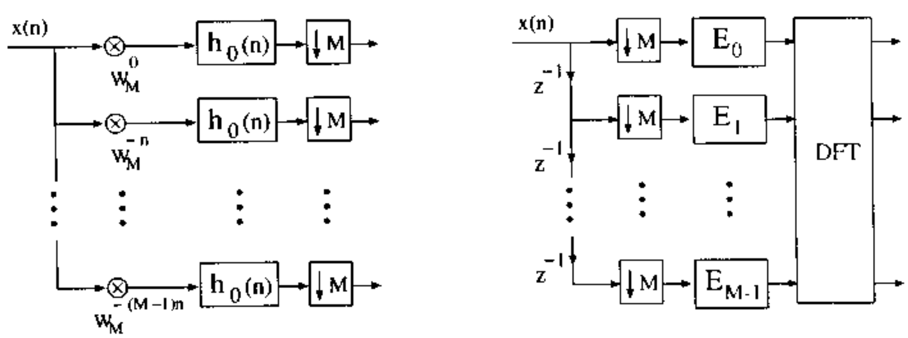


Figure 9.2: Direct form and polyphase form: DFT bank built from one filter.

Cosine-modulated Filter Banks

Cosine modulation replaces the complex DFT by the real DCT. Remember that there are four different cosine transforms. They use different extensions; Types II and IV are appropriate for filter banks. (Types I and III give undesirable bandwidths for the first and last filters. For the same reason, the filter length is constrained to be $2M$ or more generally $2KM$.) We use Type IV with the factors $k + \frac{1}{2}$ and $n + \frac{1}{2}$ in the frequencies.

The DCT matrix C^{IV} is symmetric and orthogonal. We shift the frequencies to achieve perfect reconstruction, starting from the lowpass prototype $\boldsymbol{p}(n)$:

$$\boldsymbol{h}_k(n) = \boldsymbol{f}_k(n) = \boldsymbol{p}(n)\sqrt{\frac{2}{M}}\cos\left[\left(k + \tfrac{1}{2}\right)\left(n + \frac{M+1}{2}\right)\frac{\pi}{M}\right]. \qquad (9.3)$$

The shift of n by $\frac{M}{2}$ has the effect of centering $\boldsymbol{p}$. That simplifies the conditions on this prototype filter — "*the window*" — to achieve perfect reconstruction. The PR conditions are beautiful for filter length $2M$. We write them now without proof:

$$\textbf{\textit{Even symmetry}} \quad p(n) = p(N-n) \tag{9.2}$$

$$\textbf{\textit{Orthogonality}} \quad p^2(n) + p^2(n+M) = 1. \tag{9.3}$$

These are *precisely analogous* to the conditions on the continuous-time cosines in Chapter 8. There the window was subject to $g^2(t) + g^2(t+1) = 1$. Section 9.4 will extend the discrete-time requirements to filter lengths $2KM$, and establish orthogonality and perfect reconstruction. There are analogous conditions in continuous time — when each window overlaps several other windows.

We can already see the major advantages of cosine modulation:

- Simplicity of design: one filter $p(n)$ only
- Symmetry and orthogonality
- Very fast implementation.

The simplicity is crucial when M is large. That is the central point — to accept and indeed to welcome restrictions that simplify the structure. The implementation is fast because the DCT is fast. Ultimately this is because the FFT is fast.

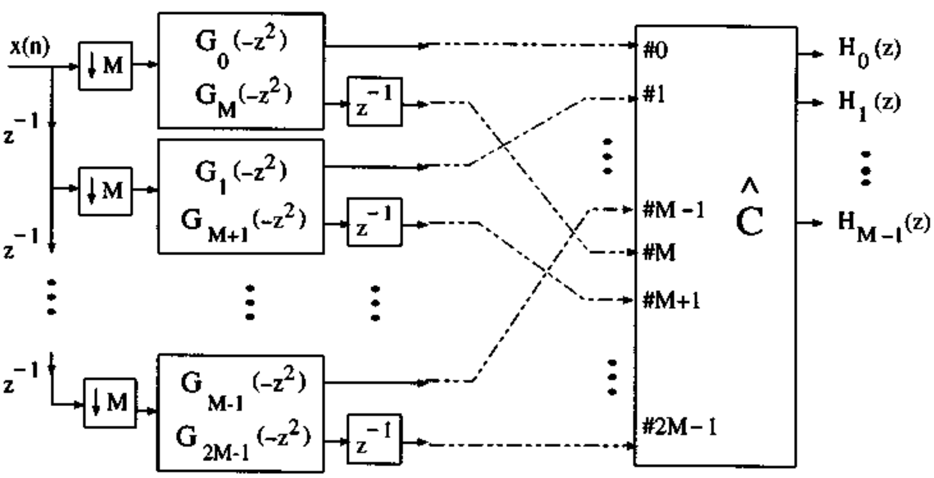


Figure 9.3: Cosine modulation by the DCT: k and $M+k$ are twins.

Remember that the M-point DCT is created from the $2M$-point DFT. In our applications the DCT has real outputs, reducing the full (complex) computation by half. The analysis bank will involve

- $2M$ polyphase filters
- $2M$ complex modulators with real inputs
- one $2M$-point DFT with real outputs.

Figure 9.3 is directly based on the DCT, and emphasizes how phases k and $M+k$ are twinned. Each pair of phases is like a 2-channel orthogonal filter bank! (On which the reader is by now an expert.) The good implementations have a delay chain of length $2M$ and decimators $(\downarrow M)$

coming first. They generate *two* blocks of length M in Figure 9.3, as the double-length DFT requires.

The cosine-modulated basis functions are *not* linear phase. They are even around $\omega = 0$ and odd around $\omega = \pi$, which produces Type-IV orthogonality. The first four basis functions when $M = 32$ are drawn in Section 8.3.

Figure 9.4a shows the idealized frequency response $P(e^{j\omega})$ of the lowpass prototype filter — the window $\boldsymbol{p}$. Then Figure 9.4b shows the phase shifts from cosine modulation. Notice especially how each cosine (the sum of two exponentials) produces two copies of $P(e^{j\omega})$. One copy is shifted left and one copy is shifted right. A good prototype will have passband approximately $|\omega| \le \frac{\pi}{2M}$. The design of that symmetric window subject to (9.2) gives a good M-channel filter bank — structured by cosine modulation.

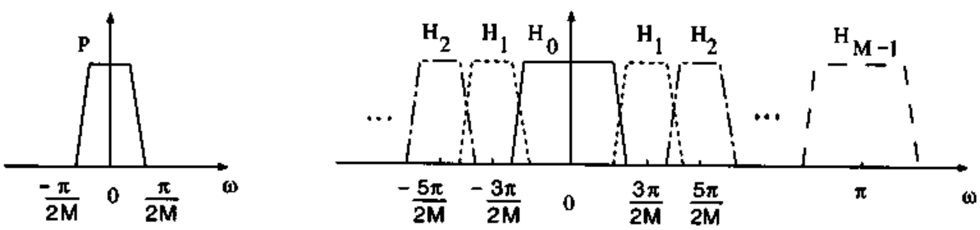


Figure 9.4: Idealized frequency response of the prototype and the cosine-modulated filters.

The original references for the discrete-time filter bank are in [Mr]. The key to the continuous-time construction was found by [Coifman-Meyer]. More recent references are in Section 9.4, where we also pursue the possibility of *biorthogonality*. In that case the analysis window $\widetilde{p}(n)$ is dual to the synthesis window $p(n)$:

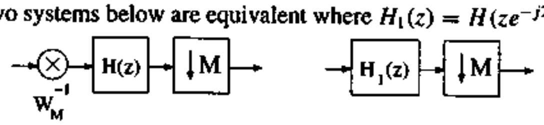
$$\widetilde{p}(n) = \frac{p(n)}{p^2(n) + p^2(n+M)}.$$

This corresponds to $\widetilde{g}(t) = g(t)/(g^2(t) + g^2(t+1))$ in continuous time.

It seems fair to say that the discrete-time construction is always a little more delicate. For orthogonality, we check sums instead of integrals. The sums often depend on equal spacing in time or frequency. Where time-varying windows and window lengths were no problem in Section 8.4, they are more difficult for discrete filters. But their potential value is so great that the effort is worth making.

Problem Set 9.1

1. Draw the DCT basis functions h_0, h_1, h_2, h_3 when $M = 8$. They are drawn in Section 8.3 for $M = 32$.
2. Suppose the prototype p in Figure 9.4 is an ideal brick wall. What are the coefficients $p(n)$ and what is $P(\omega)$?
3. Invent a prototype $p(n)$ that satisfies the PR conditions (9.2) and (9.3), with $M = 2$ and then general M. The filter has $N + 1$ coefficients.
4. Verify that the two systems below are equivalent where $H_1(z) = H(ze^{-j2\pi/M})$.



9.2 Polyphase Form: *M* Channels

Digital filter banks divide the signal into M subbands, then process the subbands and reconstruct. The analysis bank splits the input signal and the synthesis bank recombines it. The essential information is extracted from the subband signals in the *processing block*. Its form varies and depends on the applications. In an audio/video compression system, the spectral contents of the subband signals are coded depending on their energies. In a radar system, the subband signals might be used to null out a narrow-band interference adaptively. Other applications are image analysis and enhancement, robotics, computer vision, echo-cancellation, voice privacy and communications.

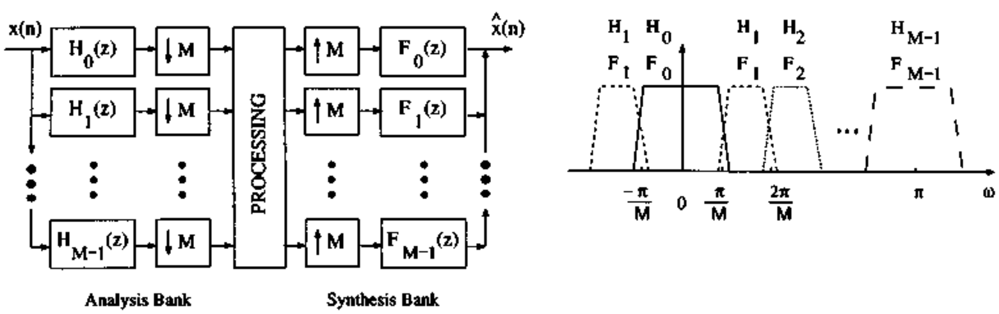


Figure 9.5: A maximally-decimated uniform filter bank with M channels.

Figure 9.5 illustrates a maximally decimated M-channel filter bank. The frequency responses $H_k(e^{j\omega})$ and $F_k(e^{j\omega})$ have passbands as shown. The analysis filters $H_k(z)$ channelize the input signal $x(n)$ into M subband signals, which are downsampled (decimated) by a factor M. At the receiving end, the M subband signals are decoded, interpolated and recombined by the synthesis filters $F_k(z)$. The decimator, which decreases the sampling rate, and the expander, which increases the rate, are denoted by $(\downarrow M)$ and $(\uparrow M)$.

Reconstruction Error

Since the filters $H_k(z)$ are not ideal, the filtered signals are not bandlimited to π/M. Therefore the aliases from downsampling will overlap. They depend on the stopband attenuation of $H_k(e^{j\omega})$ and their transition bands. The interpolated (upsampled) signals have M images of the compressed spectrum (assuming that no processing has been done after the analysis bank). These images are filtered by the synthesis filters $F_k(z)$.

There are two types of errors in the reconstructed output signal $\hat{x}(n)$: *distortions* (magnitude and phase) and *aliasing*. The nonideal filtering characteristics of $H_k(z)$ and $F_k(z)$ contribute to both distortion and aliasing. The changes in sampling rates (downsampling and upsampling) contribute to the aliasing error. A system with no aliasing error is "alias-free". We compute the z-transform relation between input and output for $M = 3$:

- The analysis filters produce $H_0(z)X(z)$ and $H_1(z)X(z)$ and $H_2(z)X(z)$.
- $(\downarrow 3)$ and $(\uparrow 3)$ then produce two aliases in each channel $k = 0, 1, 2$:

$$\tfrac{1}{3}\left(H_k(z)X(z) + H_k(zW)X(zW) + H_k(zW^2)X(zW^2)\right) \quad \text{with } W = e^{-j2\pi/3}.$$

- The synthesis filters multiply those z-transforms by $F_k(z)$. Adding the nine terms, and collecting the terms for $X(z)$ and its aliases $X(zW)$ and $X(zW^2)$, the total output from the filter bank is

$$\widehat{X}(z) = T_0(z)X(z) + T_1(z)X(zW) + T_2(z)X(zW^2). \tag{9.4}$$

The perfect reconstruction conditions are $T_0(z) = z^{-\ell}$ (no amplitude distortion) and $T_1(z) = T_2(z) = 0$ (no aliasing). These three equations are written out explicitly in (9.10) below. For general M, the input-output relation is

$$\widehat{X}(z) = \sum_{k=0}^{M-1} T_k(z)X(zW^k) \qquad \text{where} \qquad T_k(z) = \frac{1}{M}\sum_{\ell=0}^{M-1} F_\ell(z)H_\ell(zW^k). \tag{9.5}$$

Since the transfer functions $T_1(z)$, $T_2(z)$, $\ldots$, $T_{M-1}(z)$ multiply the shifted versions of the input spectrum, they are the *aliasing transfer functions. The distortion function* $T_0(z)$ multiplies the original spectrum. Then $T_0(z)X(z)$ is the output when all aliasing is cancelled. The objective is to find a set of PR filters $H_k(z)$ and $F_k(z)$:

$$\textit{Perfect Reconstruction:} \quad \begin{cases} T_0(z) = z^{-\ell} & \text{(no distortion for } k = 0) \\ T_k(z) = 0 & \text{(alias-free for } 1 \le k < M). \end{cases} \tag{9.6}$$

Paraunitary and biorthogonal filter banks (with additional properties such as linear phase and cosine modulation) satisfy all conditions in (9.6). The conventional Pseudo-QMF bank cancels aliasing at adjacent bands ($T_1(z) = 0$). The "Near PR" banks have $T_0(z) = z^{-\ell}$ and $T_k(z)$ near zero. The DFT filter bank cancels all aliasing components, but suffers from distortions! It satisfies the last $M - 1$ conditions in (9.6) but not the first. In discussing a specific M-channel filter bank, one has to keep in mind its reconstruction properties.

Note: For a two-channel filter bank, perfect reconstruction requires

$$\begin{cases} T_0(z) = \frac{1}{2}[F_0(z)H_0(z) + F_1(z)H_1(z)] = z^{-\ell} \\ T_1(z) = \frac{1}{2}[F_0(z)H_0(-z) + F_1(z)H_1(-z)] = 0. \end{cases} \tag{9.7}$$

Solving for $F_k(z)$ yields the synthesis filters from the analysis filters:

$$F_0(z) = \frac{2z^{-\ell}}{\Delta(z)}H_1(-z) \quad \text{and} \quad F_1(z) = -\frac{2z^{-\ell}}{\Delta(z)}H_0(-z) \tag{9.8}$$

where $\Delta(z) = H_0(z)H_1(-z) - H_0(-z)H_1(z)$. An FIR system has ℓ delays:

$$\text{Determinant} \quad \Delta(z) = 2z^{-\ell}, \quad F_0(z) = H_1(-z), \quad F_1(z) = -H_0(-z). \tag{9.9}$$

This simple relationship ***does not extend*** to filter banks with more channels. The perfect reconstruction equations for a three-channel bank are

$$\begin{cases} T_0(z) = F_0(z)H_0(z) + F_1(z)H_1(z) + F_2(z)H_2(z) = 3z^{-\ell} \\ T_1(z) = F_0(z)H_0(zW) + F_1(z)H_1(zW) + F_2(z)H_2(zW) = 0 \\ T_2(z) = F_0(z)H_0(zW^2) + F_1(z)H_1(zW^2) + F_2(z)H_2(zW^2) = 0 \end{cases} \tag{9.10}$$

where $W = e^{-j2\pi/3}$. Writing this system in matrix form, the solution for $F_k(z)$ is

$$\begin{bmatrix} F_0(z) \\ F_1(z) \\ F_2(z) \end{bmatrix} = \begin{bmatrix} H_0(z) & H_1(z) & H_2(z) \\ H_0(zW) & H_1(zW) & H_2(zW) \\ H_0(zW^2) & H_1(zW^2) & H_2(zW^2) \end{bmatrix}^{-1} \begin{bmatrix} 3z^{-\ell} \\ 0 \\ 0 \end{bmatrix}. \quad (9.11)$$

This inverse transpose of the modulation matrix $\boldsymbol{H}_m(z)$ gives

$$\begin{bmatrix} F_0(z) \\ F_1(z) \\ F_2(z) \end{bmatrix} = \frac{3z^{-\ell}}{\Delta(z)} \begin{bmatrix} H_1(zW)H_2(zW^2) - H_1(zW^2)H_2(zW) \\ H_2(zW)H_0(zW^2) - H_2(zW^2)H_0(zW) \\ H_0(zW)H_1(zW^2) - H_0(zW^2)H_1(zW) \end{bmatrix} \quad (9.12)$$

where $\Delta(z)$ is the determinant of $\boldsymbol{H}_m(z)$. The synthesis filter $F_k(z)$ depends on *two* filters $H_j(z)$ for $j \neq k$. This relationship complicates the design for M channels.

The number of parameters grows linearly with M. To simplify the design and implementation, explicit relations between the filters are often imposed:

Paraunitary	Synthesis is time-reversed from analysis: $F_k(z) = z^{-N}H_k(z^{-1})$.
Linear Phase	$H_k(z) = \pm z^{-N}H_k(z^{-1})$ (symmetric or antisymmetric).
DFT Filter Bank	$H_k(z)$ comes from $H_0(zW^k)$.
Cosine Modulation	$H_k(z)$ and $F_k(z)$ are DCT modulations of ***one*** prototype.
Pairwise Mirror Image	$H_{M-1-k}(z) = H_k(-z)$ for odd M and $z^{-N}H_k(-z^{-1})$ for even M. The frequency responses are symmetric about $\frac{\pi}{2}$.

Each relation reduces the number of parameters by 2 (or by M, for DFT and cosine modulation). But the requirements might prevent our design methods from finding a good solution. The only two-channel linear-phase paraunitary filter bank is Haar's $H_0(z) = 1+z^{-1}$ and $H_1(z) = 1-z^{-1}$. By imposing both properties, we have limited our solution.

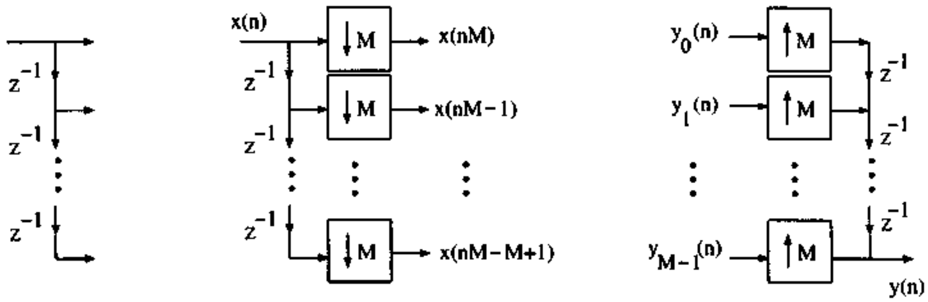
The choices depend on the applications. Cosine modulation is used in audio compression and telecommunications. It is efficient and has high stopband attenuation. Linear phase is important in image compression and signal detection.

Before developing filter banks for $M > 2$, we generalize the delay chain, the serial-parallel and parallel-serial converter, and the polyphase form.

Serial-Parallel and Parallel-Serial Converters

The figure below shows the block diagram of a delay chain. The transfer function from the input to the kth output is z^{-k}. By itself, the delay chain is not very interesting. However, cascading the delay chain with downsamplers or upsamplers will yield serial-parallel or parallel-serial converters.

Delay chain cascaded with downsamplers	=	Serial to Parallel (S/P) converter
Upsamplers cascaded with delay chain	=	Parallel to Serial (P/S) converter



Delay chain Serial to parallel converter Parallel to serial converter

The output at the kth branch of the S/P converter is $x(nM-k)$, which implies that the input sequence is selected in a *counter-clockwise* fashion. The branches select signals in the order $0,\ M-1,\ M-2,\ \ldots,\ 2,\ 1,\ 0,\ M-1,\ldots$. When $x(n)$ is a causal signal, the output of the S/P converter ($M=4$) is:

branch	0	$x(0)$	$x(4)$	$x(8)$	$x(12)$	$\cdots$
	1	0	$x(3)$	$x(7)$	$x(11)$	$\cdots$
	2	0	$x(2)$	$x(6)$	$x(10)$	$\cdots$
	3	0	$x(1)$	$x(5)$	$x(9)$	$\cdots$

output rate = (input rate)/M

A P/S converter is a cascade of expanders and a *reverse-ordered* delay chain. The output $y(n)$ is an interleaved combination of the signals $y_k(n)$. Thus its rate is M times the rate of $y_k(n)$. The signals $y_k(n)$ are selected in a *clockwise* fashion. Assuming that $y_k(n)$ are causal, and $M=3$, the serial output is

$$y_2(0)\ y_1(0)\ y_0(0)\ y_2(1)\ y_1(1)\ y_0(1)\ y_2(2)\ y_1(2)\ y_0(2)\ \cdots$$

Polyphase Representation of a Filter Bank

An analysis filter $H_k(z)$ is the sum of M phases $H_{kl}(z)$:

$$H_k(z)=\sum_{\ell=0}^{M-1} z^{-\ell}\, H_{k,\ell}(z^M), \qquad h_{k,\ell}(n)=h_k(Mn+\ell). \tag{9.13}$$

The four phases ($M=4$) of $1+3z^{-1}-4z^{-2}+7z^{-3}+6z^{-4}-3z^{-5}+z^{-6}$ are

$$\left\{\begin{array}{l} H_{k,0}(z)=1+6z^{-1} \\ H_{k,1}(z)=3-3z^{-1}\end{array}\right. \qquad \left\{\begin{array}{l} H_{k,2}(z)=-4+z^{-1} \\ H_{k,3}(z)=7.\end{array}\right. \tag{9.14}$$

Consider a lowpass $H(z)$ of length 81. Figure 9.6(a) shows $|H(e^{j\omega})|$ with center frequency at $\frac{\pi}{4}$. The magnitudes (nearly constant) and phases (nearly linear) of the four components $H(z)$ are plotted in 9.6(b) and (c). The phase responses are such that $|H|\approx 1$ in the passband and $H\approx 0$ in the stopband. Since magnitudes are approximately equal, the phase angles accomplish this task. Figure 9.6(d) plots the offsets $\Delta_\ell(\omega)=-10\omega-\phi_\ell(\omega)$. We observe that $\Delta_\ell(\omega)$ is nearly $-\ell\omega/4$ (in general $-\ell\omega/M$). The polyphase components provide *fractional delays* so that $H(z)$ is a good lowpass filter.

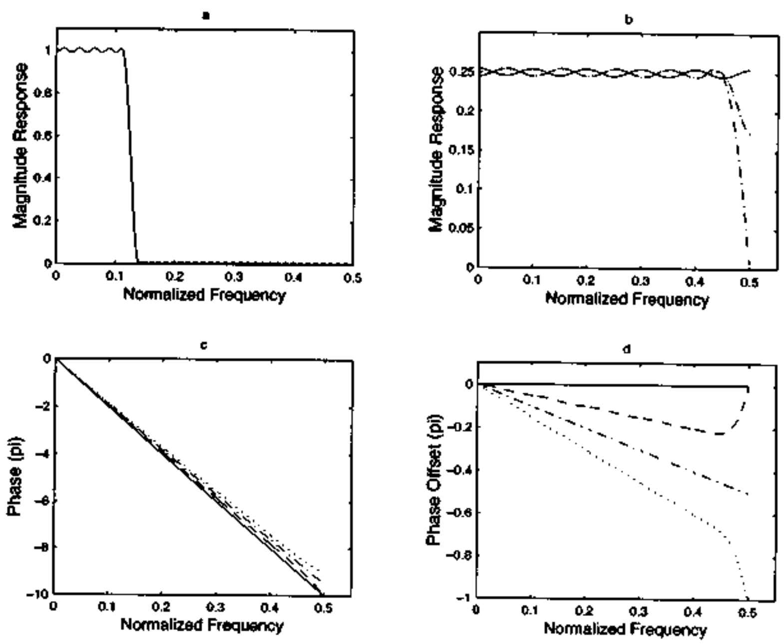


Figure 9.6: (a) Magnitude response of a lowpass filter, (b) Magnitude response of the polyphase filters, (c) Phase responses of the polyphase filters, (d) Phase offset.

There are M^2 polyphase components since each $H_k(z)$ has M components. Grouping these components in row k, the analysis transfer function is

$$\begin{bmatrix} H_0(z) \\ H_1(z) \\ \vdots \\ H_{M-1}(z) \end{bmatrix} = \underbrace{\begin{bmatrix} H_{0,0}(z^M) & H_{0,1}(z^M) & \cdots & H_{0,M-1}(z^M) \\ H_{1,0}(z^M) & H_{1,1}(z^M) & \cdots & H_{1,M-1}(z^M) \\ \vdots & \vdots & \ddots & \vdots \\ H_{M-1,0}(z^M) & H_{M-1,1}(z^M) & \cdots & H_{M-1,M-1}(z^M) \end{bmatrix}}_{\boldsymbol{H}_p(z^M)} \begin{bmatrix} 1 \\ z^{-1} \\ \vdots \\ z^{-(M-1)} \end{bmatrix}. \tag{9.15}$$

$\boldsymbol{H}_p(z)$ is the **polyphase matrix** of the analysis bank. The mapping between $\boldsymbol{h}_k(n)$ and $\boldsymbol{H}_p(z)$ is one to one. The impulse response of $H_{k,\ell}(z)$ is $\boldsymbol{h}_{k,\ell}(n) = \boldsymbol{h}_k(Mn + \ell)$.

Example 9.1. Suppose that the analysis filters of a 3-channel filter bank are

$$\begin{aligned} H_0(z) & & 1 + 2z^{-1} + 3z^{-2} + 4z^{-3} + 5z^{-4} + 6z^{-5} + 7z^{-6} \\ H_1(z) &= & 1 - 2z^{-1} + 3z^{-2} - 4z^{-3} + 5z^{-4} - 6z^{-5} + 7z^{-6} \\ H_2(z) & & 1 + 2z^{-1} - 3z^{-2} + 4z^{-3} + 5z^{-4} - 6z^{-5} + 7z^{-6} \end{aligned}$$

The corresponding polyphase transfer matrix is

$$\boldsymbol{H}_p(z) = \begin{bmatrix} 1 + 4z^{-1} + 7z^{-2} & 2 + 5z^{-1} & 3 + 6z^{-1} \\ 1 - 4z^{-1} + 7z^{-2} & -2 + 5z^{-1} & 3 - 6z^{-1} \\ 1 + 4z^{-1} + 7z^{-2} & 2 + 5z^{-1} & -3 - 6z^{-1} \end{bmatrix}.$$

The synthesis filter has a Type-II representation (phase ℓ before channel k):

$$F_k(z) = \sum_{\ell=0}^{M-1} z^{-(M-1-\ell)} F_{\ell,k}(z^M) \qquad f_{\ell,k}(n) = f_k(Mn + M - 1 - \ell). \tag{9.16}$$

For $F_k(z) = 1 + 3z^{-1} - 4z^{-2} + 7z^{-3} + 6z^{-4} - 3z^{-5} + z^{-6}$, the four components that enter column k of the Type-II polyphase matrix $\boldsymbol{F}_p(z)$ are

$$F_{0,k}(z) = 7 \quad F_{1,k}(z) = -4 + z^{-1} \quad F_{2,k}(z) = 3 - 3z^{-1} \quad F_{3,k}(z) = 1 + 6z^{-1}.$$

The corresponding synthesis bank can be rewritten as

$$\begin{bmatrix} F_0(z) \\ F_1(z) \\ \vdots \\ F_{M-1}(z) \end{bmatrix} = \underbrace{\begin{bmatrix} F_{0,0}(z^M) & F_{1,0}(z^M) & \cdots & F_{M-1,0}(z^M) \\ F_{0,1}(z^M) & F_{1,1}(z^M) & \cdots & F_{M-1,1}(z^M) \\ \vdots & \vdots & \ddots & \vdots \\ F_{0,M-1}(z^M) & F_{1,M-1}(z^M) & \cdots & F_{M-1,M-1}(z^M) \end{bmatrix}}_{\boldsymbol{F}_p^T(z^M)} \begin{bmatrix} z^{-(M-1)} \\ z^{-(M-2)} \\ \vdots \\ 1 \end{bmatrix}. \tag{9.17}$$

Transposing, $\boldsymbol{F}_p(z)$ is the polyphase transfer matrix of the synthesis bank. Using $\boldsymbol{H}_p(z)$ (Type-I polyphase) and $\boldsymbol{F}_p(z)$ (Type-II polyphase), one can redraw Figure 9.5(a) as Figure 9.7(a). Then decimators move to the left of $\boldsymbol{H}_p(z^M)$ by the *Noble Identity*. Similarly the expanders move to the right of $\boldsymbol{F}_p(z^M)$.

A few words on the implementation efficiency of Figure 9.7(b). The input is blocked into M vectors by a Serial/Parallel converter (implemented as cascade of delay chain and decimators). The blocks are filtered by $\boldsymbol{F}_p(z)\boldsymbol{H}_p(z)$ and then recombined using a Parallel/Serial converter. The total number of nonzero coefficients in $\boldsymbol{H}_p(z)$ and $\boldsymbol{F}_p(z)$ is the same as that in $H_k(z)$ and $F_k(z)$. The main difference is a more efficient rate of operation. The filtering in the polyphase form is done at the input rate divided by M.

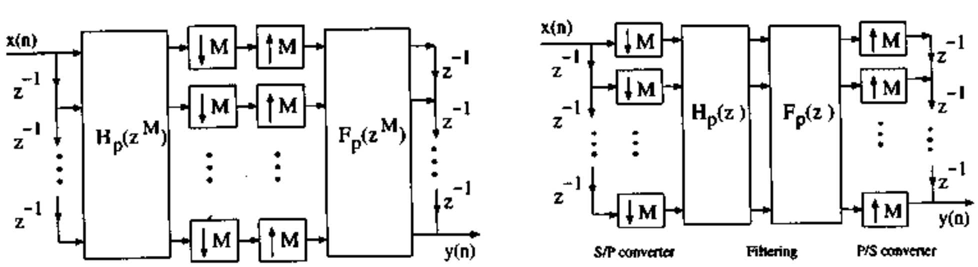


Figure 9.7: Polyphase representations of an M-channel uniform filter bank.

Modulation Matrix

Writing the PR conditions (9.6) in matrix-vector form displays the aliasing component matrix, which is the transposed modulation matrix $H_m^T(z)$:

$$\begin{bmatrix} H_0(z) & H_1(z) & \cdots & H_{M-1}(z) \\ H_0(zW) & H_1(zW) & \cdots & H_{M-1}(zW) \\ \vdots & \vdots & \ddots & \vdots \\ H_0(zW^{M-1}) & H_1(zW^{M-1}) & \cdots & H_{M-1}(zW^{M-1}) \end{bmatrix} \begin{bmatrix} F_0(z) \\ F_1(z) \\ \vdots \\ F_{M-1}(z) \end{bmatrix} = M \begin{bmatrix} T_0(z) \\ T_1(z) \\ \vdots \\ T_{M-1}(z) \end{bmatrix}$$

The term ***modulation*** comes from the fact that row k of H_m is obtained by modulating row zero (shift in frequency by $2\pi k/M$). The term ***aliasing component*** comes from the fact that row k of H_m^T determines the aliasing transfer function $T_k(z)$—which should be zero for $k > 0$. The theory of filter banks can be derived using either $H_m(z)$ or $H_p(z)$. Although they are equivalent, $H_p(z)$ is preferred because it is used in the implementation.

To connect $H_m(z)$ to $H_p(z)$ we need two matrices. One is the diagonal delay matrix $D(z) = diag(1, z^{-1}, \ldots, z^{-(M-1)})$. The other is the M-point DFT matrix F_M which gives the modulations:

$$\text{Modulation and Polyphase} \quad H_m(z) = H_p(z^M)D(z)F_M. \tag{9.18}$$

Proof. The first row of $H_m(z)$ contains the responses $H_0(zW^k)$ of the first filter. The first row of $H_p(z^M)$ contains the phases H_{0j} of that same filter. To assemble phases of any function, we multiply by delays:

$$\begin{bmatrix} H_0(z) & H_0(zW) & \cdots \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix} = \begin{bmatrix} H_{00}(z^M) & H_{01}(z^M) & \cdots \\ \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} 1 & 1 & \cdots \\ z^{-1} & (zW)^{-1} & \cdots \\ \vdots & \vdots & \vdots \end{bmatrix}$$

Factoring the delays $1, z^{-1}, \ldots, z^{-(M-1)}$ from the rows of the last matrix leaves F_M. The whole matrix multiplication becomes exactly $H_m(z) = H_p(z^M)D(z)F_M$.

The diagonal delay matrix is clearly paraunitary. The Fourier matrix $F_M/\sqrt{M}$ is unitary (thus paraunitary, but complex). Then (9.18) shows that $H_p(z)$ is paraunitary when $H_m(z)/\sqrt{M}$ is paraunitary.

Theorem 9.1 *The equivalent conditions in the z-domain for an orthogonal M-channel filter bank are*

1. *The polyphase matrix is paraunitary:* $H_p^T(z^{-1})H_p(z) = I$.

2. *The modulation matrix divided by* $\sqrt{M}$ *is paraunitary:* $\overline{H}_m^T(z^{-1})H_m(z) = MI$.

It is satisfying that each condition has a direct proof, on its own. The direct proof for polyphase is to see the analysis bank ending with $H_p(z)$, and the synthesis bank starting with $H_p^T(z^{-1})$. That meeting produces I at the center of the filter bank. In polyphase form, the whole bank generates blocks of M samples, filters with I, and reconstructs the signal from the blocks. This is perfect reconstruction.

The direct proof of the modulation condition comes from following each channel instead of each phase. When we did this with synthesis filters F_k, the perfect reconstruction condition $T_k(z) = z^{-l}\delta(k)$ was

$$[F_0(z) \;\; F_1(z) \;\cdots\; F_{M-1}(z)]\, H_m(z) = M[z^{-l} \;\; 0 \;\cdots\; 0]. \tag{9.19}$$

This is the top row of $\boldsymbol{F}_m(z)\boldsymbol{H}_m(z)$. The other rows come from modulating the F's:

$$\begin{bmatrix} F_0(z) & F_1(z) & \cdots \\ F_0(zW) & F_1(zW) & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} H_0(z) & H_0(zW) & \cdots \\ H_1(z) & H_1(zW) & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix} = M \begin{bmatrix} z^{-l} & & \\ & (zW)^{-l} & \\ & & \ddots \end{bmatrix}. \tag{9.20}$$

This is perfect reconstruction with causal filters. It applies to all cases—orthogonal or biorthogonal. In the orthogonal case the filters are $F_k(z) = z^{-l}H_k(z^{-1})$. So bring the diagonal matrix in (9.20) to the other side to find $\overline{\boldsymbol{H}}_m^T(z^{-1})\boldsymbol{H}_m(z) = M\boldsymbol{I}$. Note that $\boldsymbol{H}_m$ has complex coefficients because of the powers of W. The polyphase matrix $\boldsymbol{H}_p$ has real coefficients, because all filters are real.

Perfect Reconstruction: Polyphase Form

We have just given the perfect reconstruction condition in modulation form. It was written explicitly for $M = 3$ in equation (9.10), by following the signal through three channels. Equation (9.19) is the same statement for M channels. Equation (9.20) expresses that PR condition in terms of the matrices $\boldsymbol{H}_m(z)$ and $\boldsymbol{F}_m(z)$. Now we express perfect construction in terms of $\boldsymbol{H}_p(z)$ and $\boldsymbol{F}_p(z)$.

The reader will not be surprised by the result. The beauty of the polyphase form is the way it handles the algebra, for any number of filters.

Theorem 9.2 *An M-channel filter bank gives perfect reconstruction if*

$$\boldsymbol{F}_p(z)\boldsymbol{H}_p(z) = z^{-L}\boldsymbol{I}. \tag{9.21}$$

The overall delay of the system is $l = M - 1 + LM$, so that $T_0(z) = cz^{-l}$. The analysis and synthesis filters are ***biorthogonal.***

A first proof begins with the modulation matrices in (9.20). Then the identity (9.18) converts to polyphase matrices. The result is the PR condition (9.21).

For a second proof, imagine the analysis-synthesis cascade *with the polyphase matrices in the middle*. The bank begins with an S/P converter and ends with a P/S converter. When the product of polyphase matrices is $z^{-L}\boldsymbol{I}$, the whole filter bank is a simple delay. This is perfect reconstruction.

Note: The most general case allows reordering of the channels. The product of the Type-I analysis matrix $\boldsymbol{H}_p(z)$ and the Type-II synthesis matrix $\boldsymbol{F}_p(z)$ is

$$\boldsymbol{F}_p(z)\boldsymbol{H}_p(z) = z^{-L}\begin{bmatrix} \boldsymbol{0} & \boldsymbol{I}_{M-r} \\ z^{-1}\boldsymbol{I}_r & \boldsymbol{0} \end{bmatrix}. \tag{9.22}$$

The overall delay is increased by r. We mention that z^{-1} is present below the diagonal whenever the system is alias-free; this gives the "pseudo-circulant" of Problem 4. For $r = 0$ we return to the fundamental case (9.21), when the product is $z^{-1}\boldsymbol{I}_M$.

Example 9.2. Suppose $\boldsymbol{v}$ is any unit column vector: $\boldsymbol{v}^T\boldsymbol{v} = 1$. Then choose

$$\boldsymbol{H}_p(z) = \boldsymbol{I} - \boldsymbol{v}\boldsymbol{v}^T + z^{-1}\boldsymbol{v}\boldsymbol{v}^T \quad \text{with} \quad \boldsymbol{H}_p^{-1}(z) = z\boldsymbol{v}\boldsymbol{v}^T + \boldsymbol{I} - \boldsymbol{v}\boldsymbol{v}^T. \tag{9.23}$$

You can verify immediately that the product is I. The matrix $H_p(z)$ gives an orthogonal analysis bank, and the causal matrix $F_p(z) = z^{-1}H_p^{-1}(z)$ gives the synthesis bank. These degree-one matrices are the building blocks for *all* paraunitary matrices. This is the great factorization theorem began by Belevitch and completed by Vaidyanathan [V, p. 273]:

$$\textit{Every paraunitary } H_p(z) \textit{ factors into } \left(\prod_{j=1}^{L}\left[I - v_jv_j^T + v_jv_j^Tz^{-1}\right]\right)Q. \tag{9.24}$$

Example 9.3. Suppose $w^Tv = 1$. Choose the analysis polyphase matrix

$$H_p(z) = I - vw^T + z^{-1}vw^T \quad \text{with} \quad H_p^{-1}(z) = zvw^T + I - vw^T. \tag{9.25}$$

Again the product is I. But H_p^{-1} is not H_p^T. The causal matrix $F_p(z) = z^{-1}H_p^{-1}(z)$ gives the *biorthogonal* synthesis bank (not orthogonal unless $v = w$). These degree-one matrices do *not* give building blocks for the most general biorthogonal filter banks. [PhVaid] have identified the correct subclass.

DFT Filter Banks

The analysis filters are all modulations $H_k(z) = H_0(zW^k)$ of the lowpass filter. The idealized responses are shown in Figure 9.8. Notice that the frequency allocations are very different from Figure 9.5. Similarly the synthesis filters are $F_k(z) = F_0(zW^k)$. All filters have linear phase if H_0 and F_0 have linear phase. But the filter coefficients are complex for $M > 2$, because $W = e^{-j2\pi/M}$ is complex.

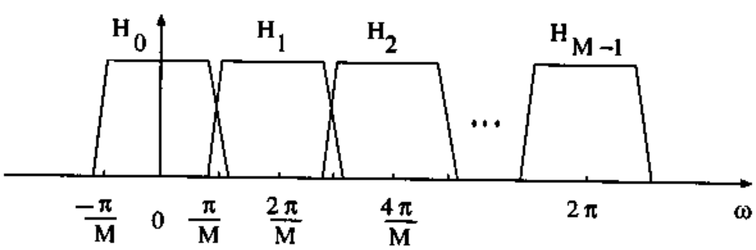


Figure 9.8: Idealized frequency response of the DFT filter bank.

This DFT filter bank is an excellent example for the use of polyphase matrices. Figure 9.2 showed the result of reordering the steps of subsampling and filtering. The lowpass filter coefficients $h_0(n)$ on the left were separated into M phases on the right: $H_0(z) = E_0(z^M) + z^{-1}E_1(z^M) + \cdots$. The polyphase matrix for the whole analysis bank is

$$H_p(z) = [DFT]\,\mathrm{diag}(E_0(z), E_1(z), \ldots, E_{M-1}(z)). \tag{9.26}$$

The implementation of a DFT filter bank is shown in Figure 9.9. The input signal $x(n)$ is blocked into vectors of M components by the delay chain and downsamplers. The subband signals are then filtered by the polyphase components $E_\ell(z)$ of $H_0(z)$ and passed through an M-point DFT. The polyphase filters have real coefficients, so the inputs to the DFT are real. If the

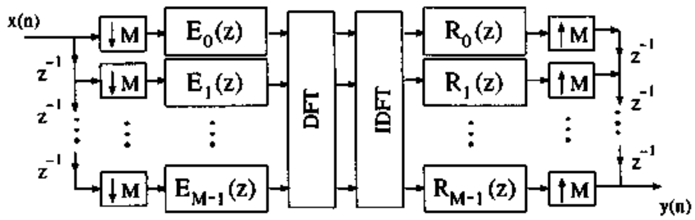


Figure 9.9: Polyphase representation of M-channel uniform DFT filter bank.

length of the lowpass filter is N, the computation load for the analysis bank is N multiplications and an M-point DFT for every M input samples.

The synthesis bank reverses these steps. We use the Type-II polyphase form of the lowpass filter $F_0(z) = R_{M-1}(z^M)+\cdots+z^{-(M-1)}R_0(z^M)$. Then the polyphase matrix $\boldsymbol{F}_p(z)$ has the IDFT multiplied by the diagonal matrix with entries $R_0(z), \ldots, R_{M-1}(z)$. The product of analysis and synthesis is a *diagonal matrix* when the DFT and IDFT cancel:

$$\boldsymbol{F}_p(z)\boldsymbol{H}_p(z) = \operatorname{diag}(R_0(z)E_0(z), R_1(z)E_1(z), \ldots, R_{M-1}(z)E_{M-1}(z)). \tag{9.27}$$

This filter bank is alias-free if all diagonal elements are equal. It gives perfect reconstruction if those equal elements are pure delays:

$$\textbf{\textit{PR condition}} \quad R_k(z)E_k(z) = z^{-L} \quad \text{for all } k. \tag{9.28}$$

But these filters $R_k(z) = z^{-L}E_k^{-1}(z)$ are IIR rather than FIR. For stability we would want each $E_k(z)$ to be minimum phase (zeros inside the unit circle). We also hope for linear phase. It is impossible to have both.

Theorem 9.3 [Nguyen-Vaidyanathan] *The polyphase components of a linear phase filter $H_0(z)$ cannot all have minimum phase.*

The anti-aliasing option is to choose FIR synthesis filters $R_k(z)$:

$$R_k(z) = E_0(z)E_1(z)\cdots E_{M-1}(z) \quad \text{omitting } E_k(z). \tag{9.29}$$

These filters are very long, about $M-1$ times as long as the analysis filters. The products $R_k(z)E_k(z)$ are all equal and aliasing is cancelled. The distortion function $T_0(z)$ is the product with a delay:

$$T_0(z) = z^{-l}E_0(z)E_1(z)\cdots E_{M-1}(z). \tag{9.30}$$

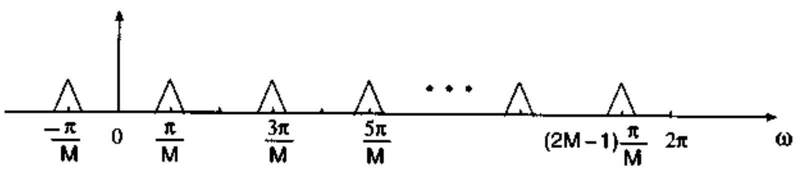
Perfect reconstruction is not possible. The natural question is whether one can design $H_0(z)$ such that either the magnitude or phase distortion is cancelled. Is there any choice of $H_0(z)$ such that T_0 is an allpass function (thus, no magnitude distortion) or a linear-phase function (thus, no phase distortion)? The answers are direct and beautiful:

If $H_0(z)$ has linear phase then $T_0(z)$ has linear phase.

If the components $E_k(z)$ are allpass then $T_0(z)$ is allpass.

In either case the output needs to be equalized. The equalizer for the allpass case should be an allpass filter that equalizes the phase distortion. The equalizer for the linear phase case should be a linear phase filter that equalizes the magnitude distortion.

For these alias-free filter banks, with synthesis filters from (9.29), we note that $H_0(z)$ and $F_0(z)$ cannot both be good lowpass filters. If they were good, there would be almost no overlap between $H_k(z)$ and $H_{k+2}(z)$, and between $F_k(z)$ and $F_{k+2}(z)$. Then the products $F_k(z)H_k(zW)$ do not overlap and their sum $T_1(z)$ cannot be zero! A typical $T_1(z)$ from good filters is drawn below:



Summary The polyphase method gives a direct approach to the analysis of DFT filter banks. Unfortunately, the results are almost all negative. We cannot even cancel aliasing. The DFT bank is fast, but not good in reconstruction. The DCT bank in Section 9.4 is also fast, and it reconstructs perfectly.

Properties of Paraunitary Filter Banks and Matrices

Power Complementary $\sum_{k=0}^{M-1} \left|H_k(e^{j\omega})\right|^2 = 1$ and $\sum_{k=0}^{M-1} H_k(z^{-1})H_k(z) = 1$.

This assures that there is no magnitude distortion: $T_0(z) = 1$. It does not determine the aliasing transfer functions. Thus paraunitary implies power complementary but power complementary does not imply paraunitary.

Time Reversal $F_k(z) = z^{-N} H_k(z^{-1})$

The magnitude responses satisfy $|F_k(z)| = |H_k(z)|$. $F_k(z)$ has maximum phase if $H_k(z)$ has minimum phase. If $H_k(z)$ is linear phase, then so is $F_k(z)$. The modulation matrix $\boldsymbol{H}_m(z)$ and the AC matrix $\boldsymbol{H}_m^T(z)$ are paraunitary. This follows from $\boldsymbol{H}_m(z) = (\text{DFT})\boldsymbol{D}(z)\boldsymbol{H}_p^T(z^M)$.

Spectral Factor of M-th Band Filter The rows of $\boldsymbol{H}_m(z)$ are paraunitary: $P_k = \widetilde{H}_k H_k$ has

$$\sum_{l=0}^{M-1} \widetilde{H}_k(zW^l)H_k(zW^l) = \sum_{l=0}^{M-1} P_k(zW^l) = 1.$$

Then P_k is an M-th band filter. Notice that $P_k(z)$ is linear phase by definition. Thus, $H_k(z)$ is a spectral factor of an M-th band linear phase filter.

Columnwise orthogonality The kth and ℓth columns are orthogonal. The elements in column k are power complementary: $\sum_m H_{mk}(z^{-1})\, H_{mk}(z) = 1$.

Determinant is an allpass function Let $\boldsymbol{H}(z)$ be a square paraunitary matrix and let $A(z)$ denote its determinant. Then $\widetilde{\boldsymbol{H}}(z)\, \boldsymbol{H}(z) = \boldsymbol{I}$ implies that $\widetilde{A}(z)A(z) = 1$. Consequently, $A(z)$ is an allpass function. For an FIR paraunitary system, $A(z)$ is a delay.

Submatrices and cascades Any columns of a paraunitary matrix $H(z)$ give a paraunitary submatrix. Any cascade $H_1(z)H_2(z)$ is also paraunitary.

Problem Set 9.2

1. Show that $H = I - 2vv^T$ is a symmetric orthogonal matrix if $v^Tv = 1$. This is a *Householder reflection* ($\det H = -1$) in the plane perpendicular to v. Show that $H = I - vv^T + z^{-1}vv^T$ is a paraunitary matrix.

2. What is the entry in row k, column l, of the modulation matrix $H_m(z)$?

3. Verify that $F_p(z)H_p(z) = z^{-1}I$ and find the coefficients in all four filters:

$$F_p(z) = \begin{bmatrix} 3+4z^{-1} & -3-2z^{-1} \\ -2-2z^{-1} & 2+z^{-1} \end{bmatrix} \quad H_p(z) = \begin{bmatrix} 2+z^{-1} & 3+2z^{-1} \\ 2+2z^{-1} & 3+4z^{-1} \end{bmatrix}.$$

Is this an orthogonal bank? Is $H_p(z)$ paraunitary? What are the determinants?

4. A two-channel bank is ***alias-free*** if $T_1(z) = F_0(z)H_0(-z) + F_1(z)H_1(-z) = 0$. Verify that $F_m(z)H_m(z)$ is diagonal. Then substitute (9.18) to prove that $F_p(z)H_p(z)$ is a *pseudo-circulant matrix*:

$$F_m(z)H_m(z) = \begin{bmatrix} T_0(z) & \\ & T_0(-z) \end{bmatrix} \quad F_p(z)H_p(z) = \begin{bmatrix} T_{0,\text{even}} & T_{0,\text{odd}} \\ z^{-1}T_{0,\text{odd}} & T_{0,\text{even}} \end{bmatrix}.$$

This pattern extends to all M-channel alias-free filter banks. When the aliasing functions $T_1, \ldots, T_{M-1}$ are zero, the product $F_m(z)H_m(z)$ is a diagonal matrix with entries $T_0(z)$, $T_0(zW), \ldots$. The product $F_p(z)H_p(z)$ is again a pseudocirculant. The prefix "pseudo" comes from the extra factor z^{-1} multiplying all entries below the main diagonal.

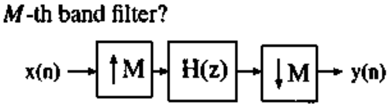
5. For the following analysis filters, find $H_p(z)$ and its determinant. Find synthesis filters $F_k(z)$ such that the overall system is PR.

   (a) $H_0(z) = 1 + 3z^{-1} - 2z^{-2}$, $H_1(z) = z^{-1} + 2z^{-2}$, $H_2(z) = z^{-2}$

   (b) $H_0(z) = 1 + z^{-1} + z^{-2} + z^{-3} + 2z^{-4} - z^{-5}$, $H_1(z) = 1 + z^{-1} + z^{-2} + z^{-3} + 3z^{-4} + z^{-6} + 2z^{-7} - z^{-8}$, $H_2(z) = z^{-1} + z^{-2} - 2z^{-4} + z^{-5}$

6. Let $H_0(z) = 1 + z^{-1} + z^{-2} - 0.5z^{-3}$, $H_1(z) = z^{-1} + z^{-2} - 0.25z^{-4}$, and $H_2(z) = z^{-2}$. Find the PR synthesis filters in polyphase form $F_p(z)$ and modulation form $F_m(z)$ using (9.12).

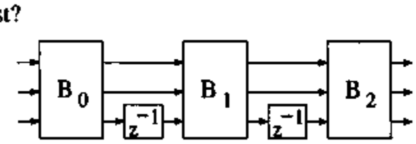
7. Find the relation between $Y(z)$ and $X(z)$ below. Is it LTI? What is the system if $H(z)$ is an odd-length linear phase M-th band filter?



8. What is the polyphase matrix for a Serial/Parallel converter?

9. Let $H_1(z)$ and $H_2(z)$ be paraunitary. Show that the cascade system $H_1(z)H_2(z)$ is paraunitary. How about the system $\frac{1}{2}(H_1(z) + H_2(z))$?

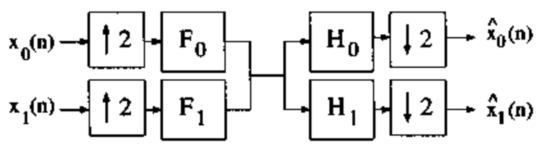
10. Show that the 3×3 polyphase matrix is paraunitary, given that B_k are orthogonal matrices. Suppose the delays z^{-1} are replaced by the first-order allpass section $A(z) = \frac{a^* + z^{-1}}{1 + az^{-1}}$. Find $H_p(z)$ and its determinant. Find the synthesis polyphase matrix $F_p(z)$ that cancels aliasing. Does PR synthesis exist?



11. A *transmultiplexer* is the reverse of a filter bank (synthesis first). Two or more signals $x_k(n)$ are multiplexed and sent over a high bandwidth channel. At the receiving end, the signal is demultiplexed. The outputs $\widehat{x}_k(n)$ suffer from distortion and crosstalk. A PR transmultiplexer cancels crosstalk and reconstructs the signals $x_k(n)$ exactly.

(a) Express $\widehat{X}_k(z)$ in terms of $X(z)$ and the filters. What are the conditions on $H_k(z)$ and $F_k(z)$ such that the transmultiplexer is PR?

(b) Suppose $(H_k'(z),\ F_k'(z))$ yield a PR filter bank, and a transmultiplexer is constructed by $H_k(z) = H_k'(z),\ F_k(z) = z^{-1}F_k'(z)$. Show that this choice yields a PR transmultiplexer.



12. Let $(H_k(z),\ F_k(z))$ be an M-channel PR filter bank. What are the exact conditions on L such that $(H_k(z^L),\ F_k(z^L))$ is also PR?

13. Let $(H_k(z),\ F_k(z))$ be the filters in a paraunitary filter bank. Define $H_k'(z) = H_k(ze^{j\theta})$ and show that the filter bank is still paraunitary.

9.3 Perfect Reconstruction, Linear Phase, Orthogonality

There is much to recommend the time domain. The filter matrix $\boldsymbol{H}_b$ in block form — all M analysis channels together — is a *block Toeplitz matrix*. These matrices are easy to understand and analyze, if you remember that each entry is an $M \times M$ block. With filters of length $2M$, we have two blocks on each row as shown:

$$\boldsymbol{H}_b = \begin{bmatrix} \cdot & \cdot & \cdot & & & \\ & \underline{h}(1) & \underline{h}(0) & & & \\ & & \underline{h}(1) & \underline{h}(0) & & \\ & & & \underline{h}(1) & \underline{h}(0) & \\ & & & & \cdot & \cdot \end{bmatrix} \tag{9.31}$$

This is the Lapped Orthogonal Transform (LOT) of Malvar. It is simply a filter bank with some overlapping but not much. We could have less overlapping or more:

1. No overlapping: $\boldsymbol{H}_b$ is block diagonal = block transform (DFT and DCT).
2. One overlap: $\boldsymbol{H}_b$ is block 2-diagonal = Lapped Orthogonal Transform (LOT).
3. Filter lengths BM: $\boldsymbol{H}$ has B block diagonals = general case.

We are not introducing new filter banks. These are standard M-channel analysis banks, including the downsampling operation $(\downarrow M)$. The only novelty is to interleave the channels, so that we watch all M channels at once. Previously $(\downarrow 2)\boldsymbol{H}_0$ from the lowpass channel was written above $(\downarrow 2)\boldsymbol{H}_1$ from the highpass channel:

$$\begin{bmatrix} (\downarrow 2)\boldsymbol{H}_0 \\ (\downarrow 2)\boldsymbol{H}_1 \end{bmatrix} = \begin{bmatrix} h_0(3) & h_0(2) & h_0(1) & h_0(0) & & \\ & & h_0(3) & h_0(2) & h_0(1) & h_0(0) \\ & & & & \cdots & \cdots \\ h_1(3) & h_1(2) & h_1(1) & h_1(0) & & \\ & & h_1(3) & h_1(2) & h_1(1) & h_1(0) \\ & & & & \cdots & \cdots \end{bmatrix} \tag{9.32}$$

Interleaving rows gives our block Toeplitz matrix, with block size $M = 2$:

$$H_b = \begin{bmatrix} h_0(3) & h_0(2) & h_0(1) & h_0(0) & & \\ h_1(3) & h_1(2) & h_1(1) & h_1(0) & & \\ & & h_0(3) & h_0(2) & h_0(1) & h_0(0) \\ & & h_1(3) & h_1(2) & h_1(1) & h_1(0) \\ & & & & \cdots & \cdots \\ & & & & \cdots & \cdots \end{bmatrix} \quad (9.33)$$

This is the time domain block form when $M = 2$ and $B = 2$. The filter lengths are $BM = 4$. The z-transform is the polyphase matrix

$$H_p(z) = \underline{h}(0) + z^{-1}\underline{h}(1) = \begin{bmatrix} h_0(0) & h_0(1) \\ h_1(0) & h_1(1) \end{bmatrix} + z^{-1} \begin{bmatrix} h_0(2) & h_0(3) \\ h_1(2) & h_1(3) \end{bmatrix}.$$

Note The block $h(k)$ contains the k-th coefficient from each phase of each filter. The underbar in $\underline{h}(k)$ is used to emphasize that this $M \times M$ block is the coefficient of z^{-k} in $H_p(z)$. When we form blocks, *the column order inside each block is to be reversed.* Thus $h_0(0)$ is to the left of $h_0(1)$ in the block, where it was to the right in (9.33). The orthogonality conditions on $\underline{h}(0)$ and $\underline{h}(1)$ are immediate in the time and z-domains:

Theorem 9.4 *The Lapped Orthogonal Transform* (LOT) *requires*

$$\begin{aligned} \underline{h}(0)^T\underline{h}(0) + \underline{h}(1)^T\underline{h}(1) &= I_{M\times M} \\ \text{(orthogonality of tails)} \quad \underline{h}(1)^T\underline{h}(0) &= 0_{M\times M} \end{aligned} \quad (9.34)$$

Also $H_bH_b^T = I$. *This moves the transposes to the second factors in* (9.34).

The general case with B blocks per row (and filter lengths BM) will produce B block equations from $H_b^TH_b = I$. You could say that these equations are "Condition O" in the time domain. The paraunitary requirements were "Condition O" in the polyphase domain and modulation domain. Notice the difference from the two-channel case. There Condition O was applied only to one filter H_0. The highpass filter H_1 was determined by an alternating flip. Here, unless we impose a special structure on the bandpass filters $H_1, \ldots, H_{M-1}$, we must include them all in Condition O.

The DFT filter bank does impose such a structure — but it makes orthogonality impossible (except in the block transform case $B = 1$ which is pure DFT with no filters). The DCT filter bank also imposes a structure. It produces all M filters from the knowledge of one filter. But this time, for the DCT bank, *orthogonality is possible.* In that case Condition O is no longer double-shift orthogonality, it is M-shift orthogonality.

For the most general LOT, multiply (9.34) by $\underline{h}(0)$ to find $\underline{h}(0)\underline{h}(0)^T\underline{h}(0) = \underline{h}(0)$. We freely use $\underline{h}(0)\underline{h}(1)^T = 0$ and $\underline{h}(0)^T\underline{h}(1) = 0$ to obtain

$$\begin{aligned} \underline{h}(0) &= \underline{h}(0)\underline{h}(0)^T(\underline{h}(0) + \underline{h}(1)) &= PQ \\ \underline{h}(1) &= \underline{h}(1)\underline{h}(1)^T(\underline{h}(0) + \underline{h}(1)) &= (I - P)Q. \end{aligned} \quad (9.35)$$

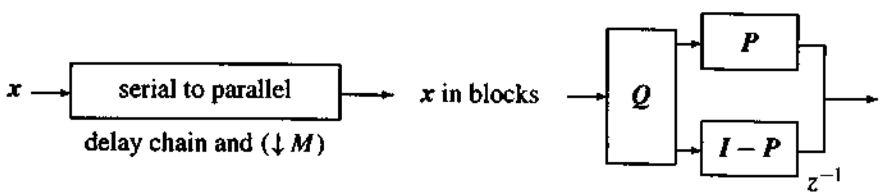
The matrix $Q = \underline{h}(0) + \underline{h}(1)$ is orthogonal because (9.34) gives $Q^TQ = I$. The matrix $P = \underline{h}(0)\underline{h}(0)^T$ is a projection matrix. It is symmetric and $P^2 = \underline{h}(0)\underline{h}(0)^T\underline{h}(0)\underline{h}(0)^T =$

$\underline{h}(0)\underline{h}(0)^T = P$. Similarly, $\underline{h}(1)\underline{h}(1)^T$ is a projection, and the two projections add to I. The general solution (9.35) for the blocks in H_b was found by Heller and Tolimieri:

$$\underline{h}(0) = PQ \text{ and } \underline{h}(1) = (I - P)Q \quad \text{with } P = \text{projection}, Q = \text{orthogonal}.$$

The special case $P = I$ gives *one* block. Then H_b is block diagonal and it represents a simple orthogonal block transform. The general LOT can choose any projection matrix P and orthogonal matrix Q. Notice that $Q = H_p(1)$.

Note 1. The polyphase matrix has the convenient form $[P + (I - P)z^{-1}]Q$. This is fast if Q and P are fast. Below, we choose Q = DCT matrix:



Note 2. The *biorthogonal* lapped transform with $B = 2$ has a very similar pattern (Problem 2). The matrix $P = \underline{h}(0)\underline{f}(0)$ still equals P^2. In this case, $P = P^T$ and $Q^TQ = I$ are not required. This gives the most general PR bank with filter lengths $2M$. The biorthogonal BOLT [PhVaid] extends this design to $B > 2$.

Note 3. For a paraunitary matrix $H_p(z)$, the degree of the determinant is the *Smith-McMillan degree*. When P has rank r, the degree is $L = M - r$. For the very special case of diagonal P, with r ones and $M - r$ zeros, the factors are clear:

$$H_p(z) = \begin{bmatrix} I_r & 0 \\ 0 & I_{M-r}z^{-1} \end{bmatrix} Q \text{ has determinant } z^{-(M-r)}.$$

Note 4. *For fast implementation, Q often starts with the DCT matrix* (the matrix C of cosines). The order M is even. Half the DCT rows and columns are symmetric, half are antisymmetric. It is natural to think of separating those parts and maintaining linear phase. Figure 9.10 does this separation by reordering the DCT rows to put C_{even} above C_{odd}. Then Malvar uses Haar butterflies and delays to separate $\underline{h}(1)$ from $\underline{h}(0)$. At the end we may apply different orthogonal matrices U and V. The blocks $\underline{h}(0)$ and $\underline{h}(1)$ are then

$$\begin{bmatrix} \underline{h}(0) & \underline{h}(1) \end{bmatrix} = \begin{bmatrix} U & V \end{bmatrix} \begin{bmatrix} C_e - C_o & (C_e - C_o)J \\ C_o - C_e & (C_e - C_o)J \end{bmatrix}. \tag{9.36}$$

The free parameters in this LOT are U and V. In practice, those must also allow fast multiplication. Malvar [Mr, p. 167] suggests $U = I$ and V = product of plane rotations or V = product of DST-IV and transposed DCT-II. The LOT is traditionally chosen to be linear phase and to start with the DCT-II.

Figure 9.11 shows the frequency responses of the LOTs. They are obtained by optimizing the coding gain (part a) and the stopband attenuation (part b). We show four basis functions (impulse responses) of the LOT that has high coding gain. Notice that they are symmetric and antisymmetric from the DCT. In summary, LOT is a linear phase paraunitary filter bank where the filter length is $2M$.

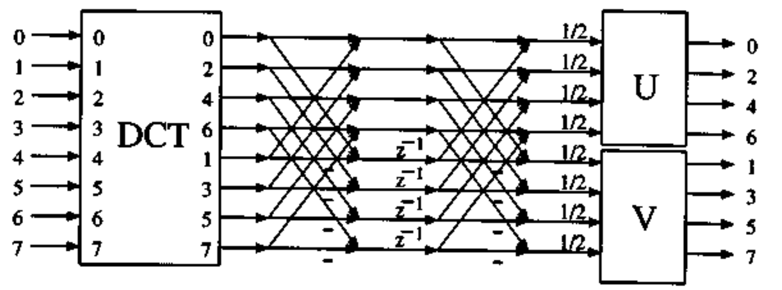


Figure 9.10: Polyphase transfer matrix of the Lapped Orthogonal Transform.

Longer Filters and GenLOT

You recognize that the two blocks ($B = 2$) of the LOT cannot give a sharp cutoff in frequency with great attenuation. The filters need to be longer. The time domain matrix H_b is always block Toeplitz when the input signal is blocked into M samples at a time. If the filter lengths are BM, there are B blocks in each row of H_b (and $B = 2$ for LOT):

$$H_b = \begin{bmatrix} \cdots & \cdots & \cdots & & & \\ & \underline{h}(B-1) & \cdots & \underline{h}(0) & & \\ & & \underline{h}(B-1) & \cdots & \underline{h}(0) & \\ & & & \cdots & \cdots & \cdots \end{bmatrix}. \tag{9.37}$$

The polyphase matrix for this FIR analysis bank with decimators ($\downarrow M$) is

$$H_p(z) = \underline{h}(0) + \underline{h}(1)z^{-1} + \cdots + \underline{h}(B-1)z^{-(B-1)}. \tag{9.38}$$

Note again that all these blocks are $M \times M$. The conditions for orthogonality come directly from $H_b^T H_b = I$ and from $H_p^T(z^{-1})H_p(z) = I$:

Theorem 9.5 *The analysis bank with filter length BM is orthogonal if*

$$\sum_{k=0}^{B-1-l} \underline{h}(k)^T \underline{h}(k+l) = \delta(l)I. \tag{9.39}$$

A block transform has $B = 1$ and only diagonal blocks in H_b. The constant polyphase matrix is $H_p(z) = \underline{h}(0)$. This gives bad results at block boundaries after compression. Overlapping blocks are much smoother but the filters have BM coefficients, which gives a huge design problem until we narrow it by structural decisions. One good decision is to start the filter bank with the DCT-II.

The simplest filters to follow the DCT are *block Haar and block diagonal and block delay*:

$$W = \frac{1}{\sqrt{2}}\begin{bmatrix} I & I \\ I & -I \end{bmatrix} \text{ and } Q_i = \begin{bmatrix} U_i & 0 \\ 0 & V_i \end{bmatrix} \text{ and } D(z) = \begin{bmatrix} I & 0 \\ 0 & z^{-1}I \end{bmatrix}.$$

The blocks are of order $\frac{M}{2}$ and these matrices are orthogonal. Therefore, all products are paraunitary. The product also has linear phase, if Q at the beginning of the filter bank separates even and odd rows of the cosine matrix C^{II} (symmetric and antisymmetric cosine basis functions). The GenLOT is a cascade of these special filters:

$$\textbf{\textit{The GenLOT has }} H_p(z) = \left(Q_{B-1}WD(z)W\right)\cdots\left(Q_1 WD(z)W\right)Q. \tag{9.40}$$

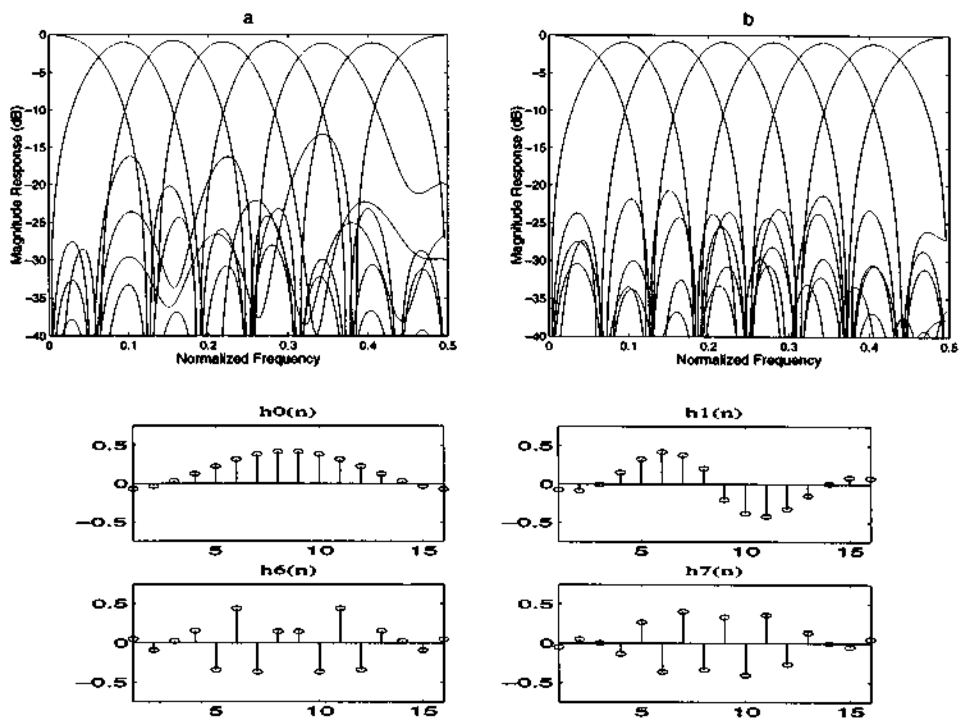


Figure 9.11: Frequency responses of the LOT. (a) High coding gain. (b) High stopband attenuation. Four of the eight symmetric-antisymmetric basis functions in (a).

The implementation flow graph for the analysis bank is in Figure 9.12 (see *GenLOT* in the Glossary for the details of K_1). The DCT-based LOT is the case with $B = 2$. The DCT itself is the case with $B = 1$ and $\boldsymbol{H}_p = \boldsymbol{Q}$. This framework covers all orthogonal filter banks with $\frac{M}{2}$ symmetric and antisymmetric channels — linear phase! The design problem is always to choose fast $\boldsymbol{Q}_i$ that also achieve sharp frequency discrimination.

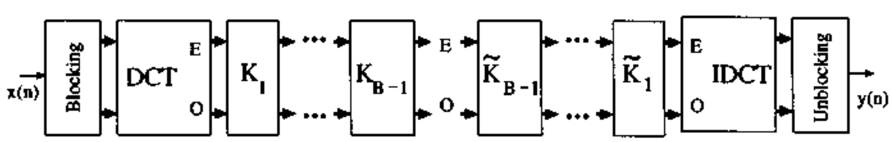


Figure 9.12: Polyphase transfer matrix of GenLOT.

Figure 9.13 shows the frequency responses of $H_k(z)$. The coding gain is 9.36 dB (left) and 23 dB attenuation (right). GenLOT design is included in

http://saigon.ece.wisc.edu/~waveweb/QMF.html

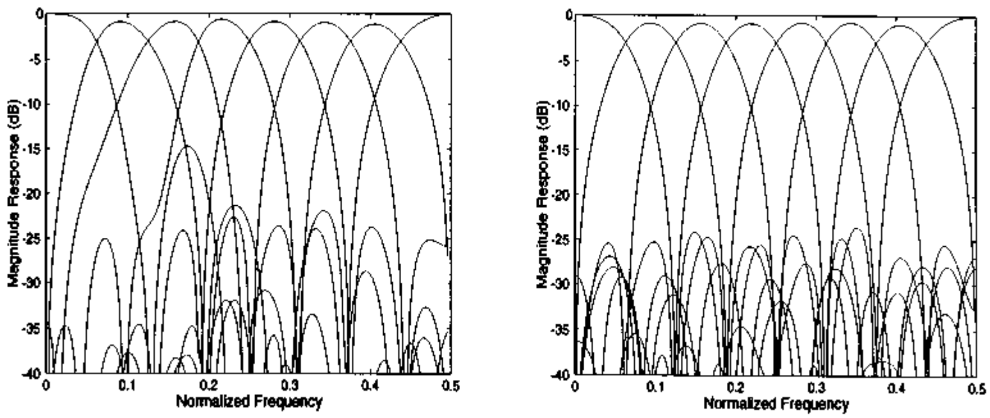


Figure 9.13: Frequency response plots for the 8-channel GenLOT (length 32) with high coding gain and high stopband attenuation.

Products of Rotations and Delays

Every $M \times M$ orthogonal matrix Q is a product of $\frac{1}{2}M(M-1)$ plane rotations (*Givens rotations*). R_{ij} gives rotation by θ_{ij} in the plane of axes i and j:

$$R_{ij} = \begin{bmatrix} 1 & & & & \\ & c & & s & \\ & & 1 & & \\ & -s & & c & \\ & & & & 1 \end{bmatrix} \quad \begin{array}{l} c = \cos(\theta_{ij}) \text{ and } s = \sin(\theta_{ij}) \\ \text{in rows and columns } i \text{ and } j. \end{array}$$

The angles θ_{ij} can be design variables (but strongly nonlinear). The rotations in the GenLOT stay within even channels (to give U_i) or odd channels (to give V_i). The *Givens factorization* of any causal FIR paraunitary matrix is parallel to the *Householder factorization* in (9.24). Where Householder uses reflections, Givens uses a sequence of rotations:

$$H_p(z) = R_L D(z) \cdots R_2 D(z) R_1 D(z) Q. \tag{9.41}$$

Here $D(z) = \text{diag}(1, 1, \ldots, z^{-1})$ delays only the last channel. Its Smith-McMillan degree is 1, so the degree of $H_p(z)$ is L. Q is any unitary matrix. The matrices R_k can be chosen as *products of* $M-1$ *rotations only*, in neighboring channels $(1, 2), (2, 3), \ldots, (M-1, M)$. The total number of parameters (plane rotations) is then $L(M-1)$ plus $\frac{1}{2}M(M-1)$ for Q. This is the same total as in the product (9.24) from Householder reflections.

Example 9.4. Consider a three-channel orthogonal filter bank whose polyphase transfer matrix has McMillan degree 2. Figure 9.14 shows the corresponding lattice structure. Each rotation is a simple butterfly connecting channels i and j. The total number of angles is $4 + 3 = 7$, and the filters have length 9.

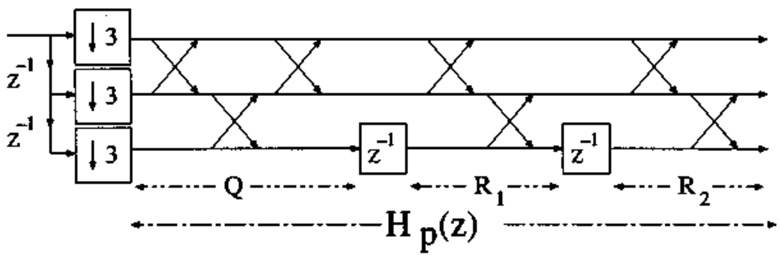


Figure 9.14: A three-channel paraunitary filter bank with McMillan degree 2.

Pairwise Mirror Image (PMI) Filter Banks

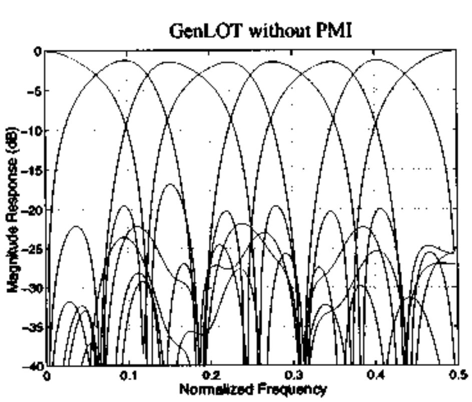
A further simplification of the GenLOT and other filter banks is to make each filter H_{M-1-k} a "mirror image" of H_k with respect to $\omega = \frac{\pi}{2}$:

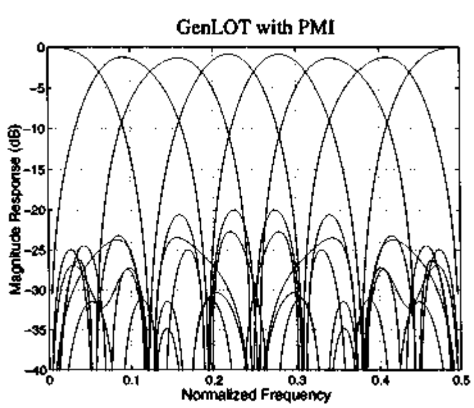
$$|H_{M-1-k}(\omega)| = |H_k(\pi - \omega)|. \tag{9.42}$$

This reduces the number of design parameters by a factor of 2. When M is odd, say $M = 3$, the middle filter is its own mirror image and $H_1(z)$ is a function of z^2. The other two channels have $H_2(z) = H_0(-z)$. There is a convenient lattice structure [NgVa3], and the PMI property is structurally imposed.

In the GenLOT this pairwise property connects the orthogonal matrices U_i and V_i. The polyphase matrix satisfies $\boldsymbol{J}\boldsymbol{H}_p(z) = \boldsymbol{H}_p(z)\,\boldsymbol{\Gamma}$, where $\boldsymbol{\Gamma} = \text{diag}(1, -1, \ldots, 1, -1)$. The matrices $\boldsymbol{V}_i$ are equal to $\boldsymbol{\Gamma}\boldsymbol{U}_i\boldsymbol{\Gamma}$, except the last of the $\boldsymbol{V}$'s is $\boldsymbol{J}\boldsymbol{U}_i\boldsymbol{\Gamma}$.

We compare the frequency responses of two 8-channel GenLOT's with $B = 3$ and filter length 24. The second set of responses has the pairwise mirror image (PMI) property. It displays better stopband attenuation (measured in dB), which was the objective function in the design process.





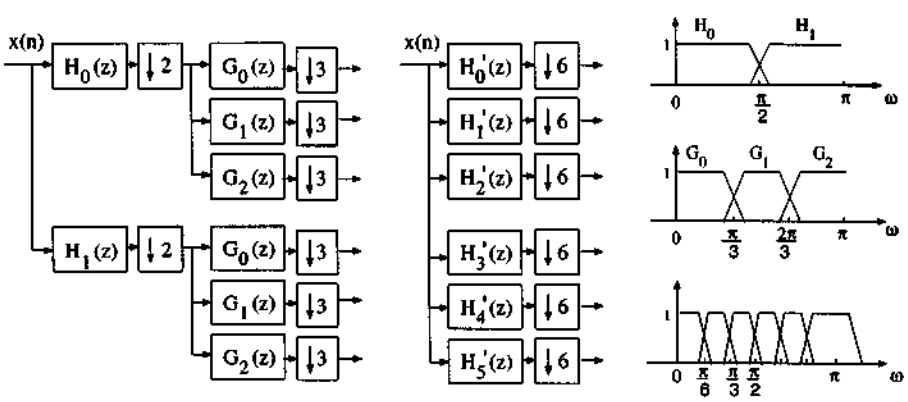
Permissible Lengths and Symmetry for Linear Phase

The filter lengths in GenLOT are the same and the number of symmetric and antisymmetric filters are equal. The conditions on the filter lengths $K_i M + \beta$ and the symmetries for *biorthogonal* filter banks are summarized below [TranNg]:

Case	*Symmetry/Antisymmetry*	*Lengths*	*Sum of Lengths*
M even, β even	$\frac{M}{2}$ S & $\frac{M}{2}$ A	$\sum K_\ell$ is even	$2mM$
M even, β odd	$(\frac{M}{2}+1)$ S & $(\frac{M}{2}-1)$ A	$\sum K_\ell$ is odd	$2mM$
M odd, β even	$(\frac{M+1}{2})$ S & $(\frac{M-1}{2})$ A	$\sum K_\ell$ is odd	$(2m+1)M$
M odd, β odd	$(\frac{M+1}{2})$ S & $(\frac{M-1}{2})$ A	$\sum K_\ell$ is even	$(2m+1)M$

Tree-structured Filter Banks

A popular method to design M-channel filter banks is to cascade smaller systems. Wavelet packets use two-channel systems. The six-channel filter bank below is obtained from cascading a two-channel system with two three-channel systems:



Typical relations between $H_k'(z)$ and $H_k(z)$ and $G_k(z)$ are $H_0'(z) = H_0(z)G_0(z^2)$ and $H_5'(z) = H_1(z)G_2(z^2)$. The 6-channel filter bank is PR if and only if the 2-channel and 3-channel filter banks are PR.

One obtains *nonuniform* filter banks by cascading systems with different decimation factors [HoVaid]. A nonuniform bank with decimation factors $(6, 6, 6, 4, 4)$ uses a 3-channel and a 2-channel system at the second tree level. This cascade is drawn in the Glossary under *Tree-Structured Filter Banks*. We emphasize the simplicity of these designs.

Summary The analysis bank is a block Toeplitz multiplication in the time domain. It is a polyphase matrix multiplication in the z-domain. That matrix contains the blocks from the Toeplitz form, just as $H(z)$ contains the coefficients from one filter. The polyphase matrix extends the familiar idea of $H(z)$ from one filter to M filters. This is the efficient order that gives the polyphase form:

- (direct time domain) Apply the analysis filters and then $(\downarrow M)$.
- (more efficient order) Put the input in blocks and filter in parallel.
- (polyphase in z-domain) Multiply by $\boldsymbol{H}_p(z) = \sum \underline{\boldsymbol{h}}_p(n)z^{-n}$.

For $M \times M$ blocks, the kth row contains coefficients from the kth filter. Column n has the M coefficients starting with $h_k(nM)$.

In the LOT case, with filter length $2M$ and two blocks per row, the first M coefficients $h_k(0), \ldots, h_k(M-1)$ go into the main diagonal block $\underline{h}(0)$. The other coefficients $h_k(M), \ldots, h_k(2M-1)$ go into the subdiagonal block $\underline{h}(1)$.

The synthesis bank is also a block Toeplitz matrix. The filter lengths and the number of blocks could be different, and the blocks in $\boldsymbol{F}_b$ are transposed. The coefficients from the kth synthesis filter are in the kth *column* of the blocks (not the kth row). Thus the two $M \times M$ blocks $\underline{f}(0)$ and $\underline{f}(1)$ in the length $2M$ case would be

$$\begin{bmatrix} f_0(0) & \cdots & f_{M-1}(0) \\ \vdots & & \vdots \\ f_0(M-1) & \cdots & f_{M-1}(M-1) \end{bmatrix} \text{ and } \begin{bmatrix} f_0(M) & \cdots & f_{M-1}(M) \\ \vdots & & \vdots \\ f_0(2M-1) & \cdots & f_{M-1}(2M-1) \end{bmatrix}.$$

The simplicity in the time domain resides in the fact that the whole filter bank becomes a matrix multiplication: $\widehat{x} = \boldsymbol{F}_b\boldsymbol{H}_b\boldsymbol{x}$. For $M > 2$ there is enough freedom to maintain orthogonality, rather than biorthogonality, while achieving other good properties. In this case we ask $\boldsymbol{H}_b$ to be orthogonal. We also ask for fast implementation, which leads us to GenLOT (with the DCT matrix). Cosine modulation is the alternative described next.

Problem Set 9.3

1. Show that $\underline{h}(0) = \boldsymbol{PQ}$ and $\underline{h}(1) = (\boldsymbol{I}-\boldsymbol{P})\boldsymbol{Q}$ produce an orthogonal block Toeplitz $\boldsymbol{H}_b$, for any $\boldsymbol{P}$ = symmetric projection matrix ($\boldsymbol{P}^2 = \boldsymbol{P}$) and $\boldsymbol{Q}$ = orthogonal matrix ($\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}$).

2. When the synthesis bank is anti-causal, the PR condition is $\boldsymbol{F}_b\boldsymbol{H}_b = \boldsymbol{I}$:

$$\boldsymbol{F}_b = \begin{bmatrix} \underline{f}(0) & \underline{f}(1) & 0 & \\ & \underline{f}(0) & \underline{f}(1) & \\ & & \cdot & \end{bmatrix} \text{ and } \boldsymbol{H}_b = \begin{bmatrix} \underline{h}(0) & & \\ \underline{h}(1) & \underline{h}(0) & \\ 0 & \underline{h}(1) & \cdot \end{bmatrix}.$$

   (a) $\boldsymbol{F}_b\boldsymbol{H}_b = \boldsymbol{I}$ and $\boldsymbol{H}_b\boldsymbol{F}_b = \boldsymbol{I}$ give what six PR conditions on the blocks?

   (b) Deduce $\underline{f}(0) = (\underline{f}(0) + \underline{f}(1))\underline{h}(0)\underline{f}(0)$ and $\underline{h}(0) = \underline{h}(0)\underline{f}(0)(\underline{h}(0) + \underline{h}(1))$. Write the corresponding equations for $\underline{f}(1)$ and $\underline{h}(1)$.

   (c) Deduce also that $\boldsymbol{P} = \underline{h}(0)\underline{f}(0)$ equals $\boldsymbol{P}^2$ and $\underline{h}(1)\underline{f}(1)$ equals $\boldsymbol{I} - \boldsymbol{P}$.

   (d) Show finally that $\underline{f}(0) + \underline{f}(1)$ is the inverse of $\boldsymbol{Q} = \underline{h}(0) + \underline{h}(1)$. Then PR for filter banks with $B = 2$ blocks has this form with $\boldsymbol{P}^2 = \boldsymbol{P}$:

$$\underline{f}(0) = \boldsymbol{Q}^{-1}\boldsymbol{P},\ \underline{f}(1) = \boldsymbol{Q}^{-1}(\boldsymbol{I}-\boldsymbol{P}),\ \underline{h}(0) = \boldsymbol{PQ},\ \underline{h}(1) = (\boldsymbol{I}-\boldsymbol{P})\boldsymbol{Q}.$$

3. Verify that $\boldsymbol{F}_p(z)\boldsymbol{H}_p(z) = \boldsymbol{I}$ if $\boldsymbol{P}^2 = \boldsymbol{P}$ in the formulas

$$\boldsymbol{F}_p(z) = \boldsymbol{Q}^{-1}(\boldsymbol{P} + (\boldsymbol{I}-\boldsymbol{P})z) \text{ and } \boldsymbol{H}_p(z) = (\boldsymbol{P} + (\boldsymbol{I}-\boldsymbol{P})z^{-1})\boldsymbol{Q}.$$

4. If $\boldsymbol{P}^2 = \boldsymbol{P}$ has rank r, it can be diagonalized to give r ones and $M - r$ zeros. Show that $\det \boldsymbol{H}_p(z) = z^{-(M-r)} \det \boldsymbol{Q}$ which confirms that the filter bank is FIR.

5. By choosing 2×2 matrices P and Q in Problem 2 give examples of (a) an orthogonal bank and (b) a biorthogonal bank. What are the lowpass and highpass analysis coefficients in your examples?

6. Let $H_0(z) = \sum_{l=0}^{M-1} z^{-l} H_{0,l}(z^M)$ and $H_k(z) = H_0(zW^k)$. If $H_{0,l}(z)$ are allpass, the analysis filters $H_k(z)$ in this DFT bank are power complementary:

$$\begin{bmatrix} H_0(z) \\ \vdots \\ H_{M-1}(z) \end{bmatrix} = [\text{IDFT}] \begin{bmatrix} H_{0,0}(z^M) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & H_{0,M-1}(z^M) \end{bmatrix} \begin{bmatrix} 1 \\ \vdots \\ z^{-(M-1)} \end{bmatrix}.$$

This is not PR! With $F_0(z) = H_0(z)$, the DFT bank has no distortion. It only has aliasing.

7. Let $\underline{h}_k(z)$ and $\underline{h}_\ell(z)$ be columns of a paraunitary matrix $H(z)$. They are mutually orthogonal. Show that the elements $H_{k,m}(z)$ in the kth column are power complementary filters: $\sum H_{k,m}(z^{-1}) H_{k,m}(z) = 1$.

8. We could design M linear phase M-th band filters $P_k(z)$ and find their spectral factors $H_k(z)$. What are the properties of this analysis bank? It is not necessarily PR. Does it have aliasing or amplitude or phase distortions?

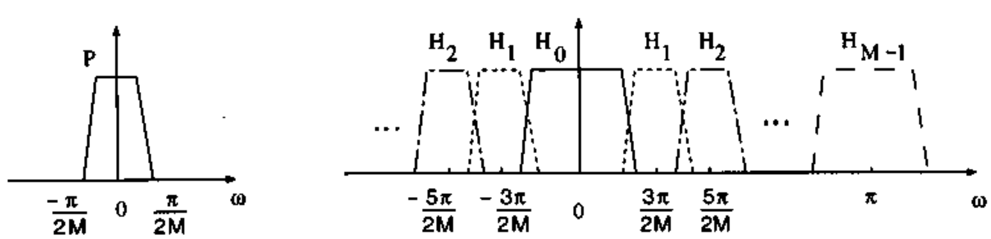
9. Verify that the GenLOT in equation (9.40) has linear phase.

10. Find the relations between the analysis functions $H_k'(z)$ and their factors $H_{ij}(z)$ for the 5-channel system in the Glossary (Tree-structured Filter Bank). When the H_{ij} are ideal filters, sketch the ideal responses of $H_k'(z)$.

11. With filter lengths $3M$, the time domain matrix H_b will have blocks $\underline{h}(0), \underline{h}(1), \underline{h}(2)$ in each row. Write down the $B = 3$ equations for orthogonality, corresponding to the two equations in (9.34).

9.4 Cosine-modulated Filter Banks

The idea of modulation keeps developing and improving. It is a way to build the whole filter bank around one filter—the prototype filter $p(n)$. Instead of modulating by exponentials, we modulate by cosines. An exponential will shift the frequency in one direction. The cosine is a sum of two exponentials, so the frequency shift goes two ways. The frequency band is partitioned symmetrically by cosine modulation, as shown below.



The original construction [Rothweiler] was chosen to cancel aliasing between adjacent subbands. In the z-domain, only the first aliasing error $T_1(z)$ was exactly zero. See *Cosine-modulated Filter Bank* in the Glossary. The aliasing between other channels, such as k and $k+2$, will be small when the prototype response dies quickly in the stopband (good attenuation at $|\omega| = \frac{\pi}{2M}$). But now, with better designs, all aliasing is gone. The filter bank can give perfect reconstruction.

Another step forward is in the length of the filters. The maximum length was originally $2M$, in the Modulated Lapped Transform (MLT). Now the length can be $2KM$, in the Extended Lapped Transform (ELT). Those developments by Malvar [Mr] have parallels from other authors as the efficiency of cosine modulation is appreciated. A very active area is the construction of *time-varying filter banks*, in which the filter changes length as the signal passes through. Only with a simple basic design could we maintain orthogonality while the overlapping filters change with time.

Cosine modulation can be analyzed in two domains. In the time domain, certain constant matrices are *orthogonal*. In the z-domain, certain polynomial matrices are *paraunitary*. It seems right to do both, but perhaps to emphasize the time domain. We begin there with the orthogonality conditions on these filter coefficients:

$$h_k(n) = f_k(n) = p(n)\sqrt{\frac{2}{M}}\cos\left[\left(k+\tfrac{1}{2}\right)\left(n+\frac{M+1}{2}\right)\frac{\pi}{M}\right] \tag{9.43}$$

Does $\boldsymbol{p} = (1, 1, \ldots, 1)$ give the ordinary DCT block transform? *No, because of the frequency shift*. The M in the numerator of (9.43) changes the phase of the DCT-IV. It also affects the orthogonality. In fact the usual orthogonality is gone. We see this directly for $M = 2$, by comparing the matrix $\boldsymbol{C}^{\text{IV}}$ with the new matrix $\boldsymbol{E}_0$:

$$(\boldsymbol{C}^{\text{IV}})_{kn} = \cos\left[\left(k+\tfrac{1}{2}\right)\left(n+\tfrac{1}{2}\right)\frac{\pi}{2}\right] = \begin{bmatrix} \cos(\frac{\pi}{8}) & \cos(\frac{3\pi}{8}) \\ \cos(\frac{3\pi}{8}) & \cos(\frac{9\pi}{8}) \end{bmatrix} = \begin{bmatrix} c & s \\ s & -c \end{bmatrix}$$

$$(\boldsymbol{E}_0)_{kn} = \cos\left[\left(k+\tfrac{1}{2}\right)\left(n+\tfrac{3}{2}\right)\frac{\pi}{2}\right] = \begin{bmatrix} \cos(\frac{3\pi}{8}) & \cos(\frac{5\pi}{8}) \\ \cos(\frac{9\pi}{8}) & \cos(\frac{15\pi}{8}) \end{bmatrix} = \begin{bmatrix} s & -s \\ -c & c \end{bmatrix}$$

Here $c = \cos(\frac{\pi}{8})$ and $s = \sin(\frac{\pi}{8})$. Then immediately $(\boldsymbol{C}^{\text{IV}})^T(\boldsymbol{C}^{\text{IV}}) = \boldsymbol{I}$ and $\boldsymbol{C}^{\text{IV}}$ is orthogonal. But $\boldsymbol{E}_0$ is not orthogonal:

$$\boldsymbol{E}_0^T\boldsymbol{E}_0 = \begin{bmatrix} s & -c \\ -s & c \end{bmatrix}\begin{bmatrix} s & -s \\ -c & c \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = \boldsymbol{I} - \boldsymbol{J}. \tag{9.44}$$

Here and always $\boldsymbol{J}$ is the reverse identity matrix. This convention is approaching the status of $\delta(n)$ and δ_{ij}, where definitions need not be repeated. Together with $\boldsymbol{E}_0$ in modulating a lapped transform comes the matrix $\boldsymbol{E}_1$ that saves orthogonality. With $M = 2$ this has the rest of the cosines:

$$(\boldsymbol{E}_1)_{kn} = \cos\left[\left(k+\tfrac{1}{2}\right)\left(n+2+\tfrac{3}{2}\right)\frac{\pi}{2}\right] = \begin{bmatrix} \cos(\frac{7\pi}{8}) & \cos(\frac{9\pi}{8}) \\ \cos(\frac{21\pi}{8}) & \cos(\frac{27\pi}{8}) \end{bmatrix} \tag{9.45}$$

$$\boldsymbol{E}_1^T\boldsymbol{E}_1 = \begin{bmatrix} -c & -s \\ -c & -s \end{bmatrix}\begin{bmatrix} -c & -c \\ -s & -s \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \boldsymbol{I} + \boldsymbol{J} \tag{9.46}$$

$$\boldsymbol{E}_1^T\boldsymbol{E}_0 = \begin{bmatrix} -c & -s \\ -c & -s \end{bmatrix}\begin{bmatrix} s & -s \\ -c & c \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \tag{9.47}$$

This pattern is extended in the Lemma below to all sizes M and to a longer sequence $\boldsymbol{E}_0, \ldots, \boldsymbol{E}_{2K-1}$. Here, we stay with $M = 2$ and $K = 1$ to see orthogonality all the way. Our prototype filter will be $\boldsymbol{p} = (1, 1, 1, 1)/\sqrt{2}$. Then the $2KM = 4$ coefficients in the $M = 2$ channels will be the cosines in E_0 and E_1 divided by $\sqrt{2}$. The time-domain matrix H_b, with the analysis channels interleaved to give the block diagonal form, has E_0 along the main diagonal and E_1 beside it:

$$H_b = \frac{1}{\sqrt{2}}\begin{bmatrix} \cdot & & & \\ E_1 & E_0 & & \\ & E_1 & E_0 & \\ & & & \cdot\ \cdot \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} \cdot & \cdot & & & & \\ -c & -c & s & -s & & \\ -s & -s & -c & c & & \\ & & -c & -c & s & -s \\ & & -s & -s & -c & c \\ & & & & \cdot & \cdot \end{bmatrix} \tag{9.48}$$

Notice how double-shift orthogonality has become *block orthogonality* (and here $M = 2$). In the block form, this orthogonality comes from $E_1^T E_0 = 0$ off the diagonal and $E_1^T E_1 + E_0^T E_0 = 2I$ on the diagonal. The J's cancel in $I + J + I - J$. So we have an orthogonal filter bank.

The prototype filter p was symmetric, as we always assume. But cosine modulation shifted the phase. The individual filters $(-c, -c, s, -s)$ and $(-s, -s, -c, c)$ are not linear phase. And they do not have the accuracy of the Daubechies filters. *We do not get high accuracy from cosine modulation.* In this present form we do not even get $p = 1$! The lowpass filter applied to the alternating vector $x(n) = (-1)^n$ does not give zero output. Nevertheless, the coding gain is adequate for compression, and the implementation is extremely fast.

M Channels with Filter Lengths $2KM$

In general we have $2K$ matrices $E_0, \ldots, E_{2K-1}$ and each matrix is $M \times M$:

$$(E_\ell)_{kn} = \sqrt{\frac{2}{M}}\cos\left[\left(k+\tfrac{1}{2}\right)\left(n + \ell M + \frac{M+1}{2}\right)\frac{\pi}{M}\right]. \tag{9.49}$$

These are the pieces of the extended cosine matrix $E = \left[E_0\, E_1\, \cdots\, E_{2K-1}\right]$, which would give the filter coefficients if the prototype filter had all $p(n) = 1$. We deal first with this case of a "box window". Then we allow other windows and discover the orthogonality conditions on $p(n)$.

The rows of E are orthogonal. Piece by piece we have the following identities.

Lemma: *The pieces of the extended cosine matrix E satisfy*

$$\begin{matrix} E_{2s} = (-1)^s E_0 \\ E_{2s+1} = (-1)^s E_1 \end{matrix} \quad \text{and} \quad \begin{matrix} E_\ell^T E_{\ell+2s} = (-1)^s \left[I + (-1)^{\ell+1} J\right] \\ E_\ell^T E_{\ell+2s+1} = \mathbf{0}. \end{matrix} \tag{9.50}$$

Proof: With $\ell = 2s$, we have added $(k+\frac{1}{2})(2sM)\frac{\pi}{M}$ to the arguments of the cosine in (9.49). This is the same as adding $s\pi$ which multiplies the cosine by $(-1)^s$. Similarly E_{2s+1} is a shift by $s\pi$ from E_1. The two identities on the left are proved. The identities on the right then quickly reduce to the first cases

$$E_0^T E_0 = I - J \quad \text{and} \quad E_1^T E_1 = I + J \quad \text{and} \quad E_1^T E_0 = \mathbf{0}. \tag{9.51}$$

Those were checked above for $M = 2$. For larger M, the $(n, \acute{n})$ entry of $E_0^T E_0$ is

$$\frac{2}{M}\sum_{k=0}^{M-1}\cos\left[\left(k+\tfrac{1}{2}\right)\left(n+\frac{M+1}{2}\right)\frac{\pi}{M}\right]\cos\left[\left(k+\tfrac{1}{2}\right)\left(\acute{n}+\frac{M+1}{2}\right)\frac{\pi}{M}\right] =$$

$$\sum_{k=0}^{M-1}\frac{1}{M}\cos\left[\left(k+\tfrac{1}{2}\right)\left(n-\acute{n}\right)\frac{\pi}{M}\right] + \frac{1}{M}\cos\left[\left(k+\tfrac{1}{2}\right)\left(n+\acute{n}+M+1\right)\frac{\pi}{M}\right]. \tag{9.52}$$

On line two, we wrote $(\cos a)(\cos b)$ using $a+b$ and $a-b$. Its first sum is $\delta(n-\acute{n})$, which gives the matrix I. Its second sum is zero, except when $n+\acute{n}=M-1$ and all terms are $\cos\pi=-1$. This gives the matrix $-J$ in $E_0^T E_0$. Note that the reverse identity has 1's when $n+\acute{n}=M-1$, because the numbering starts at zero. The identities $E_1^T E_1 = I + J$ and $E_1^T E_0 = 0$ have similar proofs.

Orthogonality Conditions on the Prototype Filter $p(n)$

Now multiply the cosines by $p(n)$ to get the true filter coefficients $h_k(n)$. The cosine matrix E is M by $2KM$, with columns indexed by n. It is multiplied by the diagonal matrix $P = \operatorname{diag}(p(0), p(1), \ldots, p(2KM-1))$. Their product EP splits into $2K$ square blocks $E_\ell \underline{P}_\ell$ of size M:

$$EP = \begin{bmatrix} E_0 & \cdots & E_{2K-1} \end{bmatrix} \begin{bmatrix} \underline{P}(0) & & \\ & \ddots & \\ & & \underline{P}(2K-1) \end{bmatrix} = \begin{bmatrix} E_0\underline{P}_0 & \cdots & E_{2K-1}\underline{P}_{2K-1} \end{bmatrix} \quad (9.53)$$

The entries of EP are cosines multiplied by $p(n)$. In other words we have $h_k(n)$.

The square block $\underline{P}_\ell$ contains the prototype coefficients $p(\ell M), \ldots, p(\ell M + M - 1)$. We are simply writing out a typical block row of the analysis bank matrix H_b. That row is EP as in (9.53), except the columns come in reverse order because $h_k(0)$ is at the right end of the row. This has no effect on orthogonality.

Under what condition is H_b an orthogonal matrix? The center block of $H_b^T H_b$ must be the identity and the off-diagonal blocks of $H_b^T H_b$ must be zero:

$$\sum_{0}^{2K-1} (E_\ell \underline{P}_\ell)^T (E_{\ell+i}\underline{P}_{\ell+i}) = \delta(i) I. \quad (9.54)$$

Inside those sums are the products $E_\ell^T E_{\ell+i}$ that are given by the Lemma. Substituting the identity (9.50) yields the orthogonality condition on the matrices $\underline{P}_\ell$. That gives us the coefficients $p(n)$ of the (symmetric) prototype filter.

Theorem 9.6 *The cosine modulated filter bank is orthogonal if and only if the diagonal blocks $\underline{P}_l$ are double-shift orthogonal:*

$$\sum \underline{P}_\ell^T \underline{P}_{\ell+2s} = \delta(s) I. \quad (9.55)$$

For the coefficients this means that we shift by double blocks:

$$\sum_{\ell=0}^{2K-2s-1} p(n+\ell M)\, p(n+\ell M + 2sM) = \delta(s) \;\; \text{for } n = 0, \ldots, \tfrac{M}{2} - 1. \quad (9.56)$$

The special case $K=1$ has only two blocks $\underline{P}_0$ and $\underline{P}_1$. So there is only $s=0$:

$$p^2(n) + p^2(n+M) = 1. \quad (9.57)$$

That corresponds exactly to the condition $g^2(t)+g^2(t+1)=1$ on the window function in Section 8.4. That continuous-time theory only allowed neighboring windows to overlap. In discrete time this is $K=1$.

The next discrete case $K = 2$ has blocks $\underline{P}_0, \underline{P}_1, \underline{P}_2, \underline{P}_3$. Now there are conditions from $s = 0$ and $s = 1$:

$$p^2(n) + p^2(n+M) + p^2(n+2M) + p^2(n+3M) = 1$$

$$p(n)p(n+2M) + p(n+M)p(n+3M) = 0. \tag{9.58}$$

We can design $4M$ lowpass coefficients to satisfy these conditions!

Proof: Theorem 9.6 is our main result. The analysis of cosine modulation led here. The conditions (9.54) for odd i are automatic because $\boldsymbol{E}_\ell^T \boldsymbol{E}_{\ell+i} = 0$ in (9.50). The conditions for even $i = 2s$ require the substitution of $(-1)^s[\boldsymbol{I} + (-1)^{\ell+1}\boldsymbol{J}]$ for $\boldsymbol{E}_\ell^T \boldsymbol{E}_{\ell+2s}$. The $\boldsymbol{J}$'s cancel when we sum on ℓ because the prototype filter is symmetric. (You see why the filter length $2KM$ is an *even* multiple of M, to have an even number of $\boldsymbol{J}$'s with alternating sign.) Then (9.55) and (9.56) reduce to $\sum \underline{P}_\ell^T \underline{P}_{\ell+2s}$. For orthogonality this must give $\delta(s)$ in (9.57) and (9.58).

Two symmetric solutions to $p^2(n) + p^2(n+M) = 1$ are the sine windows

$$p(n) = \sin\left[\frac{n\pi}{2(M-1)}\right] \quad \text{and} \quad p(n) = -\sin\left[\left(n + \tfrac{1}{2}\right)\frac{\pi}{2M}\right]. \tag{9.59}$$

Malvar observed that the latter is the only prototype for which the resulting nth filter has frequency response $\delta(n)$ at $\omega = 0$. (The accuracy is $p = 1$.) The zeroth filter reproduces a DC input and the other filters null it, with no DC leakage. This normalization was expected in the rest of the book and is desirable in image coding. Malvar also gives a family of solutions when $K = 2$.

Polyphase and Lattice Structure

The polyphase coefficients are the blocks $\boldsymbol{E}_0\underline{P}_0, \ldots, \boldsymbol{E}_{2K-1}\underline{P}_{2K-1}$ along each row of the time domain matrix $\boldsymbol{H}_b$. The filter bank is orthogonal when this polyphase matrix $\sum \boldsymbol{E}_\ell \underline{P}_\ell z^{-\ell}$ is paraunitary:

$$\boldsymbol{H}_p^T(z^{-1})\boldsymbol{H}_p(z) = \left(\sum \underline{P}_\ell^T \boldsymbol{E}_\ell^T z^\ell\right)\left(\sum \boldsymbol{E}_\ell \underline{P}_\ell z^{-\ell}\right) = \boldsymbol{I}. \tag{9.60}$$

The constant term in the product is $\sum \underline{P}_\ell^T \boldsymbol{E}_\ell^T \boldsymbol{E}_\ell \underline{P}_\ell$. Equation (9.55) makes this the identity matrix $\boldsymbol{I}$. The coefficient of z^{-i} in the product is $\sum \underline{P}_\ell^T \boldsymbol{E}_\ell^T \boldsymbol{E}_{\ell+i} \underline{P}_{\ell+i}$. Equation (9.55) makes this the zero matrix. Under these conditions $\boldsymbol{H}_p(z)$ is paraunitary.

For $K = 1$ the polyphase matrix simplifies to $\boldsymbol{E}_0\underline{P}_0 + z^{-1}\boldsymbol{E}_1\underline{P}_1$. It is paraunitary, after using $\boldsymbol{E}_1^T\boldsymbol{E}_0 = \boldsymbol{0}$ and $\boldsymbol{E}_1^T\boldsymbol{E}_1 = \boldsymbol{I} + \boldsymbol{J}$ and $\boldsymbol{E}_0^T\boldsymbol{E}_0 = \boldsymbol{I} - \boldsymbol{J}$, when

$$\underline{P}_1^T \underline{P}_1 + \underline{P}_0^T \underline{P}_0 = \boldsymbol{I}. \tag{9.61}$$

This pairs off the prototype filter coefficients into equation (9.57).

Let $G_k(z), 0 \le k \le 2M-1$ be the polyphase components of the prototype filter $P(z)$. The pairs $(G_k(z), G_{M+k}(z))$ of a paraunitary cosine-modulated filter bank satisfy

$$G_k(z^{-1})\,G_k(z) \;+\; G_{M+k}(z^{-1})\,G_{M+k}(z) \;=\; \frac{1}{2M}. \tag{9.62}$$

This agrees with the orthogonality condition in a two-channel filter bank. The conditions (9.62) extend to filters of any length [NgKoil]. Thus, the cosine-modulated bank can be implemented by two-channel paraunitary filter banks in parallel, as depicted in the Glossary (see *Tree-structured Filter Bank*). This is efficient because of the lattice structures associated with the pair $[G_k(z), G_{k+M}(z)]$ and the matrix $\widehat{C}$.

Example 9.5. Figure 9.15 shows the frequency responses of an 8-channel PR cosine-modulated filter bank. The filter length is 128. The resulting filters have stopband attenuation about 80 dB. The design procedure is based on QCLS. Chapter 10 will elaborate more on this formulation.

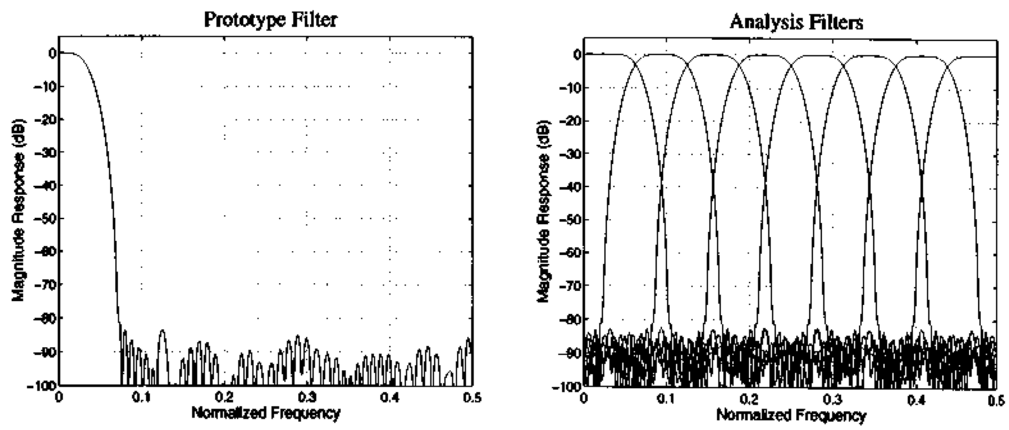


Figure 9.15: Frequency response plots for an 8-channel PR cosine-modulated filter bank.

Note 1 A pseudo-QMF bank is a cosine-modulated filter bank where the first aliasing transfer function $T_1(z)$ is cancelled. The other $T_2(z), \ldots, T_{M-1}(z)$ are minimized by constraining the stopband cutoff frequency $\omega_c \leq \frac{\pi}{M}$ of $P(z)$ such that there is no overlap between $H_k(z)$ and $H_{k\pm2}(z)$. Consequently, the aliasing error is comparable to the stopband attenuation of the prototype filter. The distortion function $T_0(z)$ is not a delay. The design procedure is to find a prototype filter that minimizes the following weighted objective function:

$$\alpha \int_{\omega_s}^{\pi} \left|P(e^{j\omega})\right|^2 d\omega + (1-\alpha) \int_0^{\pi} \left|T_0(e^{j\omega}) - e^{-jn_0\omega}\right|^2 d\omega$$

where $0 \leq \alpha \leq 1$. One often has to trade off aliasing with distortion.

Note 2 The prototype filter $P(z)$ for an NPR Pseudo-QMF bank is a *linear phase spectral factor of a $2M$-th band filter*. There is no distortion. The aliasing errors can be minimized by high attenuation.

Example 9.6. Figure 9.16 shows the frequency responses of a 16-channel NPR cosine-modulated filter bank. The filter length is 256. By constraining $P(z)$ to be the spectral factor of a 32-band filter, the resulting filter bank has no magnitude nor phase distortion. The only distortion here is the aliasing, which is less than or equal to -72 dB.

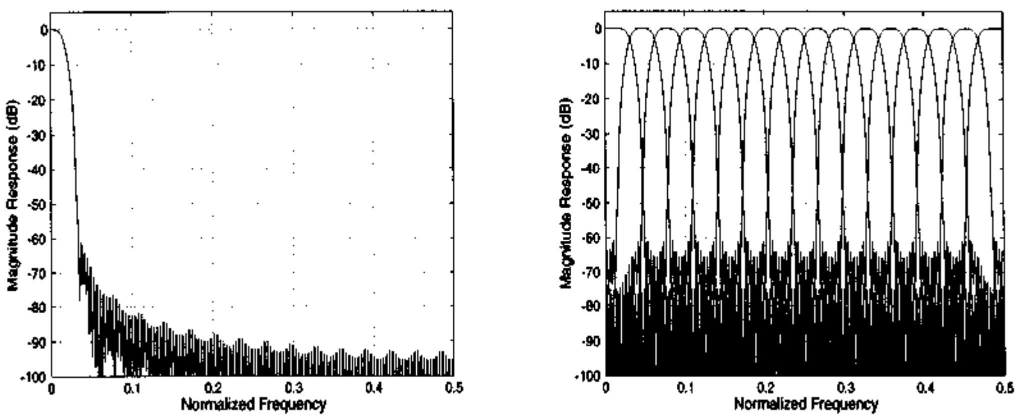


Figure 9.16: Frequency response plots for the 16-channel NPR cosine-modulated filter bank. Prototype filter $P(z)$ (left) and Analysis filters $H_k(z)$ (right).

Problem Set 9.4

1. Verify that Malvar's windows (9.59) satisfy $p^2(n) + p^2(n+M) = 1$ for orthogonality.
2. Construct explicitly (by hand or by Matlab) the 3×3 matrices C^{IV}, E_0, E_1, E_2 for $M = 3$. Check by Matlab whether the identities (9.51) still hold for odd M.
3. For prototypes $p(n)$ in synthesis and $\widetilde{p}(n)$ in analysis (so that $H_k \neq F_k$), what will be the conditions for a perfect reconstruction cosine-modulated filter bank (filter length $2M$ and eventually $2KM$)?
4. Let $p(n)$ have linear phase. Show that $h_k(n)$ in (9.43) cannot have linear phase.
5. Let $p(n)$ be a spectral factor of a $2M$-th band filter. Show that the distortion function $T_0(z) = \sum_{k=0}^{M-1} F_k(z)H_k(z)$ is a delay.
6. Find the conditions on the prototype filters $p(n)$ for a *biorthogonal* cosine-modulated filter bank. Cheung's MIT thesis has shown that with symmetry, there are $M/2$ extra free parameters compared to the orthogonal case.

9.5 Multidimensional Filters and Wavelets

Images are two-dimensional. Processing those images is an extremely important application of subband filtering. We certainly need two-dimensional filters! Their construction is coming late in the book because it is either quite easy or quite hard — depending on the type of filter:

1. *Separable*: Products of one-dimensional filters (easy).
2. *Nonseparable*: Genuinely two-dimensional filters (hard).

Each filter bank is associated with a subsampling matrix M. In d dimensions this matrix is $d \times d$. In one dimension it contains the usual number M. In two dimensions we consider the two leading possibilities:

$$\text{(separable) } M_s = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \quad \text{(nonseparable quincunx) } M_q = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}.$$

These matrices have eigenvalues $|\lambda| > 1$, as required. Their determinants are $M_s = 4$ and $M_q = 2$. The lightface symbol M still represents the number of channels. M is also the number of scaling functions plus wavelets.

In two dimensions, a filter is a two-dimensional convolution $\boldsymbol{y} = \boldsymbol{h} * \boldsymbol{x}$. In the $\omega = (\omega_1, \omega_2)$ and $z = (z_1, z_2)$ domains, we multiply by $H(\omega_1, \omega_2)$ and $H(z_1, z_2)$:

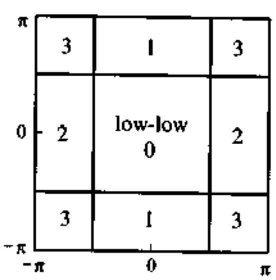
$$\boldsymbol{Hx}\,(n_1, n_2) = \sum_{k_1}\sum_{k_2} h(k_1, k_2)\, x(n_1 - k_1, n_2 - k_2)$$

$$H(\omega)\,X(\omega) = \left(\sum\sum h(k_1, k_2) e^{-i(k_1\omega_1 + k_2\omega_2)}\right)\left(\sum\sum x(n_1, n_2) e^{-i(n_1\omega_1 + n_2\omega_2)}\right)$$

$$H(z)\,X(z) = \left(\sum\sum h(k_1, k_2) z_1^{-k_1} z_2^{-k_2}\right)\left(\sum\sum x(n_1, n_2) z_1^{-n_1} z_2^{-n_2}\right).$$

The frequency response $H(\omega_1, \omega_2)$ has period 2π in both variables. A set of four brick wall filters will cover the period square with no overlap. These ideal filters are (0) low-low, (1) low-high, (2) high-low, and (3) high-high:

Support regions for four ideal 2D filters



Those ideal filters are IIR. A set of four FIR separable filters is easily constructed from one-dimensional filters $\boldsymbol{h}_{low}$ and $\boldsymbol{h}_{high}$:

$$\boldsymbol{h}_0(n_1, n_2) = \boldsymbol{h}_{low}(n_1)\,\boldsymbol{h}_{low}(n_2) \qquad \boldsymbol{h}_2(n_1, n_2) = \boldsymbol{h}_{high}(n_1)\,\boldsymbol{h}_{low}(n_2)$$

$$\boldsymbol{h}_1(n_1, n_2) = \boldsymbol{h}_{low}(n_1)\,\boldsymbol{h}_{high}(n_2) \qquad \boldsymbol{h}_3(n_1, n_2) = \boldsymbol{h}_{high}(n_1)\,\boldsymbol{h}_{high}(n_2).$$

This is the easy construction. After the filters, we sample by $(\downarrow 2)$ *in each direction*. This subsampling corresponds to the separable matrix $\boldsymbol{M}_s = 2\boldsymbol{I}$:

$$(\downarrow \boldsymbol{M}_s)\,\boldsymbol{y}(n_1, n_2) = \boldsymbol{y}(\boldsymbol{Mn}) = \boldsymbol{y}(2n_1, 2n_2). \qquad (9.63)$$

We are keeping one sample out of four. The four filters in four channels give critical sampling. The polyphase matrix will be 4×4 (generally $M \times M$).

Compare with the nonseparable quincunx filter bank. With $M_q = 2$ filters, the quincunx rule keeps the samples for which $n_1 + n_2$ is even:

$$(\downarrow \boldsymbol{M}_q)\,\boldsymbol{y}(n_1, n_2) = \boldsymbol{y}(\boldsymbol{M}_q\boldsymbol{n}) = \boldsymbol{y}(n_2 + n_1, n_2 - n_1). \qquad (9.64)$$

Figure 9.17 shows the lattice of integers and the quincunx sublattice of samples:

Notice how $M_s = 4$ separable sublattices cover the whole lattice. Similarly $M_q = 2$ quincunx sublattices (*staggered grids*) cover all mesh points. The separable lattices are strongly oriented in the horizontal and vertical directions! The quincunx lattice has an extra symmetry at 45° and −45°. It is closer to isotropic, meaning independent of direction. A diagonal edge in an

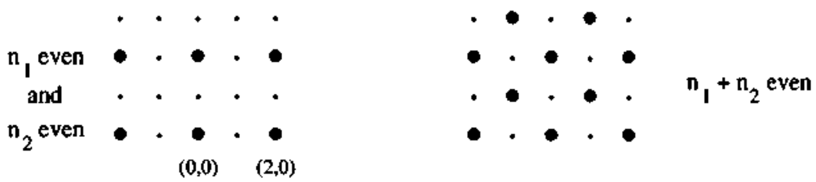


Figure 9.17: The sampling lattices for M_s (separable) and M_q (quincunx).

image will be captured far better by the quincunx. But those filters are harder to design, if we want PR and especially if we want orthogonality.

The reader will see how other sampling matrices M give the lattice points Mn. Downsampling keeps all values $y(Mn)$ on this lattice. Then the upsampling step $(\uparrow M)$ assigns a zero when (n_1, n_2) is not in the lattice. As always, *$(\uparrow M)$ is the transpose of $(\downarrow M)$*. If we apply both, we get the identity operator on the lattice and otherwise zero:

$$(\uparrow M)\,(\downarrow M) y(n_1, n_2) = \begin{cases} y(n_1, n_2) & \text{on the lattice} \\ 0 & \text{for other } (n_1, n_2). \end{cases} \tag{9.65}$$

A synthesis filter in each channel completes the filter bank: all normal.

Example 1 (Separable Haar): The Haar filter will be typical of separable filters. The four filters have coefficients $\pm\frac{1}{4}$ and the low-low response is $H_0(\omega_1, \omega_2) = \frac{1}{4}(1 + e^{-i\omega_1})(1 + e^{-i\omega_2})$. We "block" the input four samples at a time, with $x(0,0)$, $x(0,1)$, $x(1,0)$, $x(1,1)$ in the zeroth block. Or in practice, the signal goes through ordinary Haar in the x direction and then in the y direction. The two-dimensional Haar bank is a block transform with 4×4 blocks:

$$\text{Haar block} = \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & -1 & -1 \end{bmatrix}$$

This is just the two-dimensional DFT matrix. It is also the polyphase matrix. One constant term only, because Haar has no overlap. It is an orthogonal filter bank!

The Haar block shows a tensor product $H_{2\times2} \otimes H_{2\times2}$ of one-dimensional DFT's. The matrix $H_{2\times2} = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ is multiplied by each of its four entries to give the four 2×2 subblocks. (A tensor product has order $N_1 N_2$ when the matrices have orders N_1 and N_2. The matrix of order N_1 appears N_2 times on each block row of the tensor product, multiplied by entries from the matrix of order N_2.) For separable filters, the 4×4 polyphase matrix is always the tensor product $H_p(z_1) \otimes H_p(z_2)$ of 2×2 one-dimensional polyphase matrices.

Every separable filter bank inherits the properties of its one-dimensional factors. The bank is orthogonal with accuracy (p, p) if the factors are orthogonal with accuracy p. *The filters cannot also be linear phase* (except for Haar). It is possible [KarVet] to achieve linear phase and orthogonality with nonseparable filters.

Example 2 (Nonseparable): Quincunx sampling has one alias term because $M - 1 = 1$. When $(\downarrow M_q)$ is followed by $(\uparrow M_q)$ as in (9.65), that alias is $X(-z_1, -z_2)$. The quincunx modulation matrix is 2×2:

$$\boldsymbol{H}_m(z_1, z_2) = \begin{bmatrix} H_0(z_1, z_2) & H_0(-z_1, -z_2) \\ H_1(z_1, z_2) & H_1(-z_1, -z_2) \end{bmatrix}.$$

Orthogonality requires that $\boldsymbol{H}_m(z_1, z_2)\, \boldsymbol{H}_m^T(z_1^{-1}, z_2^{-1}) = 2\boldsymbol{I}$. The $(1, 1)$ entry is

$$H_0(z_1, z_2) H_0(z_1^{-1}, z_2^{-1}) + H_0(-z_1, -z_2) H_0(-z_1^{-1}, -z_2^{-1}) \equiv 2. \tag{9.66}$$

The product filter $P(z_1, z_2)$ must be halfband! The highpass H_1 comes from H_0 by an *alternating flip* in z_1 and z_2:

$$H_1(z_1, z_2) = \text{ odd 2D delay of } H_0(-z_1^{-1}, -z_2^{-1}). \tag{9.67}$$

All looks familiar. But in two dimensions there is one enormous difference. *We cannot factor most product filters.* Even if the symmetric polynomial $P(\omega_1, \omega_2)$ is nonnegative, it may be impossible to express it as the square $|H(\omega_1, \omega_2)|^2$ of a polynomial. The idea of computing zeros of $P(z)$ and separating them into two factors is strictly one-dimensional. We must design the filters directly, and not by factorization.

A cascade structure [VK, p. 182] can give orthogonality or linear phase:

$$\boldsymbol{H}_p(z_1, z_2) = \boldsymbol{R}_{2j} \begin{bmatrix} 1 & \\ & z_2^{-1} \end{bmatrix} \boldsymbol{R}_{2j-1} \begin{bmatrix} 1 & \\ & z_1^{-1} \end{bmatrix} \cdots \boldsymbol{R}_0.$$

The filter bank is orthogonal if the matrices $\boldsymbol{R}$ are orthogonal. The shortest lowpass filter $\boldsymbol{h}$ has parameters a, b, c in its eight nonzero coefficients:

$$\boldsymbol{h} = \begin{bmatrix} & -b & -ab & \\ -c & -ac & -a & 1 \\ & abc & -abc & \end{bmatrix}. \tag{9.68}$$

With constraints on a, b, c this is analogous to the Daubechies D_4 filter.

Example 3 (McClellan transformation): A convenient way to design linear phase 2D filter banks is to begin with a symmetric centered 1D filter:

$$H(\omega) = \sum_{-L}^{L} h(n) \cos n\omega = \sum_{-L}^{L} h(n) T_n(\cos\omega).$$

The Tchebycheff polynomial T_n produces $\cos n\omega$ from powers of $\cos\omega$. For example $\cos 2\omega = 2\cos^2\omega - 1$, so that $T_2(x) = 2x^2 - 1$. The McClellan transformation replaces $\cos\omega$ by a symmetric 2D filter response $F(\omega_1, \omega_2)$. Then $H(\omega_1, \omega_2) = \sum h(n) T_n(F(\omega_1, \omega_2))$ is still symmetric. In the quincunx case we choose $F = \frac{1}{2}(\cos\omega_1 + \cos\omega_2)$. More general polynomials than T_n have been used effectively in [TayKing].

By designing the 1D coefficients in $H(\omega)$ we get a good linear phase 2D filter $H(\omega_1, \omega_2)$. When the former gives perfect reconstruction with FIR inverse, so does the latter. (The 2D determinant is a monomial when the 1D determinant is. We cannot have orthogonality too.) The accuracy is also preserved, because a factor $(1+\cos\omega)^p$ transforms to $(1+\frac{1}{2}(\cos\omega_1+\cos\omega_2))^p$. There are p zeros at the alias frequency (π, π) in the (ω_1, ω_2) plane.

Note that a separable filter with response $H(\omega_1)H(\omega_2)$ has p zeros at all three of the alias frequencies $(0, \pi)$ and $(\pi, 0)$ and (π, π). The zeros are needed as in 1D for stability of the iterated lowpass filter and convergence to the scaling function.

Dilation Equation and Wavelets

The lowpass filter has coefficients $h_0(k_1, k_2)$. When we iterate with rescaling, the cascade algorithm hopes to converge to the scaling function. The limiting equation when $\phi^{(i+1)}$ and $\phi^{(i)}$ approach $\phi(t)$ is the dilation equation:

$$\phi(t_1, t_2) = M \sum \boldsymbol{h}_0(n_1, n_2)\phi(\boldsymbol{M}\boldsymbol{t} - \boldsymbol{n}). \tag{9.69}$$

Here $\boldsymbol{t} = (t_1, t_2)$ and $\boldsymbol{n} = (n_1, n_2)$ are column vectors. The matrix $\boldsymbol{M}$ has determinant M. When we change variables from $\boldsymbol{Mt} - \boldsymbol{n}$ to $\boldsymbol{\tau}$, this determinant preserves the double integral:

$$M \iint \phi(\boldsymbol{Mt} - \boldsymbol{n})dt_1dt_2 = \iint \phi(\boldsymbol{\tau})d\tau_1 d\tau_2. \tag{9.70}$$

The lowpass normalization is still $\sum\sum \boldsymbol{h}_0(n_1, n_2) = 1$. This filter leads to the scaling function. The other $M - 1$ filters lead to $M - 1$ wavelets by the usual wavelet equation:

$$w_k(\boldsymbol{t}) = M \sum\sum \boldsymbol{h}_k(n_1, n_2)\phi(\boldsymbol{Mt} - \boldsymbol{n}), \quad 1 \le k < M. \tag{9.71}$$

Orthogonality will mean that the translates $\phi(\boldsymbol{t} - \boldsymbol{n})$ at scale zero are an orthonormal basis for the space V_0. The wavelet translates $w_1(\boldsymbol{t} - \boldsymbol{n}), \ldots, w_{M-1}(\boldsymbol{t} - \boldsymbol{n})$ are an orthonormal basis for W_0. The wavelet translates and dilates $M^{j/2}w_k(\boldsymbol{M}^j\boldsymbol{t} - \boldsymbol{n})$ are an orthonormal basis for the whole finite energy space $L^2(\mathbf{R}^2)$.

Notice the dilation matrix $\boldsymbol{M}$! We nearly wrote 2^j instead of $\boldsymbol{M}^j$. This would be correct only for the matrix $\boldsymbol{M} = 2\boldsymbol{I}$ that gives separable filters. The iteration then gives a separable scaling function and three separable wavelets:

$$\begin{array}{llll} \text{low-low:} & \phi(\boldsymbol{t}) = \phi(t_1)\phi(t_2) & \text{high-low:} & w_2(\boldsymbol{t}) = w(t_1)\phi(t_2) \\ \text{low-high:} & w_1(\boldsymbol{t}) = \phi(t_1)w(t_2) & \text{high-high:} & w_3(\boldsymbol{t}) = w(t_1)w(t_2) \end{array} \tag{9.72}$$

It is pleasant to verify (Problem 3) that separable filters give these separable solutions to the dilation equation and wavelet equation. With accuracy p in the one-dimensional filter and $1, t, \ldots, t^{p-1}$ in its scaling space V_0, the separable 2D filter (product filter) will inherit this accuracy. All p^2 of the polynomials $(t_1)^r(t_2)^s$ will be in V_0, for $r < p$ and $s < p$. In the special case of splines, $\phi(\boldsymbol{t})$ is a product $\phi(t_1)\phi(t_2)$ of one-dimensional B-splines. The coefficients $\boldsymbol{h}(n_1, n_2)$ are products of binomial coefficients. These are just the coefficients in $H(z_1, z_2) = (1+z_1^{-1})^p(1+z_2^{-1})^p$.

Note that a quincunx filter does not need this factor. It is $[1 + \frac{1}{2}(z_1^{-1} + z_2^{-1})]^p$ that gives accuracy p for quincunx. We would construct $\phi(t_1, t_2)$ by the cascade algorithm. Then $(t_1)^r(t_2)^s$ is locally in V_0 if $r + s < p$. For $p = 2$ we only need the linear polynomials $1, t_1, t_2$ and not the extra t_1t_2 that comes with separable filters.

Remark We believe that ***box splines*** could lead to multidimensional algorithms. They are generated by filters for which $H(z_1, z_2)$ has simple factors. The factorization of H is a fundamental 2D difficulty, and box splines are direct constructions with known factors. The future will determine whether these (or other) multidimensional filters are successful.

Bounded Domains

For a bounded interval in one dimension, Section 8.5 proposed several constructions of boundary functions. Local support and the multiresolution property $V_j \subset V_{j+1}$ and even the polynomial accuracy p were preserved. The same properties are desirable in two dimensions but not so easy to achieve.

We note one immediate difficulty. The boundary scaling functions $\phi_b(t)$ were differences between monomials t^k and combinations $\sum y_k\phi(t-k)$ of interior functions. The combination reproduces t^k exactly except near the end of the interval. In one dimension, $\phi_b(t)$ will have local support. But in two dimensions the support will be a *thin ring* along the boundary — not local! A more careful construction is needed.

The new paper [CoDaDe] gives a simple local construction of boundary functions for a square domain, starting from separable wavelets. On a general domain their method is less simple. It seems better than earlier constructions, and this problem must be faced in solving partial differential equations by a wavelet method (Section 11.6).

Problem Set 9.5

1. Show that the quincunx upsampling $\boldsymbol{u}(\boldsymbol{n}) = \boldsymbol{x}(\boldsymbol{M}_q^{-1}\boldsymbol{n})$ yields

$$U(\omega_1,\omega_2) = \tfrac{1}{2}\left[X(\omega_1,\omega_2) + X(\omega_1+\pi,\omega_2+\pi)\right].$$

Express this also in the (z_1, z_2) domain. Which inputs give $\boldsymbol{u} = \boldsymbol{0}$?

2. For $(\downarrow\ \boldsymbol{M}_s)$ the first row of the modulation matrix contains $H_0(z_1,z_2)$, $H_0(-z_1,z_2)$, $H_0(z_1,-z_2)$, $H_0(-z_1,-z_2)$. Using products $H_0(z_1)H_0(z_2)$ of one-dimensional filters, write out the 4×4 modulation matrix $\boldsymbol{H}_m(z_1,z_2)$. How is it related to the 2×2 matrices $\boldsymbol{H}_m(z_1)$ and $\boldsymbol{H}_m(z_2)$?

3. With the separable filters in equation (9.72), show that the separable scaling function $\phi(t_1)\phi(t_2)$ solves the 2D dilation equation with $\boldsymbol{M} = 2\boldsymbol{I}$.

4. Find $M = \det \boldsymbol{M}$ and draw the lattices of vectors $\boldsymbol{Mn}$ for

$$\boldsymbol{M}_{\text{all}} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \text{ and } \boldsymbol{M} = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \text{ and } \boldsymbol{M}_{\text{hex}} = \begin{bmatrix} 1 & 2 \\ -1 & 2 \end{bmatrix}$$

5. For $\boldsymbol{M}_s$ and $\boldsymbol{M}_q$ draw the *Voronoi cell* of points whose closest lattice point is the origin. What is the area of the Voronoi cell?

6. If the lowpass filter $\boldsymbol{h}_0(n_1,n_2)$ satisfies the quincunx orthogonality condition (9.66), show that $H_1(z_1,z_2) = -z_1^{-1}H_0(-z_1^{-1},-z_2^{-1})$ gives an orthogonal highpass filter. The modulation matrix should have $\boldsymbol{H}_m\boldsymbol{H}_m^T = 2\boldsymbol{I}$.

7. A 2D analogue of the hat function $\phi(t)$ is the bilinear tent $\phi(x)\phi(y)$. Find its Fourier transform as a function of ω_x and ω_y.

8. Another 2D analogue of the hat function is the Courant finite element $C(x,y)$. It is linear in each triangle obtained from grid lines $x = k_1$, $y = k_2$, $y - x = k_3$. Draw the six triangles around the origin and show that the function that is zero outside, with $C(0,0) = 1$, has transform

$$\widehat{C}(\omega_1,\omega_2) = \left(\frac{\sin(\omega_1/2)}{\omega_1/2}\right)\left(\frac{\sin(\omega_2/2)}{\omega_2/2}\right)\left(\frac{\sin(\omega_1+\omega_2)/2}{(\omega_1+\omega_2)/2}\right).$$

9. Take the Fourier transform of the dilation equation (9.69). By recursion find the infinite product formula (notice the transpose) $\widehat{\phi}(\boldsymbol{\omega}) = \prod H((\boldsymbol{M}^{-j})^T\boldsymbol{\omega})$.

10. What is the z-transform of $(\uparrow M)(\downarrow M)\boldsymbol{y}$?