Signal Processing Systems (521279S)
Part 5(c) : Filter banks

Version 1.0

November 25, 2025

Goals of the study. The document considers efficient implementation of filter banks. Shortly, the main idea is to modulate a prototype filter. There exists a relationship between modulation and polyphase decomposition, which allows derivation of computationally efficient structures.

Contents

1 Filter banks and basics of their implementation **2**
- 1.1 Critical sampling 2
- 1.2 Prototype-based design 3
- 1.3 Modulation representation 3
- 1.4 Link to polyphase decomposition 4

2 Two-channel filter banks **6**
- 2.1 Perfect reconstruction 7
- 2.2 QMF banks 7
- 2.3 Polyphase implementation 8
  - 2.3.1 General case 8
  - 2.3.2 Special case: QMF banks 9

3 Some M-channel filter bank structures **10**
- 3.1 Tree-structured filter banks 10
- 3.2 Parallel-structured filter banks 11
  - 3.2.1 Complex modulated filter banks 11
  - 3.2.2 Cosine-modulated filter banks 12
  - 3.2.3 Example: MPEG-1 audio 13

History
V1.0 11.10.2019 .

1 Filter banks and basics of their implementation

A filter bank consists of two or more filters which have directly adjacent frequency bands. There are two kinds of filter banks:

1. *Analysis filter banks* decompose signal spectra in a certain number of bands;
2. *Synthesis filter banks* do the inverse operation reconstructing the signal from the decomposed spectra.

A subband coding (SBC) filter bank consist of an analysis filter bank, which is followed by a synthesis filter bank. Fig. 3 in Chapter 2 illustrates the two-channel case.[1]

The idea of SBC is that when some subbands of a signal are more important than others, but also the less important subbands contain some important information that cannot be simply thrown away by filtering, we must follow a strategy where the number of bits allocated to subbands depends on their importance. We do it by splitting the whole signal to subband components (this is called analysis), and encode (quantize) subbands separately. As a result, available bit bandwidth can be exploited more efficiently and (perceptual) quality of the reconstructed signal is higher.

1.1 Critical sampling

If spectrum of a signal is limited to a band having width $< \pi/M$, the signal can be downsampled by the factor M, and the original signal is still recoverable according to the Nyquist sampling theorem.

In filter banks, *critical sampling* refers to systems where input and output sample rates are equal. For example, if we have an analysis filter bank, which has M output channels each producing sampled values at the rate corresponding to the input sample rate, then critical sampling refers to the downsampling of output channels by M.

As filters in practice are not ideal, frequency responses of analysis filter bank subband filters are typically overlapped. This implies that downsampling operation can cause aliasing effects. However, it is possible to design filter banks so that aliasing effects are negligible or even completely cancelled in the synthesis.[2] When aliasing is completely cancelled, and transfer characteristics of the system correspond merely to a delay, the SBC filter bank is said to provide *perfect reconstruction* (see Sec. 2.1).

[1] When synthesis is followed by analysis, we have a transmultiplexer for frequency division multiplexing (FDM)

[2] Sometimes this is called frequency domain alias cancellation (FDAC).

1.2 Prototype-based design

Transfer functions of subband filters can be designed more or less separately in order to reach the goal of perfect reconstruction. However, such design problems can be very complex. In addition, computational complexity of the designs can become high, if separate implementation for each subband filter is needed.

Alternative approach is to have a single low-pass prototype filter, and derive M separate transfer functions from it by modulation. Such designs are called *modulated filter banks*. In general, perfect reconstruction is not achievable with such designs. However, the approach leads to computationally efficient designs. Underlying ideas behind the modulated filter banks, the *modulation representation* and its link to the polyphase representation, are explained next.

1.3 Modulation representation

Modulation of a z-transform $X(z)$ (representing a signal or a system) is done by multiplication of the independent variable z with the number W_M^k, where $W_M = \exp(-j2\pi/M)$:

$$X_k^{(m)}(z) = X(zW_M^k). \tag{1}$$

The name "modulation" can be understood by considering the effect of the operation in the time domain. If $X(z)$ is the z-transform of $x(n)$, then $X_k^{(m)}(z)$ is the z-transform of

$$x(n)\cos(2\pi kn/M) + jx(n)\sin(2\pi kn/M).$$

Also, multiplication of the independent variable z changes the signal spectrum as illustrated in Fig. 1. At this point you may also be interested to take a look at the figures of Sec. 3.2.

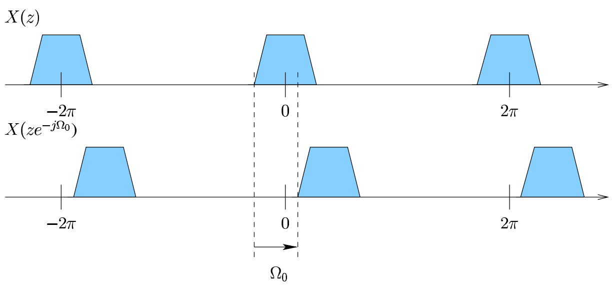


Figure 1: Change of the spectrum, when z is multiplied by $\exp(-j\Omega_0)$.

Modulation representation of a z-transform is composed of the complete set of modulated transforms, which are collected to a column vector:

$$\mathbf{x}^{(m)}(z) = \begin{bmatrix} X_0^{(m)}(z) \\ X_1^{(m)}(z) \\ \vdots \\ X_{M-1}^{(m)}(z) \end{bmatrix}. \tag{2}$$

Note that $X_0^{(m)}(z) = X(z)$.

Example 1. Let us consider the FIR filter

$$H(z) = a_0 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3} + a_4 z^{-4} + a_5 z^{-5}.$$

Its modulation representation for $M = 2$ is

$$\mathbf{h}^{(m)}(z) = \begin{bmatrix} H(z) \\ H(-z) \end{bmatrix} = \begin{bmatrix} a_0 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3} + a_4 z^{-4} + a_5 z^{-5} \\ a_0 - a_1 z^{-1} + a_2 z^{-2} - a_3 z^{-3} + a_4 z^{-4} - a_5 z^{-5} \end{bmatrix}$$

because $W_2 = \exp(-j\pi) = -1$. ■

1.4 Link to polyphase decomposition

There exists a link between polyphase and modulation representations, both of order M, which can be written as

$$\mathbf{x}^{(p)}(z) = \frac{1}{M}\mathbf{W}_M \mathbf{x}^{(m)}(z) \tag{3}$$

where $\mathbf{x}^{(p)}(z)$ and $\mathbf{x}^{(m)}(z)$ are the polyphase and modulation representation of $X(z)$, and $\mathbf{W}_M$ is the so-called DFT matrix.

Discrete Fourier transform (DFT) can be computed by multiplying a vector of sampled values by $\mathbf{W}_M$ which explains its name. $\mathbf{W}_M$ is composed of the elements $W_M^{k\lambda}$, $\lambda \in \{0, 1, \ldots, M-1\}$ corresponding to the rows $\lambda + 1$ and $k \in \{0, 1, \ldots, M-1\}$ corresponding to the columns $k + 1$. We can write

$$\mathbf{W}_M = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W_M^1 & W_M^2 & \cdots & W_M^{M-1} \\ 1 & W_M^2 & W_M^4 & \cdots & W_M^{2(M-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & W_M^{M-1} & W_M^{2(M-1)} & \cdots & W_M^{(M-1)(M-1)} \end{bmatrix}. \tag{4}$$

Example 2. Let us reconsider the FIR filter of Example 1. For $M = 2$, we get

$$\mathbf{W}_M = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Thus, according to (3) we get

$$\underbrace{\begin{bmatrix} a_0 + a_2 z^{-2} + a_4 z^{-4} \\ z^{-1}(a_1 + a_3 z^{-2} + a_5 z^{-4}) \end{bmatrix}}_{\mathbf{h}^{(p)}(z)} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \underbrace{\begin{bmatrix} H(z) \\ H(-z) \end{bmatrix}}_{\mathbf{h}^{(m)}(z)}$$

which is clearly true. ■

We see that $\mathbf{W}_M$ is symmetric. It is also unitary up to a scaling constant: $\mathbf{W}_M^H \mathbf{W}_M = M\mathbf{I}$ where the superscript ^{H} denotes transjugation (transposition followed by complex conjugation of elements). Due to symmetry, transjugation is equivalent to complex conjugation: $\mathbf{W}_M^H = \mathbf{W}_M^*$. Therefore, we get

$$\mathbf{x}^{(m)}(z) = \mathbf{W}_M^* \mathbf{x}^{(p)}(z) \tag{5}$$

This important equation shows how *modulation representation components can be calculated from the polyphase representation components*.

When a modulated filter bank is expressed using the modulation representation, we can use this equation to represent it in terms of polyphase components, which subsequently leads to efficient implementation (Fig. 2). Note that if the outputs are critically downsampled (by factor M) we can apply noble identities, and move the downsampling operations before the filters $H_\lambda^{(p)}(z^M)$ as z^M appears within the parentheses. Furthermore, note that $\mathbf{W}_M^*$ represents IDFT. When M is a power of two, FFT can be used in the filter bank implementation.

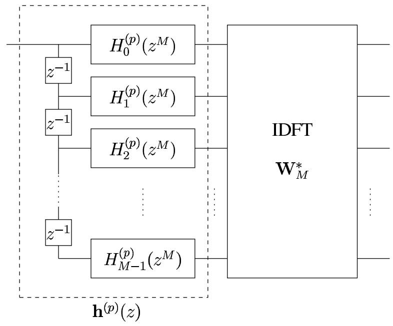


Figure 2: A system structure implementing a vector of modulated transfer functions, $\mathbf{h}^{(m)}(z) = \left[H_0^{(m)}(z),\ H_1^{(m)}(z), \ldots,\ H_{M-1}^{(m)}(z)\right]^T$. Based on the equation (5).

2 Two-channel filter banks

Taken a look at the underlying theory behind filter banks, let us consider the simplest case, two-channel filter banks. Two-channel subband coding (SBC) filter bank is illustrated in Fig. 3. In analysis, downsampling by $M = 2$ is applied to the outputs of the low-pass filter $H_0(z)$ and high-pass filter $H_1(z)$ so sampling is critical.

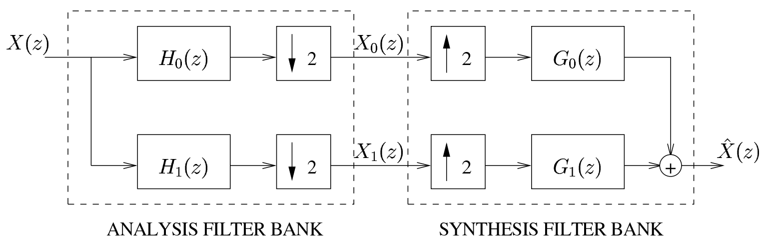


Figure 3: Two-channel SBC filter bank.

Based on the Z-transforms of downsampling and upsampling operations (see Table 1 of introduction “Part 4(a): Multirate techniques”), the analysis filter bank can be represented as

$$\begin{bmatrix} X_0(z) \\ X_1(z) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} H_0(z^{1/2}) & H_0(-z^{1/2}) \\ H_1(z^{1/2}) & H_1(-z^{1/2}) \end{bmatrix} \begin{bmatrix} X(z^{1/2}) \\ X(-z^{1/2}) \end{bmatrix}, \tag{6}$$

and the two-channel synthesis filter bank can be written as

$$\hat{X}(z) = \begin{bmatrix} G_0(z) & G_1(z) \end{bmatrix} \begin{bmatrix} X_0(z^2) \\ X_1(z^2) \end{bmatrix} \tag{7}$$

2.1 Perfect reconstruction

Perfect reconstruction means that the signal is not distorted in a SBC system, that is, impulse response of the system consists of an impulse. Let us find out the conditions for it in terms of the z-transform.

Substituting (6) into (7) gives the relationship between the output $\hat{X}(z)$ and the input $X(z)$which can be written as

$$\hat{X}(z) = F_0(z)X(z) + F_1(z)X(-z) \tag{8}$$

where

$$F_0(z) = \frac{1}{2}\left(G_0(z)H_0(z) + G_1(z)H_1(z)\right) \tag{9}$$

$$F_1(z) = \frac{1}{2}\left(G_0(z)H_0(-z) + G_1(z)H_1(-z)\right) \tag{10}$$

Note that the function $F_0(z)$ multiplies $X(z)$, and therefore it describes the transfer characteristics of the filter bank. The function $F_1(z)$ represents transfer function for the alias component $X(-z)$.

Based on this expansion, we can state two conditions for perfect reconstruction:

1. The alias component $X(-z)$ should not affect the output $\hat{X}(z)$. Therefore, one should have $F_1(z) = 0$.

2. Transfer characteristics for $X(z)$ should be a pure delay[3], that is, $F_0(z) = z^{-k}$.

Design aim in two-channel filter banks is that these conditions are satisfied as closely as possible, or even exactly.

2.2 QMF banks

In quadrature mirror filter (QMF) banks, the transfer functions are defined as

$$H_0(z) = H(z) \tag{11}$$

$$H_1(z) = H(-z) \tag{12}$$

$$G_0(z) = 2H(z) \tag{13}$$

$$G_1(z) = -2H(-z) \tag{14}$$

[3]You may see formulations where $F_0(z) = 1$. In such a case, filters are modelled as zero-phase filters which are noncausal. Corresponding causal realizable filters are obtained by shifting the impulse responses $h(n)$ of the filters so that $h(n) = 0$ for all $n < 0$.

where $H(z)$ is a low-pass prototype filter.

It is easy to see that QMF banks are alias-free. Therefore, the condition for perfect reconstruction, in terms of the prototype filter, is

$$H^2(z) - H^2(-z) = -z^k.$$

It can be shown that in the case of a linear phase FIR prototype filter $H(z)$, the frequency response of the filter must be power complementary, that is,

$$|H(e^{j\Omega})|^2 + |H(e^{j(\pi-\Omega)})|^2 = 1. \tag{15}$$

where Ω denotes the angular frequency. Thus, $H_0(z)$ and $H_1(z)$ are mirror images over $\Omega = \pi/2$. The name of the filter bank follows from these properties. The property (15) cannot be exactly satisfied in practical systems. However, it can be approximated well with numeric methods. Crochiere & Rabiner (1983) list optimized FIR prototypes for various filter lengths.

2.3 Polyphase implementation

2.3.1 General case

Let us see now how polyphase ideas can be applied in the case of SBC banks. In the general case, we must first express $H_0(z)$ and $H_1(z)$ in a polyphase form for $M = 2$:

$$H_0(z) = H_{00}(z^2) + z^{-1}H_{01}(z^2)$$
$$H_1(z) = H_{10}(z^2) + z^{-1}H_{11}(z^2)$$

Then, we use the identity 3 which gives us a structure where all the polyphase filters are operated at the reduced sample rate (Fig. 4).

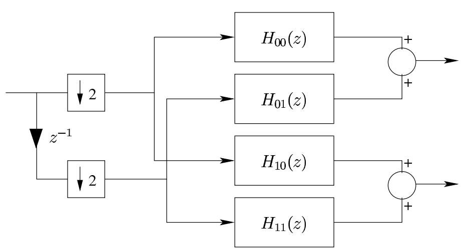


Figure 4: Polyphase implementation of the two-channel analysis bank in the general case.

2.3.2 Special case: QMF banks

In the case of QMF banks we can derive a more efficient version. In fact, the QMF bank is a 2-channel complex modulated filter bank (see Sec. 3.2.1), where a lowpass prototype $H(z)$ is modulated. The theory presented in Sec. 1.4 can be applied here, so let us use it:

1. First, we express the QMF bank in terms of the modulation representation (2). Because $M = 2$ we have

$$\mathbf{h}^{(m)}(z) = \begin{bmatrix} H(z) \\ H(-z) \end{bmatrix}.$$

   Based on the equations (11)-(14) we get the relationship

$$\begin{bmatrix} H_0(z) \\ H_1(z) \end{bmatrix} = \mathbf{h}^m(z)$$

   for the analysis filter bank, and

$$\begin{bmatrix} G_0(z) & G_1(z) \end{bmatrix} = \mathbf{h}^m(z)^T \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$$

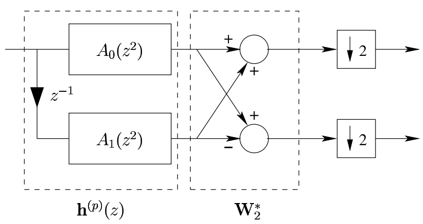
   for the synthesis.

2. Then, modulation representation is expressed in terms of polyphase representation according to the equation (5). Let the polyphase decomposition of $H(z)$ be $A_0(z^2) + z^{-1}A_1(z^2)$. We get

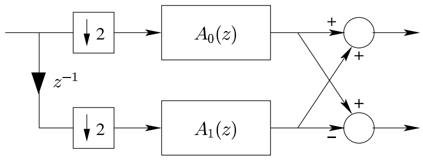
$$\mathbf{h}^{(m)}(z) = \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{\mathbf{W}_2^*} \mathbf{h}^{(p)}(z)$$

   where $\mathbf{h}^{(p)}(z) = \left[A_0(z^2)\ \ z^{-1}A_1(z^2)\right]^T$.

3. Signal flow graphs for analysis and synthesis can now be drawn. For analysis, we get first



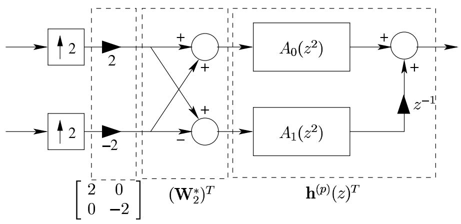
and after application of the first and third noble identities, we get



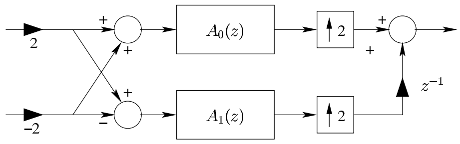
In the case of synthesis, our starting point is

$$\begin{bmatrix} G_0(z) & G_1(z) \end{bmatrix} = \mathbf{h}^p(z)^T \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{(\mathbf{W}_2^*)^T} \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$$

which gives



Then, fourth and sixth noble identities give



Similar systematic derivation can be done for other kinds of modulated filter banks, when subband filters are expressed first in terms of prototype modulations.

3 Some M-channel filter bank structures

3.1 Tree-structured filter banks

Two-channel filter banks can be used in a hierarchical manner. For example, the low-pass and high-pass outputs of an analysis filter bank can both be directed to other analysis filter banks in order to refine analysis (Fig. 5).

Tree-structured analysis and synthesis is used in the JPEG2000 standard . The method is based on the 1-D discrete wavelet transform which is applied separably

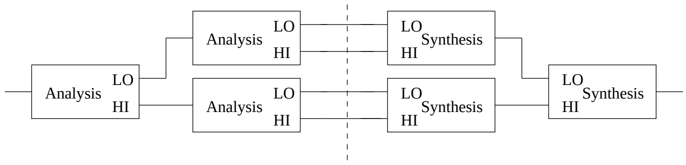


Figure 5: An example of a tree-structured SBC filter bank.

to the image data (Christopoulos *et al.* 2000). The filter bank is a 2-D octave filter bank where low-pass information is passed to further analysis. There are both reversible (i.e. perfect reconstruction) and irreversible filters that are applied in the framework.

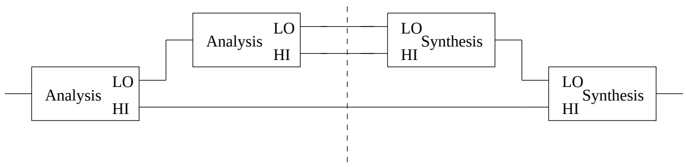


Figure 6: A two-level octave SBC filter bank (1-D).

3.2 Parallel-structured filter banks

The M-channel filter bank, where the subband filters are arranged in parallel is an alternative to the tree structure. An advantage here is that M can here be any integer number. In practice, modulated filter banks are used in many cases due to simplified design process. Complex modulated filter banks (DFT filter banks) and cosine-modulated filter banks are examples of such banks.

3.2.1 Complex modulated filter banks

Basic theory of complex modulation was given in Sec. 1.3 and Sec. 1.4. Complex modulated filter banks perform complex filtering according to each transfer function appearing in the vector

$$\mathbf{h}^{(m)}(z) = \left[H_0^{(m)}(z),\ H_1^{(m)}(z), \ldots,\ H_{M-1}^{(m)}(z)\right]^T .$$

The components are defined as $H_k^{(m)}(z) = H(zW_M^k)$ where $H(z)$ is the prototype filter. The bandwidth of the low-pass prototype is π/M. The frequency scheme is illustrated in Fig. 7. When critical sampling (downsampling by M) is used an efficient implementation can be based on polyphase decomposition (Fig. 2) and application of noble identities.

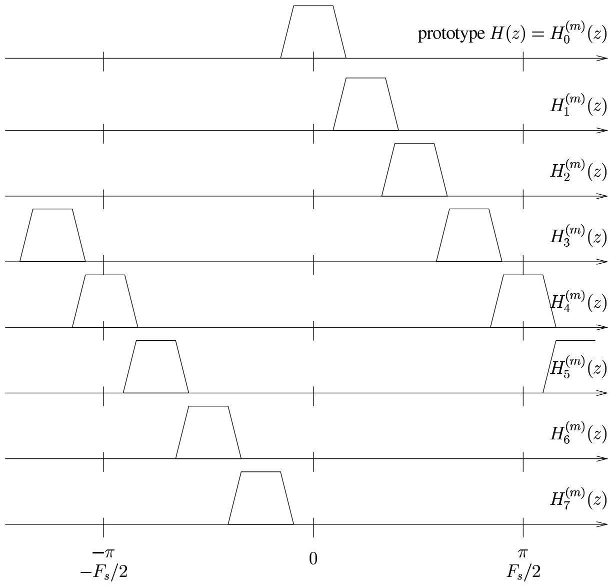


Figure 7: Frequency scheme of an 8-channel complex-modulated filter bank.

3.2.2 Cosine-modulated filter banks

Cosine-modulated filter banks are relatives of the complex modulated filter banks, and can also be explained in terms of modulation representation. The main properties of the cosine-modulated filter bank are:

- The bandwidth of the prototype is $\pi/2M$.
- In both analysis and synthesis filter banks, $2M$ complex modulated filters are used.
- The channel filters are derived from the prototype by complex modulation at frequencies $(2k+1) \times 2\pi/4M$, where $k \in \{0, 1, \ldots, 2M-1\}$. Note that the modulation frequencies are multiples of $2\pi/4M$.
- We know that $\cos(x) = (e^{jx} + e^{-jx})/2$. The relationship between the complex modulation and cosine modulation is based on this equality.

The complex-modulated filters appear as pairs $(k, 2M-k-1)$, $k = 0, 1, \ldots, M-1$ whose impulse responses are complex conjugates of each other (see Fig. 8). When channel filters of such a pair are added, it gives a channel filter with a *real* impulse response. If the impulse response of the prototype is $h(n)$ then the impulse response of the filter for the channel $k \in \{0, 1, \ldots, M-1\}$ is

$$h_i(n) = h(n)\cos(n(2k+1) \times 2\pi/4M)$$

It may be interesting to compare this formula to the MDCT studied in the previous exercise. Filter banks based on cosine function are designed in both cases, but the background theories are different.

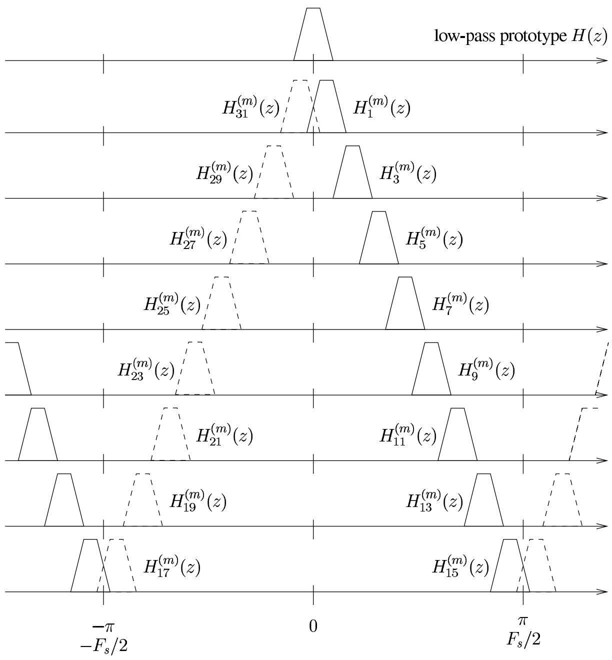


Figure 8: Frequency scheme of an 8-channel cosine-modulated filter bank.

3.2.3 Example: MPEG-1 audio

MPEG-1 audio coding uses a polyphase filter bank as a basis for subband coding (extended with MDCT/TDAC in Layer3 (MP3) codecs). The polyphase filter bank is an example of a cosine-modulated filter banka, and used to transform the signal to 32 critically sampled subbands (Shlien 1994).

The file **MPEG1_demo.mdl** which is a Matlab Simulink model demonstrates the operation of the analysis filter bank. The model contains two analysis filters.

The upper one performs filtering for just one subband whereas the lower one calculates all the subbands using the polyphase approach. Computational advantage of the polyphase implementation should be clear. Note also how cosine modulation is visible in the impulse responses of the subband filters.

References

Christopoulos, C., Skodras, A. & Ebrahimi, T. (2000) The JPEG2000 still image coding system: an overview. IEEE Transactions on Consumer Electronics 46(4):1103–1127.

Crochiere, R. E. & Rabiner, L. R. (1983) Multirate Digital Signal Processing. Prentice Hall, Inc., Englewood Cliffs, New Jersey.

Shlien, S. (1994) Guide to MPEG-1 audio standard. IEEE Transactions on Broadcasting 40(4):206–218.