Modulated Filter Bank Design with Nilpotent Matrices

Gerald Schuller   Wim Sweldens

Lucent Technologies Bell Laboratories
600 Mountain Avenue, Murray Hill, NJ 07974
{schuller,wim}@bell-labs.com

ABSTRACT

We present a technique based on nilpotent matrices for building filter banks with FIR filters and perfect reconstruction. The general design method can be used to design bi-orthogonal filters with unequal filter lengths between analysis and synthesis. This is useful for audio or image coding applications. We can also explicitly control the overall system delay of causal filter banks. The design method is based on a factorization of the polyphase matrices into factors with nilpotent matrices. These factors guarantee mathematical perfect reconstruction of the filter bank, and lead to FIR filters for analysis and synthesis. Using matrices with nilpotency of higher order than 2 leads to FIR filter banks with unequal filter length for analysis and synthesis. The general theory is then applied to the design of cosine modulated filter banks. This leads leads to an efficient implementation, and it is shown that in this case the filters have to have the same length for analysis and synthesis.

Keywords: Modulated filter banks, system delay, low delay, FIR filter banks, critical sampling

1. INTRODUCTION

Many applications, like speech and audio or image coding, require a time frequency representation of a signal (analysis), and reconstructing the signal from this representation (synthesis). Traditionally block transforms have been used, but by now it is well known that filter banks are both more general and more powerful.

In audio coding, filter banks are used to obtain a redundancy reduction, as well as a irrelevance reduction through the application of perceptual models. The perceptual model generates the limits below which the quantization error is inaudible. This needs to be done in the frequency- as well as in the time-domain. E.g. the quantization error before an acoustical event like a click or an attack has to be much smaller than after that event, in order to be inaudible. Otherwise the quantization error may be audible as a "pre-echo". For this reason filter banks with non-symmetric impulse responses or with a low system delay are desirable. Further a possibility to obtain filter banks with unequal impulse response lengths for analysis and synthesis improves the flexibility to adapt to the perceptual limits. So can long analysis filters lead to a good frequency selectivity for a good redundancy reduction, and short synthesis filters to a limited temporal noise spread and improved irrelevance reduction. A similar reasoning can also be applied to image coding.

Most existing design methods lead to orthogonal filter banks, in which case analysis and synthesis filters are time reversed versions of each other. Orthogonal filter banks suffer from three important disadvantages in audio coding, and analog reasons also apply to image coding:

1. Both analysis and synthesis filters have to have the same length L.
2. A delay proportional to the length of the filters $(L-1)$ is needed to obtain a causal system.
3. Orthogonal filters spread the quantization noise symmetrically in time around an acoustic event. This is not a good match to the psycho-acoustic properties of the ear, cf. the pre-echo problem.

In this paper we present a general method for building biorthogonal filter banks using nilpotent matrices which avoids these problems.

1. The design method for general filter banks allows for different lengths in analysis and synthesis filters.
2. It allows careful control over the delay.

3. The resulting filter banks enable a spread of the quantization noise non-symmetrical around some signal event.

In this paper we first present the general idea of using nilpotent matrices in the design of filter banks and show how this leads to an efficient factorization of the polyphase matrix. Later we look at the special case of cosine modulated filter bank and show how the nilpotent matrices lead to an efficient implementation. Several examples are included.

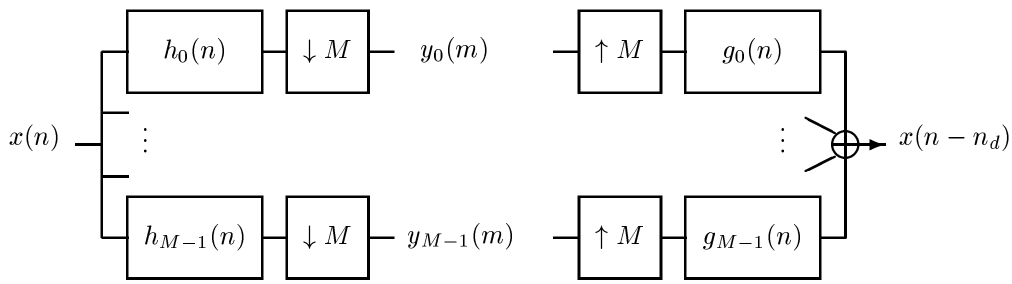


Figure 1. An M - channel filter bank with critical downsampling, perfect reconstruction and a system delay of n_d samples

2. POLYPHASE DESCRIPTION

For an M-band analysis/synthesis filter bank, as depicted in Fig. 1, the input can be represented by an M-dimensional vector $\mathbf{x}(m)$ composed of the downsampled input components

$$\mathbf{x}(m) = [x(mM+M-1), x(mM+M-2), \ldots, x(mM)]^t.$$

Its z-transform is the vector $\mathbf{X}(z)$. The polyphase description for an M-band filter bank with input signal $\mathbf{X}(z)$, the subband signal $\mathbf{Y}(z)$, and the reconstructed signal $\hat{\mathbf{X}}(z)$ is[10]

$$\mathbf{Y}(z) = \mathbf{E}(z)\mathbf{X}(z) \quad \text{and} \quad \hat{\mathbf{X}}(z) = \mathbf{R}(z)\mathbf{Y}(z).$$

Here $\mathbf{E}(z)$ is the $M \times M$ analysis polyphase matrix, $\mathbf{R}(z)$ the synthesis polyphase matrix. This formulation has the advantage, that the filtering and up/down sampling can simply be written as matrix multiplications. The resulting signal flow structure can be seen in Fig. 2. This structure also has the advantage, that the down/up samplers are in the beginning/end, so that all the signal processing operations take place at the lower sampling rate. But maybe most important is, that perfect reconstruction can be obtained simply by matrix inversion.

The filter bank is perfect reconstructing (PR) if

$$\mathbf{R}(z) = z^{-d}\mathbf{S}^{n_t}(z)\mathbf{E}^{-1}(z), \tag{1}$$

where $\mathbf{S}$ is a shift matrix, which shifts the elements of a signal vector by one sample:

$$\mathbf{S}(z) := \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & \vdots \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & \vdots & \vdots & \ddots & 1 \\ z & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

The output of the filter bank is delayed compared to its input by d signal blocks of length M minus n_t samples from the shift matrix, to allow for causal filters. Moreover there is an additional blocking delay of $M-1$ samples, resulting from assembling signal blocks of length M. The total system delay n_d is thus $n_d = d \cdot M - n_t + M - 1$ samples.

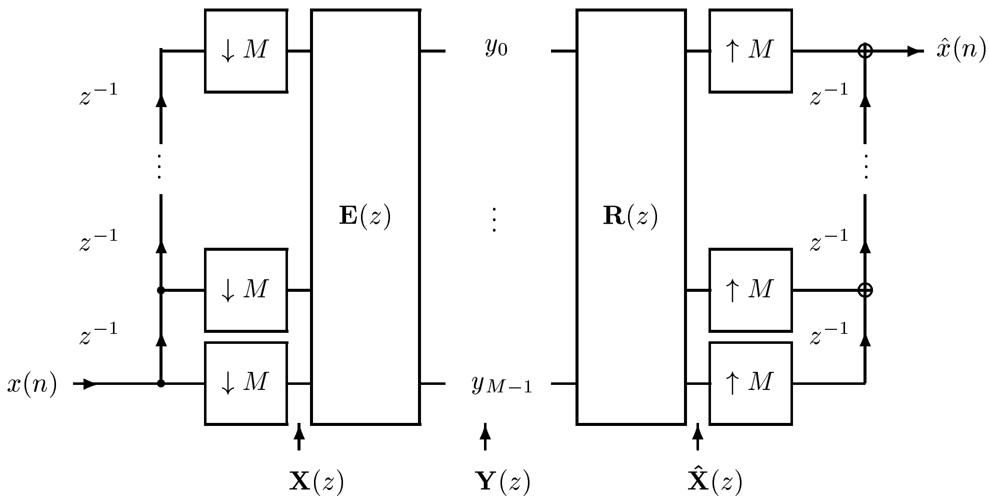


Figure 2. Polyphase representation of an M - channel filter bank with critical downsampling.

3. NILPOTENT MATRICES

In this section we mention some general properties of nilpotent matrices.

DEFINITION 3.1. *We say that a square matrix* $\mathbf{A}$ *is nilpotent of order* l *(*$l > 1$*) if* $\mathbf{A}^l = \mathbf{0}$*, where* l *is the smallest integer with this property.* It is clear from the definition that a nilpotent matrix has determinant zero, and that all eigenvalues are zero. A nilpotent matrix $\mathbf{A}$ can be characterized by its Jordan normal form[8,3]:

$$\mathbf{A} = \mathbf{T}\,\mathbf{J}\,\mathbf{T}^{-1}$$

with $\mathbf{T}$ non singular,

$$\mathbf{J} = \begin{bmatrix} \mathbf{D}_1 & & \\ & \mathbf{D}_2 & \mathbf{0} \\ \mathbf{0} & & \ddots \end{bmatrix} \quad \text{and} \quad \mathbf{D}_k = \begin{bmatrix} 0 & 1 & & \\ & 0 & 1 & \\ 0 & & & \ddots \end{bmatrix}.$$

Here $\mathbf{D}_k$ is a size $n_k \times n_k$ matrix with ones on the first upper diagonal and zeros everywhere else, so that $\sum_k n_k = M$. It is easy to see that $\mathbf{D}_k$ is nilpotent of order n_k. The nilpotency order l of $\mathbf{A}$ is thus determined by the biggest matrix $\mathbf{D}_k$, $l = \max_k(n_k)$. Since $n_k \leq M$ it follows that

$$l \leq M. \tag{2}$$

4. FACTORIZATION

A common approach in filter bank design is to build the polyphase matrix as a product or cascade of simple, canonical matrices which are easily inverted, e.g., the lattice factorization[10,14] in the orthogonal case. The inverse can then easily be found by inverting each of the building blocks.

Here we consider building blocks of the type $\mathbf{I} + \mathbf{A}(z)$, where $\mathbf{A}(z)$ is nilpotent of order l, now only in a more general form as matrix of polynomials. These are thus unimodular matrices.[10] The determinant of such a matrix is always equal to one and its inverse can immediately be found as

$$(\mathbf{I} + \mathbf{A}(z))^{-1} = \mathbf{I} + \sum_{i=1}^{l-1} (-\mathbf{A})^i(z). \tag{3}$$

Formally one can think of this as a Taylor expansion of the inverse. With general matrices this sum would have an infinite number of terms, leading to IIR filters which may not be stable. But with nilpotent matrices, the sum has only $l-1$ terms, as higher powers of $\mathbf{A}(z)$ are zero. This means that if the analysis consists of FIR filters, $\mathbf{A}(z)$ consists of finite polynomials (Laurent polynomials), and the inverse also leads to FIR filters.

Clearly we are interested in matrices $\mathbf{A}(z)$ which are inexpensive to apply and simplify the design process. Therefore we restrict ourselves to matrices $\mathbf{A}(z)$ of the form $z^p \cdot \mathbf{A}$ with $\mathbf{A}$ a nilpotent matrix and $p \in \{1, -1\}$. It can be shown that these two types provide enough generality for practical filter design. This leaves two types of matrices:

$$\mathbf{L}(z) := \mathbf{I} + z^{-1}\mathbf{A} \quad \text{and} \quad \mathbf{H}(z) := \mathbf{I} + z\mathbf{A}. \tag{4}$$

The polyphase matrix can then be written as the product

$$\mathbf{E}(z) = \mathbf{V} \prod_{i=1}^{\nu} \mathbf{L}_i(z) \prod_{j=1}^{\mu} \mathbf{H}_j(z) \mathbf{S}^{n_a}(z). \tag{5}$$

Each $\mathbf{L}_i(z)$ and $\mathbf{H}_j(z)$ can have a different matrix $\mathbf{A}$, and $\mathbf{V}$ is an invertible matrix.

Observe that the degree of the inverse of $\mathbf{I} + \mathbf{A}(z)$ is increased if $l > 2$. This means that the synthesis side has longer filters than the analysis side, or vice versa. This is quite an unusual property. Common design methods for perfect reconstruction filter banks either have the same filter lengths for analysis and synthesis (even though they can differ between bands), or they lead to IIR filters on one side. The use of nilpotent matrices provide the flexibility for a tradeoff of the filter length between analysis and synthesis.

The form of the building blocks $\mathbf{I} + \mathbf{A}(z)$ also has the advantage, that its inverse leads to causal filters if the matrix $\mathbf{A}(z)$ is causal. This property can be used to control the system delay of a filter bank. E.g. if the analysis only consists of the causal $\mathbf{L}(z)$ matrices, then the inverse for synthesis can also be realized with causal filters. This means that even though there is no restriction on the degree of the polyphase matrix and hence the length of the filters, the system delay is only the blocking delay of $M-1$ samples.

The building block $\mathbf{H}(z)$ is useful because it allows to design more general filter banks. E.g. orthogonal filter banks can be obtained by having equal numbers of the $\mathbf{H}(z)$ and $\mathbf{L}(z)$ matrices in the cascade. Note that the matrix $\mathbf{H}(z)$ needs a multiplication with z^{-1} to become causal, and the inverse $\mathbf{H}^{-1}(z)$ needs a multiplication with z^{-l+1}. Hence the additional system delay needed to make the $\mathbf{H}_j(z)$ and $\mathbf{H}_j(z)^{-1}$ matrices causal is equal to $l_j M$, where l_j is the nilpotency order of $\mathbf{A}_j$. Thus the total system delay of (5) including the shift matrix and blocking delay is $M - 1 - n_a - n_s + \sum_{j=1}^{\mu} l_j M$.

Since we need $l > 2$ for unequal filter lengths (3), and since $M \geq l$ (2), we need $M > 2$ for this to happen. Observe that cosine modulated filter banks can be described as a system with nested 2-band filter banks.[10] Hence for this case they have to have the same filter length for analysis and synthesis.

5. COSINE MODULATED FILTER BANKS

This section shows how the factorization with nilpotent matrices can be applied to the design of cosine modulated filter banks. The impulse responses of the cosine modulated filter bank we consider are of the form

$$h_k(n) = h(n) \cdot \cos\left(\frac{\pi}{M}(k+0.5)(n+0.5+n_a)\right), \tag{6}$$

$$g_k(n) = h'(n) \cdot \cos\left(\frac{\pi}{M}(k+0.5)(n+0.5-M+n_s)\right), \tag{7}$$

$k = 0, \ldots, M-1$, $n = 0, \ldots, K-1$, where $h(n)$ and $h'(n)$ are the analysis and synthesis baseband prototype filters, respectively. The values n_a and n_s can be chosen as:

$$n_a = n_s = \frac{n_t}{2}, \tag{8}$$

with n_t from (1). For odd n_t the exponent n_a can be chosen as $n_s + 1$.

It can be shown,[7,12] that the polyphase matrix of this type of filter bank can be represented as the product

$$\mathbf{E}(z) = \mathbf{T} \cdot \mathbf{F}(z) \mathbf{S}^{n_a}(z) \tag{9}$$

$$\mathbf{R}(z) = \mathbf{S}^{n_s}(z) z^{-d} \mathbf{F}^{-1}(z) \cdot \mathbf{T}^{-1}. \tag{10}$$

The elements of the matrix **T** are

$$[\mathbf{T}]_{k,n} := \cos(\frac{\pi}{M}(k+0.5)(n+0.5)), \qquad 0 \le n, k < M.$$

This is the well known DCT type IV. The filter matrix $\mathbf{F}(z)$ then has a sparse, "bi-diagonal" form

$$\begin{bmatrix} \bullet & & 0 & & \bullet \\ & \ddots & & \iddots & \\ & & \bullet \;\; \bullet & & 0 \\ 0 & & \bullet \;\; \bullet & & \\ & \iddots & & \ddots & \\ \bullet & & 0 & & \bullet \end{bmatrix}$$

or more specifically

$$\mathbf{F}(z) = \mathbf{diag}[P_0(-z^2), \ldots, P_{M-1}(-z^2)] \cdot \mathbf{J} + \\ + z^{-1} \cdot \mathbf{diag}[P_{2M-1}(-z^2), \ldots, P_M(-z^2)] \tag{11}$$

$$z^{-d}\mathbf{F}^{-1}(z) = \mathbf{diag}[P'_0(-z^2), \ldots, P'_{M-1}(-z^2)] \cdot \mathbf{J} \\ - z^{-1}\mathbf{diag}[P'_M(-z^2), \ldots, P'_{2M-1}(-z^2)], \tag{12}$$

with the time shifted polyphase representation of the prototypes,

$$P_k(z) = \sum_{m=0}^{\infty} h(m2M + k - n_a) z^{-m} \tag{13}$$

$k = 0, \ldots, 2M-1,$

$$P'_k(z) = \sum_{m=0}^{\infty} h'(m2M + k - n_s) z^{-m}. \tag{14}$$

That the filter matrix has this bi-diagonal form also means that this modulated filter bank can be viewed as a set of nested 2-band filter banks followed by the cosine transform matrix **T**, as can also be seen in.[7,10]

These modulated filter banks can now also be represented by (5). Comparing (5) and (9) shows that **T** in (9) must be part of **V** in (5), or more precisely, it must be a product with a diagonal coefficient matrix **D**

$$\mathbf{V} = \mathbf{T} \cdot \mathbf{D},$$

where **D** has non-zero coefficients only on its diagonal, in order to keep the structure of $\mathbf{F}(z)$.

$\mathbf{F}(z)$ must be represented by the product of the $\mathbf{L}(z)$ and $\mathbf{H}(z)$ matrices. Since $\mathbf{F}(z)$ has a bi-diagonal form, the **A** matrices also have to have a bi-diagonal form. Since they also have to be nilpotent, they can only have non-zero elements on the anti-diagonal. Since **A** matrices, where the non-zero elements are not contiguous on the anti-diagonal, may lead to non-contiguous impulse responses (13,14), this leaves two cases for the **A** matrices:

$$\mathbf{A} = \begin{bmatrix} 0 & & & & a_0 \\ & & & \iddots & \\ & & a_{M/2-1} & & \\ & 0 & & & \\ & & & & \\ 0 & & & & 0 \end{bmatrix}$$

or

$$\mathbf{A} = \begin{bmatrix} 0 & & & 0 \\ & & 0 & \\ & & a_{M/2-1} & \\ & \cdot^{\cdot^{\cdot}} & & \\ a_0 & & & 0 \end{bmatrix}$$

with some real or complex coefficients a_i, $i = 0, \ldots M/2-1$. Observe that both are nilpotent of order $l = 2$. In order to increase the degree of the resulting polyphase matrix with each Maximum-Delay or Zero-Delay matrix, these two types have to alternate in the product, as can also easily seen in the signal flow graphs. The signal flow graph of the resulting causal matrices $\mathbf{L}(z)$ and $\mathbf{H}(z)z^{-1}$ are shown in Fig. 4, 5, 6, and 7. An example for a modulated filter bank with 2 Zero-Delay matrices can be seen in Fig. 8 for the analysis and Fig. 9 for the synthesis. Note the similarity to Lifting[9] and ladder structures.[4] Observe that the Zero-Delay matrices don't introduce any additional system delay, and hence the system delay only consists of the blocking delay, so that $n_d = M-1$. In the traditional case of orthogonal or paraunitary filter banks this would necessitate filter with impulse responses of length M. But in the example this low system delay is obtained even though the resulting filter length is $2.5M$. The proposed structure with Maximum-Delay and Zero-Delay matrices is also suitable for designing time-varying filter banks.[13] Fig. 10 and 11 show examples for impulse responses and magnitude responses of baseband prototypes (the functions $h_k(n)$ and $g_k(n)$, which are in identical in the example, $h_k(n) = g_k(n)$). The example compares two filter banks with $M = 128$ bands. The dashed line is for an orthogonal filter bank, which results for one Maximum-Delay and one Zero-Delay Matrix in the product (5), and with $n_a = n_s = M/2 = 64$. The resulting system delay is $n_d = 255$ and the filter length is 256 taps. This filter bank can also be viewed as an alternative implementation of the traditional orthogonal filter banks.The solid line in Fig. 10 is for a filter bank which has one Maximum-Delay and three Zero-Delay matrices in (5), and also $n_a = n_s = M/2 = 64$. The resulting system delay is again $n_d = 255$, but the resulting filter length is 512 taps. This shows that the length of the filter impulse responses can be increased without increasing the resulting system delay. That property can be used to improve the frequency response, as can be seen in Fig. 11. The lower curve is for the filter bank with 512 taps, which has an about 20 dB higher stopband attenuation.

REFERENCES

1. A. Akansu, M.J.T. Smith (editors) *Subband and Wavelet Transforms* Kluwer Academic Publishers, Norwell, MA, 1996.
2. I. Daubechies and W. Sweldens. Factoring wavelet transforms into lifting steps. *J. Fourier Anal. Appl.*, 4(3):247–269, 1998.
3. W. Greub. *Linear Algebra.* Springer Verlag, New York, 1975.
4. T. A. C. M. Kalker and I. Shah. Ladder Structures for multidimensional linear phase perfect reconstruction filter banks and wavelets. In *Proceedings of the SPIE Conference on Visual Communications and Image Processing* (Boston), pages 12–20, 1992.
5. S.-M. Phoong, C. W. Kim, P. P. Vaidyanathan, and R. Ansari. A new class of two-channel biorthogonal filter banks and wavelet bases. *IEEE Trans. Signal Process.*, 43(3):649–665, 1995.
6. G. Schuller and T. Karp. Causal FIR filter banks with arbitrary system delay. In *DSP Workshop, Bryce Canyon, Utah*, 1998.
7. G. D. T. Schuller and M. J. T. Smith. New framework for modulated perfect reconstruction filter banks. *IEEE Trans. Signal Process.*, 44(8):1941–1954, 1996.
8. J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis.* Springer Verlag, New York, 1980.
9. W. Sweldens. The lifting scheme: A custom-design construction of biorthogonal wavelets. *Appl. Comput. Harmon. Anal.*, 3(2):186–200, 1996.
10. P. P. Vaidyanathan. *Multirate Systems and Filter Banks.* Prentice Hall, Englewood Cliffs, NJ, 1992.
11. G. Schuller: "A Low Delay Filter Bank for Audio Coding with Reduced Pre-Echoes", 99th AES Convention, New York, NY, Oct. 6-9, 1995
12. G. Schuller: "A New Factorization and Structure for Cosine Modulated Filter Banks with Variable System Delay", Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, Nov. 3-6, 1996
13. G. Schuller: "Time-Varying Filter Banks with Variable System Delay" ICASSP 97, Munich, Germany, Apr. 1997
14. H. S. Malvar: "Extended Lapped Transforms: Properties, Applications, and Fast Algorithms", IEEE Transactions on Signal Processing, Vol.40, NO.11, Nov. 1992.

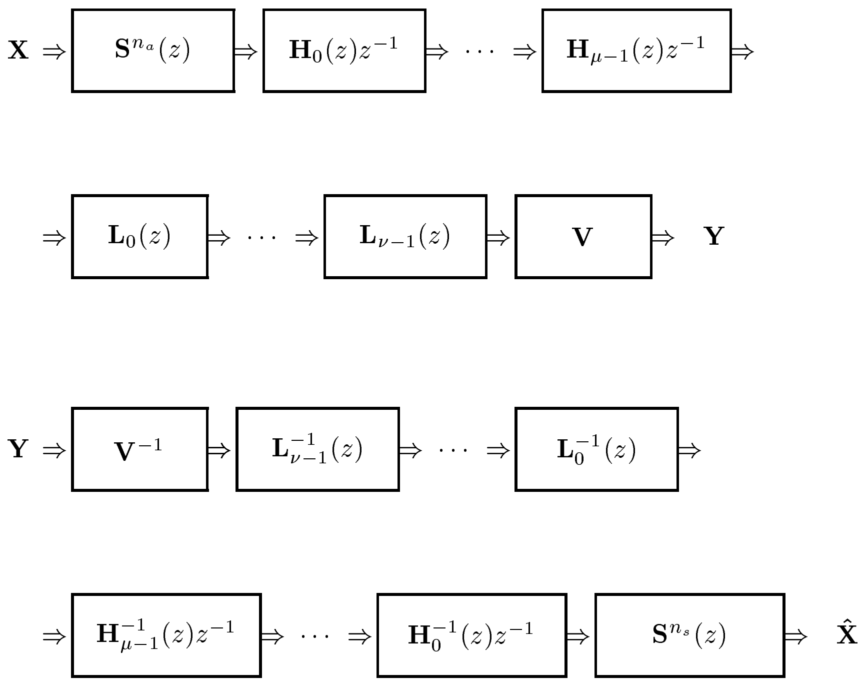


Figure 3. The block diagram of the filter bank consisting of Zero-Delay and Maximum-Delay matrices. The analysis filter bank is above, the synthesis filter bank below.

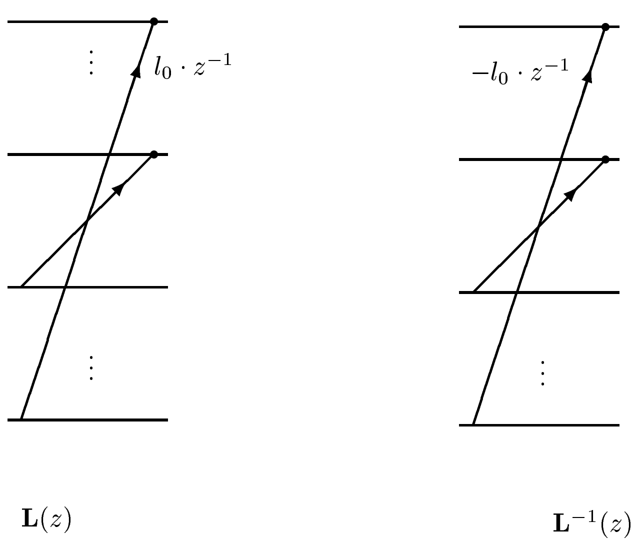
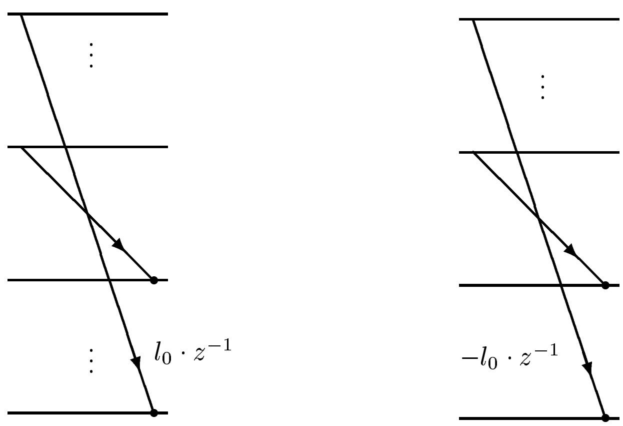
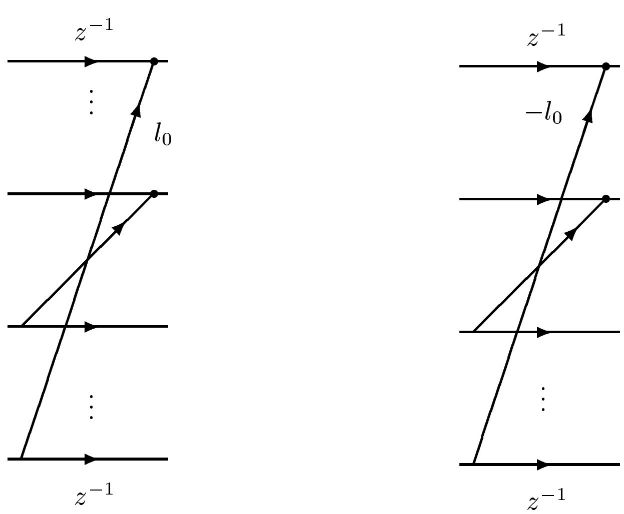


Figure 4. The signal flow structure of the Zero-Delay matrices.



$\mathbf{L}(z)$ $\mathbf{L}^{-1}(z)$

Figure 5. The second structure of the Zero-Delay matrices, where the coefficients are on the other half of the anti-diagonal.



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

Figure 6. The structure of the Maximum-Delay matrices.

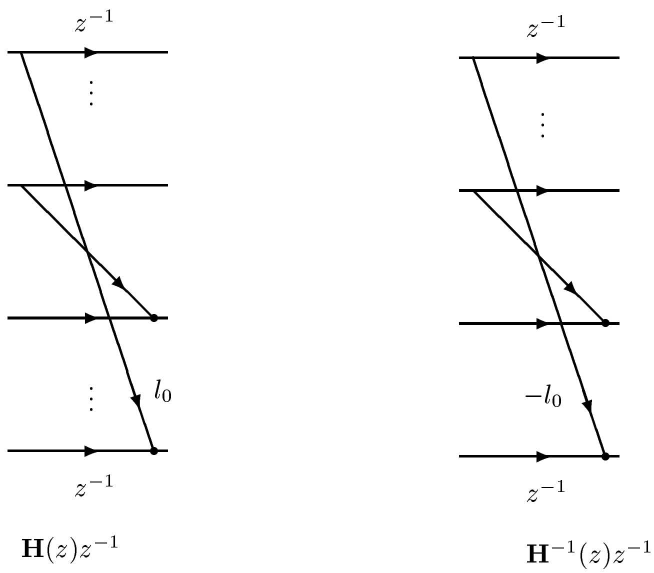


Figure 7. The second structure of the Maximum-Delay matrices.

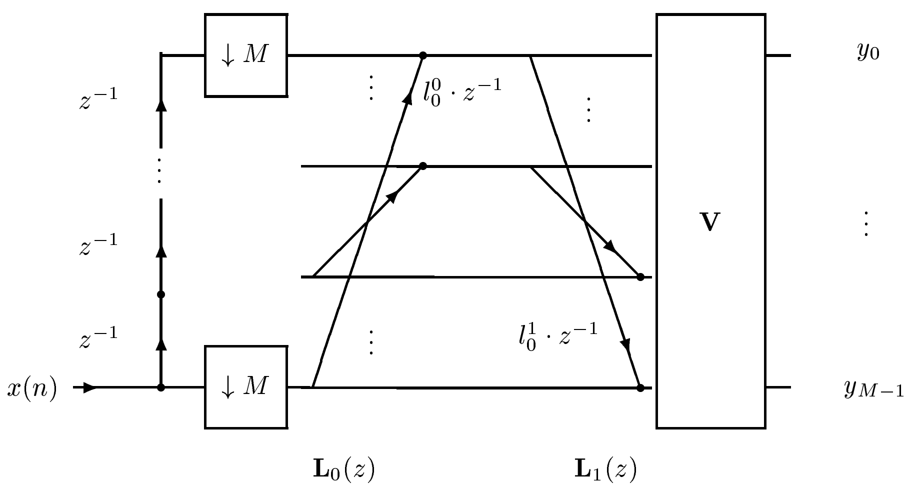


Figure 8. Example of an analysis filter bank with two Zero-Delay matrices.

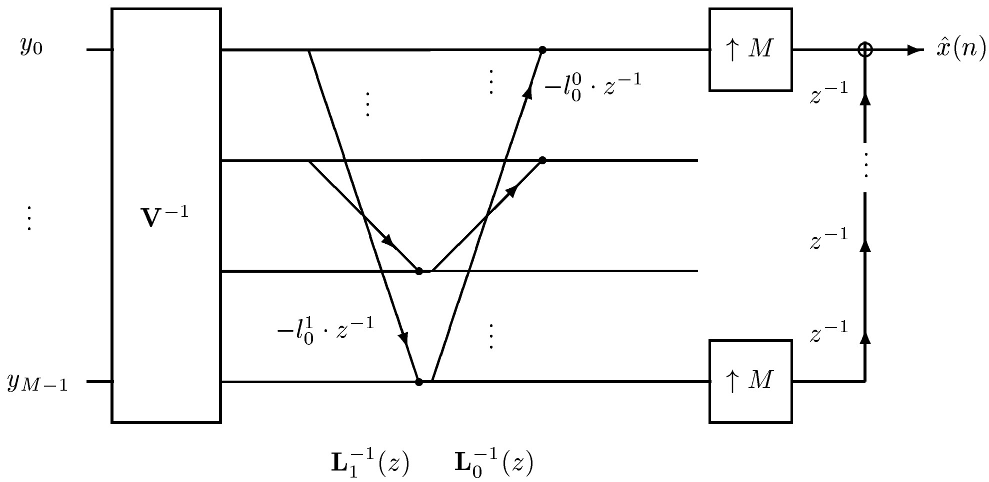


Figure 9. Example of a synthesis filter bank with two Zero-Delay matrices.

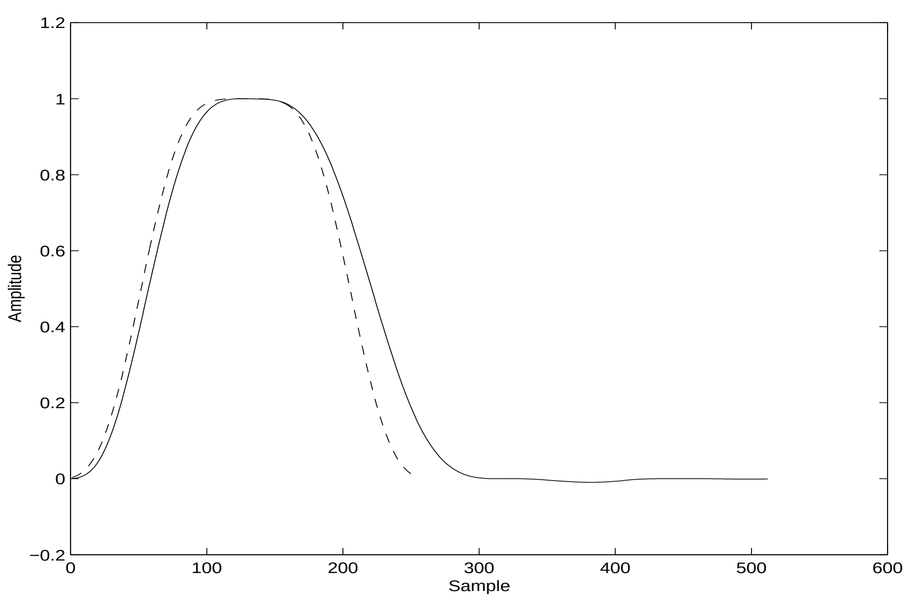


Figure 10. Impulse responses of baseband prototypes, for analysis and synthesis, both with 128 bands and 255 samples delay. The dashed line is for a orthogonal filter bank with one Maximum-Delay and one Zero-Delay matrix and filter length 256 taps. The solid line is for a filter bank with one Maximum-Delay and three Zero-Delay matrices and a filter length of 512 taps.

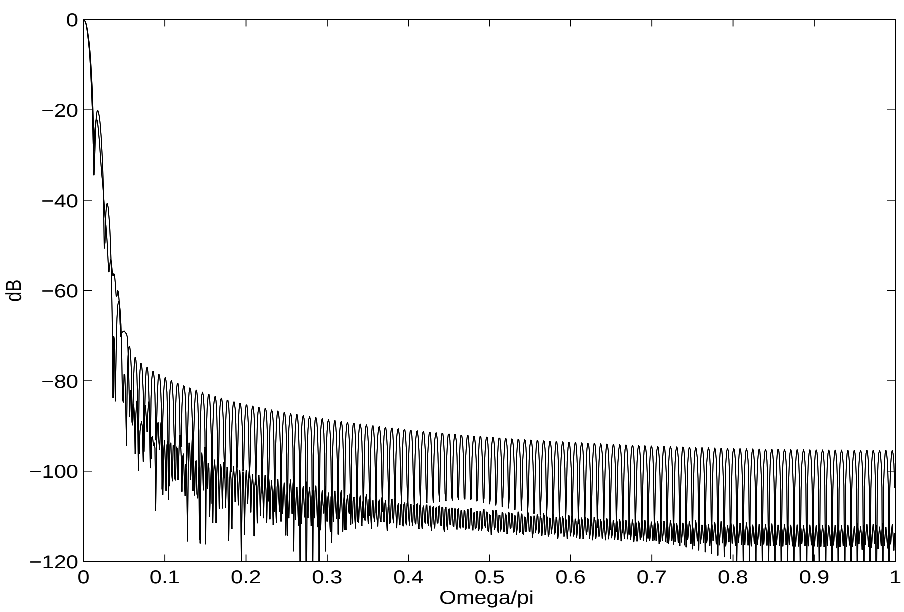


Figure 11. Magnitude responses of baseband prototypes, for analysis and synthesis, both with 128 bands and 255 samples delay. The upper line is for the filter length of 256 taps, the lower line is for the filter length of 512 taps.