A TUTORIAL ON MULTIRATE DIGITAL FILTER BANKS

P. P. Vaidyanathan, Dept. EE, Caltech, Pasadena, CA 91125.

(Invited paper)

Abstract: A number of topics in multirate digital signal processing such as decimation and interpolation, polyphase structures, power-complementary and Mth band filters, and digital filter banks are reviewed here. Maximally decimated digital filter banks are emphasized, and the problem of perfect-reconstruction reviewed. The role of FIR lossless systems in this context is also reviewed.

I. INTRODUCTION

Over the last decade, there has been a steady growth of interest in the *multirate processing* of digital signals. Key developements in this connection include the polyphase formulation [1], multi-stage decimation and interpolation [2], and multirate analysis/synthesis systems [3-7] commonly called the quadrature mirror filters. Applications of multirate systems include subband coding of speech and image signals [3,10,15-17] spectrum analysis [3,11] and voice privacy systems [12], to name a few. In this paper, we review some of the important concepts in multirate signal processing, with emphasis on filter banks. [2]

II. BASIC BUILDING BLOCKS

Two of the most basic building blocks used in multirate systems are the decimator and the interpolator shown in Fig. 1. A M-fold decimator takes an input sequence $x(n)$ and retains only samples that occur at times which are multiples of M. The input output relation is therefore $y(n) = x(Mn)$. As shown in standard references [13, page 82] and [3, page 34], the z transform of $y(n)$ is

$$Y(z) = \frac{1}{M}\sum_{k=0}^{M-1} X(z^{1/M}W^k) \tag{1}$$

where $W \triangleq e^{-j2\pi/M}$. With $z = e^{j\omega}$, (1) becomes

$$Y(e^{j\omega}) = \frac{1}{M}\sum_{k=0}^{M-1} X(e^{j(\omega-2\pi k)/M}) \tag{2}$$

which shows that $Y(e^{j\omega})$ is a sum of M uniformly shifted versions of an M-fold stretched version of $X(e^{j\omega})$. Fig. 2(a) depicts this idea for $M = 2$. Next, an M-fold interpolator, indicated by the symbol of Fig. 1(b), has the following input-output relationship:

$$y(n) = \begin{cases} x(\frac{n}{M}) & n = \text{mul of } M \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

In other words, the output $y(n)$ is obtained by inserting $M-1$ zero-valued samples between adjacent samples of $x(n)$. In the transform domain this is equivalent to

$$Y(z) = X(z^M) \quad \text{or} \quad Y(e^{j\omega}) = X(e^{j\omega M}). \tag{4}$$

So $Y(e^{j\omega})$ is merely a compressed version of $X(e^{j\omega})$. Since $X(e^{j\omega})$ is periodic with period 2π, the function $Y(e^{j\omega})$ has period $2\pi/M$ (see Fig. 2(b)).

If a decimator is followed by an interpolator as in Fig. 3(a), the overall effect is as illustrated in Fig. 3(b). We now have multiple copies of the input spectrum $X(e^{j\omega})$, and successive copies can possibly overlap.

Aliasing and Imaging. If $x(n)$ is bandlimited to $-\pi/M < \omega < \pi/M$, (more generally $\alpha < \omega < \alpha + 2\pi/M$) then there is no overlap among the terms on the RHS of (2). So the shaded areas in Fig. 3(b) are absent, and we can recover $x(n)$ from the signal $y(n)$ of Fig. 3(a) by using a lowpass (more generally, bandpass) filter with passband region $|\omega| < \pi/M$ (more generally $\alpha < \omega < \alpha + 2\pi/M$). If $x(n)$ does not conform to this bandwidth requirement, there is an overlap between various terms in (2). This results in appearance of the shaded areas in Fig. 3(b); we cannot recover $x(n)$ from $y(n)$ of Fig. 3(a) anymore.

The multiple appearance of a compressed version of $X(e^{j\omega})$ caused by an interpolator (Fig. 2(b)) is termed *imaging* (the repetitions are *images*). Thus a decimator can cause aliasing and an interpolator causes imaging.

Decimation and Interpolation filters. A decimation filter is typically a lowpass filter preceding a (M-fold) decimator, to bandlimit the signal and reduce the aliasing effect. An interpolation filter, on the other hand, *follows* an interpolator, the purpose being to eliminate images. These are lowpass filters with cutoff $\approx \pi/M$.

III. DIGITAL FILTER BANKS

An *anlaysis filter bank* is a collection of filters $H_k(z)$,$0 \le k \le M-1$ which splits a signal $x(n)$ into M subband signals $x_k(n)$, $0 \le k \le M-1$ (Fig. 4(a)). For the case of $M = 3$, typical responses of $H_k(z)$ are shown in Fig. 4(c). In this plot, it is assumed that the coefficients of $H_k(z)$ are real so that there is magnitude-symmetry with

[1] Work supported in part by the National Science Foundation grants DCI 8552579 and MIP 8604456.

[2] Boldfaced letters denote matrices. The notations $\mathbf{A}^T$,$\mathbf{A}^*$ and $\mathbf{A}^\dagger$ denote transposition, conjugation and transposed-conjugation respectively. $\mathbf{H}_*(z)$ denotes conjugation of coefficients of $\mathbf{H}(z)$ with z unchanged. Finally, $\tilde{\mathbf{H}}(z)$ stands for $\mathbf{H}_*^T(z^{-1})$.

CH2458-8/88/0000-2241$1.00 © 1988 IEEE

respect to π. The analysis bank splits the spectrum of $x(n)$ into consecutive subbands. A synthesis bank (Fig. 4(b)) is a set of filters $F_k(z), 0 \leq k \leq M-1$, which combines M signals $v_k(n), 0 \leq k \leq M-1$, into one signal $\hat{x}(n)$, typically called the *reconstructed signal.*

Polyphase structures

Consider a simple transfer function of the form $H(z) = 1+z^{-1}$ which has the frequency response magnitude $|H(e^{j\omega})| = 2\cos(\omega/2)$ which is lowpass. Note that we can write $H(z)$ in the form

$$H(z) = E_0(z^2) + z^{-1}E_1(z^2) \tag{5}$$

where $E_0(z) = E_1(z) = 1$. The functions $E_0(z)$ and $E_1(z)$ are clearly allpass (though trivial). This opens up a question: can an *arbitrary* lowpass filter be realized as a sum of allpass polyphase components? The answer is a qualified *yes* as we shall see.

Let $h(n)$ be the impulse response of a 'good' lowpass filter with cutoff at $\pi/2$. Define $e_0(n) = h(2n)$ and $e_1(n) = h(2n+1)$. Clearly the transforms $H(z), E_0(z)$ and $E_1(z)$ are related as in (5). We can interpret $e_0(n)$ and $e_1(n)$ to be the outputs of decimators whose input sequences are $h(n)$ and $h(n+1)$ respectively. Accordingly, we expect $E_0(e^{j\omega})$ and $E_1(e^{j\omega})$ to be stretched versions of $H(e^{j\omega})$. Since the latter is lowpass with cutoff at $\pi/2$, we see that $E_0(z)$ and $E_1(z)$ are approximately allpass. In summary, a good lowpass filter with cutoff frquency $\pi/2$ can be represented as in (5) where $E_0(z)$ and $E_1(z)$ are approximately allpass. More generally, a lowpass filter with cutoff frequency π/M can be written as

$$H(z) = \sum_{\ell=0}^{M-1} z^{-\ell}E_\ell(z^M) \tag{6}$$

where $E_\ell(z)$ are approximations to allpass functions.

In order to get more 'feeling' for the operation of (6), let us write $E_\ell(e^{j\omega})$ for each ℓ as $E_\ell(e^{j\omega}) = e^{j\phi_\ell(\omega)}$ which represents the allpass nature. From here we see that in the passband region $|\omega| < \pi/M$ the phase responses of the 'allpass' functions $z^{-\ell}E_\ell(z^M)$ are aligned while in the stopband, the phases are such that the sum (6) has nearly zero magnitude. This is the origin of the *polyphase concept.*

We shall now make formal definitions in a more general manner so as to expand the usefulness of the polyphase concept. Let $H(z)$ be *any* transfer function with impulse response $h(n)$. Define the M sequences

$$e_\ell(n) \triangleq h(\ell + Mn), 0 \leq \ell \leq M-1, \tag{7}$$

and let

$$E_\ell(z) \triangleq \sum_{n=-\infty}^{\infty} e_\ell(n)z^{-n}. \tag{8}$$

Then $E_\ell(z), 0 \leq \ell \leq M-1$ are said to be the M polyphase components of $H(z)$. By definition of these components (6) is valid, and will be called the (M-component) polyphase representation of $H(z)$. A second type of polyphase representation is of the form $H(z) = \sum_{\ell=0}^{M-1} z^{-(M-1-\ell)}R_\ell(z^M)$. This representation is more convenient in dealing with interpolation filters and *synthesis bank filters.*

In order to appreciate the usefulness of the polyphase representation, we indicate two useful identities [3] in Fig. 5. These hold for any $G(z)$.

Decimation and interpolation filter implementation. Consider Fig. 6(a) which represents a decimation filter followed by an M-fold decimator. Suppose $H(z)$ is written in the form (6). We can then redraw Fig. 6(a) as in Fig. 6(b) (use the identity in Fig. 5(a)). In this implementation, each filter $E_\ell(z)$ is operating at M times lower rate (compared to the input sampling rate). By using a similar developement, an interpolation filter following an interpolator can be implemented efficiently. Useful *commutator models* for these can be found in [3, Ch. 3].

The uniform DFT filter bank

An advantage of the polyphase representation is the amazing efficiency with which we can construct a filter bank at the cost of essentially one single filter. Consider Fig. 7 which represents a system with one input $x(n)$ and M outputs $y_k(n)$, $0 \leq k \leq M-1$. Here the signals $y_k(n)$ are related to the signals $x_\ell(n)$ by $y_k(n) = \sum_{\ell=0}^{M-1} W^{-k\ell}x_\ell(n)/M$. If we define $H_k(z) \triangleq Y_k(z)/X(z)$ for $0 \leq k \leq M-1$, we then see that the relation $H_k(z) = H_0(zW^k)$ is true. In other words, $H_k(e^{j\omega})$ is a shifted version of the prototype response $H_0(e^{j\omega})$. In this way, M filters with uniformly shifted frequency responses that together fill the entire range $0 \leq \omega < 2\pi$ are obtained. Note that $H_0(z)$ is equal to $\sum_{\ell=0}^{M-1} z^{-\ell}\, E_\ell(z^M)/M$. The cost of the filter bank in Fig. 7 is equal to the cost of the prototype filter $H_0(z)$ plus the IDFT-computational cost. This is almost always much less that M times the cost of implementing $H_0(z)$.

A more general scheme can be derived from Fig. 7 by replacing the DFT with a generalized DFT (GDFT, [3]). This results in analysis filters with the 0th filter not necessarily centered at $\omega = 0$. By appropriate choice of the GDFT matrix, the coefficients of $H_k(z)$ can be made equal to the conjugates of those of $H_{M-1-k}(z)$ for $0 \leq k \leq M-1$. By combining pairs of filters $H_k(z)$ and $H_{M-1-k}(z)$, it is possible to obtain a filter bank with $M/2$ filters (M even), having real coefficients.

Power complementary filters. The set $H_k(z)$, $0 \leq k \leq M-1$ is said to be a power complementary (PC) set if $\sum_{k=0}^{M-1} |H_k(e^{j\omega})|^2 = c$ for all ω, where $c \neq 0$ is a constant. By analytic continuation we conclude that such a system satisfies $\sum_{k=0}^{M-1} \tilde{H}_k(z)H_k(z) = c$ for all z. Consider now an analysis filter bank as in Fig. 4(a), where the filters form a PC set. If this analysis bank is followed by a synthesis bank as in Fig. 4(b) (i.e., if $v_k(n) = x_k(n)$) with the synthesis filters chosen as $F_k(z) = z^{-n_0}\tilde{H}_k(z)$ then the output of the synthesis bank is given by $\hat{X}(z) = z^{-n_0}\sum_{k=0}^{M-1} \tilde{H}_k(z)H_k(z)X(z)$ which reduces to $cz^{-n_0}X(z)$. In other words, the subband signals $x_k(n)$ can be combined in this manner to get back the original $x(n)$. If $H_k(z)$ is FIR, then the above choice of $F_k(z)$ results in stable (in fact FIR) synthesis filters whose impulse response coeffi-

cients are obtained by time reversal of the coefficients of $H_k(z)$. Such error-free recovery of $x(n)$ becomes a more challenging problem if we have M fold decimators following the subband signals. This topic is handled in Sec. V.

Mth band transfer functions. Let $H(z)$ be a zero-phase FIR filter, so that the impulse response $h(n)$ is symmetric with respect to $n = 0$. This is said to be an Mth band filter [8] if $h(nM) = 0$ for $n \neq 0$. The polyphase component $E_0(z)$ is a constant, equal to $h(0)$. For arbitray-phase FIR filters and IIR filters, the Mth band definition can be appropriately extended. Mth band filters are also called Nyquist filters and find applications in interpolation and intersymbol interference minimization. A key mathematical property [8] of these filters is the following: $\sum_{k=0}^{M-1} H(zW^k) = Mh(0)$. As a simple application of this property, imagine that the analysis-bank filters $H_k(z)$ in Fig. 4(a) have been chosen to be

$$H_k(z) = H(zW^k) \tag{9}$$

(as in a uniform DFT bank) where $H(z)$ is an Mth band lowpass filter. Then the original signal $x(n)$ can be recovered simply by adding the subband signals (and scaling the result).

Examples where filter banks arise

Perhaps the earliest use of filter banks was in subband coding techniques, which proved to be very promising in low bit-rate speech coding [15,3,17]. Subsequently the same ideas have also been applied in image coding [10], and in voice privacy systems [12]. Another situation where these systems arise (either as conceptual models or as practical structures) is in spectrum analysis [3], particularly in the context of short-time Fourier transformation [11]. Finally, the filter bank framework has recently been used [32] for derivation of new forms of sampling theorems, with possible applications again in signal compression.

The most commonly used type of filter bank is the analysis/synthesis system of Fig. 8, which is also called the maximally decimated quadrature mirror filter (QMF) bank [4-6,14,17,18]. The adjective *maximally* derives from the fact that decimation by a larger factor (than M) will result in loss of information about the signal.

IV. IMPERFECTIONS IN FILTER BANKS

In the QMF bank of Fig. 8, a common requirement is to make $\hat{x}(n)$ close to $x(n)$ in some well defined sense. It is not possible to avoid aliasing because the filters $H_k(z)$ are not ideal bandpass filters. The filters $F_k(z)$ can however be chosen to effect a cancelation between alias components. From the key relations (1),(4) we get [4-7,14]

$$\hat{X}(z) = \frac{1}{M}\sum_{\ell=0}^{M-1} X(zW^\ell)\sum_{k=0}^{M-1} H_k(zW^\ell)F_k(z). \tag{10}$$

The terms with $\ell \neq 0$ represent the alias components. If these can be successfully cancelled off by some means then the system is time-invariant with transfer function

$$T(z) = \frac{\hat{X}(z)}{X(z)} = \frac{1}{M}\sum_{k=0}^{M-1} H_k(z)F_k(z). \tag{11}$$

This function is called the Overall Transfer Function (OTF) or the Distortion Transfer Function (DTF). If $T(z)$ is all-pass, the signal $\hat{x}(n)$ is free from amplitude distortion; if $T(z)$ has linear-phase then $\hat{x}(n)$ is free from phase distortion. Finally, if $T(z)$ is a pure delay (i.e., $T(z) = cz^{-n_0}$) then $\hat{x}(n)$ is free from linear distortions of all kinds. We then have $\hat{x}(n) = cx(n - n_0)$, which is called the perfect-reconstruction property. [3]

The Alias-Component matrix. Let us define $H_{\ell k}(z) \triangleq H_k(zW^\ell)$. The matrix $\mathbf{H}(z) = [H_{\ell k}(z)]$ is called the alias-component (AC-)matrix. Using this, the conditions for absence of aliasing can be written as a matrix vector equation, as in [4, page 316]. If the analysis bank is of the uniform DFT type, $\mathbf{H}(z)$ turns out to be a circulant, which enables one to obtain simple closed form expressions for synthesis filters, as shown in [18].

In the two-band case case (10) reduces to

$$\hat{X}(z) = \frac{1}{2}[H_0(z)F_0(z) + H_1(z)F_1(z)]X(z) + \frac{1}{2}[H_0(-z)F_0(z) + H_1(-z)F_1(z)]X(-z) \tag{12}$$

The second term (which has $X(-z)$) represents aliasing. Clearly, if the synthesis filters are confined as $F_0(z) = H_1(-z)$ and $F_1(z) = -H_0(-z)$ this term vanishes and

$$T(z) = \frac{1}{2}[H_0(z)H_1(-z) - H_1(z)H_0(-z)]. \tag{13}$$

The earliest types of QMF banks employed this idea and in addition confined the analysis filters such that

$$H_1(z) = H_0(-z) \tag{14}$$

so that if $H_0(z)$ is lowpass then $H_1(z)$ is highpass. With this constraint we have

$$T(z) = \frac{1}{2}[H_0^2(z) - H_0^2(-z)]. \tag{15}$$

Under these conditions, if $H_0(z)$ is a linear-phase FIR filter this ensures that $T(z)$ has linear-phase, eliminating phase distortion. On the other hand, if we wish to eliminate only amplitude distortion (and tolerate phase distortion) how shall we proceed? Let us express $H_0(z)$ as

$$H_0(z) = E_{00}(z^2) + z^{-1}E_{01}(z^2). \tag{16}$$

The distortion function (15) can then be expressed as $T(z) = 2z^{-1}E_{00}(z^2)E_{01}(z^2)$. This gives us a nice hint: if we wish

[3] In practice, even if all linear distortions are eliminated, $\hat{x}(n)$ will suffer from errors caused by nonlinear operations such as quantization and coding. For the rest of this paper, we shall ignore this error.

to force $T(z)$ to be allpass (rather than a linear-phase function), this can be done by restricting $E_{00}(z)$ and $E_{01}(z)$ to be allpass. In other words, if the polyphase components of $H_0(z)$ are allpass, phase distortion is absent.

In summary if $H_0(z)$ and $H_1(z)$ are such that

$$H_0 = \frac{1}{2}[a_0(z^2) + z^{-1}a_1(z^2)],\ H_1 = \frac{1}{2}[a_0(z^2) - z^{-1}a_1(z^2)] \tag{17}$$

where $a_0(z)$ and $a_1(z)$ are IIR allpass, then after canceling aliasing, $\hat{x}(n)$ is free from amplitude distortion since $T(z) = \frac{1}{2}z^{-1}a_0(z^2)a_1(z^2)$. The specific forms (17) of IIR filters are not very restrictive. It is known, for example, that Butterworth, Chebyshev and elliptic half-band digital filters of odd order are in this form [19-22].

For the M band case, Rothweiler [17] has proposed a scheme for alias-cancelation by making certain valid assumptions about the behavior of alias-components. To understand this, first note that when the output of the filter $H_k(z)$ is decimated by M, the alias-components are $H_k(zW^\ell)X(zW^\ell)$, $1 \le \ell \le M-1$. Of these, some components (Type 1) fall within the stopband of $F_k(z)$, and some (Type 2) fall within the *transition band* of $F_k(z)$. If $F_k(z)$ is a good bandpass filter with essentially the same spectral occupancy as $H_k(z)$, we can assume that Type 1 components have been well-attenuated. It then remains only to cancel off the Type 2 components. Now, if each of the analysis (and synthesis) filters has real coefficients (so that $|H_k(e^{j\omega})|$ is symmetric with respect to π), there is an overlap of the Type 2 alias components for successive values of k. This fact in ingeniously exploited in [17] for canceling these components, in the FIR case. The resulting filter bank system is approximately alias free, and has the usual distortion function (11).

A problem of natural interest in design of QMF banks is this: how do we simultaneously eliminate aliasing, amplitude and phase distortions completely? A novel technique for this was reported by Smith and Barnwell [23] (see also [24]) for the $M = 2$ case. For $M \neq 2$, if M is a power of two, one can build a binary tree structure (23), which is free of all distortions. For the case of arbitrary M, a procedure has been reported in [7] for designing perfect reconstruction QMF banks. It turns out that the results in [23] can also be interpreted using the *lossless AC-matrix* framework developed in [7]. To economize space we shall therefore move on directly to this topic.

V. PERFECT-RECONSTRUCTION FIR FILTER BANKS

The method proposed in [23] for two-band FIR perfect-reconstruction filter banks results in an AC-matrix with a particular algebraic property called *losslessness*. This observation was made in [7], and was then used to design FIR perfect reconstruction systems for arbitrary M. Other interesting results on perfect reconstruction, such as [25], will not be elaborated here for want of space.

Perfect reconstruction for arbitrary M

Fig. 9 shows a maximally decimated M band analysis/synthesis system with analysis filters $H_k(z) = z^{-k}$ and synthesis filters $F_k(z) = z^{-(M-1-k)}$. This system serves as a theoretical model, which can be used to understand and analyze more complicated QMF banks. First notice that this is a perfect reconstruction system because, each of the M branches transmits one out of M subsequences of $x(n)$, and these subsequences are nicely interlaced by the interpolator and the chain of delays. It can be formally proved for this system that $\hat{x}(n) = x(n-M+1)$.

Next refer back to Fig. 8. Each analysis filter $H_k(z)$ can be expressed in terms of its polyphase components as $H_k(z) = \sum_{\ell=0}^{M-1} z^{-\ell}E_{k\ell}(z^M)$. Similarly each synthesis filter can be expressed as $F_k(z) = \sum_{\ell=0}^{M-1} z^{-(M-1-\ell)}R_{\ell k}(z^M)$. Defining the $M \times M$ matrices $\mathbf{E}(z) \triangleq [E_{k\ell}(z)]$ and $\mathbf{R}(z) \triangleq [R_{\ell k}(z)]$, we can redraw Fig. 8 as in Fig. 10(a). Here $\mathbf{E}(z)$ and $\mathbf{R}(z)$ are termed as the polyphase component matrices. By using the identities in Fig. 5, we can reconfigure Fig. 10(a) as in Fig. 10(b) where $\mathbf{P}(z) \triangleq \mathbf{R}(z)\mathbf{E}(z)$.

Since Fig. 10(b) is merely a redrawing of Fig. 8, it holds for any maximally decimated QMF bank. If $\mathbf{P}(z)$ happens to be of the form $\mathbf{P}(z) = cz^{-r}\mathbf{I}$ then the delay cz^{-r} can be moved past the interpolators, and the system reduces to the one in Fig. 9 (with the additional delay of cz^{-rM} connected to the output). Conclusion: a QMF bank gives rise to perfect reconstruction, if its $\mathbf{P}(z)$ matrix is of the form $z^{-r}c\mathbf{I}$! The reconstructed signal is then given by $\hat{x}(n) = cx(n - rM - M + 1)$. In fact the *most general form* of $\mathbf{P}(z)$ which is necessary and sufficient for perfect reconstruction is given in [7], and differs from the above form only by inconsequential permutations. For the rest of the paper, we shall consider perfect-reconstruction to be synonymous to the above condition on $\mathbf{P}(z)$.

How shall we enforce the condition $\mathbf{P}(z) = cz^{-r}\mathbf{I}$ in practice? One procedure would be to define $\mathbf{R}(z) = \mathbf{E}^{-1}(z)$ and then construct the expressions for $F_k(z)$. The difficulty of this approach in general is that, even if the entries of $\mathbf{E}(z)$ are FIR, the entries of $\mathbf{R}(z)$ can be IIR, which in addition are not guaranteed to be stable. A partial solution to this difficulty is provided by consrtaining $\mathbf{E}(z)$ to be such that det $\mathbf{E}(z)$ is a delay. This guarantees that $\mathbf{R}(z)$ and hence $F_k(z)$ are FIR. But the entries of $\mathbf{R}(z)$ in general are polynomials of much higher order than those of $\mathbf{E}(z)$. This difficulty can be overcome by constraining $\mathbf{E}(z)$ to be a *lossless matrix,* which we shall describe next.

Lossless transfer matrices. [4] A $p \times r$ rational transfer matrix $\mathbf{H}(z)$ is said to be lossless if it is stable[5] and satisfies $\mathbf{H}^\dagger(e^{j\omega})\mathbf{H}(e^{j\omega}) = c\mathbf{I}$ for some constant c. So a lossless matrix $\mathbf{H}(z)$ is unitary on the unit circle. This implies, by analytic continuation, that $\tilde{\mathbf{H}}(z)\mathbf{H}(z) = c\mathbf{I}$ for *all*

[4] Here it will be helpful to review the notations defined in footnotes of Sec. I.

[5] We consider only causal systems [13], so stability is equivalent to restricting the poles of all entries of $\mathbf{H}(z)$ to be inside the unit circle.

values of z. It is important to notice that losslessness of $\mathbf{H}(z)$ implies that of $\mathbf{H}^T(z)$ if and only if $p = r$.

An immediate application of this definition is the following: consider the QMF bank of Fig. 8 with FIR $H_k(z)$. Suppose the $M \times M$ polyphase component matrix $\mathbf{E}(z)$ in Fig. 10(a) is lossless. Then $\tilde{\mathbf{E}}(z)\mathbf{E}(z) = c\mathbf{I}$ so that $\mathbf{E}^{-1}(z) = c^{-1}\tilde{\mathbf{E}}(z) = c^{-1}\mathbf{E}_*^T(z^{-1})$. As a result, if the synthesis-bank polyphase-component matrix $\mathbf{R}(z)$ is taken as

$$\mathbf{R}(z) = z^{-r}\mathbf{E}_*^T(z^{-1}) \tag{18}$$

then the QMF bank is a perfect reconstruction system. Here r is large enough to avoid non causal entries in $\mathbf{R}(z)$. With $\mathbf{R}(z)$ defined as in (18), it is easy to show [7,page 482] that the synthesis filters are FIR of the same length as $H_k(z)$:

$$F_k(z) = dz^{-\beta}H_k(z^{-1}) \tag{19}$$

where d and β are constants of no fundamental importance.

Other consequences of losslessness. If $\mathbf{E}(z)$ is lossless then the AC-matrix $\mathbf{H}(z)$ (see Sec. V) is lossless (and in fact the converse is also true; for proof see [7, page 481]). As a result the kth column of $\mathbf{H}(z)$ satisfies the property

$$\sum_{\ell=0}^{M-1} \tilde{H}_k(zW^\ell)H_k(zW^\ell) = c \quad \text{for all } z \tag{20}$$

for $0 \le k \le M-1$. Since $\mathbf{H}(z)$ is square, $\mathbf{H}^T(z)$ is lossless as well, so the 0th *row* of $\mathbf{H}(z)$ satisfies

$$\sum_{k=0}^{M-1} \tilde{H}_k(z)H_k(z) = c \quad \text{for all } z. \tag{21}$$

Eqn (21) says that the set of analysis filters forms a power-complementary set (Sec. III), whereas (20) says that each analysis filter $H_k(z)$ is a spectral factor of an Mth band filter $G_k(z)$. (Just define $G_k(z) = \tilde{H}_k(z)H_k(z)$; see Sec. III).

Determinant of a lossless matrix. If $\mathbf{E}(z)$ is lossless, then the determinant $\mathcal{E}(z) \triangleq \det \mathbf{E}(z)$ is stable and all-pass [7, page 491]. In particular if $\mathbf{E}(z)$ is FIR then we have

$$\mathcal{E}(z) \triangleq \det \mathbf{E}(z) = cz^{-(K-1)} \tag{22}$$

where c is some constant and $K > 0$ is an integer.

The degree of a lossless transfer matrix. The degree of any transfer matrix, by definition, is the minimum number of scalar delay elements (i.e., building blocks of the form z^{-1}) required to implement $\mathbf{E}(z)$. For example, if $\mathbf{E}(z) = z^{-1}\mathbf{I}$ then the degree is M because one scalar delay is required to connect each input to the corresponding output. For an $M \times M$ lossless system, the degree of $\mathbf{E}(z)$ is equal to the degree of $\mathcal{E}(z)$, i.e. [26,27],

$$\deg \mathbf{E}(z) = \deg \det \mathbf{E}(z) \tag{23}$$

This result is not true for arbitrary (i.e., non-lossless) $\mathbf{E}(z)$. Notice finally, that if $\mathbf{E}(z)$ is FIR and lossless, then its degree is equal to $K-1$ where K is the integer appearing in the determinant (22).

The design problem. In order to design an FIR QMF bank based on the above ideas, the impulse response coefficients $h_k(n)$, $0 \le k \le M-1$ of the analysis filters should be chosen such that a) the filters $H_k(z)$ have good stopband attenuation, and b) the matrix $\mathbf{E}(z)$ is lossless. This is then a constrained optimization problem. The coefficients of the matrix $\mathbf{E}(z)$ should be optimized to maximize an objective function such as

$$\Phi \triangleq \sum_{k=0}^{M-1} \int_{\text{stopband}} |H_k(e^{j\omega})|^2 d\omega \tag{24}$$

under the consraint that $\mathbf{E}(z)$ be lossless. Notice that losslessness of $\mathbf{E}(z)$ restricts $H_k(z)$ to satisfy (21) so that minimization of (24) also ensures good passbands for $H_k(z)$.

Parameterization of lossless matrices

It is clear then, that the success of the design problem hinges upon our ability to write down a closed form expression for $M \times M$ FIR lossless matrices. The closed form expression should have the following features: First, every $M \times M$ FIR lossless matrix of degree $\le K-1$ should be obtainable from the expression for appropriate values of the parameters. Second, as long as the parameters in the expressions are bounded in a well-defined way, the expressions should represent a lossless system. This will enable us to optimize the parameters on a computer without worrying about losing the losslessness of $\mathbf{E}(z)$. Finally, the number of parameters should be minimum (for a give M and $K-1$) so as to minimize the optimization effort.

Such a parameterization using angles has been outlined in [28], based on a state-space approach. For *any* set of real-valued parameters, the expression faithfully represents a lossless system. The only disadvantage of the parameterization in [28] is that, it involves computation of cosines and sines of these angles. Several such computations are required during each evaluation of the objective function (24) resulting in fairly large computer time for the constrained optimization to converge.

An improved parameterization. We now outline a different parameterization, which is free from this disadvantage, because it does not involve angles. This leads to faster and easier-to-implement optimization programs.

As a first step, let us construct a mathematical expression for the most general degree-one $M \times M$ causal FIR lossless system. Clearly $\mathbf{E}(z)$ must have the form $\mathbf{E}(z) = \mathbf{e}(0) + z^{-1}\mathbf{e}(1)$. This is required to be unitary for $z = e^{j\omega}$, in particular for $z = 1$. So we can write $\mathbf{E}(z) = (1 - z^{-1})\mathbf{S} + \mathbf{R}$ where $\mathbf{R}$ is a constant unitary matrix. We can show that this $\mathbf{E}(z)$ satisfies $\tilde{\mathbf{E}}(z)\mathbf{E}(z) = \mathbf{I}$ if and only if $\mathbf{S}\mathbf{S}^\dagger + \mathbf{S}\mathbf{R}^\dagger = \mathbf{0}$. After simplification we conclude that $\mathbf{E}(z)$ is lossless if and only if it can be expressed as $\mathbf{E}(z) = [\mathbf{I} - \mathbf{S}\mathbf{S}^\dagger + z^{-1}\mathbf{S}\mathbf{S}^\dagger]\mathbf{R}$. Further simplification is possible: it can be shown [26] than the rank of $\mathbf{S}$ cannot exceed one, so that $\mathbf{S}\mathbf{S}^\dagger$ can be rewritten as $\mathbf{v}\mathbf{v}^\dagger$ for some $M \times 1$ vector $\mathbf{v}$. It can also be shown that the norm of $\mathbf{v}$ should be unity. A more complete treatement is given in [26], proving the following lemma.

Lemma: A degree-one $M \times M$ causal FIR system is lossless if and only if it can be written in the form

$$\mathbf{E}(z) = [\mathbf{I} - \mathbf{v}\mathbf{v}^\dagger + z^{-1}\mathbf{v}\mathbf{v}^\dagger]\mathbf{R}. \tag{25}$$

where $\mathbf{R}$ is $M \times M$ unitary and $\mathbf{v}$ is a column vector of unit norm.

If we now multiply degree-one systems of this form, we can construct FIR lossless systems of higher degree. In fact more is true: *any* FIR lossless systems of any degree can be obtained in this manner. To be more precise, we have the following result [26]:

Theorem: An $M \times M$ causal FIR system of degree $K-1$ is lossless if and only if it can be written in the form

$$\mathbf{E}(z) = \mathbf{G}_{K-1}(z)\mathbf{G}_{K-2}(z) \ldots \mathbf{G}_1(z)\mathbf{R} \tag{26}$$

with

$$\mathbf{G}_n(z) \triangleq [\mathbf{I} - \mathbf{v}_n\mathbf{v}_n^\dagger + z^{-1}\mathbf{v}_n\mathbf{v}_n^\dagger] \tag{27}$$

where $\mathbf{R}$ is $M \times M$ unitary, and $\mathbf{v}_n$ are $M \times 1$ unit-norm vectors.

According to the theorem, every $M \times M$ causal FIR lossless system of degree $K-1$ can be written as in (26) and conversely, the form (26) implies that $\mathbf{E}(z)$ is such a lossless system!

Number of design-freedoms. How many degrees of freedom can we exercise while designing a degree $K-1$ FIR lossless system? We see that each vector $\mathbf{v}_n$ has M complex components, and $\mathbf{R}$ has M^2 complex entries. So there are $2M(K-1) + 2M^2$ real numbers involved in (26). However, the number of freedoms is smaller than this for two reasons. First, $\mathbf{v}_n$ has unit norm, and so has only $2M-1$ real-valued freedoms. Second, $\mathbf{R}$ is unitary and hence has only M^2 real-valued freedoms (see [29]). The number of design freedoms which should be optimized while designing the analysis filters is therefore equal to $N_f = M^2 + (K-1)(2M-1)$. If the filter coefficients are restricted to be real, then this number becomes $N_{f,real} = \binom{M}{2} + (K-1)(M-1)$.

Summarizing, the perfect-reconstruction QMF analysis bank design would proceed as follows: confine $\mathbf{E}(z)$ to be as in (26). Optimize the N_f (or $N_{f,real}$) parameters to minimize (24). This should result in optimal analysis filters. The synthesis filters are then obtained from (19). Fig. 11 shows the analysis-filter responses for a design example of this type (with $M = 3$), with real-coefficient filters. The three FIR analysis filters have length 56 each. The optimization took 213 cpu seconds on a VAX 11/750 (corresponding to about 4 minutes 'real time' during peak load hours). The initialization of the $N_{f,real}$ parameters was performed by using the eigenfilters approach [30].

VI. RELATION TO OTHER TOPICS

It has been shown in [9] that the set of necessary and sufficient conditions for the QMF bank of Fig. 8 to be alias free is that the matrix $\mathbf{P}(z)$ should have an algebraic structure called the *pseudocirculant* structure. This pseudocirculant property is closely related to the theory of block-processing [34-36]. To be more specific, the $M \times M$ blocked version [36] of any scalar transfer function must be a pseudocirculant. This same pseudocirculant property also arises as a necessary and sufficient condition for a linear periodically time varying system to be time *invariant*. These relations are studied in [9].

There is a close relation between *sampling theorems* and the theory of signal reconstruction in the filter bank of Fig. 8. This relation is studied in [32], by using the polyphase framework.

References:

[1] M. Bellanger, G. Bonnerot and M. Coudreuse, "Digital filtering by polyphase network: application to sampling rate alteration and filter banks," IEEE Trans. ASSP, pp. 109-114, April 1976.

[2] R. E. Crochiere and L. R. Rabiner, "Interpolation and decimation of digital signals: a tutorial review," Proc. IEEE, pp. 300-331, March 1981.

[3] R. E. Crochiere and L. R. Rabiner, *Multirate digital signal processing*, Englewood Cliffs, NJ: Prentice Hall, 1983.

[4] M. J. T. Smith and T. P. Barnwell, III, "A new filter-bank theory for time-frequency representation," IEEE Trans. ASSP, pp. 314-327, March 1987.

[5] M. Vetterli, "A theory of multirate filter banks," IEEE Trans. ASSP, pp. 356-372, March 1987.

[6]P. P. Vaidyanathan, "Quadrature mirror filter banks, M-band extensions and perfect-reconstruction techniques," IEEE ASSP magazine, pp. 4-20, July 1987.

[7]P. P. Vaidyanathan, "Theory and design of M-channel maximally decimated quadrature mirror filters with arbitrary M, having perfect reconstruction property," IEEE Trans. ASSP, pp. 476-492, April 1987.

[8] F. Mintzer, "On half-band, third-band and Nth band FIR filters and their design," IEEE Trans. ASSP, pp. 734-738, Oct. 1982.

[9] P. P. Vaidyanathan and S. K. Mitra, "Polyphase networks, block digital filtering, LPTV systems and alias-free QMF banks: a unified approach based on pseudo circulants," IEEE Trans. ASSP, 1988.

[10] J. W. Woods, and Sean D. O'Neil, "Subband coding of images," IEEE Trans. ASSP, pp. 1278-1288, Oct. 1986.

[11] M. R. Portnoff, "Time-frequency representation of digital signals and systems based on short-time fourier analysis," IEEE Trans. ASSP, pp. 55-69, Feb 1980.

[12] R. V. Cox, D. E. Bock, K. B. Bauer, J. D. Johnston, and J. H. Snyder, "The analog voice privacy systems," Proc. ICASSP, pp. 341-344, April 1986.

[13] A. V. Oppenheim and R. W. Schafer, *Digital signal processing*, Prentice Hall, Inc., Englewood Cliffs, NJ, 1975.

[14] T. Ramstad, "Analysis/synthesis filter banks with critical sampling," Int. Conf. DSP, Florence, Sept. 1984.

[15] A. Croisier, D. Esteban, and C. Galand, "Perfect channel splitting by use of interpolation/decimation/ tree decomposition techniques," Int. Symp. on Info., Circuits and Systems, Patras, Greece, 1976.

[16] C. R. Galand and H. J. Nussbaumer, "New quadrature mirror filter structures," IEEE Trans. ASSP, pp. 522-531,

June 1984.
[17] J. H. Rothweiler, "Polyphase quadrature filters: a new subband coding technique," Proc. ICASSP, pp. 1980-1983, April 1983.
[18] K. Swaminathan and P. P. Vaidyanathan, "Theory and design of uniform DFT, parallel QMF banks," IEEE Trans. CAS, pp. 1170-1191, Dec. 1986.
[19] P. P. Vaidyanathan, S. K. Mitra and Y. Neuvo, "A new approach to the realization of low sensitivity IIR digital filters," IEEE Trans. ASSP, pp. 350-361, April 1986.
[20]P. P. Vaidyanathan, P. Regalia and S. K. Mitra, "Design of doubly complementary IIR digital filters using a single complex allpass filter, with multirate applications," IEEE Trans. CAS, pp. 378-389, April 1987.
[21] L. Gaszi, "Explicit formulas for lattice wave digital filters," IEEE Trans. CAS, pp. 68-88, Jan. 1985.
[22] T. Saramaki, "On the design of digital filters as a sum of two allpass filters," IEEE Trans. CAS, Nov. 1985.
[23] M. J. T. Smith and T. P. Barnwell III, "Exact reconstruction techniques for tree-structured subband coders," IEEE Trans. ASSP, pp. 434-441, June 1986.
[24] F. Mintzer, "Filters for distortion-free two-band multirate filter banks," IEEE Trans. ASSP, June 1985.
[25] M. Vetterli, "Filter banks allowing for perfect reconstruction," Signal Processing, pp. 219-244, April 1986.
[26] P. P. Vaidyanathan and Z. Doganata, "The role of lossless systems in modern digital signal processing," in preparation.
[27] V. Belevitch, *Classical network theory*, Holden Day, Inc., San Francisco, 1968.
[28] P. P. Vaidyanathan, Z. Doganata and T. Q. Nguyen, "More results on the perfect reconstruction problem in *M*-band parallel QMF banks," Proc. ISCAS, May 1987.
[29] F. D. Murnaghan, *The unitary and rotation groups*, Spartan Books, Washington, D.C., 1962.
[30] P. P. Vaidyanathan, T. Q. Nguyen and T. Saramaki, "Improved approach for design of perfect reconstruction FIR QMF banks, with lossless lattice strutures," Proc. ICASSP, April 1988.
[31] P. P. Vaidyanathan and Phuong-Quan Hoang, "Lattice structures for optimal design and robust implementation of two-channel perfect reconstruction QMF banks," IEEE Trans. ASSP, pp. 81-94, Jan 1988.
[32] P. P. Vaidyanathan, and Vincent C. Liu, "Sampling theorems in the context of multirate and polyphase digital filter bank structures," IEEE Trans. ASSP, 1988 (to appear).
[33] B. D. O. Anderson and S. Vongpanitlerd, *Network analysis and synthesis*, Prentice Hall, Inc., 1973.
[34] T. G. Marshall, Jr., "The polyphase transform and its applications to block-processing and filter-bank structures," Proc. ISCAS, pp. 1103-1109, May 1987.
[35] R. A. Meyer and C. S. Burrus, "A unified analysis of multirate and PTV digital filters," IEEE Trans. CAS, pp. 162-168, March 1975.
[36] S. K. Mitra and R. Gnanasekaran, "Block implementation of recursive digital filters-new structures and properties," IEEE Trans. CAS, pp. 200-207, April 1978.

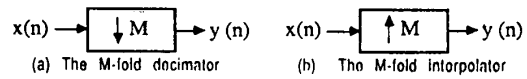


Fig. 1. The basic multirate building blocks.

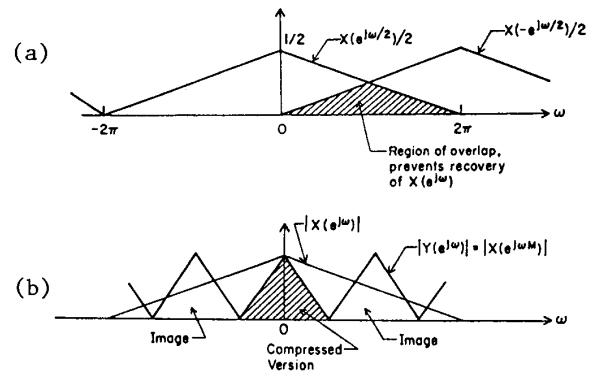


Fig. 2. a)Effect of decimation by two, b)effect of interpolation.

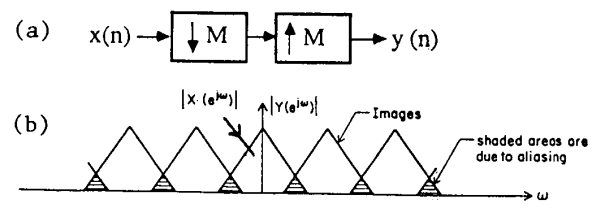


Fig. 3. A decimator followed by an interpolator.

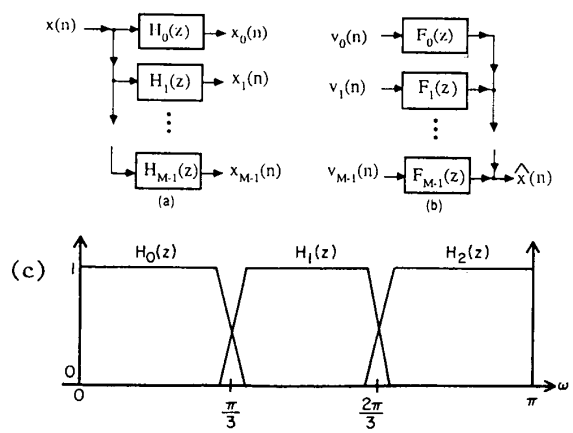


Fig. 4. Analysis and synthesis banks.

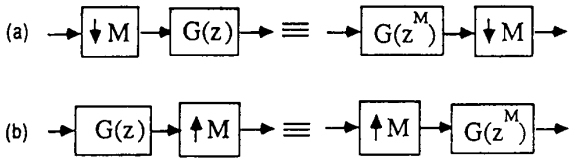


Fig. 5. Useful multirate identities.

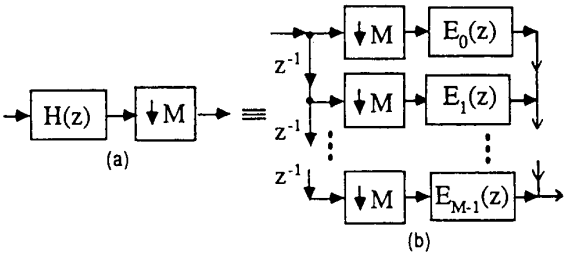


Fig. 6. Efficient implementation of a decimation filter, using the polyphase network approach.

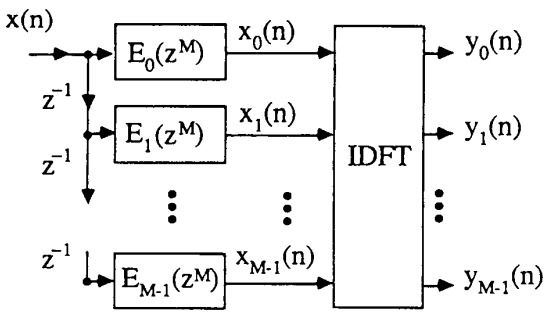


Fig. 7. A uniform DFT analysis-filter bank.

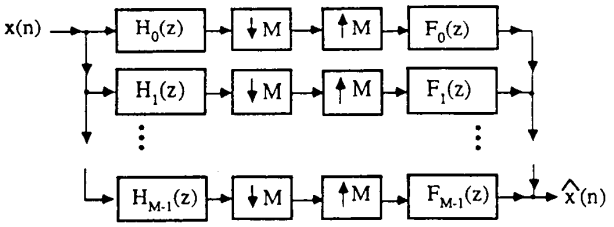


Fig. 8. The M-band maximally-decimated analysis/synthesis (QMF) bank.

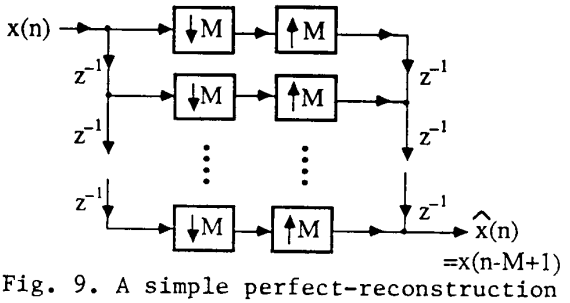
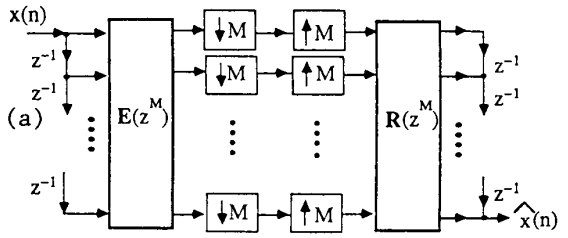


Fig. 9. A simple perfect-reconstruction analysis/synthesis system.



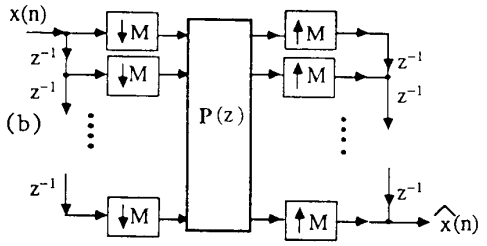


Fig. 10. Redrawing of Fig. 9, with polyphase matrices.

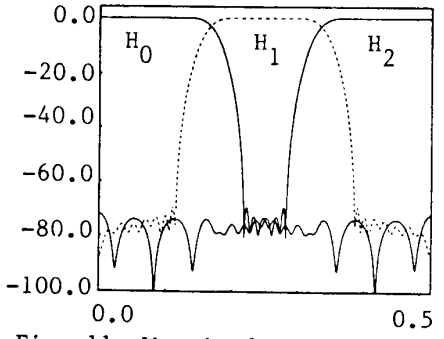


Fig. 11. Magnitude response in DB versus normalized frequency, for optimized analysis filters.