Efficient and Multiplierless Design of FIR Filters with Very Sharp Cutoff via Maximally Flat Building Blocks

P. P. VAIDYANATHAN, MEMBER, IEEE

Abstract — A new design technique for linear-phase FIR filters, based on maximally flat building blocks, is presented. The design technique does not involve iterative approximations and is, therefore, fast. It gives rise to filters that have a monotone stopband response, as required in some applications. The technique is partially based on an interpolative scheme. Implementation of the obtained filter designs requires a much smaller number of multiplications than maximally flat (MAXFLAT) FIR filters designed by the conventional approach. A technique based on FIR spectral transformations to design new multiplierless FIR filter structures is then advanced, and multiplierless implementations for sharp cutoff specifications are included.

I. INTRODUCTION

A NUMBER OF methods for the design of optimum linear-phase FIR filters are known today [1]–[3]. The methods that lead to optimal filters are generally derived by iterative approximation techniques. The well-known Parks–McClellan program [3] can be used to design FIR filters for a wide range of applications, and improved versions of the algorithm that speed up the iteration process are also known [4].

In contrast to these methods, there exist noniterative techniques for the design of FIR filters. Most of these are useful for obtaining "quick designs." The main disadvantage of such techniques is that the resulting filters generally require larger number of multiplications for implementation as compared to the filters based on optimization techniques. Examples of this type of design procedure are the windowing technique [5], the Chebyshev-function based techniques [6], and the MAXFLAT techniques [7].

In this paper, we are primarily concerned with the maximally flat (MAXFLAT) FIR filters which can be designed "quickly" using Kaiser's method [8], without the use of iterative algorithms. For a given filter order and center of transition band, these filters are maximally-flat at $\omega = 0$ and $\omega = \pi$ (for low- and high-pass filters) and are valuable in situations where the gain is required to fall in a monotone fashion (as a function of frequency) in the entire stopband. A typical example where a monotone response is desirable is in the design of filter banks based on the DFT

algorithm, where interchannel-interference is an important consideration.

As mentioned above, "MAXFLAT FIR filters" have a monotone gain response in the entire frequency band $0 \leqslant \omega \leqslant \pi$, and can be easily designed. However, as pointed out by Kaiser [8], the main disadvantage of these filters is that the required filter order is approximately inversely proportional to the square of the transition bandwidth; the design procedure typically leads to filters of much higher order than those designed to have equiripple approximation error, i.e., in implementation of these designs the number of multiplications required per computed output sample is large. Another property of these filters is that many of the impulse response coefficients, $h(n)$, are very small, i.e., $h(n)$ generally tends to have a large dynamic range.

The aim of this paper is to introduce a new technique for the design of filters with monotone frequency response. The design procedure is based on maximally flat building blocks and does not involve iteration; it is, therefore, fast. In implementation we employ an "interpolation scheme" that dramatically reduces the total number of multiplications. In addition, these techniques lead to new "multiplierless" FIR digital filter structures, which can achieve sharp-cutoff responses, with stopband attenuation exceeding 100 dB (equivalent to 17-bit dynamic range) almost everywhere.

In Section II the concept of maximally flat FIR filter design is briefly reviewed. In Section III we introduce the "maximally flat interpolators" and then describe a new technique for efficient filter design. In Section IV numerical examples are presented that demonstrate the power of the new technique. New multiplierless implementations of these designs are introduced in Section V.

II. MAXFLAT FIR FILTERS

A zero-phase FIR filter transfer function $H(z)$ of order N can be written as

$$H(z) = h_0 z^{N/2} + h_1 z^{N/2-1} + \cdots + h_N z^{-N/2} \qquad (1)$$

where $h_n = h_{N-n}$. Odd-order filters cannot be used to realize highpass functions whereas even-order filters can be used to realize almost any type of filter [2]. Let us, therefore, assume that N is even.

The frequency response of the filter can be written in equivalent forms as

$$H(e^{j\omega}) = \sum_{n=0}^{N/2} c_n \cos^n \omega = h_{N/2} + 2\sum_{n=1}^{N/2} h_{(N/2)-n} \cos n\omega.$$

$$(2)$$

$H(e^{j\omega})$ is real and represents the "gain" of the filter. The filter is said to be maximally flat if the quantity $H(e^{j\omega})$ has the following property:

$$\left.\frac{\partial^n H(e^{j\omega})}{\partial \omega^n}\right|_{\omega=0} = 0, \quad \text{for } n = 1,2,\cdots,2L-1 \quad (3a)$$

$$\left.\frac{\partial^n H(e^{j\omega})}{\partial \omega^n}\right|_{\omega=\pi} = 0, \quad \text{for } n = 1,2,\cdots,2K-1 \quad (3b)$$

where the quantities K and L are related as

$$N \triangleq \text{filter order} = 2(K+L-1). \quad (3c)$$

Thus for a given N and K, L represents the maximum degree of flatness at $\omega = 0$. Similarly, for a given N and L, K represents the maximum degree of flatness at $\omega = \pi$.

It can be shown, based on the results in [7] that such a low-pass function is given by

$$H(\omega) = \left(\cos \frac{\omega}{2}\right)^{2K} \sum_{n=0}^{L-1} d(n)\left(\sin \frac{\omega}{2}\right)^{2n} \quad (4a)$$

where the coefficients $d(n)$ are given by

$$d(n) = \frac{(K-1+n)!}{(K-1)!\,n!}. \quad (4b)$$

This result is obtained by recognizing that the problem of designing a maximally flat filter is precisely a Hermite-interpolation problem, for which the solution is well known. From (4) it is seen that $H(z)$ can be written as

$$H(z) = \left(\frac{1+\dfrac{z+z^{-1}}{2}}{2}\right)^{K} \sum_{n=0}^{L-1} d(n)\left(\frac{1-\dfrac{z+z^{-1}}{2}}{2}\right)^{n} \quad (5)$$

and can, therefore, be implemented with binary shifts and only L nontrivial multipliers $d(n)$. However, the multiplier coefficients $d(n)$ span a huge dynamic range and are not convenient to implement. In [8], Kaiser overcomes this problem by computing the actual impulse response $h(n)$ as the inverse $N+1$ point DFT of the frequency response of (4a), and using the coefficients $h(n)$ as the multipliers. In general, we, therefore, have $N/2+1$ multipliers, even though a large number of them are very small in value. Depending upon the level of stopband attenuation that one is willing to sacrifice, some of these impulse response coefficients can be set to zero.

In order to design a linear-phase MAXFLAT FIR filter, one begins with a specification of the center of transition band β and the transition bandwidth δ (both are in multiplies of π in this paper) and computes the two integers K and L such that the resulting low-pass filter has a gain of 0.95 at $\omega = (\beta - \delta/2)\pi$ and 0.05 at $\omega = (\beta + \delta/2)\pi$. The order of the filter, $N = 2(K+L-1)$ grows as
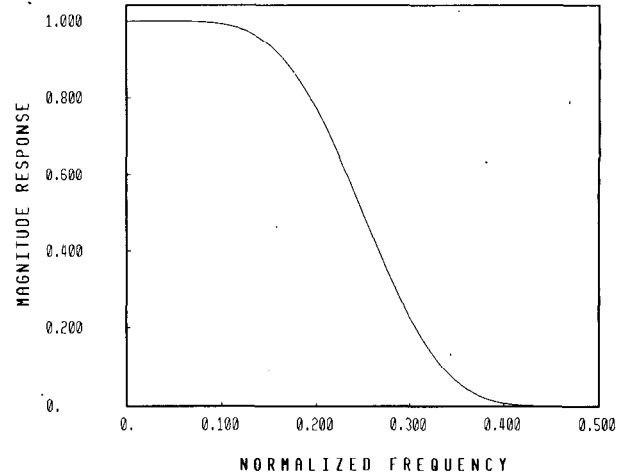

Fig. 1.   The maximally flat interpolator $I(z)$.

$1/\delta^2$ and is, therefore, much larger than for an equiripple filter approximation with 26 dB (0.05) stopband loss.

Finally, note that if a causal system is desired, then the zero-phase filter can be converted to a linear-phase filter by setting

$$H_c(z) = z^{-N/2}H(z) = \left(\frac{1+z^{-1}}{2}\right)^{2K}$$
$$\cdot \sum_{n=0}^{L-1} z^{-(L-1-n)}(-1)^n d(n)\left(\frac{1-z^{-1}}{2}\right)^{2n}. \quad (6)$$

III. THE "MAXFLAT INTERPOLATOR"

A low-pass MAXFLAT FIR filter with $\beta = 0.5$ and $\delta = 0.4$ has a frequency response:

$$I(e^{j\omega}) = \cos^6 \frac{\omega}{2}\left[1 + 3\left(\sin^2 \frac{\omega}{2} + 2\sin^4 \frac{\omega}{2}\right)\right]. \quad (7a)$$
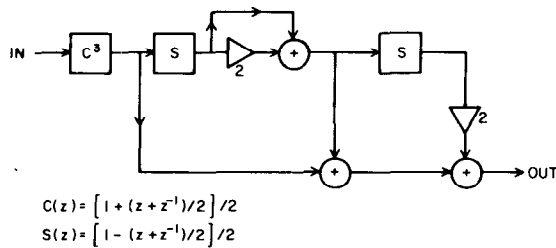
This filter clearly has $K = 3$, $L = 3$, $N = 10$, and though it requires $N/2+1 = 6$ multipliers theoretically, it can be implemented with only one multiplier of value equal to "3." (The remaining operations are shifting and adding):

$$I(z) = \left(\frac{z+z^{-1}+2}{4}\right)^3$$
$$\cdot \left\{1 + 3\left[\frac{2-z-z^{-1}}{4} + 2\left(\frac{2-z-z^{-1}}{4}\right)^2\right]\right\}. \quad (7b)$$

A plot of $I(e^{j\omega})$ is shown in Fig. 1. This filter will be called a "maximally flat interpolator" for reasons to be made clear soon. Note that the multiplier of value "3" can be implemented as "2+1" and, therefore, $I(z)$ is essentially "multiplierless." The implementation of $I(z)$ is shown in Fig. 2.

As mentioned earlier, N grows as $1/\delta^2$. By employing an indirect design approach, this can be avoided. The scheme we employ here is based on the "IFIR approach" introduced in [9]. We first design a filter with a new set of β and δ given by

$$\hat{\beta} = M\beta \qquad \hat{\delta} = M\delta \quad (8)$$

$C(z) = [1 + (z + z^{-1})/2]/2$

$S(z) = [1 - (z + z^{-1})/2]/2$

Fig. 2.   Multiplierless implementation of $I(z)$.

Fig. 3.   The interpolation technique.

where M is a judiciously chosen integer, $M \geqslant 1$. (Choice of M is discussed at the end of this section.) For example, if the design specifications of the low-pass filter are such that $\beta = 0.15$ and $\delta = 0.1$, then a direct design procedure would lead to the following results:

$$K = 104 \qquad L = 6 \qquad N = 218 \qquad (9)$$

and the implementation of this filter $H(z)$ would require 110 multipliers. We, however, proceed as follows: first design a low-pass filter $H_1(z)$ with $\beta = 0.30$ and $\delta = 0.2$, leading to

$$K = 27 \qquad L = 7 \qquad N = 66 \qquad (10)$$

which requires only 34 multipliers. The resulting filter $H_1(z)$ has frequency response, as shown in Fig. 3(a). We then replace each delay in the structure with two delays in cascade leading to a multiband filter, as shown in Fig. 3(b). The corresponding impulse response sequence is

$$h_1(0), 0, h_1(1), 0, h_1(2), 0, h_1(3), \cdots .$$

Finally, we suppress the unwanted passband around $\omega = \pi$ and retain only the desired passband by cascading with the maximally flat interpolator $I(z)$. This operation is equivalent to replacing the zero-valued samples in the impulse response with a weighted average of surrounding impulse response coefficients. This justifies the name "interpolator" for $I(z)$. (Note that there is no explicit interpolation of signals in the circuit. The sampling rate is the same throughout.) The most crucial thing to be noticed is that the passband of $I(z)$ is much wider than the desired passband of $H_1(z^2)$ and moreover $I(z)$ is maximally flat. Therefore, it does not deteriorate the response in the desired passband of $H_1(z^2)$. In addition, $I(z)$ suppresses the undesired passband of $H_1(z^2)$ very satisfactorily, once again because of maximal flatness. In fact, this suppression can be improved by using $I^2(z)$ or even $I^3(z)$ as the "interpolator." Each $I(z)$ requires only additions and shifts, and in the above example, we have cut down the number of multipliers from 110 to 34. Notice also that we can trivially get a highpass filter by using $I(-z)$ rather than $I(z)$, in conjunction with $H_1(z^2)$.

A Family of Useful "Interpolators"

As mentioned earlier, the coefficients $d(n)$ span a large dynamic range for large K and L. In all the interpolators we plan to use, we, therefore, wish to keep K and L small so that the interpolators are simple to implement. In addition, we take the value of "β" of the interpolators to be
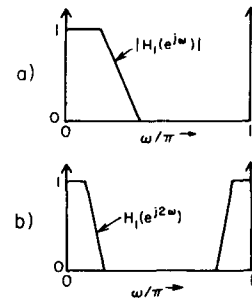
close to 0.5, so that these interpolators are most appropriate for the choice $M = 2$ in (8). For higher values of M, the interpolators can be appropriately modified to generate other efficient interpolators, as demonstrated in the examples of Section IV. With this rationale, we list the following four "basic interpolators," which we found to be most useful. The symbols C and S stand for

$$C(\omega) = \cos^2 \frac{\omega}{2} \qquad S(\omega) = \sin^2 \frac{\omega}{2} . \qquad (11)$$

1) *Interpolator $I(z)$:*

$$\beta = 0.5, \quad \delta = 0.4, \quad K = 3, \quad L = 3, \quad N = 10$$

$$I(e^{j\omega}) = C^3(\omega)(1 + 3S(\omega) + 6S^2(\omega))$$

$$= C^3(\omega)(1 + 3(S(\omega) + 2S^2(\omega))). \qquad (12)$$

2) *Interpolator $J(z)$:*

$$\beta = 0.6, \quad \delta = 0.5, \quad K = 2, \quad L = 4, \quad N = 10$$

$$J(e^{j\omega}) = C^2(\omega)(1 + 2S(\omega) + 3S^2(\omega) + 4S^3(\omega)). \qquad (13)$$

3) *Interpolator $K(z)$:*

$$\beta = 0.4, \quad \delta = 0.4, \quad K = 4, \quad L = 2, \quad N = 10$$

$$K(e^{j\omega}) = C^4(\omega)(1 + 4S(\omega)). \qquad (14)$$

4) *Interpolator $L(z)$:*

$$\beta = 0.5, \quad \delta = 0.5, \quad K = 2, \quad L = 2, \quad N = 6$$

$$L(e^{j\omega}) = C^2(\omega)(1 + 2S(\omega)). \qquad (15)$$

Notice that all these interpolators are "multiplierless" in the sense described earlier.

A further motivation for the choice of the above interpolators with low values of K and L is that, if an implementation as described in (4a) is used with large K and L, the roundoff noise generated by the interpolators becomes large. This can be seen by noting that, near the passband edge of $H(z)$ (equation (4a)), the quantity "$\cos^{2K}(\omega/2)$" is very small for large K, and hence the summation in (4a) is large-valued to compensate for a passband droop. This would result in a large noise-gain, if K and L were large.

Note that the technique of replacing z with z^M and then cascading an interpolator to suppress extra passbands is dealt with in [9] and is called the "IFIR (interpolated FIR) approach." The idea of improving the efficiency of implementation of narrow-band FIR filters, based on interpolation and decimation techniques, is known [11]. In this paper, we do not explicitly deal with changes of sampling
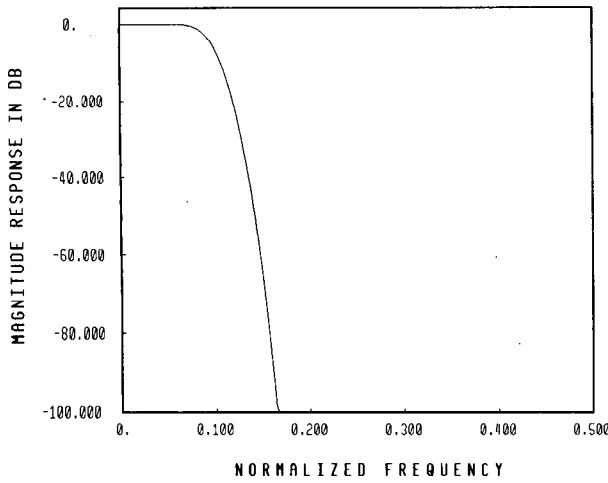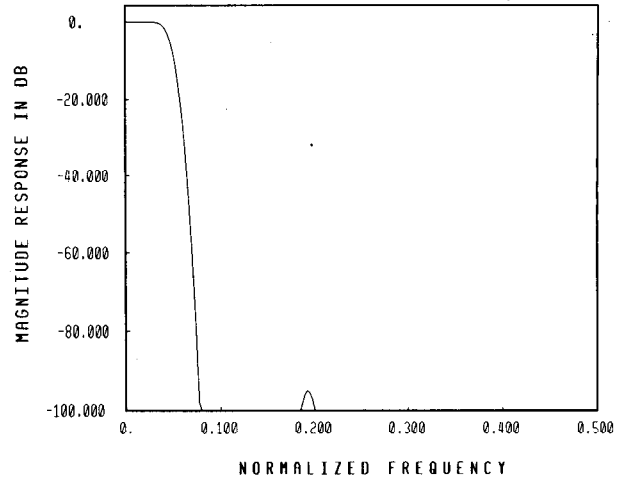
Fig. 4.   Example 1, indirect design.

Fig. 5.   Example 2, indirect design.

rate in our system, and the method in this section is closer to the one discussed in [9].

Choice of the Parameter M in (8)

The integer M should be small enough so that $\hat{\beta} < 1$, otherwise $H_1(z)$ resembles an allpass filter. The value of M can, in theory, be chosen as large as possible, subject only to this constraint. A larger M implies a larger value of $\hat{\delta}$, hence a smaller order of $H_1(z)$. However, a large M also implies that the passbands in $H_1(z^M)$ are closely spaced, and the suppression of unwanted passbands by means of "interpolators" becomes a difficult problem. Our experience shows that as long as $(\hat{\beta} + \hat{\delta}/2)$ is less than 0.75 the designs are satisfactory.

IV. DESIGN EXAMPLES BASED ON MAXIMALLY FLAT INTERPOLATORS

Example 1: Let the filter specifications be such that $\beta = 0.2$ and $\delta = 0.1$. If a direct design of $H(z)$ is attempted, this leads to:

$$K = 161 \qquad L = 17 \qquad N = 354 \qquad (16)$$

requiring 178 multipliers. If we first design a maximally flat filter $H_1(z)$ with $\beta = 0.4$ and $\delta = 0.2$, we require

$$K = 17 \qquad L = 9 \qquad N = 50. \qquad (17)$$

The frequency response of the filter that results by cascading $H_1(z^2)$ and $I^3(z)$ is shown in Fig. 4 and requires only 26 multipliers. The amount of saving is, therefore, nearly a factor of 7.

Note that, in this example, the integer M is taken equal to 2. If we had used $M = 3$, then $(\hat{\beta} + \hat{\delta}/2)$ would become equal to 0.75, and the suppression of unwanted passbands in $H_1(z^3)$ would be an expensive operation.

Example 2: Let the specifications on $H(z)$ be $\beta = 0.101$ and $\delta = 0.058$. A direct design would lead to

$$K = 547 \qquad L = 14 \qquad N = 1120 \qquad (18)$$

requiring 561 multipliers.

The specifications on β and δ can be met indirectly as follows: Design a maximally flat filter $H_1(z)$ with $\beta = 0.4$

and $\delta = 0.2$, requiring 26 multipliers, and then construct the overall filter as follows:

$$H^2(z) = H_1(z^4)\left[J(z^4)J(z^2)J(z)\right]^4 \qquad (19)$$

to suppress all the undesired passbands completely. The total number of multipliers required is only 26. In fact, it is $26 + 4 \times 3 = 38$, even after counting multipliers of value "3." Thus the saving is $561/38 \sim 15$ fold! Fig. 5 shows the response obtained by the new approach. Note that there is a small stopband ripple of height -95 dB, and apart from this, we have a monotone response.

In this example, the given specifications correspond to a narrowband filter, and therefore, a large value of M has been possible. This is because the separation between adjacent passbands in $H_1(z^M)$ is sufficiently large.

Example 3: Equiripple Designs: For equiripple filters, the filter order increases as the reciprocal of the transition bandwidth (rather than as reciprocal squared) [12]. Consequently, the "interpolation" scheme leads to savings that are not as dramatic as for maximally flat filters. For example, an equiripple filter $G(z)$ with the following characteristics:

Passband edge $= 0.15\pi$,   Stopband edge $= 0.25\pi$,

Passband ripple (peak to peak) $=$ Stopband ripple

$$= 0.05$$

has order equal to 31. The same specifications can be met by first designing an equiripple filter $G_1(z)$ with

Passband edge $= 0.3\pi$,   Stopband edge $= 0.5\pi$

requiring an order $= 17$, and then constructing

$$H(z) = G_1(z^2)I(z).$$

The total number of multipliers required by $H(z)$ is, therefore, 9 as against 16 required by $G(z)$. The saving is only a factor of 2, which is not as dramatic as for maximally flat filters.

Fig. 6 shows the responses of $G(z)$ and $H(z)$. Notice that $H(z)$ is actually even better than $G(z)$ in the stopband. Moreover, the filter $H(z)$ is equivalent to the filter designed in Example 1, as far as ω_p, ω_s and ripples are
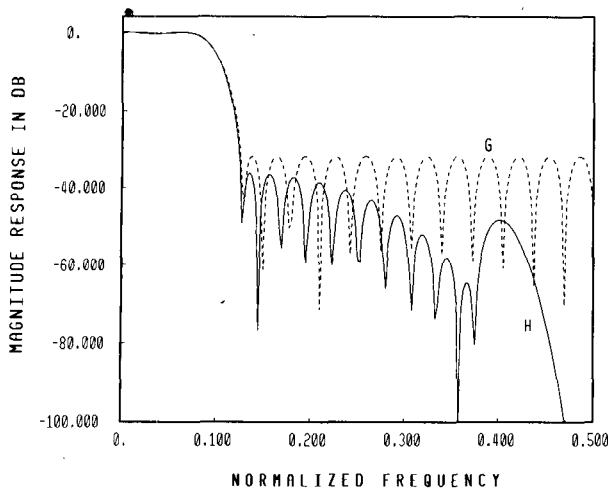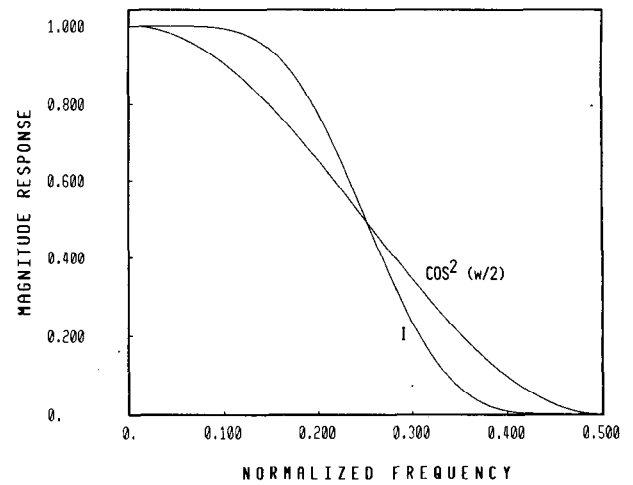
Fig. 6.   Example 3, indirect design.

Fig. 7. $I(e^{j\omega})$ and $\cos^2(\omega/2)$.

concerned. In Example 1, we however required 26 multi-pliers, which is the price paid for a monotone stopband attenuation, exceeding 100 dB almost everywhere.

V.   MULTIPLIERLESS LINEAR-PHASE FIR FILTERS

In this section, we present a new scheme for designing multiplierless FIR filters, based on maximally flat building blocks.

Frequency Transformations in FIR Filters

For IIR filters, it is well known that a valid frequency transformation can be achieved by replacing delays in a prototype filter with stable allpass functions. This replace-ment can also be done on FIR filter prototypes, but the resulting filters are not FIR and certainly are not linear-phase filters. We, therefore, require a different kind of spectral transformation scheme. The idea of frequency transformations in FIR filters which preserve the FIR nature is known [10], and we extend and modify these ideas in this section, to suit our purposes. The multiplier-less structures we wish to introduce are based on the frequency transformation concept.

The frequency response of a linear-phase FIR filter of even order can always be expressed entirely in terms of the function $\cos^2(\omega/2)$, which corresponds to the building block

$$C(z) = \left[1 + (z + z^{-1})/2\right]/2. \tag{20}$$

See, for example, (7). On the unit circle, this function has the property: $0 \leqslant C(e^{j\omega}) \leqslant 1$. If each building block $C(z)$ is now replaced by a function $F(z)$ which also has the property

$$0 \leqslant F(e^{j\theta}) \leqslant 1, \quad \text{for all } \theta \tag{21}$$

then this constitutes a valid frequency transformation, because for each new frequency θ we can compute a prototype frequency ω from

$$\omega = 2\cos^{-1}\left\{\left[F(e^{j\theta})\right]^{1/2}\right\} \tag{22}$$

In view of (21), it is clear that (22) gives rise to a real

valued ω, and, therefore, each new frequency corresponds to a unique prototype frequency. Consequently, replacing the building block $C(z)$ with $F(z)$ is a valid frequency transformation. As long as $F(z)$ is itself linear-phase FIR (and the prototype is linear-phase FIR), the resulting filter is linear-phase FIR.

Nesting of Maximally Flat Building Blocks

Consider now, the frequency response of the maximally flat, multiplierless interpolator $I(z)$, shown in Fig. 1, and given in (7a). Fig. 7 shows a plot of $I(e^{j\omega})$ along with a plot of $\cos^2(\omega/2)$. Notice that both functions are positive, monotone in ω, and are in the range $(0,1)$. Moreover, $I(e^{j\omega})$ is much flatter around $\omega = 0$ and π, and falls much more sharply than $\cos^2(\omega/2)$, around $\omega = \pi/2$.

Given a structure for $I(z)$, if we now replace each building block of the form (20) with $I^2(z)$, we obtain a new transfer function:

$$I_I(z) = I(z)|_{C(z) \leftarrow I^2(z)} \tag{23}$$

that has a much sharper cutoff than $I(z)$ itself. The transformation of (23) maps the unit circle onto the unit circle in a one-to-one manner and is a valid lowpass to low-pass frequency transformation. Fig. 8 shows plots of $I_I(e^{j\omega})$ and $I(e^{j\omega})$. Notice that $I_I(z)$ is multiplierless. (Note that $C(z)$ could also have been replaced with $I(z)$ rather than $I^2(z)$. There are several other possibilities leading to a large family of transfer functions.)

In a similar manner, we can nest any maximally flat, multiplierless building block into another and generate a large number of sharp-cutoff, multiplierless filters. Thus we have $I_J(z)$, $J_J(z)$, $J_I(z)$, $K_I(z)$, $\cdots$, and so on. In general, the notation $H_G(z)$ stands for a linear-phase FIR transfer function obtained as follows:

$$H_G(z) = H(z)|_{C(z) \leftarrow G^2(z)}. \tag{24}$$

In addition, we can do further levels of nesting, leading to further sharpening of the frequency response. This in-creases the number of additions (and shifts) to a large number, but is still "multiplierless" in principle. Fig. 9
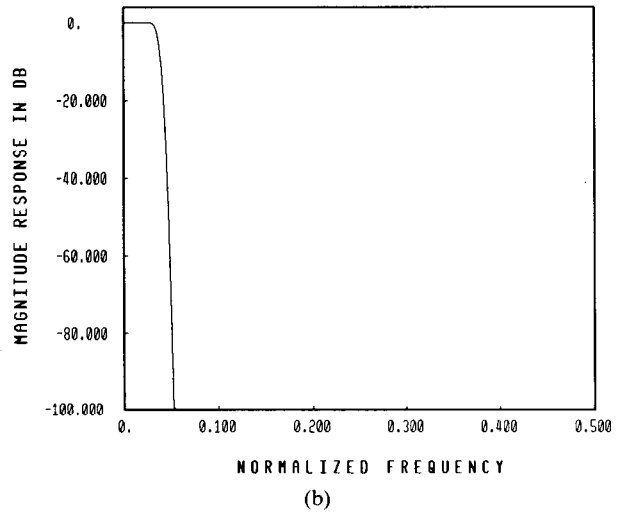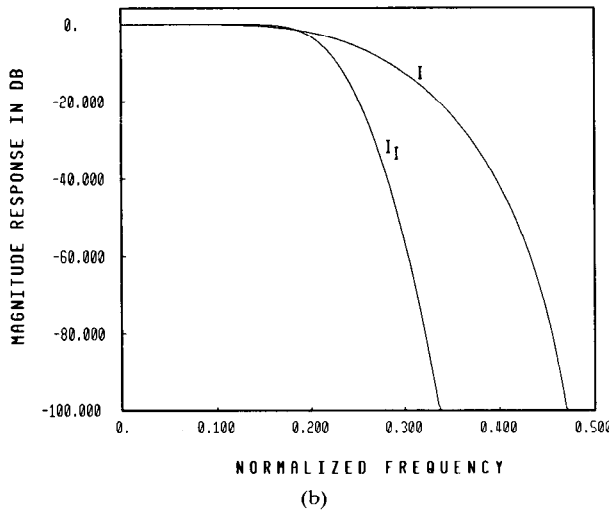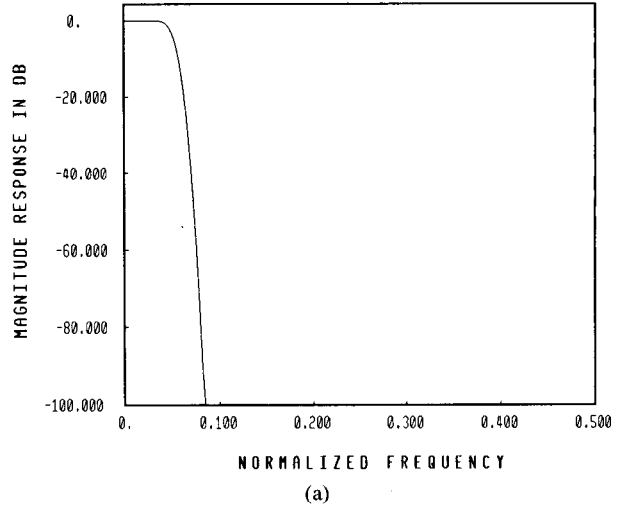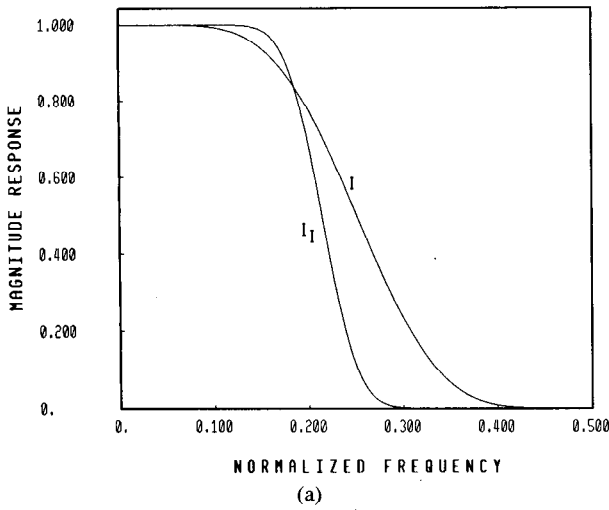
Fig. 8. (a) $I(e^{j\omega})$ and $I_I(e^{j\omega})$. (b) $I(e^{j\omega})$ and $I_I(e^{j\omega})$ in decibels.

Fig. 9. Six levels of nesting $I(e^{j\omega})$.

Fig. 10. (a) $|H_1(e^{j\omega})|$ in decibels. (b) $|H_2(e^{j\omega})|$ in decibels. (c) $|H_3(e^{j\omega})|$ in decibels. (d) $|H_4(e^{j\omega})|$ in decibels.
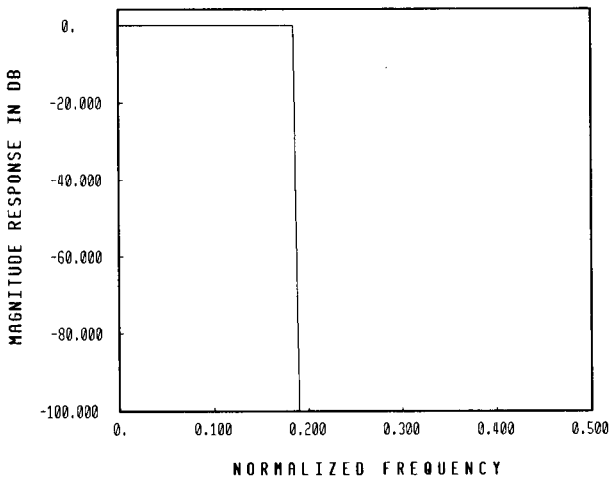
Design of Sharp-Cutoff, Narrowband, Linear-Phase Multiplierless FIR Filters

The idea of "interpolated FIR filters" can be combined with the nesting concept to generate a large class of multiplierless FIR transfer functions. For example, consider the function:

$$H_1(z) = I_I(z)I_I(z^2)I_I(z^4). \qquad (25a)$$

It is easy to see that this gives rise to a low-pass filter, and the frequency response is plotted in Fig. 10(a). Similarly, the frequency response corresponding to the following transfer functions:

$$H_2(z) = I_I(z)I_I(z^2)I_I(z^4)J_J(z^8) \qquad (25b)$$

$$H_3(z) = I_I(z^4)J_J(z^8) \qquad (25c)$$

$$H_4(z) = \left(1 - I_I(z^2)\right)I_I(z^4)I_I(z^8)J_J(z^{16}) \qquad (25d)$$

are shown in Fig. 10. The responses of the building blocks $I_I(z)$ and $J_J(z)$ are shown in Fig. 11.

All the frequency responses shown in Fig. 10 have stopband attenuation exceeding 100 dB everywhere, and

shows 6 levels of nesting $I(z)$ into itself, which demonstrates that the theoretical "box-car" low-pass filter can be approached simply by successive nestings. We do not, however, intend to recommend multiple nestings, as the complexity of the implementation grows exponentially with number of nesting levels.

(c)

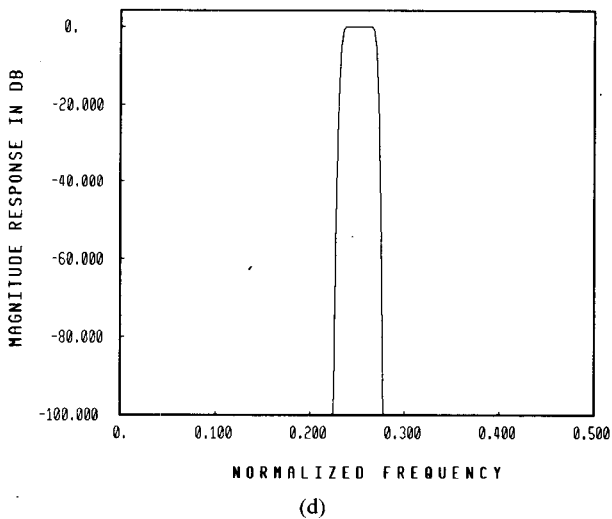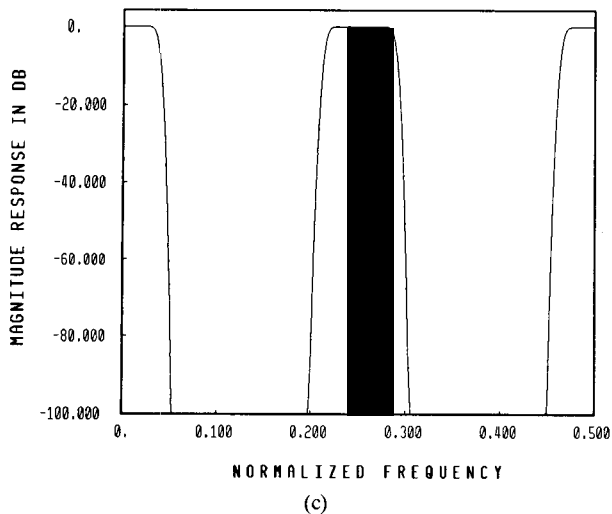Fig. 11. $I_I(e^{j\omega})$ and $J_J(e^{j\omega})$.
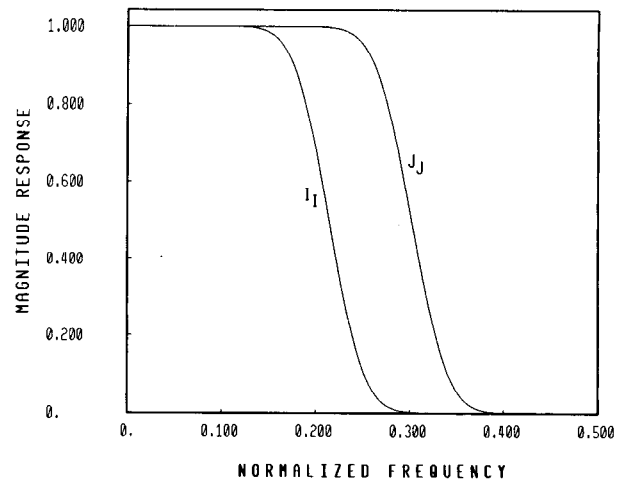

(d)

Fig. 10.   *Continued*·

the cutoff rates are very sharp. All the standard responses, such as lowpass, highpass, bandpass, and multiband can be generated by suitable combinations of the multiplierless building blocks.

General Design Procedure

Our designs are based on the four building blocks $I(z)$, $J(z)$, $K(z)$, and $L(z)$, defined by (12)–(15). The following operations are performed on these building blocks, in order to generate other filters:

1) Replace z with z^M, M = integer.
2) Replace $\cos^2(\omega/2)$ with one of the building blocks.
3) Cascade two building blocks.

These three operations can be combined and repeated in various possible ways, giving rise to a large number of possibilities. Considering, for example, the design of lowpass filters, one can achieve a passband cutoff that is an integral submultiple of the cutoff frequency of a building block. (If cutoff at an arbitrary frequency is desired, one can adjust the sampling frequency slightly, if this is practical.) A more elegant procedure for a systematic de-

sign, starting from conventional tolerance specifications, is possible. This is based on the observation that operations #1 and #2 are essentially frequency transformations, which can be used to "prewarp" the specifications. The details of a complete design procedure are currently under study.

Filter Complexity

It was mentioned earlier that the filters derived in this section are multiplierless. From the building-block transfer functions, it is clear that only additions and binary shifts are involved. In order to make a fair comparison with the complexity of a conventional implementation with multipliers, we proceed in the following way: In order to meet a given specification in the frequency domain, an equiripple filter can be designed using McClellan's algorithm [3]. Assume that this involves M_e multipliers. The precision required per multiplier coefficient of the equiripple filter is then estimated. Let this be b bits. Then, assuming that all internal signals in the equiripple implementation require b_1 bits each, we can assume each multiplier to have a complexity equivalent to $(b-1)$ binary shifts and $(b-1)$ additions of b_1-bit signals.

Thus for the equiripple approximation, we come up with two numbers: total number of binary shifts and total number of additions. For the multiplierless filters introduced in this section, these same numbers can be computed, leading to a fair comparison of complexity. (Notice that this comparison is not meaningful under situations where parallel multipliers are available, in which case, the time taken for a multiplication is not very different from addition time.)

Table I shows the number of binary shifts, additions, and delay elements involved in the implementation of various building blocks, introduced in this paper. Notice that $I_I(z^2)$ requires twice the number of delays as $I_I(z)$, but the same number of binary shifts and additions as $I_I(z)$. Similar comments hold for other derived building blocks.

TABLE I
COMPLEXITY OF BUILDING BLOCKS

Type of Building Block	# of Adds	# of Shifts	# of Delays
$I(z)$	13	4	10
$J(z)$	14	6	10
$K(z)$	11	2	10
$L(z)$	7	2	6
$I_I(z)$	135	42	100
$J_J(z)$	147	63	100
$K_K(z)$	112	21	100
$L_L(z)$	44	13	36

Example of a Complete Comparison

The multiplierless filter

$$H_5(z) = I_I(z)I_I(z^2)J_J(z^4) \qquad (26)$$

whose frequency response is shown in Fig. 12 is nearly equivalent to an equiripple filter of order 174, also shown in the same figure.

The equiripple filter has about 100-dB attenuation in the stopband, whereas the multiplierless filter has monotone attenuation exceeding 100 dB in the stopband. The complexity comparison of these two filters is shown in Table II. The equiripple filter requires about 10 times more number of shifts and about 4 times more additions than the multiplierless filter. The multiplierless filter, however, requires more memory (delay elements). The order of the multiplierless filter (number of delays) is (not surprisingly) much higher because the equiripple approximation is known to yield a lower bound on the filter order.

Note that in the above comparison, the internal signal precision for the equiripple design and that for the multiplierless design are assumed to be equal, viz., b_1 bits. This assumption presupposes that, for a given permissible roundoff noise at the filter output, the required internal signal precision is the same for both implementations. A careful analysis of roundoff noise propagation in a scaled-version of the new circuits is essential in order to justify this assumption. This is currently under study.

It is instructive to compare the performance of the monotone filters in the above examples with conventional IIR Butterworth filters. The transfer function $H_5(z)$ of (26) was found to be equivalent in performance to a 28th-order IIR Butterworth filter. The transfer function $H_2(z)$ (equation (25b)) was found to be equivalent to a 35th-order Butterworth, and so on. This demonstrates the extent to which the above design schemes can sharpen the filter response.

VI. CONCLUDING REMARKS

The new design technique mentioned in this paper is based on closed-form formulas [8] and is, therefore, fast and simple. At the same time, the number of multipliers required is very small even for stringent specifications. As mentioned earlier, a direct design of MAXFLAT FIR filters with small δ generally leads to a large dynamic range
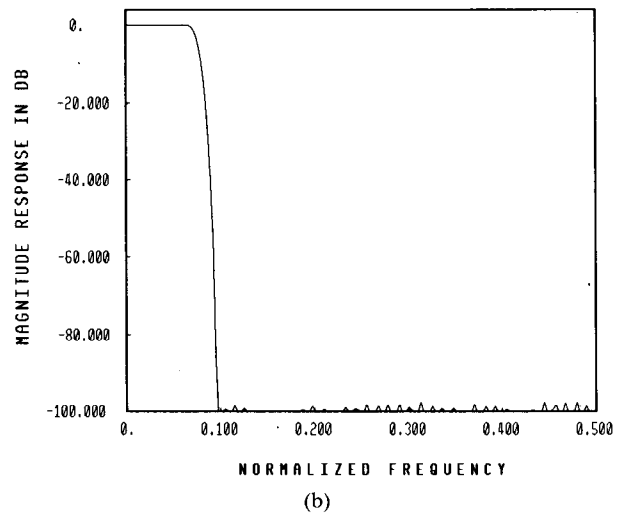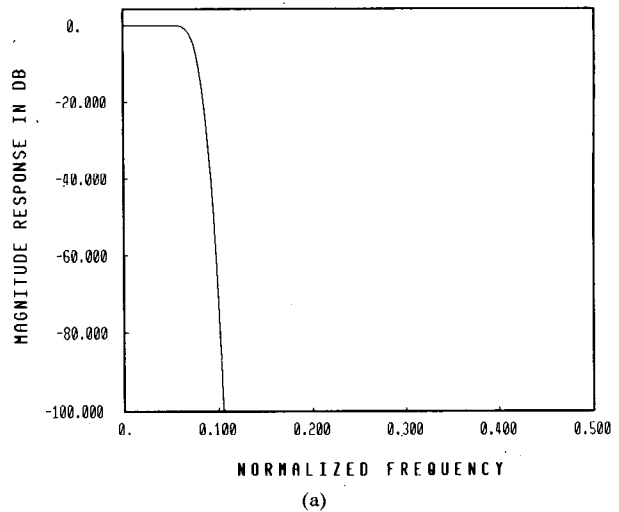

(a)

(b)

Fig. 12. (a) $|H_5(e^{j\omega})|$ in decibels. (b) Equivalent equiripple filter.

TABLE II
EXAMPLE OF COMPLEXITY COMPARISON

Type of Filter	# of Adds	# of Shifts	# of Delays
Multiplierless Filter	417	147	700
Equiripple Filter	1670	1496	174

in the impulse response coefficients. The indirect approach advanced in this paper partially overcomes this problem. Finally, it should be mentioned that the filters designed according to the new technique are not themselves maximally flat, even though they are based on maximally flat building blocks. However, if suitably designed, they do lead to a monotone frequency response, which is a requirement in many applications.

The multiplierless filters introduced in this paper require a larger number of delay units than the optimal (equiripple) designs. However, the throughput, i.e., the number of samples that can be processed per second, can be much higher than for an optimal design, as seen from Table II.

The group delay, in terms of number of samples, is higher than that for an equiripple design, but the group delay in "seconds" may turn out to be much smaller, depending upon the exact architecture of the implementation.

A crucial consideration in the actual implementation of a digital filter is the propagation of roundoff noise. (Note that if a digital filter transfer function is expected to produce about -100-dB stopband attenuation, it is not meaningful to employ a structure that generates roundoff noise at a comparable level.) The design approach developed in this paper leads to a "hierarchical structure" (because of frequency transformations), based on maximally flat building blocks. The propagation of noise in such structures can be carefully controlled by proper internal scaling and certain elegant error-cancellation schemes. A complete investigation in this direction is in order, and is currently underway.

It should finally be noticed that the idea of multiple-use of a basic filter for obtaining a sharper response has been studied by Kaiser et al. [13], leading to the concept of "amplitude change function." The methods in Section V of our paper can be looked upon as an extension of this idea.

ACKNOWLEDGMENT

REFERENCES

[1] L. R. Rabiner, "Linear program design of finite impulse response (FIR) digital filters," IEEE Trans. Audio Electroacoust., vol. AU-20, pp. 280–288, Oct. 1972.

[2] J. H. McClellan and T. W. Parks, "A unified approach to the design of optimum FIR linear phase digital filters," IEEE Trans. Circuit Theory, vol. CT-20, pp. 697–701, Nov. 1973.

[3] J. H. McClellan, T. W. Parks, and L. R. Rabiner, "FIR linear phase filter design program," in Programs for Digital Signal Processing, New York: IEEE Press, pp. 5.1-1–5.1-12, 1979.

[4] A. Antoniou, "New improved method for the design of weighted-Chebyshev, nonrecursive digital filters," IEEE Trans. Circuits Systems, vol. CAS-30, pp. 740–750, Oct. 1983.

[5] J. F. Kaiser, "Nonrecursive digital filter design using the I_0-sinh window function," in Proc. 1974 IEEE Symp. Circuits and Systems, pp. 20–23, Apr. 1974.

[6] J. J. Hill, R. Linggard, and A. G. J. Holt, "An analytical approach to the design of nonrecursive digital filters," IEEE Trans. Acoustics, Speech, Signal Processing, vol. ASSP-23, pp. 383–385, Aug. 1975.

[7] O. Herrmann, "On the approximation problem in nonrecursive digital filter design," IEEE Trans. Circuit Theory, vol. CT-18, pp. 411–413, May 1971.

[8] J. F. Kaiser, "Design subroutine (MXFLAT) for symmetric FIR lowpass digital filters with maximally flat pass and stop bands," in Programs for Digital Signal Processing, New York: IEEE Press, pp. 5.3-1–5.3-6, 1979.

[9] Y. Neuvo and S. K. Mitra, "Interpolated FIR digital filters," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-32, pp. 563–570, June 1984.

[10] A. V. Oppenheim, W. F. G. Mecklenbrauker, and R. M. Mersereau, "Variable cutoff linear phase digital filter," IEEE Trans. Circuits Syst., vol. CAS-23, pp. 199–203, April 1976.

[11] L. R. Rabiner and R. E. Crochiere, "A novel implementation for narrowband FIR digital filters," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-23, pp. 457–464, Oct. 1975.

[12] L. R. Rabiner and B. Gold, Theory and Application of Digital Signal Processing. Englewood Cliffs, NJ: Prentice Hall, 1975.

[13] J. F. Kaiser and R. W. Hamming, "Sharpening the response of a symmetric nonrecursive filter by multiple use of the same filter," IEEE Trans. Acoustics, Speech, Signal Processing, vol. ASSP-25, pp. 415–422, Oct. 1977.

✦

P. P. Vaidyanathan (S'80–M'83), for a photograph and biography please see page 224 of this issue.