

# Optimal Design of Linear-Phase FIR Digital Filters with Very Flat Passbands and Equiripple Stopbands

P. P. VAIDYANATHAN, MEMBER, IEEE

**Abstract**—A new technique is presented for the design of digital FIR filters, with a prescribed degree of flatness in the passband, and a prescribed (equiripple) attenuation in the stopband. The design is based entirely on an appropriate use of the well-known Remez-exchange algorithm for the design of weighted Chebyshev FIR filters. The extreme versatility of this algorithm is combined with certain “maximally flat” FIR filter building blocks, in order to generate a wide family of filters. The design technique directly leads to structures that have low passband sensitivity properties.

## I. INTRODUCTION

ONE OF THE important problems in digital filtering that has been considered by a number of authors in the past is the design of linear-phase FIR filters with a very flat passband response and an equiripple stopband response. Darlington [1] has considered certain general transformation techniques for handling these problems. Steiglitz [2] employs a linear programming approach for the design of such FIR filters, by imposing constraints on the derivatives of the frequency response. As pointed out by Steiglitz [2], linear programming is a very general tool for filter design and has additional flexibility over the Remez-exchange algorithm. In [3] Kaiser and Steiglitz point out the existence of numerical difficulties in the design of such FIR filters based on linear programming.

FIR filters of this class find applications in several filtering problems, as pointed out by Kaiser and Steiglitz [3]. For example, assume that we have a signal  $x(n)$  band limited to  $\omega_p$ , but immersed in noise, and that we wish to remove the noise with as little signal distortion as possible. Assume that most of the signal energy is concentrated at lower frequencies (around  $\omega = 0$ ) even though  $\omega_p$  may not be small. It is then desirable to have a lowpass filter that is very flat around  $\omega = 0$ , rather than a filter having equiripple error over the entire passband range  $(0, \omega_p)$ . As another example, consider a long-distance communication channel with a number of “repeater stations” in between. The filters in these stations are essentially in cascade. If these filters have equiripple passbands, the passband error clearly accumulates as the signal travels through the entire chan-

nel. Under such a situation, a filter with a very flat passband (at least in the region where the signal is expected to have most of its energy) is, therefore, preferred [15].

Filters with such specifications can be designed by appropriate modifications of the McClellan–Parks algorithm [4], [6], i.e., by suitable choice of the weighting function of the equiripple error. Other direct approaches that have also been described in the literature [2], [3] are based on a linear programming formulation, and Steiglitz has presented a Fortran source code in [2], which achieves this purpose. However, all of the design procedures (for these types of filters) that are known hitherto lead to *impulse response coefficients* as outputs of the design procedure. Thus the flatness or accuracy of the frequency response around  $\omega = 0$  is in general very sensitive to the quantization of the filter coefficients. Moreover, the structural simplicity offered by the fact that the frequency response is very flat around  $\omega = 0$  is not exploited in an actual hardware implementation of the resulting filter, or even during the filter design phase. Furthermore, as observed by Kaiser and Steiglitz [2], [3], the methods based on linear programming sometimes lead to numerical problems during the design phase. Rabiner has noted in an earlier work [16] that the design of even a moderately high-order FIR filter using linear programming can involve long computational time.

The purpose of this paper is to advance a new technique for the design of linear-phase FIR filters with equiripple stopbands and with a prescribed degree of flatness in the passbands. The proposed technique is based on the McClellan–Parks algorithm [4] for FIR filter design. No other optimization program or linear programming is involved. As a result, there are no numerical difficulties during the design phase. Instead of giving rise to the impulse response coefficients of the resulting filter, the design procedure directly leads to a new filter structure that has very low “passband sensitivity” (which is crucial in the implementation of filters with very flat passbands). Thus in spite of quantization of the multiplier coefficients in the implementation, the flatness of the passband remains unaffected, hence the passband continues to be highly satisfactory. Moreover, the total number of nontrivial multipliers required in the new implementation is smaller than the corresponding number for a direct design approach. In Section II the method is introduced along with

Manuscript received August 9, 1984; revised December 2, 1984. This work was supported by the National Science Foundation under Grant ECS 84-04245.

The author is with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125.

numerical examples. For the design of narrow passband filters, a number of improved methods are described in Sections III and IV based on the IFIR approach [5]. In Section V, bandpass filter designs are considered. In Section VI we discuss certain implementation considerations, including coefficient sensitivity and roundoff noise. In our developments we show how the flatness at  $\omega = 0$  can actually be used to our advantage, leading to simplifications in the actual design phase, and also during filter implementation.

II. THE PROPOSED APPROACH

Our approach in this paper is based on the McClellan-Praks method [4] for the design of weighted Chebyshev FIR filters. A good review of this algorithm, and improved versions thereof can be found in [7]. The versatility of the algorithm is based on the formulation of a weighted error function:

$$E(\omega) = W(\omega)[D(\omega) - P(\omega)] \quad (1)$$

where  $P(\omega)$  is a sum of cosines.  $D(\omega)$  is either the desired frequency response, or a functionally weighted version thereof, depending upon the "type" [4] of FIR filter sought.  $W(\omega)$  is the "weight" of the approximation error. The algorithm [6] essentially finds  $P(\omega)$ , such that  $E(\omega)$  is equiripple, thus minimizing the peak weighted error for a given filter order. Based on the software available in [6], a wide class of filters can be designed by suitably prescribing the functions  $W(\omega)$  and  $D(\omega)$ . All the techniques to be presented in this paper are based on appropriate choices of these functions.

In this section, we consider low-pass designs. We wish to design a low-pass linear-phase FIR transfer function  $G(z)$  such that  $G(e^{j\omega})$  has a tangency of  $M - 1$  at  $\omega = 0$  (i.e., the first  $M - 1$  derivatives of  $G(e^{j\omega})$  are zero at  $\omega = 0$ ). Assume that the stopband is required to be equiripple with peak error of  $\delta_2$  and that the permissible peak error at the passband edge  $\omega_p$  is  $\delta_1$  (Fig. 1). We wish to force the tangency requirement at zero frequency by extracting an appropriate building block that would represent this tangency. Notice that, if the desired magnitude of the filter response were zero rather than unit at the point of desired tangency, the tangency could have been forced trivially. For example, if one desired to design a filter  $H(z)$  with magnitude response equal to zero at  $\omega = \pi$ , and tangency  $M - 1$  at  $\omega = \pi$ , this can be achieved by designing  $H(z)$  as  $H(z) = H_1(z)H_2(z)$  where

$$H_2(z) = \left(\frac{1+z^{-1}}{2}\right)^M$$

( $H_1(z)$  is appropriately designed in order to meet other requirements). Based on this elementary observation, we wish to force the desired tangency of  $G(z)$  at  $\omega = 0$  where its magnitude is required to be unity. In order to accomplish this, we first design a low-pass filter  $H(z)$  with "complementary" specifications, as shown in Fig. 2. The desired tangency at  $\omega = \pi$  is easily forced by decomposing

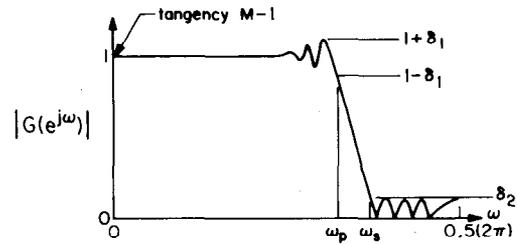


Fig. 1. The desired low-pass specifications.

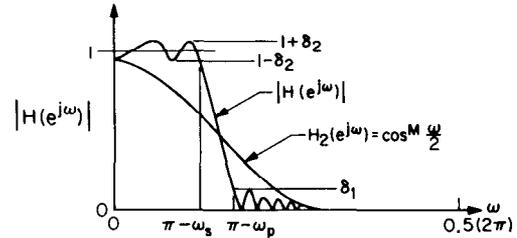


Fig. 2. The complementary specifications.

$H(z)$  as

$$H(z) = H_1(z)H_2(z) \quad (2)$$

where

$$H_2(z) = \left(\frac{1+z^{-1}}{2}\right)^M \quad (3)$$

The transfer function  $H_1(z)$  can now be designed by using the McClellan-Parks algorithm by choosing the desired response to be

$$D(\omega) = \begin{cases} |1/H_2(e^{j\omega})|, & 0 \leq \omega \leq \pi - \omega_S \\ 0, & \pi - \omega_p \leq \omega \leq \pi. \end{cases} \quad (4)$$

The weighting function  $W(\omega)$  for weighted-equiripple error should now be chosen so that the cascaded complementary filter  $H(z)$  of (2) has an equiripple passband. This can be accomplished by choosing  $W(\omega)$  to have large values close to  $\omega = 0$  and smaller values near the passband edge  $\pi - \omega_S$ . More specifically,  $W(\omega)$  should be proportional to  $|H_2(e^{j\omega})|$  for  $0 \leq \omega \leq \pi - \omega_S$ . Next, the peak stopband ripple of  $|H(e^{j\omega})|$  is required to be  $\delta_1$ . Since  $H_2(z)$  is already providing some attenuation in the stopband of  $H(z)$ , the stopband weighting in  $\pi - \omega_p \leq \omega \leq \pi$  for the approximation of  $H_1(z)$  can be decreased proportionately. If we wish to have a constant stopband weighting for simplicity, the following choice of  $W(\omega)$  is most appropriate for the design of  $H_1(z)$  in (2).

$$W(\omega) = \begin{cases} |H_2(e^{j\omega})|, & 0 \leq \omega \leq \pi - \omega_S \\ \frac{\delta_2}{\delta_1} |H_2(e^{j(\pi - \omega_p)})|, & \pi - \omega_p \leq \omega \leq \pi \end{cases} \quad (5)$$

Note that the choice of  $W(\omega)$  in (5) ensures that the cascaded filter  $H(z)$  has an equiripple passband response, and that the ratio of the peak passband error to the peak stopband error is  $\delta_2/\delta_1$ , for the complementary filter  $H(z)$ . Assuming that  $H(z)$  has been designed as above, if

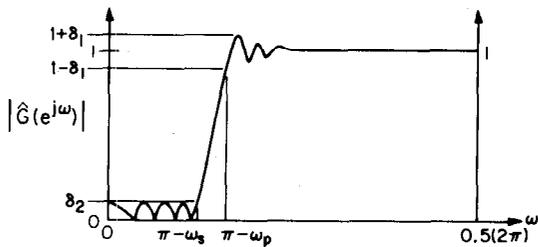


Fig. 3. The response of  $\hat{G}(z)$ .

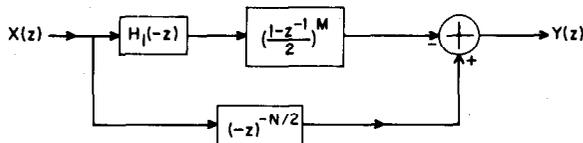


Fig. 4. Overall implementation of  $G(z)$ .

we now construct the function  $\hat{G}(z) = z^{-N/2} - H(z)$ , the resulting response is shown in Fig. 3. If each delay unit in the circuit for  $\hat{G}(z)$  is replaced with  $(-z^{-1})$ , the required transfer function is thus obtained:

$$G(z) = (-z)^{-N/2} - H(-z). \quad (6)$$

Fig. 4 shows the overall implementation of  $G(z)$ . ( $N$  is assumed even in (6).)

In the rest of the paper,  $H_1$  will be referred to as the "prewarped equiripple filter." In summary, the filter section  $H_2(z)$  takes care of the flatness of the passband of  $G(z)$ , whereas the section  $H_1(z)$  takes care of the equiripple stopband of  $G(z)$ . Thus the flatness requirement in the passband has been exploited to "extract" a portion  $H_2(-z)$  of the transfer function, which can be implemented separately without using multipliers.

*Choice of the Filter Order*

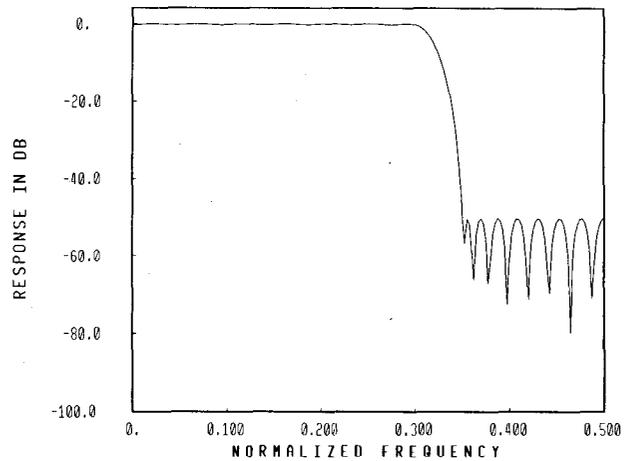
Letting  $N_1$  denote the order of  $H_1(z)$ , the overall filter order is  $N = N_1 + M$  and should be even, so that the "complementation" of (6) can be realized without the use of "fractional" delays.

Given a certain "degree of tangency" equal to  $M - 1$  at  $\omega = 0$  and given the tolerances  $\delta_1$  and  $\delta_2$ , it only remains to find  $N_1$ . It is clear that the overall order  $N$  is greater than the order  $N_e$  of the conventional optimal filter, equiripple in both bands with tolerances  $\delta_1$  and  $\delta_2$ . (This follows essentially from the alternation theorem [4].) An estimate of  $N_e$  can be found as [8]

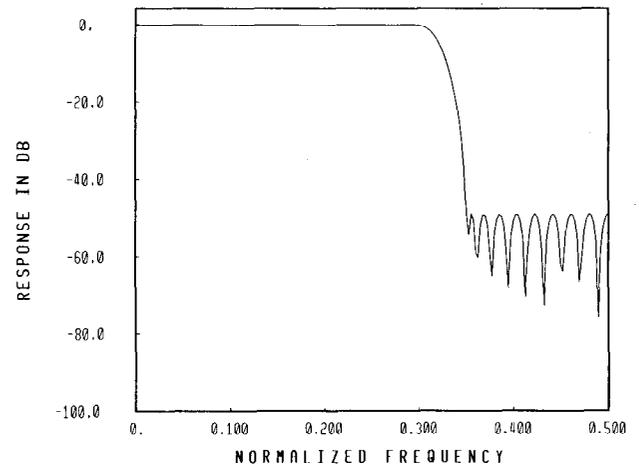
$$N_e = \frac{-10 \log_{10} \delta_1 \delta_2 - 13}{14.6 \Delta f} \quad (7)$$

where  $\Delta f = (\omega_s - \omega_p)/2\pi$ . One can therefore set  $N_1 = N_e$  and obtain a design, and then increase  $N_1$  if the specifications on tolerances are not satisfied. In almost all examples that we tried,  $N_1 = N_e$  was sufficient.

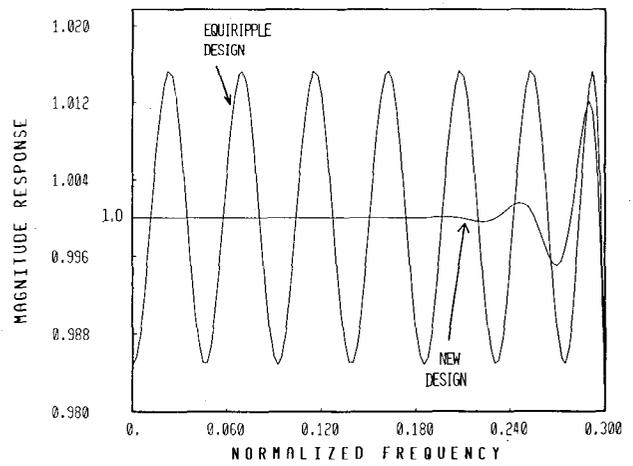
Referring to Fig. 4, the number of multiplications required per computed output sample is equal to the number of multipliers in  $H_1(z)$  because  $H_2(z)$  is multiplierless. Thus, even though the overall filter order is  $N_1 + M$ , the number of multiplications involved is only about  $(N_1/2 + 1)$ .



(a)



(b)



(c)

Fig. 5. Example 2.1. (a) Equiripple filter of Order 44. (b) The new design. (c) Passband details for Example 2.1.

*Example 2.1*

Consider the following low-pass specification:

$$\begin{aligned} \omega_p &= (0.3)2\pi, & \omega_s &= (0.35)2\pi \\ \delta_1 &= 0.016, & \delta_2 &= 0.2\delta_1 \quad (\text{corresponding to } -50 \text{ dB}) \end{aligned}$$

Order of tangency ( $M - 1$ ) at zero frequency = 15.

The estimated order of a conventional equiripple design  $G_e(z)$  is 42 from (7). An order,  $N_e = 44$  was actually

required for  $G_e(z)$ . Assuming that the required order  $N_1$  of  $H_1(z)$  is also equal to 44, and taking  $M=16$ , the filter  $G(z)$  was designed as described earlier. Fig. 5 shows the frequency response magnitudes of the new design  $|G(e^{j\omega})|$  and the equiripple design  $|G_e(e^{j\omega})|$ . Note that  $|G(e^{j\omega})|$  is extremely flat at  $\omega=0$ , as expected, and furthermore that the tolerance specifications are met. The only motivation for comparing the response with that of an equiripple design is to demonstrate that, for the same ripple size at the passband edge and in the stopband, and for the same number of multipliers ( $=23$ ), considerable flatness in the passband has been achieved by the new design. Clearly, however,  $G(z)$  whose overall order is  $44+16=60$  cannot compete with an equiripple filter of order 60 requiring 31 multipliers.

#### Example 2.2

The above example is a rather wide-band design. Consider now a narrow-band low-pass design. Let us assume that, starting from certain specifications, we have arrived at the following parameters:

$$\begin{aligned}\omega_p &= (0.1)2\pi \\ \omega_s &= (0.14)2\pi \\ N_1 &= 44 \\ M-1 &= \text{order of tangency} = 7 \\ \delta_2 &= 0.2\delta_1.\end{aligned}$$

The narrow-band nature of  $G(z)$  implies that  $H(z)$  is a wide-band low-pass filter (see Fig. 2). The desired response  $D(\omega)$  which is input to the "McClellan-Parks" program (eq. (4)) is  $D(\omega) = 1/\cos^8(\omega/2)$  in the passband of  $H(z)$ , and spans a huge dynamic range. The large dynamic range of the function  $D(\omega)$  in the passband of  $H(z)$  causes the impulse response coefficients  $h_{1,n}$  of  $H_1(z)$  to have large magnitudes. These numbers with large magnitudes are required to add up to unity because  $D(0)=1$ . Thus the passband accuracy of an actual implementation of  $H(e^{j\omega})$  (and hence the stopband sensitivity of  $G(e^{j\omega})$ ) depends crucially upon the number of bits assigned to the coefficients  $h_{1,n}$ .

It should be noticed that the dynamic range of the coefficients  $h_{1,n}$ , which is about  $139/0.45 = 308.9$  is itself large. However, this issue is not as crucial as the fact that many of the coefficients  $h_{1,n}$  themselves have large magnitudes, even though they are required to satisfy the constraint  $\sum_{n=0}^{N_1} h_{1,n} = 1$ . (If this constraint is not satisfied, then the overall transfer function  $G(z)$ , defined as (6) has a poor stopband response around  $\omega = \pi$ ). Thus for a given dynamic range of the coefficients  $h_{1,n}$ , and under the constraint  $\sum_{n=0}^{N_1} h_{1,n} = 1$ , we would like the magnitudes  $|h_{1,n}|$  to be small numbers. This can be achieved by reducing the dynamic range of  $D(\omega) = |1/H_2(e^{j\omega})|$ , in the passband of  $H(z)$ .

An additional disadvantage of the large dynamic range of  $D(\omega)$  in the passband of  $H(z)$  is that a considerable amount of signal scaling is necessary in order to avoid internal signal overflow in a fixed-point implementation. Equivalently, for a given amount of internal signal preci-

sion, this results in a large roundoff noise at the filter output.

The above sensitivity and roundoff noise problems have root in the fact that  $G(z)$  is a narrow-band transfer function. In the next two sections we indicate methods for overcoming this problem. One simple means of overcoming these is based on the "interpolation approach" to be described in the next section: essentially, a narrow-band filter can be designed by first designing a wide-band filter and then suitably manipulating it. Section IV describes a different approach for handling the above-mentioned problems, based on a new choice for  $H_2(z)$ . Section VI includes a formal study of coefficient sensitivity and roundoff noise in all the new circuits that are introduced in this paper.

### III. NARROW-BAND LOW-PASS DESIGN

Neuvo *et al.* [5] have advanced a method (the Interpolated FIR or IFIR approach) for the design of narrow-band FIR filters, in terms of a prototype wide-band filter. In order to understand this method, consider a narrow-band specification for  $G(e^{j\omega})$ , as shown in Fig. 6(a). Instead of designing  $G(z)$  directly, one can first design a filter with specifications as shown in Fig. 6(b) where the frequency axis has been "stretched" by a factor of 2. In the resulting transfer function  $G_1(z)$ , if each delay element is replaced by two units of delay, then the response is as shown in Fig. 6(c). If now the passband around  $\omega = \pi$  is suppressed without affecting the passband around  $\omega = 0$ , then the specifications of Fig. 6(a) are met. This suppression is done by cascading  $G_1(z^2)$  to a simple, multiplierless filter with transfer function  $G_2(z)$ . The transfer function  $G_2(z)$  has been referred to as an "interpolator" in [5] because  $G_2(z)$  has the effect of replacing the zero-valued samples in the impulse response of  $G_1(z^2)$  with a weighted average of adjacent samples. However, no explicit sampling rate change is involved anywhere in the implementation. The idea can be extended in an obvious manner to the case where the "stretching" is done by a larger (integral) factor  $J$ .

The primary motivation for the introduction of the IFIR approach in [5] was that the "stretched filter"  $G_1(z)$  has a wider transition band (by factor  $J$ ) and requires about  $1/J$  times as many multipliers as a direct design would require. The interpolator  $G_2(z)$  itself is multiplierless and hence, simple to implement. The overall computational complexity of the implementation is clearly much less, as compared to a direct design.

In this section we wish to use the IFIR approach for a different reason: we wish to convert the narrowband design problem ( $G(z)$ ) to a wide-band problem ( $G_1(z)$ ), so that the function  $\hat{H}(z)$ , now defined as

$$\hat{H}(z) = z^{-N/2} - G_1(-z) \quad (8)$$

has a narrower passband (and a wider transition band, equal to that of  $G_1(z)$ ) as compared to the function  $H(z)$  defined according to (6).  $\hat{H}(z)$  is then designed as the product  $\hat{H}_1(z)\hat{H}_2(z)$ , where  $\hat{H}_2(z)$  is given by the right-hand side of (3). Thus the IFIR technique has the usual

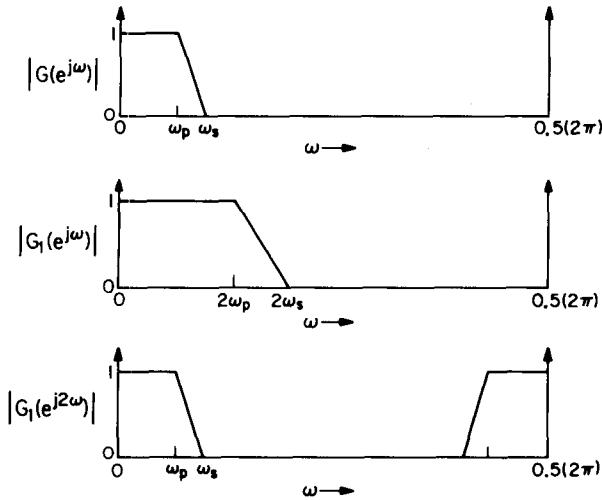


Fig. 6. The interpolated FIR approach (IFIR).

advantage of reducing the total number of multipliers (because the transition band of  $\hat{H}(z)$  is now wider) and the advantage of requiring only well-conditioned coefficients in  $\hat{H}(z)$ . In addition, the roundoff noise is considerably reduced, as demonstrated in Section VI.

#### Suppression of the Unwanted Passband at $\omega = \pi$

In the technique described in [5], the unwanted passband of  $G_1(z^2)$  is suppressed by cascading  $G_1(z^2)$  with an interpolator  $G_2(z)$  where, typically

$$G_2(z) = \left( \frac{1+z^{-1}}{2} \right)^2 \quad (9)$$

(for  $J=2$ ). The attenuating effect of  $|G_2(e^{j\omega})| = \cos^2(\omega/2)$ , in the desired passband of  $G_1(z^2)$  is compensated for, by suitably predistorting the specifications for  $G_1(z)$ .

In this paper, however, we wish to employ a new class of interpolators for  $G_2(z)$ , called the *maximally flat* interpolators, for two reasons. First, we wish the cascade  $G_1(z^2)G_2(z)$  to have the same degree of tangency as  $G_1(z^2)$  at zero frequency. This preserves the degree of flatness originally intended at  $\omega=0$ . Second, we do not wish to predistort the passband specifications of  $G_1(z)$  to compensate for the attenuation due to  $G_2(z)$ . Even though  $G_2(z)$  should be very flat around  $\omega=0$  and  $\pi$ , it need not have a narrow transition band. It can therefore be implemented in an inexpensive manner. In view of these considerations, we use the maximally flat FIR filters introduced and discussed in [9], [10].

#### Maximally Flat FIR Filters as Interpolators

An FIR low-pass frequency-response magnitude with degree of tangency  $2K-1$  at  $\omega=\pi$  and  $2L-1$  at  $\omega=0$  is given by [10]

$$|I_{K,L}(e^{j\omega})| = \cos^{2K} \left( \frac{\omega}{2} \right) \sum_{n=0}^{L-1} d(n) \sin^{2n} \left( \frac{\omega}{2} \right) \quad (10)$$

where

$$d(n) = \frac{(K-1+n)!}{(K-1)!n!} \quad (11)$$

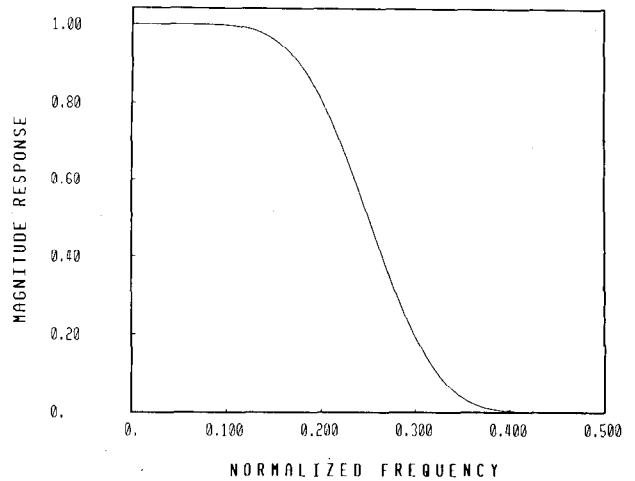
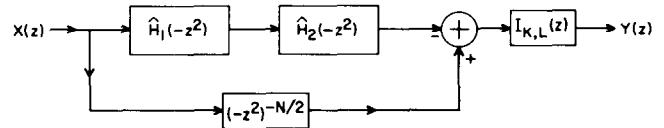
Fig. 7. The maximally flat interpolator,  $I_{4,4}(z)$ .

Fig. 8. Overall circuit for the IFIR based design.

A typical response for  $K=L=4$  is shown in Fig. 7. Notice that  $d(n)$  are integers, and for small  $K$  and  $L$ ,  $I_{K,L}(z)$  can be implemented very efficiently with few additions and no multiplications [11], [12]. For example, with  $K=L=4$ :

$$I_{4,4}(z) = C^4(z) [1 + 4S(z) + 10S^2(z) + 20S^3(z)] \quad (12)$$

where

$$C(z) = \left( \frac{1+z^{-1}}{2} \right)^2 \quad S(z) = - \left( \frac{1-z^{-1}}{2} \right)^2 \quad (13)$$

If  $G_2(z)$  is chosen to be  $I_{K,L}(z)$  with  $L=M/2$  where  $(M-1)$  is the degree of tangency of  $|G_1(e^{j\omega})|$  at  $\omega=0$ , then the order of tangency of  $G_1(z^2)G_2(z)$  is equal to  $M-1$ , at  $\omega=0$ . Furthermore, the flatness of  $|I_{K,L}(e^{j\omega})|$  at  $\omega=\pi$  ensures that the unwanted passband of  $G_1(z^2)$  is completely suppressed. The overall circuit is shown in Fig. 8.

We now proceed to demonstrate the technique with an example.

#### Example 3.1

Consider once again the narrow-band specifications of Example 2.2. Let us "stretch" the specified  $\omega_p$  and  $\omega_s$  and obtain the new specifications as follows:

$$\omega'_p = (0.2)2\pi$$

$$\omega'_s = (0.28)2\pi$$

$$N_1 = 44$$

$$M-1 = 7$$

$$\delta_2 = 0.2\delta_1.$$

The transfer function  $G_1(z)$  meeting the modified specifications can be designed as

$$G_1(z) = (-z)^{-N/2} - \hat{H}(-z) \quad (14)$$

TABLE I  
COEFFICIENTS OF  $H_1(z)$  IN EXAMPLE 2.2

| Filter length = 45           |           |           |
|------------------------------|-----------|-----------|
| ***** IMPULSE RESPONSE ***** |           |           |
| $h(0)$                       | = 0.45    | = $h(44)$ |
| $h(1)$                       | = -2.25   | = $h(43)$ |
| $h(2)$                       | = 5.92    | = $h(42)$ |
| $h(3)$                       | = -10.14  | = $h(41)$ |
| $h(4)$                       | = 10.93   | = $h(40)$ |
| $h(5)$                       | = -2.50   | = $h(39)$ |
| $h(6)$                       | = -19.78  | = $h(38)$ |
| $h(7)$                       | = 55.37   | = $h(37)$ |
| $h(8)$                       | = -95.67  | = $h(36)$ |
| $h(9)$                       | = 125.03  | = $h(35)$ |
| $h(10)$                      | = -126.60 | = $h(34)$ |
| $h(11)$                      | = 91.46   | = $h(33)$ |
| $h(12)$                      | = -26.50  | = $h(32)$ |
| $h(13)$                      | = -44.22  | = $h(31)$ |
| $h(14)$                      | = 88.04   | = $h(30)$ |
| $h(15)$                      | = -80.04  | = $h(29)$ |
| $h(16)$                      | = 19.76   | = $h(28)$ |
| $h(17)$                      | = 62.41   | = $h(27)$ |
| $h(18)$                      | = -116.68 | = $h(26)$ |
| $h(19)$                      | = 101.87  | = $h(25)$ |
| $h(20)$                      | = -16.36  | = $h(24)$ |
| $h(21)$                      | = -89.64  | = $h(23)$ |
| $h(22)$                      | = 139.35  | = $h(22)$ |

TABLE II  
COEFFICIENTS OF  $H_1(z)$  IN EXAMPLE 3.1

| Filter length = 45           |           |           |
|------------------------------|-----------|-----------|
| ***** IMPULSE RESPONSE ***** |           |           |
| $h(0)$                       | = 0.0054  | = $h(44)$ |
| $h(1)$                       | = -0.0089 | = $h(43)$ |
| $h(2)$                       | = -0.0184 | = $h(42)$ |
| $h(3)$                       | = 0.0651  | = $h(41)$ |
| $h(4)$                       | = -0.0450 | = $h(40)$ |
| $h(5)$                       | = -0.0605 | = $h(39)$ |
| $h(6)$                       | = 0.0846  | = $h(38)$ |
| $h(7)$                       | = 0.0743  | = $h(37)$ |
| $h(8)$                       | = -0.1627 | = $h(36)$ |
| $h(9)$                       | = -0.0646 | = $h(35)$ |
| $h(10)$                      | = 0.2670  | = $h(34)$ |
| $h(11)$                      | = 0.0259  | = $h(33)$ |
| $h(12)$                      | = -0.4002 | = $h(32)$ |
| $h(13)$                      | = 0.0531  | = $h(31)$ |
| $h(14)$                      | = 0.5646  | = $h(30)$ |
| $h(15)$                      | = -0.1798 | = $h(29)$ |
| $h(16)$                      | = -0.7716 | = $h(28)$ |
| $h(17)$                      | = 0.3533  | = $h(27)$ |
| $h(18)$                      | = 1.0584  | = $h(26)$ |
| $h(19)$                      | = -0.5421 | = $h(25)$ |
| $h(20)$                      | = -1.5222 | = $h(24)$ |
| $h(21)$                      | = 0.5666  | = $h(23)$ |
| $h(22)$                      | = 2.3158  | = $h(22)$ |

where

$$\hat{H}(z) = \hat{H}_1(z)\hat{H}_2(z), \quad \hat{H}_2(z) = \left(\frac{1+z^{-1}}{2}\right)^8 \quad (15)$$

and  $\hat{H}_1(z)$  is designed as described in Section II.

Finally, the desired original specifications are met by designing  $G(z)$  as  $G(z) = G_1(z^2)I_{4,4}(z)$ . Table II shows the impulse response coefficients of the filter  $\hat{H}_1(z)$ . These numbers have much smaller magnitudes than those in Table I, even though they still add up to unity (at  $\omega = 0$ ). Consequently, the passband sensitivity of  $H(z)$  and hence the stopband sensitivity of  $G(z)$  are much better than in Example 2.2. (This is demonstrated in Section VI. Moreover, it is shown in Section VI that the roundoff noise of this improved design is much less.)

Fig. 9 shows the response  $|G(e^{j\omega})|$  and also the response of a direct equiripple design of order 44, with same  $\delta_1$  and  $\delta_2$ . The new design has a much sharper transition band,

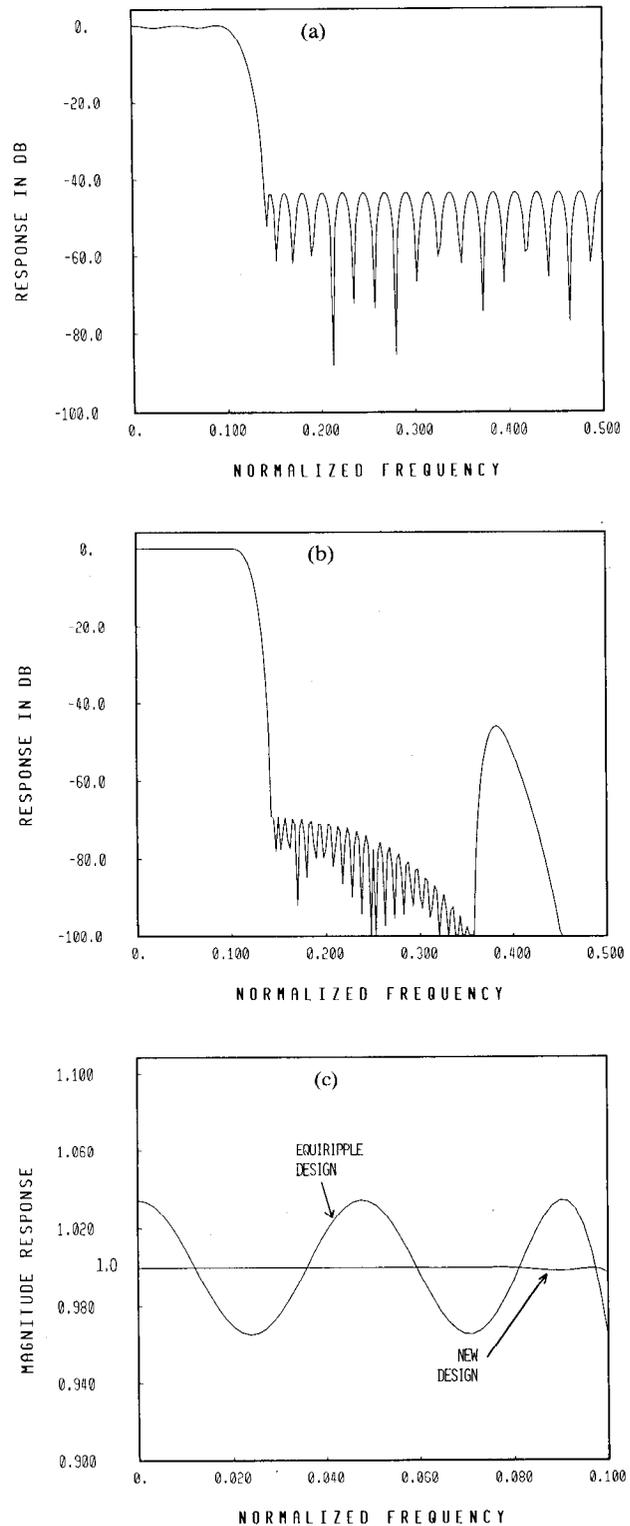


Fig. 9. Example 3.1. (a) Equiripple filter of Order 44. (b) New design, based on IFIR approach. (c) Passband details for Example 3.1.

and in general, “overmeets” the specifications everywhere. This is primarily because, in view of the IFIR approach,  $\hat{H}_1(z)$  needs to have an order of only 22 rather than the actually employed 44.

Example 3.2

This is a repetition of Example 3.1, with the difference that  $\hat{H}_1$  is now taken to be of order 22. Fig. 10 shows the

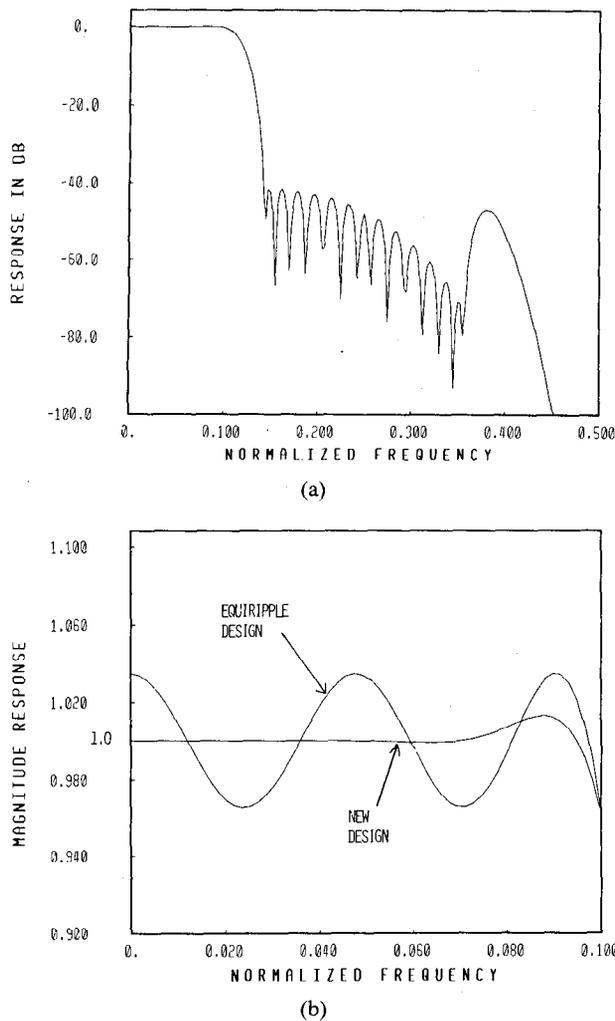


Fig. 10. Example 3.2. (a)  $|G(e^{j\omega})|$  in decibels. (b)  $|G(e^{j\omega})|$  passband details.

TABLE III  
COEFFICIENTS OF  $H_1(z)$  IN EXAMPLE 3.2

| Filter length = 23           |           |           |
|------------------------------|-----------|-----------|
| ***** IMPULSE RESPONSE ***** |           |           |
| $h(0)$                       | = -0.0033 | = $h(22)$ |
| $h(1)$                       | = 0.2258  | = $h(21)$ |
| $h(2)$                       | = -0.7527 | = $h(20)$ |
| $h(3)$                       | = 0.9295  | = $h(19)$ |
| $h(4)$                       | = -0.0792 | = $h(18)$ |
| $h(5)$                       | = -0.8750 | = $h(17)$ |
| $h(6)$                       | = 0.2411  | = $h(16)$ |
| $h(7)$                       | = 1.1625  | = $h(15)$ |
| $h(8)$                       | = -0.4529 | = $h(14)$ |
| $h(9)$                       | = -1.6563 | = $h(13)$ |
| $h(10)$                      | = 0.5340  | = $h(12)$ |
| $h(11)$                      | = 2.4616  | = $h(11)$ |

response  $|G(e^{j\omega})|$  of the overall filter, which meets all specifications and requires only 12 multipliers. Table III shows the impulse response coefficients of the filter  $\hat{H}_1(z)$ , which once again have magnitudes much less than those in Table I. Note that the IFIR approach does not lead to an equiripple stopband, but all the stopband specifications are satisfied however.

#### IV. OTHER APPROACHES FOR NARROW-BAND LOW-PASS DESIGN

Let us once again get back to the method described in Section II. For a narrow-band  $G(z)$ , the function  $H(z)$  has a wide passband, and this causes  $|H_1(e^{j\omega})|$  to span a huge dynamic range in  $0 \leq \omega \leq \pi - \omega_s$ . This is primarily because  $|H_2(e^{j\omega})|$  is very small at  $\omega = \pi - \omega_s$ . An obvious way of avoiding the above difficulty is to choose  $|H_2(e^{j\omega})|$  to be very flat not only at  $\omega = \pi$  but also at  $\omega = 0$ . The flatness at  $\omega = 0$  helps to ensure that, at the band edge ( $\omega = \pi - \omega_s$ ),  $|H_2(e^{j\omega})|$  is not "very small."

As an illustration, once again consider the specifications as in Example 2.2. Instead of choosing  $H_2(z)$  as  $[(1+z^{-1})/2]^8$  one can pick  $H_2(z)$  itself to be a maximally flat filter, viz.,  $H_2(z) = I_{4,4}(z)$ , which ensures the desired flatness at  $\omega = \pi$  of  $H(z)$ . Table IV shows the impulse response coefficients of the filter  $H_1(z)$  of order 44, designed with  $H_2(z) = I_{4,4}(z)$ . Notice that the coefficients are excellent numbers (in terms of magnitudes). Fig. 11 shows the frequency response plots of the overall filter  $G(z)$  defined as in (6). This approach to "narrow-band design" gives rise to an equiripple stopband, unlike the approach discussed in Section III.

In all discussions to follow, we wish to use the following terminology to refer to the various methods described so far.

*Method 1:* A direct, conventional, equiripple design.

*Method 2:* The approach described in Section II (Fig. 4).

*Method 3:* The narrow-band approach described in Section III (Fig. 8).

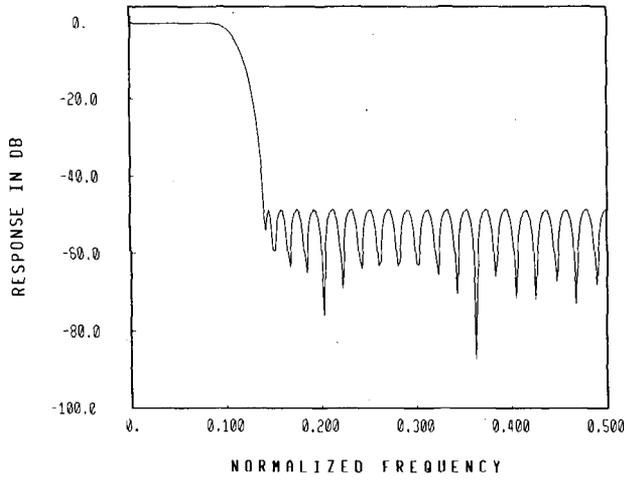
*Method 4:* The narrow-band approach described in this section (Fig. 4 with new  $H_2(z)$ ).

The approach described above (Method 4) can, in fact, be combined with the IFIR approach described in Section III (Method 3) to obtain a reduction in the order of  $H_1(z)$  by a factor of 2. In addition, the impulse response coefficients of  $H_1(z)$  are also well conditioned. This is referred to as Method 5.

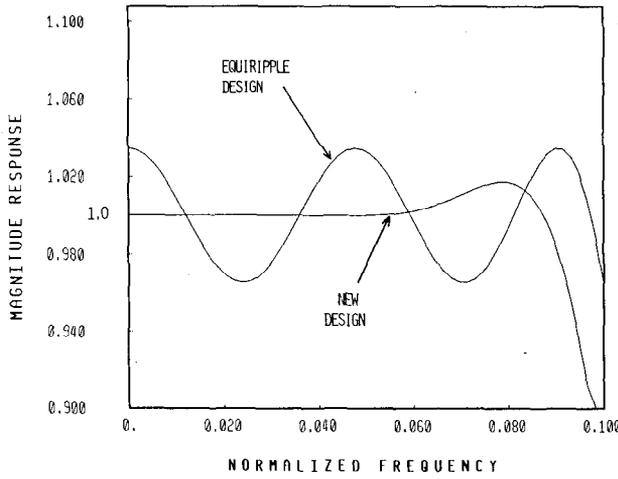
It should be noticed that the dynamic range spanned by the impulse response coefficients in Tables I-IV is as follows, for the various examples considered so far:

- (Table I) Example 2.2 139/0.45 = 308.9
- (Table II) Example 3.1 2.32/0.0054 = 429.6
- (Table III) Example 3.2 2.46/.0033 = 745.5
- (Table IV) Example of Section IV 8.667/.0034 = 2549.1.

Thus even though the methods of Sections III and IV give rise to a higher dynamic range of  $|h_{1,n}|$  (or  $|\hat{h}_{1,n}|$  as appropriate), the absolute values of these coefficients are smaller as compared to Example 2.2, as seen from Tables I-IV. As a result, for narrow-band designs, the methods of Sections III and IV give rise to filters with much lower stopband sensitivity of  $G(e^{j\omega})$  with respect to the coefficients  $h_{1,n}$  (or  $\hat{h}_{1,n}$  as they case may be), compared to the methods of Section II. In Section VI these statements are



(a)



(b)

Fig. 11. Example in Section IV. (a)  $|G(e^{j\omega})|$  in decibels. (b)  $|G(e^{j\omega})|$ , passband details.

TABLE IV  
COEFFICIENTS OF  $H_1(z)$  IN EXAMPLE OF SECTION IV

| Filter length = 45           |           |           |
|------------------------------|-----------|-----------|
| ***** IMPULSE RESPONSE ***** |           |           |
| $h(0)$                       | = 0.0164  | = $h(44)$ |
| $h(1)$                       | = -0.0376 | = $h(43)$ |
| $h(2)$                       | = -0.0034 | = $h(42)$ |
| $h(3)$                       | = 0.2570  | = $h(41)$ |
| $h(4)$                       | = -0.9305 | = $h(40)$ |
| $h(5)$                       | = 2.1960  | = $h(39)$ |
| $h(6)$                       | = -4.0476 | = $h(38)$ |
| $h(7)$                       | = 6.1742  | = $h(37)$ |
| $h(8)$                       | = -7.9400 | = $h(36)$ |
| $h(9)$                       | = 8.5442  | = $h(35)$ |
| $h(10)$                      | = -7.3526 | = $h(34)$ |
| $h(11)$                      | = 4.2860  | = $h(33)$ |
| $h(12)$                      | = -0.0662 | = $h(32)$ |
| $h(13)$                      | = -3.8764 | = $h(31)$ |
| $h(14)$                      | = 5.9077  | = $h(30)$ |
| $h(15)$                      | = -4.9716 | = $h(29)$ |
| $h(16)$                      | = 1.2943  | = $h(28)$ |
| $h(17)$                      | = 3.4078  | = $h(27)$ |
| $h(18)$                      | = -6.5313 | = $h(26)$ |
| $h(19)$                      | = 5.9121  | = $h(25)$ |
| $h(20)$                      | = -1.2576 | = $h(24)$ |
| $h(21)$                      | = -4.8122 | = $h(23)$ |
| $h(22)$                      | = 8.6666  | = $h(22)$ |

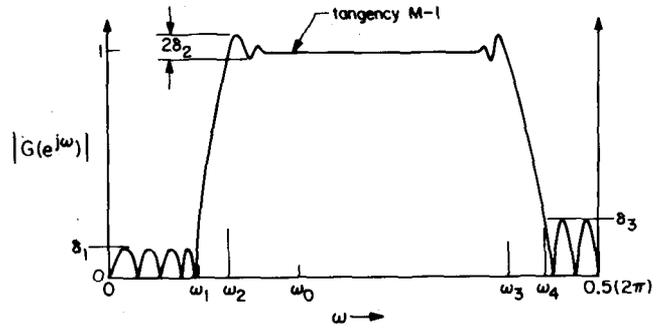


Fig. 12. Bandpass specifications.

justified by quantitative analysis, and by computer simulation of coefficient-quantization effects.

V. BANDPASS FILTERS WITH VERY FLAT PASSBAND AND EQUI RIPPLE STOPBANDS

Consider the bandpass specifications shown in Fig. 12. Let  $\omega_0$  denote the frequency in the passband where a tangency of  $M-1$  is desired. This can be achieved by designing the desired transfer function  $G(z)$  as in (6) where  $H(z) = H_1(z)H_2(z)$ , and

$$H_2(z) = (1 - 2z^{-1} \cos \omega_0 + z^{-2})^M \tag{16}$$

$H_1(z)$  is once again designed in a weighted-Chebyshev manner. For this, the specifications are as follows:

*Desired Response:*

$$D(\omega) = \begin{cases} \frac{1}{(\cos \omega - \cos \omega_0)^M}, & 0 \leq \omega \leq \omega_1 \\ & \text{and } \omega_4 \leq \omega \\ 0, & \omega_2 \leq \omega \leq \omega_3 \end{cases} \tag{17}$$

*Weighting Function:*

$$W(\omega) = \begin{cases} \frac{1}{\delta_1} \cdot |(\cos \omega - \cos \omega_0)|^M, & 0 \leq \omega \leq \omega_1 \\ \frac{1}{\delta_2} \cdot |(\cos \hat{\omega} - \cos \omega_0)|^M, & \omega_2 \leq \omega \leq \omega_3 \\ \frac{1}{\delta_3} \cdot |(\cos \omega - \cos \omega_0)|^M, & \omega_4 \leq \omega \leq \pi \end{cases} \tag{18}$$

where the constant  $\hat{\omega}$  is typically  $\omega_2$  or  $\omega_3$ . (The choice of  $\hat{\omega}$  affects the distribution of relative errors in the three bands.) Note that  $D(\omega)$  is negative in the upper passband for odd  $M$ , but  $W(\omega)$  is defined to be always positive.

*Example 5.1*

The frequency response of a filter of the above type with  $M = 2$ ,  $N_1 = \text{order of } H_1(z) = 74$ ,  $\omega_0 = (0.2)2\pi$ , and the

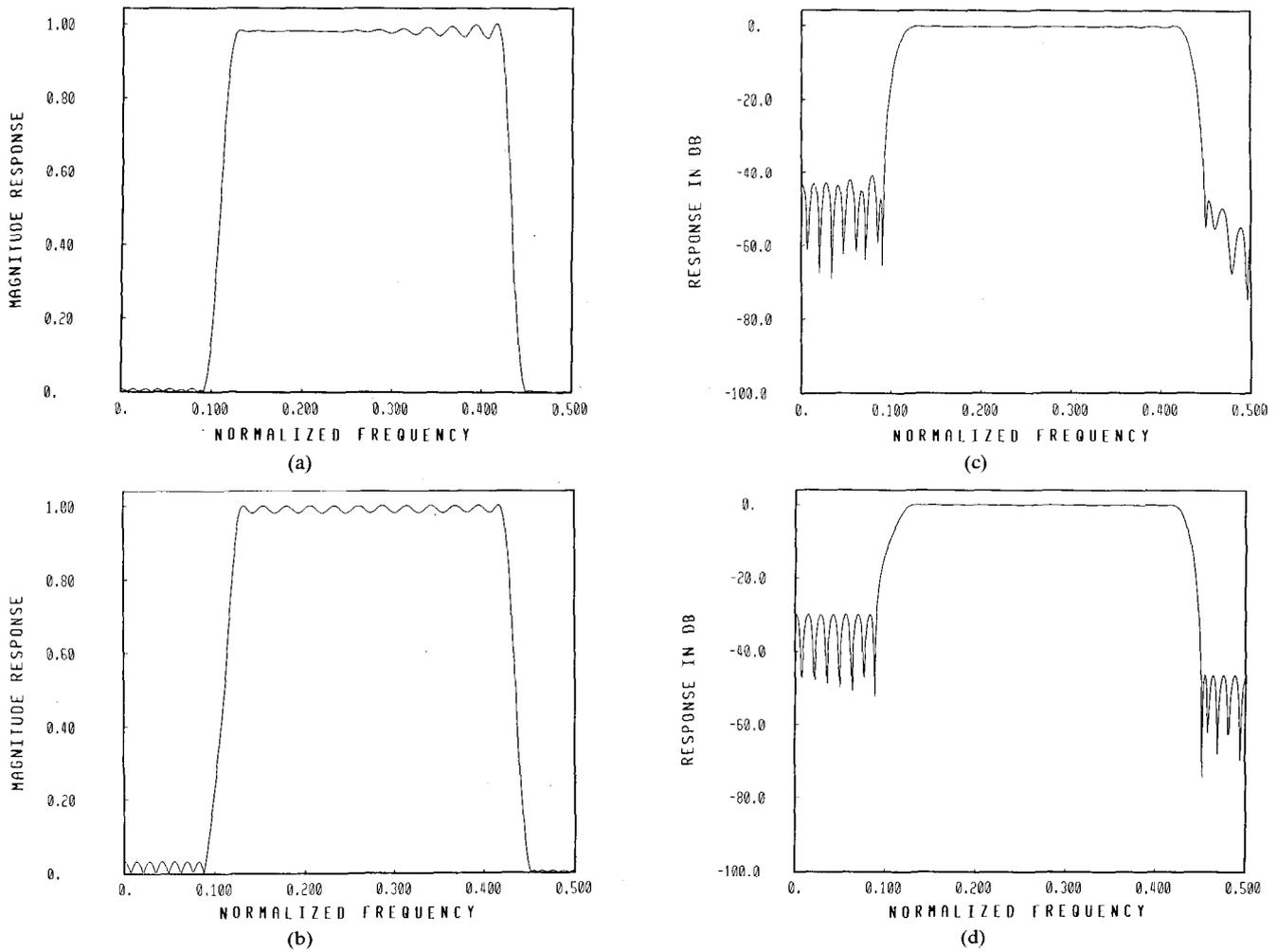


Fig. 13. Example 5.1 (a) New design. (b) Equiripple design. (c) New design. (d) Equiripple design.

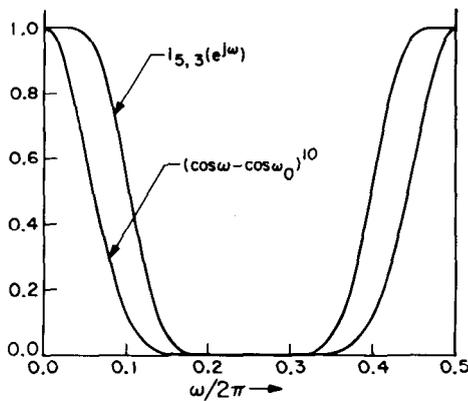


Fig. 14.  $(\cos \omega - \cos \omega_0)^{10}$  versus  $I_{5,3}(e^{j\omega})$ .

following band edges:

$$\omega_1 = (0.09)2\pi$$

$$\omega_2 = (0.128)2\pi$$

$$\omega_3 = (0.42)2\pi$$

$$\omega_4 = (0.45)2\pi$$

is shown in Fig. 13. Also shown is a conventional equirip-

ple filter of order 74. Notice the flatness around  $\omega_0 = (0.2)2\pi$ , which is evident from Fig. 13(a). The choice of the above value of  $\omega_0$  has been made arbitrarily in this example, and  $\omega_0$  actually provides an additional design flexibility.

For large values of  $M$ , the desired function  $D(\omega)$  of (17) spans a very large dynamic range in the passbands of  $H_1(z)$ , causing numerical problems mentioned earlier in Section II. Such problems can be avoided by replacing  $H_2(z)$  with a function that has the desired flatness at  $\omega_0$ , but in addition, also flattens out near  $\omega = 0$  and  $\omega = \pi$ . It is simple to construct such a function, based on the maximally flat building blocks used in (10), Section III.

**Example 5.2**

Let us assume that we have a specification as in Example 5.1, except that flatness is required around  $\omega_0 = (0.25)2\pi$  and that the degree of tangency is  $M - 1 = 9$ . The function  $(\cos \omega - \cos \omega_0)^{10}$  is not well-conditioned for the purpose. Instead, we found that  $I_{5,3}(z^2)$ , which also gives the required degree of flatness, is much more appropriate. Fig. 14 shows normalized plots of  $I_{5,3}(e^{j2\omega})$  and  $(\cos \omega - \cos \omega_0)^{10}$ . Fig. 15(a) and 15(c) show the frequency response magnitude  $|G(e^{j\omega})|$  of the final design with  $H_2(z) = (1 -$

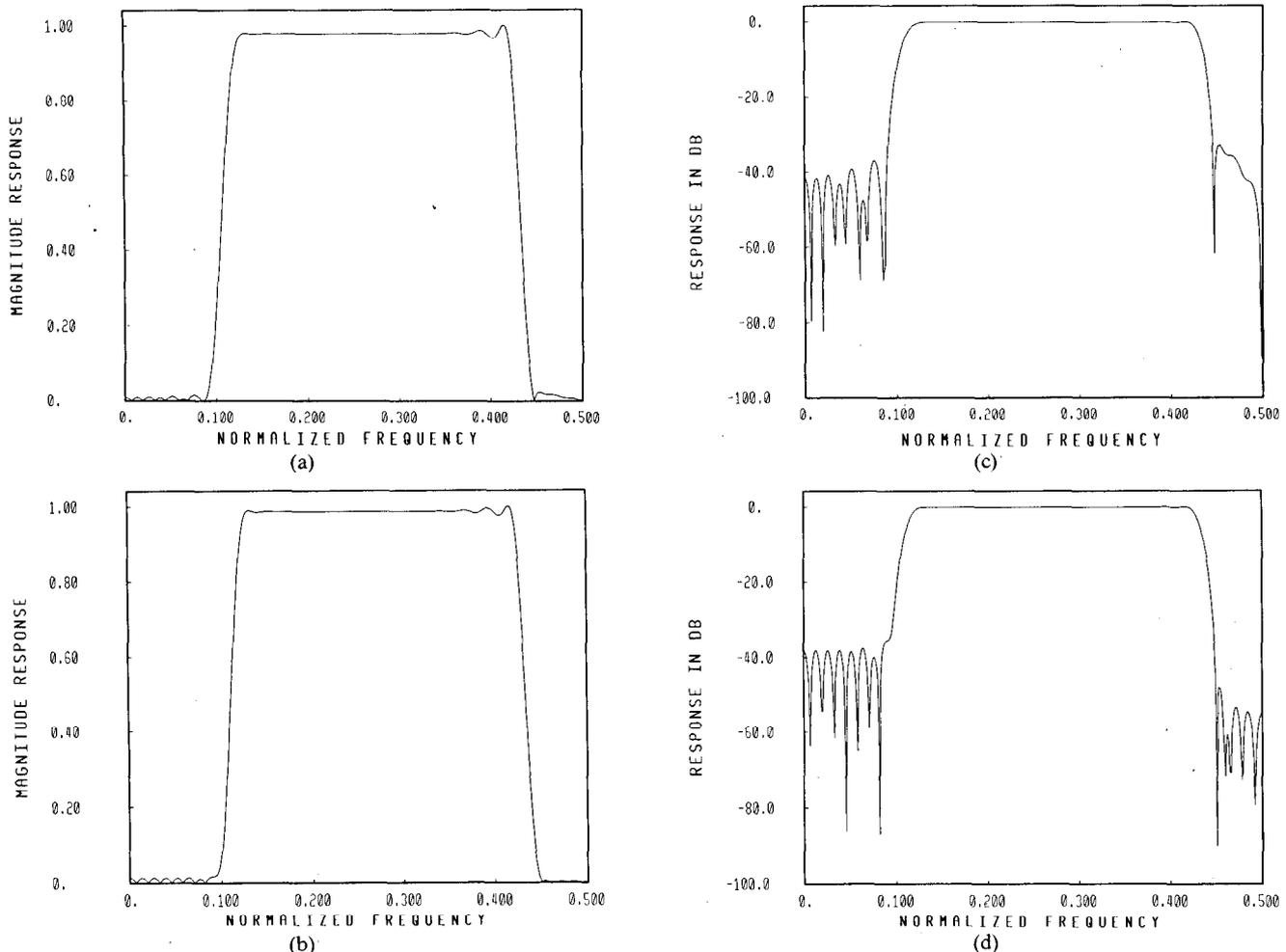


Fig. 15. Example 5.2. (a) Design with  $H_2 = (\cos \omega - \cos \omega_0)^{10}$ . (b) Design with  $H_2 = I_{5,3}(e^{j\omega})$ . (c) Design with  $H_2 = (\cos \omega - \cos \omega_0)^{10}$ . (d) Design with  $H_2 = I_{5,3}(e^{j\omega})$ .

$2z^{-1} \cos \omega_0 + z^{-2})^{10}$ . This is not an acceptable design, because the stopband response is poor. Fig. 15(b) and (d) show the response of the final design with  $H_2(z) = I_{5,3}(z^2)$ . Here, both the stopband and passband are excellent.

VI. IMPLEMENTATION CONSIDERATIONS

The methods described in this paper for FIR filter design naturally lead to structures of the form in Fig. 4 or Fig. 8 where  $H_2(z)$  and  $I_{K,L}(z)$  are multiplierless. The number of multiplications per computed output sample is, therefore,  $N_1/2 + 1$  or  $(N_1 + 1)/2$  depending on the parity of  $N_1$  where  $N_1$  is the order of  $H_1$ .

Coefficient Sensitivity

Quantization of the multipliers in  $H_1(z)$  and  $\hat{H}_1(z)$  does not seriously affect the passband of the desired transfer function  $G(z)$ . This is because the passband shaping of  $G(z)$  is essentially governed by the flatness introduced by  $H_2(z)$ . It is also obvious that the "degree of tangency" at  $\omega = 0$  is unaffected by such coefficient quantization. The stopband sensitivity of  $G(z)$  depends upon the exact method of design, as discussed in Sections II-IV.

An approximate quantitative measure of coefficient sensitivity can be obtained by proceeding along similar

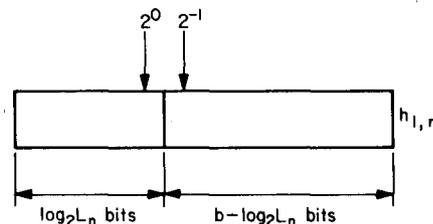


Fig. 16. Bit assignment for the coefficient  $h_{1,n}$ .

lines as in [13]. Thus for Method 2 designs (Fig. 4) we have

$$G(z) = (-z)^{-N/2} - H_1(-z)H_2(-z). \quad (19)$$

In an actual implementation, only the coefficients of  $H_1(-z)$  need to be quantized, because of the simple, multiplierless form of  $H_2(z)$ . Thus the quantization error in the frequency response is

$$\begin{aligned} E(e^{j\omega}) &= G(e^{j\omega}) - G_q(e^{j\omega}) \\ &= H_2(-e^{j\omega}) [H_{1q}(-e^{j\omega}) - H_1(-e^{j\omega})] \end{aligned} \quad (20)$$

where subscript  $q$  refers to quantized implementation. Assume  $N_1$  to be even for notational simplicity. The multi-

plier coefficients in a direct form implementation are the impulse response coefficients  $(h_{1,0}, \dots, h_{1,N_1/2})$  of  $H_1(z)$ . For a fixed-point implementation, the  $n$ th coefficient  $h_{1,n}$  is represented as shown in Fig. 16, where  $L_n$  depends upon the absolute magnitude of  $h_{1,n}$ . Thus the variance of  $|E(e^{j\omega})|$  can be expressed as

$$\sigma_2^2(\omega) = \frac{\Delta^2}{12} \left( \sin^2 M \left( \frac{\omega}{2} \right) \right) \cdot \left\{ L_{N_1/2}^2 + 4 \sum_{n=1}^{N_1/2} L_{((N_1/2)-n)}^2 \cos^2 \omega n \right\} \quad (21)$$

where  $\Delta = 2^{-b}$ , and the quantization rule is assumed to be rounding. Similarly for Method 3 designs (Fig. 8) we have

$$\sigma_3^2(\omega) = |I_{K,L}(e^{j\omega})|^2 (\sin^2 M \omega) \frac{\Delta^2}{12} \cdot \left\{ L_{N_1/2}^2 + 4 \sum_{n=1}^{N_1/2} L_{((N_1/2)-n)}^2 \cos^2 2\omega n \right\}. \quad (22)$$

The quantities  $L_n$  in (21) and (22) play an important role when judging the sensitivity of the circuit. Referring to Fig. 16,  $L_n$  is the integer part of  $h_{1,n}$  where  $h_{1,n}$  is represented by a  $b$ -bit binary number. For a given value of  $b$ , the accuracy with which the numbers  $h_{1,n}$  add up to unity (corresponding to  $H_1(e^{j0})$ ) depends upon the accuracy of the fractional parts, which are represented by  $b_1 = b - \log_2 L_n$  bits. Thus if the absolute value of  $h_{1,n}$  is large, then  $L_n$  is large and hence  $b_1$  is small. In conclusion, if the absolute values of  $h_{1,n}$  are large under the constraint that  $\sum_0^{N_1} h_{1,n} = 1$ , then the passband sensitivity of  $H(z)$  (i.e., stopband sensitivity of  $G(z)$ ) with respect to the multiplier coefficients  $h_{1,n}$  is poor.

The main conclusions that can be drawn from the above equations are as follows:

1) For both Method 2 and Method 3 designs, the sensitivity is very low around  $\omega = 0$ , in view of the high degree  $(2M-1)$  of tangency of  $\sigma^2(\omega)$  at  $\omega = 0$ . Thus the passband continues to be excellent in spite of quantization. For Method 3 designs, this statement is also true around  $\omega = \pi$ .

2) For Method 2 designs, nothing can, in general, be said about the stopband sensitivity. However, for narrow-band designs, the coefficients  $h_{1,n}$  have large magnitudes (as seen in Section II), hence the quantities  $L_n$  are typically large, leading to poor stopband sensitivities.

3) For Method 3 designs, recall that the narrowband problem is converted into a wideband problem and, therefore,  $|h_{1,n}|$  are not large (i.e.,  $L_n$  are not large). Thus the coefficient sensitivities are in general excellent in the entire stopband, even though the ultimate design is indeed narrow-band low pass.

In summary, if we use Method 2 for wide-band designs and Method 3 for narrow-band designs, then the passband sensitivities are very low and the stopband sensitivities are quite acceptable. The sensitivity properties of filters based on Methods 4 and 5 are similar to those of Method 3, i.e., the passband is excellent and stopband is quite acceptable.

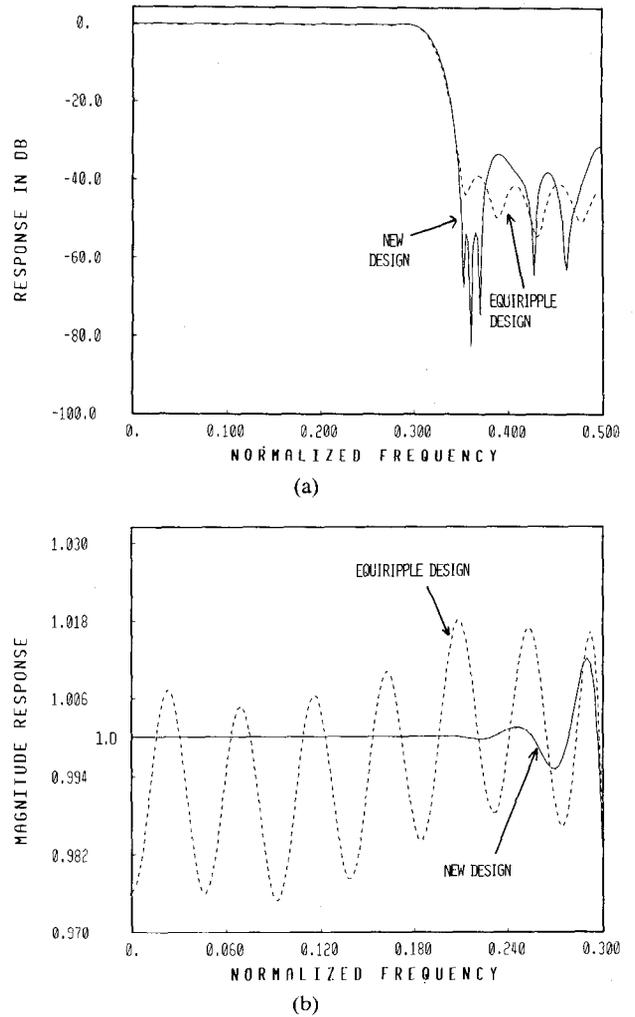


Fig. 17. Coefficient sensitivity for Example 2.1 (5 bits per multiplier). (a)  $|G(e^{j\omega})|$  in decibels. (b)  $|G(e^{j\omega})|$ , passband details.

For Method 5 designs, in addition, the stopband sensitivity around  $\omega = \pi$  is very low.

To demonstrate these conclusions, consider Example 2.1 again, which is a Method 2 design. Fig. 17 shows the frequency response of an implementation with 5 bits per multiplier. (The standard equiripple design is also shown, for 5 bits per multiplier.) The passband sensitivity is clearly excellent, and the stopband sensitivity is not any more severe than the direct form conventional equiripple designs.

As a second illustration, consider the design in Example 2.2, where the stopband sensitivity is expected to be poor (because Method 2 was used for *narrow-band design*). For 8 bits per multiplier, the passband response is still excellent, but the stopband response is simply unacceptable (Fig. 18). The same filter, designed using Method 3 with  $N_1 = 22$  (Example 3.2) has response as shown in Fig. 19, with 8 bits per multiplier coefficient. The response is clearly excellent in both the passband and the stopband.

#### Roundoff Noise in Fixed-Point Implementations

First consider Method 2 designs (Fig. 4). The noise model is shown in Fig. 20. It is assumed that  $H_1(-z)$  is implemented in direct form with one quantizer following

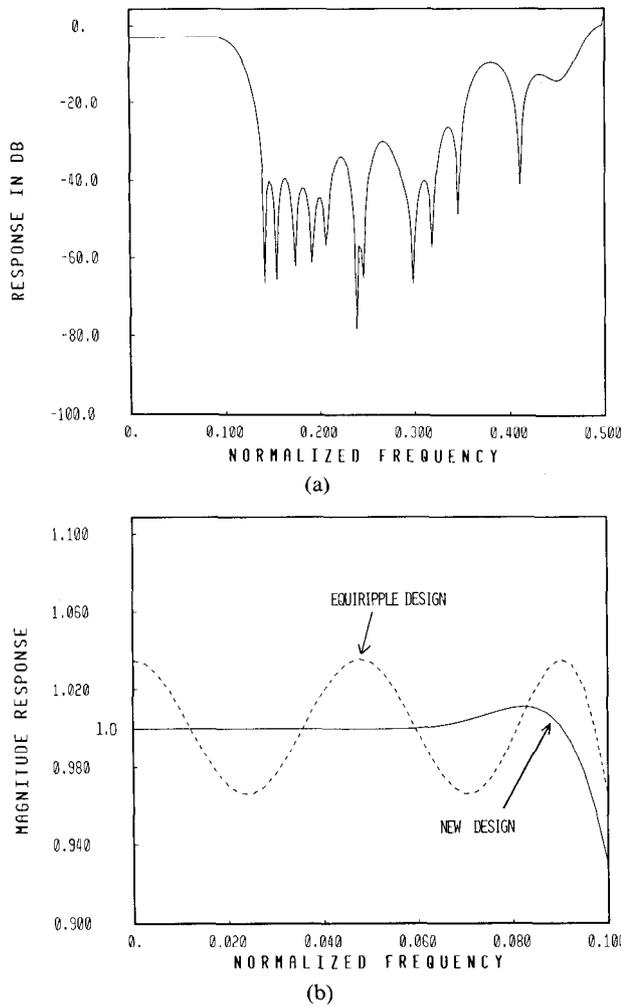


Fig. 18. Coefficient sensitivity for Example 2.2 (8 bits per multiplier). (a)  $|G(e^{j\omega})|$  in decibels. (b)  $|G(e^{j\omega})|$ , passband details.

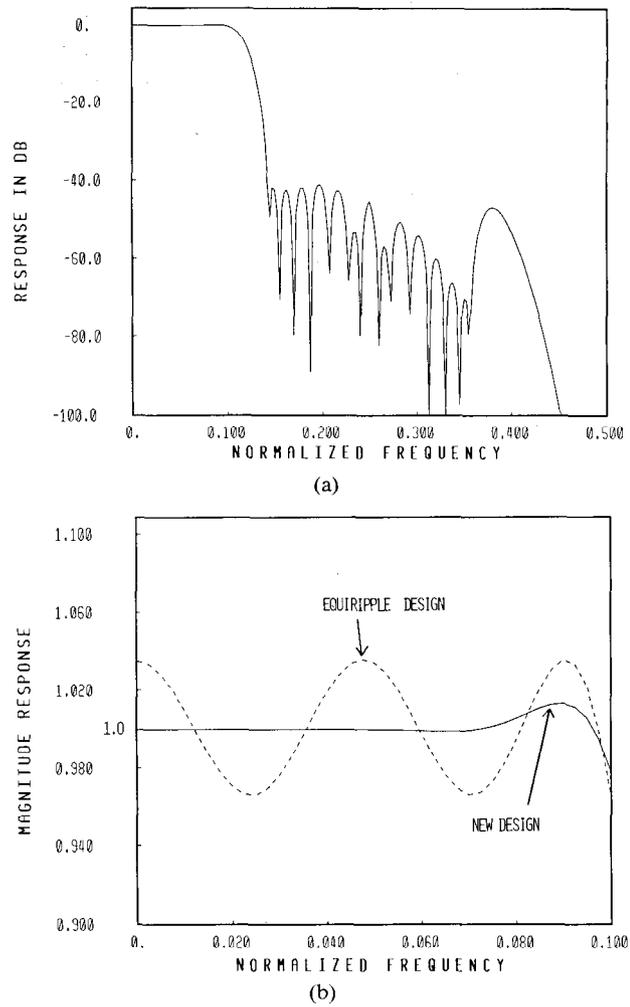


Fig. 19. Coefficient sensitivity for Example 3.2 (8 bits per multiplier). (a)  $|G(e^{j\omega})|$  in decibels. (b)  $|G(e^{j\omega})|$ , passband details.

all additions. The noise sources  $e_k(n)$  are assumed uncorrelated with each other, and each source is assumed to be white. The variance due to the roundoff noise source  $e_1(n)$  is  $\sigma_{e_1}^2 = \Delta^2/12$ , where  $\Delta = 2^{-b}$ . Assuming that each multiplier of value "half" is implemented by a right shift operation, and assuming two's complement truncation for such operations, the variance due to the sources  $e_k(n)$ ,  $k > 1$  is  $\sigma_k^2 = \Delta^2/16$ . We first derive expressions for noise variance under unscaled conditions, and then incorporate scaling, to make the expressions meaningful. Thus the total output noise variance for the unscaled implementation is

$$\sigma_0^2 = \frac{\Delta^2}{12} \left\{ \|F_1\|^2 + \frac{3}{4} \sum_{k=2}^{M+1} \|F_k\|^2 \right\} = \frac{\Delta^2}{12} A \quad (23)$$

where  $F_k(z)$  is the noise transfer function for the source  $e_k(n)$ :

$$F_k(z) = \left( \frac{1-z^{-1}}{2} \right)^{M-k+1}, \quad 1 \leq k \leq M+1 \quad (24)$$

In (23), the symbol  $\|\cdot\|$  stands for  $L_2$ -norm. Notice that the noise source  $e_1(n)$  has a noise transfer function with  $M$

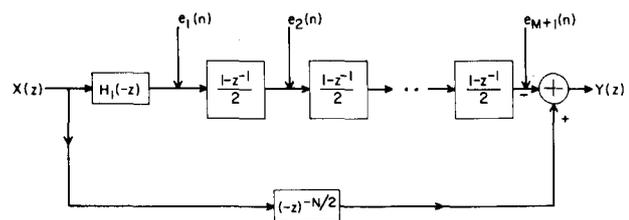


Fig. 20. Noise model for Method 2 design.

zeros at the passband frequency  $\omega = 0$ . In general, the noise transfer function for  $e_k(n)$  has  $M - k + 1$  zeros at the frequency  $\omega = 0$ . This leads us to expect a low noise-gain which is generally the case. (Note that this feature would not be available if  $H_2$  were to precede  $H_1$  in Fig. 20.)

The noise amplification factor  $A$  depends only on  $M$ . Table V shows the gain  $A$  for various  $M$ . For Method 3 designs (narrow-band filters, with IFIR approach, Fig. 8), the noise model is shown in Fig. 21. The presence of the interpolator  $I_{K,L}(z)$  helps to reduce the noise considerably. Most of the noise sources now have noise transfer functions with multiple zeros at  $\omega = 0$  and  $\omega = \pi$ . More specifically, the noise transfer function  $F_k(z)$  for the source

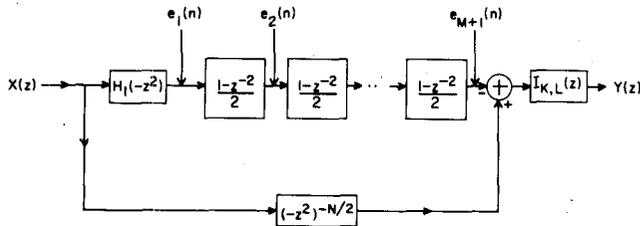


Fig. 21. Noise model for Method 3 design.

TABLE V  
PERTAINING TO NOISE GAIN FOR METHOD 2 DESIGNS

| M  | A     | A in db |
|----|-------|---------|
| 2  | 1.500 | 1.761   |
| 4  | 1.914 | 2.820   |
| 6  | 2.256 | 3.533   |
| 8  | 2.553 | 4.070   |
| 10 | 2.819 | 4.501   |
| 12 | 3.062 | 4.861   |
| 14 | 3.288 | 5.169   |
| 16 | 3.499 | 5.439   |
| 18 | 3.698 | 5.679   |
| 20 | 3.886 | 5.896   |

$e_k(n)$  is

$$F_k(z) = \left( \frac{1-z^{-2}}{2} \right)^{M-k+1} I_{K,L}(z), \quad 1 \leq k \leq M+1. \quad (25)$$

The total output noise variance  $\sigma_0^2$  can be computed in the usual manner. The noise gain  $\sigma_0^2/(\Delta^2/12)$  now depends upon  $M$ ,  $K$ , and  $L$ . Table VI shows the noise gain  $A$  (in decibels) for the circuits of Fig. 21 (design Method 3), for various values of  $K$  and  $M$  ( $L$  is assumed to be equal to  $K$ ).

In a fixed-point implementation, roundoff noise analysis is meaningful only if the internal signals are scaled such that there is no internal numerical overflow. The above expressions, and the results of Tables V and VI should be accordingly scaled, in order to obtain a reliable estimate of roundoff noise. For this, consider Fig. 20 again. Each of the building blocks  $(1-z^{-1})/2$  is already scaled in the sense of "sum-scaling" [17]. One can introduce additional scale factors to convert this into  $L_2$ -scaling, thereby reducing the noise level (but increasing overflow probability). Let us assume for simplicity that this is not done. It then remains only to scale the signal at the output of  $H_1$ . This can be done by adopting one of a number of well-known scaling policies, such as sum-scaling, peak scaling,  $L_2$ -scaling, etc. [17].

Notice that, if a narrow-band design is done based on the methods of Section II rather than Section III, then the gain of  $H_1(z)$  at its passband edge  $\pi - \omega_s$  is large, and this requires a large scaling factor for the output signal of  $H_1(z)$ . In turn, this implies a large output roundoff noise. Thus for narrow-band designs, only the methods of Sec-

TABLE VI  
THE QUANTITY  $A$  IN DECIBELS PERTAINING TO NOISE GAIN FOR METHOD 3 DESIGNS

| M  | K=L    |        |        |        |  |
|----|--------|--------|--------|--------|--|
|    | 2      | 4      | 6      | 8      |  |
| 2  | -2.524 | -2.181 | -2.017 | -1.916 |  |
| 4  | -1.660 | -1.302 | -1.121 | -1.006 |  |
| 6  | -1.079 | -0.718 | -0.528 | -0.406 |  |
| 8  | -0.641 | -0.282 | -0.088 | 0.040  |  |
| 10 | -0.289 | 0.066  | 0.262  | 0.393  |  |
| 12 | 0.006  | 0.356  | 0.553  | 0.686  |  |
| 14 | 0.260  | 0.604  | 0.801  | 0.936  |  |
| 16 | 0.484  | 0.822  | 1.018  | 1.154  |  |
| 18 | 0.683  | 1.016  | 1.211  | 1.347  |  |
| 20 | 0.864  | 1.191  | 1.385  | 1.521  |  |

TABLE VII  
ROUND OFF NOISE GAINS FOR DESIGN EXAMPLES

| Example | Noise gain in dB |             |
|---------|------------------|-------------|
|         | $L_2$ -scaling   | Sum-scaling |
| 2·1     | 13·03            | 26·31       |
| 3·1     | 11·80            | 23·86       |
| 3·2     | 12·70            | 23·96       |

tions III and IV should be used in order to keep the roundoff noise low. The actual roundoff noise gains in decibels for the various design examples presented earlier are summarized in Table VII, both for sum-scaling and  $L_2$ -scaling at the output of  $H_1$ . For  $L_2$ -scaling, the noise level is thus seen to be low, corresponding to only about 2 bits of noise gain. For sum-scaling (which leads to a completely overflow-proof implementation), the noise gains are somewhat higher, as expected.

Recall that Example 2.2 demonstrates a narrow-band design using Method 2, which is not recommended. As mentioned earlier, this is expected to have a high noise gain. Indeed, even with  $L_2$ -scaling, the noise gain in this example was found to be 58 dB (corresponding to about 10 bits!). The same example, based on Method 3 (Examples 3.1 and 3.2) has much smaller noise gain, as seen from Table VII.

All the above analyses assume that the interpolator  $I_{K,L}(z)$  does not generate noise. Indeed, for low  $K$  and  $L$ , it is possible to implement a scaled version of  $I_{K,L}(z)$  with very low roundoff noise, but the details are beyond the scope of this paper.

## VII. CONCLUDING REMARKS

The design of bandstop filters, with very flat passbands and equiripple stopbands, can be done by an obvious extension of the techniques of Sections II-IV. However, the choice of  $H_2(z)$  must be done carefully to avoid the problem of generating ill-conditioned coefficients in the filter  $H_1(z)$ .

The methods introduced in this paper do not require any optimization program other than the widely used computer program by McClellan, Parks, and Rabiner. The passbands are "flat" to a prescribed degree, but not necessarily monotone. The design technique directly leads to structures that have low sensitivity properties. A restriction of our method is that the order  $N$  of the overall filter should be even, so that (6) is physically realizable. At this point in time, it is not clear how to adapt our techniques for the design of differentiators. In contrast, the methods advanced by Steiglitz and Kaiser [3] based on linear programming can be used also for differentiator design. In addition, band-pass type of filters designed in [2] are indeed monotone in the entire passband, which is sometimes a requirement.

It should finally be noticed that the general philosophy of efficient FIR filter design, based on a cascade of the type of (2) and judicious use of the McClellan-Parks algorithm, has already been reported in [5] and [14].

#### ACKNOWLEDGMENT

The author wishes to thank Dr. J. F. Kaiser, Bell Communications Research, Inc., for his encouragement of the ideas contained in this paper. Encouraging comments by Prof. Y. Nuevo, Tampere University of Technology, are also gratefully acknowledged. The modifications of the McClellan-Parks programs for Method 2 design were incorporated by J. Trosino, first-year graduate student, Caltech, Pasadena, CA.

#### REFERENCES

- [1] S. Darlington, "Filters with Chebyshev stopbands, flat passbands, and impulse responses of finite duration," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 966-975, Dec. 1978.
- [2] K. Steiglitz, "Optimal design of FIR digital filters with monotone passband response," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 643-649, Dec. 1979.
- [3] J. F. Kaiser and K. Steiglitz, "Design of FIR filters with flatness constraints," in *IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Boston, MA, pp. 197-199, Apr. 1983.

- [4] 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.
- [5] Y. Neuvo, C.-Y. Dong, and S. K. Mitra, "Interpolated finite impulse response filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, pp. 563-570, June 1984.
- [6] 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, 1979, pp. 5.1.1-5.1.13.
- [7] A. Antoniou, "New improved method for the design of weighted-Chebyshev, nonrecursive, digital filters," *IEEE Trans. Circuits Syst.*, vol. CAS-30, pp. 740-750, Oct. 1983.
- [8] J. F. Kaiser, "Nonrecursive digital filter design using the  $I_0$ -sinh window function," *Proc. 1974 IEEE Int. Symp. on Circuits and Systems*, pp. 20-23, Apr. 1974.
- [9] O. Herrmann, "On the approximation problem in nonrecursive digital filter design," *IEEE Trans. Circuit Theory*, vol. CT-18, pp. 411-413, May 1971.
- [10] 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.
- [11] P. P. Vaidyanathan, "Fast design of efficient FIR digital filters, based on maximally flat building blocks," in *Proc. Ann. Princeton Conf. on Information Science and Systems*, Mar. 1984.
- [12] P. P. Vaidyanathan, "On maximally-flat linear-phase FIR filters," *IEEE Trans. Circuits Syst.*, Sept. 1984.
- [13] D. S. K. Chan and L. R. Rabiner, "Analysis of quantization errors in the direct form for finite impulse response digital filters," *IEEE Trans. Audio Electroacoust.*, vol. AU-21, pp. 354-366, Aug. 1973.
- [14] J. W. Adams and A. N. Willson, Jr., "A new approach to FIR digital filters with fewer multipliers and reduced sensitivity," *IEEE Trans. Circuits Syst.*, vol. CAS-30, pp. 277-283, May 1983.
- [15] R. W. Hamming, *Digital Filters*, Englewood Cliffs, NJ: Prentice Hall, 1977.
- [16] L. R. Rabiner, "Linear programming design of FIR digital filters," *IEEE Trans. Audio Electroacoust.*, vol. 20, pp. 280-288, Apr. 1974.
- [17] D. S. K. Chan and L. R. Rabiner, "Theory of roundoff noise in cascade realizations of finite impulse response digital filters," *Bell Syst. Tech. J.*, Vol. 52, pp. 329-345, Mar. 1973.

✱

P. P. Vaidyanathan (S'80-M'83), for a photograph and biography please see page 224 of the March 1985 issue of this TRANSACTIONS.