THE TRANSACTIONS OF THE IECE OF JAPAN, VOL. E **69**, NO. **5** MAY **1986**

PAPER

Iterative Design of Constrained IIR Digital Filters Requiring No Initial Values

Hisakazu KIKUCHI†, Hiromichi WATANABE†
and Takeshi YANAGISAWA††, *Members*

SUMMARY An iterative design of constrained recursive digital filters is developed. The designing scheme requires no initial values. The constraints are subjected to degrees of both numerator and denominator, transmission zeros and poles, if any, and passband and stopband shaping. The resulting filter completes a prescribed magnitude of either passband or stopband ripples. The optimality property of the filters is examined in detail with emphasis on specifications. The designing scheme involves the elliptic design as a special case. Illustrative examples are also given.

1. Introduction

As the digital integration of communication systems has been extensively developing, design techniques of constrained digital filters have received a lot of attention(1)-(11). For example, the stopband of an optimal filter for a harmonic suppression is shaped to match the magnitude response of the output stage amplifier(1). Shenoi and Agrawal(2) have given a comprehensive introduction to the similar problem in a digital switching system and digital radio. Although a combination of an equalizer and a lowpass filter is still used, the passband shaped filter eliminates the need of a separate equalizer.

Typically, it is possible to meet those requirements by designing a constrained filter which approximates an arbitrary function. A fully analytical method is not yet available, and iterative techniques had to be introduced. There are already a number of techniques for designing those filters. The Fletcher-Powell algorithm, which was employed in Ref. (3) and (4) to solve directly a nonlinear optimization problem, suffers from uncontrollability of band edges and initial guess of a starting point. The methods using linear programming(5),(6) show a good behavior of convergence, but they require rather long computational time and also restricted to the design of lower order filters. Analytical techniques can be combined with iterative procedures. One of the examples, applied to linear programming, displays good magnitude responses(7), whereas the algorithm and the optimality are not so clear and more detailed study may be needed. Another combining scheme(2) can be used for the design of a lowpass filter but the other types of filters.

On the other hand, several variations of the Remez algorithm(8)-(11) have been applied to a large class of rational approximation problems for lowpass filters. Although the method in Ref. (8) is very general, initialization is complicated.

Recently, Liang and de Figueiredo(11) have shown an interesting relationship, which shall appear later on. Their method yields a filter with the prescribed ripple ratio rather than the ripple itself. Furthermore it only allows to design lowpass filters.

The proposed technique is an extension of their method, and the detailed examination will prove the optimality property. While avoiding the problem of selecting initial values, this paper addresses the approximation problem when the magnitude characteristic is subjected to the following constraints:

1) Some of the transmission zeros and the poles of a filter may be prescribed in advance.

2) The desired passband and stopband shapes are not necessarily flat. Furthermore, each band may have multiple sizes of ripples independent of other bands.

3) The orders of a numerator and a denominator may be different. Especially, this type of filters finds its application in multirate filtering(12).

Figure 1 illustrates a typical case of possible constraints. The presented method using the Remez algorithm has the following significant features:

1) It is possible to design a band-selective recur-

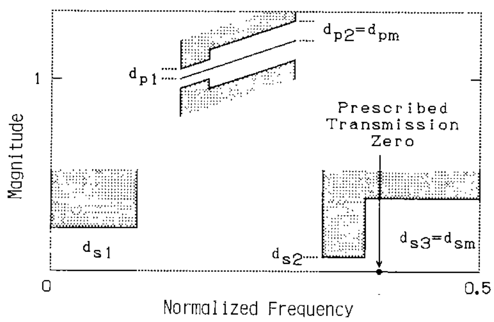


Fig. 1 Tolerance schematic of a filter.

Manuscript received December 23, 1985.
†The authors are with the Faculty of Engineering, Niigata University, Niigata-shi, 950-21 Japan.
††The author is with the Faculty of Engineering, Tokyo Institute of Technology, Tokyo, 152 Japan.

sive digital filter with constraints.

2) It completes either a given passband ripple or a given stopband ripple.

3) It needs no initial values, because the consistency of boundary conditions with a non-trivial solution is incorporated.

4) The iterative procedure converges fast, usually in a few design cycles.

The remainder of this paper is organized as follows: First, the definite statement of the problem is described and is followed by a solving algorithm. Then we discuss some points of the algorithm and the optimality property. Finally several example are given to illustrate the iterative design with emphasis on specifications.

2. Statement of the Problem

An IIR digital filter is described by a rational transfer function

$$H(z)=\frac{T_z(z)F(z)}{T_p(z)G(z)} \quad (1)$$

where two polynomials $T_z(z)$ and $T_p(z)$ correspond to constraints given in advance. Both polynomials may of course be unities. Especially a useful application for requiring $T_p(z)$ arises in the design of a Nyquist filter[13].

Given $T_z(z)$, $T_p(z)$, the degrees of $F(z)$ and $G(z)$, the passband ripples, and the stopband ripples, the problem is to find unknown polynomials $F(z)$ and $G(z)$ in such a way that the stopband ripples are minimized under preserving their ratios.

The Remez algorithm provides an optimal polynomial in the minimax sense[14],[15]. The optimality is guaranteed by the alternation theorem of a cosine polynomial. (For simplicity of notation, we write $H(e^{j2\pi f})$ as $H(f)$.)

Obviously $G(z)$ cannot have any poles outside the unit circle to ensure the stability. Hence $G(z)$ has none of the mirror image symmetries. This leads to the fact that $G(f)$ will never be a cosine polynomial of f. Thus, with the algorithm, one cannot design a denominator of an IIR digital filter.

Now, instead of Eq. (1), we consider the alternative formulation

$$\hat{H}(z)=\frac{T_z(z)F(z)}{T_p(z)\hat{G}(z)}, \quad (2)$$

where

$$\hat{G}(z)=G(z)G(z^{-1})z^{\deg G(z)}. \quad (3)$$

Since $\hat{G}(z)$ becomes a mirror image or an anti-mirror image polynomial, its magnitude response turns out to be a linear combination of cosine functions. Hence it is possible to apply the Remez algorithm to $\hat{G}(z)$.

The formulation requires extra computation of root finding, and a half of roots of $\hat{G}(z)$ inside the unit circle forms a solution for $\hat{G}(z)$. It should be noted that this computation is easily carried out by several available alogorithms. Particularly the Durand-Kerner-Aberth method[16] allows us to get simultaneously all roots of an arithmetic equation, and every root has almost the same error in computational precision. That is, the method can afford to attain a good balance of numerical precision in the roots. This fact is very desirable for the application to stringent specifications.

3. Description of the Algorithm

As a prelude to the problem formulation, several notations are introduced. We write the allowable ripples for the i-th passband and stopband as d_{pi} and d_{si}, and their maxima as d_{pm} and d_{sm}, respectively. $W(f)$ and $D(f)$ denote a weighting function and a desired function, respectively. e_p and e_s stand for deviations of the denominator or the numerator in each design cycle.

The two polynomials $F(z)$ and $\hat{G}(z)$ for a numerator and a denominator are designed on a compact subset B_s union B_p of [0, 0.5], where B_s refers to the stopband (s) and B_p does to the passband(s). Let $F(f)$ and $\hat{G}(f)$ be linear combinations of m and n cosine functions, respectively. That is, $F(f)$ is a trigonometric polynomial of m-th degree.

It is attractive to use the Remez algorithm, because its convergence rate is quadratic[14]. Since the algorithm works effectively on a set of linear expressions, linearization of the formulation is required. Dividing the design problem into two parts to deal with the numerator and the denominator separately, we can linearize the problem.

The final issue is the role assignment between the numerator and the denominator. We made the decision according to tradition: the denominator responds in the passband and the numerator does in the stopband mainly.

Thus we get the linearized formulation definitely and it is described in the following subsections.

3.1 Formulation for the Numerator

The designing equation for the numerator is formulated in such a way that the weighted error between the desired function and the numerator of interest oscillates $m+2$ times between $\pm e_s$:

$$W(f_k)[D(f_k)-F(f_k)]=(-1)^k e_s, \quad \text{for } k=0, 1, \cdots, m+1 \quad (4)$$

where $D(f)$ and $W(f)$ on the stopband(s) B_s take the forms

$$D(f)=0, \quad (5)$$

$$W(f)=(d_{sm}/d_{si})T_z(f)/T_p(f)\hat{G}(f)^{1/2}. \quad (6)$$

To get a non-trivial solution, we introduce boundary conditions of $F(z)$ within the passband(s) B_p.

$$D(f_x)=T_p(f_x)\hat{G}(f_x)^{1/2}/T_z(f_x) \tag{7}$$

$$W(f_x)=(e_s/d_{pi})/D(f_x) \tag{8}$$

To give a non-zero value to the desired function, such one must be specified on a point f_x in the passbands. The number of those conditions depends on the degrees of the numerator and the denominator as well as the band specifications. Thus it is possible to determine the unique number. The topic is, however, omitted at this point, because it needs long discussion and it will be understood naturally in the whole context of the subsection 4.3. Each f_x typically takes the midpoint of each passband. In the case that the passband includes the critical frequency, 0 or 0.5, f_x is fixed at that frequency.

3.2 Formulation for the Denominator

In the parallel way to the numerator, the formulation for the denominator implies that the weighted difference between the target function and the denominator alternates $n+2$ times between $\pm e_p$:

$$W(f_k)[D(f_k)-\hat{G}(f_k)]=(-1)^k e_p,$$
$$\text{for } k=0,1,\cdots,n+1 \tag{9}$$

where $D(f)$ and $W(f)$ on the passband(s) B_p are of the forms

$$D(f)=[F(f)T_z(f)/T_p(f)]^2, \tag{10}$$

Table 1 Expressions of a numerator and a denominator.

case	degree	m or n	F(f) or $\hat{G}$(f)
1	even	M/2	$\sum_{k=0}^{m} a(k)\cos(2\pi fk)$
2	odd	(M+1)/2-1	$\cos(\pi f)\sum_{k=0}^{m} a(k)\cos(2\pi fk)$
3	even	M/2-1	$\sin(2\pi f)\sum_{k=0}^{m} a(k)\cos(2\pi fk)$
4	odd	(M+1)/2-1	$\sin(\pi f)\sum_{k=0}^{m} a(k)\cos(2\pi fk)$
denominator		N	$\sum_{k=0}^{n} b(k)\cos(2\pi fk)$

$$W(f)=(d_{pm}/d_{pi})/D(f). \tag{11}$$

Boundary conditions of $\hat{G}(z)$ for a non-trivial solution are described on some f_x that belong to the stopband(s) B_s.

$$D(f_x)=[F(f_x)T_z(f_x)/T_p(f_x)d_{si}]^2 \tag{12}$$

$$W(f_x)=1/d_{si}^2 D(f_x) \tag{13}$$

Typically each f_x takes each midpoint of adjacent stopbands. When the stopband includes the critical frequency, 0 or 0.5, and if that frequency is assigned to one of transmission zeros in advance, then f_x is fixed near the critical frequency.

Additional attention must be paid for the practical expressions of $W(f)$, $D(f)$ and $F(f)$. They have to be modified so that they will be strictly cosine polynomials of f. Although the detailed modification is found in Ref. (15), we give its summary in Table 1, where M and N stand for degrees of $F(z)$ and $G(z)$, respectively, to avoid confusion and to complete the later discussion.

3.3 The Iterative Procedure

At the beginning of the iterative procedure, we set as $F(f)=1$, $\hat{G}(f)=1$ and $e_s=e_p=1$, whatever the specifications are.

We define a relative error e_r by

$$e_r=|e_{p\text{last}}-e_{p\text{new}}|/e_{p\text{new}}. \tag{14}$$

For a fixed small number ε, if

$$e_r<\varepsilon, \tag{15}$$

then the first design cycle is completed. Otherwise the strategy, that is to solve Eqs.(4) and (9) alternately, is repeated so as to satisfy Eq.(15). Each set of Eqs.(4) and (9), is solved by using the Remez algorithm[15].

Furthermore if d_{pm} and $e_{p\text{new}}/2$ are sufficiently close, that is, if the inequalities

$$0.99d_{pm}<e_{p\text{new}}/2<1.01d_{pm} \tag{16}$$

hold, then the algorithm finds a final solution completely. Otherwise reformulation of the problem

$$d_{si}\longleftarrow d_{si}(e_{p\text{new}}/2d_{pm})^{1/2},$$
$$\text{for every } d_{si} \text{ that belongs to } B_s \tag{17}$$

is made to meet the prescribed ripples in the passbands. This replacement must take place before the next iterative procedure starts. The deduction of Eq.(17) is deferred until the next section.

The iterative procedure is summarized below.

(Step 1) Repeat the two strategies below, until $e_r<\varepsilon$ is established.

Solve the simultaneous Eq.(4).
Solve the simultaneous Eq.(9).

(Step 2) If Eq.(16) is satisfied, stop the iteration.

Otherwise renew d_{si} with the substitution of Eq. (17) and go to step 1.

Let us turn our attention to the alternative problem. To meet the ripple requirement of the stopband instead of the passband, we need two slight modifications. One of them is concerning to the convergence criterion. To complete the prescribed stopband attenuations within 1%, we use the inequalities

$$0.99 d_{sm} < \delta < 1.01 d_{sm} \tag{18 a}$$

$$\delta = T_z(f) F(f) / T_p(f) \hat{G}(f)^{1/2} \tag{18 b}$$

instead of Eq. (16). These are evaluated at one of the edges of the stopband with the maximum deviation.

The other is of reformulation of the problem to meet the prescribed stopband ripples. In this case, we replace Eq. (17) with

$$d_{si} \longleftarrow d_{si}(d_{sm}/\delta)$$

for every d_{si} that belongs to B_s. (19)

4. Discussion

4.1 Order Estimation

The problem formulation demands to estimate the order required for a given task. Although there is not a general solution, instead a few methods are available. For example, the required order of an FIR digital filter is estimated by a Chebyshev solution(17). When the order of the numerator is much higher than that of the denominator, the higher order may fall into this estimation.

In most cases, we have observed that an elliptic filter can give a good estimate for the required order of a recursive digital filter. Antoniou has given the evaluation formula, and the detail is found in Ref. (18).

It is acceptable for our designing scheme to take into account of a large class of specifications. It can realize the multiple sizes of ripples in each passband and/or stopband, but there is no legitimate way to deal with those specifications. To exploit the estimation formula by Antoniou, it is necessary to know each unique value of ripples for both stopbands and the passbands, respectively. To this end, ingnoring the sloped shaping in the tolerance scheme, we simply take a geometric mean of the multiple moduli of the tolerant ripples in the whole passband and the whole stopband, respectively. A geometric mean is one of the most familiar averagings and it does not exceed the arithmetic counterpart. Thus this facility alleviates the hard work of order estimation.

The design method by means of the frequency transformations(18) provides neither bandpass nor bandstop filters with odd-orders. On the other hand, our designing scheme can deal with a transfer function whose order is different in its numerator and denominator. As a consequence, when the estimated value for required order of a bandpass/bandstop filter has its integer part of even, the latter method can afford to clear the given specifications with less complexity than the former one. Even if the estimated value has an odd integer part, it is often enough for the latter design to fulfill the requirements with less complexity.

4.2 Freedom from Initial Values

The boundary conditions are quite important. In a design problem for an FIR digital filter, one polynomial is only optimized on a union of disjoint closed subsets, B_p and B_s. Hence a trivial solution cannot result.

In the design problem for an IIR digital filter, on the contrary, the numerator and the denominator are optimized alternatively on B_s and B_p, respectively. Thus a proper constraint must be imposed to sustain each polynomial in a significant value on the complementary bands. Otherwise the polynomial will unfortunately go to a constant value and will result in a trivial solution. Note that these boundary conditions must have the consistency whose existence is expected when a solution has been found. This discussion leads to the explicit expressions of the boundary conditions.

Next, we consider the relationship

$$(e_p/2) e_s^2 = \text{const.} \tag{20}$$

This has been shown by Liang and de Figueiredo after tedious calculation(11).

It is easy to understand Eq. (20) intuitively. First, the optimization for $F(f)$ yields e_s and the optimization for $\hat{G}(f)$ yields e_p. Since $\hat{G}(f)$ is a squared magnitude of $G(f)$, the error reduces to approximately its half in terms of $G(f)$. Second, the design strategy, that is to solve Eqs. (4) and (9) alternatively, gives the optimum solution with maximum attainable ripples in both the passband and the stopband. Third, in the same problem, if a passband ripple greater than the former e_p is obtained, the stopband will dissipate the resulting margin. Therefore e_s^2 will be inversely proportional to $e_p/2$. This completes intuitive confirmation of Eq. (20).

By exploiting this relationship effectively, we can obtain the way to reach the prescribed ripple of either a passband or a stopband. Let d_{pm} and d_{sm} refer to a set of the maximum ripples of passbands and stopbands that completes either of the allowable ripples in both bands. Such a set will result from a design. Note that it is different from the set of d_{pm} and d_{sm} that is given in advance. To avoid confusion, a subscript 0 is added to the latter set of parameters.

Equation (20) suggests the existence of the next relationship

$$d_{pm} d_{sm}^2 \propto (e_p/2) d_{sm0}^2 \tag{21}$$

under a given set of d_{pm0} and d_{sm0}. If we desire to complete each d_{pi}, we have one degree of freedom of manipulating d_{si}. This manipulation can exclusively reflect in the design procedure through Eqs.(12) and (13), but it keeps the other designing formulations unchanged. The absolute value of d_{si} works on the design independently, whereas the every other parameter works in numerical ratios. Thus, from Eq.(21), it is found that each d_{si} should be varied to the one which is proportional to $d_{smo}(e_p/2d_{pm})^{1/2}$. Moreover, the new d_{si} should be caused to come from the last result of the design process which reflects in the next reformulation. Hence each d_{si} must be changed as follows:

$$d_{si}=d_{sio}(e_p/2d_{pm})^{1/2}. \tag{22}$$

Equation(17) is an algorithmic representation of Eq.(22).

As to fitting the stopband tolerance, note that e_s is not equal to the stopband ripple but is proportional to it. Thus Eq.(18 b) is used to assess the resulting attenuation. Of course d_{sm} reflects to e_s in a proportional fashion. This discussion leads to Eq.(19).

In addition to the above points, the Remez algorithm always converges, when it is applied to a polynomial. The main part of the algorithm constitutes of searching local extrema in conjunction with interpolation which exhibits the proper ripples. If the interpolation is possible to match a solution, the unique solution always results from the convergence. Owing to those facts, our procedure needs no initial values.

In contrast, the algorithm for designing lowpass filters, presented in Ref.(11), needs an initial value at a special frequency, 0.5. The initial guess of it is an issue with slight difficulty. Furthermore that algorithm makes the designed filter fit the prescribed ripple ratio rather than the ripple itself.

4.3 The Optimality Property

The other topic to be discussed is the optimality property. We have divided the design problem into the two separate but dependent steps, while it is originally the single problem where both the numerator and the denominator should be optimized simultaneously. Hence the resulting solution may not be optimum with respect to its magnitude response in the sense of the Chebyshev approximation.

In general[19], a trigonometric polynomial of m-th degree has at most $m+1$ local extrema in the closed interval [0, 0.5]. In the case of a lowpass or a highpass filter, since there are two band edges except for both 0 and 0.5, the polynomial can have at most $m+3$ local minima and maxima, including the band edges.

Turning our attention to the problem formulation, the Remez algorithm, implemented by Parks et al.[15], always finds $m+2$ local extrema of error function in the design of a numerator. Thus, in each case of a lowpass filter and a highpass filter, $F(f)$ has at least $m+1$ local minima and maxima on the closed subset B_s, and the additional one at $f_{xp}\in B_p$, where the boundary condition is specified. That is, $F(f)$ has at least $m+2$ local minima and maxima on the closed subset $B_s\cup f_{xp}$ which contains all of the stopband edges. Because $F(f)$ is optimized on $B_s \cup f_{xp}$, and f_{xp} is a point that belongs to the passband and is duplicate for one of the band edges for itself. Consequently, the weighted $F(f)$ exhibits $m+1$ alternations between $\pm e_s$ on the stopband B_s.

The same discussion is true for the denominator $\hat{G}(f)$. The Remez algorithm finds $n+2$ local extrema. $\hat{G}(f)$ has at least $n+1$ local minima and maxima on B_p, and an additional one at f_{xs}, where $\hat{G}(f)$ is sustained by the boundary condition. Hence the weighted $\hat{G}(f)$ shows $n+1$ alternations between $1\pm e_p$ on the passband B_p.

If the passband ripples are rather small in magnitude, thus if we can consider

$$(1\pm e_p)^{1/2}=1\pm e_p/2, \tag{23}$$

$\hat{G}(f)^{1/2}$ also oscillates, $n+1$ times, between $1\pm e_p/2$ on B_p. Therefore, $F(f)/\hat{G}(f)^{1/2}$ exhibits at least $m+1$ and $n+1$ alternations on B_s and B_p, respectively. Hence, according to the alternation theorem[20], $F(f)/\hat{G}(f)^{1/2}$ is the best unique solution in the weighted Chebyshev approximation.

The situations of bandpass filters are more complicated. We start with the discussion on a bandpass filter with an even-degree denominator. In the first stopband that is preceding to a passband, the denominator decreases monotonically from a value sustained by a boundary condition. The denominator shows $n+1$ alternations in the passband, before it increases monotonically in the last stopband. Since this behavior is natural, any more boundary conditions are not needed. While a boundary condition is also necessary for the numerator, this is same with the case of lowpass/highpass filters. Thus the optimality property is also valid in this case.

When a bandpass filter, however, has a denominator of odd-degree, the optimality property may be disappeared. If we set a boundary condition exclusively in the first stopband, the denominator will decrease monotonically after $n+1$ alternations in the passband. Note that $n+1$ is an even integer in this case. If the conditioning is so bad, it will go to zero. Since this zero results in a pole on the unit circle, the unstable filter will happen.

The difficulty can be avoided by sustaining $\hat{G}(f)$ at both f_{1s} and f_{2s}, which are the points of the preceding and following stopband, respectively. While the resulting solution is possible to have at least $n+2$ local extrema on $B_p \cup f_{1s} \cup f_{2s}$, it displays only at least n times alternations on B_p. Thus, $F(f)/\hat{G}(f)^{1/2}$ exhibits only at least $m+n+1$ alternations on $B_p \cup B_s$. Due to lack of a single alternation, the solution may no longer be optimum. There is still an exceptional possibility. As long as the Remez algorithm yields the extraripple numerator[14] in fortune, the optimum solution can be obtained.

Regarding to a bandstop filter, the particular expression of a numerator lays a restriction to its feasibility. Table 1 in the Sect. 3 shows the four cases of possible expressions for a numerator. The case 1 is acceptable to a constituent candidate of a bandstop filter, because the other three expressions produce transmission zeros at 0 and/or 0.5.

To get an optimum bandstop filter, it is enough to choose the case 1 as a numerator. When the numerator is sustained by a boundary condition on one frequency of passbands, it will show a monotonic behavior in the passbands. Both rising and falling behaviors in the last passband are acceptable to the numerator of interest. As a consequence, there are $m+1$ alternations in the stopband. The denominator has also a single boundary condition in the stopband, and will exhibit $n+1$ alternations in a whole of the passbands. Therefore we can always obtain the optimum bandstop filter with a numerator of even-degree.

Finally all of the above discussion is possible to apply to the other types of specifications. For example, it is true for a filter with prescribed transmission zeros. Since the zeros given in advance does not, of course, contribute the optimizing parameters, the complete equiripple response cannot occur, even though we intend so it is. Note that, however, it is still the optimum weighted Chebyshev approximation with respect to $F(f)/\hat{G}(f)^{1/2}$, when it is one of the cases of a lowpass/highpass filter or a bandpass filter with a denominator of even-degree and a bandstop filter.

Even if passband and stopband shapings are added to the constraints of a filter, the optimality property discussed above is completely conserved.

4.4 Limitations of the Method

All of the limitations of our design originates in three facts. One of them is the fact that it is strictly available only to optimize the limited number of parameters in our problem formulation. The approximating function has the form of $F(f)/\hat{G}(f)^{1/2}$, and the total number of adjustable parameters are limited in $m+n+2$, where m is almost a half degree of $F(z)$.

By contrast, in more general, we can consider the other formulation of $F(z)F(z^{-1})/\hat{G}(z)|_{z=e^{j2\pi f}}$, as an optimization problem in terms of a fully magnitude squared function. This formulation has $M+N+2$ parameters. While N and n are equal to each other, it should be noted that M is almost twice as many as m. Thus our solution may not be the optimum, as long as we test the optimality property in terms of a fully magnitude squared function.

However, the more general form has a more difficult problem: The optimization process has to be carried out under forcing the numerator to be nonnegative. Our scheme does not suffer from this problem and gains the computational saving.

The second reason for the limitations is due to the Remez algorithm. It allows us to find an optimum solution under prescribing the band edges. Although it always finds one of optimum polynomials in a wide sense, it is not necessarily possible to find an extraripple solution. Thus if a very particular problem is considered, and when the specifications are appropriate in terms of band edges, we always have only the opportunity to get the optimum solution.

We attribute the third reason to a fact that Eq. (20) is valid for the limited applications. If the desired filter has a rather wide transition band, not too narrow passband and stopband, and not too small passband and stopband ripples, that is right.

5. Examples

Here are several examples to illustrate the effectiveness of the iterative design described in the preceding sections. All of the examples are designed with the aids of a 16-bit personal computer, NEC PC-9801E, while ε in Eq. (15) takes 0.05.

(Example 1) Let us consider the following specifications for a seventh-order lowpass filter. Its passband edge is 0.2 and the stopband edge is 0.22 in the normalized frequency. The desired passband ripple is 0.5 dB, and the stopband attenuation is desired as much as possible. To make a comparison with the elliptic filters, any other constraints are not imposed.

The design procedure attained the stopband attenuation of 51 dB after five and a half minutes. In this case, $m=3$, $n=7$. The magnitude response is shown in Fig. 2. By inspection of the response, there are twelve alternations and it is found that the filter is optimum according to the alternation theorem.

We also designed the elliptic filter with the same specifications by the computer program of Ref. (21), however, any considerable differences could not be found. This demonstrates that the elliptic design falls into a special case of our design.

Our improved technique has another flexible advantage; The lowpass filter has a transmission zero at 0.5. In the another design technique[11], on the contrary, such one cannot be realized because of necessity of an initial

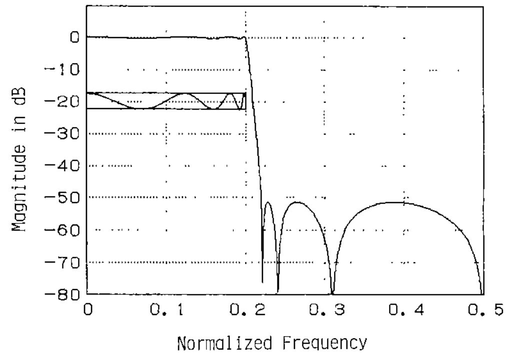


Fig. 2 Magnitude response of the lowpass filter.

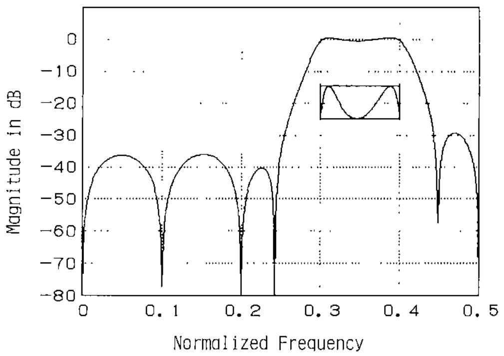


Fig. 3 Magnitude response of the bandpass filter with the prescribed transmission zero at 0.2.

value at 0.5.

(Example 2) Let us consider the bandpass filter with a prescribed transmission zero at 0.2. The passband spans the interval from 0.3 to 0.4, and the allowable ripple in the passband is fixed at 1dB. The stopband edges are 0.25 and 0.44. The attenuations are specified with 40 dB in the lower band and 30 dB in the upper band, respectively. The numerator, including a pair of the prescribed zeros, is twice the order of a denominator which is five.

After three minutes, the design procedure converged in the number of two iterations. Figure 3 shows the magnitude response, and it has only five alternations in the passband. Since the order of this bandpass filter is odd, and $m=3$, $n=5$, this is the case that the optimum one in terms of a fully magnitude response does not occur in nature.

(Example 3) Once again consider a bandpass filter with different magnitudes of ripples in one passband and one stopband. The specifications are as follows: The lower stopband is up to 0.1 with the attenuation of 45 dB. The passsband spans the interval from 0.15 to 0.25, where two different ripples of 1 dB and 2 dB are desired in the lower region up to 0.2 and in the upper region from 0.2, respectively. Similarly, two different attenuations of 65 dB and 60 dB are specified in the sub-bands from 0.3 to 0.35 and above 0.35, respectively. The latter two sub-bands constitute the upper stopband. While the order of the numerator is nine, that of the denominator is eight.

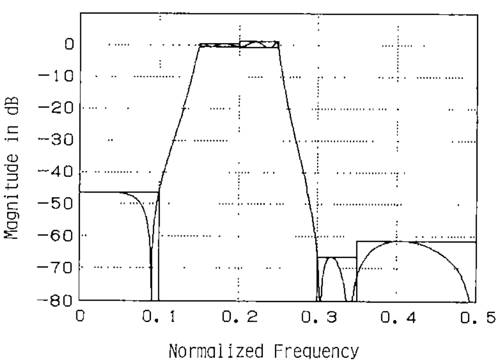


Fig. 4 Magnitude response of the bandpass filter with different ripples in one passband and one stopband.

The optimum magnitude response show in Fig. 4 was obtained after four minutes computation. The optimality is evident, because the response exhibits five and nine alternations in the stopband and the passband, respectively, and $m=4$, $n=8$.

(Example 4) A lowpass filter has been also designed to compensate the droop of 6 dB/ oct roll-off, that may arise in the preceding system, in the stopband. The orders of both the numerator and the denominator are six. The passband up to 0.17 is desired to have the 1 dB ripple. The stopband is beyond 0.2 and must be shaped to match the inverse characteristic of the given droop.

The procedure resulted in the optimum magnitude response shown in Fig. 5 after 2.5 minutes. The attenuation of the filter is higher by 2 dB than the optimum flat -band filter which is equivalent to the elliptic counterpart. This is one solution to the problem presented in Ref.(1).

(Example 5) Finally a bandstop filter has been designed to fit the stopband tolerance. The given attenuation is 60 dB over the frequencies from 0.2 to 0.35. The two sizes of the passband ripple take 0.1 dB and 1 dB over the first passband, and they are assigned to the lower region up to 0.1 and the subsequent region up to 0.15, respectively. The higher passband above 0.4 is specified with 1 dB.

The order estimation with the modified facility has

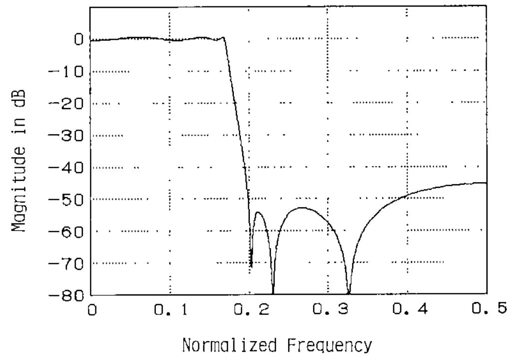


Fig. 5 Magnitude response of the lowpass filter with the stopband shaping.

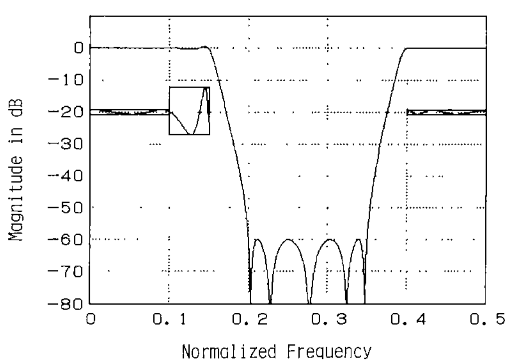


Fig. 6 Magnitude response of the bandstop filter to fit the stopband tolerance.

given a value of 10.8 as the order enough to clear the specifications. Letting the common order be ten among the numerator and the denominator, the design has produced the optimum magnitude response depicted in Fig.6. In the drawing of the response, the additional plot of the passbands are enlarged by 20 times along the vertical axis. Two values of the passband ripples have been actually minimized to 0.073 dB and 0.73 dB, respectively.

6. Concluding Remarks

An iterative procedure for the design of IIR digital filters using the Remez algorithm has been developed. It is possible to design lowpass, highpass, bandpass and bandstop filters with constraints. The constraints are subjected to degrees of numerator and denominator, transmission zeros and poles and passband and stopband shaping.

The designed filter completes a prescribed passband or stopband ripple. In certain applications, it can exhibit the optimality property with respect to a magnitude response. We have studied the individual cases about the optimality in detail and have shown definitely a class of the optimum weighted Chebyshev approximations. In addition, the method can produce the elliptic filters as a special case without constraints.

The procedure needs no initial values. The reasons are twofold. First, the consistency with a solution has led to the boundary conditions for a non-trivial one. Secondly, the Remez alogorithm always converges for a cosine polynomial.

The design procedure converges fast. This is highly dependent on the fact that the convergence rate of the Remez algorithm is quadratic.

A computer program has been implemented in a high-level interpreter language, BASIC, and partially machine code programs on a 16-bit personal computer. It will find its applications in laboratory and personal uses.

To get rid of the design limitations and to complete the automatic design, much more experimental and theoretical studies are required. One of the very interesting issues is a tradeoff between the degrees of a numerator and a denominator through an extraripple filter. Another one is to relax tight requirements for hardware implementations by making use of the non-optimum filters.

Acknowledgements

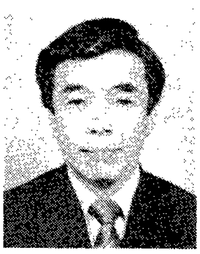
The authors are greatly indebted to Asst. Prof. A. Nishihara of Tokyo Institute of Technology, and Prof. M. Tsuboi of Niigata University, for their helpful discussion and stimulating encouragements. The authors also wish to thank Mr. Y. Tomidokoro for his contribution in machine code programming.

References

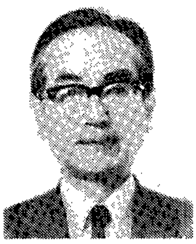
(1) N. O. Sokal: "Unsolved theoretical problem in design of optimal lowpass filter for harmonic suppression in radio-transmitter output", IEEE Trans. Circuits & Syst., **CAS-27**, p. 235 (March 1980).
(2) K. Shenoi and B. P. Agrawal: "A design algorithm for constrained equiripple digital filters", IEEE Trans. Acoust. Speech & Signal Process., **ASSP-30**, pp. 206-211 (April 1982).
(3) K. Steiglitz: "Computer-aided design of recursive digital filters", IEEE Trans. Audio, Electroacoust., **AU-18**, pp. 123-129 (June 1970).
(4) A. G. Deczky: "Synthesis of recursive digital filters using the minimum p-error criterion", IEEE Trans. Audio, Electroacoust., **AU-20**, pp. 257-263 (Oct. 1972).
(5) L. R. Rabiner, N. Y. Graham and H. D. Helms: "Linear programming design of IIR digital filters with arbitrary magnitude function", IEEE Trans. Acoust. Speech & Signal Process., **ASSP-22**, pp. 117-123 (April 1974).
(6) D. E. Dudgeon: "Recursive filter design using differential correction", IEEE Trans. Acoust. Speech & Signal Process., **ASSP-22**, pp. 443-448 (Dec. 1974).
(7) H. Dubois and H. Leich: "On the approximation problem for recursive digital filters with arbitrary attenuation curve in the pass-band and the stop-band", IEEE Trans. Acoust. Speech & Signal Process., **ASSP-23**, pp. 202-207 (April 1975).
(8) A. G. Deczky: "Equiripple and minimax (Chebyshev) approximations for recursive digital filters", IEEE Trans. Acoust. Speech & Signal Process., **ASSP-22**, pp. 98-111 (April 1974).
(9) F. Brophy and A. C. Salazar: "Synthesis of spectrum shaping digital filters of recursive design", IEEE Trans. Circuits & Syst., **CAS-22**, pp. 197-204 (March 1975).
(10) H. G. Martinez and T. W. Parks: "Design of recursive digital filters with optimum magnitude and attenuation poles on the unit circle", IEEE Trans. Acoust. Speech & Signal Process., **ASSP-26**, pp. 150-156 (April 1978).
(11) J. K. Liang and R. J. P. de Figueiredo: "An efficient iterative algorithm for designing optimal recursive digital filters", IEEE Trans. Acoust. Speech & Signal Process., **ASSP-31**, pp. 1110-1120 (Oct. 1983).
(12) M. Bellanger: "Digital Processing of Signals", John Wiley & Sons, New York (1984).
(13) T. Kida, Y. Fukuda and S. Suzuki: "Design of interpolatory digital Nyquist filters based on the minimum squared error criterion at sampling points", Proc. IEEE Int. Symp. Circuits and Syst., pp. 79-82 (June 1985).
(14) T. W. Parks and J. H. McClellan: "Chebyshev approximation for nonrecursive digital filters with linear phase", IEEE Trans. Circuit Theory, **CT-19**, pp. 189-194 (March 1972).
(15) J. H. McClellan, T. W. Parks and L. R. Rabiner: "A computer program for designing optimum FIR linear phase digital filters", IEEE Trans. Audio, Electroacoust., **AU-21**, pp. 506-526 (Dec. 1973).
(16) M. Iri: "Suuchi Keisan", Asakura Shoten, Tokyo (1981).
(17) O. Herrmann, L. R. Rabiner and D. S. K. Chan: "Practical design rules for optimum finite impulse response low-pass digital filters", Bell Syst. Tech. J., **52**, pp. 769-799 (July-Aug. 1973).
(18) A. Antoniou: "Digital Filters: Analysis and Design", McGraw-Hill, New York (1979).
(19) A. V. Oppenheim and R. W. Schafer: "Digital Signal Processing", Prentice-Hall, New Jersey (1975).
(20) E. W. Cheney: "Introduction to Approximation Theory",

McGraw-Hill, New York (1966).
(21) A. H. Gray Jr. and J. D. Markel: "A computer program for designing digital elliptic filters", IEEE Trans. Acoust. Speech & Signal Process., **ASSP-24**, pp. 529-538 (Dec. 1976).

Hisakazu Kikuchi was born in Niigata, Japan, on March 16, 1952. He received the B. E. and M. E. degrees in electronic engineering from Niigata University, Niigata in 1974 and 1976, respectively. From 1976 to 1979 he worked at the System Laboratory, FUJITSU LIMITED, Tokyo. Since 1979, he has been a research associate of the Department of Electronic Engineering, Niigata University. His research interests are in digital signal processing, CAD for digital filters and circuit theory. He is a member of IEEE.



Hiromichi Watanabe was born in Hokkaido, Japan, on August 22, 1943. He received the M. S. and Dr. Eng. degrees from Tokyo Institute of Technology, Tokyo in 1969 and 1974, respectively.
He joined Niigata University, Niigata in 1974 and has engaged in research of circuit theory and electronic circuits.



Takeshi Yanagisawa was born in Nagasaki Prefecture, Japan, on March 14, 1931. He received the B. E., M. E. and Dr. Eng. degrees in Electronic Engineering from Tokyo Institute of Technology, Tokyo in 1953, 1955 and 1958, respectively. He joined Tokyo Institute of Technology in 1958, and currently, he is a professor at the Department of Physical Electronics. Since 1980, he also serves as a director of the International Cooperation Center for Science and Technology at the same Institute. His research interests are in the fields of circuit theory, electronic circuits and CAD for electronic circuits.
Dr. Yanagisawa is a member of the Institute of Electrical Engineers of Japan.