

In terms of the new variables, (22) becomes

$$\begin{aligned}\hat{b} &= R_N H_{22}^T h_b = R_N \begin{bmatrix} H_{23}^3 \\ H_{33}^3 \end{bmatrix}^T \begin{bmatrix} h_b^3 \\ h_c^3 \end{bmatrix} \\ &= R_N [(H_{23}^3)^T h_b^3 + (H_{33}^3)^T h_c^3].\end{aligned}\quad (82)$$

Substituting (81) and (82) into (74) and (75) yields the desired recursive relations

$$\begin{aligned}\hat{b}^3 &= \hat{b} + \Delta \hat{b}^3 = \hat{b} - R_N (H_{23}^3)^T [I - H_{23}^3 R_N (H_{23}^3)^T]^{-1} \\ &\quad \cdot (h_b^3 - H_{23}^3 \hat{b})\end{aligned}\quad (83)$$

$$\hat{a}^3 = \hat{a} + \Delta \hat{a}^3 = \hat{a} - H_{12} \Delta \hat{b}^3 \quad (84)$$

and

$$\hat{a}_\Delta^3 = h_b^3 - H_{23}^3 \hat{b}^3. \quad (85)$$

Once again, the new estimates \hat{a}^3 and \hat{b}^3 are equal to the sum of the old estimates and their corresponding correction terms. After the \hat{b}^3 -vector is found, the newly added vector \hat{a}_Δ^3 is then obtained from (85).

IV. CONCLUSION

The problem under consideration has been to design a recursive digital filter which closely approximates the desired discrete impulse response in the least-square sense. In particular, sequential refinement schemes have been developed using matrix inversion lemmas for three distinct cases in which K , M ,

and N are allowed to vary separately. By the use of matrix inversion lemmas, the problem of inverting a potentially high dimensional matrix has been simplified considerably; thus without having to repeat the entire calculation, the new sets of filter coefficients can be obtained recursively based on the old estimates and the new data. In this manner, one can refine the preliminary filter design by increasing the dimension numbers successively with a minimum computational effort. The combined use of these sequential schemes will greatly enhance the performance of a recursive digital filter.

REFERENCE

- [1] J. F. Kaiser, "Digital filters," in *System Analysis by Digital Computer*, F. F. Kuo and J. F. Kaiser, Eds. New York: Wiley, 1966, pp. 218-285.
- [2] B. Gold and C. M. Radar, *Digital Processing of Signals*. New York: McGraw-Hill, 1969.
- [3] J. L. Shanks, "Recursion filters for digital processing," *Geophysics*, vol. 32, pp. 33-51, Feb. 1967.
- [4] A. J. Monroe, *Digital Processes for Sampled Data Systems*. New York: Wiley, 1962.
- [5] N. Morrison, *Introduction to Sequential Smoothing and Prediction*. New York: McGraw-Hill, 1969.
- [6] C. S. Burrus and T. W. Parks, "Time domain design of recursive digital filters," *IEEE Trans. Audio Electroacoust.* (Special Issue on Digital Filtering), vol. AU-18, pp. 137-141, June 1970.
- [7] R. Penrose, "On the generalized inverse of a matrix," *Proc. Cambridge Phil. Soc.*, vol. 51, pp. 406-413, 1955.
- [8] Y. C. Ho, "Method of least squares and optimal filtering theory," Rand Memo. RM-3329-PR, 1962.
- [9] R. K. Lee, *Optimal Estimation Identification and Control*. Cambridge, Mass.: M.I.T. Press, 1964.
- [10] A. P. Sage, *Optimum Systems Control*. Englewood Cliffs, N.J.: Prentice-Hall, 1970.
- [11] S. A. Hovanessian and L. A. Pipes, *Digital Computer Methods in Engineering*. New York: McGraw-Hill, 1969.

The Predictability of Certain Optimum Finite-Impulse-Response Digital Filters

LAWRENCE R. RABINER AND OTTO HERRMANN

Abstract—Some of the properties of optimal solutions to the finite-impulse-response low-pass filter design problem are discussed. These solutions are optimum in the sense of discrete Chebyshev approximation over a union of closed compact sets, i.e., the error of approximation exhibits at least $(N + 3)/2$ alternations (of equal amplitude) over the frequency ranges of interest, where N is the duration of the filter impulse response in samples. It has been shown that, in certain special cases, the solution can exhibit $(N + 5)/2$ alternations of equal amplitude. These solutions have been called extraripple filters because of

the extra alternation that is present. How these extraripple solutions can, within bounds, be scaled to yield additional solutions, which are still optimal over new frequency ranges, is shown. Thus an infinite number of optimal low-pass filters may be obtained directly from a finite number of extraripple solutions. An interpretation of the various types of optimal filters, in terms of locations of the zeros of the z -transform polynomial, is also given.

INTRODUCTION

A GOOD DEAL of attention has been focused recently on the design of optimal finite-impulse-response (FIR) low-pass digital filters with linear phase [1]-[7]. In particular, the problem of designing an optimal approximation (in the Chebyshev sense) to a low-pass filter with specified passband

Manuscript received August 18, 1972; revised September 25, 1972.
L. R. Rabiner is with Bell Telephone Laboratories, Inc., Murray Hill, N.J.

O. Herrmann is with Bell Telephone Laboratories, Inc., Murray Hill, N.J. He is on leave from the University of Erlangen, Nuernberg, Germany.

and stopband cutoff frequencies and a specified ratio of passband-to-stopband ripple for a fixed-order filter has been completely solved. Although computational techniques for designing such filters have been developed and shown to be exceedingly fast [8], several theoretical issues concerning the solutions still remain. For example, depending on the passband cutoff frequency, the amplitude response of the optimal solution may exhibit either $(N+3)/2$ or $(N+5)/2$ equal amplitude ripples, where N is the duration of the filter's impulse response. In some cases there is one ripple which is smaller than all the other ripples. It is the purpose of this paper to explain these and other theoretical issues. For a further discussion of the general optimal filter design problem the reader is referred to [1], [4], and [5].

CHARACTERISTICS OF OPTIMAL LOW-PASS FILTERS

Let $\{h(n), n = -(N-1)/2, \dots, (N-1)/2\}$ be the impulse response of the digital filter. (N is assumed odd throughout this paper.) The impulse response satisfies the symmetry condition

$$h(n) = h(-n), \quad 0 \leq n \leq \frac{N-1}{2} \quad (1)$$

to give the desired linear phase [1]. The frequency response of the filter can be written as

$$H(e^{j2\pi f}) = h(0) + \sum_{n=1}^{\frac{N-1}{2}} 2h(n) \cos(2\pi fn). \quad (2)$$

The response of the ideal low-pass filter $D(e^{j2\pi f})$ can be written as

$$D(e^{j2\pi f}) = \begin{cases} 1, & 0 \leq f \leq F_p \\ 0, & F_s \leq f \leq 0.5 \end{cases} \quad (3)$$

where F_p and F_s are the normalized passband and stopband cutoff frequencies, i.e., $0 \leq F_p, F_s \leq 0.5$.

Fig. 1 shows plots of both the frequency response ($H(e^{j2\pi f})$) and the approximation error ($H(e^{j2\pi f}) - D(e^{j2\pi f})$) of a typical equiripple approximation of a low-pass filter. The maximum errors of approximation in the passband and stopband are δ_1 and δ_2 , respectively. The normalized width of the transition band is $\Delta f = F_s - F_p$.

The criterion for optimal amplitude characteristics is that for specified F_p and F_s , the *weighted* approximation error can have either $(N+3)/2$ or $(N+5)/2$ extrema, where the weighted error is defined as

$$E(e^{j2\pi f}) = W(e^{j2\pi f}) [D(e^{j2\pi f}) - H(e^{j2\pi f})] \quad (4)$$

where $W(e^{j2\pi f})$ is a weighting function which allows the designer to specify the relative magnitude of the error in the passband and stopband and is defined by

$$W(e^{j2\pi f}) = \begin{cases} \frac{1}{K} = \frac{\delta_2}{\delta_1}, & 0 \leq f \leq F_p \\ 1, & F_s \leq f \leq 0.5. \end{cases} \quad (5)$$

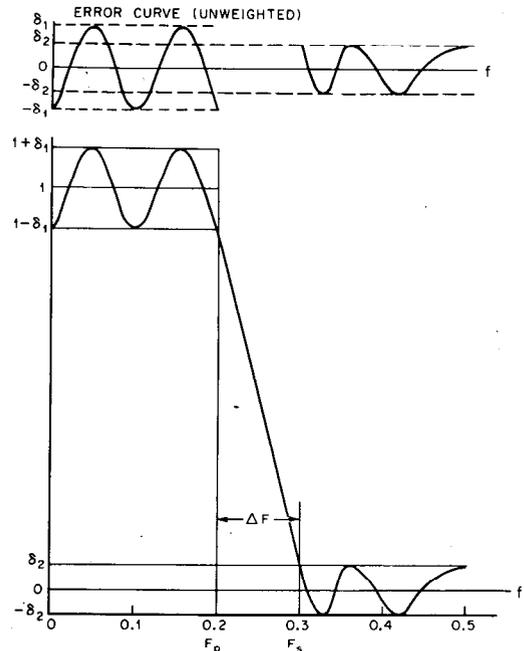


Fig. 1. Frequency response and error curve for an equiripple low-pass filter.

Thus for the optimum filter, the extrema of the weighted error curve must alternate in sign, and be equal in magnitude, i.e.,

$$E(e^{j2\pi F_i}) = -E(e^{j2\pi F_{i+1}}), \quad i = 1, 2, \dots, (N+1)/2 \quad (6)$$

where $F_i, i = 1, \dots, (N+3)/2$, are the frequencies at which the error extrema occur.

Filters which have $(N+5)/2$ extrema have been called extraripple filters [5] and have been shown to be minima along the curve of transition width (ΔF) versus passband cutoff (F_p) for fixed values of δ_1 and δ_2 [7]. Since these extraripple filters have in some sense an extra degree of freedom, it seems that it should be possible to scale these filters (in frequency) and obtain new filters (with different F_p and F_s) which still satisfy the optimality criterion discussed above. In the remainder of this paper we discuss such a scaling procedure and illustrate its application with examples. An interpretation of the curve of ΔF versus F_p for fixed δ_1 and δ_2 in terms of the motion of the zeros of the filter is also provided.

SCALED EXTRARIPPLE FILTERS

Consider the transformation

$$x = 0.5 - 0.5 \cos(2\pi f). \quad (7)$$

The interval $0 \leq f \leq 0.5$ is mapped to the interval $0 \leq x \leq 1.0$. It is readily shown that the trigonometric polynomial $H(e^{j2\pi f})$ (2) is converted to a standard polynomial in x of the form

$$P(x) = H(e^{j2\pi f}) \Big|_{x=0.5-0.5 \cos(2\pi f)} \\ = \sum_{n=0}^{\frac{N-1}{2}} a(n) x^n \quad (8)$$

where the sequence $\{a(n)\}$ is related to the sequence $\{h(n)\}$ in a direct manner [4]. The cutoff frequencies F_p and F_s are mapped to the values X_p and X_s as defined by (7).

The linear transformation

$$x' = \alpha x + \beta \tag{9}$$

can now be applied to $P(x)$ to yield a new polynomial $P(x')$ of the form

$$\begin{aligned} P(x') &= P(x) \Big|_{x'=\alpha x + \beta} \\ &= \sum_{n=0}^{N-1} b(n)(x')^n \end{aligned} \tag{10}$$

which is identical in form to (8) and thus corresponds to an FIR filter.

The mapping of (9) has two unspecified constants α and β which determine the interval in x , which is mapped to the interval $0 \leq x' \leq 1$. Two possibilities are considered.

Case 1: (a) The point $x = 0$ is mapped to $x' = 0$. (b) The point $x = X_p$ is mapped to $x' = X'_p$ with $X'_p \geq X_p$.

Case 2: (a) The point $x = 1$ is mapped to $x' = 1$. (b) The point $x = X_p$ is mapped to $x' = X'_p$ with $X'_p \leq X_p$.

Case 1 leads to the following values for α and β :

$$\begin{aligned} \beta &= 0 \\ \alpha &= X'_p / X_p. \end{aligned} \tag{11}$$

Case 2 leads to the following values for α and β :

$$\begin{aligned} \beta &= \frac{X_p - X'_p}{X_p - 1} \\ \alpha &= \frac{X'_p - 1}{X_p - 1}. \end{aligned} \tag{12}$$

For the case of extraripple filters this linear scaling can preserve the necessary conditions for the optimality of the filter. This process is illustrated in Fig. 2. Fig. 2(a) shows $P(x)$ for an extraripple filter with cutoff frequencies X_p and X_s . $P(x)$ has $(N + 1)/2$ extrema in this example, where $N = 11$. (The error curve has $(N + 5)/2$ extrema since it has extrema at $x = X_p$ and $x = X_s$.) By using the linear transformation of Case 1, the curve $P(x')$ is obtained as shown in Fig. 2(b). In this case $P(x')$ still has $(N + 1)/2$ extrema but the extremum at $x = 1$ is not of value $+\delta_2$ but instead is smaller. However the filter still satisfies the optimality criterion discussed earlier and is thus an optimal filter.

By making X'_p (the point which X_p maps to) successively larger, the amplitude of the extremum at $x = +1$ gets successively smaller until it just becomes equal to $-\delta_2$ (i.e., the extremum disappears). This case is illustrated in Fig. 2(c). At this point $P(x')$ has exactly $(N - 1)/2$ extrema, all of equal amplitude. Scaling beyond this point will produce a filter which no longer satisfies the optimality criterion.

Fig. 3 illustrates what happens in Case 2. Fig. 3(a) again shows $P(x)$ for the extraripple filter. By using the linear trans-

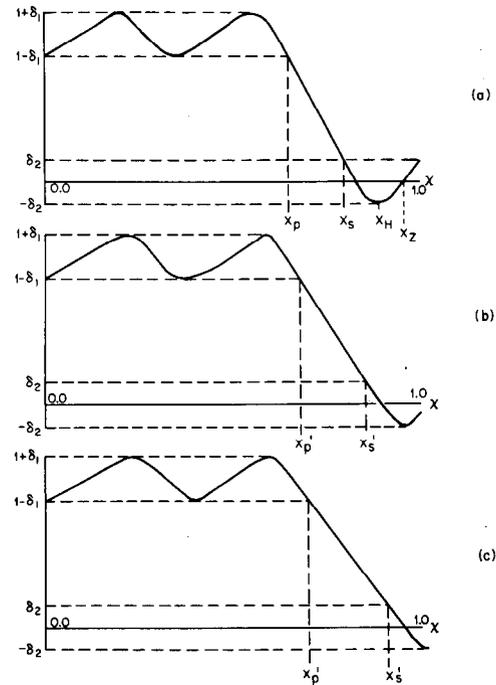


Fig. 2. Linear scaling of the polynomial $P(x)$ with $x = 0$ mapped to $x' = 0$ and $x = X_p$ mapped to $x' = X'_p (X'_p > X_p)$.

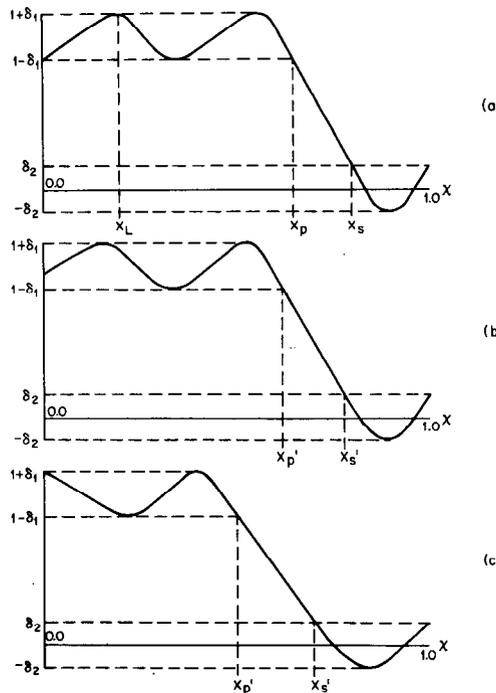


Fig. 3. Linear scaling of the polynomial $P(x)$ with $x = 1$ mapped to $x' = 1$ and $x = X_p$ mapped to $x' = X'_p (X'_p < X_p)$.

formation of Case 2, the curve $P(x')$ still has $(N + 1)/2$ extrema but the extremum at $x = 0$ is not of value $1 - \delta_1$ but instead is larger. Again the scaled filter still satisfies the optimality criterion discussed earlier and is thus an optimal filter (although with different cutoff frequencies).

By making X'_p successively smaller, the amplitude of the extremum at $x = 0$ gets successively larger until it just becomes equal to $1 + \delta_1$. This case is illustrated in Fig. 3(c). As before, at this point $P(x')$ has exactly $(N - 1)/2$ extrema, all of equal amplitude. Scaling beyond this point will produce a filter which no longer satisfies the optimality criterion.

Since the scaling is linear it is easy to see that $\Delta X'$, the transition bandwidth of the scaled filters, must *always* be greater than ΔX of the extraripple filter. For Case 1, for example, the stopband value X_s is scaled to X'_s defined as

$$X'_s = \frac{X'_p}{X_p} X_s. \quad (13)$$

Therefore,

$$\Delta X' = X'_s - X'_p = \frac{X'_p}{X_p} X_s - X'_p \quad (14)$$

$$= \frac{X'_p}{X_p} (X_s - X_p) \quad (15)$$

$$= \Delta X \cdot \frac{X'_p}{X_p} > \Delta X \quad (16)$$

since $X'_p > X_p$.

Similarly for Case 2,

$$X'_s = \frac{X_s(X'_p - 1) + (X_p - X'_p)}{X_p - 1}. \quad (17)$$

Therefore,

$$\Delta X' = X'_s - X'_p \quad (18)$$

$$= \frac{(X'_p - 1)}{(X_p - 1)} (X_s - X_p) \quad (19)$$

$$= \Delta X \left(\frac{X'_p - 1}{X_p - 1} \right) > \Delta X \quad (20)$$

since $X'_p < X_p$. Thus the fact that extraripple filters are minima along the curve of ΔF versus F_p is verified by the scaling argument. (The mapping from x to f is monotonic so the above statement concerning ΔX versus X_p can be extended to ΔF versus F_p .)

It is straightforward to determine the range of scaling which can be used in Cases 1 and 2 and which still preserves the optimality criterion. In Case 1, the most scaling which can be tolerated is the case where the largest extremum value [denoted as X_H in Fig. 2(a)] is mapped to $x' = 1.0$. (Of course $x = 0$ is still mapped to $x' = 0$.) These conditions give

$$\begin{aligned} \beta &= 0 \\ \alpha &= 1/X_H. \end{aligned} \quad (21)$$

The new cutoff frequencies X'_p and X'_s are determined as

$$\begin{aligned} X'_p &= X_p/X_H \\ X'_s &= X_s/X_H. \end{aligned} \quad (22)$$

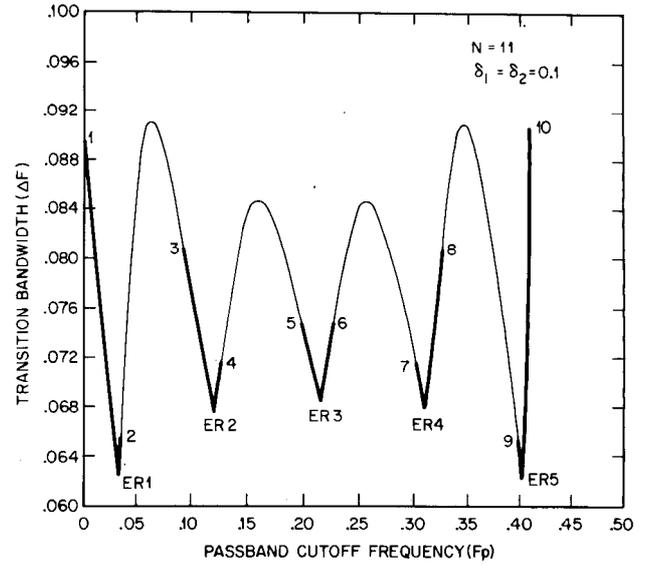


Fig. 4. Transition bandwidth versus passband cutoff frequency for filters with $N = 11$, $\delta_1 = \delta_2 = 0.1$, showing in heavy lines the regions of the curve which are scaled extraripple filters.

The new filter cutoff frequencies (in the f -space) F'_p , F'_s are related to X'_p and X'_s by the inverse of (7), which is

$$f = \frac{\cos^{-1}(1 - 2x)}{2\pi}. \quad (23)$$

In Case 2, the most scaling which can be tolerated is the case where the smallest extremum value [denoted as X_L in Fig. 3(a)] is mapped to $x' = 0$ ($x = 1$ is still mapped to $x' = 1$). These conditions give

$$\begin{aligned} \alpha &= 1/(1 - X_L) \\ \beta &= -X_L/(1 - X_L). \end{aligned} \quad (24)$$

The cutoff frequencies X'_p and X'_s are determined as

$$\begin{aligned} X'_p &= \frac{X_p - X_L}{1 - X_L} \\ X'_s &= \frac{X_s - X_L}{1 - X_L}. \end{aligned} \quad (25)$$

EXAMPLES

According to the preceding discussion, a fairly large class of optimal filters may be determined from the set of extraripple designs for fixed values of δ_1 and δ_2 . In this section examples are presented to demonstrate this effect.

Fig. 4 shows a plot of ΔF versus F_p for filters of duration $N = 11$ samples with $\delta_1 = \delta_2 = 0.1$. This curve is identical to the one in [7]. The five points corresponding to extraripple designs are denoted by ER1, ER2, ..., ER5. The parts of the curve which are drawn in more heavily than the rest of the curve represents the class of optimal filters which are just scaled versions of the extraripple filters. Table I gives values of F_p , F_s , F_L , and F_H (X_p , X_s , X_L , and X_H) for the extraripple filters, as well as values of F'_p , F'_s , X'_p , and X'_s for the filters which represent the maximum tolerable scaling. These filters are labeled as points 1-10 on Fig. 4.

TABLE I
DATA ON FILTERS FOR $N = 11, \delta_1 = \delta_2 = 0.1$

<u>ER1</u>	$F_p = .034406$ $X_p = .011638$ $F_L = .034406$ $X_L = .011638$	$F_s = .096799$ $X_s = .089663$ $F_H = .4034091$ $X_H = .9107104$
<u>Point 1</u>	$X'_p = 0.$ $F'_p = 0.$	$X'_s = .078944$ $F'_s = .090656$
<u>Point 2</u>	$X'_p = .012779$ $F'_p = .036060$	$X'_s = .098454$ $F'_s = .101593$
<u>ER2</u>	$F_p = .1213330$ $X_p = .138395$ $F_L = .0795455$ $X_L = .061161$	$F_s = .1891370$ $X_s = .3134196$ $F_H = .4090909$ $X_H = .920627$
<u>Point 3</u>	$X'_p = .082266$ $F'_p = .092598$	$X'_s = .268692$ $F'_s = .173456$
<u>Point 4</u>	$X'_p = .150327$ $F'_p = .126737$	$X'_s = .340442$ $F'_s = .198307$
<u>ER3</u>	$F_p = .215756$ $X_p = .393247$ $F_L = .0909091$ $X_L = .079373$	$F_s = .2842440$ $X_s = .6067526$ $F_H = .4090909$ $X_H = .920627$
<u>Point 5</u>	$X'_p = .340935$ $F'_p = .198473$	$X'_s = .572848$ $F'_s = .273271$
<u>Point 6</u>	$X'_p = .427152$ $F'_p = .226729$	$X'_s = .659065$ $F'_s = .301527$
<u>ER4</u>	$F_p = .3108630$ $X_p = .686580$ $F_L = .0909091$ $X_L = .079373$	$F_s = .378667$ $X_s = .861605$ $F_H = .4204545$ $X_H = .938839$
<u>Point 7</u>	$X'_p = .659558$ $F'_p = .301693$	$X'_s = .849673$ $F'_s = .373263$
<u>Point 8</u>	$X'_p = .731308$ $F'_p = .326544$	$X'_s = .917734$ $F'_s = .407402$
<u>ER5</u>	$F_p = .4032010$ $X_p = .9103372$ $F_L = .0965909$ $X_L = .089290$	$F_s = .465594$ $X_s = .988362$ $F_H = .465594$ $X_H = .988362$
<u>Point 9</u>	$X'_p = .901546$ $F'_p = .398407$	$X'_s = .987221$ $F'_s = .463940$
<u>Point 10</u>	$X'_p = .921056$ $F'_p = .409344$	$X'_s = 1$ $F'_s = .5$

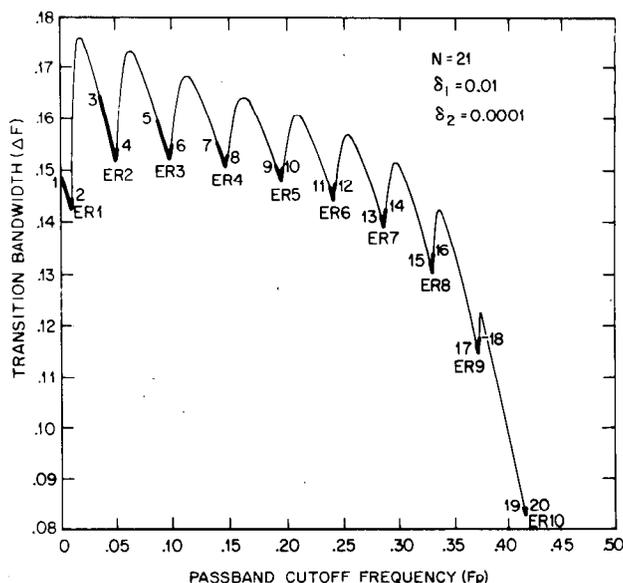


Fig. 5. Transition bandwidth versus stopband cutoff frequency for filters with $N = 21, \delta_1 = 0.01, \delta_2 = 0.0001$, showing in heavy lines the regions of the curve which are scaled extraripple filters.

Fig. 5 shows a plot of ΔF versus F_p for filters of duration $N = 21$ samples with $\delta_1 = 0.01, \delta_2 = 0.0001$, i.e., $K = 100$. The ten extraripple solutions are denoted as ER1-ER10 in Fig. 5. The heavy parts of the curve again represent the class of optimal filters which are scaled versions of the extraripple filters. The values corresponding to the maximum tolerable scaling are denoted as points 1-20 in Fig. 5. Table II gives numerical values (in both f and x) for all these points as well as for the extraripple designs.

REGIONS BETWEEN SCALED FILTERS

As seen in Figs. 4 and 5, the scaling procedures explain a fairly large region of the curve of transition bandwidth as a function of passband cutoff frequency. However, there still remain regions between the scaled filters (the heavy lines in Figs. 4 and 5) which cannot be accounted for by a simple scaling argument. The behavior of $P(x)$, the polynomial representing the filter, in these regions can be explained by the following argument. Referring to Fig. 2, it is seen that at the extreme point of linear scaling [Fig. 2(c)], the last ripple of $P(x)$ is scaled to $x = 1.0$. Additional linear scaling takes this ripple out of the region $0 \leq x \leq 1.0$ and hence gives a filter which no longer satisfies the optimality criterion. However, the position of the last extremum of $P(x)$ can be mapped to a value beyond $x = 1.0$ with an amplitude less than $-\delta_2$ (in this case), at the same time preserving the value of $-\delta_2$ for the polynomial at $x = 1.0$. The effect of this mapping is to increase both X_p and X_s [9] and hence generate the filters in the region between the linearly scaled solutions.

As the position to which the last extremum is mapped tends toward $x = \infty$, the amplitude of the extremum also increases to infinity. At the point where the ripple is mapped to infinity (with infinite amplitude), the filter obtained corresponds to the next lower order extraripple filter [7].

The last ripple may also be mapped to negative values of x in the range $-\infty \leq x \leq 0$. This region corresponds to parts of

TABLE II
DATA ON FILTERS FOR $N = 11$, $\delta_1 = 0.01$, $\delta_2 = 0.0001$

<u>ER1</u>	$F_p = 0.0096290$	$F_s = 0.1517370$	<u>ER6</u>	$F_p = 0.2416280$	$F_s = 0.3860150$
	$X_p = 0.0009148$	$X_s = 0.2105396$		$X_p = 0.4737107$	$X_s = 0.8771566$
	$F_L = 0.0096290$	$F_H = 0.4583333$		$F_L = 0.0476190$	$F_H = 0.4672619$
	$X_L = 0.0009148$	$X_H = 0.9829629$		$X_L = 0.0222136$	$X_H = 0.9894592$
<u>Point 1</u>	$X'_p = 0.$	$X'_s = 0.2098167$	<u>Point 11</u>	$X'_p = 0.4617544$	$X'_s = 0.8743659$
	$F'_p = 0.$	$F'_s = 0.1514546$		$F'_p = 0.2378141$	$F'_s = 0.3846684$
<u>Point 2</u>	$X'_p = 0.0009307$	$X'_s = 0.2141887$	<u>Point 12</u>	$X'_p = 0.4787572$	$X'_s = 0.8865011$
	$F'_p = 0.0097121$	$F'_s = 0.1531571$		$F'_p = 0.2432362$	$F'_s = 0.3906229$
<u>ER2</u>	$F_p = 0.0501950$	$F_s = 0.2027310$	<u>ER7</u>	$F_p = 0.2866440$	$F_s = 0.4257060$
	$X_p = 0.0246614$	$X_s = 0.3536736$		$X_p = 0.6141061$	$X_s = 0.9465058$
	$F_L = 0.0357143$	$F_H = 0.4583333$		$F_L = 0.0476190$	$F_H = 0.4732143$
	$X_L = 0.0125361$	$X_H = 0.9829629$		$X_L = 0.0222136$	$X_H = 0.9929355$
<u>Point 3</u>	$X'_p = 0.0122793$	$X'_s = 0.3454684$	<u>Point 13</u>	$X'_p = 0.6053393$	$X'_s = 0.9452905$
	$F'_p = 0.0353451$	$F'_p = 0.1999923$		$F'_p = 0.2837837$	$F'_s = 0.4248510$
<u>Point 4</u>	$X'_p = 0.0250888$	$X'_s = 0.3598036$	<u>Point 14</u>	$X'_p = 0.6184753$	$X'_s = 0.9532399$
	$F'_p = 0.0506318$	$F'_s = 0.2047677$		$F'_p = 0.2880740$	$F'_s = 0.4306204$
<u>ER3</u>	$F_p = 0.0984610$	$F_s = 0.2511880$	<u>ER8</u>	$F_p = 0.3301360$	$F_s = 0.4605160$
	$X_p = 0.0926686$	$X_s = 0.5037322$		$X_p = 0.7412511$	$X_s = 0.9846922$
	$F_L = 0.0446429$	$F_H = 0.4583333$		$F_L = 0.0476190$	$F_H = 0.4821429$
	$X_L = 0.0195414$	$X_H = 0.9829629$		$X_L = 0.0222136$	$X_H = 0.9968561$
<u>Point 5</u>	$X'_p = 0.0745847$	$X'_s = 0.493841$	<u>Point 15</u>	$X'_p = 0.7353728$	$X'_s = 0.9843444$
	$F'_p = 0.0880496$	$F'_s = 0.248040$		$F'_p = 0.3280075$	$F'_s = 0.4600677$
<u>Point 6</u>	$X'_p = 0.0942747$	$X'_s = 0.512463$	<u>Point 16</u>	$X'_p = 0.7435889$	$X'_s = 0.9877977$
	$F'_p = 0.0993392$	$F'_s = 0.253968$		$F'_p = 0.3309868$	$F'_s = 0.4647663$
<u>ER4</u>	$F_p = 0.1472180$	$F_s = 0.2980290$	<u>ER9</u>	$F_p = 0.3725390$	$F_s = 0.4869850$
	$X_p = 0.1990819$	$X_s = 0.6486078$		$X_p = 0.8480444$	$X_s = 0.9983291$
	$F_L = 0.0476190$	$F_H = 0.4613095$		$F_L = 0.0476190$	$F_H = 0.4910714$
	$X_L = 0.0222136$	$X_H = 0.9852983$		$X_L = 0.0222136$	$X_H = 0.9992134$
<u>Point 7</u>	$X'_p = 0.1808865$	$X'_s = 0.6406248$	<u>Point 17</u>	$X'_p = 0.8445922$	$X'_s = 0.9982912$
	$F'_p = 0.1398340$	$F'_s = 0.2953744$		$F'_p = 0.3710155$	$F'_s = 0.4868379$
<u>Point 8</u>	$X'_p = 0.2020524$	$X'_s = 0.6582857$	<u>Point 18</u>	$X'_p = 0.8487120$	$X'_s = 0.9991150$
	$F'_p = 0.1483987$	$F'_s = 0.3012659$		$F'_p = 0.3728352$	$F'_s = 0.4905293$
<u>ER5</u>	$F_p = 0.1950670$	$F_s = 0.3431450$	<u>ER10</u>	$F_p = 0.4166480$	$F_s = 0.4991710$
	$X_p = 0.3308291$	$X_s = 0.7762028$		$X_p = 0.9329834$	$X_s = 0.9999932$
	$F_L = 0.0476190$	$F_H = 0.4642857$		$F_L = 0.0476190$	$F_H = 0.4991710$
	$X_L = 0.0222136$	$X_H = 0.9874639$		$X_L = 0.0222136$	$X_H = 0.9999932$
<u>Point 9</u>	$X'_p = 0.3156267$	$X'_s = 0.7711185$	<u>Point 19</u>	$X'_p = 0.9314609$	$X'_s = 0.9999931$
	$F'_p = 0.1898935$	$F'_s = 0.3412113$		$F'_p = 0.4156840$	$F'_s = 0.4991617$
<u>Point 10</u>	$X'_p = 0.3350290$	$X'_s = 0.7860568$	<u>Point 20</u>	$X'_p = 0.9329897$	$X'_s = 1.0$
	$F'_p = 0.1964854$	$F'_s = 0.3469382$		$F'_p = 0.4166520$	$F'_s = 0.5$

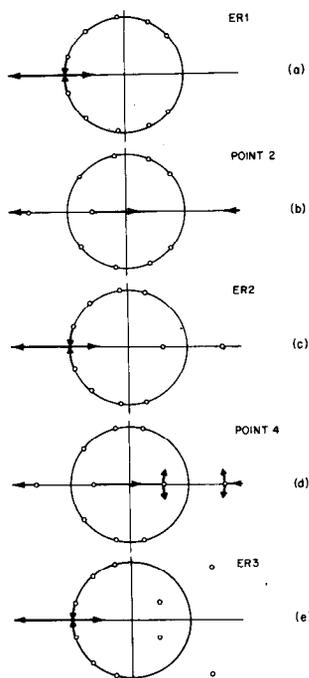


Fig. 6. The positions and motions of the zeros of the filters of Fig. 4, as a function of passband cutoff frequency.

the curve beyond the linear scaling of Fig. 3(c), where the first extremum has been scaled to $x = 0$. The mapping mechanism below $x = 0$ is similar to that described above beyond $x = 1$.

Although this explanation accounts for the behavior of the optimal solution in the gaps between scaled extraripple filters, numerical solutions are no easier to obtain in this manner than by direct solution of the discrete Chebyshev approximation problem with the appropriate parameters.

MOVEMENT OF ZEROS

It is interesting to consider the positions of the zeros of the optimal filters as a function of F_p . Fig. 6 shows a sequence of plots of the positions of the zeros at points ER1, 2, ER2, 4, ER3 of Fig. 4. For this case $N = 11$, so there are 10 zeros in the z -plane.

For filters with F_p less than or equal to the passband cutoff of the first extraripple filter, there are five complex conjugate pairs of zeros, all on the unit circle as seen in Fig. 6(a). The heavy lines in Fig. 6(a) show the movement of one pair of zeros as F_p is increased. The pair of zeros closest to $z = -1$ move along the unit circle to form a double zero at $z = -1$. For the filter corresponding to this point, both the frequency response and its derivative are zero at $f = 0.5$. In terms of scaling extraripple filter 1, the point X_Z of Fig. 2(a) (where $P(X_Z) = 0$) between the last and next to last extrema is scaled to $x' = 1$ or $f = 0.5$.

As F_p is increased to point 2 on Fig. 4, the zeros move to the positions indicated in Fig. 6(b). The zeros closest to $z = -1$ are now a real pair of mirror-image zeros on the negative real axis in the z -plane. The heavy lines in Fig. 6(b) show the move-

ment of the pair of real zeros as F_p increases. The real zeros split further apart until one is at $z = -\infty$ and the other is at $z = 0$. This case corresponds to an extraripple filter of length 9, as discussed elsewhere [7].

As F_p increases, the real zeros come closer together on the positive real axis until the situation of Fig. 6(c), corresponding to filter ER2, is reached. As shown by the heavy lines in Fig. 6(c), at this point, as F_p increases the pair of zeros on the unit circle closest to $z = -1$ begin moving along the unit circle to $z = -1$ at which point they split up and drift apart along the negative real axis. Fig. 6(d) shows the position of the zeros at the point 4 in Fig. 4. There are now 4 zeros on the real axis.

Further increases in F_p move the zeros on the negative real axis further apart until one is at $z = 0$ and the other is at $z = -\infty$ corresponding to another $N = 9$ extraripple filter. These zeros then come back together along the positive real axis until they merge with the other pair of zeros on the positive real axis and split apart to form a quadruplet with the required complex conjugate and mirror-image symmetry. This situation is illustrated in Fig. 6(e) which shows the positions of the zeros at ER3 in Fig. 4.

Much the same behavior of the zeros is obtained as F_p increases even further. The zeros closest to $z = -1$ on the unit circle merge at $z = -1$, split apart along the negative real axis, and return along the positive real axis to form either a real pair of zeros or merge with another real pair of zeros to form a quadruplet of zeros.

In terms of the motion of the zeros, it is easy to understand the equivalence between extraripple filters of length $(N - 2)$ samples and optimal (not extraripple) filters of length N samples. Furthermore, the movement of zeros off the unit circle (i.e., stopband zeros) to become either complex quadruplet zeros or real axis zeros (i.e., passband zeros) is readily understood from the above discussion.

SUMMARY

This paper has discussed how the set of $(N - 1)/2$ extraripple filters may be scaled (within some specified bounds) to yield an infinite number of optimal filters with new passband and stopband cutoff frequencies. Examples were given to illustrate the effects in specific cases. A detailed study of the movement of the filter zeros, as a function of passband cutoff frequency, provided a reasonable explanation for several noted properties of optimal filters.

REFERENCES

- [1] L. R. Rabiner, "Technique for designing finite-duration impulse-response digital filters," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 188-195, Apr. 1971.
- [2] L. R. Rabiner, B. Gold, and C. A. McGonegal, "An approach to the approximation problem for nonrecursive digital filters," *IEEE Trans. Audio Electroacoust.*, vol. AU-18, pp. 83-106, June 1970.
- [3] O. Herrmann, "On the design of nonrecursive digital filters with linear phase," *Electron. Lett.*, vol. 6, no. 11, pp. 328-329, 1970.
- [4] E. Hofstetter, A. V. Oppenheim, and J. Seigel, "A new technique for the design of nonrecursive digital filters," in *Proc. 5th Annu. Princeton Conf. Information Sciences and Systems*, pp. 64-72, 1971.
- [5] T. W. Parks and J. H. McClellan, "Chebyshev approximation for nonrecursive digital filters with linear phase," *IEEE Trans. Circuit Theory*, vol. CT-19, pp. 189-194, Mar. 1972.

- [6] L. R. Rabiner, "The design of finite impulse response digital filters using linear programming techniques," *Bell Syst. Tech. J.*, vol. 51, pp. 1177-1198, July-Aug. 1972.
- [7] T. W. Parks, L. R. Rabiner, and J. H. McClellan, "On the transition width of finite impulse-response digital filters," *IEEE Trans. Audio Electroacoust.*, vol. AU-21, pp. 1-4, Feb. 1973.
- [8] T. W. Parks and J. H. McClellan, "A program for the design of linear phase finite impulse response digital filters," *IEEE Trans. Audio Electroacoust.*, vol. AU-20, pp. 195-199, Aug. 1972.
- [9] J. Siegel, "Design of nonrecursive approximations to digital filters with discontinuous frequency responses," Ph.D. dissertation, M.I.T., Cambridge, Mass., June 1970.

Analog and Digital Filtering in Multiplex Communication Systems

CARL F. KURTH

Abstract—The design and realization of filters in communication systems have been influenced by various technologies over the past years. Beyond the traditional *LC*-filter technology, a sophisticated crystal-filter and, more recently, a mechanical-filter technology have been developed. Each filter technology is applied in multiplex communication systems according to its inherent technical and economical advantages.

With the economical acceptability of integrated circuits, two more filter technologies begin to compete with the established filter concepts. Active filters in thin-film realization show promising aspects for certain applications and for mass production. Digital-filter techniques seem to be understood well enough that it is worth looking for their application in multiplex systems.

The various filter techniques, new and old, are set into the perspective of their economical use in multiplex communication systems.

INTRODUCTION

IN MODERN communication technology, for economic reasons, every effort is being made to use only one pair of wires or one wide-band communication link for simultaneous transmission of many individual channels. This necessitates the construction of multiplex systems in which the signals of individual channels are processed to appear as a combined signal at the output terminal of those systems.

These multiplex communication systems are based on a hierarchy, which means that a smaller number of channels on separate wires is combined into a group of channels on one wire. This is continued in further steps by combining several groups of channels into a higher order group until the necessary number of channels is combined at the output terminal of a system. This process is indicated in Fig. 1. Each individual combination is accomplished by a modulation step which interleaves the channels in the frequency or in the time domain without interference. According to the two different methods

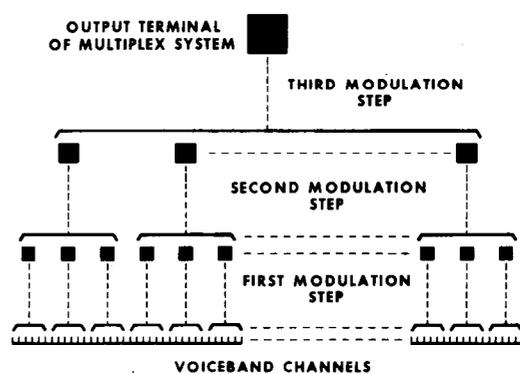


Fig. 1. Hierarchy of multiplex communication system.

of interleaving, two different hierarchies exist, which are referred to as frequency division multiplexing (FDM) and time division multiplexing (TDM) [1].

FREQUENCY DIVISION MULTIPLEXING

FDM is based on amplitude modulation and consecutive linear filtering of one sideband which is used for further transmission. Consequently, this is the area where continuous filters are needed. According to the particular level in the hierarchy the requirements on bandwidth and attenuation of the filters vary [2]–[7]. As indicated in Fig. 2, the filters in higher modulation steps have a large bandwidth; however, the requirements on the attenuation are less stringent since unwanted modulation products are further away from the pass-band. The filters with the most stringent requirements on selectivity follow the first modulation step necessary to generate the single sideband signal for each individual channel. In modern FDM communication systems the channels are staggered over the frequency band in 4-kHz slots; thus the individual channel filters have relatively steep attenuation slopes as indicated in Fig. 3 [4], [8], [11]. Within each 4-kHz slot a bandwidth of 3.2 kHz is used for transmission. One of these