RICE UNIVERSITY

CHEBYSHEV APPROXIMATION FOR NON-RECURSIVE

DIGITAL FILTERS
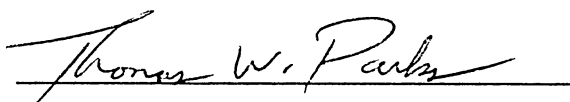

by

James H. McClellan

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

Master of Science

Thesis Director's Signature

Houston, Texas

December 1971

ABSTRACT

CHEBYSHEV APPROXIMATION FOR NON-RECURSIVE
DIGITAL FILTERS

by James H. McClellan

An efficient procedure for the design of finite
length impulse response filters with linear phase is pre-
sented. The algorithm obtains the optimum Chebyshev
approximation on separate intervals corresponding to pass
and/or stop bands, and is capable of designing very long
filters. This approach allows the exact specification of
arbitrary band-edge frequencies as opposed to previous
algorithms which could not directly control pass and stop
band locations and could only obtain $\frac{N-1}{2}$ different band-
edge locations for a length N low-pass filter, for fixed
δ_1 and δ_2.

As an aid in practical application of the algorithm,
several graphs are included to show relations among the
parameters of filter length, transition width, band-edge
frequencies, passband ripple, and stopband attenuation.

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor, Dr. T. W. Parks, who provided the invaluable ideas necessary to start this research, and my wife, Carolyn, who has tolerated my long hours of work in completing this research.

I. INTRODUCTION

The filtering of time-signals by RLC networks is a familiar problem in Electrical Engineering. A similar operation can be performed by a digital computer if the signal is sampled prior to filtering and reconstructed afterwards. It is well known that the mathematical form of an analog filter is the ratio of two polynomials in the Laplace frequency variable s. Likewise, a digital filter can be expressed as the ratio of two polynomials in the z-transform variable z.

However, a key feature of digital filters is that realization of a polynomial in z (called a non-recursive filter) is possible, in contrast to analog filters which always have a denominator polynomial. This realization is made practical by the Fast Fourier Transform (FFT) algorithm which enables high order polynomials to be implemented efficiently.

Non-recursive filters have some special properties which make them important for digital signal processing. First, since a non-recursive filter has no poles, it is not necessary to worry about the stability of the filter. In the case of recursive filters (filters with poles)

stability is a major concern. Secondly, a non-recursive filter is less sensitive to realizations in finite length arithmetic than a recursive filter. These two properties are important because the recursive filter is a much more compact design in terms of the number of filter coefficients needed. Typically, a recursive filter might have ten coefficients whereas a non-recursive filter might have one or two hundred. Thus, there is a tradeoff between the complexity of the filter and the problems of stability and coefficient quantization. Thirdly, exactly linear phase can be obtained with a non-recursive filter. This is important in applications where the delay distortion must be kept small. Finally the design of non-recursive filters is easier because the problem is linear, whereas recursive filter design is a non-linear problem.

The techniques for designing low pass non-recursive filters are classified as least squares designs and Chebychev (or min-max) designs. In the former, the basic approach is to truncate the Fourier series of the rectangular low pass shape

$$R(F) = \begin{cases} 1 & 0 \le F \le F_c \\ 0 & F_c \le F \le 0.5 \end{cases}$$

and then use the Fourier series coefficients as the impulse

response coefficients. In the latter, the approach is to interpolate to a function which has the equiripple property of Chebychev approximation, namely, the error reaches its maximum value n times where n is the number of approximating functions. The recent review articles [1], [2] survey the status of available design techniques.

The least squares designs are handicapped by the Gibbs phenomenon which results from trying to approximate the discontinuity in $R(F)$ at $F = F_c$. The windowing technique [3] improves this situation by multiplying the frequency response by a carefully chosen window function to smooth out the overshoot. Thus the impulse response is convolved with the impulse response of the window function. However, the "optimum" window functions are very complicated.

The Chebychev norm is a more natural norm for filter design since the object is to minimize the peak value of the error. The frequency sampling method [4] was the first procedure to use this norm. By allowing only one, two, or three of the parameters in the design to vary the maximum stopband error is minimized subject to the constraint that the maximum passband deviation be less than some fixed value.

The idea of Hermann [5], which was made computationally feasible by Hofstetter [6], is to fix the passband

deviation δ_1 and the stopband deviation δ_2 and interpolate the approximating function through the points $1 \pm \delta_1$ in the passband and $\pm \delta_2$ in the stopband. The resulting filter "looks" like an optimal Chebychev design since it has the equiripple property. In view of the results which will be presented here, this filter is a special case of the optimum filters and will be referred to as an extraripple filter.

Since the goal is to design an optimum min-max filter, it is important to classify optimum filters and then view the previous results from this perspective. A new procedure results which includes the extraripple filters as a special case.

II. CHEBYCHEV APPROXIMATION FOR
NON-RECURSIVE FILTERS

A. Problem Formulation

The detailed description of the new procedure described in this paper is in terms of low pass filters. Modifications for the general bandpass case are included. Linear phase, digital filters of length $2n + 1$ have a transfer function

$$G(Z) = \sum_{k=0}^{2n} h_k \, Z^{-k} \tag{1}$$

with $h_k = h_{2n-k}$. The frequency response $\hat{G}(F)$ is obtained with

$$Z = \exp(j\omega T) = \exp(j 2\pi F) \tag{2}$$

where F is the normalized frequency variable, $F = \omega T / 2\pi$.

$$\hat{G}(F) = \exp(-j 2\pi n F) \sum_{k=0}^{n} d_k \, \cos 2k\pi F$$

$$= \exp(-j 2\pi n F) \, H(F) \tag{3}$$

where $d_{n-k} = 2h_k$, $k=0,\ldots,n-1$ and $d_0 = h_n$.

The frequency domain design problem is to find the $\{h_k\}$, k=0,...,2n, or equivalently the d_k, k=0,...,n, so that H(F) has the desired characteristic shown in Figure 1. The parameters in the standard tolerance scheme of Figure 1 are defined as follows:

δ_1 = allowed passband deviation.

δ_2 = allowed stopband deviation.

F_p = desired passband edge.

F_s = desired stopband edge.

There are several ways to meet the prescribed tolerances. The most straightforward approach is to approximate the rectangle function

$$R(F) = \begin{cases} 1 & 0 \le F < \dfrac{F_p + F_s}{2} \\ \\ 0 & \dfrac{F_p + F_s}{2} \le F \le 0.5 \end{cases}$$

The discontinuity in R(F) causes difficulty in both the least square and Chebyshev approximation and has led to procedures where the region between F_p and F_s is filled in with a continuous function such as a straight line with slope of $-(F_s - F_p)^{-1}$ or a section of a cosine. Another approach is to use a weighted least square approximation

with zero weighting in the transition region between F_p and F_s. The theory of weighted Chebyshev approximation on an interval will not allow such a zero weighting function. An indirect approach which ignores the error in a transition region makes use of the equiripple property of Chebyshev approximation by specifying the number of pass and stop band ripples n_p and n_s [5], [6].

This method requires the specification of n_p, n_s, δ_1 and δ_2. A limitation of this approach is that the relationship between the desired band-edge frequencies F_p and F_s and an _a priori_ choice of n_p and n_s is unknown. A trial and error procedure is necessary to meet the F_p and F_s specifications if indeed they can be met by _any_ set of n_p and n_s.

A new approach described in this paper which meets the tolerance scheme of Figure 1 and allows specification of band-edge frequencies formulates the problem as one of Chebyshev approximation on two disjoint intervals (three for the bandpass case). For a filter of length $2n + 1$, the n possible extraripple filters arise as special cases of this approach when appropriate values of F_p and F_s are specified. Rather than minimize the error in the stop band only as in the frequency-sampling method [4], this approach does optimization on several bands. An efficient algorithm will be developed which uses interpolation as in [6] and

achieves comparable speed.

The filter design problem is stated as a weighted Chebyshev approximation problem by defining the passband B_p and stopband B_s as

$$B_p = \{F: 0 \leq F \leq F_p\} \tag{4}$$

$$B_s = \{F: F_s \leq F \leq 0.5\} \ . \tag{5}$$

The desired form for a low-pass H(F) in (3) is then

$$D(F) = \begin{cases} 1 & F \ \epsilon \ B_p \\ 0 & F \ \epsilon \ B_s \ . \end{cases} \tag{6}$$

The weight function

$$W(F) = \begin{cases} \dfrac{1}{K} & F \ \epsilon \ B_p \\ 1 & F \ \epsilon \ B_s \end{cases} \tag{7}$$

allows the designer freedom to specify the relative magnitude of the error in the two bands.

With

$$\Omega = B_p \ \cup \ B_s \ , \tag{8}$$

the union of pass and stop bands, the problem of designing
a lowpass linear phase filter of length 2n+1 becomes the
problem of finding the $\{d_k\}$ k = 0,...,n in

$$H(F) = \sum_{k=0}^{n} d_k \cos 2k\pi F \tag{9}$$

which minimize

$$\underset{F \in \Omega}{\text{maximum}} \quad W(F)|D(F) - H(F)| \tag{10}$$

The next section develops an efficient algorithm to
solve the above problem and the general bandpass problem
where there are three disjoint intervals.

B. Design Procedure

The theory of Chebyshev approximation on compact
sets has established that the problem as formulated in (10)
has a unique solution. More important, however, is the
fact that necessary and sufficient conditions which charac-
terize the best approximation are given by the following
alternation theorem which uses the fact that the set of
functions $\{\cos 2\pi kF\}$ k = 0,...,n satisfies the Haar condi-
tion [7].

Theorem: Let Ω be any closed subset of $[0, 1/2]$. In order that

$$H(F) = \sum_{k=0}^{n} d_k \cos 2\pi kF$$

be the unique best approximation on Ω to $D(F)$ it is necessary and sufficient that the error function $E(F) = W(F)[D(F) - H(F)]$ exhibit on Ω at least $n + 2$ "alternations" thus: $E(F_i) = -E(F_{i-1}) = \pm ||E||$ with $F_0 < F_1 < ... < F$ and $F_i \in \Omega$. Here $||E|| = \max_{F \in \Omega} |E(F)|$.

The extraripple filters in [5], [6] are optimal in terms of this theory and are special cases obtained with the algorithm to be described. To see this, consider an example filter of length 11 with $n_p = 3$ and $n_s = 3$. (Figure 2). For simplicity, let $\delta_1 = \delta_2$. A length 11 filter is an approximation with the set of 6 functions {1, $\cos 2\pi F, ..., \cos 10\pi F$} (n=5). If we consider that the approximation is done on $[0, F_p] \cup [F_s, 0.5]$, then the alternation theorem says that there must be at least 7 alternations on this set. A careful count, including the points 0, F_p, F_s, and 0.5, of the alternations shows that there are 8 alternations in the sense of the alternation theorem. Therefore, the extraripple filters are optimum filters on the set $[0, F_p] \cup [F_s, 0.5]$. Note, however, that this set on which the approximation is best is not known beforehand,

but rather is a result of the design procedure. The fact
that there are n + 3 alternations for the extraripple
filters is true in general (this motivates the name) and
tells us intuitively that the extraripple designs are a
very special case of Chebyshev approximation. In fact,
there are only n different extraripple filters for fixed
δ_1 and δ_2 which means that for an <u>arbitrary</u> location of F_p
and F_s and choice of weight function the optimum design
will usually have only n + 2 alternations.

The extraripple design algorithm proposed by
Hofstetter is capable of designing high order filters very
efficiently; in fact the computation time is roughly pro-
portional to n^2. Hofstetter mentions that the algorithm
is similar to the classical Remes exchange algorithm. As
will be shown, the computational aspects of the Hofstetter
algorithm can be merged with the ascent property of the
Remes exchange method [8] to give the following new algo-
rithm which converges quadratically to the best weighted
Chebyshev approximation on the set Ω.

The algorithm as indicated in Figure 3, begins with
an initial guess of n + 2 points and exchanges points until
it obtains the n + 2 points of alternation (extremal fre-
quencies) of the alternation theorem. This is accomplished
in the following way.

At the i-th step of the algorithm there are n + 2

points on which the error function is forced to have magnitude $|\rho|$ with alternating signs.

For a given set of n + 2 frequencies $\{F_k\}$ k=0,..., n+1, this requires the solution of the n + 2 equations

$$W(F_k) \ (H(F_k) - D(F_k)) = -(-1)^k \rho \quad k=0,...,n+1 \qquad (11)$$

where H(F), D(F) and W(F) are defined in (6), (7) and (9). Equation (11) may be written matrix form as follows:

$$
\begin{bmatrix}
1 & \cos 2\pi F_0 & \cos 4\pi F_0 \cdots \cos 2\pi n F_0 & \dfrac{1}{W(F_0)} \\[2ex]
1 & \cos 2\pi F_1 & & \dfrac{-1}{W(F_1)} \\[2ex]
\cdot & & & \dfrac{1}{W(F_2)} \\[2ex]
\cdot & & & \cdot \\[2ex]
\cdot & & & \cdot \\[2ex]
\cdot & & & \cdot \\[2ex]
1 & \cos 2\pi F_{n+1} & & \dfrac{(-1)^{n+1}}{W(F_{n+1})}
\end{bmatrix}
\begin{bmatrix}
d_0 \\[2ex] d_1 \\[2ex] \cdot \\[2ex] \cdot \\[2ex] \cdot \\[2ex] d_n \\[2ex] \rho
\end{bmatrix}
=
\begin{bmatrix}
D(F_0) \\[2ex] D(F_1) \\[2ex] \cdot \\[2ex] \cdot \\[2ex] \cdot \\[2ex] \cdot \\[2ex] D(F_{n+1})
\end{bmatrix}
$$

$$(12)$$

The system (12) is always non-singular [7] and thus there is a unique solution for ρ and the d's. However, the direct solution of (12) is time consuming. In fact as

pointed out by Temes [9], it is the most difficult part of
the algorithm. It is more efficient to first calculate ρ
analytically as

$$\rho = \frac{a_0 D(F_0) + a_1 D(F_1) + \ldots + a_{n+1} D(F_{n+1})}{\dfrac{a_0}{W(F_0)} - \dfrac{a_1}{W(F_1)} + \dfrac{a_2}{W(F_2)} \ldots (-1)^{n+1} \dfrac{a_{n+1}}{W(F_{n+1})}}$$

where

$$a_k = \prod_{\substack{j=0 \\ j \neq k}}^{n+1} \frac{1}{(x_k - x_j)}$$

and

$$x_i = \cos 2\pi F_i .$$

Next, Lagrange's interpolation formula in the barycentric
form is used to interpolate $H(F)$ on the n + 1 points F_0
$\ldots F_n$ to the values

$$c_k = D(F_k) - (-1)^k \frac{\rho}{W(F_k)} \qquad k = 0, \ldots, n$$

$$H(F) = \frac{\displaystyle\sum_{k=0}^{n} \frac{b_k}{(x - x_k)} c_k}{\displaystyle\sum_{k=0}^{n} \frac{b_k}{(x - x_k)}}$$

where

$$b_k = \prod_{\substack{j=0 \\ j \neq k}}^{n} \frac{1}{(x_k - x_j)} \quad .$$

Note H(F) will also interpolate to $D(F_{n+1}) - (-1)^{n+1}$ $\frac{\rho}{W(F_{n+1})}$ since it satisfies (11). Then the error function E(F) = W(F)[D(F) - H(F)] is evaluated on Ω. However, it is only necessary to evaluate E(F) at a finite number of points. A grid of 20n equally spaced points has proved to be sufficiently dense in Ω. If the error function is such that $|E(F)| \leq |\rho|$ for F ϵ Ω then we have found the best approximation. If $|E(F)| > |\rho|$ for some F ϵ Ω then a new set of n + 2 frequencies must be chosen as candidates for the extremal points. The philosophy of the Remes method [8] is to choose these new frequencies such that $|\rho|$ will be increased at the next iteration. If the new points are chosen to be the peaks of the error curve (i.e., points where $|E(F)| \geq |\rho|$ and E(F) is a local extremum) then $|\rho|$ is forced to increase and ultimately converge to its upper bound which corresponds to the solution of the problem.

The local extrema are determined by finding those points where $H(F_{k+1}) - H(F_k)$ changes sign. The endpoints F_p and F_s are always local extrema. Among these extrema we want to choose the n + 2 points where $|E(F)|$ is the greatest with a search. This search is very easy to implement because of the following properties of the error

function.

Theorem: For a filter of length 2n + 1 (the approximation
is being done with n + 1 functions) the i-th error curve
must exhibit n + 2 or n + 3 peaks thus: sgn $E(F_k)$ =
-sgn $E(F_{k+1})$ k = 0, 1,...,n+1 with $|E(F_k)| \geq |\rho|$; the F_k
are the local maxima and minima of the error curve; two of
the critical peaks are located at F_p and F_s; at least one
of the endpoints is a peak; and both endpoints are peaks if
there are n + 3 peaks.

Figure 4 shows the two cases which can occur for a
length 11 filter.

Proof: The properties of E(F) depend on the fact that a
function of the form

$$P(F) = \sum_{k=1}^{n} a_k \sin2\pi kF$$

has zeros at F = 0 and F = 1/2 and at most n - 1 zeros in
the open interval (0,1/2). For a given set of points the
optimal ρ has been calculated and

$$H(F) = \sum_{k=0}^{n} g_k \cos2\pi kF$$

has been fit through n + 1 points. It is well known that
on the set Ω which is the union of disjoint closed inter-
vals that the relative min and max occur at points where

$$\frac{dE}{dF} = \frac{dH}{dF} = 0 \text{ or at the endpoints.}$$

$$\frac{dH}{dF} = \sum_{k=1}^{n} \alpha_k \sin 2\pi kF$$

has at most n + 1 zeros in the set Ω and the two boundary points F_p and F_s are the only other candidates for extremal points. Hence, there are at most n + 3 peaks of the error curve. Also, the only way that there can be n + 3 peaks is for both endpoints, 0 and 1/2, to be peaks. Since the error curve interpolates n + 2 points in Ω where $|E(F)|$ = $|\rho|$ with alternating sign, there must be at least n + 2 local extrema where $|E(F)| \geq |\rho|$ with alternating sign. If there are n + 3 peaks, compare the magnitude of the error at 0 and 1/2 and delete the point with smaller error to obtain the new set of n + 2 extremal points.

This method can be used to design bandpass filters by doing the approximation problem on three disjoint intervals. The search is complicated by the fact that the error curve can have n + 2, n + 3, n + 4 or n + 5 extremum points. In each case n + 2 points of alternation can be chosen by the search procedure.

C. Results

The design algorithm is implemented by specifying

the filter order n, the weighting factor K, and F_p and F_s which fix the transition width; then the deviations δ_2 and also $\delta_1 = K\delta_2$ are minimized. Figure 5 shows experimental results of the relationships between these parameters.

Figure 5a shows δ_2 versus F_p for a length 29 filter. The local minima at .13765 and .17265 correspond to Hofstetter designs. (See Table 1.) Between the maxima at .1195 and the minima at .13765 there is a difference of 3.32db, indicating what is to be gained by an optimal location of the transition region.

Figure 5b shows the control over δ_2 exerted by the weighting factor. The a priori choice of K is a part of the design procedure. However, if the specifications require a certain stopband attenuation δ_2 while not exceeding a given passband ripple δ_1, then the natural choice for K is the ratio δ_1/δ_2. It is also possible to fix F_p and F_s and obtain $\frac{n_f-1}{2}$ extraripple filters by varying K.

Figure 5c indicates how much is to be gained by increasing the length of the filter, n_f, and Figure 5d shows that reducing the transition width results in larger deviations in the 2 bands.

One principal merit of the finite-length linear phase filter designs is the speed with which they can be obtained. Figure 6 shows experimental data from a

18

single-precision program run on the Burroughs B-5500 computer. In all cases the number of iterations never exceeded 10 and was usually about 6 or 7. (A length 95 filter was designed in 200 seconds with 10 iterations.) Most of the time is spent in the evaluation of the interpolation formula on 20n points. This operation is proportional to n^2 and any efforts to improve the speed would have to be concentrated in this part of the algorithm. The barycentric interpolation formula is used for this reason since there are n fewer operations than with the standard Lagrange interpolation formula.

The length of filters which can be designed is limited by the precision of the machine in the calculation of the a_k coefficients. For the large n many of the differences $(x_i - x_k)$ are very small and thus

$$a_k = \prod_{i \neq k} \frac{1}{(x_i - x_k)}$$

is very large and overflow results. Double precision can be used to design extremely long filters but the time will be increased by about a factor of four.

III. SUMMARY

A technique for the weighted Chebyshev design of non-recursive linear phase digital filters has been presented. Central to this approach is the approximation of the ideal shape on two disjoint intervals. This allows the exact specification of the band edge frequencies, whereas previous procedures obtained the band edge indirectly. Furthermore, the extraripple filters were shown to be special cases of the new design procedure. Finally, an efficient algorithm has been developed to implement the method.

A limitation inherent in this procedure is that it is not possible to independently specify all of the desired parameters. Several graphs have been included based on computational experience showing the relations between parameters. These empirical relationships are an aid to design and show the best that can be attained.

19

REFERENCES

1. L. R. Rabiner, "Techniques for designing finite-duration impulse-response digital filters" IEEE Trans. Com. Tech., vol. COM-19, April 1971, pp. 188-195.

2. H. D. Helms, "Digital filters with equiripple or mini-max responses," IEEE Trans. Audio Electroacoustics, vol. AU-19, March 1971, pp. 87-93.

3. J. F. Kaiser and F. F. Kuo, "System Analysis by Digital Computer," (Wiley, 1966), chap. 7.

4. L. R. Rabiner, B. Gold, and C. A. McGonegal, "An approach to the approximation problem for nonrecursive digital filters," IEEE Trans. Audio Electroacoustics, vol. AU-18, June 1970, pp. 83-106.

5. O. Hermann, "Design of nonrecursive digital filters with linear phase," Electron. Lett., 1970, 6, pp. 328-329.

6. E. Hofstetter, A. V. Oppenheim, and J. Siegel, "A new technique for the design of non-recursive digital filters," presented at the Fifth Annual Princeton Conference on Onformation Sciences and Systems, March 1971.

7. E. W. Cheney, "Introduction to Approximation Theory," New York, McGraw-Hill, 1966, pp. 72-100.

8. E. Ya Remes, "General computational methods of Tchebycheff Approximation," Kiev, 1957 Atomic Energy Commission Translations No. 4491, pp. 1-85.

9. G. C. Temes and J. A. C. Bingham, "Iterative Chebyshev approximation technique for network synthesis," IEEE Trans. Circuit Theory, vol. CT-14, March 1967, pp. 31-37.

TABLE I.   Comparison of Weighting Coefficients Obtained
from Hofstetter's Algorithm and New Algorithm
for a Length 29 Filter

Design Method	Hofstetter	New Design
Parameters Specified	$n_p = 6$, $n_s = 9$ $\delta_1 = .0098747$ $\delta_2 = .00098747$	$F_p = .17265$ $F_s = .26265$ $K = 10$
h_0	-.0033633	-.003357
h_1	-.0060507	-.0060291
h_2	-.0008578	-.00082454
h_3	.0097200	.0097418
h_4	.0083195	.0083190
h_5	-.011336	-.011334
h_6	-.021879	-.021844
h_7	.0047474	.0047995
h_8	.039863	.039882
h_9	.017794	.01773
h_{10}	-.058357	-.058361
h_{11}	-.073671	-.073620
h_{12}	.072324	.072385
h_{13}	.306575	.306583
h_{14}	.422471	.422443
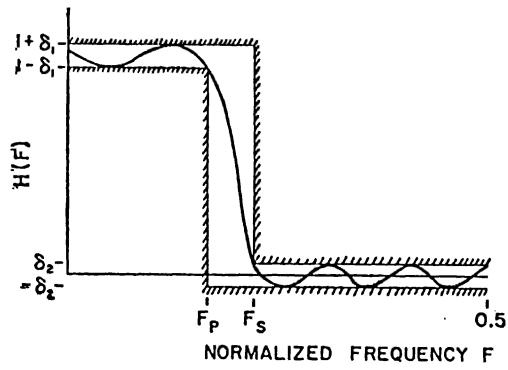
Fig. 1. Desired low-pass filter characteristic.
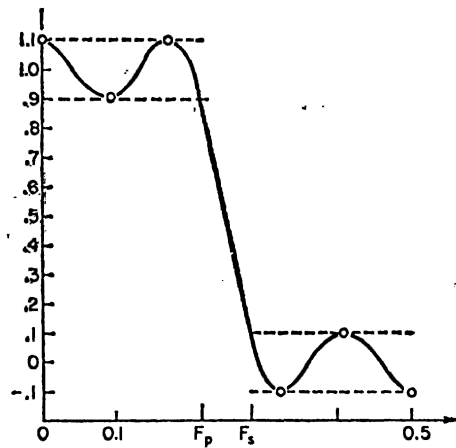
Fig. 2. Extraripple filter showing application
of alternation theorem.
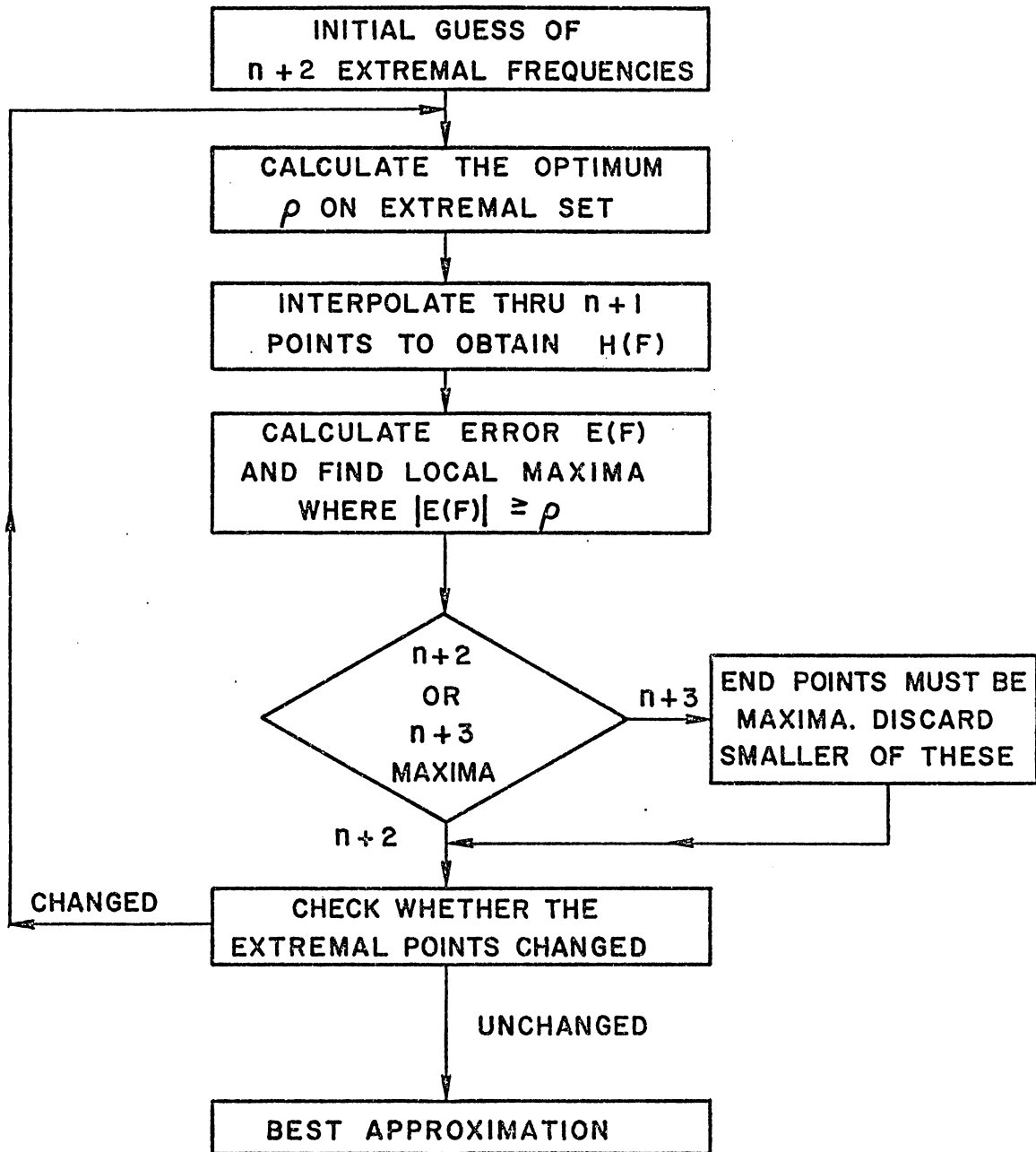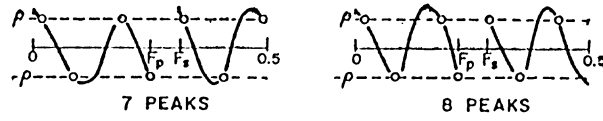
FIG. 3. BLOCK DIAGRAM OF THE DESIGN ALGORITHM
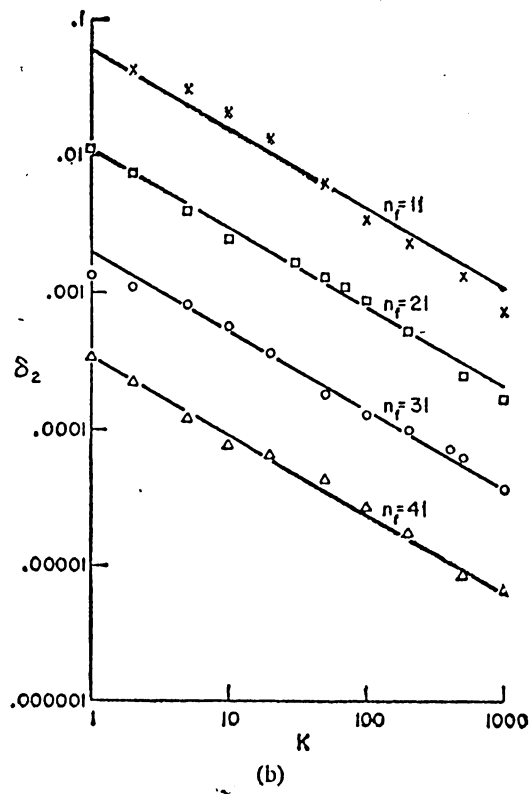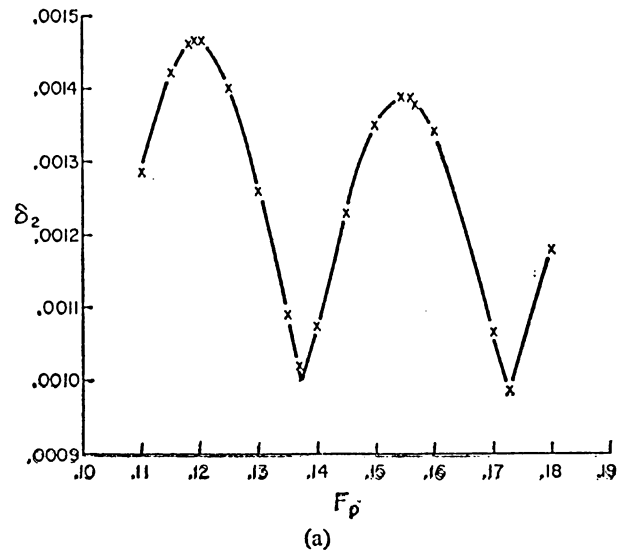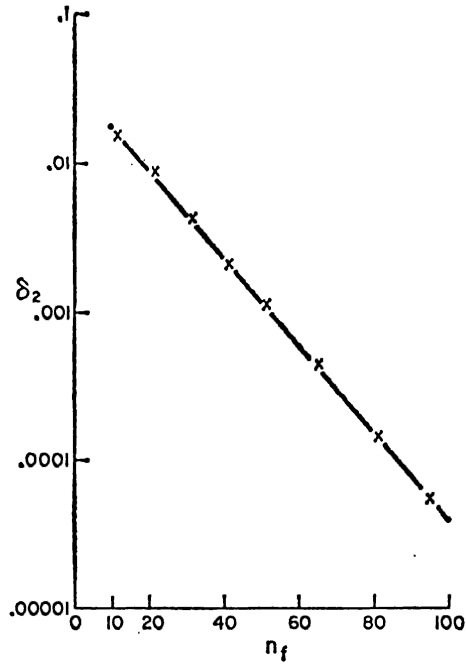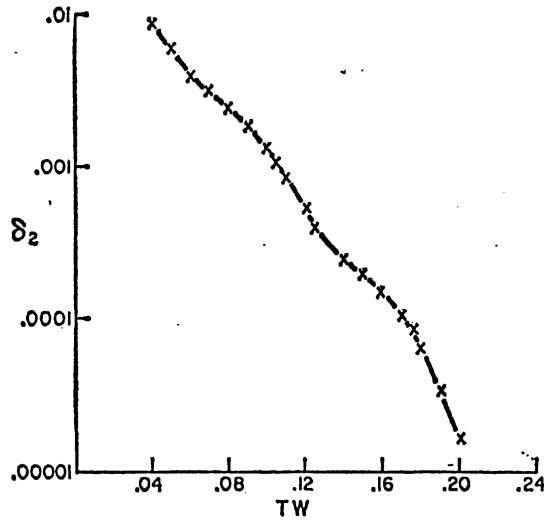
7 PEAKS   8 PEAKS

Fig. 4. Two error-curve possibilities for length 11 filter.

(a)

(b)

Fig. 5. (a) Influence of passband edge F_p on stopband deviation δ_2 for fixed transition width TW = 0.09 and weighting factor K = 10 with length 29 filter. (b) Control of stopband deviation δ_2 exerted by weighting factor K for filter lengths n_f = 11, 21, 31, and 41, with F_p = 0.2 and F_s = 0.3. (c) Dependence of stopband deviation δ_2 on filter length for passband edge F_p = 0.2, stopband edge F_s = 0.24, and weighting factor K = 50. (d) Effect of transition width TW = $F_s - F_p$ on stopband deviation δ_2 for passband edge F_p = 0.2, weighting factor K = 50, and filter length n_f = 21.
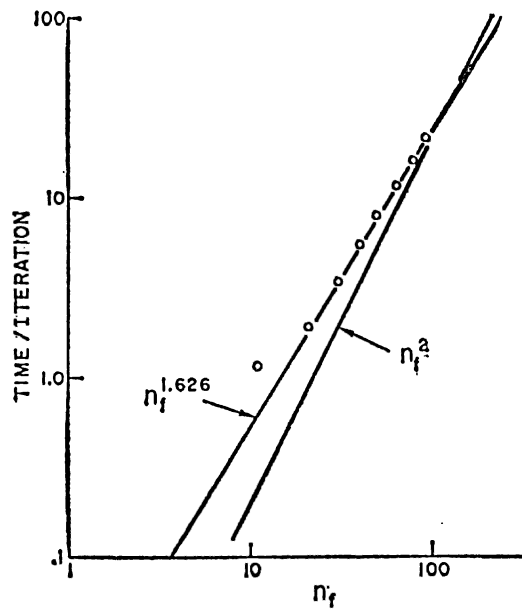
Fig. 6. Computation time as a function of filter length.