[2] ——, "The generalized direct-coupled cavity linear phase filter," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-18, pp. 308–313, June 1970.
[3] ——, "The generalized interdigital linear phase filter," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-18, pp. 301–307, June 1970.
[4] R. E. Baum, "A modification of Brune's method for narrow-band filters," *IRE Trans. Circuit Theory*, vol. CT-5, pp. 264–267, Dec. 1958.
[5] J. D. Rhodes, "Explicit formulas for element values in elliptic function prototype networks," *IEEE Trans. Circuit Theory*, vol. CT-18, pp. 264–276, Mar. 1971.
[6] M. Hibino, Y. Ishizaki, and H. Watanabe, "Design of Chebyshev filters with flat group-delay characteristics," *IEEE Trans. Circuit Theory*, vol. CT-15, pp. 316–325, Dec. 1968.
[7] J. O. Scanlan and J. D. Rhodes, "Unified theory of cascade synthesis," *Proc. Inst. Elec. Eng.*, vol. 117, pp. 665–670, Apr. 1970.
[8] J. D. Rhodes, "Matched-filter theory for Doppler-invariant pulse compression," *IEEE Trans. Circuit Theory*, vol. CT-19, pp. 53–59, Jan. 1972.

Chebyshev Approximation for Nonrecursive Digital Filters with Linear Phase

THOMAS W. PARKS, MEMBER, IEEE, AND JAMES H. McCLELLAN, STUDENT MEMBER, IEEE

Abstract—An efficient procedure for the design of finite-length impulse response filters with linear phase is presented. The algorithm obtains the optimum Chebyshev approximation on separate intervals corresponding to passbands and/or stopbands, 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 stopband locations and could only obtain $(N-1)/2$ 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.

I. Introduction

Finite-duration impulse response filters have received recently renewed attention for several reasons. The finite-length filter offers the advantages of stability, exactly linear phase, and high-speed implementation with the fast Fourier transform. New algorithms have become available to design high-order finite-duration impulse response filters with linear phase which allow very good frequency response characteristics to be obtained.

As discussed in two recent review papers [1], [2], the techniques available include the windowing technique [3], the frequency sampling method [4], [5], and several algorithms for the design of min–max or Chebyshev-type approximations to the desired frequency response [6]–[8].

The linear programming approach [6] and the nonlinear optimization procedure of Herrmann [7] are relatively slow algorithms and are limited to the design of filters with few parameters. The interpolation technique proposed by Hofstetter *et al.* [8] offers a much more efficient algorithm capable of designing longer filters. The algorithms in [7], [8] result in exactly the same filter and will be called an extraripple design in this paper.

The detailed description of the new procedure described here 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 $h=h_{2n-k}$. The frequency response $\hat{G}(F)$ is obtained with

$$Z = \exp(j\omega T) = \exp(j2\pi F) \tag{2}$$

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

$$\hat{G}(F) = \exp(-j2\pi nF) \sum_{k=0}^{n} d_k \cos 2k\pi F = \exp(-j2\pi nF) H(F) \tag{3}$$

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

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

Manuscript received June 29, 1971; revised October 6, 1971. This work was supported by NSF Grant GK23697.

The authors are with Rice University, Houston, Tex.

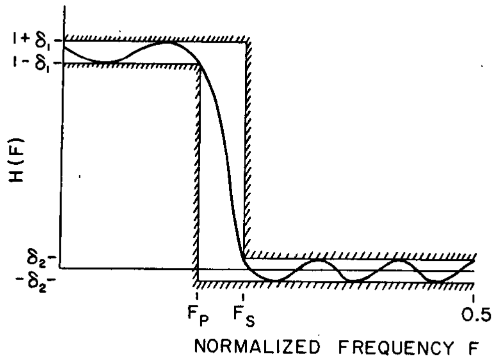

Fig. 1. Desired low-pass filter characteristic.

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 the Chebyshev approximation by specifying the number of pass- and stopband ripples n_p and n_s [7], [8]. 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 Fig. 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 on the stopband 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 [8] 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

$$B_p = \{F\colon 0 \le F \le F_p\} \tag{4}$$

$$B_s = \{F\colon F_s \le F \le 0.5\}. \tag{5}$$

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

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

The weight function

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

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

With

$$\mathfrak{F} = B_p \cup B_s \tag{8}$$

the union of pass- and stopbands, the problem of designing a low-pass linear phase filter of length $2n+1$ becomes the problem of finding the $\{d_k\}$, $k=0, \cdots, n$, in

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

which minimizes

$$\max_{F \in \mathfrak{F}} W(F) \mid D(F) - H(F) \mid . \tag{10}$$

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

II. Design Procedure

The theory of the 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 characterize the best approximation are given by the following alternation theorem which uses the fact that the set of functions $\{\cos 2\pi kF\}$, $k=0, \cdots, n$, satisfies the Haar condition [9].

Theorem

Let $\mathfrak{F}$ be any closed subset of $[0, \frac{1}{2}]$. In order that $H(F) = \sum_{k=0}^{n} d_k \cos 2\pi kF$ be the unique best approximation on $\mathfrak{F}$ to $D(F)$, it is necessary and sufficient that the error function $E(F) = W(F)[D(F) - H(F)]$ exhibit on $\mathfrak{F}$ at least $n+2$ "alternations." Thus, $E(F_i) = -E(F_{i-1}) = \pm \|E\|$ with $F_0 < F_1 < \cdots < F_{n+1}$ and $F_i \in \mathfrak{F}$. Here $\|E\| = \max_{F \in \mathfrak{F}} |E(F)|$.

The extraripple filters in [7], [8] 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 $n_s = 3$ (Fig. 2). For simplicity, let $\delta_1 = \delta_2$. A length 11 filter is an approximation with the set of six functions $\{1, \cos 2\pi F, \cdots, \cos 10\pi F\}$, $n=5$. 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 seven 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 eight 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

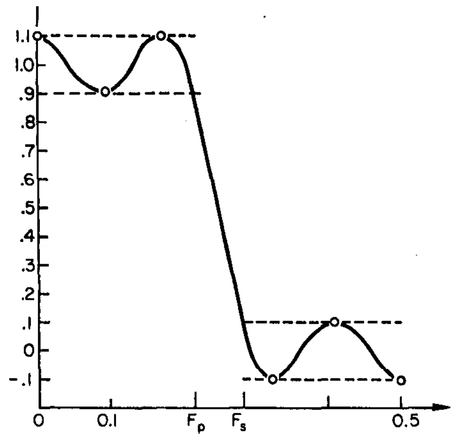

Fig. 2. Extraripple filter showing application of alternation theorem.

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 *arbitrary* 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 proportional 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 [10] to give the following new algorithm which converges quadratically to the best-weighted Chebyshev approximation on the set $\mathfrak{F}$.

The algorithm as indicated in Fig. 3 begins with an initial guess of $n+2$ points and exchanges points until it obtains the $n+2$ points of alternation (extremal frequencies) of the alternation theorem. This is accomplished as follows.

At the ith step of the algorithm there are $n+2$ points on which the error function is forced to have magnitude $|\rho|$ with alternating signs.

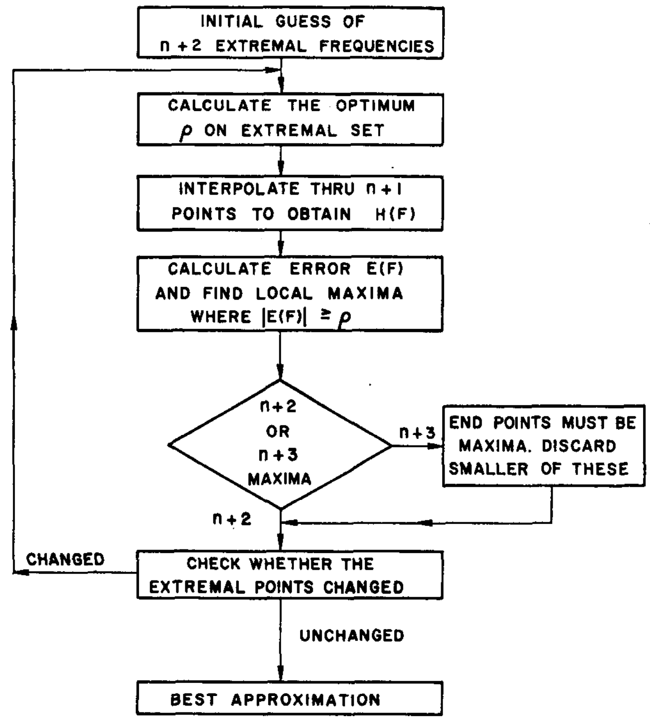

Fig. 3. Block diagram of design algorithm.

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

$$W(F_k)(H(F_k) - D(F_k)) = -(-1)^k\rho, \qquad k = 0, \cdots, n+1 \tag{11}$$

where $H(F)$, $D(F)$, and $W(F)$ are defined in (6), (7), and (9).

Equation (11) may be written matrix form as follows:

$$\begin{bmatrix} 1 & \cos 2\pi F_0 & \cos 4\pi F_0 & \cdots & \cos 2\pi n F_0 & \dfrac{1}{W(F_0)} \\ 1 & \cos 2\pi F_1 & & & & \dfrac{-1}{W(F_1)} \\ \cdot & & & & & \dfrac{1}{W(F_2)} \\ \vdots & & & & & \vdots \\ 1 & \cos 2\pi F_{n+1} & & & & \dfrac{(-1)^{n+1}}{W(F_{n+1})} \end{bmatrix} \begin{bmatrix} d_0 \\ d_1 \\ \cdot \\ \vdots \\ d_n \\ \rho \end{bmatrix} = \begin{bmatrix} D(F_0) \\ D(F_1) \\ \cdot \\ \vdots \\ D(F_{n+1}) \end{bmatrix}. \tag{12}$$

The system (12) is always nonsingular [9], and thus there is a unique solution for ρ and the d_k. However, the direct solution of (12) is time consuming. In fact as pointed out by Temes and Bingham [11], it is the most difficult part of the algorithm. It is more efficient to first calculate ρ analytically as

$$\rho = \frac{a_0 D(F_0) + a_1 D(F_1) + \cdots + a_{n+1} D(F_{n+1})}{a_0/W(F_0) - a_1/W(F_1) + \cdots + (-1)^{n+1} a_{n+1}/W(F_{n+1})}$$

where

$$a_k = (-1)^k \prod_{i=0, i\neq k}^{n+1} \frac{1}{(x_i - x_k)}$$

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

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

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

$$H(F) = \frac{\sum_{k=0}^{n} [\alpha_k/(x - x_k)]c_k}{\sum_{k=0}^{n} [\alpha_k/(x - x_k)]}$$

where

$$\alpha_k = (-1)^k \prod_{i=0, i\neq k}^{n} \frac{1}{(x_i - x_k)}.$$

Note that $H(F)$ will also interpolate to $D(F_{n+1})-(-1)^{n+1} \cdot \rho/W(F_{n+1})$ since it satisfies (11). Then the error function $E(F)=W(F)[D(F)-H(F)]$ is evaluated on $\mathfrak{F}$. 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 $\mathfrak{F}$. If the error function is such that $|E(F)| \leq |\rho|$ for $F \in \mathfrak{F}$, then we have found the best approximation. If $|E(F)| > |\rho|$ for some $F \in \mathfrak{F}$, then a new set of $n+2$ frequencies must be chosen as candidates for the extremal points. The philosophy of the Remes method [10] 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 ith error curve must exhibit $n+2$ or $n+3$ peaks. Thus, sgn $E(F_k)=-$sgn $E(F_{k+1})$, $k=0, 1, \cdots, n+1$, 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. Fig. 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 $P(F)=\sum_{k=1}^{n} a_k \sin 2\pi kF$ has zeros at $F=0$ and $F=\frac{1}{2}$ and at most $n-1$ zeros in the open interval $(0, \frac{1}{2})$. For a given set of points the optimal ρ has been calculated and $H(F)=\sum_{k=0}^{n} g_k \cos 2\pi kF$ fit through $n+1$ points. It is well known that on the set $\mathfrak{F}$, which is the union of disjoint closed intervals, the relative min and max occur at points where $dE/dF=dH/dF=0$ or at the endpoints. $dH/dF=\sum_{k=1}^{n} b_k \sin 2\pi kF$ has at most $n+1$ zeros in the set $\mathfrak{F}$ 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 $\frac{1}{2}$, to be peaks. Since the error curve interpolates $n+2$ points in $\mathfrak{F}$ 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 $\frac{1}{2}$ and delete the point with smaller area to obtain the new set of $n+2$ extremal points.

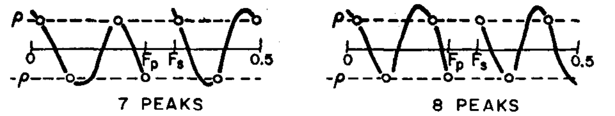

Fig. 4. Two error-curve possibilities for length 11 filter.

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	$n_p=6$, $n_s=9$ $\delta_1=0.0098747$ $\delta_2=0.00098747$	$F_p=0.17265$ $F_s=0.26265$ $K=10$
h_0	-0.0033633	-0.003357
h_1	-0.0060507	-0.0060291
h_2	-0.0008578	-0.00082454
h_3	0.0097200	0.0097418
h_4	0.0083195	0.0083190
h_5	-0.011326	-0.011334
h_6	-0.021879	-0.021844
h_7	0.0047474	0.0047995
h_8	0.039863	0.039882
h_9	0.017794	0.017773
h_{10}	-0.058357	-0.058361
h_{11}	-0.073671	-0.073620
h_{12}	0.072324	0.072385
h_{13}	0.306575	0.306583
h_{14}	0.422471	0.422443

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.

III. RESULTS

The design algorithm is implemented by specifying the filter order n_f, 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. Fig. 5 shows experimental results of the relationships between these parameters.

Fig. 5(a) shows δ_2 versus F_p for a length 29 filter. The local minima at 0.13765 and 0.17265 correspond to Hofstetter designs (Table I). Between the maxima at 0.1195 and the minima at 0.13765 there is a difference of 3.32 dB, indi-

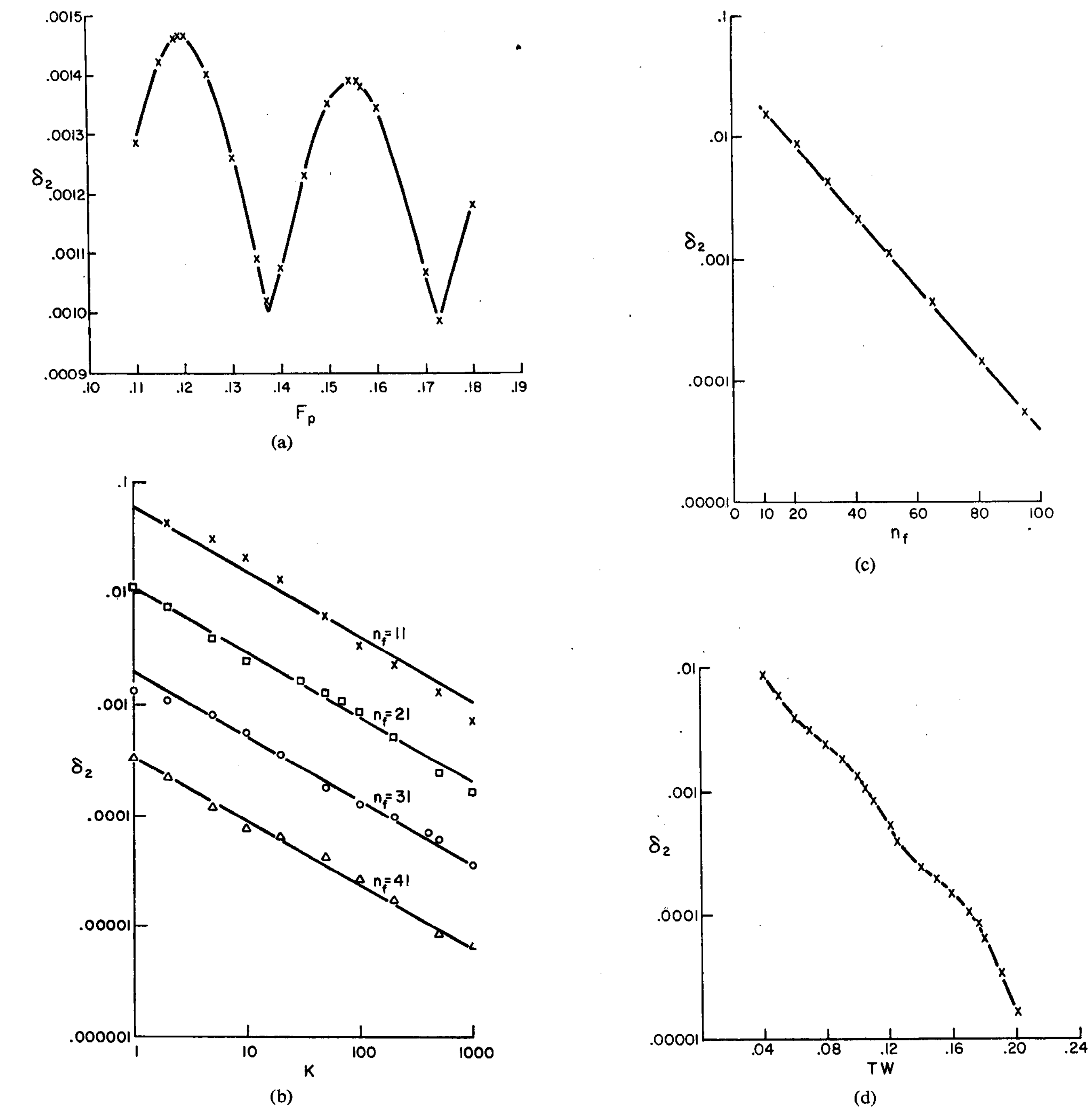

Fig. 5. (a) Influence of passband edge F_p on stopband deviation δ_2 for fixed transition width $TW=0.09$ and weighting factor $K=10$ with length 29 filter. (b) Control of stopband deviation δ_2 exerted by weighting factor K for filter lengths $n_f=11$, 21, 31, and 41, with $F_p=0.2$ and $F_s=0.3$. (c)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

cating what is to be gained by an optimal location of the transition region.

Fig. 5(b) shows the control over δ_2 exerted by the weighting factor. The *a priori* choice of K is 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 $(n_f-1)/2$ extraripple filters by varying K.

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

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

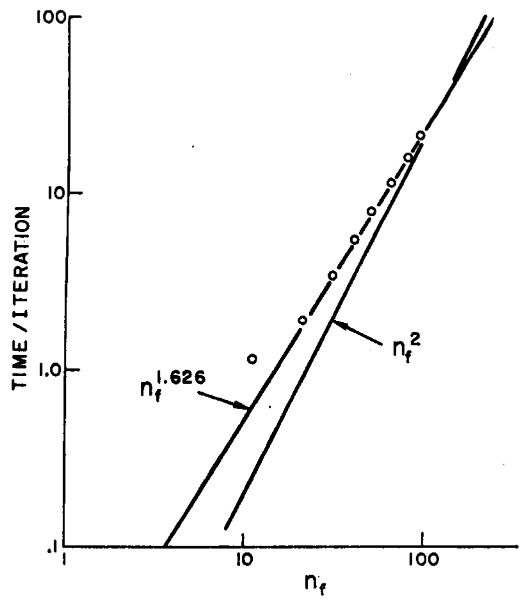

Fig. 6. Computation time as a function of filter length.

would have to be concentrated on 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 $a_k = \prod_{i \neq k} [1/(x_i - x_k)]$ 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 4.

IV. Conclusion

A technique for the weighted Chebyshev design of nonrecursive 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.

References

[1] L. R. Rabiner, "Techniques for designing finite-duration impulse-response digital filters," *IEEE Trans. Commun. Technol.*, vol. COM-19, pp. 188–195, Apr. 1971.

[2] H. D. Helms, "Digital filters with equiripple or minimax responses," *IEEE Trans. Audio Electroacoust.*, vol. AU-19, pp. 87–93, Mar. 1971.

[3] J. F. Kaiser and F. F. Kuo, *System Analysis by Digital Computer.* New York: Wiley, 1966, ch. 7.

[4] 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.

[5] B. Gold and K. L. Jordan, Jr., "A direct search procedure for designing finite duration impulse response filters," *IEEE Trans. Audio Electroacoust.*, vol. AU-17, pp. 33–36, Mar. 1969.

[6] D. W. Tufts and J. T. Francis, "Designing digital low-pass filters—comparison of some methods and criteria," *IEEE Trans. Audio Electroacoust.*, vol. AU-18, pp. 487–494, Dec. 1970.

[7] O. Herrmann, "Design of nonrecursive digital filters with linear phase," *Electron. Lett.*, pp. 328–329, 1970.

[8] E. Hofstetter, A. V. Oppenheim, and J. Siegel, "A new technique for the design of non-recursive digital filters," presented at the 5th Annu. Princeton Conf. Information Sciences and Systems, Mar. 1971.

[9] E. W. Cheney, *Introduction to Approximation Theory.* New York: McGraw-Hill, 1966, pp. 72–100.

[10] E. Ya. Remes, "General computational methods of Tchebycheff approximation," Kiev, 1957 (Atomic Energy Commission Translation 4491, pp. 1–85).

[11] G. C. Temes and J. A. C. Bingham, "Iterative Chebyshev approximation technique for network synthesis," *IEEE Trans. Circuit Theory*, vol. CT-14, pp. 31–37, Mar. 1967.

Correspondence

A Note on Multivariable Positive Real Functions

N. K. BOSE

In a recent article [1] Vowels presents some theorems and tables, and correctly asserts that "greater use should be made of the close relation between positive real functions and Hurwitz polynomials." However, he fails to mention that Reza and Bose [2] presented and published a paper about four years ago which exhibited and exploited to advantage some of the links between realizability theory and the theory of equations. It was shown that known results in the theory of polynomials of a single variable could be proved relatively easily by primarily using network-theoretic concepts. Feasibility of obtaining new results like the generation of various Hurwitz polynomials from a given one, by constructing suitable network models, was made evident.

Some of the "deduced positive real functions," and the obtained "Hurwitz positive pairs" in Vowels' paper, and many others can be proved using the ideas in Reza and Bose's paper [2]. Consequently, the

Manuscript received June 15, 1971; revised October 5, 1971.
The author was with the Department of Electrical Engineering, American University of Beirut, Beirut, Lebanon. He is now with the Department of Electrical Engineering, University of Pittsburgh, Pittsburgh, Pa. 15213.