Exchange Algorithms that Complement the Parks-McClellan Algorithm for Linear-Phase FIR Filter Design *

Ivan W. Selesnick and C. Sidney Burrus

The Computer and Information Technology Institute and
Electrical and Computer Engineering Department
Rice University, Houston, TX 77251-1892
(713) 285-5484, selesi@ece.rice.edu, csb@ece.rice.edu

October 2, 1995

Abstract

This paper describes an exchange algorithm for the frequency domain design of linear-phase FIR equiripple filters where the Chebyshev error in each band is specified. The algorithm is a hybrid of the algorithm of Hofstetter, Oppenheim and Siegel and the Parks-McClellan algorithm. The paper also describes a modification of the Parks-McClellan algorithm where either the passband or the stopband ripple size is specified and the other is minimized.

Contact Address:
Ivan W. Selesnick
Electrical and Computer Engineering - MS 366
Rice University,
Houston, TX 77251-1892
Phone: (713) 527-8101 x3508
FAX: (713) 524-5237
email: selesi@ece.rice.edu

*This research has been supported by DARPA under an NDSEG fellowship, by NSF grant MIP-9316588, and by BNR.

List of Figures

List of Tables

1 Introduction

The Chebyshev norm is widely used for the design of linear-phase FIR filters for two reasons. (i) For certain applications it is a meaningful error criterion as shown by Weisburn, Parks and Shenoy [25] and (ii) there is an excellent program, the Parks-McClellan (PM) program [14, 16, 21], that designs optimal filters according to this norm. Recall that in this approach to the design of digital filters, the band edges are specified and the weighted Chebyshev error is minimized. In this short paper we revisit equiripple filter design and describe a complement to the PM program. In the program we propose, the weighted Chebyshev error is specified and the width of the transition region is variable.

For example, consider the usual ideal lowpass filter. Figure 1 shows a typical equiripple frequency response. The passband and stopband edges are denoted by ω_p and ω_s. The Chebyshev errors in the passband and stopband are denoted by δ_p and δ_s. The figure also shows the half-magnitude frequency, ω_o. The filter in fig. 1 was designed to posses a specified half-magnitude frequency of $\omega_o = 0.4\pi$ with specified Chebyshev errors of $\delta_p = \delta_s = 0.05$. The passband and stopband edges are 'induced' by these specifications.

While the PM program allows the filter designer to specify ω_p, ω_s and the ratio δ_p/δ_s, it does not allow the user to specify both δ_p and δ_s directly. Although the user can indirectly control δ_p and δ_s with the PM program by iteratively adjusting the band edges and the ratio δ_p/δ_s [17, 19], the exchange algorithm below obtains the specified δ_p and δ_s directly.

The algorithm described below is similar to the algorithm of Hofstetter, Oppenheim and Siegel [12, 13, 20]. Although their algorithm predates the PM program and produces equiripple filters with specified δ_p and δ_s, it is not widely used because it permits limited control over the location of the band edges and only produces extra-ripple filters. Moreover, the location of the band edges can be controlled only indirectly in the Hofstetter algorithm: only the number of extremal frequencies in each band can be specified. Their algorithm is interesting, however, because extra-ripple filters have a minimum transition width property [20, p143][13, 21]. More specifically, for a fixed filter length and fixed δ_p and δ_s, extra-ripple filters locally minimize the transition band width as a function of, say, the passband edge. The design formulation leading to these extra-ripple filters was originally described by Herrmann and Schüßler [9, 11] and is also described in [20, 21]. The algorithm of Hofstetter et al can also be used for the design of maximum-ripple multiband filters having a specified ripple size in each band.

In [23] Shpak and Antoniou present an interesting modification to the PM program in which the Chebyshev error in respective bands are not constrained by a ratio. In this way they are able to obtain extra-ripple filters without sacrificing the ability to specify the band edges. This technique can be used to prevent transition region anomalies that sometimes occur in the design of optimal Chebyshev multiband filters [18]. By maintaining the ability to specify band edges in the design of extra-ripple filters, however, the Shpak-Antoniou algorithm gives less control over the weighted Chebyshev error. Clearly, for a fixed filter length, there is a tradeoff between the ability to specify the weighted Chebyshev error and the ability to specify the band edges.

Another approach to the filter design problem uses linear programming [8, 24]. Steiglitz, Parks and Kaiser have presented a very general and flexible program [24] that meets Chebyshev and other constraints. In this paper, we discuss multiple exchange Remez-type algorithms which are more efficient than linear programming methods.

The algorithms described below produce lowpass and bandpass linear-phase FIR filters having a specified Chebyshev error in each band and a single transition region frequency, such as the half-magnitude frequency, to control the location of the transition region. They are hybrids of the algorithm of Hofstetter, Oppenheim and Siegel [12, 13] and the PM algorithm. Like those algorithms, it employs a reference set of frequencies. On each iteration (*i*) an interpolation problem is solved and (*ii*) the reference set is updated. The efficient computational techniques used for the PM program [1, 2, 3, 4, 6, 23] can also be used for the algorithms below.

2 Equiripple Filter Design

The frequency response of a linear-phase FIR filter is given by the discrete-time Fourier transform of its impulse response and can be written as $H(\omega) = A(\omega)e^{-jM\omega}$ where $A(\omega)$ is a real valued periodic function of ω called the amplitude and $M = (N-1)/2$ for length-N filters [15].

We first describe a version of the standard PM program for lowpass filter design in which either δ_p or δ_s is specified and the other is minimized. We then describe the exchange algorithm for lowpass and bandpass filter design in which the Chebyshev error in each band is specified.

2.1 The PM Program with an Affine Relation between δ_p and δ_s

The usual PM program can be modified so that it achieves a specified Chebyshev error in one band and minimizes the Chebyshev error in the other. This can be achieved by imposing an affine relationship between δ_p and δ_s, where δ_p and δ_s denote the Chebyshev errors of the *realized* frequency response amplitude. Recall that by the standard PM algorithm for lowpass filter design, the user specifies a linear relationship between δ_p and δ_s: the user specifies N, ω_p, ω_s and K, and obtains a filter satisfying $\delta_p = K\delta_s$. However, the PM algorithm can be modified as described here so that the user specifies N, ω_p, ω_s, K_p, K_s, η_p and η_s, and obtains a filter satisfying the following affine relationship between δ_p and δ_s:

$$\delta_p = K_p\delta + \eta_p \tag{1}$$

$$\delta_s = K_s\delta + \eta_s. \tag{2}$$

(So that the problem is well posed, the parameters K_p, K_s, η_p and η_s supplied by the user must all be nonnegative and satisfy the inequalities: $K_p + \eta_p > 0$, $K_s + \eta_s > 0$ and $K_p + K_s > 0$.) The modified PM program minimizes δ. When η_p and η_s are both taken to be 0 this becomes the usual linear relationship permitted by the PM program. However, if $K_s = \eta_p = 0$, then the stopband ripple size δ_s of the resulting equiripple filter has the specified value η_s and the passband ripple size δ_p is minimized.

The first modification that needs to be made to the usual PM program is the interpolation step. Given a reference set of $M+2$ frequencies that includes ω_p and ω_s, let $\omega_1, \ldots, \omega_q = \omega_p$ denote those in the passband and let $\omega_s = \omega_{q+1}, \ldots, \omega_{M+2}$ denote those in the stopband, listed, in each case, in ascending order. The linear system of equations to be solved on each iteration is given by eq (1), eq (2), and by

$$
\begin{aligned}
A(\omega_i) &= 1 + (-1)^{i+c}\delta_p & \text{for } 1 \le i \le q \\
A(\omega_i) &= (-1)^{i+c}\delta_s & \text{for } q+1 \le i \le M+2
\end{aligned}
\tag{3}
$$

where c is chosen to equal 0 or 1, whichever yields the equation $A(\omega_p) = 1 - \delta_p$. The system of equations given by (1, 2, 3) is linear in δ, δ_p, δ_s and the filter coefficients. As in the PM algorithm, this system can be solved efficiently using interpolation formulas.

Consider, for example, the design of a length-7 filter with $\omega_p = 0.2\pi$ and $\omega_s = 0.5\pi$. Here $M = 3$ and $A(\omega)$ can be written as $A(\omega) = \sum_{k=0}^{3} a_k \cos k\omega$. Figure 2 shows a typical frequency response amplitude during the course of the algorithm. In this case, eqs (1, 2, 3) can be written as

$$
\begin{bmatrix}
1 & \cos\omega_1 & \cos 2\omega_1 & \cos 3\omega_1 & -1 & 0 & 0 \\
1 & \cos\omega_2 & \cos 2\omega_2 & \cos 3\omega_2 & 1 & 0 & 0 \\
1 & \cos\omega_3 & \cos 2\omega_3 & \cos 3\omega_3 & 0 & -1 & 0 \\
1 & \cos\omega_4 & \cos 2\omega_4 & \cos 3\omega_4 & 0 & 1 & 0 \\
1 & \cos\omega_5 & \cos 2\omega_5 & \cos 3\omega_5 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & -K_p \\
0 & 0 & 0 & 0 & 0 & 1 & -K_s
\end{bmatrix}
\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ \delta_p \\ \delta_s \\ \delta \end{bmatrix}
=
\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ \eta_p \\ \eta_s \end{bmatrix}
\tag{4}
$$

where $\omega_1 = 0$, $\omega_p = \omega_2 = 0.2\pi$, $\omega_s = \omega_3 = 0.5\pi$, $\omega_4 = 0.75\pi$, $\omega_5 = \pi$ are indicated by circular marks in fig. 2. Comparing system (4) to the corresponding system for the usual PM algorithm, we note that there are two additional variables and two additional equations.

It is important to note that either δ_p or δ_s given by the solution to (1, 2, 3) may be negative. This generally occurs, if at all, during the early iterations of the algorithm. When it does occur, however, it is important to modify the interpolation step in a simple way to ensure convergence. If δ_s given by eqs (1, 2, 3) is negative on some iteration, then it is necessary to repeat the interpolation step for that iteration using different interpolation equations. The equations in this case are given by

$$
\begin{aligned}
A(\omega_i) &= 1 + (-1)^{i+c}\delta_p & \text{for } 1 \le i \le q \\
A(\omega_i) &= 0 & \text{for } q+1 \le i \le M+2 \\
\delta_s &= 0. &
\end{aligned}
\tag{5}
$$

In the length-7 example above, these equations can be written as

$$
\begin{bmatrix}
1 & \cos\omega_1 & \cos 2\omega_1 & \cos 3\omega_1 & -1 \\
1 & \cos\omega_2 & \cos 2\omega_2 & \cos 3\omega_2 & 1 \\
1 & \cos\omega_3 & \cos 2\omega_3 & \cos 3\omega_3 & 0 \\
1 & \cos\omega_4 & \cos 2\omega_4 & \cos 3\omega_4 & 0 \\
1 & \cos\omega_5 & \cos 2\omega_5 & \cos 3\omega_5 & 0
\end{bmatrix}
\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ \delta_p \end{bmatrix}
=
\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}
\tag{6}
$$

Similarly, if $\delta_p < 0$ on some iteration, then the interpolation step for that iteration must be repeated with the equations

$$
\begin{aligned}
A(\omega_i) &= 1 & \text{for } 1 \le i \le q \\
A(\omega_i) &= (-1)^{i+c}\delta_s & \text{for } q+1 \le i \le M+2 \\
\delta_p &= 0. &
\end{aligned}
\tag{7}
$$

On the following iteration, after the reference set is updated, the eqs (1, 2, 3) are again used. We also note that δ_p and δ_s given by the solution to eqs (1, 2, 3) can not both be negative, because if they were, then $A(\omega)$ would have more extrema than is possible for a degree M cosine polynomial.

The procedure to update the reference set from one iteration to the next is the multiple exchange of the PM algorithm: Let S be the set obtained by appending ω_p and ω_s to the set of extrema of $A(\omega)$ in $[0, \pi]$. S will have either $M+2$ or $M+3$ frequencies and will include both 0 and π. If S has $M+2$ frequencies, then take the new reference set to be S. If S has $M+3$ frequencies, then remove either 0 or π from S according to the following rule: If $\omega = 0$ is a local maximum of $A(\omega)$ then let $\alpha = 1$, otherwise set $\alpha = -1$. If $\omega = \pi$ is a local maximum of $A(\omega)$ then let $\beta = 1$, otherwise set $\beta = -1$. If

$$(A(0) - 1)\alpha - \delta_p < A(\pi)\beta - \delta_s \tag{8}$$

then remove 0 from S, otherwise remove π from S, and take the new reference set to be the resulting set S. The expressions on each side of the inequality indicate the amount by which the error exceeds its intended value. α and β must be chosen appropriately because both the magnitude and the sign of this value is important: negative values appear in the design of filters possessing a scaled extra-ripple. The rule states that the frequency to be retained in S is the one at which the error exceeds its intended value the most. The reference set, in the example, is updated by updating only ω_4. Its new value is indicated by the x mark in fig. 2.

A flowchart for this algorithm is shown if fig. 3. This modification of the PM algorithm is easily incorporated and permits (i) the specification of ω_p and ω_s and (ii) the affine constraint on δ_p and δ_s. It is useful because by taking $K_s = \eta_p = 0$, or $K_p = \eta_s = 0$, the filter obtained by this algorithm achieves a specified Chebyshev error in one band and minimizes the Chebyshev error in the other. The algorithm works for other choices of K_p, K_s, η_p, η_s, but the meaning of arbitrary values for these parameters is unclear. The general affine constraint is a convenient way to solve the problem, without requiring special cases.

The following example illustrates the situation in which δ_p is negative on some iteration. The parameters are chosen to be $N = 19$, $\omega_p = 0.4\pi$, $\omega_s = 0.5\pi$, $K_p = 1$, $Ks = 0$, $\eta_p = 0$ and $\eta_s = 0.05$. In this case $M = 9$. When the reference set is initialized to be 4 equally spaced frequencies in the passband and 7 equally spaced frequencies in the stopband, the solution to eqs (1, 2, 3) gives $\delta_p = -1.9619$, $\delta_s = 0.05$ and the amplitude $A(\omega)$ shown in fig. 4(a). Because $\delta_p < 0$, the algorithm we describe requires that the iteration be repeated using eq (7) and the same reference set. The solution to eq (7) gives $\delta_p = 0$, $\delta_s = 0.0096$ and the amplitude shown in fig. 4(b). The reference set is then updated and on the next interpolation step, δ_p is positive, and remains positive for the duration of the algorithm. The amplitudes associated with this and next iteration are shown in figs. 4(c) and 4(d). The algorithm converges in only a few more iterations.

2.2 A New Equiripple Lowpass Filter Design Algorithm: Specified δ_p, δ_s and ω_o

In order to exactly achieve specified values for δ_p and δ_s with the PM program, it is necessary to iteratively adjust the parameters ω_p, ω_s, and the ratio δ_p/δ_s. We propose an algorithm for the design of equiripple lowpass filters that allows (i) the explicit specification of δ_p and δ_s and (ii) the specification of the half-magnitude frequency. The band edges ω_p and ω_s can not be explicitly specified – they are 'induced' by the specified values of δ_p, δ_s and ω_o as in the algorithm of Hofstetter et al. [12]. As above, on each iteration, the filter interpolating the appropriate values over the reference set is computed and the reference set is updated.

As in the PM program, extremal frequencies may migrate from one band to another during the course of the algorithm. Therefore the initialization of the reference set is not critical for convergence, but does affect the speed of convergence. The reference set here, however, does not contain two band edges as in the PM program, instead, it contains the half-magnitude frequency, ω_o. Therefore, the reference set contains $M+1$ frequencies, not $M+2$ as in the PM program. The circular marks in fig 1 indicate the reference set frequencies upon convergence for a length 21 filter. The resulting filter satisfies the alternation property for the correct choice of band edges, so it could have been designed using the PM program if the band edges had been known in advance.

The algorithm proceeds by computing the filter that alternately interpolates $1+\delta_p$, $1-\delta_p$ over the reference set frequencies in the passband, alternately interpolates δ_s, $-\delta_s$ over the reference set frequencies in the stopband, and interpolates 0.5 at ω_o. As in the PM program, interpolation formulas can be used to find the filter efficiently. Note that, because δ_p and δ_s have been explicitly specified, δ does not have to be computed at each iteration.

Suppose $\omega_1, \ldots, \omega_{M+1}$, listed in increasing order, is a reference set of $M+1$ frequencies in $[0, \pi]$ that includes ω_o. Let $\omega_1, \ldots, \omega_{q-1}$ denote those in the passband (to the left of ω_o) and let $\omega_{q+1}, \ldots, \omega_{M+1}$ denote those in the stopband (to the right of ω_o). The linear system of equations to be solved on each iteration is given by

$$\begin{aligned} &A(\omega_i) = 1 + (-1)^{i+c}\delta_p \quad \text{for } 1 \le i \le q-1 \\ &A(\omega_o) = 0.5 \\ &A(\omega_i) = (-1)^{i+c+1}\delta_s \qquad \text{for } q+1 \le i \le M+1 \end{aligned} \tag{9}$$

where c is chosen to equal 0 or 1, whichever yields the equation $A(\omega_{q-1}) = 1 + \delta_p$.

Consider, for example, the design of a length 9 filter with $\omega_o = 0.4\pi$, $\delta_p = \delta_s = 0.1$. Figure 5 shows a typical amplitude response before convergence is attained. In this case, the equations (9) can be written as

$$\begin{bmatrix} 1 & \cos\omega_1 & \cos 2\omega_1 & \cos 3\omega_1 & \cos 4\omega_1 \\ 1 & \cos\omega_2 & \cos 2\omega_2 & \cos 3\omega_2 & \cos 4\omega_2 \\ 1 & \cos\omega_3 & \cos 2\omega_3 & \cos 3\omega_3 & \cos 4\omega_3 \\ 1 & \cos\omega_4 & \cos 2\omega_4 & \cos 3\omega_4 & \cos 4\omega_4 \\ 1 & \cos\omega_5 & \cos 2\omega_5 & \cos 3\omega_5 & \cos 4\omega_5 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} 1-\delta_p \\ 1+\delta_p \\ 0.5 \\ -\delta_p \\ \delta_p \end{bmatrix} \tag{10}$$

where $\omega_1 = 0$, $\omega_2 = 0.1\pi$, $\omega_o = \omega_3 = 0.4\pi$, $\omega_4 = 0.7\pi$, $\omega_5 = \pi$ are indicated by circular marks in fig. 5.

The procedure to update the reference set from one iteration to the next is similar to the multiple exchange of the PM algorithm: Let S be the set obtained by appending ω_o to the set of extrema of $A(\omega)$ in $[0, \pi]$. S will have either $M+1$ or $M+2$ frequencies and will include both 0 and π. If S has $M+1$ frequencies, then take the new reference set to be S. If S has $M+2$ frequencies, then remove either 0 or π from S according to the following rules:

1. If $A(\omega)$ has no extrema in the open interval $(0, \omega_o)$, then remove 0 from S.

2. If $A(\omega)$ has no extrema in the open interval (ω_o, π), then remove π from S.

3. Otherwise, let ω_a be the extrema of $A(\omega)$ in $(0, \omega_o)$ closest to 0, and let ω_b be the extrema of $A(\omega)$ in (ω, π) closest to π. If

$$\delta_s |A(0) - A(\omega_a)| < \delta_p |A(\pi) - A(\omega_b)| \tag{11}$$

then remove 0 from S, otherwise remove π from S.

Take the new reference set to be the resulting set S. The reference set, in fig. 5 for example, is updated by updating ω_2, ω_4 and ω_5. Their new locations are indicated by the x marks in fig. 5. It should be noted that case(1) and case(2) only occur when ω_o is taken to be near 0 or π relative to the filter length (in these cases, the reference set upon convergence contains no frequencies in one of the bands).

Note that any transition region frequency can be fixed. Instead of the half-magnitude frequency, the half-power frequency, the passband edge, or stopband edge can be fixed by respectively imposing $A(\omega_o) = 1/\sqrt{2}$, $A(\omega_o) = 1 - \delta_p$, or $A(\omega_o) = \delta_s$. It should also be noted that if the half-magnitude is taken to be too close to either 0 or π relative to the filter length, then there will exist no filter with the specified Chebyshev error in each band. Either the passband or stopband will be too narrow.

When the specified ripple sizes are achievable, this algorithm produces exactly the same lowpass filters as does the PM program, however, it allows one to specify a different set of parameters in the design process.

2.3 Bandpass Filter Design

The design of multiband filters achieving a specified Chebyshev error with specified half-magnitude frequencies requires more care than the design of lowpass filters with this approach. There a three reasons for this. (*i*) There is generally more than one equiripple filter satisfying these constraints. (*ii*) The procedure for updating the reference set of frequencies is less obvious because the optimal filter may have scaled extra ripples at frequencies other than 0 and π. (*iii*) The transition region of a multiband filter designed by the PM program may contain large undesirable peaks [18]. Despite these aspects of the multiband case, the algorithm below remains simple, robust, and rapid as long as the specified Chebyshev error is not taken too small relative to the filter length.

We consider the design of bandpass filters and denote the Chebyshev errors of the first stopband, the passband, and the second stopband by δ_1, δ_2 and δ_3 respectively. The half-magnitude frequencies are denoted by ω_a and ω_b. When the band edges are not explicitly specified, the non-uniqueness of the bandpass filter achieving a specified Chebyshev error with specified half-magnitude frequencies is easily ascertained. The specifications can be summarized by 5 values: δ_1, δ_2, δ_3, ω_a and ω_b. However, the PM algorithm requires 6 values for bandpass filters: 4 band edges and 2 ratios, δ_2/δ_1 and δ_3/δ_1. Therefore, in a complement to the PM program, we require an additional constraint.

We have chosen to require that the derivative of $A(\omega)$ at the half-magnitude frequencies are equal in magnitude and opposite in sign. We chose this constraint because (*i*) in a sense, it weights the widths of the transition regions equally, because the width of the transition region is related to the slope of $A(\omega)$ at the corresponding half-magnitude frequency, and, (*ii*) it can be easily incorporated into a simple exchange algorithm.

The algorithm for producing equiripple bandpass filters is similar to the algorithm above for lowpass design. Excluding the two half-magnitude frequencies, the reference set for the bandpass case contains $M-2$ frequencies, and is updated by locating the local extrema of the new frequency response amplitude $A(\omega)$. The interpolation step consists of finding the filter that alternately interpolates $1+\delta_i$, $1-\delta_i$ over the reference set frequencies in band i, interpolates 0.5 at ω_a and at ω_b, and for which $A'(\omega_a) = -A'(\omega_b)$. This filter can be found by solving a system of linear equations or by modifying the usual interpolation formulas.

It follows from the interpolation step that there will be at least $M-2$ local extrema of $A(\omega)$, however, there may be as many as $M+1$. Because the reference set must contain $M-2$ extremal frequencies, it will therefore be necessary to exclude 0, 1, 2 or 3 local minima and maxima when updating the reference set. The rule we use for updating the reference set is most easily stated by describing which local extrema are *not* included. Suppose $\omega_1, \ldots, \omega_L$ are the local extrema of $A(\omega)$ listed in order.

1. To exclude 1 local extremum ($L = M-1$), use the same update rule used for the lowpass case.

2. To exclude 2 local extrema ($L = M$), find the index i that minimizes

$$(E(\omega_i) - E(\omega_{i+1}))(-1)^{i+s} \tag{12}$$

   where $s=1$ if ω_1 is a local maxima and $s=0$ if ω_1 is a local minima. $E(\omega)$ denotes the error function: $E(\omega) = (A(\omega) - D(\omega))/\delta(\omega)$ where $D(\omega)$ is 1 for $\omega < \omega_o$, 0 for $\omega > \omega_o$ and where $\delta(\omega)$ is δ_p for $\omega < \omega_o$, δ_s for $\omega > \omega_o$. If $1 < i < M-2$, then exclude ω_i and ω_{i+1} from the reference set. If $i=1$ or $i=L$, then exclude ω_i and use the procedure above for excluding 1 local extremum.

3. To exclude 3 local extrema ($L = M+1$), use the procedure for excluding 1 extremum, followed by the procedure for excluding 2 extrema.

By following this simple procedure for updating the reference set the algorithm rapidly converges and, like the PM algorithm, is capable of producing equiripple filters with extra ripples at frequencies other than 0 and π. Figure 6 shows a bandpass frequency response with three scaled extra ripples. The reference set frequencies upon the convergence of the algorithm are indicated with circular marks. Again, although this filter was not obtained by the PM program, it could have been if the resulting band edges were known in advance. That is, the filter in fig 6 is an optimal Chebyshev filter for the correct choice of band edges.

For some specifications, the filters produced by this algorithm for bandpass filter design are not optimal Chebyshev filters for any choice of band edges. Specifically, this algorithm can produce filters possessing a pair of adjacent scaled extra ripples that straddle a half-magnitude frequency. The filter in fig 3, for example, was obtained with this algorithm. Although the alternation property is satisfied on the extremal reference set frequencies, it is not an optimal Chebyshev filter for any choice of band edges because the induced band edges can not be included in the reference set without destroying the alternation property. Nevertheless, this filter does achieve the specified Chebyshev error and has narrow transition bands of approximately equal width.

When the specified Chebyshev error is taken to be very small relative to the filter length this algorithm may occasionally produce filters with undesirable transition region behavior or may fail to converge. This is due to the necessarily wide transition regions associated with very small ripple sizes. When the transition regions are wide, the half-magnitude and derivative constraints become inappropriate since they no longer accurately reflect the behavior of the frequency response throughout the transition region. However, it should be noted that optimal Chebyshev multiband filters having very wide transition regions may also possess undesirable transition region behavior.

Indeed, the behavior of the frequency response of optimal Chebyshev multiband filters can be quite different than that of two-band filters. In [18] Rabiner, Kaiser and Schafer give three strategies for avoiding nonmonotonic transition region behavior: (*i*) modify the stopband edge frequencies, (*ii*) modify the error weighting function, and (*iii*) design maximal ripple filters only. Shpak and Antoniou [23] address the occurrence of transition region ripples by employing extra δ variables. By doing this they are able to obtain extra-ripple filters and can avoid some of the undesirable behaviors of multiband equiripple filters while maintaining specified band edges. The method described in this paper, however, takes a different direction. Instead of introducing extra δ variables, we give up the explicit control over the band edges, employ half-magnitude frequencies, and explicitly control the Chebyshev error in each band.

It should also be noted that all the exchange algorithms discussed in this paper can be adopted for the design of minimum phase FIR filters. Grenez describes a simple modification of the PM program for constrained Chebyshev approximation that can be used to design linear-phase filters with nonnegative frequency response amplitudes [7]. If in each iteration of the exchange algorithm, the stopband interpolation condition $A(\omega_i) = -\delta$ is replaced by $A(\omega_i) = 0$, then the resulting frequency response amplitude will be nonnegative. The FIR filter can then be spectrally factored to obtain a minimum phase filter. This technique is especially useful when the stopband ripple sizes of a multiband filter are unequal. For multiband filters for which the stopband ripple sizes are equal and for two-band filters, the classical technique of raising the amplitude and spectrally factoring the filter can be employed [10].

3 Conclusion

Optimal Chebyshev linear-phase FIR filters are usually found by fixing the filter length N and the band edges and by minimizing the weighted Chebyshev error. Another approach is to fix N, δ_p, δ_s and a single transition region frequency and to adjust the transition width. The same approach can be applied to the design of bandpass filters that achieve a specified Chebyshev error in each band and have transition regions of approximately equal width.

Table 1 classifies four approaches to the design of equiripple filters. The approaches under 'Nonextra-ripple' produce filters that may or may not possess extra-ripples, depending on the specifications. The approaches under 'Extra-ripple' are able to produce filters that are constrained to possess extra-ripples. (Recall, however, that the Shpak-Antoniou algorithm is a generalization of the PM algorithm and a *variable* number of extra ripples can be specified.) This table clarifies the relationship among previously reported exchange algorithms for equiripple linear-phase filter design and the way in which the algorithms described in this paper relate to them.

Matlab programs are available from the authors and electronically on the World Wide Web at URL *http://www-dsp.rice.edu*.

4 Acknowledgment

We would like to acknowledge our discussions with professor Hans W. Schüßler and Markus Lang and to thank them for their helpful comments.

References

[1] J. W. Adams and A. N. Willson, Jr. On the fast design of high-order FIR digital filters. *IEEE Trans. on Circuits and Systems*, 32(9):958–960, September 1985.

[2] A. Antoniou. Accelerated procedure for the design of equiripple nonrecursive digital filters. *IEE Proceedings, Part G*, 129(1):1–10, February 1982.

[3] A. Antoniou. New improved method for the design of weighted-Chebyshev, nonrecursive, digital filters. *IEEE Trans. on Circuits and Systems*, 30(10):740–750, October 1983.

[4] F. Bonzanigo. Some improvements to the design programs for equiripple FIR filters. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, pages 274–277, 1982.

[5] IEEE DSP Committee, editor. *Selected Papers In Digital Signal Processing,II.* IEEE Press, 1976.

[6] S. Ebert and U. Heute. Accelerated design of linear or minimum phase FIR filters with a Chebyshev magnitude response. *IEE Proceedings, Part G*, 130(6):267–270, December 1983.

[7] F. Grenez. Design of linear or minimum-phase FIR filters by constrained Chebyshev approximation. *EURASIP Signal Processing*, 5:325–332, May 1983.

[8] H. D. Helms. Digital filters with equiripple or minimax responses. *IEEE Trans. Audio Electroacoust.*, 19:87–94, March 1971. Also in [22].

[9] O. Herrmann. Design of nonrecursive filters with linear phase. *Electron. Lett.*, 6(11):328–329, May 28 1970. Also in [22].

[10] O. Herrmann and H. W. Schuessler. Design of nonrecursive filters with minimum phase. *Electron. Lett.*, 6(11):329–330, May 28 1970. Also in [22].

[11] O. Herrmann and H. W. Schuessler. On the design of nonrecursive digital filters. *IEEE Arden House Workshop*, January 1970.

[12] E. Hofstetter, A. Oppenheim, and J. Siegel. A new technique for the design of nonrecursive digital filters. In *Proc. of the Fifth Annual Princeton Conference on Information Sciences and Systems*, pages 64–72, October 1971. Also in [22].

[13] E. Hofstetter, A. Oppenheim, and J. Siegel. On optimum nonrecursive digital filters. In *Proc. of the Ninth Allerton Conference on Circuits and System Theory*, pages 789–798, 1971. Also in [22].

[14] 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.*, 21:506–526, December 1973. Also in [5].

[15] T. W. Parks and C. S. Burrus. *Digital Filter Design.* John Wiley and Sons, 1987.

[16] T. W. Parks and J. H. McClellan. Chebyshev approximation for nonrecursive digital filters with linear phase. *IEEE Trans. on Circuit Theory*, 19:189–94, March 1972.

[17] T. W. Parks and J. H. McClellan. On the transition region width of finite impulse-response digital filters. *IEEE Trans. Audio Electroacoust.*, 21:1–4, February 1973.

[18] L. Rabiner, J. Kaiser, and R. Schafer. Some considerations in the design of multiband finite-impulse-response digital filters. *IEEE Trans. on Acoust., Speech, Signal Proc.*, 22(6):462–472, December 1974.

[19] L. R. Rabiner. Approximate design relationships for lowpass FIR digital filters. *IEEE Trans. Audio Electroacoust.*, 21:456–460, October 1973. Also in [5].

[20] L. R. Rabiner and B. Gold. *Theory and Application of Digital Signal Processing.* Prentice Hall, 1975.

[21] L. R. Rabiner, J. H. McClellan, and T. W. Parks. FIR digital filter design techniques using weighted Chebyshev approximation. *Proc. IEEE*, 63:595–610, April 1975. Also in [5].

[22] L. R. Rabiner and C. M. Rader, editors. *Digital Signal Processing.* IEEE Press, 1972.

[23] D. J. Shpak and A. Antoniou. A generalized Remez method for the design of FIR digital filters. *IEEE Trans. on Circuits and Systems*, 37(2):161–174, February 1990.

[24] K. Steiglitz, T. W. Parks, and J. F. Kaiser. METEOR: A constraint-based FIR filter design program. *IEEE Trans. on Signal Processing*, 40(8):1901–1909, August 1992.

[25] B. A. Weisburn, T. W. Parks, and R. G. Shenoy. Error criteria for filter design. In *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, volume 3, pages 565–568, April 1994.

Table 1: Exchange Algorithms for Equiripple Filters

	Nonextra-ripple	Extra-ripple
Band edges ω_p,ω_s specified	PM [14, 16]	SA [23]
Weighted Chebyshev error δ_p,δ_s specified	New	HS [9, 11], HOS [12]

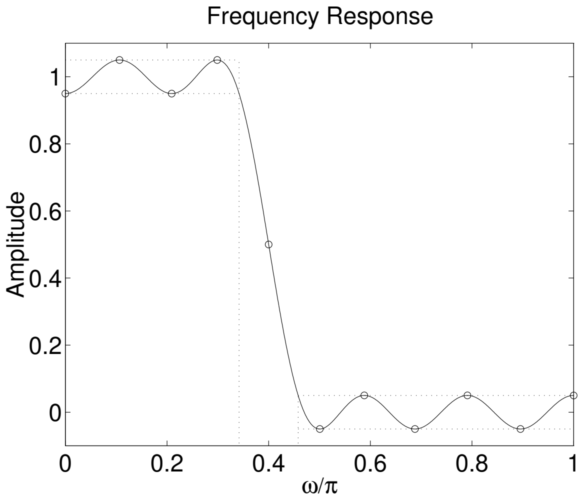


Figure 1: $N = 21$, $\delta_p = \delta_s = 0.05$, $\omega_o = 0.4\pi$, $\omega_p = 0.3418\pi$, $\omega_s = 0.4580\pi$

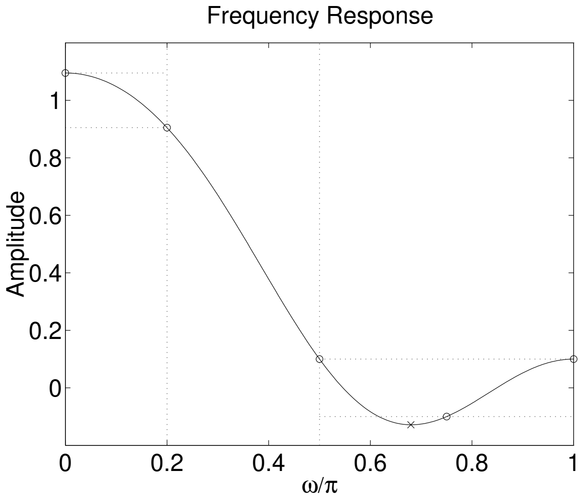


Figure 2: $N = 7$, $\omega_p = 0.2\pi$, $\omega_s = 0.5\pi$

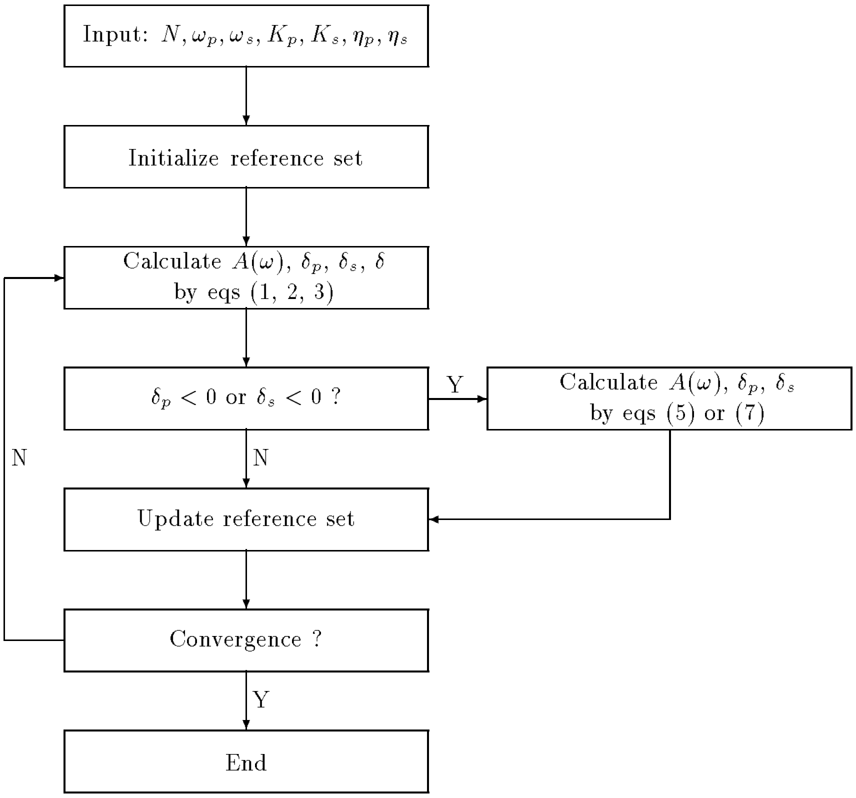
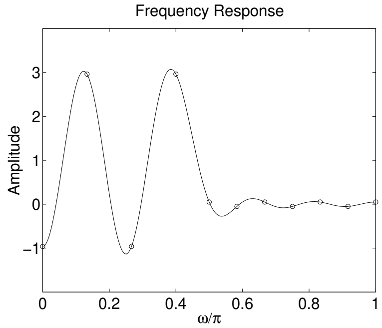
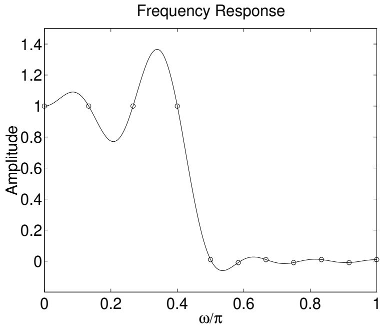


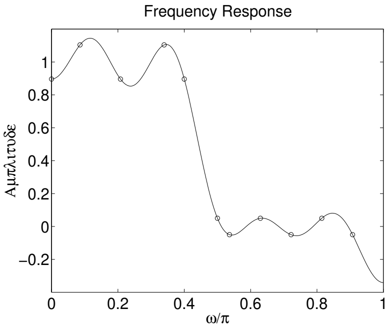
Figure 3: Flowchart for the PM algorithm for lowpass filter design modified to include an affine constraint between δ_p and δ_s.



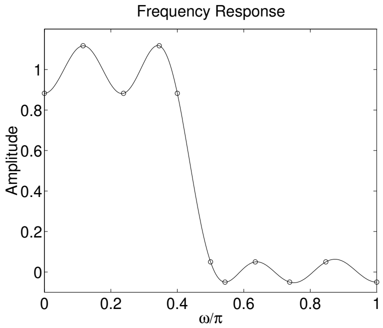
(a) $\delta_p = -1.9619$, $\delta_s = 0.05$



(b) $\delta_p = 0$, $\delta_s = 0.0096$



(c) $\delta_p = 0.1038$, $\delta_s = 0.05$



(d) $\delta_p = 0.1175$, $\delta_s = 0.05$

Figure 4: $N = 19$, $\omega_p = 0.4\pi$, $\omega_s = 0.5\pi$ (a) iteration 1 (b) iteration 1b (c) iteration 2 (d) iteration 3.

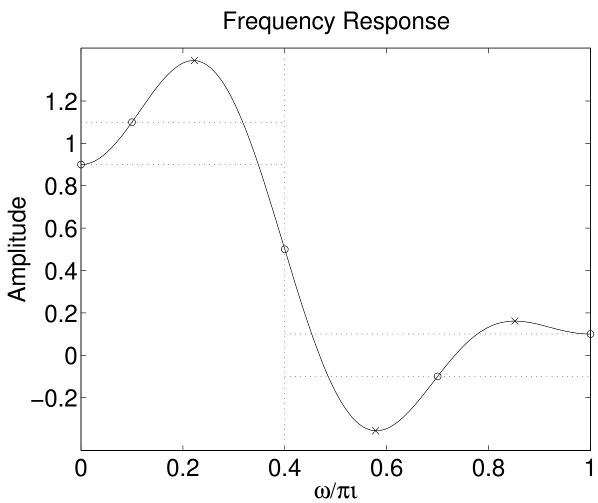


Figure 5: $N = 9$, $\omega_o = 0.4\pi$, $\delta_p = \delta_s = 0.1$

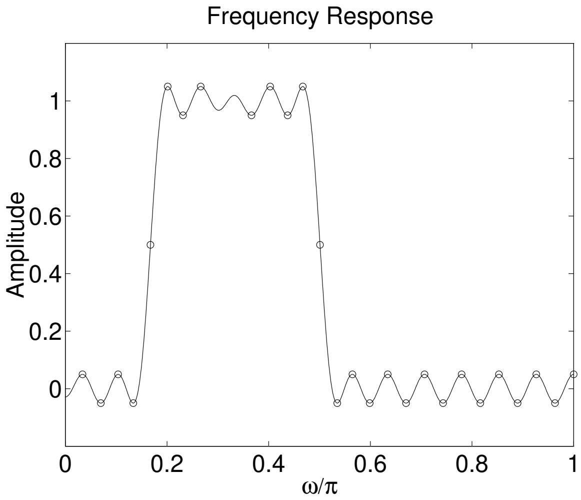


Figure 6: $N = 55$, $\delta_1 = \delta_2 = \delta_3 = 0.05$. The specified half-magnitude frequencies were $\omega_a = 0.1675\pi$, $\omega_b = 0.501\pi$. The 'induced' band edges are 0.1480π, 0.1870π, 0.4815π, and 0.5205π.

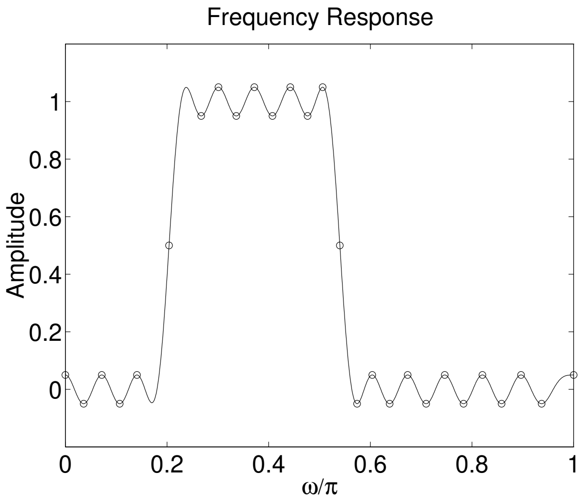


Figure 7: $N = 55$, $\delta_1 = \delta_2 = \delta_3 = 0.05$. The specified half-magnitude frequencies were $\omega_a = 0.204\pi$, $\omega_b = 0.54\pi$. The 'induced' band edges are 0.1844π 0.2235π 0.5205π 0.5595π.