# Chebyshev or Equal Ripple Error Approximation Filters

C. Sidney Burrus

This work is produced by OpenStax-CNX and licensed under the Creative Commons Attribution License 3.0†

If one poses the FIR filter design problem by requiring the maximum error over certain bands of frequencies be minimized, we call the resulting filter a Chebyshev filter or an equal ripple filter. The fact that the minimization of the Chebyshev or $L_\infty$ error results in an equal ripple error comes from the alternation theorem. This very powerful theorem allows one to minimize the Chebyshev error by directly constructing an equal ripple approximation with the proper number of ripples. That is the basis of several very effective algorithms, including the Remez exchange algorithm.

There are several ways one could pose the Chebyshev FIR filter design problem. For a simple length-N linear phase, lowpass filter with a transition band, if one considers the length N, the passband ripple $\delta_p$, the stopband ripple $\delta_s$, and the transition bandwidth $\Delta = \omega_s - \omega_p$, then one can fix or constrain any three of them and minimize the fourth. Or, as Parks and McClellan do, fix the band edges, $\omega_p$ and $\omega_s$, and the ratio of $\delta_p$ and $\delta_s$ and minimize one of them.

The Chebyshev error measure is often used for approximation in digital filter design. This is particularly true when the signals and/or noise are narrow band or single frequency or when one wants to minimize worst case possibilities. Theoretical justification for its use has been given by Weisburn, Parks, and Shenoy [93]. For FIR filter design, the Parks-McClellan formulation of the filter design problem and application of the Remez exchange algorithm is most commonly used [48], [49]. It is a particularly interesting and powerful method that should be studied and understood to be fully utilized.

Linear programming was used earlier [88], [26], [64] but dropped out of favor when the Parks-McClellan algorithm was introduced. It is now becoming more popular again because of more powerful computers, better algorithms [82], [6], and linear programming’s ability to allow a variety of constraints [80].

Still another approach to achieving a Chebyshev approximation is to minimize the $p^{th}$ power of the error using a large value of $p$ or to use an iterative scheme that solves a weighted least squared error with the weights at each stage determined by the error of the previous stage [15]. Still another design method that produces an equal ripple error approximation uses a constrained least squared error criterion [77], [76] which results in a Chebyshev solution if tight constraints are imposed.

The early work by Herrmann and Schüssler [27], [29] and the algorithm by Hofstetter, Oppenheim, and Siegel [31], [32] posed and solved a similar problem but they had only approximate control of $\omega_s$ (or $\omega_p$ or $\omega_s$) and always achieved the “extra ripple” design. Given the proper specifications, the Parks-McClellan algorithm could design any filter that the Hofstetter-Oppenheim-Siegel algorithm could, but the opposite is not true. This seems to be one of the reasons the Hofstetter-Oppenheim-Siegel algorithm is not commonly used.

---

*Version 1.3: Nov 17, 2012 5:58 pm -0600
†http://creativecommons.org/licenses/by/3.0/
http://cnx.org/content/m16895/1.3/
1 The Linear Phase FIR Filter Chebyshev Approximation Problem

The Chebyshev error is defined as the maximum difference between the actual and desired response over a band or several bands of frequencies. This is

\[ \epsilon = \max_{\omega \in \Omega} |A(\omega) - A_d(\omega)| \]  

where \( \Omega \) is the union of the bands of frequencies that the approximation is over \([17],[20]\). The approximation problem in filter design is to choose the filter coefficients to minimize \( \epsilon \).

One way to minimize \( \epsilon \) is to set up the frequency response in four equations for the four types of linear phase FIR filters as done in Equation 34 from FIR Digital Filters\(^1\), Equation 40 from FIR Digital Filters\(^2\), and the corresponding sine expressions. An alternative approach \([48]\) uses the fact that all four can be obtained from the odd-length, even-symmetry type 1 and uses only Equation 34 from FIR Digital Filters\(^3\).

From one of these frequency response representations together with powerful Alternation Theorem several optimization schemes can be developed.

If the amplitude response for odd \( L \) is expressed as a sum of \( R \) cosine terms

\[ A(\omega) = \sum_{n=0}^{R-1} a(n) \cos(\omega n) \]  

or for even \( L \)

\[ A(\omega) = \sum_{n=1}^{R} a(n) \cos(\omega (n - 1/2)) \]

with \( R = M + 1 = \frac{L+1}{2} \) for odd length-\( L \) and \( R = L/2 \) for even length-\( L \), as derived in Equation 34 from FIR Digital Filters\(^4\) and Equation 40 FIR Digital Filters\(^5\), then

**Theorem 1**

If \( A(\omega) \) is the linear combination of \( R \) cosine functions given in (2) or (3), the necessary and sufficient conditions for \( A(\omega) \) to be the least Chebyshev error approximation to \( A_d(\omega) \) over \( \omega \in \Omega \) are: The error function, \( \epsilon(\omega) = A(\omega) - A_d(\omega) \) have at least \( R + 1 \) extremal frequencies in \( \Omega \). The extremal frequencies are ordered points \( \omega_1 < \omega_2 < \cdots < \omega_{R+1} \) such that \( \epsilon(\omega_k) = -\epsilon(\omega_{k+1}) \) and \( \max_{\omega \in \Omega} |\epsilon(\omega)| = |\epsilon(\omega_k)| \) for \( k = 1, 2, \ldots, R + 1 \).

The alternation theorem \([48],[59]\) states that the minimum Chebyshev error has at least \( R + 1 \) extremal frequencies. This is stated mathematically by

\[ A(\omega_k) = A_d(\omega_k) + (-1)^k \delta \]  

for \( k = 0, 1, 2, \ldots, R \), where the \( \omega_k \) are the ordered extremal frequencies where the equal ripple error has maximum value. In other words, the optimal solution to the linear phase FIR filter design problem will have an equal ripple error function with a required number of ripples. How all of these characteristics relate can be rather complicated and good designs require experience \([28]\). When applied to other approximation problems, care must be taken to ensure the approximating functions satisfy the “Haar conditions” or other restrictions \([17],[49],[20],[59]\).

---

\(^1\) "FIR Digital Filters", [34] <http://cnx.org/content/m16889/latest/#uid1009>

\(^2\) "FIR Digital Filters", [40] <http://cnx.org/content/m16889/latest/#uid1009>

\(^3\) "FIR Digital Filters", [34] <http://cnx.org/content/m16889/latest/#uid1009>

\(^4\) "FIR Digital Filters", [34] <http://cnx.org/content/m16889/latest/#uid1009>

\(^5\) "FIR Digital Filters", [40] <http://cnx.org/content/m16889/latest/#uid1009>
2 Chebyshev Approximation by Linear Programming

It is possible to pose the Chebyshev approximation problem in filter design as a linear programming optimization problem [64], [89], [79], [42]. The error definition in (1) can be written as an inequality by

\[ A_d(\omega) - \delta \leq A(\omega) \leq A_d(\omega) + \delta \]  

where the scalar \( \delta \) is minimized.

The inequalities in (5) can be written as

\[ A \leq A_d + \delta \]  

(6)

\[ -A \leq -A_d + \delta \]  

(7)

or

\[ A - \delta \leq A_d \]  

(8)

\[ -A - \delta \leq -A_d \]  

(9)

which can be combined into one matrix inequality using Equation 48 from FIR Digital Filters\(^6\) by

\[
\begin{bmatrix}
C & -1 \\
-C & -1
\end{bmatrix}
\begin{bmatrix}
a \\ \delta
\end{bmatrix}
\leq
\begin{bmatrix}
A_d \\
-A_d
\end{bmatrix}.
\]

(10)

If \( \delta \) is minimized, the optimal Chebyshev approximation is achieved. This is done by minimizing

\[ \epsilon = \begin{bmatrix} 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} a \\ \delta \end{bmatrix} \]

(11)

which, together with the inequality of (10), is in the form of the dual problem in linear programming [19], [43], [81].

This can be solved using the \texttt{lp()} command from the Optimization Toolbox with Matlab [23], the Meteor software system [80], CPLEX [8], or Karmarkar’s algorithm [6], [35]. The Matlab \texttt{lp} command is implemented in an m-file using a form of quadratic programming algorithm that is not well suited to our filter design problem. Meteor is a linear programming system using the simplex algorithm written in Pascal by Ken Steiglitz especially for filter design. It has been compiled on a variety of computers and efficiently designs filters over 100 in length. The Karmarkar program written by Lang is a relatively short m-file that is not particularly fast but is robust and can design filters on the order of length-100. CPLEX is a proprietary program that can be used alone or called from Fortran programs and is particularly robust and fast.

A Matlab program that applies its linear programming function \texttt{lp.m} to (10),(11) for linear phase FIR filter design is given by:

```matlab
% lpdesign.m Design an FIR filter from L, f1, f2, and LF using LP.
% L is filter length, f1 and f2 are pass and stopband edges, LF is
% the number of freq samples. L is odd. Uses lp.m
% csb 5/22/91
L1 = fix(LF*f1/(.5-f2+f1)); L2 = LF - L1;  %No. freq samples in PB, SB
Ad = [ones(L1,1); zeros(L2,1)];  %Samples of ideal response
f = [0:L1-1]*f1/(L1-1), ([0:L2-1]*(.5-f2)/(L2-1) + f2)];  %Freq samples
```

\(^6\)FIR Digital Filters", [48] <https://cnx.org/content/m16889/latest/>
\[ M = \frac{(L-1)}{2}; \]
\[ C = \cos(2\pi(f[0:M])); \quad \text{%Freq response matrix} \]
\[ CC = [C, -\text{ones}(LF,1); -C, -\text{ones}(LF,1)]; \quad \text{%LP matrix} \]
\[ AD = [\text{Ad}; -\text{Ad}]; \]
\[ c = [\text{zeros}(M+1,1);1]; \quad \text{%Cost function} \]
\[ x0 = [\text{zeros}(M+1,1);\text{max}(AD)+1]; \quad \text{%Starting values} \]
\[ x = 1p(c,CC,AD,[1,1,x0]); \quad \text{%Call the LP} \]
\[ d = x(M+2); \quad \text{%delta or deviation} \]
\[ a = x(1:M+1); \quad \text{%Half impulse resp.} \]
\[ h = [a(M+1:-1:2);2*a(1);a(2:M+1)]./2; \quad \text{%Impulse response} \]

This program has numerical problems for filters longer than 10 or 20 and is fairly slow. The \text{lpc()} function uses an algorithm that seems not well suited to the equations required by filter design. It would be nice to have Meteor written in Matlab, both to show how the Simplex algorithm works, and to have an efficient LP filter design system in Matlab. The above program has been tested using Karmarkar’s algorithm [6], [66], [82] as implemented in Matlab by Lang [35]. It proved to be robust and reliable for lengths up to 100 or more. It was faster than the Matlab function but slower than Meteor or Cplex. Its use should be further investigated.

Direct use of quadratic programming and other optimization algorithms seem promising [22], [39], [52], [55], [53], [54], [57], [91], [90], [94]

3 Chebyshev Approximations using the Exchange Algorithms

A very efficient algorithm which uses the results of the alternation theorem is called the Remez exchange algorithm. Remez [65], [17], [39] showed that, under rather general conditions, an algorithm that takes a starting estimate of the location of the extremal frequencies and exchanges them with a new set calculated at each iteration will converge to the optimal Chebyshev approximation. The efficiency of this algorithm comes from finding the optimal solution by directly constructing a function that satisfies the alternation theorem rather than minimizing the Chebyshev error as done by the linear programming technique. The Remez exchange algorithm has proven to be well suited to the design of linear phase FIR filters [44], [47], [30].

A particularly useful FIR filter design implementation of the Remez exchange is called the Parks-McClellan algorithm and is described in [49], [63], [62], [48]. It has been implemented in Fortran in [30], [62], [18], [48] and in Matlab in a program at the end of this material. The Matlab program is particularly helpful in understanding how the algorithm works, however, because it does not use any special tricks, it is limited to lengths of 60 or so. Extensions and details can be found in [43], [10], [21], [78], [33], [24], [23], [71], [73], [72], [5]. This is a robust, efficient algorithm that significantly changed DSP when Parks and McClellan first described it in 1972 and has undergone important improvements. Examples are illustrated in [62], [46].

3.1 The Basic Parks-McClellan Formulation and Algorithm

Parks and McClellan formulated the basic Chebyshev FIR filter design problem by specifying the desired amplitude response \( A(\omega) \) and the transition band edges, then minimizing the weighted Chebyshev error over the pass and stop bands. For the basic lowpass filter illustrated in Figure 1, the pass band edge \( \omega_p \) and the stop band edge \( \omega_s \) are specified, the maximum passband error is related to the maximum stop band error by \( \delta_p = K \delta_s \) and they are minimized.
Notice that if there is no transition band, i.e. $\omega_p = \omega_s$, that $\delta_p + \delta_s = 1$ and no minimization is possible. While not the case for a least squares approximation, a transition band is necessary for the Chebyshev approximation problem to be well-posed. The effects of a small transition band are large pass and stopband ripple as illustrated in Figure 3b.

The alternation theorem states that the optimal approximation for this problem will have an error function with $R + 1$ extremal points with alternating signs. The theorem also states that there exists $R + 1$ frequencies such that, if the Chebyshev error at those frequencies are equal and alternate in sign, it will be minimized over the pass band and stop band. Note that there are nine extremal points in the length-15 example shown in Figure 1, counting those at the band edges in addition to those that are interior to the pass and stopbands. For this case, $R = (L + 1) / 2$ which agree with the example.

Parks and McClellan applied the Remez exchange algorithm [49] to this filter design problem by writing $R + 1$ equations using Equation 37 from FIR Digital Filters and Equation 1 from Design of IIR Filters by Frequency Transformations evaluated at the $R + 1$ extremal frequencies with $R$ unknown cosine parameters.

---

**Figure 1:** Amplitude Response of a Length-15 Optimal Chebyshev Filter

---

[FIR Digital Filters](http://cnx.org/content/m16889/latest/#uid1009)

[Properties of IIR Filters](http://cnx.org/content/m16898/latest/#uid1)
a(n) and the unknown ripple value, \( \delta \). In matrix form this becomes

\[
\begin{bmatrix}
A_d(\omega_0) & A_d(\omega_1) & A_d(\omega_2) & A_d(\omega_3) & \cdots & A_d(\omega_R)
\end{bmatrix}
\begin{bmatrix}
\cos(\omega_0) & \cos(\omega_1) & \cos(\omega_2) & \cos(\omega_3) & \cdots & \cos(\omega_R)
\end{bmatrix}
\begin{bmatrix}
1 & -1 & 1 & -1 & \cdots & 1
\end{bmatrix}
\begin{bmatrix}
a(0) \\
a(1) \\
a(2) \\
\vdots \\
a(R - 1)
\end{bmatrix} =
\begin{bmatrix}
a(R) \\
\delta
\end{bmatrix}
\]

These equations are solved for \( a(n) \) and \( \delta \) using an initial guess as to the location of the extremal frequencies \( \omega_i \). This design is optimal but only over the guessed frequencies, and we want optimality over all of the pass and stopbands. Therefore, the amplitude response of the filter is calculated over a dense set of frequency samples using Equation 34 from FIR Digital Filters\(^9\) and a new set of estimates of the extremal frequencies is found from the local minima and maxima and these are used to replace the initial guesses (they are exchanged). This process is iteratively performed until the guaranteed convergence is achieved and the optimal filter is designed.

The detailed steps of the Parks-McClellan algorithm are:

1. Specify the ideal amplitude, \( A_d(\omega) \), the stop and pass band edges, \( \omega_p \) and \( \omega_s \), the error weight \( K \) where \( \delta_p = K \delta_s \), and the length \( L \).
2. Choose \( R + 1 \) initial guesses for the extremal frequencies, \( \omega_i \), in the bands of approximation, \( \Omega \). This is often done uniformly over the pass and stop bands, including \( \omega = 0, \omega_p, \omega_s, \) and \( \pi \).
3. Calculate the cosine matrix at the current \( \omega_i \) and solve (12) for \( a(n) \) and \( \delta \) which are optimal over these current extremal frequencies, \( \omega_i \).
4. Using the \( a(n) \) or the equivalent \( h(n) \) from step 3, evaluate \( A(\omega) \) over a dense set of frequencies. This amplitude response will interpolate \( A(\omega_i) = A_d(\omega_i) \pm \delta \) at the extremal frequencies.
5. Find \( R + 1 \) new extremal frequencies where the error has a local maximum or minimum and has alternating sign. This includes the band edges.
6. If the largest error is the same as \( \delta \) found in step 3, then convergence has occurred and the optimal filter has been designed, otherwise, exchange the old extremal frequencies \( \omega_i \) used in step 2 and return to step 3 for the next iteration.
7. This iterative algorithm is guaranteed to converge to the unique optimal solution using almost any starting points in step 2.

This iterative procedure is called a multiple exchange algorithm because all of the extremal frequencies are updated each iteration. If only the frequency of the largest error is updated each iteration, it is called a single exchange algorithm which also converges but much more slowly. Some modification of the Parks-McClellan method or the Remez exchange algorithm will not converge as a multiple exchange, but will as a single exchange.

The Alternation theorem states that there will be a minimum of \( R + 1 \) extremal frequencies, even for multiband designs with arbitrary \( A_d(\omega) \). If \( A_d(\omega) \) is piece-wise constant with \( T \) transition bands, one can derive the maximum possible number of extremal frequencies and it is \( R + 2T \). This comes from the maximum number of maxima and minima that a function of the form (2) or (3) can have plus two at the edges of each transition band. For a simple lowpass filter with one passband, one transition band, and one stopband, there will be a minimum of \( R + 1 \) extremal frequencies and a maximum of \( R + 2 \). For a bandpass filter, the maximum is \( R + 4 \). If a design has more than the minimum number of extremal frequencies, it is called an extra ripple design. If it has the maximum number, it is called a maximum ripple design.

\(^9\)FIR Digital Filters\(^{,} \) [34] <http://cnx.org/content/m16889/latest/#uid1009>
It is interesting to note that at each iteration, the approximation is optimal over that set of extremal frequencies and \( \delta \) increased over the previous iteration. At convergence, \( \delta \) has increased to the maximum error over \( \Omega \) and that is the minimum Chebyshev error.

At each iteration, the exchange of a proper set of extremal frequencies with alternating signs of the errors is always possible. One can show there will never be too few and if there are too many, one uses those corresponding to the largest errors.

In step 4 it is suggested that the amplitude response \( A(\omega) \) be calculated over a dense grid in the pass and stop bands and in step 5 the local extremes are found by searching over this dense grid. There are more accurate methods that use bisection methods and/or Newton’s method to find the extremal points.

In step 2 it is suggested that the simultaneous equation of (12) be solved. Parks and McClellan [50] use a more efficient and numerically robust method of evaluating \( \delta \) using a form of Cramer’s rule. With that \( \delta \), an interpolation method can be used to find \( a(n) \). This is faster and allows longer filters to be designed than with the linear algebra based approach described here.

For the low pass filter, this formulation always has an extremal frequency at both pass and stop band edges, \( \omega_p \) and \( \omega_s \), and at \( \omega = 0 \) and/or at \( \omega = \pi \). The extra ripple filter has \( R + 2 \) extremal frequencies including both zero and pi. If this algorithm is started with an incorrect number of extremal frequencies in the stop or pass band, the iterations will correct this. It is interesting and informative to plot the frequency response of the filters designed at each iteration of this algorithm and observe how the correction takes place.

The Parks-McClellan algorithm starts with fixed pass and stop band edges then minimizes a weighted form of the pass and stop band error ripple. In some cases it may be more appropriate to fix one of the ripples and minimize the other or to fix both ripples and minimize the transition band width. Indeed Schüssler, Hofstetter, Tufts, and others [29], [27], [31], [32] formulated some of these ideas before Parks and McClellan developed their algorithm. The DSP group at Rice has developed some modifications to these methods and they are presented below.

### 3.2 Examples of the Parks-McClellan Algorithm

Here we look at several examples of filters designed by the Parks-McClellan algorithm. The examples here are length-15 with that shown in Figure 2a having a passband \( 0 < f < 0.3 \), a transition band \( 0.3 < f < 0.5 \), and a stopband \( 0.5 < f < 1 \). The number of cosine terms in the frequency response formula is \( R = 8 \), therefore, the alternation theorem says we must have at least \( R + 1 \) extremal points. There are four in the passband, counting the one at zero frequency, the minimum, the maximum, and the minimum at the band edge. There are five in the stop band, counting the ones at the band edge and at \( f = 1 \). So, the number is nine which is at least \( R + 1 \). However, in Figure 2c, there are ten extremal points but that is also at least 9, so it also is optimal. For a low pass filter, the maximum number of extremal points is \( R + 2 \) and that is what this filter has. This special case is called the “maximum ripple” case.
Figure 2: Amplitude Response of Length-15 Optimal Chebyshev Filters

It is possible to have ripples that do not touch the maximum value and, therefore, are not considered extremal points. That is illustrated in Figure 3a. The effects of a narrow transition band are illustrated in Figure 3c. Note the zero locations for these filters and how they relate to the amplitude response.
To illustrate some of the unexpected behavior that optimal filter designs can have, consider the bandpass filter amplitude response shown in Figure 4. Here we have a length-31 Chebyshev bandpass filter with a stopband $0 < f < 0.2$, a transition band $0.2 < f < 0.25$, a passband $0.25 < f < 0.5$, another transition band $0.5 < f < 0.68$, and a stopband $0.68 < f < 1$. The asymmetric transition bands cause large response in the transition band around $f = 0.6$. However, this filter is optimal since the deviation occurs in part of the frequency band that is not included in the optimization criterion. If you think you don’t care what happens in the transition bands, you may change your mind with this kind of behavior.
3.3 The Modified Parks-McClellan Algorithm

If one wants to fix the pass band ripple and minimize the stop band ripple [73], equation (12) is changed so that the pass band ripple is added to the appropriate top part of the vector \( A_d \) of the desired response and the unknown stop band is kept in the lower part of the last column of the cosine matrix \( C \).

\[
\begin{bmatrix}
A_d(\omega_0) \\
A_d(\omega_1) \\
\vdots \\
A_d(\omega_p) \\
A_d(\omega_s) \\
\vdots \\
A_d(\omega_R)
\end{bmatrix}
+ \begin{bmatrix} \delta_p \end{bmatrix} =
\begin{bmatrix}
\cos(\omega_0 0) & \cos(\omega_0 1) & \cdots & \cos(\omega_0 (R - 1)) & 0 \\
\cos(\omega_1 0) & \cos(\omega_1 1) & \cdots & \cos(\omega_1 (R - 1)) & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\cos(\omega_p 0) & \cos(\omega_p 1) & \cdots & \cos(\omega_p (R - 1)) & 0 \\
\cos(\omega_s 0) & \cos(\omega_s 1) & \cdots & \cos(\omega_s (R - 1)) & 1 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\cos(\omega_R 0) & \cos(\omega_R 1) & \cdots & \cos(\omega_R (R - 1)) & \pm 1
\end{bmatrix}
\begin{bmatrix}
a(0) \\
a(1) \\
\vdots \\
a(2) \\
a(R - 1) \\
\delta_s
\end{bmatrix}
\]

(13)

Iteration of this equation will keep the pass band ripple \( \delta_p \) fixed and minimize the stop band ripple \( \delta_s \). A problem with convergence occurs if one of the \( \delta \)’s becomes negative during the iterations. A modification to the basic exchange has been developed to give reliable convergence [73].

3.4 The Hofstetter, Oppenheim, and Siegel Algorithm

This algorithm [31], [32], [73] came into existence in order to design the filters posed by Herrmann and Schüssler [29], [27] where both the pass and stop band ripple sizes, \( \delta_p \) and \( \delta_s \), are fixed and the location of
the transition band is not directly controlled. This problem results in a maximum ripple design which, for the lowpass filter, requires extremal frequencies at both $\omega = 0$ and $\omega = \pi$ but does not use either pass or stop band frequencies $\omega_p$ or $\omega_s$. This results in $R$ extremal frequencies giving $R$ equations to find the $R$ values of $a(n)$.

$$
\begin{bmatrix}
A_d(\omega_0) & \delta_p \\
A_d(\omega_1) & -\delta_p \\
\vdots & \vdots \\
A_d(\omega_{p-1}) & \pm \delta_p \\
A_d(\omega_{s+1}) & \delta_s \\
\vdots & \vdots \\
A_d(\omega_{R-1}) & \pm \delta_s
\end{bmatrix}
+ 
\begin{bmatrix}
cos(\omega_00) & cos(\omega_01) & \cdots & cos(\omega_0(R-1)) \\
cos(\omega_10) & cos(\omega_11) & \cdots & cos(\omega_1(R-1)) \\
\vdots & \vdots & \ddots & \vdots \\
cos(\omega_{p-1}0) & cos(\omega_{p-1}1) & \cdots & cos(\omega_{p-1}(R-1)) \\
cos(\omega_{s+1}0) & cos(\omega_{s+1}1) & \cdots & cos(\omega_{s+1}(R-1)) \\
\vdots & \vdots & \ddots & \vdots \\
cos(\omega_{R-1}0) & cos(\omega_{R-1}1) & \cdots & cos(\omega_{R-1}(R-1))
\end{bmatrix}
\begin{bmatrix}
a(0) \\
a(1) \\
\vdots \\
a(R-1)
\end{bmatrix}.
$$

This algorithm is iterated as a multiple exchange, keeping the number of ripples in the pass and stop band constant, to give an optimal extra ripple filter. The location and width of the transition band is controlled only by the choice of how the number of initial ripples are divided between the pass and stop band. The final filter may not have the transition located where you want it. Indeed, no solution may exist with the desired location of the transition band.

The designs produced by the HOS algorithm are always maximum ripple but this comes with a loss of accurate control over the location of the transition band. The algorithm is not, strictly speaking, an optimization algorithm. It is an interpolation algorithm. The Chebyshev error is not minimized, the designed amplitude interpolates the specified error ripples. However, although not directly minimized, the transition band width of these designs seems to be minimized [63], [51], [62]. Extra or maximum ripple designs seem to be efficient in using all the zeros to produce small ripple size and narrow transition bands, however, the loss of accurate control over the location of the transition bands becomes even more problematic with multiple transition band designs. Perhaps some compromise methods can be devised that use some of the efficiency of the maximum ripple approximations with some of the control of other methods. The next two design methods are of that type.

3.5 The Shpak and Antoniou Algorithm

Shpak and Antoniou [78] propose decoupling the size of the pass and stopband ripple sizes in order to have control over the pass and stop band edges and have an extra ripple design. The Parks-McClellan design has the ripple sizes related with a fixed weight $\delta_p = K \delta_s$, the modified Parks-McClellan design fixes one ripple size and minimizes the other, the Hoffstetter, Oppenheim, and Siegel design fixes both ripple sizes but cannot set the transition band edges. The Shpak-Antoniou design fixes the transition band edges and gives a maximum ripple design with minimum ripple but the relationship of the pass and stopband ripple is uncontrolled.

This method has two ripple sizes, $\delta_p$ and $\delta_s$, appended to the $a(n)$ vector similar to the single $\delta$ used in (12) or (13). This allows controlling an additional extremal frequency and results in an extra ripple approximation. This can become somewhat complicated for multiple transition bands but seems very flexible [5].

http://cnx.org/content/m16895/1.3/
3.6 The New Equal Ripple Design Formulation and Exchange Algorithm

Because the arguments in the Weisburn, Parks, and Shenoy paper [93] require the assumption of no signal or noise energy in the transition band, it is now more obvious that a narrow transition band is very desirable. For this reason it may be better to fix the pass and stop band peak error, \( \delta_p \) and \( \delta_s \), and the transition band center frequency \( \omega_0 \), then minimize the transition band width rather than fixing the pass and stop band edges, \( \omega_p \) and \( \omega_s \), then minimizing \( \delta_p \) and \( \delta_s \). Two methods have been recently developed to address this point of view. The first is a new exchange algorithm that is in some ways a combination of the Parks-McClellan and Hofstetter-Oppenheim-Siegel algorithms [65] and the second is a limiting case for a constrained least squares method based on Lagrange multipliers [12], [74], [77], [76] using tight constraints.

For problems where the signal and noise spectra are such that a specific frequency \( \omega_0 \) that separates the desired passband from the desired stopband can be specified but specific separate transition band edges, \( \omega_p < \omega_s \), cannot, we formulate [73] a design method where the pass and stop band ripple sizes, \( \delta_p \) and \( \delta_s \), are specified along with the separation frequency, \( \omega_s \). The algorithm described below will interpolate the specified ripple sizes exactly (as the HOS algorithm does) but will allow exact control over the location of \( \omega_0 \) by not requiring maximum ripple. Although not set up to be an optimization procedure, it seems to minimize the transition band width. This formulation suits problems where there is no obvious transition band ("don't care band") having no signal or noise energy to be passed or rejected.

The optimal Chebyshev filter designed with this new algorithm is generally not extra ripple and, therefore, will have an extremal frequency at \( \omega = 0 \) or \( \omega = \pi \) as the Parks-McClellan formulation does. Because we are trying to minimizing the transition band width, we do not specify both the edges, \( \omega_p \) and \( \omega_s \), but only one of them or, perhaps, the center of the transition band, \( \omega_0 \). This results in \( R \) equations which are used to find the \( R \) coefficients \( a(n) \). The equations are formulated by adding the alternating peak pass and stop band ripples to the \( A_d \) in (12) and not having the special last column of \( C \) nor the unknown \( \delta \) appended to \( a \) as was done by Parks and McClellan in (12). The resulting equation to be iterated in our new exchange algorithm has the form

\[
\begin{bmatrix}
A_d(\omega_0) & \delta_p \\
A_d(\omega_1) & -\delta_p \\
\vdots & \vdots \\
A_d(\omega_0) & 0 \\
A_d(\omega_{s+1}) & -\delta_s \\
\vdots & \vdots \\
A_d(\omega_{R-1}) & \pm\delta_s
\end{bmatrix}
= \begin{bmatrix}
\cos(\omega_0) & \cos(\omega_0) & \cdots & \cos(\omega_0) \\
\cos(\omega_0) & \cos(\omega_0) & \cdots & \cos(\omega_0) \\
\vdots & \vdots & \ddots & \vdots \\
\cos(\omega_0) & \cos(\omega_0) & \cdots & \cos(\omega_0) \\
\cos(\omega_{s+1}) & \cos(\omega_{s+1}) & \cdots & \cos(\omega_{s+1}) \\
\vdots & \vdots & \ddots & \vdots \\
\cos(\omega_{R-1}) & \cos(\omega_{R-1}) & \cdots & \cos(\omega_{R-1})
\end{bmatrix}
\begin{bmatrix}
a(0) \\
a(1) \\
\vdots \\
a(R-1)
\end{bmatrix}
\]

The exchange algorithm is done as by Parks and McClellan finding new extremal frequencies at each iteration, but with fixed ripple sizes in both pass and stop bands. This new algorithm reduces the transition band width as done by the Hofstetter, Oppenheim, and Siegel method but with the transition band location controlled and without requiring the extra ripple solution. Note that any transition band frequency could be fixed. It could be \( A_d(\omega_0) = 1/2 \) to fix the half-power point. It could be \( A_d(\omega_p) = 1 - \delta_p \) to fix the pass band edge. Or it could be \( A_d(\omega_s) = \delta_s \) to fix the stop band edge.

Extending this formulation and algorithm to the multiple transition band case complicates the problem as the solution may not be unique or may have anomalous behavior in one of the transition bands. Details of the solution to this problem are given in [73].
3.7 Estimations of $N$, the Length of Optimal Chebyshev FIR Filters

All of the design methods discussed so far have assumed that $N$, the length of the filter, is given as part of the specifications. In many cases, perhaps even most, $N$ is a parameter that we would like to minimize. Often specifications are to meet certain pass and stopband ripple specifications with given pass and stopband edges and with the shortest possible filter. None of our methods will do that. Indeed, it is not clear how to do that kind of optimization other than by some sort of search. In other words, design a set of filters of different lengths and choose the one that meet the specifications with minimum length.

Fortunately, empirical formulas have been derived that give a good estimate of the relationship of the length of an optimal Chebyshev FIR filter for given pass and stopband ripple and transition band edges [62], [63]. Kaiser’s formula is

$$N = \frac{-20\log_{10}(\sqrt{\delta_p \delta_s}) - 13}{14.6(f_s - f_p)} + 1$$

and it is fairly accurate for average filter specifications (neither wide nor narrow bands).

3.8 Examples of Optimal Chebyshev Filters

In order to better understand the nature of an optimal Chebyshev and to see the power of the Parks-McClellan algorithm, we present the design of a length-21 linear phase FIR bandpass filter. To see the effects of the design specifications, we will fix the two pass band edges and the upper stop band edge, then look at the effects of varying the lower stop band edge. The Matlab program that generated the designs is:

```matlab
N = 20;
M = [0 0 1 1 0 0];
W = [7.5 10 7.5];
ff = [0:512]/512; k=0;
%for fk = .10:.02:.34
% k = k+1;
clf;
for k = 1:6
    fk = .1 + .02*(k-1);
    F = [0 fk .35 .8 .85 1];
    b = firpm(N,F,M,W);
    axis([0 1 0 1.2]);
    AA = abs(fft(b,1024)); AA = AA(1:513);
    dd = max(AA(1:50));
    ddd = dd*(W(1)/W(2));
    subplot(3,2,k); plot(ff,AA,'r'); hold;
    plot([0 F(2) F(2) F(5) F(5) 1],[dd dd 0 0 dd dd],'b');
    plot([0 F(3) F(3) F(4) F(4) 1],[0 0 1-ddd 1-ddd 0 0],'b');
    plot([0 F(3) F(3) F(4) F(4) 1],[0 0 1+ddd 1+ddd 0 0],'b');
title('L-21 Chebyshev Filter, f_s = 0.1');
ylabel('Magnitude $|H(\omega)|$');
pause;
end; hold off;
```
The results are shown in Figures 5 and 6.

**Figure 5:** Amplitude Response of Length-21 Optimal Chebyshev Bandpass Filter with various Stop Band Edges
Note the large transmission peaks in the transition band of Figures Figure 5a, b, and c that result from the two transition bands being very different in width. As the lower transition band narrows, this peak grows smaller and eventually disappears in Figure 5f. Note that there are two extremal points in the lower stop band of Figure 5b and seven in the pass band, while there are three in the lower stop band of Figures Figure 5c and d and six in the pass band. But, there are always twelve total (thirteen for a case between Figures Figure 5b and c). In Figure 6d, there are only five extremal points in the pass band but twelve total. The same filter is optimal for the conditions given in Figures Figure 6a, b, and c. Much can be learned about optimal filters by running experiments in Matlab. Remember, all of these are optimal for the specifications

Figure 6: Amplitude Response of Length-21 Optimal Chebyshev Bandpass Filter with various Stop Band Edges
given.

4 Chebyshev Approximation using Approximation

It is possible to approximate the effects of Chebyshev approximation by minimizing the \( p^{th} \) power of the error. For large \( p \) this is close to the results of a true Chebyshev approximation. This is a variation on a method called Lawson's method. This approach is described in [13], [14], [15] using the iterative reweighted least squared (IRLS) error method and looks attractive in that it can use different \( p \) in different frequency bands. This would allow, for example, a least squared error approximation in the passband and a Chebyshev approximation in the stopband. The IRLS method can also be used for complex Chebyshev approximations [86].

5 Characteristics of Optimal Chebyshev Filters

Examples of expected and unexpected results of optimality. Rabiner's work will be used here. The non-unique designs for certain multiband designs will be illustrated.

6 Complex Chebyshev Approximation

Algorithms that directly use the alternation theorem, such as the standard Remez multiple exchange algorithm, are difficult to apply to the complex approximation or 2-D approximation problem because the concept of "alternation" is difficult to define and the number of ripples in an optimal solution is more difficult to determine [92], [84], [83], [9], [40], [41], [85]. Work has been done on the complex approximation problem at Rice by Parks and Chen [16] and by Burrus, Barreto, and Selesnick [15], [7], at Erlangen by Schuessler, Preuss, Schulist, and Lang [60], [61], [67], [68], [70], [69], at MIT by Alkhairy et al [3], [4], at USC by Tseng and Griffiths [86], [87], at Georgia Tech by Karam and McClellan [34], at Cornell by Burnside and Parks [11], and by Potchinkov and Reemtsen at Cottbus [53], [54], [57], [58], [56]. The work done by Adams which uses an implementation of a constrained quadratic programming algorithm might be useful here [1], [2]. Lang has extended and further developed this constrained approach [36], [37], [38] and Selesnick is applying it to IIR filter design [75]. Tseng gives a good summary of complex approximation in [87].

7 Conclusions and Discussions of Chebyshev Design

By adding the Chebyshev filter design methods described above to the Parks-McClellan algorithm, one has a rather complete set of approaches to equal ripple filter designs that allows a wide variety of specifications. The new exchange algorithm which minimizes the transition band width while allowing the specification of the center or either edge of the transition band edge may fit many design environments better than the traditional Parks-McClellan. An alternative approach which specifies the pass and stop band peak error yet has no zero weighted transition band will be presented in Constrained Least Squares Design\(^\text{10}\). Matlab programs are available for the Parks-McClellan algorithm, the modified Parks-McClellan algorithm, the Hofstetter-Oppenheim-Siegel algorithm, the new minimum transition band design algorithm, and the constrained least squares algorithm. They are written with a common format and notation to easily see how they are programmed and how they are related. This book generally presents the lowpass case. The bandpass and multi-band cases use the same ideas but are a bit more complicated and are discussed in more detail in the references.

\(^{10}\)"Constrained Approximation and Mixed Criteria": Section Constrained Least Squares Design

<http://cnx.org/content/m16923/latest/#uid3>

http://cnx.org/content/m16895/1.3/
References


http://cnx.org/content/m16895/1.3/


http://cnx.org/content/m16895/1.3/


