# On the Transition Width of Finite Impulse-Response Digital Filters

THOMAS W. PARKS, LAWRENCE R. RABINER, and
JAMES H. MC CLELLAN

*Abstract*—Several properties of finite-duration impulse-response
(FIR) digital filters designed to have the maximum possible number
of ripples are discussed and illustrated with examples. Such filters
have been called extraripple filters. Among the properties of such
filters are as follows.

1) Extraripple low-pass filters with fixed passband ripple $\delta_1$ and
stopband ripple $\delta_2$ achieve the local minimum of transition width in
the class of linear phase filters with fixed impulse-response duration
of $N$ samples.

2) For the case $\delta_1 = \delta_2$ the minimum transition width is roughly
independent of $F_p$, the passband cutoff frequency.

3) For the case $\delta_2 < \delta_1$, the minimum transition width decreases
with increasing bandwidth.

Several figures are included to show the relation between the
transition width and bandwidth for low-pass filters.

## Introduction

The problem of designing finite impulse-response
(FIR) low-pass digital filters with linear phase has re-
cently been studied from several points of view [1]–[6].
While the problem of obtaining optimum magnitude
characteristics has been solved, several questions remain
concerning the width of the transition band (stopband
edge minus passband edge) for filters with optimum
magnitude response. This paper attempts to answer
these questions.

Fig. 1 shows plots of the frequency response and the
error curve of a typical equiripple low-pass filter. The
passband cutoff frequency is $F_p$ and the stopband cutoff
frequency is $F_s$. The maximum deviation in the pass-
band, or passband ripple is $\delta_1$, and the stopband ripple
is $\delta_2$. The width of the transition band is $\Delta F = F_s - F_p$.
If the duration of the filter impulse response is $N$
samples ($N$ odd) then, for given $F_p$ and $F_s$, a filter with
optimum magnitude characteristics can have either
$(n+2)$ or $(n+3)$ maxima in the *weighted* error curve
[5] where

$$n = \frac{N - 1}{2} \qquad (1)$$
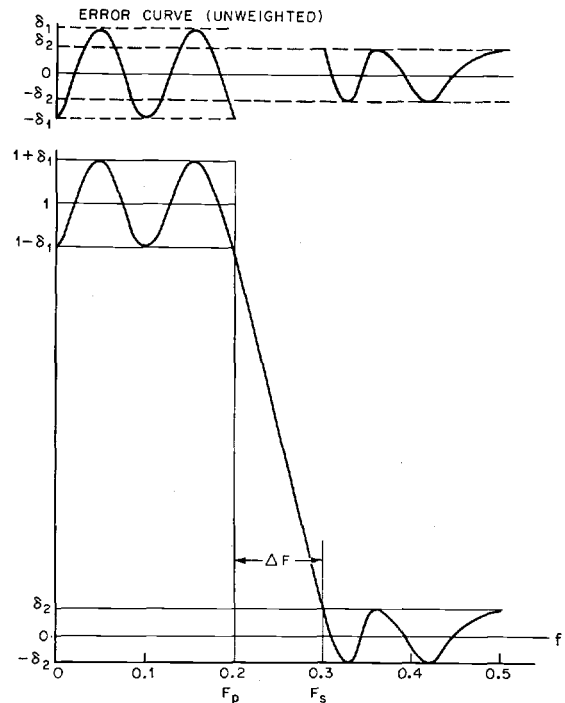
and the weighted error curve is defined by

Fig. 1. Frequency response and error curve for an optimal
low-pass filter defining the basic filter parameters.

$$e(f) = (d(f) - g(f))W(f) \qquad (2)$$

where $d(f)$ is the desired response defined as

$$d(f) = \begin{cases} 1, & 0 \le f \le F_p \\ 0, & F_s \le f \le 0.5. \end{cases} \qquad (3)$$

$W(f)$ is a weight function specifying the relative magni-
tude of the error in the passband and stopband and is
defined by

$$W(f) = \begin{cases} \dfrac{1}{K} = \dfrac{\delta_2}{\delta_1}, & 0 \le f \le F_p \\ 1, & F_s \le f \le 0.5 \end{cases} \qquad (4)$$

and $g(f)$ is the actual filter amplitude response which is
defined as

$$g(f) = \sum_{k=0}^{n} d(k) \cos (2\pi fk) \qquad (5)$$

where the $d(k)$ are related to the filter's impulse response
$\{h(k)\}$ by the relation

$$d(n - k) = 2h(k), \qquad k = 0, 1, \cdots, n - 1$$
$$d(o) = h(n). \qquad (6)$$

The $h(k)$ obey the symmetry condition

$$h(k) = h(2n - k), \qquad k = 0, 1, \cdots, n. \qquad (7)$$

For the optimum magnitude filter, the maxima of the
weighted error curve must alternate in sign and be equal
in magnitude, i.e.,

$$e(F_i) = - e(F_{i+i}), \qquad i = 1, 2, \cdots, n + 1 \qquad (8)$$

where $F_i$ are the frequencies at which the error maxima occur.

It has been conjectured that filters with $n+3$ maxima (the so-called extraripple filters) have minimum transition width [4]. This question of transition width for extraripple filters was more fully explored in a paper by Hofstetter et al. [7], which has motivated this paper. By mapping the interval $0 \le f \le \frac{1}{2}$ to the interval $-1 \le x \le 1$ by the transformation $x = \cos 2\pi f$, these authors converted (5) from a trigonometric polynomial in $f$ to a standard polynomial in $x$. They obtained the frequency response

$$P(x) = G(f)\big|_{x=\cos (2\pi f)} = \sum_{k=0}^{n} a_k x^k. \qquad (9)$$

Using this formulation Hofstetter et al. were able to design a filter with $(n+2)$ maxima with a smaller transition bandwidth than the same length extraripple filter $(n+3$ maxima). Both filters had the same values for $\delta_1$ and $\delta_2$. This "counterexample" precluded any arguments for the optimality of extraripple filters according to the authors.

Based on these observations a more detailed evaluation of the behavior of transition bandwidth as a function of passband cutoff frequency was made. This paper describes the relation between transition bandwidth $\Delta F$ and passband cutoff frequency $F_p$ for linear phase filters of fixed duration with optimum magnitude and fixed values for $\delta_1$ and $\delta_2$. It will be shown that the extraripple filters are local minima on the curve of $\Delta F$ versus $F_p$ and the influence of $K = \delta_1/\delta_2$ will be described.

### Local Minimum Property of Extraripple Filters

Within the class of linear phase filters with fixed duration $N = 2n+1$ samples, and with optimum magnitude ($\delta_1$ and $\delta_2$ fixed), the filters with $n+3$ alternations in the error curve *locally* have minimum transition width $\Delta F$. This property can be shown by first noting that by the results in [5], filters with optimum magnitude must have either $(n+2)$ or $(n+3)$ alternations. Furthermore, if one of the error maxima is missing, it can only be missing from one of the endpoints ($f=0$, $\frac{1}{2}$, or equivalently $x = \pm 1$), for otherwise more than one alternation point would be lost and the resulting filter would not have optimum magnitude. Finally, if one of the endpoint maxima are missing, the same scaling argument as used in [7] shows that the linear compression to give $(n+3)$ error maxima reduces $\Delta F$. Thus the extraripple filters represent local minima on the curve of $\Delta F$ versus $F_p$.

### The Relation Between $\Delta F$ and $F_p$

In order to better illustrate the detailed behavior of transition width, several figures are included for various choices of filter length, and weighting factor $K = \delta_1/\delta_2$.
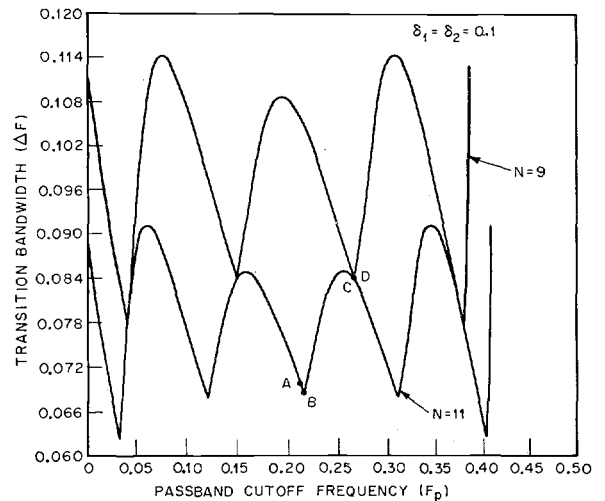


Fig. 2. Transition width as a function of passband cutoff frequency for $K=1$, $N=9$, and $N=11$.

Fig. 2 shows a plot of $\Delta F$ versus $F_p$ for optimum filters of length $N=9$ and of length $N=11$. For this example $\delta_1 = \delta_2 = 0.1$ (i.e., $K=1$). Since the filter design parameters are $F_p$, $F_s$, and $K$, an iterative procedure was used to adjust $F_s$ (or equivalently $\Delta F$) to achieve $\delta_1 = \delta_2 = 0.1$ to within some desired accuracy. Generally, on the order of 10 iterations were required to achieve the desired resolution ($\delta_1 = \delta_2 = 0.1$ within 0.01 percent). As seen in Fig. 2, the extraripple filters both for $N=9$ and $N=11$ correspond to local minima of the curves. (Table I gives the specifications of these extraripple filters.) In between these local minima are points with greater transition width corresponding to filters with $(n+2)$ alternations. It is interesting to note that extraripple length 9 filters are essentially optimum length 11 filters, and occur between the local minima for the $N=11$ curve. It is also interesting to note that there are optimum length 9 filters with *smaller* transition widths than optimum length 11 filters. However, for a given passband edge frequency, $F_p$, length 9 optimum filters *always* have greater transition widths than optimum length 11 filters.

To illustrate several of the above points, filters corresponding to four of the points in the curves of Fig. 2 have been examined in greater detail. These points are denoted in Fig. 2 by the letters $A$, $B$, $C$, and $D$. Fig. 3 shows the magnitude frequency response and $z$-plane plots of the zero positions for these four filters. Filter $A$ is an optimum filter ($N=11$) with 7 equiripple alternations ($n+2$) and one alternation which is not equal in amplitude to the other ripple peaks. Thus filter $A$ is optimum, but not equiripple. Filter $B$ is a length 11 extraripple filter with 8 equiripple alternations ($n+3$). Filter $C$ is an optimum length 11 filter with 7 equiripple alternations. Filter $D$ is a length 9 extraripple filter, which is also an optimum length 11 filter. A comparison of filters $C$ and $D$ shows they are essentially the same filter.

The $z$-plane plots of Fig. 3 show how the zeros move

**TABLE I**

Extraripple Filters for $N=9$ and $N=11$ $(\delta_1=\delta_2=0.1)$

| | | | | | |
|---|---|---|---|---|---|
| $n_p$ | Number of passband ripples | | | | |
| $n_s$ | Number of stopband ripples | | | | |
| $F_p$ | Passband cutoff frequency | | | | |
| $F_s$ | Stopband cutoff frequency | | | | |
| $\Delta F$ | Width of transition band | | | | |

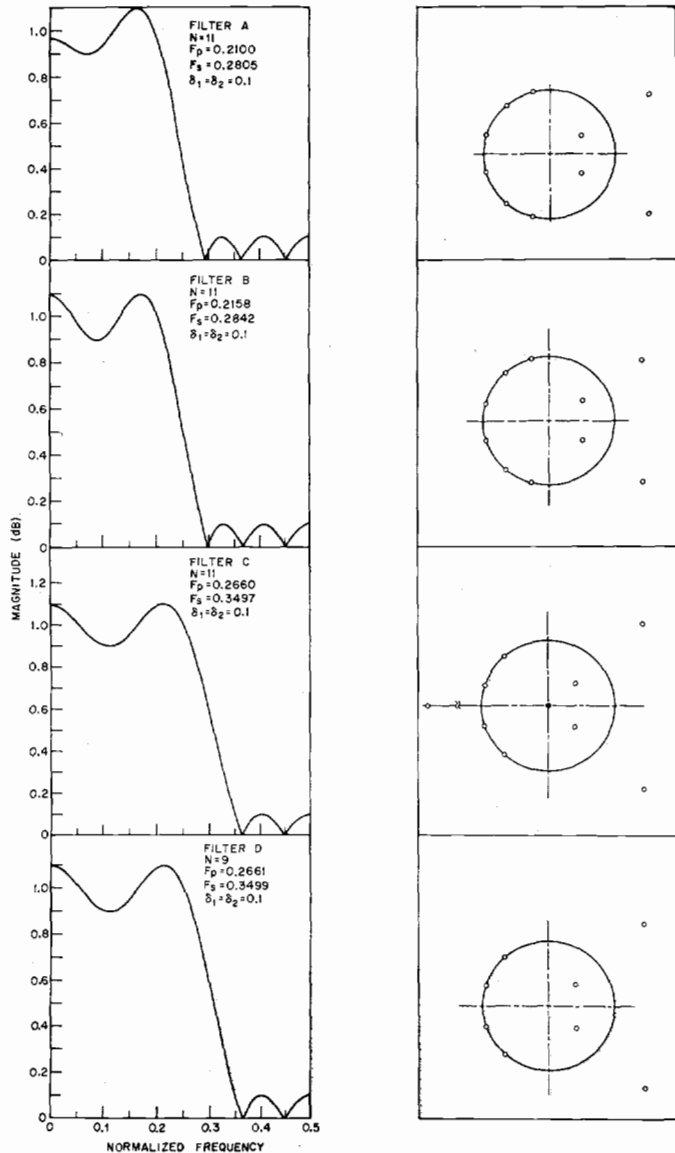| $N$ | $n_p$ | $n_s$ | $F_p$ | $F_s$ | $\Delta F$ |
|---|---|---|---|---|---|
| 11 | 1 | 5 | 0.034406 | 0.096799 | 0.062393 |
| 9 | 1 | 4 | 0.042644 | 0.119961 | 0.077317 |
| 11 | 2 | 4 | 0.121333 | 0.189137 | 0.067804 |
| 9 | 2 | 3 | 0.150119 | 0.233909 | 0.083790 |
| 11 | 3 | 3 | 0.215756 | 0.284244 | 0.068488 |
| 9 | 3 | 2 | 0.266091 | 0.349881 | 0.083790 |
| 11 | 4 | 2 | 0.310863 | 0.378667 | 0.067804 |
| 9 | 4 | 1 | 0.380039 | 0.457356 | 0.077317 |
| 11 | 5 | 1 | 0.403201 | 0.465594 | 0.062393 |



Fig. 3. Frequency response and $z$-plane zero positions of four of the filters of Fig. 2.
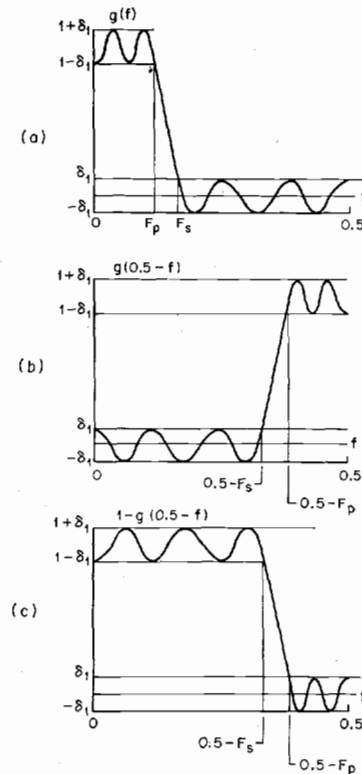


Fig. 4. Construction for demonstrating the symmetry of the curve of transition width versus passband cutoff frequency for $K=1$.

as we proceed from filter $A$ to filter $D$. Filter $A$ has 3 complex conjugate sets of zeros on the unit circle and one quadrouplet of zeros off the unit circle. Filter $B$ has essentially the same zero position plot as filter $A$ since their frequency characteristics are very close. In filter $C$, the pair of zeros on the unit circle closest to $z=-1$ has become a mirror-image real pair of zeros on the negative real axis in the $z$ plane. One zero is approximately at $z=0$; its mirror-image partner is almost at $z=-\infty$ (actually the magnitudes of these zeros are 0.002 and 495.5, respectively). In filter $D$, the zeros on the real axis are gone having been cancelled by poles at $z=0$. Thus a filter with a 9 point impulse response and essentially identical frequency-response characteristics is the result, as seen earlier.

Fig. 2 also shows that for $K=1$ the plot of $\Delta F$ versus $F_p$ is symmetrical around $F_p=0.21576$. (In this case $F_s=0.28424$—i.e., $(F_p+F_s)/2=0.25$.) An explanation of this symmetry is given with the help of Fig. 4. Part (a) of this figure shows the frequency response of an optimum filter $g(f)$ with parameters $F_p$, $F_s$, $\delta_1$, and $\delta_2$. Part (b) of this figure shows the frequency response of $g(0.5-f)$, which is essentially a high-pass filter. Finally, part (c) shows the frequency response of $1-g(0.5-f)$ which is again an optimum low-pass filter with the same $\delta_1$, $\delta_2$, and transition width as the filter of (a), but with passband cutoff frequency $0.5-F_s$, and stopband cutoff frequency $0.5-F_p$. Thus the curves of Fig. 2 are symmetrical.
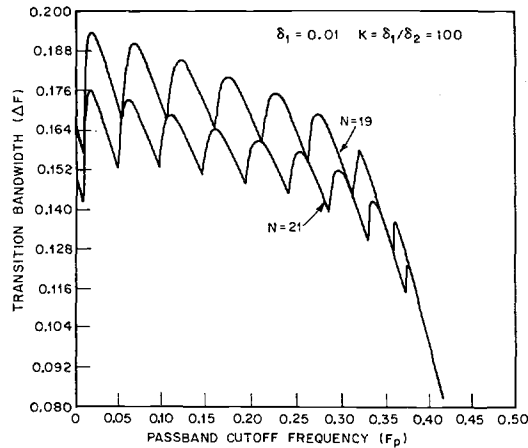
A plot of transition width as a function of passband

Fig. 5.   Transition width versus passband cutoff frequency
for $K = 100$, $N = 19$, and $N = 21$.



Fig. 6.   An explanation of Hofstetter's counter example.

cutoff frequency is greatly different for $K > 1$ than for $K = 1$. The differences are illustrated in Fig. 5. In this figure are plotted curves of $\Delta F$ versus $F_p$ for optimum filters with $N = 19$ and $N = 21$, and $\delta_1 = 0.01$, $K = 100$. (The data for these curves were obtained in a way similar to the data for Fig. 2.) This figure shows that as $F_p$ is increased, transition width decreases quite dramatically. For the widest bandwidth filters the transition width is on the order of one-half the size required for the narrowest bandwidth filters. An intuitive explanation of this behavior is that as $F_p$ is increased, and stopband ripples are moved into the passband, they are allowed to grow by a factor of $K$ (100 in this case), thus allowing a smaller $\Delta F$. For $K < 1$, $\Delta F$ increases with increasing $F_p$. This figure also shows that the symmetry evident in Fig. 2 is lost. The argument used to justify the symmetry of Fig. 2 is invalid when $K \neq 1$.

Finally, with the aid of Fig. 6 the counterexample presented by Hofstetter et al. can be readily explained. The two local minima in this figure correspond to extraripple filters with $n_p$, the number of passband ripples, equal to 8, $n_s$, the number of stopband ripples, equal to 3, and $n_p = 9$, $n_s = 2$ ($N = 21$, $\delta_1 = 0.01$, $K = 100$). The transition widths for these filters at the local minima are 0.130 and 0.114, respectively. Hofstetter et al. found an optimum filter, (denoted as point $A$ in Fig. 6) which was not extraripple, with a transition bandwidth of 0.115. Since the transition bandwidth of this filter is smaller than that of the extraripple filter ($n_p = 9$, $n_s = 2$). Hofstetter et al. concluded that the optimum filter is not necessarily an extraripple design. As seen in Fig. 5 they could also have used an $N = 19$ extraripple filter with
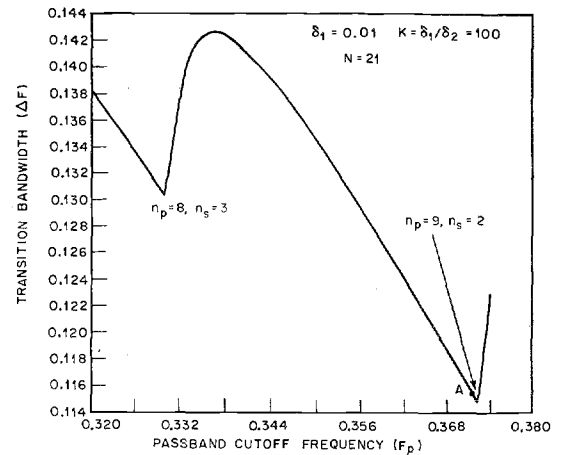
$F_p = 0.358$, $\Delta F = 0.1266$ as a counterexample. This filter also has a smaller $\Delta F$ than their $N = 21$ extraripple filter.

## Conclusions

This paper has attempted to explain several of the properties of the transition bandwidth of optimum linear-phase low-pass filters. The behavior of transition bandwidth as a function of passband cutoff frequency, and the ratio of passband to stopband ripple has been illustrated using several examples. Finally, the counterexample of Hofstetter et al. is explained in terms of the above theory.

## 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] L. R. Rabiner, B. Gold, and C. A. McGonegal, "An approach to the approximation problem for nonrecursive digital filters," IEEE Trans. Audio Electroacoust. (Special Issue on Digital Filtering), vol. AU-18, pp. 83–106, June 1970.
[3] O. Herrmann, "On the design of nonrecursive digital filters with linear phase," Electron. Lett., vol. 6, no. 11, pp. 328–329, 1970.
[4] E. Hofstetter, A. V. Oppenheim, and J. Siegel, "A new technique for the design of nonrecursive digital filters," in Proc. 5th Annu. Princeton Conf. Information Sciences and Systems, pp. 64–72, 1971.
[5] T. W. Parks and J. H. McClellan, "Chebyshev approximation for nonrecursive digital filters with linear phase," IEEE Trans. Circuit Theory, vol. CT-19, pp. 189–194, Mar. 1972.
[6] L. R. Rabiner, "The design of finite impulse response digital filters using linear programming techniques," Bell Syst. Tech. J., vol. 51, pp. 1177–1198, July–Aug. 1972.
[7] E. Hofstetter, A. Oppenheim, and J. Siegel, "On optimum nonrecursive digital filters," presented at the 9th Annu. Allerton Conf. Circuit and System Theory, Oct. 1971.