Minimax design of recursive digital filters

C. CHARALAMBOUS
(Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, Canada)

The application of two new algorithms for minimax optimization due to Charalambous and Bandler is investigated. The application is to the problem of finding the coefficients of a recursive digital filter to meet arbitrary specifications of the magnitude or the group delay characteristics. Unlike the original minimax algorithm due to Bandler and Charalambous in which a sequence of least pth optimizations as p tends to infinity is taken, the two new algorithms do not require the value of p to do this. Instead, a sequence of least pth optimization problems is constructed with *finite* values of p in the range $1 < p < \infty$. A criterion is given under which the order of the filter can be increased by growing filter sections. A general computer program has been developed, based on the ideas presented.

(Received on the 11th June 1973)

1 INTRODUCTION

A number of papers have been published on the design of recursive digital filters[1-7] of which the last three deal with iterative methods of design.

The Steiglitz approach[5] uses the optimization algorithm due to Fletcher-Powell[8] to minimize a least square error criterion in the frequency domain.

Deczky[6] uses the Fletcher-Powell algorithm to optimize a least pth objective function. The largest value of p reported in his paper was ten.

Bandler and Charalambous[9,10] proposed a generalized least pth objective function which alleviates the ill-conditioning resulting from the use of extremely large values of p. Bandler and Bardakjian[7] used this objective function in conjunction with the algorithms due to Fletcher-Powell and Fletcher[11] in digital filter design with values of p up to 10 000.

Very recently, Charalambous and Bandler[10,12] proposed two new algorithms for minimax approximation. Unlike their original algorithm in which a sequence of least pth optimizations as $p \to \infty$ is taken, the two new algorithms do not require the value of p to tend to infinity. Instead a sequence of least pth optimization problems with *finite* values of p in the range $1 < p < \infty$ is constructed. In their paper, they prove that both algorithms will converge to the minimax optimum.

This paper describes the application of the two new minimax algorithms in choosing the coefficients of a recursive digital filter to meet arbitrary specifications of the magnitude or the group delay characteristics. Based on the generalized least square optimum, the order of the filter can be increased by growing filter sections to meet the design specifications. The pole inversion technique due to Steiglitz[5] was used to meet the stability constraints. The program written is an extension of the program due to Popović[13]. The digital computer used for all the numerical results was a CDC 6400.

2 THEORY

2.1 *The form of the transfer function H(z)*

The transfer function of a recursive digital filter under consideration, is given by

$$H(z) = A \prod_{k=1}^{K} \frac{1 + a_k z^{-1} + b_k z^{-2}}{1 + c_k z^{-1} + d_k z^{-2}} \tag{1}$$

and can be realized in a cascade form[5] which is the most commonly used form of implementation. For stability of the filter, the poles of the transfer function should lie within the unit circle in the z-domain.

2.2 *The error functions*

Define real error functions related to the 'upper' and 'lower' specifications, respectively, as follows:

$$e_u(\underline{\phi}, \psi) \triangleq w_u(\psi)\,(F(\underline{\phi}, \psi) - S_u(\psi)) \tag{2}$$

$$e_l(\underline{\phi}, \psi) \triangleq w_l(\psi)\,(F(\underline{\phi}, \psi) - S_l(\psi)) \tag{3}$$

where

$F(\underline{\phi}, \psi)$ is the approximating function (actual response).
$S_u(\psi)$ is an upper specified function (desired response bound).
$S_l(\psi)$ is a lower specified function (desired response bound).
$w_u(\psi)$ is an upper positive weighting function.
$w_l(\psi)$ is a lower positive weighting function.
$\underline{\phi}$ is a vector containing the n independent parameters.
ψ is an independent variable (e.g., frequency or time).

If $S_u = S_l = S$, $w_u = w_l = w$, then we have a single specified function and $e_u = e_l = e$.

In practice, we will evaluate all the functions at a finite, discrete set of values of ψ taken from one or more closed intervals. Therefore, we will let

$$e_{ui}(\underline{\phi}) = e_u(\underline{\phi}, \psi_i) \qquad i \in I_u \tag{4}$$

$$e_{li}(\underline{\phi}) = e_l(\underline{\phi}, \psi_i) \qquad i \in I_l \tag{5}$$

I_u and I_l are appropriate index sets.

Therefore, for our problem we have

$$z_i = e^{j\psi_i \pi} \tag{6}$$

where

$$\psi_i = \frac{2f_i}{f_s} \tag{7}$$

f_s is the sampling frequency, f_i is the value of the frequency at the ith point and

$$\underline{\phi} = [a_1 b_1 c_1 d_1 \ldots a_K b_K c_K d_K A]^T \tag{8}$$

Thus,

$$n = 4K + 1.$$

2.3 The approximating function

The approximating function in our case will be either the amplitude or the group delay of the digital filter. From Equation (1)

$$H(\underline{\phi}, e^{j\theta}) = A \prod_{k=1}^{K} \left(\frac{a_k + (1 + b_k)\cos\theta + j(1 - b_k)\sin\theta}{c_k + (1 + d_k)\cos\theta + j(1 - d_k)\sin\theta} \right) \tag{9}$$

where $\theta = \pi\psi$. After algebraic manipulation, the amplitude of the filter is given by

$$|H(\underline{\phi}, \theta)| = A \prod_{k=1}^{K} \frac{N_k}{D_k}$$

$$= A \prod_{k=1}^{K} \left(\frac{1 + a_k^2 + b_k^2 + 2b_k(2\cos^2\theta - 1) + 2a_k(1 + b_k)\cos\theta}{1 + c_k^2 + d_k^2 + 2d_k(2\cos^2\theta - 1) + 2c_k(1 + d_k)\cos\theta} \right)^{1/2} \tag{10}$$

The group delay is given by

$$\tau(\underline{\phi}, \theta) = -\frac{\mathrm{d}}{\mathrm{d}\theta} [\angle H(\underline{\phi}, e^{j\theta})]$$

$$= -\sum_{k=1}^{K} \frac{\mathrm{d}}{\mathrm{d}\theta} \left(\tan^{-1} \left(\frac{(1 - b_k)\sin\theta}{a_k + (1 + b_k)\cos\theta} \right) \right)$$

$$+ \sum_{k=1}^{K} \frac{\mathrm{d}}{\mathrm{d}\theta} \left(\tan^{-1} \left(\frac{(1 - d_k)\sin\theta}{c_k + (1 + d_k)\cos\theta} \right) \right)$$

$$= -\sum_{k=1}^{K} \frac{1 - b_k^2 + a_k(1 - b_k)\cos\theta}{1 + a_k^2 + b_k^2 + 2b_k(2\cos^2\theta - 1) + 2a_k(1 + b_k)\cos\theta}$$

$$+ \sum_{k=1}^{K} \frac{1 - d_k^2 + c_k(1 - d_k)\cos\theta}{1 + c_k^2 + d_k^2 + 2d_k(2\cos^2\theta - 1) + 2c_k(1 + d_k)\cos\theta} \tag{11}$$

2.4 Variation of $|H(\underline{\phi}, \theta)|$ with respect to the parameters

Let us see how the amplitude of the digital filter varies with respect to the parameters.

(a)
$$\frac{\partial|H(\underline{\phi}, \theta)|}{\partial a_j} = A \frac{\prod_{k=1, k \neq j}^{K} N_k}{\prod_{k=1}^{K} D_k} \cdot \frac{\partial N_j}{\partial a_j}$$

$$= \frac{|H(\underline{\phi}, \theta)|}{N_j^2} [a_j + (1 + b_j)\cos\theta] \tag{12}$$

$$\frac{\partial^2 |H(\underline{\phi}, \theta)|}{\partial a_j^2} = \frac{|H(\underline{\phi}, \theta)|}{N_j^4} (1 - b_j)^2 \sin^2\theta \geqslant 0 \tag{13}$$

Equation (13) shows that $|H(\underline{\phi}, \theta)|$ is convex as a function of $a_j (1 \leqslant j \leqslant K)$.

(b)
$$\frac{\partial|H(\underline{\phi}, \theta)|}{\partial b_j} = \frac{|H(\underline{\phi}, \theta)|}{N_j^2} [-1 + 2\cos^2\theta + a_j\cos\theta + b_j] \tag{14}$$

$$\frac{\partial^2 |H(\underline{\phi}, \theta)|}{\partial b_j^2} = \frac{|H(\underline{\phi}, \theta)|}{N_j^4} (a_j + 2\cos\theta)^2 \sin^2\theta \tag{15}$$

Equation (15) shows that $|H(\underline{\phi}, \theta)|$ is *convex* as a function of $b_j (1 \leqslant j \leqslant K)$.

(c)
$$\frac{\partial|H(\underline{\phi}, \theta)|}{\partial c_j} = -\frac{|H(\underline{\phi}, \theta)|}{D_j^2} [c_j + (1 + d_j)\cos\theta] \tag{16}$$

$$\frac{\partial^2 |H(\underline{\phi}, \theta)|}{\partial c_j^2} = \frac{|H(\underline{\phi}, \theta)|}{D_j^4} [2(c_j + (1 + d_j)\cos\theta)^2 - (1 - d_j)^2 \sin\theta^2] \tag{17}$$

From Equation (16) it can be seen that the gradient of $|H(\underline{\phi}, \theta)|$ with respect to c_j changes sign only once, and at the point where this change occurs the second derivative is negative semi-definite. This shows that we have a pseudo-concave function.

(d)
$$\frac{\partial|H(\underline{\phi}, \theta)|}{\partial d_j} = -\frac{|H(\underline{\phi}, \theta)|}{D_j^2} [-1 + 2\cos^2\theta + c_j\cos\theta + d_j] \tag{18}$$

$$\frac{\partial^2 |H(\phi, \theta)|}{\partial d_j^2} = \frac{|H(\phi, \theta)|}{D_j^4} [3(-1 + 2\cos^2\theta + c_j \cos\theta + d_j)^2 - D_j^2] \quad (19)$$

As in case (c) we have a pseudoconcave function. The author has heard in a private communication with Bandler and Bardakjian that they reached similar results[14].

2.5 Minimax objective

For the minimax approximation problem we want to minimize the following objective function

$$M_f(\underline{\phi}) = \max_{i,j} [e_{ui}(\underline{\phi}), -e_{lj}(\underline{\phi})]\ i \epsilon I_u,\ j \epsilon I_l \quad (20)$$

It is important to note that if $M_f(\underline{\phi}) > 0$, then the specifications are violated and if $M_f(\underline{\phi}) < 0$, then they are satisfied. If $M_f(\underline{\phi}) = 0$, then the specifications are just met. Let us define for convenience,

$$f_i(\underline{\phi}) \triangleq \begin{cases} e_{ui}(\underline{\phi}) & i \in I_u \\ -e_{li}(\underline{\phi}) & i \in I_l \end{cases} \quad (21)$$

$$I \triangleq I_u \cup I_l \quad (22)$$

then our objective function becomes

$$M_f(\underline{\phi}) = \max_{i \in I} f_i(\underline{\phi}) \quad (23)$$

The problem of minimax optimization consists of finding a point $\overset{v}{\underline{\phi}}$ such that

$$M_f(\overset{v}{\underline{\phi}}) \leqslant M_f(\underline{\phi}) \quad (24)$$

for all points $\underline{\phi}$, at least in the neighbourhood of $\overset{v}{\underline{\phi}}$. Very recently Charalambous and Bandler proposed two algorithms to solve this problem[10,12].

2.5.1 Algorithm 1

1. Assume the starting point $\underline{\phi}^\circ$ is given; set $\xi^1 = \min[0, M_f(\underline{\phi}^\circ)]$, $r = 1$.
2. Minimize, with respect to $\underline{\phi}$, the function

$$U(\underline{\phi}, \xi^r) = M(\underline{\phi}, \xi^r) \left(\sum_{i \epsilon L} \left(\frac{f_i(\underline{\phi}) - \xi^r}{M(\underline{\phi}, \xi^r)} \right)^q \right)^{1/q} \quad \text{for } M(\underline{\phi}, \xi^r) \neq 0$$

$$= 0 \text{ for } M(\underline{\phi}, \xi^r) = 0 \quad (25)$$

where

$$M(\underline{\phi}, \xi^r) \triangleq \max_{i \epsilon I} (f_i(\underline{\phi}) - \xi^r) = M_f(\underline{\phi}) - \xi^r \quad (26)$$

$$q = p \operatorname{sign} M(\underline{\phi}, \xi^r) \begin{cases} p > 1 \text{ if } M(\underline{\phi}, \xi^r) > 0 \\ p \geqslant 1 \text{ if } M(\underline{\phi}, \xi^r) < 0 \end{cases} \quad (27)$$

and

$$L = \begin{cases} J(\underline{\phi}, \xi^r) \triangleq \{ i | f_i(\underline{\phi}) - \xi^r \geqslant 0,\ i \epsilon I \} \text{ if } M(\underline{\phi}, \xi^r) > 0 \\ I \text{ if } M(\underline{\phi}, \xi^r) < 0 \end{cases} \quad (28)$$

3. Let $\overset{vr}{\underline{\phi}}$ denote the optimum parameter vector at the rth step.

   Set

$$\xi^{r+1} = M_f(\overset{vr}{\underline{\phi}}) \quad (29)$$

4. Convergence criterion: If $|\xi^{r+1} - \xi^r| < \eta$, stop; else set $r = r + 1$ and go to 2. η is a small prescribed positive number.

2.5.2 Algorithm 2

1. As in algorithm 1
2. As in algorithm 1
3. If $M(\overset{v}{\underline{\phi}}{}^r, \xi^r) < 0$ go to 4; else set

$$\xi^{r+1} = \xi^r + \lambda^r M(\overset{v}{\underline{\phi}}{}^r, \xi^r)$$
$$= (1 - \lambda^r)\xi^r + \lambda^r M_f(\overset{v}{\underline{\phi}}{}^r) \quad (30)$$

   where

$$0 < \lambda^r < 1 \quad (31)$$

   and go to 5.

4. Set

$$\xi^{r+1} = M_f(\overset{v}{\underline{\phi}}{}^r)$$

5. Convergence criterion: If $|\xi^{r+1} - \xi^r| < \eta$, stop; else set $r = r + 1$ and go to 2.

Charalambous and Bandler proved that both algorithms will converge to the minimax optimum. It is important to note that for both algorithms the value of p can be kept constant in the range $1 < p < \infty$ unlike the usual case where p has to tend to infinity.

The gradient vector of the objective function with respect ϕ is:

$$\underline{\nabla} U(\underline{\phi}, \xi^r) = \left(\sum_{i \epsilon L} \left(\frac{f_i(\underline{\phi}) - \xi^r}{M(\underline{\phi}, \xi^r)} \right)^q \right)^{1/q-1} \times \sum_{i \epsilon L} \left(\frac{f_i(\underline{\phi}) - \xi^r}{M(\underline{\phi}, \xi^r)} \right)^{q-1} \underline{\nabla} f_i(\underline{\phi}) \quad (32)$$

$$\ldots \text{ for } M(\underline{\phi}, \xi^r) \neq 0$$

where

$$\underline{\nabla} \triangleq \left[\frac{\partial}{\partial \phi_1}\ \frac{\partial}{\partial \phi_2} \cdots \frac{\partial}{\partial \phi_n} \right]^T \quad (33)$$

If $f_i(\underline{\phi})$ for $i \in I$ are continuous with continuous first partial derivatives then the objective function Equation (25) is continuous with continuous first partial derivatives, except when $M(\underline{\phi}, \xi^r) = 0$ and two or more maxima are equal, in which case U is continuous but the gradients are discontinuous and therefore very efficient gradient methods can be used to optimize Equation (25), such as the one proposed by Fletcher[11].

Algorithm 2 is the same as algorithm 1 as long as

$$M_f(\overset{v}{\underline{\phi}}) < 0.$$

In algorithm 1 we try to push the maximum away from the level ξ^r at the rth step and in algorithm 2 we try to predict the value of

$$M_f(\overset{v}{\underline{\phi}})$$

by increasing the value of ξ^r from zero appropriately. Since

$$M(\overset{v}{\underline{\phi}}{}^r, \xi^{r+1}) = 0$$

when we are using algorithm 1, step 3 of algorithm 1 and step 4 of algorithm 2 are replaced by:

$$\xi^{r+1} = M_f(\overset{v}{\underline{\phi}}{}^r) + \epsilon \qquad (34)$$

where ϵ is a small number. This is done to avoid the possibility of discontinuous first partial derivatives. (Also $\xi^1 = \min[0, M_f(\underline{\phi}^0) + \epsilon]$.) It is important to note that in algorithm 2 we might not have to consider all the simple points, in which case time is saved in computing the gradient vector.

3 QUALITATIVE DESCRIPTION OF THE MINIMAX ALGORITHMS

Consider a system of m real nonlinear functions

$$f_i(\underline{\phi}), \quad i \in \mathrm{I}$$

Let $M_f(\underline{\phi})$ be the maximum of this set of functions at the point $\underline{\phi}$. Suppose that the original functions are shifted by a margin ξ. Then, we obtain a new set of functions

$$f_i'(\underline{\phi}, \xi) = f_i(\underline{\phi}) - \xi, \quad i \in \mathrm{I}$$

The maximum of the new set of functions is denoted by $M(\underline{\phi}, \xi)$ $(=M_f(\underline{\phi}) - \xi)$. The least pth objective function defined by Equation 25 is formed by using only the nonnegative set of the new functions when $M(\underline{\phi}, \xi) > 0$ (Index set J), and by using all the new set of functions when $M(\underline{\phi}, \xi) < 0$ (Index set I).

It is important to note that the value of ξ does not affect the minimax optimum but can play an important role for the optimum point of Equation 25. For example, if $\xi = M_f(\overset{v}{\underline{\phi}})$ where $\overset{v}{\underline{\phi}}$ is the minimax optimum, then, any finite value of p will yield the minimax solution (obviously in this case $M(\underline{\phi}, \xi) \geqslant 0$ for any $\underline{\phi}$ and therefore $q = p$. Also, at the optimum point $U = 0$). For both algorithms presented we try to force ξ to tend to $M_f(\overset{v}{\underline{\phi}})$ as the sequence of optimizations increases. In algorithm 1 for $r \geqslant 2$, we set the new value of ξ equal to the value of M_f at the previous optimum reached. Then, by minimizing Equation 25 we try to push M_f away from the artificial level ξ. We do this repeatedly until, we cannot decrease M_f any further, which means that we are at the minimax optimum.

In this case $\xi = M_f(\overset{v}{\underline{\phi}})$. In algorithm 2 we try to predict the level $M_f(\overset{v}{\underline{\phi}})$ by increasing ξ from zero appropriately (for algorithm 2 we have to assume that $M_f(\overset{v}{\underline{\phi}}) \geqslant 0$)). Suppose that $r = 1$, then the value of $M_f(\overset{v}{\underline{\phi}})$ will lie between zero and $M_f(\overset{v}{\underline{\phi}}{}^1)$ and therefore we set ξ^2 between zero and $M_f(\overset{v}{\underline{\phi}}{}^1)$ (see Equation 30). Suppose that we minimize Equation 25 with this new value of $\xi (= \xi^2)$. If the objective function (Equation 25) becomes negative then we proceed with algorithm 1, otherwise the value of $M_f(\overset{v}{\underline{\phi}})$ will lie between ξ^2 and $M_f(\overset{v}{\underline{\phi}}{}^2)$. Therefore the value of ξ^3 is set between ξ^2 and $M_f(\overset{v}{\underline{\phi}}{}^2)$. Assuming that we stay with algorithm 1 we can see that ξ increases all the time. When, at the optimum point of Equation 25, the objective function U is zero, it means that $\xi = M_f(\overset{v}{\underline{\phi}})$. A detailed mathematical proof of the convergence of the algorithms is given in Reference 12.

4 STABILITY

The poles of the transfer function $H(\underline{\phi}, z)$ must be inside the unit circle in the z-domain, in order for the filter to be stable. Suppose that $H(\underline{\phi}, z)$ has a real pole at $z = a$. Consider the function given in Reference 5

$$P(z) = \frac{z - a}{z - 1/a} \qquad (35)$$

then

$$|P(z)| = |a| \text{ when } |z| = 1. \qquad (36)$$

If we have complex poles these must occur in pairs as complex conjugates and the function

$$P(z) = \frac{(z - z_p)(z - z_p^*)}{(z - 1/z_p)(z - 1/z_p^*)} = \left(\frac{z - z_p}{z - 1/z_p^*}\right)\left(\frac{z - z_p^*}{z - 1/z_p}\right) \qquad (37)$$

has the following property:

$$|P(z)| = |z_p|^2 \text{ when } |z| = 1. \qquad (38)$$

where superscript * is used to denote the complex conjugate.

Equations (35)–(38) show that the magnitude characteristic of the filter remains unchanged by the pole inversion technique, as long as the value of A is replaced by the value

$$\frac{\text{Value of } A \text{ before pole inversion}}{\Pi\ |\text{poles outside the unit circle}|} \qquad (39)$$

5 EXAMPLES

For all the examples considered, we used the value of $p = 2$ for $r = 1$ and $p = 10$ for $r \geqslant 2$. For algorithm 2, $\lambda = 0.55$

for $r = 2$, $\lambda = 0.6$ for $r \geqslant 3$. The value of ϵ used is 10^{-8}. Stability constraints are checked before we increase the value of r and before termination. The algorithm due to Fletcher[11] was used in all the examples.

5.1 Example 1: A wide-band differentiator

Suppose that we have the following design specification for the magnitude characteristic

$$S(\psi) = \psi, \quad 0 \leqslant \psi \leqslant 1$$

$w(\psi) = 1$. We want to find the minimax optimum for a one-section design.

Solution: It was decided to use 25 sample points of ψ, namely,

$$\psi = 0.0, 0.85\ (0.05); \qquad i = 1, \ldots, 18$$

$$\psi = 0.87, 0.97\ (0.02); \qquad i = 19, \ldots, 24$$

$$\psi = 1.0 \qquad ; \qquad i = 25$$

Starting point:

$$\underline{\phi}^0 = [-0.33 \quad -0.67 \quad 0.86 \quad 0.102 \quad 0.366]^T$$

Tables 1 and 2 show the progress of algorithm 1 and algorithm 2, respectively.

From Table 1 it can be seen that for $r \geqslant 2$, $q = -10$. Also, it is interesting to note how M_f^r decreases and asymptotically approaches the minimax optimum value.

From Table 2 it can be seen that for all values of r the value of q is positive; which means that we stayed with algorithm 2. Also, note how M_f^r decreases and how ξ^r increases, and both tend to the same limit.

From Tables 1 and 2 it can be seen that, although after five optimizations the value of M_f for both algorithms agrees to three significant figures (6.56×10^{-3}), the optimum parameter vectors are different. This means that we might have two minimax optima. After perturbing the optimum parameter vectors, we returned to the same points and therefore concluded that there are two minimax optima, but with the same value of

$$M_f(\overset{v}{\underline{\phi}}).$$

At the optimum points reached, the poles and zeros have the following values:

Algorithm 1: Poles: −0.16852, −0.77147
Zeros: 1.0217, −0.74072

Algorithm 1: Poles: −0.16852, −0.77147
Zeros: 1.0217, −0.74072

TABLE 1. The progress of algorithm 1

p	$\xi^r \times 10^3$	Solution $\overset{v}{\phi}{}^r$	Values of ψ at which maxima occur	Absolute maxima $(\times 10^3)$	$M_f^r(\times 10^3)$	Number of function evaluations
		−0.33494	0.0	3.1570×10^{-5}		
		−0.66506	0.25	3.0477		
2	0	0.85374	0.6	3.8683	11.381	18
		0.10189	0.85	4.9124		
		0.36622	0.95	7.4772		
			1	11.381		
		−0.25305	0.0	4.1955		
		−0.72341	0.25	6.1265		
−10	11.381	0.93126	0.65	6.3264	6.8702	138
		0.12687	0.89	6.2614		
		0.36683	0.97	6.8702		
			1.0	6.7455		
		−0.28075	0.0	6.4091		
		−0.75607	0.25	6.5449		
−10	6.8702	0.93963	0.65	6.5370	6.5712	79
		0.12988	0.89	6.5519		
		0.36022	0.97	6.5712		
			1.0	6.5622		
		−0.28092	0.0	6.5501		
		−0.75673	0.25	6.5563		
−10	6.5712	0.93998	0.65	6.5563	6.5573	77
		0.13000	0.89	6.5564		
		0.36013	0.97	6.5573		
			1.0	6.5569		
		−0.28093	0.0	6.5565		
		−0.75675	0.25	6.5567		
−10	6.5573	0.9400	0.65	6.5567	6.5568	82
		0.13001	0.89	6.5567		
		0.36012	0.97	6.5568		
			1.0	6.5567		

TABLE 2. The progress of algorithm 2

p	$\xi^r \times 10^3$	Solution $\overset{v}{\underline{\phi}}{}^r$	Values of ψ at which maxima occur	Absolute maxima $(\times 10^3)$	$M_f^r(\times 10^3)$	Number of function evaluations
		−0.33494	0.	3.1570×10^{-5}		
		−0.66506	0.25	3.0477		
2	0	0.85374	0.6	3.8683	11.381	18
		0.10189	0.85	4.9124		
		0.36622	0.95	7.4772		
			1.	11.381		
		−0.23879	0.	6.4448		
		−0.72496	0.25	6.5438		
10	6.2598	0.93961	0.65	6.5351	6.5732	90
		0.12987	0.89	6.5501		
		0.36786	0.97	6.5732		
			1.	6.5621		
		−0.23833	0.	6.5158		
		−0.72502	0.25	6.5525		
10	6.4478	0.93988	0.65	6.5510	6.5622	29
		0.12997	0.89	6.5537		
		0.36790	0.97	6.5622		
			1.	6.5582		
		−0.23817	0.	6.5417		
		−0.72502	0.25	6.5553		
10	6.5165	0.93996	0.65	6.5551	6.5586	23
		0.1300	0.89	6.5554		
		0.36791	0.97	6.5586		
			1.	6.5572		
		−0.23813	0.	6.5506		
		−0.72502	0.25	6.5561		
10	6.5418	0.93998	0.65	6.5561	6.5574	21
		0.13000	0.89	6.5562		
		0.36792	0.97	6.5571		
			1.	6.5574		

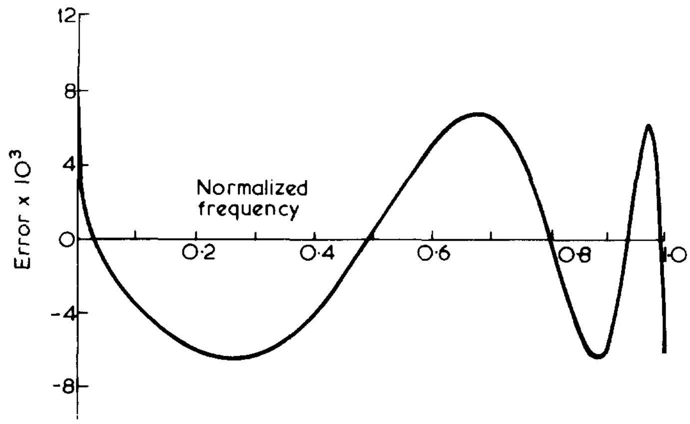

FIGURE 1. Optimum error curve for Example 1, a wide-band differentiator.

Algorithm 2: Poles: −0.16852, −0.77146
Zeros: 0.97883, −0.74070

If no excess phase is required, the optimum reached by algorithm 2 should be used, because the zeros are inside the unit circle in the z-domain. Note that the poles are the same, for both optima reached.

Figure 1 shows the optimum error curve, after five sequence of optimizations (the error curves for both algorithms are indistinguishable after five steps). Note that 101 equidistant values of ψ were used for the error evaluation and plotting.

5.2 *Example 2: A linear discriminator*

Suppose that we have the following design specification for the magnitude characteristic[5]

$$S(\psi) = |1 - 2\psi|, 0 \leqslant \psi \leqslant 1$$

$w(\psi) = 1$ and we want to find the minimax optimum for two-section design.

Solution: It was decided to use 41 sample points of ψ, namely,

$\psi = 0.0, 0.05\ (0.01)$; $i = 1, \ldots, 6$

$\psi = 0.07, 0.46\ (0.03)$; $i = 7, \ldots, 20$

$\psi = 0.5, 0.54\ (0.04)$; $i = 21, 22$

$\psi = 0.57, 0.67\ (0.1)$; $i = 23, 24$

$\psi = 0.63, 0.93 (0.03)$; $i = 25, 35$

$\psi = 0.95, 1.0\ (0.01)$; $i = 36, 41$

Starting point:

$$\underline{\phi}^0 = [0.\ \ 1.\ \ 0.\ \ -0.15\ \ 0.\ \ -0.68\ \ 0.\ \ -0.72\ \ 0.37]^T$$

TABLE 3. The progress of algorithm 1 for a linear discriminator

q	$M_f^r \times 10^3$	Number of function evaluations
2	10.598	23
−10	6.3307	82
−10	6.1870	92
−10	6.1854	110

Tables 3 and 4 show the progress of algorithm 1 and algorithm 2 respectively. The same comments as given in Section 5.1 hold for the progress of both algorithms. We have two minimax optima, with the same value of

$$\overset{v}{M}_f(\underline{\phi}).$$

At the optimum points reached, the poles and zeros have the following values:

Algorithm 1: Poles: 0.40715, −0.40715
0.87592, −0.87591
Zeros: $1.7985 \times 10^{-6} \pm j\ 1.0101$
0.85737, −0.85736

Algorithm 2: Poles: 0.40715, −0.40715
0.87592, −0.87592
Zeros: $1.4475 \times 10^{-6} \pm j\ 0.98998$
0.85737, −0.85737

Again, as in Example 1, if no excess phase is required the optimum reached by algorithm 2 should be used.

Figure 2 shows the optimum error curve and Figure 3 the specified function and the approximating function at the optimum points. 0.201 equidistant values of ψ were

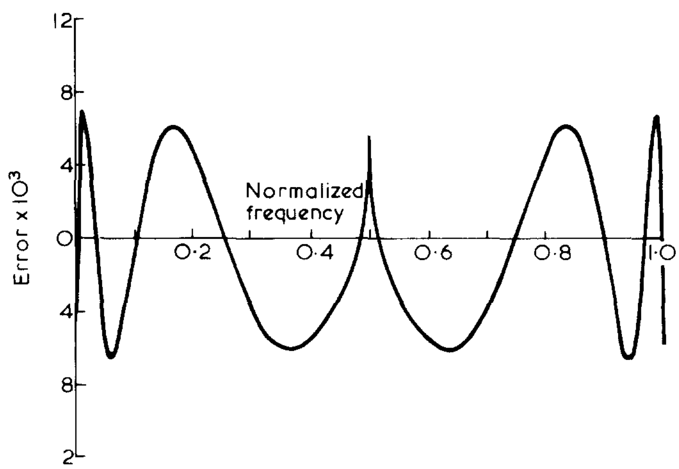

FIGURE 2. Optimum error curve for Example 2, a linear discriminator.

TABLE 4. The progress of algorithm 2 for the same linear discriminator as in Table 3.

q	$\xi^r \times 10^3$	$M_f^r \times 10^3$	Number of function evaluations
2	0		
10	5.8291	6.1902	113
10	6.0458	6.1872	30
10	6.1306	6.1860	27
10	6.1639	6.1856	29

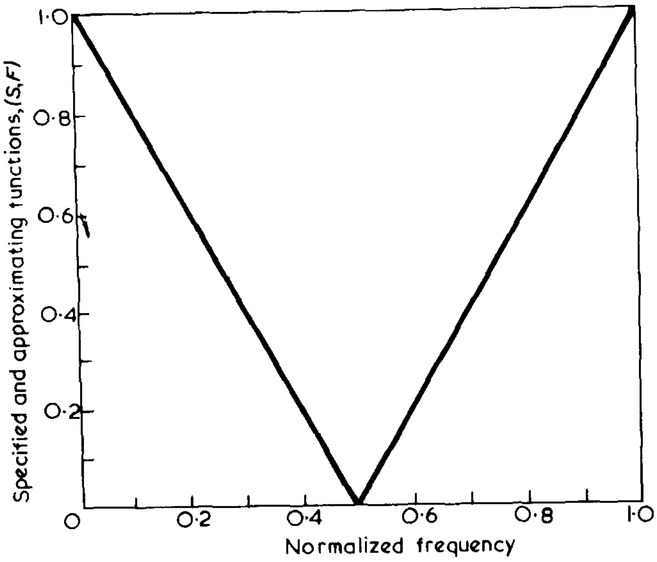

FIGURE 3. Optimized response for Example 2, a linear discriminator.

used for the evaluation and plotting of the error curve and the approximating function. Algorithm 1 took about 73 seconds to reach the optimum point while algorithm 2 took only 34 seconds.

5.3 Example 3: Low-pass filter

Design a low pass recursive digital filter of the cascade form, for a 10 KHz sampling rate, with the smallest number of sections, such that its magnitude characteristic will satisfy the following specification[7,15].

$f = 0{,}900\ (100);\quad S_u(f) = 1.1;\quad S_l(f) = 0.9$

$f = 1200;\quad S_u(f) = 0.1$

$f = 1500,\ 5000\ (500);\quad S_u(f) = 0.1.$

Solution: The specifications can be prescribed as:

$\psi = 0.,\ 0.18\ (0.02);\quad S_u(\psi_i) = 1.1,\ S_l(\psi_i) = 0.9;\ i = 1, . , 10$

$\psi = 0.24;\quad S_u(\psi_i) = 0.1;\quad i = 11$

$\psi = 0.3,\ 1.\ (0.1);\quad S_u(\psi_i) = 0.1;\quad i = 12, . , 19$

Before solving the above problem, it is worth mentioning a very important result concerning the objective function Equation (25). If the objective function corresponding to the optimum point of the first optimization ($r = 1$) with a particular value of p, is positive (negative), then, at the optimum point of any other value of p greater than one, the objective function will be positive (negative), even with $p = \infty$, i.e. minimax. Thus, even a least-square solution will tell us whether we can satisfy the specifications in the minimax sense.

By using the above important result we can solve our problem in the following way:

Find the least-square optimum ($r = 1$) with one-section. If the objective function is negative (i.e. the specification is satisfied) increase the value of p to ten (not necessarily) and carry on with algorithm 1. If the objective function is positive (specification violated), reset $\xi^1 = 0$, grow another section and do another least-square optimization. This procedure is repeated until we grow enough sections to satisfy the specification (assuming this is possible).

Starting from the point

$$\phi^0 = [0 \quad 1 \quad -1 \quad 0.5 \quad 0.1]^T \qquad (M_f(\phi^0) = 0.5)$$

and using the above ideas, the results shown in Table 5 were obtained. From this, it can be seen that after 30 function evaluations ($q = 2$) it was possible to realize that the minimax optimum will not satisfy the specification (without actually calculating minimax optimum point which is time-consuming). Growing another section, and resetting $\xi^1 = 0$, the objective function became negative after 26 more function evaluations (not shown in the table), which means that a two-section filter can satisfy the specification.

5.4 Example 4: All-pass filter

In this example, we consider the synthesis of a digital all-pass filter with prescribed group delay

$$S(\psi) = 15\psi, \quad 0.1 \leqslant \psi \leqslant 0.9$$

and $w(\psi) = 1/15\psi$.

Solution: The problem was discretized into 33 uniformly spaced sample points of ψ, namely,

$$\psi = 0.1,\ 0.9\ (0.025), \qquad i = 1, \ldots, 33$$

For all-pass digital filters

$$H(z) = A \prod_{k=1}^{K} \frac{b_k z^2 + a_k z + 1}{z^2 + a_k z + b_k} \tag{40}$$

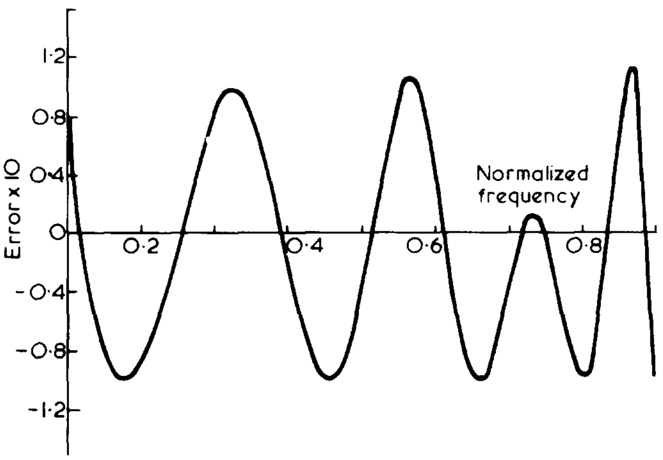

FIGURE 4. Optimum error curve for Example 4, an all-pass filter.

In this case, it can be easily shown that the group delay is given by

$$\tau = 2 \sum_{k=1}^{K} \frac{1 - b_k^2 + a_k(1 - b_k)\cos\theta}{1 + a_k^2 + b_k^2 + 2b_k(2\cos^2\theta - 1) + 2a_k(1 + b_k)\cos\theta} \tag{41}$$

The approximating function is given by

$$F(\underline{\phi}, \psi) = \tau(\underline{\phi}, \psi) - \tau_0 \tag{42}$$

where τ_0 is a constant group delay which is considered as a variable and

$$\underline{\phi} = [a_1 b_1 \ldots a_K b_K \tau_0]^T \tag{43}$$

Thus $n = 2K + 1$.

Using a four-section filter and starting from the point

$$[-0.55 \quad 0.39 \quad 0.33 \quad 0.5 \quad 1.05 \quad 0.58 \quad 1.52 \quad 0.67 \quad 1.12]^T$$

algorithm 1 and algorithm 2 generated the sequence shown in Tables 6 and 7 respectively. From Table 7, it can be seen that at the second step we switched from algorithm 2 to algorithm 1. This is because

$$\xi^2 > M_f(\overset{v}{\underline{\phi}}).$$

Figure 4 shows the optimum weighted error curve and Figure 5 the specified function and the approximating function at the optimum point. One hundred and sixty one equidistant values of ψ were used for plotting the functions shown in the figures. From Figure 4 it can be seen that there is a ripple which is not active at the maximum optimum. Algorithm 1 took about 82 seconds to reach the optimum point while algorithm 2 took only 47 seconds.

6 CONCLUSIONS

The two new algorithms for minimax approximation were found very efficient in optimizing recursive digital filters. Using the generalized least-square optimum, we immediately know whether or not our filter can satisfy the design specifications in the minimax sense. If it cannot, we

TABLE 5. The results obtained in trying to meet the design specification

q	$\xi^r \times 10^2$	Solution $\overset{v}{\underline{\phi}}{}^r$	$U(\overset{v}{\phi}{}^r, \xi^r) \times 10^2$	$M_f^r \times 10^2$	Number of function evaluations
2	0	−0.85815 1.00000 −1.50027 0.70905 0.15755	24.439	15.755	30
−2	0	−1.41237 1.00000 −1.57530 0.88292 0.61803 1.00000 1.43047 0.56271 0.02603	−1.6925	−6.2827	187
−10	−6.2827	−1.40496 1.00000 −1.58462 0.88793 −0.19570 1.00000 −1.45306 0.58177 0.03547	0.95070	−7.3536	212

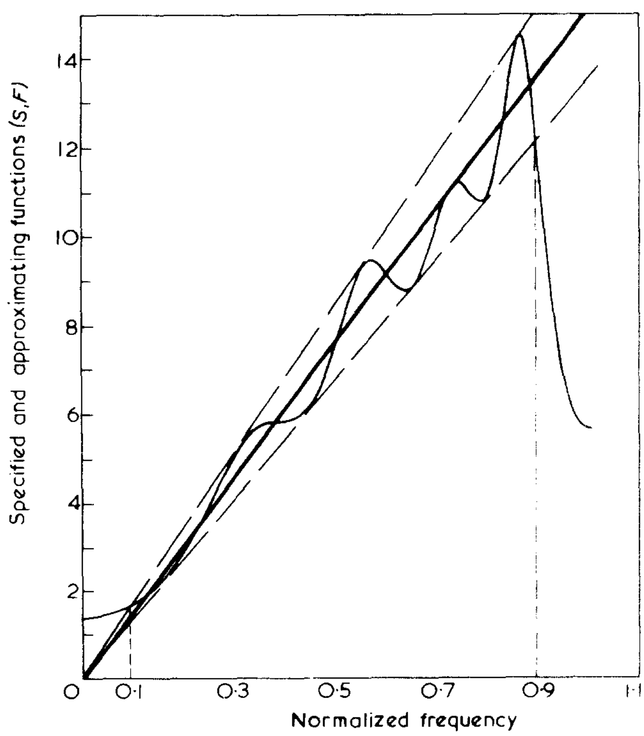

FIGURE 5. Optimized response for Example 4, an all-pass filter.

TABLE 6. Using a four section filter, a certain start point and algorithm 1, the following sequence is generated

q	$M_f^r \times 10$	Number of function evaluations
2	1.88818	16
−10	1.03057	48
−10	0.98686	61
−10	0.98400	94
−10	0.98381	101

$$\underline{\overset{v}{\phi}}^5 = \begin{bmatrix} -0.58452 \\ 0.39747 \\ 0.31472 \\ 0.53533 \\ 1.03848 \\ 0.57649 \\ 1.53043 \\ 0.68982 \\ 1.14618 \end{bmatrix}$$

TABLE 7. As for Table 6 but using algorithm 2

q	ξ^r	$M_f^r \times 10$	Number of function evaluations
2	0.	1.88818	16
−10	1.03850	0.98738	140
−10	0.98738	0.98403	53

$$\underline{\overset{v}{\phi}}^3 = \begin{bmatrix} -0.58445 \\ 0.39746 \\ 0.31473 \\ 0.53515 \\ 1.03849 \\ 0.57649 \\ 1.53040 \\ 0.68977 \\ 1.14636 \end{bmatrix}$$

can grow filter sections until the specifications are satisfied (assuming that this is possible). In practice, as the results indicate, only one optimization with $p = 2$, and two optimizations with $p = 10$ will give very good minimax results. For the examples considered, algorithm 2 took less time than algorithm 1 to approach very close to a minimax optimum. For a realistic problem, the time required for the algorithms to reach very close to a minimax optimum will be of the order of one minute of a CDC 6400 computer.

It is to be noted, that in Examples 1 and 2 we obtained two optima, having the same poles and the same $|H(\theta)|$, but with some of the zeros having different values. The reason for two optima is the fact that the value of $|H(\phi, \theta)|$ remains unchanged by inverting the zeros, as long as $\bar{A}$ is replaced by:

$$\text{Value of } A \text{ before zero inversion} \times \Pi \, |\text{zeros to be inverted}| \quad (44)$$

The zeros with different values are the inverse of each others (e.g. $0.97883 \approx 1/1.0217$ for example 1) and also Equation (44) is satisfied (e.g. $0.36792 \approx 0.36012 \times 1.0217$ for example 1).

7 ACKNOWLEDGEMENT

The author thanks Dr J. W. Bandler, Dr S. S. Haykim, B. L. Bardakjian for useful discussions and J. R. Popović, for without her assistance the developed computer package would not have been possible. This work was supported by the National Research Council of Canada.

REFERENCES

1 Kaiser, J. F. 'Digital Filters', in *Systems Analysis by Digital Computer*, F. F. Kuo and J. F. Kaiser, (Eds) John Wiley, New York (1966), Ch. 7.

2 Haykim, S. S. 'A unified treatment of recursive digital filtering', *IEEE Trans. Automatic Control*, Vol AC-17, (February 1972), pp 113–116

3 Sablatash, M. 'The state of the art in approximation techniques for digital filter design', *1971 Symposium on Digital Filtering*, Imperial College of Science and Technology, London (August 31–September 2) 1971

4 Helms, H. D. 'Digital filters with equiripple or minimax responses', *IEEE Trans, Audio Electroacoustics*, Vol AU-19, (March 1971) pp 87–93

5 Steiglitz, K. 'Computer-aided design of recursive digital filters', *IEEE Trans, Audio Electroacoustics*, Vol AU-18, (June 1970) pp 123–129

6 Deczky, A. G. 'Synthesis of recursive digital filters using the minimum p-error criterion', *IEEE Trans. Audio Electroacoustics*, Vol AU-20 (October 1972) pp 257–263

7 Bandler, J. W. and Bardakjian, B. L. 'Least *p*th optimization of recursive digital filters', IEEE *Trans Audio Electroacoustics* Vol AV-21 (October 1973) pp 460–470

8 Fletcher, R. and Powell, M. J. D. 'A rapidly convergent descent method for minimization', *Comput. J.*, Vol 6, (June 1963) pp 163–168

9 Bandler, J. W. and Charalambous, C. 'Practical least *p*th optimization of networks', *IEEE Trans. Microwave Theory Tech.*, Vol MTT-20 (December 1972) pp 834–840

10 Charalambous, C. 'Nonlinear least *p*th approximation and nonlinear programming with applications in the design of networks and systems', PhD Dissertion, McMaster University, Hamilton, Ontario, Canada (February 1973)

11 Fletcher, R. 'A new approach to variable metric algorithms', *Comput. J.*, Vol 13 (August, 1970) pp 317–322

12 Charalambous, C. and Bandler, J. W. 'New algorithms for minimax optimization', *Nonlinear Programming Symp., Washington, DC, USA* March 14th-16th, 1973

13 Popović, J. R. 'General programs for least *p*th and near minimax approximation', M.Eng.Dissert., McMaster University, Hamilton, Ontario, Canada (1972)

14 Bandler, J. W. and Bardakjian, B. L. 'Private communication', McMaster University, Hamilton, Ontario, Canada

15 Suk, M. and Mitra, S. K. 'Computer-aided design of digital filters with finite word length', IEEE *Trans Audio Electroacoustics* Vol AV-20 (December 1972) pp 356–363