LEAST p^{th} POWER DESIGN OF COMPLEX FIR 2-D FILTERS USING THE COMPLEX NEWTON METHOD

Theagenis J. Abatzoglou* and Amin G. Jaffer**

*Lockheed Research Laboratories
Palo Alto, California USA
**Hughes Aircraft Company
Fullerton, California USA

ABSTRACT

A design of 2-D complex FIR filters is proposed by minimizing the p^{th} power norm used to measure the deviation of the FIR filter response from a desired filter response. The solution of this problem cannot be obtained in closed form except for p=2; for arbitrary p>2 we present an approach which treats the problem from a complex variable point of view. An iterative scheme is described based on the complex Newton method to find the solution. It has the feature that, starting with p=2, the value of p is increased after each iteration. Because the objective function is convex any local extremum is the global minimum. Convergence can be attained after a moderate number of iterations. A characterization theorem for factorization of 2-D FIR filters in terms of 1-D filters is derived. This has strong implications for large order 2-D filter design. Two filter design examples are included.

1. INTRODUCTION

Complex coefficients 1-D FIR filters have recently been developed, [1], [2], for radar/sonar clutter supression and other applications where nonsymmetrical filter response is desired. Complex coefficient FIR 1-D filter design has been treated by a few authors; more notably Preuss derived the theory and algorithm for the Min-Max criterion, [1] and created a Remez-type algorithm. Jaffer in [2], uses a least squares to obtain the complex coefficients of a FIR filter in closed form. A characterization of affine phase complex FIR filters is also given in [2]. Using the least p^{th} power filter design is not new. Fahmy and Lampropoulos, [3], proposed a modified p^{th} power criterion to design real 2-D filters however their Newton iterates were not computed in closed form nor any comments were made relative to the behaviour of the solution for large values of p. More recently the Iterative Reweighted Least Squares (IRLS) technique has been applied, [4], [5], to compute the least p^{th} power filter design for real FIR filters. This approach is essentially a rearrangement of the Newton method and it is unusual in the sense that the power p is gradually increased in the successive iterations. Alkhairy et al, [6], developed an efficient algorithm for computing real 1-D FIR filters to approximate complex valued filter response in the minmax sense.

The emphasis of this paper is complex 2-D FIR filter design for 2<p and derivation of some of their characteristic properties. Our aim is the efficient computation of the least p^{th} power and use it as an approximation to the minmax design for which there is no characterization in 2D. We derive structure results as well as efficient algorithms for the filter formation. We use the complex Newton method, [7] to minimize the p^{th} power norm of the difference between the constructed and desired filter response. We treat the problem entirely from a complex variables point of view and do not decompose the filter coefficients into its real and imaginary components. This we believe adds elegance and compactness to our solution. The Newton method is implemented for increasing p, starting at 2=p and increasing by a constant factor on each iteration until the desired value of p is reached.. Thus successive Newton iterates can remain within their region of convergence with respect to the p^{th} power norm objective function. A factorization theorem is presented for 2-D FIR filters in terms of corresponding 1-D filters. This is valid for p=2, however we conjecture to be true for all p. This can be used to build large order 2-D filters. A few examples are presented on 2-D filter design with p=60 and nonfactorable weighting function. Here our proposed approach manages to equalize the ripple in both the stopband and passband.

2. LEAST p^{th} POWER DESIGN OF COMPLEX 2_D FIR FILTERS USING THE COMPLEX NEWTON METHOD

The filter design will be carried out without any constraints on the coefficients, however to assure the filter is linear phase, hermitian symmetry on the coefficients may be imposed, [8].

Let $z_D(f_1,f_2)$ be the desired frequency response of the filter, $z_D(f_1,f_2)$ can be expressed as

$$z_D(f_1,f_2) = a_D(f_1,f_2)e^{j\phi(f_1,f_2)}$$

where $a_D(f_1,f_2)$ and $e^{j\phi(f_1,f_2)}$ are the desired amplitude and frequency response of the filter. Then the filter is formed by minimizing the weighted p^{th} power error:

$$G_p(\mathbf{H}) = \int_0^1\int_0^1 W(f_1,f_2)\left|z_D(f_1,f_2) - H(f_1,f_2)\right|^p df_1 df_2$$

Where

$$H(f_1,f_2) = \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} h_{n,m} e^{-2\pi j n f_1} e^{-2\pi j m f_2}$$

be the frequency response of a 2-D complex FIR filter with coefficients $\{h_{n,m}\}_{n,m}$.

0-7803-2431-5/95 $4.00 © 1995 IEEE

$W(f_1, f_2)$ is a weighting function that balance tradeoffs between passbands and stopbands. Also usually $W(f_1, f_2) = 0$, on the transition bands; p is a power larger or equal to 2. When $p = 2$ we our criterion is the least squares design; as $p \Rightarrow \infty$ we get the minmax design. The filter frequency response can be expressed in terms of the filter coefficients as

$$H(f_1, f_2) = \vec{e}_{N,M}(f_1, f_2)^* \vec{h}$$

where

$$\vec{h} = \begin{bmatrix} h_{0,0} \\ h_{1,0} \\ \vdots \\ h_{N-1,0} \\ h_{0,1} \\ h_{1,1} \\ \vdots \\ h_{N-1,1} \\ \vdots \\ h_{0,M-1} \\ \vdots \\ h_{N-1,M-1} \end{bmatrix}; \quad \vec{e}_{N,M}(f_1, f_2) = \begin{bmatrix} 1 \\ e^{2\pi j f_1} \\ \vdots \\ e^{2\pi j (N-1) f_1} \\ e^{2\pi j f_2} \\ e^{2\pi j (f_1 + f_2)} \\ \vdots \\ e^{2\pi j ((N-1) f_1 + f_2)} \\ \vdots \\ e^{2\pi j ((M-1) f_2)} \\ \vdots \\ e^{2\pi j ((N-1) f_1 + (M-1) f_2)} \end{bmatrix}$$

It should be noted that the $(N(k-1)+n)^{th}$ entry is $e^{2\pi j((n-1) f_1 + (k-1) f_2)}$

2.1 The Case $p = 2$ (Least Squares Design)

The necessary and sufficient conditions for a complex 2-D FIR filter to have linear phase were derived in [8]. The least squares solution of this type of filter has been derived in [8] for the case that MN is even. A derivation for arbitrary M and N is given in [9].

Let

$$\vec{c} = \int_0^1\int_0^1 W(f_1, f_2) z_D(f_1, f_2)\, \vec{e}_{N,M}(f_1, f_2)\, df_1 df_2$$

$$E = \int_0^1\int_0^1 W(f_1, f_2)\, \vec{e}_{N,M}(f_1, f_2)\, \vec{e}_{N,M}(f_1, f_2)^*\, df_1 df_2$$

then the solution as given in [9], is

$$\vec{h} = \frac{1}{2} E^{-1}\left(\vec{c} + J_{NM}\bar{\vec{c}}\right)$$

Here J_{NM} is the exchange matrix of size NMxNM. From now on: $\bar{\vec{c}}$ is the complex conjugate of $\vec{c}$ and $\vec{c}^*$ is its complex conjugate transpose. When there are no constraints imposed on the filter response the least squares solution is obtained from :

$$\vec{h} = E^{-1}\vec{c}$$

It was shown in [2] for the 1-D case and in [9] for 2-D one that expressions are identical when the desired filter response has linear phase. It is also possible for the linear phase case to derive the solution in terms of the top half of the vector of filter coefficients, [8], [9].

The computation load of inverting the matrix E of size MNxMN can be substantial if MN is large; however as it has been observed in [8] and [9] that E is Toeplitz-block-Toeplitz and fast algorithms exist for inverting it or solving the system of equations: $E\vec{h} = \bar{\vec{c}}$ in $O(\mathbf{min}(M, N) N^2 M^2)$ flops.

We will elaborate further in the applicability of these results when p>2.

2.2 Case $2 < p$

This is an important area of research because a filter design for some 2<p gives filter characteristics in between p=2 and $p \Rightarrow \infty$. In particular for large p the characteristics of the filtere expected to resemble that of the minmax design, with equiripples in the passband and stopband. Thus selecting a solution for large p one expects approximation to the minmax solution whose characterization is rather abstract and very difficult to implement numerically.

Previously, several authors have treated p^{th} power minimization mainly in 1-D and for real FIR filters, [4] and [10]. The real 2-D case has been treated in part , [3], however the error function used was the sum of the p^{th} powers of the absolute values of real and imaginary components which is not equal to the magnitude of the error raised to the p^{th} power. Although the real Newton method was used here the Hessian was not computed in closed form nor any special properties of the structure of the Hessian matrix was adressed. Also the computational speed of the algorithm was not optimized.

Characterization of the minmax solution is by no means an easy matter. Even for the real 1D case its foundations lay on a deep theorem in approximation theory by Weierstrass. The complex 1D case was treated in [1] and it is already more complicated than the real one. Here due to the lack of a straighforward chracterization in 2D real or complex we propose the use of the least p^{th} power design to approximate minmax 2D filters.

A key advantage of the p^{th} power approach is that the objective function to be minimized is convex, twice differentiable. This impliesthat any extremeum is necessarily the global minimum. Thus a descent algorithm based on Newton's method can be made to converge to the true solution, by scaling back the update if necessary, no matter where it is started, [11]. The convergence rate is quadratic on a neighborhood of the solution.

The objective function to be minimized is

$$G_p(\vec{h}_L) = \int_0^1\int_0^1 W(f_1, f_2)\left|z_D(f_1, f_2) - \vec{e}_{N,M}(f_1, f_2)^* \vec{h}\right|^p df_1 df_2$$

$$= \int_0^1\int_0^1 W(f_1, f_2)\left[\Delta(\vec{h}_L)\bar{\Delta}(\vec{h}_L)\right]^{p/2} df_1 df_2$$

where

$$\Delta(\vec{h}) = z_D(f_1, f_2) - \vec{e}_{N,M}(f_1, f_2)^* \vec{h}$$

is the discrepancy between the desired and produced filter response.

Computation of complex first and second partials derivatives of $G_p(\vec{h})$ with respect to $\vec{h}$ are obtained next by using formulas from, [7],. These will be used in the implementation of the complex Newton method for finding the minimum of a function.

The Newton recursion for $\vec{h}$ say $\left\{\vec{h}^{(k)}\right\}_k$ is written as

$$\vec{h}^{(k+1)} = \vec{h}^{(k)} +$$

$$\left(\bar{F}_{\vec{h}^T\vec{h}} F_{\vec{h}^*\vec{h}}{}^{-1} F_{\vec{h}^T\vec{h}} - \bar{F}_{\vec{h}^*\vec{h}}\right)^{-1}\left(\bar{\nabla} G_{\vec{h}} - \bar{F}_{\vec{h}^T\vec{h}} F_{\vec{h}^*\vec{h}}{}^{-1} \nabla G_{\vec{h}}\right)$$

where the matrices and gradients of the second and first derivatives of $G_p(\vec{h})$ are defined by

$$\nabla G_{\bar{h}} = \frac{\partial G_p(\vec{h})^T}{\partial \vec{h}}\bigg|_{\vec{h}=\vec{h}^{(k)}}$$

$$F_{\vec{h}^T\vec{h}} = \frac{1}{2}\left(\frac{\partial^2 G_p(\vec{h})}{\partial \vec{h}^T \partial \vec{h}} + \frac{\partial^2 G_p(\vec{h})^T}{\partial \vec{h}^T \partial \vec{h}}\right)\bigg|_{\vec{h}=\vec{h}^{(k)}}$$

$$F_{\vec{h}^*\vec{h}} = \frac{\partial^2 G_p(\vec{h})^T}{\partial \vec{h}^* \partial \vec{h}}\bigg|_{\vec{h}=\vec{h}^{(k)}}$$

We call $F_{\vec{h}^*\vec{h}}$ and $F_{\vec{h}^T\vec{h}}$ the conjugated and unconjugated Hessians respectively. Here $\vec{h}$ and its complex conjugate $\bar{\vec{h}}$ are treated as independent complex vectors.

It was proved in [9] that in the p=2 case the unconstrained filter coefficient solution when the desired response is linear phase is identical to linear phase constrained solution. The advantage of using the constraints explicitly is the dimensionality reduction of the problem. Here we will treat the unconstrained fillter design case. These gradients and Hessians can be computed in closed form, [7], in terms the error and weight functions as follows:

$$\nabla G_{\bar{h}} = -\frac{p}{2}\int_0^1\!\!\int_0^1 W(f_1,f_2)\left|\Delta(\vec{h}^{(k)})\right|^{p-2}\bar{\Delta}(\vec{h})\vec{e}_{N,M}(f_1,f_2)^* df_1 df_2$$

$$F_{\vec{h}^T\vec{h}} = \frac{p}{2}\left(\frac{p}{2}-1\right)\int_0^1\!\!\int_0^1 W(f_1,f_2)\left[\left|\Delta(\vec{h}^{(k)})\right|^{p-4}\bar{\Delta}(\vec{h})^2\bar{E}_{N,M}(f_1,f_2)\right]df_1 df_2$$

and

$$F_{\vec{h}^*\vec{h}} = \left(\frac{p}{2}\right)^2\int_0^1\!\!\int_0^1 W(f_1,f_2)\left[\left|\Delta(\vec{h}^{(k)})\right|^{p-2}E_{N,M}(f_1,f_2)\right]df_1 df_2$$

$$E_{N,M}(f_1,f_2) = \begin{bmatrix} F & \bar{g}_1F & \bar{g}_2F & \cdots & \bar{g}_{M-1}F \\ g_1F & F & \bar{g}_1F & \ddots & \vdots \\ g_2F & g_1F & F & \ddots & \bar{g}_2F \\ \vdots & \ddots & \ddots & \ddots & \bar{g}_1F \\ g_{M-1}F & \cdots & g_2F & g_1F & F \end{bmatrix}$$

is a Toeplitz-block-Toeplitz matrix

$F = \vec{e}_N(f_1)\vec{e}_N(f_1)^*; g_k = \mathbf{exp}(2\pi jkf_2)$

F is a Hermitian Toeplitz matrix

$$\bar{E}_{N,M}(f_1,f_2) = \begin{bmatrix} \bar{F} & \bar{g}_1\bar{F} & \bar{g}_2\bar{F} & \cdots & \bar{g}_{M-1}\bar{F} \\ \bar{g}_1\bar{F} & \bar{g}_2\bar{F} & \bar{g}_3\bar{F} & \ddots & \vdots \\ \bar{g}_2\bar{F} & \bar{g}_3\bar{F} & \bar{g}_4\bar{F} & \ddots & \bar{g}_{2M-4}\bar{F} \\ \vdots & \ddots & \ddots & \ddots & \bar{g}_{2M-3}\bar{F} \\ \bar{g}_{M-1}\bar{F} & \cdots & \bar{g}_{2M-4}\bar{F} & \bar{g}_{2M-3}\bar{F} & \bar{g}_{2M-2}\bar{F} \end{bmatrix}$$

$\bar{E}$ is Hankel-block-Hankel matrix

$\bar{F} = \vec{e}_N(f_1)\vec{e}_N(f_1)^T; \bar{g}_k = \mathbf{exp}(2\pi jkf_2)$

$\bar{F}$ is a Hankel matrix

At first glance it would appear the computational load to obtain the Hessians numerically is great; however if one exploits the Hermitian Toeplitz-block-Toeplitz matrix property of E only the integrals for the first row of F and the first row and column of $\{g_kF\}_{k=1}^{M-1}$ be computed. The total number of such integrals is only [N+(M-1)(2N-1)]. For the calculation of the unconjugated Hessian only the integrals for the first and last column of $\{\bar{g}_k\bar{F}\}_{k=0}^{2M-2}$ are needed; these are (2M-1)(2N-1) of them. This results from the Hankel-block-Hankel property of $F_{\vec{h}^T\vec{h}}$

Implementation of the complex Newton algorithm is accomplished with increasing p and starting at p=2 in order to remain within its region of convergence at each iteration. As done in [4], the typical change in the value of p is a constant factor, A sequence for p may be:

$p_0 = 2;$

For $k \geq 1;\ p_k = \mathbf{min}\{\alpha p_{k-1}, p\}$, where $1 < \alpha \leq 1.5$

3 2-D FACTORIZATION THEOREM FOR FIR FILTERS IN RECTANGULAR COORDINATES

The design of 2-D FIR filters can be very computationally intensive especially for large order M and N. It is reasonable to have criteria to create optimal 2-D filters from 1-D ones. we state a theorem valid for p=2. We think a generalized version should hold for arbitrary p.

2-D Factorization Theorem:

Suppose both the desired 2-D filter response and weight function factor in terms of corresponding 1-D functions, then the 2-D filter coefficients are factorable in terms of the corresponding 1-D coefficients. If the desired frequency response in non factorable but the weighting function is, then the 2-D filter solution is partially factorable in terms of matrix components of the 1-D solution.

Proof:

The main observation here is that the complex gradient and Hessians are the linear combination of terms like: $\vec{e}_{N,M}(f_1,f_2)$, $\bar{E}_{N,M}(f_1,f_2)$ and $E_{N,M}(f_1,f_2)$. These can be factored in terms of Kronecker product as:

$\vec{e}_{N,M}(f_1,f_2) = \vec{e}_M(f_2)\otimes\vec{e}_N(f_1)$,

$$\bar{E}_{N,M}(f_1,f_2) = \vec{e}_{N,M}(f_1,f_2)\vec{e}_{N,M}(f_1,f_2)^T = \left(\vec{e}_M(f_2)\otimes\vec{e}_N(f_1)\right)\left(\vec{e}_M(f_2)\otimes\vec{e}_N(f_1)\right)^T = \left(\vec{e}_M(f_2)\otimes\vec{e}_N(f_1)\right)\left(\vec{e}_M(f_2)^T\otimes\vec{e}_N(f_1)^T\right) = \left(\vec{e}_M(f_2)\vec{e}_M(f_2)^T\right)\otimes\left(\vec{e}_N(f_1)\vec{e}_N(f_1)^T\right)$$

and

$$E_{N,M}(f_1,f_2) = \vec{e}_{N,M}(f_1,f_2)\vec{e}_{N,M}(f_1,f_2)^* = \left(\vec{e}_M(f_2)\otimes\vec{e}_N(f_1)\right)\left(\vec{e}_M(f_2)\otimes\vec{e}_N(f_1)\right)^* = \left(\vec{e}_M(f_2)\otimes\vec{e}_N(f_1)\right)\left(\vec{e}_M(f_2)^*\otimes\vec{e}_N(f_1)^*\right) = \left(\vec{e}_M(f_2)\vec{e}_M(f_2)^*\right)\otimes\left(\vec{e}_N(f_1)\vec{e}_N(f_1)^*\right)$$

The solution for p=2 is given by:

$\vec{h} = E^{-1}\vec{c}$

E can be factored as

$$E = \int_0^1\!\!\int_0^1 W(f_1,f_2)\left(\vec{e}_M(f_2)\vec{e}_M(f_2)^*\right)\otimes\left(\vec{e}_N(f_1)\vec{e}_N(f_1)^*\right)df_1 df_2 = \int_0^1 W_2(f_2)\vec{e}_M(f_2)\vec{e}_M(f_2)^* df_2 \otimes \int_0^1 W_1(f_1)\vec{e}_N(f_1)\vec{e}_N(f_1)^* df_1 = E_2\otimes E_1$$

Also $\vec{c}$ can be factored as

$$\vec{c} = \int_0^1\!\!\int_0^1 W(f_1,f_2)z_D(f_1,f_2)\,\vec{e}_{N,M}(f_1,f_2)df_1 df_2 = \int_0^1 W_2(f_2)z_D(f_2)\vec{e}_M(f_2)\vec{e}_M(f_2)^* df_2 \otimes \int_0^1 W_1(f_1)z_D(f_1)\vec{e}_N(f_1)\vec{e}_N(f_1)^* df_1 =$$

$\left(\vec{c}_2 \otimes \vec{c}_1\right)$

The solution can now be written as;

$$E^{-1}\vec{c} = \left(E_2 \otimes E_1\right)^{-1}\left(\vec{c}_2 \otimes \vec{c}_1\right) =$$

$$\left(E_2^{-1} \otimes E_1^{-1}\right)\left(\vec{c}_2 \otimes \vec{c}_1\right) = \left(E_2^{-1}\vec{c}_2\right) \otimes \left(E_1^{-1}\vec{c}_1\right)$$

This now implies the 2-D solution can be factored in terms of corresponding 1-D solutions.

If the desired resonse is factorable but the weighting is not the solution can still be expressed as:

$$\vec{h} = E^{-1}\vec{c} = \left(E_2 \otimes E_1\right)^{-1}\vec{c} = \left(E_2^{-1} \otimes E_1^{-1}\right)\vec{c}$$

Thus the most demanding operation, inversion of E, is still factorable.

NUMERICAL RESULTS

Two bandpass filters were derived using the technique described here. In both the passband extends from 0.4 to 0.6 in both frequencies. The stopband is the region outside the square [0.25,0.75]x[0.25,0.75]. The weighting function in Figure 1 is 1 on the passband and uniformly equal to 10 in the stopband, thus is not factorable. The constant factor was 1.2, p=60, M=12 and N=10. Convergence occured after 42 iterations. Figure 2 has the response of a similar filter whose weight function is the product of 1 dimensional weight functions which are equal to 1 on the passband and 3 on the stopband. Convergence occured in 47 iterations. Notice filter 1 has lower more uniform sidelobes while filter 2 has less ripple in the passband

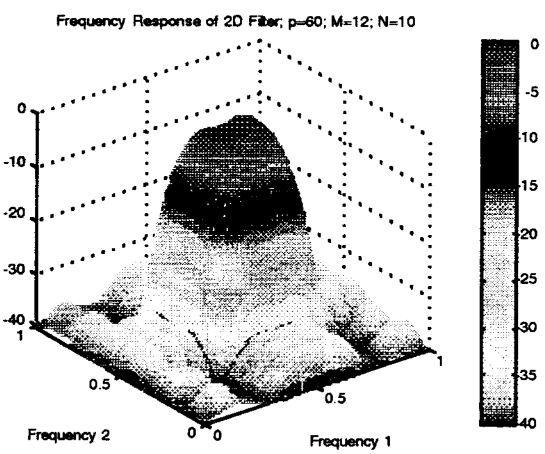

Figure 1: Passband filter response with non factorable weighting

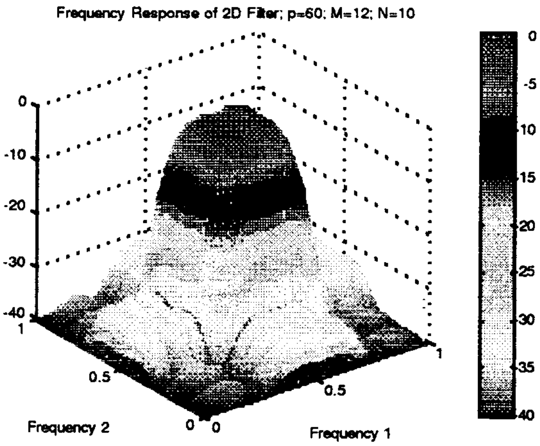

Figure 2: Passband filter response with factorable weighting

REFERENCES

[1] K. Preuss, "On the design of FIR filters by complex Chebyshev approximation," IEEE Trans. Acoust.., Speech and Signal Processing, vol. 40, No. 5, pp. 702-712, May 1989.

[2] A. G. Jaffer and W. E.. Jones, "Constrained Least Squares Design and Characterization of Affine Phase Complex FIR Filters," Proceedings of 1993 Asilomar Conference on Signals, Systems, and Computers, pp. 685-691, November 1993. Modified version of paper submitted to IEEE Trans. Signal Processing for publication.

[3] G. A. Lampropoulos and M. M. Fahmy, "A New Technique for the Design of Two-Dimensional FIR and IIR Filters," IEEE Transactions in Acoustics, Speech and Signal Processing, Vol. 33, No. 1, pp. 268_279, February 1985.

[4] J. A. Barreto and C. S. Burrus, "L_p- Complex Approximation using Iterative Reweighted Least Squares FIR Digital Filters," Proceedings of 1994 International Conference in Acoustics, Speech and Signal Processing, Vol. III, pp. 545.548, Adelaide, Australia.

[5] C. S. Burrus, J. A. Barreto and I. W. Selesnick, "Iterative Reweighted Least Squares Design of FIR Filters", IEEE Trans. Sign. Proc., Vol. 42, pp. 2926-2936, Nov. 94.

[6] A. S. Alkhairy, K. G. Christian and J. S. Lim, "Design and Characterization of Optimal FIR Filters with Arbitrary Phase," IEEE Trans. Sign. Proc., Vol. 41, No. 2, pp. 559-572, February 1993

[7] T. J Abatzoglou, J. M. Mendel and G. A. Harada, "The Constrained Total Least Squares and its Application to Harmonic Superresolution," IEEE Trans. Acous. Sp. and Sign. Proc., Vol. 39, N0. 5, pp.1070-1087, May 1991.

[8] T. J. Abatzoglou, "Least pth Power Design and Characterization of Affine Phase Complex 2-D Filters and the Min-Max Approximation," presented at the 28th annual Asilomar Conference on Signals, Systems and Computers, November 1994.

[9] A. G., Jaffer and T. J. Abatzoglou, "Weighted Least-Squares Design of Linear-Phase and Arbitrary 2-D Complex FIR Filters," presented at the 1995 IEEE ICASSP, May 1995.

[10] W. S. Lu, Y. Cui and R. L. Kirlin, "Least p^{th} optimization for the Design of 1-d Filters with Arbitrary Amplitude and Phase Responses," IEEE ICASSP Proc. 1993, Vol. III, pp. 61-64, Minneapolis, Minnesota, April 1993

[11] E. W. Cheney, Introduction to Approximation Theory, 2^{nd} edition, New York, Chelsea Publishing Co., 1982