Comments on "The Dual Parameterization Approach to... FIR Filter Design..."[1]

$$\begin{array}{ll}
\underset{h}{\text{minimize}} & \frac{1}{2}h^T\Phi h + p^T h + W_p f_p \\
\text{subject to} & \begin{bmatrix} \phi^T(f) \\ -\phi^T(f) \end{bmatrix} h - \begin{bmatrix} \delta_p + 1 \\ \delta_p - 1 \end{bmatrix} \leq 0, \qquad \text{for all } f \in [0, f_p] \\
& \begin{bmatrix} \phi^T(f) \\ -\phi^T(f) \end{bmatrix} h - \begin{bmatrix} \delta_s \\ \delta_s \end{bmatrix} \leq 0, \qquad \text{for all } f \in [f_s, 0.5]
\end{array}$$

In partial fulfilment of the requirements for

MSE313

Optimization by Vector Space Methods

Prof. D. G. Luenberger

D. Osborn, TA

Stanford University

2000

Jon Dattorro

[1]H. H. Dam, K. L. Teo, et al., "The Dual Parameterization Approach to Optimal Least Square FIR Filter Design Subject to Maximum Error Constraints", *IEEE Transactions on Signal Processing*, vol.48, no.8, pp.2314-2320, August 2000

1 Background and Overview

The literature is truly vast on the subject of finite impulse response (FIR) digital filter design. Any prudent contributor to this already saturated topic should be extraordinarily precise about what is offered, and how it is novel or different from similar past designs. The paper under scrutiny offers minimization of peak deviation, from the ideal lowpass magnitude response, using a weighted least squares technique. But this criterion for design is pedestrian and was in fact *the* criterion for FIR design prior to 1972. Then Parks and McClellan [4] implemented the efficient Remez multiple exchange algorithm that applies the result of Tonelli [3, ch.5.9] (often called *the alternation theorem*) to the Chebyshev design criterion; i.e., the minimization of maximum peak deviation. The Chebyshev criterion provides a globally optimal design in the sense that, for a given length of filter impulse response, the pass and stopband peak deviation is minimized with respect to any other design. That peak minimization is achieved by spreading the deviation evenly across the respective frequency bands. Hence, the characteristic equi-ripple response. The Parks/McClellan technique reigns today as the most widely used FIR filter design technique.

Personal experience with filter design indicates that least squares design techniques, in particular, have trouble with extremely narrowband lowpass filters.[2] [1] The trouble is typically manifest by a passband that is nowhere flat, in any sense, and that misses the allowable passband deviation by wide margin. On the other hand, the Parks/McClellan Chebyshev-based design is known to be quite successful when designing extremely narrowband filters.[3]

Given knowledge of the heritage of unsuccessful least-squares design techniques, I do not expect the particular technique presented in the scrutinized paper to perform any better than any other least-squares technique. Indeed, the authors show design examples having at most a 10:1 ratio of stopband to passband width. [Figure 2, pg.2318] Given the recent date of publication and the authors' awareness of Chebyshev techniques introduced 28 years earlier, I consider their omission of a narrowband design to be misleading.

The novelty of the paper seems only to be the approach to solving the semi-infinite quadratic optimization problem whose derivation is the filter design constraints indexed on continuous frequency. The filter design is, then, a vehicle that motivates the technique of solution. What is clever is the observation that there exists a solution of the dual problem having only a finite number of constraints. The basis for that claim is Caratheodory's dimensional theorem. A heuristic iterative procedure is proposed that sifts the only necessary constraints from the semi-infinite set.

[2]Narrowband filters find application in audio [2] where, for example, the new breed of sigma-delta one-bit analog to digital converters require a decimation in sample rate by factors now as much as 1000:1.

[3]The only anomaly I know of is the introduction of a small spike at the first and last samples of the symmetrical impulse response.

2 Synopsis

2.1 Section II

The filter design begins in section II with a zero-phase assumption thus obviating complex arithmetic.[4] A cost function is formulated simply as a weighted least-squares difference between the zero phase response and the ideal response. The specifications on peak deviation are transformed to semi-infinite affine constraints. Thus the filter design problem is abstracted to a convex quadratic semi-infinite programming problem, called the *primal* problem:

$$\begin{array}{ll}\underset{h}{\text{minimize}} & \frac{1}{2}h^T\Phi h + p^T h + W_p f_p \\ \text{subject to} & \begin{bmatrix} \phi^T(f) \\ -\phi^T(f) \end{bmatrix} h - \begin{bmatrix} \delta_p + 1 \\ \delta_p - 1 \end{bmatrix} \preceq 0, \qquad \text{for all } f \in [0, f_p] \\ & \begin{bmatrix} \phi^T(f) \\ -\phi^T(f) \end{bmatrix} h - \begin{bmatrix} \delta_s \\ \delta_s \end{bmatrix} \preceq 0, \qquad \text{for all } f \in [f_s, 0.5] \end{array}$$

where $h, p, \phi(f) \in \mathbf{R}^N$, $\Phi \in \mathbf{R}^{N\times N}$, $W_p, W_s, f_p, f_s, \delta_p, \delta_s \in \mathbf{R}$. The solution h represents the impulse response of the filter designed this way. The positive definite matrix Φ and the vector p are fixed by the pass/stopband weights, cutoff frequencies, and peak deviations: respectively $\{W_p, W_s, f_p, f_s, \delta_p, \delta_s\}$. The vector $\phi(f)$ is fixed by the length of the impulse response N.

2.2 Section III

Section III asserts: [Theorem 3.4, pg.2316] If $h^\star$ is the optimal solution of the primal semi-infinite problem, then there exists a solution of the dual problem having only a finite number of constraints. The basis for that claim is Caratheodory's dimensional theorem. [5]

The derivation of the dual problem ensues in the traditional manner. The Lagrangian duality theorem is invoked and attributed to Luenberger. [3, ch.8.6] Because Slater's constraint qualification (the existence of a strictly feasible point) is presumed to hold, and because the primal problem is strictly convex, then the dual problem has a unique optimal cost that is identical to that of the primal. The dual problem is:[5]

$$\begin{array}{ll}\underset{\Lambda_p,\Lambda_s}{\text{maximize}} & \mathcal{L}(\Lambda_p, \Lambda_s) \\ \text{subject to} & \Lambda_p(f) \succeq 0, \qquad \text{for all } f \in [0, f_p] \\ & \Lambda_s(f) \succeq 0, \qquad \text{for all } f \in [f_s, 0.5] \end{array}$$

where

[4]That is a standard trick in linear phase FIR filter design; the filter impulse response is made causal when the design procedure is finished.

[5]The authors stated the dual problem as a minimization problem. I now depart from the authors' notation to clarify and simplify their presentation.

$$\mathcal{L}(\Lambda_p,\Lambda_s) \;=\; -\frac{1}{2}\left(p+\int_0^{f_p}\left[\begin{array}{cc}\phi(f) & -\phi(f)\end{array}\right]d\Lambda_p(f)+\int_{f_s}^{0.5}\left[\begin{array}{cc}\phi(f) & -\phi(f)\end{array}\right]d\Lambda_s(f)\right)^T\cdot$$

$$\Phi^{-1}\left(p+\int_0^{f_p}\left[\begin{array}{cc}\phi(f) & -\phi(f)\end{array}\right]d\Lambda_p(f)+\int_{f_s}^{0.5}\left[\begin{array}{cc}\phi(f) & -\phi(f)\end{array}\right]d\Lambda_s(f)\right)$$

$$-\int_0^{f_p}\left[\begin{array}{cc}\delta_p+1 & \delta_p-1\end{array}\right]d\Lambda_p(f)-\int_{f_s}^{0.5}\left[\begin{array}{cc}\delta_s & \delta_s\end{array}\right]d\Lambda_s(f)\;+\;W_pf_p \tag{1}$$

is the strictly concave Lagrangian dual function, and where $\Lambda_p(f),\Lambda_s(f)\in\mathbf{R}^2$ are vectors of Lagrange multipliers. The primal optimal solution can be found from the dual;

$$h^\star \;=\; -\Phi^{-1}\left(p+\int_0^{f_p}\left[\begin{array}{cc}\phi(f) & -\phi(f)\end{array}\right]d\Lambda_p^\star(f)+\int_{f_s}^{0.5}\left[\begin{array}{cc}\phi(f) & -\phi(f)\end{array}\right]d\Lambda_s^\star(f)\right)$$

The dual problem still has semi-infinite constraints. The authors invoke Caratheodory's dimensional theorem [Theorem 3.4, pg.2316] and claim that the dual problem is equivalent to the discretized dual problem:

$$\begin{array}{lll}\underset{\Lambda_p,\Lambda_s,f_d}{\text{maximize}} & \mathcal{L}_d(\Lambda_p,\Lambda_s,f_d) & \\ \text{subject to} & \Lambda_p(f_d)\ge 0, & \text{for all } f_d\in[f_{p,1},\ldots,f_{p,m_p}] \\ & \Lambda_s(f_d)\ge 0, & \text{for all } f_d\in[f_{s,1},\ldots,f_{s,m_s}]\end{array}$$

where $m_s+m_p\le N$, and where $\mathcal{L}_d(\Lambda_p,\Lambda_s,f_d)$ is simply (1) discretized in frequency so that the integrals are replaced by summations. Similarly, the primal solution may now be expressed,

$$h_d^\star \;=\; -\Phi^{-1}\left(p+\sum_{i=1}^{m_p}\left[\begin{array}{cc}\phi(f_{p,i}) & -\phi(f_{p,i})\end{array}\right]\Lambda_p^\star(f_{p,i})+\sum_{i=1}^{m_s}\left[\begin{array}{cc}\phi(f_{s,i}) & -\phi(f_{s,i})\end{array}\right]\Lambda_s^\star(f_{s,i})\right)$$

The new variable f_d (an unknown set of at most N discrete frequencies) introduced into this formulation makes the dual problem tractable. But f_d is troublesome because it makes the discretized dual problem non-convex, hence admitting the existence of local minima. The authors' plan is to supply a good initial guess of the discrete frequencies f_d. By solving the semi-infinite dual problem over a dense enough grid, they determine which of the gridded primal constraints are active by *complementary slackness* (from the Kuhn-Tucker conditions). [3, ch.9.4] The frequencies corresponding to the active set are selected as the initial guess of f_d; what we shall call the estimated *Caratheodory frequencies*.

The discretized dual problem has a corresponding discretized primal problem that *is* convex:

$$
\begin{array}{lll}
\underset{h_d}{\text{minimize}} & \frac{1}{2}h_d^T \Phi h_d + p^T h_d + W_p f_p & \\
\text{subject to} & \begin{bmatrix} \phi^T(f_d) \\ -\phi^T(f_d) \end{bmatrix} h_d - \begin{bmatrix} \delta_p + 1 \\ \delta_p - 1 \end{bmatrix} \leq 0, & \text{for all } f_d \in [f_{p,1}, \ldots, f_{p,m_p}] \\
& \begin{bmatrix} \phi^T(f_d) \\ -\phi^T(f_d) \end{bmatrix} h_d - \begin{bmatrix} \delta_s \\ \delta_s \end{bmatrix} \leq 0, & \text{for all } f_d \in [f_{s,1}, \ldots, f_{s,m_s}]
\end{array}
$$

> *Theorem 3.5: Let $h_d^\star$ be an optimum solution for [the discretized primal problem]. If $h_d^\star$ satisfies the semi-infinite constraints, then it is the optimal solution for [the primal problem].*

That is to say, under the stated condition, $h^\star = h_d^\star$.

The proposed design procedure can be summarized as follows:

- Step 1) Obtain a good initial guess of the Caratheodory frequencies f_d by gridding the semi-infinite dual problem and finding the active primal constraints.
- Step 2) Solve the discretized dual problem using the estimated Caratheodory frequencies f_d, treating them as known. Then solve the discretized dual problem again, treating f_d as unknown.
- Step 3) Check if Theorem 3.5 is satisfied. If not, start all over again using a finer gridding.

3 Assessment

Most good traditional techniques for FIR filter, such as the Parks/McClellan [4] and least-squares technique, [1] grid the frequency domain as a necessary step in their solution process. Clearly, the authors are attempting to reduce the required grid density in the interest of computational efficiency. They have described a technique to divine the Caratheodory frequencies, and by so doing may also have increased the accuracy of the traditional weighted least-squares design. But as they themselves point out, they provide no proof that their iterative design procedure converges, hence there is no bound on the required grid density. Their empirical evidence shows that they are achieving lower grid density when compared to other gridding techniques, but they provide no data. They omitted the design example of an extremely narrowband design; that is the acid test.[6]

[6] Had their technique worked on such a critical example, it is likely that they would have been eager to announce it.

References

[1] J. W. Adams. Fir digital filters with least-squares stopbands subject to peak-gain constraints. *IEEE Transactions on Circuits and Systems*, 39:376–388, April 1991.

[2] J. Dattorro et al. Decimation filter as for a sigma-delta analog-to-digital converter. *United States Patent No.5,027,306*, June 1991.

[3] D. G. Luenberger. *Optimization by Vector Space Methods*. Wiley, 1969.

[4] T. W. Parks and J. H. McClellan. Chebyshev approximation for nonrecursive digital filters with linear phase. *IEEE Transactions on Circuit Theory*, 19:189–194, March 1972.

[5] R. Tyrrell Rockafellar. *Convex Analysis*. Princeton University Press, 1970.