DESIGN OF NONLINEAR-PHASE FIR DIGITAL FILTERS: A SEMIDEFINITE PROGRAMMING APPROACH

W.-S. Lu

Department of Electrical and Computer Engineering
University of Victoria
Victoria, BC, Canada V8W 3P6

ABSTRACT

The design of nonlinear-phase FIR digital filters is considered and it is shown that the design problem can be formulated as a *semidefinite programming* (SDP) problem. Specifically, we consider an equiripple design in which the phase response in the stopbands and transition bands are not required to be linear, and a design which approximates the desired frequency response in passband in equiripple sense and the desired frequency response in stopband in least-squares sense.

1. INTRODUCTION

We consider the problem of designing a nonlinear-phase, FIR digital filter that approximates a desired frequency response (both magnitude and phase responses) in the passbands in the Chebyshev sense, and approximates desired (zero) magnitude response in the stopbands in the least-squares sense. Consideration of such designs has been justified by many, see for example Adams [1]. Design of nonlinear-phase equiripple FIR filters is also considered in the paper. The term "nonlinear-phase design" is referred to a filter design in which the phase response in stopbands and transition bands are *not* required to be linear. As expected, the phase-response relaxation in the stopbands and transition bands from strict linearity has been found useful in enhancing the performance of the filter designed [2][3].

The design method proposed here is based on *semidefinite programming* (SDP), a relatively new optimization methodology which has been a topic of intensive research in the past several years [4]–[10]. There are several ways a SDP problem can be formulated, and the one which turns out to be convenient for filter design purposes is given as

$$\text{minimize} \quad c^T x \tag{1a}$$

$$\text{subject to} \quad F(x) \succeq 0 \tag{1b}$$

$$F(x) = F_0 + \sum_{i=1}^{n} x_i F_i \tag{1c}$$

In (1), $x \in R^{n\times 1}$ is the variable, $c \in R^{n\times 1}$, $F_i \in R^{n\times n}$ ($i = 0,\ 1,\ \ldots,\ n$) are given constant matrices with F_i symmetric, and $F(x) \succeq 0$ denotes that $F(x)$ is positive semidefinite. Note that the constraint matrix $F(x)$ in (1) is *affine* with respect to x. SDP includes both linear and quadratic programming (QP) as special cases, and it represents a broad and important class of convex programming problems. More important, many interior-point methods which have proven efficient for linear programming, have recently been generalized to SDP [6][8].

Concerning filter design problems, we notice that although there exist efficient optimization methods such as Remez exchange algorithm for the design of equiripple linear-phase FIR filter [11] and quadratic programming based algorithms for the design of linear-phase FIR filters with equiripple passbands and peak-constrained least-squares stop-bands (EPPCLSS) [1], extensions of these techniques to the nonlinear-phase case seem not at all trivial [12][3]. The objective of this paper is to indicate that SDP may serve as a suitable framework for the design of nonlinear-phase equiripple as well as EPPCLSS FIR filters.

2. PROBLEM FORMULATION

Consider the transfer function of an N-tap FIR filter

$$H(z) = \sum_{k=0}^{N-1} h_k z^{-k} \tag{2}$$

and denote its frequency response by

$$H(\omega) = \sum_{k=0}^{N-1} h_k e^{-jk\omega} = h^T[c(\omega) - js(\omega)] \tag{3}$$

where $h = [h_0 \cdots h_{N-1}]^T$, $c(\omega) = [1\ \cos\omega \cdots \cos(N-1)\omega]^T$, and $s(\omega) = [0\ \sin\omega \cdots \sin(N-1)\omega]^T$. Here we do *not* assume any symmetry in h. Let $H_d(\omega)$ be the desired frequency response, which is usually complex-valued. In an equiripple design, one seeks to find coefficient vector h that

0-7803-5471-0/99/$10.00©1999 IEEE

solves the optimization problem

$$\underset{h}{\text{minimize}}\ \underset{\omega\in\Omega}{\text{maximize}}\, W(\omega)|H(\omega)-H_d(\omega)| \qquad (4)$$

where Ω is a compact region on $[-\pi,\ \pi]$. The minimax problem in (4) can be reformulated as

$$\underset{h}{\text{minimize}} \quad \delta \qquad (5a)$$

$$\text{subject to } W^2(\omega)|H(\omega)-H_d(\omega)|^2 \le \delta \text{ for } \omega\in\Omega \quad (5b)$$

Now let

$$H_d(\omega) = H_r(\omega) - jH_i(\omega)$$

with $H_r(\omega)$ and $H_i(\omega)$ real, and use (3) to write the left-hand side of the constraint in (5b) as

$$W^2(\omega)|H(\omega)-H_d(\omega)|^2 = \alpha_1^2(\omega)+\alpha_2^2(\omega) \qquad (6)$$

where

$$\begin{aligned} \alpha_1(\omega) &= h^T c_w(\omega) - H_{rw}(\omega) \\ \alpha_2(\omega) &= h^T s_w(\omega) - H_{iw}(\omega) \\ c_w(\omega) &= W(\omega)c(\omega) \\ s_w(\omega) &= W(\omega)s(\omega) \\ H_{rw}(\omega) &= W(\omega)H_r(\omega) \\ H_{iw}(\omega) &= W(\omega)H_i(\omega) \end{aligned}$$

Constraint (5b) then becomes

$$\delta - \alpha_1^2(\omega) - \alpha_2^2(\omega) \ge 0 \qquad \omega\in\Omega \qquad (7)$$

It can be shown that (7) is equivalent to

$$\Delta(\omega) = \begin{bmatrix} \delta & \alpha_1(\omega) & \alpha_2(\omega) \\ \alpha_1(\omega) & 1 & 0 \\ \alpha_2(\omega) & 0 & 1 \end{bmatrix} \succeq 0 \qquad \omega\in\Omega \quad (8)$$

If we denote $x = [\delta\ \ h^T]^T$, then the linear dependence of $\alpha_1(\omega)$ and $\alpha_2(\omega)$ on h implies that matrix $\Delta(\omega)$ in (8) is affine w.r.t. x. Therefore, if $\{\omega_i,\ i=1,\ \ldots,\ M\}\subseteq\Omega$ is a set of grid points that are sufficiently dense in Ω, then a discretized version of (5) can be described as

$$\text{minimize} \quad c^T x \qquad (9a)$$

$$\text{subject to} \quad F(x)\succeq 0 \qquad (9b)$$

where $c = [1\ 0\ \cdots\ 0]^T$, and $F(x) = \text{diag}\{\Delta(\omega_1),\ \Delta(\omega_2),\ \cdots,\ \Delta(\omega_M)\}$. Obviously, $F(x)$ in (9b) is affine w.r.t. x, hence (9) is a SDP problem. Note that $F(x)$ is a *tridiagonal matrix* of size $3M\times 3M$, which becomes increasingly sparse with M.

In an EPPCLSS type design, one seeks to find h which minimizes the weighted L_2 error function

$$e(h) = \int_\Omega W(\omega)|H(\omega)-H_d(\omega)|^2\,d\omega \qquad (10a)$$

subject to constraints

$$\begin{aligned} |H(\omega)-H_d(\omega)|^2 &\le \delta_p \qquad \omega\in\Omega_p \qquad (10b) \\ |H(\omega)|^2 &\le \delta_a \qquad \omega\in\Omega_a \qquad (10c) \end{aligned}$$

where Ω_p and Ω_a denote the unions of passbands and stopbands, respectively. Simple manipulations of the integral in (10a) yields

$$e(h) = h^T P h - 2h^T q + c_0 \qquad (11)$$

where

$$P = \int_\Omega W(\omega)[c(\omega)\ \ s(\omega)]\,[c(\omega)\ \ s(\omega)]^T\,d\omega$$

is positive definite for a compact Ω,

$$q = \int_\Omega W(\omega)[H_r(\omega)c(\omega) + H_i(\omega)s(\omega)]\,d\omega$$

and

$$c_0 = \int_\Omega |H_d(\omega)|^2\,d\omega$$

Let $P^{1/2}$ be the symmetric square root of P, i.e., $P^{T/2} = P^{1/2}$ and $P^{1/2}P^{1/2} = P$. Then (11) can be written as

$$e(h) = ||P^{1/2}h - P^{-1/2}q||^2 - (||P^{-1/2}q||^2 - c_0)$$

Hence

$$e(h)\le\delta$$

is equivalent to

$$\delta + c_1 - ||P^{1/2}h - P^{-1/2}q||^2 \ge 0 \qquad (12)$$

where $c_1 = ||P^{-1/2}q||^2 - c_0$. It can be shown that (12) holds if and only if

$$\Gamma_0 = \begin{bmatrix} \delta + c_1 & h^T P^{1/2} - q^T P^{-1/2} \\ P^{1/2}h - P^{-1/2}q & I_N \end{bmatrix} \succeq 0 \qquad (13)$$

where I_N is the $N\times N$ identity matrix. Note that matrix Γ_0 in (13) is *affine* w.r.t. δ and h, and does not depend on ω. Similar to the way we treat constraint (5b), it can be shown that constraint in (10b) is equivalent to

$$\Gamma(\omega) = \begin{bmatrix} \delta_p & \beta_1(\omega) & \beta_2(\omega) \\ \beta_1(\omega) & 1 & 0 \\ \beta_2(\omega) & 0 & 1 \end{bmatrix} \succeq 0 \qquad \omega\in\Omega_p \qquad (14)$$

where $\beta_1(\omega) = h^T c(\omega) - H_r(\omega)$ and $\beta_2 = h^T s(\omega) - H_i(\omega)$, and that constraint in (10c) is equivalent to

$$\Phi(\omega) = \begin{bmatrix} \delta_a & \gamma_1(\omega) & \gamma_2(\omega) \\ \gamma_1(\omega) & 1 & 0 \\ \gamma_2(\omega) & 0 & 1 \end{bmatrix} \succeq 0 \qquad \omega\in\Omega_a \qquad (15)$$

where $\gamma_1(\omega) = h^T c(\omega)$ and $\gamma_2(\omega) = h^T s(\omega)$. Now if $\{\omega_i^{(p)}, i = 1, \ldots, M_p\} \subseteq \Omega_p$ $\{\omega_i^{(a)}, i = 1, \ldots, M_a\} \subseteq \Omega_a$ are the sets of grid points in the passbands and stopbands, respectively, on which the constraints (14) and (15) are imposed, then a discretized version of minimizing (10a) subject to (10b) and (10c) can be formulated as

$$\text{minimize} \quad c^T x \qquad (16a)$$

$$\text{subject to} \quad F(x) \succeq 0 \qquad (16b)$$

where $x = [\delta\ \delta_p\ \delta_a\ h^T]^T$, $c = [1\ w_p\ w_a\ 0\ \cdots\ 0]^T$ with scalar weights w_p and w_a, and $F(x) = \text{diag}\ \{\Gamma_0,\ \Gamma(\omega_1^{(p)}),\ \cdots,\ \Gamma(\omega_{M_p}^{(p)}),\ \Phi(\omega_1^{(a)}),\ \cdots,\ \Phi(\omega_{M_a}^{(a)})\}$. Clearly, $F(x)$ in (16b) is affine w.r.t. x, therefore (16) is a SDP problem.

Presently, there is only a limited number of software packages available that can be used to solve the SDP problem as formulated in (9). One of them is the LMI Control Toolbox from MathWorks Inc. [13]. The toolbox works with MATLAB and is aimed primarily at solving design problems arising from control engineering by using linear matrix inequality (LMI) techniques [7]. The LMI toolbox includes a command named `mincx` which implements the projective method proposed in [6][9] for solving SDP problems. For the sake of completeness, the next section offers a brief review of several key elements of the projective method.

3. PROJECTIVE METHOD FOR SDP

In this section, we give a brief review of the projective method proposed in [6][9] which has proven effective for SDP problems. To this end we define the set of all positive semidefinite matrices of size $n \times n$ by $\mathcal{K}$ and define the set of all positive definite matrices of size $n \times n$ by $\mathcal{S}$. A set $\mathcal{C}$ is said to be a *cone* if $x \in \mathcal{C}$ implies $\alpha x \in \mathcal{C}$ for all $\alpha > 0$. A set $\mathcal{C}$ is said to be a *convex cone* if $\mathcal{C}$ is a cone and is convex. Evidently, both $\mathcal{K}$ and $\mathcal{S}$ are convex cones and $\mathcal{S}$ can be viewed as the interior of $\mathcal{K}$. Given a positive definite matrix P, an inner product can be introduced in $\mathcal{S}$ as

$$< X, Y >_P = \text{tr}\,(PXPY) \qquad (17)$$

which induces the norm

$$||X||_P = [\text{tr}\,(PXPX)]^{1/2}$$

Note that if P is the identity matrix, then the above norm is reduced to the Frobenius norm

$$||X||_I = (\text{tr}\,X^2)^{1/2} = ||X||_{Fro}$$

An important concept involved in the development of the projective method is the Dikin ellipsoid [9] which is defined, for a fixed positive definite X, as the set

$$D(X) = \{Y : ||Y - X||_{X^{-1}} < 1\} \qquad (18)$$

It can be shown that the Dikin ellipsoid can be characterized as

$$D(X) = \{Y : ||X^{-\frac{1}{2}} Y X^{-\frac{1}{2}} - I||_{Fro} < 1\}$$

and that for a positive definite X, the Dikin ellipsoid $D(X)$ is always contained in cone $\mathcal{S}$. Therefore, $D(X)$ provides a region around X in which a search for a better point can be carried but without losing positive definiteness.

With the inner product defined in (17), one can consider the orthogonal projection of a positive definite X in $\mathcal{S}$ onto a subspace $\mathcal{E}$. In the context of SDP, $\mathcal{E}$ is the range of the linear map $\mathcal{F}$ associated with the LMI constraint in 1(c), i.e.,

$$\mathcal{F}x = \sum_{i=1}^{n} x_i F_i \qquad (19)$$

and subset $\mathcal{E}$ is characterized by

$$\mathcal{E} = \{Y : Y = \mathcal{F}x,\ x \in R^n\} \qquad (20)$$

The orthogonal projection of a given positive definite X onto subspace $\mathcal{E}$ with respect to the matric $<, >_P$ can be defined as the unique solution of the least-squares problem

$$\underset{Y \in \mathcal{E}}{\text{minimize}}\ ||Y - X||_P = \underset{x \in R^n}{\text{minimize}}\ ||\mathcal{F}x - X||$$

If we denote this orthogonal projection by X^+, then X^+ can be characterized by the optimality condition

$$< X^+ - X,\ Y >_P = 0 \qquad \text{for } Y \in \mathcal{E} \qquad (21)$$

which is equivalent to

$$< P(X^+ - X)P,\ Y >_{Fro} = 0 \qquad \text{for } Y \in \mathcal{E} \qquad (22)$$

Like any interior-point optimization method, the projective method starts at a *strictly feasible* initial point x_0 in the sense that matrix $F(x_0)$ in (1) is positive definite. For the sake of simplicity, we assume that the SDP problem at hand does have a strictly feasible initial point and that the linear objective function in (1a) has a finite lower bound.

As a preliminary step of the projective method, the problem in (1) is reformulated as the *homogeneous* problem

$$\text{minimize} \quad f(x) = \frac{\tilde{c}^T \tilde{x}}{\tilde{d}^T \tilde{x}} \qquad (23a)$$

$$\text{subject to:} \quad \tilde{\mathcal{F}}\tilde{x} \succeq 0 \qquad (23b)$$

$$\tilde{d}^T \tilde{x} \neq 0 \qquad (23c)$$

where

$$\tilde{x} = \begin{bmatrix} x \\ \tau \end{bmatrix}, \tilde{\mathcal{F}}\tilde{x} = \begin{bmatrix} \mathcal{F}x + F_0 & 0 \\ 0 & \tau \end{bmatrix}, \tilde{c} = \begin{bmatrix} c \\ 0 \end{bmatrix}, \tilde{d} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

The projective method can now be described in terms of the following algorithm.

Projective Algorithm for the SDP Problem in (23)

Step 1 Input a strictly feasible initial point $\tilde{x}_0$ and tolerance ε and compute $X_0 = \tilde{\mathcal{F}}\tilde{x}_0$.
Set $x_0^* = \tilde{x}_0$, and evaluate the objective function at x_0^* as f_0^*.

Step 2 Compute the orthogonal projection $X_k^+ = \tilde{\mathcal{F}}\tilde{x}_k$ of X_k onto $\mathcal{E}$ w.r.t. metric $< \cdot, \cdot >_{X_k^{-1}}$ and check its positive definiteness.
If $X_k^+ \succ 0$, go to Step 3; otherwise, set $f_k^* = f_{k-1}^*$, $Y_k = X_k^+ - X_k$, and go to Step 4.

Step 3 Reduce $f(x)$ in (23a) until $||X_k - X_k^+(f)||_{X_k^{-1}} \geq 0.99$ subject to $X_k^+(f) \succ 0$. where $X_k^+(f)$ denotes the orthogonal projection of X_k onto subspace $\mathcal{E}(f) = \{X = \tilde{\mathcal{F}}x : (\tilde{c} - f\tilde{d})^T x = 0\}$. Denote the resulting point by x_k^* and $f_k^* = f(x_k^*)$, and set $Y_k = X_k^+(f_k^*) - X_k$.

Step 4 If $f_{k-1}^* - f_k^* \leq \varepsilon$, then stop and output x_k^* as the solution; otherwise generate X_{k+1} using

$$X_{k+1}^{-1} = X_k^{-1} - \gamma_k X_k^{-1} Y_k X_k^{-1}$$

where γ_k is selected such that $X_{k+1}^{-1} \succ 0$ and $\det(X_{k+1}^{-1}) \geq \beta \det(X_k^{-1})$ for some fixed $\beta > 1$. Repeat from Step 2.

4. AN ILLUSTRATIVE EXAMPLE

Figure 1 shows the amplitude response of a 91-tap equiripple FIR filters designed by the proposed method to approximate a lowpass frequency response with normalized $\omega_p = 0.2375$, $\omega_a = 0.2625$, and group delay = 40. The LMI Control Toolbox was used to perform the design with 36 iterations and 190 Kflops.

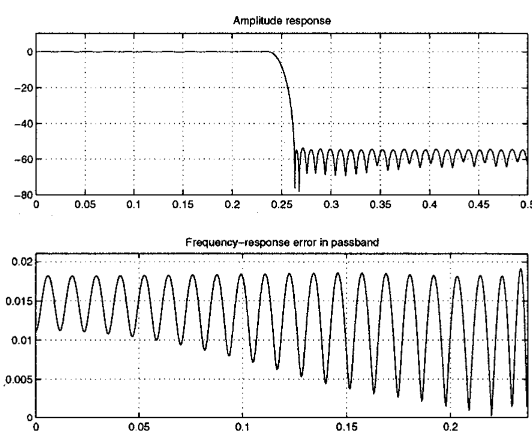

Figure 1

Acknowledgement

The author is grateful to Micronet, Networks of Centres of Excellence Program, and the Natural Sciences and Engineering Research Council of Canada for supporting this work.

5. REFERENCES

[1] J. W. Adams, "FIR digital filters with least-squares stopbands subject to peak-gain constraints," *IEEE Trans. Circuits Syst.*, vol. 39, pp. 376-388, April 1991.

[2] M. Lang and J. Bamberger, "Nonlinear phase FIR filter design according to the L_2 norm with constraints for complex error," *Signal Processing,* vol. 36, pp. 31-40, March 1994.

[3] M. C. Lang, "An iterative reweighted least squares algorithm for constrained design of nonlinear phase FIR filters," in *Proc. ISCAS '98*, May 1998.

[4] F. Alizadeh, *Combinatorial Optimization with Interior Point Methods and Semidefinite Matrices*, Ph.D. thesis, Computer Science Department, University of Minnesota, Minneapolis, MN, 1991.

[5] F. Alizadeh, "Interior point methods in semidefinite programming with applications to combinatorial optimization," *SIAM J. Optimization*, vol. 5, pp. 13-51, 1995.

[6] Y. Nesterov and A. Nemirovskii, *Interior Point Polynomial Methods in Convex Programming*, SIAM, Philadelphia, 1994.

[7] S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakvishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM, Philadelphia, 1994.

[8] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton, "Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability, and numerical results," *SIAM J. Optimization*, vol. 8, pp. 746-768, 1998.

[9] A. Nemirovskii and P. Gahinet, "The projective method for solving linear matrix inequalities," *Proc. Amer. Contr. Conf.*, pp. 840-844, Baltimore, MD., 1994.

[10] L. Vandenberghe and S. Boyd, "Semidefinite Programming," *SIAM Review*, vol. 38, pp. 49-95, March 1996.

[11] A. Antoniou, *Digital Filters: Analysis, Design, and Applications*, 2nd ed., New York: McGraw Hill, 1993.

[12] D. Burnside and T. W. Parks, "Optimal design of FIR filters with the complex Chebyshev error criteria," *IEEE Trans. Signal Processing*, vol. 43, pp. 605-616, March 1995.

[13] P. Gahinet, A. Nemirovskii, A. J. Laub, and M. Chilali, Manual of LMI Control Toolbox, The MathWorks Inc., Natick, MA, 1995.