Constrained FIR Filter Design by the Method of Vector Space Projections

Khalil C. Haddad, Henry Stark, *Fellow, IEEE*, and Nikolas P. Galatsanos

***Abstract*—A new technique for designing linear and arbitrary-phase finite-impulse response (FIR) filters with various types of constraints is proposed. The approach is based on the method of vector space projections. We describe the constraint sets and their associated projections that capture the properties of the desired filters. In filter design, as in many other engineering problems, one is primarily interested in meeting design constraints, i.e., finding a feasible solution, not necessarily an optimum one. Vector space projection methods are well-suited for this purpose. We furnish numerous examples of FIR filter design by vector space projections, including the important and difficult arbitrary phase/magnitude problem. Examples that demonstrate the advantages and flexibility of this method over other known methods are furnished.**

***Index Terms*—All-pass filters, convex projection, FIR filters, linear and quadratic constraints.**

I. Introduction

THE design of finite-impulse response (FIR) digital filters is a very basic problem in digital signal processing. Thus, it has received a lot of attention in the last 30 years. In a typical filter-design problem, the classical constraints are passband fluctuation, transition-band behavior, stopband attenuation, and filter length, i.e., support of the impulse response. When linear phase is also required, probably the most widely used approach is that of the well-known *McClellan-Parks* (MP) procedure [1]. These filters are optimal in the mini-max sense, i.e., for a given set of specifications, the largest error is minimized. However, the MP algorithm is based on polynomial factorization, and thus, is not easily extended to the multi-dimensional case. Moreover, it cannot incorporate additional constraints placed on the filter design.

In many filter design problems, constraints in addition to the classical ones are required. For example, one might require that the transient part of the step response be constrained within given amplitude limits. A second example is the design of Lth band filters where every Lth impulse-response components is constrained to having zero value except for the central value [2]. Also, in some cases, there are derivative constraints on the passband response of the filter [3] and so on.

An early and powerful design method for finding feasible solutions, i.e., solutions consistent with the imposed constraints is *linear programming* (LP) [4]. A disadvantage of linear programming, however, is that the required number of computations needed to arrive at a solution is rather large. Another disadvantage is that linear programming cannot easily handle nonlinear constraints. In [5] an alternative method for the design of linear-phase, FIR filters known as the *eigenfilter method* (EM) is presented. The idea behind EM is to first formulate a quadratic error measure $\mathbf{v}^T\mathbf{P}\mathbf{v}$ between the desired and the actual design, where $\mathbf{P}$ is a real- symmetric positive-definite matrix, and $\mathbf{v}$ is related to the filter impulse response. Then, one tries to minimize the total error by computing the eigenvectors and eigenvalues of $\mathbf{P}$ and pick the eigenvector that corresponds to the smallest eigenvalue in view of the well-known Rayleigh principle [6]. The eigenvector represents the filter coefficients. Usually, $\mathbf{P}$ is a weighted linear combination of several positive-definite matrices, e.g., $\mathbf{P} = \delta\mathbf{P}_p + \gamma\mathbf{P}_s + (1-\delta-\gamma)\mathbf{P}_e$, where $\delta \geq 0, \gamma \geq 0$, and $(\gamma+\delta) \leq 1$. The control parameters δ and γ assign priority weights to contributions in the passband and stopband; respectively. The positive-definite matrices $\mathbf{P}_p, \mathbf{P}_s$ and $\mathbf{P}_e$ are associated with energy constraints on the passband, stopband, and the unwanted signal, respectively. A disadvantage of EM is that the choice of appropriate values for δ and γ (considering that they should preserve the desired specifications as much as possible) is not obvious. The advantage of EM over linear programming is that the former is general enough to incorporate frequency and time-domain, as well as linear and quadratic, constraints.

More recently, methods based on convex optimization have been proposed for the design of FIR filters. In this approach, a change of variables leads to constraints being placed on the autocorrelations coefficients of the filter. Thus, the filter design problem is converted to a convex optimization problem. The coefficients of the filter are then recovered from the auto-correlations coefficients via spectral factorization. The advantage of this design approach is that it can incorporate different types of convex constraints (linear and nonlinear). Among others, magnitude bounds on Fourier transforms can be handled in this framework. Furthermore, it brings to bear to the filter design problem new efficient interior-point methods for convex optimization. For a recent review of this approach and additional references, see [7].

In some problems, the phase of the FIR filter needs to be a nonlinear function of frequency. Examples are found in phase-equalization, pulse shaping for chirp radar and others. A number of papers [8]–[11] present algorithms that address the general problem of designing FIR filters subject to prescribed magnitude and phase responses. In some of these algorithms the approach is to express the desired phase and magnitude as complex

Manuscript received February 1999; revised April 2000. This paper was recommended by Associate Editor J. Dias.

K. C. Haddad is with Lucent Technologies, Holmdel, NJ 07733 USA.

H. Stark and N. P. Galatsanos are with the Department of Electrical and Computer Engineering, Armour College of Engineering, Illinois Institute of Technology, Chicago, IL 60616-3793 USA.

Publisher Item Identifier S 1057-7130(00)06583-6.

Cartesian components and operate on the real and imaginary components independently [9], [10]. The final filter coefficients are formed from the resultant real and imaginary coefficients. Chen and Parks [8] approximate the complex-valued response by a real-valued function and the resulting errors in magnitude and group delay are made approximately equi-ripple. Their method, however, requires a large computer memory and the design-time increases exponentially with increasing time and frequency grid-density. Chit and Mason [12] used the *double adaptive system* (DAS) in approximating complex-valued specifications. Their method is based on least-mean-square minimization and a weight-adapting scheme designed specifically to give the filter Chebyshev characteristics. In Nguyen's EM procedure [13], the desired complex-valued function are approximated in a least-squares sense. The author claims that this method yields filters with performances better than the ones obtained with either the DAS [12] or the LP approaches [11].

To the best of our knowledge, Abou-Taleb and Fahmy [14] were the first to apply projection-like methods to an optimal (mini-max) 2-D FIR filter design. Their results are important, since the MP procedure is based on the alternation theorem and does not find a direct extension to the 2-D case. This is because the set of cosine functions used in 2-D approximation do not satisfy the Haar condition on the domain of interest, and the Chebyshev approximation does not have a unique solution. Techniques that employ exchange algorithms [15], [16] have been developed for the 2-D case at the expense of increased analytic complexity.

In an interesting recent paper, Cetin *et al.* [17] used an iterative Fourier transform algorithm to design zero-phase FIR filters. Upon examination, their algorithm is essentially a special case of vector space *projections known as projection onto convex sets* (POCS). The algorithm was derived heuristically, without explicitly defining the constraint sets and deriving their associated projectors. Moreover, the heuristic nature of this approach does not obviously lend itself to the design of filters with other constraints and with arbitrary phase.

In this paper, we consider the design of a class of FIR filters by *vector space projection methods* (VSPM's). We examine in detail the convexity of the prescribed constraint sets and rigorously derive their associated projectors. In our first example, we present the VSPM formulation of the FIR linear-phase design problem. In our second example, we demonstrate the flexibility of VSPM by imposing additional linear and nonlinear constraints on the filter design. Finally, we apply VSPM to the design of the general FIR filter subject to arbitrary magnitude and phase constraints including constraints of a nonconvex nature. In all cases, we compare our results with those of existing methods.

Before continuing with the specifics of VSPM applied to the FIR filter design problem, we should like to remind the reader of the fundamental advantages of VSPM. VSPM can handle any number of constraints including linear, convex and nonconvex types. In handling nonconvex constraints, we must weaken the notion of inner-product convergence to *summed distance error* (SDE) convergence. VSPM finds *feasible* solutions (solutions that satisfy all constraints) rather than optimal ones. In general, feasible solutions are simpler and less computationally expensive and are perfectly acceptable for a variety of engineering design problems such as filter design. Finally, VSPM can easily be extended to multi-dimensional filter design problems, unlike some other methods such as convex optimization that would require a (difficult) multi-dimensional spectral decomposition.

II. VSPM BACKGROUND

The VSPM deals with the problem of finding a mathematical object (for example, a signal, function, image, etc.) in a proper vector space that satisfies multiple constraints. When all the constraint sets are convex and have a nonempty intersection, there exists a powerful theory in finding the object that satisfies all the constraints. This subset of VSPM, mentioned in Section I, is called *projection onto convex sets* (POCS), which we describe below.

The theory of convex projection developed by Bregman [18] and Gubin *et al.* [19] was first applied to image processing by Youla and Webb [20]. The reader is referred to [21] for an introduction to this method as well as to its nonconvex extensions. Here we provide only the basic idea.

To begin with, assume that all the objects of interest are elements of a complete inner product space $\mathbf{H}$, i.e., a *Hilbert space*. Now consider a convex set $C \subset$ H; then, for any element $\mathbf{x} \in$ H, the projection $P\mathbf{x}$ of $\mathbf{x}$ onto C is the element of C closest to $\mathbf{x}$. If C is closed and convex, $P\mathbf{x}$ exists and is uniquely determined by $\mathbf{x}$ and C from the minimality criterion

$$||\mathbf{x} - P\mathbf{x}|| = \min_{g \in C} ||\mathbf{x} - \mathbf{g}||. \tag{1}$$

This rule, which assigns to every $\mathbf{x} \in$ H its nearest neighbor in C, defines the (in general) nonlinear projection operator P: H $\rightarrow C$ without ambiguity. A convenient Hilbert space for FIR filter design is R^M, the Euclidean space of M-vectors with real components. In this space, the inner product is taken as $\langle \mathbf{x}, \mathbf{y} \rangle = \Sigma_{i=1}^{M} x(i)y(i)$ and the induced norm is

$$||\mathbf{x}|| = \langle \mathbf{x}, \mathbf{x} \rangle^{1/2} = \left[\sum_{i=1}^{M} x^2(i) \right]^{1/2}. \tag{2}$$

The basic idea of POCS is as follows. Every known property of the unknown $\mathbf{x} \in$ H will restrict $\mathbf{x}$ to lie in a closed convex set C_i in H. Thus, for m known properties there are m closed convex sets $C_i, i = 1, 2, \ldots, m$ and $\mathbf{x} \in C_0 \equiv \cup_{i=1}^{m} C_i$. Then the problem is to find a point of C_0 given the sets C_i and projection operators P_i projecting onto $C_i, i = 1, 2, \ldots, m$. Based on fundamental theorems given by Opial [22] and Gubin *et al.* [19], the sequence $\{\mathbf{x}_k\}$ generated by the recursion relation

$$\mathbf{x}_{k+1} = P_m P_{m-1} \ldots P_1 \mathbf{x}_k, \qquad k = 0, 1, \ldots \tag{3}$$

or more generally by

$$\mathbf{x}_{k+1} = T_m T_{m-1} \ldots T_1 \mathbf{x}_k, \qquad k = 0, 1, \ldots \tag{4}$$

where $T_i \equiv I + \mu_i(P_i - I), 0 < \mu_i < 2$ are so-called *relaxed projectors* (they are not true projectors unless $\mu_i = 0$) converges weakly to a point in C_0. The $\mu_i, i = 1, \ldots, m$, are relaxation pa-

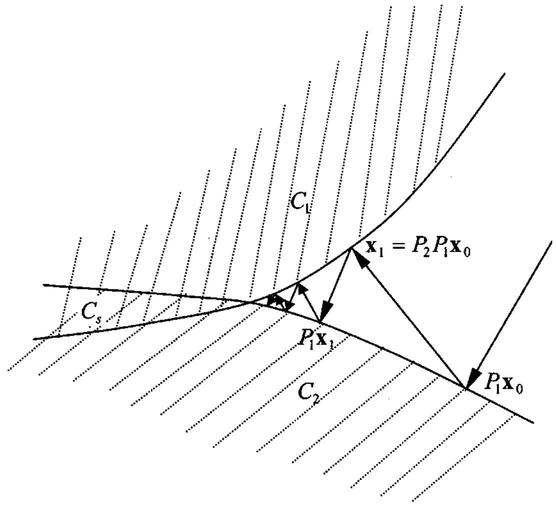


Fig. 1. Trajectory of iteration in POCS with two sets. The set C_s is the solution region and x_0 is an arbitrary starting point.

rameters and can be used to accelerate the rate of convergence of the algorithm; I is the identity operator. However, determining the optimum values of the μ, i.e., the ones that gives the fastest convergence, is generally a difficult problem and for, other than linear subspaces, experience has shown that good results are obtained when they are set to values somewhat arbitrarily between one and two. The algorithm in (2) for $m = 2$ is shown graphically in Fig. 1.

When sets are *nonconvex*, the extraordinary convergence properties of the method of VSPM no longer apply. However, there exists a fundamental theorem, which is quite useful in dealing with nonconvex sets. This theorem states that, in any problem involving not more than two constraint sets, *summed distance error* (SDE) convergence will always take place, even if nonconvex sets are involved. The SDE of a point $\mathbf{x}$ from the sets $C_i, i = 1, 2, \ldots, m$ is defined by $J(\mathbf{x}) \equiv \Sigma_{i=1}^{m}\, d(\mathbf{x}, C_i)$, where $d(\mathbf{x}, C) = \inf_{y \in C}\, \|\mathbf{x} - \mathbf{y}\|$. For more details on VSPM involving nonconvex sets (see [21]).

III. Design of Classical Linear-Phase FIR Filters Using VSPM

In this section, we describe linear-phase FIR filter design using VSPM. Consider the design of a FIR low-pass filter with linear-phase and impulse response $h(0), h(1), \ldots, h(N-1), h(N-1) \neq 0$ and $h(n) = 0, n > N-1$. We call N the filter *length*. This filter is required to meet the following specifications: in the passband, the magnitude $|H(\omega)|$ of the filter transfer function $H(\omega)$ must lie between $1-\alpha$ and $1+\alpha$, and in the stopband, $|H(\omega)|$ cannot exceed β. We put no constraints on the behavior of the filter in the transition band. Thus, if $A(\omega) \equiv |H(\omega)|$ is the magnitude and $\varphi(\omega) \equiv \arg\{H(\omega)\}$ is the phase, we require that $A(\omega) \in [1-\alpha, 1+\alpha]$ and that $\varphi(\omega) = -\omega(N-1)/2$ for $\omega \in \Omega_p \equiv \{\omega\colon 0 \leq \omega \leq \omega_p\}$. In addition, we require that $A(\omega) \leq \beta$ for $\omega \in \Omega_s \equiv \{\omega\colon \omega_s \leq \omega \leq \pi\}$, where Ω_p and Ω_s are the passband and stopband, respectively.

As stated earlier, our Hilbert space is R^M, where $M \gg N$ to insure a high-resolution Fourier transform without aliasing. In this problem, an appropriate cluster of constraint sets are C_1, C_2 and C_3 defined by

$$\begin{aligned} C_1 \equiv &\{\mathbf{h} \in R^M\colon h(n) = h(N-1-n), \\ &\text{for } n = 0, 1, \ldots, N-1 h(n) = 0, N \leq n \leq M-1\} \end{aligned} \tag{5}$$

$$\begin{aligned} C_2 \equiv &C_2(\alpha) \equiv \{\mathbf{h} \in R^M\colon 1-\alpha \leq A(\omega) \leq 1+\alpha \\ &\varphi(\omega) = -\omega(N-1)/2, \text{ for } \omega \in \Omega_p\} \end{aligned} \tag{6}$$

and

$$C_3 \equiv C_3(\beta) \equiv \{\mathbf{h} \in R^M\colon A(\omega) \leq \beta \text{ for } \omega \in \Omega_s\}. \tag{7}$$

In words, C_1 is the set of all sequences of length M with at most N nonzero coefficients with appropriate symmetry that imply a Fourier transform with *linear-phase*. The set C_2 is the set of all sequence whose Fourier magnitude is appropriately constrained in the passband and whose phase is linear in that band. Also, C_3 is the set of all sequences whose Fourier transform magnitude is appropriately constrained in the stopband. Note that it might have been tempting to use a Fourier magnitude constraint set, say C_2', given by

$$\begin{aligned} C_2' &\equiv C_2'(\alpha) \\ &\equiv \{\mathbf{h} \in R^M\colon 1-\alpha \leq A(\omega) \leq 1+\alpha \text{ for } \omega \in \Omega_p\}. \end{aligned} \tag{8}$$

However, this set is not convex, and hence, its involvement in a projection algorithm could leads to traps.[1] Given a choice, it is better to use convex rather than nonconvex sets because of guaranteed convergence of the sequence of iterates in the former (assuming the set intersection is not empty).

1) Convexity of C_1: Let $h_1(n), h_2(n) \in C_1$ and define $h_3(n) = \mu h_1(n) + (1-\mu)h_2(n)$ for $0 \leq \mu \leq 1$. Since $h_1(n) = h_1(N-1-n)$ and $h_2(n) = h_2(N-1-n)$, we have $h_3(n) = \mu h_1(N-1-n) + (1-\mu)h_2(N-1-n) = h_3(N-1-n)$. Hence, the set is convex. The proof that C_1 is closed is given on [21, p. 225]. Furthermore, it is easy to show that set C_1 is a linear subspace. For example, for $M = 3$ and $N = 2$ C_1 is the set of points on the line defined by the vector $(a, a, 0)$.

2) Projection onto C_1: To simplify matters, assume that all vectors are real. Let $\mathbf{g} \equiv (g(0), \ldots, g(M-1))^T$ be an arbitrary vector in H, $\mathbf{h} \equiv (h(0), \ldots, h(M-1))^T$ be any vector in C_1 and $\mathbf{h}^*$ the projection of $\mathbf{g}$ onto C_1. We deal with column vectors, hence the transposition ^{T}. Then

$$\mathbf{h}^* = \arg \min_{\mathbf{h} \in C_1} \sum_{n=0}^{M-1} [g(n) - h(n)]^2 \tag{9}$$

where $h^*(n) \equiv 0$ for $N \leq n \leq M-1$.

With $J \equiv \Sigma_{n=0}^{M-1}\, [g(n) - h(n)]^2$, the projection is easily computed. Taking into account that $h^*(n) = 0$ for $n = N, \ldots M-1$, we write (assuming that N is even) the Lagrange functional as

$$J \equiv \sum_{n=0}^{N/2-1} \{[g(n) - h(n)]^2 + [g(n+N/2) - h(n+N/2)]^2\} \tag{10}$$

[1] A trap is a fixed point of the algorithm that is not a solution. Traps do not appear in problems involving only convex sets.

and use the fact that $h(n+N/2) = h(N/2-n-1)$. Then with $\partial J/\partial h(l) = 0$ for $l = 0, 1, \ldots, (N/2) - 1$ we obtain

$$h^*(l) = \frac{g(l) + g(N-1-l)}{2}. \tag{11}$$

This clearly shows that $h^*(l) = h^*(N-1-l)$. Thus, the projection $\mathbf{h}^* = P_1\mathbf{g}$ of $\mathbf{g}$ onto C_1 becomes

$$h^*(l) = \begin{cases} \dfrac{g(l) + g(N-1-l)}{2}, & \text{for } l = 0, 1, \ldots N-1 \\ 0, & \text{elsewhere.} \end{cases} \tag{12}$$

3) Convexity of C_2: Let $\mathbf{h}_1$ and $\mathbf{h}_2 \in C_2$. Then $\mathbf{h}_3 \equiv \mu\mathbf{h}_1 + (1-\mu)\mathbf{h}_2 \leftrightarrow [\mu A_1(\omega) + (1-\mu)A_2(\omega)] \exp^{j\varphi(\omega)}$. The notation $\mathbf{g} \leftrightarrow G(\omega)$ or $G(\omega) \leftrightarrow \mathbf{g}$ implies a Fourier transform pair. Thus, the phase of $H_3(\omega) \leftrightarrow \mathbf{h}_3$ is $\varphi(\omega)$ and since $A_1(\omega)$ and $A_2(\omega)$ are lower and upper bounded by $1-\delta$ and $1+\delta$, respectively, so is $A_3(\omega) \equiv \mu A_1(\omega) + (1-\mu)A_2(\omega)$ for any $0 \le \mu \le 1$. Since C_2, as defined, includes all its limit points, it is closed. The geometric representation of C_2 in the frequency domain is that of a radial cut of length 2α, at an angle $-\omega(N-1)/2$, in an annulus of inner diameter $1-\alpha$ and outer diameter $1+\alpha$.

4) Projection onto C_2: The projection of an arbitrary vector $\mathbf{g} \in R^M$ with Fourier transform $G(\omega) = |G(\omega)|e^{j\theta_G(\omega)}$ can easily be computed using the Lagrange multiplier method. We obtain $\mathbf{h}^* = P_2\mathbf{g} \leftrightarrow H^*(\omega)$, where

$$H^*(\omega) = \begin{cases} (1+\alpha)\exp^{j\varphi(\omega)}, & \text{if cond.A} \\ (1-\alpha)\exp^{j\varphi(\omega)}, & \text{if cond.B} \\ |G(\omega)|\cos[\theta_G(\omega) - \varphi(\omega)]\exp^{j\varphi(\omega)}, & \text{if cond.C} \\ G(\omega), & \text{if } \omega \in \Omega_p^c \end{cases} \tag{13}$$

where conditions A, B, and C apply for all frequencies for $\omega \in \Omega_p$ and where

$$\begin{aligned} &\text{cond. A is: } |G(\omega)|\cos[\theta_G(\omega) - \varphi(\omega)] \ge 1+\alpha \\ &\text{cond. B is: } |G(\omega)|\cos[\theta_G(\omega) - \varphi(\omega)] \le 1-\alpha \\ &\text{cond. C is: } 1-\alpha \le |G(\omega)|\cos[\theta_G(\omega) - \varphi(\omega)] \le 1+\alpha. \end{aligned} \tag{14}$$

In the definition of $H^*(\omega), \Omega_p^c$ is the set of all frequencies in $[0, \pi] \notin \Omega_p$.

5) Convexity of C_3: Let $\mathbf{h}_1$ and $\mathbf{h}_2$ be $\in C_3$. Then $\mathbf{h}_3 \equiv \mu\mathbf{h}_1 + (1-\mu)\mathbf{h}_2 \leftrightarrow [\mu H_1(\omega) + (1-\mu)H_2(\omega)] \equiv H_3(\omega)$ and we must show that $|H_3(\omega)| \le \beta$. But for any two complex numbers A and B we have $|A+B| \le |A| + |B|$. Since $|H_1(\omega)|$ and $|H_2(\omega)|$ are bounded by β, it follows that $|H_3(\omega)| \le \beta$. The set C_3 can be represented in the complex frequency domain as a circle with radius β, centered at the origin. Since it includes its own boundary, it is closed.

6) Projection onto C_3: The projection of an arbitrary $\mathbf{g} \leftrightarrow G(\omega)$ onto C_3 is easily computed with the method of Lagrange multipliers as

$$\mathbf{h}^* = P_3\mathbf{g} \leftrightarrow \begin{cases} \beta G(\omega)/|G(\omega)|, & \text{for } |G(\omega)| > \beta, \omega \in \Omega_s \\ G(\omega), & \text{for } |G(\omega)| \le \beta, \omega \in \Omega_s \\ G(\omega), & \text{elsewhere.} \end{cases} \tag{15}$$

The FIR filter-design algorithm is given by

$$\mathbf{h}_{k+1} = P_1P_2P_3\mathbf{h}_k, \mathbf{h}_0 \text{ arbitrary.} \tag{16}$$

A good choice for the starting point $\mathbf{h}_0$ is $\mathbf{h}_0 \leftrightarrow H_0(\omega) = H_0(\omega) = H_{\text{ideal}}(\omega)$ with $H_{\text{ideal}}(\omega) = 1$, for $\omega \in \Omega_p$ and $H_{\text{ideal}}(\omega) = 0$ elsewhere. In Section VI, we furnish numerical results in which the VSPM algorithm in (16) is compared with the MP algorithm.

IV. Design of Classical Linear-Phase FIR Filters Subject to Additional Constraints

As we mentioned earlier, it is possible to design a linear-phase FIR filter subject to additional constraints. Here, we consider the design of a linear-phase FIR filter whose response $\mathbf{a} \in R^M$ to a known input is restricted to lie within certain bounds. For this problem, key sets are of the form

$$\begin{aligned} C_4(n) \equiv &\{\mathbf{h} \in R^M \colon b_1(n) \le (\mathbf{s} * \mathbf{h})_n \le b_2(n) \\ h(n) =& 0, n > N-1\} \\ n =& 0, 1, \ldots, N+L-2 \end{aligned} \tag{17}$$

where $\mathbf{s}$ is the given input, with components $s(0), s(1), \ldots, s(L-1) \neq 0$ and $s(l) = 0, l \ge L, *$ denotes convolution, $(\mathbf{s} * \mathbf{h})_n$ denotes the response at time n, and $b_1(n)$ and $b_2(n)$ represent the desired lower and upper bounds, respectively, on the response at time n. The explicit form of $a(n) \equiv (\mathbf{s} * \mathbf{h})_n$ is

$$a(n) = \sum_{i=0}^{N-1} h(i)s(n-i), \quad \text{for } n = 0, 1, \ldots, L+N-2 \tag{18}$$

where $s(-1) = s(-2) = \ldots = 0$. The components of $a(n), n \ge N+L-1$ are zero.

The entire system can be written in matrix form as

$$\hat{\mathbf{a}} = \mathbf{S}\hat{\mathbf{h}} \tag{19}$$

where

$$\begin{aligned} \hat{\mathbf{a}} &= (a(0), a(1), \ldots, a(L+N-2))^T \\ \hat{\mathbf{h}} &= (h(0), h(1), \ldots, h(N-1))^T \end{aligned}$$

and $\mathbf{S}$ is an $L+N-1 \times N$ matrix in (20), shown at the bottom of the next page.

From (18), it is not difficult to see that we can write $a(n)$ as $a(n) = \mathbf{s}_n^T\hat{\mathbf{h}}$, where $\mathbf{s}_n^T$ is the vector whose elements are the nth row of $\mathbf{S}$. Then $C_4(n)$ is equivalent to

$$C_4(n) \equiv \{\hat{\mathbf{h}} \in R^N \colon b_1(n) \le \mathbf{s}_n^T\hat{\mathbf{h}} \le b_2(n)\}, \quad n = 0, 1, 2, \ldots, L+N-2 \tag{21}$$

where, in the interest of saving notation, we omit introducing new notation for the set whose elements are in the reduced space.

The projection and proof of convexity for this set are given on [21, pp. 94–99]. We repeat the projection here for completeness. For any point $\mathbf{g} = (\hat{\mathbf{g}}, \mathbf{g}^c)^T \in R^M$ where

$$\begin{aligned} \hat{\mathbf{g}} &= (g(0), g(1), \ldots g(N-1))^T \in R^N \\ \mathbf{g}^c &= (g(N), g(N+1), \ldots, g(M-1))^T \end{aligned}$$

and its projection $\mathbf{h}^* = (\hat{\mathbf{h}}^*, 0, 0, \ldots, 0)^T \in R^M$, we obtain, for $n = 0, 1, \ldots, N-1$

$$\hat{\mathbf{h}}^* = P_4\hat{\mathbf{g}} = \begin{cases} \hat{\mathbf{g}}, & \text{if } b_1(n) \le \langle \mathbf{s}_n^T, \hat{\mathbf{g}} \rangle \le b_2(n) \\ \hat{\mathbf{g}} + \dfrac{b_1(n) - \mathbf{s}_n^T \hat{\mathbf{g}}}{\|\mathbf{s}_n\|^2}\mathbf{s}_n, & \text{if } \langle \mathbf{s}_n^T, \hat{\mathbf{g}} \rangle < b_1(n) \\ \hat{\mathbf{g}} = \dfrac{\mathbf{s}_n^T \hat{\mathbf{g}} - b_2(n)}{\|\mathbf{s}\|^2}\mathbf{s}_n, & \text{if } \langle \mathbf{s}_n^T, \hat{\mathbf{g}} \rangle > b_2(n). \end{cases} \tag{22}$$

The FIR filter design algorithm is given by

$$\mathbf{h}_{k+1} = P_1P_2P_3P_4\mathbf{h}_k, \mathbf{h}_0 \text{ arbitrary}. \tag{23}$$

Another example is the problem of designing a linear phase FIR filter with quadratic constraints. Consider a linear-phase low-pass FIR filter $\mathbf{h}$ with $\mathbf{h} \in R^M$. Let the input signal be, as in the previous example, an unwanted waveform $\mathbf{s}$, i.e., a finite-duration sequence of length L with values $s(0), s(1), \ldots, s(L-1) \neq 0, s(l) = 0, l > L-1$.

The time of occurrence of $\mathbf{s}$ is unknown, but its shape is known. We would like $\mathbf{h}$ to be such that, in addition to being a low-pass filter with given specifications, it constrains the energy of the output signal $\mathbf{a}$ due to $\mathbf{s}$. Thus, with $a(n)$ representing the components of $\mathbf{a}$, the sequence

$$a(n) = \sum_{i=0}^{N-1} h(i)s(n-i), \qquad n = 0, 1, \ldots, L+M-2 \tag{24}$$

represents the $(L+M-1)$ point output-sequence in response to the L-point input-sequence $\mathbf{s}$. Equations (18)–(20) of the previous example apply here as well, and to restrict the energy in $\mathbf{a}$, a useful constraint set is

$$\begin{aligned} C_5 &\equiv C_5(\sigma, \mathbf{d}) \\ &\equiv \{\mathbf{h} \in R^M : \|\mathbf{S}\hat{\mathbf{h}} - \mathbf{d}\| \le \sigma, h(n) = 0, n \ge N\} \end{aligned} \tag{25}$$

where, as before, the vector $\hat{\mathbf{h}}$ consists of the first N components of $\mathbf{h}$. The projection $\mathbf{h}^*$ of any $\mathbf{g} = (\hat{\mathbf{g}}, \mathbf{g}^c) \in R^M$, where $\hat{\mathbf{g}} = (g(0), g(1), \ldots g(N-1))^T \in R^N$, will have the form $\mathbf{h}^* = (\hat{\mathbf{h}}, \mathbf{0})$ where, as before, $\hat{\mathbf{h}} = (h(0), h(1), \ldots, h(N-1))^T$. Since all the components of $\mathbf{h}^*$ above the $(N-1)$st are restricted to be zero, a set equivalent to C_5 is the reduced set

$$C_5 \equiv C_5(\sigma, \mathbf{d}) \equiv \{\hat{\mathbf{h}} \in R^N : \|\mathbf{S}\hat{\mathbf{h}} - \mathbf{d}\| \le \sigma.\} \tag{26}$$

In words, the set C_5 is the set of all real-valued impulse responses, $\hat{\mathbf{h}}$, whose responses to the signal $\mathbf{s}$ lies within a sphere of radius σ centered at $\mathbf{d}$

By studying the relation

$$\|\mathbf{S}\hat{\mathbf{h}}\| \le \sigma \tag{27}$$

we conclude that C_5 has the form of an ellipsoid, and therefore, is convex.

1) Projection onto C_5: The computation of the projection of an arbitrary element $\hat{\mathbf{g}} \in R^N$ onto C_5 involves finding the extremum of the Lagrange functional

$$J(\hat{\mathbf{h}}) \triangleq \|\hat{\mathbf{g}} - \hat{\mathbf{h}}\|^2 + \lambda(\|\hat{\mathbf{h}} * \mathbf{s} - \mathbf{d}\|^2 - \sigma^2) \tag{28}$$

where, as usual, λ is the Lagrange multiplier. Differentiation followed by some algebra determines that the minimum of J has the form

$$\hat{\mathbf{h}}_\lambda = (\mathbf{I} + \lambda\mathbf{S}^T\mathbf{S})^{-1}(\lambda\mathbf{S}^T\mathbf{d} + \hat{\mathbf{g}}) \tag{29}$$

where $\mathbf{S}$ is as in (20), $\mathbf{S}^T\mathbf{S}$ has dimension $N \times N$, $\mathbf{I}$ is an $N \times N$ identity matrix, and $\mathbf{d}$ is an $N + L - 1 \times 1$ prescribed vector.

The projection $\hat{\mathbf{h}}^*$ in the reduced space is given as

$$\hat{\mathbf{h}} = P_5\hat{\mathbf{g}} = \begin{cases} \hat{\mathbf{g}}, & \text{if } \|\mathbf{S}\hat{\mathbf{g}} - \mathbf{d}\| \le \sigma \\ \hat{\mathbf{h}}_{\lambda^*}, & \|\mathbf{S}\hat{\mathbf{g}} - \mathbf{d}\| > \sigma, \text{ where } \lambda^* > 0 \\ & \quad \text{is chosen so that } \|\mathbf{S}\hat{\mathbf{h}}_\lambda - \mathbf{d}\| = \sigma \end{cases} \tag{30}$$

and the projection $\mathbf{h}^* \in R^M$ is given by $\mathbf{h}^* = (\hat{\mathbf{h}}^*, \mathbf{0})^T$.

The computation of λ^* is facilitated by the recognition that $\varphi(\lambda) \equiv \|\mathbf{S}\hat{\mathbf{h}}_\lambda - \mathbf{d}\|^2$ is a monotonically decreasing function of λ for $\lambda > 0$. This is directly demonstrable by showing that, for $\lambda > 0, \varphi'(\lambda) < 0$. The actual computation of λ^* can be done by a Newton–Raphson type algorithm. Finding the extremum of J is readily done using Parseval's theorem in the frequency domain. We shall omit the details in the interest of brevity. For

$$\mathbf{S} = \begin{bmatrix} s(0) & 0 & 0 & \cdots & 0 & 0 & 0 \\ s(1) & s(0) & 0 & \cdots & 0 & 0 & 0 \\ s(2) & s(1) & s(0) & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ s(L-1) & s(L-2) & s(L-3) & \cdots & \cdots & s(L-N+1) & s(L-N) \\ 0 & s(L-1) & s(L-2) & \cdots & \cdots & \cdots & s(L-N+1) \\ 0 & 0 & s(L-1) & \cdots & \cdots & \cdots & s(L-N+2) \\ \vdots & \vdots & \vdots & s(L-1) & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s(L-1) & s(L-2) & s(L-3) \\ 0 & 0 & 0 & \cdots & 0 & s(L-1) & s(L-2) \\ 0 & 0 & 0 & \cdots & 0 & 0 & s(L-1) \end{bmatrix} \tag{20}$$

further details, see [23] and [24]. The FIR filter-design algorithm is given by

$$\mathbf{h}_{k+1} = P_1 P_2 P_3 P_5 \mathbf{h}_k, \mathbf{h}_0 \text{ arbitrary.} \tag{31}$$

V. Design of FIR Filters with Arbitrary Magnitude and Phase Using VSPM

Consider the design of a FIR filter with arbitrary phase and magnitude that meets the following specifications: for a specific frequency ω, the magnitude $|H(\omega)|$ and $\varphi(\omega)$ phase of the filter's frequency response $H(\omega)$ must be in $[a(\omega)-\delta, a(\omega)+\delta]$ and $[\alpha(\omega)-\varepsilon, \alpha(\omega)+\varepsilon]$, respectively. The first step in designing a filter that meets these constraints is to define the appropriate sets. Define first

$$C_6 \equiv C_6(\delta,\varepsilon) \equiv \{\mathbf{h} \in R^M\colon a(\omega)-\delta \le |H(\omega)| \le a(\omega)+\delta \text{ and } \alpha(\omega)-\varepsilon \le \varphi(\omega) \le \alpha(\omega)+\varepsilon\} \tag{32}$$

$$C_7 \equiv \{\mathbf{h} \in R^M\colon h(N-1) \neq 0 \text{ and } h(n)=0, \text{ for } n = N, N+1, \ldots M-1\} \tag{33}$$

where $\varphi(\omega) = \arg H(\omega)$. While the set C_6 is nonconvex and, therefore, *guaranteed* strong convergence is not a possibility, experience has shown that excellent results are still possible when projecting onto such sets since SDE convergence will always take place when only two sets are involved [21]. The set C_7 is the set of all N-length impulse responses in R^M with the first N elements nonzero and all the rest zero. This set is convex and the projection $\mathbf{h}^*$ of $\mathbf{g}$ onto C_7 is given by

$$P_7\mathbf{g} = \mathbf{h}^* = \begin{cases} h^*(n) = g(n), & 0 \le n \le N-1 \\ 0, & \text{else.} \end{cases} \tag{34}$$

The projection onto C_6 is more difficult and is best computed from geometrical considerations.

1) Projection onto C_6: To prevent cumbersome notations, let $H \equiv H(\omega), a \equiv a(\omega), \varphi \equiv \varphi(\omega)$ and $\alpha \equiv \alpha(\omega)$. Let us assume we need to project a trial solution $\mathbf{g} \leftrightarrow G(\omega) \equiv x$ onto the set C_6. Since the contour of the set C_6 is made of sharp corners and curves with discontinuous derivatives, we must be careful to partition to space into regions whose vectors will be projected on various parts of the contours C_6. Indeed, the mathematical description of the appropriate projection operation will depend on the location of the point in the complex-plane defined by Re$[H]$ and Im$[H]$. We partition the complex-plane into nine disjoint regions as shown in Fig. 2 and geometrically described in Table I. For the sake of brevity, in what follows we provide the detailed calculation of the projection only for the cases where $x \in$ Region V and again when $x \in$ Region VI. For the other regions the calculations are similar and are given without derivation in Table II.

2) Projection from Region V(VIII): The Lagrange functional to be minimized for this projection is

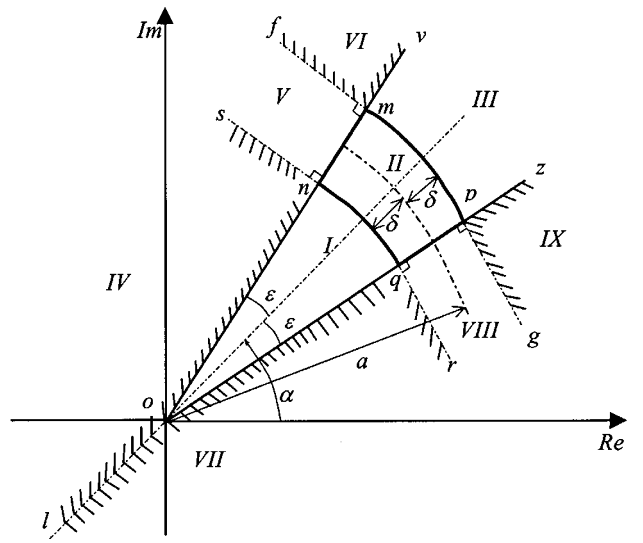


Fig. 2. Tolerance around $H(\omega)$ of magnitude a and phase α in the complex-plane.

TABLE I
REGION DESCRIPTION OF FIG. 2

Region	Description
I	Interior of bounded region (o,n,q)
II	Interior of bounded region (m,n,q,p)
III	Partially unbounded region (v,m,p,z)
IV	Partially unbounded region (l,o,n,s)
V	Partially unbounded region (s,n,m,f)
VI	Partially unbounded region (f,m,v)
VII	Partially unbounded region (l,o,q,r)
VIII	Partially unbounded region (r,q,p,g)
IX	Partially unbounded region (z,p,g)

TABLE II
PROJECTIONS DEFINITION OF P_6

Region	Projection
I	$(a-\delta)\exp[j(\arg x)]$
II	x
III	$(a+\delta)\exp[j(\arg x)]$
IV	$(a-\delta)\exp[j(\alpha+\varepsilon)]$
V	$\lvert x\rvert\cos(\arg x-\alpha-\varepsilon)\exp[j(\alpha+\varepsilon)]$
VI	$(a+\delta)\exp[j(\alpha+\varepsilon)]$
VII	$(a-\delta)\exp[j(\alpha-\varepsilon)]$
VIII	$\lvert x\rvert\cos(\alpha-\varepsilon-\arg x)\exp[j(\alpha-\varepsilon)]$
IX	$(a+\delta)\exp[j(\alpha-\varepsilon)]$

$$J(x,H) \equiv J = |x-H|^2 + \lambda\left[\tan^{-1}\left(\frac{H_I}{H_R}\right) - (\alpha+\varepsilon)\right] \tag{35}$$

where H_I and H_R denote the imaginary and real parts of $H, \alpha+\varepsilon$ is the constraint on the argument ∂_H of H and x is a point in Region V described by $x - x_R + jx_I$. We can rewrite J as

$$J = (x_R - H_r)^2 + (x_I - H_I)^2 + \lambda\left[\tan^{-1}\left(\frac{H_I}{H_R}\right) - (\alpha+\varepsilon)\right] \tag{36}$$

and set the derivatives

$$\frac{\partial J}{\partial H_R} = 0 \tag{37}$$

and

$$\frac{\partial J}{\partial H_I} = 0 \tag{38}$$

in order to find H^*, the Fourier transform of $\mathbf{h}^*$. From (37) and (38), we obtain

$$2(x_R - H_R) = -\lambda \frac{H_I}{H_R^2 + H_I^2} \tag{39}$$

$$2(x_I - H_I) = \lambda \frac{H_R}{H_R^2 + H_I^2.} \tag{40}$$

Dividing (39) by (40) and recognizing that

$$\begin{aligned} x_R &= |x|\cos(\theta_x), \quad x_I = |x|\sin(\theta_x) \\ \arg(H) &= \alpha + \varepsilon, \quad H_R = |H|\cos(\theta_x) \\ H_I &= |H|\cos(\theta_x) \end{aligned}$$

we finally obtain that

$$|H| = |x|\cos(\theta_x - \alpha - \varepsilon) \tag{41}$$

where $\theta_x \equiv \arg(x)$. Thus

$$H = x\cos(\theta_x - \alpha - \varepsilon)e^{j(\alpha+\varepsilon)}. \tag{42}$$

3) Projection from Region VI(IX): For any point $x \in R_6$, its projection must either be on the contour segment (n, m) or the contour segment (m, p) in Fig. 2. Any point P on the (m, n) contour is given by

$$P = (a+\delta)e^{j(\alpha+\varepsilon)} - pe^{j(\alpha+\varepsilon)} \tag{43}$$

where $p \geq 0$ to be determined. The point x *may be* decomposed as

$$x = [|x|\cos(\theta_x - (\alpha+\varepsilon)) + j|x|\sin(\theta_x - (\alpha+\varepsilon))]e^{j(\alpha+\varepsilon)}. \tag{44}$$

Thus

$$\begin{aligned} |x-P|^2 = &|x|^2 + (a+\delta)^2 + p^2 - (a+\delta)\cos[\theta_s - (\alpha+\varepsilon)] \\ &+ |x|p\cos[\theta_x - (\alpha+\varepsilon)] - 2(a+\delta)p. \end{aligned} \tag{45}$$

To find the projection on (n, m), we minimize $|x-P|^2$ as a function of p. We can rewrite (45) as $|x-P|^2 = \sigma + \varphi(p)$, where $\phi(p)$ is the p-dependent part. Thus, we seek to make $\phi(p)$ as small as possible. But $\phi(p) = p^2 + 2p[|x|\cos[\theta_x - (\alpha+\varepsilon)] - (a+\delta)]$ and for any point in $R_6, |x|\cos[\theta_x - (\alpha+\varepsilon)] \geq a+\delta$. Thus, $\phi(p) = p^2 + 2Kp$ where

$$K = |x|\cos[\theta_x - (\alpha+\varepsilon)] - (a+\delta) \geq 0. \tag{46}$$

This is a parabola with minimum at $p_{\min} = -K$ and smallest value for $p \geq 0$ at $p = 0$. Therefore, the projection onto (n, p) is the point $P^* = (a+\delta)e^{j(\alpha+\varepsilon)}$.

Assume next that the projection of x is on the contour segment (m, p). Any point P on (m, p) can be written as $P = (a+\delta)e^{j(\theta_P)}$. Then

$$|x-P|^2 = |x|^2 + (a+\delta)^2 - |x|p\cos(\theta_x - \theta_P) \tag{47}$$

where $\alpha + \varepsilon \leq \theta_x \leq \alpha + \varepsilon + (\pi/2)$ and $\alpha - \varepsilon \leq \theta_P \leq \alpha + \varepsilon$. We seek the smallest value of $\theta_x - \theta_P$ to minimize $|x-P|^2$. For *any* θ_x, this occurs at $\theta_P = \alpha + \varepsilon$. Therefore, $P^* = (a+\delta)e^{j(\alpha+\varepsilon)}$, the same as before. The projections for other regions are obtained similarly and are tabulated in Table II.

The FIR filter-design algorithm is given by

$$\mathbf{h}_{k+1} = P_6 P_7 \mathbf{h}_k, \mathbf{h}_0 \text{ arbitrary.} \tag{48}$$

VI. Numerical Results

It is clear from our previous discussion that VSPM filter-design requires the frequency-domain implementation of the constraints in C_2, C_3 and C_6. These constraints are realized on a grid of discrete frequencies. An M-length DFT is implemented by the FFT algorithm for $M = 2^k$. Thus, we create M discrete frequencies $\omega_i = (2\pi/M)i, i = 0, 1, \ldots M-1$ over the interval [0, 2π]. Below we give several examples of FIR filter design using VSPM. In the first example, we design a simple linear-phase low-pass FIR filter and compare its performance to that obtained with the MP method. Example 2 consists of designing a low-pass FIR filter with bounds on the overshoot and undershoot of the step response. Example 3 consists of designing a low-pass FIR filter with an output-energy constraint on a prescribed, undesirable, signal and comparing the VSPM design with the EM design. Example 4 consists of designing an all-pass FIR filter with a prescribed nonlinear phase signal and comparing the VSPM design with the designs of LP, DAS, and EM. In all of the following examples, $M = 1024$ was used except for Example 4, in which the $M = 4096$ was used for better accuracy. The stopping criterion of the iterative VSPM in the following examples is given by: *stop when* $\|\mathbf{h}_{k+1} - \mathbf{h}_k\| < 10^{-6}$. When this condition is met we say, somewhat arbitrarily, that convergence has occurred.

Example 1: VSPM versus MP—Low-pass Linear-phase FIR Filter Design

We design a low-pass filter using the sets C_1 with finite impulse response of length $N = 31$, $C_2(\alpha)$ with $\alpha = 0.0243$ and $C_3(\beta)$ with $\beta = 0.0243$. The passband and stopband edge frequencies are $\omega_p = 0.4\pi$ and $\omega_s = 0.5\pi$. The result is compared with the MP algorithm with the same passband/stopband frequencies computed using MATLAB with equal error weights on both the stopband and the passband. Fig. 3 shows the resulting frequency responses of both the MP (broken line) and the VSPM (solid line) results. Figs. 4 and 5 show the passband and stopband of the two algorithms in details. The main result is that the response of the filter generated by VSPM (solid line) is as good as the MP one (broken lines). For this example, the VSPM algorithm in (11) required 4,000 iterations for convergence.

Example 2: Filter Design with Constraints on the Step Response

In this example, we use VSPM to design a filter that constrains the overshoot and undershoot of the truncated *step re-*

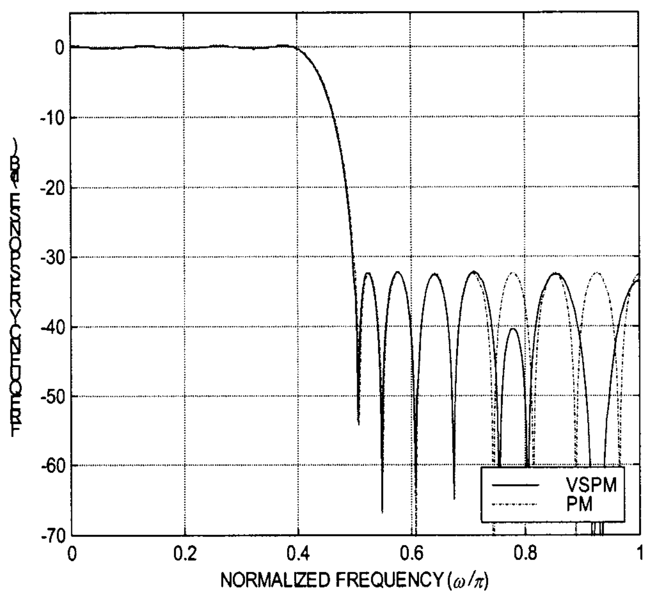


Fig. 3. Frequency response of VSPM- and PM-designed filters.

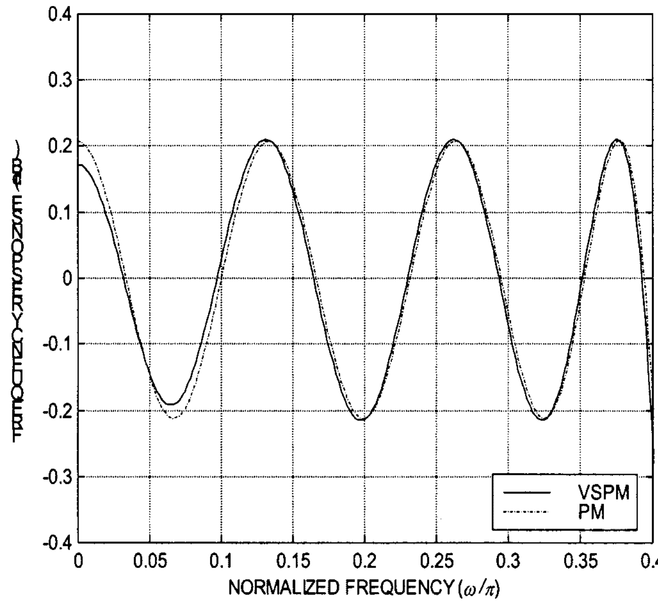


Fig. 5. Passband response of VSPM- and PM-designed filters in details.

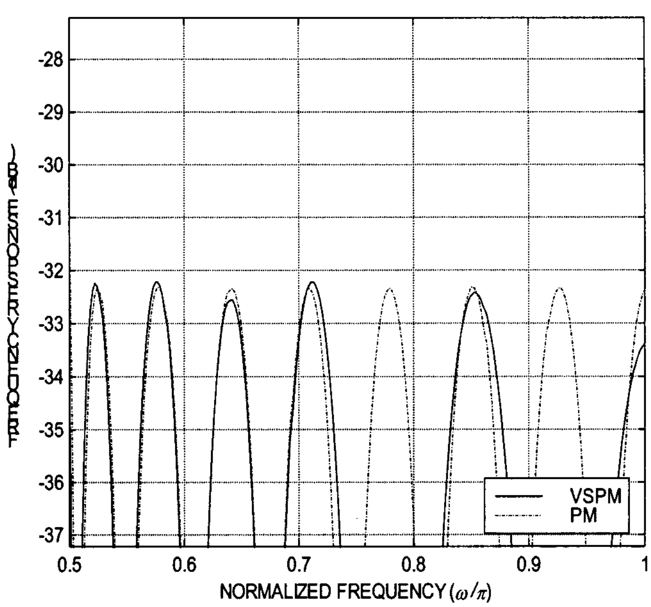


Fig. 4. Stopband response of VSPM- and PM-designed filters in details.

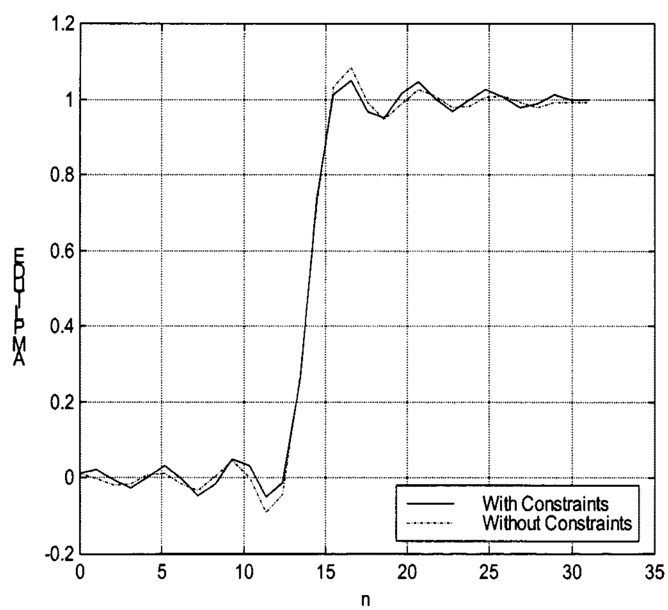


Fig. 6. Step response with and without constraints.

sponse $\mathbf{a}$ of the filter in Example 1. Thus, we use the sets C_1 with $N = 31$, $C_2(\alpha)$ with $\alpha = 0.13$ and $C_3(\beta)$ with $\beta = 0.13$, and the same passband and stopband edge frequencies as in Example 1. We chose values of α and β that still allowed a nonempty intersection of all the constraints sets. However, now we involve the sequence of sets $\{C_4(n)\}$ of (21). The truncated step-response $\mathbf{a}$ is given by

$$\mathbf{a} \equiv \begin{bmatrix} a(0) \\ a(1) \\ \vdots \\ a(30) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} h(0) \\ h(1) \\ \vdots \\ h(30) \end{bmatrix} \tag{49}$$

while the $b_1(n)$ and $b_2(n)$ that define set $C_4(n)$ are given by

$$\begin{aligned} b_1(n) &= -0.005 \\ b_2(n) &= 0.005, \quad \text{for } n = 1, \ldots 13 \\ b_1(n) &= 1 - 0.005 \\ b_2(n) &= 1 + 0.005, \quad \text{for } n = 18, \ldots 31. \end{aligned} \tag{50}$$

The sets $C_4(n)$ are not applied for $n > 30$. Also, we let the response be *unconstrained* during its monotone rise. The broken line in Fig. 6 represents the step response of the filter in Example 1 without applying a step response constraint. The solid

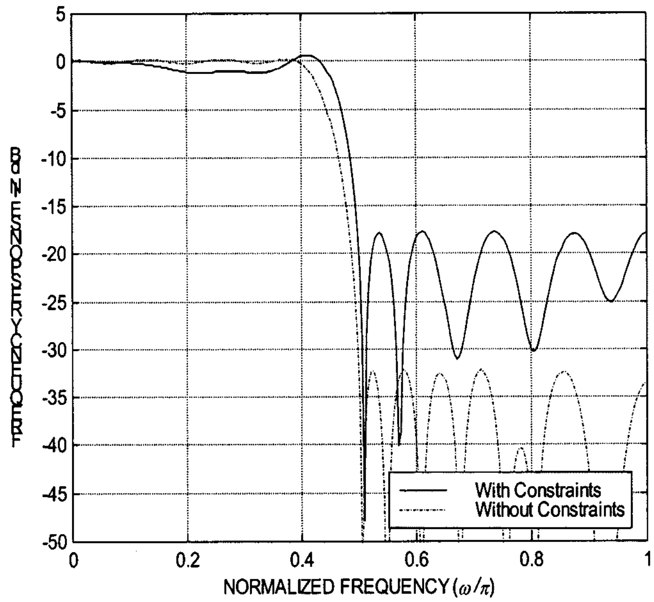


Fig. 7. Frequency response of both filters designed with and without constraints.

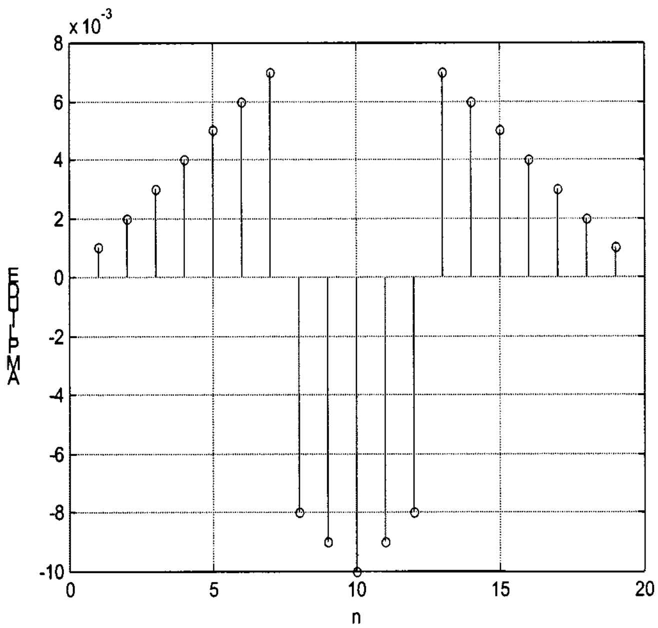


Fig. 8. Appearance of the unwanted signal $s(n)$.

line represents the step response with the step-response constraint applied. Notice that when the constraints are involved, the overshoot and undershoot do not exceed the tolerances prescribed. Fig. 7 shows the frequency response of the filter with and without the step-response constraint. Note that a price has been paid for achieving a superior step-response: the frequency response is inferior to when no step-response constraints are applied. For this example, the VSPM algorithm in (23) required 3,000 iterations for convergence.

Example 3: VSPM versus EM with the Energy Output Constraints

In this example, we compare the VSPM (actually POCS) with the EM. Let the undesired waveform $\mathbf{s}$ be as in Fig. 8.

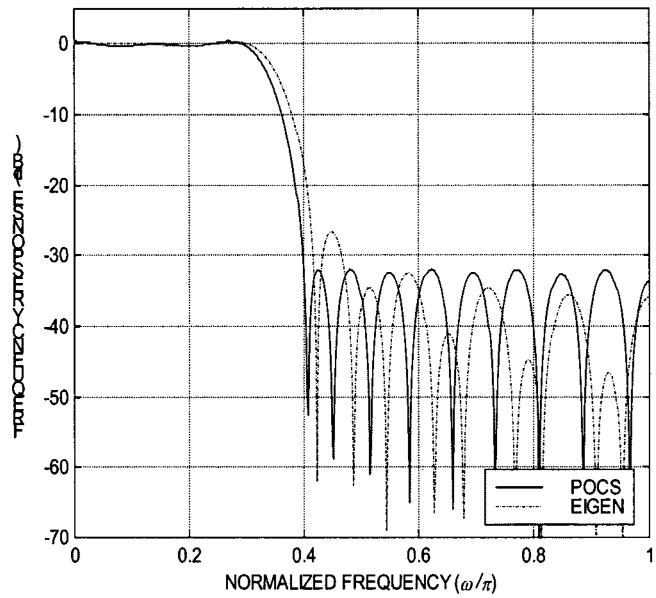


Fig. 9. Frequency response of both EM- and VSPM-designed filters with $C_5(\sigma), \sigma^2 = 5.56 \times 10^{-4}$.

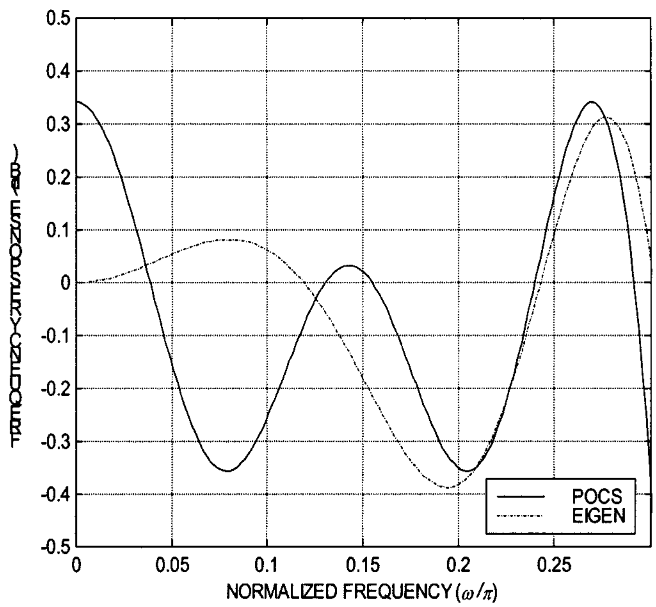


Fig. 10. Passband response of Fig. 9 in detail.

The energy of $\mathbf{s}$ is $||\mathbf{s}||^2 = 6.7 \times 10^{-4}$. Using the eigenfilter method, we designed a low-pass FIR filter with $N = 29$. The passband and stopband edge frequencies are $\omega_p = 0.3\pi$ and $\omega_s = 0.4\pi$. After several trials with different arbitrary δ and γ, the passband and stopband control parameters, we selected these $(\delta = 0.5, \gamma = 0.005)$ in such a way as to preserve as much as possible the desired specifications. The energy of the response to the input signal $\mathbf{s}$ turned to be 5.56×10^{-4}.

Next, we designed an FIR filter using VSPM while keeping the same specifications. We used the set C_5, defined in (26), $\sigma^2 = 5.56 \times 10^{-4}$ and the sets $C_2(\alpha)$ with $\alpha = 0.4015$ and $C_3(\beta)$ with $\beta = 0.025$. Fig. 9 shows the resulting frequency response of both methods. The broken and solid lines represent the EM-and VSPM-based responses, respectively. Fig. 10 shows the passband in detail: note that VSPM produced a filter with

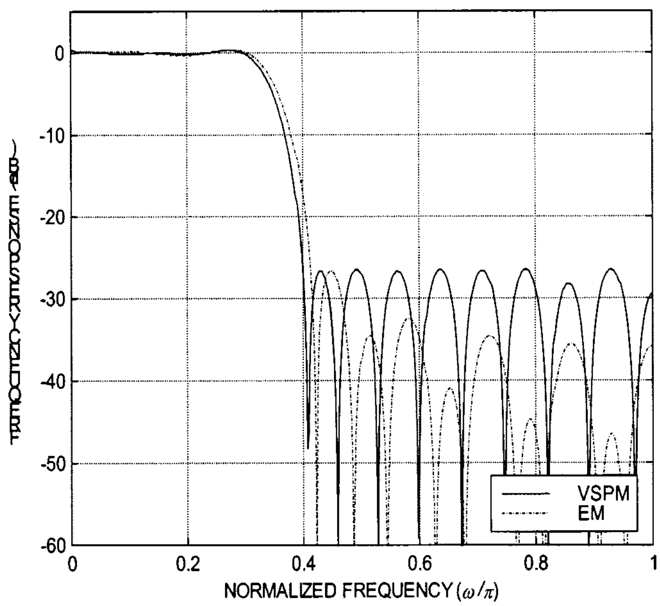


Fig. 11. Frequency response of both EM- and VSPM-designed filters with $\delta = 0.5, \gamma = 0.005$ and VSPM with $C_5(\sigma), \sigma^2 = 5.56 \times 10^{-4}$.

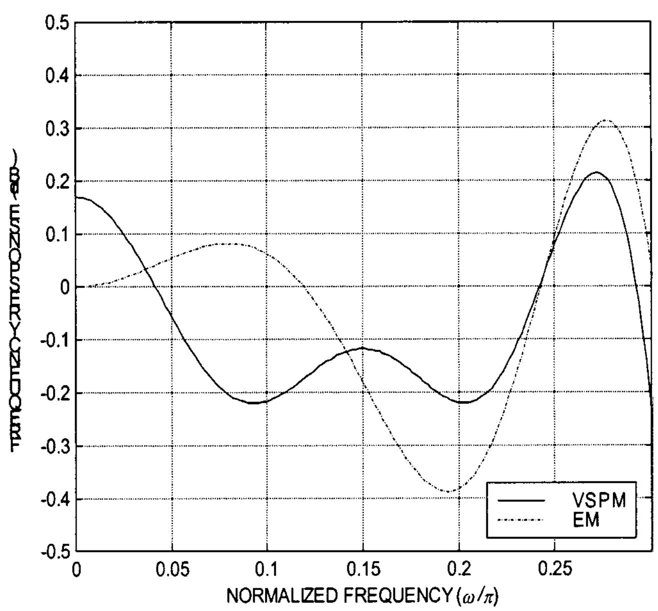


Fig. 12. Passband response of Fig. 11 in more detail.

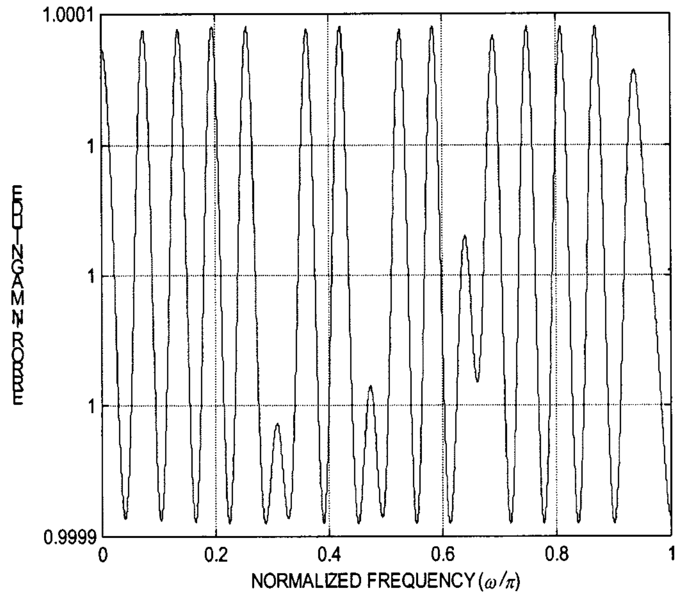


Fig. 13. Magnitude error of the all-pass filter with magnitude deviation $\delta = 9 \times 10^{-5}$.

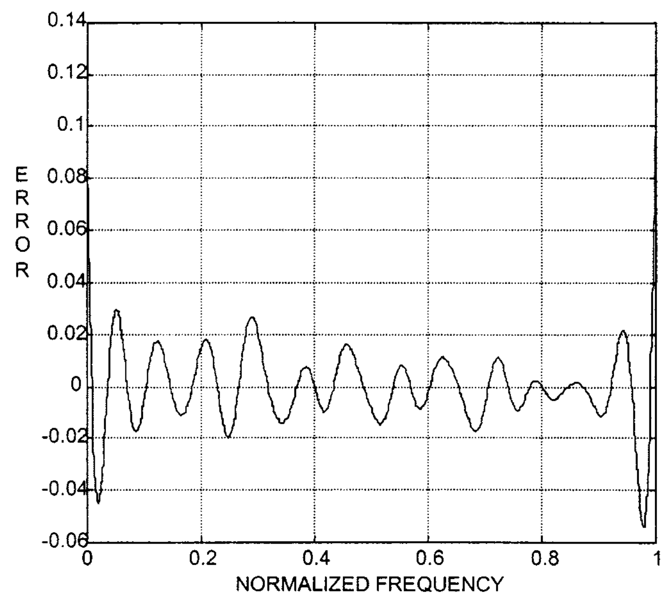


Fig. 14. Group-delay error of the all-pass filter with phase deviation $\varepsilon = 3 \times 10^{-3}$.

better attenuation in the stopband than the EM and comparable fluctuations in the passband. Moreover, the constraints were not satisfied by the EM at the edge frequencies, in contrast to the behavior of the VSPM.

As another experiment, we modify (relax) the conditions of this example by letting the VSPM response in the stopband be as large as the EM (i.e., in the first stopband lobe). The design involves set $C_5(\sigma)$ with $\sigma^2 = 5.56 \times 10^{-4}$, set $C_2(\alpha)$ with $\alpha = 0.025$, and $C_3(\beta)$ with $\beta = 0.047$. Fig. 11 shows the frequency response of both methods. Fig. 12 shows the passband in detail, and clearly demonstrates the superior performance of the VSPM-designed filter. For this example, the VSPM algorithm in (31) required 5,000 iterations for convergence.

Example 4: Phase Compensation Using All-Pass FIR Filter: Comparison of Methods

Assume that in a particular phase-compensation problem, the required phase response of the filter is given by [11]

$$\arg\ H(\omega) = -\left(\frac{N-1}{2}\right)\omega + 2\pi(1-\cos\omega) \tag{51}$$

which corresponds to the following group-delay function

$$\tau(\omega) = -\frac{d\arg H}{d\omega} = \left(\frac{N-1}{2}\right) - 2\pi\sin\omega. \tag{52}$$

We try to approximate an all-pass filter, i.e., $|H(\omega)| = 1$, that has the desired phase characteristics. The filter parameters

TABLE III
PEAK ERROR IN MAGNITUDE AND GROUP DELAY

	LP	DAS 16		EM	VSPM
		Complx.-valued Apprx.	Real-valued Apprx.		
$\max\lvert E_M(\omega)\rvert$	9.31×10^{-4}	9.2×10^{-4}	2.88×10^{-3}	2.08×10^{-4}	0.95×10^{-4}
$\max\lvert E_\tau(\omega)\rvert$	Not available			0.14	0.14

that we select are: filter length $N = 61$, magnitude deviation $\delta = 9 \times 10^{-5}$, and phase deviation $\varepsilon = 3 \times 10^{-3}$ [see (32)]. We can clearly see from Figs. 13 and 14 that the VSPM yields an all-pass filter with a small magnitude error (0.95×10^{-4}). The peak error of the group delay are largest (0.14) at $\omega = 0$ and $\omega = \pi$ whereas they are very small at other frequencies. Table III summarizes the peak error of the four approaches: LP [11], DAS [12], EM [13], and VSPM. The results verify that VSPM yields a filter with better performance. (The group delay peak-error is not available in [11], [12]). For this example, the VSPM algorithm in (48) required 3,500 iterations for convergence.

VII. CONCLUDING REMARKS

In this paper we have reviewed VSPM's and used these methods to design several important classes of FIR filters. In particular, we used VSPM to design linear and arbitrary-phase/magnitude FIR filters subject to various design constraints. We furnished several examples and demonstrated the advantages of VSPM over existing methods.

The main advantages of using VSPM for filter design are:

1) while VSPM generally does not yield optimal solutions, it will furnish solutions that meet all design constraints (assuming that they are consistent) using the same set of mathematical tools;
2) VSPM is easily extended to the design of multi-dimensional filters, an extension that is difficult for other methods.

ACKNOWLEDGMENT

The authors wish to thank the anonymous reviewers whose careful reading and detailed comments on the original manuscript significantly improved the presentation.

REFERENCES

[1] J. H. McClellan and T. Parks, "A unified approach to the design of optimum FIR linear-phase digital filters," *IEEE Trans. Circuit Theory*, vol. CT-20, pp. 697–701, Nov. 1973.

[2] F. Mintzer, "On half-band, third-band, and Nth-band FIR filters and their design," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-30, pp. 734–738, 1982.

[3] J. F. Kaiser and K. Stieglitz, "Design of FIR filters with flatness constraints," *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, pp. 197–200, 1983.

[4] L. R. Rabiner, "Linear program design of finite impulse response (FIR) digital filters," *IEEE Trans. Audio Electroacoust.*, vol. AU-20, pp. 280–288, 1972.

[5] P. P. Vaidyanathan and T. Q. Nguyen, "Eigenfilters: A new approach to least-square FIR filter design and applications including Nyquist filters," *IEEE Trans. Circuits Syst.*, vol. CAS-34, pp. 11–23, Jan. 1987.

[6] B. Nobel and J. W. Daniel, *Applied Linear Algebra*. Englewood Cliffs, NJ: Prentice-Hall, 1977.

[7] S. P. Wu, S. Boyd, and L. Vanderberghe, "FIR filter design via spectral factorization and convex optimization," in *Applied Computational Control, Signals and Circuits*, B. Datta, Ed: Birkhauser, 1997, ch. 2, pp. 51–81.

[8] X. Chen and T. W. Parks, "Design of FIR filters in the complex domain," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, no. 2, pp. 144–153, Feb. 1987.

[9] L. G. Cuthbert, "Optimizing nonrecursive digital filters to nonlinear phase characteristics," *Radio Electron. Eng.*, vol. 44, no. 12, pp. 646–651, Dec. 1974.

[10] A. G. J. Holt, J. Attikiouzel, and R. Bennet, "Iterative technique for designing nonrecursive digital filter with nonlinear phase characteristics," *Radio Electron. Eng.*, vol. 46, no. 12, pp. 589–592, Dec. 1976.

[11] K. Steiglitz, "Design of FIR digital phase networks," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, no. 2, pp. 171–176, Apr. 1981.

[12] N. N. Chit and J. S. Mason, "Complex Chebyshev approximation for FIR digital filters," *IEEE Trans. Signal Processing*, vol. 39, pp. 49–54, Dec. 1991.

[13] T. Q. Nguyen, "The design of arbitrary FIR digital filters using the eigenfilter method," *IEEE Trans. Signal Processing*, vol. 41, Mar. 1993.

[14] A. Abou-Taleb and M. M. Fahmy, "Design of FIR two-dimensional digital filters by successive projections," *IEEE Trans. Circuits Syst.*, vol. CAS-31, pp. 801–805, 1984.

[15] R. M. Merserau, D. B. Harris, and H. S. Hersey, "An efficient algorithm for the design of two-dimensional digital filters," in *Proc. 1974 Int. Symp. Circuits and Systems*, 1975, pp. 443–446.

[16] B. Harris and R. M. Merserau, "A comparison of algorithms for minimax design of two-dimensional linear phase FIR digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-25, pp. 492–500, 1977.

[17] A. E. Cetin, O. N. Gerek, and Y. Yardimici, "Equiripple FIR filter design by the FFT algorithm," *IEEE Signal Processing Mag.*, pp. 60–64, Mar. 1997.

[18] L. M. Bregman, "Finding the common point of convex sets by the method of successive projections," *Dokl. Akad. Nauk. USSR*, vol. 162, no. 3, p. 487, 1965.

[19] L. G. Gubin, B. T. Polyak, and E. V. Raik, "The method of projections for finding the common point of convex sets," *USSR Comput. Math. Phys.*, vol. 7, no. 6, p. 1, 1967.

[20] D. C. Youla and H. Webb, "Image reconstruction by the method of projections onto convex sets—Part I," *IEEE Trans. Med. Imag.*, vol. M1-1, p. 95, 1982.

[21] H. Stark and Y. Yang, *Vector Space Projection Methods: A Numerical Approach to Signal and Image Processing, Neural Nets and Optics*. New York: Wiley, 1998.

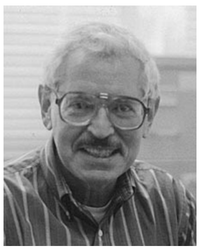
[22] Z. Opial, "Weak convergence of the sequence of successive approximation for nonexpansive mappings," *Bull. Am. Math. Soc.*, vol. 73, p. 591, 1967.

[23] H. Stark and E. T. Olsen, "Projection-based image restoration," *J. Opt. Soc. Am. A*, vol. 9, no. 11, pp. 1914–1919, Nov. 1992.

[24] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos, "Projection-based spatially-adaptive reconstruction of block transform compressed images," *IEEE Trans. Image Processing*, vol. 4, pp. 896–908, July 1995.

Khalil C. Haddad was born in Beirut, Lebanon. He received the B.E. degree form American University of Beirut, the M.Sc. degree from Florida Institute of Technology, Melbourne, and the Ph.D degree from Illinois Institute of Technology, Chicago, in 1984, 1988, and 1998 respectively, all in electrical engineering.

During his Ph.D. studies, he worked in industry. Since 1999, he has been with Lucent Technologies, Holmdel, NJ. His main research interests are in digital signal processing and digital communication.



Henry Stark (M'75–SM'80–F'88) received the B.S.E.E. degree from the City College of New York in 1961, and the M.S.E.E. and Ph.D. degrees from Columbia University, New York, in 1964 and 1968, respectively.

Prior to joining the Engineering and Applied Science Department, Yale University, New Haven, CT, in 1970, he was an Engineer at the Bendix Corporation and Columbia University Electronic Research Labs (later known as the Riverside Research Institute). From 1969 to 1970, he was a Senior Lecturer at the Technion in Israel. He left Yale University in 1977 to join the Department of Electrical, Computer, and Systems Engineering at Rensselaer Polytechnic Institute, where he was promited to Full Professor. In 1988, he was selected as Chair of Electrical and Computer Engineering, Illinois Institute of Technology, and awarded the Bodine Distinguished Chair of Electrical and Computer Engineering. He has published numerous books and papers in optics, signal processing, and probability and random processes, and is currently completing a third edition of *Probability, Random Processes and Estimation Theory* with co-author John Woods.

In 1980, Dr. Stark was awarded a Best Paper Award by the Engineering and Medicine Society of the IEEE for his work on Direct Fourier Tomography. This year, he was awarded the Esther Hoffman Beller Medal for his contributions to optical engineering and optical engineering education. He is a Fellow of the Optical Society of America.

Nikolas P. Galatsanos received the Diploma in electrical engineering from the National Technical University of Athens, Athens, Greece, in 1982, and the M.S.E.E. and Ph.D. degrees from the Electrical and Computer Engineering Department, University of Wisconsin at Madison, in 1984 and 1989, respectively.

Since the fall of 1989, he has been on the faculty of the Electrical and Computer Engineering Department, Illinois Institute of Technology, where he is currently an Associate Professor. His research interests include inverse problems for visual communication and medical imaging applications, and the application of vector space projection methods to signal and image processing problems. He is co-editor of the book *Image Recovery Techniques for Image and Video Compression and Transmission* (Norwell, MA: Kluwer, 1998).

Dr. Galatsanos has served as an Associate Editor for the IEEE TRANSACTIONS ON IMAGE PROCESSING and currently serves as an Associate Editor for the IEEE SIGNAL PROCESSING MAGAZINE.