OPTIMIZED ORTHOGONAL AND BIORTHOGONAL WAVELETS USING LINEAR PARAMETERIZATION OF HALFBAND FILTERS

W.-S. Lu and A. Antoniou

Department of Electrical and Computer Engineering
University of Victoria
Victoria, BC, Canada V8W 3P6
email: wslu@ece.uvic.ca, andreas@ece.uvic.ca

ABSTRACT

A new method for the construction of compactly supported orthogonal and biorthogonal wavelets that can be optimized for specific applications is proposed. Unlike existing methods for the parameterization of orthogonal wavelets, the new method is based on a linear parameterization of all halfband filters. The linear dependence of the halfband filters on a "free" parameter vector considerably facilitates the subsequent numerical optimization through which optimized wavelets are constructed. Biorthogonal wavelets are obtained by factorization of a polynomial representing the parameterized halfband filters while orthogonal wavelets are constructed by imposing nonnegativity on the parameterized halfband filters followed by spectral factorization. The well-known Daubechies orthogonal wavelets and the Cohen-Daubechies-Feauveau biorthogonal wavelets can be viewed as subsets in the parameterized orthogonal and biorthogonal wavelet classes, respectively. Applications of this parameterization-based optimization of wavelets to signal coding are discussed.

1. INTRODUCTION

The compactly supported orthogonal wavelets proposed by Daubechies [1] have been a subject of study since 1988. These wavelets are not only orthogonal but also have the maximum number of vanishing moments. It is well known that there is a direct connection between the vanishing moments and the accuracy of approximating a smooth signal using a linear combination of orthogonal basis functions $\phi(t-k)$, where $\phi(t)$ is the mother scaling function [2]. In certain applications one would trade a certain number of vanishing moments for other features that are more desirable.

This paper is concerned with a method for the design of orthogonal and biorthogonal wavelets with maximum coding gain. The design is accomplished using *linear* parameterization of halfband filters. Parameterizing orthogonal wavelets is not new. Reference [3] describes a factorization method for the subgroup of unitary matrices that are equal to the identity matrix at $z=1$ and shows that the factorization can be used to parameterize various classes of wavelets. Reference [4] presents a method for the parameterization of $M \times M$ causal FIR lossless systems of degree N, and the method was extended in [5] to characterize compactly supported dyadic orthogonal wavelets. However, the coefficients of wavelets are parameterized *nonlinearly* and, as is evidenced in Eqns. (27), (28) of [5], the degree of nonlinearity increases with the filter length.

In this paper, we present a method for the design of optimized orthogonal and biorthogonal wavelet filters. The method is based on a *linear* characterization of halfband lowpass filters. Because of the linearity of this characterization, the constraints involved in the design turn out to be *linear* and, in the case of designing subband coders for maximum coding gain, the objective function is also *linear*. A primal Newton-barrier method that is especially suited for solving the resulting linear programming problem is presented in detail.

2. LINEAR CHARACTERIZATION OF HALFBAND FILTERS

An FIR filter is called a *halfband* filter if its transfer function can be expressed as

$$P(z) = C(z)C(z^{-1}) \tag{1}$$

and

$$P(z) + P(-z) = 2 \qquad \text{for all } z \tag{2}$$

A particularly interesting case is when $C(z)$ in (1) assumes the form

$$C(z) = \left(\frac{1+z^{-1}}{2}\right)^L B_1(z) \tag{3}$$

where $B_1(z)$ is a Kth-order polynomial in z^{-1}. In this case, $C(z)$ has at least L zeros at $\omega = \pi$ and it characterizes an orthogonal, lowpass, analysis wavelet filter with a finite gain at $\omega = 0$ and at least L vanishing moments [2].

A key step in our approach to the parameterization of the wavelet filter represented by $C(z)$ is to find a linear characterization of $P(z)$ that satisfies (2). By (1) and (2), it follows that if the zero-phase transfer function $P(z)$ satisfies the condition $P(e^{j\omega}) \geq 0$ for all ω and assumes the form

$$P(z) = \left(\frac{z^{-1}}{4} + \frac{1}{2} + \frac{z}{4}\right)^L B(z) \tag{4}$$

with

$$B(z) = \sum_{k=-K}^{K} \hat{b}_k z^k, \qquad \hat{b}_k = \hat{b}_{-k} \tag{5}$$

then (1) and (3) hold where polynomial $B_1(z)$ can be obtained by factorization, i.e.,

$$B(z) = B_1(z)B_1(z^{-1}) \tag{6}$$

If we denote

$$B(z) = \sum_{k=-K}^{K} \hat{b}_k z^k = z^{-K} \sum_{k=0}^{2K} b_k z^k, \; b_{2K-k} = b_k \tag{7}$$

0-7803-4455-3/98/$10.00 © 1998 IEEE

and

$$\left(\frac{z^{-1}}{4}+\frac{1}{2}+\frac{z}{4}\right)^L = z^{-L}\sum_{l=0}^{2L} a_l z^l,\ a_{2L-l}=a_l \tag{8}$$

then $P(z)$ in (4) can be written as

$$P(z) = z^{-N}\sum_{s=0}^{2N}\Big(\sum_{\substack{l+k=s\\ l\geq 0,k\geq 0}} a_l b_k\Big) z^{s-N}$$

where $N = K + L$ is assumed to be an odd integer. It follows that $P(z)$ satisfies (2) if and only if

$$\sum_{\substack{l+k=s\\ l\geq 0,k\geq 0}} a_l b_k = \begin{cases} 1 & s = N \\ 0 & s = 1,\ 3,\ \ldots,\ N-2 \end{cases} \tag{9}$$

We now write (9) as

$$\mathbf{A}\mathbf{b} = \mathbf{m} \tag{10}$$

where $\mathbf{A} \in R^{\frac{N+1}{2}\times(K+1)}$ is determined by the a_l's in (8) which depend on the filter's order N and the number of desired vanishing moments L,

$$\mathbf{b} = [b_0\ b_1\ \cdots\ b_K]^T \in R^{K+1}$$

and

$$\mathbf{m} = [0\ \cdots\ 0\ 1]^T \in R^{(N+1)/2}$$

By comparing the number of independent parameters in $B(z)$ with the number of equations in (10), the degrees of freedom one has for a given number of vanishing moments is

$$\eta = \frac{K-L+1}{2} \tag{11}$$

It can be readily verified that $\eta \geq 1$ if and only if

$$N+1 \geq 2(L+1) \tag{12}$$

When (12) is satisfied, the system of linear equations in (10) is underdetermined and the solutions of (10) can be parameterized as

$$\mathbf{b} = \mathbf{b}_0 + \mathbf{V}_\eta \boldsymbol{\phi} \tag{13}$$

where $\mathbf{b}_0$ is the last column of the Moore-Penrose pseudo inverse of $\mathbf{A}$, $\mathbf{V}\eta$ consists of the last η columns of $\mathbf{V}$ which is obtained from the singular value decomposition of $\mathbf{A}$ [6], namely, $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$, and $\boldsymbol{\phi} \in R^\eta$ is a *free* parameter vector which can be used to optimize the wavelet filter for a particular application.

The parameterization of $\mathbf{b}$ given by (13) has several desirable features. First, if the maximum number of vanishing moments is $L = (N+1)/2$, then $\eta = 0$ and (13) assumes the form $\mathbf{b} = \mathbf{b}_0$. This $\mathbf{b}$ leads to the well-known Daubechies orthogonal wavelet filter. Second, if condition (12) is satisfied, then the frequency response of the halfband filter can be expressed as

$$P(e^{j\omega}) = 2^{-L}(1+\cos\omega)^L \mathbf{h}_1(\omega)^T(\mathbf{b}_0 + \mathbf{V}_\eta\boldsymbol{\phi}) \tag{14}$$

where

$$\mathbf{h}_1(\omega) = [1\ \ 2\cos\omega\ \cdots\ 2\cos K\omega]^T \tag{15}$$

This leads to

$$P(e^{j\omega}) = P_0(e^{j\omega}) + \mathbf{h}^T(\omega)\boldsymbol{\phi} \tag{16}$$

where

$$P_0(e^{j\omega}) = 2^{-L}(1+\cos\omega)^L \mathbf{h}_1^T(\omega)\mathbf{b}_0 \tag{17a}$$

is the "fixed" component of $P(e^{j\omega})$, and

$$\mathbf{h}(\omega) = 2^{-L}(1+\cos\omega)^L \mathbf{V}_\eta^T \mathbf{h}_1(\omega) \tag{17b}$$

An important feature of (16) is that the frequency response $P(e^{j\omega})$ is parameterized *linearly.* As will be seen in Sec. 3, this linear characterization of $P(e^{j\omega})$ considerably simplifies the optimization required to construct the wavelet filters that are most suitable for a given application. Finally, we note that $\mathbf{b}_0$ and $\mathbf{V}_\eta$ in (13) are completely determined by the a_l's in (8). Consequently, for the given filter length $N+1$ and the number of vanishing moments L, $P_0(e^{j\omega})$ and $\mathbf{h}(\omega)$ in (16) are *application-independent* and need to be evaluated only once.

3. DESIGN OF OPTIMIZED ORTHOGONAL WAVELETS FOR CODING

Below we develop a suitable design algorithm.

3.1. Design Algorithm

We begin with a remark on the nonnegativity of polynomial $B(z)$ on the unit circle. By (5), the factorization in (6) exists if and only if $B(e^{j\omega}) \geq 0$ for $\omega \in [0,\ \pi]$. From (4), it is quite clear that $B(e^{j\omega})$ is nonnegative if and only if $P(e^{j\omega})$ is so. By (15)–(17), we see that $P(e^{j\omega})$ is nonnegative if and only if

$$\mathbf{h}_1^T(\omega)\mathbf{V}_\eta \geq -\mathbf{h}_1^T(\omega)\mathbf{b}_0 \qquad \text{for } \omega \in [0,\ \pi] \tag{18}$$

In practice, (18) is discretized to become

$$\mathbf{D}\boldsymbol{\phi} \geq \mathbf{p} + \varepsilon\mathbf{e} \tag{19}$$

where

$$\mathbf{D} = \begin{bmatrix} \mathbf{h}_1^T(\omega_1) \\ \mathbf{h}_1^T(\omega_2) \\ \vdots \\ \mathbf{h}_1^T(\omega_M) \end{bmatrix} \mathbf{V}_\eta \in R^{M\times\eta},\ \mathbf{p} = -\begin{bmatrix} \mathbf{h}_1^T(\omega_1) \\ \mathbf{h}_1^T(\omega_2) \\ \vdots \\ \mathbf{h}_1^T(\omega_M) \end{bmatrix} \mathbf{b}_0 \in R^M$$

with $\{\omega_i,\ 0 \leq i \leq M-1\}$ being the frequencies on $[0,\ \pi]$ at which (18) is evaluated. The term $\varepsilon\mathbf{e}$, where ε is a small positive scalar and $\mathbf{e} = [1\ 1\ \cdots\ 1]^T$, is included in (19) in order to assure the nonegativity of $P(e^{j\omega})$ after discretization. Note that the inequality constraints in (19) are *linear* with respect to parameter vector $\boldsymbol{\phi}$. This linearity is of critical importance in developing efficient algorithms for the design of optimized orthogonal and biorthogonal wavelet filters.

Consider the 2-channel orthogonal subband coder illustrated in Fig. 1, where blocks Q model quantization-related signal manipulations, $C(z)$ and $D(z)$ represent the analysis filters, with $C(z)$ given by (3), (6) and (13), and

$$D(z) = z^{-N}C(-z^{-1})$$

The synthesis filters are represented by $\tilde{C}(z)$ and $\tilde{D}(z)$ which are also determined using $C(z)$ as

$$\tilde{C}(z) = z^{-N}C(z^{-1}) \quad \text{and} \quad \tilde{D}(z) = C(-z)$$

For 2-channel orthogonal filter banks, the coding gain maximization problem [7] – [8] is to maximize the variance

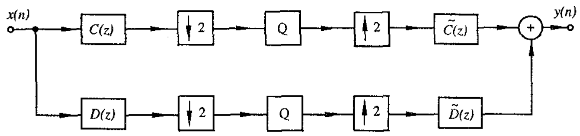

Figure 1. Two-channel orthogonal subband coder.

$$\sigma_y^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |C(e^{j\omega})|^2 S_{xx}(e^{j\omega})d\omega \tag{20}$$

where $S_{xx}(e^{j\omega})$ denotes the power spectral density (PSD) of a zero-mean widesense-stationary random input, and $C(z)$ is given by (3), (6), and (13). Now if (19) holds, then $P(e^{j\omega})$ is nonnegative and (1) and (16) imply that

$$\sigma_y^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} P(e^{j\omega})S_{xx}(e^{j\omega})d\omega = -\mathbf{c}^T\boldsymbol{\phi} + \lambda \tag{21a}$$

where

$$\mathbf{c} = -\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathbf{h}(\omega)S_{xx}(e^{j\omega})d\omega \tag{21b}$$

and λ is a constant given by

$$\lambda = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_0(e^{j\omega})S_{xx}(e^{j\omega})d\omega$$

Therefore, the coding maximization problem can be formulated as the linear programming problem [9]

$$\underset{\boldsymbol{\phi}}{\text{minimize}} \quad \mathbf{c}^T\boldsymbol{\phi} \tag{22a}$$

$$\text{subject to} \quad \mathbf{D}\boldsymbol{\phi} \geq \mathbf{p} + \varepsilon\mathbf{e} \tag{22b}$$

A design algorithm can now be constructed as follows:

Algorithm

Step 1 Input data: filter length $N+1$ (even) and number of vanishing moments L with $N+1 \geq 2L$.

Step 2 Compute the a_l's using (8) and construct matrix $\mathbf{A}$.

Step 3 If $N+1 = 2L$, then set $\mathbf{b} = \mathbf{b}_0$, where $\mathbf{b}_0$ is the last column of $\mathbf{A}^{-1}$, and go to Step 5.

Step 4 If $N+1 > 2L$, evaluate $\mathbf{b}$ using (13) where ϕ is the minimizer of the linear programming problem in (22).

Step 5 Construct $B(z)$ using (7) and perform spectral factorization in (6) to obtain $B_1(z)$.

Step 6 Construct $C(z)$ using (3).

3.2. Primal Newton-Barrier Algorithm

The linear programming problem in (22) can be solved by using the well-known simplex method [9]. Alternatively, it can be converted into a standard-form linear programming problem using a slack-variable technique and then it can be solved using either the simplex method or interior-point methods. In [10][11], an efficient interior-point method known as the primal Newton-barrier (PNB) method was proposed for standard-form linear programming problems. Although problem (22) can be converted into a standard form, the introduction of slack variables increases the problem size and, consequently, degrades the efficiency of the method. This is particularly the case for the problem in (22) where the number of constraints M is much larger than the dimension of ϕ, i.e., η. In what follows, we describe a PNB algorithm that is directly applicable to the problem in (22).

The PNB method modifies the linear objective function in (22a) by adding a logarithmic barrier function that incorporates the constraints in (22b), i.e.,

$$\underset{\boldsymbol{\phi}}{\text{minimize}}\ f(\boldsymbol{\phi}) = \mathbf{c}^T\boldsymbol{\phi} + \tau\sum_{i=1}^{M}\log(\mathbf{d}_i^T\boldsymbol{\phi} - p_i) \tag{23}$$

where $\mathbf{d}_i^T$ is the ith row of $\mathbf{D}$, p_i is the ith entry of $\mathbf{p}$, and $\tau > 0$ is known as the *barrier parameter.* If one starts with an initial $\boldsymbol{\phi}_0$ strictly inside the feasible region, namely, $\mathbf{D}\boldsymbol{\phi}_0 > \mathbf{p}$, then the barrier function is well defined and its effect on the original problem depends largely on the magnitude of τ. For a fixed $\tau > 0$, the PNB algorithm solves the problem in (23) and uses the solution as the initial point for the same problem with the magnitude of τ reduced. As $\tau \to 0$, the algorithm converges to the solution of the original linear program. An attractive feature of this method is that the inequality constraints in (22b) are eliminated by the introduction of the barrier term in (23) so that the linear programming problem is converted into a sequence of unconstrained nonlinear optimization problems which, as will be seen below, are easy to solve. The gradient and Hessian of $f(\boldsymbol{\phi})$ are given by

$$\nabla f(\boldsymbol{\phi}) = \mathbf{c} - \tau\sum_{i=1}^{M}\frac{\mathbf{d}_i}{\mathbf{d}_i^T\boldsymbol{\phi} - p_i} \tag{24}$$

$$\nabla^2 f(\boldsymbol{\phi}) = \tau\sum_{i=1}^{M}\frac{\mathbf{d}_i\mathbf{d}_i^T}{(\mathbf{d}_i^T\boldsymbol{\phi} - p_i)^2} \tag{25}$$

Since $M \gg \eta$, $\nabla^2 f(\boldsymbol{\phi})$ is positive definite. In other words, $f(\boldsymbol{\phi})$ is strictly convex in the entire interior of the feasible region. At iterate ϕ_k, an excellent search direction is given by Newton's method [12] as

$$\mathbf{s}_k = -[\nabla^2 f(\boldsymbol{\phi})]^{-1}\nabla f(\boldsymbol{\phi}) = (\mathbf{R}_k^T\mathbf{R}_k)^{-1}\Big(\sum_{i=1}^{M}\mathbf{r}_k^{(i)} - \frac{\mathbf{c}}{\tau}\Big) \tag{26}$$

where

$$\mathbf{R}_k = \text{diag}\left\{\frac{1}{\mathbf{d}_1^T\boldsymbol{\phi}_k - p_1}, \ \ldots, \ \frac{1}{\mathbf{d}_M^T\boldsymbol{\phi}_k - p_M}\right\}\cdot\mathbf{D}$$

and $\mathbf{r}_k^{(i)}$ is the ith column of $\mathbf{R}_k^T$. Along search direction $\mathbf{s}_k$, the objective function $f(\boldsymbol{\phi}_k + \alpha\mathbf{s}_k)$ is strictly convex on interval $[0,\ \alpha_{\max}]$ where $\alpha_{\max} > 0$ is determined as

$$\alpha_{\max} = \min\left\{\frac{\mathbf{d}_i^T\boldsymbol{\phi}_k - p_i}{-\mathbf{d}_i^T\mathbf{s}_k} \quad \text{for } i \text{ with } \mathbf{d}_i^T\mathbf{s}_k < 0\right\} \tag{27}$$

Therefore, on interval $[0,\ \alpha_{\max}]$ function $f(\boldsymbol{\phi}_k + \alpha\mathbf{s}_k)$ is unimodal, and a line search method such as a quadratic interpolation method can be used to find the $\alpha_k \in [0,\ \alpha_{\max}]$ that minimizes $f(\boldsymbol{\phi}_k + \alpha\mathbf{s}_k)$. The new iterate is determined as $\boldsymbol{\phi}_{k+1} = \boldsymbol{\phi}_k + \alpha_k\mathbf{s}_k$. The iteration continues until $||\alpha_k\mathbf{s}_k||$ becomes less than a given tolerance $\varepsilon_{\text{inner}}$. The solution obtained is then used as the initial point for the next round of iterations where the barrier parameter is set to $\sigma\tau$ where σ is typically in the range [0.05, 0.1]. The algorithm is considered to have converged when the difference between two consecutive solutions of the subproblem in (23) is less than a given tolerance $\varepsilon_{\text{outer}}$.

4. THE BIORTHOGONAL CASE

In the biorthogonal case, (1) is replaced by

$$P(z) = z^N C(z)\tilde{C}(z) \tag{28}$$

and $P(z)$ is required to satisfy (2). If both $C(z)$ and $\tilde{C}(z)$ contain factor $2^{-L}(1+z^{-1})^L$, we can write

$$C(z) = \left(\frac{1+z^{-1}}{2}\right)^L B_1(z), \quad \tilde{C}(z) = \left(\frac{1+z^{-1}}{2}\right)^L \tilde{B}_1(z) \tag{29}$$

and (4) holds with

$$B(z) = z^K B_1(z)\tilde{B}_1(z) \tag{30}$$

In this case, Eqns. (7)–(17) hold and the variance is given by (21). For biorthogonal wavelet filters, $P(e^{j\omega})$ does not need to be nonnegative; hence the linear constraints in (22b) are not present. Instead, constraints on $P(z)$ may be imposed in order to shape the frequency response in the passband or stopband or both, depending upon the application at hand. Because of the linear dependence of $P(e^{j\omega})$ on parameter vector ϕ, these constraints can be expressed as linear inequalities with respect to ϕ; hence the design can also be formulated as a linear programming problem of the type in (22).

5. EXAMPLES

Two MATLAB© test signals, namely, "handel" comprising 73,113 samples and "chirp" comprising 13,129 samples were used as the input signals to design an orthogonal wavelet filter represented by $C(z)$. As is shown in Fig. 2 (dashed line), the PSD of signal "handel" is largely in the low-frequency region. With $N = 5$, $L = 1$, $M = 250$ and $\varepsilon = 1.5 \times 10^{-4}$, the PNB algorithm took eight iterations to converge. The amplitude response of the optimized $C(z)$ is shown in Fig. 2. The PSD signal of "chirp" is largely in the high-frequency region as seen in Fig. 3 (dashed line). With $N = 5$, $L = 0$, $M = 250$, and $\varepsilon = 6 \times 10^{-5}$, the PNB took nine iterations to converge. As is expected, the optimized $C(z)$ acts like a highpass filter, as seen in Fig. 3.

ACKNOWLEDGMENT

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

REFERENCES

[1] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm. Pure Appl. Math.*, vol. 41, pp. 909-996, 1988.

[2] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley, MA: Wellesley-Cambridge Press, 1996.

[3] D. Pollen, "$SU_I(2, F[z,\ 1/z])$ for F a subfield of C," *J. Amer. Math. Soc.*, vol. 3, no. 3, pp. 611-624, July 1990.

[4] P. P. Vaidyanathan, T. Q. Nguyen, Z. Doğanata, and T. Saramäki, "Improved technique for design with lossless polyphase matrices," *IEEE Trans. ASSP*, vol. 37, pp. 1042-1055, July 1989.

[5] H. Zou and A. H. Tewfik, "Parameterization of compactly supported orthogonal wavelets," *IEEE Trans. Signal Processing*, vol. 41, pp. 1428-1431, March 1993.

[6] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed., The Johns Hopkins University Press, 1989.

[7] P. P. Vaidyanathan, "Theory of optimal orthogonal filter banks," in *Proc. ICASSP*, vol. 3, pp. 1487-1490, 1996.

[8] P. P. Vaidyanathan, "Filter banks with maximum coding gain and energy compaction," in *Proc. the 29th Asilomar Conf.*, pp. 36-40, Nov. 1996.

[9] P. E. Gill, W. Murray, and M. H. Wright, *Numerical Linear Algebra and Optimization*, vol. 1, Reading, MA: Addison-Wesley, 1991.

[10] M. H. Wright, " Interior methods for constrained optimization," *Acta Numerica*, pp. 341-407, Cambridge Univ. Press, Cambridge, UK, 1992.

[11] P. E. Gill, W. Murray, M. A. Saunders, J. A. Tomlin, and M. H. Wright, "On projected Newton barrier methods for linear programming and an equivalence to Karmarkar's projective method," *Math. Programming*, vol. 36, pp. 183-209, 1986.

[12] R. Fletcher, *Practical Methods of Optimization*, 2nd ed., New York: John Wiley, 1987.

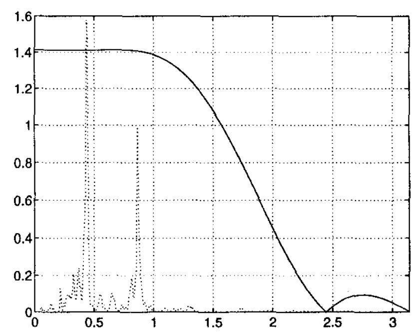

Figure 2. PSD of signal "handel" (dashed line) and the associated $|C(e^{j\omega})|$ (solid line).

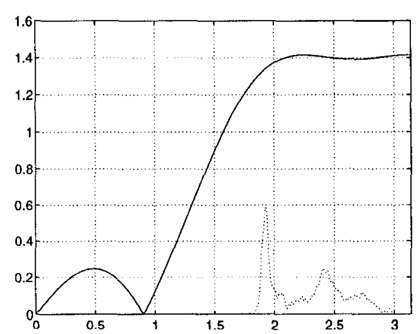

Figure 3. PSD of signal "chirp" (dashed line) and the associated $|C(e^{j\omega})|$ (solid line).