Unconstrained IIR Filter Design Method Using Argument Principle Based Stability Criterion

Aimin Jiang and Hon Keung Kwan
Department of Electrical and Computer Engineering
University of Windsor
401 Sunset Avenue, Windsor, Ontario, Canada N9B 3P4
E-mail: jiang13@uwindsor.ca, kwan1@uwindsor.ca

Abstract—In this paper, we propose an unconstrained IIR digital filter design method in the weighted least-squares (WLS) sense. Unlike some other design methods which utilize approximation techniques to simplify the nonlinear approximation error functions, this method expresses the design problem strictly, and utilizes the general-purpose optimization algorithms to find the minimum point. In order to guarantee the stability of designed filters, a stability criterion is developed from the argument principle of complex analysis. Unlike other frequency-domain stability criteria, it is both sufficient and necessary. Two examples are also presented to demonstrate the effectiveness of the proposed method.

I. INTRODUCTION

When both magnitude and phase (or group delay) responses are simultaneously under consideration, the IIR digital filter design problem becomes nonconvex. The nonconvexity property of the design problem is incurred by the nonlinearity of its frequency response and stability issue. Therefore, approximation techniques have to be deployed. In [1], the denominator part of the approximation error is discarded. This operation can lead to a quadratic approximation error function, and accordingly simplifies the design procedure. However, the optimal result cannot be definitely attained in the minimax sense. Some iterative algorithms [2]-[6] adopt the Steiglitz-McBride (SM) iterative scheme [7]. The main idea behind the SM iterative scheme is to replace the denominator of the approximation error by the previous counterpart, and, thus, the error function at current iteration can be formulated in a quadratic form. Lots of numerical examples have been presented in literatures to demonstrate the effectiveness of this design scheme. Nevertheless, the convergences of the iterative procedures using the SM iterative scheme have not been guaranteed so far. In [8], the Gauss-Newton method is introduced to design IIR digital filters. By using the first-order Taylor series approximation of the frequency response, the approximation error can be cast as a (convex) quadratic function of filter coefficients at each iteration. Like the design methods using the SM iterative scheme, the convergence cannot be strictly guaranteed.

Stability issue is another difficulty encountered during the design procedure. As the denominator order exceeds 2, the stability domain cannot be strictly expressed as a convex set. Some design approaches employ explicit stability criteria in the design procedures, such as the positive-realness [1], [2], [4]-[6] and the Rouché’s theorem [8] based stability criteria. These criteria can be formulated as linear or quadratic inequalities. However, they are only sufficient conditions for the stability, which means some stable filters could be excluded from their admissible solutions. Some other design methods utilize implicit stability criteria. For examples, some design methods [9]-[11] first implement an FIR digital filter which satisfies the specifications, and then approximate the FIR filter by an IIR digital filter in some optimal sense using model-reduction techniques. In general, these approximation procedures can substantially guarantee the stability of designed IIR filters. However, using this strategy, it is difficult to design filters with accurate cutoff frequencies. Recently, a stability criterion based on the argument principle of complex analysis has been introduced in [12]. This criterion is a both sufficient and necessary condition for the stability. In order to incorporate it into the iterative procedure, the higher-order expansion components are truncated, and the resulting stability constraint becomes a linear equality constraint. However, large numbers of simulation results show that this linearized constraint could be invalid in some situations. As an attempt to resolve this problem, in this paper, the argument principle based stability criterion is linearly combined with the weighted least-squares (WLS) cost function as a barrier function.

The paper is organized as follows: In Section II, the design problem in the WLS sense is formulated. In Section III, the argument principle is first reviewed, and then the stability criterion based on the argument principle is developed. In Section IV, an unconstrained WLS design method is proposed. Simulations and conclusions are finally presented in Sections V and VI, respectively.
II. PROBLEM FORMULATION

Let $D(\omega)$ denote the ideal frequency response. The transfer function of an IIR digital filter is defined by

$$H(z) = \frac{P(z)}{Q(z)} = \frac{\sum_{n=0}^{N_p} p_n z^{-n}}{1 + \sum_{n=1}^{N_q} q_n z^{-n}} = \frac{\sum_{n=0}^{N_p} p_n z^{-n}}{\sum_{n=1}^{N_q} q_n z^{-n}} \varphi_n(z)$$

(1)

where $p = [p_0, p_1, \ldots, p_{N_p}]^T$, $q = [q_0, q_1, \ldots, q_{N_q}]^T$ ($q_0 = 1$), and $\varphi(z) = [1, z^{-1}, \ldots, z^{-L}]$. Our design task is to find a stable IIR digital filter, which can best approximate $D(\omega)$ in the WLS sense. We assume all the coefficients are real values.

In the WLS sense, the cost function of the IIR filter design problem can be expressed as

$$J_{WLS}(p, q) = \int_{-\pi}^{\pi} W(\omega) \| E(\omega) \|^2 d\omega$$

(2)

where $W(\omega)$ represents a prescribed nonnegative weighting function, and the complex error $E(\omega)$ is defined by

$$E(\omega) = H(e^{j\omega}) - D(\omega)$$

(3)

Due to the existence of the denominator polynomial $Q(e^{j\omega})$ in $H(e^{j\omega})$, $J_{WLS}(p, q)$ is neither a linear nor quadratic function of filter coefficients. Thus, nonlinear optimization techniques have to be deployed to find the minimum point. On the other hand, there is more than one global minimum point on the error-performance surface. Therefore, the globally minimum point cannot be definitely achieved at each design.

III. ARGUMENT PRINCIPLE BASED STABILITY CRITERION

In this section, the argument principle will be first reviewed. Then the stability criterion based on the argument principle will be developed.

A. Argument Principle

Suppose $f(z)$ is analytic in the region $R$ enclosed by a contour $C$ in the $z$-plane except a finite number of poles. Let $N_z$ be the number of zeros and $N_p$ the number of poles of $f(z)$ in $R$, where each zero and pole is counted according to its multiplicity. Then we have

$$N_z - N_p = \frac{1}{2\pi j} \int_C \frac{f'(z)dz}{f(z)}$$

(4)

This result is called the argument principle.

For developing the stability criterion, $f(z)$ is written as

$$f(z) = z^M Q(z) = \sum_{n=M}^{M} q_n z^{-n}, \quad q_0 = 1$$

(5)

Here, the contour $C$ is an origin-centered circle, i.e., $C = \{z: |z| = R, 0 \leq R \leq 1\}$, where $R$ denotes the specified maximum pole radius. Then according to the argument principle described above, if all poles lie inside $C$ if and only if the following condition is satisfied

$$M = \frac{1}{2\pi j} \int_C \frac{f'(z)dz}{f(z)}$$

(6)

The integral in (6) is computed counterclockwise along $C$. Taking (5) into (6), we have

$$\frac{1}{2\pi j} \int_C \frac{f'(z)dz}{f(z)} = \frac{1}{2\pi} \int_C d\arg f(z)$$

(7)

$$= \frac{1}{2\pi} \int_C d\arg J_r e^{\gamma} = \int_C d\arg Q(z) = 0$$

(8)

Thus, the stability criterion (8) of an IIR digital filter can be stated as: An IIR digital filter with the denominator $Q(z)$ is stable, if and only if the total change in the argument of $Q(z)$ is equal to 0, when the integral is carried out along $C$ counterclockwise.

B. Stability Criterion

The denominator polynomial can be expressed as $Q(re^{j\omega}) = Q_d(re^{j\omega}) + jQ_I(re^{j\omega})$, where $Q_d(re^{j\omega})$ and $Q_I(re^{j\omega})$ denote real and imaginary parts of $Q(re^{j\omega})$, respectively. Then the argument of $Q(re^{j\omega})$ can be computed by

$$\arg Q(re^{j\omega}) = \arctan \frac{Q_I(re^{j\omega})}{Q_d(re^{j\omega})}$$

(9)

By taking differentials with respect to $\omega$ on both sides of (9), we can obtain

$$\frac{d}{d\omega} \arg Q(re^{j\omega}) = \frac{q^T A \Psi(re^{j\omega}) q}{|Q(re^{j\omega})|^2}$$

(10)

where

$$A = \text{diag} \{0, 1, \ldots, M\}$$

(11)

$$\Psi(re^{j\omega}) = \varphi_d(re^{j\omega}) \varphi_d^T(re^{j\omega})$$

(12)

In (11), $\text{diag} \{a_1, a_2, \ldots, a_L\}$ represents a diagonal matrix of size $L$-by-$L$ with $a_i$ ($i = 1, 2, \ldots, L$) on its $i$th diagonal. By computing the integral of (8) over $[-\pi, \pi]$, the stability criterion can be strictly expressed as

$$J_{\arg}(r, q) = \int_{-\pi}^{\pi} \tau_{\arg}(\omega) d\omega = 0$$

(13)

where

$$\tau_{\arg}(\omega) = \frac{q^T G_{\arg}(\omega) q}{|Q(re^{j\omega})|^2}$$

(14)

$$G_{\arg}(\omega) = \frac{A \Psi(re^{j\omega}) + \Psi(re^{j\omega}) A}{2}$$

(15)

If there are $L \leq M$ roots of $Q(z)$ lie outside $C$ and $(M-L)$ roots inside $C$, it can be easily verified that $J_{\arg}(r, q) = 2\pi L$. Then, given a denominator $q$, $J_{\arg}(r, q)$ has a stair shape with respect to $r$. 

867
IV. UNCONSTRAINED DESIGN METHOD

In order to incorporate (13) into the design procedure, we construct a new cost function by linearly combining the pure WLS cost function $J_{\text{WLS}}(p, q)$ with $J_a(r, q)$, i.e.,

$$J(p, q) = (1 - \lambda)J_{\text{WLS}}(p, q) + \lambda J_a(r, q)$$

where $\lambda \in [0, 1]$. As mentioned earlier, when all poles lie inside $C$, $J_a(r, q)$ is always equal to 0. As any pole moves outside $C$, $J_a(r, q)$ will increase rapidly. Then, the poles outside $C$ will be penalized by the penalty term $J_a(r, q)$. We will use the general purpose optimization algorithms to solve this unconstrained design problem. Normally, such algorithms require users to provide subroutines to calculate the value, the gradient, and the Hessian of the cost function at a given point. Therefore, here we provide the formulas to calculate the gradient and the Hessian of (16):

$$\nabla_J p, q = 2(1 - \lambda) \int_{-\pi}^{\pi} W(\omega) E(\omega) \phi_\omega(\omega) d\omega$$

$$\nabla_J^2 p, q = -2(1 - \lambda) \int_{-\pi}^{\pi} W(\omega) H(\omega) E(\omega) \phi_\omega(\omega) d\omega$$

$$\nabla_J^3 p, q = 2(1 - \lambda) \int_{-\pi}^{\pi} W(\omega) E(\omega) \phi_\omega(\omega) d\omega$$

where $\phi_\omega(\omega)$ is given by (2) and $W(\omega)$ is the weight function. The integrals in (17)-(21) can be simplified. Using the finite sum to replace the integral in (17), we obtain

$$
\begin{align*}
J_{\text{WLS}}(p, q) &= \sum_{l=0}^{N} W(q) E(l) Q(l) d\omega \\
&= \sum_{l=0}^{N} W(q) E(l) Q(l) e^{j\omega l} d\omega \\
&= \frac{1}{L} \sum_{l=0}^{N} W(q) E(l) Q(l) e^{j\omega l} d\omega \\
&= \frac{1}{L} \sum_{l=0}^{N} W(q) E(l) Q(l) e^{j\omega l} \\
&= \frac{1}{L} \sum_{l=0}^{N} W(q) E(l) Q(l) e^{j\omega l} \\
&= \frac{1}{L} \sum_{l=0}^{N} W(q) E(l) Q(l) e^{j\omega l}
\end{align*}
$$

In the similar way, the computations of $J_a(r, q)$ and all the integrals in (17)-(21) can be simplified. Using the finite sum to replace the integral in (17), we obtain

$$
\begin{align*}
\frac{\partial J(p, q)}{\partial p} &= 2(1 - \lambda) \int_{-\pi}^{\pi} W(\omega) E(\omega) Q(\omega) e^{j\omega l} d\omega \\
&= \frac{2(1 - \lambda)}{L} \sum_{l=0}^{N} W(q) E(l) Q(l) e^{j\omega l}
\end{align*}
$$

Note that the summation above can be replaced by an inverse discrete Fourier transform (IDFT), if $L > N$. Then the computational cost can be further saved by applying the inverse FFT algorithm to calculate the summation. Similar techniques can be employed to compute the integrals in (18)-(21).

V. SIMULATIONS

In this section, two examples are presented to illustrate the effectiveness of the proposed method. We implement the proposed method in MATLAB. The unconstrained design problem is solved by the unconstrained optimization command fminunc. Both gradient and Hessian of the cost function are utilized to find the solution. In our designs, the initial guesses used in fminunc are always chosen as $p = [1, 1, \ldots, 1]^T$ and $q = [1, 0, 0, \ldots, 0]^T$. In order to evaluate the performances, peak and least-squares $(L_2)$ errors of magnitude

<table>
<thead>
<tr>
<th>Method</th>
<th>Passband (Peak/ L2 in dB)</th>
<th>Stopband (Peak/ L2)</th>
<th>Mag (Peak/ L2 in dB)</th>
</tr>
</thead>
</table>

Table I. Peak and $L_2$ Errors of Example 1.
(Mag) and group delay (GD) responses are measured over the frequency bands of interest.

A. Example 1

The first example is to design a halfband highpass filter. The specifications are the same as those proposed in [4]. The ideal frequency response is given by

$$D(\omega) = \begin{cases} e^{-j12\omega} & 0.525\pi \leq \omega < \pi \\ 0 & 0 \leq \omega \leq 0.475\pi \end{cases}$$

The filter orders are set as $M = N = 14$. The weighting function is chosen as $W(\omega) = 10$ for $\omega \in [0.525\pi, \pi]$, $W(\omega) = 1$ for $\omega \in [0, 0.475\pi]$, and 0 otherwise. Other parameters are set as $r = 0.98$ and $\lambda = 10^{-7}$. The magnitude and group delay responses of the filter designed by the proposed method are shown in Fig. 1. The maximum pole radius of the obtained filter is 0.9247. Using the same specifications, we also design the filter by the WLS method proposed in [4]. The maximum pole radius of the obtained filter is 0.9322. The design results are also given in Table II. For comparison, all the peak and $L_2$ errors are summarized in Table I.

B. Example 2

The second example taken from [12] is to design a lowpass filter. The ideal frequency response is given by

$$D(\omega) = \begin{cases} e^{-j12\omega} & 0 \leq \omega \leq 0.4\pi \\ 0 & 0.56\pi \leq \omega < \pi \end{cases}$$

The filter orders are chosen as $N = 15$ and $M = 4$. The weighting function is set as $W(\omega) = 1$ for $\omega \in [0, 0.4\pi]$, $W(\omega) = 2.6$ for $\omega \in [0.56\pi, \pi]$, and 0 otherwise. As the author proposed in [12], the admissible maximum pole radius is set as $r = 0.84$. The parameter $\lambda$ is selected as $10^{-6}$. The frequency responses of the designed filter are given in Fig. 2. The maximum pole radius of the obtained filter is 0.7054. We also adopt the WLS method using the linearized argument principle based stability constraint [12] to design the filter under the same specifications. The maximum pole radius of the obtained filter is 0.7233. The design results are also given in Table II. The peak and $L_2$ errors are summarized in Table II. It can be seen that the proposed method can achieve better performances in the WLS sense.

VI. CONCLUSIONS

In this paper, an unconstrained WLS method has been proposed for designing stable IIR digital filters. Unlike some other design methods, this method does not utilize any approximation technique. In order to guarantee the stability of designed filters, an argument principle based stability criterion is developed. This stability criterion is both sufficient and necessary for the stability. In order to find a stable IIR digital filter, a new cost function is constructed by linearly combining the stability criterion with the pure WLS cost function. This stability criterion serves as a barrier function to control the poles’ positions. Utilizing the general-purpose optimization algorithms, the design problems without any constraints can be efficiently solved. Two examples are presented to illustrate the effectiveness of the proposed method.