# A New Approach to the Realization of Low-Sensitivity IIR Digital Filters

P. P. VAIDYANATHAN, MEMBER, IEEE, SANJIT K. MITRA, FELLOW, IEEE, AND YRJÖ NEUVO, SENIOR MEMBER, IEEE

*Abstract*—A new implementation of an IIR digital filter transfer function is presented that is structurally passive and, hence, has extremely low pass-band sensitivity. The structure is based on a simple parallel interconnection of two all-pass sections, with each section implemented in a *structurally lossless* manner. The structure shares a number of properties in common with wave lattice digital filters. Computer simulation results verifying the low-sensitivity feature are included, along with results on roundoff noise/dynamic range interaction. A large number of alternatives is available for the implementation of the all-pass sections, giving rise to the well-known wave lattice digital filters as a specific instance of the implementation.

## I. INTRODUCTION

DIGITAL filtering has been a major area of research activity for the last several years, beginning with the pioneering work by Kaiser [1] in the 1960's. A large number of contributions has appeared in the literature [2]–[31] concerning the design and the practical problems associated with an actual implementation. Among these, a number of contributions emphasizes the importance of low-sensitivity digital filter structures requiring very few bits per multiplier coefficient because such structures lead to efficient implementations. A major step in this direction was made in 1971 when Fettweis [2] introduced the concept of wave digital filters. These filters are obtained from classical doubly terminated *LC* networks, and inherit the robustness properties into the digital domain [3], [4]. A specific class of wave filters, called the wave lattice digital filters [8], are obtained by translating continuous-time *LC* lattice filters into the digital domain and are particularly known for extremely low pass-band sensitivity. The resulting digital filters can be looked upon as a parallel connection of all-pass filters, each all-pass branch being synthesized by an ingenious cascading of suitably chosen wave adaptors. Wegener [23] and Gaszi [27] have also made studies in this same direction.

The use of wave filters in multirate applications is also

known [34]. In addition, Constantinides has noticed the use and applicability of all-pass sections for such applications [28], [29], and has proposed efficient filter designs based on parallel combinations of all-pass functions of the form $A(z^2)$ [and, in general, $A(z^N)$], where the overall low-pass design has a cutoff frequency of about $\pi/2$. Constantinides has also clearly argued the reason for the low-sensitivity properties of these structures by establishing their equivalence to continuous-time doubly terminated lossless lattice networks [28]. Related results in this direction have also been reported by Ansari and Liu [30].

It is clear, therefore, that all the above-mentioned contributions derive their excellent robustness properties because of their close relation to the continuous-time doubly terminated lossless networks. In a recent contribution [20], it has been shown that low-sensitivity digital filters can also be designed directly in the $z$ domain, without recourse to continuous-time prototype networks. Such an independent $z$-domain approach also gives rise to wave digital filters, orthogonal filters, and digital lattice structures as special cases [20]–[22]. Basic to such a synthesis procedure are the concepts of structural boundedness and the LBR two-pair extraction [20]. Structural boundedness is essentially a generalization of the pseudopassivity concept [3]. The pseudopassivity of a wave digital filter is the consequence of passivity of internal building blocks, whereas structural boundedness does not necessarily require that the internal building blocks be (passive) equivalents of continuous-time passive elements.

The purpose of this paper is to develop a procedure for synthesizing a digital filter transfer function as a parallel connection of two all-pass sections without any reference to continuous-time *LC* synthesis. Certain digital filter transfer functions (to be clearly spelled out) can be decomposed in this fashion, and are shown to satisfy the "structural passivity requirement" so that they have low pass-band sensitivity. Moreover, any digital filter transfer function for which we can find a wave digital lattice filter can also be directly handled in the $z$ domain by making use of the proposed approach.

In our proposed method, each of the two all-pass sections can be implemented as a cascade of first- and second-order building blocks. The second-order section can be efficiently implemented with only two multipliers, whereas the first-order section requires only one multiplier, as described in [19]. The theory underlying the re-

alization method is based on a very useful property satisfied by certain digital transfer functions that have transmission zeros either on the unit circle or in reciprocal pairs. The corresponding implementations have extremely low coefficient sensitivity in the passband because of the inherent *structural boundedness*. This structural boundedness arises out of the *structural losslessness* of the all-pass networks proposed in [19]. Although the theory can be derived by translating certain well-known power-balance concepts in classical filter theory [8], [35], our derivation is entirely in the $z$ domain. We feel that this leads to a self-contained presentation, emphasizing the simplicity of the concepts involved.

In Section II, the role of *structural boundedness* [20] in low-sensitivity implementations is reviewed, and then it is shown that a parallel connection of all-pass filters leads to such boundedness, under certain conditions. In Section III, we derive conditions under which a digital filter transfer function can be expressed as a sum of two stable all-pass functions. Fortunately, these conditions turn out to be mild, and are satisfied under most situations. Section IV includes a general discussion of implementation considerations. In Section V, design examples are presented, together with computer simulations to demonstrate the excellent pass-band sensitivity properties of the implementations. Finally, in Section VI, roundoff noise/dynamic range interaction in the new structures is analyzed, and a method is proposed for obtaining a cascade of optimal all-pass sections having the largest signal-to-roundoff-noise ratio under scaled conditions.

## II. REVIEW OF STRUCTURAL BOUNDEDNESS

Consider the transfer function of an $N$th-order IIR filter:

$$G(z) = \frac{P(z)}{D(z)} = \frac{p_0 + p_1 z^{-1} + \cdots + p_N z^{-N}}{1 + d_1 z^{-1} + \cdots + d_N z^{-N}} \quad (1)$$

where the coefficients $p_i$ and $d_i$ are real. We wish to design a structure, with multiplier coefficients $m_0, m_1, \cdots$, such that the sensitivity of $|G(e^{j\omega})|$ with respect to each $m_i$ is very small in the passband. Let us now assume that the structure is such that, regardless of the actual values of the multipliers $m_i$, the quantity $|G(e^{j\omega})|$ is bounded above by a fixed constant, say, unity

$$|G(e^{j\omega})| \leq 1 \quad \text{for all } \omega. \quad (2)$$

In other words, assume that the structure forces an upper bound on $|G(e^{j\omega})|$, regardless of what the values of the multipliers are, as long as the multipliers are within a certain range. Such implementations are called *structurally passive* or *structurally bounded* [20]. Now, assume that a transfer function $G(z)$ has been implemented in a *structurally bounded* manner. Assume further that $|G(e^{j\omega_k})|$ is unity for certain frequencies $\omega_k$ in the passband. If now a multiplier $m_i$ is perturbed, the quantity $|G(e^{j\omega_k})|$ can only decrease, regardless of the sign of the perturbation. Thus, a plot of $|G(e^{j\omega_k})|$ with respect to $m_i$ has zero slope at the nominal values of $m_i$, and this holds for each multiplier
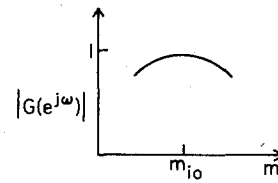


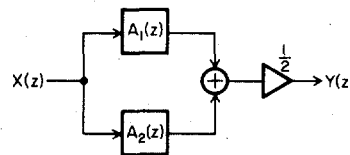Fig. 1. Demonstrating the low-pass-band-sensitivity property.



Fig. 2. Parallel connection of all-pass filters.

and each $\omega_k$ (see Fig. 1). In effect, we have zero first-order sensitivity at frequencies $\omega = \omega_k$:

$$\left. \frac{\partial |G(e^{j\omega_k})|}{\partial m_i} \right|_{m_i = m_{i_0}} = 0 \quad \forall i, \forall k. \quad (3)$$

In addition, if we have a number of closely spaced points in the passband where $|G(e^{j\omega})| = 1$, then we can expect low sensitivity over the passband. It is therefore clear that the fundamental requirement for low pass-band sensitivity is structural boundedness. Note that the above argument is analogous to the well-known *Orchard's argument* [33] in classical filter theory.

A stable transfer function $G(z)$ with real coefficients satisfying (2) is called a bounded real (BR) function. In addition, if (2) holds with equality for all $\omega$, then $G(z)$ is called a lossless bounded real (LBR) transfer function, more commonly known as a stable all-pass function. If a structure is such that the transfer function remains BR in spite of parameter quantization, then it has a low-sensitivity property.

Let us now consider a stable all-pass function $A_1(z)$ of order $m$:

$$A_1(z) = \frac{a_m + a_{m-1} z^{-1} + \cdots + z^{-m}}{1 + a_1 z^{-1} + \cdots + a_m z^{-m}} \quad (4)$$

where $a_k$ are real. Notice that the numerator and denominator are mirror images of each other. There exists a number of well-known structures [12], [18], [19], [29] which implement such functions in such a manner that, in spite of multiplier quantization, the mirror-image property is preserved. Consequently, the all-pass property, i.e., the equality

$$|A_1(e^{j\omega})| = 1, \quad (5)$$

holds regardless of quantization. Such implementations will be called *structurally lossless*.

Next consider an interconnection of two stable all-pass filters $A_1$ and $A_2$ as shown in Fig. 2. The overall transfer function is

$$G(z) = \tfrac{1}{2}[A_1(z) + A_2(z)] \quad (6)$$

where $A_1(z)$ is as in (4) and $A_2(z)$ is given by

$$A_2(z) = \frac{b_n + b_{n-1}z^{-1} + \cdots + z^{-n}}{1 + b_1 z^{-1} + \cdots + b_n z^{-n}}. \quad (7)$$

Since $A_1(z)$ and $A_2(z)$ are all-pass functions, we can write

$$A_1(z) = z^{-n_1} \frac{\hat{D}_1(z)}{D_1(z)}, \quad A_2(z) = z^{-n_2} \frac{\hat{D}_2(z)}{D_2(z)} \quad (8)$$

where $n_1$, $n_2$ are nonnegative integers, and $\hat{D}_k(z)$ denotes the mirror image of $D_k(z)$. Assume that $A_1(z)$ and $A_2(z)$ are in minimal form, i.e., $D_k(z)$ and $\hat{D}_k(z)$ do not have uncancelled common zeros. Then (6) implies

$$G(z) = \frac{1}{2} \left[ \frac{z^{-n_1} D_2(z) \hat{D}_1(z) + z^{-n_2} D_1(z) \hat{D}_2(z)}{D_1(z) D_2(z)} \right]. \quad (9)$$

Assume that $(1 - \alpha z^{-1})$ is a factor of $D_1(z)$, and is at the same time a factor of the numerator in (9). This would imply that $(1 - \alpha z^{-1})$ is necessarily a factor of $D_2(z) \hat{D}_1(z)$. But since $A_1(z)$ is assumed to be in minimal form, this implies that $(1 - \alpha z^{-1})$ must be a factor of $D_2(z)$ rather than $\hat{D}_1(z)$. Thus, if we assume that $D_1(z)$ and $D_2(z)$ do not have common factors, and that $A_k(z)$ are in minimal form, then there is no pole-zero cancellation in (6), and $G(z)$ has order $N = m + n$.

Next, it is clear from (6) that on the unit circle,

$$G(e^{j\omega}) = \frac{1}{2}[e^{j\theta_1(\omega)} + e^{j\theta_2(\omega)}] \quad (10)$$

where $\theta_1(\omega)$ and $\theta_2(\omega)$ are real-valued functions of $\omega$. Thus,

$$|G(e^{j\omega})| = \frac{1}{2}|1 + e^{j[\theta_2(\omega) - \theta_1(\omega)]}| \quad (11)$$

which shows that $|G(e^{j\omega})|$ cannot be all-pass unless $\theta_2(\omega) = \theta_1(\omega)$ for all $\omega$. Equivalently, unless $A_1(z) = A_2(z)$, $G(z)$ is not all-pass.

Equation (11) reveals other important information: thus, if $A_1(z)$ and $A_2(z)$ are implemented such that they remain all-pass in spite of parameter quantization, then $|G(e^{j\omega})| \leq 1$ for all $\omega$, in spite of this quantization. Accordingly, the structural losslessness of $A_k(z)$ induces structural boundedness of $G(z)$. It therefore follows that Fig. 2 leads to a simple low-sensitivity implementation.

Thus, given an arbitrary BR digital filter transfer function $G(z)$, if it is possible to decompose it as in (6), we have a simple means of achieving low pass-band sensitivity. In the next section, we investigate this possibility.

### III. IMPLEMENTATION OF A DIGITAL TRANSFER FUNCTION AS A SUM OF TWO ALL-PASS FUNCTIONS

Let us consider a typical $N$th-order BR transfer function $G(z) = P(z)/D(z)$ as in (1), with $|G(e^{j\omega})|$ as in Fig. 3. Assume that $P(z)$ has linear phase. This is typical of most digital filter transfer functions (because the zeros are usually located on the unit circle) and is therefore only a mild restriction. To be specific, let $P(z)$ be symmetric, i.e., $p_k = p_{N-k}$. Now consider another transfer function $H(z)$:
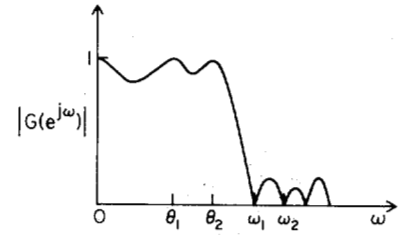


Fig. 3. A typical low-pass response.

$$H(z) = \frac{Q(z)}{D(z)} = \frac{q_0 + q_1 z^{-1} + \cdots + q_N z^{-N}}{1 + d_1 z^{-1} + \cdots + d_N z^{-N}} \quad (12)$$

where $H(z)$ is defined such that

$$|H(e^{j\omega})|^2 = 1 - |G(e^{j\omega})|^2. \quad (13)$$

In other words, $H(z)$ is complementary to $G(z)$. In terms of the $z$ variable, we therefore have

$$\tilde{P}(z) P(z) + \tilde{Q}(z) Q(z) = \tilde{D}(z) D(z) \quad (14)$$

where $\tilde{P}(z) = P(z^{-1})$, and so on. Let us now make the following nontrivial assumption: $G(z)$ is such that $Q(z)$ satisfying (14) is antisymmetric, i.e., $q_k = -q_{N-k}$. Now, by symmetry of $P(z)$ and antisymmetry of $Q(z)$, we have

$$\tilde{P}(z) \triangleq P(z^{-1}) = z^N P(z) \quad (15a)$$

$$\tilde{Q}(z) \triangleq Q(z^{-1}) = -z^N Q(z). \quad (15b)$$

Hence, (14) becomes

$$[P(z) + Q(z)][P(z) - Q(z)] = D(z) z^{-N} D(z^{-1}). \quad (16)$$

Moreover,

$$P(z^{-1}) + Q(z^{-1}) = z^N[P(z) - Q(z)]. \quad (17)$$

Hence, the zeros of $P(z) + Q(z)$ are the reciprocals of the zeros of $P(z) - Q(z)$. As the given transfer function $G(z)$ is stable, none of its poles are on the unit circle, hence,

$$P(z) + Q(z) \neq 0, \quad |z| = 1. \quad (18)$$

Let $z_1, z_2, \cdots, z_r$ be the zeros of $P(z) + Q(z)$ inside the unit circle, and let $z_{r+1}, \cdots, z_N$ be those outside. Then, from (16) it is clear that

$$D(z) = \prod_{k=1}^{r} (1 - z^{-1} z_k) \prod_{k=r+1}^{N} (1 - z^{-1} z_k^{-1}). \quad (19)$$

Thus, we can rewrite (16) as

$$[P(z) + Q(z)][P(z) - Q(z)]$$

$$= \left[ \prod_{k=1}^{r} (1 - z^{-1} z_k) \prod_{k=r+1}^{N} (1 - z^{-1} z_k^{-1}) \right]$$

$$\cdot \left[ \prod_{k=1}^{r} (z^{-1} - z_k) \prod_{k=r+1}^{N} (z^{-1} - z_k^{-1}) \right] \quad (20)$$

which enables us to identify

$$P(z) + Q(z) = \alpha \prod_{k=1}^{r} (1 - z^{-1} z_k) \prod_{k=r+1}^{N} (z^{-1} - z_k^{-1})$$
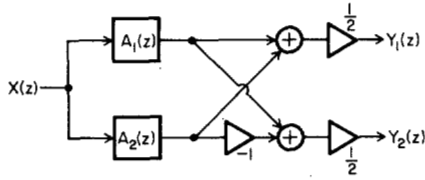
$$(21)$$

Fig. 4. Realizing the doubly complementary pair.

$$P(z) - Q(z) = \frac{1}{\alpha} \prod_{k=1}^{r} (z^{-1} - z_k) \prod_{k=r+1}^{N} (1 - z^{-1}z_k^{-1})$$

(22)

where $\alpha$ is a real constant. This leads to the equations

$$G(z) + H(z) = \frac{P(z) + Q(z)}{D(z)} = \alpha \prod_{k=r+1}^{N} \left( \frac{z^{-1} - z_k^{-1}}{1 - z^{-1}z_k^{-1}} \right)$$

(23)

$$G(z) - H(z) = \frac{P(z) - Q(z)}{D(z)} = \frac{1}{\alpha} \prod_{k=1}^{r} \left( \frac{z^{-1} - z_k}{1 - z^{-1}z_k} \right).$$

(24)

Thus, we have the following set of equations:

$$G(z) + H(z) = \alpha A_1(z)$$

(25a)

$$G(z) - H(z) = \frac{1}{\alpha} A_2(z)$$

(25b)

where $A_1(z)$ and $A_2(z)$ are stable all-pass functions of orders $N - r$ and $r$, respectively,

$$A_1(z) = \prod_{k=r+1}^{N} \left( \frac{z^{-1} - z_k^{-1}}{1 - z^{-1}z_k^{-1}} \right),$$

$$A_2(z) = \prod_{k=1}^{r} \left( \frac{z^{-1} - z_k}{1 - z^{-1}z_k} \right).$$

(26)

In view of the condition (13), it is easily verified that $\alpha^2 = 1$. As the sign of $\alpha$ is not of consequence, we finally arrive at

$$G(z) = \tfrac{1}{2}[A_1(z) + A_2(z)]$$

(27)

$$H(z) = \tfrac{1}{2}[A_1(z) - A_2(z)].$$

(28)

This then leads to the implementation of $G(z)$ as a parallel combination of stable all-pass functions, as desired (Fig. 2). In addition, we can simultaneously obtain the *complementary function* $H(z)$, with an extra digital adder (Fig. 4). If $G(z)$ is a low-pass function, then $H(z)$ is a high-pass function and vice versa. As $G(z)$ and $H(z)$ satisfy (13), they are said to form a *power complementary* pair. In addition, (25) says that $G$ and $H$ are complementary with respect to an all-pass function. Such a pair $[G(z), H(z)]$ is said to be *doubly complementary* [31]. The results of this section can be summarized in the following manner.

*Lemma 3.1:* Let $G(z) = P(z)/D(z)$ be a BR function of order $N$ and let $P(z)$ be symmetric, i.e., $p_k = p_{N-k}$. In

addition, let $G(z)$ be such that there exists an antisymmetric polynomial $Q(z)$ (i.e., $q_k = -q_{N-k}$) such that (14) holds. Under these conditions, $G(z)$ can be implemented as in (27) where $A_1(z)$ and $A_2(z)$ are stable all-pass functions. Furthermore, the function $H(z) = Q(z)/D(z)$ is BR, and is doubly complementary with respect to $G(z)$.

The conditions of the lemma are easily satisfied in most filtering applications. For example, let the BR function $G(z)$ be such that

1) $N$ is odd;
2) $(\partial^k |G(e^{j\omega})|/\partial\omega^k)|_{\omega=0} = 0$ for $k = 1, 2, \cdots n_0$, $n_0$ = some odd integer, i.e., $|G(e^{j\omega})|$ has odd-order tangency at zero frequency;
3) $|G(e^{j0})| = 1$; and
4) there are $(N - n_0)/2$ frequencies in the range $0 < \omega < \pi$ where $|G(e^{j\omega})| = 1$.

Under these conditions, the requirements of the lemma are satisfied. Note that classical low-pass Butterworth, Chebyshev, and elliptic digital filters [32], [35] of odd order always satisfy these requirements and can be implemented as in Fig. 2. However, optimality of the transfer function in any classical sense is not a prerequisite for obtaining the implementation of Fig. 2. This is demonstrated in Section V by means of a design example.

## IV. IMPLEMENTATION CONSIDERATIONS

The implementation of the specified transfer function $G(z)$ based on the structures of Fig. 2 can be achieved in a number of ways. The simplest approach is to realize the two all-pass filters $A_1(z)$ and $A_2(z)$ as cascades of second-order and possibly a first-order all-pass section. If these all-pass sections are realized in structurally lossless form, the overall realization is guaranteed to be structurally bounded, and hence exhibits low pass-band sensitivity with respect to each multiplier coefficient. A catalog of minimum multiplier first-order and second-order all-pass realizations has been advanced by Mitra and Hirano [19]. (Four possible first-order sections and twenty-four possible second-order sections are cataloged in [19].) Each of the realizations proposed in [19] remains all-pass independent of the actual value of the multiplier coefficients. Or in other words, the proposed all-pass sections are structurally lossless.[1] For given $A_1(z)$ and $A_2(z)$, in general, there are a large number of equivalent realizations. For example, for a 5th-order $G(z)$ with two complex pole pairs and one real pole, there are altogether 2304 ($= 24 \times 24 \times 4$) equivalent realizations, all exhibiting very low pass-band sensitivity. Other practical considerations such as product roundoff error may be used to select an optimum realization out of all such equivalent structures. We defer a discussion on roundoff error/dynamic range analysis until Section VI.

Instead of implementing the all-pass sections as cascades of second-order sections, one can also implement them in the form of lattice structures [12]–[14], of which

[1]Notice that this structural losslessness is different from pseudolosslessness [3], which is a wave digital counterpart of the well-known lossless positive real property [35].
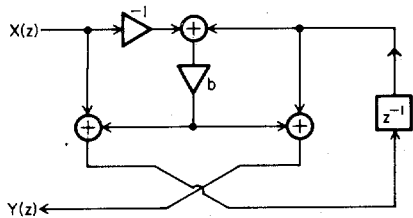
Fig. 5. First-order all-pass section, based on wave adaptor terminated in a delay.
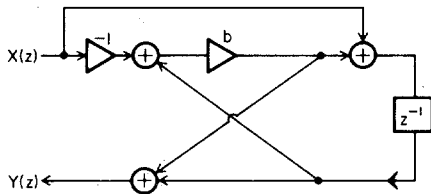


Fig. 6. First-order all-pass section, based on Gray–Markel one-multiplier lattice, terminated in a delay.

there exist several well-known versions. This could lead up to other possible advantages. (For example, the normalized cascaded lattice is automatically scaled in an $L_2$ sense, at all the internal nodes. Although it requires 4 multipliers per first-order section, the 4 multipliers can be combined into a single planar rotation element, which can in turn be implemented by a cordic-processor element. Moreover, it is possible to avoid zero-input limit cycles with such lattice structures.)

In particular, if we use the one-multiplier first-order lattice sections of Gray and Markel [12], then the resulting structures are precisely the same as the wave lattice digital filter building blocks [8], [36]. This is because a simple redrawing of the adaptors used in [36] (reproduced in Fig. 5) reveals that they are structurally equivalent to the Gray–Markel one-multiplier lattice sections [12], reproduced in Fig. 6. Unlike a four-multiplier Gray–Markel cascaded lattice, the one-multiplier cascade is not inherently scaled. Accordingly, the analysis presented by Wegener [23] should be suitably applied in order to scale down the input signal, so that internal signal overflow is avoided.

## V. Design Examples

Given a BR transfer function $G(z)$ as in (1), that satisfies the requirements of lemma 3.1, the design procedure is as follows: form the antisymmetric polynomial $Q(z)$ satisfying (14), and then compute the zeros of the polynomial $P(z) + Q(z)$. Let $z_1, z_2, \cdots, z_r$ be the zeros inside the unit circle, and let $z_{r+1}, \cdots, z_N$ be those outside. The all-pass functions $A_1(z)$ and $A_2(z)$ are now formed according to (26), and the implementation of Fig. 2 is thus accomplished.

As an example, consider the BR transfer function

where $k = 0 \cdot 13494$, so that $|G(e^{j\omega})|_{\max} = 1$. Clearly, the polynomial $P(z)$ is given by

$$P(z) = 0 \cdot 13494(1 + 1 \cdot 73306z^{-1} + 2 \cdot 83075z^{-2}$$
$$+ 2 \cdot 83075z^{-3} + 1 \cdot 73306z^{-4} + z^{-5})$$

whereas

$$D(z) = 1 - 0 \cdot 7004z^{-1} + 1 \cdot 42787z^{-2} - 0$$
$$\cdot 57995z^{-3} + 0 \cdot 40866z^{-4} - 0 \cdot 05463z^{-5}.$$

Computing $\tilde{D}(z)D(z) - \tilde{P}(z)P(z)$, we obtain its antisymmetric spectral factor $Q(z)$

$$Q(z) = 0 \cdot 26989(1 - 2 \cdot 63479z^{-1} + 4 \cdot 09366z^{-2}$$
$$- 4 \cdot 09366z^{-3} + 2 \cdot 63479z^{-4} - z^{-5}).$$

In order to find $Q(z)$, it is not necessary to find the roots of the polynomial $\tilde{D}(z)D(z) - \tilde{P}(z)P(z)$, nor is it necessary to employ general spectral factorization algorithms. In the appendix, a computationally simple procedure is outlined for finding $Q(z)$.

Next, the zeros of $P(z) + Q(z)$ are determined. These are given by

$$z_1 = 0 \cdot 155661$$

$$z_2 = 0 \cdot 109659 + j0 \cdot 924586$$

$$z_3 = z_2^*$$

$$z_4 = 0 \cdot 401930 + j1 \cdot 51943$$

$$z_5 = z_4^*.$$

Of these, $z_1$, $z_2$, and $z_3$ are inside the unit circle, whereas $z_4$ and $z_5$ are outside the unit circle. We therefore construct the two all-pass functions as

$$A_1(z) = \frac{(1 - z_4z^{-1})(1 - z_5z^{-1})}{(-z_4 + z^{-1})(-z_5 + z^{-1})}$$

$$A_2(z) = \frac{(-z_1 + z^{-1})(-z_2 + z^{-1})(-z_3 + z^{-1})}{(1 - z_1z^{-1})(1 - z_2z^{-1})(1 - z_3z^{-1})}.$$

Thus, $A_1(z)$ is a second-order section, whereas $A_2(z)$ is a cascade of a first-order section $A_{21}(z)$ and a second-order section $A_{22}(z)$ with

$$A_{21}(z) = \left(\frac{-z_1 + z^{-1}}{1 - z_1z^{-1}}\right),$$

$$A_{22}(z) = \frac{(-z_2 + z^{-1})(-z_3 + z^{-1})}{(1 - z_2z^{-1})(1 - z_3z^{-1})}.$$

The appropriate all-pass functions are, therefore,

$$A_1(z) = \frac{0 \cdot 40482 - 0 \cdot 32542z^{-1} + z^{-2}}{1 - 0 \cdot 32542z^{-1} + 0 \cdot 40482z^{-2}} \quad (30)$$

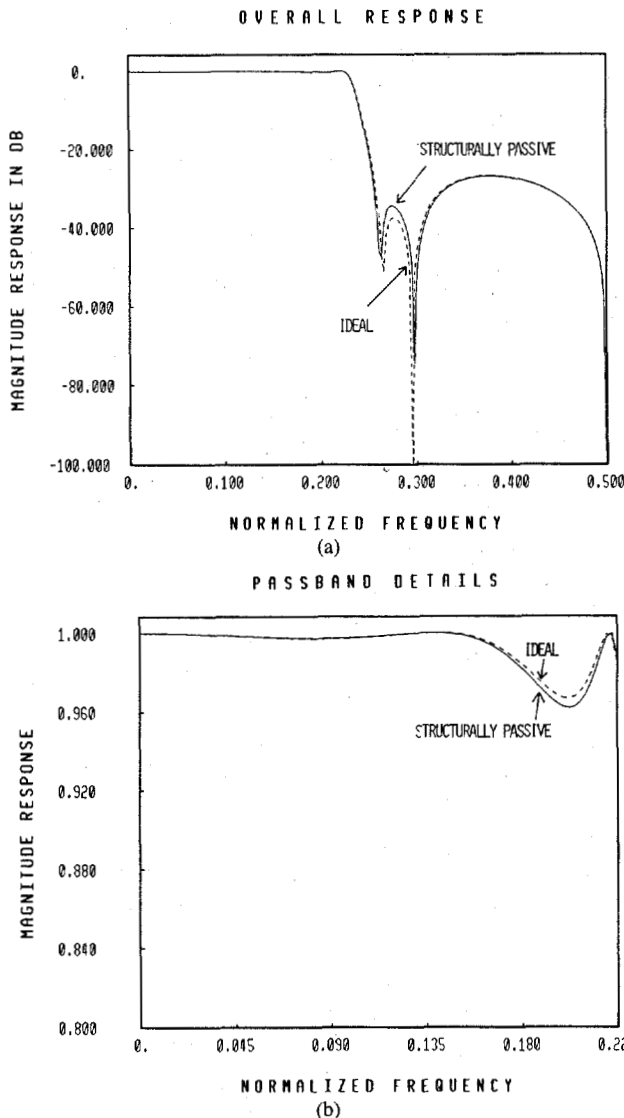$$G(z) = \frac{P(z)}{D(z)} = k\frac{1 + 1 \cdot 73306z^{-1} + 2 \cdot 83075z^{-2} + 2 \cdot 83075z^{-3} + 1 \cdot 73306z^{-4} + z^{-5}}{1 - 0 \cdot 7004z^{-1} + 1 \cdot 42787z^{-2} - 0 \cdot 57995z^{-3} + 0 \cdot 40866z^{-4} - 0 \cdot 05463z^{-5}} \quad (29)$$

Fig. 7.  (a) The new implementation with 3 bits per multiplier. (b) The new implementation with 3 bits per multiplier.



Fig. 8.  (a) The direct form with 3 bits per multiplier. (b) The direct form with 3 bits per multiplier.

and

$$A_2(z) =$$

$$\frac{-0 \cdot 13494 + 0 \cdot 90102 z^{-1} - 0 \cdot 37498 z^{-2} + z^{-3}}{1 - 0 \cdot 37498 z^{-1} + 0 \cdot 90102 z^{-2} - 0 \cdot 13494 z^{-3}}.$$

(31)

In order to study the sensitivity properties, the coefficients of the all-pass filters $A_1(z)$ and $A_2(z)$ were quantized to as low as 3 binary bits of mantissa in canonic sign digit code (SD code), and the structure of Fig. 2 simulated. Note that each multiplier has, therefore, a complexity equivalent to two addition operations. Fig. 7 shows the relevant frequency responses. In all the plots, the dashed curve indicates the ideal (infinite-precision) response. The excellent sensitivity properties of the structurally passive implementation are evident from the response plots, particularly in the passband. In the present example, $G(z)$ has been chosen to be a filter transfer function that is not op-
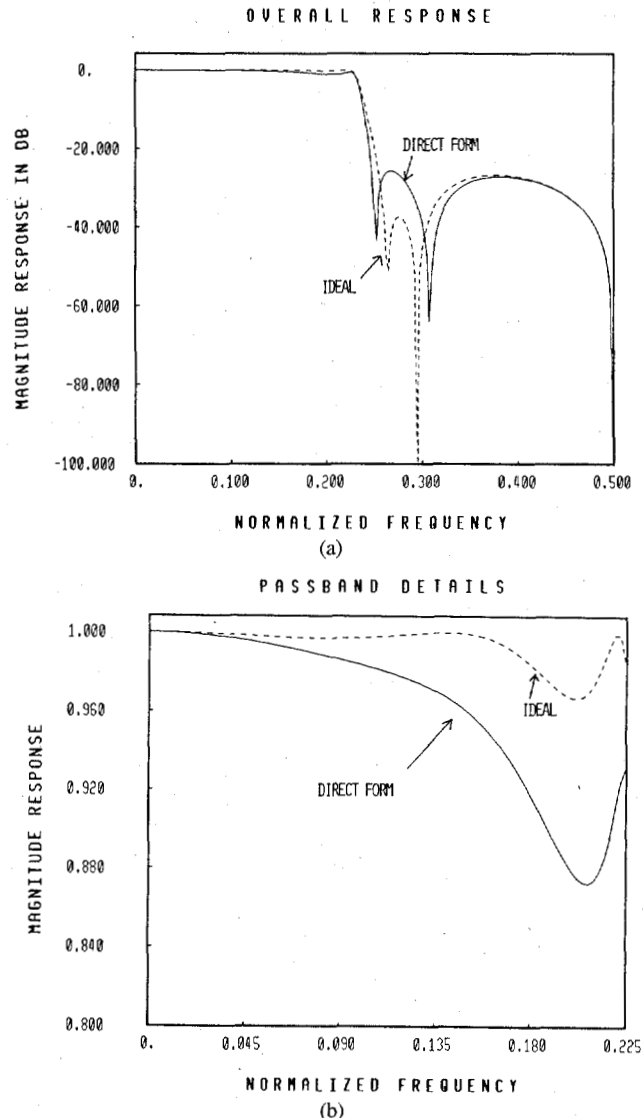
timal in any classical sense, as seen from the ideal response plots in Fig. 7. This has been purposely chosen so as to emphasize the point we made earlier that, in order to obtain an implementation as in Fig. 2, $G(z)$ need not necessarily be optimal.

As a comparison, the transfer function of (29) was also implemented in direct form, with the same amount of parameter quantization (3 bits of SD code per mantissa). Fig. 8 shows the relevant frequency response plots. Not surprisingly, the performance is unacceptable.

In the new structures, since $A_1(z)$ and $A_2(z)$ are all-pass functions, they require only 2 and 3 multipliers, respectively. Thus, a total of 5 multiplications is involved, per computed output sample. In contrast, the direct form requires 7 multipliers [even after taking into account the symmetry of the numerator $P(z)$] and, in addition, requires more precision for each multiplier.

As a further demonstration, the all-pass functions of (30) and (31) were implemented with only 2 bits of SD

code per multiplier mantissa. (Note that each multiplier is then equivalent in complexity to one addition operation.) The resulting quantized all-pass functions are given by

$$A_1(z) = \frac{0 \cdot 375 - 0 \cdot 3125z^{-1} + z^{-2}}{1 - 0 \cdot 3125z^{-1} + 0 \cdot 375z^{-2}} \qquad (32)$$

and

$$A_2(z) =$$

$$\frac{-0 \cdot 1328125 + 0 \cdot 875z^{-1} - 0 \cdot 375z^{-2} + z^{-3}}{1 - 0 \cdot 375z^{-1} + 0 \cdot 875z^{-2} - 0 \cdot 1328125z^{-3}} \cdot$$

$$\qquad (33)$$

Notice that the coefficients involved in a direct implementation of $A_1(z)$ and $A_2(z)$ are

$$0 \cdot 3125 = (0 \cdot 0101)_2$$

$$0 \cdot 375 = (0 \cdot 011)_2$$

$$0 \cdot 375 = (0 \cdot 011)_2$$

$$0 \cdot 875 = (1 \cdot 00 - 1)_2$$

$$0 \cdot 1328125 = (0 \cdot 0010001)_2.$$

Fig. 9 shows the resulting frequency responses obtained with the implementation of Fig. 2. Notice that the passband behavior continues to be excellent. In fact, the quantized response has *less* error in the passband and stopband than the ideal response! This can be explained by the fact that the quantized response has a wider transition band. Moreover, the ideal response being not optimal in any way, the above improved behavior is not surprising. As each multiplier is equivalent to one addition operation in complexity, the total complexity of the quantized circuit of Fig. 2 is now only 16 addition operations (equivalent to a single 17-bit multiplier coefficient!). With this low complexity, the structure still achieves about 30 dB stopband attenuation, and $0 \cdot 1$ dB peak pass-band ripple. This example, therefore, demonstrates the excellent potentiality of the circuit of Fig. 2 from a sensitivity viewpoint. (For completion we note that, with a quantization level of 2 bits of SD code per multiplier mantissa, the direct form structure became unstable in this example.)

*Comment on Stop-Band Sensitivity*

The pass-band sensitivity of the complementary filter $H(z)$ is expected to be excellent (for the same reason that the pass-band sensitivity of $G(z)$ is excellent). Now, in spite of parameter quantization, (27) and (28) hold, hence, (13) holds for each frequency. Thus, the stop-band sensitivity of $G(z)$ is expected to be good. However, in terms of decibels, a small pass-band error in $H(z)$ corresponds to a large stop-band error in $G(z)$, particularly in the region of low pass-band and stop-band error. Fig. 10 shows a plot of $\alpha$ versus $\beta$ where

$$\alpha = -20 \log_{10} a, \qquad \beta = -20 \log_{10} b,$$

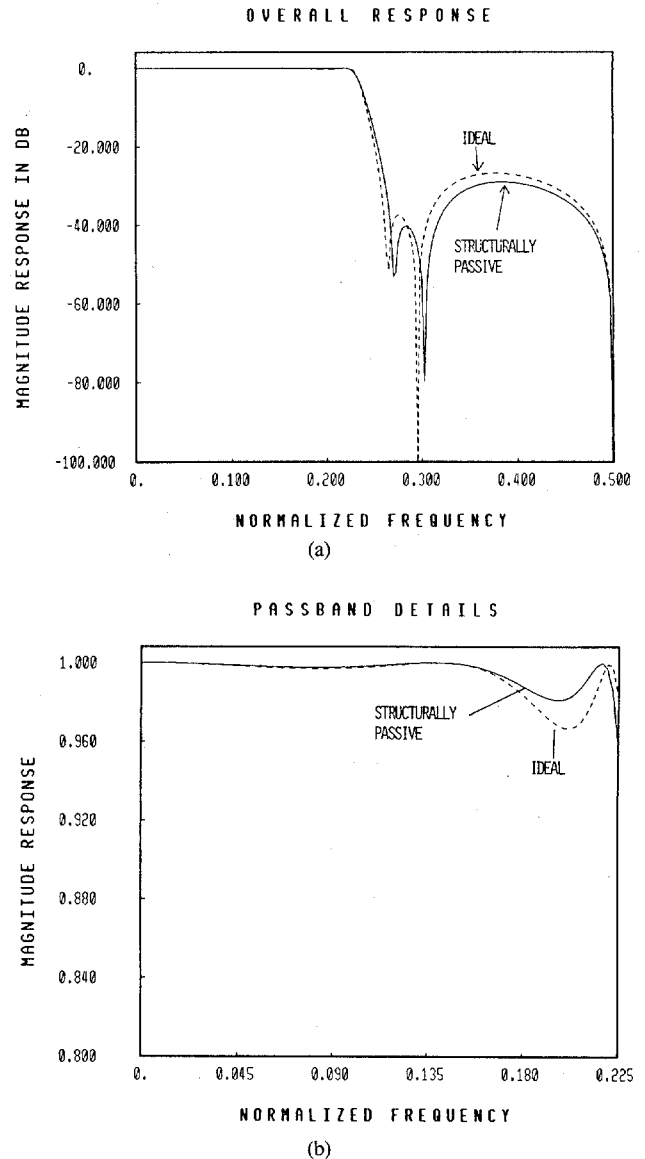$$a^2 + b^2 = 1. \qquad (34)$$



Fig. 9. (a) The new implementation, with 2 bits per multiplier. (b) The new implementation, with 2 bits per multiplier.
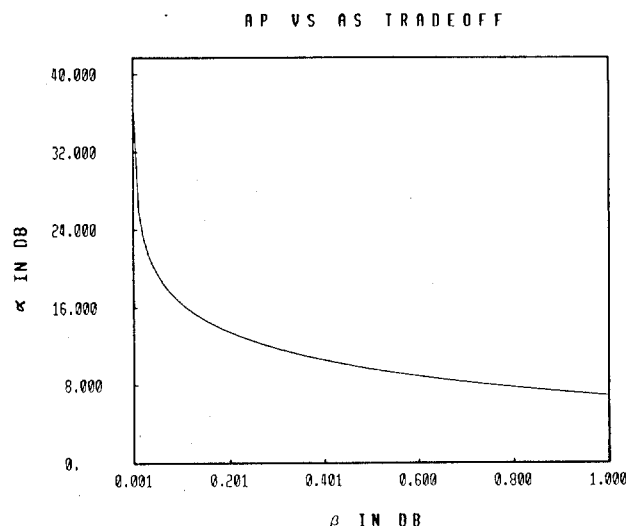


Fig. 10. The tradeoff between pass-band and stop-band errors.

Notice that the quantity $\alpha$ decreases very sharply for small changes in $\beta$, in the region of large $\alpha$. This figure demonstrates that, if $G(z)$ has large stop-band attenuation, then low *pass-band* sensitivity of $H(z)$ does not necessarily imply low *stop-band* sensitivity of $G(z)$.

## VI. ROUNDOFF NOISE AND DYNAMIC RANGE

Referring to the configuration of Fig. 2, each all-pass filter can actually be implemented as a cascade of second-order sections and possibly a first-order section leading to the structure of Fig. 11. In fact, such an implementation is, in general, better than a direct form implementation of $A_1(z)$ and $A_2(z)$. For first- or second-order all-pass functions, a number of minimum-multiplier implementations has been derived in [19]. All of these implementations are "structurally lossless" in the sense described earlier. For each of these structures, a complete fixed-point roundoff noise analysis is performed in [19]. Based on these tabulated results, it is a simple matter to calculate the roundoff noise gain for the structure of Fig. 2, under a fixed-point implementation.

In an actual fixed-point implementation, it is necessary to avoid overflow of certain internal signals in the stucture. When the implementation is based on 2's complement arithmetic, it is sufficient [32], [37] to ensure that the *inputs* to internal multipliers are within the permissible dynamic range of the implementation, which we assume to be the range $[-1, 1)$. Let us furthermore assume each internal signal to be represented by $b$ bits (plus a sign bit).

Let us now consider a cascade of all-pass sections as shown in Fig. 12. Each section satisfies $|A_{1k}(e^{j\omega})| = 1$ for all $\omega$. Let $m_{kl}$ represent the $l$th multiplier in the implementation of the $k$th all-pass function $A_{1k}(z)$. Let us now define the following transfer functions:

$T_{kl}$ = transfer function from the input of the $k$th

   all-pass section $A_{1k}(z)$ to the input of $m_{kl}$     (35)

$S_{kl}(z)$ = transfer function from the output of $m_{kl}$

   to the output of the section $A_{1k}(z)$     (36)

$P_{kl}(z)$ = transfer function from the input of the

   overall system to the input of $m_{kl}$     (37)

$Q_{kl}(z)$ = transfer function from the output of

   $m_{kl}$ to the filter output.     (38)

Clearly then,

$$P_{kl}(z) = A_{11}(z) \cdots A_{1,k-1}(z) T_{kl}(z) \qquad (39)$$

$$Q_{kl}(z) = S_{kl}(z) \cdots A_{1,k+1}(z) \cdots A_{1,r}(z). \qquad (40)$$

As a result,

$$|P_{kl}(e^{j\omega})| = |T_{kl}(e^{j\omega})| \qquad (41)$$

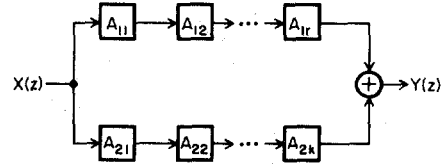$$|Q_{kl}(e^{j\omega})| = |S_{kl}(e^{j\omega})|. \qquad (42)$$



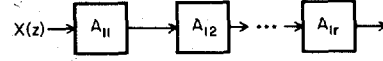Fig. 11. Implementation of the circuit of Fig. 2 in terms of low-order all-pass sections.



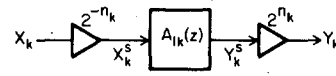Fig. 12. Pertaining to the scaling of internal signals.



Fig. 13. Insertion of scaling factors.

Assuming as usual that the input $x(n)$ is scaled in the $L_2$ sense,[2] i.e., $\|X(e^{j\omega})\|_2 = 1$, we can avoid overflow at the input of any multiplier in the $k$th all-pass section of Fig. 12 (i.e., make sure it is in the range $([-1, 1)$ for all time $n$) simply by inserting into the cascade the multipliers $2^{-n_k}$ and $2^{n_k}$ as shown in Fig. 13 where

$$2^{n_k} \geq \alpha_k \triangleq \sup_l \|T_{kl}\|_2. \qquad (43)$$

Now, the internally generated quantizer error has variance $\Delta^2/12$ where $\Delta = 2^{-b}$ and, hence, the roundoff noise variance contaminating $Y_k(z)$, due to quantization inside the $k$th all-pass section, is

$$\beta_k^2 \frac{\Delta^2}{12} \cdot 2^{2n_k} \qquad (44)$$

where

$$\beta_k^2 = \sum_l \|S_{kl}\|_2^2. \qquad (45)$$

Thus, the obvious effect of avoiding internal overflow is an increased roundoff noise, which is the result of the known roundoff noise/dynamic range interaction [37]. Basically, the introduction of scale factors as in Fig. 13 ensures that the overall signal gain of the all-pass cascade is the same as that of the unscaled structure. However, scaling does result in a decreased signal-to-noise ratio because of increased noise. The overall noise variance at the output of the cascade of Fig. 12 is

$$\sigma_1^2 = \frac{\Delta^2}{12} \sum_{k=1}^{r} 2^{2n_k} \cdot \beta_k^2. \qquad (46)$$

Thus, the structure of Fig. 11 has total roundoff noise variance (under scaled conditions) $\sigma_1^2 + \sigma_2^2$, where $\sigma_1^2$ is as in (46) and $\sigma_2^2$ is the corresponding contribution from the all-pass chain $A_{21} A_{22} \cdots A_{2k}$. It is clear that each all-pass section should be independently chosen so as to have the smallest value of the product $\alpha_k^2 \beta_k^2$ or, in practice, the

---

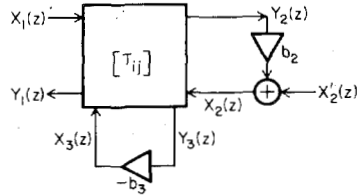[2]The symbol $\|\cdot\|_2$ stands for the usual $L_2$ norm [37].

Fig. 14. Multiplier-extraction in second-order sections.

smallest value of

$$p_k = 2^{2n_k} \beta_k^2. \tag{47}$$

Moreover,

$$F \triangleq 1/(\sigma_1^2 + \sigma_2^2)$$

represents a true figure of merit, under scaled conditions. Notice that the ordering of the sections in the all-pass cascade does not affect the roundoff noise/dynamic range performance.

Based on a digital multiplier-extraction approach, four distinct first-order all-pass sections and twenty-four distinct second-order all-pass sections are cataloged in [19]. Each of these sections has the property that the input–output relation is an all-pass function, in spite of parameter quantization. Accordingly, the entire cascade of Fig. 12 remains all-pass in spite of multiplier quantization, ensuring that the overall transfer function in Fig. 11 is structurally bounded, leading to low pass-band sensitivity. Next, for each all-pass section reported in [19], a roundoff noise analysis is also included, and the values of $\beta_k^2$ are tabulated [19, Tables I and II]. It therefore only remains to compute the scaling parameters $\alpha_k^2$ of (43). This computation can be done with little additional effort, simply by making judicious use of the entries in Tables I and II of [19].

In order to explain how this can be done, consider a digital filter structure with two-multipliers $b_2$ and $-b_3$, drawn in the form of a three-pair as shown in Fig. 14. Let the 3-input 3-output multiplierless system in Fig. 14 be described by the transfer matrix $[T_{ij}]$, i.e., let

$$\begin{bmatrix} Y_1(z) \\ Y_2(z) \\ Y_3(z) \end{bmatrix} = \begin{bmatrix} T_{11}(z) & T_{12}(z) & T_{13}(z) \\ T_{21}(z) & T_{22}(z) & T_{23}(z) \\ T_{31}(z) & T_{32}(z) & T_{33}(z) \end{bmatrix} \begin{bmatrix} X_1(z) \\ X_2(z) \\ X_3(z) \end{bmatrix}. \tag{48}$$

It is easily verified that

multiplier $b_2$. Inspection of (49) reveals that if each $T_{ij}(z)$ is replaced with $T_{ji}(z)$, then it leads to (50). In other words, the scaling transfer functions are precisely equal to the corresponding noise transfer functions of the transposed structures. The noise gains $\beta_k^2$ are tabulated in [19] for each structure and its transpose. Accordingly, the scaling norms $\alpha_k^2$ can be found by inspection of Table II in [19]. Similar discussions hold for first-order all-pass sections.

To take a specific example, consider type-$2A$ all-pass sections. Then, from the entries of Table II in [19],

$$\beta_k^2 = (1 + b_2^2)R_1 \tag{51}$$

whereas

$$\alpha_k^2 = \max(R_5, R_6). \tag{52}$$

Similarly, for a type-$3A_t$ section,

$$\beta_k^2 = \rho R_4 + R_5 \tag{53}$$

whereas

$$\alpha_k^2 = \max(R_1, R_1) = R_1. \tag{54}$$

(Notice that each entry in Table II of [19] is a sum of two terms; one corresponding to each one of the two multipliers.)

In Table I of this paper, we have tabulated for convenience the quantities $\beta_k^2$ and $\alpha_k^2$ for all the relevant second-order all-pass sections. From this table, it is easy to find the all-pass section with smallest $p_k$. If the designer has the additional restriction that the number of delays should be minimal, then the choice of second-order all-pass sections is restricted to types $2A$, $2A_t$, $3A$, and $3A_t$.

*Comments on First-Order All-Pass Sections*

It is readily verified that the four first-order sections reported in [19] have the values of $\alpha_k^2$ and $\beta_k^2$ as shown in Table II. The product $\alpha_k^2\beta_k^2$ is the same for types $1A$ and $1A_t$. Accordingly, types $1A$ and $1A_t$ are equivalent (and so are types $1B$ and $1B_t$) as far as noise/dynamic range properties are concerned. Types $A$ are better than types $B$ by a factor of 2, which corresponds only to 3 dB (i.e., one-half a binary bit). Accordingly, type $B$ should always be preferred in view of canonicity in delays.

*Numerical Example*

For the example of Section V, the overall transfer function $G(z)$ (6) requires a second-order section $A_1(z)$, and a

$$\frac{Y_1(z)}{X_2'(z)}\bigg|_{X_1(z)=0} = \frac{T_{12}(z) + b_3[T_{12}(z)T_{33}(z) - T_{13}(z)T_{32}(z)]}{1 - b_2 T_{22}(z) + b_3 T_{33}(z) - b_2 b_3[T_{22}(z)T_{33}(z) - T_{23}(z)T_{32}(z)]} \tag{49}$$

whereas

$$\frac{Y_2(z)}{X_1(z)}\bigg|_{X_2'=0} = \frac{T_{21}(z) + b_3[T_{21}(z)T_{33}(z) - T_{31}(z)T_{23}(z)]}{1 - b_2 T_{22}(z) + b_3 T_{33}(z) - b_2 b_3[T_{22}(z)T_{33}(z) - T_{23}(z)T_{32}(z)]}. \tag{50}$$

Equation (49) represents the noise transfer function from the output of $b_2$ to the output of the all-pass section, whereas (50) represents the scaling transfer function for

third-order section $A_2(z)$, where $A_2(z) = A_{21}(z) \cdot A_{22}(z)$. The first-order section $A_{21}(z)$ is best implemented with type $1B$ (or $1B_t$) which is canonic in delays. For this sec-

TABLE I
CATALOG OF SECOND-ORDER ALL-PASS SECTIONS
THE QUANTITIES $R_k$, $\rho$, $b_2$ HAVE SIGNIFICANCE AS DEFINED IN [19]

| Serial Number | Type of Structure | $\beta_k^2$ Noise Gain | $\alpha_k^2$ (Scaling Information) |
|---|---|---|---|
| 1 | 2A | $(1+b_2^2)R_1$ | $\max(R_5, R_6)$ |
| 2 | 2B | $(1+b_2^2)R_2$ | $\max(R_7, R_8)$ |
| 3 | 2C | $(1+b_2^2)R_3$ | $\max(R_9, R_{10})$ |
| 4 | 2D | $(1+b_2^2)R_4$ | $\max(R_{11}, R_{12})$ |
| 5 | $2A_t$ | $R_5 + R_6$ | $\max(R_1, b_2^2 R_1)$ |
| 6 | $2B_t$ | $R_7 + R_8$ | $\max(R_2, b_2^2 R_2)$ |
| 7 | $2C_t$ | $R_9 + R_{10}$ | $\max(R_3, b_2^2 R_3)$ |
| 8 | $2D_t$ | $R_{11} + R_{12}$ | $\max(R_4, b_2^2 R_4)$ |
| 9 | 3A(3E) | $R_1 + R_1$ | $R_5$ |
| 10 | 3B(3F) | $R_2 + R_2$ | $R_7$ |
| 11 | 3C(3G) | $R_3 + R_3$ | $R_9$ |
| 12 | 3D(3H) | $R_4 + R_4$ | $R_{11}$ |
| 13 | $3D_t(3H_t)$ | $\rho R_1 + R_{11}$ | $R_4$ |
| 14 | $3C_t(3G_t)$ | $\rho R_2 + R_9$ | $R_3$ |
| 15 | $3B_t(3F_t)$ | $\rho R_3 + R_7$ | $R_2$ |
| 16 | $3A_t(3E_t)$ | $\rho R_4 + R_5$ | $R_1$ |

Note: Types $2A$, $2A_t$, $3A$, $3A_t$ are canonic in delays.

TABLE II
CATALOG OF FIRST-ORDER ALL-PASS SECTIONS

| Type | 1A | 1B | $1A_t$ | $1B_t$ |
|---|---|---|---|---|
| $\alpha_k^2$ | $\frac{1}{1-b_1^2}$ | $\frac{2}{1+b_1}$ | $2$ | $\frac{2}{1-b_1}$ |
| $\beta_k^2$ | $2$ | $\frac{2}{1-b_1}$ | $\frac{1}{1-b_1^2}$ | $\frac{2}{1+b_1}$ |
| $\alpha_k^2 \beta_k^2$ | $\frac{2}{1-b_1^2}$ | $\frac{4}{1-b_1^2}$ | $\frac{2}{1-b_1^2}$ | $\frac{4}{1-b_1^2}$ |
| # of delays | 2 | 1 | 2 | 1 |

tion, we have

$$\beta_k^2 = 2 \cdot 369$$

$$\alpha_k^2 = 1 \cdot 731$$

$$\beta_k^2 \alpha_k^2 = 4 \cdot 10. \tag{55}$$

Sections $A_1(z)$ and $A_{22}(z)$ have several choices as tabulated in Table III. For both $A_1(z)$ and $A_{22}(z)$, the product $\alpha_k^2 \beta_k^2$ turns out to be smallest for type-$3D_t$ sections. With these

TABLE III
EXHAUSTIVE SEARCH FOR BEST SECOND-ORDER SECTION IN NUMERICAL EXAMPLE

| Serial Number | Type of Structure | $A_{22}(z)$ | | | | $A_1(z)$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $\beta_k^2$ | $\alpha_k^2$ | $\beta_k^2 \alpha_k^2$ | $2^{n_k} \geq \alpha_k$ | $\beta_k^2$ | $\alpha_k^2$ | $\beta_k^2 \alpha_k^2$ | $2^{n_k} \geq \alpha_k$ |
| 1 | 2A | 15·75 | 37·15 | 585 | 8 | 3·72 | 9·55 | 35·53 | 4 |
| 2 | 2B | 9·56 | 34·38 | 329 | 8 | 3·44 | 6·56 | 22·57 | 4 |
| 3 | 2C | 7·55 | 32·66 | 247 | 8 | 2·15 | 3·16 | 6·79 | 2 |
| 4 | 2D | 4·26 | 31·51 | 135 | 8 | 1·40 | 4·29 | 6·00 | 2 |
| 5 | $2A_t$ | 38·22 | 15·02 | 574 | 4 | 11·0 | 3·36 | 36·9 | 2 |
| 6 | $2B_t$ | 36·51 | 9·12 | 333 | 4 | 9·25 | 3·11 | 28·8 | 2 |
| 7 | $2C_t$ | 34·82 | 7·20 | 251 | 4 | 6·26 | 1·94 | 12·1 | 2 |
| 8 | $2D_t$ | 33·51 | 4·08 | 137 | 2 | 6·28 | 1·26 | 7·91 | 2 |
| 9 | 3A(3E) | 30·05 | 1·07 | 32·2 | 1 | 6·72 | 1·45 | 9·75 | 2 |
| 10 | 3B(3F) | 18·24 | 2·13 | 38·9 | 2 | 6·23 | 2·69 | 16·76 | 2 |
| 11 | 3C(3G) | 14·40 | 2·16 | 31·1 | 2 | 3·89 | 3·10 | 12·06 | 2 |
| 12 | 3D(3H) | 8·16 | 2·00 | 16·3 | 2 | 2·53 | 2·00 | 5·06 | 2 |
| 13 | $3D_t(3H_t)$ | 2·27 | 4·08 | 9·26 | 2 | 3·19 | 1·26 | 4·02 | 2 |
| 14 | $3C_t(3G_t)$ | 2·32 | 7·20 | 16·7 | 4 | 4·21 | 1·94 | 8·17 | 2 |
| 15 | $3B_t(3F_t)$ | 2·25 | 9·12 | 20·5 | 4 | 3·38 | 3·11 | 10·51 | 2 |
| 16 | $3A_t(3E_t)$ | 1·15 | 15·03 | 17·3 | 4 | 1·89 | 3·36 | 6·35 | 2 |

choices for all-pass sections, the total (scaled) roundoff noise gain is

$$4 \cdot 10 + 9 \cdot 26 + 4 \cdot 02 = 17 \cdot 38 \tag{56}$$

corresponding to 12 dB, i.e., about 2 bits of signal deterioration.

Next, if we prefer to use second-order sections that are canonic in delays, then the best possible choice is type $3A_t$ for both $A_1(z)$ and $A_{22}(z)$. The resulting total scaled output roundoff noise gain is now

$$4 \cdot 10 + 17 \cdot 3 + 6 \cdot 35 = 27 \cdot 75 \tag{57}$$

which corresponds to $14 \cdot 4$ dB, i.e., between 2–3 bits of signal deterioration.

In Table III, we have also indicated the appropriate $2^{n_k}$-approximation for each $\alpha_k$. These quantities give the important information necessary in Fig. 13 to avoid internal signal overflow. Notice that the use of $2^{n_k}$ instead of $\alpha_k$ leads to a change in the actual scaled noise gain. For example, (57) above becomes

$$2 \cdot 369 \times 4 + 1 \cdot 15 \times 16 + 1 \cdot 89 \times 4 = 35 \cdot 436$$

which is equivalent to $15 \cdot 5$ dB noise gain (i.e., about 3 bits).

Type-$1B$ and type-$3A_t$ sections, which are most appropriate for this numerical example, are shown in Fig. 15.

## VII. CONCLUDING REMARKS

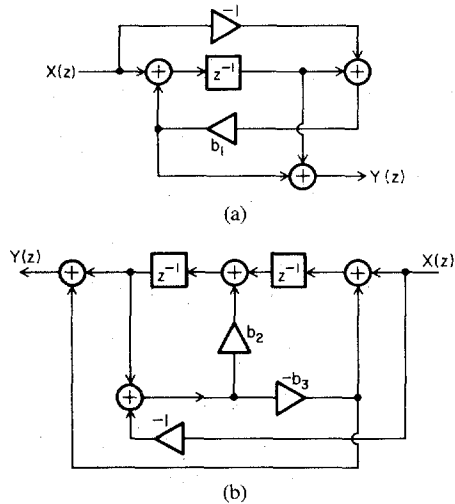The main emphasis in this paper has been an independent $z$-domain approach for the synthesis of low-pass-

Fig. 15. The first- and second-order all-pass building blocks for the numerical example. (a) Type-$1B$ section. (b) Type-$3A_t$ section.

band-sensitivity digital filter structures. The resulting structures are in the form of a parallel connection of two all-pass sections, and each section can be implemented in a variety of ways. One of these is the simple cascade of first- and second-order sections, each section being chosen from a catalog of structurally lossless circuits, in order to optimize the roundoff-noise/dynamic range performance. As an alternative, the all-pass functions can also be synthesized using the Gray and Markel approach [12]–[14], or in the form of a wave digital cascade [8]. The latter approach leads to the wave digital lattice filters, which have been extensively studied [23], [27] and are well known for their excellent properties under quantized environment.

## APPENDIX

The computation of the spectral factor $Q(z)$ is made particularly simple because of the antisymmetric nature of $Q(z)$. From (16) it is clear that

$$Q^2(z) = P^2(z) - z^{-N}D(z^{-1})D(z). \qquad (A.1)$$

The polynomial on the right-hand side of (A.1) is known since $P(z)$ and $D(z)$ are known. Indicate this polynomial by

$$R(z) = \sum_{n=0}^{2N} r_n z^{-n}. \qquad (A.2)$$

Then $Q^2(z) = R(z)$ and, hence, the coefficients $q_n$ are related to $r_n$ as

$$r_n = \sum_{k=0}^{n} q_k q_{n-k}. \qquad (A.3)$$

Thus, $q_n$ can be computed recursively as follows:

$$q_0 = \sqrt{r_0}, \quad q_1 = r_1/2q_0 \qquad (A.4)$$

$$q_n = \frac{r_n - \sum_{k=1}^{n-1} q_k q_{n-k}}{2q_0}, \quad 2 \le n \le N. \qquad (A.5)$$

In fact, because of the antisymmetry of $q_n$, it is necessary only to compute one-half the number of coefficients in (A.5).

## ACKNOWLEDGMENT

## REFERENCES

[1] J. F. Kaiser, in *System Analysis by Digital Computer*, F. F. Kuo and J. F. Kaiser, Eds. New York: Wiley, 1966.

[2] A. Fettweis, "Digital filter structures related to classical filter networks," *Arch. Elek. Ubertrangungstechnik*, vol. 25, pp. 79–81, Feb. 1971.

[3] —, "Pseudopassivity, sensitivity, and stability of wave digital filter," *IEEE Trans. Circuit Theory*, vol. CT-19, pp. 668–673, Nov. 1972.

[4] A. Fettweis and K. Meerkotter, "Suppression of parasitic oscillations in wave filters," *IEEE Trans. Circuits Syst.*, vol. CAS-22, pp. 239–246, Mar. 1975.

[5] M. N. S. Swamy and K. Thyagarajan, "A new type of wave digital filters," *J. Franklin Inst.*, vol. 300, no. 1, pp. 41–58, July 1975.

[6] R. E. Crochiere, "Digital ladder structures and coefficient sensitivity," *IEEE Trans. Audio Electroacoust.*, vol. AU-20, pp. 240–246, Oct. 1972.

[7] A. Fettweis, "On sensitivity and roundoff noise in wave digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-22, pp. 383–384, Oct. 1974.

[8] —, "Wave digital lattice filters," *Int. J. Circuit Theory Appl.*, vol. 2, pp. 203–211, June 1974.

[9] A. G. Constantinides, "Alternative approach to design of wave digital filters," *Electron. Lett.*, vol. 10, pp. 59–60, 1974.

[10] S. S. Lawson and A. G. Constantinides, "The design of digital filters using linear transformations on classical $LC$ two-ports," in *Proc. Network Theory Symp.*, 1975, pp. 369–378.

[11] D. A. Vaughan-Pope and L. T. Bruton, "Transfer function synthesis using generalized doubly terminated two-pair networks," *IEEE Trans. Circuits Syst.*, vol. CAS-24, pp. 79–88, Feb. 1977.

[12] A. H. Gray and J. D. Markel, "Digital lattice and ladder filter synthesis," *IEEE Trans. Audio Electroacoust.*, vol. AU-21, pp. 491–500, Dec. 1973.

[13] —, "A normalized digital filter structure," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-23, pp. 268–277, June 1975.

[14] A. H. Gray, Jr., "Passive cascaded lattice digital filters," *IEEE Trans. Circuits Syst.*, vol. CAS-27, pp. 337–344, May 1980.

[15] R. C. Agarwal and C. S. Burrus, "New recursive digital filter structures having very low sensitivity and roundoff noise," *IEEE Trans. Circuits Syst.*, vol. CAS-22, pp. 921–926, Dec. 1975.

[16] J. Szczupak and S. K. Mitra, "On digital filter structures with low coefficient sensitivities," *Proc. IEEE*, vol. 66, pp. 1082–1083, Sept. 1978.

[17] Y. C. Lim and S. R. Parker, "FIR filter design over a discrete powers-of-two coefficient space," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 583–591, June 1983.

[18] S. K. Mitra, P. S. Kamat, and D. C. Huey, "Cascaded lattice realization of digital filters," *Int. J. Circuit Theory Appl.*, vol. 5, pp. 3–11, 1977.

[19] S. K. Mitra and K. Hirano, "Digital all-pass networks," *IEEE Trans. Circuits Syst.*, vol. CAS-21, pp. 688–700, Sept. 1974.

[20] P. P Vaidyanathan and S. K. Mitra, "Low pass-band sensitivity digital filters: A generalized viewpoint and synthesis procedures," *Proc. IEEE*, pp. 404–423, Apr. 1984.

[21] —, "Passivity properties of low sensitivity digital filter structures," *IEEE Trans. Circuits Syst.*, vol. CAS-32, pp. 217–224, Mar. 1985.

[22] —, "Very low-sensitivity FIR filter implementation using 'structural passivity' concept," *IEEE Trans. Circuits Syst.*, vol. CAS-32, pp. 360–364, Apr. 1985.

[23] W. Wegener, "On the design of wave digital lattice filters with short coefficient wordlength and optimal dynamic range," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 1091–1098, Dec. 1978.

[24] P. DeWilde and E. Deprettere, "Orthogonal cascade realization of

real multiport digital filters," *Int. J. Circuit Theory Appl.*, vol. 8, pp. 245-277, 1980.

[25] D. Henrot and C. T. Mullis, "A modular and orthogonal digital filter structure for parallel processing," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Apr. 1983, pp. 623-626.

[26] S. K. Rao and T. Kailath, "Orthogonal digital filters for VLSI implementation," *IEEE Trans. Circuits Syst.*, vol. CAS-31, pp. 933-945, Nov. 1984.

[27] L. Gazsi, "Explicit formulas for lattice wave digital filters," *IEEE Trans. Circuits Syst.*, vol. CAS-32, pp. 68-88, Jan. 1985.

[28] A. G. Constantinides and R. A. Valenzuela, "A class of efficient interpolators and decimators with applications in transmultiplexers," in *Proc. IEEE Int. Symp. Circuits Syst.*, Rome, Italy, May 1982, pp. 260-263.

[29] —, "An efficient and modular transmultiplexer design," *IEEE Trans. Commun.*, vol. COM-30, pp. 1629-1641, July 1982.

[30] R. Ansari and B. Liu, "A class of low-noise computationally efficient recursive digital filters with applications to sampling rate alterations," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 90-97, Feb. 1985.

[31] Y. Neuvo and S. K. Mitra, "Complementary IRR digital filters," in *Proc. IEEE Int. Symp. Circuits Syst.*, Montreal, P. Q., Canada, May 1984, pp. 234-237.

[32] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.

[33] H. J. Orchard, "Inductorless filters," *Electron. Lett.*, pp. 224-225, Sept. 1966.

[34] A. Fettweis, J. A. Nossek, and K. Meerkotter, "Reconstruction of signals after filtering and sampling-rate reduction," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, San Diego, CA, Mar. 1984, pp. 11.7.1-11.7.4.

[35] G. C. Temes and J. W. LaPatra, *Circuit Synthesis and Design*. New York: McGraw-Hill, 1977.

[36] J. K. J. Van Ginderdeuren, H. J. DeMan, N. F. Gonclaves, and W. A. M. Van Noije, "Compact NMOS building blocks and a methodology for dedicated digital filter applications," *IEEE Trans. Solid-State Circuits*, vol. SC-18, pp. 306-316, June 1983.

[37] L. B. Jackson, "On the interaction of roundoff noise and dynamic range in digital filters," *Bell Syst. Tech. J.*, vol. 49, pp. 159-184, 1970.
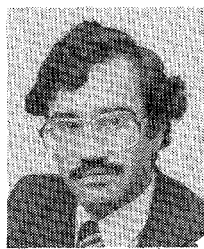
**P. P. Vaidyanathan** (S'80-M'83) was born in Calcutta, India, on October 16, 1954. He received the B.Sc. (Hons.) degree in physics, and the B.Tech and M.Tech degrees in radiophysics and electronics from the University of Calcutta, India, in 1974, 1977, and 1979, respectively, and the Ph.D. degree in electrical and computer engineering from the University of California, Santa Barbara, in 1982.

He was a Postdoctoral Fellow at the University of California, Santa Barbara, from September 1982 to February 1983. Since March 1983 he has been with the California Institute of Technology, Pasadena, as an Assistant Professor of Electrical Engineering. His main research interests are in digital signal processing, linear systems, and filter design.

Dr. Vaidyanathan served as the Vice Chairman of theTechnical Program Committee for the 1983 IEEE International Symposium on Circuits and Systems. He currently serves as an Associate Editor for the IEEE TRANS-ACTIONS ON CIRCUITS AND SYSTEMS. He was the recipient of the Award for Excellence in Teaching at the California Institute of Technology for 1983-1984.

**Sanjit K. Mitra** (S'59-M'63-SM'63-SM'69-F'74) received the B.S. (Hons.) degree in physics from the Utkal University, India, in 1953; the M.Sc. (Tech.) degree in radio physics and electronics from the University of Calcutta, India, in 1956; the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1960 and 1962, respectively.

He was a member of the Faculty of Cornell University, Ithaca, NY, from 1962 to 1965, and a member of the Technical Staff of Bell Laboratories from 1965 to 1967. He joined the Faculty of the University of California, Davis, in 1967 and transferred to the Santa Barbara (UCSB) campus in 1977 as a Professor of Electrical and Computer Engineering. From July 1979 to June 1982 he served as the Chairman of the Department of Electrical and Computer Engineering at UCSB. He is a Consultant to the Lawrence Livermore National Laboratory, Livermore, CA, and the Los Alamos National Laboratory, Los Alamos, NM, and is the Consulting Editor for the *Electrical/Computer Science and Engineering Series* (New York: Van Nostrand Reinhold). He has also served as a Consultant to Ampex Corporation, Fairchild Semiconductor, Siliconix, HDR Sciences, the U.S. Army, General Motors Delco Division, and the World Bank. He has held visiting appointments at the Indian Institute of Technology, New Delhi, India; Kobe University, Japan; University of Erlangen-Nuernberg, West Germany; the Australian National University, Canberra, Australia; Tampere University of Technology, Tampere, Finland; and the Federal University of Rio de Janeiro, Brazil. He has published a number of papers in active and passive networks, and digital filters, and is the author of *Analysis and Synthesis of Linear Active Networks* and *An Introduction to Digital and Analog Integrated Circuits and Applications*, Editor of *Active Inductorless Filters*, Co-editor of *Modern Filter Theory and Design* (with G. C. Temes), Co-editor of *Two-Dimensional Digital Signal Processing* (with M. P. Ekstrom). He holds two patents in active filters.

Dr. Mitra was an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS, and served on the Editorial Boards of the PROCEEDINGS OF THE IEEE and IEEE Press, and presently is on the Editorial Board of *Circuits, Systems and Signal Processing*. He has been a member of the Administrative Committee of the IEEE Circuits and Systems Society, Local Arrangements Chairman (1969), General Chairman (1974), and Program Chairman (1983) of the Annual IEEE International Symposium on Circuits and Systems and is presently the President-Elect of the IEEE Circuits and Systems Society. He is the recipient of the 1973 F. E. Terman Award and 1985 AT&T Foundation Award of the American Society of Engineering Education, a Visiting Professorship from the Japan Society for Promotion of Science in 1972, and Senior Fulbright Award for Brazil in 1984. He is a Fellow of the American Association for the Advancement of Science and a member of the American Society for Engineering Education, Sigma Xi, and Eta Kappa Nu.
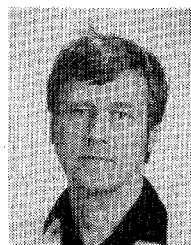
**Yrjö Neuvo** (S'70-M'74-SM'82) was born in Turku, Finland, on July 21, 1943. He received the Diploma Engineer and Licentiate of Technology degrees from the Helsinki University of Technology, Helsinki, Finland, in 1968 and 1971, respectively, and the Ph.D. degree in electrical engineering from Cornell University, Ithaca, NY, in 1974.

He held various research and teaching positions at the Helsinki University of Technology, the Academy of Finland, and Cornell University, from 1968 to 1976. Since 1976 he has been a Professor of Electrical Engineering at the Tampere University of Technology, Tampere, Finland. During the academic year 1981-1982 he was with the University of California, Santa Barbara, as a Visiting Professor. His main research interests are in the areas of digital filters and microcomputer systems.

Dr. Neuvo is a member of Phi Kappa Phi and the Finnish Academy of Technical Sciences.