Digital Filtering by Polyphase Network: Application to Sample-Rate Alteration and Filter Banks

MAURICE G. BELLANGER, MEMBER, IEEE, GEORGES BONNEROT, AND MICHEL COUDREUSE

Abstract—**The digital filtering process can be achieved by a set of phase shifters with suitable characteristics. A particular set, named polyphase network, is defined and analyzed. It permits the use of recursive devices for efficient sample-rate alteration. The comparison with conventional filters shows that, with the same active memory, a reduction of computation rate approaching a factor of 2 can be achieved when the alteration factor increases. A more substantial gain can be obtained in the direct realization of a uniform bank of recursive filters through combination of the polyphase network with a discrete Fourier transform (DFT) computer; savings in hardware also result from the low sensitivity of the structure to coefficient word lengths.**

I. Introduction

The optimization of digital processing requires the minimization of the computation rate and, consequently, of the signal sampling frequency. This implies that at every stage of the processing, the sampling rate should be kept to a value not exceeding twice the highest frequency component contained in the baseband signal. Conventional digital filter structures do not in general comply with this basic rule. For example, after low-pass filtering, some of the input signal components have been removed and, if input and output filter sampling frequencies are identical, there are some redundances in the digital processing. However, there is a device which does not exhibit this kind of redundance in the same conditions—the phase shifter. As it does not cancel any signal component, if the sampling frequency is minimal at the input, it still is minimal at the output. This suggests a way to achieve efficient sample-rate alteration and, consequently, optimize digital processing by recursive devices which basically require identical input and output sampling frequencies.

The purpose of this paper is to introduce, for digital filtering, a structure based on phase shifters and referred to as polyphase network, which is amenable to efficient sample-rate alteration and can use recursive devices unlike techniques presented in [1] and [2]. It turns out, however, that the polyphase network structure has its highest efficiency when combined with a discrete Fourier transform (DFT) computer to realize a bank of filters. It then permits a direct implementation of a bank of recursive filters without any approximation, thus considerably generalizing the window fast Fourier transform (FFT) technique for spectrum analysis [3, pp. 389-390]. After definition a simple and general design technique will be presented. It relies on a decomposition of the Z transfer function of the filter to be realized, which has been introduced in [4] and analyzed with more detail in [5]. Some specific properties of the structure will also be discussed before applications are considered.

Manuscript received March 23, 1974; revised May 21, 1975.
The authors are with Télécommunications Radioélectriques et Teléphoniques, Plessis Robinson, France.

II. Digital Filtering by Polyphase Network

It is common knowledge that the filtering process can be obtained by cancellation of the undesired signal components at the output of a summation device connected to a set of phase shifters. The phase shifters are all-pass and rotate the signal component phases in such a way that the components to be kept have the same phase at the input of the summation device and, therefore, add in amplitude, while the components to be removed have different phases and vanish after summation. Several sets of phase shifters having various characteristics can be imagined to achieve this filtering process. The present work concentrates on a particular class, referred to as polyphase networks, which will be shown to have attractive properties with respect to computation rate and implementation when implemented digitally.

Let a polyphase network of order N, N integer, be defined as a set of N all-pass devices ϕn $(n = 0, 1, \cdots, N-1)$ having the following phase characteristics:

$$\varphi n(f) = -\frac{2\pi}{N} \cdot n \cdot \left[\frac{f}{fr} + \frac{1}{2}\right]$$

where fr is a reference frequency and $[x]$ denotes the greatest integer in x.

These characteristics are shown in Fig. 1. To show that such a set of $\phi n (n = 0, 1, \cdots, N-1)$, followed by a summation device, is capable of performing the filtering process, let us apply a signal simultaneously to all the ϕn. The components of the N output signals, which are in the frequency band $(0, fr/2)$, all have the same phase and, consequently, add in amplitude in the summation device. In the frequency band $(fr/2, 3fr/2)$, the components of the signal coming from $\phi 1$ have been rotated of $-2\pi/N$ with respect to those of the signal coming from ϕo; in the same band the components of the signal coming from ϕn have been rotated of $-n \cdot 2\pi/N$. Thus the addition of the signals coming from all the ϕn, $(n = 0, 1, \cdots, N-1)$ involves the cancellation of the components in the band $(fr/2, 3fr/2)$.

The same applies to signal components in the frequency bands $(3fr/2, 5fr/2)$, $(5fr/2, 7fr/2)$, etc., and finally only the signal components contained in the frequency bands $[i \cdot (N - \frac{1}{2}) \cdot fr, i \cdot (N + \frac{1}{2}) \cdot f_r]$ with i integer are kept after summation. Consequently, the set of ϕn $(n = 0, 1, \cdots, N-1)$ followed by a summation device has a frequency response which is that of a low-pass filter with cutoff frequency $fr/2$ and periodicity $N \cdot fr$ on the frequency axis.

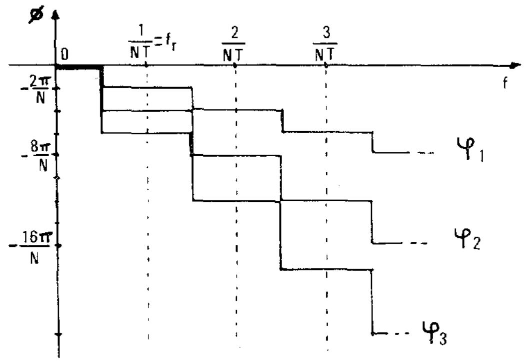


Fig. 1. Phase characteristics in a polyphase network.

This might seem a rather complicated way of performing filtering, but the reason for concentrating on such a specific structure is that it lends itself to digital processing at low rate, as will be shown later on. In fact, each function $\varphi n(f)$ can be decomposed into a sum of two terms:

$$\varphi n(f) = -\frac{2\pi}{N} \cdot n \cdot \frac{f}{fr} + \frac{2\pi}{N} \cdot n \cdot \left(\frac{f}{fr} - \left[\frac{f}{fr} + \frac{1}{2} \right] \right).$$

The first term characterizes the delay $n/N \cdot fr$. The second has the periodicity fr on the frequency axis. Let us assume it is implemented as a device Hn. Such a device is all-pass and has a phase versus frequency response with periodicity fr on the frequency axis. Thus its impulse response is a discrete time function consisting in samples with $1/fr$ spacing. Fig. 2(b) shows the phase versus frequency response of the delay $1/N \cdot fr$ and Fig. 2(c) shows the phase versus frequency response of the device $H1$. The sum of these responses produces the desired staircase phase characteristic $\varphi_1(f)$.

The polyphase network decomposed using this method is shown in Fig. 3, with $T = 1/N \cdot fr$, along with the summation device. The fact that the whole system has a frequency response which has the period $N \cdot fr$ on the frequency axis permits such a system to be realized as a discrete system with sampling rate $N \cdot fr = 1/T$ and Z transfer function $H(z)$ with $z = \exp(j \cdot 2\pi \cdot f \cdot T)$. The delays in Fig. 3 have the transfer functions $Z^{-n} (n = 0, 1, \cdots, N-1)$. The Hn, which have an impulse response consisting of samples with $1/fr = NT$ spacing, have a transfer function which can be expressed in terms of Z^N. Finally, the whole system of Fig. 3 can be described by the following equation:

$$H(Z) = \sum_{n=0}^{N-1} Z^{-n} \cdot Hn(Z^N).$$

As the system behaves like a low-pass filter, $H(Z)$ is also the transfer function of a digital low-pass filter with sampling frequency $N \cdot fr = 1/T$ and cutoff frequency $fr/2$.

The problem arises now of the design of the Hn set. They can be designed as all-pass devices approximating an ideal phase characteristic. However, that approach might become tricky with increasing N and will not be considered here. On the contrary, a simple general technique will be provided leading to devices approximating ideal all-pass and phase characteristics.

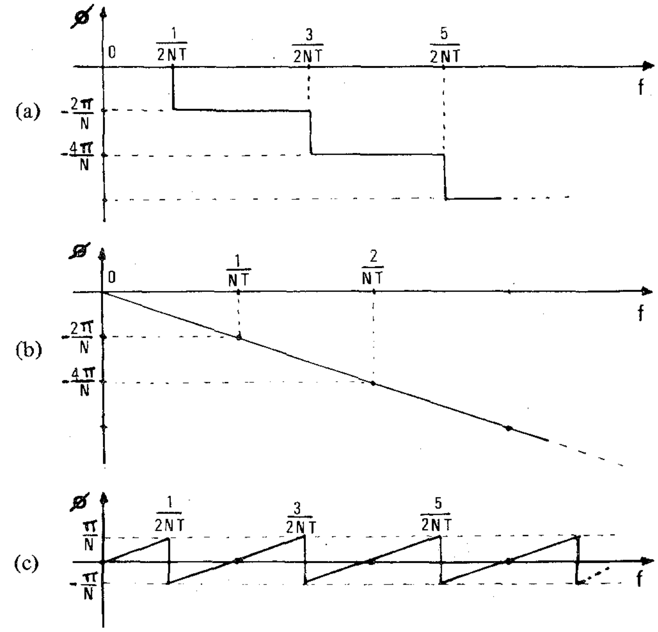


Fig. 2. (a) Phase characteristic $\varphi_1(f)$. (b) Phase characteristic of the delay T. (c) Phase characteristic of H_1.

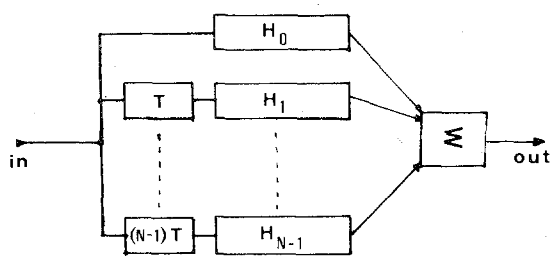


Fig. 3. Block diagram of the polyphase network of order N.

III. Design of a Polyphase Network from the Associated Low-Pass Filter

The starting point for the design of the polyphase network will be the transfer function of the associated filter.

If the design is optimal, the transfer function $\Sigma_{n=0}^{N-1} Z^{-n} \cdot Hn(Z^N)$ should give an optimal low-pass filter whose transfer function $H(Z)$ can be obtained by conventional design techniques and expressed in terms of its poles P_k and zeros Z_k in the Z plane by the following equation:

$$H(Z) = A \times \frac{\prod_{k=1}^{K} (Z - Z_k)}{\prod_{k=1}^{K} (Z - P_k)},$$

A being a constant and K an integer.

What remains to be found is the relation between the poles and zeros of $H(Z)$ and the coefficients of the $Hn(Z^N)$ in order to get the identity

$$A \times \frac{\prod_{k=1}^{K} (Z - Z_k)}{\prod_{k=1}^{K} (Z - P_k)} = \sum_{n=0}^{N-1} Z^{-n} Hn(Z^N). \tag{1}$$

This will be obtained by considering that the denominator of $H(Z)$ can become a function of Z^N using the identity

$$\frac{1}{Z - P_k} = \frac{Z^{N-1} + P_k Z^{N-2} + \cdots + P_k^{N-1}}{Z^N - P_k^N}.$$

The transfer function $H(Z)$ can then be rewritten

$$H(Z) = A \times \frac{\prod_{k=1}^{K} (Z - Z_k)(Z^{N-1} + P_k Z^{N-2} + \cdots + P_k^{N-1})}{\prod_{k=1}^{K} (Z^N - P_k^N)}.$$

The numerator can be expressed in terms of Z^{-1} as a polynomial of degree KN:

$$H(Z) = \frac{\sum_{i=0}^{KN} a_i Z^{-i}}{\sum_{k=1}^{K} (1 - P_k^N Z^{-N})}.$$

A decimation process can be used to represent the numerator polynomial of the variable Z with degree KN as N polynomials of the variable Z^N with degree K multiplied by factors Z^{-n} with $n = 0, 1, \cdots, N - 1$. Then the decomposition of $H(Z)$ according to (1) is obtained if $Ho(Z^N)$ is taken as

$$H_o(Z^N) = \frac{\sum_{k=0}^{K} a_{kN}(Z^{-N})^k}{\prod_{k=1}^{K} (1 - P_k^N Z^{-N})} \tag{2}$$

and $Hn(n = 1, 2, \cdots, N - 1)$ is taken as

$$H_n(Z^N) = \frac{\sum_{k=0}^{K-1} a_{kN+n}(Z^{-N})^k}{\prod_{k=1}^{K} (1 - P_k^N Z^{-N})}. \tag{3}$$

It is worth pointing out that the above derivation is still valid if the filter cutoff frequency fc deviates from half the reference frequency fr of the polyphase network; for example, if $fc < fr/2$, the $Hn(Z^N)$ are no longer essentially all-pass, but attenuate the signal components in the frequency band (fc, fr).

Let us illustrate the decomposition process by an example in which $N = 4$. The filter to be realized by a polyphase network structure of an order of 4 has the following specifications: cutoff frequency $fc = 0.425 \times fr$, in-band ripple 0.2 dB and out-of-band attenuation over 64 dB from the frequency $fr/2$ on.

The filter sampling frequency is $fs = 4 \times fr$ and conventional design techniques show that the specifications are met with an elliptic filter of order $K = 8$; the locations in the Z plane of the poles P_k obtained are shown in Fig. 4 and the frequency response is represented in Fig. 5(a).

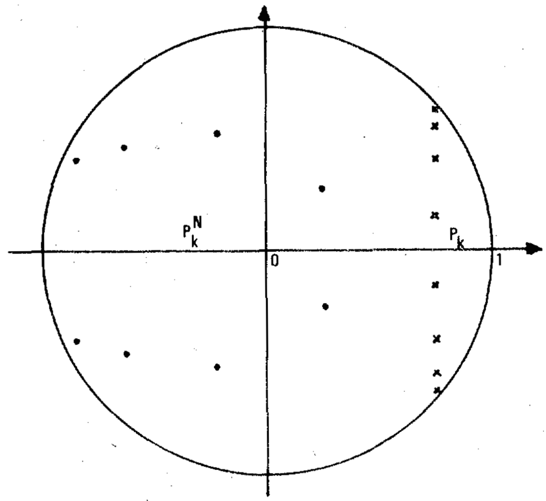


Fig. 4. Poles of the filter to be realized by polyphase network P_k(X) in the Z plane and $P_k^N(\bullet)$.

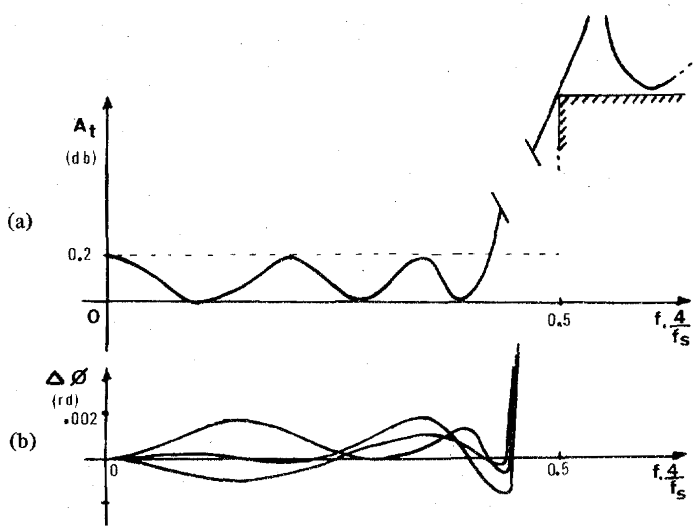


Fig. 5. (a) Frequency response of the associated filter. (b) Phase differences between phase shifters and delays $\varphi_n(f) + n \cdot 2\pi \cdot f/fs$.

According to (2) and (3) the poles of the $Hn(n = 0, 1, 2, 3)$ in the Z^N plane are obtained by raising the $P_k(k = 1, \cdots, 8)$ to the Nth power. The coefficients are computed according to the above formulas and the phase versus frequency characteristics are evaluated. It is verified that the phases of the $Hn(n = 1, 2, 3)$ relative to H_O approximate the previously mentioned sawtooth shapes. The extent to which the phase shifters match the delays in each path of the polyphase network to approximate the staircase responses of Fig. 1 is illustrated in Fig. 5(b).

At this stage some remarks seem to be relevant. First, all the $Hn(Z^N)$ have the same poles which, therefore, do not influence the degree of cancellation of the signal components to be removed and, consequently, the filter attenuation. Second, the amplitude versus frequency response of each path is close to unity, which is a desirable feature in practice. Finally, a simple technique has been provided to design a polyphase network from the associated low-pass filter. The implementation of the structure obtained will be dealt with in the next section.

IV. Implementation

The polyphase network is defined by the Z transfer function of each path: $Hn(Z^N)$ ($n = 0, 1, \cdots, N-1$). For the implementation, conventional procedures can be used. The direct forms are usually avoided because of the filter response sensitivity to the accuracy of the representation of the denominator coefficients, as shown in [6]. On the contrary, it has been demonstrated in [7] that, compared to either the cascade or the parallel form, direct forms are attractive structures for the realization of finite impulse response (FIR) filters. The authors share the opinion that the same still holds for a class of infinite impulse response (IIR) filters; those having low sensitivity to denominator coefficients. It is particularly true if the digital filter is connected to an analog-to-digital (A/D) converter. Then the dynamic range of the digital filter is generally much larger than that of the A/D converter and, if a direct form is used, half the multiplications involved can be carried out with signal-sample word lengths corresponding to the small dynamic range, which compensates for an increase of coefficient word lengths, provided that increase is not too significant. Other advantages of direct forms include reduced complexity of overhead circuity and the possibility of a unique overflow detection device. The $Hn(Z^N)$ are likely to belong to such a class, as can be expected from the fact that their poles are those of the associated filter raised to the Nth power. Thus the sensitivity to denominator coefficients is considerably reduced. Anyway, the penalty in denominator coefficient bits due to the choice of a direct form instead of the cascade form can be estimated.

Let $Hn^*(Z^N)$ be the transfer function of the phase shifter in the nth path after rounding of the denominator coefficients to a quantization size of q_1 and $d(Z^N)$ be the error polynomial introduced in this process. It is known, for example, from [8] that

$$H_n^*(Z^N) = H_n(Z^N)\left[1 - \frac{d(Z^N)}{D(Z^N)} \cdot \frac{1}{1 + \dfrac{d(Z^N)}{D(Z^N)}}\right]$$
$$\approx H_n(Z^N)\left[1 - \frac{d(Z^N)}{D(Z^N)}\right].$$

Assume K is an even number, the denominator $D(Z^N)$ of $Hn(Z^N)$ can be factorized $D(Z^N) = \Pi_{k=1}^{K} D_k(Z^N)$ and $Hn(Z^N)$ decomposed into $K/2$ second-order sections in which rounding of the denominator coefficients to a quantization size of q_2 introduces error polynomials $d_k(Z^N)$ ($k = 1, 2, \cdots, K/2$). Then

$$H_n^*(Z^N) \approx H_n(Z^N)\left[1 - \sum_{k=1}^{K/2} \frac{d_k(Z^N)}{D_k(Z^N)}\right].$$

A rather high statistical estimate of $d(Z^N)$ is given by $q_1/2\sqrt{3} \times K$. The same formula for $d_k(Z^N)$ leads to $q_2/\sqrt{3}$ [7]. The rounding errors in direct form are at least similar to those in cascade form if

$$\frac{q_2}{q_1} \geqslant \frac{K}{2 \cdot \displaystyle\sum_{k=1}^{K/2} \frac{D(Z^N)}{D_k(Z^N)}}.$$

Then an order of magnitude of the number of extra denominator coefficient bits for the direct forms compared to the cascade forms is given by

$$b = \log_2 \left(\frac{K}{2 \cdot \displaystyle\min_{(|Z|=1)} \left| \sum_{k=1}^{K/2} \frac{D(Z^N)}{D_k(Z^N)} \right|} \right).$$

In the above-mentioned example $b = 3$ is judged sufficiently small to make the direct forms attractive in the implementation of the paths of the polyphase network.

V. Application to Sample-Rate Alteration

The polyphase network structure can perform digital filtering in the general case, but it might seem too complicated a structure to be competitive with conventional procedures. In fact, it offers significant savings in computation rate and hardware in two important applications, sample-rate alteration and, chiefly, realization of filter banks.

Several techniques have been suggested to reduce the computation speed in sample-rate alteration [1], [2]. The gain mainly comes from the use of FIR filters, which, compared to IIR filters, involves an increase of the storage needed. However, it is desirable in some applications to minimize the storage or to achieve a minimum phase or group delay; a polyphase network approach can meet these goals and, however, bring a significant reduction of the computation rate using recursive devices.

The main purpose of Sections II and III was to show that, starting from an intuitive decomposition of the filtering process, a realization of a low-pass filter with Z transfer function $H(Z)$ could be obtained by a set of N devices with transfer functions $Hn(Z^N)$ connected to delay elements with transfer functions Z^{-n}, where $n = 0, 1, \cdots, N-1$.

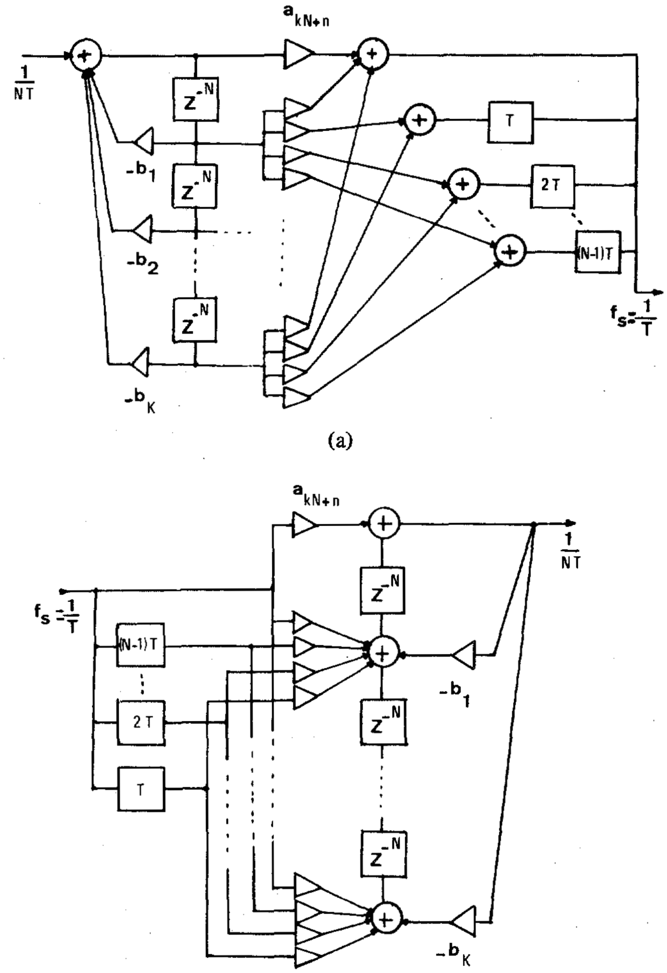
The decomposition obtained is expressed by the equality

$$H(Z) = \sum_{n=0}^{N-1} Z^{-n} H_n(Z^N).$$

In the time domain the fact that the Hn are functions of Z^N implies that the signal samples coming out of the corresponding devices are weighted sums of input samples and previous output samples with N sampling period spacing. Consequently, if a sampling-rate alteration by a factor of N is desired between input and output of the filter, it is sufficient to carry out the multiplications and additions in the devices Hn at the low-sampling rate. A property similar to that of finite impulse response filters has been obtained.

Moreover, it can be seen from (2) and (3) that the $Hn(Z^N)$ all have the same denominator. Then the corresponding computations can be carried out only once, assuming proper ordering of the operations. This will be illustrated in the cases of sampling-rate reduction or increase by a factor of N, which can be achieved by polyphase networks with N paths. For the sampling-rate increase case, the recursive operations can be carried out first on the input signal at the low-sampling rate; the adequate procedure is given in Fig. 6(a). The multiplication rate is $(KN + K + 1) \times fs/N$, fs being the high sampling rate, which is to be compared with $2K \cdot fs$ for the conventional filter; the storage for signal samples is the same in both cases, but the read-only memory hardware for the coefficients is increased in the polyphase network approach. The procedure for sample-rate reduction is depicted in Fig. 6(b). It uses the transpose configuration and has similar advantages.

Finally, the computation rate for large N can be cut by approximately half in the polyphase network approach at the expense of an increased read-only memory hardware.



(b)

Fig. 6. (a) Structure for sample-rate increase. (b) Structure for sample-rate reduction.

VI. Realization of Uniform Filter Banks

In fact the application of the preceding section is a particular case of a more general one; which is the realization of a uniform filter bank through combination of a polyphase network and a DFT computer. An application of this technique in communications has been described in [9]. The principle will be recalled here.

A uniform filter bank is defined and analyzed in [10]. It consists of a set of N filters, of which $N - 1$ are frequency-shifted versions of a basic filter. Following notations of Section III, the mth filter of the bank corresponds to a frequency shift of $m \times fs/N$, which implies that its Z transfer function is obtained through replacement of Z by $Z \cdot \exp(j \cdot 2\pi m/N)$ in (1), for example. The new transfer function reads

$$B_m(Z) = \sum_{n=0}^{N-1} Z^{-n} \exp\left(-j2\pi \frac{m \cdot n}{N}\right) H_n(Z^N).$$

It is worth noting that the $Hn(Z^N)$ are the same for all the filters of the bank and, consequently, a factorization can take place which is the key to substantial computation rate savings. The complete bank of filters is defined by the following matrix equation in which $W = \exp(-j \cdot 2\pi/N)$:

$$\begin{bmatrix} Bo \\ B_1 \\ \\ \vdots \\ B_{N-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & W & W^2 & \cdots & W^{N-1} \\ 1 & & & & \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & W^{N-1} & \cdots & & W^{(N-1)^2} \end{bmatrix} \begin{bmatrix} Ho(Z^N) \\ Z^{-1}H_1(Z^N) \\ \\ \vdots \\ Z^{-(N-1)}H_{N-1}(Z^N) \end{bmatrix}.$$

In this matrix equation the DFT matrix appears along with the polyphase network of the previous sections. Using this configuration, it turns out that the extra computations needed to obtain a uniform bank of filters instead of a low-pass filter are only those contained in the DFT. The block diagram of the circuitry associated with this equation is given in Fig. 7.

In this application all the Hi still have the same denominator coefficients, but the recursive part can no longer be shared as it processes signal samples which are different in the various paths. Consequently, the computation rate in the polyphase network is $2K \times fs$, if the N outputs of the filter bank are sampled at the rate fs/N each. A bank of N separate conventional filters providing outputs sampled at the rate fs each would require N times more multiplications. Besides, the low sensitivity of the polyphase structure to the accuracy of representation of the denominator coefficients described in Section III still significantly simplifies the hardware.

The case in which N is a power of 2 deserves special attention. The amount of multiplications and additions in the real signal DFT is about $N \cdot \log_2 (N/2)$. The computation rate in

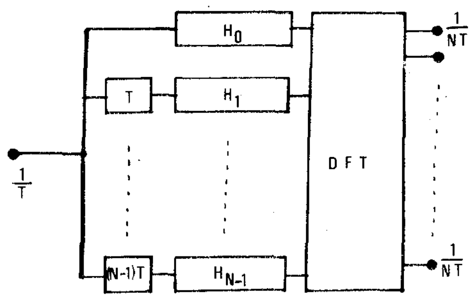


Fig. 7. Bank of filters using DFT polyphase network combination.

the filter bank, if outputs are sampled at fs/N, amounts to $fs \cdot [2K + \log_2 (N/2)]$. This illustrates the considerable advantage of the approach for large values of N.

VII. Discussion and Conclusion

The approach to digital filtering which has been presented shows that some of the attractive properties for sample-rate alteration of FIR filters can be extended to IIR filters as well. It generalizes the window FFT technique for spectrum analysis, allowing the channel filter to be a true IIR filter. Besides, the polyphase presentation corresponds to a physical description of how the system really operates. If a sine wave is fed to the input of the system of Fig. 3, for example, the various phase shifts at the output of the $Hn(n = 0, 1, \cdots, N-1)$ can be visualized.

However, some open questions remain. For example, how that technique could be used for nonuniform filter banks or whether other transfer functions $Hn(Z^N)$ leading to attractive structures could be found to satisfy (1).

To summarize the results of this paper, let us state that the polyphase network approach to sample-rate alteration can cut by nearly half the multiplication and addition rate when the alteration factor is large, compared with conventional filters. The active memory hardware is similar, but a more coefficient read-only memory is required. The main application of this structure, however, lies in the realization of a uniform bank of filters through combination of the polyphase network with a DFT computer. The low sensitivity of this structure to coefficient word lengths combined with the significant computation rate reduction achieved can considerably simplify the hardware. The basic reason for these reductions of the multiplication and addition rate is that in the polyphase network approach, the sampling rate is kept to the lowest value throughout the processing.

Acknowledgment

The authors wish to acknowledge the contributions to this work by the reviewers and Dr. R. W. Schafer, whose numerous comments and suggestions helped to improve the paper considerably.

References

[1] R. W. Schafer and L. R. Rabiner, "A digital signal processing approach to interpolation," *Proc. IEEE*, vol. 61, pp. 692-702, June 1972.

[2] M. G. Bellanger, J. L. Daguet, and G. P. Lepagnol, "Interpolation, extrapolation, and reduction of computation speed in digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-22, pp. 231-235, Aug. 1974.

[3] L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.

[4] S. Darlington, "On digital single-sideband modulators," *IEEE Trans. Circuit Theory*, vol. CT-17, pp. 409-414, Aug. 1970.

[5] H. W. Thomas and N. P. Lutte, "Z-transform analysis of non uniformly sampled digital filters," *Proc. Inst. Elec. Eng.* (London), vol. 119, pp. 1559-1567, Nov. 1972.

[6] J. F. Kaiser, "Digital Filters," in *Systems Analysis by Digital Computer*, F. F. Kuo and J. F. Kaiser, Eds. New York: Wiley, 1966, ch. 7.

[7] D. S. K. Chan and L. R. Rabiner, "Analysis of quantization errors in the direct form for FIR filters," *IEEE Trans. Audio Electroacoust.*, vol. AU-21, pp. 354-366, Aug. 1973.

[8] F. Bonzanigo and F. Pellandini, "Problèmes de réalisation des filtres digitaux," *AGEN*, no. 9, pp. 56-71, 1969.

[9] M. G. Bellanger and J. L. Daguet, "TDM-FDM transmultiplexer: Digital polyphase and FFT," *IEEE Trans. Commun. (Special Issue on Communications in Europe)*, vol. COM-22, pp. 1199-1205, Sept. 1974.

[10] R. W. Schafer and L. R. Rabiner, "Design of digital filter banks for speech analysis," *Bell Syst. Tech. J.*, vol. 50, pp. 3097-3115, Dec. 1971.