Synthesis and Realization of

Noncausal Digital Filters

Chok-Ki CHAN

A thesis submitted in partial fulfilment of

the requirements for the degree of Doctor of Philosophy,

the Chinese University of Hong Kong

Department of Electronics

The Chinese University of Hong Kong

April 1984

For my wife and family

ACKNOWLEDGEMENTS

It is my pleasure to express my gratitude to my supervisor, Professor C. F. Chen, for his most stimulating guidance, valuable advice and encouragement throughout the course of this research work.

I would also like to express my gratitude to the donor of the Cheng Yick Chi Graduate Fellowship whose financial assistant and support are proved to be valuable.

Thanks are also due to Dr. J. S. L. Wong and Mr. D. P. Kwok for their continual encouragement. Moreover, I am particularly grateful to Dr. C. P. Kwong and Dr. P. C. Ching for their helpful discussions and criticisms.

Finally, I am indebted to Miss Rhoda Lam, Miss Adeline Siu and Miss Winnie Lau for their patience in typing the manuscript.

Abstract

In the field of digital filter design, most of
the research efforts have been concentrated in the
synthesis and realization of causal filters. Noncausal
digital filters receive far less attention mainly due to
the common notion that they are physically unrealizable.
By making use of the flexibility of digital computer,
techniques for the realization of noncausal filter have
been developed recently. In this thesis, the synthesis
and realization methods for noncausal filters are
considered.

The synthesis problem is tackled by decompositing
the noncausal filter into a causal subfilter and a purely
noncausal subfilter connected in parallel or in series.
For zero phase filters, the subfilters can be made identical
for most cases. The frequency response relationships are
given to facilitate the design process in practical
problems.

A sample-by-sample approach to noncausal filter
realization has been developed with two new techniques
introduced. Compared with the conventional block processing
approach, the new methods have the advantages of smaller
basic group delay, small memory size requirements and
exactly known phase error.

The purely noncausal part of the filter is realized by a nonrecursive subfilter or a recursive one. The nonrecursive realization is based on a FIR filter design method using Wiener-Lee decomposition technique and the unit circle real part function. The convergency of the FIR subfilter is guaranteed since many practical filter responses satisfy Dirichlet conditions. The recursive realization is based on the continued fraction expansion often used as a model reduction technique. For narrowband lowpass or highpass filters, recursive realization is more computationally efficient than the nonrecursive one. Many practical examples are presented. Comparisons are also made with the conventional technique.

CONTENTS

Chapter 1

INTRODUCTION

In the past two decades, the field of digital
signal processing has grown enormously to provide concrete
theorical background for many design and implementation *theoretical*
problems in a huge number of areas [O2]. The first attempt
to provide a comprehensive theory was due to Gold and Radar
in their classic book on digital signal processing [G2].
Since then, many textbooks ~~by various authors~~, such as *those by*
Oppenheim and Schafer [O1], Cappellini et al [C4], Rabiner
and Gold [R3] and Chen [C7], have deal with both the theory
and applications of digital signal processing in great depth.
Moreover, the rapid advance of high speed digital integrated
electronics has widened the application areas of digital
signal processing from the low frequency end, such as seimic,
biomedical and sonar signals, to the high frequency end, like
speech and radar signals [O2].

In practical problems, the digital system design
engineer would inevitably encounter the problem of
realizability from time to time. The problem is so
fundamental that it is widely discussed in the literature.
Naturally, it leads to the classification of the causal
(nonanticipative) and the noncausal (anticipative) digital
systems. A most common notion is that only causal systems
are physically realizable; or, in other words, noncausal
systems are unrealizable and therefore have little practical
values. Logically, based on this classification, many of
the research and development works *were* concentrate their efforts

only in the design, implementation and application of causal digital systems. Usually, most design methodologies restrict, at the very beginning, their designs to only causal type of transfer functions without ever consider noncausal type of functions. Nevertheless, many well established analytical tools, such as difference equations and the two-sided z-transform, can handle both causal and noncausal problems [01]. The concept of unrealizability has impeded the development of noncausal digital system design, not to mention its applications, to real world problems.

Despite being unrealizable, noncausal systems are useful in many aspects. In system analysis, noncausal functions frequently appear in some intermediate steps of mathematical operations. This usually happens when causal functions are being recombined or decomposed. A well known analysis procedure is the decomposition of a causal sequence h(k) into even and odd parts (Fig. 1.1), where,

$$h(k) = h_e(k) + h_0(k), \qquad \text{all } k \qquad (1.1)$$

$$\text{with } h_e(k) = \frac{1}{2}(h(k) + h(-k))$$

$$\text{and } h_0(k) = \frac{1}{2}(h(k) - h(-k)) \qquad (1.2)$$

Both $h_e(k)$ and $h_0(k)$ are noncausal sequences, which are particularly important in providing proofs for many of the causal digital system theorems, such as the Hilbert transform [01].
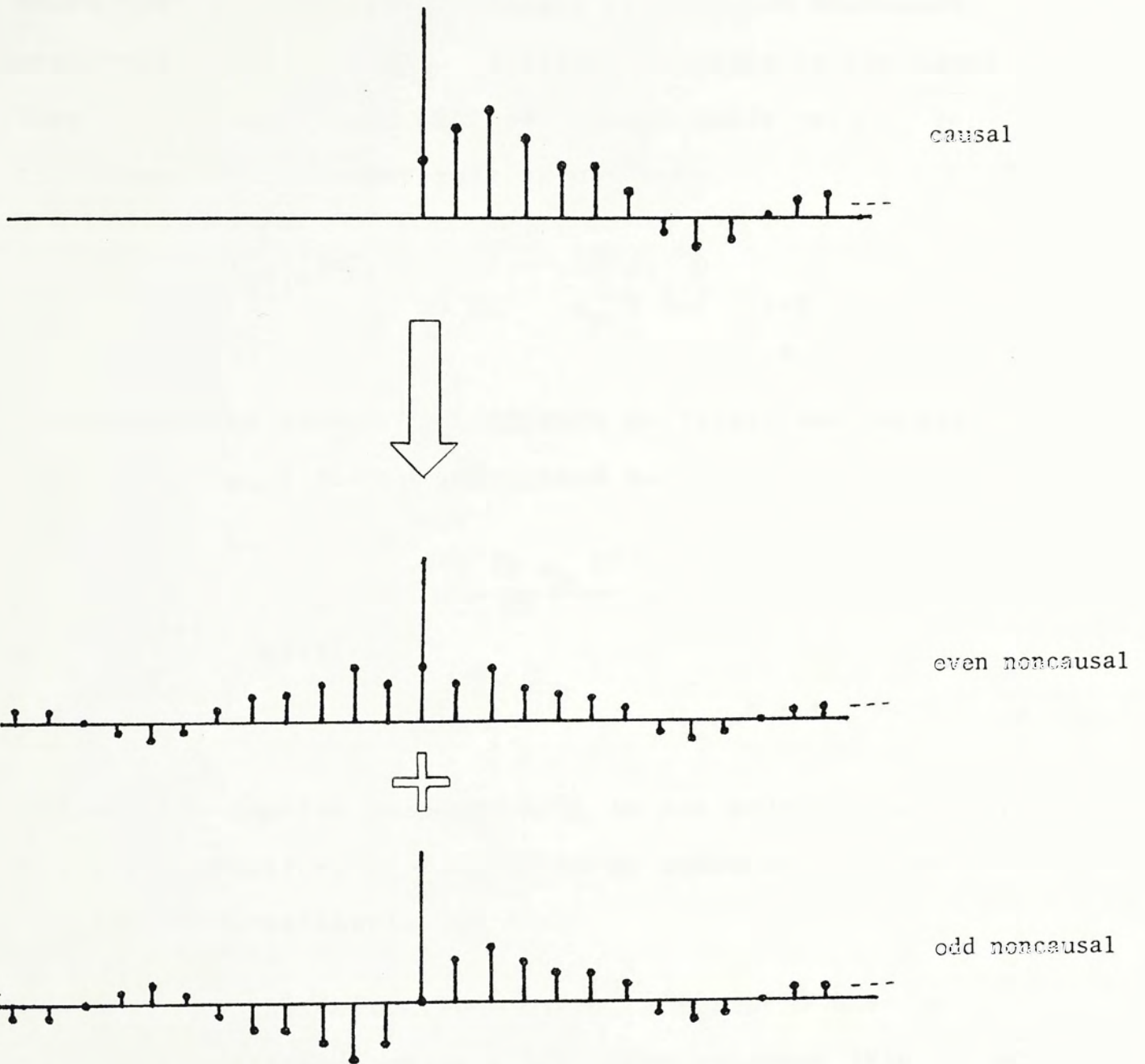
2

causal

even noncausal

odd noncausal

Fig. 1.1   Decomposition of *a* causal sequence into even
and odd sequences

3

In the field of digital filter design, it is well known that many noncausal filters can provide benchmark performance [C4, O1, R3]. A classic example is the ideal lowpass digital filter with zero phase shift (Fig. 1.2a). The frequency characteristic is given by,

$$H(e^{j\omega T}) = \begin{cases} 1, & |\omega| < \omega_p \\ 0, & \omega_p < |\omega| < \pi/T \end{cases} \qquad (1.3)$$

Performing the inverse z-transform on (1.3), the impulse response (Fig. 1.2b) is determined as,

$$h(k) = \begin{cases} \dfrac{\sin (k \omega_p T)}{k\pi}, & k \neq 0 \\ \dfrac{\omega_p T}{\pi}, & k = 0 \end{cases} \qquad (1.4)$$

Indeed, the impulse response $h(k)$ is not equal to zero for $k < 0$ and therefore is considered as noncausal in nature and physically unrealizable.

Another often encountered example is the ideal Hilbert transformer having a frequency response (Fig. 1.3a) of

$$H(e^{j\omega T}) = \begin{cases} -j, & 0 < \omega < \pi/T \\ j, & -\pi/T < \omega < 0 \end{cases} \qquad (1.5)$$

The impulse response (Fig. 1.3b) is an odd sequence given by,

$$h(k) = \begin{cases} \dfrac{2 \sin^2 (k\pi/2)}{k\pi}, & k \neq 0 \\ 0, & k = 0 \end{cases} \qquad (1.6)$$
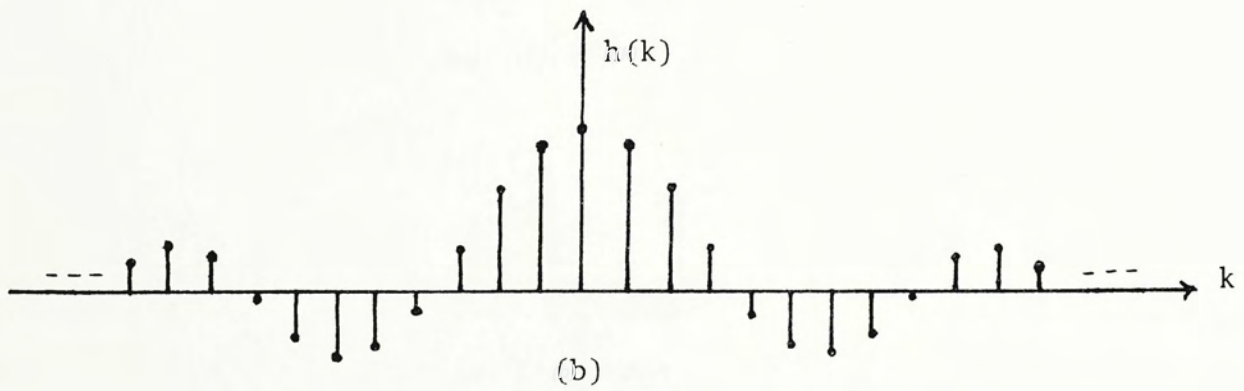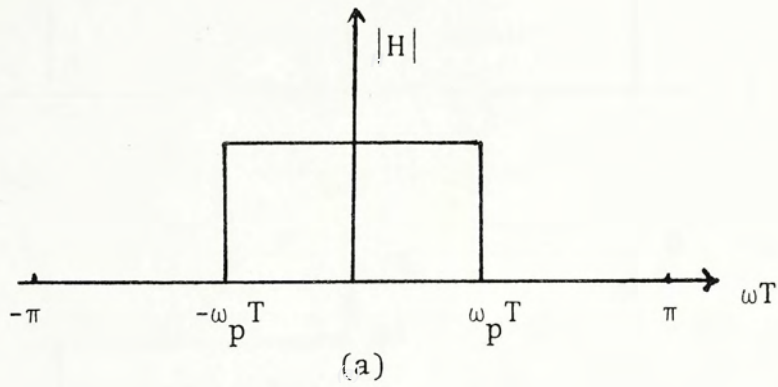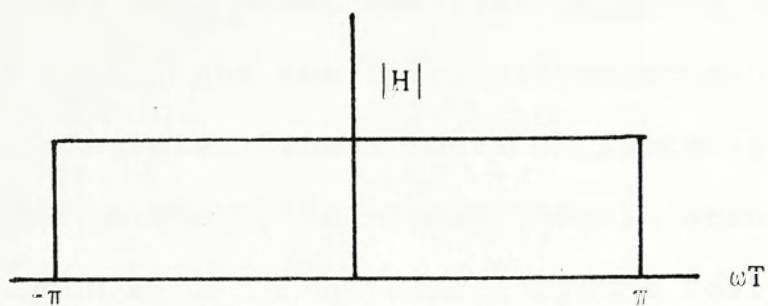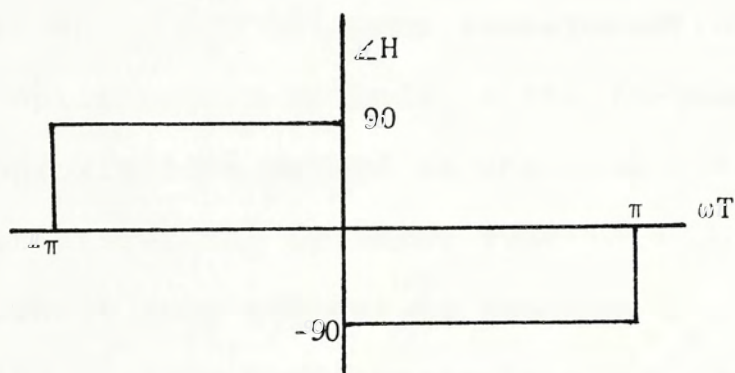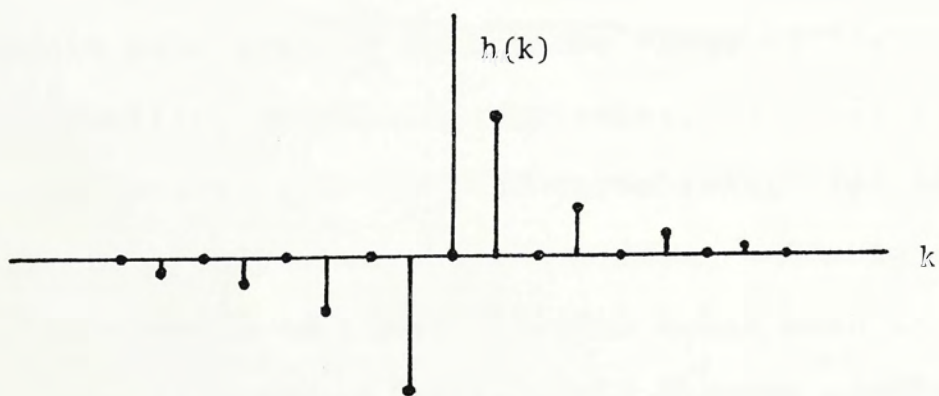
Fig. 1.2 Ideal zero phase lowpass digital filter:
(a) magnitude response; (b) impulse response.

Fig. 1.3 Ideal Hilbert transformer: (a) magnitude response;
(b) phase response; (c) impulse response.

Clearly, the ideal Hilbert transformer is also noncausal.
In fact, there are many other familiar examples, such as the
ideal bandpass filter and the ideal differentiator, that are
all noncausal in nature. These ideal characteristics are
often considered as the optimum unachievable standard
responses on which many of the causal system designs are
based. However, most of the well known causal filter design
methods, for example, the bilinear transformation, impulse
invariant and optimization methods in the frequency domain,
or the Pade approximation method in the time domain, confine,
at the earliest stage, the transfer functions to be of the
causal type without ever attempting to consider noncausal type
of functions [R3]. This unfortunate fact has led to the
relative undevelopment of noncausal filter design techniques
after all these years.

you haven't said that yet!

With so many desirable features, noncausal digital
systems would have applied to a wider range of fields should
the unrealizability constraint not exist. In reality, the
anticipative nature (the cause of unrealizability) of these
systems is true only when the independent variable of the
system is ~~an element of~~ time. In some areas such as image
processing the independent variable is a "space element" and,
therefore, the anticipatory difficulty no longer exists. In
fact, the application of noncausal processing techniques for
distortionless image signal filtering is well known [C4].

Even when time is the independent variable, it is
still possible to implement noncausal systems in an

approximate manner largely due to the flexibility of digital computers. Recently, a two-pass technique for implementing zero phase cascade noncausal filters has been developed [K1, C1]. It makes use of the fact that sampled digital data in the memory can be processed in both the forward and reversed direction. Therefore, even the unrealizability constraint is truth in a strict sense, realizable approximations or implementation procedures can be derived for noncausal systems. However, in the literature, research works in the design, realization and application of noncausal systems are very rare.

In this thesis, the problems in the design and realization of noncausal systems are being investigated. New design concepts and realization techniques are proposed. The techniques are based on the decomposition of a noncausal system into a causal and a purely noncausal subsystems connected either in parallel or in cascade. The problem of stability is considered with respect to the decomposited subsystems. The characteristics of resulting noncausal filters from different realization techniques are studied. Results from the proposed methods are compared with the conventional method.

1.1 Fundamental differences between causal and noncausal filters

In the literature, causal digital filters are further classified into the finite impulse response (FIR) filters and the infinite impulse response (IIR) filters. The most general form of the one-sided z-transform of IIR filters can be written as

$$H(z) = \frac{\displaystyle\sum_{i=0}^{M} b_i z^{-i}}{\displaystyle\sum_{i=0}^{N} a_i z^{-i}} \, , \qquad M \leq N \qquad (1.7)$$

where a_N is not equal to zero and there is no common factor between numerator and denominator. The FIR transfer function has the general form,

$$H(z) = \sum_{k=0}^{N} h(k) z^{-k} \qquad (1.8)$$

where $h(N)$ is not equal to zero.

The noncausal digital transfer function is given by the two-sided z transform of its impulse response,

$$H(z) = \sum_{k=-\infty}^{\infty} h(k) z^{-k} \qquad (1.9)$$

which may sometimes be written as a rational function with a region of convergence defined by $R_+ < |z| < R_-$.

It is interesting to look at the pole-zero patterns of this three types of digital filters. Due to the

stability constraint, the poles of the IIR filter must all
be located inside the unit circle while the zeros can be
anywhere in the z plane (Fig. 1.4). Since the coefficients
are real, both poles and zeros are in conjugate pairs. For
the FIR filter, the zeros can lie anywhere in z plane while
the only pole is located at z = 0. In general, the zeros
exist in conjugate pairs (Fig. 1.5a), however, if the linear
phase constraint is imposed, the zeros also have to be   *why!*
reciprocal pairs (Fig. 1.5b). For both causal IIR and FIR   *why?*
filters, all zeros must lie inside the unit circle if minimum
phase is required.

The noncausal filter, however, can have poles   *not consistent*
anywhere on the z plane (excluding) the unit circle (Fig. 1.6) *with the figure*
outside
due to stability criteria. There is no restriction on the
location of zeros.

In many practical filter design problems, the
spectral shaping ability of the filter is of utmost importance.
In general, the frequency response can be obtained from the
pole-zero locations. Let $D_i(e^{j\omega})$ be the vector magnitude
from the pole p_i to the point $e^{j\omega}$ on the unit circle and
$\theta_i(e^{j\omega})$ be the angle of the vector from p_i to $e^{j\omega}$, while
$c_i(e^{j\omega})$ denotes the magnitude due to the zero z_i and $\psi_i(e^{j\omega})$
denotes the angle. Then the frequency response can be
written as

$$H(e^{j\omega}) = \frac{\prod\limits_{i=1}^{M} c_i(e^{j\omega})}{\prod\limits_{i=1}^{N} D_i(e^{j\omega})} \exp[\sum\limits_{i=1}^{M} \psi_i(e^{j\omega}) - \sum\limits_{i=1}^{N} \theta_i(e^{j\omega})]$$

$$(1.10)$$

1

as shown in Fig. 1.7.

The spectral shaping ability of FIR filters are greatly limited by the fact that only zeros can be moved freely to construct the desire frequency response. Thus, given the same frequency specification, the FIR filter order is usually very high (even though, efficient implementation is possible due to the coefficient symmetry in the linear phase case or by making use of the fast convolution algorithm) [R3]. In the case of causal IIR filters, in addition to the freedom of positioning zeros, the poles are free to move within the unit circle. Hence, the spectral shaping ability of IIR filters is better; in other words, the order of IIR filters is in general lower than that of FIR filters when both filters are designed to satisfy the same frequency requirements.

The freedom of positioning both poles and zeros is greatest for the noncausal filter. The only restriction is that poles must not lie on the unit circle. Therefore, noncausal filters have the best ability to control both the magnitude and phase responses given the same number of poles and zeros. In fact, it is well known that causal IIR filters cannot have exact linear phase characteristic while noncausal filters, theoretically, can be designed to have exact linear phase [R3].
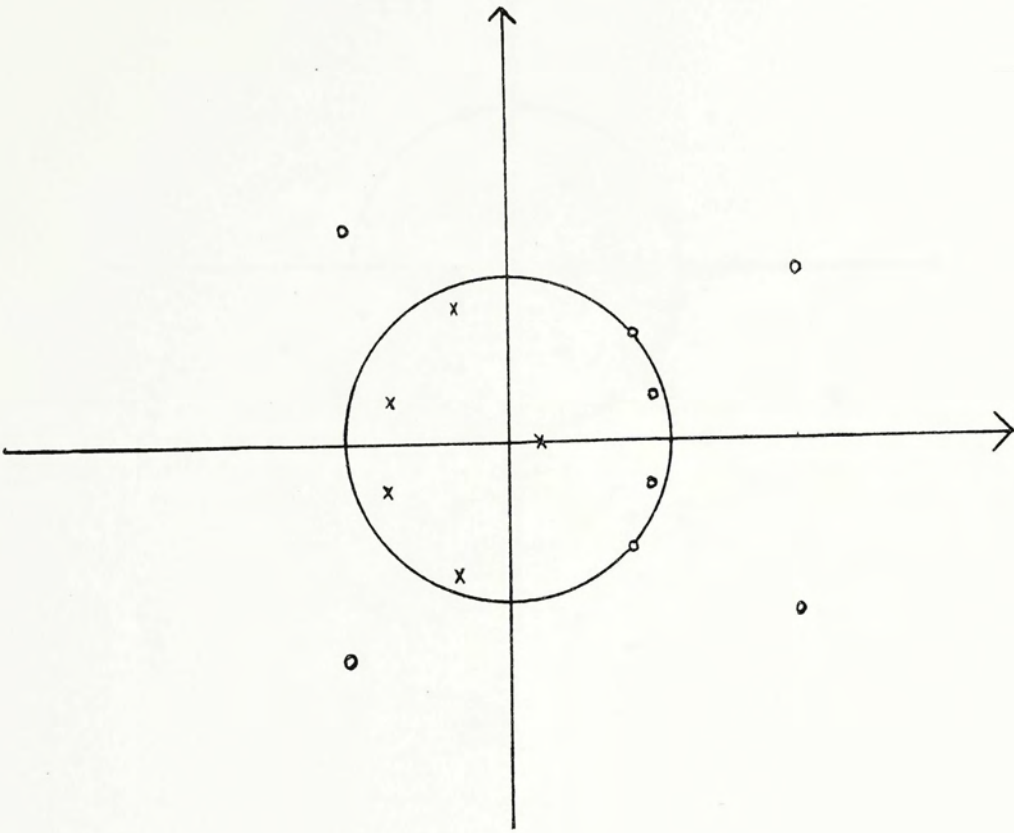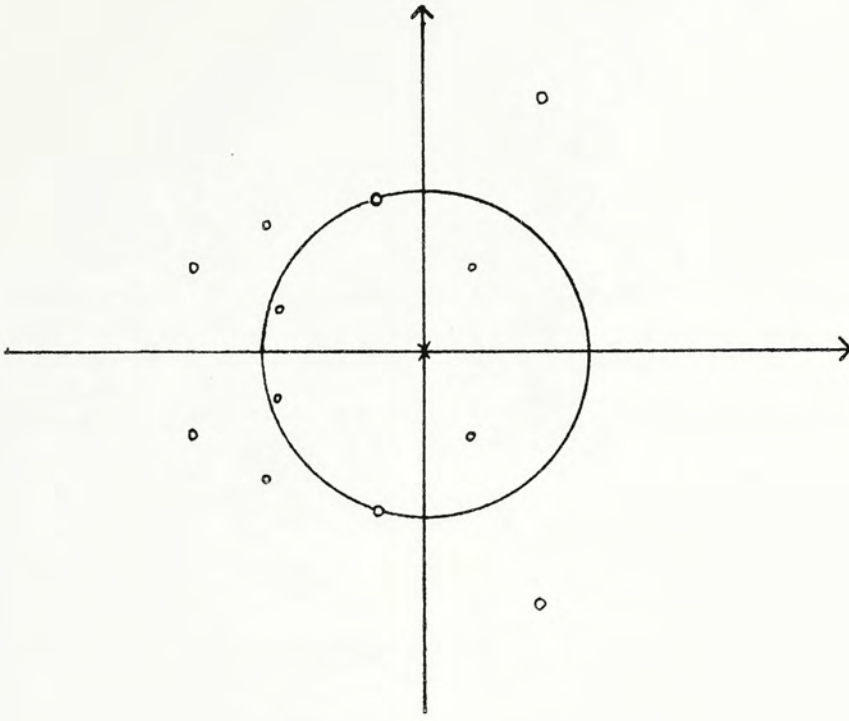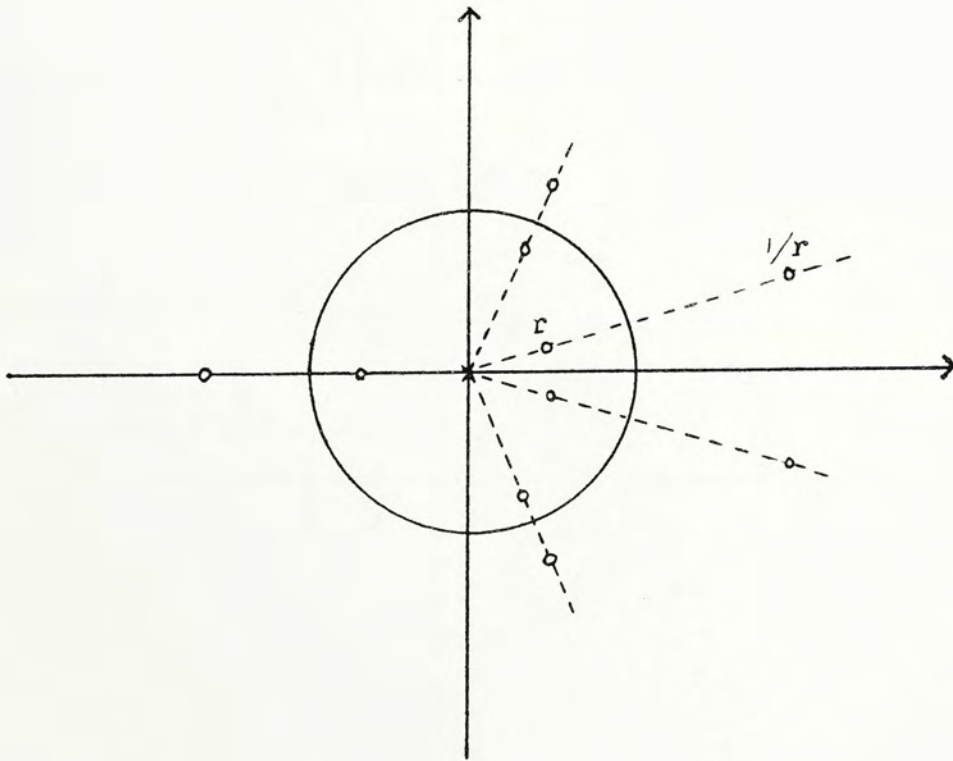
11

Fig. 1.4   Pole zero pattern of causal recursive
digital filter

(a)

(b)

Fig. 1.5a pole zero patterns of causal nonrecursive

filter (a) general case (b) linear phase case
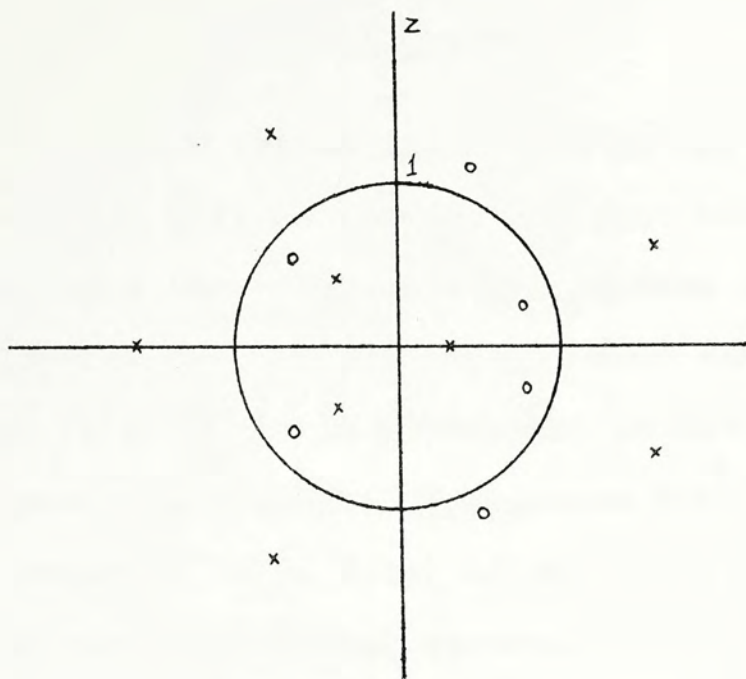
Fig. 1.6   pole zero pattern of noncausal filter
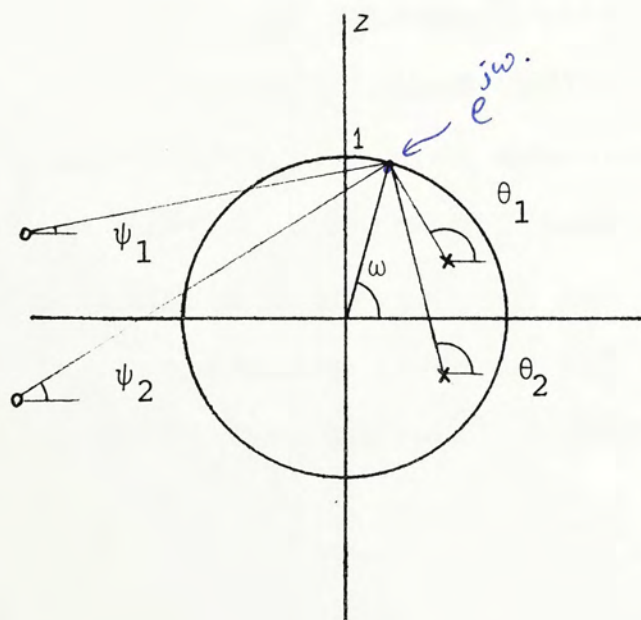
Fig. 1.7   Geometric evaluation of frequency response

from pole zero diagram

Chapter 2

Noncausal Systems

A noncausal system is defined as one of which the output depends not only on present and past values of input but also on future input values. The impulse response of a noncausal system has, therefore, non-zero samples for negative time values. It is convenient to define three different types of sequences. A sequence h(k) is referred as a causal sequence (Fig. 2.1a) if all its values are zero for k < 0. A purely noncausal sequence (Fig. 2.1b) is defined as one having zero values for k > 0. A noncausal sequence (Fig. 2.1c) refers to one having some non-zero values for both k < 0 and k > 0. Therefore, the noncausal system's impulse response is a noncausal sequence while that for a causal system is a causal sequence.

In this thesis, we concerned with the design and realization of linear time invariant (LTI) digital noncausal filters. The analysis and design of noncausal filters is not well developed even though the application of such filters in distortionless processing is not uncommon [C1, C4, K1]. It is necessary to establish the stability conditions before going into further details of design and realizability.

2.1 Stability criteria

A noncausal transfer function H(z) is given by the two-sided z transform of its impulse response h(k),

15

$$H(z) = \sum_{k=-\infty}^{\infty} h(k) \, z^{-k} \qquad (2.1)$$

For every bounded input sequence, the output of a stable noncausal system is also bounded. Thus, the noncausal LTI system is stable if and only if the impulse response is absolutely summable,

$$\sum_{k=-\infty}^{\infty} |h(k)| < \infty \qquad (2.2)$$

It is possible to separate the noncausal impulse response $h(k)$ into two causal impulse responses $g_1(k)$ and $g_2(k)$ by [C1],

$$g_1(k) = \begin{cases} h(k) & k > 0 \\ \frac{1}{2}(k), & k = 0 \\ 0 & k < 0 \end{cases}$$

$\leftarrow h(k)$

$$g_2(-k) = \begin{cases} 0, & k > 0 \\ \frac{1}{2}h(k), & k = 0 \\ h(k) & k < 0 \end{cases} \qquad (2.3)$$

Then we have,

$$h(k) = g_1(k) + g_2(-k) \qquad (2.4)$$

Now the stability condition (2.2) becomes,

$$\sum_{k=0}^{\infty} |g_1(k)| < \infty$$

and
$$\sum_{k=0}^{\infty} |g_2(k)| < \infty \qquad (2.5)$$

which are equivalent to the individual causal system's stability criteria. Taking the z-transform of (2.4), we get

$$H(z) = G_1(z) + G_2(z^{-1}) \qquad (2.6)$$

It follows that if the poles of $G_1(z)$ and $G_2(z)$ all lie inside the unit circle then the noncausal system is stable.

Now the stability criteria for noncausal system can be restated. If the noncausal system $H(z)$ is decomposited into two parallel connected systems consisting of a causal system $G_1(z)$ and a purely noncausal system $G_2(z^{-1})$, then $H(z)$ is stable if both $G_1(z)$ and $G_2(z)$ have all their poles inside the unit circle.

2.2 Design methodology

Even though many design techniques have been developed for causal filters [R3], the design of noncausal filters is relatively undeveloped possibly due to their limited application areas and inherent design difficulties. Some of the noncausal design techniques are discussed in this section.

The design problem is tackled by decompositing the noncausal system into parallel or cascade connecting systems consisting of a causal and a purely noncausal parts. The concausal transfer function $H(z)$ becomes $H(z) = H_1(z) H_2(z^{-1})$ if the cascade connection is used, and $H(z) = G_1(z) + G_2(z^{-1})$ in case of parallel connection (Fig. 2.2). Both $H_1(z)$ and $G_1(z)$ are causal systems while
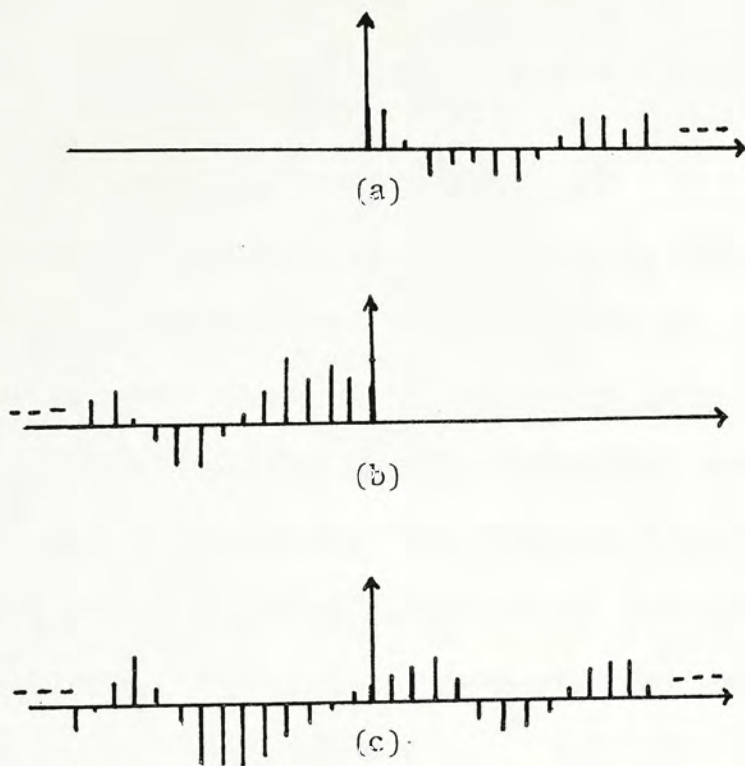
17

(a)

(b)

(c)

Fig. 2.1 Types of sequences: (a) causal; (b) purely noncausal; (c) noncausal.
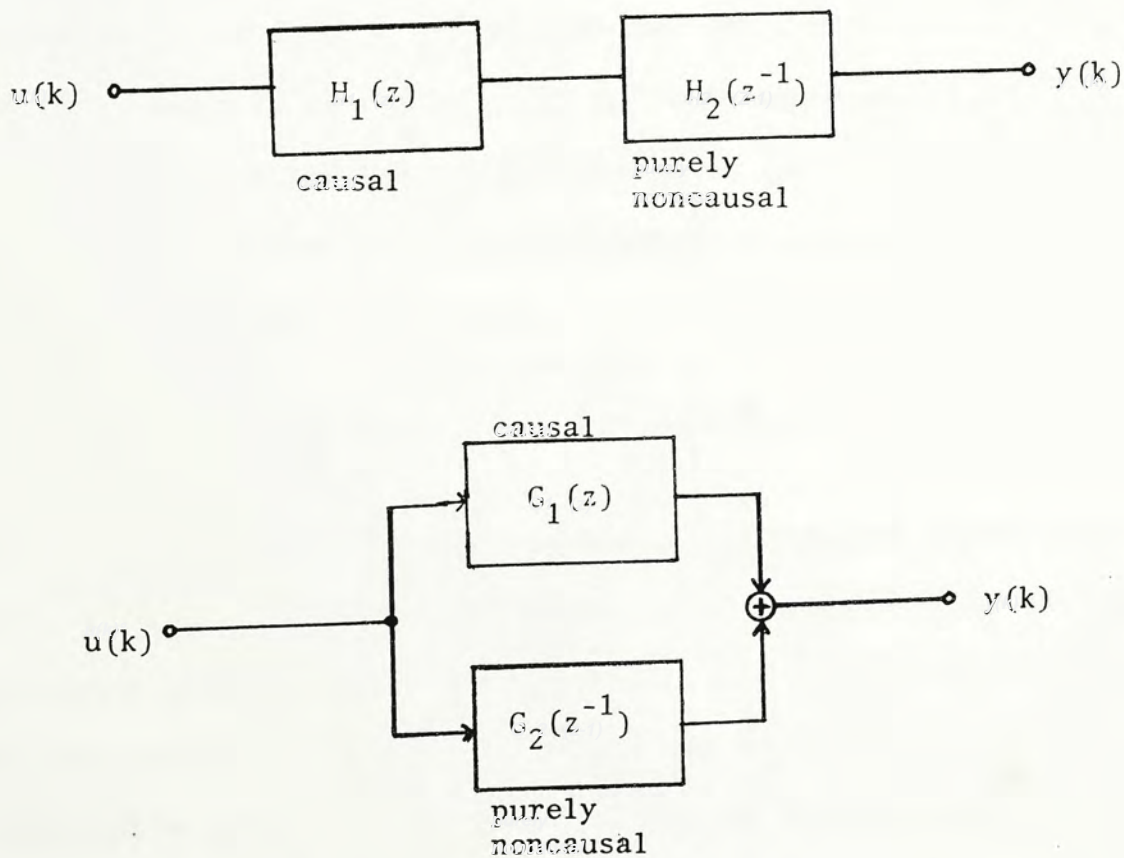

Fig. 2.2 Decomposition of a noncausal system: (a) cascade structure; (b) parallel structure.

18

$H_2(z^{-1})$ and $G_2(z^{-1})$ are purely noncausal systems.

Applying the transformation $z^{-1} \rightarrow z$ to the purely noncausal systems, we obtain $H_2(z)$ and $G_2(z)$ which are both causal. Based on specifications in time or frequency domains, $H_2(z)$ and $G_2(z)$ can then be designed by many existing causal techniques. Finally, the purely noncausal systems $H_2(z^{-1})$ and $G_2(z^{-1})$ are obtained by the inverse transformation $z \rightarrow z^{-1}$. In effect, the decomposition procedure reduces the noncausal design problem into a causal one which is easily solved.

For parallel connection the decomposition is easily obtained by (2.3) in the time domain (Fig. 2.3). $G_1(z)$ and $G_2(z)$ are in nonrecursive forms. If recursive realization is desirable it can be achieved by causal time domain design techniques such as Pade' approximations [B1, Y1] or orthogonal-function approximations [S3, F1]. To obtain the cascade systems, we equate their transfer functions with the parallel recursive ones,

$$H_1(z) \; H_2(z^{-1}) = \overline{G}_1(z) + \overline{G}_2(z^{-1}) \qquad (2.7)$$

It is easy to see that the poles of $H_1(z)$ are identical to that of $\overline{G}_1(z)$ and those of $H_2(z)$ are same as that of $\overline{G}_2(z)$. The zeros are found by solving the numerator polynomials. The assignment of zeros to $H_1(z)$ and $H_2(z^{-1})$ can be done arbitrarily provided that the order of numerators is less than or equal to that of denominators for both transfer functions. If minimal phase for the subfilters is desirable,
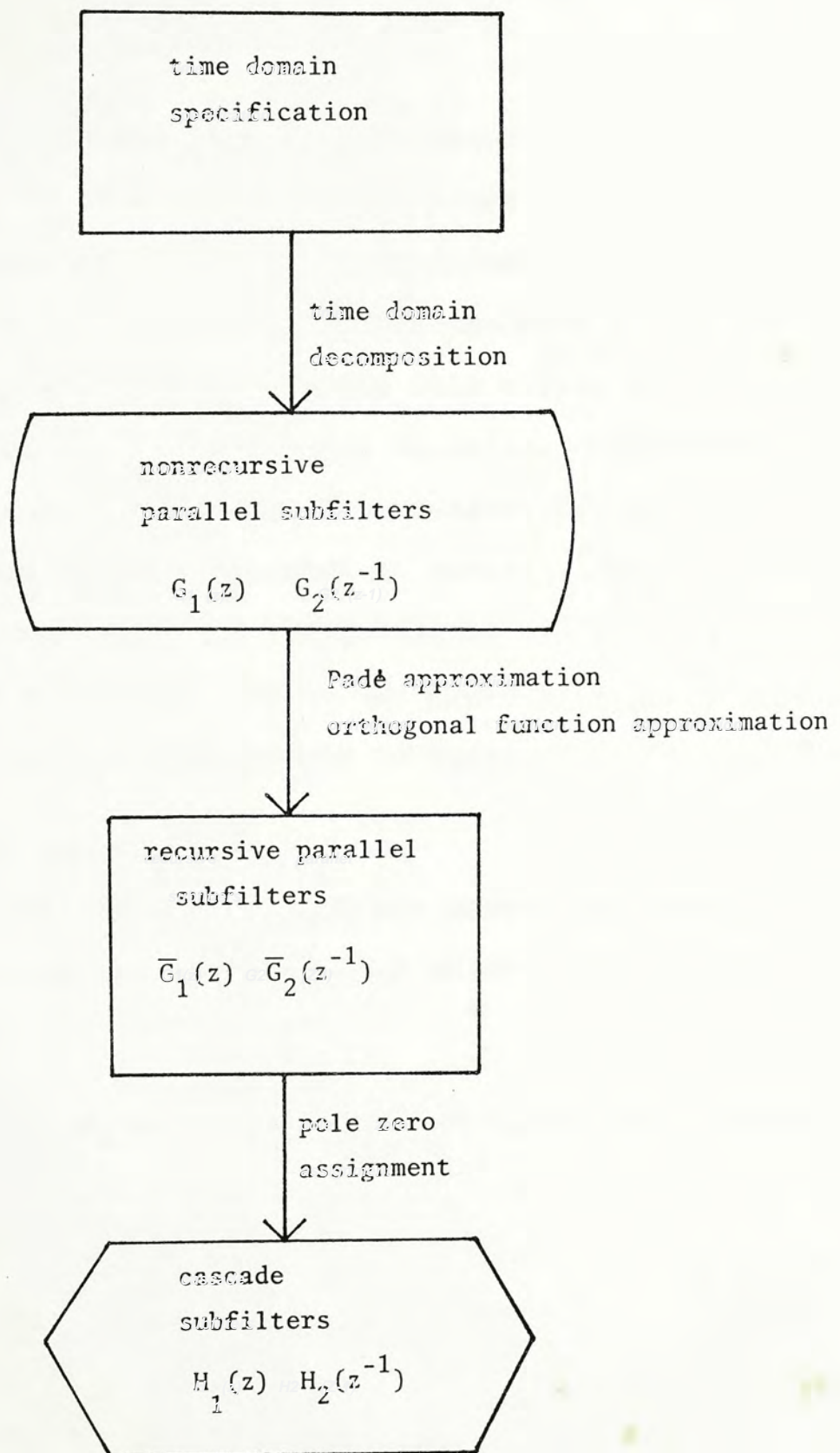
Fig. 2.3 Time domain design

all the zeros inside the unit circle can be assigned to $H_1(z)$ and others to $H_2(z^{-1})$. However, the phase of the overall system is unaffected.

If the noncausal transfer function $H(z)$ is given, the poles and zeros of $H_1(z)$ and $H_2(z^{-1})$ are directly obtained from those of $H(z)$ (Fig. 2.4) since $H(z) = H_1(z) \, H_2(z^{-1})$. The poles inside the unit circle are assigned to $H_1(z)$ while those outside unit circle belong to $H_2(z^{-1})$ because of the stability and causality constraints. Zeros may be assigned in any manner provided that both transfer functions remain recursive in nature. The $\overline{G}_{1,2}(z)$ of the parallel connection is determined by (2.7) using partial fraction expansion. Again the poles of $H_1(z)$ belong to $\overline{G}_1(z)$ while that of $H_2(z)$ belong to $\overline{G}_2(z)$.

2.3 Even and odd sequences

It is well known that for any stable noncausal sequence $h(k)$ it can be separated into an even and an odd parts,

$$h(k) = h_e(k) + h_0(k) \qquad k = 0, \pm 1, \pm 2 \ldots \qquad (2.8)$$

The even sequence $h_e(k)$ is given by

$$h_e(k) = \tfrac{1}{2}[h(k) + h(-k)] \qquad (2.9)$$

and the odd sequences $h_0(k)$,

$$h_0(k) = \tfrac{1}{2}[h(k) - h(-k)] \qquad (2.10)$$

for all integer k.

21
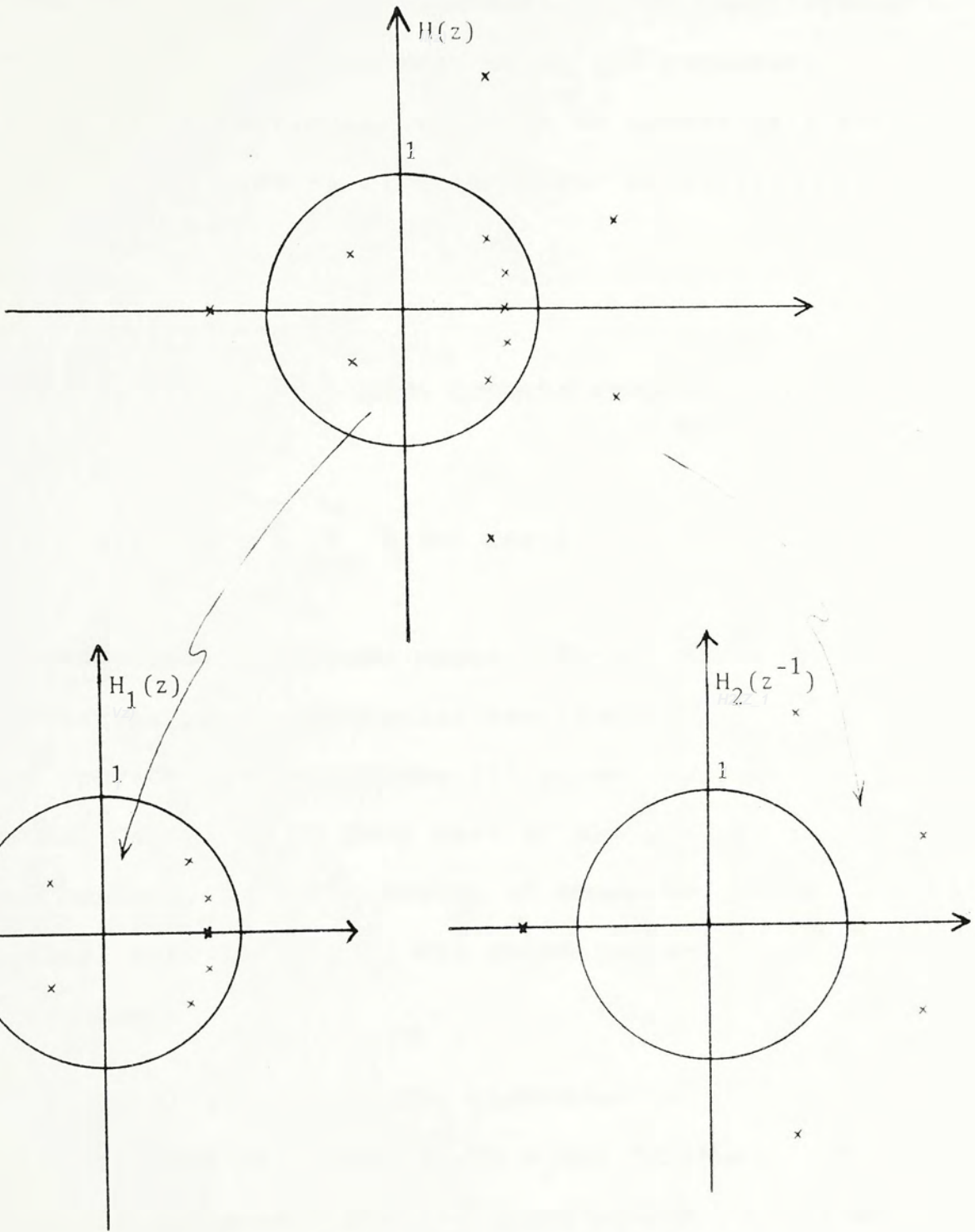
Fig. 2.4 Pole assignment for cascade noncausal system.

As mentioned in the introduction, an ideal system's *filter's ?*

impulse response is either an even or an odd sequence,

therefore, it is of particular interest to investigate the

design of noncausal systems with symmetric or antisymmetric

impulse response.

2.3.1 Even impulse response

For a system with even impulse response, the

frequency is *response.*

$$H(e^{j\omega T}) = 2 \sum_{k=0}^{\infty} h_e(k) \cos k\omega T \qquad (2.11)$$

which is purely real with zero phase. This class of *?*

noncausal systems is of particular importance due to their

ability to provide distortionless filtering needed in many

applications [C4, R3]. In fact most of the past research *previous*

works have emphasized in the design of noncausal filters

with identical subfilters [K1, C1] which guaranty an even *ee*

impulse response.

In parallel connection, according to (2.3), the

subfilter $G_1(z)$ must be identical to $G_2(z)$ in order to give

an even impulse response. The frequency response is given

by

$$H(e^{j\omega T}) = G_1(e^{j\omega T}) + G_1(e^{-j\omega T}) = 2R_e[G_1(e^{j\omega T})]$$

$$(2.12)$$

which is purely real and equal to twice the subfilter's real

part response. The symmetry impulse response can be obtained

by (2.4).

For cascade noncausal filter $H(z)$, the subfilters, $H_{1,2}(z)$ can be obtained from,

$$H(z) = H_1(z) \; H_2(z^{-1}) = G_1(z) + G_1(z^{-1}) \qquad (2.13)$$

Assuming that $G_1(z)$ is in rational form, i.e. $G_1(z) = \frac{N(z)}{D(z)}$, then (2.13) becomes,

$$H_1(z) \; H_2(z^{-1}) = \frac{N(z)}{D(z)} + \frac{N(z^{-1})}{D(z^{-1})}$$

$$= \frac{N(z) \; D(z^{-1}) + N(z^{-1}) \; D(z)}{D(z)D(z^{-1})} \qquad (2.14)$$

where the order of the numerator polynomial $N(z)$ is less than or equal to the order n of the denominator $D(z)$. It is clear from the stability and causality constraints that the poles of $H_1(z)$ and $H_2(z^{-1})$ are identical to the zeros of $D(z)$ and $D(z^{-1})$ respectively. Let $P(z) = N(z) \; D(z^{-1})$ then $P(z^{-1}) = N(z^{-1}) \; D(z)$, thus

$$H(z) = \frac{P(z) + P(z^{-1})}{D(z) \; D(z^{-1})} \qquad (2.15)$$

and $P(z) = P_n z^n + \ldots P_1 z^1 + P_0 + P_{-1} z^{-1} + \ldots + P_{-n} z^{-n}$

$$(2.16)$$

If the order of $N(z)$ is less than $D(z)$, then some of the P_i will be zero. Collecting common terms, the numerator becomes

$$Q(z) = P(z) + P(z^{-1}) = (P_n + P_{-n})z^n + \ldots$$

$$+ (P_1 + P_{-1})z + 2P_0$$

$$+ (P_{-1} + P_1)z^{-1} + \ldots$$

$$+ (P_{-n} + P_n)z^{-n} \qquad (2.17)$$

$Q(z)$ can be rewritten as,

$$Q(z) = q_n z^n + \ldots + q_1 z + q_0 + q_1 z^{-1} + \ldots$$

$$+ q_n z^{-n} \qquad (2.18)$$

which is a polynomial with symmetric coefficients. Applying the transformation, $z \rightarrow z^{-1}$, to (2.18), it is observed that the transformed polynomial is identical to the original polynomial, that is,

$$Q(z) = Q(z^{-1}) \qquad (2.19)$$

Therefore, if z_i is a complex zero of $Q(z)$, then from (2.19) it is concluded that $1/z_i$ is also a zero of $Q(z)$. Since the coefficients q_i are all real, the complex conjugates z_i^* and $1/z_i^*$ must also be complex zeros of $Q(z)$. Then a general elementary factor of $Q(z)$ must be of the form,

$$Q_1(z) = Q_{i_1}(z) \, Q_{i_1}(z^{-1})$$

$$= (z - z_i)(z - z_i^*)(z^{-1} - z_i)(z^{-1} - z_i^*)$$

$$= q_{i_2} z^2 + q_{i_1} z + q_{i_0} + q_{i_1} z^{-1} + q_{i_2} z^{-2} \qquad (2.20)$$

25

which is indeed a polynomial with symmetric coefficients. The complex conjugate and reciprocal pairs of zeros of $Q_i(z)$ is plotted in Fig. 2.5 to show the positional symmetry. If all the zeros of the numerator $Q(z)$ are in reciprocal pairs, then (2.14) becomes

$$H_1(z) \ H_2(z^{-1}) = \frac{Q_1(z) \ Q_1(z^{-1})}{D(z) \ D(z^{-1})} \tag{2.21}$$

which implies $H_1(z) = H_2(z)$. The cascade noncausal filter with even impulse response can therefore be decomposed into identical subfilters when the reciprocal zeros are assigned according to (2.20). In the degenerate case where the complex zeros are on the unit circle (i.e. the reciprocal becomes itself), the cascade filter cannot be decomposed into identical subfilters unless they are all even order zeros. Except for this degenerate case, we can always construct the cascade filter by identical subfilters. If minimal phase subfilter is desirable, all the zeros inside unit circle are assigned to $H_1(z)$.

The impulse response of the cascade noncausal filter with identical subfilters is given by the convolution of the subfilter's impulse responses $h_1(j)$, that is,

$$h(k) = h(-k) = \sum_{j=0}^{\infty} h_1(j) \ h_1(j + k), \text{ for all } k$$

$$\tag{2.22}$$

(2.22) is readily recognized as the autocorrelation of the subfilter's impulse response, which is indeed an even

function. The frequency response is given by,

$$H(e^{j\omega T}) = H_1(e^{j\omega T}) \, H_1(e^{-j\omega T}) = |H_1(e^{j\omega T})|^2 \qquad (2.23)$$

which is real and equal to the square of subfilter's magnitude response. Due to the ease and simplicity in design and realization, only cascade noncausal filter with identical subfilters is considered here.

2.3.2 Odd impulse response

A noncausal system with an odd impulse response has a frequency response of,

$$H(e^{j\omega T}) = 2j \sum_{k=0}^{\infty} h_0(k) \sin k\omega T \qquad (2.24)$$

which is purely imaginary. In order words, the magnitude response may have a 90° or -90° phase shift. This class of noncausal filter is particularly suitable for realizing Hilbert transformation and ideal differentiators which have purely imaginary frequency responses [01].

The parallel connected filter is easily obtained by,

$$H(z) = G_1(z) - G_1(z^{-1}) \qquad (2.25)$$

where the purely noncausal subfilter is simply the inversion of the original subfilter. Its frequency response is

$$H(e^{j\omega T}) = 2j \, I_m[G_1(e^{j\omega T})] \qquad (2.26)$$

which is purely imaginary as expected.

27

For cascade connection, the subfilters are not the same. It can be shown that, however, it is possible to construct a pair of needed subfilters that differs by only a factor of $(z - z^{-1})$.

Consider a cascade noncausal filter $H(z)$ with odd impulse response given by (2.25),

$$H(z) = H_1(z) \ H_2(z^{-1}) = \frac{N(z)}{D(z)} - \frac{N(z^{-1})}{D(z^{-1})}$$

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

$$(2.27)$$

It is clear that the poles of the cascade subfilters $H_1(z)$ and $H_2(z^{-1})$ are identical with the zeros of $D(z)$ and $D(z^{-1})$ respectively. With $P(z) = N(z) \ D(z^{-1})$, we have,

$$H(z) = \frac{P(z) - P(z^{-1})}{D(z) \ D(z^{-1})} \qquad (2.28)$$

where $P(z)$ is in the same form as (2.16).

Collecting common terms, it is eas~~ily~~ $\overset{y}{}$ to see that the numerator polynomial $Q(z)$ has antisymmetric coefficients,

$$Q(z) = P(z) - P(z^{-1})$$

$$= q_n z^n + \ldots + q_1 z - q_1 z^{-1} - \ldots - q_n z^{-n}$$

$$(2.29)$$

where q_0 is identically zero. Applying the transformation,

28

$z \to z^{-1}$, it is observed that $Q(z) = -Q(z^{-1})$. As far as the zeros are concerned, this is similar to (2.19), thus the zeros of the antisymmetric polynomials must also be in both reciprocal and complex conjugate pairs. However, in order to obtain a polynomial of the form in (2.29), the degenerate zeros at ± 1 (its reciprocal and complex conjugate are itself) must also be included. Multiplying a general elementary factor of $Q(z)$ with the factors $(1 - z^{-1})$ and $(1 + z)$, we have

$$(1 - z^{-1}) \, Q_{i_1}(z) \, (1 + z) \, Q_{i_1}(z^{-1})$$

$$= (1 - z^{-1})(z - z_i)(z - z_i^*) \cdot (1 + z)(z^{-1} - z_i)$$
$$\cdot (z^{-1} - z_i^*)$$

$$= q_{i_3} z^3 + q_{i_2} z^2 + q_{i_1} z - q_{i_1} z^{-1} - q_{i_2} z^{-2} - q_{i_3} z^{-3}$$

$$(2.30)$$

which is in the desired form of (2.29). Therefore, except for the case where complex zeros on the unit circle are of odd order, $Q(z)$ can always be decomposed to (Fig. 2.6),

$$Q(z) = (1 - z^{-1}) \, Q_1(z) \, (1 + z) \, Q_1(z^{-1}) \qquad (2.31)$$

which implies that the cascade subfilters are in the form,

$$H(z) = \frac{(1 - z^{-1}) \, Q_1(z)}{D(z)} \cdot \frac{(1 + z) \, Q_1(z^{-1})}{D(z^{-1})}$$

$$= \underbrace{(1 - z^{-1}) \, H_1(z)}_{\text{causal}} \cdot \underbrace{(1 + z) \, H_1(z^{-1})}_{\text{purely noncausal}} \qquad (2.32)$$
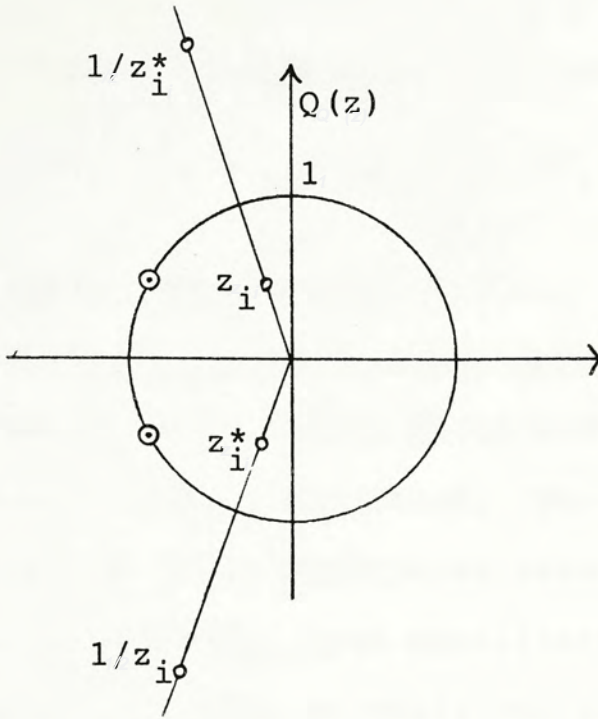
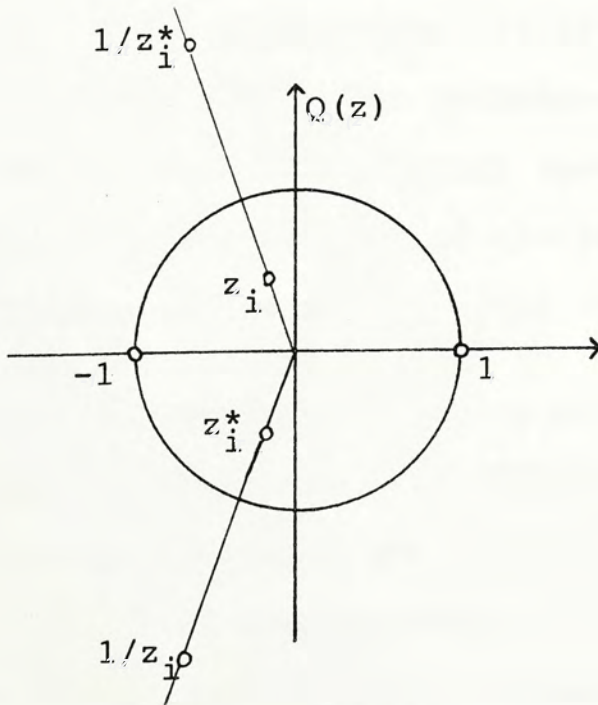Fig. 2.5   Zero positions for noncausal system with even

impulse response

Fig. 2.6   Zero positions for noncausal system with odd

impulse response

Thus the subfilters are identical up to a factor of $(z - z^{-1})$.

The frequency response is then given by

$$H(e^{j\omega T}) = j\ 2\ \sin\ \omega T\ |H_1(e^{j\omega T})|^2 \qquad (2.33)$$

2.4 Practical design considerations

In many filter synthesis problems, the frequency
response requirements are usually given and the filter
transfer function is to be determined. Due to the fact that
a noncausal filter is first decomposed into causal and
purely noncausal subfilters, both subfilters have to be
designed independently based on their own specifications.
Thus, the frequency response requirements for the subfilters
have to be derived from the given overall specifications.
For the general case, where the noncausal filter's impulse
response is asymmetric, there are no relationships between
the causal and noncausal subfilters. It is, therefore,
inherently difficult to derive the individual subfilters'
response requirements from the original specifications.
This design difficulty has prohibited the practical use of
the noncausal filters in the most general form.

However, for noncausal filters with symmetric or
antisymmetric impulse responses, the relationships between
the causal and purely noncausal subfilters have been
derived for both parallel and cascaded connections (section
2.3). The subfilters requirements can thus be easily
obtained from the given filter specifications. Since these
types of noncausal filters also have zero phase

3.

characteristics, they are of particular interest for many practical applications.

In both even and odd impulse response filters, the subfilters for cascade connection are in general easier to design than that of parallel connection. The specified overall frequency requirements become real part (eq. 2.12) or imaginary part requirements (eq. 2.26) for the parallel connected subfilters. Synthesis of digital filters from real part or imaginary part response is uncommon even though it has been developed by Guillemin [G1] for passive analog networks. However, for the case of cascade connected subfilters, the given response requirements become magnitude response specifications (eq. 2.23 and 2.33) while the phase responses are immaterial. Design methods based on magnitude response are well developed (see, for example, [R3, A1]), thus the subfilters are readily constructed using familiar methods. Therefore, for many applications, design simplicity favors the realization of cascade structures rather than parallel ones.

Chapter 3

Conventional Block Processing Technique

While the noncausal filter design problem is
solved by decomposing the filter into causal subfilters
which can then be designed by many well-known methods in
time domain or frequency domain, however the difficulties
in realization remain the main obstacles in many
application areas. In both parallel and cascade
connections, direct implementation of the purely noncausal
subfilter is physically impossible due to its anticipatory
nature. The classical realization approach makes use of
the fact that sequences of finite length can be processed
in reversed order [K1, C1]. For processing infinite input
sequence, the sequence is segmented into finite length
segments for the time reversal processing. The
conventional method is therefore essentially a block
processing approach to noncausal filtering.

The following sections is devoted to describe
and analyse the conventional approach in some detail.

3.1 Time reversal of signals

The purely noncausal subfilter can be rendered
causal by a simple transformation of $z \rightarrow z^{-1}$. The method
was first proposed by Kormylo et al [K1] and further
generalized and improved by Czarnach [C1].

Consider the input/output relation of the purely
noncausal transfer function given by,

$$Y(z) = H(z^{-1}) \ U(z) \qquad\qquad (3.1)$$

Applying the transformation $z \to z^{-1}$ to both sides of (3.1), we obtained,

$$Y(z^{-1}) = H(z) \ U(z^{-1}) \qquad\qquad (3.2)$$

where $H(z)$ is a causal and stable subfilter which is physically realizable. The transformation on the input and output sequences is equivalent to a time reversal of the sequences. By processing the input sequence in reversed order with the causal subfilter $H(z)$, a time reversed output is obtained. The final output $Y(z)$ is then obtained by applying the inverse transformation on $Y(z^{-1})$ which is again a time reversal operation.

The time reversal process is physically possible only if the sequence is of finite length. However, the transformation inherently assumes that sequences are of infinite length. Therefore, it is natural that errors are introduced when infinitely long seqeunces are segmented to finite length sequences in order to perform the time reversal process. The nature of this error is examined in section 3.4.

The implementation of the time reversal operation is most easily done in the computer by reindexing the sequences where no samples are being shifted physically.

3.2 Parallel and cascade realization

As mentioned in chapter 2, the noncausal filter can be decomposed into causal and purely noncausal parts,

34

thus there are two ways of implementation — the parallel and the cascade methods. The implementations of the zero phase noncausal filter for both parallel and cascade connections are shown in Fig. 3.1. The time reversal process is indicated by $k \rightarrow -k$.

If infinite length sequences can be processed without segmentation, thán the parallel and cascade realizations are equivalent as far as the ouptut is concerned. In practice, only truncated finite length sequences can be processed and the two realization methods produce different output errors. In actual applications we are also interested in the difference in other aspects of the realization methods, such as memory size requirement and computation efficiency. Moreover, the design methodology may also dictate the use of the parallel or cascade structure since, as noted in chapter 2, frequency domain design favors cascade form while time domain specifications naturally lead to parallel form. This fact is reflected by the dominance of cascade realization in previous works [K1, C1].

Without rigorous proofs, the general differences in memory size requirement and computation efficiency for both realization methods are established in the following example. 1 example. to est. the general property?

An noncausal filter consisting of identical subfilters, $H(z) = \frac{z}{z - 0.4}$, is realized by both methods.
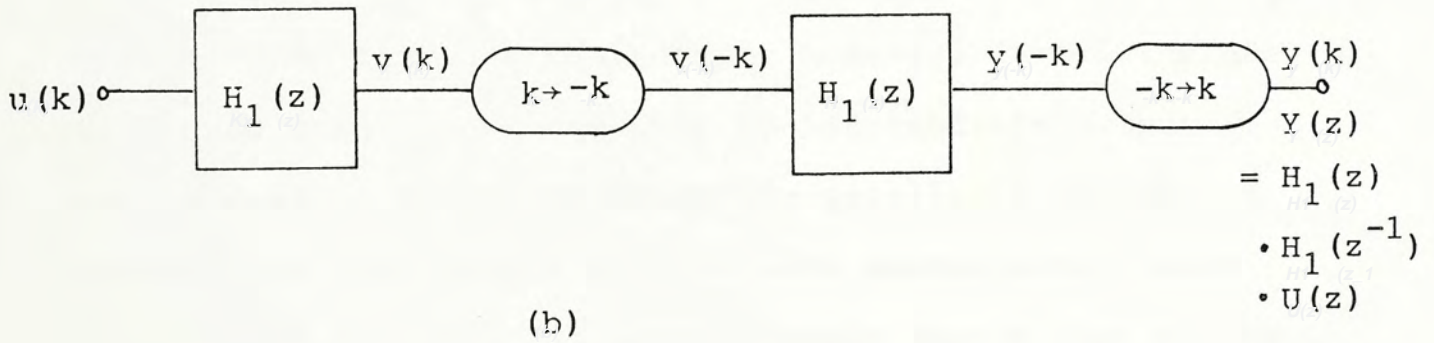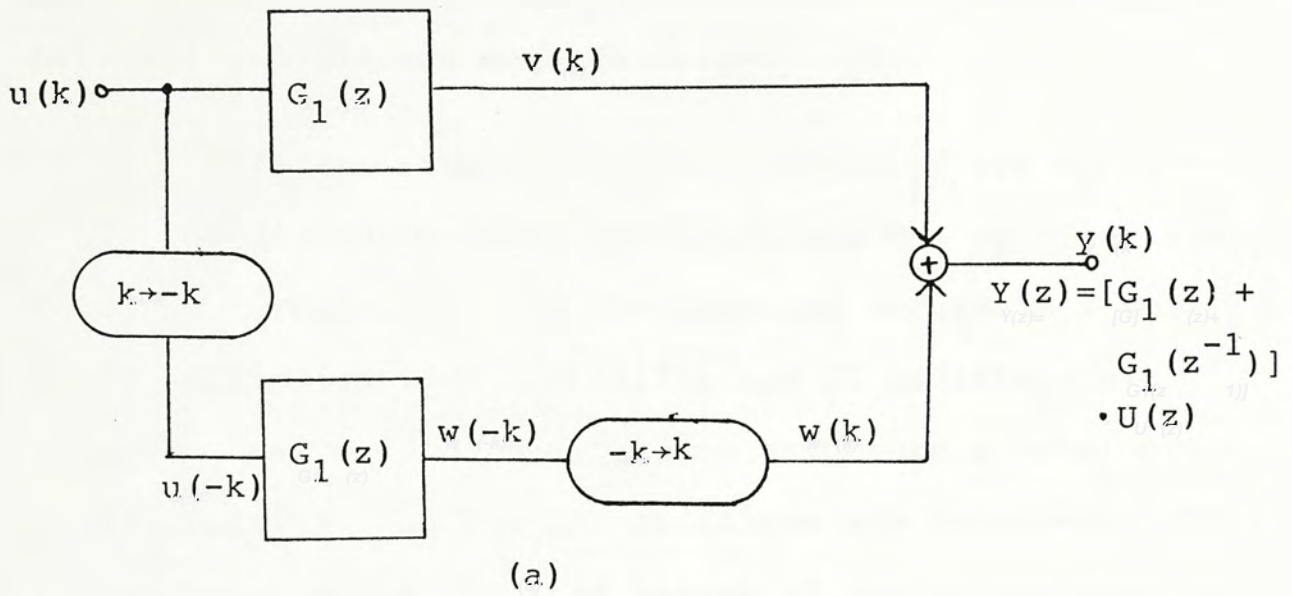
35

Fig. 3.1 Noncausal recursive filters realized by block processing approach: (a) parallel sturcutre (b) cascade structure.

36

A finite length input sequence, u(k) = 1 for $0 \leq k \leq 4$ is being processed; results for output sequence y(k) in the range $-4 \leq k \leq 10$ are shown in Figs. 3.2 and 3.3 for the parallel and cascade methods respectively.

To generate each output sample of the subfilter H(z), one multiplication, one shift and one addition are required. Therefore, for the parallel structure a total of 20 multiplications, 20 shifts and 25 additions are needed. However, for the cascade structure a total of 26 multiplications, shifts and additions are necessary. Thus, in general, if the required length of ouptut sequence is longer than the input sequence the parallel structure needs less mathematical operations.

The storage requirement, however, favors the use of cascade structure since only 11 intermediate output samples need to be stored while for parallel structure 20 intermediate samples are stored. The needed memory size for parallel structure is approximately double that of the cascade realization.

However, in many cases, the required ouptut sequence is usually of the same length L as the input sequence. Then, in this example, the cascade realization requires 2L multiplications, additions and shift operations; the memory size needed for intermediate result storage is L. For parallel realization, again 2L multiplications and shift operations are needed but the number of additions increased to 3L. The storage requirement is 2L, which is
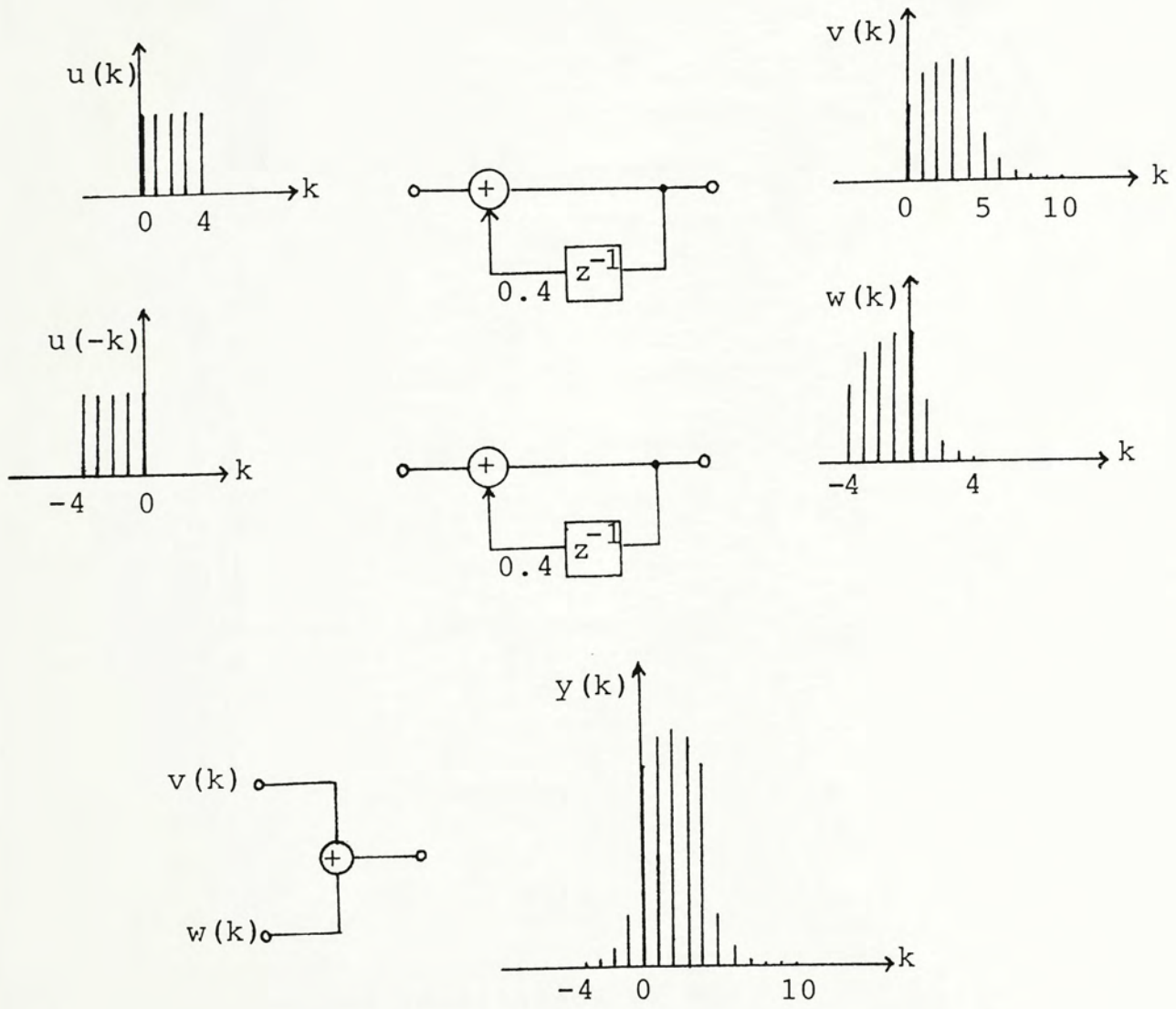
37

Fig. 3.2 Parallel realization of zero phase
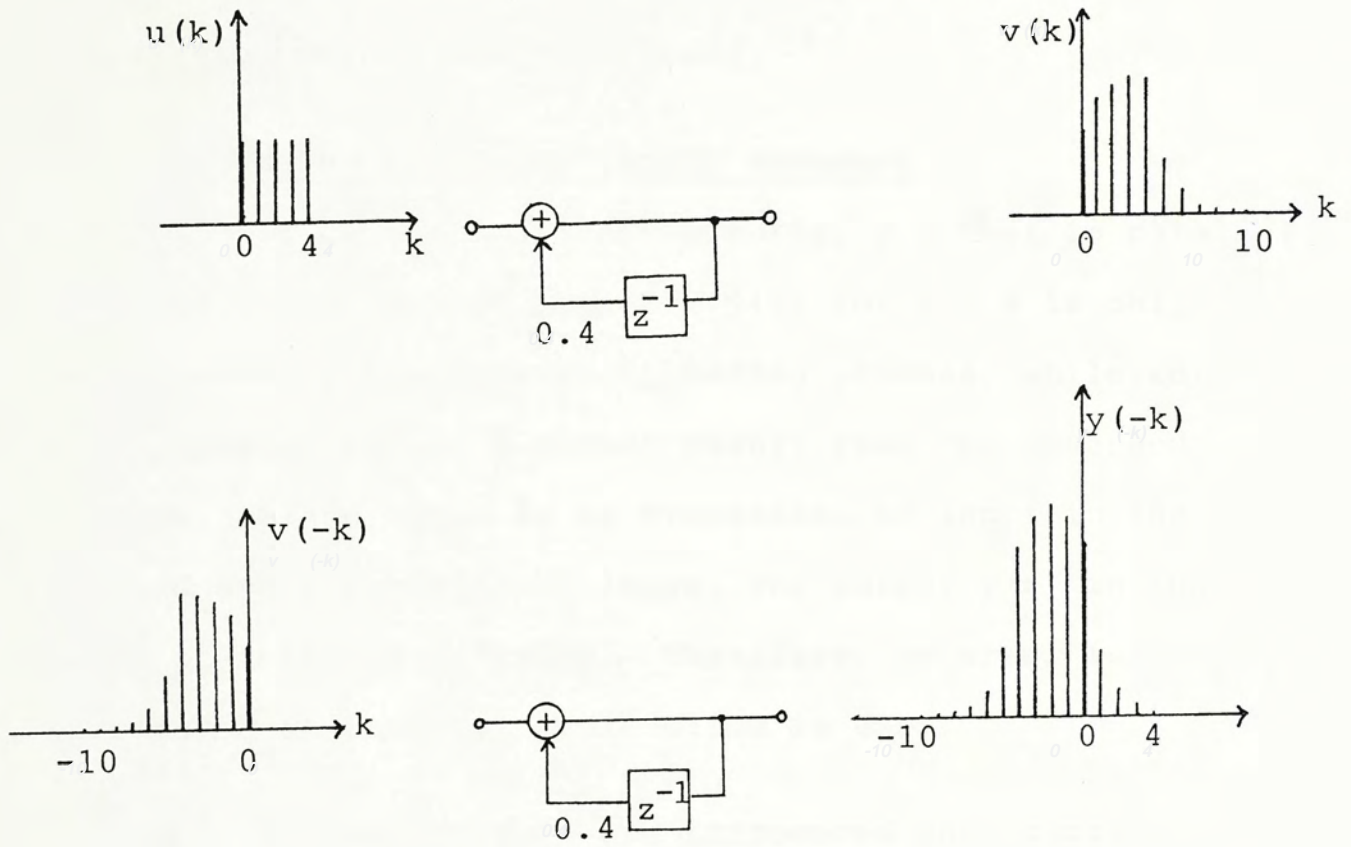noncausal filter for finite length
sequence

Fig. 3.3 Cascade realization of zero phase noncausal filter

for finite length sequence

twice that needed by the cascade structure. Therefore, it is more economical in both memory size requirement and number of computations to realize a cascade structure.

In the next section, the errors for both realization method are considered.

3.3 Processing of finite length sequence

It can be observed form Fig. 3.2 that in parallel realization the output sequence $y(k)$ for $k > 4$ is only contributed by the forward filtering process, while for $k < 0$, the output is a direct result from the reversed process. Since there is no truncation of input in the forward and reversed subfilters, the output $y(k)$ in the range of interest is exact. Therefore, no error is introduced if parallel realization is used.

However, errors are introduced when cascade realization is used because the forward subfilter's output is truncated at a certain time k for the time reversal operation and processing by the reversed subfilter (Fig. 3.3). Czarnach [C1] proposed a state space method to eliminate this truncation error.

Let the state space representations of the forward and reversed subfilter be given by (3.3) and (3.4) respectively.

$$x_1(k + 1) = A_1 x_1(k) + b_1 u(k)$$

$$v(k) = c_1 x_1(k) + d_1 u(k) \qquad (3.3)$$

$$x_2(k + 1) = A_2 x_2(k) + b_2 v(-k)$$

$$y(-k) = c_2 x_2(k) + d_2 v(-k) \qquad (3.4)$$

where $x_{1,2}(k)$ are state vectors and $A_{1,2}$, $b_{1,2}$, $c_{1,2}$, $d_{1,2}$ are the system matrices.

The finite length input sequence $u(k)$ is zero for $k < 0$ and $k > k_0$. Suppose the output for the forward subfilter is truncated at $k = k_1 \geq k_0$ then for any $k > k_1$, the output $v(k)$ of this subfilter is uniquely given by:

$$v(k) = c_1 A_1^{k-k_1-1} x_1(k_1 + 1) \qquad \text{for } k > k_1 \qquad (3.5)$$

Reversing $v(k)$ and applying to the reversed subfilter, we get the state vector $x_2(k)$ for $k \leq -k_1$

$$x_2(k) = \sum_{\lambda=0}^{\infty} A_2^\lambda b_2 c_1 A_1^{-k-k_1+\lambda} x_1(k_1 + 1) \qquad (3.6)$$

Since the output of the forward system is truncated at k_1, the state vector $x_2(-k_1)$ given by (3.6) contains all the loss information due to the truncation. Thus, we obtained from (3.6)

$$x_2(-k_1) = \left\{ \sum_{\lambda=0}^{\infty} A_2^\lambda b_2 c_1 A_1^\lambda \right\} x_1(k_1 + 1) \qquad (3.7)$$

The matrix summation in (3.7) is a fixed matrix depending only on the subfilters. Thus the loss information due to truncation is passed on to the initial state of the reversed subfilter from the final state of the forward subfilter. If the initial state of the reversed subfilter is chosen according to (3.7), no error is introduced and

the output $y(k)$ in the range of interest is exact.

Even though the matrix $T = \sum_{\lambda=0}^{\infty} A_2^{\lambda} b_2 c_1 A_1^{\lambda}$ is determined once the subfilters are fixed, however, to compute the initial state $x_2(-k_1)$ of the reversed subfilter from (3.7), n^2 multiplications and additions are needed if the subfilters' order is n. The computation cost for elimination of truncation error is therefore expensive ~high. for high order subfilters.

In summary, the output $y(k)$ are exact for both parallel and cascade realizations provided that (3.7) is implementated into the cascade structure. If the recursive subfilters are of order n and canonical form realization is used, then each output sample requires 2n+1 multiplications and additions.

When the length of both input and output sequences is L, then $(4n + 2)L$ multiplications and $(4n + 3)L$ additions are required for parallel realization; while, for cascade realization with error elimination, $(4n + 2)L + n^2$ multiplications and additions are needed. In most applications L is much larger than n (e.g. typical values are n = 8, L = 1000), therefore, the numbers of computations are similar for both realizations. Finally, it should be mentioned that the memory size for parallel realization is 2L which is double that of cascade method while the processing time is similar.

3.4 Processing of infinite length sequence

In most filtering problems, the input sequence is of infinite length. The time reversal procedure of the conventional technique requires the segmentation of the input sequence into finite length ones. Segmentation necessarily introduces transient errors since when a segment is being processed the effect of future segments are not known. To reduce the transient effects, similar overlap-save algorithms were proposed by Kormylo et al [K1] and Czarnach [C1]. The main idea is to overlap the segments to be processed by a certain amount and discard the initial and/or the end transients.

For cascade realization the overlap-save algorithm is depicted as shown in Fig. 3.4. The i th segment $u_i(k)$ overlaps with $u_{i-1}(k)$ and $u_{i+1}(k)$ by an amount $q = q_1 + q_2$ at both ends. The sequence is then processed in the forward and reversed direction. The output sequence $y_i(k)$ is formed by discarding q_1 samples at the beginning and q_2 sample at the end. The error depends on the overlapping length in relation to the length of the main energy portion of the impulse response $h(k)$. For the case, $q = q_2$, Czarnach [C1] has shown that the maximum error is given by

$$|e_{max}| = M \sum_{k=-\infty}^{-q-1} |h(k)| \qquad (3.8)$$

where M is the upper bound of input $|u(k)|$. For a certain accuracy requirement, q can be determined by (3.8). In
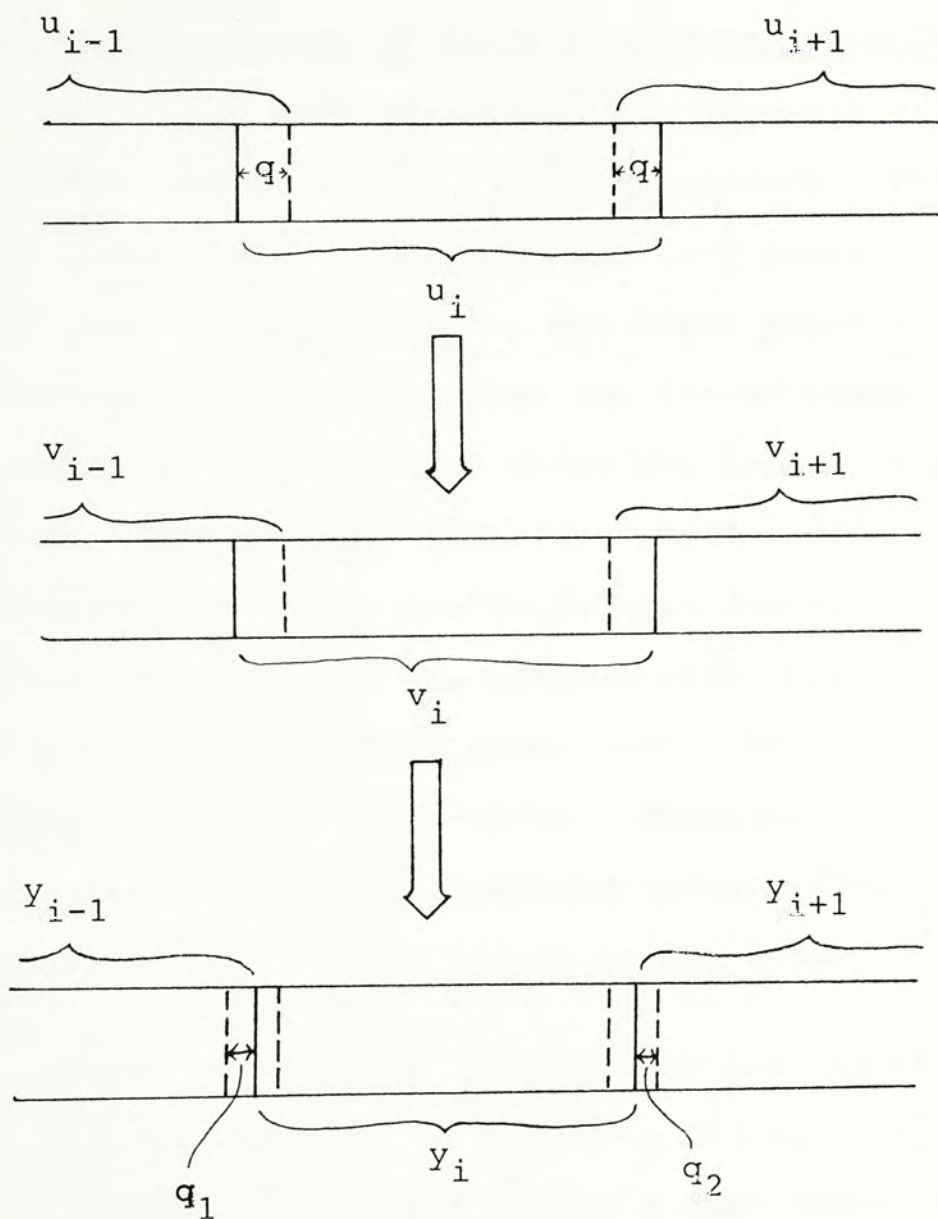
Fig. 3.4   Overlap-save algorithm for processing infinite
length sequence

practice the segment length is usually chosen to be several times that of q.

3.5 Main characteristics of block processing approach

One of the main advantages of noncausal filtering is that it can provide almost ideal performance. The most useful and widely known example is the zero phase filters. For finite length input sequence, the block processing technique gives exactly zero phase (or linear phase due to finite processing time) results which are desirable for many applications. However, for infinite length sequence, both phase and magnitude errors are introduced due to segmentation. Even though the maximum error bound for each output sample can be computed from (3.8), there is no similar formula for the phase error. Therefore, one of the disadvantages of block processing technique is that the phase error due to segmentation is not known.

Comparing with the classical optimal linear phase FIR filters, the block processing noncausal filters is computationally more efficient for a wide range of filter responses [C1]. This is due to the fact that all filtering is done recursively rather than nonrecursively in a block processing noncausal filter. Since it is well known that optimum FIR filters in general require less computation than similar phase equalized causal recursive filters [R1], the block processing noncausal filters are the most computational efficient to provide linear phase characteristics in many applications.

45

However, the block processing approach requires large amount of memory (in the order of thousands) to store the intermediate results. This also implies that the processing delay time is extremely long. If there are L samples in each segment, the delay time is at least 2LT if the processing time for each sample is *assumed to be* approximately one sampling time T. Comparing with similar linear phase causal filters, this delay time is very much longer.

Chapter 4

Sample-by-sample approach using all zero approximation

The conventional block processing technique has three main disadvantages, namely, very long group delay time, huge memory size requirement and unknown phase distortion due to segmentation. In many applications where fast processing is mandatory, the conventional approach to noncausal processing cannot be applied.

A sample-by-sample approach to noncausal filtering is described in this chapter. The new approach has the advantages of short basic group delay time, small memory size requirement and exactly known phase response. Therefore, noncausal processing can now be applied to a wider application areas.

4.1 Nonrecursive approximation

The main difficulty in the realization of noncausal filters arises chiefly from the purely noncausal part $H_1(z^{-1})$. Without the limitations of the block processing approach, a nonrecursive filter of length N can be used to approximate $H_1(z^{-1})$. The nonrecursive approximation can be rendered causal by adding sufficient delays.

Consider an N-1th order all zero approximation $F(z)$ to the stable and causal filter $H_1(z)$. In general, $F(z)$ is of the form

$$F(z) = \sum_{i=0}^{N-1} f(i)\ z^{-i} \qquad (4.1)$$

47

with the frequency response,

$$F(e^{j\omega T}) \simeq H_1(e^{j\omega T}) \tag{4.2}$$

The purely noncausal transfer function is obtained by applying the transformation $z \rightarrow z^{-1}$ on (4.1),

$$F(z^{-1}) = \sum_{i=0}^{N-1} f(i) z^i \tag{4.3}$$

Adding N-1 unit delays, the FIR transfer function becomes causal, that is,

$$\overline{F}(z^{-1}) = z^{-(N-1)} F(z^{-1}) = \sum_{i=0}^{N-1} f(i) z^{i-N+1} \tag{4.4}$$

with $\quad |\overline{F}(e^{-j\omega T})| = |F(e^{-j\omega T})| \cong |H_1(e^{-j\omega T})|$

Since the phase of $H_1(e^{j\omega T})$ is nonlinear, the coefficients of FIR filter have no symmetric constraint .

For cascade realization, the delayed causal FIR subfilter is cascaded with the causal IIR subfilter. The overall transfer function becomes (Fig. 4.1a)

$$H(z) = \overline{F}(z^{-1}) H_1(z) \tag{4.5}$$

For parallel realization, however, delays are added to both the causal and purely noncausal subfilters (Fig. 4.1b), that is,

$$H(z) = z^{-(N-1)} (G_1(z) + F(z^{-1}))$$

$$= z^{-(N-1)} G_1(z) + \overline{F}(z^{-1}) \tag{4.6}$$

48
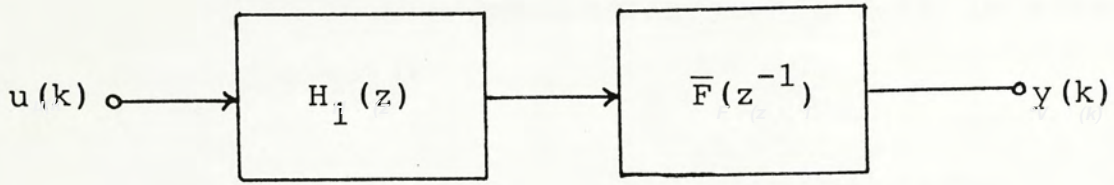
(a)

(b)

Fig. 4.1 Noncausal filter realization by FIR and
IIR subfilters: (a) cascade structure;
(b) parallel structure.

where F(z) is the all zero approximation to $G_1(z)$.
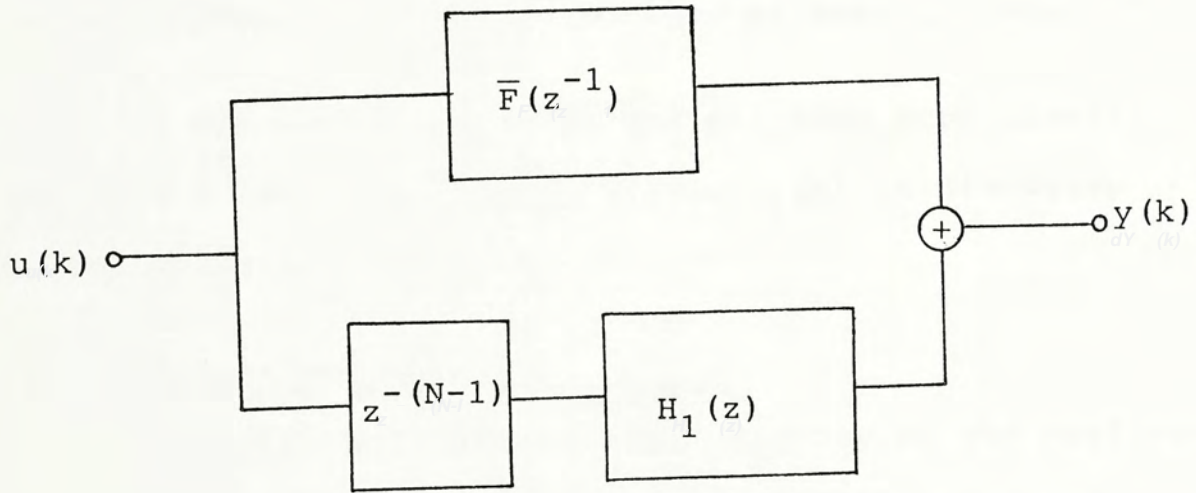Therefore, some extra delay is added also to the causal
subfilter $G_1(z)$. The overall response has a magnitude
approximates that of the original response with an added
linear phase of $-(N-1)\omega T$.

Now H(z) is both causal and stable, thus can be
implemented by many well established implementation
techniques [R3]. The implementation does not require any
segmentation or time reversal of signals and therefore is
essentially a sample-by-sample processing of signals. The
frequency response $H(e^{j\omega T})$ can also be exactly computed.

The construction of the all zero approximation
F(z) from a causal recursive filter $H_1(z)$ is discussed in
the following sections.

4.2 Real part sufficiency

In the literature, vast majority of the previous
research works on FIR digital filter design have been is
concentrated on linear phase FIR filters owning to their
superior performance in providing distortionless processing
[R2, L2, P1]. Most design methods are based on the
assumption that the filter coefficients are symmetrical (or   how do you show this
antisymmetrical) which necessarily implies that the resulting
FIR filter is linear phase. However, several papers have
proposed design methods for particular classes of nonlinear
phase FIR filter. The design of minimum phase FIR filters
is first considered by Herrman et al. [H3]. While in the

50

application areas of phase equalization and chirp processing, all pass FIR phase network design has been proposed by Steiglitz [S1]. Moreover, Goldberg et al. [G3] have investigated methods for designing optimal nonlinear phase FIR filters which can be implemented efficiently. For the general case of designing FIR filters to approximate both the magnitude and phase responses, methods have been proposed using optimization and iterative techniques [C8, H2]. Even though, we can construct the all zero approximation $F(z)$ to the recursive filter $H_1(z)$ using the general optimization or iterative techniques which require long computation time, a more efficient and flexible method is proposed. Before we discuss the details of the new design method, we would like to establish some properties of digital systems concerning the real part of a digital transfer function on unit circle. These properties are used to accomplish the design of similar (in term of frequency response) IIR and FIR digital filters.

In this classical work on passive network, Guillemin [G1] has shown that the j-axis real part of a transfer function is sufficient for the construction of the system. It is further shown that a continuous system can be designed based on real part function specifications rather than the conventional magnitude and phase specifications [G1]. For a causal, stable and linear time invariant (LTI) digital system, similar conclusions can be made on the transfer function's real part on unit circle.

Consider a digital system function given by

51

$$H(z) = \sum_{n=-\infty}^{\infty} h(nT) \, z^{-n} \qquad (4.7)$$

where $h(nT)$ is a real sequence defined for $n = 0, \pm 1, \pm 2, \ldots$
and the sampling time, $T > 0$.
On the unit circle, $z = e^{j\omega T}$, we have

$$H(z^{-1}) = H^*(z), \qquad (4.8)$$

where $H^*(z)$ is the complex conjugate of $H(z)$.

Thus, the real part of the system function on the unit

circle is given by,

$$H_R(\omega T) = \frac{1}{2}[H(z) + H(z^{-1})] \Big|_{z \, = \, e^{j\omega T}} \qquad (4.9)$$

and the imaginary part,

$$H_I(\omega T) = \frac{1}{2j}[H(z) - H(z^{-1})] \Big|_{z \, = \, e^{j\omega T}} \qquad (4.10)$$

It is noted that both H_R and H_I are real function of the

real variable ωT.

The impulse response of the system is given by
[01],

$$h(nT) = \frac{1}{2\pi j} \oint_C H(z) \, z^{n-1} \, dz \qquad (4.11)$$

where C is a counter-clockwise closed curve in the region

of convergence of $H(z)$. For a stable and causal LTI system,

the region of convergence is $|z| \geq 1$. Choosing C to be the

unit circle, $z = e^{j\omega T}$, we have

$$h(nT) = \frac{T}{2\pi} \int_{-\pi/T}^{\pi/T} H(e^{j\omega T}) \; e^{jn\omega T} \; d\omega \qquad (4.12)$$

Writing $H(e^{j\omega T}) = H_R(\omega T) + j \; H_I(\omega T)$ and noting that $H_R(\omega T)$ is an even function of ωT while $H_I(\omega T)$ is an odd function of ωT, we have,

$$h(nT) = \frac{T}{\pi} \int_0^{\pi/T} H_R(\omega T) \; \cos \; (n\omega T) \; d\omega$$

$$- \frac{T}{\pi} \int_0^{\pi/T} H_I(\omega T) \; \sin \; (n\omega T) \; d\omega \qquad (4.13)$$

Replacing n by −n into the integrals of (4.13), we observe that the first integral of (4.13) is an even function of n while the second integral is an odd function of n.

For a causal system, $f(nT) = 0$ for negative integer n. Thus, for n = −1, −2, −3, ..., we have

$$\int_0^{\pi/T} H_R(\omega T) \; \cos \; (n\omega T) \; d\omega = \int_0^{\pi/T} H_I(\omega T) \; \sin \; (n\omega T) \; d\omega$$

$$(4.14)$$

Then, for n = 1, 2, 3, ..., we obtain

$$\int_0^{\pi/T} H_R(\omega T) \; \cos \; (n\omega T) \; d\omega = - \int_0^{\pi/T} H_I(\omega T) \; \sin \; (n\omega T) \; d\omega$$

$$(4.15)$$

Thus, (4.13) becomes

$$h(0) = \frac{T}{\pi} \int_0^{\pi/T} H_R(\omega T) \, d\omega \, ,$$

and

$$h(nT) = \frac{2T}{\pi} \int_0^{\pi/T} H_R(\omega T) \cos(n\omega T) \, d\omega \, , \qquad (4.16)$$

or

$$h(nT) = \frac{-2T}{\pi} \int_0^{\pi/T} H_R(\omega T) \sin(n\omega T) \, d\omega \, ,$$

$$n = 1, 2, 3, \ldots, \qquad (4.17)$$

which shows that the real part alone suffices to determine the impulse response of the causal digital system [01]. Unlike the analog case it is noted that the right side of (4.16) is a discrete time function while the left side is a continuous function.

The digital system transfer function is obtained by applying the z-transform to (4.16),

$$H(z) = \frac{T}{\pi} \int_0^{\pi/T} H_R(\omega T) \, d\omega$$

$$+ \frac{2T}{\pi} \sum_{n=1}^{\infty} z^{-n} \int_0^{\pi/T} H_R(\omega T) \cos(n\omega T) \, d\omega \qquad (4.18)$$

Equation (4.18) essentially shows that we can uniquely construct the digital system from a given real part function alone if the LTI system is stable and causal. Moreover, the imaginary part is given by the well-known Hilbert transform [01]. Therefore, by relating the real part functions (on j-axis for analog systems or on unit circle for digital system) of two transfer functions, good approximation in both magnitude and phase response is

obtained. In fact, design methods have been developed based
on unit circle real part functions in rational trigonometric
forms [V1, D1]. If the real part function is given in
other functional forms or graphical forms, (4.16) can be
computed numerically to give a nonrecursive filter. Since
(4.16) is the integral of a real function over a real
variable, it is more efficient to evaluate on computer. If
a recursive filter is desirable it can be obtained from the
nonrecursive approximation by many well known methods
[B1, Y1].

The unit circle real part function can also be
used to perform the inverse z transformation which is
conventionally obtained by the contour integral method, power
expansion method and partial-fraction expansion method [C7].
While the last method is only suitable for rational system
function, the first two methods can be applied to general
cases. However, both methods involve contour integrals
which are difficult to solve. In such cases, the real
part function can first be computed,

$$H_R(\omega T) = \text{Re}[H(e^{j\omega T})]$$ (4.19)

and the inverse transform computed by (4.16) numerically.

Another important application of the real part
function is in the area of digital system modelling. The
transform integral (4.16) is particular easy to apply when
the real part response of the system is given either in
graphical form or from frequency measurement data.

In the next section, the real part sufficency property is ultilized to construct the nonrecursive approximation for the purely noncausal subfilter.

4.3 Wiener-Lee decomposition

Since one of the most popular methods of designing IIR filter (i.e. the causal subfilter) is the bilinear transformation technique, this section derives a similar technique for designing FIR filters. If the same analog filter is used for both the IIR and FIR filters, their responses will be approximately the same satisfying the condition (4.2) and thus suitable for realizing noncausal filters.

The design of FIR filter based on the bilinear transformation of an analog filter function is first considered by Sallai [S2]. The analog filter function is of the special form,

$$K(s) = \frac{\sum\limits_{\ell=0}^{R} a_\ell s^\ell}{(s + 1)^M} \quad , \qquad R < M \qquad (4.20)$$

Transforming by, $s = \dfrac{1 - z^{-1}}{1 + z^{-1}}$, (4.20) becomes

$$K(s) = \sum_{\ell=0}^{R} a_\ell (1 - z^{-1})^\ell (1 + z^{-1})^{M-\ell}, \quad R < M, \qquad (4.21)$$

which is a FIR digital filter transfer function. However, (4.20) is not a typical analog transfer function therefore many well known analog filters cannot be transformed into

FIR digital filters. The difficulty can be solved by decomposing the continuous filter transfer function into Wiener-Lee decomposed form which under the bilinear transformation becomes a nonrecursive digital filter.

Consider a general analog filter function of the form,

$$H(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + \ldots + b_1 s + b_0}{a_p s^p + a_{p-1} s^{p-1} + \ldots + a_1 s + a_0} \; , \; p \geq m$$

(4.22)

Equation (4.22) can be decomposed into the Wiener-Lee form [L1, C2]

$$G(s) = \frac{g_0}{2} + \sum_{n=1}^{\infty} g_n \left(\frac{1-s}{1+s}\right)^n$$

(4.23)

Lee [L1] has shown that by equating the real parts of (4.22) and (4.23) along the j-axis, the coefficients g_n are given by the familiar Fourier cosine transform. On the j-axis the value of a general term in (4.23) is expressed as

$$\left(\frac{1-s}{1+s}\right)^n \bigg|_{s = j\bar{\omega}} = e^{-jn\phi}$$

(4.24)

with unity magnitude and a phase of $-n\phi$. Equating the phase, the relation between ϕ and the analog frequency $\bar{\omega}$ is

$$\phi = 2 \tan^{-1} \bar{\omega} \qquad \text{or} \qquad \bar{\omega} = \tan \frac{\phi}{2}$$

(4.25)

In the ϕ domain, the value of (4.23) along the imaginary

axis becomes,

$$G(s)\Big|_{s = j \tan \frac{\phi}{2}} = \frac{g_0}{2} + \sum_{n=1}^{\infty} g_n \exp(-jn\phi) \qquad (4.26)$$

Matching the j-axis real parts of (4.22) and (4.26) in the

ϕ domain, we obtain the familiar Fourier cosine series

expansion,

$$\overline{H}_R(\phi) = H_R(j \tan \frac{\phi}{2}) = \frac{g_0}{2} + \sum_{n=1}^{\infty} g_n \cos n\phi \qquad (4.27)$$

where the subscript R denotes the real part of the function.

The coefficients are then given by,

$$g_n = \frac{1}{\pi} \int_0^{2\pi} \overline{H}_R(\phi) \cos n\phi \, d\phi \,, \quad n = 0, 1, 2, \ldots, \qquad (4.28)$$

From (4.28), it is clear that g_n are always real.

Since the real part alone is suffice to determine

the transfer function (section 4.2) and that it satisfies

Dirichlet conditions, F(s) converges to H(s) for $n \to \infty$. The

nonrecursive filter is then obtained by the familiar bilinear

transformation of (4.23),

$$\overline{G}(z) = G(s)\Big|_{s = \frac{1 - z^{-1}}{1 + z^{-1}}} = \frac{g_0}{2} + \sum_{n=1}^{\infty} g_n z^{-n} \qquad (4.29)$$

with digital frequency response

$$\overline{G}(e^{j\omega T}) = G(j\overline{\omega})\Big|_{\overline{\omega} = \tan \frac{\omega T}{2}} \qquad -\frac{\omega}{T} < \omega \leq \frac{\pi}{T} \qquad (4.30)$$

The frequency response is the same as that of an IIR filter $H_1(z)$ obtained from the bilinear transformation of the same analog filter $H(s)$, that is,

$$\overline{G}(e^{j\omega T}) = H_1(e^{j\omega T}) \qquad (4.31)$$

where $\qquad H_1(z) = H(s)\Big|_{s = \dfrac{1 - z^{-1}}{1 + z^{-1}}}$

It should be noted that the matching of j-axis real parts of (4.27) is equivalent to the matching of the real parts of $H_1(z)$ and $\overline{G}(z)$ on the unit circle, $z = e^{j\omega T}$. Therefore, the real part sufficiency of digital system is analogous to that of analog system as shown in section 4.2.

Finally, the FIR filter is obtained by truncation using a window function $W(n)$,

$$f(n) = g_n W(n) , \qquad n = 1, 2, \ldots, N$$

and $\qquad f(0) = \dfrac{g_0}{2} W(0) \qquad (4.32)$

It is necessary to investigate the fall-of rate of g_n in order to determine the FIR filter order N. Since all well known analog filter functions have well-behave j-axis real part that satisfies Dirichlet conditions, the coefficents fall off at least as rapidly as $1/n$ [M1]. In most cases the fall-off rate is faster since the real part response may have high order derivatives that satisfy Dirichlet conditions. Therefore, very good approximations to $H_1(e^{j\omega T})$ can be achieved by relatively low order FIR

transfer functions independent of the type of window functions used.

The coefficients can be computed numerically on a computer by (4.28) or by expanding (4.31) into a Laurent series about z = 0 [K2]. While (4.28) is suitable for a general class of functions (such as functions given in graphical forms), the latter method is particular easy to carry out if rational function is given.

By making use of the Wiener-Lee decomposition, the well-known bilinear transformation technique of IIR filter design is extended to FIR filter design. It is now possible to design the FIR and IIR subfilters of a noncausal filter from an analog filter using bilinear transformation.

4.4 Design procedures

A noncausal cascade filter suitable for sample by sample processing has the transfer function of (4.5) which is repeated here,

$$H(z) = \overline{F}(z^{-1}) \; H_1(z) \qquad\qquad (4.33)$$

Neglecting the linear phase shift, the real frequency response is approximately given by

$$H(e^{j\omega T}) \cong |H_1(e^{j\omega T})|^2 \qquad\qquad (4.34)$$

Once a magnitude specification in the frequency domain and the group delay ripple requirement are specified, a step to step approach is taken to design both the IIR subfilter

60

$H_1(z)$ and the FIR subfilter $F(z)$ based on the bilinear transformation technique.

The design procedure consists of six major steps (Fig. 4.2):

(1) obtain an analog magnitude specification by halfing the original log magnitude scale and prewarping the digital frequencies into analog frequencies by,

$$\overline{\omega}_x = \tan \frac{\omega_x T}{2} \qquad (4.35)$$

(2) design of the analog filter from design tables;

(3) computation of IIR filter $H_1(z)$ from the analog filter by the well known bilinear transformation algorithms [C3, R3];

(4) computation of the first N terms of the Wiener-Lee decomposed form coefficients of the analog transfer function;

(5) time reversal of the impulse response of the FIR filter $F(z)$ and the addition of N-1 delays to obtain $\overline{F}(z^{-1})$; and,

(6) cascading $H_1(z)$ and $\overline{F}(z^{-1})$ to form the noncausal filter $H(z)$.

In step (4), it is necessary to determine the FIR filter order N from the group delay ripple requirement by an iterative method. The convergency properties guarantee that the group delay ripple decreases as N increases. While both compexity and group delay of the filter increase with N, the smallest N should be determined such that the group
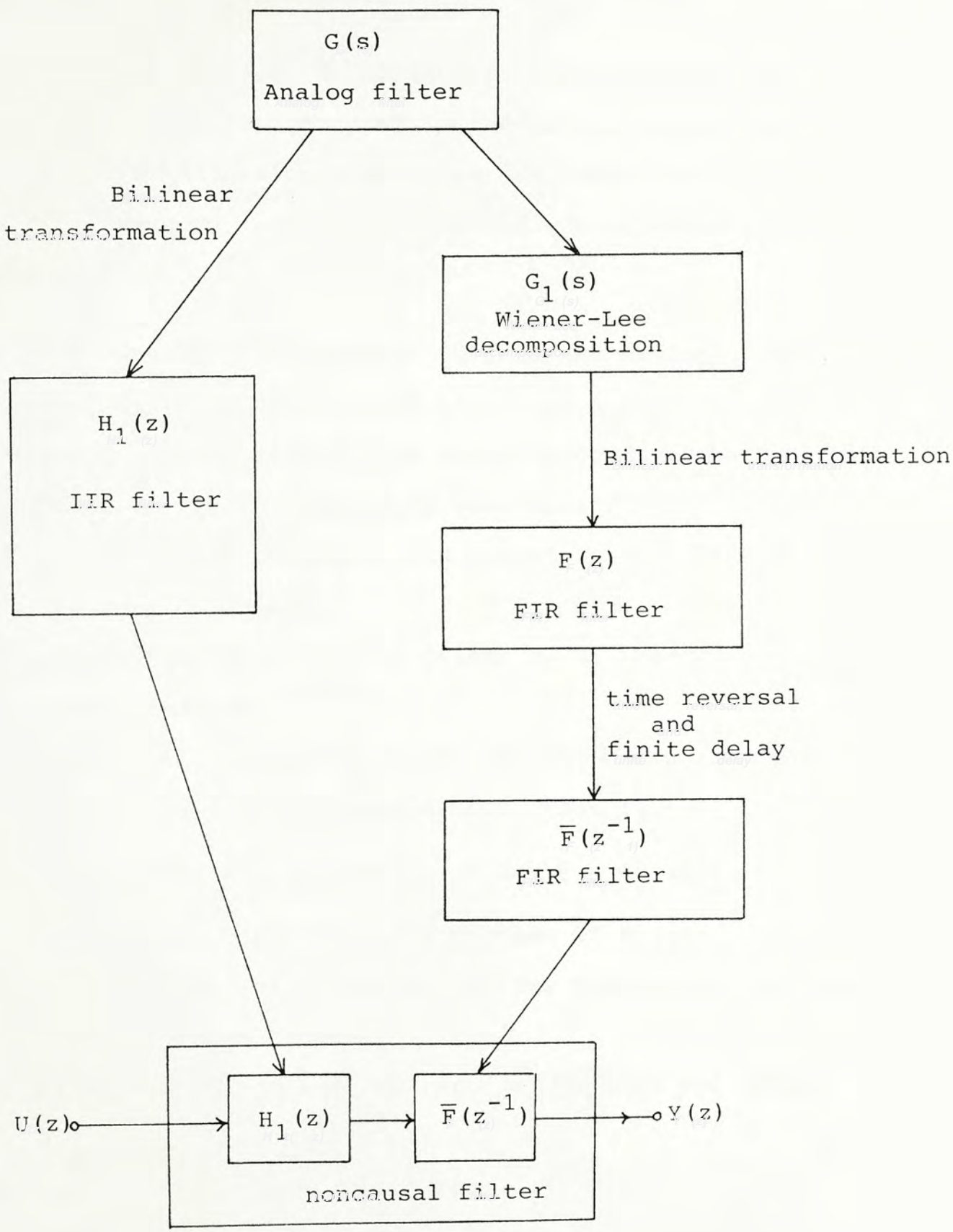
61

Fig. 4.2 Design procedure for sample-by-sample
cascade noncausal filter

delay tolerance is satisfied.

For bandpass and highpass filter design, the design has to be incorporated in the analog domain using well established analog transformation techniques since digital frequency transformations are not suitable for FIR filters.

The above procedures outline the design of cascade noncausal filter consisting of identical subfilter. For parallel noncausal filter, the design procedures become:

(1) based on digital magnitude specification and obtain an analog filter following the procedures (1) and (2) of the cascade design;

(2) obtain $H_1(z)$ from analog filter by bilinear transformation;

(3) construct the cascade filter and decompose into parallel form by partial fraction method; that is,

$$H(z) = H_1(z) H_1(z^{-1}) = G_1(z) + G_1(z^{-1}) \qquad (4.36)$$

where $G_1(z)$ contains all the poles of $H_1(z)$;

(4) expand $G_1(z)$ into a Laurent series $F(z)$ about $z = 0$ and retain first N terms;

(5) time reversal of impulse response of $F(z)$ and adding N-1 delays to form $\overline{F}(z^{-1})$;

(6) add N-1 delay to $G_1(z)$ to form $\overline{G}_1(z)$; and,

(7) form the parallel noncausal filter by connecting $\overline{G}_1(z)$ and $\overline{F}(z^{-1})$ in parallel.

It should be noted that the magnitude response is approximately given by $|H_1(e^{j\omega T})|^2$.

4.5   Filter characteristics and desgin examples

By using many design examples, we would like to investigate some common properties of the noncausal filters designed by the proposed methods. Since, in practice, the cascade realization has the advantages of ease in design and low cost implementation, we would only consider this realization in this section. The design examples cover many types of conventional filters, namely, lowpass, bandpass, highpass and bandstop filters. A comparison is being made in Example 2 with the design example using the conventional block processing approach [C1].

4.5.1   Lowpass filters
Example 1 :

A lowpass cascade noncausal filter is designed with the following specifications:

(i)   cutoff at frequency, $\omega_p T = 1.0$, with less than 0.6 dB passband ripple;

(ii)   stopband cutoff at frequency, $\omega_s T = 1.38$, with minimum stopband attenuation = 52 dB; and,

(iii) a passband group delay ripple of not more than 3%.

Prewarping frequencies by (4.35), the analog frequencies are given as $\bar{\omega}_p = 0.546$ and $\bar{\omega}_s = 0.825$. Taking the square root of the magnitude specification, the passband ripple and minimum stopband attenuation are then 0.3 dB and 26 dB respectively.

From an analog filter design table [Z1], a fourth-order elliptic filter is chosen based on the analog specifications. The filter characteristics satisfy both the $\bar{\omega}_p$ and $\bar{\omega}_s$ requirements and having a passband ripple of 0.18 dB and a minimum stopband attenuation of 27 dB.

With the analog transfer function, an IIR and a 51th order FIR filters are then computed by the described methods on a computer. Their magnitude and phase responses are shown in Fig. 4.3a and b, respectively. The 51 term FIR filter's magnitude and phase responses approximate almost exactly in the passband, however, they differ significantly in both the transition and stopband. Yet the passband ripple and minimum stopband attenuation become 0.24 dB and 28 dB respectively, both satisfying the specification. However, the stopband frequency is relaxed to 1.43.

The impulse response of the FIR filter is shown in Fig. 4.4. The convergency of the function is evident from the rapidly diminishing Wiener-Lee decomposition form coefficients. It is observed that the coefficient fall off faster than the reciprocal of n. Therefore, neglecting high order terms will have little effects on the main characteristics of the function.

Time reversing the FIR impulse response and adding N-1 unit delays, we obtained $\bar{F}(z^{-1})$. The composite filter is then obtained by cascading the IIR and FIR filters.
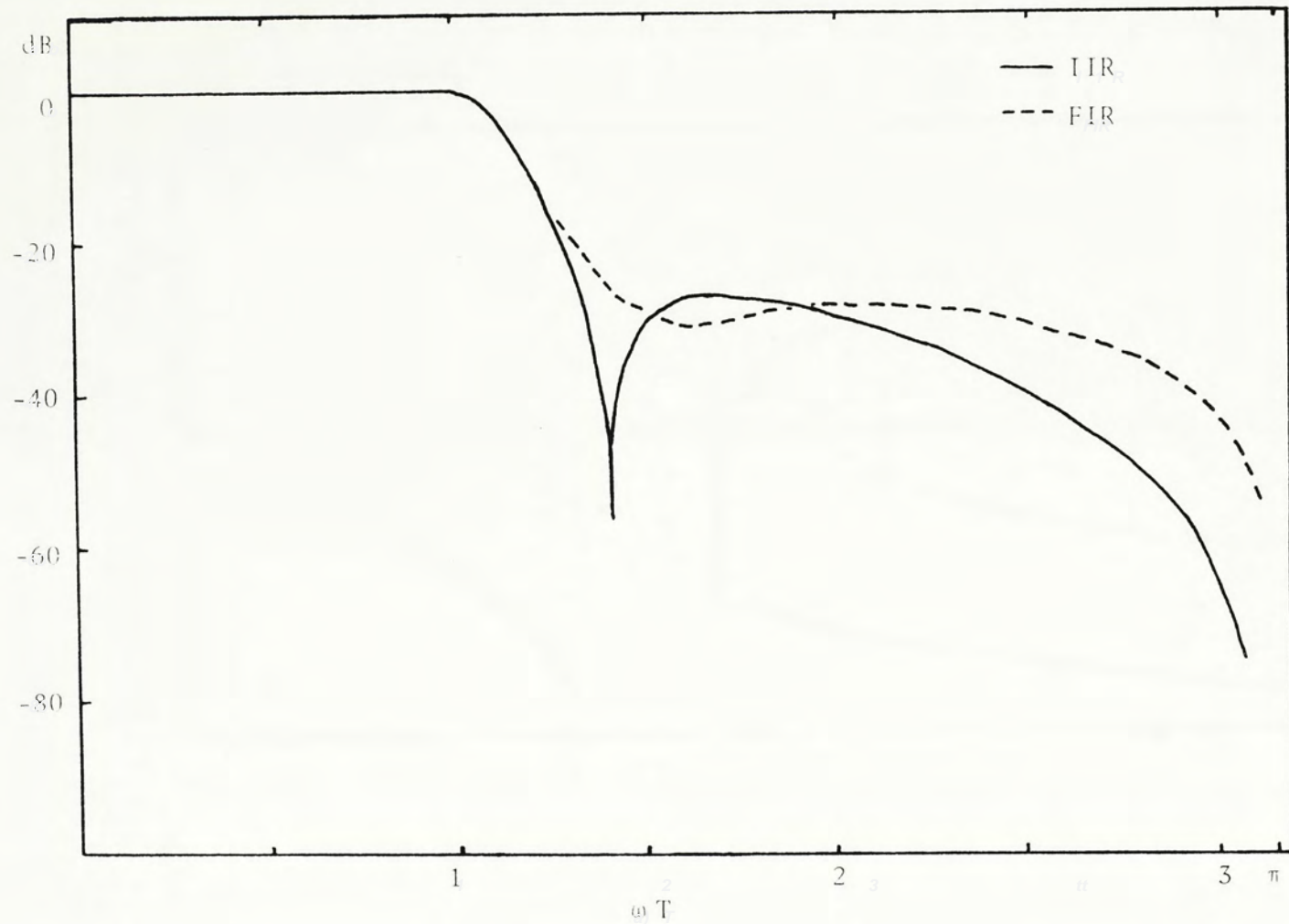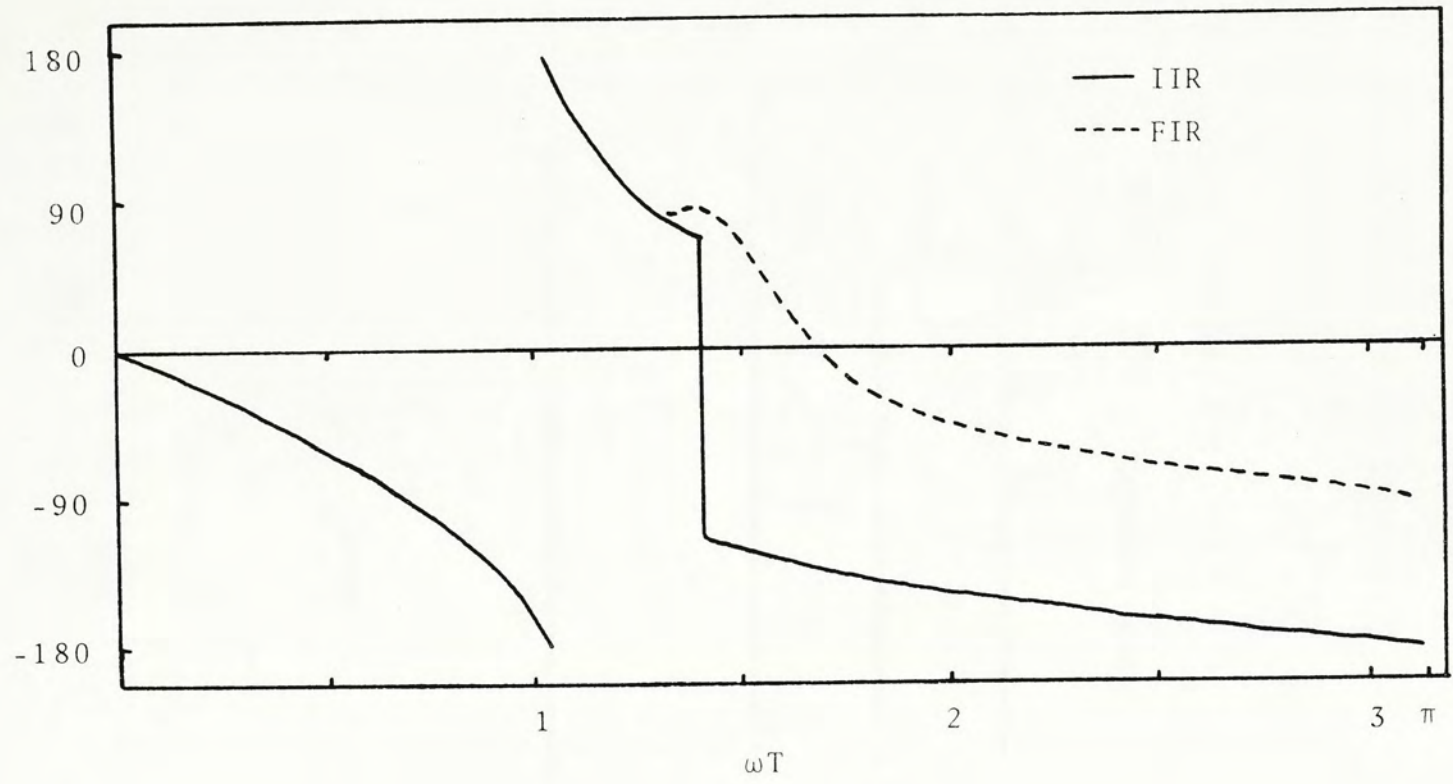
Fig. 1.3a

Fig. 4.3b

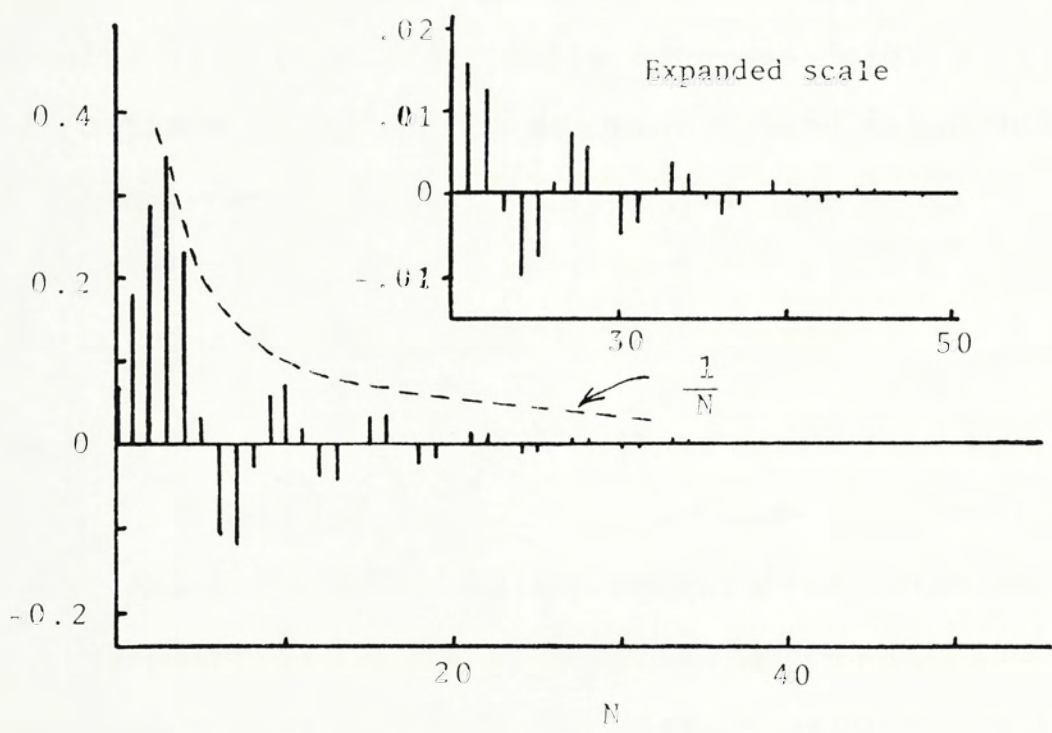Magnitude and phase responses of $H_1(z)$ and $F(z)$

Fig. 4.4 Wiener-Lee coefficients

The magnitude response and the group delay characteristics of the composite filter is shown in Fig. 4.5a and b. The passband ripple and minimum stopband attenuation are 0.4 dB and 57.4 dB respectively satisfying the specification. The stopband cutoff frequency is 1.35 which is better than the specified 1.38.

The passband normalized group delay of the composite filter is essentially constant (Fig. 4.5b) with 0.32% maximum ripple. The transition band group delay also has a near constant feature. Instead of zero group delay of the theoretical noncausal filter, a group delay of 50 samples is introduced.

Example 2 :

In this example, it is intended to compare the general characteristics of the proposal realization method with the convention block processing approach. The comparison is made between the present example and the filter example shown in Czarnach [C1]. The present lowpass filter is designed based on the following specification :

(i) $\omega_p T = 0.5\pi$, $\omega_s T = 0.54\pi$ with 0.5 dB passband ripple
        and 70 dB stopband attenuation; and

(ii)   a passband group delay ripple of less than 3%.

A 7th order elliptic filter is design for the subsystems using a table [Z1]. The passband ripple is 0.04 dB and stopband attenuation is 42.7 dB. Based on this elliptic filter, an IIR and a FIR filter are constructed.
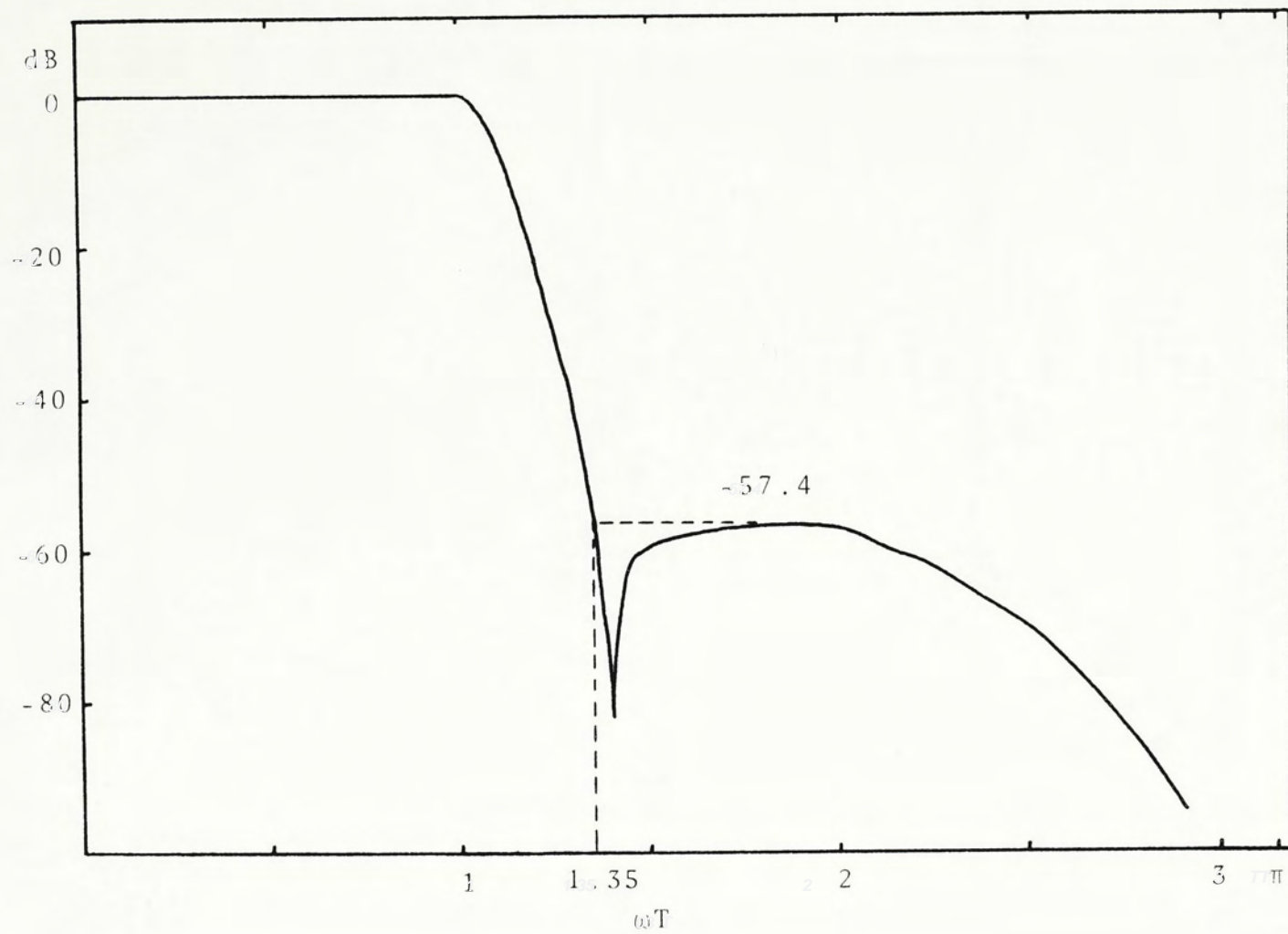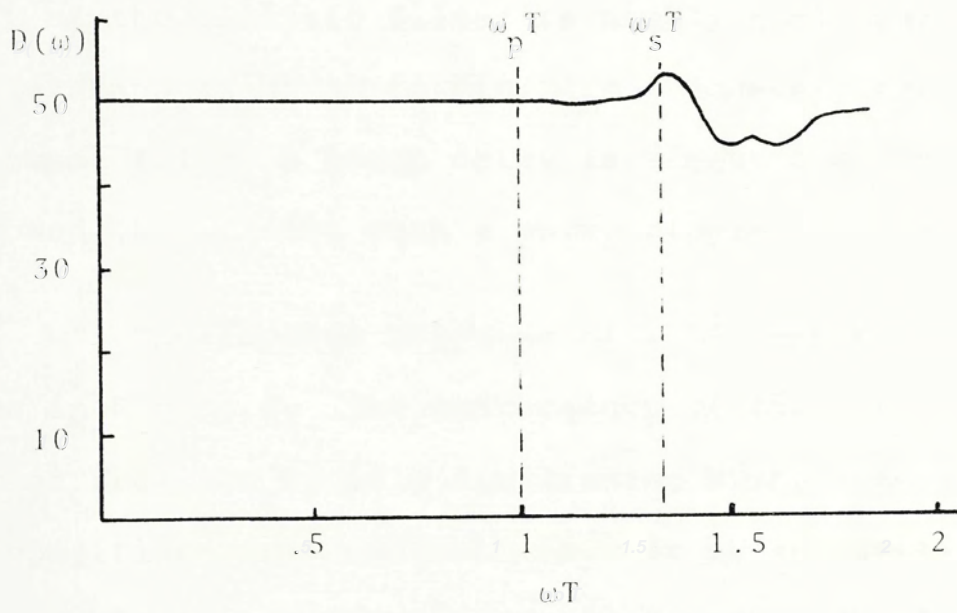
Fig. 4.5a

Fig. 4.5b

Time reversing the FIR impulse response and adding N - 1 delays, we obtained $\bar{F}(z^{-1})$. The composite filter is then obtained by cascading the IIR and FIR subfilters.

The magnitude response and the group delay characteristics of a 70 term FIR filter cascading with an IIR filter are shown in Fig. 4.6a and b. The passband ripple and minimum stopband attenuation are 0.46 dB and 75 dB respectively satisfying the specification. The group delay of the elliptic filter is highly nonlinear within the passband as shown in Fig. 4.7. However, the composite noncausal filter's group delay is almost flat in the passband (Fig. 4.6b) with a delay ripple of less than 3%.

The impulse response of a 25 term FIR filter is shown in Fig. 4.8. The convergency of the function is evident from the rapidly diminishing Wiener-Lee decomposition form coefficients. It is interesting to investigate the effect of the FIR filter order N on the composite filter parameters, such as, passband ripple, stopband attenuation and passband group delay ripple. Fig. 4.9 shows the dependence of these parameters on N. With increasing N, all filter parameters converge to ideal values rapidly.

For a similar filter using the block processing approach [C1], a segmentation of L = 2000 is needed with an overlapping length q of 400 points. The normalized group delay is equal at least 4000 samples. Including the wasted

overlapping sample points, the number of real
mutliplications needed for each output sample is 20.

The differences between the present filter and
the filter shown in Czarnach [C1] is shown in Table 4.1.
For the present example, it is noted that the group delay
is 57 times shorter, the memory size needed is 26 times
smaller while the number of multiplication is 4.2 times
more than that of the filter realized by block precessing
technique.

	sample-by-sample approach	block processing approach
group delay	69	4000
memory size	77	2000
multiplication per sample	84	20

Table 4.1   Comparison between sample-by-sample
approach and block processing approach

It is also noted that the proposed method is more
flexible. The group delay errors and magnitude response
errors are adjustable with the FIR filter order N.   If the
linear phase requirement is not very strict, N can be
fairly small. For the block processing approach, however,
the error in phase has not been investigated.

In conclusion, the new approach has the advantages
of short basic group delay time, small memory size
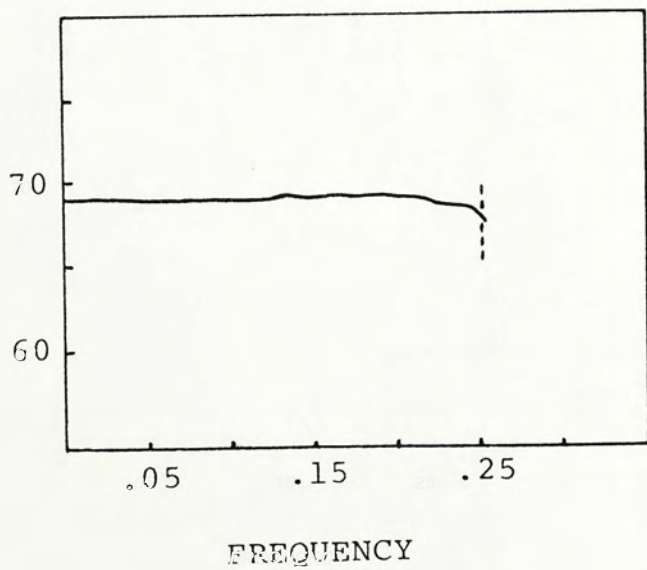
FREQUENCY

(a)

FREQUENCY
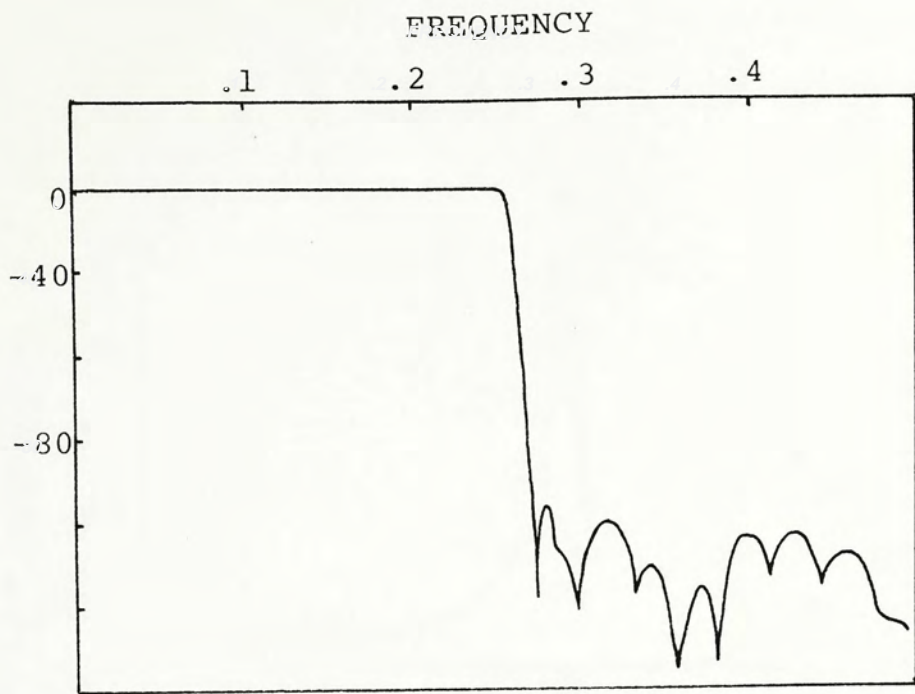
Fig. 4.6   Noncausal filter (N=70); (a) Magnitude response;
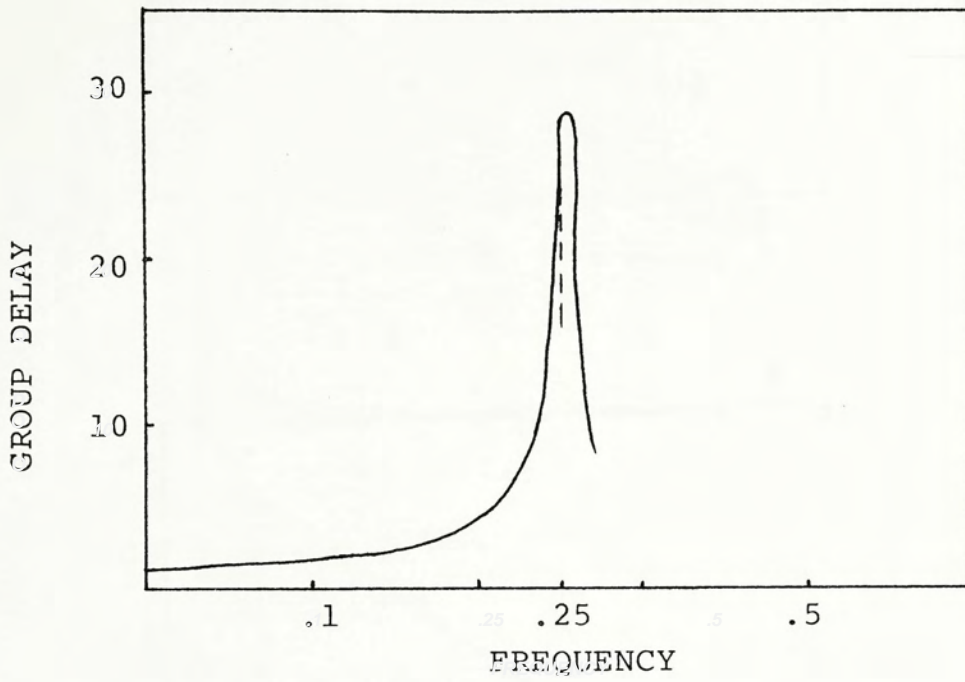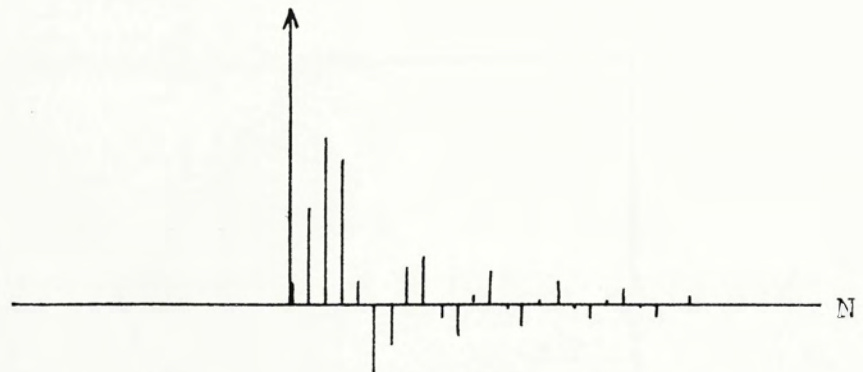(b) group delay response.

Fig. 4.7   Group delay of 7th order elliptic filter

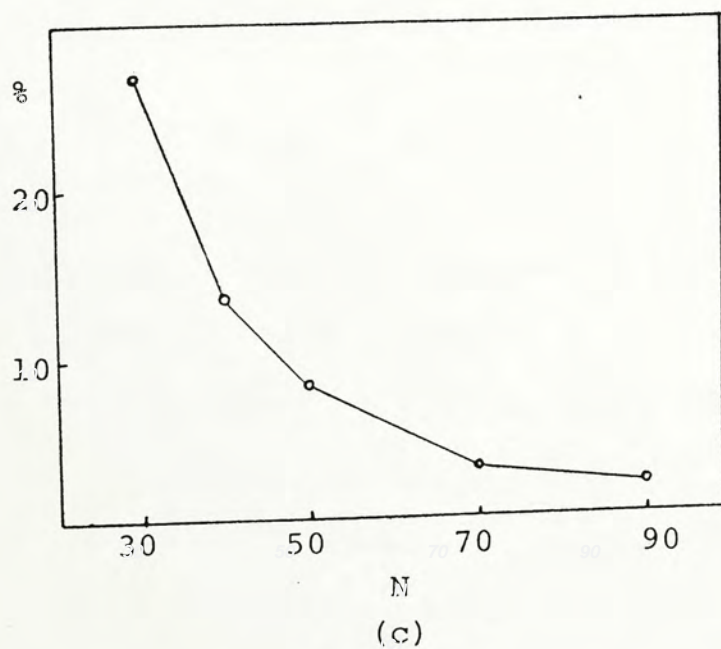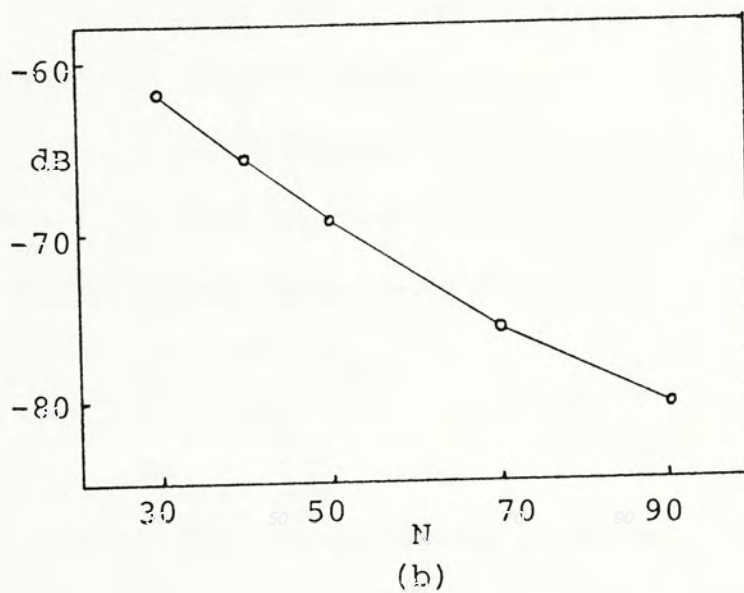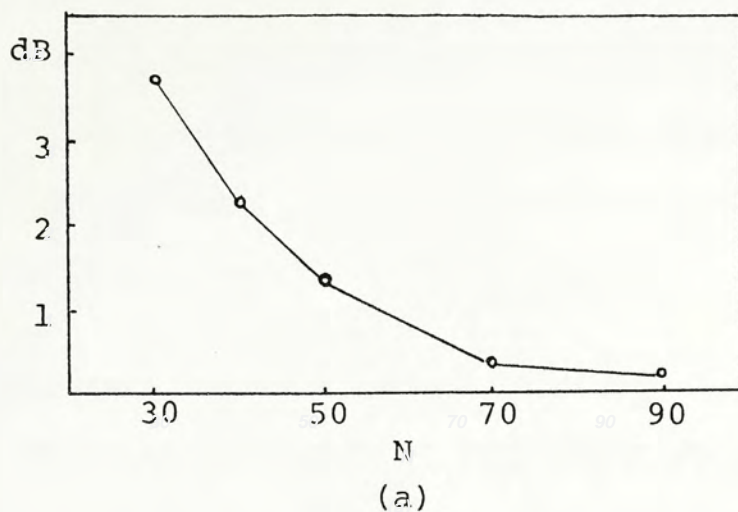Fig. 4.8   Impulse response of F(z)

Fig. 4.9 Effect of FIR order N on: (a) passband ripple;
(b) minimum stopband attenuation; (c) passband
group delay ripple.

requirement and exactly known phase response. The main
disadvantage is that the number of multiplication per
output sample can be fairly large if the group delay
requirement is very strict (e.g. less than 1% ripple).

4.5.2 Bandpass filter

To design a bandpass noncausal filter by the
proposed method, it is necessary that the analog prototype
filter is of bandpass type. This is due to the fact that
digital frequency transformation is not suitable for the
FIR subfilter of the noncausal filter. Usually after the
lowpass analog filter is selected from filter design table,
a lowpass to bandpass transformation [R3],

$$s \rightarrow \frac{s^2 + \overline{\omega}_\ell \, \overline{\omega}_u}{s(\overline{\omega}_u - \overline{\omega}_\ell)} \qquad (4.36)$$

is applied to obtain the bandpass prototype filter. The
analog upper and lower cutoff frequencies, $\overline{\omega}_u$ and $\overline{\omega}_\ell$, are
computed by (4.35) from the digital frequency
specifications.

Example 3 :

A wideband bandpass noncausal filter is designed
based on the following specifications :
(i)   low cutoff frequency $F_\ell = 0.15$, upper cutoff
       frequency $F_u = 0.35$; passband ripple of less than
       0.4 dB and minimum stopband attenuation of 80 dB.
(ii)   passband group delay ripple of less than 1%.
The filter is realized by cascade structure.

77

According to the design procedure as outlined in Section 4.4, a 4th order elliptic lowpass filter is selected from table with a passband ripple of 0.18 dB and a minimum stopband attenuation of 41 dB. The analog cutoff frequencies given by (4.35) are $\bar{\omega}_\ell = 0.51$ and $\bar{\omega}_u = 1.96$. The 8th order bandpass filter is then obtained by applying the analog frequency transformation (4.36) on the lowpass filter.

Following the procedures, an IIR and a 60 term FIR bandpass filters are constructed based on bilinear transformation. The noncausal filter's magnitude response and passband group delay are plotted in Fig. 4.10a and b. The passband ripple and minimum stopband attenuation is 0.36 dB and 81 dB respectively. The passband group delay is essentially constant (Fig. 4.10b) with a ripple of 0.64% satisfying the specifications.

To observe the convergency of the FIR subfilter, the passband group delay ripple is plotted against N as shown in Fig. 4.11. The group delay ripple decreases rapidly towards zero as N is increased. Comparing with that of the lowpass case (Fig. 4.9) it is not surprising to note that they both have a common exponentially decreasing characteristics.

Example 4 :

Using the same analog lowpass filter, a narrowband bandpass noncausal filter is designed with cutoff frequencies at $F_\ell = 0.15$ and $F_u = 0.25$. Going
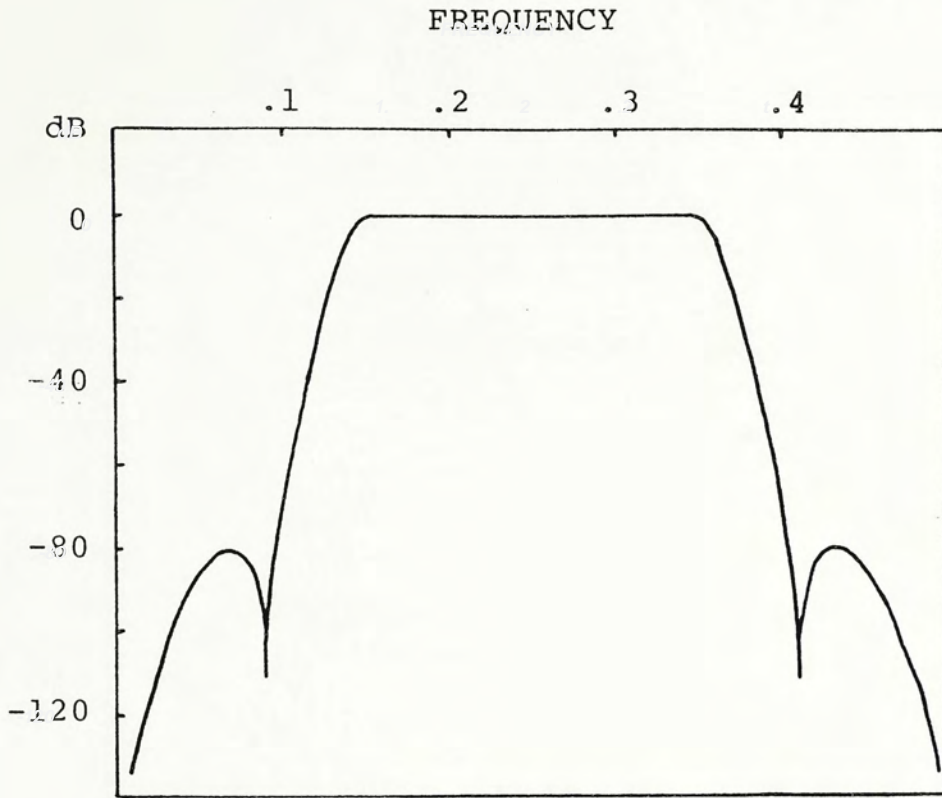
(a)
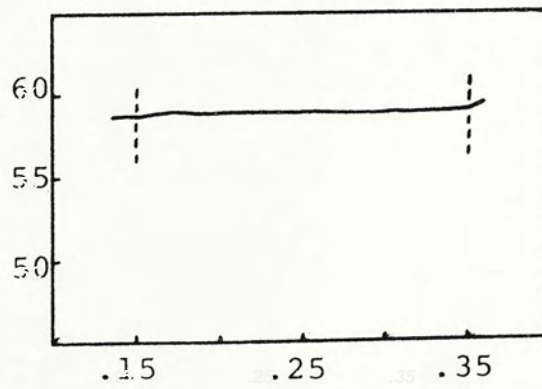
(b)

Fig. 4.10   Bandpass noncausal filter (N=60); (a) magnitude

response; (b) group delay

Fig. 4.11   Effect of N on group delay ripple

(a)

(b)

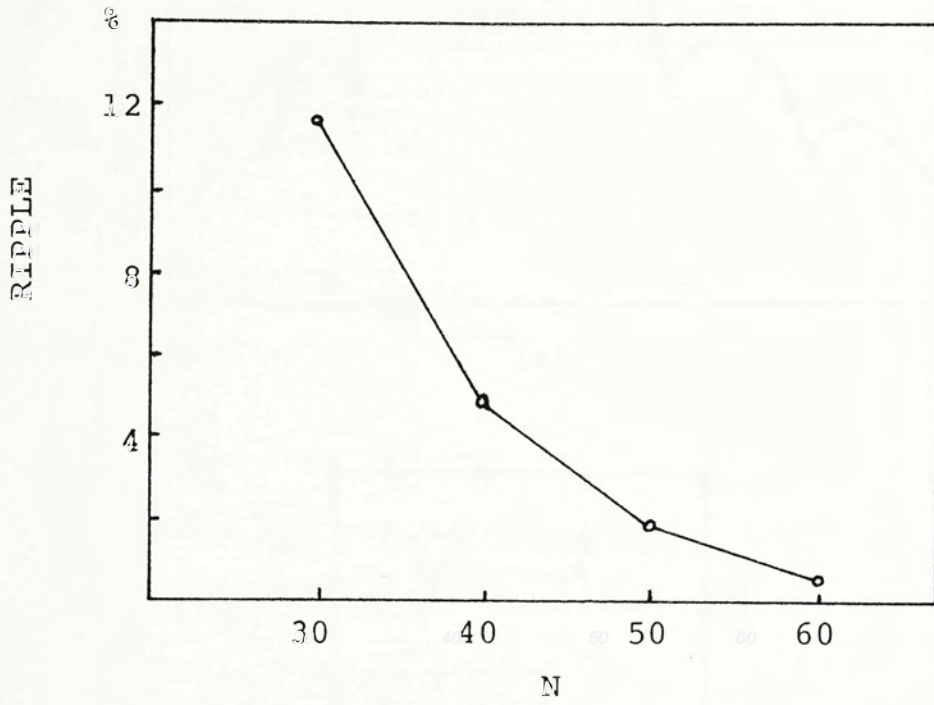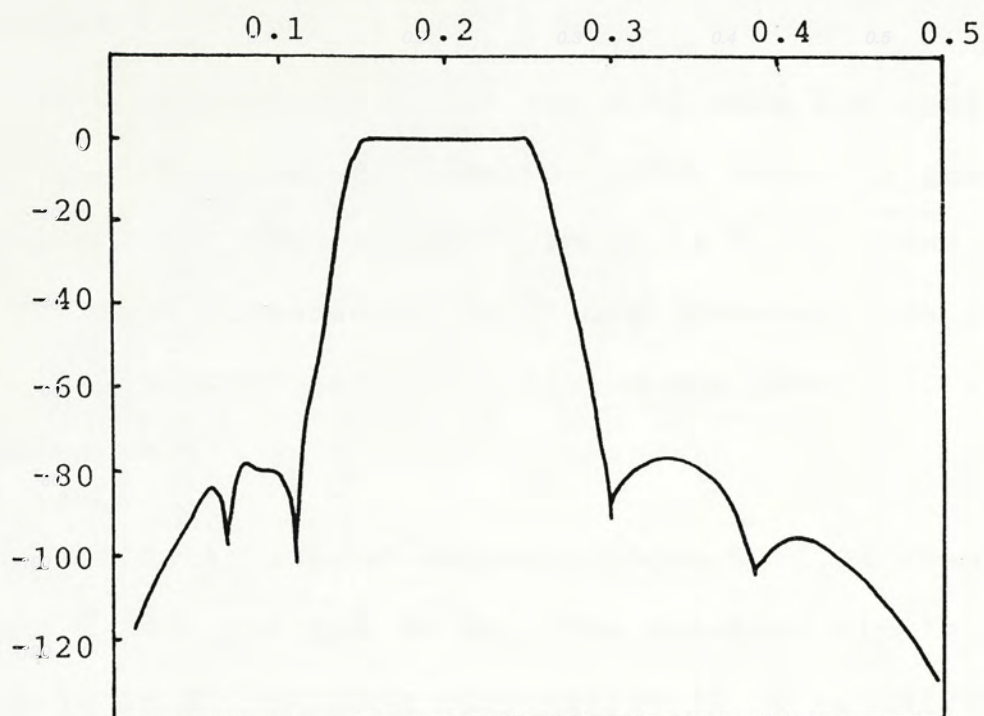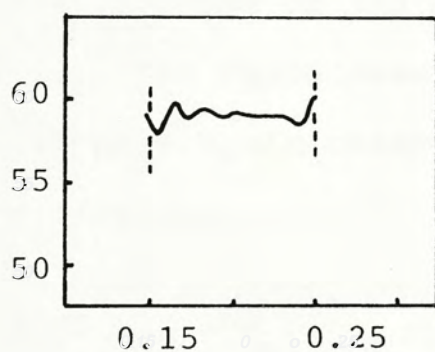Fig. 4.12   Bandpass noncausal filter (N = 60);
            (a) magnitude response; (b) group delay.
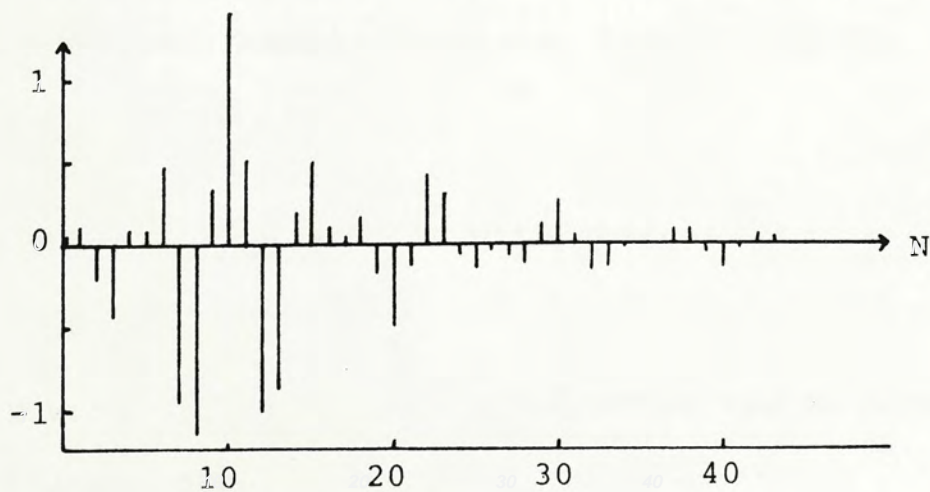
Fig. 4.13   Impulse response

through the same procedure, the bandpass noncausal filter
is realized by cascading an IIR and a 60 term FIR subfilter.
The magnitude response and passband group delay is shown in
Fig. 4.12a and b. The passband ripple is 0.72 dB and
minimum stopband attenuation is 77 dB. However, the group
delay has significant ripple (4.6%) in the passband, which
is not acceptable.

The delay ripple decreases down to 1.3% when the
FIR length N is increased to 80. The passband ripple then
becomes 0.43 dB and minimum attenuation 80 dB satisfying
the specifications.

The impulse response of the bandpass FIR subfilter
is shown in Fig. 4.13. The rapid decay of the Wiener-Lee
coefficient with N is once again observed in consistent
with theoretical predictions.

4.5.3 Highpass filter

Like the bandpass filter case, the design of
highpass noncausal filter can be obtained from an analog
highpass prototype filter. The highpass prototype filter
is derived using the analog frequency transformation,

$$s \leftarrow \frac{\overline{\omega}_p}{s} \qquad (4.37)$$

Other procedures follow exactly with that of the bandpass
case.

However, it is possible to derive the highpass
noncausal filter based on a corresponding lowpass

noncausal design using the digital lowpass to highpass transformation,

$$z \rightarrow -z \qquad\qquad (4.38)$$

Unlike other kinds of digital frequency transformation, the nonrecursive property of the FIR filter remains unchanged under the transformation. The highpass cutoff frequency, F_{hp}, is related to the lowpass cutoff frequency, $F_{\ell p}$, by,

$$F_{hp} = 0.5 - F_{\ell p} \qquad\qquad (4.39)$$

Computationwise, it is easier to implement the digital transformation (4.38) on computer than the corresponding analog transformation (4.37). Nevertheless, both methods give the same result.

Example 5 :

A 7th order lowpass elliptic filter is shown as the prototype filter with 0.04 dB passband ripple and 42 dB minimum stopband attenuation. The digital highpass cutoff frequency F_{hp} is at 0.3; thus, the corresponding lowpass cutoff frequency $F_{\ell p}$ is at 0.2. Following the same procedures in the lowpass design, a lowpass noncausal filter is first obtained. Applying the digital frequency transformation (4.38), the desirable highpass design is obtained. The highpass noncausal filter consists of an IIR subfilter cascading with a 50 term FIR subfilter. The magnitude response and group delay characteristics are shown in Fig. 4.14a and b. Since the FIR order is relatively low, large ripples are observed near the

FREQUENCY
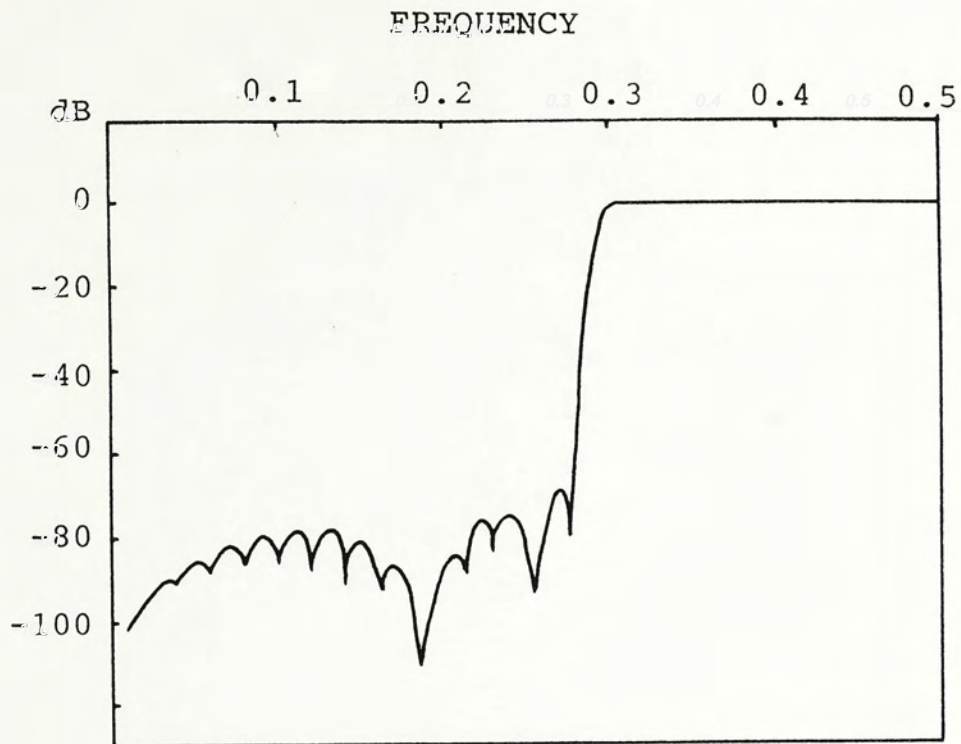
(a)
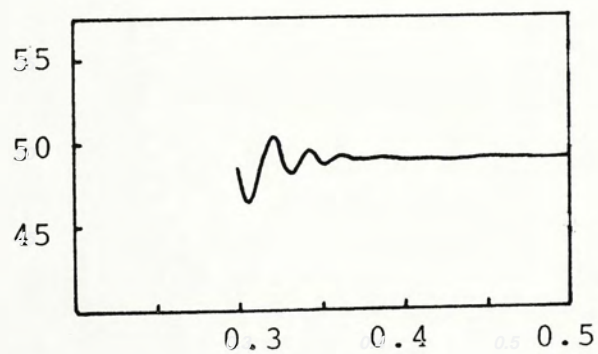
(b)

Fig. 4.14 Highpass noncausal filter (N = 50);

(a) magnitude response, (b) group delay.
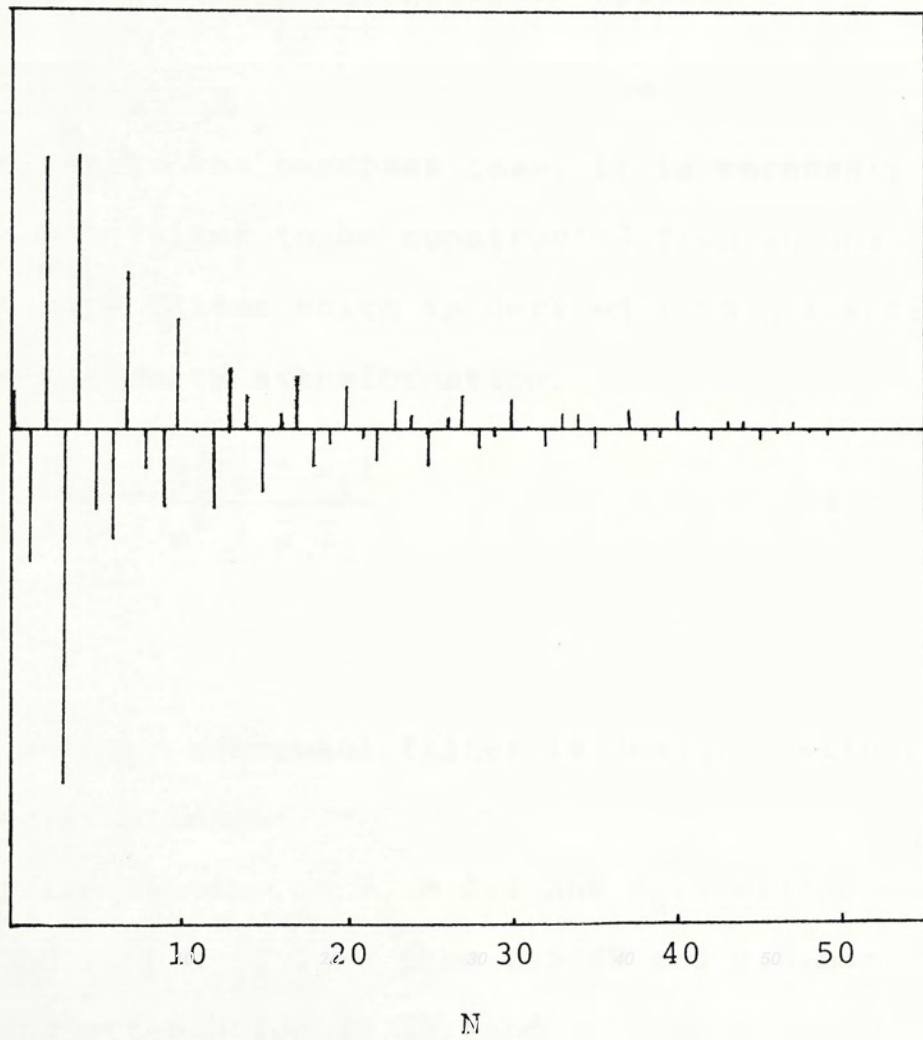
Fig. 4.15   Impulse response

passband edge. The passband group delay ripple is 8.3%.

The impulse response of the highpass FIR subfilter (Fig. 4.15) is exactly the same as that of the lowpass FIR subfilter except that all the odd order terms are reversed in sign. Thus, the same convergency properties is expected for the highpass filter.

4.5.4 Bandstop filters

Similar to the bandpass case, it is necessary for the noncausal filter to be constructed from an analog bandstop prototype filter which is derived from a lowpass filter by the frequency transformation,

$$s \to \frac{s(\overline{\omega}_u - \overline{\omega}_\ell)}{s^2 + \overline{\omega}_u \overline{\omega}_\ell} \tag{4.40}$$

Example 6 :

A bandstop noncausal filter is designed with the following specifications:

(i)   cutoff frequencies at $F_\ell = 0.1$ and $F_u = 0.4$;

(ii)   passband ripple of less than 0.5 dB and minimum stopband attenuation 90 dB; and

(iii) passband group delay ripple of less than 3%.

Prewarping the cutoff frequencies by (4.35), the analog cutoff frequencies are $\overline{\omega}_\ell = 0.32$ and $\overline{\omega}_u = 3.08$. From filter table, a 4th order lowpass filter is selected with passband ripple 0.18 dB and minimum stopband attenuation of 50 dB. The bandstop filter is then derived

by (4.40). Applying the bilinear transformation to the bandstop filter, the IIR and FIR subfilters are obtained. The magnitude response and passband group delay characteristics of a 40 term FIR filter cascading with an IIR filter are shown in Fig. 4.16a and b. The passband ripple is 0.48 dB while minimum stopband attenuation is 96 dB. The passband group delay are nearly constant with a ripple of 2.34%.

Plotting the passband group delay ripple of this filter against N (Fig. 4.17), the convergence properties of the nonrecursive realization method are again evident.

.1   .2   .3   .4

0

−40

−80

−120

(a)

45

40

35

.1   .2   .3   .4
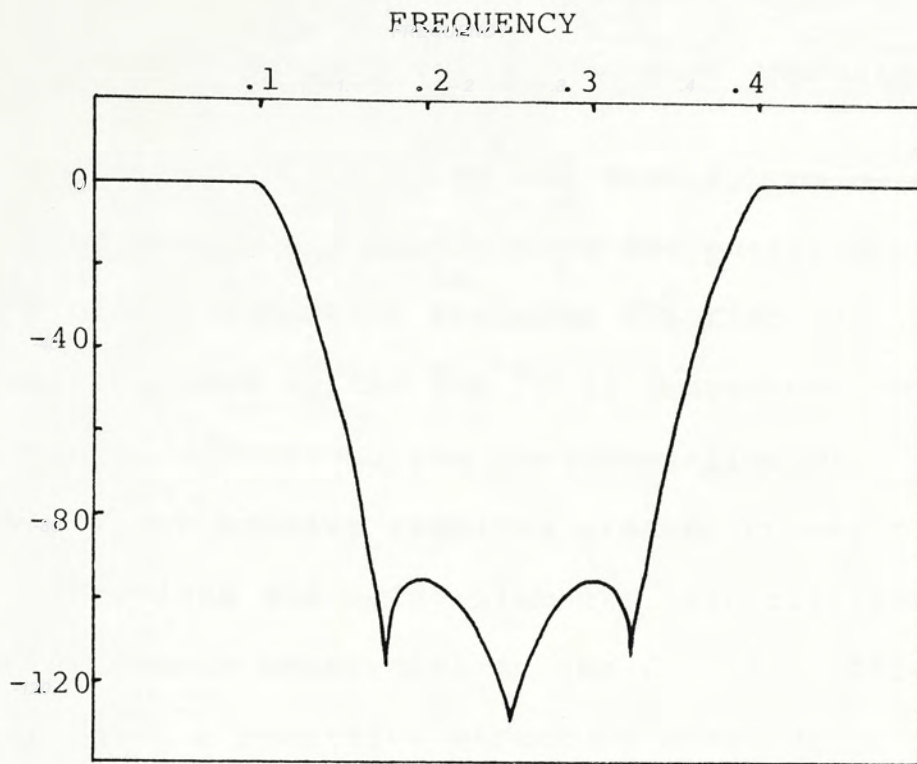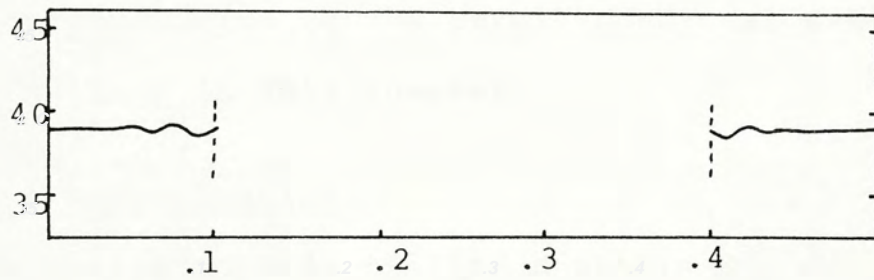
(b)

Fig. 4.16   Bandstop noncausal filter (N=40).

%

15

10

5

20   30   40

N
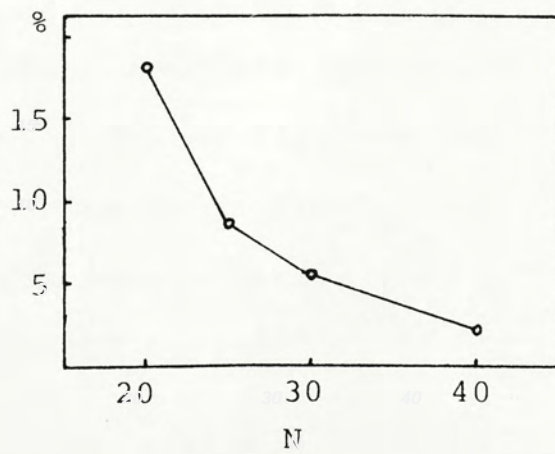
Fig. 4.17   Effect of N on group delay ripple

Chapter 5

Sample-by-sample approach using pole-zero approximation

In the last chapter, an all zero approximation method has been derived to approximate the purely noncausal part $H_1(z^{-1})$ of the noncausal transfer function $H(z)$. The resulting nonrecursive filter $\overline{F}(z^{-1})$ is guaranteed to be both stable and causal. However, the implementation of nonrecursive filter usually requires greater number of arithmetic operations and delay elements, especially when there is no symmetric constraint on the filter coefficients. On the other hand, a recursive structure does not in general have these disadvantages. Therefore, methods of constructing a recursive approximation to the purely noncausal function $H_1(z^{-1})$ are explored in this chapter.

5.1 Recursive Approximation

The design goal is to find a stable and causal recursive function $R(z)$ such that both the magnitude and phase responses closely approximate that of the purely noncausal function $H_1(z^{-1})$ except for a possible linear phase shift. Accurate phase response is especially important when the original noncausal filter $H(z)$ has zero phase characteristics. The problem is particularly difficult since all the poles of $R(z)$ must locate inside the unit circle while that of $H_1(z^{-1})$ all lie outside the unit circle.

A general approach to the problem is to match some characteristics of the impulse responses of $R(z)$ and $H_1(z^{-1})$. The Laurent series of $H_1(z^{-1})$ is given by,

89

$$H_1(z^{-1}) = \sum_{k=-\infty}^{0} h(k) \, z^{-k} \tag{5.1}$$

Multiplying (5.1) by $z^{-(N-1)}$, we obtain a delayed series,

$$z^{-(N-1)} H_1(z^{-1})$$

$$= \sum_{k=0}^{N-1} h(k-N+1) \, z^{-k} + \sum_{k=-\infty}^{0} h(k-N+1) \, z^{-k} \tag{5.2}$$

which contains a positive and a negative series. The positive time series is causal and therefore can be approximated by a recursive function.

A nth order recursive filter in general has the form,

$$R(z) = \frac{b_0 + b_1 z^{-1} + \ldots + b_m z^{-m}}{1 + a_1 z^{-1} + a_2 z^{-2} + a_n z^{-n}} \,, \qquad n \geq m \tag{5.3}$$

where not all a_i are equal to zero. Requiring that $R(z)$ be causal, the Laurent expansion of $R(z)$ about $z = 0$ is,

$$R(z) = \sum_{k=0}^{\infty} r(k) \, z^{-k} \tag{5.4}$$

It is possible to match (5.4) with the positive time series of (5.2) and obtain $R(z)$. This is equivalent to the time domain design of recursive digital filters problem with many well known methods [B1, B2, H1, S3, Y1]. Before some of the time domain design methods are described, it is necessary to point out that the approximation can never be exact. The frequency response of a delayed noncausal function given by (5.2) is,

90

$$H_1(e^{-j\omega}) \, e^{-j(N-1)\Omega} = \sum_{k=0}^{N-1} h(k-N+1) \, e^{-j\omega k} + \sum_{k=-\infty}^{0} h(k-N+1) \, e^{-j\omega k}$$

$$= \overline{F}(e^{j\omega}) + E(e^{j\omega}) \qquad\qquad (5.5)$$

where $\overline{F}(z)$ is the stable and causal all zero approximant while $E(e^{j\omega})$ is the error term. Since only $\overline{F}(z)$ is used to obtain $R(z)$, even if the approximation is exact there is at least an error of $E(e^{j\omega})$. As being discussed in last chapter, however, this error can be made arbitarily small since $\overline{F}(e^{j\omega})$ converges to $H_1(e^{-j\omega})$ as N increases for most practical transfer functions satisfying Dirichlet conditions.

5.2 Pade' Approximation and Least Square Technique

The earliest major development in time domain design of recursive digital filters appears to be due to Burrus et al [B1]. The approximation of the nonrecursive power series $\overline{F}(z)$ by a rational function $R(z)$ is often referred to as the Pade' approximation technique.

Rewritting the N term nonrecursive approximation to the delayed purely noncausal function $H_1(z^{-1})$, we have

$$\overline{F}(z) = \sum_{k=0}^{N-1} f_k \, z^{-k} \qquad\qquad (5.6)$$

where $f_k = h(k-N+1)$. The time domain problem is to find the coefficients a_i and b_i such that the first N terms of (5.4) match as close as possible with f_k. If $N = n + m + 1$, exact solution for a_i and b_i can be obtained by equating $r(k)$ with f_k for $k = 0, 1, \ldots N-1$.

Setting $r(k) = f_k$, for $k = 0, 1, \ldots N-1$ and $r(k) = 0$ for $k > N-1$, the inverse z-transform of $R(z)$ is

$$\frac{b_0 + b_1 z^{-1} + \ldots + b_m z^{-m}}{1 + a_1 z^{-1} + \ldots + a_n z^{-n}} = f_0 + f_1 z^{-1} + f_2 z^{-2} + \ldots$$

(5.7)

Writting (5.7) in matrix form, this becomes

$$
\begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_m \\ 0 \\ \vdots \\ 0 \end{bmatrix}
=
\begin{bmatrix}
f_0 & 0 & 0 & \cdots & 0 \\
f_1 & f_0 & 0 & & \cdot \\
f_2 & f_1 & f_0 & & \cdot \\
\cdot & & & & \cdot \\
\cdot & & & & \cdot \\
\cdot & & & & \cdot \\
f_{N-1} & f_{N-2} & & & f_{N-n-1}
\end{bmatrix}
\begin{bmatrix} 1 \\ a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_n \end{bmatrix}
$$

(5.8)

This equation is partitioned into the following form,

$$
\begin{bmatrix} b_0 \\ b_1 \\ \cdot \\ \cdot \\ b_m \end{bmatrix}
\quad
\begin{bmatrix}
f_0 & 0 & \cdots & 0 \\
\cdot & & & \cdot \\
\cdot & & & \cdot \\
\cdot & & & \cdot \\
f_m & & &
\end{bmatrix}
$$
$$
=
\begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_n \end{bmatrix}
$$
$$
\begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}
\quad
\begin{bmatrix}
f_{m+1} & & & \\
\cdot & & & \\
\cdot & & & \\
\cdot & & & \\
f_{N-1} & \cdots & & f_{N-n-1}
\end{bmatrix}
$$

(5.9)

or

$$\begin{bmatrix} b \\ \hline 0 \end{bmatrix} = \begin{bmatrix} F_1 \\ \hline F_2 \end{bmatrix} [a] \qquad\qquad (5.10)$$

Since $N = m + n + 1$, F_2 has n rows and n + 1 columns which guarantee a nontrivial solution to the equation $[a] [F_2] = 0$. The coefficients b_i are then given by

$$[b] = [F_1] [a] \qquad\qquad (5.11)$$

Even though, the first N points of the impulse responses of $\overline{F}(z)$ and $R(z)$ are matched exactly, there are no constraints on the values of the impulse response of $R(z)$ for k > N. In fact, the Pade' approximants, if not unstable, often exhibit very significant tail components outside the specified region [H1]. In such case, the frequency response is poorly approximated and cannot be used to realize the noncausal filter.

Different techniques has been used by various authors to deal with the stability and tail problems of the Pade' approximants [B2, B3, H1]. Iterative procedures are often used to minimize a weighted mean square error over a finite number of time samples. Long computation time is required especially when designing a high order recursive filter. In some cases, the approximant could not be obtained due to ill-conditioning of the matrix [H1].

The major drawback of most time domain design methods when applying to the present problem is that the minimization in time samples' error may not guarantee a satisfactory frequency characteristics required for the noncausal filters.

The exact frequency response is given by

$$\overline{F}(e^{j\omega T}) = \sum_{k=0}^{N-1} f_k \, e^{jk\omega T} \tag{5.12}$$

Suppose an error ε occurs at the pth time sample f_p, the frequency response will have an error term,

$$E_p(e^{j\omega T}) = \varepsilon e^{jp\omega T} \quad , \quad 0 < p < N-1 \tag{5.13}$$

However if the same error occurs at the qth time sample f_q, then the error in frequency response is

$$E_q(e^{j\omega T}) = \varepsilon e^{jq\omega T} \quad , \quad 0 < q < N-1 \text{ and } q \neq p \tag{5.14}$$

Comparing (5.13) and (5.14) it can be seen that the not only the magnitude of the time sample errors but also their locations on the time axis contribute to the frequency response errors. Thus when the weighted sum of square error in time samples is used as a performance index in the iteration procedure, the frequency domain error is unpredictable since the locations of the time errors have not been taken into account. The normal optimization procedure is therefore not suitable for the present design problem.

A new method suitable for the design of a recursive digital filter from the all zero approximation of the purely noncausal subfilter is developed in the next section.

94

5.3 Continued Fraction Expansion

A method suitable for obtaining a recursive transfer function $R(z)$ from the nonrecursive approximation $\overline{F}(z)$ of a delayed impulse response of $H_1(z^{-1})$ can be derived based on the continued fraction expansion technique often used in control theory [C5, C6, S4] and analog network synthesis [S3, T1, V2]. The model reduction technique has been applied to solve the stability problem of digital recursive filters designed by time domain methods [K3].

Consider a N term nonrecursive approximation to the delayed purely noncausal function of $H_1(z^{-1})$ given by (5.6). The polynomial $\overline{F}(z)$ in the z domain is transformed to a rational function $F(s)$ in the s domain by the familiar bilinear transformation,

$$F(s) = \overline{F}(z)\bigg|_{z = \frac{1 + s}{1 - s}}$$

$$= \sum_{k=0}^{N-1} f_k \left(\frac{1 - s}{1 + s}\right)^k$$

$$= \frac{d_{N-1}s^{N-1} + d_{N-2}s^{N-2} + \ldots + d_0}{c_{N-1}s^{N-1} + c_{N-2}s^{N-2} + \ldots + c_1 s + 1} \qquad (5.15)$$

The coefficients c_i are easily obtained by a binomial expansion since they originate from the expansion of $(1 + s)^{N-1}$. The coefficients d_i are readily found by a matrix transformation of f_i [C3],

95

$$[d_0, d_1, \ldots d_{N-1}] = [f_{N-1}, f_{N-2}, \ldots f_0] \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1N} \\ t_{21} & & & \cdot \\ \vdots & & & \cdot \\ t_{N1} & & \cdots & t_{NN} \end{bmatrix}$$

(5.16)

The matrix elements t_{ij} can be generated by a simple algorithm [C3]. Every elements of first column are unity. The first row is the standard binomial coefficients of the expansion of $(1 - s)^{N-1}$. Other elements of the matrix T are then given by the formula,

$$t_{ij} = t_{i-1,j} + t_{(i-1),(j-1)} + t_{i,j-1}$$

$$i = 2, 3, \ldots N$$

$$j = 2, 3, \ldots N \qquad (5.17)$$

A model reduction technique based on the continued fraction is then applied to F(s). Assuming $\overline{F}(z)$ is a lowpass filter, the numerator and denominator polynomials of (5.15) are first rearranged in ascending order and then expanded into a continued fraction about s = 0 by repeated division,

$$F(s) = \cfrac{1}{K_1 + \cfrac{s}{K_2 + \cfrac{s}{K_3 + \cfrac{s}{K_4 + \cdot_{\cdot_{\cdot}}}}}}$$

(5.18)

If a mth order recursive filter is desirable, $2m + 1$ quotients in (5.18) are retained and the inverse procedure performed. The rational function in the s domain is then transformed back to the z domain by the inverse bilinear transformation. The resulting m^{th} order recursive filter R(z) is the desired approximation to the purely noncausal subfilter. The design procedures are summarized as shown in Fig. 5.1.

Even though the continued fraction technique have a long history of application, however, the physical meaning of this operation on a transfer function is first pointed out by Chen [C5]. In feedback concept, the continued fraction expansion (5.18) corresponds to a combination of many feedback and feed forward blocks (Fig. 5.2). The outermost loop is the most dominant and corresponds to the steady state solution. As the quotients in (5.18) descend lower and lower in position, they are less and less important in terms of their influence on the performance of the system. Thus, the simplified transfer function is obtained by discarding the inner loops of the system.

In the time domain, the continued fraction expansion and time moments matching methods are similar in that they both match the first few time moments of the system [B4]. In fact, once the time moments of the system are known the quotients of the continued fraction expansion can be computed and vice versa [B4]. In the frequency domain, the matching of the first few time moments is
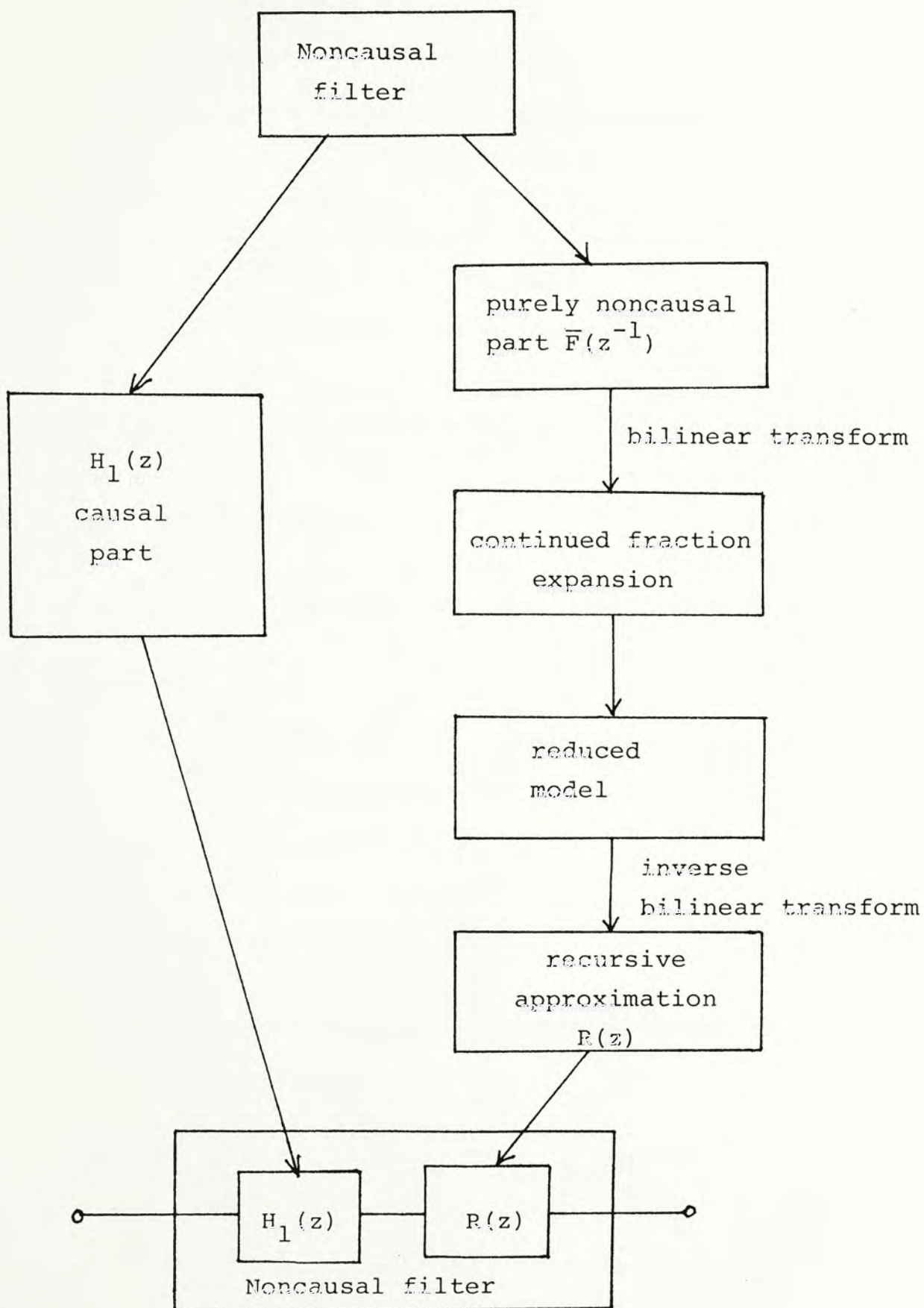
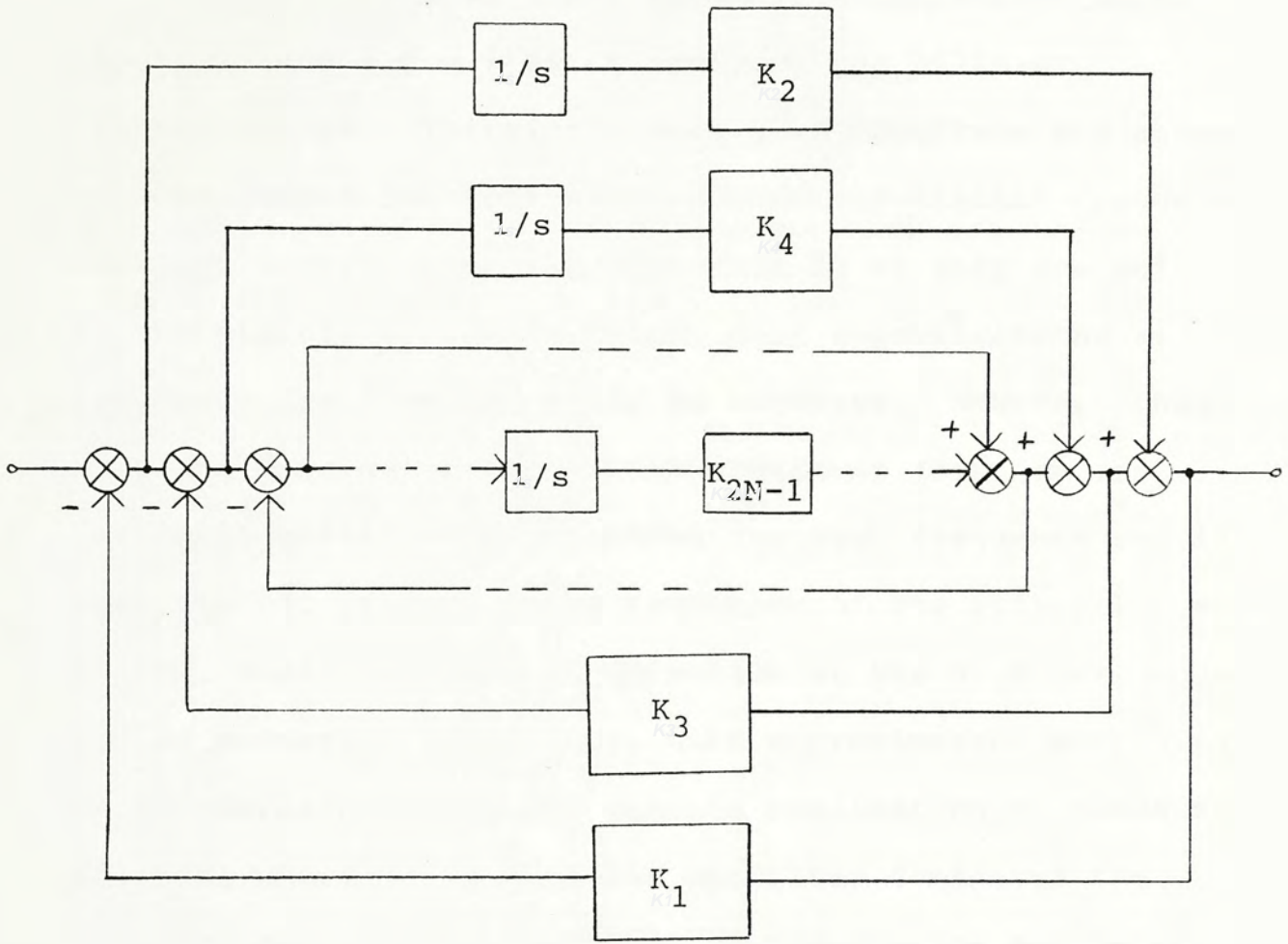Fig. 5.1 Design procedure for recursive realization of noncausal filter

Fig. 5.2 Block diagram of continued fraction expansion.

equivalent to the matching of the first few coefficients of a Taylor series expansion of the transfer function about s = 0 [B5]. This, in the z domain, is equivalent to an expansion about z = 1 after applying the bilinear transformation. Therefore, very good magnitude and phase response approximations are expected for digital frequency close to zero. Since the expansion is at only one point in the digital frequency range, poor approximations at frequency far from the point is expected. However, when more and more quotients of the continued fraction are retained, better approximations for high frequency end is anticipated. For moderate reduction of the original lowpass filter, small stopband attentuation at the high frequency end is expected. Therefore, this approximation method is particularly suitable for cascade realization of noncausal filters, where the causal IIR subfilter dominates the stopband behavior. The saving of computation and delay elements for recursive realization over the nonrecursive one may be large especially for narrowband lowpass filters. This fact is illustrated by examples in section 5.4.

The recursive cascade noncausal filter is then given by

$$H(z) = R(z) H_1(z) \qquad (5.19)$$

where $H_1(z)$ is the causal subfilter.

For highpass filter, the numerator and denominator polynomials of the transfer function are arranged in

descending order as in (5.15) and expanded into continued fraction about $s = \infty$, which with the bilinear transformation is equivalent to the point $z = -1$ in the z domain. Therefore, good approximation in the high frequency end is guaranteed even though it is poor in the low frequency region.

It has been seen that the model reduction technique results in good approximations either in the low frequency end or the high frequency end but not both. If good approximations in both low and high frequency regions are required, it can be achieved by carrying out the continued fraction expansion alternately at $s = 0$ and $s = \infty$ [C9]. That is to say, the first quotient of the continued fraction is obtained from the constant terms of the numerator and denominator polynomial; the second quotient is then computed from the coefficients of the highest order terms, and so on. That is,

$$
F(s) = \frac{f_{21} + f_{22}s + \ldots + f_{2n}s^{n-1}}{f_{11} + f_{12}s + \ldots + f_{1n}s^{n-1}}
$$

$$
= \cfrac{1}{\dfrac{f_{11}}{f_{21}} + s\,\dfrac{f_{3,n-1}s^{n-2} + f_{3,n-2}s^{n-3} + \ldots + f_{31}}{f_{2n}s^{n-1} + f_{2,n-1}s^{n-2} + \ldots + f_{21}}}
$$

$$
= \cfrac{1}{\dfrac{f_{11}}{f_{21}} + \cfrac{s}{\dfrac{f_{2n}}{f_{3,n-1}}s + \dfrac{f_{41} + f_{42}s + \ldots + f_{4,n-1}s^{n-2}}{f_{31} + f_{32}s + \ldots + f_{3,n-1}s^{n-2}}}}
$$

(5.20)

101

By repeating this sequence of expansion, we have

$$F(s) = \cfrac{1}{K_1 + \cfrac{s}{K_2 s + \cfrac{1}{K_3 + \cfrac{s}{K_4 s + \ddots}}}}$$ (5.21)

where

$$K_1 = \frac{f_{11}}{f_{21}} \; , \quad K_2 = \frac{f_{2n}}{f_{3,n-1}} \; , \quad K_3 = \frac{f_{31}}{f_{41}} \; \ldots \; \text{etc.}$$

The reduced model is obtained by retaining the first few quotients of (5.20) and performing the inverse procedures. Since the Taylor series expansion is carried out both at $s = 0$ and $s = \infty$, good approximations are obtained for both low and high frequencies. This method is particularly suitable for bandstop filter approximations since the passbands are located at both low and high frequency ends. A bandstop noncausal filter is designed in Example 3 of section 5.4 to illustrate this fact.

5.4 Design considerations and examples

Before carrying out the design it should be pointed out that the main goal of recursive realization of the purely noncasual subfilter is to save both the number of arithmetic operations and delay elements. This may not always be achieved and, sometimes, nonrecursive realization is more efficient. In order to have a recursive subfilter $R(z)$ more efficient than the FIR subfilter $\overline{F}(z)$, it is

necessary that $N-1 \geq 2m$ where m is the order of $R(z)$.
Thus, it is necessary to reduce the order of the original
transfer function by more than one half and still maintain
a reasonably good approximation in the frequency of interest.
Otherwise, nonrecursive realization should be used instead.

After a frequency specification is given, the
cascade recursive noncausal filter is then designed following
the procedures as shown in Fig. 5.1. The order m of the
recursive approximation $R(z)$ to the purely noncausal transfer
function $H_1(z^{-1})$ is usually also specified. An obvious
question immediately emerges as how to choose the length N
of $\overline{F}(z)$ that will give a $R(z)$ which is a good approximant
of $H_1(z^{-1})$. It is noted that $\overline{F}(z)$ itself is an approximation
of $H_1(z^{-1})$ with a frequency error of $E(e^{j\omega})$ given by (5.5)
Since $R(z)$ is derived from $\overline{F}(z)$, we would expect that $R(z)$.
be, at most, as good as $\overline{F}(z)$. Thus to reduce error, an
instinctive reaction is to increase N enormously so that
$\overline{F}(z)$ converges to $H_1(z^{-1})$. However, since the order m of
$R(z)$ is fixed if the order of $\overline{F}(z)$ is too large, the
approximation will be too poor to be useful. In the other
extreme when N is too small, then even if $R(z)$ is an
excellent approximation to $\overline{F}(z)$, the frequency error $E(e^{j\omega})$
will be unacceptable large. Therefore, it is necessary to
select N so that $R(z)$ is considered to be an acceptable
approximation to $H_1(z^{-1})$. An example is used to illustrate
this fact.

Example 1 :

A recursive lowpass noncausal filter with linear phase is designed to investigate some properties of the proposed method. A 4th order analog elliptic filter with 0.18 dB and 37 dB minimum stopband attenuation is selected as the prototype filter. The digital passband frequency is at 0.1 and a 10th order recursive approximation $R(z)$ to the noncausal filter is desired.

Based on a 30th order FIR approximation, $R(z)$ is derived using the proposed method by retaining 21 quotients. Cascading $R(z)$ with the causal subfilter $H_1(z)$, we obtain the recursive noncausal filter. The magnitude responses of $H(z)$, $R(z)$ and $H_1(z)$ are shown in Fig. 5.3a while the group delay of $H(z)$ in the passband is plotted in Fig. 5.3b. It can be seen that the rational approximation $R(z)$ is excellent in the passband but is poor in the stopband with an attenuation of only 6 dB. Since $R(z)$ is cascaded with $H_1(z)$, the overall magnitude response in the stopband is dominated by $|H_1(e^{j\omega T})|$. The basic group delay of the filter is approximately 30 sample times which, as expected, is similar to the order of the FIR subfilter $\overline{F}(z)$. The percentage group delay ripple in the passband is 3.7%.

Since linear phase is one of the main design criteria, it is interesting to investigate the effect of N, the length of $\overline{F}(z)$, on the passband group delay ripple while keeping the order m of $R(z)$ fixed. The result is plotted in Fig. 5.4. It is noted that an optimal point is reached

when N is equal to 30 which gives a smallest delay ripple of 3.71%. It is observed that for N smaller than 30, the delay ripples for both noncausal filters realized by $\overline{F}(z)$ or $R(z)$ are the same. However, for N larger than 30, the delay ripple of $\overline{F}(z)$ decreases with N while that of $R(z)$ grows rapidly. In this case, this indicates that a 10th order rational approximation is no longer good enough for a FIR subfilter of order larger than 30.

Comparing the recursive realization with the nonrecursive one, $R(z)$ requires 21 real multiplications and 10 delay elements when implemented by canonical form while $\overline{F}(z)$ needs 31 real multiplications and 30 delay elements when direct form is used. Therefore, much savings are obtained by the recursive realization in this example.

To further improve the group delay performance, the recursive filter order m has to be increased. It is interesting to plot the percentage passband group delay ripple against N while using m as a parameter (Fig. 5.5). The solid line indicates the delay ripple when the filter is realized by the FIR subfilter. The dashed lines indicate the results from recursive realization with m as a parameter. For m = 14, the recursive approximation is exact for N less than 40. The delay ripple departs significantly from the FIR ones when N is greater than 40. The magnitude response and group delay characteristics of a filter derived with m = 14 and N = 40 are shown in Fig. 5.6a and b. The magnitude response of $R(z)$ with m = 14 is similar to the
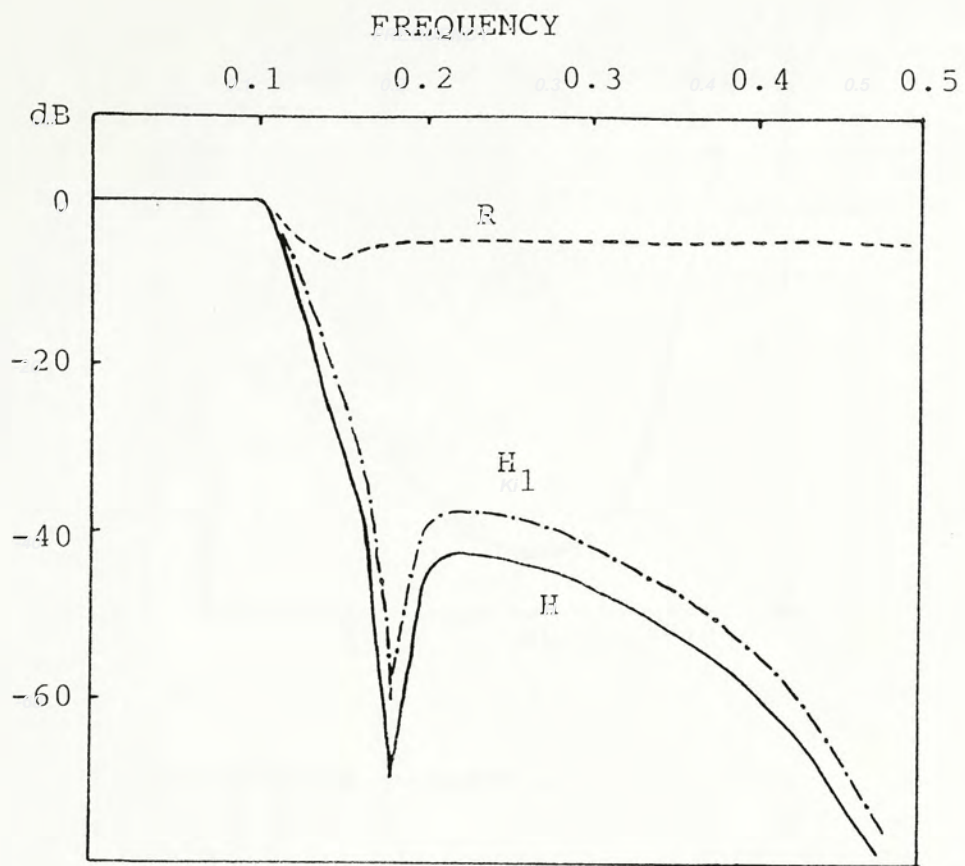
(a)
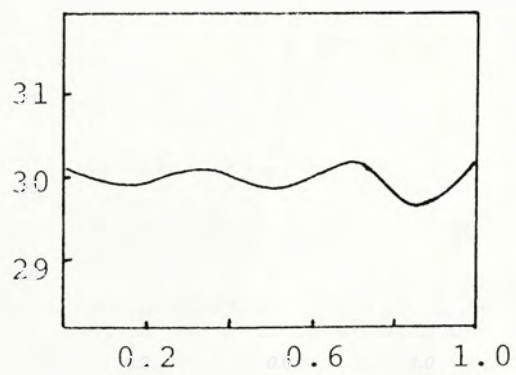
(b)

Fig. 5.3   Lowpass noncausal filter (N = 31, m = 10)

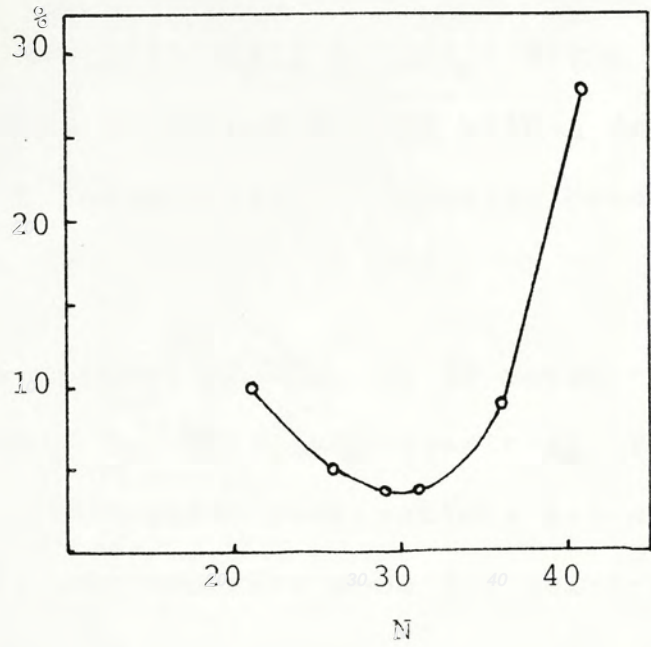(a) magnitude response, (b) group delay.

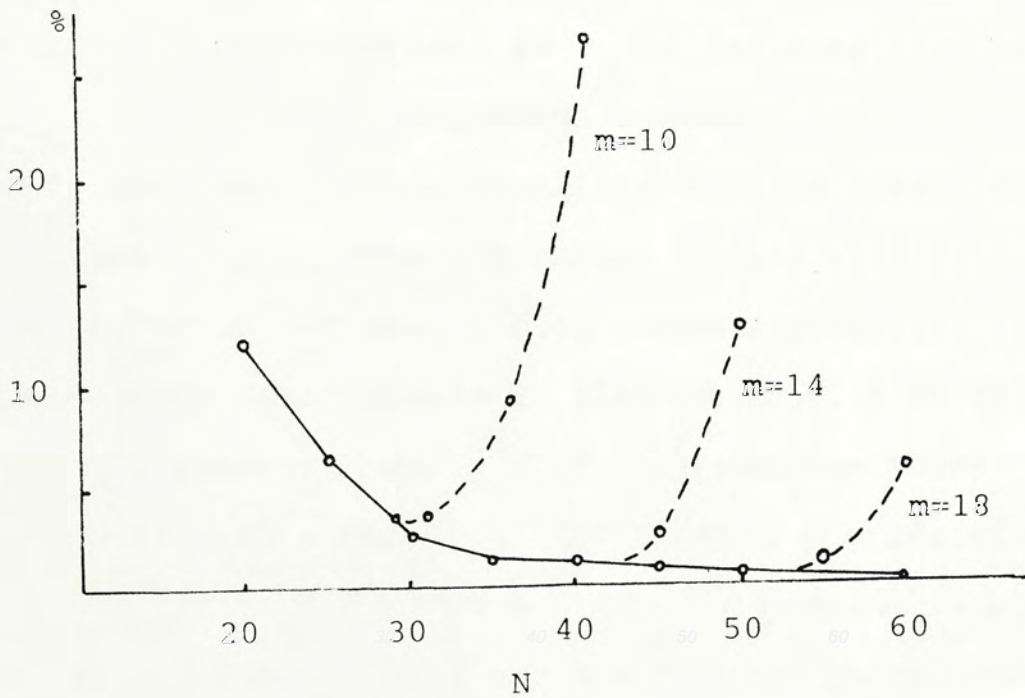Fig. 5.4   Effect of N on passband group delay ripple.

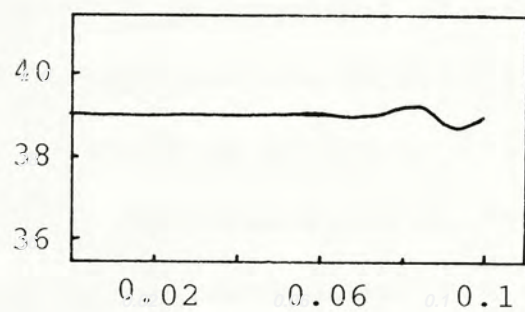Fig. 5.5   Group delay ripple against N with m as a

parameter.

optimal one with m = 10. The delay ripple, however, improves to 1.42% which is significantly better. For m = 18, the optimal point occurs at around N = 52 with a delay ripple of only 0.55%. For N larger than 52, similar behavior is observed.

For the optimal points, it is noted that the ratio N/m is approximately 3. This indicates that, for this particular filter, recursive realizations are always more efficient than the nonrecursive ones for sufficiently good results.

In this example, the passband frequency is at $F_p = 0.1$ which is close to the point of expansion for the continued fraction. Therefore, the number of retained quotients to maintain the passband performance is small. However, it can be forseen that if F_p is far away from zero, the order of recursive approximation required for sufficiently good delay characteristics will be high. To illustrate this point another noncausal filter with $F_p = 0.2$ is designed based on the same analog lowpass prototype filter. The passband group delay ripple is plotted (Fig. 5.7) against N with m as a parameter. For m = 10, the minimum point occurs at N = 18 with a ripple of 4.9%. When m = 14, the minimum ripple is about 2.4% at N = 23. A minimum ripple of 1.5% is achieved for m = 18 and N = 30. At these optimal points the ratio N/m is approximately 1.7. For a N/m ratio less than 2, the nonrecursive realization is in general more efficient than corresponding recursive one. This implies

108
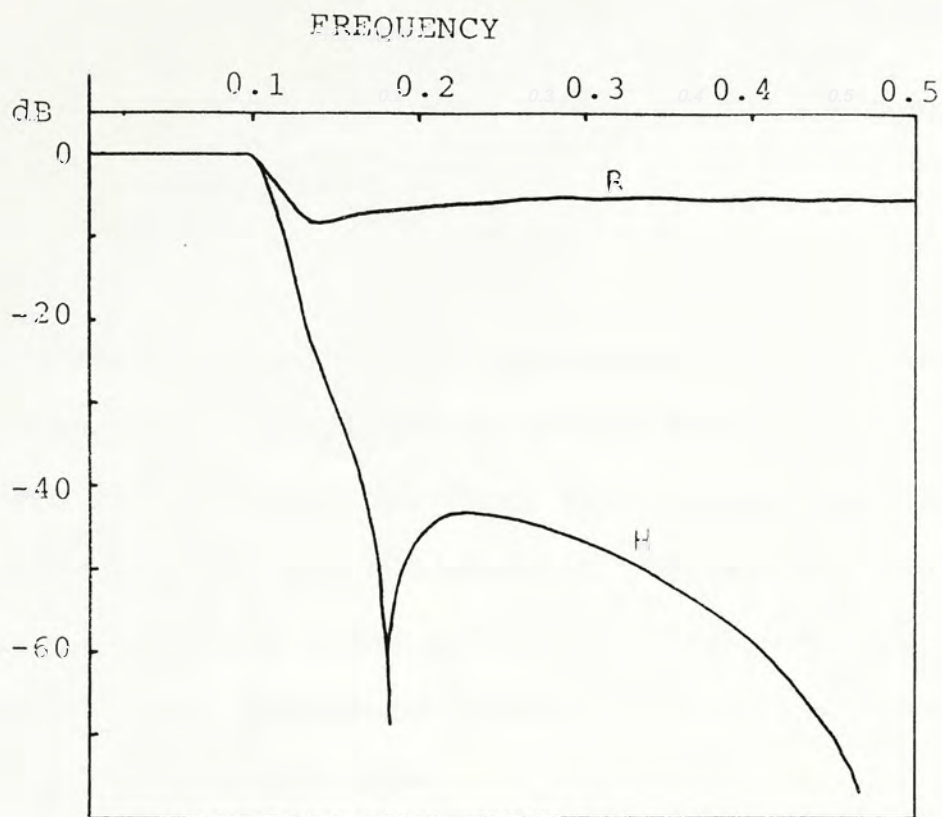
(a)

(b)

Fig. 5.6   Lowpass noncausal filter (N = 40);
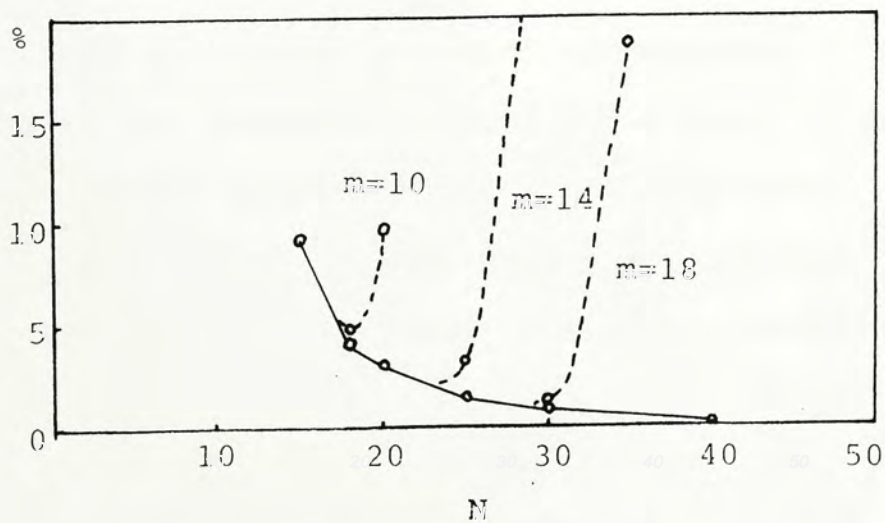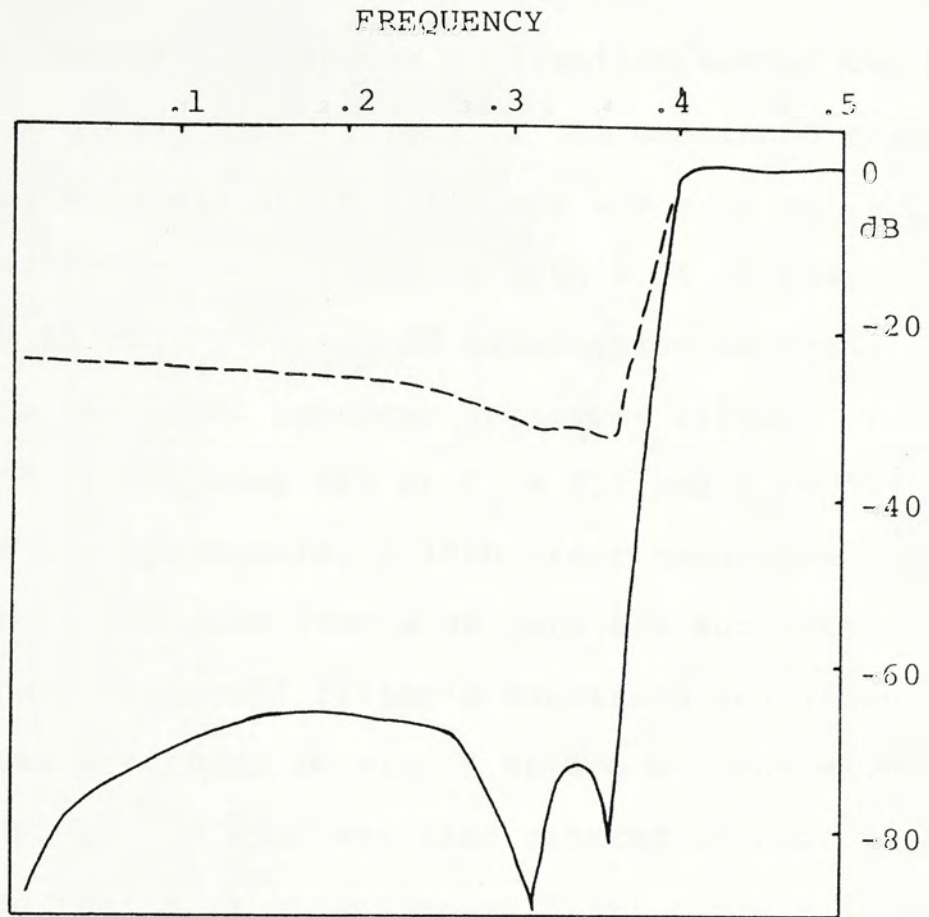(a) magnitude response, (b) group delay.

Fig. 5.7   Group delay ripple against N with m as a

parameter.

that the proposed recursive realization is more efficient
only for narrowband lowpass filter.

Example 2 :

A recursive highpass noncausal filter is designed
to demonstrate that the proposed method can be applied to
the highpass case if the continued fraction is expanded at
$s = \infty$. The numerator and denominator polynomials are then
arranged in descending order before the expansion is carried
out. Thus the same continued fraction expansion program can
be used for the highpass case.

A 5th order highpass elliptic prototype filter is
selected with 40 dB minimum stopband attenuation and 0.18 dB
passband ripple. The IIR and the FIR subfilters with
F_p = 0.4 are then obtained by bilinear transformation method.
A 19th order recursive approximation to the purely noncausal
part is then derved from a 39th order FIR subfilter using
the proposed method for highpass filters. The cascade
filter's magnitude response and passband group delay
characteristics are plotted in Fig. 5.8a and b. The
recursive filter is basically a good approximation of the
FIR subfilter. The passband group delay ripple is 8.2%.
To improve the delay ripple, N has to be increased. In
this example, the filter is essentially a narrowband
highpass one, therefore, efficient recursive realization
is expected.

FREQUENCY

(a)
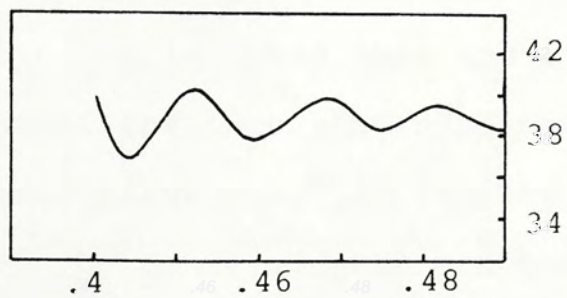

(b)

Fig. 5.8   Lowpass noncausal filter (N=40, m=19);

(a) magnitude response, (b) group delay.

Example 3 :

The proposed recursive realization method can be applied to design bandstop filters if the continued fraction is expanded alternately about s = 0 and s = ∞ as in (5.20). A 4th order elliptic lowpass filter with 0.18 dB passband ripple and 50 dB minimum stopband attenuation is transformed by (4.40) to a 8th order bandstop prototype filter. The desired cutoff frequencies are at F_ℓ = 0.1 and F_u = 0.4. Following the same procedure, a 18th order recursive approximation is obtained from a 30 term FIR subfilter. The linear phase noncausal filter's magnitude and group delay responses are shown in Fig. 5.9a and b. The magnitude responses of $H_1(z)$ and R(z) are also plotted in Fig. 5.9a. It is observed that R(z) approximates $H_1(z)$ closely in both the lower and upper passband. However, the approximation is poor in the stopband. The overall stopband attenuation is mainly due to $H_1(z)$. The group delay is approximately constant in the passband with a ripple of 5.5%.

In this example, it is noted that the ratio N/m is 1.7 which indicates that the recursive realization is not as efficient as the nonrecursive one. In reality, efficient recursive realization can be achieved only for bandshop filter with narrow passbands at both low and high frequencies.

112
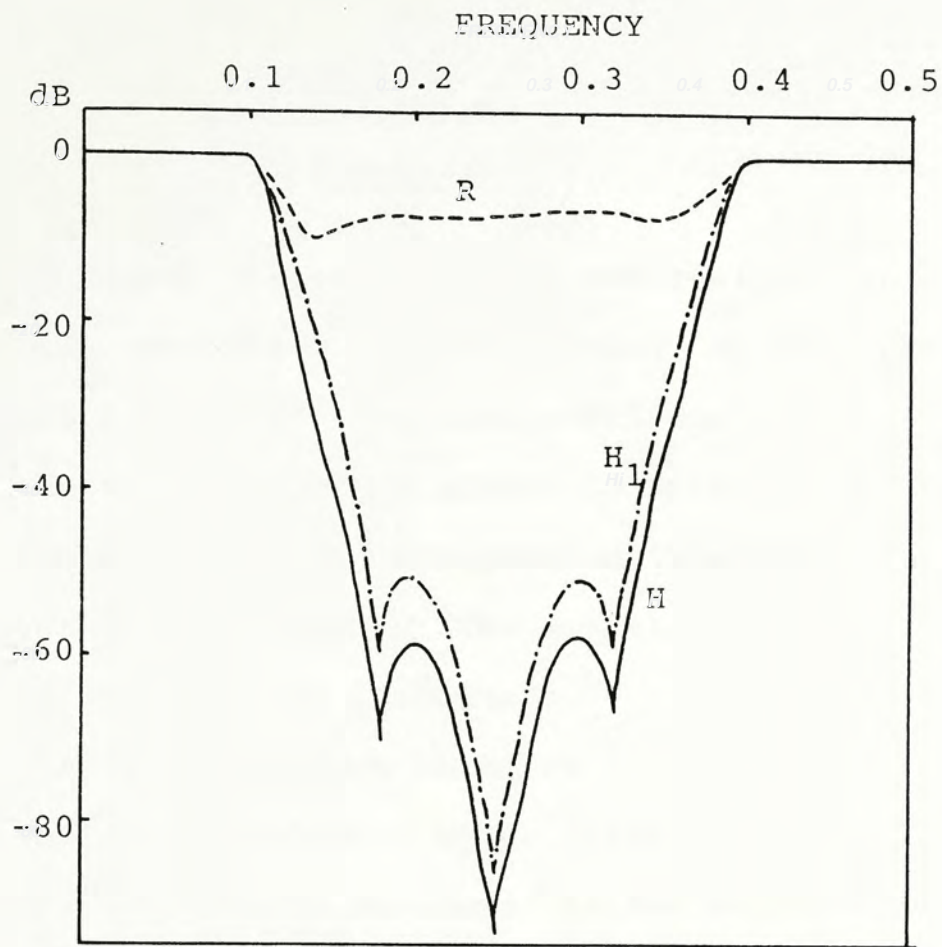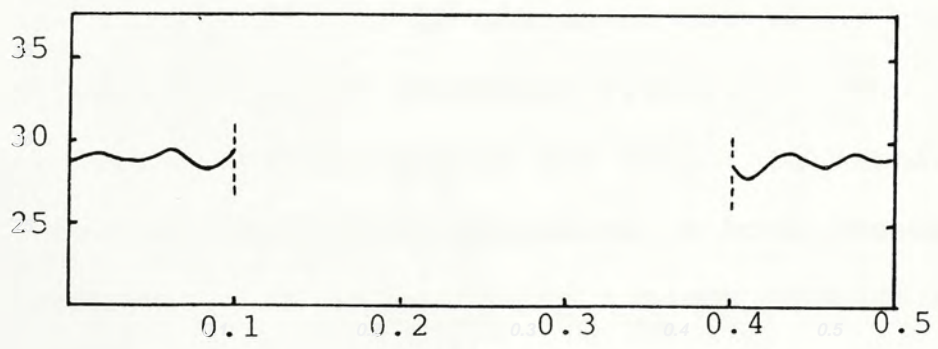
Fig. 5.9 Bandstop noncausal filter (N = 29, m = 18);

(a) magnitude response, (b) group delay.

113

Chapter 6

Conclusion

Noncausal filter synthesis and realization techniques are considered in this thesis. By decompositing the noncausal filter into a causal subfilter and a purely noncausal subfilter connected either in parallel or in series, the noncausal synthesis problem is reduced into a causal filter design problem. The stability criteria are expressed in terms of the subfilters. The decomposition of zero phase noncausal filters turns out to be consisting of identical subfilters for most cases. The frequency response relationships between the noncausal filter and the subfilters are given which greatly facilitate the synthesis process in a practical problem.

Realization is one of the major obstacles that prohibits the application of noncausal filtering. The conventional approach makes use of the fact that sampled signals in the computer can be processed in both forward and reversed directions. It is basically a block processing approach which requires large memory size and very long processing time. A sample-by-sample approach to the realization problem has been developed. Two new methods are introduced. The resulting filter has a small basic group delay and small memory size requirements; thus, it is particularly suitable for applications where fast processing is mandatory.

The purely noncausal part of the noncausal filter is realized by either a nonrecursive filter or a recursive one in the sample-by-sample technique. The nonrecursive realization is based on a FIR filter design method using Wiener-Lee decomposition technique and the unit circle real part function of the digital filter. The real part function can also be applied for numerical evaluation of the inverse z-transform and digital system modelling. The resulting noncausal filter consists of a recursive subfilter and a nonrecursive subfilter. The convergency for the nonrecursive subfilter is guaranteed for many practical filter responses since they satisfy Dirichlet conditions. Many practical examples have been shown for lowpass, highpass, bandpass and bandstop filters. Comparisons between the block processing technique and nonrecursive realization show that much shorter group delay and smaller memory size requirement characterize the latter method while the former is more computation efficient. The proposed method is also more flexible in design and the resulting phase error is exactly known.

The recursive realization is based on the well known continued fraction expansion often used as a model reduction technique in control problems. The recursive approximation to the purely noncausal part is obtained by applying the reduction technique to the original FIR approximation. The resulting filter is linear phase in the passband. For narrowband lowpass or highpass filters, recursive realization is shown to be more computation efficient than the nonrecursive one while the basic group

delay characteristics and passband behavior are preserved. If the continued fraction is expanded about s=0 and s=∞ alternately, good approximations are obtained in both low and high frequencies. Examples are presented to point out the characteristics of the recursive realization technique.

Since many ideal filters, often referred as unachievable standards, are noncausal in nature, the development in noncausal filter synthesis and realization techniques are important to provide solutions for many application areas. It is hoped that this work would generate more attentions to the relatively undeveloped field of noncausal filtering.

REFERENCES

[A1]   Antoniou A., Digital Filters: Analysis and Design,
        McGraw Hill, 1979.

[B1]   Burrus C. S. and Parks T. W., "Time domain design
        of recursive digital filters," IEEE Trans. Audio
        Electroacoust., vol. AU-18, pp. 137-141, June 1970.

[B2]   Brophy F. and Salazar A. C., "Recursive digital
        filter synthesis in the time domain," IEEE Trans.
        Acoust., Speech, Signal Processing, vol. ASSP-22,
        no. 1, pp. 45-55, Feb. 1974.

[B3]   Brophy F. and Salazar A. C., "Considerations of the
        Pade approximant technique in the synthesis of
        recursive digital filters," IEEE Trans. Audio
        Electroacoust., vol. AU-21, pp. 500-505, Dec. 1973.

[B4]   Bosley M. J., Kropholler H. W. and Lees F. P., "On
        the relation between the continued fraction expansion
        and moments matching methods of model reduction,"
        Int. J. Control, vol. 18, no. 3 pp. 461-474, 1973.

[B5]   Blinchikoff H. J. and Zverev A. I., Filtering in the
        Time and Frequency Domains, John Wiley & Sons, 1976.

[C1]   Czarnach R., "Recursive processing by noncausal
        digital filters," IEEE Trans. Acoust., Speech,
        Signal Processing, vol. ASSP-30, no. 3, pp. 363-370,
        June 1982.

[C2] Chen C. F. and Haas I. J., Elements of Control
      Systems Analysis, Prentice Hall, New Jersey, 1968,
      pp. 194-203.

[C3] Chen C. F. and Tsay Y. T., "A new formula for the
      discrete-time system stability test," Proceedings
      of IEEE, 65, pp. 1200-1202, 1977.

[C4] Cappelini V., Constantinides A. G. and Emiliani P.,
      Digital Filters and Their Applications, 1978.

[C5] Chen C. F. and Shieh L. S., "A novel appraoch to
      linear model simplification," Int. J. Control, vol. 8,
      no. 6, pp. 561-570, 1968.

[C6] Chen C. F., Chang C. Y. and Hau K. W., "Model
      reduction using the stability equation method and
      the continued-fraction method," Int. J. Control,
      vol. 32, no. 1, pp. 81-94, 1980.

[C7] Chen C. T., One-Dimensional Digital Signal Processing,
      Marcel Dekker, Inc., N. Y., 1979.

[C8] Cuthbert L. G., "Optimizing non-recursive digital
      filter to non-linear phase characteristics," The
      Radio and Electronic Engineer, 44, no. 12, pp. 645-51,
      Dec. 1974.

[C9] Chuang S. C., "Application of continued-fraction
      method for modelling transfer functions to give
      more accurate initial transient response," Electronics
      Letters, vol. 6, no. 26, pp. 861-863, Dec. 31, 1970.

[D1]   Dutta Roy S. C., "Comments on 'On the construction
       of a digital transfer function from its real part
       on unit circle'," Proceedings of IEEE, vol. 71,
       no. 8, pp. 1009-1010, Aug. 1983.

[F1]   Friedman D. H., "On approximating an FIR filter
       using discrete orthonormal exponentials," IEEE Trans.
       Acoust., Speech, Signal Processing, vol. ASSP-29,
       no. 4, pp. 923-926, Aug. 1981.

[G1]   Guillemin E. A., Synthesis of Passive Networks, John
       Wiley & Sons, New York, 1977, pp. 279-321.

[G2]   Gold B. and Rader C. M., Digital Processing of
       Signals, McGraw-Hill, Inc., 1969.

[G3]   Goldberg E., Kurshan R. and Malah D., "Design of
       finite impulse response digital filters with nonlinear
       phase response," IEEE Trans. Acoust., Speech, Signal
       Processing, vol. ASSP-29, no. 5, pp. 1003-1010,
       Oct. 1981.

[H1]   Hastings-James R. and Mehra S. K., "Extensions of
       the Pade-Approximant technique for the design of
       recursive digital filters," IEEE Trans. Acoust.,
       Speech, Signal Processing, vol. ASSP-25, no. 6,
       pp. 501-509, Dec. 1977.

[H2]   Holt A. G. J., Attikiouzel J. and Bennett R.,
        "Iterative technique for designing non-recursive
        digital filter to non-linear phase characteristics,"
        The Radio and Electronic Engineer, vol. 46, no. 12,
        pp. 589-592, Dec. 1976.

[H3]   Herrmann O. and Schuessler H. W., "Design of
        nonrecursive digital filters with minimum phase,"
        Electronics Letters, vol. 6, pp. 329-330, May 28, 1970.

[K1]   Kormylo J. J. and Jain V. K., "Two-pass recursive
        digital filter with zero phase shift," IEEE Trans.
        Acoust. Speech, Signal Processing, vol. ASSP-22,
        pp. 384-387, Oct. 1974.

[K2]   Kwong C. P., "Simple method of computation of Wiener-
        Lee decomposition," Electronics Letters, vol. 19,
        no. 18, pp. 747-748, September 1983.

[K3]   Kwong C. P., "Model reduction approach to digital
        recursive filter design," Proc. IERE (Hong Kong
        Section) Workshop on Advanced Microprocessor and
        Digital Signal Processing, pp. 56-65, September 1982.

[L1]   Lee Y. W., Statistical Theory of Communication,
        John Wiley & Sons, New York, 1960, pp. 487-491.

[L2]   Lim Y. C. and Parker S. R., "FIR filter design over
        a discrete powers-of-two coefficient space," IEEE
        Trans., Speech, Signal Processing, vol. ASSP-31,
        no. 3, pp. 583-591, June 1983.

[M1] McGillem C. D. and Cooper G. R., Continuous and
      Discrete Signal and System Analysis, Holt, Rinehart
      & Winston, 1974.

[O1] Oppenheim A.V. and Schafer R. W., Digital Signal
      Processing, Prentice Hall, New Jersey, 1975.

[O2] Oppenheim A. V., Applications of Digital Signal
      Processing, Prentice-Hall, Englewood, N.J., 1978.

[P1] Patney R. K. and Dutta Roy S. C., "Design of linear-
      phase FIR filters using pseudo-Boolean methods,"
      IEEE Trans. Circuits and Systems, vol. CAS-26, no. 4,
      April 1979.

[R1] Rabiner L. R., Kaiser J. F., Herrmann O. and Dolan
      M. T., "Some comparisons between FIR and IIR digital
      filters," Bell Syst. Tech. J., vol. 53, pp. 305-331,
      Feb. 1974.

[R2] Rabiner L. R. "Techniques for designing finite-
      duration impulse-response digital filters," IEEE
      Trans. Commun. Technol., vol. COM-19, pp. 188-195,
      April 1971.

[R3] Rabiner L. R. and Gold B., Theory and Application of
      Digital Signal Processing, Prentice Hall, New Jersey,
      1975.

[S1] Steiglitz K., "Design of FIR digital phase networks,"
      IEEE Trans. Acoust., Speech, Signal Processing,
      vol. ASSP-29, no. 2, pp. 171-176, April 1981.

[S2] Sallai G. Y., "Approximation of FIR Digital Filter by Bilinear Transformation," Int. J. Cir. Theor. Appl., vol. 7, pp. 267-275, 1979.

[S3] Su K. L., Time Domain Synthesis of Linear Networks, Prentice Hall, New Jersey, 1971.

[S4] Shieh L. S. and Goldman M. J., "Continued fraction expansion and inversion of the Cauer third form," IEEE Trans. Circuits and Systems, vol. CAS-21, no. 3, pp. 341-345, May 1974.

[T1] Temes G. C. and Lapatra J. W., Circuit Synthesis and Design, McGraw Hill, 1977.

[V1] Vaidyanathan P. P. and Mitra S. K., "On the construction of a digital transfer function from its real part on unit circle," Proceedings of IEEE, vol. 70, no. 2, pp. 198-199, Feb. 1982.

[V2] Van Valkenburg M. E., Analog Filter Design, Holt, Rinehart and Winston, 1982.

[Y1] Yahagi T., "New methods for the design of recursive digital filters in the time domain," IEEE Trans. Acoust., Speech, Signal Proceeding, vol. ASSP-29, no. 2, pp. 245-254, April 1981.

[Z1] Zverev A. I., Handbook of Filter Synthesis, Wiley, New York, 1967.