

Chapter 5: Discrete Filter Structure

- Realization of LTI-DT systems in either software or hardware (FIR and IIR)
- Cascade, parallel & lattice structures are of particular importance which exhibit robustness in finite-word length implementation.
- Quantization effects in the implementation of digital filters using finite precision arithmetic.

Structures for the realization of Discrete-time systems:

An LTI DT system are characterized by general LCCD equation

$$y[n] = - \sum_{k=1}^N a_k y[n-k] + \sum_{k=0}^M b_k x[n-k]$$

Taking z-transform and characterizing in the rational system function,

$$H(z) = \frac{\sum_{k=0}^M b_k z^{-k}}{1 + \sum_{k=1}^N a_k z^{-k}}$$

In this characterization, we obtain the zeros and poles of the system function, which depend on the choice of the system parameters $\{b_k\}$ and $\{a_k\}$ & which determine the frequency response characteristics of the system.

Major factors that influence the choice of specific realization

- Computational complexity (no. of arithmetic operations)
- Memory requirements (no. of locations required to store system parameters, i/p, o/p, etc)
- Finite-word-length effects in the computations (quantization effects)

A. Structure for FIR systems:

$$y[n] = \sum_{k=0}^{M-1} b_k x[n-k]$$

Or, the system function

$$H(z) = \sum_{k=0}^{M-1} b_k z^{-k}$$

The unit sample response of the FIR system is identical to the coefficients (b_k), that is,

$$h[n] = \begin{cases} b_n, & 0 \leq n \leq M-1 \\ 0, & \text{otherwise} \end{cases}$$

1. Direct Form Structure:

$$y[n] = \sum_{k=0}^{M-1} h[k]x[n-k]$$

$$\text{or, } y[n] = \sum_{k=0}^{M-1} b_k x[n-k]$$

This structure requires

- $M-1$ memory locations
- M multiplications & $M-1$ additions per output point.

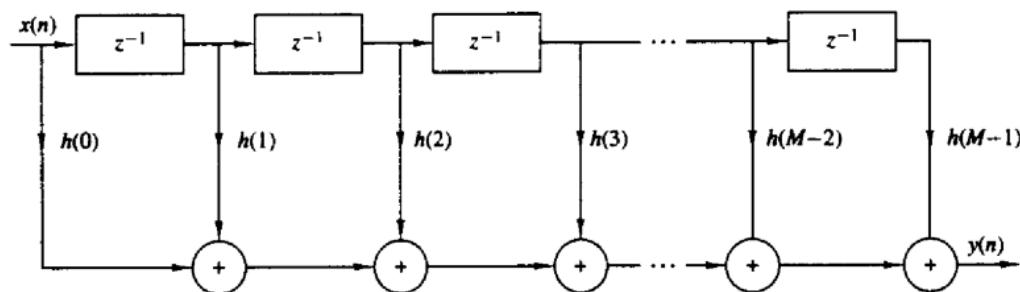


Figure 1: Direct Form Structure of FIR system.

This structure is often called a *transversal* or *tapped-delay line filter*.

2. Cascade-Form Structure:

The cascade realization follows from difference equation,

$$H(z) = \prod_{k=0}^{M-1} b_k z^{-k}$$

Factorizing in 2nd order FIR systems so that,

$$H(z) = \prod_{k=1}^K H_k(z), \quad k = 1, 2, 3, \dots, K$$

$$H_k(z) = b_{k0} + b_{k1}z^{-1} + b_{k2}z^{-2}$$

and K is the integer part of $(M + 1)/2$

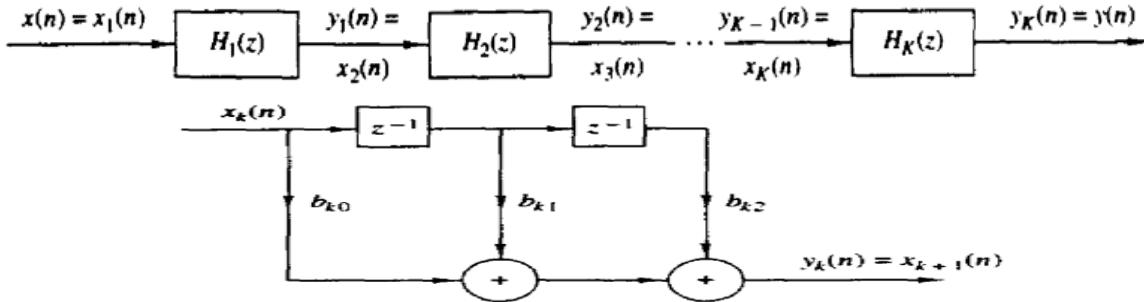


Figure 2: Cascade Realization of an FIR System

3. Frequency-Sampling Structures:

In frequency sampling realization, the parameters that characterize the filter are the values of the desired frequency response instead of the impulse response $h[n]$. To derive the frequency-sampling structure, we specify the desired frequency response at a set equally spaced frequencies,

$$\begin{aligned}\omega_c &= \frac{2\pi}{M}(k + \alpha), \quad k = 0, 1, 2, \dots, \frac{M-1}{2}, \quad M \text{ odd} \\ k &= 0, 1, 2, \dots, \frac{M}{2}-1, \quad M \text{ even}, \quad \alpha = 0 \text{ or } \frac{1}{2}\end{aligned}$$

And solve for the unit sample response $h[n]$ from these equally spaced frequency specifications.

Thus we can write the frequency response as

$$H(\omega) = \sum_{n=0}^{M-1} h[n] e^{-j\omega n}$$

And the values of $H(\omega)$ at frequencies $\omega_k = (2\pi/M)(k + \alpha)$ are simply

$$H(k + \alpha) = H\left(\frac{2\pi}{M}(k + \alpha)\right) = \sum_{n=0}^{M-1} h[n] e^{-j\frac{2\pi}{M}(k + \alpha)n}, \quad k = 0, 1, 2, \dots, M - 1$$

The set of values $\{H(k + \alpha)\}$ are called the frequency samples of $H(\omega)$. In the case where $\alpha = 0$, $\{H(k)\}$ corresponds to the M-point DFT of $\{h[n]\}$.

It is a simple matter to invert above equation and express $h[n]$ in terms of the frequency samples.

$$h[n] = \sum_{k=0}^{M-1} H(k + \alpha) e^{j\frac{2\pi}{M}(k + \alpha)n}, \quad n = 0, 1, 2, \dots, M - 1$$

when $\alpha = 0$, it is simply the IDFT of $\{H(k)\}$. Now taking z-transform, we have

Digital Signal Analysis & Processing

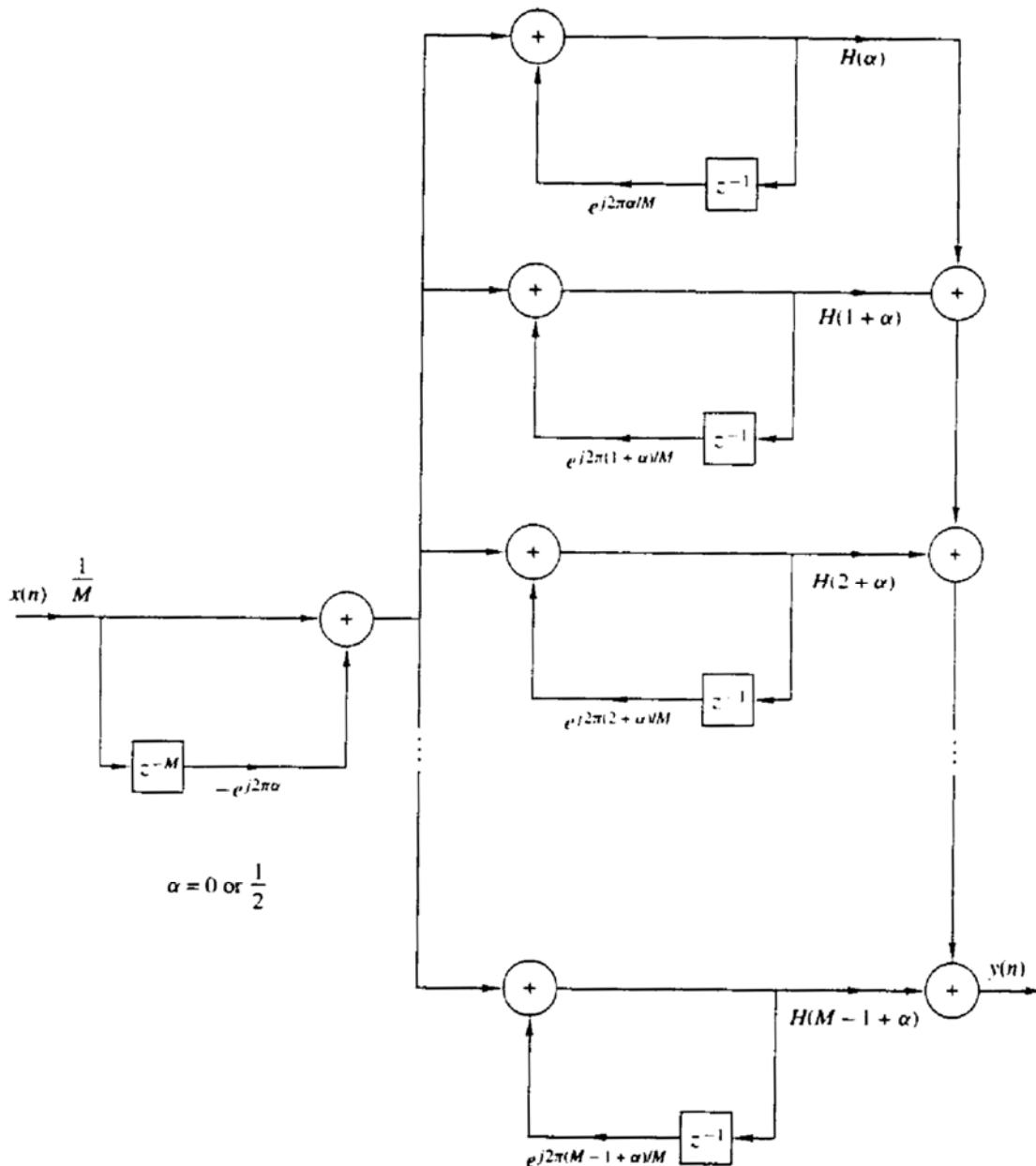
Chapter 5: Discrete Filter Structure

$$H(z) = \sum_{n=0}^{M-1} h[n]z^{-n} = \sum_{n=0}^{M-1} \left[\sum_{k=0}^{M-1} H(k + \alpha) e^{j\frac{2\pi}{M}(k+\alpha)n} \right] z^{-n}$$

By interchanging the order of the two summations and performing the summation over the index n, we obtain

$$H(z) = \sum_{k=0}^{M-1} H(k + \alpha) \left[\frac{1}{M} \sum_{n=0}^{M-1} \left(e^{j\frac{2\pi}{M}(k+\alpha)} z^{-1} \right)^n \right] = \frac{1 - z^{-M}}{M} \sum_{k=0}^{M-1} \frac{H(k + \alpha)}{1 - e^{\frac{j2\pi(k+\alpha)}{M}} z^{-1}}$$

Thus the system function is characterized by the set of frequency samples $\{H(k + \alpha)\}$ instead of $\{h[n]\}$. The realization is illustrated in figure below:



4. Lattice structure of FIR System:

Lattice filters are used extensively in digital speech processing and in the implementation of adaptive filters.

Let us consider FIR system functions

$$H(z) = A_m(z), \quad m = 0, 1, 2, \dots, M-1 \quad \text{--- I}$$

Where, $A_m(z)$ is the polynomial,

$$A_m(z) = \sum_{k=0}^m \alpha_m(k)z^{-k}, \quad m \geq 1 \quad \text{--- II}$$

The subscript m on the polynomial $A_m(z)$ denotes the degree of the polynomial. For mathematical convenience, we define $\alpha_m(0) = 1$. If $\{x[n]\}$ is input sequence to the filter $A_m(z)$ & $\{y[n]\}$ is the output sequence, we have,

$$y[n] = x[n] + \sum_{k=1}^m \alpha_m(k)x[n-k] \quad \text{--- III}$$

Now, suppose that we have a filter of order 1 i.e. $m=1$, then

$$y(n) = x(n) + \alpha_1(1)x(n-1) \quad \text{--- IV}$$

This output can be obtained from a 1st order or single-stage lattice filter, as in figure below:

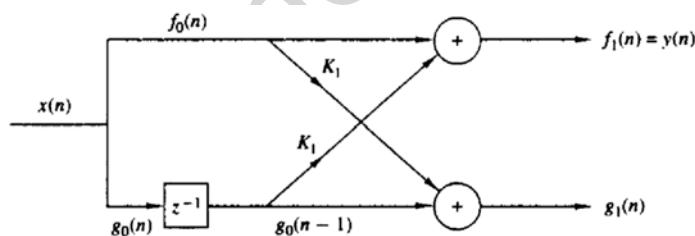


Figure 3: Single Stage Lattice Filter.

Where $k_1=\alpha_1(1)$ is called a reflection coefficient. From Figure 3,

$$\begin{aligned} f_0(n) &= g_0(n) = x[n] \\ f_1(n) &= f_0(n) + k_1 g_0(n-1) = x(n) + k_1 x(n-1) \\ g_1(n) &= k_1 f_0(n) + g_0(n-1) = k_1 x(n) + x(n-1) \end{aligned} \quad \text{--- V}$$

Digital Signal Analysis & Processing

Chapter 5: Discrete Filter Structure

Now, for m=2,

$$y(n) = x(n) + \alpha_2(1)x(n-1) + \alpha_2(2)x(n-2) \dots VI$$

By cascading two lattice stages as shown in Figure 4, It is possible to obtain the output y(n)

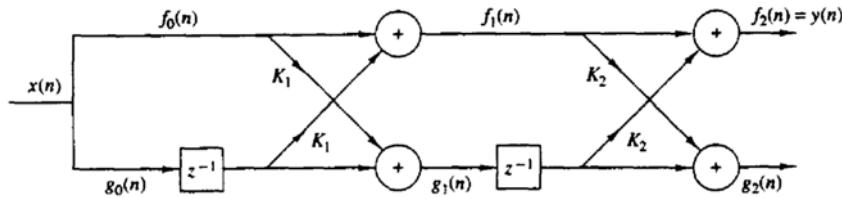


Figure 4: Two stage-lattice filter.

Output from 2nd stage is

$$\begin{aligned} f_2(n) &= f_1(n) + k_2 g_1(n-1) \\ g_2(n) &= k_2 f_1(n) + g_1(n-1) \end{aligned} \dots VII$$

From equations V and VII,

$$\begin{aligned} f_2(n) &= x(n) + k_1 x(n-1) + k_2 \{k_1 x(n-1) + x(n-2)\} \\ i.e. y(n) &= x(n) + k_1(1+k_2)x(n-1) + k_2 x(n-2) \end{aligned} \dots VIII$$

Now equating equations VI and VIII,

$$\begin{aligned} \alpha_2(1) &= k_1(1+k_2); \alpha_2(2) = k_2 \\ i.e. k_1 &= \frac{\alpha_2(1)}{1+k_2} = \frac{\alpha_2(1)}{1+\alpha_2(2)} \end{aligned} \dots IX$$

Now by method of induction,

$$\begin{aligned} f_m(n) &= f_{m-1}(n) + k_m g_{m-1}(n-1) \\ g_m(n) &= k_m f_{m-1}(n) + g_{m-1}(n-1) \end{aligned} \dots X$$

Hence, (M-1) stage lattice filter is,

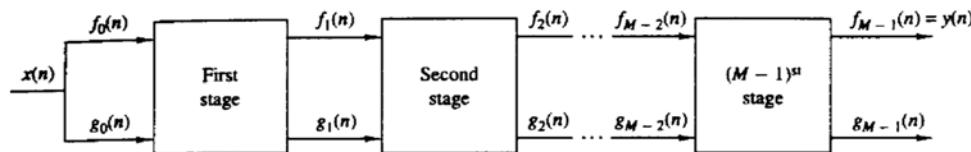


Figure 5: (M-1) stage lattice

Then, the output of the (M-1) stage filter corresponds to the output of an (M-1) order FIR , i.e.

$$y[n] = f_{M-1}[n]$$

As a consequence of the equivalence between an FIR filter & a lattice filter, the output $f_m(n)$ of an m-stage lattice filter can be expressed as,

$$f_m(n) = \sum_{k=0}^m \alpha_m(k)x(n-k), \alpha_m(0) = 1 \dots XI$$

Digital Signal Analysis & Processing

Chapter 5: Discrete Filter Structure

which is convolution sum. In z-domain,

$$F_m(z) = A_m(z)X(z)$$
$$\text{or, } A_m(z) = \frac{F_m(z)}{X(z)} = \frac{F_m(z)}{F_0(z)} \quad \dots \text{--- XII}$$

The other output component from the lattice, namely $g_m(n)$, can also be expressed in the form of a convolution sum as,

$$g_m(n) = \sum_{k=0}^m \beta_m(k)x(n-k), \quad \dots \text{--- XIII}$$

where, the filter coefficients $\{\beta_m(k)\}$ are associated with a filter that produces $f_m(n) = y(n)$ but operates in reverse order.

i.e $\beta_m(k) = \alpha_m(m-k)$, $k = 0, 1, 2, \dots, m$ with $\beta_m(m) = 1$.

In z-domain,

$$B_m(z) = \sum_{k=0}^m \beta_m(k)z^{-k} = \sum_{k=0}^m \alpha_m(m-k)z^{-k}$$

Taking $m - k = l$,

$$B_m(z) = \sum_{l=m}^0 \alpha_m(l)z^{-(m-l)} = z^{-m}A_m(z^{-1})$$

This shows that the zeros of $B_m(z)$ are simply the reciprocals of zeros of $A_m(z)$. i.e. $B_m(z)$ is reverse polynomials of $A_m(z)$.

Let us write the lattice equation in z-domain,

$$F_0(z) = G_0(z) = X(z)$$
$$F_m(z) = F_{m-1}(z) + k_m z^{-1}G_{m-1}(z)$$
$$G_m(z) = k_m F_{m-1}(z) + z^{-1}G_{m-1}(z)$$

If we divide each equation by $X(z)$,

$$A_0(z) = B_0(z) = 1$$
$$A_m(z) = A_{m-1}(z) + k_m z^{-1}B_{m-1}(z)$$
$$B_m(z) = k_m A_{m-1}(z) + z^{-1}B_{m-1}(z)$$

- Characterization of class-m FIR filters in direct form requires $m(m+1)/2$ filter coefficients, $\{\alpha_m(k)\}$, the lattice-form characterization requires only the m reflection coefficients $\{k_i\}$.
- The addition of stages to the lattice does not alter the parameters of the previous stages.

I. Conversion of lattice coefficient to direct form filter coefficients:

The direct form FIR filter coefficients $\{\alpha_m(k)\}$ can be obtained from the lattice coefficients $\{k_i\}$ by using the following relations:

$$\begin{aligned}
 A_0(z) &= B_0(z) = 1 \\
 A_m(z) &= A_{m-1}(z) + k_m z^{-1} B_{m-1}(z) \\
 B_m(z) &= k_m A_{m-1}(z) + z^{-1} B_{m-1}(z) \\
 &= z^{-m} A_m(z^{-1}), m = 1, 2, 3, \dots, M-1 \\
 A_m(z) &= 1 + \sum_{k=1}^m \alpha_m(k) z^{-k}
 \end{aligned}$$

II. Conversion of direct form FIR coefficient to lattice coefficients

Suppose, we are given the polynomial $A_m(z)$ and we wish to determine the corresponding lattice filter parameters $\{k_i\}$. For the m -stage lattice we immediately obtain the parameter $k_m = \alpha_m(m)$. To obtain k_{m-1} we need the polynomial $A_{m-1}(z)$.

We have

$$\begin{aligned}
 A_m(z) &= A_{m-1}(z) + k_m z^{-1} B_{m-1}(z) \quad \dots i \\
 B_m(z) &= k_m A_{m-1}(z) + z^{-1} B_{m-1}(z) \quad \dots ii
 \end{aligned}$$

From (i) and (ii)

$$A_m(z) = A_{m-1}(z) + k_m z^{-1} \{B_m(z) - k_m A_{m-1}(z)\}/z^{-1}$$

After some calculation

$$A_{m-1}(z) = \frac{A_m(z) - k_m B_m(z)}{1 - k_m^2}$$

Alternative formula is given by

$$\begin{aligned}
 k_m &= \alpha_m(m) \\
 \alpha_{m-1}(k) &= \frac{\alpha_m(k) - \alpha_m(m) \alpha_m(m-k)}{1 - \alpha_m^2(m)}, m = M-1, M-2, \dots, 1
 \end{aligned}$$

B. Structures for IIR systems:

$$y[n] = - \sum_{k=1}^N a_k y[n-k] + \sum_{k=0}^M b_k x[n-k]$$

In Z-domain,

$$\begin{aligned}
 H(z) &= \frac{\sum_{k=0}^M b_k z^{-k}}{1 + \sum_{k=1}^N a_k z^{-k}} = H_1 H_2 \\
 H_1(z) &= \sum_{k=0}^M b_k z^{-k} \rightarrow \text{All-zero FIR System}
 \end{aligned}$$

$$H_2(z) = \frac{\sum_{k=0}^M b_k z^{-k}}{1 + \sum_{k=1}^N a_k z^{-k}} \rightarrow \text{All-pole IIR System}$$

1. Direct form structure

$$y[n] = - \sum_{k=1}^N a_k y[n-k] + \sum_{k=0}^M b_k x[n-k]$$

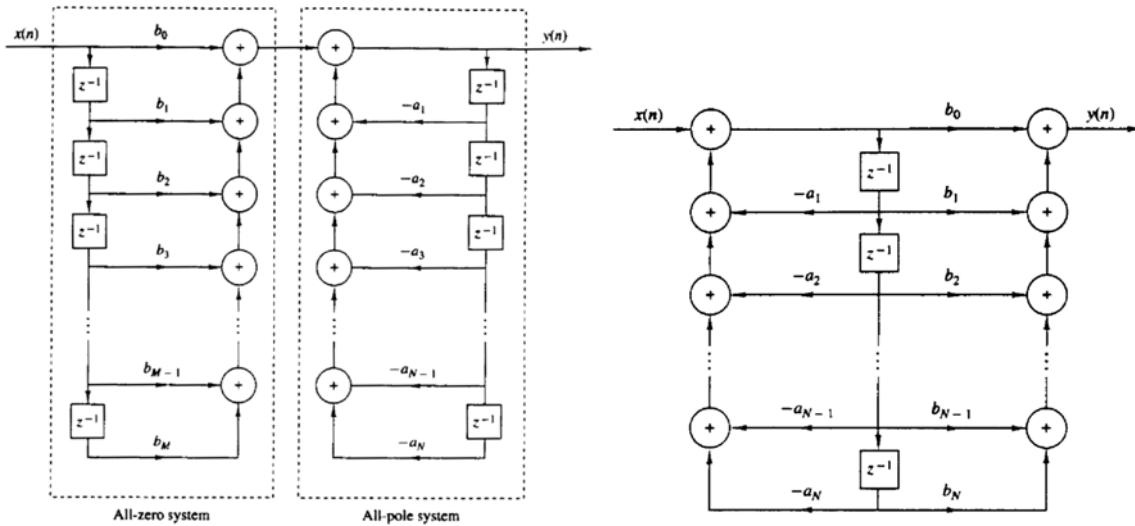


Figure : (a): Direct form I (b) Direct form II ($N=M$)

2. Cascade form structure:

The system can be factored into a cascade of 2nd order subsystems, such that $H(z)$ can be expressed as,

$$H(z) = \prod_{k=1}^L H_k(z) \quad (\text{assume } N \geq M)$$

Where L is the integer part of $(N+1)/2$.

$H_k(z)$ has the general form,

$$H_k(z) = \frac{b_{k0} + b_{k1}z^{-1} + b_{k2}z^{-2}}{1 + a_{k1}z^{-1} + a_{k2}z^{-2}}$$

The coefficients $\{a_{ki}\}$ and $\{b_{ki}\}$ in the 2nd order subsystems are real. This implies that in forming the second order subsystems we should group together a pair of complex-conjugate poles and we should group together a pair of complex-conjugate zeros.

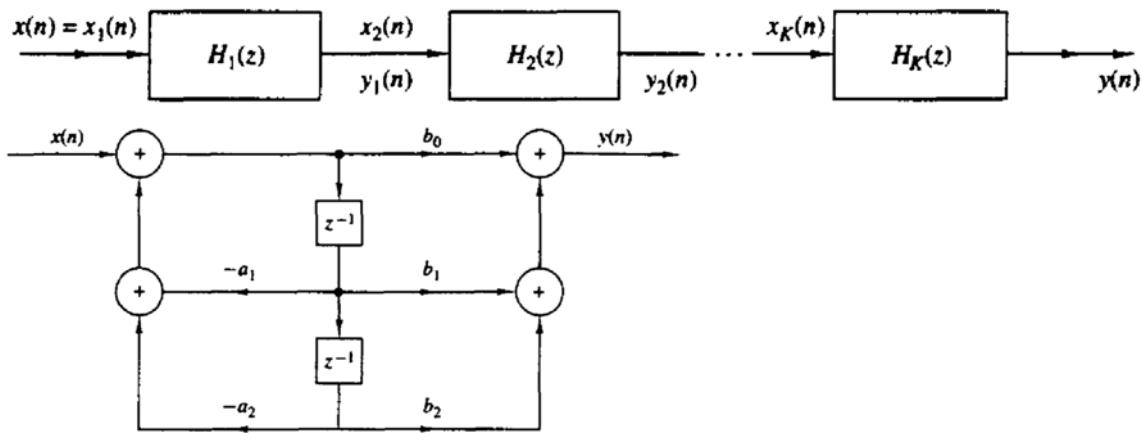


Figure 6: (a) Cascade of 2nd -order subsystem (b) Realization of each 2nd-order section.

3. Parallel form structure

A parallel form realization of an IIR system can be obtained by performing a partial-fraction expansion of $H(z)$.

$$H(z) = C + \sum_{k=1}^N \frac{A_k}{1 - p_k z^{-1}}$$

Where $\{p_k\}$ are the poles, $\{A_k\}$ are the coefficients (or residues) in the partial fraction expansion, and the constant $C = b_N/a_N$.

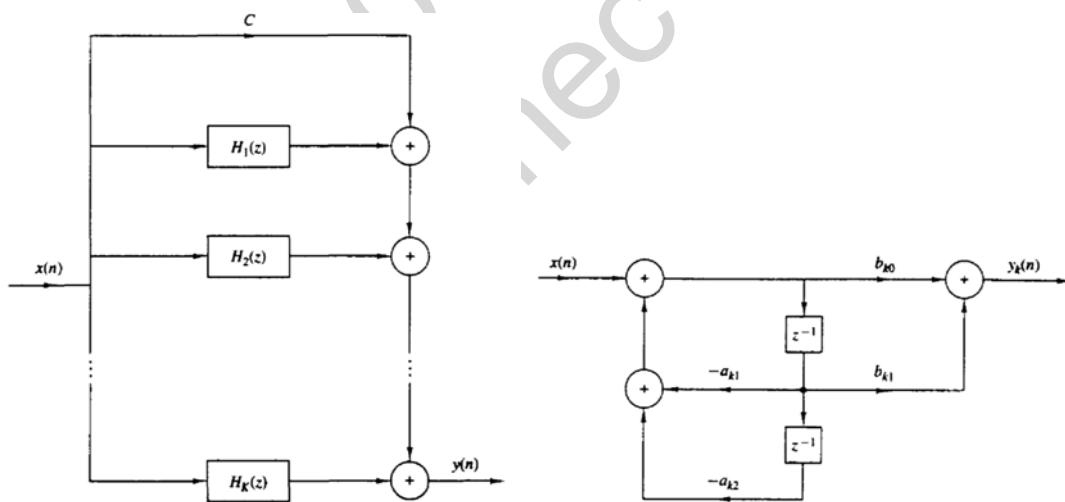


Figure 7: (a) Parallel structure of IIR system. (b) Structure of 2nd -order section in a parallel IIR system realization.

In general, some of the poles of $H(z)$ may be complex valued. In such a case, the corresponding coefficients A_k are also complex valued. To avoid multiplications by complex numbers, we can combine pairs of complex-conjugate poles to form two-pole subsystems. In addition, we can combine, in an arbitrary manner, pairs of real-valued poles to form two-pole subsystems. Each of these subsystems has the form

$$H_k(z) = \frac{b_{k0} + b_{k1}z^{-1}}{1 + a_{k1}z^{-1} + a_{k2}z^{-2}}$$

Where the coefficients $\{b_{ki}\}$ and $\{a_{ki}\}$ are real-valued system parameters. The overall function can now be expressed as

$$H(z) = C + \sum_{k=1}^K H_k(z)$$

Where K is the integer part of $(N+1)/2$. When N is odd, one of the $H_k(z)$ is really a single-pole system (i.e. $b_{k1}=a_{k2}=0$).

Obtain the direct, the cascade and the parallel form realizations of the following IIR filter

$$H(z) = \frac{3(2z^2 + 5z + 4)}{(2z + 1)(z + 2)}$$

4. Lattice and Lattice-Ladder Structures for IIR Systems:

- All pole IIR System \rightarrow Lattice Structure
- Pole-zero IIR System \rightarrow Lattice-Ladder Structure

Lattice Figure of IIR System:

Let us begin with an all pole system with system function

$$H(z) = \frac{1}{1 + \sum_{k=1}^{N-1} a_k z^{-k}} = \frac{1}{A_N(z)}$$

The difference equation for this IIR system is

$$y[n] = x[n] - \sum_{k=1}^{N-1} a_N(k)y[n-k]$$

If we interchange the roles of input and output, we obtain

$$x[n] = y[n] - \sum_{k=1}^{N-1} a_N(k)x[n-k]$$

Digital Signal Analysis & Processing

Chapter 5: Discrete Filter Structure

$$Or, y[n] = x[n] + \sum_{k=1}^{N-1} a_N(k)x[n-k]$$

Which describes an FIR system having the system function $H(z) = A_N(z)$.

Now, we take the all-zero lattice filter and redefine the input as $x[n] = f_N(n)$ and output as $y[n] = f_0(n)$.

These are exactly the opposite of the definitions for the all-zero lattice filters. These definitions dictate that the quantities $\{f_m(n)\}$ can be computed in descending order. This computation can be accomplished by rearranging the recursive equation and thus solving for $f_{m-1}(n)$ in terms of $f_m(n)$.

$$i.e. f_{m-1}(n) = f_m(n) - k_m g_{m-1}(n-1), \quad m = N, N-1, \dots, 2, 1$$

The equation for $g_m(n)$ remains unchanged

$$i.e. g_m(n) = k_m f_{m-1}(n) + g_{m-1}(n-1),$$

The result of these changes in the set of equations

$$\begin{aligned} f_N(n) &= x(n) \\ f_{m-1}(n) &= f_m(n) - k_m g_{m-1}(n-1), \quad m = N, N-1, \dots, 2, 1 \\ g_m(n) &= k_m f_{m-1}(n) + g_{m-1}(n-1) \\ y[n] &= f_0(n) = g_0(n) \end{aligned}$$

Which correspond to the structure.

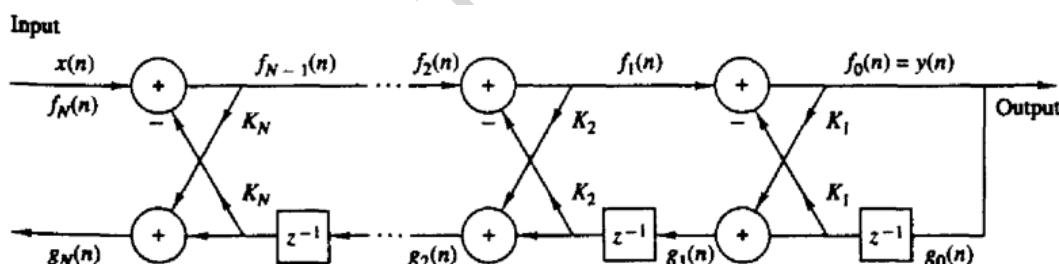


Figure: Lattice structure for an all pole IIR system.

In practical applications the all-pole lattice structure has been used to model the human voice tract and a stratified (def: deposited or arranged in horizontal layers) earth. In such cases the lattice parameters, $\{k_m\}$ have the physical significance of being identical to reflection coefficients in the physical medium. This is the reason that the lattice parameters are often called reflection coefficients. In such applications, a stable model of the medium requires that the reflection coefficients, obtained by performing measurements on output signals from the medium, be less than unity.

C. Lattice Ladder Structures:

Let us consider an IIR system with system function

$$H(z) = \frac{\sum_{k=0}^M c_M(k)z^{-k}}{1 + \sum_{k=1}^N a_N(k)z^{-k}} = \frac{C_M(z)}{A_N(z)} \quad \dots \quad (I)$$

Where, the notation for the numerator polynomial has been changed to avoid confusion with our previous development. We assume that $N \geq M$.

The direct form II structure is,

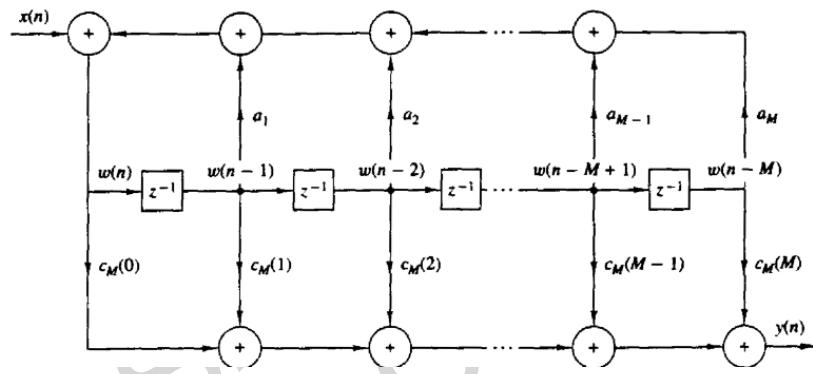


Figure : Direct form-II realization of IIR system.

$$y[n] = - \sum_{k=1}^N a_N(k)y[n-k] + \sum_{k=0}^M c_M(k)x[n-k] \quad \dots \quad (II)$$

$$w[n] = - \sum_{k=1}^N a_N(k)y[n-k] + x[n] \quad \dots \quad (III)$$

$$y[n] = \sum_{k=0}^M c_M[k]w[n-k] \quad \dots \quad (IV)$$

Note that eqn (III) is the input-output of an all-pole IIR system and that eqn (IV) is the input-output of an all-zero system. Furthermore, we observe that the output of the all-zero system is simply a linear combination of delayed outputs from the all-pole system.

Digital Signal Analysis & Processing

Chapter 5: Discrete Filter Structure

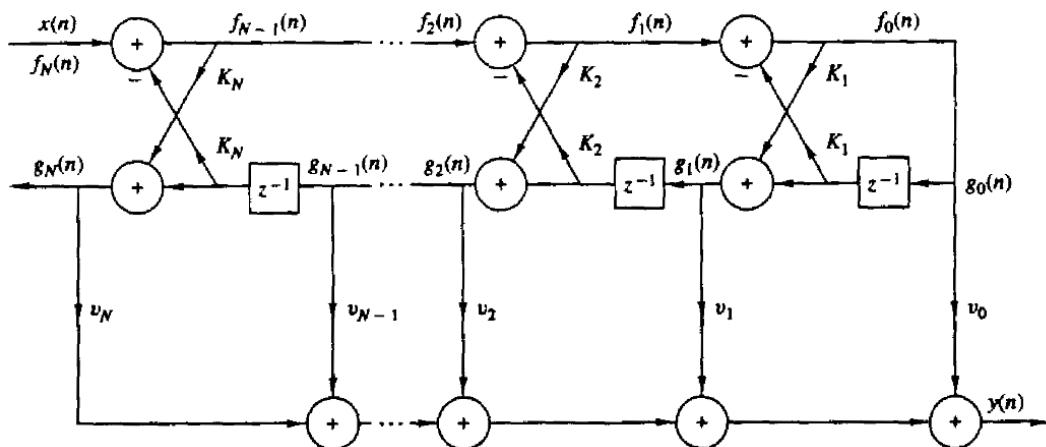


Figure: Lattice-Ladder structure for the realization of a pole-zero system.

$$y[n] = \sum_{m=0}^M v_m g_m(n) \quad \text{--- (V)}$$

Where $\{v_m\}$ are the parameters that determine the zeros of the system. The system function corresponding to eqn (V) is

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{m=0}^M v_m G_m(z)}{X(z)} \quad \text{--- (VI)}$$

Since $X(z) = F_N(z)$ & $F_0(z) = G_0(z)$

$$\begin{aligned} H(z) &= \sum_{m=0}^M v_m \frac{G_m(z)}{G_0(z)} \frac{F_0(z)}{F_N(z)} \\ &= \sum_{m=0}^M v_m \frac{B_m(z)}{A_N(z)} \\ &= \frac{\sum_{m=0}^M v_m B_m(z)}{A_N(z)} \quad \text{--- (VII)} \end{aligned}$$

If we compare eqn (I) with (VII), we conclude that

$$C_M(z) = \sum_{m=0}^M v_m B_m(z) \quad \text{--- (VIII)}$$

This is the desired relationship that can be used to determine the weighting coefficients $\{v_m\}$. Thus, we have demonstrated that the coefficients of the numerator polynomial $C_M(z)$ determine the ladder parameters $\{v_m\}$, whereas the coefficients in the denominator polynomial $A_N(z)$ determine the lattice parameters $\{k_m\}$.

The ladder parameters are determined from (III) which can be expressed as,

$$C_m(z) = \sum_{k=0}^{m-1} v_k B_k(z) + v_m B_m(z)$$

or $C_m(z) = C_{m-1}(z) + v_m B_m(z)$

Thus $C_m(z)$ can be computed recursively from the reverse polynomials $B_m(z)$, $m = 1, 2, 3, \dots, M$. Since $\beta_m(m) = 1$ for all m , the parameters v_m , $m = 0, 1, \dots, M$ can be determined by first noting that,

$$v_m = C_m(m), \quad m = 0, 1, 2, \dots, M$$

$$\therefore C_{m-1}(z) = C_m(z) - v_m B_m(z)$$

Quantization of filter coefficients and effects on location of Pole and Zeros:

In the realization of FIR and IIR filters in hardware or in software on a general purpose computer, the accuracy with which filter coefficients can be specified is limited by the word length of the computer or the register provided to store the coefficients. Since, the coefficients used in implementing a given filter are not exact, the poles and zeros of system function will, in general, be different from the desired poles and zeros. Consequently, we obtain a filter having a frequency response of the filter with unquantized coefficients.

The sensitivity of the filter frequency response characteristics to quantization of the filter coefficients is minimized by realizing a filter having a large number of poles and zeros as an interconnection of second order filter sections. This leads to the parallel form and cascade form realizations in which the basic building blocks are second order filter sections.

Pole Perturbation:

Consider a general IIR filter with system function

$$H(z) = \frac{\sum_{k=0}^M b_k z^{-k}}{1 + \sum_{k=1}^N a_k z^{-k}}$$

The direct form realization of IIR filter with quantized coefficients has the system function

$$\bar{H}(z) = \frac{\sum_{k=0}^M \bar{b}_k z^{-k}}{1 + \sum_{k=1}^N \bar{a}_k z^{-k}}$$

Where, the quantized coefficients $\{\bar{b}_k\}$ and $\{\bar{a}_k\}$ can be related to the unquantized coefficients $\{b_k\}$ and $\{a_k\}$ by the relation

Digital Signal Analysis & Processing

Chapter 5: Discrete Filter Structure

$$\begin{aligned}\bar{a}_k &= a_k + \Delta a_k \quad k = 1, 2, 3, \dots, N \\ \bar{b}_k &= b_k + \Delta b_k \quad k = 1, 2, 3, \dots, M\end{aligned}$$

$\{\Delta a_k\}$ and $\{\Delta b_k\}$ represent quantization errors.

The denominator of $H(z)$ may be expressed in the form

$$D(z) = 1 + \sum_{k=1}^N a_k z^{-k} = \prod_{k=1}^N (1 - p_k z^{-1})$$

Where $\{p_k\}$ are the poles of $H(z)$. Similarly, we can express denominator of $\bar{H}(z)$ as,

$$\bar{D}(z) = \prod_{k=1}^N (1 - \bar{p}_k z^{-1})$$

Where $\bar{p}_k = p_k + \Delta p_k$, $k = 0, 1, 2, \dots, N$ and Δp_k is the error or **perturbation** resulting from the quantization of filter coefficient.

The perturbation error Δp_i can be expressed as,

$$\Delta p_i = \sum_{k=1}^N \frac{\partial p_i}{\partial a_k} \Delta a_k$$

Where $\frac{\partial p_i}{\partial a_k}$, the partial derivative of p_i with respect to a_k represents the incremental change in the pole p_i due to change in the coefficient a_k . Thus, the total error Δp_i is expressed as a sum of the incremental errors due to changes in each of the coefficients $\{a_k\}$.

The partial derivatives $\frac{\partial p_i}{\partial a_k}$, $k = 1, 2, \dots, N$ can be obtained by differentiating $D(z)$ with respect to each of $\{a_k\}$. We have,

$$\left(\frac{\partial D(z)}{\partial a_k} \right)_{z=p_i} = \left(\frac{\partial D(z)}{\partial z} \right)_{z=p_i} \left(\frac{\partial p_i}{\partial a_k} \right)$$

Then

$$\frac{\partial p_i}{\partial a_k} = \frac{\left(\frac{\partial D(z)}{\partial a_k} \right)_{z=p_i}}{\left(\frac{\partial D(z)}{\partial z} \right)_{z=p_i}}$$

$$\left(\frac{\partial D(z)}{\partial a_k} \right)_{z=p_i} = -z^{-k} \Big|_{z=p_i} = -p_i^{-k}$$

Digital Signal Analysis & Processing

Chapter 5: Discrete Filter Structure

$$\left(\frac{\partial D(z)}{\partial z} \right)_{z=p_i} = \left\{ \frac{\partial}{\partial z} \left[\prod_{l=1}^N (1 - p_l z^{-1}) \right] \right\}_{z=p_i}$$

$$= \left\{ \sum_{k=1}^N \frac{p_k}{z^2} \prod_{l=1, l \neq k}^N (1 - p_l z^{-1}) \right\}_{z=p_i} = \frac{1}{p_i^N} \prod_{l=1, l \neq k}^N (p_i - p_l)$$

$$\frac{\partial p_i}{\partial a_k} = - \frac{p_i^{N-k}}{\prod_{l=1, l \neq k}^N (p_i - p_l)}$$

$$\text{Thus, } \Delta p_i = - \sum_{k=1}^N \frac{p_i^{N-k}}{\prod_{l=1, l \neq k}^N (p_i - p_l)} \Delta a_k$$

This expression provides a measure of sensitivity of the i^{th} pole to changes in the coefficients $\{a_k\}$.

References:

1. J. G. Proakis, D. G. Manolakis, "Digital Signal Processing, Principles, Algorithms and Applications", 3rd Edition, Prentice-hall, 2000. Chapter 9.