Unit : IV FIR filter design.

Realization of FIR filters:-

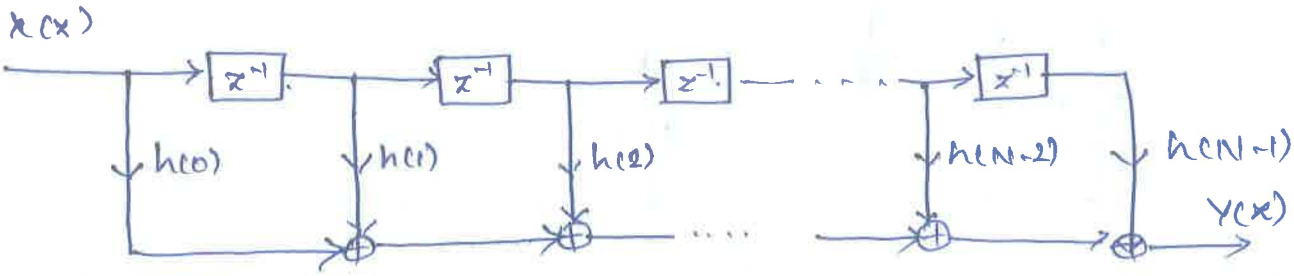
Transversal structure:- (Direct form)

The system function of an FIR filter can be written as

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

$$= h(0) + h(1) z^{-1} + h(2) z^{-2} + \dots + h(N-1) z^{-(N-1)}$$

$$Y(z) = h(0) X(z) + h(1) z^{-1} X(z) + h(2) z^{-2} X(z) + \dots + h(N-1) z^{-(N-1)} X(z)$$



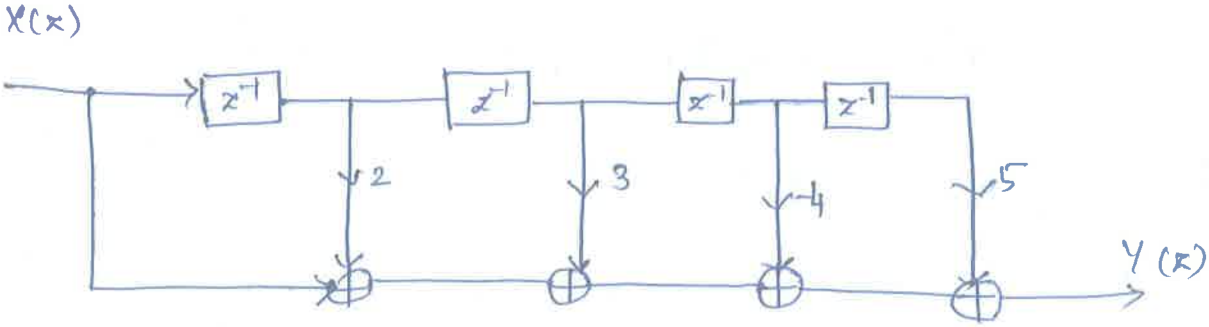
Problem:-

1) Determine the direct form realization of system function

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

Soln.

$$Y(z) = X(z) + 2z^{-1} X(z) + 3z^{-2} X(z) - 4z^{-3} X(z) + 5z^{-4} X(z)$$



Cascade form realization.

2) Obtain the cascade realization of system function

$$H(z) = (1 + 2z^{-1} - z^{-2})(1 + z^{-1} - z^{-2})$$

Soln:

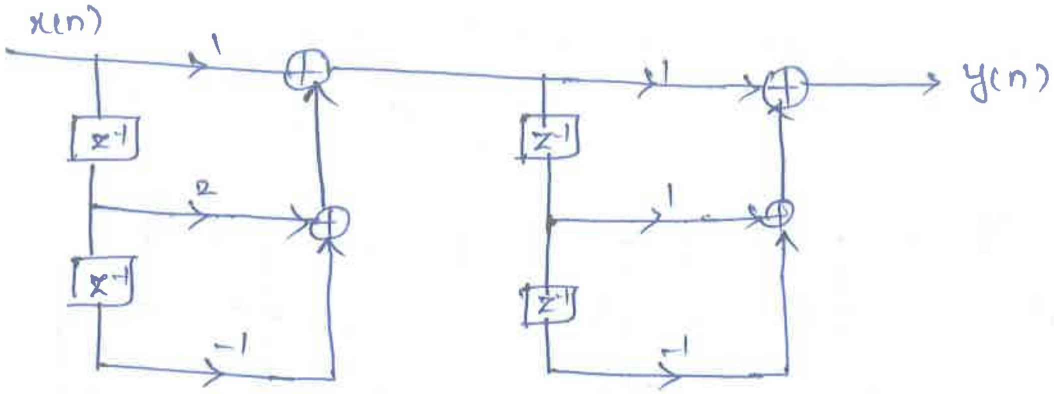
$$H(z) = H_1(z)\,H_2(z)$$

$$H_1(z) = 1 + 2z^{-1} - z^{-2}.$$

$$H_1(z) \Rightarrow Y_1(z) = X_1(z) + 2z^{-1}X_1(z) - z^{-2}X_1(z)$$

$$H_2(z) = 1 + z^{-1} - z^{-2}.$$

$$Y_2(z) = X_2(z) + z^{-1}X_2(z) - z^{-2}X_2(z)$$



Exercises:

1. Obtain the cascade realization of system function.

$$H(z) = 1 + \tfrac{5}{2}z^{-1} + 2z^{-2} + 2z^{-3}$$

2. Obtain direct form, cascade form realization for the system function.

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

Linear phase realization:-

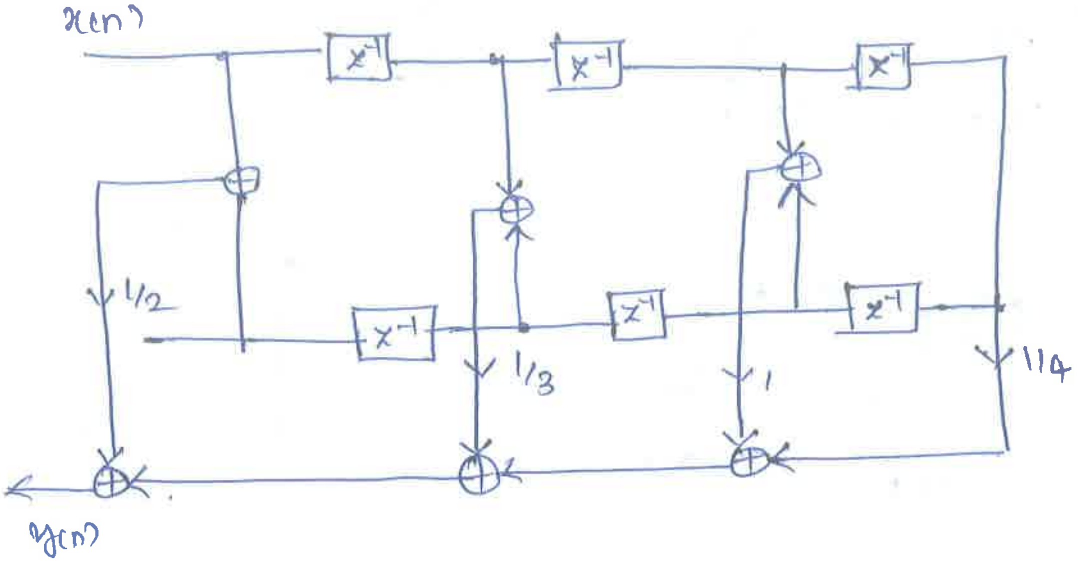
For a linear phase FIR filter $h(n) = h(N-1-n)$

Realize the system function

$$H(z) = \frac{1}{2} + \frac{1}{3}z^{-1} + z^{-2} + \frac{1}{4}z^{-3} + z^{-4} + \frac{1}{3}z^{-5} + \frac{1}{2}z^{-6}$$

$N = 7$

$$h(0) = h(6) = 1/2$$
$$h(1) = h(5) = 1/3$$
$$h(2) = h(4) = 1$$
$$h(3) = 1/4$$



Polyphase structures:-

1. Realize the following system with polyphase structures:-

$$\cancel{H(z) = 1 - 3z^{-2} - 9z^{-4} + 7z^{-6}}$$

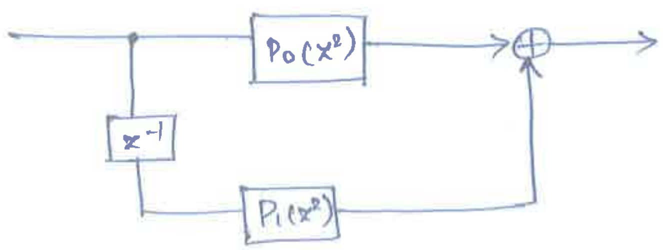
$$H(z) = 1 + 4z^{-1} - 3z^{-2} + 6z^{-3} - 9z^{-4} + 5z^{-5} + 7z^{-6}$$

$$H(z) = P_0(z^2) + z^{-1} P_1(z^2) \quad \text{——— (1)}$$

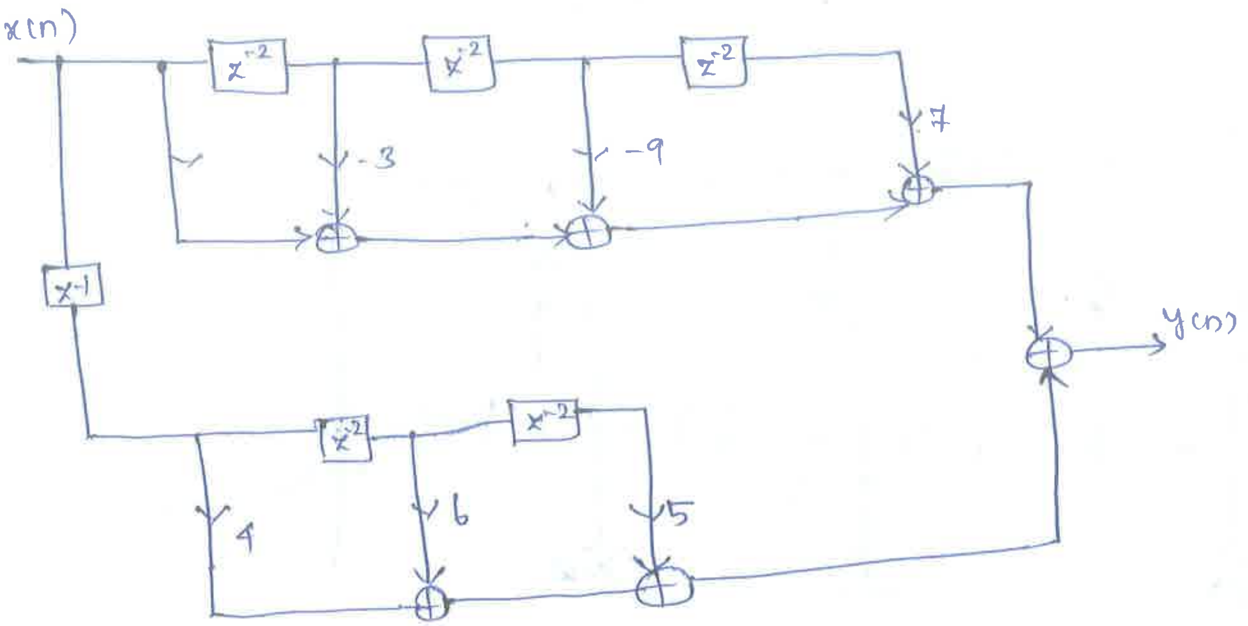
$$P_0(z^2) = 1 - 3z^{-2} - 9z^{-4} + 7z^{-6}$$

$$P_1(z^2) = 4 + 6z^{-2} + 5z^{-4}$$

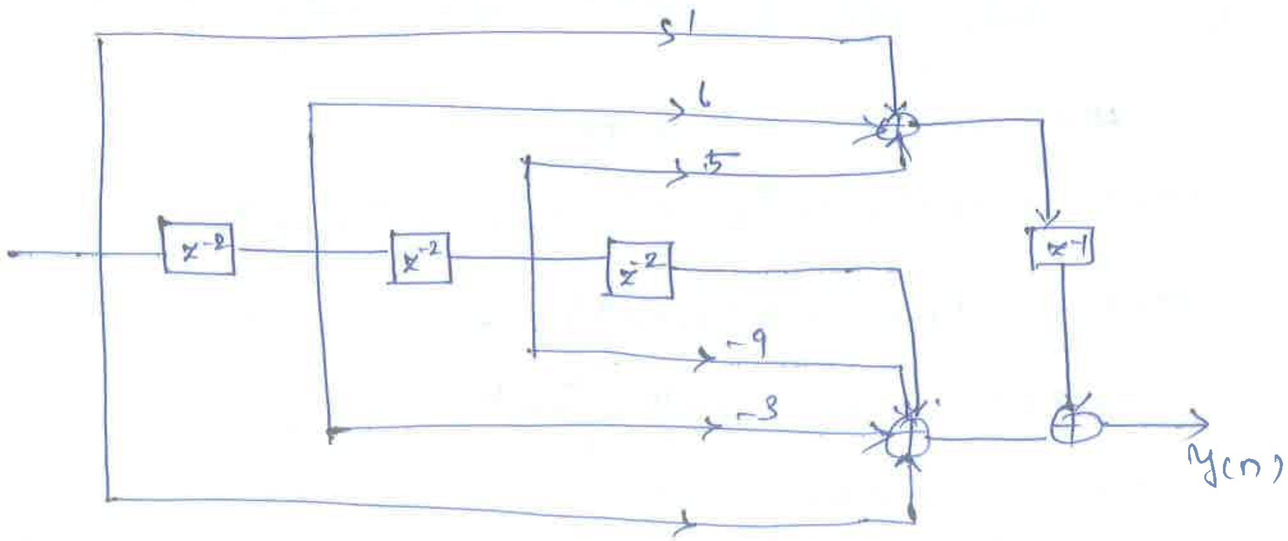
Polyphase structure can be drawn as



$P_0(z^2)$ and $P_1(z^2)$ can be realized in direct form.



Canonical polyphase realization.



4. Design an ideal LPF with a frequency response

$$H_d(e^{j\omega}) = 1 \quad \text{for} \quad -\pi/2 \le \omega \le \pi/2.$$

$$= 0 \quad \text{for} \quad \pi/2 \le |\omega| \le \pi$$

Find the values of $h(n)$

Soln:

$$h_d(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(e^{j\omega})\, e^{j\omega n}\, d\omega$$

$$= \frac{1}{2\pi} \int_{-\pi/2}^{\pi/2} e^{j\omega n}\, d\omega$$

$$= \frac{1}{2\pi} \left[\frac{e^{j\omega n}}{jn} \right]_{-\pi/2}^{\pi/2}$$

$$= \frac{1}{2\pi j n} \left[e^{j\pi/2 n} - e^{-j\pi/2 n} \right]$$

$$h_d(n) = \frac{1}{\pi n} \sin \tfrac{\pi}{2} n \qquad -\infty < n < \infty$$

Truncating $h_d(n)$ to 11 samples we have

$$h_d(n) = \frac{\sin \frac{\pi}{2} n}{\pi n} \quad \text{for } |n| \le 5 \qquad \text{--- (1)}$$

$n = 0 \rightarrow$ in equation (1)

$$h(0) = \lim_{n \to 0} \frac{\sin \frac{\pi}{2} n}{\pi n} = \frac{1}{\pi} \left[\frac{\pi}{2} \right]$$

$$h(0) = \frac{1}{2}$$

$$h(1) = h(-1) = \frac{\sin \pi/2}{\pi} = \frac{1}{\pi} = 0.3183$$

$$h(2) = h(-2) = 0$$

$$h(3) = h(-3) = -0.106$$

$$h(4) = h(-4) = 0 \; ; \; h(5) = h(-5) = 0.06366.$$

The transfer function of the filter is given by

$$H(z) = h(0) + \sum_{n=1}^{\frac{N-1}{2}} [h(n)\,(z^{n}+z^{-n})]$$

$$= 0.5 + 0.3183\,(z+z^{-1}) - 0.106(\,z^{3}+z^{-3}) + 0.06366(z^{5}+z^{-5})$$

The transfer function of the realizable filter is

$$H'(z) = z^{-(N-1/2)}\;H(z)$$

$$= z^{-5}\left[0.5 + 0.3183\,(z+z^{-1}) - 0.606(z^{3}+z^{-3}) + 0.06366(z^{5}+z^{-5})\right]$$

$$H'(z) = 0.06366 - 0.106z^{-2} + 0.3183z^{-4} + 0.5\,z^{-5} + 0.3183z^{-6} - 0.106z^{-8} + 0.06366z^{-10}.$$

(6).

From the above equation.

$h(0) = h(10) = 0.06366$; $h(1) = h(9) = 0$; $h(2) = h(8) = -0.106$

$h(3) = h(7) = 0$ $\quad h(4) = h(6) = 0.3183$ $\quad h(5) = 0.5$

The frequency response is given by

$$\overline{H(e^{jw})} = \sum_{n=0}^{5} a(n)\cos wn.$$

$$a(0) = h\left(\frac{N-1}{2}\right) = h(5) = 0.5$$

$$a(n) = 2h\left(\frac{N-1}{2} - n\right)$$

$$a(1) = 2h(5-1) = 2h(4) = 0.6366$$

$a(2) = 0$

$a(3) = -0.212$

$a(4) = 0$

$a(5) = 0.127$

$$\bar{H}(e^{jw}) = 0.5 + 0.6366\cos w - 0.212\cos 3w + 0.127\cos 5w$$

The desired frequency response is sampled at sufficient no of points.

$H_d(e^{jw})$ = Ideal desired frequency response

$H(k)$ = DFT of sequence obtained by sampling $H_d(e^{jw})$

$h(n)$ = Impulse response of FIR filter.

Procedure for Type I design:

1. Choose the ideal frequency response $H_d(e^{jw})$
2. Sample $H_d(e^{jw})$ at N-points by taking $w = w_k = 2\pi k/N$ where $k = 0, 1, 2 \dots N-1$ to generate the sequence $H(k)$

$$H(k) = H_d(e^{jw})\Big|_{w = 2\pi k/N}$$

3. Compute the N-samples of impulse response $h(n)$

When N is odd

$$h(n) = \frac{1}{N}\left[H(0) + 2\sum_{k=1}^{\frac{N-1}{2}} Re\left[H(k)\, e^{j2\pi nk/N}\right]\right]$$

N is even

$$h(n) = \frac{1}{N}\left[H(0) + 2\sum_{k=1}^{N/2-1} Re\left[H(k)\, e^{j2\pi nk/N}\right]\right]$$

here $H(N/2) = 0$

4. Take Z Transform of the impulse response $h(n)$ to get the filter transfer function $H(z)$

1) Determine the co-efficients of a linear phase FIR filter of length $N = 15$ which has a symmetric unit sample response that satisfies the conditions

$$H\left(\frac{2\pi k}{15}\right) = \begin{cases} 1 & k = 0, 1, 2, 3 \\ 0.4 & k = 4 \\ 0 & k = 5, 6, 7 \end{cases}$$

$$\alpha = \frac{N-1}{2}$$

$N = 15$ $\quad \alpha = 7$. $\qquad H(k) = H_d(\omega)\Big|_{\omega=\omega_k}$.

$$\omega = \omega_k = 2\pi k/15 =$$

$$H_d(\omega)\Big|_{\omega_k} = 1\, e^{-j\alpha\omega_k} = e^{-j7\times 2\pi k/15} \qquad k=0,1,2,3$$

$$= 0.4\, e^{-j\alpha\omega_k} = 0.4\, e^{-j7\times 2\pi k/15} \qquad k=4$$

$$= 0 \qquad k=5,6,7.$$

The samples of impulse response $h(n)$ are given by

$$h(n) = \frac{1}{N}\left[H(0) + 2\sum_{k=1}^{N-1/2} \mathrm{Re}\left[H(k)\, e^{j2\pi nk/N}\right]\right]$$

$$= \frac{1}{15}\left[H(0) + 2\sum_{k=1}^{7} \mathrm{Re}\left[H(k)\, e^{j2\pi nk/15}\right]\right]$$

$$= \frac{1}{15}\left[H(0) + 2\sum_{k=1}^{3} \mathrm{Re}\left[H(k)\, e^{j2\pi nk/15}\right] + 2\,\mathrm{Re}\left[H(4)\, e^{j2\pi n 4/15}\right]\right]$$

$$= \frac{1}{15}\left[1 + 2\sum_{k=1}^{3} \mathrm{Re}\left[e^{-j7\cdot 2\pi k/15}\, e^{j2\pi nk/15}\right] * 2\,\mathrm{Re}\left[0.4\, e^{-j7\, 2\pi\times 4/15}\, e^{j8\pi n/15}\right]\right]$$

$$= \frac{1}{15}\left[1 + 2\sum_{k=1}^{3} \mathrm{Re}\left[e^{j2\pi k/15\,(n-7)}\right] + 2\,\mathrm{Re}\left[0.4\, e^{j8\pi/15\,(n-7)}\right]\right]$$

$$= \frac{1}{15}\left[1 + 2\sum_{k=1}^{3} \cos 2\pi k/15\,(n-7) + 0.8\cos 8\pi/15\,(n-7)\right]$$

$$= \frac{1}{15}\left[1 + 2\cos\frac{2\pi(n-7)}{15} + 2\cos\frac{4\pi(n-7)}{15} + \frac{2\cos 6\pi(n-7)}{15} + 0.8\cos\frac{8\pi(n-7)}{15}\right]$$

$n=0$ $h(0) = -0.0141$

$h(1) = -0.0019$

$h(2) = 0.04$

$h(3) = 0.0122$

$h(4) = -0.0914$

$h(5) = -0.0181$

$h(6) = 0.3130$

$h(7) = 0.5\cancel{2}$

$h(8) = 0.3130$

$h(9) = -0.981$

$h(10) = -0.0914$

$h(11) = 0.0122$

$h(12) = 0.04$

$h(13) = -0.0019$

$h(14) = -0.0141$

Magnitude response :-

$$A(\omega) = h\left(\frac{N-1}{2}\right) + \sum_{n=1}^{\frac{N-1}{2}} 2h\left(\frac{N-1}{2} - n\right)\cos\omega n.$$

Frequency Sampling method of designing FIR filters:-

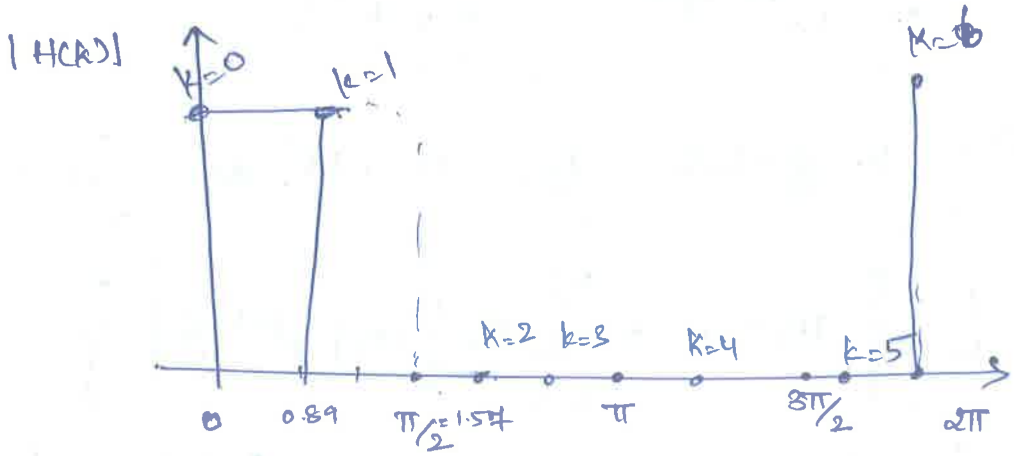
1. Determine the filter co-efficients h(n) obtained by sampling $H_d(e^{j\omega}) = e^{-j(N-1)\omega/2}$ $\quad 0 \le \omega \le \pi/2$

$\quad = 0 \qquad \pi/2 \le \omega \le \pi$

for $N = 7$

Soln: The initial point is located at $\omega = 0$ and the spacing between 2 points is $\frac{2\pi}{N} = \frac{2\pi}{7}$

∴ The ideal magnitude response.



$\omega = 2\pi k/7$ $\quad \Rightarrow k=0 \quad \omega = 0$

$\Rightarrow k=1 \quad \omega = 2\pi/7 = 0.8977.$

$k=2 \quad \omega = \frac{4\pi}{7} = 1.7952.$

$k=3 \quad \omega = 6\pi/7 = 2.6928.$

$k=4 \quad \omega = 8\pi/7 = 3.5904$

$k=5 \quad \omega = 10\pi/7 = 4.4879$

$k=6 = \omega = 12\pi/7 = 5.3856$

$|H(k)| = 1 \quad$ for $k = 0, 1, 6$

$\quad\quad 0 \quad$ for $k = 2, 3, 4, 5$

$$\therefore \quad \theta(k) = -\left(\frac{N-1}{N}\right)\pi k = \frac{-6}{7}\pi k \quad \text{for} \quad k = 0, 1, 2, 3$$

$k = \frac{N+1}{2}$.

$$\theta(k) = (N-1)\pi - \left(\frac{N-1}{N}\right)\pi k \qquad k = \frac{N+1}{2} \ldots N-1$$

$$= 6\pi - \frac{6\pi}{7}\pi k \qquad k = 4, 5, 6$$

$e^{-j\theta(k)}$

$$H(k) = e^{-j6\pi k/7} \qquad k = 0, 1$$

$$= 0 \qquad \text{for } k = 2, 3, 4, 5$$

$$= e^{-j6\pi/7(k-7)} \qquad \text{for } k = 6.$$

The filter co efficients for N- odd are given by

$$h(n) = \frac{1}{N}\left\{ H(0) + 2\sum_{k=1}^{\frac{N-1}{2}} Re\left[H(k)\, e^{j2\pi kn/7}\right]\right\}$$

$n = 0, 1, \ldots N-1$

$$= \frac{1}{7}\left\{1 + 2\, Re\left[e^{-j6\pi/7} \cdot e^{j2\pi kn/7}\right]\right\}$$

$$h(n) = \frac{1}{7}\left\{1 + 2\cos\frac{2\pi}{7}(n-3)\right\}$$

$$h(0) = h(6) = \frac{1}{7}\left(1 + 2\cos 6\pi/7\right) = -0.11456$$

$$h(1) = h(5) = \frac{1}{7}\left(1 + 2\cos\frac{4\pi}{7}\right) = 0.07928$$

$$h(2) = h(4) = \frac{1}{7}\left(1 + 2\cos\frac{2\pi}{7}\right) = 0.321$$

$$h(3) = \frac{1}{7}(1+2) = 0.42857$$

Finite word length effects

The effects due to finite precision representation of numbers in digital system are commonly referred to as finite word length effects.

They are

i) Errors due to quantization of input data by A/D converter

ii) Errors due to quantization of filter co-efficients

iii) Errors due to rounding the product in multiplication

iv) Errors due to overflow in addition.

v) Limit cycles.

Radix Number system:-

Any number can be represented as

$$N = \sum_{i=-A}^{B} d_i r^{-i}$$

A - no of integer digits

B - No of fraction digits.

r - Radix or Base

d_i - i^{th} digit of the number

For benary number.

$$N = \sum_{i=-A}^{A} d_i \, 2^{-i}$$

Example :-

$(178.25)_{10}$ => $r = 10$

$A = -2$

$B = 2$

$$\therefore N = \sum_{i=-2}^{2} d_i \, 10^{-i}$$

$$= (d_{-2}\,10^{2}) + (d_{-1}\,10^{1}) + (d_0 \times 10^{0}) + (d_1\,10^{-1}) + (d_2 \times 10^{-2})$$

$$= 1 \times 10^{2} + 7 \times 10^{1} + 8 \times 10^{0} + 2 \times 10^{-1} + 5 \times 10^{-2}$$

$(111.11)_2$ => $$N = \sum_{i=-2}^{2} d_i \, 2^{-i}$$

$$= d_{-2}\,2^{-2} + d_{-1}\,2^{-1} + d_0\,2^{0} + d_1\,2^{-1} + d_2\,2^{-2}$$

$$= 1 \times 2^{-2} + 1 \times 2^{-1} + 1 \times 2^{0} + 1 \times 2^{-1} + 1 \times 2^{-2}.$$

Two methods of representing binary numbers are.

* fixed point representation
* floating point representation

Fixed point representation:

- The digits alloted for integer part and fraction part are fixed, and so the position of the benary point is fixed.

Floating point representation:-

The binary point can be shifted to desired position so that number of digits in the integer part and fraction part of a number can be varied.

Fixed point representation:-

There are three different formats for representing negative binary fraction numbers. They are

* Sign- magnitude
* One's Complement
* Two's Complement

There is only one way of representing positive numbers.

$$\sum_{i=1}^{B} d_i \, 2^{-i}$$

d_0 - is set to zero to represent the the positive sign.

Sign magnitude format:-

- In this format the negative value of a given number differ only in sign bit.

$$\text{Negative binary fraction number} = 1(1 \times 2^0) + \sum_{i=1}^{B} d_i 2^{-i}$$

The range of decimal fraction numbers that can be represented is

$$-\left[1 - 2^{-(B-1)}\right] \text{ to } \left[1 - 2^{-(B-1)}\right]$$

with step size $\frac{1}{2^B - 1}$ B no of bits.

Example:-

1. Convert $+0.125_{10}$ and -0.125_{10} to sign magnitude format of binary and verify the result by converting the binary to decimal:-

$$.125 \longrightarrow +0.001 \xrightarrow[\text{sign}]{\text{Append}} 0.001 \xrightarrow{\text{Remove dot}} 0001_2$$

$$-0.125 \longrightarrow -.001 \longrightarrow 1.001 \longrightarrow 1001_2.$$

Binary to decimal

$$0001_2 \longrightarrow .001_2 \longrightarrow +.125$$

$$1001_2 \longrightarrow -.001_2 \longrightarrow -0.125.$$

One's complement format:-

Negative binary fraction number in one's complement

$$= 1 \times 2^0 + \sum_{i=1}^{B} (1 - d_i)\, 2^{i}.$$

Example:-

2. Convert $+0.125_{10}$ and -0.125_{10} to 1's complement ~~sign magnitude~~ format of binary..

Soln:

$$+0.125_{10} \longrightarrow +.001 \longrightarrow 0.001 \longrightarrow 0001_2$$

$$-0.125_{10} \longrightarrow -.001 \xrightarrow{\text{Complement}} -.110. \longrightarrow 1.110 \longrightarrow 1110_2$$

Two's complement format:-

Range -1 to $+\left[1 - 2^{-(B-1)}\right]$ step size $= \frac{1}{2^B - 1}$

Example :-

3. convert +0.125 and −0.125 to two's complement format of binary.

$\cdot 125 \rightarrow +\cdot 001 \rightarrow 0.001 \rightarrow 0001_2$

$-\cdot 125 \rightarrow -\cdot 001 \rightarrow -\cdot 110 \rightarrow -\cdot 111 \rightarrow 1.111 \rightarrow 1111_2$

Floating point Representation :-

The floating point representation is employed to represent large range of numbers in a given binary word size.

The floating point number is represented as

$$N_F = M \times 2^E$$

M – Mantissa and it will be in binary fraction format. Range $0.5 \leq M < 1$

E – Exponent it is either positive or negative integer.

the left most bit of mantissa is used to represent sign

the left most bit " exponent " " " sign.

Example :-

$5_{10} \rightarrow 101_2 \xrightarrow[\text{exponent}]{\text{append}} 101.0 \times 2^0 \xrightarrow{\text{Normalize}} \cdot 1010 \times 2^{+3_{10}}$ — Convert to binary — $\cdot 1010 \times 2^{11_2}$

$\cdot 1010 \times 2^{11_2} \xrightarrow[\text{sign}]{\text{append}} 0.1010 \times 2^{011} \xrightarrow[\text{dot}]{\text{Remove}} 01010 \times 2^{011}$

$5 \rightarrow 01010011$

$\cancel{-5 \rightarrow .101 \rightarrow -.101 \times 2^{0}} \rightarrow \frac{-101.0}{\cancel{-1.0110}}$

$-5 \rightarrow 101 \rightarrow 101.0 \times 2^{0} \rightarrow -.1010 \times 2^{+3} \rightarrow -.1010 \times 2^{011}$

$11010 \times 2^{011} \leftarrow 1.1010 \times 2^{011}$

$-5 \rightarrow (11010011)_2$

$0.125 \rightarrow .001_2 \rightarrow +.001 \times 2^{0} \rightarrow .1000 \times 2^{-2_{10}} \rightarrow .1000 \times 2^{-10}$

$01000 \times 2^{110} \leftarrow 0.1000 \times 2^{110}$

$+0.125 \rightarrow (01000\,110)_2$

$-0.125 \rightarrow -.001 \rightarrow -.001 \times 2^{0} \rightarrow .1000 \times 2^{-2} \rightarrow .1000 \times 2^{-10}$

$11000 \times 2^{110} \leftarrow 1.1000 \times 2^{110}$

$-0.125 \rightarrow (11000\,110)_2$

Quantization by truncation and rounding :-

- In fixed point or floating point arithmetic the size of the result of an operation may be exceeding the size of binary used in the number system.
- In such cases the low order bits has to be eliminated in order to store the result.
- The two methods of eliminating these low order bits are truncation and rounding.

- The effect of rounding and truncation is to introduce an error.

Quantization steps :-

When 'B' bit binary is selected to represent the decimal numbers, then 2^B binary codes are possible.

Each step of binary number is also called quantization step.

$$q = \frac{R}{2^B} = \frac{1-(-1)}{2^B} = \frac{2}{2^B} = 2\frac{1}{2^{B-1}} = \frac{1}{2^b}$$

B → size of binary including sign bit

b → " " " excluding " ".

Truncation :-

- It is the process of reducing the size of binary number by discarding all bits less significant than the least significant bit that is retained.

fixed point numbers :-

i) +ve number range of error :- $-2^{-b} < e \leq 0$

ii) 1's complement negative number :- $0 \leq e < 2^{-b}$

iii) 2's " " " :- $-2^{-b} < e \leq 0$

$$N_{tf} = N_f + N_f \varepsilon_t$$

N_f → unquantized N_{tf} → truncated floating point

ε_t → relative error.

$$e_t = \frac{N_{tf} - N_f}{N_f}$$

2's complement	+ve mantissa	$-2 \times 2^{-b} < \varepsilon_t \le 0$
2's complement	-ve "	$0 \le \varepsilon_t < 2^{-b} \times 2$
1's complement	+ve and -ve mantissa	$-2 \times 2^{-b} < \varepsilon_t \le 0$
Sign magnitude	+ve & -ve mantissa	$-2 \times 2^{-b} < \varepsilon_t \le 0$

Rounding:

- It is the process of reducing the size of a binary number to finite word size of b-bits such that the rounded b-bit number is closest to the original unquantized number.

Range of error fixed point

$$\frac{-2^{-b}}{2} \le e_r \le \frac{2^{-b}}{2}.$$

Range of error floating point $-2^{-b} \le \varepsilon_r \le 2^{-b}$.

Quantization error:

The quantization error for rounding will be in the range $-q/2$ to $q/2$.

The error due to rounding is treated as a random variable

$$E(e) = \frac{1}{q/2 - (-q/2)} \int_{-q/2}^{q/2} e\, de$$

$$= \frac{1}{q}\left[\frac{e^2}{2}\right]_{-q/2}^{q/2}$$

$$= \frac{1}{2q}\left[\left(\frac{q}{2}\right)^2 - \left(-\frac{q}{2}\right)^2\right]$$

$$= 0$$

$$\sigma_e^2 = E\{e^2\} - E^2\{e\} = E\{e^2\}$$

$$= \frac{1}{\frac{q}{2}-\left(-\frac{q}{2}\right)}\int_{-q/2}^{q/2} e^2\,de = \frac{1}{q}\left[\frac{e^3}{3}\right]_{-q/2}^{q/2} = \frac{1}{3q}\left[\frac{q^3}{8}+\frac{q^3}{8}\right]$$

$$= \frac{1}{3q} \times \frac{\cancel{2}q^3}{8_4} = \frac{q^2}{12}$$

$$\sigma_e^2 = \frac{1}{12}\left(\frac{R}{2^B}\right)^2 = \frac{R^2}{12}\cdot 2^{-2B}$$

$B \rightarrow$ size of binary including sign bit

The variance of error signal is also called steady state noise power due to i/p quantization.

Steady state o/p noise power (variance) due to the quantization error signal:-

- The quantized i/p signal of a digital system can be represented as a sum of unquantized signal $x(n)$ and error signal $e(n)$

$$\therefore \quad x_q(n) = x(n) + e(n)$$

$h(n)$ is the impulse response of the system and $y'(n)$ is the response or o/p of the system due to input and error signal.

$$y'(n) = x_q(n) * h(n)$$

$$= [x(n) + e(n)] * h(n)$$

$$= [x(n) * h(n)] + [e(n) * h(n)]$$

$$y'(n) = y(n) + \varepsilon(n)$$

$y(n) = x(n) * h(n) =$ o/p due to i/p signal $x(n)$

$\varepsilon(n) = e(n) * h(n) =$ o/p due to error signal $e(n)$

Steady state o/p noise power due to i/p quantization error $\quad \sigma_{eoi}^2 = \sigma_e^2 \sum_{n=0}^{\infty} h^2(n)$

$$\sigma_{eoi}^2 = \sigma_e^2 \frac{1}{2\pi j} \oint_c H(z) H(z^{-1}) z^{-1} dz.$$

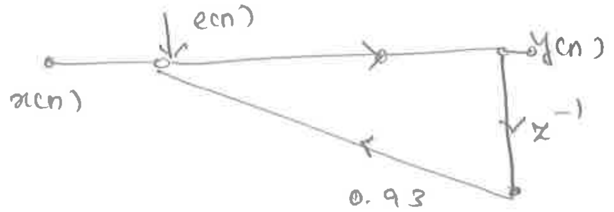
$$\sigma_{eoi}^2 = \sigma_e^2 \sum_{i=1}^{N} \left[\text{Res} \left[H(z) H(z^{-1}) z^{-1} \right] \right|_{z=P_i}$$

$$= \sigma_e^2 \sum_{i=1}^{N} \left[(z - P_i) H(z) H(z^{-1}) z^{-1} \right] \Big|_{z=P_i}$$

Since the closed contour integration is around the unit circle $|z| = 1$, only the residues for the poles that lie inside the unit circle in z-plane are considered.

Problems:

1) For the recursive filters shown in fig, the i/p $x(n)$ has a peak value of 10V represented by 6-bits. Compute the variance of o/p due to A/D conversion process.



Soln:

$$\text{Quantization step size } q = \frac{R}{2^B}$$

$$q = \frac{10}{2^6} = 0.15625$$

$$\text{Variance of error signal } \sigma_e^2 = \frac{q^2}{12} = \frac{0.15625^2}{12} = 2.0345 \times 10^{-3}$$

The difference equation of the above system without error is

$$y(n) = 0.93y(n-1) + x(n)$$

On taking z-Transform of above equation we get,

$$Y(z) = 0.93z^{-1}Y(z) + X(z)$$

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{1-0.93z^{-1}}$$

Output noise power due to A/D process.

$$\sigma_{eoi}^2 = \sigma_e^2 \frac{1}{2\pi j} \oint H(z)H(z^{-1})z^{-1}dz$$

$$H(z)H(z^{-1})z^{-1} = \frac{1}{1-0.93z^{-1}} \times \frac{1}{1-0.93z} \times z^{-1}$$

$$= \frac{z^{-1}}{\left(1-\frac{0.93}{z}\right)(-0.93)\left(z-\frac{1}{0.93}\right)}$$

$$= \frac{-1.0753z^{-1}}{\left(\frac{z-0.93}{z}\right)\left(z-1.0753\right)}$$

$$= \frac{-1.0753}{(z-0.93)(z-1.0753)}$$

$P_1 = 0.93 \qquad P_2 = 1.0753$

here $P_1 = 0.93$ is the only pole that lies inside the unit circle in z. plane.

The steady state o/p noise power due to i/p quantization error signal is given by,

$$\sigma_{eoi}^2 = \sigma_e^2 \frac{1}{2\pi j} \oint H(z) H(z^{-1}) z^{-1} dz.$$

$$= \sigma_e^2 \sum_{i=1}^{N} Res \left[H(z) H(z^{-1}) z^{-1} \right] \Big|_{z=P_i}$$

$$= \sigma_e^2 \sum_{i=1}^{N} \left[(z-P_i) H(z) H(z^{-1}) z^{-1} \right] \Big|_{z=P_i}$$

$$\sigma_{eoi}^2 = \sigma_e^2 \cancel{(z-0.93)} \times \frac{-1.0753}{\cancel{(z-0.93)}(z-1.0753)} \Big|_{z=0.93}$$

$$= \sigma_e^2 \times \frac{-1.0753}{0.93-1.0753} = 7.4006 \, \sigma_e^2$$

$$= 7.4006 \times 2.0345 \times 10^{-3}$$

$$\sigma_{eoi}^2 = 0.0151.$$

Exercise:

1. An LTI system is characterized by the difference equation $y(n) = 0.68 y(n-1) + 0.15 x(n)$. The i/p signal $x(n)$ has a range $-5V$ to $+5V$ represented by 8-bits. Find the quantization step size, variance of the error signal and variance of the quantization noise at the o/p.

$q = 0.0390625$

$\sigma_e^2 = 1.2716 \times 10^{-4}$

$\sigma_{eoi}^2 = 5.328 \times 10^{-6}$

2) The o/p of an A/D converter is applied to a digital filter with the system function $H(z) = \frac{0.45z}{z-0.72}$. Find the o/p noise power for the digital filter, when the i/p signal is quantized to 7 bits.

$q = 0.015625$

$\sigma_e^2 = 2.0345 \times 10^{-5}$

$\sigma_{eoi}^2 = 8.5551 \times 10^{-6}$

Quantization filter co-efficients:-

- The filter co-efficients are quantized to the word size of the register used to store them either by truncation or by rounding.
- The sensitivity of the filter frequency response characteristics to quantization of the filter co-efficients is minimized by realizing the filter having a large no. of poles and zeros as an interconnection of second order sections.
- It is possible to prove that the co. efficient quantization has less effect in cascade realization when compared to parallel realization.

Problems:-

1) For second order IIR filter $H(z) = \frac{1}{(1-0.5z^{-1})(1-0.45z^{-1})}$

Study the effect of shift in pole location with 3-bit co-efficient representation in direct and cascade form.

$$H(z) = \frac{1}{(1-0.5z^{-1})(1-0.45z^{-1})} = \frac{1}{z^{-1}(z-0.5)(z-0.45)}$$

$$H(z) = \frac{z}{(z-0.5)(z-0.45)}$$

$\therefore \; P_1 = 0.5 \qquad P_2 = 0.45$

i) Direct form realization:

$$H(z) = \frac{1}{1 - 0.95z^{-1} + 0.225z^{-2}}$$

quantizing the co-efficients by truncation.

$$.95_{10} \longrightarrow .1111 \longrightarrow .111 \longrightarrow .875$$

$$.225 \longrightarrow .0011 \longrightarrow .001 \longrightarrow .125$$

$\bar{H}(z)$ be the transfer function of the IIR system after quantizing the co-efficients.

$$\bar{H}(z) = \frac{1}{1 - 0.875z^{-1} + 0.125z^{-2}}.$$

Let us examine the poles of the sy after co-efficient quantization.

$$\bar{H}(z) = \frac{1}{z^{-2}(z^2 - 0.875z + 0.125)}$$

$$= \frac{z^2}{(z - 0.695)(z - 0.18)}$$

$\therefore \; P_{q1} = 0.695 \qquad P_{q2} = 0.18$

We can observe the poles of $\bar{H}(z)$ deviate very much from the original poles.

Cascade realization:

$$H(z) = \frac{1}{(1-0.5z^{-1})(1-0.45z^{-1})}$$

In cascade the system can be realized as cascade of first order sections.

$$H(z) = H_1(z) \cdot H_2(z)$$

$$H_1(z) = \frac{1}{1-0.5z^{-1}} \qquad H_2(z) = \frac{1}{1-0.45z^{-1}}$$

$$.5 \rightarrow .1000 \longrightarrow .100 \longrightarrow .5$$

$$.45 \rightarrow .0111 \longrightarrow .011 \longrightarrow .375$$

$$\overline{H_1(z)} = \frac{1}{1-0.5z^{-1}} \qquad \overline{H_2(z)} = \frac{1}{1-0.375z^{-1}}$$

$$\overline{P_{c1}} = 0.5 \qquad \overline{P_{c2}} = 0.375$$

On comparing the poles we can say that one of the poles is same and the other pole is close to the original pole.

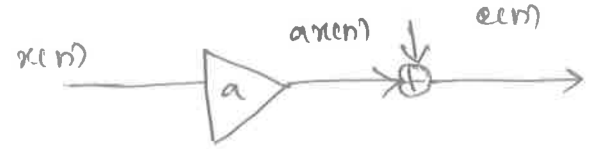
Exercise:

1. Discuss the effect of co-efficient quantization on pole locations of the following IIR system, when it is realized in cascade form and in direct form-I. Assume a word length of 4-bits through Truncation.

$$H(z) = \frac{1}{1-0.7z^{-1}+0.12z^{-2}}$$

Product Quantization error:-

The error due to quantization of the o/p' of multiplier is referred to as product quantization error.



$Q[ax(n)] = ax(n) + e(n)$

$ax(n)$ = unquantized product

$e(n)$ = product quantization error signal

Output noise power due to product Quantization:-

$e_k(n)$ = Error sig from k^{th} noise source

$h_k(n)$ = Impulse response " " "

$T_k(z) = Z\{h_k(n)\}$ Noise Transfer Function (NTF) for k^{th} noise source.

σ_{ek}^2 = Variance of k^{th} noise source

σ_{ekop}^2 = O/P noise power or variance due to k^{th} noise source.

$$\sigma_{ek}^2 = \frac{q^2}{12} \text{ or } \frac{2^{-2B}}{12}$$

$$\sigma_{ekop}^2 = \sigma_{ek}^2 \frac{1}{2\pi j} \oint T_k(z) T_k(z^{-1}) z^{-1} dz.$$

$$\therefore \sigma_{ekop}^2 = \sigma_{ek}^2 \sum_{i=1}^{N} Res \left[T_k(z) T_k(z^{-1}) z^{-1}\right]\Big|_{z=P_i}$$

$$= \sigma_{ek}^2 \sum_{i=1}^{N} \left[(z-P_i) T_k(z) T_k(z^{-1}) z^{-1}\right]\Big|_{z=P_i}$$

The total steady state noise variance at the output of the system due to product quantization errors is given by the sum of the o/p noise variances due to all the noise sources.

Problems:

1) In the IIR system given below the products are rounded to 4-bits $H(z) = \dfrac{1}{(1-0.35z^{-1})(1-0.62z^{-1})}$

Find the o/p round off noise power in

a) direct form realization b) Cascade realization.

Soln:

1) Direct form realization:

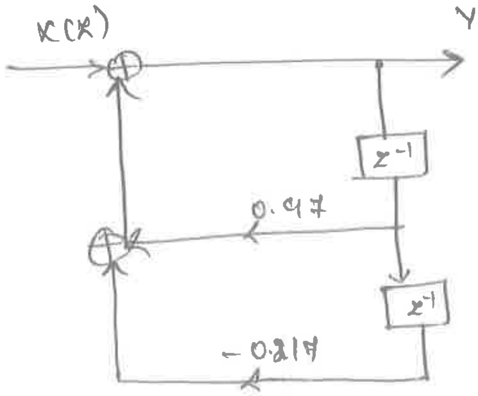
$$H(z) = \frac{1}{1-0.62z^{-1}-0.35z^{-1}+0.217z^{-2}} = \frac{1}{1-0.97z^{-1}+0.217z^{-2}}$$

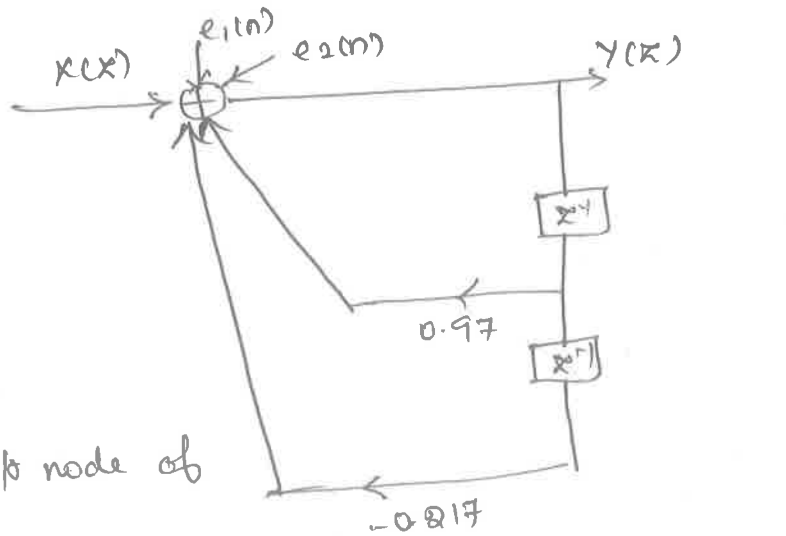
$$H(z) = \frac{Y(z)}{X(z)} = \frac{1}{1-0.97z^{-1}+0.217z^{-2}}.$$

$$Y(z)-0.97z^{-1}Y(z)+0.217z^{-2}Y(z) = X(z).$$

$$Y(z) = X(z)+0.97z^{-1}Y(z) - 0.217z^{-2}Y(z)$$

Direct form structure:- noise model.





both noise sources are at the i/p node of the s/y.

NTF for noise signal $e_1(n) = T_1(z) = H(z) = \dfrac{1}{(1-0.35z^{-1})(1-0.62z^{-1})}$

$T_2(z) = H(z) = \dfrac{1}{(1-0.35z^{-1})(1-0.62z^{-1})}$

$$\sigma_{e1op}^2 = \sigma_e^2 \frac{1}{2\pi j} \oint T_1(x)\, T_1(x^{-1})\, x^{-1}\, dx$$

$$T_1(x) \cdot T_1(x^{-1}) \cdot x^{-1} = \frac{4.6083 \cdot x}{(x-0.35)(x-0.62)(x-2.8571)(x-1.6129)}$$

$P_1 = 0.35 \quad P_2 = 0.62$.

$$\sigma_{e1op}^2 = \sigma_e^2 \times \sum_{i=1}^{2} \left[(x-P_i)\, T_1(x)\, T_1(x^{-1})\, x^{-1} \right] \Big|_{x=P_i}$$

$$\sigma_{e1op}^2 = \sigma_e^2 \times \sum \left[\frac{\cancel{(x-0.35)}\ 4.6083x}{\cancel{(x-0.35)}\,(x-0.62)(x-2.8571)(x-1.6129)} \right] \Bigg|_{x=0.35} + \left[\frac{\cancel{(x-0.62)}\ 4.6083x}{(x-0.35)\,\cancel{(x-0.62)}\,(x-2.8571)(x-1.6129)} \right] \Bigg|_{x=0.62}$$

$$= \sigma_e^2 \times (-1.8867 + 4.7641)$$

$$= 1.3021 \times 10^{-3} \times (-1.8867 + 4.7641)$$

$$\sigma_{e1op}^2 = 3.7467 \times 10^{-3} = \sigma_{e2op}^2$$

Total o/p noise power due to all the noise sources in direct form realization.

$$\sigma_{etop}^2 = \sigma_{e1op}^2 + \sigma_{e2op}^2 = 2 \times 3.7467 \times 10^{-3} = 7.4934 \times 10^{-3}.$$

Cascade Realization:

$$H(x) = \frac{1}{(1-0.5x^{-1})(1-0.45x^{-1})}$$

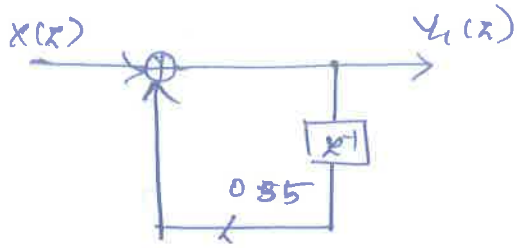
$$H_1(x) = \frac{1}{1-0.5x^{-1}} \qquad H_2(x) = \frac{1}{1-0.45x^{-1}}$$

$$H(z) = \frac{1}{(1-0.35z^{-1})(1-0.62z^{-1})}$$

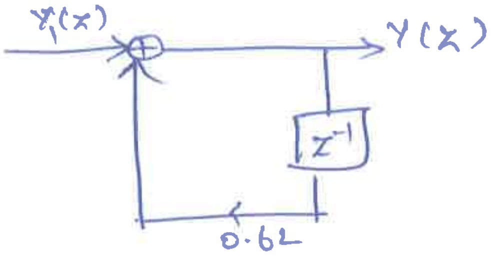
$$H(z) = H_1(z) \cdot H_2(z)$$

$$H_1(z) = \frac{1}{1-0.35z^{-1}} \qquad H_2(z) = \frac{1}{1-0.62z^{-1}}$$

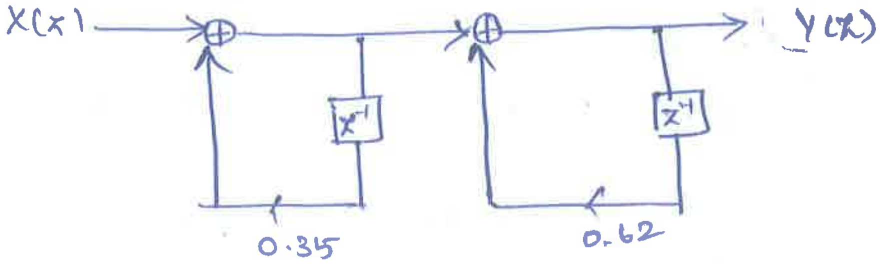
$$H_1(z) = \frac{Y_1(z)}{X(z)} = \frac{1}{1-0.35z^{-1}}$$



$$H_2(z) = \frac{Y(z)}{Y_1(z)} = \frac{1}{1-0.62z^{-1}}$$



On cascading.



NTF for signal $e_1(n) \Rightarrow T_1(z) = \dfrac{1}{(1-0.35z^{-1})(1-0.62z^{-1})}$

NTF for signal $e_2(n) \rightarrow T_2(z) = \dfrac{1}{1-0.62z^{-1}}$.

σ^2_{e1op} for $e_1(n)$ and σ^2_{e2op} for $e_2(n)$

σ^2_{e1op} is same as that of cascade form Realization.

$$\therefore \sigma^2_{e1op} = 3.7467 \times 10^{-3}.$$

$$\sigma^2_{e2op} = \sigma_e^2 \frac{1}{2\pi j} \oint T_2(z) T_2(z^{-1}) z^{-1} dz.$$

Q.

$$= \sigma_e^2 \sum_{i=1}^{N} \text{Res}\left[T_2(z)\, T_2(z^{-1}) z^{-1}\right] \Big|_{z=P_i}.$$

$$= \sigma_e^2 \sum_{i=1}^{N} (z-P_i) \cdot T_2(z)\, T_2(z^{-1}) z^{-1} \Big|_{z=P_i}$$

$$T_2(z) T_2(z^{-1}) z^{-1} = \frac{1}{1-0.62z^{-1}} \cdot \frac{1}{1-0.62z} \cdot z^{-1}$$

$$= \frac{z^{-1} \cdot z.}{(z-0.62)\,(-0.62)\,(z-1.6129)}$$

$$= \frac{-1.6129}{(z-0.62)(z-1.6129)}$$

$P_1 = 0.62.$

$$= \sigma_e^2 \left[\frac{\cancel{(z-0.62)}\; -1.6129}{\cancel{(z-0.62)}\,(z-1.6129)} \right] \Bigg|_{z=0.62}$$

$$= 1.3021 \times 10^{-3} \times \frac{-1.6129}{0.62-1.6129} = 2.1152 \times 10^{-3}.$$

$$\sigma_{eTop}^2 = \sigma_{top,q}^2 + \sigma_{e2op}^2$$

$$= 3.7467 \times 10^{-3} + 2.1152 \times 10^{-3}.$$

$$\boxed{\sigma_{eTop}^2 = 5.8619 \times 10^{-3}.}$$

Cascading $H_2(z) \cdot H_1(z)$

$$\sigma_{eTop}^2 = 5.2306 \times 10^{-3}.$$