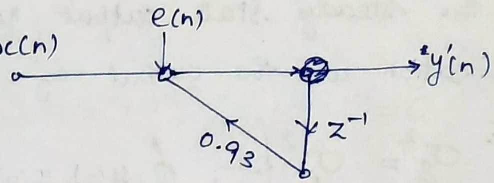
H.W: For the recursive filter shown in fig below — the input $x(n)$ has a peak value of 10V, represented by 6 bits. Compute the variance of output due to ADC process.



Soln:

hint: $q = \frac{\text{range of signal}}{\text{No. of quantization levels}}$

range is $\pm 5V = 10V$

No. of quantization levels $= 2^b = 2^6$

diff. eqn from fig: $y(n) = 0.93\, y(n-1) + x(n)$

Coefficient Quantization Error:

$\Rightarrow$ error occurs due to quantization of filter coefficients.

When digital filters are designed, coefficients are evaluated with infinite precision. But the size of the binary register used to store the coefficients is of finite duration. Hence the filter coefficients are quantized to the word size of the register used to store them either by truncation or by rounding.

Generally the location (or the value) of poles and zeros of the digital filters directly depends on the value of filter coefficients. When we quantize the filter coefficients, the value of poles & zeros will be modified, hence the location of poles and zeros will be shifted from the desired location. This will create deviations in the frequency response of the system. Hence we obtain a filter having a frequency response that is different from the frequency response of the filter with unquantized coefficients.

If the poles of the desired filter are close to the unit circle (ie, in narrow band filters particularly), then those of the filter with quantized coefficients may lie just outside the unit circle leading to instability.

When we realize the higher order filter (i.e, filter with large number of poles and zeros) as an interconnection of second order sections (i.e cascade or parallel realization), sensitivity of the filter frequency response characteristics to quantization of filter coefficients is minimized, Because cascade and parallel form realization contains first and second order sections.

Finally we can prove that the coefficient quantization has less effect in cascade realization when compared to parallel realization.

Pbm: Consider a second order IIR filter with

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

Find the effect of shift in pole locations with 3-bit coefficient representation in direct and cascade form.

Solution: $b=3$

no. of poles are 2.

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

$$= \frac{z^2}{z^2 - 0.95z + 0.225}$$

$$H(z) = \frac{z^2}{(z-0.5)(z-0.45)}$$

poles are $\boxed{P_1 = 0.5, \; P_2 = 0.45}$

__Direct form__ :

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

Coefficients are : 0.95, 0.225

Let us truncate the coefficient by truncation,

$$(0.95)_{10} = (0.1111)_2 \xrightarrow[\text{to 3 bits}]{\text{Truncate}} (0.111)_2 = (0.875)_{10}$$

$$(0.225)_{10} = (0.0011)_2 \xrightarrow[\text{to 3-bits}]{\text{truncate}} (0.001)_2 = (0.125)_{10}$$

quantized coefficients are : 0.875, 0.125

$\therefore$ Transfer function of IIR filter after quantizing the coefficients,

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

Let the poles of $\overline{H(z)}$, $\bar{P}_1$ & $\bar{P}_2$.

roots of the quatratic equation of $\overline{H(z)}$,

$$z^2 - 0.875z + 0.125 = 0$$

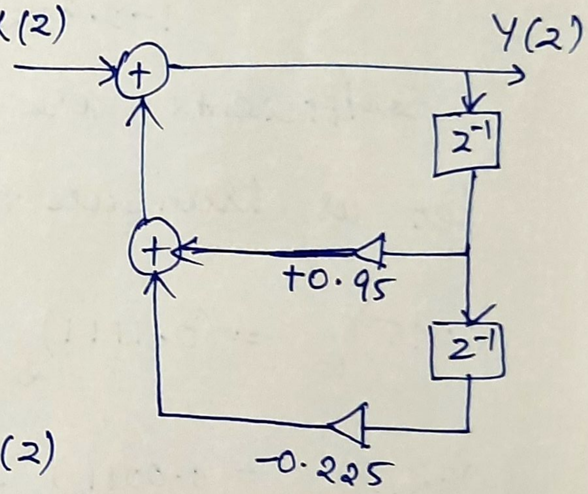
$$z = \frac{0.875 \pm \sqrt{0.875^2 - 4(0.125)}}{2a}$$

$$= 0.695 \text{ (or) } 0.18$$

$$\boxed{\overline{P_1} = 0.695 \quad ; \quad \overline{P_2} = 0.18}$$

Due to quantization, poles of $\overline{H(z)}$ deviate very much from the poles of $H(z)$.

<u>Given</u> $H(z)$ in direct form realization

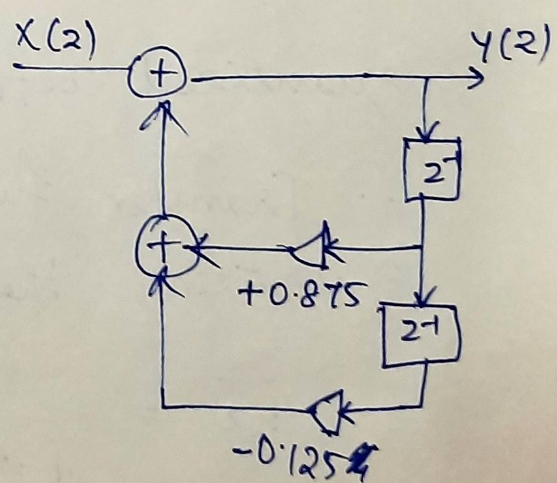


$$H(z) = \frac{1}{1 - 0.95z^{-1} + 0.225z^{-2}}$$

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

$$Y(z) = 0.95z^{-1}Y(z) - 0.225z^{-2}Y(z) + X(z)$$

After Quantization:



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

$$\frac{Y(z)}{X(z)} = \text{''}$$

$$Y(z) - 0.875z^{-1}Y(z) + 0.125z^{-2}Y(z) = X(z)$$

$$Y(z) = 0.875z^{-1}Y(z) - 0.125z^{-2}Y(z) + X(z)$$

Cascade realization:

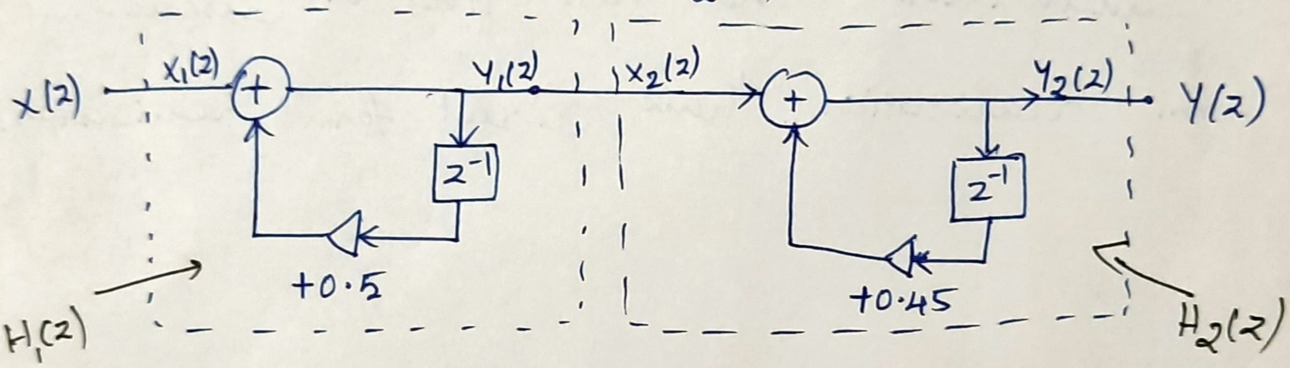
$$H(z) = \frac{1}{1-0.95z^{-1}+0.225z^{-2}}$$

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

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

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

$$\frac{Y_1(z)}{X_1(z)} = \frac{1}{1-0.5z^{-1}}, \quad \frac{Y_2(z)}{X_2(z)} = \frac{1}{1-0.45z^{-1}}$$



Let us quantize the coefficients of $H_1(z)$ & $H_2(z)$ by truncation,
(0.5 & 0.45)

$$(0.5)_{10} = (0.1000)_2 \xrightarrow[\text{to 3 bits}]{\text{truncate}} (0.100)_2 = (0.5)_{10}$$

$$(0.45)_{10} = (0.0111)_2 \xrightarrow[\text{to 3 bits}]{\text{truncate}} (0.011)_2 = (0.375)_{10}$$

After quantization,

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

draw the cascade form structure

quantized coefficients are : 0.5 & 0.375

before quantization	After quantization
$P_1 = 0.5$, $P_2 = 0.45$	$\bar{P}_1 = 0.5$, $\bar{P}_2 = 0.375$.

If we compare the poles before & after quantization, one of the pole is same & other pole is very close to orignal pole.

Conclusion:

Quantization has less effect in cascade form realization than direct form realization.