COEFFICIENT QUANTIZATION ERROR

–

–

Co-efficient quantization error

- We know that the IIR Filter is characterized by the system function

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

- After quantizing ,

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

Where $[a_k]_q = a_k + \Delta a_k$

$[b_k]_q = b_k + \Delta b_k$

- The quantization of filter coefficients alters the positions of the poles and zeros in z-plane.
  1. If the poles of desired filter lie close to the unit circle, then the quantized filter poles may lie outside the unit circle leading into instability of filter.
  2. Deviation in poles and zeros also lead to deviation in frequency response.

****

Consider a second order IIR filter with $H(z) = \frac{1.0}{(1-0.5z^{-1})(1-0.45z^{-1})}$ **find the effect on quantization on pole locations of the given system function in direct form and in cascade form. Take b=3bits. [Apr/May-10] [Nov/Dec-11]**

Solution:

Given that,

$$H(z) = \frac{1.0}{(1-0.5z^{-1})(1-0.45z^{-1})}$$

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

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

The roots of the denominator of H(z) are the original poles of H(z). let the original poles of H(z) be p_1 and p_2.

Here $p_1=0.5$ and $p_2=0.45$

Direct form I:

$$H(z)=\frac{1.0}{(1-0.5z^{-1})(1-0.45z^{-1})}$$

$$H(z)=\frac{1}{1-0.5z^{-1}-0.45z^{-1}+0.225z-2}$$

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

Let us quantize the coefficients by truncation.

$$.95_{10}\xrightarrow[\text{Binary}]{\text{Convert to}}.1111_2\xrightarrow[\text{3-bits}]{\text{Truncate to}}.111_2\xrightarrow[\text{decimal}]{\text{Convert to}}.875_{10}$$

$$.225_{10}\xrightarrow[\text{Binary}]{\text{Convert to}}.0011_2\xrightarrow[\text{3-bits}]{\text{Truncate to}}.001_2\xrightarrow[\text{decimal}]{\text{Convert to}}.125_{10}$$

Let $\overline{H}(z)$ be the transfer function of the IIR system after quantizing the coefficients.

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

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

On cross multiplying the above equation we get,

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

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

<u>Cascade form</u>:

Given that

$$H(z)=\frac{1.0}{(1-0.5z^{-1})(1-0.45z^{-1})}$$

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

$H(z)=H_1(z)+H_2(z)$

Where,

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

Let us quantize the coefficients of $H_1(z)$ and $H_2(z)$ by truncation.

$$.5_{10}\xrightarrow[\text{Binary}]{\text{Convert to}}.1000_2\xrightarrow[\text{3-bits}]{\text{Truncate to}}.100_2\xrightarrow[\text{decimal}]{\text{Convert to}}.5_{10}$$

$$.45_{10}\xrightarrow[\text{Binary}]{\text{Convert to}}.0111_2\xrightarrow[\text{3-bits}]{\text{Convert to}}.011_2\xrightarrow[\text{decimal}]{\text{Convert to}}.375_{10}$$

let , $\overline{H_1}(z)$ and $\overline{H_2}(z)$ be the transfer function of the first-order sections after quantizing the coefficients.

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

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

$$let, \overline{H_1}(z) = \frac{Y_1(z)}{X(z)} = \frac{1}{1-0.5z^{-1}}$$

$$Y_1(z) - 0.5z^{-1}Y_1(z) = X(z)$$

$$Y_1(z) = X(z) + 0.5z^{-1}Y_1(z)$$

$$\text{let}, \overline{H_2}(z) = \frac{Y(z)}{Y_1(z)} = \frac{1}{1-0.375z^{-1}}$$

on cross multiplying the above equation we get,

$$Y(z) - 0.375z^{-1}Y(z) = Y_1(z)$$

$$Y(z) = Y_1(z) + 0.375z^{-1}Y(z)$$

****