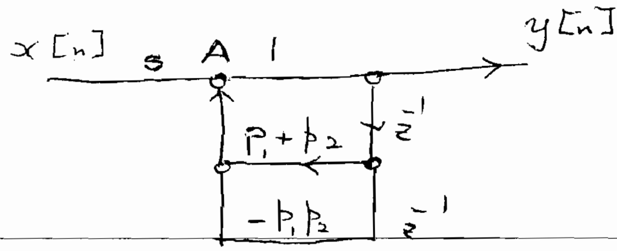
ECE-539, Digital Signal Process.

Spring 2011, BALU SANTHANAM

Example: Scaling & Overflow

Consider the direct form implementation of the two-pole system:

$$H(z) = \frac{1}{(1-p_1 z^{-1})(1-p_2 z^{-1})}, \quad |z| > \max\{|p_1|, |p_2|\}$$



The node A in this structure represents a likely point in this implementation for overflow due to quantization of sum of products. This can be averted by scaling the input by a factor s using the scaling paradigms described in class. Towards this purpose, we evaluate the different norms of $h[n]$

$$H(z) = \frac{1}{(1 - p_1 z^{-1})(1 - p_2 z^{-1})}$$

$$= \frac{k_1}{1 - p_1 z^{-1}} + \frac{k_2}{1 - p_2 z^{-1}}$$

$$k_1(1 - p_2 z^{-1}) + k_2(1 - p_1 z^{-1}) = 1$$

Comparing coefficients:

$$k_2\left(1 - \frac{p_1}{p_2}\right) = 1$$

$$k_2 = \frac{p_2}{p_2 - p_1}, \quad k_1 = \frac{-p_1}{p_2 - p_1}$$

$$h[n] = \frac{-p_1}{p_2 - p_1}(p_1)^n u[n] + \frac{p_2}{p_2 - p_1}(p_2)^n u[n]$$

$$= \frac{p_2^{n+1} - p_1^{n+1}}{p_2 - p_1} u[n]$$

$$\sum_{n=0}^{\infty} |h[n]|^2 = \frac{\dfrac{1 + p_1 p_2}{1 - p_1 p_2}}{(1 - p_1^2)(1 - p_2^2)}$$

(From previous results)

3-0235 — 50 SHEETS — 5 SQUARES
3-0236 — 100 SHEETS — 5 SQUARES
3-0237 — 200 SHEETS — 5 SQUARE
3-0137 — 200 SHEETS — FILLER

COMET

$$\sum_{n=0}^{\infty} |h[n]| = \sum_{n=0}^{\infty} \left| \frac{p_2^{n+1} - p_1^{n+1}}{p_2 - p_1} \right|$$

Assuming $p_2 > p_1$

$$\sum_{n=0}^{\infty} |h[n]| = \frac{1}{p_2 - p_1} \sum_{n=0}^{\infty} p_2^{n+1} - \frac{1}{p_2 - p_1} \sum_{n=0}^{\infty} p_1^{n+1}$$

$$= \frac{p_2}{p_2 - p_1} \frac{1}{1 - p_2} - \frac{p_1}{p_2 - p_1} \frac{1}{1 - p_1}$$

$$= \frac{1}{p_2 - p_1} \left\{ \frac{p_2}{1 - p_2} - \frac{p_1}{1 - p_1} \right\}$$

$$= \frac{1}{p_2 - p_1} \frac{p_2 \cancel{- p_1 p_2} - p_1 \cancel{+ p_1 p_2}}{(1 - p_1)(1 - p_2)}$$

$$= \frac{1}{(1 - p_1)(1 - p_2)}$$

$\Rightarrow$ The ℓ^1 norm scale factor is given by $S = (1 - p_1)(1 - p_2)$

$\Rightarrow$ This will ensure that there is no overlap at the node A

$\Rightarrow$ For p_1 and p_2 close to the unit circle $(1 - p_1)(1 - p_2)$ may be very small reducing the SNR

→ The ℓ^2 norm scale factor is given by:

$$S_2 = \sqrt{\frac{(1-p_1^2)(1-p_2^2)}{\left(\frac{1+p_1 p_2}{1-p_1 p_2}\right)}}$$

This will not be as severe as ℓ^1 norm scaling but increases the likelihood of overflow at node A.

→ This bounds the energy of the output to node A to 1.

→ As can be seen SNR & input scaling work against each other.