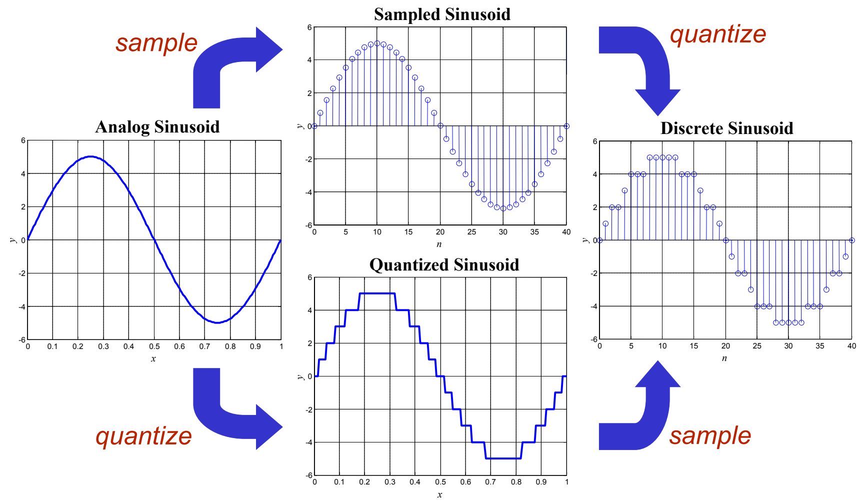
ECE 4270
Fundamentals of Digital Signal Processing

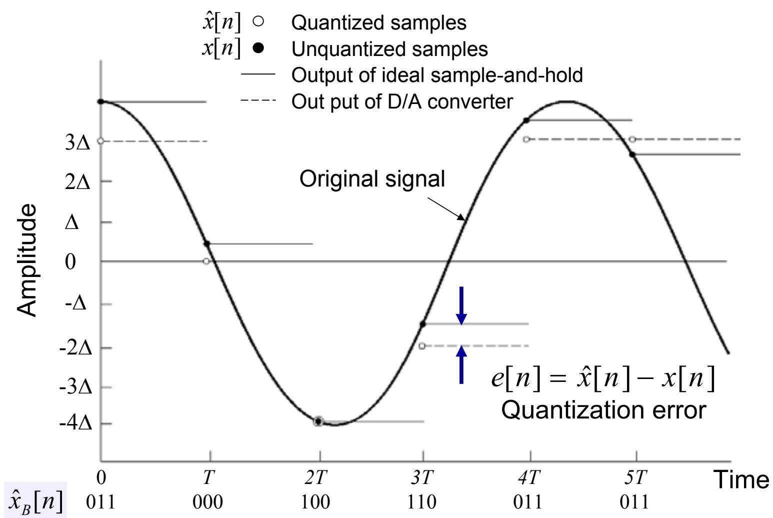
Lecture 11: Quantization, & Frequency Response of Rational Systems

School of Electrical and Computer Engineering
Georgia Institute of Technology
Summer, 2004

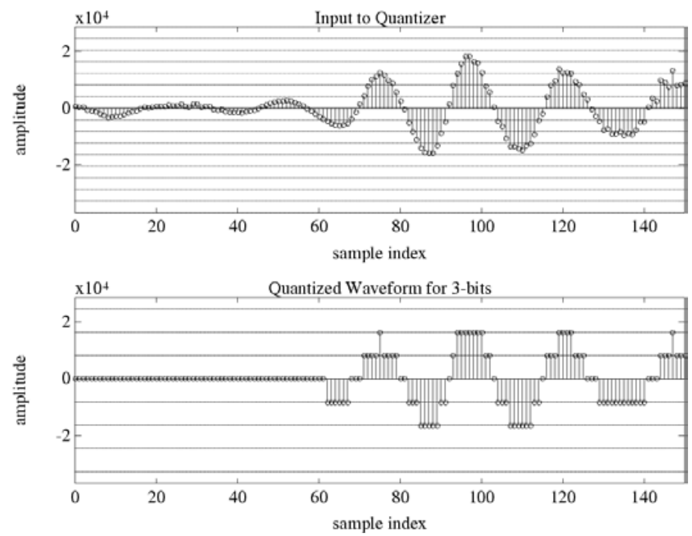
Discrete Signals



A-to-D and D-to-A Conversion



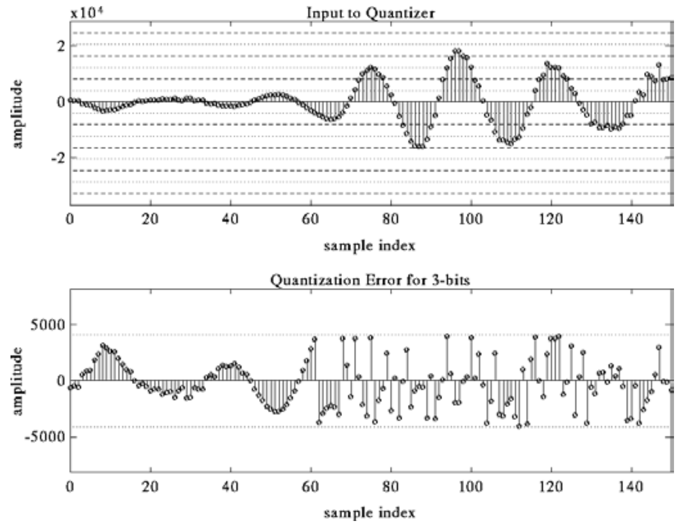
3-Bit Speech Quantization



Input to quantizer

3-bit quantized waveform

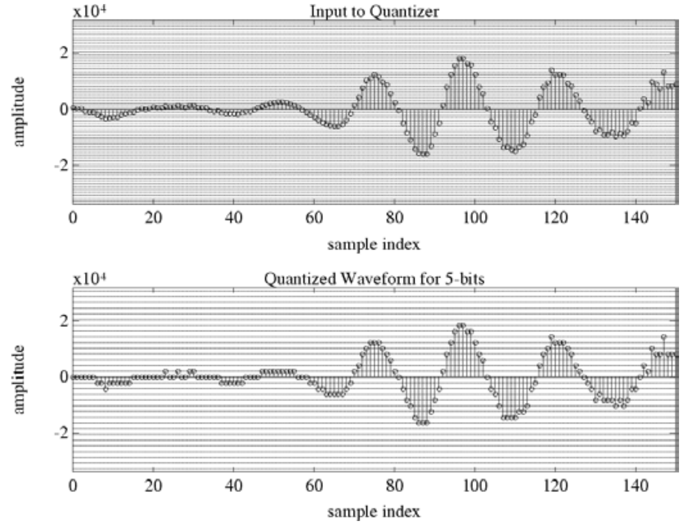
3-Bit Speech Quantization Error



Input to quantizer

3-bit quantization error

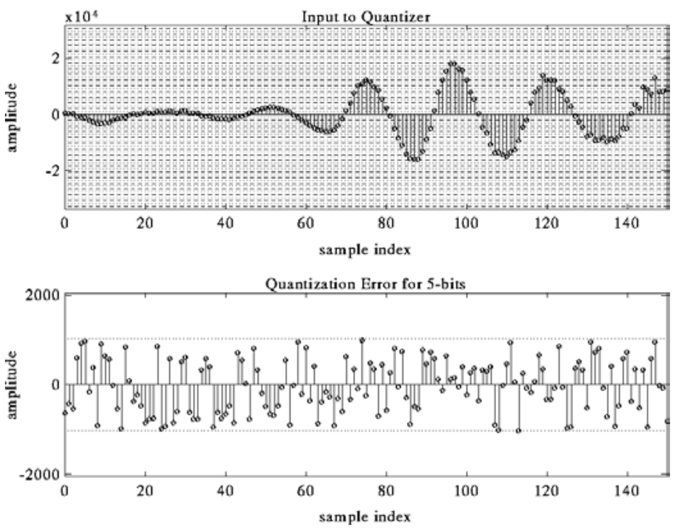
5-Bit Speech Quantization



Input to quantizer

5-bit quantized waveform

5-Bit Speech Quantization Error

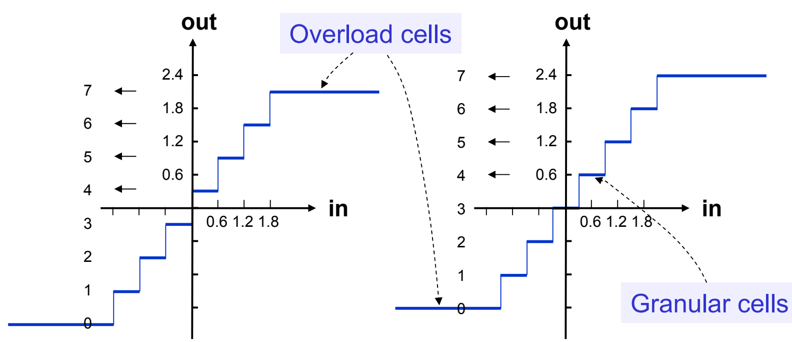


Input to quantizer

5-bit quantization error

Scalar Quantization & Quantization Error

granular noise → error in granular cells
overload noise → error in overload cells



Mid-rise uniform staircase quantizer

Mid-tread uniform staircase quantizer

Vector Quantization

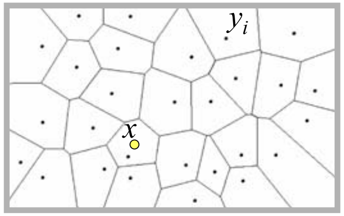
- Distortion function:

  $\delta(x, y)$

  measures the dissimilarity between x and y.

- Centroid of a region Y:

$$\bar{y} = \min_{y} \int_Y \delta(x, y) dx$$



- Voronoi region:

$$Y_i = \{x;\ \delta(x, y_i) = \min_{j=1,2,..,N} \delta(x, y_j)\}$$

Granular Quantization Error

- Each sample is quantized and each sample has a quantization error defined as

$$e[n] = \hat{x}[n] - x[n]$$

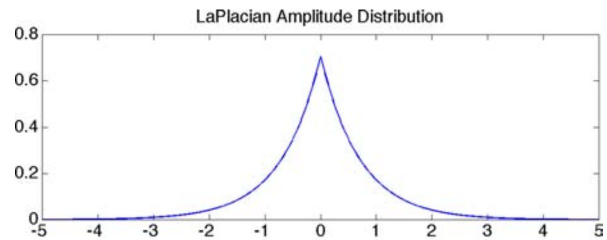
- Since each sample falls in an interval of length Δ, and the quantized sample falls in the middle of that interval,

$$-(\Delta / 2) < e[n] \le (\Delta / 2).$$

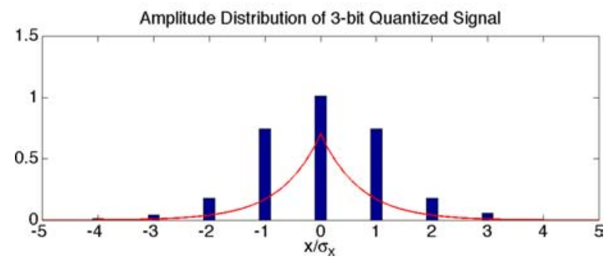
- We call this "quantization noise" because it seems to vary randomly. Clearly, the strength (power) of this noise is proportional to Δ; *i.e.*,

$$\sigma_e^2 = K\Delta^2$$

Typical Amplitude Distributions



Laplacian distribution is often used as a model for speech signals

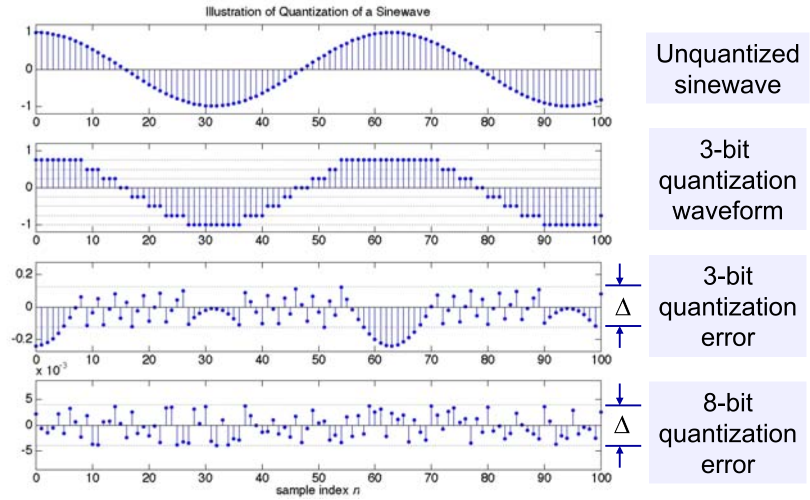


A 3-bit quantized signal has only 8 different values

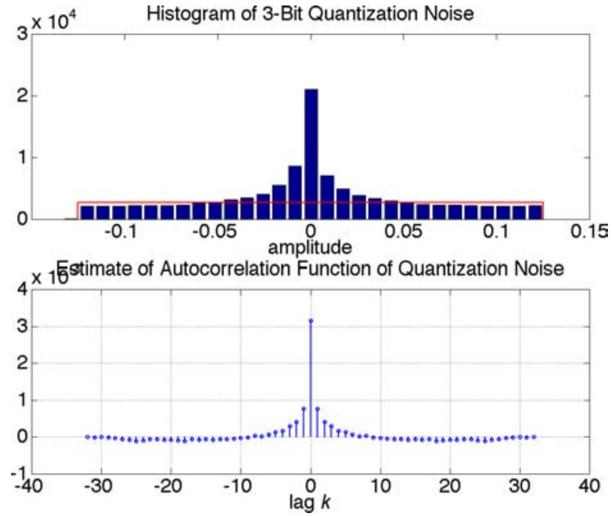
Probabilistic Model for Quantization

- We observed that the quantization error has very complicated variations that suggest a random or noise-like character.
- Random signals are represented by probability distributions and averages such as
  - Mean and mean-square (average power)
  - Histograms
  - Autocorrelation function
  - Power spectrum
- This is a good way to think about quantization noise.

Quantization of a Sine Wave

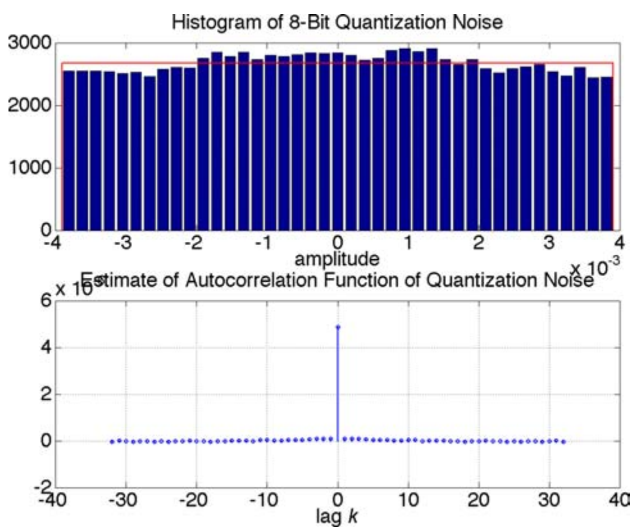


3-Bit Quantization Noise



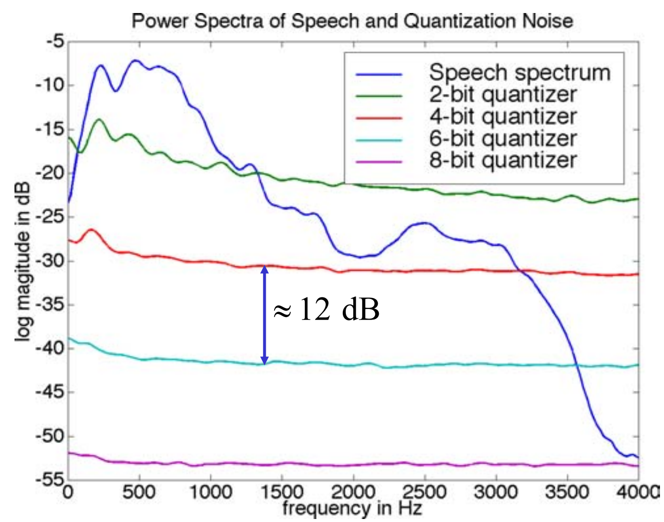
$$\Delta = 2/2^3 = 0.25$$

8-Bit Quantization Noise

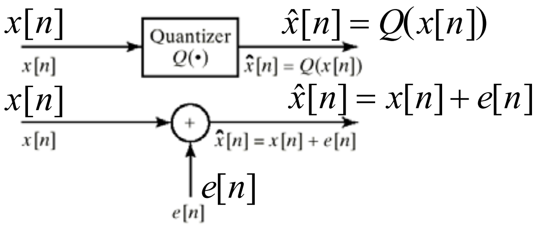


$$\Delta = 2/2^8 = 0.0078$$

Spectra of Quantization Noise



Linear Noise Model



- Error is uncorrelated with the input.
- Error is uniformly distributed over the interval

$$-(\Delta/2) < e[n] \le (\Delta/2).$$

- Error is stationary white noise, (*i.e.* flat spectrum)

$$\Phi_{ee}(\omega) = \sigma_e^2 = \int_{-\Delta/2}^{\Delta/2} \frac{1}{\Delta} e^2 de = \frac{\Delta^2}{12}, \quad |\omega| \le \pi$$

Quantizer Signal-to-Noise Ratio

- Assume $2^{(B+1)}$ levels and amplitude range $2X_m$. Then using a probabilistic analysis we obtain

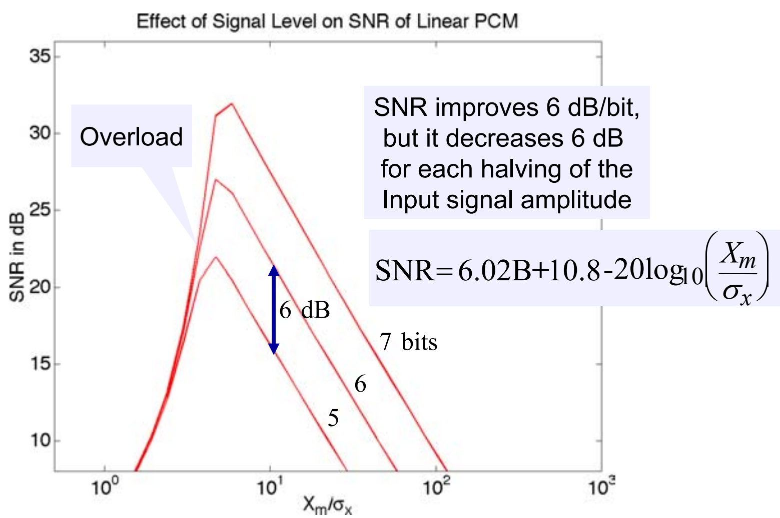
$$\Rightarrow \underbrace{\Delta = \frac{2X_m}{2^{(B+1)}} = 2^{-B} X_m}_{\text{step size}} \quad \Rightarrow \quad \underbrace{\sigma_e^2 = \frac{2^{-2B} X_m^2}{12}}_{\text{noise power}}$$

- Therefore the quantizer SNR is:

$$\text{SNR} = 10\log_{10}\left(\frac{\sigma_x^2}{\sigma_e^2}\right) = 10\log_{10}\left(\frac{12\, 2^{2B} \sigma_x^2}{X_m^2}\right)$$

$$= 6.02B + 10.8 - 20\log_{10}\left(\frac{X_m}{\sigma_x}\right) \quad \text{(in dB)}$$

Variation of SNR with Signal Level



Analysis of Fine Quantization

- Input is random variable $\mathbf{x}$ with pdf $p_{\mathbf{x}}(x)$
- Distortion is defined as MSE, i.e.

$$\delta(x, y) = \|x - y\|^2$$

- Quantizer Q, $Q(x) = y_i = \min_{j=1,2,..,N} \delta(x, y_j)$
- Reproduction codebook C, $C = \{y_j, j = 1, 2, .., N\}$
- Input space partition Y, $Y = \{Y_j, j = 1, 2, ..., N\}$
- Average distortion

$$D(Q) = E[\|\mathbf{x} - Q(\mathbf{x})\|^2] = \sum_{j=1}^{N} \int_{Y_j} \|x - y_j\|^2 p_{\mathbf{x}}(x) dx$$

Analysis of Fine Quantization (Large N)

Bennett assumptions

- N is very large
- Partition is Voronoi – reproduction values are the regional centroids
- Partition is regular – covers the whole space such that overload cells have nearly zero probability measure
- Cell volumns $V(Y_j) = \int_{Y_j} dx$ are all bounded and tiny
- $p_{\mathbf{x}}(x)$ is smooth, meaning Riemann sums approach Riemann integral and mean value theorem applies

Analysis of Fine Quantization (Large N)

$$D(Q) = \sum_{j:V(Y_j)<\infty} \int_{Y_j} \left\|x - y_j\right\|^2 p_{\mathbf{x}}(x)dx + \sum_{j:V(Y_j)=\infty} \int_{Y_j} \left\|x - y_j\right\|^2 p_{\mathbf{x}}(x)dx$$

$$D(Q) \approx \sum_{j:V(Y_j)<\infty} \int_{Y_j} \left\|x - y_j\right\|^2 p_{\mathbf{x}}(x)dx$$

$$\approx \sum_{j=1}^{N} p_{\mathbf{x}}(y_j) \int_{Y_j} \left\|x - y_j\right\|^2 dx$$

since $p_{\mathbf{x}}(x)$ is smooth and mean value theorem applies

But, $P_{\mathbf{x}}(Y_j) = \int_{Y_j} p_{\mathbf{x}}(x)dx \approx p_{\mathbf{x}}(y_j)\int_{Y_j} dx = p_{\mathbf{x}}(y_j)V(Y_j)$

$$D(Q) \approx \sum_{j=1}^{N} P_{\mathbf{x}}(Y_j) \int_{Y_j} \frac{\left\|x - y_j\right\|^2}{V(Y_j)} dx$$

← Relates to "normalized moment" – for further reading if interested

Quantization Summary

- Quantization of signal values and results of computation is unavoidable in a digital system.
- We can analyze quantization error using a random noise model.
- The more bits in the number representation, the lower the noise. An oft-stated theorem is that "the signal-to-noise ratio increases 6 dB with each added bit"
  - however, remember that if the signal level decreases while keeping the quantizer step size the same, it is like throwing away bits!

A/D & D/A Conversion

Digital-to-Analog Conversion

- Reconstruction from sampled sequence using ideal lowpass filtering

$$X_r(j\Omega) = X(e^{j\Omega T})H_r(j\Omega)$$

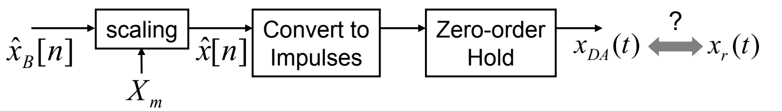
Discrete-time Fourier transform

Ideal lowpass filter $H_r(j\Omega) = \begin{cases} T, & |\Omega| < \pi/T \\ 0, & |\Omega| > \pi/T \end{cases}$

Ideal discrete to continuous conversion

$$x_r(t) = \sum_{n=-\infty}^{\infty} x[n]\frac{\sin[\pi(t-nT)/T]}{\pi(t-nT)/T}$$

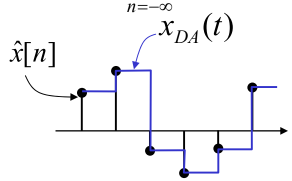
- Realizable implementation in circuitry – D/A converter



D/A Conversion - II

$$x_{DA}(t) = \sum_{n=-\infty}^{\infty} \hat{x}[n]h_0(t-nT) = x_0(t) + e_0(t)$$

$$x_0(t) = \sum_{n=-\infty}^{\infty} x[n]h_0(t-nT) \qquad e_0(t) = \sum_{n=-\infty}^{\infty} e[n]h_0(t-nT)$$



$$\hat{x}[n] = x[n] + e[n]$$

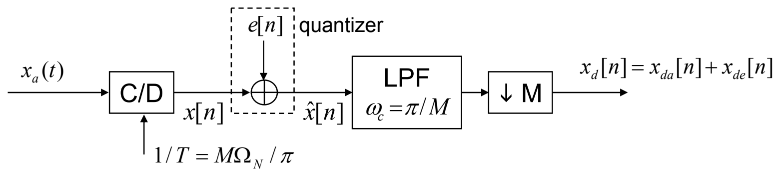
quantization noise ($e[n]$); original signal $x_a(nT)$ ($x[n]$)

$$h_0(t) = \begin{cases} 1, & 0 < t < T \\ 0, & \text{otherwise} \end{cases} \Longleftrightarrow H_0(j\Omega) = \frac{2\sin(\Omega T/2)}{\Omega}e^{-j\Omega T/2}$$

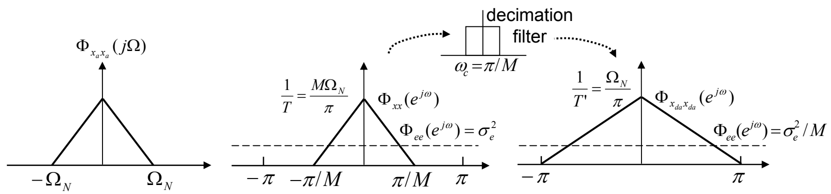
See Fig.4.54

Quantization noise aside, we need to compensate for the effect of zero order hold circuit to avoid spectral distortion ➔ $\tilde{H}_r(j\Omega) = H_r(j\Omega)/H_0(j\Omega)$

Over-Sampling & Its Effect on Q-Noise



- M is called oversampling ratio (above Nyquist rate)
- M affects signal-to-quantization-noise ratio

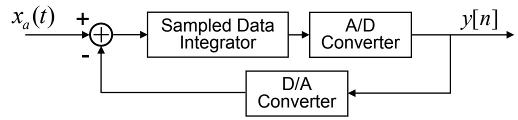


Power spectrum of analog signal

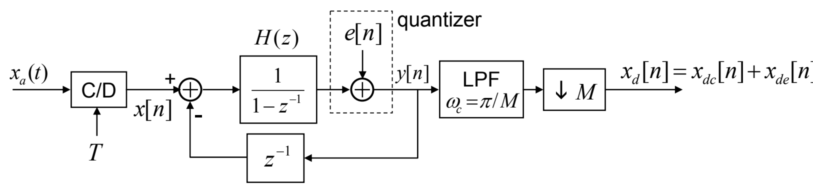
Power spectral density of over-sampled signal and q-noise

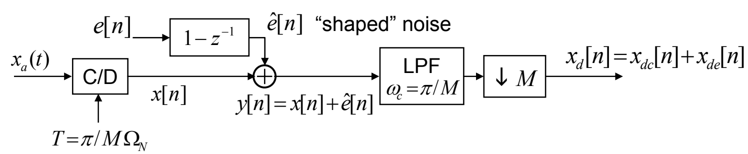
Power spectral density of signal and q-noise after down-sampling

Oversampled A/D with Noise Shaping



Delta-Sigma modulator;
Sigma-Delta A/D/A converter

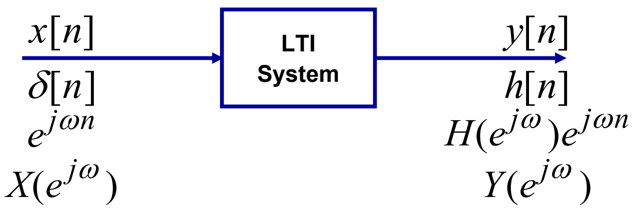




Frequency Response of LTI Systems

Convolution Theorem

$$y[n]=\sum_{k=-\infty}^{\infty} x[k]h[n-k] \Leftrightarrow Y(e^{j\omega})=X(e^{j\omega})H(e^{j\omega})$$

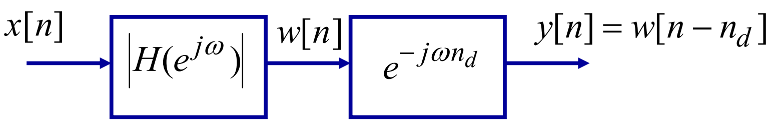


$$Y(e^{j\omega})=H(e^{j\omega})X(e^{j\omega})$$

$$|Y(e^{j\omega})|=|H(e^{j\omega})||X(e^{j\omega})|$$

$$\angle Y(e^{j\omega})=\angle H(e^{j\omega})+\angle X(e^{j\omega})$$

Linear Phase Means Delay



- Since we can consider the effects of magnitude and phase separately, it follows that linear phase of the form $\angle H(e^{j\omega})=-\omega n_d$ implies delay of n_d samples.
- Thus an ideal lowpass filter with delay has

$$h_{\text{lp}}[n]=\frac{\sin\omega_c(n-n_d)}{\pi(n-n_d)} \Leftrightarrow H(e^{j\omega})=\begin{cases} e^{-j\omega n_d} & |\omega|<\omega_c \\ 0 & \omega_c<|\omega|<\pi \end{cases}$$

Frequency Response Functions

$$H(e^{j\omega})=|H(e^{j\omega})|e^{j\angle H(e^{j\omega})}$$

- Log-magnitude (in dB)

$$20\log_{10}|H(e^{j\omega})|$$

- Phase (in radians)

$$\angle H(e^{j\omega})=\arg\left[H(e^{j\omega})\right]$$

- Group delay (in samples)

$$\tau(\omega)=\text{grd}\left[H(e^{j\omega})\right]=-\frac{d}{d\omega}\left\{\angle H(e^{j\omega})\right\}$$

Group Delay of a Linear Phase System

- Frequency response of a linear phase system is of the form

$$H\left(e^{j\omega}\right)=\left|H\left(e^{j\omega}\right)\right|e^{-j\omega n_d}$$

- What is the group delay?

$$\tau(\omega)=-\frac{d}{d\omega}\left\{\angle H(e^{j\omega})\right\}=-\frac{d}{d\omega}\left(-\omega n_d\right)$$

$$= n_d \text{ samples}$$

Rational System Functions

- Consider a general difference equation of the form

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

- Rational system function of a causal and stable LTI system

$$H(z)=\frac{\sum_{k=0}^{M} b_k z^{-k}}{\sum_{k=0}^{N} a_k z^{-k}}=\frac{b_0}{a_0}\frac{\prod_{k=1}^{M}(1-c_k z^{-1})}{\prod_{k=1}^{N}(1-d_k z^{-1})}$$

Causal $\Rightarrow$ $|z|>\max_k |d_k|$

Stable $\Rightarrow$ $\max_k |d_k|<1$

Impulse Response

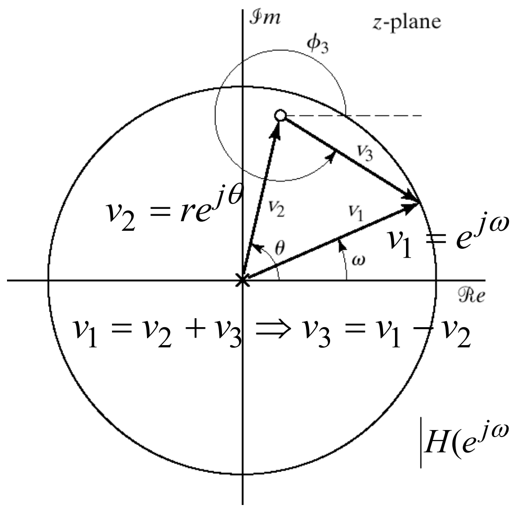
- We can make a partial fraction expansion of the rational system function:

$$H(z)=\underbrace{\left[\sum_{r=0}^{(M-N)} B_r z^{-r}\right]}_{\text{if } M\ge N}+\sum_{k=1}^{N}\frac{A_k}{1-d_k z^{-1}} \qquad |z|>\max_k |d_k|$$

- The inverse z-transform gives the impulse response

$$h[n]=\underbrace{\left[\sum_{r=0}^{(M-N)} B_r \delta[n-r]\right]}_{\text{if } M\ge N}+\sum_{k=1}^{N} A_k d_k^n u[n]$$

Pole-Zero Plot



$$H(z)=(1-re^{j\theta}z^{-1})$$

$$=\frac{z-re^{j\theta}}{z}$$

$$H(e^{j\omega})=\frac{e^{j\omega}-re^{j\theta}}{e^{j\omega}}$$

$$=\frac{v_3}{v_1}$$

$$\left|H(e^{j\omega})\right|=\frac{\left|e^{j\omega}-re^{j\theta}\right|}{\left|e^{j\omega}\right|}=\frac{|v_3|}{|v_1|}=|v_3|$$

Example

- System function:

$$H(z) = \frac{1}{\left(1 - re^{j\theta} z^{-1}\right)\left(1 - re^{-j\theta} z^{-1}\right)} = \frac{1}{1 - 2r\cos\theta z^{-1} + r^2 z^{-2}}$$

- Difference equation:

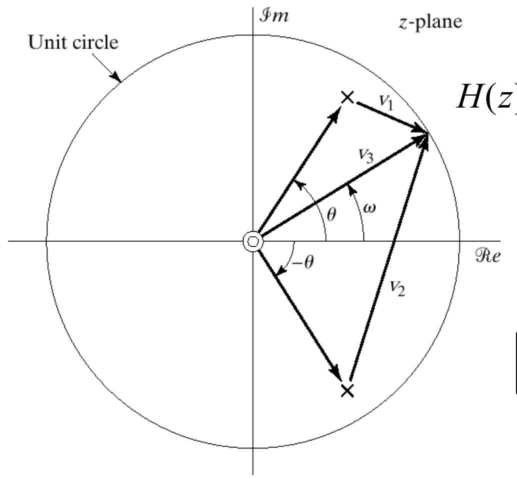
$$y[n] = 2r\cos\theta y[n-1] - r^2 y[n-2] + x[n]$$

- Impulse response:

$$H(z) = \frac{\frac{1}{1 - e^{-j2\theta}}}{\left(1 - re^{j\theta} z^{-1}\right)} + \frac{\frac{1}{1 - e^{j2\theta}}}{\left(1 - re^{-j\theta} z^{-1}\right)}$$

$$h[n] = \frac{r^n \sin[\theta(n+1)]}{\sin\theta} u[n]$$

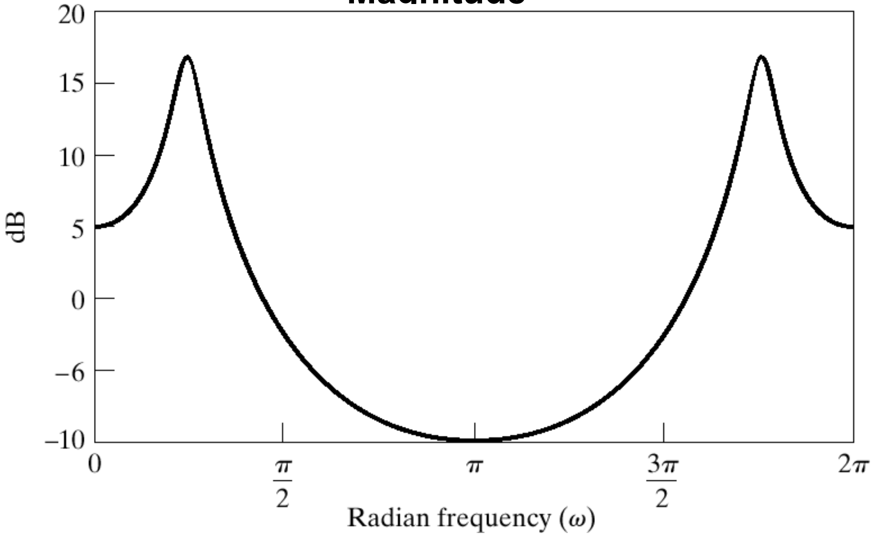
Example



$$H(z) = \frac{z^2}{\left(z - re^{j\theta}\right)\left(z - re^{-j\theta}\right)}$$

$$\left|H(e^{j\omega})\right| = \frac{|v_3|^2}{|v_1| \cdot |v_2|}$$

Magnitude



Phase and Group Delay