CT111 Introduction to Communication Systems
Lecture 11: Quantization

Yash M. Vasavada

Associate Professor, DA-IICT, Gandhinagar

16 February 2018

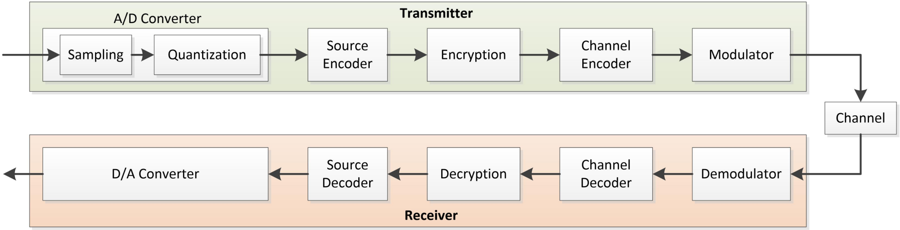
Overview of Today's Talk

Overview of Today's Talk

Overview of Today's Talk

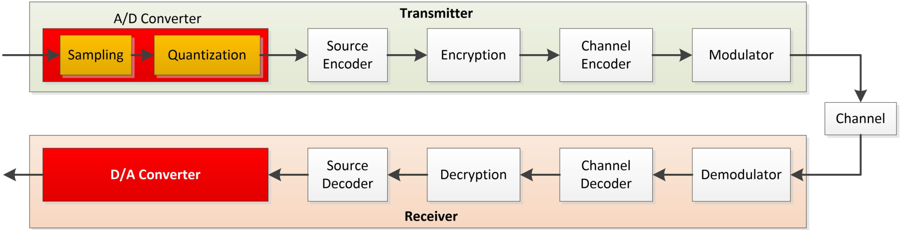
Overview of Today's Talk

Digital Communication Transceiver
Block Diagram



Digital Communication Transceiver
Block Diagram

- We will now be looking at the process of quantization of an analog information source



Discrete-Time Representations of Continuous-Time Signals

- Examples of Continuous Time (C-T) signals: video, voice, image, etc.
- Sampling theorem says that if the C-Tsignal is band limited, it can be exactly recovered by its time domain samples taken sufficiently close together
  - → Sampling analog signals makes them discrete in time
  - → For band-limited signal, the ideal sampling scheme introduces no distortion

Sampling Theorem

- Let $x(t)$ be a band-limited signal with Fourier Transform $X(f)$ which is zero if $|f| > B$
- Time domain signal $x(t)$ can be perfectly reconstructed from its uniformly spaced samples provided these samples are taken at a rate $R > 2B$. Here, $2B$ is called the Nyquist Rate
- If time-domain samples are collected at a rate less than $2B$, aliasing occurs and it is not possible to perfectly reconstruct the analog signal from its samples
- We have studied the proof of sampling theorem in an earlier lecture.

Digital Representations of C-T Signals

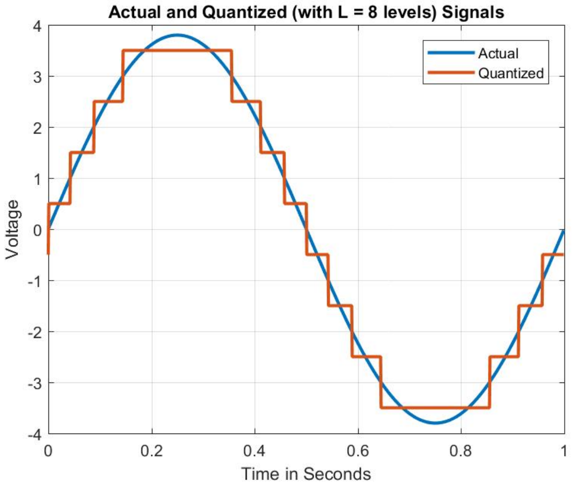
- Quantization of sampled analog signals makes the samples discrete in amplitude
  - → Quantization introduces some distortion. There is a trade off between the bandwidth requirement and distortion
- In quantization, the amplitude of an analog signal within $x_{\max}$ volts and $x_{\min}$ volts is converted to one of L levels: $\{\tilde{x}_1, \ldots, \tilde{x}_k, \ldots, \tilde{x}_L\}$

An Example
of a Quantizer

An example quantizer with $L = 8$.

k	x_{k-1}	x_k	$\tilde{x}_k$	**Output Bits**
1	-4	-3	-3.5	000
2	-3	-2	-2.5	001
3	-2	-1	-1.5	010
4	-1	0	-0.5	011
5	0	1	0.5	100
6	1	2	1.5	101
7	2	3	2.5	110
8	3	4	3.5	111

An Example
Quantization of a Sinusoidal Signal



Notations (Contd...)

Quantization

- Rate R of a quantizer is the number of *bits* required to represent a sample
- *Uniform quantizer* is the one in which $\tilde{x}_k - \tilde{x}_{k-1} = \Delta, \forall k \in 1, \ldots, L-1$
  - $\rightarrow$ $R = \log_2 L$ bits / sample for uniform quantizer
  - $\rightarrow$ In the previous example, $R = 3$ bits / sample
- In a *non-uniform quantizer*, the quantization regions are of unequal widths
  - $\rightarrow$ when would a non-uniform quantizer be preferred over a uniform quantizer?
- Uniform versus Non-uniform quantization is sometimes also called linear versus nonlinear quantization coding

Distortion

due to Quantization

- Every quantizer introduces distortion into the signal.
- Let us introduce the notion of a distortion function $d(x, \tilde{x})$
- $d(x, \tilde{x})$ between two numbers x and $\tilde{x}$ can be any function, but (intuitively) it should at least be a non-decreasing function of $|x - \tilde{x}|$
- We want to minimize the average distortion D defined as follows:

$$
\begin{aligned}
D = E\left[d\left(X, \tilde{X}\right)\right] &= \int_{-\infty}^{\infty} d(x, \tilde{x}) p(x)\, dx \\
&= \sum_{k=1}^{L} \int_{x_{k-1}}^{x_k} d(x, \tilde{x}) p(x)\, dx
\end{aligned}
$$

Mean Squared Error or MSE

in the Quantization Process

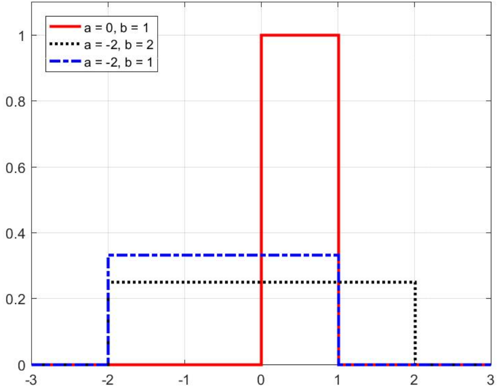
- A common choice for the distortion function: $d(x, \tilde{x}) = (x - \tilde{x})^2$
- $D = MSE = E\left[\left(X - \tilde{X}\right)^2\right] = \sum_{k=1}^{L} \int_{x_{k-1}}^{x_k} (x - \tilde{x})^2 \, p(x) \, dx$
- An interpretation of MSE: $\tilde{X} = f_Q(X) = x + \tilde{n}$, where $\tilde{n} = x - \tilde{x}$ is the noise introduced by the quantization operation. In this case, MSE $= E\left[\tilde{n}^2\right]$ may be interpreted as the power of the quantization noise
- $\left(\frac{P_S}{P_N}\right) = \frac{\text{Signal Power}}{\text{Noise Power}} = \frac{E\left[X^2\right]}{D}$
- When quantization noise $\tilde{n}$ is uniformly distributed, quantization-induced SNR is given approximately as $6 \times R$ dB.

MSE
An Example

- Let $L = 8$, $\{\tilde{x}_1 = -3.5, \tilde{x}_2 = -2.5, \ldots, \tilde{x}_L = 3.5\}$
- Let $p(x) = \begin{cases} \frac{1}{8}, & -4 \leq x < 4 \\ 0, & \text{otherwise} \end{cases}$
- $E\left[X^2\right] =?$
- $\text{MSE} = D =?$

Recall:
Uniform PDF

- $p(x) = \begin{cases} \dfrac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{else} \end{cases}$



Recall:
Uniform PDF

- Mean: $m_x = \int_a^b x\, p(x)\, dx = \frac{1}{b-a} \int_a^b x\, dx = \frac{a+b}{2}$
- Variance: $\sigma_x^2 = \int_a^b (x - m_x)^2\, p(x)\, dx = \frac{(b-a)^2}{12}$
- Probability:

$$P(a_1 \le x < b_1) = \int_{a_1}^{b_1} p(x)\, dx = \frac{b_1 - a_1}{b - a}, \quad a < a_1, b_1 < b$$

MSE
An Example

- $E\left[X^2\right] = \dfrac{(b-a)^2}{12} = \dfrac{64}{12}$, since $b = -a = 4$
- Distortion is evaluated as follows:

$$
\begin{aligned}
D &= \sum_{k=1}^{L} \int_{x_{k-1}}^{x_k} (x - \tilde{x})^2 \, p(x) \, dx \\
&= \sum_{k=1}^{8} \int_{-5+k}^{-4+k} (x - (-4.5 + k))^2 \, (1/8) \, dx \\
&= \frac{1}{12}
\end{aligned}
$$

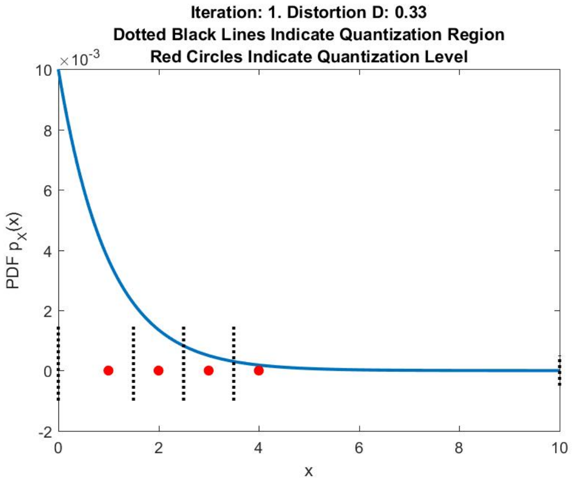
- SNR: $\left(\dfrac{P_S}{P_N}\right) = \dfrac{64/12}{1/12} = 64$ or 18 dB

Quantizer Design

- Consider the distortion measure: $D = \int_{-\infty}^{\infty} (x - \tilde{x})^2 \, p(x) dx$
- Objective: minimize D given a certain rate R
- A general strategy:
  - $\rightarrow$ Make $(x - \tilde{x})^2$ where $p(x)$ is large
  - $\rightarrow$ Tolerate larger $(x - \tilde{x})^2$ where $p(x)$ is small
- i.e., concentrate quantization levels in the regions of large PDF.

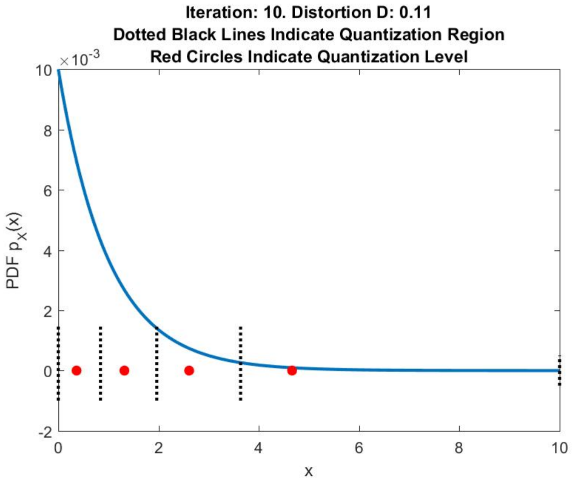
Uniform versus Nonuniform Quantization

Exponential PDF. Uniform. Not Good.



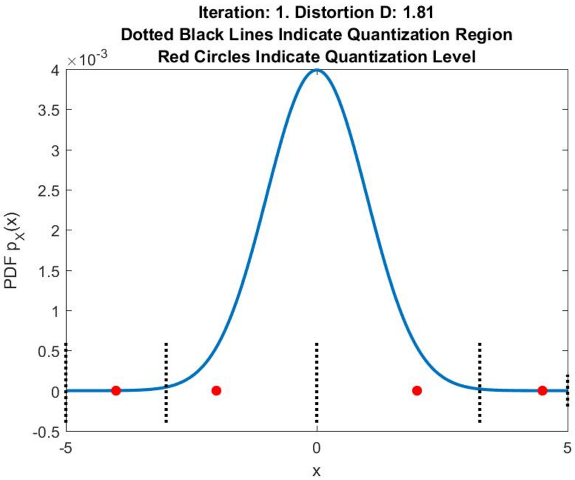
Uniform versus Nonuniform Quantization

Exponential PDF. Nonuniform. Good.



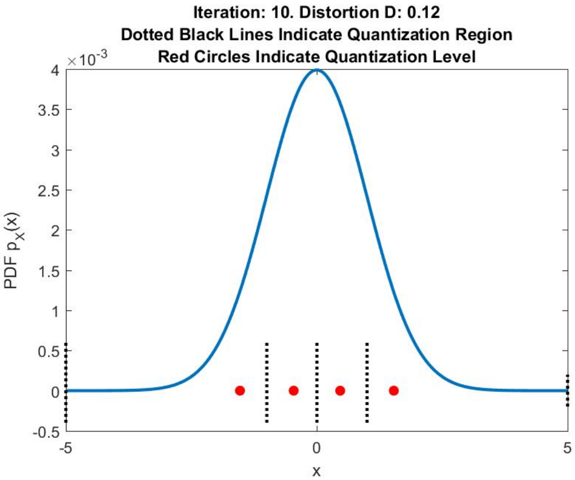
Uniform versus Nonuniform Quantization

Gaussian PDF. Uniform. Not Good.



Uniform versus Nonuniform Quantization

Gaussian PDF. Nonuniform. Good.



Speech Encoding

- All speech coding techniques require quantization.
- Many of them employ additional properties of speech

  [Temporal Waveform Coding:] attempts to represent time domain samples of speech

  [Spectral Waveform Coding:] attempts to represent spectral characteristics of speech

  [Model based Coding:] replicates a model of the process by which the Human auditory tract generates the speech

Speech Encoding

- All speech coding techniques require quantization.
- Many of them employ additional properties of speech

  [Temporal Waveform Coding:] attempts to represent time domain samples of speech

  [Spectral Waveform Coding:] attempts to represent spectral characteristics of speech

  [Model based Coding:] replicates a model of the process by which the Human auditory tract generates the speech

Speech Encoding

- All speech coding techniques require quantization.
- Many of them employ additional properties of speech

  [Temporal Waveform Coding:] attempts to represent time domain samples of speech

  [Spectral Waveform Coding:] attempts to represent spectral characteristics of speech

  [Model based Coding:] replicates a model of the process by which the Human auditory tract generates the speech

Temporal Waveform Coding

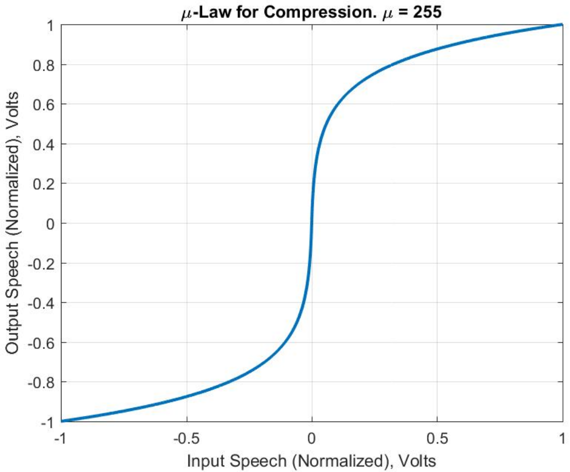
1. *Pulse Code Modulation (PCM):* directly quantizes the samples of speech
   - *Companding:*
     - → In human speech, probability of low amplitudes is greater than that of the high amplitudes.
     - → Need to use non-uniform quantization.
     - → An alternate is nonlinear compander followed by the uniform quantizer
     - → Nonlinearity for μ-law companding: $y = \frac{\log(1 + \mu|x|)}{\log(1 + \mu)}$
   - Example: 64 kbps PCM for telephone lines:
     - → Speech signal (with a bandwidth of 4 kHz) is sampled at 8 ksps
     - → Encodes each sample with 8 bit uniform quantizer
     - → Uses companding with $\mu = 255$

Temporal Waveform Coding

1. *Pulse Code Modulation (PCM):* directly quantizes the samples of speech
   - *Companding:*
     - In human speech, probability of low amplitudes is greater than that of the high amplitudes.
     - Need to use non-uniform quantization.
     - An alternate is nonlinear compander followed by the uniform quantizer
     - Nonlinearity for μ-law companding: $y = \dfrac{\log(1+\mu|x|)}{\log(1+\mu)}$
   - Example: 64 kbps PCM for telephone lines:
     - Speech signal (with a bandwidth of 4 kHz) is sampled at 8 ksps
     - Encodes each sample with 8 bit uniform quantizer
     - Uses companding with $\mu = 255$

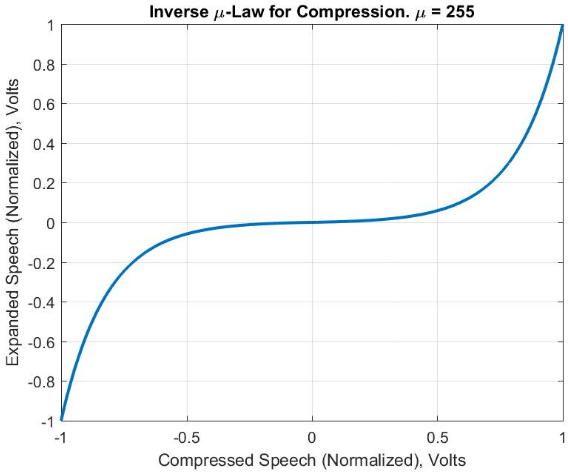
Speech Encoding

μ Law used at Speech Encoder



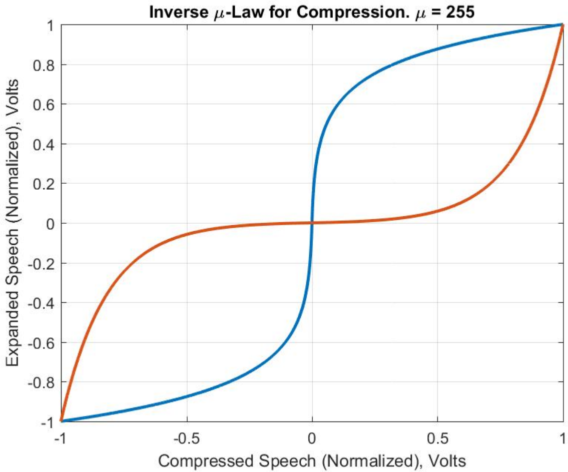
Speech Encoding

Inverse μ Law used at Speech Decoder



Speech Encoding

μ Law and Inverse μ Law



Temporal Waveform Coding

2. *Differential PCM (DPCM):* speech samples are strongly correlated from one sampling instant to the next. DPCM quantizes the difference between the current sample and the predicted value of this sample using the past samples.
3. *Adaptive DPCM:*
   - Perceptual quality is determined by the relative accuracy of the quantization to the size of speech
   - Adaptive quantizers vary the step size between the quantization levels depending on whether the speech is loud or soft
   - Adaptive DPCM at 32 kbps is a common and easily implemented, and used in Digital European Cordless Telephone or DECT standard

Temporal Waveform Coding

2. *Differential PCM (DPCM):* speech samples are strongly correlated from one sampling instant to the next. DPCM quantizes the difference between the current sample and the predicted value of this sample using the past samples.
3. *Adaptive DPCM:*
   - Perceptual quality is determined by the relative accuracy of the quantization to the size of speech
   - Adaptive quantizers vary the step size between the quantization levels depending on whether the speech is loud or soft
   - Adaptive DPCM at 32 kbps is a common and easily implemented, and used in Digital European Cordless Telephone or DECT standard

Temporal Waveform Coding

2. *Differential PCM (DPCM):* speech samples are strongly correlated from one sampling instant to the next. DPCM quantizes the difference between the current sample and the predicted value of this sample using the past samples.
3. *Adaptive DPCM:*
   - Perceptual quality is determined by the relative accuracy of the quantization to the size of speech
   - Adaptive quantizers vary the step size between the quantization levels depending on whether the speech is loud or soft
   - Adaptive DPCM at 32 kbps is a common and easily implemented, and used in Digital European Cordless Telephone or DECT standard

Temporal Waveform Coding

4. *Delta Modulation:*
   - An extreme case of PCM in which signal is oversampled and $R = 1$ bit/sample, i.e., needs a very simple one bit quantizer
   - Can suffer from granular noise (when the incoming signal does not change much over time) or slope overload noise (if the signal changes too fast)
   - Adaptive Delta Modulation can produce reasonable quality speech at 16 kbps

Spectral Waveform Coding

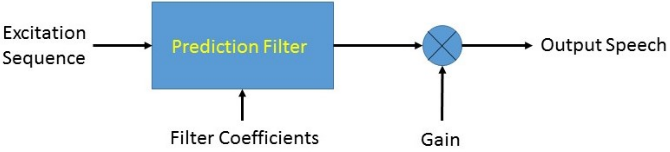
1. Subband Coding
   - Human perception of speech quality depends on the frequency band. Higher MSE may be tolerated at very low and very high frequencies
   - Subband coders filter the speech signal into multiple bands using Quadrature Mirror Filters (QMFs)
   - Filtered signals are quantied using standard PCM, with a different value of R for different bands
2. Adaptive Transform Coding
   - Correlated time domain signals are transformed into frequency domain samples using Fourier Transformation
   - These frequency domain signals are represented using their perceptual importance

Spectral Waveform Coding

1. Subband Coding
   - Human perception of speech quality depends on the frequency band. Higher MSE may be tolerated at very low and very high frequencies
   - Subband coders filter the speech signal into multiple bands using Quadrature Mirror Filters (QMFs)
   - Filtered signals are quantied using standard PCM, with a different value of R for different bands
2. Adaptive Transform Coding
   - Correlated time domain signals are transformed into frequency domain samples using Fourier Transformation
   - These frequency domain signals are represented using their perceptual importance

Linear Predictive Speech Coding

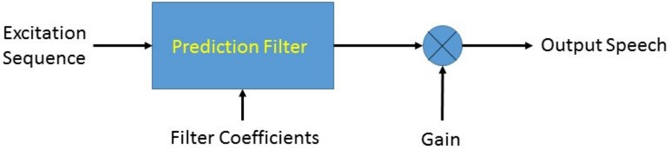
- Human speech is modeled as noise (air from the lungs) exciting a linear filter (throat, vocal cords and the mouth)
- Following are quantized and transmitted to the receiver: excitation sequence, filter coefficients, gain



- VSELP (Vector Sum Excited Linear Prediction)
  - → 20 ms frames with 159 bits per frame. Data rate of approx 8 kbps.
  - → Two stage VQ is used to quantize the excitation sequence
  - → Bits (such as those representing the filter gain) are more critical for perceptual quality. These are protected by error correction coding

Linear Predictive Speech Coding

- Human speech is modeled as noise (air from the lungs) exciting a linear filter (throat, vocal cords and the mouth)
- Following are quantized and transmitted to the receiver: excitation sequence, filter coefficients, gain



- VSELP (Vector Sum Excited Linear Prediction)
  - → 20 ms frames with 159 bits per frame. Data rate of approx 8 kbps.
  - → Two stage VQ is used to quantize the excitation sequence
  - → Bits (such as those representing the filter gain) are more critical for perceptual quality. These are protected by error correction coding

Relative Performance

Table: Vocoder Comparison

Type	Rate (kbps)	Complexity (MIPS)	Delay (ms)	Quality
PCM	64	0.01	0	High
ADPCM	32	0.1	0	High
Subband	16	1	25	High
VSELP	8	$\sim$100	35	Fair