CS 414 – Multimedia Systems Design

Lecture 3 – Digital Audio Representation

Klara Nahrstedt

Spring 2008

Administrative

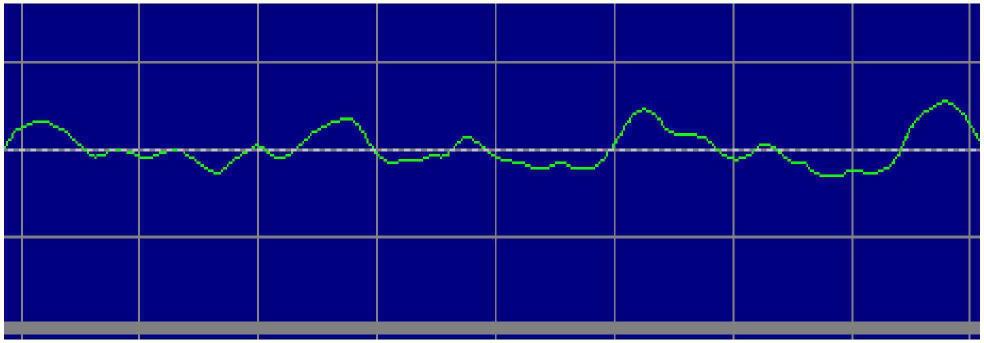
- Form Groups for MPs
  - Deadline January 20 to email TA

Key Questions

- How can a continuous wave form be converted into discrete samples?
- How can discrete samples be converted back into a continuous form?

Characteristics of Sound

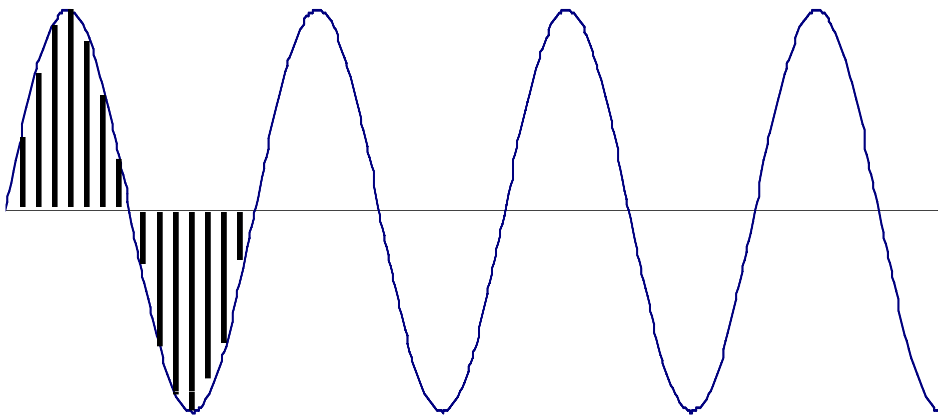
- Amplitude
- Wavelength (*w*)
- Frequency (λ)
- Timbre



- Hearing: [20Hz – 20KHz]
- Speech: [200Hz – 8KHz]

Digital Representation of Audio

- Must convert wave form to digital
  - sample
  - quantize
  - compress

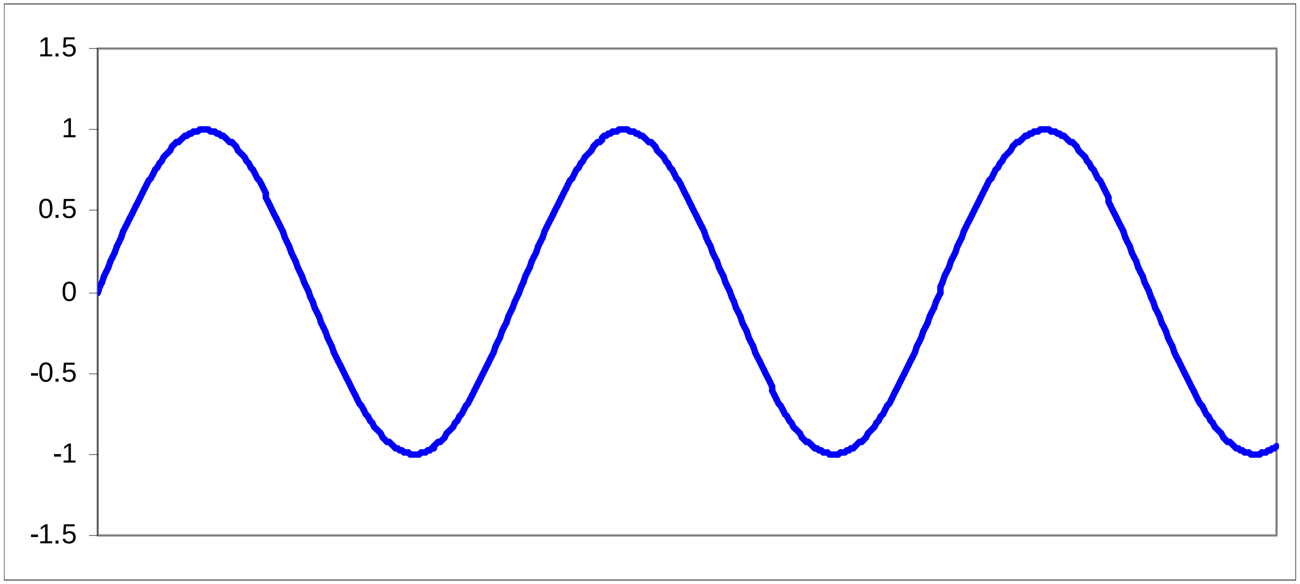


Sampling (in time)

- Measure amplitude at regular intervals
- How many times should we sample?

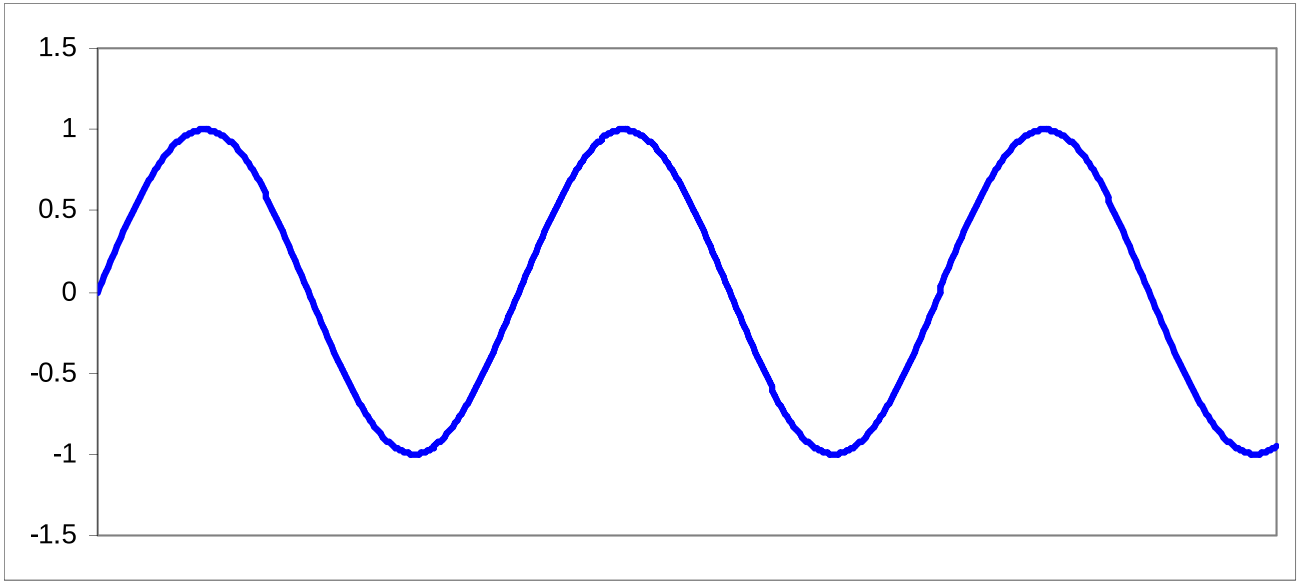
Example

- Suppose we have a sound wave with a frequency of 1 cycle per second



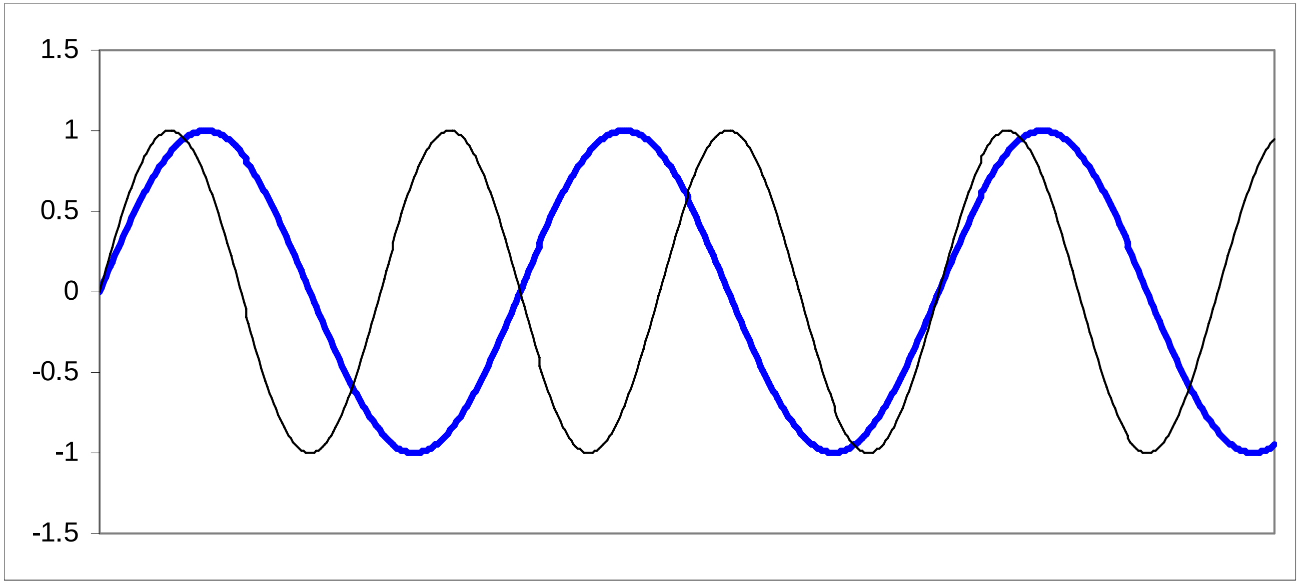
Example

- If we sample at one cycle per second, where would the sample points fall?



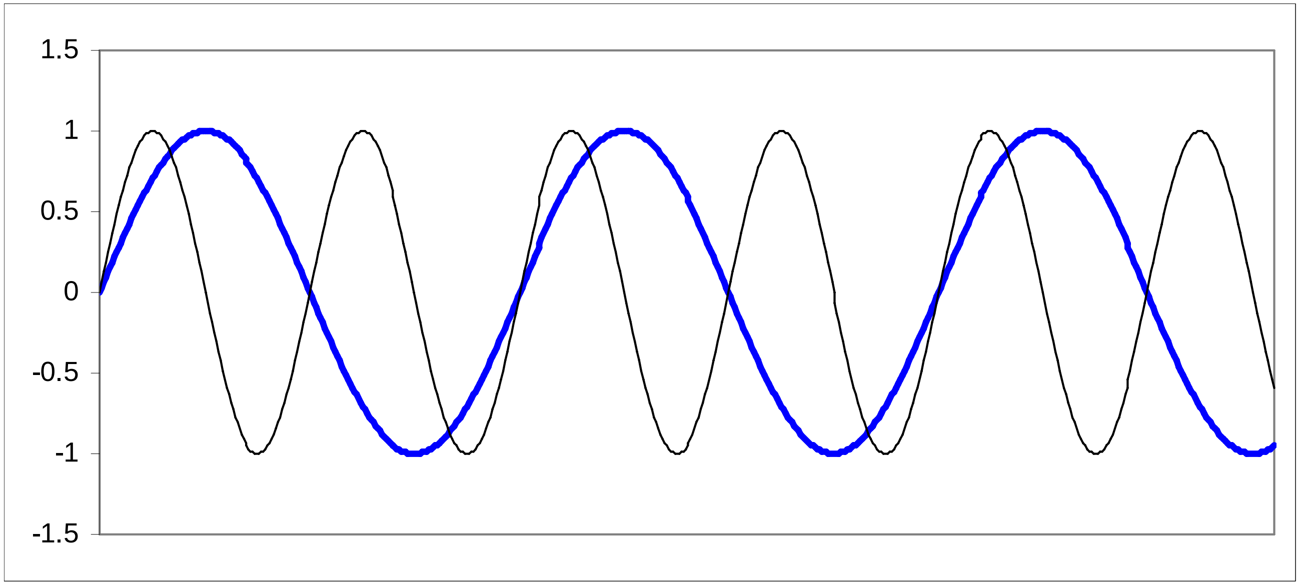
Example

- If we sample at 1.5 cycles per second, where will the sample points fall?



Sampling - Example

- If we sample at two cycles per second, where do the sample points fall?



Nyquist Theorem

For lossless digitization, the sampling rate should be at least twice the maximum frequency response.

- In mathematical terms:

$$f_s > 2*f_m$$

- where f_s is sampling frequency and f_m is the maximum frequency in the signal

Nyquist Limit

- max data rate = 2 H $\log_2 V$ *bits/second,* where
  H = bandwidth (in Hz)
  V = discrete levels (bits per signal change)

- Shows the maximum number of bits that can be sent per second on a *noiseless* channel with a bandwidth of H, if V bits are sent per signal
  - Example: what is the maximum data rate for a 3kHz channel that transmits data using 2 levels (binary) ?
  - (2x3,000xln2=6,000bits/second)

Limited Sampling

- But what if one cannot sample fast enough?

Limited Sampling

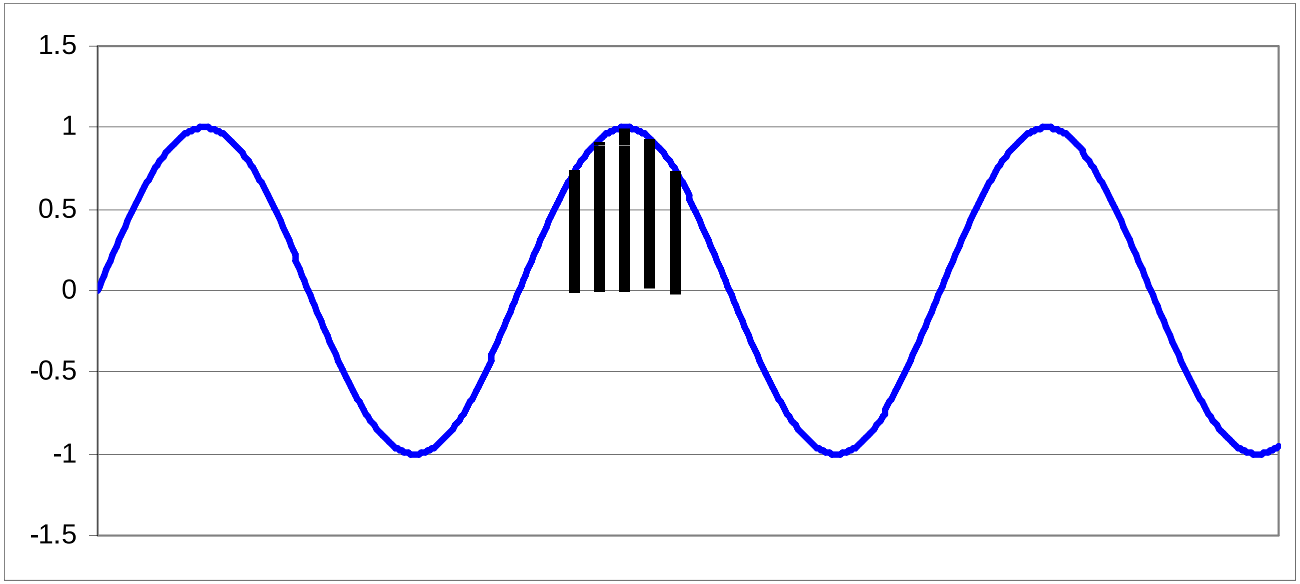
- Reduce signal frequency to half of maximum sampling frequency
  - low-pass filter removes higher-frequencies
  - e.g., if max sampling frequency is 22kHz, must low-pass filter a signal down to 11kHz

Sampling Ranges

- Auditory range 20Hz to 22.05 kHz
  - must sample up to to 44.1kHz
  - common examples are 8.000 kHz, 11.025 kHz, 16.000 kHz, 22.05 kHz, and 44.1 KHz
- Speech frequency [200 Hz, 8 kHz]
  - sample up to 16 kHz
  - but typically 4 kHz to 11 kHz is used

Sampling Rates	Used As...
8000	Telephony Standard, Popular in UNIX Workstations
11000	Quarter of CD rate, Popular on Macintosh
16000	G.722 Standard (Federal Standard)
18900	CD-ROM XA Rate
22000	Half CD rate, Macintosh rate
32000	Japanese HDTV, British TV audio, Long play DAT
37800	CD XA Standard
44056	Professional audio industry
44100	CD Rate
48000	DAT Rate

Quantization



Quantization

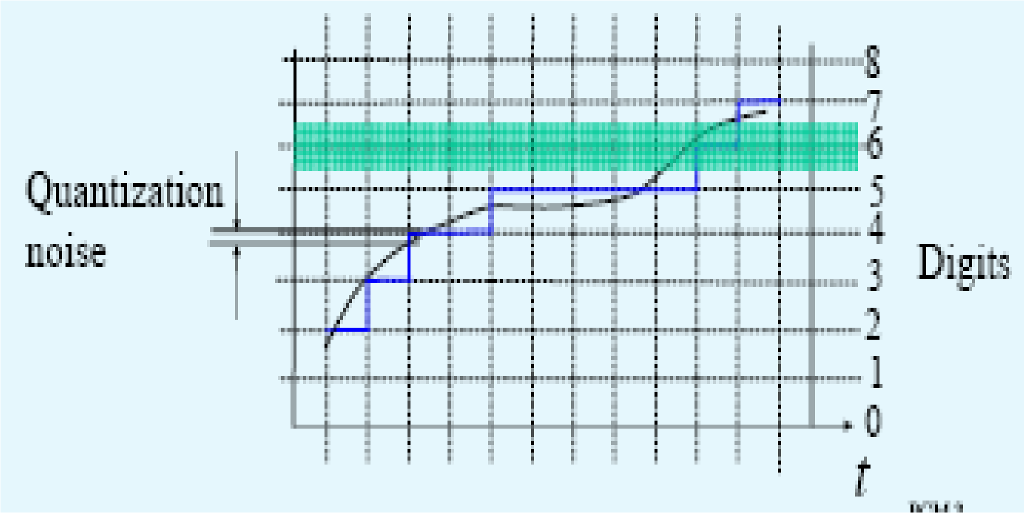
- Typically use
  - 8 bits = 256 levels
  - 16 bits = 65,536 levels
- How should the levels be distributed?
  - Linearly? (PCM)
  - Perceptually? (u-Law)
  - Differential? (DPCM)
  - Adaptively? (ADPCM)

Pulse Code Modulation

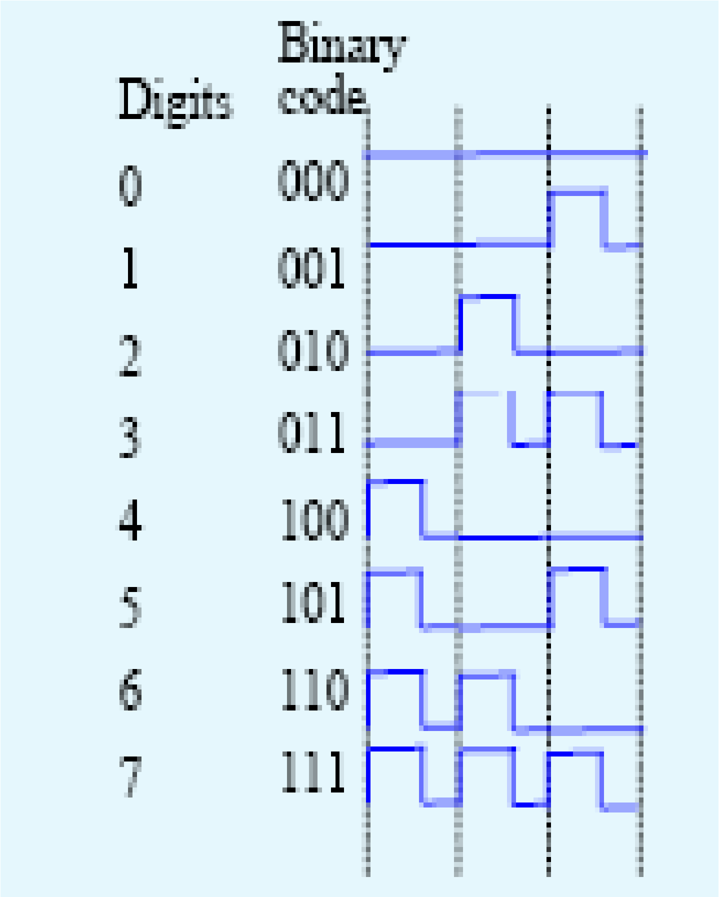
- Pulse modulation
  - Use discrete time samples of analog signals
  - Transmission is composed of analog information sent at different times
  - Variation of pulse amplitude or pulse timing allowed to vary continuously over all values
- PCM
  - Analog signal is quantized into a number of discrete levels

PCM Quantization and Digitization

Quantization

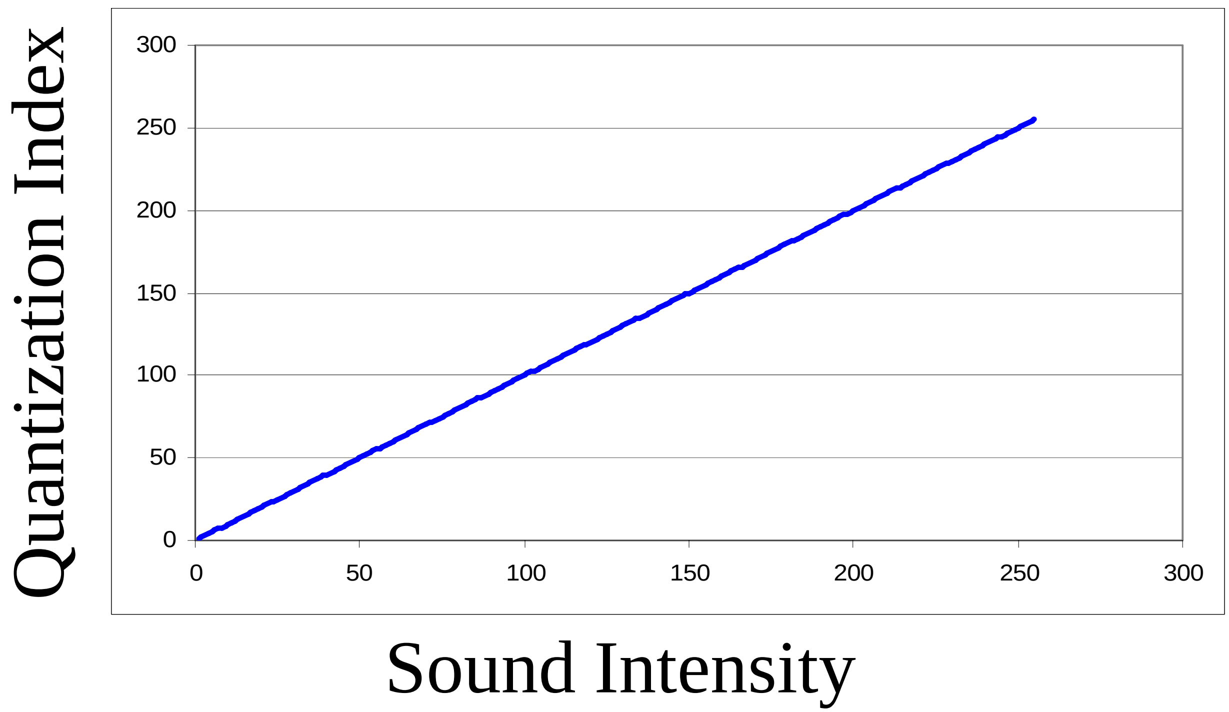


Digitization

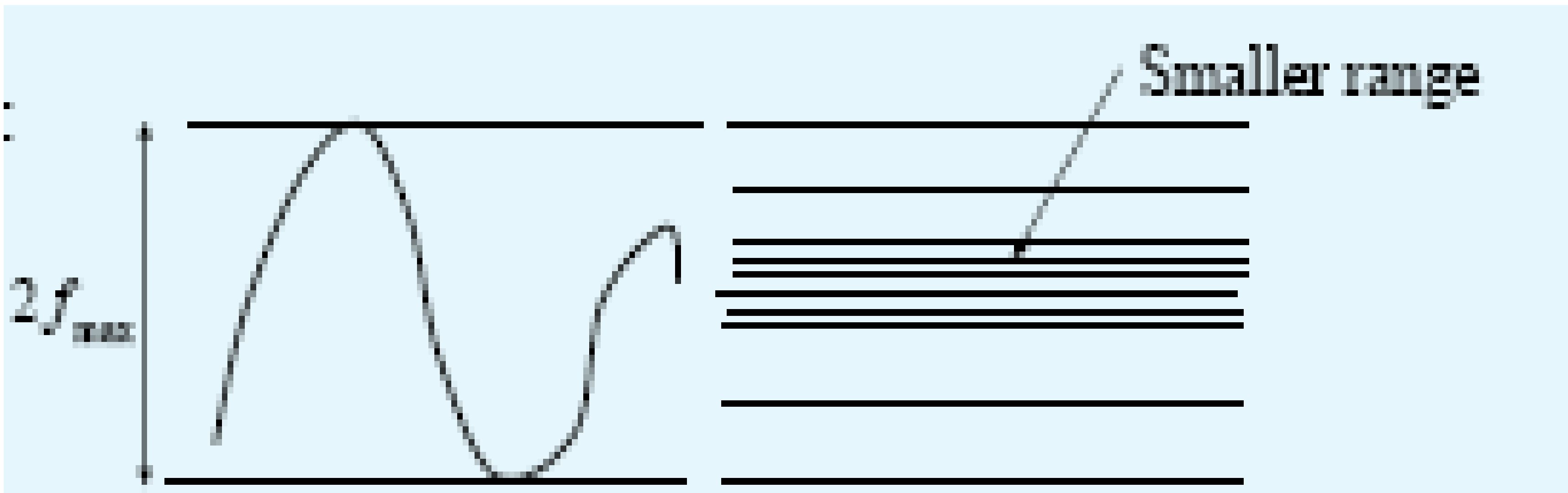


Linear Quantization (PCM)

- Divide amplitude spectrum into N units (for $\log_2 N$ bit quantization)

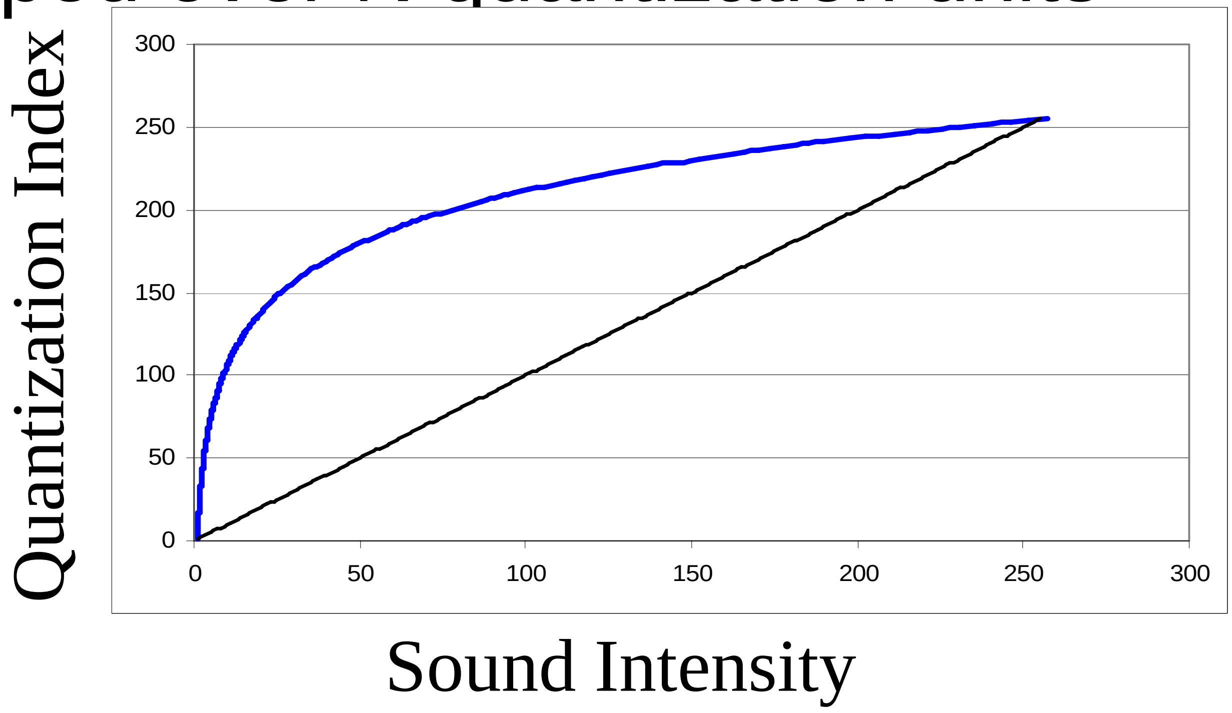


Non-uniform Quantization



Perceptual Quantization (u-Law)

- Want intensity values logarithmically mapped over N quantization units



Differential Pulse Code Modulation (DPCM)

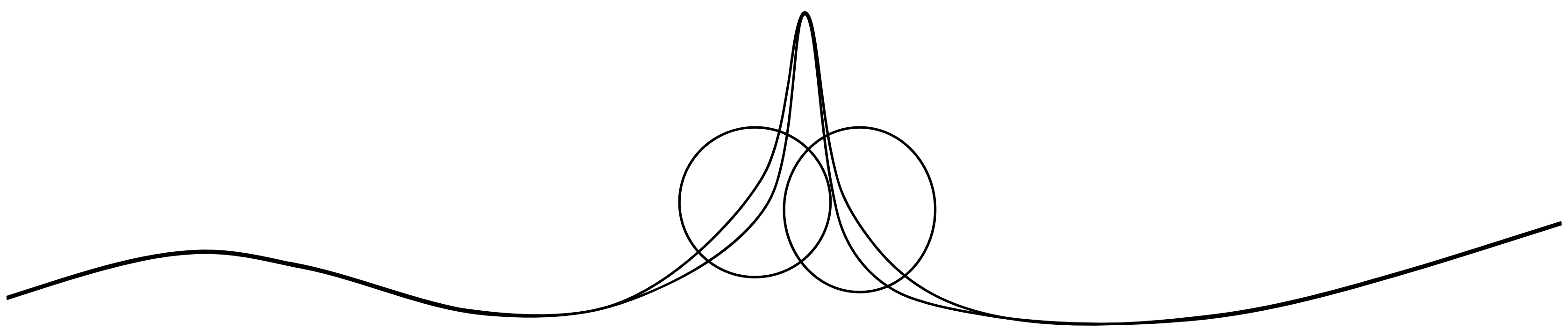
- What if we look at sample differences, not the samples themselves?
  - $d_t = x_t - x_{t-1}$
  - Differences tend to be smaller
    - Use 4 bits instead of 12, maybe?

Differential Quantization (DPCM)

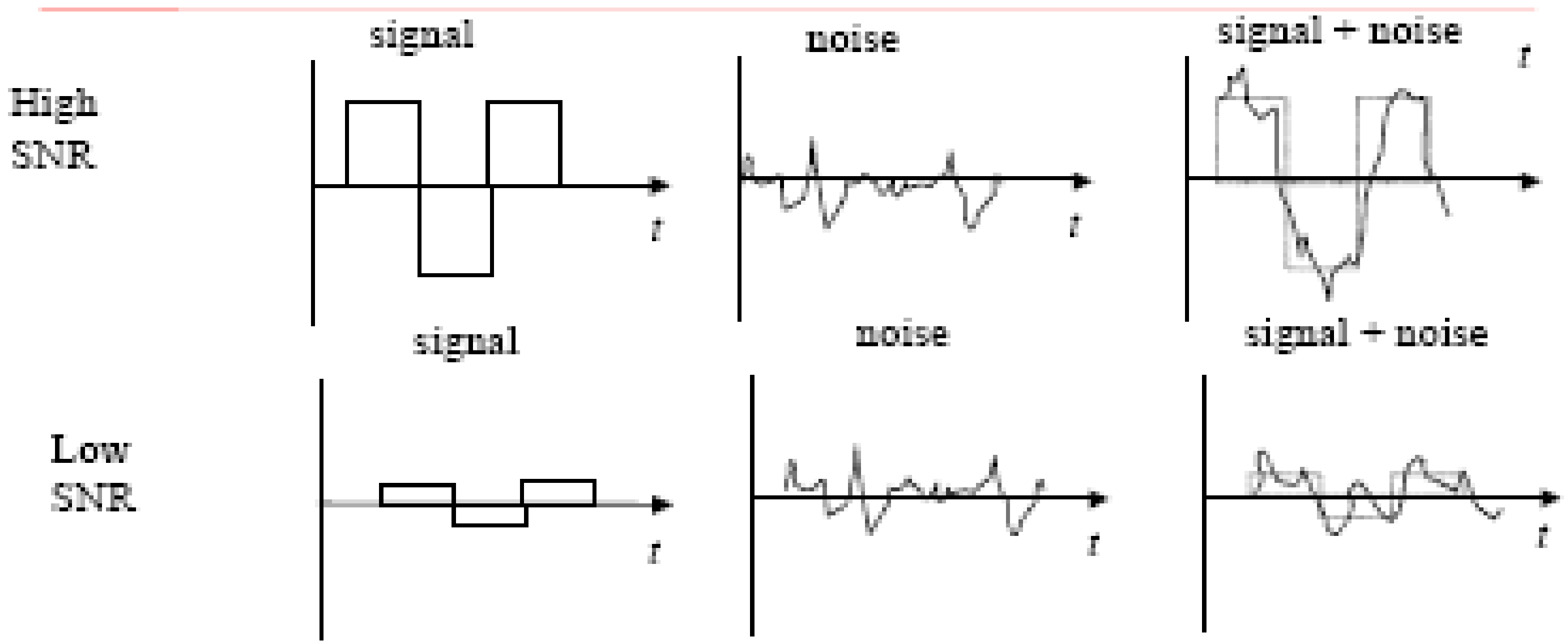
- Changes between adjacent samples small
- Send value, then relative changes
  - value uses full bits, changes use fewer bits
  - E.g., 220, 218, 221, 219, 220, 221, 222, 218,.. (all values between 218 and 222)
  - Difference sequence sent: 220, -2,+1, -1, 0, +1, +2, -2, ..
  - Result: originally for encoding sequence 0..255 numbers need 8 bits;
  - Difference coding: need only 3 bits

Adaptive Differential Pulse Code Modulation (ADPCM)

- Adaptive similar to DPCM, but adjusts the width of the quantization steps
- Encode difference in 4 bits, but vary the mapping of bits to difference dynamically
  - If rapid change, use large differences
  - If slow change, use small differences



Signal-to-Noise Ratio



$$\text{SNR} = \frac{\text{Average Signal Power}}{\text{Average Noise Power}}$$

$$\text{SNR (dB)} = 10 \log_{10} \text{SNR}$$

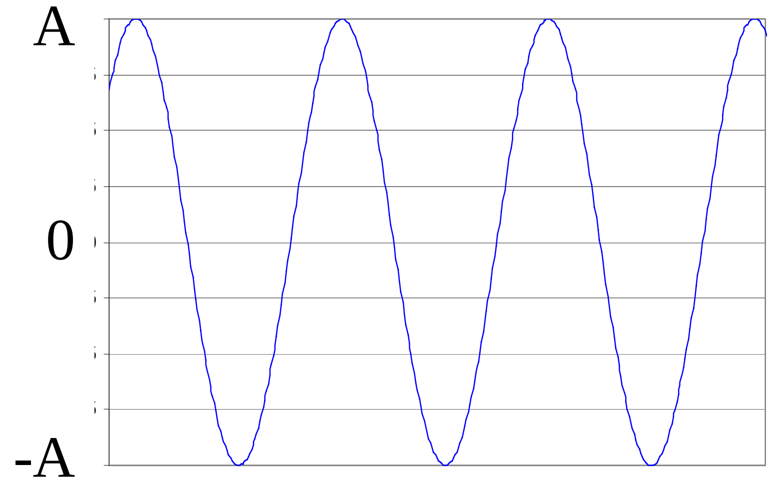
Signal To Noise Ratio

- Measures strength of signal to noise

SNR (in DB)= $10\log_{10}\left(\frac{Signal\ energy}{Noise\ energy}\right)$

- Given sound form with amplitude in [-A, A]

- Signal energy = $\frac{A^2}{2}$

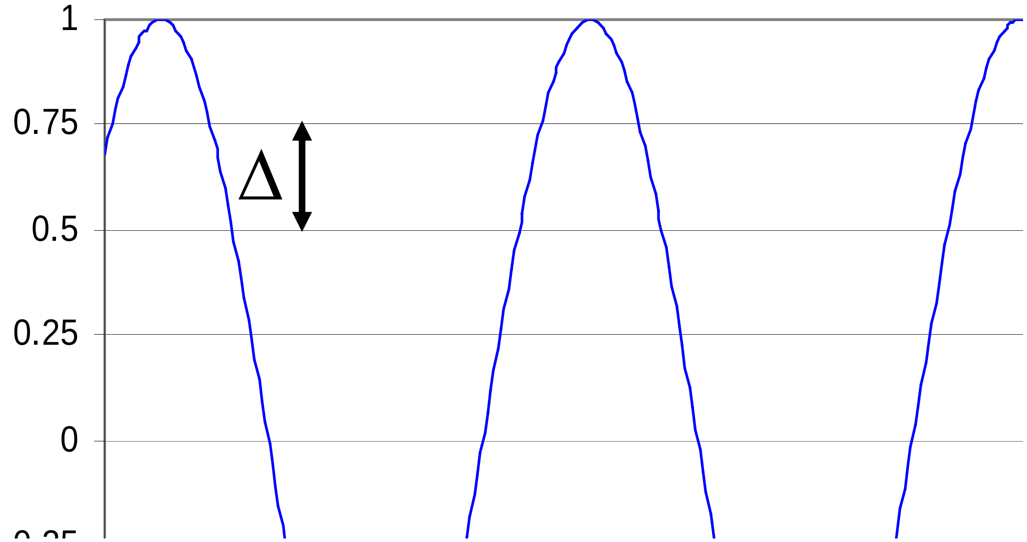


Quantization Error

- Difference between actual and sampled value
  - amplitude between [-A, A]
  - quantization levels = N

$$\Delta = \frac{2A}{N}$$

- e.g., if A = 1, N = 8, Δ = 1/4



Quantization Error

- For uniform quantizer
  the quantization error (q) bounded by $\left|\frac{\Delta}{2}\right|$
- Probability density function of quantization error *q*
- *pdf (q)* = $\left\{ \begin{matrix} \frac{1}{\Delta} if & \frac{-\Delta}{2} < q < \frac{\Delta}{2} \\ 0 & otherwise \end{matrix} \right\}$
- Noise energy = $\int_{\frac{-\Delta}{2}}^{\frac{\Delta}{2}} q^2 p(q)\, dq \quad = \frac{1}{\Delta} \int_{\frac{-\Delta}{2}}^{\frac{\Delta}{2}} q^2\, dq \quad = \frac{\Delta^2}{12}$

Compute Signal to Noise Ratio

- Signal energy = $\dfrac{A^2}{2}$; Noise energy = $\dfrac{\Delta^2}{12}$; $\Delta = \dfrac{2A}{N}$
- Noise energy = $\dfrac{A^2}{3 \cdot N^2}$
- Signal to noise = $10 \log \dfrac{3N^2}{2}$
- Every bit increases SNR by ~ 6 decibels

Example

- Consider a full load sinusoidal modulating signal of amplitude *A*, which utilizes all the representation levels provided
- The average signal power is $P = A^2/2$
- The total range of quantizer is *2A* because modulating signal swings between –*A* and *A*. Therefore, if it is $N=16$ (4-bit quantizer), $\Delta = 2A/2^4 = A/8$
- The quantization noise is $\Delta^2/12 = A^2/768$
- The S/N ratio is $(A^2/2)/(A^2/768) = 384$; *SNR (in dB) 25.8 dB*

Data Rates

- Data rate = sample rate * quantization * channel

- Compare rates for CD vs. mono audio
  - 8000 samples/second * 8 bits/sample * 1 channel = 8 kBytes / second
  - 44,100 samples/second * 16 bits/sample * 2 channel = 176 kBytes / second ~= 10MB / minute

Comparison and Sampling/Coding Techniques

mode	bits	sample per sec	noise level	freq. (Hz)	mono storage (bytes/ sec)	stereo storage (bytes/ sec)
PCM	16	44,100	v. low	20 - 20K	88,200	176,400
A-Law	8	44,100	low	20 - 20K	44,100	N/A
ADPCM	16	high	low	20 - 18K	22,050	N/A
PCM	8	22,050	low	20 - 9.2K	22,050	44,100
A-Law	8	22,050	low	20 - 9.2K	22,050	44,100
ADPCM	16	music	low	20 - 7K	11,025	22,050
PCM	8	11,025	high	20 - 4.5K	11,025	22,050
A-Law	8	11,025	low	20 - 4.5K	11,025	22,050
ADPCM	16	speech	low	20 - 3K	5,500	N/A
A-Law	16	8,000	low	20 - 3K	8,000	16,000

Summary

Audio Device	Frequency Response (Bandwidth)	Signal-to-Noise Ratio	Total Harmonic Distortion
CD	20 Hz - 20,000 Hz	98dB	0.005%
Cassette tape	20 Hz - 17,000 Hz	75dB	0.01%
FM Radio	20 Hz - 15,000 Hz	75dB	0.01%
AM Radio	50 Hz - 5,000 Hz	60dB	0.1%
Telephone	300 Hz - 3400 Hz	42dB	Poor