## Lecture 3

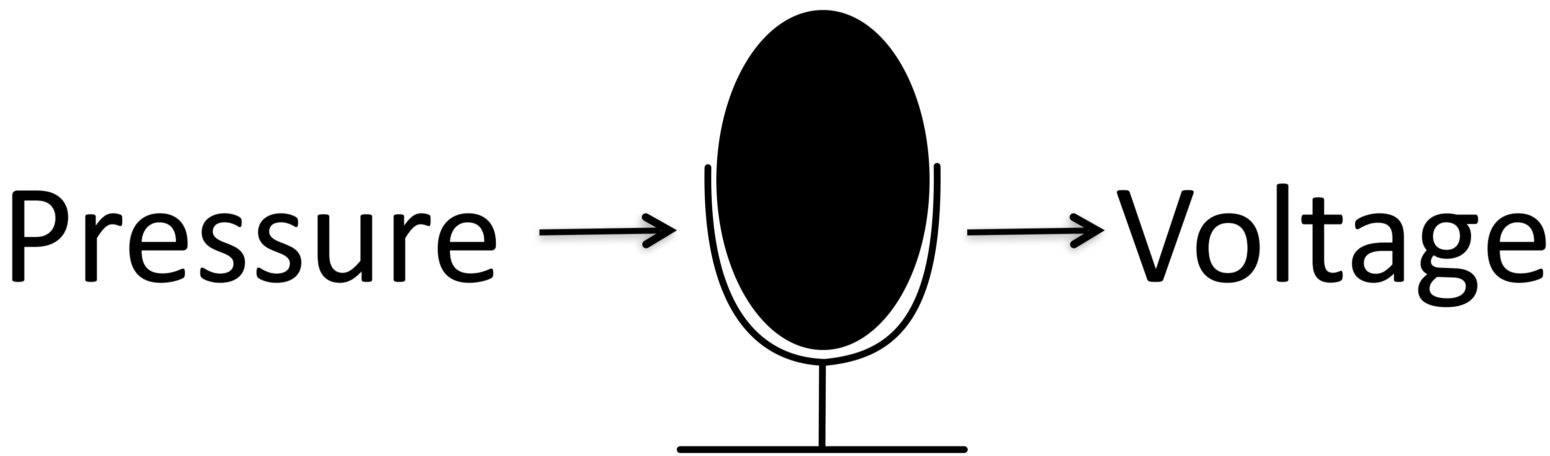
# Audio Capture and Representation

Ref: Fundamentals of Multimedia

# Audio is a wave phenomenon

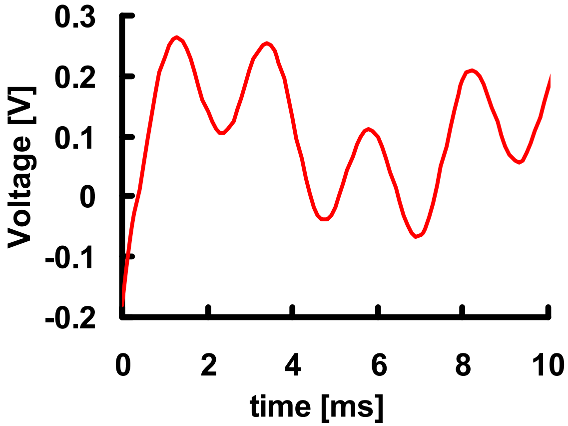
- A speaker diaphragm moves back and fourth and generates longitudinal pressure waves
- Ear perceives that as sound
- No air, no sound

# Pressure to Voltage Conversion

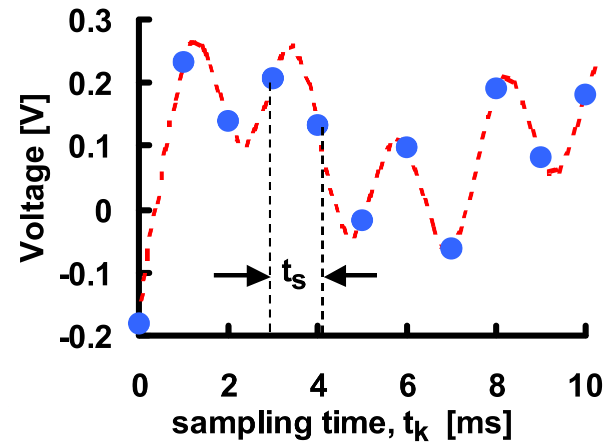


# Analogue

# Digital



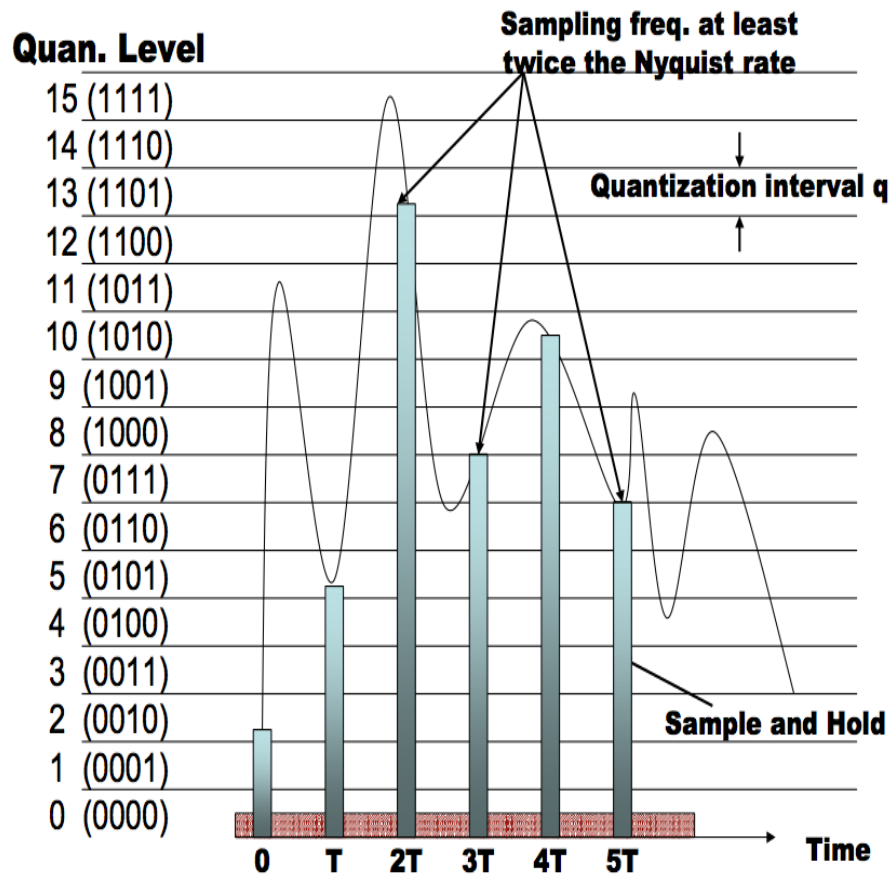

→ **ADC** →




-When you go from analog to digital, you need to discretise both the time axis and the amplitude axis!

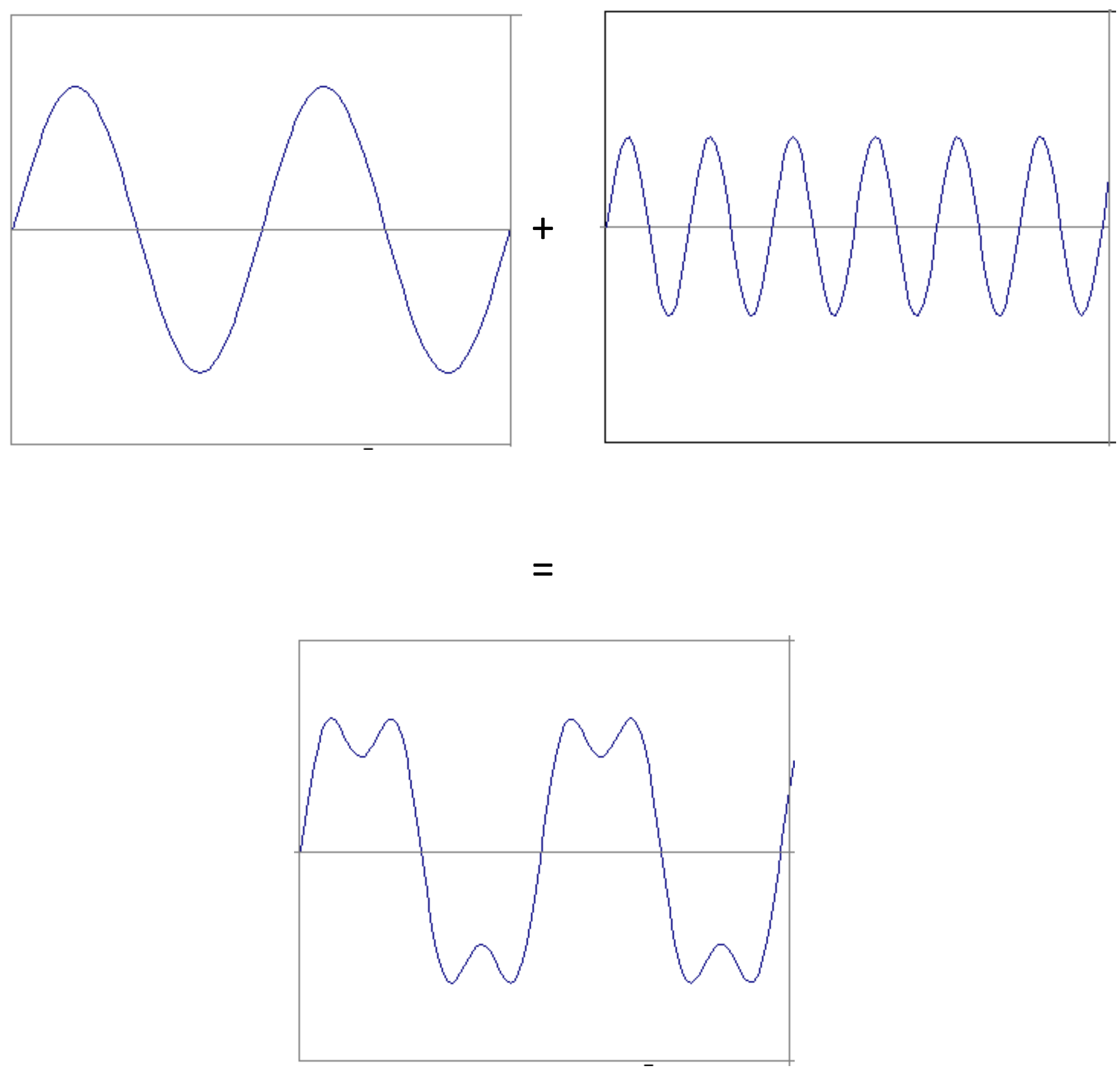
# Discretization

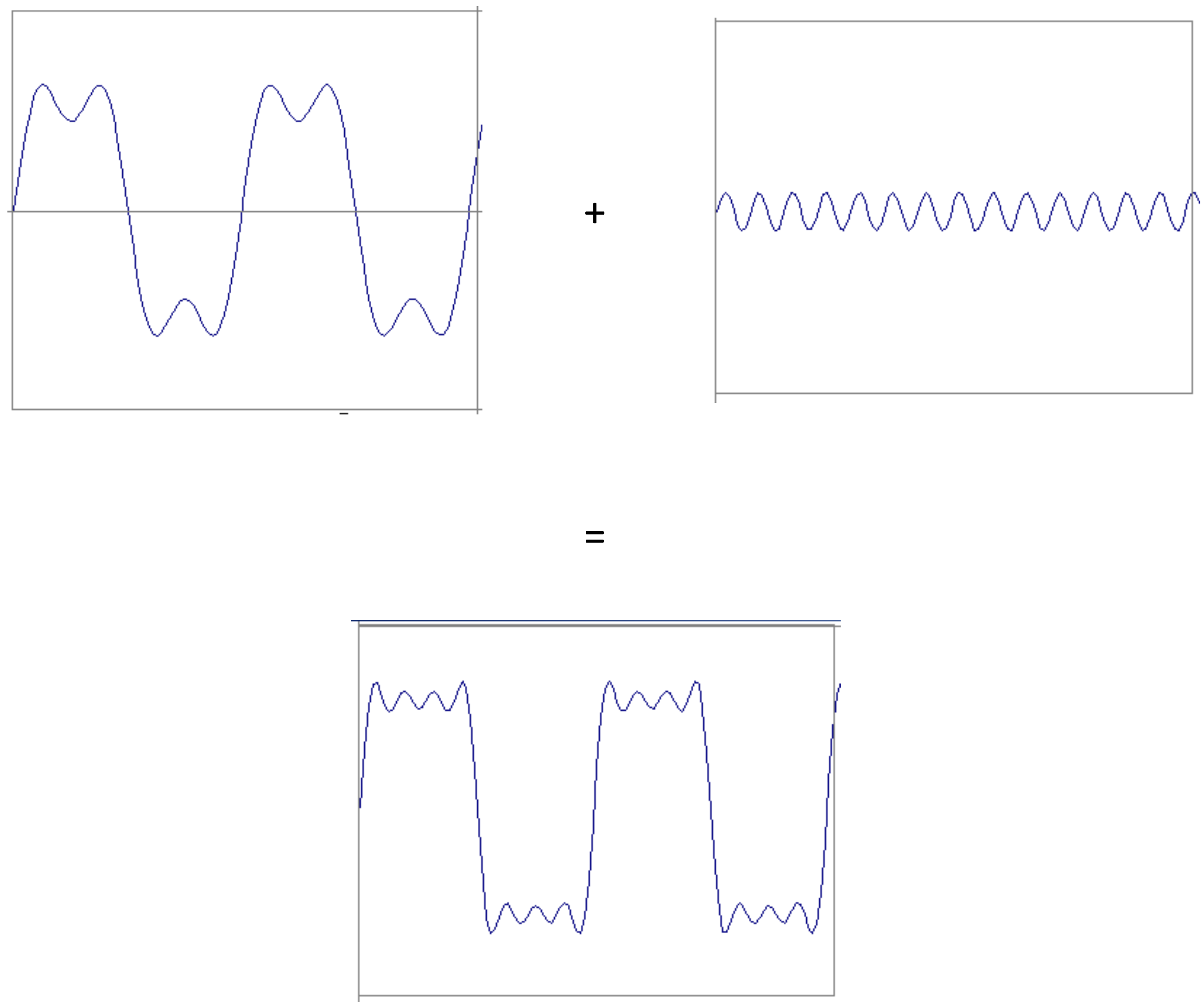
- Time axis
  - sampling
- Amplitude axis
  - quantization

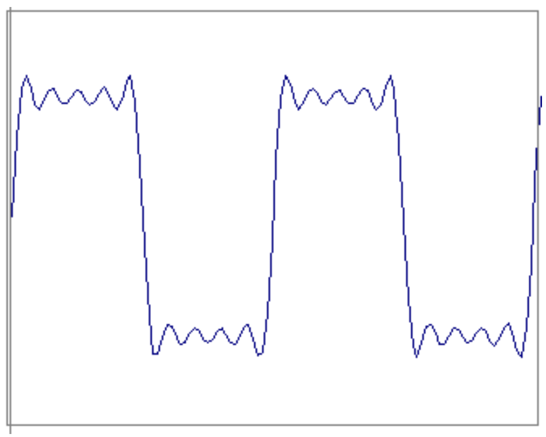



Ref: Dr. Wang Ye

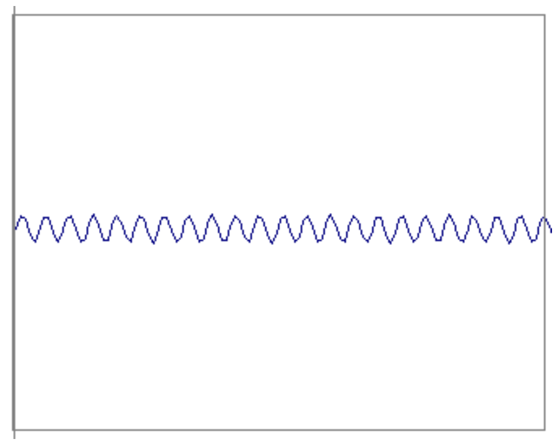
# Frequency Components of Audio Signal



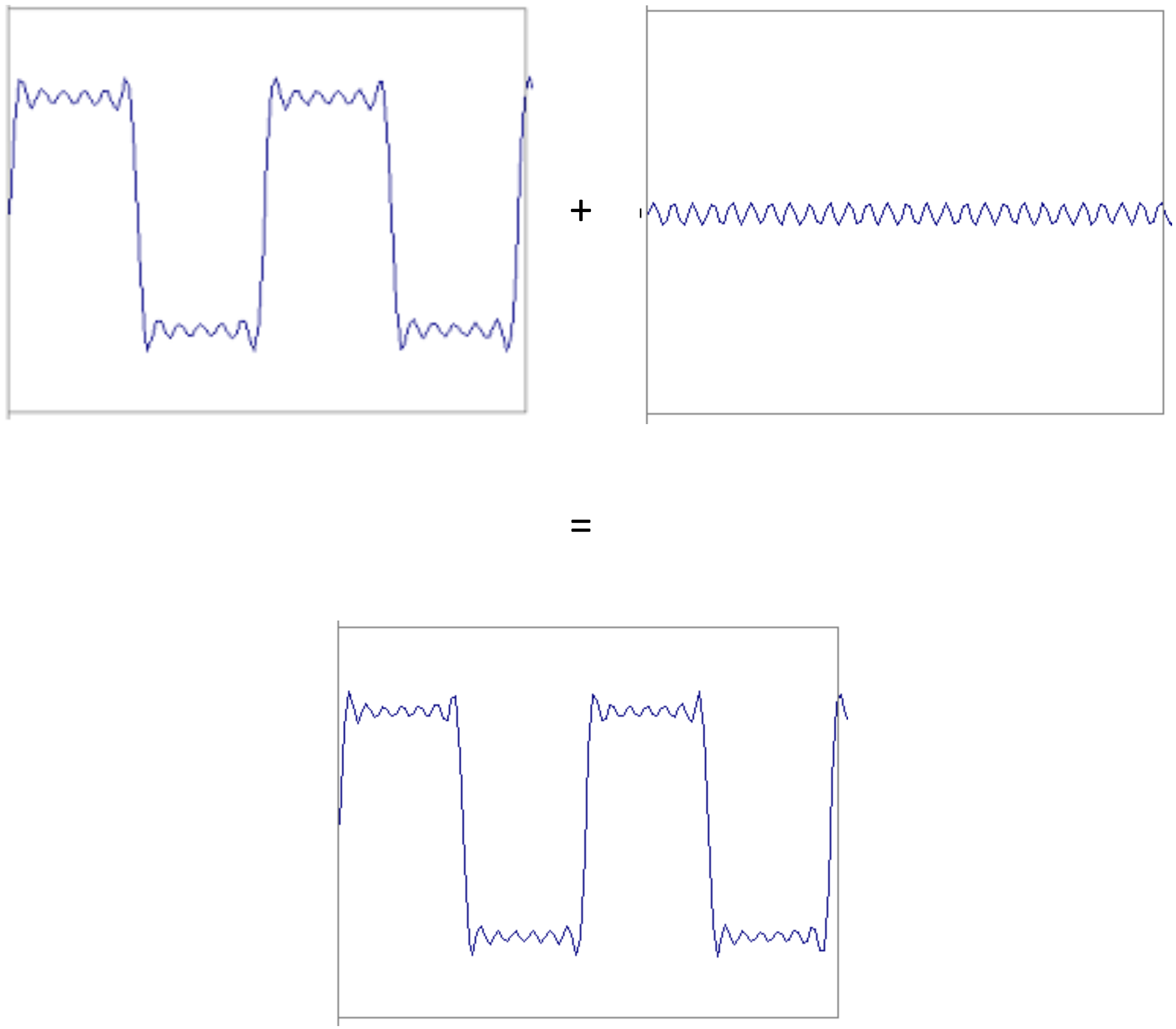




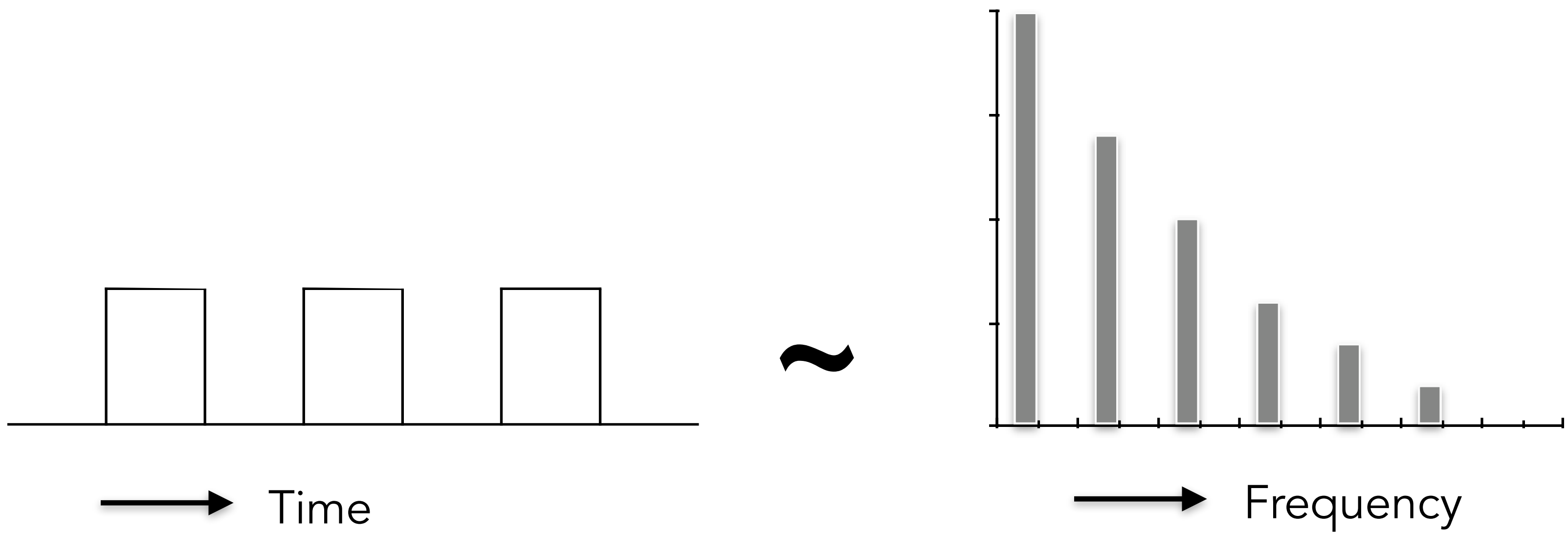
+



=

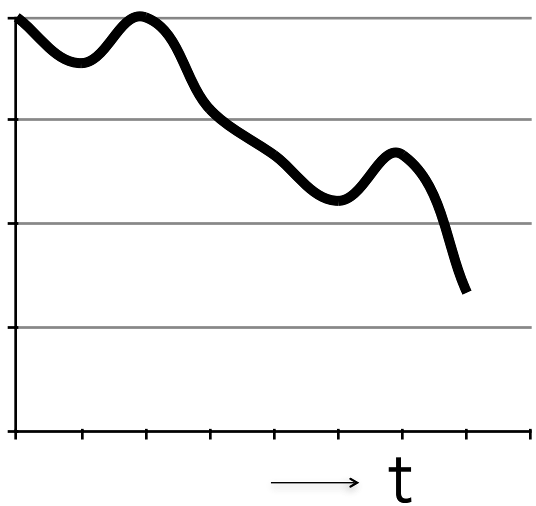


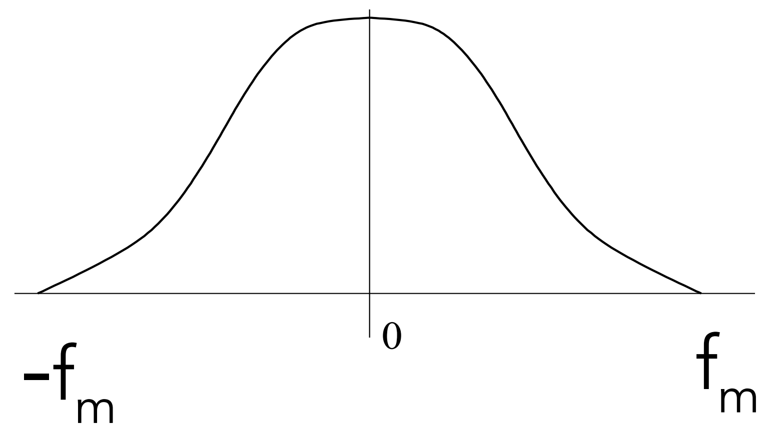
# Time domain To Frequency Domain




Any periodic signal can be represented in terms of its constituent frequencies!

# Time Domain To Frequency Domain




A non-periodic signal is represented with Continuous Frequency Response in Frequency Domain!

# Sampling Theorem

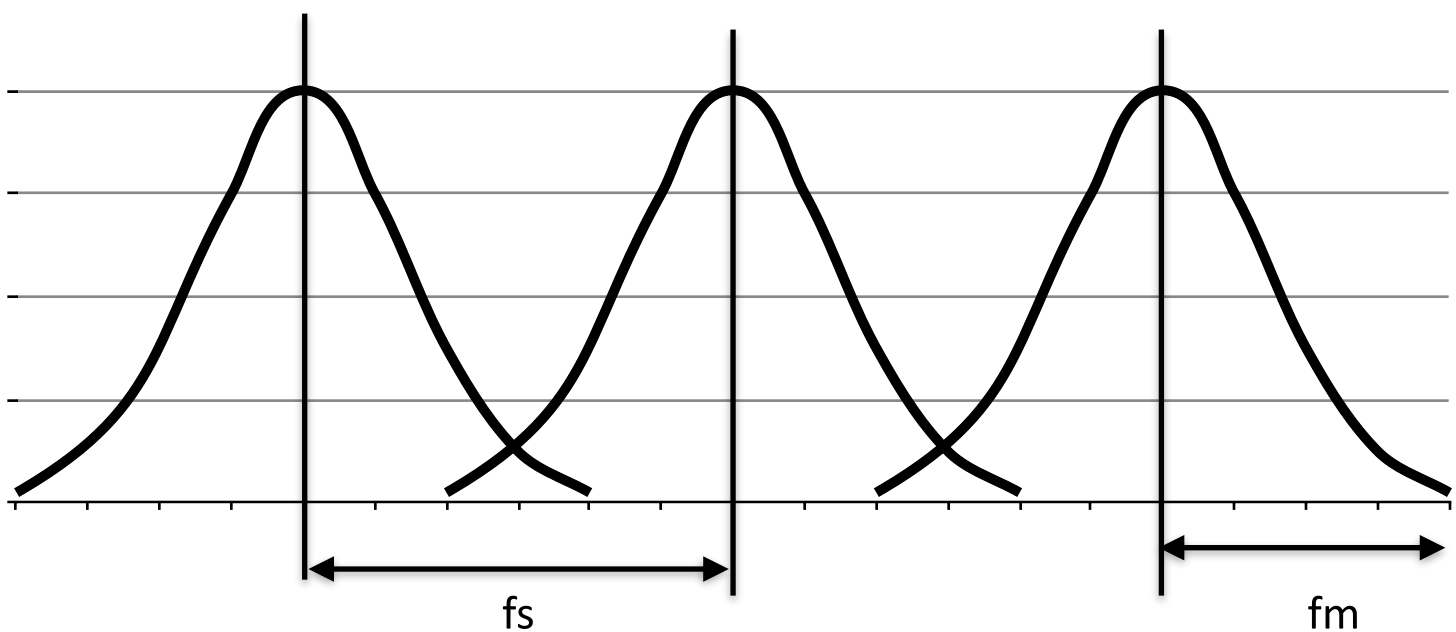
A signal can be reconstructed faithfully if it is sampled at a rate at least twice the maximum frequency component in it!

Nyquist rate = $2f_m$

Nyquist frequency = fs/2

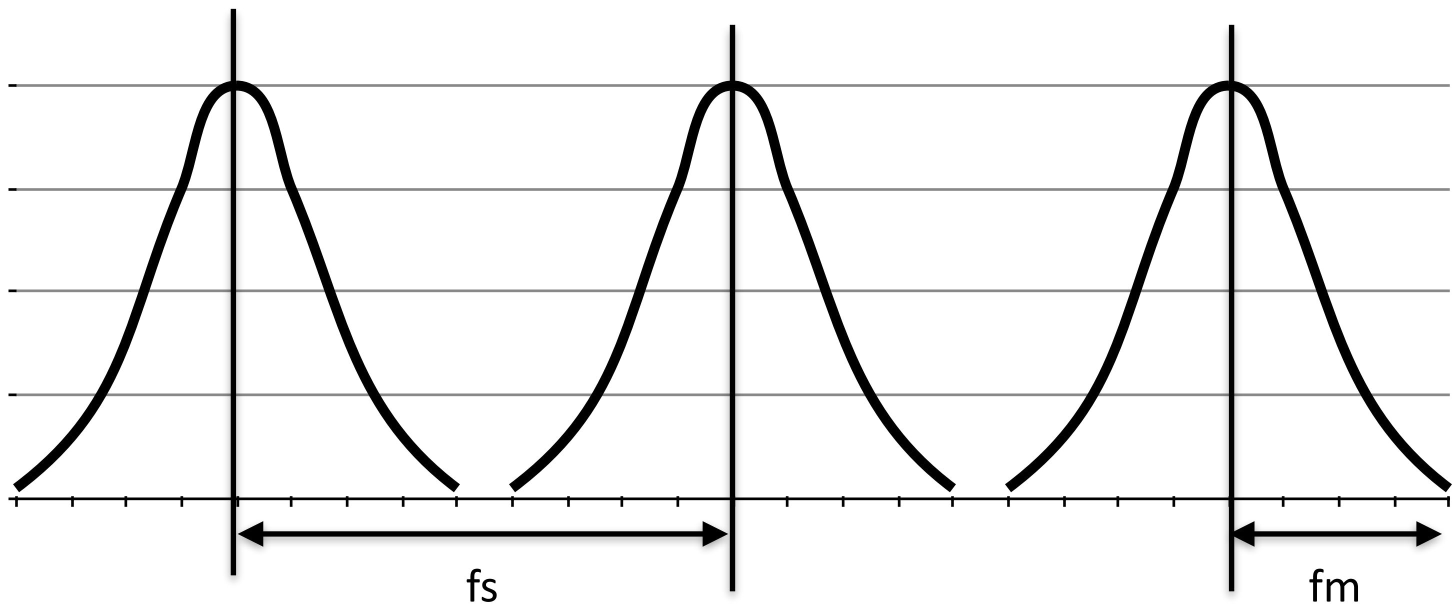
# Aliasing

- When the sampling rate is less than Nyquist, it leads to aliasing!
- Low frequency appears as high frequency



# No-Aliasing

- The original signal can be reconstructed faithfully using a low pass filter



How do you ensure no aliasing?

Low pass filter the signal before passing to ADC!

# What should be sampling rate for audio?

Audible Range: 20 Hz to 20 kHz

Sample (music) >40k samples

Voice Range: 0 to 4 kHz!

Sample(speech) > 8k samples

# Quantization

- Representing large set of values with a smaller set of values (called code words).
- The large set may have continuous values also (i.e. infinite set)

# Signal to Quantization Noise Ratio

$$SQNR_{db} = 6.02NDB$$

# Exercise: Quantize the following 5 bit signals into 2 bit signals!

## {23, 12, 9, 5, 31, 16, 19, 4, 13, 22}

- There are four code-words: 0, 1, 2, 3
- Obtain interval each code-word represents to obtain codes
- Obtain representative value corresponding to each code-word to decode
- Put all this information in a table.

# Linear Vs Non-linear Quantization

- Linear Quantization: equal step
- Non-linear Quantization: unequal steps

# Pulse Code Modulation

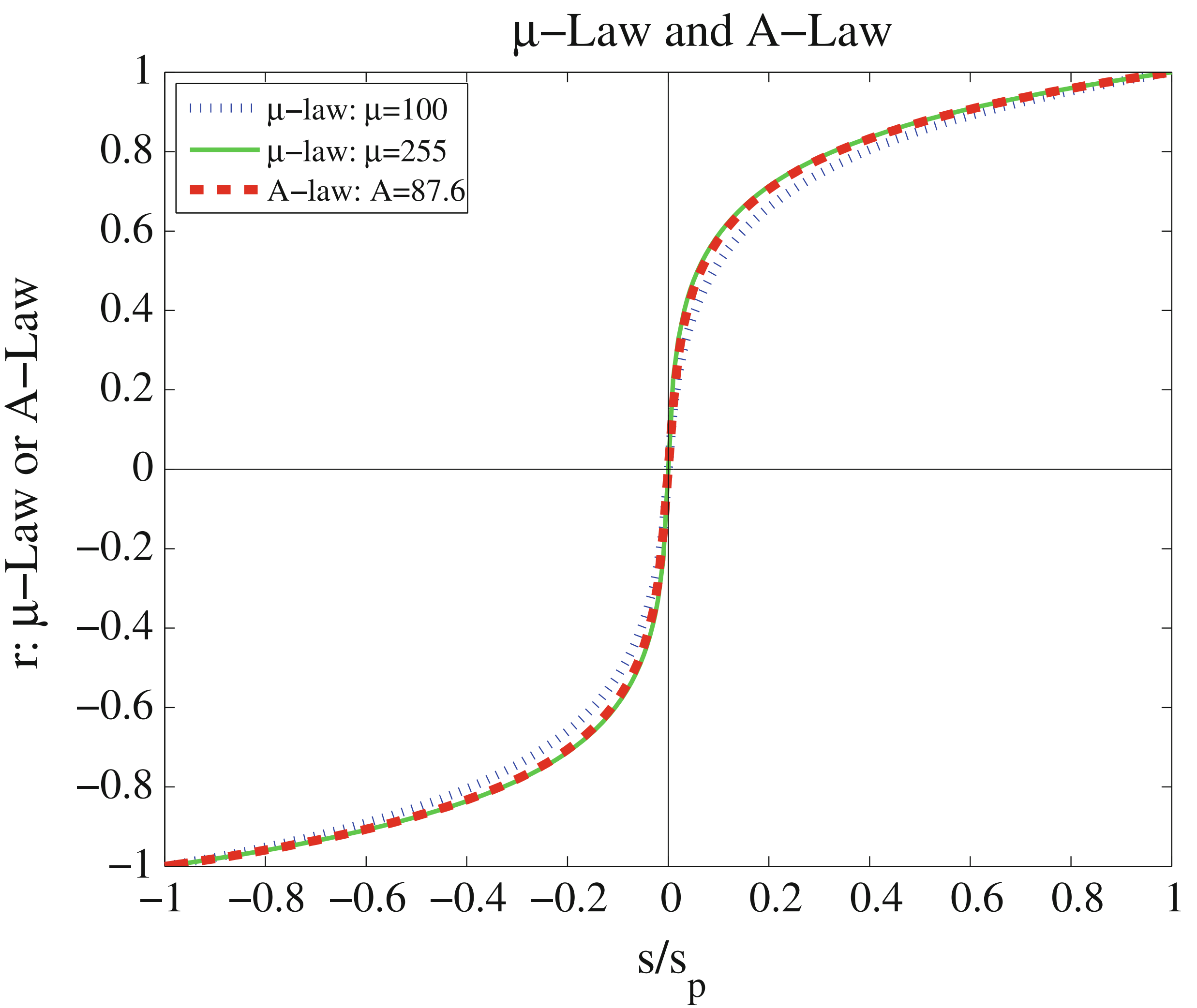
- Each audio sample is represented by an integer code-word.
- Linear PCM uses linear quantization and non-linear PCM  uses non-linear quantization
- Non-linear PCM is also called companding

# Weber's Law

$$\Delta\text{Response} \propto \frac{\Delta\text{Stimulus}}{\text{Stimulus}}$$

100 200 500 600 1200 1300

# Audio Perception Models



# Audio Perception Models

μ-law Companding

$$r = \frac{\operatorname{sign}(s)}{\ln(1+\mu)} \ln\left\{1 + \mu\left|\frac{s}{s_{\mathrm{p}}}\right|\right\}$$

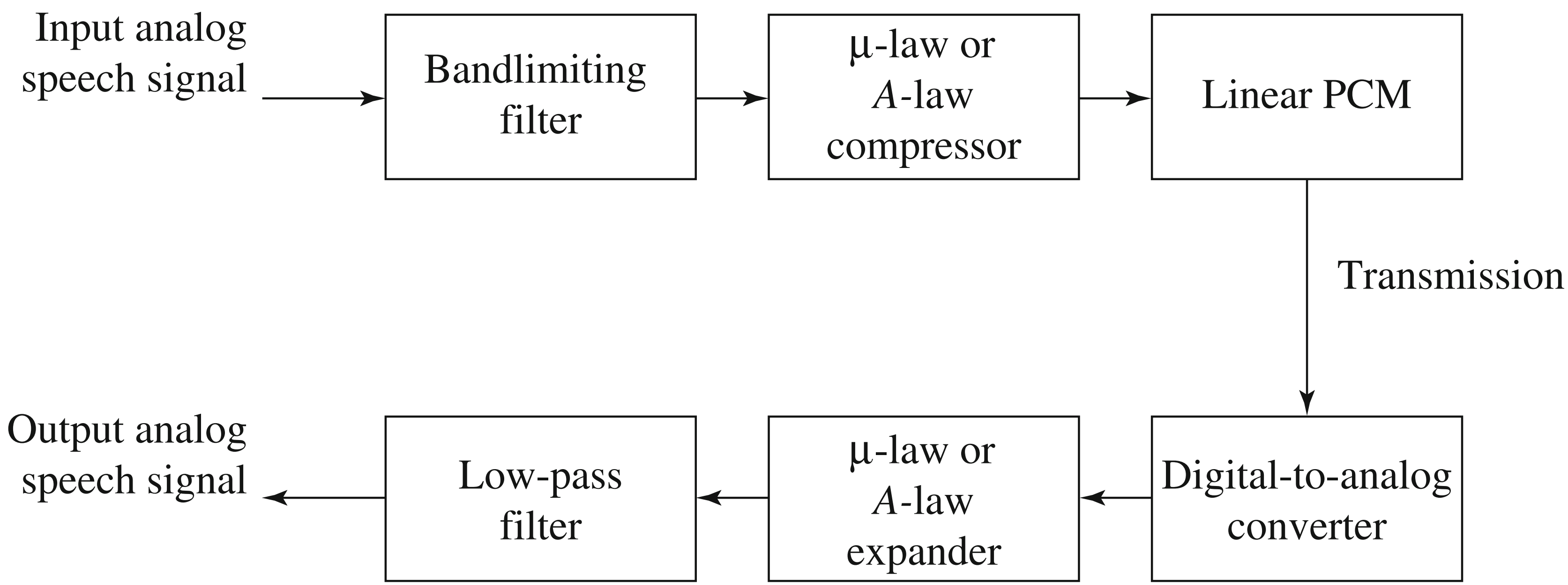
A-law Companding!

$$r = \begin{cases} \frac{A}{1+\ln A}\left(\frac{s}{s_{\mathrm{p}}}\right), & \left|\frac{s}{s_{\mathrm{p}}}\right| \leq \frac{1}{A} \\ \frac{\operatorname{sign}(s)}{1+\ln A}\left[1 + \ln A\left|\frac{s}{s_{\mathrm{p}}}\right|\right], & \frac{1}{A} \leq \left|\frac{s}{s_{\mathrm{p}}}\right| \leq 1 \end{cases}$$

$$\text{where } \operatorname{sign}(s) = \begin{cases} 1 & \text{if } s > 0, \\ -1 & \text{otherwise} \end{cases}$$

# How to Implement Companding?

1. Use non-uniform quantization steps in the ADC internally
2. Use a 12 bit ADC and a software lookup table to get 8 bit codes
3. Use additional nonlinear analog circuit before linear ADC


Input analog speech signal
Bandlimiting filter
μ-law or A-law compressor
Linear PCM
Transmission
Output analog speech signal
Low-pass filter
μ-law or A-law expander
Digital-to-analog converter

Companding in mainly used in Telephony!

| Quality | Sampling rate (kHz) | Bits per sample | Mono/ Stereo | Bitrate (if uncompressed) (kB/s) | Signal bandwidth (Hz) |
|---|---|---|---|---|---|
| Telephone | 8 | 8 | Mono | 8 | 200–3,400 |
| AM radio | 11.025 | 8 | Mono | 11.0 | 100–5,500 |
| FM radio | 22.05 | 16 | Stereo | 88.2 | 20–11,000 |
| CD | 44.1 | 16 | Stereo | 176.4 | 5–20,000 |
| DVD audio | 192 (max) | 24 (max) | Up to 6 channels | 1,200.0 (max) | 0–96,000 (max) |