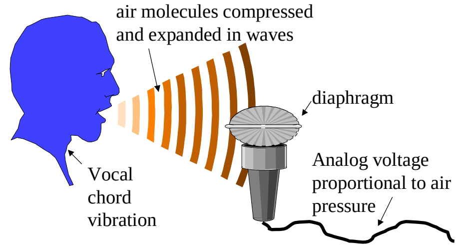
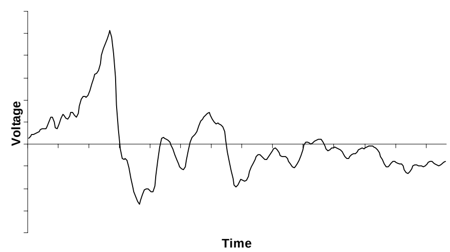
Audio Basics

- Analog to Digital Conversion (Pulse Code Modulation)
- Sampling
- Quantisation

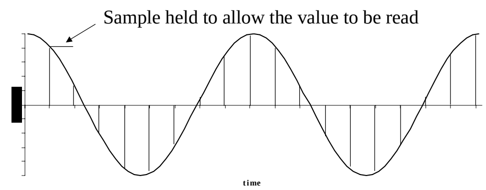
Analog Audio



A 5ms Example Waveform



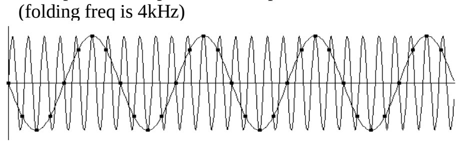
Sample & Hold (digitising the time axis)



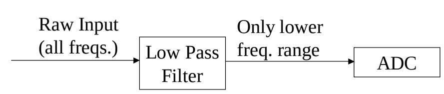
Sample-rate: the number of samples per second
eg telephone: 8000 samples/sec CD: 44100 samples/sec

Aliases (Undersampling)

- Alias freq = abs(actual freq – closest integer multiple of sampling freq)
- Folding freq: the frequency above which aliasing occurs (half the sampling rate)
- Can be useful! (eg software-defined radios)
- Example: 7kHz input, 8kHz sample rate: 1kHz alias, (folding freq is 4kHz)



Analog low pass filter



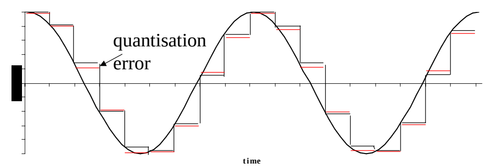
- Applied before sampling (Analog-to-Digital Converter)
- Removes HIGH frequencies (above a set threshold)
- A.k.a.: anti-aliasing filter

Nyquist-Shannon sampling theorem

- First formulated by Harry Nyquist in 1928 ("Certain topics in telegraph transmission theory").
- formally proved by Claude E. Shannon in 1949 ("Communication in the presence of noise"). Mathematically, the theorem is formulated as a statement about the Fourier transform.
- sampling rate ≥ 2 x bandwidth (for no loss of information)
- Nyquist freq. (rate): highest freq that can accurately be represented
- Nyquist interval: time between samples
- Example: 20kHz range required (~limit of human hearing).
  - By Nyquist, sample rate must be ≥ 40000 samples/sec

Quantisation (digitising the voltage axis)

—— = quantised value

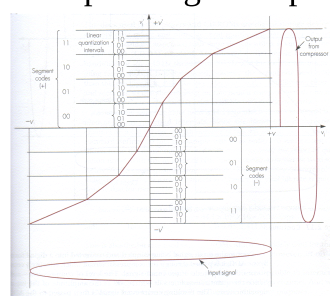


Precision: the number of bits used to represent the voltage
eg telephone: 8-bit (256 levels) CD: 16-bit (65536 levels)

Companding

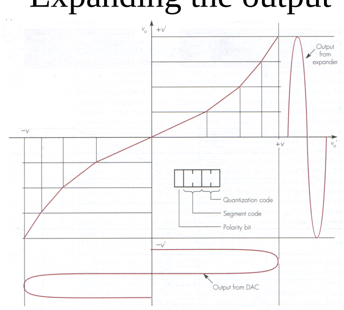
- Abreviation of compressing/expanding
  - In terms of the dynamic range (in dB – decibels)
  - dB = 10 log P, (dBm is a decibel relative to 1mW).
- Rationale for companding:
  - Low amplitudes are more frequent
  - SNR (Signal to noise ratio) is lower at low amplitudes
  - Human ear is more sensitive to low amplitudes
- So, quantise on a logarithmic scale (more levels for low amplitudes, less for high amplitudes)

Compressing the input



source: "Multimedia Communications" by Fred Halsall, page 91

Expanding the output



source: "Multimedia Communications" by Fred Halsall, page 92

ITU-T standard G.711

- μ-law (pronounced mu-law) is used in US and Japan
- A-law is used in EU, rest of the world, and international routes