Foundations of Audio Technology

Nimal Skandhakumar

Faculty of Technology
University of Sri Jayewardenepura

2020

Partially based on:

- MusonicX: Music Technology Foundations - University of Adelaide
  - https://www.edx.org/course/music-technology-foundations
  - https://www.youtube.com/channel/UCCbr_KhAzyMGeja81KjolRA

Technology in Sound

- Creation of sound
  - Bells, Musical Instruments, Synthesizers
- Recreation of sound
  - Musical Notation, Recording and playback, Magnetic tapes, PCM recording
- Analog vs. Digital

1877

Phonographic Cylinder

1883

Piano (Music) Roll

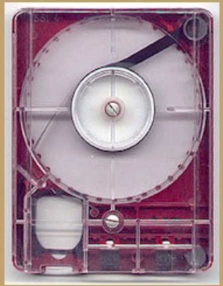
1895

Gramophone Record

1930s

Reel-to-Reel

1962



4-Track (Stereo-Pak)

1963

Compact Cassette

1965

8-Track

1982

Compact Disc

1993

MP3

1997

DVD

2006

Blu-ray Disc

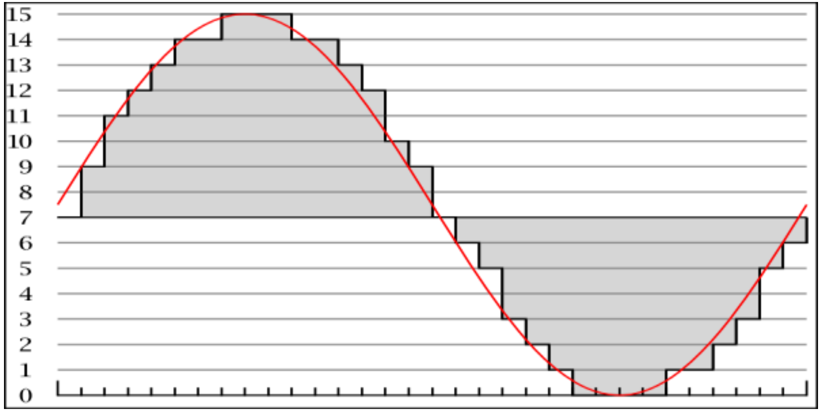
Recording Sound – Evolution of Mediums

Audio Technology

- Analog audio involves the creation or imitation of a continuous wave.
  - Vinyl, tape, analog synths, etc.
- Digital audio data is the actual representation of sound.
- Stored in the form of samples:
  - Samples represent the amplitude of sound at a discrete point in time.
  - Quality of digital recording depends on the sampling rate
    – the number of samples taken per second.
  - Remember ADC & DAC?

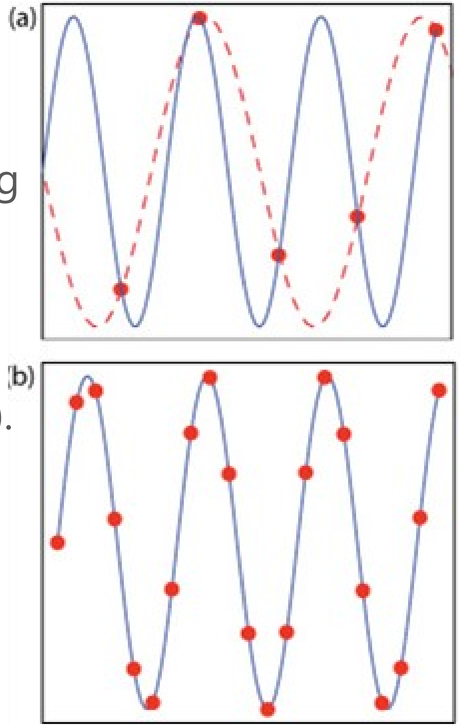
Digitising Analog Audio

- Sampling:
  - Divide the horizontal axis (time) into discrete pieces
- Quantization:
  - Divide the vertical axis (signal strength - voltage) into pieces.
  - For example, 8-bit quantization divides the vertical axis into 256 levels.



Digitising Analog Audio

- Nyquist Theorem
  - Nyquist theorem is used to calculate the optimum sampling rate in order to obtain good audio quality.
  - For Lossless digitization, the sampling rate should be at least twice the maximum frequency responses.
  - Digitally sampled audio has a bandwidth of (20Hz - 20KHz).
  - By sampling at twice the maximum frequency (40KHz) we could have achieved good audio quality.
  - CD audio slightly exceeds this, resulting in an ability to represent a bandwidth of around 22050 Hz.



Digitising Analog Audio

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

Audio Filtering

- Prior to sampling and analog-to-digital conversion, the audio signal is also usually filtered to remove unwanted frequencies.
- The frequencies kept depend on the application:
  - For speech, typically from 50Hz to 10kHz is retained, and other frequencies are blocked by the use of a band-pass filter that screens out lower and higher frequencies.
  - An audio music signal will typically contain from about 20Hz up to 20kHz.

Audio Quality vs. Data Rate

Quality	*Sample Rate (kHz)*	*Bits per Sample*	*Mono / Stereo*	*Data Rate (kBytes/sec) (uncompressed)*	*Frequency Band*
Telephone	8	8	Mono	8	200-3400 Hz
AM Radio	11.025	8	Mono	11.0	540-1700 KHz
FM Radio	22.050	16	Stereo	88.2	
CD	44.1	16	Stereo	176.4	20-20000 Hz
DAT	48	16	Stereo	192.0	20-20000 Hz

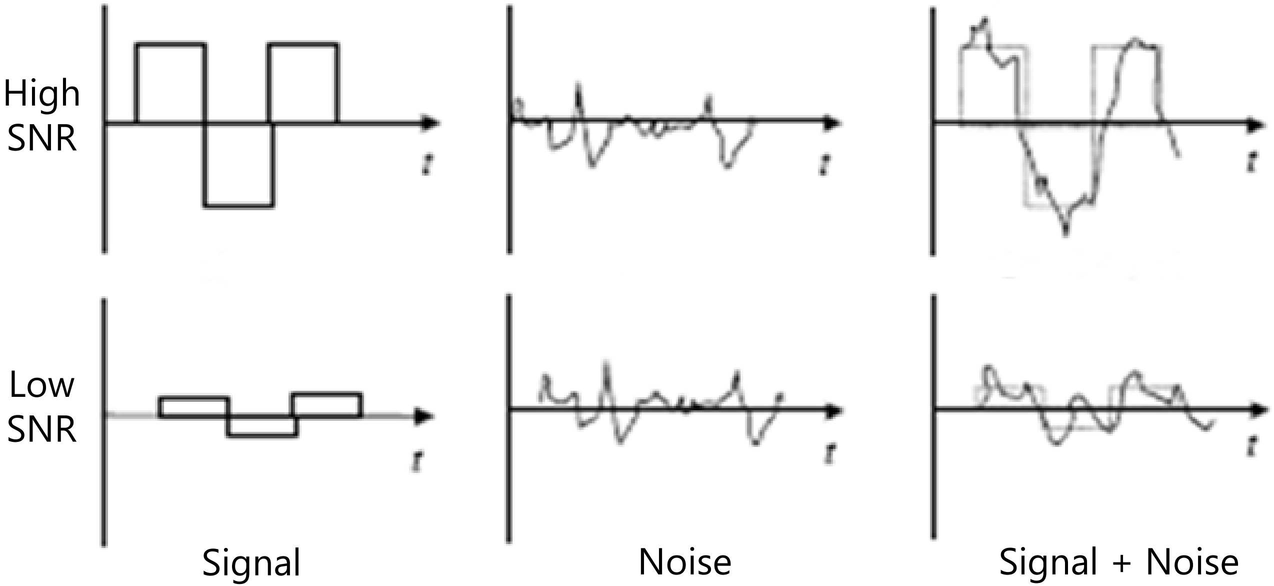
Digitising Analog Audio

- Sampling date
  - Samples per second
  - Determines frequency range that can be recorded
  - CD: 44,100 Hz
  - Pro: 48,000 Hz
- Bit depth
  - Quality of each sample
  - Determine the quality of dynamic range
  - CD: 16 bit
  - Pro: 24 bit

Digital Audio Processing

- A completed recording almost always needs to be edited.
- Basic sound editing operations include,
    - trimming, splicing and assembly
    - volume adjustments
    - working on multiple tracks
- Additional available sound editing operations include format conversion, resampling or downsampling, fade-ins and fade-outs, equalization, time stretching, digital signal processing, and reversing sounds.

Digital Audio Noise



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

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

Signal-to-Noise Ratio

Audio Application	Frequency Response	SNR (dB)
Plain Old Telephone System or POTS	300Hz to 3kHz	40
AM Radio, LP Records	100Hz to 5kHz	50
FM Radio, Cassettes	50Hz to 15kHz	70
Consumer Stereo System, CD Player	20Hz to 20kHz	90
Professional Audio Equipment	5Hz to 24kHz	120

Digital Audio File Formats

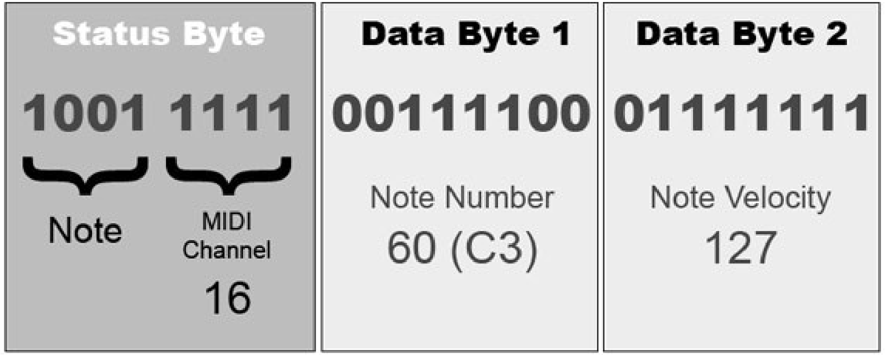
- Uncompressed
  - Recording, editing
  - Maintain full information
  - .wav, .aiff
- Compressed
  - Transfer, stream
  - Reduce file size
  - .mp3, .ogg

Digital Audio Compression

- Compression can be either,
  - Lossless:
    The quality of uncompressed audio equals the quality of original audio file
  - Lossy:
    The quality of uncompressed audio is lower than the original (e.g. mp3, wma)
- For compressing and decompressing matching algorithm pairs are used
  - Algorithms = Codecs (MP3, WMA, OGG, APE, FLAC)

MIDI Audio

- **M**usical **I**nstrument **D**igital **I**nterface is a protocol that enables computer, synthesizers, keyboards, and other musical devices to communicate with each other.
- MIDI is a shorthand representation of music stored in numeric form.
- A sequencer and sound synthesizer is required to create MIDI scores.
- It is not digitised sound.
- MIDI is device dependent.



Practicals

Class Project 1

Recording & Editing Sounds

1. Record 3 sounds from your local environment using a mobile phone.
2. Edit and clean the audio samples using a digital editing software.
3. Compile all 6 clips into a single audio file.
4. Upload this to LMS as a compressed audio file.

Doing is learning