MUSIC IN REAL LIFE VS
MUSIC IN DIGITAL WORLD

AIST2010 Lecture 2
OUTLINE

SOUND AND MUSIC IN REAL LIFE

DIGITIZING SOUND AND MUSIC DATA

MUSICAL INSTRUMENTS

MUSIC HARDWARE AND SOFTWARE
Vibration and sound
- The vibration of air particles from the source into your ears
  - Or via other media (e.g. water)
  - Vibrations detected by human auditory system ➔ “audio” signals
- Different vibrations give rise to different kinds of sounds
  - Some are harmonious, some are noise, some are high, some are low

Music
- A clever combination of sounds and voids

Representing sounds
- It’s always an abstraction, e.g.
  - Waveform (amplitude vs. time)
  - Spectrum (frequency vs. amp. vs. time)
For periodic signals, frequency measured in Hz → cycle per second (cps)
Note: There are other ways to represent the spectrum too.
BASIC PERIODIC WAVEFORMS

A pitch is heard for periodic waveforms (at certain frequencies)
- Sine wave
- Cosine wave
  - Phase-shifted sine wave
- Triangular wave
- Sawtooth wave
- Square wave

Commonly used in vintage synthesizers (i.e. electronics)

Image from: https://moogfoundation.org/learning-synthesis/synthesis-fundamentals/
**Human sound perception is subjective**
- General range: ~20 Hz – ~20,000 Hz (20 kHz)
- Hearing range narrows with age, especially at high frequencies
- How well do animals hear? [https://lsu.edu/deafness/HearingRange.html](https://lsu.edu/deafness/HearingRange.html)

**Human is more sensitive to ~2 to 4 kHz**
- A not-so-flat “frequency response”
- For the sake of human communication?
- Perception of loudness is also affected by age
MEASURING LOUDNESS

Sound power = (amplitude)²
- Sound intensity = sound power per unit area (watt per sq. metre)

“Threshold of hearing” (TOH) vs “Threshold of pain” (TOP)
- \( I_{TOH} \) \( \overset{\text{def}}{=} 10^{-12} \text{ W/m}^2 \)
- TOP is roughly \( 10^{13} \) times of the intensity of TOH

Note: Power depends on distance!
MEASURING LOUDNESS

Decibel (dB) is a logarithmic unit to measure the ratio

$$\text{dB}(I) \overset{\text{def}}{=} 10 \cdot \log_{10} \left( \frac{I}{I_{\text{TOH}}} \right)$$

- i.e. TOP is 130 dB when TOH is 0 dB
- This value relative to TOH is also called the sound pressure level (SPL), i.e. dB\text{SPL}
MEASURING LOUDNESS

Besides the intensity, the **amplitude** (which we discuss more) can also be represented in dB, usually we assume amplitude is 1 at 0dB

\[ A_{dB}(x) = 20 \cdot \log_{10} x \]

Do not mess up with:

- dB\(_A\) or dB(A) represents the loudness perceived by human ear
  - Compensation using the A-weighting basing on equal-loudness curves

This term is 20 because intensity is amplitude squared
MEASURING PITCH

Frequency range on a piano keyboard

“Middle C” = 261 Hz
“A440” = 440 Hz

- Double frequency = “octave” higher
- The next occurrence of the same letter-name in music
MEASURING PITCH

Pitch, like dynamics, is perceived in a logarithmic manner

- Feeling of doubling amplitude
  - The change from 0.25 to 0.5 is similar to that from 0.1 to 0.2, a doubling of volume

- Feeling of doubling frequency
  - The change from 100 to 200 Hz is similar to that from 250 to 500 Hz, an octave higher

Perceptual scales of pitch

- Mel-scale
- Critical bands (Bark scale)
- Equivalent rectangular bandwidth (ERB) rate
- log Frequency

Empirical scales!
ASPECTS OF MUSIC

When sounds and rests are cleverly put together, we have music

- Melody
  - combination of pitches (at different time)
- Rhythm
  - combination of long and short sounds
- Harmony (chords)
  - combination of pitches (at the same time)
- Harmonic progression
  - combination of chords (at different time)

- Dynamics
  - the expression in loudness
- Structure
  - how music sections are arranged
- Texture
  - the amount of different/similar sounds at times
- Timbre
  - different shapes of waveforms
  - different spectrum

There can be more!
How can we represent audio/sound/music in a computer?

- Two main ways
  - **Faithful** representation of what can be heard: *audio recording and playback*
    - To improve the audio representation, research is on the improving sound fidelity
  - **Reconstruction** of music data (e.g. melody, rhythm, etc.): *audio analysis and resynthesis*
    - To improve the symbolic representation, research is on the analysis and synthesis methods
To represent audio signals in the digital world, digitizing is needed

- **ADC** (analog-digital conversion) for recording
- **DAC** (digital-analog conversion) for playing back

**Sampling:** the representation in time

- The higher sampling rate the better
  - CD quality: 44.1 kHz → 44100 samples every second

**Quantization:** the representation in amplitude

- The more dynamic levels the better
  - CD quality: 16-bit → \(2^{16}=65536\) different levels of amplitude
**SAMPLING RATE**

**How many samples per second is the best?**
- Storage requirements vs. sound quality

**Sampling Theorem**
- Signal reconstructable if frequency is not higher than
  - **Nyquist frequency**: $\Omega = F_s/2$, where $F_s$ is the sampling rate
  - Otherwise artifacts called **aliasing** will form (an extra freq. component)

**Common rates**

<table>
<thead>
<tr>
<th>Analog telephone: 8,000 Hz</th>
<th>Pro digital video equipment: 48,000 Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low quality audio: 22,050 Hz</td>
<td>DVD-Audio, Blu-ray audio: 96,000 Hz</td>
</tr>
<tr>
<td>Audio CD, MP3: <strong>44,100 Hz</strong></td>
<td>Super-Audio CD: 2,822,400 Hz</td>
</tr>
</tbody>
</table>

Image from: *Fund. of Music Processing*, p. 62

440Hz signal at 22,050Hz sampling rate

Aliasing occurs at (b) and (c)
DYNAMIC LEVELS

The dynamic level is controlled by the bit depth
- Typically 8-, 16-, and 24-bit, corresponding to \(2^8\), \(2^{16}\), and \(2^{24}\) levels
- Quantization error between analog value and quantized value

Represented as “Amplitude”
- Integer vs. floating
- Signed vs unsigned

Common bit depths
- Audio CD: 16-bit (~96 dB of range)
- Pro audio: 24-bit (~144 dB)
Different audio perception on our left and right ear helps determining the location of the sound source

- "Stereo" sound, with 2 channels
  - Some professional recordings are shaped with Head Related Transfer Function (HRTF)

Surround sound systems
- 5.1 or 7.1 with subwoofer
- Extra channel = extra storage

HKBU LIATe theatre with 24.2-channel sound system

Image from: http://liate.hkbu.edu.hk/theatre.html
STORING AUDIO DATA

The raw audio data, bit by bit, is called linear pulse-code modulation “linear PCM” encoding (usually as .wav files)

- Good enough for sound editing and manipulation
- Consuming too much storage
  - CD quality: 44.1kHz 16-bit 2-channel = 80 minutes consuming ~700 MB

Compressed audio can fulfil requirements of the general public

- Lossy: MP3, AAC, WMA, …
- Lossless: FLAC, ALAC (.m4a), …
- Audio quality often measured as bandwidth requirement
  - Spotify “normal quality”: AAC format at ~96 kbps
DEALING WITH MUSIC DATA

Musical Instrument Digital Interface (MIDI)
- Standardized in 1983
- Widespread technology to connect various musical hardware and software

Control data instead of actual audio data
- Symbolic representation
- Playback = Resynthesis on-the-fly
- Basic messages: “Note On/Off, note number, velocity”
- Only decides when and what should happen, but not HOW
  - Synthesizer will make the actual sound (e.g. a violin sound, or a piano sound)
BASIC MIDI MESSAGES

“Note On” denotes the starting point, where “Note Off” is to stop

Calculating the note number \((0-127)\)

\[
p = 69 + 12 \times \log_2 \left( \frac{f}{440 \text{ Hz}} \right)
\]

Velocity \((0-127)\): How fast the key is pressed (correlates to playing intensity)
MIDI CHANNELS

16 MIDI channels on a MIDI connection

- Driving 16 different instruments simultaneously
- Each having their own set of notes (note number, velocity, on/off)
  - More than one note can happen at one time in a channel, but note on/off will correspond to note number
- “Polyphony”
- What if you want a whole orchestra?
DEALING WITH MUSIC DATA

After “noteon”, the note can be further modified, e.g. with 7-bit values of

- Pitch bend
- Modulation
- Aftertouch/Pressure
- Control change (CC)
  - Damper pedal, sostenuto, expression, portamento, …

As always, the result depends on the synthesizer!

https://www.midi.org/midi/specifications-old/category/reference-tables

**MIDI 2.0 draft** is just announced in January 2019…

- Increasing bit precision: finer control of music expression
DISPLAYING MUSIC DATA

Musicians read music, just like we read text
- Different music notation software have their own proprietary standard

An “open” format: MusicXML
- Storing music notes like MIDI
- Yet it also stores how it should be rendered and displayed
  - E.g. bar lines, key signatures, …
REAL WORLD MUSICAL INSTRUMENTS

How are audio signals generated in real life?

- Always vibrations…

Musical instruments are made to create regulated vibrations

- Usually with a controllable pitch (periodic waveform)
- Different instruments show different spectra (frequency components)
- The vibration may be amplified with an acoustic body, e.g.
  - Sound box of a violin
  - Sound board of the piano
REAL WORLD MUSICAL INSTRUMENTS

Piano ➔ Chordophones
- Keys driving hammers to hit steel strings and make them vibrate

Strings ➔ Chordophones
- Friction between bow and strings create vibrations

Winds ➔ Aerophones
- Air flow creates vibrations inside the tube of various length

Percussion ➔ Membranophones/Idiophones
- Too many possibilities! In general, an object vibrates when being struck

Note: We are not going to cover physics in this course!
Since the development of electricity, musicians started to explore into new ways of expressing themselves, e.g.

- **Telharmonium (1897)**
  - Tonewheels generates electric musical notes with electromagnets

- **Theremin (1920)**
  - Frequency/amplitude of the sound is controlled using the capacitance between hands and antennae

- **Analog synthesizers (1960 beyond)**
  - Sounds are generated and modified using op-amp, resistors, filters, and other electronics
INTERFACING WITH COMPUTERS

Audio interfaces
- “Sound cards”: providing signal conversion from analog to digital, and vice versa
- DAC + ADC

MIDI interfaces
- Input/output of MIDI messages between computers and instruments

Control surfaces
- Providing faders and buttons for direct control over DAW software UI
With computers, now we can easily make music with software programs

- MIDI = prominent tool for music

**MIDI controllers are the interface for musicians to make MIDI messages**

- Most controllers are in the shape of a piano keyboard
  - 88-key, 72-key, …
  - Weighted keys for realism of touching and expressiveness

- Other shapes
  - Winds, mallet percussion, drums, strings, …
There is a research conference for everything...

  - Check out publications if you are curious
    - [https://www.nime.org/archives](https://www.nime.org/archives)

Roli Seaboard, LinnStrument, …

- Multi-dimensional, polyphonic
MUSIC SOFTWARE

Synthesizers

- Earliest software synthesizers simply emulate the analog synthesizers
- Now there are more complex kinds of synthesis which is only made possible by computer programs
- **Keyboard workstations** are controllers with an embedded system for synthesis
- Sometimes exists as plugins for DAWs

Vocaloid

- A special kind of sound synthesizer, which can mimic **human voice** by speech synthesis with controls of pitch and other vocal expressions
**MUSIC SOFTWARE**

**Digital Audio Workstation (DAW)**
- Allowing musicians to put together the tasks for recording, editing, and producing of audio data
- Most of DAW supports VST plugins as a standard way to extend features
- To name a few: Pro Tools, Logic Pro, Reason, GarageBand

**Live Performance**
- Besides using keyboard workstations, a computer with proper software can be used for live performance with maximum flexibility with MIDI controllers and VST plugins
- Examples: MainStage, Ableton Live
Sheet music is still a major medium among musicians.

**Notation** software allow easy typesetting with computer keyboard or MIDI controllers:
- Examples: Sibelius, Finale, MuseScore
- Some DAW can also output typeset music

While text can be recognized by OCR technology, handwritten or printed music can also be imported with proper recognition software:
- More and more popular on mobile devices
LECTURE REVIEW

We have discussed:

- How sounds are produced, and perceived by human
- Ways to digitize sound and music into computer systems
  - Audio vs. symbolic representations
- Music instruments and controllers
- Music hardware and software
READ FURTHER

Section 1.3 “Audio Representation”, Fundamentals of Music Processing

Chapter 1 “Audio and Music Signals”, Computer Music Instruments

Section 8.1.1 “The MIDI protocol”, Computer Music Instruments