**L1**

Applications of audio ML:

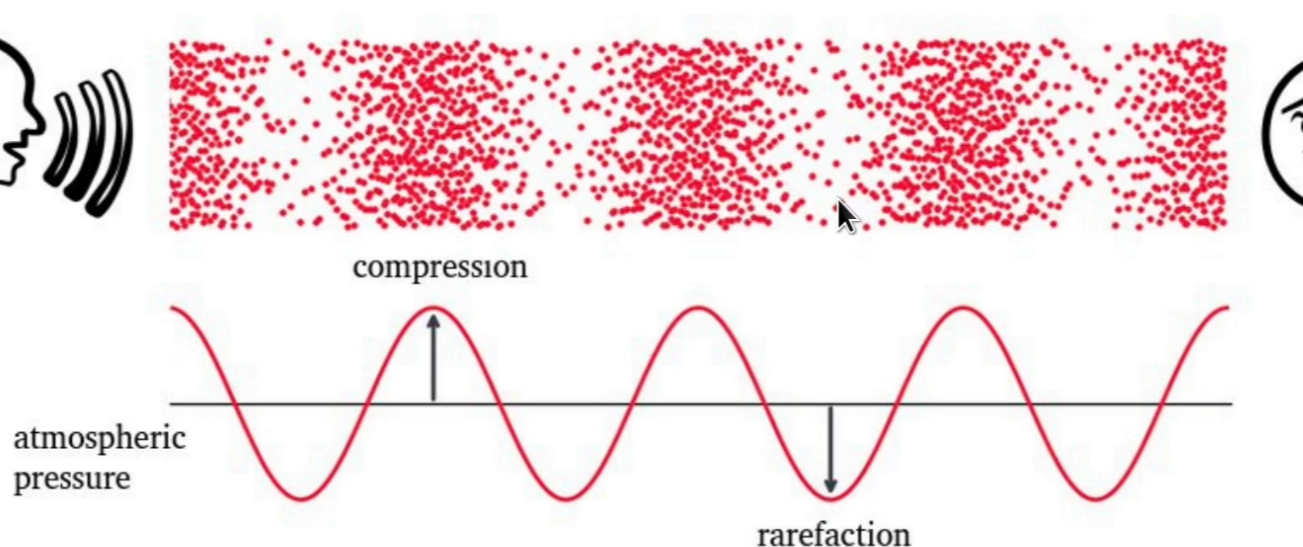
* Classification, speech recognition, denoising/upsampling, music information retrieval
  + Music has sub applications: instrument recognition, etc.

Topics:

* Sound waves, DAC/ADC, time-frequency domain audio features
* Audio transformations
  + Fourier/STFT, Constant-Q, Mel Spectrograms, Chromograms

**L2**

Sound is produced by vibrations, causing air molecules to oscillate. Changes in air pressure create a mechanical wave.



A diagram of different types of waves

Description automatically generated

Equation for a simple sine wave

y(t) = Amplitude \* sin(2pi \* frequency \* time + phi (phase))

The **period** is between each peak or dips. **Frequency** is the inversion of the period, number of cycles per second. **Amplitude** is the vertical distance between 0 and the peak or dip of the current point. **Phase** tells us the position of the waveform at time 0, and we are able to shift the waveform.

Perceptually, **larger amplitudes** will have **larger sounds**. This makes sense, because amplitude measure change in air pressure, which in perception results in higher noise.

**Pitch** is the perception of frequency. It’s logarithmic and is perceived similar if differ by a power of 2. Beginning concept of octave. Each increase in octave, the same note will have double the frequency in Hz.

Pitch to Frequency Equation:

Ratio between two subsequent semitones:

**Cents** are a subdivision of semitones. There are 1200 cents in Octaves, and 100 cents in semitone.

**L3**

Intensity, loudness, and timbre

**Sound power** is the rate at which energy is transferred. Energy per unit of time emitted by a sound source in all directions. Measured in watts (W).

**Sound intensity** is sound power per unit area. Measured in watts/meters squared (W/m2)

Questions: How much power is there in thunder, concerts? Answer: Only 1 Watt.

Humans can perceive sounds with very small intensities.

* Threshold of hearing (TOH): 10-12 W/m2
* Threshold of pain (TOP): 10 W/m2
* Humans can perceive an incredible amount of intensity range before reaching TOP.

**Intensity level** is measured in decibels (dB) and is logarithmic in scale.   
**Decibel** is a ratio between two intensity values.   
**Intensity of Reference** (TOH) is used as a baseline to create a decibel measurement.

Equation for intensity level: :

At 0 decibel, the threshold of hearing is reflected. Every 3 decibel, intensity doubles.

A table with numbers and text

Description automatically generated

From: Fundamentals of Music Processing Fuller

**Loudness** is a subject measure of intensity. It is the perception of a sound’s intensity. Depends on duration/frequency of a sound and age. It is measured in phons.

**Phons** are a measure of perceived loudness, even at different intensity levels. At lower frequencies, sounds are perceived less loud even at the same intensity.

**Timbre** is the ‘color’ of sound. The remaining difference between two sounds with the same intensity, frequency and duration. Words such as ‘bright, dark, dull, harsh, etc.’ describe timbre. Multi-dimensional measure of sound.

* Sound envelope
* Harmonic content
* Amplitude/frequency modulation

**Sound Envelope**

Attack-Decay-Sustain-Release Model

A diagram of a line with arrows and a letter

Description automatically generated

**Harmonic Content**

**Complex Sound** is a superposition of sinusoids. A **partial** is a sinusoid used to describe a sound.

**Fundamental frequency** is the lowest partial.

**Harmonic partial** is a frequency that is a multiple of the fundamental frequency.

* i.e. f1 = 440 hz, f2 = 2 \* 440 = 880, f3 = 3 \* 440 = 1320, …
* Inharmonicity indicates a deviation from a harmonic partial.

**Spectrograms** show the intensity of a frequency at a specific time in the duration of an audio sample.

* One reason why two sounds may be different, despite having the same intensity, frequency and duration, is the distribution of harmonic partials. This can be shown with a spectrogram.

**Frequency Modulation** is periodic variation in frequency, aka vibrato. Apply a messenger signal with frequency modulation on a carrier signal to achieve effect.

**Amplitude Modulation** is a periodic variation in amplitude, aka tremolo. Apply a messenger signal with amplitude modulation on a carrier signal to achieve effect.