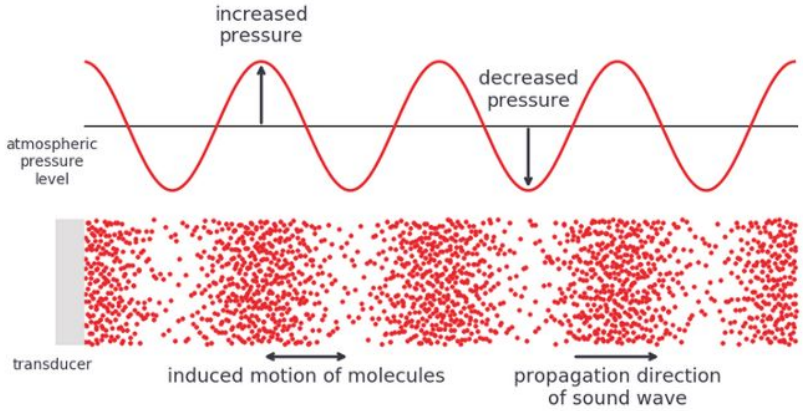
How does machine understand the audio?

CONTENT

SOUND
IN PHYSICAL WORLD

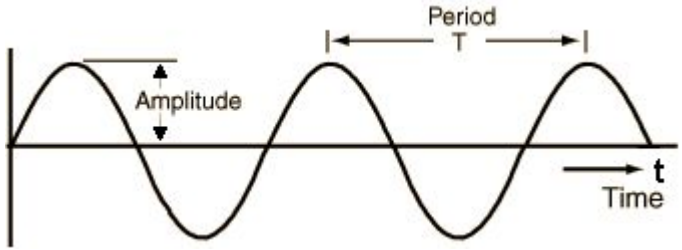
In physical world,

- Sound is generated when objects vibrate and cause air molecules to bump into others.
- Oscillation of the air molecules generates sound waves (aka mechanical waves)
- The sound wave travel through a medium and transfer energy



Some fundamental features are ...

- Period and Frequency
  - Period is the time to complete one cycle
  - Frequency is the number of cycle in one second
- Amplitude
  - Amplitude is the intensity of pressure



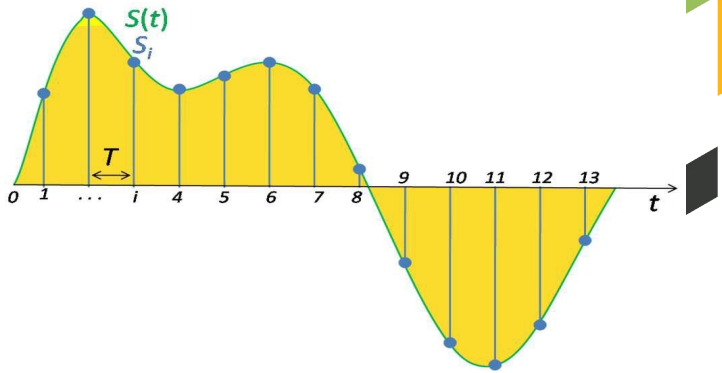
FROM PHYSICAL WORLD TO DIGITAL WORLD

Analog to Digital Conversion (ADC)

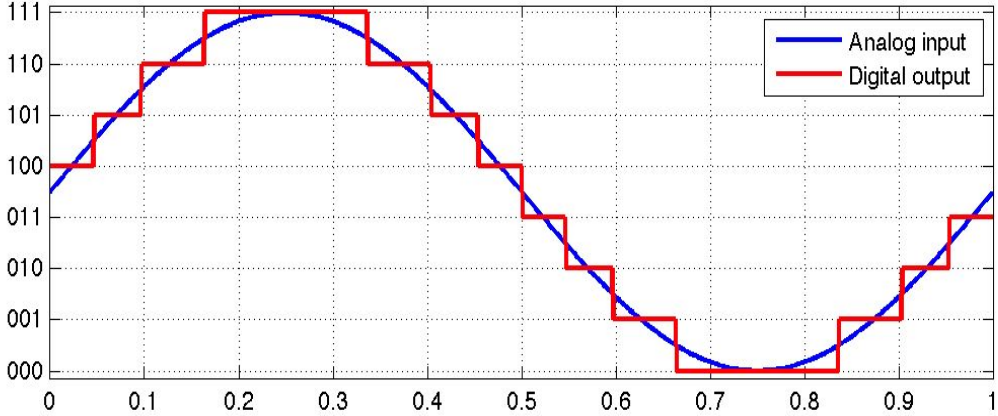
- All audio is found as an analog signal
  - Analog signal is a continuous graph of time
  - Amplitude of a sound contains infinite values
- Storing a raw analog signal would be nearly impossible, requiring infinite storage
- Analog to Digital Conversion contains:
  - Sampling
  - Quantization

Sampling vs Quantization

- Sampling:
  - Split x axis into period of time



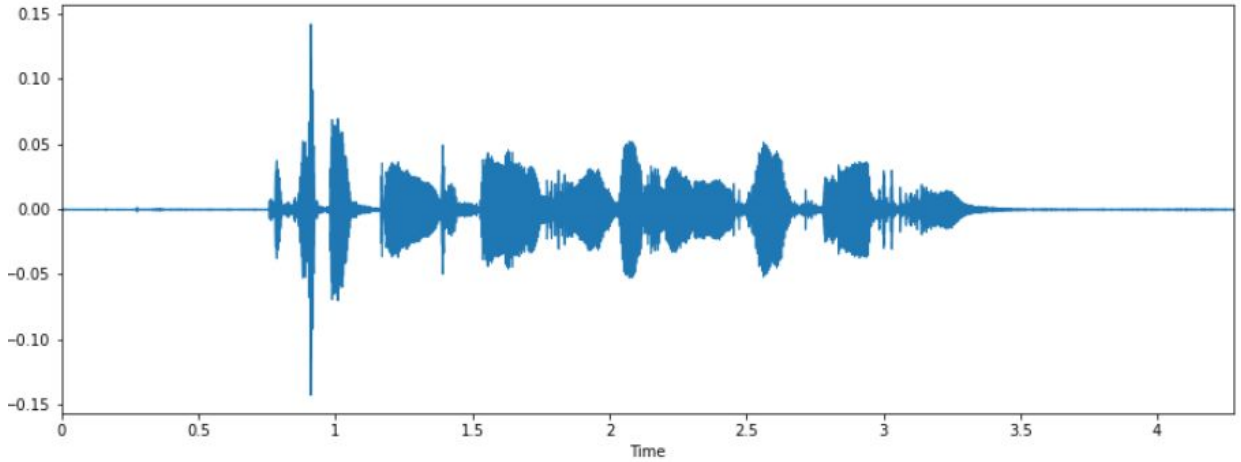
- Quantization
  - Split y axis into period of value



SPECTROGRAMS

Time domain

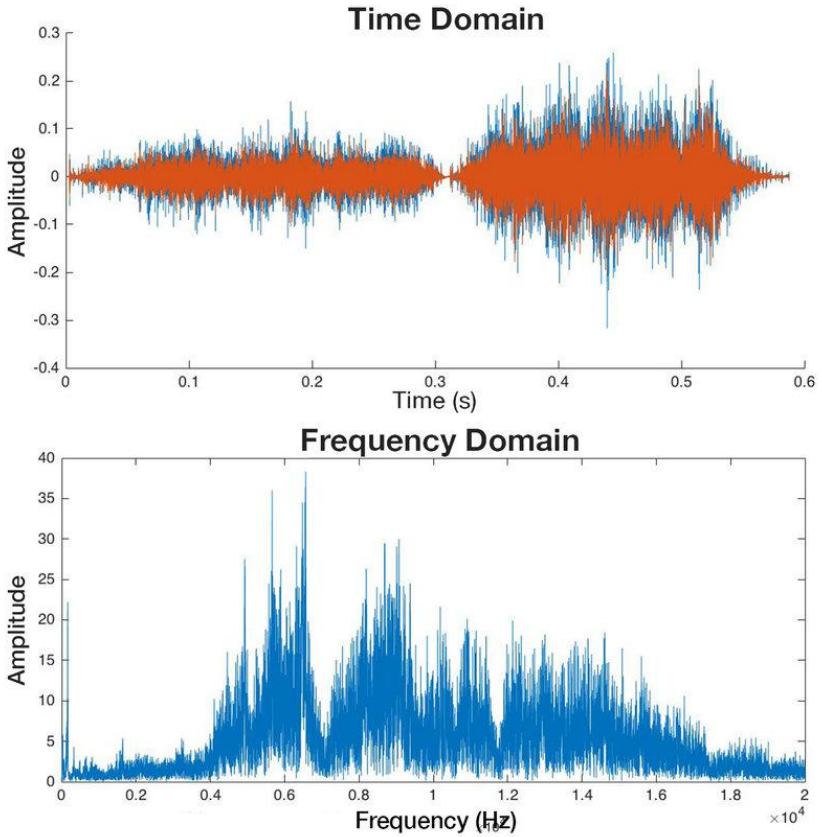
- Time domain show the relationship between
  - Time (x-axis)
  - Amplitude (y-axis)



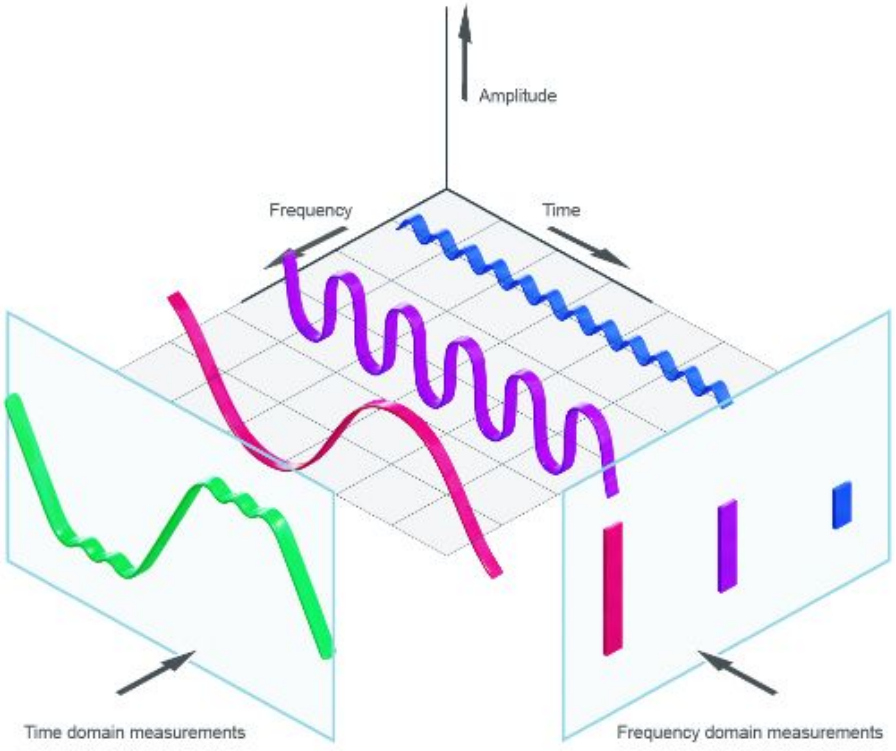
But, what's about Frequency?

Frequency domain

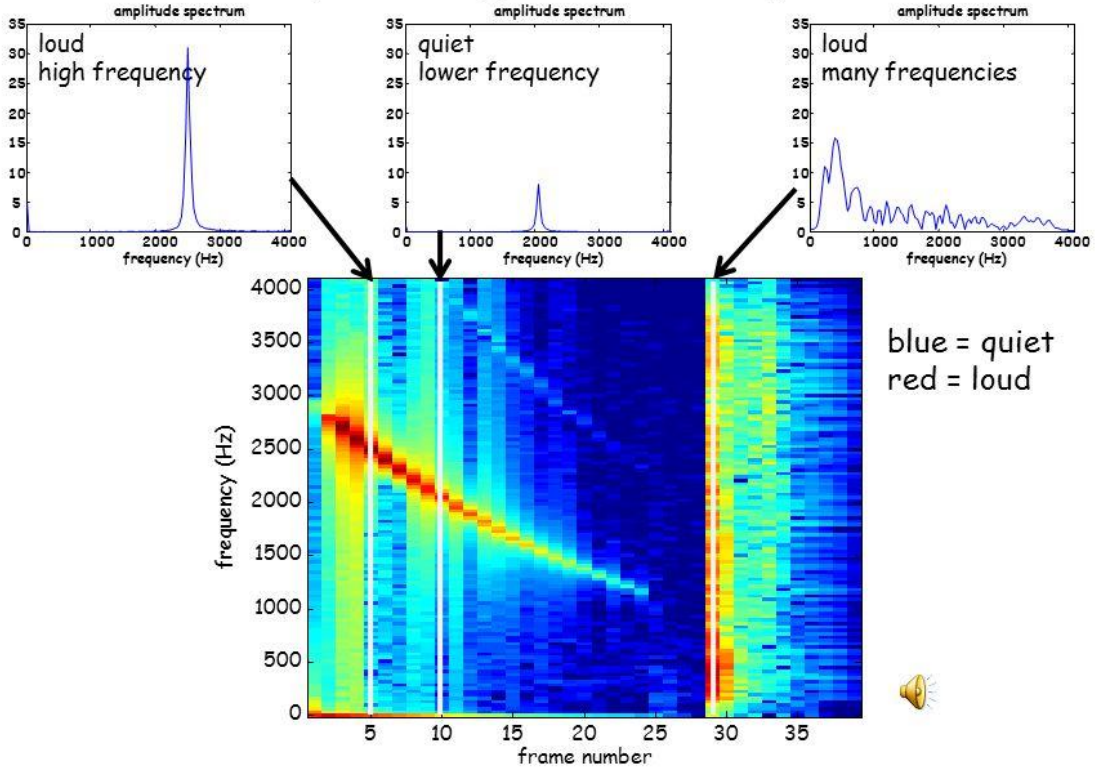
- Frequency domain show the relationship between
  - Frequency (x-axis)
  - Amplitude (y-axis)



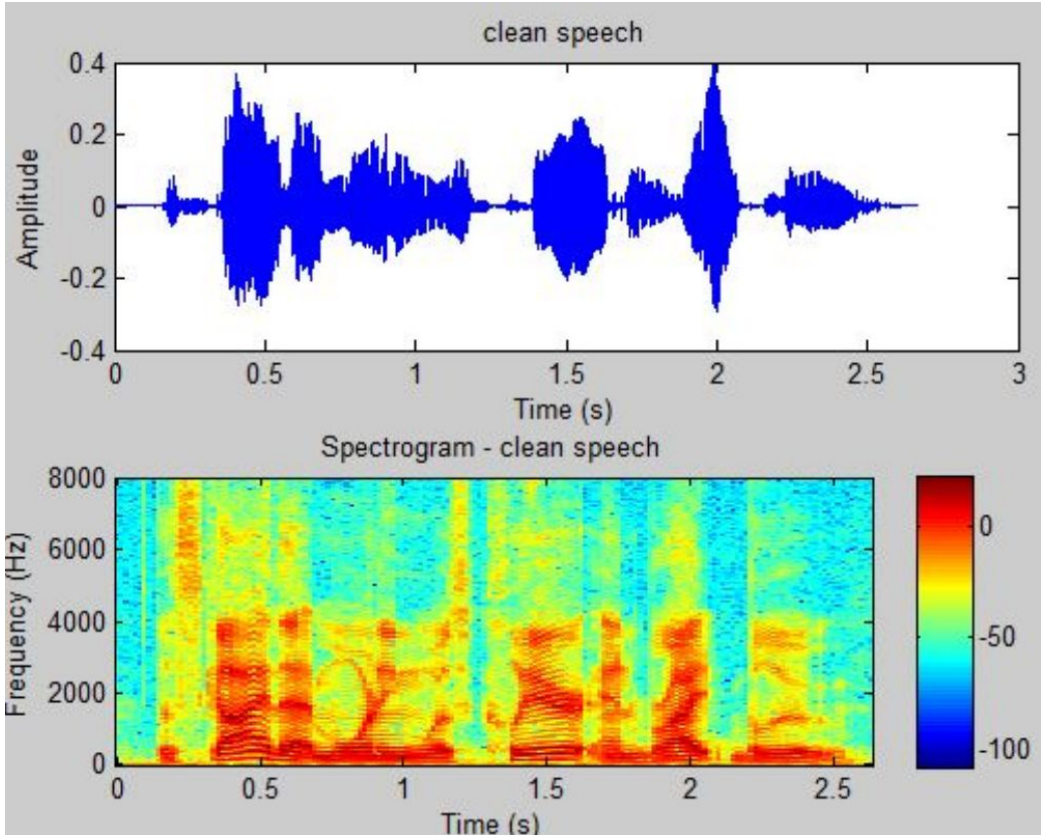
Time domain vs Frequency domain



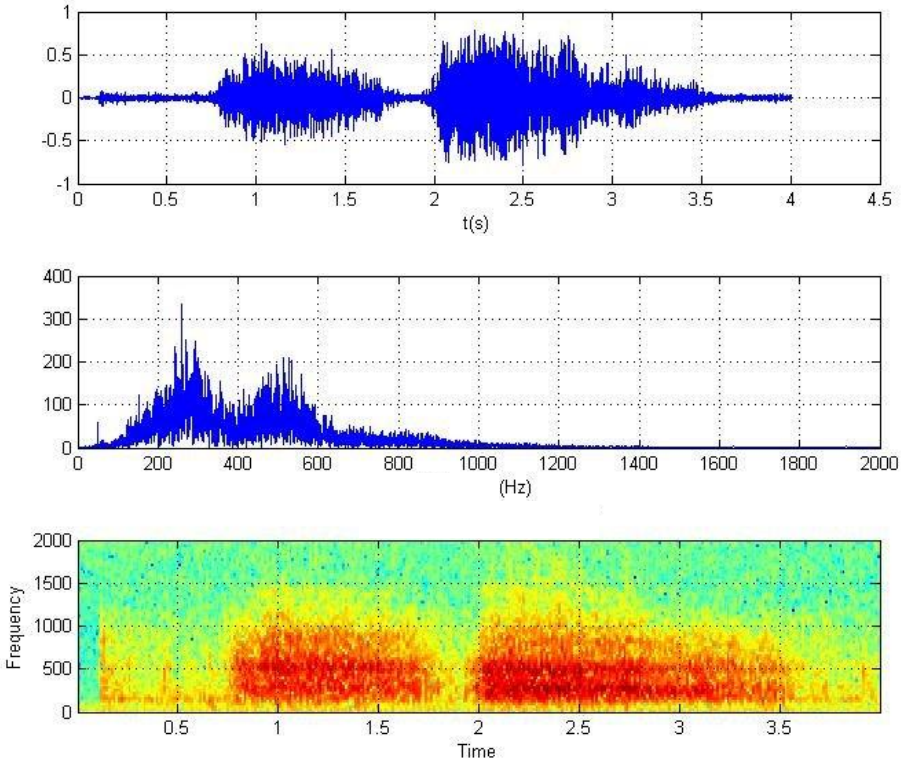
Spectrograms with Frequency



Spectrograms with Time



Spectrograms



EXAMPLE

Spectrograms example


PIXTA
VIETNAM