DSP Lecture 5

Speech Signal Processing

Tal Rosenwein

Lecture Overview

1. FIR & IIR Filters
2. Classification of Signals
3. Energy / RMS calculation

Lecture Goals

1. Learn basic concepts about audio processing
2. Learn basic concepts about speech signal processing

Lecture Overview

1. **FIR & IIR Filters**
2. Classification of Signals
3. Energy / RMS calculation

FIR & IIR Filters: FIR

Differential equation: $y[n] = b_0x[n] + b_1x[n-1] + b_2x[n-2] + ... + b_Nx[n-N]$

Transfer function (Z): $H(Z) = \frac{B(Z)}{Z^n} = b_0 + b_1Z^{-1} + b_2Z^{-2} + ... + b_NZ^{-N}$

Impulse response: $h[0] = b_0,\ h[1] = b_1,\ h[2] = b_2,\ ...,\ h[N] = b_N,\ h[N+1] = 0, ...$

- **Finite Impulse Response - h[n] is finite**
- **FIR filters are always stable (by definition- all poles are at the origin, z=0)**
- **Also known as ‘moving average filters’ (MA), or ‘all zeros’ filters**
- **Pros:**
  - **Stable**
  - **Linear phase**
  - **Easy to implement and construct using closed-form formulas**
- **Cons:**
  - **In order to get a sharp response, large order is required.**
  - **Larger delay compared to IIR**

FIR & IIR Filters: IIR

Differential equation: $y[n] + a_1 y[n-1] + ... + a_N y[n-N] = b_0 x[n] + b_1 x[n-1] + ... + b_N x[n-N]$

Transfer function (Z): $H(Z) = \frac{B(Z)}{A(Z)} = \frac{b_0 + b_1 Z^{-1} + b_2 Z^{-2} + ... + b_N Z^{-N}}{1 + a_1 Z^{-1} + a_2 Z^{-2} + ... + a_N Z^{-N}}$

Differential equation: $y[n] = \underbrace{b_0 x[n] + b_1 x[n-1] + ... + b_N x[n-N]}_{'MA'} \underbrace{- a_1 y[n-1] - a_2 y[n-2] - ... - a_N y[n-N]}_{'AR'}$

- **Infinite Impulse Response - h[n] is infinite**
- **IIR filters iff all poles lie inside the unit circle**
- **Also known as ‘AutoRregressive Moving Average filters’ (ARMA)**
- **Pros:**
  - **Obtain sharper frequency response compared to same-order FIR filters (by moving the poles from origin)- lower computational complexity**
  - **Lower latency compared to FIR**
- **Cons:**
  - **May cause phase issues**
  - **Harder to implement (no-close form in some cases).**
  - **Overflow may be an issue due to recursiveness (infinite) of the response.**

Lecture Overview

1. FIR & IIR Filters
2. **Classification of Signals**
3. Energy / RMS calculation

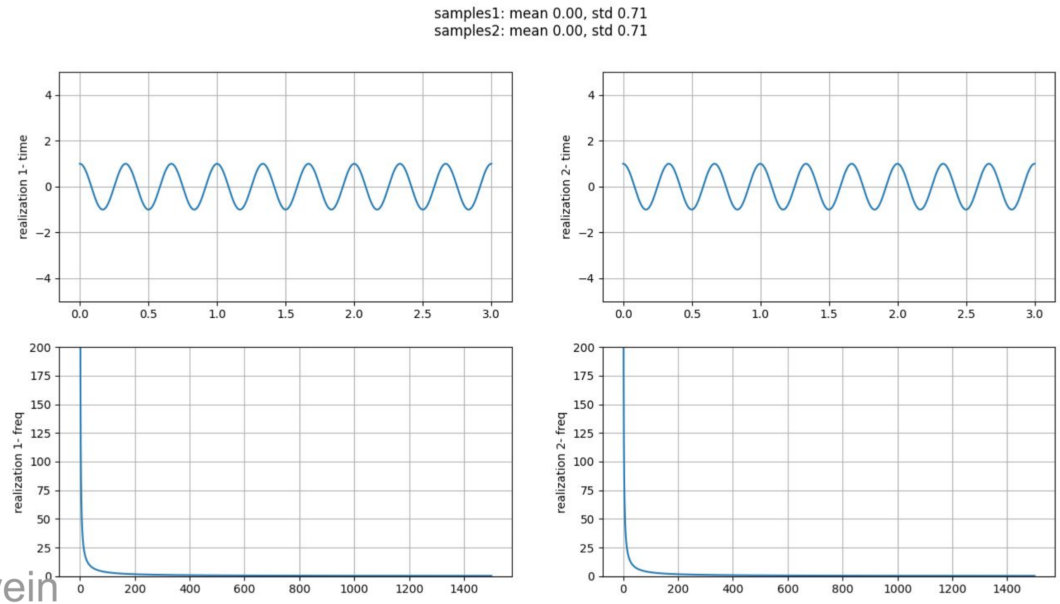
Classification of Signals

Signals can be **classified** in various ways as we saw in previous lectures**:**

- Continuous vs discrete
- Periodic vs Aperiodic

We will now discuss on two classes:

- Probabilistic Vs Deterministic.



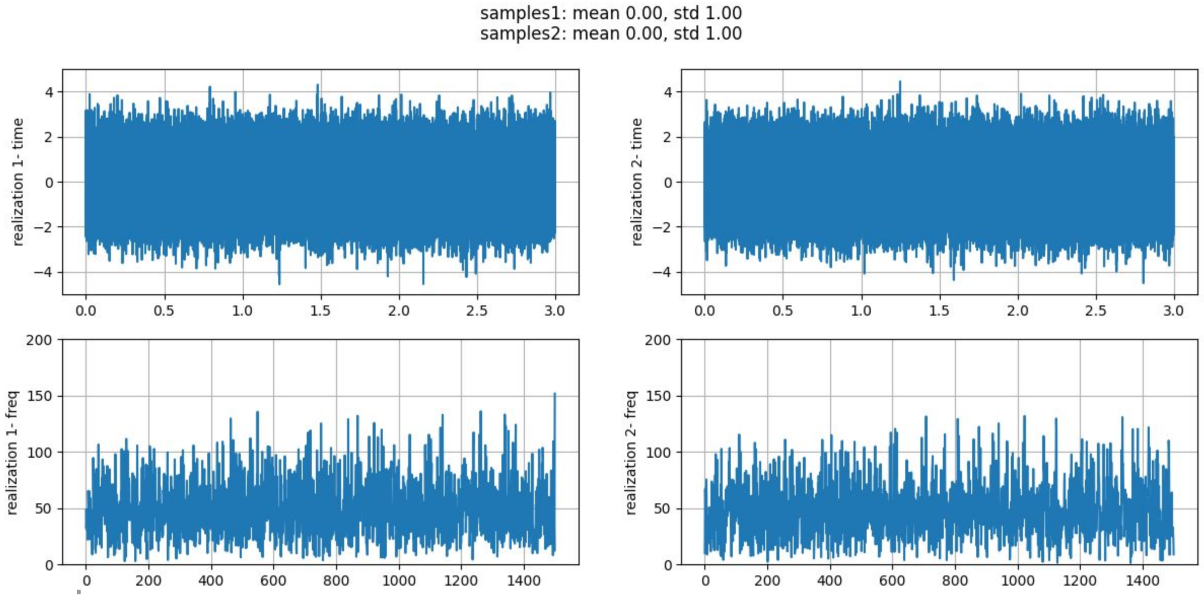
Classification of Signals

Signals can be **classified** in various ways as we saw in pervious lectures**:**

- Continuous vs discrete
- Periodic vs Aperiodic

We will now discuss on two classes:

- Probabilistic Vs Deterministic.



Classification of Signals

Signals can be **classified** in various ways as we saw in pervious lectures**:**

- Continuous vs discrete
- Periodic vs Aperiodic

We will now discuss on two classes:

- Probabilistic Vs Deterministic.
- Stationary Vs. non-stationary.

Stationarity is a way of **describing the characteristics of the signal generating process**, which further gives us two categories.

The difference between **stationary and non-stationary signals** is that the **statistical** (joint probability distribution) **properties of a stationary process signal do not change with time**. Consequently, parameters such as **mean and variance** also **do not change over time**.

So, a continuous time random process $\{X_t\}$ which is WSS has the following restrictions on its mean function $m_X(t) \triangleq \mathrm{E}[X_t]$ and autocovariance function $K_{XX}(t_1,t_2) \triangleq \mathrm{E}[(X_{t_1} - m_X(t_1))(X_{t_2} - m_X(t_2))]$:

$$\begin{array}{ll} m_X(t) = m_X(t+\tau) & \text{for all } \tau \in \mathbb{R} \\ K_{XX}(t_1,t_2) = K_{XX}(t_1 - t_2, 0) & \text{for all } t_1,t_2 \in \mathbb{R} \\ \mathrm{E}[|X(t)|^2] < \infty & \text{for all } t \in \mathbb{R} \end{array} \quad \textbf{(Eq.3)}$$

Signals

Stationary:

- White noise (probabilistic & stationary)
- Sin / cos (deterministic & stationary)

Non-Stationary:

- Speech (probabilistic & non-stationary)

Parameter of Comparison	Stationary Signals	Non-Stationary
Time	The time period for the stationary signal remains constant at all times.	The time period for a non-stationary signal varies with time and is not constant.
Frequency	The frequency of a stationary signal remains constant across the process	The frequency of a Non-stationary wave changes constantly during the process.
Spectral Contents	Spectral content for Stationary signals are constant	Spectral contents are dynamic and keep changing in case of the non-stationary signal.
Fourier Equation	Fourier transform is good at representing stationary signals	Fourier transform is non-good at representing non-stationary signals.
Examples	Single-tone sinewave constant frequency, Multitone sine wave of constant frequency	Speech signals, Multitone sinewave of varied frequency

Source 1, 2

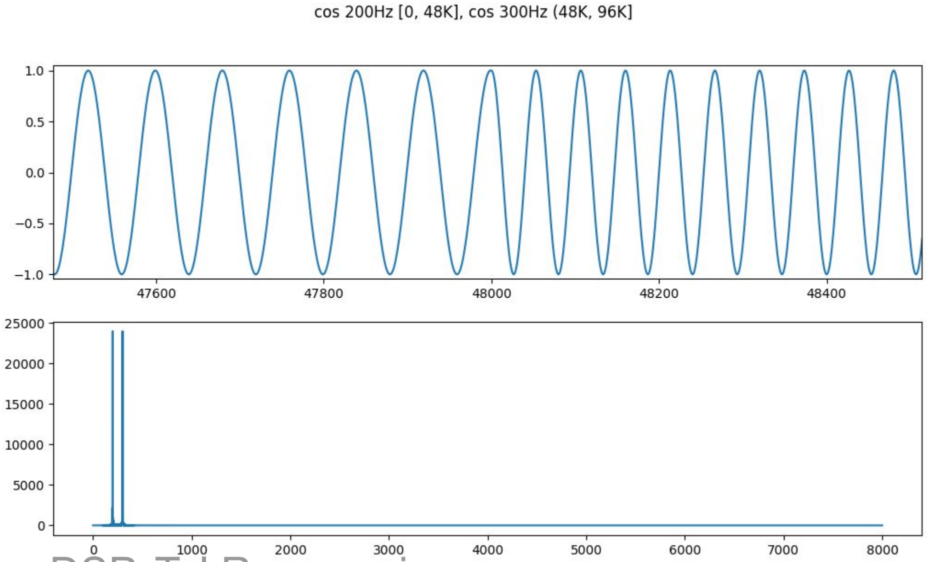
Classification of Signals

Signals can be **classified** in various ways as we saw in pervious lectures**:**

- Continuous vs discrete
- Periodic vs Aperiodic

We will now discuss on two classes:

- Probabilistic Vs Deterministic.
- Stationary Vs. non-stationary.



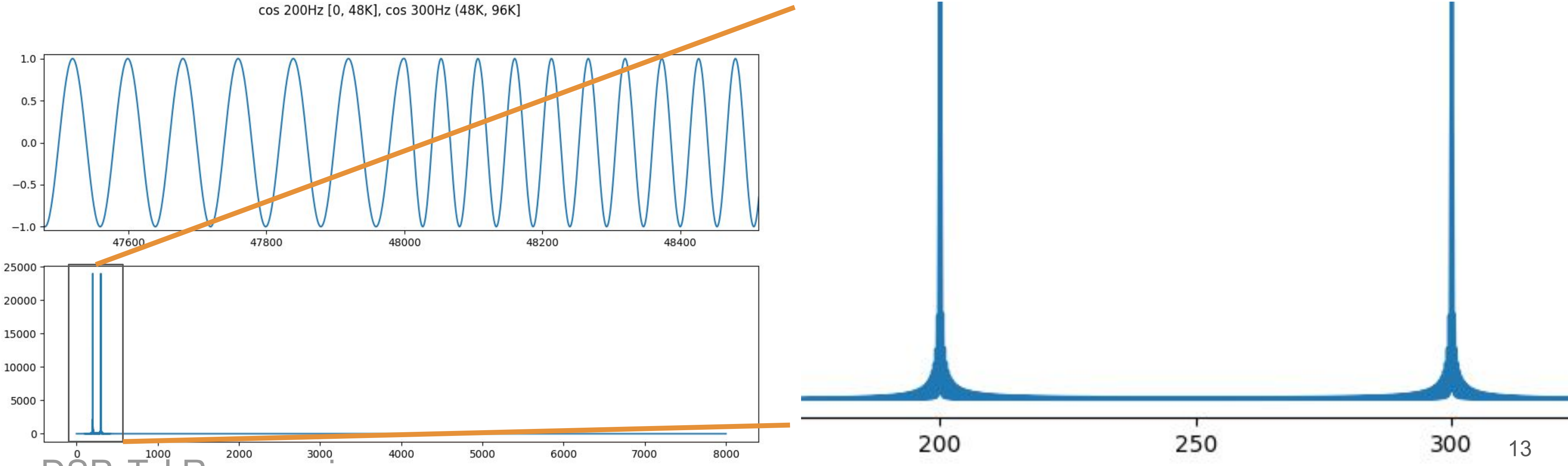
Classification of Signals

Signals can be **classified** in various ways as we saw in pervious lectures**:**

- Continuous vs discrete
- Periodic vs Aperiodic

We will now discuss on two classes:

- Probabilistic Vs Deterministic.
- Stationary Vs. non-stationary.



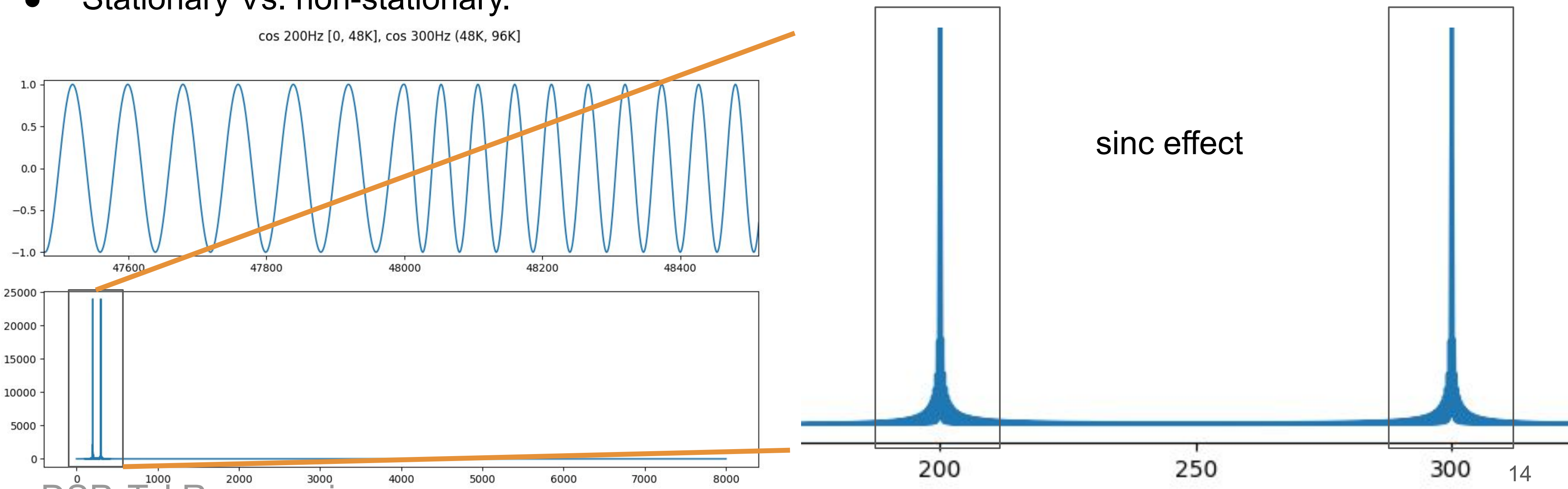
Classification of Signals

Signals can be **classified** in various ways as we saw in pervious lectures**:**

- Continuous vs discrete
- Periodic vs Aperiodic

We will now discuss on two classes:

- Probabilistic Vs Deterministic.
- Stationary Vs. non-stationary.



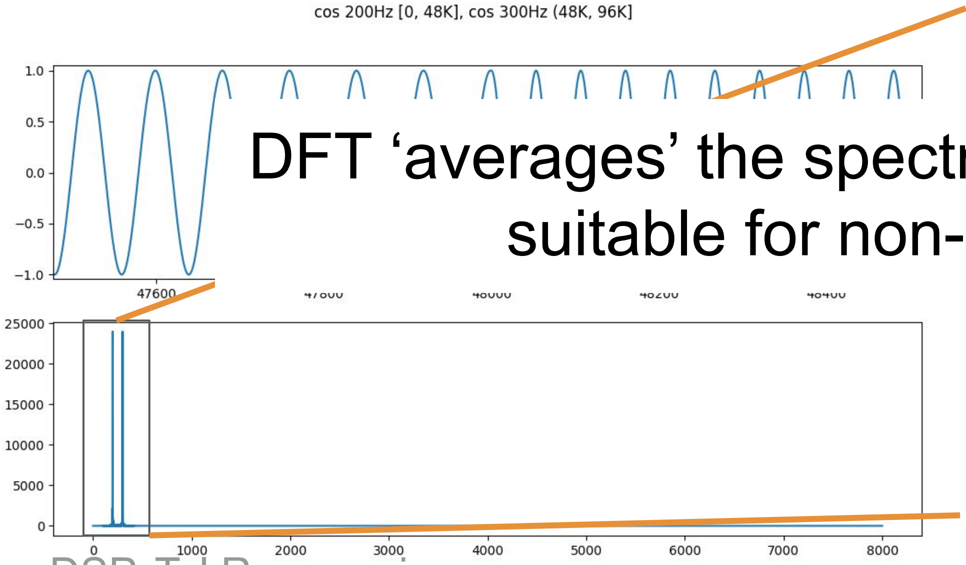
Classification of Signals

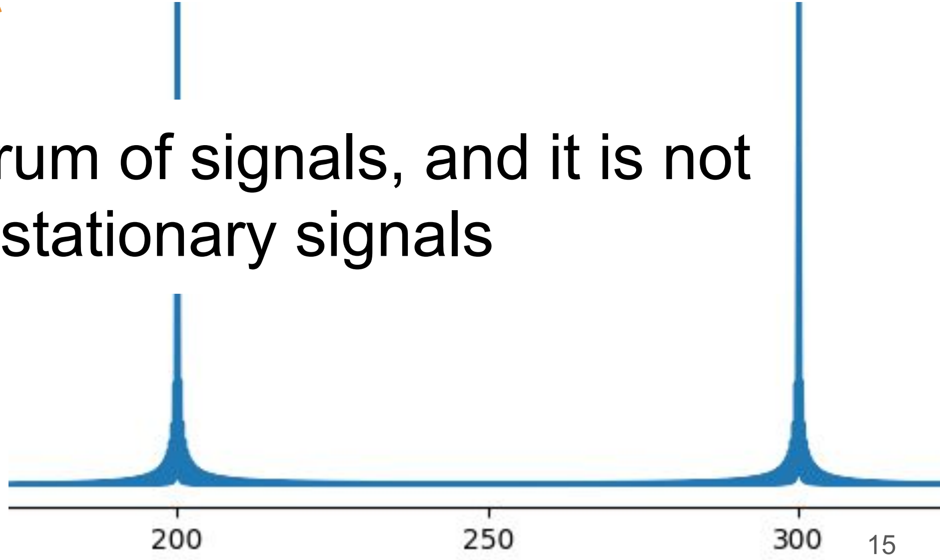
Signals can be **classified** in various ways as we saw in pervious lectures:

- Continuous vs discrete
- Periodic vs Aperiodic

We will now discuss on two classes:

- Probabilistic Vs Deterministic.
- Stationary Vs. non-stationary.





DFT 'averages' the spectrum of signals, and it is not suitable for non-stationary signals

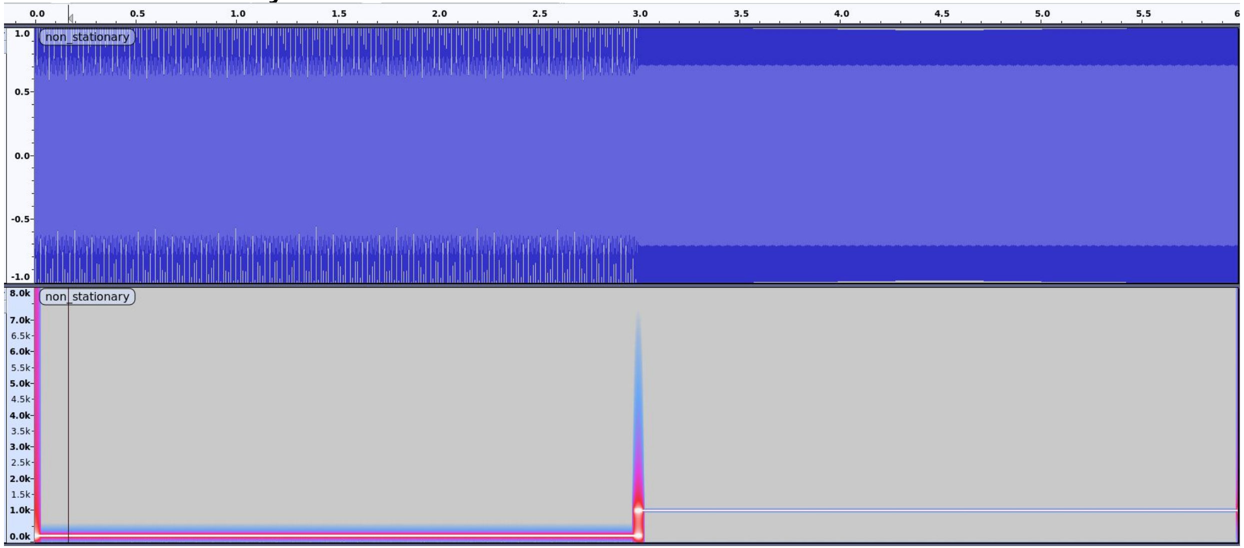
Classification of Signals

Signals can be **classified** in various ways as we saw in pervious lectures**:**

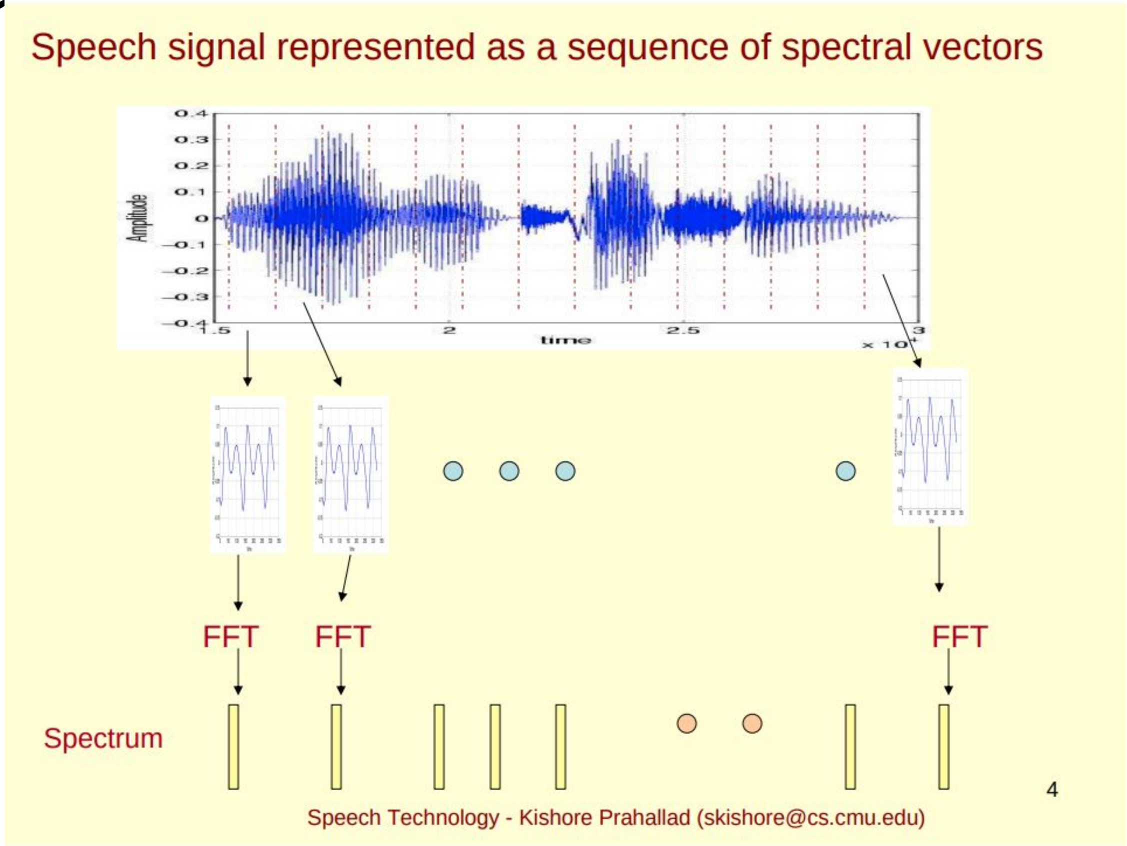
- Continuous vs discrete
- Periodic vs Aperiodic

We will now discuss on two classes:

- Probabilistic Vs Deterministic.
- Stationary Vs. non-stationary.

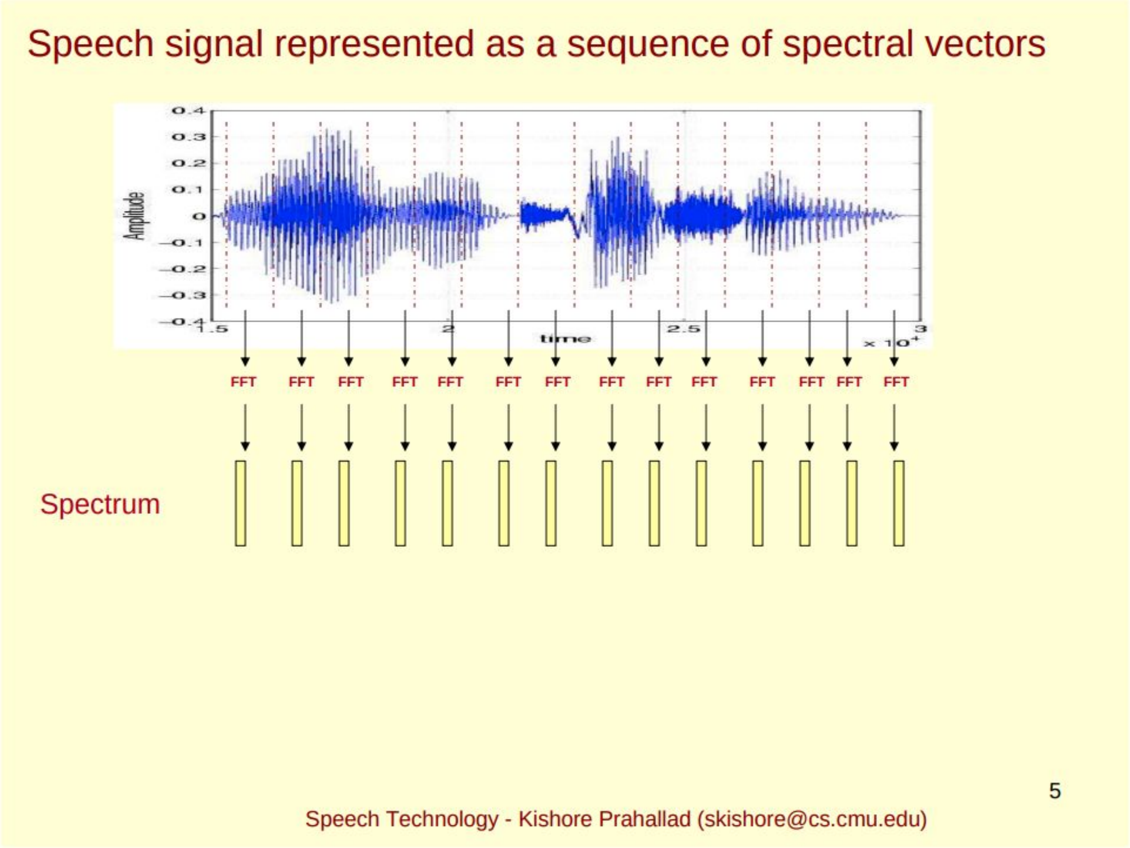


Classification of Signals



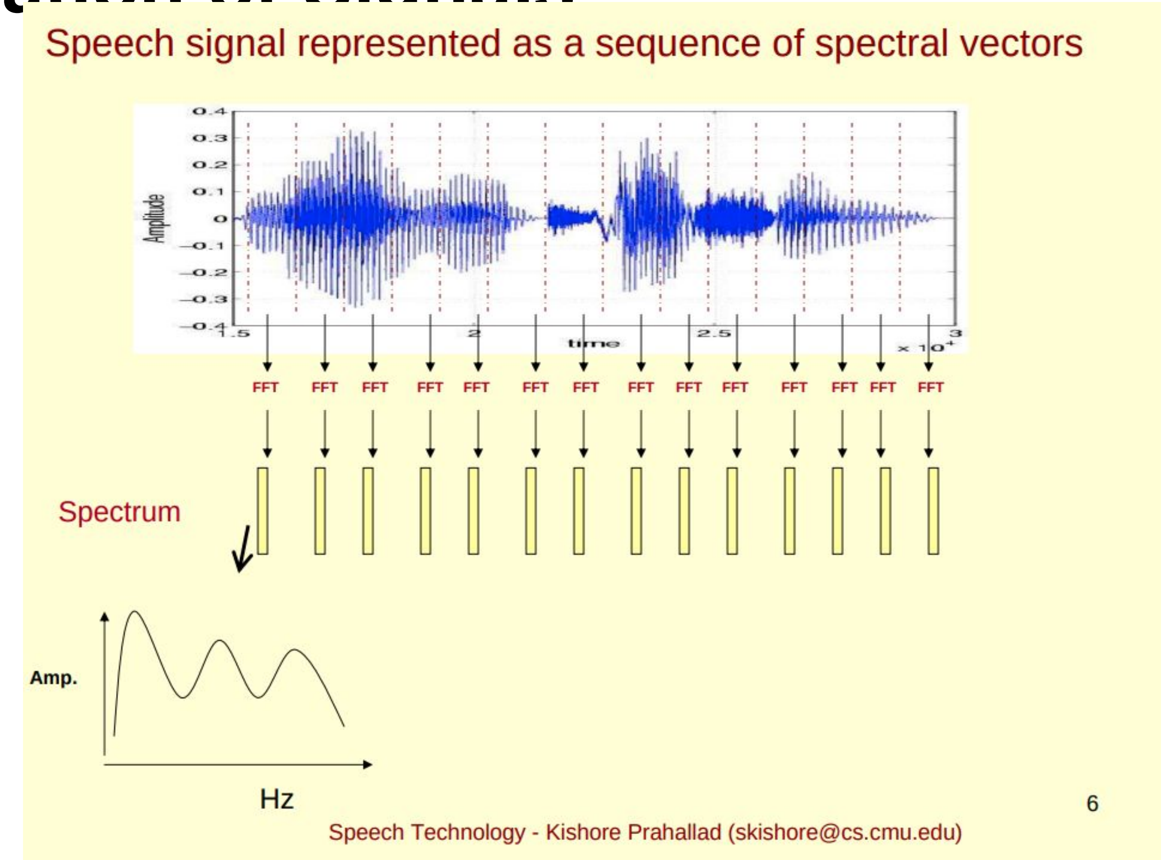
Slide taken from Kishore Prahallad

Classification of Signals



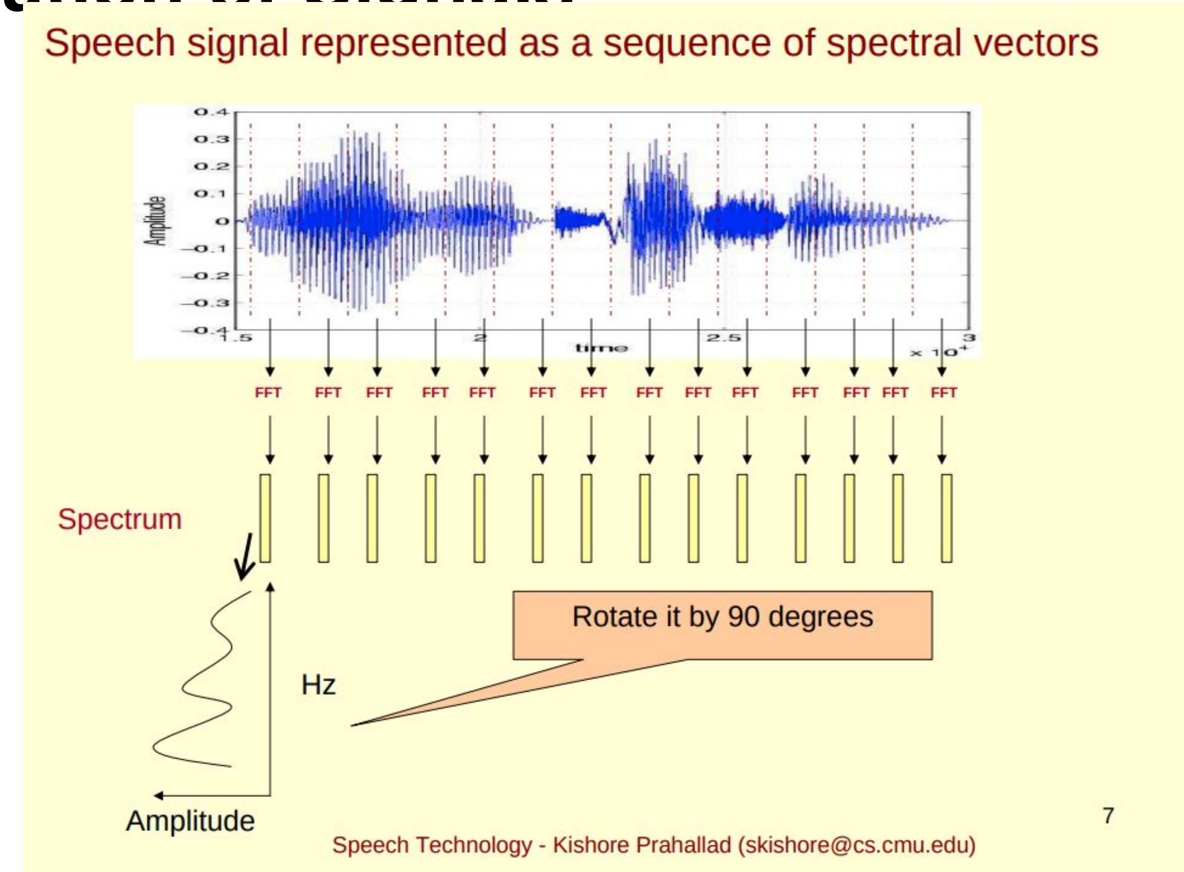
Slide taken from [Kishore Prahallad]

Classification of Signals

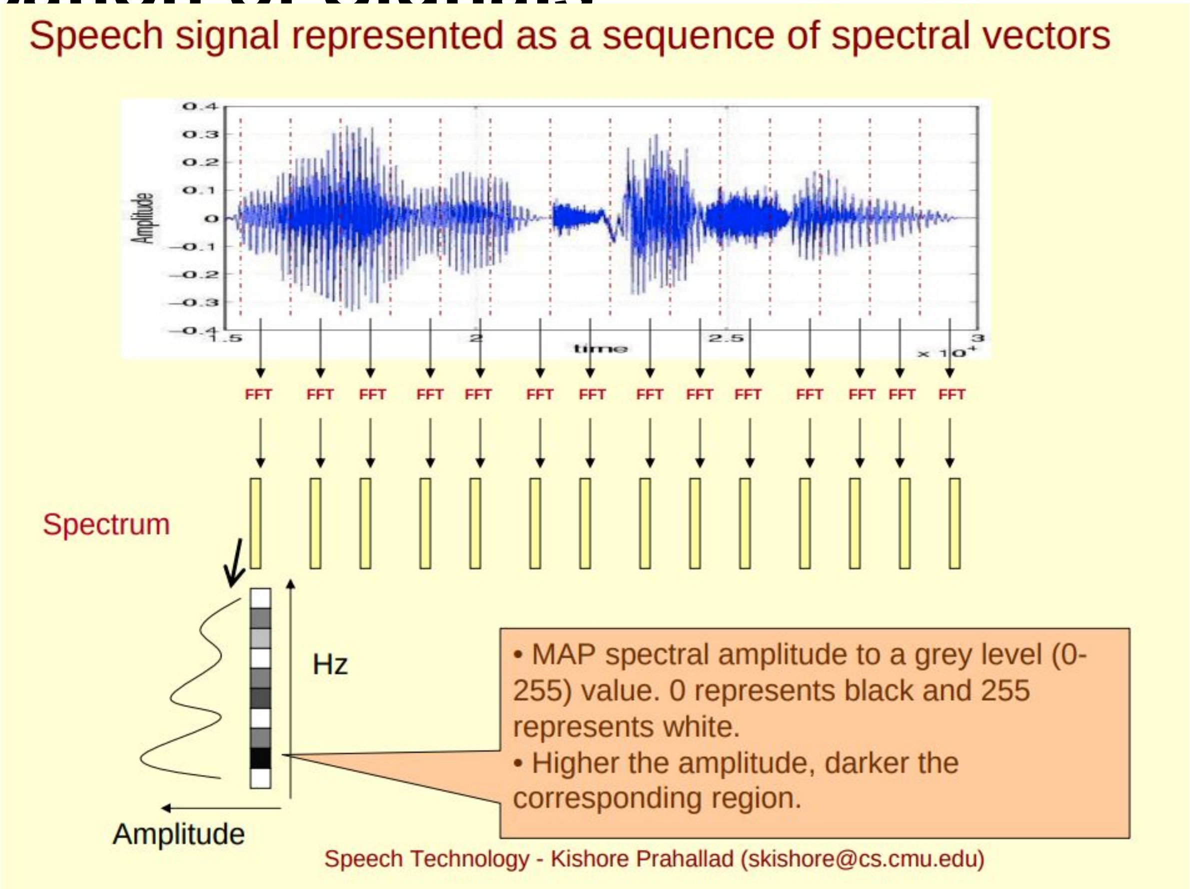


Slide taken from [Kishore Prahallad]

Classification of Signals

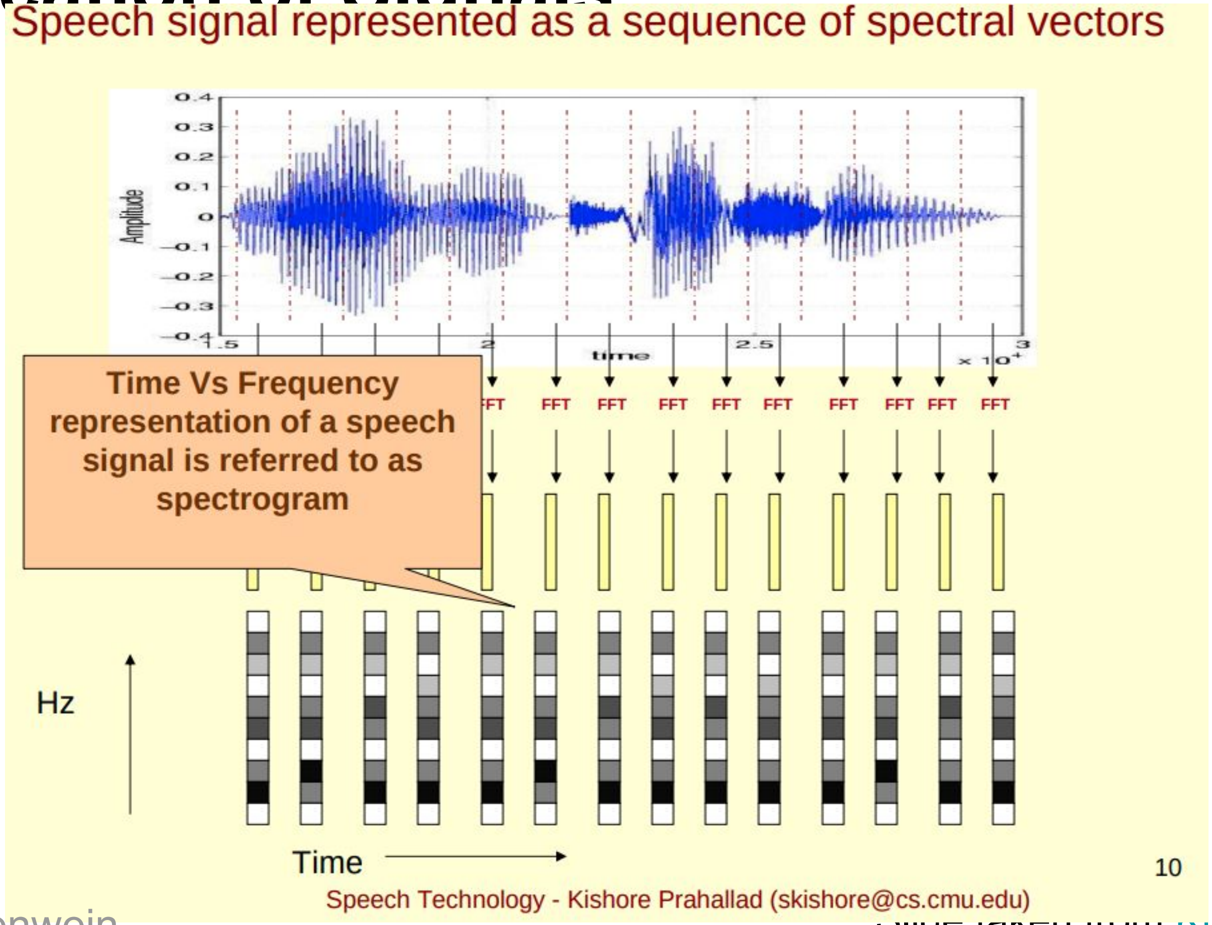


Classification of Signals



Slide taken from [Kishore Prahallad]

Classification of Signals

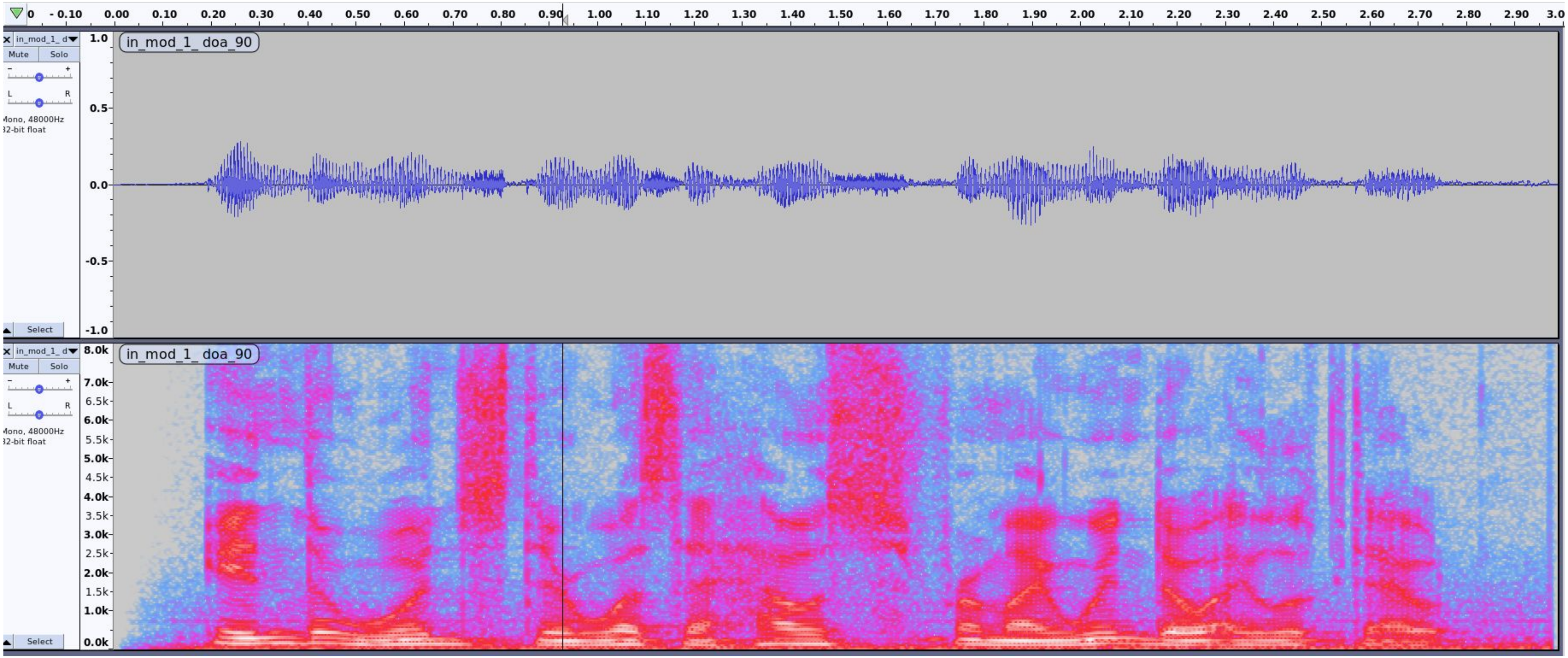


Slide taken from Kishore Prahallad

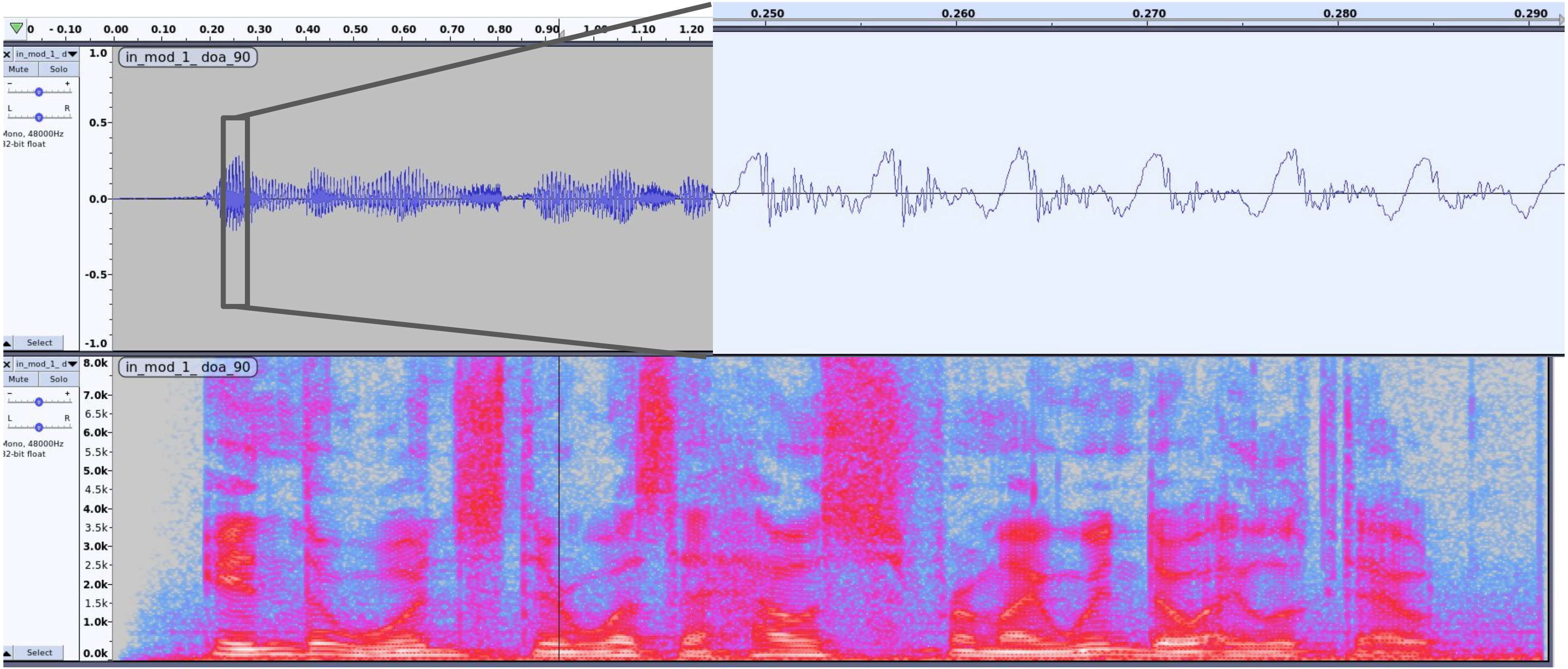
Classification of Signals

Temporal / Frequency Resolution Tradeoff

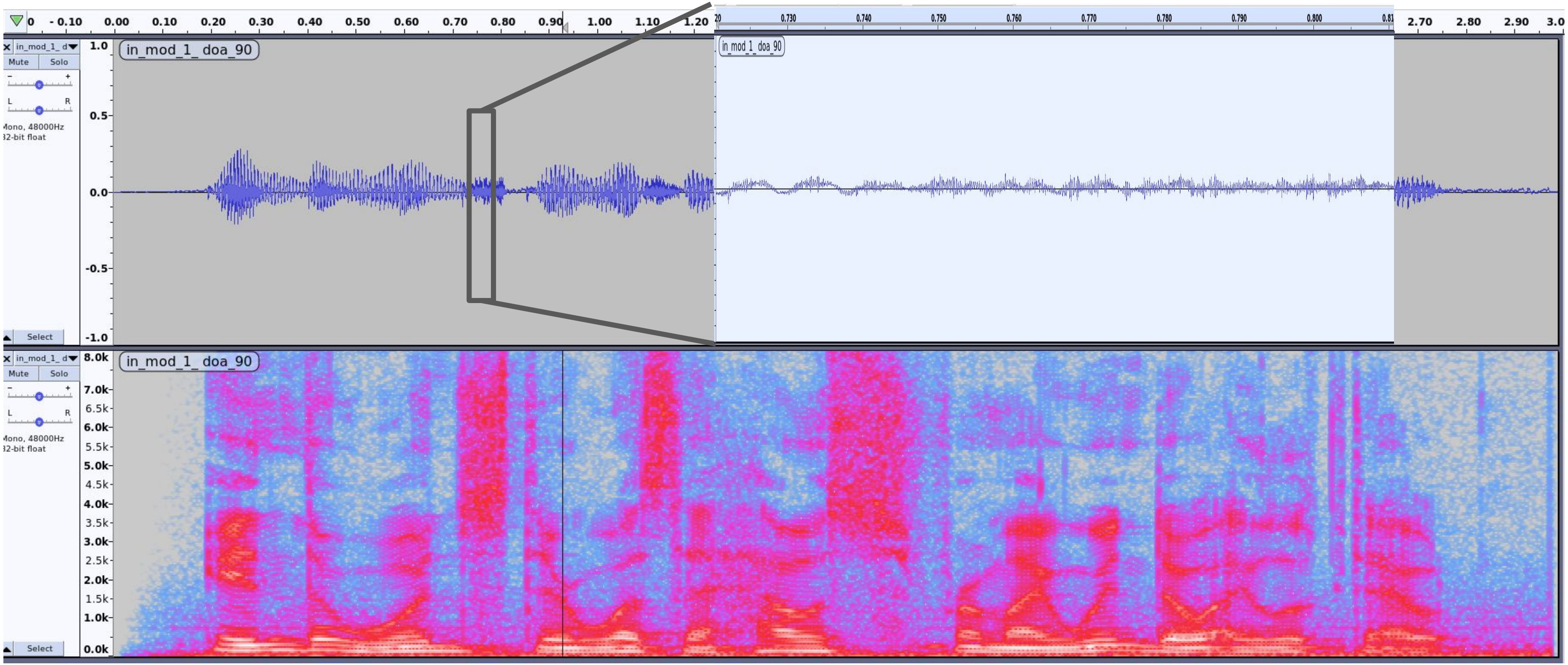
Classification of Signals



Classification of Signals



Classification of Signals



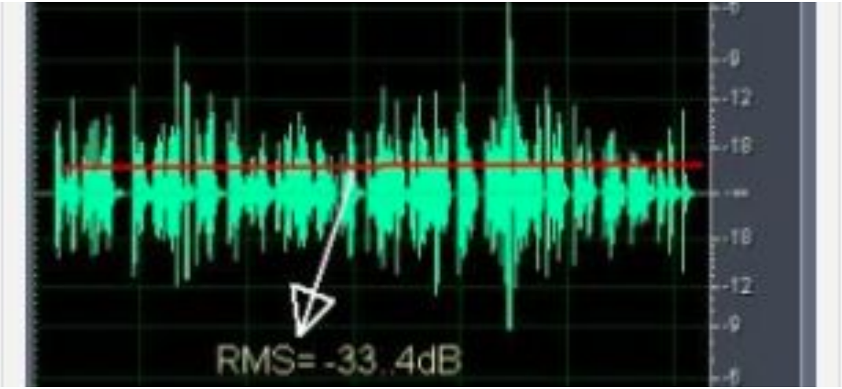
Lecture Overview

1. FIR & IIR Filters
2. Classification of Signals
3. **Energy / RMS calculation**

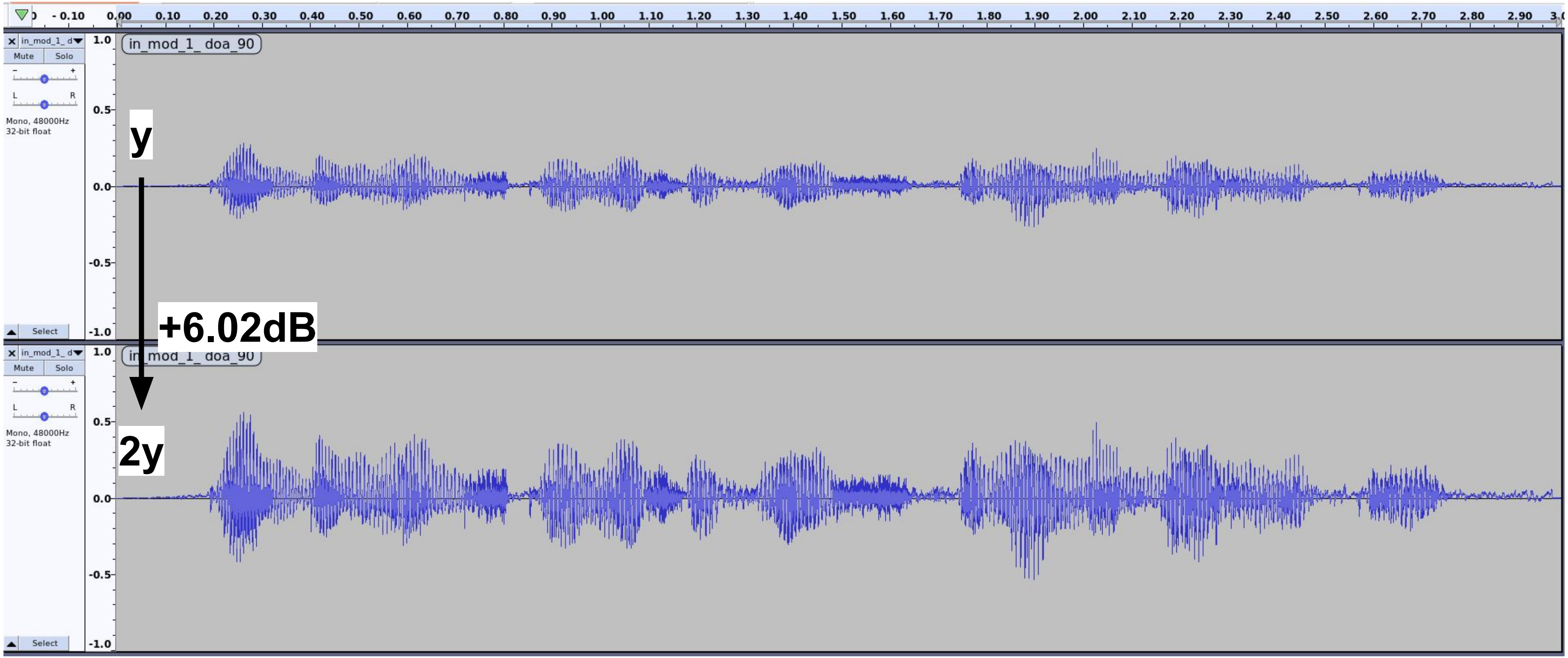
Energy & RMS

- Sometimes we want to represent some “trends” of signal, but its average has a (zero) constant value.
- 2 accepted measures for an ‘alternative’ average: RMS and Energy
  - RMS: Root-Mean-Squared. Most books define this as the “amount of AC power that produces the same heating effect as an equivalent DC power”
  - Energy: defined as the area under the squared magnitude of the considered signal
- Usually calculated in chunks with hop.

$$E_s = \sum_{n=-\infty}^{\infty} |x(n)|^2 \qquad RMS = \sqrt{\frac{1}{n}\sum_i x_i^2}$$

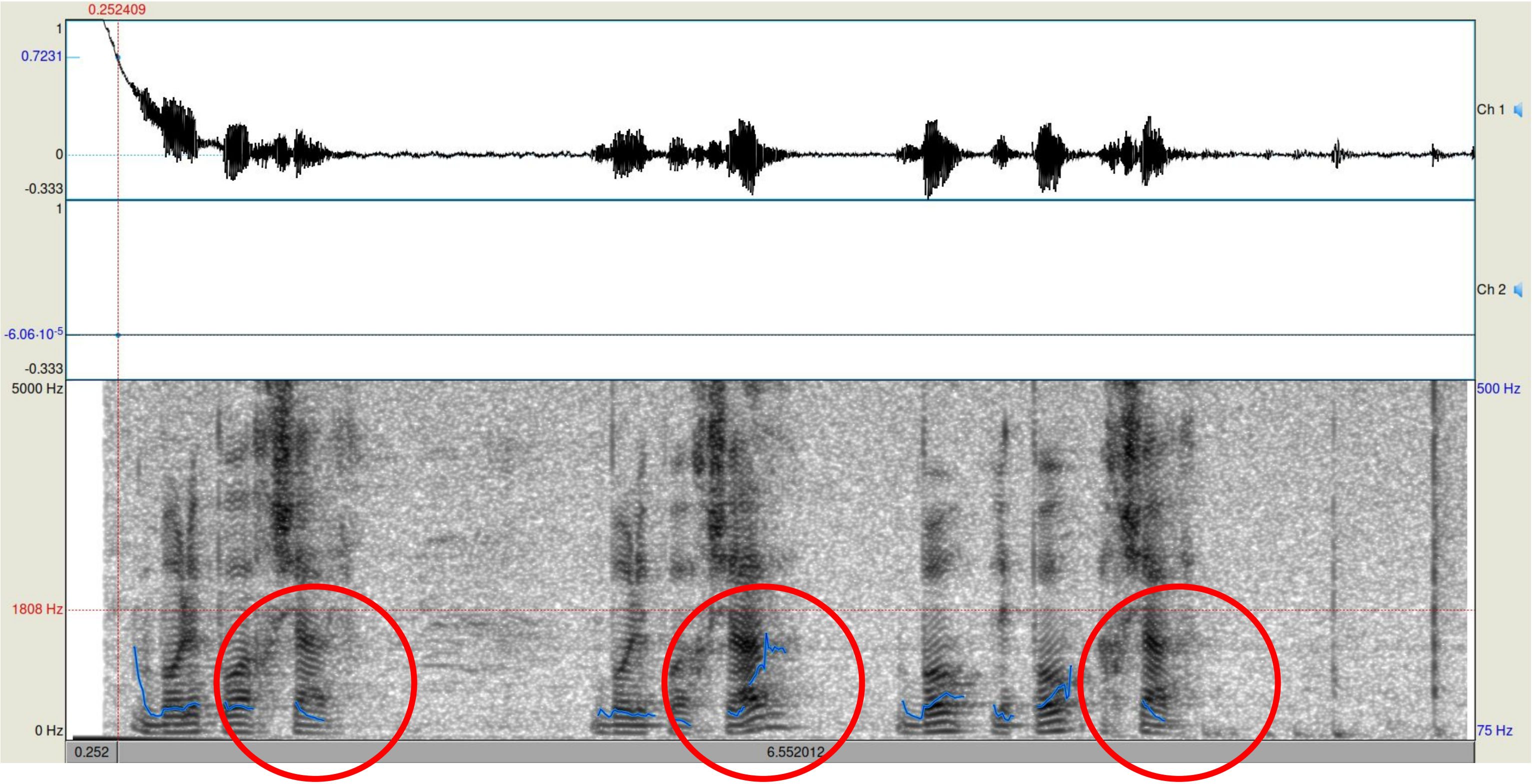


Energy & RMS



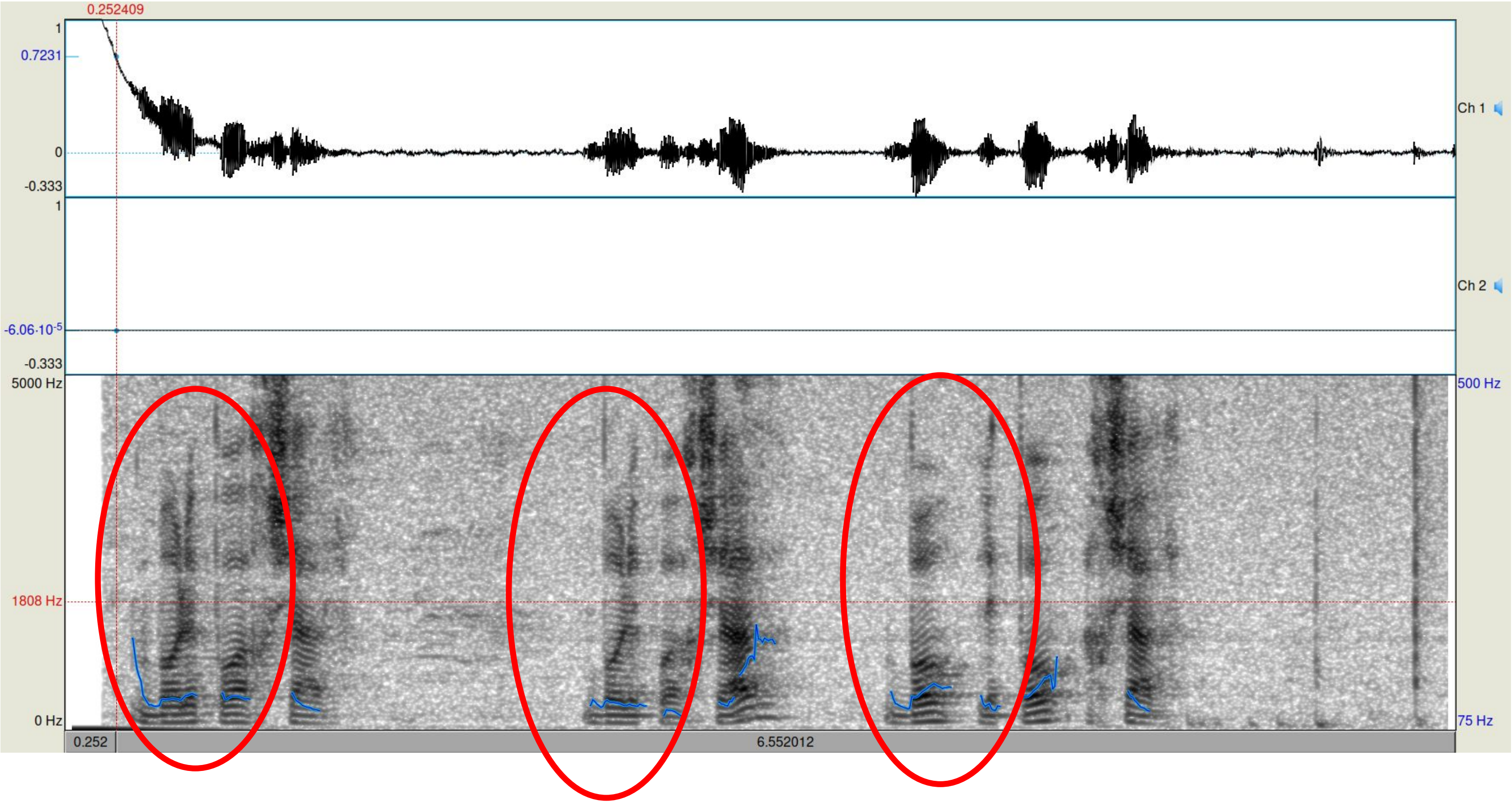
in_mod_1_doa_90
y
+6.02dB
2y
in_mod_1_doa_90
Mono, 48000Hz
32-bit float

Energy & RMS



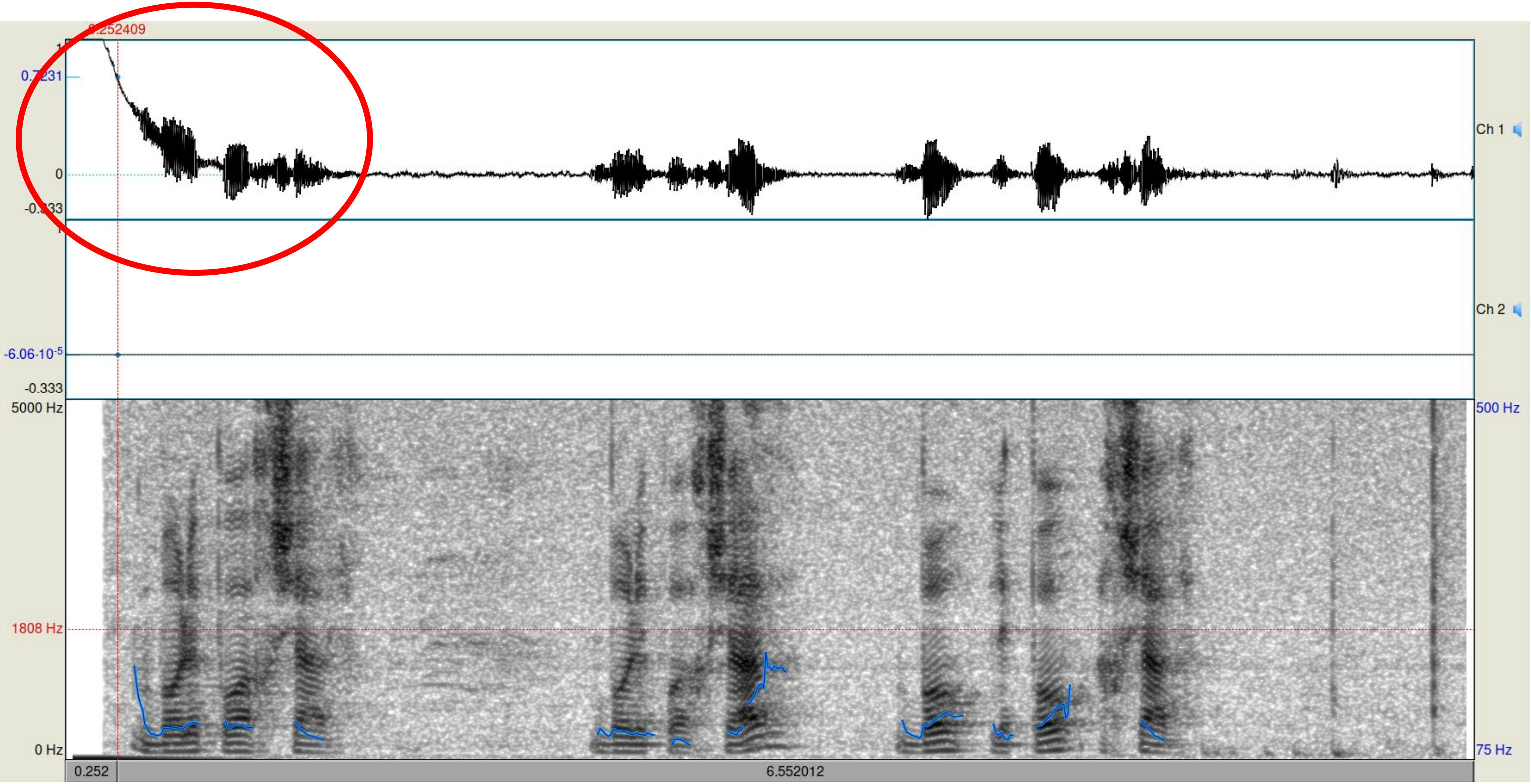
0.252409
1
0.7231
0
-0.333
1
-6.06·10-5
-0.333
Ch 1
Ch 2
5000 Hz
1808 Hz
0 Hz
500 Hz
75 Hz
0.252
6.552012

Energy & RMS

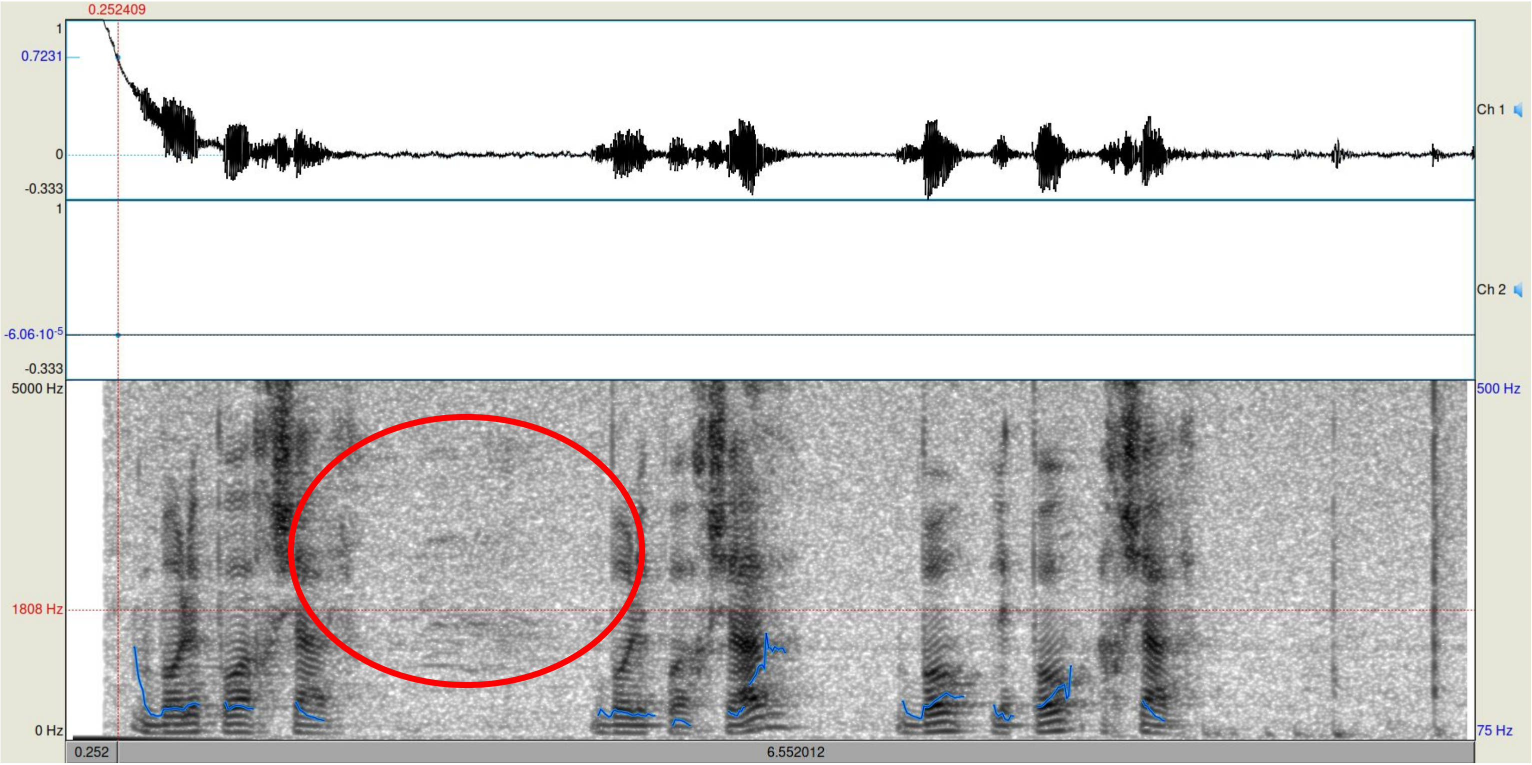


0.252409
1
0.7231
0
-0.333
Ch 1
1
Ch 2
-6.06·10-5
-0.333
5000 Hz
500 Hz
1808 Hz
0 Hz
75 Hz
0.252
6.552012

Energy & RMS


0.252409
1
0.7231
0
-0.333
Ch 1
Ch 2
-6.06·10-5
-0.333
5000 Hz
1808 Hz
0 Hz
500 Hz
75 Hz
0.252
6.552012

Energy & RMS


0.252409
1
0.7231
0
-0.333
1
-6.06·10-5
-0.333
Ch 1
Ch 2
5000 Hz
1808 Hz
0 Hz
500 Hz
75 Hz
0.252
6.552012

Code

```python
def show_spectrogram(fpath: str,
                     win_length_sec: float = 0.025,
                     hop_length_sec: float = 0.01,
                     ):
    # Load the audio file
    audio, sr = librosa.load(fpath, sr=48000)

    # Compute the STFT
    win_length_samples = int(win_length_sec * sr)
    hop_length_samples = int(hop_length_sec * sr)
    n_fft = win_length_samples
    stft = librosa.stft(audio,
                        n_fft=n_fft,
                        win_length=win_length_samples,
                        hop_length=hop_length_samples)
    stft_db = librosa.amplitude_to_db(abs(stft), ref=np.max)

    # Display the STFT
    librosa.display.specshow(stft_db, sr=sr, y_axis='linear', x_axis='time')
    plt.colorbar(format='%+2.0f dB')
    plt.xlabel('Time (s)')  # Custom x-axis label
    plt.ylabel('Frequency (Hz)')  # Custom y-axis label
    plt.title('STFT of Audio Signal')
    plt.show()
```