



**FPT SCHOOL OF BUSINESS  
& TECHNOLOGY**

# Digital Signal Processing

Các thuộc tính của tín hiệu Audio

Phd. Trần Thanh Trúc

# M3 - Lý thuyết: Các thuộc tính của tín hiệu Audio

# Các thuộc tính của tín hiệu Audio

## Nội dung

---

- Sóng âm thanh
- Các đặc tính cơ bản của audio:
  - Rms, spectral centroid
- Chuyển đổi audio:
  - Chuyển đổi Fourier và STFT
  - Chuyển đổi Spectrograms
- Thư viện cần dùng: librosa

# Các thuộc tính của tín hiệu Audio

## Mục tiêu

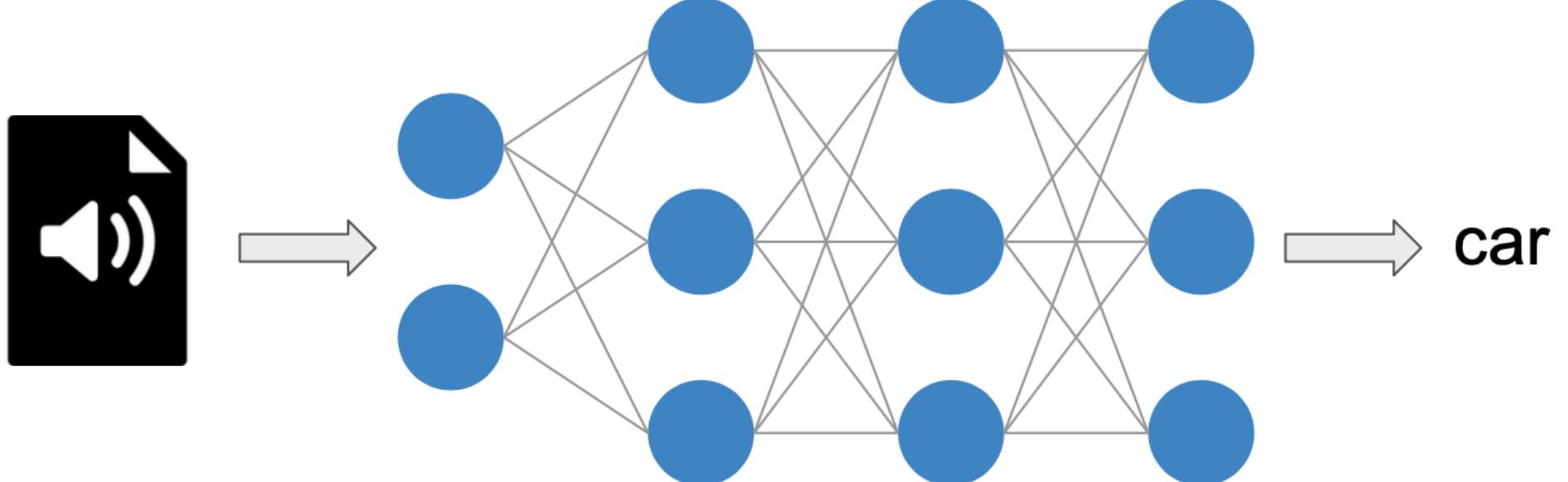
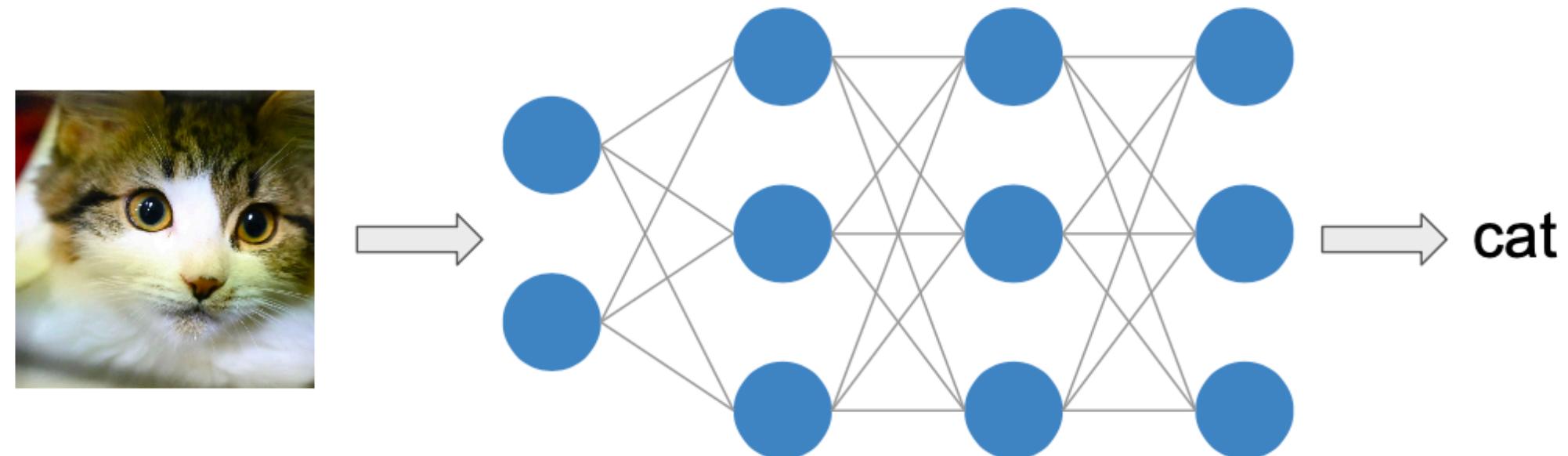
---

- Hiểu biết sâu hơn về dữ liệu audio
- Biết về đặc tính tần số/thời gian của audio
- Trích các đặc tính của âm thanh thô
- Biết được các đặc tính nào có thể sử dụng cho học máy
- Sử dụng thư viện librosa

# Các thuộc tính của tín hiệu Audio

## Các ứng dụng sử dụng thuộc tính của tín hiệu audio

Bài toán phân loại audio



# Các thuộc tính của tín hiệu Audio

## Các ứng dụng sử dụng thuộc tính của tín hiệu audio

- Một số ứng dụng dựa vào tín hiệu audio
  - Phân loại audio
  - Nhận dạng giọng nói
  - Khử nhiễu audio
  - Trích thông tin từ audio: âm thanh nhạc cụ, phân loại mood, ....

# Các thuộc tính của tín hiệu Audio

## Đặc tính công suất, cường độ của audio

- **Công suất audio** = công suất tại nơi phát âm thanh (đơn vị Watt)
- **Cường độ âm thanh** = công suất đo được tại điểm thu /m<sup>2</sup> (đơn vị đo W/m<sup>2</sup>)
- **Ngưỡng nghe** của tai người (ngưỡng thấp):

$$TOH = 10^{-12} \text{W/m}^2$$

- **Ngưỡng chói tai** (chịu đựng)

$$TOP = 10 \text{W/m}^2$$

# Các thuộc tính của tín hiệu Audio

## Mức cường độ

---

- Mức cường độ (intensity level)
  - Được ở thang logarithm
  - Được đo bằng decibels (dB)
  - Được sử dụng để xác định mức TOH

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

# Các thuộc tính của tín hiệu Audio

## Mức cường độ

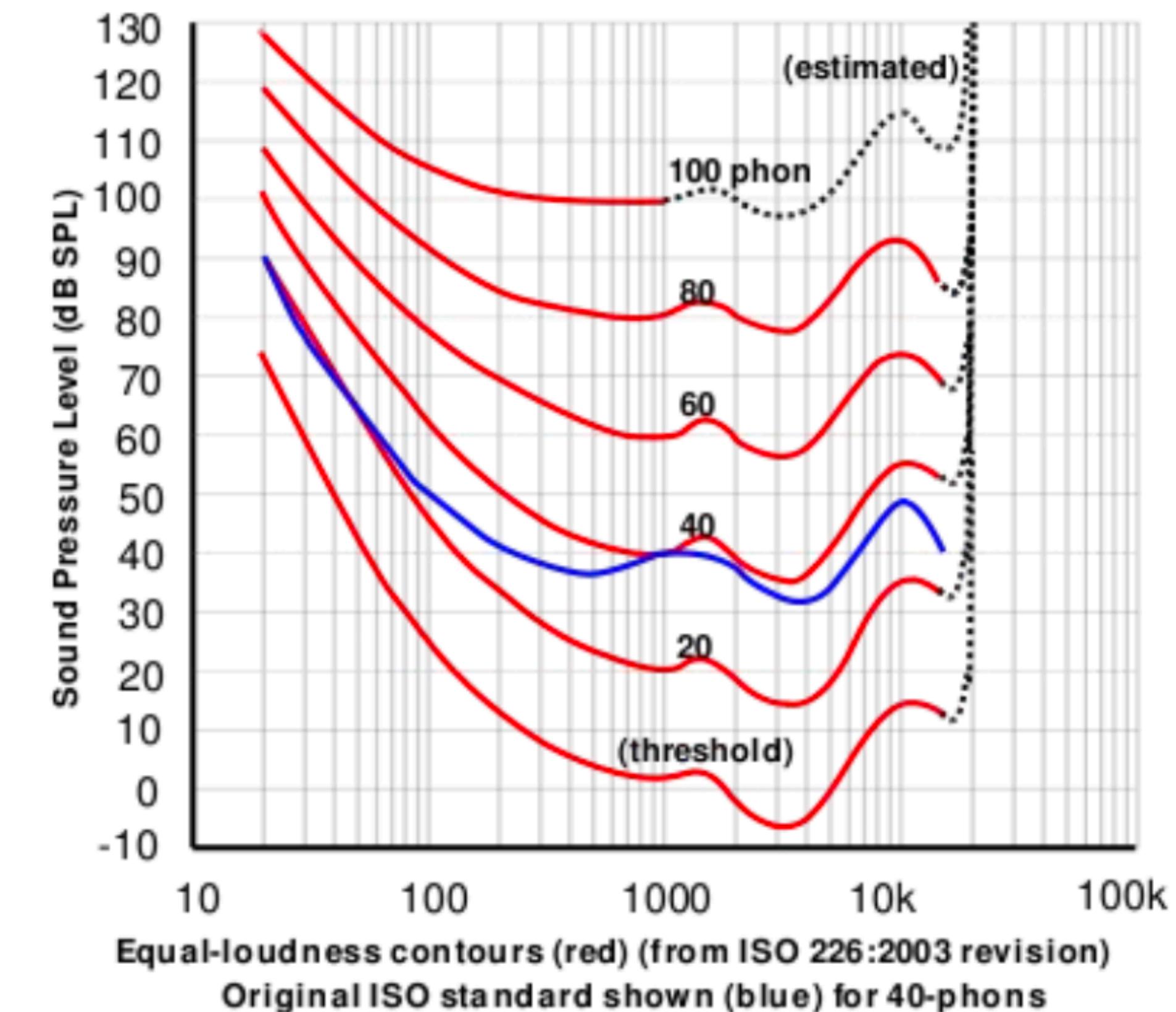
Source	Intensity	Intensity level	$\times$ TOH
Threshold of hearing (TOH)	$10^{-12}$	0 dB	1
Whisper	$10^{-10}$	20 dB	$10^2$
Pianissimo	$10^{-8}$	40 dB	$10^4$
Normal conversation	$10^{-6}$	60 dB	$10^6$
Fortissimo	$10^{-2}$	100 dB	$10^{10}$
Threshold of pain	10	130 dB	$10^{13}$
Jet take-off	$10^2$	140 dB	$10^{14}$
Instant perforation of eardrum	$10^4$	160 dB	$10^{16}$

Table 1.1 from [Müller, FMP, Springer 2015]

# Các thuộc tính của tín hiệu Audio

## Độ ồn (Loudness)

- Là khái niệm có tính cảm nhận chủ quan về cường độ audio
- Phụ thuộc vào thời gian tồn tại, tần suất lặp lại của audio
- Phụ thuộc vào độ tuổi
- Đo bằng đại lượng *phons*



# Các thuộc tính của tín hiệu Audio

## Âm sắc (Timbre)

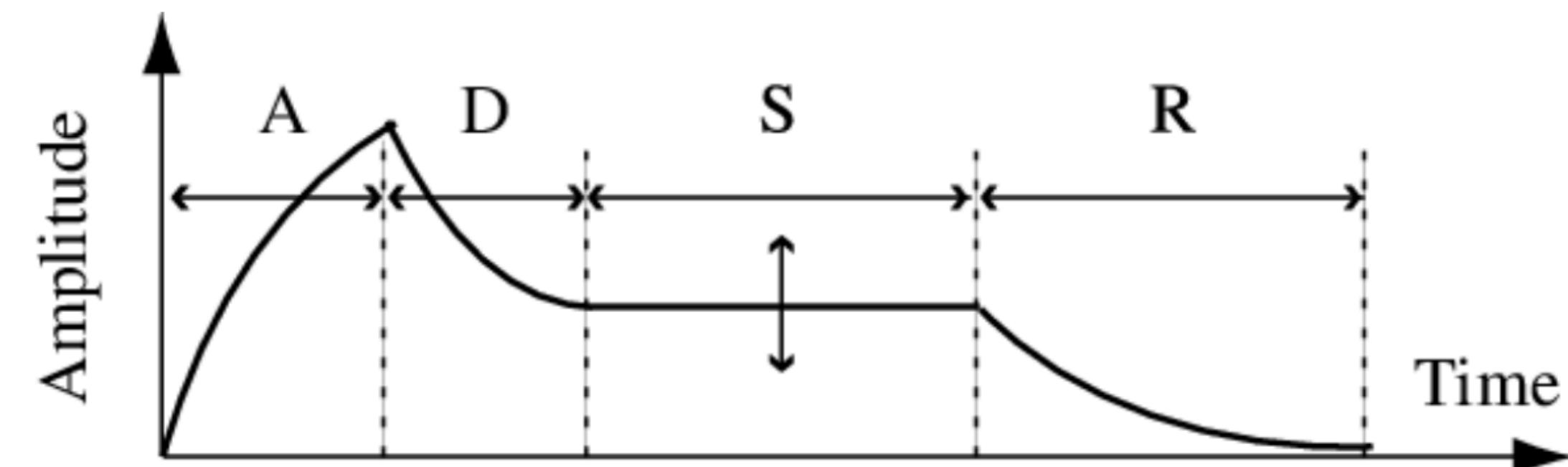
---

- Âm sắc = màu sắc của audio
- Âm sắc thể hiện sự khác nhau giữa 2 âm với cùng cường độ, tần số và thời lượng
- Thể hiện bằng các từ như: âm sáng, tối, buồn, hùng hồn, ấm áp
- Âm sắc được hình thành từ nhiều chiều khác:
  - Biên âm thanh (sound envelope)
  - cái hài âm thanh
  - Điều chế tần số, biên độ

# Các thuộc tính của tín hiệu Audio

## Biên âm thanh

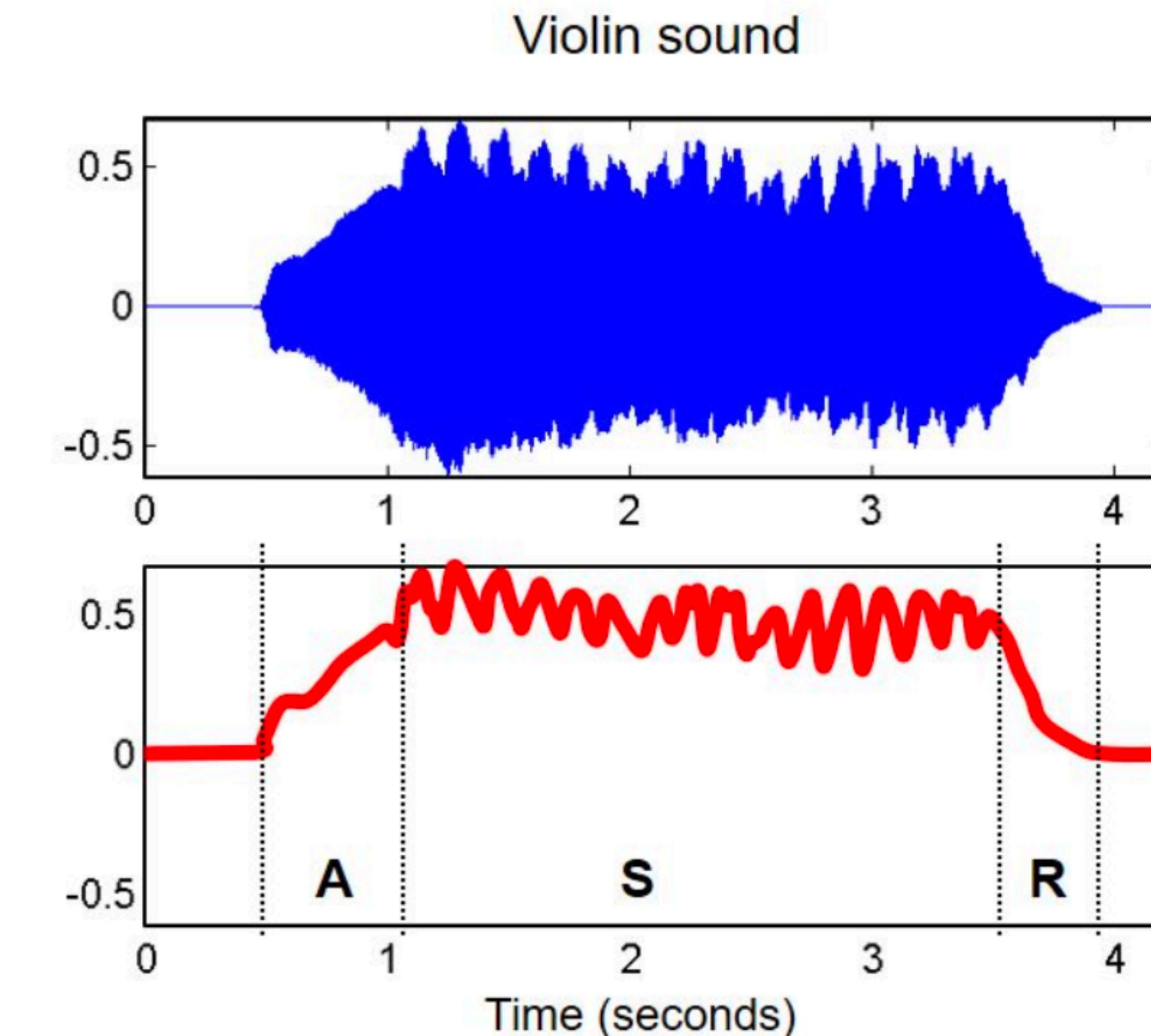
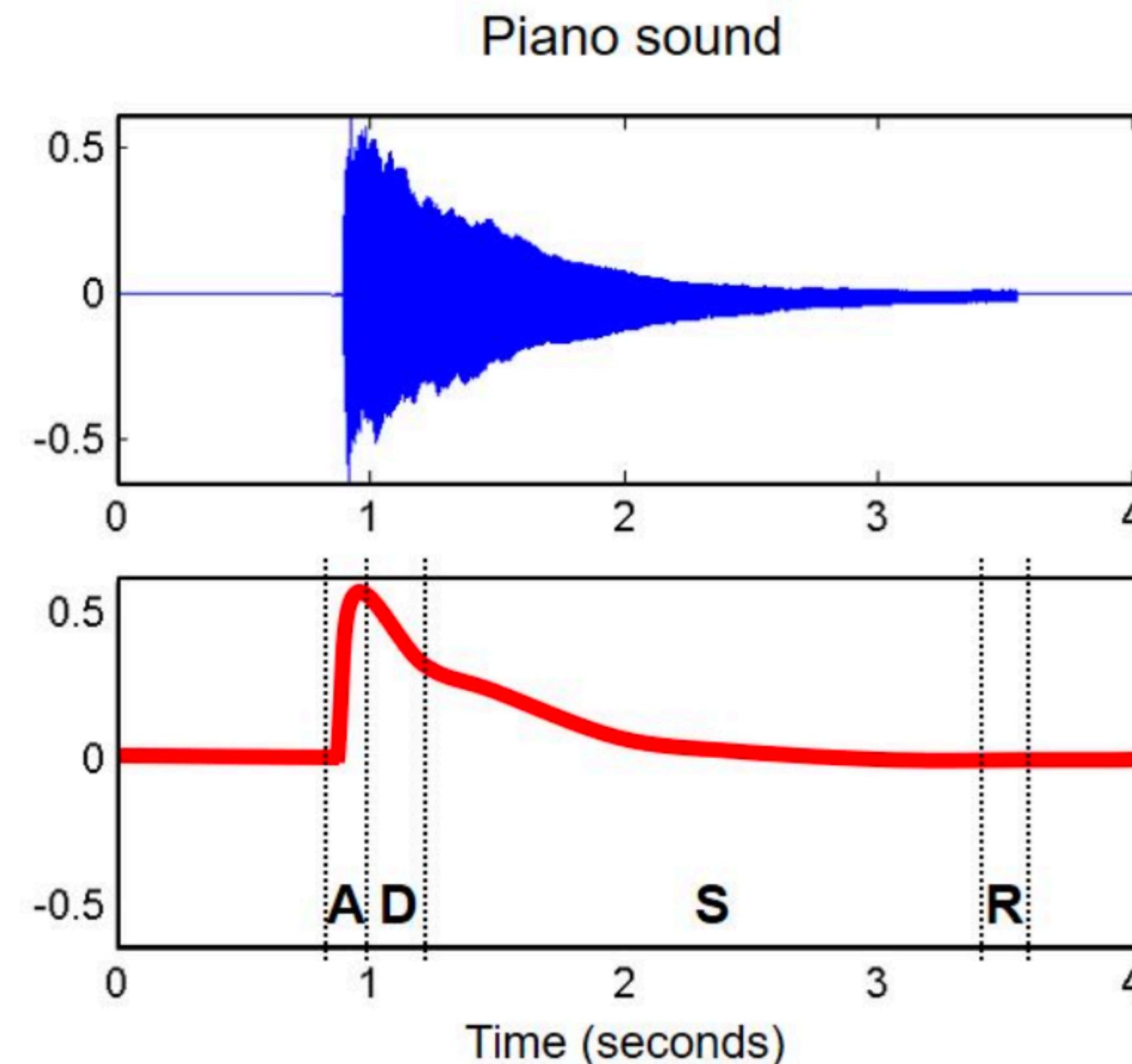
- Được phân thành: Attack - Decay - Sustain - Release (ADSR)



# Các thuộc tính của tín hiệu Audio

## Biên âm thanh

- Được phân thành: Attack - Decay - Sustain - Release (ADSR)



# Các thuộc tính của tín hiệu Audio

## Âm thanh phức hợp

---

- Audio phức hợp = chồng chập nhiều sóng sin. Trong đó chỉ có một phần sóng sin chứa âm thanh mong muốn.
- Phần âm thanh thấp nhất được gọi là tần số cơ bản
- Các hài âm thanh khác là các bộ số của tần số cơ bản
$$f_1 = 440, f_2 = 2 \cdot 440 = 880, f_3 = 3 \cdot 440 = 1320, \dots$$
- Các âm thanh phức hợp không hài hòa sẽ chứa các tần số không còn là bộ số của tần số cơ bản

# Các thuộc tính của tín hiệu Audio

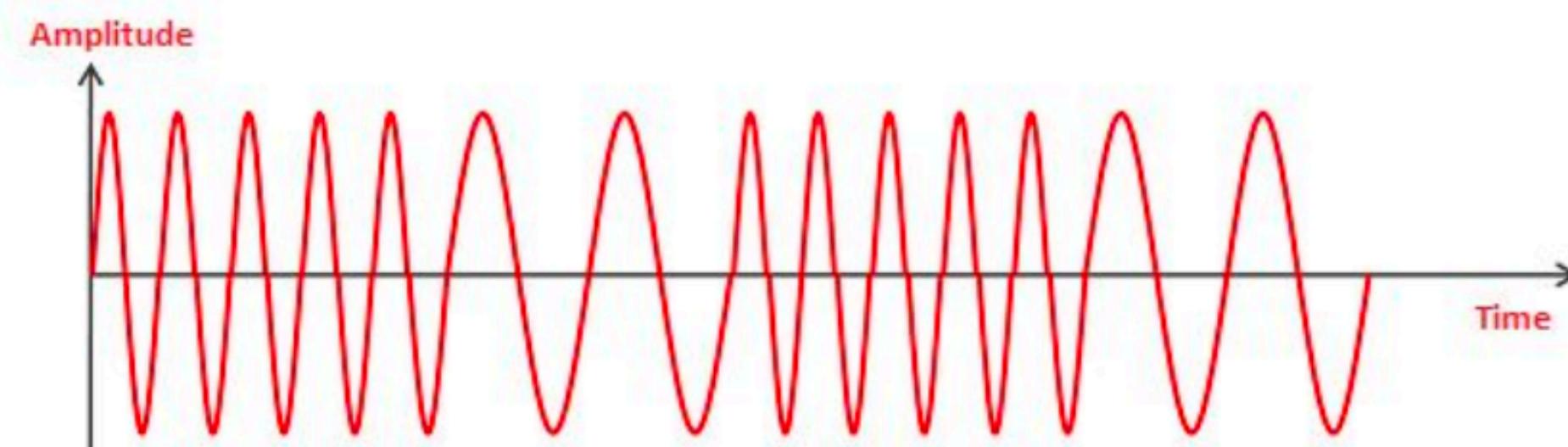
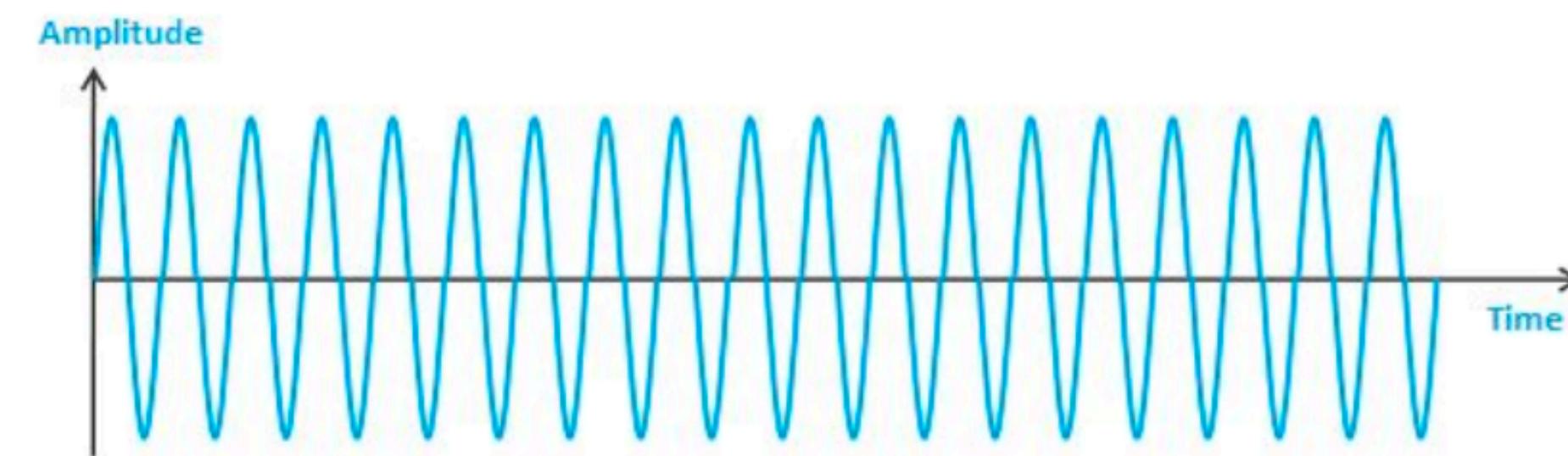
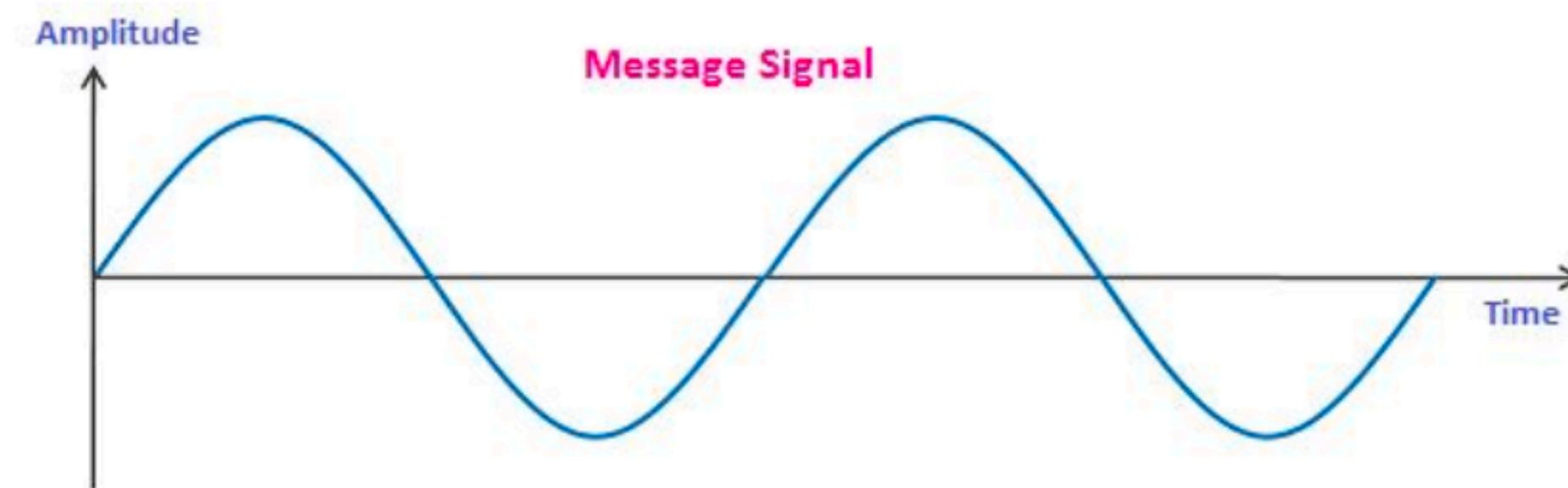
## Âm thanh phức hợp

---



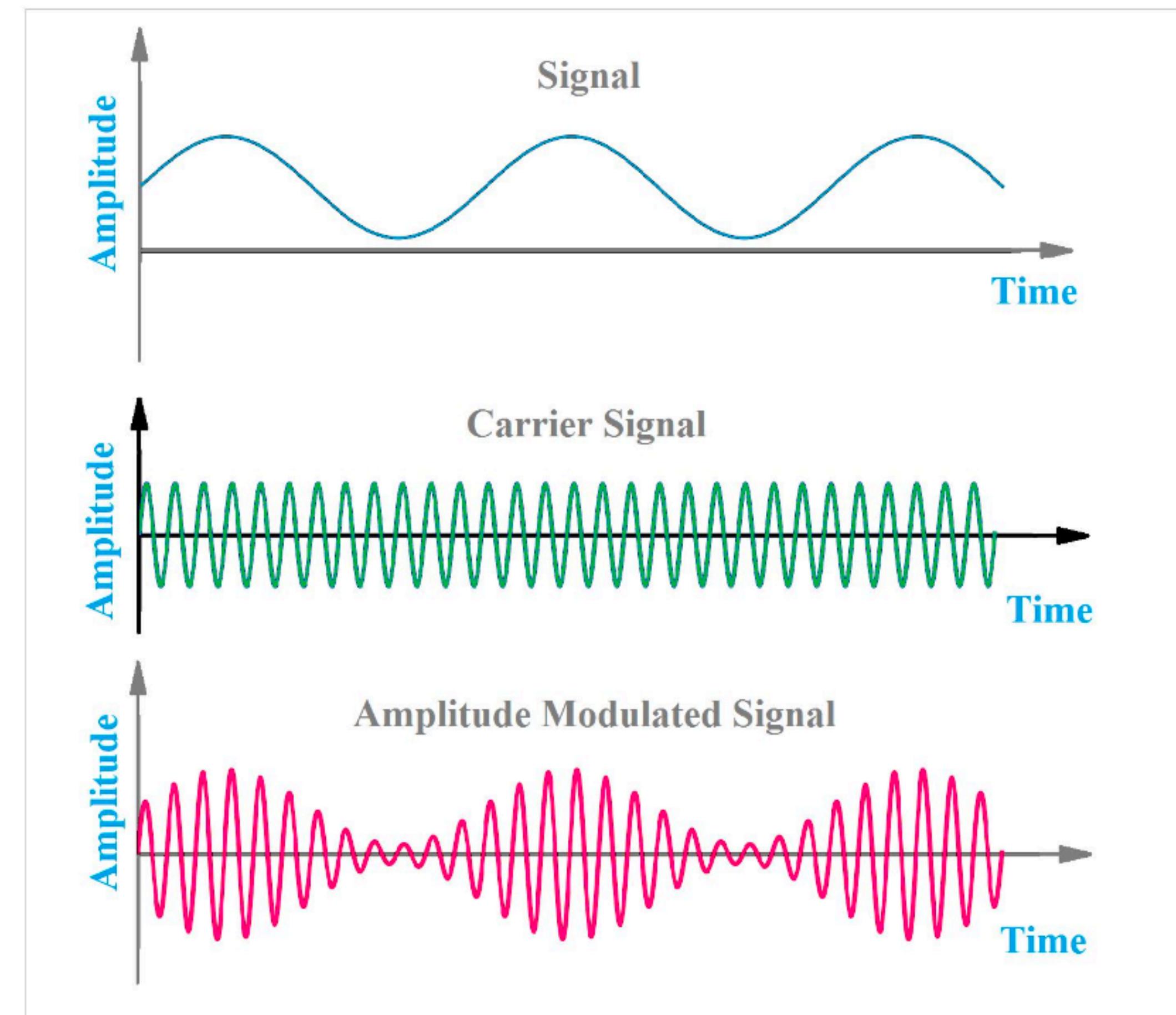
# Các thuộc tính của tín hiệu Audio

## Điều chế tần số



# Các thuộc tính của tín hiệu Audio

## Điều chế biên độ



# Các thuộc tính của tín hiệu Audio

## Audio số chuẩn hiện nay

$$f_{\text{Nyquist}} = 22050 \text{ Hz}$$

$$f_{\text{sampling}} = 2 \times f_{\text{Nyquist}} = 44100 \text{ Hz}$$

Encoding bit-length = 16 bits/ sample

$$\text{SQNR} = 6.02 \times Q = 16dB$$



# Các thuộc tính của tín hiệu Audio

## Các thuộc tính được sử dụng trong học máy

- Phân loại các thuộc tính audio
  - Các mức độ về abstraction
  - Phạm vi về thời gian
  - Về các thể loại nhạc
  - Miền tín hiệu
  - Các giải pháp về học máy

# Các thuộc tính của tín hiệu Audio

## Phân loại các thuộc tính audio > Mức độ abstraction



- Mức abstract cao
  - Ví dụ: nhạc cụ, thể loại nhạc (genre: rock, jazz,...), key (dạng midi), tempo, giai điệu, hợp âm (chord), mood
- Mức abstract trung bình
  - Ví dụ: pitch (cao độ của nốt nhạc), tiếng beat (nhịp), onset (thời điểm bắt đầu của bản nhạc, âm thanh), mô hình biến đổi (fluctuation pattern), đặc trưng MFCC.
- Mức abstract thấp:
  - Đường bao biên độ, năng lượng, tâm phổ audio (spectral centroid), dòng phổ (spectral flux), mức chuyển đổi zero (zero-crossing rate)



# Các thuộc tính của tín hiệu Audio

## Phân loại các thuộc tính audio > Phạm vi về thời gian

---

- Phạm vi về thời gian (temporal scope)
  - Phạm vi tức thời (trong khoảng 50ms)
  - Mức đoạn (segment level) (vài dây)
  - Mức toàn thể (cả bài nhạc)

# Các thuộc tính của tín hiệu Audio

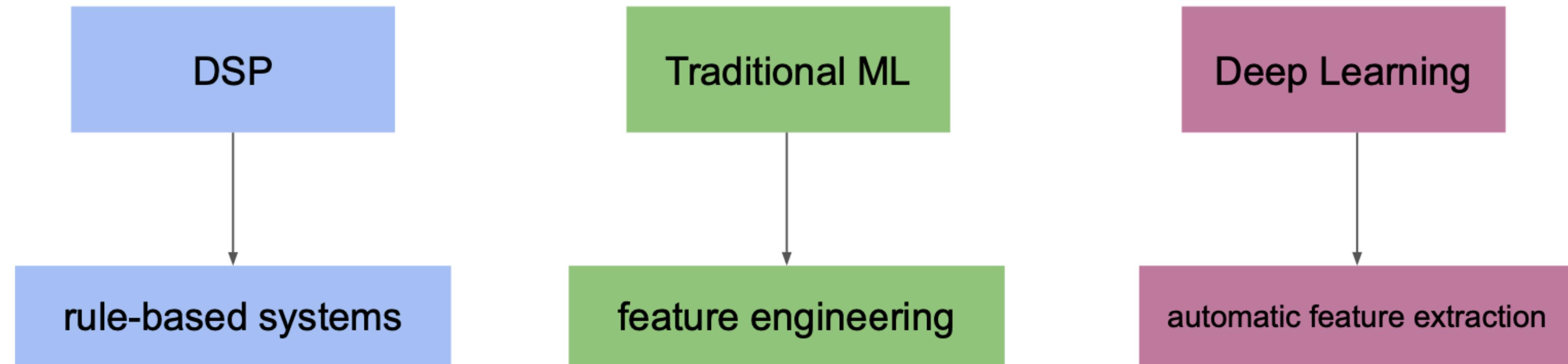
## Phân loại các thuộc tính audio > Phạm vi về âm nhạc (music)

---

- Phạm vi về âm nhạc:
  - Nhịp điệu (beat)
  - Âm sắc (timbre)
  - Nốt chính (pitch)
  - Hoà âm (harmony)
  - ....

# Các thuộc tính của tín hiệu Audio

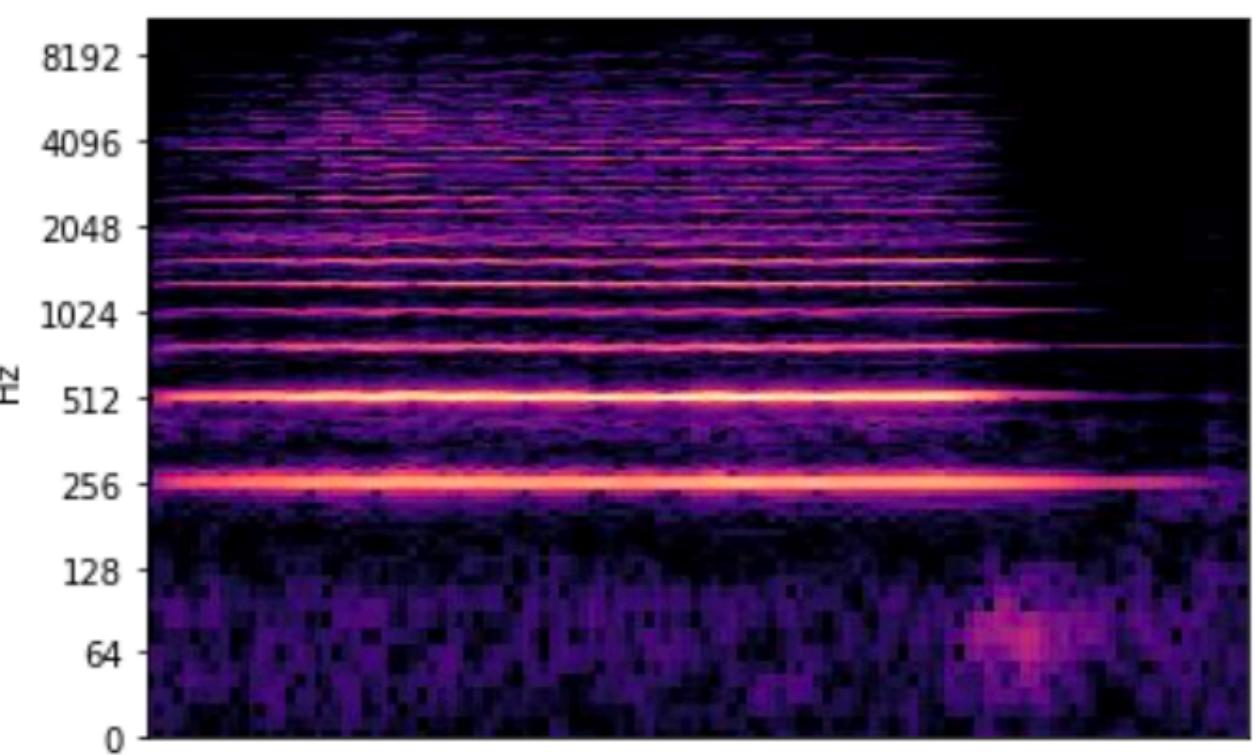
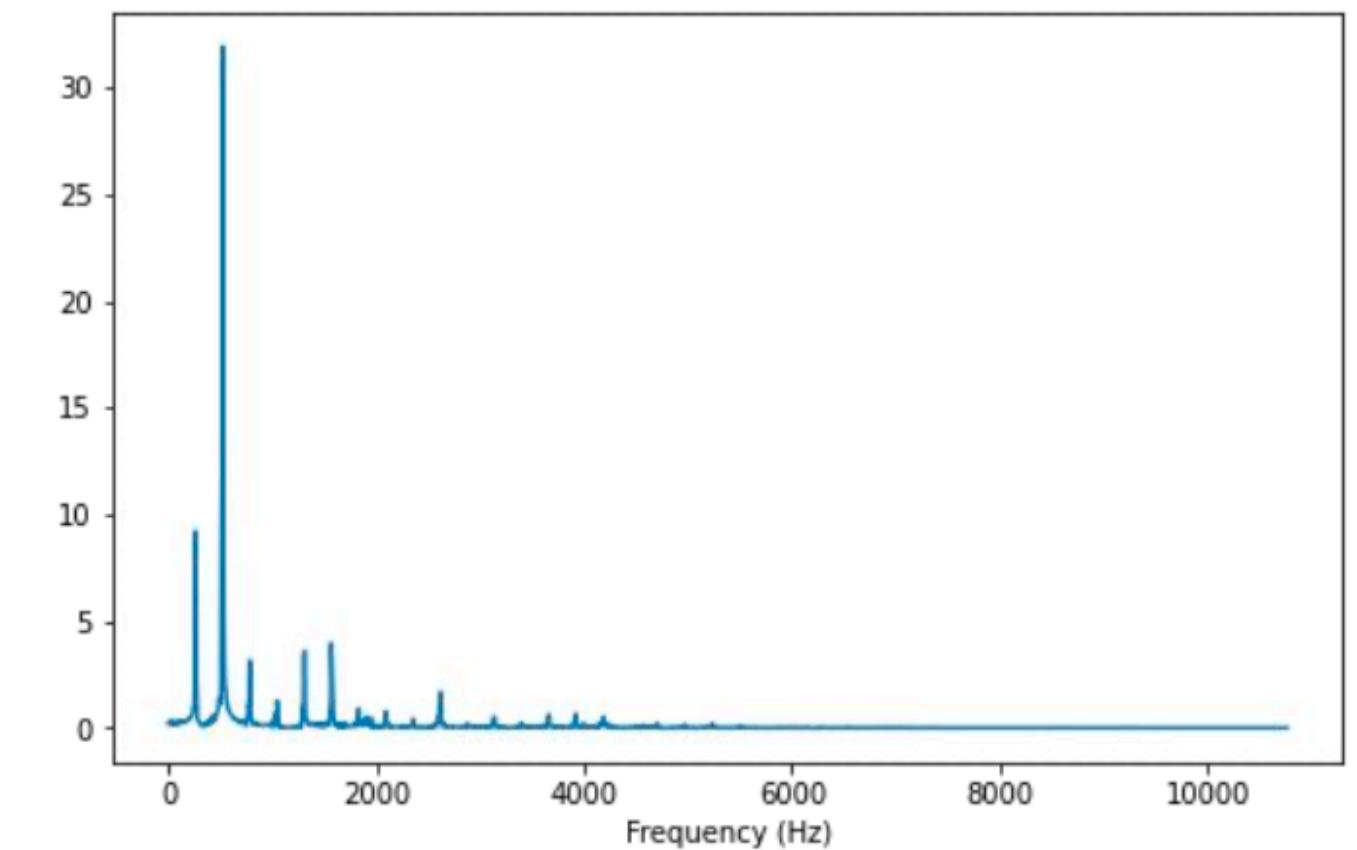
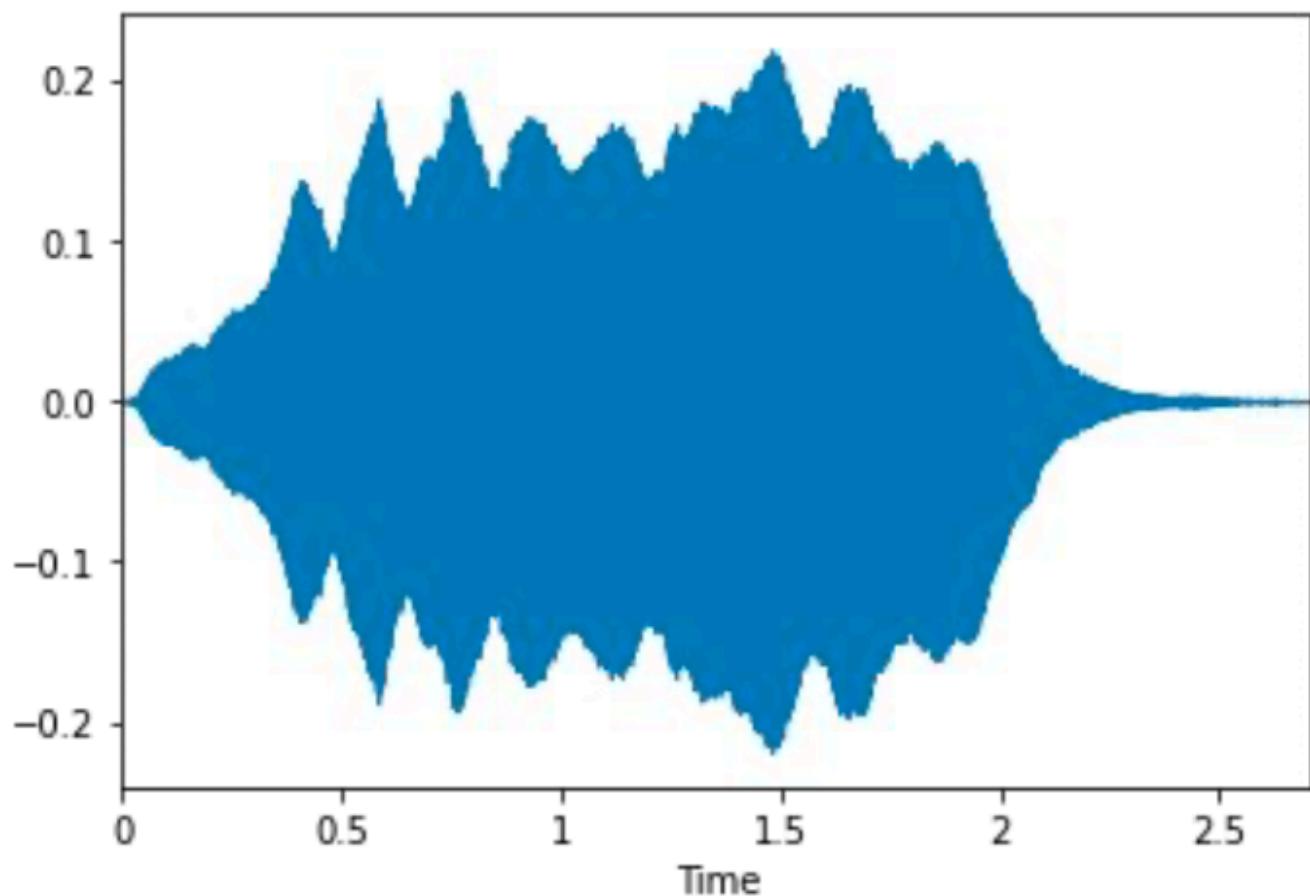
## Đặc tính âm thanh sử dụng vào AI



# Các thuộc tính của tín hiệu Audio

## Phạm vi miền tín hiệu

- Miền thời gian
- Miền tần số
  - Công suất in band
  - Tâm phổ
  - Dòng phổ
- Biểu diễn dạng time-frequency
  - Spectrogram
  - Mel-Spectrogram
  - Constant-Q-Transform



# Các thuộc tính của tín hiệu Audio

## Đặc tính âm thanh sử dụng vào AI > Học máy truyền thống

- Đặc tính sử dụng trong học máy truyền thống

- Amplitude envelope
- Root-mean square energy
- Zero crossing rate
- Band energy ratio
- Spectral centroid
- Spectral flux
- Spectral spread
- Spectral roll-off

Amplitude envelope  
Zero crossing rate  
Spectral flux



Traditional  
ML algorithm

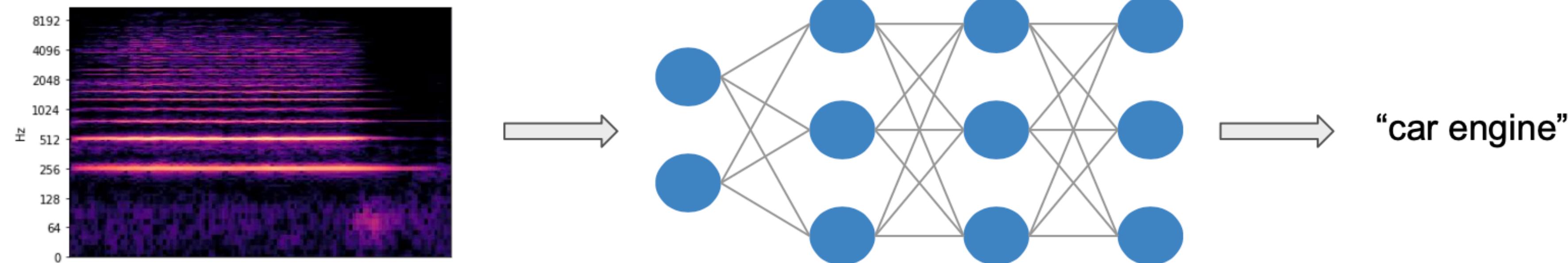


“car engine”

# Các thuộc tính của tín hiệu Audio

## Đặc tính âm thanh sử dụng vào AI

- Đặc tính sử dụng vào học sâu



# Quiz :

# Quiz

## Nội dung

---

- Các lý thuyết cơ sở về audio