

# Xử Lý Tiếng Nói

Nguyễn Đức Hoàng Hạ  
2022



# Nội dung

- Khái niệm về tín hiệu số
- Một số tín hiệu cơ bản và phân loại dạng tín hiệu số
- Các phép toán cơ bản trên tín hiệu
- Biến đổi Fourier
- Standing wave, harmonics, formants and speech features



# Trắc nghiệm lý thuyết toán về số phức

<https://www.menti.com/>



# Tín hiệu là gì?

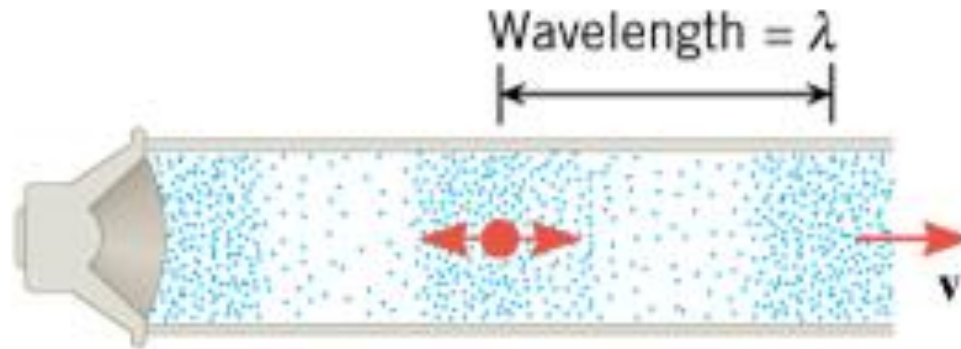
Wikipedia: “Trong lý thuyết thông tin, một chuyên ngành của toán học ứng dụng và kỹ thuật điện/điện tử, **tín hiệu là một đại lượng vật lý chứa đựng thông tin hay dữ liệu và có thể truyền đi được.**”

Ví dụ:

- 1) Âm thanh là dạng tín hiệu về dao động áp suất trong không khí
- 2) Hình ảnh là dạng tín hiệu về cường độ sáng và màu sắc tại một vị trí.



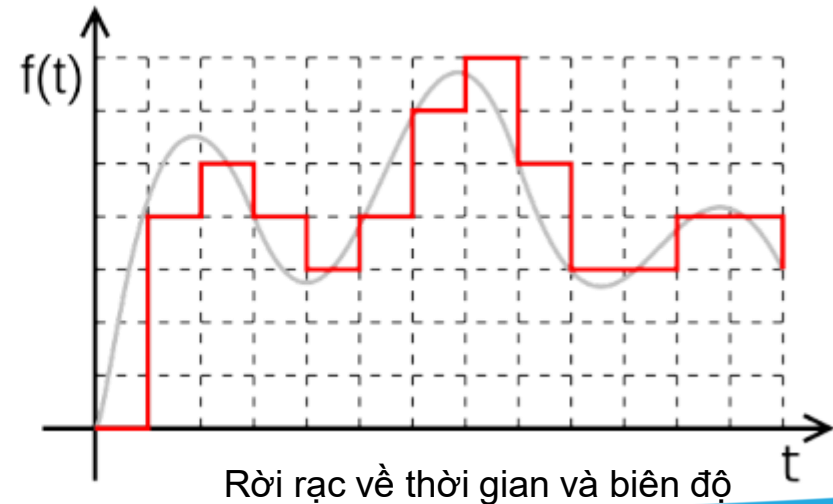
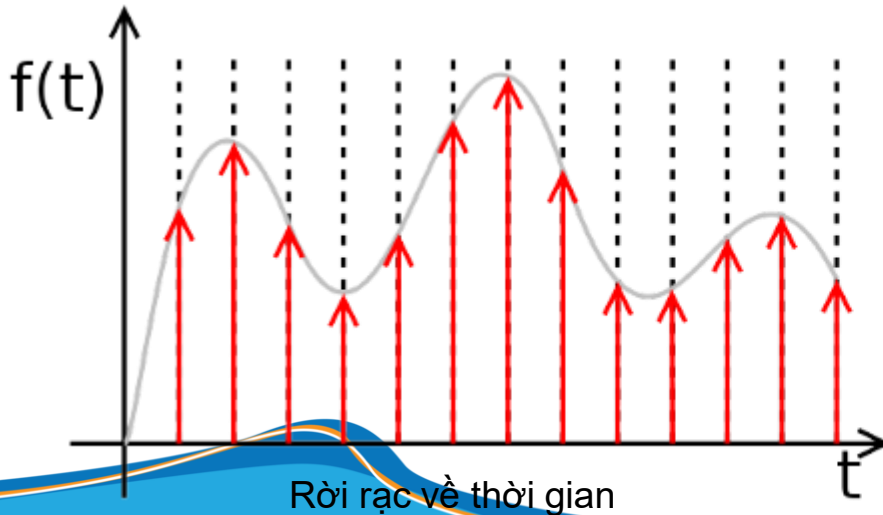
# Âm thanh



# Tín hiệu số (Digital Signal)

Analog signal: **liên tục** về thời gian và biên độ

Digital signal: **rời rạc** về thời gian và biên độ



# Tín hiệu số

Rời rạc hóa trên miền thời gian:

- + Lấy mẫu tín hiệu thay thế cho toàn bộ tín hiệu tương tự
- + Đơn giản hóa vấn đề tính toán

Rời rạc hóa giá trị biên độ:

- + Lưu trữ trên thiết bị số
- + Xử lý trên tín hiệu số



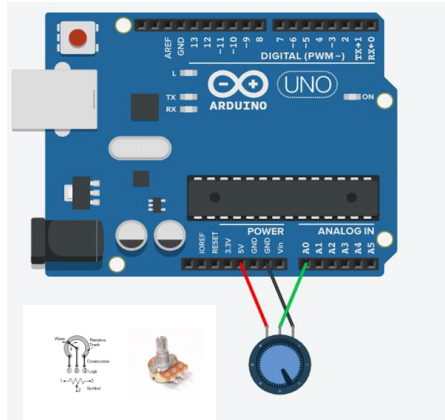
# Ưu điểm của tín hiệu số so với tín hiệu tương tự

- + Lưu trữ dễ dàng
- + Xử lý tính toán đơn giản hơn
- + Truyền và nhận tín hiệu chính xác và hiệu quả hơn



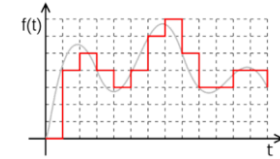


# Ví dụ AnalogRead trên <https://www.tinkercad.com/>



```
int sensorValue = 0;
void setup()
{
  pinMode(A0, INPUT);
  Serial.begin(9600);
}
void loop()
{
  // read the analog in value:
  sensorValue = analogRead(A0);
  // print the results to the serial monitor:
  Serial.print("sensor = ");
  Serial.println(sensorValue);

  delay(200); // Wait for 200 millisecond(s)
}
```



ADC 10 bits → [0,1023]

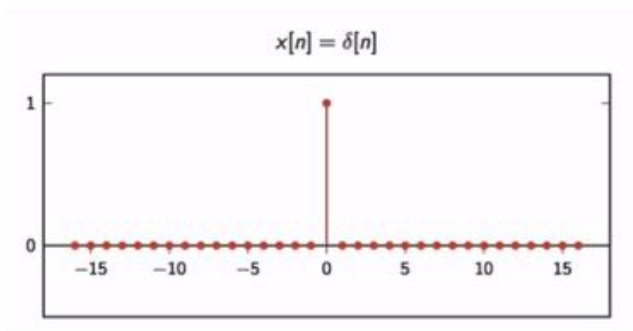


## Bài 2.1: Xử lý tín hiệu số

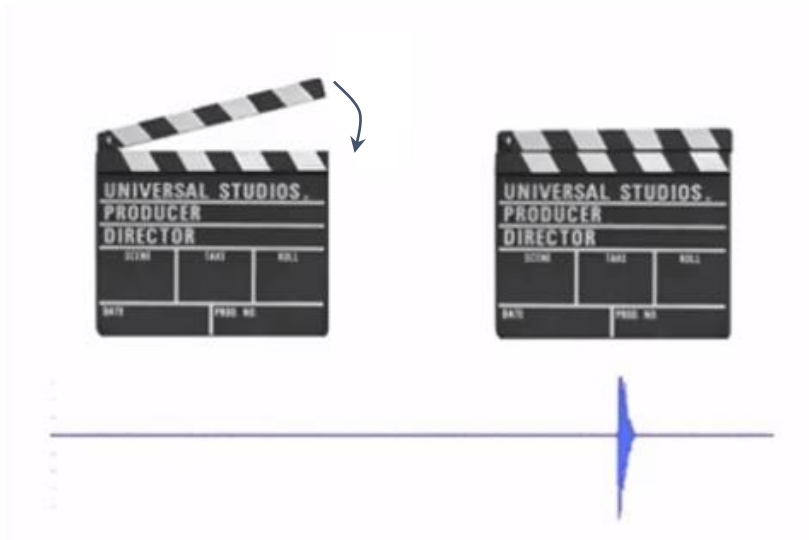
- Tín hiệu số
- **Một số tín hiệu cơ bản và phân loại dạng tín hiệu số**
- Các phép toán cơ bản trên tín hiệu
- Biến đổi Fourier



# The delta signal

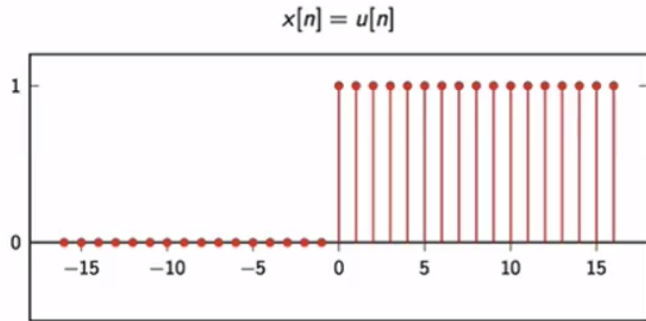


The delta signal



Ví dụ trong đồng bộ audio và video

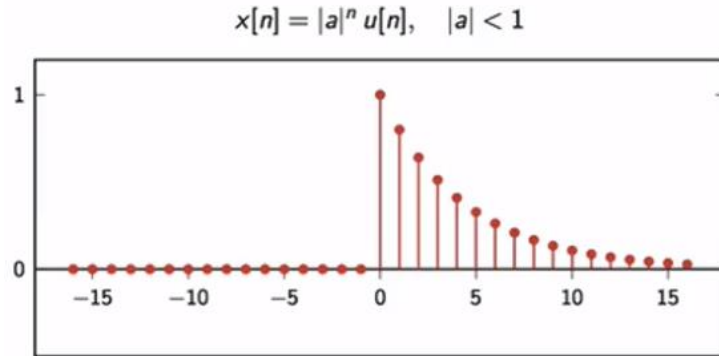
# The unit step



The unit step



# The exponential decay

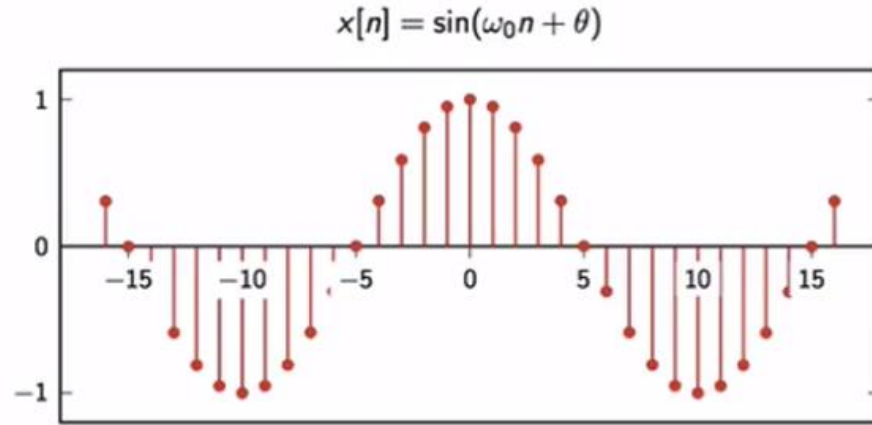


The exponential decay

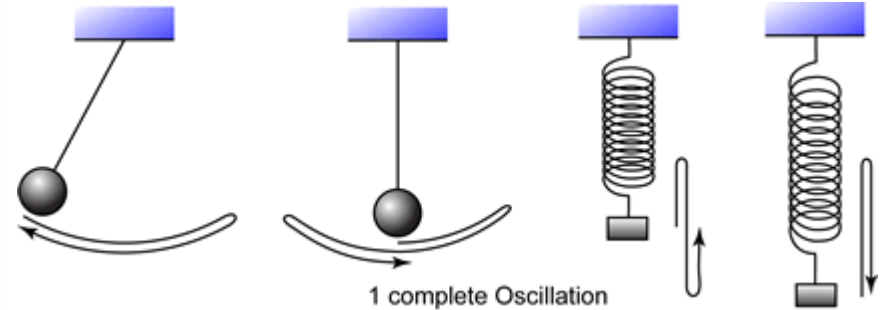


Ví dụ trong bài toán trao đổi nhiệt

# Tín hiệu sin/cos

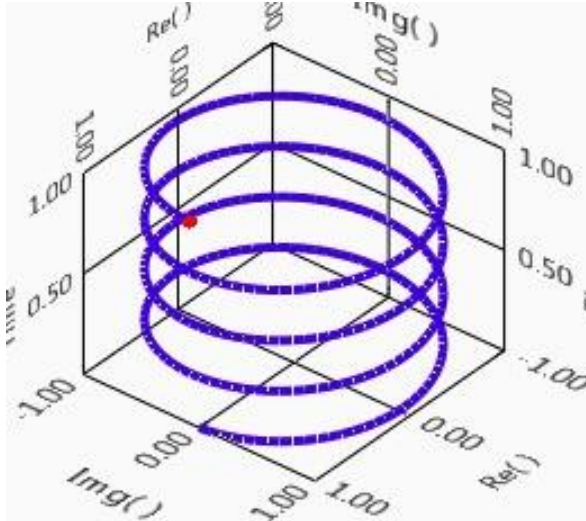


The sinusoid

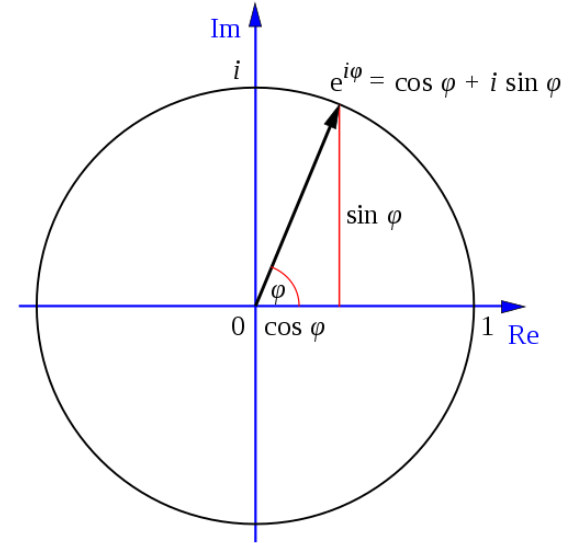


Ví dụ về dao động điều hòa

# Complex exponential signals



$$z(t) = Ae^{j(\omega_0 t + \phi)}$$



Euler's formula

# Phân loại dạng tín hiệu





# Phân loại dạng tín hiệu

1. Finite length
2. Infinite length
3. Periodic
4. Finite support



# Finite-length signals

$$x[n], n = 0, 1, \dots, N - 1$$

Biểu diễn ở dạng vector

$$\mathbf{x} = [x_0 x_1 \dots x_{N-1}]^T = \begin{bmatrix} x_0 \\ x_1 \\ \dots \\ x_{N-1} \end{bmatrix}$$

# Infinite-length signals

$$x[n], n \in \mathbb{Z}$$

Được dùng trong lý thuyết, để tổng quát hóa tín hiệu

# Periodic signals

$$x[n] = x[n + kN]$$
$$n, k, N \in \mathbb{Z}$$

# Periodic signals

$$x[n] = x[n + kN]$$
$$n, k, N \in \mathbb{Z}$$

Chuỗi tín hiệu được lặp lại với chu kỳ  $N$   
Thông tin trên tín hiệu tương đương với chuỗi tín hiệu có length =  $N$   
Tuy nhiên, chuỗi có chiều dài vô hạn

# Finite support signals

$$\tilde{x}[n] = \begin{cases} x[n] & \text{if } 0 \leq n < N \\ 0 & \text{otherwise} \end{cases}$$



# Finite support signals

$$\tilde{x}[n] = \begin{cases} x[n] & \text{if } 0 \leq n < N \\ 0 & \text{otherwise} \end{cases}$$

Thông tin tín hiệu tương tự trong chuỗi tín hiệu với length = N  
Là một cách để liên kết giữa tín hiệu có giới hạn chiều dài và không giới hạn chiều dài

# Các phép toán cơ bản trên tín hiệu

1. Scaling

$$y[n] = \alpha x[n]$$

2. Sum

$$y[n] = x[n] + z[n]$$

3. Product

$$y[n] = x[n] \cdot z[n]$$

4. Delay

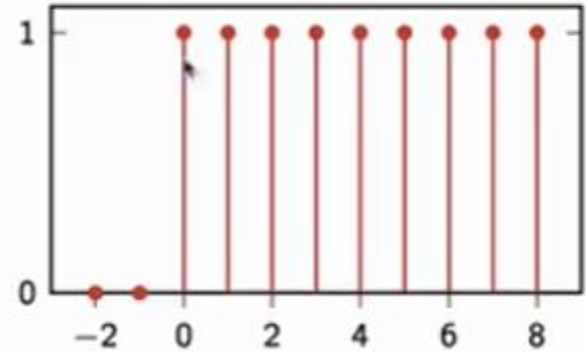
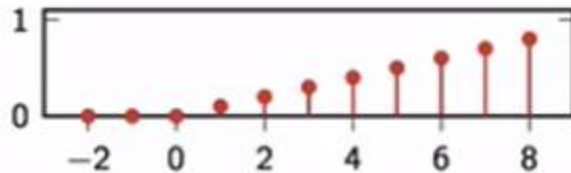
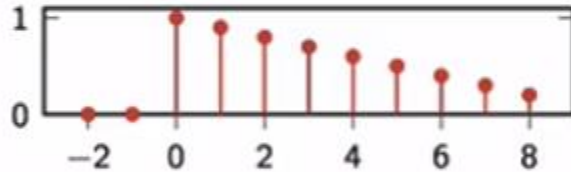
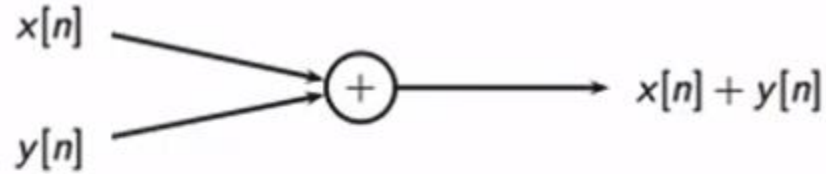
$$y[n] = x[n - k]$$



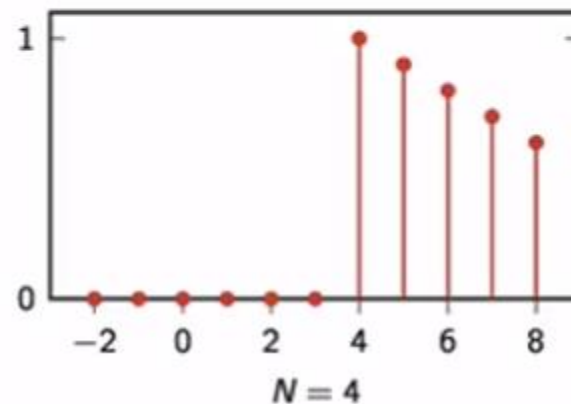
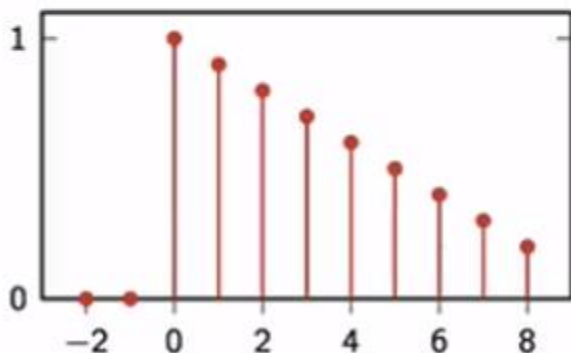
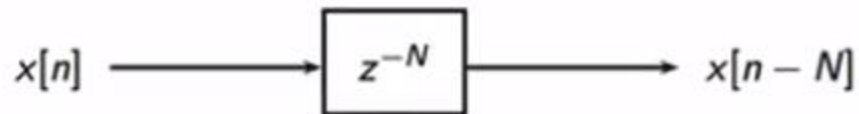
# Biểu diễn các phép toán qua sơ đồ



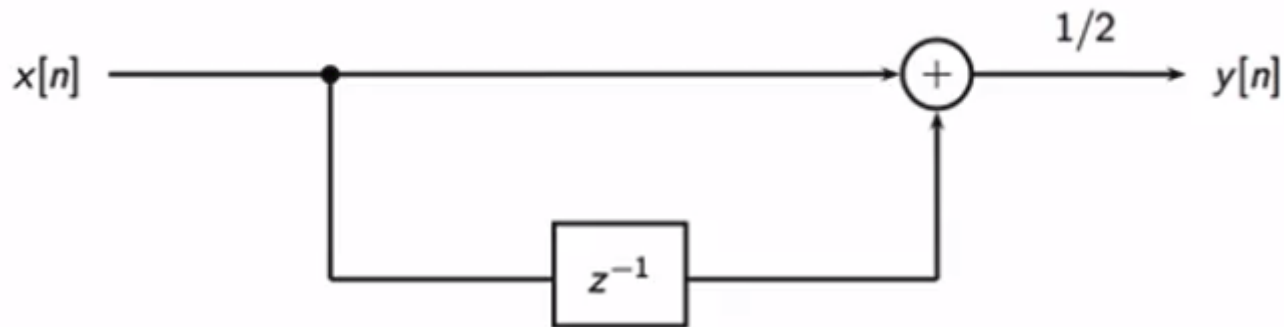
# Building block: add



# Building block: delay

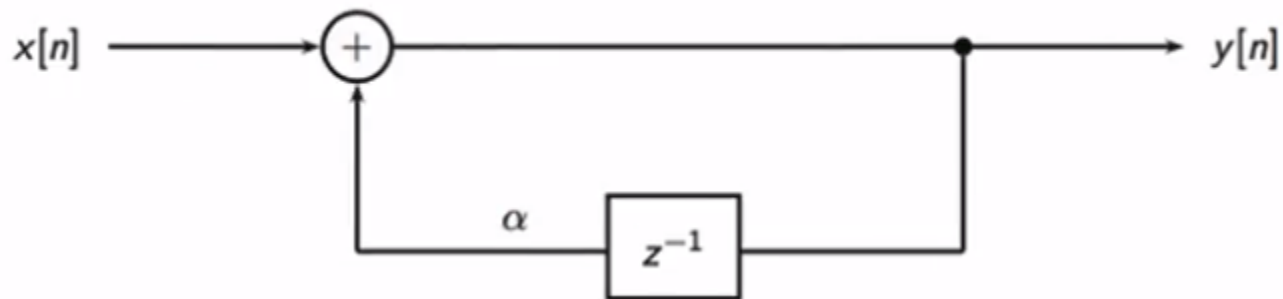


# Bài tập



Viết lại công thức tính cho sơ đồ trên.

# Bài tập



Viết lại công thức tính cho sơ đồ trên.

# Fourier transform

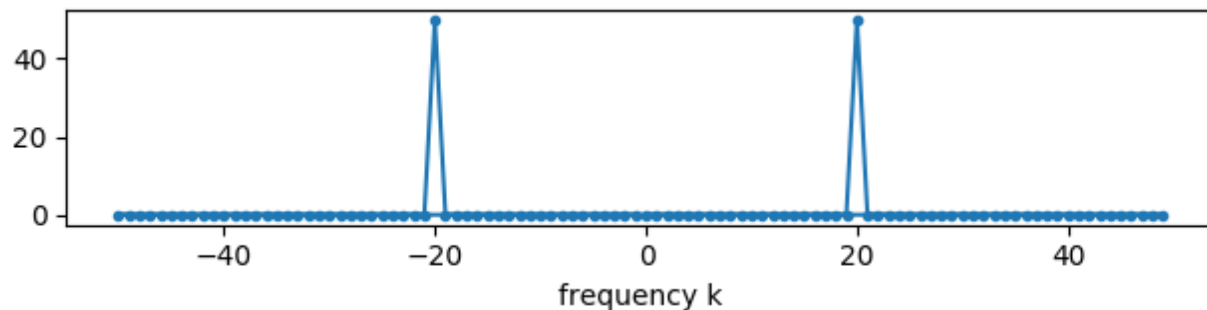
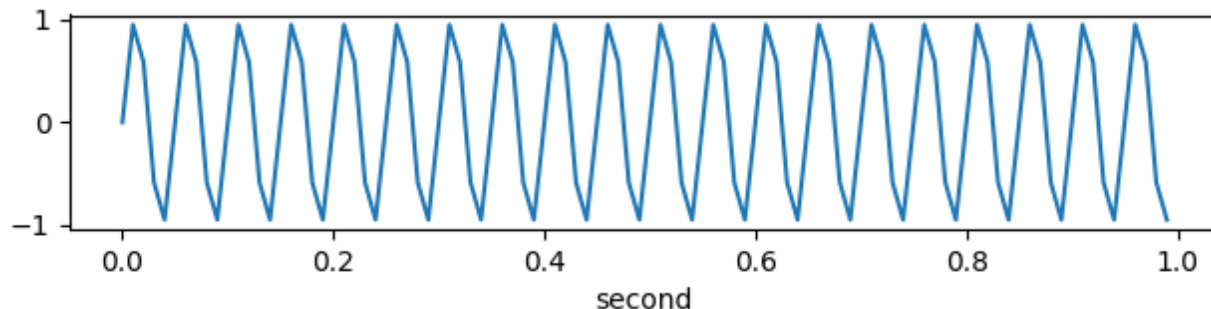
Wikipedia: "Biến đổi Fourier hay chuyển hóa Fourier, được đặt tên theo nhà toán học người Pháp Joseph Fourier, là **phép biến đổi một hàm số hoặc một tín hiệu theo miền thời gian sang miền tần số.**"

Biến đổi Fourier rời rạc:

$$\begin{aligned} X_k &= \sum_{n=0}^{N-1} x_n \cdot e^{-i2\pi kn/N} \\ &= \sum_{n=0}^{N-1} x_n \cdot [\cos(2\pi kn/N) - i \cdot \sin(2\pi kn/N)], \end{aligned}$$

# Ví dụ 1 về biến đổi Fourier

$$x[t] = \sin(2\pi kt), k = 20$$

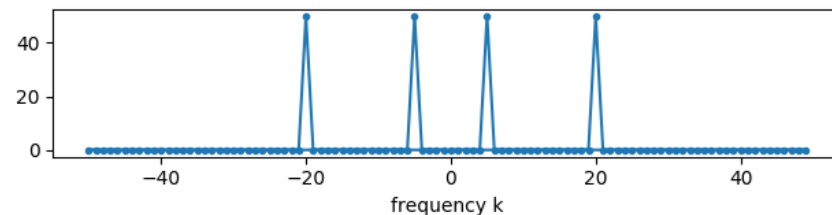
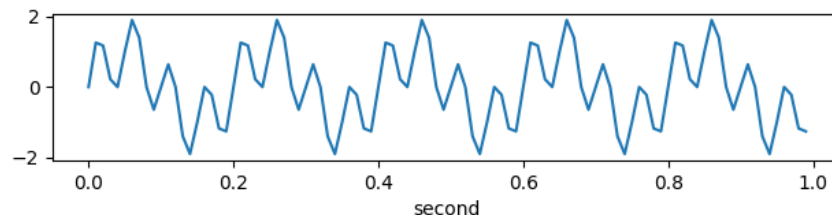


# Ví dụ 2 về biến đổi Fourier

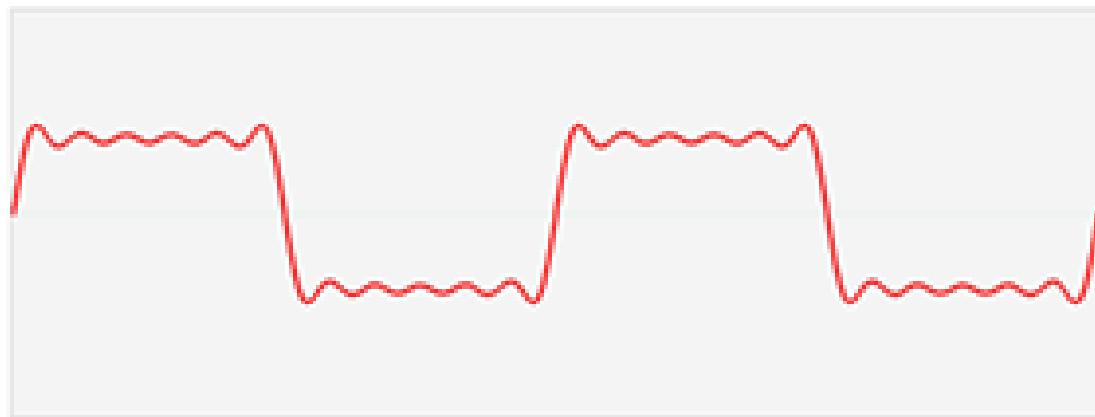
$$x[t] = \sin(2\pi k_1 t) + \sin(2\pi k_2 t)$$

$$k_1 = 5$$

$$k_2 = 20$$







# Biến đổi Fourier của hàm $f(x)$

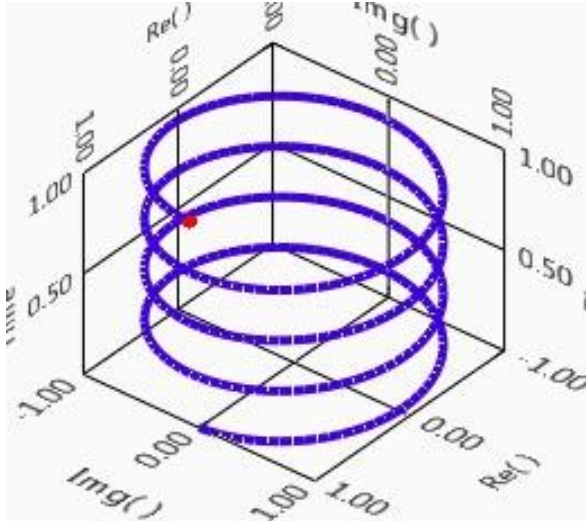
Thuận

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x \xi} dx,$$

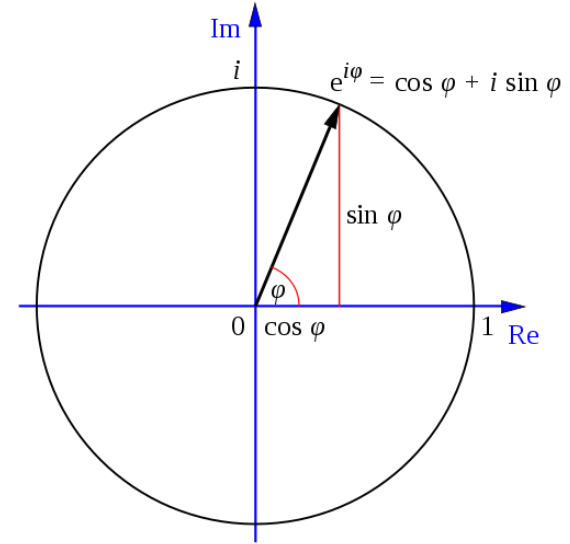
Ngược

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi) e^{2\pi i x \xi} d\xi,$$

# Nhắc lại về complex exponential signals



$$z(t) = Ae^{j(\omega_0 t + \phi)}$$



Euler's formula

# Một số tính chất của biến đổi Fourier

([https://en.wikipedia.org/wiki/Fourier\\_transform](https://en.wikipedia.org/wiki/Fourier_transform))



# Một số tính chất của biến đổi Fourier

## Linearity

For any complex numbers  $a$  and  $b$ ,

if

$$h(x) = af(x) + bg(x),$$

then

$$\hat{h}(\xi) = a \cdot \hat{f}(\xi) + b \cdot \hat{g}(\xi).$$



# Một số tính chất của biến đổi Fourier

Translation / time shifting

For any real number  $x_0$ ,

if

$$h(x) = f(x - x_0),$$

then

$$\hat{h}(\xi) = e^{-2\pi i x_0 \xi} \hat{f}(\xi).$$



# Một số tính chất của biến đổi Fourier

## Modulation / frequency shifting

For any real number  $\xi_0$ , if  $h(x) = e^{2\pi i x \xi_0} f(x)$ , then  $\hat{h}(\xi) = \hat{f}(\xi - \xi_0)$ .



# Một số tính chất của biến đổi Fourier

## Time scaling

For a non-zero real number  $a$ , if  $h(x) = f(ax)$ , then

$$\hat{h}(\xi) = \frac{1}{|a|} \hat{f}\left(\frac{\xi}{a}\right).$$

The case  $a = -1$  leads to the *time-reversal* property, which states: if  $h(x) = f(-x)$ , then  $\hat{h}(\xi) = \hat{f}(-\xi)$ .





# Một số tính chất của biến đổi Fourier

## Convolution

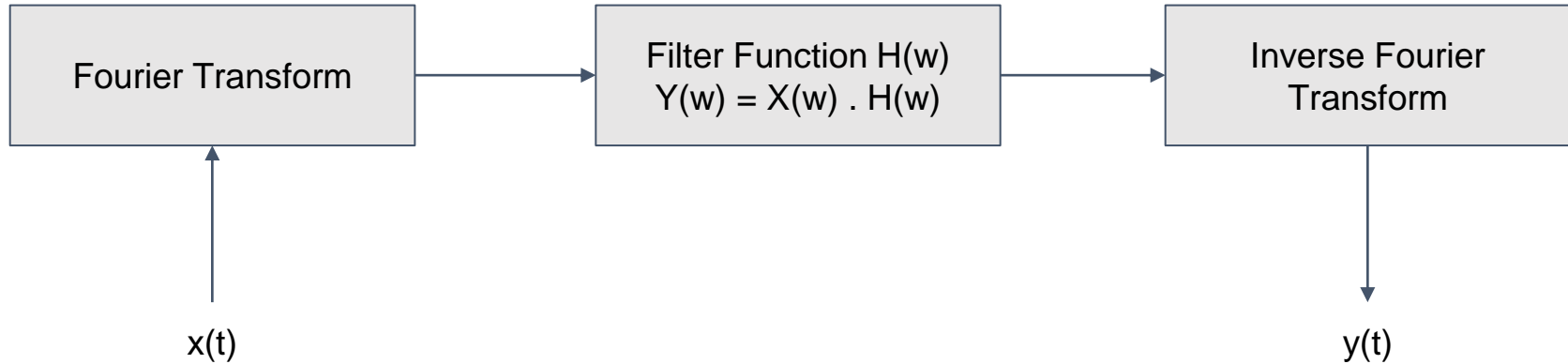
$$h(x) = (f * g)(x) = \int_{-\infty}^{\infty} f(y)g(x - y) dy,$$

where  $*$  denotes the convolution operation, then:

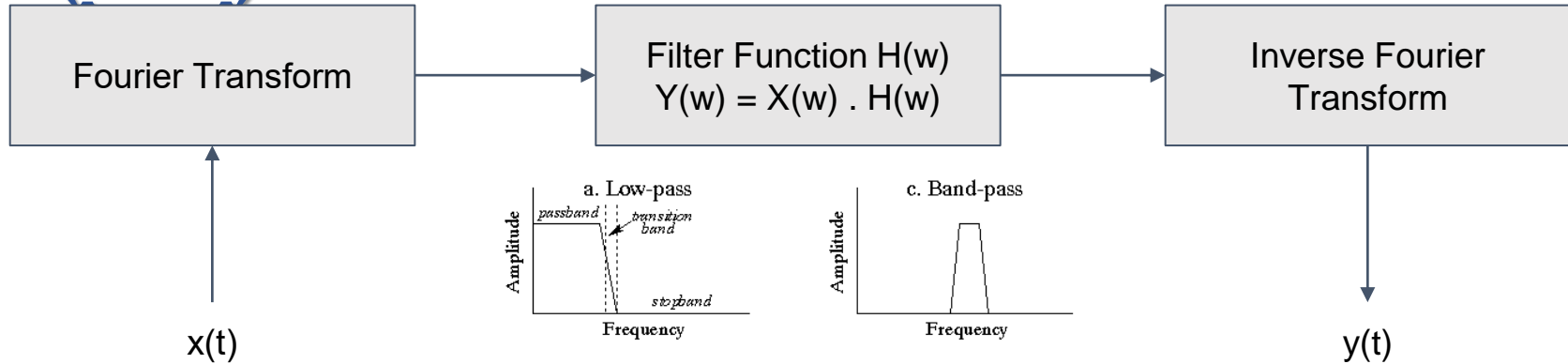
$$\hat{h}(\xi) = \hat{f}(\xi) \cdot \hat{g}(\xi).$$



# Ứng dụng trong thiết kế bộ lọc trên miền tần số

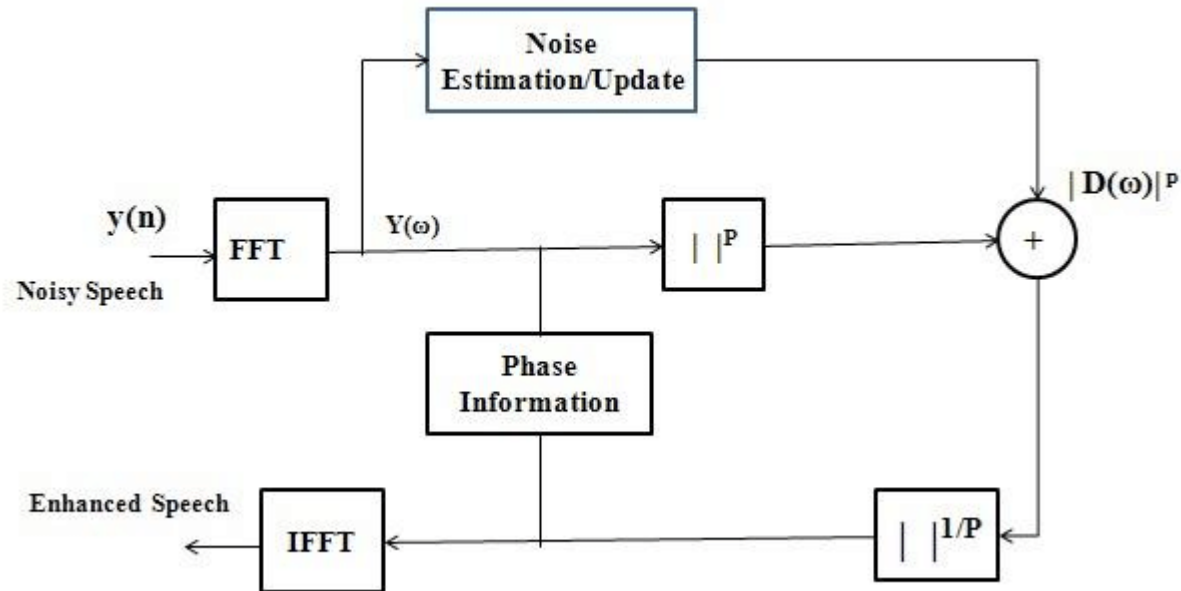


# Ứng dụng trong thiết kế bộ lọc trên miền



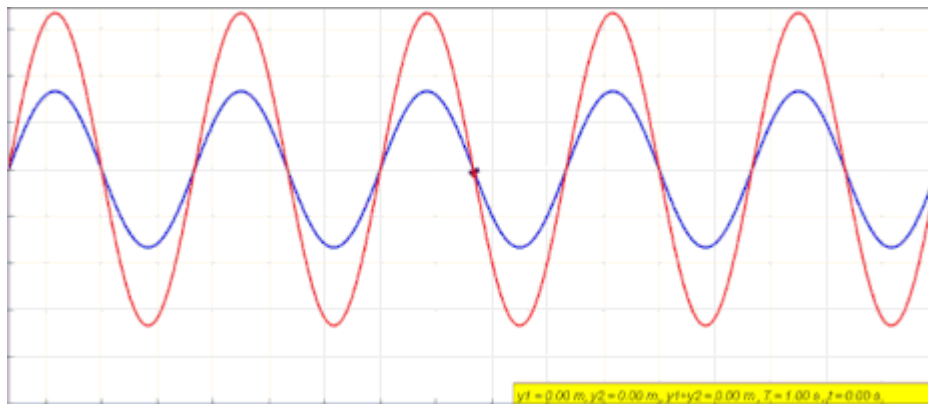
Ví dụ các bộ lọc lý tưởng

# Ứng dụng trong lọc nhiễu



# Sóng dừng/đứng

Wiki: Sóng đứng là một loại sóng dao động theo thời gian nhưng có biên độ đỉnh sóng không di chuyển trong không gian



Thí nghiệm về sóng đứng: <https://www.youtube.com/watch?v=-gr7KmT0rx0>

# Sóng dừng/đứng - Dây thanh quản

Nội soi cách phát âm:

<https://youtu.be/-1Fv7IPmJo4?t=92>

Vocal fold vibration - slow motion

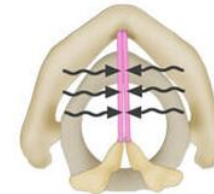
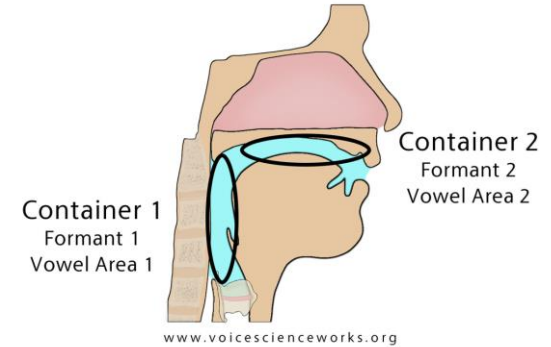
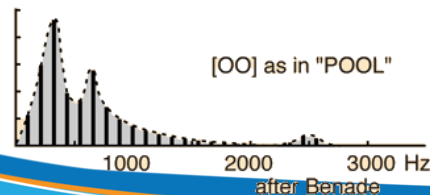
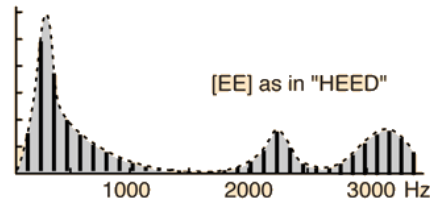
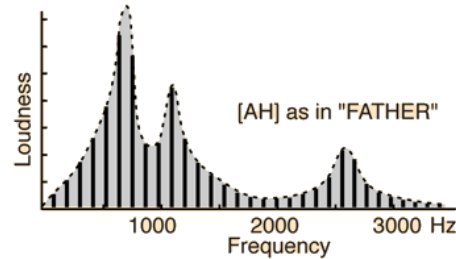
[https://www.youtube.com/watch?v=Drns\\_eV9wWg](https://www.youtube.com/watch?v=Drns_eV9wWg)



# Harmonics vs Formants

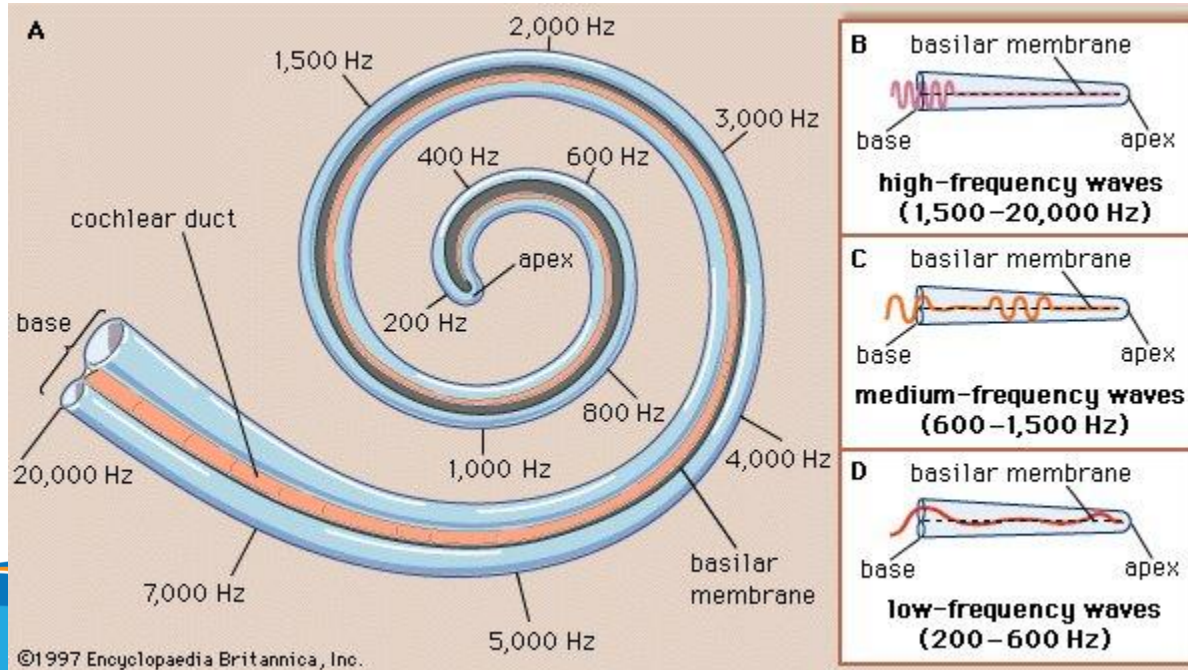
<https://www.voicescienceworks.org/harmonics-vs-formants.html>

<http://hyperphysics.phy-astr.gsu.edu/hbase/Music/vowel.html>



# Inner ear

<https://www.britannica.com/science/inner-ear>





# The mel frequency scale

$$\text{mel} = 2595 * \log_{10} (1 + \text{hertz} / 700)$$

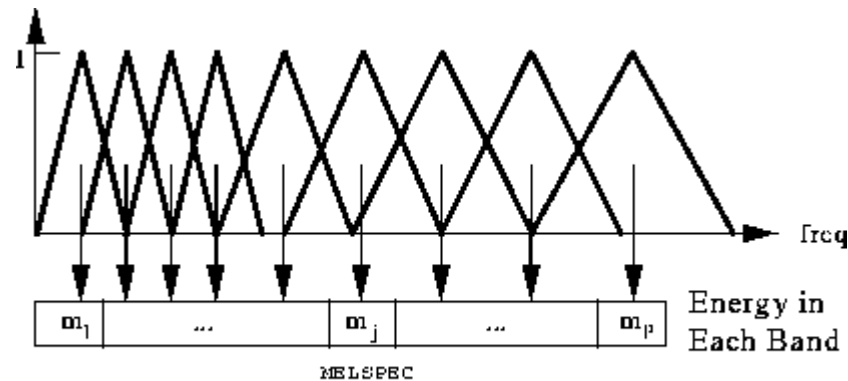


Fig. 5.3 Mel-Scale Filter Bank

# Bài tập về nhà

Khảo sát đặc trưng Mel spectrogram

<https://librosa.org/doc/main/generated/librosa.feature.melspectrogram.html#librosa.feature.melspectrogram>

