AI in Audio and Signal Processing

Instructor: Mehdi TOUMI

Course Overview

- **Course Structure:**
  - Lectures: two 4-hour sessions
  - 3 Major Projects
- **Objectives:**
  - Understand the fundamental principles of signal and audio processing
  - Identify the challenges and opportunities of applying AI to signal and audio processing
  - Describe and apply common AI techniques for signal and audio processing
  - Design and implement AI systems for specific signal and audio processing tasks
  - Evaluate the performance of AI systems for signal and audio processing

Syllabus Review

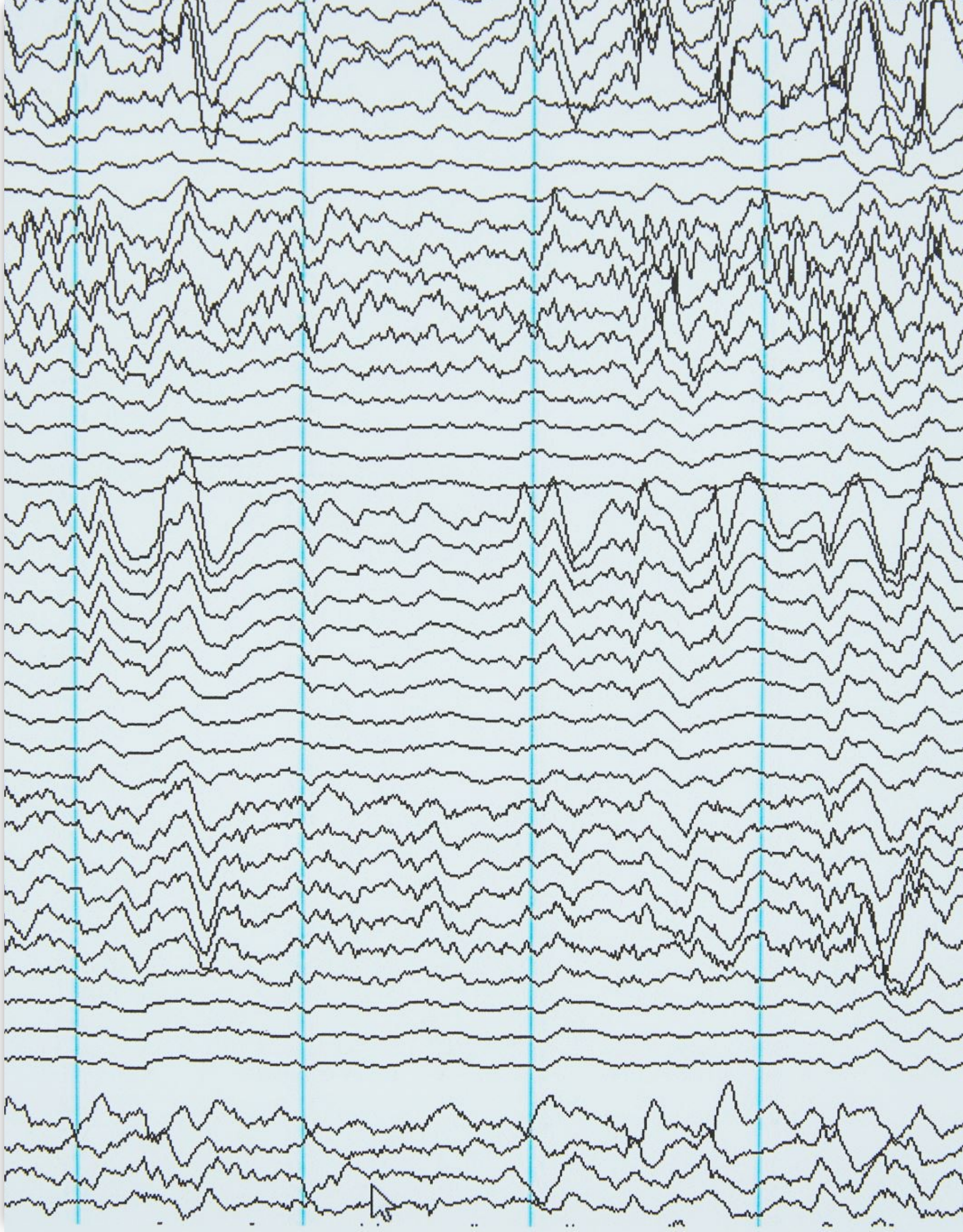
- **Topics Covered:**
  - Basics of signal processing
  - Audio signal characteristics
  - Machine learning and deep learning techniques
  - Applications in music information retrieval, and more
- **Grading:**
  - Projects: 90%
  - Attendance and Participation: 10%
- **Key Dates:** Project deadlines

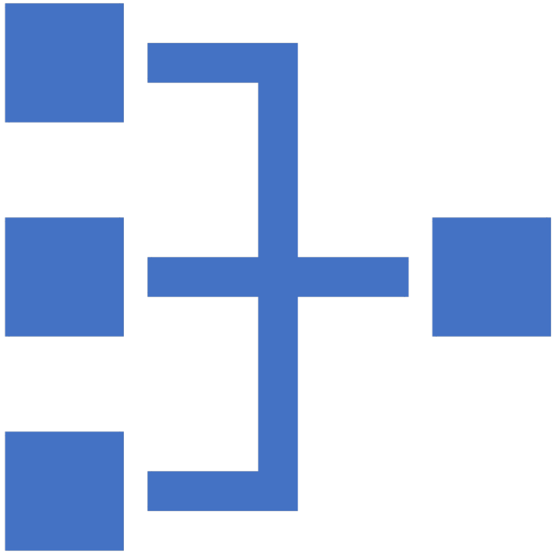
Signal Processing

- What is signal processing

- Types of Signals : Continuous (Analog) vs Discrete (Digital)

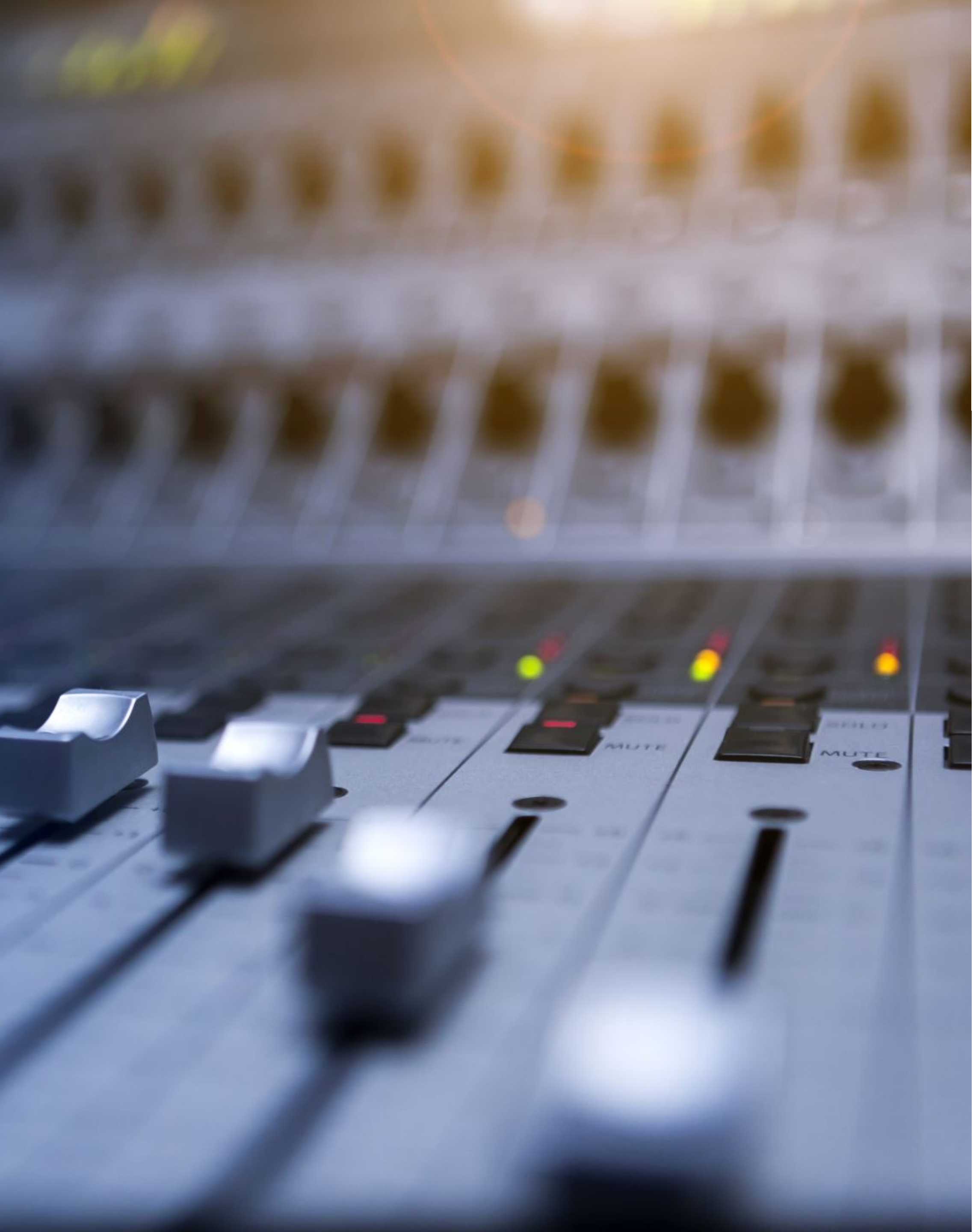
- Basic Properties of a signal:
    1. Amplitude
    2. Frequency
    3. Phase
    4. Periodicity





Key Concepts in Signal Processing

- Fourier Theorem and Fourier Transform
- Nyquist Rate/Theorem and Aliasing
- Filtering
- Convolution



Basics of Audio Signals

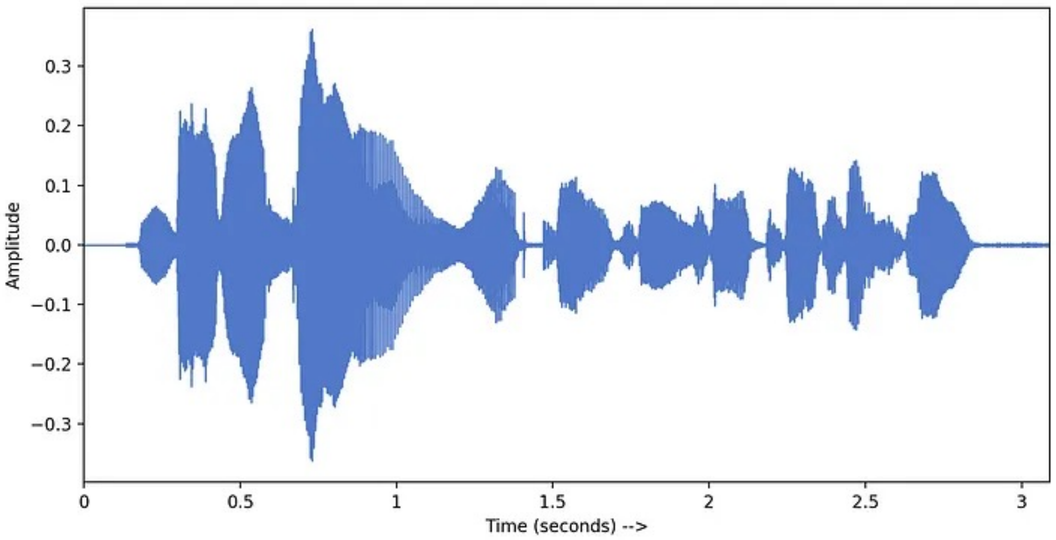
- Nature of Sound:

Audio signal representation

Time-Domain Representation:

Shows how the signal changes over time

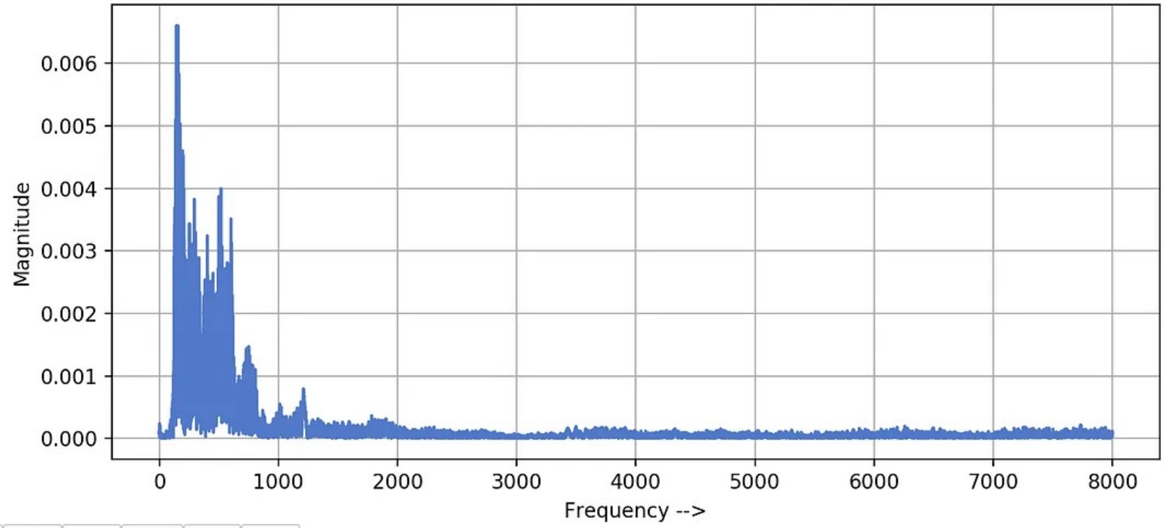
Example: Waveforms



Frequency-Domain Representation:

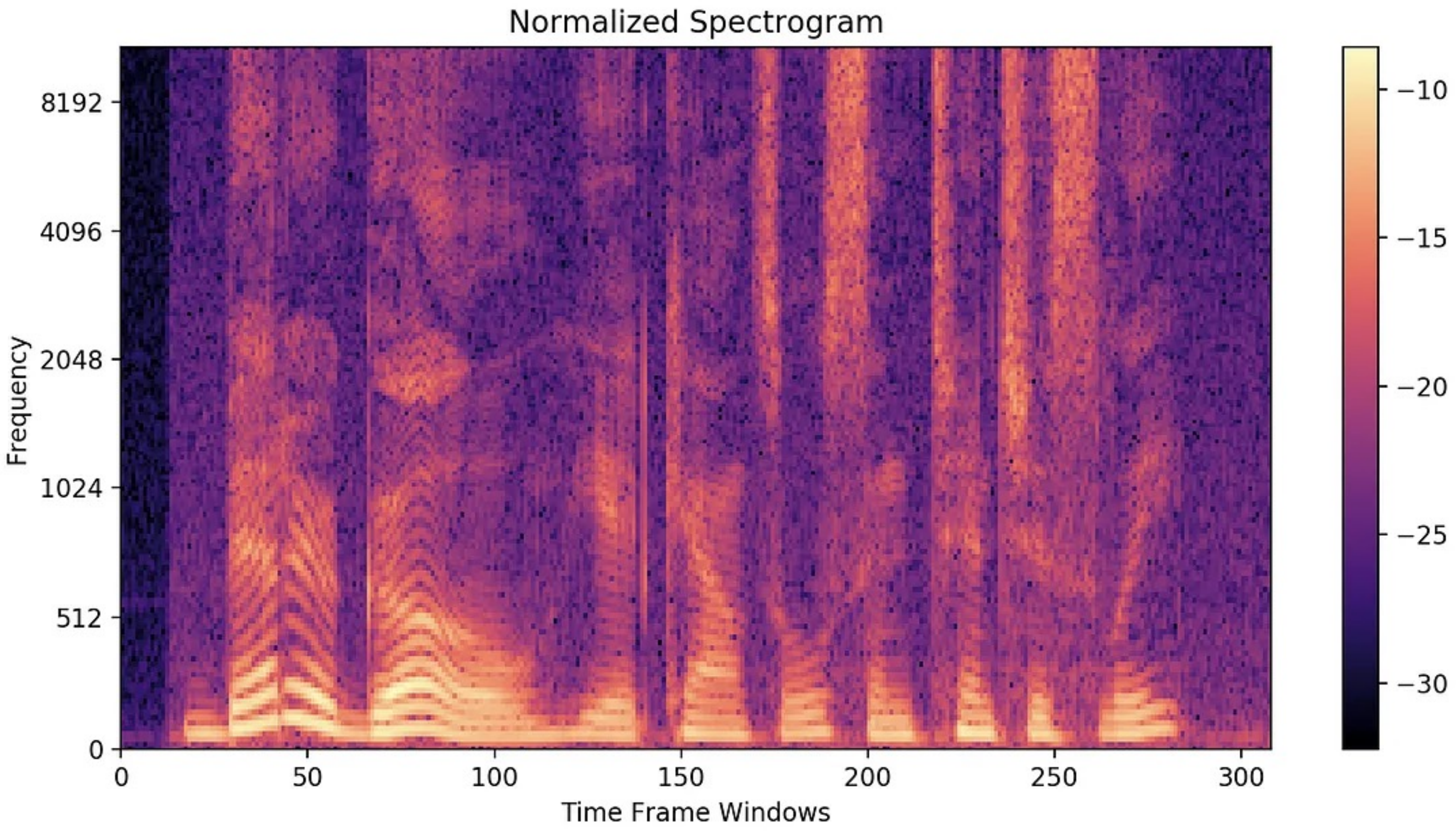
Shows the signal's frequency components

Example: Fourier Transform



Pictures by [TWD]

Spectrograms



Beneficial at speech recognition

Pictures by [TWD]

Human Auditory System:

- **Human Auditory System:**

  Anatomy and Physiology:

  Outer, middle, and inner ear

  How the ear processes sound waves

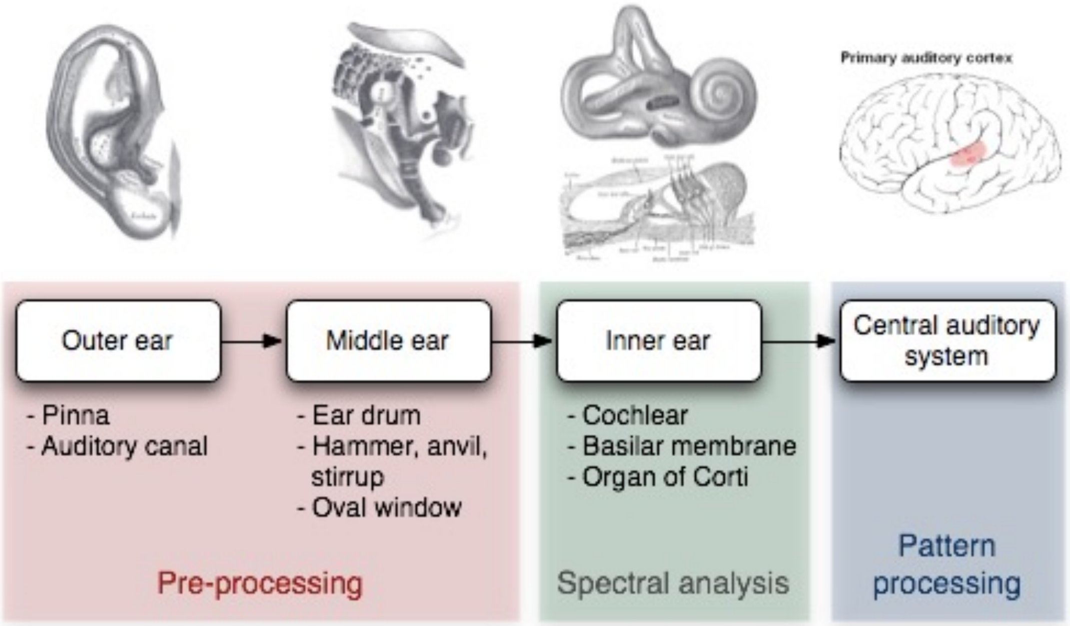
- **Common Audio Signal Processing Tasks:**

  Audio Filtering:

  Low-pass, high-pass, band-pass, and band-stop filters

  Noise Reduction:

  Techniques for reducing noise in audio signals



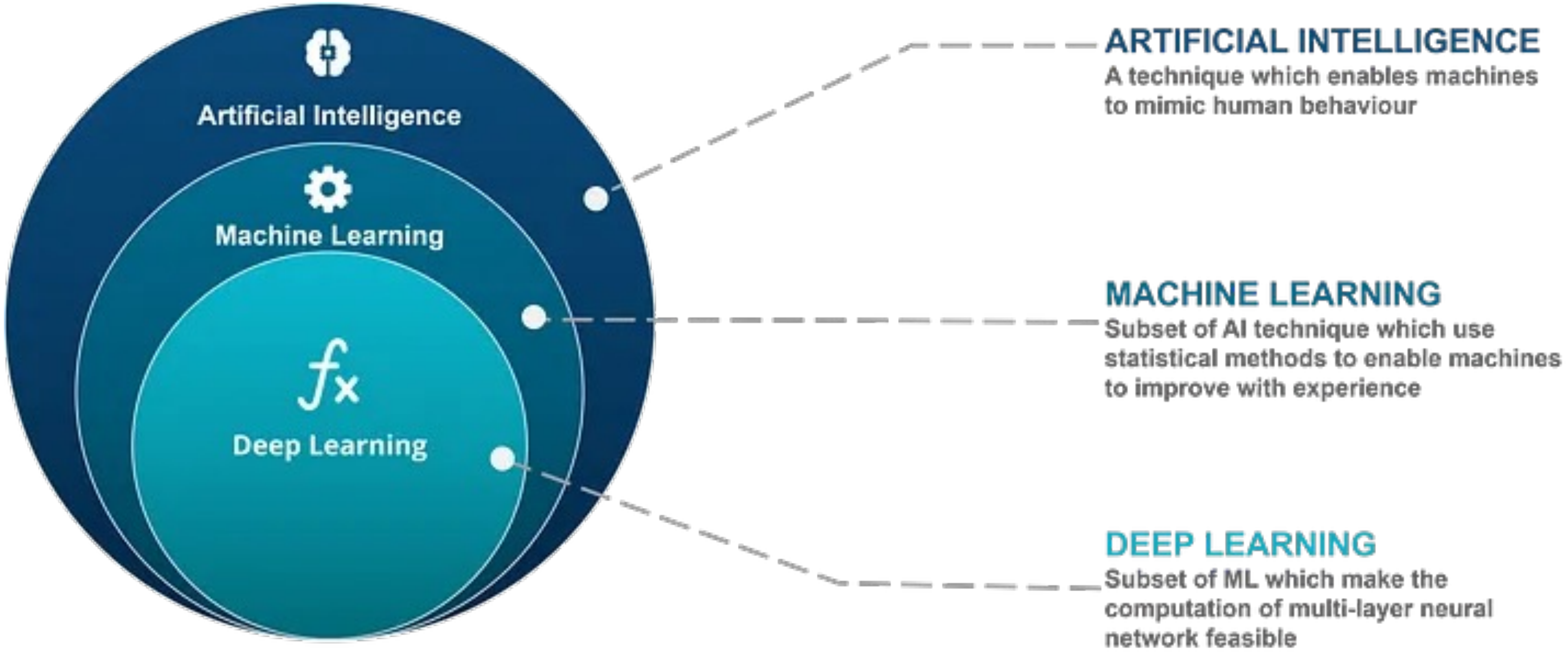
Picture by [TWD]

Introduction to AI Applications in Audio Processing

- Objectives:
  - Overview of AI and its relevance to audio signal processing
  - Introduction to basic AI techniques used in audio processing
  - Discussion of real-world applications

- AI overview:

  Different branches of AI



[Illustrated by Medium]()

Basic AI Techniques in Audio Processing:

Machine Learning Algorithms:

Supervised learning, unsupervised learning, reinforcement learning

Common algorithms (SVM, k-NN, decision trees)

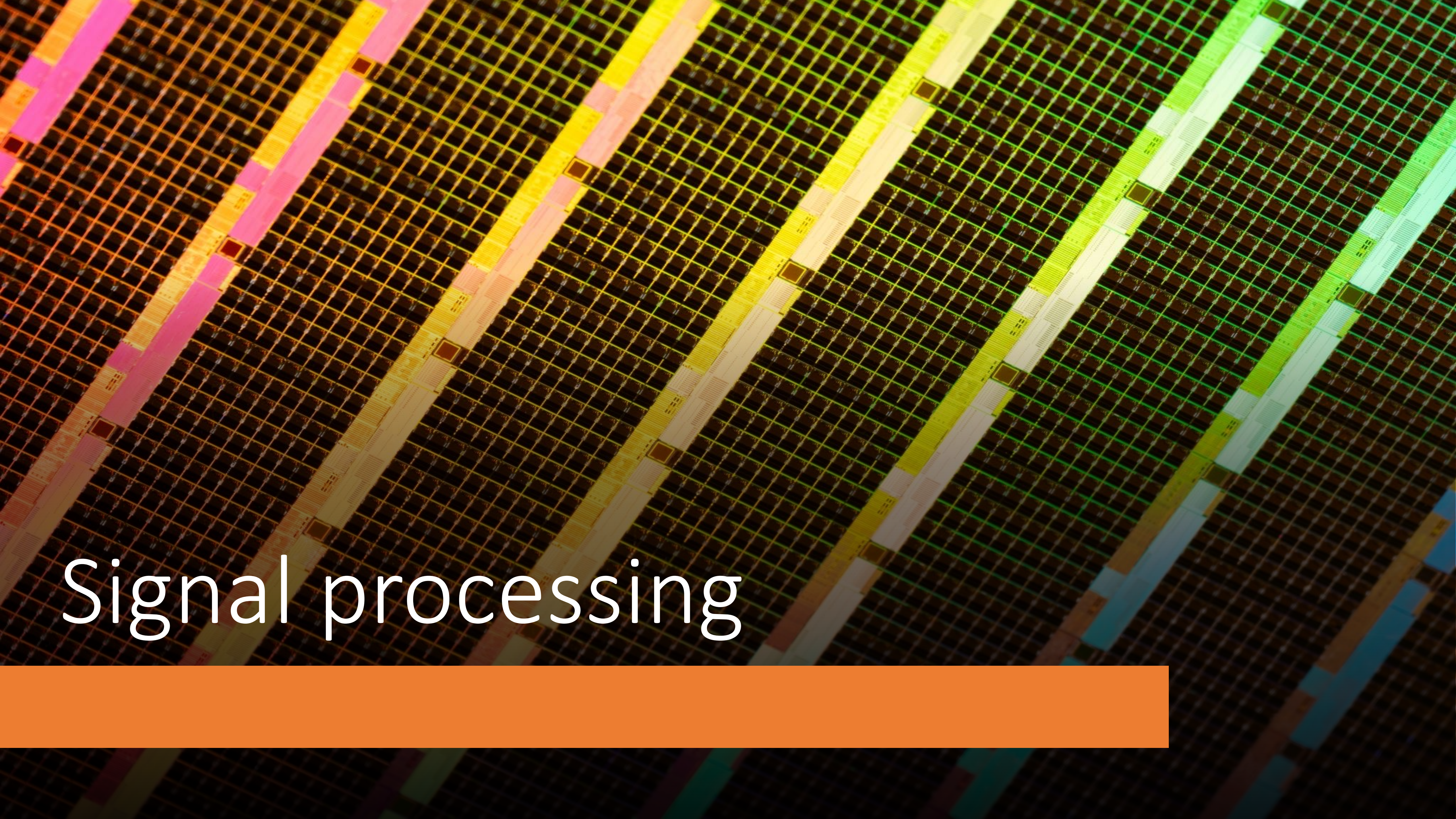
Classification vs regression

Deep Learning Architectures: Neural networks, CNNs, RNNs

Introduction to frameworks (Python, TensorFlow, Keras, Matlab, Google colab)

Real-World Applications:

- **Speech Recognition:**
  - Overview of ASR systems
  - Examples (Siri, Google Assistant)
- **Music Information Retrieval:**
  - Genre classification, artist identification
  - Examples (Spotify, Shazam)
- **Audio Enhancement:**
  - Noise reduction, audio restoration



Signal processing

Fourier Theorem

- The Fourier theorem states that any periodic function f(t) with period T can be represented as a sum of sine and cosine functions (or equivalently, complex exponentials). This representation is known as the Fourier series. For a function f(t) with period T, the Fourier series is given by:

$$f(t) = a_0 + \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{2\pi nt}{T}\right) + b_n \sin\left(\frac{2\pi nt}{T}\right)\right)$$

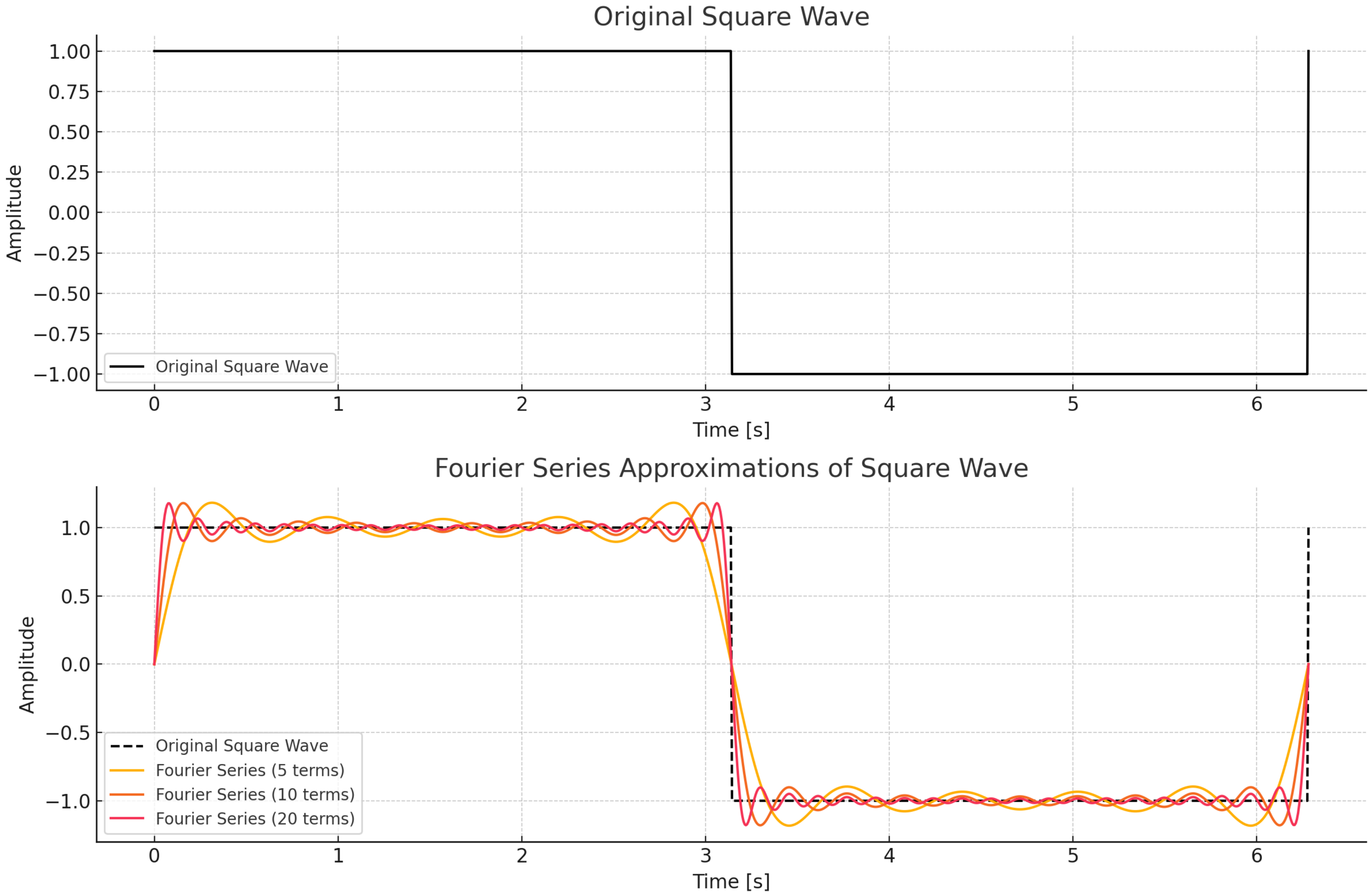
- Alternatively, using Euler's formula, it can be written in terms of complex exponentials:

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{i\frac{2\pi nt}{T}}$$

- where the coefficients are calculated as follows:

$$a_0 = \frac{1}{T}\int_0^T f(t)\,dt$$
$$a_n = \frac{2}{T}\int_0^T f(t)\cos\left(\frac{2\pi nt}{T}\right)\,dt$$
$$b_n = \frac{2}{T}\int_0^T f(t)\sin\left(\frac{2\pi nt}{T}\right)\,dt$$
$$c_n = \frac{1}{T}\int_0^T f(t)e^{-i\frac{2\pi nt}{T}}\,dt$$

Fourier Theorem



video(0to4:50)

Transition to Fourier Transform

- In the limit as T→∞, these integrals span from −∞ to +∞, and the discrete sum becomes an integral:

$$F(\omega) = \mathcal{F}\{f(t)\} = \int_{-\infty}^{\infty} f(t)e^{-i\omega t}\,dt$$

F(ω) is the Fourier Transform of f(t),

ω is the angular frequency (in radians per second),

i is the imaginary unit.

- The inverse Fourier Transform, which converts the frequency-domain representation back to the time domain, is given by:

$$f(t) = \mathcal{F}^{-1}\{F(\omega)\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(\omega)e^{i\omega t}\,d\omega$$

Fourier Transform Properties

- Linearity: $$\mathcal{F}\{af(t) + bg(t)\} = aF(\omega) + bG(\omega)$$
- Time Shifting: $$\mathcal{F}\{f(t - t_0)\} = F(\omega)e^{-i\omega t_0}$$
- Frequency Shifting: $$\mathcal{F}\{f(t)e^{i\omega_0 t}\} = F(\omega - \omega_0)$$
- Scaling: $$\mathcal{F}\{f(at)\} = \frac{1}{|a|}F\left(\frac{\omega}{a}\right)$$
- Time Reversal: $$\mathcal{F}\{f(-t)\} = F(-\omega)$$
- Convolution: $$\mathcal{F}\{f(t) * g(t)\} = F(\omega)G(\omega)$$
- Modulation: $$\mathcal{F}\{f(t)g(t)\} = \frac{1}{2\pi}(F(\omega) * G(\omega))$$
- Parseval's Theorem: $$\int_{-\infty}^{\infty} |f(t)|^2 \, dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} |F(\omega)|^2 \, d\omega$$

Derivations

$$\mathcal{F}\{af(t)+bg(t)\} = \int_{-\infty}^{\infty}[af(t)+bg(t)]e^{-i\omega t}\,dt$$
$$= a\int_{-\infty}^{\infty} f(t)e^{-i\omega t}\,dt + b\int_{-\infty}^{\infty} g(t)e^{-i\omega t}\,dt$$
$$= aF(\omega) + bG(\omega)$$

$$\mathcal{F}\{f(t-t_0)\} = \int_{-\infty}^{\infty} f(t-t_0)e^{-i\omega t}\,dt$$

Let $u = t - t_0$, then $du = dt$:

$$= \int_{-\infty}^{\infty} f(u)e^{-i\omega(u+t_0)}\,du$$
$$= e^{-i\omega t_0}\int_{-\infty}^{\infty} f(u)e^{-i\omega u}\,du$$
$$= e^{-i\omega t_0}F(\omega)$$

$$\mathcal{F}\{e^{i\omega_0 t}f(t)\} = \int_{-\infty}^{\infty} e^{i\omega_0 t}f(t)e^{-i\omega t}\,dt$$
$$= \int_{-\infty}^{\infty} f(t)e^{-i(\omega-\omega_0)t}\,dt$$
$$= F(\omega-\omega_0)$$

$$\mathcal{F}\{f(at)\} = \int_{-\infty}^{\infty} f(at)e^{-i\omega t}\,dt$$

Let $u = at$, then $du = a\,dt$ or $dt = \frac{du}{a}$:

$$= \int_{-\infty}^{\infty} f(u)e^{-i\omega\frac{u}{a}}\frac{du}{a}$$
$$= \frac{1}{|a|}\int_{-\infty}^{\infty} f(u)e^{-i\frac{\omega}{a}u}\,du$$
$$= \frac{1}{|a|}F\left(\frac{\omega}{a}\right)$$

$$\cos(\omega_0 t) = \frac{1}{2}(e^{i\omega_0 t} + e^{-i\omega_0 t})$$

Then,

$$f(t)\cos(\omega_0 t) = f(t)\cdot\frac{1}{2}(e^{i\omega_0 t} + e^{-i\omega_0 t})$$

Using linearity and the frequency shift property:

$$\mathcal{F}\{f(t)\cos(\omega_0 t)\} = \frac{1}{2}\left(\mathcal{F}\{f(t)e^{i\omega_0 t}\} + \mathcal{F}\{f(t)e^{-i\omega_0 t}\}\right)$$
$$= \frac{1}{2}\left(F(\omega-\omega_0) + F(\omega+\omega_0)\right)$$

Usual Fourier Transforms and Their Derivations

- Integral and Sifting Properties of a Dirac function:

$$\int_{-\infty}^{\infty} \delta(t)\, dt = 1$$

$$\int_{-\infty}^{\infty} f(t)\delta(t - t_0)\, dt = f(t_0)$$

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)\, d\tau$$

Delta Function $\delta(t)$:

$$\mathcal{F}\{\delta(t)\} = 1$$

Derivation:

By definition, the delta function picks out the value of the integrand at zero:

$$\int_{-\infty}^{\infty} \delta(t)e^{-i\omega t}\, dt = e^0 = 1$$

Usual Fourier Transforms and Their Derivations

Exponential $e^{i\omega_0 t}$:

$$\mathcal{F}\{e^{i\omega_0 t}\} = 2\pi\delta(\omega - \omega_0)$$

Derivation:

$$\int_{-\infty}^{\infty} e^{i\omega_0 t} e^{-i\omega t}\, dt = \int_{-\infty}^{\infty} e^{i(\omega_0 - \omega)t}\, dt = 2\pi\delta(\omega - \omega_0)$$

Rectangular Pulse $\Pi(t)$:

$$\mathcal{F}\{\Pi(t)\} = \frac{\sin(\omega/2)}{\omega/2}$$

Derivation:

A rectangular pulse can be represented as:

$$\Pi(t) = \begin{cases} 1 & |t| \le \frac{1}{2} \\ 0 & \text{otherwise} \end{cases}$$

The Fourier Transform is:

$$\int_{-1/2}^{1/2} e^{-i\omega t}\, dt = \left[\frac{e^{-i\omega t}}{-i\omega}\right]_{-1/2}^{1/2} = \frac{\sin(\omega/2)}{\omega/2}$$

The Fourier Transform of a sine wave sin(ω0t) is:

$$F(\omega) = \pi\left(\delta(\omega - \omega_0) - \delta(\omega + \omega_0)\right)$$

$$F(\omega) = \mathcal{F}\{\sin(\omega_0 t)\} = \int_{-\infty}^{\infty} \sin(\omega_0 t) e^{-i\omega t}\, dt$$

Steps to Derive the Fourier Transform:

1. **Express the Sine Wave using Euler's Formula:**

$$\sin(\omega_0 t) = \frac{e^{i\omega_0 t} - e^{-i\omega_0 t}}{2i}$$

2. **Substitute into the Fourier Transform Integral:**

$$F(\omega) = \int_{-\infty}^{\infty} \frac{e^{i\omega_0 t} - e^{-i\omega_0 t}}{2i} e^{-i\omega t}\, dt$$

This can be split into two integrals:

$$F(\omega) = \frac{1}{2i}\left(\int_{-\infty}^{\infty} e^{i(\omega_0 - \omega)t}\, dt - \int_{-\infty}^{\infty} e^{-i(\omega_0 + \omega)t}\, dt\right)$$

3. **Evaluate the Integrals:**

The integrals are of the form of the Dirac delta function $\delta(\omega)$, which is defined by:

$$\int_{-\infty}^{\infty} e^{i(\omega' - \omega)t}\, dt = 2\pi\delta(\omega' - \omega)$$

Applying this to our integrals:

$$\int_{-\infty}^{\infty} e^{i(\omega_0 - \omega)t}\, dt = 2\pi\delta(\omega_0 - \omega)$$

$$\int_{-\infty}^{\infty} e^{-i(\omega_0 + \omega)t}\, dt = 2\pi\delta(\omega_0 + \omega)$$

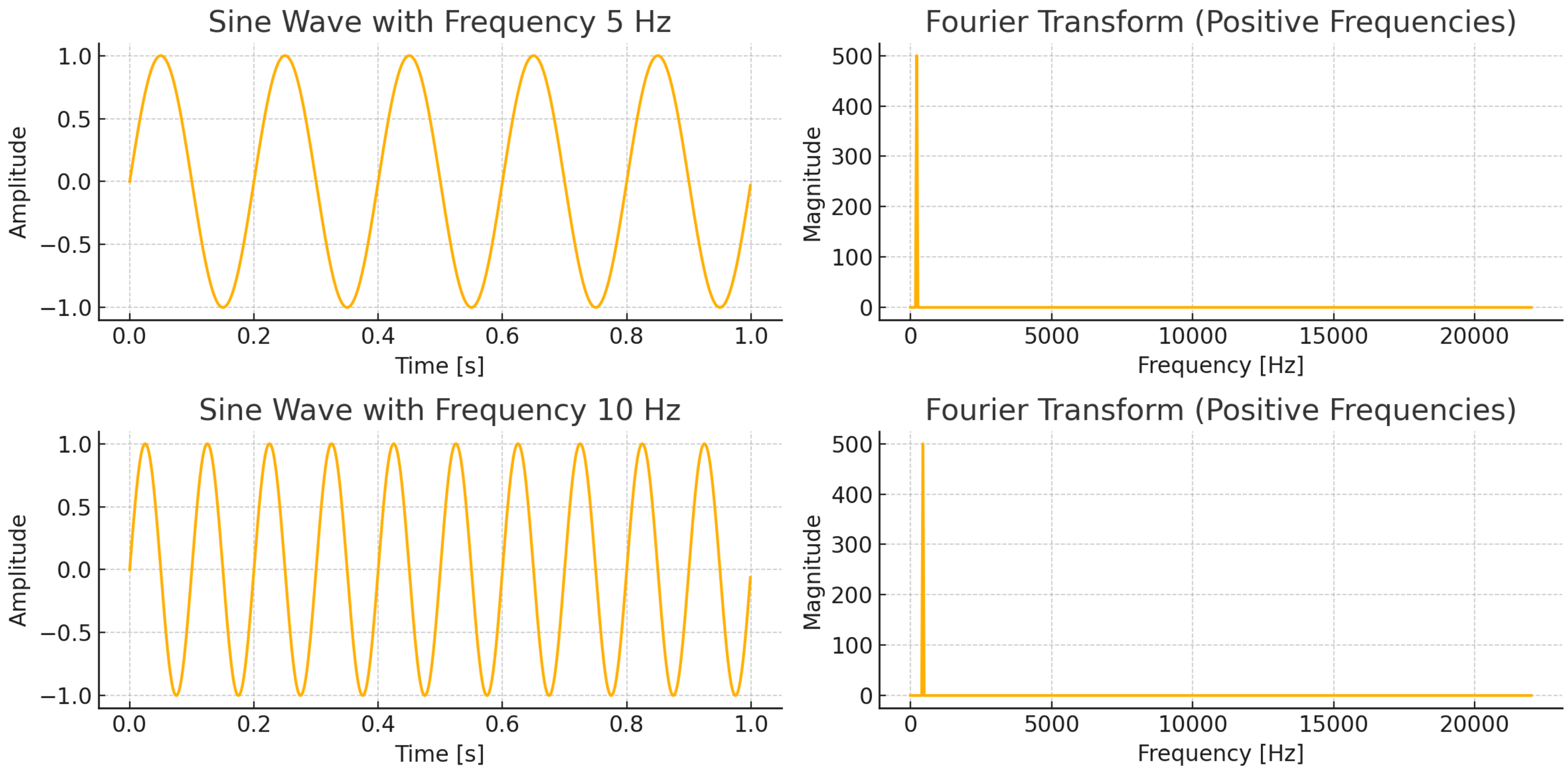
4. **Substitute Back:**

$$F(\omega) = \frac{1}{2i}\left(2\pi\delta(\omega - \omega_0) - 2\pi\delta(\omega + \omega_0)\right)$$

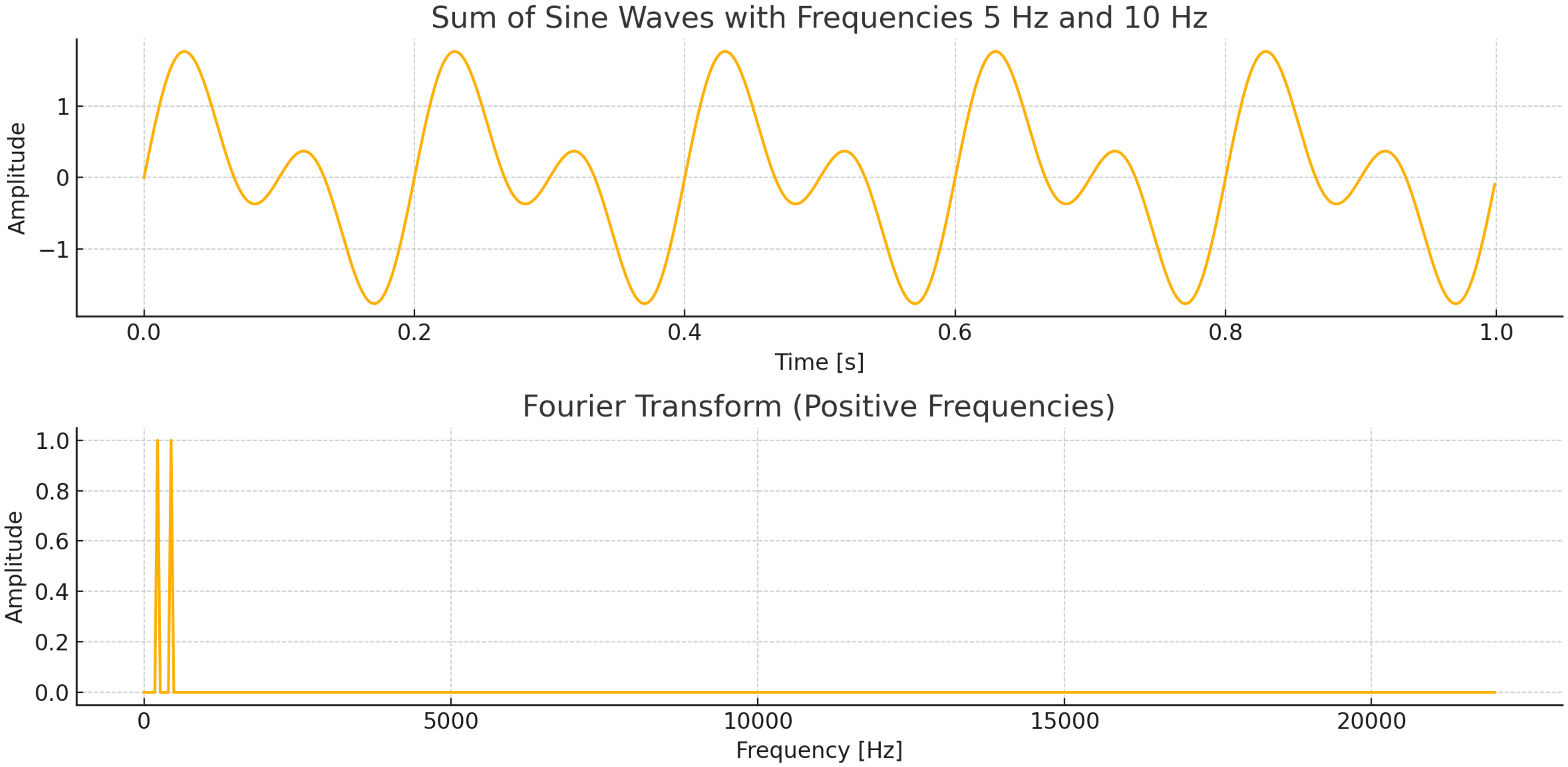
5. **Simplify the Expression:**

$$F(\omega) = \pi\left(\delta(\omega - \omega_0) - \delta(\omega + \omega_0)\right)$$

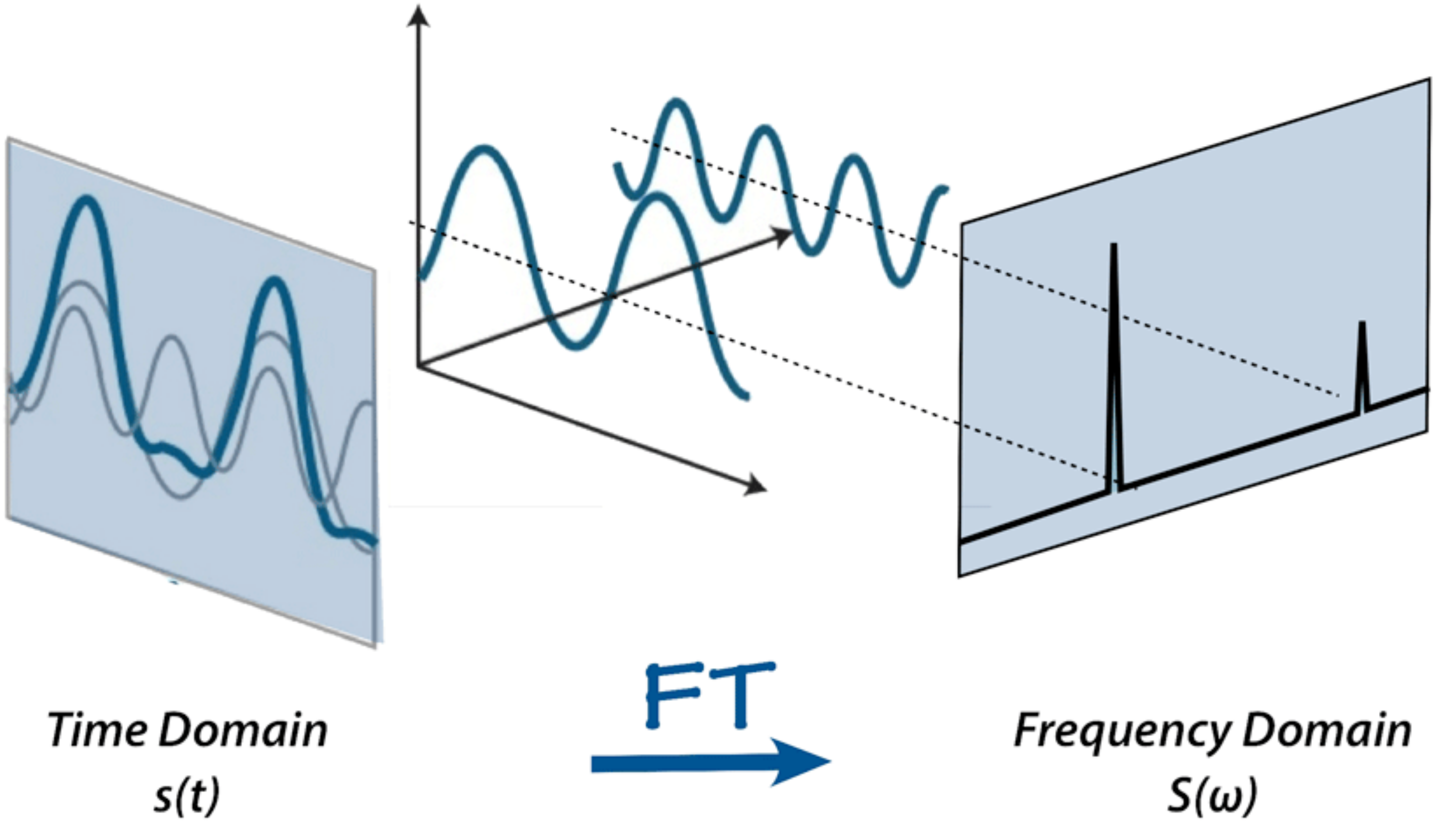
Fourier Transform



Fourier Transform



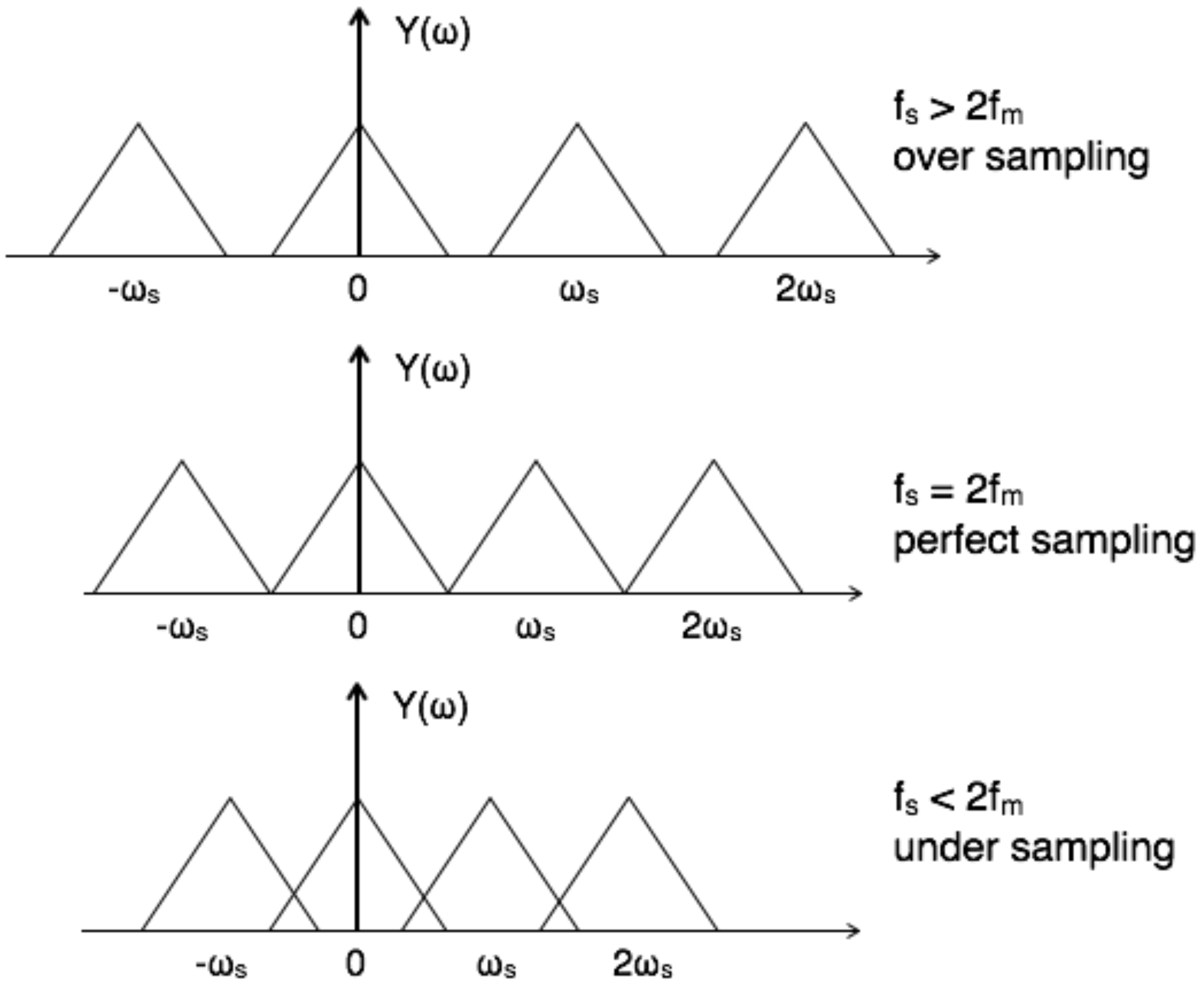
Fourier Transform



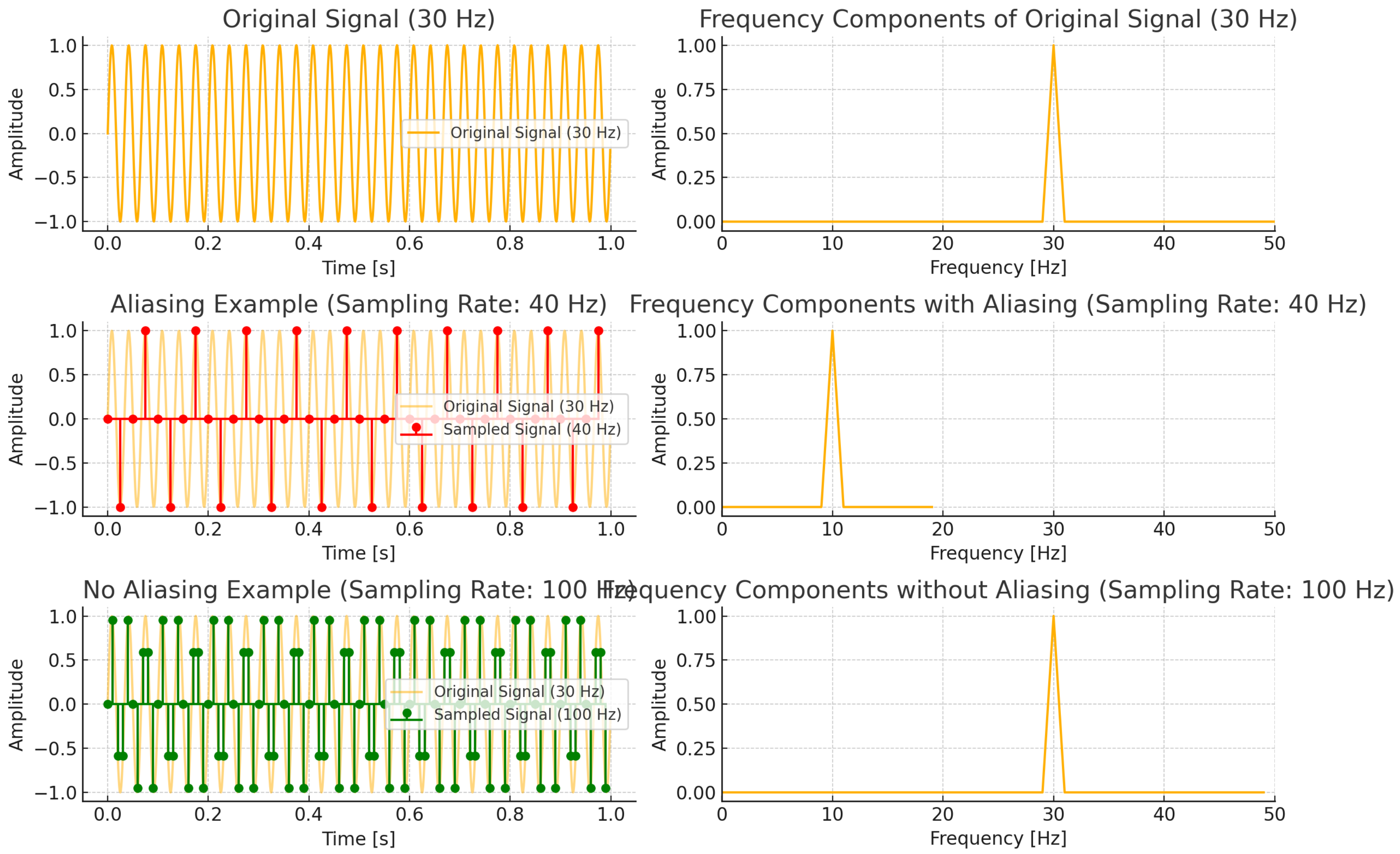
Nyquist Theorem

- The Nyquist theorem states that to accurately sample a signal, the sampling rate must be at least twice the maximum frequency present in the signal. If this condition is not met, aliasing occurs, which results in different signals becoming indistinguishable when sampled.

Sampling rate and Aliasing



Sampling in time domain



Filtering

- **FIR (Finite Impulse Response) Filters**

  Characteristics:

  - **Finite Duration:** The impulse response of an FIR filter settles to zero in a finite amount of time.
  - **Non-Recursive:** FIR filters are typically implemented using a non-recursive structure, meaning that their output depends only on the current and past input values.
  - **Linear Phase:** FIR filters can be designed to have a linear phase response, which means that all frequency components of the input signal are delayed by the same amount, preserving the wave shape of the signal in the passband.

$$y[n] = \sum_{k=0}^{N} b_k x[n-k]$$

y[n] is the output signal at time n,x[n] is the input signal at time n,bk are the feedforward filter coefficients (input weights)

Filtering

- **IIR (Infinite Impulse Response) Filters**

  Characteristics:

  - **Infinite Duration:** The impulse response of an IIR filter theoretically lasts forever because it uses feedback.
  - **Recursive:** IIR filters are typically implemented using a recursive structure, meaning that their output depends on both current and past input values as well as past output values.
  - **Non-Linear Phase:** IIR filters generally do not have a linear phase response, which can lead to phase distortion.

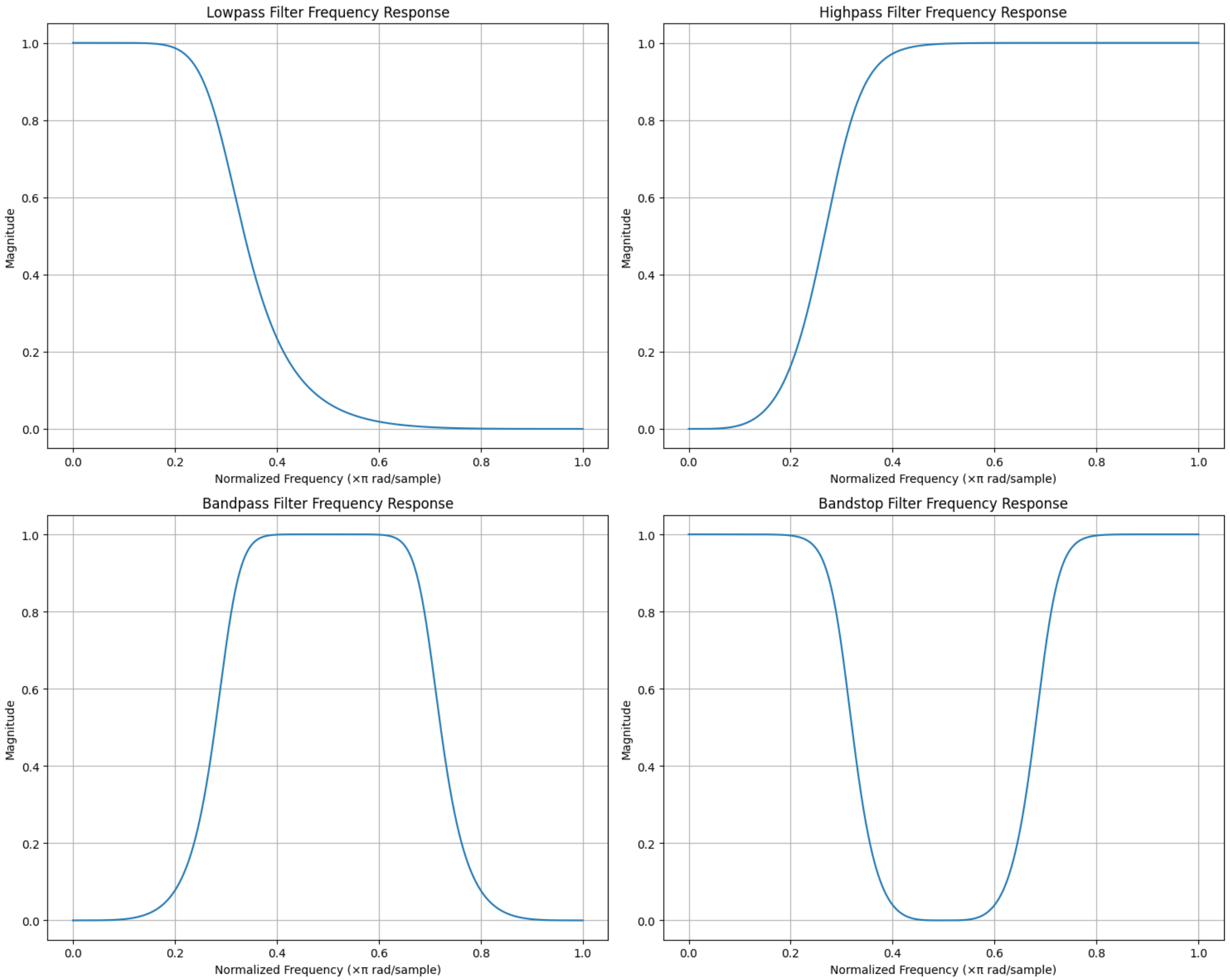
$$y[n] = \sum_{k=0}^{N} b_k x[n-k] - \sum_{j=1}^{M} a_j y[n-j]$$

aj are the feedback filter coefficients (output weights),N is the order of the feedforward part of the filter, M is the order of the feedback part of the filter.

Comparison Table

Property	FIR Filters	IIR Filters
Impulse Response	Finite	Infinite
Implementation	Non-Recursive	Recursive
Stability	Always Stable	Can be Unstable
Phase Response	Can be Linear	Non-Linear
Computational Cost	Higher	Lower
Memory Usage	Higher	Lower

Types of filters (IIR)



Types of filters (FIR)