メディア処理基礎 / Fundamentals of Media Processing

Fundamentals of Signal Processing
Part 1

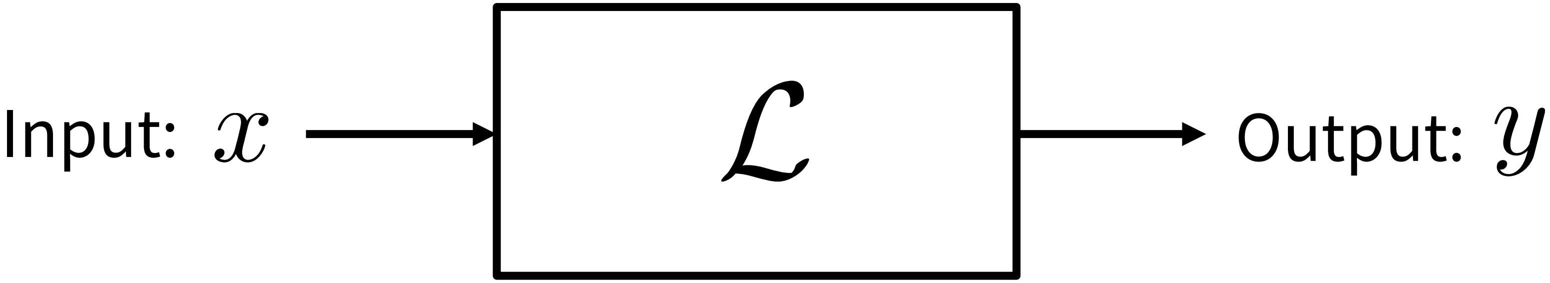
小山 翔一 / Shoichi Koyama

国立情報学研究所 / 総合研究大学院大学　准教授

Associate Professor, National Institute of Informatics / SOKENDAI

What is Signal Processing?

- Techniques for analyzing, modifying, and synthesizing signals, such as sound, images, and others



Generating output y by processing input x using mapping $\mathcal{L}$

See also https://youtu.be/R90ciUoxcJU

What is Signal Processing?

- **Noise reduction**
  - Input: speech contaminated by noise
  - Output: enhanced speech by reducing noise

Noisy speech $\longrightarrow \boxed{\mathcal{L}} \longrightarrow$ Enhanced speech

Extracting speech signal based on properties of speech and/or noise

What is Signal Processing?

- **Speech recognition**
  - Input: Human's speech
  - Output: Spoken text

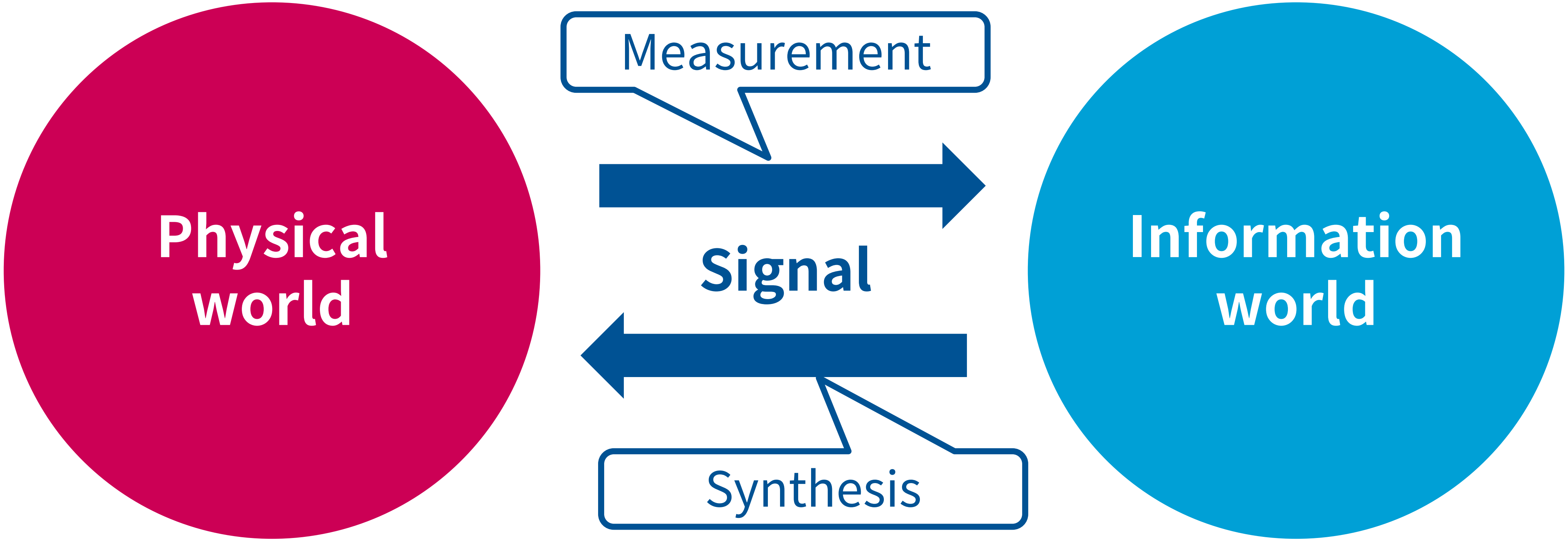
$$\text{Speech} \longrightarrow \boxed{\mathcal{L}} \longrightarrow \text{Text}$$

Conversion from speech to text is learned from a large amount of training data

SIGNAL AND LINEAR TIME-INVARIANT SYSTEM

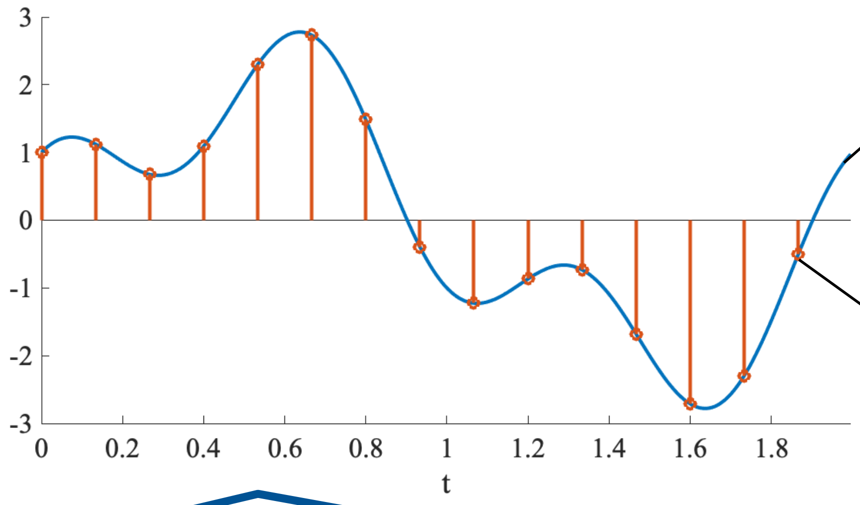
Signal

- Signal is temporal/spatial variations of physical quantities obtained by sensors or their representation by symbols
  - Speech, music, image, video, ultrasonic sonar, radiowave, brainwave, seismic wave, stock price, etc.



Signal

- Signal used in this class is time-series signal: one-dimensional signal of amplitude variation changing with time
  - Continuous-time signal / Analog signal:
    Continuous value of time and amplitude
  - Discrete-time signal / Digital signal:
    Discrete value of time and quantized value of amplitude



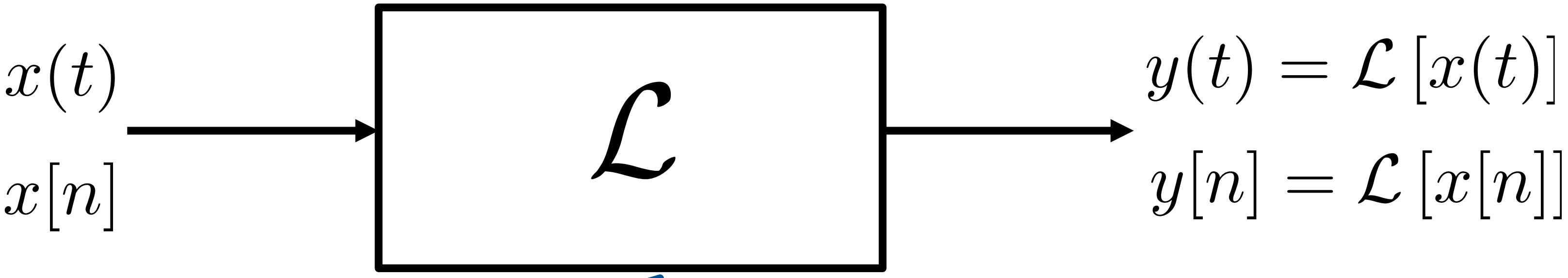
Analog signal $x(t)$ → Function

Digital signal $x[n]$ → Numerical sequence / Vector

Signal processing theory founds its basis on wide variety of mathematics

System

- System: Representation of signal processing stages and input-output characteristics



Converting input to output by mapping $\mathcal{L}$

Linear time-invariant system

- Focusing on linear time-invariant (LTI) system
- **Linearity**:
  - Superposition principle holds

$$\mathcal{L}\left[\alpha x[n] + \beta y[n]\right] = \alpha\mathcal{L}\left[x[n]\right] + \beta\mathcal{L}\left[y[n]\right]$$

$$\forall \alpha, \beta \in \mathbb{C}$$

- **Time-invariance** / **Shift-invariance**:
  - System is consistent with time change

$$y[n] = \mathcal{L}\left[x[n]\right] \Rightarrow y[n-m] = \mathcal{L}\left[x[n-m]\right], \ \forall m$$

Input-output characteristics of LTI system can be decomposed into basic elements for analysis

Impulse response

- Definition of impulse response
  - Output of LTI system $h[n]$ when input is delta function $\delta[n]$

$$h[n] = \mathcal{L}\left[\delta[n]\right]$$

$$\delta[n] \longrightarrow \boxed{\mathcal{L}} \longrightarrow h[n]$$

Here,

$$\delta[n] = \begin{cases} 1, & n = 0 \\ 0, & \text{otherwise} \end{cases}$$

Impulse response

LTI system characteristics are fully described by impulse response

- When impulse response of LTI system is $h[n]$, input signal $x[n]$ and output signal $y[n]$ have the following relationship:

$$y[n] = h[n] * x[n] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$

  - This operation is called convolution

Output signal of any input signal for LTI system can be computed if its impulse response is known

Impulse response

- Arbitrary signal is written by weighted sum of delta function

$$
\begin{aligned}
x[n] * \delta[n] &= \sum_{m=-\infty}^{\infty} x[m]\delta[n-m] \left(= \sum_{m=-\infty}^{\infty} \delta[m]x[n-m] \right) \\
&= \ldots + x[n-1]\delta[1] + x[n]\delta[0] + x[n+1]\delta[-1] + \ldots \\
&= x[n]
\end{aligned}
$$

- Thus,

$$
\begin{aligned}
y[n] &= \mathcal{L}\left[x[n]\right] \\
&= \mathcal{L}\left[\sum_{m=-\infty}^{\infty} x[m]\delta[n-m]\right] \\
&= \sum_{m=-\infty}^{\infty} x[m]\mathcal{L}\left[\delta[n-m]\right] \\
&= \sum_{m=-\infty}^{\infty} x[m]h[n-m] \\
&= x[n] * h[n]
\end{aligned}
$$

Equation above

$x[m]$ does not depend on n

$h[n] = \mathcal{L}\left[\delta[n]\right]$

Impulse response

- In continuous case, input signal $x(t)$ and output signal $y(t)$ are related by convolution with impulse response of LTI system $h(t)$

$$y(t) = h(t) * x(t) = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)\mathrm{d}\tau$$

- Impulse response $h(t)$ is output of LTI system when input is delta function $\delta(t)$

$$h(t) = \mathcal{L}[\delta(t)]$$

where $$\delta(t) = \begin{cases} \infty, & t = 0 \\ 0, & t \neq 0 \end{cases}$$

Not strictly correct representation

Convolution

- Convolution is operation to obtain function/sequence from two functions/sequences
  - Continuous system:

$$y(t) = h(t) * x(t) = \int_{-\infty}^{\infty} h(\tau)x(t-\tau)\mathrm{d}\tau$$
$$= \int_{-\infty}^{\infty} h(t-\tau)x(\tau)\mathrm{d}\tau$$

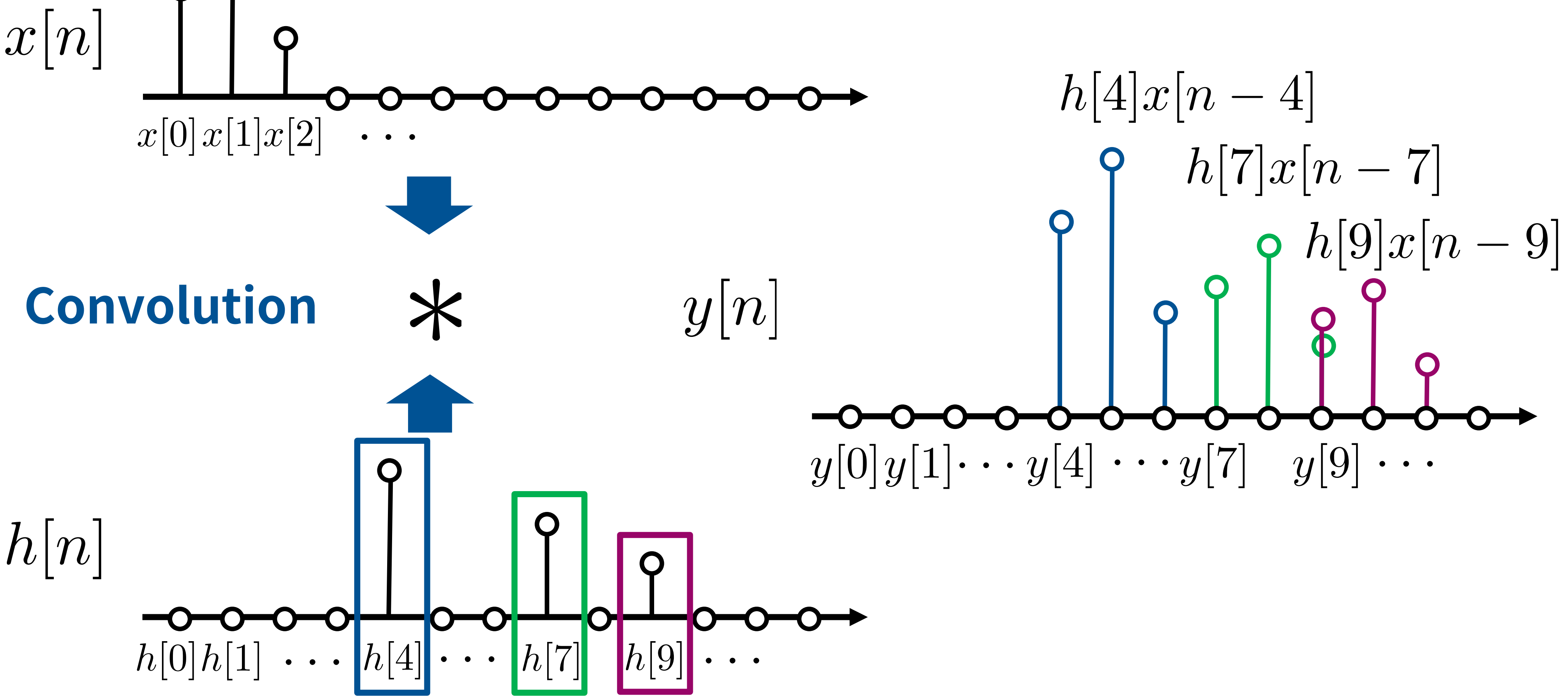
  - Discrete system:

$$y[n] = h[n] * x[n] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$
$$= \sum_{m=-\infty}^{\infty} h[n-m]x[m]$$

Convolution

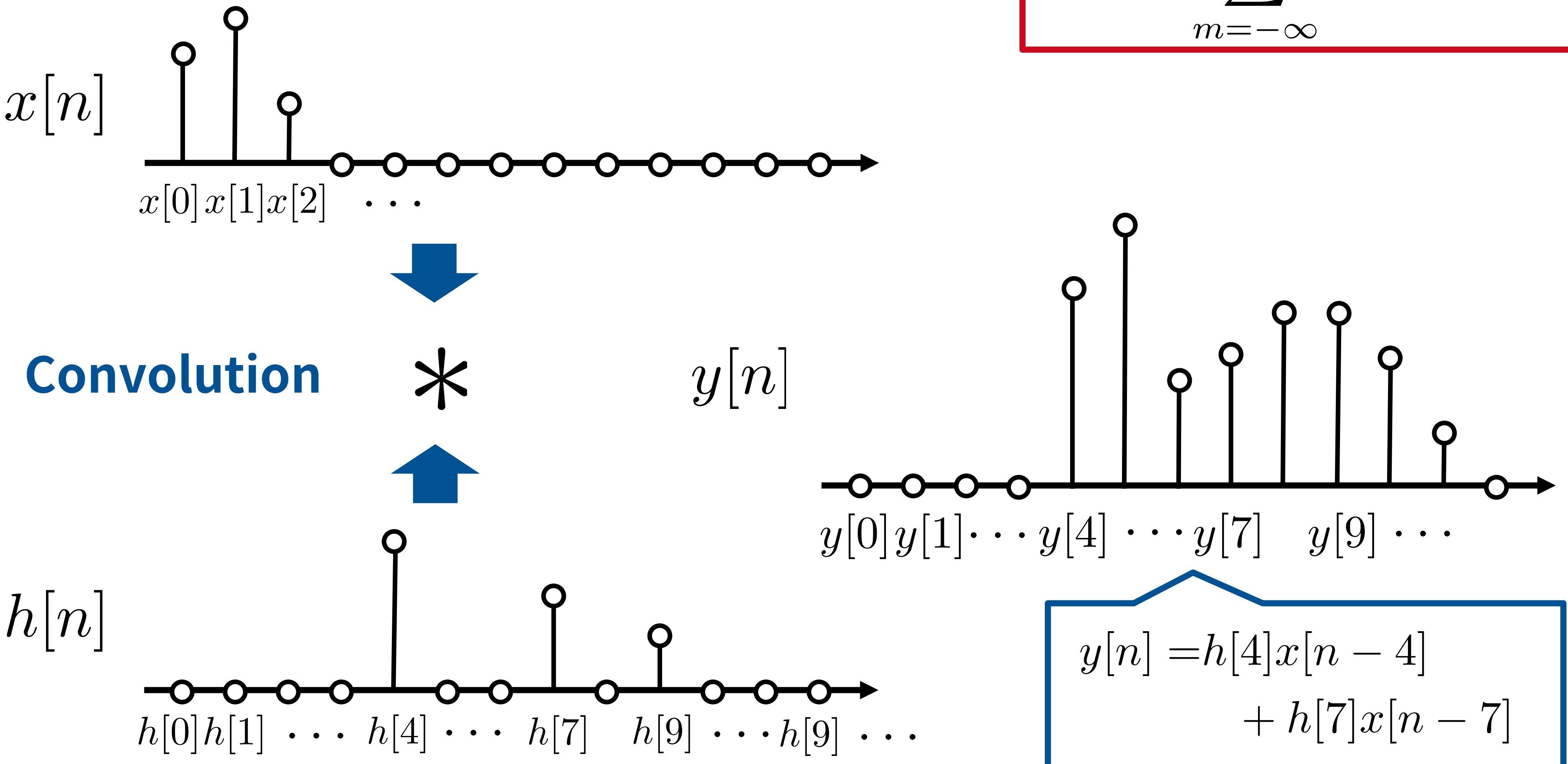
- In discrete case,

$$\boxed{y[n] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]}$$



- In discrete case,

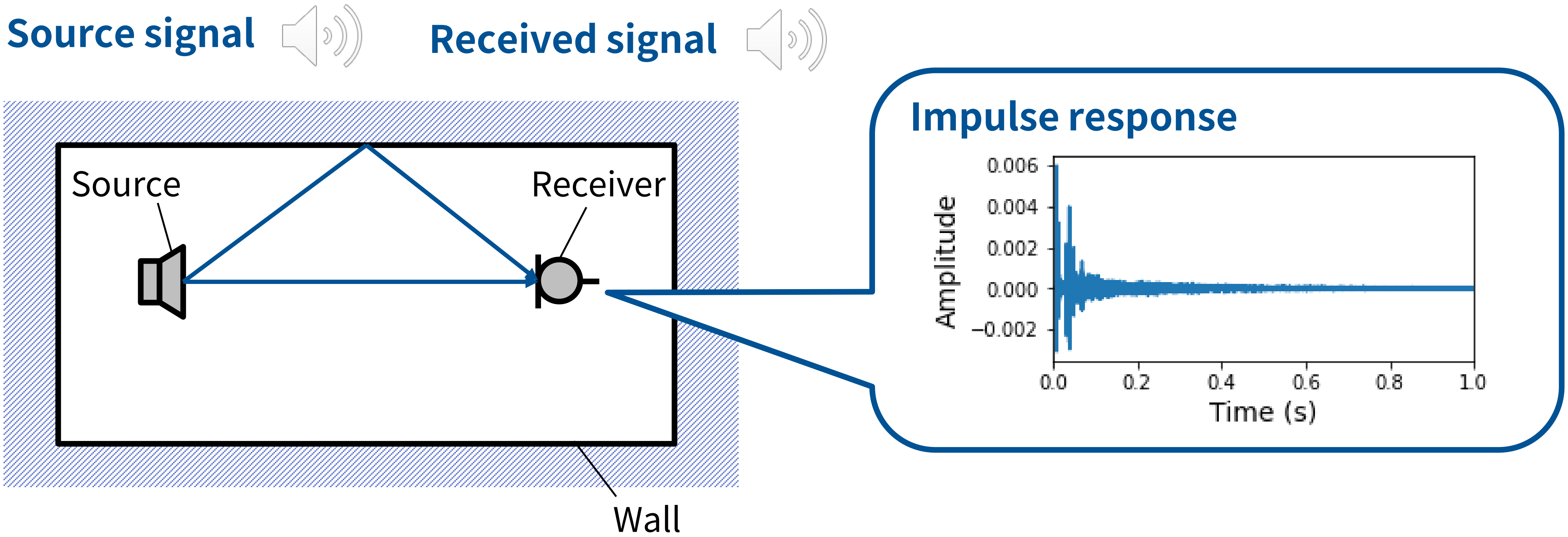
$$y[n] = \sum_{m=-\infty}^{\infty} h[m]x[n-m]$$



$$\begin{aligned} y[n] =\, & h[4]x[n-4] \\ & + h[7]x[n-7] \\ & + h[9]x[n-9] \end{aligned}$$

Convolution in acoustic signal processing

- Transfer characteristics from source (loudspeaker) to receiver (microphone) can be regarded as LTI system
  - If impulse response is measured or predicted in advance, signal at the receiver position from any source signal can be computed
  - Here, impulse response represents characteristics of sound reflections at walls



FOURIER TRANSFORM

Fourier series expansion

- Expansion representation by approximating signal by linear combination of sinusoidal signals

Fourier series expansion

Orthogonal basis expansion of continuous-time periodic signal $x(t)$ with period of T

$$x(t) = \sum_{k=-\infty}^{\infty} c_k \underbrace{\exp\left(\mathrm{j}\frac{2\pi kt}{T}\right)}$$

Complex sine-wave $e^{\mathrm{j}\varphi} = \cos\varphi + \mathrm{j}\sin\varphi$

Here,

$$c_k = \frac{1}{T}\int_{-T/2}^{T/2} x(t)\exp\left(-\mathrm{j}\frac{2\pi kt}{T}\right)\mathrm{d}t$$

(Fourier coefficient)

Fourier series expansion

➢ Another representation of Fourier series expansion

Fourier series expansion (represented by trigonometric functions)

Orthogonal basis expansion of continuous-time periodic signal $x(t)$ with period of T

$$x(t) = a_0 + \sum_{k=1}^{\infty} a_k \cos\left(\frac{2\pi kt}{T}\right) + \sum_{k=1}^{\infty} b_k \sin\left(\frac{2\pi kt}{T}\right)$$

Here,

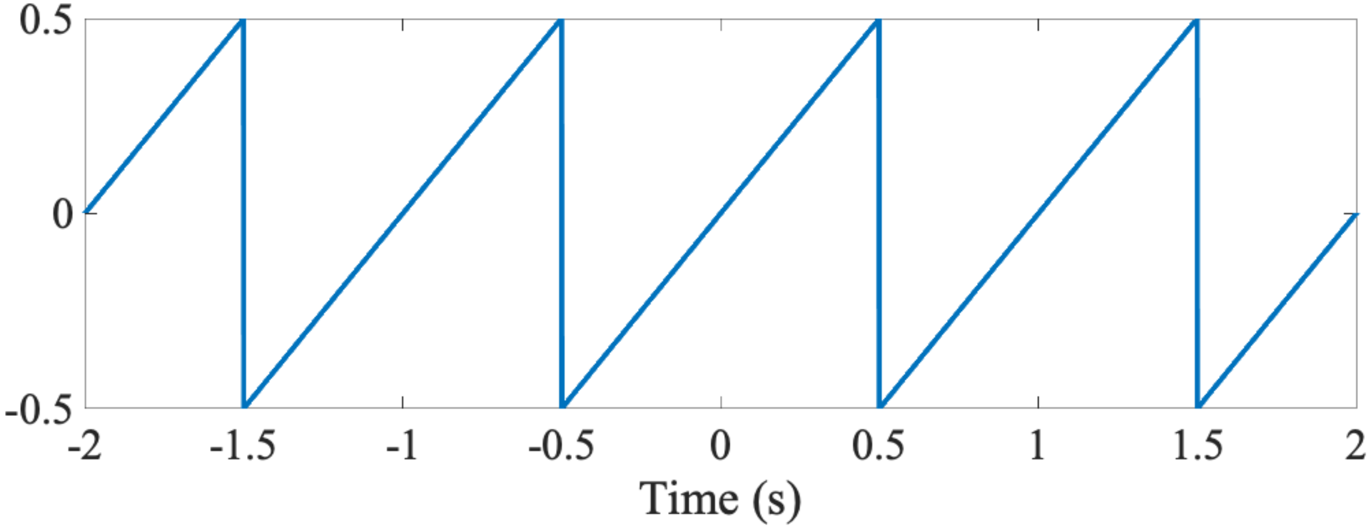
$$a_0 = \frac{1}{T}\int_{-T/2}^{T/2} x(t)\mathrm{d}t$$

$$a_k = \frac{2}{T}\int_{-T/2}^{T/2} x(t)\cos\left(\frac{2\pi kt}{T}\right)\mathrm{d}t$$

$$b_k = \frac{2}{T}\int_{-T/2}^{T/2} x(t)\sin\left(\frac{2\pi kt}{T}\right)\mathrm{d}t$$

- Saw wave

$$x(t) = \begin{cases} t - p, & \left(p - \frac{1}{2}\right) T \le t \le \left(p + \frac{1}{2}\right) T \quad (p \in \mathbb{Z}) \\ 0, & \text{otherwise} \end{cases}$$



Example

$$a_0 = \frac{1}{T}\int_{-T/2}^{T/2} x(t)\mathrm{d}t = \frac{1}{T}\int_{-T/2}^{T/2} t\mathrm{d}t = 0$$

$$a_k = \frac{2}{T}\int_{-T/2}^{T/2} x(t)\cos\left(\frac{2\pi kt}{T}\right)\mathrm{d}t = \frac{2}{T}\int_{-T/2}^{T/2} t\cos\left(\frac{2\pi kt}{T}\right)\mathrm{d}t = 0$$

Odd functions are expanded only by sine function

$$b_k = \frac{2}{T}\int_{-T/2}^{T/2} x(t)\sin\left(\frac{2\pi kt}{T}\right)\mathrm{d}t = \frac{2}{T}\int_{-T/2}^{T/2} t\sin\left(\frac{2\pi kt}{T}\right)\mathrm{d}t$$

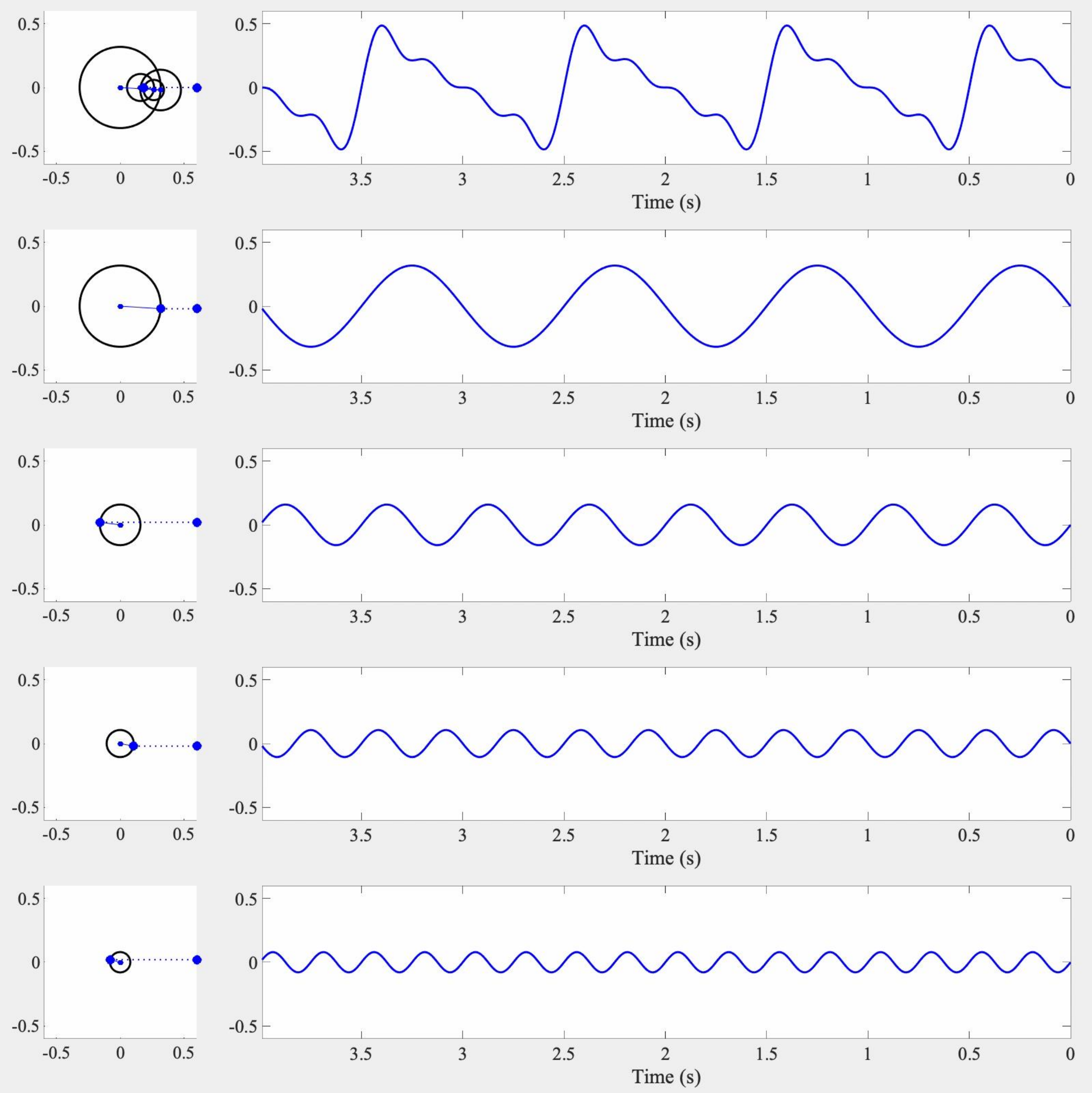
$$= \frac{2}{T}\left\{-\frac{T}{2\pi k}t\cos\left(\frac{2\pi kt}{T}\right)\Big|_{-T/2}^{T/2} + \frac{T}{2\pi k}\int_{-T/2}^{T/2}\cos\left(\frac{2\pi kt}{T}\right)\mathrm{d}t\right\}$$

$$= -\frac{T}{\pi k}\cos(\pi k)$$

$$\Rightarrow \quad x(t) = \sum_{k=1}^{\infty}\left\{-\frac{T}{\pi k}\cos(\pi k)\sin\left(\frac{2\pi kt}{T}\right)\right\}$$

$$= \frac{T}{\pi}\sin\left(\frac{2\pi t}{T}\right) - \frac{T}{2\pi}\sin\left(\frac{4\pi t}{T}\right) + \cdots$$

Example



From Fourier series to Fourier transform

- Fourier series expansion
  - Aimed at approximating signal
  - Only for periodic signals
    - By constraint of periodic signals, signal having uncountably many (i.e., continuous) degrees of freedom is represented by countably many basis functions
  - $x(t)$ and $(c_k)_{k \in \mathbb{Z}}$ are equivalent information if the series converges
    - Just a difference in perspective

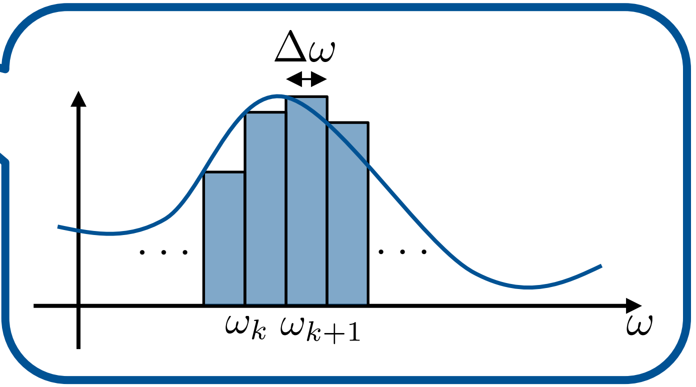
$$\boxed{x(t) = \sum_{k=-\infty}^{\infty} c_k \exp\left(\mathrm{j}\frac{2\pi k t}{T}\right)}$$

From Fourier series to Fourier transform

- Extension of Fourier series expansion to aperiodic signals
  - Replacing with $\Delta\omega = 2\pi/T,\ \omega_k = 2\pi k/T$

$$x(t) = \sum_{k=-\infty}^{\infty} \left[\frac{1}{T} \int_{-T/2}^{T/2} x(t) \exp\left(-\mathrm{j}\frac{2\pi k t}{T}\right) \mathrm{d}t \right] \exp\left(\mathrm{j}\frac{2\pi k t}{T}\right)$$

$$= \frac{1}{2\pi} \underbrace{\sum_{k=-\infty}^{\infty} \left[\int_{-T/2}^{T/2} x(t) \exp\left(-\mathrm{j}\omega_k t\right) \mathrm{d}t \right] \exp\left(\mathrm{j}\omega_k t\right) \Delta\omega}$$



  - When $T \to \infty \ (\Delta\omega \to 0)$

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} x(t) \exp\left(-\mathrm{j}\omega t\right) \mathrm{d}t \right] \exp\left(\mathrm{j}\omega t\right) \mathrm{d}\omega$$

Fourier transform

- Transformation of continuous-time signal into continuous-frequency complex function

- Fourier transform

$$X(\omega) = \int_{-\infty}^{\infty} x(t) \exp\left(-\mathrm{j}\omega t\right) \mathrm{d}t$$

$$(\omega \in \mathbb{R})$$

- Inverse Fourier transform

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega) \exp\left(\mathrm{j}\omega t\right) \mathrm{d}\omega$$

$$(t \in \mathbb{R})$$

Fourier transform

- Notations for Fourier transform and inverse Fourier transform

$$\mathcal{F}[x(t)] = X(\omega)$$

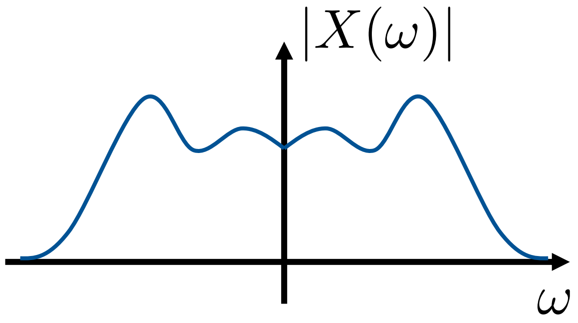
$$\mathcal{F}^{-1}[X(\omega)] = x(t)$$

- Fourier transform pair is denoted as

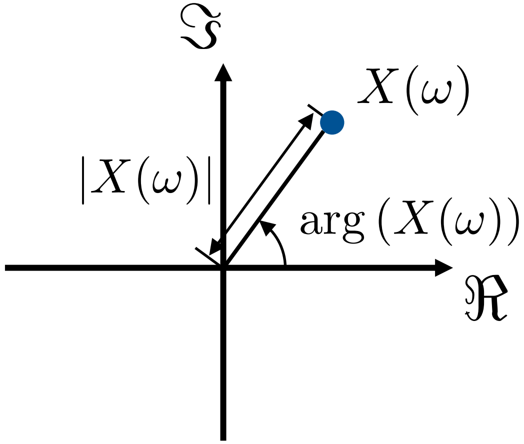
$$x(t) \xleftrightarrow{\text{FT}} X(\omega)$$

Fourier transform

- ω : (Angular) frequency
- When $X(\omega)$ is regarded as a complex function of variable ω
  - $X(\omega)$: (Angular) frequency spectrum
  - $|X(\omega)|$: Magnitude spactrum
  - $|X(\omega)|^2$: Power spectrum
  - $\arg(X(\omega))$: Phase spectrum



- When $X(\omega)$ is regarded as a complex scalar value at ω
  - $|X(\omega)|$: Magnitude
  - $|X(\omega)|^2$: Power
  - $\arg(X(\omega))$: Phase



Discrete Fourier transform

➢ Definition

- Discrete Fourier transform

$$X[k] = \sum_{n=0}^{N-1} x[n] \exp\left(-\mathrm{j}\frac{2\pi kn}{N}\right)$$

$$k \in \{0, 1, \ldots, N-1\}$$

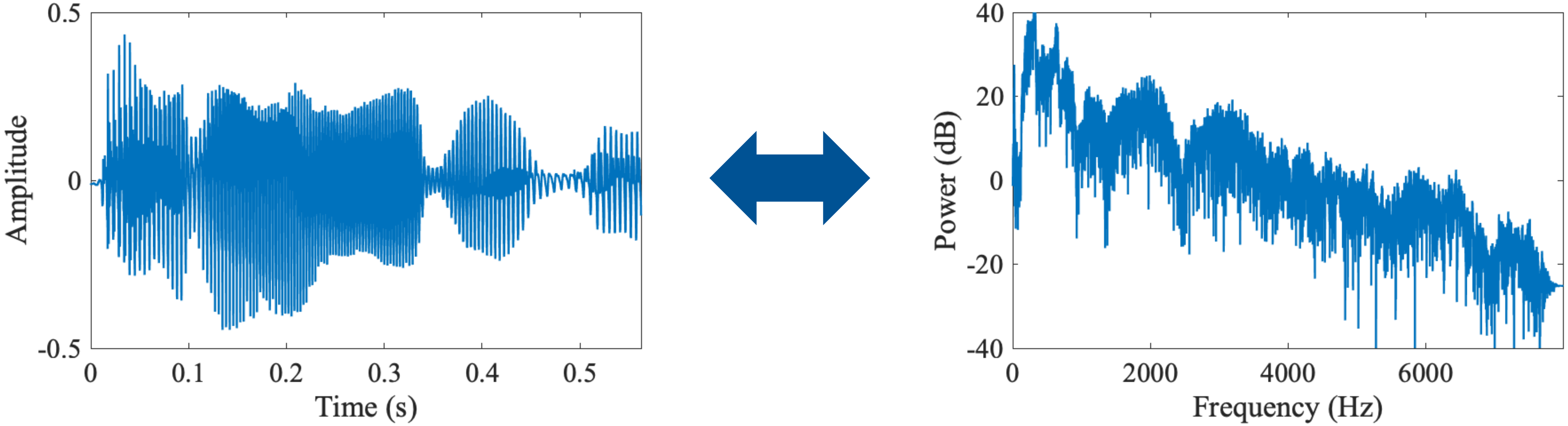
- Inverse discrete Fourier transform

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] \exp\left(\mathrm{j}\frac{2\pi kn}{N}\right)$$

$$n \in \{0, 1, \ldots, N-1\}$$

Fourier transform for frequency analysis

- From engineering perspective, Fourier transform is frequency analysis of temporal signal by decomposing it by amplitude and phase of sinewaves
- Inverse Fourier transform is waveform synthesis by generating temporal signal from amplitude and phase of sinewaves



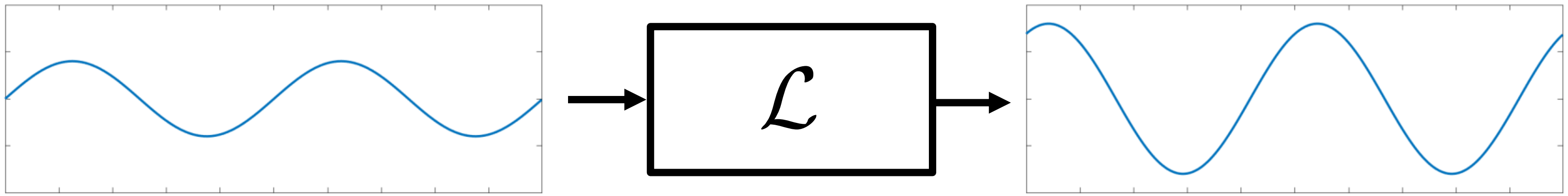
Temporal signal

Spectrum

Frequency response of LTI system

LTI system characteristics are fully described by frequency response

- By decomposing LTI system into sinewaves, input-output relationship can be represented by change of amplitude and phase at each frequency.
- Amplitude change is called gain (or amplitude response), and phase change is called phase shift (or phase response)
- Gain and phase shift for each frequency is called frequency response



Transfer function

- Input-output relationship represented by function of frequency is called transfer function
- Transfer function $H(\omega)$ at angular frequency ω is written by gain $G(\omega)$ and phase shift $\exp(\mathrm{j}\theta(\omega))$ as

$$H(\omega) = G(\omega)\exp(\mathrm{j}\theta(\omega))$$

- Output of the system when input is complex sinewave $\exp(\mathrm{j}\omega t)$ at angular frequency ω

$$\mathcal{L}[\exp(\mathrm{j}\omega t)] = H(\omega)\exp(\mathrm{j}\omega t)$$

- Input signal and output signal are related by their spectrum $X(\omega),\ Y(\omega)$

$$Y(\omega) = H(\omega)X(\omega)$$

Transfer function

- Representing input-output relationship of LTI system by using Fourier transform,

$$
\begin{aligned}
y(t) &= \mathcal{L}[x(t)] \\
&= \mathcal{L}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)\exp(\mathrm{j}\omega t)\mathrm{d}\omega\right] \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)\mathcal{L}\left[\exp(\mathrm{j}\omega t)\right]\mathrm{d}\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)H(\omega)\exp(\mathrm{j}\omega t)\mathrm{d}\omega
\end{aligned}
$$

$$\mathcal{L}\left[\exp(\mathrm{j}\omega t)\right] = H(\omega)\exp(\mathrm{j}\omega t)$$

Transfer function

- Output signal is

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} Y(\omega) \exp\left(\mathrm{j}\omega t\right) \mathrm{d}\omega$$

- By comparing with

$$y(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega) H(\omega) \exp\left(\mathrm{j}\omega t\right) \mathrm{d}\omega$$

we can obtain

$$Y(\omega) = H(\omega) X(\omega)$$

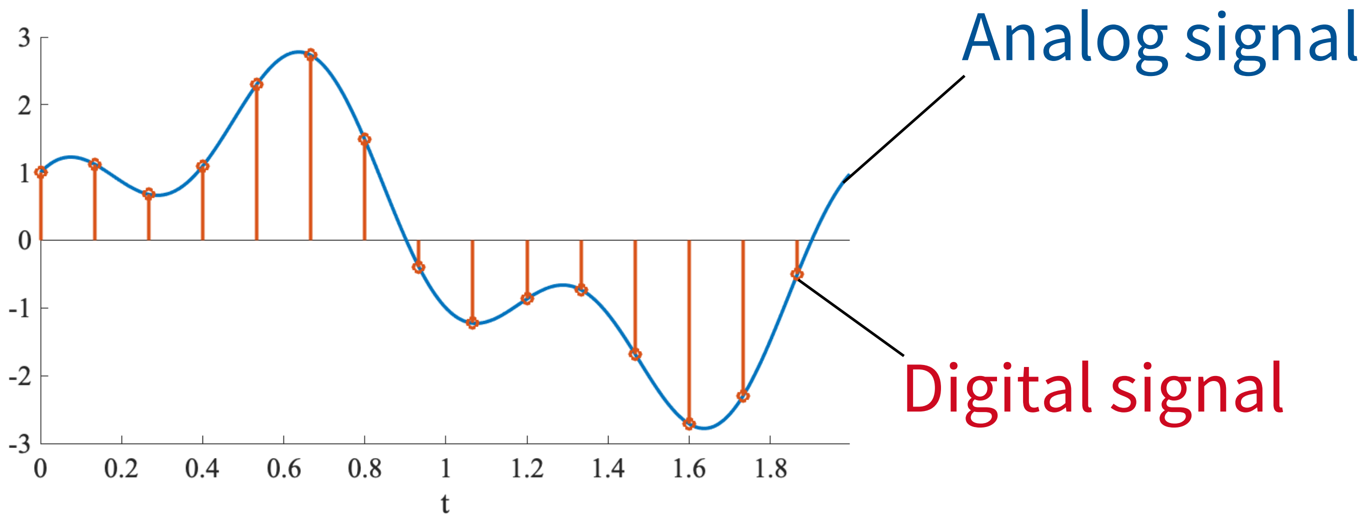
Output signal of LTI system is multiplication of
input signal and transfer function in the frequency domain

SAMPLING THEOREM

Sampling

- By discretizing continuous-time signal in the temporal axis, which is called sampling, discrete-time signal is obtained
- Time interval of sampling T is called sampling period, and its inverse $1/T$ is called sampling frequency
- Discrete-time signal $x[n]$ is written as

$$x[n] = x(nT) = \sum_{n=-\infty}^{\infty} x(t)\delta(t - nT)$$



Sampling theorem

➢ Sampling theorem

– When the upper limit of the frequency band of Fourier transform $X(\omega)$ of continuous-time signal $x(t)$ is $\omega_0 = 2\pi f_0$, continuous-time signal $x(t)$ is perfectly reconstructed from discrete-time signal $x[n]$ of sampling frequency $2f_0$ or above

– Corresponding to the condition that aliasing does not occur, sampling frequency must satisfy f_s

$$f_0 \leq \frac{f_\mathrm{s}}{2}$$

– Half of the sampling frequency is called Nyquist frequency

Sampling theorem

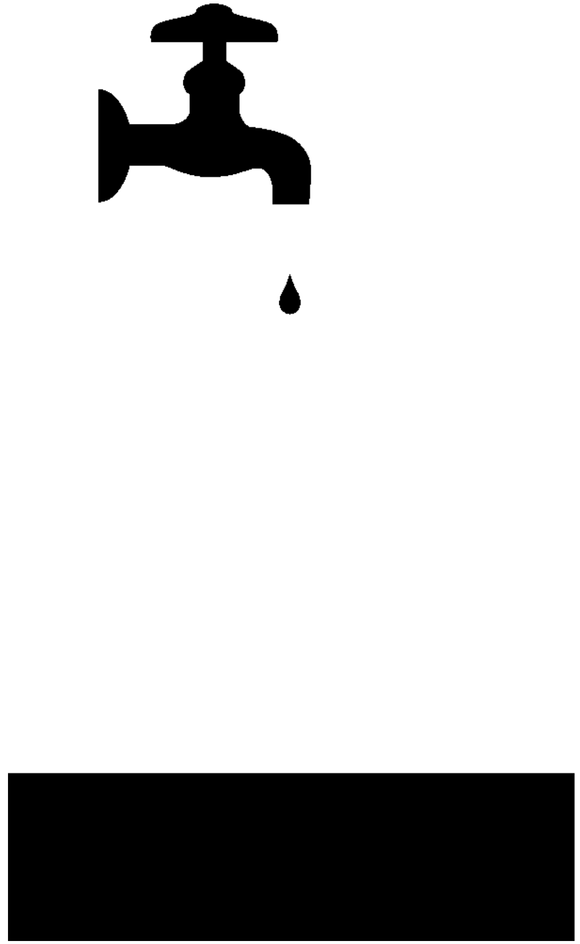
- Irradiating flash lamp to periodic waterdrop from faucet

$$T = 0.5\ \mathrm{s}$$

- Irradiating flash lamp to periodic waterdrop from faucet

$$T = 1.0\ \mathrm{s}$$

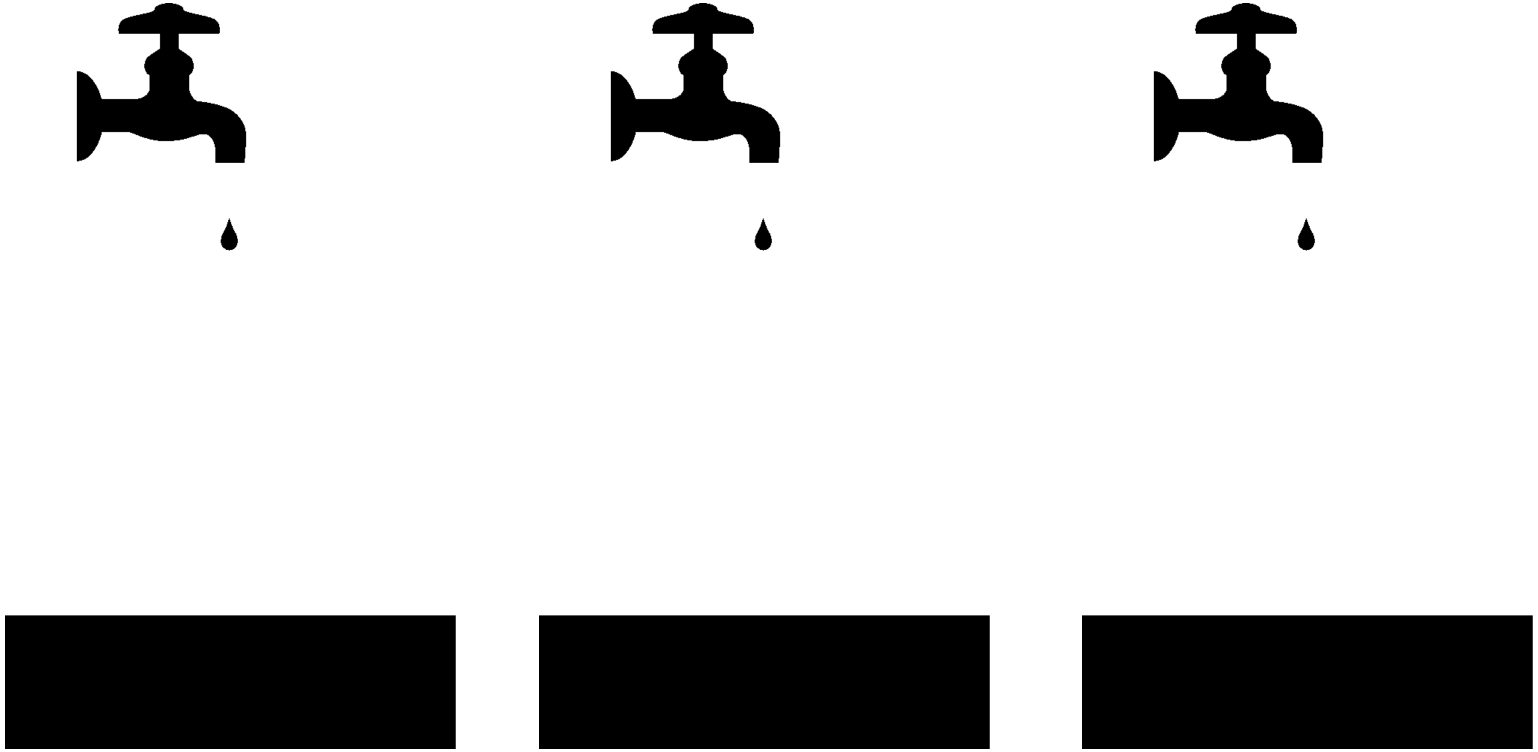
- Irradiating flash lamp to periodic waterdrop from faucet



$$T = 1.5 \text{ s}$$

Sampling theorem

- Irradiating flash lamp to periodic waterdrop from faucet



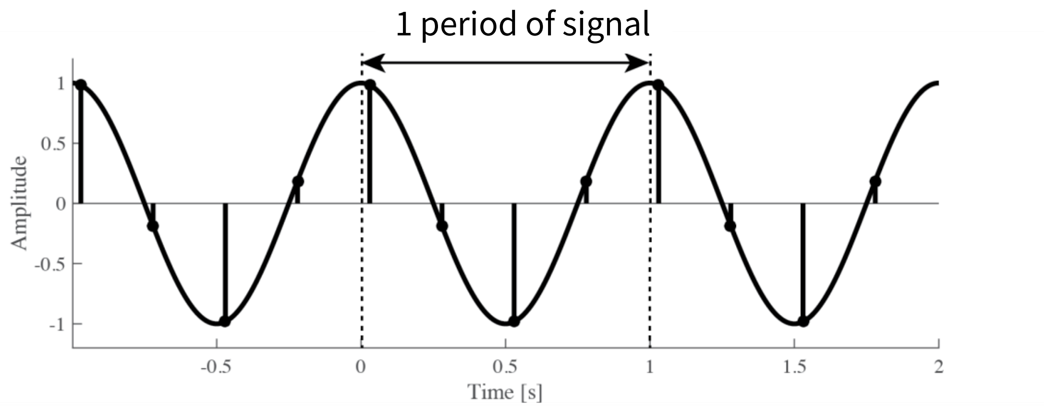
$T = 0.5$ s $\quad$ $T = 1.0$ s $\quad$ $T = 1.5$ s

日本音響学会編, ”音響学入門ペディア,” コロナ社, 2017.

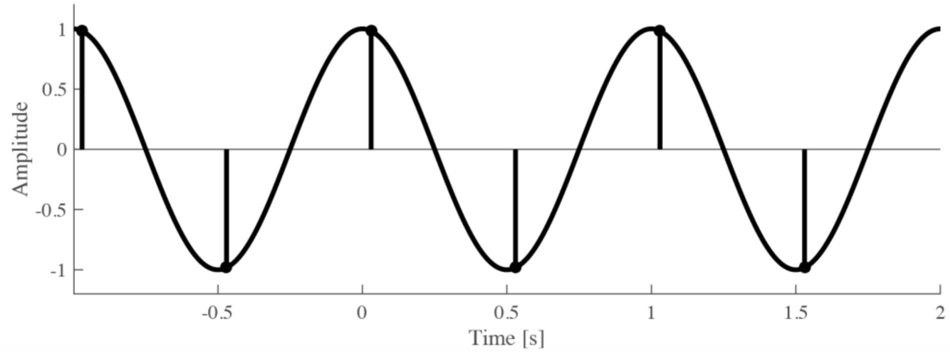
Direction of waterdrop is indistinctive
when interval of irradiation is large

- Suppose sinewave of 1 sec of period (1 Hz of frequency)

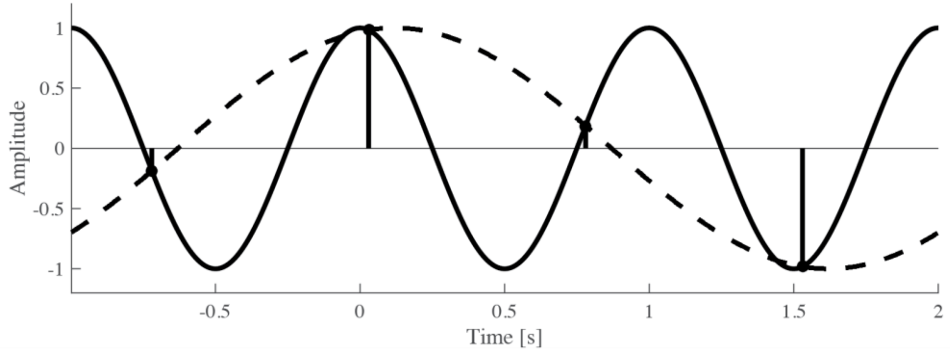
$$f_s = 4\ \mathrm{Hz}$$



$$f_s = 2\ \mathrm{Hz}$$



$$f_s = 4/3\ \mathrm{Hz}$$



Relationship between continuous and discrete signals

➢ Relationship between continuous-time and discrete-time signals in the frequency domain

$$\begin{aligned}
X_{\mathrm{D}}(\omega T) &= \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} x(t)\delta(t-nT)\exp(-\mathrm{j}\omega t)\mathrm{d}t \\
&= \int_{-\infty}^{\infty} x(t) \sum_{n=-\infty}^{\infty} \delta(t-nT)\exp(-\mathrm{j}\omega t)\mathrm{d}t \\
&= \frac{1}{2\pi} X_{\mathrm{A}}(\omega) * \mathcal{F}\left[\sum_{n=-\infty}^{\infty} \delta(t-nT)\right] \\
&= \frac{1}{2\pi} X_{\mathrm{A}}(\omega) * \frac{2\pi}{T} \sum_{n=-\infty}^{\infty} \delta\left(\omega - \frac{2\pi}{T} n\right) \\
&= \frac{1}{T} \int_{-\infty}^{\infty} X_{\mathrm{A}}(\xi) \sum_{n=-\infty}^{\infty} \delta\left(\omega - \frac{2\pi}{T} n - \xi\right) \mathrm{d}\xi \\
&= \frac{1}{T} \sum_{n=-\infty}^{\infty} X_{\mathrm{A}}\left(\omega - \frac{2\pi}{T} n\right)
\end{aligned}$$

$$\boxed{x(t) = x(nT)}$$

Convolution and multiplication

Fourier transform of delta sequence

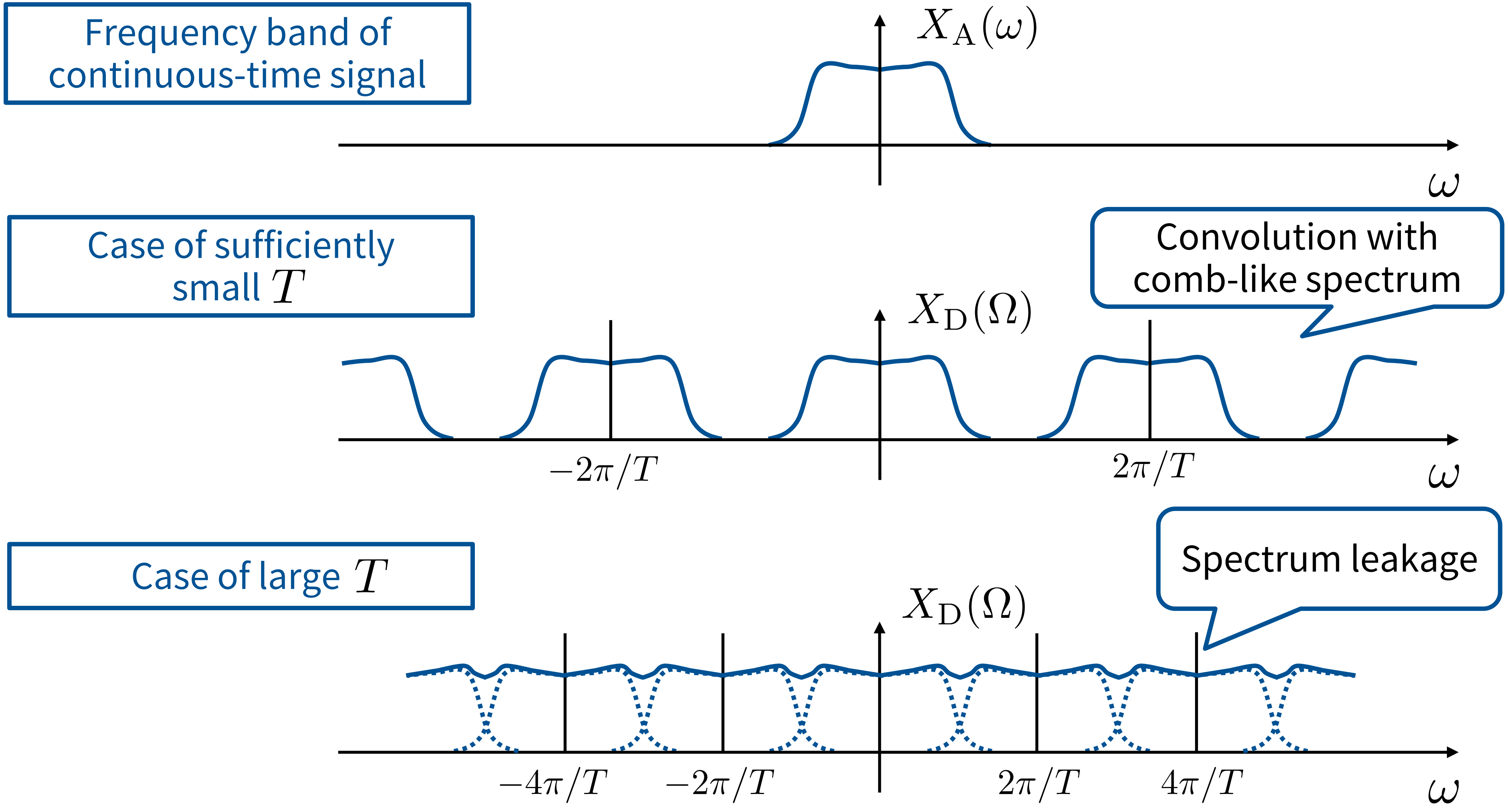
Definition of convolution

Definition of delta function

$$\Rightarrow \quad X_{\mathrm{D}}(\Omega) = \frac{1}{T} \sum_{n=-\infty}^{\infty} X_{\mathrm{A}}\left(\frac{\Omega}{T} - \frac{2\pi}{T} n\right)$$

Relationship between continuous and discrete signals

➢ $X_{\mathrm{D}}(\Omega)$ is shifted sum of $X_{\mathrm{A}}(\omega)$ at intervals of $2\pi/T$



Sampling theorem, again

➢ Sampling theorem

- When the upper limit of the frequency band of Fourier transform $X(\omega)$ of continuous-time signal $x(t)$ is $\omega_0 = 2\pi f_0$, continuous-time signal $x(t)$ is perfectly reconstructed from discrete-time signal $x[n]$ of sampling frequency $2f_0$ or above

- Corresponding to the condition that aliasing does not occur, sampling frequency must satisfy f_s

$$f_0 \leq \frac{f_\mathrm{s}}{2}$$

- Half of the sampling frequency is called Nyquist frequency

Sampling theorem

- Relationship between continuous-time signal $x(t)$ with band limitation $(-\pi/T < \omega < \pi/T)$ and discrete-time signal $x[n]$ in the time domain

$$
\begin{aligned}
x(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} X_{\mathrm{A}}(\omega)\exp(\mathrm{j}\omega t)\mathrm{d}\omega \\
&= \frac{1}{2\pi}\int_{-\pi/T}^{\pi/T} X_{\mathrm{A}}(\omega)\exp(\mathrm{j}\omega t)\mathrm{d}\omega \\
&= \frac{1}{2\pi T}\int_{-\pi}^{\pi} X_{\mathrm{A}}\left(\frac{\Omega}{T}\right)\exp\left(\mathrm{j}\frac{\Omega}{T}t\right)\mathrm{d}\Omega \\
&= \frac{1}{2\pi T}\int_{-\pi}^{\pi} T X_{\mathrm{D}}\left(\Omega\right)\exp\left(\mathrm{j}\frac{\Omega}{T}t\right)\mathrm{d}\Omega \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left[\sum_{n=-\infty}^{\infty} x[n]\exp(-\mathrm{j}\Omega n)\right]\exp\left(\mathrm{j}\frac{\Omega}{T}t\right)\mathrm{d}\Omega \\
&= \sum_{n=-\infty}^{\infty} x[n]\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi}\exp\left[\mathrm{j}\Omega\left(\frac{t}{T}-n\right)\right]\mathrm{d}\Omega\right\} \\
&= \sum_{n=-\infty}^{\infty} x[n]\frac{\sin\pi\left(\frac{t}{T}-n\right)}{\pi\left(\frac{t}{T}-n\right)}
\end{aligned}
$$

Band limitation

Change of variable $\Omega = \omega T$

Because of band limitation

$$X_{\mathrm{D}}(\Omega) = \frac{1}{T}X_{\mathrm{A}}\left(\frac{\Omega}{T}\right)$$

Sampling theorem

- Band-limited continuous-time signal $x(t)$ is perfectly reconstructed by convolution of discrete-time signal $x[n]$ and sinc function

$$x(t) = \sum_{n=-\infty}^{\infty} x[n] \operatorname{sinc}\left[\pi\left(\frac{t}{T} - n\right)\right]$$

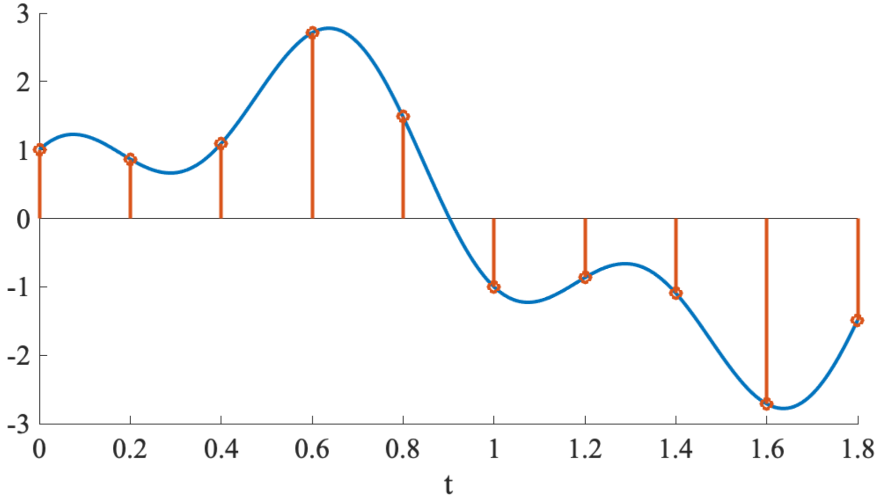
$$= x[n] * \operatorname{sinc}\left(\frac{\pi t}{T}\right)$$

➡ Sinc interpolation

  – For perfect reconstruction, discrete-time signal $x[n]$ must be defined in $n \in \mathbb{Z}$
  – Difficult in practice, but well approximated by truncation because of rapid attenuation of sinc function

Reconstruction from discrete-time signal

- Sampling of continuous-time signal



- Reconstruction of continuous-time signal by sinc interpolation