I645 Human Perceptual Systems and its Models

3. Sound representations and signal processing

2024/6/21

Masashi UNOKI

URL: http://www.jaist.ac.jp/~unoki/

Email: unoki@jaist.ac.jp

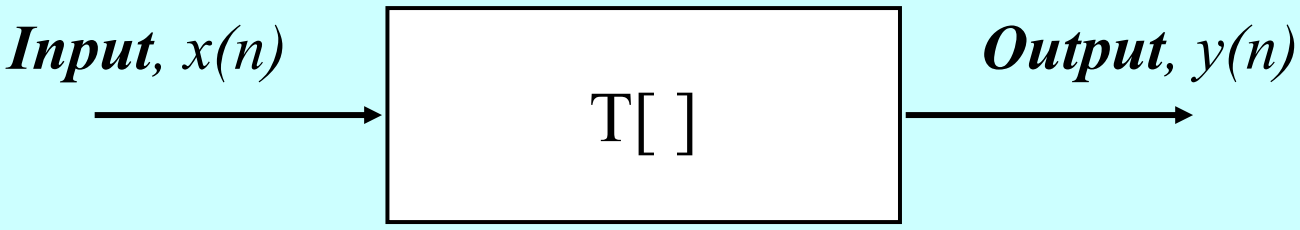
JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY
HOKURIKU 1990

0. Motivation

- ***Question 1:*** ***What is the difference between linear and nonlinear processing?***

System (processing)

$$y(n) = T\left[x(n)\right]$$



Question 2: ***Which is linear or nonlinear processing?***

Fourier transform, Laplace transform, wavelet transform, z transform, Audio systems, Human auditory system

1. Introduction

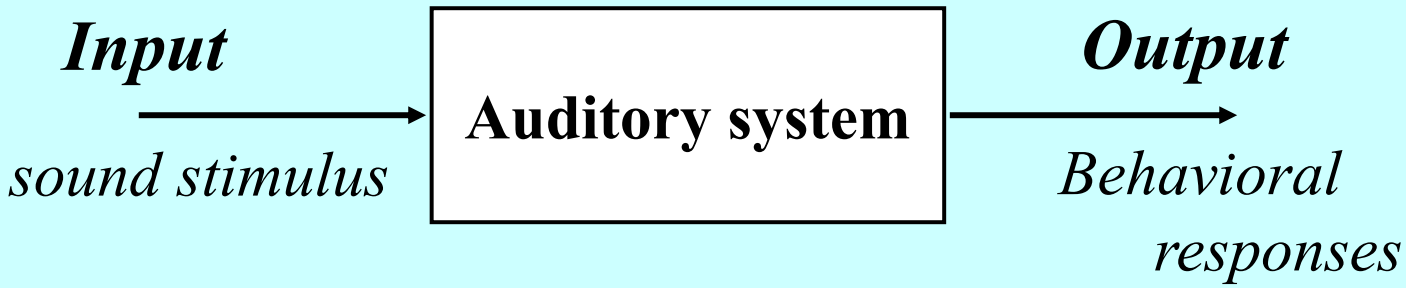
- **"Hearing"**
  - a ***crucial sense*** for humans
  - ***central*** to the interaction of human begins with other human beings
  - ***importance*** to our interactions with our environments
- ***Question: Is hearing a linear system?***

  Answer: No!
- ***How is hearing investigated?***
  - **Auditory physiology**
  - **Auditory psychophysics (Psychoacoustics)**

- **Auditory physiology**
  - Internal workings of the auditory system: how sound is processed by the cells and structures in the ear and brain
  - *Direct/indirect* measurements of the biological systems
    - by ***Surgery***
    - ***Neurophysiological*** techniques (electrical activity)
    - Modern "***brain imaging***" techniques (fMRI)

Psychoacoustics

- *Behavioral* study of hearing – behavioral in that the participant is required to make a response to the sounds that are presented.



- **Importance: to combine *physiological* experiments with experiments that employ *behavioral* techniques**

2. Sound

What is "sound"

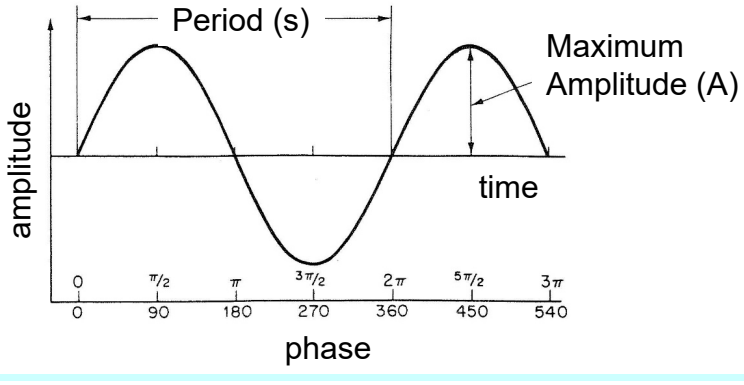
- Sound ***waveforms*** (physical meaning)
- ***Sensations*** evoked by sound (psychological meaning)

Characteristics of sound

- Variation of the air pressure $x(t) = A\sin(2\pi ft + \theta)$
- The simplest types of sound: **sine wave**

Sinusoidal wave

- A: amplitude
- f: frequency
- θ: phase



Various sounds around our life








Speech sound

- *Are these speech or non-speech?*

- **Speech sound/voice**

evoked sound by (human) glottis

3. Sound representations

Fourier analysis and spectrum representations

- Discrete Fourier series

Fourier coefficients

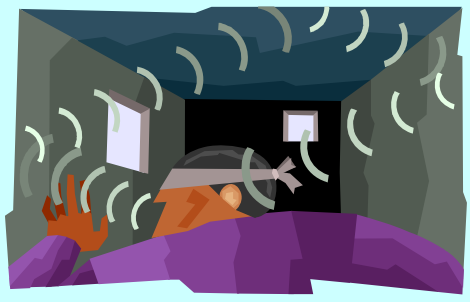
$$\begin{cases} c_k = \int_{-T/2}^{T/2} \tilde{x}_a(t) e^{-jk\omega_0 t} dt \\ \tilde{x}_a(t) = \frac{1}{T} \sum_{k=-\infty}^{\infty} c_k e^{jk\omega_0 t} \end{cases}$$

- Discrete Fourier transform

Fourier component

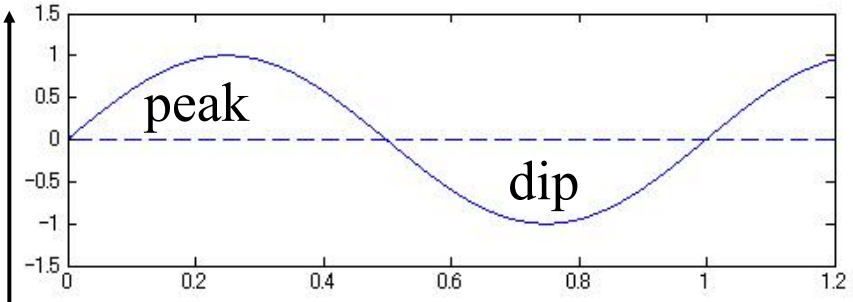
$$\begin{cases} \tilde{X}(k) = \sum_{n=0}^{N-1} \tilde{x}(n) W_N^{kn} \\ \tilde{x}(n) = \frac{1}{N} \sum_{k=0}^{N-1} \tilde{X}(k) W_N^{-kn} \end{cases}$$

Vibrations — Frequency

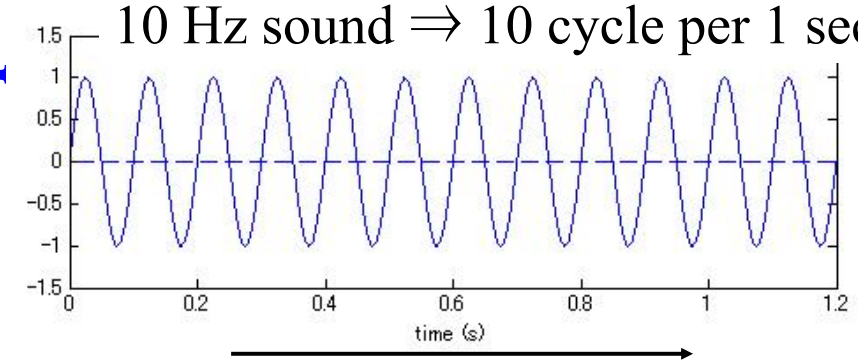


"waveform"

1 Hz sound ⇒ 1 cycle per 1 sec.



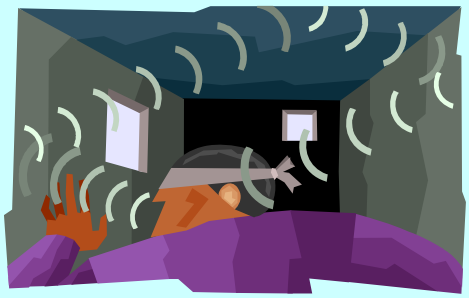
10 Hz sound ⇒ 10 cycle per 1 sec.

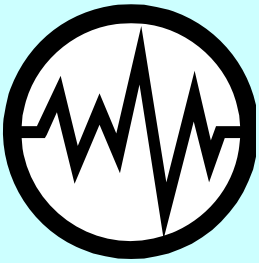


Amplitude

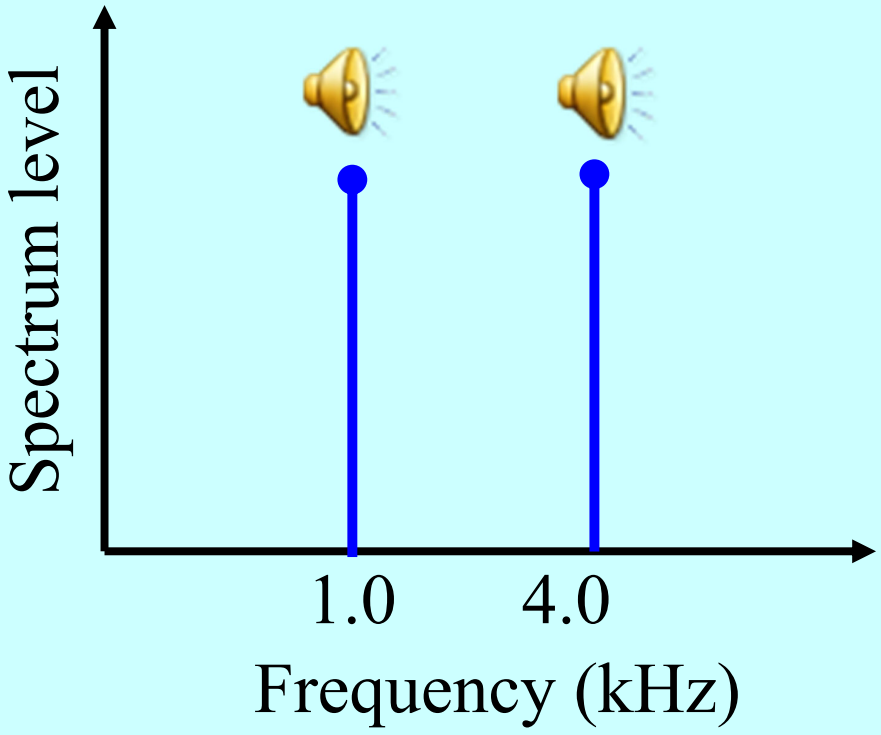
Time (s)

Vibrations - Spectrum





"waveform"



Fourier Analysis/Synthesis

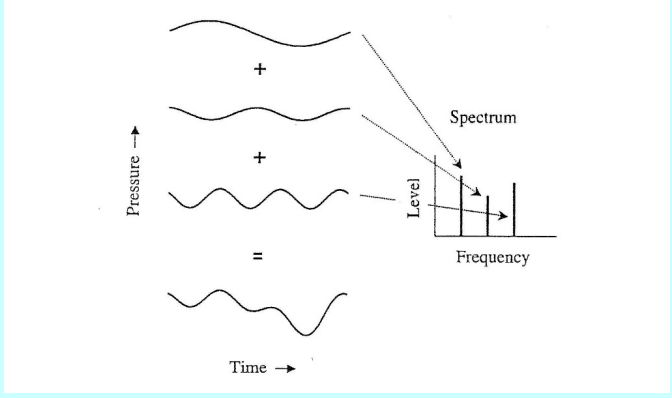
- **Concept of decomposition**

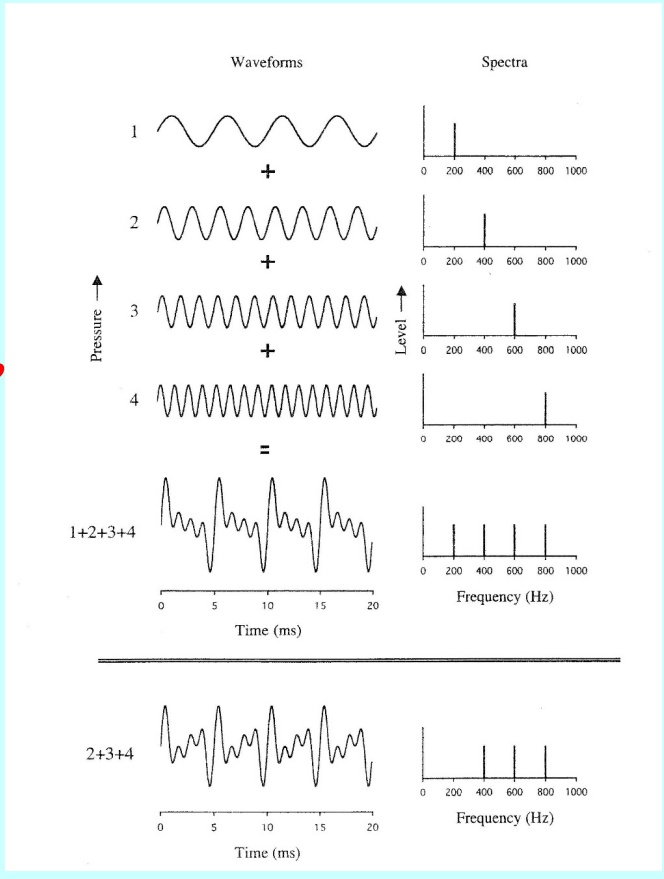
Periodic sound =

DC

+ summation of "sine"

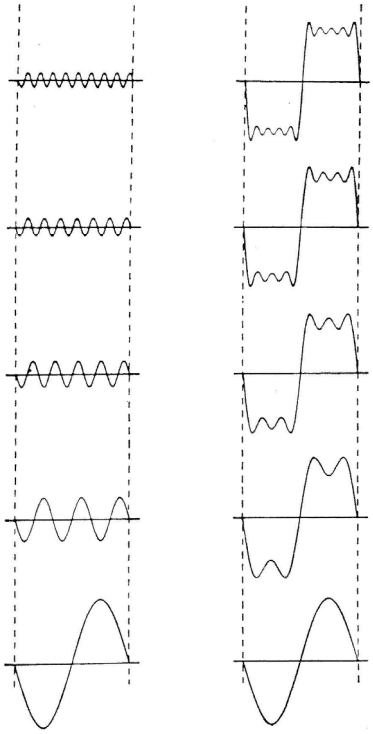
+ summation of "cosine"

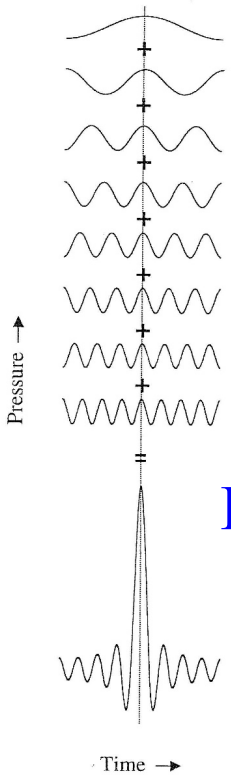




Decomposition of Rect. and Impulsive signals

Rectangular signal

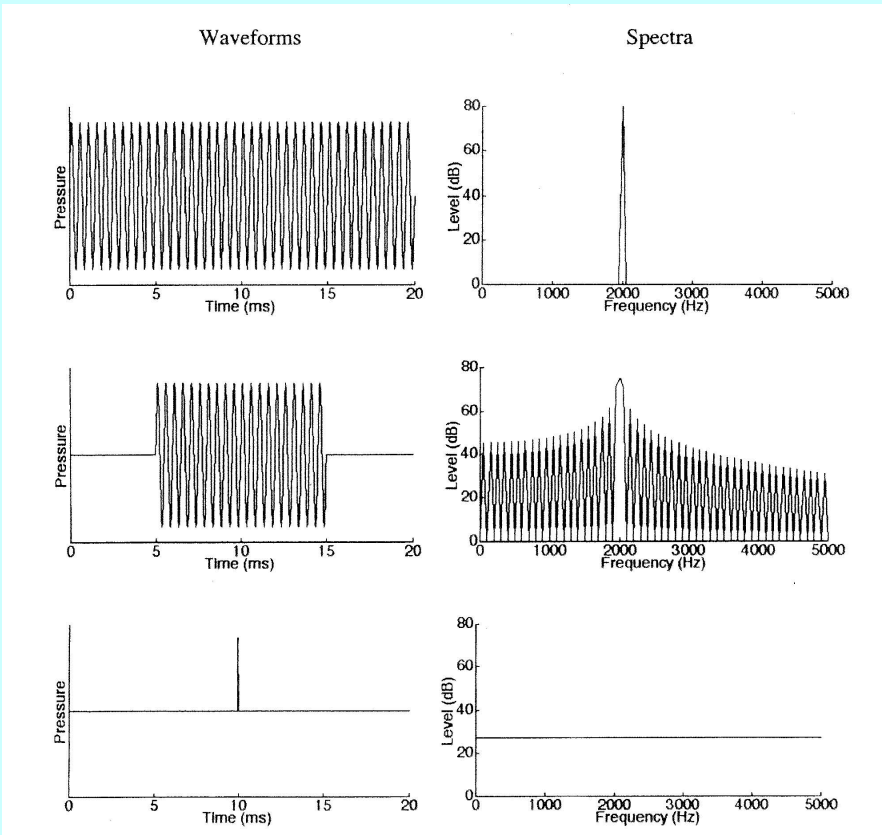




Impulsive signal

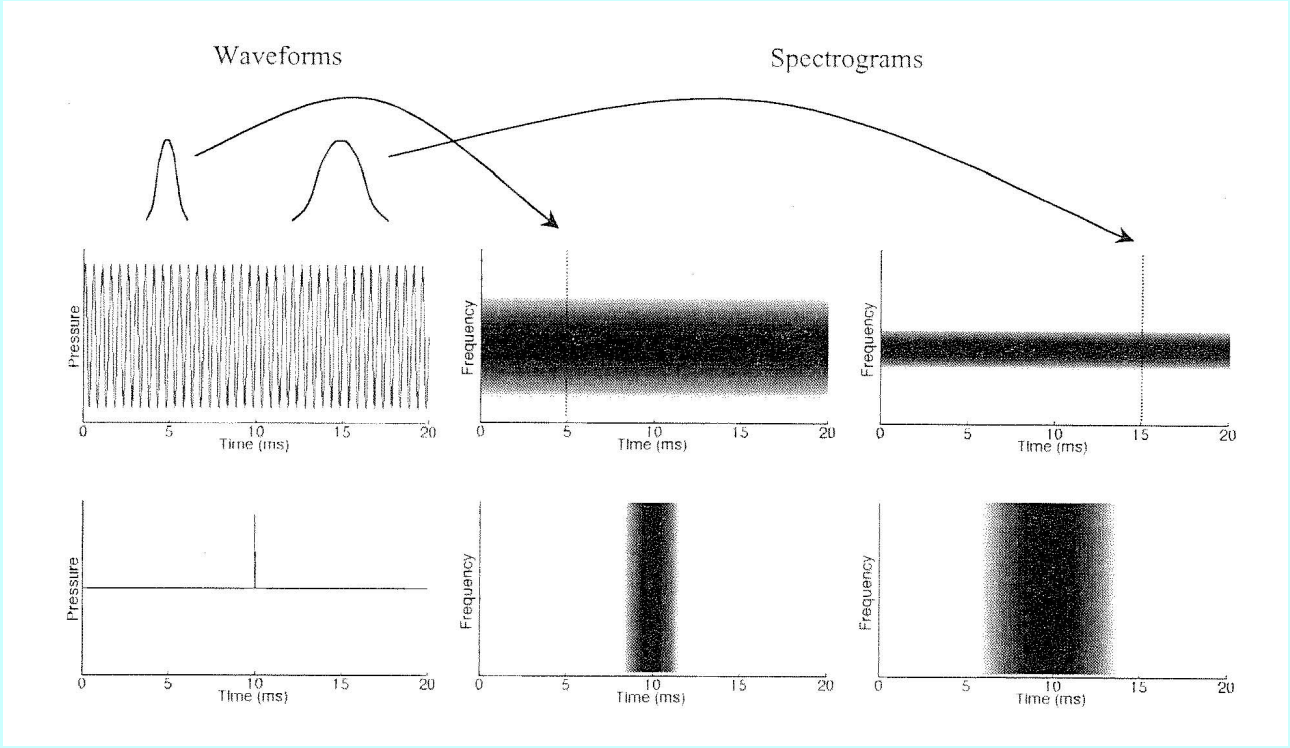
Relation between signal and spectrum ①

- ***When signal duration could be shorter,***
  - **Sine wave**
  - **Tone-burst**
  - **Impulse**



Relation between signal and spectrum ②

Windowing dependence



Relation between signal and spectrum ③

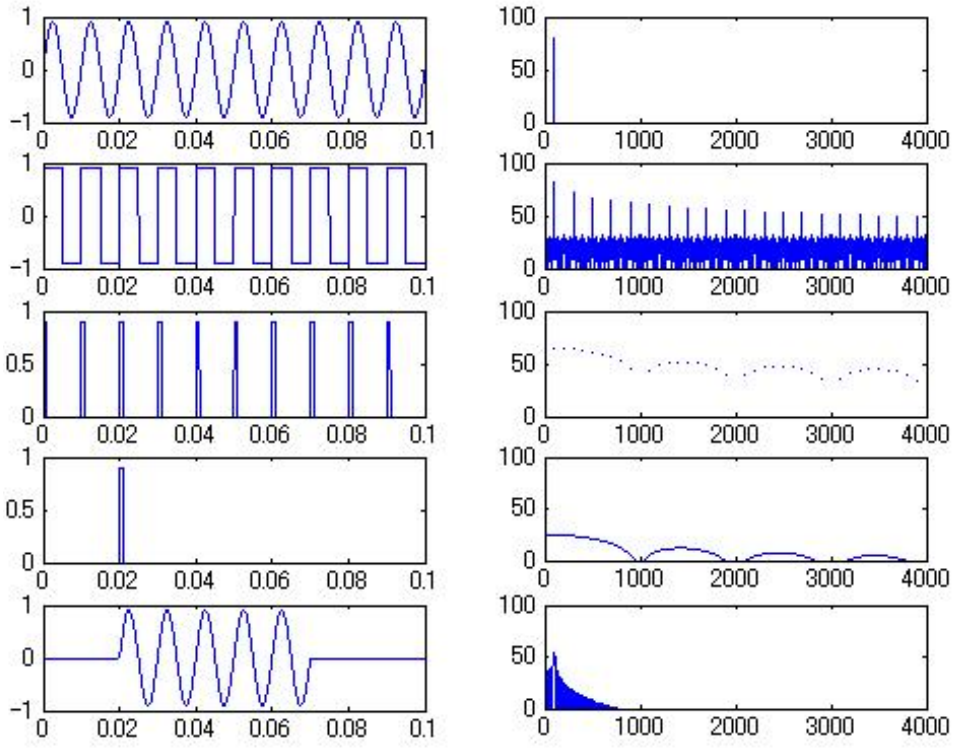
Sinusoidal wave

Rectangular tone

Pulse train

Pulse

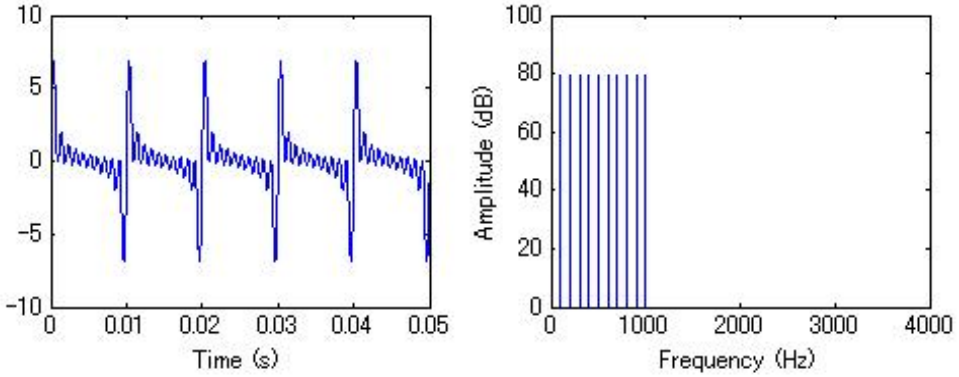
Tone-burst



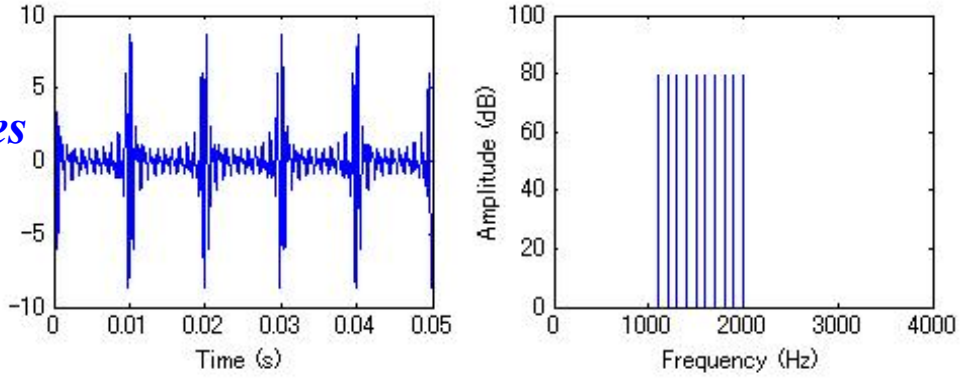
Relation between signal and spectrum ④

Harmonic sound

Lower frequencies



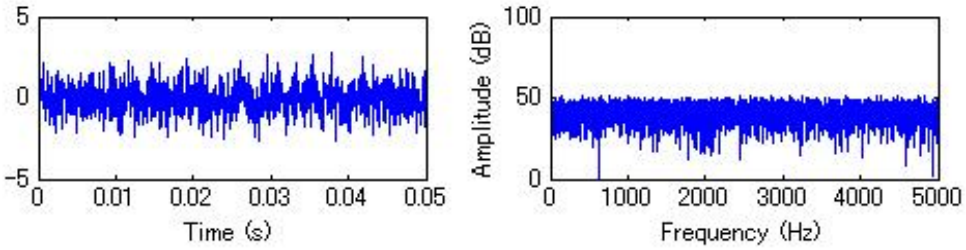
Higher frequencies



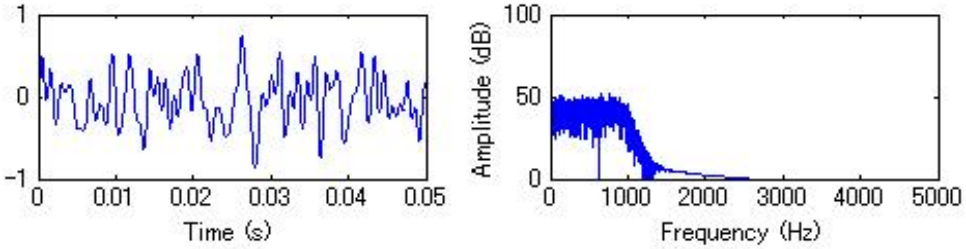
Relation between signal and spectrum ⑤

White noise

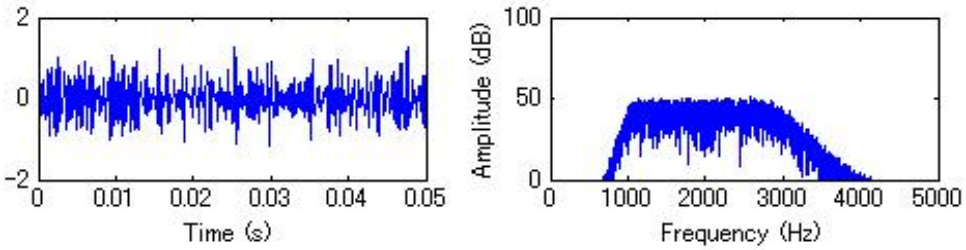
whole-band



Low-passed



Band-limited



4. Characteristics of sound

- ***Power***
  - The energy transmitted per second
- ***Sound intensity***
  - The sound power transmitted through a given unit area in a sound field (e.g., a square meter of air) every second
  - Unit: watt per square meter [w/m^2]
  - 10^{-12} W/m^2 $=2\times10^{-5}$ $N/m^2=20\mu Pa$
- ***Proportional relation***

$$I=kP^2$$

I: the intensity, P: the rms pressure

- ***Sound level in* dB**

$$\text{decibel scale} = 10\log_{10}(I_o/I_i) \quad \text{(dB)}$$

I_i: Intensity of the input, I_o: Intensity of the output

e.g., 60 dB SPL

→ 60 dB higher in the reference level of 0 dB

Intensity of 10^{-6} W/m^2

the ratio of intensities by 10 (or 1/10)

→ 10 dB ***increases*** (or ***decreases***)

the ratio of intensities by 2 (or ½)

→ 3 dB ***increases*** (or *decreases*)

Sound level in dB

decibel scale $= 20\log_{10}(P_o/P_i)$

P_i: rms pressure of the input

P_o : rms pressure of the output

- RMS value → root-mean-square value of sound pressure

SPL (Sound Pressure Level) in dB

SL (Sensational Level) in dB

e.g. SL of 60 dB will be 60 dB above the absolute threshold

- Energy density Energy per 1 Hz band

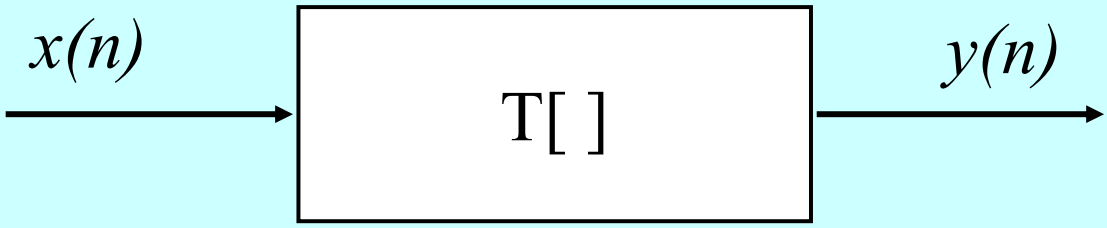
Relationship between dB, intensity, and Pressure ratios

Level SPL in dB	**Intensity** I/I_0	**Pressure** P/P_0	**Typical case**
140	10^{14}	10^7	***Gunshot***
120	10^{12}	10^6	***Loud rock group***
100	10^{10}	10^5	***Shouting***
80	10^8	10^4	***Heavy Traffic***
70	10^7	3.16×10^3	***Normal conversation***
50	10^5	316	***Quiet conversation***
30	10^3	31.6	***Library***
20	10^2	10	***Quiet Forest***
0	1	1	***Absolute Threshold***

5. Linear/nonlinear systems

Linearity

$$y(n) = T[x(n)]$$



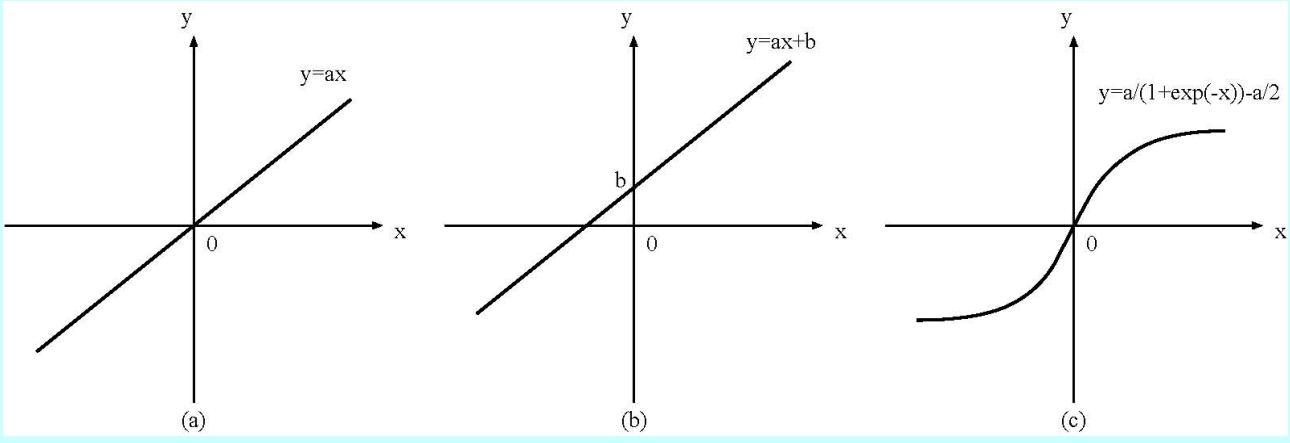
$$T[ax_1(n)] + T[bx_2(n)] = aT[x_1(n)] + bT[x_2(n)]$$
$$= ay_1(n) + by_2(n)$$

Requirement for linearity

- *Additivity* property
- *Homogeneity* property
- *Superposition* property

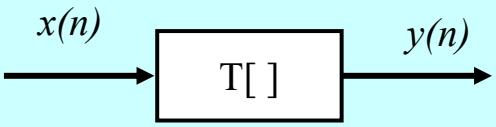
→ otherwise, ***nonlinearity***

Question: *Which is a linear system?*

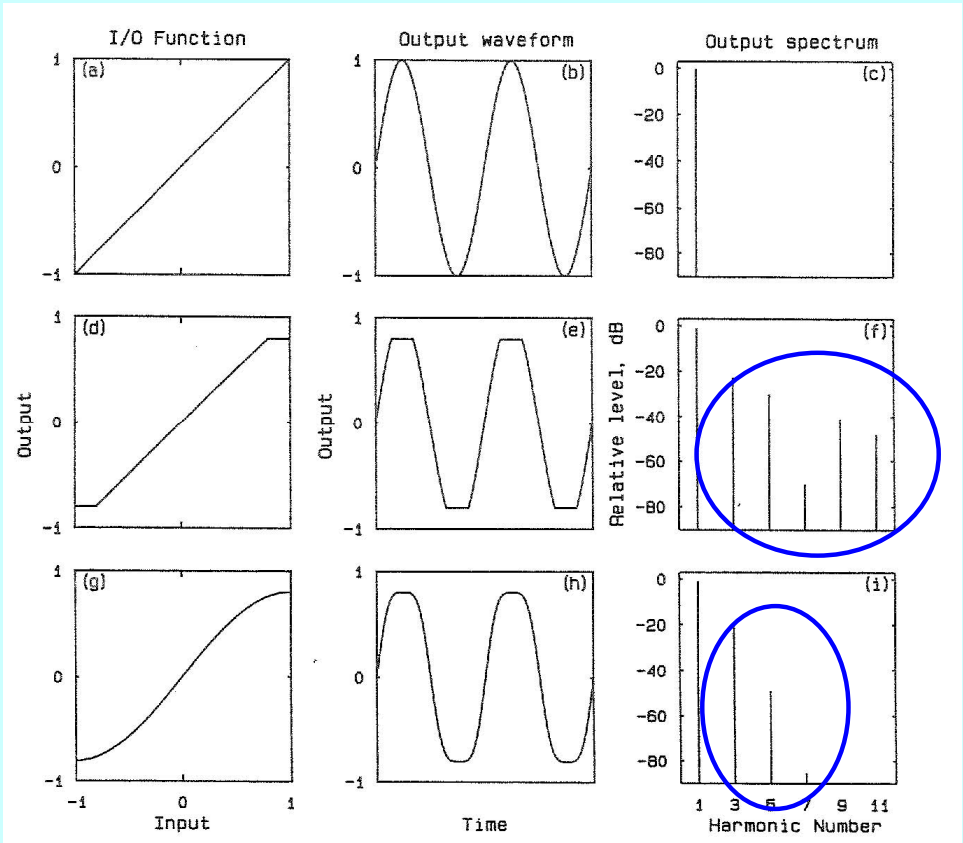


Example

$$y(n) = T\left[x(n)\right]$$

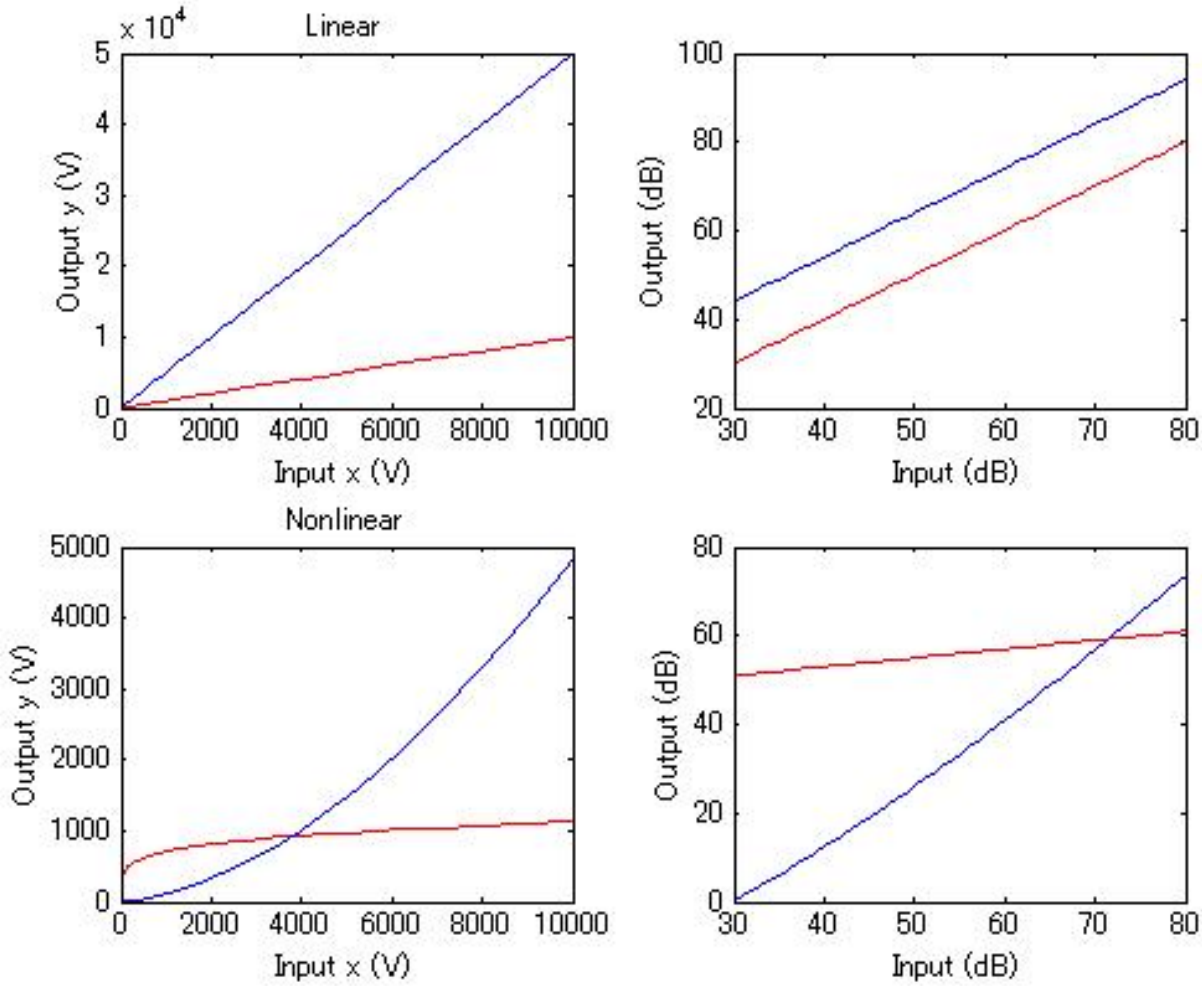


$A\sin\omega t$



Question:

Why can harmonisity be observed?



6. Filter

Types of the filter

- LPF

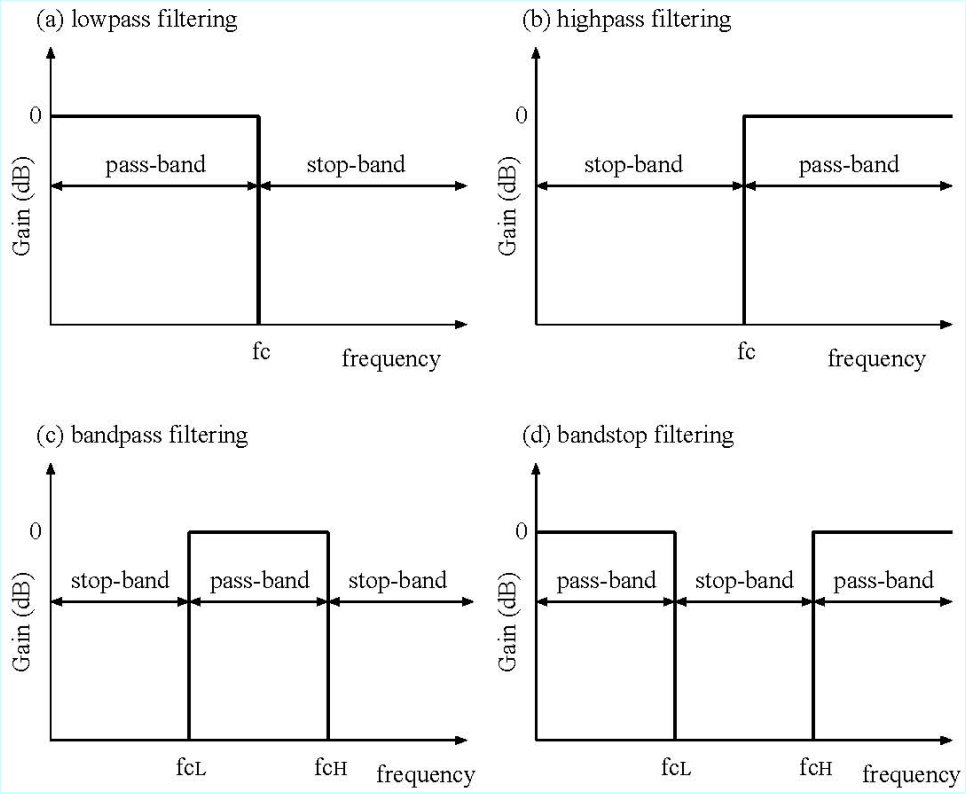
  Low-pass filter

- BPF

  Band-pass filter

- HPF

  High-pass filter



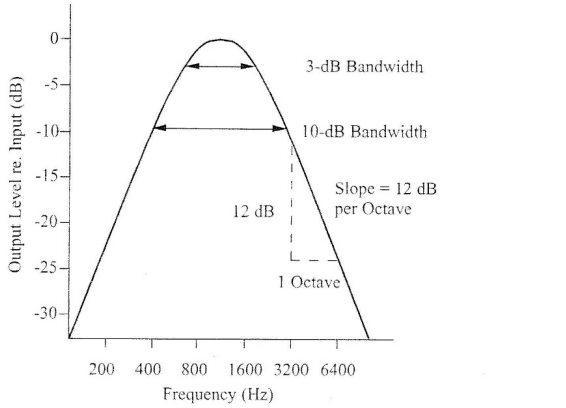
Filter characteristics

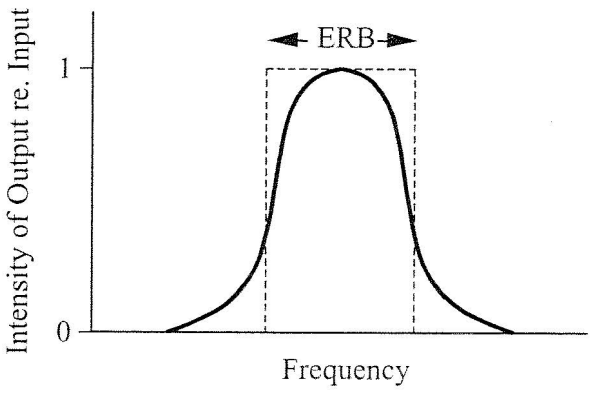
Bandwidth

- 3-dB BW (or 10-dB BW)
- ERB (Equivalent Rectangular Bandwidth)

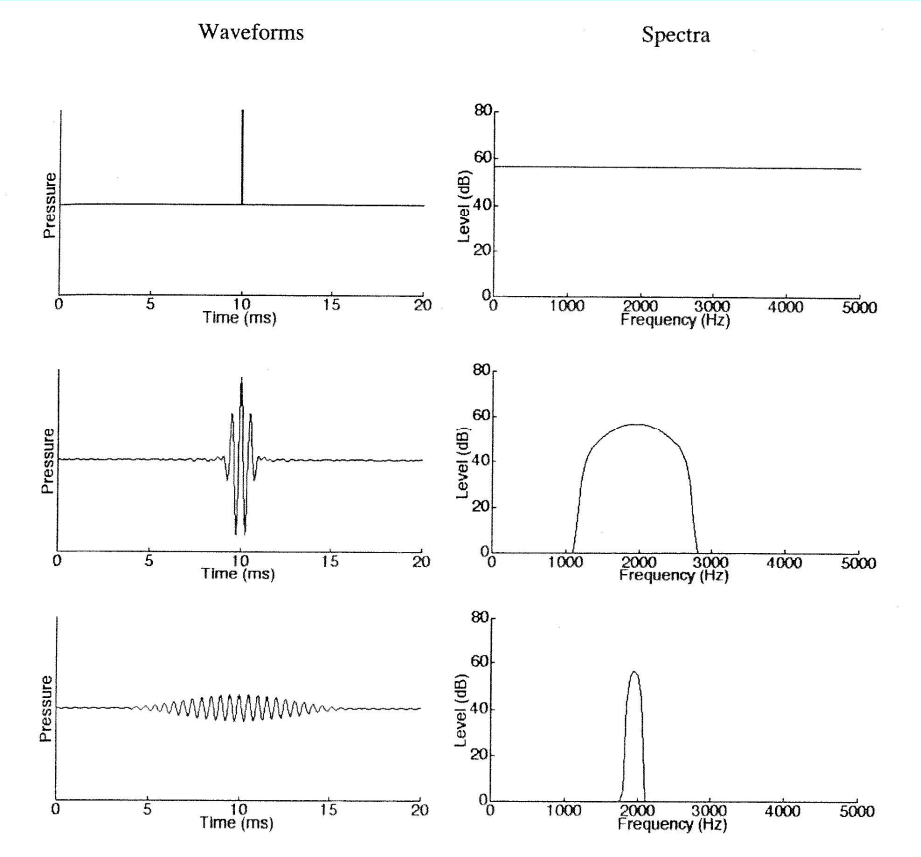
Filter-Q (tuning)

- Q=center frequency/bandwidth





Filter bandwidths and impulse responses



7. Waveform manipulation

- Effect of compression
- Effect of half-wave rectification