Basics of Sound

- **Mathematics of the pure tone**
  - $x(t) = A\sin(2\pi t / T + \phi)$
  - or $x(t) = A\sin(2\pi f t + \phi)$
    - A: amplitude
    - ϕ: phase
    - T: period
    - f: frequency
- **Phase Lead - Phase Lag**

Power, Intensity, and Decibels

- **Treat signal *x(t)* as voltage**

  By Ohm's law, the current *i(t)* = *x(t)/R*

  Then the instantaneous power is

  $$P(t) = x(t)i(t) = x^2(t) / R$$

  Energy is the integrated power over a certain time period (e.g. kilowatt vs. kilowatthour)

- **Treat signal *x(t)* as sound pressure**

  The instantaneous intensity is

  $I(t) = x^2(t)/(\rho c)$, measured in watts/m^2 ($x(t)$: pressure, Newtons/m^2 or pascals)

  ρ: the density of the medium

  c: speed of sound

Power, Intensity, and Decibels (cont.)

- **Sound Levels**

  Ratio of one sound to another (baseline), expressed as decibels (dB)

  $$L_2 - L_1 \,(\text{decibels}) = 10 \log_{10} (I_2 / I_1)$$

  - Note the use of common logarithm
  - Double intensity leads to 3 dB, and double amplitude leads to 6 dB

  - SNR: *signal-to-noise ratio*

  - Conversational speech is about 65 dB. Above 100 dB is damaging to the ear

Spectrum

- **Fourier Series**
  - For any periodic function of time, *x(t)*, with period *T*, i.e.

    $x(t+mT) = x(t)$, for all integer m

    x(t) can be represented as a Fourier series like this

    $$x(t) = A_0 + \sum_{n=1}^{\infty} [A_n \cos(\omega_n t) + B_n \sin(\omega_n t)]$$

    Furthermore,

    $\omega_n = n\omega_0 = 2\pi n / T$, where n is integer

  - "*The multiplicity of vibrational forms which can be thus produced by the composition of simple pendular vibrations is not merely extraordinarily great; it is so great that it can not be greater*." (H. Holmholtz, 1863)

Spectrum (cont.)

- **Fourier Transform**
  - For any function of time, *x(t)*, the Fourier transform *X(ω)* of *x(t)* is defined in terms of the Fourier integral:

$$X(\omega) = \int_{-\infty}^{+\infty} e^{-i\omega t} x(t) dt$$

  - The Fourier transform converts a function of time to a function of frequency
  - Inverse Fourier transform

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\omega t} X(\omega) d\omega$$

$$e^{i\omega t} = \cos(\omega t) + i \sin(\omega t)$$

Pitch and Formant

• Pitch

- Definition: "that attribute of auditory sensation in terms of which sounds may be ordered on a musical scale" (American Standards Association, 1960)

- Pitch is related to the repetition rate of the waveform:
  - For pure tone, pitch corresponds to its frequency
  - For a periodic complex tone, pitch corresponds to its fundamental frequency

• Formant

- A resonance in the vocal tract which is usually manifested as a peak in the spectral envelope of a speech sound

Three characteristics of sound sensation

- **Pitch (frequency)**

- **Loudness (intensity)**

- **Timbre (quality - spectral envelope)**