Sound

- Force - what is the energy source?
- Oscillator - what vibrates?
- Resonator - what "shapes" the sound?

Object vibration (basic mechanics)

\[ \text{total energy} = \text{kinetic energy} + \text{potential energy} \]

\[
\frac{\text{kinetic energy}}{t} = \frac{\text{potential energy}}{t} \quad \text{or} \quad t + \Delta t
\]
Air pressure (wave)

- high
- low
- medium
- medium
- high

At any point \((x, y, z)\), the pressure oscillates over time.

Time snapshot (wave)

- velocity
- high
- low
- high
- low
- high

Pressure: high

Air pressure

\(P(x, y, z, t) = \frac{I_{atm}}{I_{atm}} + I(x, y, z, t)\)

(average air pressure \(I_{atm}\) depends on temperature, altitude, weather,)

For a fixed \((x, y, z)\), we have:

\(P(x, y, z, t)\) pressure

\(P_{atm}\)

\(I_{atm}\)

\(I(x, y, z, t)\)

In general, the sounds we hear are the sum of many waves.

Absolute Threshold of Hearing \(I_0\)

\[\frac{I_0}{I_{atm}} = 10^{-9}\]

Units of \(I\): atmospheres (or Pascals)

Pain Threshold

\[\frac{I_pain}{I_{atm}} = 10^{-3}\]

Physics: Sound energy density at \((x, y, z, t)\) ~ \(I^2\).

\[\log_{10}\frac{I}{I_0} = \text{"Bel"}\]

\[10 \log_{10}\frac{I}{I_0} = 20 \log_{10}\frac{I}{I_0} = \text{"decibel"}\]

The reason for using dB is that the JND is typically ~1 dB (later....)
Examples (dB)

- Jet plane: 120 dB
- Noisy traffic: 90 dB
- Conversation (1 m): 60 dB
- Quiet room: 30 dB
- Recording studio: 10 dB
- Threshold of hearing: 0 dB

Example

If you double the sound pressure level, then what is the dB increase?

\[ \log_{10} \frac{2I}{I_0} = 2 \log_{10} 2 + \log_{10} \frac{I}{I_0} \]

\[ = \log_{10} 2 + 20 \log_{10} \frac{I}{I_0} \]

Examples

Suppose you have two sounds that are 40 dB and 20 dB. What is the intensity if you add these sounds together?

Exercise (Hint: it's not 60 dB)

\[ I(x, y, z, t) \text{ is not arbitrary.} \]

Rather, it must satisfy the wave equation:

\[ \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) I = \frac{1}{v^2} \frac{\partial^2 I}{\partial t^2} \]

**Speed of Sound**

\[ v \approx 340 \text{ m/s} \]

\[ = 34 \text{ m/μs} \text{ (milliseconds)} \]

(depends on temperature)

\[ 330 \leftrightarrow 350 \text{ (cold air / hot air)} \]

- Wave equation + boundary conditions
- \( \Rightarrow \) complicated shadow and reflection effects
- ("diffraction")

plane wave + single slit

sea waves + inlet
Example 1

3D translating sine wave

\[ I(x, y, z, t) = \sin(kx + ky + kz + \omega t) \]

\[ \lambda = \frac{\text{cycles}}{\text{meter}} \]

\[ \omega = \frac{\text{cycles}}{\text{second}} \]

\[ |V| = 340 \, \text{m/s} \]

Exercise:

Verify that \( I(x, y, z, t) = \sin(kx + ky + kz + \omega t) \)

satisfies the wave equation

if \( k_x^2 + k_y^2 + k_z^2 = \frac{\omega^2}{V^2} \).

\[ |V| = \frac{\omega}{\sqrt{k_x^2 + k_y^2 + k_z^2}} \]

\[ \frac{\omega}{V} = \frac{\text{cycles/s}}{\text{cycles/m}} \]

\[ V = \frac{\omega}{k} \]

Example 2a:

Expanding impulse function

\[ d = vt \]

One can show it also satisfies the wave equation. (Details omitted.)

3D impulse function, \( \delta(x, y, z) \)

\[ \iiint_{-\infty}^{\infty} \delta(x, y, z) \, dx \, dy \, dz = 1 \]

\[ I(x, y, z) \text{ at } t = 0 \]

Impulse Function + wave equation produces an expanding sphere.

Energy spread over thin sphere

\[ \text{area} = 4\pi R^2 \]

where \( R = \sqrt{x^2 + y^2 + z^2} \).

Thus,

\[ I^2 \sim \frac{1}{R^2} \]

\[ I \sim \frac{1}{R} \]
\[ t = 0 \]
\[ I(x, y, z, 0) \sim I_{\text{src}} \delta(x, y, z) \]
\[ t > 0 \]
\[ R = vt \]
\[ I(x, y, z, t) \sim \frac{1}{R} I_{\text{src}} \delta(t - \frac{R}{v}) \]
where \[ R = \sqrt{x^2 + y^2 + z^2} \].

**Example 2b: “point source”**

For a general source at \((x, y, z) = 0\), write it as a sum of time-shifted impulse functions:

\[ I_{\text{src}}(t) = \sum I_{\text{src}}(t') \delta(t - t') \]

**Binaural hearing:** if sound arrives from left, how long is the inter-aural delay?

\[ d = vb \Rightarrow t = \frac{d}{v} \]

\[ t = \frac{17 \text{ cm}}{340 \text{ m/s}} = 0.05 \text{ ms} \]

We can detect such small time differences. (Analogous to stereovision.)

**Naive computational model**

- Maximum likelihood:
  
  Find \( \alpha, \tau \) that minimize \[ \frac{1}{T} \sum (I(x, t) - \alpha I_r(t - \tau))^2 \]
  where \( |\tau| < 0.5 \text{ ms} \)

Suppose you find \( \tau, \alpha \). Then what?
Naive model: Cone of confusion

Modelled as a sphere.
All incoming directions \((0, \phi)\) on a cone define the same intensity and timing difference.

Naive computational model

- Maximum likelihood:
  \[
  \text{Find } \alpha, \tau \text{ that minimize } \sum_t \left( I_e(t) - \alpha I_r(t-\tau) \right)^2
  \]
  where \(|\tau| < 0.5 \text{ ms}

- Use \(\tau\) to estimate \(\phi\)

(Exercise)

Why naive?

- Head, shoulders, outer ears significantly reshape the sound waves before they are measured inside the two ears.

("head related impulse response," will be discussed in future lectures)