

# Sound Localization

---

**Binaural hearing, HRTF, and spatial perception**

**Lorenz Schwarz**

Karlsruhe University of Arts and Design (HfG)

Winter Semester 2024/25

[Course info](#)

[← Chapters](#) · [Download PDF ↓](#)

*For iPhone/iPad users, the PDF download is recommended.*

# SOUND LOCALIZATION

---

Lorenz Schwarz · HfG Karlsruhe · CC BY 4.0

# Spatial Hearing in Audio Practice

---

Spatial hearing informs how we capture, shape, and evaluate sound scenes:

- Microphone technique and stereo array selection (localization cues, stereo width)
- Panning, depth, and spatial processing (level/time differences, reverberation cues)
- Immersive and binaural production workflows
- Assessment of monitoring and room acoustics (imaging, translation)

# BINAURAL HEARING

---

Lorenz Schwarz · HfG Karlsruhe · CC BY 4.0

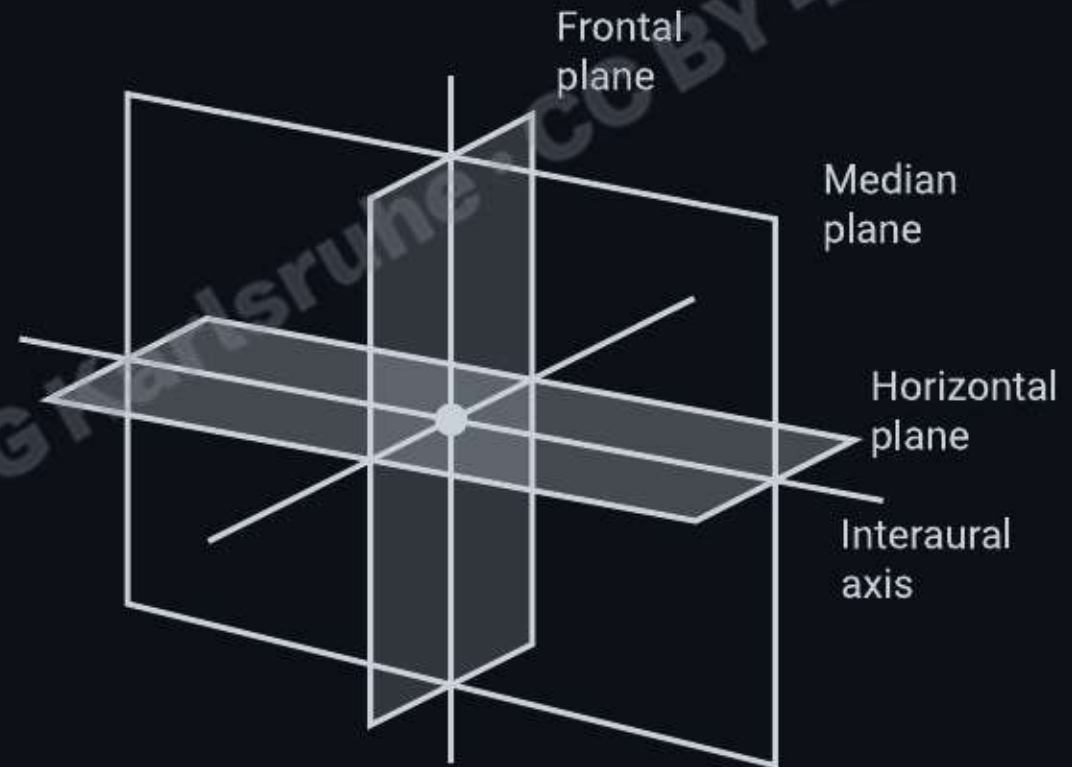
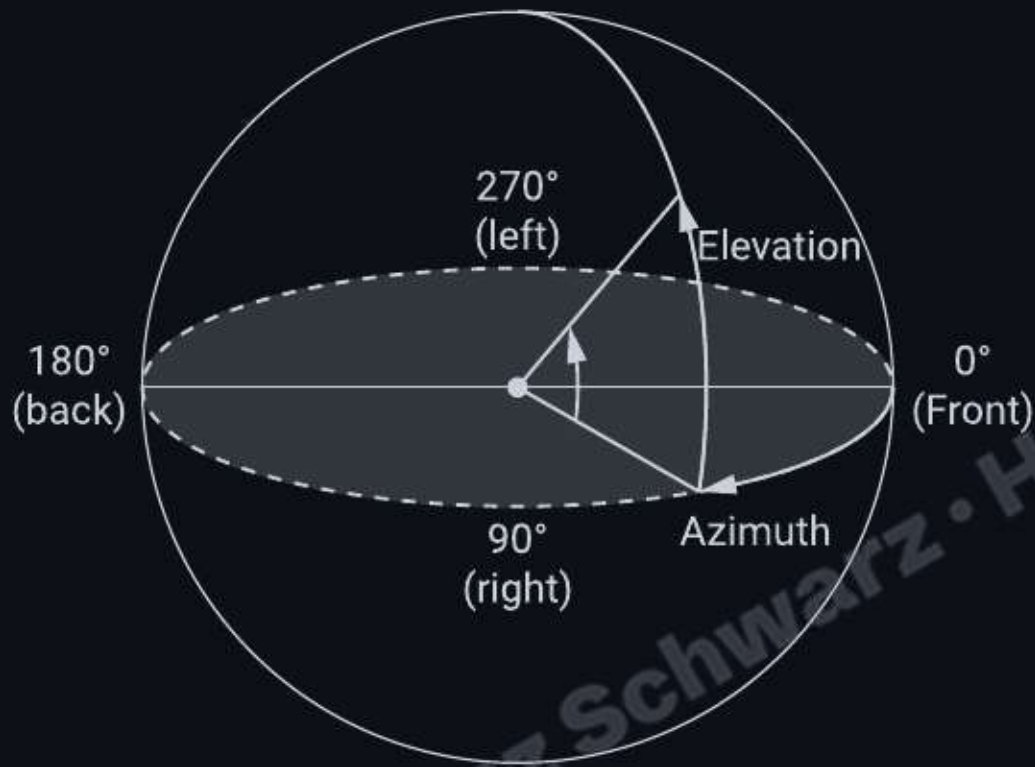
# Spatial references in binaural hearing

---

- Azimuth (horizontal angle)
- Elevation (vertical angle)
- Distance

Anatomical reference planes: frontal (front/back), horizontal (upper/lower), median (left/right)

Lorenz Schwarz | H10 Karlsruhe | CC BY 4.0



Horizontal coordinate system (left) and anatomical planes (right)

# Lateralization and localization

- Lateralization: Perceived left-right position of a sound presented over headphones, typically experienced inside the head.

▶ Lateralization (listen via headphones)

- Localization: Perception of a sound source at a specific position in external space, as with loudspeakers or real sound sources.

▶ Localization (listen via headphones)



# Binaural hearing and source localization

---

The first theory of lateral sound localization was proposed by Lord Rayleigh as "duplex theory." Later extended with Blauert's spectral cues.

- **Interaural Time Difference (ITD):**
  - Difference in arrival time between ears. Effective for low frequencies (<1500 Hz) because long wavelengths maintain phase coherence between ears.
- **Interaural Level Difference (ILD):**
  - Difference in sound intensity between ears. Effective for high frequencies (>1500 Hz) due to head shadow effect (short wavelengths are blocked by the head).
- **Spectral Cues:**
  - Additional information from how the head and pinnae filter sound (HRTF).

# Interaural time difference (ITD)



# Interaural time difference (ITD)

---

ITD helps localize low-frequency sounds.

For an ear distance of  $\lambda \approx 22 \text{ cm}$ , the natural interaural time delay is approximately  $\approx 660 \mu\text{s}$

$$f = \frac{c}{\lambda} = \frac{343 \frac{\text{m}}{\text{s}}}{0.22\text{m}} \approx 1500\text{Hz}$$

- Phase delays at low frequencies (if wavelength is greater than half the distance between the ears  $< 800\text{Hz}$ )

► ITD noise bursts (listen via headphones)

# Interaural level difference (ILD)

---

Interaural Level Difference (ILD), also known as Interaural Intensity Difference (IID), plays a key role in the localization of high-frequency sounds.

- Low frequencies bend around the head with minimal attenuation
- High frequencies are significantly attenuated due to the head shadow effect
- Level differences increase above ~1600 Hz

▶ **ILD noise bursts (listen via headphones)**

# Interaural level difference (ILD)



Lorenz Schwarz · HfG Karlsruhe · CC BY 4.0

# ITD and ILD

## Localization accuracy:

- 1° for sources in front
- 15° for sources to the sides

## Frequency ranges:

- ITD dominant for  $f < 800$  Hz
- Both contribute between 1000-1500 Hz
- ILD dominant for  $f > 1600$  Hz
- No localization below 80 Hz



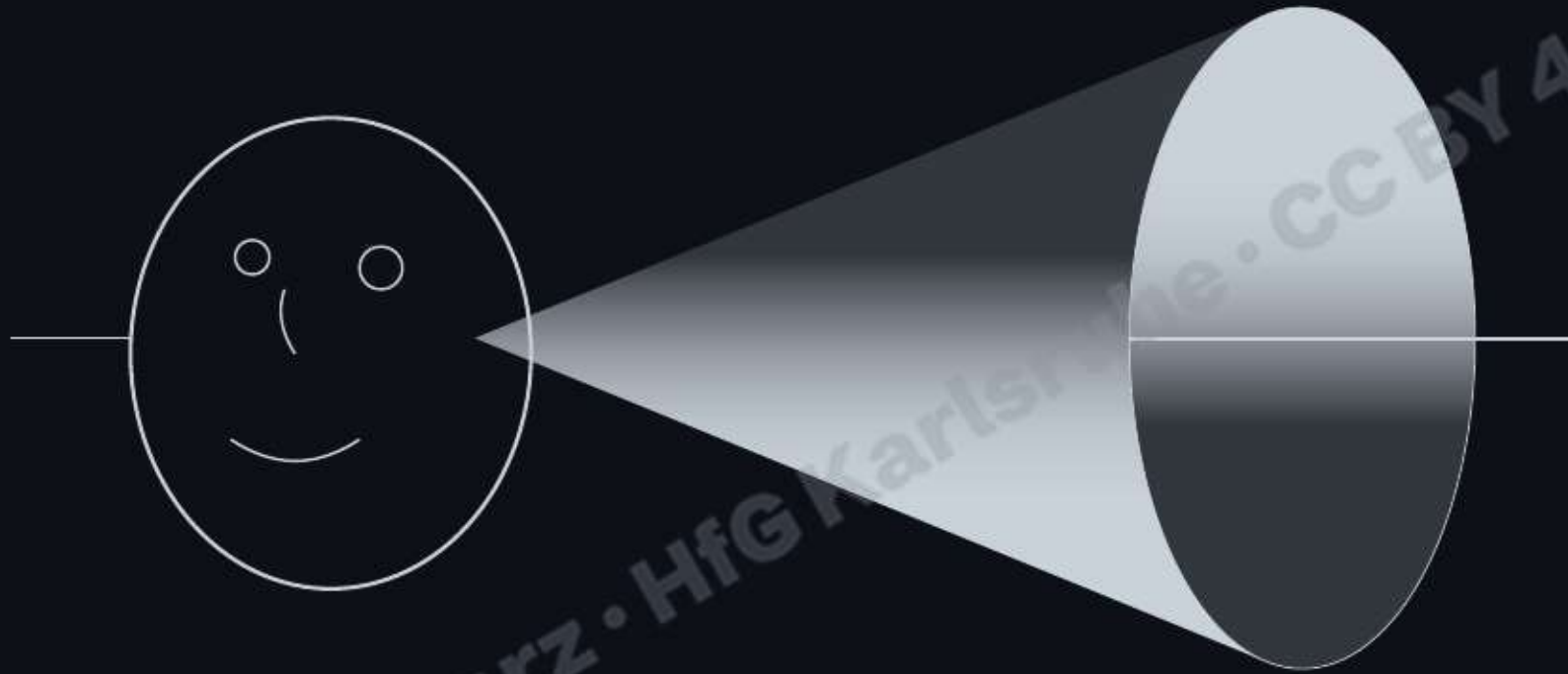
White noise at  $-90^\circ$  azimuth: maximum ITD of 0.63 ms

# Cone of confusion

---

- **Front-back ambiguity:** Cannot discriminate between sounds originating from the front or rear.
- **No elevation cues:** ITD and ILD do not provide information about vertical localization (elevation).

Lorenz Schwarz · HfG Karlsruhe · CC BY 4.0



A region in space where sound sources produce identical ITDs and ILDs, making localization ambiguous.

# Resolving ambiguity

---

- **Tilting the head:** Introduces new timing and level information to help resolve the location of the sound source.
- **Spectral cues:** The filtering effects of the pinnae and torso shape the sound spectrum, helping to distinguish elevation and front-back positioning.

Lorenz Schwarz · HfG Karlsruhe · CC BY 4.0

# SPECTRAL CUES

---

Lorenz Schwarz · HfG Karlsruhe · CC BY 4.0

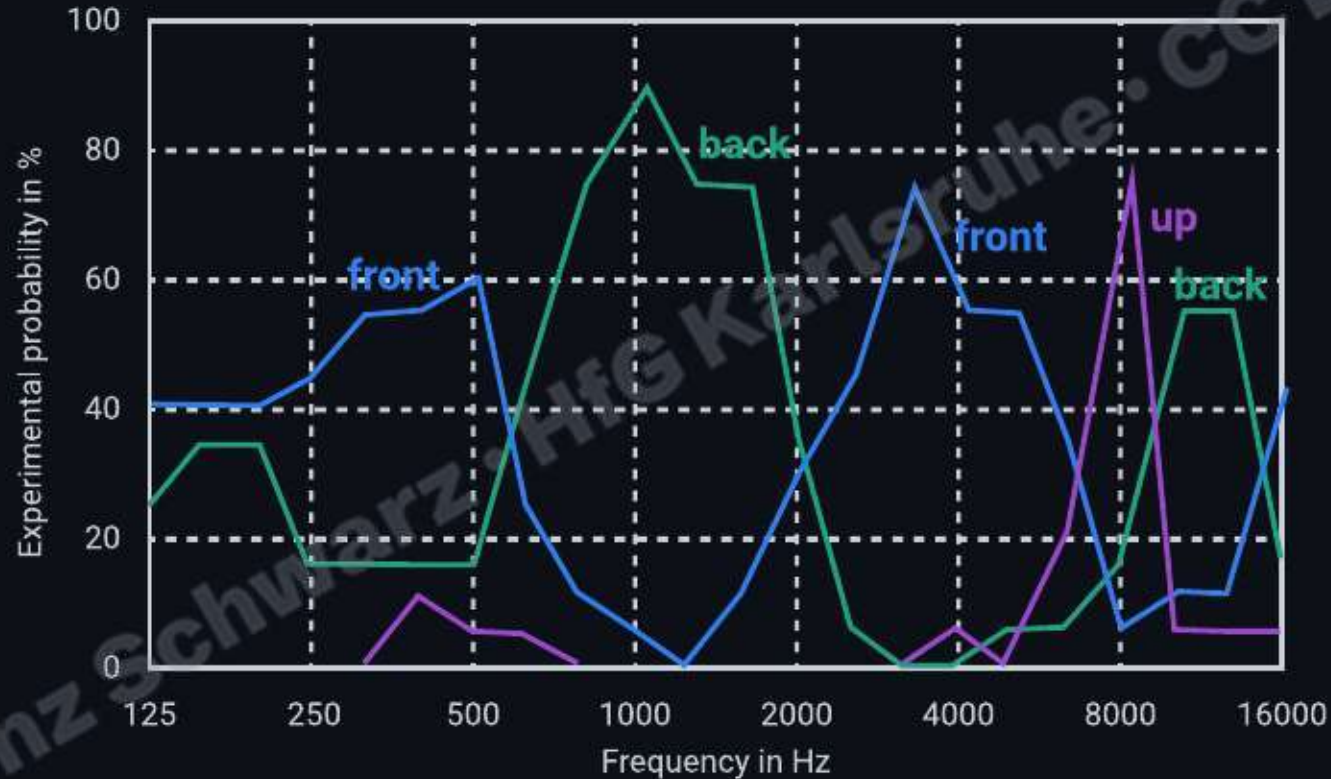
# Spectral cues

---

The outer ear, auricle (pinna), and ear canal act as a resonator system, shaping sound based on its direction of incidence.

- Frequency-dependent filtering provides spatial information, aiding in sound localization.
- Encodes vertical and front-back localization cues (median plane), primarily through frequency bands affected by the pinna.

# Directional bands (after J. Blauert)



► Pink noise back and front ear level and above (listen via headphones)

# Head related transfer function (HRTF)

---

HRTF describes how spatial audio cues are encoded in the sound reaching the ears, allowing for sound source localization.

The torso, head, and pinna act as direction-dependent filters, introducing frequency-specific alterations to the sound.

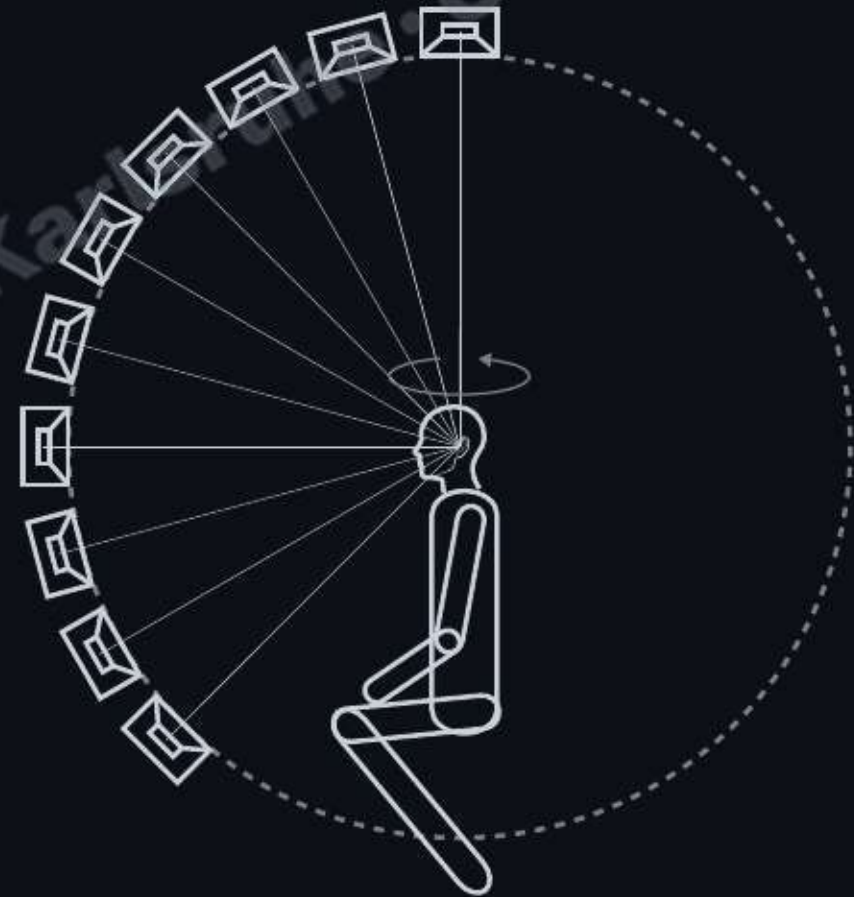
→ *This effect can be mathematically represented as a transfer function.*

▶ Clicks rotating 360° up - down (listen via headphones)

# HRTF measurement

---

HRTFs are measured at small angular increments in an anechoic chamber, with interpolation used to estimate unmeasured positions.



# Applications of HRTF

---

- 3D audio rendering
- VR & AR
- Binaural recording

→ *The auditory system adapts to a modified head-related transfer function over time.*

Lorenz Schwarz · H20 Karlsruhe · CC BY 4.0

# Franssen effect

---

The Franssen effect demonstrates how the auditory system localizes sound based on attack transients rather than sustained energy.

A tone begins in one speaker with a sharp onset, then continues from the opposite speaker. Despite the energy source switching sides, the perceived location remains anchored to the initial onset.

**Note:** This demonstration requires stereo speakers with sufficient separation. It will not work with headphones.

▶ [Play Franssen effect](#)

# Franssen effect



# Auditory distance perception

---

The auditory system has limited ability to determine distance, relying on:

- **Initial Time Delay Gap (ITDG):** Time between direct sound and first reflection
- **Direct-to-reverberant ratio:** Closer sources have more direct sound
- **Reverberation density:** More diffuse reflections indicate greater distance
- **High-frequency absorption:** Distant sounds lose high-frequency content
- **Loudness:** Closer sources perceived as louder
- **Motion parallax:** Closer sources shift position faster for moving listeners

▶ Distance close - far

# STEREOPHONIC REPRODUCTION

---

Lorenz Schwarz · HfG Karlsruhe · CC BY 4.0

# Stereophonic Sound

---

When two channels are played through separate speakers, listeners perceive a soundstage extending between those speakers.

- **Summing localization:** Brain combines signals from both speakers to create phantom sound sources
- **Phantom center:** Virtual sound source perceived between the speakers

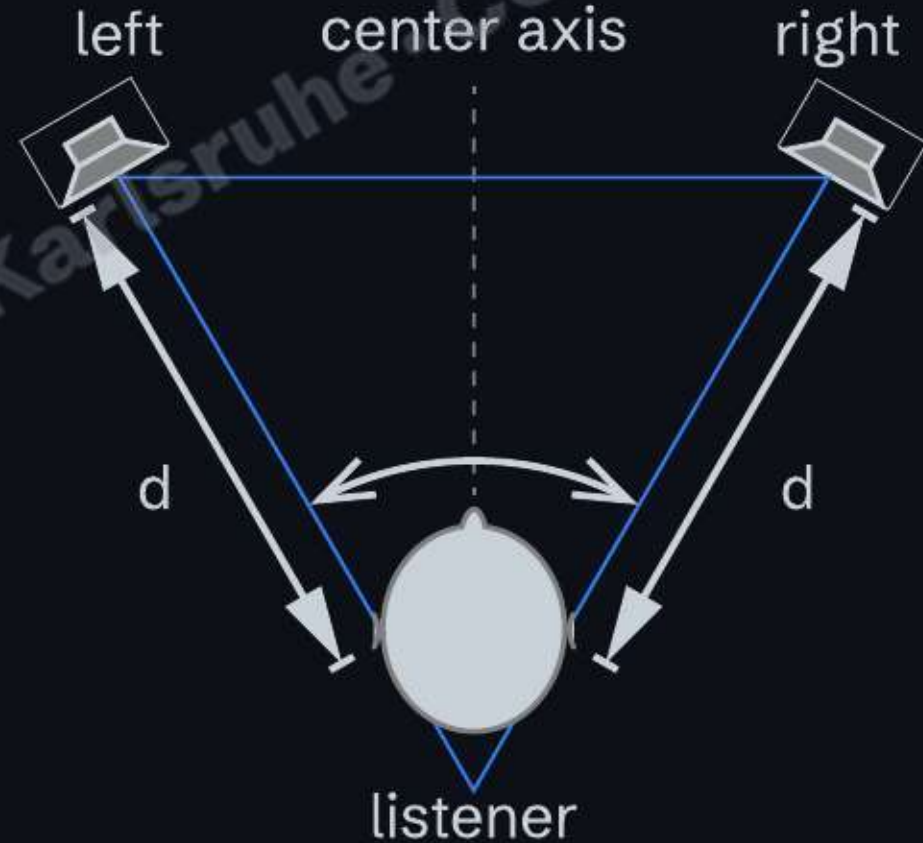
Creates an illusion of multi-directional spatial perspective.

# Artificial stereo and phantom center

The recommended placement forms an equilateral triangle: each speaker and the listener at equal distances:

- Sweet spot or reference listening position

▶ Panning from left to right



# Dual-mono signal

---

Mono source material played back through two stereo channels with identical signals on both the left and right channels.

- Identical waveform in both L and R channels
- Creates phantom center image when played through speakers
- Perceived as center image in headphones (lateralization)
- Common in broadcast, mono recordings, and centered mix elements

# Summing localization

Two speakers create a phantom sound source between them by manipulating the same binaural cues the brain uses for natural localization:

- **Amplitude panning** (ILD)
- **Time delay** (ITD)

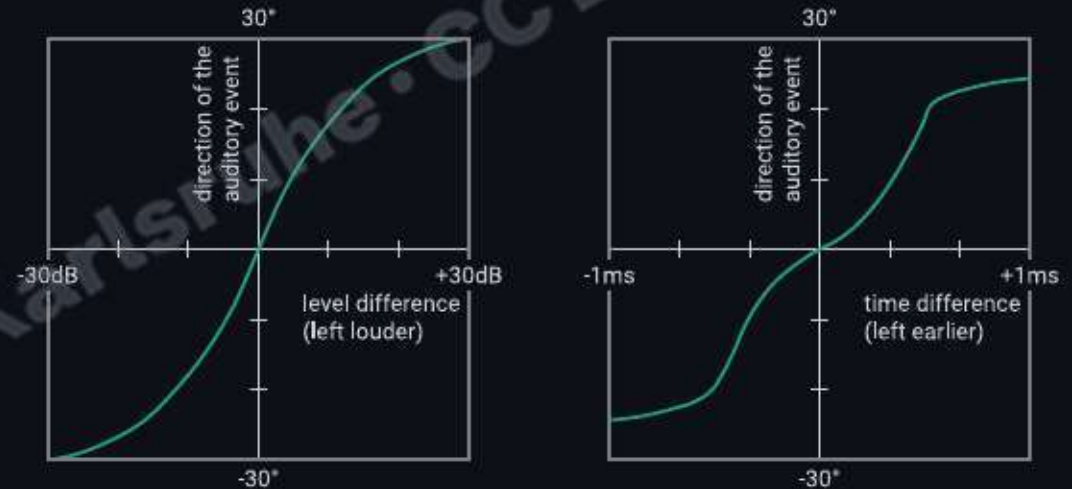


Diagram: perceived source location (Wendt, 1963)

# Stereo recording

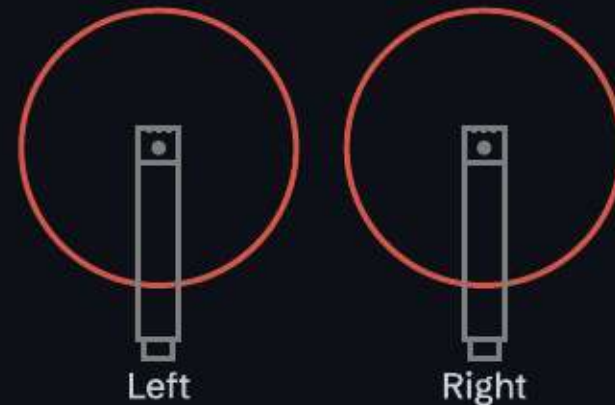
ITD and ILD are used in common stereo recording practices:

## Stereo recording techniques:

- Coincident pair (X-Y): Accurate localization
- Spaced pair (A-B): Wide stereo image



XY: level differences



A-B: time-of-arrival stereo

# Summary: sound localization

---

- ITD (time differences) for low frequencies
- ILD (level differences) for high frequencies

Ambiguity (Cone of confusion) resolved through head movement and spectral cues (HRTF) for elevation and front-back.

## Applications:

- Stereo reproduction relies on phantom imaging
- HRTF measurements enable binaural 3D audio

# Copyright and Licensing

---

Original content: © 2025 Lorenz Schwarz

Licensed under [CC BY 4.0](#). Attribution required for all reuse.

Includes: text, diagrams, illustrations, photos, videos, and audio.

**Third-party materials:** Copyright respective owners, educational use.

**Contact:** [lschwarz@hfg-karlsruhe.de](mailto:lschwarz@hfg-karlsruhe.de)

[← Chapters](#)