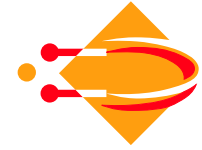


# Sound: hearing and processing

---



Introduction to

**Sound**

The human auditory system

Digital representation

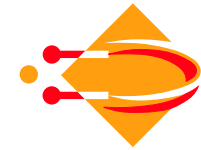
Digital processing

Mads Nielsen, ass. prof.

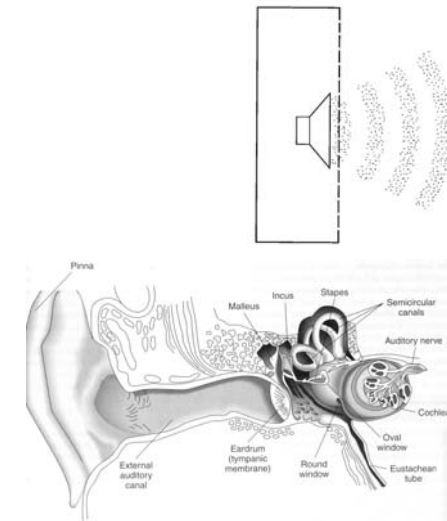
Martin Lilholm, ph.d. stud.

Torben Vaarby Laursen, teaching assistant

# Overview

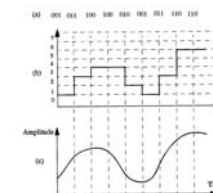


**Physics:** sound waves, sound propagation.



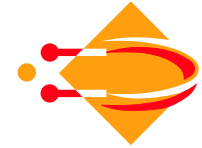
**Psychophysics:** outer ear, middle ear, inner ear, neural processes.

**Digital sound:** sampling, quantization, filtering.



# Sound waves

---



A sound is a travelling **air pressure** wave.

Formally  $p(x, t)$  is the pressure as a function of position  $x$  and time  $t$ .

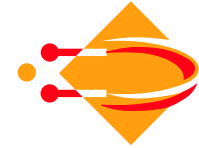
The dynamics of air pressure is governed by Newton second law:

$$F = M \cdot a \quad \Leftrightarrow \quad \frac{\partial^2 p}{\partial x^2} = \rho \frac{\partial^2 p}{\partial t^2}$$

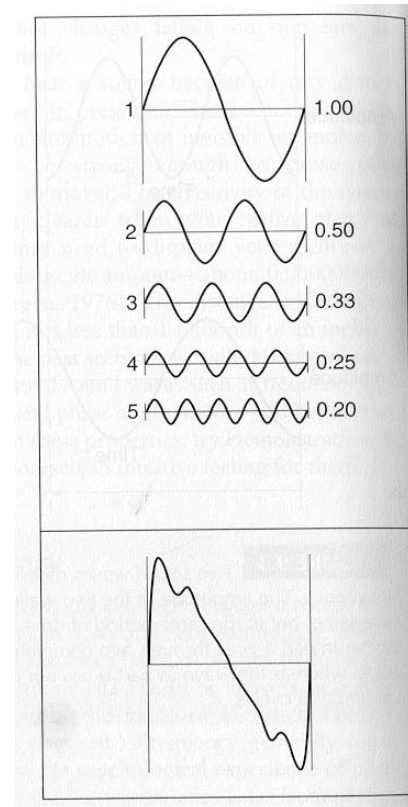
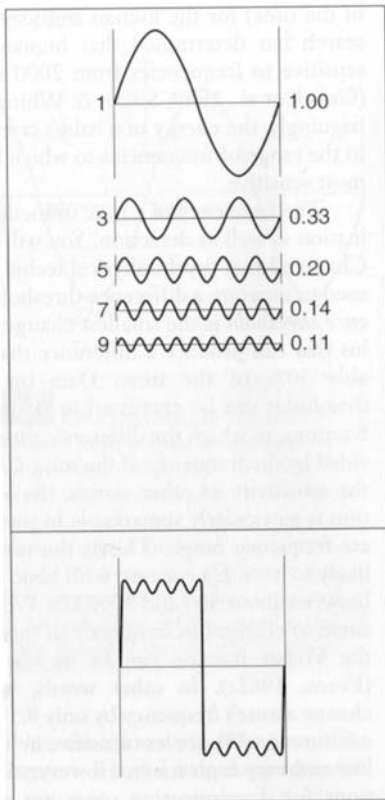
This is solved by waves  $p(x, t) = A \sin(kx - \omega t + \phi)$ .

This is a wave of amplitude  $A$  (dyne), speed  $c = \frac{\omega}{k}$  (m/s), and frequency  $f = \frac{\omega}{2\pi}$  (Hz = s<sup>-1</sup>).

# Sounds

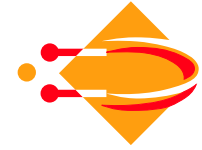


A single sound is composed of the superposition of several waves.



# Sound propagation

---



The **amplitude** of the wave corresponds to the **loudness** of a sound.

In a  $D$  dimensional space, the amplitude  $A$  decreases with distance  $r$ :

$$A(r) = \frac{A_0}{r^{D-1}}$$

Normal space is 3D, a tube is essentially 1D.

Normally loudness is expressed in **decibel (dB)** defined as

$$\text{number of dB} = 20 \log_{10} \frac{P}{P_0}$$

where  $P_0 = 0.0002 \text{ dynes/cm}^2$  (just hearable sound at 1000 Hz).

# Psychophysics

---



We must measure the sensation or perception

**Sensation:** Immediate and basic experience generated as stimuli fall on our sensory system.

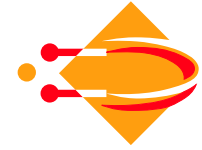
**Perception:** The interpretation of sensations, giving them meaning and organisation.

We must measure the subjective experience.

We use **indirect measurements** as we can only measure feedback from the observer such as: answer yes/no, answer stronger/fainter, more detailed answer, reaction time, ...

# Psychophysical measurements

---



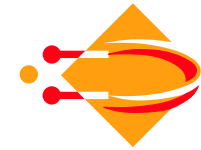
We may measure **detection thresholds** under varying noise conditions.

We may measure **Weber fraction**: just detectable difference in signal.

Mathematical tool: **Signal detection theory**

# Amplitudes

---



Some amplitudes measured in decibels.

---

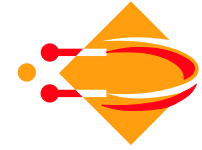
Level	dB	Example
	160	Loudest rockband on record
Intolerable	140	Jet airplane taking off
	120	Very loud thunder
Very noisy	100	Heavy automobile traffic
Loud	80	Loud music from radio
Moderate	60	Average conversation
Faint	40	Quite neighbourhood
	20	Soft whisper
Very faint	0	Softest detectable noise

---

Source: Matlin & Foley: Sensation and Perception

## Sound detection thresholds

---



Young adults can detect sounds in the interval 20-20,000 Hz. Older people may have difficulties with high frequencies.

The Weber fraction is lowest (0,3%) in the interval 500-2000 Hz. (Singing is 70-1000 Hz).

Detection thresholds is lowest (0.0002 dynes/cm<sup>2</sup>) in the interval 2000-5000 Hz. (Crying baby is 3000 Hz).

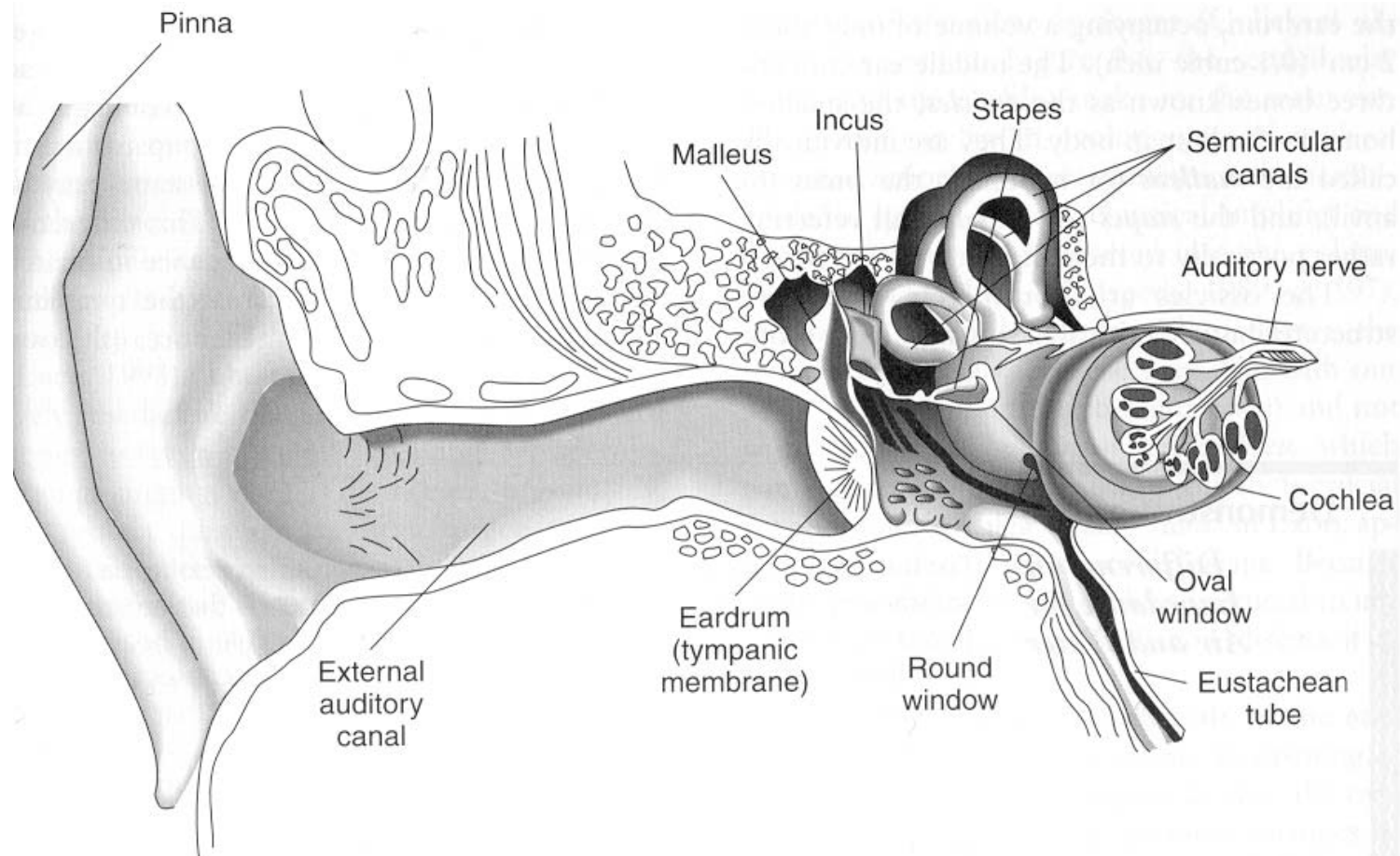
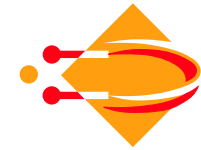
The dynamic range is  $10^7$ .

The Weber fraction in loudness is about 0.3.

The time resolution is ?

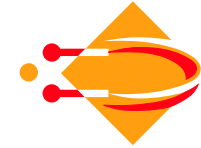
Optimally we would like to know  $\Delta t$ ,  $\Delta f$  and  $\Delta A$  as a function of  $f$  and  $A$ .

# The anatomy of the ear



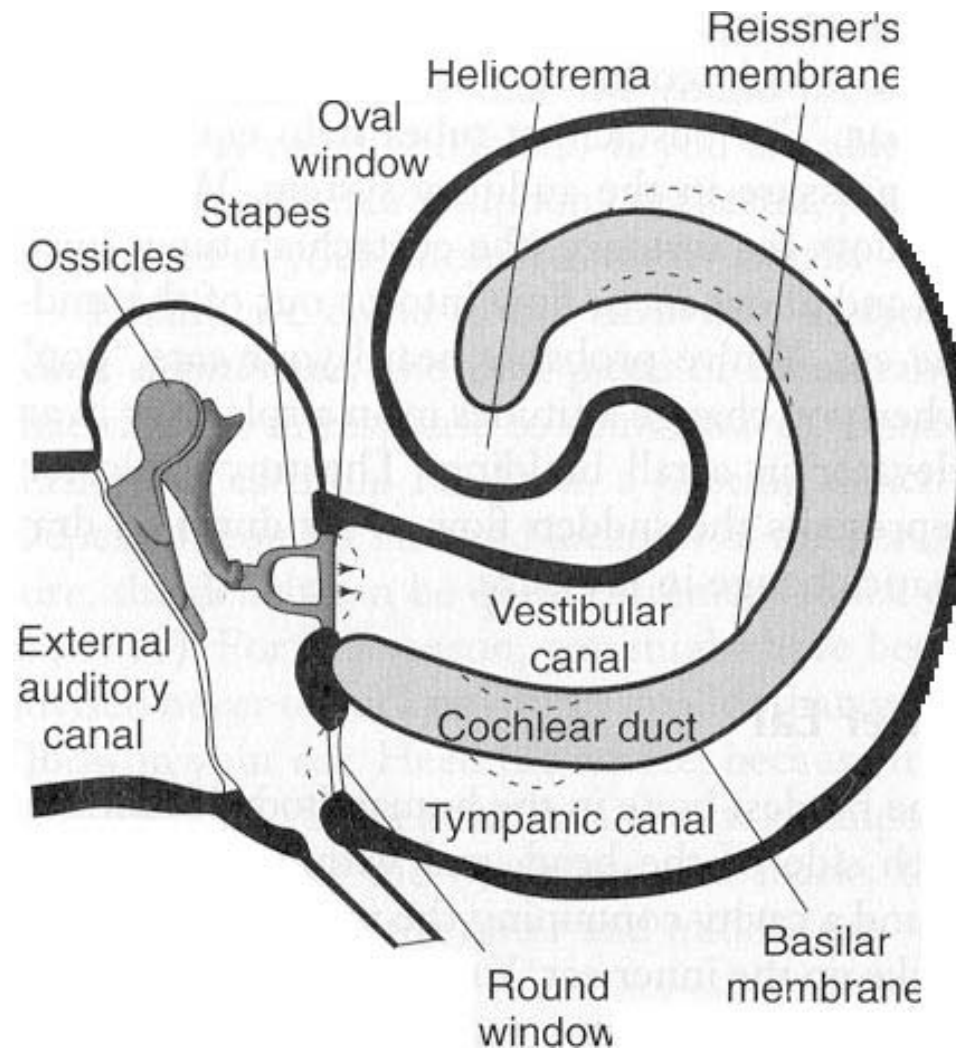
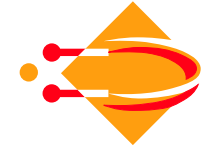
# The function of the ear

---

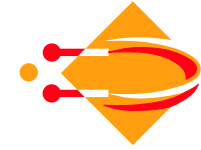


- The outer ear:
  - External tissue (pinna) slightly increase amplitude, help determine direction
  - External channel (25mm×8mm) amplifies (3000Hz 10dB)
  - Eardrum interfaces to the middle ear
  
- The middle ear (2 cm<sup>3</sup>):
  - Filled with liquid (impedance mismatch 30 dB)
  - 3 bones (hammer, anvil, stirrup) amplifies actively
  - The stirrup is connected to the oval window
  - The oval window that connects to the inner ear is smaller (17×) than the eardrum
  - The eustachian tube connected to the throat levels pressure

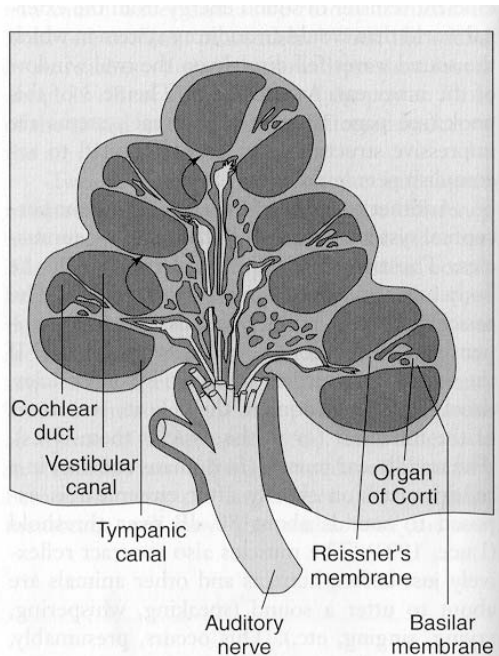
# The middle and inner ear



# The function of the inner ear



- The inner ear (mechanical to neural conversion):

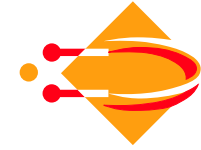


The cochlea (snail) contains receptors

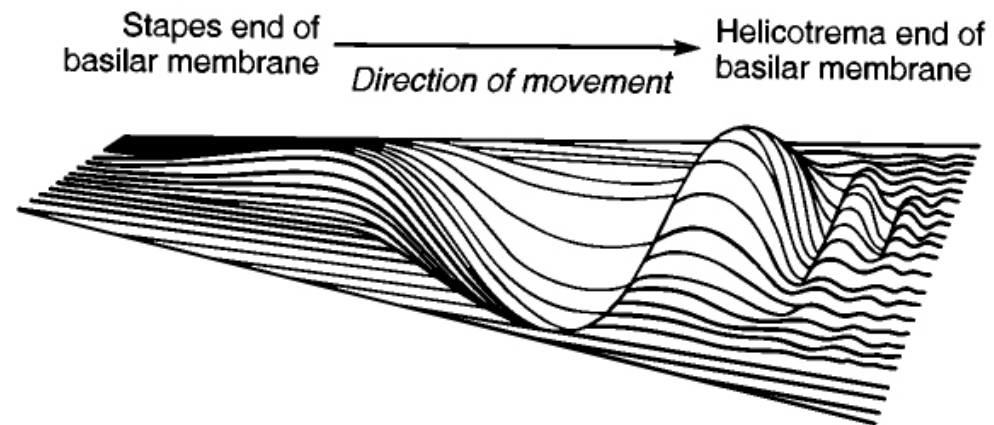
The semicircular canal deals with  
sense of orientation

The snail contains the vestibular canal,  
the tympanic canal, and the cochlear duct

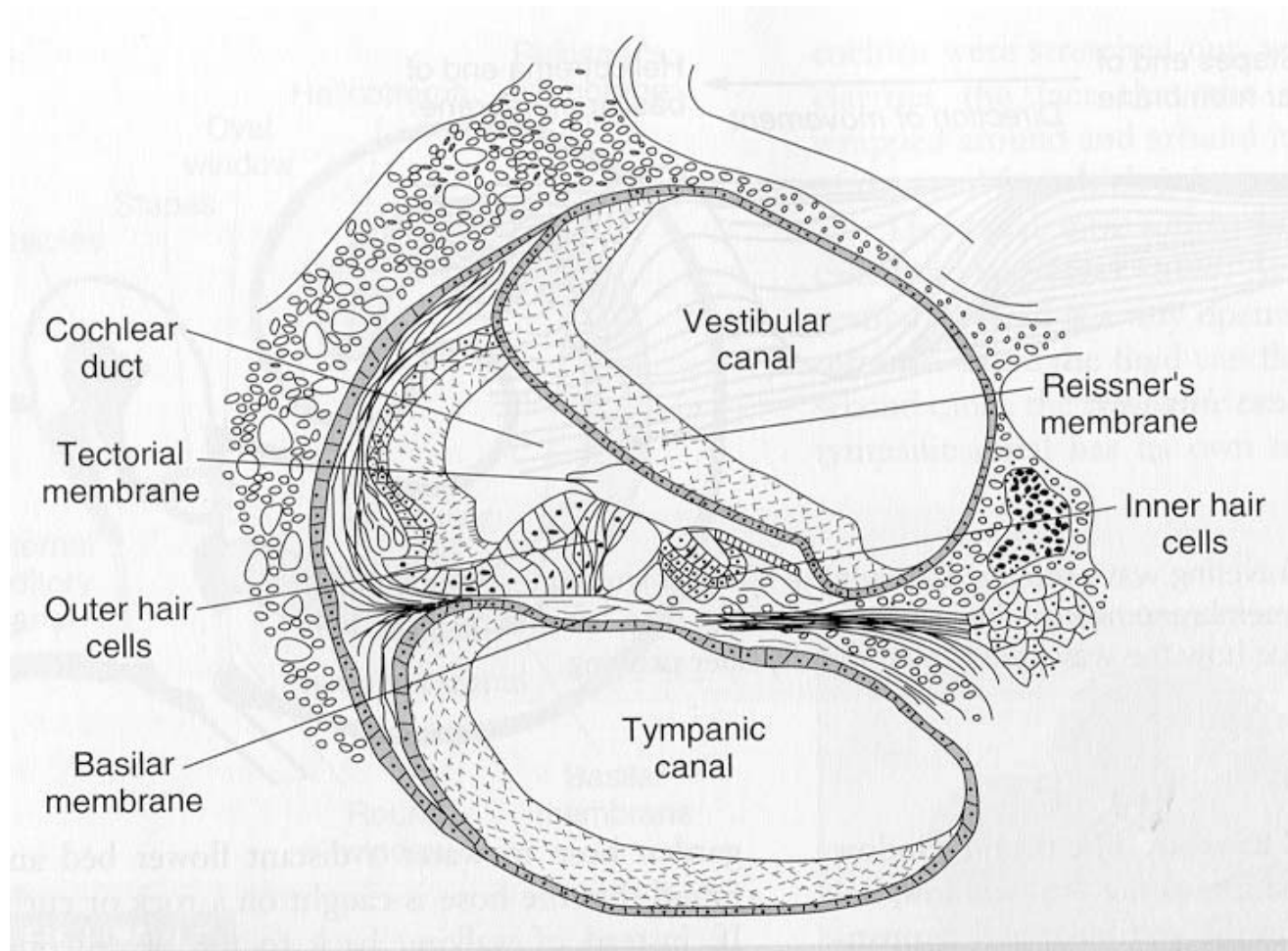
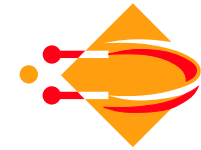
# The cochlear duct



- The cochlear duct:
  - Is separated from the tympanic canal by the basilar membrane.
  - It contains the receptors, the hair cells
  - The basilar membrane moves under pressure changes

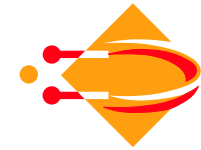


# The cochlear duct



# The mechanical to neural conversion

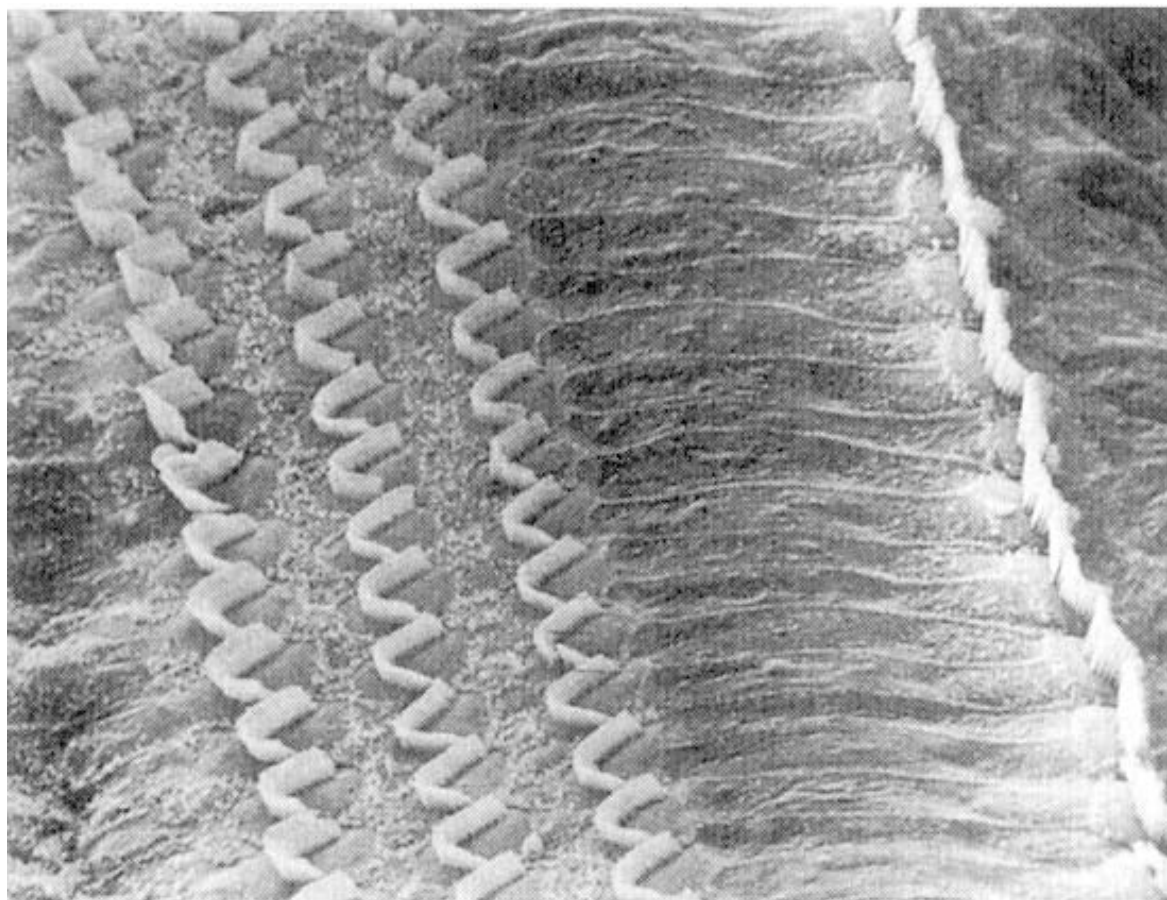
---



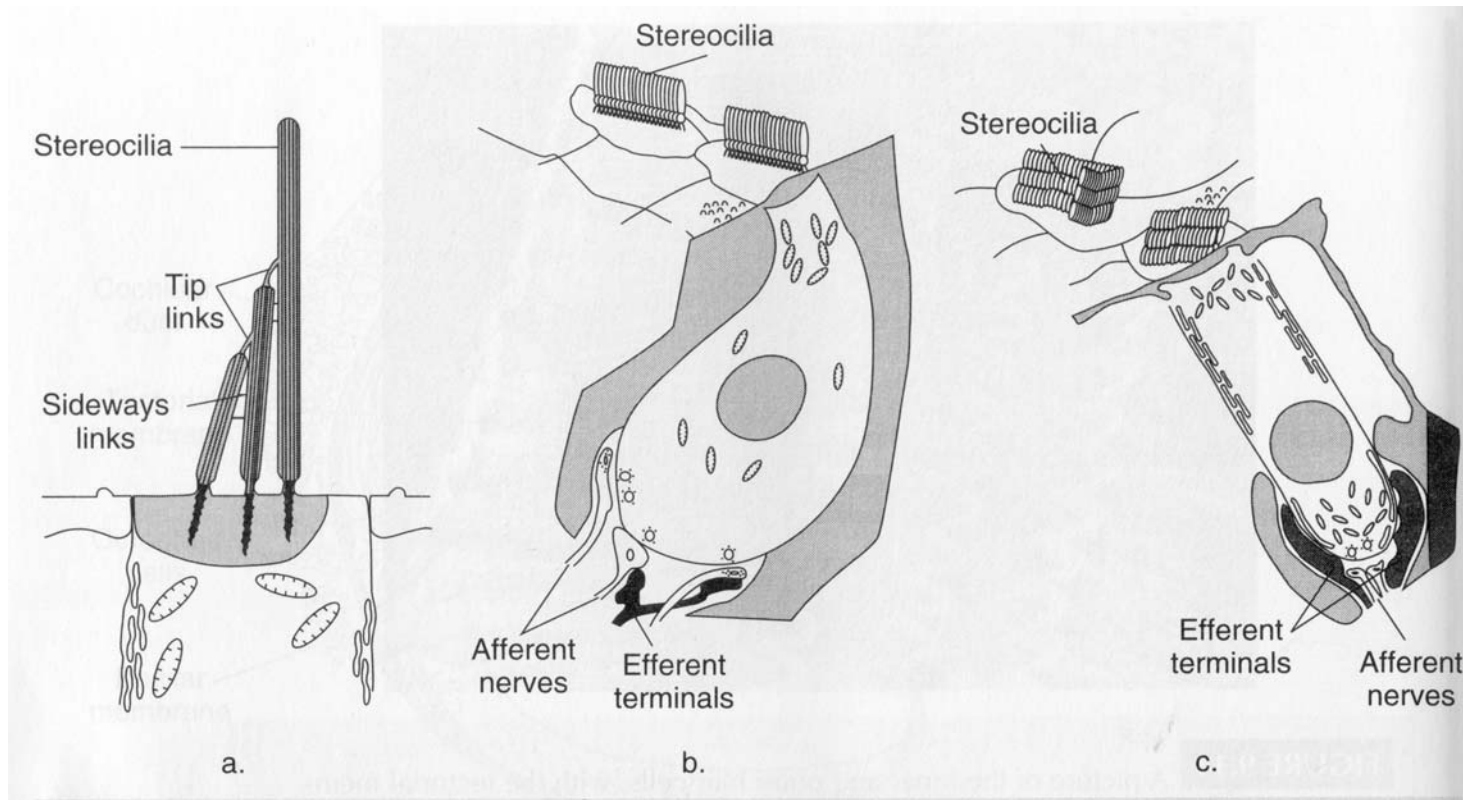
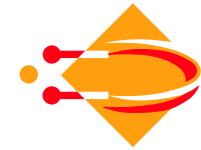
- The hair cells
  - Hair cells connect to the **auditory nerve**
  - Hairs (connected to the membranes) are called stereocilia
  - They have a thickness of 50 nm
  - A displacement of 0.3 nm produces an auditory signal
  - Functionality based upon mechanical openings of ion canals
- The inner hair cells
  - Form a single row
  - They are approximately 3,500 cells
  - They are responsible for the primary auditory signal
- The outer hair cells
  - Form three rows
  - They are approximately 12,000 cells
  - They give us the very fine frequency discrimination

# The hair cells

---

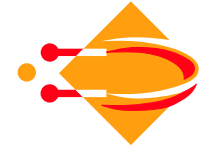


# The hair cells



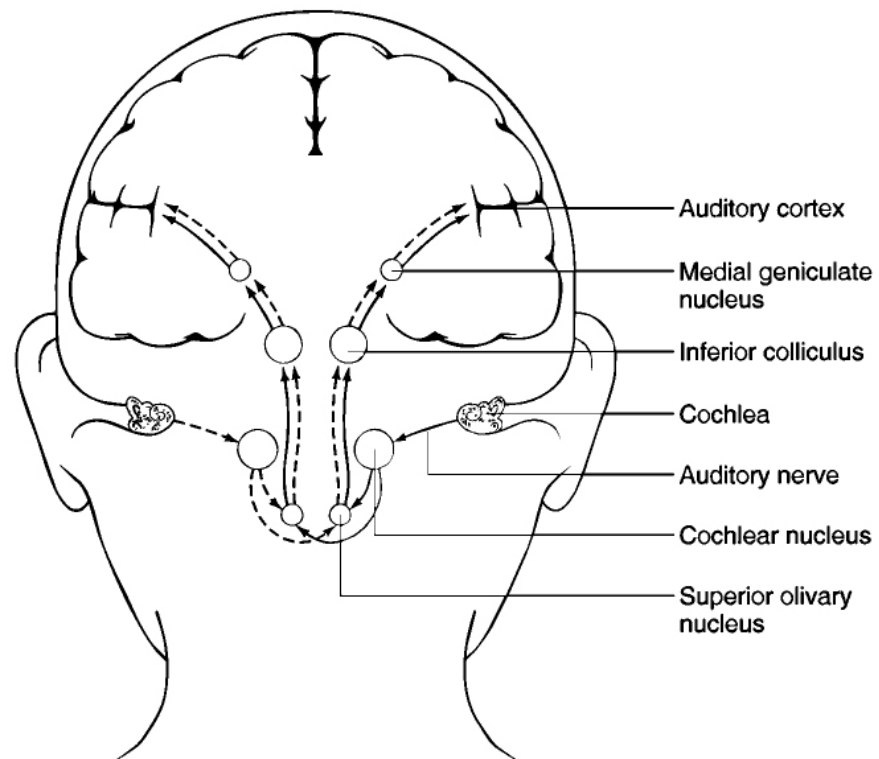
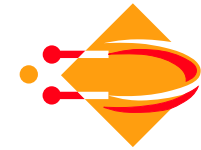
# The auditory nerve

---

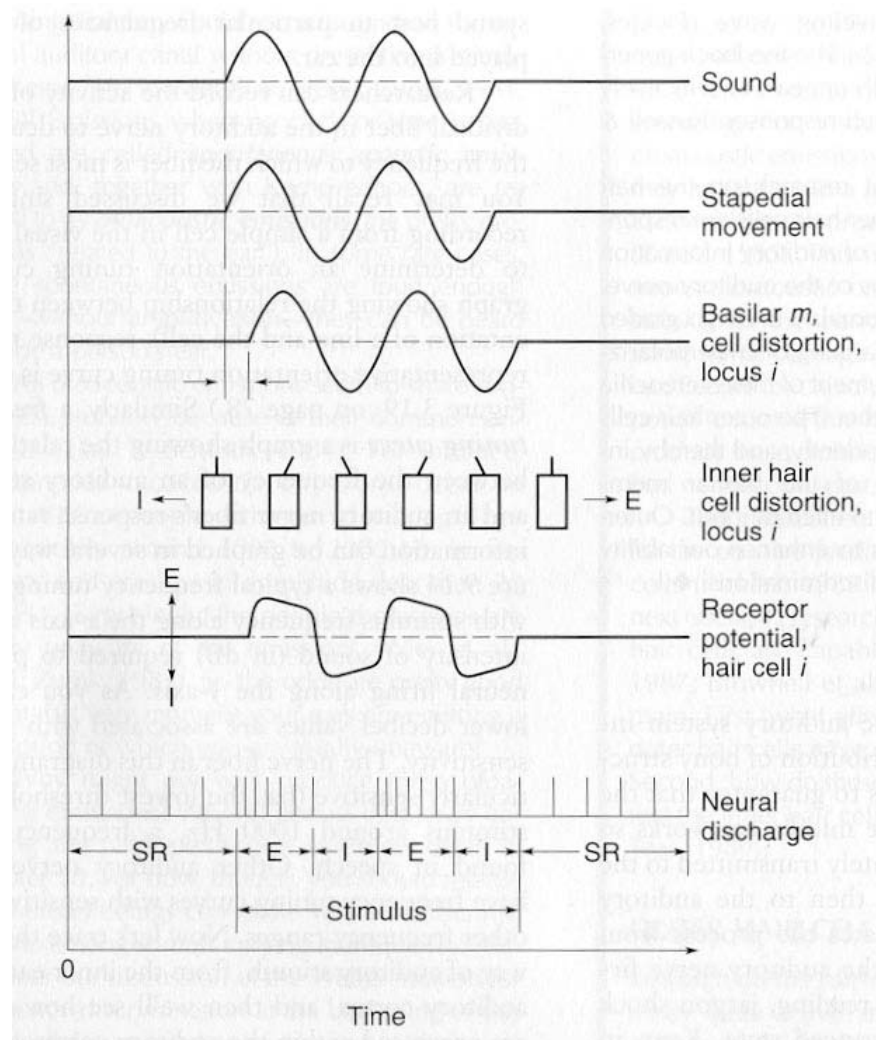
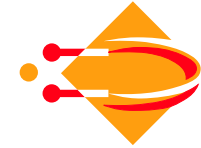


- The afferent fibres
  - Carries information from the inner ear to the auditory cortex
  - The auditory nerve contains 32,000 afferent fibres
  - The inner hair cells connect to 90-95 %
  - The 12,000 outer hair cells only connect to 5-10%
- The efferent fibres
  - Carries information from the cortex to the inner ear
  - Only few fibres present
  - Crucial for functioning

# The auditory pathway

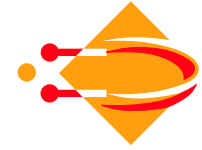


# The signal processing



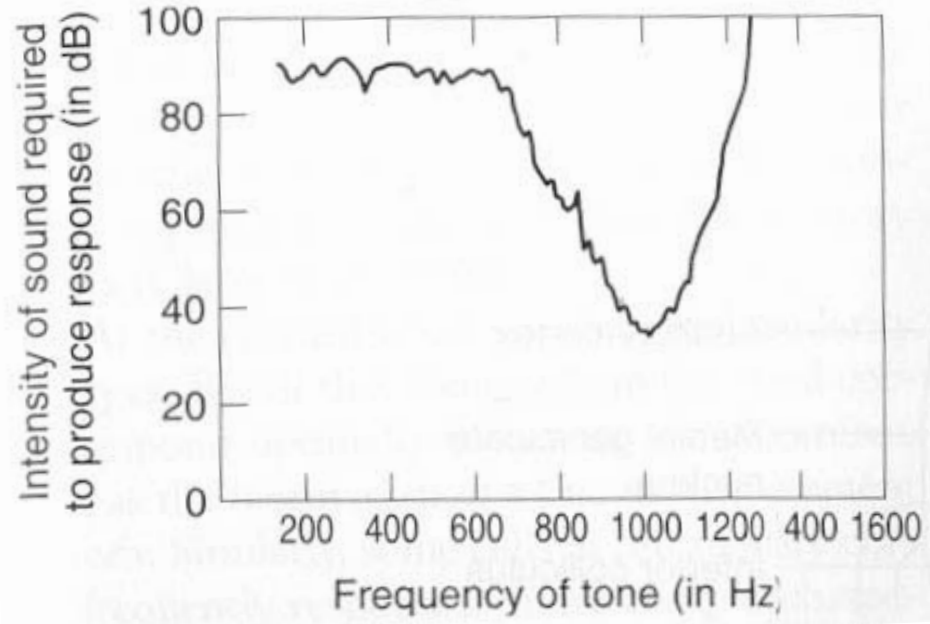
# The auditory cortex

---

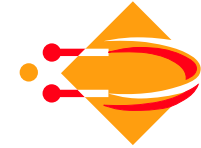


Is organized tonotopically

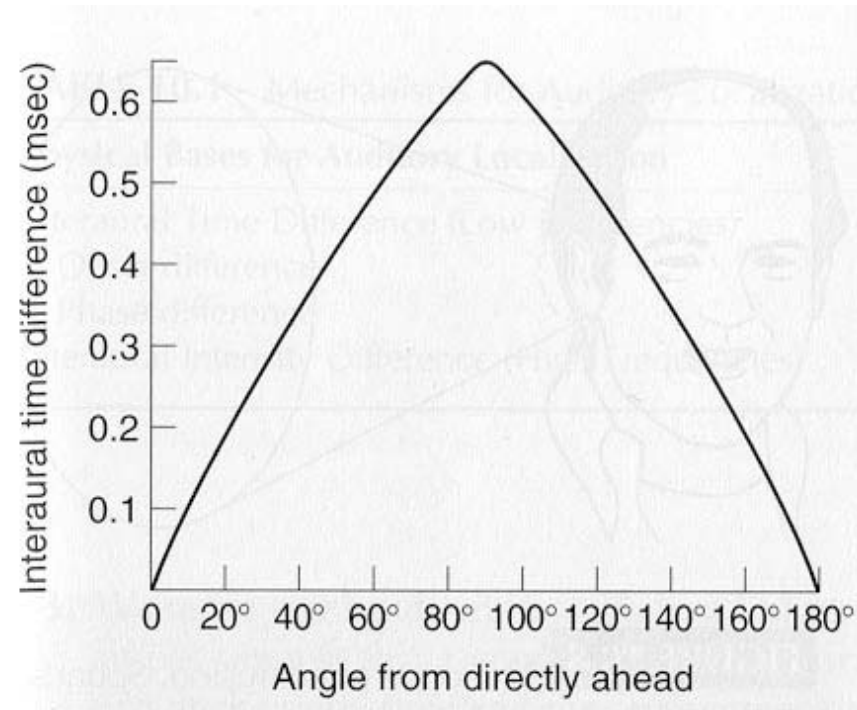
Handles sound localization and speech perception



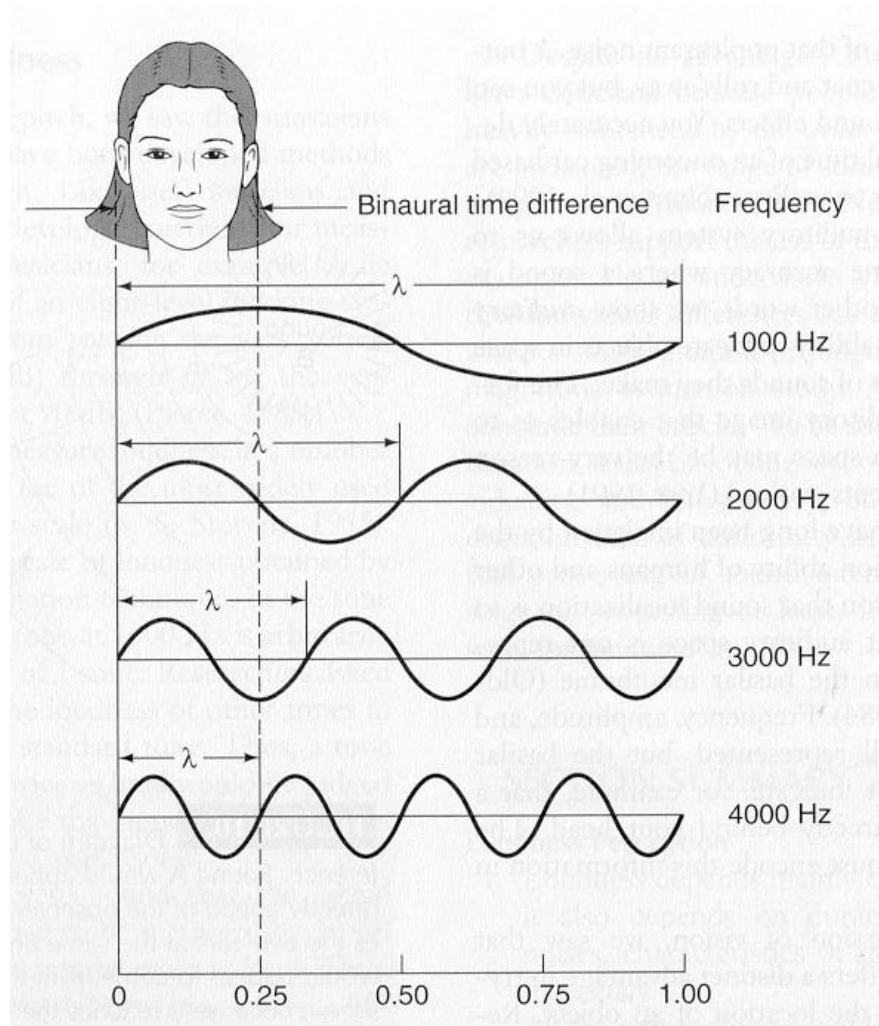
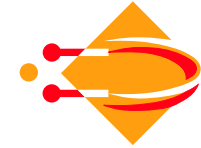
# Binaural hearing



A sound meets the ears at different time instances:

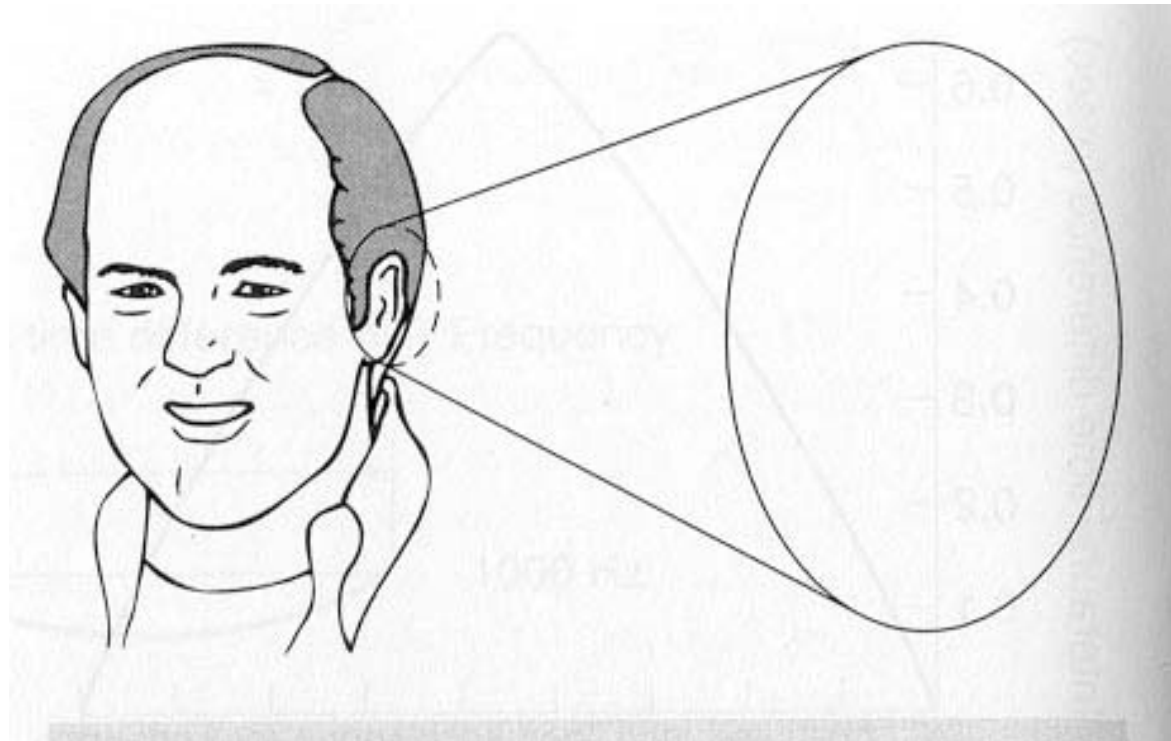
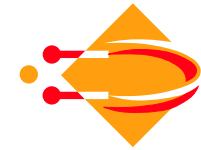


# Binaural time difference

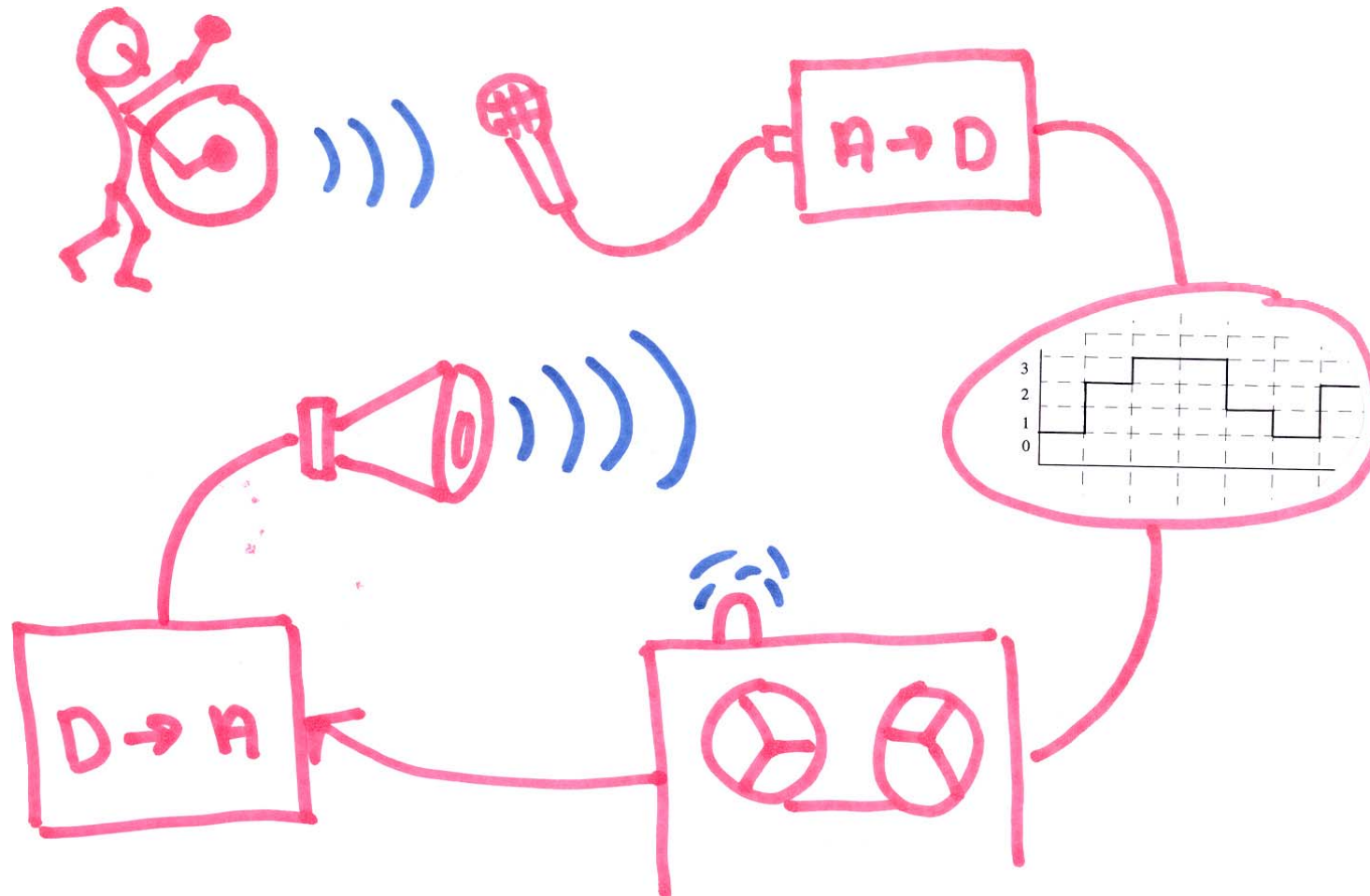
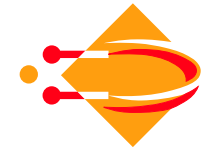


# The cone of confusion

---

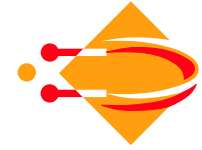


# Digital signal processing



# Analog to digital conversion

---



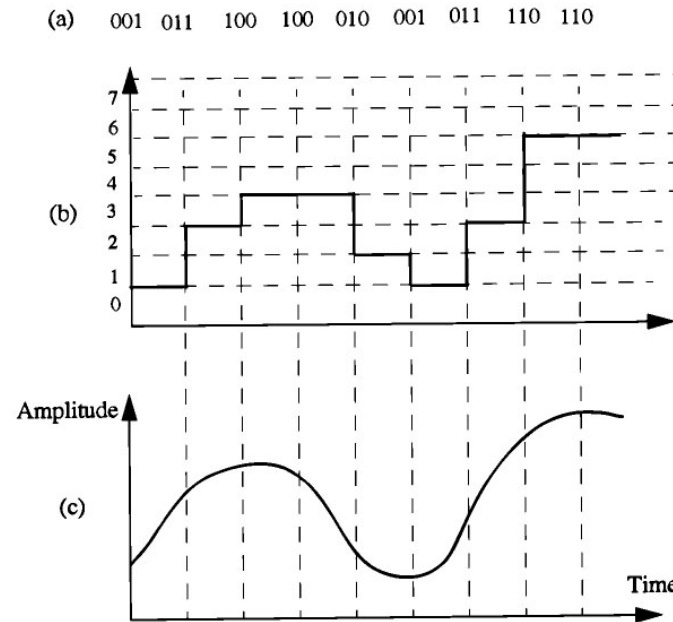
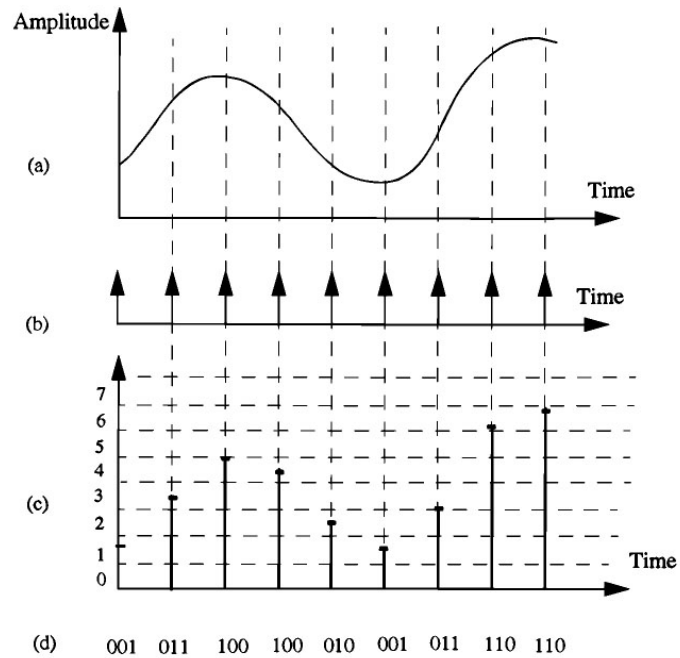
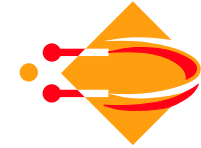
Microphone converts to electrical signals

Electrical to digital conversion by sampling and quantization

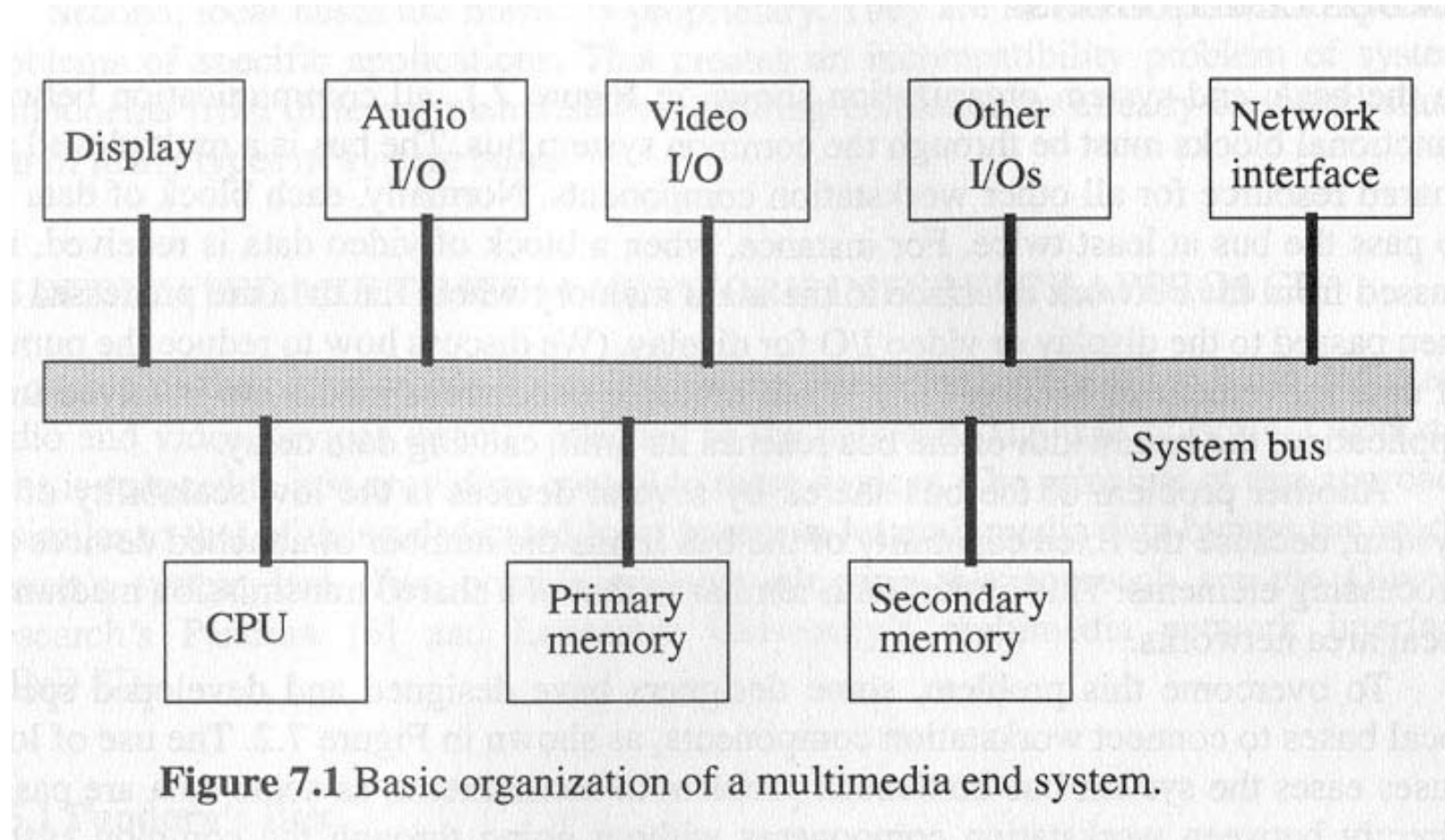
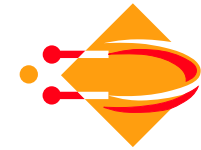
Sampling by clock on readout on A/D converter

Quantization by tree-structure of relative potentials

# Analog $\leftrightarrow$ Digital conversion



# Computer architecture



**Figure 7.1** Basic organization of a multimedia end system.

# Digital processing

---



This is a course on its own!