Module IN3031 / INM378 Digital Signal Processing and Audio Programming

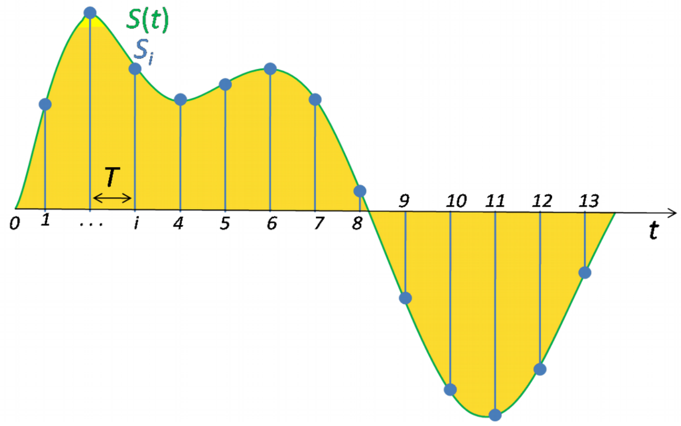
Johan Pauwels j.pauwels@city.ac.uk
based on slides by Tillman Weyde

Who am I ?

- Not Tillman Weyde :-)
  - Lecturer of this module in previous years
  - Now on a sabbatical (focus on research)
- “Visiting Lecturer”
  - This semester (also for Big Data), no regular office hours
  - Otherwise a researcher in automatic music analysis and part-time lecturer at Queen Mary University of London

What This Module Is About

- **Basics:**
  signals, sampling, frequency, spectrum
- **Theory:**
  correlation and convolution, Fourier transform
- **DSP system architectures:**
  streams, channels, filters
- **Data analysis:**
  audio and images, financial data
- **Game programming:**
  audio and music for games



DSP Functions

Typical functions needed:

- Recording: sound, image, video, sensors
- Digital sound, image, and video effects
- Noise reduction, signal enhancement/recovery
- Data compression
- Signal analysis and retrieval:
  sound, music, image, sensor, financial (…)
- Spatial audio: games, VR
- Video and 3D graphics (not part of this module)

Learning Outcomes

Knowledge and understanding:

- Appraise the principles and theories of signal processing.
- Critically evaluate how these principles and theories are used in computer software.
- Apply relevant knowledge in the creation of games and multimedia applications.

Skills

- Design the integration of music and audio in an interactive software.
- Create the music or audio elements of an interactive software.
- Implement DSP functionality in Python
- Implement signal analysis in Python

Course texts

Main texts (links to PDFs on Moodle):

Dorran, David: *Digital Signal Processing Foundations*. DIT 2015

Smith, Steven: *Digital Signal Processing: a practical guide for engineers and scientists*. Newnes, 2003. Available PDFs

Other interesting texts

Lyons, Richard G. Understanding Digital Signal Processing, 3/E. Pearson Education India, 2011. (similar to Smith)

Rocchesso, Davide: *Introduction to Sound Processing*. Florence, 2003, http://profs.sci.univr.it/~rocchess/SP/

Stevens, R. & Raybould, D.: Game Audio Implementation: A Practical Guide Using the Unreal Engine. 2011. (quite specific)

Marks, A. & Novak, J.: Game Development Essentials: Game Audio Development. 2008. (non-technical)

Labs

Tuesday 20:00-20:50, room ELG06-07

Tools:

Mainly:

Python (signal processing and analysis) on Google Colab

→ requires Google account (setup during labs)

FMOD (games engine w/ sound modules)

Office Hours/Contact

For general discussions you can use the super-module on Moodle: https://moodle.city.ac.uk/course/view.php?id=34993

Preferably post your questions on the Moodle message board. For more personal issues, contact me by email.

Week 1:
Signal Basics

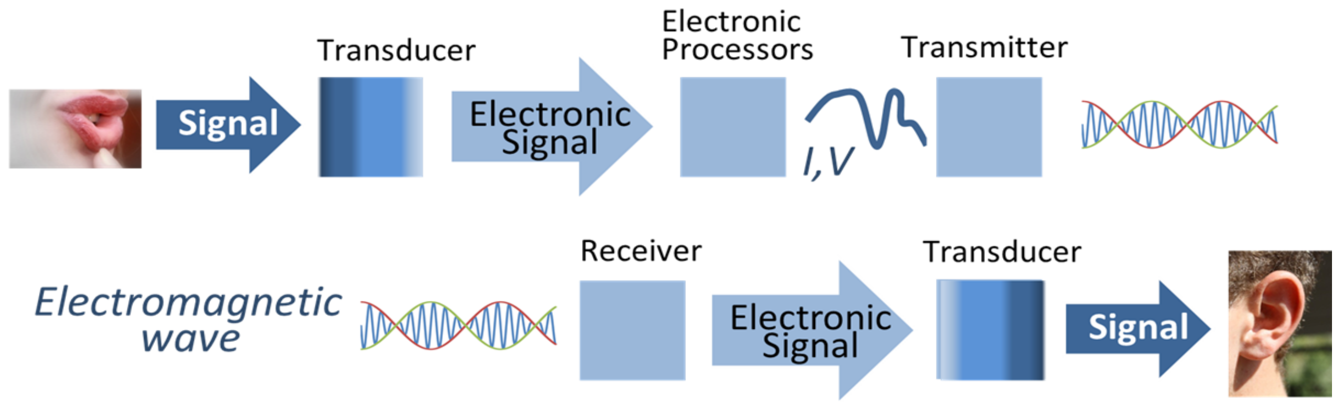
- **What is a Signal?**
  - From latin *signum* (sign): information sent through a medium, from humans or technical, natural or social processes
  - Typically represented as a uniform array or sequence of numbers, possibly higher-dimensional

Signal Processing

- What is Signal Processing?
  - Combines mathematics, physics and technology
  - Transfer, manipulation, analysis, and synthesis of information contained in signals
  - Signals are variable in time and space
    - Sound
    - Images
    - Radio
    - Sensors
    - Financial data
    - Text and symbols

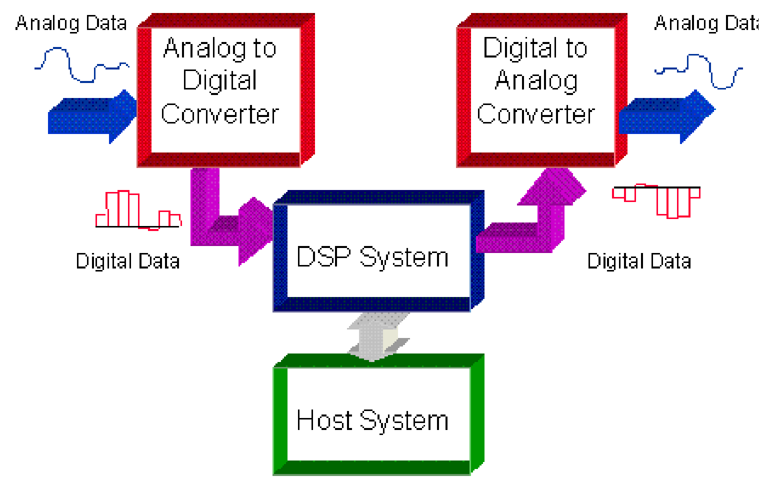
Signal Transfer (Radio)

- Analog (e.g. radio, TV, 1G mobiles)



- Digital

  (DAB, digital TV, 2G+ mobiles, computers, …)



Digital Signal Processing

- **Digital representations** of signals (in bits)
- (Specialised) digital **computers** for processing
- **Used everywhere** in tech, e.g.
  - telephony
  - television & radio
  - games
  - GPS, sensors, ...
- It's **all in your pocket**:

Signals and Waves

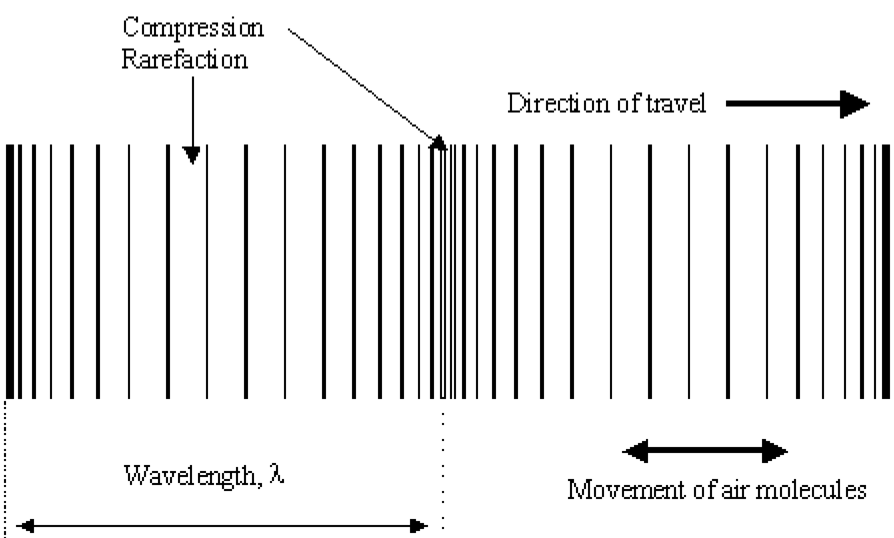
Signals

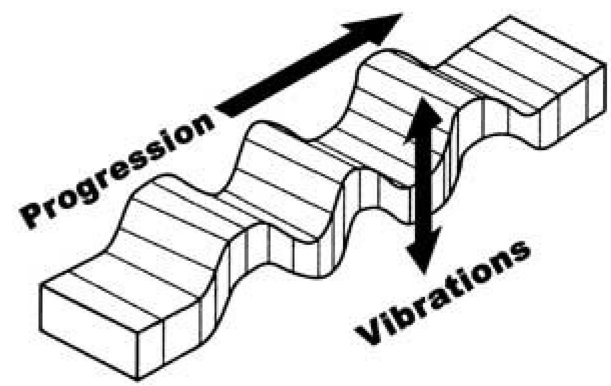
- In technology, our signals are **numeric values recorded over time or space**, e.g.
  - air pressure/movement (sound)
  - brightness (image, video)
  - acceleration, rotation (motion)
  - social or financial data

- Signals are often **recorded oscillations** (**waves**)

Physical Waves

- Movement **travels** through a **medium** (e.g. air) and the medium returns to previous state (**oscillation**).
- Movement direction depends on physical situation (compressibility, environment).
    - **longitudinal**: movement on axis of travel (air)
    - **transversal**: orthogonal movement (e.g. water)

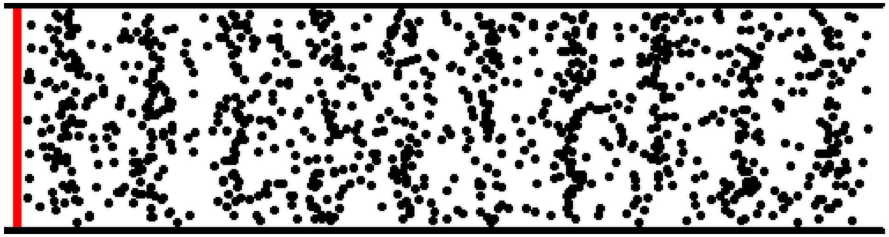




Wave Animation

Animated figure of a longitudinal wave (e.g. sound).

The wave travels, but the particles oscillate.



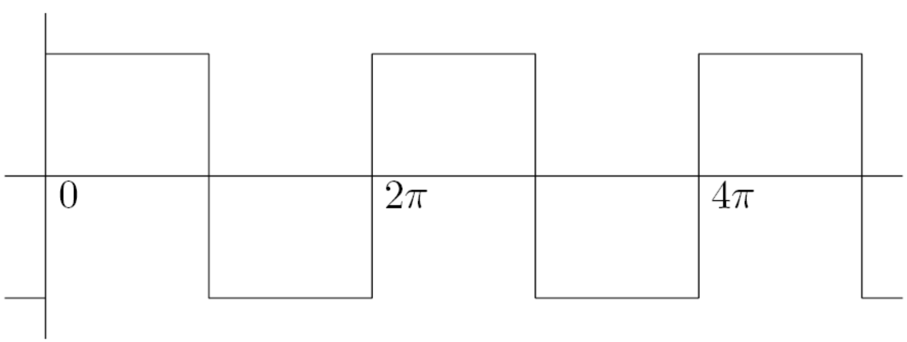
Basic Wave Properties

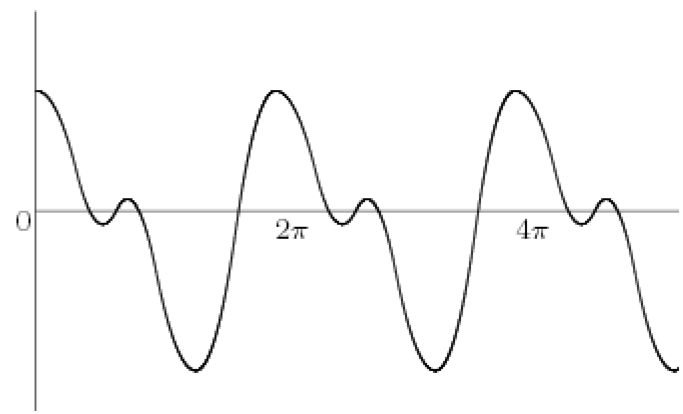
- **Frequency** - speed of oscillations: faster oscillations mean
  - smaller structure in images
  - faster movement or change
  - higher pitch in sound
- **Amplitude** - strength of oscillations: stronger oscillations mean
  - wider movement, greater change
  - louder sound, brighter light

Period and Frequency

period *p*: duration of a periodic signal's cycle

frequency *f*: number of cycles per time $f = 1 / p$

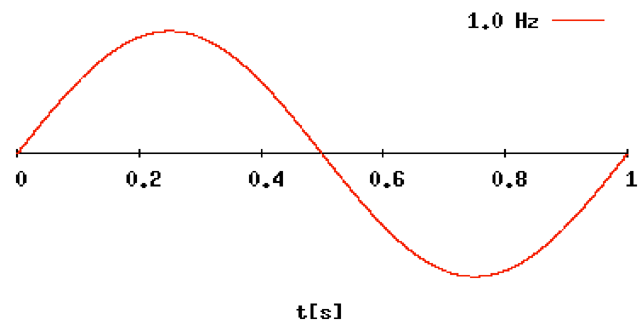




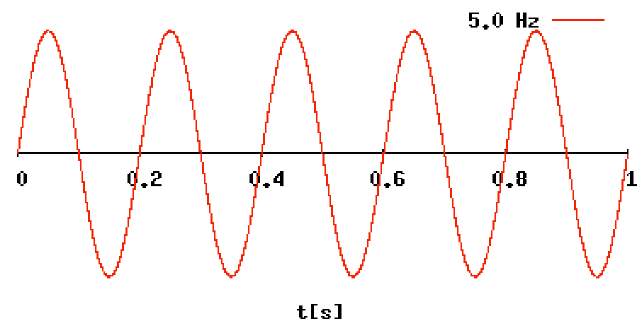
Frequency

Number of **cycles per time**.

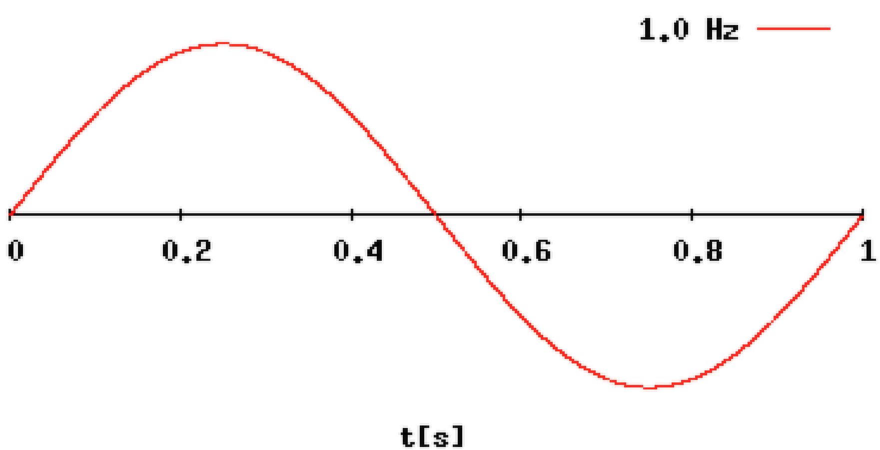
Measured in **Hertz** (Hz, 1/sec).



1 Hz equals 1 wave cycle per second

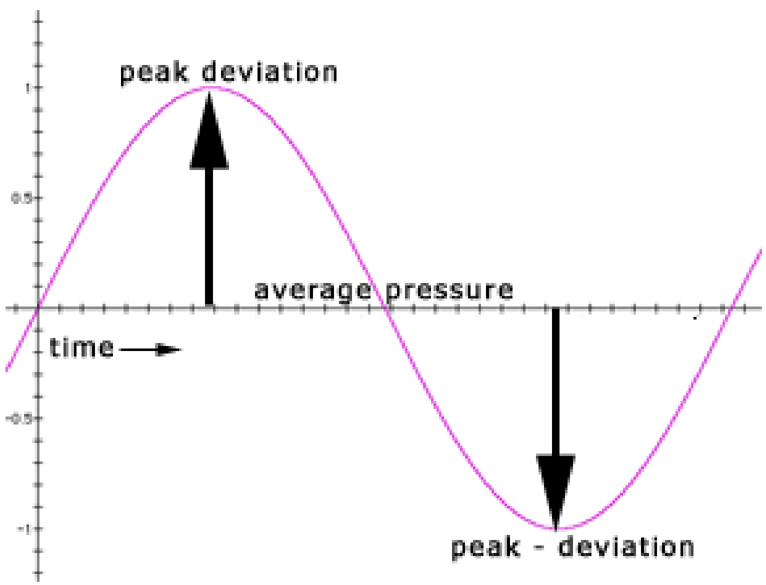


5 Hz equals five wave cycles per second



Amplitude

Amplitude: **scale** of values, often measures at crest and trough peaks, (e.g., for sound **maximal deviations** from normal air pressure)

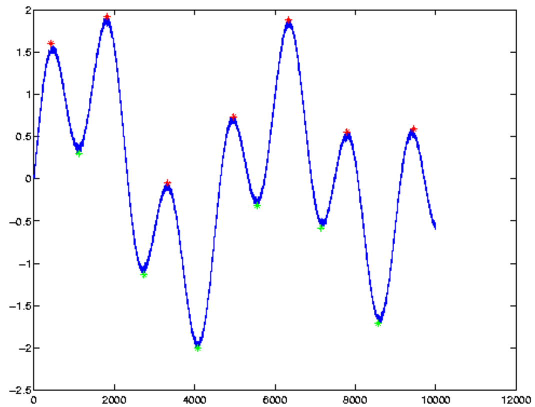


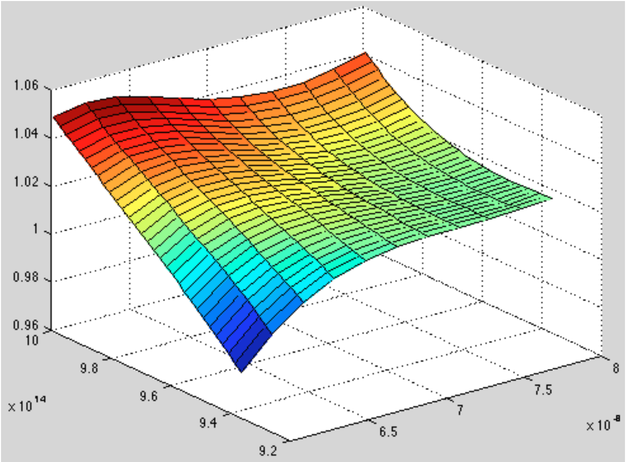
Signals: Mathematical Model

Signals are a **relevant quantity *y*** (air pressure, pixel value),

as a function (typically)
of time: ***y* = *f(t)*** (1-dimensional for audio)
or space: y = f(x,y) (2-dimensional for images)

Graphs are useful, particularly for 1D signals:





Signal Energy and Power

Two definitions:

- **Energy** of a time variant signal: defined as the sum of the squares of the signal values over all time points
  $\mathbf{energy(f) = sum_t(f(t)^2)}$
- **Power**: energy per time
  $\mathbf{power(f) = energy(f)/time = sum_t(f(t)^2)/time = mean(f(t)^2)}$

This matches physics for audio and electrical signals, not for images, values are already energies (of light).

Decibels

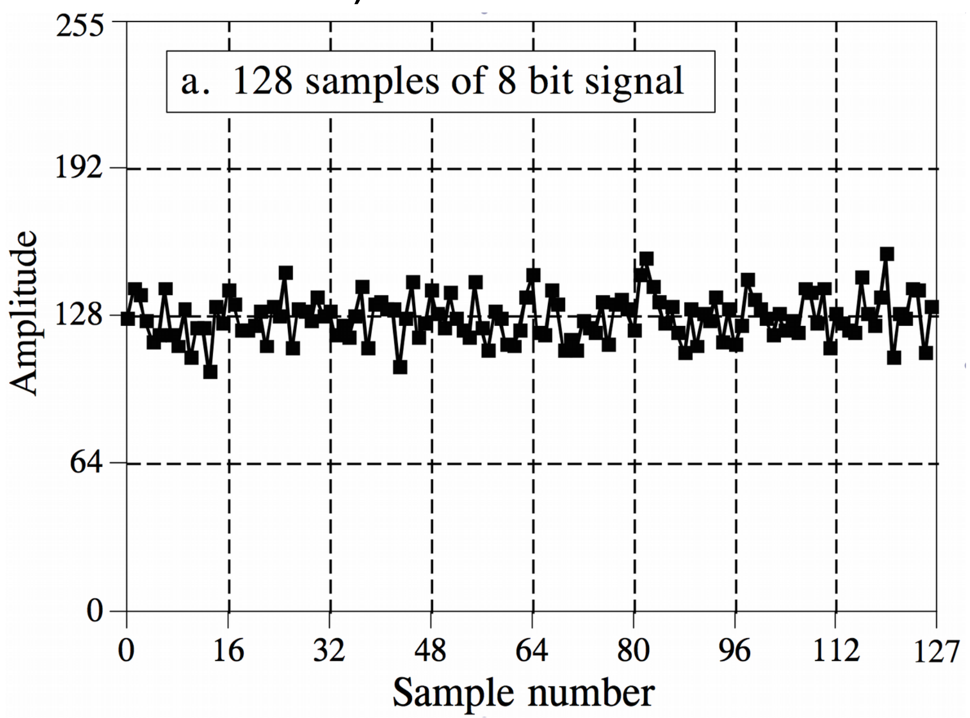
- Signals typically have a **wide range** of values, from very large to very small
- dB is a **logarithmic** expression of **ratios**, especially **useful** for **very large and small** numbers and ratios
- **Definition**:
  a/b = x dB means **$x = 10 \log_{10} (a/b)$**
- In other words:
  adding 10 dB corresponds to **multiplying by factor 10**
- **Examples:**
  +3dB ~ *2 (approximately)
  +20dB = *100 (exactly)
  create more examples ...

Digital Signals: Sampling and Quantisation

Sampling

- Digital signals are sequences of **samples** (values) **at discrete points** in time or space.

  (more details next week)



Sine & Cosine Functions in Signal Processing and Data Analysis

Sine/cosine functions sin(t)/cos(t)

- appear in **basic physical processes**

- in **audio** they are perceived as
'***pure tones***' or '*simple tones*' (no 'overtones')

- can be used to **analyse** and **generate** signals

Sine and cosine are the building blocks of harmonic signal theory.

Sine functions and simple harmonic motion

Simple oscillating system (mass ***m*** and a force growing by factor ***k*** with displacement **x** from *equilibrium point*), e.g. mass & spring, string under tension, electric LC circuit.

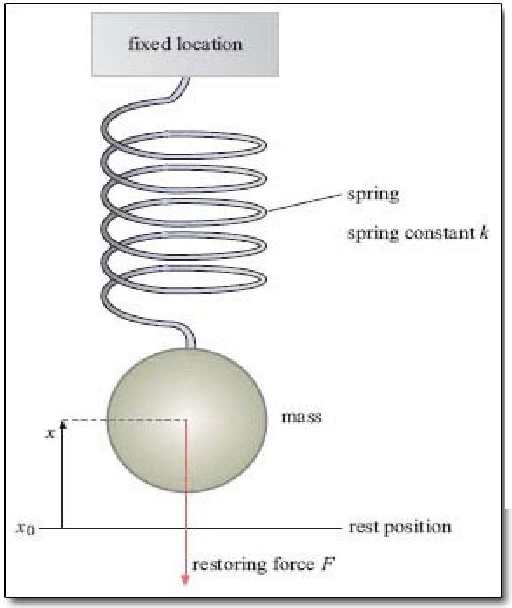
- Equation: $x = c \sin\left(\sqrt{k/m}\, t + \phi\right)$

φ depends on the **start time**

- Frequency: $f = \frac{1}{2\pi}\sqrt{\frac{k}{m}}$

2π is **period of sine**

- System frequency ***f*** depends on ***k*** and ***m***



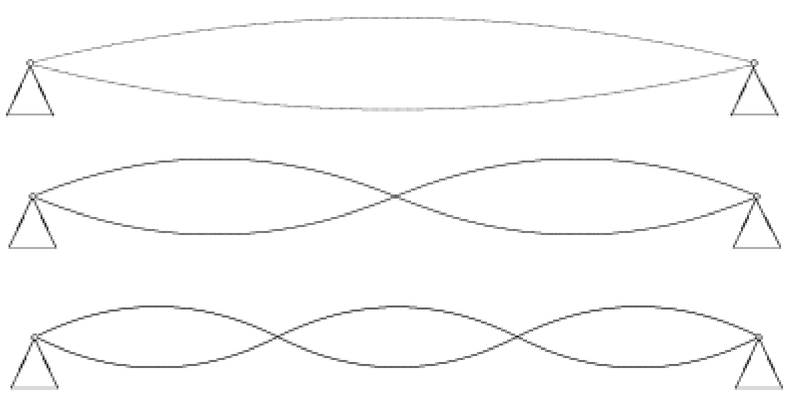
Resonance

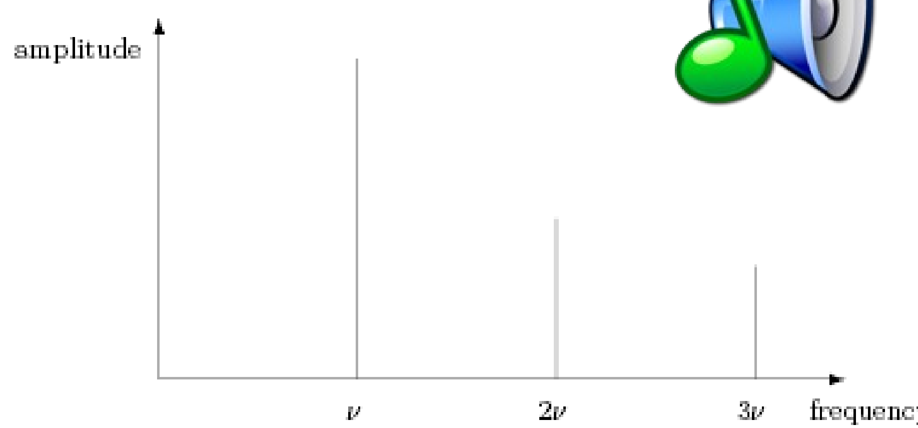
$$f = \frac{1}{2\pi}\sqrt{\frac{k}{m}}$$

- Systems oscillate easily at natural frequency (simple harmonic motion)
- Used, in musical instruments, mechanical watches, etc
- Can be modified by
  - changing ***m**, e.g.*
    - *air volume* (wind instruments, e.g. ***trombone***)
    - different string length and width (piano, **guitar**, violin)
  - changing ***k***
    - electrical capacitor (synthesizer)
    - string tension (**guitar**, …)

Complex Signals

- **real systems** oscillate at **more than one frequency**
- several frequencies are added with different intensities
  these are called **partials** (or **overtones** or harmonics)
- $s(t) = a_0 f_0(t) + a_1 f_1(t) + ... + a_n f_n(t)$





Harmonic and Inharmonic Signals

- **Harmonic** signals have **integer ratios** between **fundamental f_0** and the other **partials**
- Most **musical sounds** are (approximately) **harmonic**
- **Bells** have **typically inharmonic** sounds

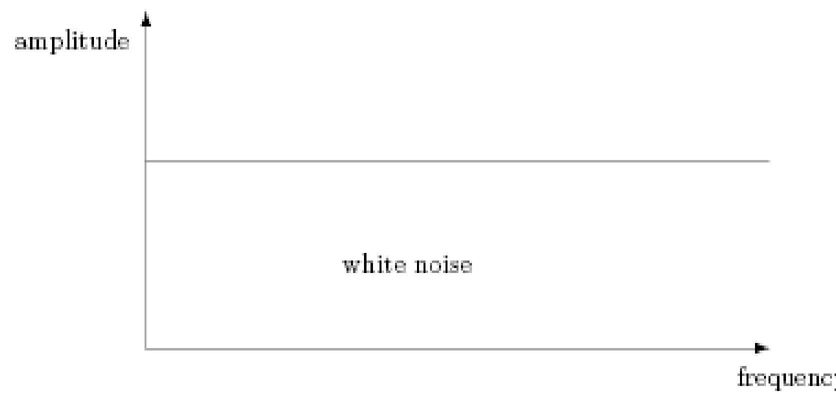
Noise

- **Tones** contain **energy** at **discrete frequency points**
- **Noise** contains **energy** at **all frequencies**

(e.g. analog radio not tuned to a station)

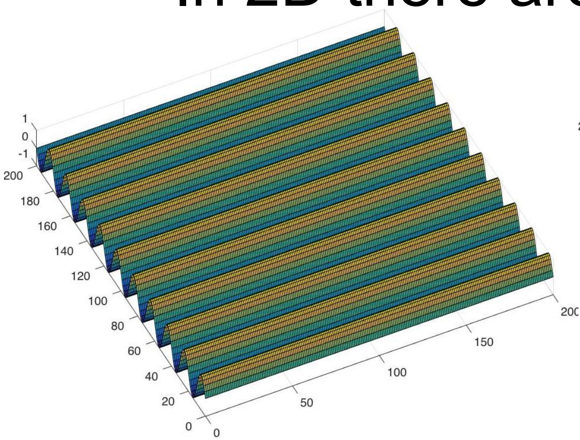
tone

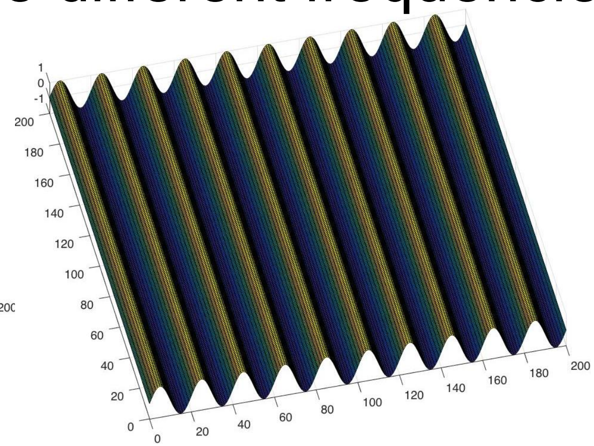
noise

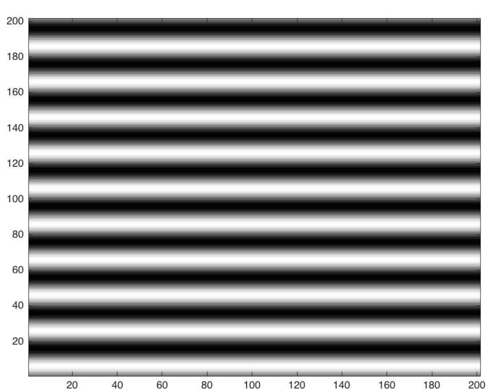


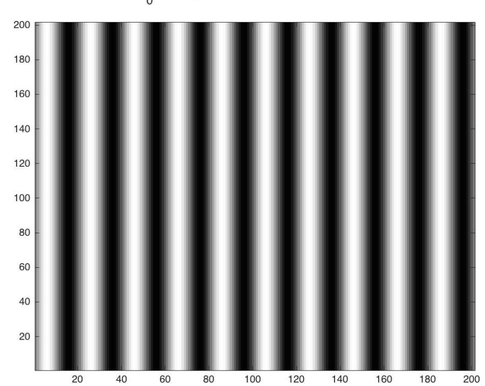
Sine Waves in 2D

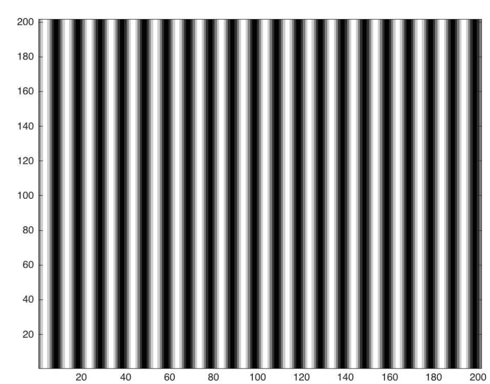
- In 2D there are different frequencies in **both dimensions**

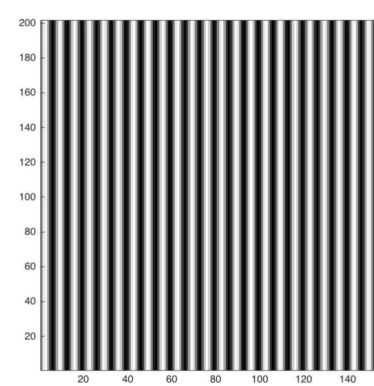






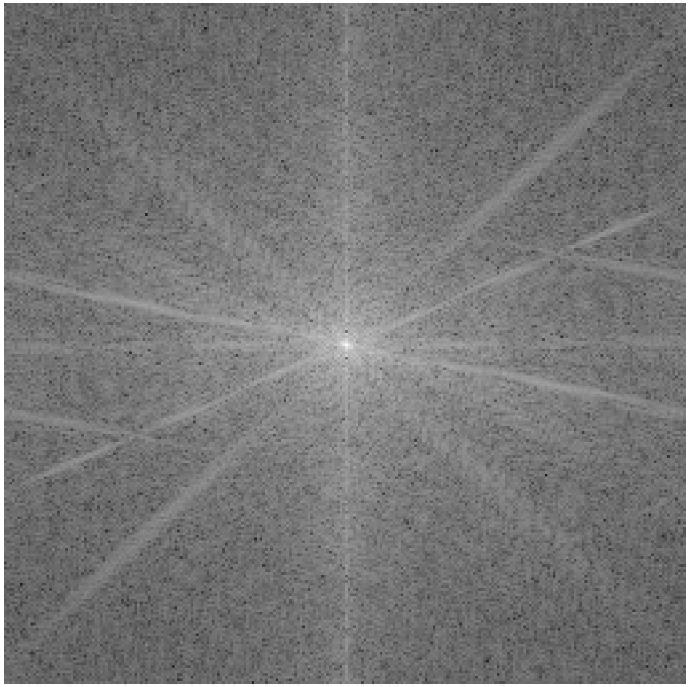






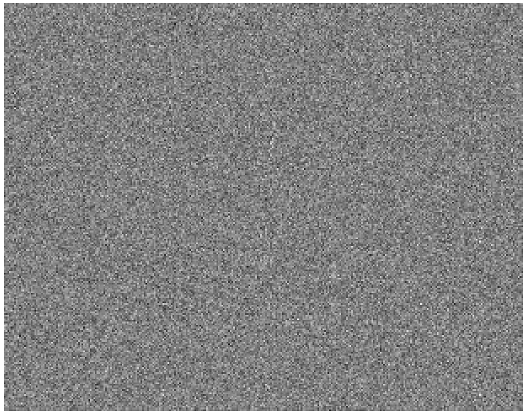
Sine Waves in 2-D

- We can relate whole **images** to **mixtures of sine waves**, but it's not as straightforward (more in later weeks)



Noise in 2D

- In **2D** there is also **noise**


noise


image

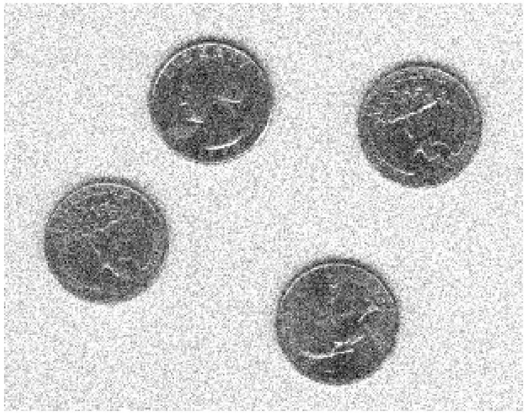


image with noise

- **Photos** taken in **low light** often contain noise

Frequencies in Audio and Music

Audio Frequencies Perceived by Humans

- Range approximately 20Hz – 20,000 Hz
- Frequencies **perceived logarithmically** (Weber's law)
  1 **octave** up corresponds to 2 x frequency
- **Sequential discrimination** accuracy **up to 3Hz**
  (i.e. tones with that frequency difference are perceived as being different when heard one after the other)

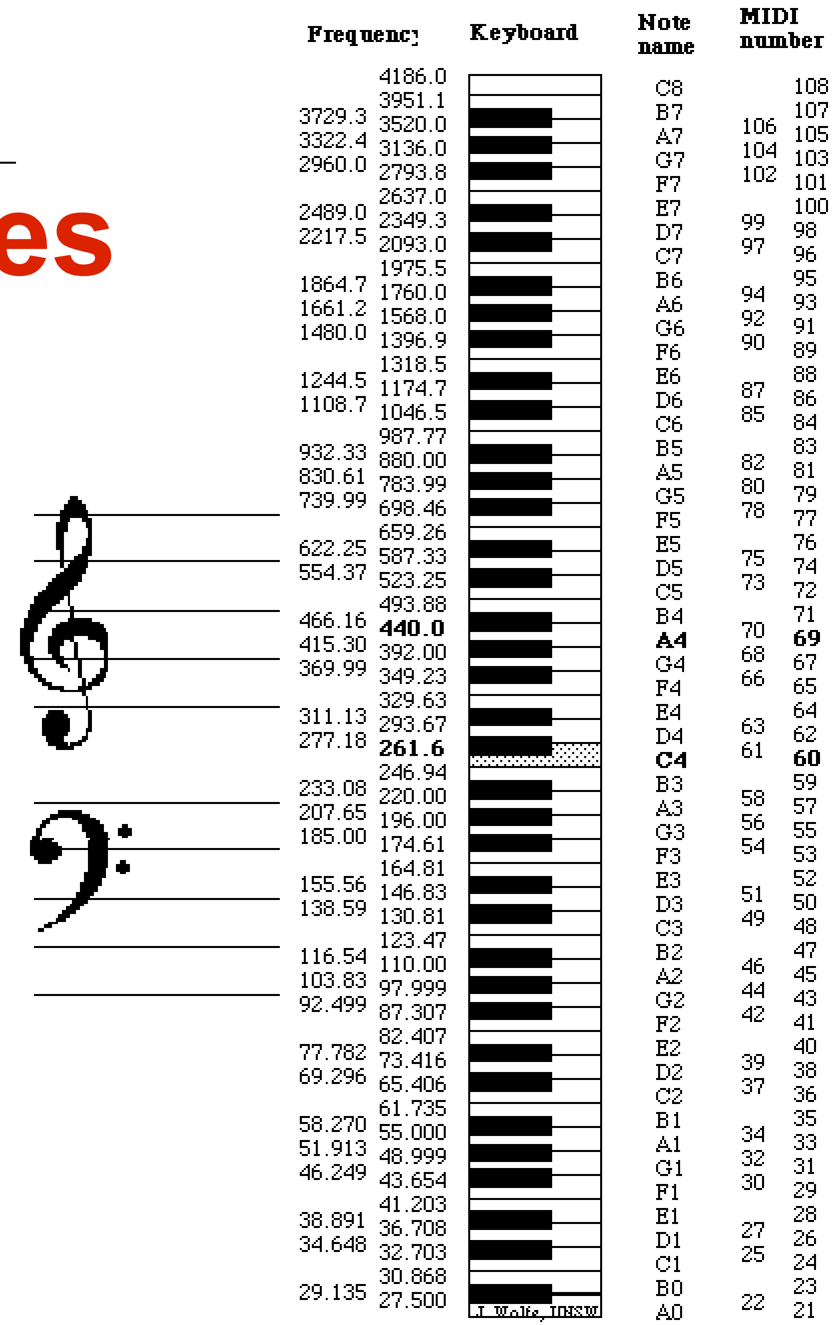
Frequencies in Music

- In music frequencies are organised as **pitches**, which correspond to one fundamental frequency each.
- In all cultures a frequency **ratio** of **2:1** (an **octave**) has a special role, these tones are perceived to be highly related
- Western music:
  - **octave** divided into 12 **semitones**
  - a **semitone** has a **ratio** of **12th root of 2** (in equal temperament, there are other variants)
  - reference note is the '**middle A**' at 440Hz

CITY UNIVERSITY OF LONDON EST 1894

Frequencies in MIDI

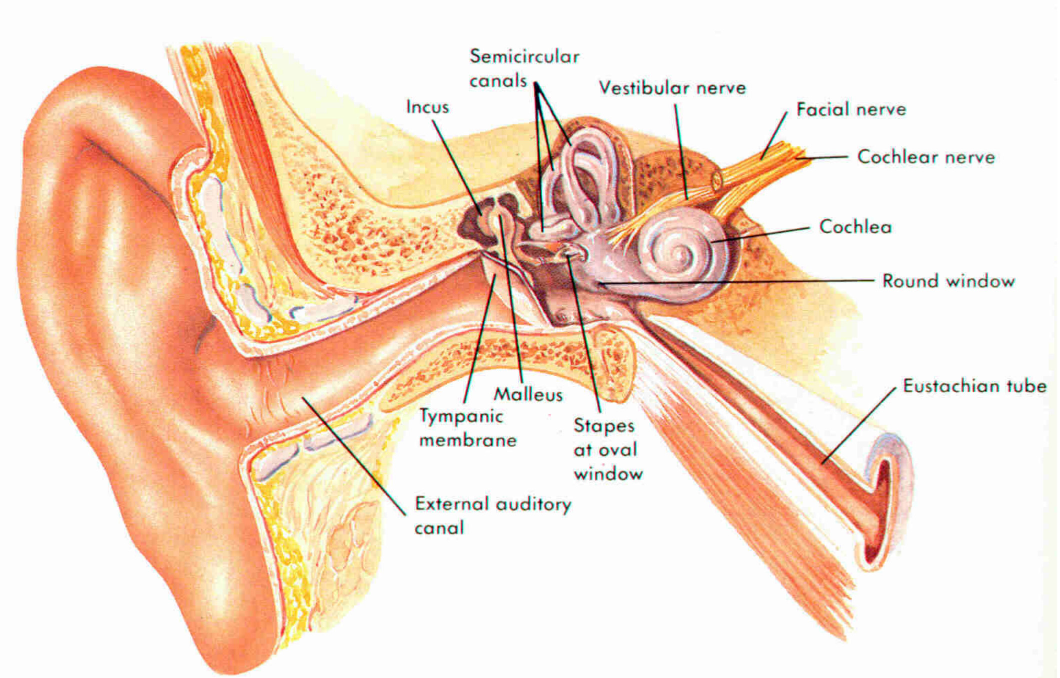
- In MIDI (Musical Instrument Digital Interface) all notes have a number.
- 'middle A' has number 69,
- Freq of MIDI number X calculated as 440 * 2^([x-69]/12)



The Human Ear

outer ear (ear flap and canal)

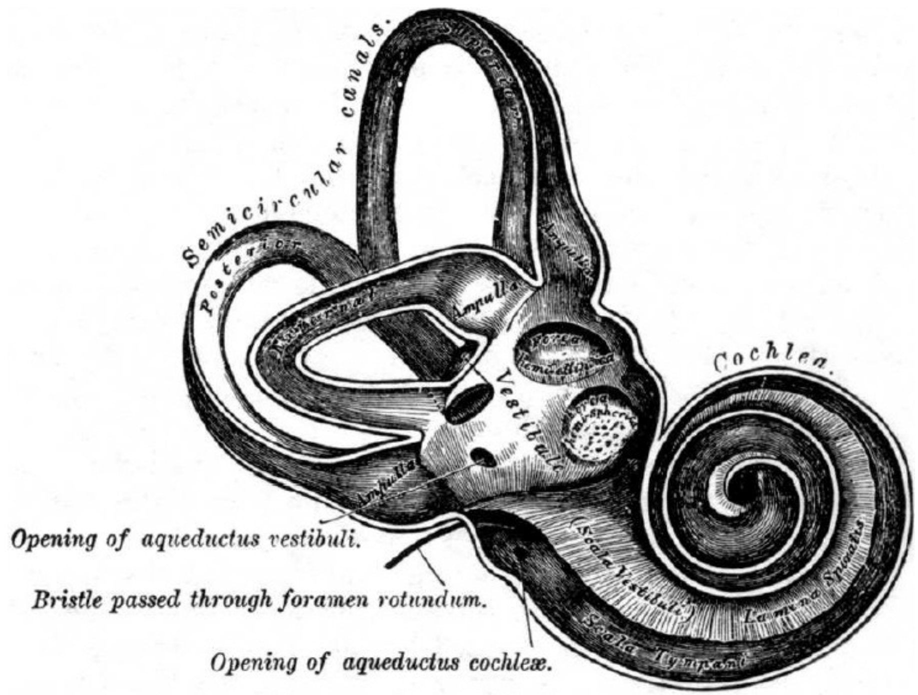
middle ear: eardrum (Tympanic membrane), hammer (Malleus), anvil (Incus), and stirrup (Stapes) transmit vibrations to the inner ear



HEARING
GROSS ANATOMY OF THE EAR—FRONTAL SECTION

The Inner Ear

- the **vestibule** (middle)
- the semicircular canals (back, sense of balance)
- the **cochlea** (front, connected to the auditory nerve)



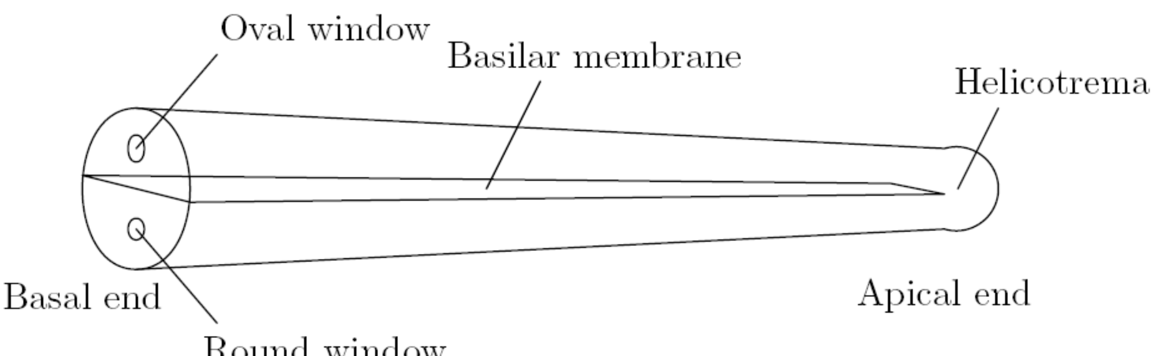
The osseous labyrinth laid open. (Enlarged.)

The Cochlea

Unrolled **length ~3cm**

Vibrations **enter oval window** transmitted by the stapes

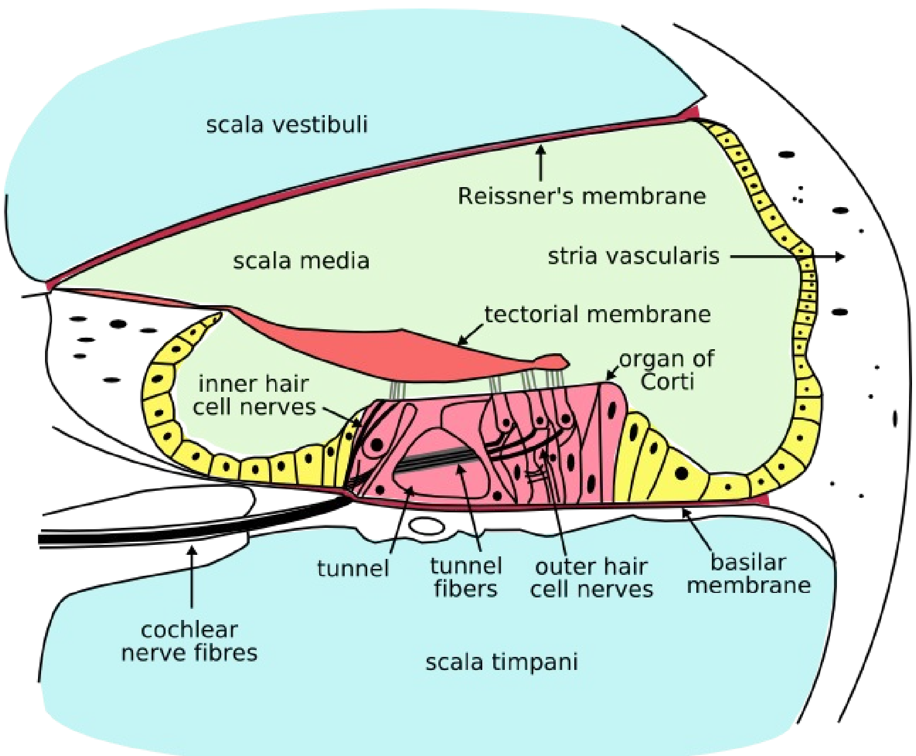
Wave **transmission** on basilar membrane **varies by freq**



The cochlea, uncoiled

Basilar Membrane

Hair cells on basilar membrane transform (mechanical) vibrations into (electro-chemical) nerve signals.

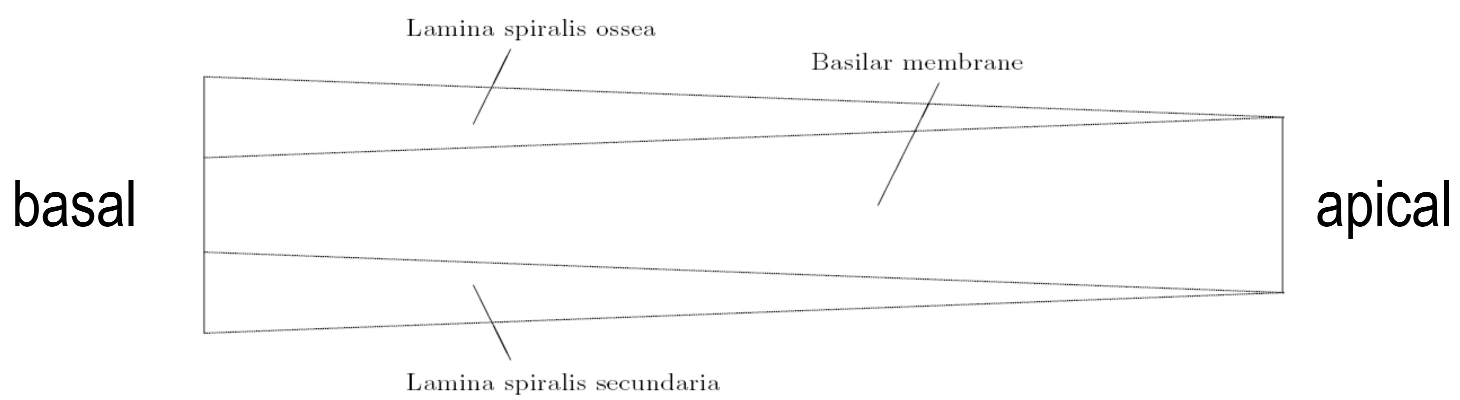


Frequency Analysis in the Cochlea

Basilar membrane widens from basal (input) to apical end

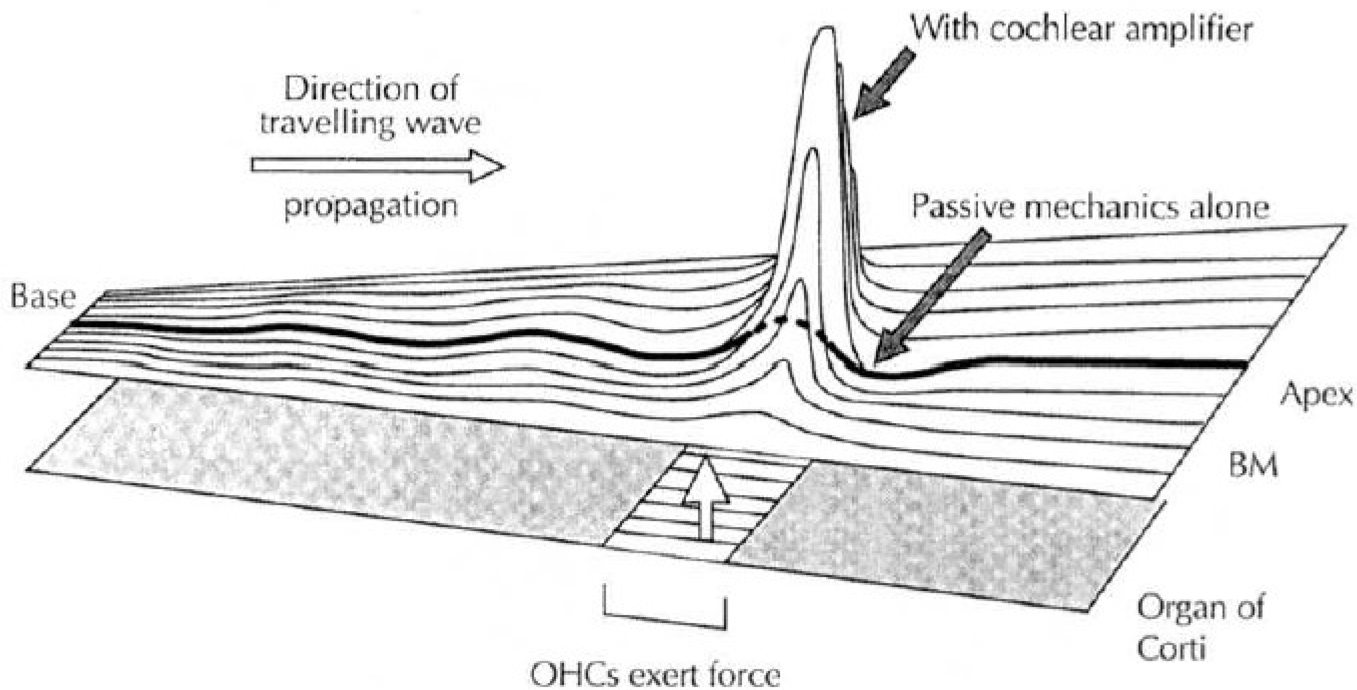
Resonance for higher frequencies at lower (basal) positions

Different hair cells 'tuned' to different frequencies



Frequency analysis in the ear

Active sharpening of frequency perception by top-down mechanisms (**cochlear amplifier**).



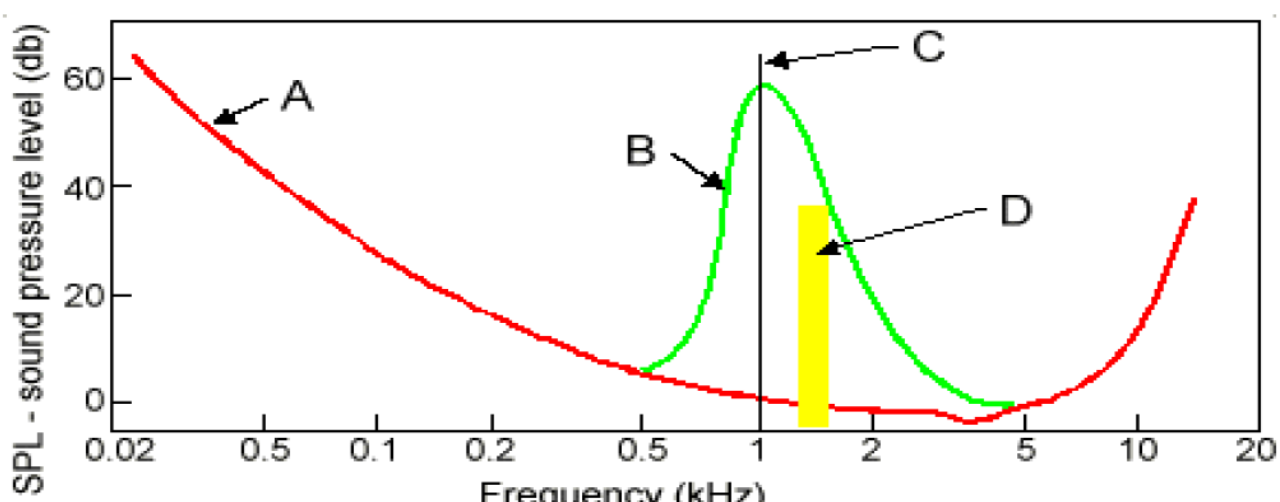
Masking

Sounds close in frequency and time mask weaker sounds.

Used in lossy compression (MP3, WMA, OggVorbis)

A: normal audible threshold; B: threshold changed by tone C

D: Masked tone



READING

Physics of waves:

http://www.physicsclassroom.com/Class/sound/soundtoc.html

Lesson 1 to 5 with tests.

Doran, basics of DSP:

https://arrow.dit.ie/cgi/viewcontent.cgi?article=1013&context=engschelecon

Read pages 2-9 and do the following quiz:

http://eleceng.dit.ie/dorran/moodle/mod/quiz/view.php?id=146

Next week:
Sampling
and Reconstruction
Signal Correlation