Perception of Sound

Augusto Sarti

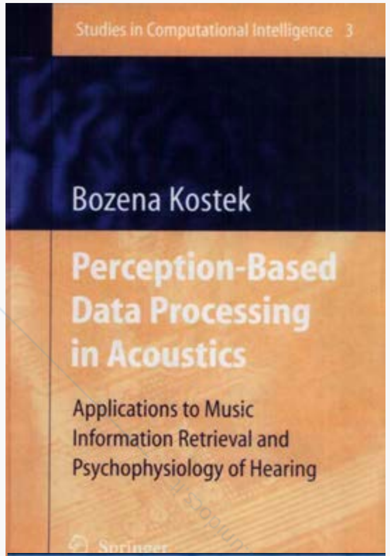
Computer Music Representations and Models

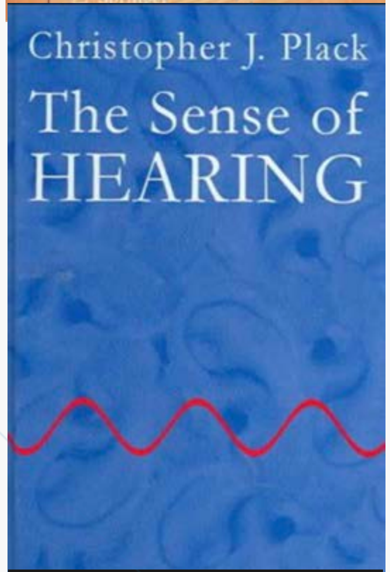
M.Sci. Music and Acoustic Engineering

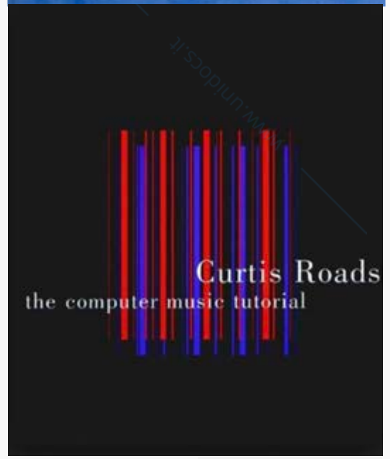
POLITECNICO MILANO 1863

Sources

- Bozena Kostek, “Perception-Based Data Processing in Acoustics - Applications to Music Information Retrieval and Psychophysiology of Hearing”, Springer-Verlag, 2005
- Carlo Drioli, Nicola Orio, “Elementi di acustica e psicoacustica” (course notes, in Italian)
- Christopher J. Plack, “The Sense of Hearing”. Routledge, 2005
- Jack Katz, Robert F. Burkard, Larry Medwetsky, “Handbook of Clinical Audiology”. Lippincott Williams & Wilkins, 2002
- Curtis Roads, “The Computer Music Tutorial”. MIT Press, 2007







POLITECNICO MILANO 1863

(*)

Acoustics: a crash overview

POLITECNICO
MILANO 1863

M.S. Music and Acoustic Engineering

www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari

Acoustics (from the greek: *akouein* = to listen)

(*)

- Study of *sound* and its behavior in the various transmitting media and environments
  - It includes effects of *absorption*, *diffraction*, *interference*, reflection and *rifraction*
- In a wider sense, acoustics is the physics of sound, from all standpoints
  - Until the beginning of the 20th century, the terms *sound* and *acoustics* referred to audible frequencies
  - Today acoustics also includes waves that are not at audible frequencies or whose levels are below or beyond human audibility limits
  - The term sonic was recently used to refer to acoustic processes that have nothing to do with auditory processes
- *Architectural acoustics* refers sound propagation in closed spaces
- Applied acoustics is also defined as *acoustic engineering*

POLITECNICO MILANO 1863

M.S. Music and Acoustic Engineering

Sound

(*)

- Any vibration of the air or any other means that is able to produce an auditory impression
  - Typically caused by the vibration of some body in the transmitting medium
- Parameters: frequency, amplitude (in relation to sound pressure and intensity (dB)), envelope, spectrum (timbre) and duration
- Characteristics: period, wavelength, (and propagation speed), speed of particles, superposition of effects, absorption, interference, diffraction, reflection and rifraction (echo and reverberation), stationary waves (frequency response), sound propagation
- Study areas
  - Acoustics: behavior of waves in elastic media, irrespective of the auditory process
  - Psychoacoustics: auditory stimula, including the intellectual process of interpreting such stimula
  - Sonology: mind process that extracts sound characteristics to improve storage, intelligibility, etc.

POLITECNICO MILANO 1863

Some pertinent definitions

(*)

- Audio signal = function
  - Domain → time (continuous)
  - Co-domain → sound wave magnitude (air pressure)
- $P_{eff}=P_0/\sqrt{2}$ (Root Mean Square of pressure variations), minimum perceived (audibility threshold) = 0.00002 Pa (N/m^2), pain threshold = 20 Pa
- Acoustic intensity: $I= (P_{eff})^2/\rho c$ (ρ=density of the transmitting medium, c=speed of sound in the medium)
- $PL=20\log(P/P_{ref})$, Pressure Level (medium's deformation)
- $IL=10\log(I/I_{ref})$, Intensity Level (energy flow)
- $SPL=20\log(P/0.00002)$, Sound Pressure Level

POLITECNICO MILANO 1863

M.S. Music and Acoustic Engineering

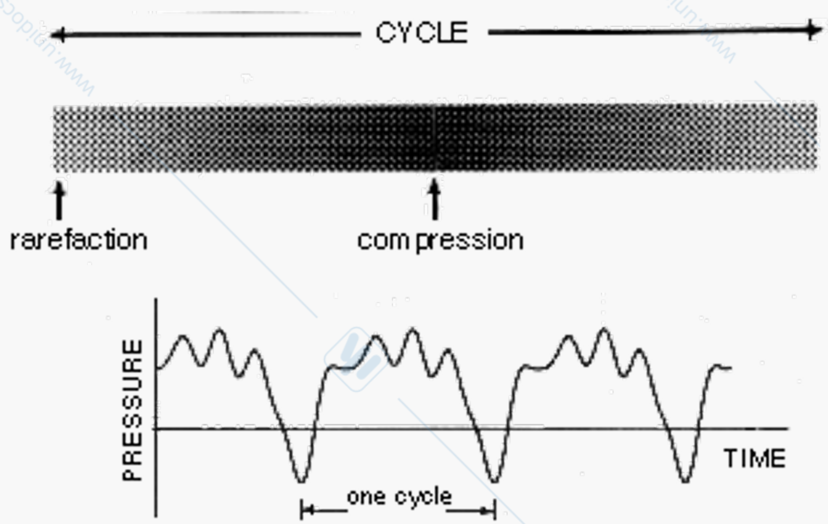
www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari

www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari

(*)

Acoustic waves

- Cyclical change of air pressure
- Acoustic waves can be
  - Longitudinal (solids, liquids, and gas)
  - Transversal (solids)
- The duration of an oscillation is called **period** (*T*) and the number of oscillations per second is called **frequency** *(f)*
  - $T = 1/f$
    - e.g., if f = 440 Hz, then T = 1/440 s = 2.27 ms

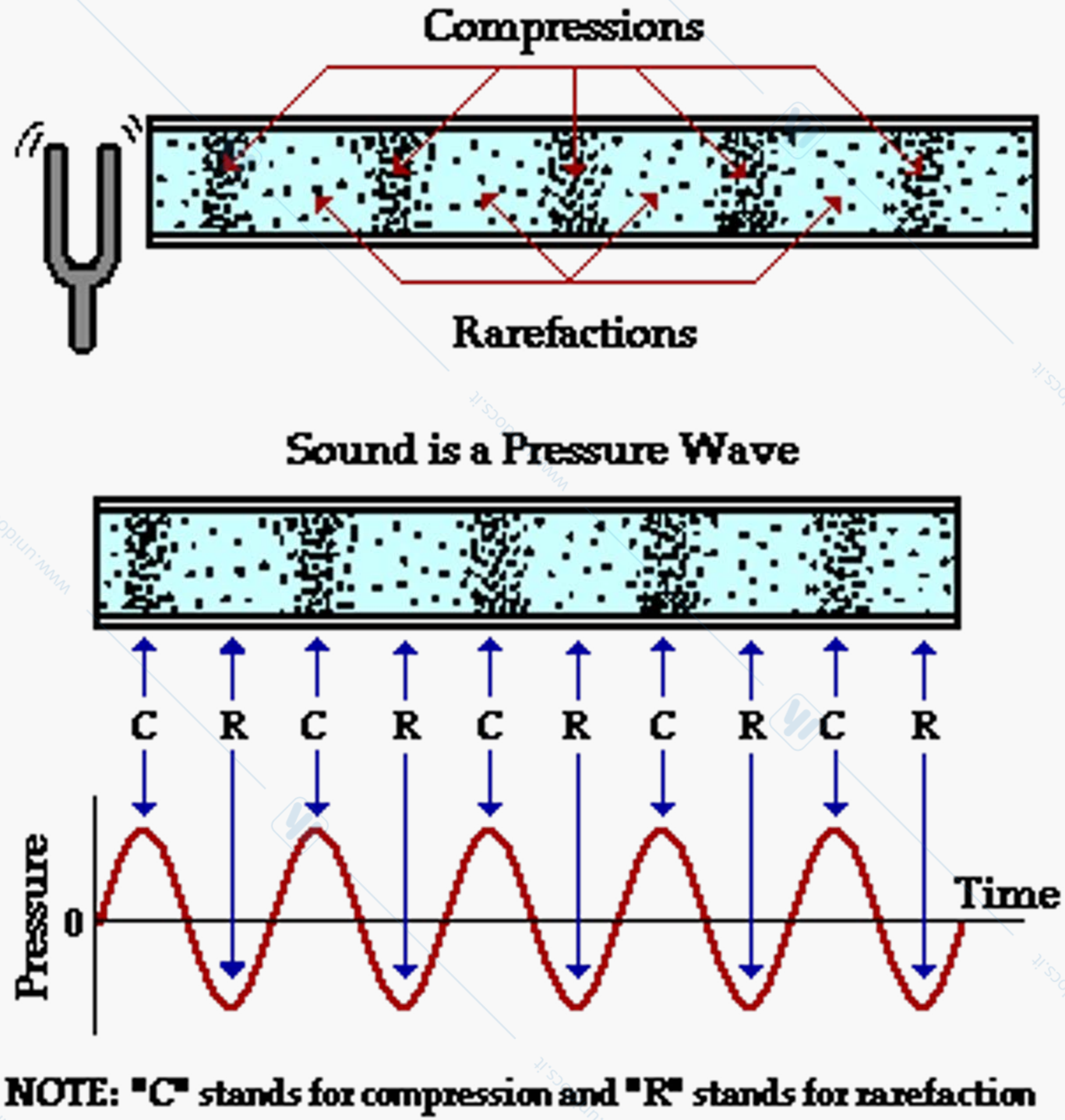


POLITECNICO
MILANO 1863

M.S. Music and Acoustic Engineering

Waves

(*)

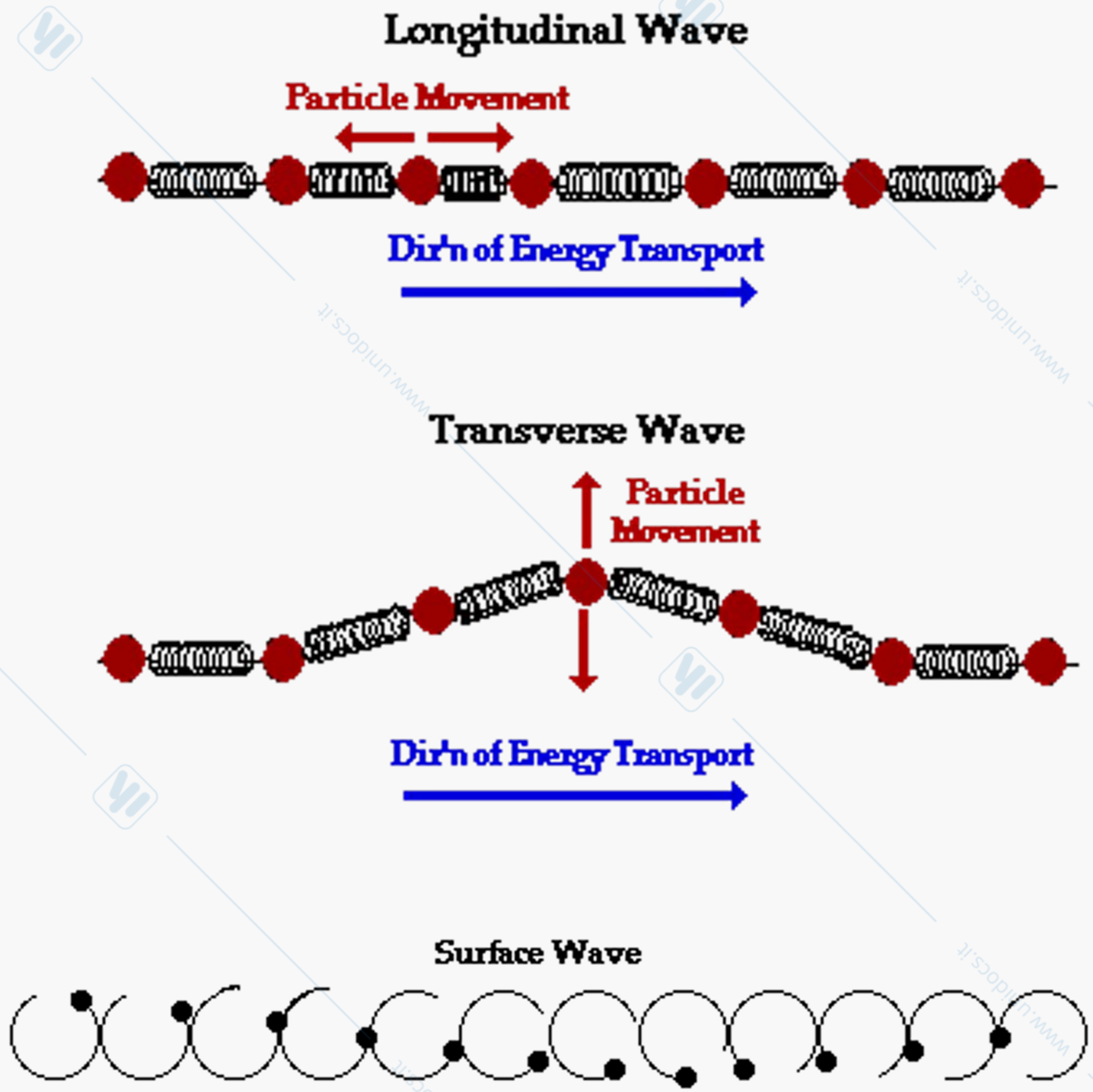


POLITECNICO MILANO 1863

Types of waves

(*)

- Longitudinal
- Transversal
- Surface waves



A surface wave is sometimes referred to as a circular wave since particles of the medium undergo a motion in a complete circle.

POLITECNICO MILANO 1863

M.S. Music and Acoustic Engineering

Interference Patterns

(*)

- Interference patterns
  - Anti-nodes: points where either thick lines meet or thin lines meet
  - Nodes: points where a thick line meets a thin line
  - Nodes group along nodal lines.
  - Anti-nodes do the same along anti-nodal lines.
- Spacing between such lines depends on wavelength.
  - Nodal and anti-nodal lines get farther apart as the wavelength increases

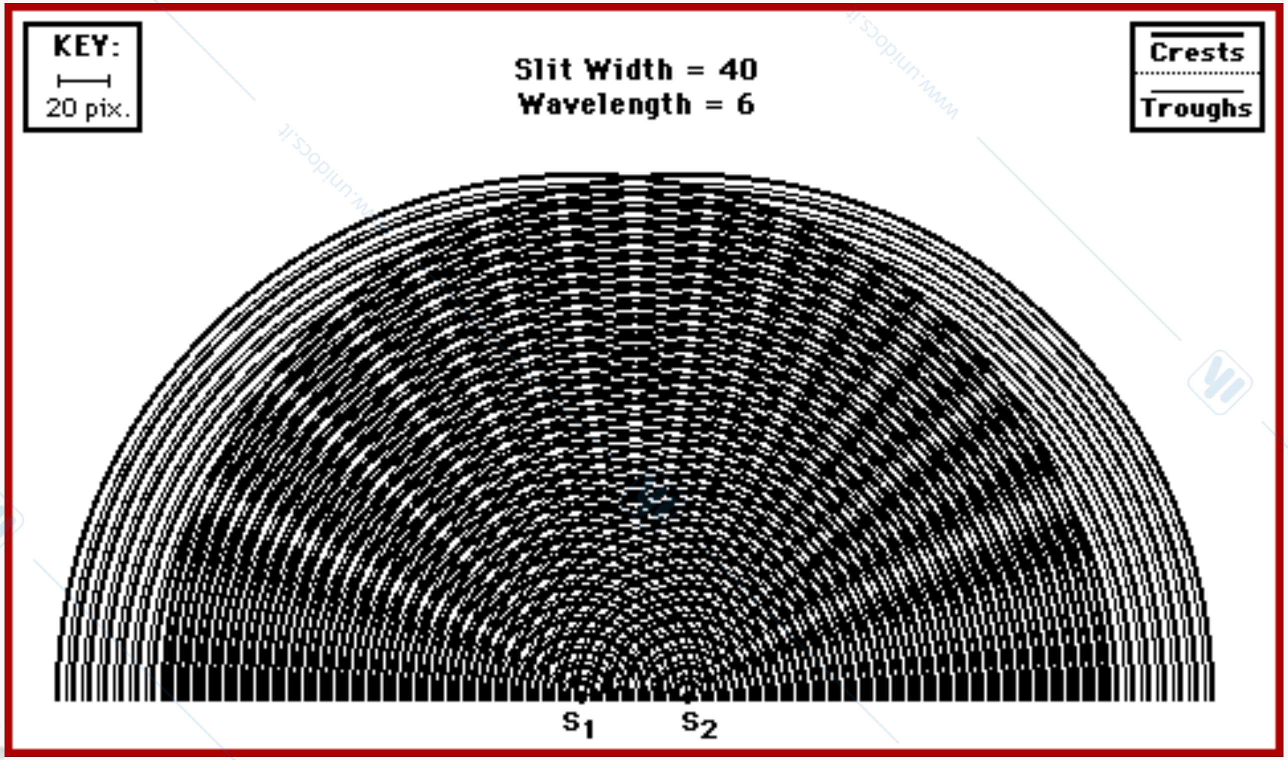
POLITECNICO MILANO 1863

M.S. Music and Acoustic Engineering

(*)

Interference Patterns

- Changing frequency



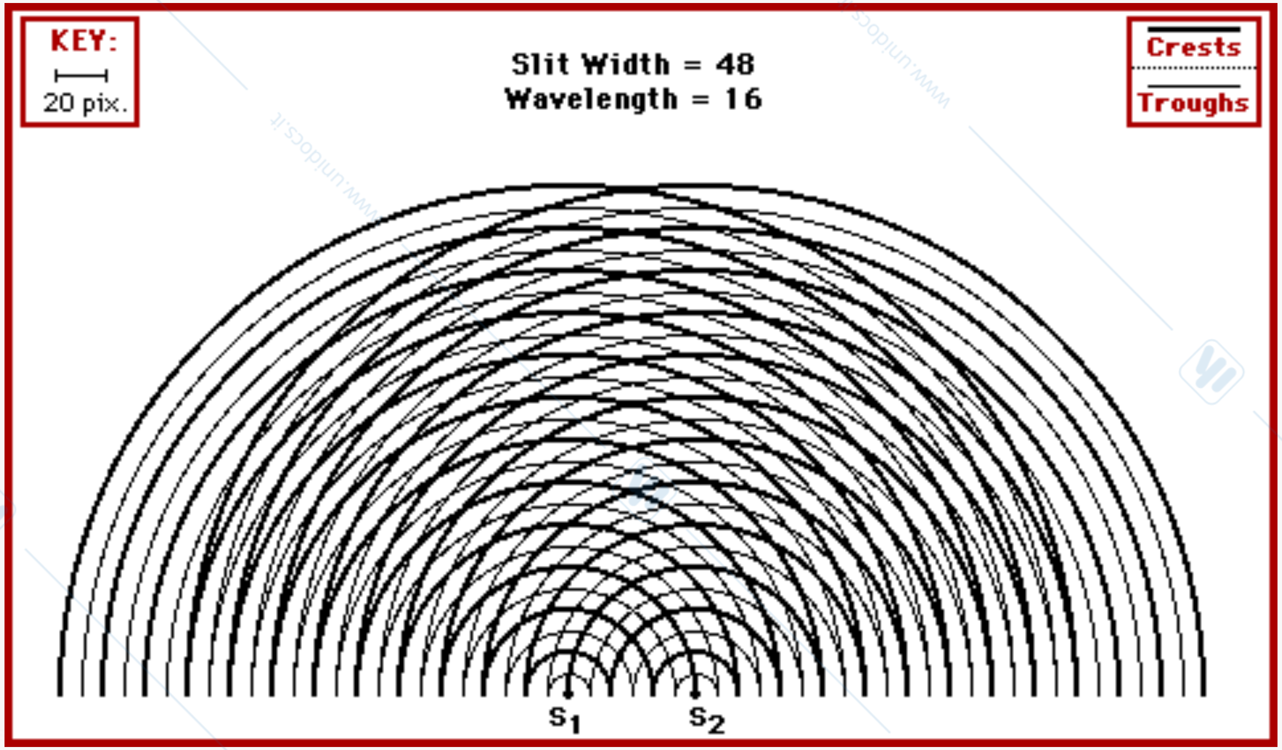
POLITECNICO MILANO 1863

M.S. Music and Acoustic Engineering

Interference Patterns

(*)

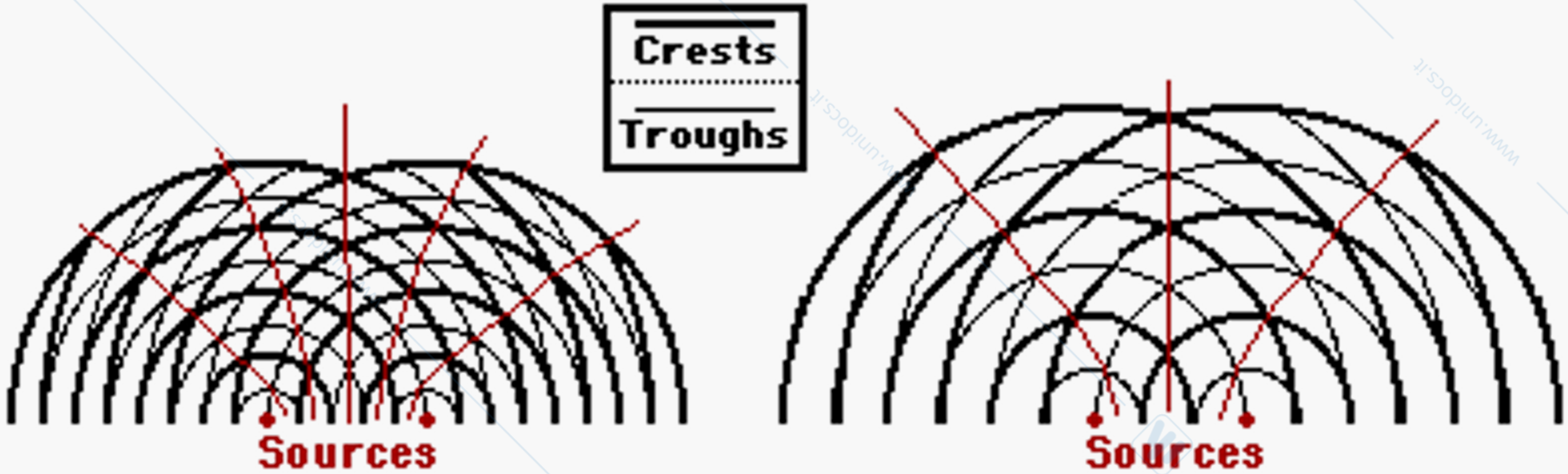
- Changing separation distance



POLITECNICO
MILANO 1863

Two-Point Source Interference

(*)



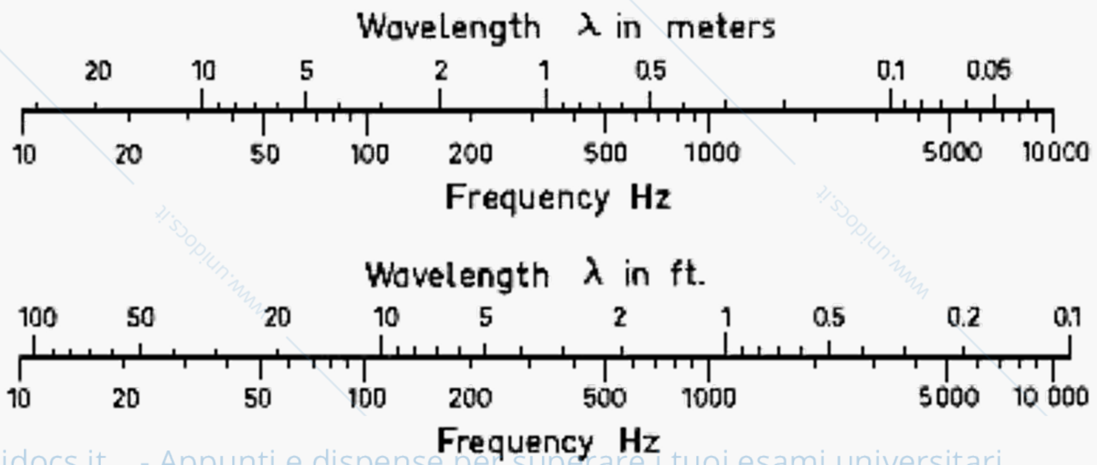
The proximity of the anti-nodal lines in a two-point source interference pattern is dependent upon the wavelength of the waves.

POLITECNICO MILANO 1863

Wavelength

- Distance btw a point on the wave and the corresponding point in the next period (cycle)
- Can be thought of as the distance covered by the sound in one cycle
- Denoted with λ.
- The **speed of sound** in a medium is
  $v = f \lambda \Rightarrow \lambda = v / f$
- Speed of sound in air: 344 m/s
  – As v is constant, we can derive λ from f and vice-versa

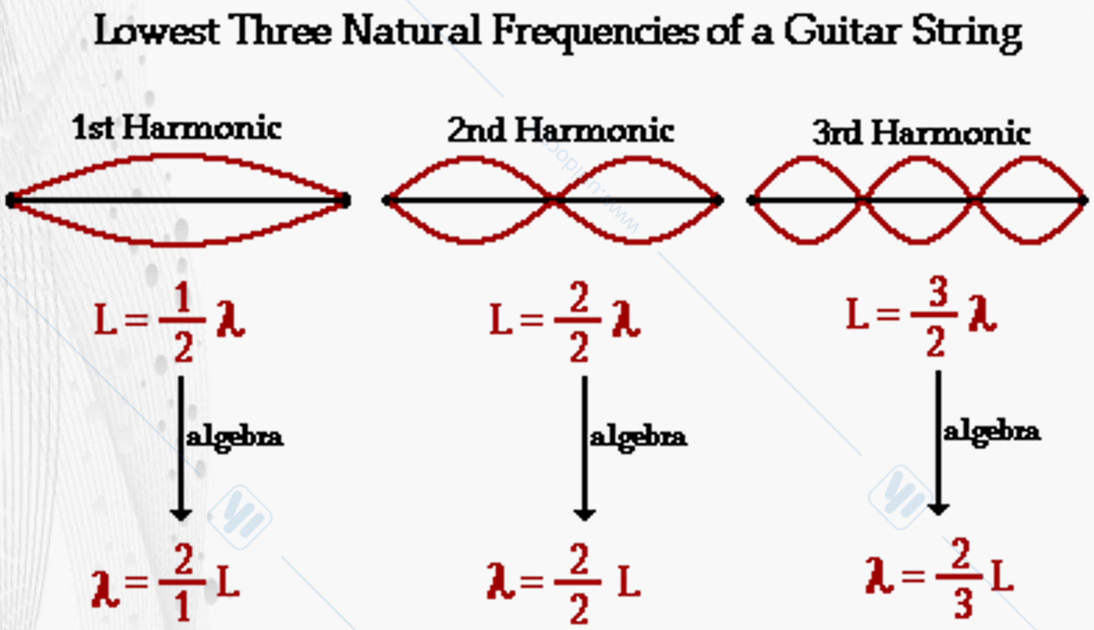
Substance	Temperature (°C)	Speed (m/sec)
CO_2	0	258
CO_2	35	274
Air	0	331.5
Air	20	344
Water Vapor	35	402
Helium	20	927
Hydrogen	0	1,270
Water	15	1,437
Steel	-	5,000

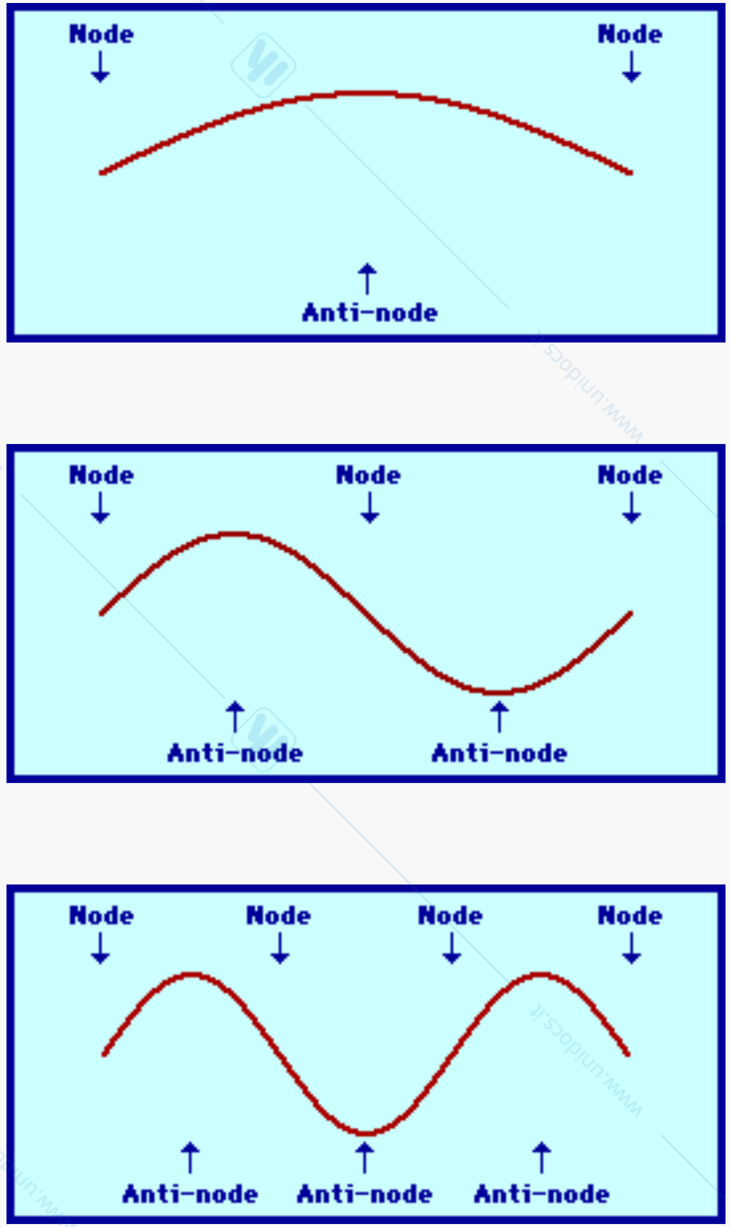


Strings

(*)

- Standing Waves on a String





POLITECNICO
MILANO 1863

(*)

Velocity of particles

- In the presence of a sound wave the molecules of some means alter their random motion and oscillate at the frequency associated to the wave, moving at a speed that depends on the sound pressure p according to

$$p = \rho \; v \; u$$

Where ρ = density of the medium,
v = speed of sound,
u = speed of particles

- The magnitude d of the oscillations depends on u according to

$$d = u / 2\pi f$$

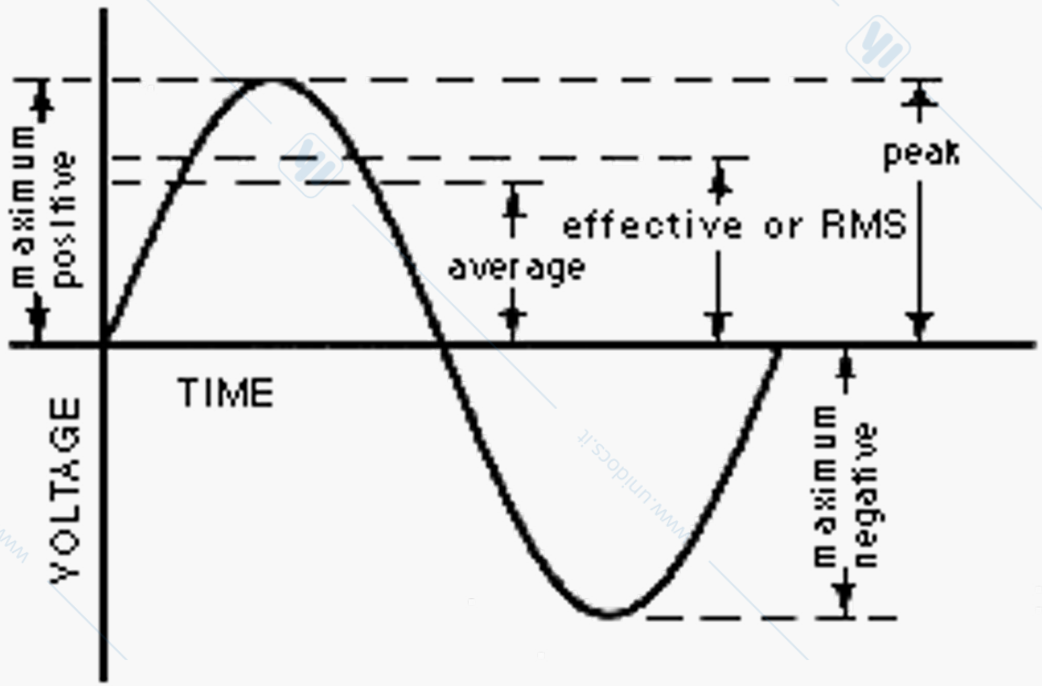
  – Small magnitude (10^{-8} m for the sound of speech at 10m)

POLITECNICO MILANO 1863

M.S. Music and Acoustic Engineering

Sound pressure

- A vibrating surface produces changes in the air pressure. The maximum magnitude P_0 of such changes is the **sound pressure**
  - Measured in pascal [Pa] = [N/m^2] = 1 millibar
- The sound waveform is describes the pressure changes
- **The Root Mean Square value** of a sinusoid is 0.707 times P_0
  - $P_{eff} = P_0/\sqrt{2}$



POLITECNICO MILANO 1863

Acoustic power

- Acoustic power: energy per unit time
  - Measured in watt [W] = [joule/sec]

- Examples of max power of typical sound sources (W)

Speech	10^{-5}
Quarrel	10^{-3}
Bass singer	0.03
Clarinet	0.05
Trumpet	0.3
Piano	0.4
Trombone	6
Symphonic orchestra	60

POLITECNICO MILANO 1863

M.S. Music and Acoustic Engineering

Acoustic power

SOURCE	POWER (WATTS)	SOURCE	POWER (WATTS)
Orchestra (75 piece) @ fff	70	Orchestra (75 piece) @ mf	0.09
Bass Drum	25	Piccolo	0.08
Pipe Organ	13	Flute	0.06
Snare Drum	12	Clarinet	0.05
Cymbals	10	French Horn	0.05
Trombone	6	Triangle	0.05
Piano	0.4	Bass Voice	0.03
Bass Saxophone	0.3	Alto Voice (pp)	0.001
Bass Tuba	0.2	Average Speech [2]	0.000024
Double Bass	0.16	Violin (@ pppp)	0.0000038

POLITECNICO MILANO 1863

Intensity

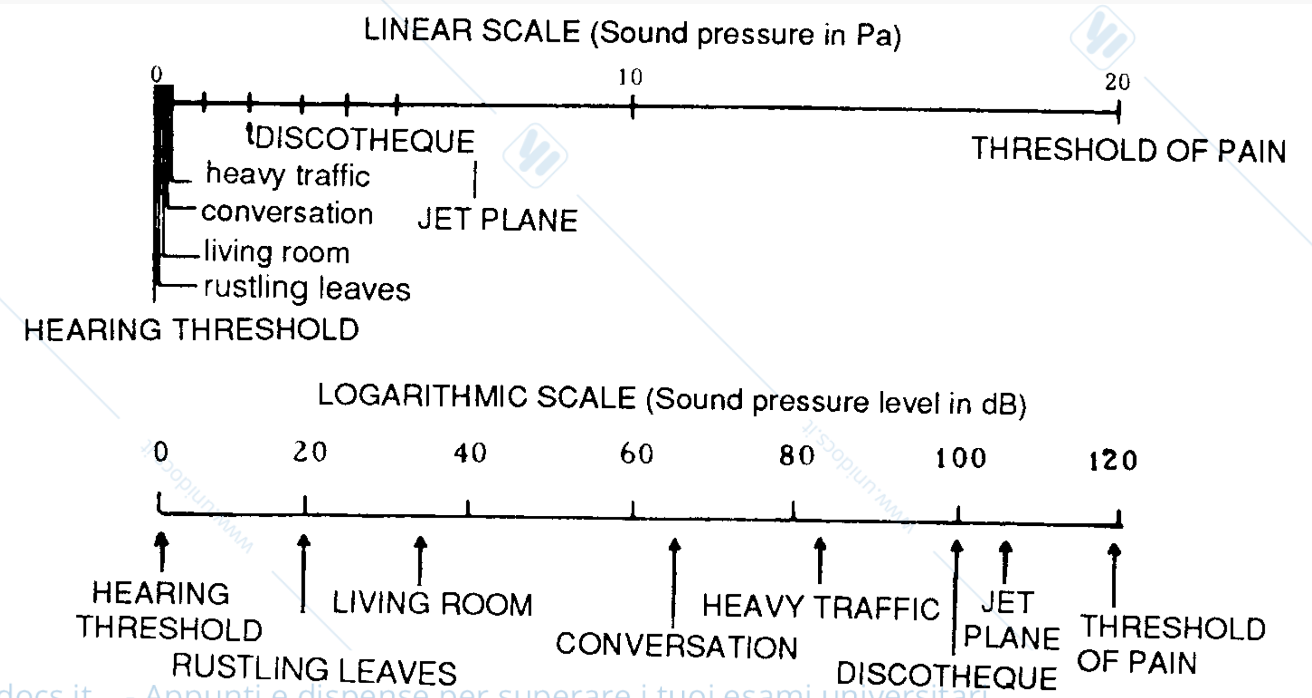
- Total average power that goes through a given surface
- Intensity

$$\mathbf{I} = (P_{eff})^2/\rho c \quad [watt/m^2]$$

(ρ=density of the medium, c=speed of sound)

- Due to the ear's extremely wide range of intensities (from 10^{-12} watt/m^2 to 1 watt/m^2) it is convenient to express such values in dB
- Pressure Level

$$\mathbf{PL} = 20 \log(P/P_{ref})$$



POLITECNICO MILANO 1863

M.S. Music and Acoustic Engineering

Intensity

Source	Intensity	Intensity Level	# of Times Greater Than TOH
Threshold of Hearing (TOH)	$1*10^{-12}$ W/m^2	0 dB	10^0
Rustling Leaves	$1*10^{-11}$ W/m^2	10 dB	10^1
Whisper	$1*10^{-10}$ W/m^2	20 dB	10^2
Normal Conversation	$1*10^{-6}$ W/m^2	60 dB	10^6
Busy Street Traffic	$1*10^{-5}$ W/m^2	70 dB	10^7
Vacuum Cleaner	$1*10^{-4}$ W/m^2	80 dB	10^8
Large Orchestra	$6.3*10^{-3}$ W/m^2	98 dB	$10^{9.8}$
Walkman at Maximum Level	$1*10^{-2}$ W/m^2	100 dB	10^{10}
Front Rows of Rock Concert	$1*10^{-1}$ W/m^2	110 dB	10^{11}
Threshold of Pain	$1*10^{1}$ W/m^2	130 dB	10^{13}
Military Jet Takeoff	$1*10^{2}$ W/m^2	140 dB	10^{14}
Instant Perforation of Eardrum	$1*10^{4}$ W/m^2	160 dB	10^{16}

POLITECNICO MILANO 1863

M.S. Music and Acoustic Engineering

Intensity

- Acoustic pressure level

  $$\mathbf{PL} = 20 \log(P/P_{ref})$$

- When $P_{ref} = 2\ 10^{-5}\ N/m^2$ (audibility threshold): Sound Pressure Level

  $$\mathbf{SPL} = 20 \log(P/\ 2\ 10^{-5}) = 20 \log P + 94$$

- Sound Power Level:

  $$\mathbf{PL} = 20 \log(W/W_0) \quad \text{where } W_0 = 10^{-12} \text{ watt}$$

- Intensity Level

  $$\mathbf{IL} = 20 \log(I/I_0) \quad \text{where } I_0 = 10^{-12} \text{ watt/m}^2$$

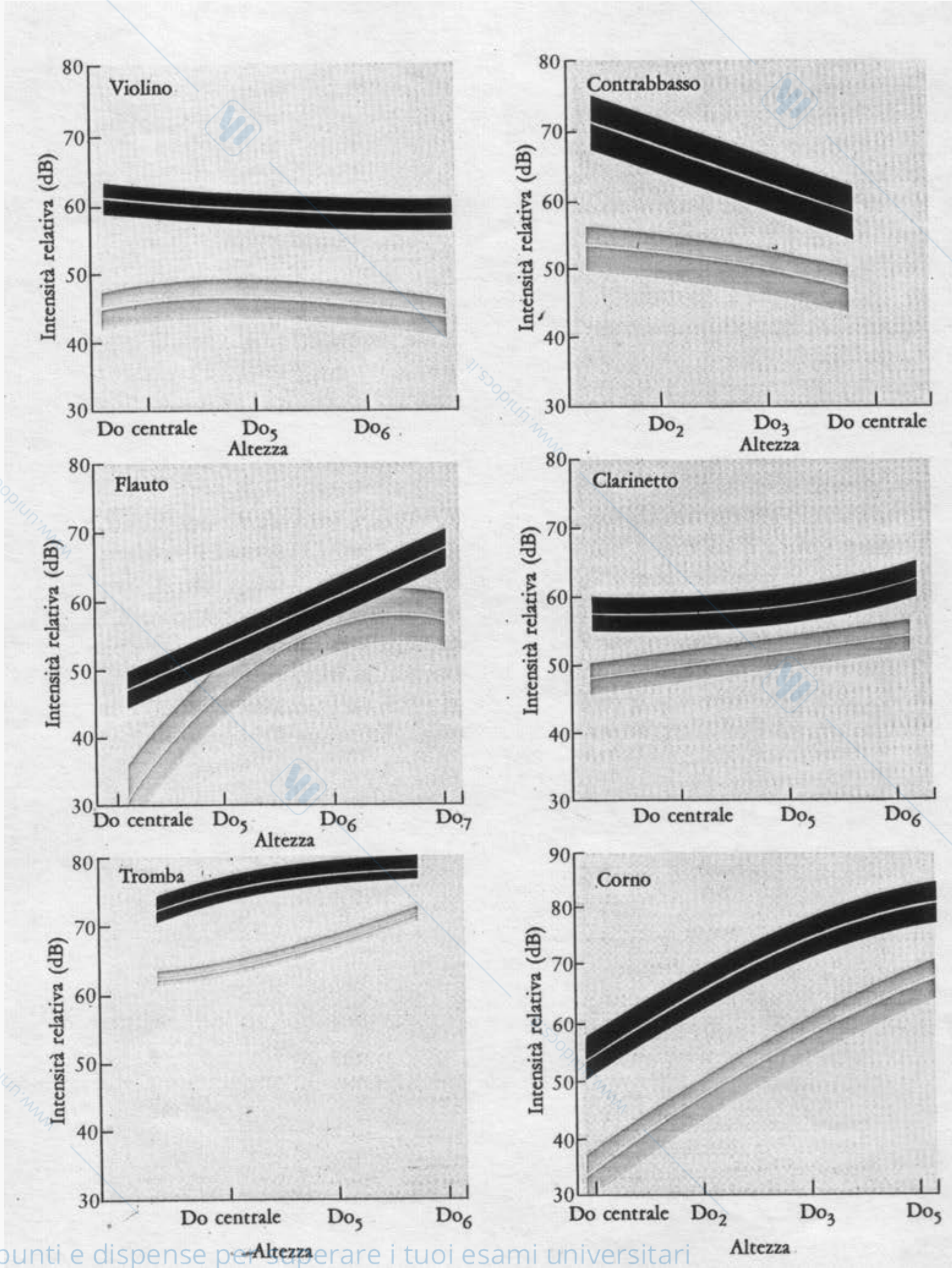
- 0 dB internationally represents the audibility threshold

- In some situations (sound recording) 0 dB corresponds to the desired level
  E.g. 6 dB descending ramp followed by a 3 dB descending ramp

POLITECNICO MILANO 1863

Intensity

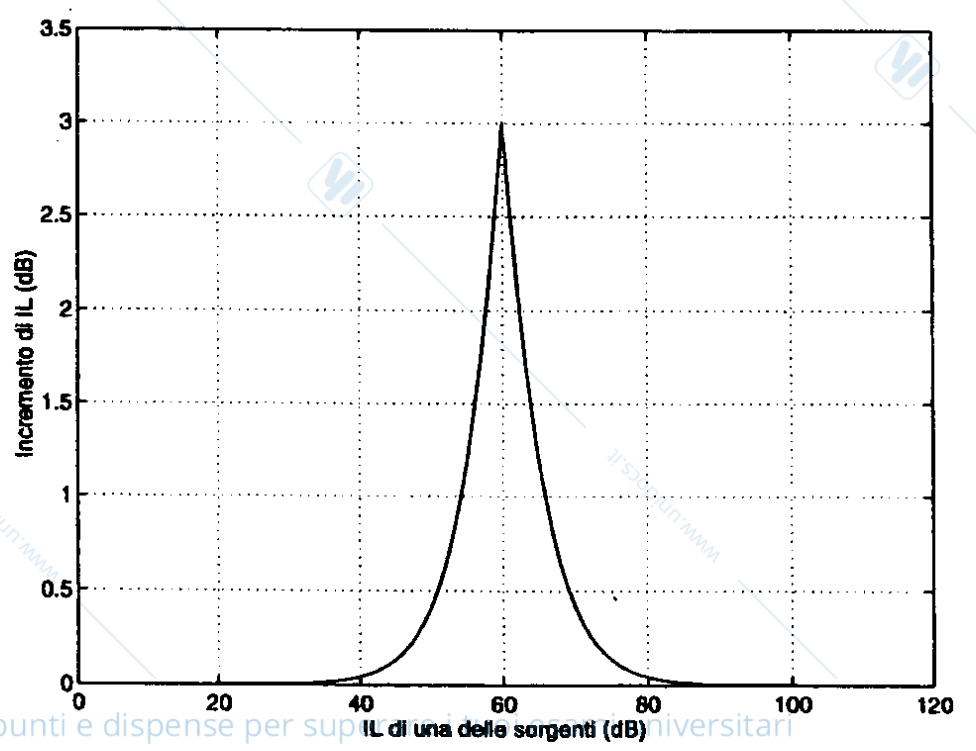
- Dynamic range of some musical instruments



(*)

Intensity: multiple sources

- Two incorrelated sources of the same intensity
  - + 3 dB Power Level
  - + 3 dB SPL
- Sources with different intensity: examples of increments of IL
  - Source a) IL= 60 db
  - Source b) IL from 0 to 120 dB

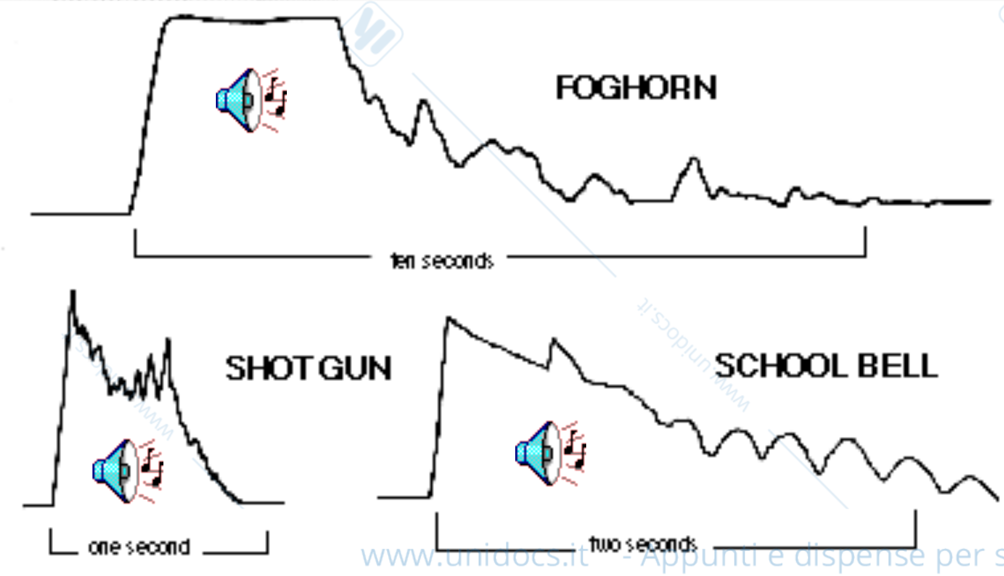


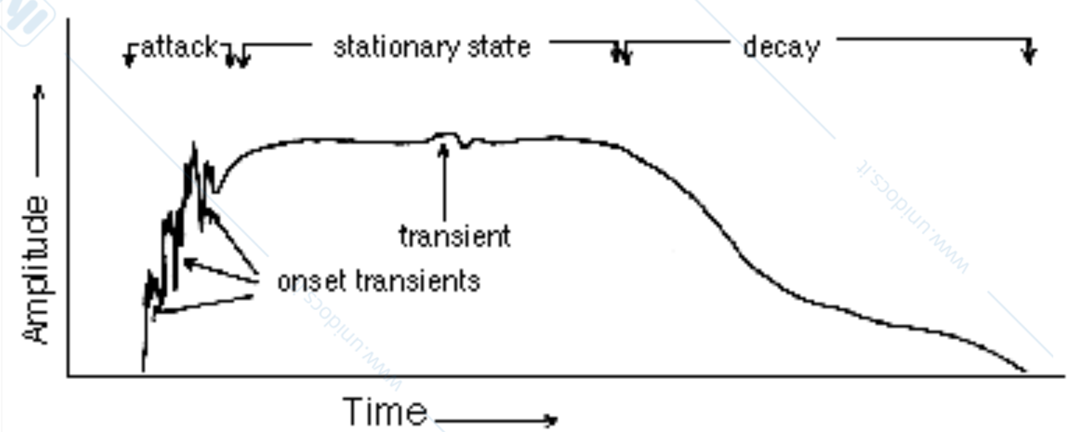
POLITECNICO
MILANO 1863

Envelope

(*)

- Temporal evolution of the macro-level of the sound pressure (waveform)
- Envelope phases:
  - *attack* (dall'istante iniziale, all'istante in cui l'inviluppo raggiunge il suo massimo)
  - *transients* (initial fluctuations)
  - *stationary state* ("sustain", from the and of the attack to the beginning of the decay)
  - *decay* (when we have loss of sound energy)
- Transients help localize the source and recognize the timbre
- Envelope generators are used in electronic music and synthesizers





www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari
www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari

(*)

Sound propagation

- Transmission of acoustic energy through a soundwave
- Propagation facts
  - Geometric diffusion – energy decay due to wavefront expansion
    - Spherical diffusion - 6 dB reduction per distance doubling
    - Cylindrical diffusion - 3 dB reduction per distance doubling
  - Atmospheric efffects
  - Surface effects

POLITECNICO MILANO 1863

M.S. Music and Acoustic Engineering

(*)

Absorption

- Loss of energy through some material
- What is not absorber is transmitted or reflected
- Absorption coefficient (α, 0÷1): percentage of absorbed energy (as a function of frequency)
  - Can be accurately measured
  - Important in the design of concert halls, theaters, recording studios, etc.
  - It depends on the level of *acoustic insulation* of the material
- Total absorption is measured in Sabines, and is calculated by multiplying the coefficient of absorption by the area of material in square feet

Absorptive wall materials	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz
Drapery (10 oz/yd2, 340 g/m2, flat against wall)	0.04	0.05	0.11	0.18	0.3	0.35
Drapery (14 oz/yd2, 476 g/m2, flat against wall)	0.05	0.07	0.13	0.22	0.32	0.35
Drapery (18 oz/yd2, 612 g/m2, flat against wall)	0.05	0.12	0.35	0.48	0.38	0.36
Drapery (14 oz/yd2, 476 g/m2, pleated 50%)	0.07	0.31	0.49	0.75	0.7	0.6
Drapery (18 oz/yd2, 612 g/m2, pleated 50%)	0.14	0.35	0.53	0.75	0.7	0.6
Fiberglass board (25mm(1″) thick)	0.06	0.2	0.65	0.9	0.95	0.98
Fiberglass board (50mm(2″) thick)	0.18	0.76	0.99	0.99	0.99	0.99
Fiberglass board (75mm(3″) thick)	0.53	0.99	0.99	0.99	0.99	0.99
Fiberglass board (100mm(4″) thick)	0.99	0.99	0.99	0.99	0.99	0.97
Open brick pattern over 75mm(3″) fiberglass	0.4	0.65	0.85	0.75	0.65	0.6
Pageboard over 25mm(1″) fiberglass board	0.08	0.32	0.99	0.76	0.34	0.12
Pageboard over 50mm(2″) fiberglass board	0.26	0.97	0.99	0.66	0.34	0.14
Pageboard over 75mm(3″) fiberglass board	0.49	0.99	0.99	0.69	0.37	0.15
Performated metal (13% open, over 50mm(2″) fiberglass)	0.25	0.64	0.99	0.97	0.88	0.92

Absorption

(*)

Reflective wall materials	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz
Brick (natural)	0.03	0.03	0.03	0.04	0.05	0.07
Brick (painted)	0.01	0.01	0.02	0.02	0.02	0.03
Concrete block (coarse)	0.36	0.44	0.31	0.29	0.39	0.25
Concrete block (painted)	0.1	0.05	0.06	0.07	0.09	0.08
Concrete (poured, rough finish, unpainted)	0.01	0.02	0.04	0.06	0.08	0.1
Doors (solid wood panels)	0.1	0.07	0.05	0.04	0.04	0.04
Glass (1/4″ plate, large pane)	0.18	0.06	0.04	0.03	0.02	0.02
Glass (small pane)	0.04	0.04	0.03	0.03	0.02	0.02
Plasterboard (12mm (1/2″) paneling on studs)	0.29	0.1	0.06	0.05	0.04	0.04
Plaster (gypsum or lime, on masonry)	0.01	0.02	0.02	0.03	0.04	0.05
Plaster (gypsum or lime, on wood lath)	0.14	0.1	0.06	0.05	0.04	0.04
Plywood (3mm(1/8″) paneling over 31.7mm(1-1/4″) airspace)	0.15	0.25	0.12	0.08	0.08	0.08
Plywood (3mm(1/8″) paneling over 57.1mm(2-1/4″) airspace)	0.28	0.2	0.1	0.1	0.08	0.08
Plywood (5mm(3/16″) paneling over 50mm(2″) airspace)	0.38	0.24	0.17	0.1	0.08	0.05
Plywood (5mm(3/16″) panel, 25mm(1″) fiberglass in 50mm(2″) airspace)	0.42	0.36	0.19	0.1	0.08	0.05
Plywood (6mm(1/4″) paneling, airspace, light bracing)	0.3	0.25	0.15	0.1	0.1	0.1
Plywood (10mm(3/8″) paneling, airspace, light bracing)	0.28	0.22	0.17	0.09	0.1	0.11
Plywood (19mm(3/4″) paneling, airspace, light bracing)	0.2	0.18	0.15	0.12	0.1	0.1

Floor materials	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz
Carpet	0.01	0.02	0.06	0.15	0.25	0.45
Concrete (unpainted, rough finish)	0.01	0.02	0.04	0.06	0.08	0.1
Concrete (sealed or painted)	0.01	0.01	0.02	0.02	0.02	0.02
Marble or glazed tile	0.01	0.01	0.01	0.01	0.02	0.02
Vinyl tile or linoleum on concrete	0.02	0.03	0.03	0.03	0.03	0.02
Wood parquet on concrete	0.04	0.04	0.07	0.06	0.06	0.07
Wood flooring on joists	0.15	0.11	0.1	0.07	0.06	0.07

Absorption

(*)

Seating Materials	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz
Benches (wooden, empty)	0.1	0.09	0.08	0.08	0.08	0.08
Benches (wooden, 2/3 occupied)	0.37	0.4	0.47	0.53	0.56	0.53
Benches (wooden, fully occupied)	0.5	0.56	0.66	0.76	0.8	0.76
Benches (cushioned seats and backs, empty)	0.32	0.4	0.42	0.44	0.43	0.48
Benches (cushioned seats and backs, 2/3 occupied)	0.44	0.56	0.65	0.72	0.72	0.67
Benches (cushioned seats and backs, fully occupied)	0.5	0.64	0.76	0.86	0.86	0.76
Theater seats (wood, empty)	0.03	0.04	0.05	0.07	0.08	0.08
Theater seats (wood, 2/3 occupied)	0.34	0.21	0.28	0.53	0.56	0.53
Theater seats (wood, fully occupied)	0.5	0.3	0.4	0.76	0.8	0.76
Seats (fabric-upholsterd, empty)	0.49	0.66	0.8	0.88	0.82	0.7
Seats (fabric-upholsterd, fully occupied)	0.6	0.74	0.88	0.96	0.93	0.85

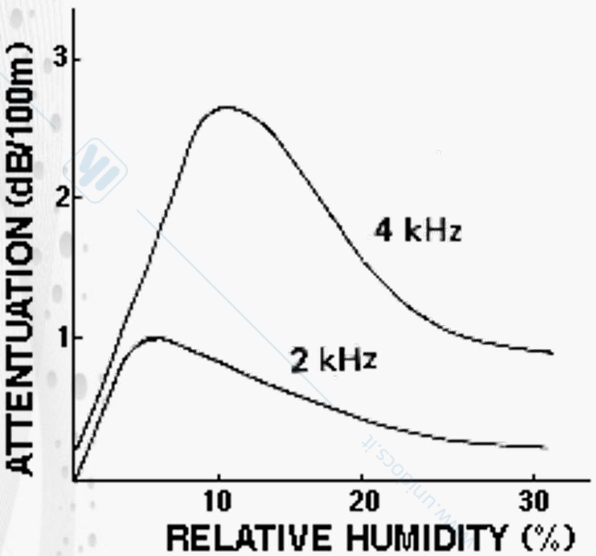
Ceiling Maeterial	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz
Plasterboard (12mm(1/2") in suspended ceiling grid)	0.15	0.11	0.04	0.04	0.07	0.08
Underlay in perforated metal panels (25mm(1") batts)	0.51	0.78	0.57	0.77	0.9	0.79
Metal deck (perforated channels,25mm(1") batts)	0.19	0.69	0.99	0.88	0.52	0.27
Metal deck (perforated channels, 75mm(3") batts)	0.73	0.99	0.99	0.89	0.52	0.31
Plaster (gypsum or lime, on masonary)	0.01	0.02	0.02	0.03	0.04	0.05
Plaster (gypsum or lime, rough finish or timber lath)	0.14	0.1	0.06	0.05	0.04	0.04
Sprayed cellulose fiber (16mm(5/8") on solid backing)	0.05	0.16	0.44	0.79	0.9	0.91
Sprayed cellulose fiber (25mm(1") on solid backing)	0.08	0.29	0.75	0.98	0.93	0.76
Sprayed cellulose fiber (25mm(1") on timber lath)	0.47	0.9	1.1	1.03	1.05	1.03
Sprayed cellulose fiber (32mm(1-1/4") on solid backing)	0.1	0.3	0.73	0.92	0.98	0.98
Sprayed cellulose fiber (75mm(3") on solid backing)	0.7	0.95	1	0.85	0.85	0.9
Wood tongue-and-groove roof decking	0.24	0.19	0.14	0.08	0.13	0.1

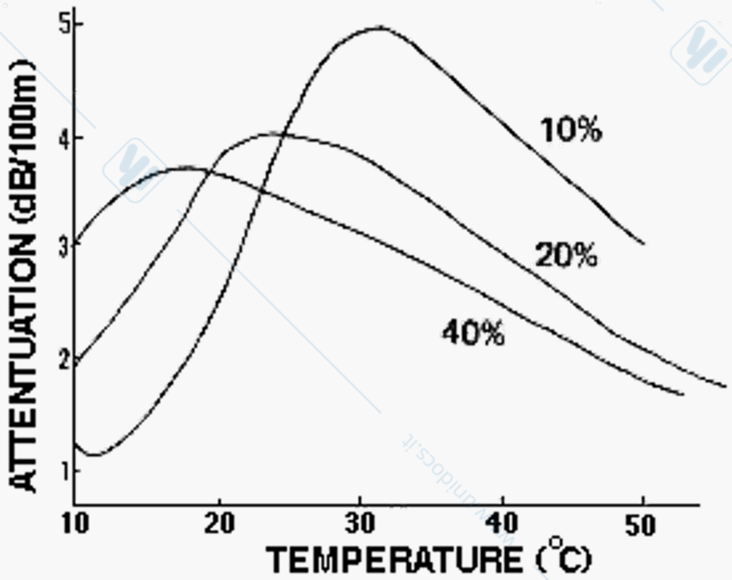
Miscellaneous surface material	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz
People-adults (per 1/10 person)	0.25	0.35	0.42	0.46	0.5	0.5
People-high school students (per 1/10 person	0.22	0.3	0.38	0.42	0.45	0.45
People-elementary students (per 1/10 person)	0.18	0.23	0.28	0.32	0.35	0.35
Ventilating grilles	0.3	0.4	0.5	0.5	0.5	0.4
Water or ice surface	0.008	0.008	0.013	0.015	0.02	0.025

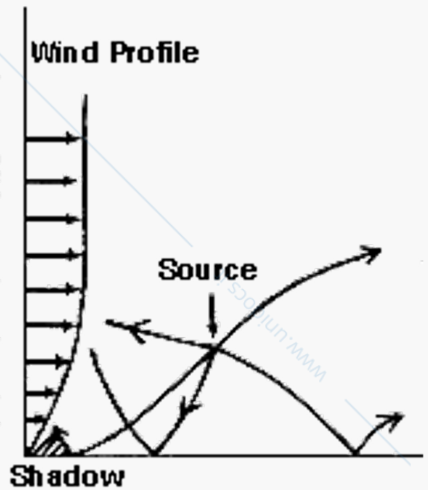
Atmospheric effects

(*)

- **Absorption due to air**
  Depends on temperature and humidity of the atmosphere (25 dB/100 m at 2 kHz for 30% of relative humidity and 20° C)
- **Temperature gradients**
  Propagation speed increases with temperature.
  - If temperature increases with altitude (sunny winter days), sound is reflected downward and propagation follows the earth curvature
  - Due to turbulence, waves propagating in the direction of the wind tend to be deflected downward. Those against the wind go upward



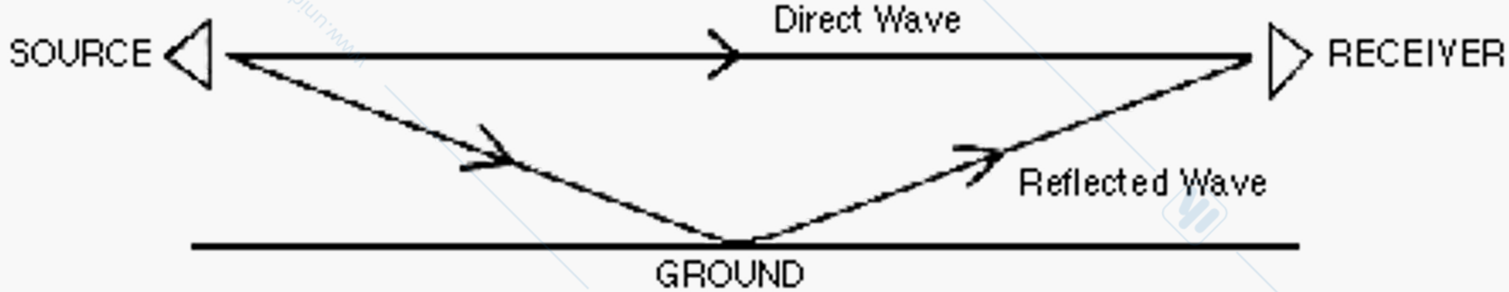




Surface effects

(*)

- **Terrain absorption**
  Dipends on surface
  - Rock surfaces absorb very little
  - Soft dirt can cause the signal to drop up to 100dB per 100m at 2 kHz
  - Higher frequencies are more absorbed than lower frequencies
- **Ground effect**
  When the source and the receiver are close to the ground, the reflected wave can interfere with the incident one in a destructive way



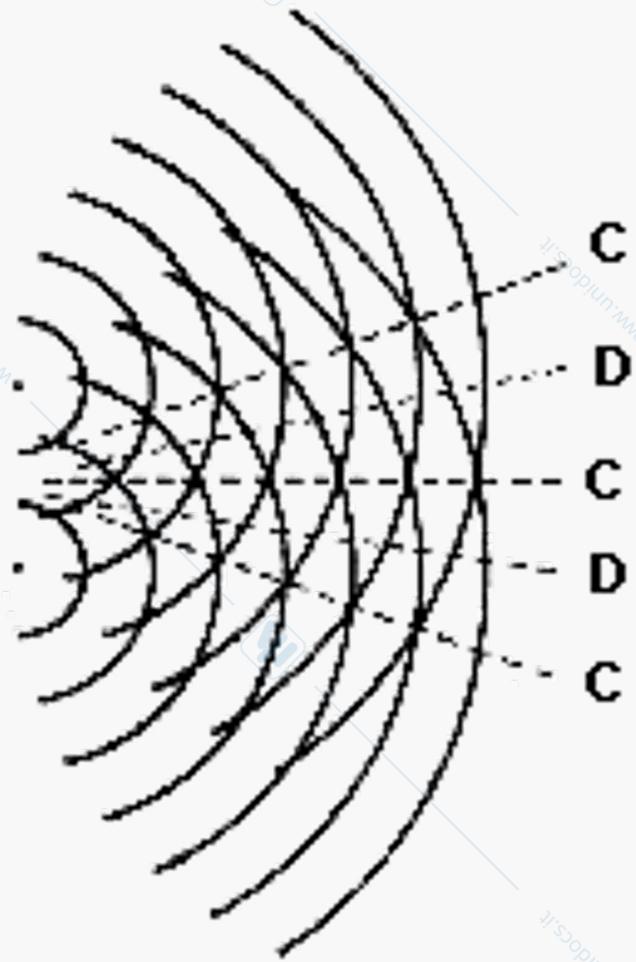
- **Attenuation due to barriers** (trees, buildings, hills)
  - Barriers can be used to reduce ground effect
  - They work better on higher frequencies (lower frequencies are diffracted at the borders
  - Up to 40 dB reduction (the rest is propagated by the atmosphere).
  - More effective near source or receiver
  - Wind or temperature gradients make them ineffective
  - **Canyon effect**. Reverberation due to multiple reflections. Makes it difficult to localize source. Increases environmental noise.

POLITECNICO MILANO 1863

Interference

(*)

- Sound waves coming from different sources (same frequency)
  - Where wave maxima meet we have constructing interference (C)
  - Where wave maxima meet with wave minima we have destructive interference (D)



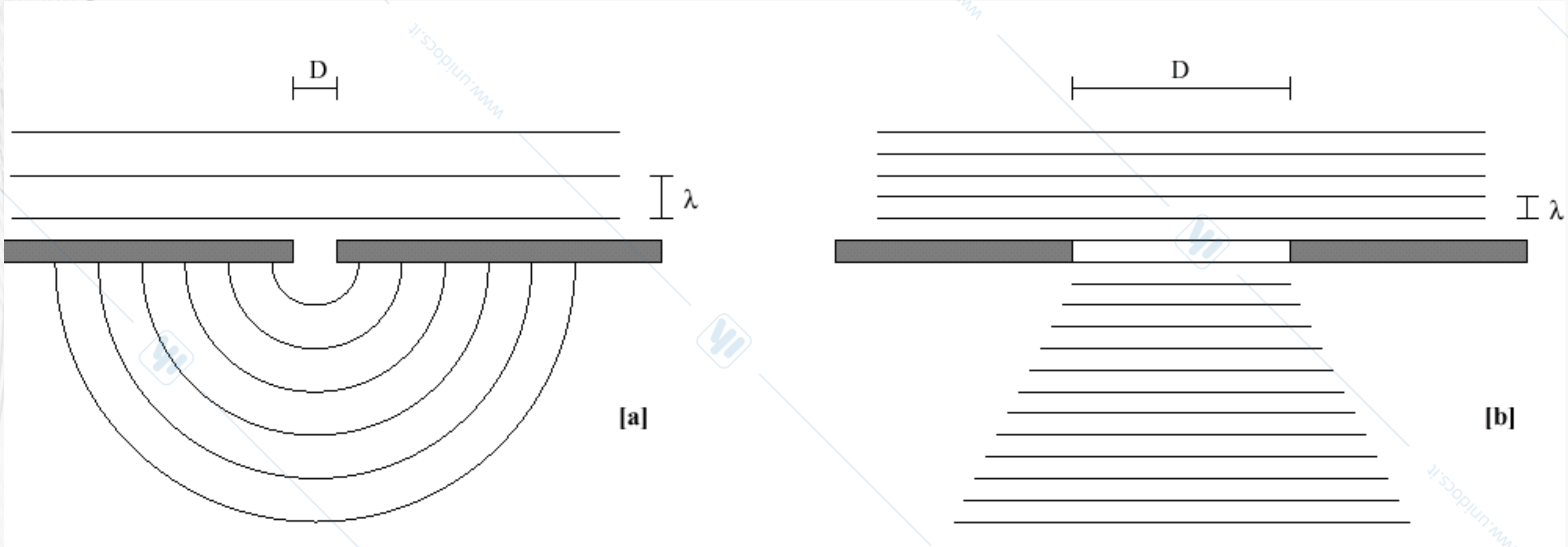
POLITECNICO
MILANO 1863

Diffraction

(*)

[a]: D< λ - isotropic source

- [b]: D> λ - directional source

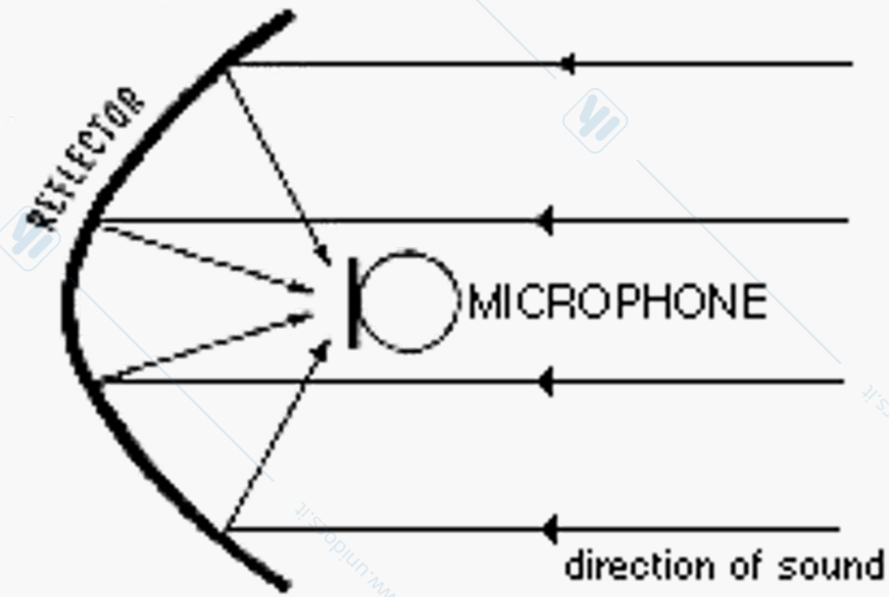


POLITECNICO
MILANO 1863

Reflection and rifraction

(*)

- Example
  - Multiple echoes recorded under a parabolic bridge in Stanley Park, Vancouver. The sound source is a stick hitting a metal pipe
- Echo
- Reverberation
- Parabolic reflector

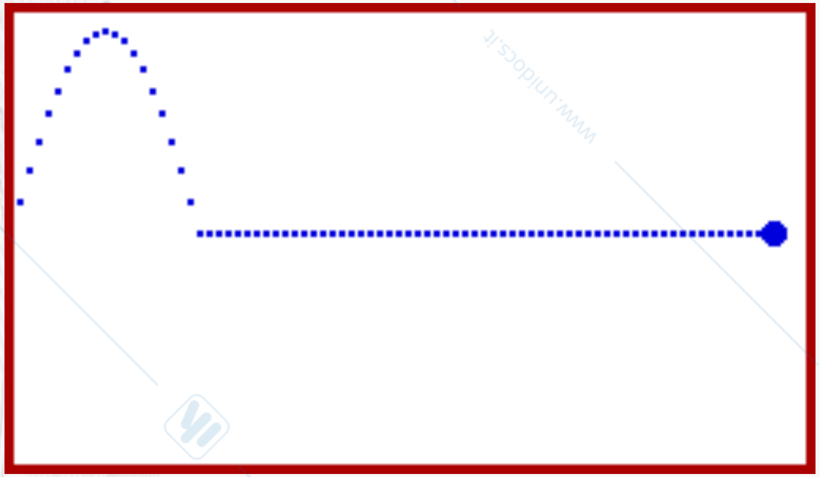


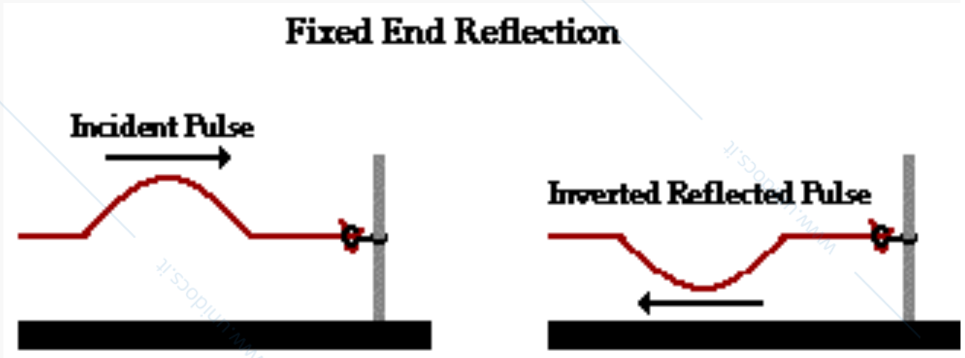
POLITECNICO
MILANO 1863

Reflection

(*)

- Impulse reflection at a fixed termination
- Impulse reflection from a less dense medium to a denser one



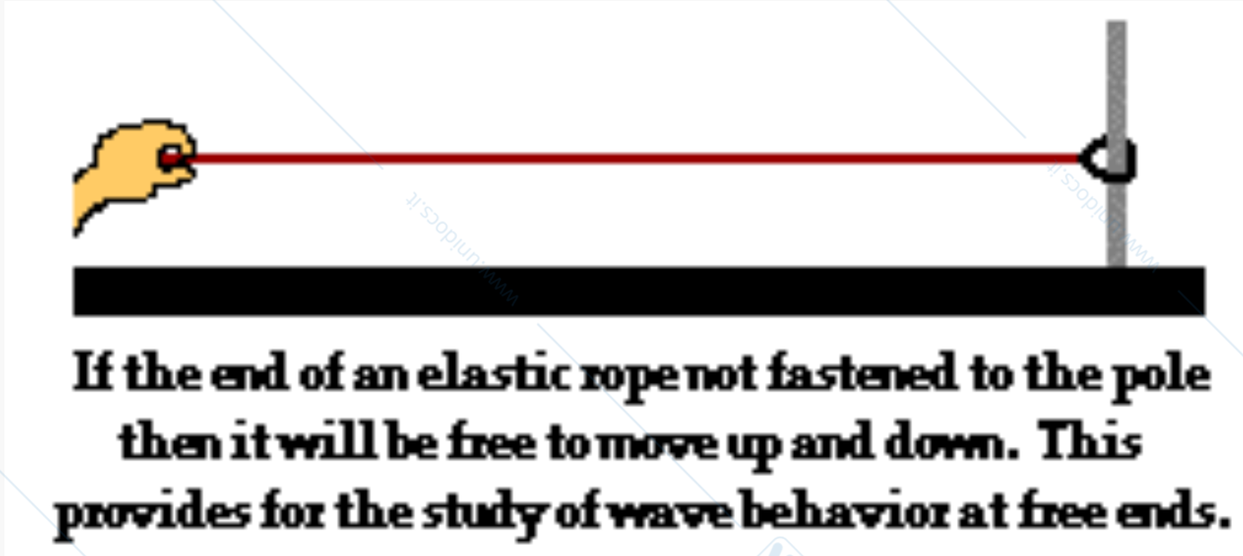


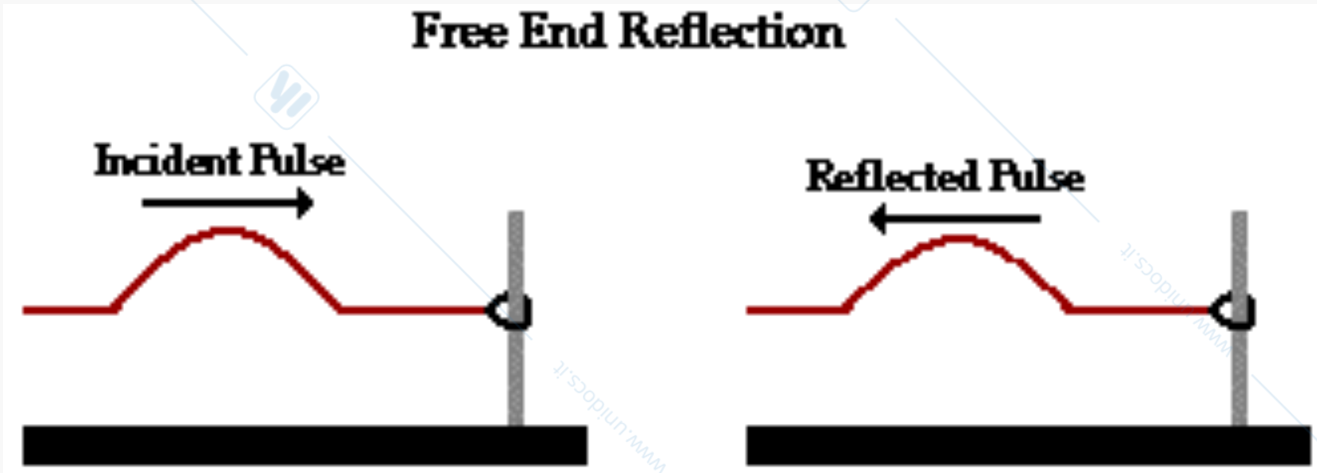
POLITECNICO MILANO 1863

Reflection

(*)

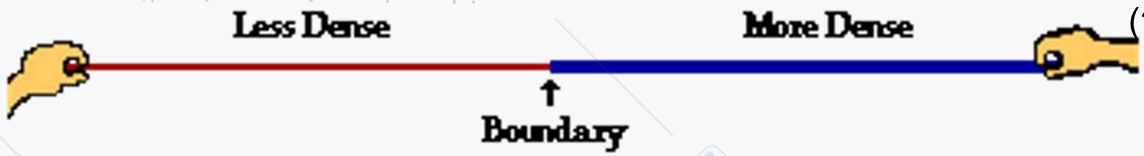
- Free termination
- From higher to lower density



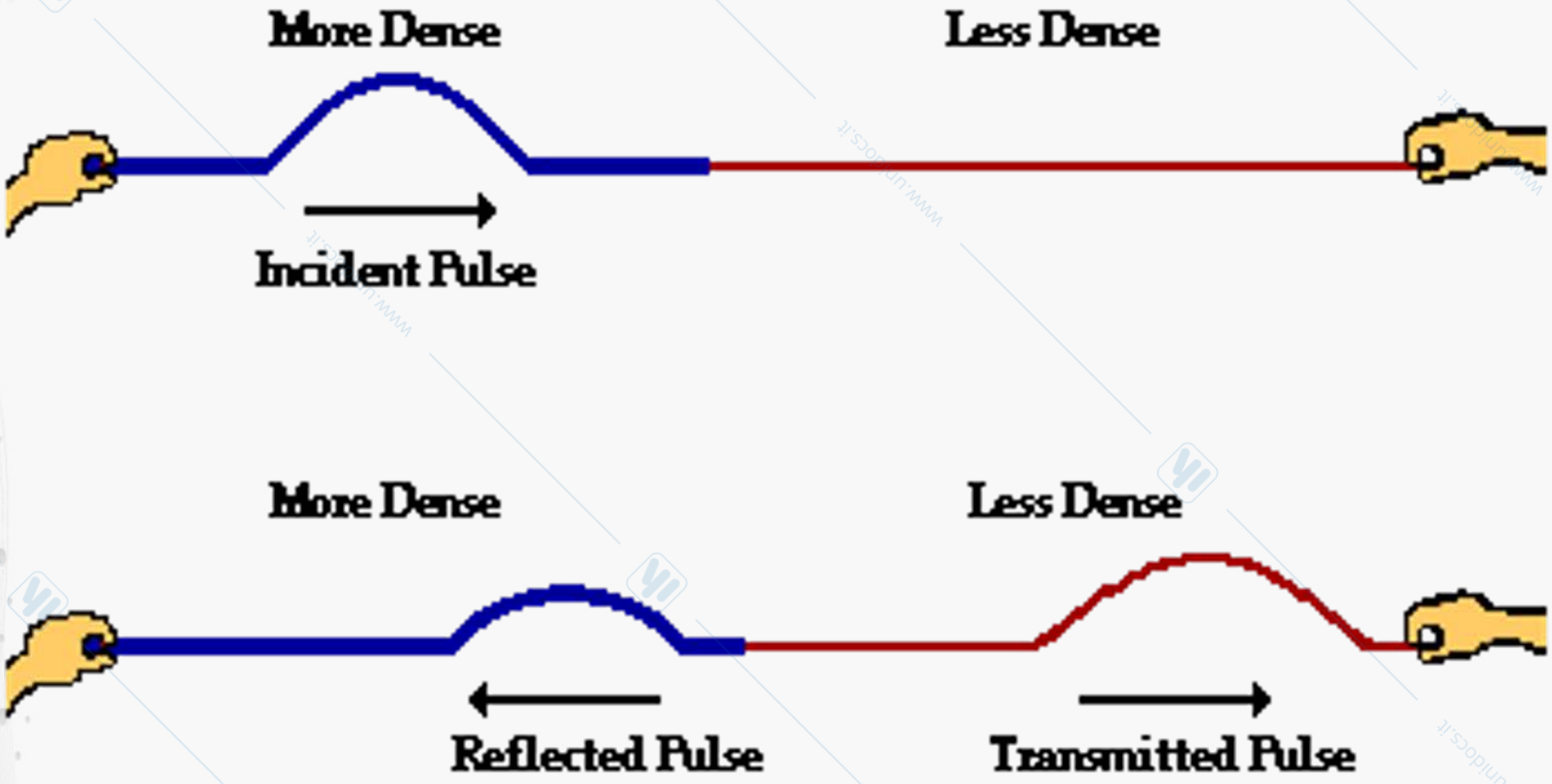


POLITECNICO
MILANO 1863

Transmission



A wave traveling from a more dense to a less dense medium…



…will be reflected off the boundary and transmitted across the boundary into the new medium. There is no inversion.

Transmission

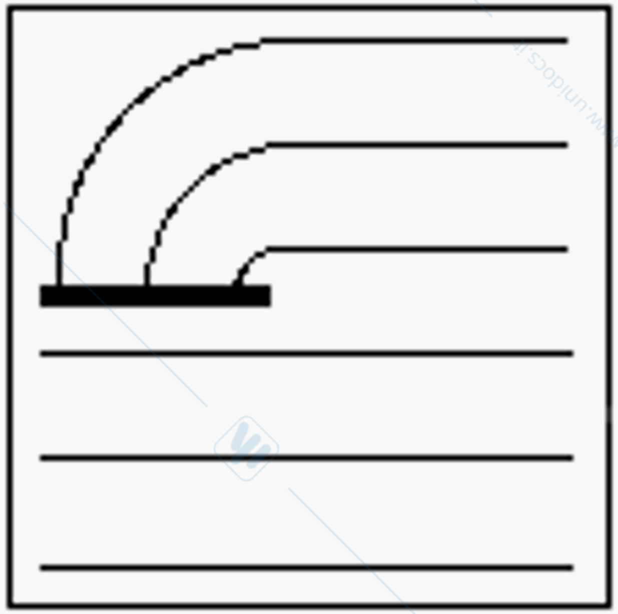
(*)

- wavelength is always greater in the least dense medium
- the frequency of a wave does not change while crossing a boundary
- the reflected pulse is inverted when a wave in a less dense medium travels toward a boundary with a denser medium
- the amplitude of the incident pulse is always greater than the amplitude of the reflected pulse

Diffraction

(*)

- Diffraction is the bending of waves around obstacles and openings
- The amount of diffraction increases with increasing wavelength.

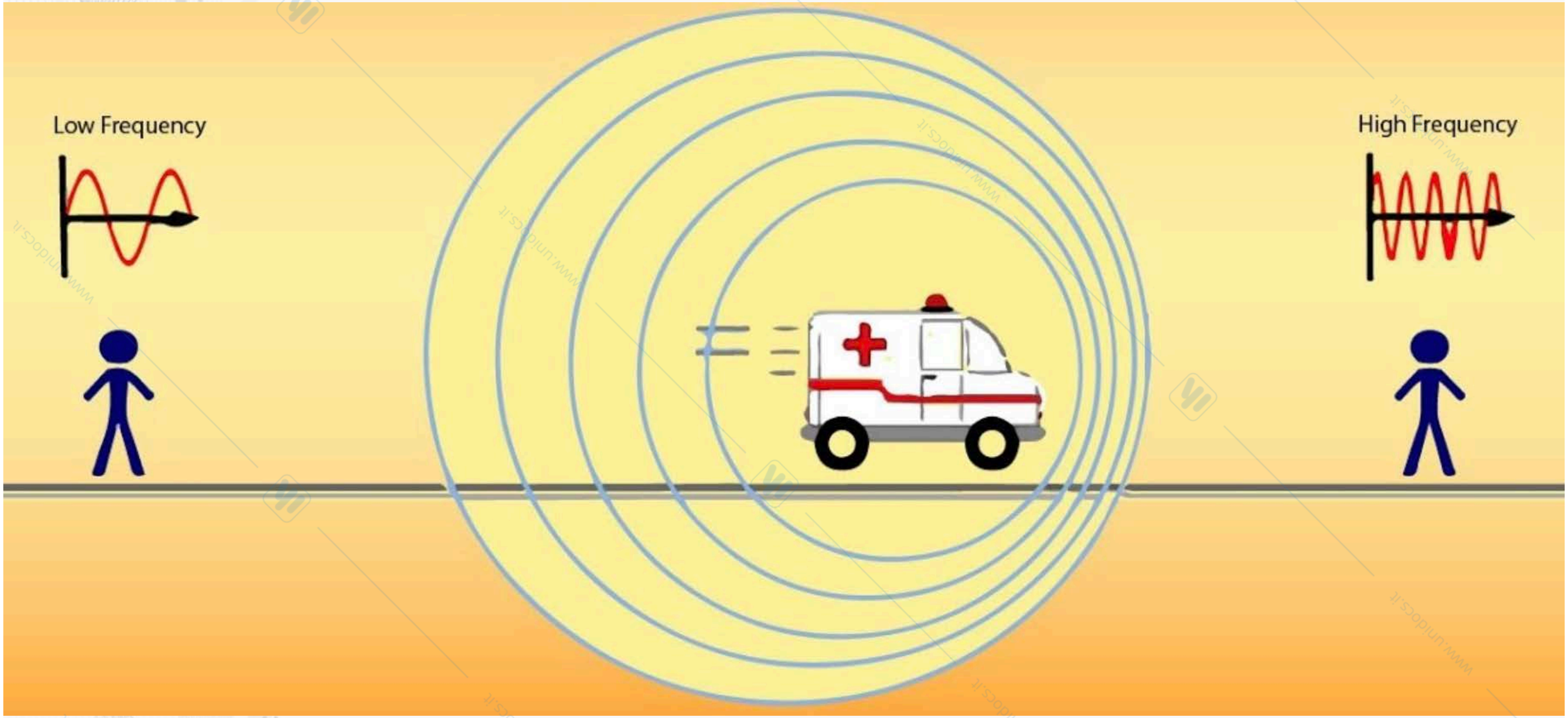


Diffraction of Waves

POLITECNICO MILANO 1863

Doppler effect

(*)



(*)

Echo

- Sound repetition due to reflections (outdoor)
- In order to be called an echo (insted of a reverberation) there have to be at least 50 ms (usually 100) of pause between original sound and reflection
- Echo at the Rolley Lake
- Ferry blast (0.25 s echo)

POLITECNICO
MILANO 1863

Reverberation

(*)

- Multiple reflections indoor or in a semi-open environment
- Similar to an echo, but repetition rate is too fast to perceive sounds as separate
- Reverberation time: time that it takes to achieve sound attenuation of 1.000.000 times (IL = –60 dB w.r.t. original sound)
- Rev. time is proportional to volume of propagation, and inversely proportional to the sum of all reflecting surfaces multiplied by their absorption coefficients
- It constitutes an important clue to human sense of balance and orientation (greatly diminished in an anechoic chamber)
- It increases the environment noise, therefore it needs to be kept under control in the design of Lo-Fi environments (workplace, classrooms, industrial spaces, etc.)

POLITECNICO MILANO 1863

Reverberation

(*)

- In symphonic music reverberation increases the sound *blend. Suitable rev. time:* btw 1s and 2s
- In Hi-Fi rooms, we often try to boost reflection in order to achieve rev. times of at least 1s. A good design seeks a type of reverbaration that adds *spatiality*, *warmth* and *coverage* to sounds
- Goal: good balance btw clarity/definition and spatiality
  - Storkyrkan church, Stockholm, Sweden
  - Canadian Pacific train station, Vancouver
  - Marble hallway, Parliament Buildings, Ottawa
  - Footsteps in covered bridge, Chatham
  - Keys, locks and heavy doors in the vaults of the National Library, Vienna, Austria
- Digital techniques allow us to control reverberation in a post-production phase or in real time (live-electronics) to simulate different environments
  - 4 examples: *bright*, *dark*, *warm* and *closed* environments

POLITECNICO MILANO 1863

www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari
www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari

(*)

Timbre

- Determined by
  - Frequency content
  - Transients
  - Presence of and relation btw partials, which can be in
    - Harmonic ratio
      - Harmonic series (up to the16th) synthesized with sinusoids
    - Inharmonic ratio (non-multiple relations, e.g. percussions)
      - Chinese chau gong (tam-tam)
  - Phase relations
    - Two sinusoids are in phase if they simultaneously arrive at the same point of their cycle
- We perceive it as a *gestaltic* perception of the whole sound rather than as a function of some of its feature or component

POLITECNICO
MILANO 1863

Elements of Psychoacoustics

POLITECNICO
MILANO 1863

M.S. Music and Acoustic Engineering

Psychoacoustics

- Psychoacoustics is the scientific study of sound perception and audiology
- It studies the psychological and physiological responses associated with sound
- It can be further categorized as a branch of psychophysics
- Psychoacoustics owes its name from a field within psychology (recognition science) which deals with human perceptions
- It is an interdisciplinary field of many areas, including psychology, acoustics, electronic engineering, physics, biology, physiology, and computer science

POLITECNICO
MILANO 1863

Overview

- Source
  - Exciter (hammer, reed, …): provides the energy
  - Resonator: where vibrations reside, characterizes the fundamental
    - Vibrational element (string, pressure waves, …)
    - Resonator (soundboard, body)
- Transmission means
  - Environment: sound propagation
  - “boundaries”: reflections, reverberations, absorption
- Receiver (ear)
  - Transformation in mechanical oscillations
  - Conversion into electrical signals
  - Processing - identification

POLITECNICO MILANO 1863

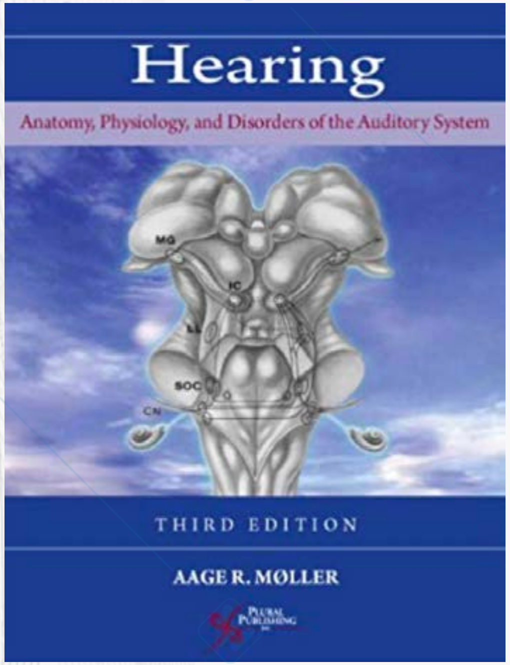
The sound pipeline...

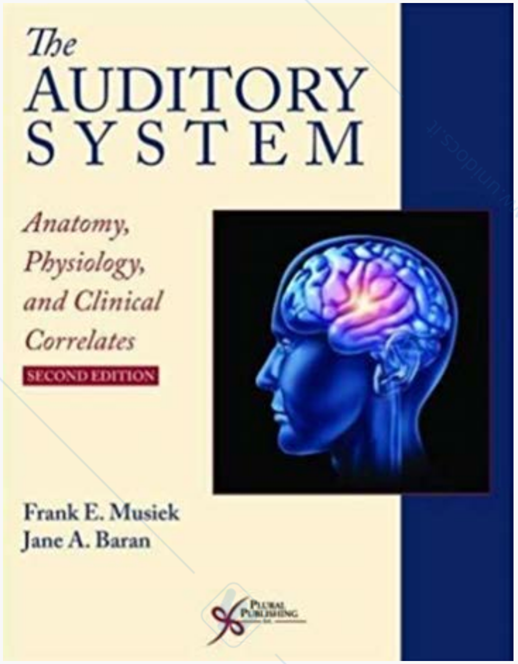
- Source
  - Exciter (hammer, reed, …): provides the energy
  - Resonator: where vibrations reside, characterizes the fundamental
    - Vibrational element (string, pressure waves, …)
    - Resonator (soundboard, body)
- Transmission means
  - Environment: sound propagation
  - “boundaries”: reflections, reverberations, absorption
- Receiver (ear)
  - Transformation in mechanical oscillations
  - Conversion into electrical signals
  - Processing - identification

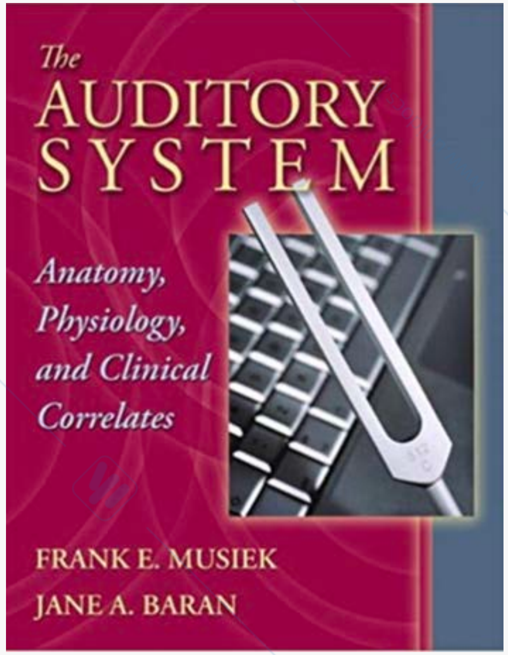
POLITECNICO
MILANO 1863

www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari
www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari

The auditory system







POLITECNICO
MILANO 1863

The auditory system

- Ear Anatomy: "How the Ear Functions" 1940 (Knowledge Builders)*
  - https://youtu.be/6-izelfgxmw
- How the ear works (Javitz)
  - https://youtu.be/qgdqp-oPb1Q
- How your ear works - Inside the Human Body: Building Your Brain (BBC One)
  - https://youtu.be/r-c5GpoD8wI

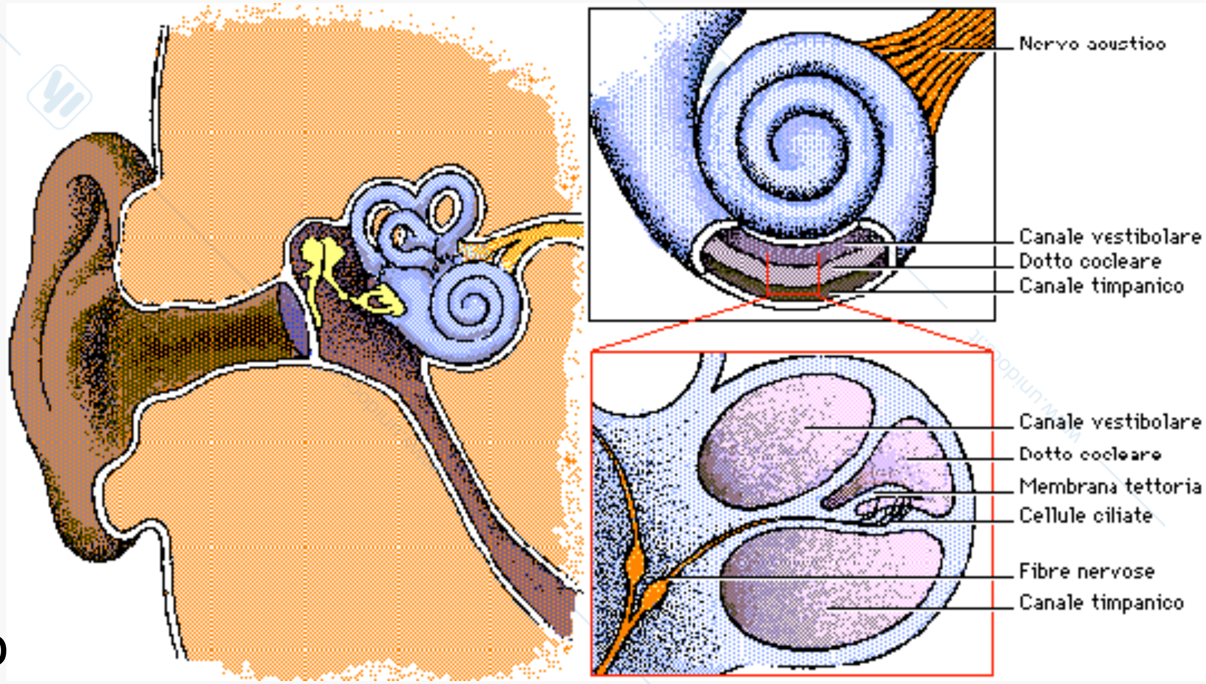
* Old but really good!

POLITECNICO MILANO 1863

M.S. Music and Acoustic Engineering

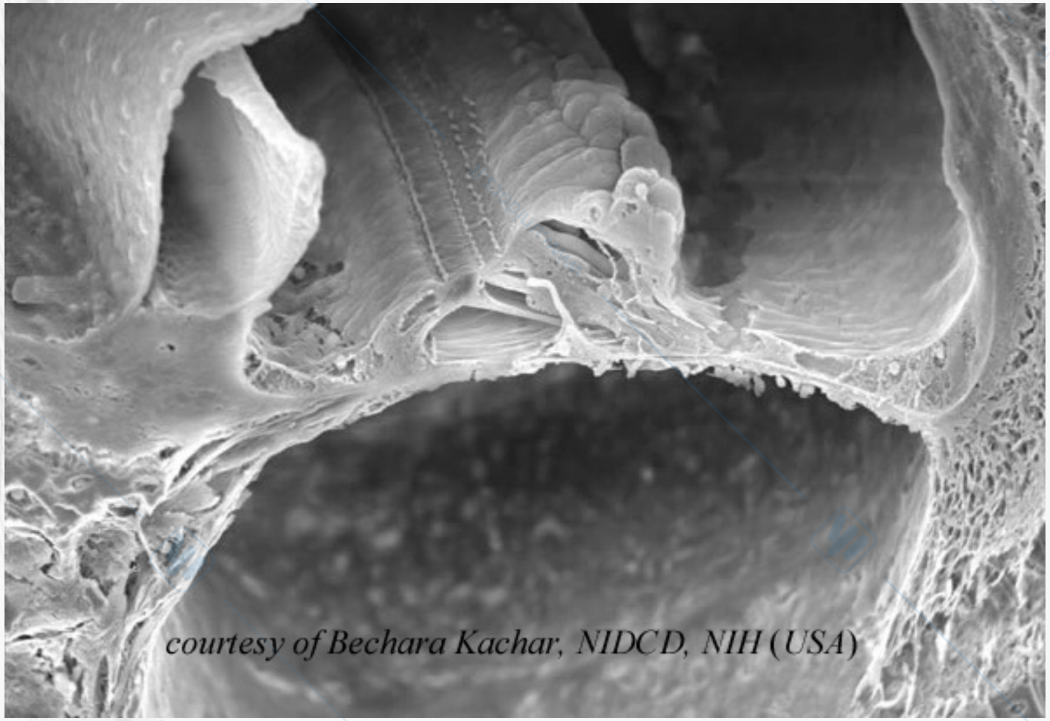
The ear

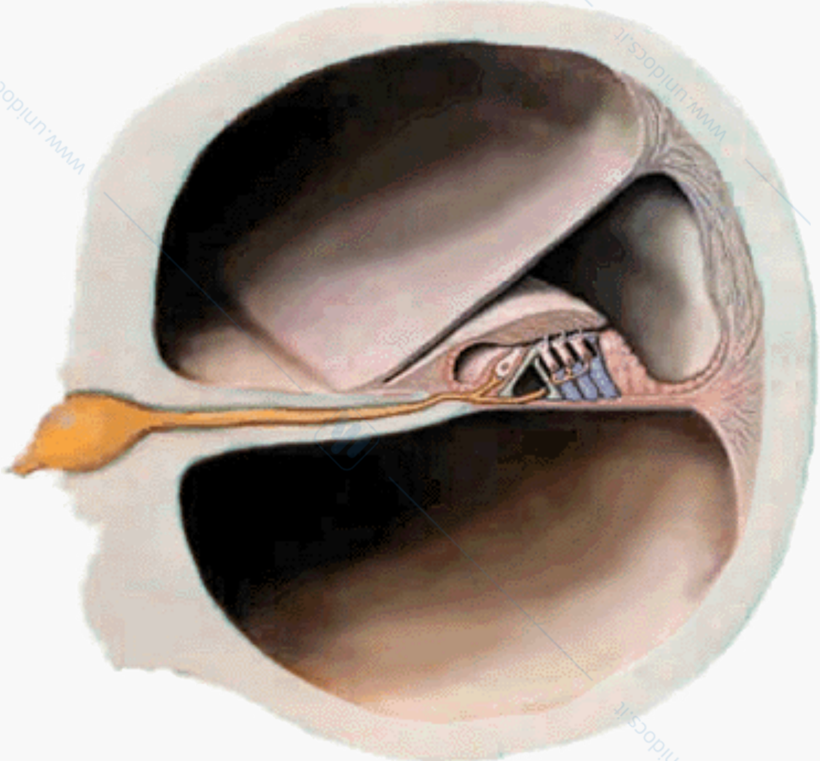
- Outer ear
  - Ear lobe (pinna)
  - Hearing duct
- Middle ear
  - The ear drum transmits the vibrations to the
  - Ear bones, which amplify the pressure (and reduces the motion)and pass it onto the oval window Inner ear
  - Oval window
  - Cochlea = converts vibrations into electrical signals
    - Filled with liquid (perilymph), the cochlea is divided by the basilar membrane
    - The liquid passes through an aperture (helicotrema) in the membrane
    - Throughout the basilar membrane are 30.000 hair cells (neural receptors), which convert the wave motion into neural pulses
    - Each region of the basilar membrane is maximally sensitive to a different frequency
    - Resonance frequency doubles every 3.5 mm



POLITECNICO
MILANO 1863

Structure of the cochlea





Superposition of pure tones: first-order effects

- First-order beats (originated in the choclea)
  - 2 sounds at f_1 and f_2
  - If f_2 is slightly higher than f_1 (up to $f_2 = f_1 + 15$ Hz), we perceive a single sound at frequency $f = (f_1+f_2)/2$, whose amplitude is modulated at half the frequency deviation

$$\sin(2\pi f_1 t) + sin(2\pi f_2 t) = 2\sin\left(2\pi \frac{f_1+f_2}{2} t\right)\cos\left(2\pi \frac{f_1-f_2}{2} t\right)$$

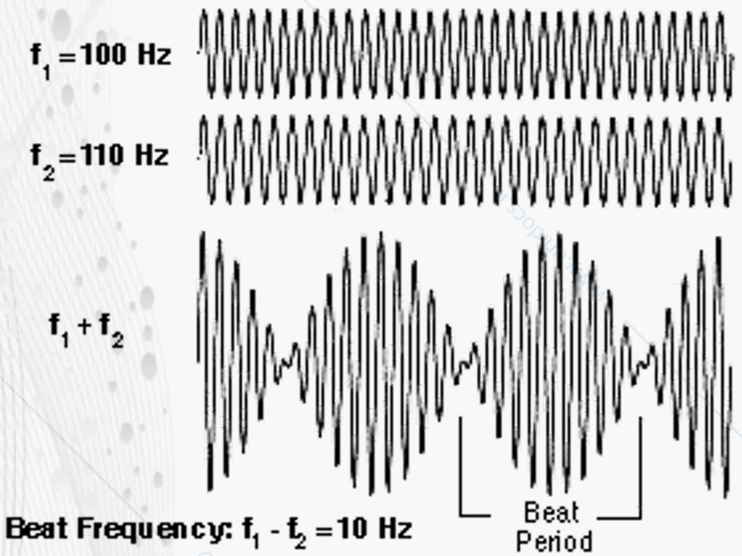
Tone frequency

Modulation (beat) fequency

POLITECNICO MILANO 1863

M.S. Music and Acoustic Engineering

Superposition of pure tones: first-order effects



Two sounds at 100 e 110 Hz
Higher tone's frequency gradually decreases to 100 Hz

- Combination sounds (originated in the cochlea)
  - Perceived even if they are note present in the stymulus
  - Arise from NL distortions in the middle ear
  - Frequencies: f_1, f_2, f_1+f_2, f_1-f_2, $2f_1$, $2f_2$

POLITECNICO MILANO 1863

Musical beats (Cook 2)

- Risset's musical beats
- 100 Hz sound with 7 harmonics of magnitude 1/f
- 7 of such sounds with fundamentals 100, 100.1, 100.2, 100.3, 100.4, 100.5, 100.6 Hz
- 7 such sounds with fundamentals 700.0, 700.7, 701.4, 702.1, 702.8, 703.5, 704.2 Hz

  synthesize the seventh partial of the previous example

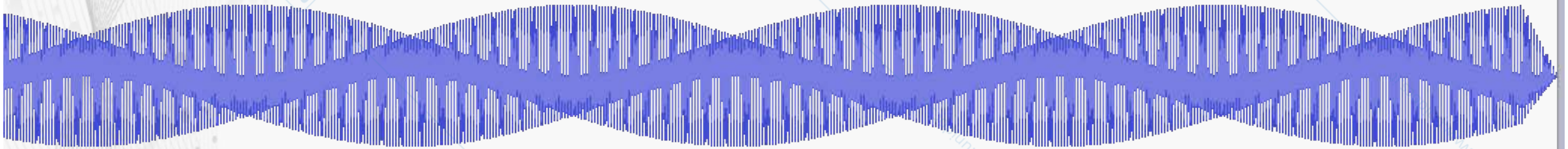
POLITECNICO
MILANO 1863

Superposition of pure tones: second-order effects

- They originate in the neural system
  - 2 sounds at f_1 and f_2
  - Se $f_2$is slighly off $2f_1$, then the phase difference is not constant
  - We perceive a tone beat at f_2-$2f_1$
  - These are no longer amplitude modulations (first order), but cyclic changes in the waveshape

POLITECNICO MILANO 1863

Superposition of pure tones: second-order effects



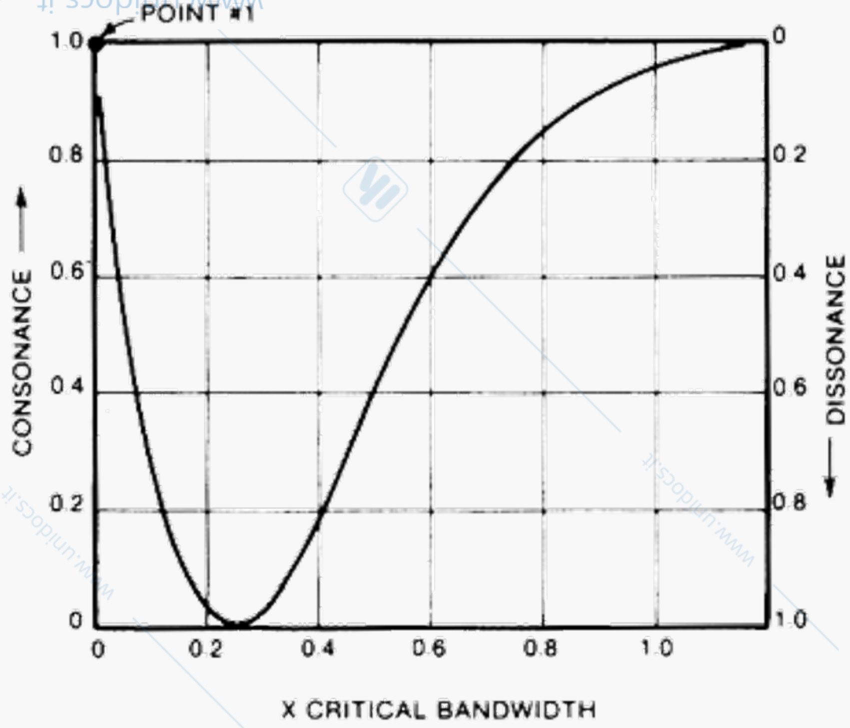
Two sounds at 100 and 201 Hz
(beat frequency = 1 Hz)

- Reconstruction of the fundamental
  - Sound with f_1, $2f_1$, $3f_1$, $4f_1$, … → perceived pitch = f_1
  - Sound with $2f_1$, $3f_1$, $4f_1$, … → perceived pitch = f_1

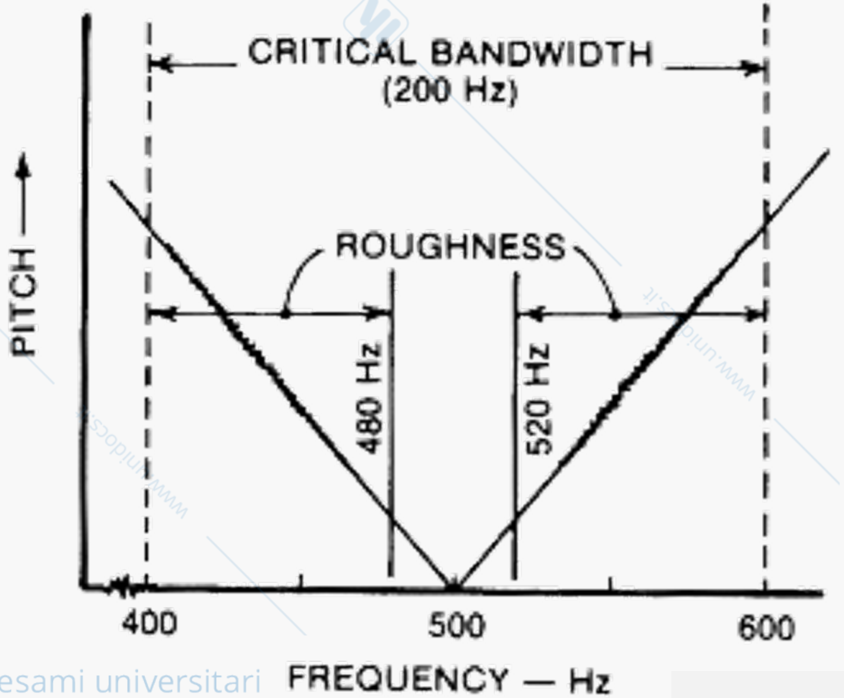
POLITECNICO MILANO 1863

Consonance

- Same frequence (pt. 1)
- About a fourth: max dissonance
- About ½ critical band: 40% consonance improvement
- ¾ critical band: 80% consonance improvement
- 1 critical band: 100% consonant



Ref.: Plomp and Levelt
J. Acous. Soc. Am. 38 (1965) 548



POLITECNICO
MILANO 1863

Pitch

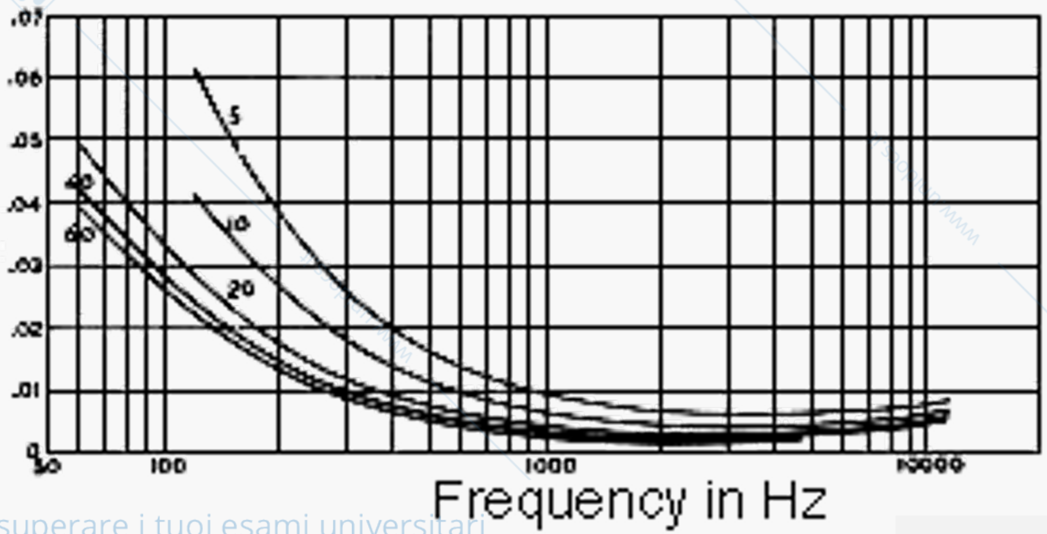
Subjective impression of frequency

- The frequency corresponding to the perceived pitch could be missing
- Being a psychoacoustic variable, it depends on subjective sensitivity (e.g. perfect pitch)
- The pitch of a tone allows us to identify the note on a musical scale
- It is generally possible to discriminate btw 1400 different levels of pitch, among which 120 are commonly used in occidental scales

Pitch

- The Just Noticeable Difference (JND) is the smallest variation of frequency that we can perceive. It depends on sound intensity and frequency range
- The ear is maximally sensitive in the 1-4KHz range (speech) and this ability increases with sound intensity
- Lowest perceivable frequency: 20-30 Hz
- Highest perceivable frequency: 15-20 kHz (depends on how long an individual was exposed to noises and loud sounds)
  - Pure tone
  - Complex tone (triangular waveshape) at the same frequency
  - 1 kHz repeated tone that initially lasts 40 ms and then it gets gradually shorted, down to 2 ms (we lose the ability to discriminate its pitch)
  - Cello tone (harmonic spectrum)
  - Gamelan tone (inharmonic tone)

Δf / f (normalized JND) as a function of f.
The No. On the curve denotes the level above hearing threshold



POLITECNICO MILANO 1863

Pitch

A simplified model for the JND

- For medium-intensity sinusoids:
  - At least 3Hz for f< 600 Hz
  - About 0.5% for f> 600 Hz = about 2.5 mel
- If we compare complex sounds, we obtain rather different results (up to one tenth)

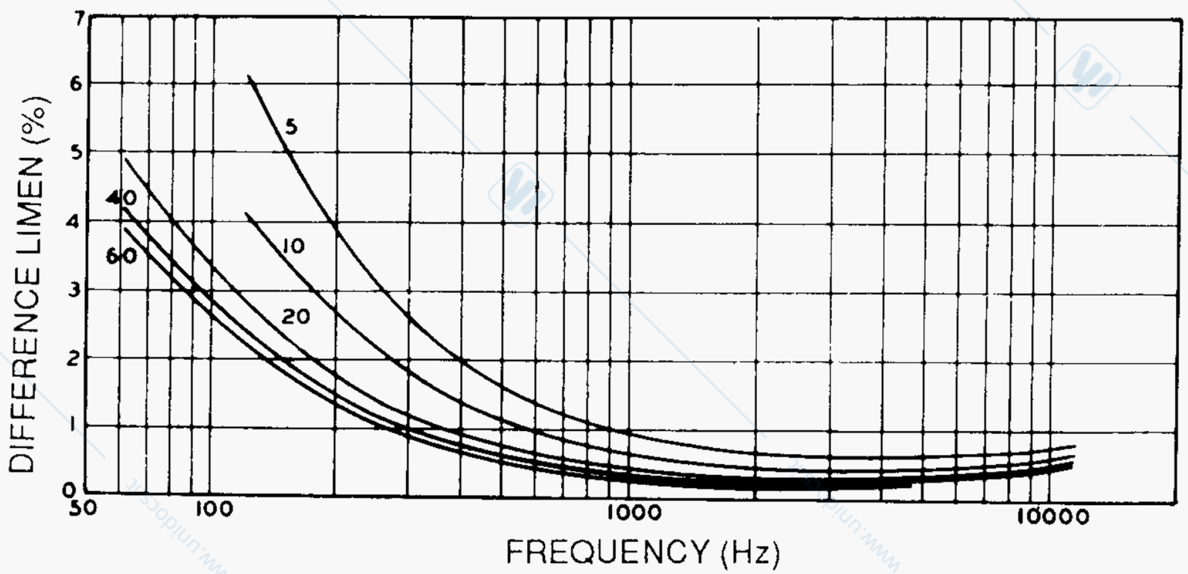


Figure 3.7 Smallest perceptible frequency variation at sound intensities shown according to measurements of Shower and Biddulph. (From Olson, 1968.)

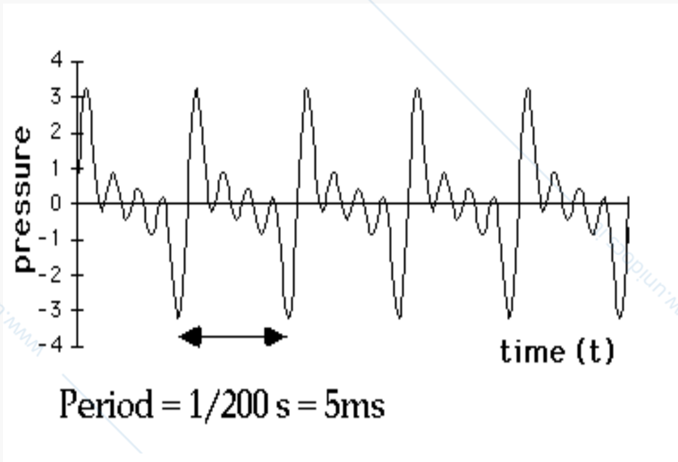
POLITECNICO MILANO 1863

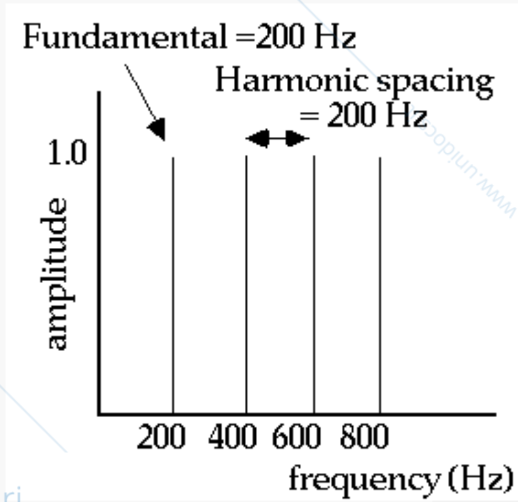
M.S. Music and Acoustic Engineering

Pitch

- **Structure**. Almost all sounds that give a sensation of pitch are periodic. Their spectrum consists of harmonics that are integer multiples of the fundamental.
  The pitch of a complex periodic tone is close to the pitch of a sine wave at the fundamental. Helmholtz claimed that the pitch is heard at the fundamental since the fundamental frequency gives the lowest frequency peak on the basilar membrane.

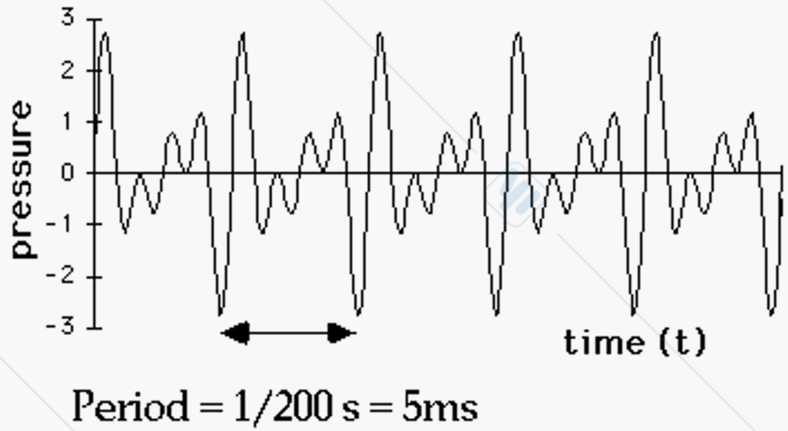
Sound (the fundamental frequency of 200 Hz is present)

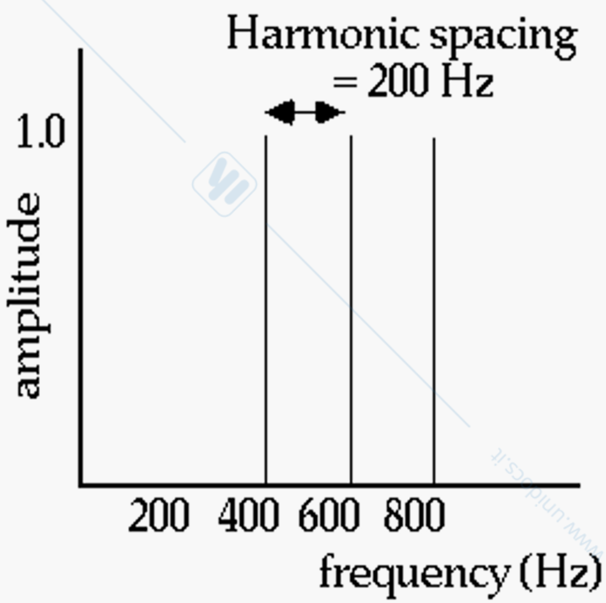




Missing fundamental

- Seebeck (and later Schouten) showed that complex periodic sounds with NO energy at the fundamental may still give a clear pitch sensation at the fundamental (cf telephone speech - the telephone acts as a high-pass filter, removing energy below about 300 Hz).
- sound which has NO fundamental
- previous sound WITH the fundamental

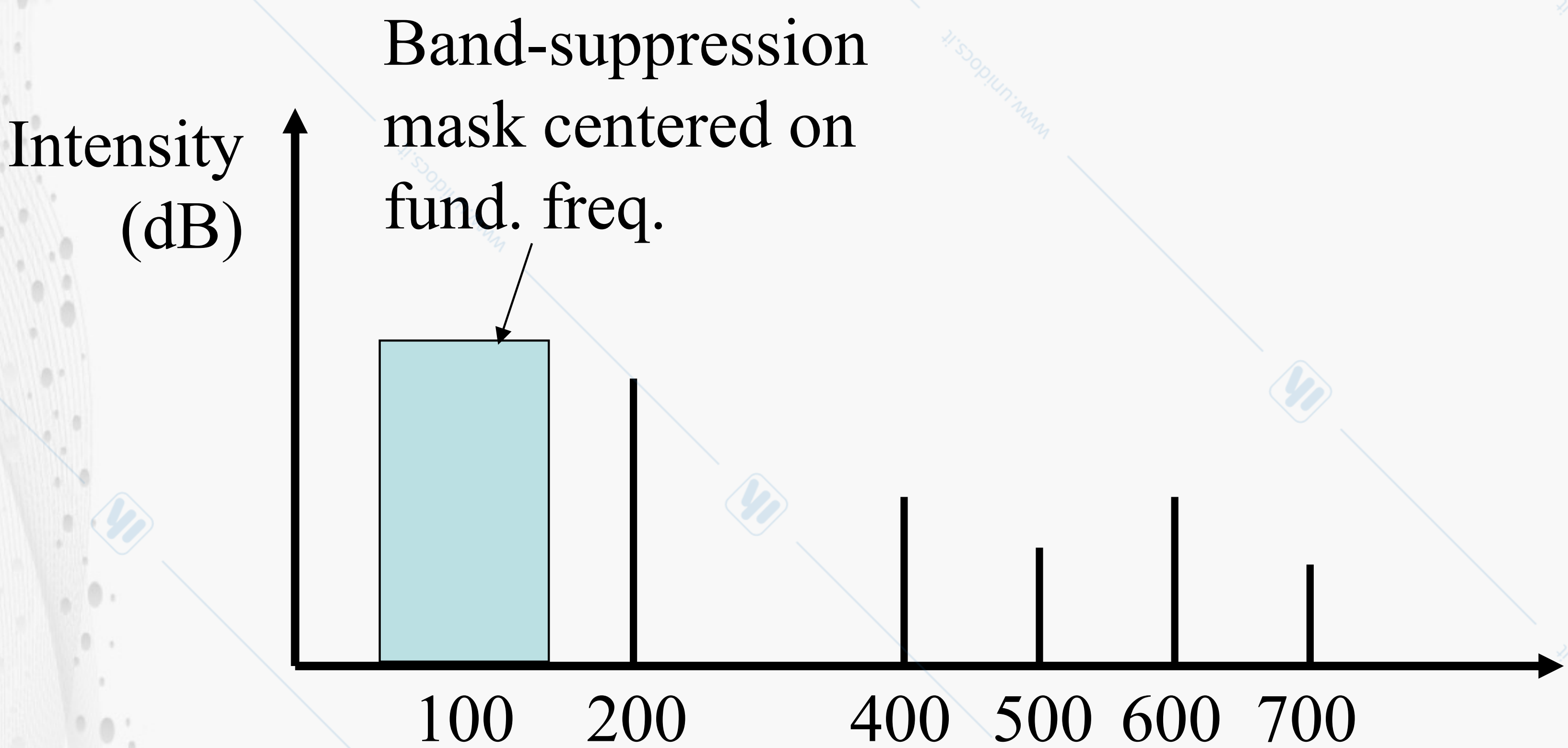




POLITECNICO
MILANO 1863

Is the fundamental really missing?

- Effect of masking on "virtual" pitch



Masking does not eliminate the pitch!

POLITECNICO MILANO 1863

(*)

Helmholtz's place theory

- Helmholtz suggested that the ear reintroduces energy at the fundamental frequency through a process of distortion that produces energy at frequencies corresponding to the difference between two components that are present in the spectrum (i.e. at the harmonic spacing). Any pair of adjacent harmonics would generate energy at the fundamental
- Helmholtz's explanation is wrong because:
  – a pitch at the fundamental is still heard in lowpass filtered masking noise that heavily masks its frequency
  – a complex sound consisting of enharmonic frequencies (eg 807, 1007, 1207) gives a pitch that is slightly higher than the difference of 200.
  – the distortion only occurs at high intensities but low intensities still give the pitch.

POLITECNICO MILANO 1863

(*)

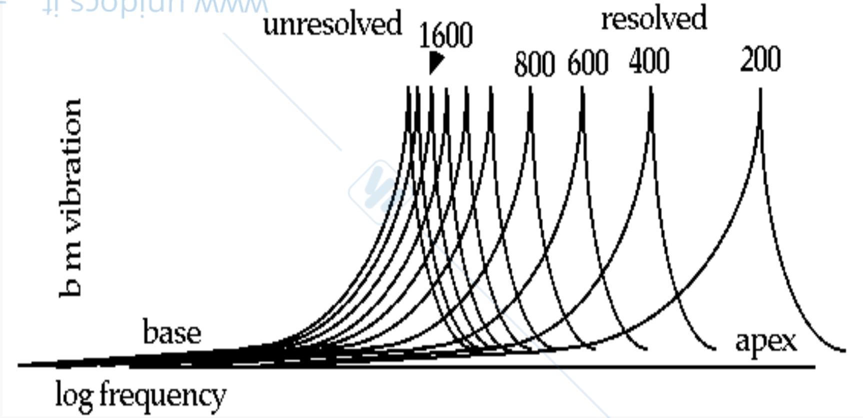
Schouten's timing theory

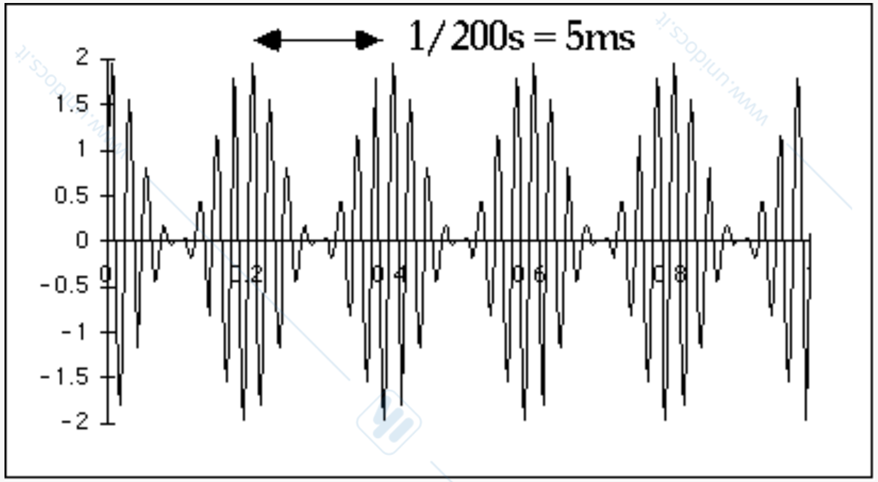
- Schouten proposed that the brain times the intervals between beats of the unresolved (see next diagram) harmonics of a complex sound, in order to find the pitch
- Schouten's theory is wrong because:
  - pitch is determined more by the *resolved* than by the *unresolved* harmonics
  - you can still hear a pitch corresponding to the fundamental when the two consecutive frequency components go to opposite ears

POLITECNICO MILANO 1863

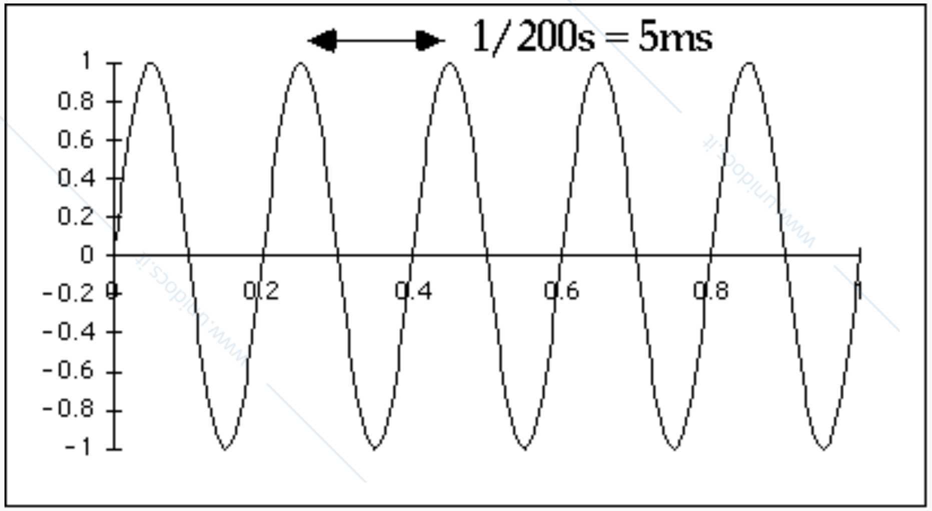
Pitch

The following diagram shows the excitation pattern that would be produced on the basilar membrane separately by individual harmonics of a 200 Hz fundamental. Notice that the excitation patterns of the higher numbered harmonics are closer together than those of the low-numbered harmonics. This is because the filters have a bandwidth which is roughly a tenth of their center frequency (and so is constant on a log scale), whereas harmonics are equally spaced in frequency on a linear scale. More harmonics then get into a high-frequency filter than into a low-frequency one. The low-numbered harmonics are resolved by the basilar membrane (giving roughly sinusoidal output in their filters); but the high-numbered harmonics are not resolved. They add together in their filters to give a complex vibration which shows *beats* at the fundamental frequency





Output of 1600 Hz filter



Output of 200 Hz filter

POLITECNICO MILANO 1863

M.S. Music and Acoustic Engineering

(*)

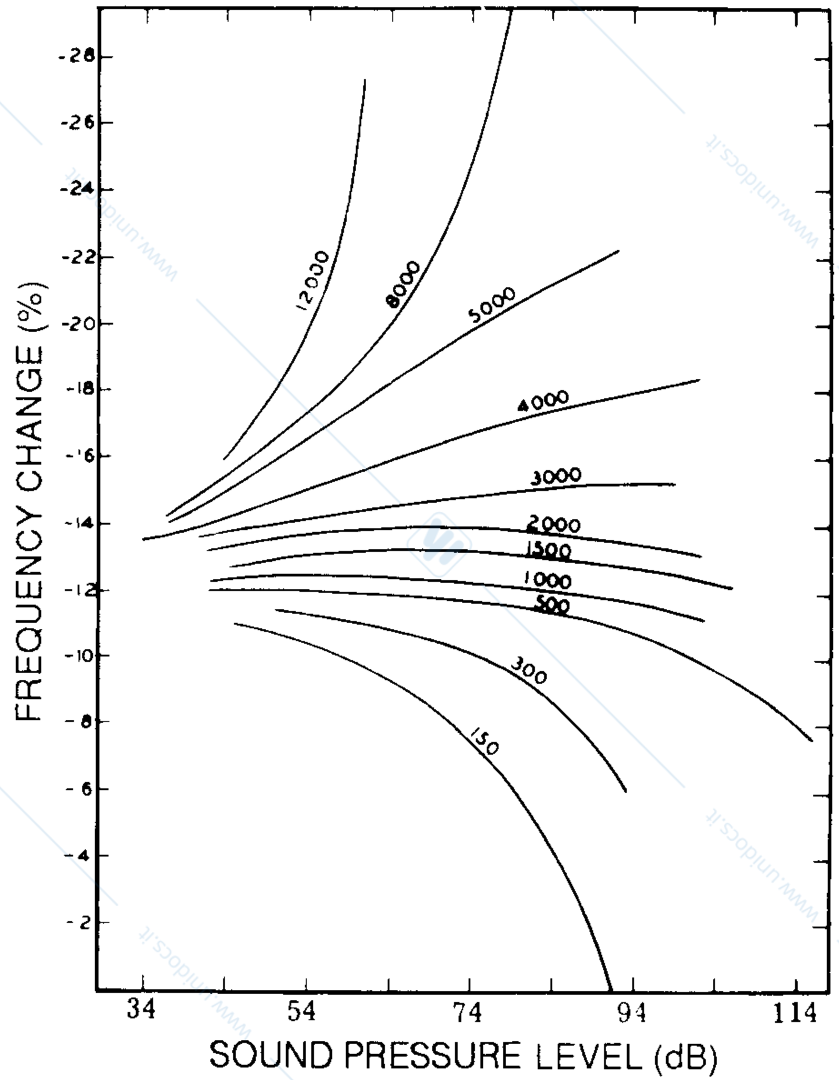
Pattern recognition theories

- Goldstein's theory states that pitch is determined by a pattern recognition process on the *resolved* harmonics from both ears. The brain finds the best-fitting harmonic series to the resolved frequencies, and takes its fundamental as the pitch.
- Goldstein's theory accounts well for most of the data, but there is also a weak pitch sensation from periodic sounds which do not contain any resolvable harmonics or from aperiodic sounds that have a regular envelope (such as amplitude modulated noise). A theory such as Schouten's may be needed in addition to Goldstein's in order to account for such effects.
- Evidence for there being two separate mechanisms for resolved and unresolved harmonics is:
  - pitch discrimination and musical pitch labeling (eg A#) is much worse for sounds consisting of only unresolved harmonics;
  - comparison of pitches between two sounds one having resolved and the other unresolved harmonics is worse than comparison of pitches between two sounds both with unresolved harmonics.

POLITECNICO
MILANO 1863

Pitch and intensity

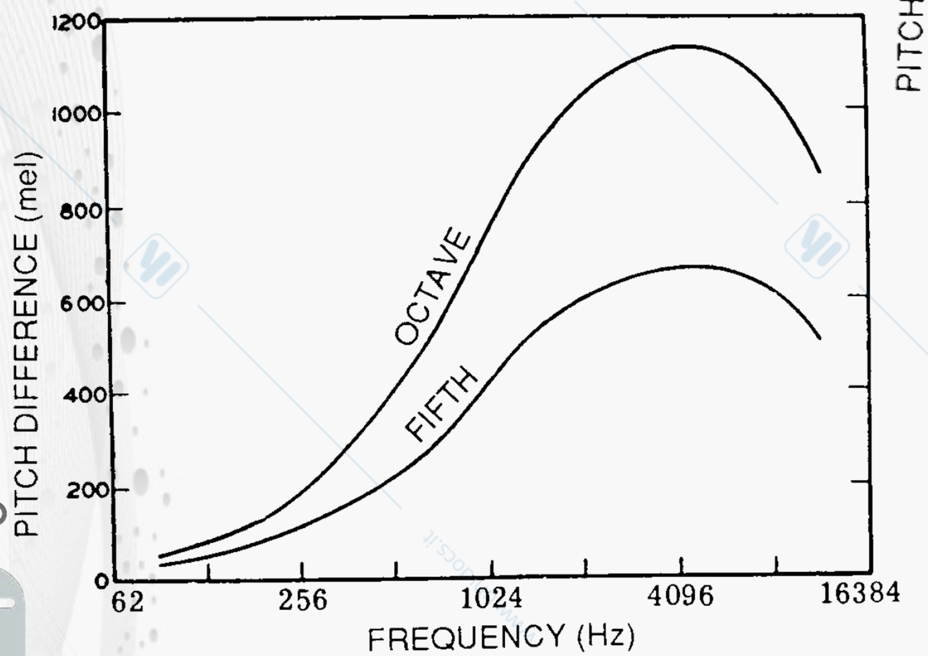
- Pitch as a function of the intensity
  - the pitch of a 150Hz sinusoid goes down of about 12%, (about two half-tones) as the intensity increases from 45dB to 90dB
- A descending curve means that frequency must increase in order to keep the perceived pitch constant

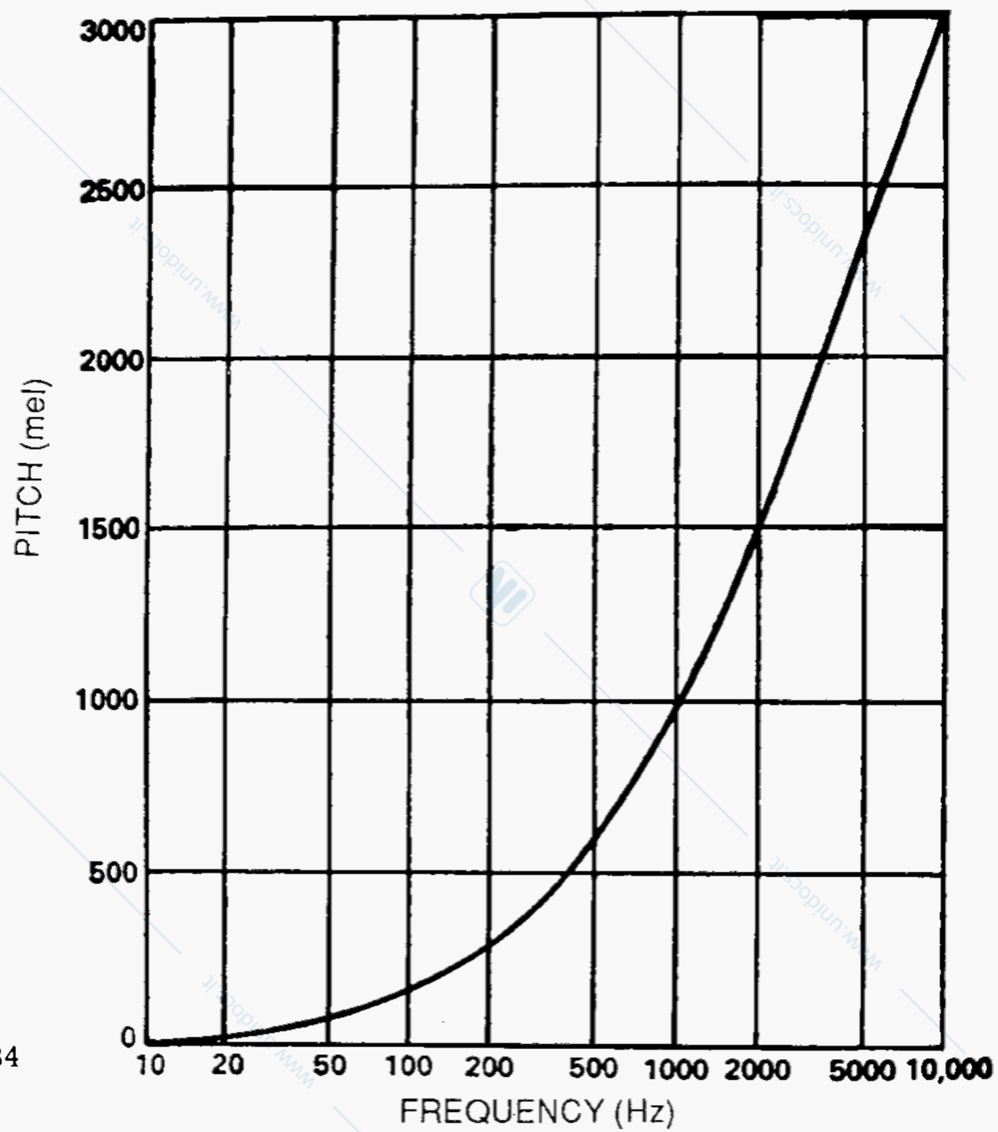


POLITECNICO MILANO 1863

Pitch and frequency

tone height (mel scale)
doubling mels means
doubling perceived pitch





POLITECNICO
MILANO 1863

Categoric perception

(*)

- Intervals
- Rhythmic sequences
- Voiced sounds

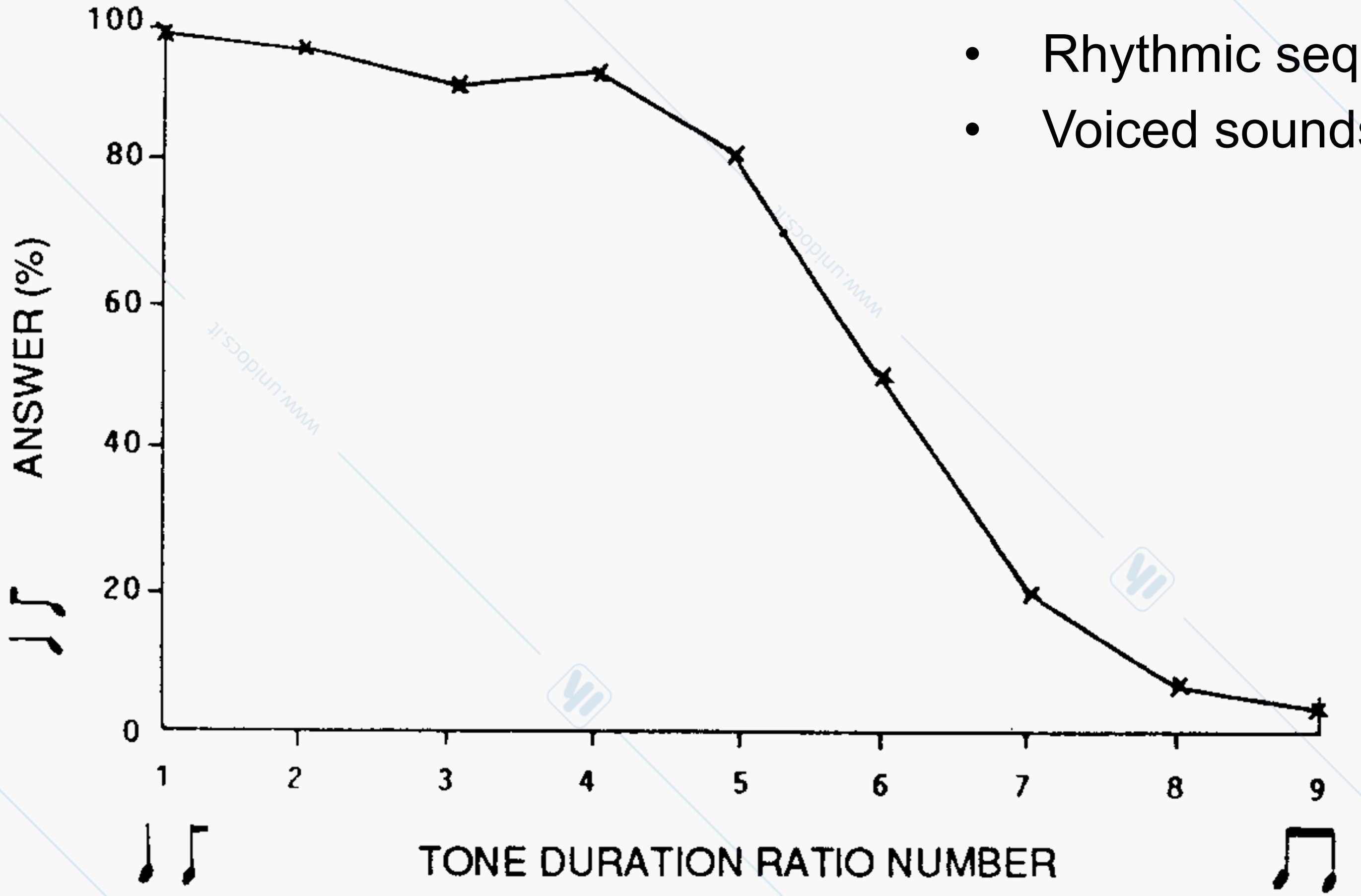


Figure 3.15 Categorization of note durations as function of tone durations in a tone pair. At the left extreme, the long note is 640 msec and the short is 320 msec, and at the right extreme both tones are 480 msec. Each step to the right on the scale corresponds to a decrease of the long tone and an increase of the short tone by 20 msec. Mostly, this small change of the durations does not cause much of a shift of categorization. However, the change from 540 + 420 to 520 + 440 msec (number 6 and 7 on the scale) produces a great effect on categorization. (From Clarke, 1987.)

POLITECNICO MILANO 1863

Critical band

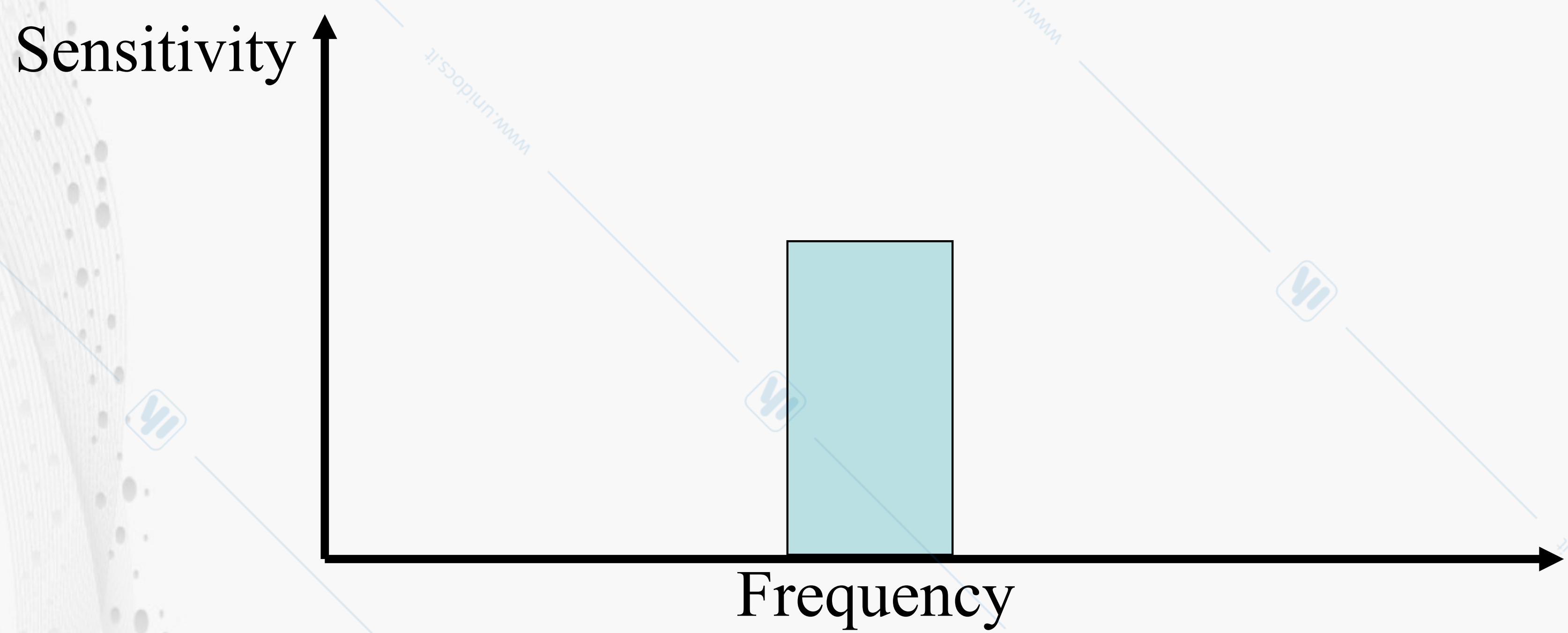
- It is the range of frequencies that activate the same hair cell
- It is the range of resonance frequencies that correspond to the hair cells of the same portion of the basilar membrane that responds to the same frequency
  - It covers about 1.2 mm of basilar membrane (about 1300 hair cells)
- It describes the ear's ability to separate *simultaneous* sounds
  - Not to be confused with the JND of the pitch
  - It does not depend on intensity
- 24 critical bands of about 1/3 octave cover the whole audible spectrum

Two sinusoids, one fixed at 400 Hz, the other going from 400 to 510Hz, which is the limit of the critical band (separation)

www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari

The critical band

- The width of an internal auditory filter

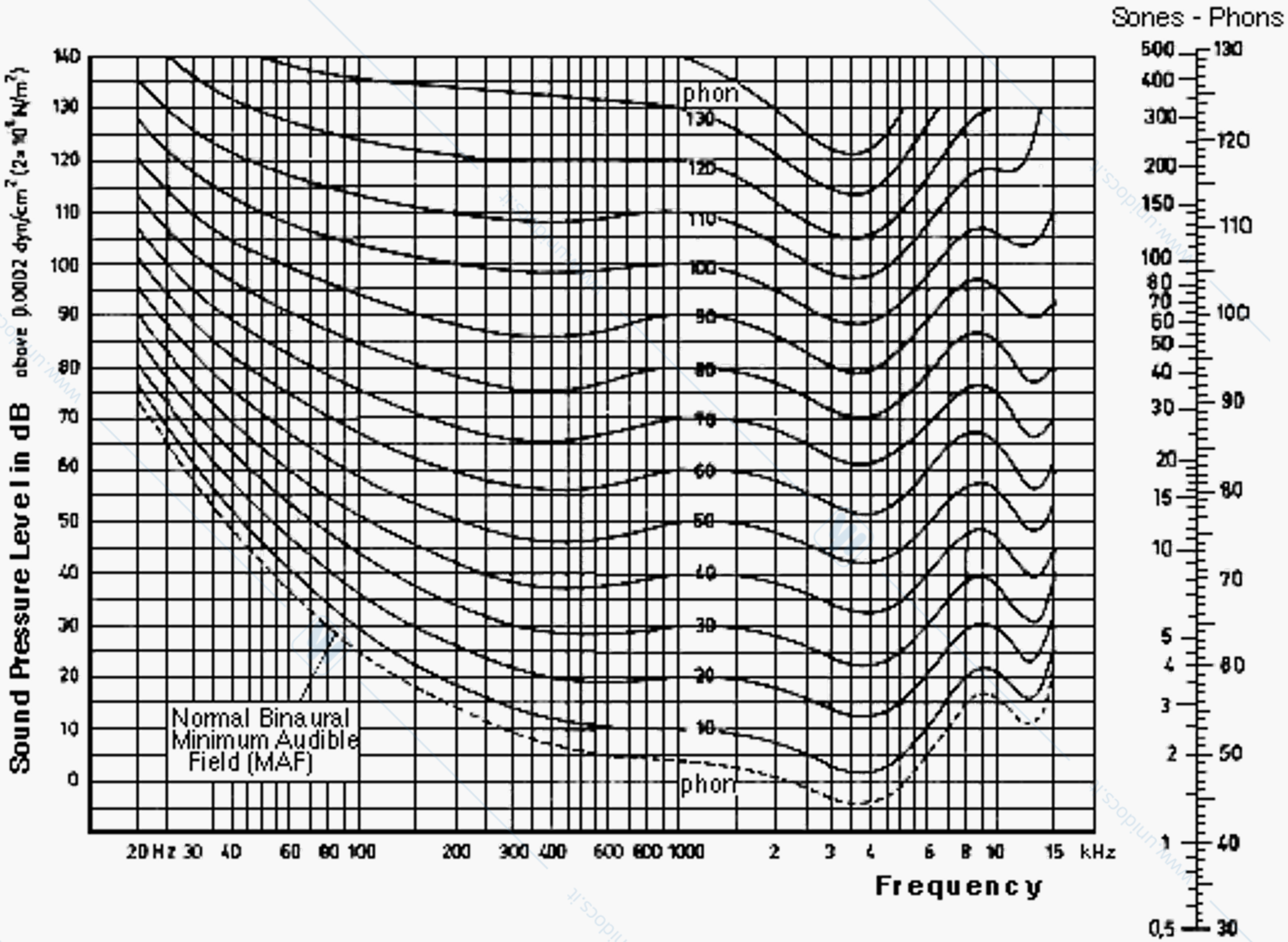


Bandwidth of frequency-specific auditory filter

POLITECNICO MILANO 1863

Acoustic intensity measurements

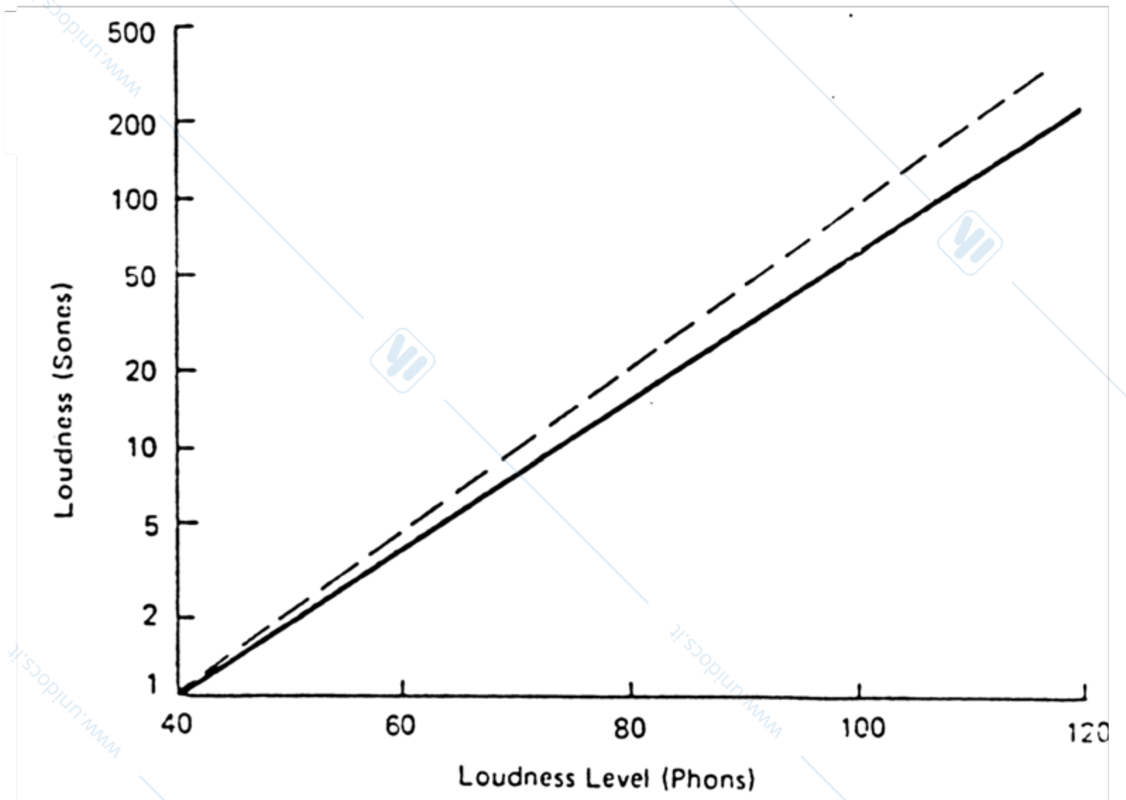
- Loudness = subjective impression of sound intensity
- Pure sounds, with the same SPL and different frequencies produce different sensations of intensity
- Fletcher and Munsen's *isophonic* curves
- Phon (Loudness Level)
  - suono at x Phon = isophonic curve at x SPL (1000 Hz)



POLITECNICO MILANO 1863

Acoustic intensity measurements

- Phone scale is not a perceptive easurement (doubling phones does not correspond to doubling perceived intensity
- Son: every 10 Son perceived intensity doubles



POLITECNICO MILANO 1863

Intensity thresholds

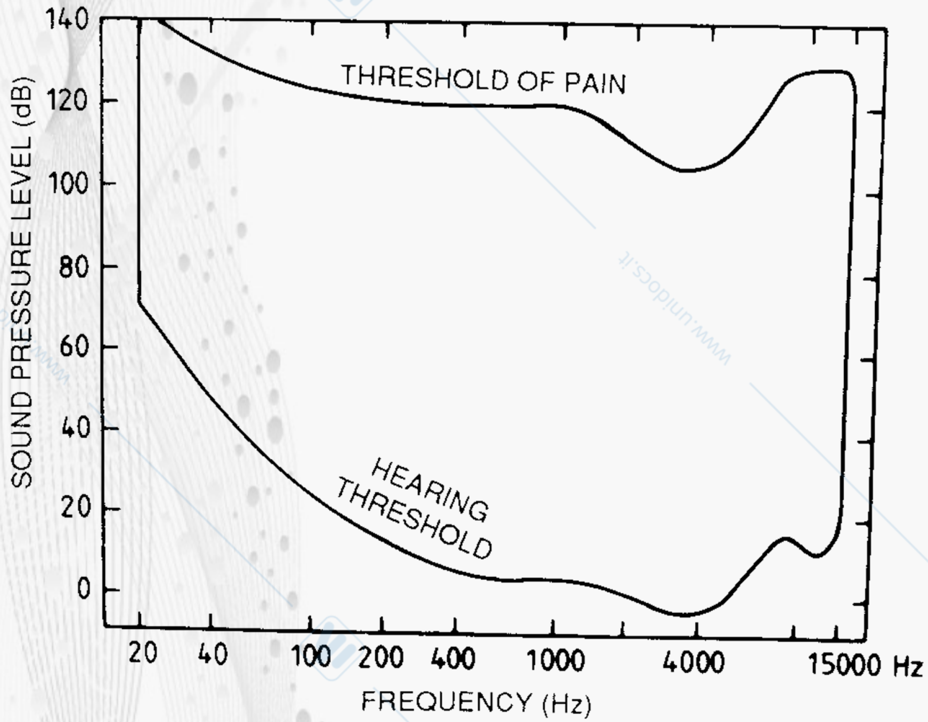


Figure 3.8 Audible frequency and sound level range.

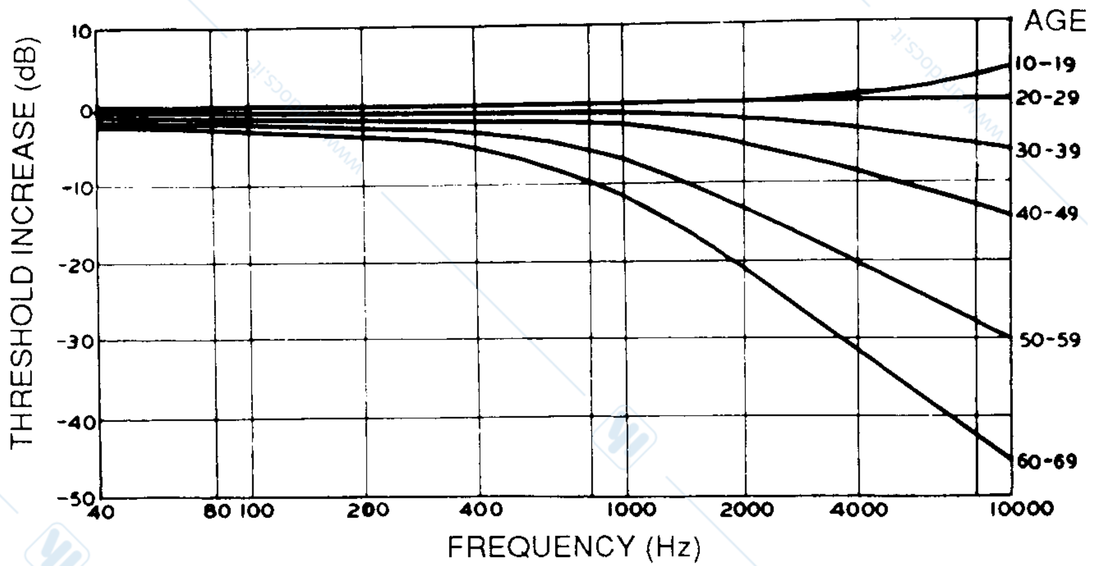


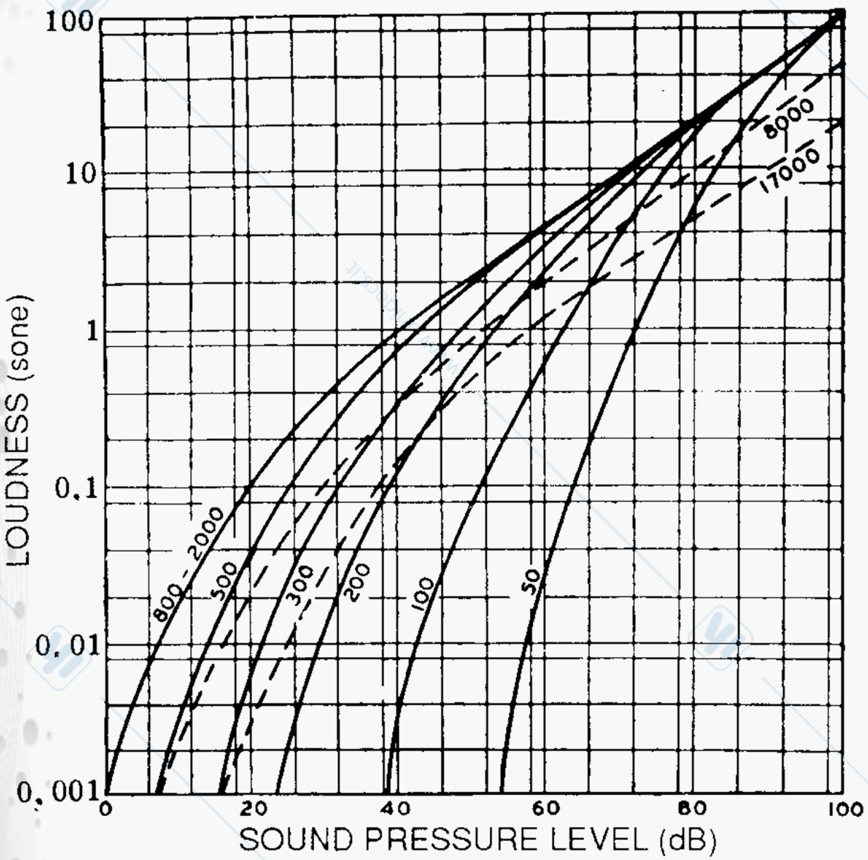
Figure 3.10 Dependence of the hearing threshold on age. (From Olson, 1968.)

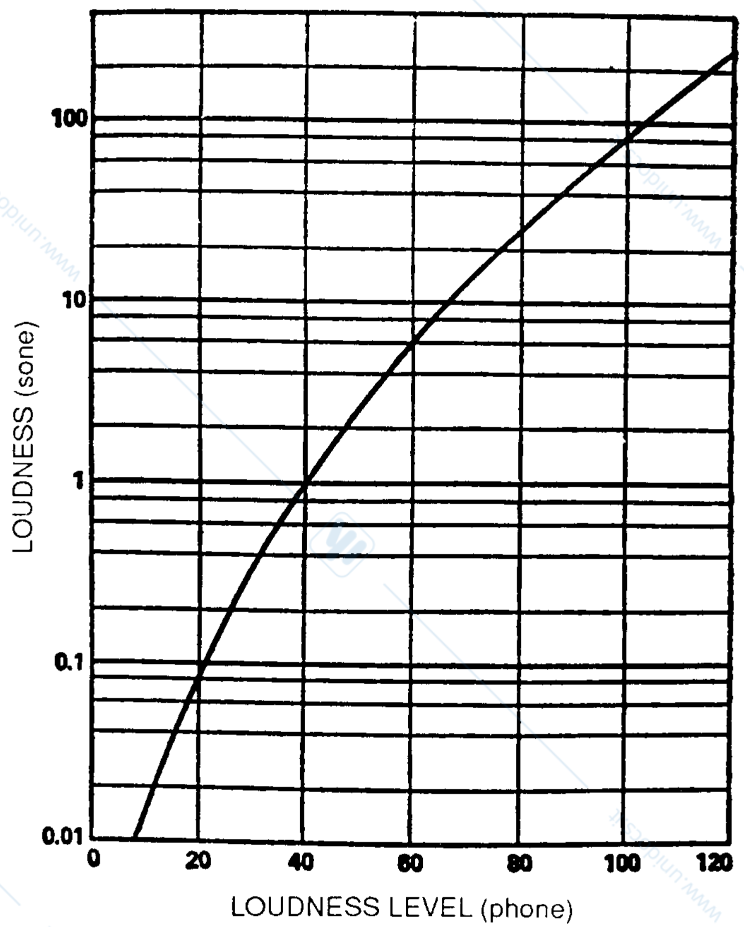
POLITECNICO
MILANO 1863

www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari
www.unidocs.it - Appunti e dispense per superare i tuoi esami universitari

Loudness in sones

- Half sones correspond to half loudness





- Equal Loudness Examples
- Which sounds Twice Louder? 5 or 10 dB
- Complex Sounds composed of partials

POLITECNICO MILANO 1863

M.S. Music and Acoustic Engineering

Threshold Shift

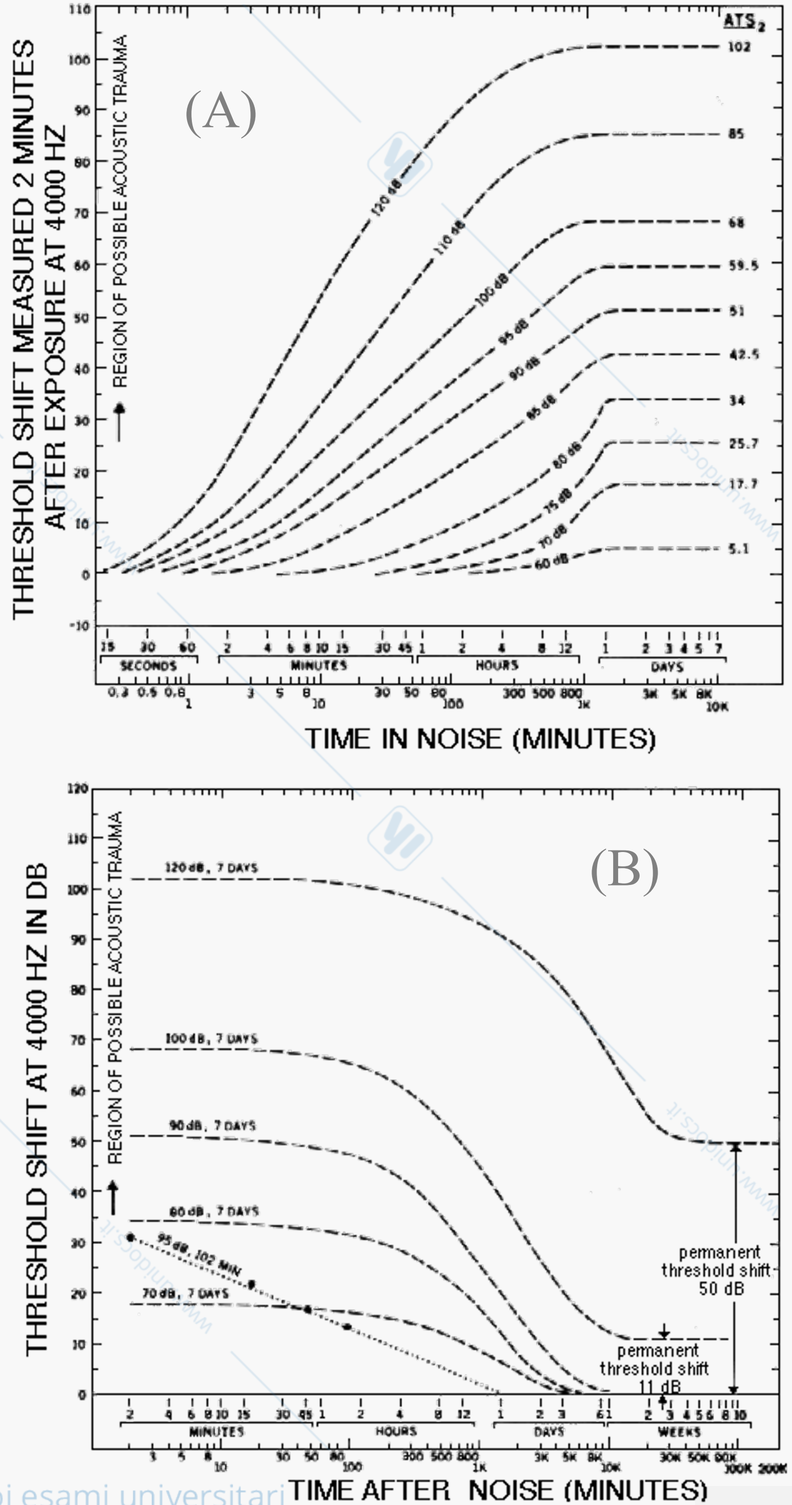
- Environmental noise causes hearing loss (hearing threshold alteration)
- Depends on age and gender
  - Males are most sensitive at low frequencies
  - Females are most sensitive at high frequencies
- The most damaging noises are between 2 and 4 kHz.
- Causes
  - The noise causes blood vessels to shrink, which reduces blood intake to hair cells
  - This causes temporary hearing loss
  - Similar to what happens when going from a brightly lit room to a dimly lit room

POLITECNICO
MILANO 1863

M.S. Music and Acoustic Engineering

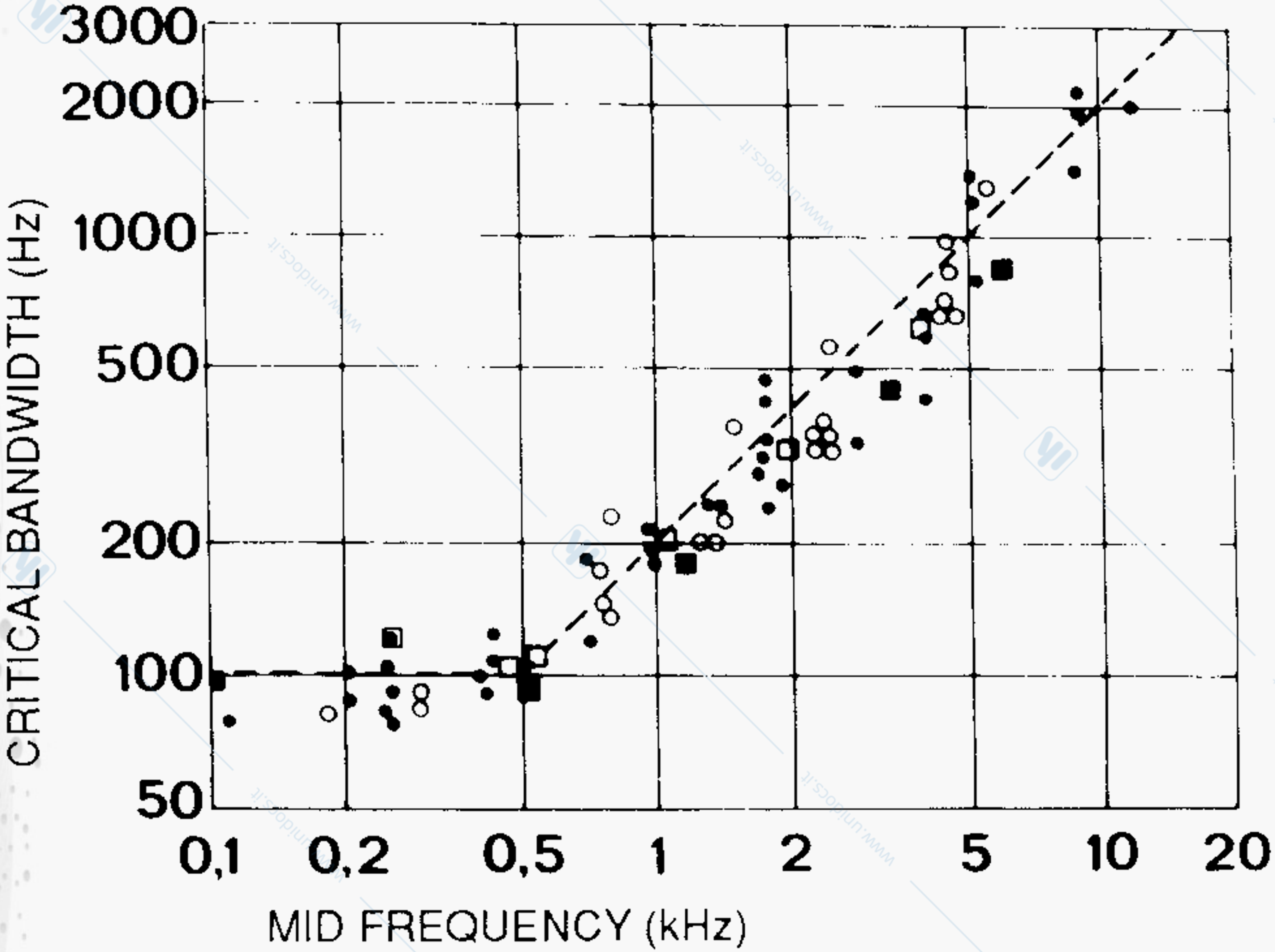
Threshold Shift

- *Temporary Threshold Shift (TTS)*
  - From 24 hours to a week
- *Permanent Threshold Shift (PTS)*
  - After being exposed to the noise for a long time
- *Chronic Threshold Shift*
  - When frequently exposed to noise without having time to recuperate, TS can turn into PTS (jack-hammer operators, etc.)



- Damage assessment (A) and recovery time (B) after exposure to a 4 kHz noise
- Curves are estimated for young healthy males, based on the results reported in Miller, "Effects of Noise on People", *Journal of the Acoustical Society of America*, vol. 56, 1974, p. 734-5]

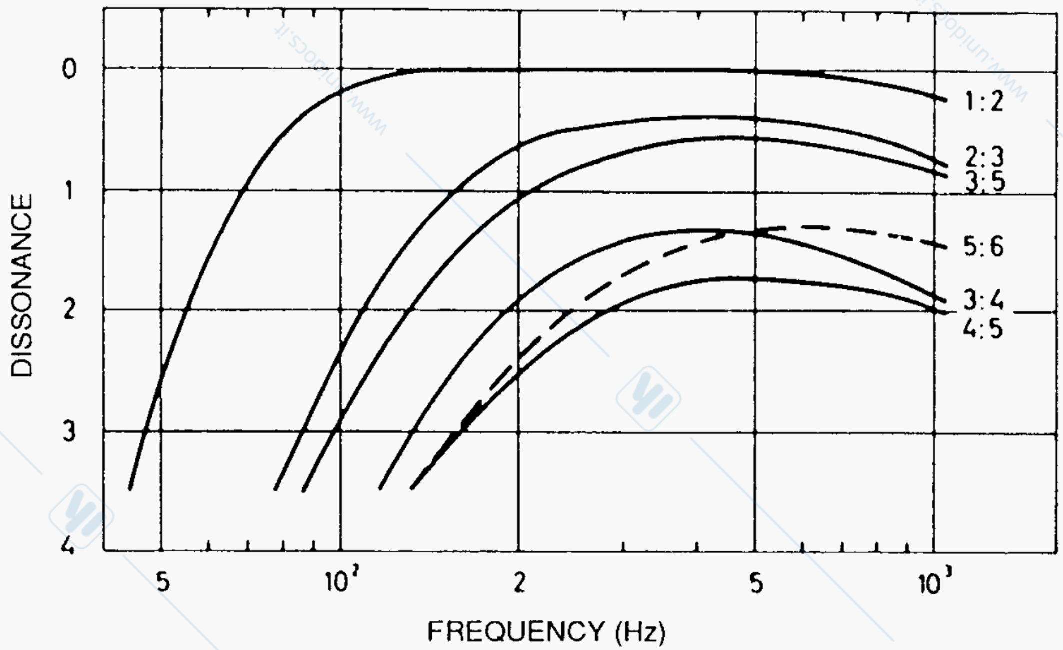
Critical bands



POLITECNICO
MILANO 1863

Consonance

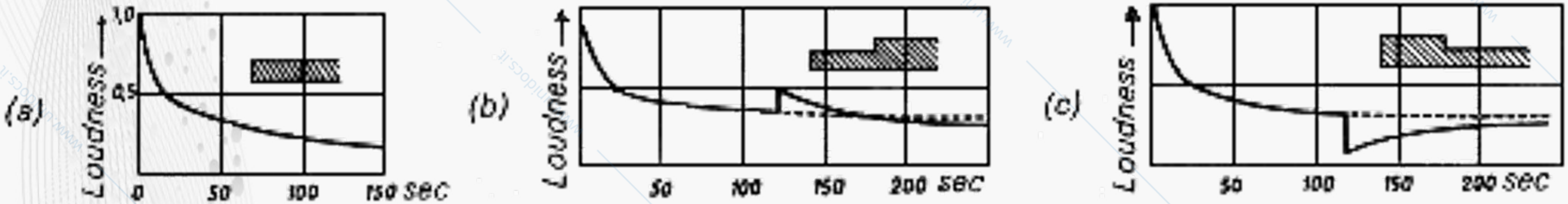
- As a function of average frequency
- Dissonance increases at lower frequencies
  - Due to critical bands



Degree of consonance of simultaneous sounds

Adaptation

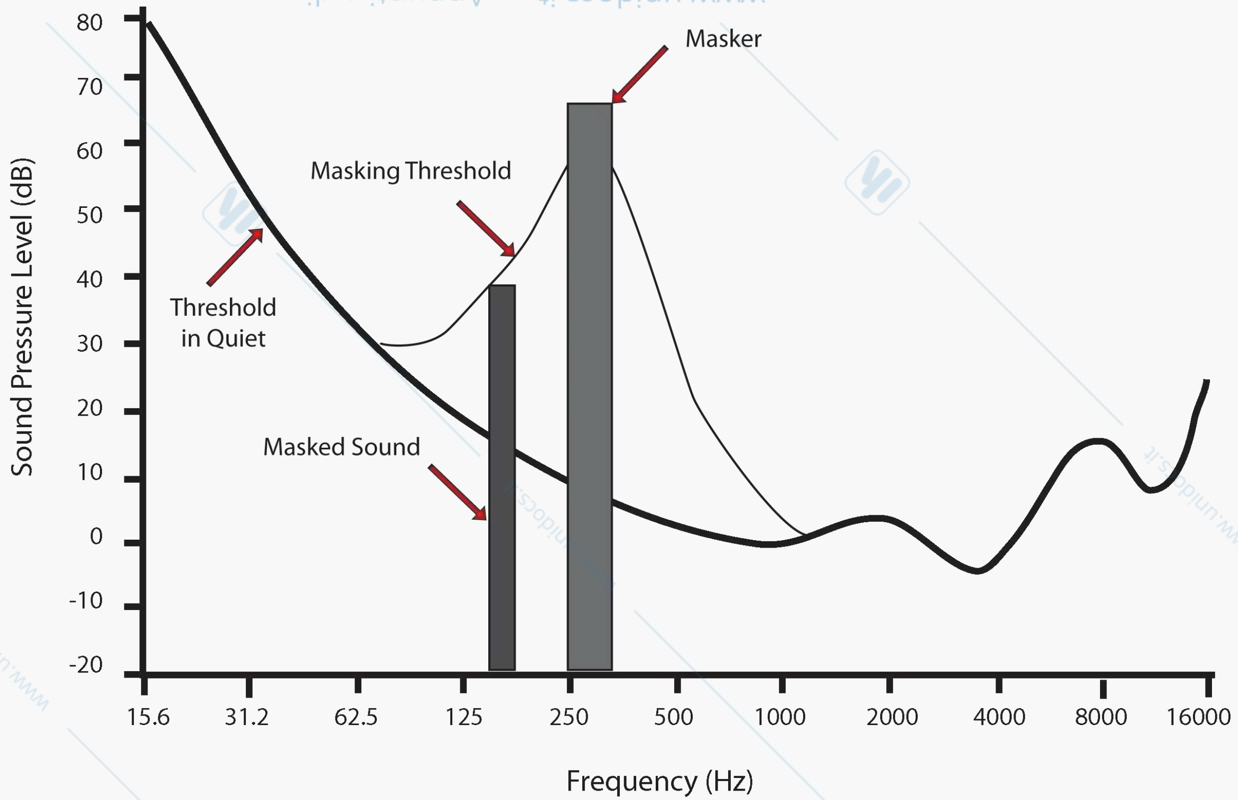
- Sensitivity loss due to the persistent exposure to a sound
  - The sound seems to gradually dim down
  - Due to neural cell's state of fatigue



- Perceived loudness reduction due to exposure to a constant tone (sinusoid), measured in a 3 minute interval [F.Winckel, *Music, Sound and Sensation,* Dover, 1967, p. 104]
  - (A): constant SPL
  - (B): SPL doubles after 2 minutes
  - (C): SPL halves after 2 minutes

POLITECNICO
MILANO 1863
M.S. Music and Acoustic Engineering

Frequency masking



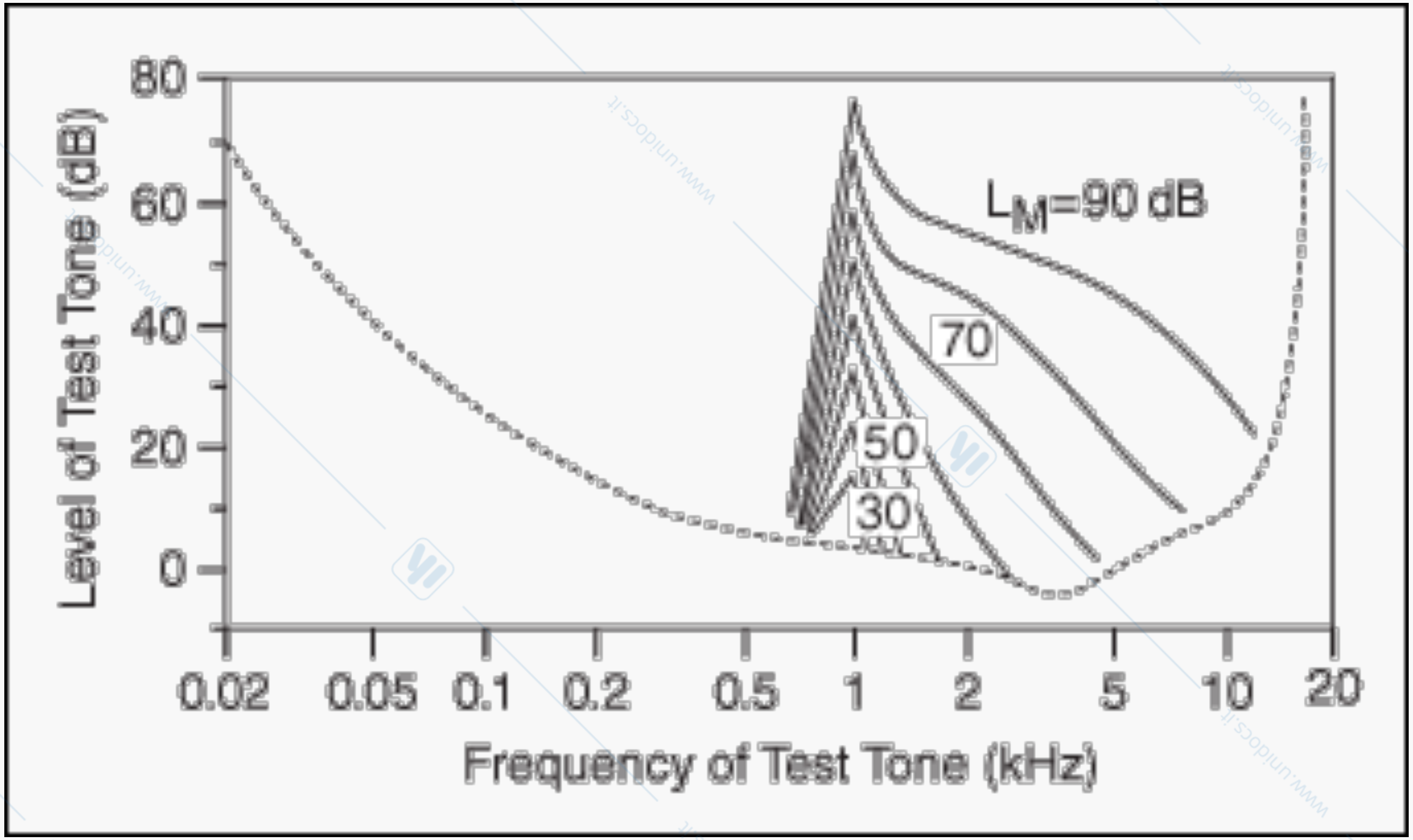
- Generate a 1 Khz masking tone at 60dB and a test tone at a different frequency (e.g. 1.1Khz) and increase the volume until test tone becomes audible
- Change the test tone frequency and repeat the experiment
- Plot the result
- The louder tone masks the test tone at lower frequencies in the close spectral proximity, but it masks the test tone even more at higher frequencies

POLITECNICO MILANO 1863

M.S. Music and Acoustic Engineering

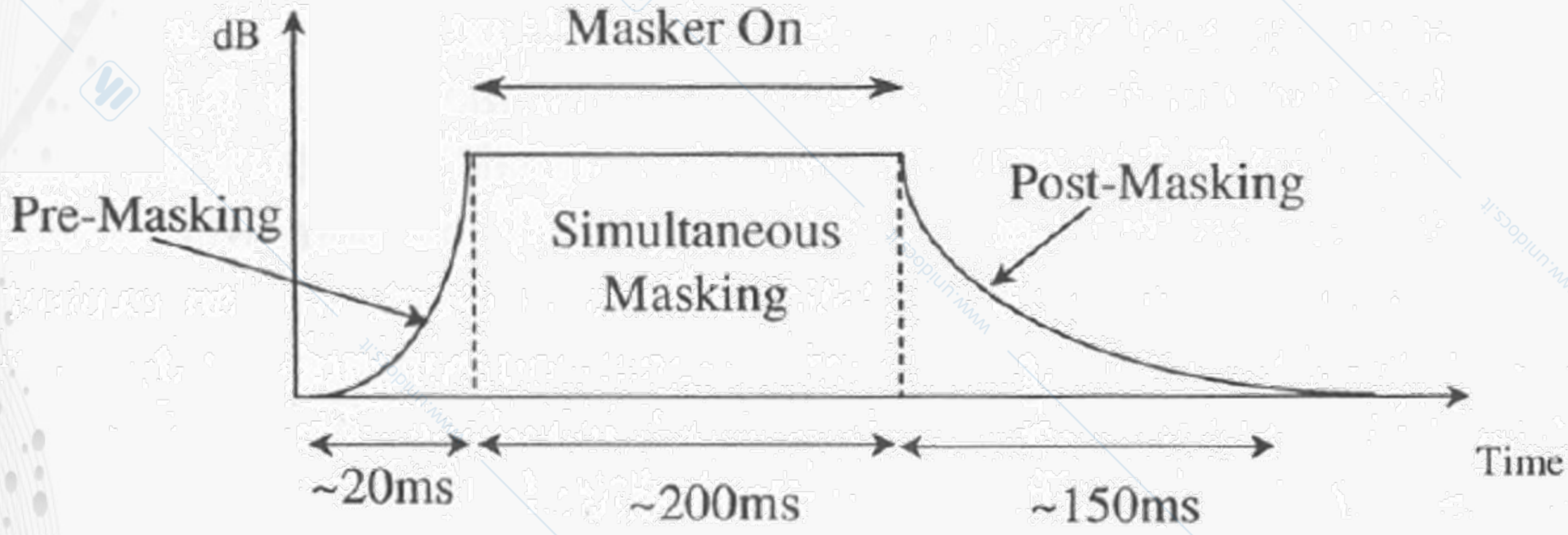
Frequency masking

- If we repeat the experiment with masking tones at various levels of loudness we obtain:



POLITECNICO MILANO 1863

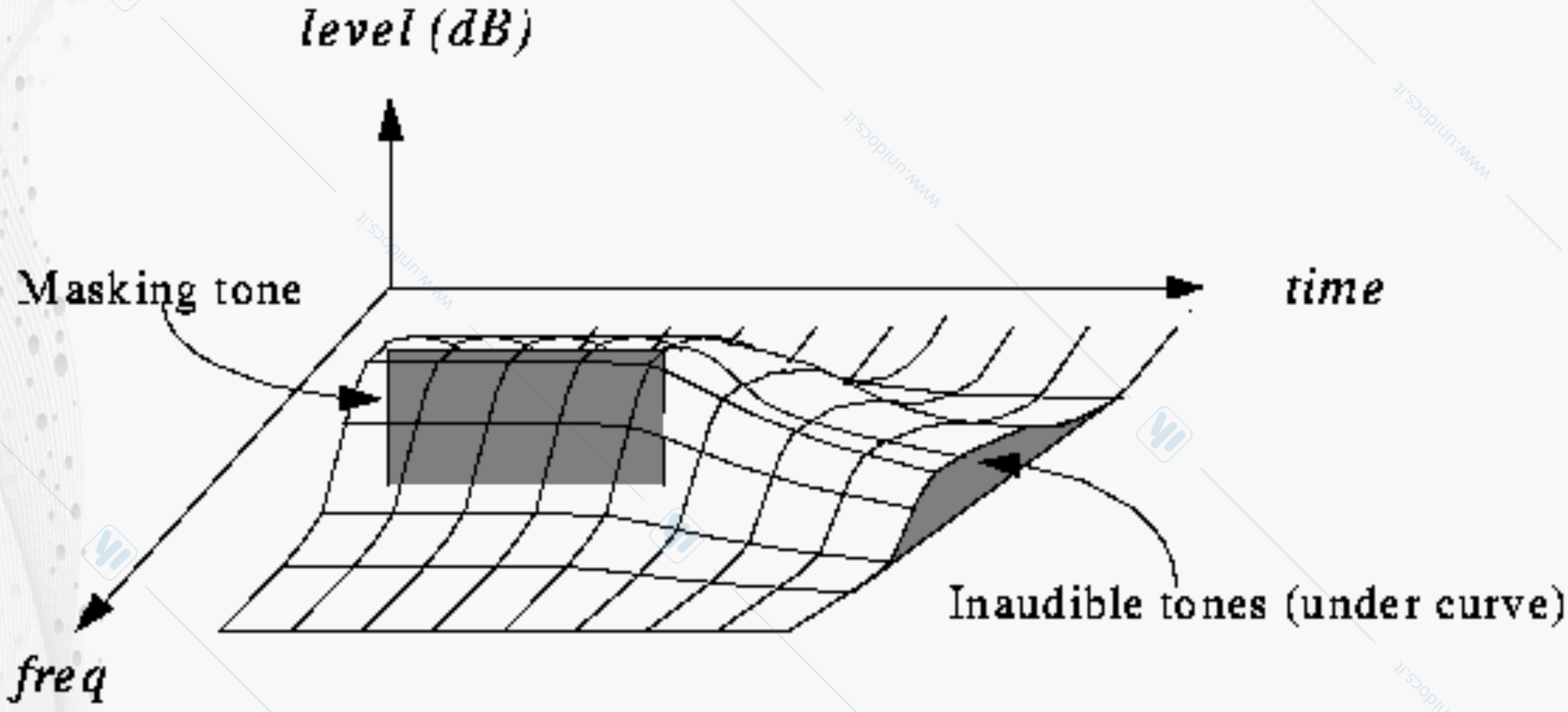
Temporal masking



- After having been exposed to a louder sound, the ear needs some time to be able to perceive a softer sound (of similar frequency)
  - Experiment:
    - generate a 1 Khz masking tone at 60 dB together with a 1.1Khz, 40dB test tone
    - Interrupt first the masking tone and then the test tone
  - The louder tone will mask softer adjacent tones within a certain variable time threshold

POLITECNICO MILANO 1863

Time-frequency masking



POLITECNICO
MILANO 1863

Masking

- Higher frequency tone masks lower frequency tone
  - 500 Hz at 0dB with –40dB tones at 300, 320, 340, 360, 380, 400, 420, 44, 460, 480 Hz
- Lower frequency tone masks higher frequency tone
  - 500 Hz at 0dB with –40dB tones at 1700, 1580, 1340, 1200, 980, 860, 740, 620 Hz

POLITECNICO MILANO 1863

Auditory Scene Analysis

References

- A.S. Bregman, P. Ahad, «Demonstrations of Auditory Scene Analysis: The Perceptual Organization of Sound»
- Bragman's auditory scene analysis website, including commented audio examples
  - http://webpages.mcgill.ca/staff/Group2/abregm1/web/downloadstoc.htm

POLITECNICO MILANO 1863

M.S. Music and Acoustic Engineering

Music theory advanced - https://youtu.be/K1zor-rq--A

Auditory Scene Analysis

- **The brain needs to group...**
  - **simultaneously**, i.e. separating out which frequency components that are present at a particular time have come from the same sound source
  - **successively**, i.e. deciding which group of components at one time is a continuation of a previous group
- Auditory streaming is the formation of **perceptually distinct apparent sound sources**
- Temporal order judgement is good *within* a stream but bad *between* streams
- Examples
  - implied polyphony
  - noise burst replacing a consonant in a sentence
  - click superimposed on a sentence or melody

Grouping Principles

- **Proximity**
  - Tones **close in frequency** will group together, so as to minimize the extent of frequency jumps and the number of streams
  - Tones with **similar timbre** will tend to group together
  - Speech sounds of **similar pitch** will tend to be heard from the same speaker
  - Sounds from different locations are harder to group together across time than those from the **same location**
- **Common fate**
  - Sounds from a common source tend to start and stop at the same time and change in amplitude or frequency together
    - Vibrato
    - Envelope
  - A single component is easy to hear out if it is the only one to change in a complex

POLITECNICO MILANO 1863

Grouping Principles

- **Good continuation**
  - Abrupt discontinuities in frequency or pitch, can give the impression of a different sound source
- **Continuity Effect**
  - Sound that is interrupted by a noise that masks it, can appear to be continuous. Alternations of sound and mask can give the illusion of continuity with the auditory system interpolating across the mask

POLITECNICO
MILANO 1863

Auditory scene analysis (31-37)

(*)

- Grouping by timbral similarity
- Bregman Streaming Example
- Yodeling (Franzi Lang)
- Streaming: Interleaved Melodies
- Warren-style Loop
- Wessel Loops
- Interleaved melodies from music

POLITECNICO MILANO 1863

(*)

Effects

- Cocktail party
  - Two speakers, one location
  - Two speakers, two locations
- Precedence effect
  - Same voice on both loudspeakers
  - Same voice, with reversed phase, on both loudspeakers

POLITECNICO MILANO 1863

M.S. Music and Acoustic Engineering

(*)

Timbre

- Depends on
  - Spectral content
  - Transients
  - Relations btw partials
    - Harmonic ratio (frequencies are multiple of the fundamental)
    - Harminic series (up to the16°) synthesized with sinusoids
    - Inharmonic ratio (frequencies are not multiple, e.g.
      - Percussions
      - Chinese Tam Tam
    - Relations btw phases
      - Two sinusoids are in phase if they arrive at the same point of the cycle at the same time
  - Timbre is perceived as a gestaltic impression of the whole sound rather than a function of some of its components

POLITECNICO MILANO 1863