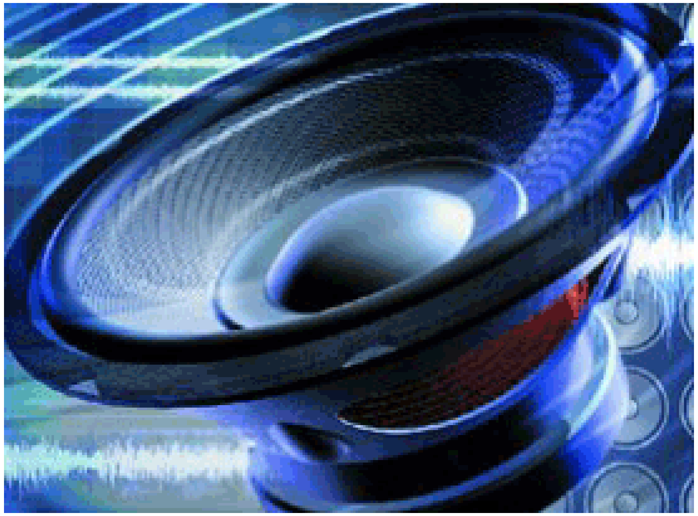
2.2 Auditory demonstrations



This section comprises a certain number of auditory examples which are not embedded in this presentation, for copyright concerns

(ref: *Auditory Demonstrations CD*, ed. Philips)

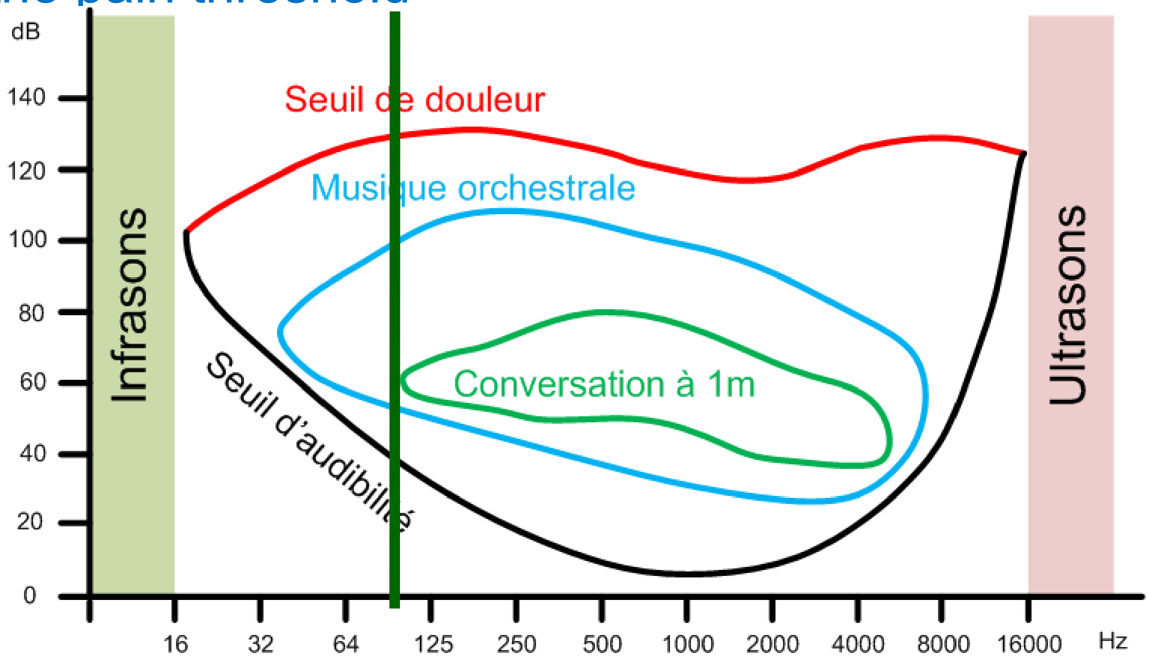
Psychoacoustics

- The goal of psychoacoustics is to establish relationships between auditory sensations and physical characteristics of acoustic stimuli

- Different stimuli (pure tones, combinations of pure tones, random sound, speech…) and listening conditions (headphone, free field, diffuse field, monaural or binaural…) can be used

The auditory area

- Represents the frequency domain of audible sounds
- It is delimited with:
  - the auditory threshold
  - the pain threshold

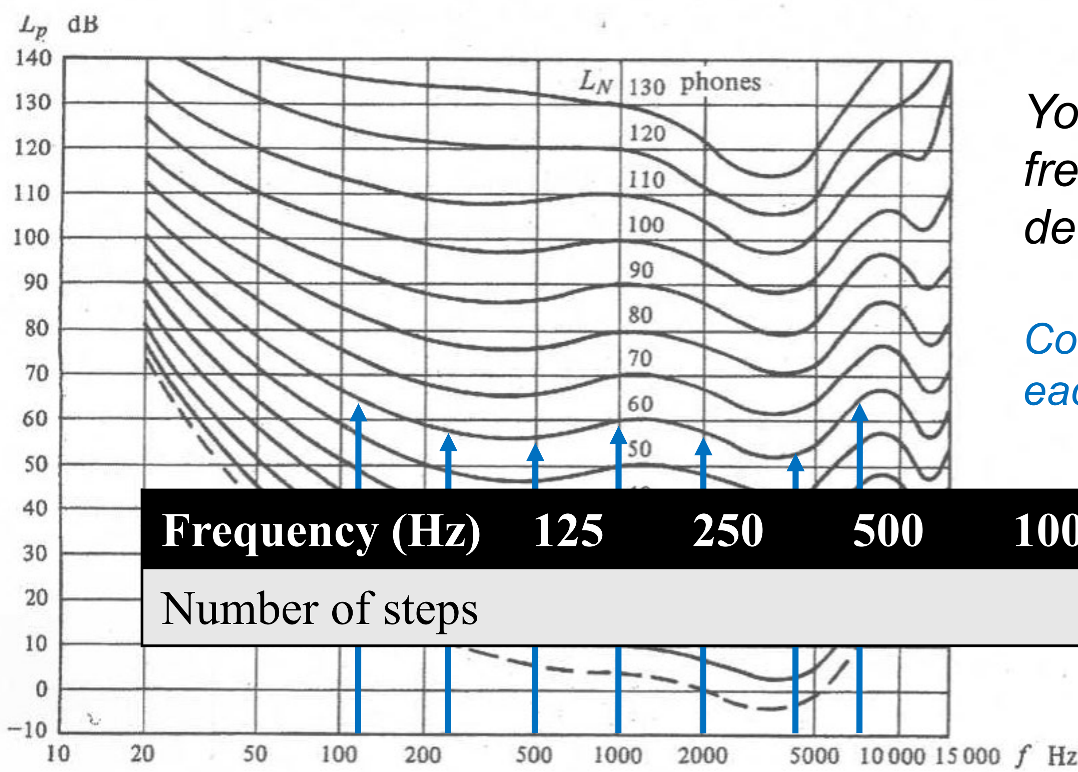


The auditory area

- Normal thresholds are established for mean subjects, aged from 18 to 30 years (« normal » audition)
- Frequencies are comprised between 16 Hz and 16 kHz
- It depends on the age and the medical state of subjects
- The bottom limit is hard to assess (experimental means, infrasound perception via mechanical ways)
- The maximum sensitivity of human ears is between 2 kHz and 4 kHz, where the auditory dynamics is the highest (140 dB)
- Hearing level HL represents the difference between the sound pressure level and the corresponding hearing threshold (for a given subject and a given sound)

LOUDNESS & SOUND LEVEL

Loudness



Frequency (Hz)	125	250	500	1000	2000	4000	8000
Number of steps							

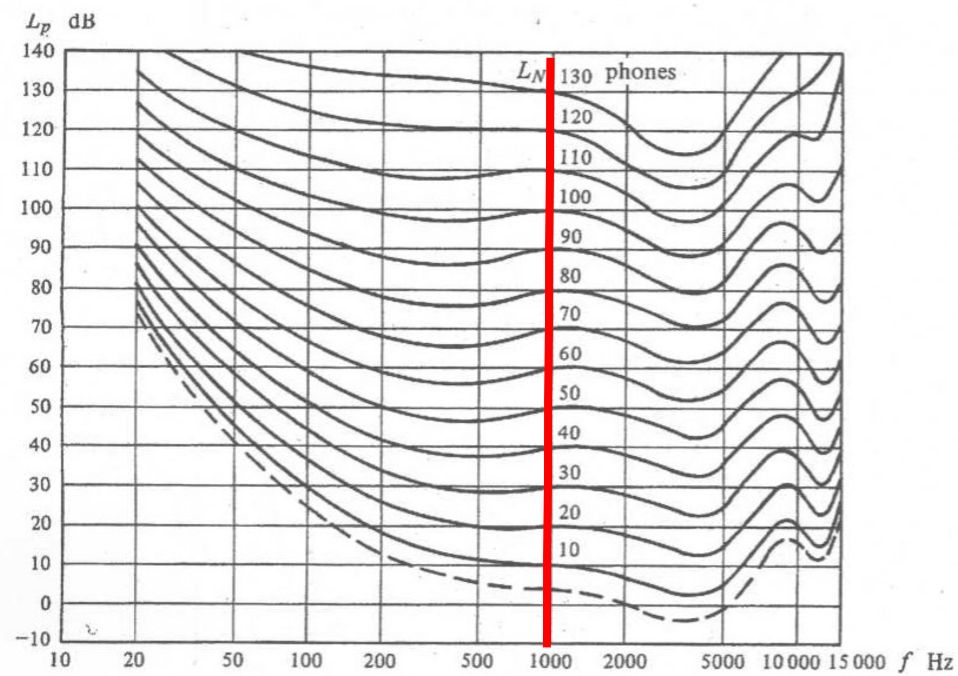
You will hear tones at several frequencies, presented in 10 decreasing steps of 5 decibels.

Count the number of steps you hear at each frequency.

Demonstration 6 (17-18)
p 21-22

Loudness

- It represents the measure of the auditory perception of sound intensity (for pure tones)
- For pure tones, it depends on:
  - frequency
  - sound pressure
  - duration (for « impulsive » sounds)
- The equal-loudness lines correspond to a same value of sound level perception
- The equal-loudness level (L_N) is expressed in phones
- The phone represents the perception levels with same loudness than the one for a pure tone at 1 kHz



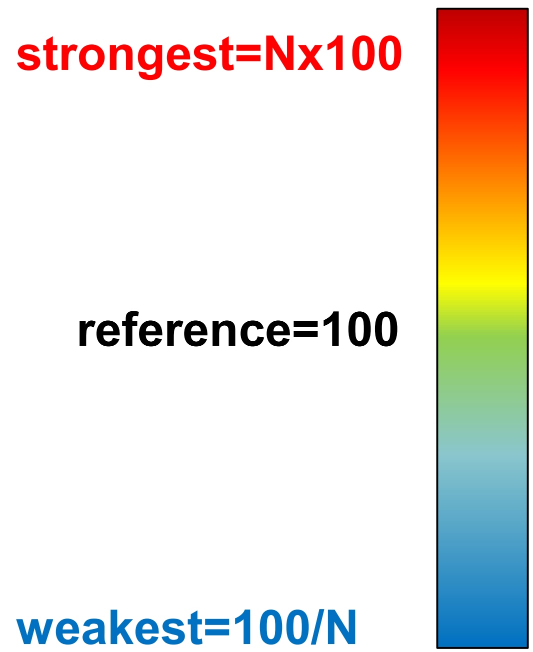
EPFL

Loudness scaling

In this experiment you will rate the loudness of 20 noise samples which are preceded by a fixed reference (to help you you can arbitrarily say reference loudness = 100).

- First you hear the reference sound, followed by the strongest and weakest noise samples

- Now the 20 samples:
*For each sample, write down a **number n reflecting its loudness relative to the reference***
*Ex: **2** if 2 times louder than the reference, or **200***
0.5** if 2 times less loud than the reference , or **50



Demonstration 7 (19-20)
p 23-24

Loudness scaling

- Relationships between the magnitude of sensation and the excitation can be written as:

  sensation~(excitation-threshold)[a]

- In the median area of audition, loudness follows the Law of Stevens:

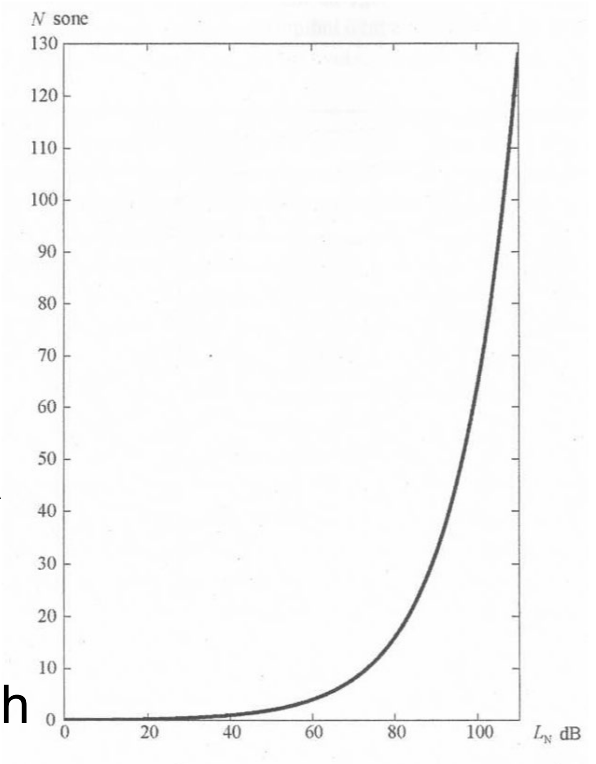
$$N \approx C.I^{0.3}, \text{ or } N \approx C.p^{0.6}$$

  C being a constant (depends on frequency)

Note: a doubling of loudness (*N*➔2*N*) corresponds to a 10 dB increase in sound pressure level (*p*➔3*p*).

- The loudness scale reference is the sone. It corresponds to the loudness of a sound of 1kHz such as *Lp*=40 dB :

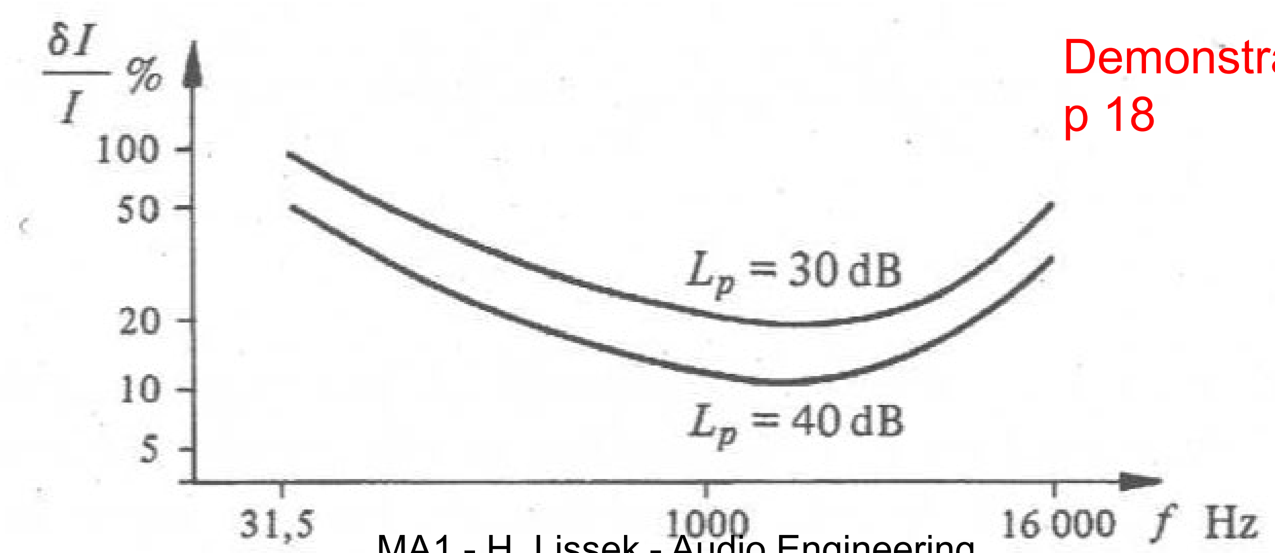
$$N = 2^{\frac{L_N - 40}{10}}$$



EPFL

Differential threshold of loudness variation

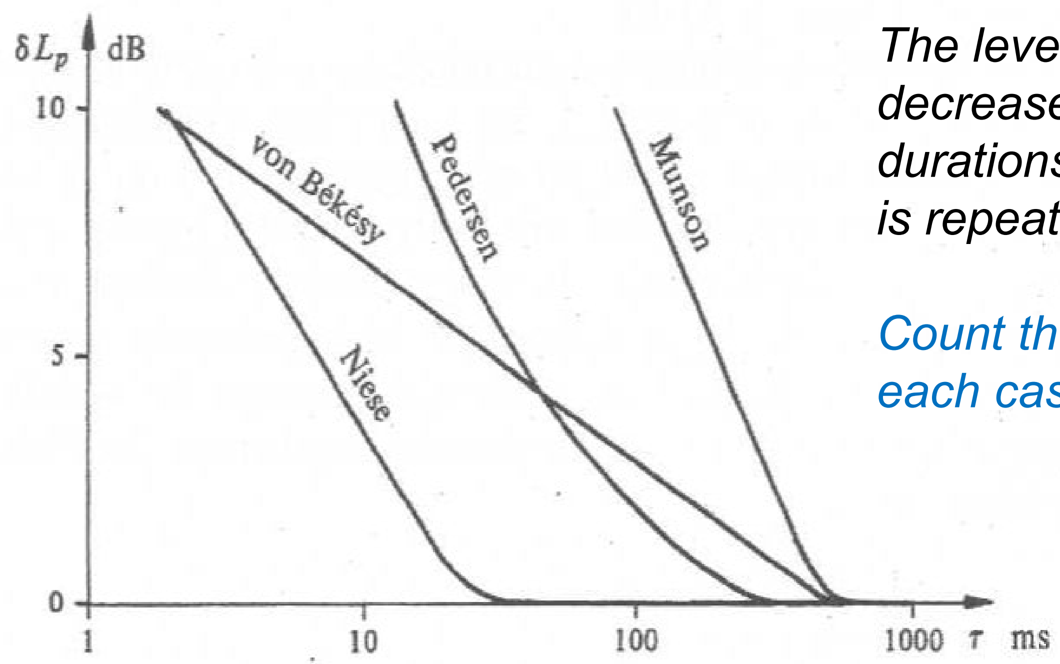
- It is the minimal variation of sound pressure or intensity that induces a variation of loudness (perceived sound level)
- It depends on the sound and the subject
- eg the loudness threshold is 1 dB at the center of audible range
- A halving of perceived loudness corresponds to -10 dB (power divided by 10)!



Demonstration 4 (8-11) p 18

Loudness of short sounds (temporal integration)

- When sound duration is decreasing, below 100 ms (for pure tones) or below 200 ms (broadband noise), loudness is decreasing after a quite-linear law



The level of a broadband noise signal decreases in 8 steps for several signal durations (7 durations). Each sequence is repeated twice.

Count the number of steps you hear in each case

Demonstration 8 (21)
p 25-26

FREQUENCY & PITCH

Pitch

- Pitch is often defined as the characteristic of a sound that makes it sound high or low, or that determines its position on a musical scale;
- The pitch scaling ranges from « low » to « high ». This is related to our ability to recognize a melody.
- Pitch of pure tones mainly depends on frequency (after a non-linear law)
- Increase of intensity increases the pitch of high-frequency sounds and decreases it for low-frequency sounds, and leave unchanged the pitch of medium-range sounds.

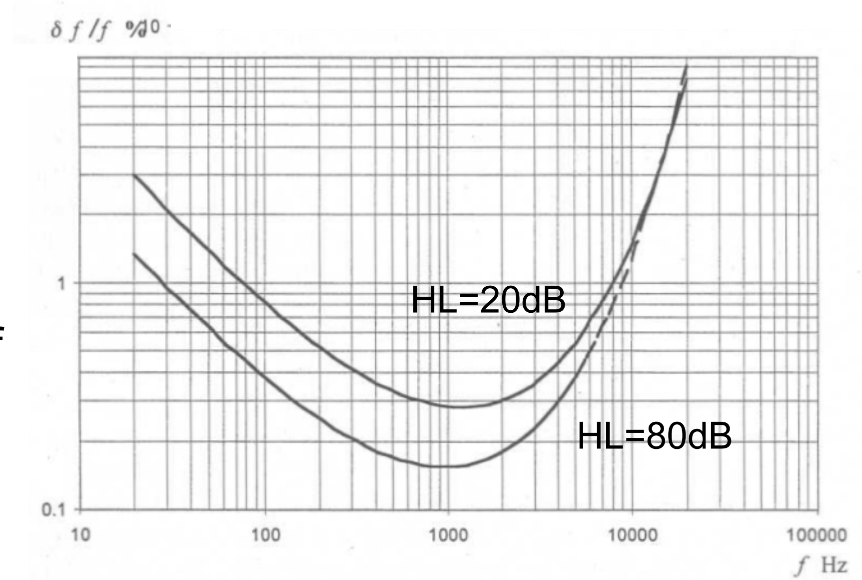
6 tone pairs are presented at various frequencies.

Compare the pitches for each tone pair

Demonstration 12 (27-28)
p 37-38

Frequency Difference Limen (FDL): threshold for pitch variation

- Characterizes thresholds for distinguishing 2 nearly equal pitches
- The FDL, or just noticeable difference (jnd) for pitch has been found to depend on **frequency**, **sound level, duration** of the tone, and **suddenness** of frequency change.
- Classical method of assessment: submit two consecutvie harmonic salves, of frequencies *f* and $f+\delta f$. The δf giving the first sensation of difference is the jnd.
- The human ear is much more sensitive to frequency variations than sound pressure variations



You wil hear ten groups of 4 tone pairs. In each group there's a small frequency difference between the tones of a pair, which are decreasing in each successive groups.

Demonstration 17 (33)
p 44-45

Harmonic pitch

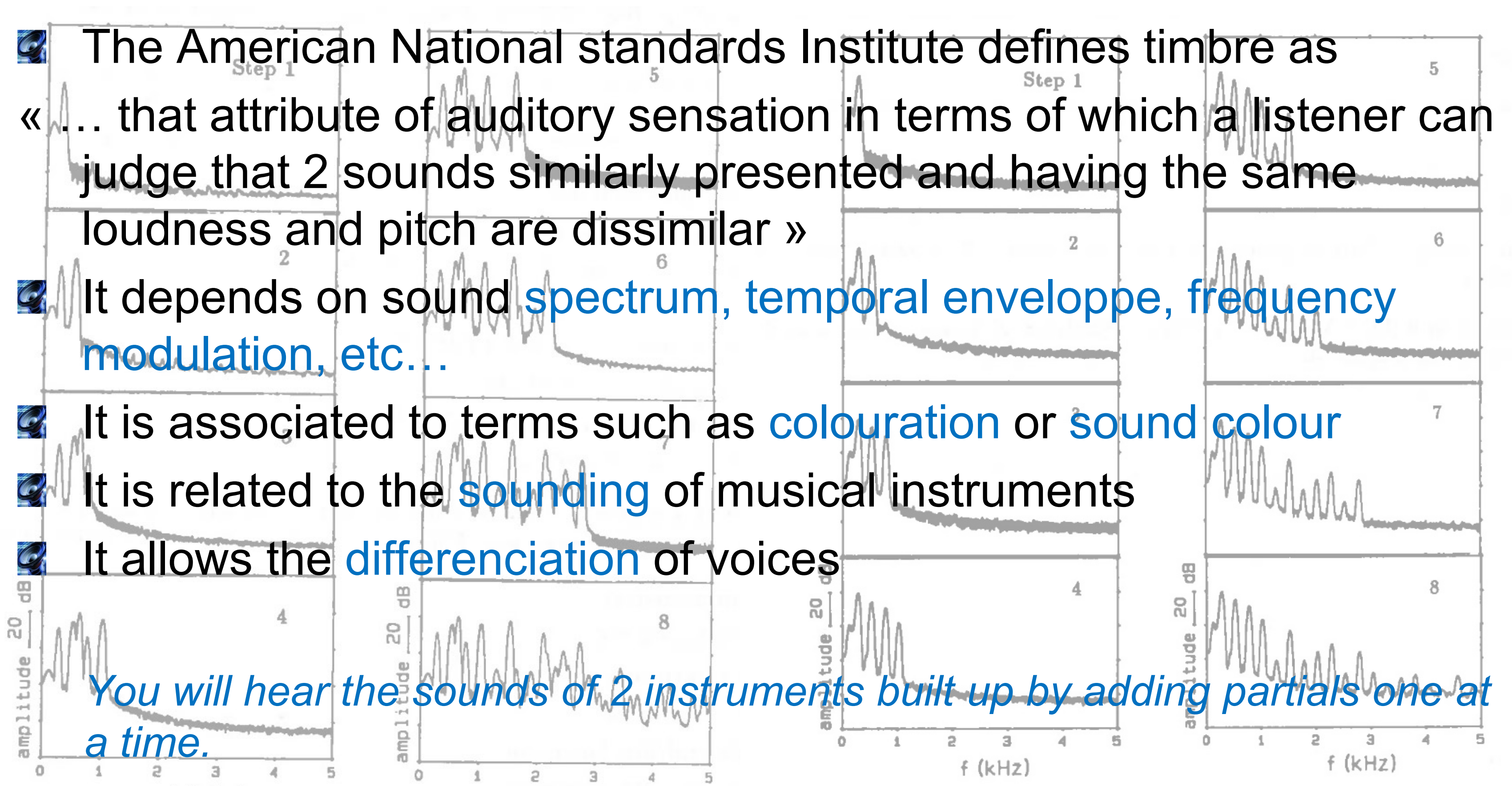
- This is also called musical pitch
- It corresponds to a logarithmic scaling of frequencies
- In most of the cultures it is based on the octave
- It is mainly a scale for people who have a « musical ear »
- The absolute « musical ear » can recognize the harmonic pitch of any tone when played alone.

A 500 Hz tone alternates with a stepwise increasing comparison near 1000 Hz.

Which step seems to represent a « correct » octave?

Demonstration 15 (31)
p 42

Timbre



- The American National standards Institute defines timbre as

« … that attribute of auditory sensation in terms of which a listener can judge that 2 sounds similarly presented and having the same loudness and pitch are dissimilar »

- It depends on sound spectrum, temporal enveloppe, frequency modulation, etc…
- It is associated to terms such as colouration or sound colour
- It is related to the sounding of musical instruments
- It allows the differenciation of voices

You will hear the sounds of 2 instruments built up by adding partials one at a time.
Guess what are the instruments

Demonstration 28 (53)
p 63-66

Waveform and timbre

- Audition of waveforms is a complex problem
- Phase differences between harmonics have audible effects
- Electroacoustics components can modify the phase relationships of harmonics, and then on audition

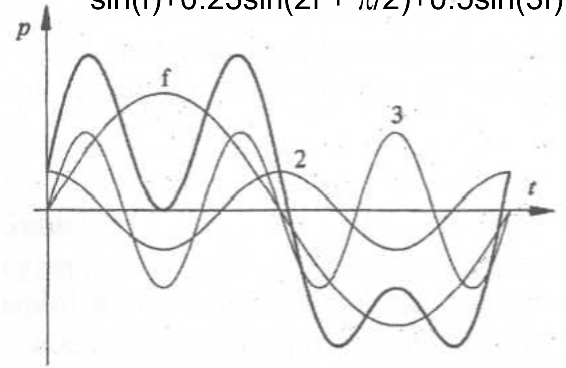
➔ Linear phase distorsion

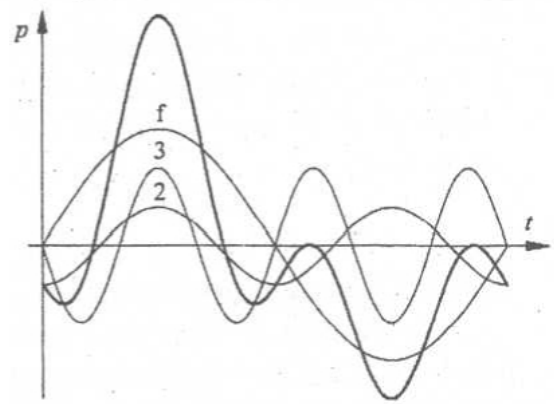
Ex1: Matlab example

Ex2: You will hear a 440 Hz pure tone plus its second harmonic added with a phase varying from -90° to +90°

This is followed by the same tones, distorted through a square-law device.

sin(f)+0.25sin(2f + π/2)+0.5sin(3f)





sin(f)+0.25sin(2f - π/2)+0.5sin(3f+π)

Demonstration 33 (67)
p 77-78

Beats, vibrato and tremolo

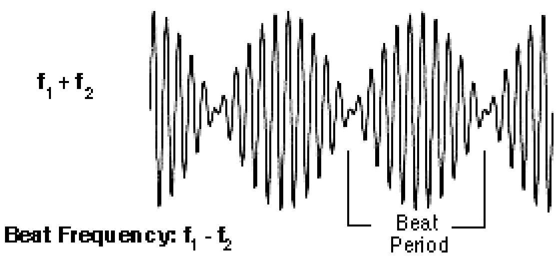
- A listener is submitted to 2 pure tones of very near pitches. He perceives a unique sound, as far as the difference of frequencies δf is small enough
  - the pitch of the resulting sound is comprised between the 2 components
  - the loudness modulates @δf between min and max values
- The vibrato is a periodic variation of frequency of a sound
- The tremolo is a variation of the magnitude of a sound

$$\sin(2\pi f_1 t) + \sin(2\pi (f_1 + \delta f) t) \approx 2\cos(2\pi \delta f t).\sin(2\pi f_1 t)$$

Two tones having frequencies of 1000 Hz and 1004 Hz are presented separately and then together.

Note the beats

Demonstration 32 (62)
p 74-75



Ear's selectivity

- As far as the frequency difference is above a certain threshold, the listener perceives distinctly two pure tones, assuming individual loudnesses are of same order
- The ear's selectivity becomes better as frequency increases

Dissonance and consonance

- The ear is sensitive to frequency ratios and not their difference
- dissonance is an unpleasant feeling of waiting/non-resolution
- concordance is a pleasant feeling of resolution. The 2 tones seem to merge
- Roughness (in savart) is a measure of inharmonicity or dissonance. It is obtained by computing the ratio between two frequency intervals, the first being a approximation of the second :

$$\text{roughness} = 1000.\log_{10}\left(\frac{(f_1/f_2)_{tested}}{(f_1/f_2)_{correct}}\right)$$

Musical scales

- The chromatic scale (tempered) results from the following constraints:
  - transposition, ability to re-build the same melody from any of the notes of the scale
  - Minimal roughness of intervals and chords
  - Limited number of notes (12 for occidental music)

Subjective sounds

- When submitted to an intense pure tone (above 60 phones), a listener perceives a certain timbre, as if he was listening to an harmonic sound
  - ➔ Ears non-linearities induce subjective harmonics
- The pitch of an harmonic sound is the one of its fundamental, even if absent.

You will hear a complex tone with 10 harmonics, first complete and then with the lower harmonics successively removed.

Does the pitch of the complex change?

Demonstration 20 (37)
p 50

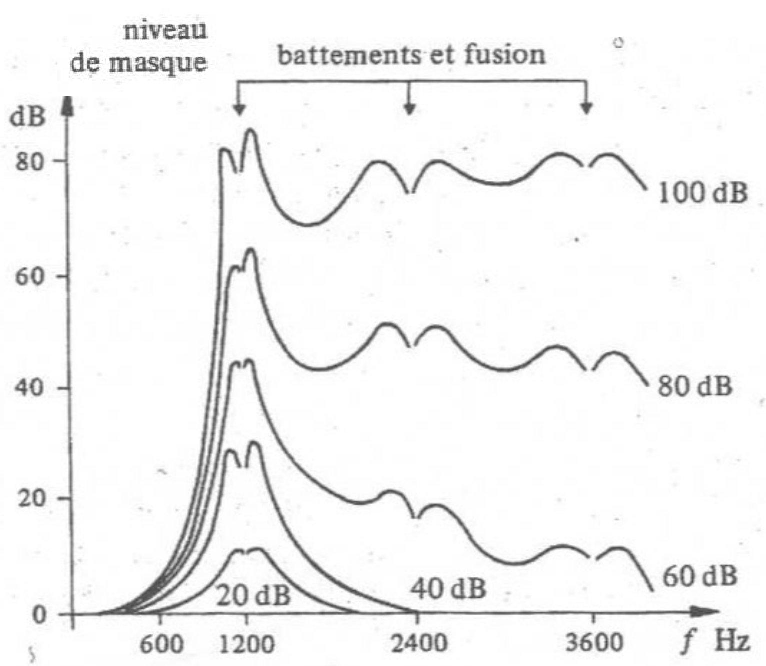
MASKING

Sound masking

- Masking refers to the fact that a sound can be made inaudible by one other
- Masking is due to the intrinsec behavior of the ear
- There are many masking types
  - « Simultaneous » masking
  - Backward masking
  - Forward masking
- A masking sound can alter the loudness and pitch of a pure tone or harmonic sound
- Masking is the property that allows the compression of sound (ex : MiniDisc, MP3…)

Simultaneous masking

- When 2 sounds of enough different levels are heard, the more intense (masker), induces a diminution of loudness of the masked sound, the less instense:



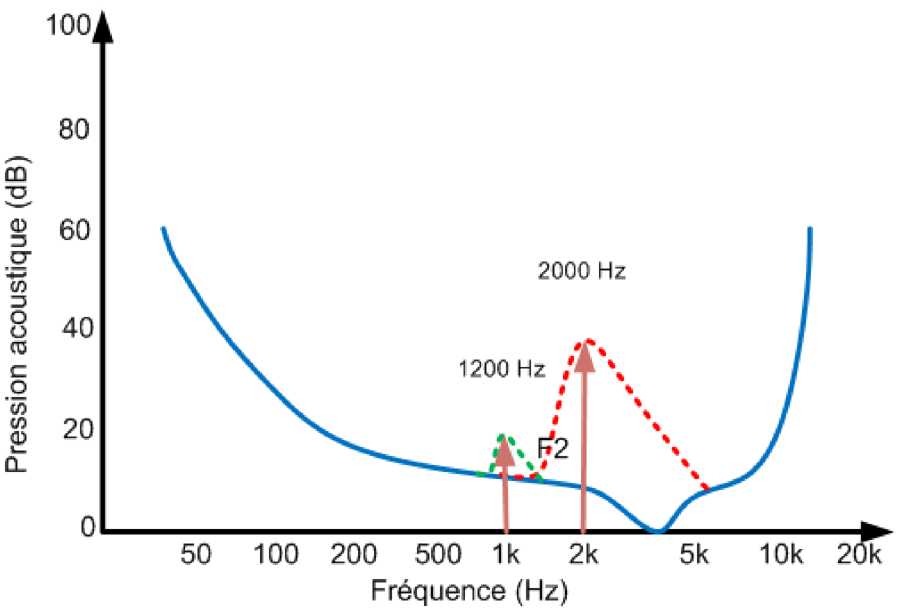
- Here the masking is a sound at f_m=1200 Hz, at different level
  - Auditory threshold of the masked sound is represented in the graph
- Masking is maximum near the masked (although merging occurs at f_m and multiples)
- It is dissymetric in frequency
- Masking effect increases much faster and its dissymetry increases when the masker's level increases

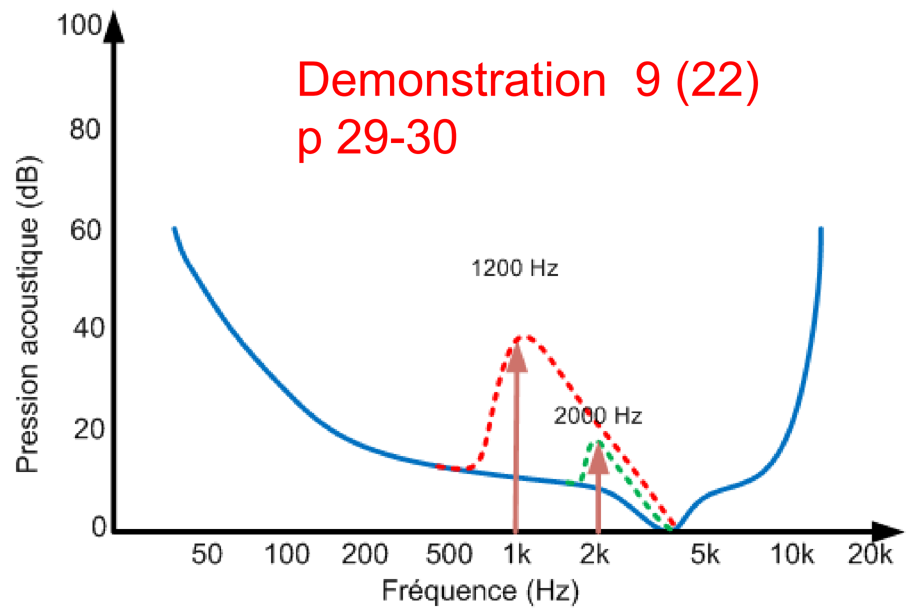
EPFL

Simultaneous masking

A masking tone alternates with the combination of masking tone plus a stepwise-decreasing test tone. First the masker is 1200 Hz and the test tone is 2000 Hz, then the masker is 2000 Hz and the test tone is 1200 Hz.

Count how many steps of the test tone can be heard in each case.





Demonstration 9 (22)
p 29-30

Simultaneous masking

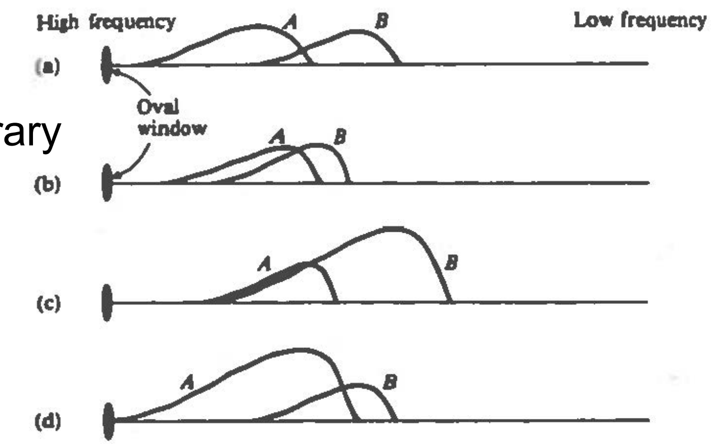
Explanation: simple model of the basilar membrane

a)The 2 excitations don't overlap

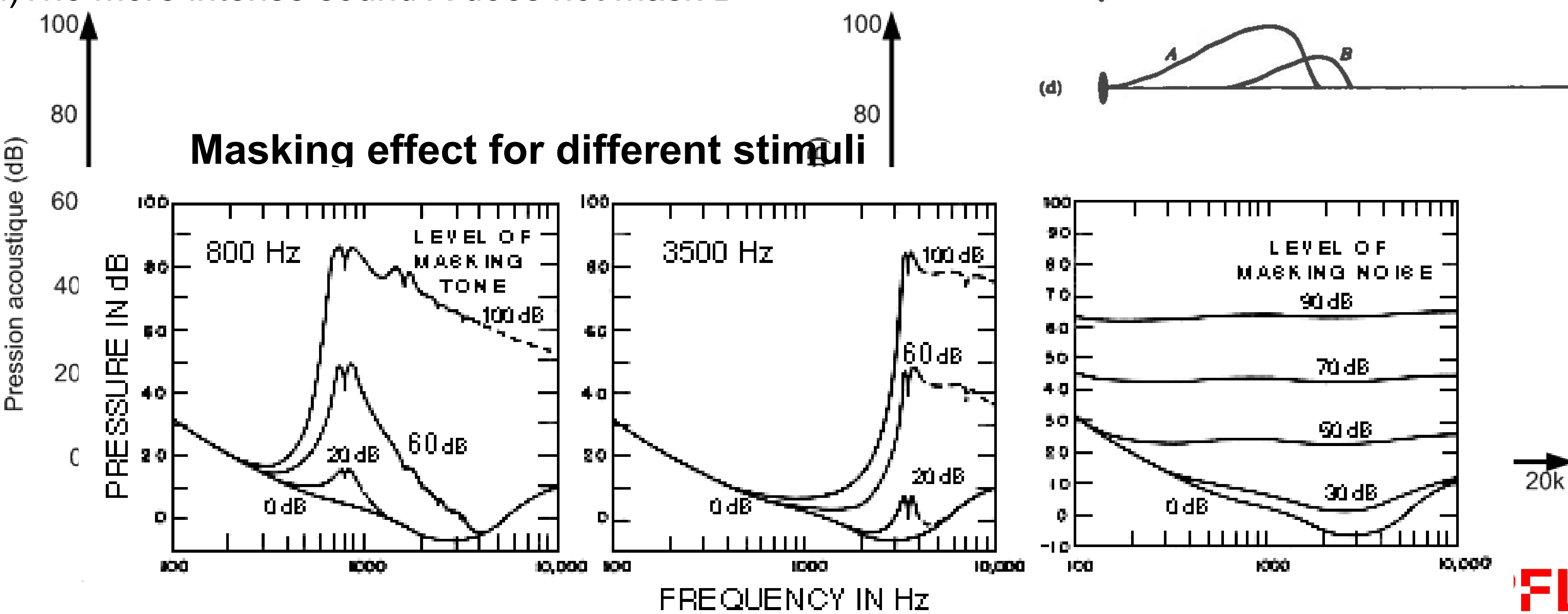
b)Significant overlap: tone B masks tone A more than the contrary

c)The more intense sound B completely masks A

d)The more intense sound A does not mask B



Masking effect for different stimuli



Temporal masking

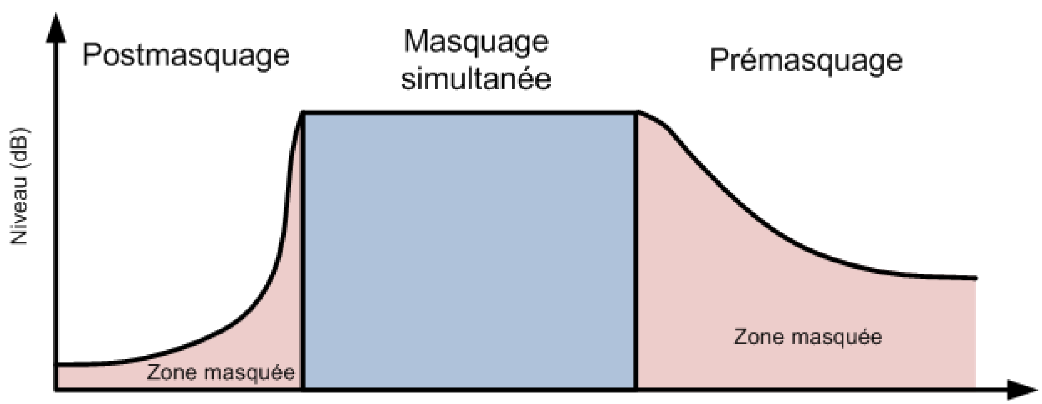
Masking can occur even if the two sounds are not simultaneous

- Forward masking: refers to the masking of a tone by a sound that ends a short time (up to about 20 or 30 ms) before the tone begins.

Suggests that recently stimulated sensors are not as sensitive as fully-rested sensors

- Backward masking: refers to the masking of a tone by a sound that begins a few milliseconds after the tone has ended.

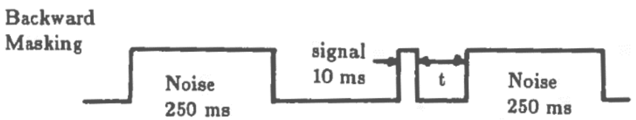
May be explained by higher processes in the nervous system?



Temporal masking

•First you hear a brief sinusoidal tone, decreasing in 10 steps of 4 decibels each.

•Now the same signal is followed by a noise burst with a brief time gap in between. It is heard alternating with the noise burst alone. For 3 decreasing time-gap values, you will hear two staircases.
Count the number of steps for which you can hear the brief signal preceding the noise.



•Now the noise burst precedes the signal. Again 2 staircases are heard for each of the same 3 time-gap values.
Count the number of steps that you can hear the signal following the noise.

Demonstration 10 (23-25)
p 31-32

Critical bands

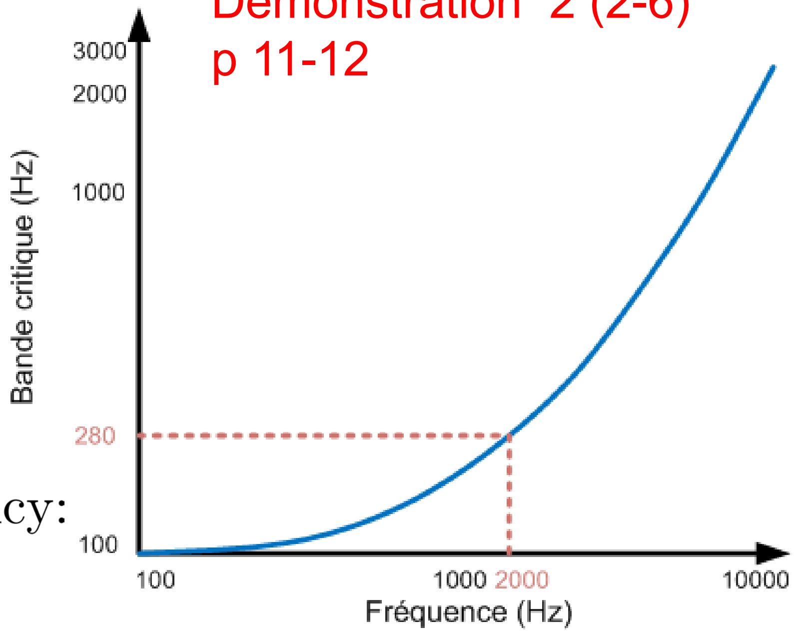
- There exists frequency bands around any tone, called critical band, inside which perception of the tone is altered
- The width of these bands depends on the frequency
- The masking effect increases with the bandwidth of the masking noise, when the bandwidth of masking noise is lower than the critical bandwidth

Critical bands

You will hear a 2000 Hz tone in 10 decreasing steps of 5 decibels.
Count how many steps you can hear.

Demonstration 2 (2-6)
p 11-12

Same question when:
The signal is masked with broadband noise
The noise has a bandwidth of 1000 Hz
The noise has a bandwidth of 250 Hz
The noise has a bandwidth of 10 Hz



Critical band as a function of frequency:

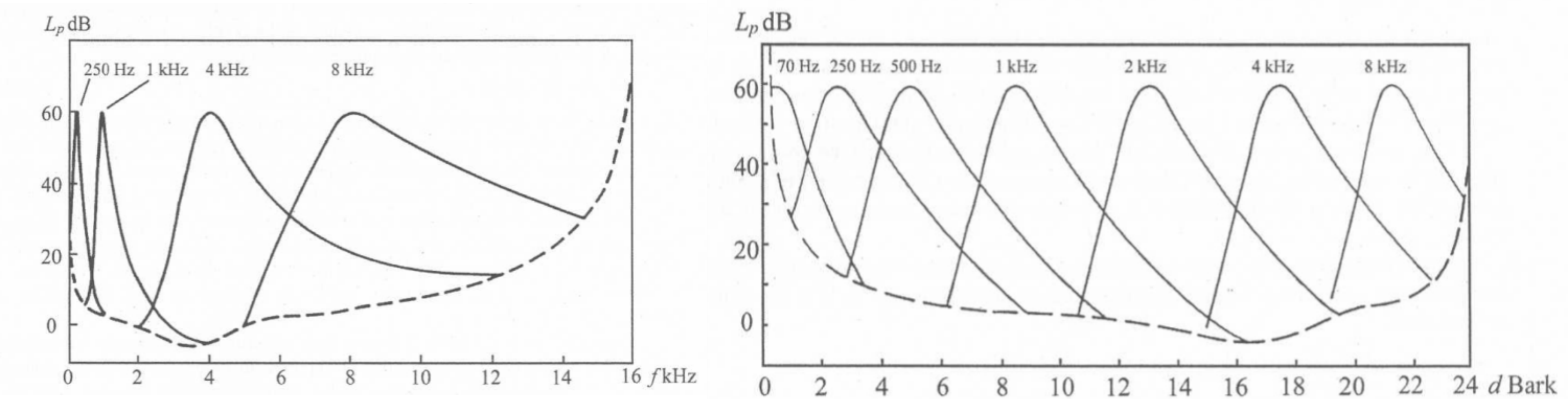
$$B_c = 25 + 75\left(1 + 1.410^{-6} f^2\right)^{0.69}$$

Cochlear filter banks

- Thanks to the tonotopy (representation of the auditory spectrum as function of the position along the cochlea), the masking effect can be computed as a function of the distance d (in Bark) from the apex of cochlea (helicotreme)

$$d = 13\arctan\left(\frac{0.76f}{1000}\right) + 3.5\arctan\left(\frac{f}{7500}\right)^2$$

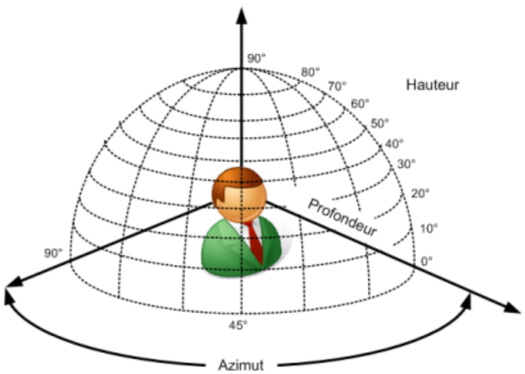
- The auditory system behaves as a filter bank with overlapping



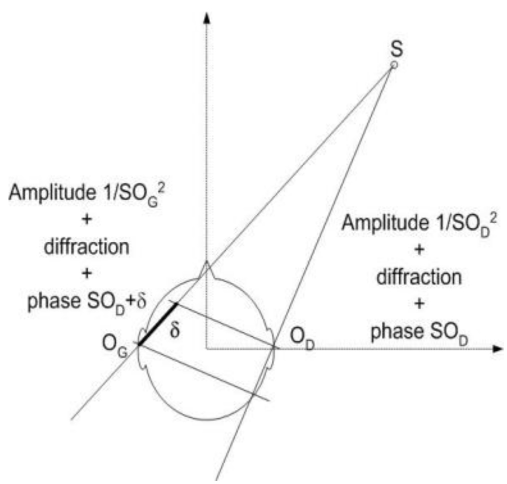
LOCALIZATION

Auditory localization

- The auditory localization is the faculty to locate a sound source in space



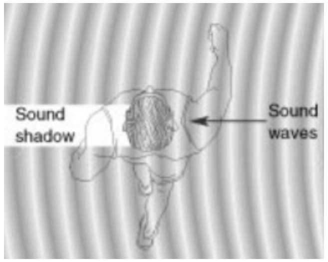
- 2 phenomena explain azimuthal localization :



  - acoustic pressure differences at the 2 ears (head diffraction)
  - acoustic path differences between the sound source and each ear

Head diffraction (HRTF)

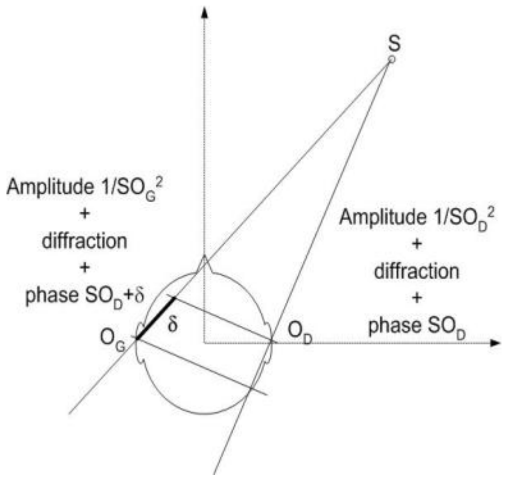
- The head is an obstacle to sound propagation and creates an « acoustic shadow» for the farest ear



- HRTFs highlight that, for a source outside the median plane of the head, there exists sound intensity differences between the ears
- Help to locate the high frequency sounds

Interaural Level Difference (ILD)

- The difference of acoustic paths induces a difference of sound intensity (see acoustic damping), denoted Interaural Level Difference (ILD), which induces a loudness and timbre alteration contributing to localize a source



Interaural Time Differences (ITD)

- The difference of acoustic paths induces a delay, denoted Interaural Time Difference (ITD), between the time of arrival at each ear
- In permanent regim, this delay corresponds to a phase shift
- The phase shift allows to localize the source only if it is lower than π rad or than ½ the wavelength (otherwise there is an ambiguity)
- Allows to localize mainly the low frequency sounds

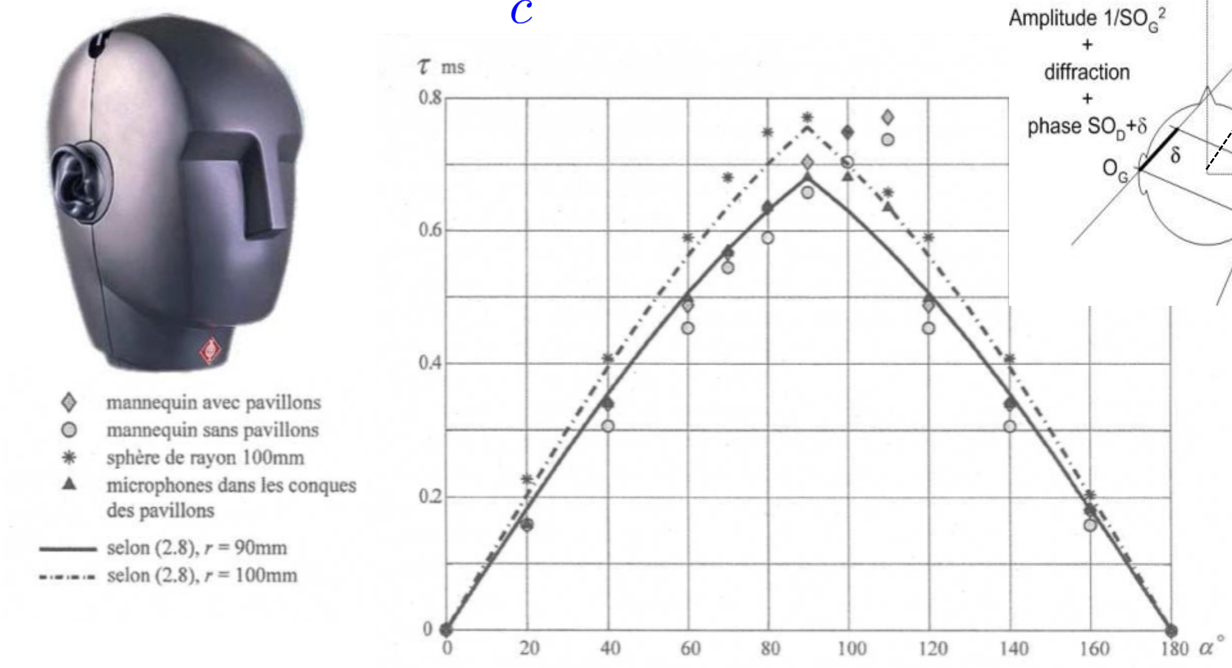
Demonstration 37 (72-73) p 88

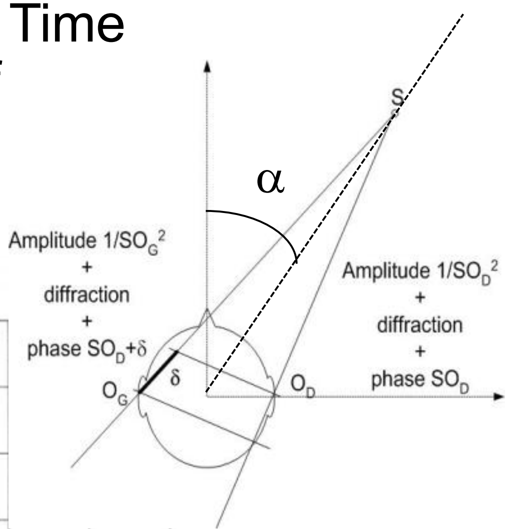
Tones of 500 Hz and 2000 Hz are heard with alternating interaural phases of plus and minus 45°
Next, the interaural arrival times of a click is varied. The apparent location of the click appears to move.
Finally, the interaural intensity differences of a 250Hz and a 4000Hz tone are varied

Interaural Time Differences (ITD)

- For a spherical head of radius *r*, the Interaural Time Difference τ can be calculated as a function of

$$\tau = \frac{r}{c}(\alpha + \sin\alpha)$$



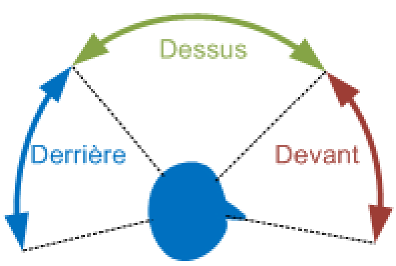


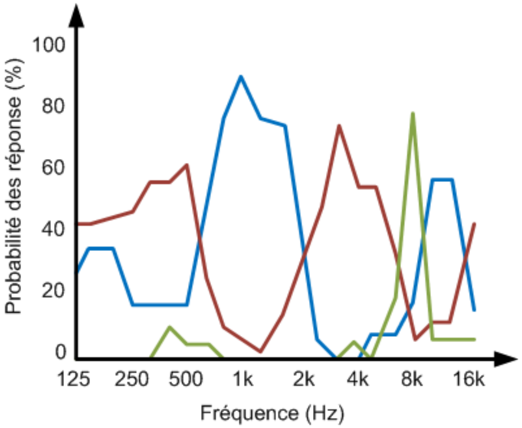
Localization properties

- Humans are capable of feeling ITD of the order of magnitude of tens of μs
- The sound source is perceived at the far left (right earphone is delayed) or at far right (left earphone is delayed): we speak of lateralization
- In normal binaural situations, the ILD and ITD are combined and allow the source localization, even more efficient if the sound is complex
- With headphones, it is possible to substitute a phenomenon for the other (eg. a delay compensated by an increase of level)

Localization in elevation

- Explained by the head and concha diffraction
- Presence of important peaks and pits in the spectrum of the incident sound at high frequencies (cf. HRTF)
- Localization in elevation is less precise than in the horizontal plane (15 to 20° errors for a source above the head)
- The spectrum plays a determinent role





Localization in distance

- It is extremely bad
- It seems based on loudness and timbre comparisons with the subject's own auditory experience
- Teh implied mechanisms:
  - Intensity variation
  - Spectrum (absorption in the air varying with low and high frequencies)
  - Acoustic field structure (plane/spherical waves)
  - The ratio between direct and reverberated sound