Digital Sound Capstone
DXARTS 460

Lecture 3:
Psychoacoustic Fundamentals

Pitch | Loudness | Timbre

Spring 2011

Instructor: Stelios Manousakis

Psychoacoustics: what is sound?

- ***Psychoacoustics*** is the study of the way humans perceive sound

Linear and logarithmic

- There are two principal modes of perception: linear and logarithmic

- In the case of ***linear*** perception, the change between two values is perceived on the basis of the ***difference*** between the values. Thus, a phenomenon is perceived linearly if a change from 0.1 to 0.2 is judged to be the same amount of increase as a change from 0.7 to 0.8.

- A phenomenon is ***logarithmic*** if a change between two values is perceived on the basis of the ***ratio*** of the two values. In this case, a change from 0.1 to 0.2 (a ratio of 1:2) would be perceived as the same amount of increase as a change from 0.4 to 0.8.

- Human auditory perception (and in other senses, such as vision and touch) is mostly logarithmic

- Therefore, we use logarithmic units to measure perceptual sensations

Psychological dimensions of sound

- Loudness -> relates to intensity
- Pitch -> relates to frequency
- Duration -> relates to time
- Timbre -> a complex dimension: relates to (steady) spectrum, formant structure, transients and other features
- Localization -> Radiation and other spatial cues

The human ear

- 3 principal parts:
  - outer ear
  - middle ear
  - inner ear
- Outer ear: a sound gathering device
- Middle ear: translates air pressure into liquid pressure mechanically with the eardrum
- Inner ear: sends electric impulses to the brain; essentially, a biological microphone

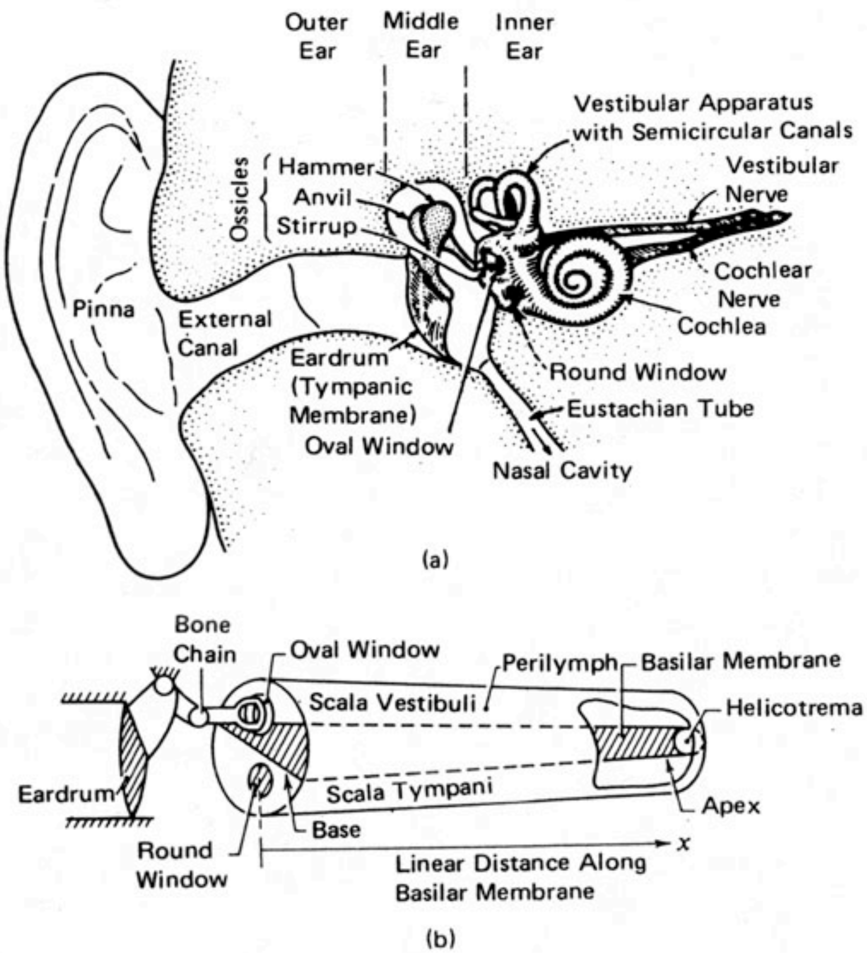


FIGURE 2.6 (a) Schematic view of the ear (Flanagan, 1972; Fig. 4) (not in scale); (b) the cochlea shown stretched out (highly simplified).

The human ear

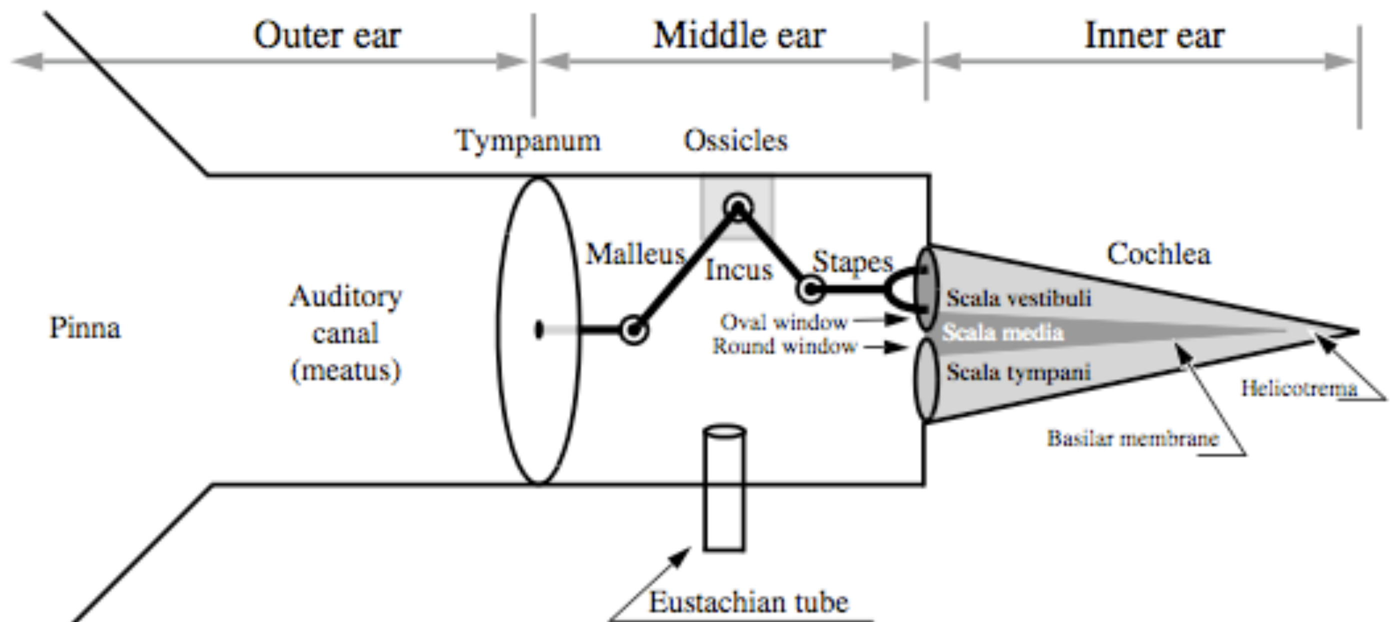


Figure 6.1
Schematic diagram of the human ear.

Intensity

- ***Intensity*** is a measure of the *power* in a sound that actually contacts an area such as the eardrum: it characterizes the ***rate*** at which energy is delivered in the audible sensation associated with amplitude.

- It is *proportional to the square of the amplitude*: that is, if the amplitude of a sound doubles, the intensity increases by a factor of 4.

- Intensity is perceived nearly logarithmically, and is expressed as power applied over an area in watts per square meter. The range of intensities that a human can perceive is bounded by 10^{12} W/m2 at the threshold of audibility and 1 W/m2 at the threshold of feeling.

- The difference in wattage between the softest perceptible sound and the threshold of pain is on the order of a millionfold.

Intensity | Decibel

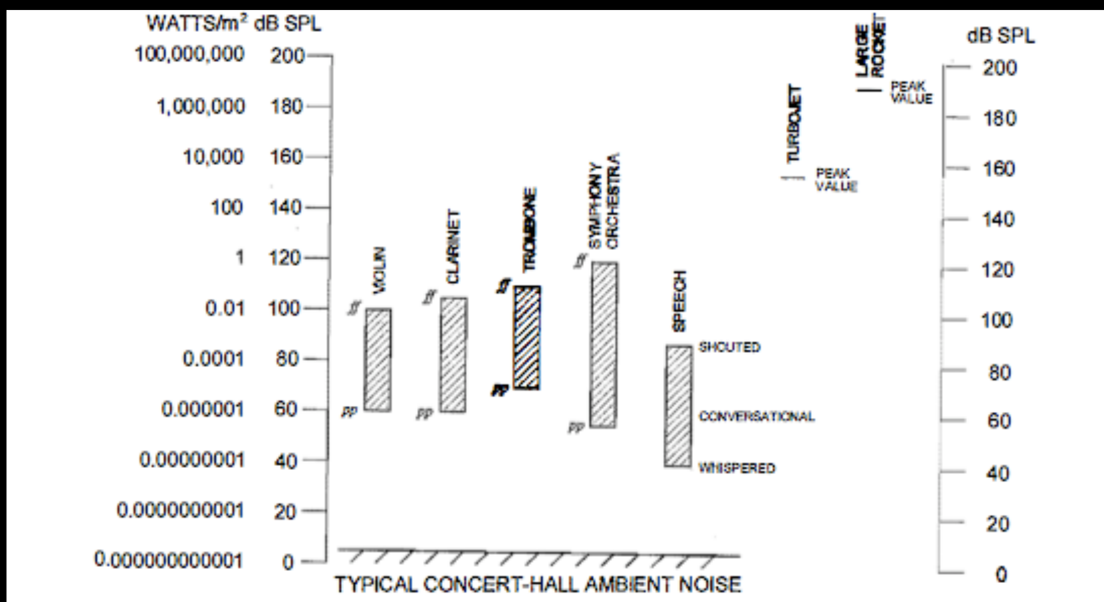
- To make a more manageable scale of measurement, the scale is therefore not only comparative but also *logarithmic.* The unit of measurement is the ***decibel*** *(dB).*

- Amplitude is also measured in dB. Recalling that power is proportional to amplitude squared, decibels in watts may be equated to decibels in amplitude.

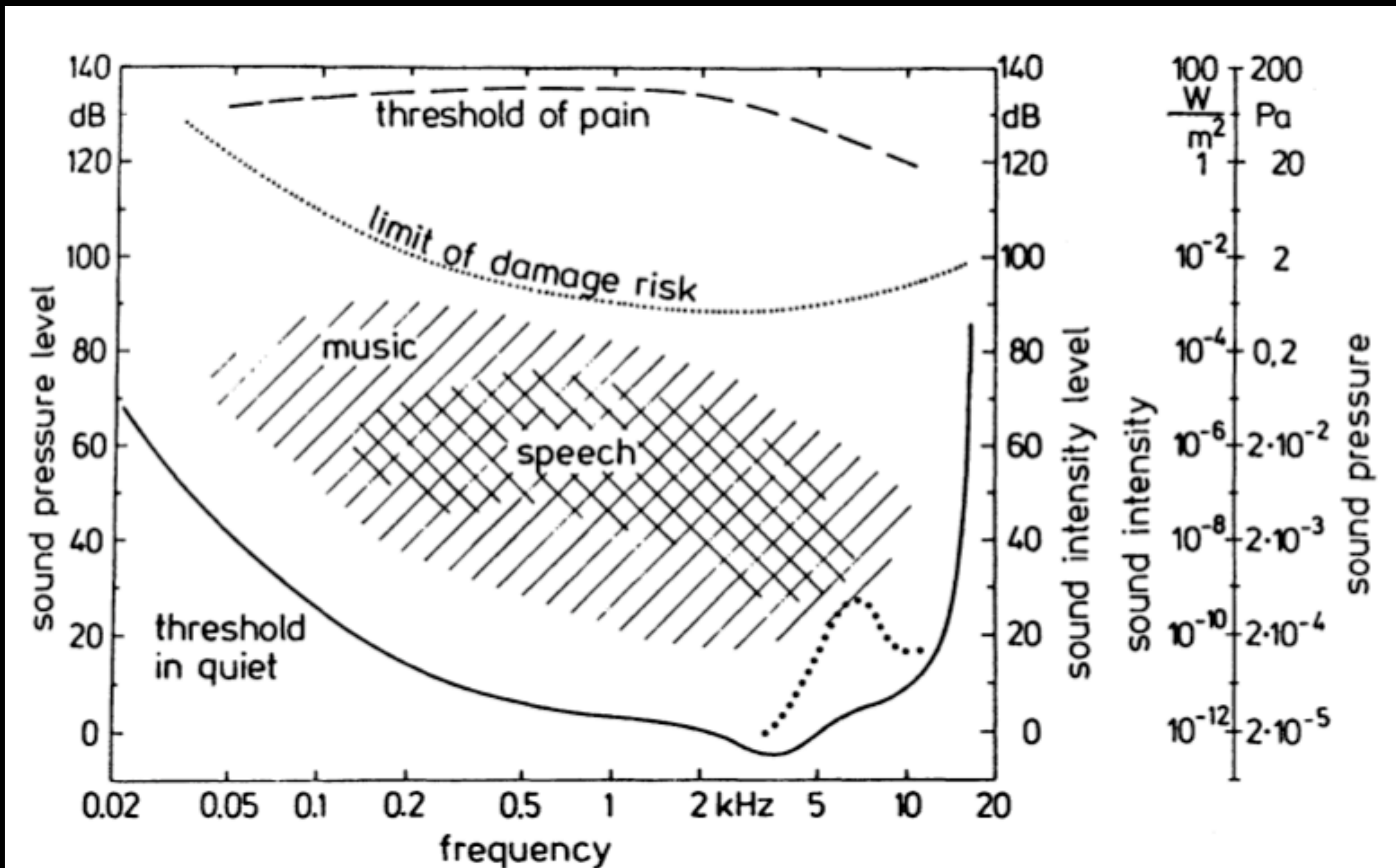
- The decibel is a way of expressing the ***relative power*** of something compared to a standard, usually electrical power, current or voltage, or sound levels: it is a logarithmic unit of relative measurements.

Decibels

- Doubling the *power* results in an increase of 3dB.
- Doubling the *pressure level* results in a change of approximately 6dB.
- In acoustics, the reference sound-pressure level is generally taken as the threshold of audibility. A sound at that level has a value of 0 dB SPL.
- Conversational speech has an intensity of approximately 60 dB SPL, while shouted speech is closer to 80 dB SPL. A sound at the threshold of feeling has a level of 120 dB SPL, and so the range of audibility (120 dB) represents a ratio in amplitudes of one to one million.



Area of audibility



Loudness

- ***Loudness*** is a measure of the subjective response to a sound's amplitude (intensity).
- ***Loudness*** correlates to the amount of energy contained in a wave, its *power (work/distance)*.
- ***Intensity*** is a measure of the power in a sound that actually contacts an area such as the eardrum.

Curves of Equal Loudness

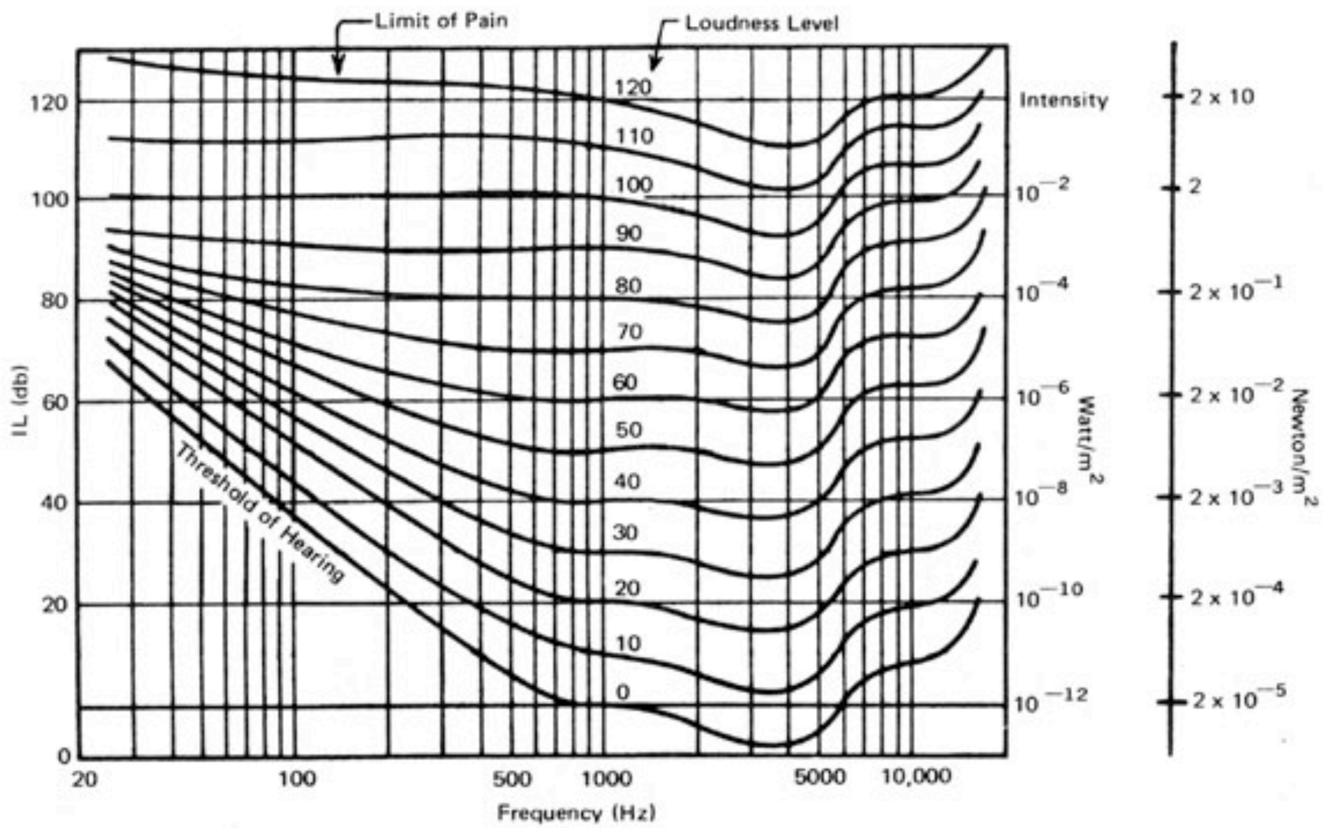


FIGURE 3.13 Curves of equal loudness (Fletcher and Munson, 1933) in a sound intensity level (IL) and frequency diagram. The corresponding scales of sound intensity and average pressure variation are also shown. Reprinted by permission from the *Journal of the Acoustical Society of America*.

Frequency range of human hearing

- The theoretical bandwidth of human hearing is 20Hz – 20,000Hz
- In practice, *the upper range* is seldom above 18 kHz for the young adult, and with noise exposure and age, it drops dramatically
- At the *low frequency end*, there is a range where the sensation of pitch starts to disappear, and below it, the sense of physical vibration takes over. This range is between 20 and 25 Hz. Below that, frequencies can be felt by the body; the brain can also distinguish individual cycles of the sound wave and are experienced as *rhythm*
- Audible wavelengths: audible sounds will have wavelengths of between about 1.7 centimeters (0.021 inches) and 17 meters (18.6 yards)
- • Physical objects in-between your ears and a sound source will act as filters. E.g.: wearing a hat with a brim, puts a notch (dip in the spectrum) at about 2kHz

Pitch

- ***Pitch*** is our subjective response to frequency.

- There is a nonlinear relationship between pitch perception and frequency.

- Thus, equivalent perceptual pitch changes correspond to equivalent changes in an ***exponent*** that describes the differences between two frequencies.

- Listeners compare tones on the basis of the musical interval separating them–that is, on the basis of the ***ratio*** of their frequencies rather than the difference between them.

Pitch

- Humans hear an equivalent pitch class with every doubling of frequency (***octave***).

- frequencies of successive octaves of concert A
  > 55 (55 * 2 ** 0) | 110 (55 * 2 ** 1) | 220 (55 * 2 ** 2) …

- Western music divides the octave in 12 equal parts: twelve tone temperament.

- Choose a starting frequency
  > multiply it by 2 ** n/12 for n = 0 to 11
  > higher octaves, doubling frequency
  > lower octaves, halving frequency

Pitch

- Twelve-Tone equal temperament in Hertz

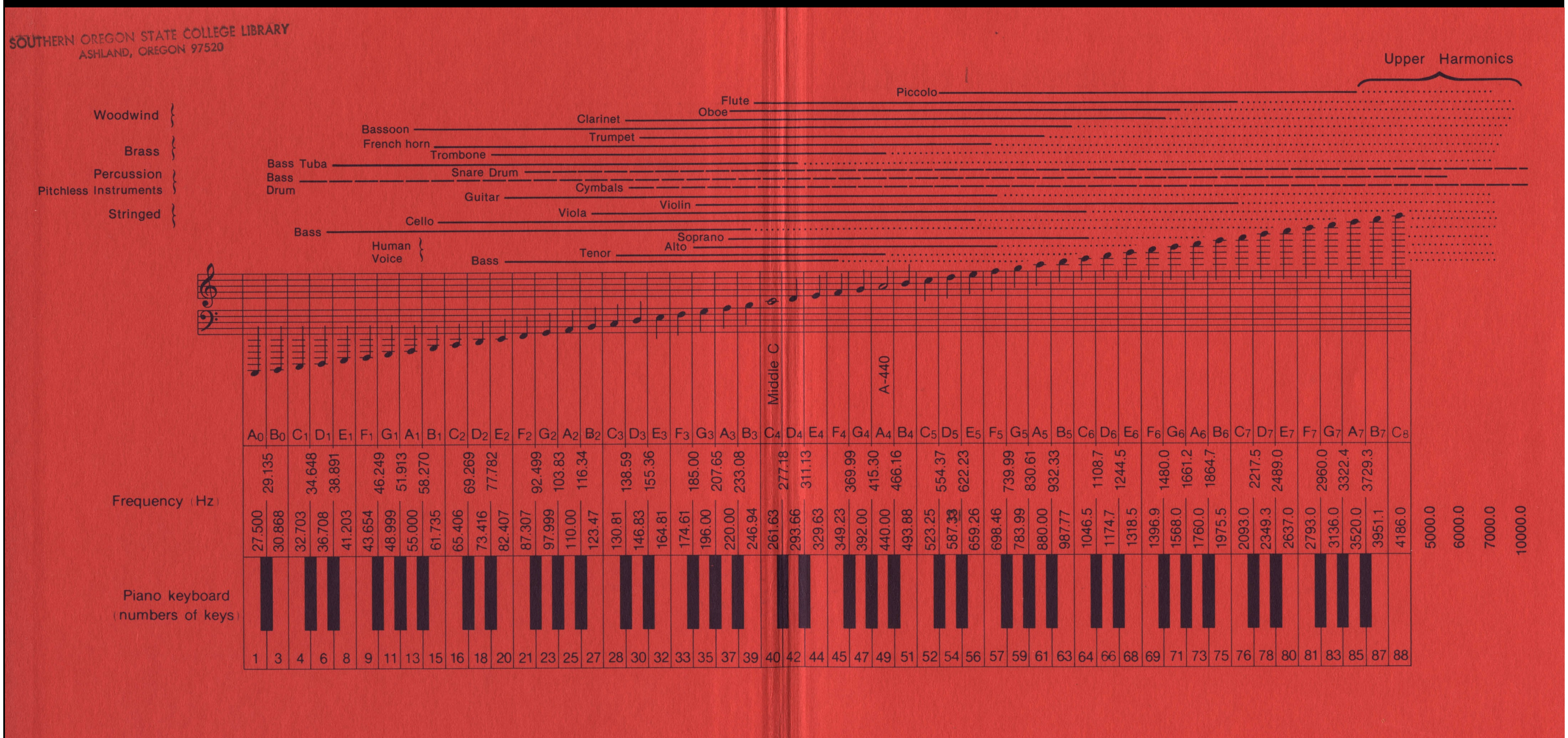
A =	220 * 2 ** 0/12 =	220Hz
A# =	220 * 2 ** 1/12 =	233Hz
B =	220 * 2 ** 2/12 =	247Hz
C =	220 * 2 ** 3/12 =	261.6Hz
C# =	220 * 2 ** 4/12 =	277Hz
D =	220 * 2 ** 5/12 =	293.6Hz
D# =	220 * 2 ** 6/12 =	311Hz
E =	220 * 2 ** 7/12 =	329.6Hz
F =	220 * 2 ** 8/12 =	349.2Hz
F# =	220 * 2 ** 9/12 =	370Hz
G =	220 * 2 ** 10/12 =	392Hz
G# =	220 * 2 ** 11/12 =	415Hz

Pitch

- Octaves within the hearing range:

0. 20Hz
1. 40Hz
2. 80Hz
3. 160Hz
4. 320Hz
5. 640Hz
6. 1.25KHz
7. 2.5kHz
8. 5KHz
9. 10KHz
10. 20KHz

Pitch


SOUTHERN OREGON STATE COLLEGE LIBRARY
ASHLAND, OREGON 97520
Upper Harmonics
Woodwind
Piccolo
Flute
Oboe
Clarinet
Bassoon
Brass
French horn
Trumpet
Trombone
Bass Tuba
Percussion
Pitchless Instruments
Snare Drum
Bass Drum
Cymbals
Stringed
Guitar
Violin
Viola
Cello
Bass
Human Voice
Soprano
Alto
Tenor
Bass
Middle C
A-440
A0 B0 C1 D1 E1 F1 G1 A1 B1 C2 D2 E2 F2 G2 A2 B2 C3 D3 E3 F3 G3 A3 B3 C4 D4 E4 F4 G4 A4 B4 C5 D5 E5 F5 G5 A5 B5 C6 D6 E6 F6 G6 A6 B6 C7 D7 E7 F7 G7 A7 B7 C8
Frequency (Hz)
29.135 34.648 38.891 46.249 51.913 58.270 69.269 77.782 92.499 103.83 116.34 138.59 155.36 185.00 207.65 233.08 277.18 311.13 369.99 415.30 466.16 554.37 622.23 739.99 830.61 932.33 1108.7 1244.5 1480.0 1661.2 1864.7 2217.5 2489.0 2960.0 3322.4 3729.3
27.500 30.868 32.703 36.708 41.203 43.654 48.999 55.000 61.735 65.406 73.416 82.407 87.307 97.999 110.00 123.47 130.81 146.83 164.81 174.61 196.00 220.00 246.94 261.63 293.66 329.63 349.23 392.00 440.00 493.88 523.25 587.33 659.26 698.46 783.99 880.00 987.77 1046.5 1174.7 1318.5 1396.9 1568.0 1760.0 1975.5 2093.0 2349.3 2637.0 2793.0 3136.0 3520.0 3951.1 4186.0
5000.0 6000.0 7000.0 10000.0
Piano keyboard (numbers of keys)
1 3 4 6 8 9 11 13 15 16 18 20 21 23 25 27 28 30 32 33 35 37 39 40 42 44 45 47 49 51 52 54 56 57 59 61 63 64 66 68 69 71 73 75 76 78 80 81 83 85 87 88

Just Noticeable Difference (JND)

- Pitches (and other sensations) must differ by a minimum threshold for us to distinguish them. This thresh- old is the *just noticeable difference (JND) of pitch.*

- The rule of JND was discovered by Ernst Heinrich Weber, in the 19th century, and has been further developed since then

Critical bands

- Critical bands can be thought of as channels of frequency-selective psychoacoustic processing that affect our perception of pitch, loudness, and masking of frequency components lying within a critical frequency distance of one another.
- ***Critical bands***, determine the ability of the ear to discriminate adjacent tones.
- The concepts of consonance and dissonance relate to critical bands
- MPEG format uses them to tune its compression

Masking

- ***Frequency masking***:
  When two sinusoids are presented simultaneously, the fainter sinusoid can be rendered inaudible, or *masked,* by the louder sinusoid if the fainter one lies within a certain frequency range of the louder one.

- ***Temporal masking***:
  \- *Forward masking* Even after a sound ends, its effect on the threshold of hearing lingers for a while. The threshold of a test signal following the masker is impaired for a period of time. Can last for as long as 100-200ms

  \- *Simultaneous masking:* The masker and the test signal are presented at the same time; this is identical to frequency domain masking

  \- *Backward masking:* A masker influences the audibility of a fainter test signal that precedes it! Sound perception is actually integrated over a time interval preceding the moment of recognition: a 'window' around 200ms. Fainter sounds lying within this interval are subject to some degree of masking regardless of their order of arrival.

Masking

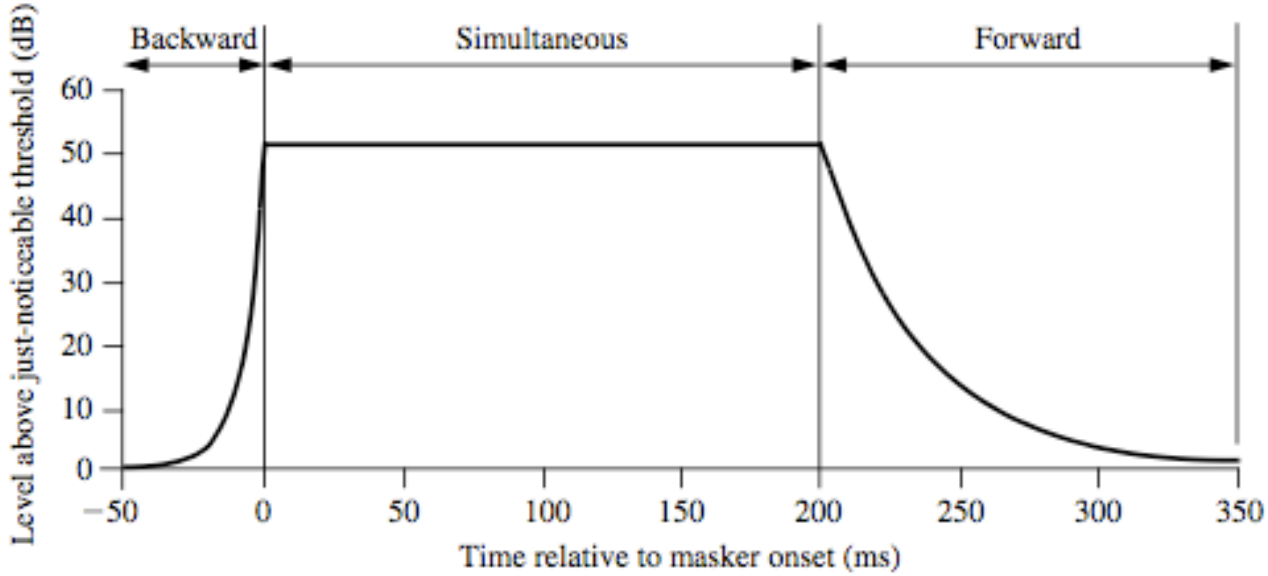


Figure 6.12
Forward and backward masking.

Pitch

- Two closely tuned pure tones will cause a **beating** sensation until the difference between their frequencies exceeds about 10 to 15 Hz.
- Within the region of beating lies the ***limit of discrimination***: a point at which the frequency separation of two tones is sufficiently large for the listener to perceive two separate tones.
- At separations just outside the range of beating, there is a region described by acousticians as ***tonal roughness***.
- This perception of pitch is associated with ***critical bands***

Pitch

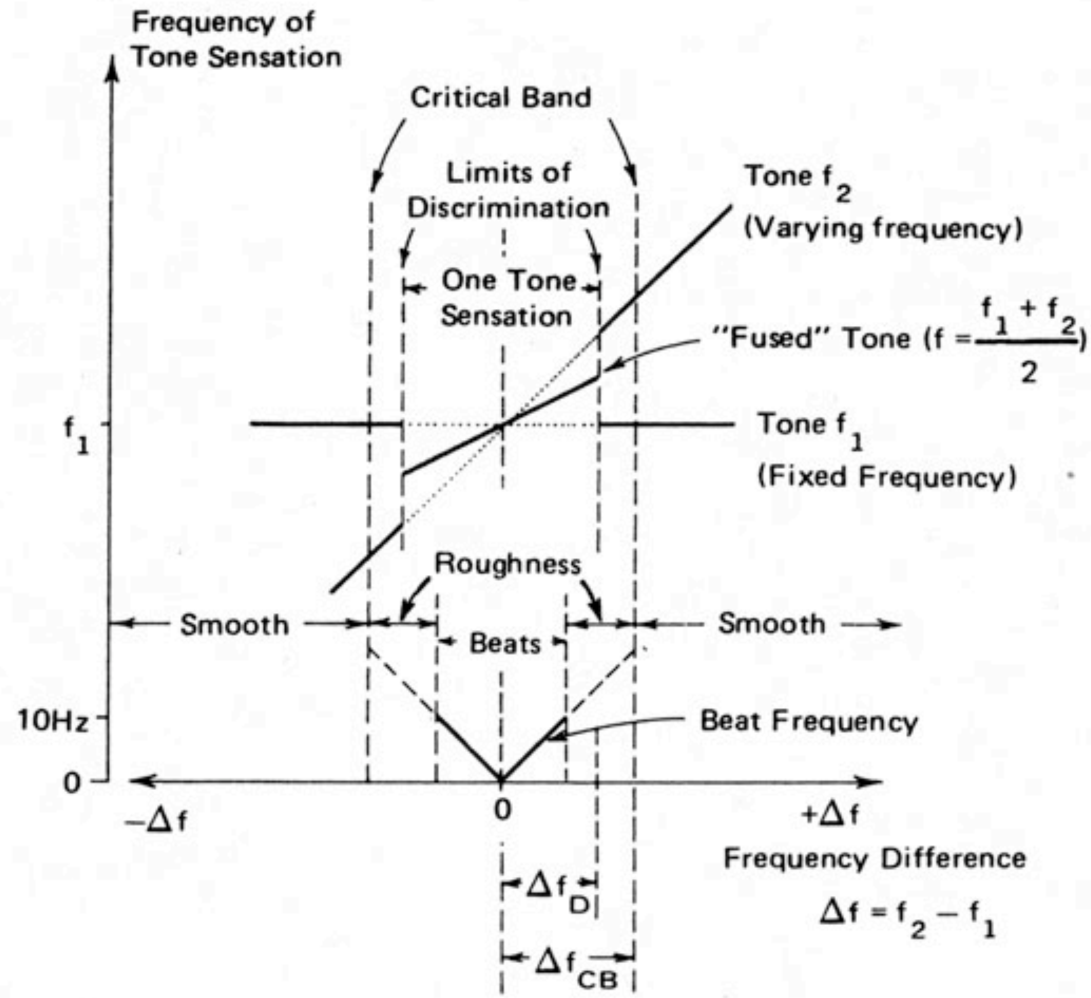
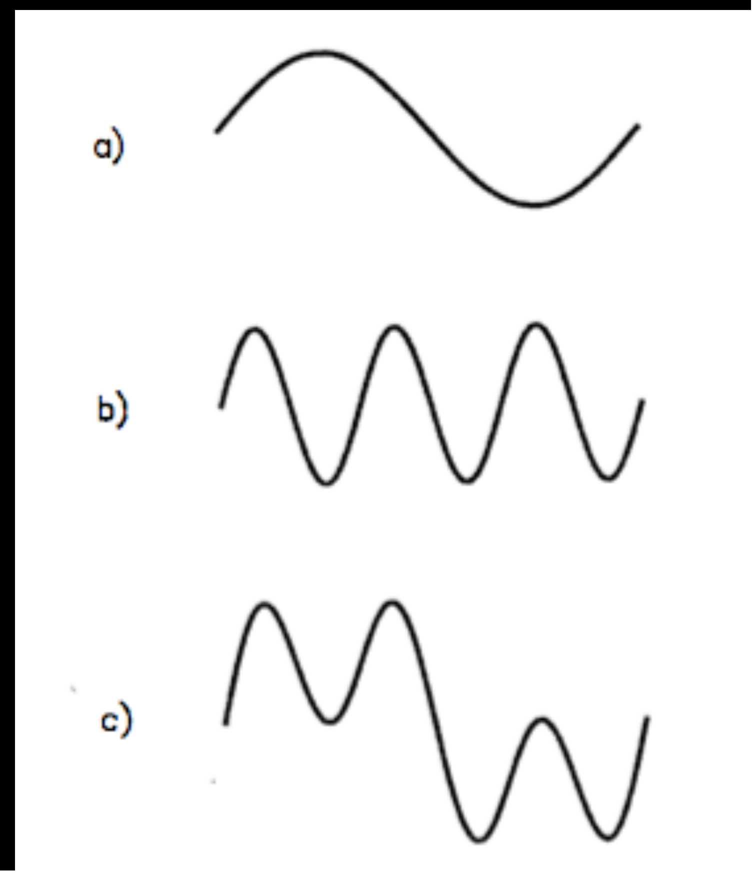


FIGURE 2.12 Schematic representation of the frequency (heavy lines) corresponding to the tone sensations evoked by the superposition of two pure tones of nearby frequencies f_1 and $f_2 = f_1 + \Delta f$.

Timbre

- Musical ***timbre*** is the characteristic tone quality of a particular class of sounds.
- A complex, multi-dimensional psychological dimension of sound. A subjective term, without one measuring unit.
- Relates to the steady spectrum of a sound, but also to formant structure, transients and other features
- Jean Baptiste Fourier theorized that any periodic waveform can be expressed as the sum of one or more sine waves.

  Figure ex.: a + b = c



Noise

- Complex timbres that are not periodic
- Spectrum is non-discrete, but is described as a distribution: random phenomena

- Many different categories of noise:

- ***white*** noise: large bandwidth with equal power distribution
http://upload.wikimedia.org/wikipedia/commons/6/66/Whitenoisesound.ogg

- ***pink*** noise: power spectral density is inversely proportional to the frequency (more psychoacoustically 'correct')
http://en.wikipedia.org/wiki/File:Pink_noise.ogg

- ***brown*** (brownian/red) noise or random walk: random movement of particles in fluid, liquid or gas
http://en.wikipedia.org/wiki/File:Brownnoise.ogg

- ***gray*** noise: an even more psychoacoustically 'correct' pink noise version
http://en.wikipedia.org/wiki/File:Gray_noise.ogg

-Also: *blue, violet, orange, green, black* noises; for more, see:
http://en.wikipedia.org/wiki/Colors_of_noise

Localization

• Characteristics of the acoustics of the space (reverberation)

• The perception of direction is partly the result of the encoding happening through the pinna of our ears.

• Important factors: Delays between the two ears and filtering of spectral content from our head

Terminology recapitulation

Psychoacoustics terms:

- Sound: air pressure vibrations translated mechanically by the ear to liquid vibrations and then to current, that then travels through a human's nervous system to the brain

- Loudness: human's response to intensity (logarithmic perception)

- Pitch: human's response to frequency (logarithmic relationship)

- Timbre: humans response to the vibration pattern of air molecules

Bibliography

- Deutsch, D. (ed.) (1999) *The Psychology of Music* (2nd ed). San Diego, CA: Academic Press
- Dodge, C. and Jerse, T. (1997). *Computer Music Synthesis, Composition, and Performance (2nd ed.).* New York: Schirmer Books
- Dowling W. J. and Harwood D. L. (1986). *Music Cognition.* San Diego, CA: Academic Press
- Everest, F. A. (2001). *The master handbook of acoustics* (4th ed.), New York, McGraw Hill
- Loy, G. (2007). *Musimathics, Volumes 1 and 2.* Cambridge, MA: MIT Press
- Moore, F. R. (1990). *Elements of Computer Music*. Englewood Cliffs, NJ: Prentice Hall
- Truax, B. (2001). *Acoustic communication*. Westport, Conn. [u.a.], Ablex Publ.