Aalto University
School of Electrical
Engineering

Communication acoustics
Ch 10: Basic psychoacoustic quantities

Ville Pulkki and Matti Karjalainen

Department of Signal Processing and Acoustics
Aalto University, Finland

September 12, 2023

This chapter

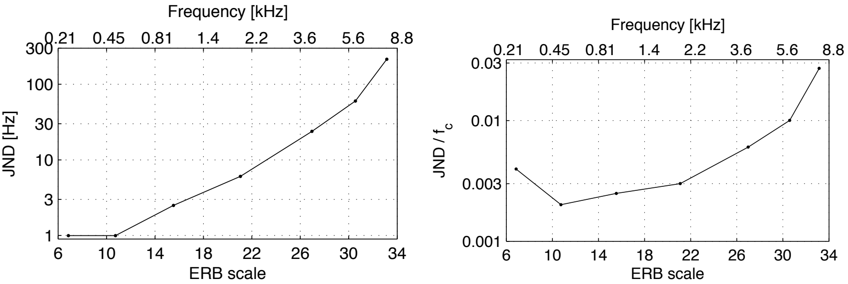
- Pitch
- Loudness
- Timbre
- Duration

Pitch

- Definition: "that auditory attribute of sound according to which sounds can be ordered on a scale from low to high" ANSI

Pitch

- Definition: "that auditory attribute of sound according to which sounds can be ordered on a scale from low to high" ANSI
- JND of frequency of two successive sinusoids



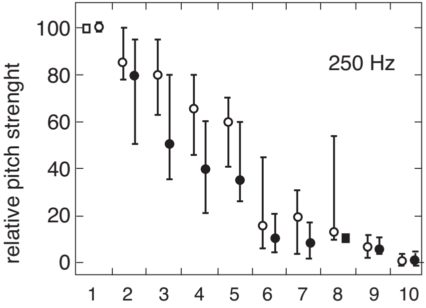
Adapted from Sek and Moore (1995)

Pitch strength

- Different sounds produce differently strong perception of pitch

Pitch strength

- Different sounds produce differently strong perception of pitch

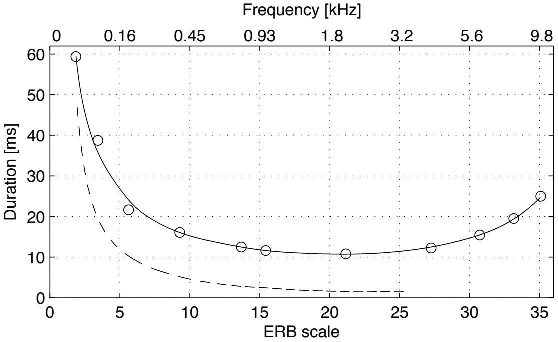


1. pure tone
2. low-pass harmonic tone (7 harmonics)
3. low-pass harmonic tone (all harmonics)
4. AM modulated tone
   (mod. freq 250 Hz, f_0 = 1kHz)
5. band-pass complex tone
   (f_0 = 250 Hz, harmonics btw 1 and 3 kHz)
6. band-pass noise (200 - 300 Hz)
7. low-pass noise (cutoff 250 Hz)
8. comb-filtered noise (delay 4 ms)
9. AM modulated noise (mod.freq 250 Hz)
10. high-pass noise (cutoff 250 Hz)

Adapted from Fastl and Stoll (1979)

Pitch perception versus duration of sound

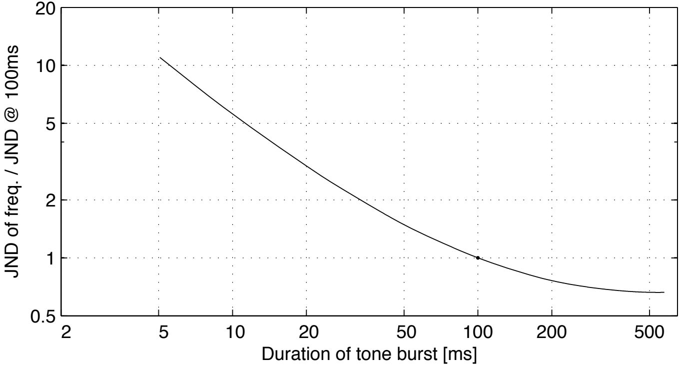
- Minimum length required for pitch perception
- Already very short tone bursts lead into perception of pitch



Adapted from Burck et al. (1935)

Pitch perception versus duration of sound

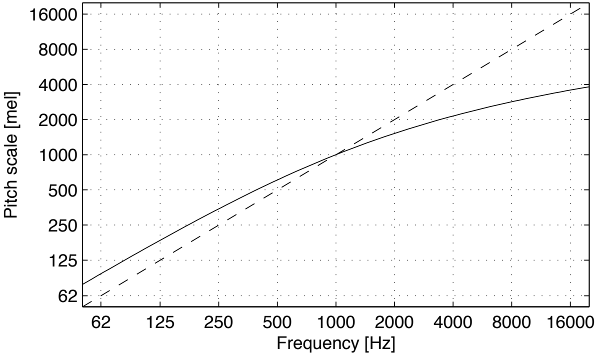
- The accuracy of pitch perception is enhanced during first 200 ms of sound



Adapted from Fastl and Zwicker (2007)

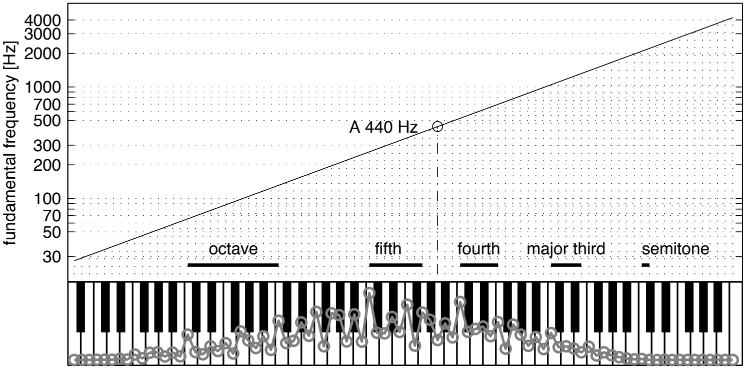
Mel frequency scale

- 'adjust the pitch of the test tone to be two times higher than the reference tone'
- Mel scale derived



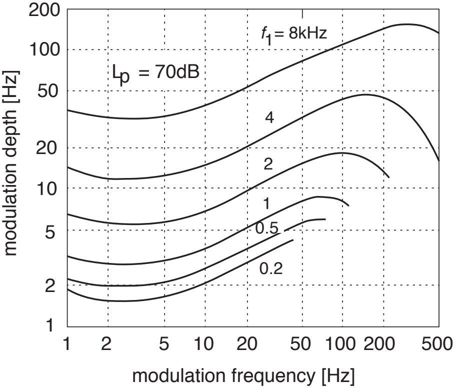
Musical scale

- Musical pitch scale is logarithmic
- (Approximate) frequency ratios: Octave = 2:1, Fifth = 3:2, Fourth 4:3, Third 5:4



Detection of frequency modulation

- Curves have different carrier frequencies



Adapted from Demany and Semal (1989)

Virtual pitch

- Although lowest harmonics are missing, a pitch is perceived to f_0
- Compare: telephone band 300Hz + 4kHz, although male voice $f_0 < 100\text{Hz}$

Pitch theories

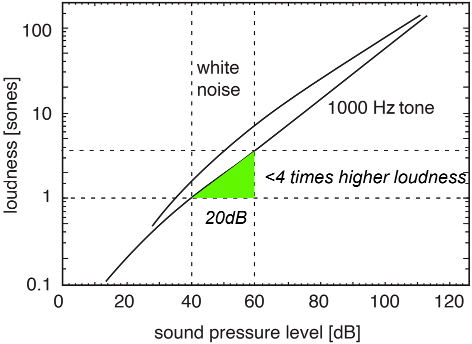
- Peak of activation at basilar membrane?
- Some kind of autocorrelation process after cochlea?
- Pitch theories have been debated for decades
- Neither theory explains fully perceptual phenomena

Loudness

- ’that attribute of auditory sensation in terms of which sounds can be ordered on a scale extending from quiet to loud" ANSI
- One of fundamental quantities in psychoacoustics
- Approach loudness with simple tests, and continue to more complicated ones

Loudness

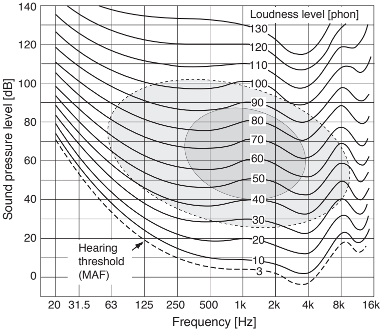
- Task: adjust sound to be ’twice as loud’, lots of subjects, repetitions, and SPLs tested
- Define loudness scale with unit [sone]
- 10dB increase in SPL leads to doubling of loudness



Reprinted from Canteretta and Friedman (1978)

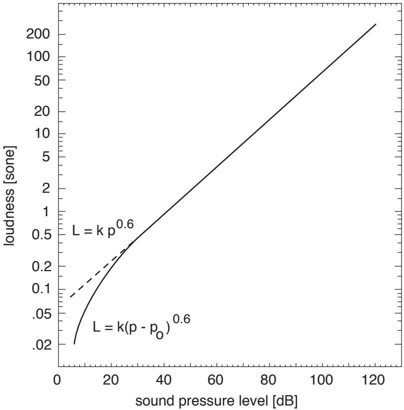
Loudness level

- Loudness level defined with reference values located at 1 kHz with 10 dB spacing in the sound pressure level
- Unit: [phon]



Connection between sound pressure, loudness and loudness level

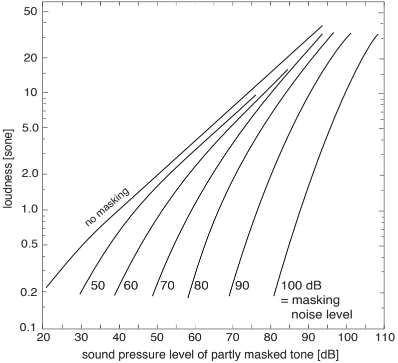
- N = loudness [sone]
- L_{L} = loudness level [phon]
- $N = 2^{(L_{\mathrm{L}}-40)/10}$
- $L_{\mathrm{L}} = 40 + 10\log_2(N)$
- $N = k \cdot (p - p_0)^{0.6}$
- Doubling loudness in sones means 10phon (= 10dB @ 1kHz) change in loudness level (or SPL)



Reprinted from Canteretta and Friedman (1978)

Loudness of tone in presence of noise

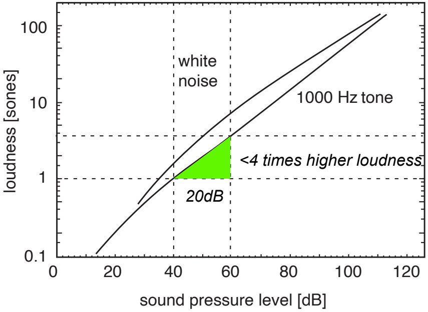
- White noise as masker with different SPLs
- Loudness decreases fast when approaching the masking threshold



Reprinted from Canteretta and Friedman (1978)

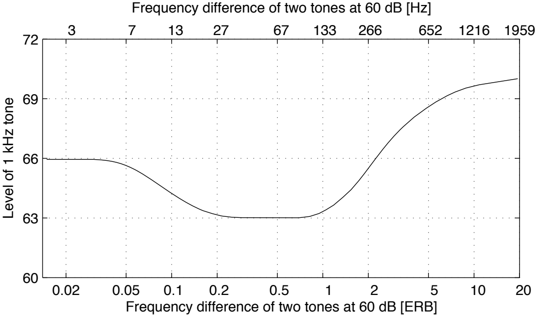
Loudness with broad-band signals

- Loudness is often affected, if the spectrum of sound changes and SPL is kept equal
- This was already seen in basic loudness listening test with sinusoids and noise



Reprinted from Canteretta and Friedman (1978)

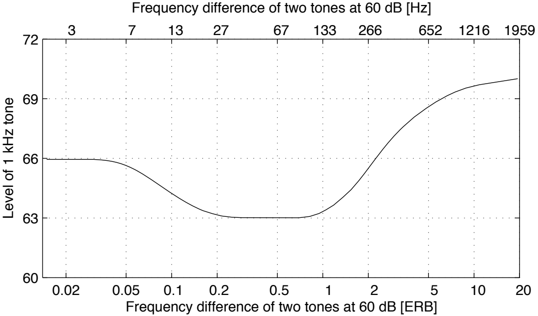
Loudness with two sinusoids



Adapted from Fastl and Zwicker (2007)

- The level of a reference tone adjusted to match the loudness with a pair of tones
- Frequency difference shown in x-axis

Loudness with two sinusoids



Adapted from Fastl and Zwicker (2007)

- The level of a reference tone adjusted to match the loudness with a pair of tones
- Frequency difference shown in x-axis
- Must be some kind of frequency integration in hearing!

A theoretic view of loudness process

- Input signal spectrum $S(f)$ is warped to auditory frequency scale z
- $S'(z) = S[f(z)]\frac{df}{dz}$

A theoretic view of loudness process

- Input signal spectrum $S(f)$ is warped to auditory frequency scale z
- $S'(z) = S[f(z)]\frac{df}{dz}$
- Signal also spreads in frequency due to frequency masking, $B(z)$ is spreading function
- $E(z) = S'(z) \star B(z)$

A theoretic view of loudness process

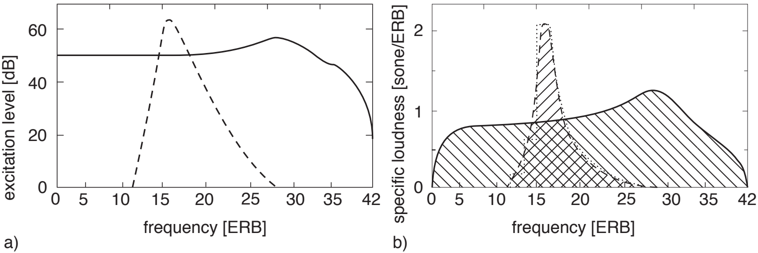
- Input signal spectrum $S(f)$ is warped to auditory frequency scale z
- $S'(z) = S[f(z)]\frac{df}{dz}$
- Signal also spreads in frequency due to frequency masking, $B(z)$ is spreading function
- $E(z) = S'(z) \star B(z)$
- Compute specific loudness $N'(z)$, kind of loudness function over frequency
- $N'(z) = c\,E(z)^{0.23}$

A theoretic view of loudness process

- Input signal spectrum $S(f)$ is warped to auditory frequency scale z
- $S'(z) = S[f(z)]\frac{df}{dz}$
- Signal also spreads in frequency due to frequency masking, $B(z)$ is spreading function
- $E(z) = S'(z) \star B(z)$
- Compute specific loudness $N'(z)$, kind of loudness function over frequency
- $N'(z) = c\,E(z)^{0.23}$
- Integrate over frequency for loudness N
- $N = \int_0^M N'(z)dz$

Excitation pattern and specific loudness

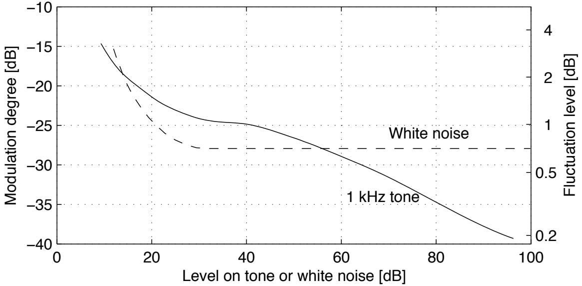
- a) excitation patterns. b) Specific loudness.
- (dashed) sinusoid, (continuous) noise



Adapted from Fastl and Zwicker (2007)

Difference threshold of loudness

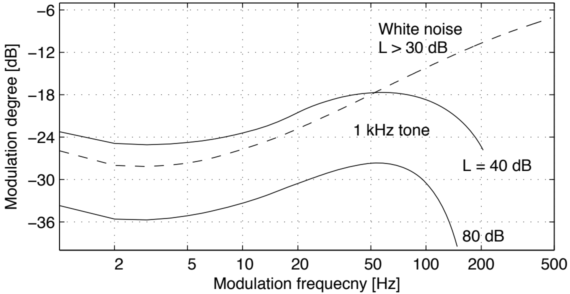
- The just noticeable level of amplitude modulation, about 1 dB with noise
- Why 1kHz value decreases continuously? Similar FM-tone JND result did not show this kind of result.



Adapted from Fastl and Zwicker (2007)

JND threshold of amplitude modulation

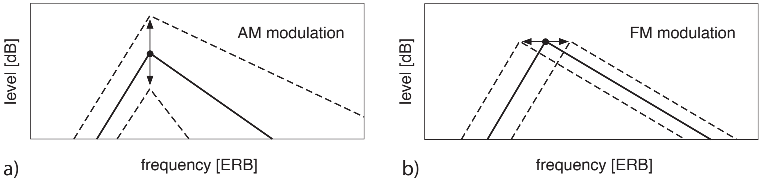
- Curves for tones with two levels and noise



Adapted from Fastl and Zwicker (2007)

Difference threshold of loudness

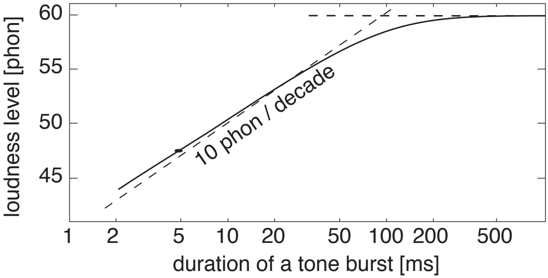
- AM causes periodic change of width of excitation pattern, especially at higher levels
- With FM this is not available
- Explains why larger level causes smaller difference thresholds



Adapted from Fastl and Zwicker (2007)

Loudness vs duration of sound

- The dependence of loudness level on duration
- Tone burst with frequency of 2kHz and a sound pressure level of 57dB



Adapted from Fastl and Zwicker (2007)

Timbre

- When two sounds have the same pitch, loudness, and duration, timbre is what makes one particular sound different from another.

Timbre

- When two sounds have the same pitch, loudness, and duration, timbre is what makes one particular sound different from another.
- Humans recognize the sound source mostly with timbre
- Closest physical explanation is magnitude spectrum and its variation with time

Timbre

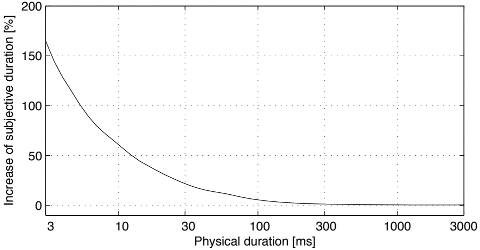
- When two sounds have the same pitch, loudness, and duration, timbre is what makes one particular sound different from another.
- Humans recognize the sound source mostly with timbre
- Closest physical explanation is magnitude spectrum and its variation with time
- Also phase spectrum has an effect

Timbre

- When two sounds have the same pitch, loudness, and duration, timbre is what makes one particular sound different from another.
- Humans recognize the sound source mostly with timbre
- Closest physical explanation is magnitude spectrum and its variation with time
- Also phase spectrum has an effect
- Complex phenomenon, not well understood or modeled
- Simple models based on (frequency-)specific loudness explain only steady noise-like sounds

Perceived duration of sound

- 1-kHz tone at an SPL of 60 dB with duration shown in x-axis
- Adjust the duration to "twice" or "half"
- Subjective duration [dura]



Adapted from Fastl and Zwicker (2007)

References

These slides follow corresponding chapter in: Pulkki, V. and Karjalainen, M. Communication Acoustics: An Introduction to Speech, Audio and Psychoacoustics. John Wiley & Sons, 2015, *where also a more complete list of references can be found.*

References used in figures:

Burck, W., Kotowski, P., and Lichte, H. (1935) Die horbarkeit von laufzeitdifferenzen. Elek. Nachr.-Techn., 12, 355 362.

Fastl, H. and Stoll, G. (1979) Scaling of pitch strength. Hearing Res., 1(4), 293 301.

Fastl, H. and Zwicker, E. (2007) Psychoacoustics – Facts and Models. Springer.

Sek, A. and Moore, B.C. (1995) Frequency discrimination as a function of frequency, measured in several ways. J. Acoust. Soc. Am., 97, 2479 2486.

Canteretta, E.C. and Fridman, M.P. (eds)(1978) Handbook of Perception. Academic Press.