Aalto University
School of Electrical Engineering

Communication acoustics
Ch 11: Further analysis in hearing

Ville Pulkki and Matti Karjalainen

Department of Signal Processing and Acoustics
Aalto University, Finland

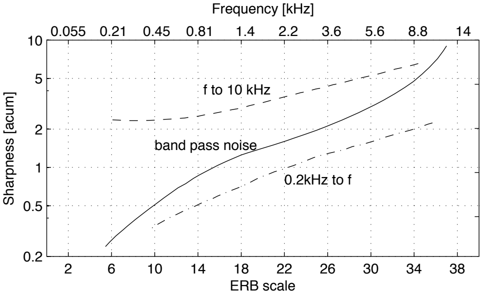
September 23, 2021

This chapter

- Subcategories of timbre
  - Sharpness
  - Fluctuation strength
  - Impulsiveness
  - Roughness
  - Tonality
- Sensitivity to magnitude and phase spectra
- Music
- Perceptual organization of sound

Sharpness

- "How sharp is the sound?"
- Sharpness of narrowband noise (solid line), high-pass filtered noise (upper cutoff is at 10 kHz), and low-pass filtered noise (lower cutoff is at 200 Hz)
- Unit: [acum]

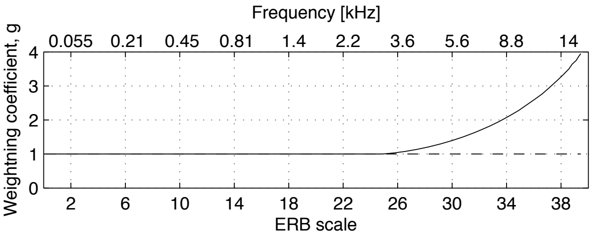


Adapted from Fastl and Zwicker (2007)

Sharpness, modeling

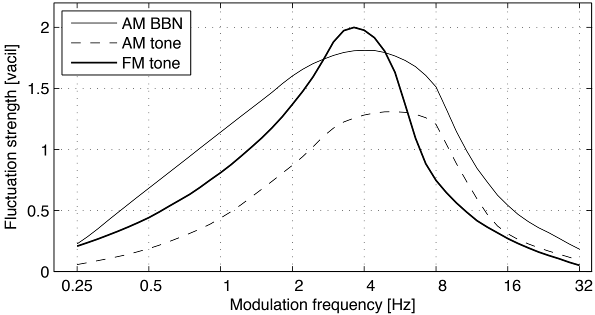
- Sharpness can be modeled as weighted average of specific loudness (auditory spectrum)

$$S = 0.11 \frac{\int_0^{42\,\mathrm{ERB}} N'(z)\,g(z)\,z\,dz}{\int_0^{42\,\mathrm{ERB}} N'(z) dz}, \tag{1}$$



Fluctuation strength

- Amplitude and frequency modulation of sounds cause perception of "fluctuation"



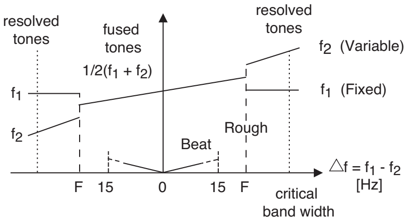
Adapted from Fastl and Zwicker (2007)

Impulsiveness

- There is no clearly defined psychoacoustic concept of impulsiveness
- Impulsiveness is related to rapid onsets in signal
- If the repetition rate of impulses is $>$ 10-15 Hz, roughness is perceived
- In noise control, impulsiveness is considered to increase hearing damage risk compared to non-impulsive sound of same energy

Roughness

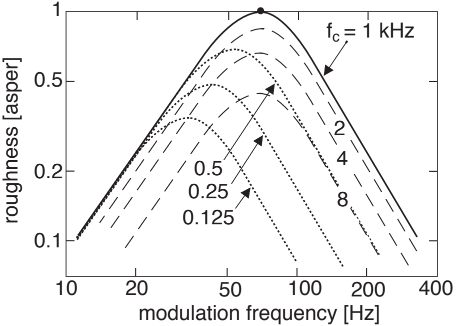
- Fast (> 15 Hz) modulation is perceived as roughness [asper]
- Addition of two tones of different frequencies creates envelope fluctuation
- Unit of roughness is asper
- Roughness of 1kHz tone, 60dB, 100% AM modulated at 70 Hz equals to 1 asper.
- When the frequency difference increases, tones start to segregate
- When the frequency difference is larger than a critical band, roughness disappears



Adapted from Roederer (1975)

Roughness

- Roughness for different carrier frequencies as a function of AM modulation frequency with 100 % modulation.



Adapted from Fastl and Zwicker (2007)

Tonality

- Tonality (tonalness) = sound exhibits voiced component(s), periodicity
- Non-tonal sound is noise-like, non-periodic
- Do not mix with musical term "tonal"
- Non-tonal (noisy) signal masks a tonal one more easily than vice versa
- Measurement necessary especially in lossy audio coding
- Tonality with varying modal density, log. distribution of frequencies (approx/critical band):

Sounds with N partials per critical band:
10 / CB 20 / CB 40 / CB 80 / CB

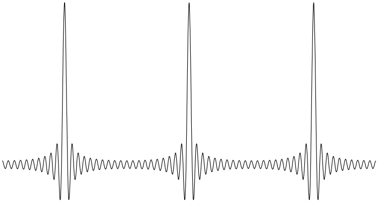
Sensitivity to magnitude spectrum

- Basic: magnitude spectrum of ear canal signal analyzed by the cochlea
- Complex adaptation processes
- Listeners try to adapt actively to acoustic transmission channel
- Similarly to visual after-effect, a negative picture is seen on black background after looking intesively to an image
- Adaptation in periphery?
- Adaptation in central brain processing?
- Mechanisms are not well known

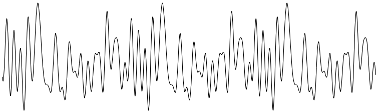
Sensitivity to phase spectrum

- Noise-like signals: no sensitivity
- Many harmonic signals: no sensitivity
- Certain "peaky" signals are sensitive
- Vowel voices, trombone, trumpet, sawtooth wave, impulse train

Zero-phase cosine signal



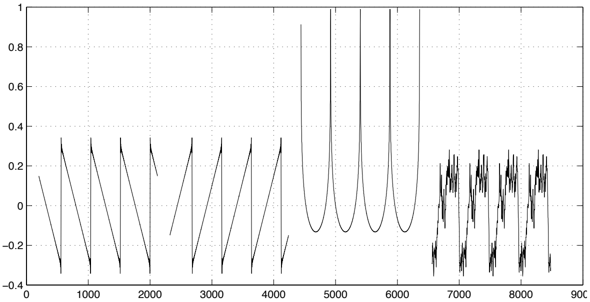
Random-phase version



Sensitivity to phase spectrum

- Effects can not be listened to in rooms
- Room reverberation destroys anyway phase spectrum
- Headphone listening, depends also on phase response of headphones
- Time-domain peaks in signal may lead into "buzzyness"
- Low-frequency phase alteration changes perceived level of bass, "bassiness"
- Seems to be a distance cue in human localization process

Examples

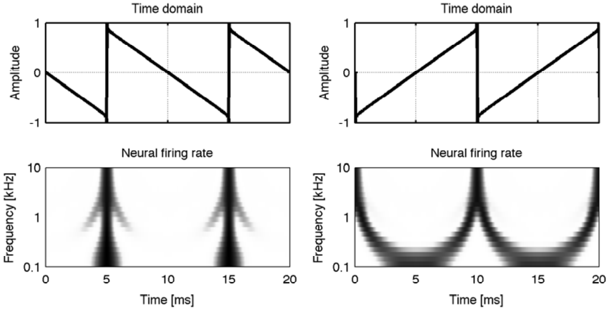


Sound examples. $f_0 = 100$Hz, 100 partials (100...10000Hz).

Saw   Saw upside down   Pulses   Random phase

Response of cochlea to sawtooth with phase modifications

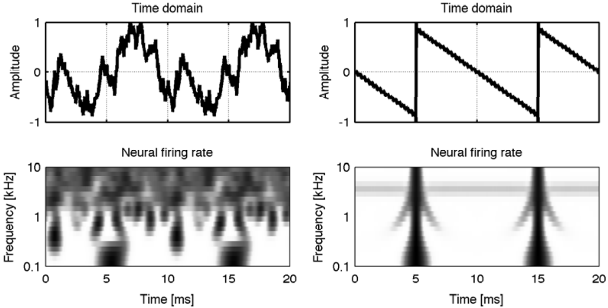
- Sawtooth, Sawtooth time-inverted



Adapted from Laitinen et al. 2013

Response of cochlea to sawtooth with phase modifications

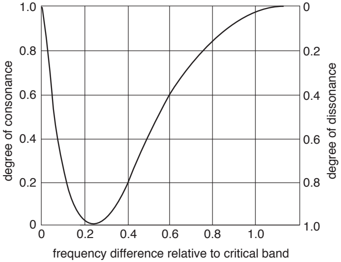
- Random phase, 3kHz sine polarity inverted



Adapted from Laitinen et al. 2013

Consonance and dissonance of harmonic tones

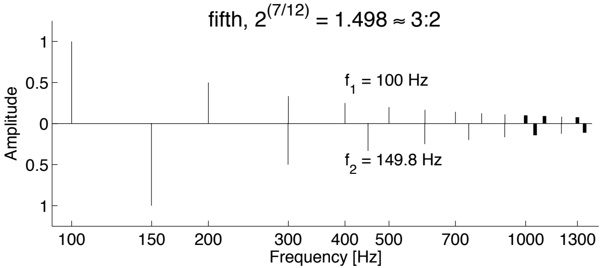
- Roughness due to interaction of partials in a sound contribute to dissonance
- Ratios of small integers are most consonant (just intonation)
- Starting point: Consonance vs. dissonance of two sinusoids



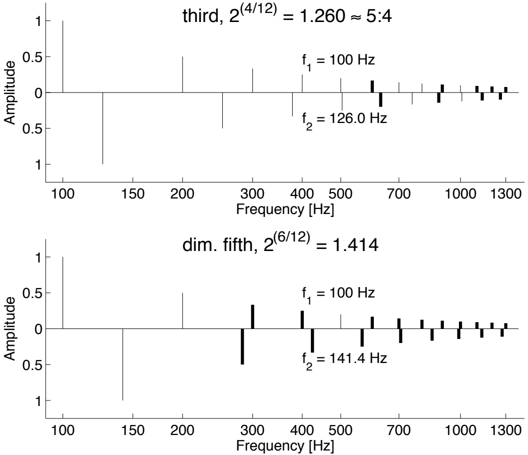
Adapted from Plomp and Levelt (1965)

Consonance and dissonance of harmonic tones

- A harmonic tone contains a number of partials
- If the partials of different tones are too close -> added dissonance

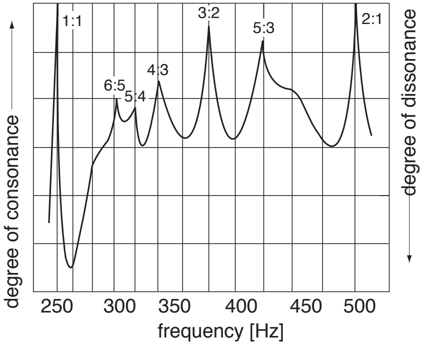


Consonance and dissonance of harmonic tones



Consonance and dissonance of harmonic tones

- Level of consonance depending on separation btw two harmonic tone complexes



Adapted from Plomp and Levelt (1965)

Consonance and dissonance of harmonic tones

- Demos with different intervals

Fifth 3/2 Fifth $\sqrt[(7/12)]{2}$

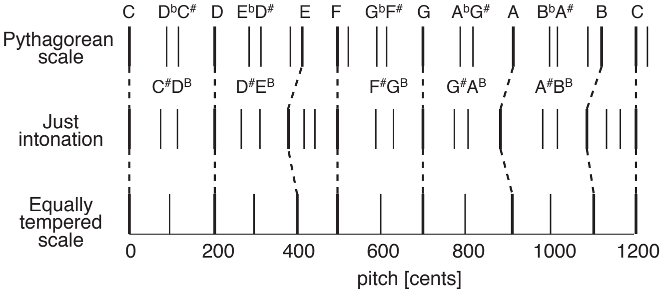
Dim. fifth $\sqrt[(6/12)]{2}$

Fourth 4/3 Fourth 4/3

Third 5/4 Third $\sqrt[(5/12)]{2}$

Scales, tuning

- Just intonation, triads tuned as 4:5:6 and octaves 2:1
- Pythagorean scale, greatest number of pure fifths and octaves.
- equal temperament, pure octave is divided into 12 semitones having frequency ratio $\sqrt[12]{2} \approx 1.05946$



Adopted from Rossing et al., 2001

Scales, tuning

	do	re	mi	fa	so	la	ti	do2
Just:	100							200
Pyth:	100							200
Equ:	100							200

Scales, tuning

	do	re	mi	fa	so	la	ti	do2
Just:	100							200
Pyth:	100							200
Equ:	100							200

	do	re	mi	fa	so	la	ti	do2	
Just:	100		125					200	do-mi
Pyth:	100				150			200	do-so
Equ:	100	112.2						200	do-(di)-re

Scales, tuning

	do	re	mi	fa	so	la	ti	do2
Just:	100							200
Pyth:	100							200
Equ:	100							200

	do	re	mi	fa	so	la	ti	do2	
Just:	100		125					200	do-mi
Pyth:	100				150			200	do-so
Equ:	100	112.2						200	do-(di)-re

	do	re	mi	fa	so	la	ti	do2	re2	
Just:	100		125		150			200		do-so
Pyth:	100	112.5			150			200	225	so - re2-re
Equ:	100	112.2	126.0					200		re-(ri)-mi

Scales, tuning

	do	re	mi	fa	so	la	ti	do2	re2	mi2	
J	100	112.5	125		150		187.5	200	225		so-ti-re2
P	100	112.5	126.6		150	168.8		200	225	253.2	re-la-mi2
E	100	112.2	126.0	133.5	149.8			200			mi — so

Scales, tuning

	do	re	mi	fa	so	la	ti	do2	re2	mi2	
J	100	112.5	125		150		187.5	200	225		so-ti-re2
P	100	112.5	126.6		150	168.8		200	225	253.2	re-la-mi2
E	100	112.2	126.0	133.5	149.8			200			mi — so

	do	re	mi	fa	so	la	ti	do2	re2	mi2	
J	100	112.5	125	133.3	150	160	187.5	200	225		do2-la-fa
P	100	112.5	126.6	133.3	150	168.8	189.9	200	225	253.2	mi-ti, do2-fa
E	100	112.2	126.0	133.5	149.8	168.2	188.8	200			so-ti

Rhythm

- Rhythm is a complex concept which refers to different temporal structures in music
- Heart beat, walking
- Some concepts in rhythm
  - Note value, length of note in time
  - Measure or bar: A rhythmic 'placeholder' which indicates a prototype repeated rhythm in music
  - Tempo: The speed of presentation
  - Beat: The accenting of specific temporal positions in a bar.
- Not very well understood dimension of music

Perceptual organization of sound

- The hearing mechanism involves certain inborn capabilities to analyse the summed sounds of the auditory environment arriving from multiple sources with or without room reflections and reverberation
- Auditory events are connected to internal representations of sources based on many cues
- Spectral, temporal cues
- Direction, distance cues

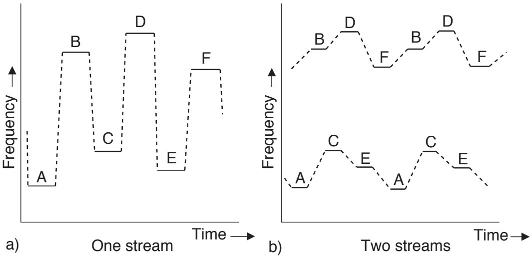
Pattern formation

Gestalt laws of grouping

- *Principle of proximity.*
- *Principle of similarity.*
- *Principle of closure.* In the case of a pure tone being interrupted sequentially by bursts of white noise, the human auditory system assumes the pure tone continues uninterrupted during the noise bursts.
- *Principle of continuity.* For example, the smooth pitch variations and smooth formant changes in speech imply to the listener that the speech originates from the same speaker and is organized into a single stream.
- *Principle of common motion.* If sensory elements move in the same direction at the same rate, they tend to be grouped as parts of a single stimulus.
- *Principle of belongingness.*

Sound Streaming and Auditory Scene Analysis

- Formation of melody line
- With slow tempo, notes with very large intervals are bound to single stream
- With fast tempo, several streams are formed with same notes



Demos

Slow tempo
Fast tempo

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:

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

Plomp, R. and Levelt, W.J. (1965) Tonal consonance and critical bandwidth. J. Acoust. Soc. Am., 38, 548-560.

Roederer, J.G. (1975) The Physics and Psychophysics of Music: An Introduction. Springer.

Rossing, T.D., Moore, F.R., and Wheeler, P.A. (2001) The Science of Sound, 3rd edn. Addison-Wesley.