Psychoacoustical Models
WS 2016/17

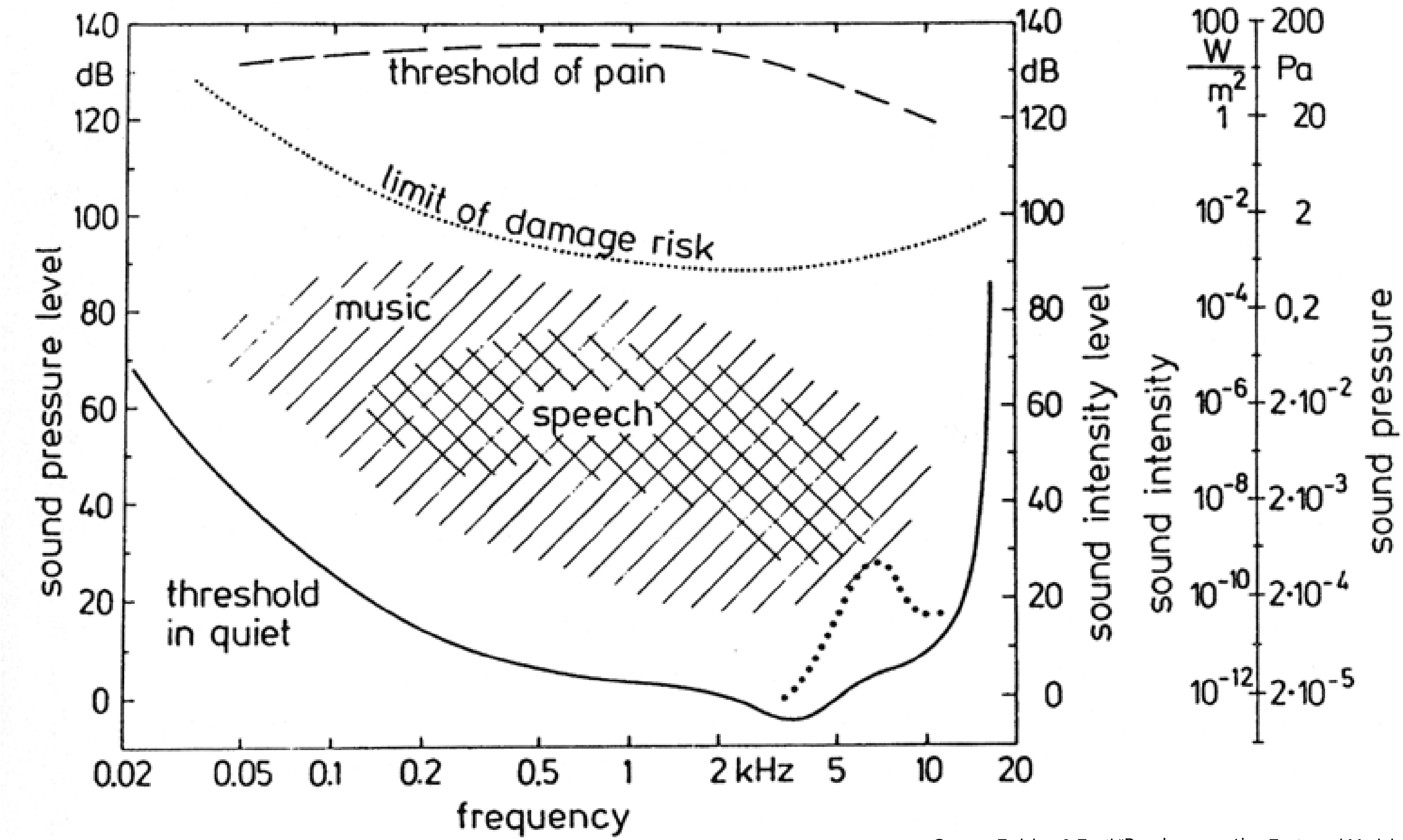
related lectures:
Applied and Virtual Acoustics (Winter Term)
Advanced Psychoacoustics (Summer Term)

TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Sound Perception

Frequency and Level Range of Human Hearing



Source: Zwicker & Fastl “Psychoacoustics Facts and Models”

Threshold in Quiet or the Absolute Threshold

- Hearing threshold of 100 persons with normal hearing for sine tones (50% curve is the median)

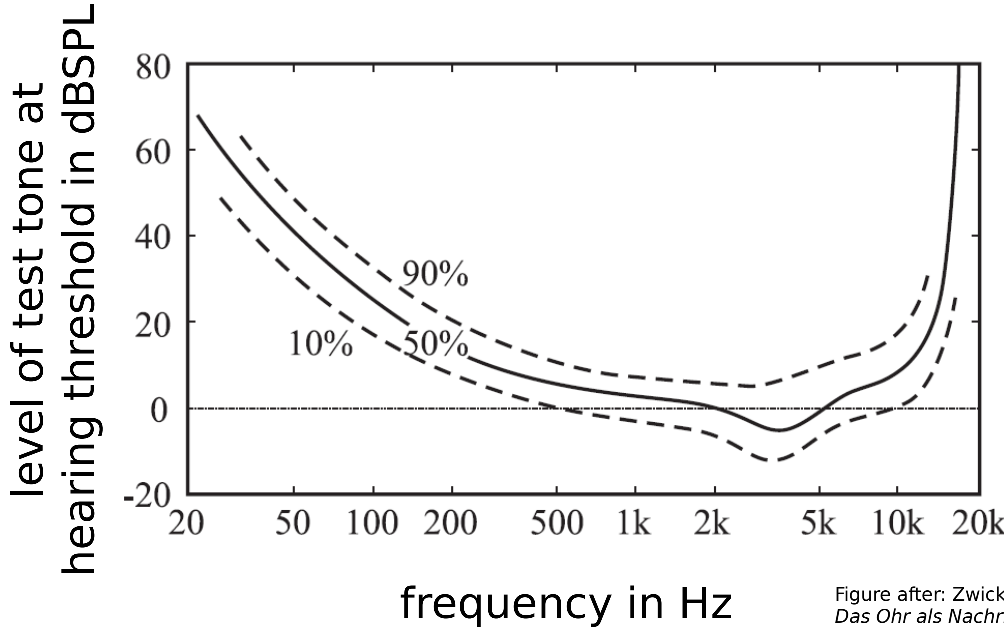


Figure after: Zwicker, E.; Feldtkeller, R. (1967). *Das Ohr als Nachrichtenempfänger*, Hirzel Verlag, Stuttgart.

- Approximations:

$$\frac{L_{Tq}}{dB} = 3.64\left(\frac{f}{kHz}\right)^{-0.8} - exp\left(-0.6\left(\frac{f}{kHz} - 3.3\right)^{2}\right) + 10^{-3}\left(\frac{f}{kHz}\right)^{4}$$

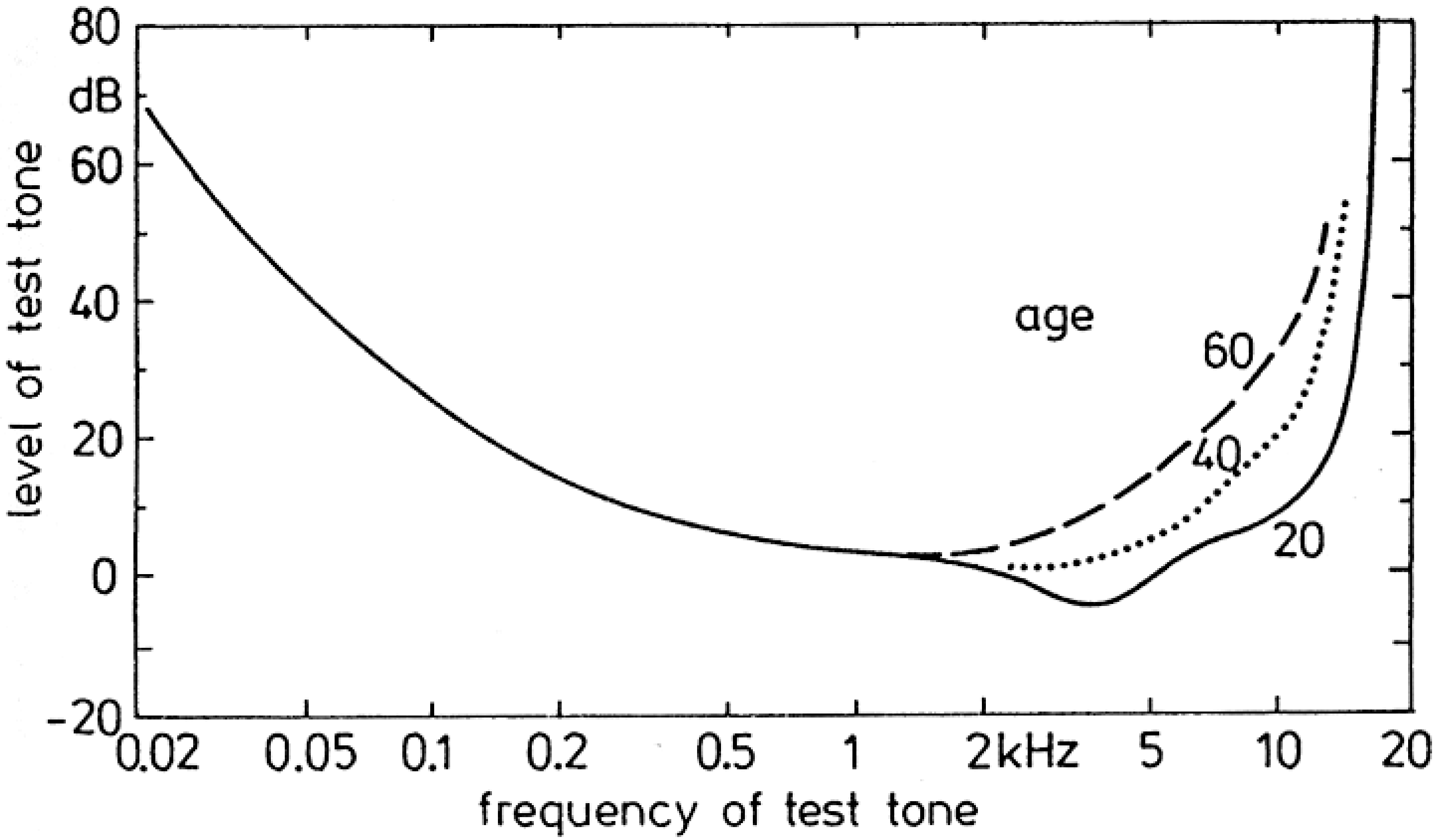
$$4\ kHz\ Signal\ with\ Amplitude\ \pm 1\ LSB\ (16\ bit)$$

Source: U. Zölzer, “Digital Audio Signal Processing”

TECHNISCHE UNIVERSITÄT ILMENAU

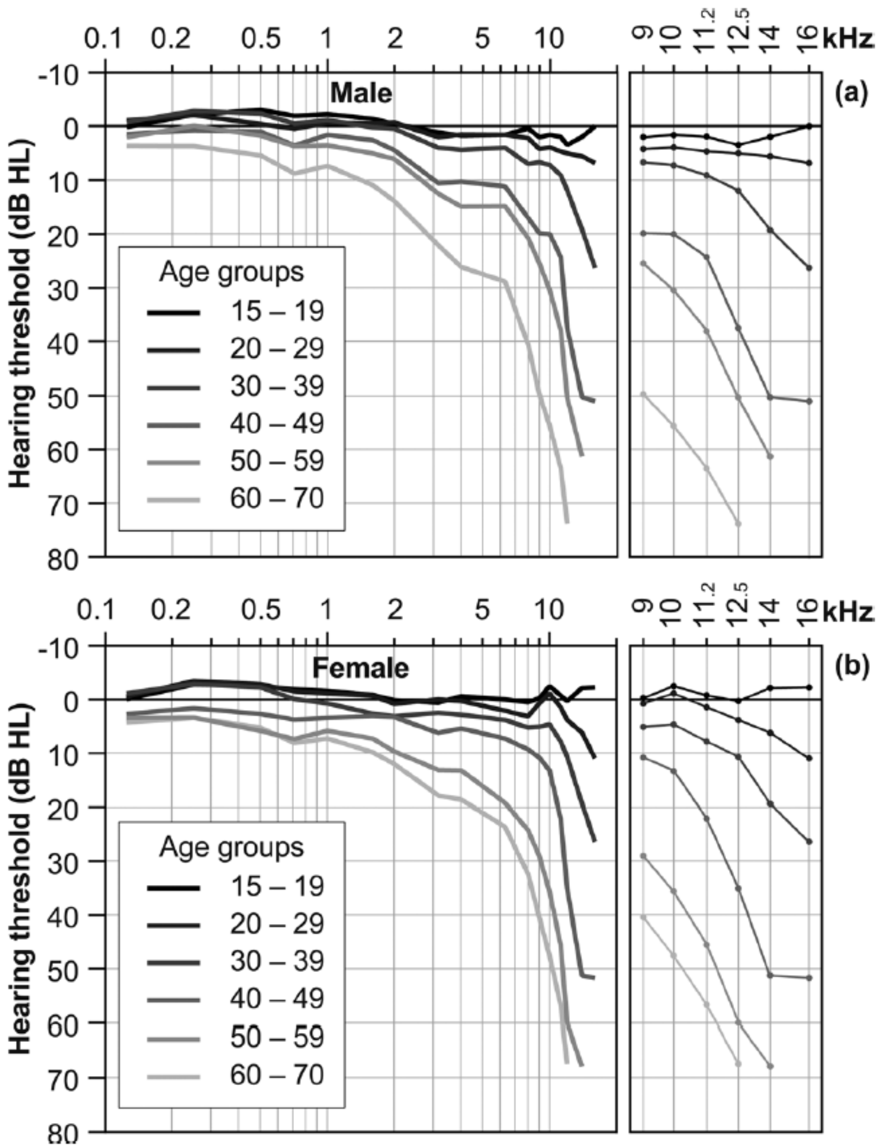
Fraunhofer IDMT

Threshold in Quiet or the Absolute Threshold



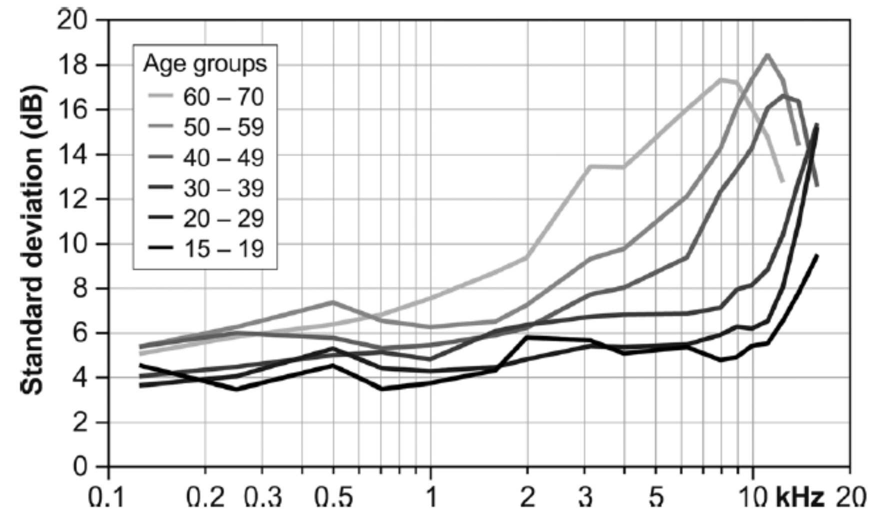
Source: Zwicker & Fastl “Psychoacoustics Facts and Models”

Hearing Threshold and age



Figures left: Average **pure-tone audiograms in dB HearingLoss** in (a) men and (b) women grouped by their age in decades (the parameter is age group in years). The extended high-frequency range is zoomed for clarity.

Figure right: Pure-tone **threshold standard deviation** of all participants as a function of frequency (the parameter is age in 10-year groups).



Figures: Milan Jilek, Daniel Suta, and Josef Syka, „Reference hearing thresholds in an extended frequency range as a function of age “, J. Acoust. Soc. Am., Vol. 136, No. 4, October 2014

Loudness

- <u>Loudness Level:</u>

  - **Loudness N:** psychological concept to describe the <u>magnitude of an auditory sensation</u>, the loudness of a sound (measured in ‘sone’)

  - **loudness level L_N** of a sound is measured in ‘**phon**’
  - **L_N** of a sound is the sound pressure of a 1 kHz tone which is as loud as the sound

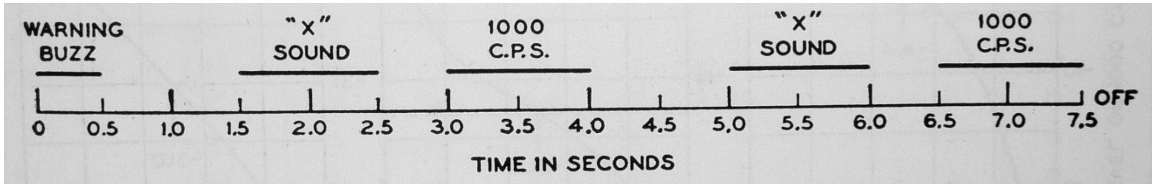


FIG. 129.—TIME SEQUENCE FOR LOUDNESS COMPARISONS.

Fig: Fletcher, “Speech and Hearing in Communication”, 1953.

Prof. Dr.-Ing. K. Brandenburg, bdg@idmt.fraunhofer.de Prof. Dr.-Ing. G. Schuller, shl@idmt.fraunhofer.de

TECHNISCHE UNIVERSITÄT ILMENAU
Fraunhofer IDMT

Loudness

- <u>Equal-Loudness Level Contours</u>

Equal loudness contours of *pure tones* in a *free sound field*.

The parameter is expressed in **loudness level**, L_N, and **loudness**, N.
Can be observed:

The sensitivity of the human ear - a function of frequency
The most sensitive to sounds around 2–4 kHz

Links to measure the sensitivity on different frequenc
  http://www.phys.unsw.edu.au/jw/hearing.html
  http://www.phys.unsw.edu.au/music/dB/loudness.html, 2010

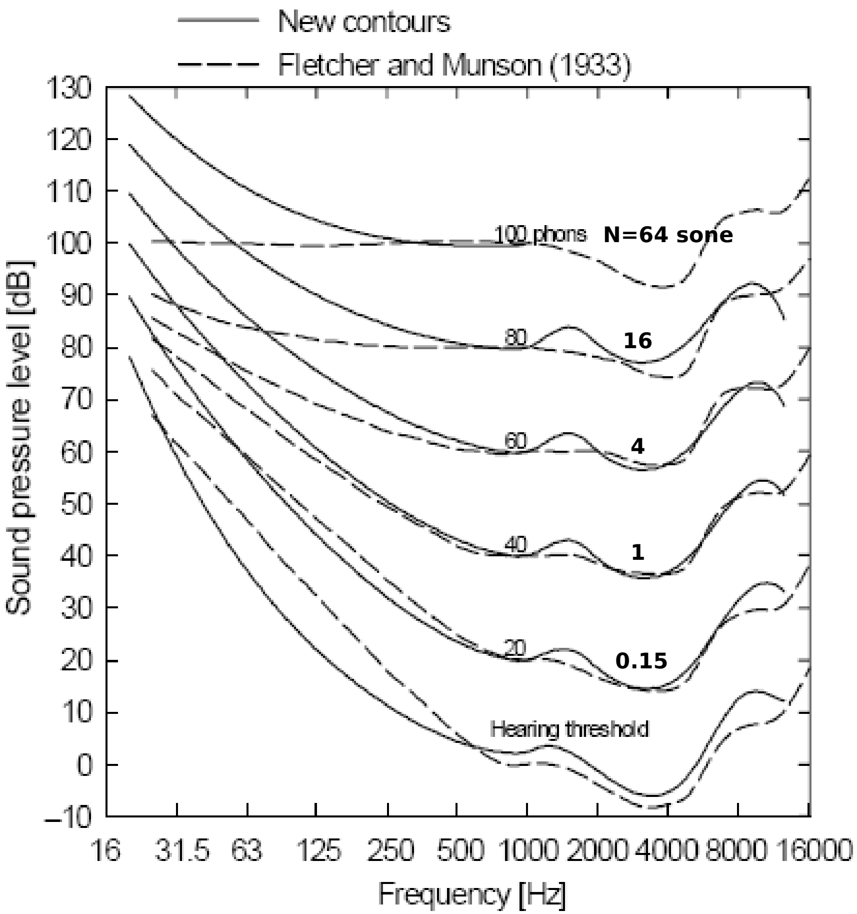


Fig: Suzuki et al., "Precise and Full-range Determination of Two-dimensional Equal Loudness Contours", 2003.

Loudness

- Loudness Scale:

  - aim: *double the number of units on this scale means magnitude of sensation is doubled*
    → relation between loudness level L_N and the loudness *N* (rule of thumb):
    $$2 \cdot N \mathrel{\widehat{=}} L_N + 10\,phon$$
  - one potential experiment:
    listen to a sound with L_{N1} and then adjust the same sound until $N_2 = 2 \cdot N_1$, then compare L_{N_1} and L_{N_2}

Prof. Dr.-Ing. K. Brandenburg, bdg@idmt.fraunhofer.de Prof. Dr.-Ing. G. Schuller, shl@idmt.fraunhofer.de

Loudness

- Loudness Scal

Source of sound	sound pressure	sound pressure level	loudness
	pascal	dB re 20 µPa	sone
threshold of pain	100	134	~ **676**
hearing damage during short-term effect	20	approx. 120	~ **256**
jet, 100 m distant	6 - 200	110 - 140	~ **128** - **1024**
jack hammer, 1 m distant / discotheque	2	approx. 100	~ **64**
hearing damage during long-term effect	6×10^{-1}	approx. 90	~ **32**
major road, 10 m distant	2×10^{-1} - 6×10^{-1}	80 - 90	~ **16** - **32**
passenger car, 10 m distant	2×10^{-2} - 2×10^{-1}	60 - 80	~ **4** - **16**
TV set at home level, 1 m distant	2×10^{-2}	ca. 60	~ **4**
normal talking, 1 m distant	2×10^{-3} - 2×10^{-2}	40 - 60	~ **1** - **4**
very calm room	2×10^{-4} - 6×10^{-4}	20 - 30	~ **0.15** - **0.4**
leaves' noise, calm breathing	6×10^{-5}	10	~ **0.02**
auditory threshold at 2 kHz	2×10^{-5}	0	**0**

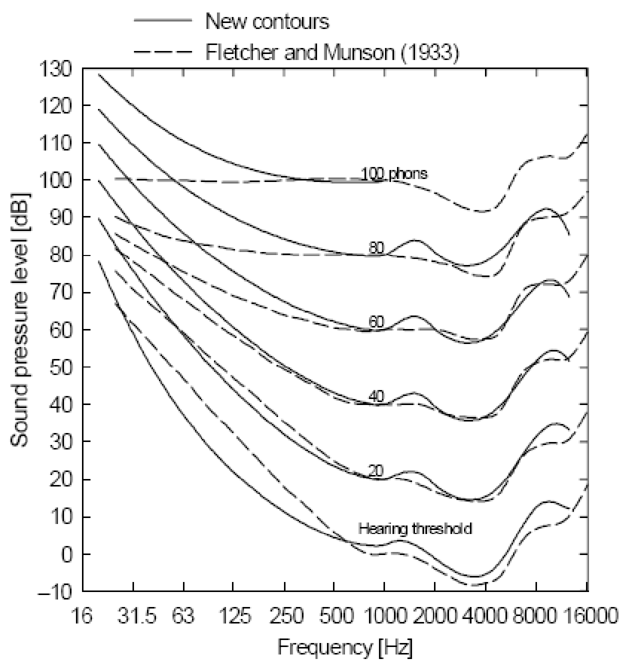
Fig: en.wikipedia.org/wiki/Sone, 2010.

Prof. Dr.-Ing. K. Brandenburg, bdg@idmt.fraunhofer.de Prof. Dr.-Ing. G. Schuller, shl@idmt.fraunhofer.de

TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Sound Examples



Example 1: Sensitivity of hearing as function of frequency

- Sweep from 0-16000 Hz (equal amplitude)

Example 2: Upper limit of hearing

- 8 kHz 10 kHz 12 kHz 14 kHz 16 kHz

- 17 kHz 18 kHz 19 kHz 20 kHz

Critical Bands

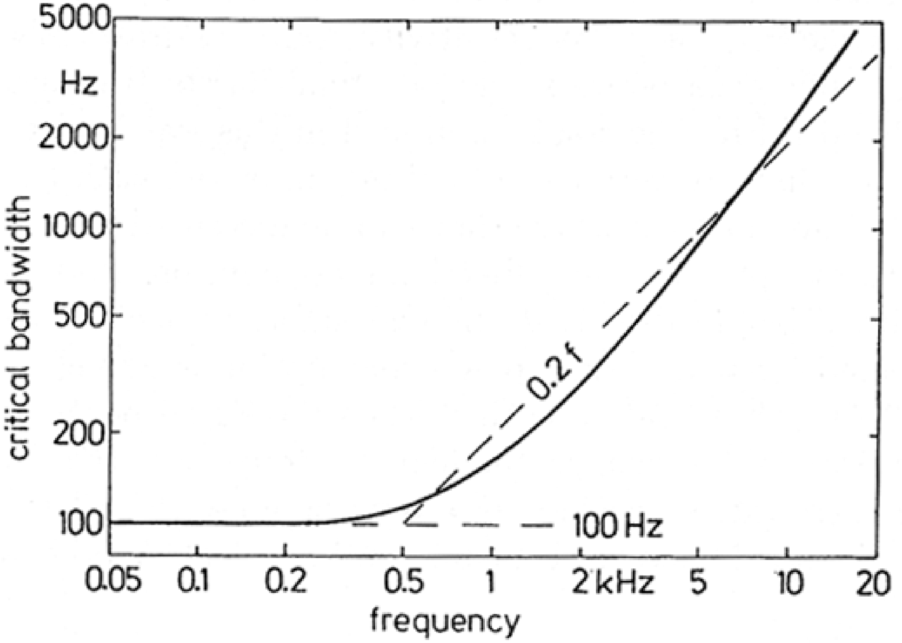
Prof. Dr.-Ing. K. Brandenburg, bdg@idmt.fraunhofer.de Prof. Dr.-Ing. G. Schuller, shl@idmt.fraunhofer.de

Frequency Grouping in Human Hearing

- Different interpretations that produce the same
  segmentation
  - Constant distance in the Cochlea
  - By using tones under the threshold in quiet,
    their intensity add up in a critical band and are now audible
  - Tones in a critical band above the threshold in quiet: their energy adds up

- Formula for the width of the critical bands
  - for frequencies < 500 Hz: Constant 100Hz
    width
  - for frequencies > 500 Hz: 0.2*frequency

Frequency Grouping Bandwidth – The Critical Bands



Source: Zwicker & Fastl "Psychoacoustics Facts and Models"

Critical bandwidth as a function of frequency, that quantifies the cochlear filter passbands.

Approximations for low and high frequency ranges are indicated by broken lines.

Prof. Dr.-Ing. K. Brandenburg, bdg@idmt.fraunhofer.de Prof. Dr.-Ing. G. Schuller, shl@idmt.fraunhofer.de

TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Excursus - Critical Bands and Loudness

- <u>Spectral effects - influence of frequency separation:</u>
  - measure the loudness level (or level of the equally loud 1 kHz tone) of 2 tones by varying the frequency separation

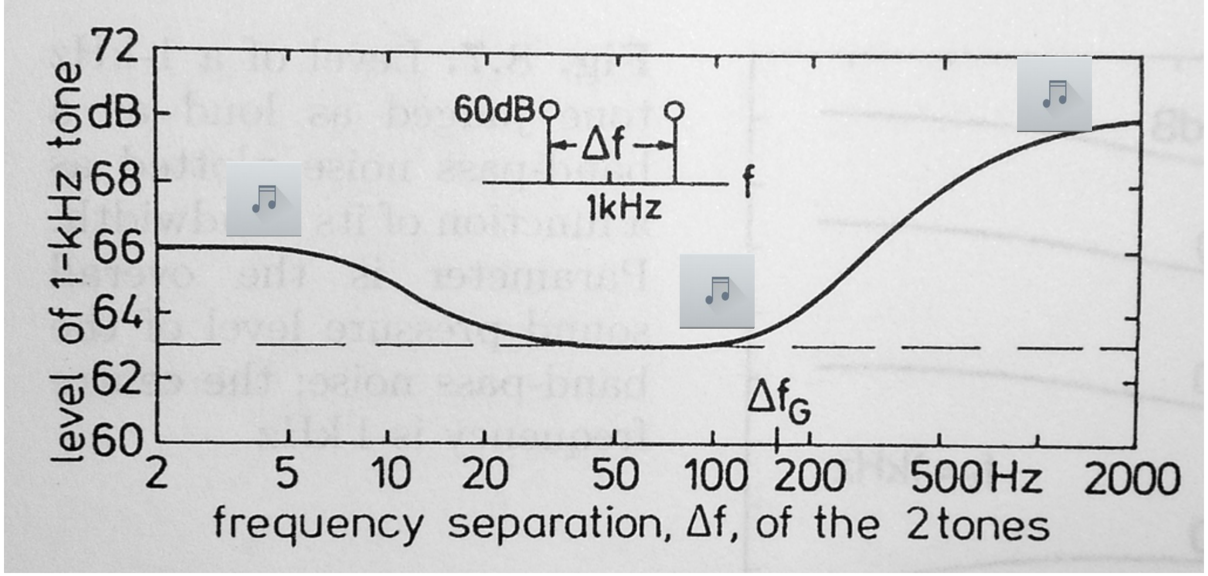
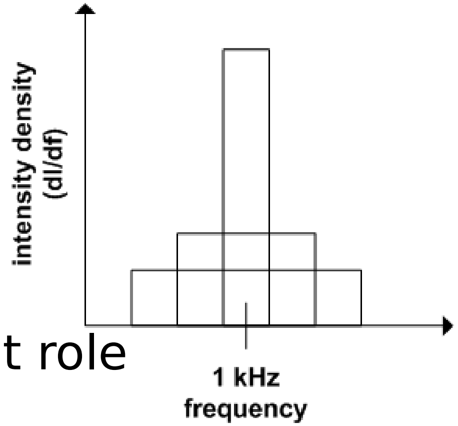


Fig: Zwicker, Fastl “Psychoacoustics - Facts and Models”, 2nd Edition, 1999.

Excursus - Critical Bands and Loudness



- Spectral effects - influence of bandwidth:
  - bandwidth of the signals plays an important role
  - sound level also influence loudness level
  - ✉ total sound intensity (SPL) have to be constant to measure loudness as function of bandwidth
  - ✉ critical bandwidth

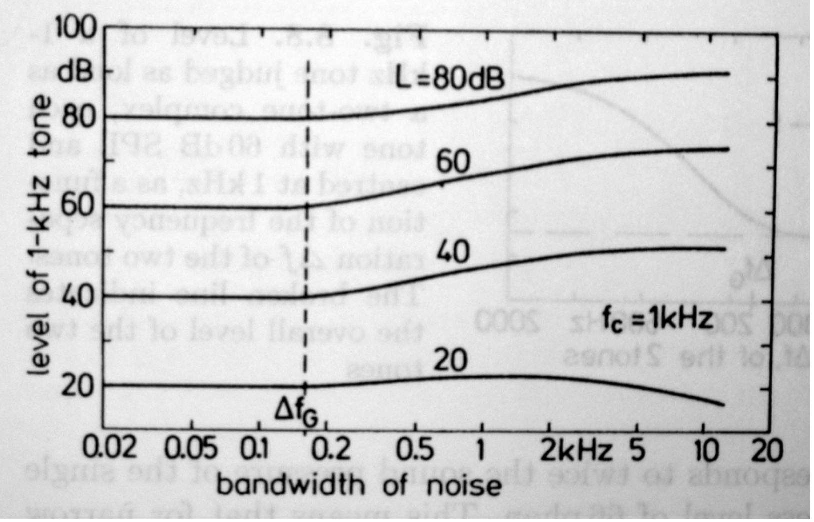


Fig: Zwicker, Fastl “Psychoacoustics - Facts and Models”, 2nd Edition, 1999.

Prof. Dr.-Ing. K. Brandenburg, bdg@idmt.fraunhofer.de Prof. Dr.-Ing. G. Schuller, shl@idmt.fraunhofer.de

TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Critical Bands: Bark Scale

- Critical-band concept used in many models and hypothesis
- unit was defined leading to so-called critical-band rate scale
- scale ranging from 0 – 24, unit “Bark”

- relation between *z* and *f* is important for under-standing many character-istics of human ear

z/Bark	f_u/Hz	f_o/Hz	Δf_G/Hz	f_m/Hz
0	0	100	100	50
1	100	200	100	150
2	200	300	100	250
3	300	400	100	350
4	400	510	110	450
5	510	630	120	570
6	630	770	140	700
7	770	920	150	840
8	920	1080	160	1000
9	1080	1270	190	1170
10	1270	1480	210	1370
11	1480	1720	240	1600
12	1720	2000	280	1850
13	2000	2320	320	2150
14	2320	2700	380	2500
15	2700	3150	450	2900
16	3150	3700	550	3400
17	3700	4400	700	4000
18	4400	5300	900	4800
19	5300	6400	1100	5800
20	6400	7700	1300	7000
21	7700	9500	1800	8500
22	9500	1200	2500	10500
23	12000	15500	3500	13500
24	15500			

Critical Bands: Bark Scale

- Critical-band concept used in many models and hypotheses
  →unit was defined leading to so-called critical-band rate scale
- scale ranging from 0 – 24, unit "Bark" (after Zwicker)
- One Bark corresponds to one critical band
- Attempt to approximate critical bands with formulas:

Critical Bandrate z:

$$\frac{z}{Bark} = 13 \arctan\left(0.76 \frac{f}{kHz}\right) + 3.5 \arctan\left(\frac{f}{7.5\ kHz}\right)^2$$

Critical Bandwidth:

$$\Delta f_B = 25 + 75\left(1 + 1.4\left(\frac{f}{kHz}\right)^2\right)^{0.69}$$

Prof. Dr.-Ing. K. Brandenburg, bdg@idmt.fraunhofer.de Prof. Dr.-Ing. G. Schuller, shl@idmt.fraunhofer.de

Masking

Prof. Dr.-Ing. K. Brandenburg, bdg@idmt.fraunhofer.de Prof. Dr.-Ing. G. Schuller, shl@idmt.fraunhofer.de

TECHNISCHE UNIVERSITÄT ILMENAU
Fraunhofer IDMT

Masking

- data compression
  - exploitation of perception in critical bands with reference to the threshold in quiet is not enough

- Basic principle:
  - a **test signal**, called a **maskee** is placed at the center frequency of the critical bandwidth
  - one **masking signal**, called **masker** (equal power and distance from maskee)
  - If the P_{maskee} is weak relative to the total power of the maskers ✉ the test signal is not audible ✉ test signal is masked
  - In order for the test signal to become audible, its power has to be raised to above a certain level – **masking threshold**.

Prof. Dr.-Ing. K. Brandenburg, bdg@idmt.fraunhofer.de Prof. Dr.-Ing. G. Schuller, shl@idmt.fraunhofer.de

TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Masking of Pure Tones by Noise - Broad-Band Noise

- broad-band noise:
  - white noise from 20 Hz - 20 kHz

- figure:
  - masking threshold for pure tones masked by broad band noise of different levels

  - uniform masking noise (**UMN**) by equalization of the 10 dB per decade slope

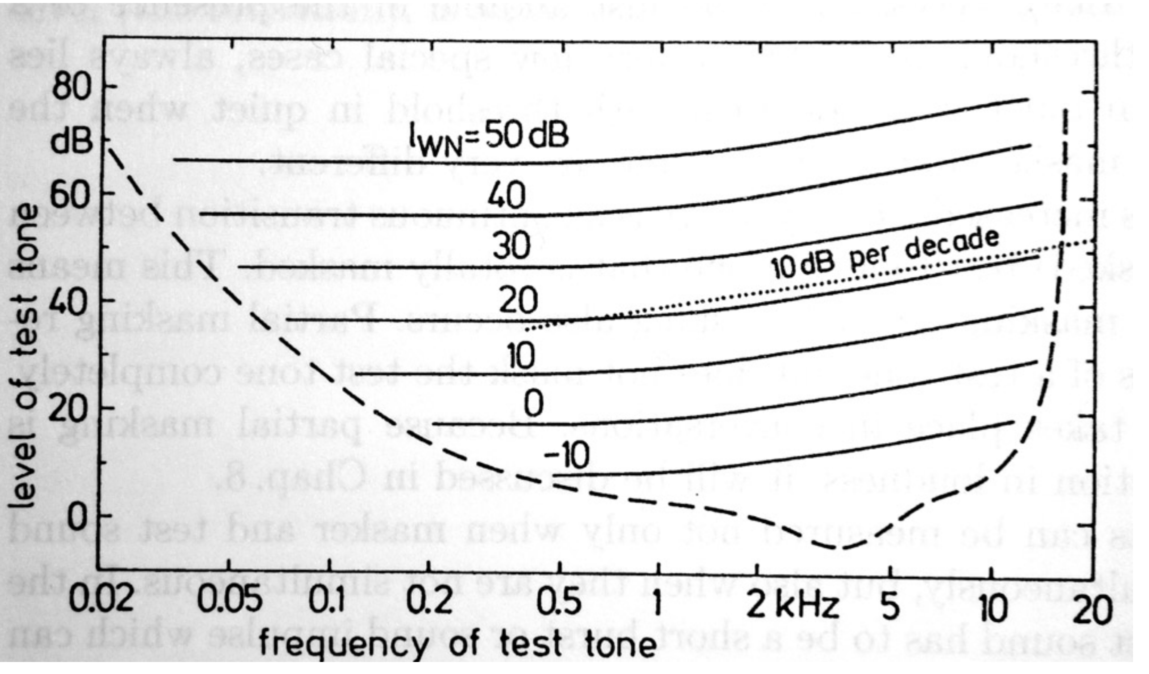


Fig: Zwicker, Fastl "Psychoacoustics - Facts and Models", 2nd Edition, 1999.

Masking of Pure Tones by Noise - Narrow-Band Noise

- narrow-band noise:
  - noise with a bandwidth equal or smaller than critical bandwidth

- figure:
  - threshold of pure tones masked by narrow-band noise for different centre frequencies

  - difference between maximum of masked threshold and test tone level

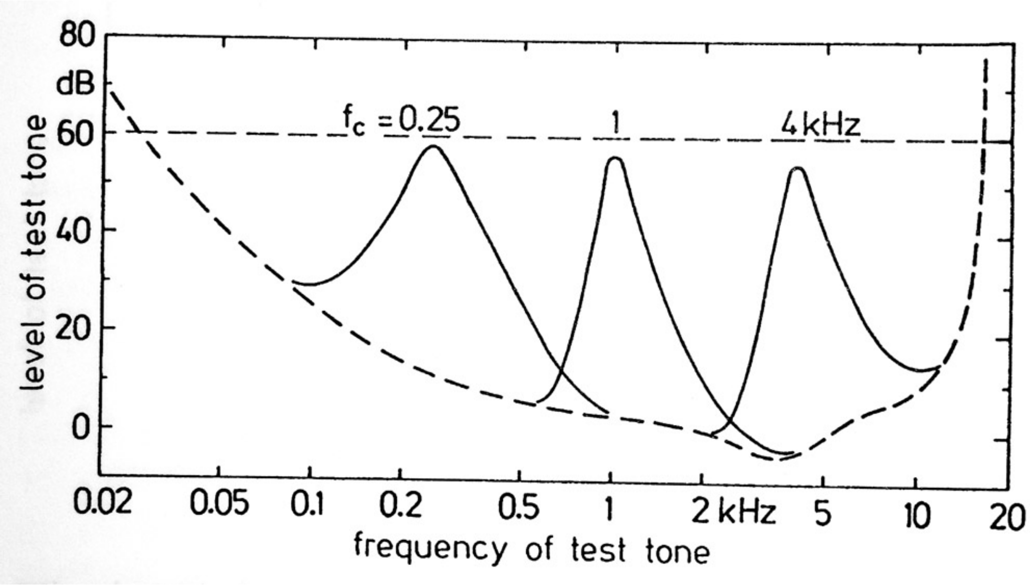


Fig: Zwicker, Fastl “Psychoacoustics - Facts and Models”, 2nd Edition, 1999.

Masking of Pure Tones by Noise - Narrow-Band Noise

- narrow-band noise:
  - noise with a bandwidth equal or smaller than critical bandwidth

- figure:
  - dependence of masked threshold on level of narrow-band noise

  - dips at higher levels ✉ nonlinear effects (difference noise caused by interactions between test tone and noise)

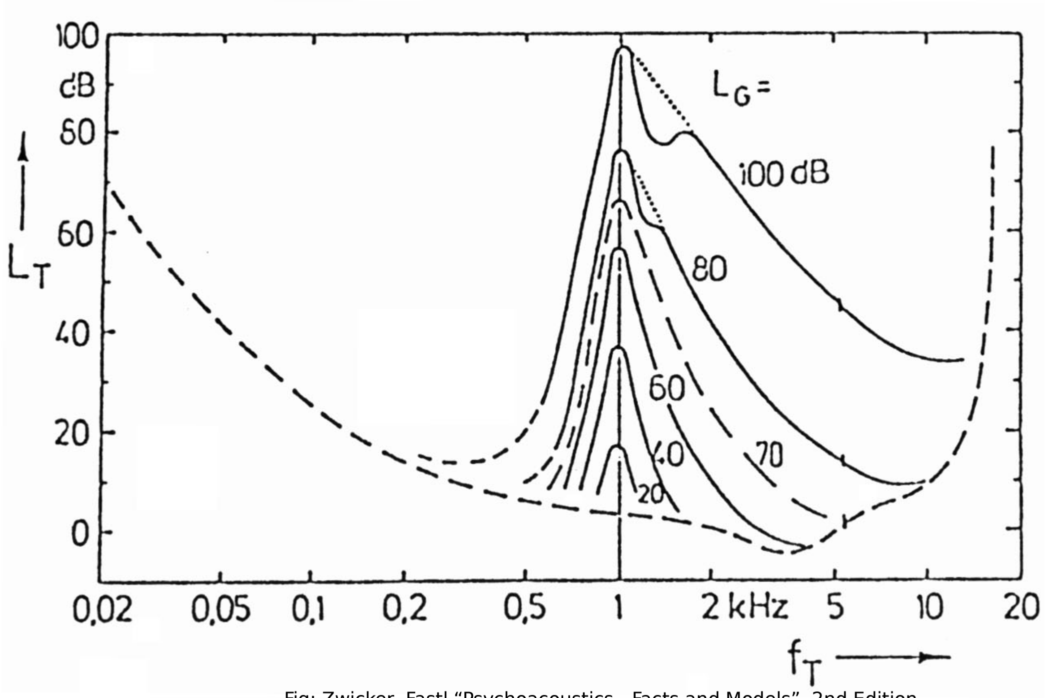


Fig: Zwicker, Fastl "Psychoacoustics - Facts and Models", 2nd Edition, 1999.

Prof. Dr.-Ing. K. Brandenburg, bdg@idmt.fraunhofer.de Prof. Dr.-Ing. G. Schuller, shl@idmt.fraunhofer.de

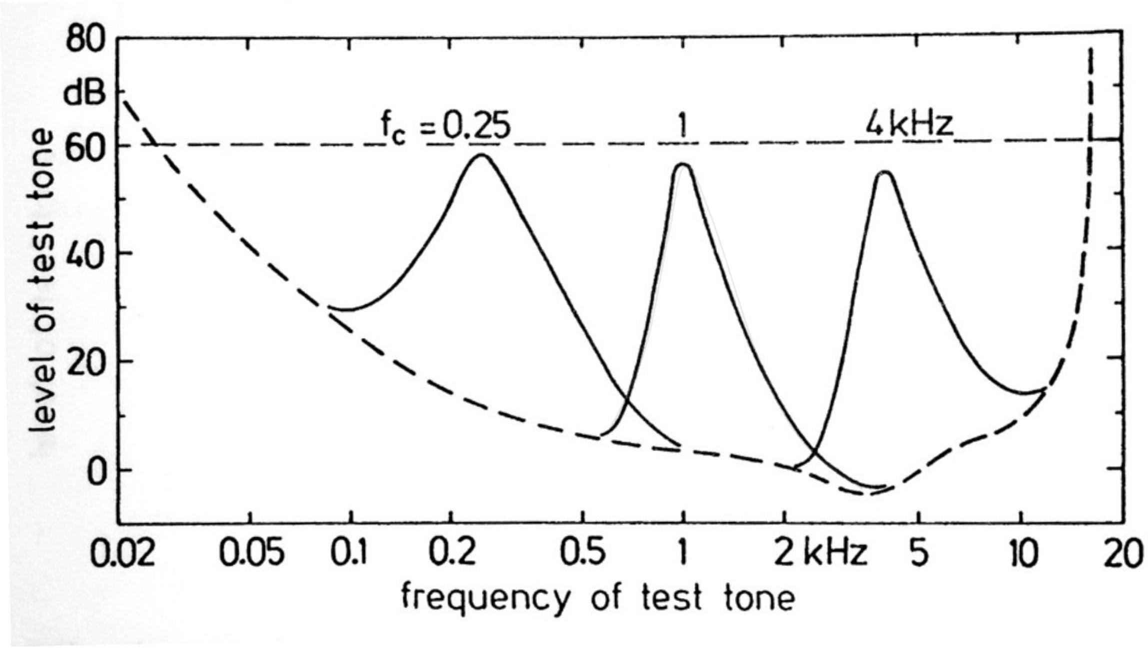
TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Test: Narrow Band Noise Masking Tone

Example 3:

- Narrow Band Noise at 1000 Hz, width 160 Hz;
- Sine tones at 600, 800, 1000, 1200, 1400, 1600 Hz at varying levels (-80 to -20 dB)



TECHNISCHE UNIVERSITÄT ILMENAU

Fraunhofer IDMT

Sound Examples: Masking with White Noise

Example 4: Masking with white noise

- 500 Hz sinusoid tone at varying amplitude ALONE
- Level: -40,-35,-30,-25,-20,-15,-10 dB

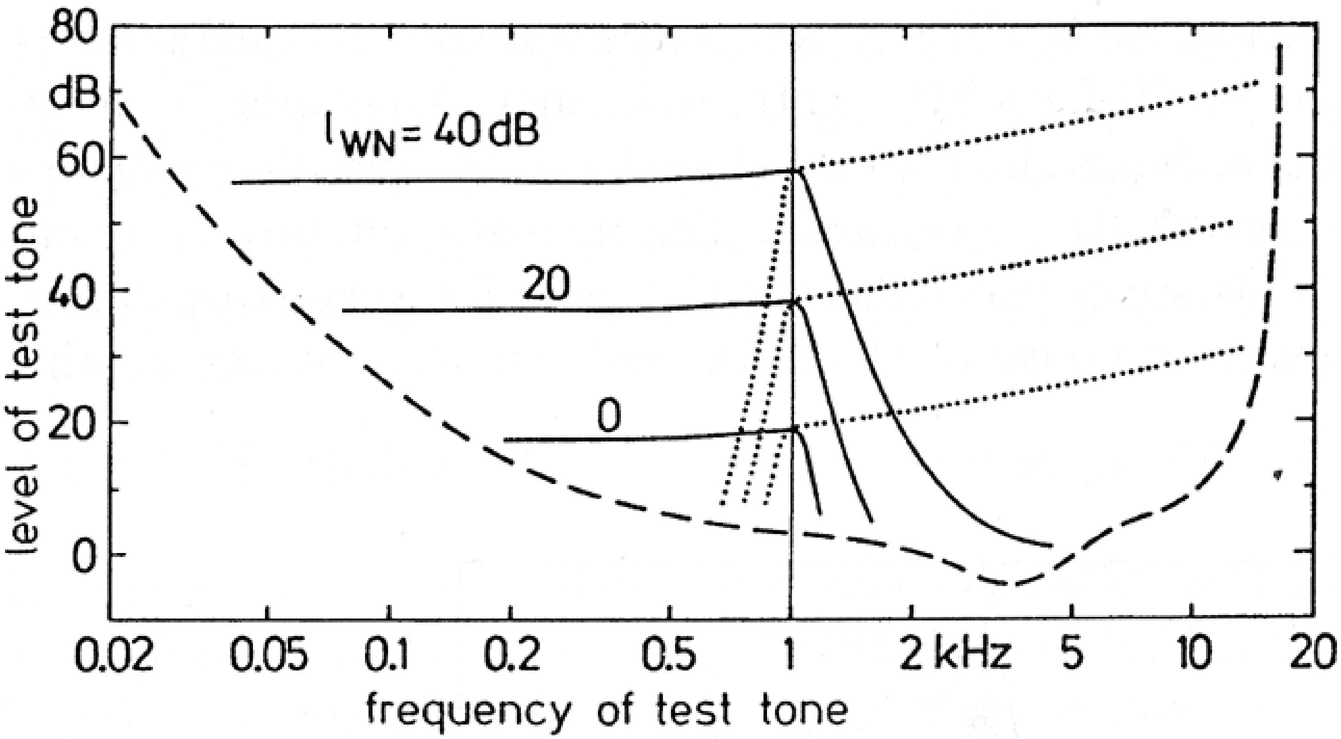
Example 5: Masking with white noise

- 500 Hz tone at varying amplitude with White Noise
- Level: -40,-35,-30,-25,-20,-15,-10 dB
- Noise Level: -50 dB

Example 6: Masking with white noise

- 5000 Hz tone at varying amplitude with White Noise
- Levels: same as Example 5

Masking of Pure Tones by Low-Pass or High-Pass Noise



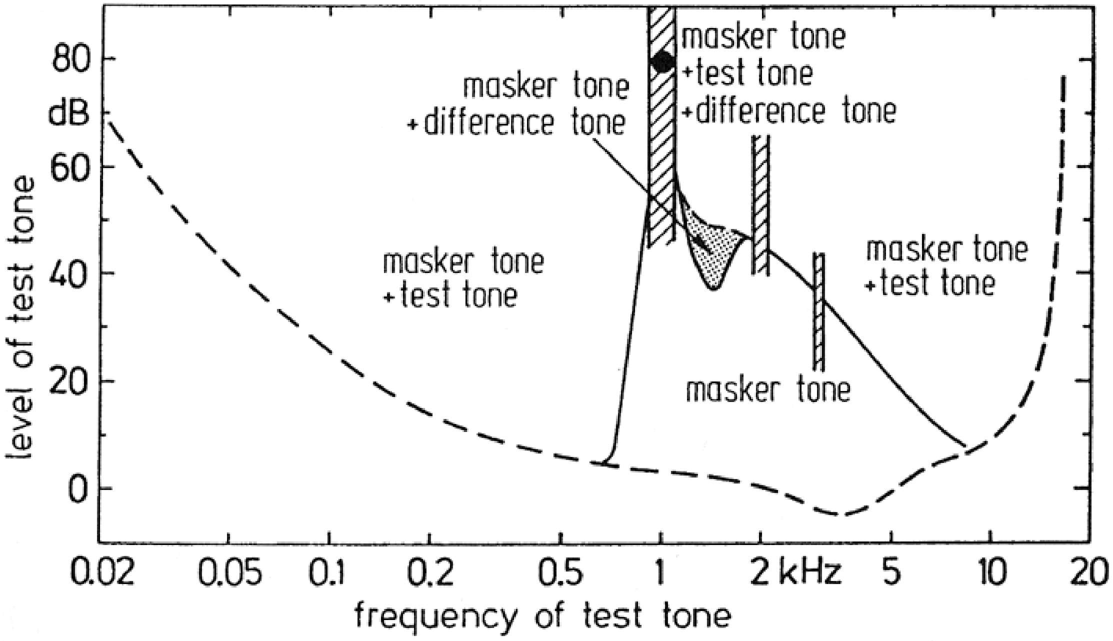
Source: Zwicker & Fastl “Psychoacoustics Facts and Models”

Masking of Pure Tones by Pure Tone

- pure tone:
  - single frequency

- figure:
  - 1 kHz masking tone with level of 80 dB
  - threshold for 'detection of anything'

  difficulties:
  - beats (hatching)
  - masker and difference tone (stippling)

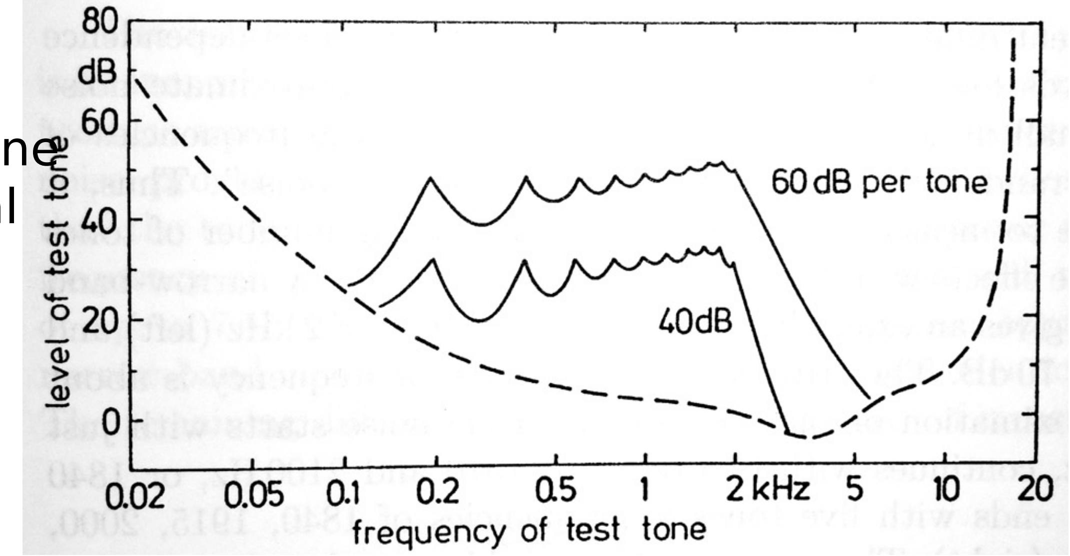


Source: Zwicker & Fastl "Psychoacoustics Facts and Models"

Masking of Pure Tone by Complex Tones

- complex tone:
    - fundamental tone with its harmonics

- figure:
    - threshold of pure tones masked by a complex tone with 200 Hz fundamental frequency and nine harmonics



Source: Zwicker & Fastl "Psychoacoustics Facts and Models"

Tonality (1)

- Tonality index α:
  - noisy signal: $\alpha = 0$
  - tonal signal: $\alpha = 1$

- System theory
  - Sharp spectral lines = Signal is periodic = Signal is predictable
  - Approximation: If the signal is predictable then it should be periodic
  - Therefore we can use prediction to approximate if a signal is tonal (by periodicity)

Prof. Dr.-Ing. K. Brandenburg, bdg@idmt.fraunhofer.de Prof. Dr.-Ing. G. Schuller, shl@idmt.fraunhofer.de

Tonality (2)

Example:

- calculate the predictability:

$$c(t,f) = \frac{(\hat{r}(t,f) - r(t,f))^2}{r(t,f)^2} + \frac{\left(\widehat{\Phi}(t,f) - \Phi(t,f)\right)^2}{\Phi(t,f)^2} \qquad \begin{array}{l} c \ \rightarrow 1 : noisy\ signal \\ c \ \rightarrow 0 : tonal\ signal \end{array}$$

  - If $c(t,f) > 1$ set it to 1 → $\alpha(t,f) = |c(t,f) - 1|$
  - amplitude of a spectral line in time and frequency $r(t,f)$
  - phase of a spectral line in time and frequency $\Phi(t,f)$
  - predicted values $\hat{r}(l,k)$ for amplitude and $\widehat{\Phi}(l,k)$ for phase

$$\hat{r}(t,f) = r(t-1,f) + \left(r(t-1,f) - r(t-2,f)\right)$$

$$\widehat{\Phi}(t,f) = \Phi(t-1,f) + \left(\Phi(t-1,f) - \Phi(t-2,f)\right)$$

Masking – Spreading Function

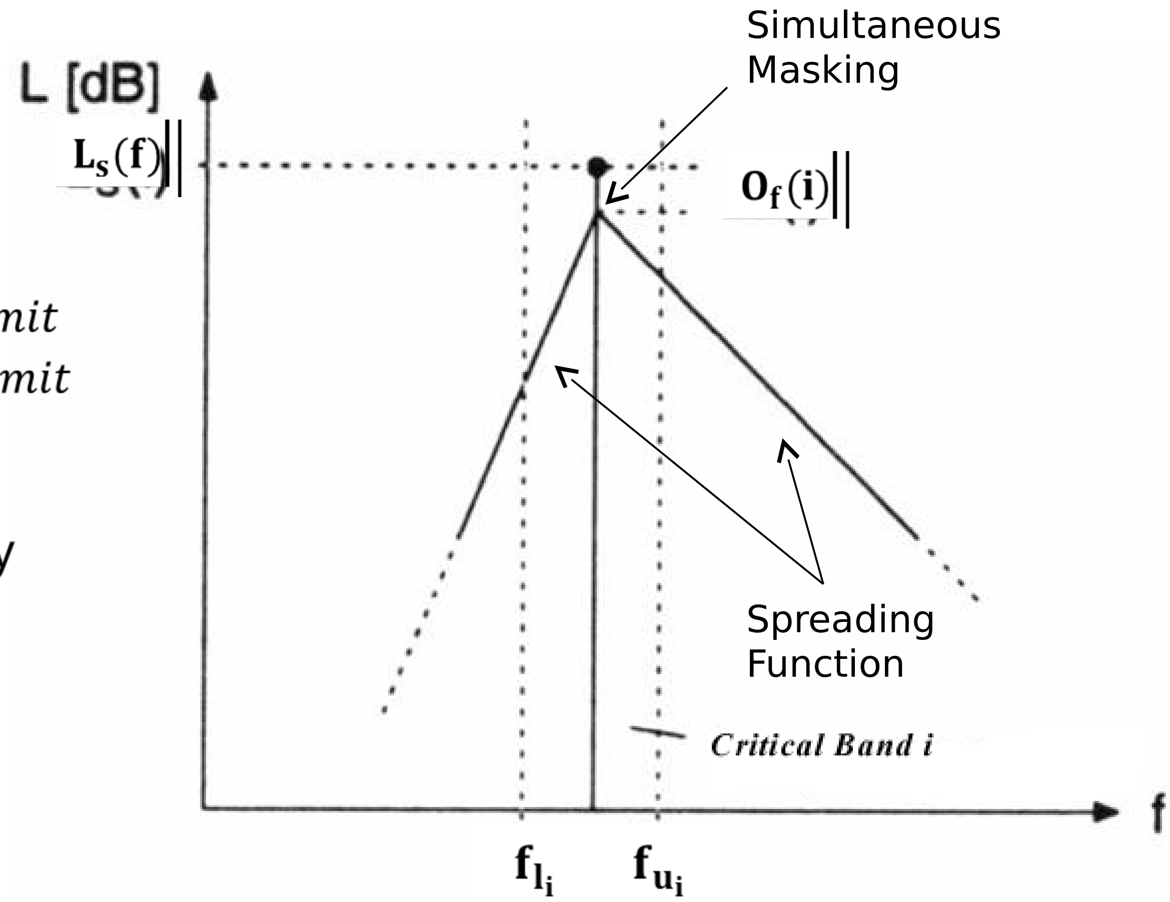
f_{l_i} ... *lower frequency limit*
f_{u_i} ... *upper frequency limit*
of Critical Band i

Power Spectral Density of signal x(t)

$$S_p(f) = X_R^2(f) + X_I^2(f)$$

Sound Pressure Level

$$L_S(f) = 10 \ \log_{10} S_p(f)$$



Source: U. Zölzer, "Digital Audio Signal Processing"

Prof. Dr.-Ing. K. Brandenburg, bdg@idmt.fraunhofer.de Prof. Dr.-Ing. G. Schuller, shl@idmt.fraunhofer.de

Calculating the Masking Threshold

Comparison of the signal level to Masking Threshold_

$$\frac{O_f(i)}{dB} = \alpha(14.5 + i) + (1 - \alpha)\alpha_v$$

α ...*Tonality Index*
α_v ...*Noise Coefficient*

$$\alpha_v = -2 - 2.05\arctan\left(\frac{f}{4\ kHz}\right) - 0.75 arctan\left(\frac{f^2}{2{,}56\ kHz^2}\right)$$

Approximation

$$\frac{O_f(i)}{dB} = \alpha(14.5 + i) + (1 - \alpha)5.5$$

Different Masking with different maskers:

- Tone masking: (14.5 + i) dB, where i is the integer number for the critical band
- Noise as a masker: 5.5 dB

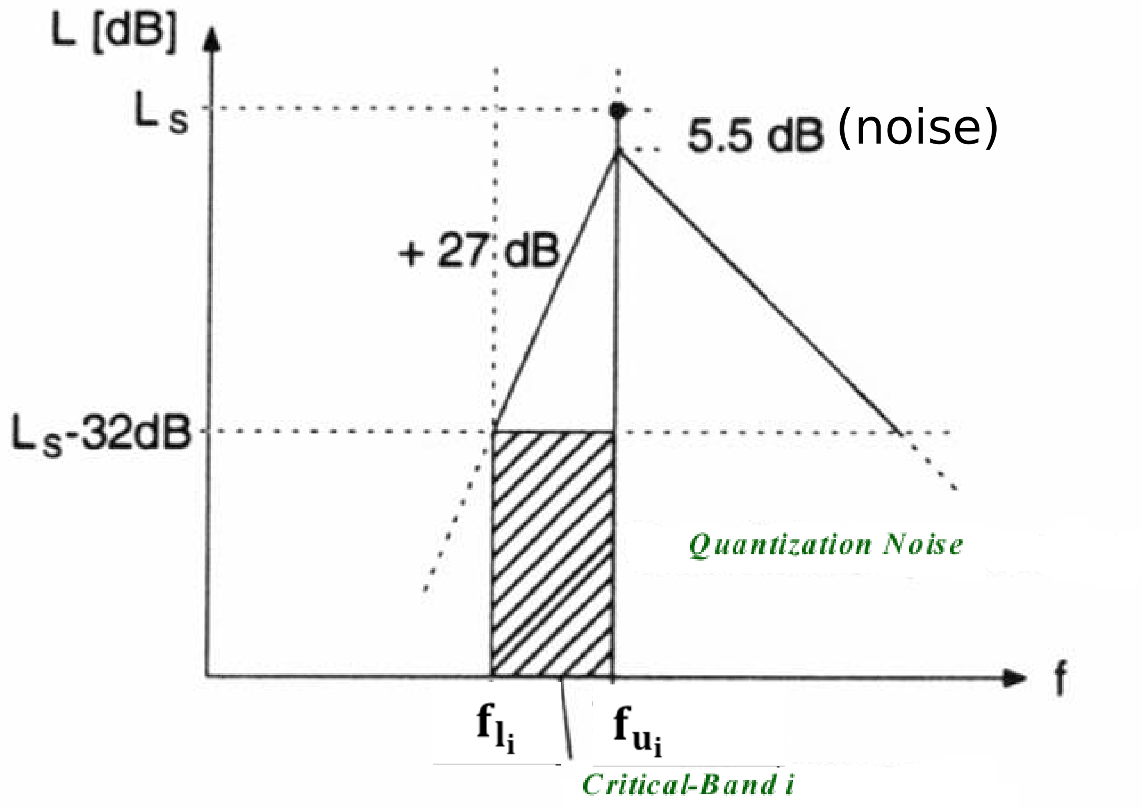
$L_s(f)$ ... *Sound Pressure Level*
$O_f(i)$ ... *Distance to Masking Threshold*

Simultaneous Masking Threshold (Power)

$$T(f) = 10^{[L_s(f) - O_f(i)]/10}$$

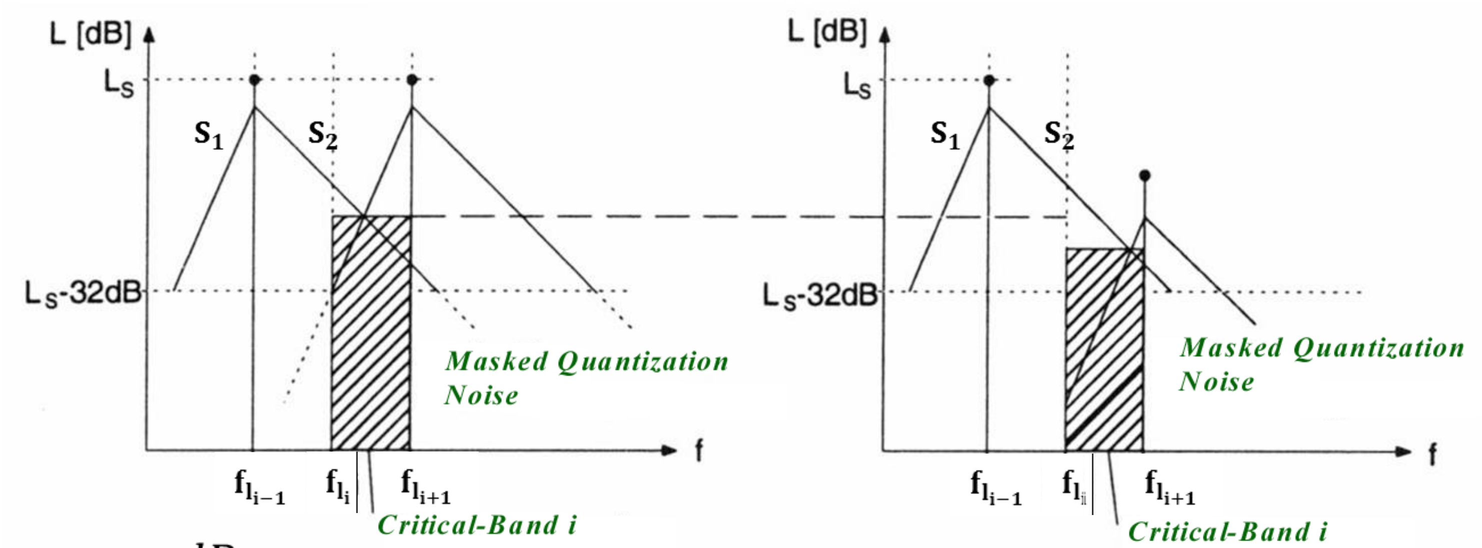
TECHNISCHE UNIVERSITÄT ILMENAU
Fraunhofer IDMT

In-Band Masking



Masking Neighboring Bands

- spread of masking due to the non-linearity of auditory filters
- resulting masking threshold = sum of power of neighbouring spreading functions
- here: value at intersection of neighbouring spreading functions taken



$$S_1 = 27\,\frac{dB}{Bark}$$

$$S_2 = 24 + 0.23\left(\frac{f}{kHz}\right)^{-1} - 0.2\,\frac{L_S(f)}{dB}\,\frac{dB}{Bark}$$

Prof. Dr.-Ing. K. Brandenburg, bdg@idmt.fraunhofer.de Prof. Dr.-Ing. G. Schuller, shl@idmt.fraunhofer.de

Sound Examples

Example 7: Dynamic range

- Bach organ music with 16 bits per sample

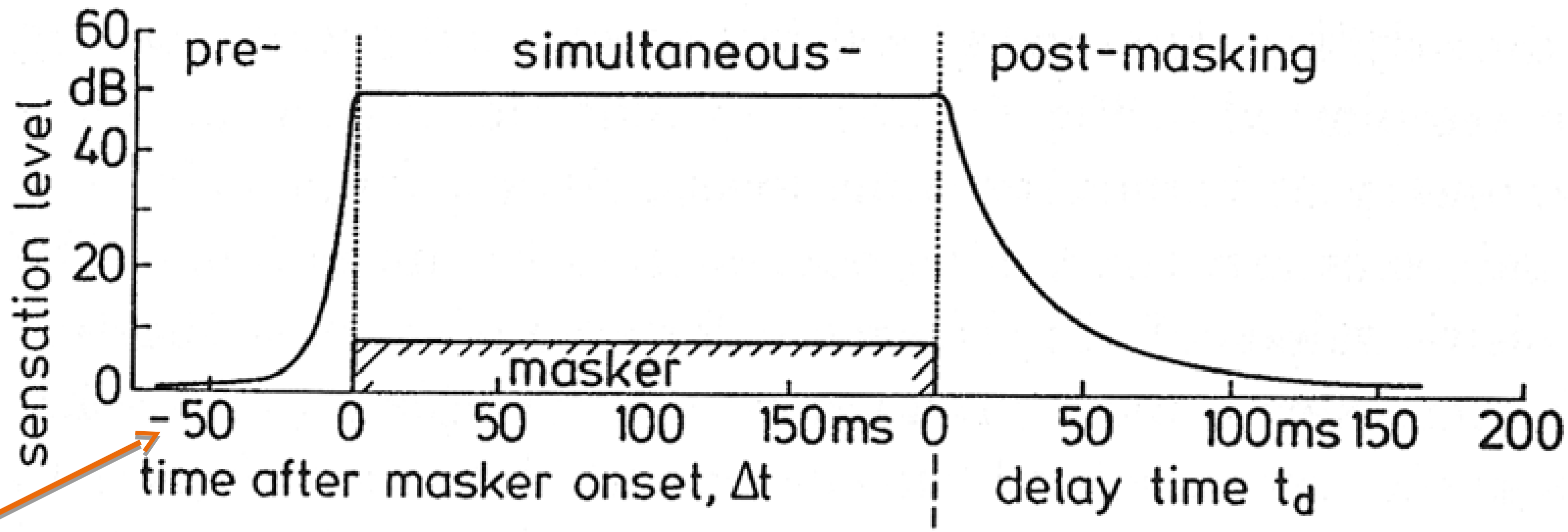
Example 8: Dynamic range

- Bach organ music with 11 bits per sample

Example 9: Dynamic range

- Bach organ music with 6 bits per sample

Temporal Masking Effects (1)



This is not correct !

Schematic drawing to illustrate and characterize the regions within which premasking, simultaneous masking and postmasking occur. Note that postmasking uses a different time origin than premasking and simultaneous masking

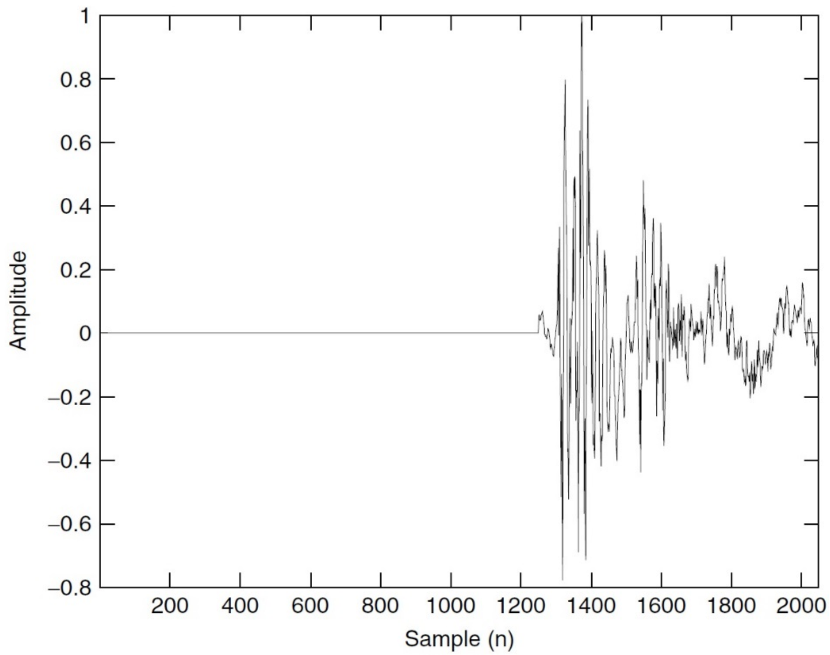
Source: Zwicker & Fastl “Psychoacoustics Facts and Models”

Temporal Masking Effects (2)

- Post-Masking: corresponds to decay in the effect of the masker ✉ expected
- Pre-Masking: appears during time before masker is switched on
  - Quick build-up time for loud maskers
  - Slower build-up time for faint test sounds
- Frequency resolution ↔ Blurring in time
- Frequency resolution in the ear ✉ Masking in time
- Because of in-ear fast processing between quiet to loud signals, we get Pre-Echoes
  - Pre-Masking: 1-5 ms
  - Post-Masking: ~100ms

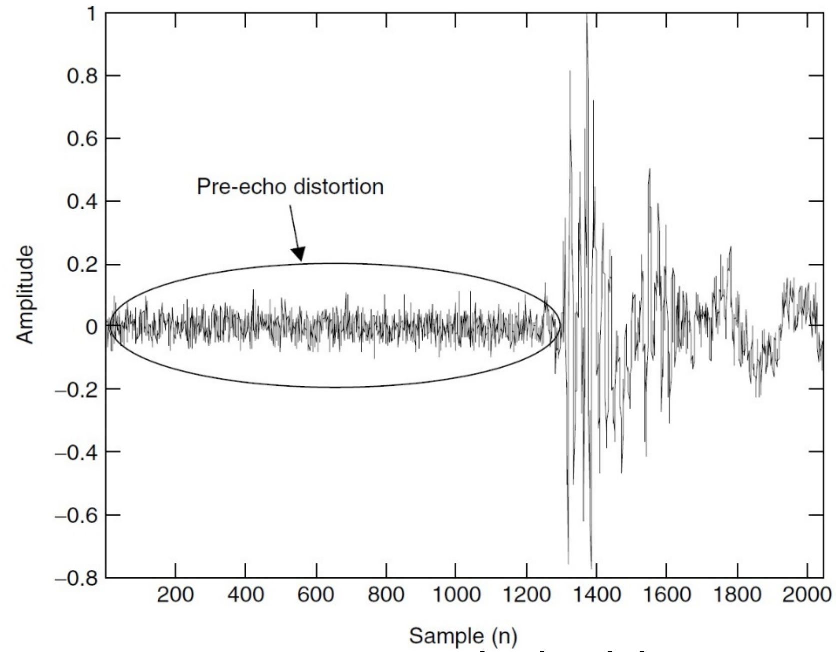
Pre-Echo: Example

Without Pre-Echo:



- Castanets original

With Pre-Echo:



- Castanets coded with a block size of 2048 samples

Source: Spanias et.al. "Audio Signal Processing and Coding"

Prof. Dr.-Ing. K. Brandenburg, bdg@idmt.fraunhofer.de Prof. Dr.-Ing. G. Schuller, shl@idmt.fraunhofer.de

TECHNISCHE UNIVERSITÄT ILMENAU
Fraunhofer IDMT

next lecture:

09.11. - Quantization and Coding