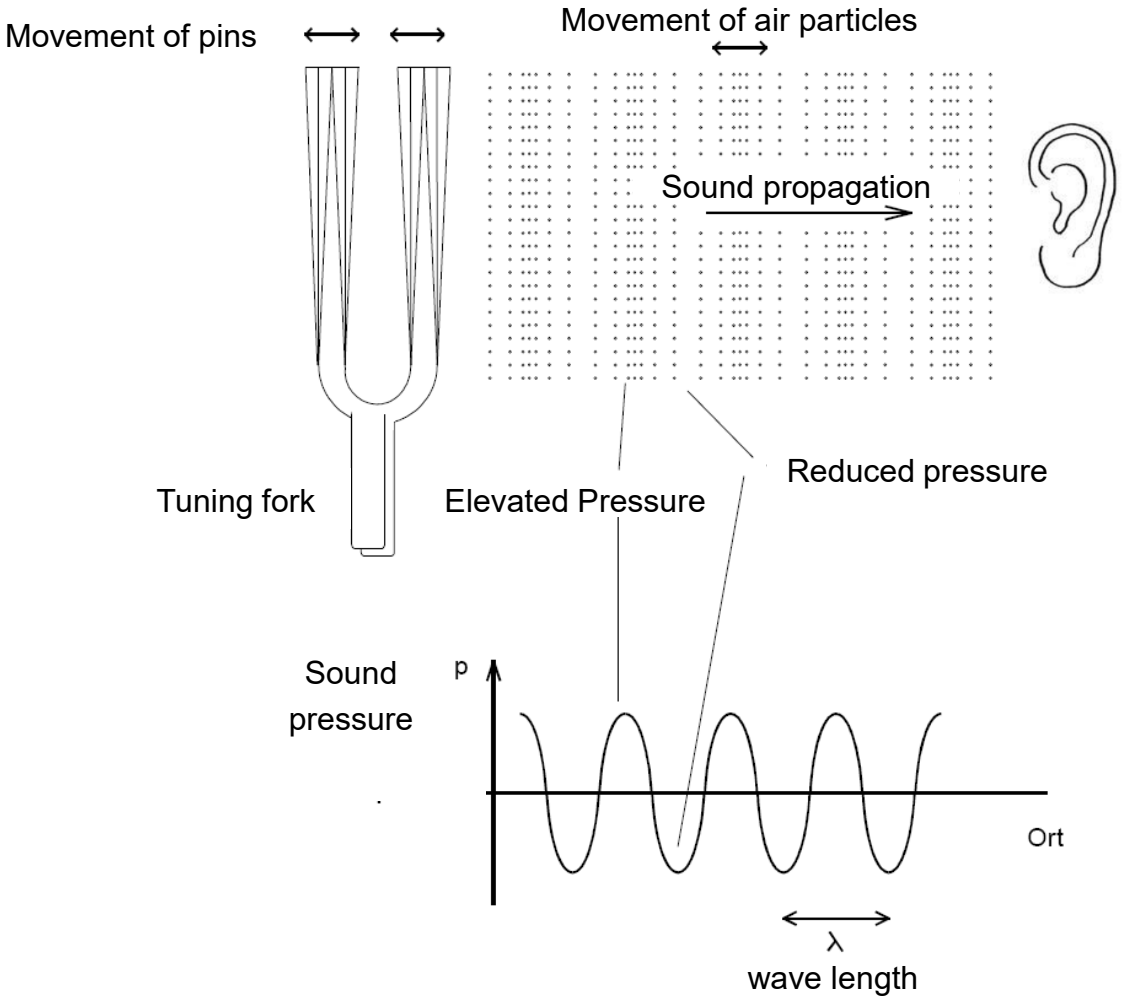
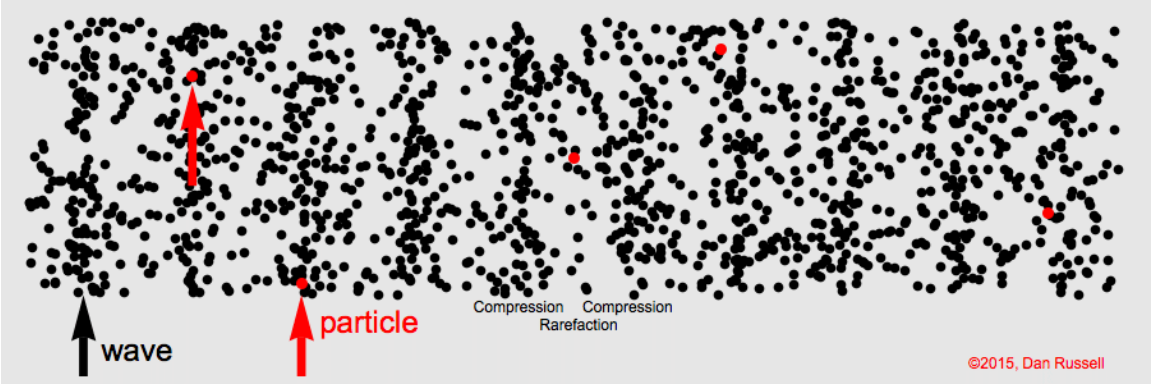
The Contents

1

- **Noise_Basic acoustics**
- Sampling_Nyquist-Shannon
- FFT_Analysis
- Psychoacoustics

Noise_ the sound field



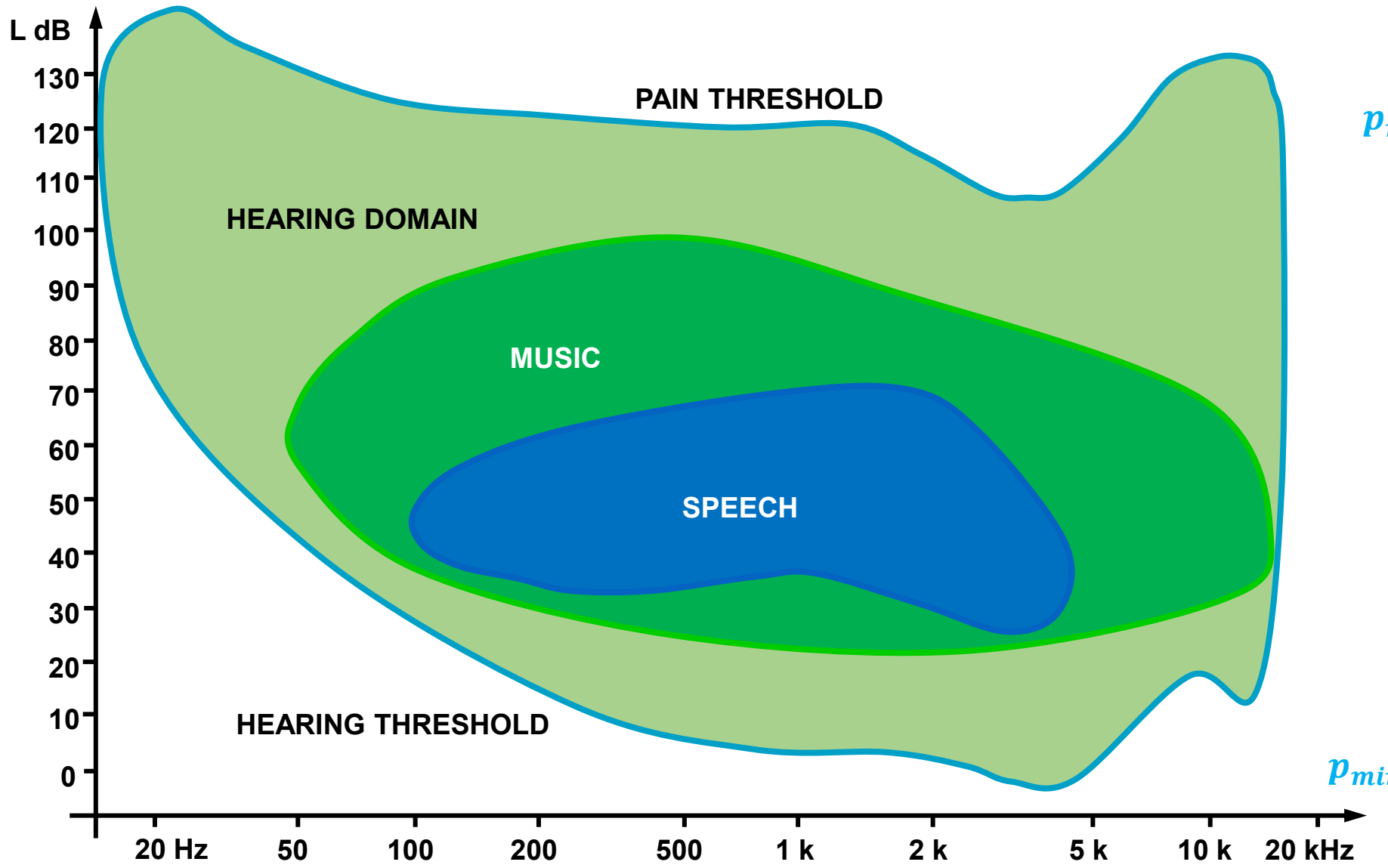


$$\lambda = c / f$$

Examples for wavelengths in air (c ≈ 343 m/s)

- f = 20 Hz, λ ≈ 17 m
- f = 20 kHz, λ ≈ 1,7 cm

Noise_ human Auditory Range



$p_{max} = 10\ pa \longrightarrow 120\ dB$

$p_{min} = 2 \cdot 10^{-5} pa \longrightarrow 0\ dB$

$f = 20 - 20k\ Hz$

Noise_ human Auditory Range

Sound pressure level of some every day scenarios

130 dB	Threshold of pain, jet plane (50 m)
120 dB	Rock concert (30 m)
110 dB	
100 dB	Disco/Club (at the ear), jackhammer (1 m)
90 dB	Main street (5 m)
80 dB	Vacuum cleaner (2 m)
70 dB	Car (10 m)
60 dB	
50 dB	Normal conversation (distance 1 m)
40 dB	
30 dB	Very quiet room (near the ear)
20 dB	
10 dB	
0 dB	Threshold of hearing

What's decibel ?
Definition…

Noise_ linear vs. logarithmic scale

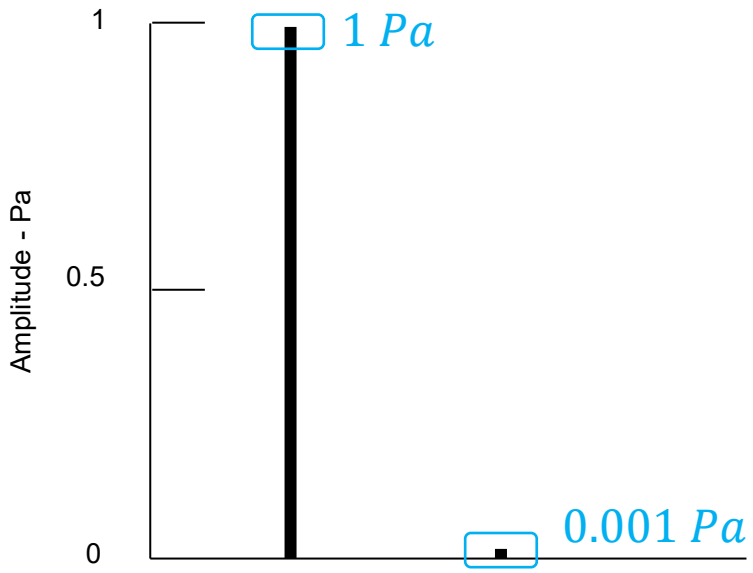
Perception of noise is limited by:

- Auditory threshold $p_{min} = 2 \cdot 10^{-5} Pa$
- Threshold of pain $p_{max} = 20\ Pa$

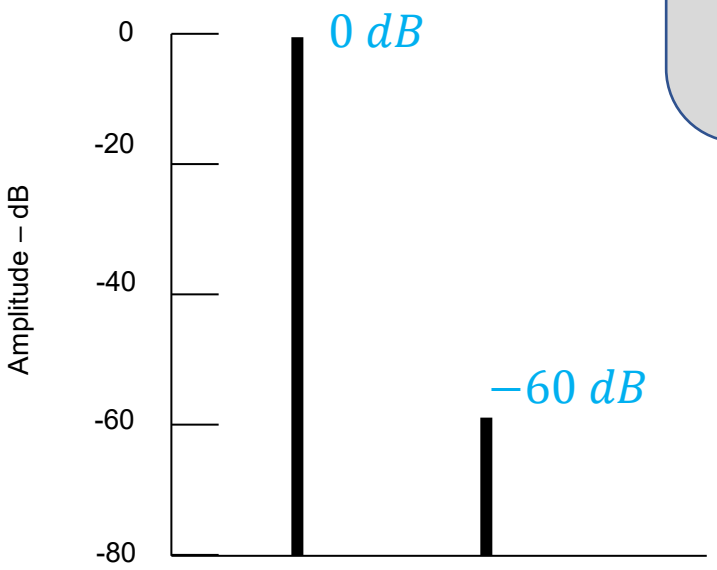
声压级:

$$L_p = 20 \cdot \log_{10}\left(\frac{p}{p_{ref}}\right)$$

$$p_{ref} = 2 \cdot 10^{-5}\mathrm{Pa} \rightarrow L_{min} = 0\mathrm{dB}$$



Linear Amplitude Scale



Logarithmic Amplitude Scale

Decibel (Log) scale makes comparing vastly different magnitudes easier!

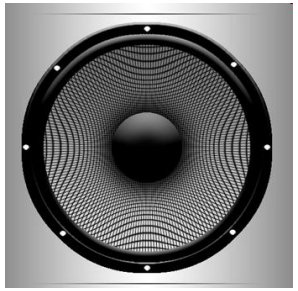
Noise_ Sound pressure level

$$\bar{p} = \sqrt{\frac{1}{T}\int_0^T [p(t)]^2\, dt}$$

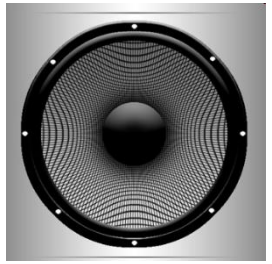
$$SPL = 10\log\left(\frac{\overline{p^2}}{{p_{ref}}^2}\right) = 20\log\left(\frac{\bar{p}}{p_{ref}}\right)$$

Sound pressure level	Sound pressure level	Situation
$0dB$	$20uPa$	HEARING THRESHOLD
$94dB$	$1Pa$	CALIBRATION
$100dB$	$2Pa$	DOUBLE
$114dB$	$10Pa$	CALIBRATION
$120dB$	$20Pa$	PAIN THRESHOLD

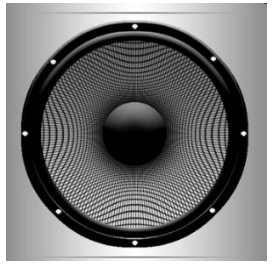
Noise_ Calculating using logarithmic scaling



$p_1 = 2 \cdot 10^{-3}\, Pa$



+



$p_1 = 2 \cdot 10^{-3}\, Pa$ $\quad$ $p_2 = 2 \cdot 10^{-3}\, Pa$

$$L_p = 20 \log\left(\frac{2 \cdot 10^{-3}\, Pa}{2 \cdot 10^{-5}\, Pa}\right)[dB]$$

$$L_p = 20 \log(100)$$

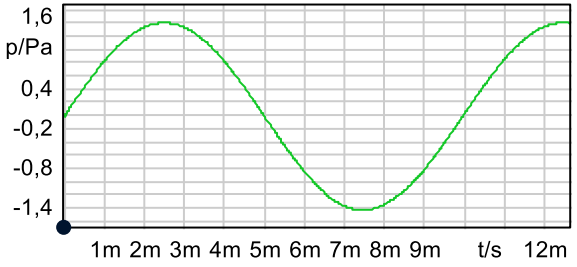
$$L_p = 40 dB$$

$$L_{P1+P2} = 40dB + 40dB\ ?$$

Noise_ Calculating using logarithmic scaling

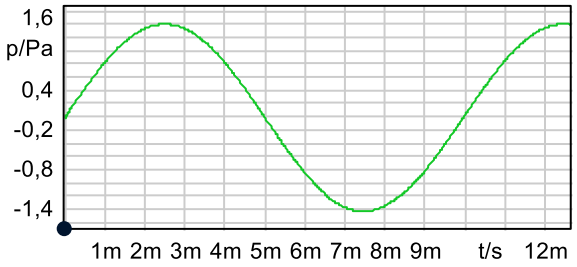
$$L_p = 20\log\left(\frac{P}{P_{\text{ref}}}\right)[dB]$$

100 dB



+

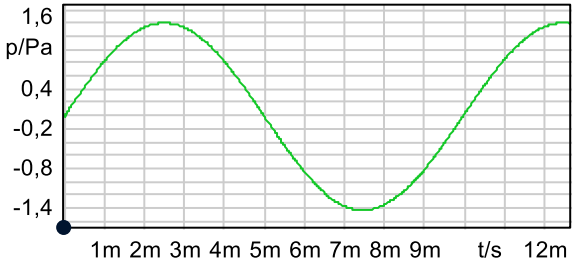
100 dB



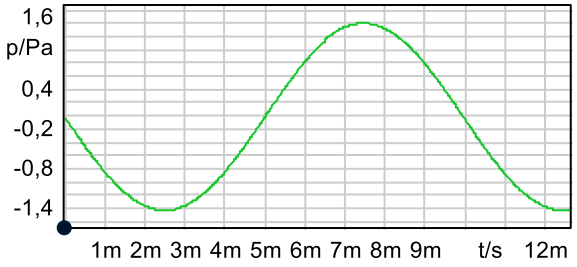
= 106 dB

$$L_p = 20\log\left(\frac{2*P}{P_{\text{ref}}}\right)[dB]$$

6dB

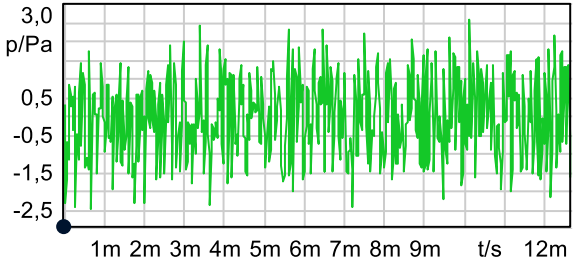


+

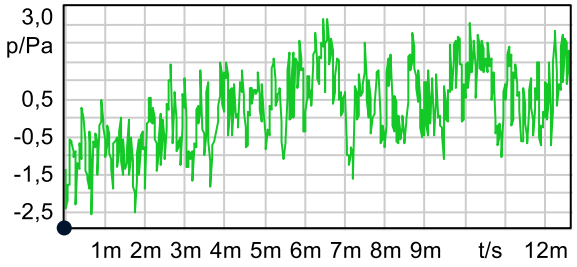


= -∞ dB

$$L_p = 20\log\left(\frac{P - P}{P_{\text{ref}}}\right)[dB]$$



+



= 103 dB

$$L_p = 10\log\left(\frac{2*P^2}{P_{\text{ref}}^{\ 2}}\right)[dB]$$

3dB

Noise_ Calculating using logarithmic scaling

- Correlation/ Calculation: Linear_ Amplitude DOUBLE

$$P_1 = P_2 = 2\,Pa \rightarrow SPL = 100dB \; SPL = 20\log\left(\frac{P_1+P_2}{P_{ref}}\right) = 106dB$$

- NON-Correlation/ SAME SPL, Calculation: Energy Superposition

$$P_1 = P_2 = 2\,Pa \rightarrow SPL = 100dB$$

$$SPL = 10\log\left(\frac{\overline{P_1{}^2}+\overline{P_2{}^2}}{P_{ref}{}^2}\right) = 10\log 2 + 10\log\left(\frac{\overline{P_1{}^2}}{P_{ref}{}^2}\right) = 3 + SPL_1 = 103\,dB$$

- Random-Signal

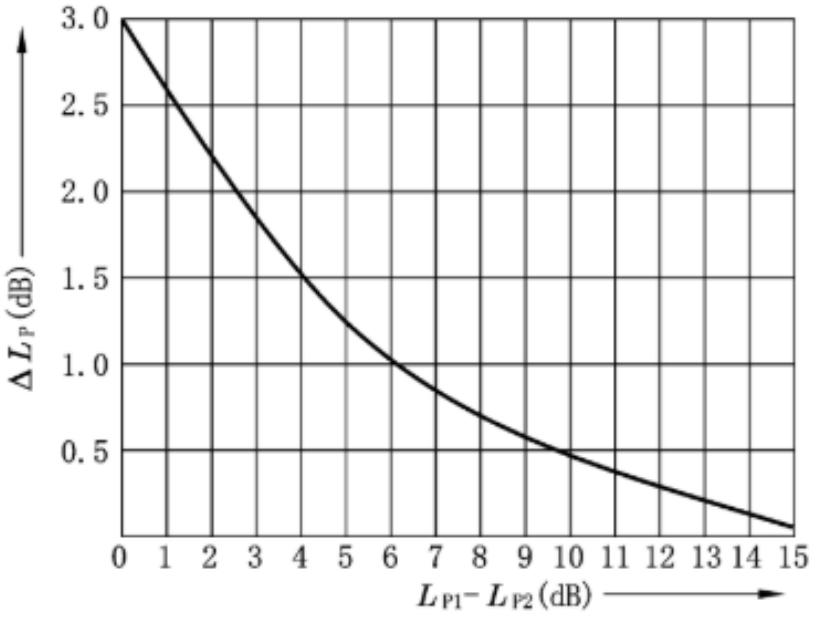
$$SPL = SPL_1 + SPL_2 + SPL_3 + \cdots$$
$$= 10\log\left(10^{\frac{SPL_1}{10}} + 10^{\frac{SPL_2}{10}} \ldots\right)$$

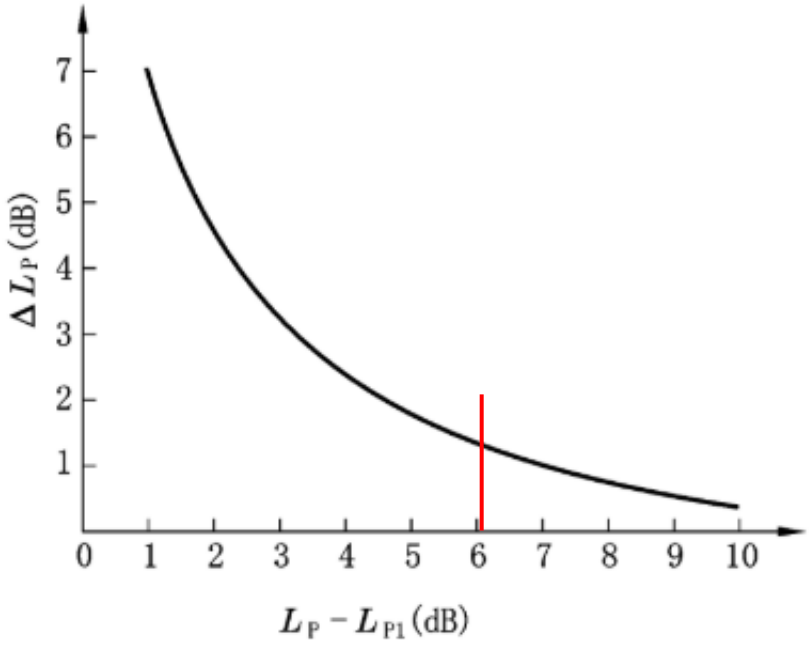
NON-Correlation ?

Phase ? SPL ?

Noise_ Calculating using logarithmic scaling

- SUMMARY





LP-LP1

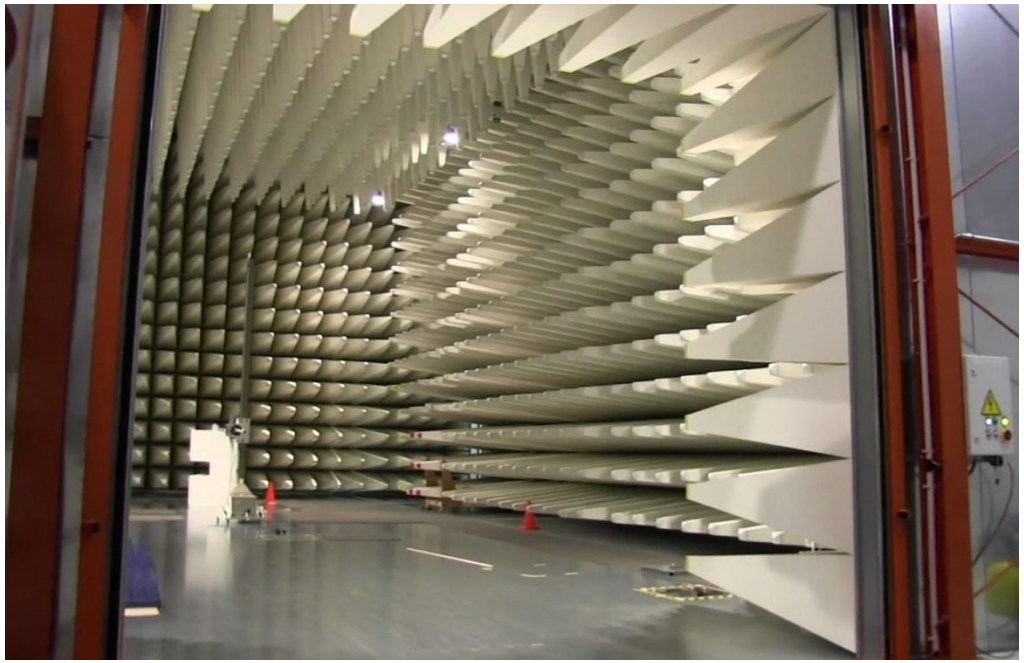
< 6dB NO ! (1dB)

6~15dB FACTOR !

>15dB YES !

Noise_ Calculating using logarithmic scaling

- Sound field
  - Free field:
    Sound propagation without reflections



  - Diffuse sound field:
    Sound incident equal from every direction

Noise_ Calculating using logarithmic scaling

- Sound field
  - Real sound field (e.g. vehicle interior)

    Mixture of diffuse and free sound field, reflection from windows but absorption from seats etc.

Noise_ Octave

$1000\ Hz$

$400\ Hz$

$1000\ Hz$

$800\ Hz$

$1000\ Hz$

$1050\ Hz$

A_4: 440 Hz (standard pitch)

A_5: 880 Hz

A_6: 1760 Hz

A_7: 3520 Hz

440 Hz Span

880 Hz Span

1760 Hz Span

Equal ratio series

$$f_{i+1} / f_i = 2^{1/n}$$

- The word octave is borrowed from music theory where it refers to a span of eight notes:
- The frequency ratio of the highest note to the lowest one for an octave is 2:1

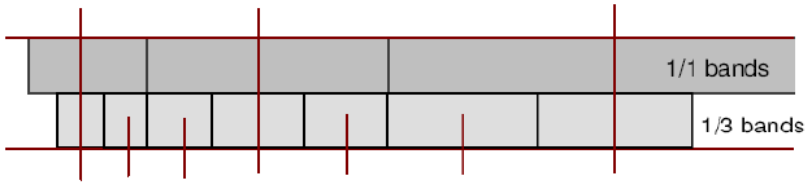
$$f_2 = f_1 \cdot 2$$

1/1 Octave

Noise_ Octave

$$f_{i+1} \,/\, f_i = 2\text{^}1/N \;\; (N = 1,2,3 \ldots)$$

- One can also divide an octave further into three equal bands:



- Band center and limiting frequencies(ICE):

$$f_{c,i} = 1000 \times \left(2^{1/N}\right)^{i} \;\; with\ i = 0, \pm 1, \pm 2 \ldots$$

$$f_{lower} = (2^{1/3})^{-1/2} f_c \quad (N = 3)$$

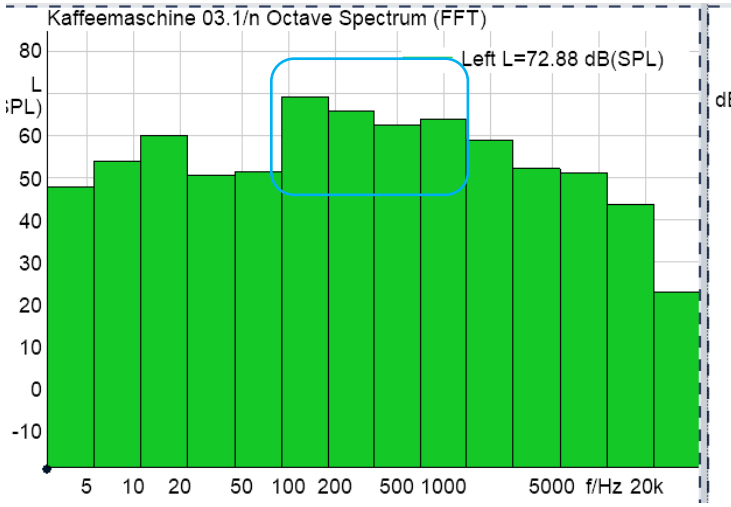
$$f_{upper} = (2^{1/3})^{1/2} f_c$$

ISO 150266(GB3240)/ Octave

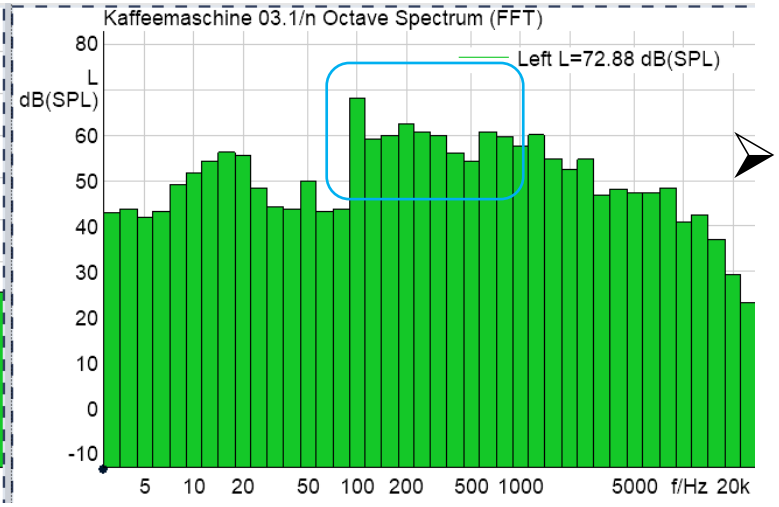
Normalized frequency	1/1 oct	1/2 oct	1/3 oct		1/1 oct	1/2 oct	1/3 oct		1/1 oct	1/2 oct	1/3 oct
16	x	x	x	160			x	1600			x
18				180		x		1800			
20			x	200			x	2000	x	x	x
22.4		x		224				2240			
25			x	250	x	x	x	2250			x
28				280				2800		x	
31.5	x	x	x	315			x	3150			x
35.5				355		x		3550			
40			x	400			x	4000	x	x	x
45		x		450				4500			
50			x	500	x	x	x	5000			x
56				560				5600		x	
63	x	x	x	630			x	6300			x
71				710		x		7100			
80			x	800			x	8000	x	x	x
90		x		900				9000			
100			x	1000	x	x	x	10000			x
112				1120				11200		x	
125	x	x	x	1250			x	12500			x
140				1400		x		14000			
160			x	1600			x	16000	x	x	x

Table 2.1 Normalized frequencies (Hz)

Noise_ 1/3Octave

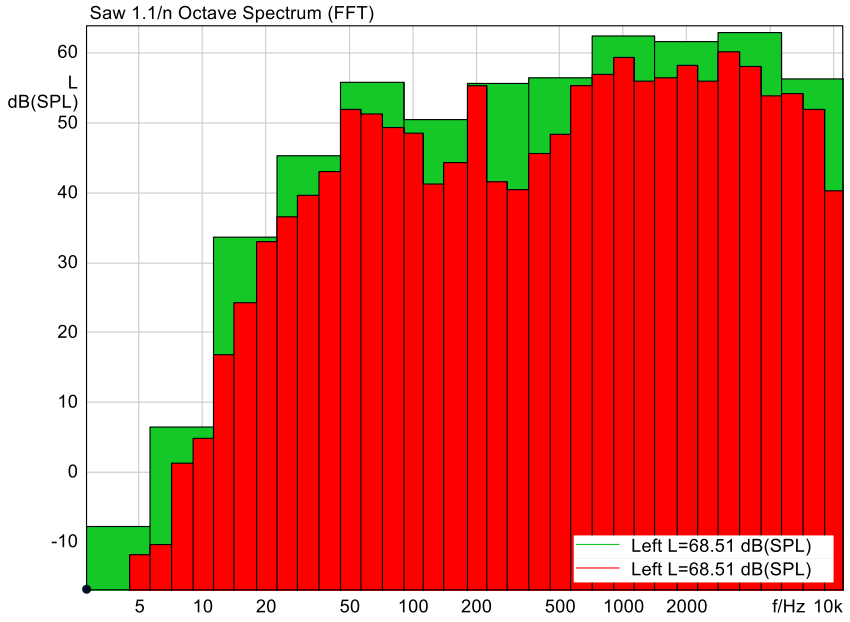


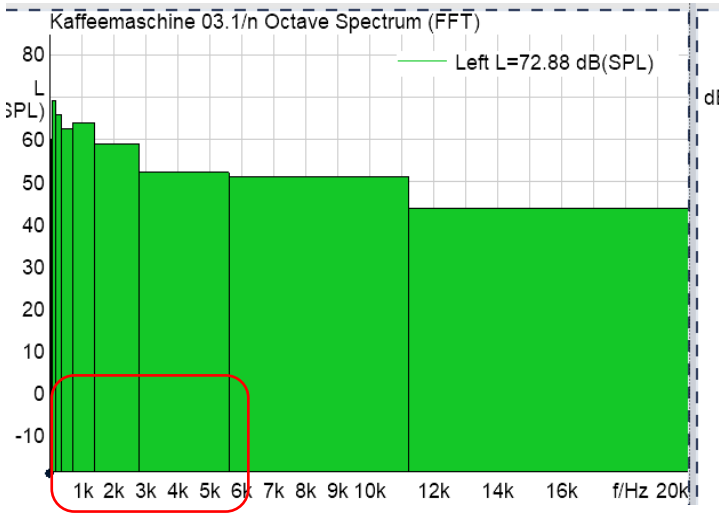
1/1 Octave



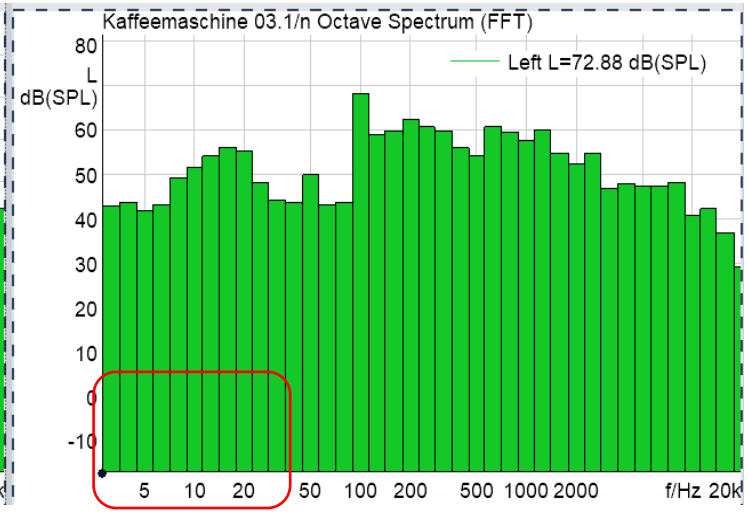
1/3 Octave

- Octave band eigenvalue outlier





Linear



Log

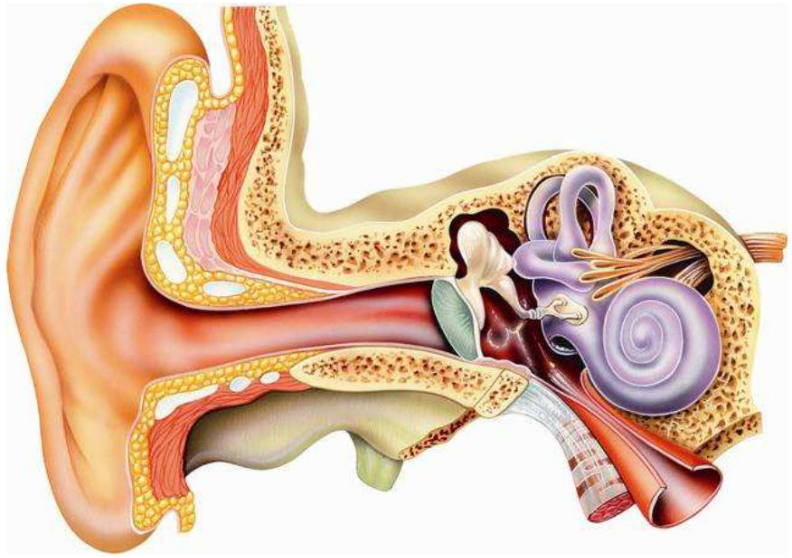
- Low frequency band features

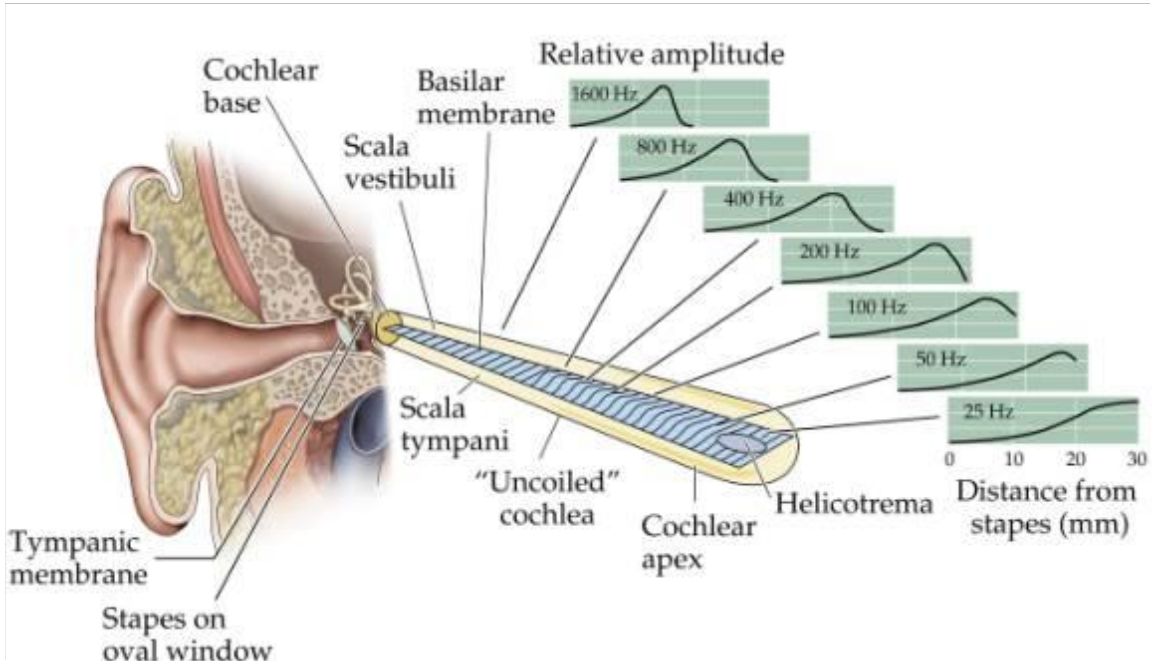
Noise_ 1/3Octave

- Concept
- Taken from music
- Approximation of the human hearing/ Critical Bands
- Generally: 1/n Octave, continuous spectral analysis calculates the frequency distribution using 1/n octave wide frequency bands
- 3rd Octave means one third of an octave (Filters defined in ANSI S1.6-1984 and DIN 45401)

Noise_ weighting

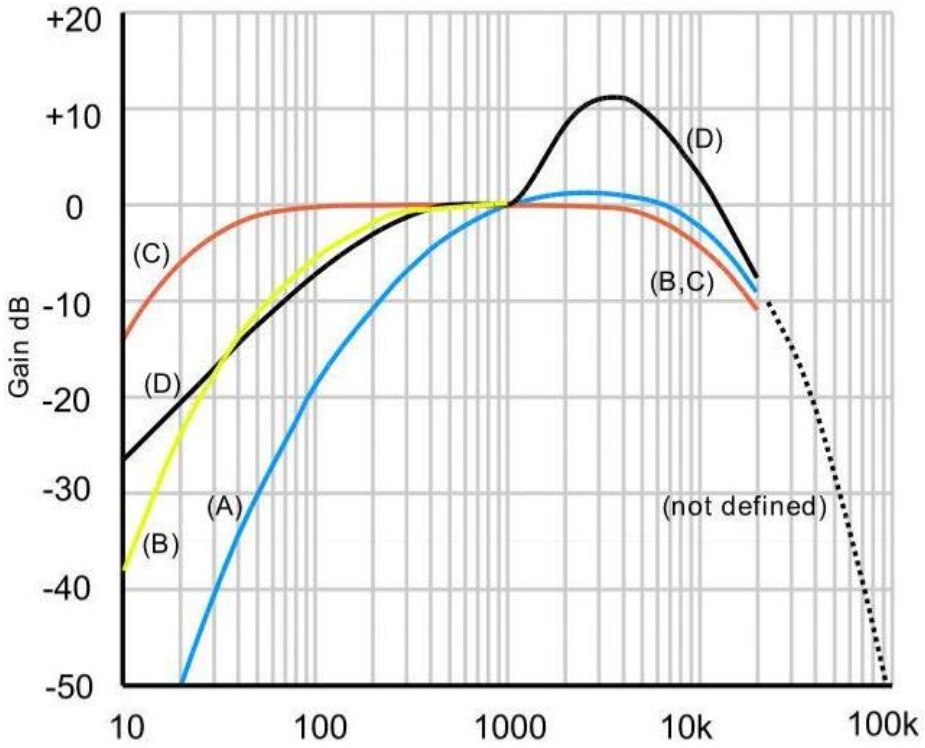
- nonlinear sound transmission
- sound perception occurs in frequency (“critical”) bands
- **sounds must “build-up”**





Time Weighting Fast Slow Frequency Weighting A, B, C, D… ?

Noise_ Frequency weighting



A-weighting (blue), B (yellow), C (red), and D-weighting (blk)

The processing of A-weighting

f_c	125	250	500	1000	2000	4000
SPL	90	90	90	90	90	90
A −weighting	-16.1	-8.6	-3.2	0	1.2	1

$$L_p = 10lg\left(10^{\frac{90}{10}} \times 6\right) = 10lg10^9 + 10lg6 = 97.8dB$$

$$L_A = 10lg\left(10^{\frac{73.9}{10}} + 10^{\frac{81.4}{10}} + 10^{\frac{80.8}{10}} + 10^{\frac{90}{10}} + 10^{\frac{91.2}{10}} + 10^{\frac{91}{10}}\right)$$

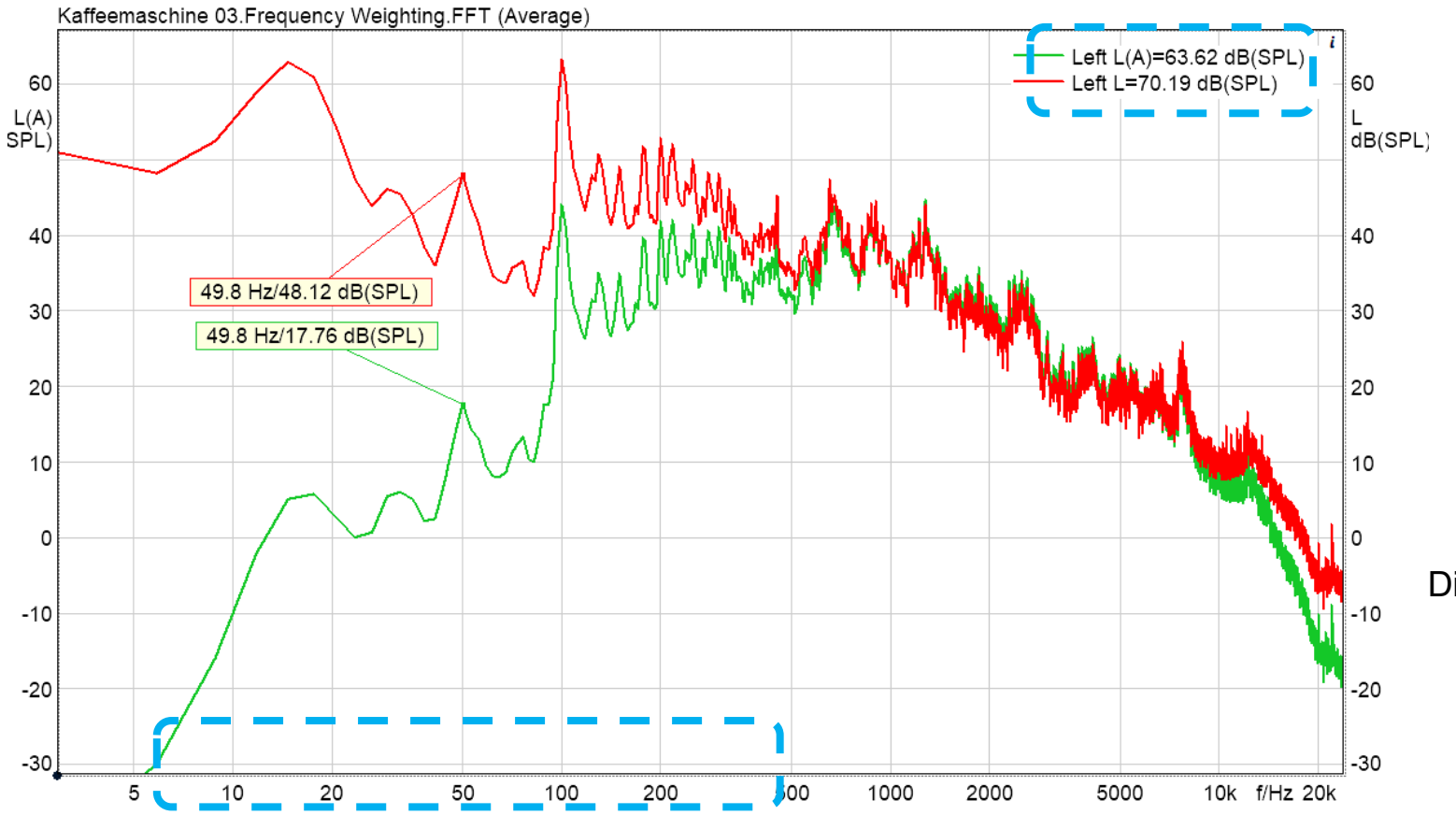
$$= 96.3dB(A)$$

$$A(f) = 20\log\left\{f_4^2 f^4\left[(f^2+f_1^2)(f^2+f_2^2)^{\frac{1}{2}}(f^2+f_3^2)^{\frac{1}{2}}(f^2+f_4^2)\right]^{-1}\right\} - A_{1000}$$

- $A_{1000} = -2.00dB, f_1 = 20.6Hz, f_2 = 107.7Hz, f_3 = 737.9Hz, f_4 = 12194Hz$

Noise_ Frequency weighting

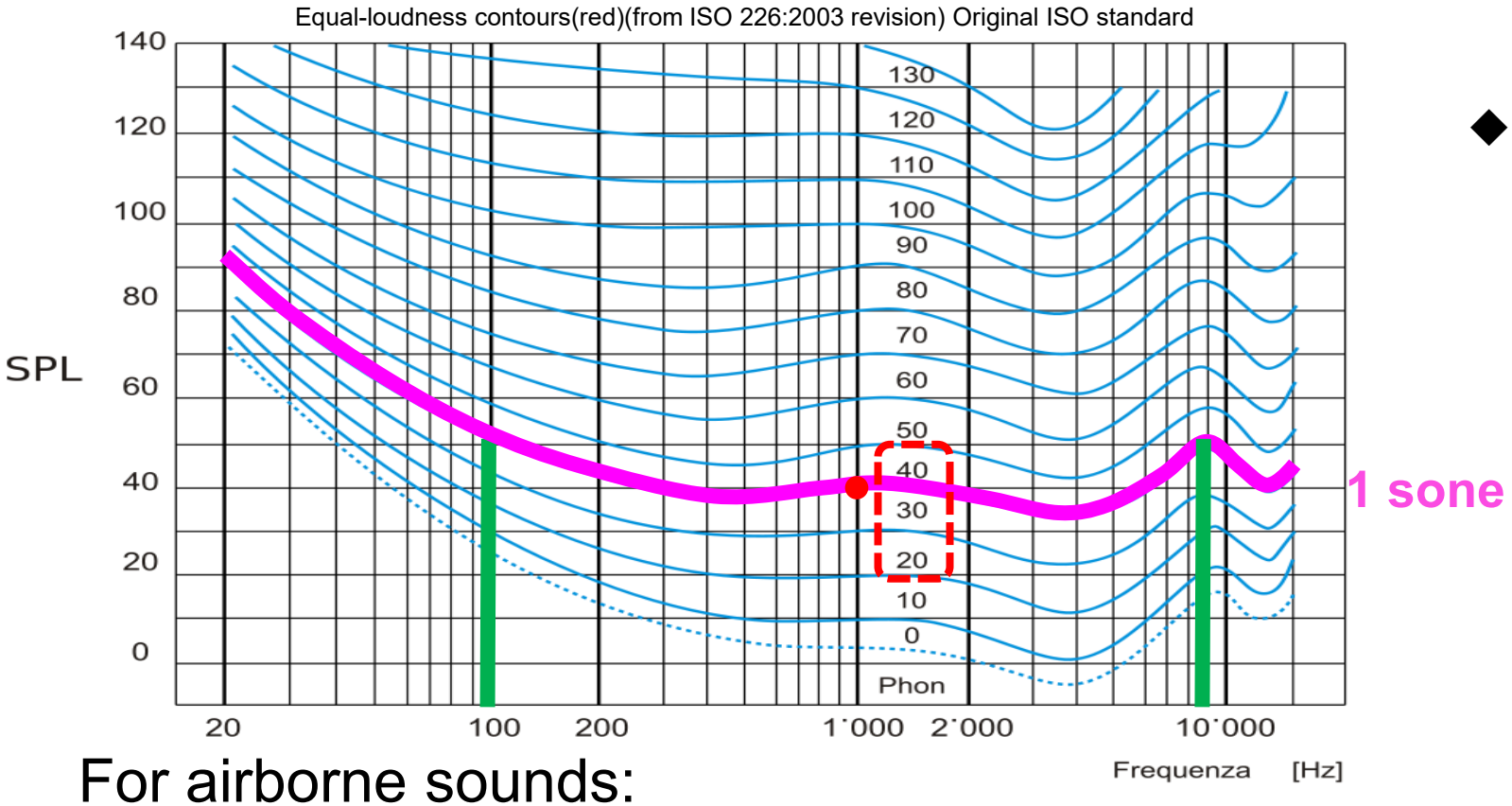
- DIN EN 61672: Representation of the sound pressure level in dB with frequency-dependent attenuation
- Difference between A-weighted and unweighted curves in an FFT-analysis



标称频率(Hz)	频率计权(dB)		
	A	C	Z
25	-44.7	-4.4	0.0
31.5	-39.4	-3.0	0.0
40	-34.6	-2.0	0.0
50	-30.2	-1.3	0.0
63	-26.2	-0.8	0.0
80	-22.5	-0.5	0.0
100	-19.1	-0.3	0.0
125	-16.1	-0.2	0.0
160	-13.4	-0.1	0.0

Difference frequencies with different weighting factors

Noise_ Frequency weighting

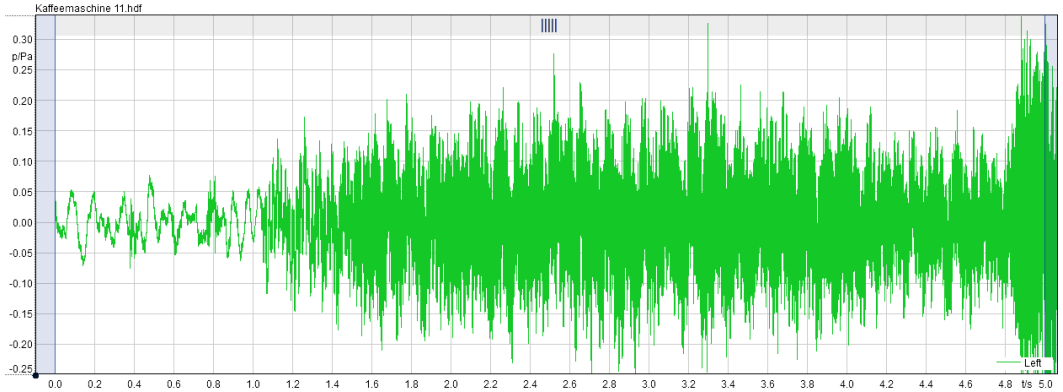


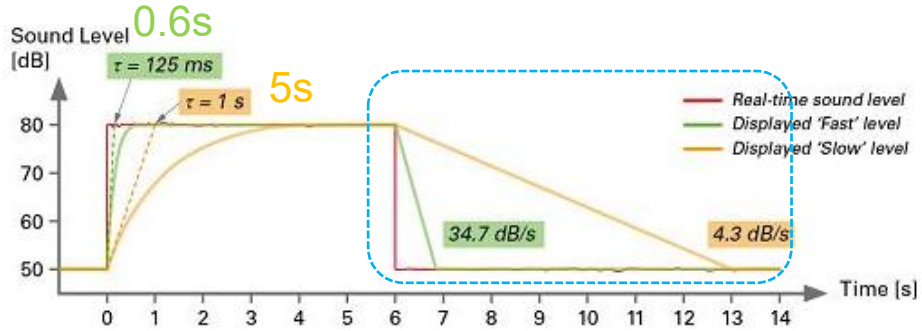
- The A-weighting curve corresponds to the constant loudness curve at approx. ***20-40 phon***, the B-weighting curve is for ***the 50-70 phon*** range and the C-weighting curve for ***80-90 phon***. The A-weighting curve is the most commonly used one.

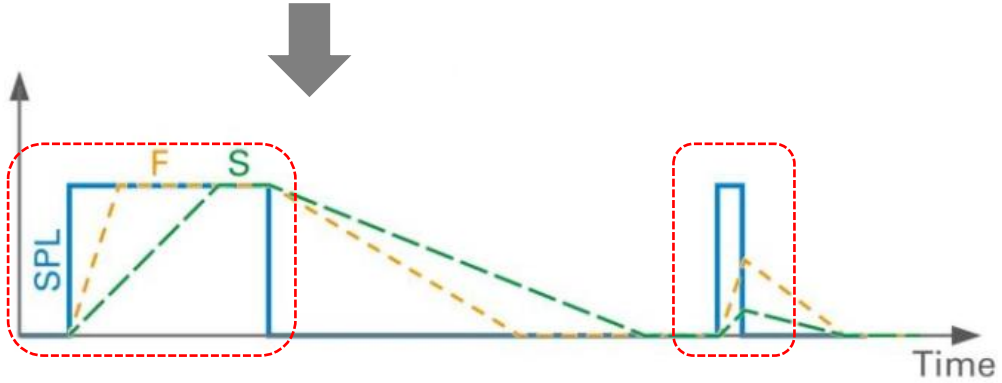
For airborne sounds:

- Equal loudness curves (top left) were created from listeners giving their opinions on how loud sounds were
  - ✓ A 50 dB tone at 100 Hz is as “loud” as a 50 dB tone at 10000 Hz
- Sound weighting curves try to equalize tones at different frequencies
  - ✓ 1000 Hz 40 dB tone is 40 dB(A), 100 Hz 50 dB tone is 20 dB(A)

Noise_ Time weighting







Hint: the calculation of level curves with Fast or Slow is performed according to **DIN EN 61672:2014-07**

$$\bar{p} = \sqrt{\frac{1}{T}\int_0^T [p(t)]^2\, dt}$$

- $LAL_{A\tau} = 20\log\left\{\frac{\left[\frac{1}{\tau}\int_{-\infty}^{t} P_A^2(\varepsilon)e^{-\frac{t-\varepsilon}{\tau}}d\varepsilon\right]^{\frac{1}{2}}}{P_0}\right\} = 10\log\left[\frac{P_T^2(t)}{P_0^2}\right]$
- Fast: the integration time is set to 125 ms.
- Slow: the integration time is set to 1000 ms.
- Manual: The integration time is the time entered in the field Time Constant [ms]. The longer the integration time, the stronger the smoothing of the curves

Time weighting &Frequency weighting as well ?

Noise_ Time weighting

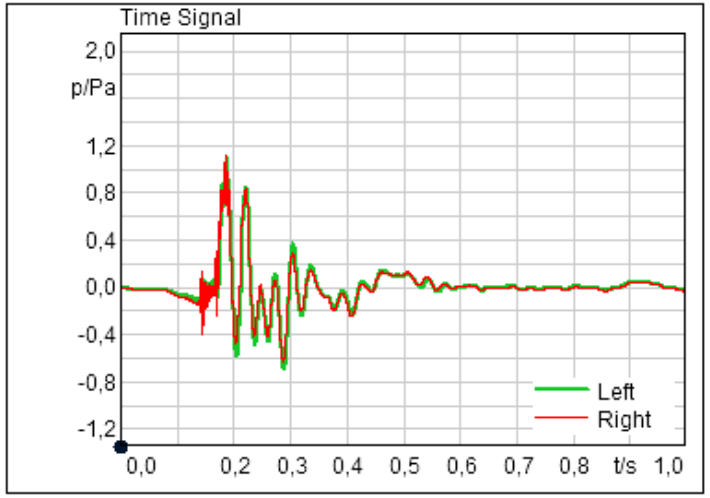


Figure 1 : Time domain signal

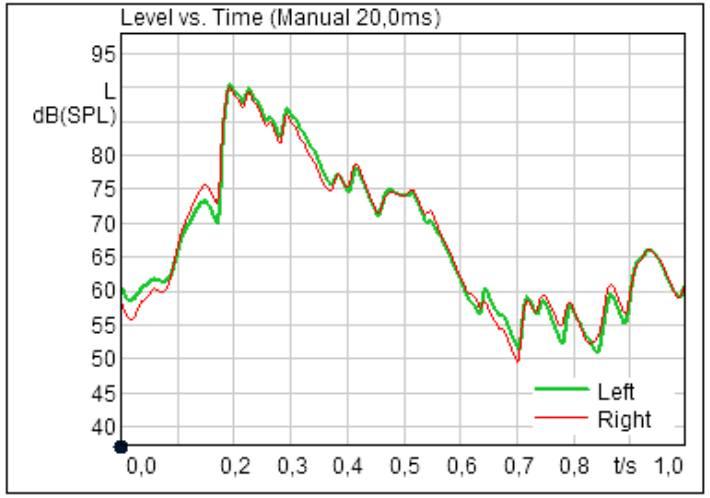


Figure 2 : Results of the analysis level vs. time for manual weighting

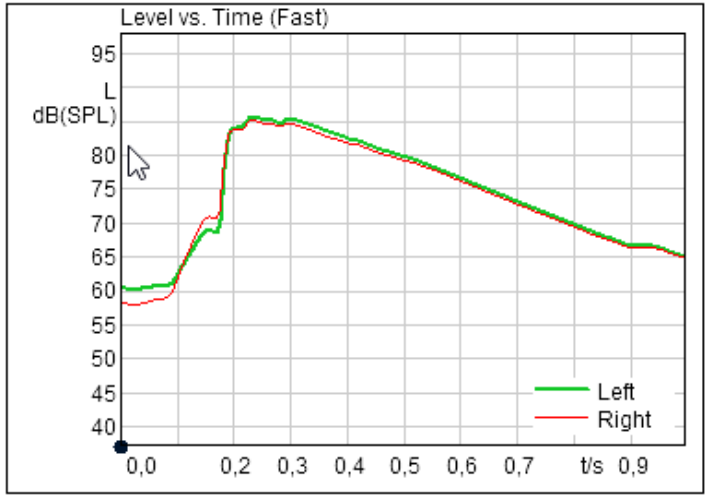


Figure 3 : Results of the analysis level vs. time for fast weighting

- The setting of the time weighting enables you to control the **smoothing** of analysis results.
- For the calculation of, for example, a Level Analysis, the signal must be integrated over time.

  This parameter determines **the duration of the exponential integration**.
- The larger the values, the smoother the resulting curves.

The Contents

2

- Noise_Basic acoustics
- **Sampling_Nyquist-Shannon**
- DSP_Analysis
- Psychoacoustics

Sampling

- Acoustic measurement setup

Sound Source

Sound field

Sensors

Sampling

analog, electrical signal

Frontend

Digitalization

Computer

Sampling

- Acoustic measurement setup

Sound Source

Sound field

Sensors

Sampling

analog, electrical signal

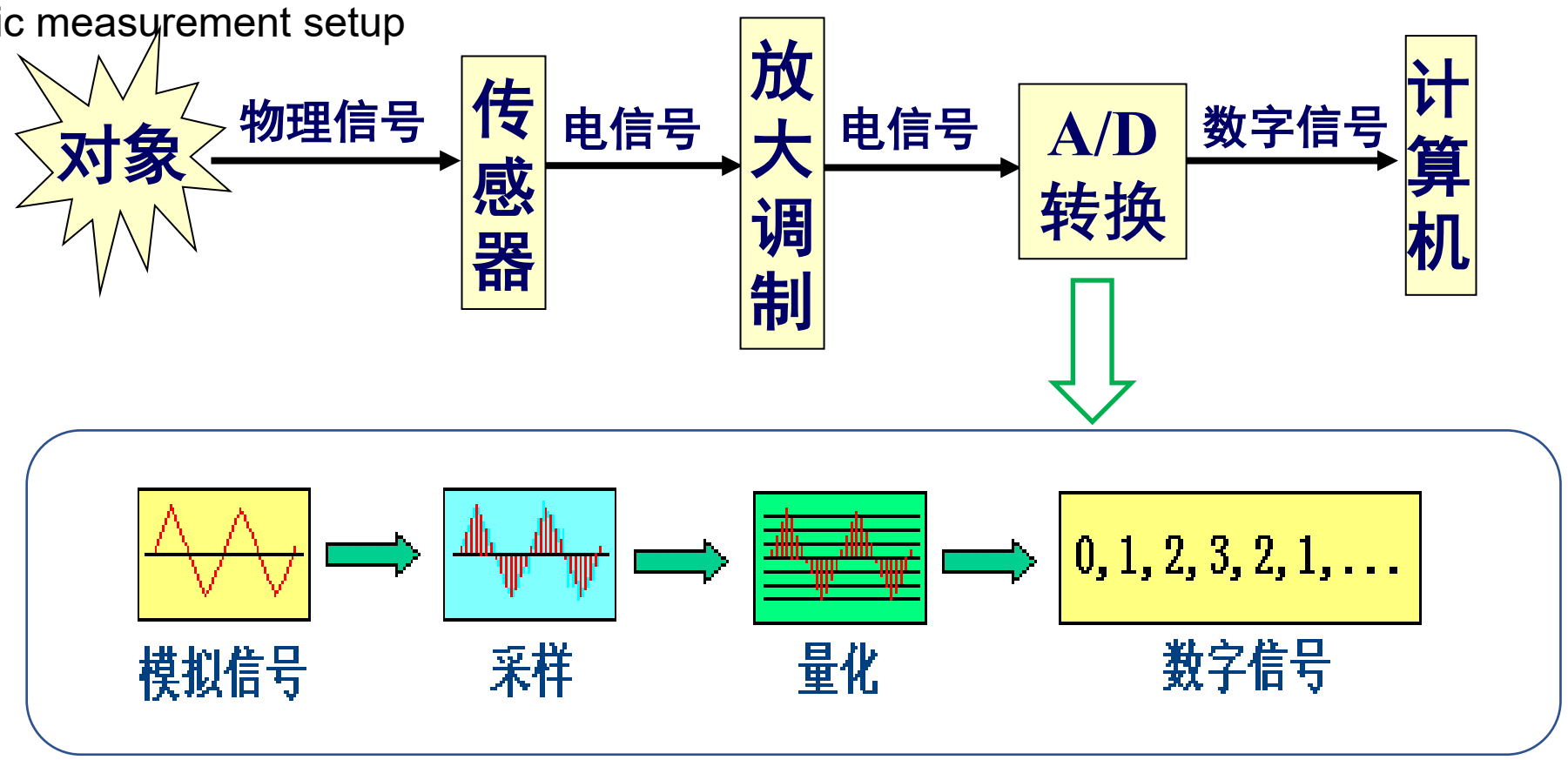
Frontend

Digitalization

Computer

Sampling

- Acoustic measurement setup



采样——利用采样脉冲序列，从信号中抽取一系列离散值，使之成为采样信号x(nTs)的过程.

量化——把采样信号经过舍入变为只有有限个有效数字的数，这一过程称为量化.

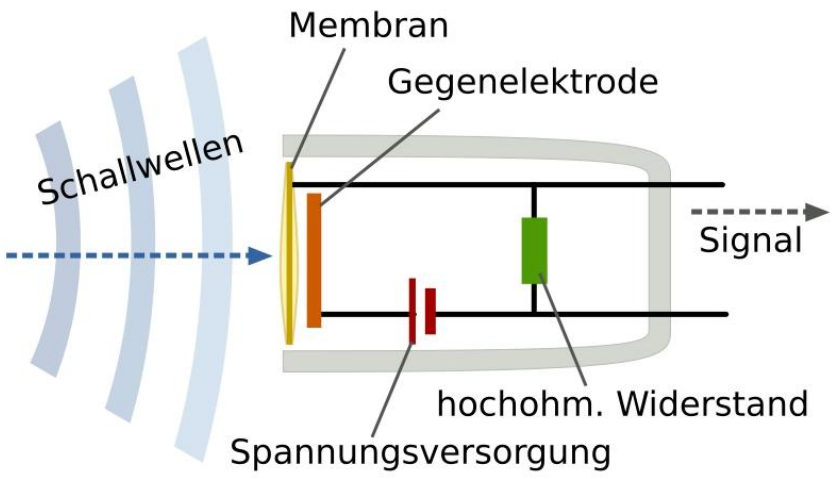
编码——将经过量化的值变为二进制数字的过程。

Sampling_ Sensor

➢ **Condenser microphone**

- Most important tool when measuring sound (**Conversion of the sound pressure to an analog voltage signal**)
- For measurements mostly condenser or electret microphones are used
- Measurement via change of capacity

$$C = \epsilon_0 \cdot \epsilon_r \cdot \frac{A}{d}$$



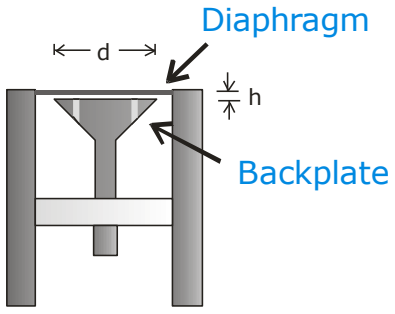
- Differentiation of microphone types by feeding voltage, diameter, sound field correction, accuracy class etc.

Sampling_ Sensor

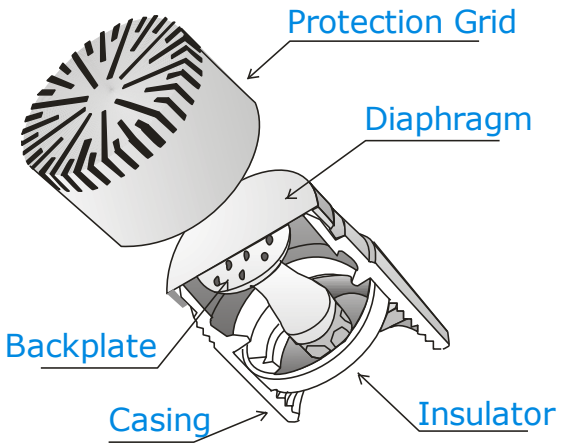
➢ Condenser microphone

- Most important tool when measuring sound (Conversion of the sound pressure to an analog voltage signal)

- For measurements mostly condenser or electret microphones are used
- Measurement via change of capacity
- Differentiation of microphone types by feeding voltage, diameter, sound field correction, accuracy class etc.



(A)



(B)

Sampling_ Sensor

➢ **Microphone**

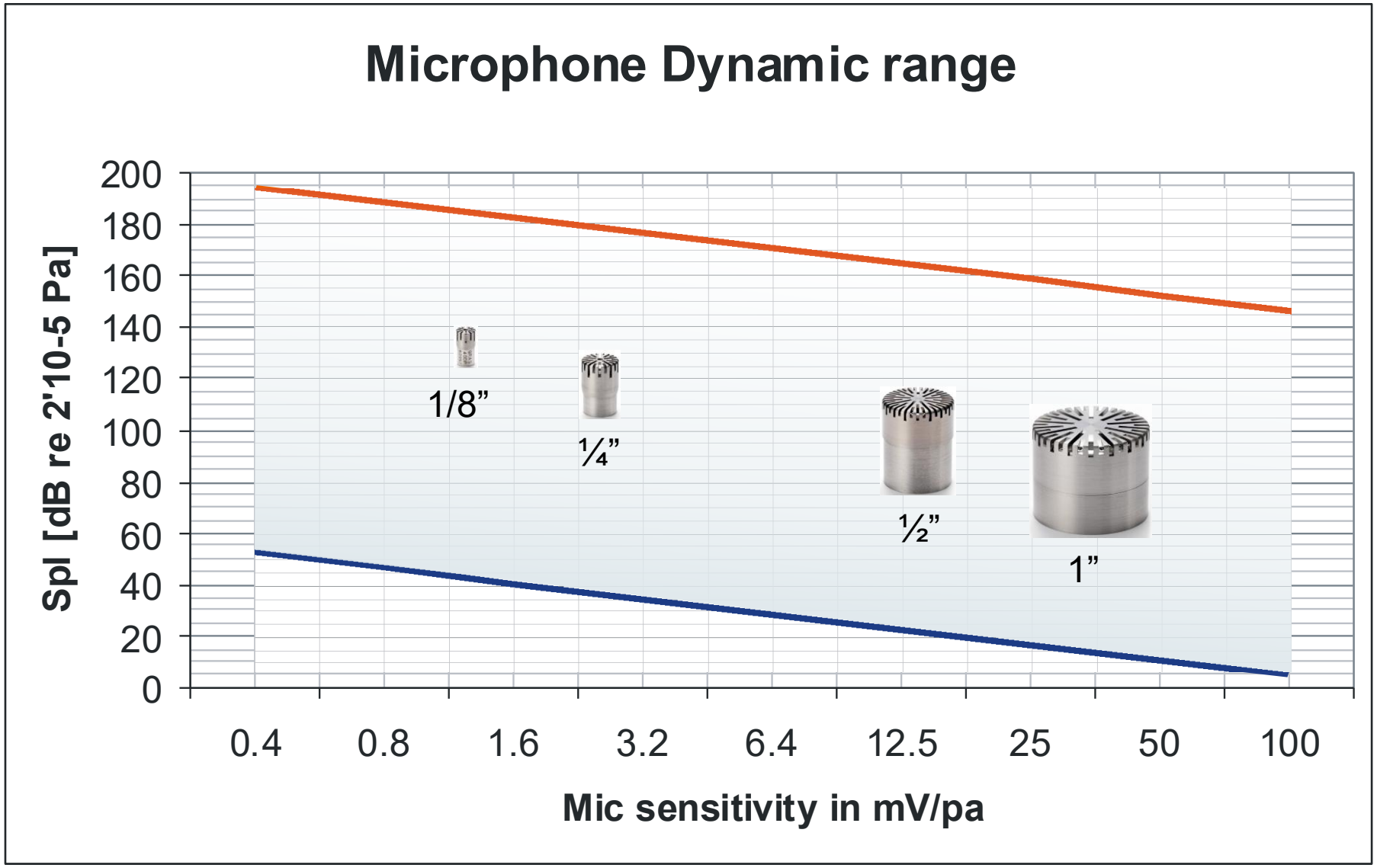
- Principles
  - 200 V Polarization voltage: LEMO-connector
    - Pro: less noise
    - Con: prone to interferences, cable bulky
  - Constant current (ICP, CCP, Deltatron): BNC-connector
    - Pro: less prone to interferences, thin cables can be handled easily
    - Con: more noise

- Membrane diameter(1“, ½“, ¼“)
  - Large: High sensitivity, low lower cut-off frequency
  - Small: Lower sensitivity, small, robust, cheap

- Sound field correction
  - Frequency adjustment for different sound fields (e.g. free field, diffuse field etc.)

Sampling_ Sensor

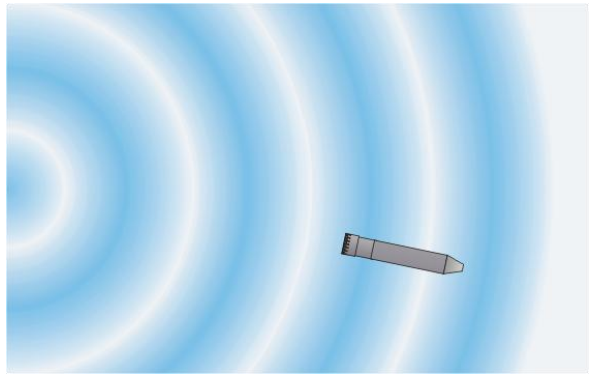
Microphone specification

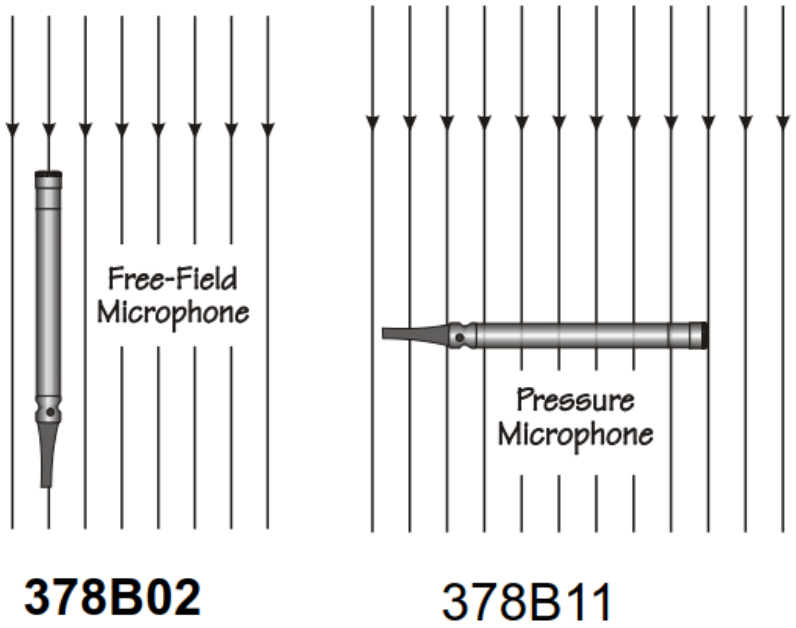


Sampling_ Sensor

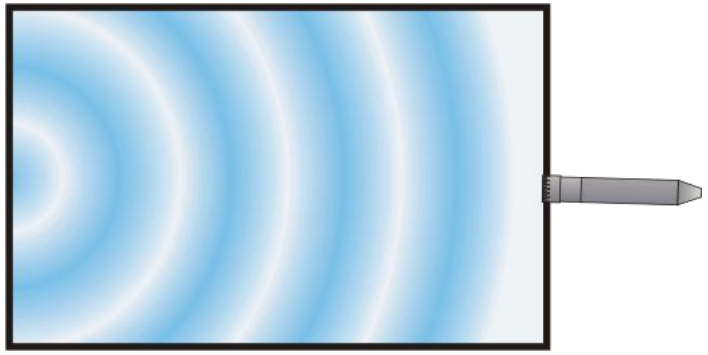
- Types of microphones

Free field microphone

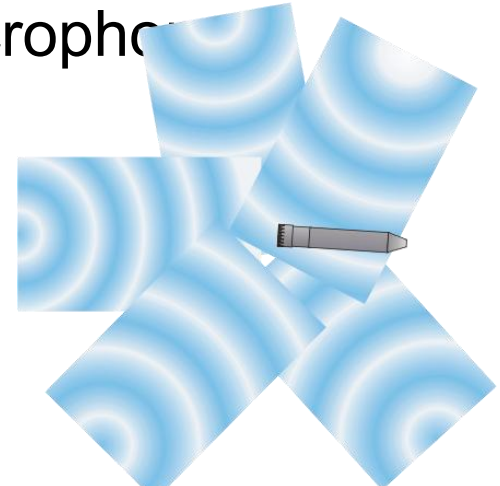


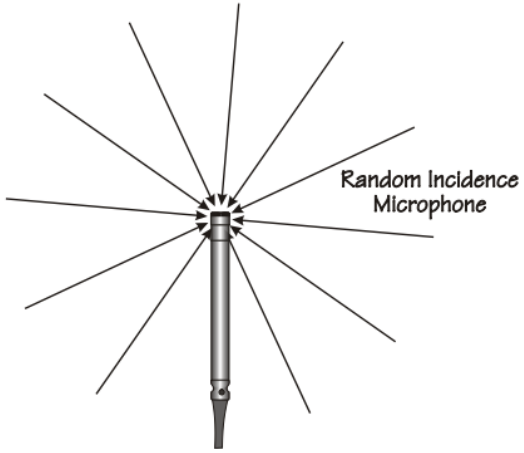


Pressure



Random
microphone



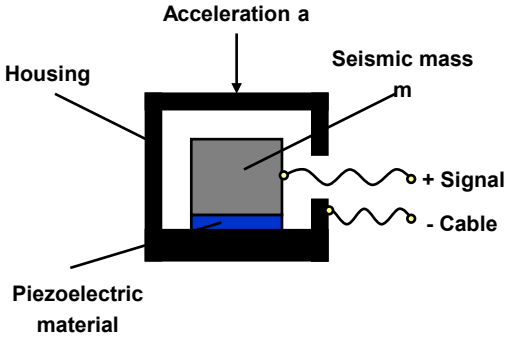


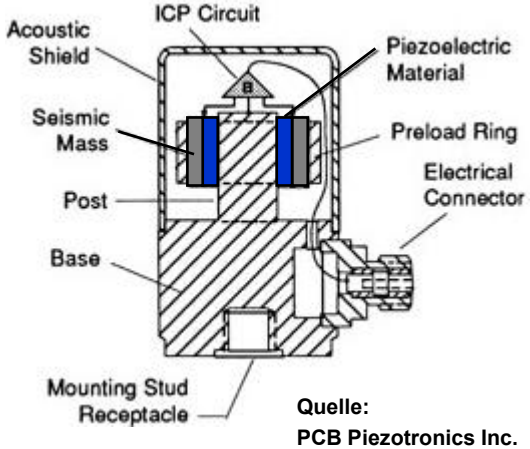
Sampling_ Sensor

- **Binaural sensors**
  - Idea: The sound signal should be recorded exactly as it would be perceived by a person
  - The human body alters the sound signal, the sound recorded at the ears is different to the sound recorded with microphones
  - One needs so choose the proper equalization (Free field, diffuse field, ID (Independent of direction))
  - The signal can be played back via headphones

Sampling_ Sensor

- **Accelerometers**
- Measurement of acceleration via measurement of inertial force to a reference mass
- Piezoelectric Accelerometers
  - Piezoceramic layer converts pressure fluctuations to electrical signals
  - The pressure fluctuation is induced by a defined seismic mass





Quelle:
PCB Piezotronics Inc.

Sampling_ Sensor

- **Accelerometers**

  Distinguishing features between sensors

  - Size
    - Big: higher sensitivity, low lower cut-off frequency, impact on the structure
    - Small: lower sensitivity, robust, little influence on the structure

  - Number of axis
    - Uni-Axial: only vibrations in one axis can be detected
    - Tri-Axial: Acceleration in orthogonal 3 axis
  - Sensor type
    - Charge-sensors: low temperature dependence, high dynamic
    - Constant current: Wiring less interference-prone

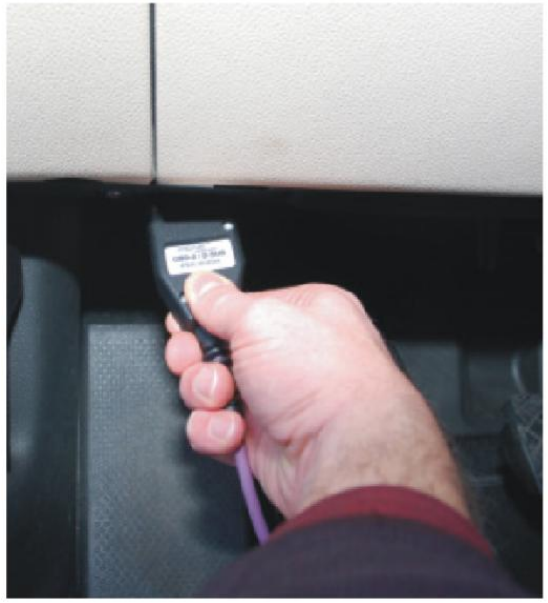
Sampling_ Sensor

- **RPM-Sensors**
  - Sensors to determine the rotating speed
  - Acquisition of motor-speed, vehicle speed etc.

  - Differentiation of sensors by principle of operation and output signal
    - Optical sensor, mostly TTL-Signal
    - Inductive sensor, mostly TTL- or Sinusoidal signal
  - Output frequency is mapped to the desired information

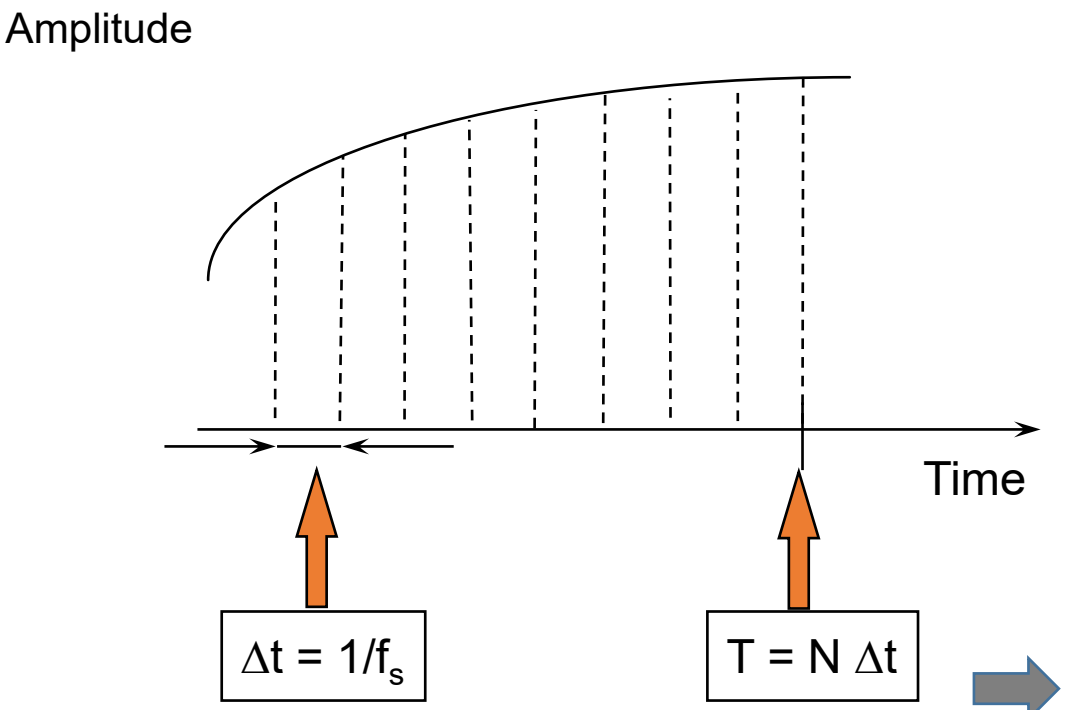
Sampling_ Sensor

- **OBD - CAN Bus**
  - Digital Interface to connect to the frontend to the vehicle data bus
  - OBD (On Board Diagnostics) provides general information regarding the drive state (Speed, RPM, Temperature,…)
  - The CAN (Controller Area Network) Interface provides additional information
  - To be able to decode CAN data a *.dbc file is needed (provided by the car manufacturer)



Sampling_ Quantization

- To store signals on computers, signals must be converted from analog to discrete time and amplitude pairs known in the digital domain. This conversion is completed by an Analog to Digital Convertor (ADC) by SAMPLING.



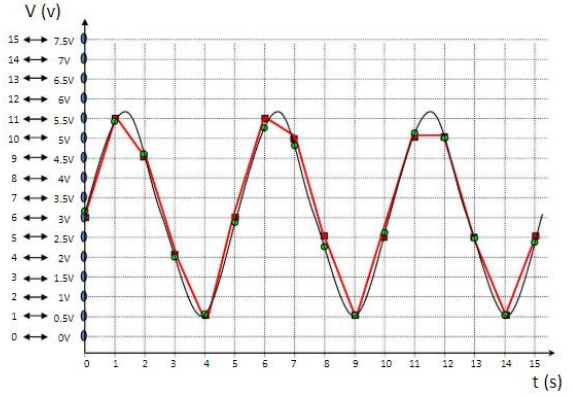
N = Spectrum size,
number of samples acquired

f_s = Sampling frequency

T = block length,
time required for analysis

Time Discretization

Amplitude Discretization

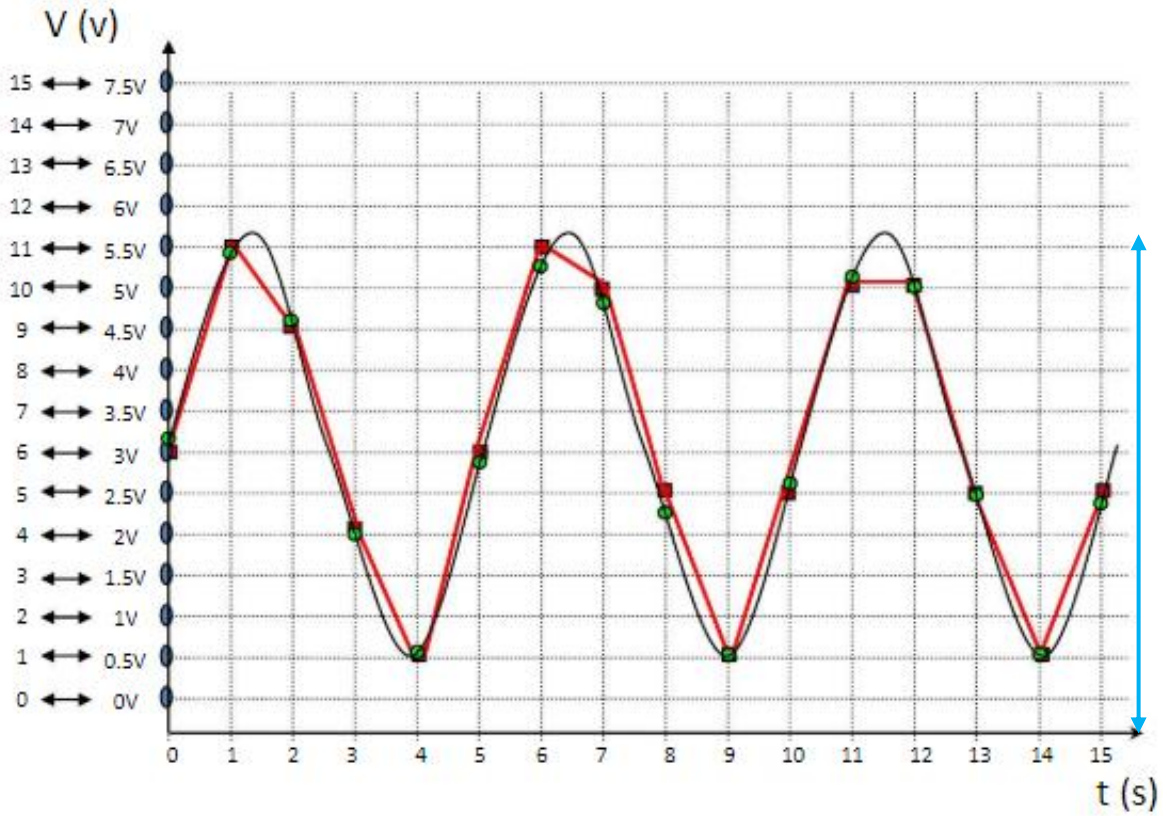


Sampling_ Quantization

- The continuous amplitude of the real time signal will be split up in discrete

levels = $2^{\text{number of bits}}$ -> QUANTIZATION → A/D Converter →

12 bit_ $2^{12} = 4096$
16 bit_ $2^{16} = 65536$
24 bit_ $2^{24} = 16777216$ ...



10v

Level of quantization

Amplitude(v) / bit	10
12 bit_$2^{12} = 4096$	$\frac{10}{2^{12}} = 0.0024$
16bit_$2^{12} = 4096$	$\frac{10}{2^{16}} = 0.00015$
24bit_$2^{24} = 16777216$	$\frac{10}{2^{24}} = 0.00000059$

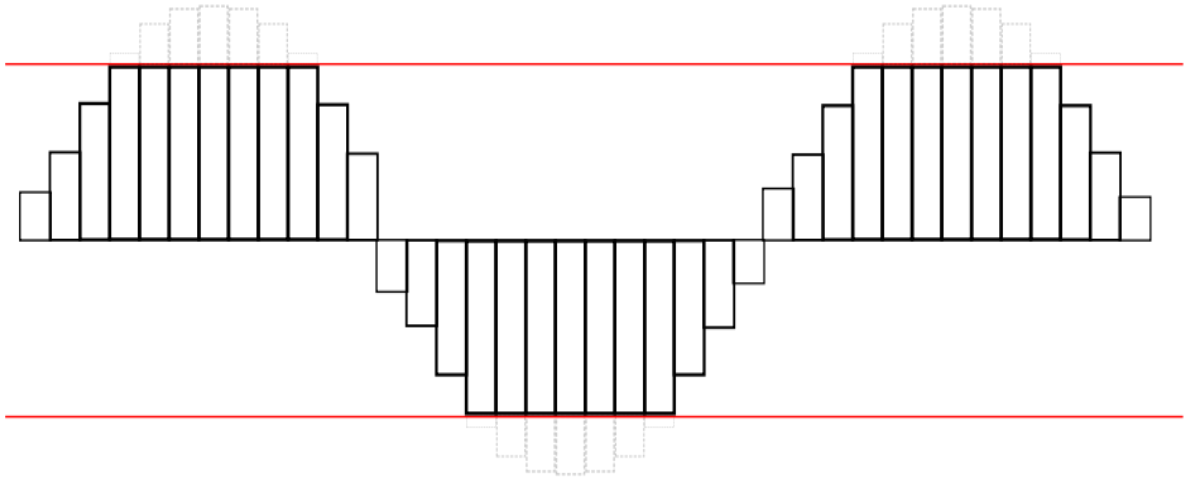
Sampling_ Quantization

- Quantization_ Range

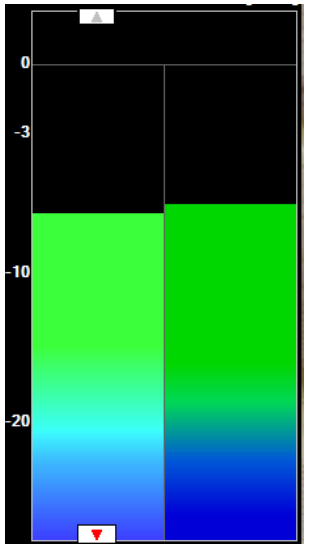
$$dB = 10lg\frac{Power_{signal}}{Power_{noise}}$$

(>63dB/ICE)

- Adjust the range with care, otherwise the signal may clip!



- Adjust the range with care, improve signal-noise ratio



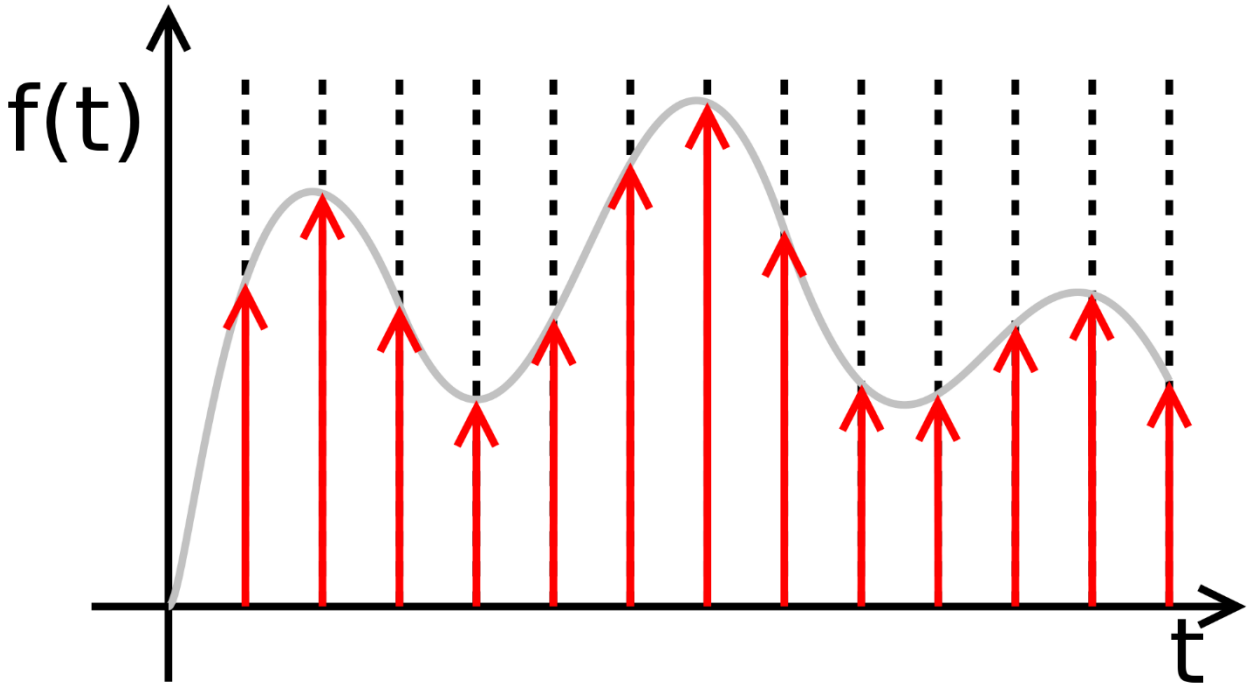
Sampling_ Nyquist –Shannon threshold

Can an Analog Signal be stored in a computer?

- No!

- Computers do not and never will have infinite memory!
- However, reasonable, highly accurate facsimiles of analog time signals from microphones, accelerometers or other transducers can be captured and stored in computer based systems.
- What makes this highly desirable is computer processing of those signals.

Sampling_ Nyquist –Shannon threshold

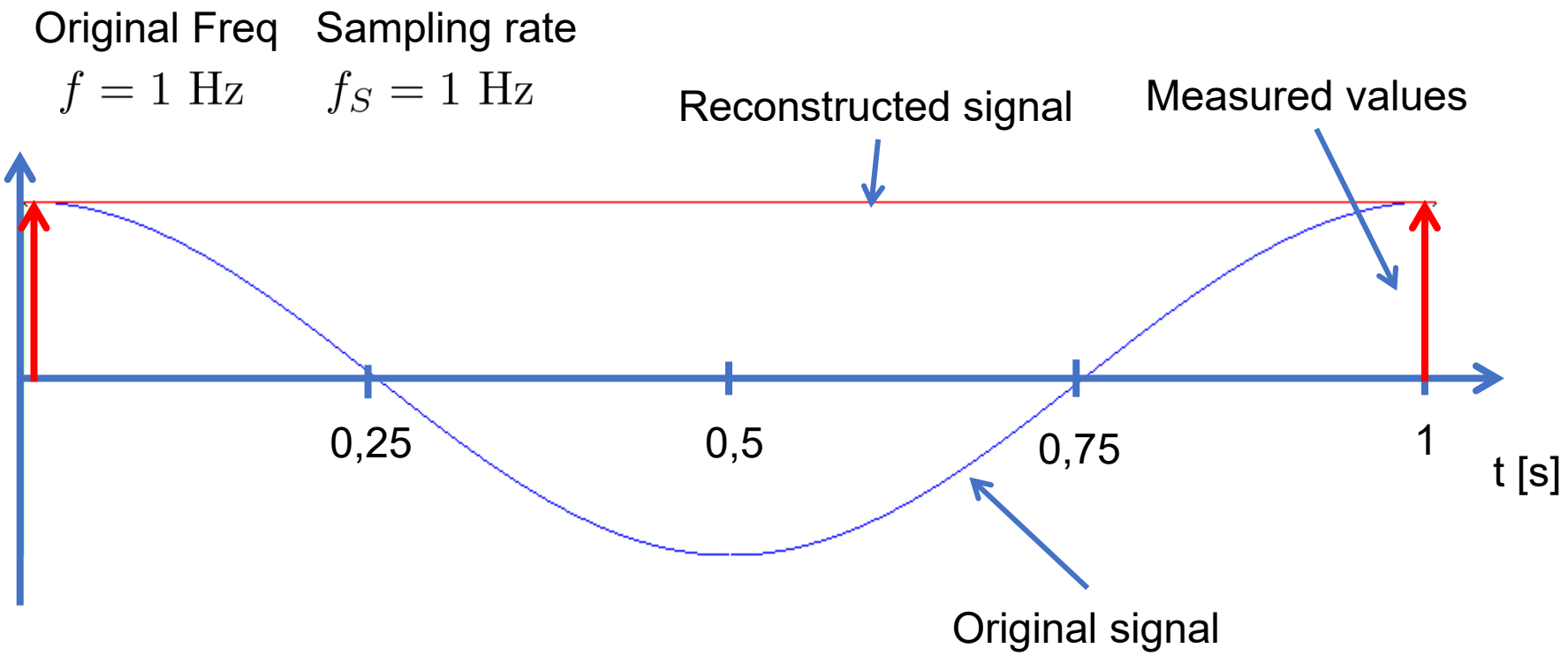


Nyquist -Shannon sampling theorem:

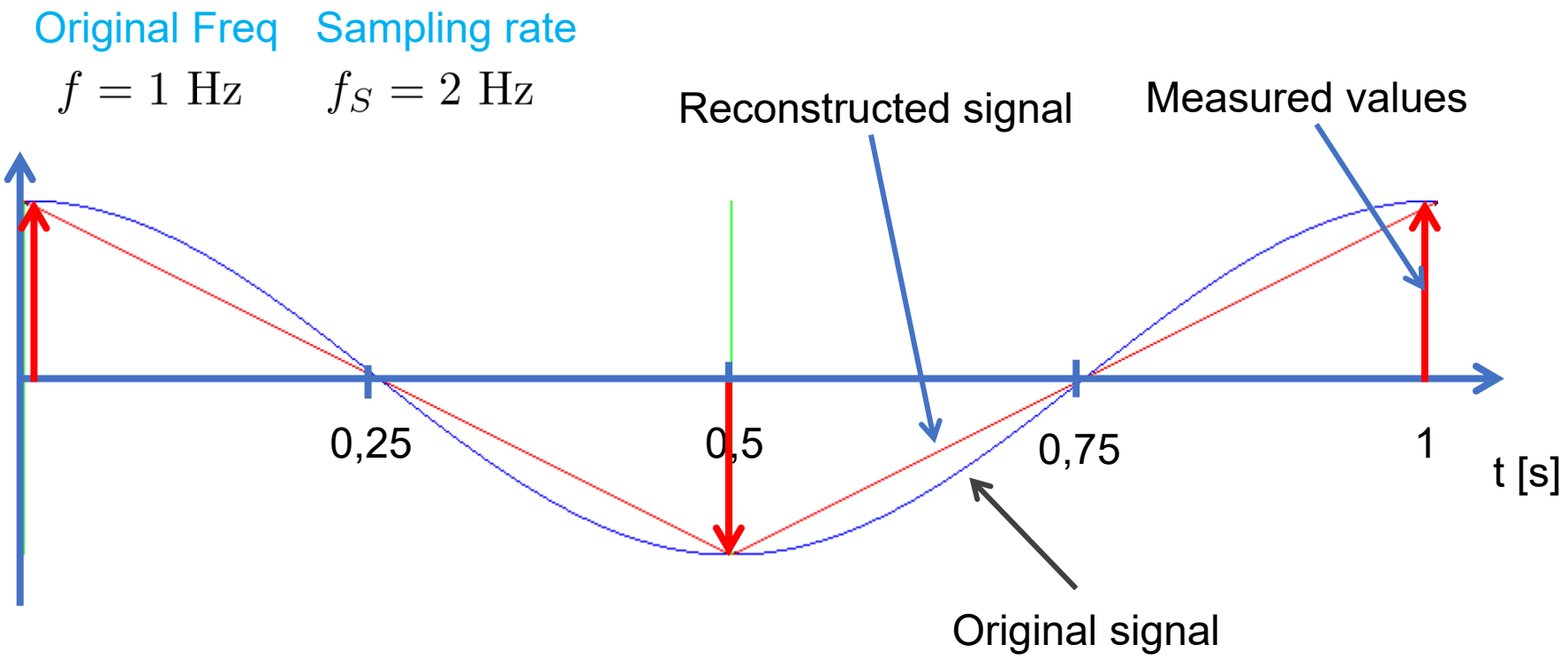
Exact reconstruction of the original signal from time discrete data for

$$f_s > 2 \cdot f_{max}$$ 2.56 ?

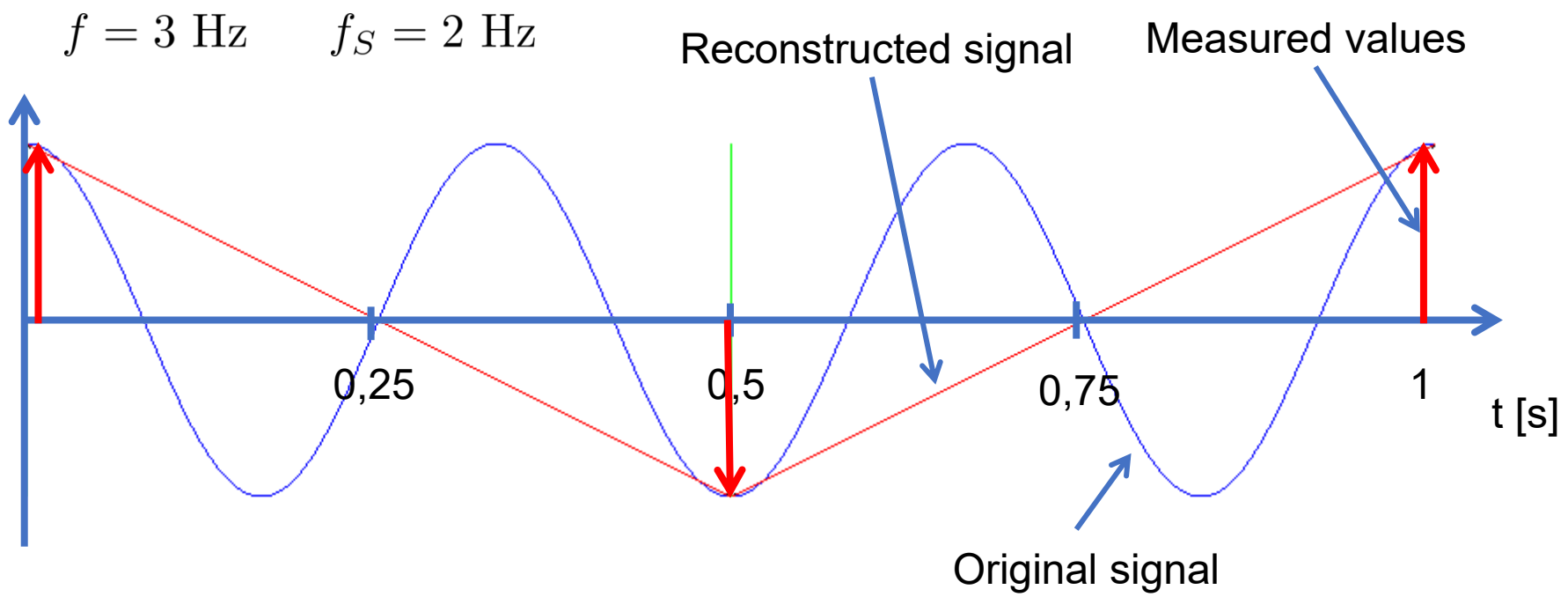
Sampling_ Nyquist –Shannon threshold



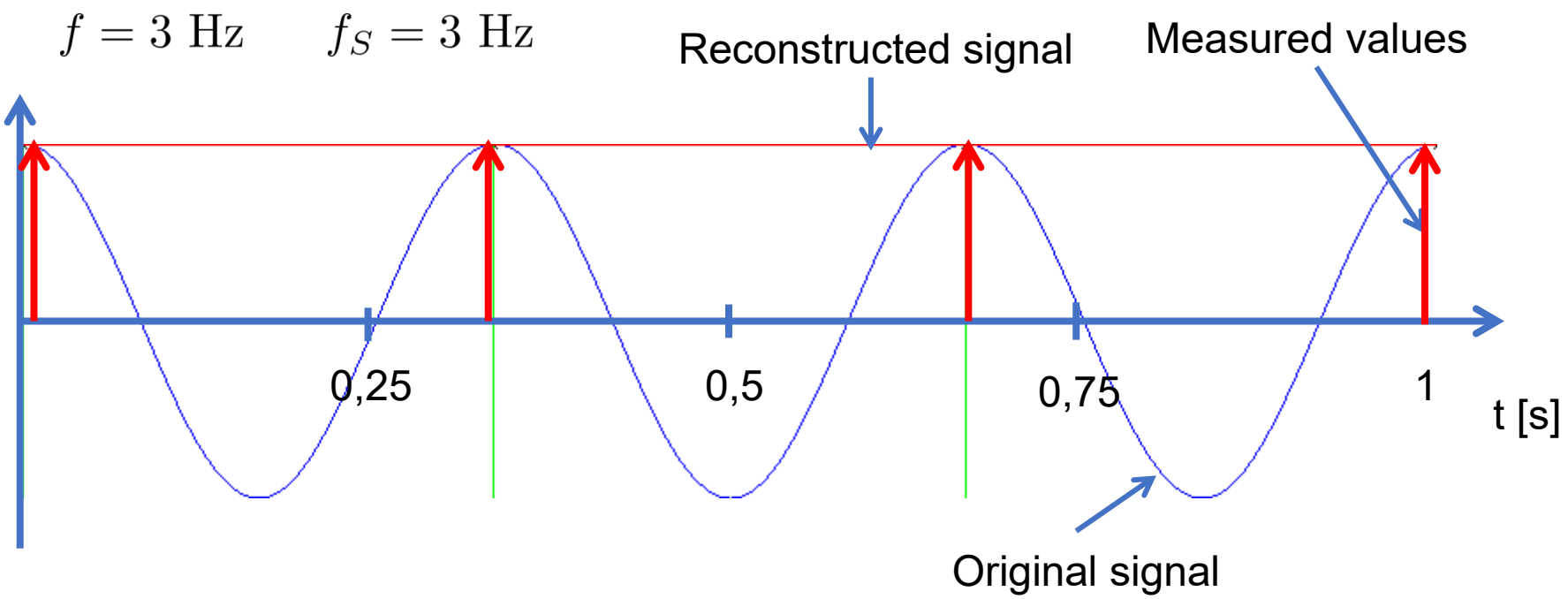
Sampling_ Nyquist –Shannon threshold



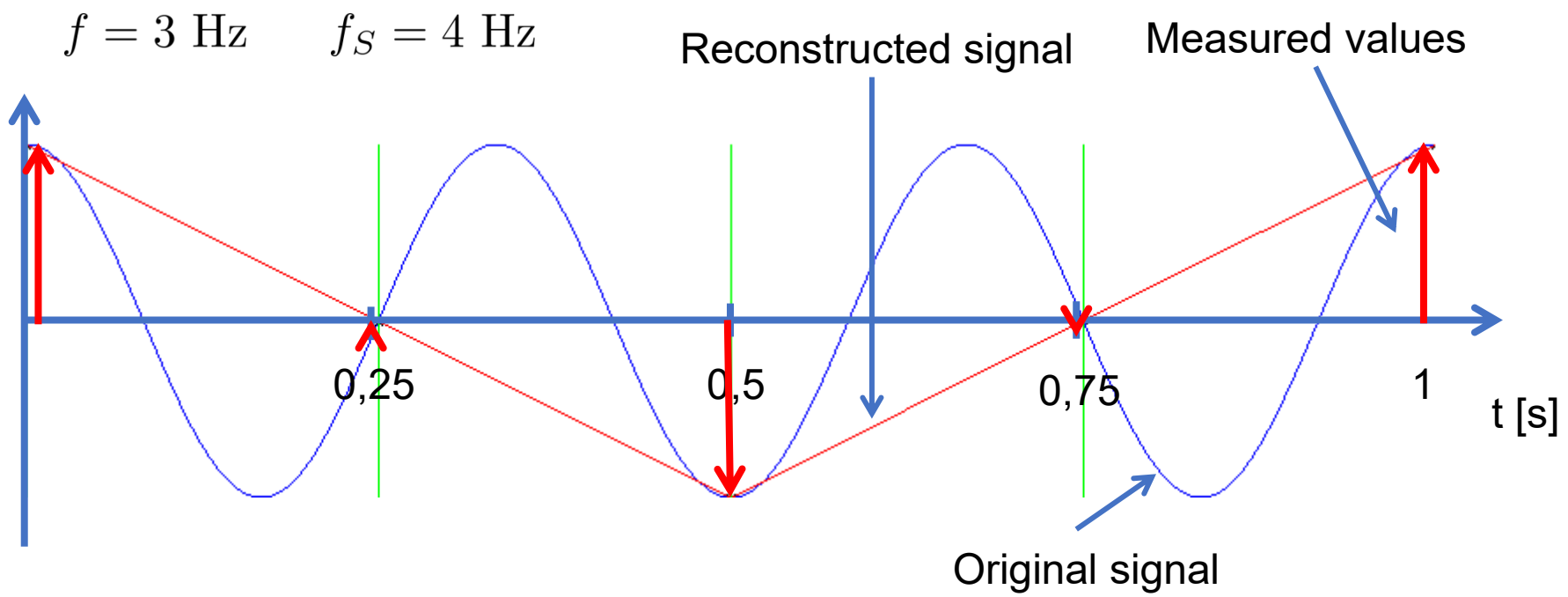
Sampling_ Nyquist –Shannon threshold



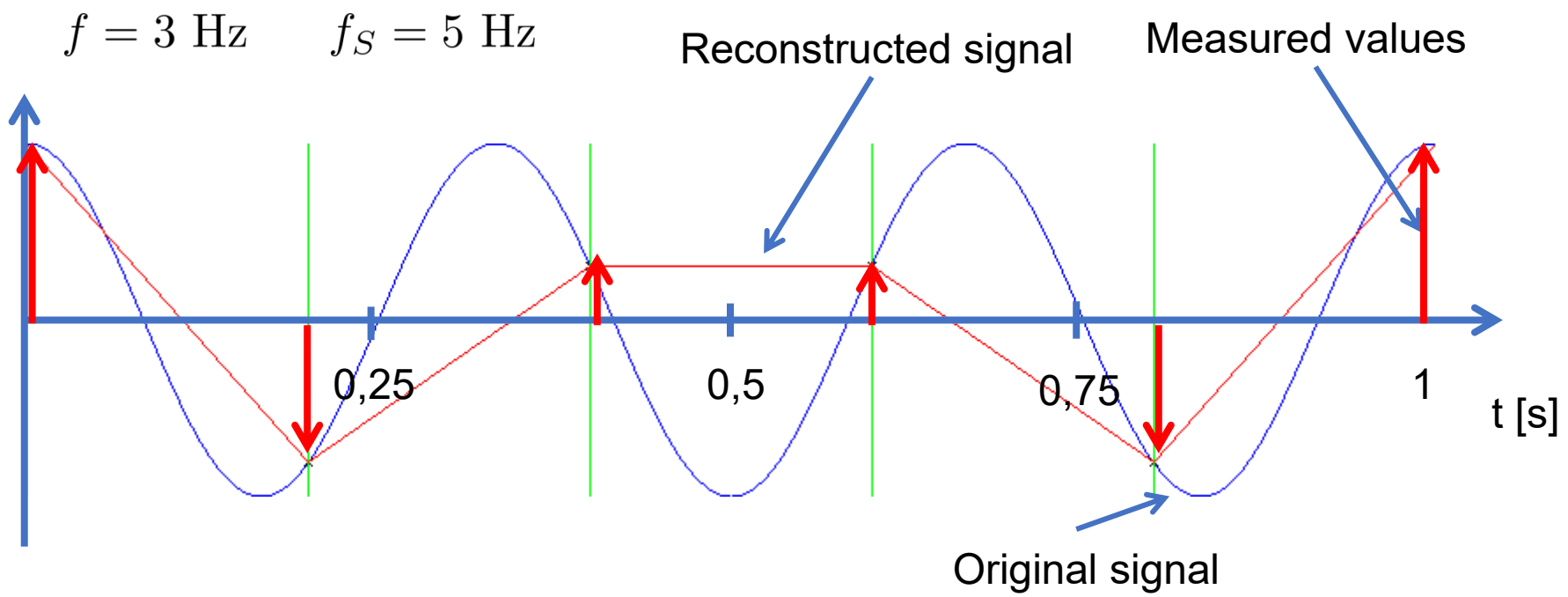
Sampling_ Nyquist –Shannon threshold



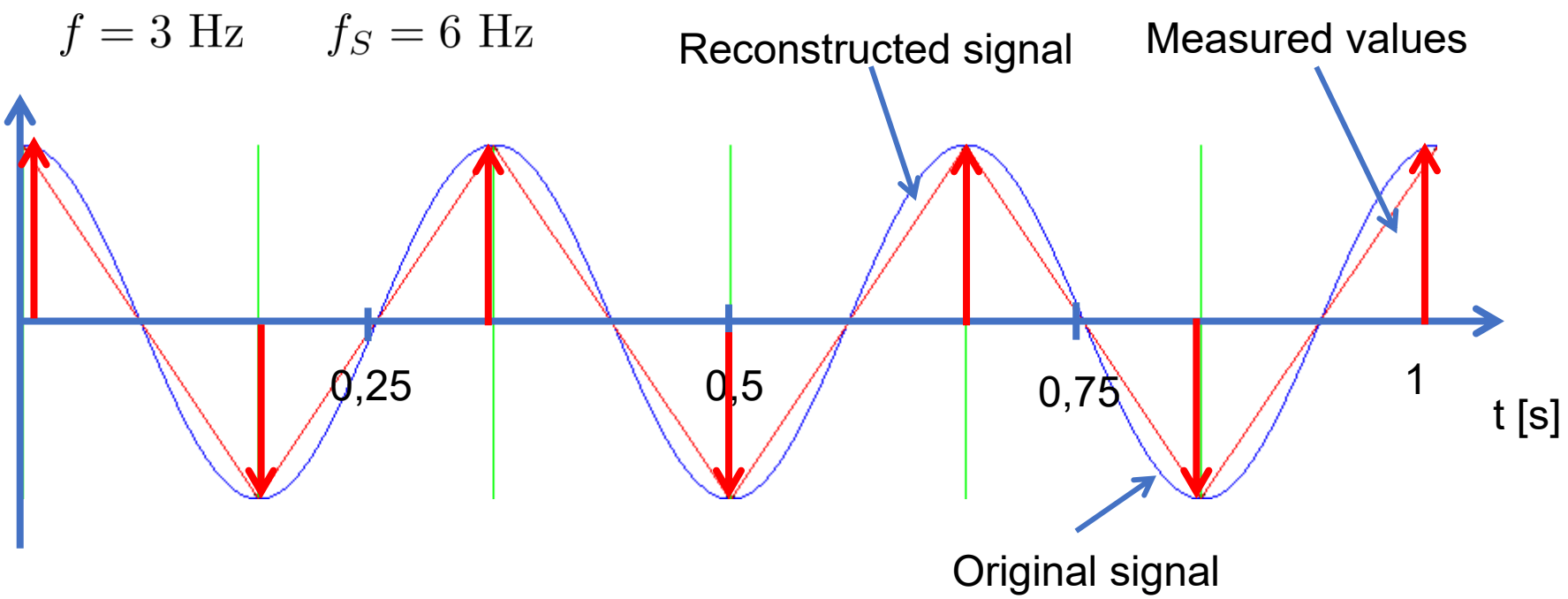
Sampling_ Nyquist –Shannon threshold



Sampling_ Nyquist –Shannon threshold

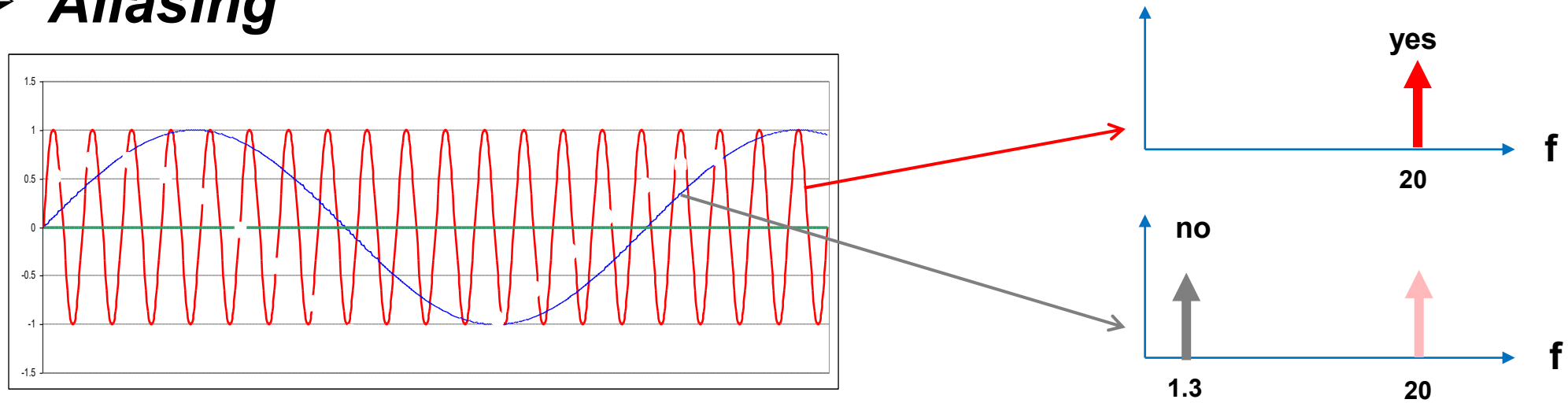


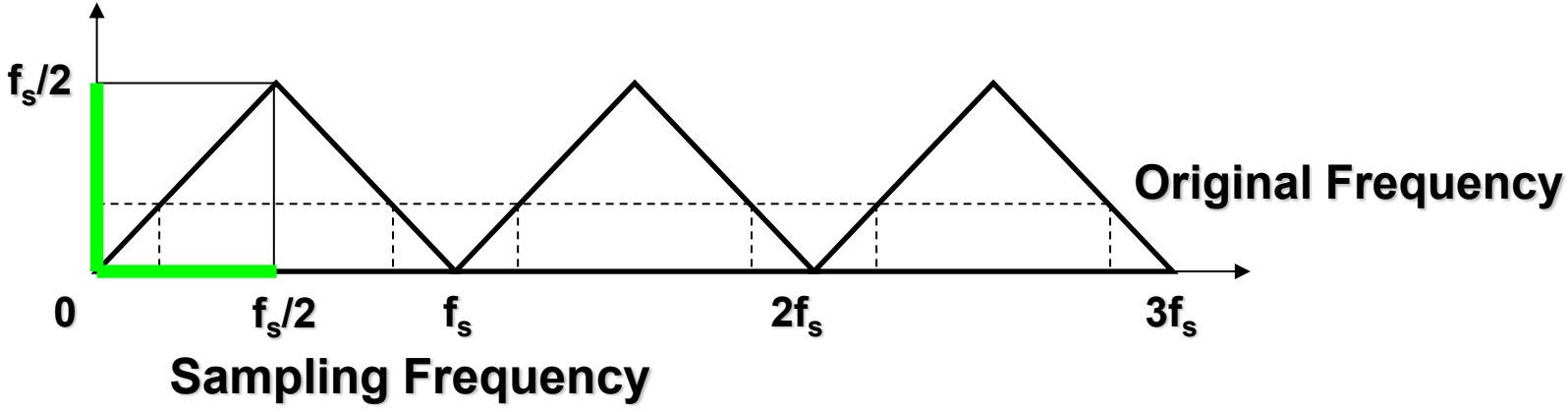
Sampling_ Nyquist –Shannon threshold



Sampling_ Nyquist –Shannon threshold

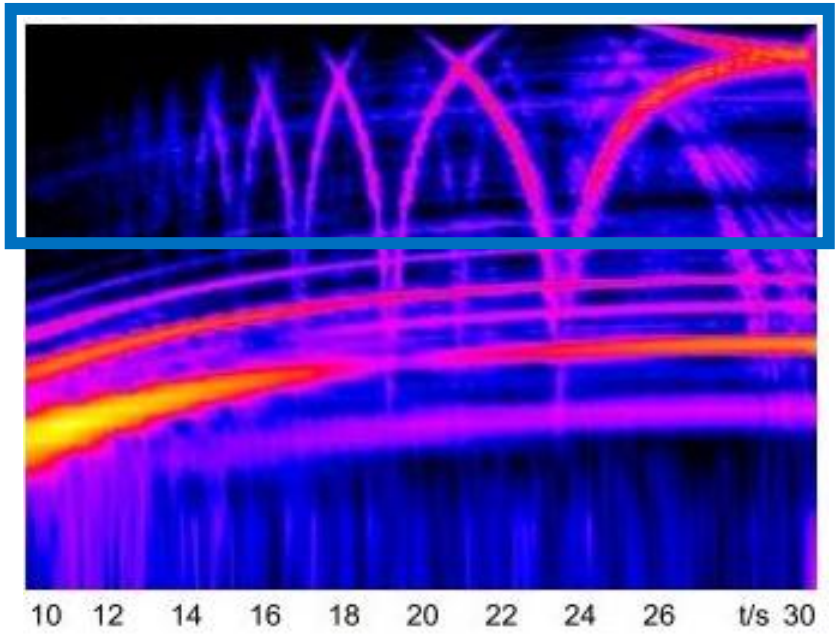
- ***Aliasing***





Sampling_ Nyquist –Shannon threshold

- ***Aliasing***



Window **Solution: Anti-aliasing filter** $2^8 = 256$ …

Sampling_ Resolution

$Sampling\ Rate(Bandwith) - F_s$

$Analysis\ Bandwith - F_a$

$Time\ Resolution - t_0$

$Frequency\ Resolution - f_0$

$Spectrum\ Size(FFT\ Size) - F_{size}$

$$F_s = 2 \times F_a \ \ldots\ldots (1)$$

$$F_a = F_{size} \times f_0 \ldots (2)$$

$$t_0 = \frac{1}{f_0} \ldots\ldots\ldots\ldots (3)$$

$$\boldsymbol{F_s} \uparrow \rightarrow \boldsymbol{F_a} \uparrow \rightarrow \boldsymbol{f}_0\ \boldsymbol{Range} \uparrow \rightarrow \boldsymbol{t}_0 \boldsymbol{Range} \uparrow$$

$$\boldsymbol{Data = Channels \times Sampling\ Rate \times Bit \times Sampling\ Time}$$

The Contents

3

- Noise_Basic acoustics
- Sampling_Nyquist-Shannon
- **DSP_Analysis**
- Psychoacoustics

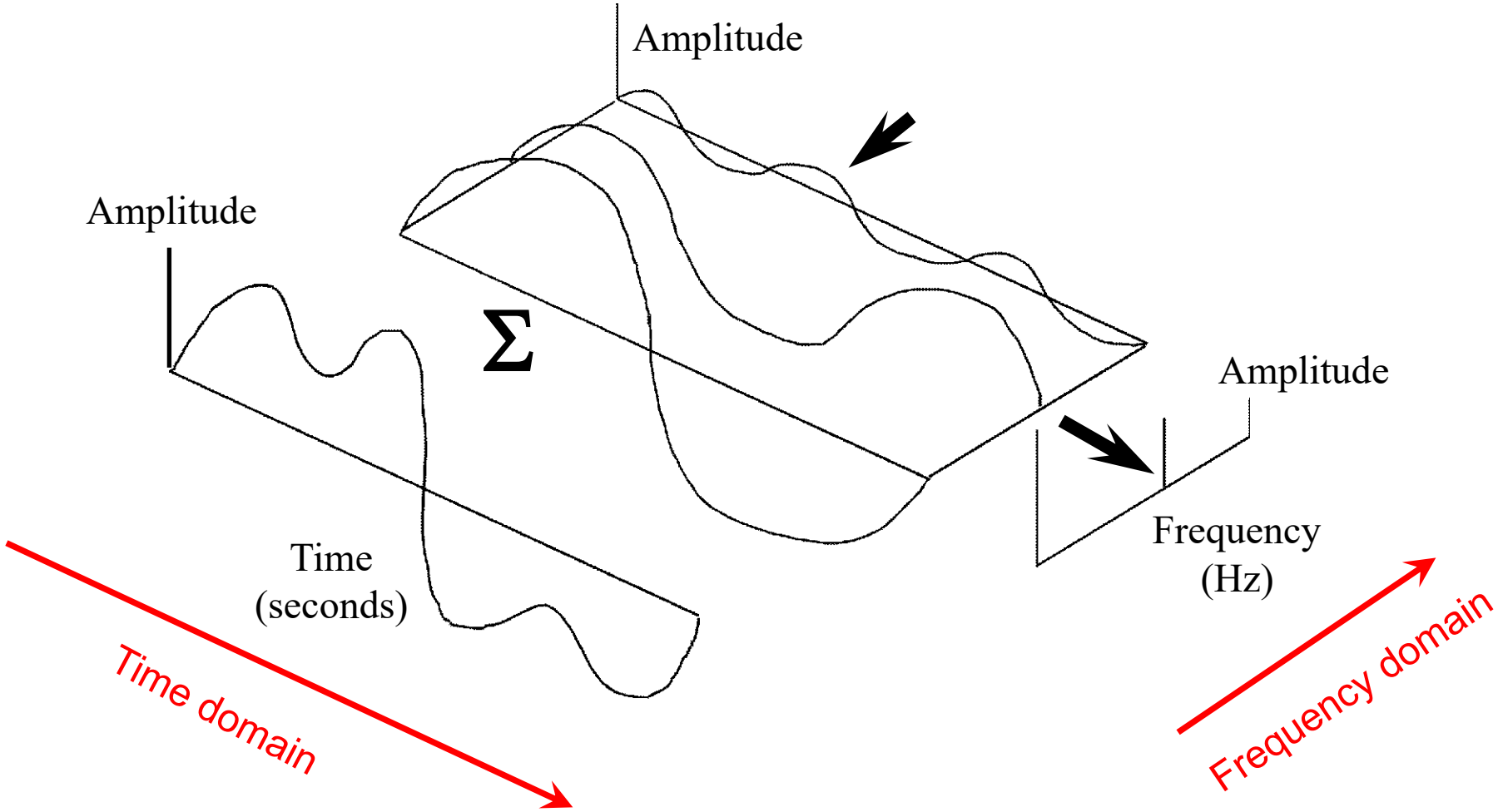
DSP_ Analysis/FFT

- History - FFT = Fast Fourier Transform
  - Named with regard to Jean Baptiste Joseph Fourier (1768-1830), who created the concept of the “Fourier series”
  - He found that any signal can be approximated using sine and cosine curves with different frequencies and amplitudes (example on next slide)

$$X(\omega) = \int_{-\infty}^{+\infty} x(t)\, e^{-j\omega t}\, dt$$

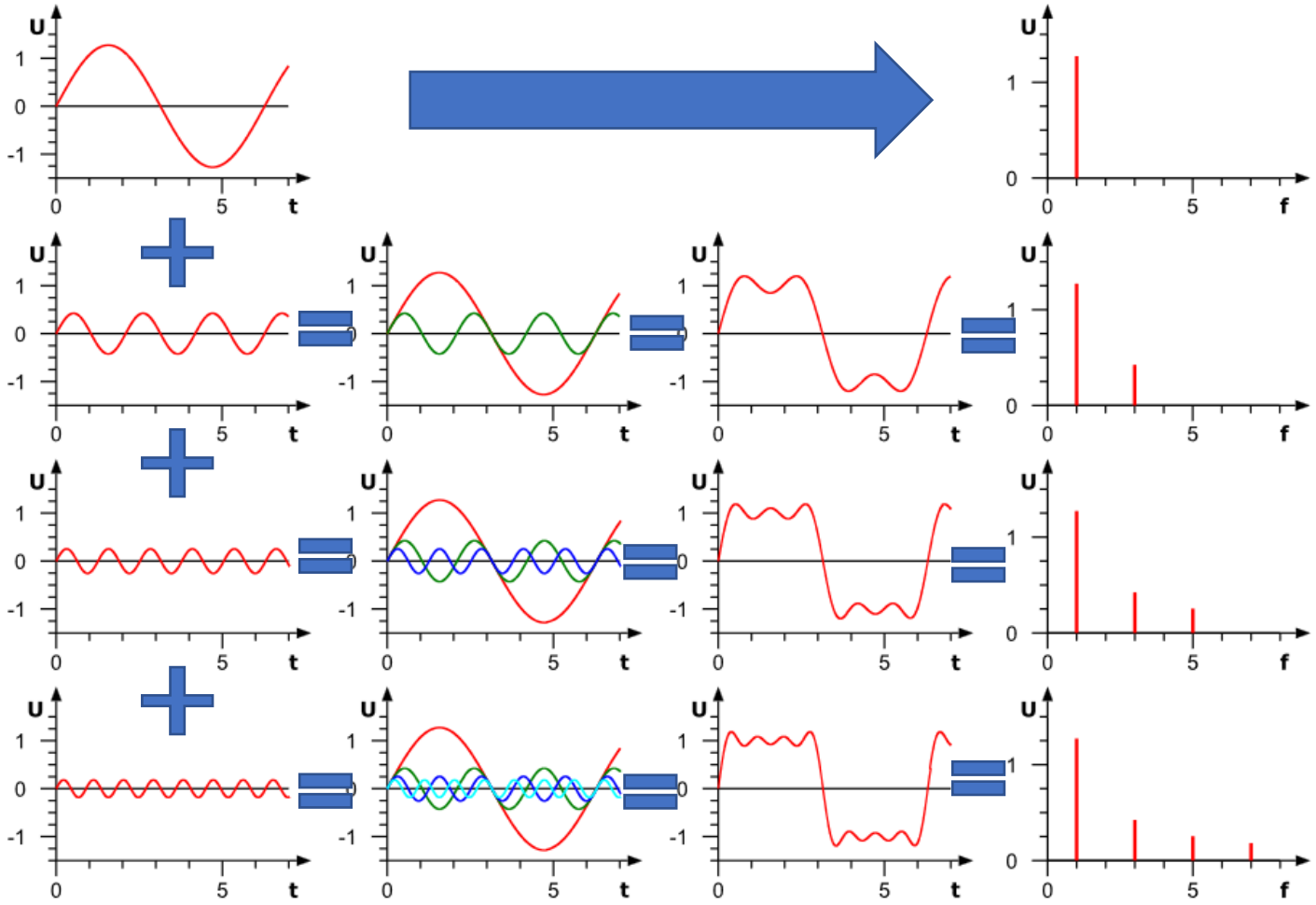
$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(\omega)\, e^{j\omega t}\, d\omega$$

DSP_ Analysis/FFT



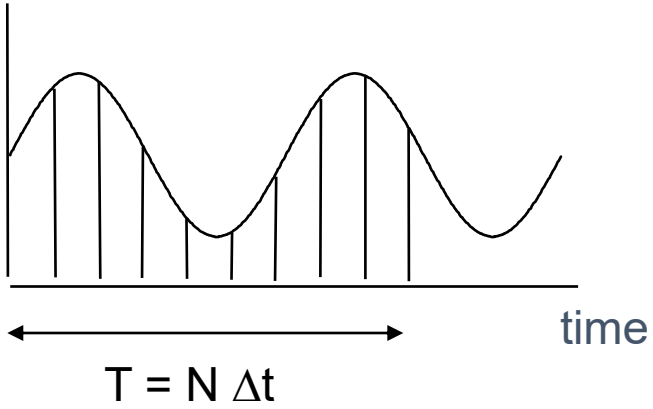
DSP_ Analysis/FFT

- FFT Analysis



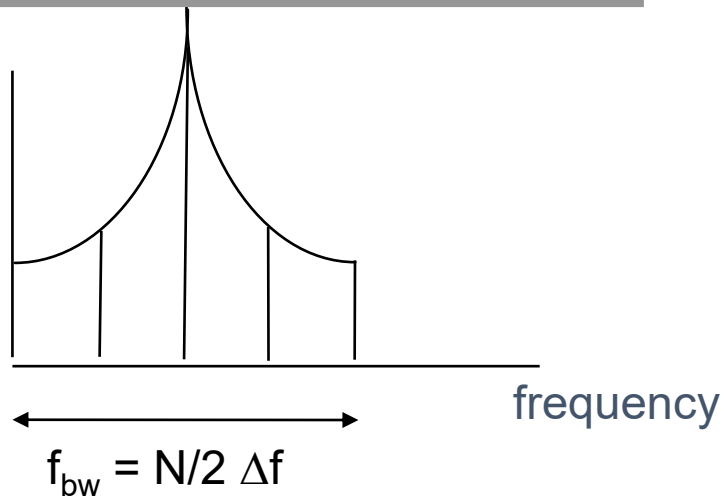
DSP_ Analysis/FFT

$f_s = Sampling\ frequency$
$2.56 f_s = f_a (analysis frequency)$



FFT

$f_a = \Delta f \times N$ (Spectral lines)
$2.56 f_s = \Delta f \times N$



Δt = Sampling interval

$\Delta t = 1/f_s$

f_{bw} = Bandwidth

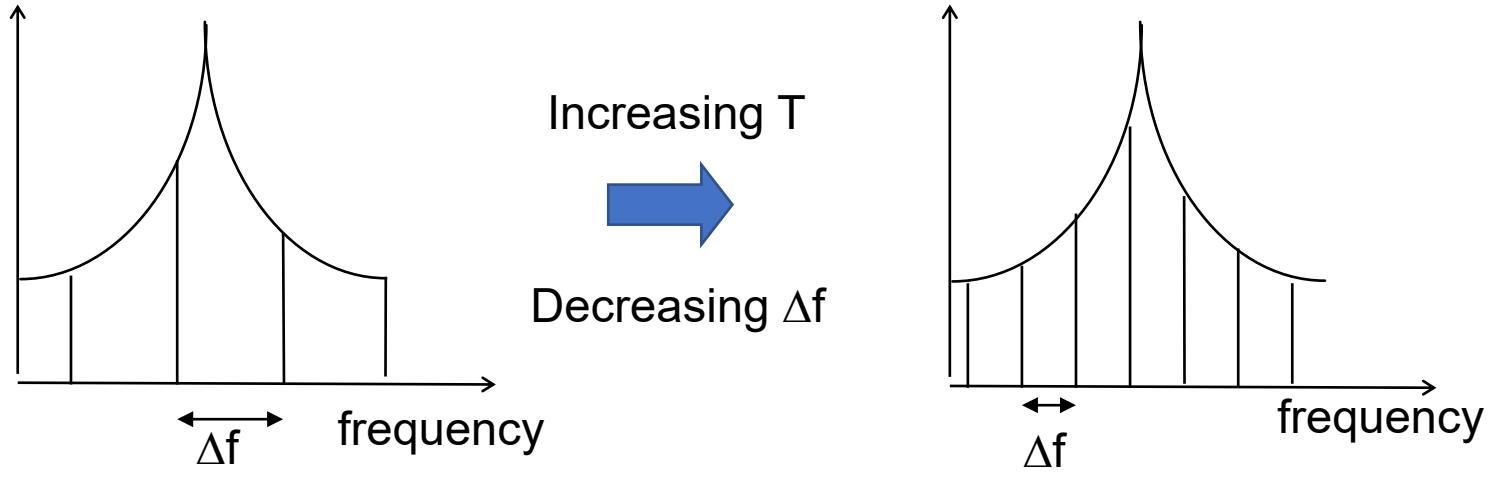
$\Delta f = 1/T$

DSP_ Analysis/FFT

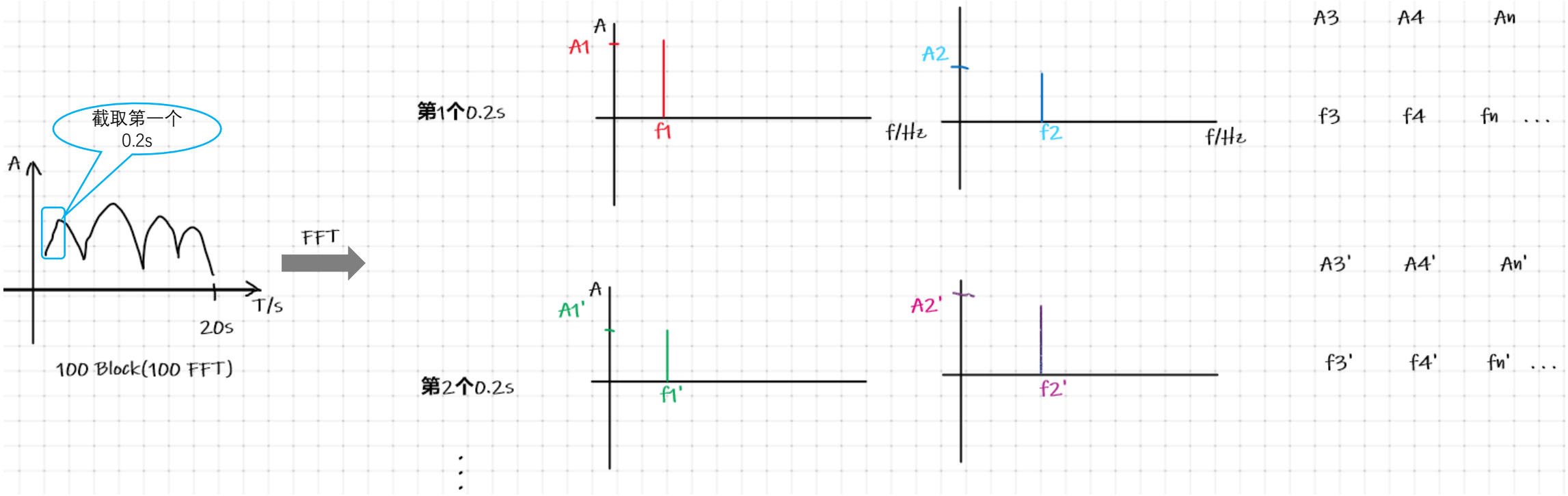
- Decreasing acquisition time T corresponds to a bigger Δf, it means lower frequency resolution.

- A finer desired frequency resolution Δf corresponds to a longer Block time T.

$$\frac{1}{T} = \Delta f$$



DSP_ Analysis/FFT



$$FFT(Average) = RMS = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2} = \sqrt{\frac{1}{100}\sum_{i=1}^{100} A_i^2}$$

FFT(Average)_提取每个Block在不同频率下的峰值计算RMS

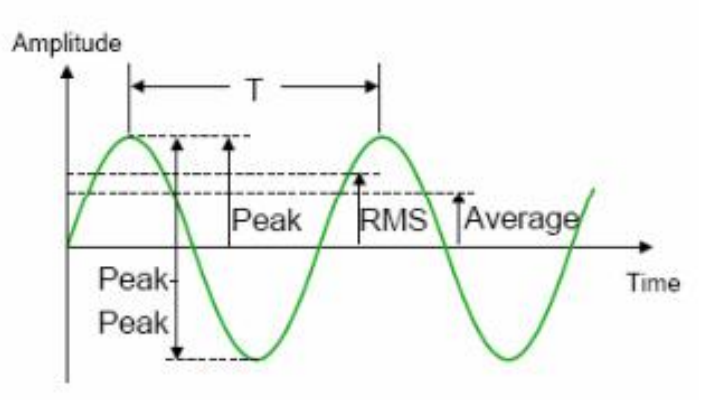
<u>Average</u> _先是A1与A1’平均B，然后B再与第三个0.2s的值A1”平均，依次类推…（原则是边FFT边平均，这样能缩短计算时间）

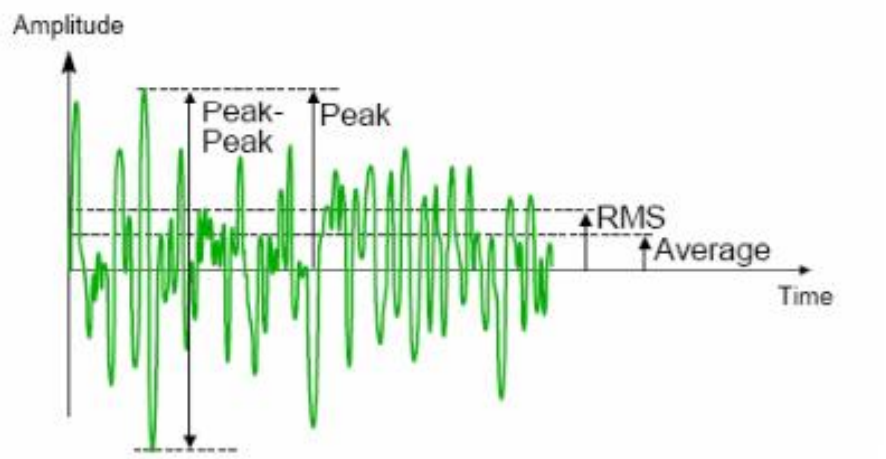
$$FFT(Peak_Hold) = \frac{1}{n}\sum_{i=1}^{n} x_{i/Max}$$

FFT(Peak_ Hold)_提取每个Block在不同频率下最大峰值计算算术平均

即把 f1对应的A1和f1’对应的A1‘比较，如A1’>A1,则取A1‘ 计算.

DSP_ Analysis/FFT

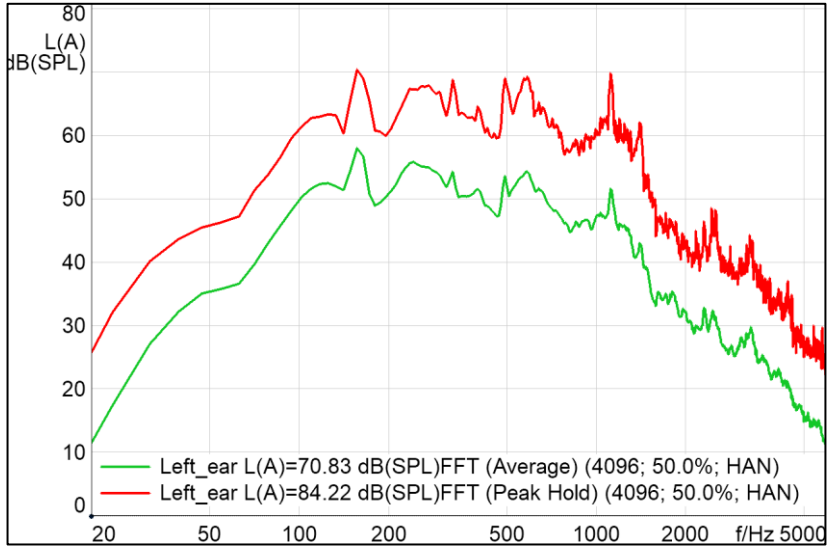




$$Average = \frac{1}{T}\int_0^T |x(t)|dt$$

$$RMS = \sqrt{\frac{1}{T}\int_0^T |x^2(t)|dt}$$

$$Peak = \sqrt{2}RMS$$



DSP_ Analysis/LEVEL

- The level analysis is a simple signal analysis to represent the amplitude of a signal as a logarithmic value in dB
  - In the most common form of the level analysis the measured sound pressure p(t) in pascal and the reference value for airborne sound $p_0 = 2 \cdot 10^{-5}$ Pa are used to determine the sound pressure level $L_p(t)$ in dB.
  - level analyses can also be used to visualize the course of a voltage or to evaluate the signal of an accelerometer. With regard to this versatile usability closely attention to the given scaling quantity is advised. The different characteristics are described on the overview page dB Scaling. However, to keep this description simple, the following text refers to the Sound Pressure Level analysis only.

$$L_p(t) = 20 \cdot \log\left(\frac{p(t)}{2 \cdot 10^{-5} Pa}\right) dB$$

DSP_ Analysis/Order

- The level analysis is a simple signal analysis to represent the amplitude of a signal as a logarithmic value in dB
  - In the most common form of the level analysis the measured sound pressure p(t) in pascal and the reference value for airborne sound $p_0 = 2 \cdot 10^{-5}$ Pa are used to determine the sound pressure level $L_p(t)$ in dB.
  - level analyses can also be used to visualize the course of a voltage or to evaluate the signal of an accelerometer. With regard to this versatile usability closely attention to the given scaling quantity is advised. The different characteristics are described on the overview page dB Scaling. However, to keep this description simple, the following text refers to the Sound Pressure Level analysis only.

$$L_p(t) = 20 \cdot \log\left(\frac{p(t)}{2 \cdot 10^{-5} Pa}\right) dB$$

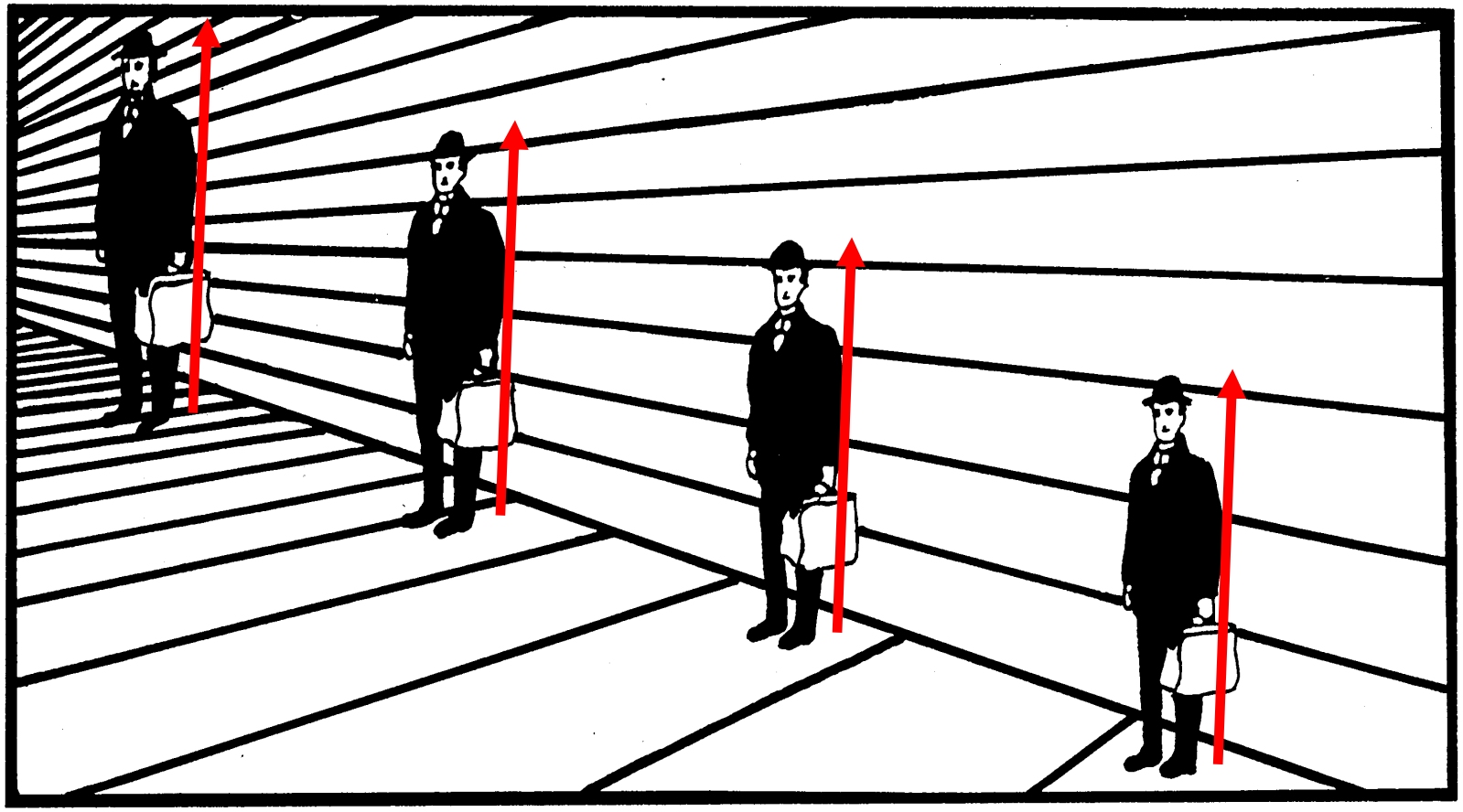
The Contents

4

- Noise_Basic acoustics
- Sampling_Nyquist-Shannon
- DSP_Analysis
- **Psychoacoustics**

Why Psychoacoustics?

- Problem: objective description of subjective impression



Why Psychoacoustics?

Sound Development trends for the vehicle

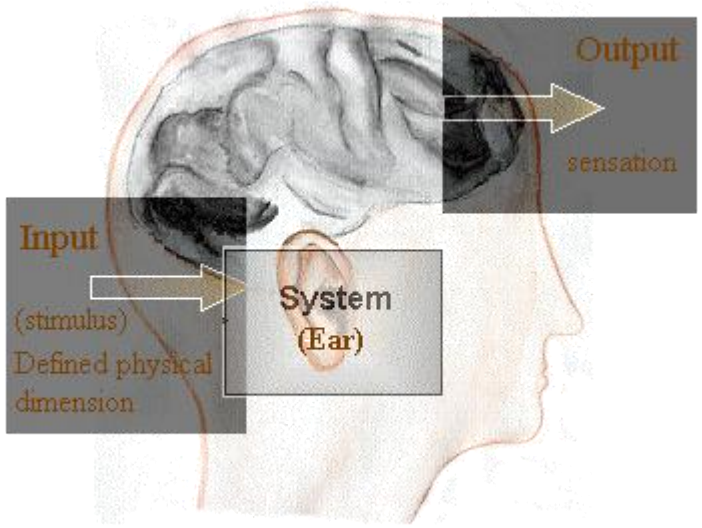
Realistic	Meeting traditional sound expectations	New sounds starting from scratch
Preservation of electric drive sound	Generation of sounds comparable to combustion engines.	Creation of new sounds suitable for electric cars.

Different background noises, different feelings

Let Your Ears Decide!

Why Psychoacoustics?

- Objective measurement → subjective impression
- Relationship between measurement properties and resulting feelings
- Needed: Principle of operation and transfer function of the human hearing



HEAD acoustics philosophy

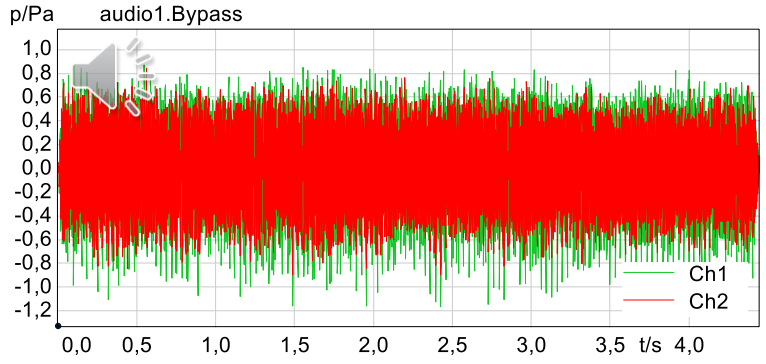
Let Your Ears Decide

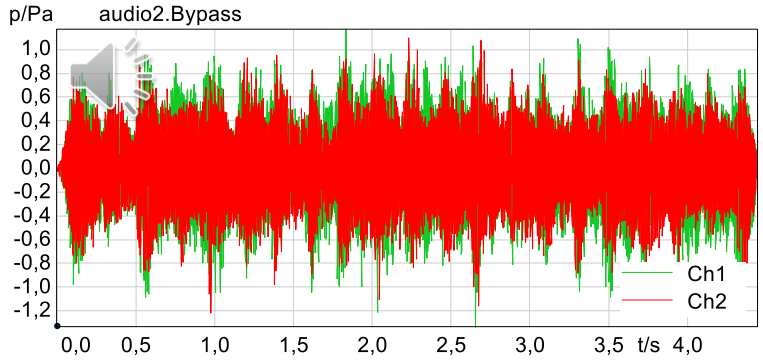
- Human beings perceive sound with two ears in a sound field => **binaurally**
- The holistic perception of a sound event always depends on the **context/ *soundscape***
- Analysis data alone does not cover all aspects of **human perception and acoustical quality**
- Nevertheless analyses with ***high correlation between*** **perception and mathematics** are required

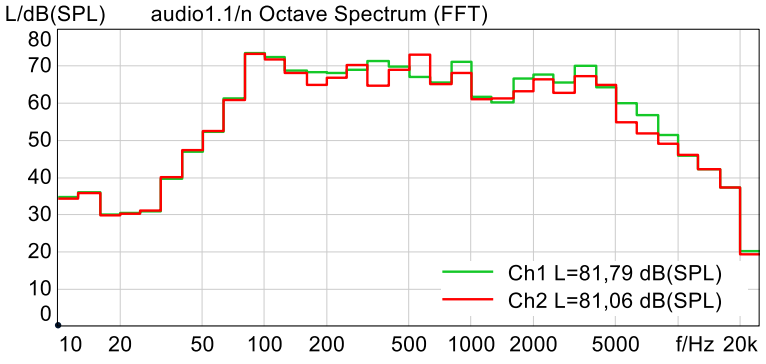
if you are dealing with acoustic applications, it is unavoidable to be able to reproduce and listen to the (binaural) time data as well as the identification and objective quantization of any kind of sound patterns.

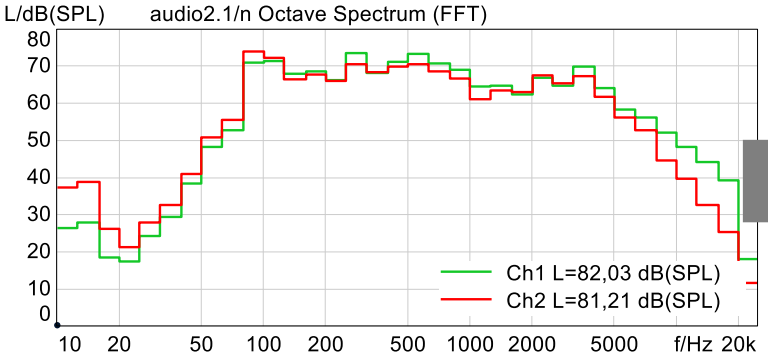
Psychoacoustics

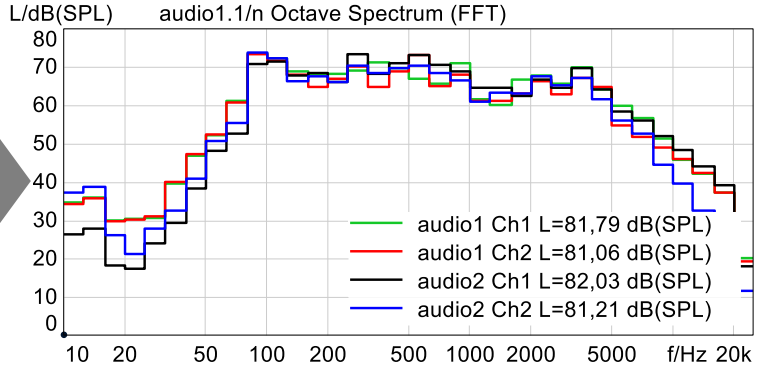
- Sound Pressure Level alone does not reflect the properties of the human hearing and the human perception of sound





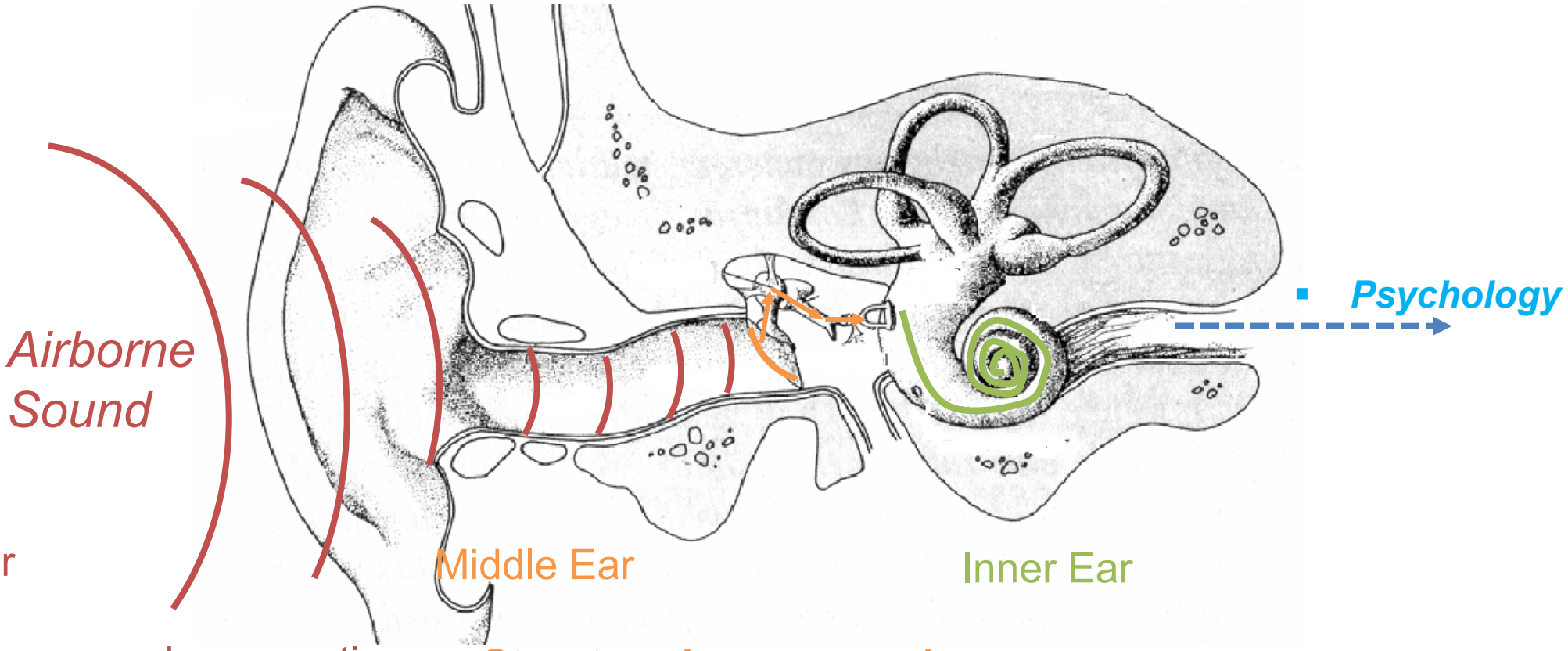






Similar spectrum and overall level!

What is Psychoacoustic?



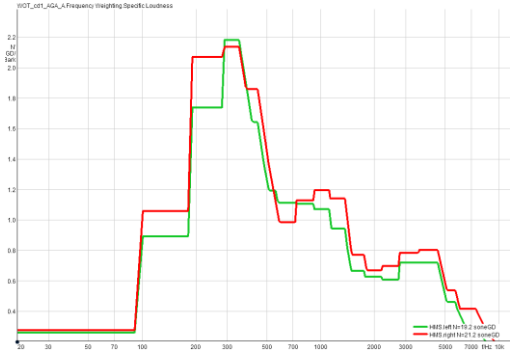
- Airborne sound propagation
- Distortion of the sound field
- Cavum Conchae and Ear Canal Resonances
- ***Physical filtering***

- **Structure-borne sound propagation**
- **Impedance matching**
- Amplification (leverage)
- Protection

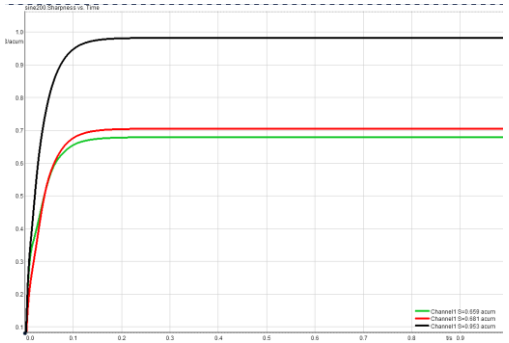
- Frequency-space transformation
- **Divides audible frequency range into 24 bands (critical bands)**
- Nerve fibers (hair cells) transmit motion into nervous signals
- **Psychoacoustics**

What is Psychoacoustic/ Parameters

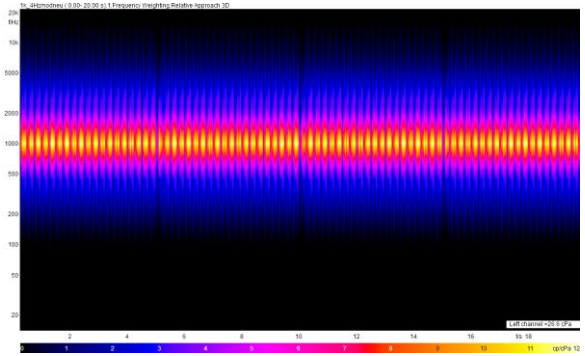
dB ?



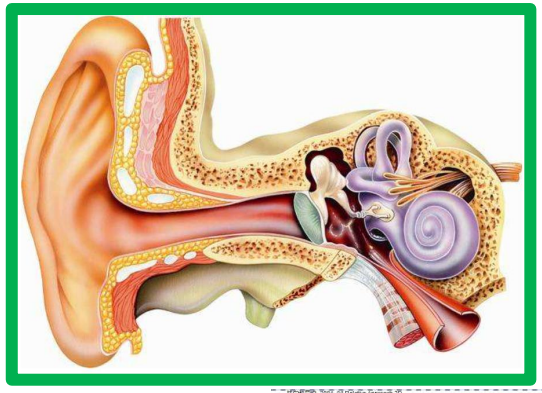
Specific Loudness

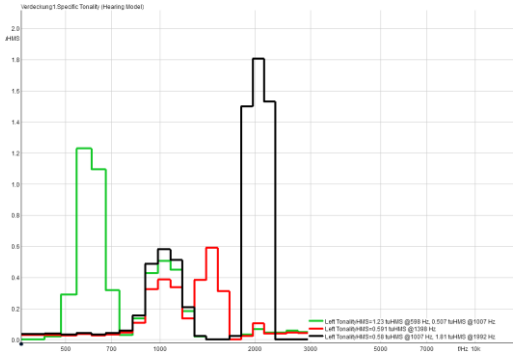


Sharpness vs. time

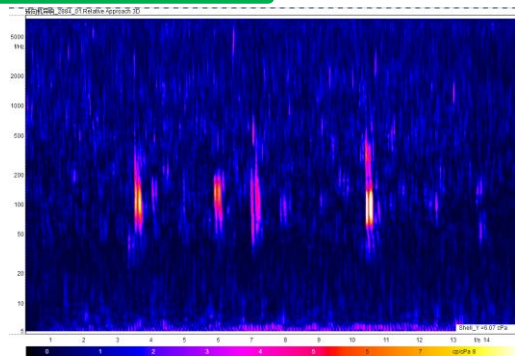


Fluctuation Strength

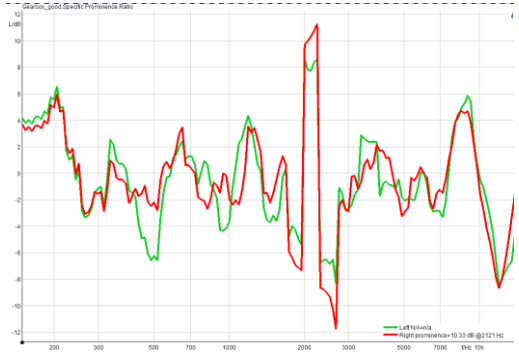




Tonality(Hearing model)



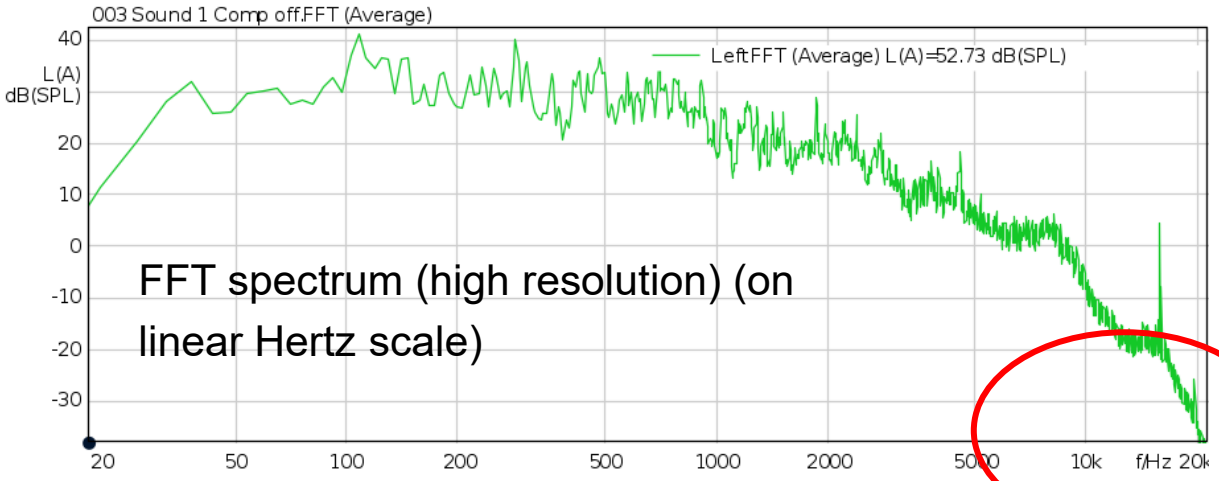
Relative Approach

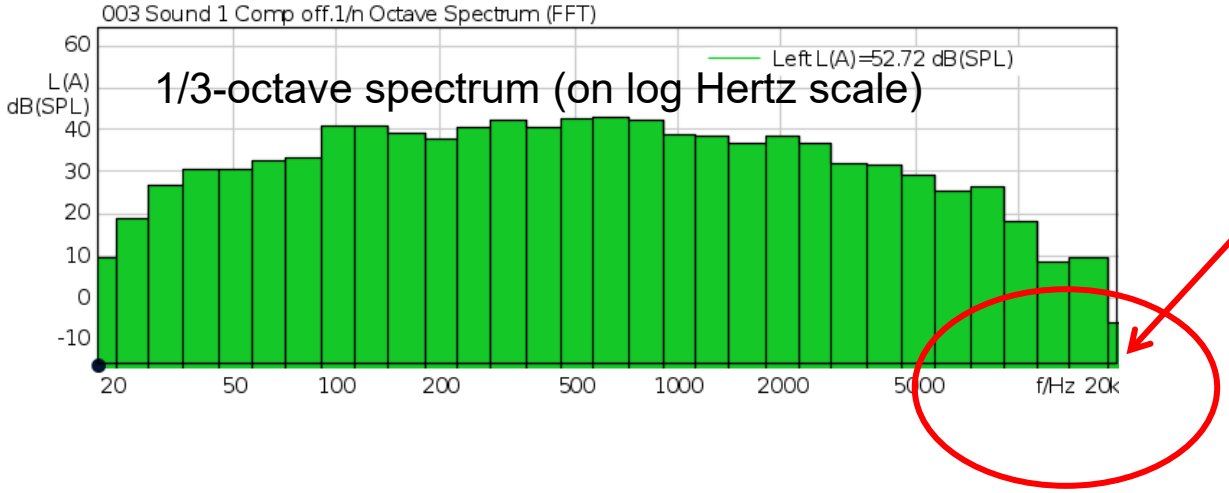


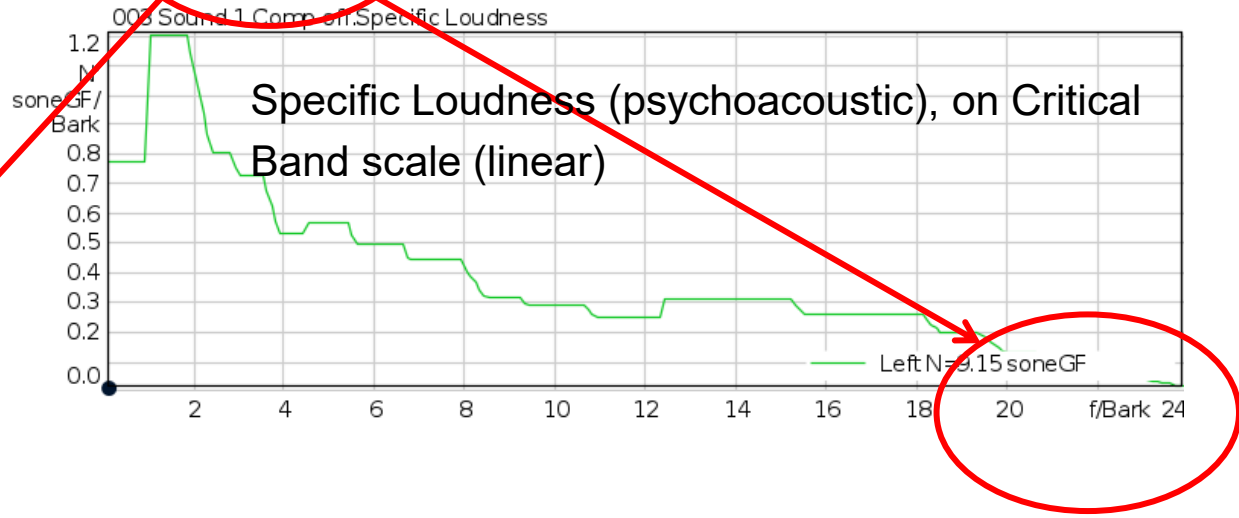
Prominence Ratio

• • •

Psychoacoustics _ What's a Critical Band?



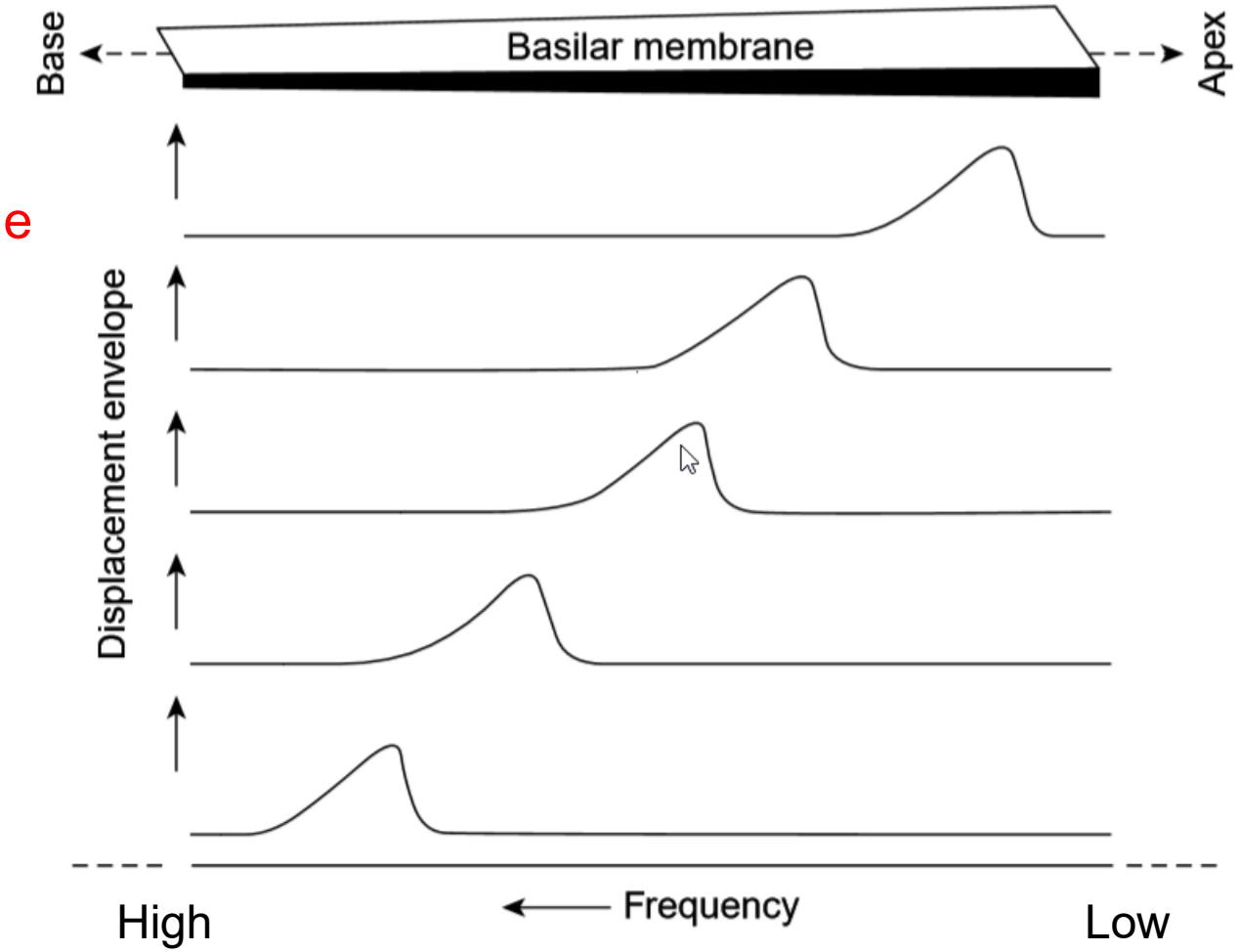




A critical bandwidth corresponds to the excitation shape on the basilar membrane of the ear's cochlea, from a pure tone or narrowband noise.

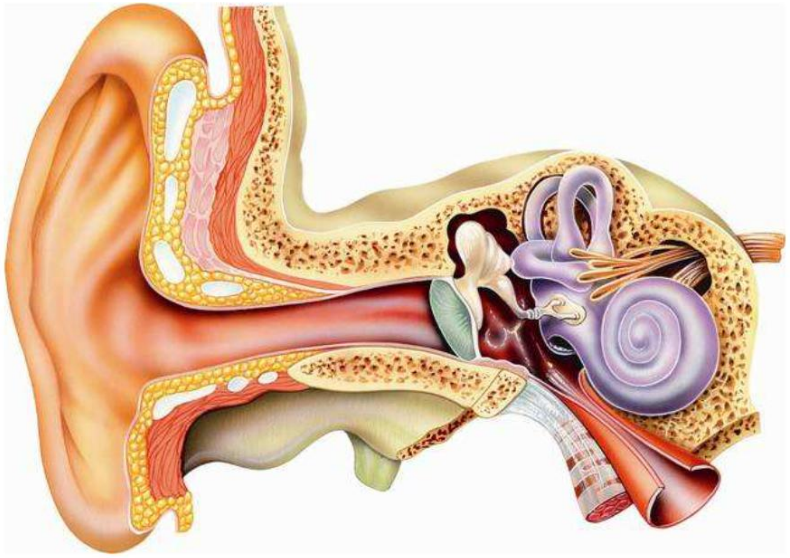
Psychoacoustics _ Critical Bands

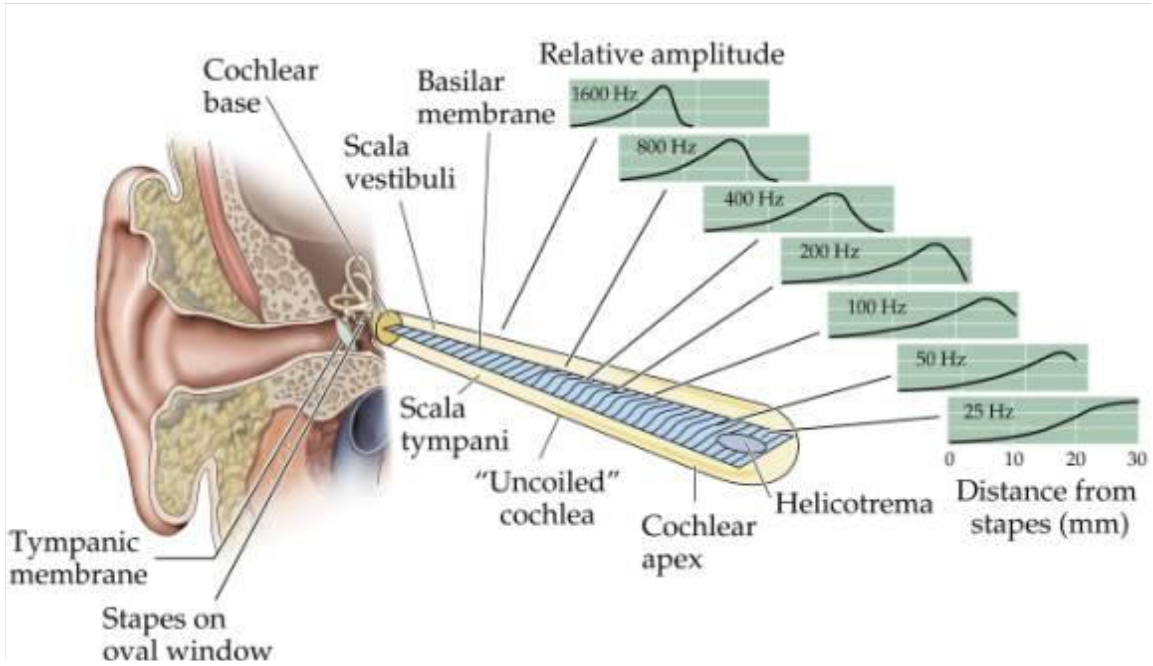
- Critical bands are phenomena created by the cochlea.
- This shape governs the ability to distinguish frequencies in the ear.
- Different parts of the cochlea respond to different frequencies.



Psychoacoustics _ Critical Bands

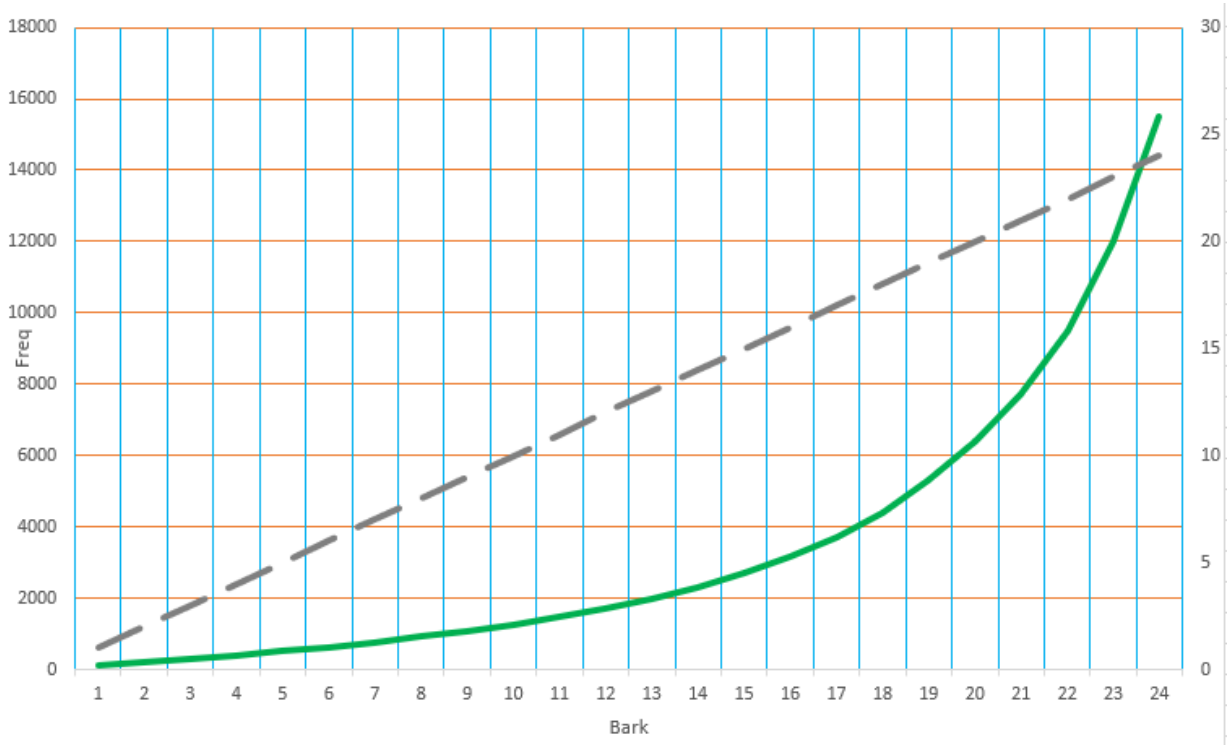
- nonlinear sound transmission
- sound perception occurs in frequency (“critical”) bands
- **sounds must “build-up”**





Psychoacoustics _ Critical Bands

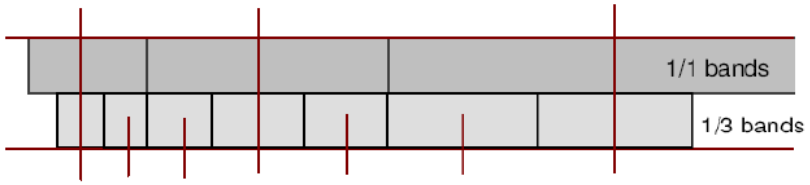
- Definition
  - Auditory Tests suggests neither a linear nor a logarithmic behavior of the hearing
  - The auditory range can be divided into 24 frequency bands
  - The width of one band grows with rising frequency



Psychoacoustics _ Critical Bands

$$f_{i+1} \,/\, f_i = 2\text{^}1/N \;\; (N = 1,2,3 \ldots)$$

- One can also divide an octave further into three equal bands:



- Band center and limiting frequencies(ICE):

$$f_{c,i} = 1000 \times \left(2^{1/N}\right)^{i} \;\; with\ i = 0, \pm 1, \pm 2 \ldots$$

$$f_{lower} = (2^{1/3})^{-1/2} f_c \quad (N = 3)$$

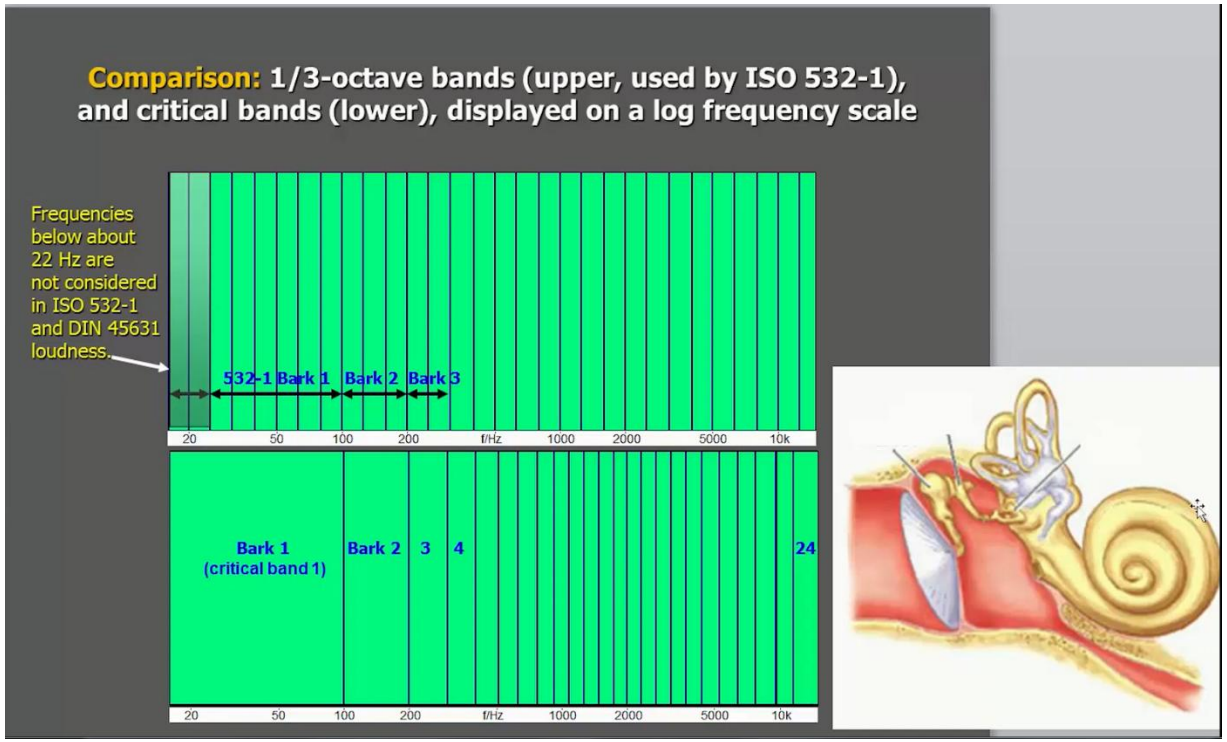
$$f_{upper} = (2^{1/3})^{1/2} f_c$$

ISO 150266(GB3240)/ Octave

Normalized frequency	1/1 oct	1/2 oct	1/3 oct		1/1 oct	1/2 oct	1/3 oct		1/1 oct	1/2 oct	1/3 oct
16	x	x	x	160			x	1600			x
18				180		x		1800			
20			x	200			x	2000	x	x	x
22.4		x		224				2240			
25			x	250	x	x	x	2250			x
28				280				2800		x	
31.5	x	x	x	315			x	3150			x
35.5				355		x		3550			
40			x	400			x	4000	x	x	x
45		x		450				4500			
50			x	500	x	x	x	5000			x
56				560				5600		x	
63	x	x	x	630			x	6300			x
71				710		x		7100			
80			x	800			x	8000	x	x	x
90		x		900				9000			
100			x	1000	x	x	x	10000			x
112				1120				11200		x	
125	x	x	x	1250			x	12500			x
140				1400		x		14000			
160			x	1600			x	16000	x	x	x

Table 2.1 Normalized frequencies (Hz)

Psychoacoustics _ Critical Bands



z [Bark]	f [Hz]	Δf [Hz]	z [Bark]	f [Hz]	Δf [Hz]	z [Bark]	f [Hz]	Δf [Hz]
0	0		**8**	920	150	**16**	3150	450
1	100	100	**9**	1080	160	**17**	3700	550
2	200	100	**10**	1270	190	**18**	4400	700
3	300	100	**11**	1480	210	**19**	5300	900
4	400	100	**12**	1720	240	**20**	6400	1100
5	510	110	**13**	2000	280	**21**	7700	1300
6	630	120	**14**	2320	320	**22**	9500	1800
7	770	140	**15**	2700	380	**23**	12000	2500
						24	15500	3500

For the conversion from hertz to Bark, the following approximation formula is available:

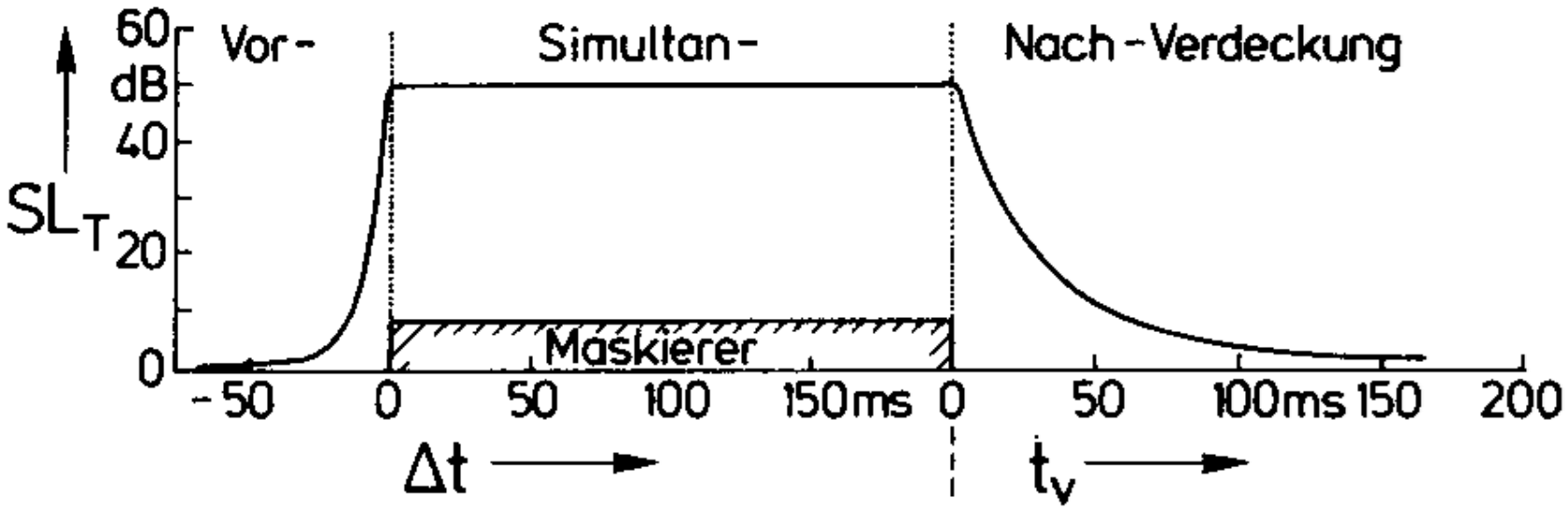
$$critical\ band\ rate\ [Bark] = (13 \cdot \tan^{-1}(0.00076 \cdot f)) + 3.5 \cdot \tan^{-1}((0.00013333 \cdot f)^2) \quad , \ with\ f = frequency\ [Hz]$$

Reciprocally, you can calculate the approximated frequency from a Bark value using the following formula:

$$frequency\ [Hz] = (2.841 \cdot e^{0.219 \cdot z} + 100) \cdot z - 32 \cdot e^{-0.15 \cdot (5-z)^2} \quad , \ with\ z = critical\ band\ rate\ [Bark]$$

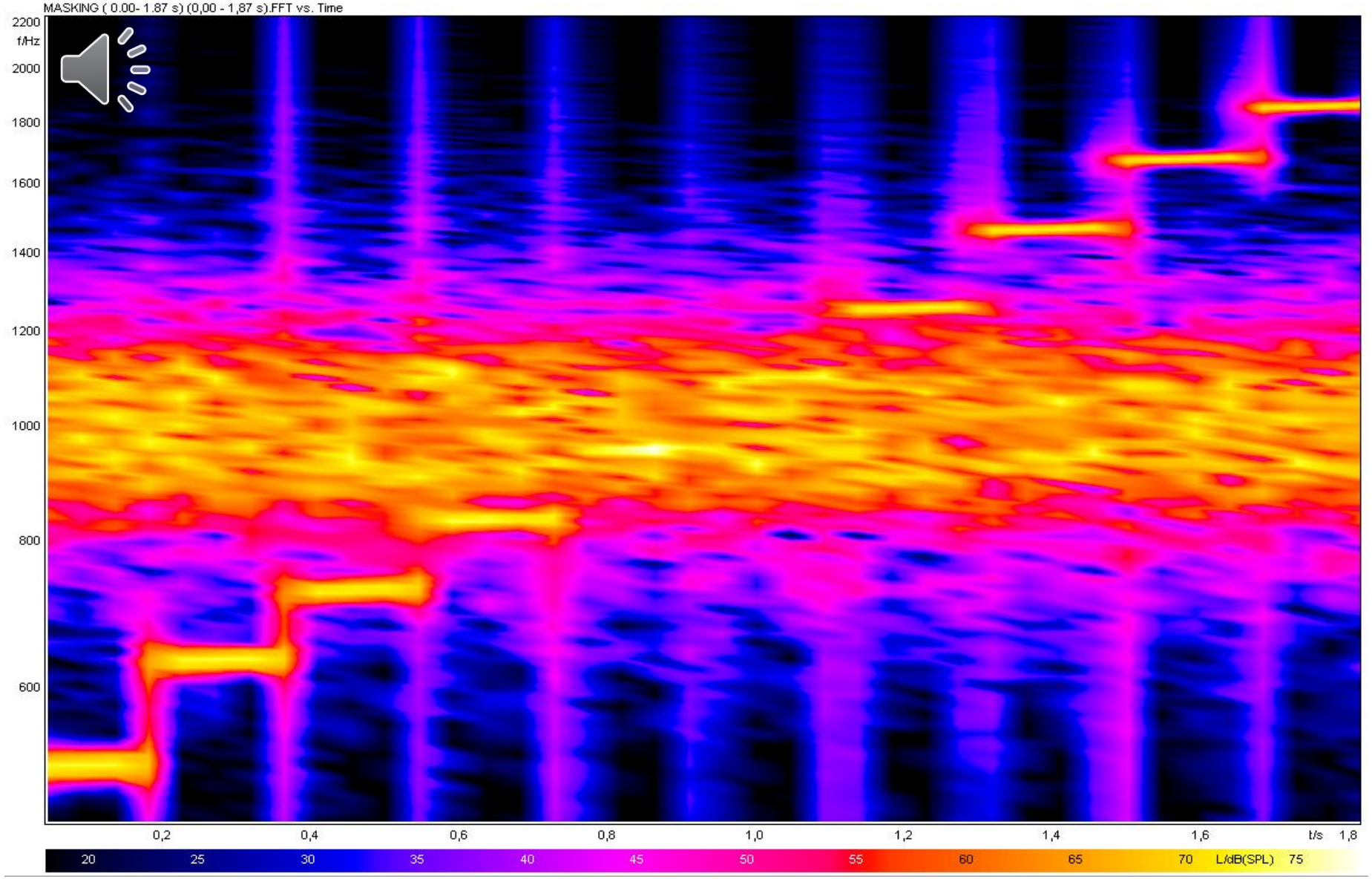
Pre- and post masking_ Temporal

- **Temporal** Masking can affect signals prior to or after the masking signal
- The longer the masking signal, the longer the effect 76 it ends

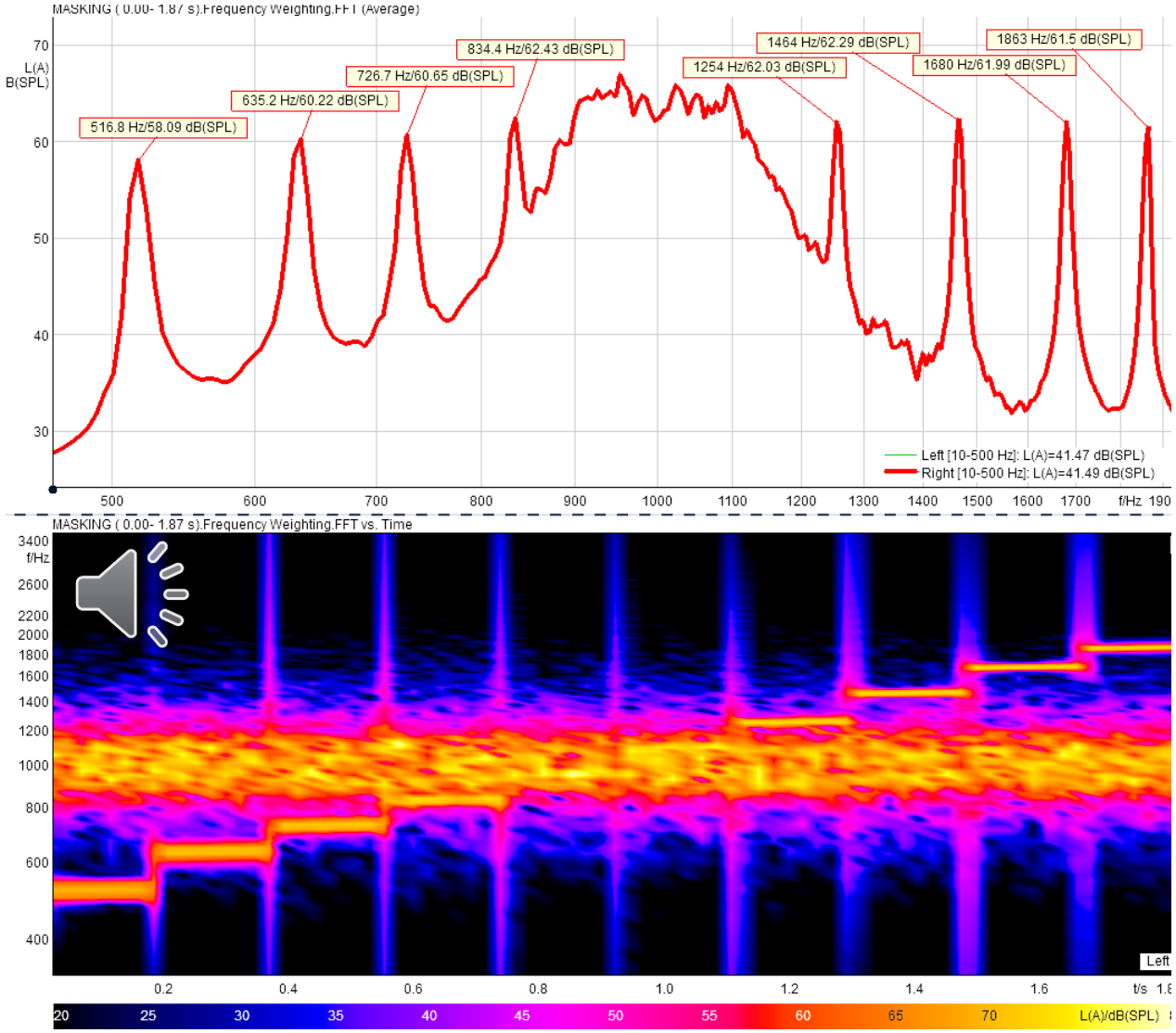


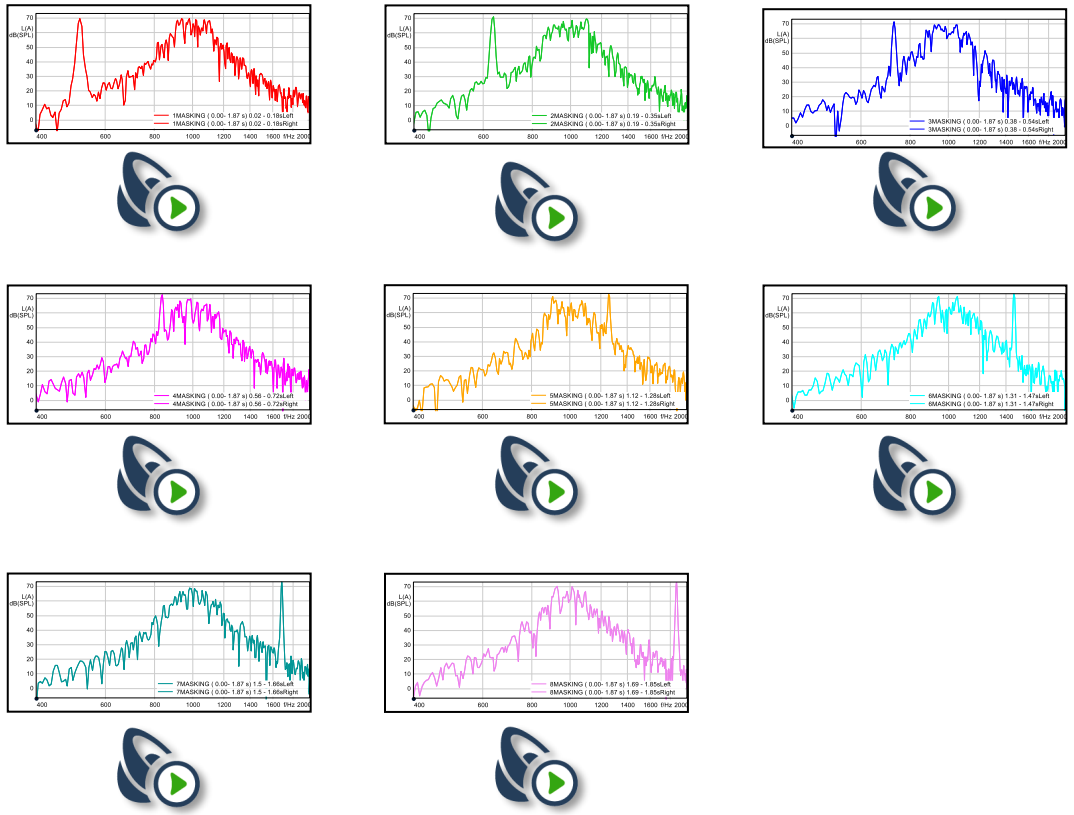
From: Zwicker, Psychoakustik

Psychoacoustics _Masking

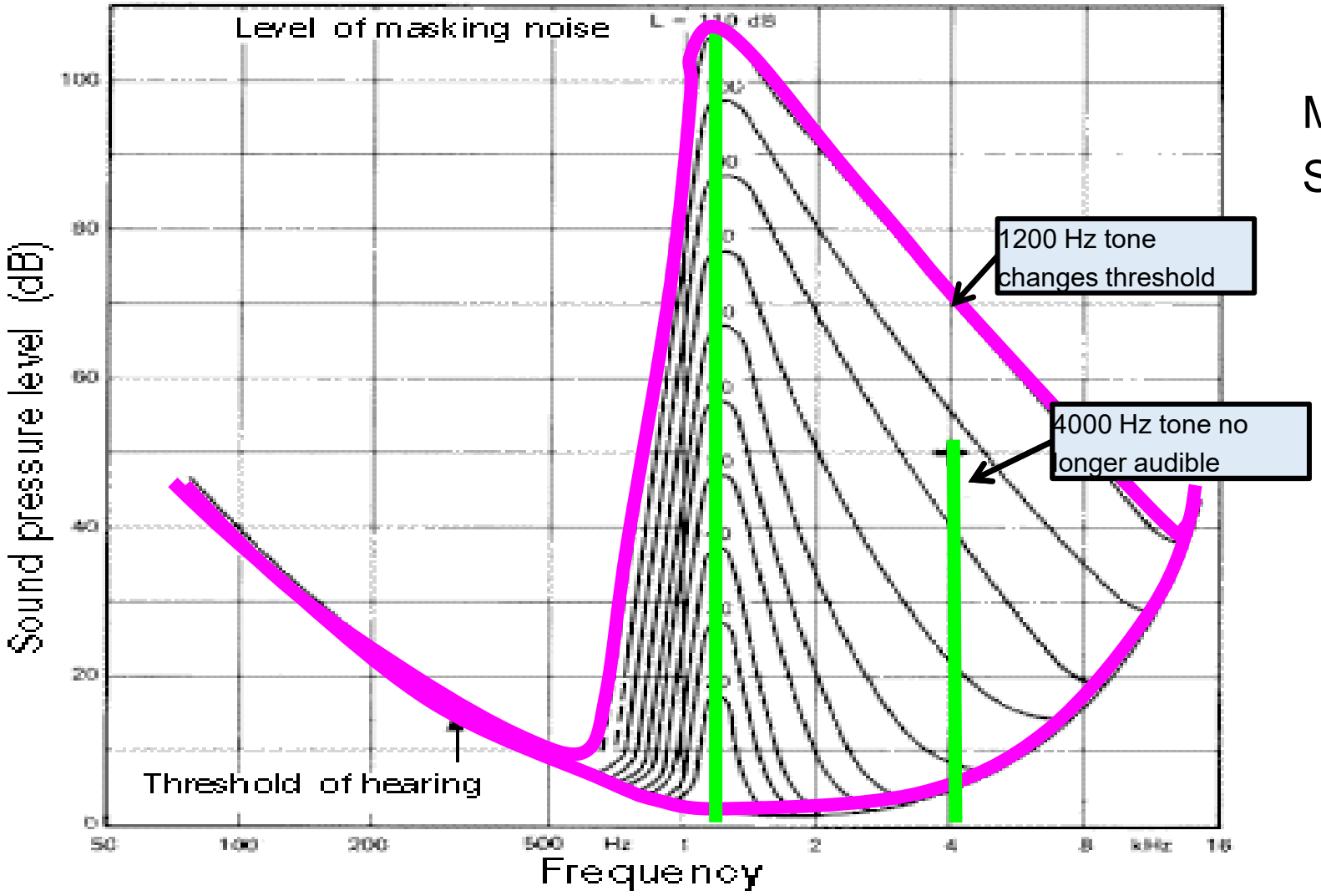


Psychoacoustics _Masking



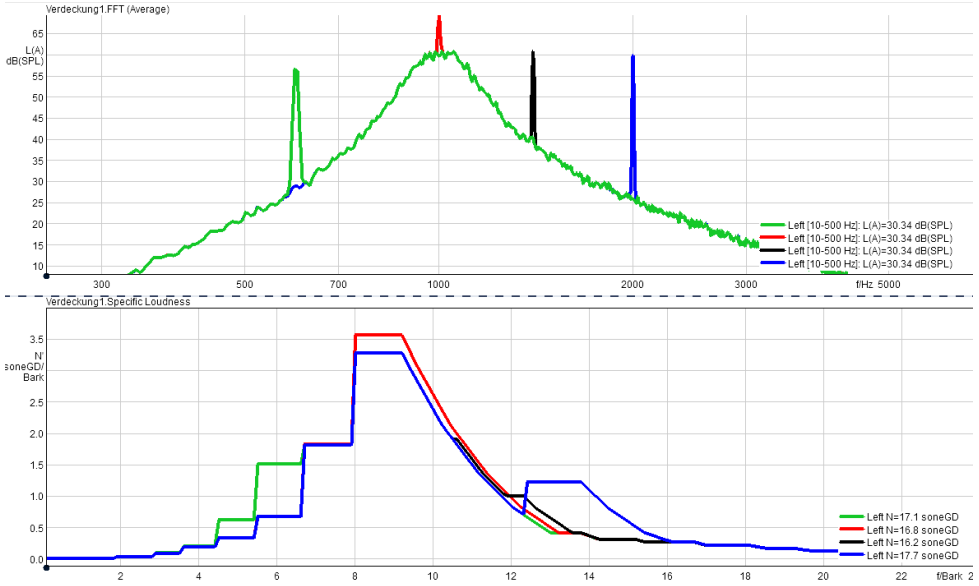


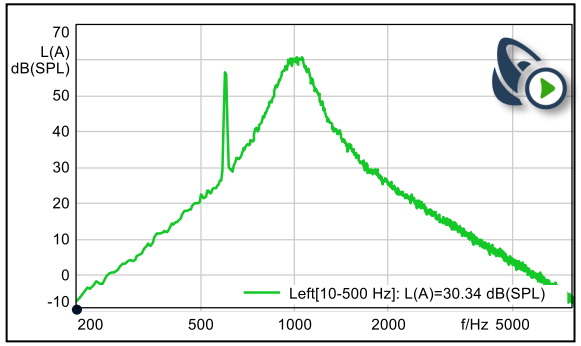
Simultaneous masking_ Frequency

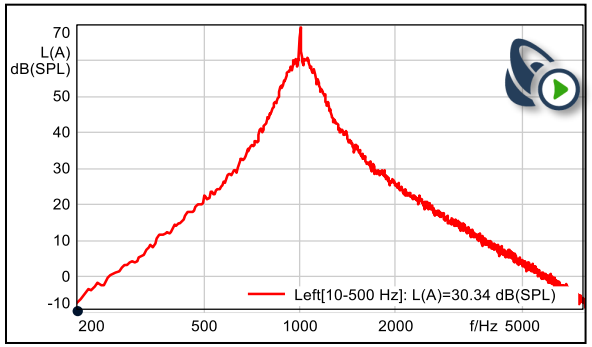


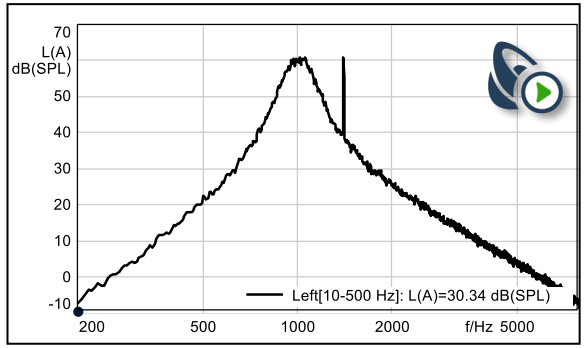
Masking for 1200 Hz Sine Tone

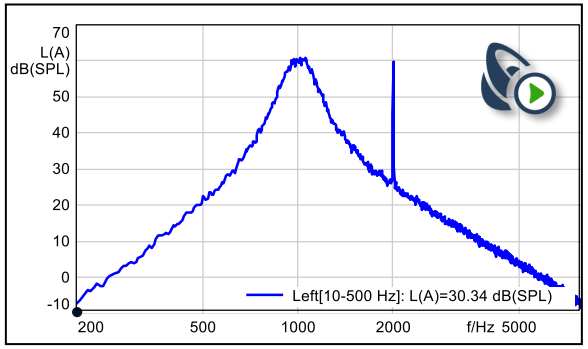
Psychoacoustics _ Masking





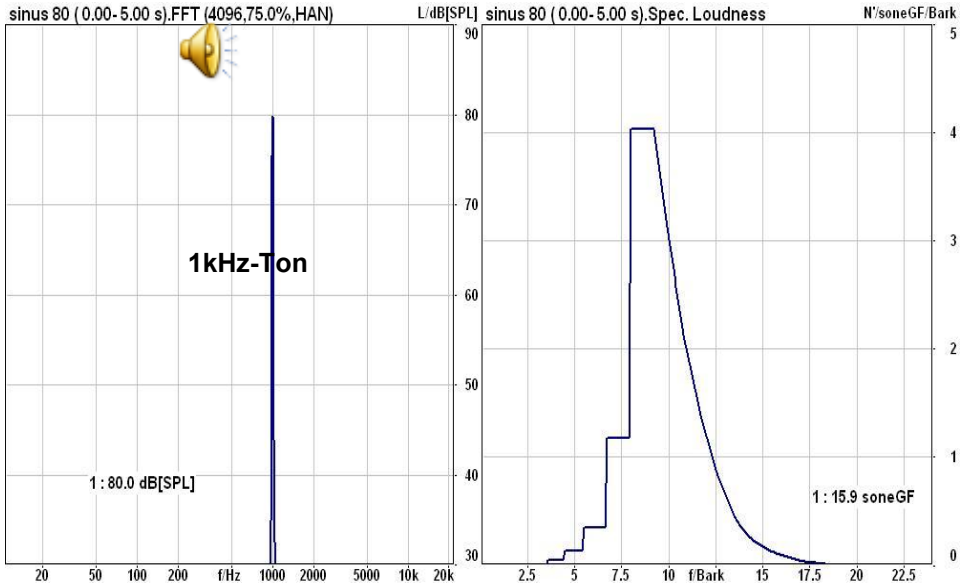


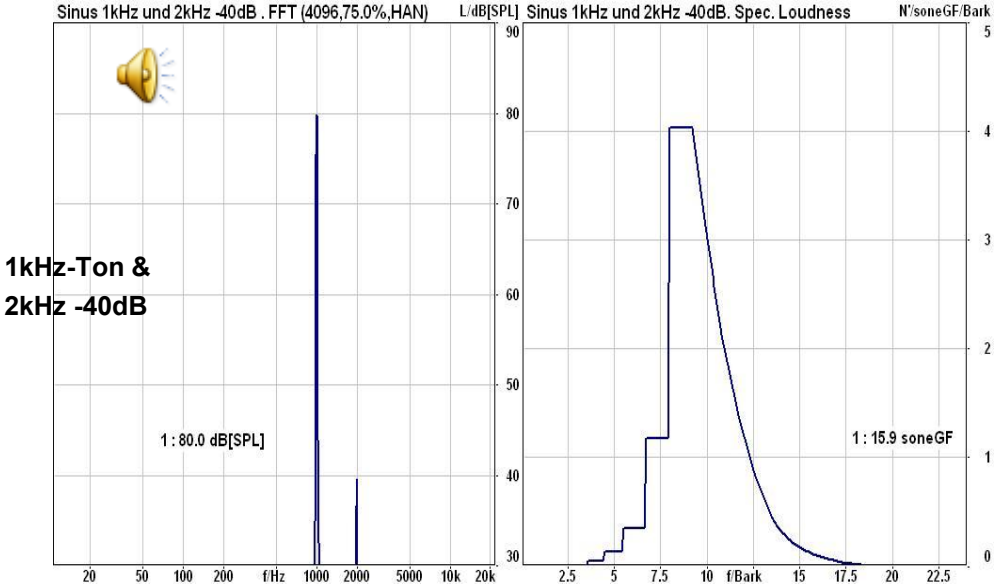


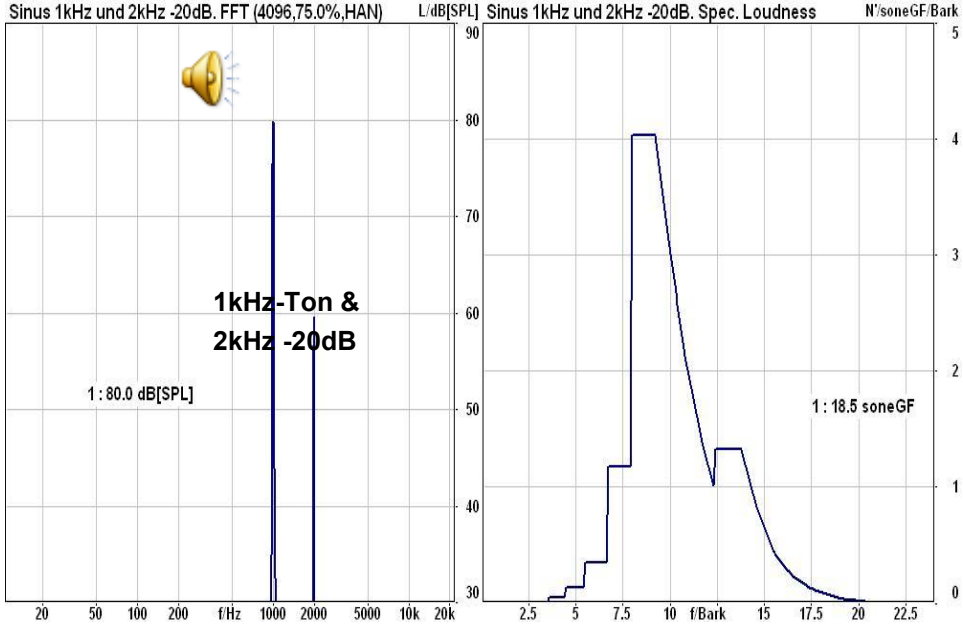


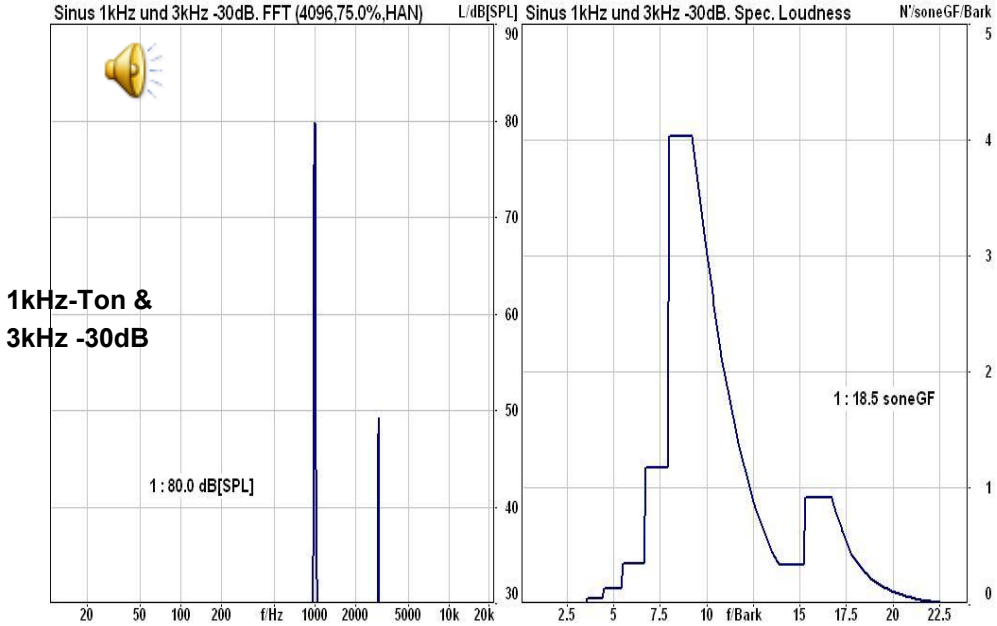
- How masking will affect psychoacoustics with frequency and time => **Broadband/ narrowband**
- The relationship between masking and human hearing with two ears in a sound field => **<u>binaurally</u>**

Simultaneous masking

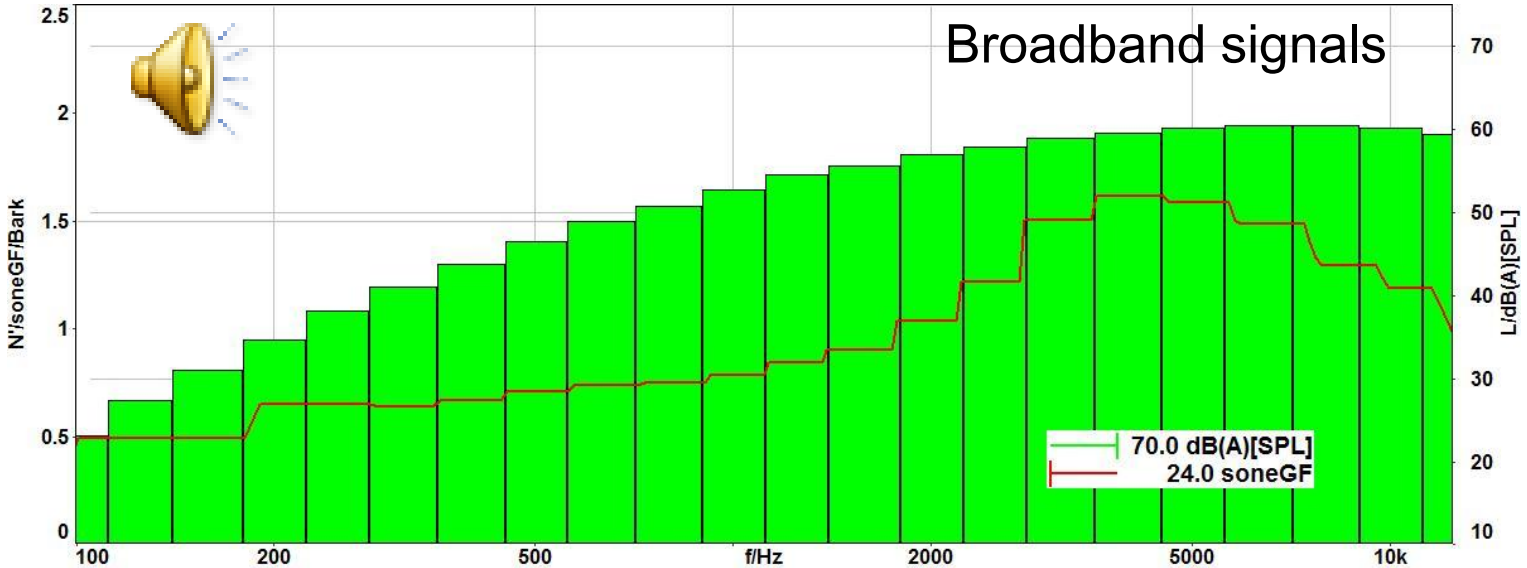


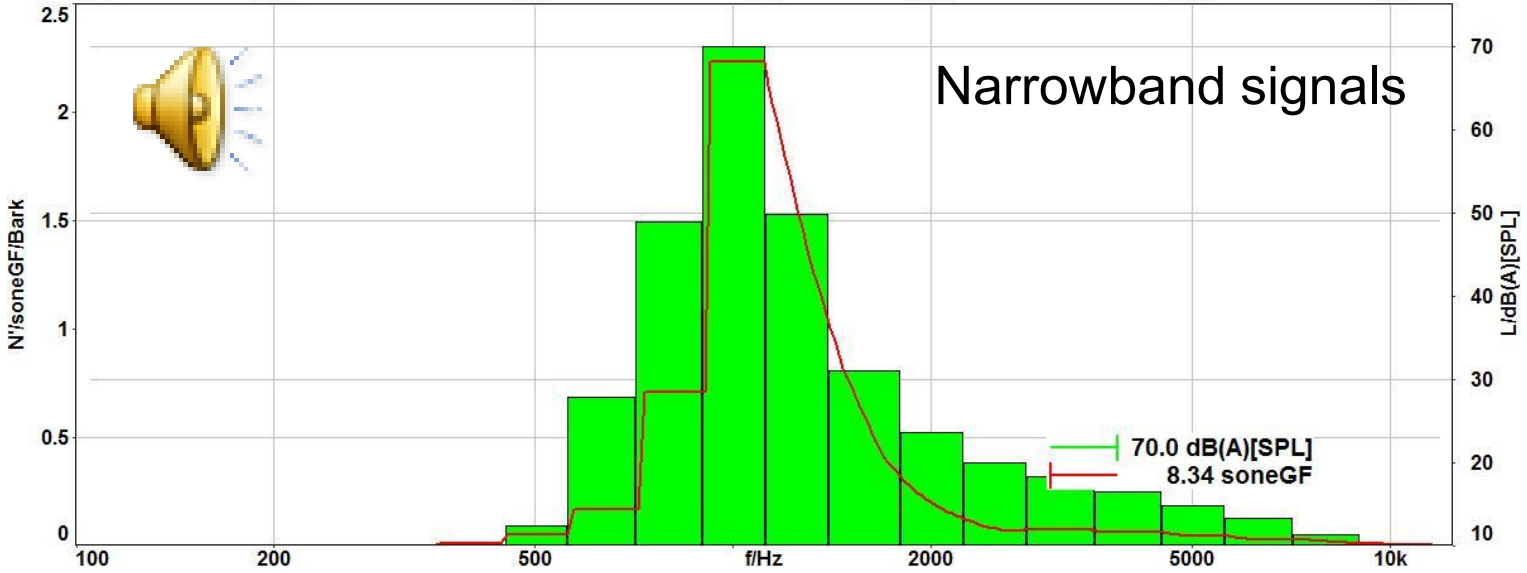






Psychoacoustics _ Masking

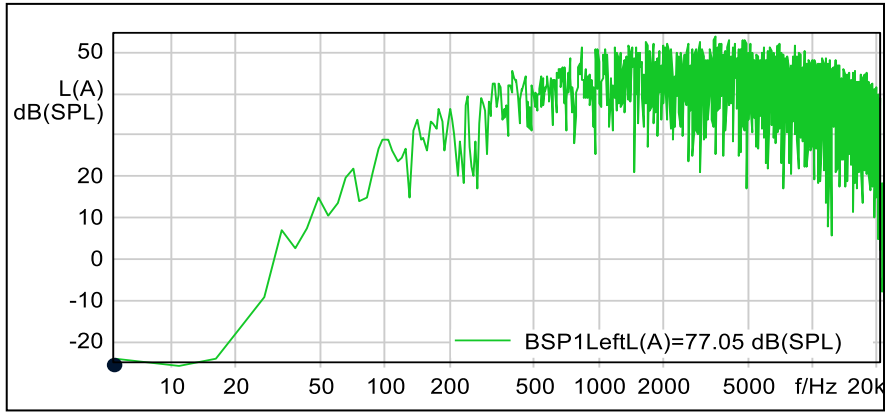


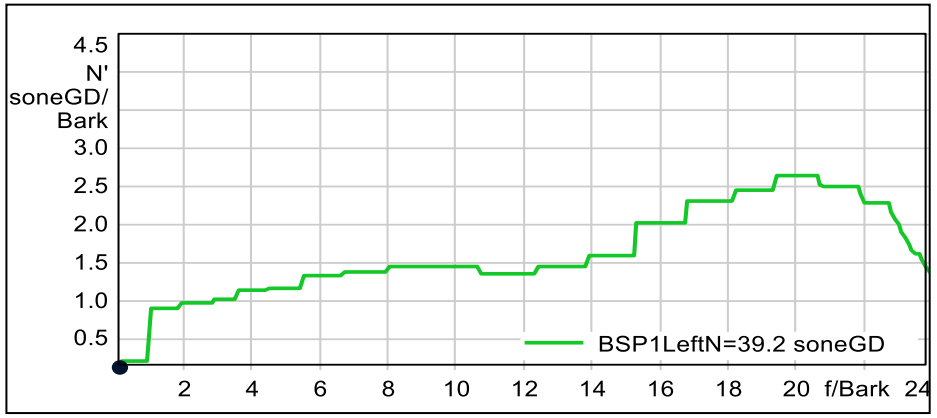


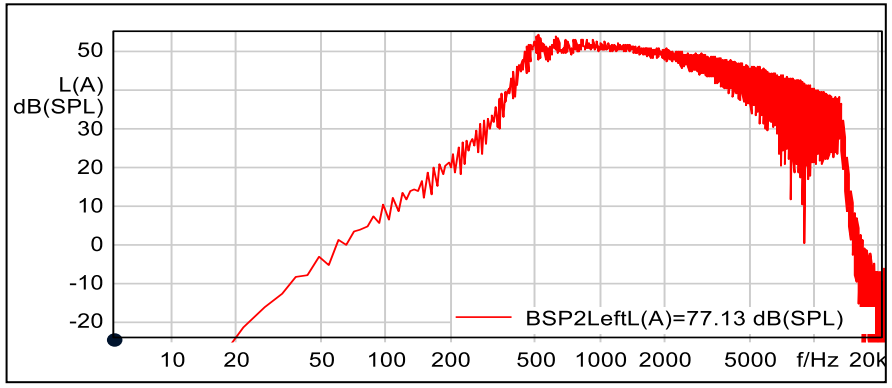
24Bark/ 200Hz? 2kHz? 4kHz?

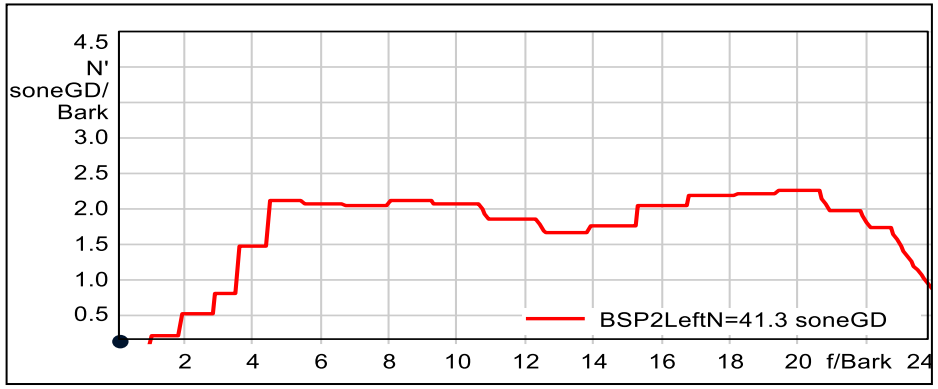
Signals	SPL	Loudness
Broadband	70dB	24 soneGF
Narrowband	70dB	8.34 soneGF

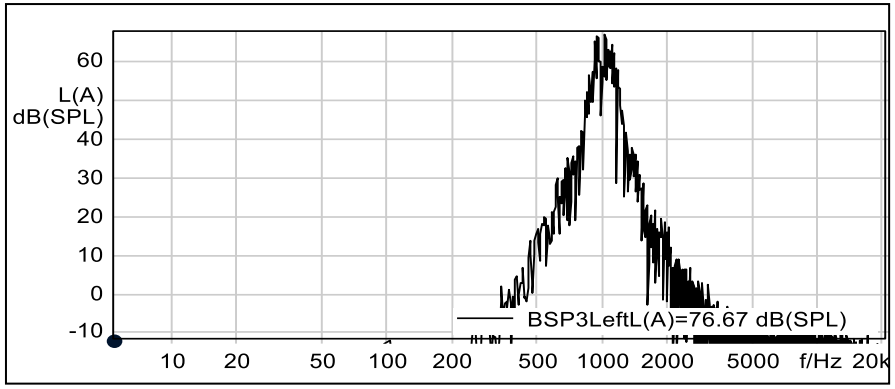
Psychoacoustics _ Loudness

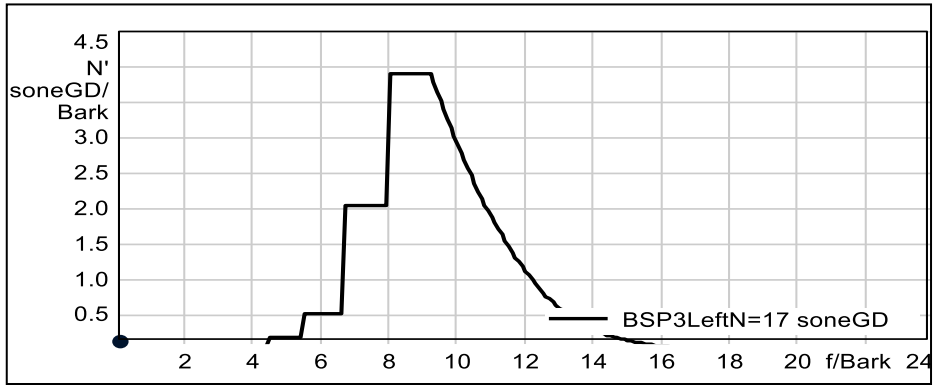










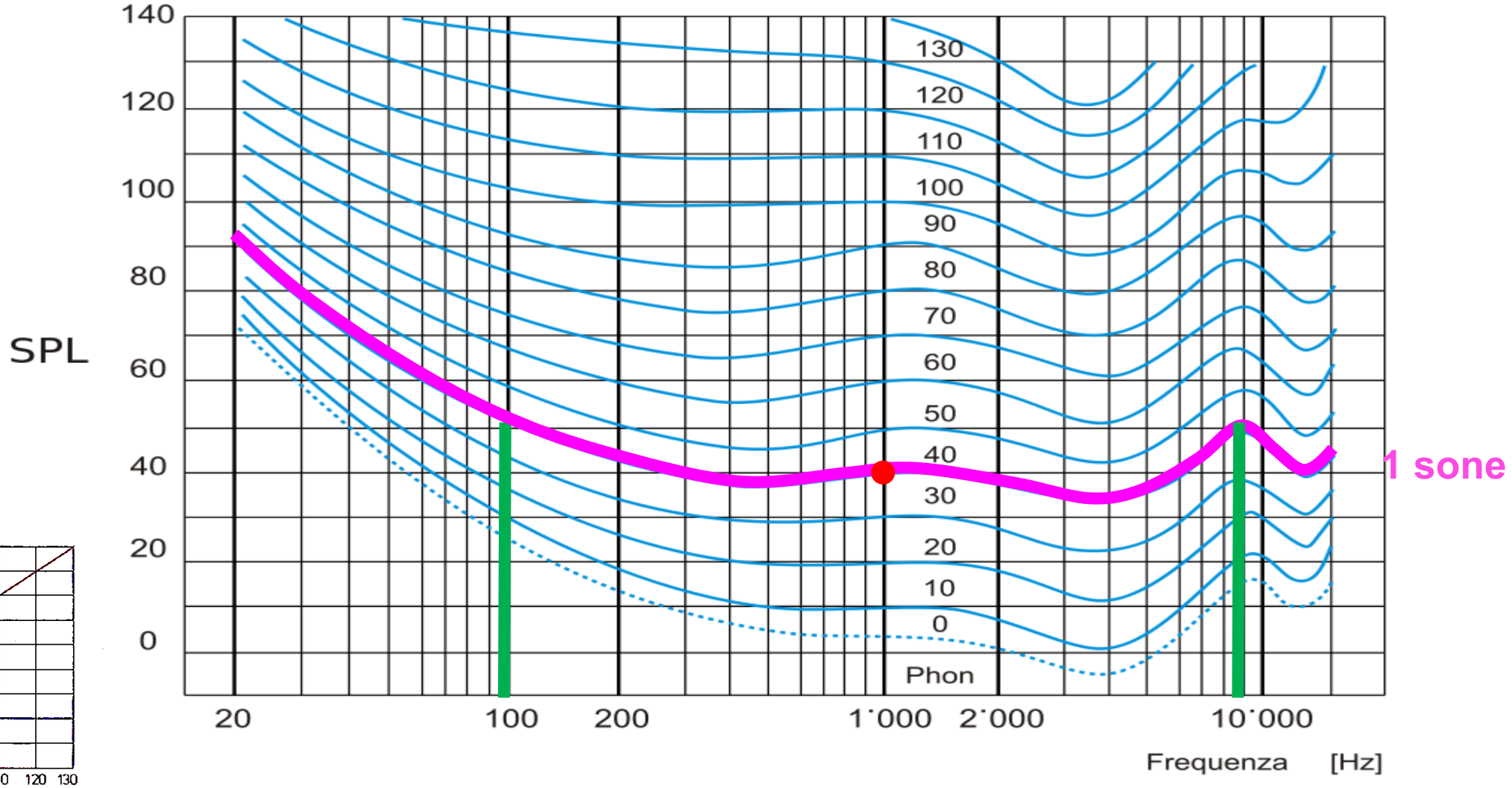


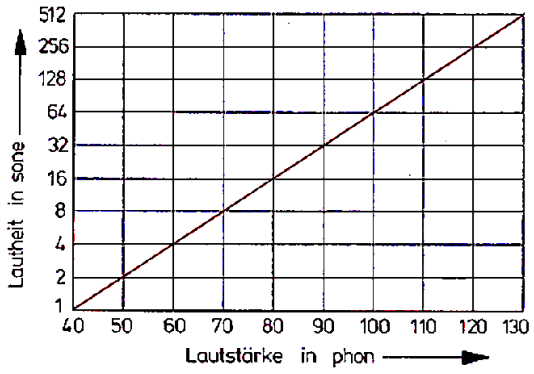
Psychoacoustics _ Loudness

- Definition
- A pure sine tone with the frequency 1 kHz and a sound pressure level of $40\ dB_{SPL}$ exhibits the loudness 1 sone
- Unit to indicate the perceived Volume
- Unit: sone or sone G(F,D) for overall loudness (F: measured in free field condition, D: measured in diffuse field condition)
- Loudness is meant to indicate the sensation linearly

Psychoacoustics _ Loudness

Equal-loudness contours(red)(from ISO 226:2003 revision) Original ISO standard





Psychoacoustics _ Loudness

标准类别	年份 / 算法	1967	1975	1980	1991	2005	2007	2010	2017
DIN 德国标准	Zwicker 稳态	*DIN 45631(1967)*	手绘变机绘		DIN 45631(1991)		扩展了瞬态算法	DIN 45631/A1 (2010)	
	Zwicker 时变								
	Stevens								
	Glasberg & Moore								
ISO 国际标准	Zwicker 稳态		ISO 532 B 1975						
	Zwicker 时变								ISO 532-1
	Stevens		ISO 532 A 1975						
	Glasberg & Moore								ISO 532-2
ANSI 美国标准	Zwicker 稳态								
	Zwicker 时变								
	Stevens			ANSI S3.4-1980					
	Glasberg & Moore					ANSI S3.4-2005	ANSI S3.4-2007		

*标红处为DIN、ISO、ANSI的最新算法，已加入ArtemiS SUITE中。

Psychoacoustics _ Loudness

- ISO532-1_Stationary sounds/***Moore Model***

$$L_N = \begin{cases} 40 + 33.22 \cdot \lg\left(\frac{N}{sone}\right) & \text{for } N \geq 1\,sone \\ 40 \cdot \left(\frac{N}{sone} + 0.0005\right)^{0.35} & \text{for } N < 1\,sone \end{cases}$$

$$L = \int_0^{24Bark} N'(z)dz$$

Time Signal → 3rd Octave Filtering → Critical Bands 24 Bark → Corrections Ear Trans F/D → Integration 24 Bark → Specific loudness

- DIN45631/1967~2010 ANSI S3.4/1980~2007 → 3rd Octave Filtering
- ANSIS3.4/1980~2007_*FFT* /Moore/Glasberg → 3rd Octave Filtering
- DIN45631/1967~2010 ISO 532-1 → Masking → Integration 24 Bark

3rd Octave Filtering
- Time Domain
- Frequency/FFT
- 25~12.5KHz

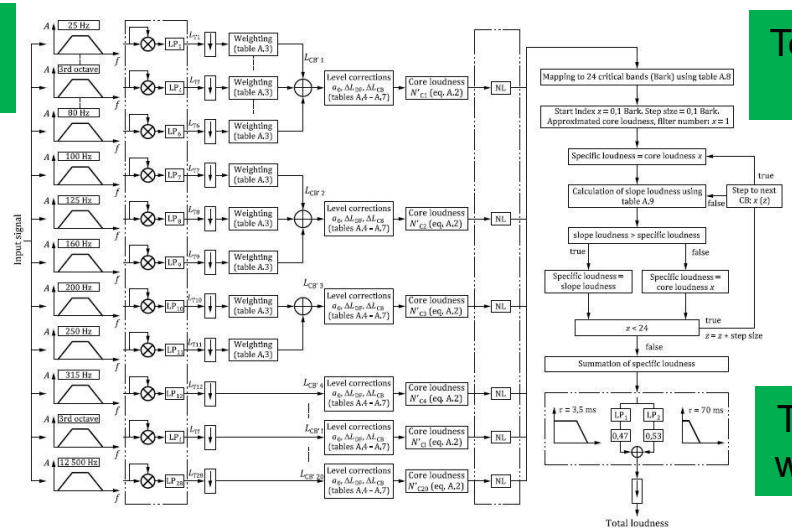
Critical Bands 24 Bark
- >300Hz 3rd Octave
- <300Hz 1/2/3 Bark

Corrections Ear Trans F/D
- Ear Trans
- Free/ Diffuse
- 4K Hz

Integration 24 Bark
- Single value Integration
- Free/ Diffuse
- 4K Hz

- ISO532-1_Time varying

Smoothing & Down-Sampling



Temporal Decay

Temporal weighting

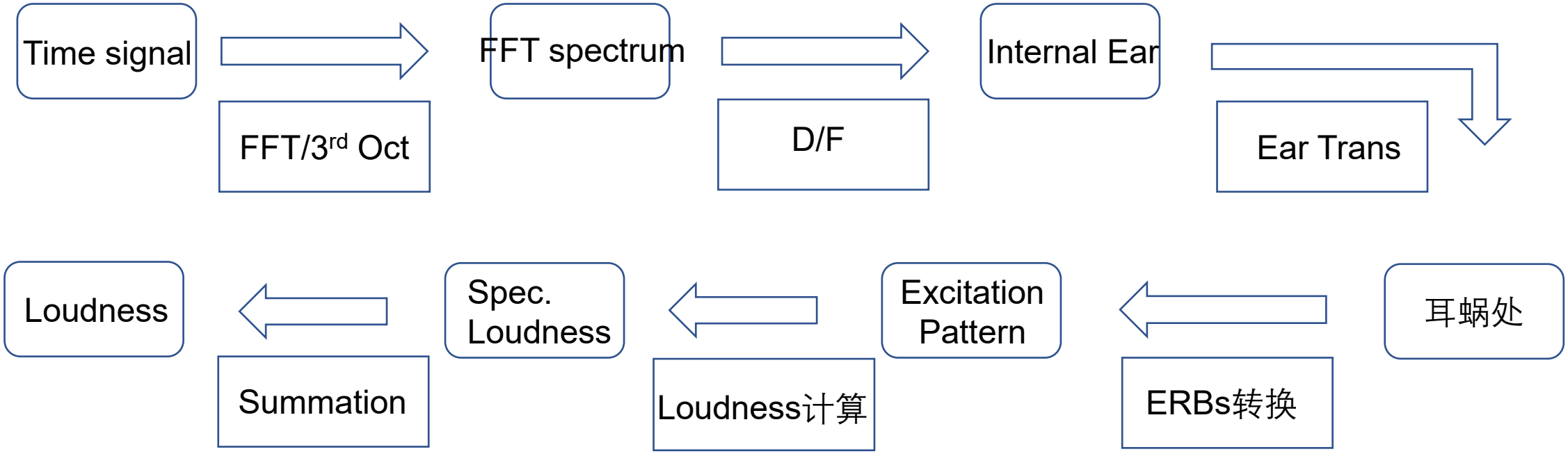
Developing

- 1st Octave
- 3rd Octave
- Freq Domain
- Time Domain
- Pure tone
- Multi-tone
- Single ear
- Double ears

…

Psychoacoustics _ Loudness

- ANSI S3.4-2007



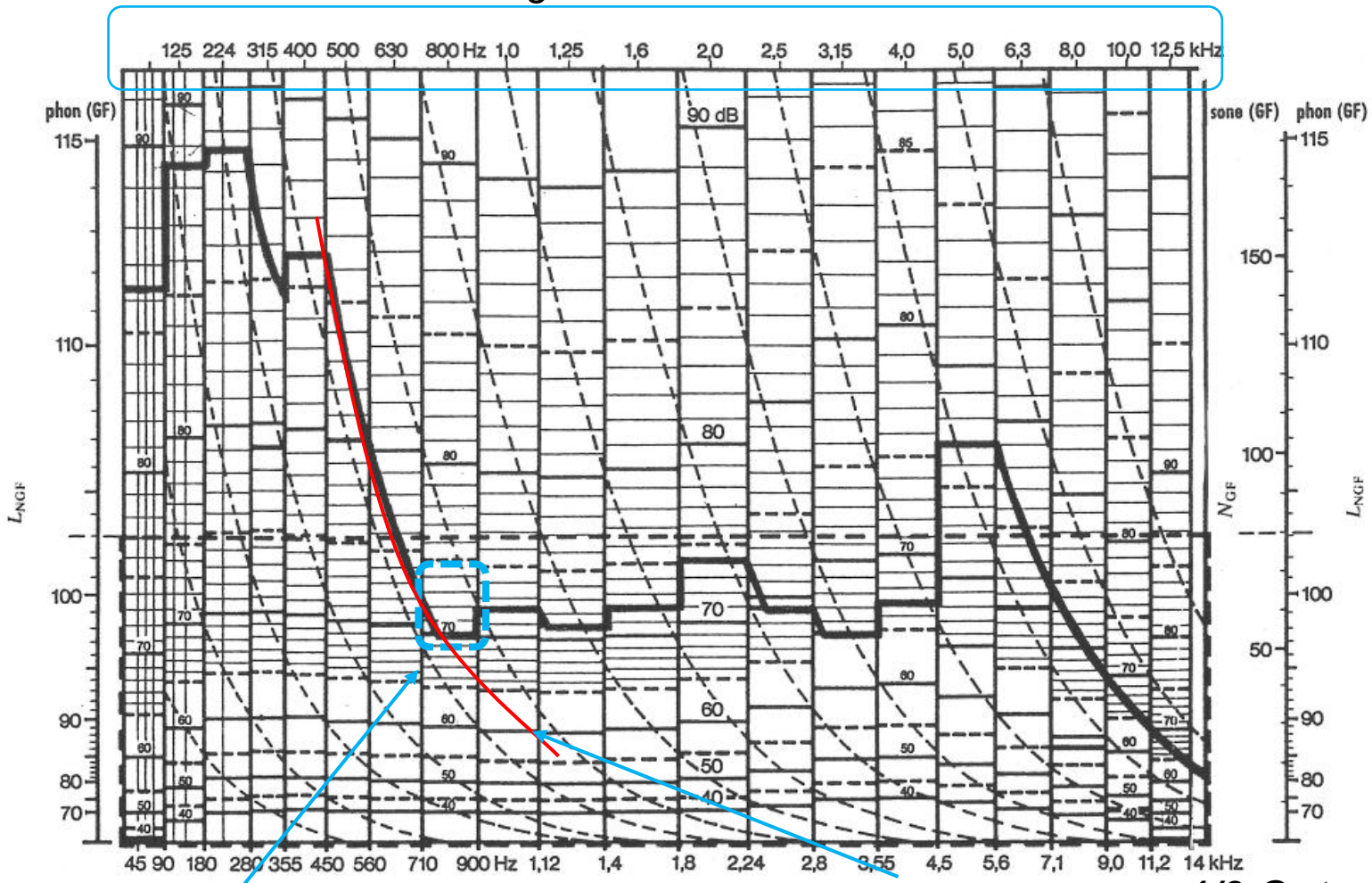
Psychoacoustics _ Loudness

- Calculation
- Hearing related measuring principle which includes masking effects
- DIN 45631-A1 ISO 532 B (for stationary Noise)
- ISO 532 A (draft)/ ANSI S3.4-2007 (FFT-Analysis)
- ISO 532 A (draft)/ ANSI S3.4-2007 (3rd oct. FFT-Analysis)

Psychoacoustics _ Loudness

- Graphical procedure/***Zwicker Model DIN45631***
- Core and slop loudness are filled into the diagram
- Overall loudness via integration of area beneath the curve

a,1/3 Octave(f_c)→ SPL



b,$f_c = 800 \rightarrow SPL = 70$

c, Masking_ compare

1/3 Octave($f_{up/low}$)

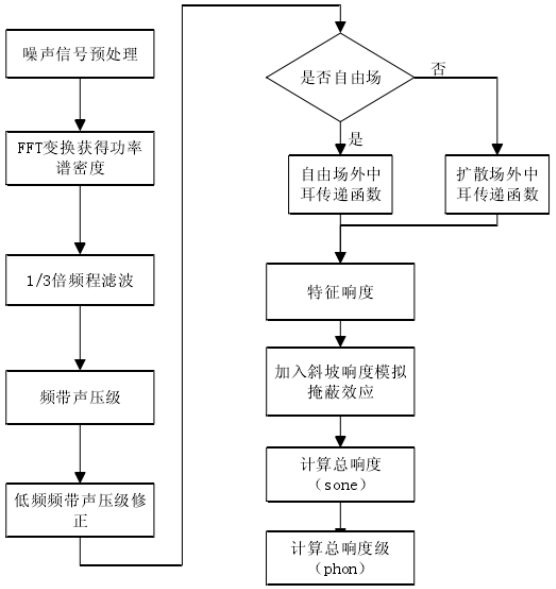


图 3-1 Zwicker 稳态响度模型算法流程图

$$N = \int_0^{24} n'(z) \cdot dz$$

$$L_N = \begin{cases} 40 + 33.22 \cdot \log_{10} \dfrac{N}{sone}, & N \geq 1\,sone \\ 40 \cdot \left(\dfrac{N}{sone} + 0.0005\right)^{0.35}, & N < 1\,sone \end{cases}$$

Psychoacoustics _ Loudness

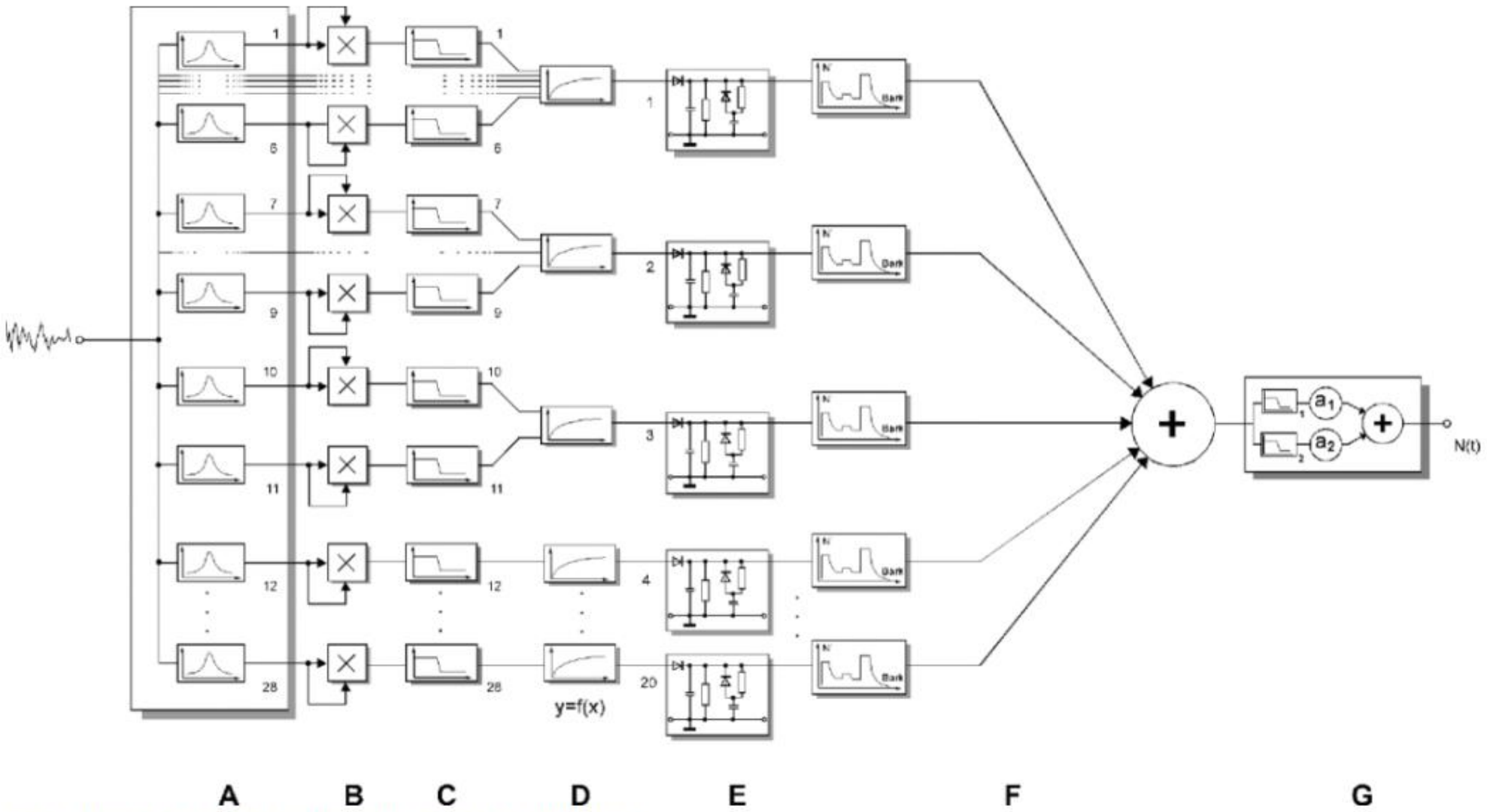


Figure 5 : Schematic representation for loudness calculation according to DIN 45631/A1

Psychoacoustics _ Loudness

- The calculation process eight steps electric

A. Calculation of the time-dependent 1/3 octave levels: a filter bank with 28 Chebyshev filters (low ripple) of 6th order is used for the calculation.
B. Calculation of intensities: the third-octave band signals are transformed to time-dependent intensity signals by squaring.
C. Time-related averaging: the temporal course of the intensities is smoothed through frequency-dependent lowpass filters.
D. Calculation of core loudness values: Calculation of the *core loudness values* is done according to the DIN. Thereby the signals of the lowpass filters 1 to 6, 7 to 9, as well as 10 and 11 are added as regarding to intensity. Those of the lowpass filters 12 to 28 are processed individually.
E. Generation of a fade-out time depending on duration by means of a diode network: a diode network described by Zwicker is utilized [2].
F. Calculation of the loudness summation: Taking 20 core loudness values, the specific loudness distribution is calculated initially. After that specific partial loudness values are summed. This complies with the method described in section DIN 45631 for stationary signals.
G. Temporal averaging of the loudness summation: The loudness summation is filtered with two lowpass filters of 1st order (time constant 3.5 and 70 ms). The following, weighted addition of these signals makes up the total loudness.

Psychoacoustics _ Sharpness

- Definition
- Noise containing high frequencies predominantly are perceived as “sharp” or “shrill”
- A measure for this sensation is sharpness
- Unit: “acum“
- **Definition**

The sharpness of 1 acum is assigned to a *narrow band* noise with a bandwidth of *160 Hz,* a middle frequency of *1 kHz and a sound pressure level of* $60\ dB_{SPL}$

Psychoacoustics _ Sharpness

- Standardization

Aures

$$S = K_1 \cdot \frac{\int_0^{24\,Bark} n'(z) \cdot e^{0.171 \cdot z} \cdot dz}{\ln\left(\frac{N}{20} + 1\right)}$$

DIN 45692

$$S = 0.11 \cdot \frac{\int_0^{24\,Bark} n'(z) \cdot z \cdot g(z) \cdot dz}{N}$$

$$g(z) = \begin{cases} 1 & , \quad z \leq 15.8\,Bark \\ 0.85 + 0.15 \cdot e^{0.42 \cdot (z - 15.8)} & , \quad z > 15.8\,Bark \end{cases}$$

von Bismarck

$$S = K_2 \cdot \frac{\int_0^{24\,Bark} n'(z) \cdot z \cdot g(z) \cdot dz}{N}$$

$$g(z) = \begin{cases} 1 & , \quad z < 14\,Bark \\ 1 + 0.003 \cdot (z - 14\,Bark)^3 & , \quad z \geq 14\,Bark \end{cases}$$

- ***n'(z)*** Specific Loudness [sone/Bark]
- ***z*** Critical Band Rate [Bark]
- ***N*** Total Loudness [sone]
- ***K_n*** Scaling factors

Standard DIN 45692

Specific loudness

Weighting function

$$S = 0,11 \cdot \frac{\int_{z=0}^{24\mathrm{Bark}} N'(z) \cdot g(z) \cdot z \, dz}{\int_{z=0}^{24\mathrm{Bark}} N'(z) dz}$$

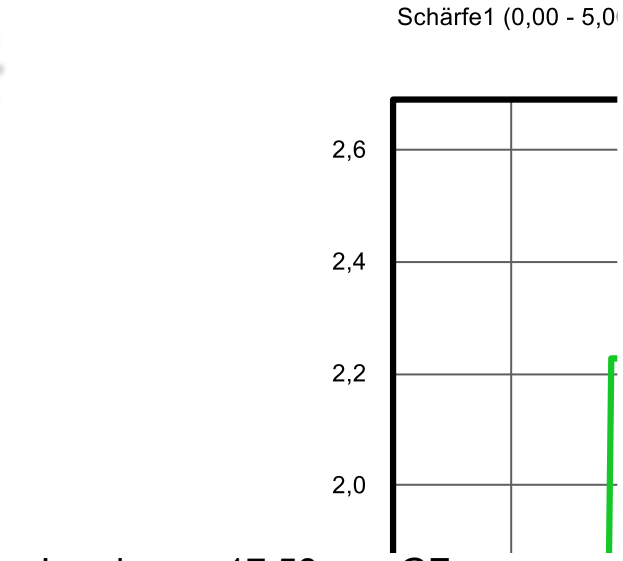
Frequency group

Overall loudness

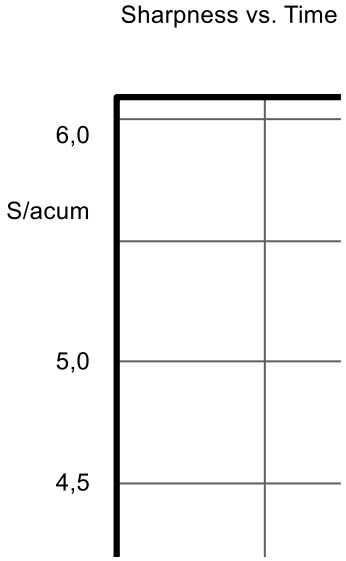
$$g(z) = 1 \quad \text{for} \quad z \leq 15,8\mathrm{Bark}$$

$$g(z) = 0,15 \cdot e^{0,24 \cdot (z - 15,8)} + 0,85 \quad \text{for} \quad z > 15,8\mathrm{Bark}$$

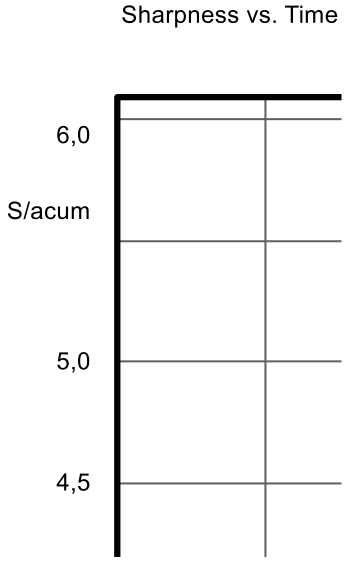
Psychoacoustics _ Sharpness



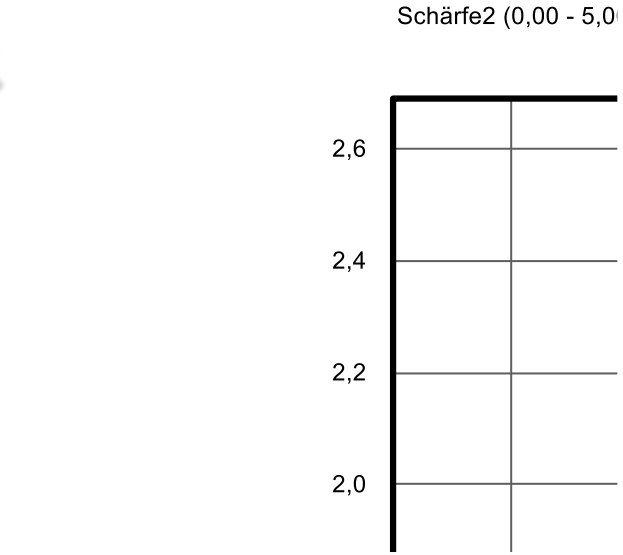
Loudness=17,58soneGF



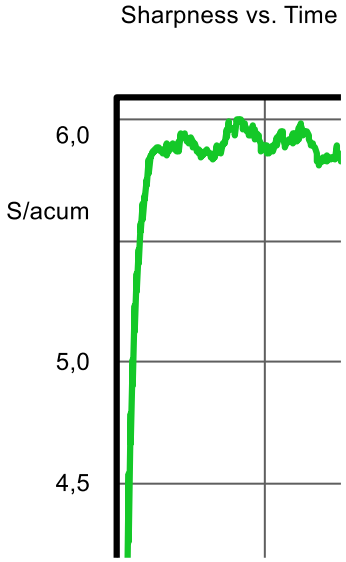
Sharpness=1,30acum Aures



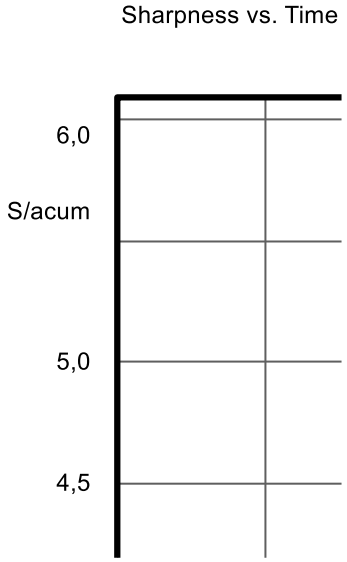
Sharpness=0,89acum v. Bismarck



Loudness=16,88soneGF



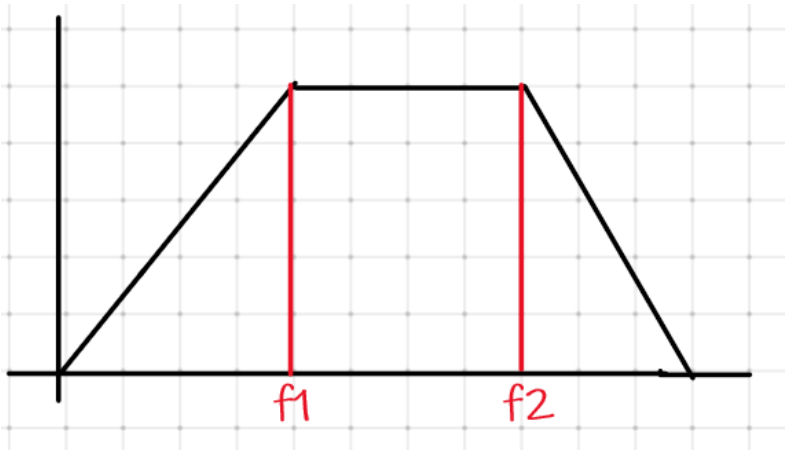
Sharpness=5,88acum Aures



Sharpness=2,98acum v. Bismarck

Psychoacoustics _ Sharpness

- Influence on sharpness
- When the noise is narrow band:
  - Spectral distribution (width of the band)
  - Middle frequency
- When the noise is wide band:
  - Rising the lower cut-off frequency→ rising sharpness
  - Lowering the lower cut-off frequency → descending sharpness



- Europe/Northern America: sharpness correlates with annoyance

- Asia: More tolerant with high sharpness

Psychoacoustics _ Sharpness

- Summary
- The sharpness strongly depends of the position of the maximum of the spectrum, the higher the sharper
- The higher the sharpness the lower the *sound quality*
- Different calculation methods, the used method has to be indicated
- *Unit*

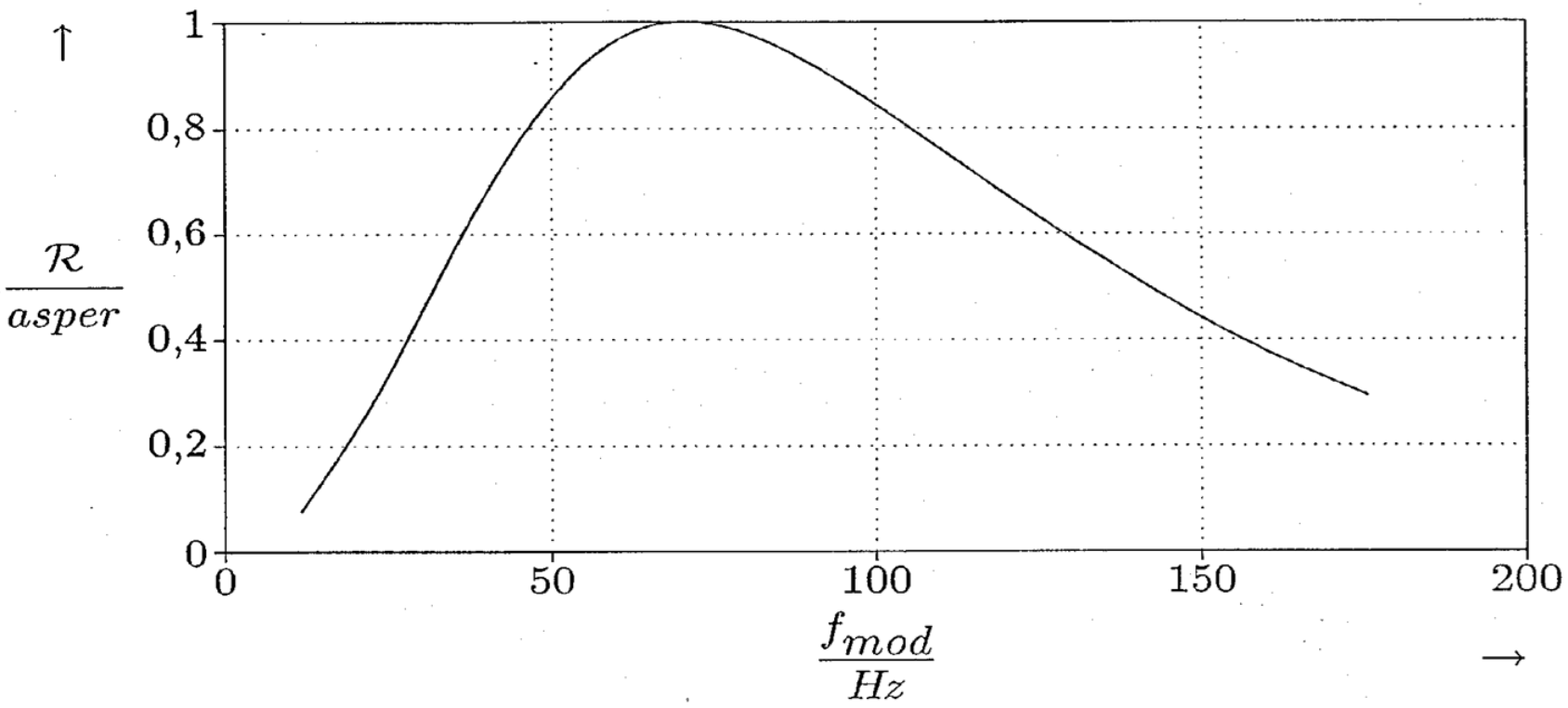
  1 accum = narrow band noise with a level of $60\ dB_{SPL}$, a middle frequency of $f_0 = 1\ kHz$ and a bandwidth of $160\ Hz$ (1 Bark)

Definition : Fluctuation strengths

- 1 vacil corresponds to a $\mathbf{1}\ \boldsymbol{kHz}$ tone with $\mathbf{60}\ \boldsymbol{dB_{SPL}}$, modulated with a modulation degree of $\boldsymbol{m} = \mathbf{1}$ and a modulation frequency $\boldsymbol{f_{mod}} = \mathbf{4}\boldsymbol{Hz}$
- The unit of fluctuation strength is “Vacil”

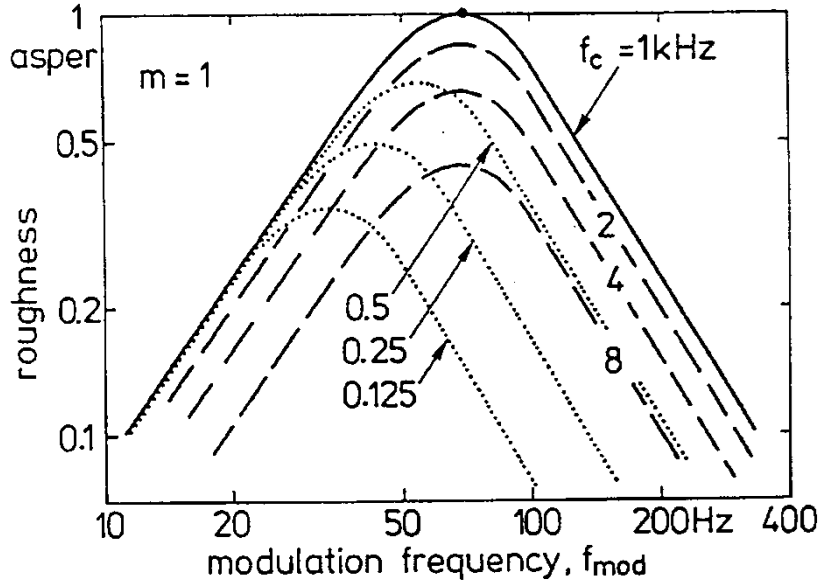
Definition :Roughness R

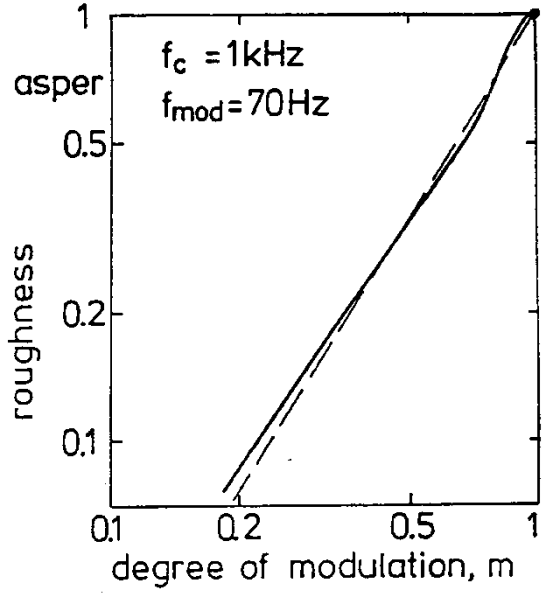
- The Unit of Roughness is „asper“.
- 1 Asper is defined as a $\mathbf{1\ kHz}$ tone of $\mathbf{60\ dB_{SPL}}$, modulated with a modulation depth **m = 1** and a modulation frequency of $\boldsymbol{f_{mod} = 70\ Hz}$



Parameter with influence on the roughness

- Modulation frequency f_{mod}
- Degree of modulation m (Amplitude modulation)
- Frequency deviation f_{dev} (Frequency modulation)





Psychoacoustics _ Fluctuation strength and roughness

- Measure for amplitude and/or frequency modulations within the signal

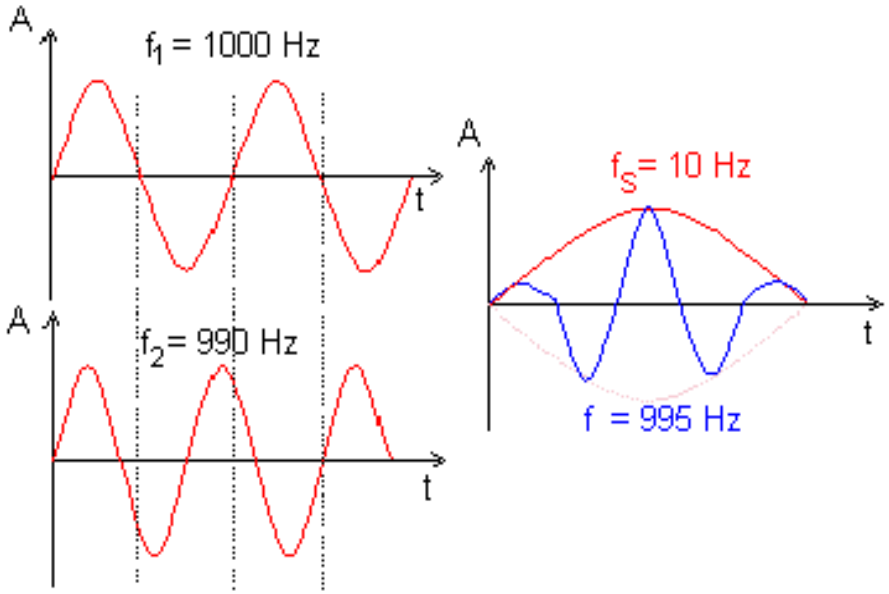
- → temporal structure of the signal

Psychoacoustics _ Fluctuation strength and roughness

- Origin

- The perceived pitch is right in the middle between the original pitches

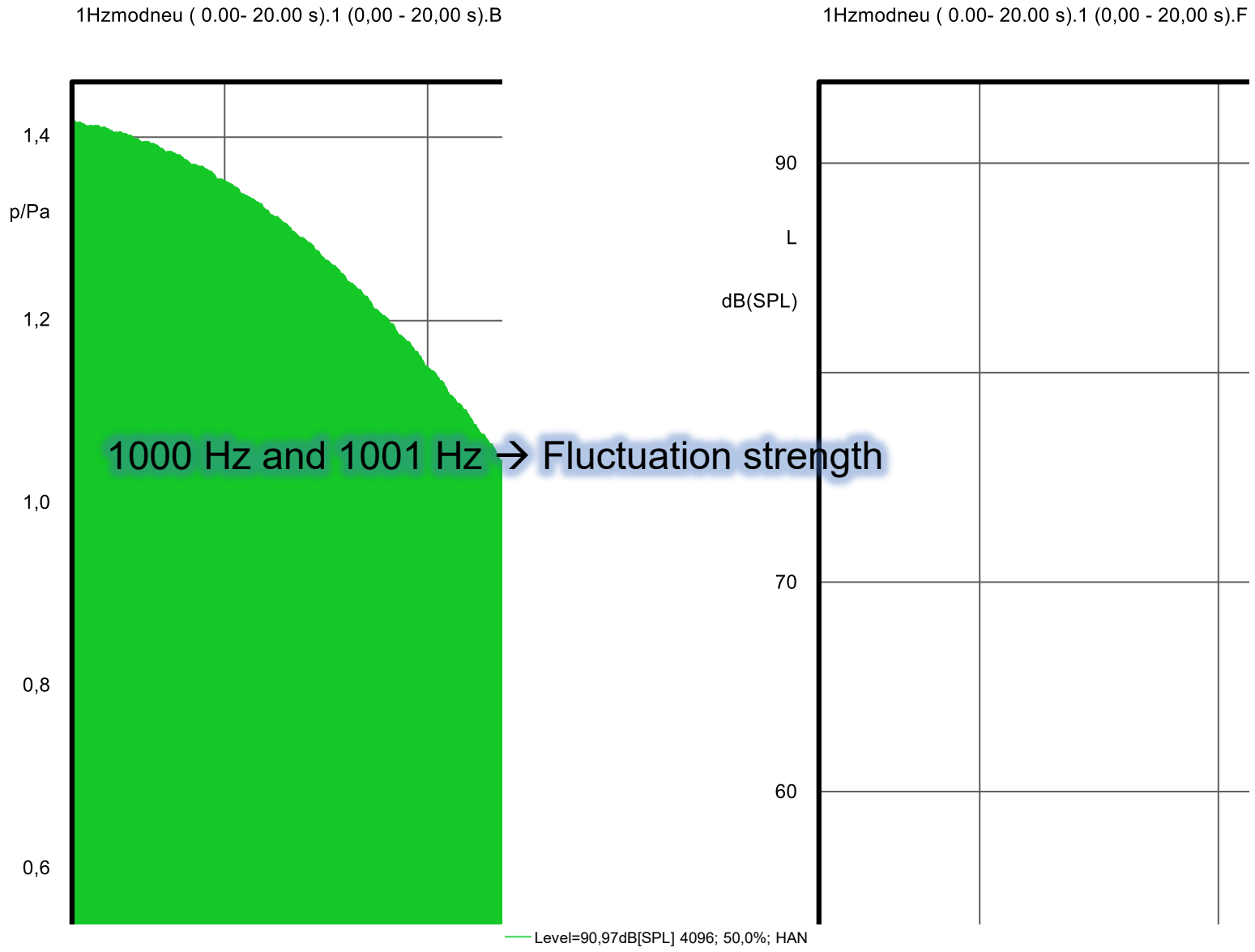
$$f = \frac{f_1 + f_2}{2}$$



- The Frequency of the beat if the gap between the tones

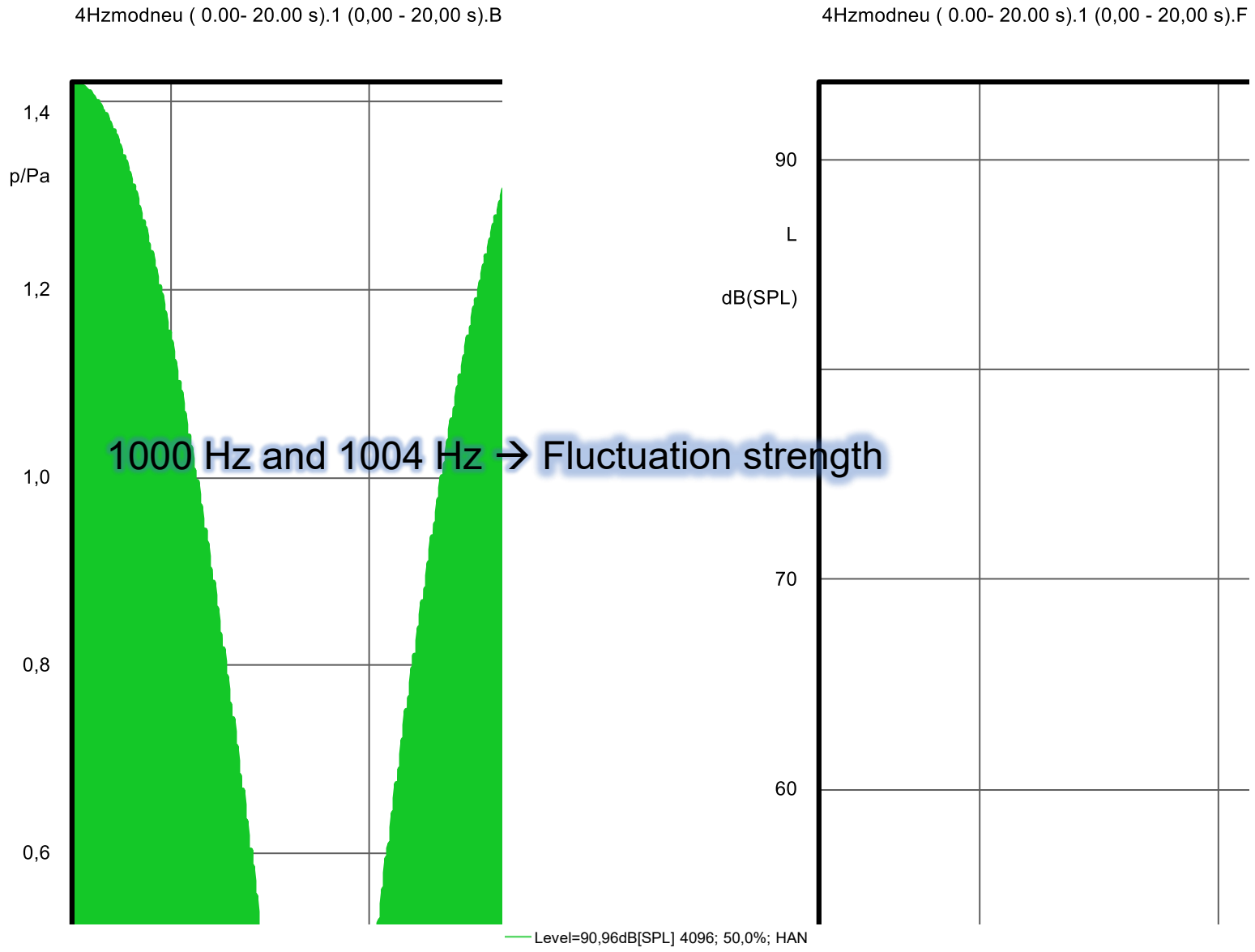
$$f_s = f_1 - f_2$$

Fluctuation strength and roughness



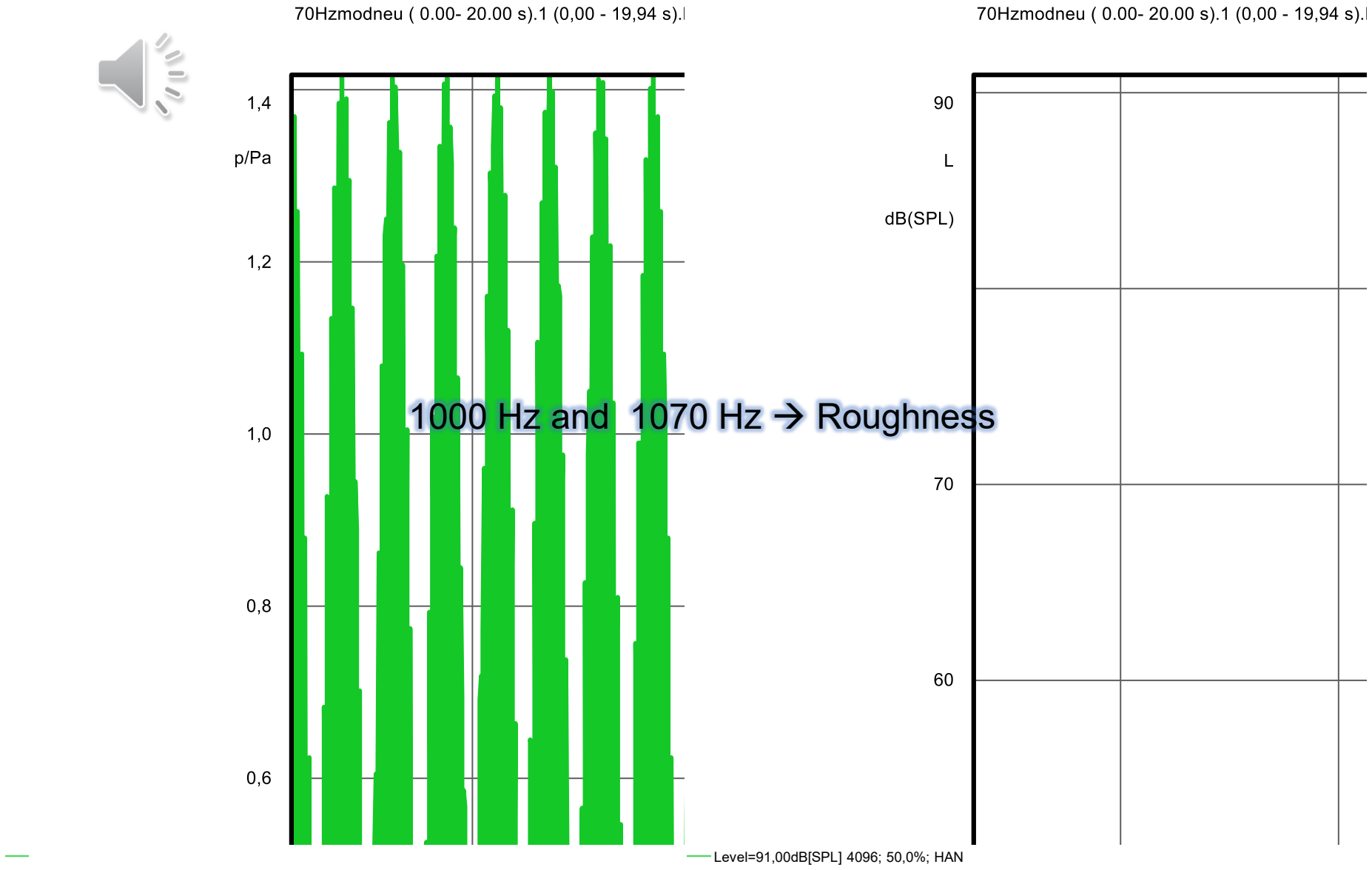
1000 Hz and 1001 Hz → Fluctuation strength

Fluctuation strength and roughness

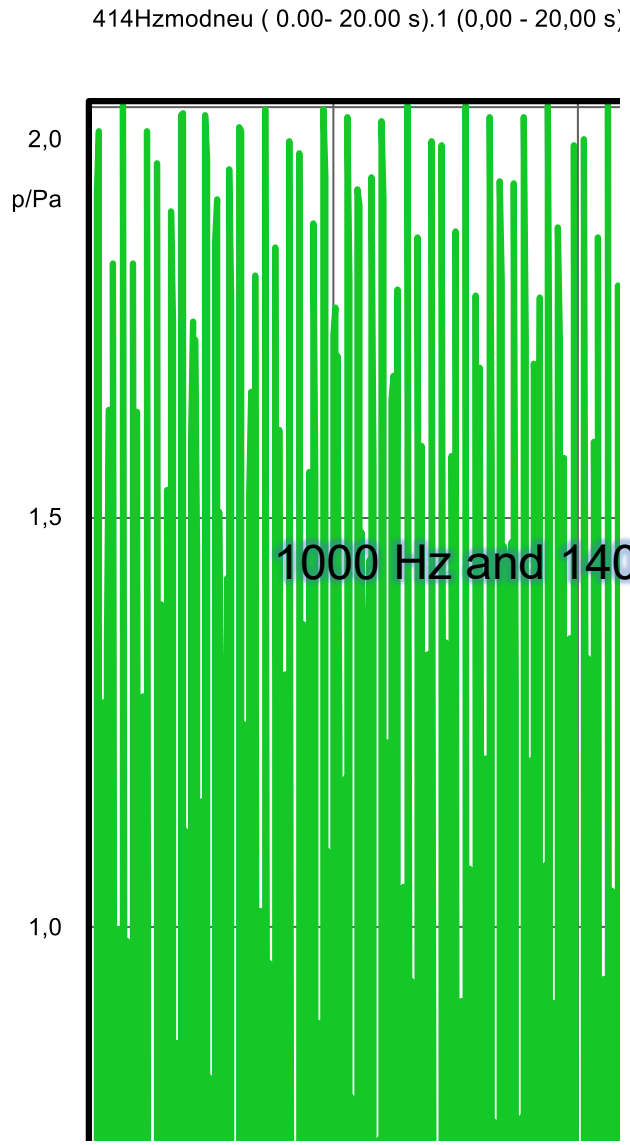


1000 Hz and 1004 Hz → Fluctuation strength

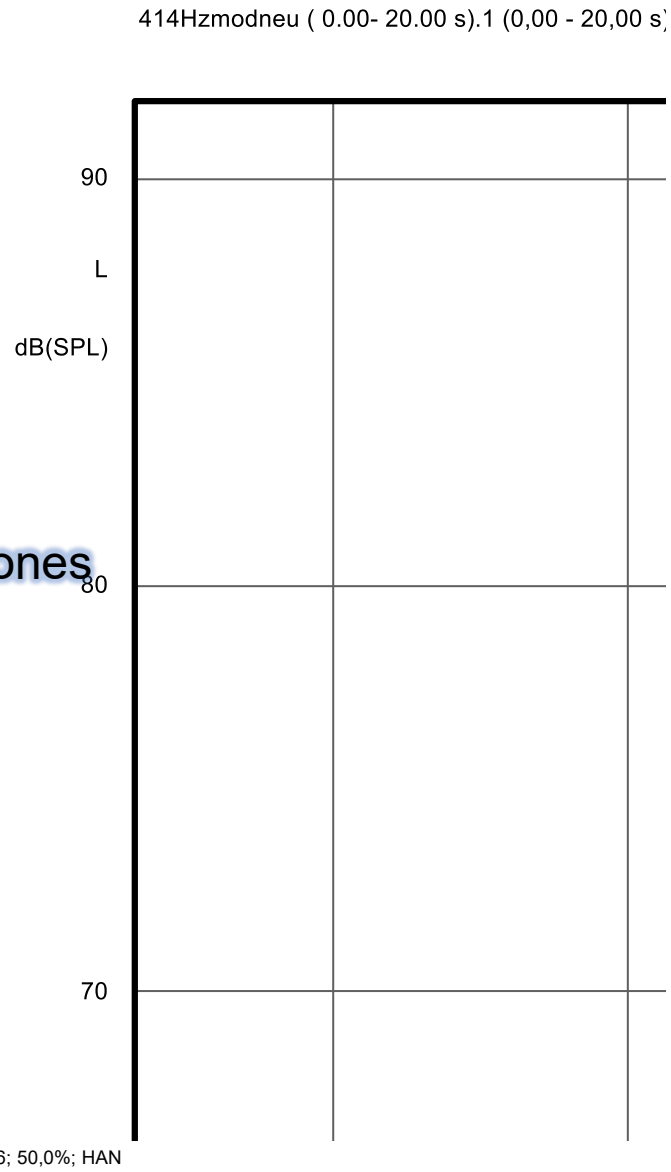
Fluctuation strength and roughness



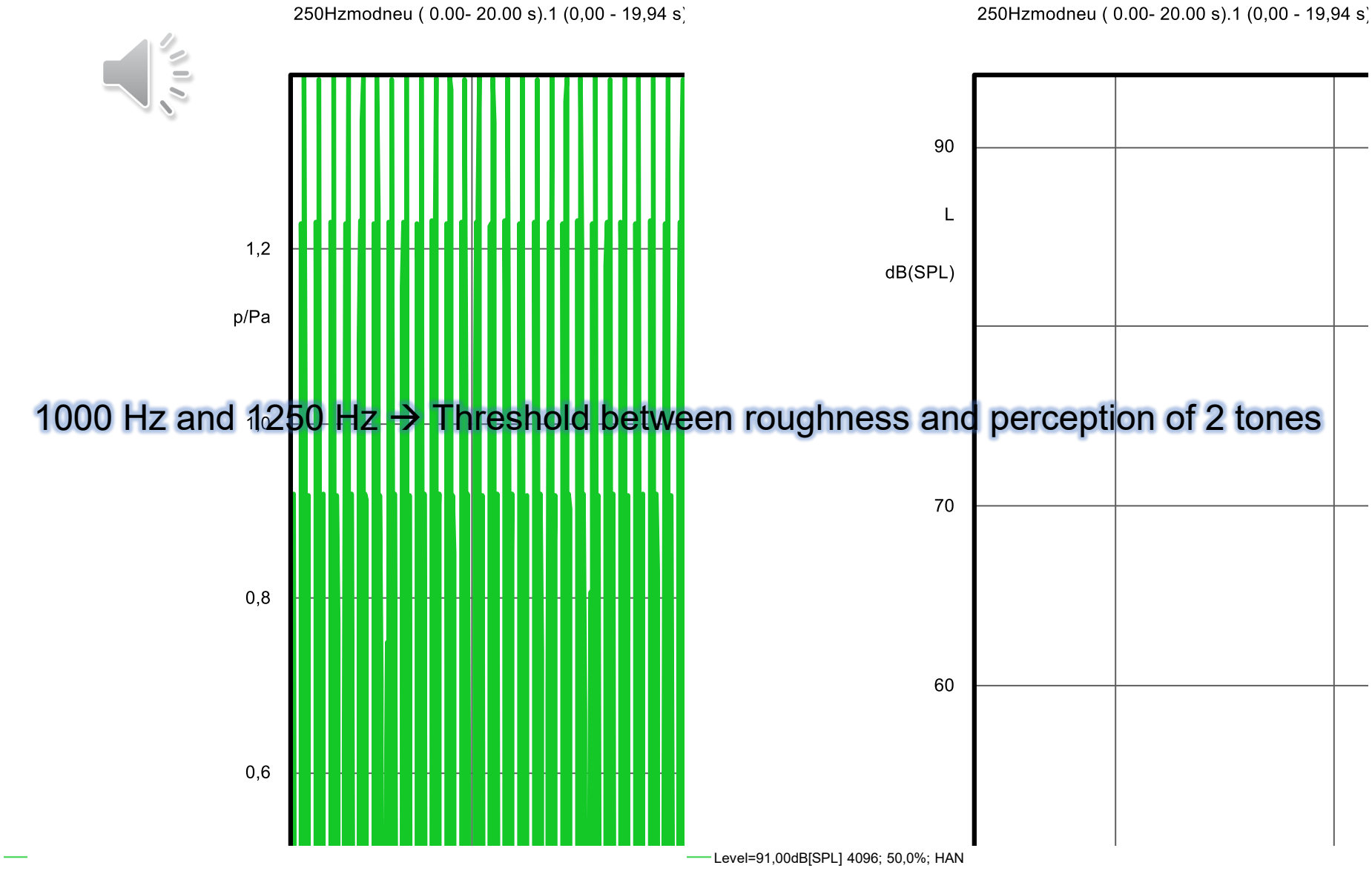
Fluctuation strength and roughness



1000 Hz and 1400 Hz → 2 tones

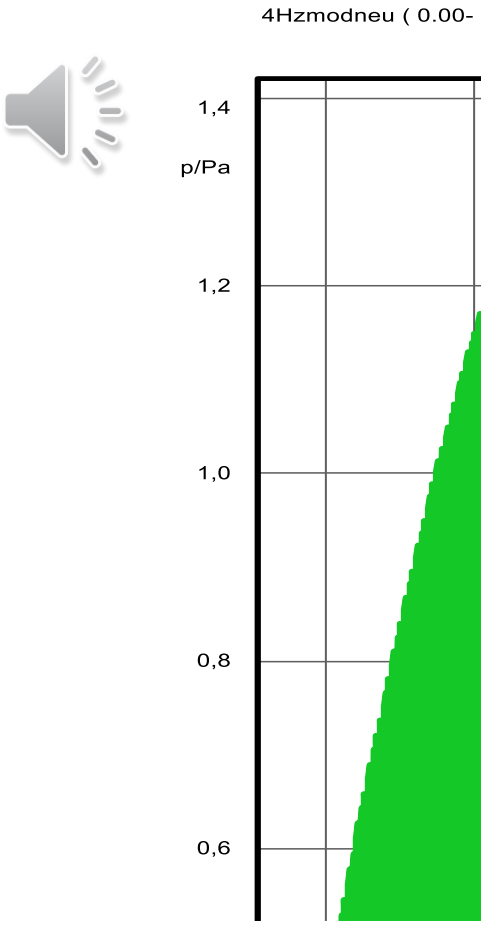


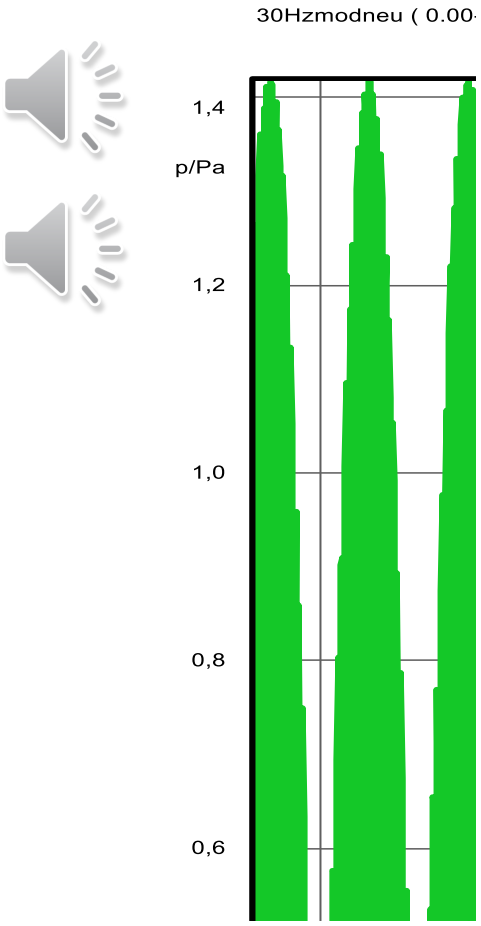
Fluctuation strength and roughness

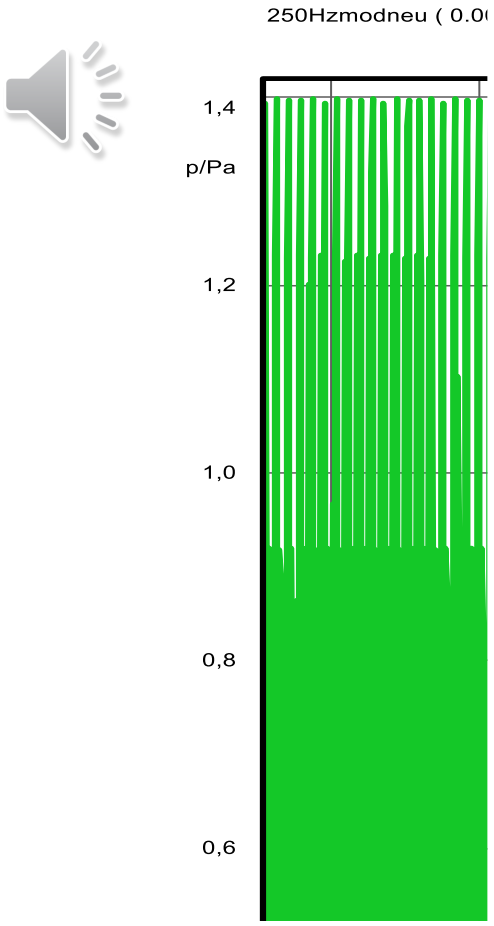


1000 Hz and 1250 Hz → Threshold between roughness and perception of 2 tones

Fluctuation strength and roughness







$25\ \text{Hz} < f_{AM}$

Fluctuation strength

$25\ \text{Hz} \leq f_{AM} \leq 250\ \text{Hz}$

Roughness

$15\ \text{Hz} \leq f_{AM} \leq 30\ \text{Hz}$

R-Roughness

$f_{AM} > 250\ \text{Hz}$

2 separate tones

Modulated signals, tone complexes

sound	frequency / Hz	modulation-rate / Hz	fluctuation	roughness
	1000	0		
	1000	1		
	1000	4		
	1000	20		
	1000	70		
	1000 & 2000	0		

Psychoacoustics _ HEAD acoustics

Non standardized Analyses

High Resolution Spectral Analysis (only HEAD acoustics)

Relative Approach 2D, 3D (only HEAD acoustics)

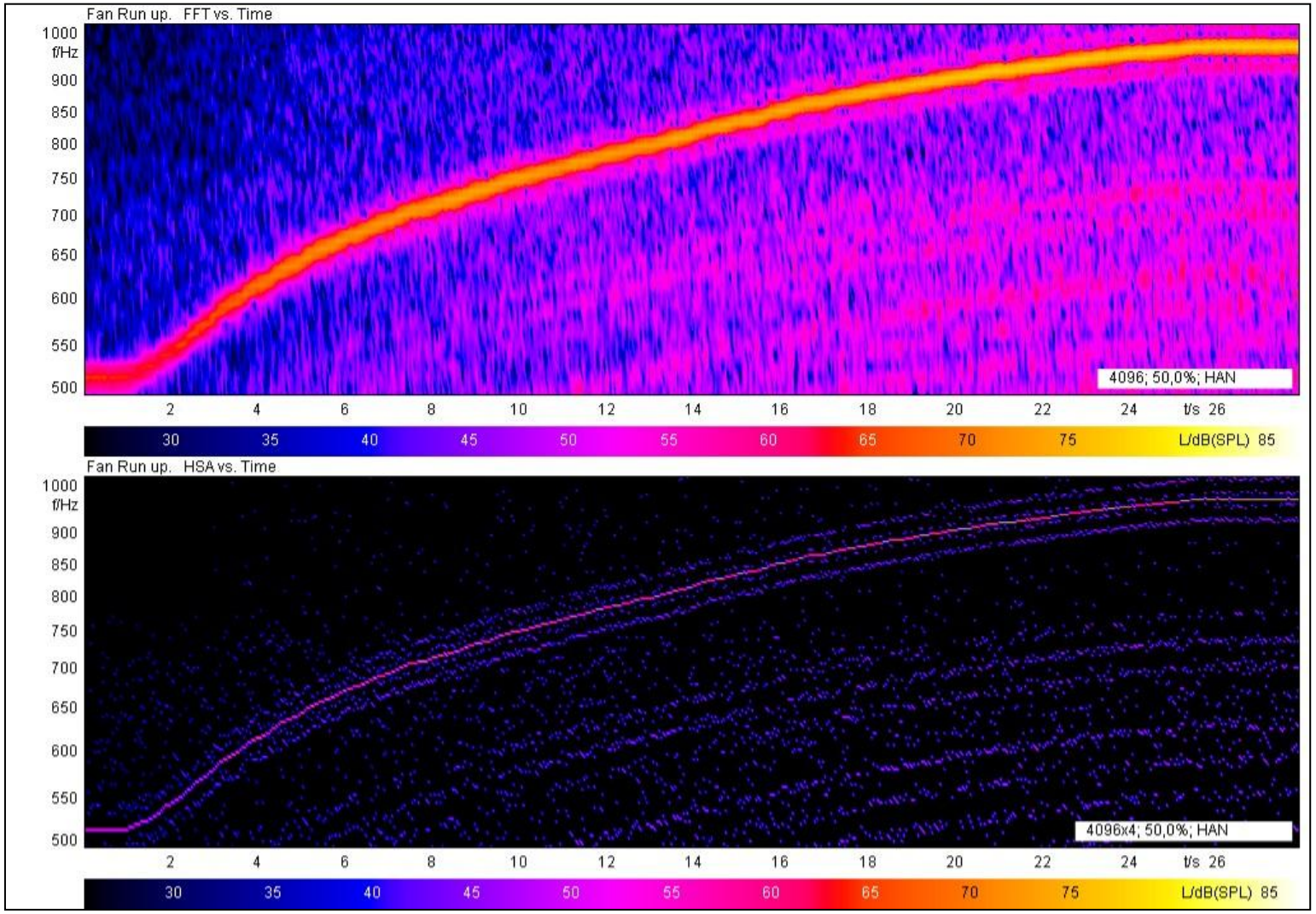
Standardized Analyses

Tonality DIN 45681

Specific Prominence Ratio (ECMA 74 -8, -9)

Tonality Hearing Model (ECMA 74-15) (only HEAD acoustics)

Psychoacoustics _ HEAD acoustics

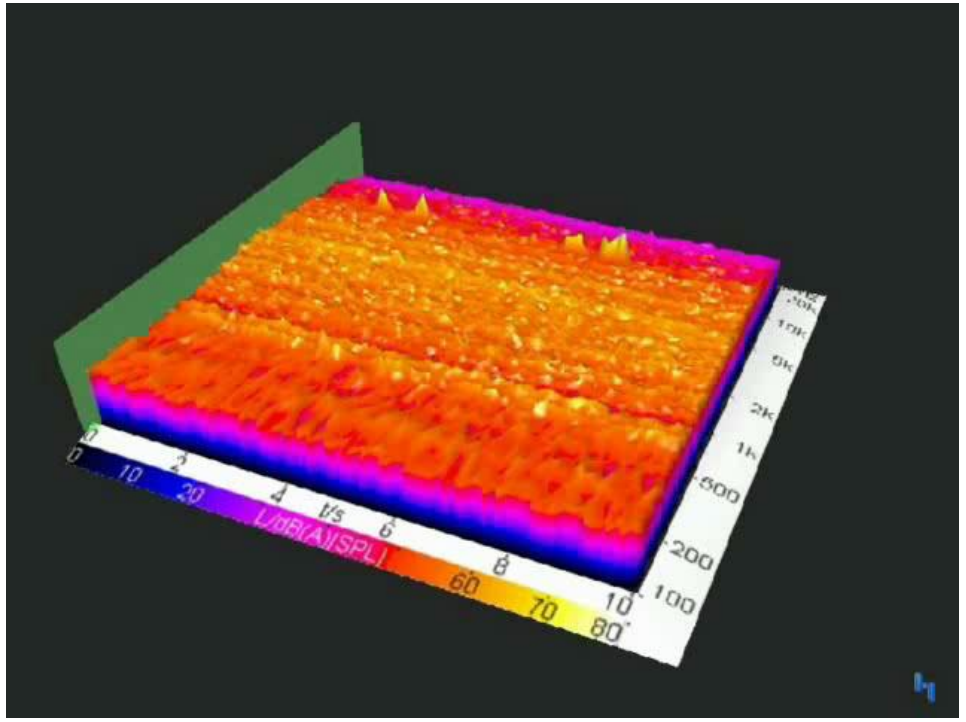


FFT vs. Time

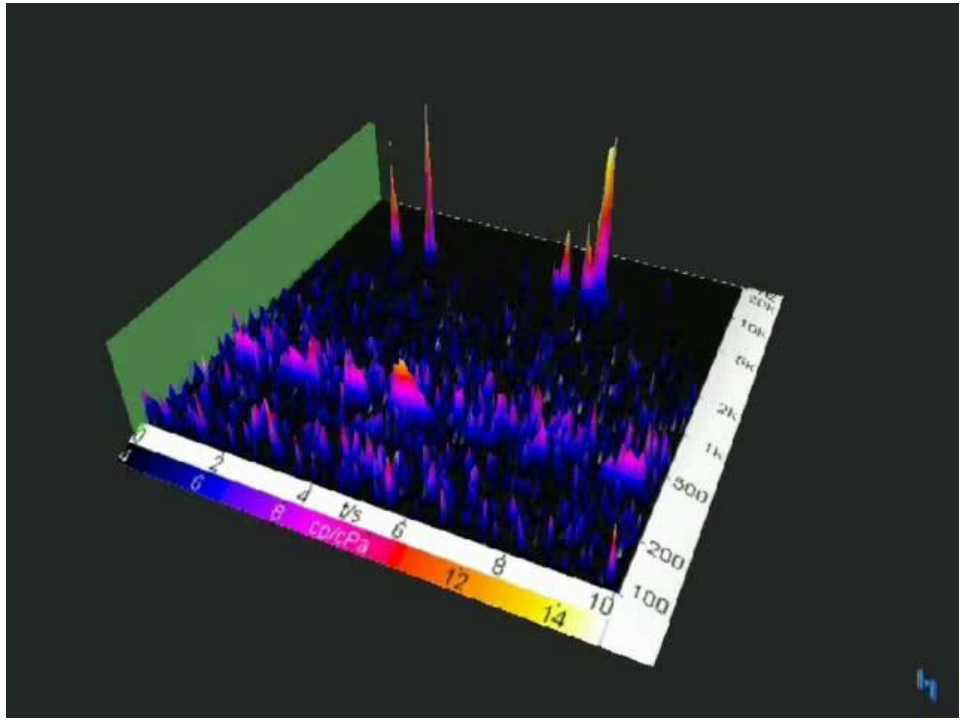
HSA

Psychoacoustics _ HEAD acoustics

HEAD acoustics
专利算法



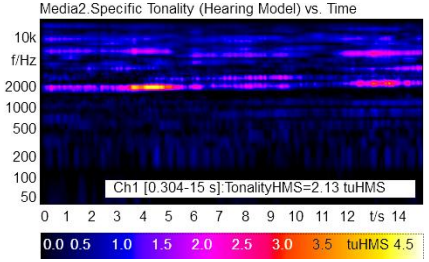
FFT vs. Time



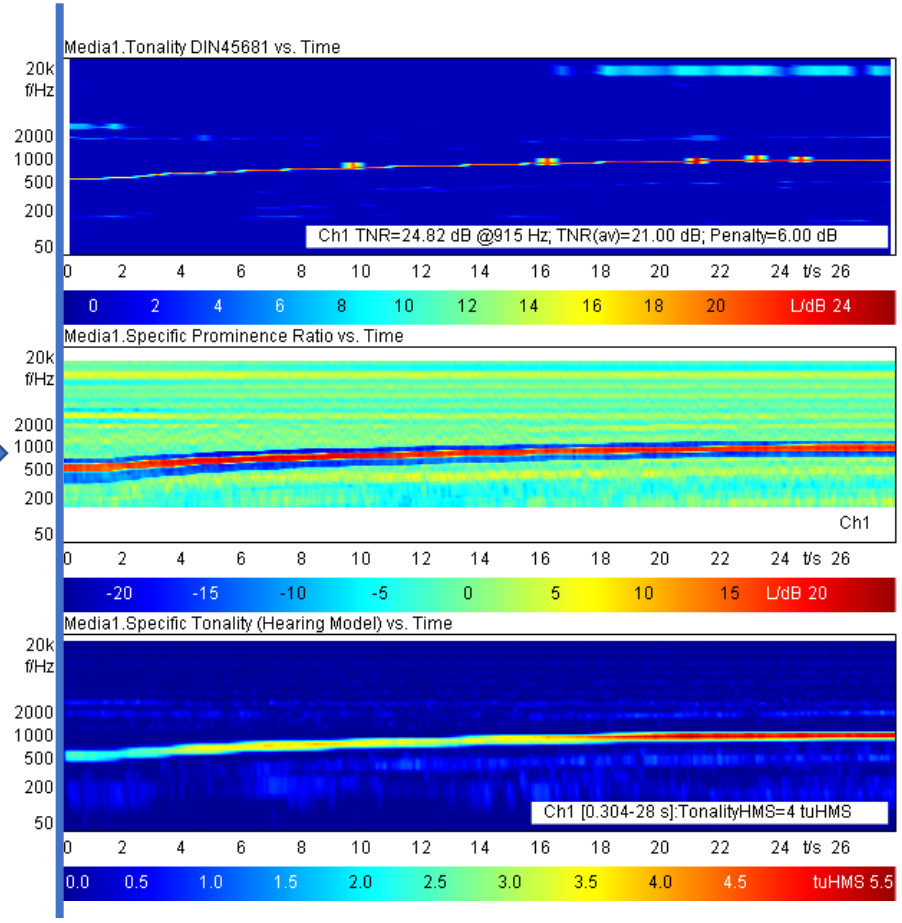
Relative Approach 3D

Tonality (Hearing Model)

Hearing Model



Tonality
(Hearing Model)



Tonality DIN 45681

Prominence Ratio

HEAD专利算法：
Tonality (Hearing Model)

基于Hearing Model的Tonality算法：

- 可以适应背景情况真实反应人耳感受
- 已写入ECMA 74(2018年)

Tonality (Hearing Model)

序号	算法名词	来源
1	Tone to Noise Ratio	ECMA 74 & ISO 7779
2	Prominence Ratio	ECMA 74 & ISO 7779
3	Tonality	DIN 45681
4	Tonality vs. time	Aures & Terhardt