AUDITIVE GESTALTUNG – SS 2010 – Lesson 1

1. Acoustics Basics

1.1 Dynamic Range - Amplitude

Bel (B) = unit of measure for ratios
dB (deciBel) = 1/10 Bel (more practical for everyday use); it uses a logarithmic scale in base 10; a dB only expresses a ratio between two quantities, hence it is a *dimensionless unit*; it can be combined with a suffix to specify a *reference level* in order to express an absolute value. The reference level can be a sound pressure level, like 0.00002 Pa; or voltage amplitude, like 775 mV; or a digital value.
Ratio in dB = 20 x log10 (U1 / U 0), where U0 is the reference level and U1 the level being measured
dB SPL = unit of measure for *Sound Pressure Level*; ref. level = 0 dB SPL, defined as 0.00002 Pa
1 Pa (Pascal) = unit of measure for *pressure* = 1 N/m² (1 Newton per square meter)
Standard atmospheric pressure at sea level = 101 325 Pa = 1 013.25 millibar (1 millibar = 100 Pa)
Threshold of hearing (or threshold of audibility) = 0 dB SPL = 0.00002 Pa (pressure deviation); note: it varies with frequency
Threshold of pain = 137.5 dB SPL = 150 Pa (pressure deviation); note: it varies with frequency
Human hearing max dynamic range: about 137,5 dB from the threshold of hearing to the threshold of pain
dBu = unit of measure for the *amplitude* of analogue audio signals (electrical voltage), with ref. level **0 dBu = 0.775 V** (775 mV). A typical standard reference level for studio equipment is **+4 dBu**
dBV = unit of measure for the *amplitude* of analogue audio signals (electrical voltage), with ref. level **0 dBV = 1 V** (1 000 mV). A typical standard reference level for consumer equipment is **-10 dBV**

<u>Examples of simple dB calculation</u>

How much difference in dB is there between an electrical signal of 2V amplitude, compared to the reference level 0 dBu?
20 x log10 (2 / 0.775) = 8.234 dB
How much difference in dB is there between an electrical signal of 2V amplitude, compared to the reference level 0 dBV?
20 x log10 (2 / 1) = 6.020 dB
How much difference in dB is there between the reference levels 0 dBV and 0 dBu?
20 x log10 (1 / 0.775) = 2,21 dB
Remember: the difference in level between -10 dBV and +4 dBu is
[14 - 2,21 = 11,79 dB] , and not 14 dB!

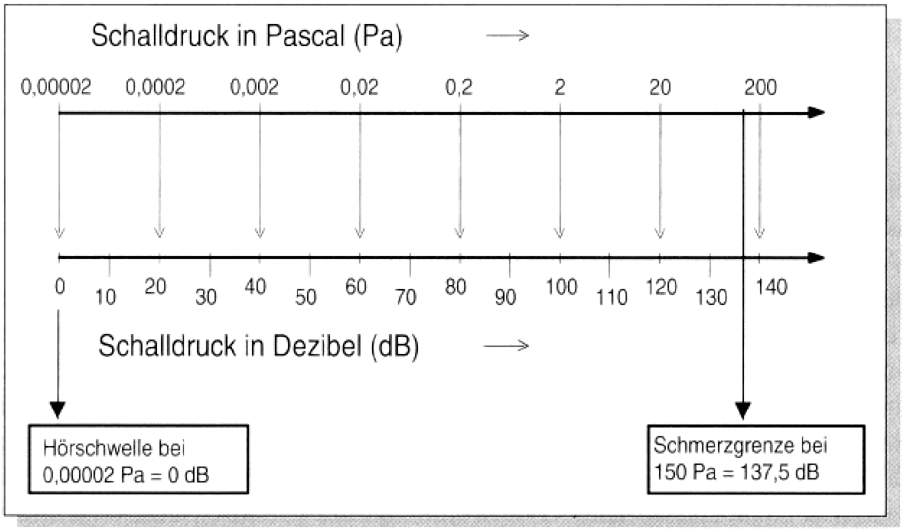
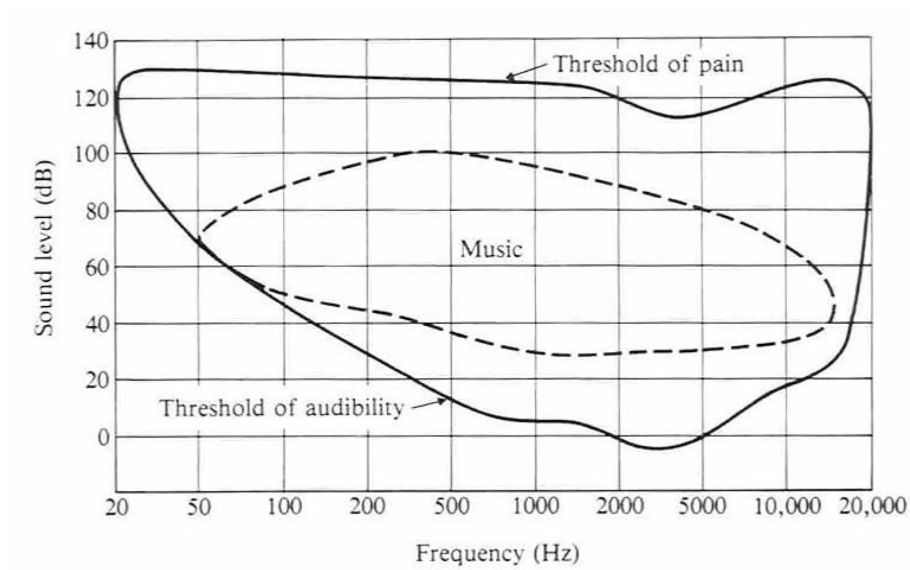


Abb. 1.3: Gegenüberstellung von Schalldruckwerten in Pascal und Dezibel

Sound pressure levels in Pascal (linear scale, Pa) and dB SPL (logarithmic scale)
(from: "das Tonstudio Handbuch" - Hubert Henle)

1.2 Frequency Range

Hz (Hertz) = unit of measure for frequency; 1 Hz (Hertz) = 1 cycle per second; 1 kHz = 1 000 Hz; 1 MHz = 1 000 kHz; 1 GHz = 1 000 MHz
Human hearing frequency range: about 16 - 20 000 Hz (for young individuals that did not suffer any hearing trauma, did not spend too much time at the disco or at rock concerts, etc. ...); with increasing age, the tendency is to lose sensitivity for the high frequency range.
Infrasound = frequencies below 16 Hz
Ultrasound = frequencies above 20 kHz
1 octave = doubling of the frequency in Hz
Human hearing frequency range = about 10 octaves bandwidth (from 16 to 16 384 Hz there are exactly 10 octaves)



The freq. range in the **hearing of animals** can vary greatly; some examples:

dog (60 - 45 000 Hz)
cat (45 - 64 000 Hz)
horse (55 - 33 500 Hz)
mouse (1 000 - 91 000 Hz)
bat (2 000 - 110 000 Hz)
beluga whale (1 000 - 123 000 Hz)
owl (200 - 12 000 Hz)
chicken (125 - 2 000 Hz)

1.3 Relation between Dynamic and Frequency Range

The **Fletcher/Munson diagram** is a measurement of *subjective loudness perception*, plotted for different dynamic ranges and across the complete frequency range

Phon = unit of measure for the *perceived loudness*; 1 Phon corresponds to 1 dB at 1 000 Hz in the dB SPL scale; it does not correspond to the dB SPL scale at any other frequency; see also the *Fletcher/Munson diagram* for reference

The **isophon** (= equal loudness) contours show what sound pressure level (in dB SPL) is required to produce a certain loudness sensation (in Phon) at different frequencies

Analysis of the isophon contours shows that *our perception of loudness changes with the frequency*: we are generally less sensitive to low (less than 100-200 Hz) and very high frequencies (more than 10 kHz), and most sensitive in the range between 2 and 5 kHz (which is very important for speech recognition).

It also shows that *our perception of loudness changes across its dynamic range*: our hearing "frequency response" is less linear at very low listening levels, and more linear at mid to high listening levels (which is why 85 dB SPL are usually used as reference for mixing and audio production).

<u>Example</u>: a subjective loudness of 90 Phon requires 90 dB SPL at 1 kHz, only 80 dB SPL at 4 kHz, bus as many as 110 dB SPL at 30 Hz!

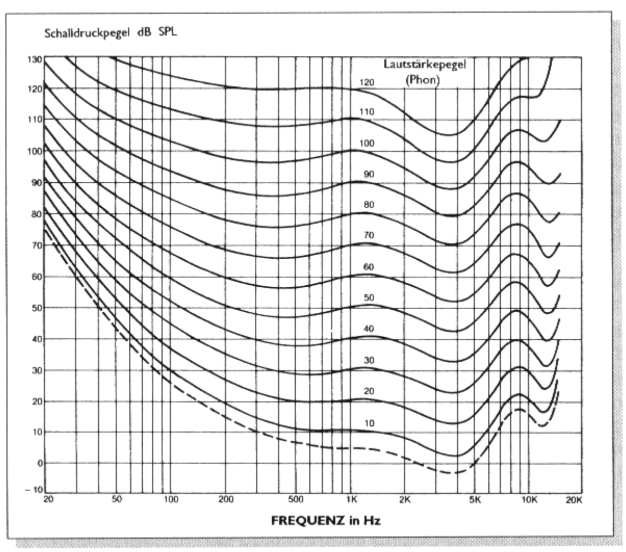


Abb. 1.4: Fletcher/Munson-Diagramm

(from: "das Tonstudio Handbuch" - Hubert Henle)

Even the difference in dB required for the subjective perception of “double as loud” is not constant: at 1000 Hz, about 10 dB are required, but at very low frequencies variations of just 6 dB can produce the same effect

This non-linearity is more pronounced at very low listening levels: this is why home and car stereo systems often have a *loudness switch*, which boosts low and high frequencies to compensate for our the ear non-linearity when listening at very low levels

That is why it is important to work at specific levels when mixing and mastering, usually around 80-85 dB SPL: at this level the ear works more linear (or less non-linear ...)

It is also good to sometimes switch from the standard listening level to very soft, or very loud levels; or to listen through a door or from a stair case, to check if the most important elements in a mix are still clearly hearable ...

1.4 Speed of Sound

Speed of sound (propagation in Earth atmosphere) at sea level, 20° C temperature = 343.8 m/s (approx. 340 m/s); at sea level, 0° C temperature = 331.8 m/s

Wavelength (in m) = c / f (where “c” is the speed of sound in m/s, and “f” the frequency in Hz); for example: the wavelength of 440 Hz = 340 / 440 = 0.772 m

Frequency (in Hz) = c / L (where “c” is the speed of sound in m/s, and “L” the wavelength in m); for example: the frequency of a sound with 6 m wavelength = 340 / 6 = 56.6 Hz

<u>Quick reference</u>: sound travels 340 m in 1 sec; 34 m in 100 ms; 3.4 m in 10 ms; 34 cm in 1 ms

Max delay between right and left ear for an incident wave: 0.5 ms
(average distance between ears = about 17 cm)

The speed of sound depends mostly on *medium density*; as both altitude and temperature affect air pressure, they affect in turn also the speed of sound in the atmosphere; higher density = higher propagation speed

Speed of sound propagation at 0° C through:
- Air: 331 m/s
- Water: 1485 m/s
- Copper: 3710 m/s
- Iron: 5100 m/s
- Wood: 3000-4000 m/s
- Glass: 5000 m/s
- Hard rubber: 1500 m/s

The **sound propagation waves in free field** (an ideal open space with no boundaries and no reflections) for a sound source that is not moving are *spherical* in shape

If there is no medium, like in outer space, no sound propagation can occur. What we generally hear in science-fiction movies such as Star Trek and Star Wars (laser and phaser cannons, explosions, engine noise, etc.) is pure fiction, and it is just done to add an emotional layer to the movie. In reality, *there is no sound in empy space* (except if it is transmitted by direct contact between two objects).

1.5 Relation between Sound Power (Intensity), Sound Pressure Level and Loudness

Sound Power or **Intensity** is a measure of the *sound energy that passes through a given area each second*
Energy per second is measured in **Watt** (1 W = 1 Joule per second)
Intensity is related to the sound pressure amplitude: specifically *the energy in a wave is proportional to the square of the pressure amplitude*
Formula: $\mathbf{I = P^2}$, where "I" is the sound power (intensity), and "P" the sound pressure amplitude
Examples:
double sound pressure = 4 times the power (intensity)
1/2 sound pressure level = 1/4 the power (intensity)
The formula to translate sound power (intensity) (in W) to dB is:
Li = 10 x log10 (I1 / I0), where "I0" is the reference intensity and "I1" the intensity being measured
The formula to translate sound pressure level (in Pa) to dB is:
Lp = 20 x log10 (U1 / U 0), where "U0" is the reference level and "U1" the level being measured
The **sound power** (intensity) *doubles every 3,01 dB* [10 x log10 (2 / 1) = 3,01 dB] ...
The **sound pressure level** (in Pa) *doubles every 6,02 dB* [20 x log10 (2 / 1) = 6,02 dB]
The **perceived loudness** (in Phon) *doubles about every 10 dB*; also: perceived loudness is frequency dependant, and varies quite a lot between individuals (see Fletcher/Munson diagram!)
Examples:
10 times the power (in W) = +10 dB SPL = 3,162 times the SPL in Pa, but is just perceived as "double as loud"
100 times the power (in W) = +20 dB SPL = 10 times the SPL in Pa,
but is just perceived as "4 times as loud"
1000 times the power (in W) = +30 dB SPL = 31,62 times the SPL in Pa,
but is just perceived as "8 times as loud"

1.6 Relation between Sound Power (Intensity), Sound Pressure Level and Distance

In a free field (an ideal open space with no boundaries and no reflections), *the sound pressure level is inversely proportional to the distance from the sound source*
Formula: $p = 1/d$ (p = sound pressure, d= distance ratio)
Examples:
double distance = 1/2 sound pressure level = -6,02 dB
4 times the distance = 1/4 sound pressure level = -12,04 dB
1/2 the distance = 1/0.5 = double sound pressure level = +6,02 dB
In a free field, *the sound power (intensity) is inversely proportional to the square of the distance from the sound source*
Formula: $i = 1/d^2$ (i = intensity, d = distance ratio)
Think it like this: the sound waves carry energy; doubling the distance, this energy is spread on an area that is 4 times as large
Examples:
double distance = 1 / (2 x 2) = 1/4 power in W = -6,02 dB
(remember: 1/2 power is only -3 dB!)
4 times the distance = 1 / (4 x 4) = 1/16 power in W = -12,04 dB
1/2 the distance = 1 / (0.5 x 0.5) = 4 times the power in W = +6,02 dB

When dealing with **sound amplification systems** (amplifiers and loudspeakers), being able to calculate power requirement for a given sound pressure level at a given distance is very important, in order to decide what equipment must be used.

Example:

A loudspeaker has an *efficiency* of 90 dB SPL /W at 1 m distance; what sound pressure level will it produce at 32 m distance, in a free field?
p = 1/d = 1/32; to calculate the ratio in dB: 20 x log10 (1/32) = 30,10 dB
90 - 30,10 = 59,90 dB. So at 32 m distance this loudspeaker just produces about 60 dB SPL
How much more power is required to still produce 90 dB SPL at 32 m distance?
For each additional 10 dB SPL we need 10 times the power (in W); for +30 dB SPL we need 1 000 times the power! If this loudspeaker needed 1 W to produce 90 dB SPL at 1 m, we are going to need 1 000 W power to produce 120 dB SPL at 1 m, corresponding to about 90 dB SPL at 32 m distance.

2. Digital Audio Recording (DAR)

2.1 Sampling Basics

Sampling: the information of an analogue signal is reduced from *continuous time*, non quantized state to *discrete time*, quantized state, resulting in a *finite amount of digital* (= numerical) *information*

This **digital information** can be *processed in real time* (by DSP systems or general purpose computer CPUs) or *stored in digital form* on a *magnetic* medium (DAT, ADAT tape, Hard Disk) or *optical* medium (CD/DVD-R) for later processing

During **sample playback** the process is inverted: the digital information is used to reconstruct the original analogue signal; provided sampling rate and quantization are accurate enough, the regenerated analogue signal will be virtually identical to the original

Sampling Frequency or **Sampling Rate** (in Hz) = *temporal accuracy*: how many samples (discrete amplitude measurements) per sec. of the analogue signal are being taken

The **frequency bandwidth** depends on the sampling rate (see Nyquist Theorem under)

Quantization (in bits) = amplitude accuracy: how accurate is the amplitude measurement of each sample

The **dynamic range** and S/N ratio depend mostly on the quantization accuracy

ADC: analogue to digital converter

DAC: digital to analogue converter

Typical **conversion time** @ 44.1 kHz sampling frequency required by a modern ADC or DAC with 128x oversampling: about 1 ms

Nyquist Theorem: the sampling rate (in Hz) must be at least 2 times the desired bandwidth (in Hz); in other words: to represent a given frequency it is necessary to have at least one sample per positive and one sample per negative phase of the wave cycle

Nyquist frequency = 1/2 the sampling rate

Aliasing = distortion in the form of signal artifacts not present in the original signal, and not related to it in a harmonic way (this is not THD, Total Harmonic Distortion!); this happens when signals above the Nyquist frequency enter the ADC without proper "antialiasing filtering" and get "mirrored" around the Nyquist frequency itself, appearing back into the hearable freq. range

Example:

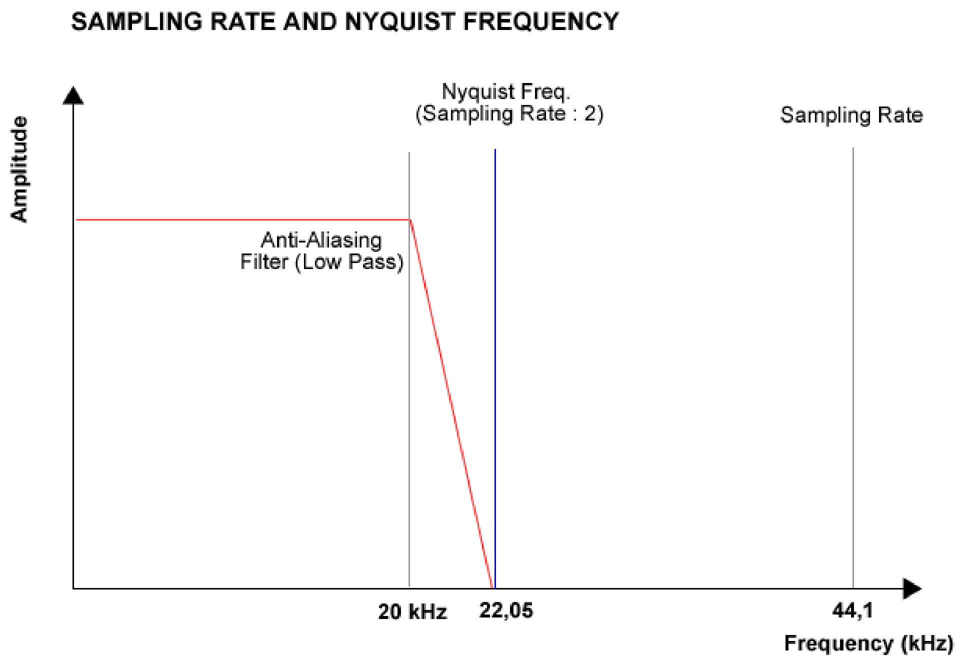
in a 44 100 sampling rate system, the Nyquist frequency is 22 050 Hz; without filtering, a signal of 30 000 Hz would be mirrored at 22 050 - (30 000 - 22 050) = 14 100 Hz

An **antialiasing filter** is generally a low pass filter with very sharp slope; as practical filters cannot be manufactured with "infinite slope", a higher sampling rate is required in order to effectively remove all frequencies above the Nyquist, without affecting the desired bandwidth

Example:
in a typical “CD quality” recording (44 100 Hz sampling rate) the antialiasing filter must not affect signals under 20 000 Hz, but must filter all signals above 22 050 Hz to avoid aliasing, leaving just 2 050 Hz bandwidth to switch from "full pass" to "full cut" operation

A similar (inverse) process occurs in the DAC: after DA conversion, an analogue low pass filter removes all undesired artifacts generated by the sampling process over 20 000 Hz

Oversampling: modern ADC and DAC use a different approach, in which the signal is first filtered by a much more “relaxed” *analogue antialiasing filter* and sampled at a much higher frequency (for example 128 x 44 100 = 5 644 800 Hz); then a *digital low pass filter* is used (which is easier and cheaper to implement than a similar analogue antialiasing filter, and can also be very accurate); finally the signal is *downsampled* to the desired rate (for example 44 100 Hz) before further processing or storage on a medium. A similar inverted process occurs at sample playback.



2.2 DAR – Dynamic and Frequency Range

Typically, the **bandwidth** of a digital recording is less than half the sampling rate.

Example:
A digital recording with *44.1 kHz sampling rate* allows for a *20 - 20 000 Hz bandwidth* (however, usually also freq. under 20 Hz can be recorded and reproduced, depending on the converters)
A digital recording with *88.2 sampling rate* could allow up to *40 000 Hz bandwidth*.

Digital recording dynamic range and signal to noise ratio (simplified formula):
S/N ratio in dB = 6 N + 1.8 (where N is the number of quantization bits)
Max. dynamic range for a 16-bit digital recording: (6 x 16) + 1.8 = 97.8 dB
Max. dynamic range for a 24-bit digital recording (in the digital domain): (6 x 24) + 1.8 = 145.8 dB

However**:** the **typical dyn. range of good quality 24-bit ADC/DAC units** is just about 110-120 dB; this means: the 24-bit signal theoretical dynamic range (145.8 dB) can only be achieved in the digital domain; actual recordings will only manage up to 110-120 dB, which is the dynamic range offered by most good quality 24-bit ADC/DAC units

dBfs = dB "full scale" = unit of measure for the *amplitude of digital audio signals*; the *reference level* is "0 dBfs", which is also the maximum signal amplitude that can be stored digitally in a typical digital audio recording system (for example: DAT, ADAT, DTRS; also 16 or 24-bit WAV/ AIFF files)
Signals louder than "0 dBfs" just produce "**clipping**" (= truncation of the waveform, hence distortion), except in a **32-bit float system** (for example: the audio engine of HDR systems such as *Cubase*, *Nuendo* and *Logic*)
In a digital audio recording system most signal levels are defined as a *negative* dBfs amount; for example, -6 dBfs = 6 dB quieter than a "full scale" signal (at 0 dB), or 6 dB from clipping
Even in systems supporting **32-bit float resolution** (where theoretically the full 24-bit resolution of the signal is maintained throughout the signal path), the output should never exceed 0 dBfs before DA conversion, or clipping occurs in the converter, causing distortion
To avoid clipping of the DA stage in a HDR systems supporting internally *32-bit float resolution*, it is usually enough to reduce the level of the master output fader, unless clipping occurs before in some plugin that does not support the 32-bit float format

2.3 DAR – Hard Disk Usage

To calculate HD usage recording at different resolutions (stereo, multitrack, 16 and 24 bit):
HD usage in Bytes/min = [Bytes of quantization per channel] x [n. of audio channels] x [sampling freq. in Hz] x 60 sec
"CD quality" recording (16-bit stereo, 44.1 kHz): 2 Bytes x 2 Ch. x 44100 x 60 = 10 584 000 Bytes/min = 10,093 MB/min
24-bit recording, 8 track, 96 kHz: 3 Bytes x 8 Ch, x 96000 x 60 = 138 240 000 Bytes/min = 131,835 MB/min
Remember: for HD manufactures, 10 584 000 equals to 10,58 MB, but for your Operating System this is just 10,093 MB
1 KB = 1024 (and not 1000) Bytes; 1 MB = 1024 KB; 1 GB = 1024 MB; 1 TB = 1024 GB; etc.
Disk manufactures want to advertise larger disk capacities, so they sell a hard disk as a 120 GB capable device, while the effective capacity as seen by the OS and programs is only 111,7 GB ...

Bibliography / Further Reading:

- EDERHOF, Andreas: *das Mikorfonbuch* – GC Carstensen 2004 (ISBN 3-910098-28-2)
- HENLE, Hubert: *das Tonstudio Handbuch* – GC Carstensen 2001 (ISBN 3-910098-19-3)
- HÖMBERG, Martin: *Recording Basics* – PPV Medien 2002 (ISBN 3-932275-21-7)

Website:

- www.digitalnaturalsound.com or www.dns-studios.com
  > Medialab & MMA > Auditive Gestaltung