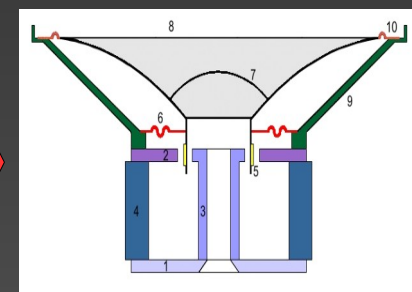


Zvuk a sluch

- Záznam
 - mikrofon (uhlíkový, dynamický, piezo, kondenzátorový, ...)
 - A/D převod (vzorkování/kvantizace, 44 kHz/16bit)
- Reprodukce
 - D/A převod (mono, stereo, 5.1)
 - Reproduktor (elektrodynamický, elektrostatický, piezo, ...)



Záznam zvuku

- Zaznamenaný zvuk je sekvence amplitud o různé frekvenci a fázi, jednokanálový nebo vícekanálový záznam
- Analogový vs. digitální záznam zvuku
- Lidské vnímání zvuku: 20 Hz – 20kHz, pro telefonní záznam postačuje 8 kHz
- Lidské ucho nejcitlivější v rozsahu 2 – 4000Hz, nejvíce informace pro porozumění řeči 500 – 2000Hz
- Dynamický rozsah (Citlivost) lidského ucha – rozdíl vnímání hlasitého a slabého zvuku 120dB 2-4 kHz, jinak je nižší
- Psycho-akustický model lidského ucha
- CBR (Constant Bitrate): Jednokanálový záznam, 16kHz, jeden vzorek 16 bit. → konstantní datový tok: 256 000 bit./s
- VBR (Variable Bitrate), ABR (Average Bitrate)

Psychoakustický model lidského ucha

- Umožňuje odstranění informace ze zvukového signálu, které jsou pro lidský sluch nezaznamatelné (neslyšitelné) a jsou tedy v signálu navíc, nemělo by dojít ke změně vnímané kvality

1) omezení frekvenčního rozsahu

- citlivost lidského ucha je omezená. Slyšitelné frekvenční pásmo je cca 20 Hz až 20 kHz podle věku.

2) frekvenční maskování

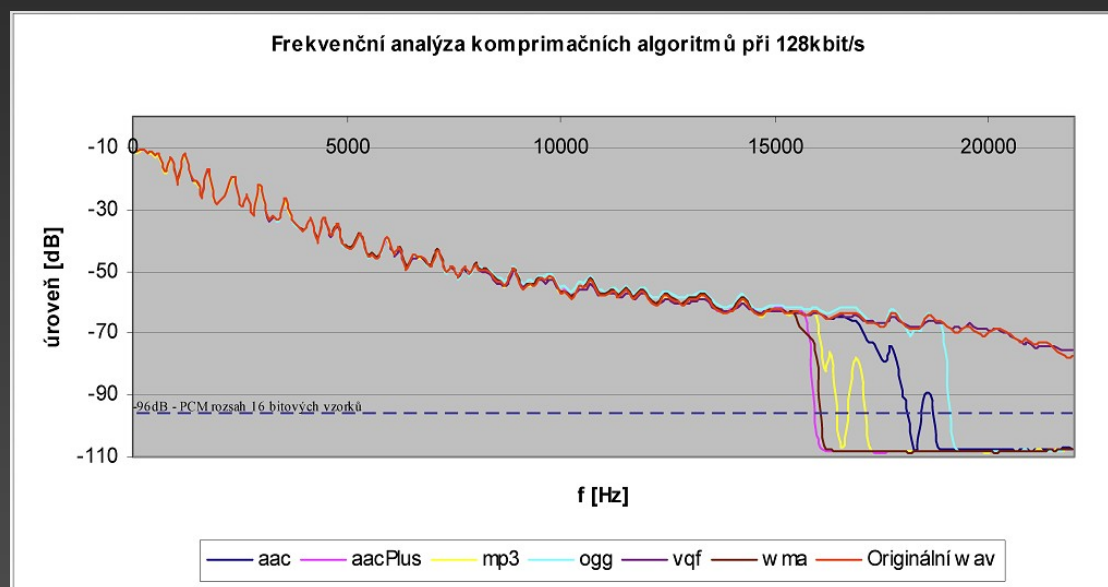
- nelineární citlivost lidského sluchu: pokud zazní silný zvukový signál, tak nelze zaznamenat slabší (frekvenčně blízký) signál, lze tedy slabší vynechat bez subjektivní ztráty kvality

3) časové maskování

- setrvačnost vnímání lidského sluchu: po zániku silnějšího zvukového signálu trvá jistou chvíli, než je lidský sluch schopný zaznamenat současně působící slabší zvukový signál

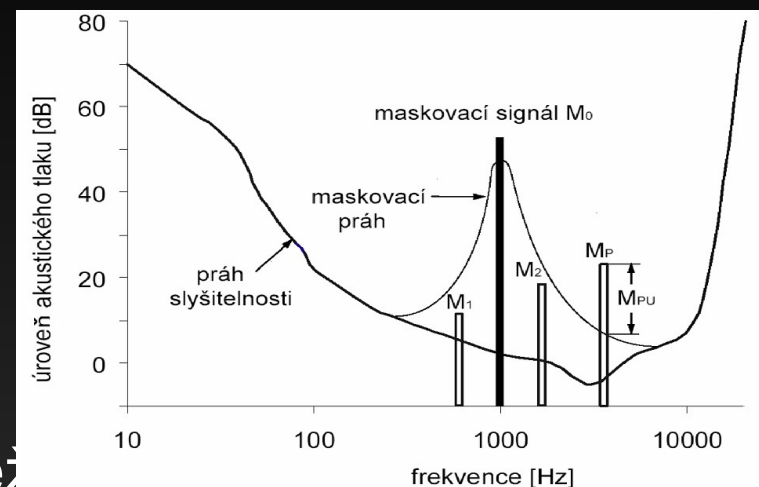
Omezení frekvenčního rozsahu

- Zvukový signál je omezen pásmovou propustí, odstraní se (především) vysoké frekvence
- CD (44 kHz) + vzorkovací teorém \rightarrow 22 kHz
 - V praxi je u CD low-pass filtr na 20 kHz
 - U metod ztrátové komprese kolem 16 kHz



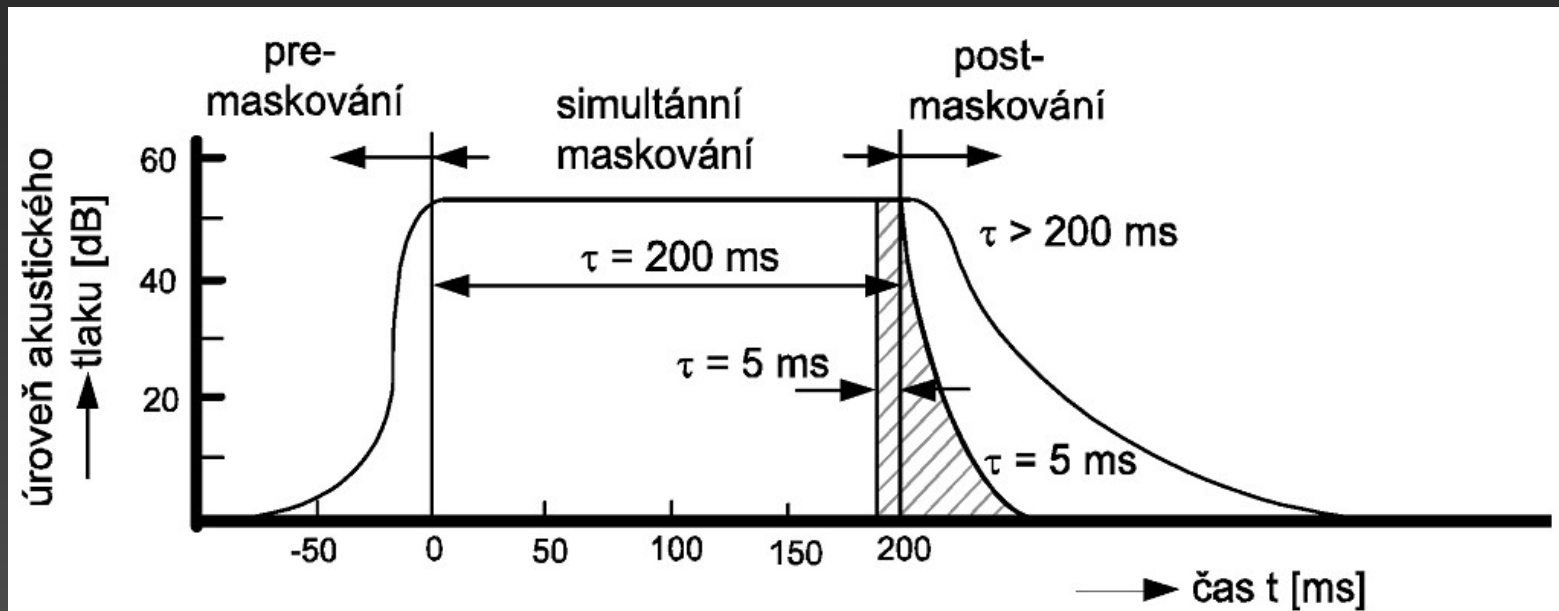
Frekvenční maskování

- Schopnost lidského sluchu odlišit dvě různé (blízké) frekvence
- Dvě frekvence jsou ve zvukovém signálu současně, jedna frekvence je slyšitelná a druhá ne
 - závisí na intenzitě (amplitudě) náležející k dané frekvenci
- Frekvenční maskování
 - soubor (interval) frekvencí nahrazen jednou frekvencí o specifické intenzitě
- Důležité je správně zvolit šířku pásma (soubor frekvencí) pro frekvenční maskování
 - proměnná – nižší frekvence užší (100 Hz) na vysokých až 4kHz



Časové maskování

- Když po hlasitější frekvenci zazní slabší, tak určitý čas je slyšitelnost této frekvence potlačena
- Funguje částečně i naopak, kdy je potlačena slyšitelnost slabší frekvence před hlasitější.



Komprese zvuku

- Neztrátová komprese
 - Rekonstruovaná data budou přesně stejná jako originální signál
 - Využívá pomocných transformací (např. stereo = společná část + rozdíly) a běžných kompresních metod jako RLE, Huffman, LZW, ZIP apod.
- Ztrátová komprese
 - Využívá frekvenční omezení a maskování + dodatečnou neztrátovou kompresi

Zvukové formáty neztrátové

- WAV (Waveform Audio Format)
 - Ve skutečnosti kontejner, ale obvykle obsahuje neztrátové audio
 - PCM (Pulse Code Modulation)
 - nejjednodušší způsob kódování dat, bez komprese, navzorkovaná data ukládána jako posloupnost celočíselných dat
 - DPCM (Diferenciální) PCM
 - ukládán rozdíl vzorku oproti předchozí hodnotě vzorku (20% datová úspora)
 - ADPCM (Adaptivní) PCM
 - DPCM + proměnná velikost vzorku

Zvukové formáty neztrátové

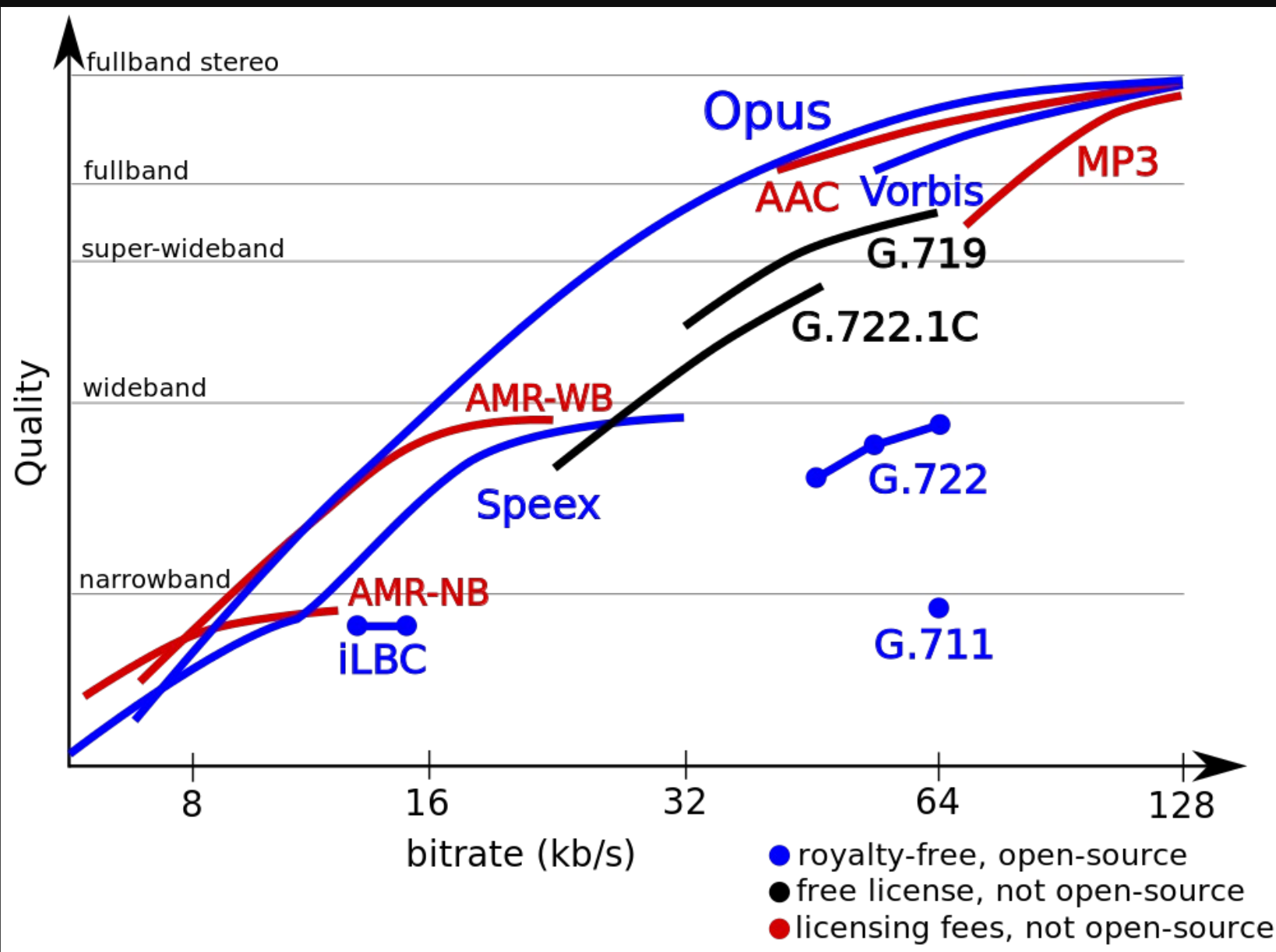


- Ogg FLAC (Free Lossless Audio Codec)
 - Silná podpora, aktuální, zdarma
 - Kompresní poměr 50-60% originálních dat
- ALAC (Apple Lossless Audio Codec)
 - Kompresní poměr 40-60% originálních dat
- WMA Lossless (Windows Media Audio)
 - Kompresní poměr 40-60% originálních dat
 - Placené, nepříliš rozšířené

Zvukové formáty ztrátové

- Využívá frekvenční omezení a maskování + dodatečnou neztrátovou kompresi
- MP3 (MPEG1 Layer 3)
 - Zvuková stopa z MPEG1 videa
 - Široká podpora
- AAC (Advanced Audio Coding)
 - Nástupce MP3, zvuková stopa z MPEG2 a MPEG4
- OGG Vorbis
 - Svobodný formát, podporovaný
- Opus
 - Nástupce OGG vorbis+speex, současná špička

Srovnání kvality ztrátových kodeků



Nízký datový tok: řeč

- Speciální kodeky
 - Omezená kvalita
 - mono (stereo)
 - bitrate < 32 kb/s
 - nízká latence
 - tolerance k výpadkům paketů
- Neuronové sítě
 - Lyra (Google)
 - EnCodec (Facebook)

