

MULTIMEDIÁLNÍ A HYPERMEDIÁLNÍ SYSTEMY

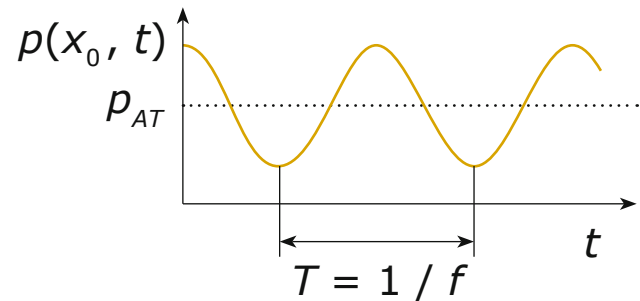
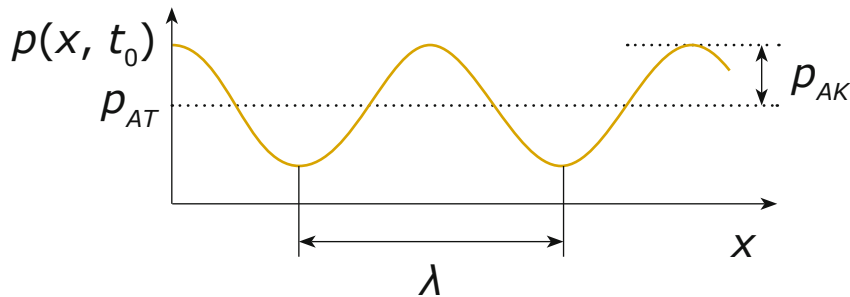
3)

Zvuk, sluch,
záznam a reprodukce zvuku

Petr Lobaz, 24.2.2015

ZVUK

- příčné kmitání molekul vzduchu



$p(x, t)$ – lokální okamžitý atmosférický tlak

p_{AT} – průměrný atmosférický tlak

$p_{AK}(x, t)$ – lokální okamžitý akustický tlak: $p(x, t) - p_{AT}$

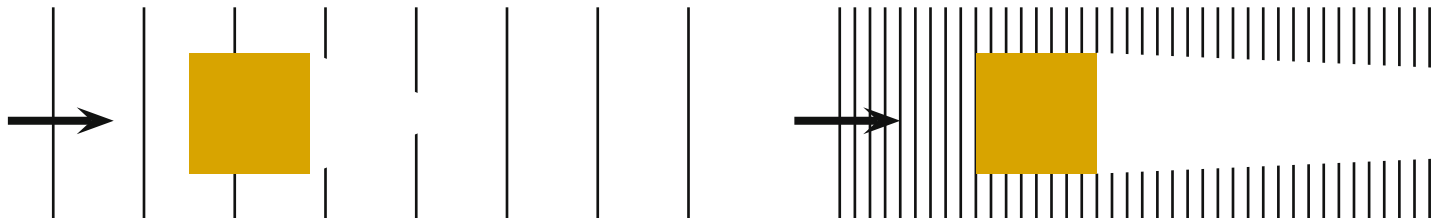
p_{AK} – amplituda akustického tlaku

λ – vlnová délka

f – (časová) frekvence, T – perioda vlnění

ZVUK

- rychlost zvuku $c \approx 340 \text{ m}\cdot\text{s}^{-1}$
 - praktičtěji $\approx 1 \text{ ft}\cdot\text{ms}^{-1}$ ($\approx 30 \text{ cm}$ za milisekundu)
 - ⇒ zpožděné odrazy zvuku od stěn umožňují odhadnout velikost místnosti
- vlnová délka $\lambda = c T = c / f$
 - pro 1000 Hz asi 1 ft ($\approx 30 \text{ cm}$)
 - ⇒ protože akustický stín vrhají pouze překážky výrazně větší než vlnová délka zvuku, rozlišíme „přímý“ zvuk od zvuku „za rohem“ podle přítomnosti vyšších frekvencí



(ZVUKOVÝ) SIGNÁL

- záznam signálu ani jeho reprodukce nejsou dokonalé
- teoretický signál $s(t)$, reálný signál $s'(t)$
 - typicky uvažujeme aditivní šum: $s'(t) = s(t) + n(t)$
- při hodnocení kvality reálného signálu porovnáváme výkon (energii) šumu a výkon (energii) teoretického signálu
 - odstup signál–šum = $\frac{\text{výkon(šum)}}{\text{výkon(t. signál)}}$
 - typicky 0,0001; 0,00001; 0,000001; ...
- praktičtější čísla určena logaritmickou škálou: decibely (dB)
 - odstup signál–šum [dB] = $10 \log_{10} \frac{\text{výkon(šum)}}{\text{výkon(t. signál)}}$
 - typicky -40 dB, -50 dB, -60 dB, ...

(ZVUKOVÝ) SIGNÁL

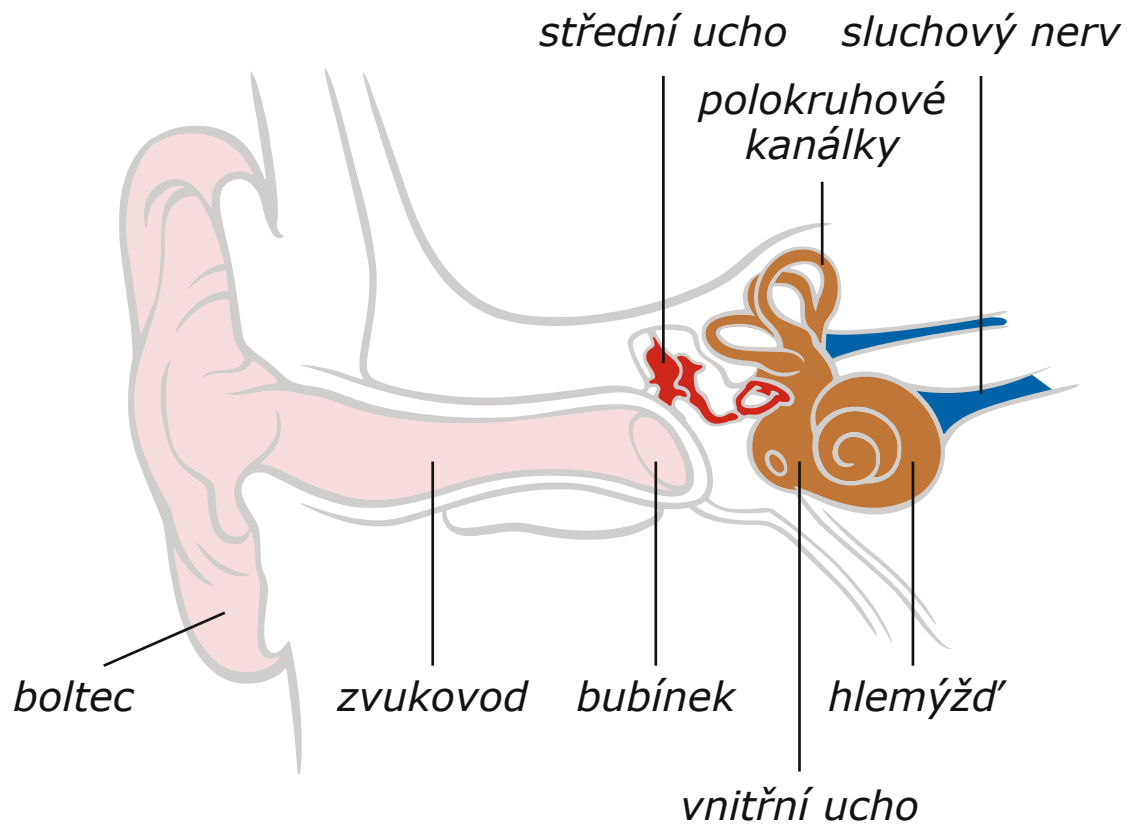
DYNAMICKÝ ROZSAH

- odstup „nejtiššího“ od „nehlasitějšího“ zaznamenaného (reprodukovatelného) signálu
- „nejtišší“ signál omezen úrovní šumu
⇒ dynamický rozsah signálu:
odstup mezi „nejhlasitějším“ signálem a šumem

(ZVUKOVÝ) SIGNÁL

- je třeba zaznamenat akustický signál $p_{AK}(x_0, t) = p_{AK}(t)$ v konkrétním bodu prostoru x_0
- „kvalitní“ záznam různě rychlé změny $p_{AK}(t)$
⇒ záznam frekvenčního rozsahu f_{min} až f_{max}
- „kvalitní“ záznam hodnot $p_{AK}(t)$ v rozsahu $p_{AK min}$ až $p_{AK max}$
⇒ „velký“ dynamický rozsah
- záznam $p'_{AK}(t)$ se liší jen minimálně od $p_{AK}(t)$
⇒ „velký“ odstup signál–šum
- význam „kvalitní“, „velký“ apod. musí být odvozen z vlastností lidského sluchu

UCHO



orig. kresba L. Chittka, A. Brockmann, licence CC 2.5

UCHO

- vnější ucho
 - boltec: rozlišení lokace zvuku „před“ a „za“
 - zvukovod: rezonátor, zlepšení citlivosti při 3–4 kHz
- bubínek
 - kmitá vlivem změny akustického tlaku
 - oddělení tlaků vnějšího a středního ucha
- střední ucho – sluchové kůstky
 - přenos kmitání bubínku do vnitřního ucha
 - zesilovač kmitání
 - okolní svaly mohou stupeň zesílení měnit

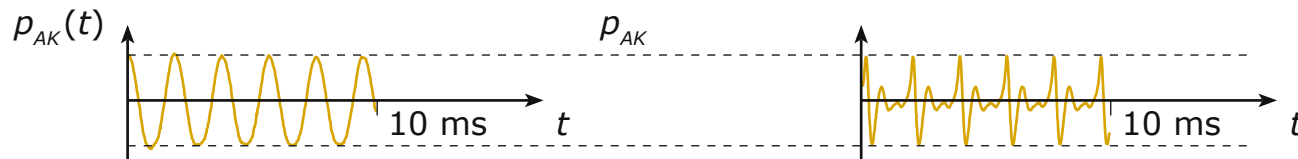
UCHO

- vnitřní ucho
 - polokruhové kanálky – pro systém držení rovnováhy (pro sluch nepodstatné)
 - hlemýžď (cochlea) s membránami – vlastní detekce zvuku
 - při dopadu zvuku se rozkmitá bazilární membrána
 - ⇒ v jistém místě rozkmit (rezonance) největší
 - ⇒ tam zareagují nervová zakončení
 - ⇒ frekvenční analýza zvuku
- sluchový (kochleární) nerv
 - přenos zvukové informace do mozku
 - při dalším zpracování zvukové informace pracuje mozek s informacemi z obou uší (binaurální sluch)

HLASITOST ZVUKU

OBJEKTIVNÍ KRITÉRIA

- hlasitost závisí na efektivní hodnotě akustického tlaku p_{AK-E}
 - energie (výkon) zvukové vlny nezávisí na amplitudě (maximální hodnotě) akustického tlaku p_{AK} , ale na jeho efektivní hodnotě p_{AK-E}
 - příklad: tyto dva průběhy určitě nevykonávají stejnou práci, ačkoliv mají stejnou amplitudu



- pro sinový průběh platí $p_{AK-E} = 0,707 p_{AK}$

- obecně pro zvuk délky T :
$$p_{AK-E} = \left(\frac{1}{T} \int_0^T (p_{AK}(t))^2 dt \right)^{1/2}$$

HLASITOST ZVUKU

- minimální slyšitelný p_{AK-E} : $p_0 = 2 \times 10^{-5} \text{ Pa}$
(ideální podmínky, stáří posluchače 15 let, frekvence cca 2 kHz, ...)
- při $p_{AK-E} \approx 50 \text{ Pa}$ vnímáme chvění vzduchu mimosluchově
 - velká hlasitost, práh bolestivosti
- vjem hlasitosti téměř logaritmický
 - úroveň akustického tlaku (sound pressure level) je vhodnější uvádět v dB
 - výkon (intenzita) I zvukové vlny $\approx konst. \times (p_{AK-E})^2$

$$\Rightarrow \text{SPL [dB]} = 10 \log_{10} \frac{konst. \times (p_{AK-E})^2}{konst. \times (p_0)^2} = 20 \log_{10} \frac{p_{AK-E}}{p_0}$$

HLASITOST ZVUKU

- rozsah SPL cca 120 dB
- extrémní SPL v praxi nepoužitelné \Rightarrow užitečný rozsah 100 dB
- rozlišitelnost jednoduchých zvuků cca 1 dB
rozlišitelnost složitých zvuku cca 3 dB

	SPL [dB]	p_{AK} [Pa]
práh slyšitelnosti (1 kHz)	0	$2,0 \cdot 10^{-5}$
tikot hodin (1 m)	30	$6,3 \cdot 10^{-4}$
běžná řeč (1 m)	50	$6,3 \cdot 10^{-3}$
hluk na ulici	80	$2,0 \cdot 10^{-1}$
hlasitý zpěv (1 m)	100	2,0
práh bolesti	130	63,0
start tryskáče (1 m, změřený p_{AK})	190	20000,0

HLASITOST ZVUKU

- intenzita zvuku klesá se vzdáleností
 - ve vzdálenosti x intenzita zvuku I
 - ve vzdálenosti kx intenzita zvuku I / k^2

⇒ rozdíl SPL = $10 \log_{10} \frac{I / k^2}{I} = -20 \log_{10} k$

⇒ ve dvojnásobné vzdálenosti ($k = 2$) **rozdíl SPL: -6,02 dB**
v poloviční vzdálenosti ($k = 0,5$) **rozdíl SPL: +6,02 dB**

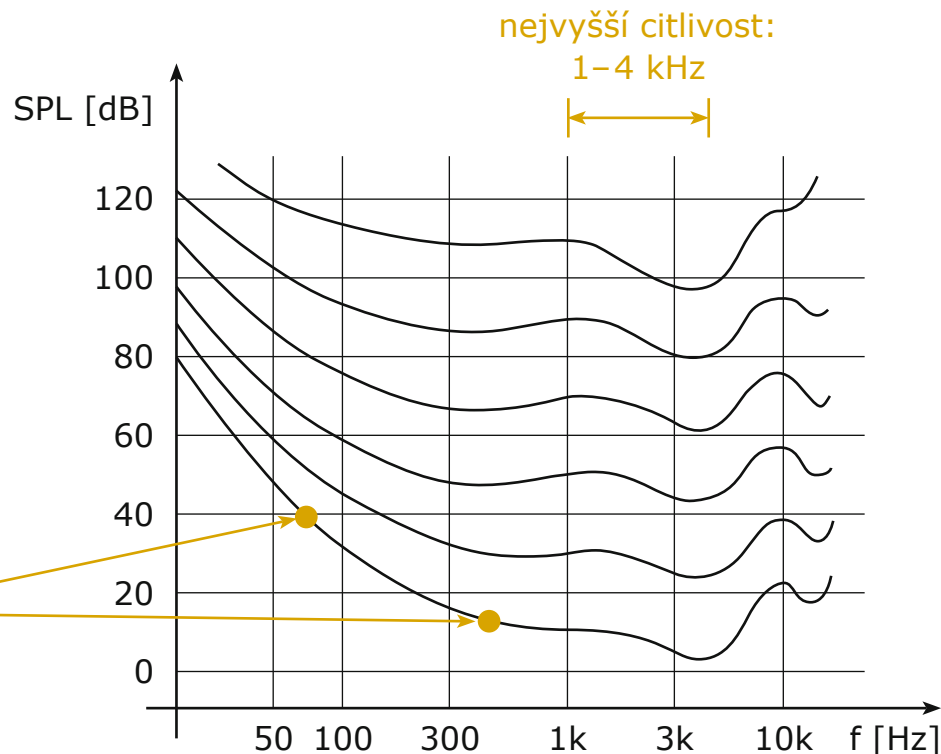
- v praxi může být pokles nižší (odrazy od stěn, ...)
- nekorelované zvuky intenzit I_1, I_2

⇒ celková intenzita $I = I_1 + I_2$
- nekorelované zvuky úrovní SPL₁, SPL₂ [dB]
 - $I_1 = konst. \times 10^{0,1 \text{ SPL}_1}$ $I_2 = konst. \times 10^{0,1 \text{ SPL}_2}$
 - ⇒ celková úroveň SPL = $10 \log_{10} (10^{0,1 \text{ SPL}_1} + 10^{0,1 \text{ SPL}_2})$

HLASITOST ZVUKU

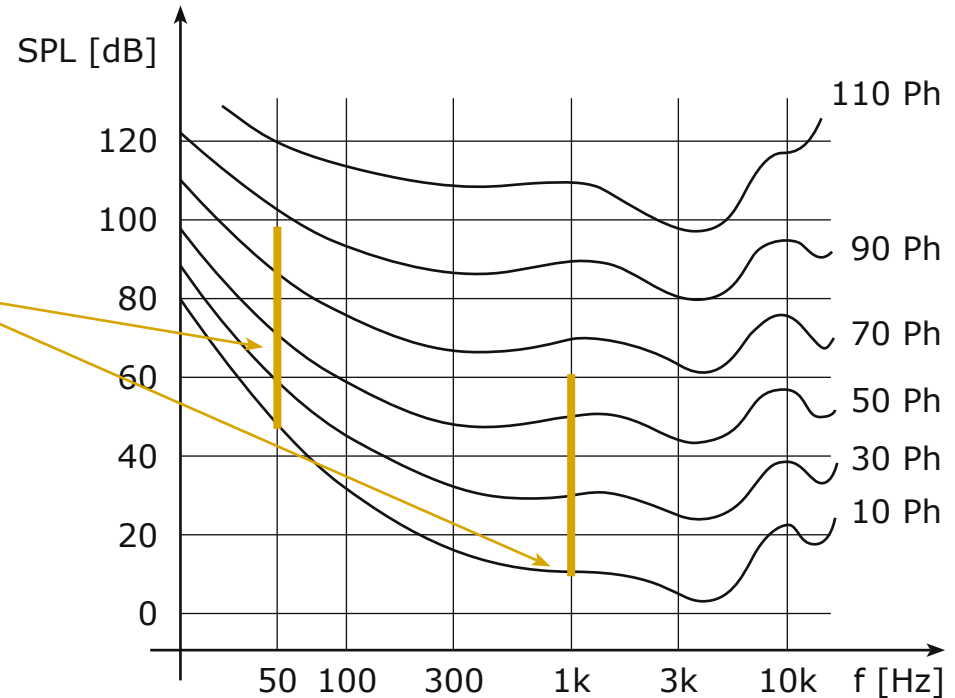
SUBJEKTIVNÍ KRITÉRIA

- 2× větší hlasitost \Rightarrow rozdíl SPL 10 dB (někdy, přibližně)
- vjem hlasitosti závisí na frekvenci
- měření závislosti Fletcher a Munson (1933), Robinson a Dadson (1956), ISO 226
- křivky konstantní hlasitosti: například tyto dva zvuky se zdají stejně hlasité



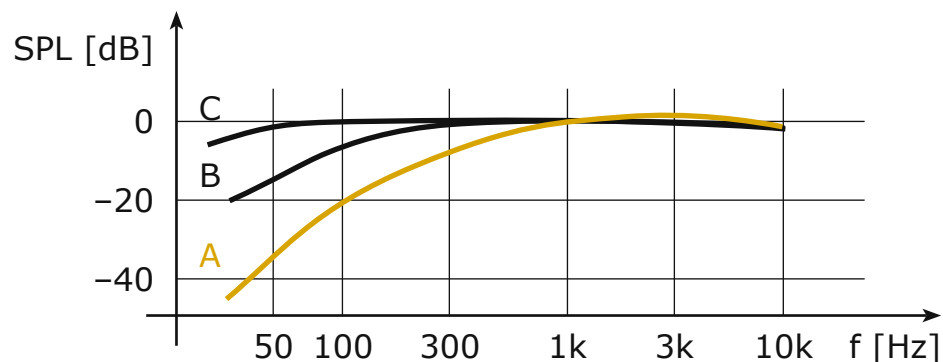
HLASITOST ZVUKU

- jednotka stejné vnímané hlasitosti: fon (Ph)
 - odpovídá SPL [dB] pro 1 kHz
- pro jiné frekvence převod na dB SPL podle křivek konstantní hlasitosti
- změna SPL 50 dB:
 - 50 Hz \Rightarrow 80 ph
 - 1 kHz \Rightarrow 50 ph
 - \Rightarrow zavádí se další jednotky, které hlasitost určují lépe



HLASITOST ZVUKU

- označení „hlasitosti“ jedním číslem – kompenzace
 - A – cca 40 fonů, v podstatě ignoruje první oktávy
 - B – cca 70 fonů
 - C – cca 100 fonů
- např. 60 dB SPL při 100 Hz odpovídá 40 dBA (úrovni SPL podle kompenzace A)
- měření hlučnosti nejčastěji podle kompenzace A



HLASITOST ZVUKU

ŠPIČKOVÁ HLASITOST / ÚROVEŇ

- pro krátké zvuky ($\ll 1$ s)
- detektor: peak power meter (PPM)
- důležité v technických aplikacích – signál nesmí nikdy překročit maximální povolenou úroveň
- pro technické účely je vhodné definovat úroveň jinak

– úroveň signálu výkonu P : $\text{dBm} = 10 \log_{10} \frac{P}{1 \text{ mW}}$
(zavedeno pro první telefony)

– úroveň signálu napětí U : $\text{dBu} = 20 \log_{10} \frac{U}{0,775 \text{ V}}$
(hifi, odvozeno z historické telefonie)

– při max. povoleném napětí U_{max} : $\text{dBFS} = 20 \log_{10} \frac{U}{U_{\text{max}}}$
(FS = full scale; **musí platit dBFS < 0**)

HLASITOST ZVUKU

PRŮMĚRNÁ HLASITOST / ÚROVEŇ

- lépe odpovídá pojmům „tichý“ vs „hlasitý“
- klasický detektor: volume unit meter (VU meter)
 - v podstatě „velmi líný“ PPM
 - býval kalibrován, aby $0 \text{ dBVU} = +4 \text{ dBu}$
 - rozsah vhodný pro měření hudby -20 až $+3 \text{ dBVU}$ (kladné hodnoty reprezentují „velmi hlasité“)
 - měl by brát zřetel na frekvenční citlivost ucha
- při sledování hlasitosti zvuku (např. TV) důležité špičky (např. výstřel) i průměr (např. hlas)
- norma definující „číselnou hodnotu hlasitosti“ např. ITU-R BS.1770

VÝŠKA ZVUKU

OBJEKTIVNÍ KRITÉRIA

- nízké frekvence – hluboké tóny
vysoké frekvence – vysoké tóny
- při $f < 20$ Hz nevnímáme spojitý tón, ale oddělené pulsy
- při $f > 20$ kHz se při rozumných SPL zvuk nevnímá
 - horní hranice klesá s věkem (cca 1 Hz / den)
 - ve věku kolem 70 let cca 2–4 kHz
- standardní frekvenční pásmo slyšitelnosti: 20 Hz – 20 kHz
- logaritmické vnímání
 - 1 oktáva = rozsah od f do $2f$
 - standardní rozsah slyšitelnosti 10 oktáv
- dvě frekvence < 2 kHz rozlišitelné, pokud se liší o cca 1 Hz
 - při 1 kHz to odpovídá cca 0,1 %; u vyšších frekvencí horší

VÝŠKA ZVUKU

komorní A 440 Hz (jedna z definic)

hlas

bas 82 – 329 Hz (2 oktávy) (základní frekvence)

baryton 110 – 370 Hz (2 oktávy)

tenor 131 – 523 Hz (2 oktávy)

alt 175 – 698 Hz (2 oktávy)

mezzosoprán 247 – 880 Hz (2 oktávy)

soprán 262 – 1397 Hz (2,5 oktávy)

housle 196 – 1975 Hz (3,3 oktávy)

flétna 262 – 2349 Hz (3,1 oktávy)

klavír 27 – 4186 Hz (7,3 oktávy)

varhany 16 – 8372 Hz (9,0 oktávy)

VÝŠKA ZVUKU

SUBJEKTIVNÍ KRITÉRIA

- typický zvuk se skládá ze základní frekvence f a harmonických frekvencí $n \times f$ (celočíselné násobky)
- rozeznávání výšky tónu podle základní frekvence a struktury harmonických frekvencí
 - např. vnímané komorní A může obsahovat jen minimum energie pro 440 Hz!
 - pokud reproduktor zvládne zvuky jen např. od 70 Hz a přehrává zvuk 100 Hz + 150 Hz + 200 Hz + 250 Hz + ..., uslyšíme tón výšky 50 Hz
 - ve zvuku houslí často chybí základní frekvence, ...
- rozpoznání cca po 30 ms od začátku zvuku
- detekuje se místy rezonance membrány vnitřního ucha

VÝŠKA ZVUKU

KRITICKÉ PÁSMO

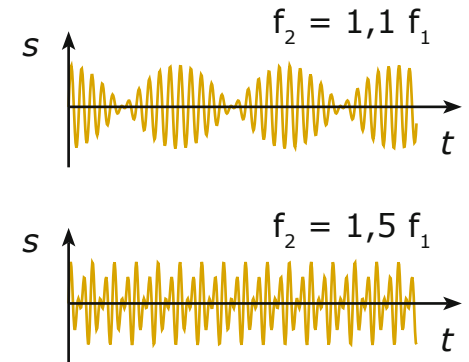
- rozsah kmitání bazilární membrány pro danou frekvenci
- šířka cca 100 Hz pro frekvenci do 500 Hz,
cca 2 kHz pro frekvenci okolo 10 kHz
- rozsah sluchu odpovídá 24 nepřekrývajícím se kritickým pásmům

- z existence kritických pásem plyne jev „maskování“
- velmi důležitý pojem ve ztrátové kompresi zvuku

SOUZVUK

- dva současně znějící tóny frekvencí f_1, f_2 :

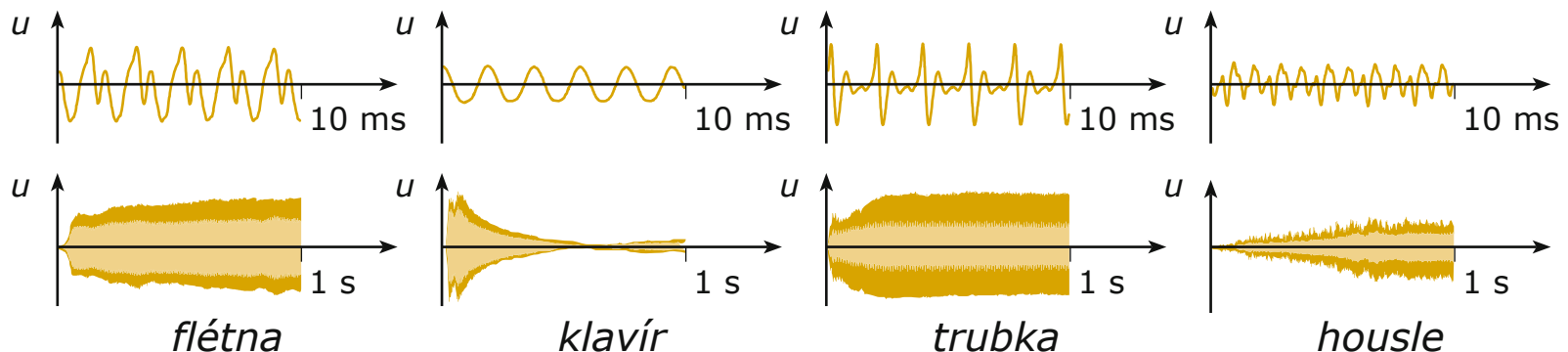
$$s(t) = \sin(2\pi f_1 t) + \sin(2\pi f_2 t) = \\ = 2 \sin\left(2\pi \frac{f_1 + f_2}{2} t\right) \cos\left(2\pi \frac{f_1 - f_2}{2} t\right)$$



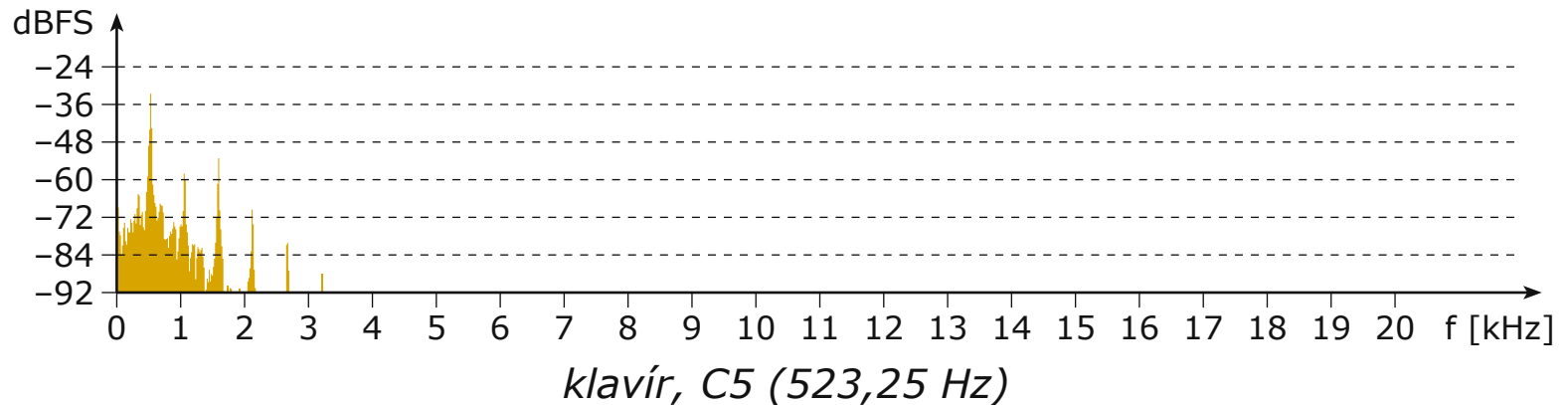
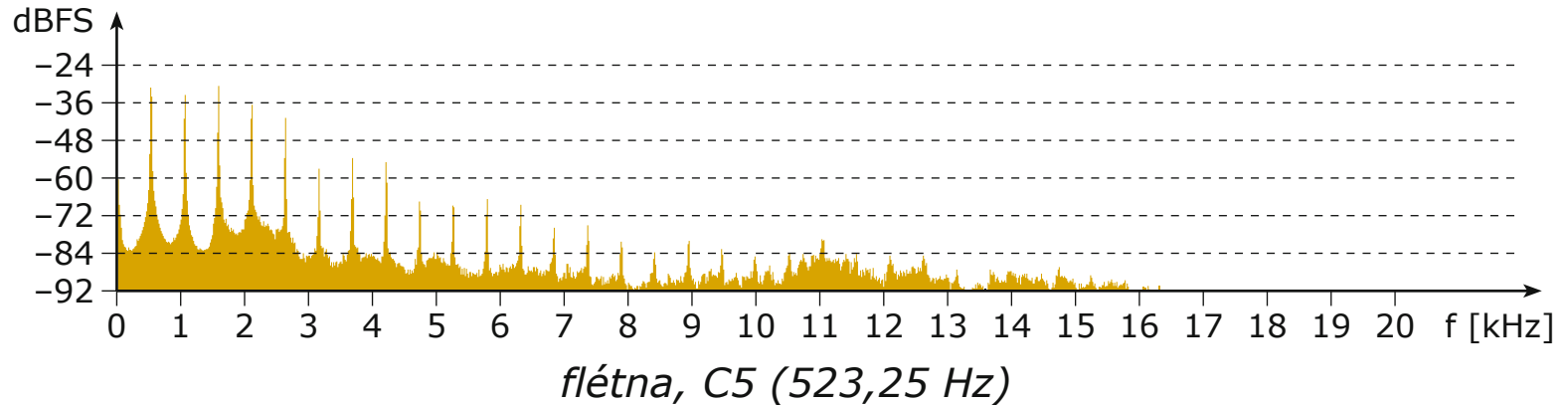
- f_1, f_2 patří do stejného kritického pásma:
 - jsou-li f_1, f_2 blízké ($f_1 - f_2 < 10$ Hz)
vjem tónu průměrné frekvence a **kolísání hlasitosti**
 - jsou-li f_1, f_2 vzdálené ($f_1 - f_2 < \text{šířka kritického pásma}$)
vjem „hrubého“ tónu
- f_1, f_2 nepatří do stejného kritického pásma:
vjem souzvuku dvou odlišných tónů

BARVA ZVUKU (TÉMBR)

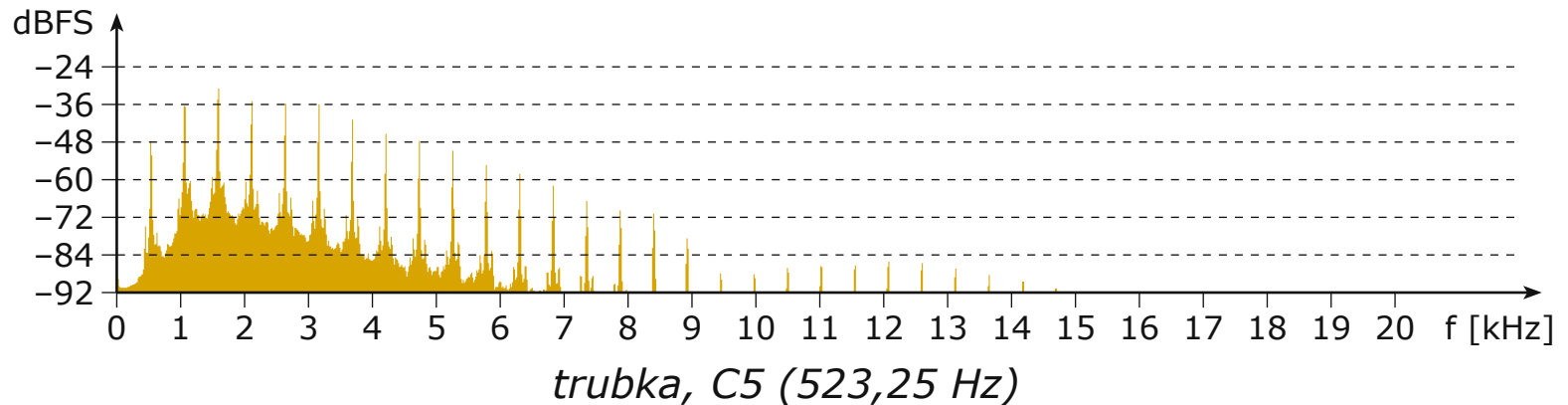
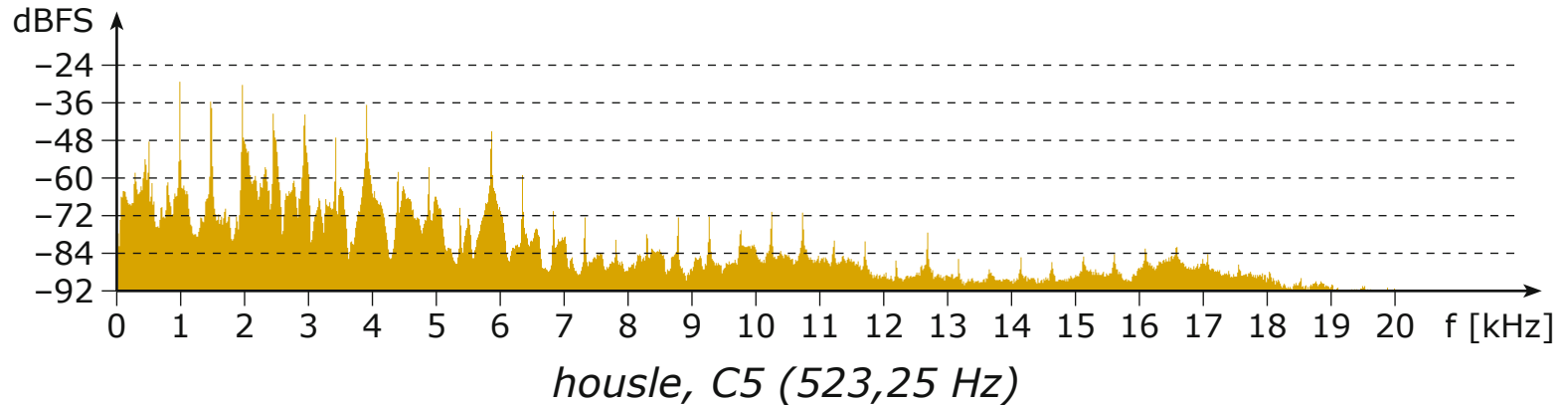
- subjektivní pojem; např. odlišení zvuku hudebních nástrojů
- dána především
 - poměrem energií (prvních) harmonických frekvencí (příliš nezávisí na jejich vzájemné fázi)
 - počátečním přechodovým (tranzientním) zvukem typicky v délce několika (10) vln základní frekvence, např. pro $f_z = 500$ Hz jde o prvních 20 ms zvuku
- příklad: tón C5 ($f_z = 523,25$ Hz $\sim 5,2$ periody za 10 ms)



BARVA ZVUKU (TÉMBR)



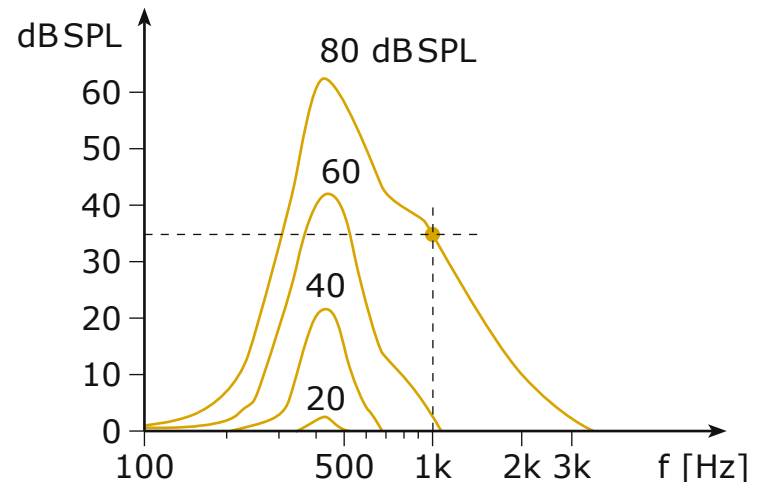
BARVA ZVUKU (TÉMBR)



MASKOVÁNÍ

FREKVENČNÍ MASKOVÁNÍ

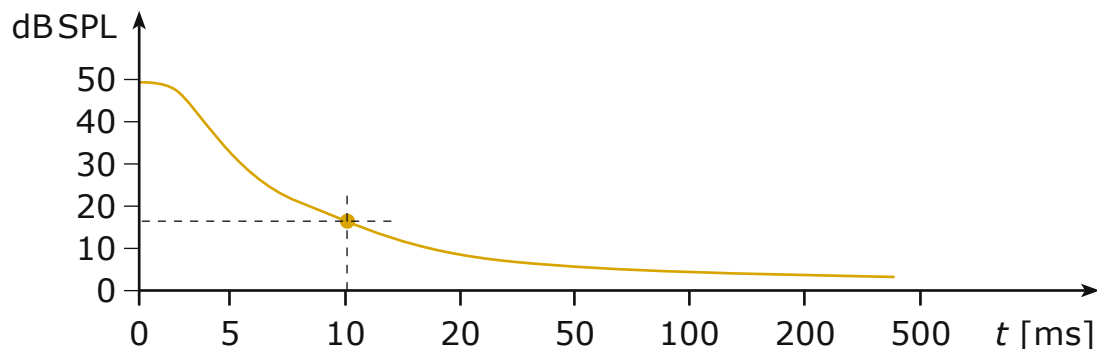
- znějí-li současně dva nekorelované zvuky, slabší může být zamaskován (není slyšet)
- závisí na rozdílu frekvencí a rozdílu úrovní
- příklad: maskovací zvuk šum, frekvence 365–455 Hz, úroveň určená grafem;
tón 1 kHz úrovně 35 dB SPL (a slabší) je maskován šumem 80 dB SPL (Egan a Hake, 1950)
- aplikace: ztrátová komprese zvuku – co není slyšet, nemusí se kódovat



MASKOVÁNÍ

ČASOVÉ MASKOVÁNÍ

- k frekvenčnímu maskování dochází také chvíli *před* i *po* zapnutí maskovacího zvuku
- příklad: maskovací zvuk schopný frekvenčně maskovat zvuk úrovně 50 dB SPL má 10 ms po skončení schopnost zamaskovat zvuk úrovně 17 dB SPL



FREKVENČNÍ ANALÝZA

- obecný periodický signál $s(t)$ periody T_Z můžeme vyjádřit součtem mnoha sinových průběhů:

$$s(t) = a_0 + \sum_{k=1}^{\infty} a_k \sin(2\pi k f_Z t + \varphi_k)$$

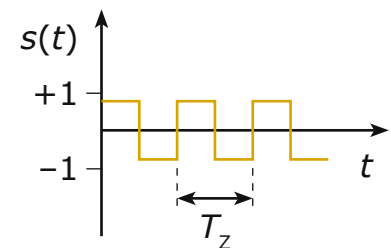
f_Z základní frekvence, $f_Z = 1 / T_Z$

a_k amplituda sinu frekvence $k f_Z$ (a_0 je „stejnoseměrný člen“)

φ_k fázový posun sinu frekvence $k f_Z$

- příklad: obdélníková vlna

$$s(t) = \sin(2\pi f_Z t) + \frac{1}{3} \sin(2\pi 3 f_Z t) + \dots + \\ + \frac{1}{2k+1} \sin(2\pi (2k+1) f_Z t) + \dots$$



FREKVENČNÍ ANALÝZA

- obecnější vyjádření pro neperiodické signály:

$$s(t) = \int_0^{\infty} a(f) \sin(2\pi f t + \varphi(f)) df$$

- alternativně po rozkladu

$$\begin{aligned} a(f) \sin(2\pi f t + \varphi(f)) &= a(f) \cos(\varphi(f)) \sin(2\pi f t) + \\ &\quad a(f) \sin(\varphi(f)) \cos(2\pi f t) \\ &= a_s(f) \sin(2\pi f t) + a_c(f) \cos(2\pi f t) \end{aligned}$$

$$s(t) = \int_0^{\infty} a_s(f) \sin(2\pi f t) + a_c(f) \cos(2\pi f t) df$$

- jak pro daný $s(t)$ zjistit $a_s(f)$ a $a_c(f)$?

FREKVENČNÍ ANALÝZA

FOURIEROVA TRANSFORMACE

- $a_s(f) = 2 \int_{-\infty}^{\infty} s(t) \sin(2\pi f t) dt$
- $a_c(f) = 2 \int_{-\infty}^{\infty} s(t) \cos(2\pi f t) dt$
- proč to funguje?

$$\int_{-\infty}^{\infty} \sin(2\pi f_1 t) \cos(2\pi f_2 t) dt = 0$$

$$\int_{-\infty}^{\infty} \sin(2\pi f_1 t) \sin(2\pi f_2 t) dt = \delta(f_1 - f_2) = \begin{cases} 0 & f_1 \neq f_2 \\ \infty & f_1 = f_2 \end{cases}$$

FREKVENČNÍ ANALÝZA

- celková amplituda $a(f)$ průběhu frekvence f , $f > 0$:

$$a(f) = \sqrt{a_s(f)^2 + a_c(f)^2}$$

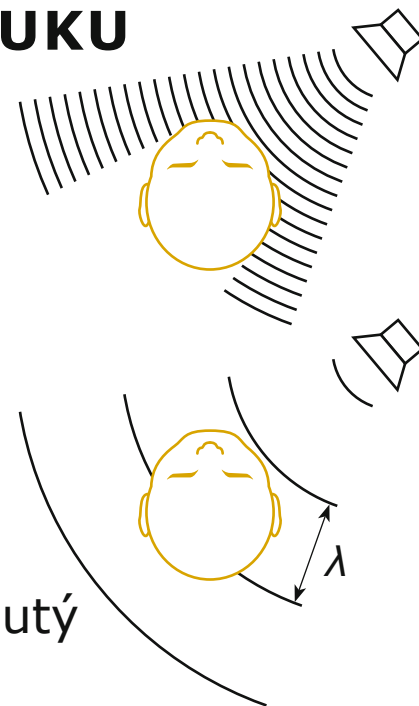
- funkci $a(f)$ říkáme *spektrum* signálu $s(t)$
- výkon průběhu frekvence $f \sim a(f)^2$

- Fourierova transformace se častěji počítá v komplexním oboru
- potom se uvažují frekvence f kladné i záporné, amplitudy $a(f)$, $a(-f)$ na sobě (jednoduše) závisí
- podrobnosti později

IDENTIFIKACE ZDROJE

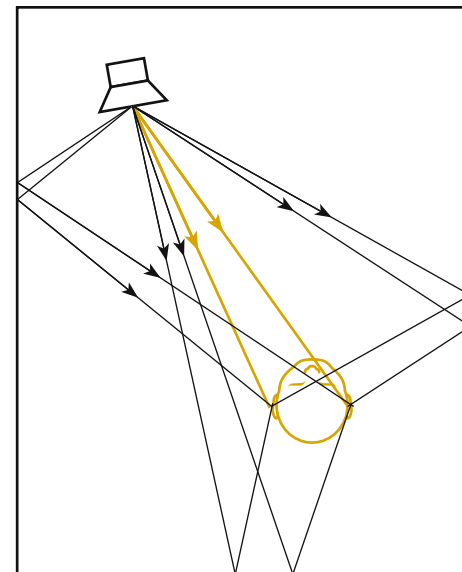
IDENTIFIKACE SMĚRU ZDROJE ZVUKU

- dobrá, rozlišovací schopnost asi 1°
- rozeznání směru cca po 1 ms
- vysoké frekvence
 - hlava vrhá akustický stín
 - λ musí být menší než hlava, tj. pro $f > 2$ kHz
 - uši vnímají různou intenzitu vysokých frekvencí \Rightarrow změna barvy zvuku
- nízké frekvence – zvuk u uší fázově posunutý
 - λ musí být větší než hlava, jinak je posuv příliš složitý
- impulsní zvuky – uši vnímají impuls v různých okamžicích



IDENTIFIKACE ZDROJE

- Haasův efekt
 - pokud uši slyší tentýž zvuk a časový rozdíl je mezi 0,7 ms a 50 ms, lokalizujeme jej podle „prvního dojmu“
⇒ dozvuk místnosti neovlivňuje lokalizaci
 - navíc vnímáme zvuk jako celek
 - při prodlevě > 50 ms vnímáme druhý zvuk jako ozvěnu



IDENTIFIKACE ZDROJE

IDENTIFIKACE VZDÁLENOSTI ZDROJE ZVUKU

- špatná
- ve větší vzdálenosti menší obsah vysokých frekvencí (vliv útlumu)
- relativně dobrý odhad velikosti prostředí
 - Haasův efekt

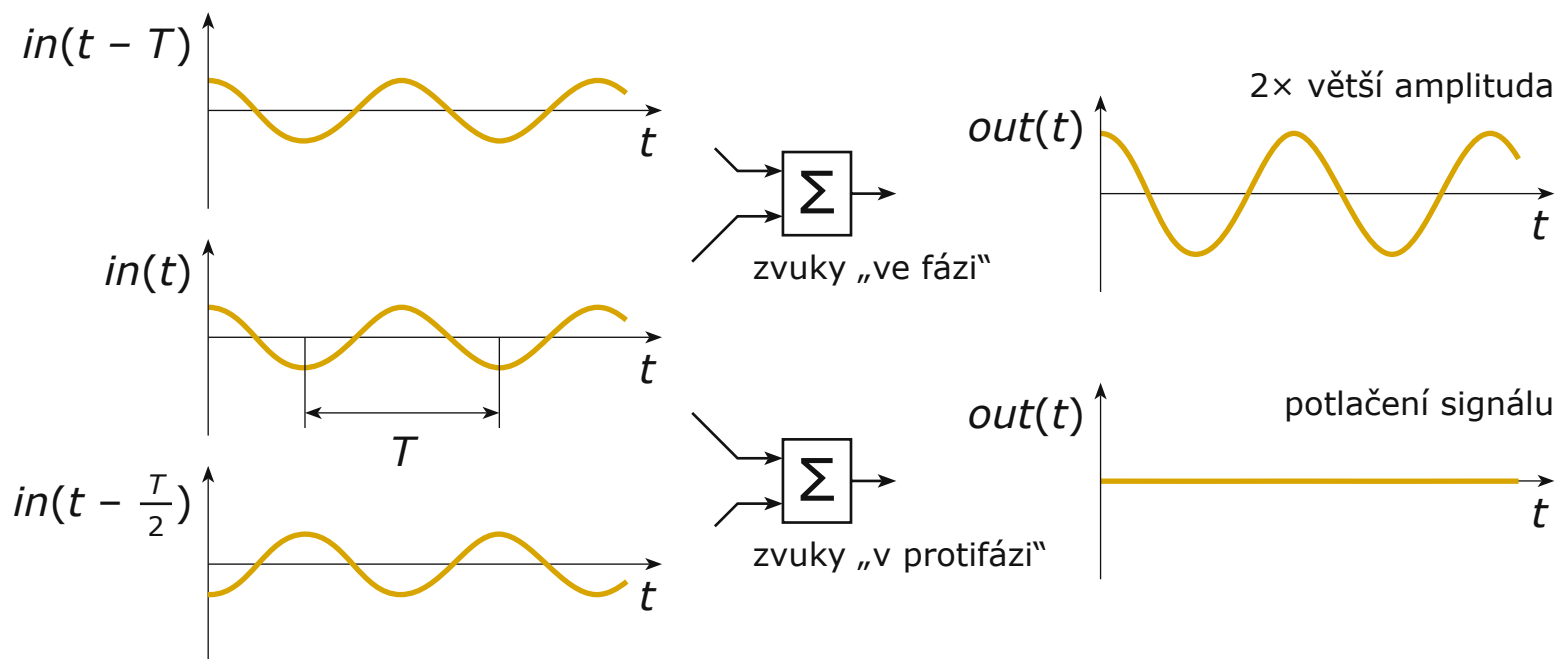
IDENTIFIKACE ZDROJE

- odlišné chování, vnímají-li uši tentýž zvuk časově (fázově) posunutý, nebo zvuky výrazně odlišnější
- efekt koktejlového večírku – jsme schopni soustředit se na konverzaci, i když je zvuk v podstatě maskovaný ruchem okolí
 - schopnost se ztrácí při poslechu záznamu



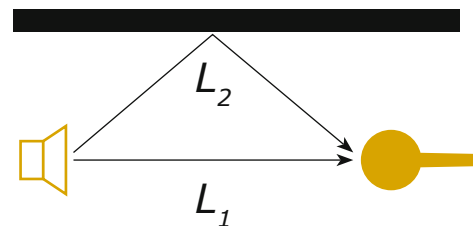
HŘEBENOVÝ FILTR

- obecné pojmenování zpracování signálu:
součet signálu se svou časově zpožděnou kopií
 $out(t) = in(t) + in(t - \Delta t)$

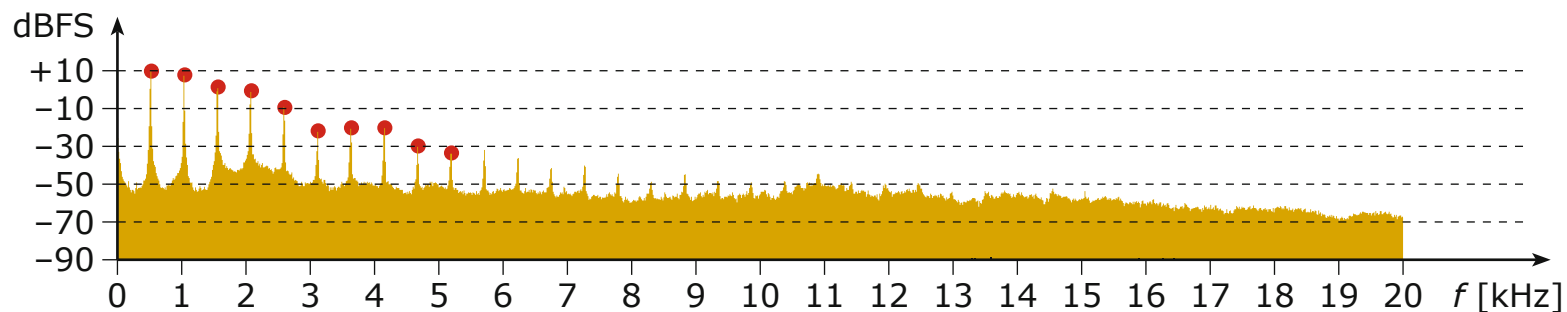


HŘEBENOVÝ FILTR

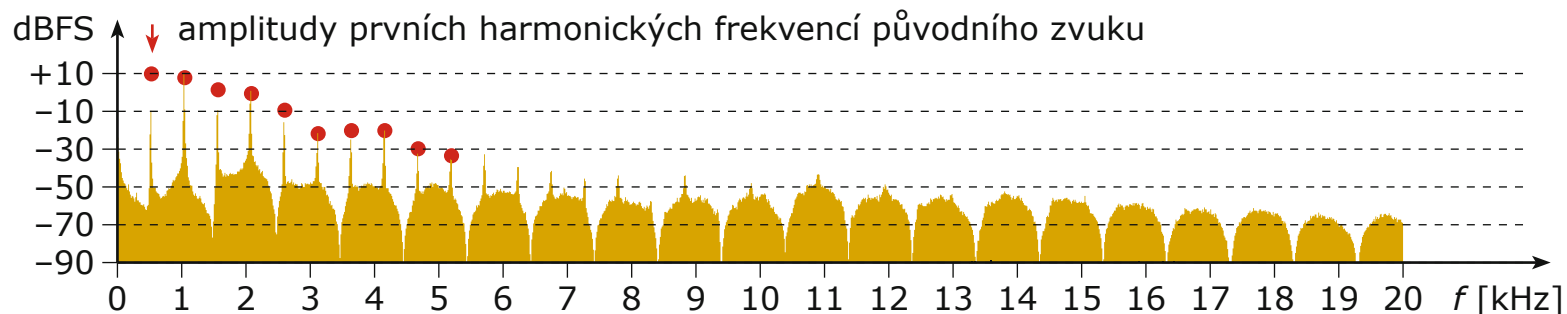
- ve zpracování zvuku *efekt hřebenového filtru*
 - např. zvuk je nahrán společně se svou zpožděnou kopií
 - typicky přímý zvuk + odraz od stěny
 - příklad: $L_1 = 3$ ft, $L_2 = 4$ ft
rozdíl zpoždění
zvuků 1 a 2 cca 1 ms
 - ⇒ pro frekvenci 1 kHz ($T = 1$ ms)
2× vyšší akustický tlak (+6 dB)
 - ⇒ pro frekvenci 500 Hz ($T = 2$ ms)
akustický tlak 0
- některé frekvence zdůrazněny, jiné potlačeny
 - změna barvy zvuku
- typický problém při mixu stereo → mono



HŘEBENOVÝ FILTR



flétna, C5 (523,25 Hz), původní zvuk



flétna, C5 (523,25 Hz), součet se zvukem zpožděným o cca 0,5/523 s

NAHRÁVACÍ ŘETĚZEC

- začátek řetězce: zdroj zvuku
konec řetězce: posluchač reprodukováného zvuku
 - Chceme dokonalou kopii původního zvuku?
 - včetně pokašlávání hudebníků?
 - včetně zvuku letadla, které náhodou prolétlo kolem?
 - včetně ruchu klimatizace v místnosti?
 - ...
 - zvuk nástroje je jiný ze vzdáleností 10 cm a 10 m;
jak má znít na nahrávce?
- ⇒ **obvykle chceme uvěřitelnou kopii zvuku**

NAHRÁVACÍ ŘETĚZEC

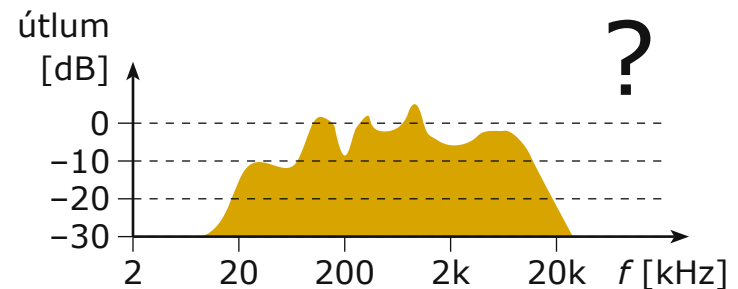
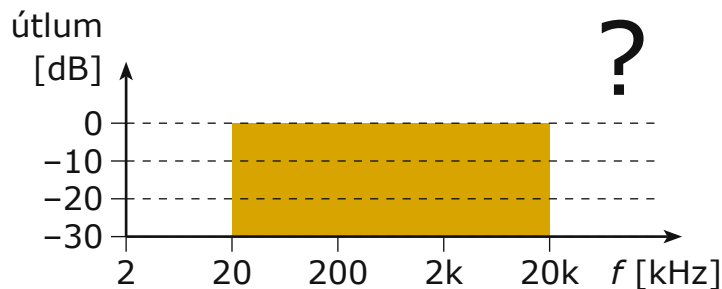
- zdroj zvuku (hlas, hudební nástroj, prostředí, ...)
 - ne každý hlas (apod.) zní na nahrávce dobře
 - důležitá volba akustického prostředí
- mikrofon / mikrofony
 - důležitá volba typu, umístění, směřování, ...
- zpracování záznamu jednoho zdroje zvuku
 - odstranění šumu, potlačení sykavek, ...
- mixáž několika zdrojů do jednoho složeného zvuku
 - volba poměrů intenzit, tvorba prostorovosti, ...
- uložení zvuku do distribuční podoby

- dekódování distribuční podoby
- reprodukce
 - velká variabilita – nahrávka má znít dobře všude

REPRODUKTORY

POSLECHOVÉ REPRODUKTORY

- musí dobře znít (značně subjektivní vlastnost)
- velká variabilita typů – hifi, výkonové koncertní, do automobilů, do telefonů, ...
- vyjma high-end mají jejich uváděné parametry stěží jakoukoliv informační hodnotu
- příklad: frekvenční rozsah 20 – 20 000 Hz
Co to znamená? Jak vypadá frekvenční charakteristika?



REPRODUKTORY

STUDIOVÉ REPRODUKTORY

- plochá frekvenční charakteristika
 - např. 60 – 20000 Hz tolerance ± 3 dB; 35 Hz pokles -6 dB
- vyrovnaná fázová charakteristika
 - všechny frekvence se zpracují se stejným zpožděním
 - ⇒ na vstupu obdélníkový průběh → na výstupu také
- rovnoměrná směrová charakteristika
 - zvuk je ve všech směrech stejně kvalitní
- velká dynamika
 - umí bez zkreslení přejít od „tichého“ k „hlasitému“
- malé harmonické zkreslení ($< 1\%$)
 - reprodukce sinového průběhu není dokonale sinová
 - zkreslení je součet harmonických frekvencí – omezit!

REPRODUKTORY

KONSTRUKCE

- reprobox obsahuje jeden reproduktor
 - minimální konstrukční problémy
 - nehodí se pro větší akustické výkony
- reprobox obsahuje více reproduktorů
 - typicky tweeter (vysoké tóny), woofer (hluboké tóny)
 - kvalitní rozdělení zvuku mezi reproduktory obtížné
- subwoofer
 - reproduktor specializovaný na velmi hluboké tóny (cca do 100 Hz)
 - do reproduktorové soustavy obvykle stačí jeden
 - ⇒ cenově výhodné (ostatní reproboxy nemusí basy umět)
 - ⇒ další konstrukční problémy při dělení zvuku

REPRODUKCE

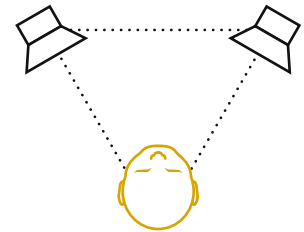
MONOFONNÍ 1.0

- jeden zdroj zvuku
- typický výstup mikrofону
- množství starých nahrávek, zvukových stop filmů
- předpokládaný způsob reprodukce u rádia, mobilních zařízení



STEREOFONNÍ 2.0

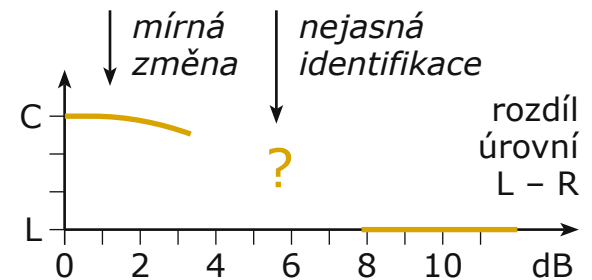
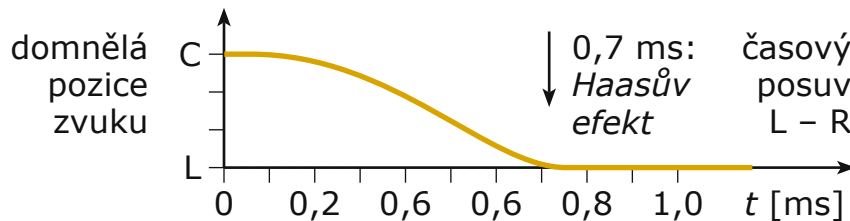
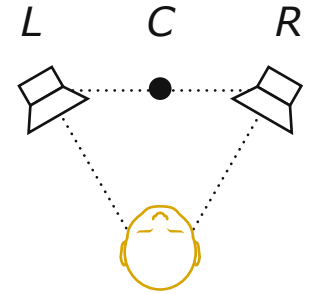
- dva zdroje zvuku (L, R)
- stereofonní obraz – virtuální zdroje zvuku
- posluchač a zdroje – rovnostranný trojúhelník, ve stejné výšce
- sluchátka – rychlejší únava, zvuk před posluchačem často vnímán jako uvnitř hlavy



REPRODUKCE

STEREOFONNÍ OBRAZ

- L+R stejný zvuk
 - vjem, jako by byl zdroj zvuku v bodu C
- L+R stejný zvuk jiné hlasitosti
 - nejasný posun virtuálního zdroje (vyjma extrémů, viz graf dole)
- L+R stejný, časově posunutý zvuk
 - přesný posun virtuálního zdroje zvuku
 - potíže s kompatibilitou monofonní reprodukce (efekt hřebenového filtru)



REPRODUKCE

LOW FREQUENCY EFFECT (LFE)

- hluboké tóny málo slyšitelné
 - ⇒ musí mít velkou amplitudu (v nahrávce např. 1,0)
 - ⇒ užitečnější zvuky mají amplitudu nižší (např. 0,1)
 - ⇒ amplituda šumu záznamu je pořád stejná (např. 0,001)
 - ⇒ užitečnější zvuky mají horší odstup signál–šum
- řešení – nízké tóny ve zvláštní mono stopě (LFE), zvláštní řízení hlasitosti
- studiový zvuk → rozdělení na „normální“ kanály a LFE
- při reprodukci dekódování „normálních“ kanálů a LFE → sloučení do zvuku obsahujícího basy i výšky → *případné* rozdělení zvuku do běžných reproduktorů a subwooferu (tzv. *bass management*)

REPRODUKCE

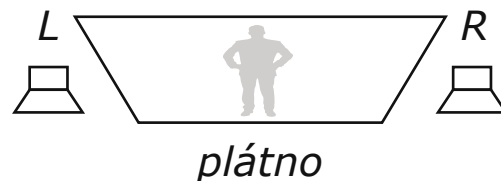
2.1

- je třeba rozlišovat, jestli se notace 2.1 vztahuje ke způsobu záznamu zvuku, nebo k reproduktorové konfiguraci
- 2.1 zvuk (L+R+LFE) můžeme přehrávat na libovolné konfiguraci (mono / stereo / sluchátka / s nebo bez subwooferu / ...)
- 2.1 konfiguraci reproduktorů (L+R+subwoofer) můžeme použít k přehrávání libovolného zvuku (1.0, 2.0, 5.1, ...)
 - vyžaduje bass management
 - často vyžaduje umělé zvýšení / snížení počtu kanálů (upmixing / downmixing)
- umístění subwooferu prakticky libovolné
 - umíme špatně identifikovat směr zdroje hlubokých tónů

REPRODUKCE

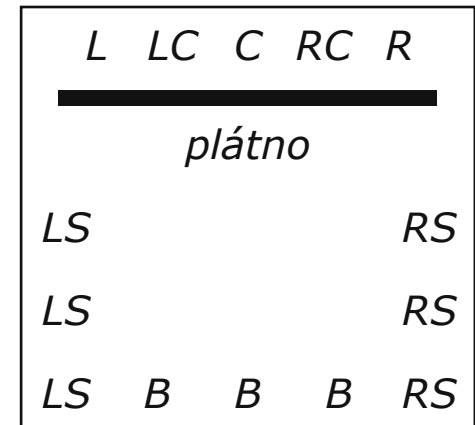
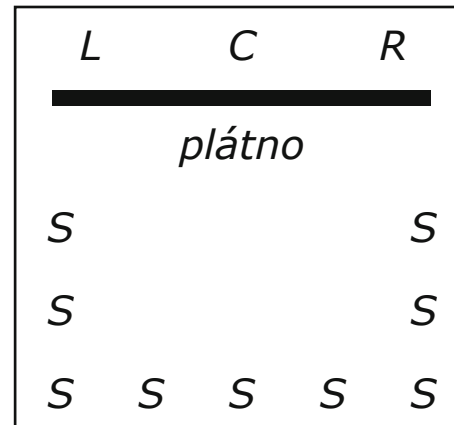
VÍCEKANÁLOVÁ REPRODUKCE

- původně určena k ozvučení kinosálů
 - reproduktory jen u plátna
 - zvuk pro diváky vpředu hlasitý, pro diváky vzadu tichý
 - ⇒ více reproduktorů po stěnách kinosálu
 - ⇒ když už jsou k dispozici, proč jim nepřiradit vlastní stopy?
-
- problém stereofonní reprodukce
 - L+R reproduktory stejný zvuk
 - A vnímá zdroj zvuku v místě, kde je figura na plátně
 - Haasův efekt: B vnímá zdroj zvuku u reproduktoru R – mimo figuru



REPRODUKCE

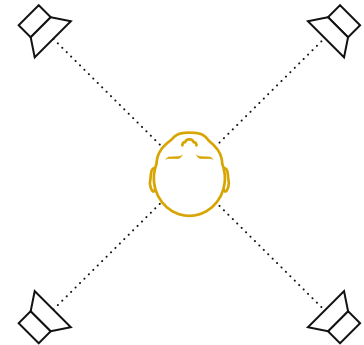
- řešení – středový kanál
 - stereofonní obraz funkční ve většině promítacího sálu
- typická soustava ozvučení kinosálu
 - vpředu center (C), left (L), right (R)
 - po stranách surround (S),
případně left surround (LS), right surround (RS)
 - případně dále
back (B),
left center (LC),
right center (RC),
...



REPRODUKCE

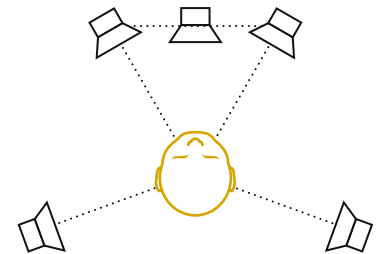
KVADROFONNÍ (4.0)

- dva alternativní systémy:
 - hudební 45°, 135°
 - filmový (reproduktory jako u 5.1, oba zadní reproduktory stejný zvuk)



SURROUND (5.1)

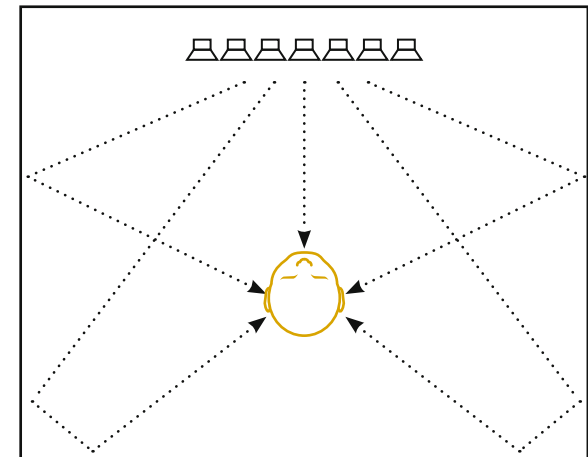
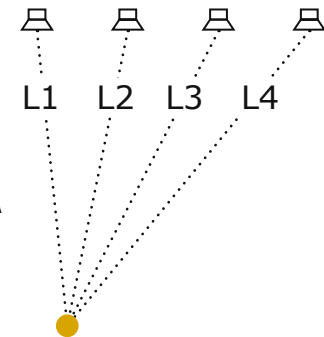
- nahrávka L, R, C (center), LS, RS (left/right surround), LFE
- rozmístění 0°, 30°, 110°
- mnoho dalších systémů (6.1, 7.1, ...)



REPRODUKCE

REPRODUKTOROVÉ POLE

- cca 40 malých reproduktorů + 2× woofer
 - v daném bodu je intenzita zvuku maximální, když se všechny dílčí zvuky setkávají „ve fázi“
 - dílčí zvuky časově posunuté, aby se kompenzovaly rozdíly vzdáleností L_1, L_2, \dots
 - je-li reproduktorů hodně, je jinde intenzita zvuku malá
- ⇒ tvorba „zvukových paprsků“
- surround efekt odrazem od stěn
 - simulace sterea, směřování zvuku k jednomu posluchači, ...



REPRODUKCE

DOWNMIXING

- počet reproduktorů menší než počet zvukových kanálů
- typicky 2.0 → 1.0 (rádio, mobilní telefon, ...)
často 5.1 → 2.0, 5.1 → Dolby Surround stereo
- maticový zápis (I_m = vstup. amplituda, O_n = výst. amp.)

$$\left. \begin{array}{l} O_1 = a_{11}I_1 + a_{12}I_2 + \dots + a_{1M}I_M \\ O_2 = a_{21}I_1 + a_{22}I_2 + \dots + a_{2M}I_M \\ \dots \\ O_N = a_{N1}I_1 + a_{N2}I_2 + \dots + a_{NM}I_M \end{array} \right\} \mathbf{O} = \mathbf{A} \mathbf{I}$$

- příklad: 5.0 → 2.0

$$L_{out} = (1 \cdot L_{in} + 0,707 \cdot C_{in} + 0 \cdot R_{in} + 1 \cdot LS_{in} + 0 \cdot RS_{in}) / 2,707$$

- C_{in} musí polovinu energie předat do L_{out} , polovinu do R_{out}
- 1/2 energie = -3 dB ⇒ amplitudu C_{in} násobit $1/\sqrt{2} = 0,707$

REPRODUKCE

- normalizační faktor 2,707 v předchozím příkladu
 - amplituda L_{out} nepřekročí povolené maximum
- při downmixingu problém s efektem hřebenového filtru
 - stereofonní obraz vzniká mj. časovými (fázovými) posuvy
 - po sloučení časově posunutých zvuků změna barvy
 - největší problém, vznikl-li výchozí zvuk upmixem

UPMIXING

- typicky 2.0 → 5.1
- výstupní kanály různé lineární kombinace vstupních kanálů
 - různě frekvenčně pozměněných
(zvuk ze zadu méně vysokých frekvencí)
 - různě časově (fázově) posunutých

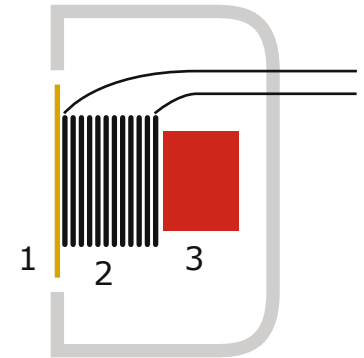
MIKROFON

- mikrofon – jedna nebo několik mikrofonních kapslí
 - mikrofonní kapsle – převod akustického tlaku na elektrické napětí
 - základní mikrofonní kapsle monofonní
 - kombinací kapslí se dosahují různé efekty
- při výběru mikrofonu musíme zvolit
 - směrovou charakteristiku mikrofonu
 - princip převodu akustické energie na elektrickou (dynamické mikrofony, kapacitní mikrofony)
 - kolik mikrofonů použít (zdroj zvuku je složitější, pro stereofonní záznam, ...)
 - typ mikrofonu (klopový, do ruky, bezdrátový, ...)

PRINCIP MIKROFONU

DYNAMICKÝ MIKROFON

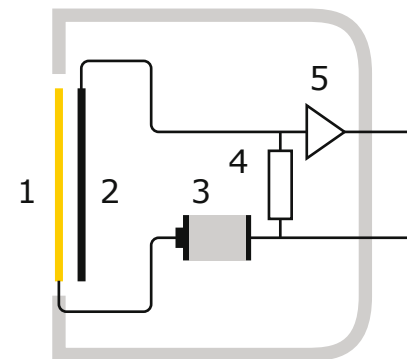
- pružná membrána (1) s cívkou (2) na zadní straně, permanentní magnet (3)
- pohyb membrány \Rightarrow pohyb cívky v magnetickém poli \Rightarrow indukce napětí
- malá membrána \Rightarrow cívka musí být malá \Rightarrow malé indukované napětí \Rightarrow náchylnost k šumu
- velká membrána \Rightarrow má velkou setrvačnost \Rightarrow špatně reaguje na vysoké frekvence \Rightarrow horší frekvenční rozsah
- typicky hůře snímá zvuk nad 10 kHz, výborný pro 1–4 kHz
- umí snímat velký akustický tlak
- nevyžaduje externí napájení
- mechanicky odolný – vhodný pro exteriér



PRINCIP MIKROFONU

KAPACITNÍ MIKROFON

- pohyblivá vodivá deska (1) a nepohyblivá vodivá deska (2) tvoří kondenzátor nabíjený např. baterií (3)
- při pohybu (1) se mění jeho kapacita
- změnu kapacity lze převést (viz odpor (4)) na malou změnu napětí ⇒ nutno zesílit (5)
- vynikající frekvenční rozsah
- výborná citlivost („hlasitější“ výstup oproti dynamickému)
- vyžaduje napájení zesilovače (baterie nebo připojení 48 V na výstupní svorky – tzv. fantomové napájení)
- velmi citlivý na zacházení, vlhkost, ...



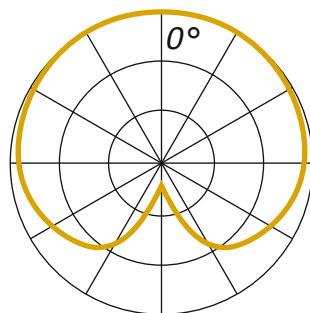
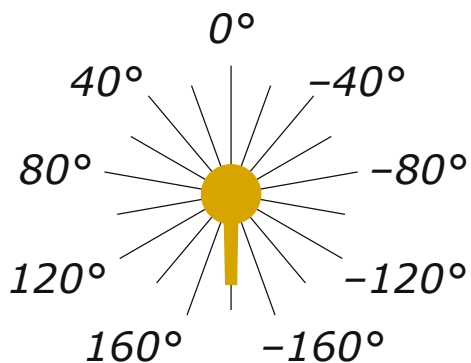
PRINCIP MIKROFONU

ELEKTRETOVÝ MIKROFON

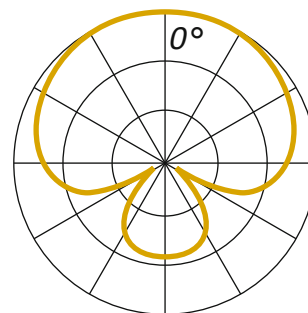
- stejný princip jako kondenzátorový
- desky kondenzátoru nabité stabilním nábojem
- levnější než původní kapacitní
- může být velmi malý a velmi levný – pro klopové mikrofony, interní mikrofony kamer apod.
- dražší typy obsahují předzesilovač signálu ⇒ vyžadují napájení

SMĚROVÁ CHARAKTERISTIKA

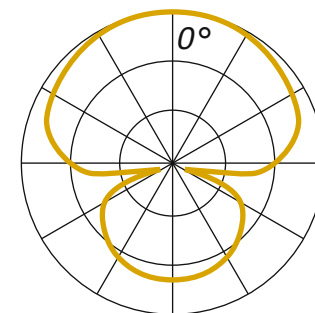
MĚŘENÍ



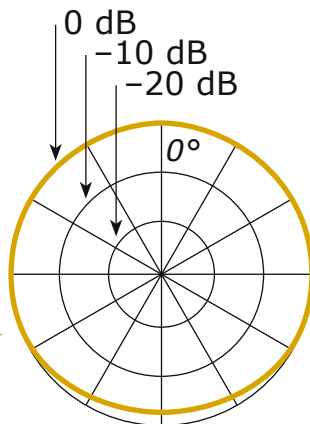
kardioidní



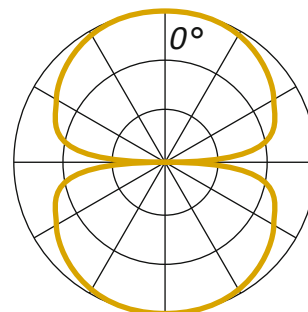
superkardioidní



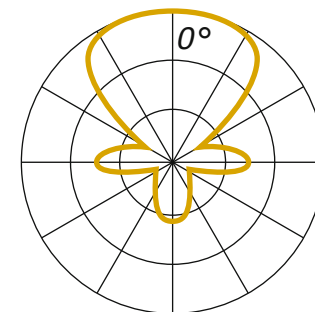
hyperkardioidní



všesměrová



osmičková



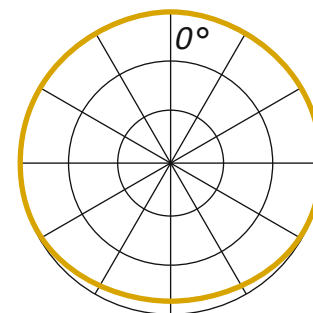
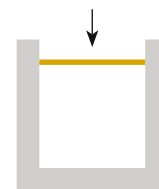
puška (lobar)

SMĚROVÁ CHARAKTERISTIKA

VŠESMĚROVÝ MIKROFON

- uzavřená mikrofonní kapsle, reaguje na akustický tlak
- často jako klopový, ale i ruční nebo studiový mikrofon
 - pro rozhovory lepší než směrový mikrofon – díky všesměrovosti nic „nepřeslechne“
- vynikající zvuk v prostředí s dozvukem
- netrpí proximity efektem
- snímá veškerý zvuk, i ten, který nechceme (ruchy z okolí)

membrána

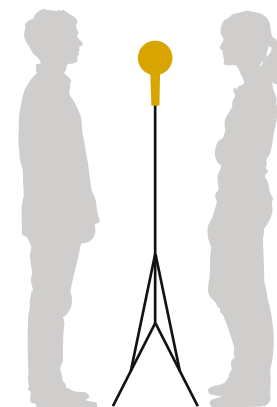
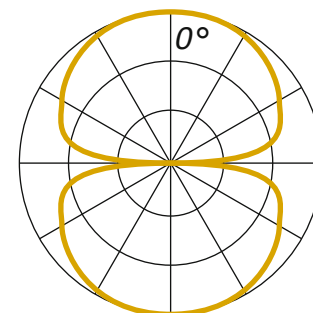
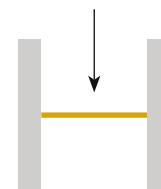


SMĚROVÁ CHARAKTERISTIKA

OSMIČKOVÝ MIKROFON

- otevřená mikrofonní kapsle, reaguje na rozdíl akustických tlaků zpředu a zezadu kapsle
- trpí proximity efektem
- zvuk přichází z úhlu 90°
 - ⇒ stejné akustické tlaky z obou stran
 - ⇒ téměř dokonalý útlum
 - využití, pokud potřebujeme silný útlum ze strany
- pokud zvuk přichází zpředu nebo zezadu, je výstup stejný, ale s opačnou polaritou
 - využití při M/S stereofonním nahrávání

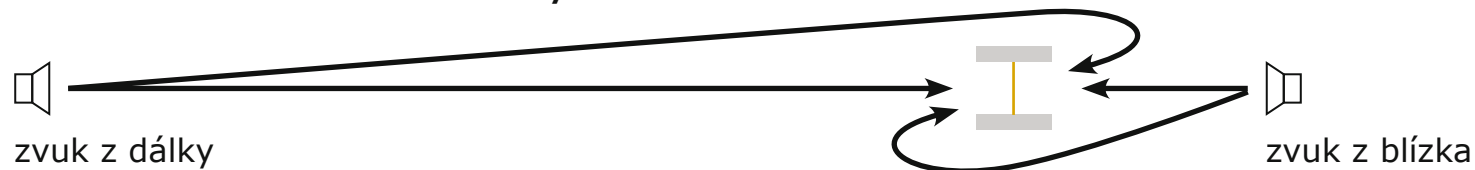
membrána



SMĚROVÁ CHARAKTERISTIKA

PROXIMITY EFEKT

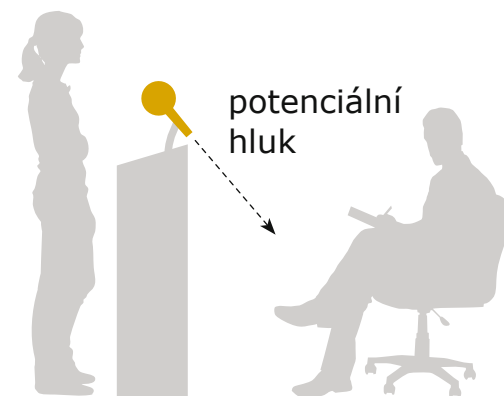
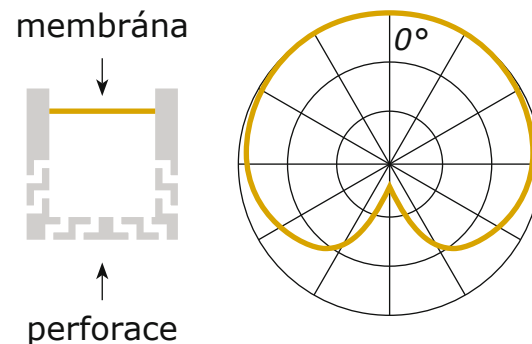
- u mikrofonů reagujících na rozdíl akustických tlaků (s otevřenou / polouzavřenou kapslí)
- zvuk z velké vzdálenosti \Rightarrow akustický tlak z obou stran membrány srovnatelný, ale kapsle stíní a zvuky jsou časově posunuté \Rightarrow detekce změny akustického tlaku
- zvuk z blízka \Rightarrow díky útlumu akustický tlak z jedné strany mnohem větší \Rightarrow zdůraznění nízkých frekvencí
 - např. hlas zní hlubší, než skutečně je
- dá se kompenzovat ekvalizací, pak ale nejsou slyšet nízké frekvence z dálky



SMĚROVÁ CHARAKTERISTIKA

KARDIOIDNÍ MIKROFON

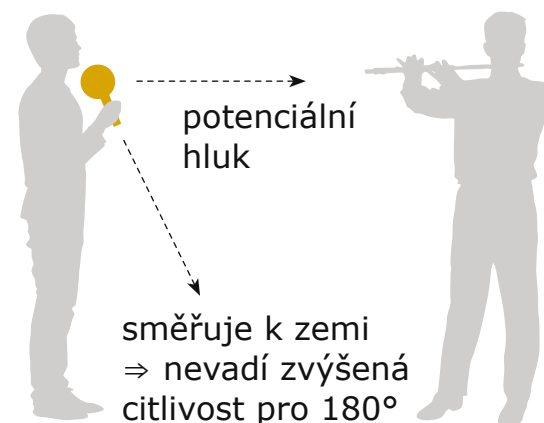
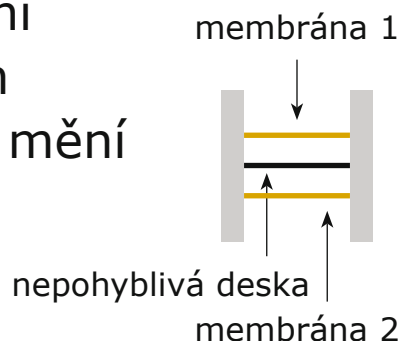
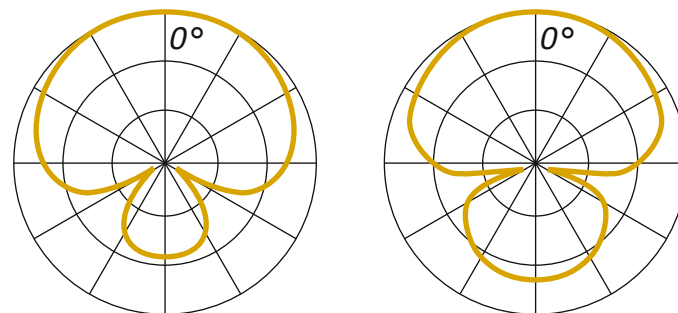
- polouzavřená mikrofonní kapsle, reaguje na rozdíl akustických tlaků zředu a zezadu kapsle
- zvuk zezadu \Rightarrow dopadá na membránu z obou stran \Rightarrow perforace taková, aby se akustické tlaky kompenzovaly
- trpí proximity efektem
- typický konferenční, jevištní mikrofon (nesnímá hluk ze sálu, zabraňuje zpětné vazbě vznikající, je-li mikrofon připojen na sálové reproduktory)



SMĚROVÁ CHARAKTERISTIKA

SUPERKARDIOIDNÍ M. HYPERKARDIOIDNÍ M.

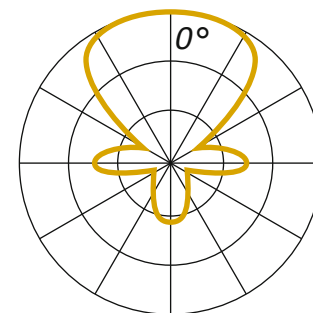
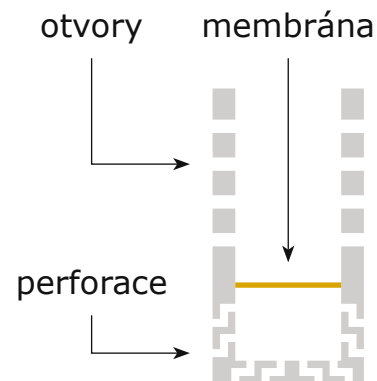
- stejný princip jako kardioidní mikrofon, konstrukcí daná úzká směrovost
⇒ dobrý reportážní mikrofon (snímá jen reportérův hlas)
- alternativně má mikrofonní kapsle dvě kapacitní membrány ⇒ jejich různým nabitím se mění směrová charakteristika



SMĚROVÁ CHARAKTERISTIKA

PUŠKOVÝ MIKROFON

- těž shotgun / lobar
- trubka s otvory + kardioidní mikrofon
- zvuk z předu \Rightarrow proniká současně otvory rozmístěnými s vhodnou roztečí \Rightarrow na membránu vše dopadá ve fázi \Rightarrow silný zvuk
- zvuk z jiné strany \Rightarrow není fázová shoda \Rightarrow silný útlum
- extrémní směrovost
- špatná frek. charakteristika
- natáčení hlasu pro film



SMĚROVÁ CHARAKTERISTIKA

SPECIÁLNÍ SMĚROVÝ MIKROFON

- v mikrofonu několik kapslí
- zvuk z jednoho směru \Rightarrow výstup z kapslí ve fázi
 \Rightarrow výsledný součet velký
- poměrem signálů z kapslí lze měnit směrovou charakteristiku
- mikrofonové pole – kapsle rozmístěny v ploše
 - stejný princip jako pole reproduktorů
 - elektronicky lze mikrofon zaměřit na daný bod v prostoru

NAHRÁVÁNÍ ZVUKU

SNÍMÁNÍ HLASU

- nevyžaduje velký frekvenční rozsah mikrofonu
- specifické problémy:
 - explozivní hlásky (b, p, d, t) – velký nárůst akustického tlaku – omezení sítkou před mikrofonem (pop filtr)
 - sykavky (c, s, z) – velký akustický tlak – omezení zpracováním signálu (de-esser)
- malá vzdálenost mikrofonu od úst – proximity efekt, nežádoucí ruchy (mlaskání, ...)
- velká vzdálenost mikrofonu od úst – vyžaduje citlivý mikrofon ⇒ bude snímat i ruchy z okolí
- typicky 15 – 60 cm

NAHRÁVÁNÍ ZVUKU

SNÍMÁNÍ HUDEBNÍCH NÁSTROJŮ

- volba vhodného typu mikrofonu
 - volba vhodné nahrávací vzdálenosti
 - volba počtu mikrofonů a součtu jejich signálů
 - volba prostředí, kde se bude nahrávat, umístění mikrofonů vzhledem k podlaze a zdem
 - případná úprava nástroje (obalení strun, přeladění, ...)
- ⇒ neexistují univerzální recepty, vyžaduje zkušenost a dobrý sluch
- zvuk se dále elektronicky upravuje (komprese, gate, ekvalizace, ...)
 - žádoucí podoba zvuku závisí na použití v mixu

STEREOFONNÍ NAHRÁVKA

UMĚLÝ STEREOFONNÍ OBRAZ

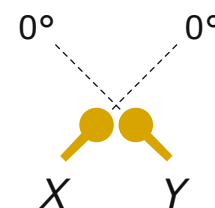
- jednotlivé nástroje (stopy) monofonní
- umělé umístění stop do prostoru – změna hlasitosti a zpoždění Left/Right při upmixu stopy 1.0 → stopa 2.0
- mixáž stereofonních stop
- je třeba dávat pozor na monofonní kompatibilitu (barva zvuku se nesmí výrazně změnit při downmixu na 1.0)

STEREOFONNÍ NAHRÁVKA

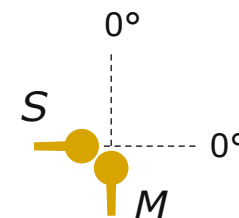
PŘIROZENÝ STEREOFONNÍ OBRAZ

- systém X/Y
 - dva kardioidní mikrofony kapslemi těsně u sebe v úhlu 90° (lze i osmičkové)
 - výstup X \rightarrow Right, Y \rightarrow Left
 - mezi zvuky X, Y není prakticky časový posuv \Rightarrow dobrá monofonní kompatibilita
- systém M/S (mid/side)
 - všesměrový (M) a osmičkový (S) mikrofony kapslemi těsně u sebe, osa osmičkového rovnoběžně se zdrojem zvuku
 - výstup M + S \rightarrow Right, M - S \rightarrow Left
 - funguje díky závislosti polarity S na umístění zdroje zvuku

zdroj zvuku



zdroj zvuku



STEREOFONNÍ NAHRÁVKA

- systém A/B
 - dva všesměrové mikrofony ve vzdálenosti m
 - typicky $m = d / 3$ až $d / 2$
typicky $m > 3v$
 - výstup A → Left, B → Right
 - kvalitní stereofonní obraz
 - problematická monofonní kompatibilita

- vylepšení monofonní kompatibility:
 - $m = 17$ cm (cca vzdálenost uší)
- dobrý stereofonní obraz
- při downmixu na mono jsou maximální útlumy vlivem efektu hřebenového filtru od sebe vzdáleny > 2 kHz
⇒ efekt hřebenového filtru není tolik slyšet

