

# PROSTOROVÁ REPRODUKCE ZVUKU



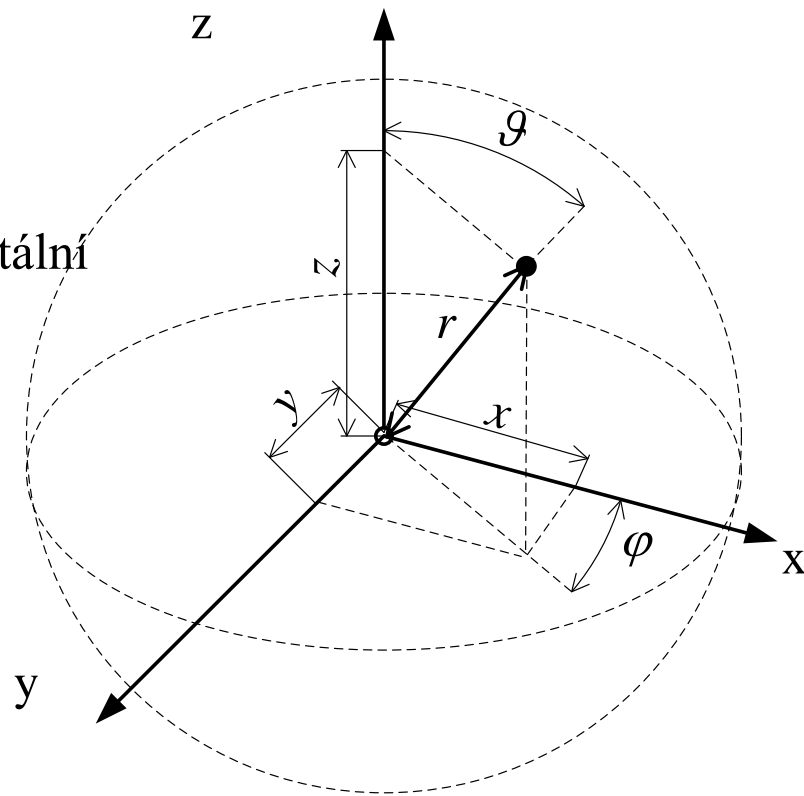
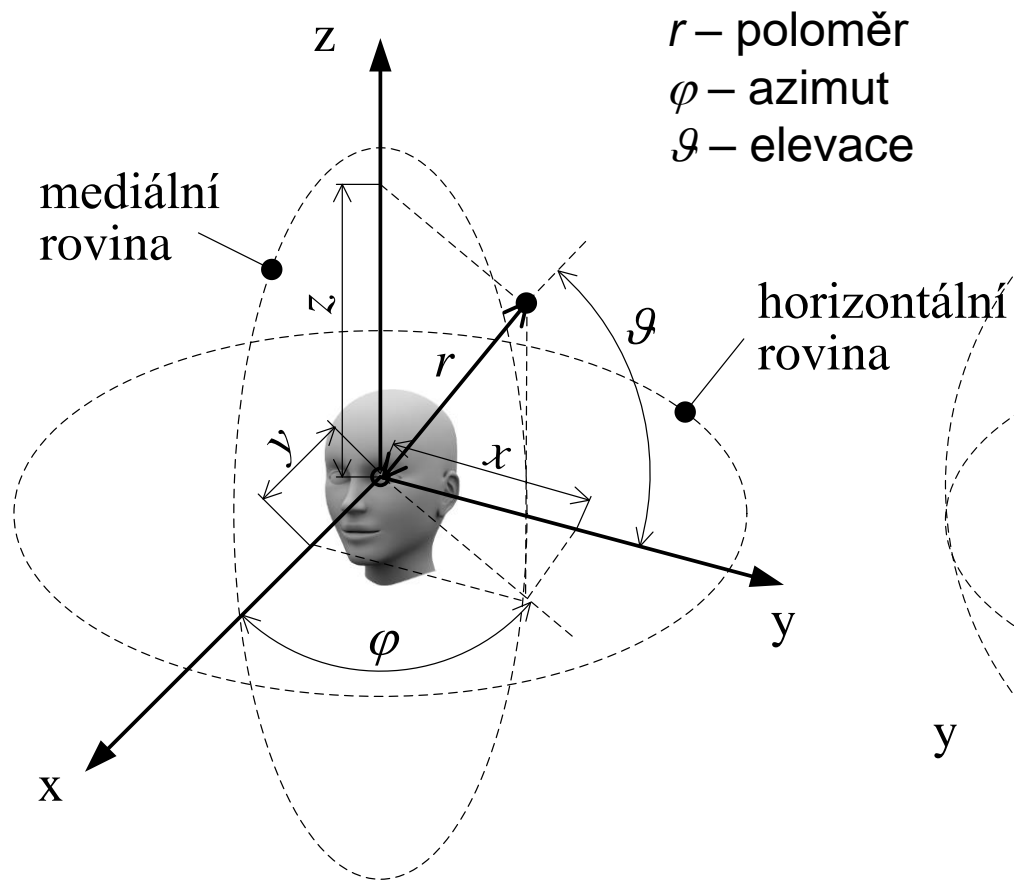
**Kurz:**      **Elektroakustika 2**

---

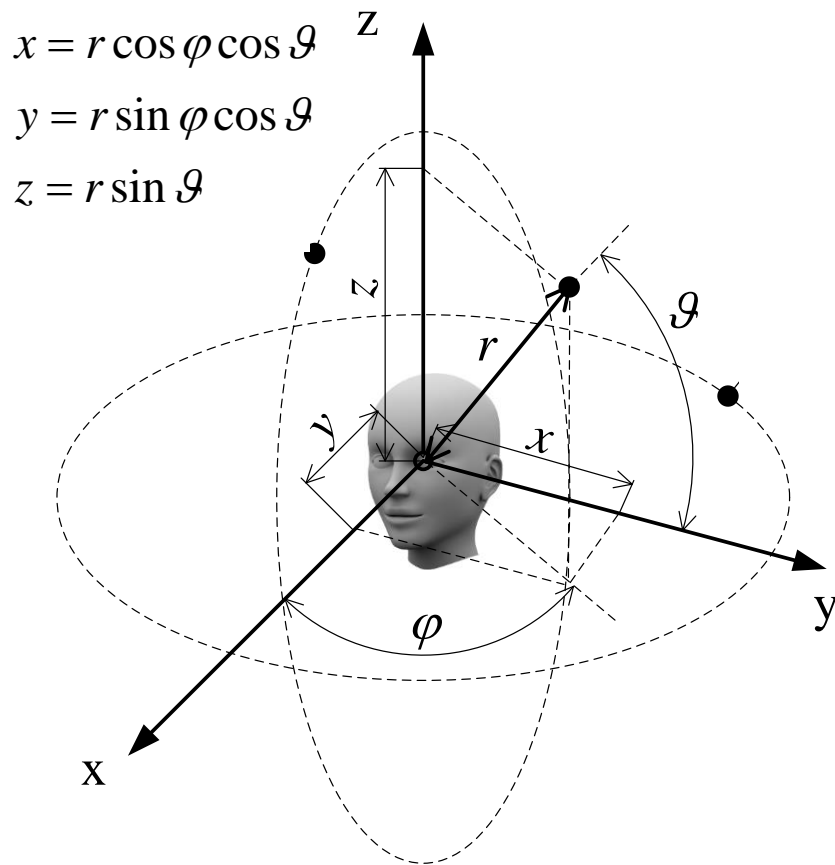
**Lektor:**    Jiří Schimmel

**Autor:**    Jiří Schimmel

# Interaurální vs. sférické souřadnice



# Interaurální vs. kartézské souřadnice



$$x = r \cos \varphi \cos \vartheta$$

$$y = r \sin \varphi \cos \vartheta$$

$$z = r \sin \vartheta$$

$$r = \sqrt{x^2 + y^2 + z^2}$$

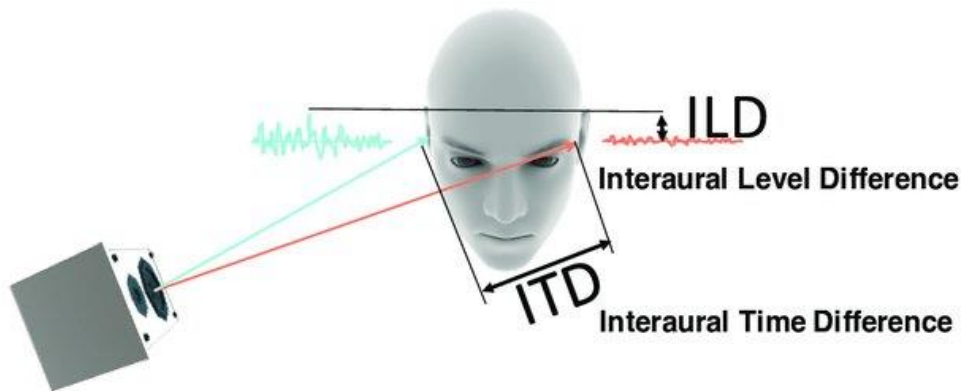
$$\varphi = \arctan \begin{cases} \arctan\left(\frac{x}{y}\right) & \text{pro } (x > 0) \text{ a } (y \geq 0) \\ \arctan\left(\frac{x}{y}\right) + \pi & \text{pro } (x < 0) \\ \arctan\left(\frac{x}{y}\right) + 2\pi & \text{pro } (x > 0) \text{ a } (y < 0) \end{cases}$$

$$\vartheta = \arcsin\left(\frac{z}{r}\right)$$

# Psychoakustika prostorové reprodukce

# Interaurální difference

- Při dopadu zvukového vlnění mimo mediální rovinu budou vlnění dopadající na obě uši rozdílná
  - 1) Interaurální rozdíl intenzit (Interaural Intensity Difference, IID)  
Interaurální rozdíl úrovní (Interaural Level Differences, ILD)
    - rozdíl hladin intenzity zvuku / akustického tlaku v přivráceném a odvráceném uchu
  - 2) Interaurální časový rozdíl (Interaural Time Difference, ITD)
    - rozdíl času dopadu zvukového vlnění do přivráceného a odvráceného ucha



# Interaurální časové rozdíly

- Binaurální rozdíl vzdáleností:

$$\Delta l = d_1 + d_2 = d \sin \varphi$$

- při započítání ohybu kolem hlavy

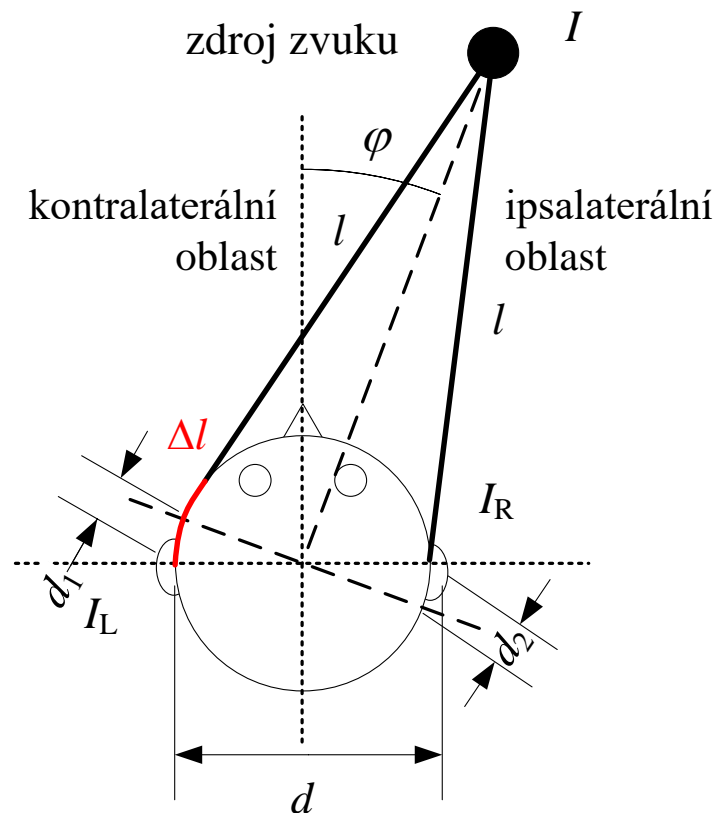
$$\Delta l' = d \frac{(\varphi + \sin \varphi)}{2}$$

- vzdálenost mezi ušima  $d = 15 \div 21$  cm  
Interaurální časová diference

$$ITD = \Delta l / c_0$$

- maximální ITD je při dopadu vlnění ze strany

$$ITD_{\max} \cong 0,44 \div 0,61 \text{ ms } (0,57 \div 0,79 \text{ ms})$$

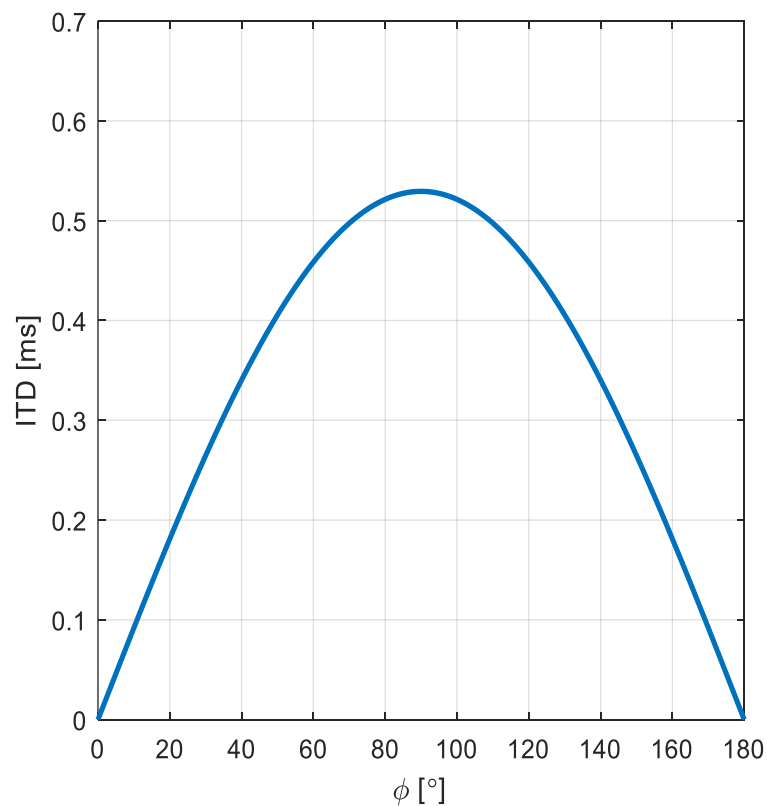


# Interaurální fázové rozdíly

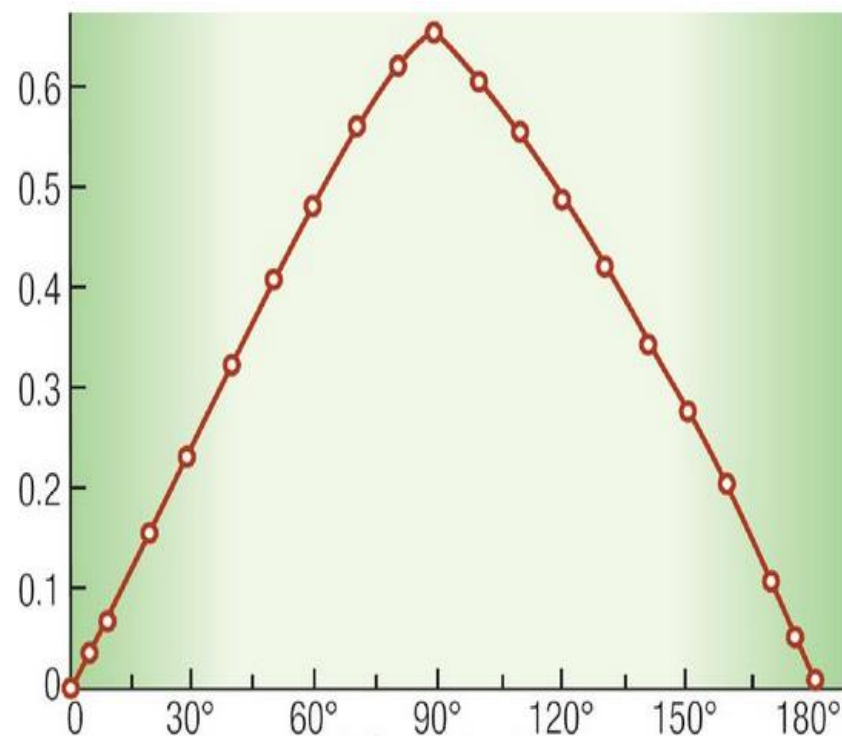
- Postřehnutelná ITD je kmitočtově závislá, na 1 kHz je postřehnutelná od 10–15  $\mu\text{s}$ , směrem k nižším a vyšším kmitočtům roste
- Pro harmonické signály znamená ITD fázový rozdíl (*Interaural Phase Difference*, IPD)
  - posune-li se vzájemně fáze zvuku dopadajících na obě uši, je zvuk vnímán na straně toho ucha, v němž fáze zvuku předbíhá
  - postřehnutelné do vlnové délky, která je dvojnásobkem binaurálního rozdílu vzdáleností pro zvuk dopadající ze strany; posune-li se fáze více než o polovinu trvání periody, zvuk přeskočí do druhého ucha

# Interaurální fázové rozdíly

teoretický průběh:  $ITD = d(\sin\varphi)/c_0$



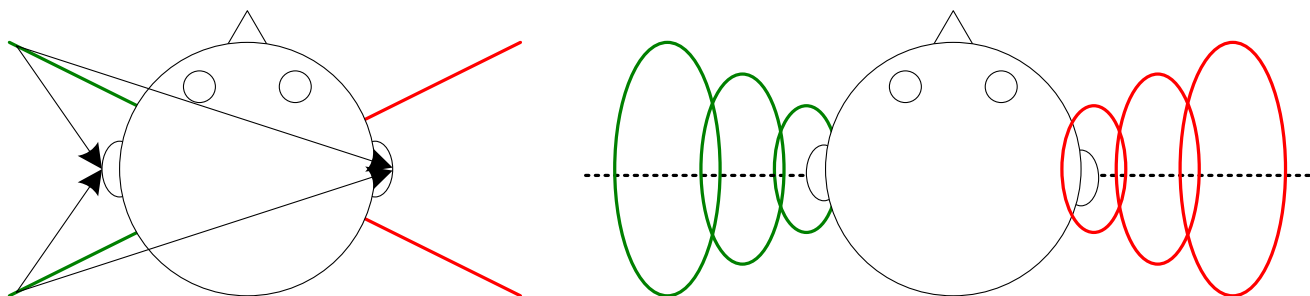
naměřená data





# Kužel nejistoty

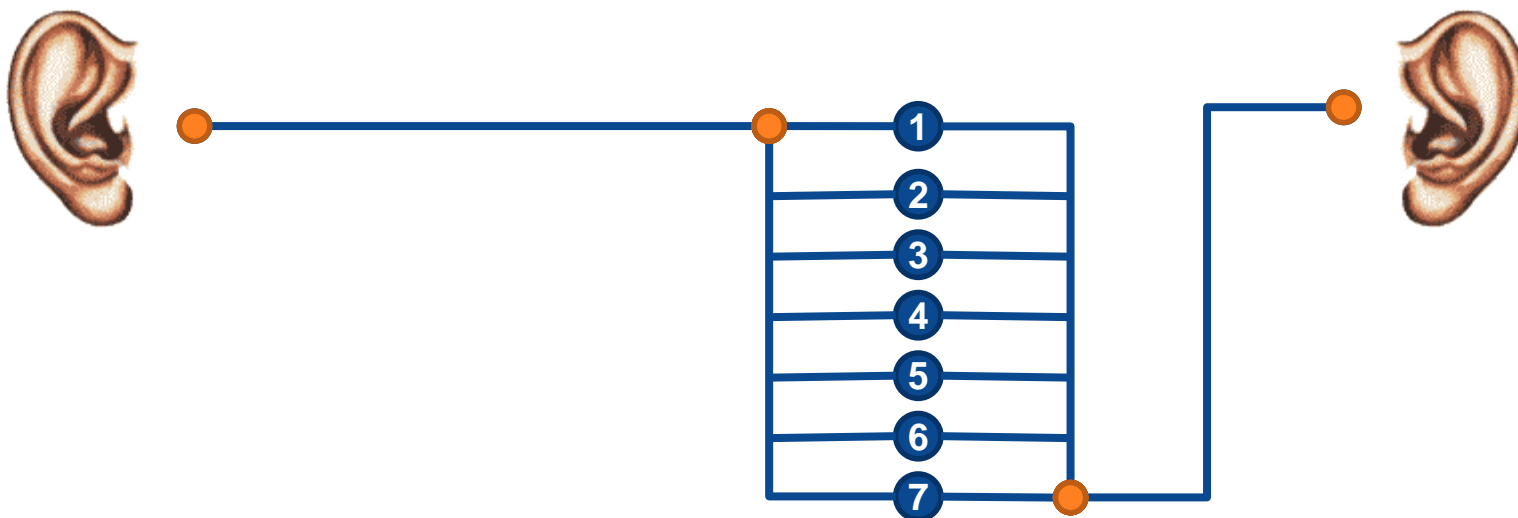
- Body se stejnou vzdáleností od obou uší leží na hyperbolách → pozice zdroje zvuku není pouze pomocí IID a ITD definována jednoznačně
- Kužel nejistoty (*cone of confusion*): spojení bodů se stejným binaurálním rozdílem ve 3D prostoru
- Pro určení pozice zdroje zvuku ve 3D prostoru jsou nutné další mechanismy analýzy sluchového vjemu



# Jeffressův model

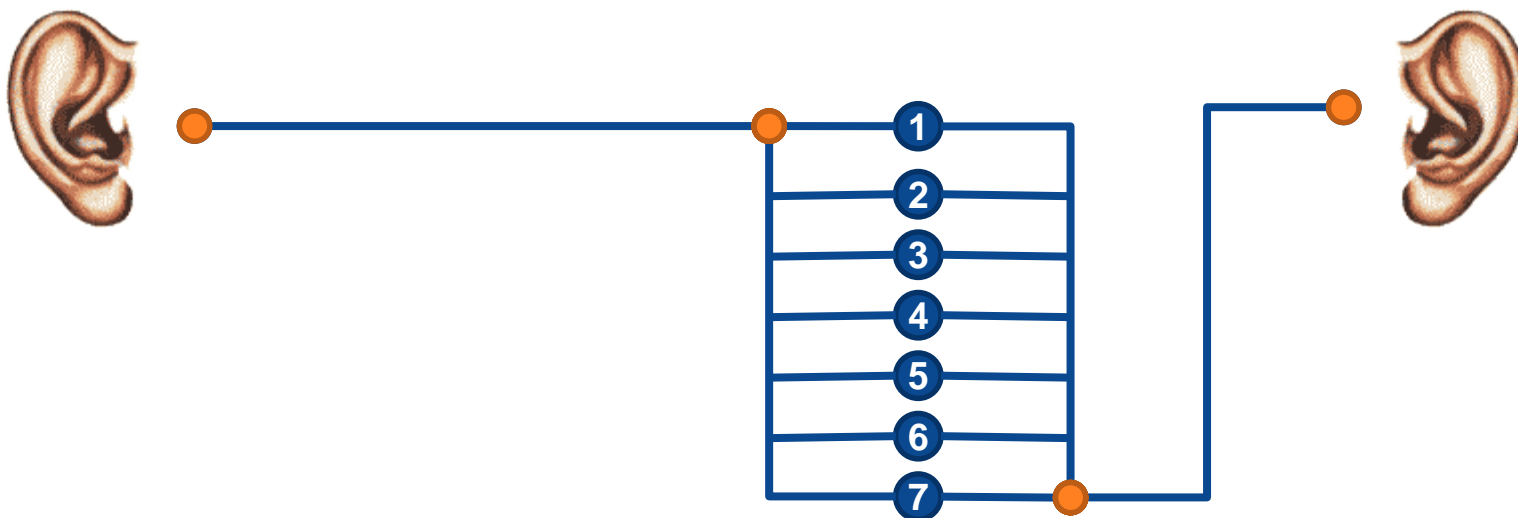
- Detekce ITD pomocí zpožďovací linky a detektoru shody

## 1. Zvukové vlnění dopadající zepředu



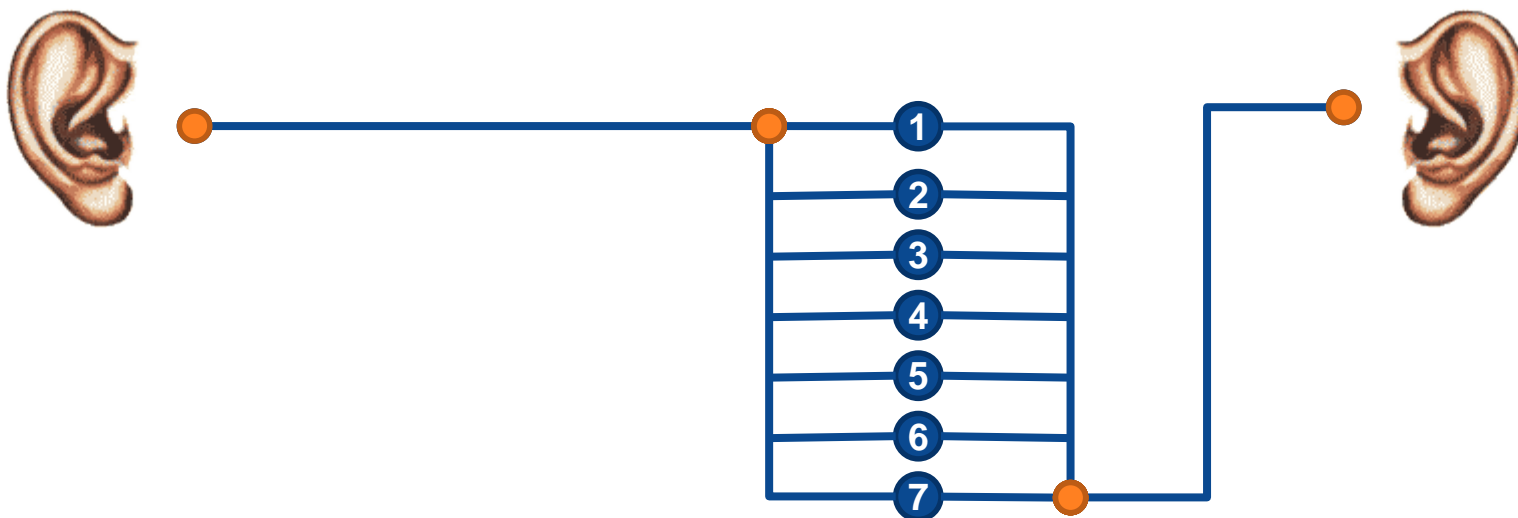
# Jeffressův model

- Detekce ITD pomocí zpožďovací linky a detektoru shody
2. Zvukové vlnění dopadající zleva



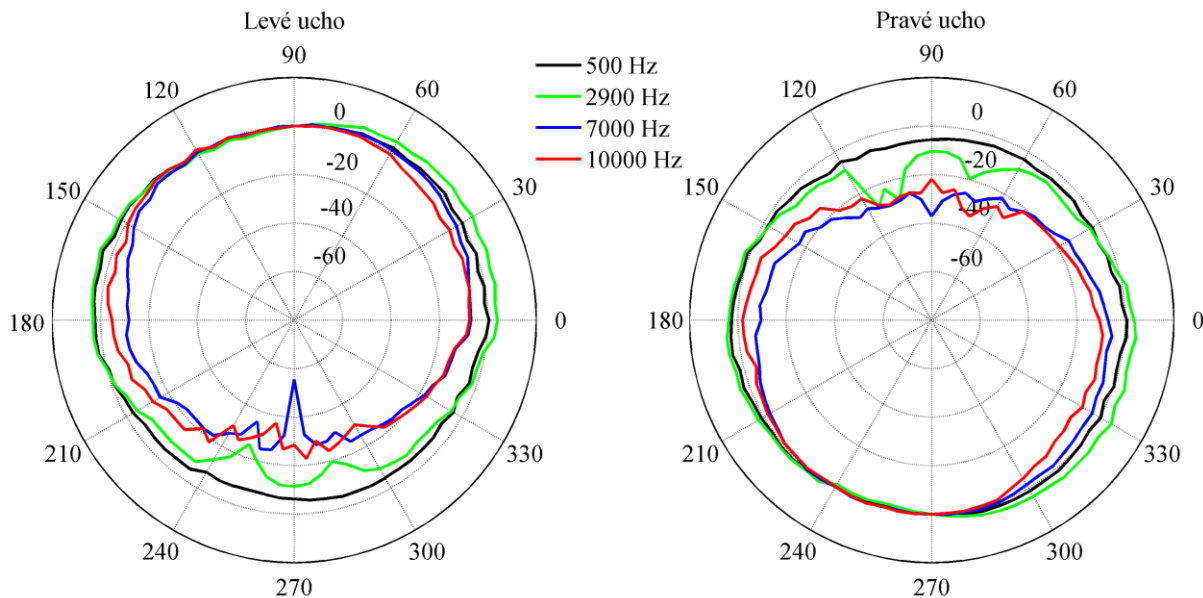
# Jeffressův model

- Detekce ITD pomocí zpožďovací linky a detektoru shody
3. Zvukové vlnění dopadající zprava

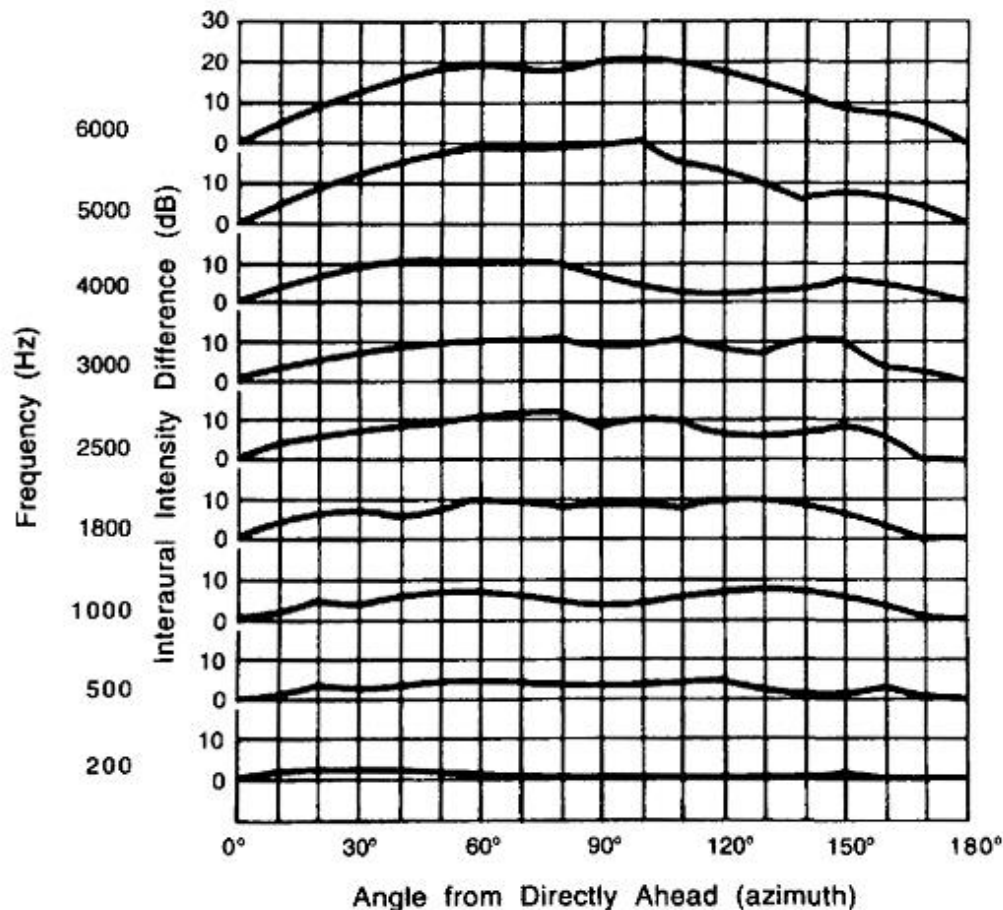


# Interaurální rozdíly úrovní

- Kratší dráha zvukového vlnění = nižší útlum šířením (velmi blízké zdroje)
- **Akustický stín hlavy:** pro vyšší kmitočty, kde jsou rozměry hlavy větší než vlnová délka vlnění
  - vlnění s vlnovou délkou větší než cca  $1/3$  poloměru hlavy, tj. od cca 540 ÷ 750 Hz



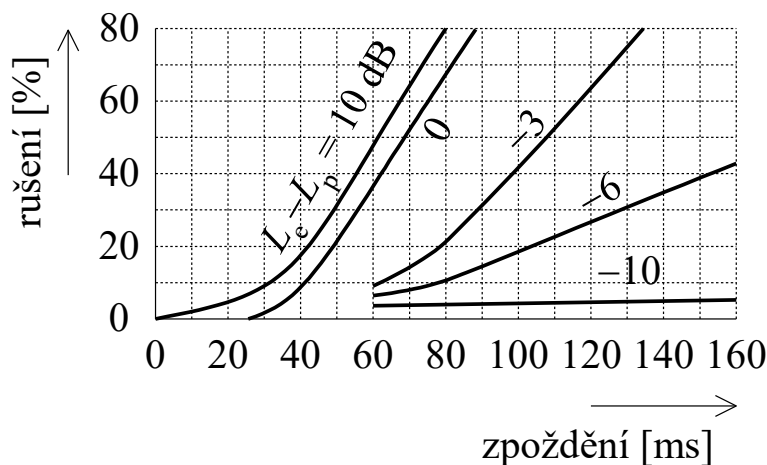
# Interaurální rozdíly úrovní



W. E. Feddersen, T. T. Sandel, D. C. Teas, and L. A. Jeffress, "Localization of High-Frequency Tones", The Journal of the Acoustical Society of America 29, 988-991 (1957)

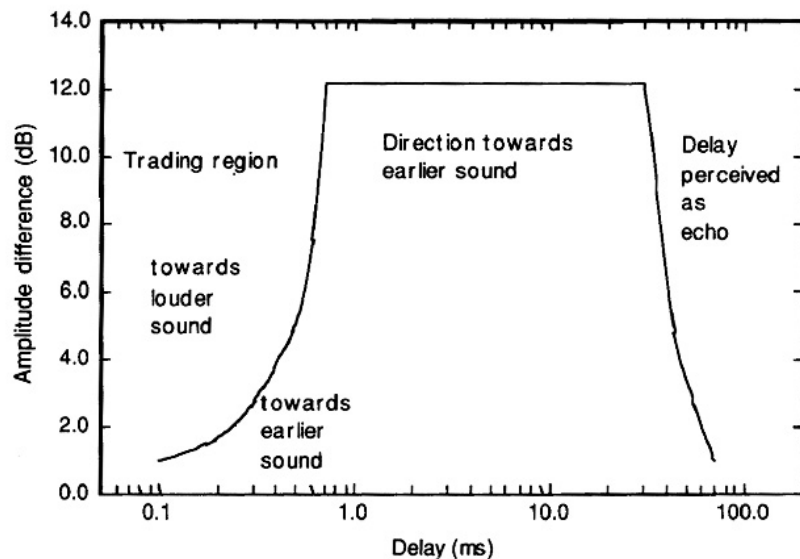
# Jev priority

- Haasův efekt, efekt předstihu, precedence effect, law of the first wavefront
- Zvuková vlna, která dorazí k uchu první, převažuje při určování směru
- Haasův experiment: přichází-li odraz k posluchači se zpožděním do 30 ms, posluchač směr odrazu nevnímá, odražený signál nepostřehne a zdroj zvuku určuje podle příchodu přímé zvukové vlny i tehdy, když hladina intenzity odražené vlny je o 7 až 10 dB vyšší než hladina intenzity přímé vlny

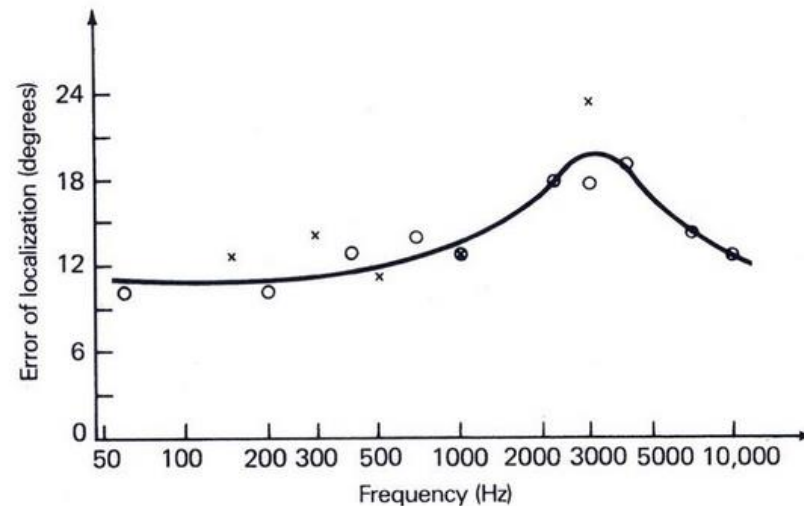


# Společné působení ILD a ITD

interakce IID a ITD



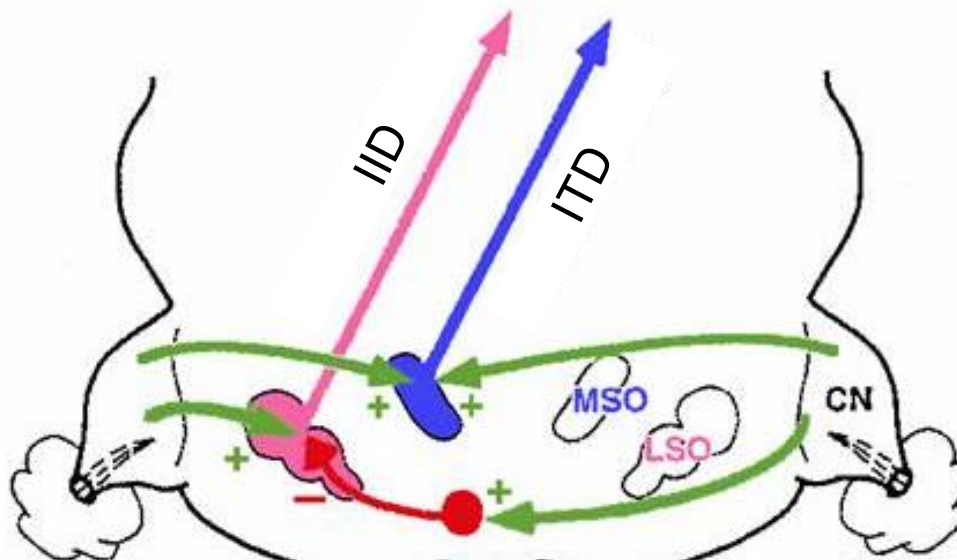
přesnost určení směru přicházejícího zvuku v závislosti na kmitočtu





# Zpracování interaurálních rozdílů mozkiem

- V prodloužené míše:
  - MSO – rozpoznání ITD (zpožďovací linka s detektory)
  - LSO – rozpoznání ILD (rozdílový člen)

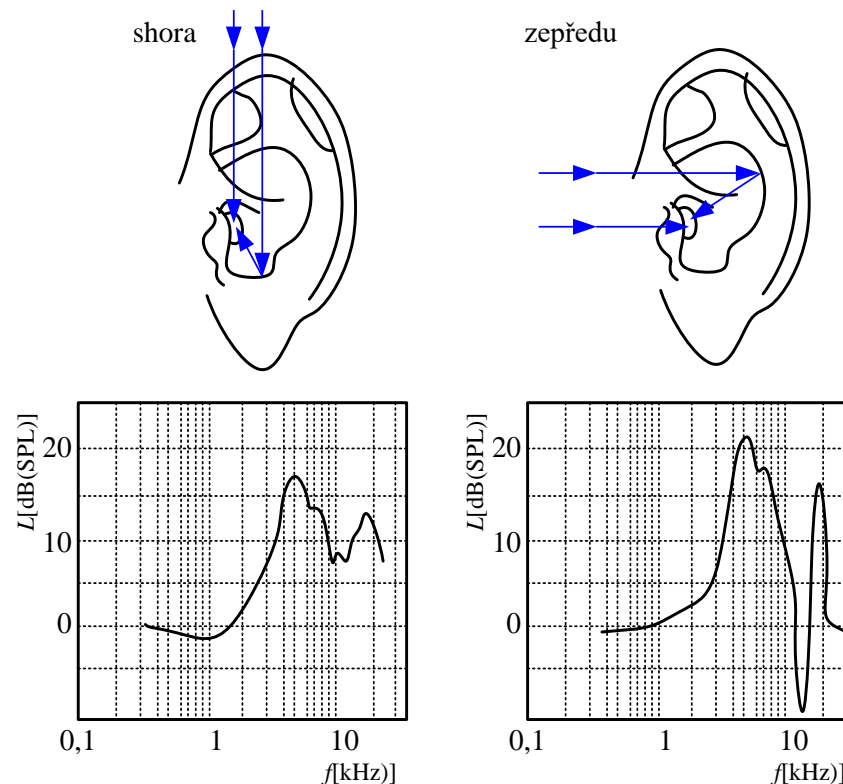


# Lokalizace zvuku v mediální rovině

- **Lateralizace:** vnímání zvuku na spojnici mezi oběma ušima uvnitř hlavy
- **Externalizace:** lokalizace zdroje zvuku vně hlavy
- Mechanismy externalizace
  - Změny spektra na vyšších kmitočtech
    - akustický stín boltce – útlum vyšších kmitočtů (od 800 Hz) vlnění přicházejících zezadu
    - odrazy od ramen – část odraženého signálu směřuje k boltci a sečte se se signálem přímým
    - odrazy ve vnějším uchu – vlny mnohonásobně odražené na vnitřní struktuře boltce se setkávají na bubínku s různými fázovými posuny
  - Pohyb zdroje zvuku – časová změna IID a ITD
  - Pohyby hlavou – časová změna IID a ITD
  - Odrazy zvuku v prostoru

# Odrazy od boltce a ramen

- Odrazy od boltce:
  - akustický směrový filtr – pro každý směr se mění dráhový rozdíl původního signálu a signálu odraženého od vnitřní struktury boltce,
  - na bubínku se setkávají s různými fázovými posuny v závislosti na vlnové délce,
  - každému směru odpovídá jiný filtr,
  - pro různé směry vznikají lokální extrémy v části přenosové funkce.
- Odrazy od ramene: část odraženého signálu směřuje k boltci a sečte se se signálem přímým.



Algazi, R., Aveando, C., Thompson, D., Avendano, C., "CIPIC HRTF Database," in Proc. IEEE WASPAA'01, New Platz, NY, 2001, <http://interface.cipic.ucdavis.edu/>.

# Interaurální koherence

- Interaurální koherence (Interaural Coherence, IC) – míra podobnosti signálu v levém a pravém uchu, maximální hodnota normalizované křížové korelační funkce:

$$IC = \max_d \frac{\sum_{n=-\infty}^{\infty} p_L(n) p_R(n+d)}{\sqrt{\sum_{n=-\infty}^{\infty} p_L^2(n) \sum_{n=-\infty}^{\infty} p_R^2(n+d)}}$$

$p_L, p_R$  – akustický tlak v levém resp. pravém uchu

$d$  je uvažováno v rozsahu  $\pm 1$  ms

$$IC \in \langle 0, 1 \rangle$$

- $IC = 1$ : signály jsou koherentní (pouze různá amplituda a fáze)
- $IC = 0$ : signály jsou zcela nezávislé

# Právě postřehnutelné rozdíly

- Právě postřehnutelné rozdíly (Just-Noticeable Differences, JND): minimální změna difference nutná k tomu, aby byla subjektivně postřehnutelná
  1. **ILD**: 0,1 až 1 dB, téměř nezávislá na kmitočtu a hladině akustického tlaku, závislá na absolutní hodnotě ILD
  2. **ITD**: silně závislá na kmitočtu, do kmitočtu 1000 Hz konstantní IPD cca 0,05 rad (cca 1/60 periody), nezávislá na hladině akustického tlaku, závislá na absolutní hodnotě ITD
  3. **IC**: silně závislé na IC, pro IC = 1 je postřehnutelná změna o 0,002, pro IC blízké 0 musí být změna 100x větší, aby byla postřehnutelná
- Pro dobu trvání podnětu kratší než cca 400 ms se právě postřehnutelné difference zvětšují

# Právě postřehnutelné rozdíly

- Minimální postřehnutelný rozdíl směru přicházejícího zvuku (Localization blur)
  - závisí na směru (nejmenší vpředu, největší na stranách)
  - v horizontální rovině pro zvuk přicházející ze předu:
    - $0,75 \div 2^\circ$  pro šumové impulsy,
    - $1 \div 4^\circ$  pro harmonické signály (závislý na kmitočtu),
    - $0,9 \div 1,5^\circ$  pro řeč,
    - $3,2^\circ$  pro širokopásmový šum
  - v horizontální rovině pro zvuk přicházející z boku:  $10^\circ$
  - v mediální rovině:  $10 \div 22^\circ$  (nejvyšší shora)

# Vliv kritických pásem na prostorové slyšení

- Při sledování právě rozpoznatelné ITD dvou skupin několika period harmonického signálu s různým kmitočtem jako funkce rozdílu těchto kmitočtů bylo zjištěno, že až do rozdílu kmitočtů blízkému šířce kritického pásma je právě rozpoznatelná ITD malá a pro vyšší rozdíly výrazně roste → lidský sluch vnímá interaurální rozdíly v kmitočtových pásmech, která se shodují resp. blíží kritickým pásmům.

# Lokalizace zvuku – shrnutí

- **Duplexní teorie** (Rayleigh, 1907): ITD určuje lokalizaci na nízkých kmitočtech, ILD na vysokých, hranice je mezi 1 a 2 kHz
- **Lokalizace v systémech prostorového zvuku** (Gerzon, 2006)
  1. nízké kmitočty: do kmitočtu 400, resp. až 700 Hz nepředstavuje hlava žádnou překážku – IID resp. ILD jsou zanedbatelné a o směru přicházejícího zvuku rozhodují ITD resp. IPD
  2. střední kmitočty: pro signály od kmitočtu 400, resp. 700 Hz je vlnová délka signálu tak malá, že ITD resp. IPD nejsou nejdůležitějším faktorem, záleží spíše na IID resp. ILD způsobené akustickým stínem hlavy
  3. vysoké kmitočty: u kmitočtů nad cca 5 kHz se projevují spektrální změny zvuku vlivem odrazů od ušních boltců



# Stereofonní reprodukce zvuku







# Stereofonní reprodukce zvuku



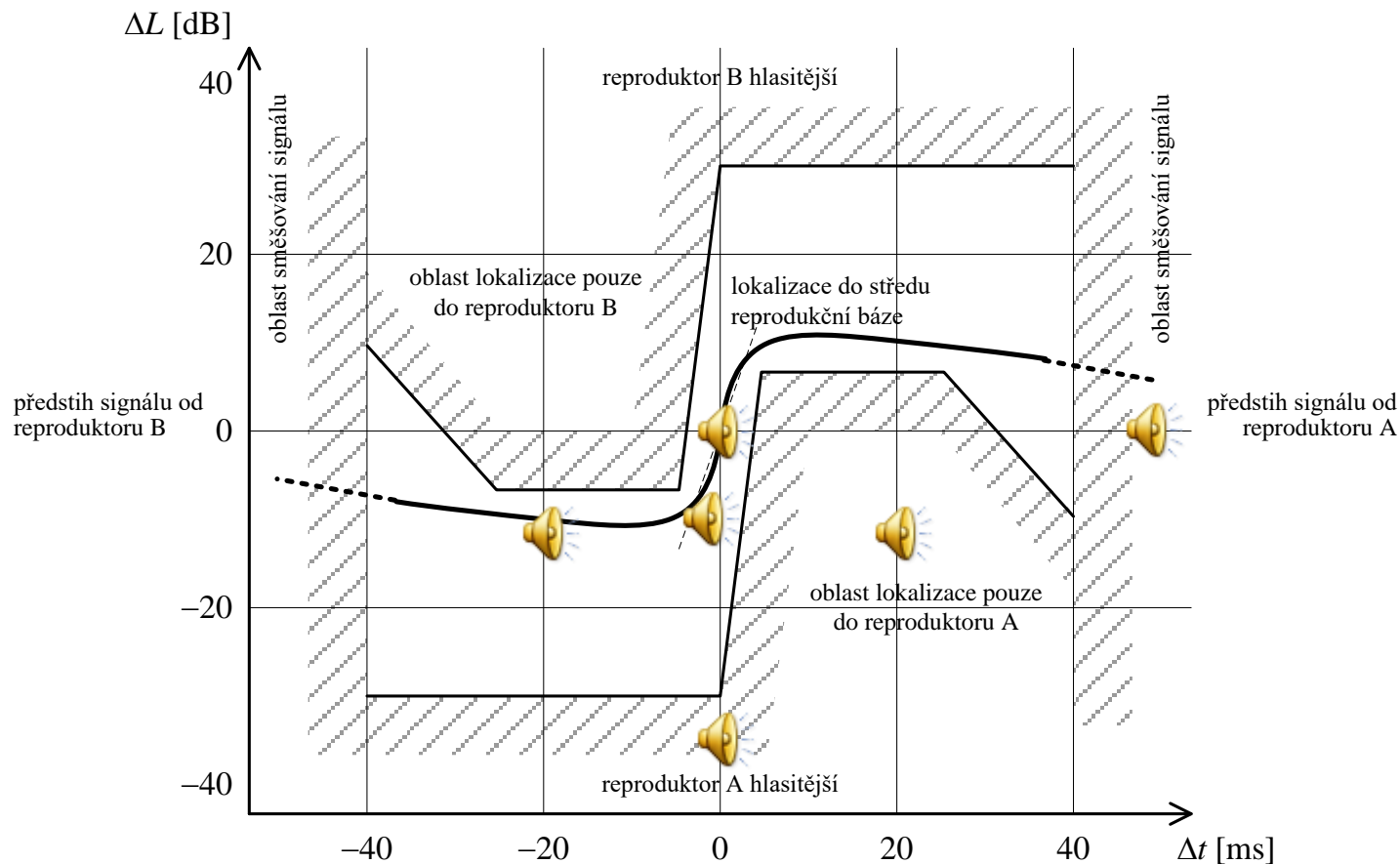
# Stereofonní reprodukce zvuku

- Definice pojmů:
  - Mezikanálová časová diference (Inter-Channel Time Difference, ICTD): rozdíl zpoždění signálu mezi levým a pravým kanálem
  - Mezikanálová diference úrovní (Inter-Channel Level Difference, ICLD): rozdíl úrovně signálu mezi levým a pravým kanálem
  - Mezikanálová koherence (Inter-Channel Coherence, ICC): koherence signálu mezi levým a pravým kanálem

# Stereofonní reprodukce zvuku

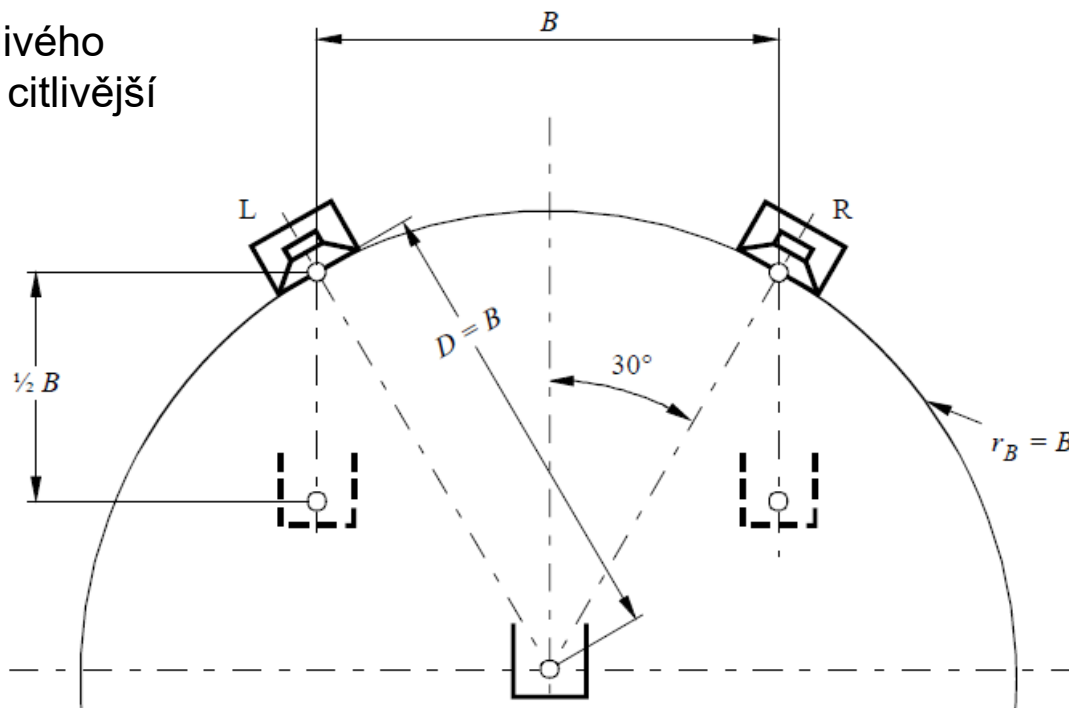
- Mechanismy interaurálních diferencí se navzájem doplňují
  - rozdíl času příchodu jednotlivých zvukových vlnění  $\Delta t = t_L - t_P$
  - rozdíl úrovně akustického tlaku těchto vlnění  $\Delta L = L_L - L_P$
- Van de Boerova závislost směrové lokalizace 
  -   $\Delta t < \text{asi } 4 \text{ ms}$ ,  $\Delta L < \text{asi } 20 \text{ dB}$  (koincidenční mikrofonní techniky)
  -   $\Delta t < \text{asi } 4 \text{ ms}$ ,  $\Delta L > \text{asi } 20 \text{ dB}$  (směr určen hlasitějším reproduktorem)
  -   $4 \text{ ms} < \Delta t < \text{asi } 30 \text{ ms}$ ,  $\Delta L < \text{asi } 6 \text{ dB}$  (Haasův jev, přizvučování)
  -   $4 \text{ ms} < \Delta t < \text{asi } 30 \text{ ms}$ ,  $\Delta L > \text{asi } 6 \text{ dB}$  (mikrofonní techniky blízké koincidence)
  -   $\Delta t > \text{asi } 40 \text{ ms}$  (směšování)

# Van de Boerova závislost směrové lokalizace

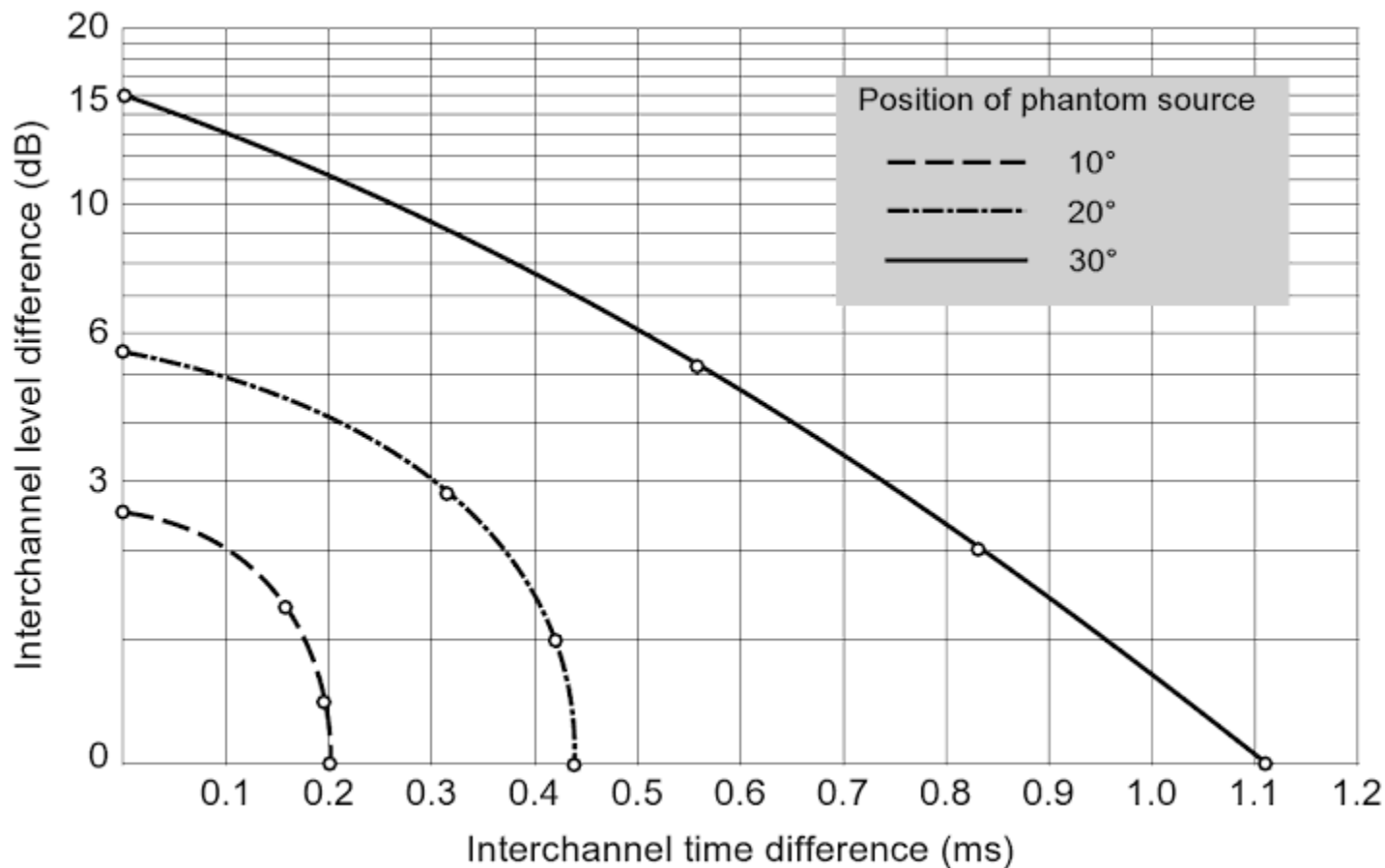


# Stereofonní reprodukce zvuku

- Ideální konfigurace – rovnostranný trojúhelník
  - posluchač ve stejné vzdálenosti od obou reproduktorů
  - reproduktory ve směru  $\pm 30^\circ$
  - pro větší úhly je pozice zdánlivého zdroje zvuku méně stabilní a citlivější na efekt rotace hlavy

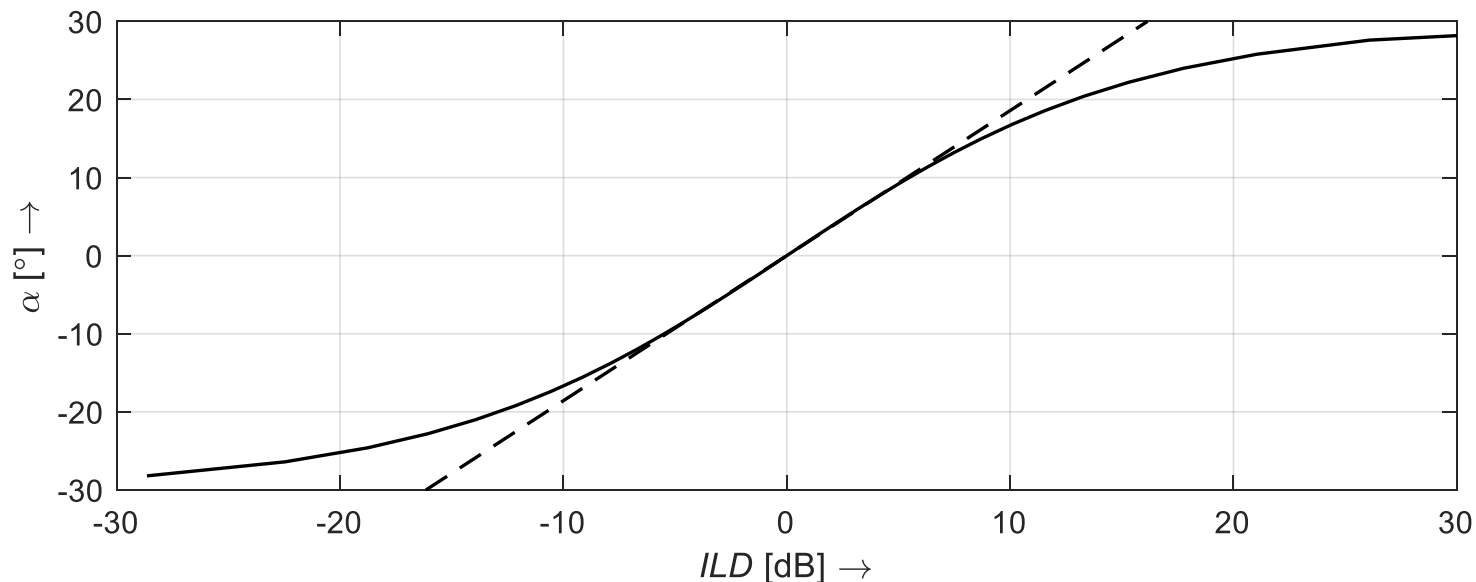


# Pozice virtuálního zdroje zvuku



# Intenzitní stereofonie

- Vjem směru přicházejícího zvuku je založen pouze na ILD
  - výpočet zesílení pro levý a pravý kanál podle tak, aby rozdíly úrovní zvukového signálu mezi kanály kopírovaly křivku ILD





# Intenzitní stereofonie

- Amplitudové panorámování
  - Sinový zákon:

$$\frac{\sin \theta}{\sin \theta_0} = \frac{g_1 - g_2}{g_1 + g_2}$$

- Tangentový zákon:

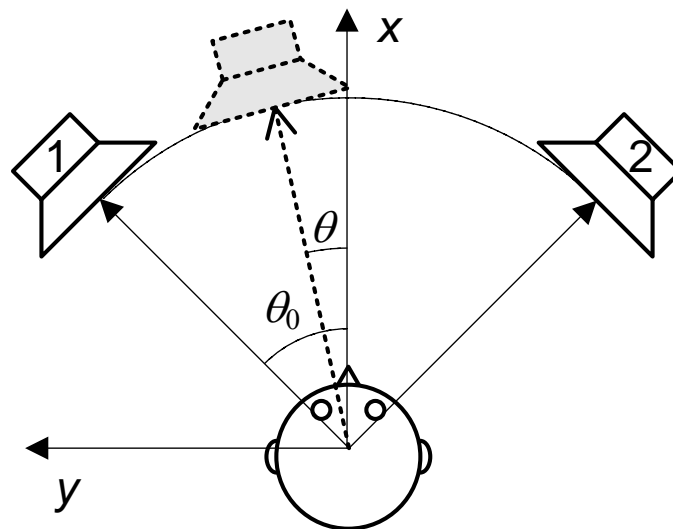
$$\frac{\tan \theta}{\tan \theta_0} = \frac{g_1 - g_2}{g_1 + g_2}$$

$g_1, g_2$  – zesílení reproduktorů

- normalizace zesilovacích činitelů na konstantní akustický výkon

$$g_1^2 + g_2^2 = C$$

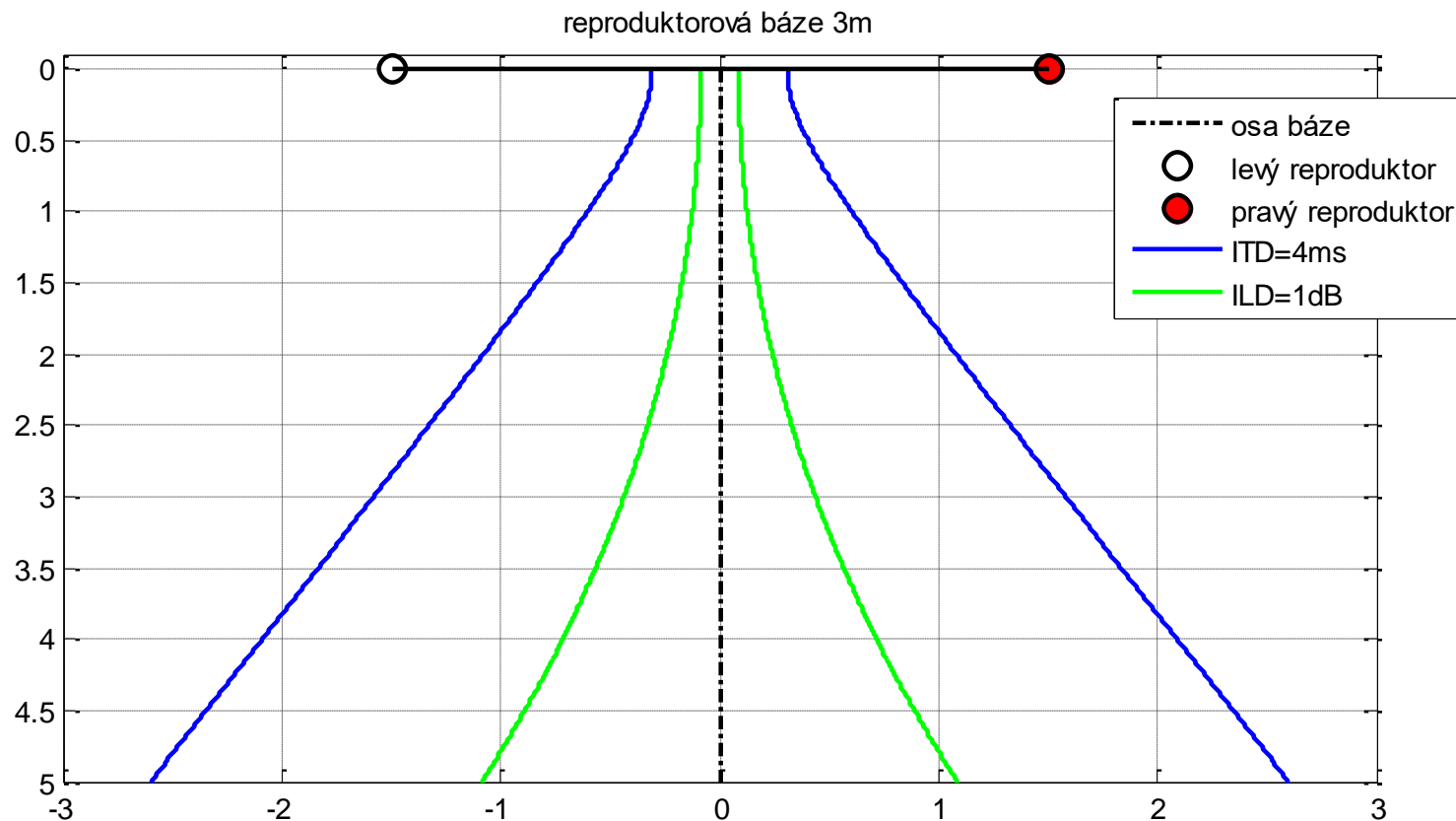
parametrem  $C$  je řízena hlasitost



# Intenzitní stereofonie

- Poslechem mimo osu reprodukční báze přidáváme vliv ITD a ovlivňujeme ILD:
  - rozdíl časového doběhu signálu od jednotlivých reproduktorů je dán hranicí jevu priority, tj. musí být  $< 4$  ms (rozdíl vzdáleností od posluchače musí být  $< 1,3$  m)
  - rozdíl hladin akustického tlaku od jednotlivých reproduktorů je dán právě postřehnutelnými ILD, tj. musí být  $< 1$  dB (poměr vzdáleností od posluchače musí být  $< 1,12$ )
- Doporučená šířka stereofonní báze: 2 – 3 m, neměla by přesáhnout 3,5 m

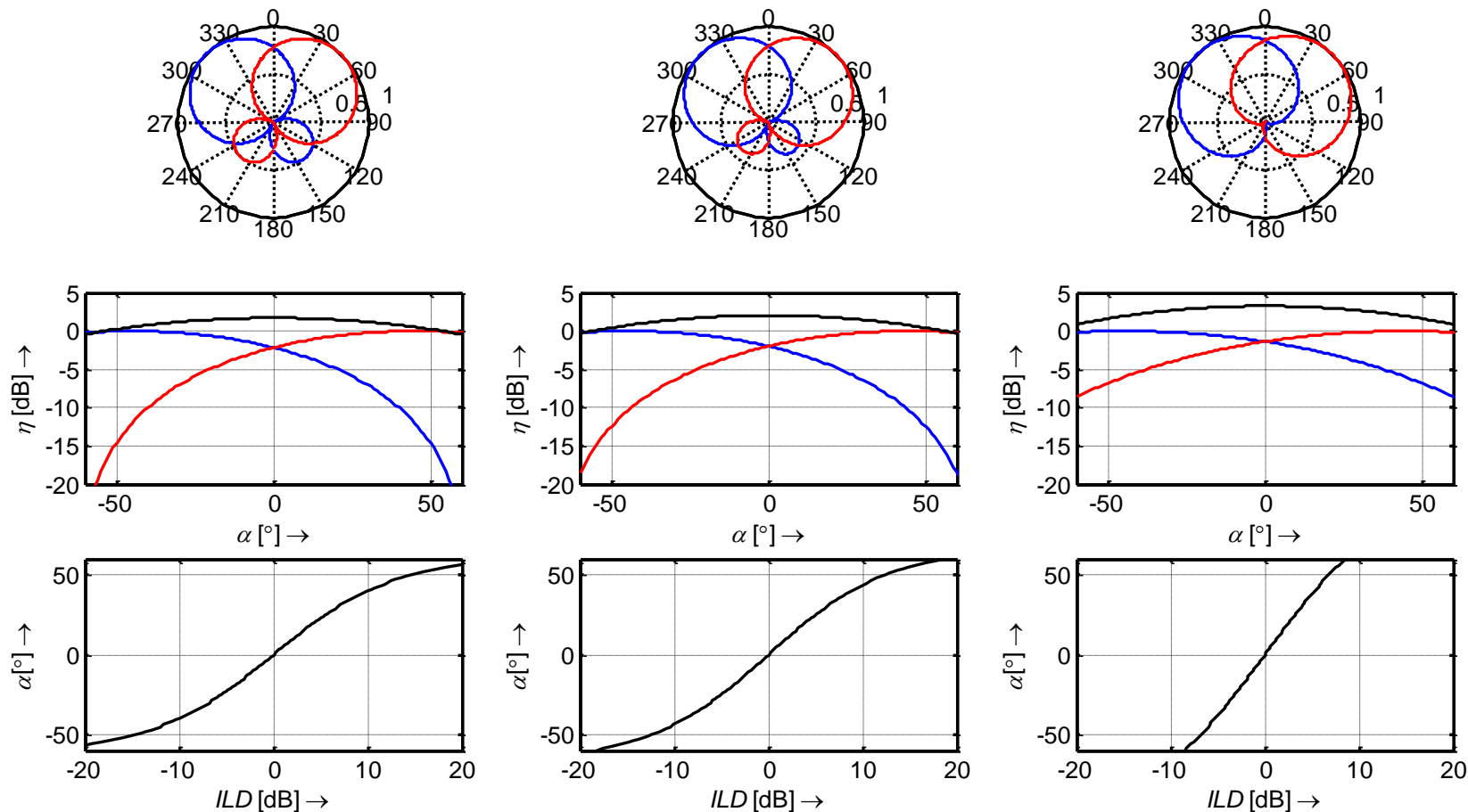
# Intenzitní stereofonie



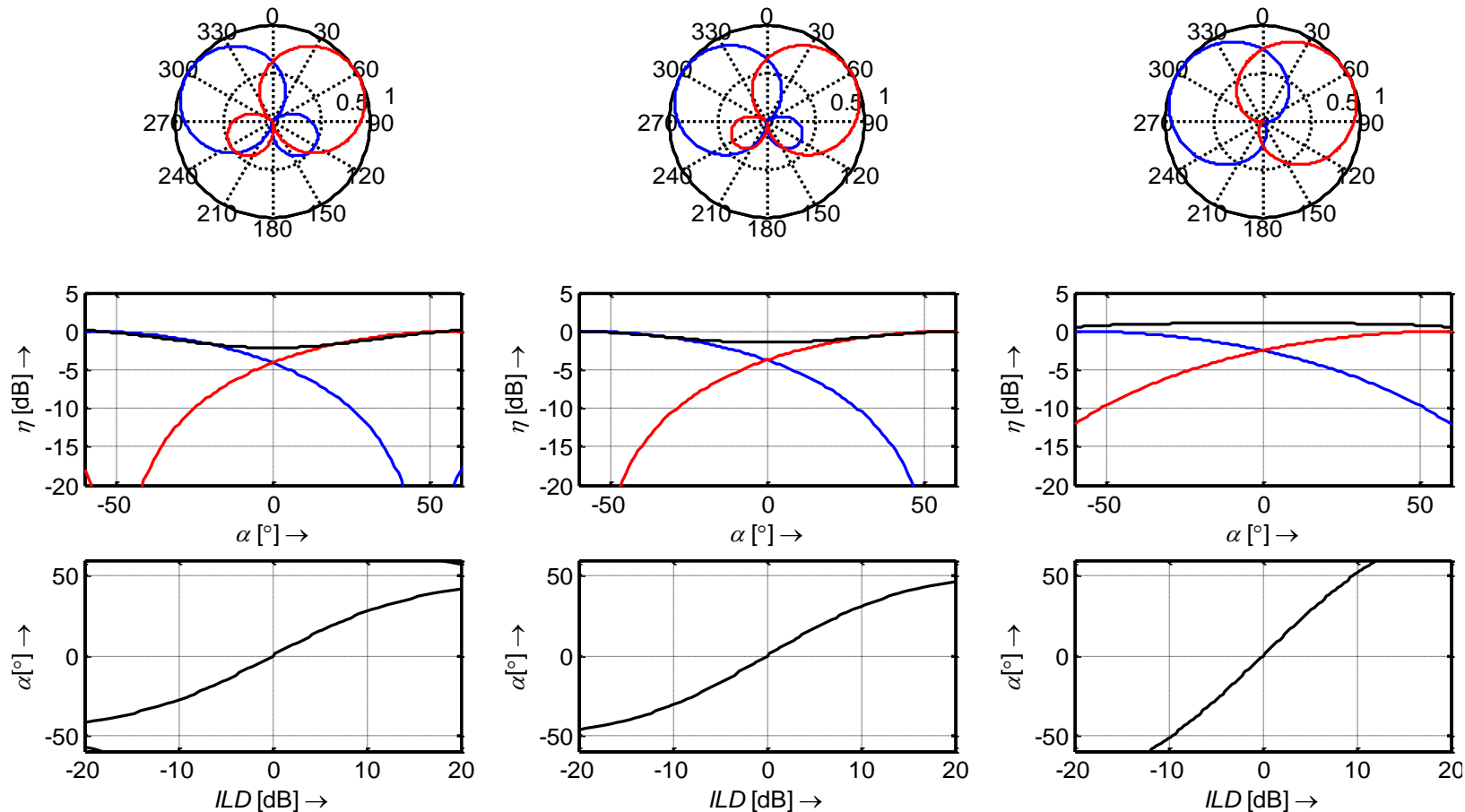
# Koincidenční mikrofonní techniky

- Zanedbatelný časový (fázový) rozdíl, uplatňuje se pouze ILD
- **XY**: pár mikrofónů umístěných v rovině, osy natočené mezi sebou o úhel:
  - kardioidní směrová charakteristika:  $\alpha = 90^\circ$  až  $120^\circ$
  - hyperkardioidní směrová charakteristika:  $\alpha = 60^\circ$  až  $120^\circ$
- **Blumleinova**: pár mikrofónů s osmičkovou směrovou charakteristikou, osy natočené o  $90^\circ$
- **MS** (Midcomponent – Side component):
  - 1. kardioidní mikrofón umístěný v ose zdroje zvuku
  - 2. mikrofón s osmičkovou charakteristikou orientovaný stranově

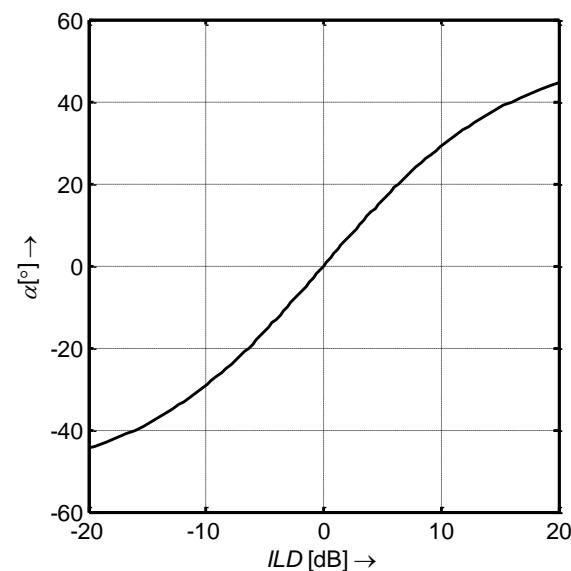
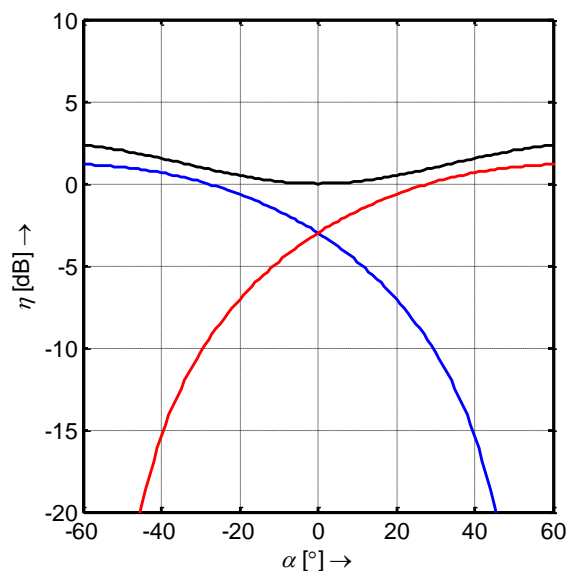
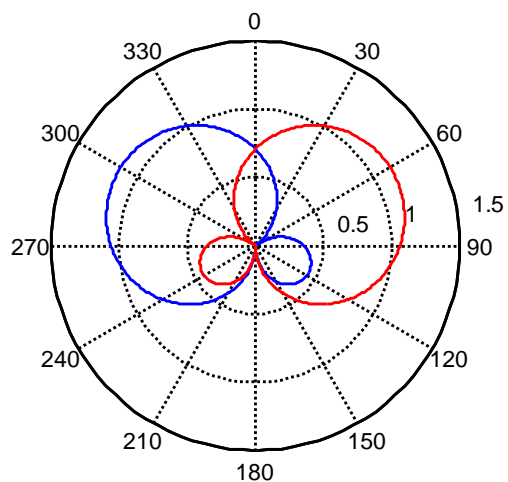
# Snímání XY – citlivost kanálů a ILD pro $\alpha = 90^\circ$



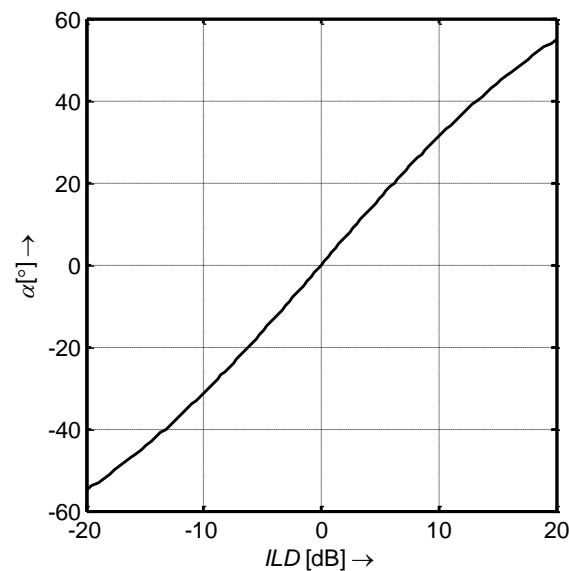
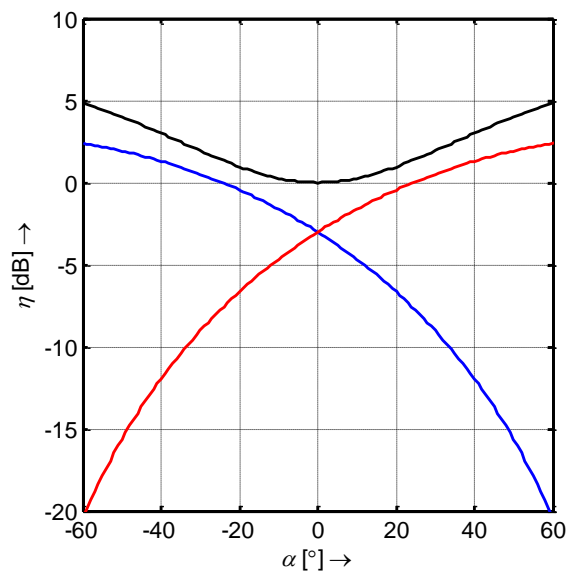
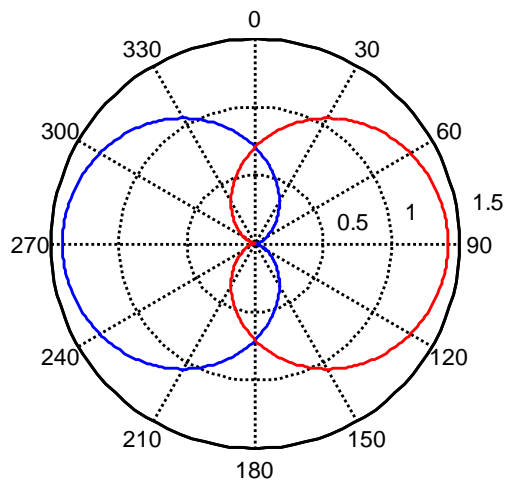
# Snímání XY – citlivost kanálů a ILD pro $\alpha = 120^\circ$



# Technika MS – kardioidní mikrofon M

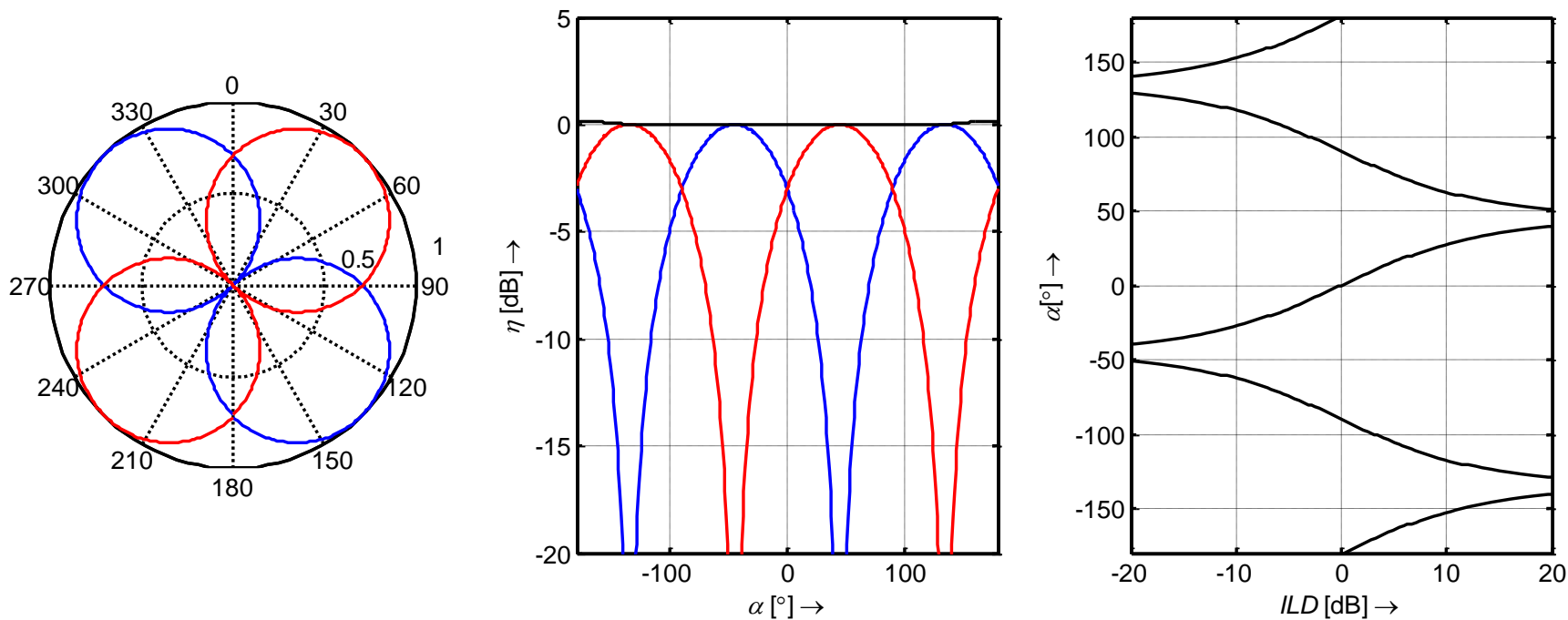


# Technika MS – všesměrový mikrofon M





# Blumleinova technika – citlivost kanálů a ILD



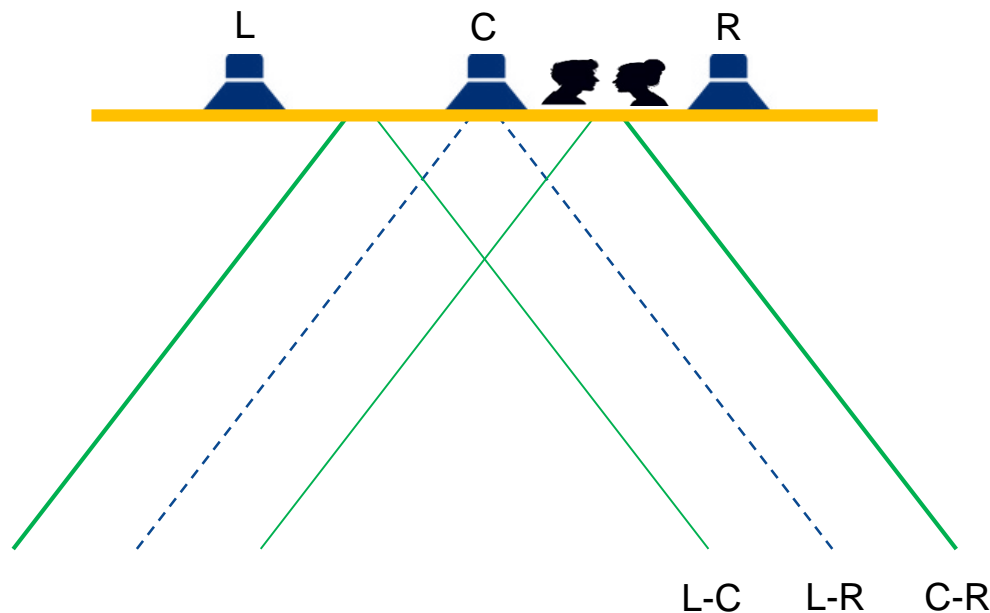
# Prostorová reprodukce zvuku

# Prostorová reprodukce zvuku



# Prostorová reprodukce zvuku

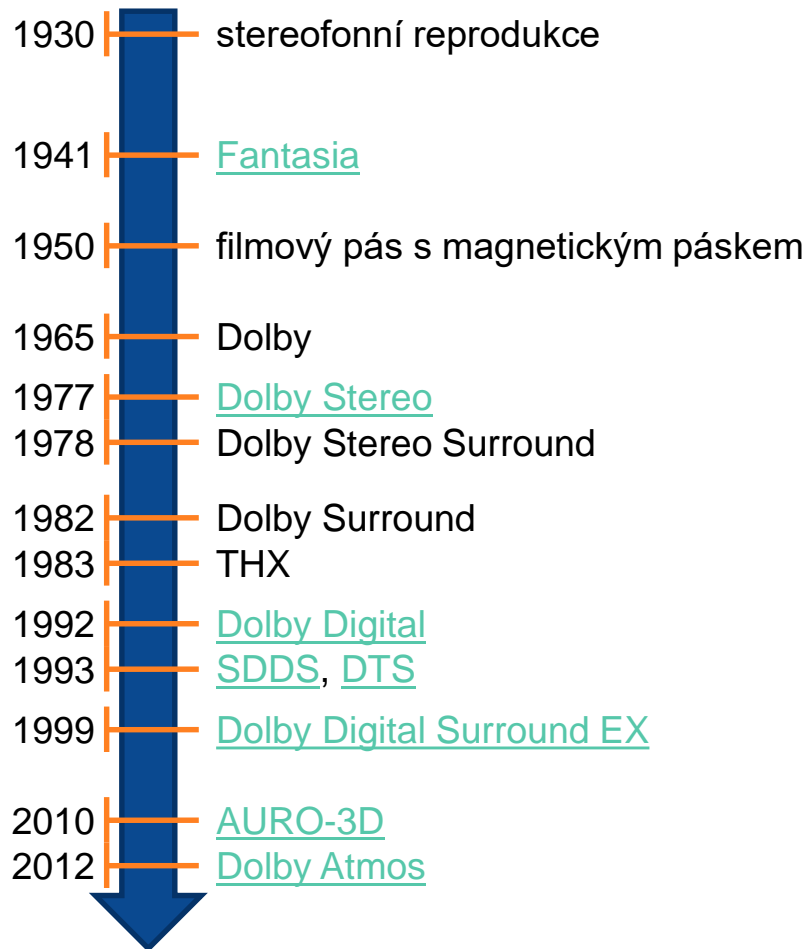
- Požadavky:
  - věrná reprodukce směru přicházejícího zvuku ve frontální části poslechového pole za účelem dobré lokalizace dialogů na plátně
  - dojem zvukového pole obklopujícího posluchače ve zbývající části prostoru



# Prostorová reprodukce zvuku

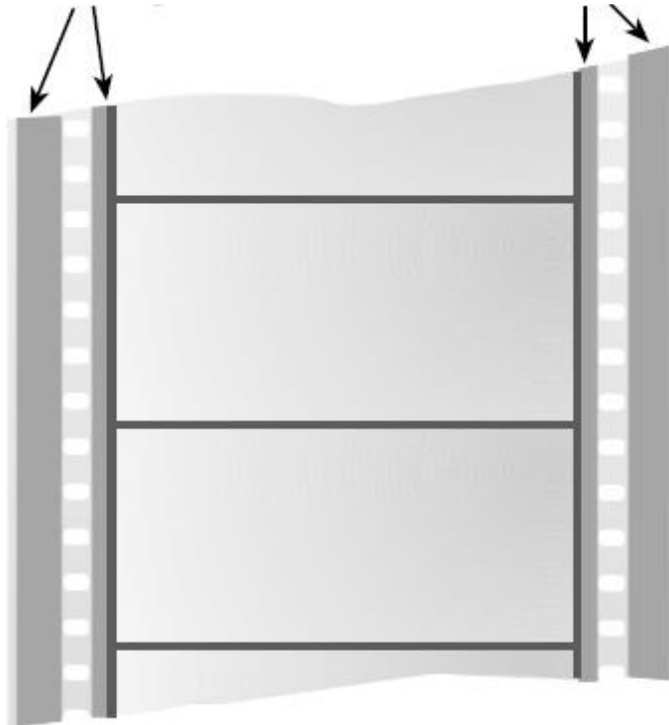
- Definice pojmů:
  - Označení ***X.Y***:  $X$  = počet kanálů systému s plným spektrem,  $Y$  = počet kmitočtově omezených kanálů systému
  - Označení ***MIN***:  $M$  = počet předních kanálů (za plátnem),  $N$  = počet surroundových (ambientních) kanálů (po stranách a vzadu)
  - ***Downmix***: převod vyššího počtu kanálů na nižší počet
  - ***Upmix***: převod nižšího počtu kanálů na vyšší počet
  - ***Škálovatelnost***: schopnost systému reprodukovat vyšší počet zaznamenaných kanálů nižším počtem reprodukčních kanálů se zhoršením pouze prostorové informace
  - ***Immersive Sound***: skutečná rekonstrukce zvukového pole v místě posluchače (Isono Auro, Dolby Atmos)

# Historie

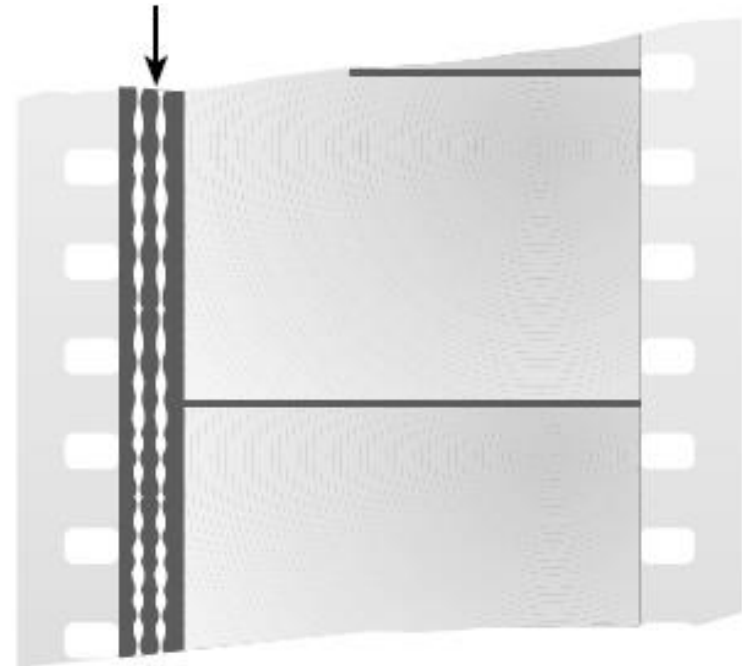


# Analogový záznam na filmový pás

magnetický záznam



optický záznam



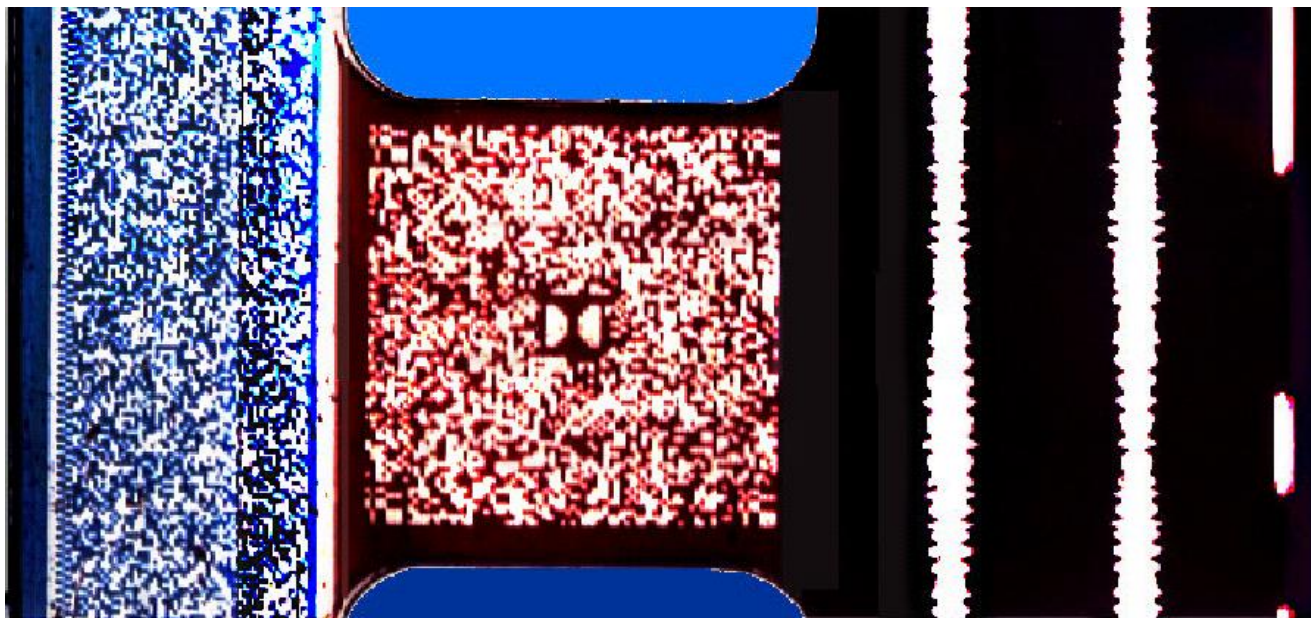
# Digitální záznam audia na filmový pás

SDDS

Dolby Digital

Dolby Stereo

DTS  
sync





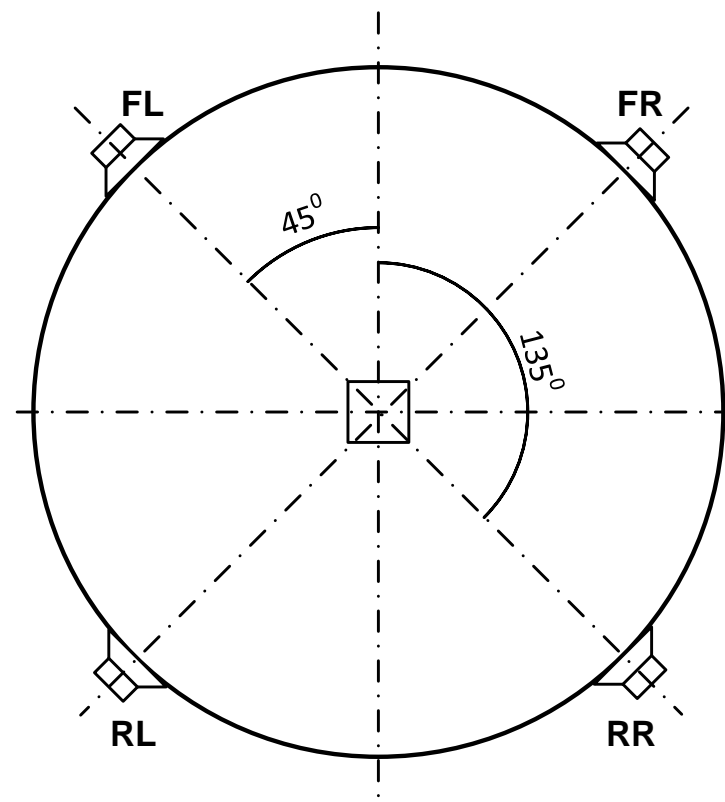
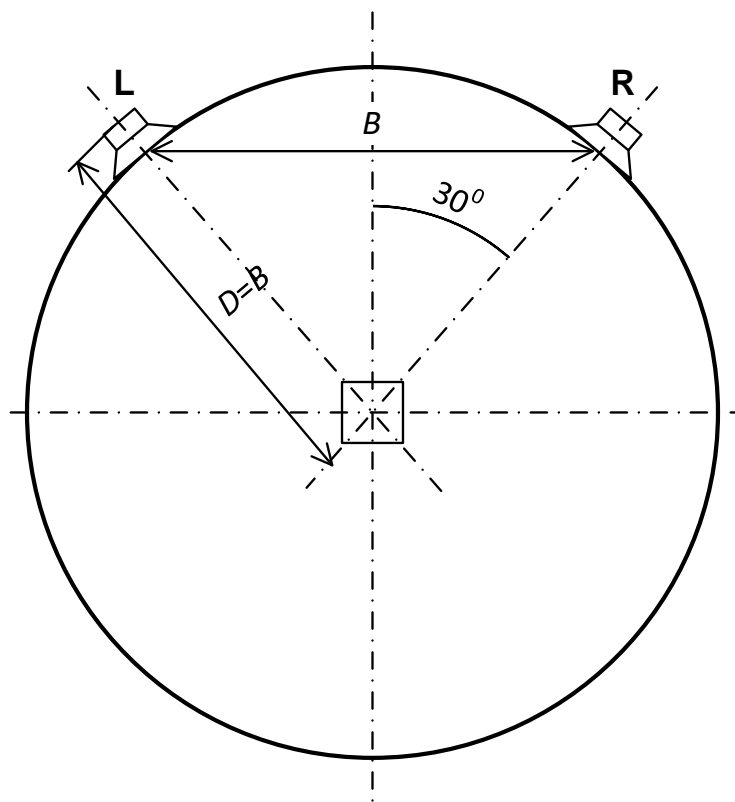
# Digitální distribuce

Digital Cinema Package: formát a distribuce digitálního zvuku a obrazu



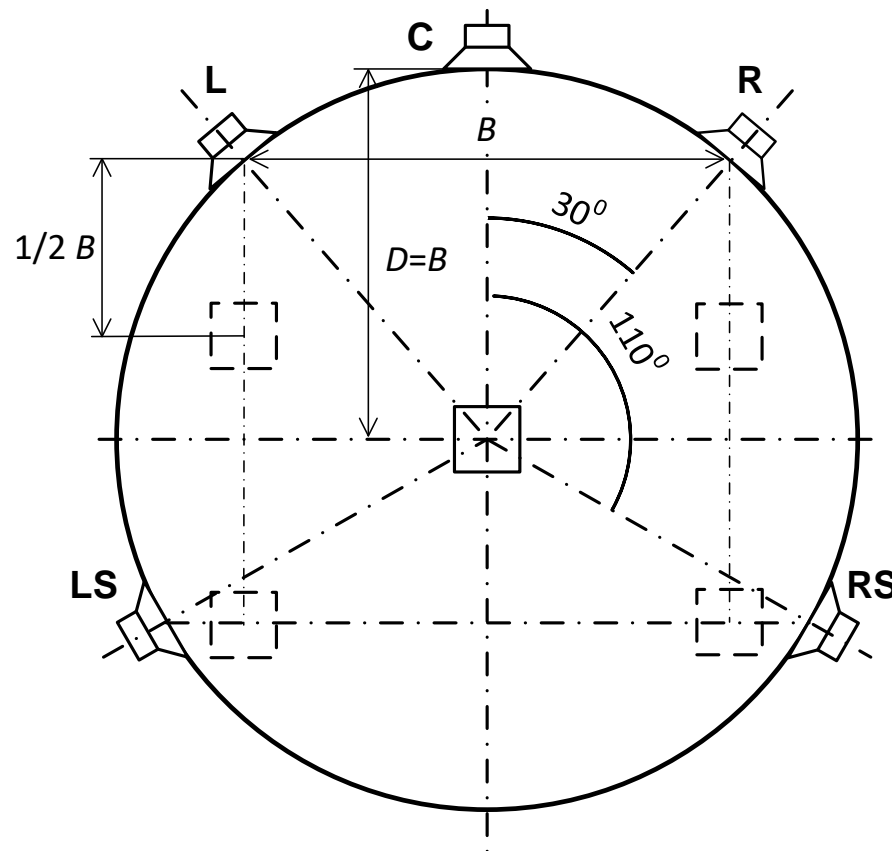
# Systémy horizontální reprodukce

- Pravidelné uspořádání – stereofonie, kvadrofonie



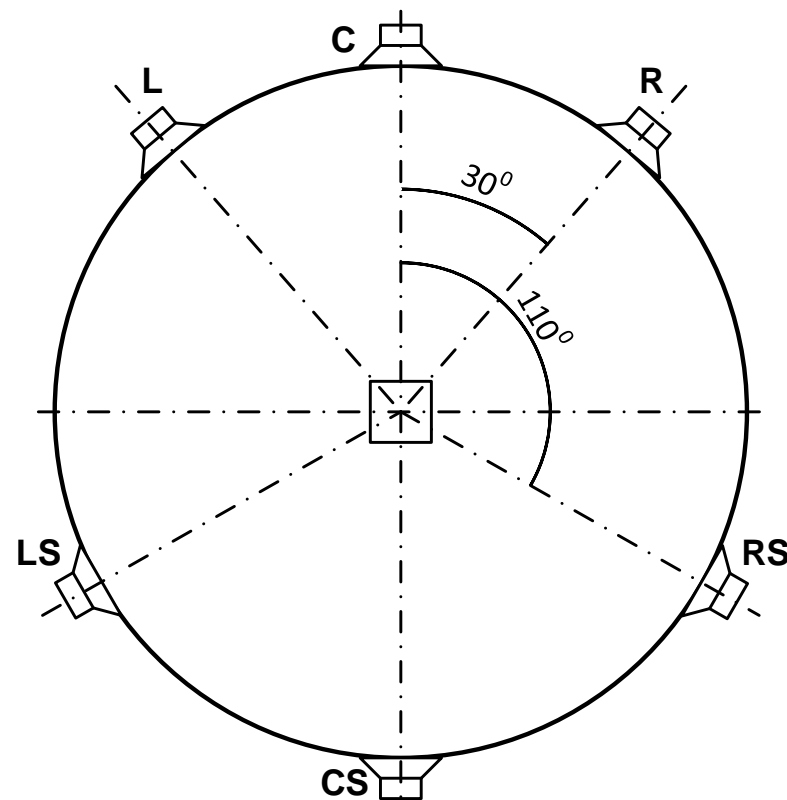
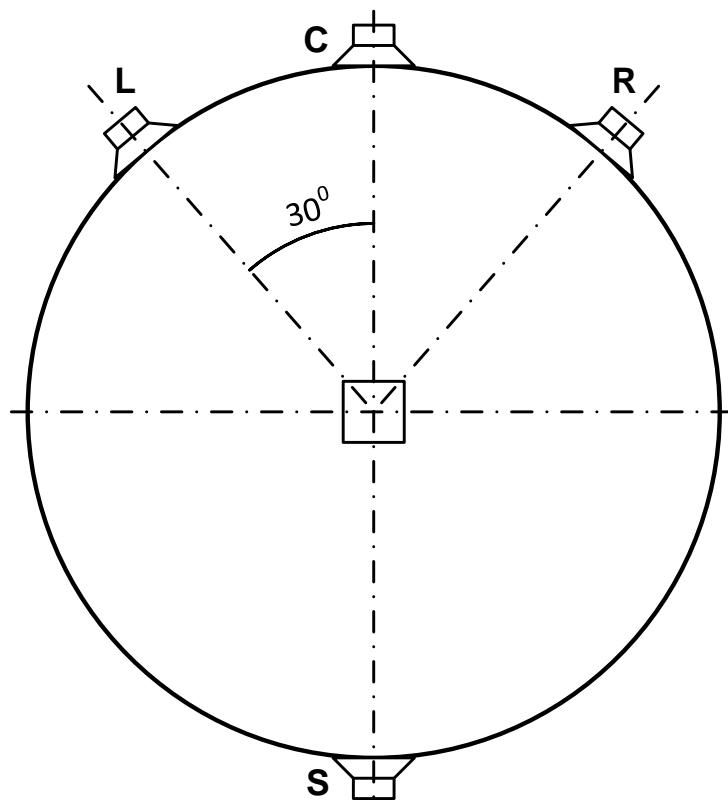
# Systémy horizontální reprodukce – ITU-R BS.775

- Vícekanálová stereofonní reprodukce s obrazem a bez.
- Pozice reproduktorů:
  - C –  $0^\circ$ ,
  - LS, RS –  $\pm 100^\circ$  až  $120^\circ$ ,
  - L, R –  $\pm 30^\circ$ .
- Platí i pro systémy monofonní a stereofonní systém a LRCS.
- LFE kanál do kmitočtu 120 Hz.



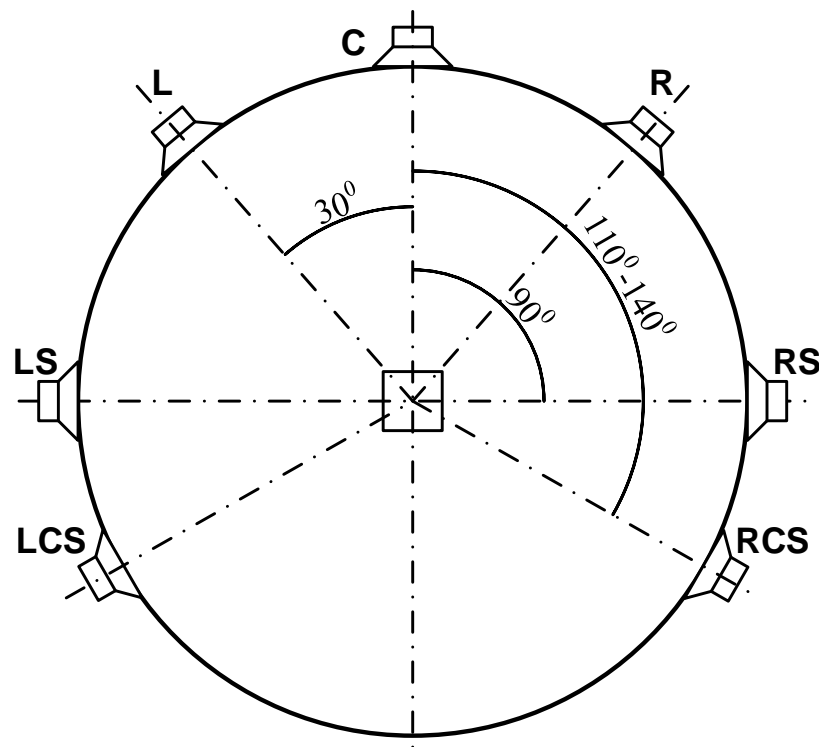
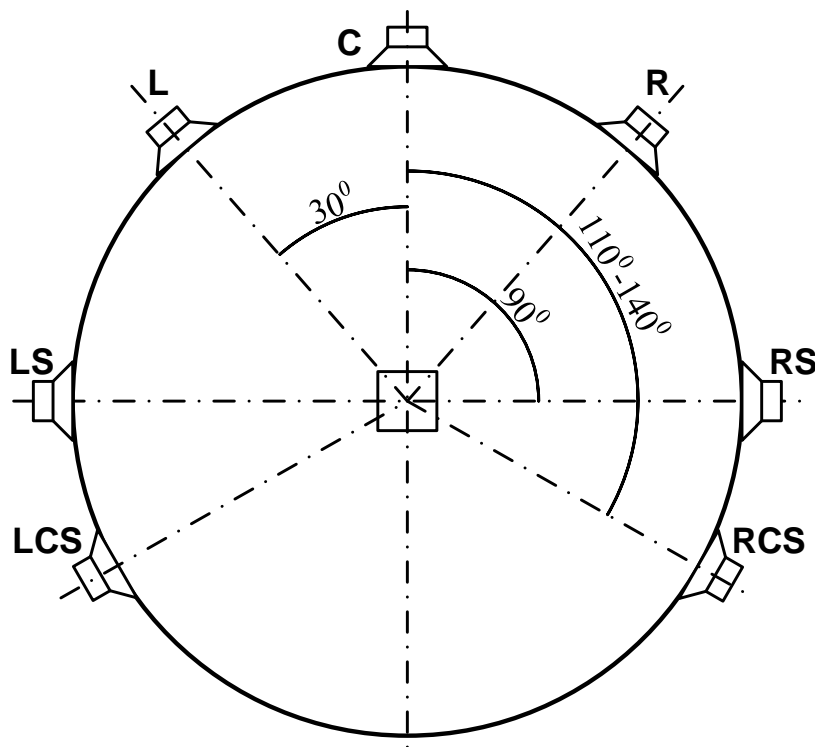
# Systémy horizontální reprodukce

- Nepravidelné uspořádání – 3/1 (LRCS), 3/3 surround



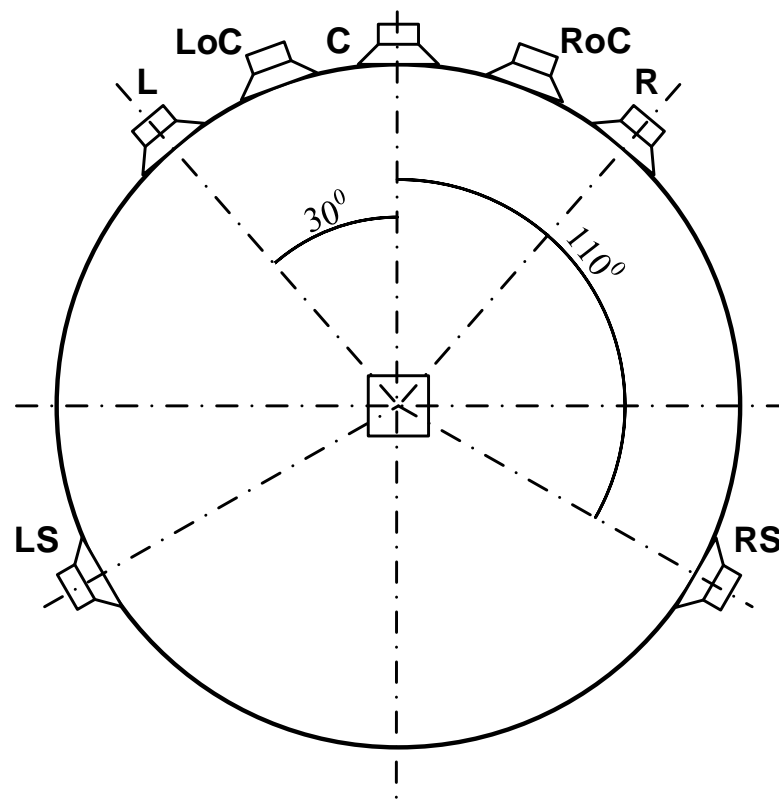
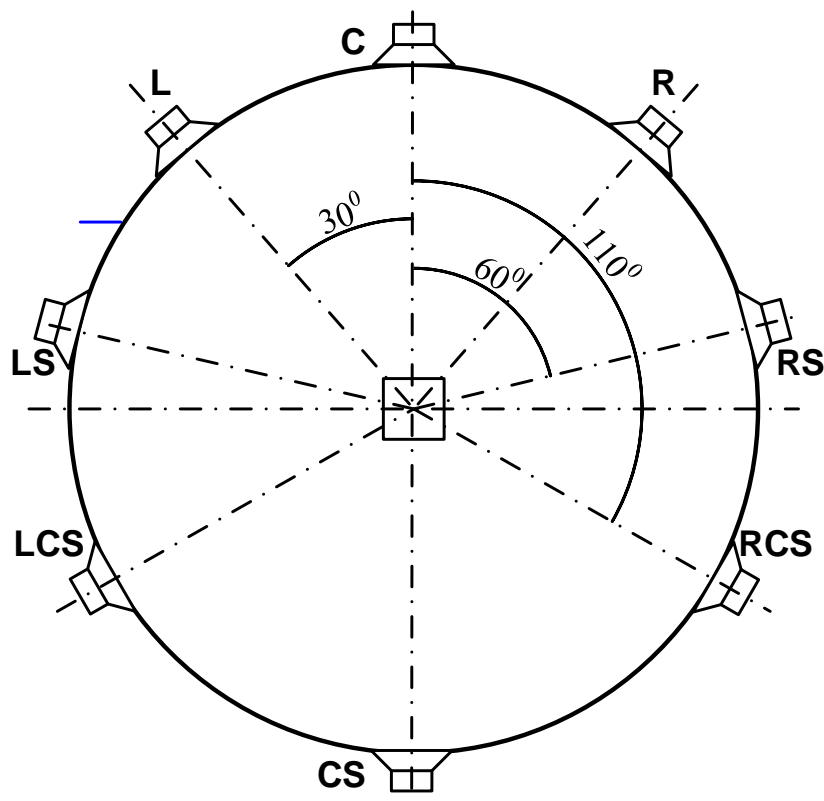
# Systémy horizontální reprodukce

- Nepravidelné uspořádání – 3/4 Logic 7, 3/4 Dolby 7.1



# Systémy horizontální reprodukce

- Nepravidelné uspořádání – 3/5 TMH, 5/2 SDDS



# Surround with Height

- Přidání dalších hladin reprodukce k horizontální hladině

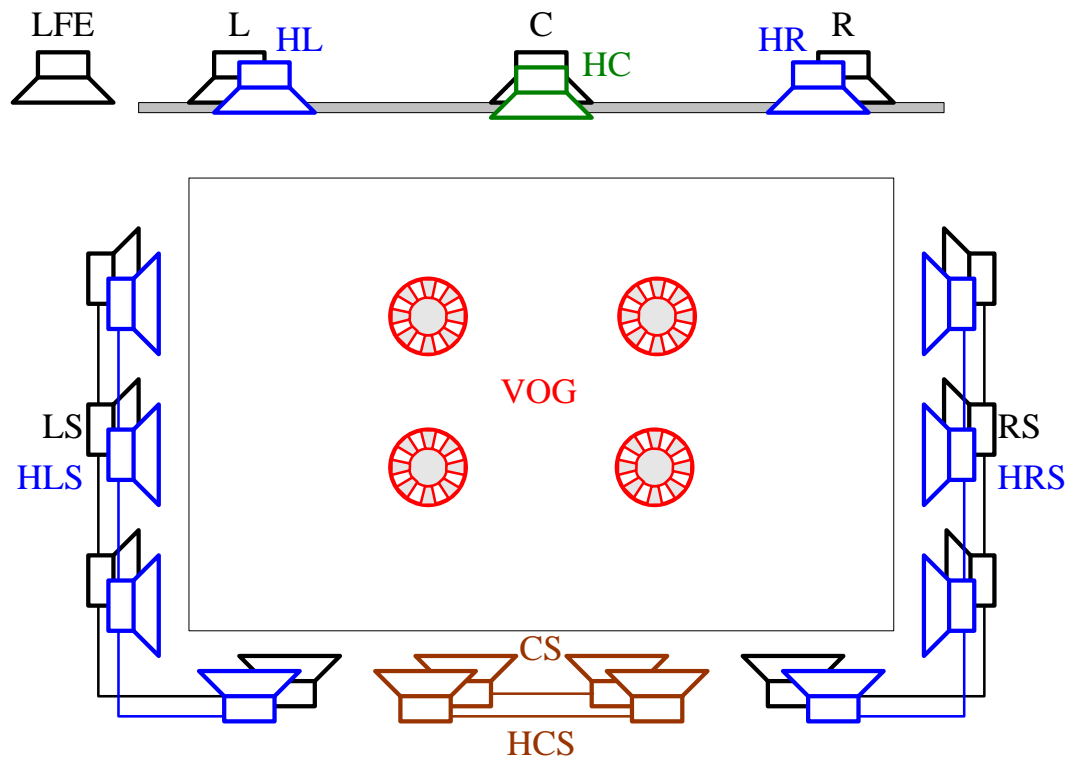


# Auro 3D

- Škálovatelný formát se třemi hladinami:
  - Lower layer – standardní horizontální rovina,
  - Height layer – úhel  $40^\circ$  vůči posluchači,
  - Top layer – nad posluchačem.
- Formáty:
  - 9.1 – L, C, R, LS, RS, HL, HR, HLS, HRS, LFE
  - 10.1 – 9.1 plus VOG
  - 11.1 – 10.1 plus HC
  - 13.1 – 11.1 plus CS a HCS



# Auro 3D



9.1 Lower layer

9.1 Height layer

Top layer (10.1, 11.1, 13.1)

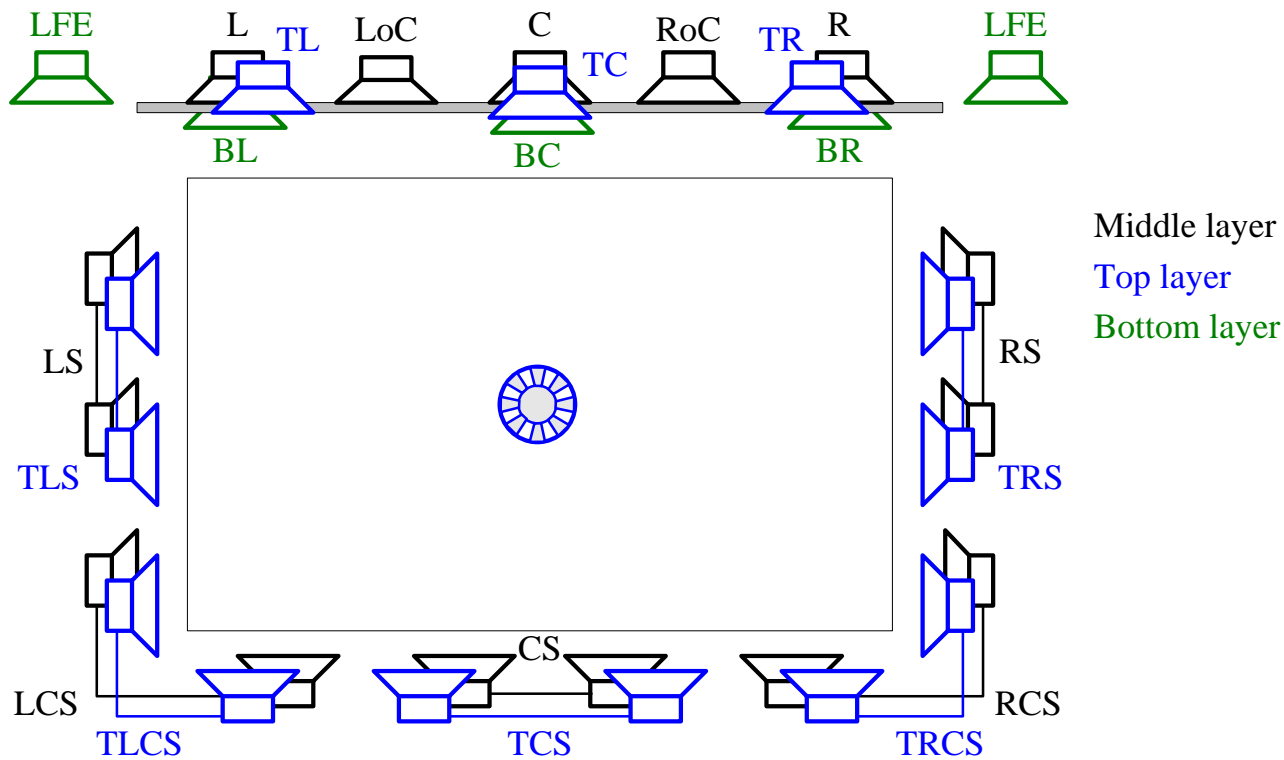
rozšíření 10.1 na 11.1

rozšíření 11.1 na 13.1

## NHK Super Hi-Vision 22.2

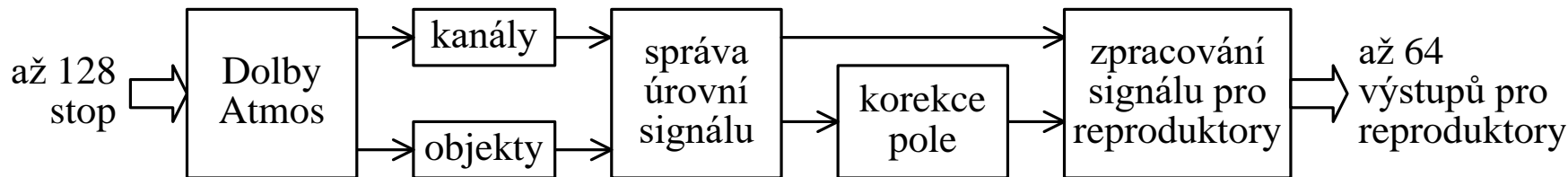
- Určeno pro HDTV Super Hi-Vision (7680x4320 bodů).
- Kanály rozděleny do 3 hladin:
  - střední: 10 kanálů, 5 předních jako SDDS, 5 surroundových jako TMH,
  - horní: 9 kanálů, stejné pozice jako kanály střední hladiny kromě levého a pravého středního, navíc kanál nad hlavou,
  - dolní: 3 kanály, levý, střední a pravý přední + 2 LFE.

# NHK Super Hi-Vision 22.2

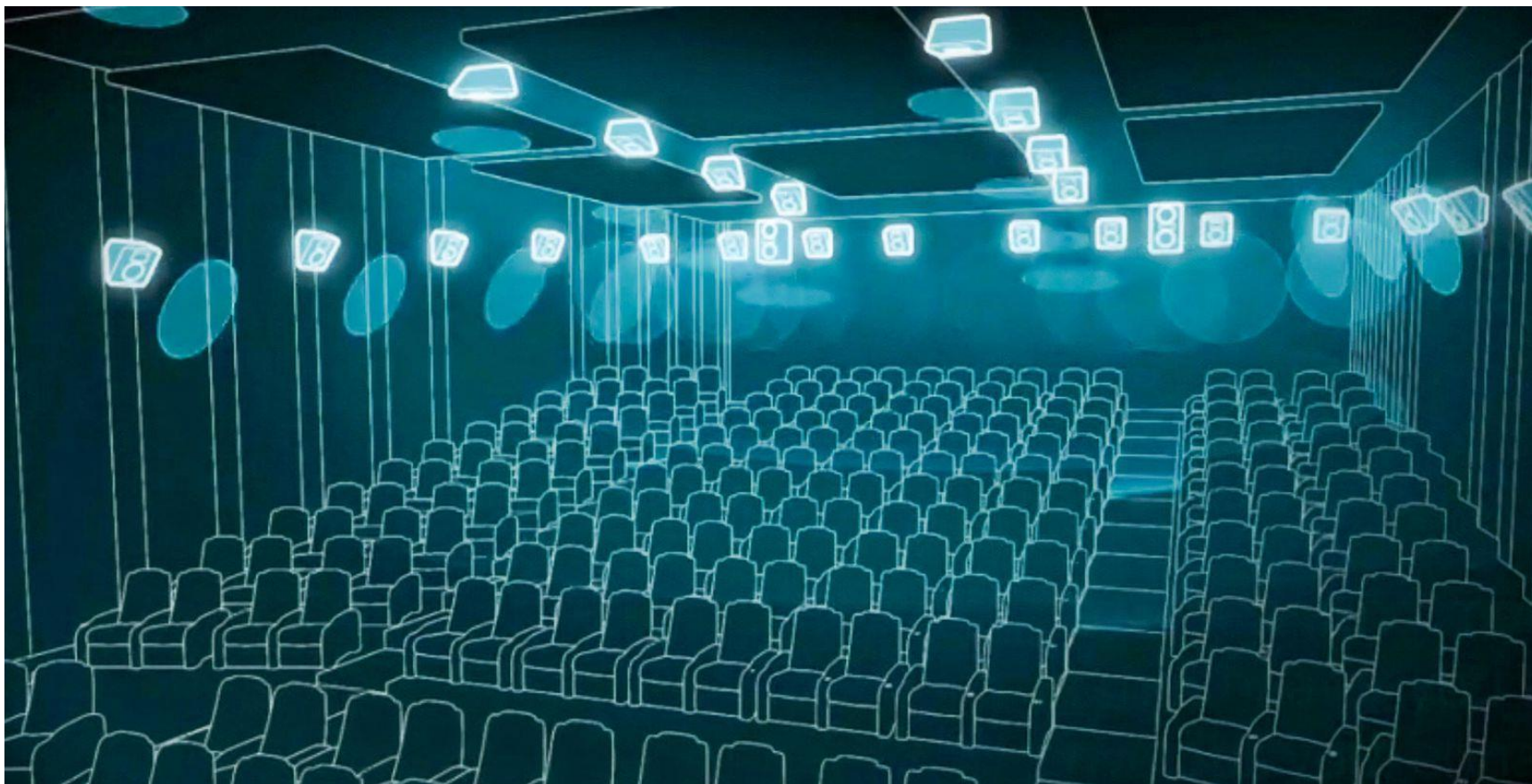


# Dolby Atmos

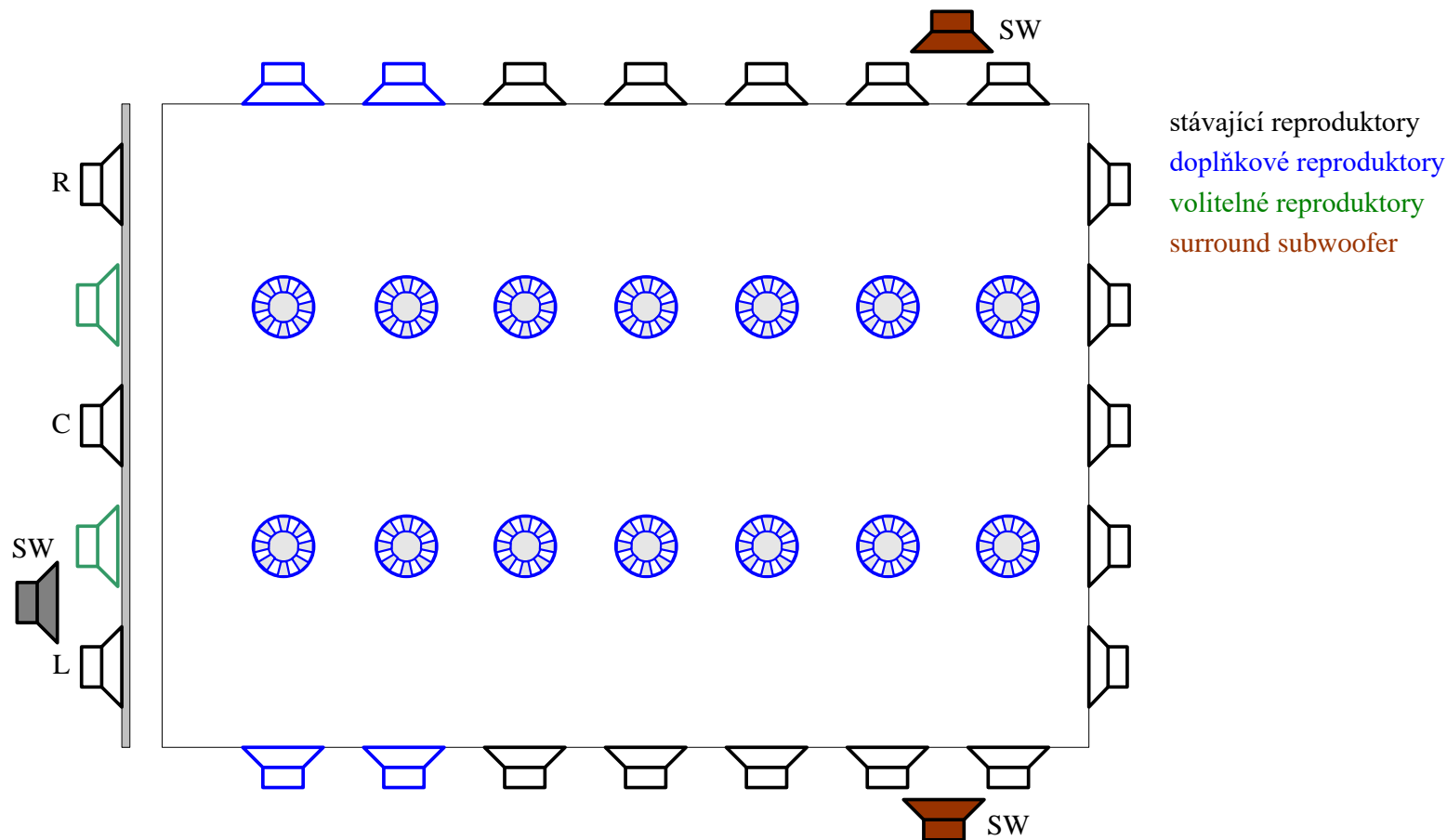
- Dolby Laboratories 2012
- Overhead Sound – hladina reprodukce nad hlavou posluchače
- Možnost použití ekvalizace místnosti
- Surround bass management – směřování nf signálu surroundových kanálů do speciálních subwooferů



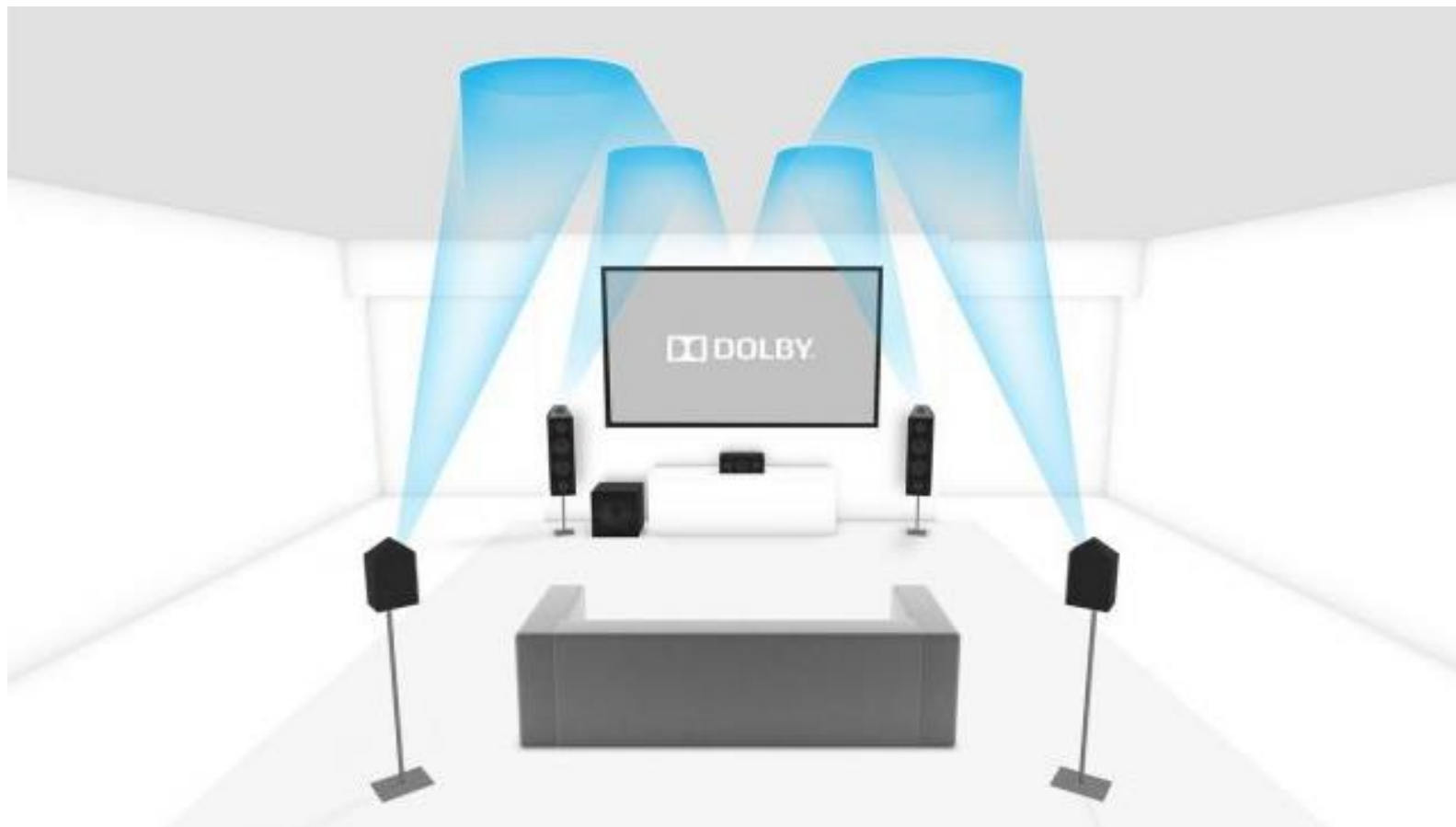
# Dolby Atmos



# Dolby Atmos

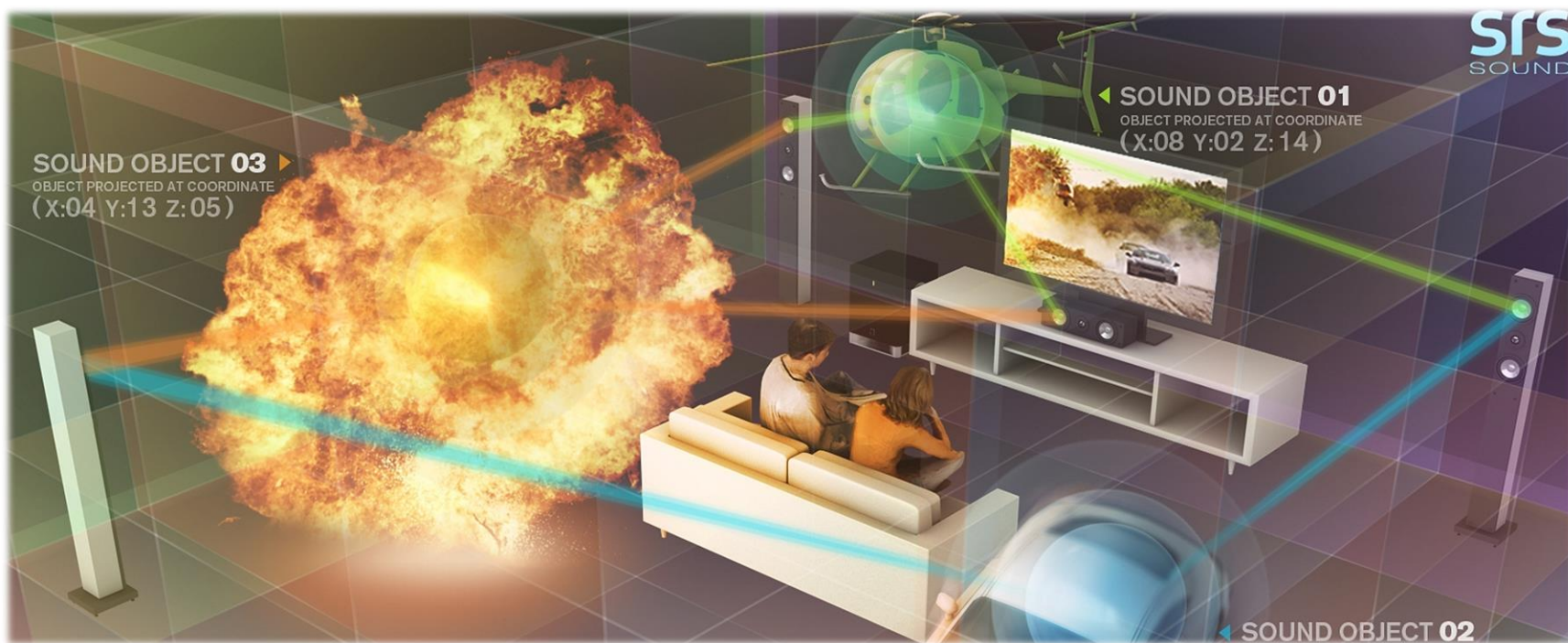


# Dolby Atmos



# Objektově založená reprodukce

- **Zvukový objekt:** zvukový signál, který je reprodukován tak, jako by pocházel z konkrétního cílového místa, které je určeno doprovodnou informací



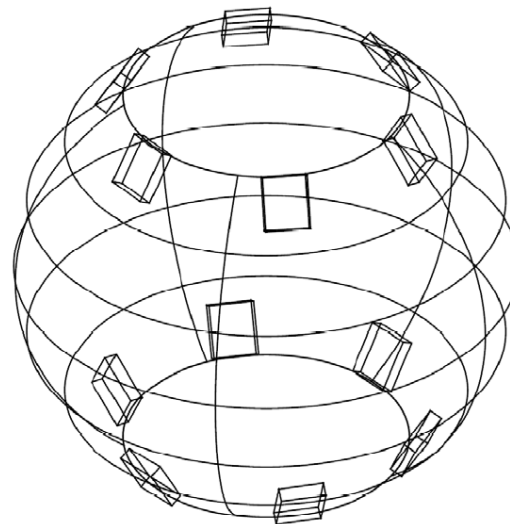
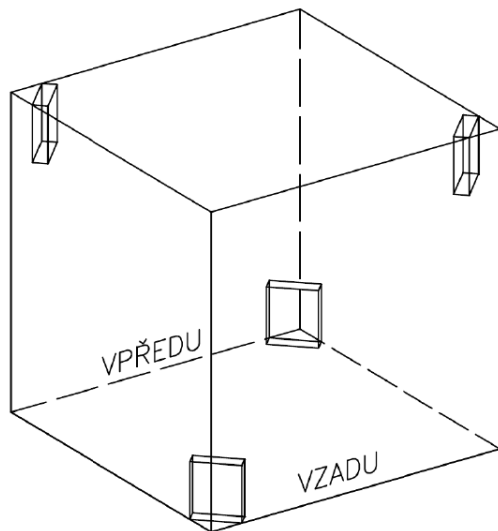


# Objektově založená reprodukce

- Skutečná pozice zvukových objektů se může v průběhu času měnit
- Tato pozice nemusí být definována předem, ale až v okamžiku reprodukce podle pozic reproduktorových soustav v poslechovém prostoru
- Definice pozice může zahrnovat interakci s uživatelem
- Komerční systémy:
  - Iosono Auro 3D
  - Dolby Atmos
- Standardy
  - MPEG-D
  - MPEG-H

# Systémy založené na matematickém popisu pole

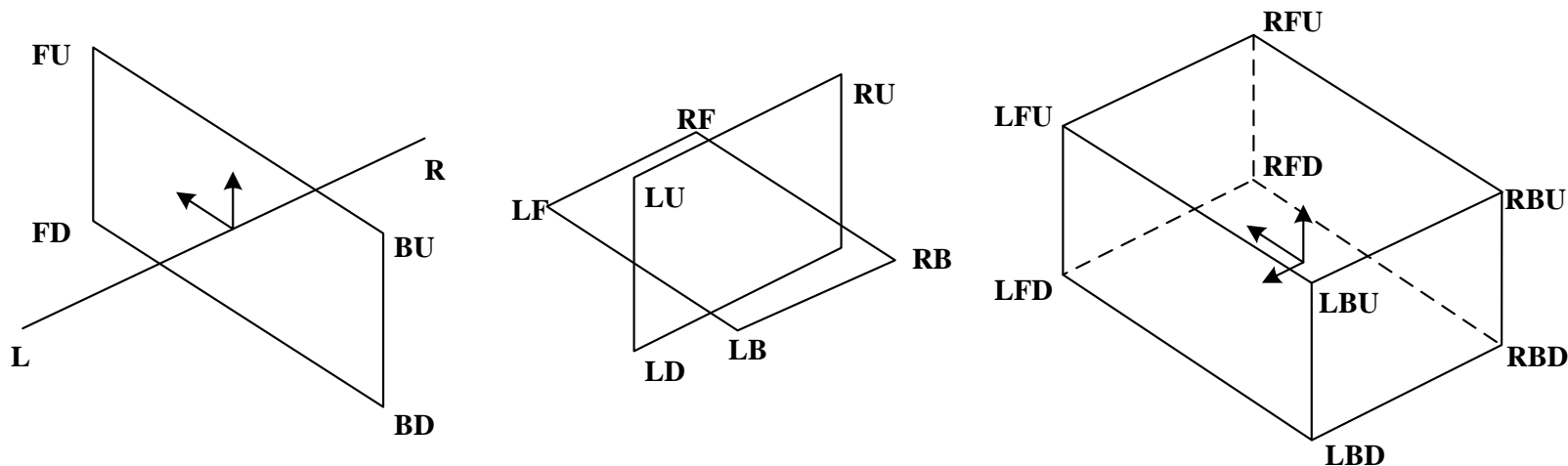
- **Perifonie:** souhrnné označení systémů schopných reprodukce směrového vjemu v horizontální i vertikální rovině
  - nejjednodušší případ: rozestavění reproduktorů do vrcholů čtyřstěnu
  - pro lepší reprodukci vysokých kmitočtů je vhodné použít 8 a více reproduktorů rozdělených alespoň do dvou úrovní, pod a nad horizontální rovinou poslechu



# Systémy založené na matematickém popisu pole

- **maticové systémy:** popsány kódovací maticí pro  $m$  vstupních signálů a  $n$  výstupních signálů, tj. konečným počtem proměnných (např. Ambisonie)
- **jádrové (kernel) systémy:** popsány nekonečným počtem proměnných, tj. popisují funkci na spojitě ploše kolem posluchače

pravidelné uspořádání zdrojů zvuku + konvence značení zdrojů



# Obecná meta-teorie sluchové lokalizace

- 1992 Michael A. Gerzon
- Člověk používá různé mechanismy pro sluchovou lokalizaci
- Kromě případů, kdy jsou jednotlivé sluchové podněty silně v konfliktu, je celkový sluchový dojem dán většinou souhlasných podnětů
- Hierarchie modelů
  - hierarchie nelinearity
  - hierarchie směrovosti
  - hierarchie pohybu hlavou

# Obecná meta-teorie sluchové lokalizace

- **Hierarchie směrovosti:** směr přicházejícího zvuku je určen tzv. směrovými kosiny (funkce souřadnic zdroje zvuku  $x, y, z$ )
- každý z modelů určuje vektor lokalizace, jehož směr určuje předpokládaný směr zvuku a jeho velikost popisuje stabilitu lokalizace
- sluch musí pro lokalizaci zdroje zvuku používat více těchto modelů současně
- pokud mají různé modely různé vektory lokalizace, je o směru přicházejícího zvuku rozhodnuto podle souhlasných výsledků nejvíce modelů

# Obecná meta-teorie sluchové lokalizace

- **Rychlostní model** (Velocity Model, model akustické rychlosti): model 1. stupně 1. řádu, určuje směr přicházejícího zvuku na základě akustického tlaku a složek akustické rychlosti

$$r_V \mathbf{r}_V = \operatorname{Re} \frac{\sum_{i=1}^n g_i \mathbf{l}_i}{\sum_{i=1}^n g_i}$$

$r_V$  – velikost vektoru rychlosti,  $\mathbf{r}_V$  – směr vektoru rychlosti,  $g_i$  – komplexní zesilovací činitele jednotlivých reproduktorů,  $\mathbf{l}_i$  – jednotkový vektor ve směru  $i$ -tého reproduktoru,  $n$  – počet reproduktorů

- imaginární část zlomku představuje fázové zkreslení (ve směru osy spojující uši by mělo v rozsahu kmitočtů 300 až 1500 Hz být minimální)
- vektor rychlosti určený tímto modelem koreluje s IPD

# Obecná meta-teorie sluchové lokalizace

- **Energetický model** (Energy Model, model proudění energie): model 2. stupně 1. řádu, určuje směr přicházejícího zvuku na základě akustického tlaku a složek akustické rychlosti

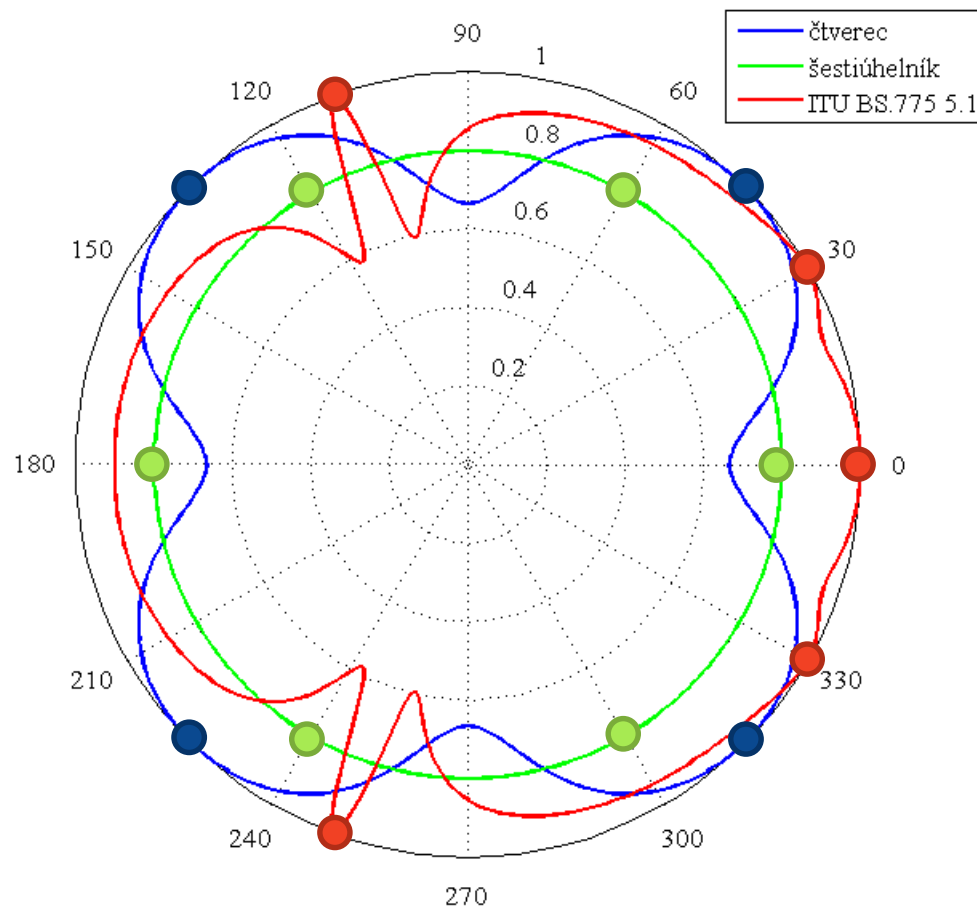
$$r_E \mathbf{r}_E = \frac{\sum_{i=1}^n (g_i g_i^*) \mathbf{l}_i}{\sum_{i=1}^n (g_i g_i^*)}$$

$r_E$  – velikost vektoru energie,  $\mathbf{r}_E$  – směr vektoru energie, \* značí komplexní sdružení

- vektor energie určený tímto modelem koreluje s IID.

# Závislost velikosti vektoru energie na azimutu

- pozice
- reproduktorů





# Obecná meta-teorie sluchové lokalizace

- **Požadavky na dekodéry prostorové reprodukce zvuku:**
  1. Směr vektoru rychlosti a vektoru energie musí být souhlasný alespoň do kmitočtu okolo 4 kHz s tím, že se azimut reprodukováného zdroje zvuku  $\varphi_V = \varphi_E$  výrazně nemění s kmitočtem
  2. na nízkých kmitočtech okolo 400 Hz musí být velikost vektoru rychlosti blízko hodnotě 1 pro všechny azimuty reprodukováných zdrojů zvuku
  3. na středních a vysokých kmitočtech okolo 700 Hz až 4 kHz musí být velikost vektoru energie  $r_E$  maximalizována na co největší části poslechového prostoru
- nejlepší lokalizace pro danou reproduktorovou konfiguraci nastává, pokud je
  - velikost vektoru rychlosti na nízkých kmitočtech rovna 1
  - velikost vektoru energie maximalizována na středních kmitočtech
  - přechod mezi těmito dvěma pásmy je mezi 300 Hz a 700 Hz.
- pokud je vektor rychlosti a energie reprodukováného zvukového zdroje shodný s vektorem rychlosti a energie reálného zdroje, jejich vjem je stejný

# Kódování prostorové informace

# Kódování stereofonního signálu

- Mezi výkonovým spektrem levého a pravého kanálu je silná korelace
- Vlastnosti zvuku v levém a pravém uchu jsou porovnávány v rámci kritických pásem
- Na vyšších kmitočtech se schopnost rozlišit pozici zdroje snižuje
- Nad 2 kHz je v rámci kritického pásma stereofonní vjem dán směrem přicházející energie
- Metody:
  - MS stereo: redukce rozdílové informace mezi kanály (side component)
  - Intensity stereo: odstranění irelevantní informace na vysokých kmitočtech
  - kombinace

# MS Stereo

1. Převod signálu levého a pravého kanálu na součtový a rozdílový signál:

$$x_M(t) = \frac{x_L(t) + x_R(t)}{\sqrt{2}}$$

$$x_S(t) = \frac{x_L(t) - x_R(t)}{\sqrt{2}}$$

2. Redukce datového toku složky S

- zeslabení
- větší kvantovací krok
- ...

3. Převod zpět na levý a pravý kanál stereofonního signálu:

$$x_L(t) = \frac{x_M(t) + x_S(t)}{\sqrt{2}}$$

$$x_R(t) = \frac{x_M(t) - x_S(t)}{\sqrt{2}}$$

# Intensity stereo

- Výpočet azimutu v každém kritickém pásmu (např. pomocí sinového nebo tangentového zákona)

$$\frac{\sin \alpha}{\sin \alpha_0} = \frac{E_L - E_R}{E_L + E_R} \quad \frac{\tan \alpha}{\tan \alpha_0} = \frac{E_L - E_R}{E_L + E_R} \quad E = \sum_{k=k_{\min}}^{k_{\max}} |X(k)|^2 = \sum_{n=0}^{N-1} x(n)^2$$

$\alpha$  – azimut zdroje zvuku,  $\alpha_0$  – azimut pozice reproduktoru ( $30^\circ$ ),  $E_L$ ,  $E_R$  – energie levého a pravého kanálu v kritickém pásmu,  $k_{\min}$ ,  $k_{\max}$  – indexy první a poslední spektrální složky kritického pásma,  $x(n)$  – vzorky signálu v kritickém pásmu,  $N$  – délka signálu

- Downmix levého a pravého kanálu do monofonního v každém kritickém pásmu

$$x_M(t) = x_L(t) + x_R(t)$$

# Intensity stereo

- Zpětné dekódování azimutu:

$$\frac{\sin \alpha}{\sin \alpha_0} = \frac{g_L - g_R}{g_L + g_R} \quad \frac{\tan \alpha}{\tan \alpha_0} = \frac{g_L - g_R}{g_L + g_R} \quad g_L^2 + g_R^2 = 1$$

$g_L, g_R$  – zesílení levého a pravého kanálu v kritickém pásmu

- Upmix do levého a pravého kanálu

$$x_L(t) = g_L x_M(t) \quad x_R(t) = g_R x_M(t)$$

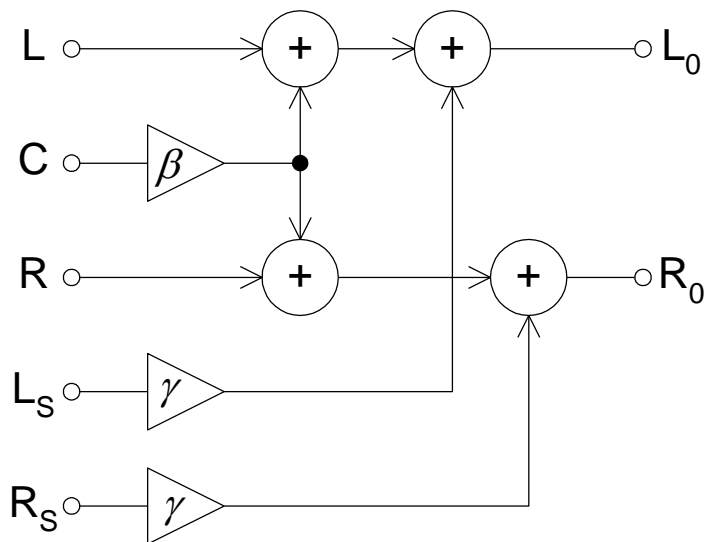
- Vlastnosti:
  - vhodné pouze pro vysoké kmitočty
  - částečná ztráta prostorové informace
  - možnost rozšíření na vícekanálový formát (coupling channels, dynamic crosstalk, generalized intensity coding): spárování kanálů s korelovaným výkonovým spektrem a přenos v režimu Intensity stereo

# Kódování vícekanálových formátů

- Maticové kódování
  - vyšší počet vstupních kanálů než zakódovaných, dekodování pomocí pseudoinverze matice
  - stejný počet vstupních a zakódovaných signálů, dekodování pomocí inverze matice, škálovatelnost
- Binaural Cue Coding
  - vyšší počet vstupních kanálů než zakódovaných, dekodování pomocí prostorových parametrů

# Maticové kódování bez redukce počtu kanálů

- Maticový kodér:



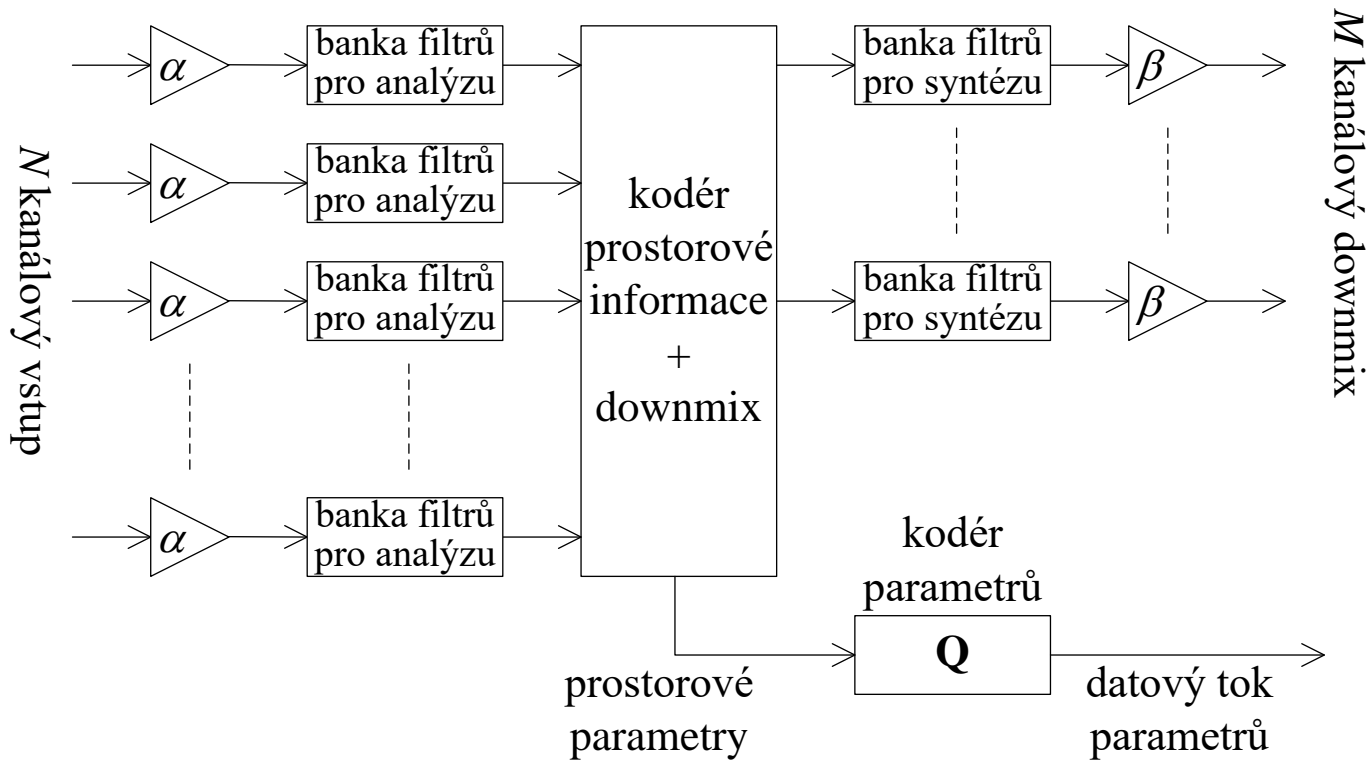
$$\begin{bmatrix} L_0 \\ R_0 \\ T_3 \\ T_4 \\ T_5 \end{bmatrix} = \mathbf{C}_M \begin{bmatrix} L \\ R \\ C \\ LS \\ RS \end{bmatrix}$$

$$\mathbf{C}_M = \alpha \begin{bmatrix} 1 & 0 & \beta & \gamma & 0 \\ 0 & 1 & \beta & 0 & \gamma \\ 0 & 0 & \beta & 0 & 0 \\ 0 & 0 & 0 & \gamma & 0 \\ 0 & 0 & 0 & 0 & \gamma \end{bmatrix}$$

LFE kanál je přenášen samostatně se vzorkovacím kmitočtem 96x nižším, do  $L_0 R_0$  není zahrnut



# Pásmový maticový kodér



# Dekódování pomocí inverzní matice

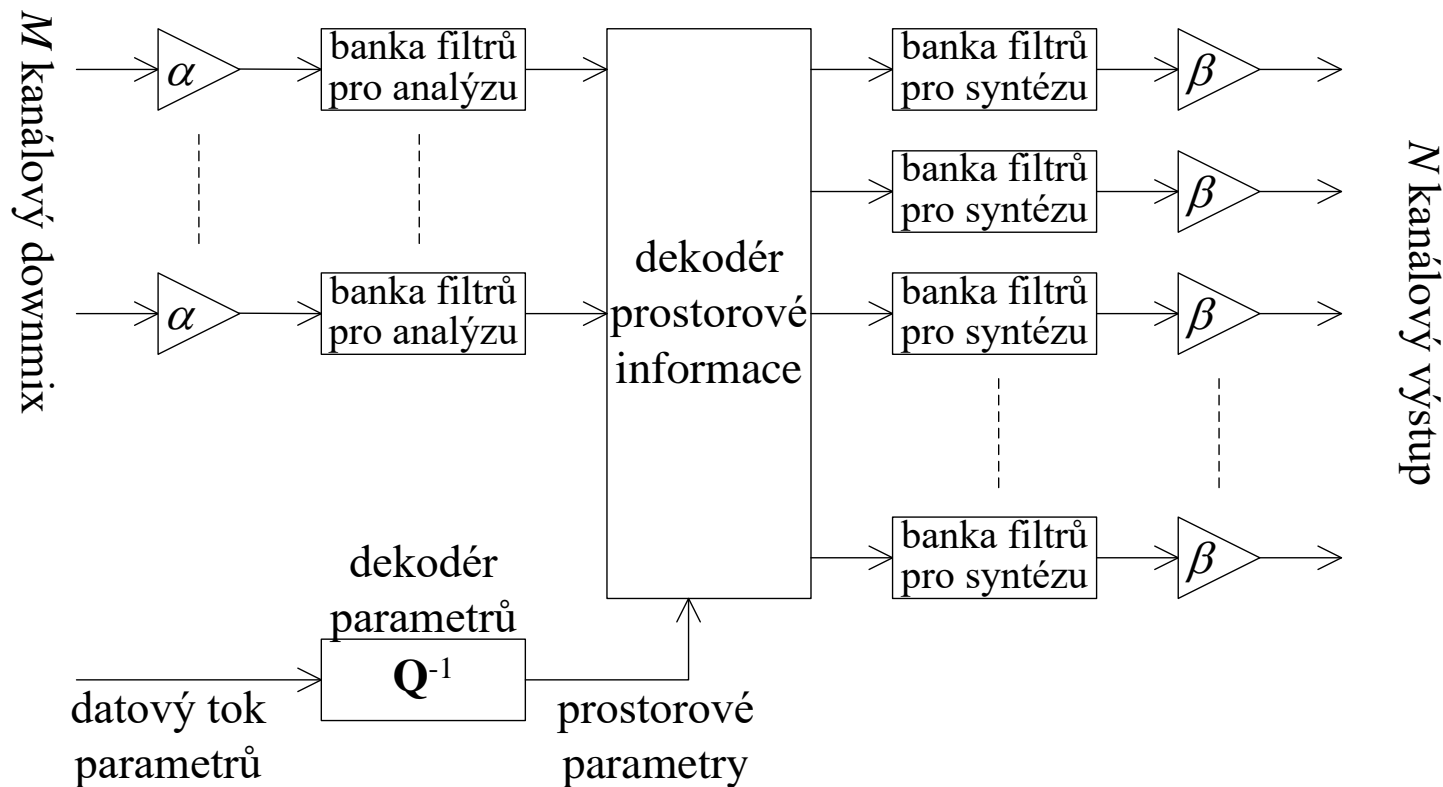
- Maticový dekodér

$$\begin{bmatrix} L \\ R \\ C \\ LS \\ RS \end{bmatrix} = \mathbf{D}_M \begin{bmatrix} L_0 \\ R_0 \\ T_3 \\ T_4 \\ T_5 \end{bmatrix}$$

$$\mathbf{D}_M = \mathbf{C}_M^{-1} = \frac{1}{\alpha} \begin{bmatrix} 1 & 0 & -1 & -1 & 0 \\ 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & 1/\beta & 0 & 0 \\ 0 & 0 & 0 & 1/\gamma & 0 \\ 0 & 0 & 0 & 0 & 1/\gamma \end{bmatrix}$$

režim	$\alpha$	$\beta$	$\gamma$
1	$1/(1+\sqrt{2})$	$1/\sqrt{2}$	$1/\sqrt{2}$
2	$2/(3+\sqrt{2})$	$1/\sqrt{2}$	$1/2$
3	1	1	1

# Pásmový maticový dekodér

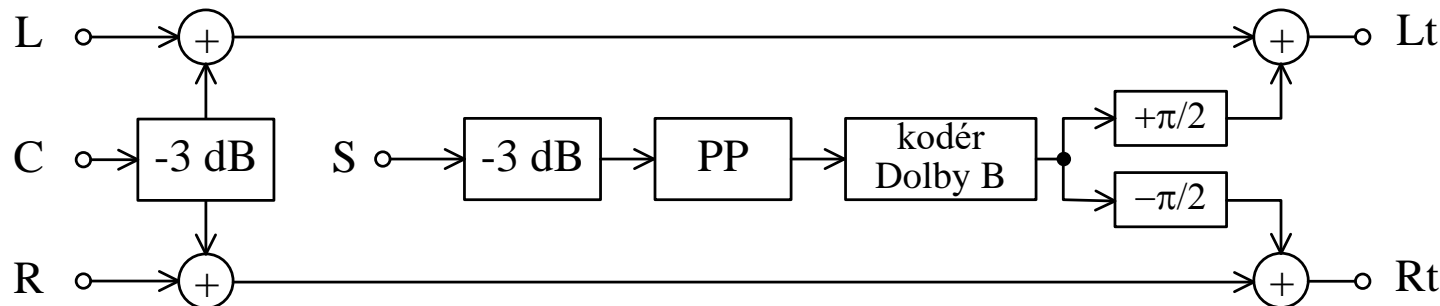


# Maticové kódování pomocí MP matice

- MP (Motion Picture) matice: kódování 4 kanálu do 2 (Dolby Surround)
  - vstupní kanály: levý, pravý, centrální, surround
  - výstupní kanály: Left total, Right Total

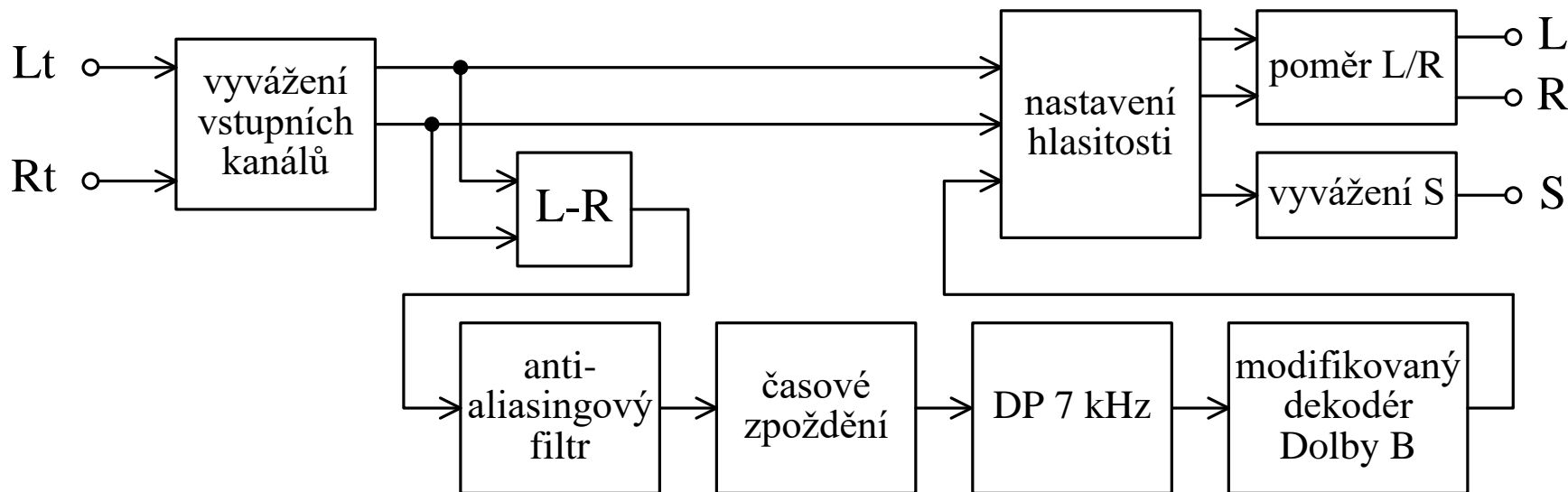
$$[L_t \quad R_t] = \mathbf{MP} \cdot \begin{bmatrix} L \\ R \\ C \\ S \end{bmatrix}$$

$$\mathbf{MP} = \begin{bmatrix} 1 & 0 & 1/\sqrt{2} & +j \\ 0 & 1 & 1/\sqrt{2} & -j \end{bmatrix}$$



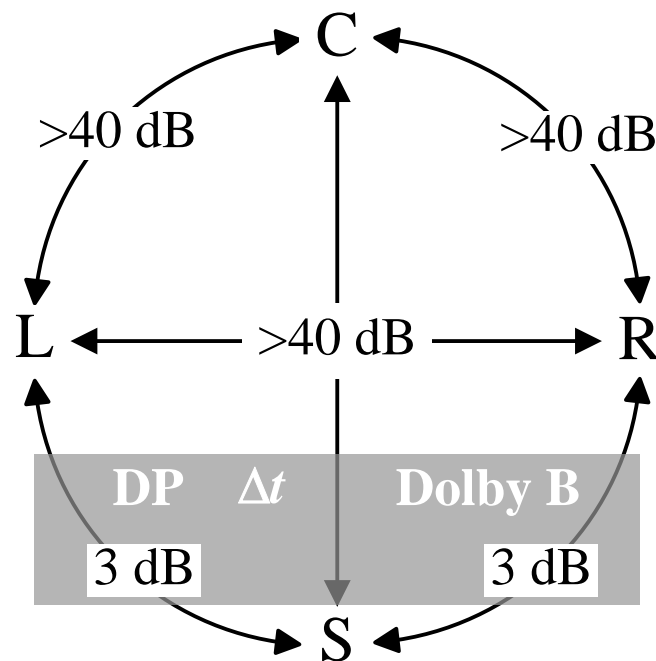
# Dekodér Dolby Surround

- Centrální kanál vytvořen fantomově
- Surroundový kanál se v L a R akusticky vyruší, do surroundového výstupu je částečně dekódován na základě dynamických změn stranové složky L a R



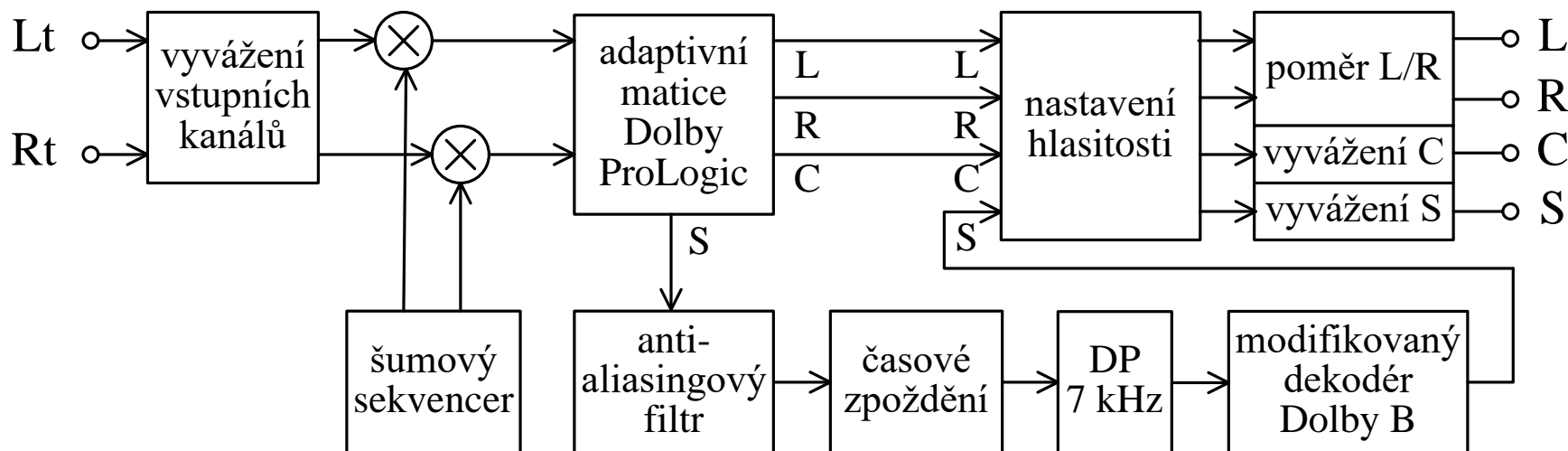
# Dekodér Dolby Surround

- Princip rozdílového kódování:
  - dominantní zvukovou složkou jsou dialogy, umisťované centrálně
  - surroundovým reproduktorem je přehráván signál, pokud je jeho intenzita v jednom z kanálů vyšší
- Separace Surroundového kanálu:
  1. časové zpoždění
  2. dolní propust 7 kHz
  3. modifikovaný dekodér Dolby B



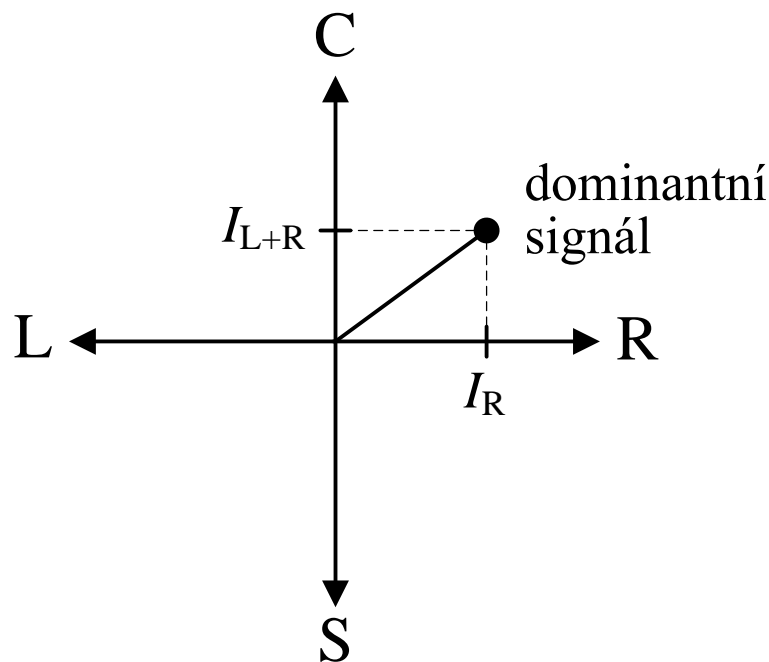
# Dekodér Dolby ProLogic

- Dekódování levého, pravého, surroundového i centrálního kanálu
- Princip konstantního výkonu: dominantní zvuky dočasně omezí schopnost posluchače rozeznat změny směru nedominantních zvuků
- Režimy dekodéru:
  - rychlý – zvýraznění směrovosti více různých dominantních zvuků zároveň
  - pomalý – zachování rozložení zvukového pole



# Adaptivní dekódovací matice dekodéru ProLogic

- Přímá cesta: kanály L a R přiváděné na výstup
- Řídicí cesta:
  1. nalezení vektoru dominantního signálu
  2. vyhodnocení amplitudy L&R a L+R&L-R
  3. logaritmizace a výpočet jejich rozdílu
  4. rozdělení řídicích signálů L/R a C/S podle polarity
- Režimy dekodéru:
  - LRCS
  - 2/1 (kanály L, R, S)
  - 3/0 (3-Stereo, kanály L, R, C)





# Dekodéry Surround a stereo signálu

- Dolby ProLogic II
  - změna rychlosti sledování dominantních složek zvuku spojitě a nezávisle v osách L/R a C/S  $\Rightarrow$  přesnější určení vektoru dominantní složky
  - režim Movie: použití stereofonního surround kanálu bez DP
  - režim Music: pro stereofonní nahrávku, lze měnit virtuální šířku středního kanálu a pozici zvukového obrazu
- Dolby ProLogic IIx
  - rozšíření ProLogic II pro formáty 6.1 a 7.1
- DTS-Neo:6
  - režim Cinema: dekodování zvukových stop MP matice
  - režim Music: rozšíření zvukového obrazu stereofonních nahrávek pro formáty prostorového zvuku do 6.1

# Formát UHJ

- Vyvinuto BBC pro škálovatelný rádiový přenos vícekanálové informace
- Kódování složek B-formátu  $W$ ,  $X$ ,  $Y$ ,  $Z$  do kanálů  $\Sigma$ ,  $\Delta$ ,  $T$  a  $Q$  (tzv. C-formát):

$$\Sigma = 0,9397W + 0,1856X$$

$$\Delta = j(-0,3420W + 0,5099X) + 0,6555Y$$

$$T = j(-0,1432W + 0,6512X) - 0,7071Y$$

$$Q = 0,9772Z$$

- Formáty přenosu:
  - BHJ – pouze kanály  $\Sigma$ ,  $\Delta$  (pro dekódování do stera)
  - THJ – BHJ rozšířené o kanál  $T$  (pro dekódování horizontální surroundové reprodukce)
  - SHJ – BHJ rozšířené o kanál  $T$  omezený do 5 kHz (pro FM přenos surroundu)
  - PHJ – všechny kanály (pro dekódování perifonie / surround with height), ekvivalentní B-formátu

# Formát UHJ

- Dekódování BHJ do sterea: jako MS formát

$$L = (\Sigma + \Delta)/2$$

$$R = (\Sigma - \Delta)/2$$

- Dekódování UHJ do B-formátu:

$$\Sigma = (L + R)/2$$

$$\Delta = (L - R)/2$$

$$W = 0,982 \Sigma + 0,197 j(0,828 \Delta + 0,768 T)$$

$$X = 0,419 \Sigma - j(0,828 \Delta + 0,768 T)$$

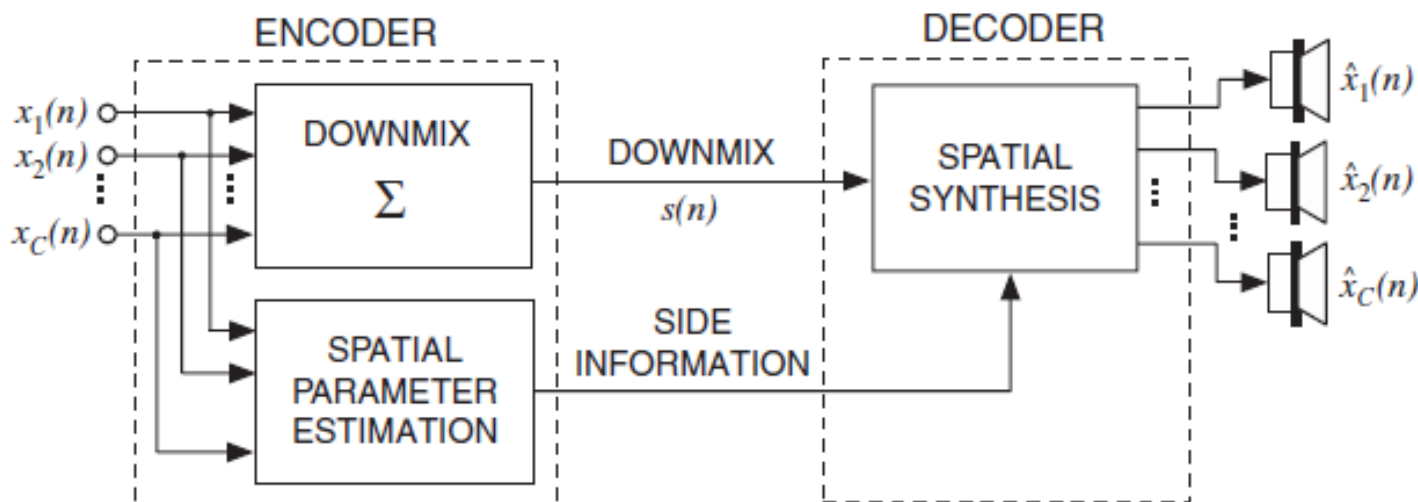
$$Y = 0,769 \Delta - 0,676 T + 0,187 jS$$

$$Z = 1,023 Q$$

Fázový posun o  $90^\circ$  ( $\pm j$ ) se realizuje fázovacími články

# Binaural Cue Coding

- Kódování: downmix vstupních kanálů do jednoho nebo více kanálů, výpočet prostorové informace (ICLD, ICTD, ICC, apod.)
- Dekódování: upmix s využitím přenesených prostorových informací



# Kódovací strom TTO (Two-To-One)

- Provádí downmix 2 signálů  $x_1, x_2$  na 1 signál  $s$

$$s_m(k) = \frac{x_{1,m}(k) + x_{2,m}(k)}{\psi_{1,b} + \psi_{2,b}} \quad (\psi_{1,b} + \psi_{2,b})^2 = \frac{p_{x_1,b} + p_{x_2,b} + 2\rho_{x_1,x_2,b} \sqrt{p_{x_1,b} p_{x_2,b}}}{p_{x_1,b} + p_{x_2,b}}$$

a prostorové parametry  $P_n$

- rozdíl úrovní signálů mezi kanály

$$\Delta L_{x_1,x_2,b} = 10 \log \frac{p_{x_1,b}}{p_{x_2,b}} \quad p_{x_i,b} = \sum_k \sum_{m=m_b}^{m_{b+1}-1} x_{i,m}(k) x_{i,m}^*(k)$$

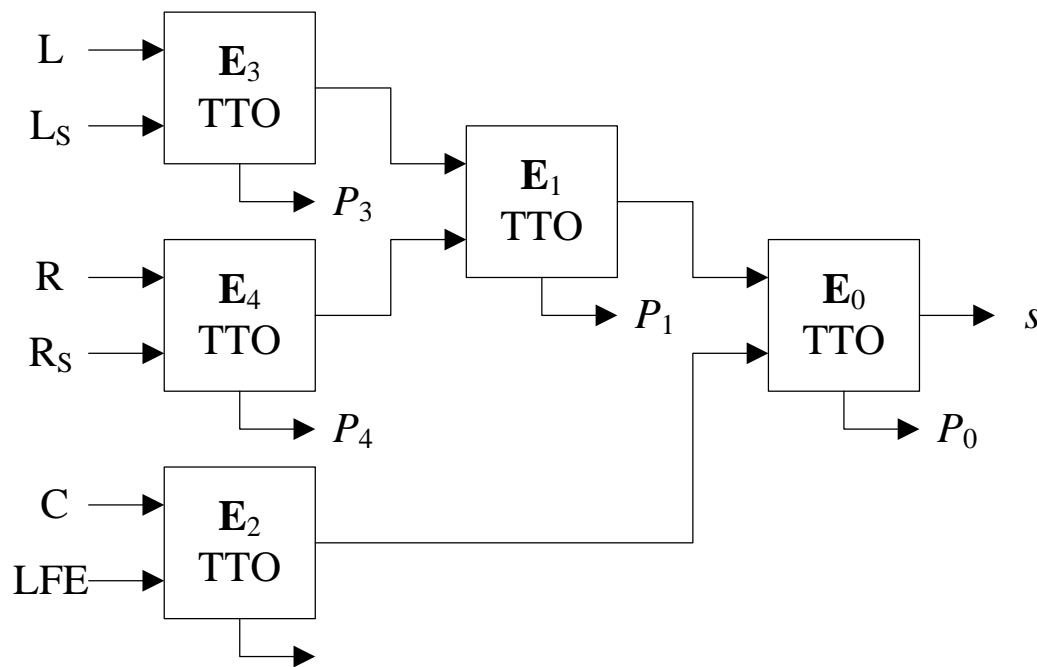
$p_{x,b}$  – výkon signálu  $x$  v pásmu  $b$ ,  $m_b$  – první spektrální koeficient pásma  $b$ ,  $k$  – pořadí vzorku

- korelace mezi kanály

$$\rho_{x_1,x_2,b} = \operatorname{Re} \left\{ \frac{\sum_k \sum_{m=m_b}^{m_{b+1}-1} x_{1,m}(k) x_{2,m}^*(k)}{\sqrt{p_{x_1,b} p_{x_2,b}}} \right\}$$

# Kódovací strom TTO (Two-To-One)

- příklad: mono downmix formátu 5.1 surround



# Kódovací strom TTT (Three-To-Two)

- provádí downmix 3 signálů  $x_L, x_R, x_C$  na 2 signály  $s_L, s_R$  a pomocný signál  $s_C$

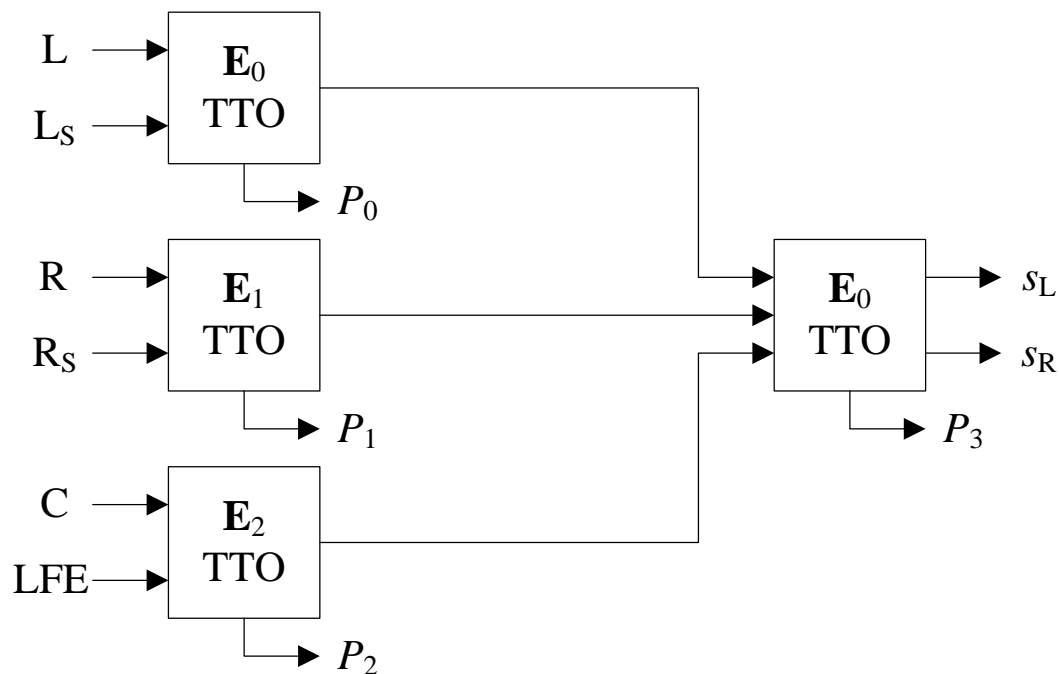
$$\begin{bmatrix} s_{L,m}(k) \\ s_{R,m}(k) \\ s_{C,m}(k) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_{L,m}(k) \\ x_{R,m}(k) \\ x_{C,m}(k)/\sqrt{2} \end{bmatrix}$$

- a prostorové parametry  $P_n$

$$\Delta L_{1,b} = 10 \log \frac{p_{x_L,b} + p_{x_R,b}}{p_{x_C,b}/2} \qquad \Delta L_{2,b} = 10 \log \frac{p_{x_L,b}}{p_{x_R,b}}$$

# Kódovací strom TTT (Three-To-Two)

- příklad: stereo downmix formátu 5.1 surround





# Přenos reziduálních signálů

- pro původní signály před downmixem platí

$$x_{1,m}(k) = \psi_{1,b} s_m(k) + d_{1,m}(k)$$

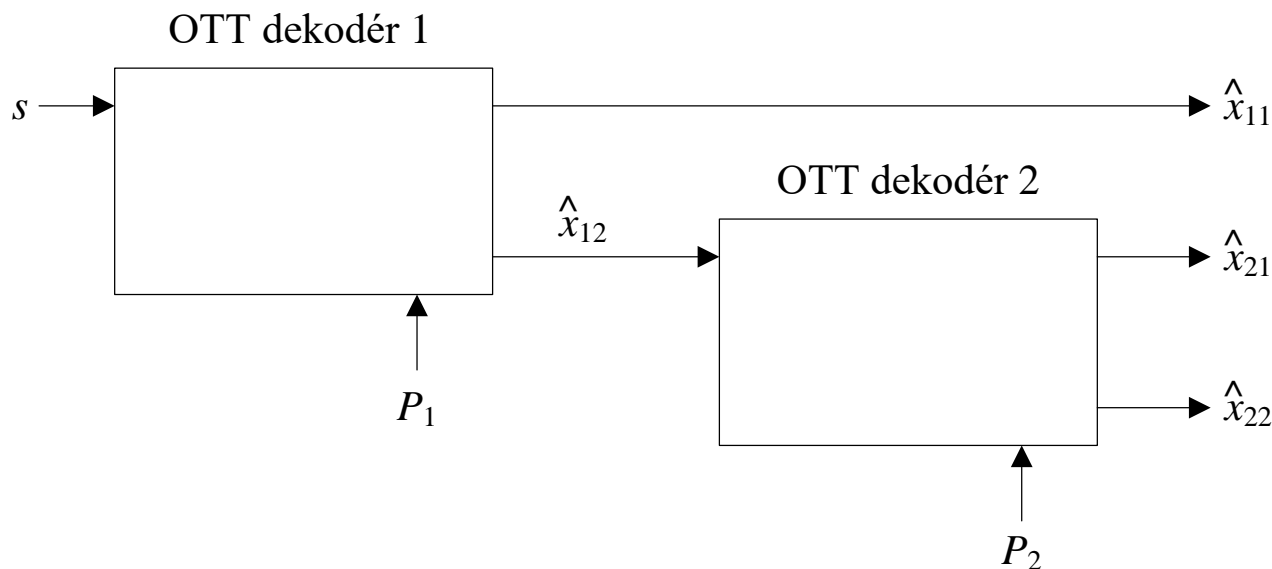
$$x_{2,m}(k) = \psi_{2,b} s_m(k) + d_{2,m}(k)$$

$d_{i,m}$  – reziduální signály

- reziduální signály mohou být zakódovány a přeneseny společně s downmixem a prostorovými parametry

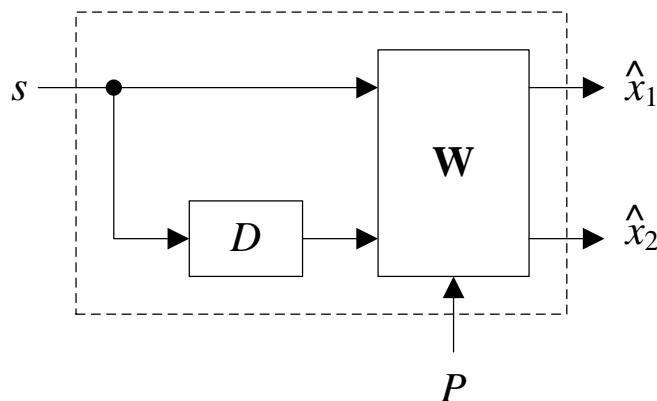
# Dekodér prostorové informace

- Stavební bloky:
  - OTT (One-To-Two) – upmix jednoho kanálu do dvou pomocí prostorových parametrů
  - TTT (Two-To-Three) – upmix dvou kanálů do tří pomocí prostorových parametrů
  - Příklad: tříkanálový výstup pomocí dvou bloků OTT



# Dekódovací strom OTT (One-To-Two)

- upmix signálu  $s$  na odhady dvou původních signálů  $x_1, x_2$  pomocí jednoho prostorového parametru  $P_n$ , dekorelátoru  $D_n$  a matice  $\mathbf{W}$ 
  1. korelace výstupních signálů musí odpovídat přenesené korelaci mezi kanály
  2. poměr výkonu výstupních signálů musí odpovídat přenesenému rozdílu úrovní
  3. součet energií výstupních signálů musí být stejný jako energie vstupního signálu



# Dekódovací strom OTT (One-To-Two)

- dekódovací matice:

$$\mathbf{W}_b = \begin{bmatrix} \lambda_{1,b} \cos(\alpha_b + \beta_b) & +1 \\ \lambda_{2,b} \cos(-\alpha_b + \beta_b) & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \lambda_{1,b} \sin(\alpha_b + \beta_b) \end{bmatrix}$$

$\lambda_{1,b}$ ,  $\lambda_{2,b}$  – relativní amplitudy výstupního signálu vůči vstupnímu

$$\lambda_{1,b} = \sqrt{\frac{10^{(\Delta L_{x_1 x_2 b} / 10)}}{1 + 10^{(\Delta L_{x_1 x_2 b} / 10)}}} \quad \lambda_{2,b} = \sqrt{\frac{1}{1 + 10^{(\Delta L_{x_1 x_2 b} / 10)}}}$$

$\alpha_b$  – proměnná závislá na korelaci mezi kanály

$$\alpha_b = \frac{1}{2} \arccos(\rho_{x_1 x_2, b})$$

$\beta_b$  – poměr downmixu a výstupního signálu dekorelátoru

$$\beta_b = \tan\left(\frac{\lambda_{2,b} - \lambda_{1,b}}{\sqrt{2}} \alpha_b\right)$$

# Dekódovací strom TTT (Two-To-Three)

- dekódovací matice:

$$\begin{bmatrix} \hat{x}_{L,m}(k) \\ \hat{x}_{R,m}(k) \\ \hat{x}_{C,m}(k) \end{bmatrix} = \begin{bmatrix} w_{11,b} & 0 \\ 0 & w_{22,b} \\ w_{31,b} & w_{32,b} \end{bmatrix} \cdot \begin{bmatrix} s_{L,m}(k) \\ s_{R,m}(k) \end{bmatrix}$$

$$w_{11,b} = \sqrt{\frac{\kappa_{1,b} \kappa_{2,b}}{\kappa_{1,b} \kappa_{2,b} + \kappa_{2,b} + 1}}$$

$$w_{22,b} = \sqrt{\frac{\kappa_{1,b}}{\kappa_{1,b} + \kappa_{2,b} + 1}}$$

$$w_{31,b} = \frac{1}{2} \sqrt{2 \frac{\kappa_{2,b} + 1}{\kappa_{1,b} \kappa_{2,b} + \kappa_{2,b} + 1}}$$

$$w_{32,b} = \frac{1}{2} \sqrt{2 \frac{\kappa_{2,b} + 1}{\kappa_{1,b} + \kappa_{2,b} + 1}}$$

$$\kappa_{i,b} = 10^{\Delta L_{i,b}/10}$$

# Prostorová reprodukce zvuku