

Univerza v Ljubljani  
Fakulteta za *matematiko in fiziko*

Govekar  
Audio



# Fizik v svetu avdia

D. Svenšek

Fakulteta za matematiko in fiziko, Univerza v Ljubljani, Slovenija

*SSS na FMF  
Ljubljana, 14. november 2008*

# Okoliščine, izzivi in rešitve, ki nastopajo pri snemanju zvoka

## Napovednik

- kaj želimo in kako to dosežemo?
- tri ključni elementi kvalitete zvoka
- zakonitosti prostora
- kondenzatorski mikrofoni
- stereo postavitve
- binauralna tehnika

## Kaj želimo?

- zelo preprosto (audiofilsko načelo): čim bolj avtentičen, naraven zvok
- obenem mora posnetek dobro funkcionirati (npr. velikost prostora, oddaljenost)
- merodajni le “naravni” zvoki (sintetična glasba tukaj nezanimiva)

## Kako to dosežemo?

- realen, dovolj velik, akustično lep prostor (ne studio)
- čim boljši mikrofoni, čim manj mikrofonov (idealno dva)
- postavitev mikrofonov je odločilna
- čim boljša oprema (preamp, AD, resolucija, najpomembneje: dober analogni del!)
- in seveda: nobenih umetnih odmevov, komprimiranja (danes vse), ekvilizacije (dovoljena malenkost) in ostalih efektov
- čim boljša oprema pri predvajanju

## anekdota ...

- **D. Svenšek pred leti** o kvaliteti zvoka s CDja in vinilk (v knjigi *Kvarkadabre Zakaj je nebo modro?*):

/.../ Do rahlega popačenja pri predvajanju bi lahko prišlo le pri zelo visokih frekvencah, zanje pa smo že tako ali tako precej slabše občutljivi. Poleg tega diskretni signal pri predvajanju spet peljemo skozi filter, ki zgladi neenakomernosti, povezane s končno zalogo diskretnih vrednosti, predvsem pa izboljša podajanje visokih frekvenc. Sinusni signal s frekvenco 20 kHz, ki smo ga uspeli vzorčiti le “na hribu in v dolini”, je tako pri predvajanju zaradi filtra spet skoraj lep sinus. Kvaliteta zvoka z zgoščenk torej ni nič slabša od tistega z vinilk, prej obratno. /.../

- **D. Svenšek danes**
- stvar je jasna, a žal irelevantna (mp3)

## Sluh – najbogatejši čut

$$L_{\text{dB}} = 10 \log_{10} \left( \frac{P_1}{P_0} \right)$$

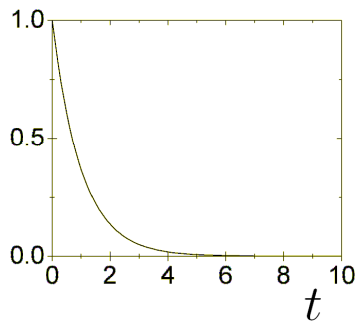
- razpon glasnosti 120 dB (to je 12 velikostnih redov ...)
- prag slišnosti 20  $\mu\text{Pa}$  (RMS) (amplituda valovanja je 0.01 nm, to je manj od Bohrovega radija!)
- A to je šele začetek. Bistveno je, da ima sluh velikansko prednost pred vidom (uklon), kar je evolucija dodobra izkoristila (detekcija faze!).

## Tri elementi dobrega zvoka

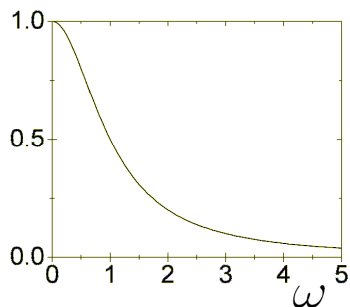
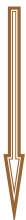
- barvna korektnost (frekvenčna “linearnost”)
- lokalizacija in realen, pravilen občutek prostora (fazna zvestost)
- tekstura (časovni potek)

# Frekvenčna “linearnost”

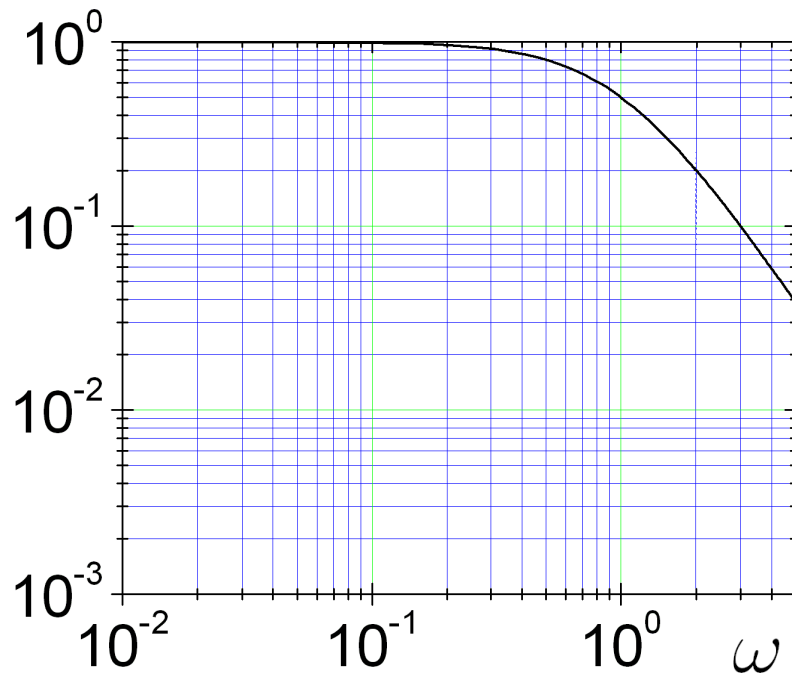
- frekvenčna prenosna funkcija  $F(\omega)$  sistema, ki ga imamo za **linearne** (☺ brez narekovajev)
- idealno:  $F(\omega)$  je konstanta, a to ne gre
- $F(\omega)$  je Fourier transform odzivne funkcije na impulz (impulse response)



$$\alpha(t) = \alpha_0 e^{-t/\tau}$$



$$F(\omega) = \alpha_0 \frac{2/\tau}{1/\tau^2 + \omega^2}$$



# Tekstura

- fina časovna ločljivost
- za linearne sisteme že določena s frekvenčnim odzivom
- običajna mera za nelinearnost: total harmonic distortion (THD) (steady-state količina)
- **dejstvo**: ojačevalca z enakim THD (lahko pod 0.01%) zvenita različno
- **tranzienti!** – tega s steady-state količino ne zajamemo (ker nas zanima nelinearnost sistema!), je pa bistveno bližje realnemu zvoku kot monofrekvenčni signal

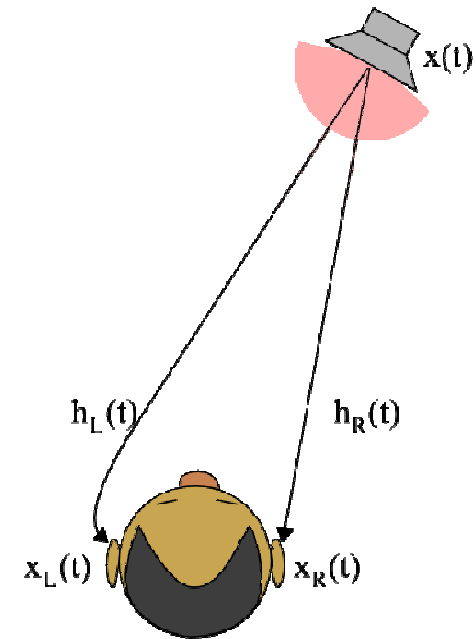
# Lokalizacija in prostor

## Kako določamo lego zvočila?

- na voljo imamo dvoje ušes in glavo ...
- razlike v jakosti
- časove/fazne razlike
- razlike v barvi zvoka
- vse skupaj opiše HRTF (head-related transfer function)
- lep primer: zvočilo v meridialni ravnini

## Kaj pa na posnetku?

- prostor izredno selektivna zadeva (evolucijska sposobnost)
- velik problem digitalije: prostor je tam, kjer je ločljivost slaba





# Prostor kot tak

- za dober posnetek je seveda odločilen že fizični prostor, ne šele tisti na posnetku
- najpomembneje: razmerje direktni zvok/prostor (tudi za poslušalca v živo)

## Odmev

- Sabinov zakon (difuzno polje)
- ključno: dovolj velik prostor → dovolj dolg, a ne premočan odmev

$$V \frac{dw}{dt} + \frac{cS_{eff}}{4} w = P$$

$$w = w_0(1 - e^{-t/\tau}) \quad w_0 = \frac{4P}{cS_{eff}}$$

$$w = w_0 e^{-t/\tau} \quad \tau = \frac{4V}{cS_{eff}}$$

čas poodmevanja:

$$t_{60} = 0,161 \frac{V}{S_{eff}}$$

## Eigenmodes

- vsak prostor ima vozle (odvisni od frekvence)
- paziti, da zvočila (solista) ne postavimo nerodno
- sprememba jakosti in barve zvoka (prezentnost)

$$L\Phi - \frac{\partial^2 \Phi}{\partial t^2} = 0$$

$$L\Phi_j + \omega_j^2 \Phi_n = 0$$

$$LH_\omega + \omega^2 H_\omega = -\delta(\mathbf{r} - \mathbf{r}_0)$$

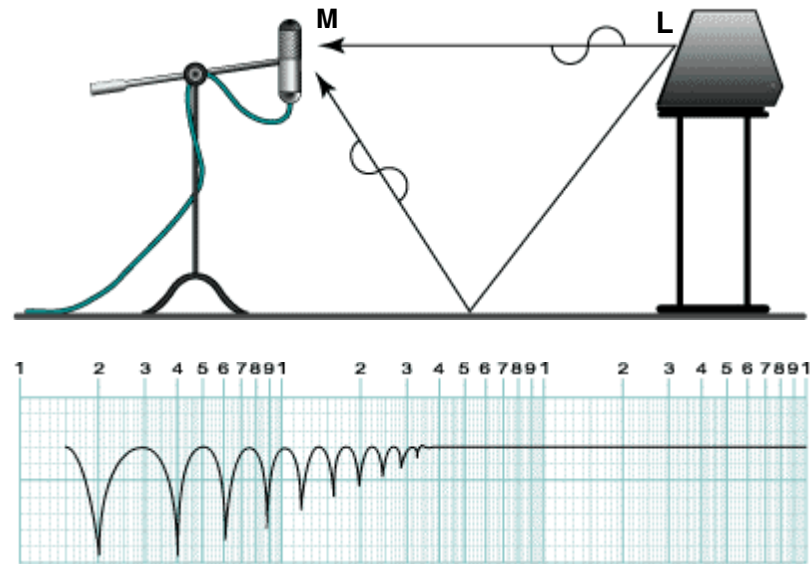
$$H_\omega(\mathbf{r}, \mathbf{r}_0) = \sum_j^n a_n \Phi_j(\mathbf{r})$$

prenosna funkcija med parom točk:

$$H_\omega(\mathbf{r}, \mathbf{r}_0) = \sum_j^n \frac{\Phi_j(\mathbf{r}_0)\Phi_j(\mathbf{r})}{\omega_j^2 - \omega^2}$$

## Comb filtering

- interferenca
- naraven pojav, fizično prisoten
- problem pri miksanju več mikrofonov (nenaravno)



## Geometrija prostora

- čimbolj kaotična (nekvadrasta), da se valovanje hitro razleze (čas. slika) oziroma lastne rešitve čimbolj enakomerno ekscitirajo (frekv. slika)
- primer slabega: majhna betonska soba
- čim manj “ostrih senc” (ostri, ravni robovi, periodični vzorci)
- kroglast, ne paličast

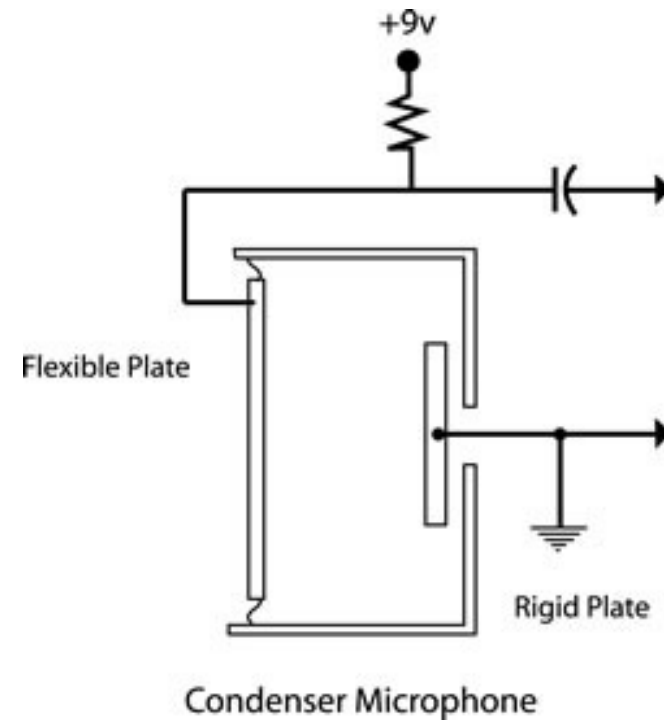
# Mikrofoni

- kondenzatorski superiorni (ne bomo o dinamičnih, ribbon, piezo ...)

$$e = CU$$

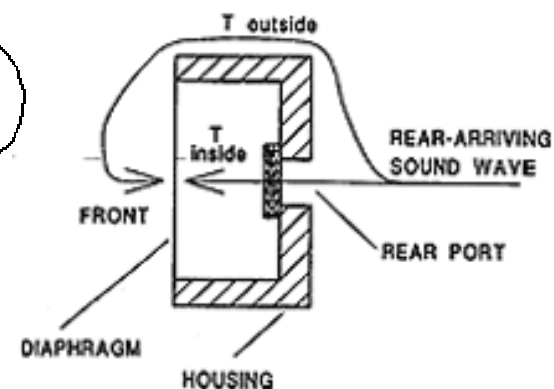
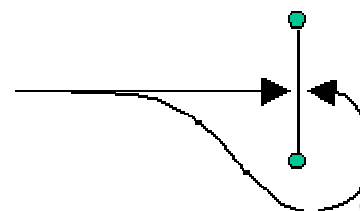
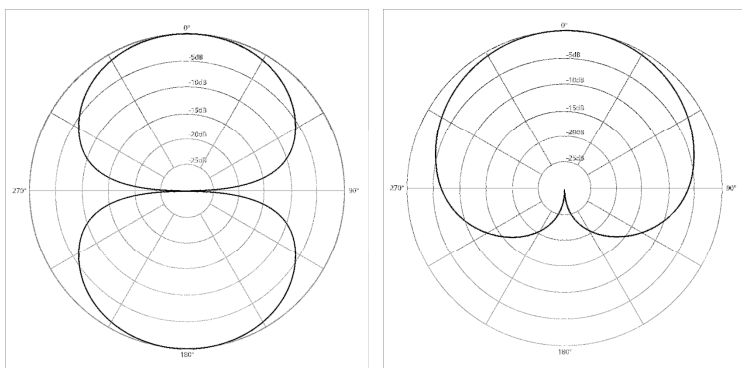
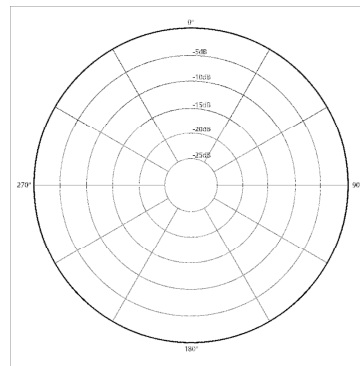
$$\frac{de}{dt} = I = \frac{dC}{dt}U$$

- spomnimo se: 20  $\mu\text{Pa}$   $\rightarrow$  amplituda  
0.01 nm



## Smerna karakteristika

- izotropna, **omnidirekionalni**:  
tlačna kapsula, najbolj flat
- anizotropna: tlačni gradient →  
neobčutljivost pri nizkih
- **bidirekcionalni** (osmica)
- **kardioid** (ledvička)



## Splošno pri mehanskih pretvornikih

- problem resonance
- želimo konstanten frekvenčni odziv, resonanca nezaželjena
- pretvornik deluje dovolj **pod prvo resonanco**

- pomembno:

linearen fazni zamik

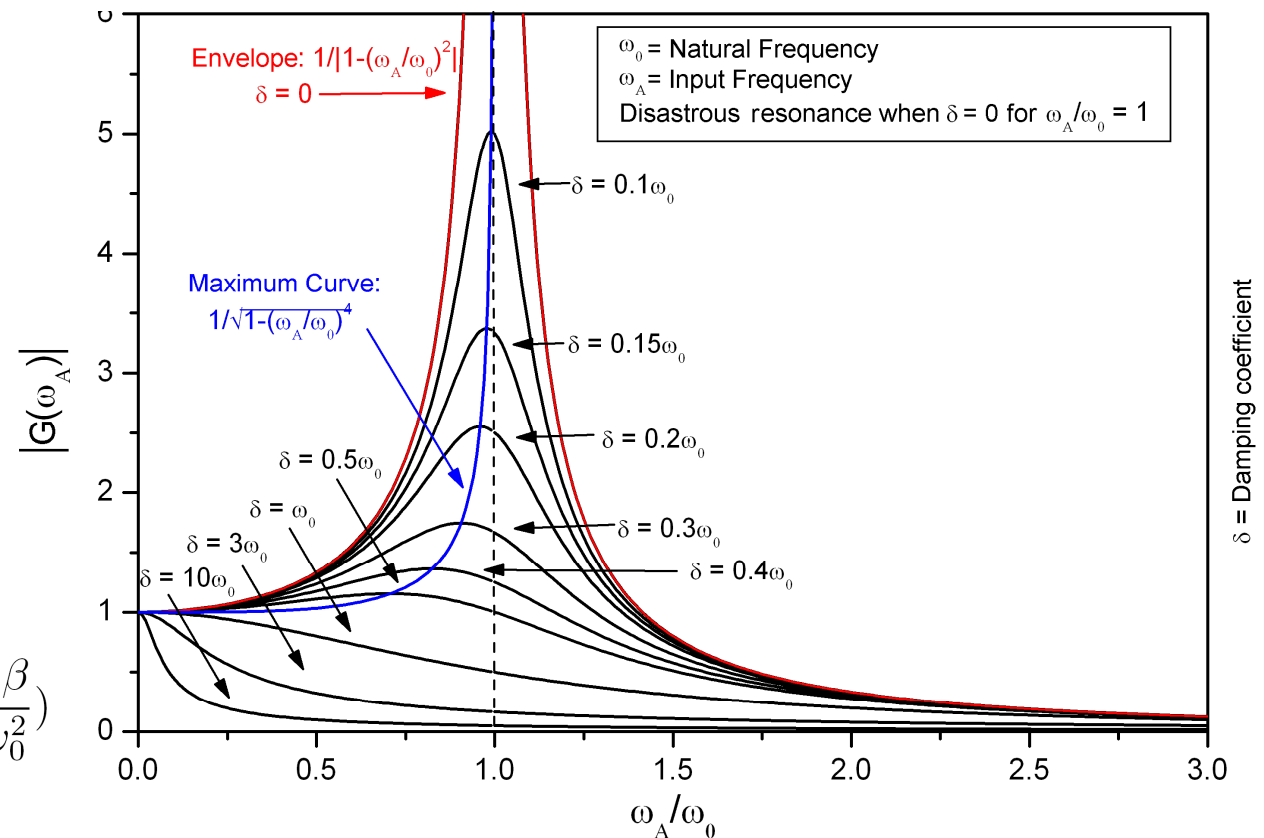
povzročča le **časovni**

**premik** signala

$$\operatorname{tg} \Delta = \frac{2\beta\omega}{\omega_0^2 - \omega^2}$$

$$\Delta \approx \frac{2\beta}{\omega_0^2} \omega$$

$$\phi(t) = \omega t + \Delta = \omega \left( t + \frac{2\beta}{\omega_0^2} \right)$$



## Omnidirekcionalni

- najpreciznejši, najbolj flat frekvenčna karakteristika (ni problema pri basih)
- merilni mikrofoni so te vrste
- pri višjih frekvencah postajajo direkcionalni, odvisno od velikosti kapsule
- prva (in za vrhunski posnetek edina) izbira, a zahtevajo dober prostor

## Težave direkcionalnih mikrofonov

- niso flat – nizke frekvence pešajo (gradient tlaka), dušenje 6 dB/oktavo
- direkcionalnost je frekvenčno odvisna → koloracija
- bližinski pojav (proximity effect): bas nenaravno narašča
- kdaj pridejo prav: slab prostor (premajhen ali preveč odmevajoč), moteče ozadje (publika, hrup), koincidenčne postavitve

## Še zanimiv efekt: vpliv odbojne površine

- pogoj na togi površini:  $v_{\perp} = 0 \implies (\nabla p)_{\perp} = 0$
- hrbet tlaka, neglede na vpadno smer (čas. slika)

oz. lastni način (frekv. slika)

- pridobimo faktor 2 v amplitudi = 6 dB
- ni koloracije zaradi vozlov prostora
- “pressure zone mic”, “Grenzflächenmikrofon”
- tudi sicer: z oddaljenostjo mikrofona od tal kontroliramo jakost nizkih frekvenc

$$p(\mathbf{r}, t) = (e^{i\mathbf{k}\cdot\mathbf{r}} + Ae^{-i\mathbf{k}'\cdot\mathbf{r}}) e^{-i\omega t} \implies A = 1$$

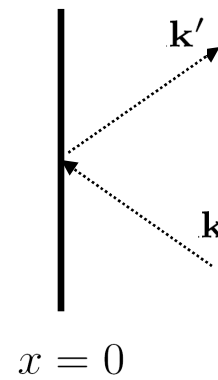
$$p(\mathbf{r}, t) = \begin{Bmatrix} \cos k_x x \\ \sin k_x x \end{Bmatrix} \begin{Bmatrix} \cos k_y y \\ \sin k_y y \end{Bmatrix} \begin{Bmatrix} \cos k_z z \\ \sin k_z z \end{Bmatrix} e^{-i\omega t}$$

$$\frac{\partial^2 p}{\partial t^2} - c^2 \nabla^2 p = 0 \quad c^2 = (\rho\chi)^{-1}$$

valovna enačba

$$\rho \frac{\partial \mathbf{v}}{\partial t} = -\nabla p \quad \text{gibalna enačba}$$

$$\chi S \frac{\partial p}{\partial t} = -\nabla \cdot \mathbf{v} \quad \text{kontinuitetna enačba}$$



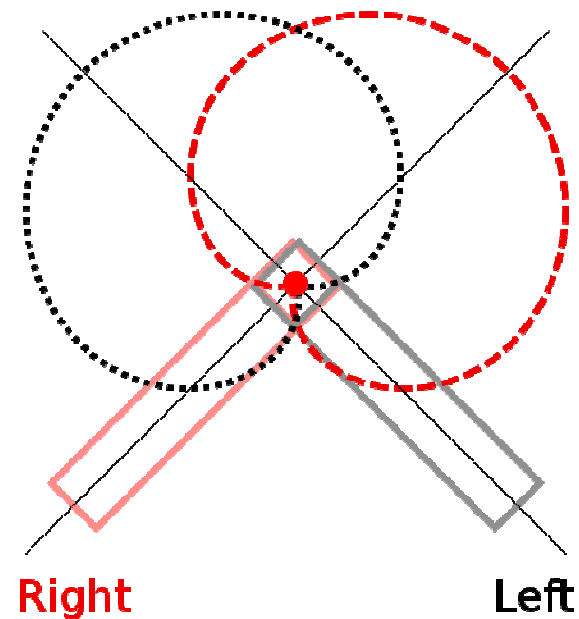


# Postavitve

- osnovno pravilo: čim manj mikrofonov! (comb filtering pri mešanju, nenaraven prostor)
- dva ali tri mikrofoni
- spot mics za večje zasedbe (prezentnost, definicija), a zelo skromno zmiksano k stereo paru + časovni zamik (delay)

## Koincidenčni XY stereo

- par navzkriž usmerjenih direkcionalnih mikrofonov v “isti” točki
- učinek sterea zaradi razlike v intenziteti med kanaloma
- natančna, sterilna lokalizacija, ni blizu naravnemu
- priročnost, monokompatibilnost
- primerno za premajhne, preodmevajoče, hrupne prostore



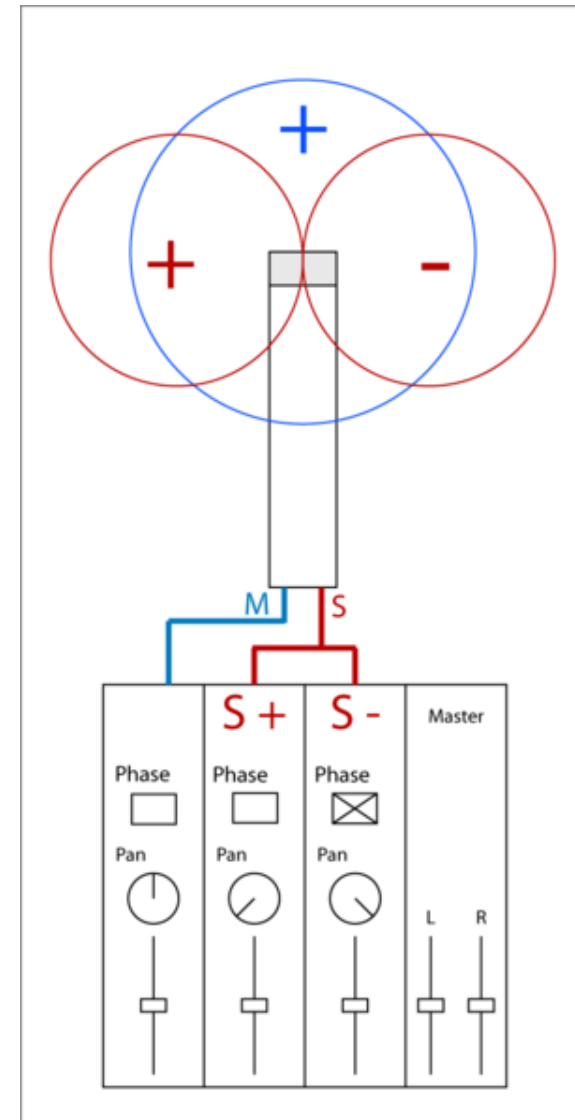
## Mid-Side (MS) stereo

- učinek sterea zaradi razlik v intenziteti
- naravnost usmerjeni kardiod (mid), počez usmerjena osmica (side), v "isti točki"
- levi in desni kanal s seštevanjem oziroma odštevanjem M in S signalov

$$\begin{aligned} L &= (M + S)/2 \\ R &= (M - S)/2 \end{aligned}$$

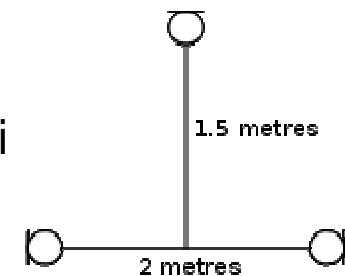
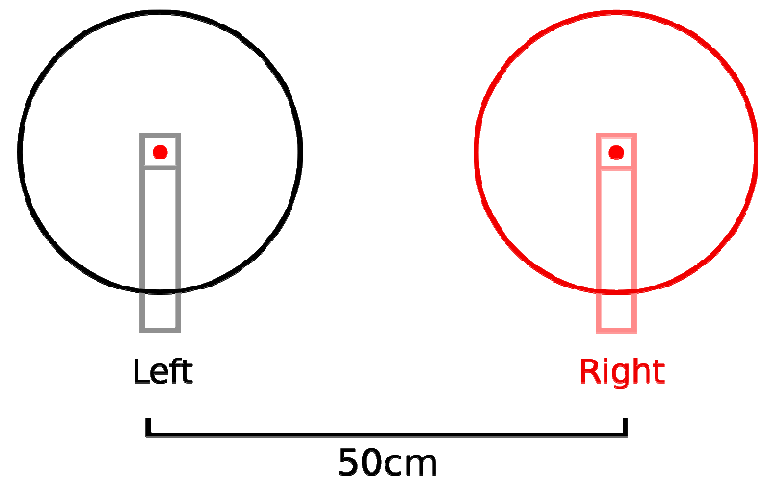
- kompaktnost, priročnost, monokompatibilnost
- običajne slabosti koincidenčnih tehnik

- 
- karaoke:  
$$\begin{aligned} M &= (L + R)/2 \\ S &= (L - R)/2 \end{aligned}$$



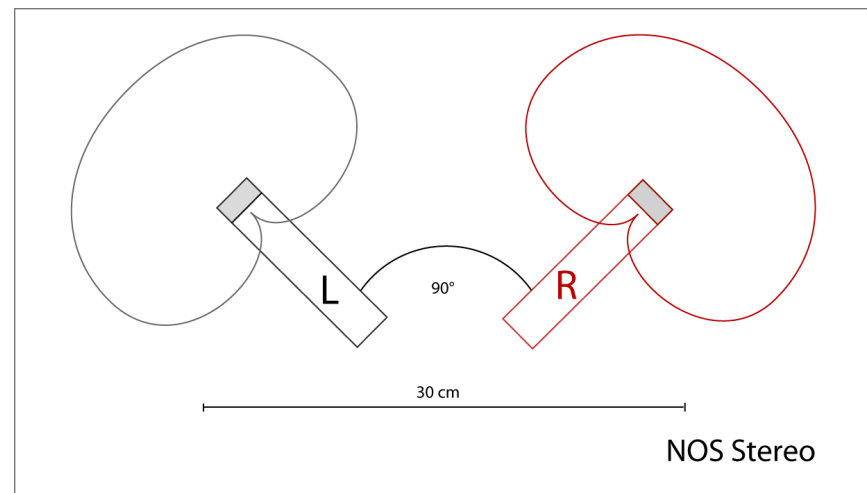
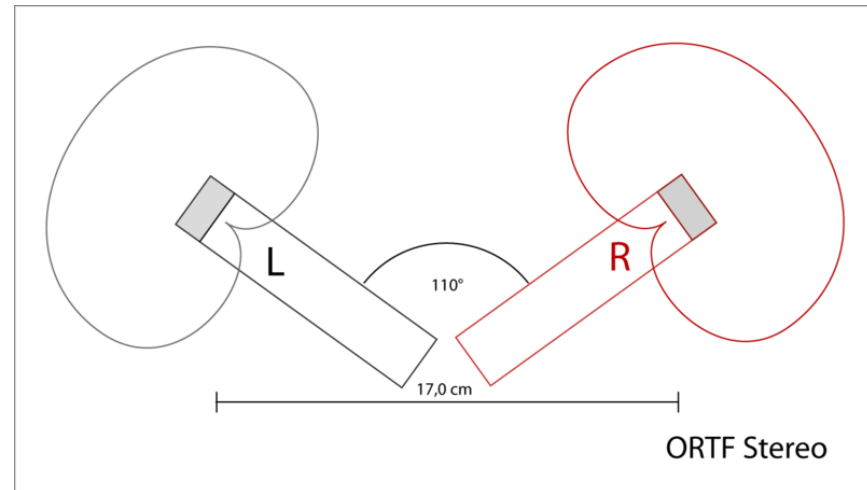
## Spaced pair (AB) stereo

- razmaknjena omnidirekionalna mikrofona
- stereo zaradi faznih (časovnih) razlik
- naraven, zračen, mehek stereo
- vse prednosti omnidirekcionalnih mikrofonov, pogoj je dober prostor
- ni monokompatibilnosti, kar danes ni problem
- lahko ima "luknjo" na sredini – možno dodati še mic na sredini (npr. Decca tree)
- za vrhunski rezultat fazno uparjena mikrofona (matched pair)
- fazna zvestost zahteva boljšo opremo, reprodukcija na slušalkah je boljša kot na zvočnikih, ker ponavadi niso fazno zvesti



## Kombinacija XY in AB: ORTF in NOS

- spaced pair direkcionalnih mikrofonov
- intenzitetni in fazni stereo
- boljša monokompatibilnost
- kompaktnost (na enem stojalu)



## Jecklin disk

- predstavnik spaced pair tehnike z vmesno oviro, absorberjem
- omnidirekionalna mikrofona, vmes disk, prevlečen z umetno peno
- tako nekoliko ločimo kanala kljub uporabi omnijev
- vmesna ovira funkcioniira podobno kot glava

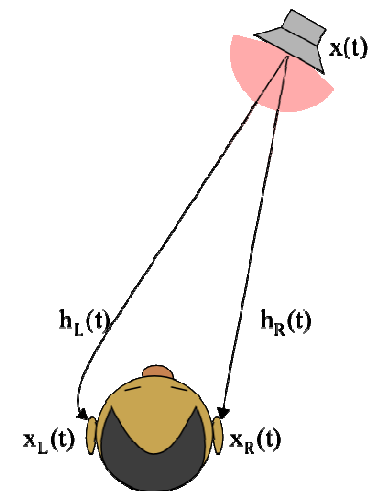


## Umetna glava (binaural)

- ko sem prvič videl, mi je šlo na smeh ...
- “vmesna ovira” je glava, omnija pa sta nameščena v ušesnih školjkah
- reprodukcija s slušalkami, skrajno avtentično
- uparjena mikrofona, pravi proporci in material glave, najpomembnejša je oblika in velikost uhljev
- reprodukcija na zvočnikih: cross talk, dvakrat čez HRTF (pri snemanju in poslušanju), možna je dekonvolucija
- glava močno spremeni barvo – efekt površine +6dB za višje frekvence. Material (tudi frizura ☺) je izredno pomemben (odbojnost, absorpcija)!
- najboljša je neumetna glava



Neumann KU100



## Efekt površine lahko ocenimo!

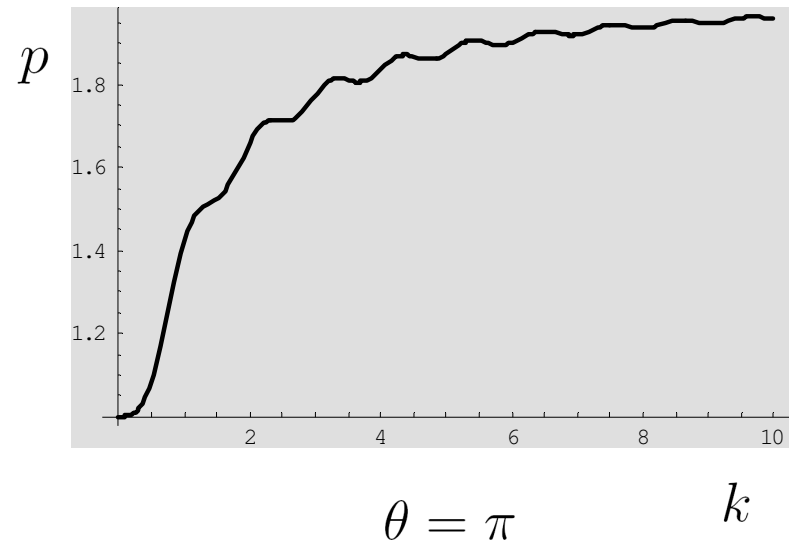
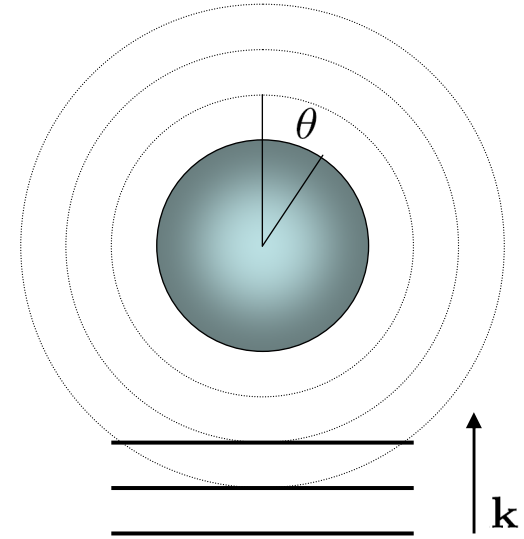
- prvi približek: sipanje ravnega vala na sferi

$$p(\mathbf{r}) = e^{ikr \cos \theta} + \sum_{l=0}^{\infty} A_{k,l} h_l^{(1)}(kr) P_l(\cos \theta)$$

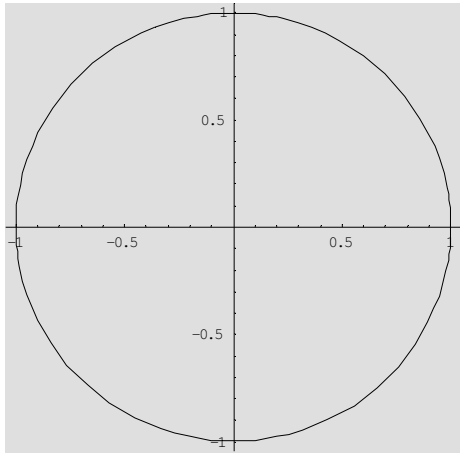
$$p(\mathbf{r}) = \sum_{l=0}^{\infty} [(2l + 1)i^l j_l(kr) + A_{k,l} h_l^{(1)}(kr)] P_l(\cos \theta)$$

$$\left. \frac{\partial p}{\partial r} \right|_{r=R} = 0$$

$$A_{k,l} = -(2l + 1)i^l \frac{j_l'(kR)}{h_l'(kR)}$$

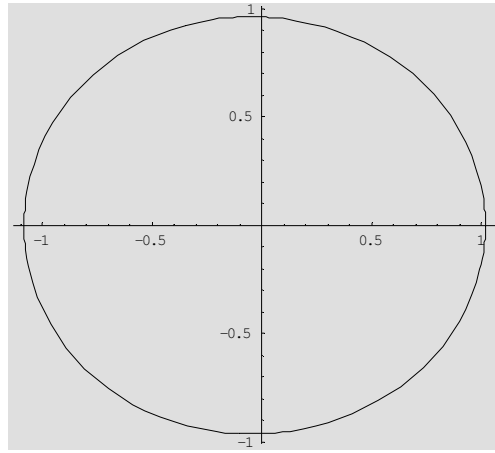






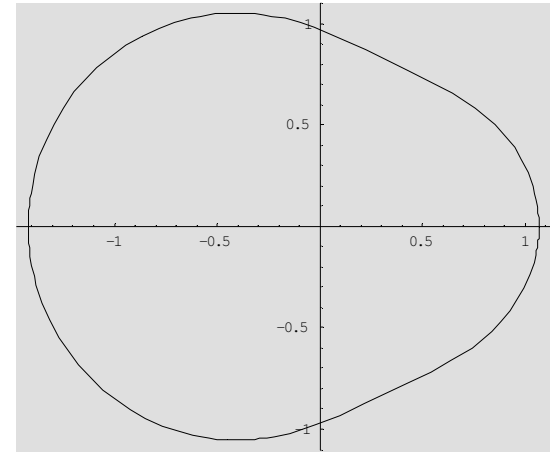
$$\lambda = 30 \cdot 2R$$

$$\nu \approx 55 \text{ Hz}$$



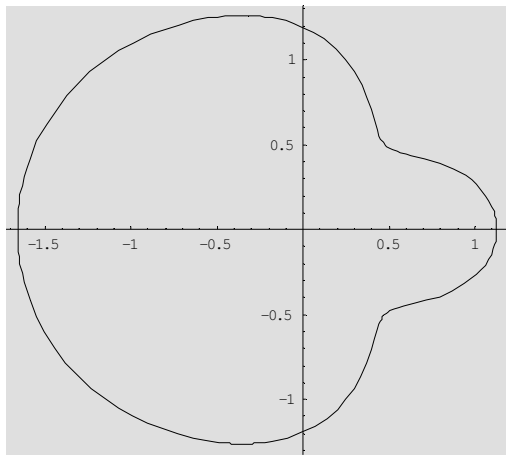
$$\lambda = 6 \cdot 2R$$

$$\nu \approx 280 \text{ Hz}$$



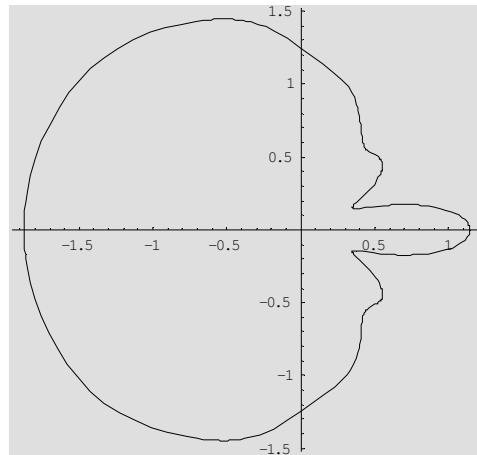
$$\lambda = 3 \cdot 2R$$

$$\nu \approx 560 \text{ Hz}$$



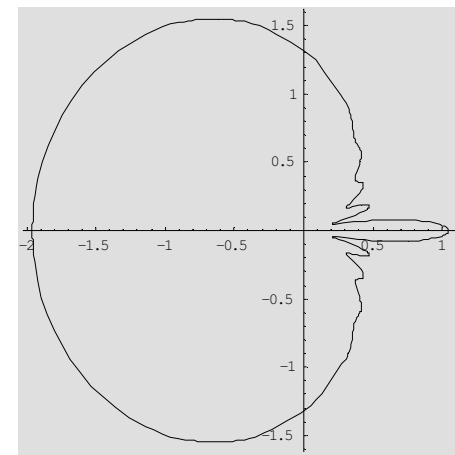
$$\lambda = 1.5 \cdot 2R$$

$$\nu \approx 1100 \text{ Hz}$$



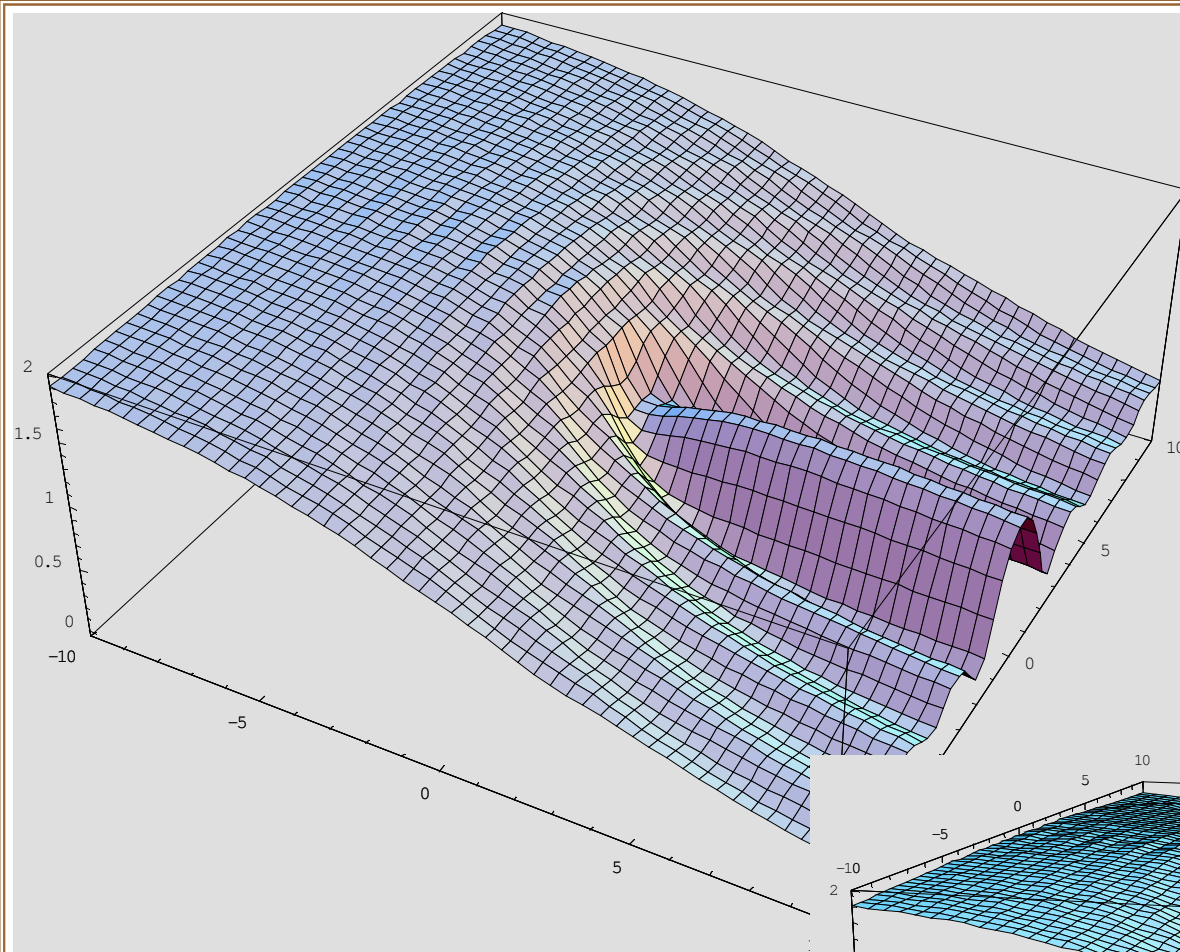
$$\lambda = 0.6 \cdot 2R$$

$$\nu \approx 2800 \text{ Hz}$$



$$\lambda = 0.3 \cdot 2R$$

$$\nu \approx 5700 \text{ Hz}$$



amplituda tlaka na  
površini kogle  
( $k, \theta$ ) prikaz

očitno, da je material  
zelo pomemben

