

STEREOSNEMANJE: PRINCIPI DVOUŠESNE ZAZNAVE ZVOKA

DANIEL SVENŠEK

Fakulteta za matematiko in fiziko
Univerza v Ljubljani

PACS: 43.66.Pn, 43.20.+g, 43.38.Md

Spregovorili bomo o izzivih, s katerimi se soočamo pri snemanju akustičnih scen s stereomikrofonskim parom. Poslušalčevo akustično izkušnjo želimo čim verneje zajeti in jo nato reproducirati z namenom, da bi v njem ponovno vzbudila občutke žive izvedbe. Omnidirekcionalni mikrofoni prekašajo direkcionalne, razmaknjene postavitve so ustreznejše od koincidenčnih. Z akustično oviro med razmaknjenima mikrofonomi lahko izboljšamo intenzitetno separacijo kanalov. V najbolj naravnem primeru je ovira lahko umetna (celo človekova) glava. Vpliv glave bomo kvalitativno pojasnili in vpeljali njena impulzni odziv (HRIR) in prenosno funkcijo (HRTF). Z vpadom zvočnega valovanja na visokoimpedančno mejno površino bomo kvalitativno razložili vpliv glave na nizke in visoke frekvence. Predstavili bomo eksaktno rešitev sipanja ravnega vala na togi krogli in iz nje izluščili prenosno funkcijo, ki rabi kot model prenosne funkcije človeške glave. Na primeru tega modela si bomo ogledali, kako kvalitativne lastnosti prenosne funkcije glave interpretiramo v luči temeljnih fizikalnih načel.

STEREO RECORDING: PRINCIPLES OF BINAURAL PERCEPTION OF SOUND

Challenges of stereo pair miking of acoustic venues are reviewed. The goal is to capture the listener's live acoustic experience as veristic as possible and then reproduce it as to arouse in him the feelings he had during the live performance. Coincident setups are disfavoured and omnidirectional spaced pair techniques are preferred to directional ones. These can be baffled to improve the intensity separation of the two channels. A natural extrapolation of the baffling is the use of a dummy (or even human) head – the binaural recording technique. The effect of the head is qualitatively discussed, the head-related impulse response (HRIR) and transfer function (HRTF) are introduced. The effect of the incidence on a high impedance boundary is discussed and related to the influence of the dummy head surface on low and high frequencies. The exact solution of the scattering of a plane wave on a hard sphere is presented and from it the hard sphere transfer function is extracted, which serves as a model for the HRTF of human head. With the help of this model, qualitative features of the HRTF are brought in connection with underlying physical principles.

Uvod

Snemanje glasbe, ki se izvaja in doživlja v akustičnem ambientu, se bistveno razlikuje od studijskega snemanja. Pri slednjem iluzijo akustičnega prostora sestavimo umetno, ob tem pa, odvisno od zvrsti glasbe, uporabljamo številne

efekte, s katerimi lahko posnetek obenem naredimo polnejši, mogočnejši, intenzivnejši, jasnejši, mehkejši . . . Nasprotno pa pri snemanju v akustičnem prostoru izhajamo iz minimalističnega načela, da je najlepša zvočna slika naravna, natanko tista, ki jo v tem ambientu doživlja poslušalec. Z akustičnim posnetkom želimo poslušalčevo izkušnjo čim verneje zajeti in jo potem tudi kar najbolj reproducirati, da bi v njem ob ponovnem poslušanju, tokrat posnetka, vzbudila enake občutke. Predpogoj za dober posnetek te vrste je akustično lep ambient in posnetku podrejena postavitev izvajalcev. Ker sta to v praksi zelo selektivni zahtevi, se dandanes tudi v akustičnih ambientih praviloma uporabljajo snemalne tehnike, ki so bližje studijskim.

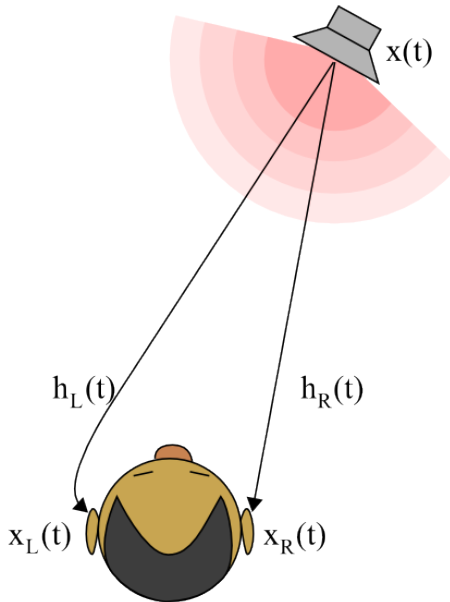
Kljub postopnemu uveljavljanju tehnik prostorskega zvoka („surround“), ki lepo zajamejo in reproducirajo ambient, kar bogati poslušalčevo izkušnjo, se bomo tukaj omejili na stereotehniko, ki je daleč najbolj razširjena. Praviloma gre za dva krajevno ločena uparjena mikrofona za levi in desni kanal, redkeje tri (npr. „Decca Tree“ z dodatnim sredinskim mikrofonom). Uparjenost mikrofonskega para pomeni, da se (kompleksni!) prenosni funkciji mikrofonom, ki ga sestavljata, karseda malo razlikujeta. Pri tem je toleranca izredno majhna, tako da se uparjenih mikrofonom ne da načrtno izdelati, ampak se iz proizvodne serije z naključno posejanimi odstopanji izbere tiste, ki se najbolj ujemajo.

V akustičnem ambientu koincidenčne tehnike (pri teh se membrani dveh usmerjenih mikrofonom nahajata praktično v isti točki) niso zaželeni, saj ambienta ne zaznajo realno. V nasprotju z direktnim zvokom z značilno močno korelacijo med kanaloma je za prostorski zvok, kot ga zaznavamo z razmaknjenima ušesoma, značilna nekoreliranost, in ravno ta daje občutek prostornosti. Ker pa koincidenčni stereopar informacijo za levi in desni kanal zajema v isti točki, sta signala po definiciji fazno povsem korelirana, neglede na to, ali gre za direktni ali prostorski zvok. Dokončno se omejimo še na omnidirekcionalne mikrofone, tj. mikrofone s krogelnosimetrično smerno karakteristiko, ki se jim morajo za vrhunski posnetek usmerjeni mikrofoni vsekakor umakniti, seveda pa so najbolj selektivni, kar se primernosti prostora tiče.

Razmaknjen par omnidirekcionalnih mikrofonom je torej osnova, ki daje najnaravnejši rezultat. V primerjavi z zaznavo z ušesi sta kanala šibkeje intenzitetno ločena, podobnejša, zaradi česar je tudi monokompatibilnost (kvaliteta monosignala, ki ga dobimo s seštevanjem signalov obeh kanalov), kolikor je ta danes sploh še pomembna, slaba. Zato se med mikrofona včasih nameščajo absorptivne ovire (npr. sfera, disk – t. i. Jecklin disk), ki desnemu mikrofonom delno zastirajo akustični pogled na levo in obratno. S tem ohranimo prednosti omnidirekcionalnih mikrofonom in hkrati močnejše intenzitetno ločimo kanala. Tako razmišljanje nas vodi do sklepa, da je najbolj realna ovira kar človeška glava. Snemalna tehnika, pri kateri omnidirekcionalna mikrofona namestimo v ušesni školjki umetne (ali tudi žive) glave, se imenuje dvoušesna (v nadaljevanju „binauralna“) tehnika. Tovrstne posnetke reproduciramo s slušalkami, pri čemer jim po realnosti in

zaznavi prostora ni para. Pri snemanju so odločilne prav vse podrobnosti, od položaja mikrofонов (na milimeter natančno), oblike in velikosti uhljev, do snovi, iz katere je narejena snemalna glava. Zaradi odboja, absorpcije, uklona in resonanc v „votlinah“ ušesne školjke glava tlačno polje ob membranah mikrofонов močno spremeni, pri določenih frekvencah se glasnost poveča za več kot 10 dB. Vemo tudi, da je višjefrekvenčni del zvoka, ki prihaja z leve, ob desnem ušesu oslavljen, saj je uklona okrog glave pri manjših valovnih dolžinah vse manj. Zavedati se moramo, da je enako kot amplituda pomembna tudi faza valovanja, saj v dani točki ob glavi prihaja do interference uklonjenega valovanja, kar daje zaznanemu zvoku značilen spektralni pečat, odvisen od smeri vpadlega valovanja glede na smer glave.

Prenosna funkcija glave



Slika 1. Impulzni odziv na lokaliziran izvir na mestih obeh ušes.

Celoten vpliv glave na tlak v ušesnih školjkah opišemo z impulznim odzivom glave (HRIR, head-related impulse response) za levo in desno uho, $h_L(t, \mathbf{r})$ in $h_R(t, \mathbf{r})$, oziroma njegovo kompleksno Fourierovo transformiranko, prenosno funkcijo glave (HRTF, head-related transfer function), $H_L(\omega, \mathbf{r})$ in $H_R(\omega, \mathbf{r})$, kjer \mathbf{r} podaja lego zvočnega izvira glede na glavo (slika 1). Naj bo $x(t, \mathbf{r})$ tlačni signal izvira pri \mathbf{r} . Tlačna signala v levem in desnem ušesu, $x_L(t)$ in

$x_R(t)$, dobimo s konvolucijo z impulznim odzivom,

$$x_{L,R}(t) = h_{L,R}(t, \mathbf{r}) * x(t, \mathbf{r}) = \quad (1)$$

$$= \int_{-\infty}^{\infty} d\tau h_{L,R}(t - \tau, \mathbf{r}) x(\tau), \quad (2)$$

pri čemer velja $h_{L,R}(t < 0, \mathbf{r}) = 0$, v frekvenčnem prostoru pa preprosto

$$x_{L,R}(\omega) = H_{L,R}(\omega, \mathbf{r}) x(\omega, \mathbf{r}), \quad (3)$$

kjer je $x(\omega, \mathbf{r})$ kompleksna amplituda tlaka izvira pri \mathbf{r} , $x_{L,R}(\omega)$ pa kompleksna amplituda tlaka v levem oziroma desnem ušesu. Impulzni odziv glave se ponavadi premeri v daljnem polju, kjer imamo opravka z ravnimi valovi in je odvisnost od oddaljenosti izvora samo še asimptotična ($\propto 1/r$), tako da je HRIR (in posledično tudi HRTF) netrivialno odvisen „le“ od smeri, ki jo opišemo s kotoma ϑ (zenitna razdalja) in φ (azimut), $\mathbf{r} = (r, \vartheta, \varphi)$. Če torej zvočilo premikamo v vodoravni ravnini, ki poteka skozi glavo (ekvatorialna ravnina), je $\vartheta = \pi/2$ in se spreminja kot φ , pri čemer naj $\varphi > 0$ pomeni, da zvok prihaja z leve.

V načelu je tlak v ušesih s HRIR ali HRTF torej popolnoma določen. V praksi seveda nastopijo problemi: nezadostno število izmerjenih smeri, raznolikost ušesnih školjk, odvisnost prenosne funkcije od točnega položaja v ušesu, vpliv bližnjega polja pri velikih valovnih dolžinah ... Če želimo v prenosni funkciji glave videti kaj več kot nabor kompleksnih podatkov v odvisnosti od kotov ϑ in φ , si moramo ogledati fizikalne zakonitosti in koncepte, ki nastopajo pri vpadu skalarne valovanja na oviro.

Vpliv površine

Valovanje, ki pod vpadnim kotom α pada na neskončno ravno površino s specifično mehansko (= akustično) impedanco Z , se odbije po odbojnem zakonu, amplituda odbitega valovanja tik ob površini pa je

$$p_1 = \frac{\frac{Z}{Z_0} \cos \alpha - 1}{\frac{Z}{Z_0} \cos \alpha + 1} p_0, \quad (4)$$

kjer je $Z_0 = \sqrt{\rho/\chi} \approx 410 \text{ kg m}^{-2} \text{ s}^{-1}$ specifična akustična impedanca zraka, p_0 pa amplituda vpadnega valovanja tik ob površini. Običajno velja $Z \gg Z_0$ (impedanca človeškega mehkega tkiva [2] npr. znaša okrog $Z \approx 1.6 \cdot 10^6 \text{ kg m}^{-2} \text{ s}^{-1}$), tako da je $p_1 \approx p_0$ praktično za vse vpadne kote ($\cos \alpha = 1$ pomeni pravokotni vpad). To pomeni, da je ob takšni površini amplituda tlaka dvakrat večja kot v vpadnem valovanju, kar predstavlja 6 dB večjo glasnost (pojav izkorišča t. i. PZM – „pressure zone microphone“, ki se ga namesti na tla ali steno).

Pri glavi sta koristna dva mejna primera. Za valovne dolžine, majhne v primerjavi s premerom glave, le-ta pomeni veliko in ravno površino, za katero velja zgornja ugotovitev. Visokofrekvenčno valovanje, ki prihaja s strani, bo torej na vpadni strani ob površini za 6 dB ojačano. Na nasprotni strani pa bo zaradi šibkega uklona pri teh valovnih dolžinah glede na vpadno valovanje celo oslabiljeno. Drugače je pri nizkih frekvencah: za valovanje z valovno dolžino, veliko v primerjavi s premerom glave, le-ta ne pomeni nobene ovire. V tem primeru torej nikjer ob glavi ni sprememb glasnosti. Značilno frekvenco, okrog katere se na vpadni strani glave zgodi prehod od nespremenjene k povečani glasnosti, ocenimo z zahtevo, da je valovna dolžina enaka „premeru“ glave, na okrog 1700 Hz. Eksaktno pa velja, da je ta frekvenca obratnosorazmerna z velikostjo glave.

Sipanje valovanja na krogli

V tem razdelku si bomo ogledali eksaktno rešitev sipanja ravnega zvočnega vala na togi krogli (slika 2). Ta podaja polno informacijo o akustičnem polju v prostoru okrog krogle. Če jo preberemo na površini krogle, pa predstavlja točno prenosno funkcijo toge krogle, ki jo ob izmerjenih HRTF kot eno od opcij običajno srečamo v računalniških programih za auralizacijo (simulacijo zvoka, ki ga slišimo). Zaradi krogelne simetrije je prenosna funkcija v tem primeru, drugače kot pri prenosni funkciji dejanske glave, odvisna le od enega kota.

Iz linearizirane Eulerjeve enačbe (Newtonov zakon za gibanje idealne tekočine),

$$\rho \frac{\partial \mathbf{v}}{\partial t} = -\nabla p, \quad (5)$$

kjer sta $p(\mathbf{r}, t)$ in $\mathbf{v}(\mathbf{r}, t)$ tlačno in hitrostno polje, ρ pa gostota zraka, linearizirane kontinuitetne enačbe za maso,

$$\frac{\partial \rho}{\partial t} + \rho \nabla \cdot \mathbf{v} = 0, \quad (6)$$

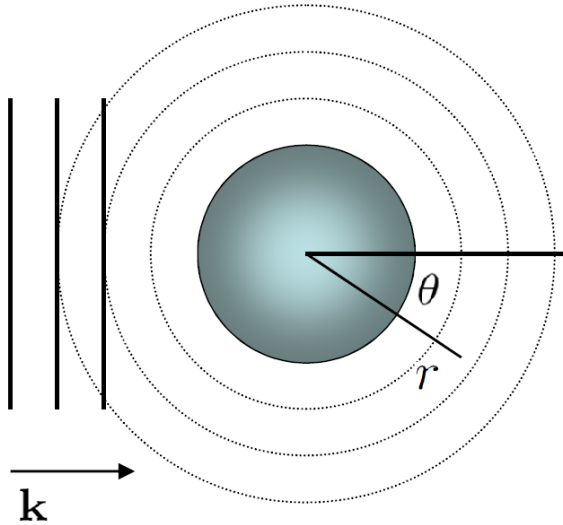
in linearizirane enačbe stanja,

$$\frac{d\rho}{\rho} = \chi dp, \quad (7)$$

kjer je χ adiabatna stisljivost, sledita valovna enačba za tlačno polje in z nastavkom $p = p_0 e^{-i\omega t}$ amplitudna enačba:

$$\nabla^2 p - \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} = 0, \quad \nabla^2 p_0 + k^2 p_0 = 0, \quad (8)$$

kjer je $c = 1/\sqrt{\rho\chi}$ hitrost valovanja in $k = \omega/c$ velikost valovnega vektorja. Rešitve slednje zapišimo v sfernih koordinatah [1]. Rešitve radialnega dela



Slika 2. Sipanje ravnega vala na togi krogli. Ravni val potuje v smeri $\theta = 0$. Rešitev za celotno tlačno polje je osnosimetrična glede na to smer in je tako odvisna le od polarnega kota θ .

so sferne Besselove in Neumannove funkcije, $j_l(kr)$ in $n_l(kr)$, s katerimi zapišemo potujoče krogelno valovanje, ki se širi iz koordinatnega izhodišča navzven, v obliki sfernih Hanklovih funkcij prve vrste, $h_l^{(1)} = j_l + in_l$. Osnosimetrične rešitve kotnega dela pa so Legendrovi polinomi $P_l(\cos \theta)$.

Tlačno polje okrog krogle (slika 2) sestavimo iz vpadajočega ravnega vala z amplitudo 1 in od površine krogle odbitih krogelnih valov,

$$p_0(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} + \sum_{l=0}^{\infty} A_{k,l} h_l^{(1)}(kr) P_l(\cos \theta), \quad (9)$$

kjer so $A_{k,l}$ koeficienti, ki jih moramo določiti. Fazo ravnega vala zapišemo kot $\mathbf{k} \cdot \mathbf{r} = kr \cos \theta$ in ga razvijemo po krogelnih funkcijah,

$$p_0(\mathbf{r}) = \sum_{l=0}^{\infty} \left[(2l+1) i^l j_l(kr) + A_{k,l} h_l^{(1)}(kr) \right] P_l(\cos \theta). \quad (10)$$

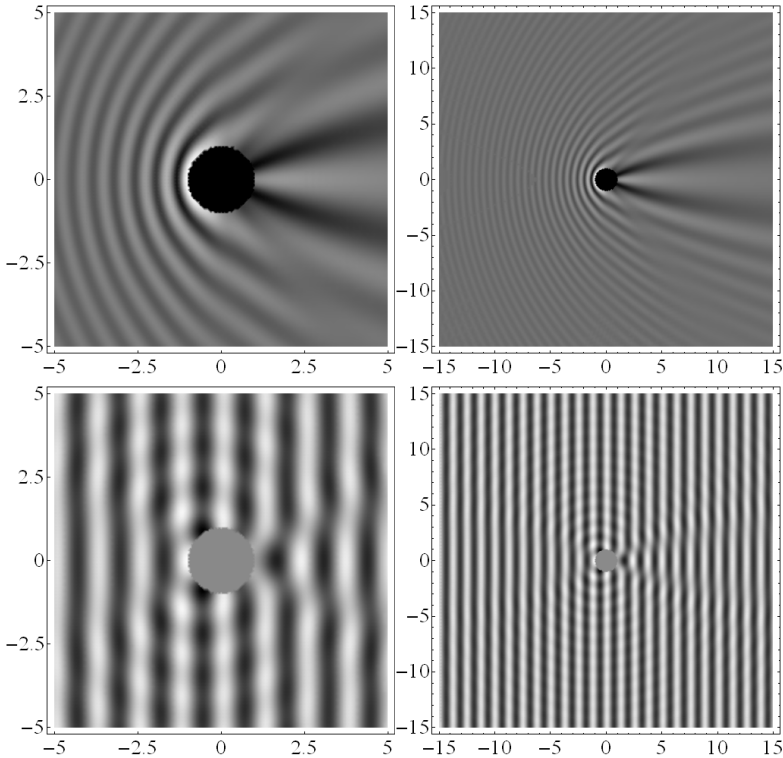
Na površini toge krogle s polmerom R mora za radialno komponento hitrosti veljati $v_r(R) = 0$, od koder iz enačbe (5) sledi robni pogoj za tlak:

$$\left. \frac{\partial p_0}{\partial r} \right|_{r=R} = 0, \quad (11)$$

z njim pa iz enačbe (10) iskani koeficienti:

$$A_{k,l} = -(2l + 1)i^l \frac{j'_l(kR)}{h'_l(kR)}. \quad (12)$$

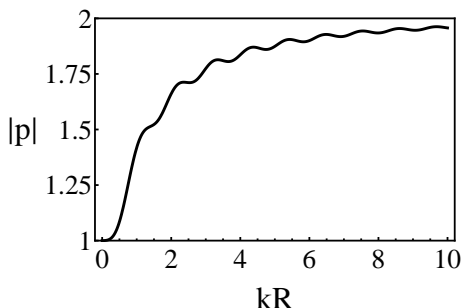
Rešitev za akustično polje, ki je podana z enačbo (9) in koeficienti (12), je vsota (interferenca) vpadajočega ravnega vala in odbitih krogelnih valov. Seštevati moramo do dovolj velikega l , da je prispevek členov z višjimi l zanemarljiv. Slika 3 (zgoraj) prikazuje amplitudo tlaka okrog krogle (absolutna vrednost kompleksnega tlaka po enačbi (9)) za $kR = 5$. Lepo je viden porast tlaka pred kroglo in uklonski minimumi za njo, kakor tudi interferenca vpadajočega ravnega vala z odbitimi krogelnimi valovi, ki z razdaljo od krogle postopno pojema (daleč stran ostane samo ravni val, tj. konstantna amplituda). Na sliki 3 (spodaj) pa je primer trenutnega tlačnega polja. Zanimiva je zakasnitev valovanja za kroglo.



Slika 3. Sipanje ravnega vala na togi krogli, ravni val s $kR = 5$ vpada z leve; (zgoraj); amplituda tlaka po enačbi (9) in trenutna slika tlačnega polja (spodaj). Desni sliki prikazujeta razmere dlje od krogle.

HRTF v modelu toge krogle

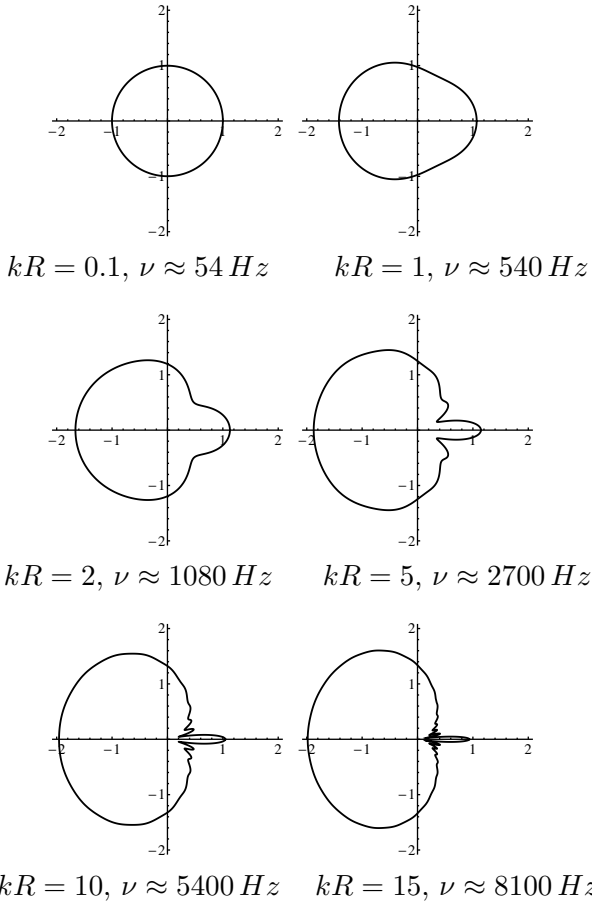
Za binauralni zajem zvoka je seveda pomemben tlak ob površini krogle, kamor namestimo mikrofona. Najprej si na sliki 4 oglejmo, kako je amplituda tlaka na površini pri $\theta = \pi$ (vpadna stran) odvisna od k , torej od frekvence $\nu = kc/2\pi$, oziroma valovne dolžine $\lambda = 2\pi/k$. Vidimo, da je pri nizkih frekvencah tlačna amplituda zares enaka amplitudi vpadajočega vala, pri visokih frekvencah pa postane skladno z napovedjo dvakrat večja. Prehod se zgodi okrog $kR \approx 2$ oziroma $\lambda \approx \pi R$. Na vpadni strani je torej visokofrekvenčni del spektra dvignjen za 6 dB.



Slika 4. Amplituda tlaka na površini krogle pri $\theta = \pi$ v odvisnosti od brezdimenzijske velikosti valovnega vektorja kR .

Slika 5 prikazuje polarne diagrame amplitude tlaka na površini krogle za naraščajoče frekvence. Ravni val vpada z leve. Na tej strani je tlak pri višjih frekvencah torej povečan do faktorja 2. Na senčni strani, kjer so višje frekvence oslabiljene, prihaja do izrazitih interferenčnih efektov, pri čemer se položaj ojačitvenih izrastkov z valovno dolžino živahno spreminja. V dani smeri so zato nekatere frekvence močno oslabiljene – za senčno stran je torej značilna izrazita koloracija zvoka. Zanimiv je uklonski preostanek pri $\theta = 0$, ki z naraščajočo frekvenco ne pojema, ampak se le tanjša. Tak centralni uklonski maksimum srečamo tudi pri uklonu svetlobe na okrogli ploščici.

Slika 6 združeno prikazuje odvisnost amplitude tlaka na površini krogle od kota θ in od frekvence v enotah kR ($kR = 2\pi\nu R/c$), ki je na grafih predstavljena z razdaljo od izhodišča. Nizkofrekvenčni del spektra se nahaja na sedlastem območju v okolici izhodišča. Ker je kompleksna amplituda vpadajočega vala kar 1, je kompleksni tlak $p(R, \theta)$ iz enačbe (9) pravzaprav že iskana kompleksna prenosna funkcija toge krogle (na grafih 4–6 smo prikazovali le amplitudo, torej absolutno vrednost kompleksnega tlaka). Paziti je treba le še pri kotni odvisnosti, saj je prenosna funkcija funkcija smeri, iz katere prihaja valovanje, in ne smeri točke na krogli, v kateri merimo tlak. Kompleksni prenosni funkciji toge krogle za levo in desno uho, torej HRTF v približku toge kroglaste glave, za zvočilo v ekvatorialni ravnini ($\vartheta = \pi/2$)



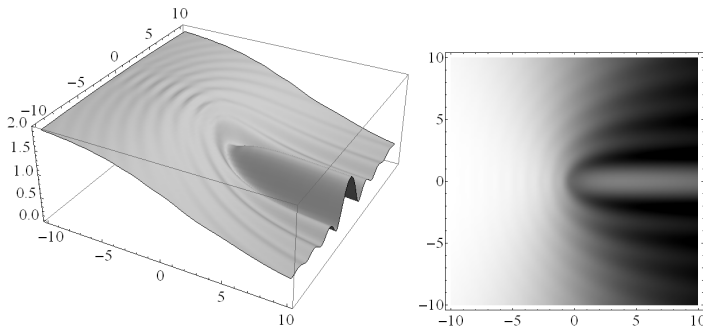
Slika 5. Amplituda tlaka na površini krogle v odvisnosti od kota θ za naraščajoče vrednosti kR . Pripisana je orientacijska vrednost frekvence za primer $R = 10 \text{ cm}$. Vpadna stran je na levi, senčna ($\theta = 0$) pa na desni.

sta tako

$$H_L(\vartheta = \frac{\pi}{2}, \varphi) = p(R, \frac{\pi}{2} + \varphi), \quad (13)$$

$$H_D(\vartheta = \frac{\pi}{2}, \varphi) = p(R, \frac{\pi}{2} - \varphi), \quad (14)$$

kjer $p(R, \theta)$ podaja enačba (9). Za zvočilo v splošni smeri (ϑ, φ) upoštevamo osno simetrijo, izračunamo kot Θ glede na os od desnega ušesa proti levemu,



Slika 6. Odvisnost amplitude tlaka na površini krogle od frekvence v enotah kR (predstavlja jo oddaljenost od izhodišča) in kota θ (kaže ga polarni kot). Ravni val vpada z leve. Oba grafa prikazujeta isto funkcijo. Pozor: to je dvodimenzionalna predstavitev tlaka na površini krogle, ne tlačno polje okrog krogle kot na sliki 3.

$\cos \Theta = \sin \vartheta \sin \varphi$, in zapišemo prenosni funkciji v splošnem:

$$H_L(\vartheta, \varphi) = p(R, \pi - \Theta), \quad (15)$$

$$H_D(\vartheta, \varphi) = p(R, \Theta). \quad (16)$$

Sklep

Zvok, zajet s snemalno glavo, je moduliran in prek odziva snemalne glave (HRIR ali HRTF) nosi „prstni odtis“ glave, ki je bila uporabljena. Binauralni posnetek se posluša s slušalkami, praviloma odprtimi, ker najmanj spremenijo prenosno funkcijo obušesnega območja. Seveda se ne moremo izogniti biološkim raznolikostim. Morda le tako, da mikrofona namestimo kar v svoji ušesi – vendar se pri tem takoj pojavi problem točne pozicije in ponovljivosti nameščanja. Binauralni posnetek, predvajan prek stereo zvočnikov, je ob poslušanju že drugič konvolviran z odzivom glave in poleg nekoliko oddaljene zvočne perspektive deluje presvetlo, zato ga je treba vsaj ekvilizirati. Večji problem je mešanje kanalov, saj zvok levega zvočnika dospe tudi v desno uho in obratno. Z dekonvolucijo je ta t. i. „cross-talk“ možno v principu odstraniti, a le za točno določen položaj poslušalčeve glave glede na geometrijo zvočnikov, individualno prilagojeno HRTF, gluho poslušalnico ... Prav tako so problematične nizke frekvence, ki jih običajno zaznavamo že v bližnjem polju. Ker je dekonvolucija inverzni problem, so vse neidealnosti toliko bolj kritične.

LITERATURA

- [1] I. Kuščer in A. Kodre, *Matematika v fiziki in tehniki*, DMFA Slovenije, Ljubljana, 1994.
- [2] I. Ogura *et al.*, v *Ultrasound in Medicine*, Vol. 4, urednika D. White in E. A. Lyons, 535–543, Plenum Press, New York, 1978.