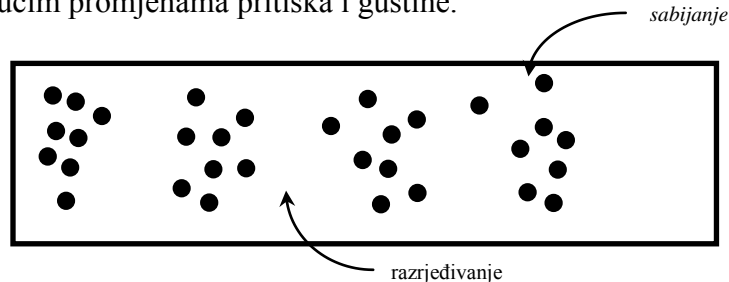


OSObine I PERCEPCIJA ZVUKA

FIZIČKE KARAKTERISTIKE ZVUČNOG TALASA

Do generisanja zvuka dolazi pri promjeni stacionarnog stanja čestica koje je praćeno odgovarajućim promjenama pritiska i gustine.



Do promjene stacionarnog stanja čestica može da dođe usljed vibriranja žice ili bubnja. Akustičke oscilacije se šire konačnom brzinom u vidu zvučnih talasa od mjesta gdje je došlo do poremećaja ravnotežnog stanja. Širenje zvučnih talasa je, dakle, uslovljeno oscilovanjem molekula materijalne sredine kroz koju se oni prostiru. Zvučni ili akustički talasi su mehanički longitudinalni talasi koji mogu da se prostiru kroz sredine svih agregatnih stanja.

Osnovne fizičke karakteristike zvučnog talasa (tzv. *objektivne karakteristike zvuka*) su: **osnovna frekvencija**, **zvučni spektar** i **intenzitet zvučnog talasa**- zvuka.

Frekvencija zvuka zavisi od talasne dužine zvučnog talasa. Talasna dužina je razdaljina koju prelazi talas u jednom periodu odašiljanja. Dakle, postoji direktna veza između talasne dužine i frekvencije. Niska frekvencija odgovara velikim talasnim dužinama i obratno. Zvuk se (prenosi) odašilja kroz vazduh sa brzinom od otprilike 340 metara u sekundi. To znači da jedan ton 1 Hz ima talasnu dužinu od 340 metara. Vrijedi da je $\lambda = \frac{v}{f}$, gdje je λ talasna dužina u metrima, v brzina prostiranja zvučnog talasa u m/s i f frekvencija u Hz. Tako, npr. frekvenciji od 20 kHz odgovara talasna dužina od 1,7 cm.

Ovo je važno znati da bi smo mogli odrediti koju visinu tona možemo pustiti u nekoj sobi. Naime, da bi smo pratili ton dužina sobe treba da bude veća ili jednaka talasnoj dužini tona. Npr., ako želimo dobar prijem tonova od 50 Hz treba nam dužina sobe koja je jednaka polovini talasne dužinu od $340/50 = 6,80$ metara, dakle 3,40 metara. Za prijem najnižeg tona iz čujnog opsega (otprilike 20 Hz), potrebna dužina prostorije iznosi 8,5 metara.

Zavisno od opsega osnovnih frekvencija, zvučne signale dijelimo na:

- Infrazvuk: 0-20 Hz
- Zvuk iz čujnog opsega: 20 Hz – 20 kHz
- Ultrazvuk: 20 kHz – 1 GHz
- Hiperzvuk: 1 GHz – 10 THz

Prema obliku zvučnog spektra zvuk se dijeli na **tonove**, koje karakteriše periodičnost promjena u sredini kroz koju se prostiru i **šumove** ili buku koji su neperiodičnog karaktera. Čist ton nastaje kada čestice sredine vrše harmonijsko oscilovanje, dok složen muzički ton nastaje superpozicijom viših harmonika. Složen ton za koga kažemo da nije muzički ima frekvencije viših tonova koji nisu cjelobrojni umnošci osnovne frekvencije. Šum predstavlja složen neperiodičan signal čije se karakteristike neprekidno mijenjaju. Prasak ili udarna buka predstavlja složen zvučni signal koji naglo nastaje i nestaje.

Amplituda zvuka je mjera odstupanja pritiska od srednje vrijednosti, stanja koje je označeno kao tišina. **Nivo zvučnog pritiska** je objektivna mjera jačine zvuka koja se definiše u odnosu na referentnu vrijednost kao: $L[dB] = 20 \log \frac{p}{p_o}$, gdje je p zvučni pritisak, a p_o zvučni pritisak koji odgovara pragu čujnosti na 1000 Hz i iznosi 10 pW/m^2 .

Multimedijalni sistemi koriste zvukove iz opsega koji čovjek može da čuje. U daljnjem razmatranju ćemo zvukove iz tog opsega nazivati *audio* a odgovarajuće talasne oblike *audio signalima*.

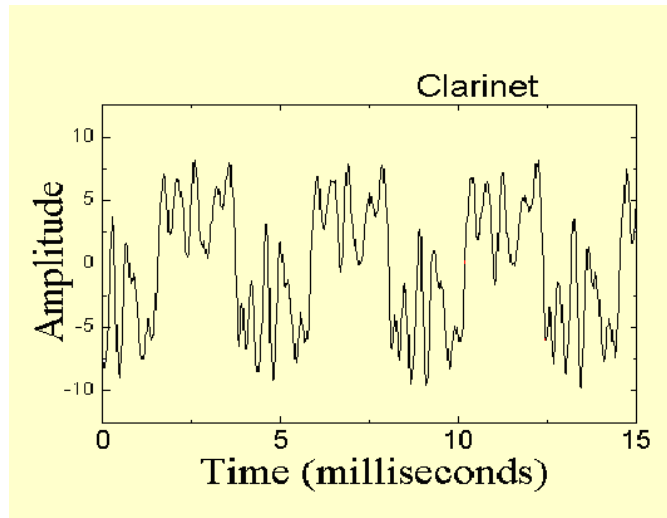
PERCEPCIJA

Svojim fizičkim veličinama zvuk djeluje na čovjeka preko organa sluha. Nastale fizičke promjene stvaraju određene subjektivne osjećaje, koje u izvjesnoj mjeri zavise od odgovarajućih fizičkih veličina zvuka i njihovih promjena. Karakteristike zvuka u domenu subjektivnog osjećaja (tzv. *subjektivne psihološke ili biofizičke karakteristike zvuka*) koje odgovaraju objektivnim veličinama su: visina tona, boja zvuka i glasnoća (subjektivna mjera jačine zvuka).

Visina tona je osjećaj na osnovu koga se može reći da je jedan ton visok ili nizak (dubok) vezana je prvenstveno za frekvenciju. Kod složenog tona bitna je frekvencija osnovnog tona. Postoji diferencijalni prag visine tona koji se može zapaziti (df/f). On je konstantan samo na srednjim frekvencijama. Na osjećaj visine tona utiče i intenzitet zvuka. Zvuk vrlo malog intenziteta na pragu čujnosti ne daje osjećaj visine tona.

Po usvojenoj akustičkoj terminologiji, koncertni ton (ton A iznad srednjeg C) ima frekvenciju od 440 Hz. Najveća muzička mjera opsega tonova je oktava. Ton koji je za oktavu viši od nekog datog tona ima zapravo dva puta veću frekvenciju, dok ton koji je za oktavu niži ima dva puta manju frekvenciju. Oktava je podijeljena u 12 polustepeni koji odgovaraju notama (C, Cis, D, Dis, E, F, Fis, G, Gis, A, Ais, H).

Boja zvuka je kvalitet koji dopušta da se prepoznaju dva zvuka iste visine i istog osjećaja jačine kada su emitovani iz dva različita zvučna izvora. Ona je određena brojem viših harmonika, odnosom njihovih intenziteta i njihovom faznom razlikom u odnosu na osnovni ton.



Prenos zvuka zavisi od sredine u kojoj se prostiru zvučni talasi. Gledano sa aspekta multimedijalnih sistema prostiranje zvučnih talasa je u vazduhu. U trenutku pojave zvučnih talasa dolazi do promjene pritiska u prostoru. Uho je prijemnik zvuka upravo zahvaljujući registrovanju promjena pritiska.

Posredstvom slušnog sistema fizikalni podražaj u obliku zvuka se pretvara u subjektivni doživljaj. Između tri karakteristike subjektivnog osjećaja - glasnoće, visine i boje, te tri karakteristike objektivnog stimulansa koje se u njima odražavaju – njegovog intenziteta (ili čujnosti), frekvencije i spektralnog sastava, moguće je uspostaviti odnos. Eksperimentalno je ustanovljeno da ista objektivna jačina zvuka neće izazvati isti subjektivni osjećaj glasnoće na različitim frekvencijama.

Karakteristike podražaja su mehaničke veličine koje se mjere u jedinicama mehanike dok se karakteristike subjektivnog osjećaja mjere metodama eksperimentalne psihologije. Sa svoje strane eksperimentalna psihologija se služi statističkim metodama. **Fonometrija**, kao tehnika mjerenja subjektivnog osjećaja sluha, u pojedinačnim slučajevima mjeri karakteristike zvučnog stimulansa pa iz tih rezultata određuje karakteristike osjećaja na osnovu opštih odnosa između jednih i drugih karakteristika.

Visina tona je subjektivna karakteristika zvuka određena prije svega frekvencijom zvuka odnosno osnovnom frekvencijom kod složenog zvuka. Manja frekvencija daje osjećaj nižeg tona, a veća frekvencija višeg. **Visina tona je proporcionalna logaritmu frekvencije**. Odavde proističe pravilo da se u svim dijagramima koristi logaritamska skala za frekvenciju.

Uho je veoma osjetljivo na promjene frekvencije. U području ispod 500 Hz najmanja promjena frekvencije koja se može primijetiti iznosi 3 Hz. Eksperimentalni rezultati pokazuju da u cijelom čujnom području ima oko 850 tonova koji se mogu razlikovati po visini. Zbog nalaženja intervala koji melodijski izgledaju jednako, formirana je tzv. melodijska skala gdje se koristi jedinica " mel", pri čemu se mel i herc poklapaju u oblasti ispod 500 Hz dok se povećanjem frekvencija javljaju sve veće promjene.

Maksimalna jačina zvuka koju čovjek može još jasno da čuje je prag bola ili gornja granica čujnosti koja pri 1000 Hz iznosi 10 W/m^2 , što odgovara pritisku koji je 10^{12} puta

veći od pritiska na pragu čujnosti. Zvuk oko nas rijetko prelazi prag bola koji dakle iznosi 120 dB do 130 dB. Npr. nivo buke u mirnoj kući iznosi 35 dB, dok je nivo buke u bučnoj ulici 70 dB.

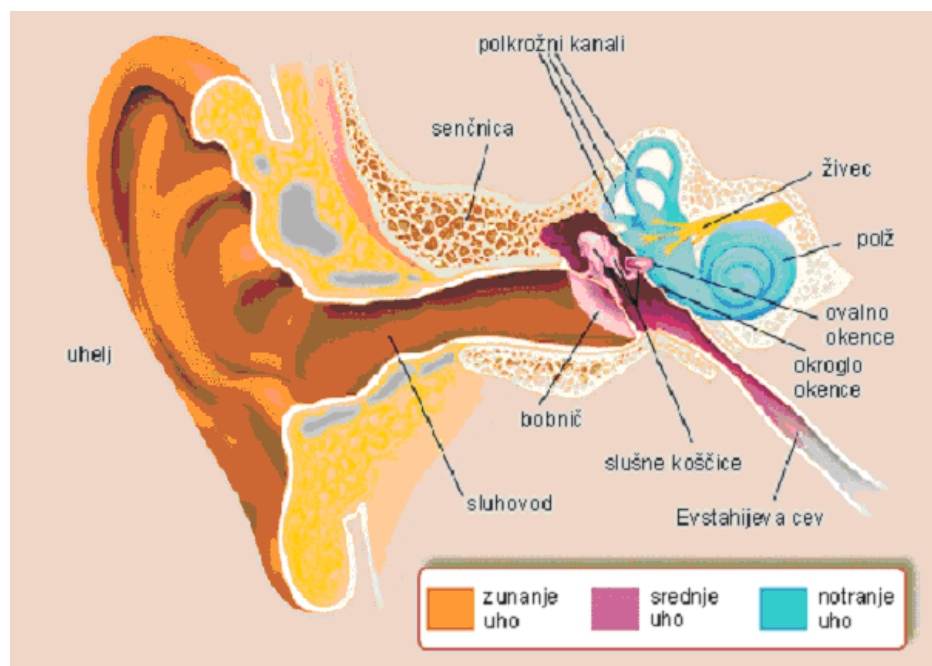
Percepcija glasnosti zvuka zavisi od frekvencije: na nižim frekvencijama zahtijeva se više snage da se isporuči jednaka primljena glasnost kao na srednjim i visokim frekvencijama.

Zvuk se ponekad više osjeća nego čuje. Npr. nivo buke na radnom mjestu viši od 90 dB dovodi do razdražljivosti i grešaka, posebno ako su u buci prisutne visoke frekvencije. Ako je nivo buke viši od 80 dB nemoguće je koristiti telefon. Eksperimenti su pokazali da nivo buke do 45 dB ne proizvodi reakciju, od 45-55 dB proizvodi sporadične žalbe, od 50-60 dB žalbe na sve strane, od 55-65 dB prijetnje, a preko 65 dB žestoke reakcije. Ovakva istraživanja mogu poslužiti kao koristan vodič za stvaraoce multimedije.

Dinamički opseg uha dakle iznosi oko 130 dB. Diferencijalni prag osjetljivosti uha iznosi 1 dB i predstavlja najmanju razliku u nivou jačine zvuka koja se može uhom razlikovati. Naravno da ovi podaci predstavljaju statističku srednju vrijednost.

Osjećaj zvuka zavisi od mnogo faktora: jačini zvuka koji je nastao u njegovom izvoru, vrsti medija u kome je generisan, vrsti medija kroz koga prolazi, o udaljenosti i smjerovima u kojima se zvuk širi, ambijentu, godinama slušaoca i drugim subjektivnim faktorima.

S anatomskog gledišta možemo čovječje uho podijeliti u tri anatomske cjeline. To su vanjsko, srednje i unutrašnje uho. Vanjsko uho, koje obuhvata ušku i zvukovod, vrši sabiranje (kolekciju) i dovodi zvukove do bubnjića. Srednje uho služi, kao aparat transmisije, za prenošenje zvukova a obuhvata: bubnjić, slušne koščiće, slušne mišiće, tubu i pneumatske prostore sljepočne kosti. Unutrašnje uho prepoznaje frekvencijski sastav zvuka i djelomično ga interpretira, a ovaj proces se završava na nivou slušnih centara u mozgu. Recimo i to da balbularne jezgre prenose podražaje do malog mozga i služe održanju podsvjesnih refleksa. Kortikalno središte, napokon, omogućava osjećaj zvuka i ravnoteže. To je, možemo reći, aparat percepcije.

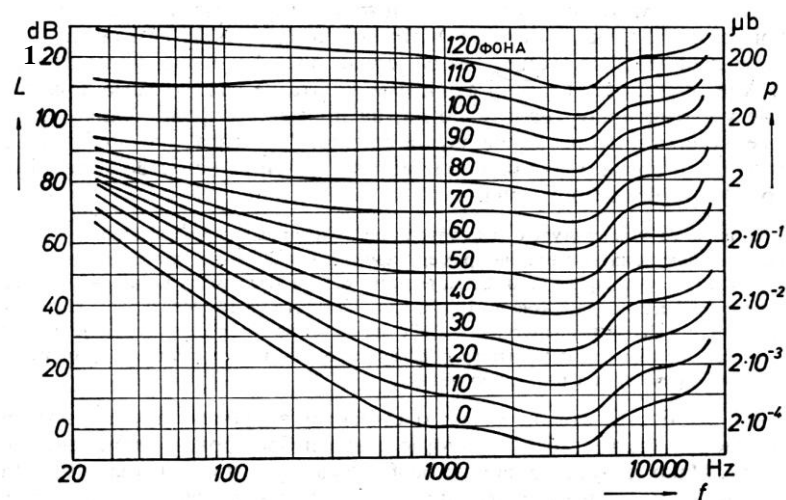


Akustički dio tog centra nalazi se u mozgu. Akustičke vibracije nailaze na impedansu uha koja je zavisna od omskog otpora (trenje), induktivnog otpora (masa), kapacitivnog otpora (rigiditet) te frekvencije. Kod povećanja mase opaža se gubitak u oblasti visokih tonova a kod povećanja rigiditeta gubitak u području dubokih tonova.

Dijelovi uha, dakle, služe kolekciji, transmisiji, selekciji i percepciji akustičkih podražaja. Sa funkcionalnog gledišta dijelimo uho u samo dva dijela: u konduktivni i perceptivni aparat. Pod konduktivnim aparatom podrazumijevamo sve dijelove uha koji služe prenošenju mehaničke vibracione energije do osjetnih stanica unutrašnjeg uha. Perceptivni aparat se završava u kortikalnim centrima.

Uho je različito osjetljivo na razne tonove. Tako je osjetljivost uha na duboke tonove vrlo slaba, dok je na visoke tonove velika i postiže kulminaciju između 2 kHz i 3 kHz, a zatim se naglo spušta prema gornjoj granici sluha. To znači da je za duboke tonove potreban znatno veći intenzitet nego za visoke tonove da bi ih ljudsko uho moglo čuti.

Glasnoća zvuka je fiziološka veličina koja opisuje osjećaj jačine zvuka u našem uhu. Odnos između objektivne jačine zvuka (L) i subjektivnog osjećaja, glasnoće (A), određuje se eksperimentalno. Nivo glasnoće se izražava bezdimenzionalnom jedinicom *fon* koja je analogna decibelu. Naime, nivo glasnoće od n fona ima onaj zvuk čija je glasnoća jednaka glasnoći koju izaziva zvuk frekvencije od 1000 Hz i nivoa zvučnog pritiska n dB iznad praga čujnosti. Fonska skala oslanja se na subjektivni osjećaj glasnoće, dok je decibelska skala usmjerena ka mjerenju jačine zvuka. Krive koje uspostavljaju vezu između objektivne jačine i subjektivnog osjećaja glasnoće dobijene su na osnovu ispitivanja grupe slušalaca koji su, slušajući sa oba uha, upoređivali glasnoće tonova različite visine sa glasnoćom referentnog tona frekvencije 1 kHz. Pri tome se jačina zvuka mijenja dok se glasnoća oba tona ne izjednači. Nivoi zvučnih pritisaka izmjereni u tim trenucima u decibelima na raznim frekvencijama određuju jednu krivu sa konstantnim subjektivnim osjećajem glasnoće. Ponavljanjem postupka za različite glasnoće dobiju se *izofonske krive*, prikazane na slici.



Izofonske krive

Psihološkim istraživanjima je utvrđeno da se Weber-Fechnerov psihofizički zakon može smatrati dovoljno tačnim kad se radi o doživljaju zvuka: ako se intenzitet nadražaja, u ovom slučaju jačina zvuka, mijenja linearno, tada se nivo čujnosti mijenja logaritamski.

Sa slike koja prikazuje izofonske krive je vidljivo da je uho standardnog slušaoca najosjetljivije u srednjefrekvencijskom području, od 2 kHz do 4 kHz.

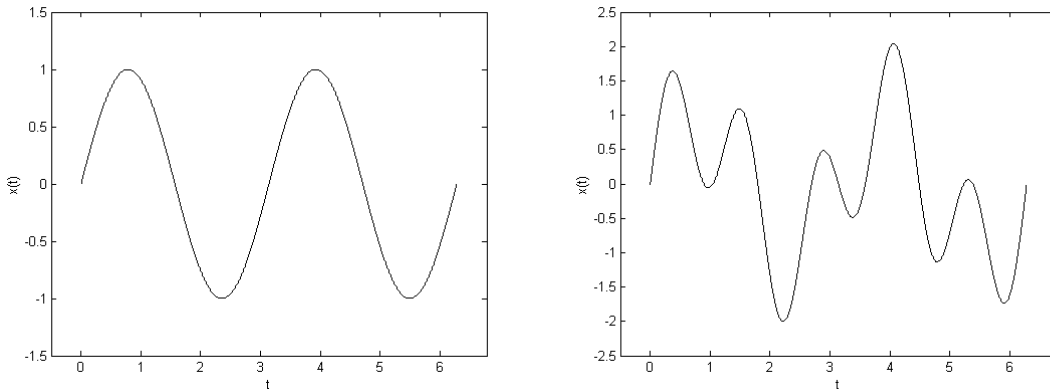
Međutim, ove krive koje prikazuju koji je intenzitet pojedinih tonova potreban kako bismo ga čuli određenom glasnoćom, mijenjaju se u realnoj situaciji. Pri pojavi tona velikog intenziteta, prag čujnosti za tonove koji se nalaze u blizini njega se povećava. Zato se dešava da susjedne tonove ne čujemo ukoliko je jedan od njih mnogo većeg intenziteta od ostalih, iako je intenzitet i tih tonova iznad praga čujnosti. Ta pojava se naziva “*audio maskiranje*”.

DIGITALIZACIJA AUDIO SIGNALA

Zvuk koji nastaje kao posljedica varijacija akustičkog pritiska se koristeći mikrofonski prevodi u električni signal (napon ili struja). Mikrofon se sastoji od dijafragme koja spojene sa namotajem žice koji se kreće u magnetnom polju. Zvučni talas pomjera dijafragmu što dovodi do kretanja namotaja žice u magnetnom polju i indukovanja struje. Struja se prenosi do zvučnika koji imaju svoju sopstvenu zavojnicu u magnetnom polju. Promjena struje uzrokuje promjene magnetnog polja koje utiču na pomijeranje druge zavojnice vezane na dijafragmu čime se mijenja vazdušni pritisak i proizvodi zvuk.

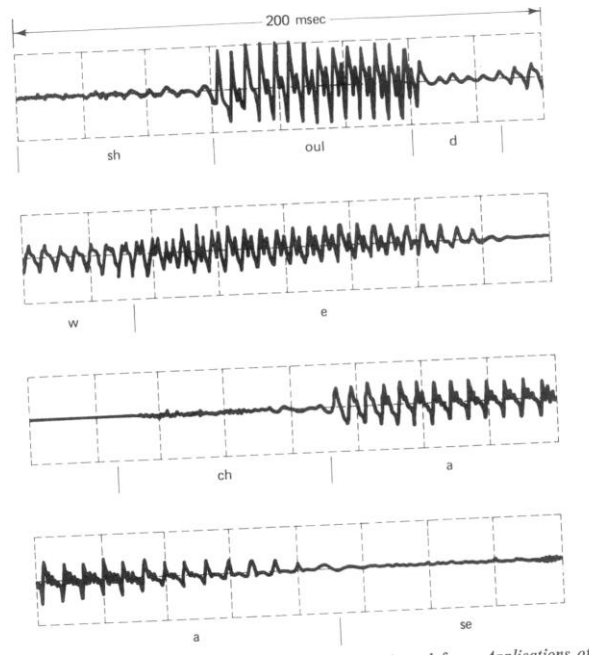
Prije nego počnemo ozbiljnije razmatranje o audio signalima, ponovimo osnovnu podjelu signala. Ako signal postoji u svakom trenutku vremena ili ne, govorimo o *kontinualnom* ili *diskretnom* signalu u vremenu, respektivno. Za signale koji mogu da poprime proizvoljnu vrijednost iz dozvoljenog opsega kažemo da su *kontinualni po amplitudi*, dok za signale čije vrijednosti amplitude pripadaju konačnom skupu kažemo da su *kvantovani*. Ako je signal kontinualan u vremenu i po amplitudi kažemo da se radi o *analognom* signalu. Za signal diskretan u vremenu i kvantovane amplitude kažemo da je *digitalan*.

Audio signali su po svojoj prirodi analogni signali.



Analogni audio signali

Kao što se može vidjeti sa sljedeće slike, različiti govorni glasovi odgovaraju različitim oblicima signala. Slično je i sa ostalim oblicima audio signala.



Primjer zapisa govornog signala

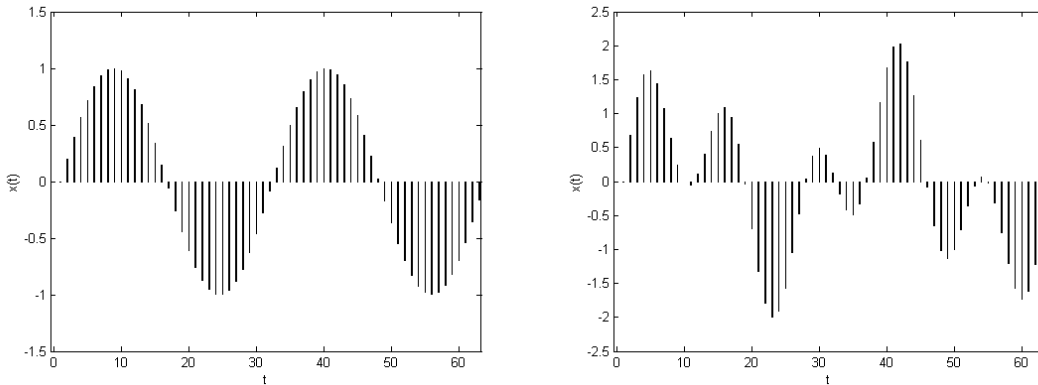
Mikrofoni emituju veoma nizak naponski nivo, oko jednog milivolta. Nivoi analognih audio signala sa kojima se radi su podijeljeni u dvije kategorije: potrošački nivo od 0,316 V i profesionalni nivo od 1,23 V i više.

Optimalni opseg frekvencija audio signala je od 20Hz do 20KHz. Dinamički opseg CD-a je 96dB, dok je dinamički opseg audio kasete 48 dB.

U multimedijalnim sistemima radimo sa digitalnim signalima. Pri digitalizaciji audio signala neophodno je da digitalna verzija audio signala dočara što vjernije analogni audio

signal. Digitalni audio signali nastaju u procesu odmjerenja (uzimanja uzoraka signala u vremenu) i kvantizacije. Vremenski intervali između dva susjedna odmjerenja ne moraju biti jednaki, iako ćemo mi u daljnjem izlaganju smatrati da se odmjerenje vrši u jednakim vremenskim intervalima.

Signal koji nastaje odmjerenjem definisan je samo u pojedinim odvojenim (diskretnim) vremenskim trenucima. Nazivamo ga diskretni signal. Njegova amplituda može, ali ne mora biti kontinualna. Amplituda odmjerača signala $x_k = x(kT_0)$ se grafički predstavlja u zavisnosti od rednog broja odmjerača k . Na slici je prikazan audio signal diskretan u vremenu sa kontinualnom amplitudom.



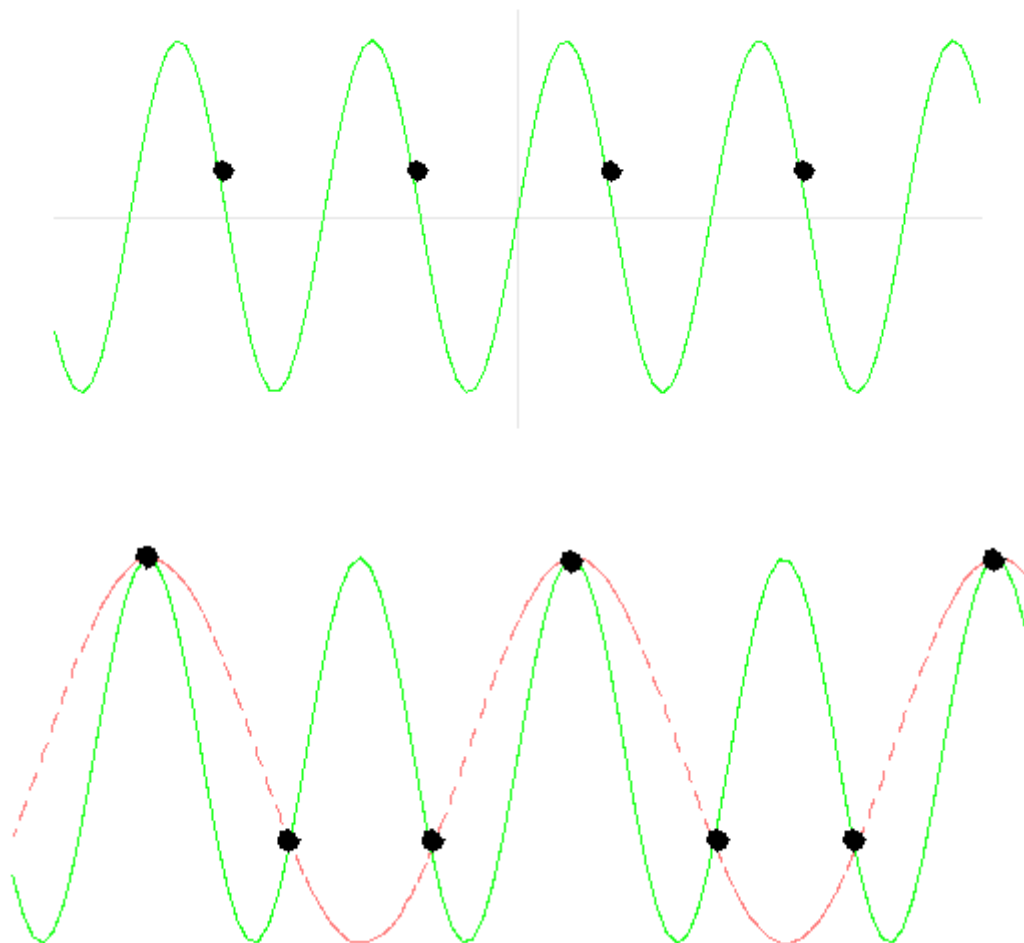
Vremenski diskretni audio signali sa kontinualnom amplitudom

Intuitivno je jasno da audio signale koji se brzo mijenjaju treba češće mjeriti. Na taj način obezbjeđujemo dovoljno informacija o signalu, odnosno osiguravamo da neka brza promjena signala ne prođe neopaženo. To znači da interval odmjerenja audio signala koji sadrže visoke tonove (brze promjene u vremenskim oblicima signala) mora biti dovoljno mali. Budući da je audio signal frekvencijski ograničen, moguće je period odabiranja izabrati tako da se na osnovu odabiraka signala može u potpunosti rekonstruisati originalni signal. Uslov koji za to mora biti ispunjen je dat poznatim *teoremom odmjerenja* (matematičko izvođenje je zasnovano na osobinama Furijeove transformacije):

$$T_0 = \frac{1}{f_0} \leq \frac{1}{2f_g}$$

gdje je T_0 period odmjerenja, f_0 frekvencija odmjerenja, a f_g granična frekvencija u spektru audio signala. Minimalna frekvencija odmjerenja f_0 se naziva Nikvistova frekvencija, a uslov koji postavlja teorema odmjerenja Nikvistov kriterij.

Iz ilustracije diskretizacije sinusoida sa sljedećih slika jasno se vidi da pri frekvenciji odmjerenja koja je manja od Nikvistove gubimo informaciju o signalu. Ako uzmemo 1 ili 1,5 odmjerača po periodu pri rekonstrukciji signala nećemo dobiti originalni signal koji smo odmjerili.



Nakon diskretizacije, neophodno je izvršiti kvantovanje audio signala. Kvantizacija je proces kojim se ulazni signal kontinualne amplitude preslikava u izlazni signal čija amplituda može da poprimi konačno mnogo različitih nivoa.

Pretpostavimo da amplituda ulaznog signala leži u opsegu $x_{\min} \leq x(t) \leq x_{\max}$, gdje $D = x_{\max} - x_{\min}$ nazivamo dinamički opseg signala, te da želimo da signal kvantujemo sa $L+1$ različitih nivoa $\hat{x}_k, k=0,1,\dots,L$. Definišimo $L+1$ nivoa ulaznog signala $x_k, k=0,1,\dots,L$ takvih da je $x_0 = x_{\min} < x_1 < x_2 < \dots < x_{L-1} < x_L = x_{\max}$.

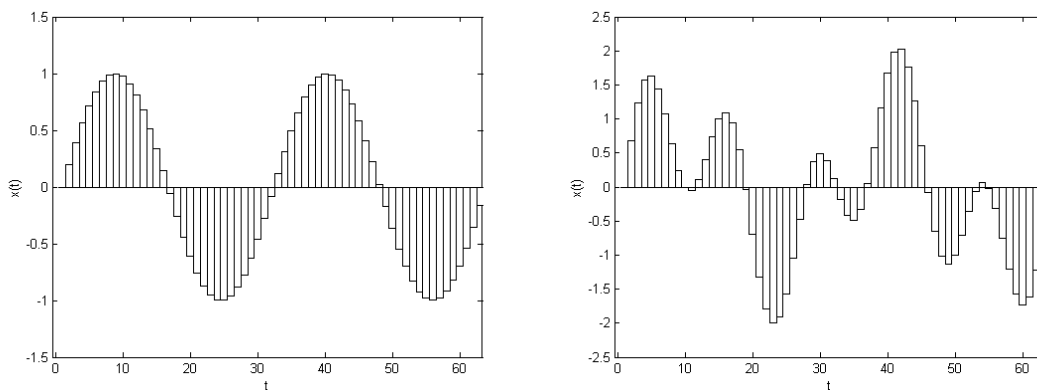
Preslikavanje ulaznog u izlazni signal se vrši na osnovu kvantizacione funkcije, na sljedeći način:

$$x(t) \in [x_k, x_{k+1}) \Rightarrow \hat{x}(t) = \hat{x}_k.$$

Zbog toga što amplituda kvantovanog signala poprima konačan broj različitih nivoa, vrijednosti signala se zapisuju (koduju) sa konačnim brojem bita B i uobičajeno je $L = 2^B$. Očigledno je da vrijednosti signala u istim trenucima vremena prije i poslije kvantizacije

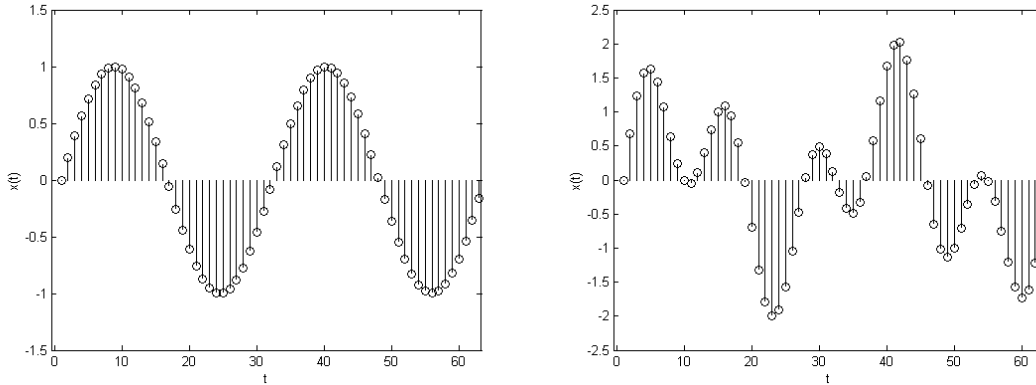
nisu iste. Greška kvantizacije koja nastupa zavisi od dinamičkog opsega i broja novoa kvantizacije. Broj bita kojim se koduje kvantovani signal je veoma važan. Pri memorisanju, procesiranju i prenosu signala poželjno je baratati sa što manjim brojem bita. Međutim, kvantovanjem signala sa malim brojem kvantizacionih nivoa gubimo informacije koje ima originalni signal. Stoga je neophodno naći kompromis između ova dva oprečna zahtjeva, tako da se odabere najmanji mogući broj bita za zapis signala, a da se pri tome sačuvaju neophodne informacije u signalu. U praksi se koristi znatno veći broj kvantizacionih nivoa, vrlo rijetko manje od $2^8=256$.

Na sljedećoj slici su dati primjeri kvantovanih audio signala. Nakon kvantizacije, vrijednosti signala odstupaju od vrijednosti originalnog analognog signala. Ovo odstupanje okarakterisano je kao kvantizacioni šum.



Kvantovani audio signali kontinualni u vremenu

Kvantizacijom diskretnog audio signala, a zatim kodovanjem kvantovanih nivoa dolazimo do digitalnog signala. **Digitalni audio signal je niz brojeva koji predstavljaju kvantovane vrijednosti analognog audio signala u diskretnim trenucima vremena.** Kodovane vrijednosti amplitude odmjeraka kvantovanog signala $\hat{x}_k = x(kT)$ se grafički predstavljaju u zavisnosti od rednog broja odmjerka k . Radi jednostavnije notacije, u daljnjem izlaganju ćemo koristiti oznaku $x(n)$ da označimo digitalni audio signal. Na sljedećoj slici su prikazani primjeri digitalnih audio signala.



Digitalni audio signali

Zadatak analogno/digitalnog (A/D) konvertora je da izvrši odmjeravanje, kvantizaciju i kodovanje analognog signala, tj. da ga prevede u digitalni oblik. Prije same A/D konverzije neophodno je ukloniti frekvencijske komponente iznad frekvencije određene Nikvistovim kriterijem. Digitalno/analogni (D/A) konvertor rekonstruiše analogni signal iz niza brojeva.

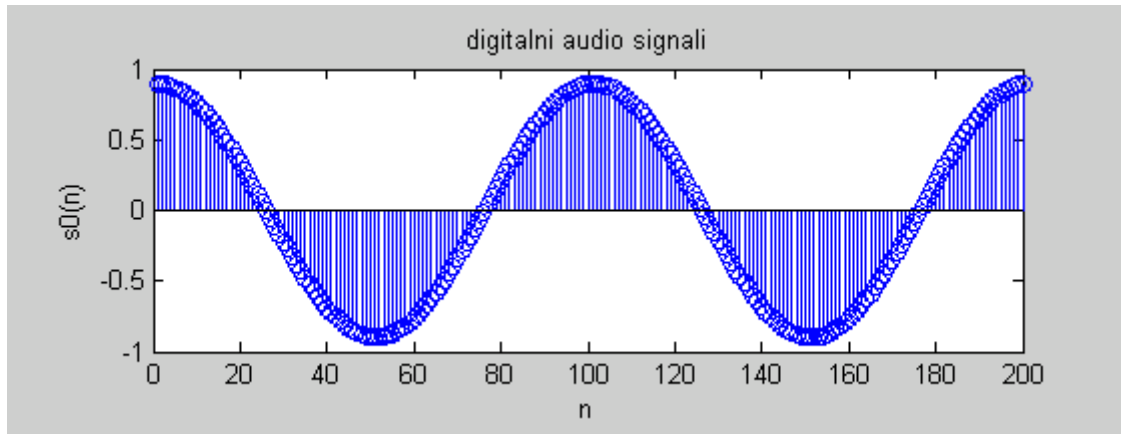
Audio signal govornog kvaliteta je moguće digitalizovati sa 8 kHz i 8 bita po odmjerku, dok je za CD-kvalitet audio signala potreban digitalni audio zapis signala digitalizovan sa 44.1 kHz i 16-bit.

Dinamički opseg koji pokriva digitalni signal je određen brojem bita koji se koristi pri zapisu odmjeraka audio signala. Svaki bit doprinosi dinamičkom opsegu sa 6 dB. Dodavanjem jednog bita udvostručava se opseg brojeva kojim predstavljamo amplitude zvučnog signala. Nivo čujnosti se ne povećava dva puta, već za $20 \log \frac{2P}{P_0} - 20 \log \frac{P}{P_0} = 20 \log 2 \approx 6 [\text{dB}]$.

Prema tome, 8 bitni audio može da pokrije dinamički opseg audio kasete (48 dB), dok je 16 bita dovoljno za dinamički opseg audio CDa (96dB), što je približno jednako dinamičkom opsegu ljudskog uha.

Primjer:

Na sljedećoj slici je prikazana digitalizovana sinusoida učestanosti 440 Hz. Audio zapis se nalazi u fajlu "440Hz.wav". Ako se amplituda sinusoide poveća 100 puta, prilikom zapisa u fajl dolazi do odsijecanja dijela signala po amplitudi. Audio zapis sinusoide od 440 Hz gdje se dešava pojava odsijecanja po amplitudi je u fajlu "odsijecanje.wav".



Osim broja bita koji su vezani sa dinamičkim opsegom, dva su još faktora koja utiču na propusnu moć i memorijske zahtjeve. To su brzina odmjerenja i izbor između mono i stereo zvuka. Npr. 8 bitni audio odmjeren sa 22 kHz zauzima 1.25 MB po minuti snimka, dok 16 bitni audio odmjeren sa 44 kHz zauzima 10 MB po minuti. Opšta formula za određivanje veličine (u bajtovima) digitalnog audio snimka je:

Frekvencija odmjerenja*trajanje snimka u sekundama*(rezolucija u bitima/8)*K, gdje je K=1 za mono, a K=2 za stereo.

Frekvencijski opsezi i brzine odmjerenja audio signala koji se koriste u praksi su:

Format	Sampling Rate	Bandwidth	Frequency Band
Telephony	8 kHz	3.2 kHz	200-3400 Hz
Teleconferencing	16 kHz	7 kHz	50-7000 Hz
Compact Disk	44.1 kHz	20 kHz	20-20,000 Hz
Digital Audio Tape	48 kHz	20 kHz	20-20,000 Hz

Primjer:

CD ROM je projektovan tako da može da reprodukuje audio signal odmjeren sa 44,1 kHz, sa 16 bita po odmjerku i 2 kanala. Za to je potrebna brzina prenosa od $44100 \cdot 16 \cdot 2 = 1411200$ bita/s = 176400 B/s = 172,2656 kB/s.

Na jedan sektor na CD ROM-u može da se upiše 2352 B podataka. Ako se koriste kodovi za korekciju greške (Raw data), što je slučaj kad se memorišu podaci, ali ne i za audio signal, u jedan sektor se efektivno upisuje 2048 B. Zato se pri određivanju brzine CD ROM-a koristi korekcionni faktor 2048/2352, te je brzina prenosa, označena sa 1x, jednaka $172,2656 \text{ kB/s} \cdot 2048/2352 = 150 \text{ kB/s}$. Brzina prenosa za Raw data iznosi 176400 B/s, odnosno 75 konsektivnih sektora u sekundi. Brzina prenosa označena sa 52x iznosi 7800 kB/s, odnosno 7,6 MB/s.

Pri zapisu audio signala bez mogućnosti korigovanja greške (RAW data), na CD sa 333000 sektora je moguće upisati $333000 \cdot 2352 \text{ B} = 764859,375 \text{ kB} = 747 \text{ MB}$, dok kapacitet pri zapisu podataka sa korištenjem kodova za korekciju greške iznosi $333000 \cdot 2048 \text{ B} = 650 \text{ MB}$. Dakle, na CD može stati $764859,375 \text{ kB} / 172,265625 \text{ kB/s} = 4440$ sekundi (74 minute) stereo zvuka.

U sljedećoj tabeli su date često korištene frekvencije odmjerenja i rezolucije, sa pripadajućim približnim veličinama datoteka.

Frekv. odmj. (kHz)	Rezolucija (bita)	Mono/Stereo	Približna brzina prenosa u minuti	Komentar
44,1	16	Stereo	10,5 MB	kvalitet kompakt diska
44,1	16	Mono	5,25 MB	dobra zamjena za visokokvalitetne snimke iz mono izvora, kao što je glas naratora
44,1	8	Stereo	5,25 MB	postiže najviši kvalitet reprodukcije na slabijim uređajima, kao što je većina zvučnih kartica za Windows računare
44,1	8	Mono	2,6 MB	odgovarajuća zamjena za snimke iz mono izvora
22,05	16	Stereo	5,25 MB	zvuči tamnije nego CD kvalitet zbog niže frekvencije uzorkovanja, ali je zvuk i dalje pun i "prisutan" zbog visoke rezolucije i sterea, preporučeno za CD ROM projekte (FM Radio)
22,05	16	Mono	2,5 MB	solidan izbor za govor
22,05	8	Stereo	2,6 MB	popularan izbor za stereo snimke razumnog kvaliteta, gdje reprodukcija u punom opsegu nije moguća
22,05	8	Mono	1,3 MB	kao kad slušate TV
11,02 5	8	Stereo	1,3 MB	na ovako niskoj frekvenciji odmjerenja skoro da nema prednosti kad se koristi stereo
11,02 5	8	Mono	650 KB	najniže što se može ići a da bude upotrebljivo
5,5	8	Stereo	650 KB	stereo ne radi
5,5	8	Mono	325 KB	kao loša telefonska veza

Primjer odmjerenja signala:

Odmjerenjem kontinualnog prostoperiodičnog signala $x(t) = \sin(\Omega_0 t)$ sa korakom odmjerenja Δt nastaje digitalni signal $x(n) = \sin(\omega_0 n)$:

$$x(t) = \sin(\Omega_0 t) = \sin(2\pi F_0 t)$$

$$x(n) = x(t) \Big|_{t=n\Delta t}$$

$$x(n) = \sin(\Omega_0 t) \Big|_{t=n\Delta t} = \sin(\Omega_0 n\Delta t) = \sin(\omega_0 n) = \sin(2\pi f_0 n)$$

Kružna učestanost Ω_0 se izražava u rad/s a frekvencija f_0 u Hz. Sa $\omega_0 = \Omega_0 \Delta t$ je označena digitalna učestanost, koja je kao i digitalna frekvencija f_0 neimenovan broj.

LINEARNA I NELINEARNA KVANTIZACIJA

Pod linearnom kvantizacijom podrazumijevamo takvo kvantovanje signala kod koga je svaka razlika amplituda koje su kodovane susjednim vrijednostima jednaka. Ovakvo uniformno kvantovanje ne uzima u obzir perceptualne karakteristike audio signala.

Postupak odmjeravanja i kodovanja signala uopšte se naziva Pulse Code Modulation (PCM), iako se često pod pojmom PCM podrazumijeva postupak kojim se signal odmjerava uniformnim korakom i tako dobijeni odmjeri kvantuju linearno. Modifikacija ove tehnike nazvana Difference Pulse Code Modulation (DPCM) koduje razlike između odmjeraka signala, umjesto same odmjerke.

Tehnika kvantovanja zasnovana na ljudskoj percepciji, jednostavno nazvana „perceptualno kodovanje“, ili μ -law (čak i u-law zbog jednostavnijeg pisanja) prvo transformiše amplitude po logaritamskoj zakonitosti, a zatim tako dobijen signal uniformno kvantuje. Na taj način se koduje više nivoa iz opsega frekvencija koje čovjek može bolje da razluči, a manje iz opsega gdje je potrebna veća razlika frekvencija kako bi čovjek registrovao promjenu visine tona.

AUDIO FORMATI

Audio fajlovi mogu sadržavati podatke o audio signalu na koje jeste ili nije primijenjena kompresija. Ranije su korišteni formati fajlova bez kompresije, što znači da su fajlovi bili mnogo veći. Uobičajeni audio formati fajlova bez kompresije su: **.au** za Unix radne stanice i **.wav** za Windows.

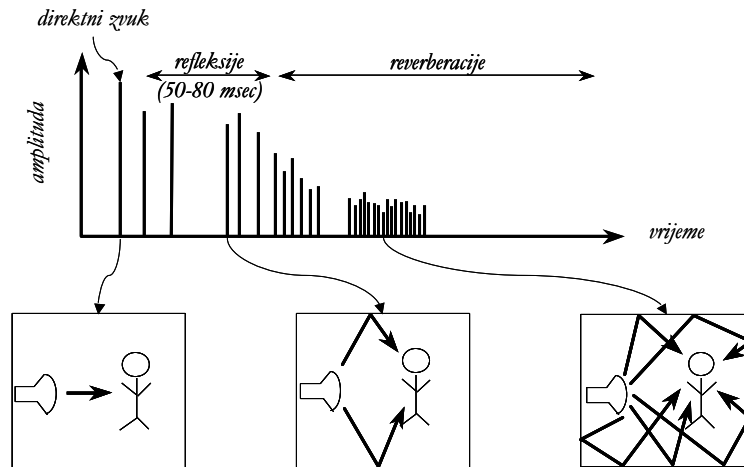
Sa kompresijom je moguće drastično smanjiti veličinu fajlova. Jedan od najjednostavnijih tipova kompresije je **DPCM** (Differential Pulse Code Modulation), gdje se koduju razlike između susjednih odmjeraka umjesto samih odmjeraka signala. **MPEG** (Motion Picture Experts Group) je tip audio i video kompresije, o kojoj će više riječi biti kasnije u toku predavanja iz ovog predmeta. **MP3** je popularni fajl format ili codec (coder/decoder) koji komprimuje fajlove na oko 10% njihove originalne veličine. MP3 je skraćenica za MPEG-1 Audio Layer 3. Što je veći broj za “audio layer”, softver koji je u njega ugrađen je sofisticiraniji. Suština ove vrste kompresije je da se iz fajla uklone sve one informacije koje nisu bitne za ljudska osjetila. Npr., svi zvuci izvan čujnog opsega se mogu eliminisati bez gubitka kvaliteta. MP3 fajlovi mogu biti kodovani sa različitim brojem bita po sekundi. Što više bita po sekundi, to je bolji kvalitet zvuka. Neophodan je kompromis između kvaliteta zvuka i veličine fajlova.

Audio format poznat kao **MOD** (skraćeno od “modules”) kombinuje aspekte MIDI i WAV fajlova. MOD fajl sadrži sekvence digitalne MIDI muzike, a uz to WAV odmjerke instrumenata koji sviraju različite note.

Streaming Audio omogućava slušanje zvučnih fajlova prije nego što se downloaduje kompletan fajl. Posebno je značajan pri radu sa velikim fajlovima, npr. kad se želi slušati cijeli koncert. Pri tome slušanje odvija istovremeno sa prenosom podataka sa web servera. Popularna aplikacija streaming audia je internet radio. Primjer streaming audio formata je Windows Media Audio (WMA).

ŠIRENJE ZVUKA U PROSTORU

Zvučni talasi koji se prostiru od izvora do prijemnika mogu da prođu različitim putevima. Ako talas na svom putu ne nailazi na prepreke govorimo o *direktnom* zvuku. *Eho* nastupa kad do slušaoca dolazi zvuk koji se odbija (prelama) od neke prepreke. Ako prelamanje proizvodi signale koji kasne manje od 100 ms govorimo o *refleksiji*, a ako je kašnjenje signala koji nastaju usljed prelamanja veće, to označavamo kao *reverberacije*. Zvuk se odbija od ravne čvrste plohe. Refleksija zvuka je bolja, ako je površina plohe glatkija i veće gustoće. Ako je ploha konveksna, dolazi do disperzije (raspršivanja) zvuka, a ako je površina konkavna, dolazi do usnopljavanja zvuka.



Difrakcija je pojava savijanja zvuka oko prepreke. Difrakcija je obrnuto prporcionalna visini tona (frekvenciji zvučnog talasa), te je više izražena na nižim frekvencijama.

Refrakcija je pojava loma zvučnog talasa, odnosno promjene smjera zvuka, a događa se uslijed promjene medija kojim se zvuk kreće (temperatura, pritisak vazduha i sl.). Tipični primjer je skretanje zvuka pod uticajem vjetra.

Apsorpcija je pojava upijanja zvuka. Prilikom udara zvučnog talasa u neku površinu jedan dio zvučne energije se predaje materiji, dok se drugi reflektuje.

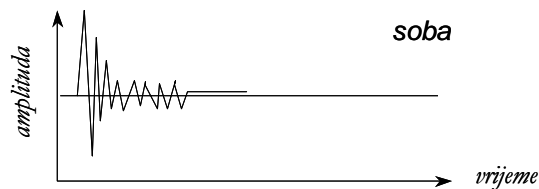
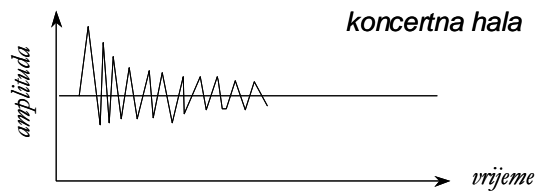
Doplerov efekt je pojava koja se očituje u promjeni visine tona zbog kretanja izvora zvuka. Ako nam se izvor zvuka približava zvučni talasi se zgušnjavaju, tj. povećava im se

frekvencija, a ako se izvor zvuka udaljava, talasi se prorjeđuju, a zvuk ima sve manju frekvenciju.

Interferencija nastaje uzajamnim djelovanjem dva talasa koji se susreću, pri čemu im se amplitude sabiraju ili oduzimaju, zavisno od faze.

Stojeći talas je pojava kad zbog intrferencije direktnog i reflektovanog talasa nastaje mjestimično poništavanje (nule) i maksimalno pojačavanje (maksimumi) zvuka. Pojava se očituje kao potpuna odsutnost zvuka u nekim dijelovima prostora. Izbjegava se tako da se zidovi prostorije ne postavljaju paralelno.

Jasno je da na kvalitet zvuka utiče ambijent u kom se nalaze izvor zvuka i slušalac. Ambijent se može posmatrati kao sistem kroz koji zvuk prolazi do slušaoca pa se, prema tome, može opisati *impulsnim odzivom*. Na sljedećim slikama su prikazani impulsni odzivi koncertne hale i obične sobe. Impulsni odziv koncertne hale je dužeg trajanja što znači da će i zvuk koji nastane da se duže čuje u koncertnoj hali nego u sobi.





NEKOLIKO PRAKTIČNIH SAVJETA

Pažljivo snimajte, vodite računa o svakom zvuku. Eliminišite pozadinske šumove. Za kvalitetan zvuk koristite gluvu sobu.

Koristite zaseban mikrofonski uređaj za svakog govornika.

Držite izbalansirane nivoe zvuka.

Obogatite zvučni zapis interesantnim zvučnim efektima.

Pažnja: ako se zvuk vezuje za animacije i ako se vrijeme trajanja muzike i animacije podesi na sporijoj mašini, može se desiti da se muzika nastavi kad se animacija završi ako se projekat izvršava na bržoj mašini (natjerati bržu mašinu da radi sporije!).