

# Uvod u digitalni audio

15. novembar 2016

Zvuk je vremenski promjenljiva mehanička deformacija koja se kroz sredinu prostire kao mehanički talas. Nastaje pri promjenama stacionarnog stanja čestica sredine, npr. pri sabijanju i širenju molekula vazduha usljed vibracija žice na gitari ili membrane zvučnika. Pri tome dolazi do oscilatornih promjena gustine sredine i vrijednosti pritiska. Oscilacije čestica sredine (akustičke oscilacije) se šire u vidu (zvučnih) talasa od mjesta gdje je došlo do poremećaja stacionarnog stanja – izvora zvuka. Zvučni talasi su mehanički longitudinalni talasi koji mogu da se prostiru kroz sredine svih agregatnih stanja. U kontekstu multimedijalnih sistema najznačajnije je prostiranje zvuka kroz vazduh.

Izvori zvuka stvaraju promjenu stacionarnog stanja sredine čime se generiše zvuk. Zvuk se može generisati na primjer vibriranjem žice instrumenta, strujanjem vazduha, vibriranjem membrane bubnja ili zvučnika, itd. Sa druge strane, prijemnici zvuka reaguju na promjene vazdušnog pritiska, kao što je to na primjer, slučaj kod ljudskog uha. Važno je napomenuti da uho nije senzor energije signala.

## 1 Objektivne i subjektivne karakteristike zvuka

Pošto su ljudi krajnji korisnici multimedijalnog sadržaja, od interesa je razmotriti ljudsku percepciju zvuka. Međutim, pošto percepcija bilo kojeg podražaja od strane čovjeka zavisi od niza faktora, kako objektivnih tako subjektivnih, moguće je razlikovati objektivne i subjektivne karakteristike zvuka. Objektivne karakteristike zvuka su fizičke osobine zvučnog talasa i iskazuju se mjerljivim fizičkim veličinama čije vrijednosti ne zavise od slušaoca. To su:

- osnovna frekvencija,
- oblik spektra zvuka i

- intenzitet zvuka.

Osnovna frekvencija zvuka je brzina oscilovanja čestica sredine kroz koju se zvuk prenosi. Prema frekvencijskom opsegu zvukove možemo podijeliti na:

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

Multimedijalni sistemi koriste zvukove iz opsega koji čovjek može da čuje. U daljnjem razmatranju ćemo zvukove iz tog opsega nazivati *audio* a odgovarajući signal *audio signal*. Osnovna frekvencija zvuka je povezana sa talasnom dužinom zvuka relacijom

$$\lambda = \frac{v}{f}, \quad (1)$$

gdje je  $f$  osnovna frekvencija zvuka, a  $v$  brzina prostiranja zvučnog talasa. Od talasne dužine zavisi potrebna veličina prostorije u kojoj se reprodukuje zvuk. Da bi se postigla vjerna reprodukcija tona sa određenom osnovnom frekvencijom, potrebno je da dimenzije prostorije budu veće ili jednake polovini njegove talasne dužine.

Na osnovu oblika spektra zvukove možemo podijeliti na tonove i šumove (buku). Za tonove su karakteristične periodične promjene vazdušnog pritiska. Pošto se za tonove mogu vezati oscilacije čestica sredine, možemo razlikovati čiste tonove kojima odgovaraju sinusoidalne oscilacije čestica sredine i složene tonove koji se sastoje iz više komponenata – osnovnog tona (koji stvarno čujemo) i niza harmonijskih gornjih tonova ili alikvotnih tonova. Složeni tonovi se mogu matematički predstaviti superpozicijom sinusoidalnih oscilacija komponenata između čijih frekvencija postoji harmonijska veza, tj. sve su umnošci frekvencije jednog, osnovnog, tona. Tonovi se koriste u muzici pa se često nazivaju i muzički tonovi.

Sa druge strane, kod šumova periodičnost promjena vazdušnog pritiska tokom vremena ne postoji, odnosno, frekvencije sinusoidalnih komponenata zvuka nisu u harmonijskoj vezi. Ipak, i ovakvi zvukovi mogu da se koriste u muzici. Na primjer, zvukovi udaraljki su šumovi. Šumovi su često neželjena pojava koja čak može da utiče i na oštećenje sluha. U ovom kontekstu često se, za ovakve zvukove, koristi termin buka.

Pošto je ljudska percepcija zvuka zasnovana na opažanju varijacija vazdušnog pritiska, mjerenje intenziteta zvuka predstavlja mjerenje zvučnog pritiska (sound pressure level). Pri mjerenju intenziteta zvuka, zvučni pritisak se izražava relativno u odnosu na referentnu vrijednost koja iznosi  $p_0 = 20 \mu\text{Pa}$ . Referentna vrijednost je izabrana tako da predstavlja približno prag čujnosti prosječnog slušaoca na 1000 Hz. Uobičajeno je da se zvučni pritisak izražava u decibelima

$$SPL = 20 \log \frac{p_{RMS}}{p_0} \text{ [dB]}, \quad (2)$$

gdje je  $p_{RMS}$  srednjekvadratna (root mean square, RMS) vrijednost zvučnog pritiska

$$p_{RMS} = \sqrt{\frac{1}{2T} \int_{-T}^T p^2(\tau) d\tau}. \quad (3)$$

Ponekad se nivo zvuka izražava nivoom intenziteta zvuka (sound intensity level) ili intenzitetom zvučnog talasa koji se definiše kao snaga zvuka (zvučnog talasa) po jedinici površine. Nivo intenziteta zvuka se takođe računa u odnosu na referentnu vrijednost koja odgovara pragu percepcije prosječnog slušaoca i izražava se u decibelima

$$SIL = 10 \log \frac{I}{I_0} \text{ [dB]}. \quad (4)$$

Za progresivne ravanske talase, referentna vrijednost nivoa intenziteta zvuka je povezana sa referentnom vrijednošću zvučnog pritiska preko karakteristične impedanse vazduha  $\rho_0 c_0 = 400 \frac{\text{kg}}{\text{sm}^2}$ . Referentna vrijednost nivoa intenziteta zvuka iznosi  $I_0 = 10^{-12} \frac{\text{W}}{\text{m}^2}$ . Referentne vrijednosti nivoa zvučnog pritiska i nivoa intenziteta zvuka su tako izabrane da vrijednosti  $SPL$  i  $SIL$  za isti zvuk budu jednake.

Nivoi zvučnog pritiska i intenziteta zvuka izražavaju se u decibelima. Korištenje decibela je pogodno u slučajevima kada je potrebno vrijednosti neke fizičke veličine, npr. snage ili intenziteta, izraziti u odnosu na neku referentnu vrijednost. Pretpostavimo da posmatrana veličina ima prirodu snage. Tada je njena vrijednost u decibelima

$$P_{dB} = 10 \log \frac{P}{P_0}, \quad (5)$$

gdje su  $P$  vrijednost veličine koja se mjeri,  $P_0$  referentna vrijednost te veličine, a  $P_{dB}$  je vrijednost veličine izražena u decibelima, relativno u odnosu na referentnu vrijednost. U ovom slučaju očigledno je da veličine  $P$  i  $P_0$  moraju biti dimenziono jednake.

Referentna vrijednost fizičke veličine odgovara vrijednosti od 0 dB. Povećanjem vrijednosti iznad referentne vrijednosti u decibelima će biti pozitivna, dok će smanjenjem vrijednosti ispod referentne vrijednosti u decibelima postati negativna. Osnovne prednosti korištenja relativne vrijednosti snage izražene u decibelima ogledaju se u smanjenju dinamičkog opsega vrijednosti mjerene veličine i transformaciji multiplikativnih veza u aditivne. Obje prednosti su posljedica korištenja logaritamske transformacije odnosa vrijednosti veličine i referentne vrijednosti.

Pošto je snaga signala proporcionalna kvadratu vrijednosti (intenziteta) signala, vrijednost u decibelima se može izračunati i kao

$$X_{dB} = 20 \log \frac{X}{X_0} \text{ [dB]}. \quad (6)$$

Kada je u pitanju intenzitet zvuka, referentne vrijednosti se biraju kao prag ljudske percepcije. Prema tome, prag čujnosti odgovara intenzitetu zvuka od 0 dB. Sa druge strane, maksimalna vrijednost zvučnog pritiska (prag bola) iznosi oko 20 Pa. Prema tome, prag bola odgovara intenzitetu zvuka od 120 dB.

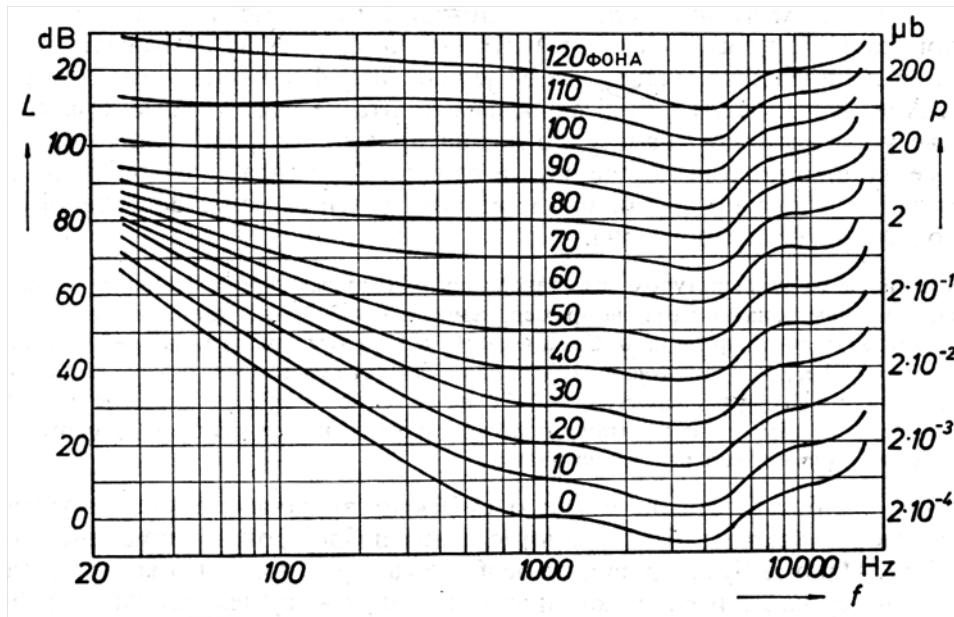
Auditorni sistem detektuje promjene vazdušnog pritiska. Slušalac ove promjene opaža kao zvuk. Percepcija zvuka sa određenim objektivnim karakteristikama je subjektivna — zavisi od slušaoca. Ipak, moguće je uočiti određene subjektivne, psihoakustičke, karakteristike zvuka. To su:

- visina tona,
- boja tona i
- glasnoća.

Opisivanje zvuka ovim karakteristikama zavisi kako od objektivnih karakteristika zvuka, tako i od uslova u kojima se nalazi slušalac, prisustva drugih zvukova, te samog slušaoca.

Visina tona (pitch) je perceptualna osobina zvuka koja omogućava uređenje muzičkih tonova na skali povezanoj sa frekvencijom. Prema visini tonovi su podijeljeni u oktave. Oktava je interval visina tonova koji odgovara udvostručenju frekvencije. Visine tonova se mogu označiti:

- slovima unutar jedne oktave — Helmholtzova notacija (C, D, E, F, G, A, H)
- slovima i brojevima — naučna notacija (slovo i broj koji označava oktavu)



Slika 1: Izofonske krive.

- brojevima koji predstavljaju frekvencije u Hz ili korištenjem MIDI standarda

$$p = 69 + 12 \log_2 \left( \frac{f}{440\text{Hz}} \right). \quad (7)$$

Druga subjektivna osobina zvuka je boja tona (timbre). Boja tona je posljedica postojanja harmonijskih gornjih tonova i u tom pogledu je povezana sa spektrom zvučnog signala. Ona omogućava razlikovanje zvukova koji potiču iz različitih izvora, a imaju jednake visine i glasnoće. Konkretno, to nam omogućava razlikovanje govornika, vrsta muzičkih instrumenata (duvački, gudački,...) itd.

Perceptualna osobina zvuka koja opisuje subjektivni osjećaj jačine zvuka je glasnoća zvuka. Glasnoća zavisi od zvučnog pritiska, frekvencije, spektralnog sadržaja, drugih tonova, slušaoca, itd. Jedinica za glasnoću je *fon*. Glasnoću od  $x$  fona ima zvuk koji na frekvenciji od 1 kHz ima nivo zvučnog pritiska od  $x$  dB. Uočeno je, međutim, da je za tonove drugačijih frekvencija potrebno varirati nivo zvučnog pritiska kako bi se dobila ista glasnoća. Fletcher i Munson su 1933. detaljno analizirali zavisnosti glasnoće od frekvencije i nivoa zvučnog pritiska i došli do familije *izofonskih krivih*, prikazanih na Slici 1 koje pokazuju intenzitete tonova različitih frekvencija koje čovjek opaža kao jednako glasne. Izofonske krive pokazuju da linearna promjena pobude (zvučnog pritiska) rezultuje logaritamskom promjenom subjektivnog

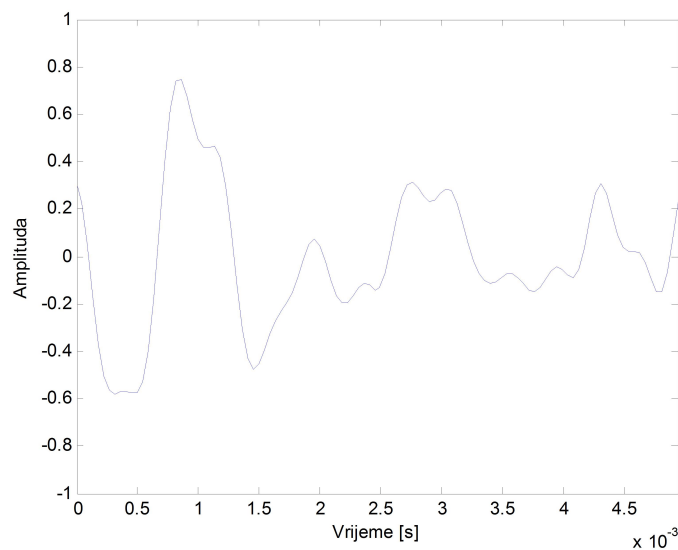
osjećaja glasnoće što je manifestacija Weber-Fechnerovog zakona prema kojem je relativna osjetljivost ljudskih čula na promjene pobude nezavisna od intenziteta pobude. Prilikom mjerenja nivoa buke u okruženju na osnovu izofonskih krivih je moguće odrediti uticaj tonova različitih frekvencija na ljudsko opažanje nivoa buke. Američko akustičko udruženje (Acoustical Society of America) razvilo je 1936. godine prvi standard za mjerače buke (American tentative standards for sound level meters for measurement of noise and other sounds Z24.3-1936) na osnovu izofonskih krivih do kojih su došli Fletcher i Munson.

Budući da se radi o mehaničkim talasima i kod prostiranja zvučnih talasa se mogu primjetiti uobičajene talasne pojave, kao što su:

- Refleksija (eho, reverberacije)
- Difrakcija (savijanje zvuka oko prepreke)
- Refrakcija (prelamanje zvučnih talasa)
- Apsorpcija (upijanje zvuka)
- Doplerov efekat (promjena visine tona zbog kretanja izvora/slušaoaca)
- Interferencija (superpozicija zvučnih talasa)
- Stojeći talas (ponišćavanje ili pojačavanje talasa)

## 2 Digitalizacija zvuka

Multimedija je danas uglavnom neodvojiva od računara. Pošto su podaci u računarima predstavljeni u digitalnom obliku, multimedijalni podaci su, takođe, u digitalnom obliku. Neki mediji (tekst, grafika, animacije) mogu nastati u digitalnom obliku, dok drugi (zvuk, slika, video) postoje u analognom obliku i, da bi se njima moglo manipulirati pomoću računara, neophodno ih je konvertovati u digitalni oblik. Kako pri ovoj konverziji ne bi došlo do gubitka informacija, potrebno je voditi računa o fizičkoj prirodi podataka. Fizički nosilac informacije je signal, npr. vrijednost vazdušnog pritiska, intenzitet svjetlosti ili električne struje. Mnogi signali sa kojima se susrećemo imaju analognu prirodu, na primjer temperatura, zvuk, intenzitet/boja svjetlosti, itd. Analogni signal je kontinualan u vremenu, odnosno, prostoru i njegove vrijednosti su kontinualne. Grafički prikaz analognog signala je dat na Slici 2. Pošto se signal matematički može predstaviti kao funkcija, onda je grafički prikaz, u stvari, grafik te funkcije. Koristi se i termin *talasni*

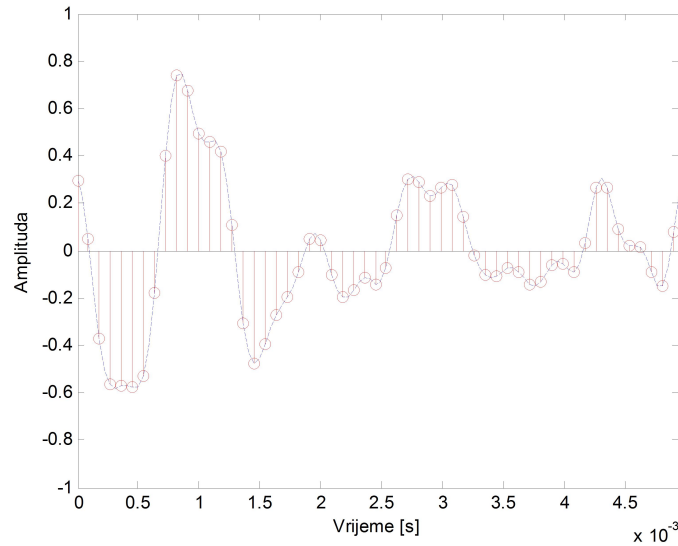


Slika 2: Analogni signal.

*oblik* signala iako nije neophodno da signal bude povezan sa nekom talasnom pojavom.

Kako bismo konvertovali analogni signal u oblik pogodan za obradu pomoću digitalnog računara potrebno je najprije pretvoriti dati signal u analogni električni signal korištenjem električnog senzora. Električni senzori su, npr. senzori temperature (termoparovi, transdjuseri), akustični senzori (mikrofoni), senzori svjetlosti (kamere), itd.

Konverzija analognih podataka u digitalni oblik se odvija u okviru procesa *digitalizacije*, odnosno, analogno-digitalne (A/D) konverzije. Digitalizacija signala obuhvata tri koraka: odmjeravanje signala, njegovu kvantizaciju i kodovanje. Odmjeravanje analognog signala je registrovanje (mjeranje) vrijednosti signala u određenim trenucima. Na ovaj način se signal predstavlja nizom njegovih numeričkih vrijednosti. Dobijeni signal se obično naziva *diskretni* signal. Diskretni signal je definisan u određenim, diskretnim, vremenskim trenucima, ali njegove vrijednosti mogu biti kontinualne. Primjer diskretnog signala prikazan je na Slici 3. Vrijednosti diskretnog signala nazivaju se *odmjeraci* signala. Zbog jednostavnije realizacije uobičajeno je da se signal odmjerava u pravilnim intervalima. Razmak između sukcesivnih odmjeraka je *period odmjeravanja*. Broj odmjeraka u jednoj sekundi je *brzina odmjeravanja* ili *frekvencija odmjeravanja* i izražava se u hercima. Rezultujući broj odmjeraka će zavisiti od brzine odmjeravanja i trajanja signala. Brzina odmjeravanja utiče na potreban memorijski prostor za čuvanje od-



Slika 3: Diskretni signal.

mjeraka signala kao i na potrebnu brzinu komunikacione veze kojom će se digitalni signal prenositi.

Odmjeravanjem analognog signala  $x_a(t)$  sa periodom odmjerenja  $\Delta t$  se dobija diskretni signal čiji su odmjerci dati sa

$$x(n) = x_a(t) |_{t=n\Delta t} . \quad (8)$$

Neka je dat kontinualni (analogni) sinusni signal  $x_a(t) = \cos(2\pi F_0 t)$  koji odmjeravamo sa frekvencijom odmjerenja  $F_s = \frac{1}{\Delta t}$ . Odmjerci rezultujućeg signala (diskretne sinusoidne) su dati sa

$$\begin{aligned} x(n) &= x_a(t) |_{t=n\Delta t} = \\ &= \cos(2\pi F_0 n\Delta t) = \\ &= \cos\left(2\pi \frac{F_0}{F_s} n\right) = \\ &= \cos(2\pi f_0 n) = \\ &= \cos(\omega_0 n) . \end{aligned}$$

Veličina  $f_0 = \frac{F_0}{F_s}$  se naziva digitalna frekvencija, a  $\omega_0 = 2\pi \frac{F_0}{F_s}$  digitalna kružna učestanost iako se ovi termini uglavnom ne koriste pretjerano strogo. Frekvencija  $F_0$  se izražava u Hercima (Hz), kružna učestanost  $\Omega_0 = 2\pi F_0$  se izražava u radijanima u sekundi (rad/s), a digitalna frekvencija i digitalna



kružna učestanost su neimenovani brojevi. Ipak, digitalna frekvencija predstavlja broj ciklusa koje signal kompletira za jedan odmjerak, a digitalna kružna učestanost promjenu faze signala u radijanima za jedan odmjerak pa se često uz njihovu vrijednost dodaju i oznake ovih veličina.

Sa povećanjem frekvencije  $F_0$  analognog signala povećava se broj ciklusa koje signal kompletira u nekom vremenskom intervalu, oscilacije signala su brže. Međutim, kod diskretnog signala broj ciklusa koje signal kompletira će se povećavati samo dok je kružna učestanost u intervalu  $0 \leq \omega_0 \leq \pi$ . Razlog za ovo je činjenica da je nemoguće razlikovati dvije diskretne sinusoidne čije se digitalne kružne učestanosti razlikuju za  $2k\pi, k \in Z$

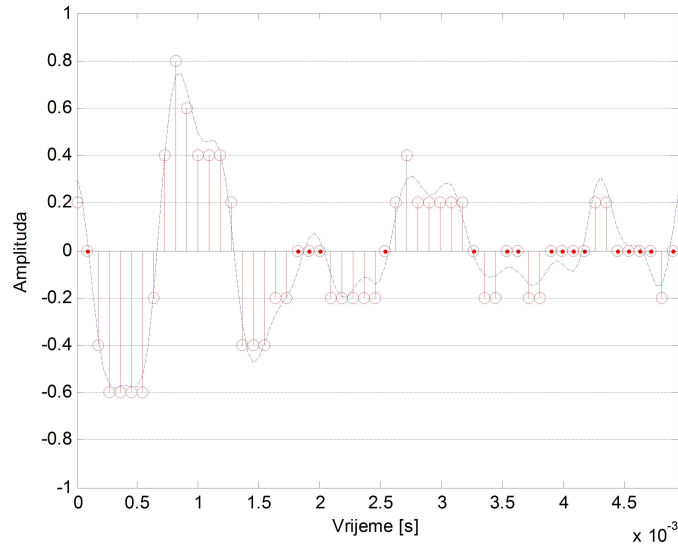
$$\begin{aligned} \cos((\omega_0 + 2k\pi)n) &= \cos(\omega_0 n) \cos(2kn\pi) - \sin(\omega_0 n) \sin(2kn\pi) = \\ &= \cos(\omega_0 n). \end{aligned}$$

Posljedica ove osobine diskretne sinusoidne je da je nemoguće razlikovati diskretizovane varijante sinusoida čije se frekvencije razlikuju za  $kF_s, k \in Z$

$$\begin{aligned} \cos(2\pi(F_0 + kF_s)n\Delta t) &= \cos(2\pi F_0 n\Delta t) \cos(2\pi kF_s n\Delta t) - \\ &\quad - \sin(2\pi F_0 n\Delta t) \sin(2\pi kF_s n\Delta t) = \\ &= \cos\left(2\pi \frac{F_0}{F_s} n\right) \cos(2kn\pi) - \\ &\quad - \sin\left(2\pi \frac{F_0}{F_s} n\right) \sin(2kn\pi) = \\ &= \cos(\omega_0 n), \end{aligned}$$

pri čemu je iskorišćeno  $F_s = \frac{1}{\Delta t}$ . Sinusoidne čije se frekvencije razlikuju za  $kF_s$  nazivaju se *aliasi*, a ova pojava *aliasing* ili *preklapanje spektra*.

Budući da je diskretni signal niz brojeva otvara se mogućnost za njegovo memorisanje i obradu korištenjem digitalnog računara. Međutim, pošto je za reprezentaciju vrijednosti odmjerka na raspolaganju konačan broj bita, nije moguće predstaviti proizvoljnu vrijednost iz skupa realnih brojeva. Zbog ovoga je neophodno diskretizovati i vrijednosti koje signal može poprimiti, odnosno, kvantizovati signal. Signal koji je diskretan i u vremenu i po vrijednosti je *digitalni* signal. Primjer digitalnog signala prikazan je na Slici 4. Kvantovani signal može poprimiti samo vrijednosti iz konačnog skupa diskretnih vrijednosti. Vrijednosti koje kvantovani signal može poprimiti su nivoi kvantizacije. Razmak između nivoa kvantizacije je korak kvantizacije. Vrijednosti signala (čak i nekvantovanog) su obično iz ograničenog skupa (određenog fizičkim zakonima) pa se može uspostaviti veza između broja nivoa kvantizacije i koraka kvantizacije. Broj nivoa kvantizacije zavisi od broja bita koji su na raspolaganju za kodovanje jednog odmjerka. Ako je na



Slika 4: Digitalni signal.

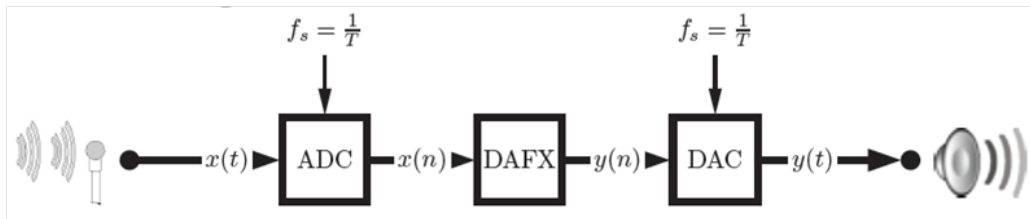
raspolaganju  $n$  bita moguće je kodovati  $2^n$  različitih vrijednosti kvantovanog signala. Na primjer, sa 8 bita se može predstaviti  $2^8 = 256$  kvantizacionih nivoa. Korak kvantizacije može biti fiksna (uniformna kvantizacija) ili promjenljiv (neuniformna kvantizacija).

Nakon kvantizacije signala potrebno je numeričke vrijednosti odmjeraka konvertovati u odgovarajuće binarne reprezentacije, odnosno, dodijeliti svakom kvantizacionom nivou kodnu riječ. Kako bi se ovo postiglo, potrebno je izabrati format u kojem će podaci biti kodovani. Ovdje se donosi odluka o korištenju koda sa fiksnom ili promjenljivom dužinom. Ovaj korak je važan, zato što su podaci u računaru predstavljeni nizovima binarnih brojeva, a smisao podacima u računarskim sistemima daju aplikacije.

Prilikom digitalizacije signala potrebno je, dakle, odgovoriti na tri pitanja:

1. Kolika je brzina odmjeravanja?
2. Koliko mora biti fina kvantizacija i da li je kvantizacija uniformna?
3. U kom formatu će biti kodovani podaci (kakav je format fajla)?

Reprodukcija digitalnog sadržaja u mnogim slučajevima zahtijeva da se digitalni signal pretvori u analogni. Na primjer, reprodukcija muzike zahtijeva da se od digitalnog ponovo dobije analogni električni signal koji se u zvučnicima pretvara u varijacije vazdušnog pritiska što čovjek opaža kao



Slika 5: Blok dijagram sistema za digitalnu obradu analognog signala.

zvuk, dok reprodukcija slike/video koristi ekran monitora da bi se generisala svjetlost promjenljivog intenziteta koja je osnova za ljudsku percepciju slike. Konverzija digitalnog signala u analogni poznata je pod nazivom digitalno/analogna (D/A) konverzija.

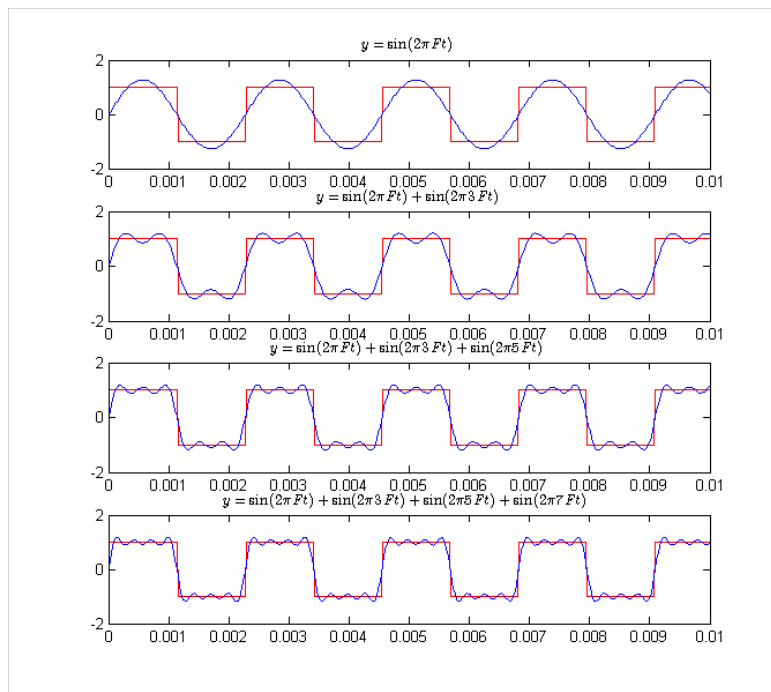
Imajući prethodno izlaganje u vidu, sistem za digitalni prenos i obradu analognog signala se može predstaviti blok dijagramom na Slici 5. Blok označen sa ADC je analogno/digitalni konvertor, na njegov ulaz se dovodi analogni električni signal, kao što je npr. signal iz mikrofona, a na njegovom izlazu se dobija digitalni signal. Sama obrada signala odvija se u bloku označenom sa DAFX. Ovim blokom je realizovan algoritam kojim se ulazni digitalni signal transformiše u izlazni. Konačno, blok označen sa DAC je digitalno/analogni konvertor na čijem se izlazu dobija analogni električni signal kojim se npr. pobuđuje zvučnik.

Do sada, međutim, nismo razmatrali pitanje mogućnosti rekonstrukcije analognog signala na osnovu raspoložive digitalne reprezentacije, kao ni pitanje kvaliteta te rekonstrukcije. U ovom cilju ponovo ćemo posebno razmotriti uticaj koraka odmjeravanja i kvantizacije signala.

## 2.1 Teorema odmjeravanja

Intuitivno je jasno da audio signale koji se brzo mijenjaju treba češće odmjeravati. Na taj način obezbjeđujemo dovoljno informacija o signalu, odnosno osiguravamo da neka brza promjena signala ne prođe neopaženo. To znači da interval odmjeravanja audio signala koji sadrže visoke tonove (brze promjene u vremenskim oblicima signala) mora biti dovoljno mali. Postavlja se pitanje da li je moguće odrediti minimalnu brzinu odmjeravanja koja će osigurati vjernu diskretnu reprezentaciju analognog signala. Pri ovome se pod vjernom reprezentacijom smatra ona reprezentacija na osnovu koje je moguće rekonstruisati originalni analogni signal.

Kako bi se izveo uslov koji mora da zadovolji frekvencija odmjeravanja kako bi se omogućila idealna rekonstrukcija odmjerenog signala, potrebno je posmatrati signal sa stanovišta Furijeove analize. Grubo posmatrano,

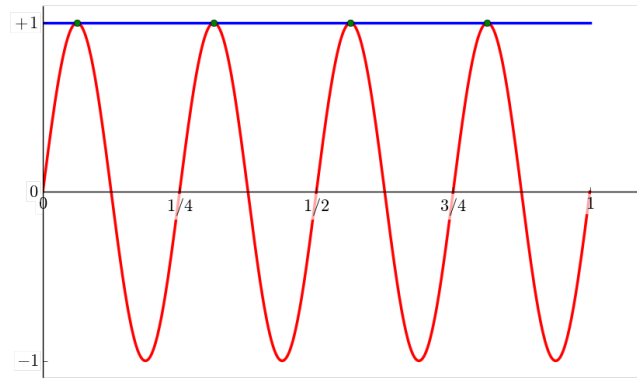


Slika 6: Aproksimacije pravougaonog signala Furijeovim redom.

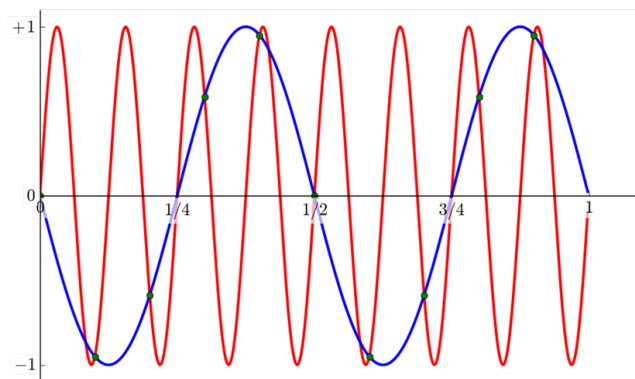
Furijeova analiza omogućava predstavljanje signala težinskom sumom sinusoida različitih frekvencija i početnih faza. Ukoliko su frekvencije korištenih sinusoida umnošci jedne, osnovne, frekvencije dobijamo periodične signale koji odgovaraju muzičkim tonovima. Komponente signala se u ovom slučaju nazivaju *harmonici*, Slika 6. Sa druge strane, ako dozvolimo da sinusoidi budu proizvoljnih frekvencija možemo formirati reprezentacije šire klase signala koji uključuju i aperiodične signale, npr. šumove. Skup frekvencija i početnih faza sinusoida na koje je razložen signal naziva se *spektar* signala.

Dakle, da bismo razmotrili pitanje određivanja frekvencije odmjerenja, dovoljno je da razmotrimo pitanje odmjerenja sinusoida. Na Slici 7 prikazan je slučaj kada se sinusoida odmjera brzinom odmjerenja koja je jednaka frekvenciji sinusoida. Očigledno, na osnovu ove reprezentacije nije moguće rekonstruisati polaznu sinusoidu, već samo konstantan signal. Na Slici 2.1 prikazani su odmjerci sinusoida frekvencije 8 Hz dobijeni frekvencijom odmjerenja od 10 Hz. Vidimo da se na osnovu ovih odmjeraka može rekonstruisati i sinusoida frekvencije 2 Hz. Pri frekvenciji odmjerenja od 10 Hz sinusoida frekvencije 2 Hz je *alias* sinusoida frekvencije 8 Hz, odnosno, na osnovu odmjeraka nije moguće razlikovati ove dvije sinusoida.

Može se dokazati da je za idealnu rekonstrukciju originalnog signala na



Slika 7: Odmjeravanje sinusoide brzinom odmjjeravanja koja je jednaka frekvenciji sinusoide.



osnovu njegovih odmjeraka potrebno da frekvencija odmjeravanja  $F_S$  bude bar dva puta veća od najviše frekvencije u spektru signala. Ova teorema je poznata pod imenom *teorema odmjeravanja*, a kriterijum

$$F_S \geq F_g, \quad (9)$$

gdje je  $F_g$  najviša frekvencija u spektru signala, se naziva Nikvistov kriterijum, prema Hariju Nikvistu koji je formulisao teoremu odmjeravanja u navedenom obliku. Teoremu su nezavisno dokazali Klod Šenon i Vladimir Koteljnikov.

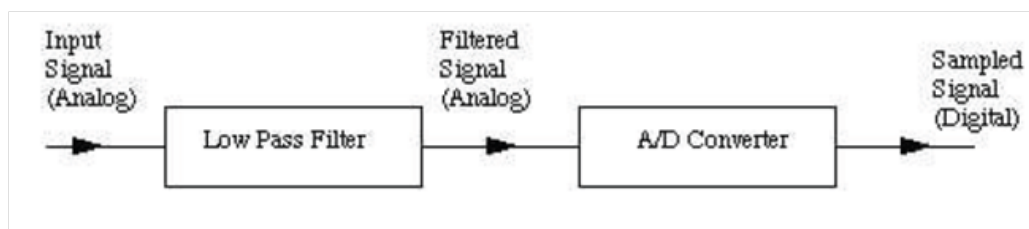
Ukoliko signal koji se odmjerava sadrži komponente čije su frekvencije veće od polovine frekvencije odmjeravanja (Nikvistova frekvencija<sup>1</sup>) dolazi do pojave koja se naziva *preklapanje spektra* ili *aliasing*. U tom slučaju će se, nakon rekonstrukcije, u signalu pojaviti komponente na frekvencijama koje su aliasi njihovih originalnih frekvencija. Ovo je posljedica ranije pokazane činjenice da sinusoide čije se frekvencije razlikuju za  $kF_s$  imaju istu diskretnu reprezentaciju, odnosno, rezultuju istim skupom odmjeraka. Originalna frekvencija signala i njen alias su simetrični u odnosu na Nikvistovu frekvenciju, pa se Nikvistova frekvencija naziva i *frekvencija presavijanja*.

Većina signala iz stvarnog svijeta nema ograničen spektar tako da uvijek prijete opasnost od preklapanja spektra. Zbog toga se prije A/D konvertora stavlja niskopropusni filter (antialiasing filter), kao na Slici 8. Zadatak ovog filtra je da oslabi harmonike na frekvencijama višim od polovine frekvencije odmjeravanja. Gubitak informacija zbog ovog filtriranja je relativno mali, pod uslovom da je frekvencija odmjeravanja dobro izabrana, zbog toga što energija realnih signala tipično opada sa porastom frekvencije. Pored toga, često postoje i ograničenja ljudske percepcije visokofrekventnih signala. Npr. ljudi čuju zvuk do približno 20 kHz tako da je frekvencija odmjeravanja od 44,1 kHz (CD kvalitet) sasvim dovoljna da bi se obezbijedila mogućnost rekonstrukcije audio signala. Postojanje predfiltra će rezultovati odsustvom visokih tonova u situacijama kada se koriste niže frekvencije odmjeravanja. Niskopropusno filtriranje je potrebno i kod promjene brzine odmjeravanja signala. Takođe, prilikom D/A konverzije signala, niskopropusni filter sa istom graničnom frekvencijom se koristi neposredno nakon D/A konvertora. Razlog za ovo je potreba da se uklone visokofrekventne komponente koje se javljaju zbog aproksimacije signala.

Dakle, da bi se odabrala frekvencija odmjeravanja audio signala potrebno je uzeti u obzir frekvencijski opseg signala kao i raspoloživu memoriju za

---

<sup>1</sup>Terminologija je ovdje zbunjujuća: brzina odmjeravanja koja je jednaka dvostrukoj najvišoj frekvenciji u spektru signala naziva se *Nikvistova brzina*, a za fiksnu frekvenciju odmjeravanja, polovina frekvencije odmjeravanja se naziva *Nikvistova frekvencija*.



Slika 8: Niskopropusni predfiltar (antialiasing filtar).

Tabela 1: Pregled često korištenih frekvencija odmjeravanja.

Telefoni, bežični mikrofoni	8 kHz
G.722 VoIP	16 kHz
CD kvalitet, MP3	44,1 kHz
Profesionalna audio oprema (npr. DAT)	48 kHz
DVD audio, Blu-ray	do 192 kHz

odmjerke, odnosno, brzinu komunikacione veze. Imajući u vidu da je maksimalna frekvencija ljudskog govora oko 4 kHz, a granica ljudskog sluha oko 20 kHz, dolazimo do zaključka da će se, frekvencija odmjeravanja audio signala, kretati u opsegu od 8 kHz naviše. U Tabeli 1 dat je pregled često korištenih frekvencija odmjeravanja u različitim primjenama. U industriji se sve češće sreću visoke frekvencije odmjeravanja od, npr. 96 kHz ili 192 kHz. Međutim, laboratorijski eksperimenti nisu pokazali prednost po pitanju kvaliteta u odnosu na brzine odmjeravanja od 44,1 kHz ili 48 kHz. Pošto čovjek ne može da čuje ultrazvuk, instrumenti i ljudski glas ne generišu zvukove čije su frekvencije iznad 20 kHz, a frekventijski opseg mikrofona takođe ne prelazi 20 kHz, sa perceptualnog stanovišta ne postoji razlog za korištenje vrlo visokih frekvencija odmjeravanja. Štaviše, previsoka frekvencija odmjeravanja, kao što je npr. 192 kHz može da rezultuje intermodulacionom distorzijom, odnosno pojavom signala unutar čujnog opsega čije frekvencije su zbir i razlika frekvencija signala iz ultrazvučnog opsega. Druga mana korištenja frekvencije odmjeravanja od 192 kHz je vrlo velika količina podataka koja se dobija što poskupljuje sistem i usporava prenos podataka kroz mrežu.

Validan razlog za korištenje više frekvencije odmjeravanja je lakše projektovanje filtara u A/D i D/A konvertorima. Međutim, ovaj problem se danas rješava tako što konvertori interno koriste višu frekvenciju odmjeravanja, a na izlazu daju decimiran signal.

## 2.2 Uticaj kvantizacije na količinu podataka i kvalitet signala

Neka amplituda signala može da poprimi vrijednosti iz opsega  $[x_{min}, x_{max}]$ . Dinamički opseg signala je  $x_{max} - x_{min}$ . Podijelimo ovaj interval na  $L$  kvantizacionih nivoa  $x_0 = x_{min} < x_1 < x_2 < \dots < x_{L-1} = x_{max}$ . Vrijednost

$$\Delta = \frac{x_{max} - x_{min}}{L - 1}, \quad (10)$$

je korak kvantizacije ili rezolucija. Na primjer, ako je  $x_{min} = 0, x_{max} = 1$  i kvantujemo signal sa  $L = 11$  nivoa, korak kvantizacije je  $\Delta = 0,1$ . Kvantizacijom se odmjerku  $x(n)$  pridružuje kvantovana vrijednost

$$x_q(n) = Q[x(n)], \quad (11)$$

koja je jednaka najbližem kvantizacionom nivou.

Broj kvantizacionih nivoa povezan je sa brojem bita  $B$  potrebnih za kodovanje odmjerača i vrijedi  $L = 2^B$ . Ako je korak kvantizacije veći (grublja kvantizacija) signal se kvantuje sa manje nivoa pa je potrebno manje bitova za kodovanje odmjerača što rezultuje manjom količinom podataka. Analogno, manji korak kvantizacije (finija kvantizacija) zahtijeva veći broj bita za kodovanje odmjerača. Za digitalni audio signal uobičajena je kvantizacija sa 8 ili 16 bita po odmjerku, što rezultuje sa 256, odnosno, 65536 kvantizacionih nivoa, respektivno.

Kvantizacija je ireverzibilan proces, nakon izvršene kvantizacije više nije moguće rekonstruisati originalnu vrijednost odmjerača. Dakle, kvantizacijom se unosi greška u reprezentaciju signala. Iako se, u ovom slučaju, ne radi o šumu uobičajeno je korištenje naziva *šum kvantizacije* ili *greška kvantizacije*. Analogija sa šumom potiče od toga što je vrijednost greške kvantizacije u suštini slučajna veličina. Šum kvantizacije je razlika između kvantovane i originalne vrijednosti odmjerača

$$e_q(n) = x_q(n) - x(n). \quad (12)$$

Očigledno, grublja kvantizacija povlači izraženiji šum kvantizacije čime je kvalitet signala lošiji, dok je za smanjenje šuma kvantizacije potrebno koristiti finiju kvantizaciju, odnosno, alocirati više bita po odmjerku.

Uticaj šuma kvantizacije na kvalitet audio signala može se ocjenjivati testovima na slušaocima što je komplikovano i dugotrajno. Zbog toga je uvedena objektivna mjera kvaliteta audio signala – odnos signal-šum

$$SNR = 10 \log \frac{P_x}{P_n} [\text{dB}], \quad (13)$$



gdje je  $P_x$  srednja snaga signala, a  $P_n$  srednja snaga, odnosno, varijansa šuma. Kao što je naznačeno jednačinom, odnos signal-šum se izražava u decibelima.

Pošto se greška kvantizacije može posmatrati kao šum kvantizacije, kvalitet kvantovanog audio signala se može izraziti pomoću odnosa signal-šum kvantizacije (Signal-to-Quantization-Noise Ratio, SQNR)

$$SQNR = 10 \log \frac{P_x}{P_q}, \quad (14)$$

gdje je  $P_q$  srednja snaga šuma kvantizacije.

Pretpostavimo da je srednja vrijednost signala koji se kvantuje jednaka nuli i da je opseg njegovih vrijednosti  $[-x_{max}, x_{max}]$ , odnosno da je  $x_{min} = -x_{max}$ . Ako se za kvantovanje signala  $x(n)$  koristi B bita, opseg vrijednosti digitalnog signala je  $-2^{B-1}$  do  $2^{B-1} - 1$ . Korak kvantizacije je

$$\Delta = \frac{2x_{max}}{2^B} = \frac{x_{max}}{2^{B-1}}. \quad (15)$$

Šum kvantizacije  $e_q(n)$  može poprimiti vrijednosti iz intervala  $[-\frac{\Delta}{2}, \frac{\Delta}{2}]$ .

Odnos signal-šum kvantizacije se može izraziti kao odnos vršne vrijednosti signala koja iznosi približno  $2^{B-1}\Delta$  i maksimalne vrijednosti šuma kvantizacije<sup>2</sup>  $\Delta/2$

$$SQNR = 20 \log \frac{2^{B-1}\Delta}{\frac{\Delta}{2}} \quad (16)$$

$$= 20 \log 2^B \quad (17)$$

$$= 20B \log 2 \quad (18)$$

$$= 6,02B[\text{dB}]. \quad (19)$$

Dakle, dodavanjem jednog bita rezolucija se povećava za oko 6 dB. Ovo znači da se sa 16 bita postiže maksimalni SQNR od 96 dB.

U prethodnoj analizi je pretpostavljen najgori slučaj. Ako se, pak, pretpostavi da je signal sinusoidalan, da je greška kvantizacije statistički nezavisna od signala i da je njena raspodjela uniformna na intervalu  $[0, \frac{\Delta}{2}]$  dobija se da je odnos signal-šum kvantizacije

$$SQNR = 1,76 + 6,02B[\text{dB}]. \quad (20)$$

Iako je prethodni rezultat izveden za sinusoidalne signale, može se pokazati se da sličan rezultat dobija za svaki signal čiji je dinamički opseg jednak dinamičkom opsegu kvantizatora.

---

<sup>2</sup>Ovo je u stvari Peak Signal-to-Quantization-Noise Ratio (PSNR).

Tabela 2: Bitske brzine za neke audio primjene.

Kvalitet	Brzina odmjeraivanja (kHz)	Biti po odmjerku	Mono/stereo	Brzina (kB/s)	Opseg frekvencija
Telefon	8	8	mono	8	200-3400
AM radio	11,025	8	mono	11,0	100-5500
FM radio	22,050	16	stereo	88,2	20-11000
CD	44,1	16	stereo	176,4	5-20000
DAT	48	16	stereo	192,0	5-20000
DVD audio	192(max)	24(max)	do 6 kanala	1200,0(max)	0-96000

Sada je moguće sumarizovati memorijske zahtjeve, odnosno, zahtjeve za brzinu prenosa audio signala. Memoriija potrebna za  $T$  sekundi mono audio signala odmjeraivanog sa frekvencijom odmjeraivanja  $F_S$  i kodovanog sa  $B$  bita je  $TF_S B/8$  bajtova, a potrebna brzina prenosa je  $BF_S$  bit/s. Stereo signal zahtijeva 2 puta više memorije, odnosno 2 puta veću brzinu prenosa. Na primjer 1 minut audio signala CD kvaliteta zahtijeva

$$60 \times 44100 \times 16 \times 2/8/1024/1024 = 10,1 \text{ MB}$$

i bitsku brzinu od<sup>3</sup>

$$44100 \times 16 \times 2/8/1000 = 176,4 \text{ kB/s}$$

Potrebne bitske brzine za neke karakteristične audio primjene date su u Tabeli 2. Navedene vrijednosti važe ako se ne koristi kompresija audio signala.

Ako se odmjerci audio signala čuvaju kao uniformno kvantovane vrijednosti, kao što je slučaj kod zapisa na CD, DVD, Blu-ray ili WAV formata audio fajlova, kažemo da se koristi linearni format. Međutim, u nekim primjenama kao što je npr. telefonija, koristi se neuniformna kvantizacija tako što se različit broj bita alocira u skladu sa osobinama ljudske percepcije zvuka. Opisani postupak konverzije analognog signala u digitalni i njegovog kodovanja kao niza bitova naziva se impulsna kodna modulacija (Pulse Code Modulation – PCM). Kvantizacija kod impulsne kodne modulacije je

<sup>3</sup>Za bitsku brzinu se koristi SI prefiks k = 1000.

neuniformna, dok varijanta pod nazivom linearna impulsna kodna modulacija (LPCM) koristi uniformnu kvantizaciju. Međutim, termin PCM se često koristi za linearnu impulsnu kodnu modulaciju.