**DOG01**

1. Opiši proces nastajanja govora i osnovne elemente govora

Govor nastaje kao poruka u mozgu u apstraktnom obliku. Zatim se pretvara u skup neuronskih signala koji upravljaju postupkom artikulacije. Tim postupkom se formira zvučni signal koji sadrži informaciju originalne poruke.  
Osnovni elementi govora su fonemi, tj. glasovi.

2. Nabroji i kratko opiši glavna područja primjene digitalne obrade govora.

Kodiranje govora u svrhu učinkovitijeg prijenosa i pohrane, sinteza govora, sustavi za transformaciju govornog signala, sustavi za dijalog između računala i čovjeka, sustavi za pomoć osobama s oštećenjem vida, sluha ili govora, prepoznavanje govora, govornika ili jezika, poboljšanje kvalitete govora, analiza govora...

3. Što je kodiranje govora i koja mu je namjena?

Kodiranje govora je postupak pretvorbe analognog govornog signala u digitalni oblik, koji je u današnje vrijeme mnogo prikladniji kako za pohranu, tako i za prijenos. Nakon te pretvorbe, nad digitalnim signalom moguće je primijeniti bilo kakav postupak za obradu, ali kodiranje se ne bavi problemom obrade, već čim učinkovitijim postupkom digitalne reprezentacije govora.

4. Opiši podjelu sustava za kodiranje obzirom na način reprezentacije govornog signala te diskutiraj potrebnu brzinu prijenosa za pojedine tipove.

Postoje reprezentacija valnog oblika govornog signala, te parametarska reprezentacija govornog signala. Također, parametarska se reprezentacija dijeli još i na reprezentaciju preko parametara pobude, te parametara vokalnog trakta.  
Za reprezentaciju valnog oblika, potrebne su brzine od 20000 bit/s, pa naviše, dok je brzina potrebna za parametarsku reprezentaciju puno manja, i iznosi oko 500-10000 bit/s za metodu analize/sinteze, i oko 75 bit/s za sintezu teksta u govor.

5. Navedi koji su sustavi kodiranja govora temeljeni na parametarskoj reprezentaciji prvi korišteni u digitalnoj mobilnoj telefoniji.

1. sjeverno-američki standard IS54 VSELP  
2. japanski standard JPD-VSELP  
3. europski standard GSM RPE-LTP koder

6. Na koji način se ostvaruje ušteda u brzini prijenosa kod kodera s parametarskom reprezentacijom?

Govorni signal u sebi sadrži popriličnu količinu reduntantne informacije. Ako se „bitna“ informacija razdvoji od „nebitne“, te ako se „nebitni“ dio opiše modelom, a „bitni“ dio kvantizira, kodira i prenese na prijemnu stranu, moguće je ostvariti istu kvalitetu reprodukcije, uz manju brzinu prijenosa informacije.

7. Koji su glavni trendovi razvoja i nove smjernice kod modernih sustava za kodiranje govora za primjene u mobilnim komunikacijama?

*Kod modernih sustava za kodiranje govora u mobilnim komunikacijama, pokušava se uz očuvanje iste kvalitete smanjiti potrebna brzina prijenosa ili poboljšati kvaliteta bez promjene potrebne brzine. To se pokušava ostvariti proširenjem frekvencijskog pojasa, kojih je trenutnih 200Hz-3.2kHz dovoljno za prijenos, ali kvaliteta je takvim uskim pojasom značajno narušena. Novim koderima tako se pojas želi proširiti na 50Hz-7kHz, a zovu se „Wide-band“ koderi.* (?)  
Bitan aspekt kod primjene govornih tehnologija je složenost algoritma, tj. problem vezan uz potrebnu procesnu moć procesora na kojem će se obrađivati. To je i direktno povezano s problemom potrošnje električne energije. Zbog toga, velika se pažnja posvećuje projektiranju učinkovitih algoritama kodiranja govora, koji osiguravaju pogodan kompromis između sažimanja i složenosti, radi smanjenja potrošnje električne energije.

8. Na koji način se ostvaruje sustav za kodiranje govora u tipičnom mobilnom telefonu?

Izvedba algoritma kodiranja govora u mobilnim telefonima je programska. Svaki mobilni telefon ima procesor - DSP (u novije vrijeme i više njih) koji izvodi algoritam kodiranja i dekodiranja u realnom vremenu. Njegova zadaća je da uzorke govornog signala koji se snima mikrofonom obradi i pretvori u digitalnu komprimiranu poruku koja se odašilje putem radio veze. Istovremeno, mora dolaznu poruku dekodirati i pretvoriti u uzorke koji se zatim pretvaraju u analogni oblik i šalju na slušalicu.

9. Nabroji glavne dijelove anatomije vokalnog trakta te objasni način kako pojedini dijelovi utječu na nastajanje govora. (?)

Prostor između glasnica – glottis, pharynx ili ždrijelo, usna šupljina, jezik, nepce, nadzubno meso, zubi, usne, velum ili resica, nosna šupljina i nosnice.  
Vokalni trakt se ponaša kao filtar koji spektralno oboji pobudni signal iz pluća, a na formiranje glasa utječu položaji jezika, usana, čeljusti, resice, frekvencija titranja glasnica, otvorenost resice, itd.

10. Navedi temeljnu klasifikaciju fonema hrvatskog jezika.

Na osnovu konfiguracije i otvora vokalnog trakta, glasovi se dijele na:  
otvorni glasovi – samoglasnici – vokali  
poluotvorni glasovi – glasnici – sonanti  
zatvorni glasovi – suglasnici – konsonanti

11. Kako se dijele samoglasnici u hrvatskom jeziku obzirom na širinu otvora i mjesto tvorbe?

Širina otvora:  
najotvoreniji – a  
srednji – e o  
najzatvoreniji – i u

Mjesto tvorbe:  
prednji – i e  
srednji – a  
stražnji – o u

12. Nabroji sonante u hrvatskom jeziku. Da li glasnice titraju prilikom izgovora sonanata.

Sonanti: j, l, lj, m, n, nj, r, v  
Glasnice titraju, i zato su oni zvučni glasovi.

13. Opiši podjelu konsonanata u hrvatskom jeziku obzirom na zvučnost i tip, te ilustriraj primjerima zvučno-bezvučnih parova.

|  |  |  |  |
| --- | --- | --- | --- |
| ZVUČNI | b d g | dž đ - | v z ž - |
| BEZVUČNI | p t k | č ć c | f s š h |
| tip glasa | ZATVORNI ili PRASKAVI ili EKSPLOZIVNI | SLOŽENI ili ZATVORNO-TJESNAČNI ili AFRIKATI | TJESNAČNI ili STRUJNI ili FRIKATIVNI |

**DOG02**

1. Objasni što utječe na nastajanje govora, vezano uz anatomiju vokalnog trakta i fizikalne procese u njemu. (?)

Utječu usnice i glasnice, te pomicanje artikulatora (tj. promjena volumena vokalnog trakta): promjena volumena usne šupljine uslijed kretanja čeljusti, jezika, usnica i resice.

2. Zbog čega je jezik jedan od najvažnijih artikulatora u nastajanju govora. Kako on utječe na mjesto tvorbe glasa (npr. za slučaj samoglasnika)?

Promjenom oblika i položaja jezika u usnoj šupljini mijenja se oblik i volumen same šupljine, a time i glas koji proizvodimo. Jezik je najznačajniji artikulator jer pri izgovoru definira dinamiku govora. Npr. na mjesto tvorbe samoglasnika utječe pomičući se naprijed-nazad:  
naprijed – e i  
sredina – a  
iza – o u

3. Koje metode se koriste za statičko i dinamičko oslikavanje anatomije vokalnog trakta i pozicije artikulatora?

Metode za statičko i dinamičko oslikavanje anatomije vokalnog trakta i pozicije artikulatora su presjek vokalnog trakta pomoću MRI, te Röntgensko skeniranje. Postoji i metoda koja se zasniva na akustičkim mjerenjima: temelji se na uzbudi vokalnog trakta pomoću vanjskog izvora.

4. Na koji način resica utječe na nastajanje glasova. Za koju klasu glasova je to od posebnog značaja?

Resica utječe na nastajanje glasova tako što može otvarati, tj. zatvarati vezu nazalnog i vokalnog trakta. Tako propušta (ili blokira) zračnu struju u nazalni otvor, a od posebnog je značaja za klasu nazalnih glasova.

5. Objasni proces nastajanja zvučnog govora na glasnicama. Na koji način svjesno utječemo na frekvenciju titranja glasnica, odnosno na činjenicu da li glasnice uopće titraju ili ne?

Glasnice predstavljaju prepreku toku zraka iz pluća prema vokalnom traktu. Zbog pritiska u plućima, zrak iz pluća prolazi kroz otvor glasnica. Prema Bernoullijevom zakonu, tlak na mjestu prepreke je niži nego lijevo i desno od nje (zato jer je protok veći). Uz pravilnu napetost glasnica, ovaj sniženi tlak će privući i spojiti glasnice i potpuno zatvoriti tok zraka. Kad tlak iza zatvorenih glasnica ponovno dovoljno naraste, glasnice se ponovno razmiču i propuštaju zrak prema vokalnom traktu. To se ciklički ponavlja i tako glasnice titraju, a mi možemo utjecati na titranje odabirom glasa, te promjenom pritiska struje iz pluća.

**DOG03**

1. Što je cilindrični model vokalnog trakta? Po čemu se on razlikuje u odnosu na stvarni oblik vokalnog trakta za pojedini glas?

To je simplificirani trodimenzionalni model vokalnog trakta koji je temeljen na pretpostavci da se vokalni trakt u određenom trenutku vremena može opisati pomoću niza međusobno spojenih cijevi bez gubitaka. Takva aproksimacija razlikuje se u odnosu na stvarni oblik vokalnog trakta time što zanemaruje gubitke uslijed trenja, toplinske vodljivosti i vibracija stijenki cijevi. Također se razlikuje u tome što ne uzima u obzir nosnu šupljinu pomoću koje se neki glasovi formiraju.

2. Koji glasovi se mogu vrlo dobro aproksimirati cilindričnim modelom, a za koje glasove takav model ipak nije dostatan?

Zbog njihove stacionarnosti u dužem vremenskom periodu, samoglasnici se mogu vrlo dobro aproksimirati cilindričnim modelom, dok on nije dostatan za nazalne glasove.

3. Navedi koji su glavni faktori pored samog oblika vokalnog trakta koji utječu na nastajanje govora.

Zračenje na usnicama, akustička veza nosne i usne šupljine, gubitci u vokalnom traktu (toplinska vodljivost, trenje), promjene oblika vokalnog trakta uslijed mekoće stijenki, te pobudni signal vokalnog trakta.

4. Nabroji osnovne zakone koji određuju fizikalne osnove nastanka i širenja zvuka u vokalnom traktu. Koji od ovih zakona su dovoljni za opis procesa ako pretpostavljamo sustav bez gubitaka?

Zakon očuvanja mase, zakon očuvanja količine gibanja, zakon očuvanja energije, zakoni termodinamike, zakoni mehanike fluida.  
Za sustav bez gubitaka u fluidu i na stijenkama (uslijed toplinske vodljivosti) dovoljni su zakoni očuvanja mase, momenta i energije.

5. Napiši parcijalne diferencijalne jednadžbe koje opisuju fizikalni proces širenja zvuka u vokalnom traktu. Koji oblik poprimaju ove jednadžbe ako je funkcija poprečnog presjeka vokalnog trakta vremenski nepromjenjiva?

Ako je funkcija poprečnog presjeka vokalnog trakta vremenski nepromjenjiva:

6. Za cilindrični model s vremenski stalnim i uniformnim presjekom A(x,t)=A, napiši pretpostavljeni oblik rješenja za protok i tlak na mjestu x u trenutku t, te potvrdi da pretpostavljena rješenja zadovoljavaju parcijalne diferencijalne jednadžbe. (?)

DOKAZ:

7. Opiši kako se fizički izvodi sustav za harmonijsku pobudu uniformne cijevi koji se ponaša kao idealni izvor protoka volumena zraka na početku cijevi. Jednadžbama pokaži kako fizičko pomicanje ovog aktuatora formira protok. (?)

Takav sustav se fizički izvodi pomoću cijevi koju pobuđuje klip čiji se otklon duž x osi mijenja po sinusnom zakonu:  
  
Takav klip je idealan izvor, i on se pod utjecajem sile pomiče lijevo-desno i tako mijenja brzinu protoka volumena zraka.

8. Opiši što je idealno zaključenje uniformne cijevi, te koji je električki ekvivalent takvog idealnog zaključenja.

Budući da je cijev otvorena na jednom kraju, sve što zrači iz nje, zrači u beskonačan prostor. To ne odgovara stvarnom sustavu jer konačna brzina protoka na izlazu cijevi ne može uzrokovati promjenu tlaka ovog beskonačnog prostora. To daje p(l,t)=0, što bi značilo da bi zvuk nestao na usnama.  
Električki ekvivalent takvog idealnog zaključenja jer kratki spoj, jer tlak na izlazu odgovara naponu na kraju linije koji je jednak nuli za bilo koju struju, a kod realnih modela, tlak i brzina protoka na kraju cijevi bit će vezani zaključnom impedancijom ZL različitom od 0, koja modelira proces zračenja na usnicama.

9. Za idealnu homogenu cijev bez gubitaka pobuđenu s harmonijskom pobudom protoka volumena zraka na početku cijevi , frekvencije Ω, napiši izraze za pretpostavljeno rješenje za pozitivno i negativno putujući val.

Gdje su K+ i K- kompleksne konstante koje određuju amplitudu i početnu fazu svakog vala.

10. Izvedi izraze za kompleksne koeficijente K+ i K- koji određuju amplitude pozitivno i negativno putujućeg vala za idealnu homogenu cijev pobuđenu s harmonijskom pobudom i idealnim zaključenjem na kraju cijevi.

Rubni uvjeti:

Za x = 0

=>

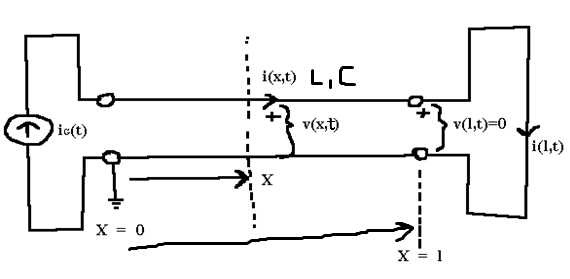
11. Napiši konačne izraze koji predstavljaju rješenje za protok i tlak u idealnoj homogenoj cijevi presjeka A i dužine L na proizvoljnom mjestu x u trenutku t, ako je pobuđena s harmonijskom pobudom i idealnim zaključenjem na kraju cijevi. Odredi i izraz za protok na kraju cijevi (x=L).

Protok na kraju cijevi (x=l):

12. Što je karakteristična akustička impedancija idealne homogene cijevi i o čemu ovisi. Napiši izraz. U kojim izrazima se javlja ova impedancija?

Karakteristična akustička impedancija je omjer zvučnog tlaka i brzine čestice u cijevi u kojoj se rasprostire zvučni val, u uvjetima slobodnog rasprostiranja vala, tj. kad nema reflektiranih valova. Ovisi o gustoći medija u kojem se nalazi, o brzini zvuka u tom mediju, te površini poprečnog presjeka cijevi.

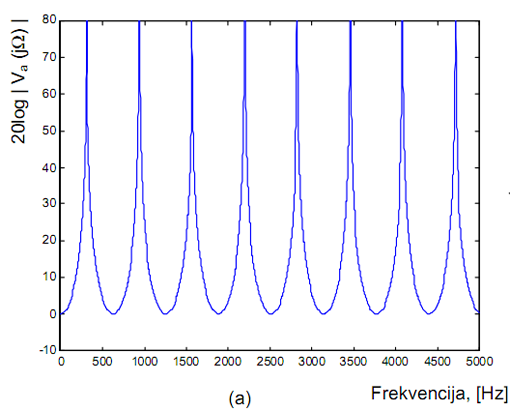
13. Za idealnu homogenu cijev, skiciraj ekvivalentni električki model, te tablično prikaži ekvivalentne varijable i veličine. Usporedno prikaži i sustav parcijalnih diferencijalnih jednadžbi za oba modela (akustički i električki).



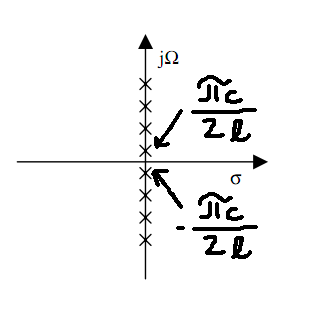
EKVIVALENTI:

|  |  |
| --- | --- |
| AKUSTIČKA VELIČINA | ANALOGNA ELEKTRIČKA VELIČINA |
| p – tlak | v – napon |
| u – brzina protoka volumena zraka | i – struja |
| - akustički induktivitet | L – induktivitet po metru |
| A/ c2 – akustički kapacitet | C – kapacitet po metru |

14. Napiši izraz za frekvencijsku karakteristiku protoka idealne homogene cijevi, te skiciraj oblik amplitudno-frekvencijske karakteristike za slučaj cijevi dužine 17.5cm i brzinu širenja zvuka 350m/s.

 - šiljci su na 500, 1500, 2500...

15. Napiši izraz za prijenosnu funkciju idealne homogene cijevi, te odredi pozicije polova ove prijenosne funkcije. Skiciraj pozicije polova u s-ravnini, te diskutiraj stabilnost.

Beskonačno mnogo polova:  
  
Polovi su na j osi => sustav je na rubu stabilnosti.

**DOG04**

1. Zbog čega općenit cilindrični model vokalnog trakta s kontinuirano promjenjivim poprečnim presjekom A(x) zamjenjujemo s modelom sa spojenim cijevima fiksnih presjeka? Skiciraj takav model.

Mijenjamo ga zato što cilindrični modeli ne vode do analitičkih rješenja, dok ovakav model to omogućuje.  


2. Skiciraj grananje pozitivno i negativno putujućih valova brzine protoka na mjestu spoja k-te i (k+1)-ve cijevi, te označi sve valove na početku i kraju obje cijevi.

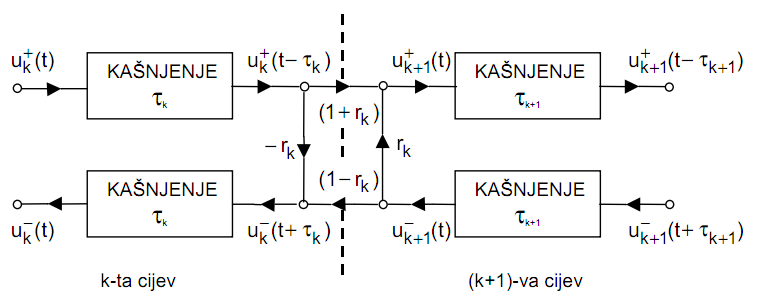


3. Postavi jednadžbe koje definiraju kontinuitet protoka i tlaka na spoju k-te i (k+1)-ve cijevi, te u ove jednadžbe uvrsti izraze za tlak i protok u te dvije cijevi. Pretpostaviti da se obje cijevi ponašaju kao obično kašnjenje (tj. nema gubitaka).

Na spoju k-te i k+1-ve cijevi:

Tlak i protok u svakoj cijevi:  
  
Analogno i za k+1...  
Uvrštavamo u prve dvije jednadžbe :

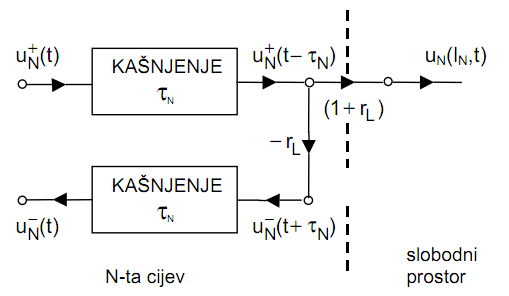
4. Napiši konačne izraze za odlazne valove iz spoja dvije susjedne cijevi izražene kao funkciju dolaznih valova u ovaj spoj i faktora refleksije rk. Napiši izraz za faktor refleksije. Prikaži i grafički tokove signala na ovom spoju.



5. Grafički skiciraj ekvivalente modele spoja dvije cijevi koji koriste dva ili samo jedan množač.

|  |  |
| --- | --- |
| Dva množača: | Jedan množač: |
|  |  |

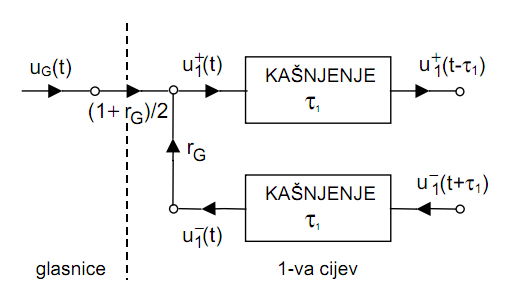
6. Prikazati model (dijagram toka signala i odgovarajuće jednadžbe) za zaključenje zadnje cijevi modela na mjestu usnica. Čime je određen faktor refleksije rL? Koliko iznosi rL za slučaj idealnog zaključenja?



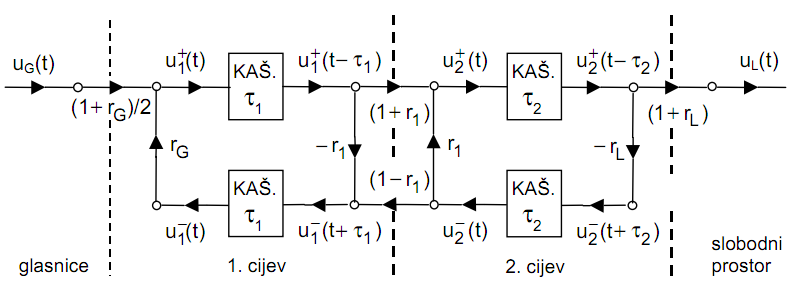
Vrijedi:

Brzina protoka:  
  
Za idealno zaključenje rL = 1.

7. Prikazati model (dijagram toka signala i odgovarajuće jednadžbe) za zaključenje na početku prve cijevi modela na mjestu glsanica. Čime je određen faktor refleksije rG?Koliko iznosi rG za slučaj idealnog zaključenja?



8. Prikaži dijagram toka signala za model vokalnog trakta s dvije spojene cijevi i zaključenjima na glasnicama i na usnicama.



9. Napiši izraz za prijenosnu funkciju protoka modela vokalnog trakta s dvije spojene cijevi sa zaključenjima na glasnicama i usnicama. Diskutiraj oblik ove prijenosne funkcije, kao i broj polova.

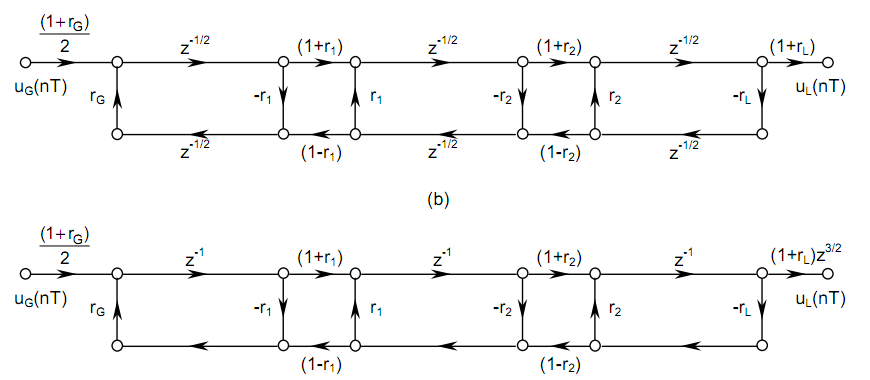
(1+rg) i (1+rl) su frekvencijske karakteristike

10. Što je potrebno učiniti s modelom vokalnog trakta sa spojenim cijevima kako bi njegovu analizu mogli provesti u vremenski diskretnoj domeni? Kako se odabire period otipkavanja takvog modela?

Potrebno je prijenosnu funkciju prebaciti u vremenski diskretnu domenu metodom jednakih impulsnih odziva:  
  
s periodom otipkavanja:

11. Izvedi prijenosnu funkciju modela sa spojenim cijevima istih dužina u vremenski diskretnoj domeni Va(z), koja se dobiva iz vremenski kontinuirane prijenosne funkcije Va(s) pomoću metode jednakih impulsnih odziva, gdje je period otipkavanja jednak dvostrukom vremenu prolaza signala kroz jedan segment modela.

12. Skiciraj dijagram toka signala vremenski diskretnog modela vokalnog trakta s dvije spojene cijevi istih dužina (obje verzije, ... s kašnjenjima u obje grane ili samo u gornjoj grani).



13. Napiši izraz pomoću kojeg je moguće odrediti prijenosnu funkciju modela sa spojenim cijevima istih dužina u vremenski diskretnoj domeni za proizvoljni broj cijevi N, te ilustriraj primjenu ovog izraza za model s dvije cijevi.

[1-rG] – glasnice  
[1 –rLz-1] – zaključenje na usnicama

N=2:

14. Diskutiraj odnos broja cijevi modela (N), dužine vokalnog trakta (L) i frekvencije otipkavanja (fs), reda prijenosne funkcije modela, broja i tipa korijena, odnosno maksimalnog broja rezonantnih karakteristika tog modela.

Broj cijevi - određuje red sustava (N cijevi: D(z) N-tog reda)  
Dužina vokalnog trakta L - određuje dužinu jedne cijevi   
Kašnjenje kroz jednu cijev -   
Frekvencija otipkavanja -

Sustav N-tog reda V(z) ima N polova, ali su polovi ili realni ili konjugirano kompleksni parovi kojih ima najviše N/2, te svaki par opisuje jednu rezonantnu karakteristiku (u intervalu od 0 do fs/2, njih najviše N/2). Nema nula brojnika (nego konstanta pomaknuta za N/2 koraka). Sustav je all-pole sustav.

15. Kojim parametrima je opisan jedan formant vremenski diskretnog modela. Ilustriraj kako je ove parametre moguće odrediti za model s dvije spojene cijevi i faktorima refleksije rL, r1i rG uz pretpostavku da model ima jedan konjugirano kompleksni par polova.

Parametri su:  
- jedan par konjugirano-kompleksnih polova  
- frekvencija otipkavanja

Određivanje parametara za N=2:

Frekvencija pojavljivanja formanata je 1 formant/kHz.  
U području [0 fs/2] govorni signal ima formanata, a svaki je određen parom konjugirano-kompleksnih polova.

16. Obrazloži zašto se red modela mora prilagođavati dužini voklanog trakta. Kako ćemo odrediti potrebni red modela ako poznamo dužinu voklanog trakta, frekvenciju otipkavanja i brzinu širenja zvuka?

Red modela se mora prilagođavati dužini vokalnog trakta jer je uz potreban broj cijevi u modelu N, s frekvencijom otipkavanja dan izrazom:

Što je veća frekvencija otipkavanja ili duži vokalni trakt, mora biti veći broj spojenih cijevi modela N.

**DOG05**

1. Što je identifikacija sustava? U kojim područjima se javlja potreba za identifikacijom sustava? Kako se dijele modeli sustava obzirom na razinu a-priori znanja o njima?

Identifikacija sustava je određivanje modela sustava na osnovu njegovog ponašanja, tj. na osnovu pobude i odziva.  
Potreba za tim se javlja u područjima automatske regulacije i upravljanja, u statistici i teoriji informacija, u komunikacijskim sustavima, te u postupcima obrade signala.  
Podjela modela sustava s obzirom na razinu a-priori znanja o njima:  
- grey box model – struktura poznata, nepoznati koeficijenti  
- black box model – sva znanja nepoznata

2. Napiši podjelu modela sustava obzirom na oblik prijenosne funkcije (brojnik i nazivnik i njihov red). Koji od tih modela se mogu odrediti bez direktnog poznavanja pobudnog signala?

Podjela s obzirom na oblik prijenosne funkcije:  
- modeli koji imaju samo brojnik: digitalni filtar s konačnim impulsnim odzivom (FIR) -> Moving Average model  
- modeli koji imaju samo nazivnik: rekurzivni filtar bez nultočaka (IIR, all-pole) -> Auto Regressive  
- modeli koji imaju i brojnik i nazivnik: općeniti digitalni IIR filtar, ima netrivijalne nule i polove -> Auto Regressive Moving Average (za određivanje koeficijenata moramo poznavati i pobudu i odziv)

Bez direktnog poznavanja pobude mogu se odrediti AR modeli jer za takve modele mogu se uvesti pretpostavke na statistička i spektralna svojstva pobude.

3. Nabroji poznatije formulacije postupka linearne predikcije. Koje su dvije temeljne formulacije detaljno obrađene u okviru predmeta?

Formulacije postupka linearne predikcije: metoda kovarijance, metoda autokorelacije, formulacija korištenjem mrežaste strukture, formulacija pomoću inverznog filtra, formulacija estimacije spektra, formulacija maksimalne vjerodostojnosti, formulacija skalarnog ili unutarnjeg produkta.  
Metoda kovarijance i metoda autokorelacije su obrađene u okviru predmeta.

4. Definiraj linearni prediktor. Napiši njegovu jednadžbu diferencija. Kako se računa pogreška predikcije. Kako se naziva filtar koji od signala dobiva pogrešku predikcije i u kojem odnosu je njegova prijenosna funkcija s prijenosnom funkcijom prediktora?

Linearni prediktor je sustav za izračunavanje linearne kombinacije prethodnih uzoraka s(n-1), s(n-2)... s(n-p) i služi za određivanje parametara govornog signala.

Jednadžba diferencija linearnog prediktora, gdje je p stupanj (red), a koeficijenti linearnog prediktora.

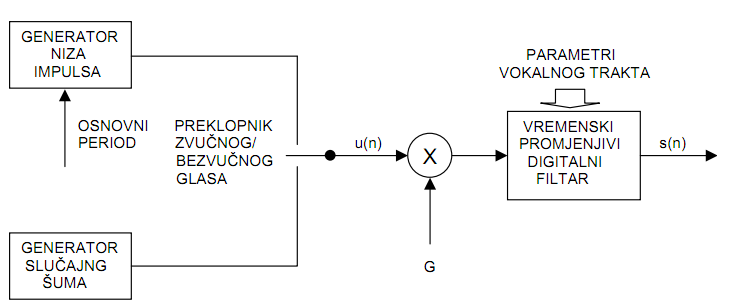
Pogreška predikcije računa se kao razlika stvarne vrijednosti n-tog uzorka i njegove predikcije:

Filtar koji od signala dobiva pogrešku predikcije zove inverzni filtar sustava. Njegova prijenosna funkcija je inverz prijenosne funkcije prediktora.

5. Koja funkcija cilja se koristi kod izračunavanja optimalnih koeficijenata linearnog prediktora? Koje su prednosti korištenja baš takve funkcije cilja?

Kao funkcija cilja kod izračunavanja optimalnih koeficijenata koristi se suma kvadrata predikcijske pogreške. Cilj dobre predikcije je odabrati koeficijente prediktora tako da pogreška predikcije svih uzoraka na intervalu bude što manja.  
Prednosti su: predstavlja kompromisno rješenje koje je „podjednako dobro“, ili „podjednako loše“ za sve uzorke tog intervala.

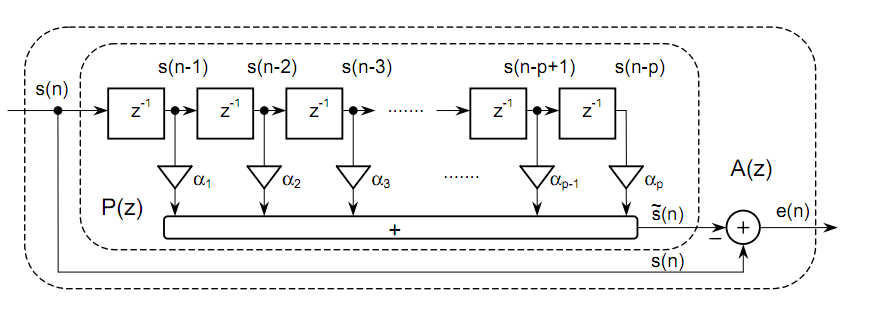
6. Skiciraj digitalni model nastajanja govornog signala. Što se modelira filtrom H(z)? Kojeg je tipa taj filtar? Napiši jednadžbu diferencija za ovaj digitalni model. Što ona povezuje?



Filtrom H(z) modelira se utjecaj oblika glotalnog pulsa, prijenosne funkcije vokalnog trakta i zračenja na usnicama.  
Filtar H(z) je IIR filtar – beskonačni impulsni odziv bez nultočaka:

Jednadžba difrerencija za H(z): povezuje uzorke govora s(n) i pobude u(n)

7. Skiciraj blok dijagram linearnog prediktora reda p. Napiši jednadžbu diferencija za signal pogreške predikcije. Diskutiraj odnos ove jednadžbe diferencija sa sličnom jednadžbom za digitalni model nastajanja govora.



Jednadžba diferencija za signal pogreške predikcije:

Usporedba

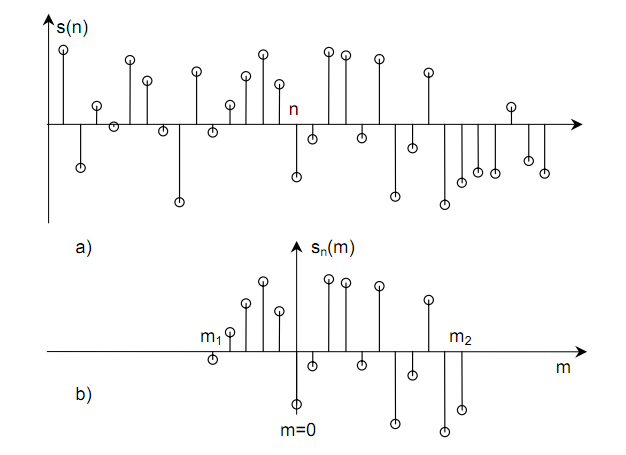
Uz pretpostavku da je govorni signal stvarno izlaz modela H(z) i pretpostavku da su koeficijenti linearnog prediktora jednaki koeficijentima modela ak: pogreška predikcije e(n) bit će identički jednaka pobudnom signalu modela Gu(n).

8. Objasni zašto možemo pretpostaviti da će postupak linearne predikcije dati koeficijente koji se mogu iskoristiti kao koeficijenti digitalnog modela nastajanja govora H(z). Za koji tip pobudnog signala bi ova pretpostavka bila savršeno zadovoljena?

Želimo odrediti skup koeficijenata prediktora iz uzoraka otipkanog govornog signala tako da se postigne dobro poklapanje kratkotrajnih spektralnih svojstava govornog signala i pretpostavljenog modela u čijem nazivniku figuriraju koeficijenti , a kako govorni signal ima vremenski promjenjiva svojstva (kvazi-stacionaran signal), koeficijenti prediktora moraju biti određeni iz kratkotrajnog segmenta signala unutar kojeg su spektralna svojstva signala „dovoljno“ stalna. Zadatak je pronaći koeficijente prediktora koji će minimizirati srednju kvadratnu pogrešku predikcije na promatranom segmentu signala. Rezultirajući optimalni parametri bit će prihvaćeni (odabrani) kao koeficijenti ak polinoma u nazivniku nepoznatog modela H(z).  
Za pobudne signale s ravnim spektrom (beskonačni idealni bijeli šum i jedinični impuls) ta je pretpostavka zadovoljena.

9. Zbog čega se analiza govora linearnom predikcijom mora provoditi na segmentima signala konačnog trajanja? Opiši i grafički ilustriraj postupak definiranja izdvojenog segmenta uvođenjem pomoćnog indeksa m.

Mora se provoditi na segmentima konačnog trajanja jer govorni signal ima vremenski promjenjiva svojstva, a unutar kratkog segmenta su spektralna svojstva „dovoljno“ stalna.  
sn(m) – dio signala s(n) u okolini indeksa m, gdje je okolina određena rasponom pomoćnog indeksa m koji se kreće od m1 do m2:  
sn(m) = s(m+n)



a) signal s(n)  
 b) izdvojeni segment sn(m) u okolini indeksa n

10. Izvedi konačno rješenje za optimalne koeficijente linearnog prediktora reda p koji minimiziraju ukupnu kvadratnu predikcijsku pogrešku na izdvojenom segmentu govornog signala definiranom proizvoljnim rasponom pomoćnog indeksa m. Koji oblik rješenja se dobiva ovim izvodom?

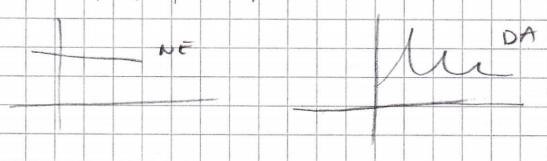
Dobiva se sustav od p linearnih jednadžbi po nepoznanicama do oblika:

11. Napravi izvod rješenja za optimalni linearni prediktor drugog reda. Izračunaj analitički rješenja za koeficijente alfa1 i alfa2.

12. Napiši matrični oblik rješenja za optimalne koeficijente linearnog prediktora reda p uz općenit raspon pomoćnog indeksa m koji izdvaja odabrani segment signala. Napiši izraze kojim se izračunavaju elementi ovog matričnog sustava.

=

13. Kako se može odrediti sumarna kvadratna pogreška predikcije linearnog prediktora reda p, na osnovu izračunatih optimalnih koeficijenata alfa i odabranih elemenata matrice sustava jednadžbi? Objasni za koji tip signala se ne može ostvariti predikcijski dobitak i zašto ne (po čemu to odmah vidimo)?

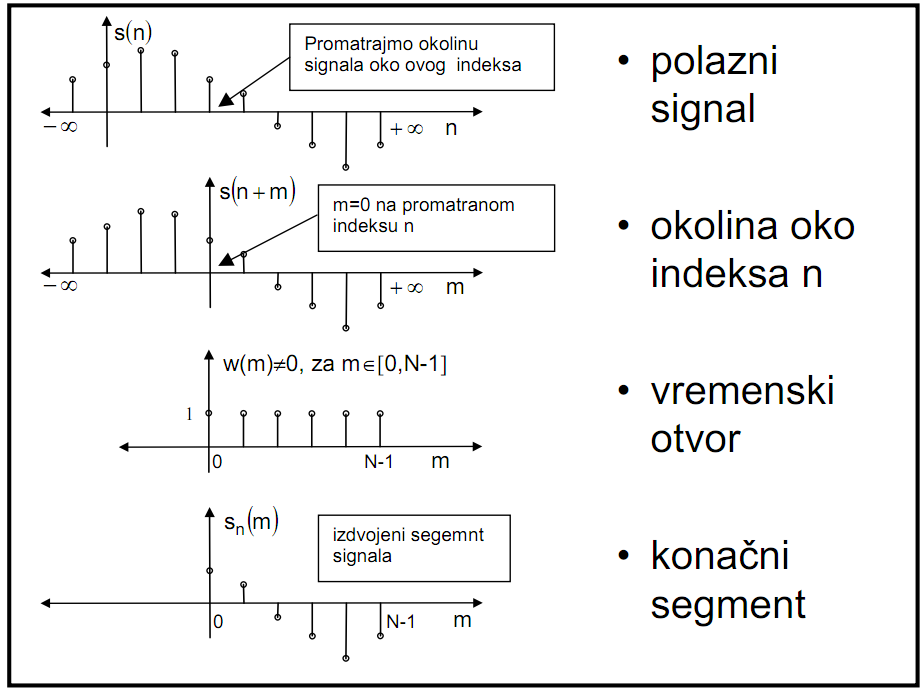
Predikcijski dobitak se ne može ostvariti za nekorelirane signale i to vidimo po spektru, tj. predikcija ima smisla samo za signale obojenog spektra:  


14. Izvedi (dokaži) svojstvo simetrije matrice sustava jednadžbi za nalaženje optimalnog linearnog prediktora.

**DOG06**

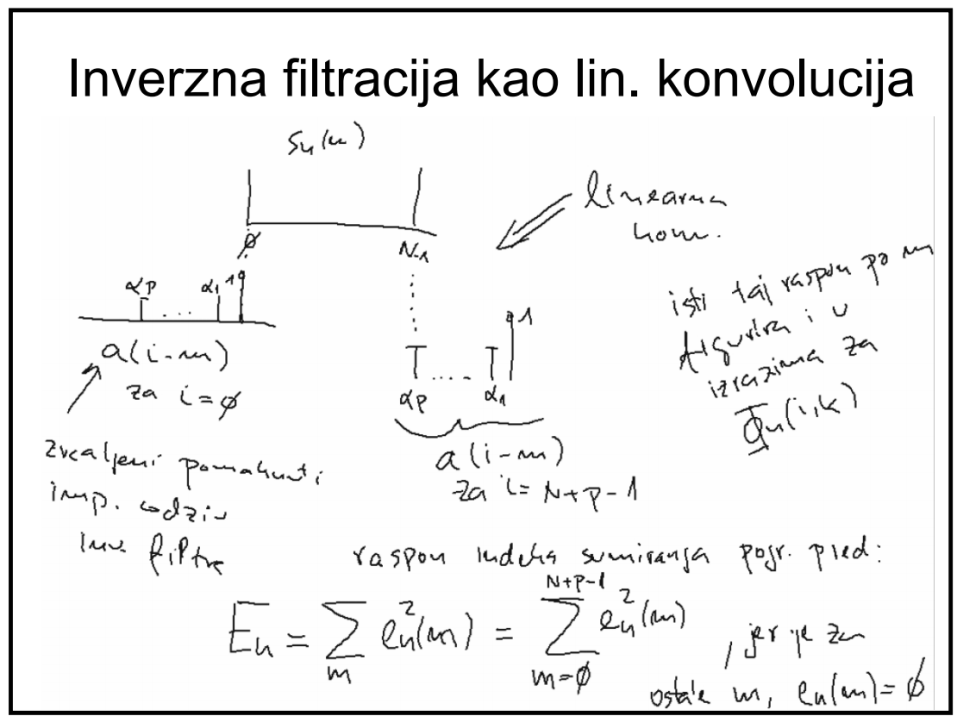
1. Objasni na koji način se kod linearne predikcije postupkom autokorelacije provodi definiranje (ograničavanje) intervala analize u pomoćnom indeksu m? Skiciraj grafički.

Pretpostavka je da je segment signala sn(m) jednak nuli izvan intervala 0 ≤ m ≤ N-1, što odgovara operaciji množenja signala s vremenskim otvorom sn(m) = s(m+n)\*w(m), gdje je w(m) vremenski otvor konačne dužine koji je jednak nuli izvan intervala 0 ≤ m ≤ N-1



2. Za vremenski ograničeni signal konačnog trajanja od N uzoraka i linearni prediktor reda p, skiciraj domenu (raspon vremenskih indeksa) za signal pogreške predikcije en(m). Koja je ukupna dužina signala en(m), tj. koliki je broj uzoraka općenito različit od nule?

Budući da je sn (m) različit od nule samo na intervalu 0 ≤ m ≤ N-1, odgovarajuća pogreška predikcije en(m) za prediktor p-tog reda bit će različita od nule samo na intervalu 0 ≤ m ≤ N-1+p.



3. Zbog čega kod autokorelacijskog postupka linearne predikcije postoji problem na rubovima? Koliko iznosi pogreška predikcije za indeks m=0, a koliko za m=N+p-1? Kakva se rješenja koriste da bi se ovi problemi rubova umanjili?

Pogreška predikcije je velikog iznosa na početku intervala (za 0 ≤ m ≤ p-1), zato što se pokušava provesti predikcija signala na temelju izvornih uzoraka koji su na silu postavljeni na nulu. Pogreška je također velika na kraju intervala (za N ≤ m ≤ N+p-1), jer se pokušava provesti predikcija uzoraka jednakih nuli na temelju uzoraka koji su različiti od nule.  
Jedan od načina kojim je moguće riješiti ovaj problem je primjena vremenskog otvora koji prigušuje signal prema rubovima segmenta sn(m), kao što je npr. Hammingov vremenski otvor.

Za m=0: En = en2 (0)  
Za m = N+p-1: En = en2(N+p-1)

4. Objasni kako i zašto se raspon sume u izrazu za elemente matrice sustava jednadžbi autokorelacijskom postupka može suziti. Koje su konačne granice koje se dobivaju takvim suženjem.

Uvodimo supstituciju m\* = m-i.

Budući da je sn (m\*) jednak nuli izvan intervala 0 ≤ m\* ≤ N-1, jednostavno se može pokazati da se granice sumacije sužavaju. Vrijedi:  
 • za donju granicu sn(m\*) = 0, m\* ≤ 0  
 • analogno za gornju sn(m\*+i-k) = 0, m\* + i -k ≥ 2N  
-> na osnovu toga stvarni raspon indeksa je 0 ≤ m\* ≤ N −1− ( i-k )

5. Izvedi izraz za elemente matrice sustava jednadžbi autokorelacijskog postupka fin(i,k). Koja funkcija se dobiva kao rješenje ovog izvoda i u kojem argumentu?

Dobije se vremenski kratkotrajna autokorelacijska funkcija (eng. short-time autocorrelation function) izračunata za pomak (i-k): Rn(i-k).  
Funkcija Rn(j) je parna funkcija u pomaku j, tj. Rn(j) = Rn(-j), vrijedi:

6. Napiši matrični oblik jednadžbe za optimalne koeficijente linearnog prediktora pomoću autokorelacijskog postupka. Napiši izraz pomoću kojeg se računaju elementi ovog sustava jednadžbi.

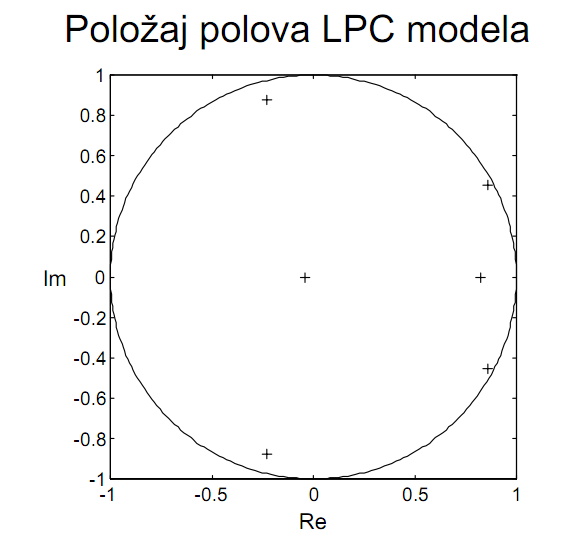
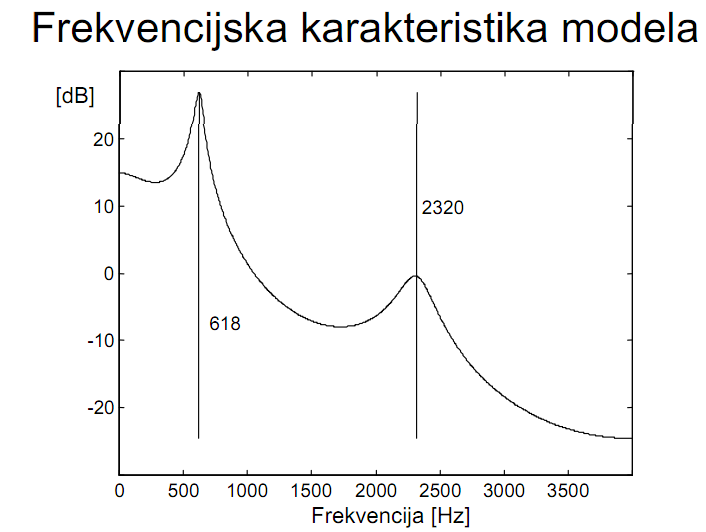
7. Koliki je broj različitih skalarnih veličina koje su potrebne za potpuno formiranje sustava jednadžbi autokorelacijskog postupka? Koje su to veličine? Kako nazivamo tip matrice takve strukture i koje je to njeno temeljno svojstvo strukture?

Postoji p autokorelacijskih koeficijenata koji definiraju cijeli sustav jednadžbi. Te veličine su koeficijenti.  
Matrica autokorelacijskih vrijednosti je Toeplitzove forme (simetrična po obje glavne dijagonale) dimenzija p\*p. Također, vrijednosti elemenata duž svake desno položene dijagonale su jednake.

8. Napiši izraz pomoću kojeg je moguće odrediti ukupnu kvadratnu pogrešku predikcije autokorelacijskog postupka na osnovu optimalnih koeficijenata alfa i koeficijenata autokorelacije za prediktor reda p.

Rn(0) je energija signala, dok su u sumi skalarni produkti vektora i [Rn(k)].

9. Diskutiraj stabilnost modela H(z) koji se dobiva autokorelacijskim postupkom. Kakve polove može imati ovaj model (koje tipove)? Ilustriraj odnos položaja konjugirano kompleksnih polova i formantnih karakteristika modela. Kako na osnovu položaja pola u kompleksnoj ravnini možemo odrediti centralnu frekvenciju i širinu pojasa formanta?

H(z) je stabilan sustav, svi polovi su unutar jedinične kružnice. Ima šest polova: dva realna i dva konjugirano-kompleksna para. Svaki konjugirano-kompleksni par polova određuje jednu rezonantnu karakteristiku koju se poklapa s jednim od formanata govornog signala. Polovi blizu jedinične kružnice imaju veliki Q-faktor i daju izražene maksimume u prijenosnoj funkciji (odgovaraju uskim formantima).

10. Zašto kažemo da LPC model modelira spektralnu ovojnicu vremenski kratkotrajnog spektra govornog signala. Kakva je točnost ovog modeliranja u vrhovima formanata (spektralni maksimumi) u odnosu na točnost modeliranja u spektralnim dolinama između formantnih karakteristika?

LPC model modelira spektralnu ovojnicu vremenski kratkotrajnog spektra govornog signala jer prediktor opisuje sporo promjenjivi dio spektra uzduž frekvencijske osi. Točnost modela najveća je u vrhovima formanata, a manja u spektralnim dolinama.

11. Objasni zašto kažemo da inverzni filtar provodi spektralno izbjeljivanje govornog signala. Koji je odnos log-spektra ulaznog signala i log-spektra signala pogreške predikcije? Kako red prediktora utječe na stupanj izbjeljivanja spektra?

Inverzni filtar provodi spektralno izbjeljivanje jer su vrhovi svih latica približno na istoj visini (unutar 15dB). Što je red prediktora viši, to je bolje izbjeljivanje. Spektar pogreške predikcije računa se kao razlika logaritamskog spektra signala i amplitudno-frekvencijske karakteristike prediktora u logaritamskoj mjeri. Spektar odgovara kvocijentu modela spektra signala i modela frekvencijske karakteristike LPC modela (H(z)).

12. Objasni u kojem su odnosu autokorelacije izdvojenog segmenta govornog signala i autokorelacija impulsnog odziva modela H(z) koji je određen autokorelacijskim postupkom nad tim izdvojenim segmentom.

Autokorelacije izdvojenog segmenta govornog signala i autokorelacija impulsnog odziva modela H(z) su identički jednake za pomake od –p do p, dok se izvan tog područja razlikuju.

**DOG07**

1. Po čemu se razlikuje postupak ograničavanja signala kod autokorelacijskog postupka i kod metode kovarijance? Zašto kod metode kovarijance ne postoje problemi s rubovima?

Kod metode kovarijance ograničavamo samo interval nad kojim računamo pogrešku predikcije, a signal ostavimo kakav jest. Nema problema s rubovima jer svi uzorci signala sudjeluju u određivanju prediktora jednakom “težinom”, te se ne koristi vremenski prozor.

2. Kako se definira raspon indeksa m kod izdvajanja signala u metodi kovarijance? Koji uzorci signala (raspon vremenskih uzoraka) su potrebni za izračunavanje svih elemenata matrice sustava jednadžbi? Napiši izraze iz kojih je to jasno vidljivo.

Odabiremo pomoćnu varijablu m’=m-i, m’=m-k:

Uzorci signala potrebni za izračunavanje sn(m) su vrijednosti van interval

3. Koja svojstva ispunjava matrica sustava jednadžbi za metodu kovarijance? Da li je Toeplitz strukture?

Matrica ispunjava svojstvo simetričnosti.  
Matrica je dimenzije p\*p koja sadrži kroskorelacije za sve pomake i i k je simetrična, ali nije Toeplitzove strukture.

4. Izvedi i opiši odnos elemenata matrice sustava jednadžbi za metodu kovarijance na desno ležećim dijagonalama. Radi čega je ovaj odnos značajan, tj. za što se može iskoristiti?

Na glavnoj dijagonali su elementi koji nose “energiju signala” – energija N susjednih elemenata uz p pomaka.

-> Bilo koja dva elementa matrice sustava koji su u odnosu i su vrlo slični po korelacijskom značenju.  
Svi elementi u desno lež. dijagonalama povezani su izrazom:

5. Opisati efikasni postupak formiranja matrične jednadžbe za metodu kovarijance, na osnovu kros-korelacija fin(i,0) za i=0 do p. Obrazloži numeričku složenost izračuna ovog stupca kros-korelacija, kao i složenost izračuna svih ostalih elemenata matrice sustava jednadžbi.

Korištenjem originalnog izraza izračunamo za i = 0, 1, …, p a sve ostale elemente sustava na glavnoj dijagonali i ispod nje mogu se odrediti rekurzivnim izrazom:  
  
Elementi iznad glavne dijagonale se pronalaze zrcaljenjem oko glavne dijagonale (zbog simetrije).  
Broj operacija za elemente koji se određuju direktnim izrazom je (p+1)\*(Nmnoženja+(N-1)zbrajanja), a za svaki preostali element, rekurzivni izraz traži p2/2 operacija.

6. Koje svojstvo mora ispunjavati matica sustava jednadžbi da bi postupak određivanja rješenja mogli provesti metodom Cholesky dekompozicije? U koji oblik se rastavlja matrica FIn? Opiši strukturu pojedinih dijelova ovog rastava.

Mora ispunjavati svojstvo da je pozitivno definitna, tj. da su njezine vlastite vrijednosti pozitivne, te da je simetrična.

* matricu rastavljamo kao produkt tri matrice V, D i VT
* svojstva tih matrica:
  + V je donja trokutasta matrica s jedinicama na glavnoj dijagonali
  + D je dijagonalna matrica
  + VT je transponirana matrica V, odnosno gornja trokutasta matrica s jedinicama na glavnoj dijagonali

7. Izvedi i opiši dva linearna sustava jednadžbi koji se dobivaju postupkom Cholesky dekompozicije. Koji je redoslijed rješavanja ova dva sustava?

Uvodimo pomoćni stupac dimenzije p, napravimo dva sustava jednadžbi, prvi u pomoćnom stupcu Y i drugi u stupcu konačnog rješenja:  
1. sustav:   
2. sustav:

Prvo odredimo Y iz prvog sustava, a zatim iz drugog.

8. Ilustriraj postupak rješavanja prvog (pomoćnog) sustava jednadžbi za metodu Cholesky dekompozicije u pomoćnoj vektorskoj varijabli Y. Kako nazivamo takav postupak rješavanja linearnog sustava? Napiši izraz za rješenje yi.

– za općenit red p vrijedi rekurznivni izraz:

Unaprijedna supstitucija – rješenje iz prošlog koraka uvrštavamo u izraz za sljedeće rješenje.

9. Ilustriraj postupak rješavanja drugog sustava jednadžbi za metodu Cholesky dekompozicije u konačnoj vektorskoj varijabli alfa. Kako nazivamo takav postupak rješavanja linearnog sustava? Napiši izraz za rješenje alfa\_i.

Nakon određivanja rješenja Y, uvrštavamo ga u drugu jednadžbu: i dobivamo konačni stupac s koeficijentima predikcije .  
Za općenit red p vrijedi rekurzivni izraz:

Koristi se unazadna supstitucija jer se kreće od zadnjeg elementa, te iterativno ide prema prvom.

10. Za primjer linearnog prediktora trećeg reda prikaži matricu sustava jednadžbi FIn čiji su elementi izraženi na osnovu elemenata pomoćne matrice V i dijagonalne matrice D (za metodu Cholesky dekompozicije). Kako nalazimo prvi dijagonalni element matrice D?

Određivanje pomoćnih matrica:  
 =

11. Napiši pomoćne izraze A i B koji se koriste za izračunavanje elemenata matrica V i D za metodu Cholesky dekompozicije. Objasni način primjene ovih izraza, tj. redoslijed kojim se dobivaju nepoznati elementi matrica D i V.

Primjena izraza: naizmjenična primjena B i A:  
 - iz izraza B odredimo d1 - zatim iz izraza A odredimo Vi,1 za i=2, …, p: V1,1=, ,   
 - iz prvog stupca i određenog elementa d1 određujemo prvi stupac matrice V:

1. d1 2. V21, V31, … Vp1 3. d2 4. V32, V42, … Vp2 itd. dok ne nađemo sve elemente.

12. Izvedi izraz za sumarnu kvadratnu pogrešku predikcije optimalnog linearnog prediktora reda p za metodu kovarijance koji je izračunat pomoću Cholesky dekompozicije. Kako se može odrediti pogreška predikcije za prediktor čiji je red p-1?

općenito

Za metodu kovarijance:

Gornjim izrazom se također može računati pogreška predikcije za prediktor reda p-1.

13. Objasni kojim postupkom se složenost postupka izračuna korelacijskog stupca fin(i,0) za i=0 do p za metodu kovarijance može dodatno smanjiti. Koja je ušteda u broju potrebnih operacija ako je širina okvira analize jednaka 250 uzoraka, a korak analize jednak 100 uzoraka.

Složenost se može smanjiti postupkom pomaka okvira analize za N/3 koraka u desno. Tada prve dvije parcijalne sume (u već imamo od prošlog okvira analize i ne računamo ih ponovno.

Time smanjujemo složenost s faktorom 3 (za preklom 2/3). Ušteda u broju operacija je tada:  
N\*(p+1) = 250\*(10+1) = 2750 MAC operacija  
Ušteda je 2750/3 = 916 operacija

14. Diskutiraj stabilnost prediktora određenog postupkom kovarijance, odnosno utjecaj valnog oblika ulaznog signala koji se modelira LPC prediktorom na stabilnost modela.

Ne postoji garancija na stabilnost prediktora. Do pojave nestabilnog LPC modela dolazi u slučajevima kada je okvirom analize zahvaćen segment signala koji se raspiruju prema kraju.

15. Kako širina vremenskog otvora analize N utječe na sličnost postupka kovarijance i autokorelacije? Koja od ove dvije metode daje točnija rješenja ako se koriste vrlo uski vremenski otvori analize koji su jednaki ili neznatno širi od reda prediktora? Koja od tih metoda može savršeno identificirati koeficijente nepoznatog all-pole modela i pod kojim uvjetima?

Ako je širina okvira analize N jako velika u odnosu na red p, tada je rezultat koji se dobiva metodom kovarijance vrlo blizak rezultatu autokorelacijskog postupka s pravokutnim otvorom. Za uske vremenske otvore, točnija rješenja dobit ćemo ako koristimo kovarijantni postupak. Kad je signal nad kojim se provodi analiza uistinu nastao kao impulsni odziv all-pole sustava, takva kovarijantna analiza na okviru od N=p uzoraka će savršeno točno identificirati nepoznate koeficijente modela.

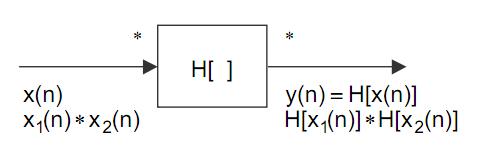
**DOG08**

1. Objasni ključne sličnosti i razlike postupka identifikacije modela vokalnog trakta pomoću linearne predikcije u usporedbi s postupkom homomorfne analize.

Model LPC određuje model vokalnog trakta na osnovu njegovod odziva (pobudni signal nije poznat). Izlaz (govor) je jednak konvoluciji pobude i impulsnog odziva modela. Dekonvolucija – razdvojiti pobudu i impulsni odziv nepoznatog sustava.  
Homomorfna analiza je postupak razdvajanja dvije konvoluirane komponente, tako da se jedna ukloni (zbog svojstva signala da ima relativno ravan spektar: bijeli šum, niz harmonika).

2. Definiraj pojam sustava koji su homomorfni za konvoluciju. skiciraj grafički i označi ulaze i izlaze takvih sustava. Što je generalizirani princip superpozicije?

To su sustavi kji podliježu poopćenom principu superpozicije, tj. sustavi H() koji zadovoljavaju svojstvo da možemo + zamijeniti s \* (generalizirani princip superpozicije).

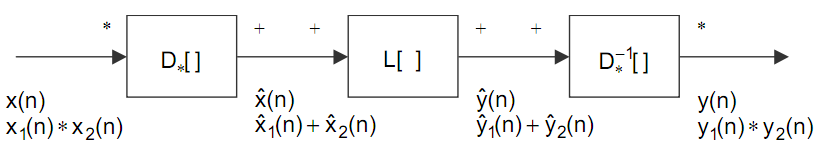


Prolaskom dvije konvoluirane komponente kroz H() na izlazu se dobiva konvolucija individualnih transformacija svake pojedine komp.

3. Što je homomorfni filtar i koja je njegova uloga?

Sustav H() tako projektiran, da transformacija jedne komponente bude bliska jediničnom impulsu , tada je na izlazu konvolucija što je jednako ; tako da se odvaja samo druga komponenta, a prva eliminira.

4. Prikaži i objasni kanonsku izvedbu sustava homomorfnih za konvoluciju. Koje su prednosti takve izvedbe?



D\*() – karakteristični sustav za homomorfnu dekonvoluciju

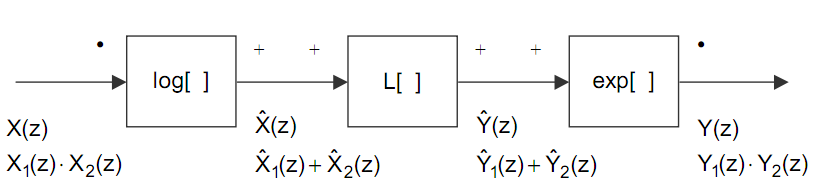
Kaskada tri sustava:  
D\*() konvoluciju pretvara u sumu, L() običan linearni sustav, D\*-1() inverzni sustav za homomorfnu dekonvoluciju, sumu pretvara u konv.

Prednosti: karakteristika ukupnog sustava određena je središnjim linearnih sustavom L() koji se projektira za željenu primjenu: sustav D\*() se ne mijenja, a podliježe generaliziranom principu superpozicije, gdje je ulaz konvolucija, a izlaz zbrajanje. Inverzni karakteristični sustav D\*-1() također je fiksiran sustav, koji transformira aditivnu kombinaciju signala natrag u konvolucijsku kombinaciju izlaza.

5. Prikaži i objasni što je karakteristični sustav za homomorfnu dekonvoluciju. Kako se definira njegov inverz?

To je sustav D\*() koji konvoluciju pretvara u sumu.

6. Skiciraj izvedbu karakterističnog sustava za homomorfnu dekonvoluciju temeljenu na z-transformaciji. Kojom operacijom se umnožak z-transformacija pretvara u aditivnu kombinaciju?



Umnožak z-transformacije pretvara se u aditivnu kombinaciju primjenom logaritma.

7. Kako se definira logaritam kompleksne veličine koji ispunjava tražene uvjete generaliziranog principa superpozicije. Dokaži da je uz takvu definiciju traženi uvjet uistinu ispunjen.

Dokaz:

8. Zbog čega problem fazne višeznačnosti predstavlja problem u definiciji kompleksnog logaritma i kako se taj problem rješava u sustavim homomorfnim za konvoluciju?

Problem fazne višeznačnosti predstavlja problem jer je imaginarni dio jednak kutu z-transformacije na jediničnoj kružnici, dok je kod numeričkog izračuna taj kut uvijek u rasponu od 0 do 2pi, pa dva ulazna argumenta X1 i X2 mogu dati jednaki kompleksni logaritam umnoška.  
To se rješava tako da fazna frekvencijska funkcija oba spektra X1 i X2 mora biti kontinuirana neparna funkcija od ; i na mjestima faznih skokova potrebno je izračunatu osnovnu faznu vrijednost nadopuniti višekratnicima od 2pi -> PHASE-UNWRAPPING.

9. Skiciraj i objasni izvedbu karakterističnog sustava za homomorfnu dekonvoluciju temeljenu na Fourierovoj transformaciji vremenski diskretnog signala (DTFT).

Fourierova transformacija također pretvara konvoluciju u umnožak, ali modelira ponašanje sustava samo u stacionarnom stanju. Izlaz karakterističnog sustava D\*() je:

10. Što je Kepstar i kako se definira? Kako se definira kompleksni, a kako realni kepstar? Tko je zaslužan za uvođenje pojma Kepstar?

Kepstar je rezultat Fourierove transformacije logaritma spektra signala. Pojam kepstar uveo je A. Michael Noll 1967. godine.

Kompleksni kepstar je inverzna Fourierova transformacija kompleksnog logaritma Fourierove transformacije , tj. izlaz karakterističnog sustava za homomorfnu dekonvoluciju:  
  
Realni kepstar je inverzna Fourierova transformacija logaritma spektra snage signala:

11. U kojem su odnosu realni i kompleksni kepstar? Napiši izraz pomoću kojeg je moguće odrediti realni kepstar iz kompleksnog kepstra.

12. Koja svojstva ispunjavaju realni i kompleksni kepstar (simetričnost, realnost, ...)?

Oba su realne sekvence, jer i običan i kompleksni logaritam spektra daju konjugirane simetrične funkcije, pa je njihova IFT realna. Realni kepstar je uvijek simetrična sekvenca oko ishodišta zbog realnosti logaritma modula c[n] = c[-n].

13. Za koju svrhu koristimo kepstralnu analizu na govornom signalu? Koji su rezultati ove analize, te kako treba odabrati linearni operator u kanonskom obliku karakterističnog sustava za homomorfnu dekonvoluciju u svrhu nalaženja traženih rezultata?

Kepstralnu analizu koristimo zbog toga da saznamo koji dio pripada pobudi, a koji modelu vokalnog trakta. To je moguće jer je taj signal jednak konvoluciji pobudnog signala i impulsnog odziva vokalnog trakta.  
Rezultat analize je odvajanje dijela pobude i dijela koji pripada vokalnom traktu, a linearni susta nužno ovisi o svojstivma njegovih ulaznih signala, pa se pri izboru linearnog operatora moraju uzeti u obzir i svojstva ulaznih signala.

14. Za općenit rekurzivni sustav sa polovima i nulama unutar i izvan jedinične kružnice, napiši opći oblik prijenosne funkcije X(z) kao umnožak korjenih faktora. Izračunaj logaritam ove prijenosne funkcije.

Logaritam:

15. Pokaži kako se u vremensku domenu (tj. u domenu kompleksnog kepstra) preslikava logaritam jednog korjenog faktora brojnika općenitog modela X(z), za k-tu nulu unutar jedinične kružnice (koristiti Taylorov razvoj).

-> niz uzoraka na pozitivnim vremenskim indeksima n čije su amplitude

Doprinos k-te nule: i

16. Napiši konačni izraz za kompleksni kepstar implusnog odziva sustava X(z) definiranog na osnovu pozicije polova i nula unutar i izvan jedinične kružnice. Koji koeficijenti modela određuju iznos kepstra na pozitivnim vremenskim indeksima, a koji na negativnim?

ak i ck na pozitivnim vremenskim indeksima, bk i dk na negativnim vremenskim indeksima.

17. Izračunaj kompleksni kepstar impulsnog odziva FIR sustava s dvije konjugirano kompleksne nule unutar jedinične kružnice, na radijusu r, po kutom +/-fi.

18. Izračunaj kompleksni kepstar impulsnog odziva all-pole IIR sustava drugog reda s dva konjugirano kompleksna pola unutar jedinične kružnice, na radijusu r, po kutom +/-fi. (?)

19. O čemu ovisi brzina opadanja modula kepstra s udaljavanjem od ishodišta? Napiši izraz kojim se definira gornja granica (garantirano opadanje). Kako izgleda kepstar sustava čiji su polovi ili nule na jediničnoj kružnici? (?)

Brzina opadanja modula kepstra s udaljavanjem od ishodišta ovisi o maksimalnim modulima svih koeficijanata za (eksponencijalno prigušenje kepstralnih uzoraka).

– maksimalni modul koeficijenata

- faktor proporcionalnosti

20. Što su sustavi minimalne faze? Koja svojstva ispunjava kompleksni kepstar impulsnog odziva sustava minimalne faze?

Sustavi minimalne faze su sustavi koji imaju samo polove i nule unutar jedinične kružnice (potpuno su opisani koeficijentima ak i ck, te faktorom pojačanja A).  
Kompleksni kepstar takvog sustava ispunjuje svojstvo da je fazno-frekvencijska karakteristika potpuno određena s amplitudno-frekvencijskom karakteristikom.

Ovaj kepstar može se odrediti direktno iz cmin[n]

21. Napiši i obrazloži kako i za koje sustave se kompleksni kepstar može jednoznačno odrediti iz realnog kepstra c[n].

Kompleksni kepstar može se jednoznačno odrediti iz realnog za sustave minimalne faze, tj. realni kepstar predstavlja parni dio kompleksnog kepstra.

22. Napiši rekurzivni izraz pomoću kojeg se uzorci kompleksnog kepstra impulsnog odziva sustava minimalne faze mogu odrediti direktno iz uzoraka tog impulsnog odziva.

23. Objasni kako analitički možemo odrediti kompleksni kepstar idealiziranog pobudnog signala vokalnog trakta koji se sastoji od niza kroneckerovih delta impulsa na razmaku Np. U kojim vremenski trenutcima se javljaju uzorci kepstra ovog signala?

Izrazimo P(z) kao produkt korjenih faktora oblika   
Kompleksni kepstar ima uzorke u vremenskim trenucima koji su cjelobrojni višekratnici Np.

24. Izvedi kompleksni kepstar i njegovu z-transformaciju za signal koji se sastoji od dva impulsa: delta[n] i alfa\*delta[n-Np]. Diskutiraj oblik i brzinu opadanja ovog kepstra.

Dobiveni beskonačni periodični niz impulsa s periodom Np i alternirajućim eksponencijalno padajućim amplitudama , amplituda impulsa teži prema nuli kako r raste (za ). Ne postoji impuls u n=0, nego tek u n=Np.

25. Objasni ukratko kako se kepstralna analiza može koristiti za utvrđivanje zvučnosti govornog signala i određivanje fundamentalne frekvencije titranja glasnica.

Analiza nad segmentima konačnog trajanja

Za zvučne glasove: V(z) ima samo polove (ak i bk su nule za svaki k).  
Kod nazalnih i bezvučnih glasova: V(z) ima i polove i nule  
 - za zvučne glasove (vokalni trakta -> FIR sustav) -> Hv(z) imat će polove unutar jedinične kružnice, a nule unutar i izvan jedinične kružnice (nije sustav minimalne faze). U kepstru šiljci na indeksima Np i njegovim cjelobrojnim višekratnicima. c(n) je simetričan oko nule.

Na osnovu pozicije izraženih maksimuma u kepstru zvučnih glasova moguće je odrediti frekvencije titranja glasnica (pitch detection), a u kepstru bezvučnih glasova šiljci ne postoje. Ako je veličina šiljka iznad nekog praga, ulazni govorni segment je zvučni pri čemu pozicija šiljka predstavlja dobru estimaciju osnovnog perioda, a ako je veličina šiljka ispod praga -> bezvučni.

26. Objasni kako izgleda i kako se odabire linearni sustav za izdvajanje pobudnog signala vokalnog trakta pomoću homomorfnog filtra. Koje svojstvo kepstra govora koristimo za ovo izdvajanje?

Koristimo svojstvo da su komponente kepstra razdvojene u vremenu, tj. dio blizak nuli (low time part) je određen s hv(n), a preostali dio koji je udaljen od nule (high time part), je određen s p(n) -> vremenska dislokacija.  
Govorni signal s(n) množi se s vremenskim otvorom w(n), pa se na tom segmentu primjenjuje D\*() čiji je izlaz kepstar cp(n), zatim na cp(n) primjenjujemo dva linearna operatora L1() i L2(), pa se ta dva dijela propuštaju kroz inverzni karakteristični sustav D\*-1(). Tako dobijemo dva izlazna signala h(n) i e(n).  
Ako pravilno odaberemo L1() i L2(), h(n) je impulsni odziv linearnog sustava, a e(n) je njegova pobuda.

**DOG09**

1. Objasni kako se kod kepstralne analize govora rješava problem nestacionarnosti govornog signala. Koju transformaciju koristimo umjesto sve-vremenske Fourierove transformacije vremenski diskretnog signala (DTFT)?

Problem nestacionarnosti rješava se u postupcima vremenski kratkotrajne analize – provodi se na kratkim odsječcima govora. Umjesto svevremenske FT, koristimo DFT i IDFT transformaciju koja je opisana sumama u oba smjera transformacije.

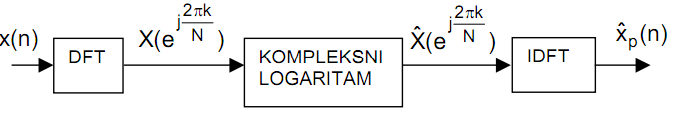
2. Koji su temeljni parametri kojima se definira vremenski kratkotrajna homomorfna analiza? Koje su tipične vrijednosti ovih parametara za analizu govornog signala?

Temeljni parametri: segment signala se izdvaja pomoću vremenskog otvora odgovarajuće širine i oblika (pomak ishodišta analize za odabrani broj uzoraka – korak analize), širina vremenskog otvora < koraka analize, dva susjedna izdvojena segmenta za dva susjedna okvira analize se nužno preklapaju, učestalost analize (broj okvira analize u sekundi), učestalost mora biti odabrana u skladu s vremenskom dinamikom izmjene spektralnih svojstava govora (i pobude vokalnog trakta).

3. Opiši i objasni numeričke postupke izračuna kepstra govora korištenjem diskretne Fourierove transformacije (DFT). Kako se bira broj uzoraka DFT-a i kako on utječe na svojstva tako izračunatog kepstra. (?)

Aproksimiramo računanje spektra pomoću DFT-a tako da zamijenomo Fourierove transformacije DFT-om:

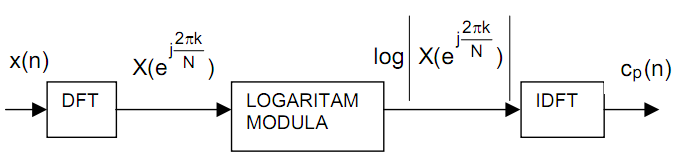
Otipkavanje u spektru -> periodičko proširenje u vremenu.



Kompleksni kepstar nalazimo (za segment signala x(n) dužine N uzoraka) kao:

Aproksimacija stvarnog kepstra:

Aproksimacija cp(n) i stvarni kepstar:



N mora biti velik (zbog aliasinga), da se osigura kontinuitet fazne karakteristike (N=512 ili više).

4. Objasni i formulom izrazi koji je odnos kepstra izračunatog pomoću svevremenske Fourierove transformacije (DTFT) izdvojenog odsječka govornog signala i kepstra izračunatog DFT-om u N točaka za isti taj segment trajanja manjeg od N uzoraka. (?)

Kepstar izračunat pomoću DFT-a možemo dobiti periodičkim ponavljanjem kepstra izračunatog korištenjem prave Fourierove transformacije svakih N uzoraka, tj. prema slijedećem izrazu:

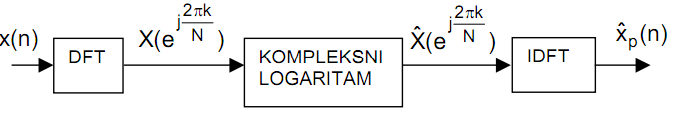
DTFT:

DFT:

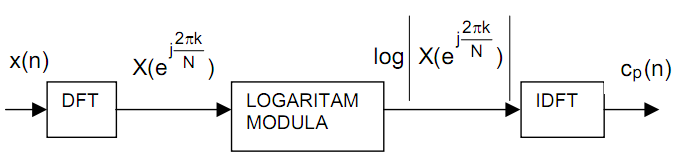
5. Koje svojstvo kepstra govornog signala opravdava korištenje DFT-a umjesto DTFTa, tj. zašto možemo ignorirati pojavu aliasinga u vremenskoj domeni? Za koji tip govornog signala je problem aliasinga izraženiji?

Pojavu aliasinga možemo ignorirati zbog toga što je model formiranja govornog signala temeljen na vrlo sporo promjenjivom linearnom sustavu pobuđenim s periodičnim nizom impulsa (ili slučajnim šumom).  
Najizraženiji je a zvučne govorne signale (sve osim nazalnih).

6. Skiciraj blok shemu i napiši izraze za postupak izračunavanja kompleksnog kepstra korištenjem DFT-a



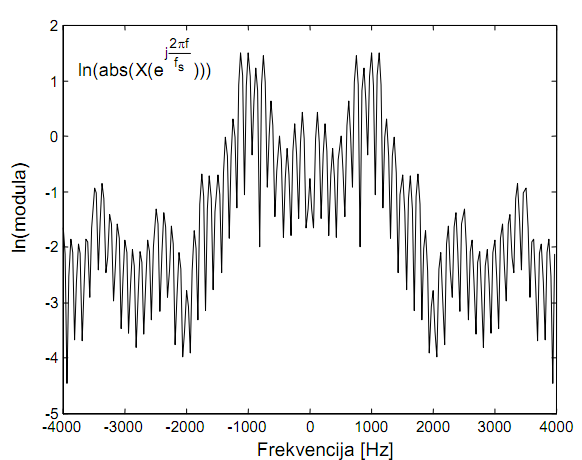
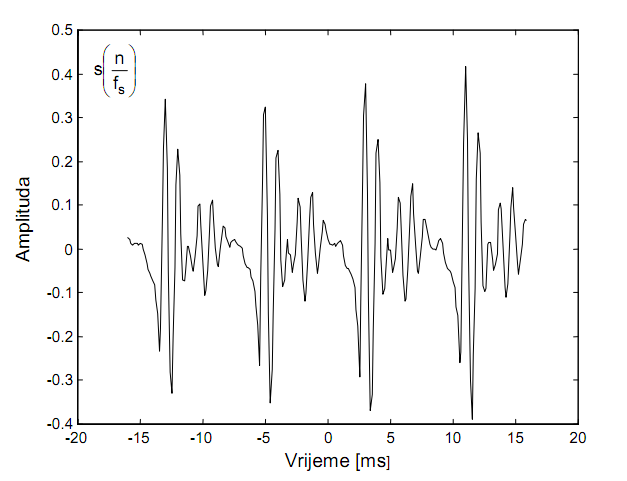
7. Skiciraj blok shemu i napiši izraze za postupak izračunavanja realnog kepstra korištenjem DFT-a. Koji je odnos ovog kepstra i kepstra koji se dobiva DTFTom?

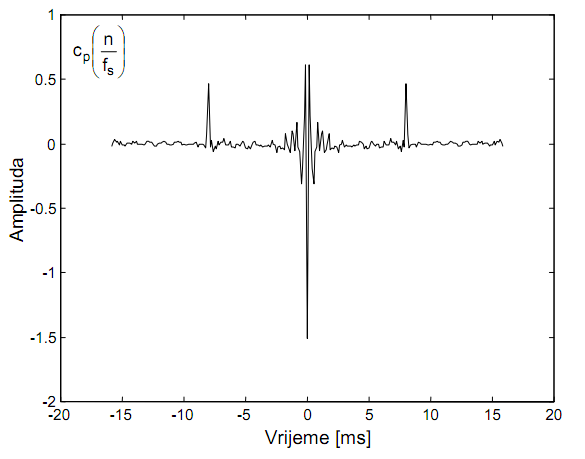


DTFT:

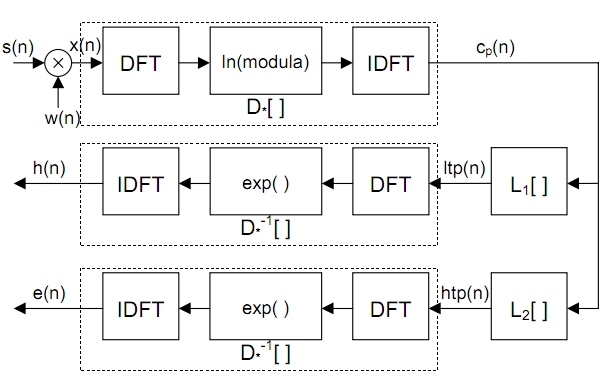
DFT:

8. Za primjer samoglasnika kvalitativno skiciraj: valni oblik izdvojenog segmenta govora, modul njegovog DFTa u log mjerilu i oblik realnog kepstra ovog segmenta. Označi koji dio realnog kepstra odgovara pobudi, a koji dio vokalnom traktu.



- veliki šiljak: pobuda, mali šiljci: vokalni trakt

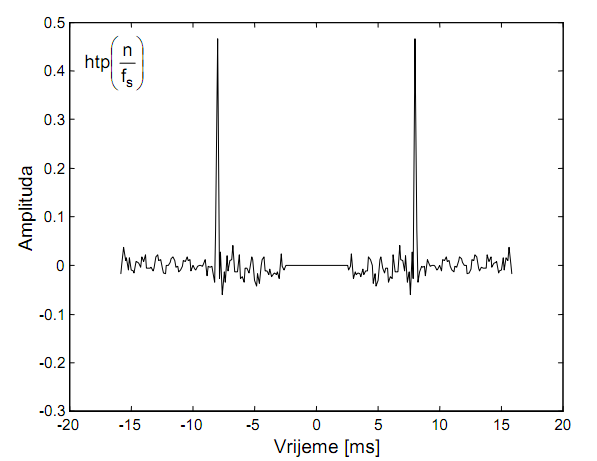
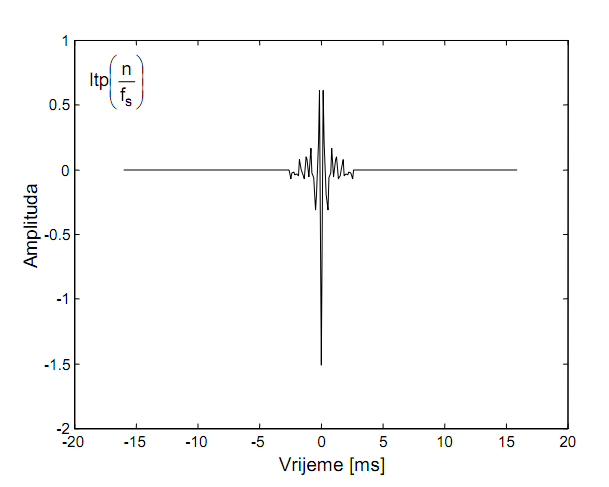
9. Skiciraj cjelokupnu blok shemu sustava za homomorfnu analizu govora kojim se govorni signal rastavlja u signal pobude i impulsni odziv vokalnog trakta. Kako su definirani linearni sustavi koji obavljaju ovo razdvajanje (L1, L2)?



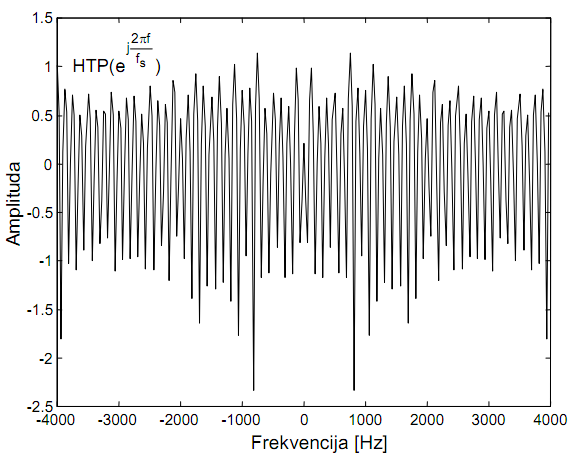
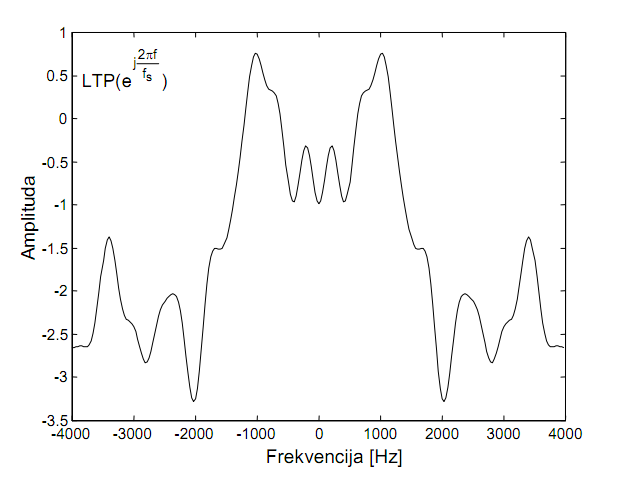
Niskovremenski dio, ltp, izdvaja dio oko ishodišta, a drugi dio je njegov komplement http:

|  |  |
| --- | --- |
|  |  |
|  |  |

10. Skiciraj oblik spektra niskovremenskog i visokovremenskog dijela kepstra samoglasnika. Što se dobiva zbrajanjem ova dva log-spektra? Skiciraj konačni vremenski oblik ova dva razdvojena dijela nakon propuštanja kroz inverzni sustav za homomorfnu dekonvoluciju. Čemu odgovara svaki pojedini od ta dva dobivena signala?



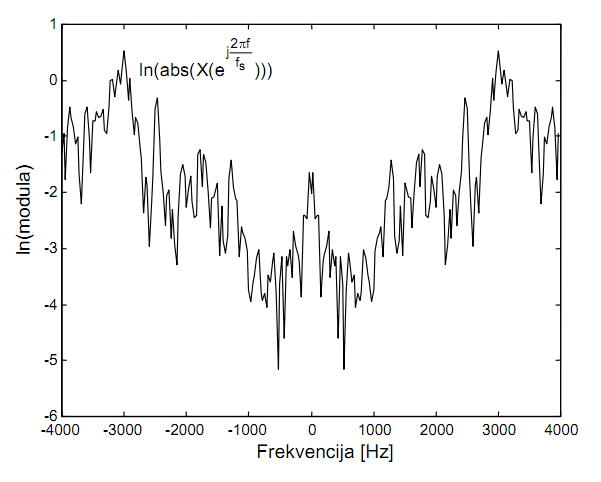
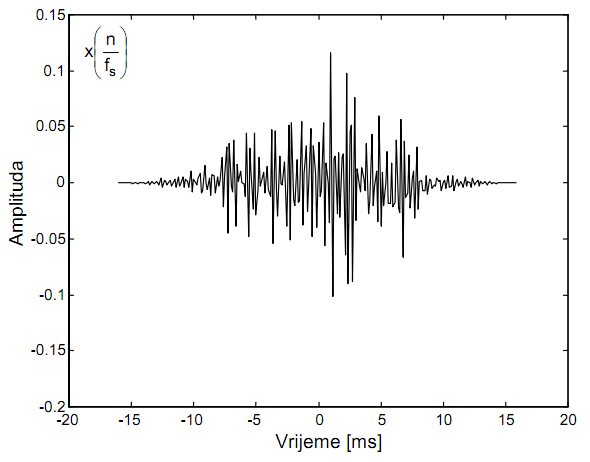
Niskovremenski (lijevo) i visokovremenski (desno) dio spektra. Njihovim zbrajanjem dobiva se jedinični operator, suma DFT-a od lpt i htp dijela jednaka je logaritmu modula DFT-a.

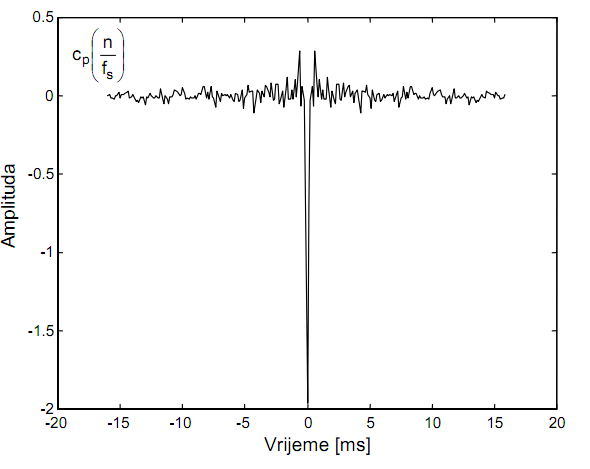
  
LTP dio kepstra opisuju sporo promjenjivu spektralnu ovojnicu govornog signala ( log – magnituda frekvencijske karakteristike modula vokalnog trakta), a HTP dio kepstra trebao bi biti jednag log spektru pobudnog signala.

11. Zbog čega se prilikom primjene sustava za homomorfnu analizu govora korištenjem realnog kepstra konvolucijom između estimirane pobude vokalnog trakta i njegovog impulsnog odziva dobiva signal koji nema jednak valni oblik kao i ulazni segment govora (tj. zašto nemamo potpunu rekonstrukciju)? U kojem smislu su ova dva signala ipak jednaki?

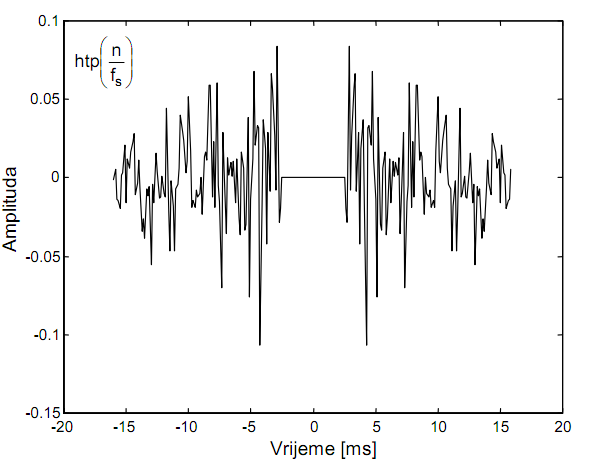
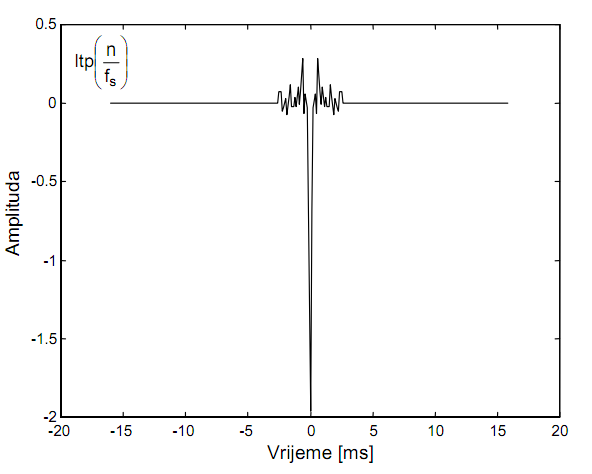
Dobiva se signal koji nema jednak valni oblik kao i ulazni segment govora jer u procesu odbacujemo faznu karakteristiku.  
Međutim, ova dva signala su ipak jednaka u smislu modula spektra: modul spektra konvolucije e(n)\*h(n) je identički jednak modulu spektra izdvojenog segmenta x(n).

12. Za primjer bezvučnog suglasnika (npr. glas 's') kvalitativno skiciraj: valni oblik izdvojenog segmenta govora, modul njegovog DFTa u log mjerilu i oblik realnog kepstra ovog segmenta. Označi koji dio realnog kepstra odgovara pobudi, a koji dio vokalnom traktu.

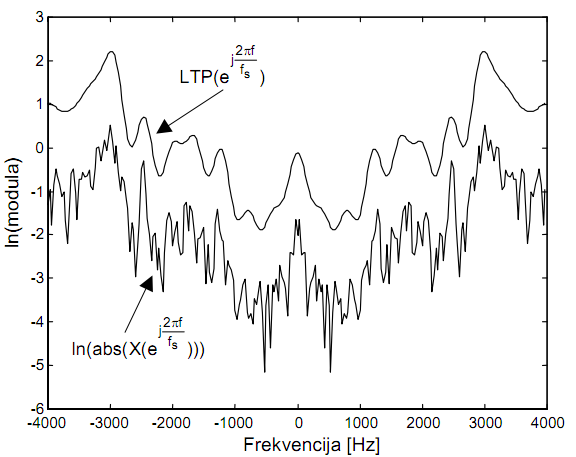


- centar: pobuda, nema šiljaka u htp dijelu

13. Skiciraj oblik spektra niskovremenskog i visokovremenskog dijela kepstra bezvučnog suglasnika. Što se dobiva zbrajanjem ova dva log-spektra? Skiciraj konačni vremenski oblik ova dva razdvojena dijela nakon propuštanja kroz inverzni sustav za homomorfnu dekonvoluciju. Čemu odgovara svaki pojedini od ta dva signala?



Zbrajanjem se dobiva spektralna ovojnica signala zajedno s prirodnim logaritmom modula DFT-a signala x(n) pomnoženog s vremenskim otvorom. LTP odgovara spektralnoj ovojnici signala (formatnoj strukturi govornog signala), a HTP dio ima oblik spektra bijelog šuma.



14. Objasni zašto se kao pobudi signal vokalnog trakta za slučaj homomorfne analize bezvučnog suglasnika korištenjem realnog kepstra dobiva signal koji izgleda kao jedinični impuls sa pribrojenim šumom male amplitude? Kako ovaj pobudni signal učiniti sličnijim stvarnoj bezvučnoj pobudi vokalnog trakta?

Kad se pobudi signal vokalnog trakta za slučaj homomorfne analize bezvučnog suglasnika korištenjem realnog kepstra dobiva se signal koji izgleda kao jedinični impuls s pribrojenim šumom male amplitude jer sve komponente DFT-a HTP dijela imaju nultu fazu, pa se idealno pribrajaju na vremenskom indeksu n=0 (koherentne su).  
Promjenom fazne karakteristike pobudnog signala (odabrali je slučajno između 0 i 2pi – fazu pojedinog spektralnog uzorka)