# *Kvantizacija filtarskih koeficijenata*

## *Kvantizacija filtarskih koeficijenata*

- Točnost s kojom se mogu specificirati koeficijenti filtra ograničena je duljinom riječi računala ili duljinom registra.
- Kvantizacija koeficijenata ima za posljedicu promjenu položaja polova i nula, tj. promjenu frekvencijske karakteristike.
- Osjetljivost frekvencijske karakteristike na kvantizaciju koeficijenata moguće je minimizirati realizacijom filtra kao spoja sekcija drugog reda (paralelna ili kaskadna realizacija).

## *Analiza osjetljivosti na kvantizaciju filtarskih koeficijenata*

Razmotrimo općeniti IIR filtar s prijenosnom funkcijom

$$H(z)=\frac{\sum_{k=0}^{M} b_k z^{-k}}{1+\sum_{k=1}^{N} a_k z^{-k}}$$

Direktna realizacija ovog filtra s kvantiziranim koeficijentima daje prijenosnu funkciju

$$\hat{H}(z) = \frac{\sum_{k=0}^{M} \hat{b}_k z^{-k}}{1+\sum_{k=1}^{N} \hat{a}_k z^{-k}}$$

kvantizirani koeficijenti

Vezu $\hat{a}_k$ i $\hat{b}_k$ s nekvantiziranim koeficijentima $a_k$ i $b_k$ možemo iskazati slijedećim relacijama

$$\hat{a}_k = a_k + \Delta a_k \qquad k = 1,2,\ldots,N$$

$$\hat{b}_k = b_k + \Delta b_k \qquad k = 0,1,\ldots,M$$

$\Delta a_k$ i $\Delta b_k$ predstavljaju grešku kvantizacije

Nazivnik prijenosne funkcije možemo prikazati i u obliku

$$A(z)=1+\sum_{k=1}^{N} a_k z^{-k} = \prod_{k=1}^{N}\left(1-p_k z^{-1}\right)$$

gdje su $p_k$ polovi prijenosne funkcije $H(z)$.

Slično bi mogli prikazati nazivnik prijenosne funkcije $\hat{H}(z)$:

$$\hat{A}(z)=\prod_{k=1}^{N}\left(1-\hat{p}_k z^{-1}\right)$$

gdje je $\hat{p}_k = p_k + \Delta p_k \qquad k = 1,2,\ldots,N$,

a $\Delta p_k$ je greška koja nastaje kao rezultat kvantizacije filtarskih koeficijenata.

Povežimo grešku $\Delta p_k$ s greškom kvantizacije $\Delta a_k$.

$$\Delta p_i = \sum_{k=1}^{N} \frac{\partial p_i}{\partial a_k}\Delta a_k$$

inkrementalna promjena pola $p_i$ zbog promjene koeficijenta $a_k$

Totalna pogreška $\Delta p_i$ dana je sumom inkrementalnih grešaka zbog promjene u svakom koeficijentu $a_k$.

Parcijalne derivacije $\frac{\partial p_i}{\partial a_k}$ možemo naći na slijedeći način:

$$\left(\frac{\partial A(z)}{\partial a_k}\right)_{z=p_i} = \left(\frac{\partial A(z)}{\partial z}\right)_{z=p_i}\left(\frac{\partial p_i}{\partial a_k}\right)$$

Tada vrijedi

$$\frac{\partial p_i}{\partial a_k} = \frac{\left(\frac{\partial A(z)}{\partial a_k}\right)_{z=p_i}}{\left(\frac{\partial A(z)}{\partial z}\right)_{z=p_i}}$$

Brojnik je

$$\left(\frac{\partial A(z)}{\partial a_k}\right)_{z=p_i} = \left(\frac{\partial\left[1+\sum_{k=1}^{N} a_k z^{-k}\right]}{\partial a_k}\right) = \left(z^{-k}\right)_{z=p_i} = p_i^{-k}$$

Nazivnik je

$$\left(\frac{\partial A(z)}{\partial z}\right)_{z=p_i} = \left\{\frac{\partial}{\partial z}\left[\prod_{l=1}^{N}\left(1-p_l z^{-1}\right)\right]\right\}_{z=p_i}$$

$$= \left\{\sum_{k=1}^{N}\frac{p_k}{z^2}\prod_{\substack{l=1\\ l\neq k}}^{N}\left(1-p_l z^{-1}\right)\right\}_{z=p_i} = \frac{1}{p_i^N}\prod_{\substack{l=1\\ l\neq k}}^{N}\left(p_i - p_l\right)$$

Sada možemo napisati

$$\frac{\partial p_i}{\partial a_k} = \frac{p_i^{N-k}}{\prod_{\substack{l=1\\ l\neq i}}^{N}\left(p_i - p_l\right)}$$

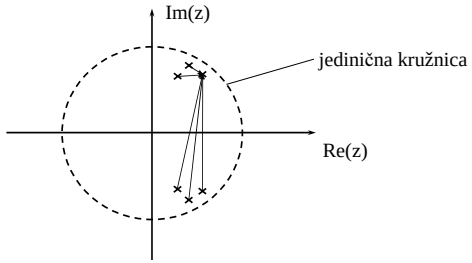
Totalna perturbacijska greška $\Delta p_i$ iznosi

$$\Delta p_i = \sum_{k=1}^{N}\frac{p_i^{N-k}}{\prod_{\substack{l=1\\ l\neq i}}^{N}\left(p_i - p_l\right)}\Delta a_k$$

Ovaj rezultat daje mjeru osjetljivosti $i$-tog pola na promjenu koeficijenata $a_k$.

Sličan rezultat se može dobiti za osjetljivost nula na promjenu koeficijenata $b_k$.

Članovi ($p_i$-$p_l$) u nazivniku predstavljaju vektore od polova $p_l$ do pola $p_i$.




Ako su polovi tijesno grupirani, duljina vektora $|p_i$-$p_l|$ će biti mala. To će imati za posljedicu veliku osjetljivost i stoga veliku perturbacijsku grešku.

Greška $\Delta p_k$ se može minimizirati maksimizacijom $|p_i$-$p_l|$. To se može postići realizacijom filtra višeg reda uz pomoć sekcija drugog reda.

U tom slučaju je svaki par kompleksnih polova realiziran nezavisno od drugih polova , te je i greška u pojedinom paru polova nezavisna od njihove udaljenosti od drugih polova filtra.

Za kaskadnu realizaciju isti argumenti vrijede i za nule.

**Kaskadna realizacija je superiornija direktnoj.**

Kod paralelne realizacije nule se određuju implicitno. Pojedina nula osjeća kvantizacijsku pogrešku u brojniku i nazivnicima svih sekcija drugog reda.
Ipak, paralelna realizacija je puno bolja nego direktna.

# *Realizacija sekcija II reda*

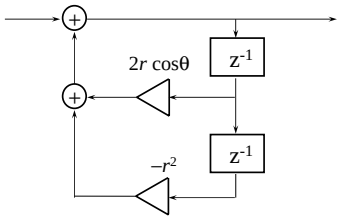
Čak i u slučaju sekcije s dva pola, struktura koja se koristi za realizaciju te filtarske sekcije igra važnu ulogu u greškama nastalim zbog kvantizacije koeficijenata. Pokažimo to kroz slijedeći slučaj.

Neka je zadan filtar s dva pola čija je prijenosna funkcija

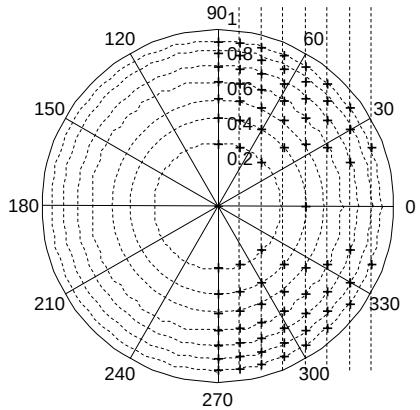
$$H(z) = \frac{1}{1 - (2r\cos\theta)z^{-1} + r^2 z^{-2}}$$

Njegovi polovi su $p_{1,2} = r \cdot e^{\pm j\theta}$

Pri direktnoj realizaciji sustav ima dva koeficijenta $a_1$=2$r$ cosθ i $a_2$=–$r^2$.

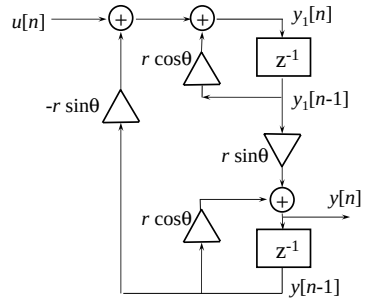



Nakon kvantizacije koeficijenata $a_1$=2$r$ cosθ i $a_2$= –$r^2$, polovi leže na mreži u z ravnini koja je definirana presjekom koncentričnih kružnica (kvantizacija $a_2$=–$r^2$) i vertikalnih pravaca (kvantizacija $a_1$=2$r$ cosθ).




4-bitna kvantizacija (3 bit + predznak)

Alternativna realizacija filtra s dva pola dana je slikom




$y_1(n)=u(n) + r\cos\theta\, y_1(n-1) - r\sin\theta\, y(n-1)$

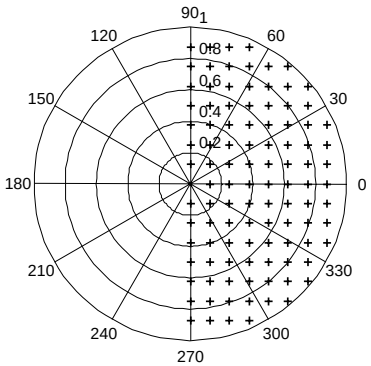
$y(n)=r\sin\theta\, y_1(n-1) + r\cos\theta\, y(n-1)$

Transformacijom ove dvije jednadžbe dolazimo do

$$H(z) = \frac{(r\sin\theta)z^{-1}}{1 - (2r\cos\theta)z^{-1} + r^2 z^{-2}}$$

U ovom obliku postoje dva koeficijenta $\alpha_1$=$r$ sinθ i $\alpha_2$=$r$ cosθ.

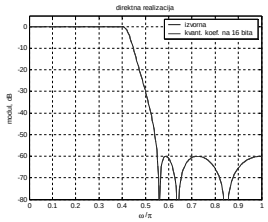
Mogući polovi sada su jednoliko raspoređeni po pravokutnoj mreži.



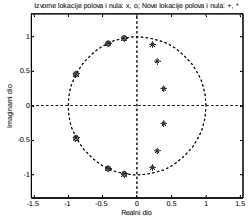

- Polovi su sada jednoliko raspoređeni unutar jedinične kružnice.
- Cijenu za to plaćamo povećanjem broja operacija.
- Idealno bi bilo odabrati strukturu koja omogućava gust raspored mogućnosti za polove u području gdje polovi inače leže.

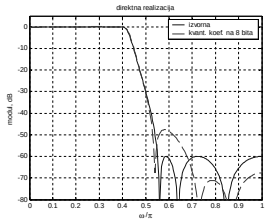
# *Primjer kvantizacije koeficijenata*

Na primjeru IIR eliptičkog filtra šestog reda biti će pokazan utjecaj kvantizacije koeficijenata.
Usporediti će se direktna i kaskadna realizacija.

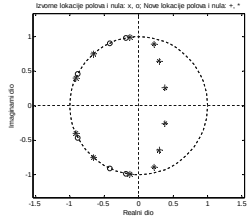


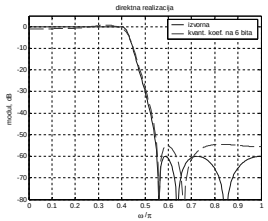

direktna realizacija




direktna realizacija

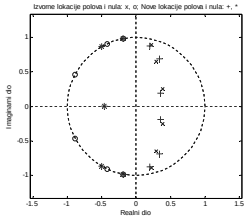


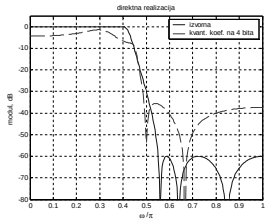




direktna realizacija

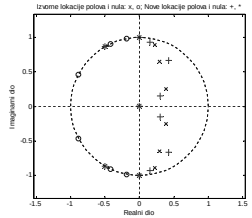
nule za kvantizirane koeficijente

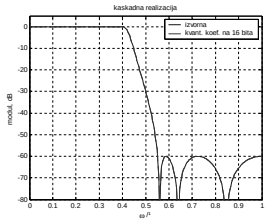
-2.1537
-0.1910 + 0.9816i
-0.1910 - 0.9816i
-0.5000 + 0.8660i
-0.5000 - 0.8660i
-0.4643



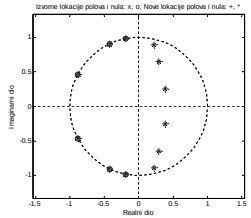



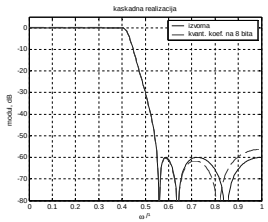

direktna realizacija



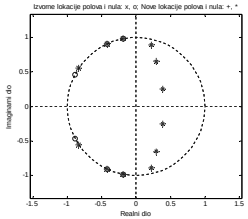



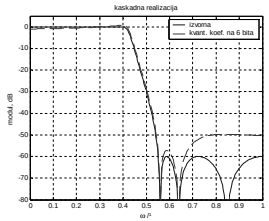

Kaskadna realizacija



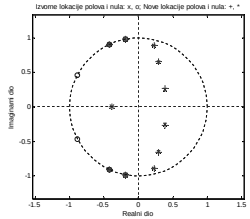



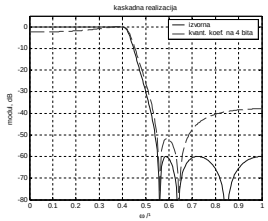

Kaskadna realizacija



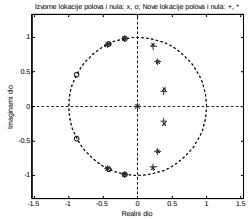



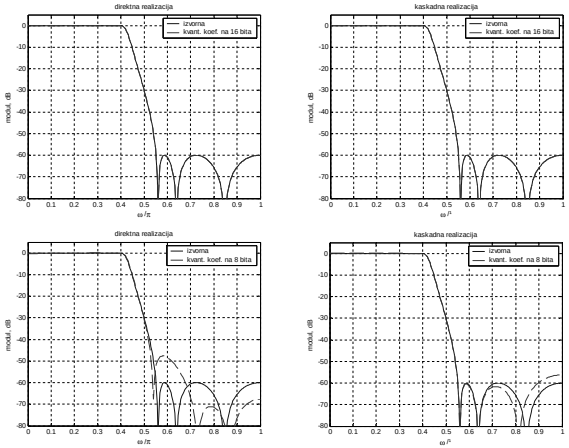

Kaskadna realizacija

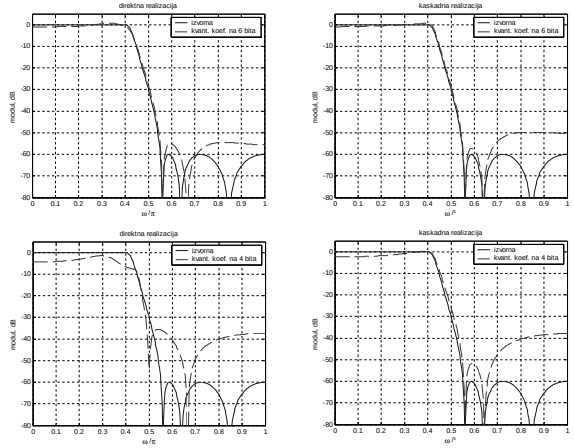


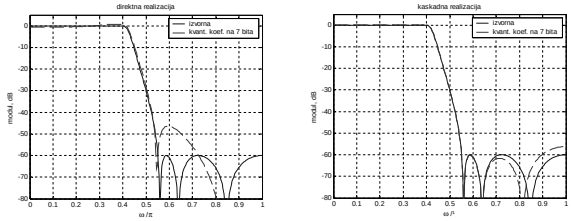




Kaskadna realizacija




direktna realizacija
izvorna
kvant. koef. na 16 bita
kaskadna realizacija
izvorna
kvant. koef. na 16 bita
direktna realizacija
izvorna
kvant. koef. na 8 bita
kaskadna realizacija
izvorna
kvant. koef. na 8 bita
modul, dB
ω /π


direktna realizacija
izvorna
kvant. koef. na 6 bita
kaskadna realizacija
izvorna
kvant. koef. na 6 bita
direktna realizacija
izvorna
kvant. koef. na 4 bita
kaskadna realizacija
izvorna
kvant. koef. na 4 bita
modul, dB
ω /π


direktna realizacija
izvorna
kvant. koef. na 7 bita
kaskadna realizacija
izvorna
kvant. koef. na 7 bita
modul, dB
ω /π