

D

Gercelov algoritam

Gercelovim (Göertzel) algoritmom [74] se izračunava proizvoljni odbirak spektra signala u $O(N)$ računskih operacija. Broj operacija za izračunavanje m odbiraka spektra signala dužine N odbiraka Gercelovim algoritmom je $O(mN)$, dok je složenost FFT-a $O(N \log N)$. Kada je od interesa samo $m \sim \log N$ odbiraka spektra broj operacija u Gercelovom algoritmu je manji od direktne primene FFT-a. Manji broj operacija je posledica toga što se u Gercelovom algoritmu izračunavaju samo odbirci spektra od interesa, dok se primenom FFT-a izračunava svih N odbiraka spektra.

Za razliku od FFT-a, gde je potrebno pohraniti sve odbirke u memoriju kako bi se započela izračunavanja, Gercelovim algoritmom se izračunavanja vrše u hodu. Vrednost odbirka se koristi samo za ažuriranje međurezultata i više nije potrebna, čime se eliminiše potreba za memorijom. Pošto se izračunavanja vrše u hodu, izračunata vrednost je dostupna odmah nakon ažuriranja međurezultata poslednjim odbirkom.

Vrednost k -tog odbirka DFT-a se izračunava iz odbiraka $x[n]$ kao:

$$X[k] = \sum_{r=0}^{N-1} x[r] W_N^{kr}, \quad (\text{D.1})$$

gde je rotacioni faktor (engl. twiddle factor) W_N :

$$W_N = e^{-j2\pi/N}. \quad (\text{D.2})$$

Vrednost izraza (D.1) se ne menja ako se pomnoži sa

$$W_N^{-kN} = e^{j(2\pi/N)kN} = e^{j2\pi k} = 1 \quad (\text{D.3})$$

i može se preurediti u

$$X[k] = W_N^{-kN} \sum_{r=0}^{N-1} x[r] W_N^{kr} = \sum_{r=0}^{N-1} x[r] W_N^{-k(N-r)}. \quad (\text{D.4})$$

Suma se dalje može preurediti u oblik

$$\begin{aligned} X[k] &= x[0] W_N^{-kN} + x[1] W_N^{-k(N-1)} + \dots + x[N-1] W_N^{-k} \\ &= [\dots [x[0] W_N^{-k} + x[1]] W_N^{-k} + x[2]] W_N^{-k} + \dots + x[N-1] W_N^{-k}. \end{aligned} \quad (\text{D.5})$$

Osnovna ideja Gercelovog algoritma je u formiranju pomoćnog niza $y_k[n]$:

$$y_k[n] = W_N^{-k} y_k[n-1] + x[n], \quad (\text{D.6})$$

koji se koristi za iterativno izračunavanje k -tog odbirka spektra $X[k]$. Pod pretpostavkom da je $y_k[0] = 0$ i $x[N] = 0$, vrednost $y_k[N]$ je:

$$\begin{aligned} y_k[N] &= \overbrace{x[N]}^0 + W_N^{-k} y_k[N-1] \\ &= [W_N^{-k} y_k[N-2] + x[N-1]] W_N^{-k} \\ &\vdots \\ &= [\cdots [x[0] W_N^{-k} + x[1]] W_N^{-k} + x[2]] W_N^{-k} + \cdots + x[N-1]] W_N^{-k} \end{aligned} \quad (\text{D.7})$$

Upoređivanjem (D.7) i (D.5) direktno sledi

$$y_k[N] = X[k], \quad (\text{D.8})$$

odnosno vrednost k -tog odbirka spektra je jednaka vrednosti $N+1$ -og člana pomoćnog niza $y_k[n]$. U izrazu (D.6) se koristi samo jedna vrednost rotacionog faktora W_N^{-k} za bilo koji broj odbiraka N , dok se u izrazu za direktno izračunavanje k -tog odbirka DFT-a (D.1) koristi N konstanti koje je potrebno čuvati u memoriji.

Pomoćni niz (D.6) bi se u principu mogao koristiti za izračunavanje k -tog odbirka spektra, ali se u praksi ne koristi u tom obliku, iz razloga koji će postati jasni kada se izračunaju pozicije polova prenosne funkcije. Primenom z transformacije na (D.6) dobijamo

$$Y(z) = z^{-1} W_N^{-k} + X(z) \quad (\text{D.9})$$

odakle je prenosna funkcija

$$H_k(z) = \frac{Y(z)}{X(z)} = \frac{1}{1 - W_N^{-k} z^{-1}} = \frac{1}{1 - e^{j2\pi k/N} z^{-1}}. \quad (\text{D.10})$$

Prenosna funkcija $H_k(z)$ ima jedan pol

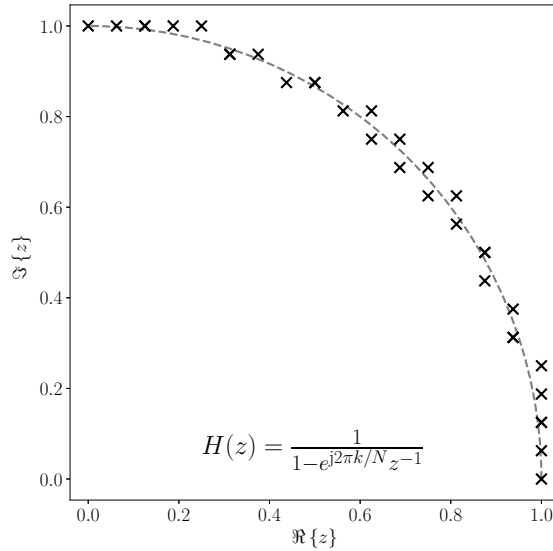
$$z_k = W_N^{-k} = e^{j2\pi k/N} = \cos(2\pi k/N) + j \sin(2\pi k/N), \quad (\text{D.11})$$

koji leži na jediničnom krugu $|z_k| = 1$ i ima istu brojčanu vrednost kao koeficijent u izrazu za pomoćni niz (D.6).

Jednakost $y_k[N] = X[k]$ pokazuje da $N+1$ -vi član niza $y_k[N]$ ima istu vrednost kao k -ti odbirak spektra izračunatog DFT-om, ali to ne znači da prenosna funkcija (D.10) ima istu frekvencijsku karakteristiku kao DFT bin [75]. Frekvencijski odziv DFT bina je oblika $\sin(x)/x$, dok (D.10) sadrži pol na jediničnom krugu $|z| = 1$ koji se ponaša kao kompleksni rezonator sa veoma selektivnim frekvencijskim odzivom.

Razlog zbog kojeg se (D.6) ne koristi u praksi je u tome što pol prenosne funkcije z_k može biti van jediničnog kruga $|z_k| > 1$ kada se koeficijent W_N^{-k} kvantizuje na konačan broj bita. Drugim rečima, izraz (D.6) je numerički nestabilan.

Na slici D.1 prikazane su pozicije polova prenosne funkcije sa koeficijentima kvantizovanim na širinu reči od pet bita. Širina reči od pet bita je izabrana radi ilustracije, jer su efekti kvantizacije koeficijenata izraženi i lako uočljivi. Sa slike se može videti da su u mnogim slučajevima pozicije polova vidno van jediničnog kruga $|z| > 1$, i u tim slučajevima je sistem nestabilan. Kada su polovi unutar jediničnog kruga $|z| < 1$ sistem je stabilan, ali sa prigušenim odzivom.



Slika D.1: Pozicije polova u numerički nestabilnom Gercelovom algoritmu sa kvantizovanim koeficijentima.

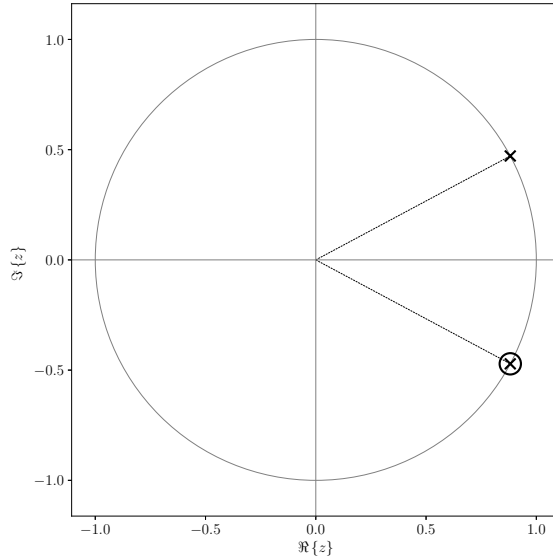
Problem numeričke stabilnosti Gercelovog algoritma se može rešiti na donekle neočekivani način, uvođenjem pola i nule $z = W_N^k$ koji se poništavaju:

$$\begin{aligned}
 H_k(z) &= \frac{1}{1 - W_N^{-k} z^{-1}} \overbrace{\frac{1 - W_N^k z^{-1}}{1 - W_N^k z^{-1}}}^{\text{pol i nula}} \\
 &= \frac{1 - W_N^k z^{-1}}{1 - 2 \cos(2\pi k/N) z^{-1} + z^{-2}}
 \end{aligned} \tag{D.12}$$

Pol prenosne funkcije $z = W_N^{-k}$ i uvedeni pol $z = W_N^k$ čine konjugovano-kompleksni par polova, što za posledicu ima da su svi koeficijenti u imeniocu realni. Pozicije polova i nule su prikazani na slici D.2.

Polovi (D.12) su [76]:

$$z_{1,2} = \frac{2\alpha \pm \sqrt{4\alpha^4 - 4}}{2} = \alpha \pm \sqrt{\alpha^2 - 1}, \tag{D.13}$$



Slika D.2: Pozicije polova i nule prenosne funkcije (D.12).

gde je α

$$\alpha = \cos\left(\frac{2\pi k}{N}\right) \quad (\text{D.14})$$

Pošto je $|\alpha| = |\cos(2\pi k/N)| \leq 1$ član $\sqrt{\alpha^2 - 1}$ se može napisati kao $j\sqrt{1 - \alpha^2}$:

$$z_{1,2} = \alpha \pm j\sqrt{1 - \alpha^2}. \quad (\text{D.15})$$

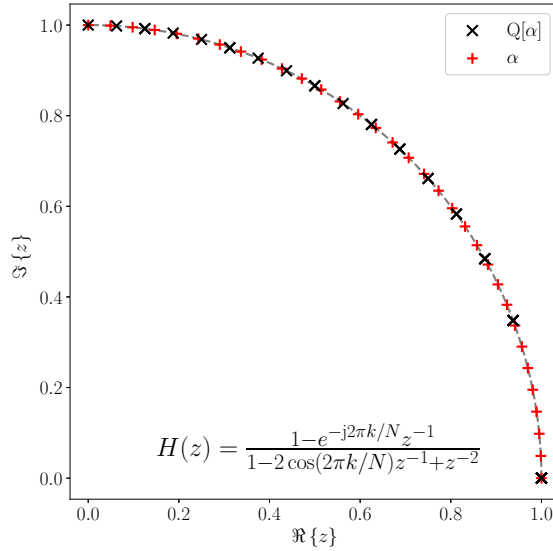
Moduo polova (D.12) je

$$|z_{1,2}| = \sqrt{\Re\{z_{1,2}\}^2 + \Im\{z_{1,2}\}^2} = \sqrt{\alpha^2 + 1 - \alpha^2} = 1. \quad (\text{D.16})$$

Iz (D.16) sledi da su polovi preuređene prenosne funkcije (D.12) uvek na jediničnom krugu, bez obzira na konačnu tačnost $\alpha = Q[\cos(2\pi k/N)]$, sve dok je $|\alpha| \leq 1$. Drugim rečima, prenosna funkcija (D.12) je numerički stabilna. Na slici D.3 su prikazane pozicije polova sa koeficijentom α kvantizovanim na širinu reči od pet bita, sa koje se može videti da svi polovi zaista leže na jediničnom krugu.

Uvođenjem pola i nule u prenosnu funkciju Gercelovog algoritma je obezbeđeno da polovi uvek leže na jediničnom krugu, bez obzira na širinu reči koeficijenta α . Širina reči α međutim određuje moguće pozicije polova na jediničnom krugu. Uporedo sa mogućim pozicijama polova kada je koeficijent α kvantizovan na pet bita, na slici D.3 su simbolom $+$ prikazane pozicije polova izračunate korišćenjem brojeva sa pokretnim zarezom u dvostrukoj preciznosti.

Zbog kvantizacije vrednosti koeficijenta očekivano je da postoji ograničen broj pozicija polova. Međutim, na prvi pogled možda nije očigledno da pozicije polova nisu



Slika D.3: Pozicije polova u numerički stabilnom Gercelovom algoritmu sa kvantizovanim koeficijentima.

ravnomerno raspoređene po jediničnom krugu. Najveći razmak između mogućih pozicija polova sa kvantizovanim koeficijentima je između pola $z = 1$ i prvog susednog, odnosno pola najbližeg realnoj osi.

Neravnomerna raspodela polova za posledicu ima da pozicije polova oko $z = 1$, koje odgovaraju DFT binovima sa malim indeksima, zahtevaju veliku širinu reči koeficijenta α . Pored toga, širina reči zavisi i od broja odbiraka N , pošto su idealne pozicije polova

$$z_k = e^{j2\pi k/N} = \cos\left(\frac{2\pi k}{N}\right) + j \sin\left(\frac{2\pi k}{N}\right) \quad (\text{D.17})$$

Realni deo najbližeg pola realnoj osi je jednak najvećem broju manjem od jedinice za zadatu širinu reči, i za označene brojeve širine reči od m bita iznosi:

$$\Re\{z\} = Q[\alpha] = 1 - 2^{-m+1} \quad (\text{D.18})$$

Iz (D.15) vrednost imaginarnog dela je

$$\Im\{z\} = \sqrt{1 - Q[\alpha]^2} = \sqrt{1 - (1 - 2^{-m+1})^2} \approx 2^{(2-m)/2} \quad (\text{D.19})$$

odakle je najmanji argument pola:

$$\theta_{\min} \approx \arctan\left(\frac{2^{(2-m)/2}}{1 - 2^{1-m}}\right) \stackrel{m \gg 1}{\approx} \frac{2^{(2-m)/2}}{1 - 2^{1-m}} \quad (\text{D.20})$$

Procena potrebne širine reči za izračunavanje k -tog, $k \sim 1$, DFT bina se može dobiti izjednačavanjem (D.20) i argumenta (D.17)

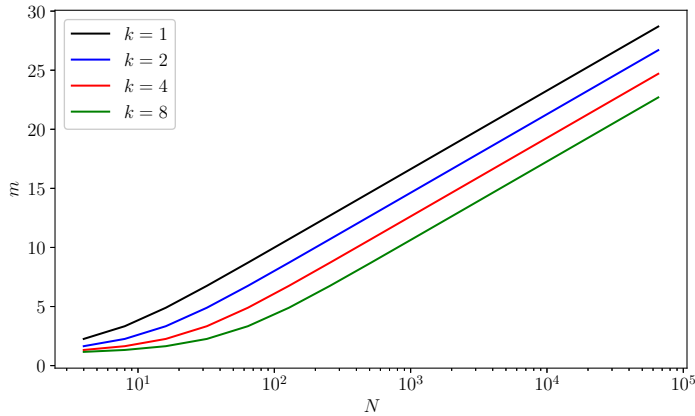
$$\frac{2\pi k}{N} = \frac{2^{(2-m)/2}}{1 - 2^{1-m}} \quad (\text{D.21})$$

i rešavanjem po m

$$m = 2 \log_2 \left(\frac{N + k \sqrt{\frac{N^2}{k^2} + 8\pi^2}}{2\pi k} \right) \quad (\text{D.22})$$

$$\approx 2 [\log_2(N) - \log_2(k) - \log_2(\pi)]$$

Minimalna širina reči u zavisnosti od broja odbiraka N je prikazana na slici D.4.



Slika D.4: Minimalna širina reči m u zavisnosti od broja odbiraka N za pozicioniranje pola na učestanosti k -tog DFT bina.

Prenosna funkcija numerički stabilnog Gercelovog algoritma se može preurediti u kaskadnu vezu IIR i FIR filtra:

$$H_k(z) = \frac{Y_k(z)}{X(z)} = \underbrace{(1 - 2 \cos(2\pi k/N) z^{-1} + z^{-2})^{-1}}_{\text{IIR}} \underbrace{(1 - W_N^k z^{-1})}_{\text{FIR}}, \quad (\text{D.23})$$

odakle sledi

$$Y_k(z) = (1 - W_N^k z^{-1}) \left[\frac{X(z)}{1 - 2 \cos(2\pi k/N) z^{-1} + z^{-2}} \right] \quad (\text{D.24})$$

Izlazni signal se dalje može predstaviti kao:

$$Y_k(z) = (1 - W_N^k z^{-1}) V_k(z), \quad (\text{D.25})$$

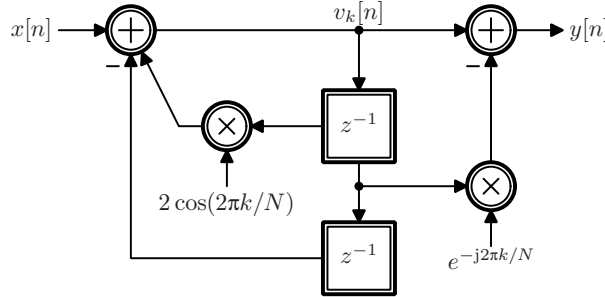
$$y_k[n] = v_k[n] - W_N^k v_k[n-1], \quad (\text{D.26})$$

gde je $V_k(z)$ izlaz IIR filtra

$$V_k(z) = \frac{1}{1 - 2 \cos(2\pi k/N) z^{-1} + z^{-2}} X(z), \quad (\text{D.27})$$

$$v_k[n] = 2 \cos(2\pi k/N) v_k[n-1] - v_k[n-2] + x[n]. \quad (\text{D.28})$$

Izraz (D.26) je potrebno izračunati samo za $n = N$ pošto je od interesa samo vrednost $y_k[N]$. Pri izračunavanju odbirka spektra Gercelovim algoritmom u ovom obliku potrebno je izračunati $v_k[n]$ (D.28) za N odbiraka ulaznog signala $x[0], \dots, x[N-1]$ i $N+1$ -vi odbirak $x[N] = 0$. Realizacija Gercelovog algoritma (D.26) je prikazana na slici D.5.



Slika D.5: Realizacija Gercelovog algoritma (D.26).

Izloženim algoritmom je za izračunavanje odbirka spektra signala potrebno dodati nulti odbirak posle poslednjeg odbirka signala. Umetanje nultog odbirka kao $N+1$ -og odbirka signala može predstavljati problem kada se Gercelov algoritam koristi za izračunavanje odbirka spektra u realnom vremenu. U tom slučaju je potrebno preurediti izraz za $y_k[N]$ kako bi se izbeglo umetanje nultog odbirka.

Kao što je već rečeno, od interesa je samo vrednost $y_k[N]$:

$$y_k[N] = v_k[N] - W_N^k v[N-1]. \quad (\text{D.29})$$

Pošto je $x[N] = 0$ imamo

$$\begin{aligned} v_k[N] &= 2 \cos(2\pi k/N) v_k[N-1] - v_k[N-2] + x[N] \\ &= 2 \cos(2\pi k/N) v_k[N-1] - v_k[N-2] \end{aligned} \quad (\text{D.30})$$

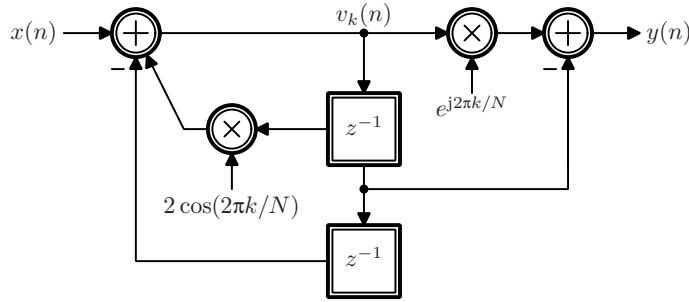
Zamenom (D.30) u (D.29) sledi

$$\begin{aligned} y_k[N] &= (2 \cos(2\pi k/N) - W_N^k) v_k[N-1] - v_k[N-2] \\ &= (\cos(2\pi k/N) + j \sin(2\pi k/N)) v_k[N-1] - v_k[N-2]. \\ &= W_N^{-k} v_k[N-1] - v_k[N-2] \end{aligned} \quad (\text{D.31})$$

Iz poslednjeg rezultata sledi da je vrednost $y_k[N]$ moguće izračunati iz $v_k[N-1]$ i $v_k[N-2]$, odnosno nije potrebno dodavati nulti odbirak. Na taj način je moguće izračunati vrednost odbirka spektra signala dužine N odbiraka u N umesto u $N+1$ taktova. Realizacija Gercelovog algoritma (D.31) je prikazana na slici D.6.

Kompleksnost izračunavanja jednog odbirka spektra DFT-om

$$X[k] = \sum_{r=0}^{N-1} x[r] W_N^{kr},$$



Slika D.6: Realizacija Gercelovog algoritma (D.31).

i Gercelovim algoritmom

$$\begin{aligned} v_k[n] &= 2 \cos(2\pi k/N) v_k[n-1] - v_k[n-2] + x[n] \\ X[k] &= W_N^{-k} v_k[N-1] - v_k[N-2] \end{aligned}$$

je $O(N)$, i oba algoritma se mogu izračunati u hodu, bez potrebe za čuvanjem svih odbiraka. Međutim, iako je u oba slučaja broj računskih operacija proporcionalan sa N , Gercelovim algoritmom se postiž značajna ušteda.

U slučaju DFT-a izračunava se suma N proizvoda odbiraka signala $x[r]$ i kompleksnih koeficijenata W_N^{kr} , za koju je potrebno izvršiti $2N$ množenja i sabiranja kada je $x[r]$ realan signal. Koeficijent $2 \cos(2\pi k/N)$ u Gercelovom algoritmu je realan i potrebno je izvršiti samo N množenja i $2N$ sabiranja za izračunavanje $v_k[N-1]$ i $v_k[N-2]$, i još 2 množenja i jedno sabiranje za izračunavanje $X[k]$ iz njih.

Ukupan broj operacija u Gercelovom algoritmu je gotovo dvostuko manji od direktnog izračunavanja odbirka spektra DFT-om. Pored toga, u Gercelovom algoritmu se koriste vrednosti $\cos(2\pi k/N)$ i $\sin(2\pi k/N)$ za zadato k i N , koje se mogu implementirati kao množači konstantnom, dok se u DFT-u koriste vrednosti za $k = 0, \dots, N-1$ i zadato N , i moraju se koristiti množači.

Primer D.1: Izračunavanje odbirka spektra Gercelovim algoritmom

Izračunavanje odbirka spektra Gercelovim algoritmom će biti pokazano na nizu prvih osam prostih brojeva:

$$x = (2, 3, 5, 7, 11, 13, 17, 19)$$

Prvi odbirak spektra signala $x(n)$ izračunat diskretnom Furijeovom transformacijom ima vrednost $X(1) = -7.5857 + 27.5564j$.

Diferencna jednačina (D.28) za $k = 1$ i $N = 8$ je

$$v_1[n] = 2 \cos(\pi/4) v_1[n-1] - v_1[n-2] = \sqrt{2} v_1[n-1] - v_1[n-2]$$

Koraci u izračunavanju $X[1]$ Gercelovim algoritmom su dati u tabeli.

n	$x[n]$	$v[n]$	$v[n-1]$	$v[n-2]$
0	2.0000	2.0000	0.0000	0.0000
1	3.0000	5.8284	2.0000	0.0000
2	5.0000	11.2426	5.8284	2.0000
3	7.0000	17.0711	11.2426	5.8284
4	11.0000	23.8995	17.0711	11.2426
5	13.0000	29.7279	23.8995	17.0711
6	17.0000	35.1421	29.7279	23.8995
7	19.0000	38.9706	35.1421	29.7279
8	0.0000	19.9706	38.9706	35.1421

Za izračunavanje odbirka spektra korišćenjem (D.29) potrebno je produžiti signal nultim odbirkom, što se vidi iz poslednjeg reda tabele. Vrednost $X[1]$ izračunata Gercelovim algoritmom je

$$\begin{aligned} y_1[N] &= v_1[N] - W_8 v_1[N-1] = 19.9706 - e^{-j\pi/4} 38.9706 \\ &= -7.5857 + 27.5564j, \end{aligned}$$

i jednaka je vrednosti izračunatoj DFT-om.

Vrednost $X[1]$ se može izračunati Gercelovim algoritmom i bez dodavanja nultog odbirka korišćenjem (D.31) i iznosi:

$$\begin{aligned} y_1[N] &= W_N^{-k} v_k[N-1] - v_k[N-2] = e^{j\pi/4} 38.9706 - 35.1421 \\ &= -7.5857 + 27.5564j, \end{aligned}$$

što je takođe tačan rezultat.

Gercelov algoritam se može iskoristiti i za efikasno izračunavanje odbiraka spektra $X[k]$ i $X[N-k]$. Odbirci spektra $X[k]$ i $X[N-k]$ su na učestanostima $\Omega = \pm\pi k/N$, i u slučaju kada je ulazni signal realan važi $X[k] = X[N-k]^*$ usled konjugovano-kompleksne simetrije spektra. Međutim, kada je ulazni signal kompleksan, spektar nije konjugovano-kompleksno simetričan i odbirci spektra $X[k]$ i $X[N-k]$ se moraju pojedinačno izračunati DFT-om.

Ukoliko primetimo da je

$$e^{j2\pi(N-k)/N} = e^{-j2\pi k/N} \quad (\text{D.32})$$

$$e^{\pm j2\pi k/N} = \cos\left(\frac{2\pi k}{N}\right) \pm j \sin\left(\frac{2\pi k}{N}\right) \quad (\text{D.33})$$

možemo zaključiti da važi $v_k[n] = v_{N-k}[n]$, odnosno da su pri izračunavanju $X[k]$ i $X[N-k]$ vrednosti na izlazu IIR filtra iste. Dovoljno je izračunati vrednosti $v_k[n]$ i

odbirci spektra $X[k]$ i $X[N - k]$ se mogu izračunati iz $v_k[N - 1]$ i $v_k[N - 2]$ kao:

$$X[k] = W_N^{-k} v_k[N - 1] - v_k[N - 2], \quad X[N - k] = W_N^k v_k[N - 1] - v_k[N - 2]. \quad (\text{D.34})$$

Broj operacija za izračunavanje oba odbirka spektra $X[k]$ i $X[N - k]$ je za jedno kompleksno množenje i sabiranje veći od slučaja kada se izračunava samo $X[k]$. Izračunavanje oba odbirka spektra DFT-om za slučaj kompleksnog izlaznog signala bi zahtevalo dvostruko veći broj operacija u odnosu na slučaj kada se izračunava vrednost samo jednog odbirka.

Primenom Gercelovog algoritma se smanjuje broj računskih operacija u odnosu na direktno izračunavanje odbirka spektra DFT-om, i eliminiše se potreba za čuvanjem N vrednosti rotacionih faktora. Međutim, iterativna priroda Gercelovog algoritma rezultuje većom greškom izračunavanja u odnosu na DFT ili FFT. U literaturi [77] je pokazano da srednja vrednost greške Gercelovog algoritma raste kao $O(N^2)$, gde je N broj odbiraka, dok greške izračunavanja DFT i FFT rastu sa $O(\sqrt{N})$ i $O(\sqrt{\log(N)})$, respektivno. Značajno brži porast greške izračunavanja Gercelovim algoritmom ograničava njegovu primenu za veliki broj odbiraka. Maksimalni broj odbiraka zavisi od prihvatljive greške izračunavanja i širine reči kojima su predstavljeni koeficijenti i međurezultati. Greška izračunavanja Gercelovog algoritma je najveća za DFT binove sa malim indeksima, kao posledica neravnomerne raspodele polova na jediničnom krugu usled kvantizacije koeficijenata.