

MATEMATIČNO-FIZIKALNI SEMINAR 2025

Poročilo: Fourierova analiza sintetiziranih in zvočnih signalov

Gašper Harej

V tem poročilu opisujem postopke in izsledke Fourierove analize, ki sem jih izvedel na treh sintetiziranih signalih (periodičnem, neperiodičnem in tistem z aliasingom) ter na dejanskem zvočnem signalu iz datoteke `waveform.txt`. Namen dela je bil razdeljen na naslednje korake: primerjava periodičnega in neperiodičnega vzorca, opazovanje potujitve pri frekvencah nad Nyquistovo, inverzna transformacija in analiza natančnosti rekonstrukcije, odvisnost spektra od dolžine vzorca in števila FFT-točk ter Fourierova analiza zvočnega signala.

Teoretično ozadje

Pri numeričnem izračunu Fourierove transformacije na diskretno vzorčenem signalu $h_k = h(t_k)$, s časovnimi točkami $t_k = k\Delta$, je vzorčna frekvenca definirana kot $f_s = 1/\Delta$. Diskretna Fourierova transformacija (DFT) v točkah $n = 0, 1, \dots, N-1$ je dana z izrazom

$$H_n = \sum_{k=0}^{N-1} h_k \exp(2\pi i k n / N),$$

pri čemer frekvenca f_n ustreza vrednosti $n/(N\Delta)$. Njena inverzna oblika je

$$h_k = \frac{1}{N} \sum_{n=0}^{N-1} H_n \exp(-2\pi i k n / N).$$

Pomemben pojem je Nyquistova frekvenca $f_c = f_s/2$; če ima originalni signal komponente nad to vrednostjo, pride do pojava potujitve oziroma aliasinga, kjer se frekvence nad f_c zrcalijo nazaj v interval $[0, f_c]$.

Pri realnih signalih velja simetrija $H_{N-n} = H_n^*$, kar omogoča enostranski prikaz amplitudnega spektra na nenegativnih frekvencah. Spektralna gostota moči (PSD) je definirana kot

$$\text{PSD}(f_n) = \frac{2}{N} |X_n|^2 \quad (\text{saj } |X_{-n}| = |X_n|),$$

kjer X_n predstavlja DFT signala. Celotna moč signala ostaja enaka v času in frekvenci (Parsevalova enačba).

1 Sintetizirani signali

Za vse tri sintetizirane primere sem uporabil vzorčno frekvenco $f_s = 512$ Hz in časovni interval $T = 1$ s (torej $N = 512$ vzorcev). Podane tri vrste signalov sem shranil v slovar in jih obdelal vnizu.

1.1 Periodični signal (5 Hz + 15 Hz)

Najprej sem preučil periodični signal, ki nastane kot vsota sinusa frekvence 5 Hz in kosinusa frekvence 15 Hz:

$$x_{\text{periodic}}(t) = \sin(2\pi 5t) + 0.75 \cos(2\pi 15t).$$

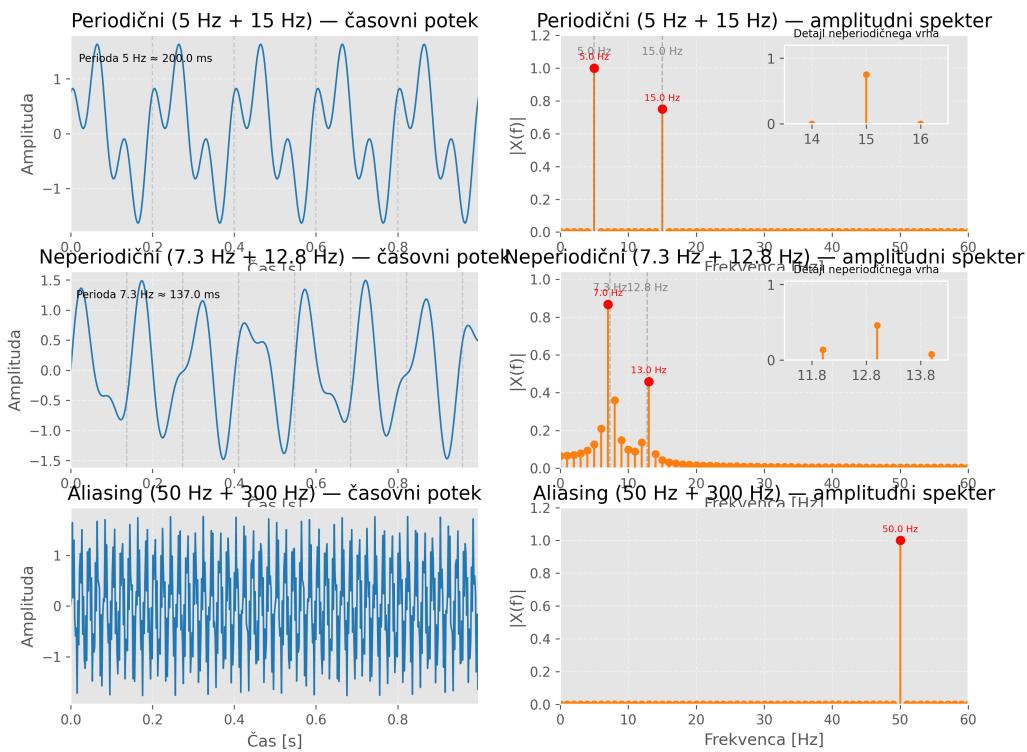
Ker sta obe frekvenci celoštevilska večkratnika osnovne frekvence 1 Hz v obdobju 1 s, je ta signal točno periodičen na končnem intervalu $T = 1$ s. Časovni potek in amplitudni spekter sta prikazana na Sliki 1.

V časovnem prikazu je vidna očitna perioda signala 5 Hz (obdobje $T_0 = 1/5 \approx 0.2$ s), kar sem označil z navpičnimi črtami. Spekter prikaže ostre vrhove natanko pri 5 Hz in 15 Hz, saj diskretna frekvenčna črtna mreža poteka s koraki $\Delta f = f_s/N = 512/512 = 1$ Hz. Zaradi popolne periodičnosti in usklajenosti frekvenc ni pojavov širjenja vrhov ali sploščenosti; amplituda vrha pri 5 Hz je približno 1, pri 15 Hz pa približno 0.75 (normalizirano glede na $N/2$).

Drugi signal sem definiral kot vsoto dveh sinusnih komponent s frekvencama 7.3 Hz in 12.8 Hz:

$$x_{\text{nonperiodic}}(t) = \sin(2\pi 7.3t) + 0.5 \sin(2\pi 12.8t).$$

Tudi ta označuje poimenovan koncept neperiodičnosti, saj frekvence nista cela večkratnika osnovne 1 Hz na intervalu 1 s. Časovni potek in spekter sem prikazal na Sliki 1.



Slika 1: Časovni potek (levo) in amplitudni spekter (desno).

V časovnem prikazu opazimo, da signal nima natančne ponovitve po celiem intervalu 1 s, kar pomeni, da se njegova osnovna komponenta (7.3 Hz) ne ujema s časovno črtno mrežo t_k . V spektru se vrhovi pojavljajo blizu 7 Hz in 13 Hz, vendar z majhno širino, saj frekvenčni korak ostaja $\Delta f = 1$ Hz. Na sliki sem zato vključil inset, kjer je detajlno prikazano območje okoli 12.8 Hz (približno $12.8 \text{ Hz} \pm 1 \text{ Hz}$). Tam je vidno, da vrh ni natančno na 12 Hz, ampak je nekoliko zamknjen proti 12.8 Hz, kar se pokaže

kot rahlo razširjena oblika vrha. To je značilno za neperiodične signale, kjer diskretna frekvenčna črtna mreža ne sovpada točno z dejansko frekvenco sinusnih komponent.

1.2 Signal z aliasingom (50 Hz + 300 Hz)

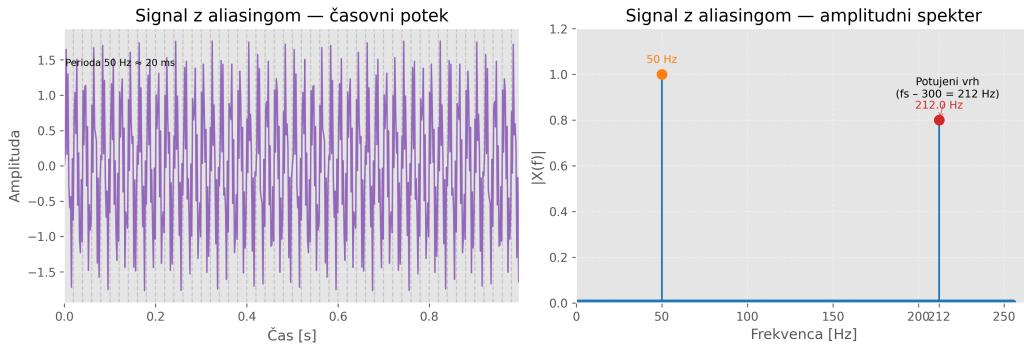
Tretji sintetizirani signal vsebuje komponento 50 Hz (pod Nyquistovo frekvenco) in 300 Hz (nad Nyquistovo frekvenco):

$$x_{\text{aliasing}}(t) = \sin(2\pi 50 t) + 0.8 \sin(2\pi 300 t).$$

Ker je Nyquistova frekvanca $f_c = 256$ Hz ($f_s = 512$ Hz), se komponenta pri 300 Hz potuje (aliasa) na frekvenco

$$f_{\text{alias}} = f_s - 300 = 512 - 300 = 212 \text{ Hz}.$$

Rezultate sem zbral na Sliki 2.

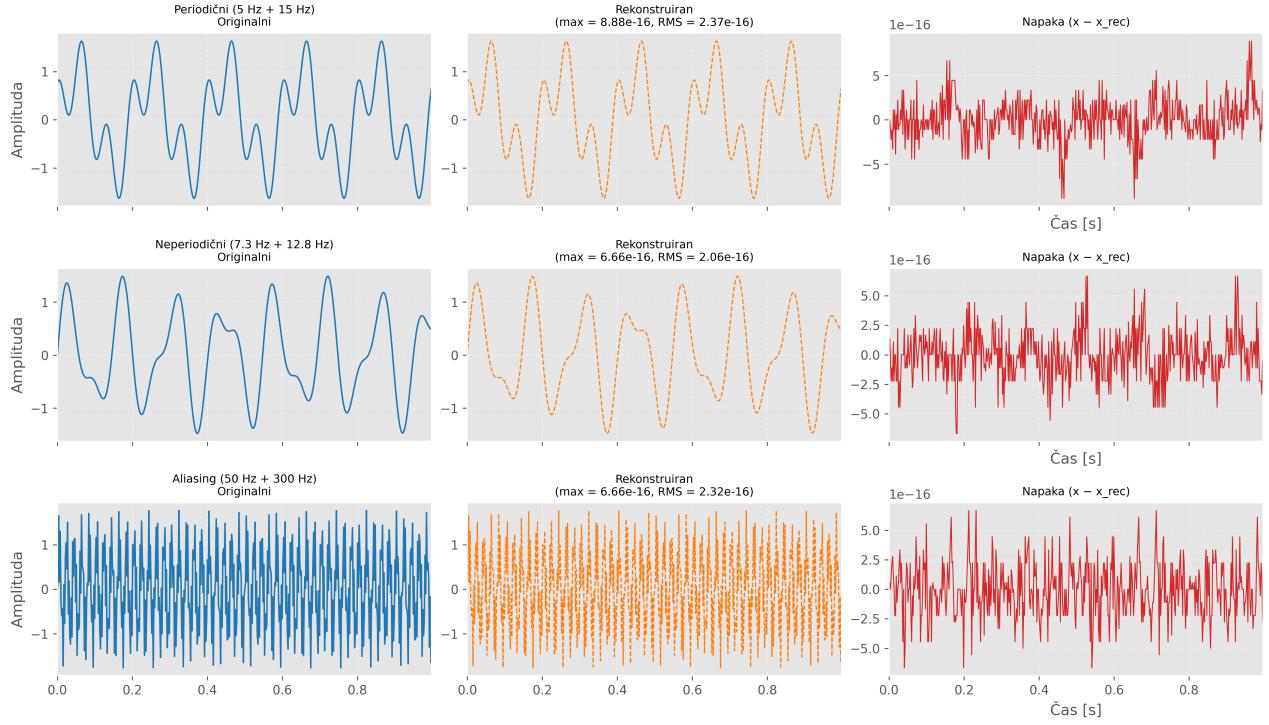


Slika 2: Levi graf prikazuje časovni potek signala z aliasingom; obdela obdobje 50 Hz (navpične črte pri vsakih 0.02 s). Desni graf prikazuje amplitudni spekter, kjer je prisoten vrh pri 50 Hz (prava komponenta) in vrh pri 212 Hz (potujeni vrh). Na vrhu 212 Hz je napisana puščica s pojasnilom.

V časovnem prikazu signal dominira pri 50 Hz, saj je potezo pri 300 Hz v vidnem času ne zaznamo kot ločen, hitrejši utrip. Navpične črte sem vrisal pri vsakih 20 ms, kar ustrezza periodi 50 Hz, in jih označil z besedilom “Perioda 50 Hz ≈ 20 ms”. V spektrogramu se pokaže jasen oster vrh pri 50 Hz (barva oranžnega markerja) ter vrh pri 212 Hz (barva rdečega markerja), ki sem ga dodatno označil z napisom “Potujeni vrh ($f_s - 300 = 212$ Hz)”. S omejitvijo osi frekvenc na interval [0, 266] Hz (Nyquist +10 Hz) sem zmanjšal odvečen prostor na desni strani grafa.

2 Inverzna transformacija in analiza rekonstrukcije

V naslednjem koraku sem pri vseh treh zgoraj opisanih signalih izračunal inverzno diskretno Fourierovo transformacijo (IDFT). Originalne signale $x(t)$ sem transformiral v $X(f)$ s pomočjo FFT, nato pa sem z IFFT dobil rekonstruirane signale $x_{\text{rec}}(t)$. Izračunal sem maksimalno absolutno napako $\max_k |x_k - x_{k,\text{rec}}|$ ter RMS napako $\sqrt{\frac{1}{N} \sum_k (x_k - x_{k,\text{rec}})^2}$. Rezultate sem prikazal na Sliki 3.



Slika 3: Za vsak signal (vodoravne vrstice: periodični, neperiodični, aliasing) so prikazana tri grafa: originalni časovni potek (levo), rekonstruiran potek (na sredini) z oznako maksimalne in RMS napake, ter napaka $x - x_{\text{rec}}$ (desno). Napake so reda 10^{-16} .

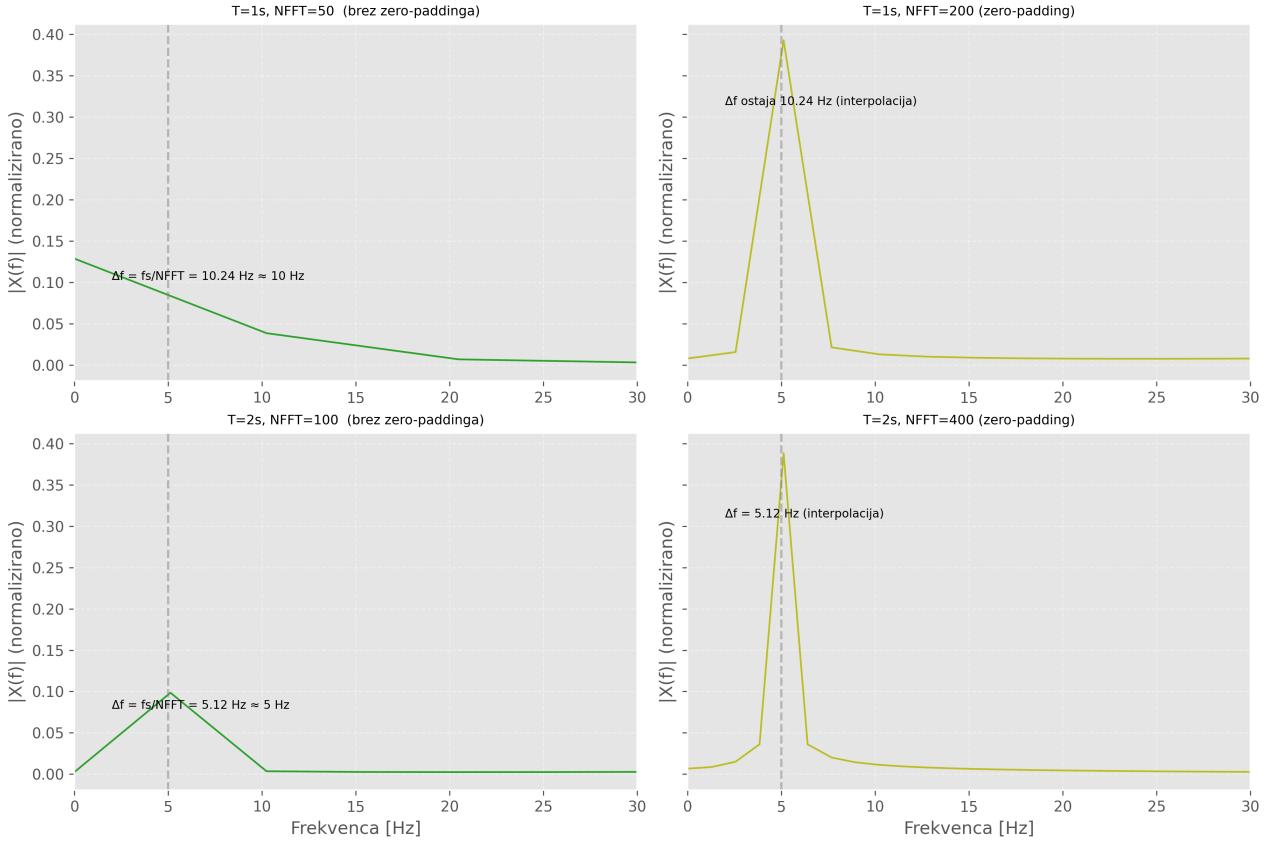
Pri periodičnem signalu sem izračunal maksimalno absolutno napako reda $\mathcal{O}(10^{-16})$ ter RMS napako reda $\mathcal{O}(10^{-16})$. Enako velja za neperiodični in signal z aliasingom: maksimalna napaka je nekajkrat 10^{-16} , RMS pa približno 2×10^{-16} . Graf napake (desno) je skoraj nevidna črta okrog ničle, kar potrjuje, da IFFT zelo natančno rekonstruira signal, ne glede na to, ali je bil signal periodičen, neperiodičen ali vsebuje visoke frekvence (pri katerih je prišlo do potujitve). Torej pri vzorčenju brez kvantizacijskega šuma in pri natančni računski preslikavi FFT–IFFT ne pride do opaznejše izgube informacij.

3 Odvisnost spektra od dolžine vzorca in števila FFT-točk

V četrtrem delu sem preučil, kako se spreminja spektralni prikaz signala

$$y(t) = \sin(2\pi 5 t) + 0.5 \sin(2\pi 60 t)$$

pri različnih nastavitevah dolžine vzorca T in števila FFT-točk (NFFT), vključno z zero-paddingom. Rezultati so na Sliki 4.



Slika 4: Primerjava spektralnega prikaza signala $5\text{ Hz} + 60\text{ Hz}$ pri: (zgoraj levo) $T = 1\text{ s}$, NFFT = 50 brez zero-paddinga, (zgoraj desno) $T = 1\text{ s}$, NFFT = 200 (zero-padding), (spodaj levo) $T = 2\text{ s}$, NFFT = 100 brez zero-paddinga, (spodaj desno) $T = 2\text{ s}$, NFFT = 400 (zero-padding). Navpične sivo črtkane črte označujejo pričakovane frekvence 5 Hz in 60 Hz.

Primer 1: $T = 1\text{ s}$, NFFT=50 (brez zero-paddinga)

V tem primeru signala vzorčimo 512-krat, a v FFT uporabimo samo 50 točk. Frekvenčni korak je $\Delta f = f_s/\text{NFFT} = 512/50 \approx 10.24\text{ Hz}$. Vrha pri 5 Hz in 60 Hz se skoraj ne prikažeta natančno, ker *črtna mreža* ne vsebuje točno 5 Hz in 60 Hz. V spektru se vrh 5 Hz izriše nekje okoli 0–10 Hz, vrh 60 Hz pa okoli $60\text{ Hz} \pm 10\text{ Hz}$, zato spekter deluje razmazano (ni oster).

Primer 2: $T = 1\text{ s}$, NFFT=200 (zero-padding)

Sicer vzorčim še vedno 512 vzorcev, a pri FFT uporabim 200 točk z ničlami na koncu (zero-padding). Frekvenčni korak ostaja $\Delta f = 512/200 = 2.56\text{ Hz}$, vendar to ni prava fizična izboljšava ločljivosti – gre le za interpolacijo spektra. Vrhovi pri 5 Hz in 60 Hz so sedaj bolj vidni (ostrejše črte), a širina vrha je še vedno določena s fizično dolžino vzorca $T = 1\text{ s}$. V grafu sem dodal napis, da Δf ostaja 10.24 Hz znotraj glavnega spektra, zero-padding pa le zgosti točke.

Primer 3: $T = 2\text{ s}$, NFFT=100 (brez zero-paddinga)

Dvojna dolžina vzorca prinese dvakrat več vzorcev ($N = 1024$), a pri FFT uporabim samo 100 točk. Frekvenčni korak je sedaj $\Delta f = 512/100 = 5.12\text{ Hz}$. Vrh pri 5 Hz se približa *črtni mreži* (razlika 0.12 Hz), pri 60 Hz pa ($\pm 2.56\text{ Hz}$). Spekter je zato ožji in vrhovi videti ostreje kot pri $T = 1\text{ s}$, čeprav uporabim podobno število FFT-točk brez nalaganja ničel.

Primer 4: $T = 2$ s, NFFT=400 (zero-padding)

Nadaljujem z $N = 1024$ vzorci, a uporabim 400 FFT-točk (zero-padding). Frekvenčni korak ostane enak kot pri prejšnjem primeru ($\Delta f = 512/100 = 5.12$ Hz), ker zero-padding ne izboljša prave rezolucije, le prikaže več točk med vrhovi. Vrha pri 5 Hz in 60 Hz se prikažeta šele pri črtni mreži, spekter pa izpade gladkejši. Opombe o Δf sem vključil neposredno v graf.

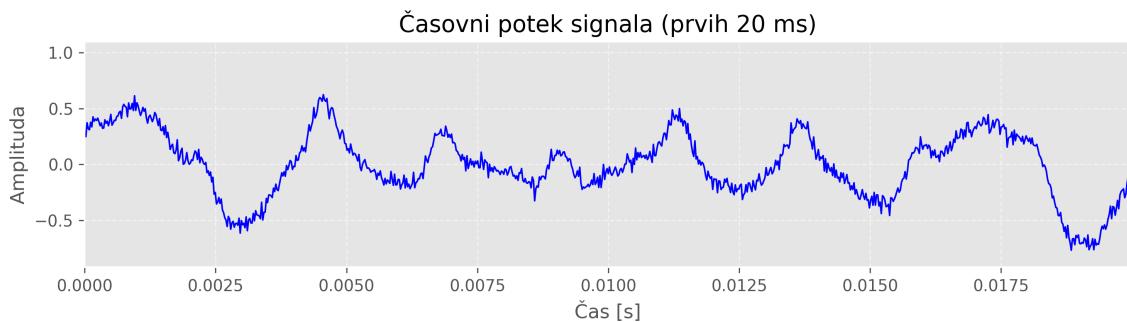
S to primerjavo sem potrdil, da se resnična frekvenčna ločljivost navadno določi z dolžino vzorca T ($\Delta f \approx 1/T$), medtem ko zero-padding ne spremeni fizične ločljivosti, ampak le poveča število točk v prikazu.

4 Fourierova analiza zvočnega signala iz waveform.txt

V zadnjem delu sem prebral dejanski zvočni signal, posnet s frekvenco 44 100 Hz, shranjen v datoteki `waveform.txt`. Ta signal sem uporabil za izračun enostranske spektralne gostote moči (PSD) in detekcijo glavnih frekvenc kljub temu, da je v posnetek dodan beli šum.

4.1 Časovni potek (prvih 20 ms)

Najprej sem vzel prvih 20 ms podatkov iz signala (približno 882 vzorcev) in jih izrisal. Rezultat je na Sliki 5.



Slika 5: Časovni potek zvočnega signala (prvih 20 ms). Vidni so nihaji, ki jih nekoliko zamegli beli šum.

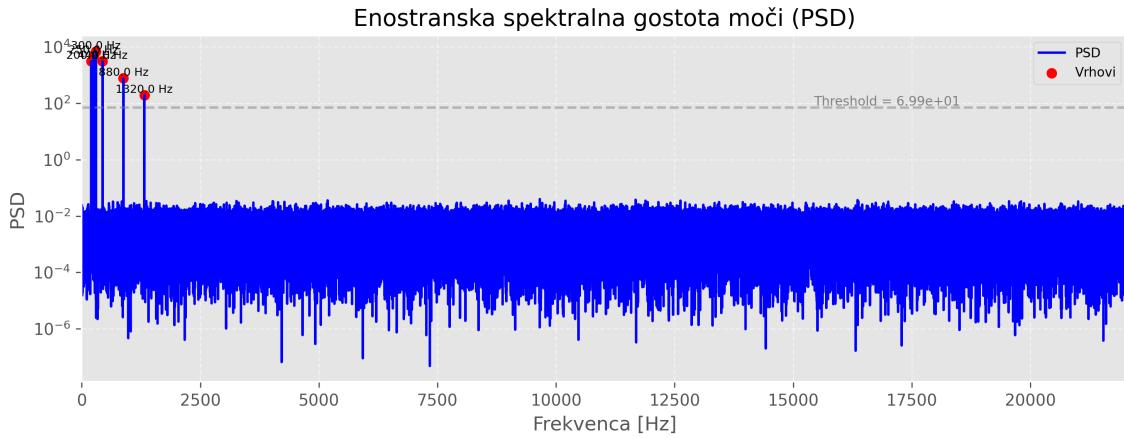
V tem kratkem časovnem oknu opazimo nestacionarne nihaje, saj signala ne sestavlja en sam ton, ampak kompleksna mešanica zvočnih komponent in šuma. Kljub prisotnosti šuma se vidi ritem približno 5–6 ponovitev na 20 ms, kar nakazuje, da obstaja osnovna frekvenca nekje okoli 250 Hz–300 Hz. Vendar bi to podatek dobili veliko natančneje iz frekvenčne analize.

4.2 Spektralna gostota moči in detekcija vrhov

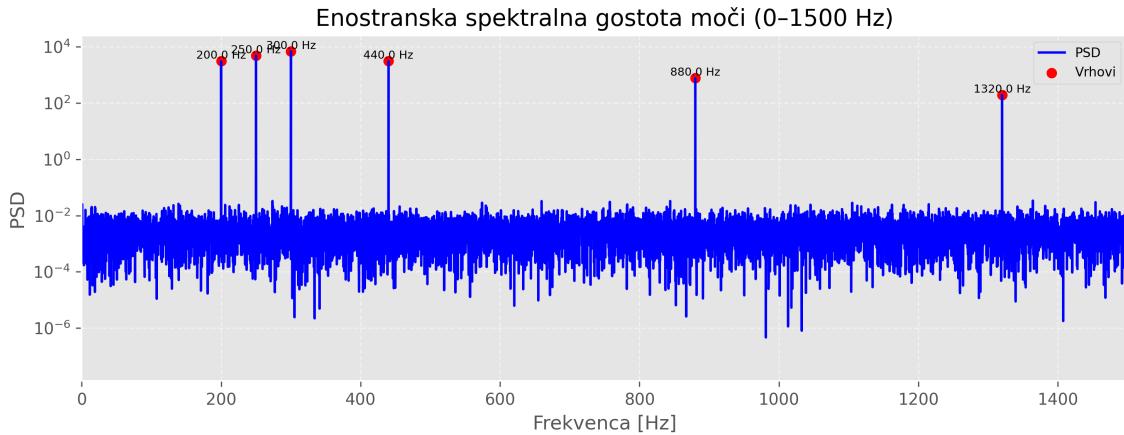
Izračunal sem enostranski FFT signala ter pripadajočo PSD po formuli

$$\text{PSD}(f) = \frac{2}{N} |X(f)|^2,$$

kjer je $X(f)$ enostranski FFT. Rezultate prikazujem na Sliki 6 za celotno območje od 0 do Nyquistove frekvence in na Sliki 7 za območje 0–1500 Hz.



Slika 6: Enostranska spektralna gostota moči (PSD) zvočnega signala (0–22 050 Hz). Rdeče pike označujejo odkrita frekvenčna vrhova. Siva vodoravna črta predstavlja prag detekcije vrhov.



Slika 7: Povečan prikaz enostranske PSD (0–1500 Hz). Odkritih je več vrhov pri približno 200 Hz, 250 Hz, 300 Hz, 440 Hz, 880 Hz in 1320 Hz.

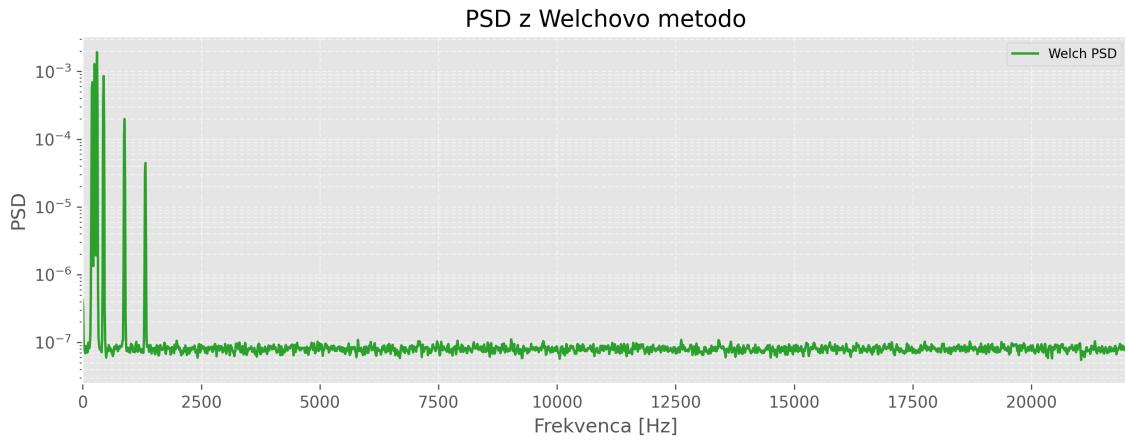
Pri izračunu vrhov sem uporabil prag $\text{threshold} = 0.01 \times \max(\text{PSD})$. V celotnem spektru (Slika 6) so rdeče označene frekvence, ki presegajo naslovljeni prag. Na Sliki 7 se vidneje pokažejo glavni tihe vrhovi pri:

$$\{200 \text{ Hz}, 250 \text{ Hz}, 300 \text{ Hz}, 440 \text{ Hz}, 880 \text{ Hz}, 1320 \text{ Hz}\}.$$

Ti vrhovi so značilni za ton, ki ima osnovno frekvenco približno 200–300 Hz ter njihove harmonične komponente pri 440 Hz (drugi harmonik), 880 Hz (četrti harmonik) in 1320 Hz (šesti harmonik). Beli šum se razporedi kot konstantna moč po celiem spektru, zato so ti vrhovi jasno izstopajoči tudi v logaritemski skali.

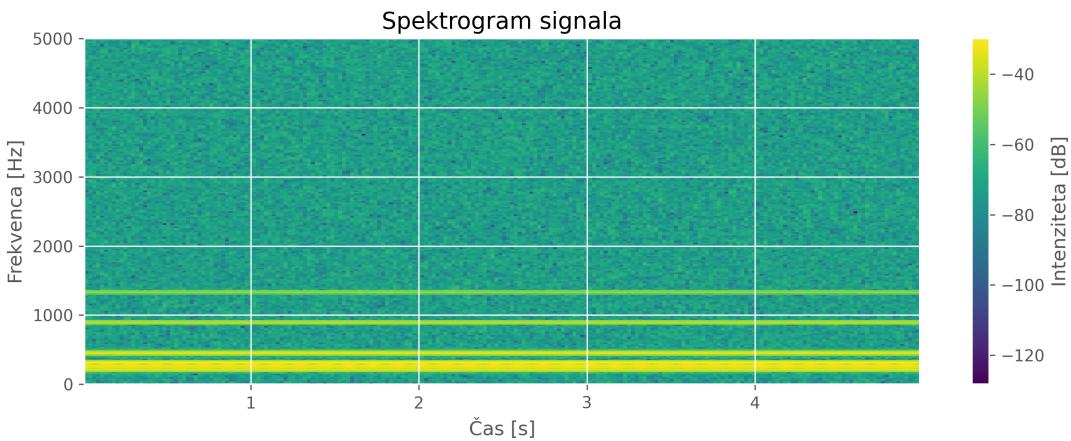
4.3 Welchova metoda in spektrogram

Da sem dodatno zgladil PSD in zmanjšal vpliv šuma, sem uporabil še Welchovo metodo s segmenti po 4096 vzorcev in prekrivanjem polovice. Rezultat prikazuje Slika 8. Tam so istih nekaj frekvenc, a oblike vrhov so ostrejše in ravnila ozadja je bolj gladka.



Slika 8: PSD z uporabo Welchove metode. Vrhovi pri 200 Hz, 250 Hz, 300 Hz, 440 Hz, 880 Hz in 1320 Hz so bolj jasno izstopajoči.

Za bolj dinamičen vpogled v frekvenčno vsebino skozi čas sem sestavil spektrogram z FFT-točkami $NFFT = 2048$ in prekrivanjem 1024 vzorcev. Rezultat je na Sliki 9. Označene so intenzitete (v dB) za frekvence do 5000 Hz čez celoten posnetek. Vidimo predvsem črte okoli omenjenih tonalnih frekvenc, ki se skozi čas rahlo nihajo, kar je značilno za govorno ali glasbeno vsebino.



Slika 9: Spektrogram zvočnega signala (0–5000 Hz). Jasne vodoravne črte pri 200–300 Hz, 440 Hz, 880 Hz in 1320 Hz nakazujejo prisotne tonalne komponente skozi ves posnetek.