

Sprawozdanie LABORATORIUM 3-4

Cyfrowa Technika Foniczna

Temat: Próbkowanie, kwantyzacja i kształtowanie widma szumu rekwantyzacji

Wykonał: Więcek Patryk, Łebkowski Mateusz

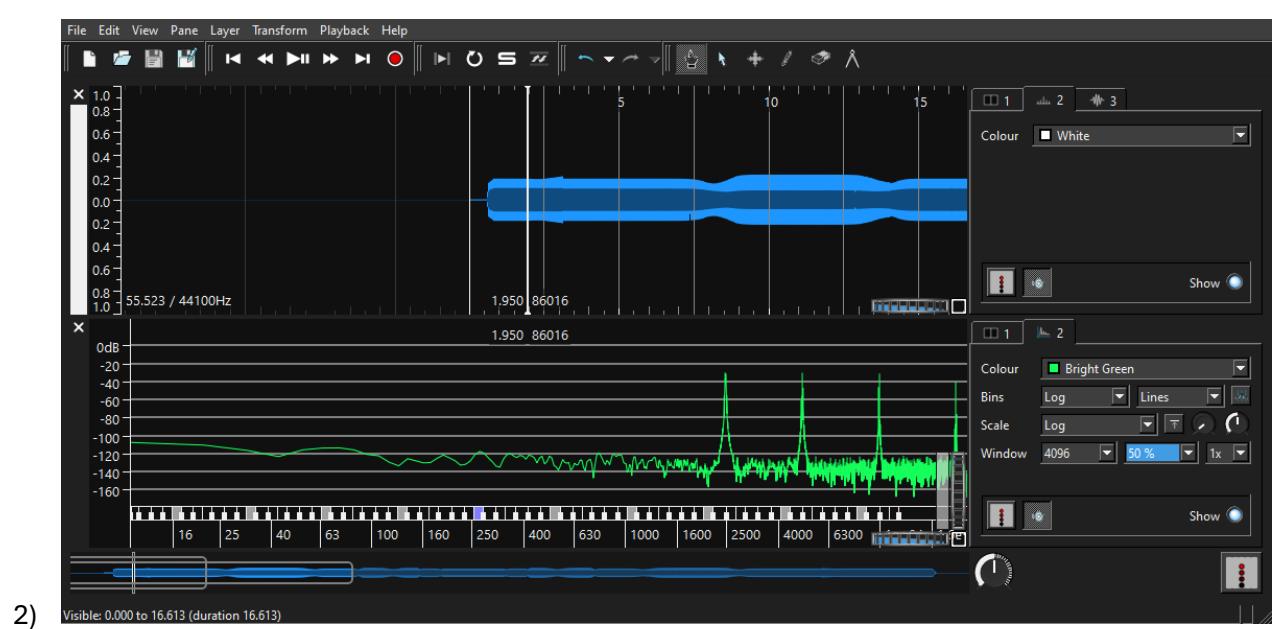
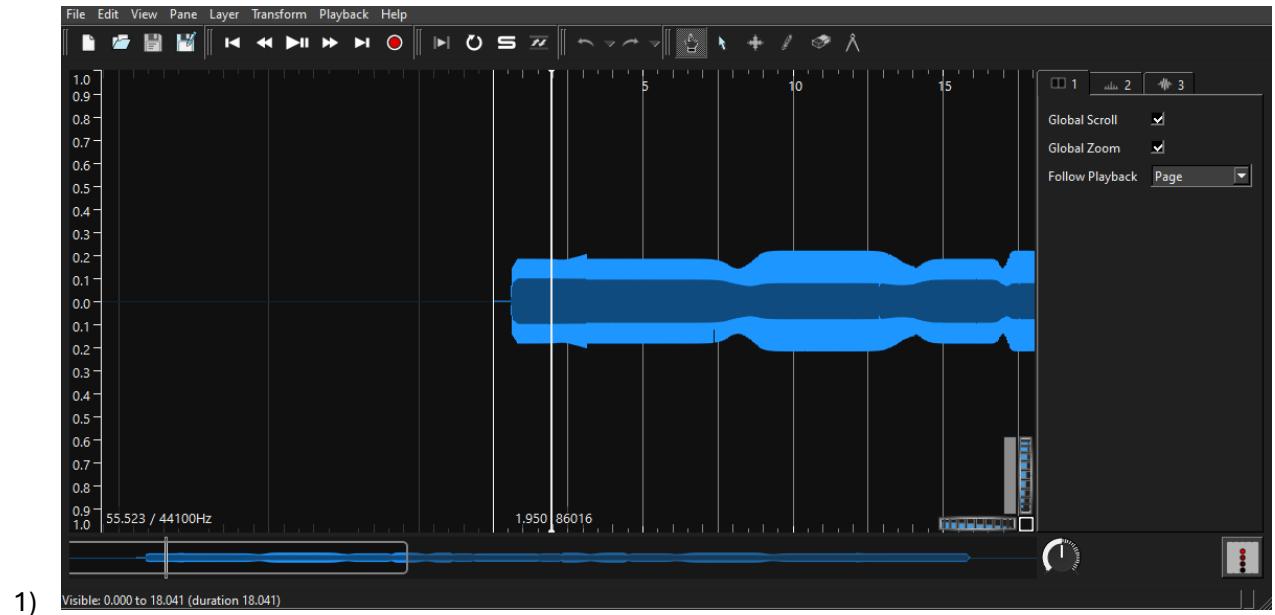
Nr Albumu: 21974, 21969

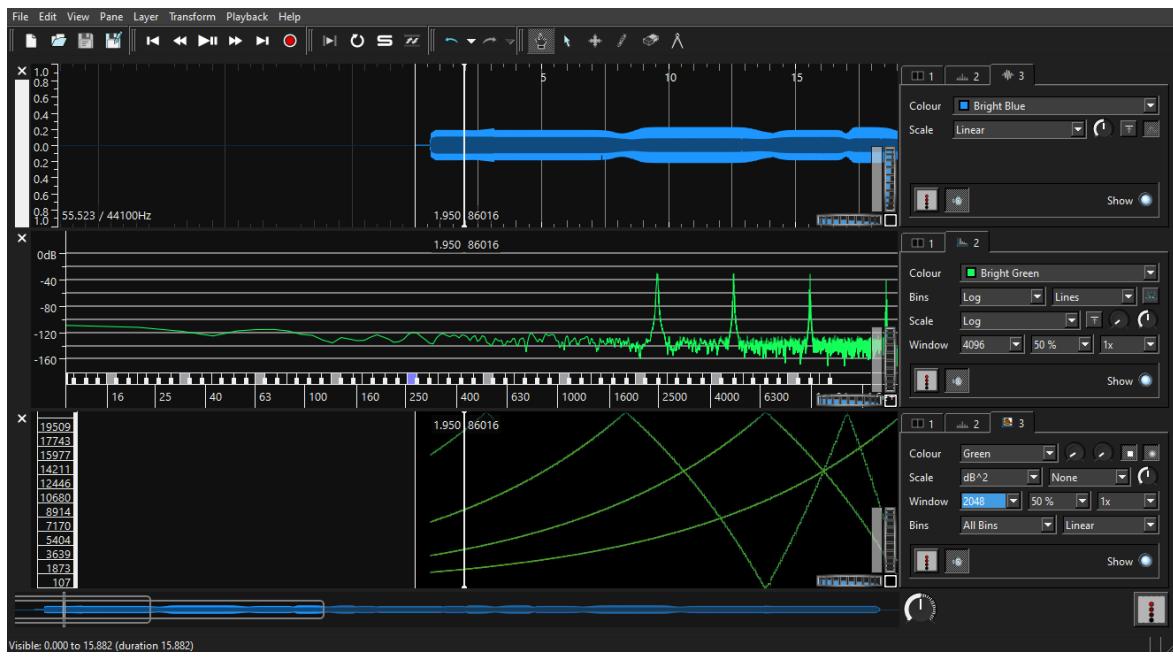
Grupa: MZ03IP1

Prowadzący: dr Marcin Lewandowski

Data: 25.01.2025r.

Zadanie 1. Próbkowanie





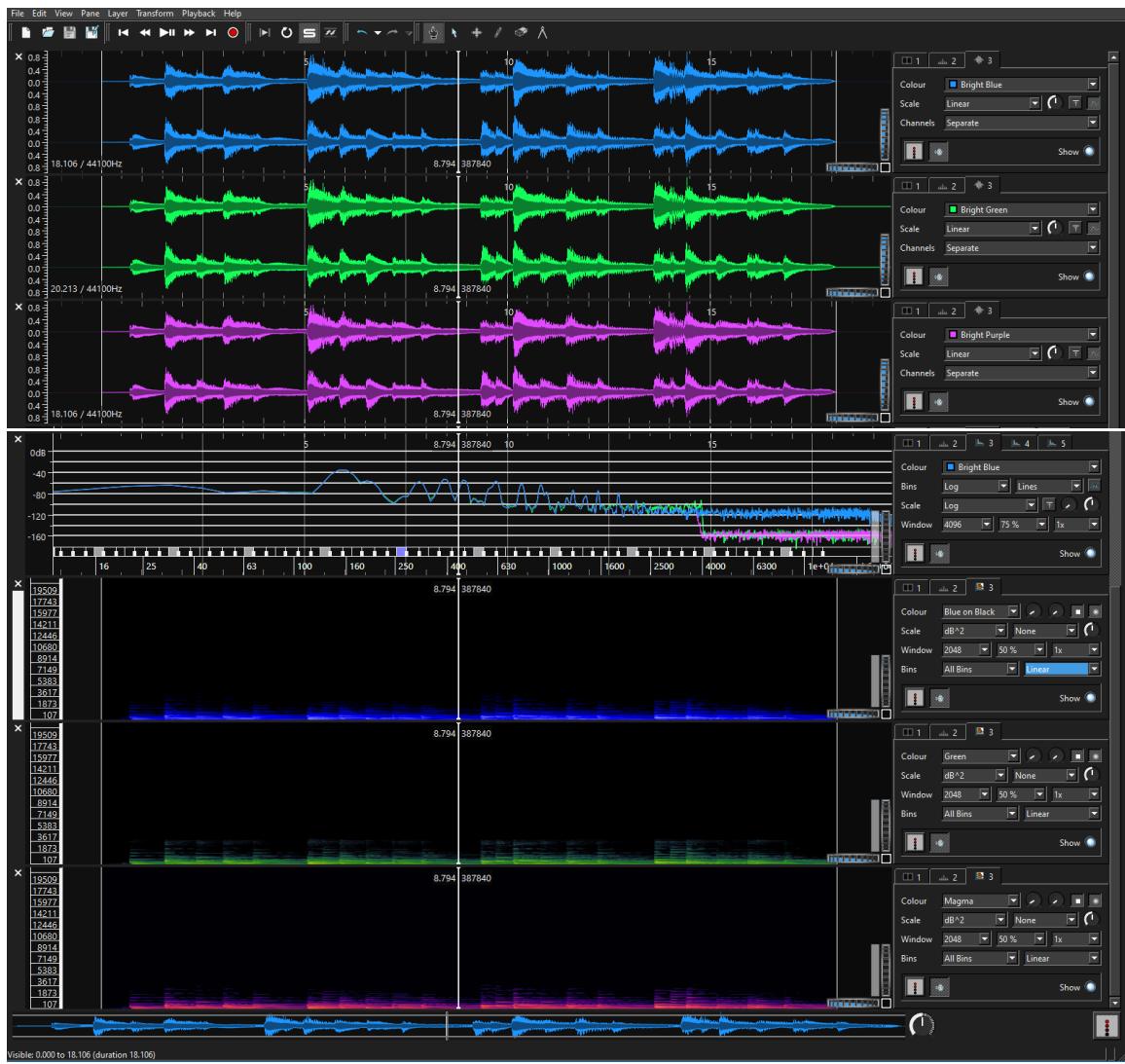
4) (Analiza w kolejnym punkcie)

- 5) W analizowanym sygnale dźwiękowym wyraźnie można zaobserwować zmiany wysokości tonów. Widmo pokazuje, jak częstotliwości składowe (2kHz, 4kHz, 8kHz, 16kHz) zmieniają się liniowo w czasie. Wrażenia słuchowe odpowiadają temu zjawisku, dźwięki stają się wyraźnie wyższe, a potem niższe.

Na spektrogramie można zauważyc przemieszczanie się poszczególnych składowych częstotliwościowych – linie przesuwają się równomiernie w górę (podczas zwiększenia częstotliwości) oraz w dół (podczas zmniejszania częstotliwości)

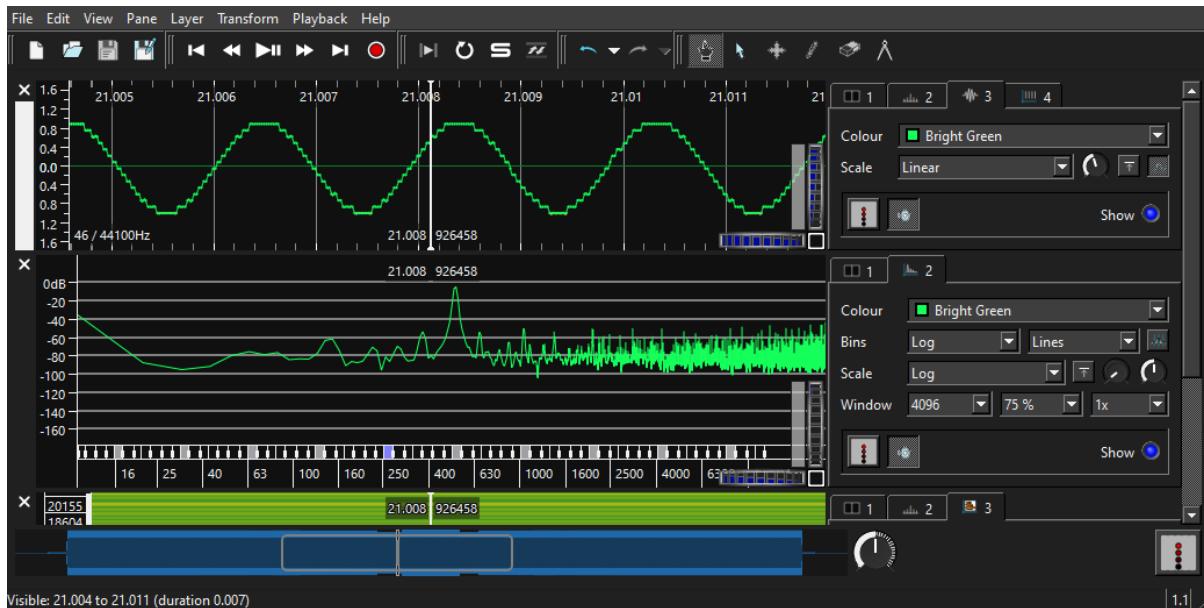
Porównując widma z wrażeniami słuchowymi widać, że dźwięki o wyższej częstotliwości są widoczne jako wyższe pasma na spektrogramie, natomiast wraz ze spadkiem częstotliwości pasma te się zmniejszają.

6)

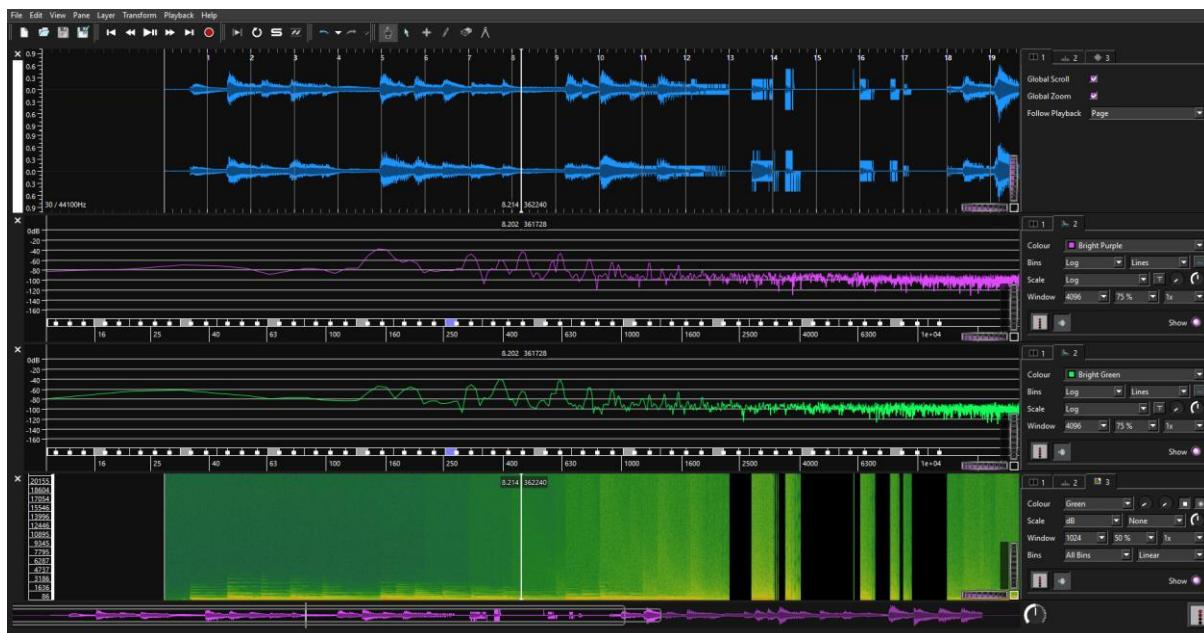


- 7) Po odsłuchaniu wszystkich utworów można stwierdzić, że plik *sample1* w porównaniu do oryginalnego utworu brzmi płasko i wydaje się stłumiony. Plik *sample2* natomiast, porównując go z oryginałem, jest pozbawiony dużej części szumów, kosztem delikatnej utraty brzmienia utworu. W mojej osobistej ocenie najprzyjemniej słucha się utworu *sample2*. Na powyższym zrzucie ekranu widać na spektrum, że w przypadku obu plików *sample* pasma dźwiękowe tracą częstotliwości powyżej 4kHz. Dodatkowo pasmo dźwiękowe pliku *sample1* w delikatnie różni się od pasma dźwiękowego oryginalnego pliku dla częstotliwości poniżej 4kHz. Różnica między tymi plikami może wynikać z występowania zjawiska aliasingu. W pliku *sample1* prawdopodobnie nie zastosowano filtra antyaliasingowego, natomiast w przypadku pliku *sample2* taki filtr został zastosowany.

Zadanie 2. Kwantyzacja

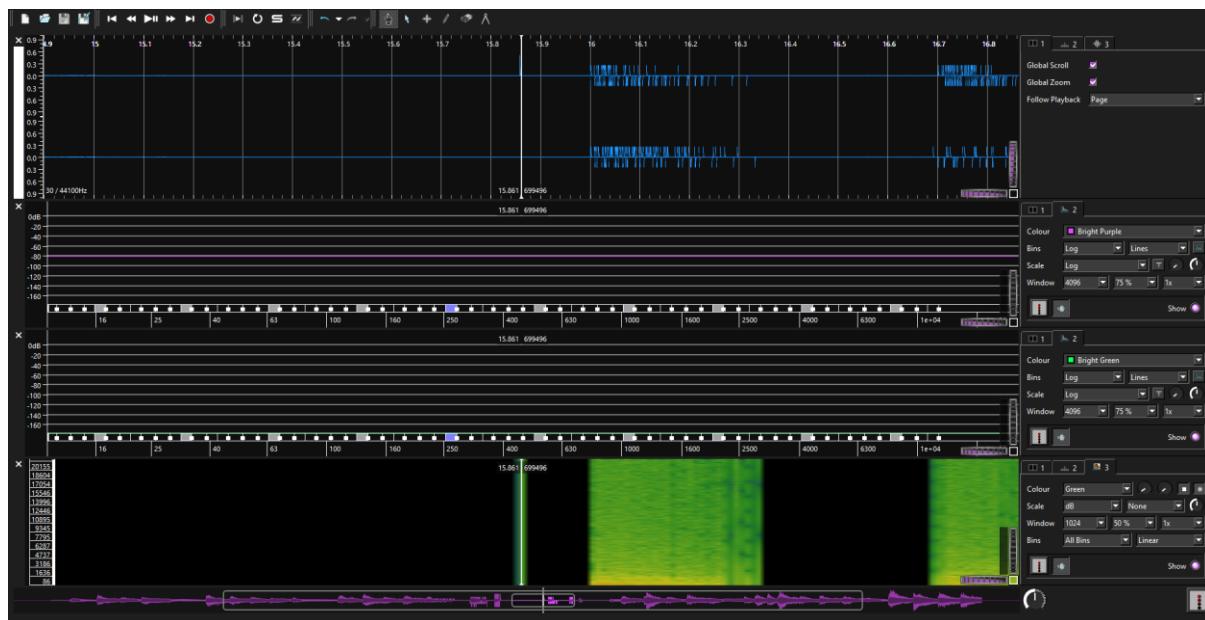


Na wyżej załączonym zrzucie ekranu uchwycono moment, w którym kwantyzacja pliku dźwiękowego *quantization_sinus_mono_loweringBitDepth.wav* jest 3 bitowa. Sygnał czasowy staje się "schodkowy" i mniej płynny, co wynika z ograniczonej liczby poziomów kwantyzacji. W widmie obserwuje się wzrost poziomu szumu a także częstotliwości. Z 3-bitową kwantyzacją sygnał staje się bardzo zniekształcony, przypominający metaliczne dźwięki, można doświadczyć charakterystycznych trzasków, "trzeszczenia" i bardzo niskiej jakości dźwięku.

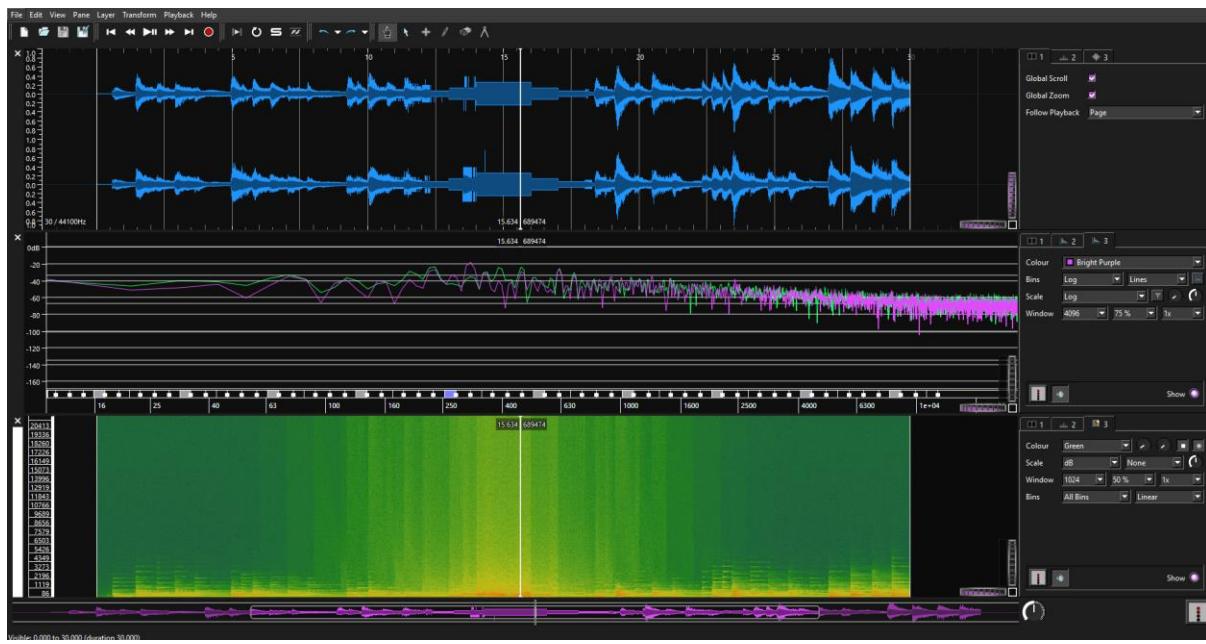


Na powyższym zrzucie ekranu pliku dźwiękowego *piano_16b_to_2b_to_16b_quantizer1.wav* widać, że w niektórych sekcjach (np. okolice 15–20 sekundy) występują wyraźne "schodki" na sygnale. Te artefakty są wynikiem obniżenia rozdzielczości kwantyzatora do 2 bitów. W tych fragmentach poziom szczegółowości jest znaczco zmniejszony, co wpływa na kształt fali

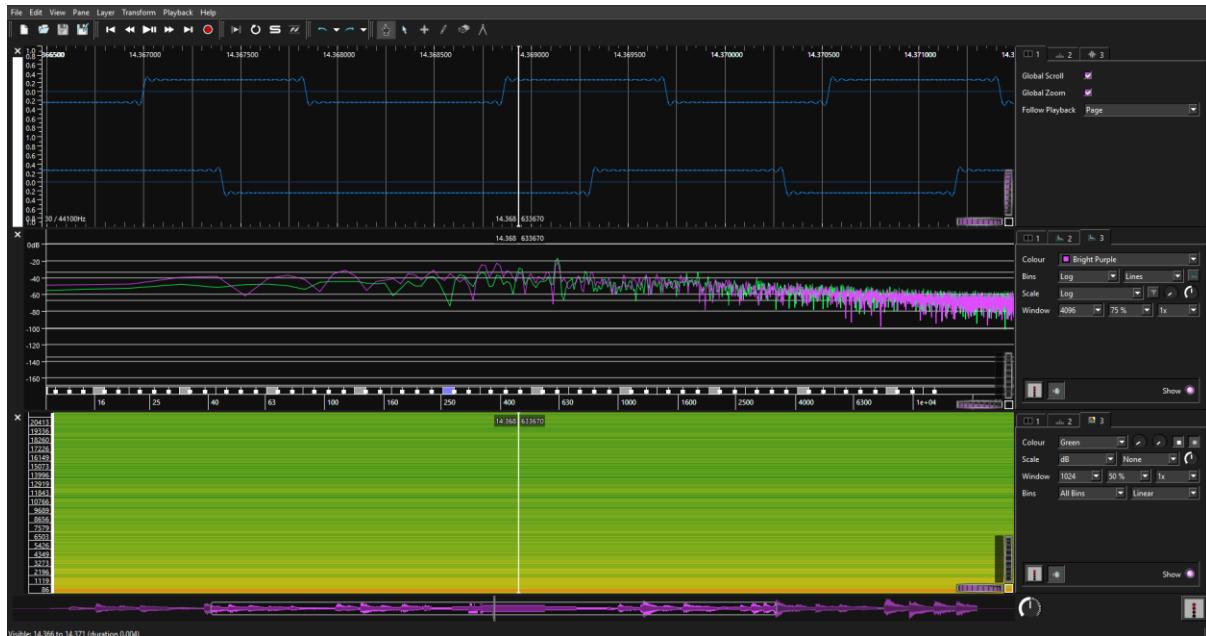
dźwiękowej. Przebieg czasowy po kwantyzacji do 16 bitów (np. w innych sekcjach) wygląda znacznie płynniej. Natomiast na widmie częstotliwości w sekcjach z 2-bitową kwantyzacją widać wzrost szumu w całym zakresie częstotliwości, co jest typowe dla szumu kwantyzacyjnego. Fragmenty z 16-bitową rozdzielczością zachowują znacznie lepszy stosunek sygnału do szumu, co widać jako czystsze widmo. W przypadku spektrogramu widoczne są pionowe ciemne linie w okolicach wyższych częstotliwości (np. przy około 17 sekundzie). Oznaczają one brak szczegółowych danych w tym zakresie częstotliwości, co wskazuje na degradację jakości dźwięku w tych fragmentach, gdy kwantyzacja była ograniczona do 2 bitów. Ponownie podczas przesłuchiwanego pliku dźwiękowego można doświadczyć "trzeszczenia" i bardzo niskiej jakości dźwięku.



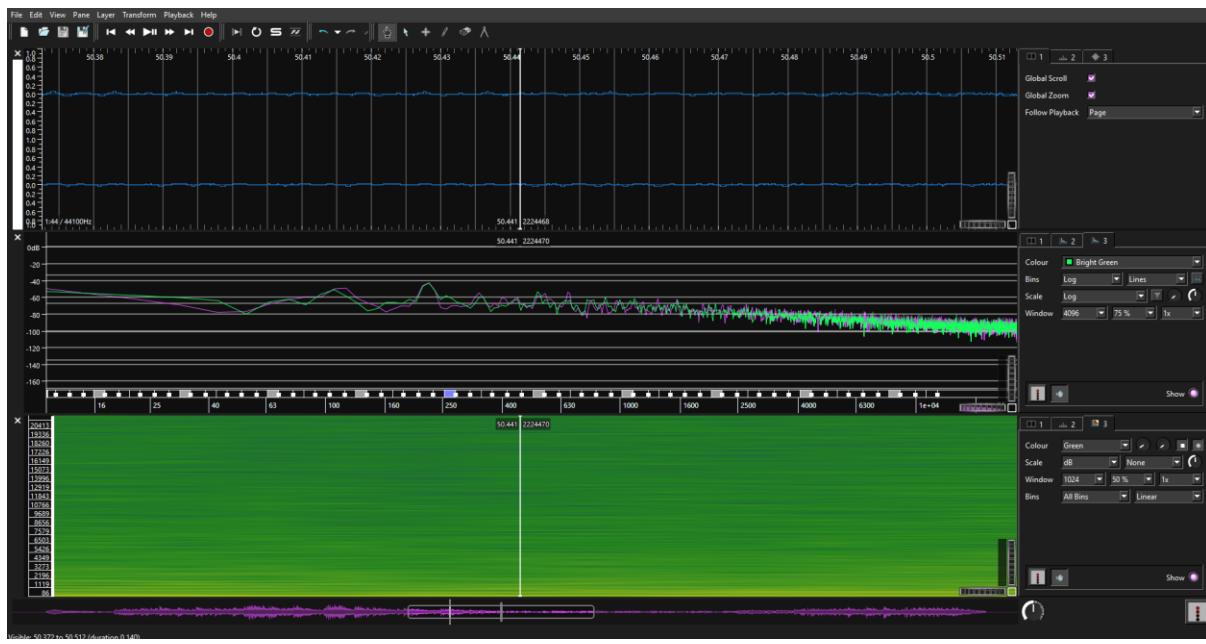
Natomiast na podstawie powyżej zamieszczonego zrzutu ekranu, można wywnioskować, że zastosowany kwantyzator to najprawdopodobniej kwantyzator mid-tread, ponieważ widoczne jest to, że sygnał osiąga "stan zerowy" jako oddzielny próg kwantyzacji.



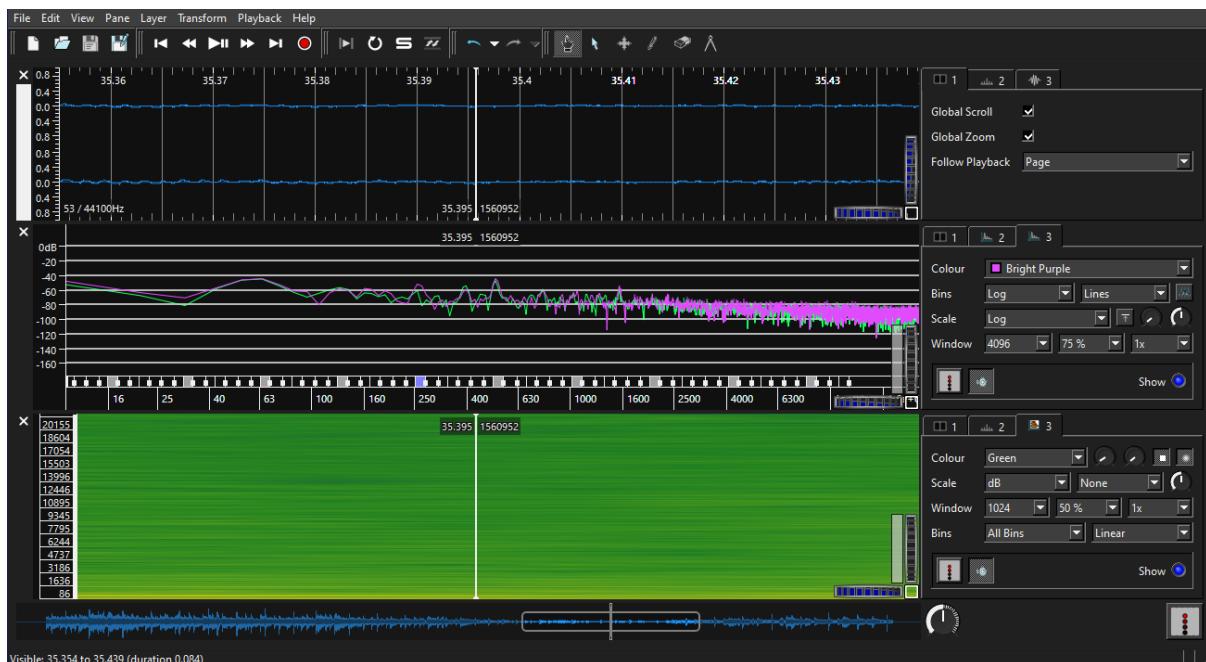
W sekcji od około 12 do 18 sekund widoczne są duże zmiany w spektrogramie (wyraźny wzrost szumów kwantyzacyjnych), co wskazuje na 2-bitową kwantyzację. W widmie czasowym widoczne są skokowe zmiany w przebiegu próbek. Te zmiany mogą wskazywać na różne poziomy kwantowania – im mniej bitów, tym bardziej skwantowany przebieg staje się schodkowy i zniekształcony. Przy niższej rozdzielczości bitowej można zauważać więcej szumów, szczególnie w zakresie wysokich częstotliwości. W spektrogramie widoczny jest wzrost szumów i mniej wyraźna separacja harmonicznych podczas przejścia na niższe bity. Podczas odsłuchu dominuje nieprzerwany szum we wcześniej wspomnianej sekcji i dźwięk staje się prawie całkowicie nieczytelny.



Na wyżej załączonym zrzucie ekranu w widoku czasowym schodkowa struktura sygnału wskazuje, że wartości przechodzą przez zero. Oznacza to, że prawdopodobnie nie ma poziomu zerowego, co jest charakterystyczne dla kwantyzatora mid-rise.

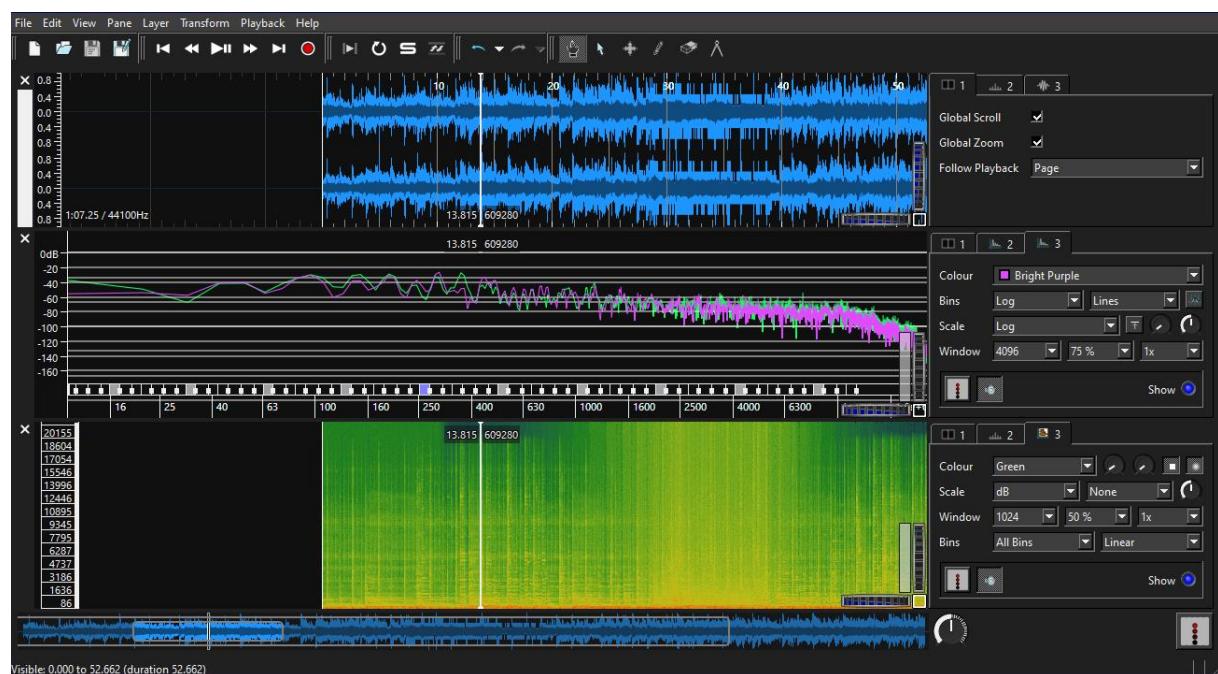


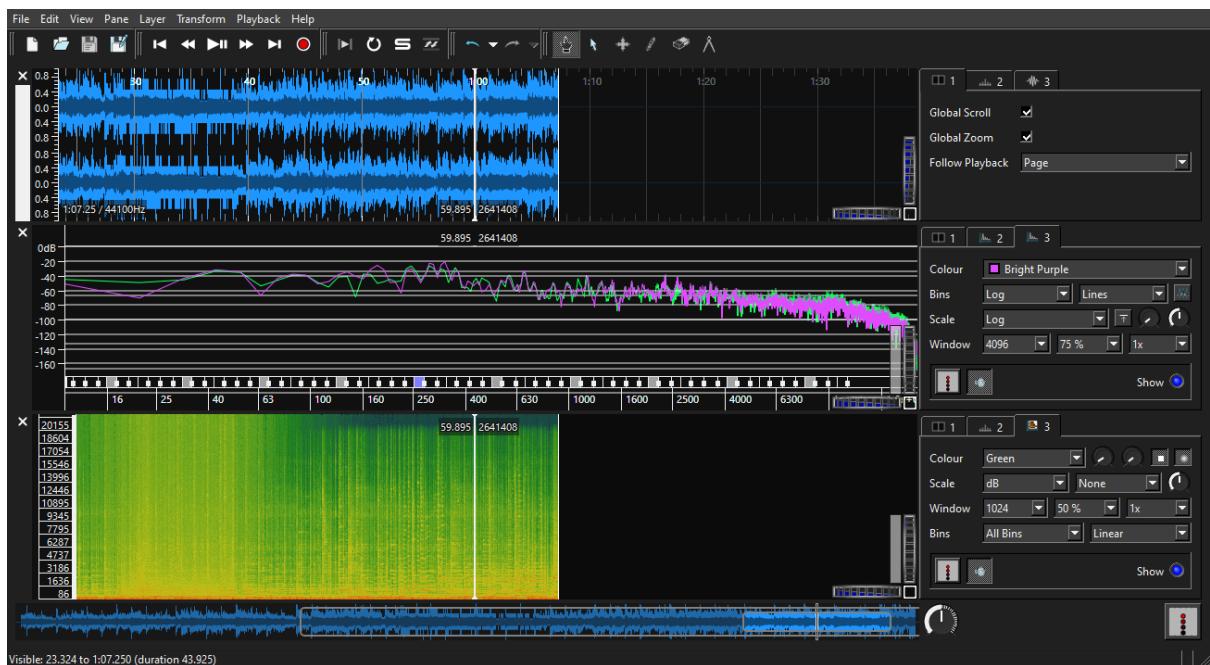
Na załączonym wyżej zrzucie ekranu z plikiem dźwiękowym *quantization_music_1_8bit_fade_error_compensated.wav* widać, że przebieg sygnału pokazuje pewną „schodkowość” charakterystyczną dla ograniczonej liczby poziomów kwantyzacji (8 bitów). W miarę obniżania poziomu sygnału „schodki” stają się bardziej widoczne, ponieważ sygnał oryginalny jest coraz bliższy błędowi kwantyzacji. Na niskim poziomie amplitudy zakłócenia od błędu kwantyzacji dominują nad sygnałem. Przy zmniejszaniu amplitudy sygnału, poziom szumu kwantyzacji pozostaje względnie stały (błąd niezależny od poziomu sygnału), przez co staje się bardziej widoczny w widmie. Szum kwantyzacji jest widoczny jako szerokopasmowe zakłócenie, które staje się bardziej dominujące w miarę obniżania amplitudy sygnału. Wyraźnie widać, że poziom szumu pozostaje na stałym poziomie, nawet gdy poziom sygnału maleje. W miarę zmniejszania poziomu sygnału przed kwantyzatorem, szum kwantyzacji staje się bardziej słyszalny. Można usłyszeć charakterystyczne „trzaski” lub „ziarnistość” dźwięku.



Na wyżej załączonym zrzucie ekranu widoczna jest „schodkowość” sygnału (zwłaszcza w miejscach o niższej amplitudzie sygnału). Ograniczona liczba poziomów kwantyzacji (8 bitów) powoduje, że drobne zmiany amplitudy nie są odwzorowywane. W miarę zmniejszania amplitudy sygnału, szумy kwantyzacji dominują nad oryginalnym sygnałem. Natomiast szum kwantyzacji utrzymuje się na względnie stałym poziomie w widmie, niezależnie od amplitudy utworu. W miejscach o niskiej staje się dominujący w zakresie średnich i wysokich częstotliwości. Na spektrogramie wyraźnie widać, że szum kwantyzacji jest obecny w całym zakresie częstotliwości i staje się bardziej zauważalny w miarę zmniejszania poziomu sygnału. Szum ma charakter szerokopasmowy i intensyfikuje się na tle oryginalnego dźwięku. Podczas odsłuchiwania przy wysokim poziomie sygnału błąd kwantyzacji jest mniej zauważalny – szum jest maskowany przez oryginalną treść dźwięku. Natomiast w miarę zmniejszania amplitudy sygnału wejściowego, błąd kwantyzacji staje się bardziej słyszalny i odbierany jako szum lub „ziarnistość” dźwięku.

W teorii, dla dobrze zaprojektowanego kwantyzatora i równomiernego rozkładu sygnału, błąd kwantyzacji może być modelowany jako dodanie niezależnego, równomiernie rozłożonego szumu o szerokim paśmie. W przypadków analizowanych plików dźwiękowych błąd kwantyzacji nie jest idealnie addytywny w niższych poziomach amplitudy. Przy bardzo niskich amplitudach, widoczna jest korelacja między sygnałem wejściowym a szumem kwantyzacji. W szczególności w pliku *quantization_music_2_8bit_fade_error_compensated.wav*, szum kwantyzacji wydaje się bardziej zintegrowany z sygnałem i mniej równomierny. W wysokim poziomie sygnału błąd kwantyzacji jest mniej zależny od sygnału wejściowego, ponieważ dominuje energiczny sygnał muzyczny. Wrażenie jest zbliżone do addytywnego szumu. Natomiast przy niskim poziomie sygnału, zależność błędu od sygnału staje się bardziej zauważalna. Sygnał oryginalny zdaje się modulować szum kwantyzacji, co wyklucza jego całkowitą niezależność. Dla obu plików błędy kwantyzacji zachowują stały charakter w zakresie rozkładu i szerokości widma. Jednak zmieniająca się percepcja błędu wynika z relacji między poziomem sygnału a błędem. Gdy sygnał słabnie, szum staje się bardziej zauważalny, ale jego właściwości widmowe pozostają spójne.





Na powyżej załączonych zrzutach ekranu z początku, środka oraz końca ścieżki dźwiękowej *quantization_music_3_8bit_to_16b_downto_3_andback_to_24b.wav* błęd kwantyzacji, zwany również szumem kwantyzacji, powstaje podczas procesu przekształcania sygnału analogowego na cyfrowy. Polega on na zaokrąglaniu wartości sygnału do najbliższych dyskretnych poziomów, co prowadzi do niewielkich różnic między sygnałem oryginalnym a skwantowanym. W idealnych warunkach, gdy sygnał ma dużą dynamikę, a liczba poziomów kwantyzacji jest wysoka, błąd ten można modelować jako addytywny szum o rozkładzie równomiernym, niezależny od sygnału wejściowego. Jednak w praktyce, zwłaszcza przy niskiej liczbie bitów (np. 3 bity), błąd kwantyzacji staje się zależny od sygnału, co prowadzi do zniekształceń nietliniowych. Podczas analizy widma sygnału skwantowanego można zauważać obecność dodatkowych składowych w zakresie wysokich częstotliwości, które nie występowały w sygnale oryginalnym. Zjawisko to wynika z nietliniowości procesu kwantyzacji. Przy niskiej rozdzielczości bitowej, kwantyzator wprowadza

zniekształcenia harmoniczne oraz efekt aliasingu, co skutkuje pojawieniem się nowych składowych częstotliwościowych. Te elementy są wynikiem interakcji między sygnałem a błędem kwantyzacji i są szczególnie zauważalne przy niższych rozdzielczościach bitowych. Podsumowując, błąd kwantyzacji nie zawsze może być traktowany jako addytywny i niezależny od sygnału wejściowego. Jego charakter zależy od liczby bitów użytych do kwantyzacji oraz od właściwości samego sygnału. Dodatkowe składowe wysokoczęstotliwościowe pojawiają się głównie w wyniku nieliniowości procesu kwantyzacji i są bardziej wyraźne przy niższych rozdzielczościach bitowych.

Zadanie 3. Dithering i kształtowanie szumu rekwantyzacji

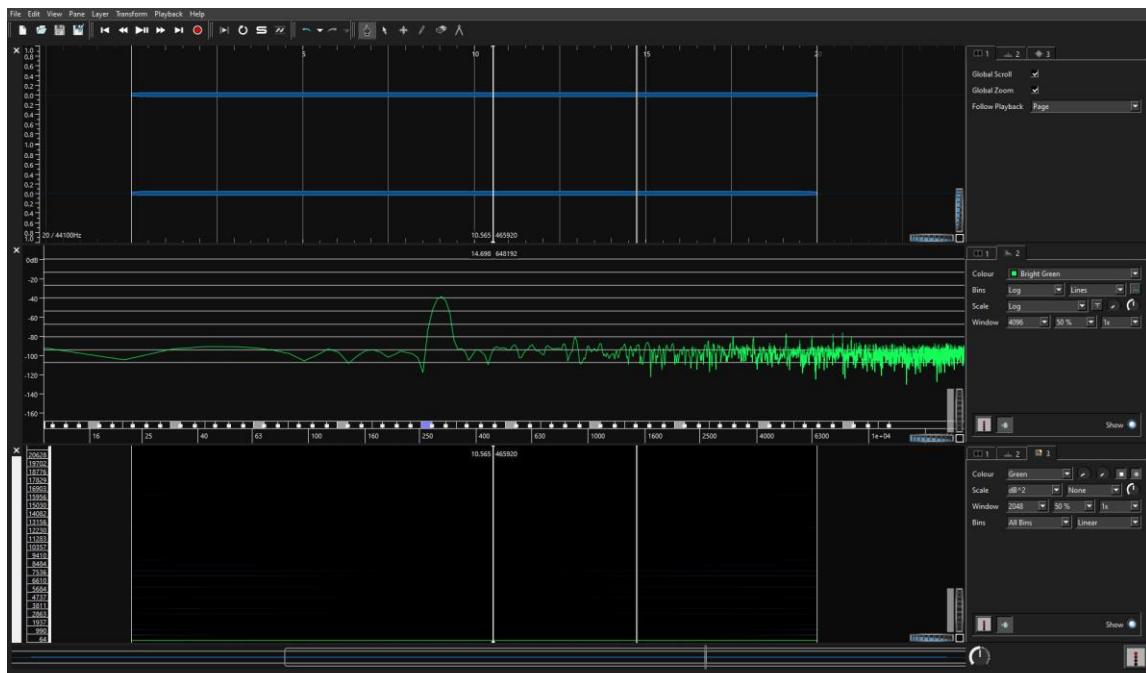


- 1)
- 2) Na powyższym zrzucie ekranu widać, że plik rozpoczyna się od sygnału skwantowanego do 8 bitów i następnie zmniejszany jest poziom sygnału przy kwantyzacji do -28dB. Następnie od około 20 sekundy nagrania włączony zostaje pierwszy typ sygnału dithera, a następnie drugi, po czym następuje powrót do pierwszego dithera. Od około 40 sekundy ustawiony zostaje łagodny kształt błędu kwantyzacji, następnie większy, a na końcu najbardziej stromy. Na koniec następuje powrót do pierwszego dithera.

W przypadku włączenia pierwszego dithera szum staje się bardziej zauważalny, a na spektrogramie widać, że energia błędu w widmie jest rozłożona równomiernie. Przy włączeniu drugiego dithera energia nadal jest rozproszona, ale nie w takim stopniu jak w przypadku pierwszego (powstają proste linie na spektrogramie, ale wyżej niż wcześniej). W przypadku łagodnego kształtowania błędu kwantyzacji zmniejszenie odczuwalności szumów, niż w przypadku pierwszego dithera, a na spektrogramie widać, że większa część energii przesuwana jest w wyższe częstotliwości.

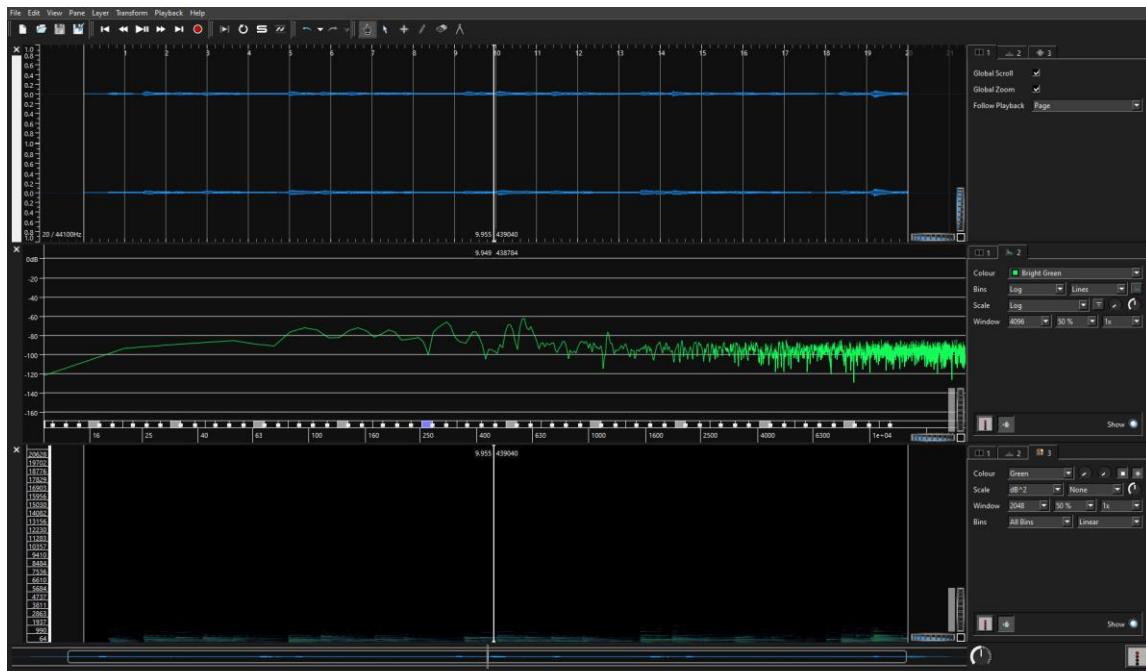
Gdy kształtowanie błędu staje się silniejsze szum jest delikatnie bardziej odczuwalny niż w poprzednim etapie, a energia jest bardziej rozproszona na niższe częstotliwości.

W etapie, w którym zastosowano najbardziej agresywne kształtowanie błędu szumy są mniej lekko mniej zauważalne niż w przypadku łagodnego kształtowania oraz widać, że większa część energii została przeniesiona na wyższe częstotliwości.



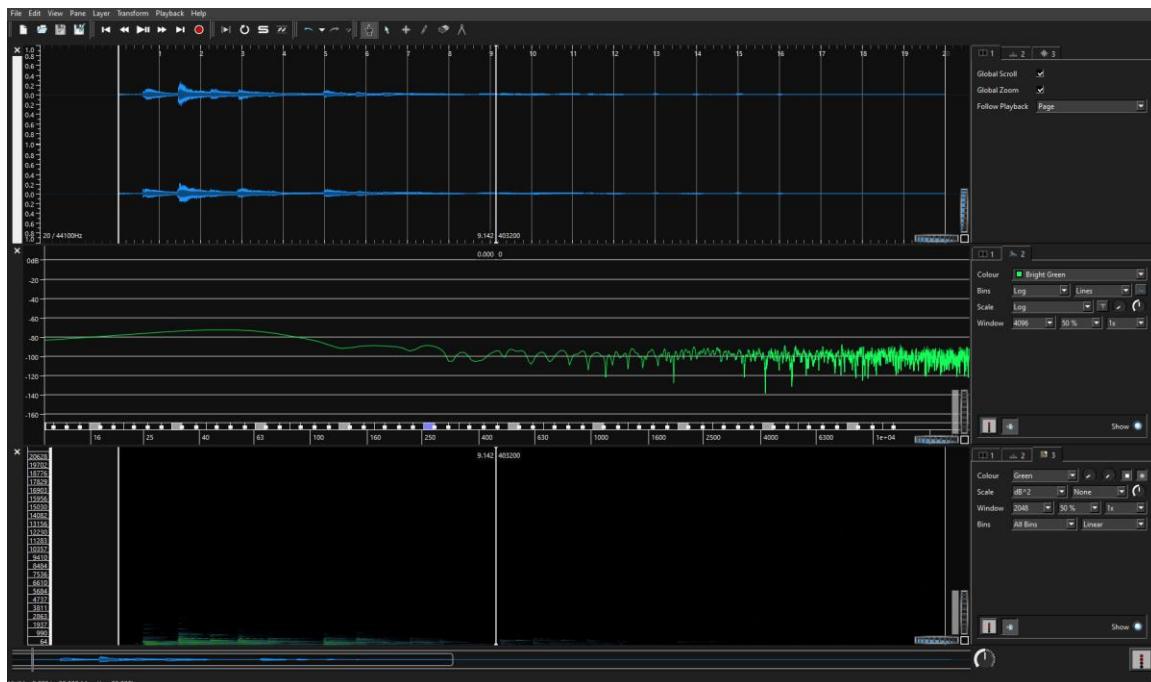
3)

- 4) Od około 6 sekundy nagrania do około 13 sekundy można zauważać skuteczną eliminację zniekształceń harmonicznych. Bardziej odczuwalne również stają się wtedy szумy. Biorąc pod uwagę, że amplituda jest zmieniana od 0 do 1LSB i z powrotem do 0, można w przybliżeniu określić, że przy amplitudzie od amplitudy około 0,5-0,6 uzyskiwane są najlepsze rezultaty.



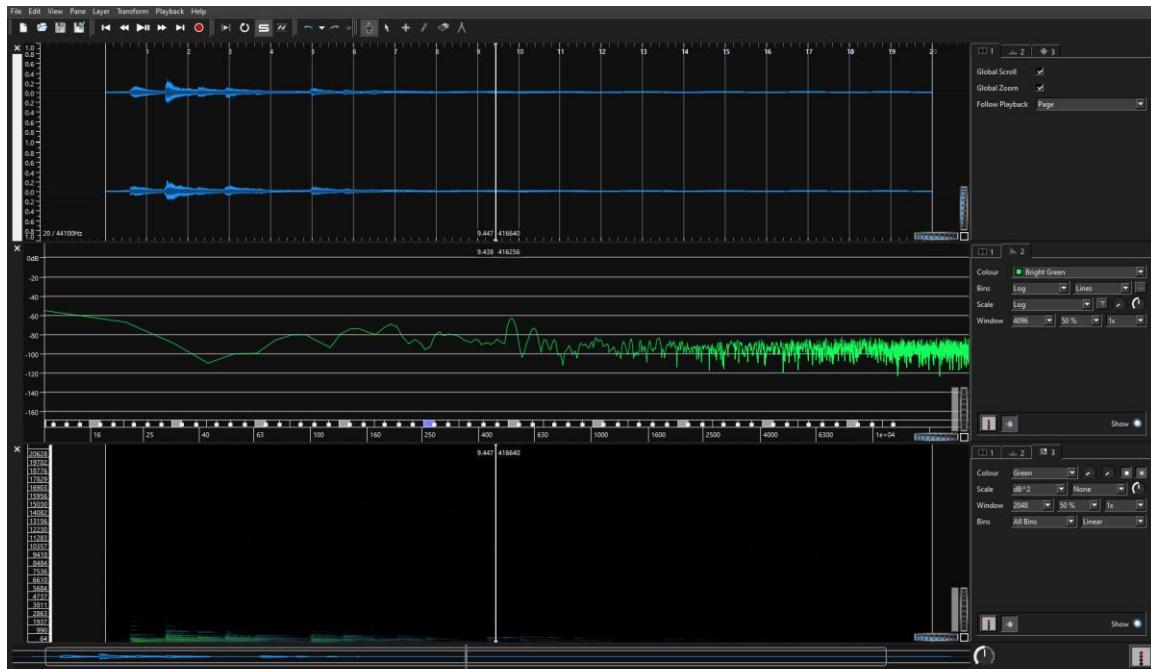
5)

- 6) Od ok. 4 – 16 sekundy nagrania można zauważać skuteczną eliminację zniekształceń harmonicznych. Porównując otrzymane rezultaty z nagraniem sinusoidalnym, w przypadku nagrania piano, niższa amplituda (już ok. 0,4-0,5) pozwala uzyskać oczekiwany efekt. Może to być spowodowane tym, że nagranie fortepianu jest bardziej złożone przez co zniekształcenia harmoniczne są trudniejsze do zauważenia.



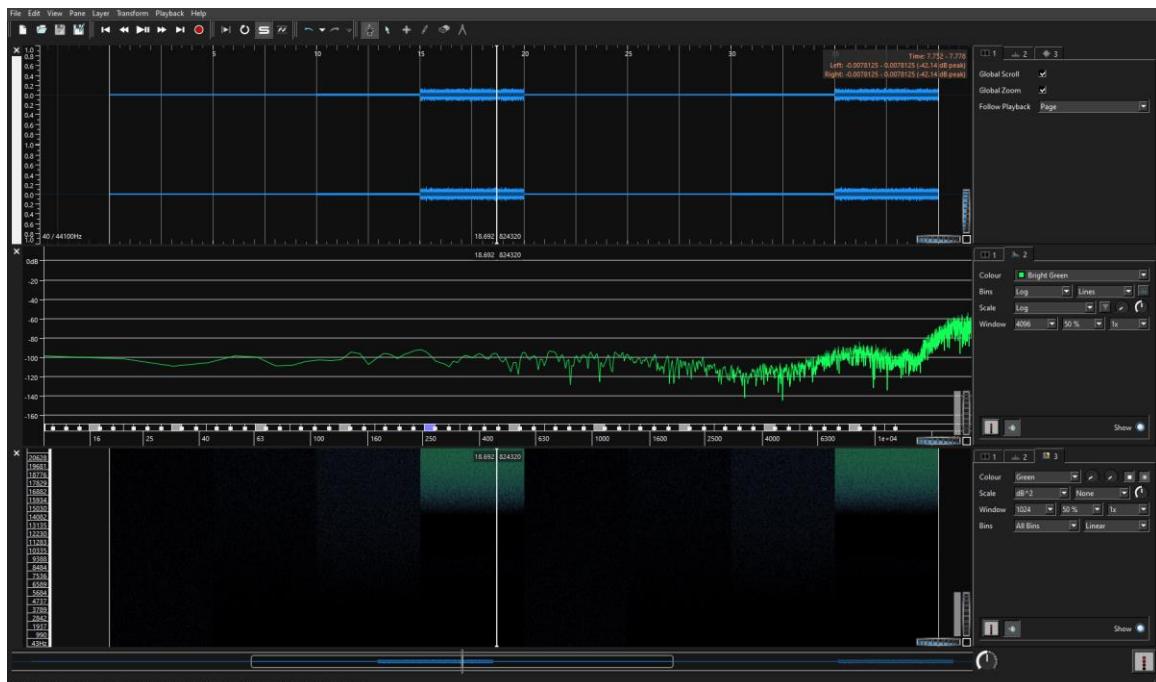
7)

- 8) Od około 12 sekundy fortepijan praktycznie przestaje być słyszalny i można zauważać falowanie. Na nagraniu samego błędu kwantyzacji wspomniane falowanie można usłyszeć już od około 8 sekundy. Dithering RPDF choć skutecznie maskuje błąd kwantyzacji, to modulacja szumów nie jest w pełni eliminowana.



9)

- 10) Podobnie jak w przypadku dithera RPDF dźwięk fortepijanu w pewnym momencie praktycznie przestaje być słyszalny, natomiast nie występuje (lub w niewielkim stopniu) wspomniane wcześniej falowanie. Oznacza to, że dither o rozkładzie TPDF sprawdza się lepiej w przypadku eliminowania błędu modulacji szumu niż RPDF.



11) Click and drag to navigate; use mouse-wheel or trackball scroll to zoom; hold Shift and drag to zoom to an area

12) W przypadku dithera TPDF 2LSB energia w widmie jest rozłożona równomiernie w całym słyszalnym paśmie. Błędy są zatem maskowane w pełnym zakresie częstotliwości. W przypadku dithera filtrowanego górnoprzepustowo TPDF 2LSB można zaobserwować zmniejszenie energii w niższych częstotliwościach, co sprawia, że szum wydaje się mniej przytaczający.

W przypadku dithera TPDF 2LSB kształtowanego filtrem 1-szego rzędu energia w widmie zostaje bardziej przesunięta ku wyższym częstotliwościami, a energia w niższych częstotliwościach jest znacznie obniżona, co dalej redukuje odczuwanie szumu. W przypadku najbardziej zaawansowanego dithera TPDF 2LSB kształtowanego funkcją 9-tego rzędu można zauważać praktycznie całkowity brak energii w niższych częstotliwościach, co sprawia, że szum jest najmniej odczuwalny w porównaniu

Zadanie 4. Jitter i błędy synchronizacji zegarów

- **Jitter1.wav:** w widmie częstotliwościowym można zaobserwować dodatkowe składowe harmoniczne wokół częstotliwości 13 kHz, będące wynikiem modulacji zegara sygnałem prostokątnym o częstotliwości 8 kHz; podczas odsłuchiwania można usłyszeć obecność dodatkowych tonów, nadających sygnałowi metaliczny (piszący) charakter;
- **Jitter2.wav:** na widmie można było zauważać obecność szerokopasmowego szumu wokół częstotliwości 13 kHz, spowodowanego modulacją zegara sygnałem szumowym; sygnał zawierał wyraźny szum, zaktócający czystość oryginalnej sinusoidy;
- **Jitter3.wav:** można było dostrzec modulację amplitudy sygnału 13 kHz, wynikającą z modulacji zegara sygnałem o niskiej częstotliwości; sygnał wykazywał powolne zmiany głośności, przypominające „pulsowanie”;
- **Jitter4.wav:** dźwięk fletu był czysty, bez żadnych zniekształceń czy dodatkowych dźwięków; brzmienie było naturalne, takie jak fletu...;
- **Jitter5.wav:** w analizie widma można zauważać dodatkowe składowe harmoniczne, podobne do tych w przypadku sygnału sinusoidalnego; dźwięk fletu stał się metaliczny i nienaturalny, co przypominało efekt z sygnałem sinusoidalnym;

- **Jitter6.wav:** widmo pokazało obecność szerokopasmowego szumu, podobnie jak wcześniej; nagranie było zakłócone przez szum, co utrudniało usłyszenie czystego dźwięku fletu;
- **Jitter7.wav:** można zauważyc zmiany w głośności dźwięku tak jak w przypadku *Jitter3.wav*; dźwięk fletu pulsował, co sprawiało, że brzmienie było niestabilne;
- **Sync1.wav:** na wykresie częstotliwości pojawiły się dodatkowe skoki, co sugeruje obecność niepożądanych dźwięków; niestety ze względu na naturalne ograniczenie słuchu, nie jestem w stanie przeprowadzić pełnej obserwacji słuchowej... na podstawie wykresu jestem w stanie określić, że faktycznie w momencie gdy urządzenie nie pracują synchronicznie styczą charakterystyczne dla tego zajścia „kliknięcia”;
- **Sync2.wav:** podobnie jak w poprzednim przypadku, widoczne były dodatkowe skoki na wykresie częstotliwości; według teorii sprzed rozpoczęcia obserwacji, dźwięk powinien być mniej zniekształcony niż w *sync1.wav*, niestety przez wcześniej wspomniane naturalne ograniczenie słuchu, nie jestem tego w stanie określić przy pomocy obserwacji, natomiast nadal można dostrzec w dźwięku niepożądane tony;
- **Sync3.wav:** w przypadku tego pliku dźwiękowego, trudno było dostrzec na wykresie częstotliwości jakiekolwiek zmiany; natomiast po zobrazowaniu widma można dostrzec niewielkie różnice między dźwiękami; w momencie odtworzenia przy wcześniej wspomnianych artefaktach, można usłyszeć delikatne „kliknięcie”/”trzask”.

Podsumowując, jitter prowadzi do wyraźnych zniekształceń, które w zależności od rodzaju modulacji (harmoniki, szum, pulsowanie) mogą zmieniać charakter dźwięku, sprawiając, że staje się on metaliczny, szumiący lub niestabilny. Brak synchronizacji zegara (w plikach *sync1.wav*, *sync2.wav*, *sync3.wav*) powoduje pojawienie się "kliknięć" lub "trzasków", które są mniej zauważalne w porównaniu do zniekształceń harmonicznych, ale również wpływają na jakość dźwięku.

Zniekształcenia będą bardziej słyszalne w *Jitter1.wav* i *Jitter5.wav*, ponieważ wprowadzenie dodatkowych harmonicznych (szczególnie w przypadku *Jitter1.wav*) zmienia naturalne brzmienie sygnału w sposób, który jest łatwy do wykrycia przez ucho ludzkie. Harmoniczne o częstotliwości zbliżonej do głównej częstotliwości (13 kHz) wprowadzają wyraźny, rytmujący charakter brzmienia, który jest bardziej zauważalny niż szum czy zmiany w amplitudzie. W przypadku instrumentów akustycznych, takich jak flet, wprowadzenie nienaturalnych harmonicznych może mieć bardzo silny wpływ na odbiór dźwięku, ponieważ zmienia on charakterystyki brzmieniowe instrumentu, sprawiając, że dźwięk staje się nieprzyjemny lub zniekształcony.