

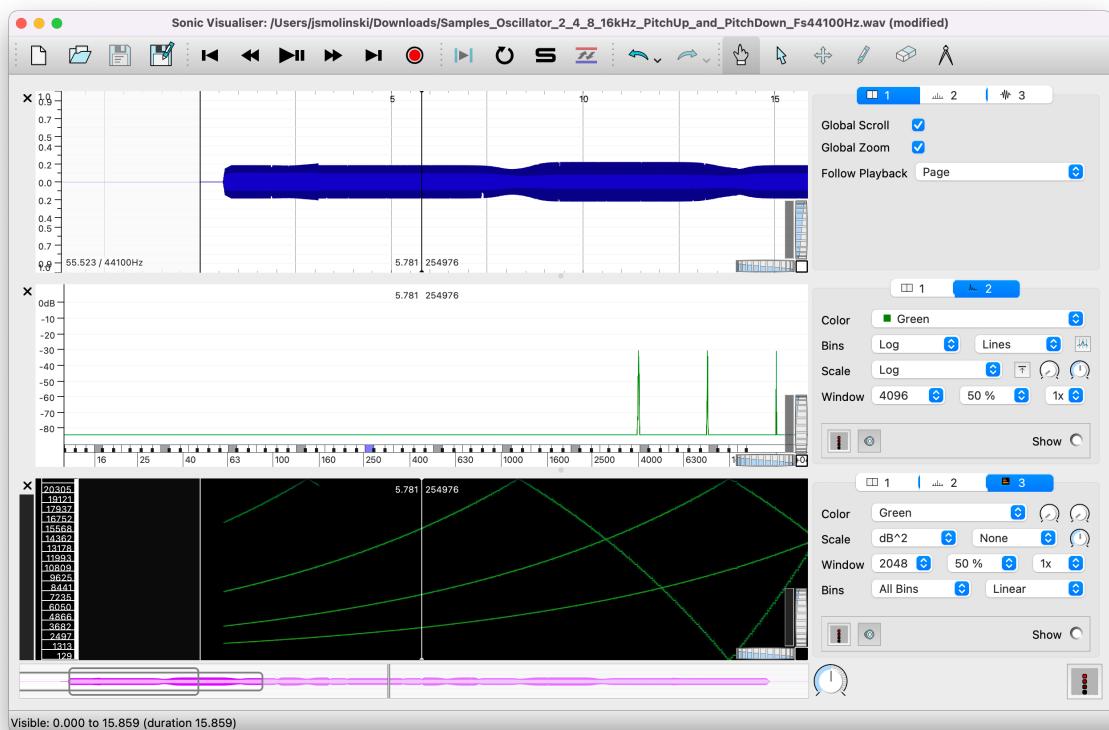
# Cyfrowa Technika Foniczna 2021Z

Sprawozdanie z Laboratorium 3 i 4

Jacek Smoliński nr ind. 15685

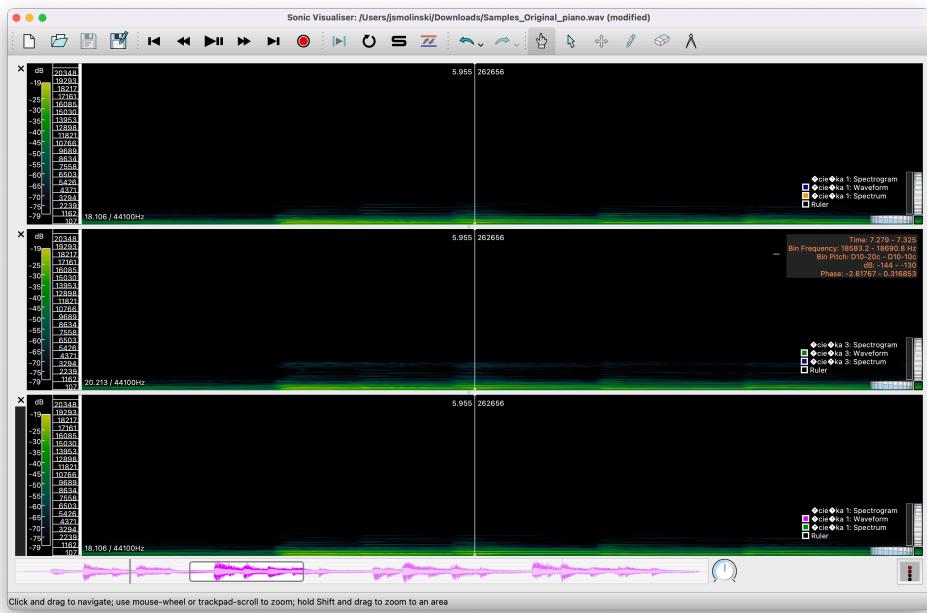
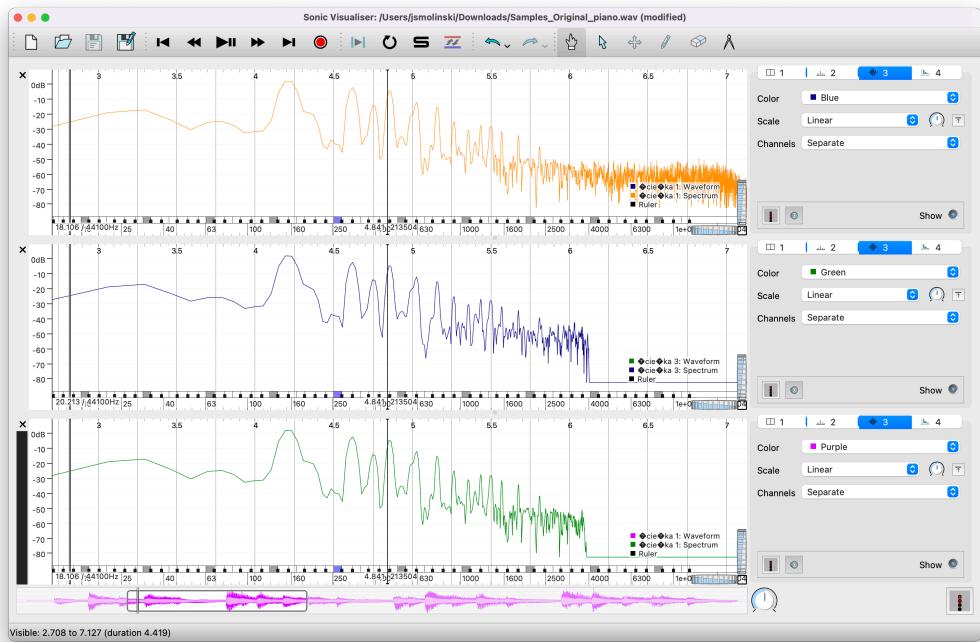
## Zadanie 1.

Załadowałem plik i dodałem panele z zadanimi opcjami:



To co rzuca się od razu w oczy, to początkowe cztery słupki na częstotliwościach 2,4,8,16 kHz w oryginalnym pliku. Tony wędrują do góry i jak można się spodziewać do połowy częstotliwości próbkowania (22.05kHz – twierdzenie Nyquista) częstotliwości są prawidłowo oddawane. Powyżej tego zakresu z racji braku stosowania żadnej filtracji w przetworniku a/c, można usłyszeć w praktyce efekt aliasingu. Tony powyżej częstotliwości 22.05kHz powodują powstawanie harmonicznych o innych częstotliwościach (spadki i chaos na spektrogramie po początkowych wzrostach). Dobrze też widać to na widmie, gdzie początkowe cztery słupki przesuwają się w prawo, a następnie odbijają i wracają w lewo (efekt niepożdanego lustra).

Następnie przeanalizowałem i porównałem kolejne dwa pliki z oryginałem



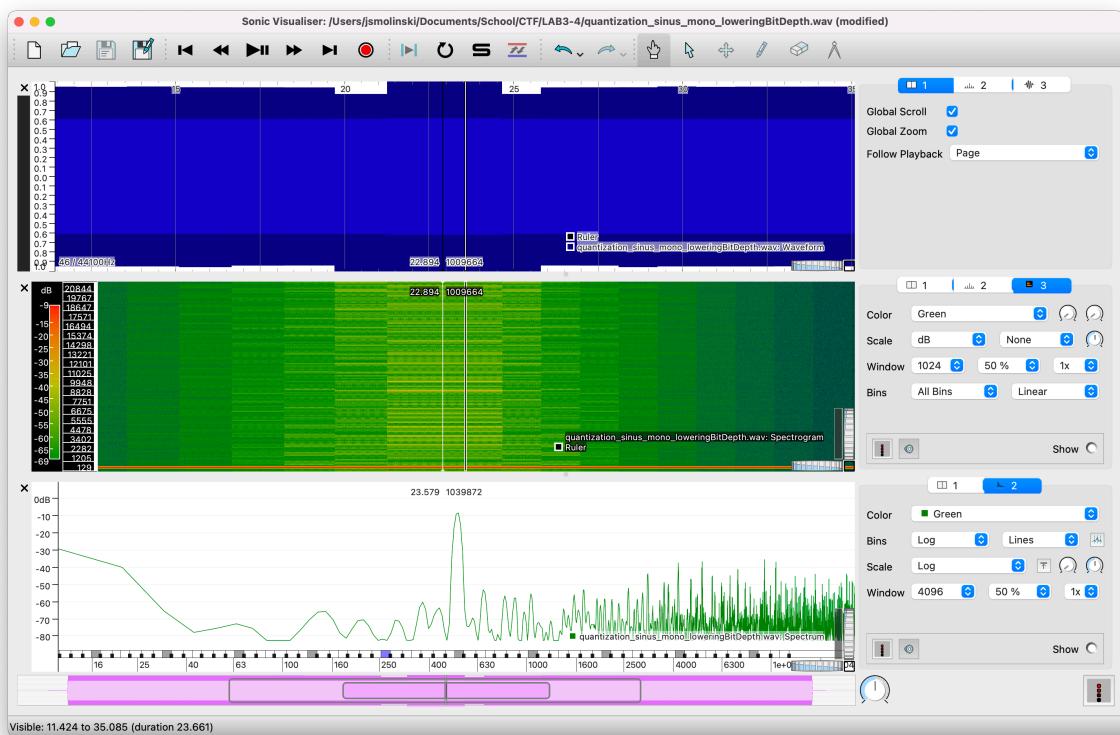
Od razu można zauważyc, że z powodu częstotliwości próbkowania, tony powyżej 4kHz są obcięte w obydwu plikach. Plik *sample1* (niebieski kolor, środkowy wykres) ma zdecydowanie gorszą jakość od pliku *sample2*. Widać to także w analizie widma, plik oryginalny i ostatni mają podobne widma (prócz obcięcia), plik środkowy ma zdecydowanie zmienione, zniekształcone widmo.

Na podstawie wykresów i odsłuchu stwierdzam, że plik *sample1* ze słyszalnymi zakłóceniami został zarejestrowany z użyciem filtra antialiasingowego (ucięcie na 4kHz i brak zakłóceń częstotliwościowych). Plik *sample2* prawdopodobnie na wyjściu został poddany filtracji antylustrzanej – znikają w nim szum i trzaski.

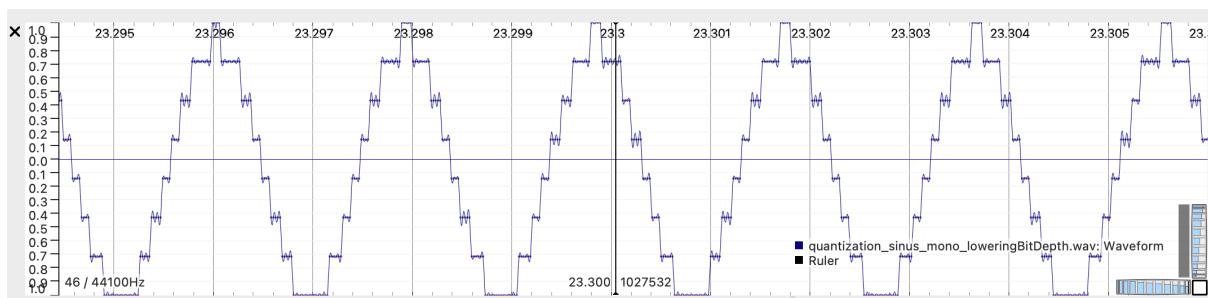
## Zadanie 2

### Zmiana liczby bitów kwantyzacji

Tak wygląda widmo i spektrogram pierwszego pliku:



Świetnie widać na nim punkty, w których zmienia się liczba bitów kwantyzacji. Około 21 sekundy zaczyna się sekcja skwantowana trzema bitami ( $2^3=8$  poziomów). Można to zobaczyć na wykresie (liczba schodków):



Patrząc również na brak poziomu zerowego, uznaję że jest to kwantyzator mid-rise.

Błąd kwantyzacji w dziedzinie częstotliwości można zaobserwować w widmie jako niepożądane częstotliwości nie należące do sygnału, w dziedzinie czasu widać „schodki”. Wrażenie towarzyszące zmniejszaniu liczby bitów kwantyzacji – zwiększenie szumu, wyostrzenie fal sinusoidalnej, zauważalne pogorszenie jakości mniej więcej od poziomu 8 bitów w dół.

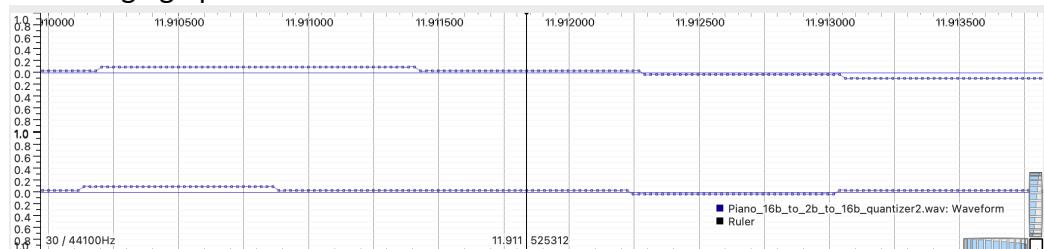
## Kwantyzator mid-tread i mid-rise

Żeby określić typ kwantyzatora należy przyjrzeć się czy występuje wartość 0 na wykresach sygnału w dziedzinie czasu. W tym celu warto spojrzeć na fragment o niskiej rozdzielcości bitowej.

Fragment pierwszego pliku audio w dziedzinie czasu:



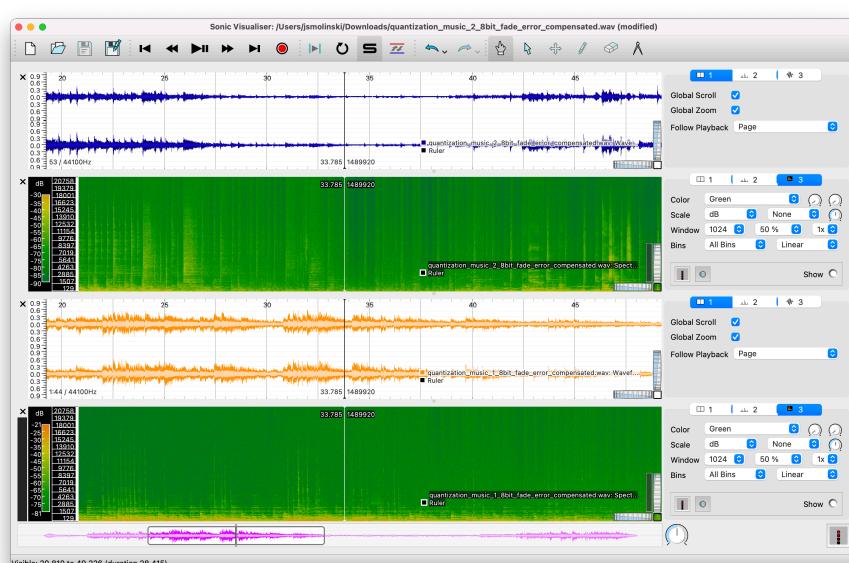
Fragmentu drugiego pliku audio w dziedzinie czasu:



Pierwszy kwantyzator jest typu mid-tread, a drugi typu mid-rise.

## Szum kwantyzacji

Porównanie dwóch plików:



Moim zdaniem błąd kwantyzacji może być traktowany jako addytywny i niezależny od sygnału wejściowego tylko w przypadku spełnienia założeń modelu liniowego. Po przesłuchaniu

próbek, stwierdzam że w wypadku tych nagrań założenia te nie zawsze są spełnione. Kwantyzator nie jest przesterowany, a krok kwantyzacji jest dostatecznie mały, natomiast mam poważne obawy co do liczby kroków kwantyzacji w krytycznych sekcjach próbek. Przy bardzo małej liczbie bitów niestety szum jest zależny od sygnału wejściowego. Drugim zarzutem jest wygładzona funkcja prawdopodobieństwa. Próbki różnią się od siebie i mają inną charakterystykę. Pierwsza próbka jest bardziej jednostajna, ma mniejszą dynamikę (amplitudę) i szum kwantyzacji zaczyna być szybciej dostrzegalny. W drugiej próbce natomiast błąd kwantyzacji dostrzegłem później, a nawet silny błąd przy niskiej rozdzielcości pozwalał na odsłuchanie linii wokalnej.

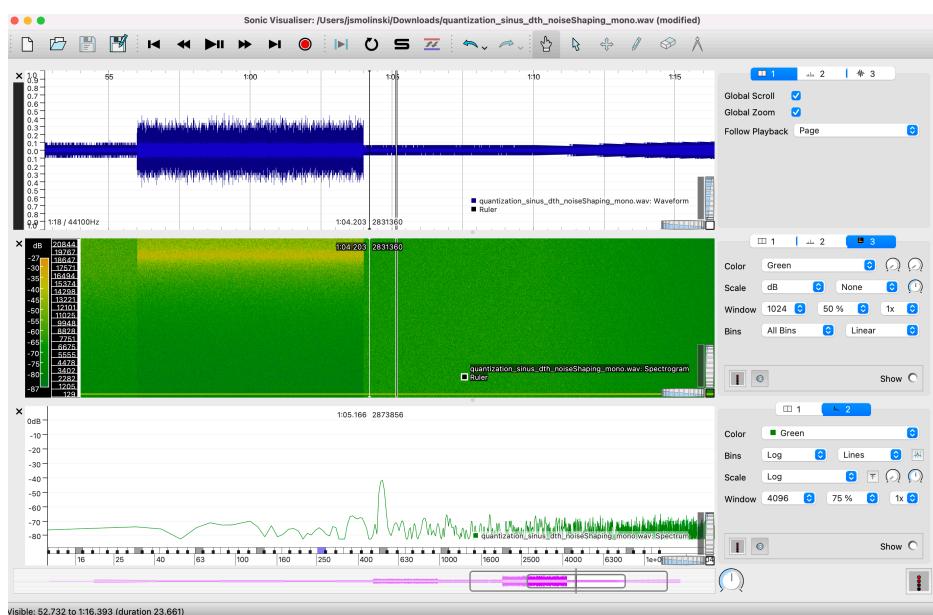
Przy początkowych fragmentach próbek można stwierdzić oczywiście że szum jest addytywny i niezależny (duża liczba kroków kwantyzacji).

W przypadku ostatniego pliku (8 bitów do 16 do 3 i do 24) mam podobne wrażenia słuchowe. Dopóki liczba kroków kwantyzacji jest odpowiednia duża, szum można traktować jako niezależny od sygnału (który tym razem z pewnością ma wygładzoną funkcję prawdopodobieństwa). Nawet przy niskiej liczbie bitów trudno (6 i mniej) to dostrzec, ale mam wrażenie, że niektóre dźwięki przynoszą więcej szumu niż inne. Co za tym idzie, przy zbyt niskiej liczbie kroków kwantyzacji błąd kwantyzacji z pewnością zależy od sygnału wejściowego.

Jeśli chodzi o dodatkową informację wysokoczęstotliwościową, jest to efekt kwantyzacji sygnału z większą liczbą bitów, niż oryginalnie sygnał został skwantowany.

## Zadanie 3

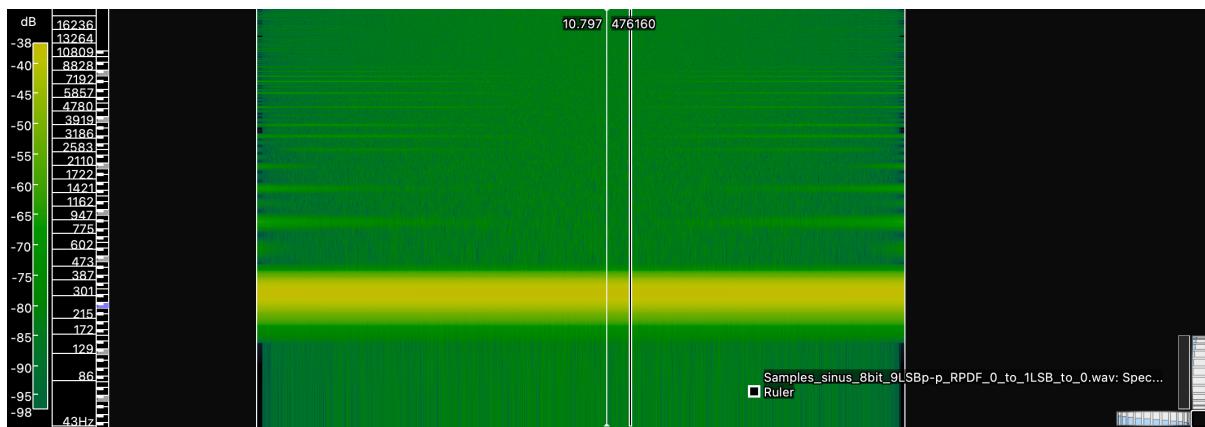
Otwarty plik z ditherem:



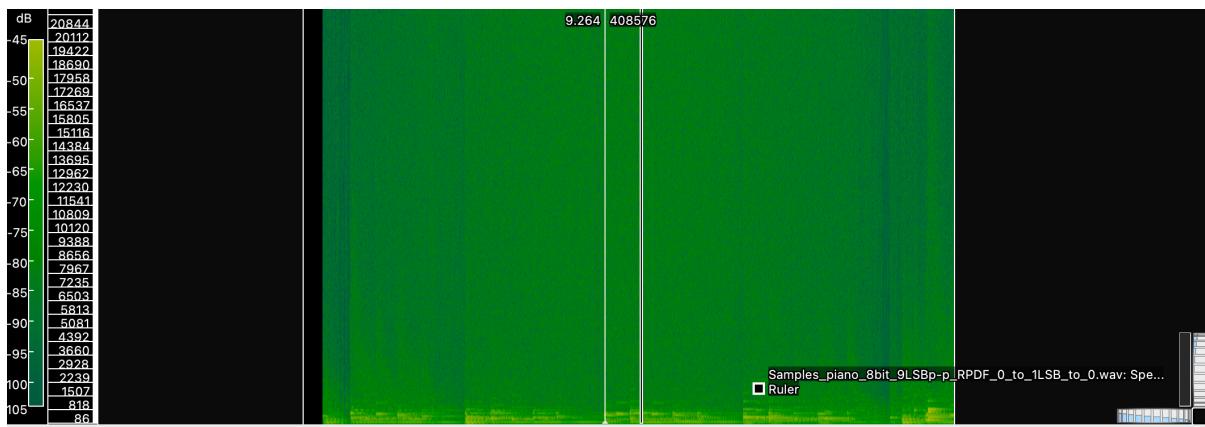
Sygnal przy zmniejszaniu jego poziomu brzmi coraz gorzej. Przy najniższym jego poziomie błąd kwantyzacji jest bardzo duży, sinus jest teraz nieprzyjemny (przypomina bardziej kwadrat). Po dodaniu pierwszego dithera słysząc bardzo silny szum. Drugi rodzaj dithera jest mniej

przyjemny, bardziej regularny (jak wiertło). Po powrocie do pierwszego dithera, rozpoczyna się kształtowanie błędu kwantyzacji. Szum tak jakby gaśnie, a sinus jest dosyć czytelny i wygładzony. Moim zdaniem przedostatni filtr H(z) daje najlepsze efekty w tym wypadku. Można zauważać że dodanie dithera i kształtowania błędu kwantyzacji, sygnał który pierwotnie był zupełnie inny i obarczony silnymi zniekształceniami, brzmi jak sygnał oryginalny o zwiększym poziomie białego szumu.

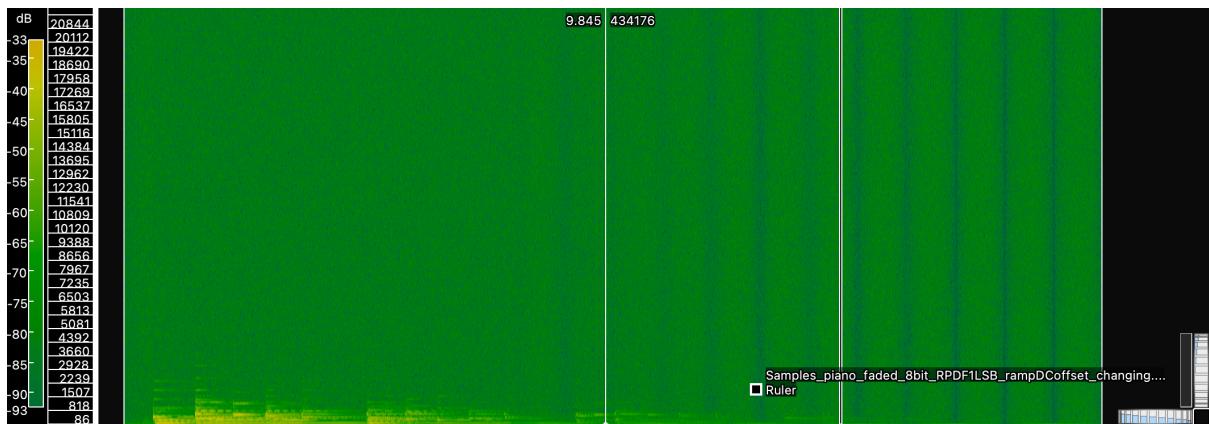
W przypadku następnego pliku, gdzie mamy do czynienia sinusem i ditherem od 0 do 1 LSB i z powrotem, widać po spektrogramie że w środkowej części nagrania (LSB=1), zakłóceń jest najmniej. Zgadza się to z moimi odczuciami, przy maksymalnej wartości LSB sygnał jest czytelny, ale zaszumiony. W wypadku innych wartości sygnał jest znacznie zniekształcony.



Dla odmiany w przypadku nagrania z takim samym ditherem, ale innym źródłem dźwięku (pianino), maksymalna wartość LSB daje zbyt duży szum, a niska skutkuje zniekształceniami i zanikaniem dźwięku (fade in/out). Średnia górna wartość wydaje się najlepszym kompromisem.



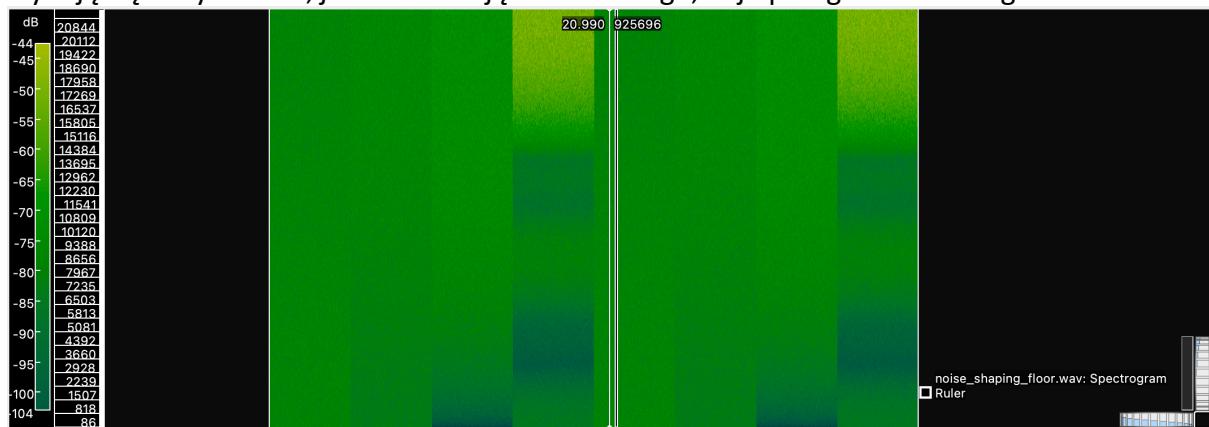
W przypadku następnej próbki z ditherem RPDF (równomiernym) i dodanym sygnałem wolnozmiennym trójkątnym, przy dużej głośności jakość jest wystarczająca. Jednak przy coraz cichszym nagraniu zaczyna być słyszalny efekt pojawiania się i znikania dźwięku (fade in/out), dużo silniejszy niż w poprzednim nagraniu. Mniej więcej od połowy przez ten efekt jakość jest niezadowalająca. Podejrzewam że to kwestia sygnału **trójkątnego** i rozkładu **równomiernego**.



Moje podejrzenie wynika stąd, że w przypadku zastosowania ditheru TPD (rozkład trójkątny) efekt ten znika i jakość jest bardzo dobra nawet przy niskiej głośności.

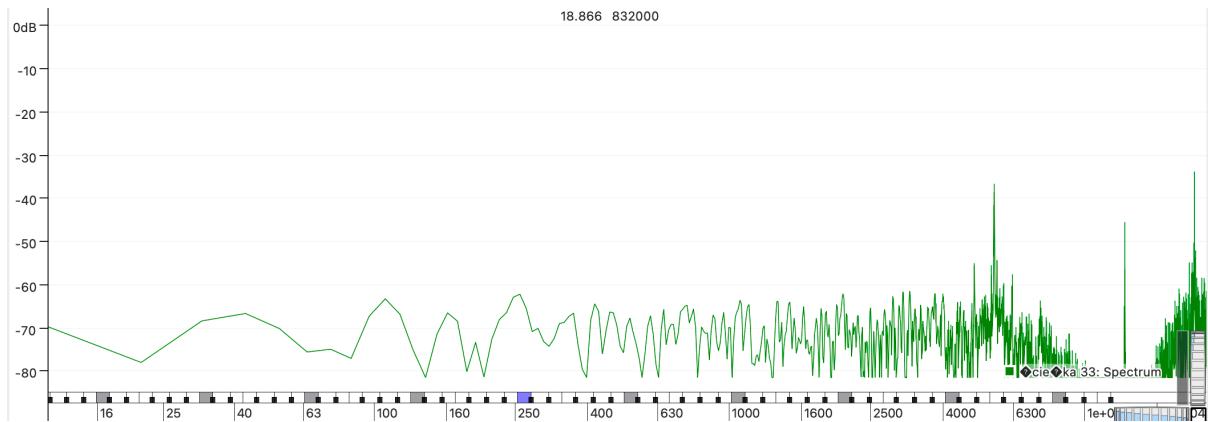


Porównałem różne rodzaje ditheru przy cyfrowej ciszy i z pewnością mogę powiedzieć że ostatni typ czyli **dither TPDF 2LSB kształtowany funkcją 9-tego rzędu ważoną charakterystykami odwrotnymi do krzywych izofonicznych** jest bardzo zaawansowaną formą ditheru. Wynika to oczywiście z samej nazwy i z bardzo przyjemnego, słabo dostrzegalnego szumu jaki generuje. Pierwszy rodzaj ditheru był dosyć ostry, dwa środkowe wydają się dosyć dobre, jednak odstają od ostatniego, najlepszego dla ludzkiego ucha ditheru.



## Jitter

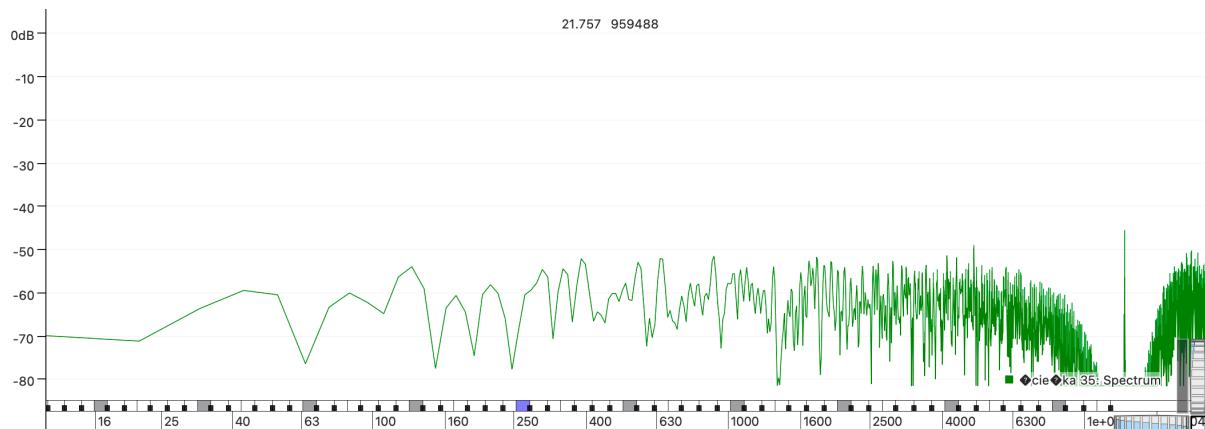
W przypadku pierwszego jittera (sygnał prostokątny) występują charakterystyczne dodatkowe harmoniczne – prążki poza głównym na 13kHz. Wrażenia dźwiękowe to ostre harmoniczne, zmieniające sinus w inny typ ostrzejszej fali – silne zniekształcenie.



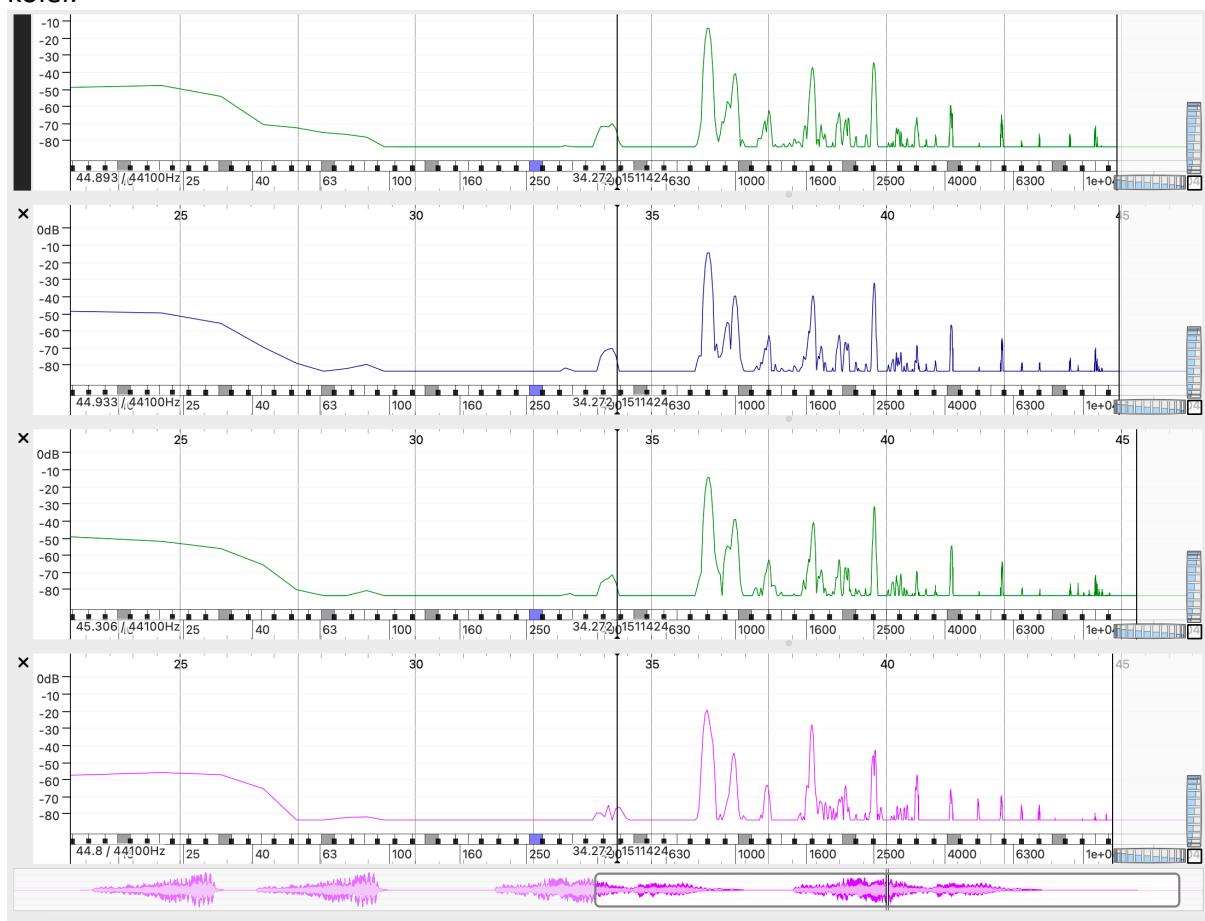
Drugi jitter wywołyany sygnałem szumowym objawia się znacznym szumem w nagraniu. Nie ma tu specyficznych składowych harmonicznych.



Trzeci jitter wywołyany przez modulację sygnałem o niskiej częstotliwości znacznie pogarsza wrażenia słuchowe. Odniosłem wrażenie pulsowania (niska częstotliwość), zniekształceń przypominających działanie popsutego wentylatora czy kompresora w lodówce. Na widmie można zauważać zwiększyony poziom szumu bez wyjątkowych składowych harmonicznych.

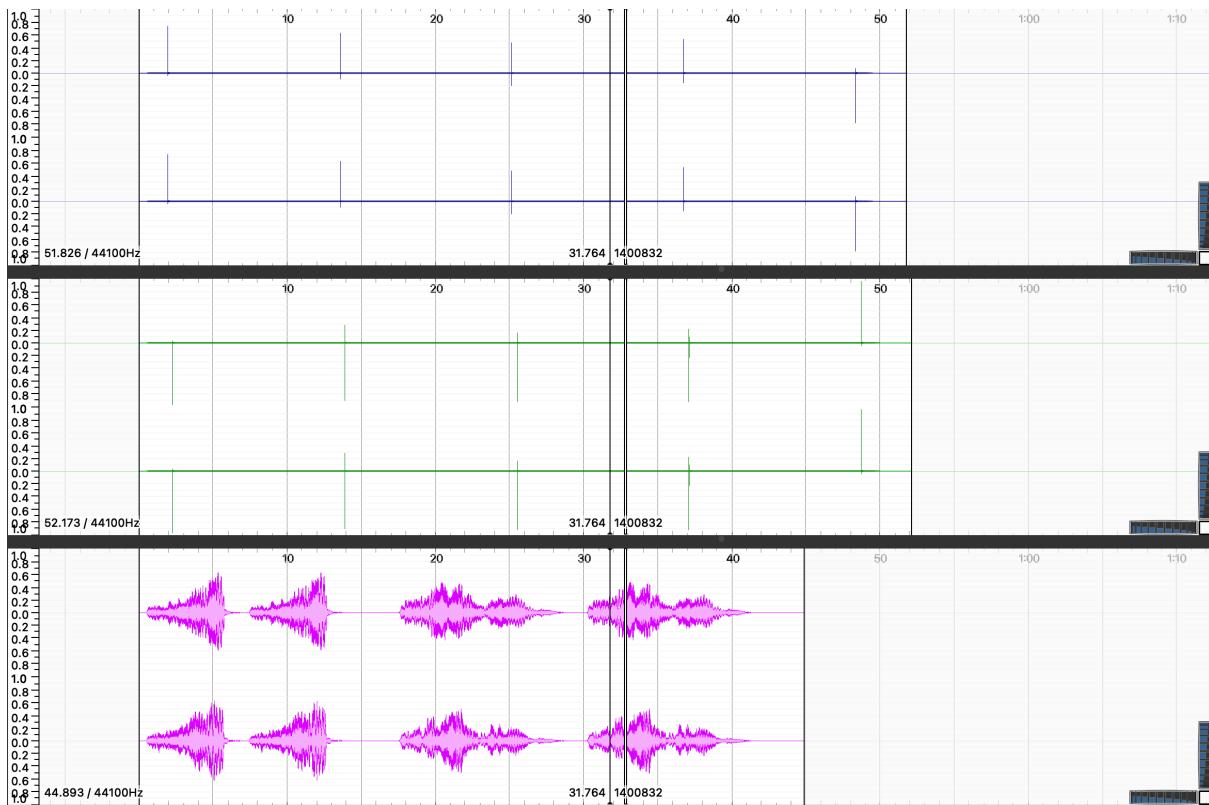


Następnie przeanalizowałem pliki jitter4/5/6/7.wav. Poniżej widmo sygnału tych plików po kolej:



Na drugim od góry wykresie widma można zauważać dodatkowe harmoniczne. Jitter ten w odbiorze był chropowaty. Trzeci wykres od góry prezentuje delikatnie zaszumione widmo – nagranie brzmiało miękko, ale można było dostrzec szum. Ostatnie nagranie sprawiało znów wrażenie delikatnego pulsowania, twardości. Tym razem wszystkie nagrania miały akceptowalną jakość i trudniej było znaleźć zniekształcenia – podejrzewam, że to w związku z dość niewysokimi częstotliwościami, porównując z pierwszymi plikami.

Ostatni podpunkt w tym zadaniu dotyczy plików sync1/2/3.wav Poniżej wykres przebiegu wszystkich trzech sygnałów:



Początkowo myślałem, że coś zrobiłem nie tak. Pierwsze dwa nagrania to regularne kliknięcia, a trzecie wydaje się być dobrej jakości. Doszło do mnie potem jednak, że kluczem do rozwiązania tego zadania jest zwrócenie uwagi na częstotliwość tonów. W pierwszym i drugim przypadku jest to sinus o wysokiej częstotliwości, a jitter czym wyższa częstotliwość tym bardziej zniekształca nagranie. Stąd podejrzewam, że najbardziej zniekształcony jest plik sync2.wav – czyli najwyższa częstotliwość. Ostatnie nagranie fletu nie jest tak zniekształcone, gdyż składa się z wielu harmonicznych, jest w szerokim spektrum, główne tony są poniżej tych w sinusach.