

# ANALIZA DŹWIĘKÓW MUZYCZNYCH

## *Dźwięk muzyczny*

---

Dźwięk muzyczny – sygnał wytwarzany przez instrument muzyczny.

Najważniejsze parametry:

- wysokość – związana z częstotliwością podstawową,
- barwa (brzmienie) – związane ze strukturą widmową sygnału
- obwiednia – zmiany głośności w czasie
- czas trwania

## Analiza czasowa i obwiednia

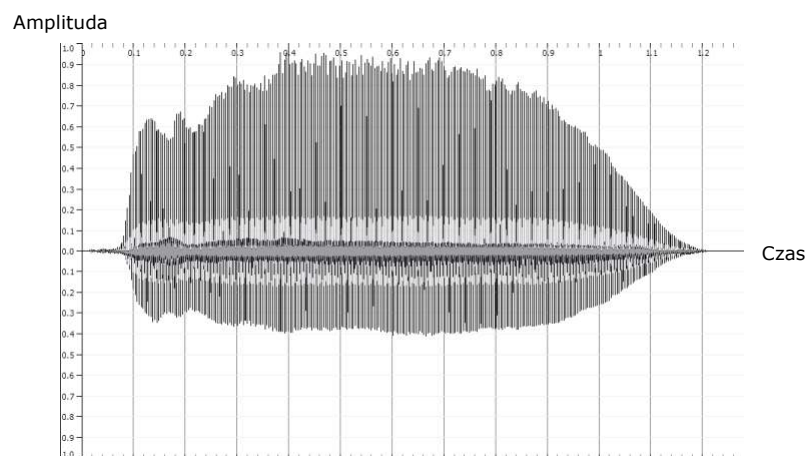
---

- Analiza czasowa – dostarcza informacji o zmienności amplitudy w czasie.
- Wynik analizy: wykres amplitudy w funkcji czasu.
- **Obwiednia sygnału** (*envelope*) – krzywa łącząca „szczyty” wykresu czasowego.
- Zmiany obwiedni odzwierciedlają zmiany głośności dźwięku.

## Analiza czasowa

---

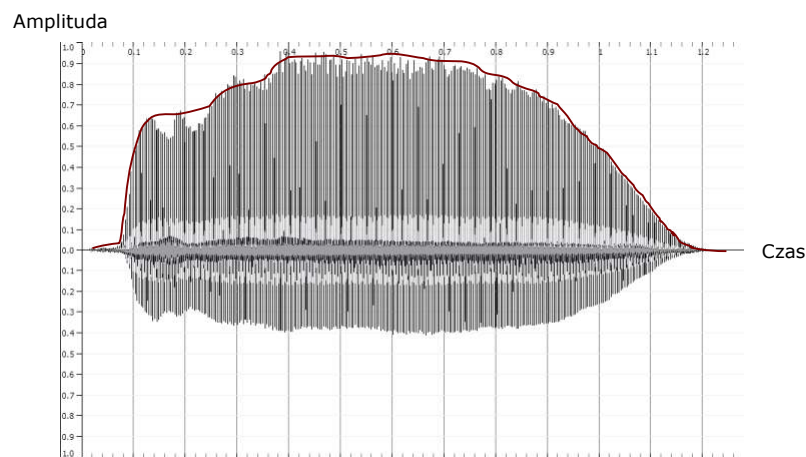
Wykres analizy czasowej dźwięku trąbki:



## Analiza czasowa - obwiednia

---

Obwiednia dźwięku trąbki:



## Wyznaczanie obwiedni

---

- Układ podążający za obwiednią (*envelope follower*) pozwala wyznaczyć obwiednię szybkozmiennego sygnału muzycznego.
- Najprostsza metoda: filtracja dolnoprzepustowa sygnału po jego konwersji na wartość bezwzględną
- Przybliżoną obwiednię można uzyskać obliczając energię lub moc sygnału
- Istnieją bardziej złożone metody, np. układ analizy wartości szczytowych (*peak hold*)

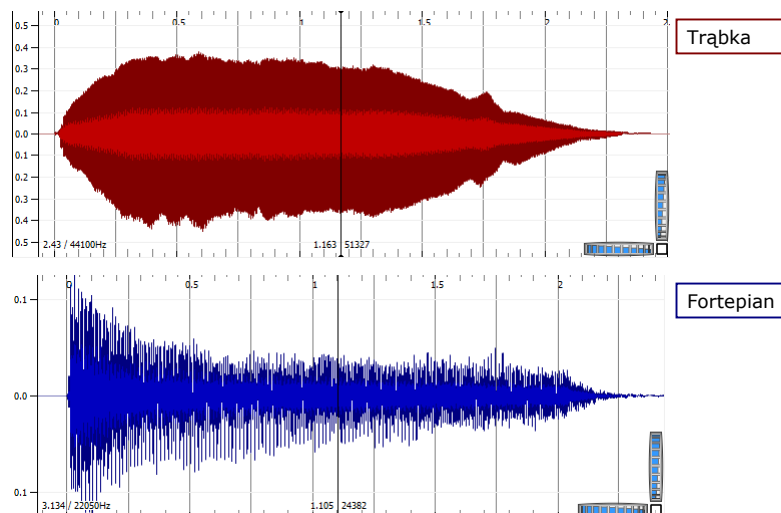
## Transjenty

W przebiegu czasowym dźwięków muzycznych możemy wyróżnić:

- stan **ustalony** (stabilny dźwięk),
- stany nieustalone – **transjenty**:
  - faza **ataku** – początek dźwięku.
  - faza **wybrzmiewania** – koniec dźwięku

Długość faz transjentowych zależy od charakteru instrumentu i sposobu gry.

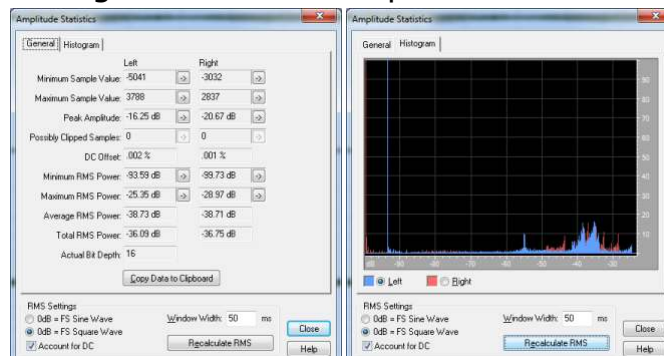
## Przykłady obwiedni



## Analiza statystyczna amplitudy

Podstawowe parametry statystyczne:

- wartość szczytowa i średnia
- moc skuteczna (*RMS power*) – minimalna, maksymalna, średnia, całkowita
- histogram rozkładu amplitud



## Analiza częstotliwościowa

- **Widmo dźwięku** – rozkład natężenia składowych dźwięku w zależności od **częstotliwości** tych składowych.
- Analiza częstotliwościowa – obliczenie widma dla wybranego fragmentu dźwięku
- Jest to widmo statyczne – dla określonego punktu w czasie
- Analiza częstotliwościowa dostarcza informacji o strukturze widmowej sygnału

## *Analiza Fouriera*

---

Przypomnienie:

- Dowolny sygnał okresowy można przedstawić w postaci sumy sygnałów sinusoidalnych o odpowiednich amplitudach i częstotliwościach (szereg Fouriera).
- Dla sygnału cyfrowego – dyskretne przekształcenie Fouriera (DFT) i jego „szybka” wersja – FFT (bloki o długości  $2^N$ )

Uwaga – terminologia:

- transformacja Fouriera – obliczanie widma
- transformata Fouriera – wynik przekształcenia

## *Analiza częstotliwościowa*

---

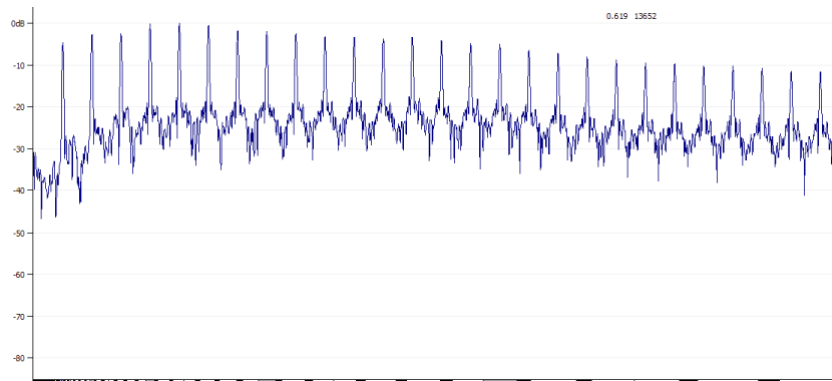
Jak obliczamy widmo:

- „wycinamy” z sygnału blok próbek za pomocą funkcji okna (jesteśmy w dziedzinie czasu),
- obliczamy przekształcenie Fouriera (FFT) - przechodzimy do dziedziny częstotliwości
- wynik przedstawiamy na wykresie: amplituda (lub poziom) w funkcji częstotliwości
- uzyskujemy widmo statyczne w wybranym punkcie czasu

## Analiza częstotliwościowa

---

Widmo statyczne dźwięku trąbki (stan ust.)



## Widmo dźwięków muzycznych

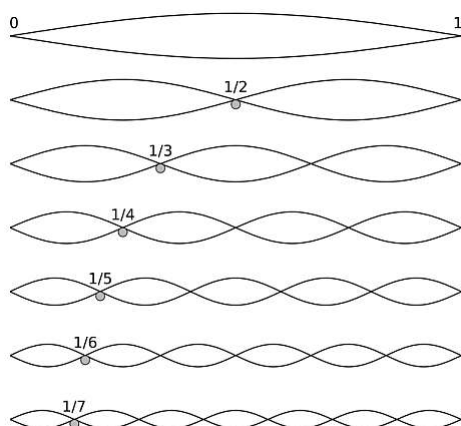
---

Większość dźwięków muzycznych charakteryzuje się następującymi cechami:

- jest **prążkowe** – dominują wyraźne maksima,
- jest **harmoniczne** – prążki tworzą szereg harmoniczny
- częstotliwość pierwszego prążka ( $f_0$ ) to **częstotliwość podstawowa** – zwykle decyduje o wysokości dźwięku
- wyższe prążki to **harmoniczne**: pierwsza ( $2f_0$ ), druga ( $3f_0$ ), trzecia ( $4f_0$ ), itd.  
– decydują o **barwie** dźwięku.

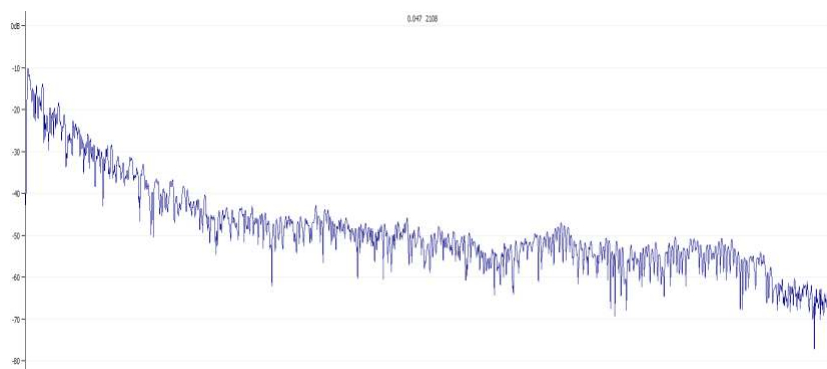
## Widmo dźwięków muzycznych

Dlaczego właśnie takie widmo?



## Dźwięki perkusyjne

Dźwięki typu perkusyjnego mają widmo szumowe.





### Analiza czasowo-częstotliwościowa

- Widmo statyczne jest obserwowane w wybranym punkcie skali czasu.
- Widmo dźwięków muzycznych jest dynamiczne - zmienne w czasie („ewoluuje”).
- Aby uzyskać informacje o barwie dźwięku, potrzebujemy zbadać widmo w różnych punktach skali czasu, dla całego dźwięku.
- Analiza czasowo – częstotliwościowa: połączenie wyników analizy częstotliwościowej dla kolejnych punktów czasu.

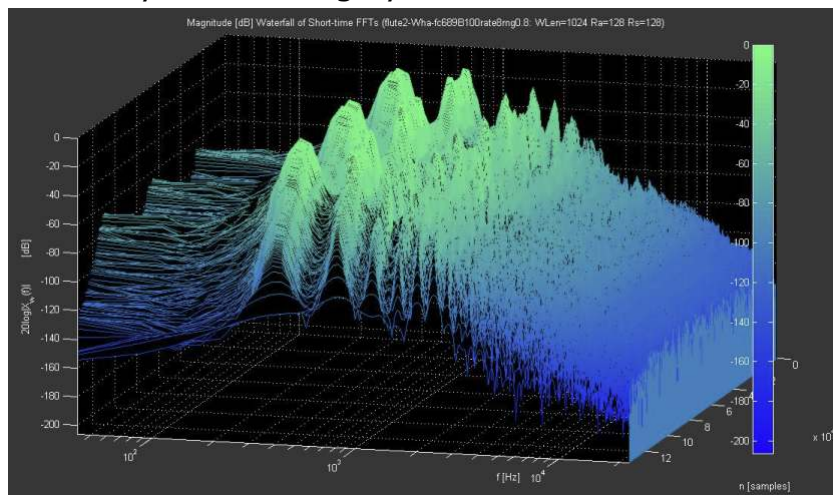
### Analiza czasowo-częstotliwościowa

Analiza FFT czasowo-częstotliwościowa:

- Dzielimy oś czasu na przedziały
- Dla każdego przedziału obliczamy statyczne widmo
- Łączymy wszystkie wyniki analizy
- Wynik przedstawiamy na trójwymiarowym wykresie: czas – częstotliwość – amplituda.
- Jak przedstawić wykres 3D na płaszczyźnie?
  - wykres typu *waterfall*
  - spektrogram

## Wykres typu waterfall

Ilustracyjny, dobrze widoczna zmienność widma, mało czytelne szczegóły



## Spektrogram

Najczęściej używany tryb wizualizacji wyników analizy czasowo-częstotliwościowej:

- oś pozioma: czas
- oś pionowa: częstotliwość
- poziom składowych (lub amplituda) przedstawiana w postaci barwy lub jasności

Przekrój pionowy w wybranym punkcie czasu daje nam wynik analizy częstotliwościowej

## Spektrogram

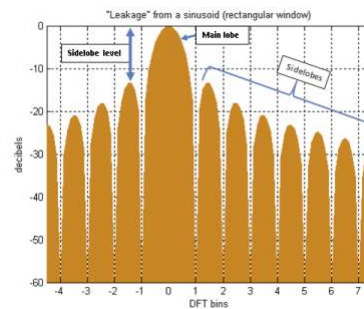


## STFT

- Wynik analizy Fouriera jest słuszny jeżeli analizujemy **jeden okres** sygnału
- To założenie nie jest spełnione, jeżeli wycinamy fragment dźwięku do analizy
- To, co robimy w rzeczywistości to **STFT** - krótkookresowa transformacja Fouriera
- Liczymy FFT dla sygnału przemnożonego przez **funkcję okna**, niezerową tylko w analizowanym fragmencie

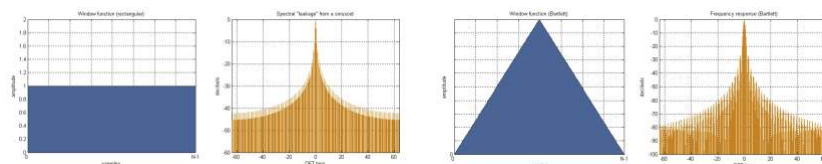
## Okno prostokątne

- Załóżmy, że po prostu wycinamy fragment dźwięku do analizy,
- czyli stosujemy okno prostokątne.
- W efekcie uzyskujemy widmo zniekształcone przez przecieki widmowe.
- Aby je zmniejszyć, stosuje się różne typy funkcji okienkujących (*windowing*)

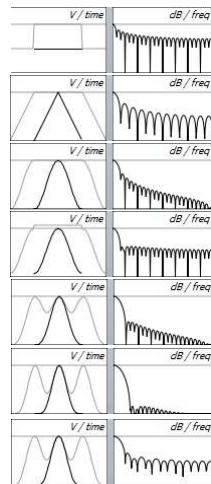


## Okno trójkątne

- Najprostszy sposób zmniejszenia przecieków widmowych: okno trójkątne (Bartletta).
- Mniejsze przecieki (listki boczne), ale zniekształcamy sobie analizowany sygnał.
- Są lepsze funkcje okna, za cenę bardziej złożonych obliczeń przy okienkowaniu.
- Nie ma okna „najlepszego” (uniwersalnego)



## Funkcje okna



Prostokątne

Trójkątne (Bartletta)

Hanna („hanning”)

Hamminga

Blackmana

Blackmana-Harrisa

Parzena

## Rozdzielczość częstotliwościowa

Jaki rozmiar okna analizy?

- Musi być  $2^N$ , gdy stosujemy FFT.
- Weźmy więc 512 próbek.
- W wyniku FFT dostajemy tyle samo wartości, pokrywających zakres od  $-F_s/2$  do  $F_s/2$ .
- Każda wartość FFT wpada do jednego z 512 przedziałów o szerokości  $F_s/512$
- Dla  $F_s = 44,1$  kHz daje to 86,13 Hz
- Tyle wynosi **rozdzielczość częstotliwościowa** analizy – nie rozróżnimy dwóch składowych oddległych o mniej niż 86,13 Hz.

## *Rozdzielczość czasowa*

---

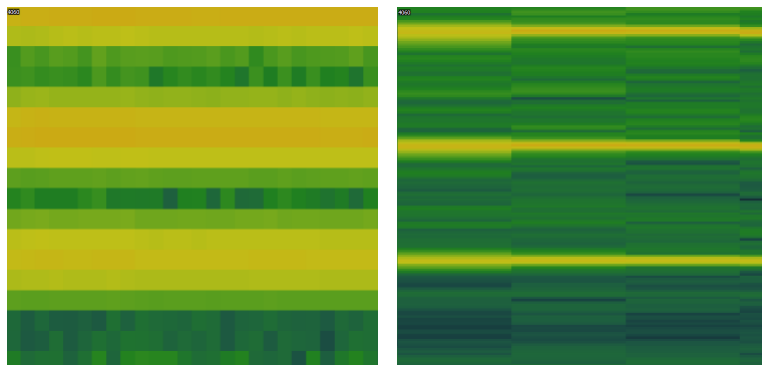
Zatem większe okno jest lepsze?

- Weźmy 4096 próbek.
- Rozdzielczość częstotliwościowa: 10,77 Hz (8 razy większa)
- Ale teraz okno pokrywa zakres 92,88 ms (wcześniej 11,61 ms)
- Nie rozróżnimy zdarzeń czasowych w odstępie mniejszym niż 92,88 ms.
- Poprawiliśmy rozdzielczość częstotliwościową, pogorszyliśmy **rozdzielczość czasową**.

## *Rozdzielczość czasowo-częstotliwościowa*

---

Spektrogram dla okna 512 i 4096 próbek, wyłączone zakładkowanie



## Zakładkowanie

---

Zakładkowanie (*overlapping*) – sąsiednie przedziały czasowe analizy widmowej pokrywają się (stosuje się najczęściej: 25%, 50%, 75% zakładki).

Korzyści:

- zwiększenie rozdzielczości czasowej
- zmniejszenie zniekształceń sygnału spowodowane funkcją okna

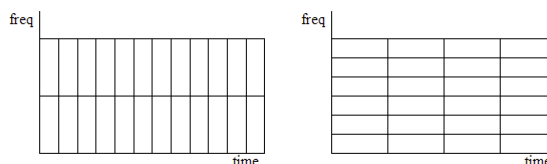
Wada: zwiększenie czasu analizy (więcej przedziałów).

## Rozdzielczość czasowo-częstotliwościowa

---

Wnioski dla analizy STFT:

- trzeba znaleźć kompromis między rozdzielczością czasowo-częstotliwościową
- nie ma uniwersalnych ustawień
- dlatego stosuje się inne metody analizy, nie bez takich ograniczeń, np. analizę falkową.



### *STFT - praktyczne rady*

---

- Rozmiar okna: 512, 1024, 2048 lub 4096
- Zakładkowanie: 50% lub 75%
- Okno: Blackmana-Harrisa, Blackmana, Hanna lub Hamminga
- Większe okno gdy widmo ma wiele składowych i jest mało zmienne
- Mniejsze okno (np. 512) i większe zakładkowanie gdy widmo jest mocno zmienne (np. w transjentach dźwięku)

### *Spektrogram a transjenty*

---

Obserwując spektrogram można wyznaczyć czasy trwania stanów transjentowych:

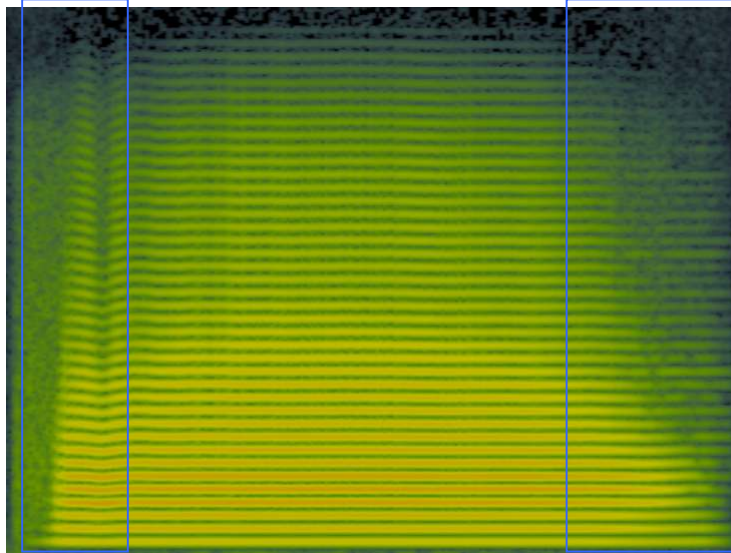
- faza ataku: stopniowe budowanie widma od niższych do wyższych składowych
- stan ustalony – stabilne widmo, może się zmieniać na skutek artykulacji
- faza wybrzmiewania: stopniowe zanikanie widma od wyższych do niższych składowych

Spektrogram daje więcej informacji o stanach transjentowych niż sama analiza czasowa



## *Spektrogram a transjenty*

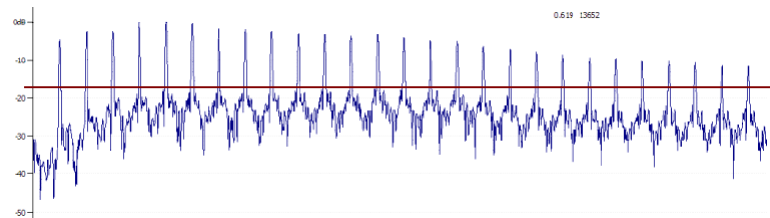
---



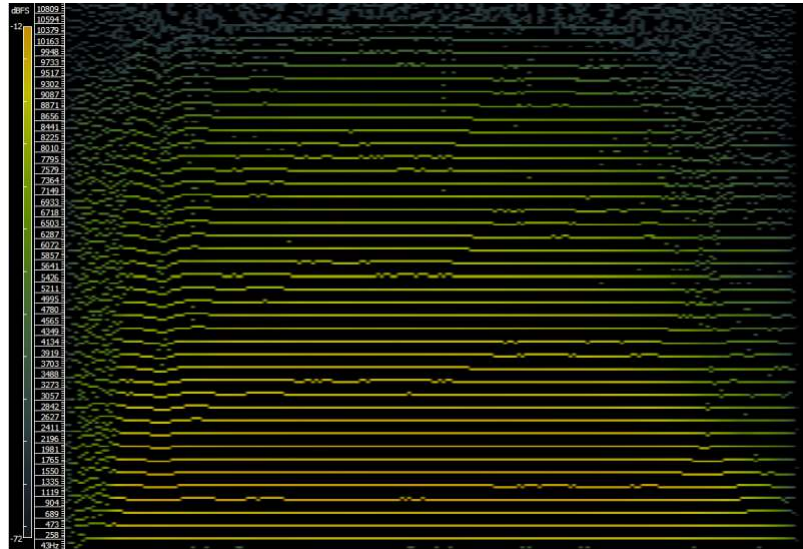
## *Analiza widma progowanego*

---

Możemy uwzględnić w analizie tylko wartości widma o poziomie większym niż pewien próg.  
Pomijamy szumy i małe składowe, lepiej uwidaczniamy prążki widmowe.



## *Analiza widma progowanego*

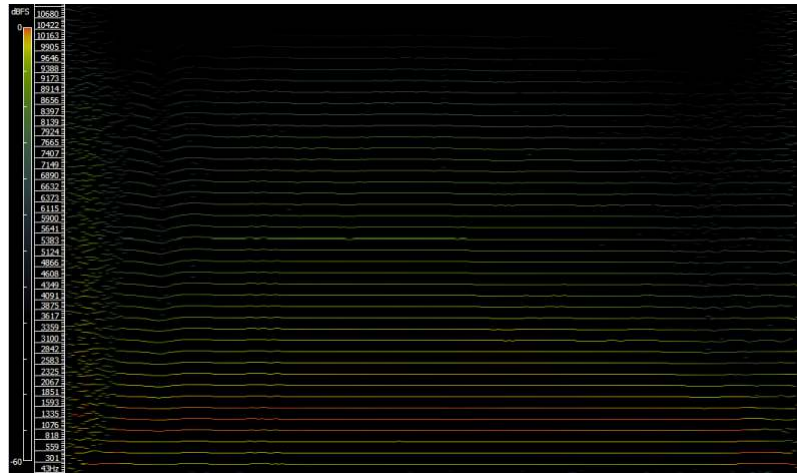


## *Analiza wartości szczytowych widma*

- Wyszukiwane są szczyty wyraźnych prążków widmowych
- Na spektrogramie uwidaczniamy poziomy tylko tych maksimów lokalnych
- Spektrogram pokazuje zmienność lokalnych maksimów widmowych
- Użyteczne dla sygnałów muzycznych o widmie harmonicznym prążkowym

## Analiza wartości szczytowych widma

---



## Analiza McAulay-Quatieri (MQ)

---

Analiza MQ stanowi rozwinięcie idei progowanego spektrogramu:

- analiza STFT w zakładkowanych ramkach
- odrzucanie składowych mniejszych niż próg
- łączenie wyników z kolejnych ramek
- wyznaczanie ciągłych ścieżek dla składowych widmowych
- odrzucanie krótkotrwałych ścieżek
- wynikiem analizy MQ jest zestaw ścieżek reprezentujących zmienność składowych widmowych, np. harmoniczných

## Analiza falkowa (wavelet)

Tak jak analiza Fouriera reprezentuje sygnał jako sumę harmoniczných o różnych częstotliwościach, analiza falkowa dokonuje dekompozycji sygnału na sumę **falek** o różnych **skalach**.

Falka jest sygnałem o ograniczonym okresie i zerowej wartości średniej.

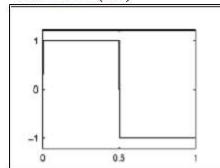
Transformacje falkowe:

- CWT – ciągła transformacja falkowa
- DWT – dyskretna transformacja falkowa  
– najczęściej używana

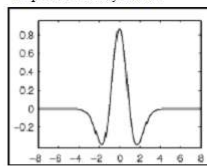
## Analiza falkowa

Typowe kształty falek

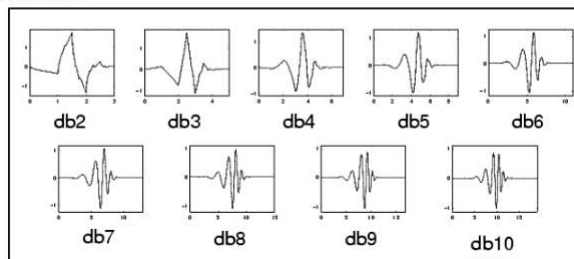
Falka Haara(db1)



Kapelusz Meksykański



Falki Daubechies

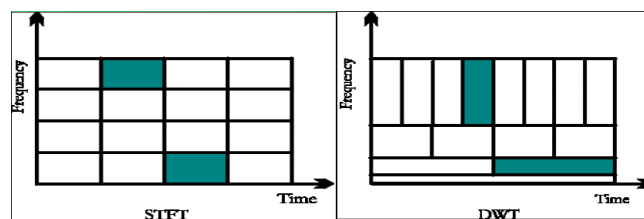


## Analiza falkowa

---

W porównaniu z analizą Fouriera:

- zmienna rozdzielczość czasowo-częstotliwościowa, większa dokładność analizy czasowej dla składowych o różnych częstotliwościach
- rozdzielczość częstotliwościowa ograniczona do liczby skal falek



## Analiza falkowa

---

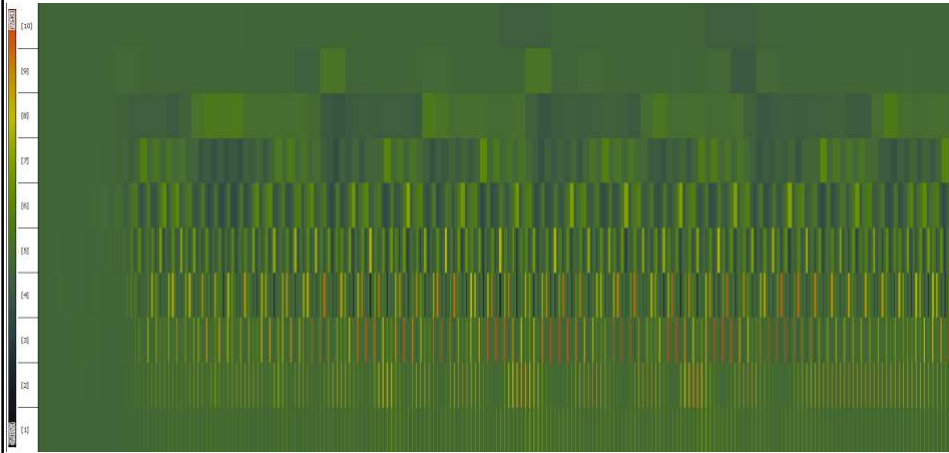
Wynik analizy falkowej DWT przedstawia się najczęściej w formie skalogramu:

- oś pozioma – czas
- oś pionowa – skala falki
- kolor – wartość współczynnika falki

Skalogram sam z siebie nie jest szczególnie „informacyjny”. Zaletą analizy falkowej jest dostarczenie dokładniejszych niż w przypadku STFT informacji o sygnale, np. do celów parametryzacji.

### *Analiza falkowa - skalogram dźwięku trąbki*

---



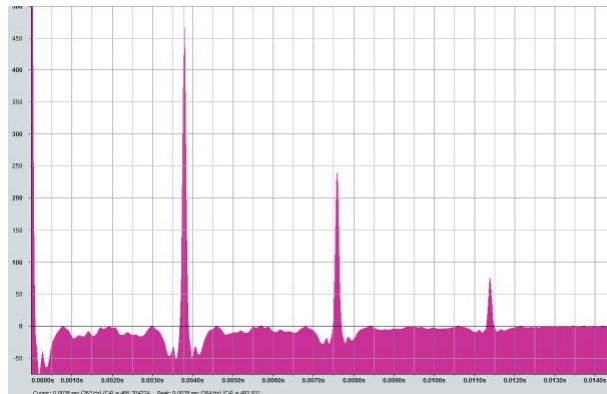
### *Analiza autokorelacyjna*

---

- Autokorelacja – miara zgodności sygnału z przesuniętą w czasie kopią tego samego sygnału
- Wynik wyznaczany dla różnych wartości przesunięcia
- W przypadku sygnałów harmoniczných, autokorelacja jest największa gdy przesunięcie jest wielokrotnością okresu
- Analiza ta jest jedną z metod wyznaczania częstotliwości podstawowej sygnału

## Analiza autokorelacyjna

Wynik analizy dla jednej ramki czasowej.  
Pierwsze maksimum odpowiada częstotliwości podstawowej 264 Hz.



## Analiza autokorelacyjna

Autokorelogram – podobny do spektrogramu,  
uwidacznia zmienność harmoniczną



## *Analiza fraz muzycznych*

---

Metody analizy dla fraz muzycznych, nie pojedynczych dźwięków:

- śledzenie wysokości (*pitch tracking*)
  - analiza tempa muzycznego
  - wykrywanie początków nut (*onset detection*)
- i wiele innych.

Zagadnienia śledzenia częstotliwości podstawowej (wysokości dźwięku) będą tematem osobnego wykładu.