Algorytmy detekcji częstotliwości podstawowej

Plan

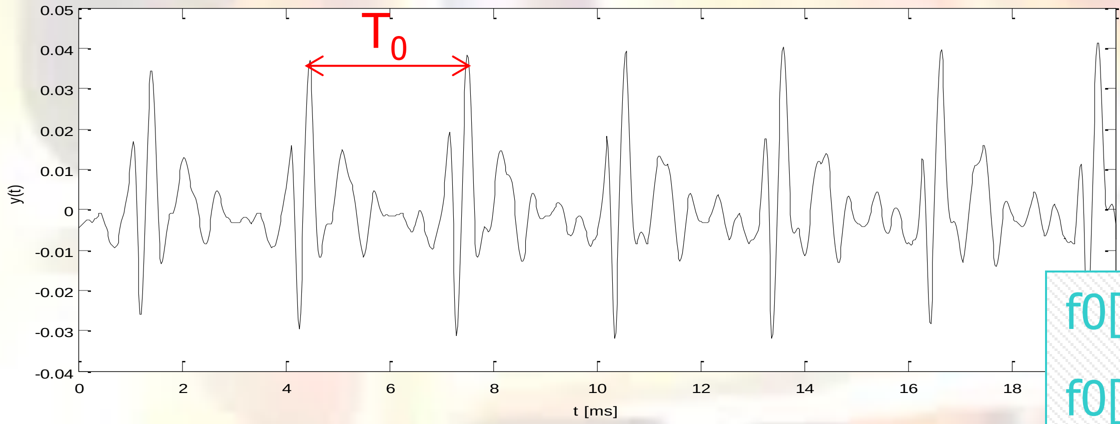
- Definicja częstotliwości podstawowej
- Wybór ramki sygnału do analizy
- Błędy oktawowe i dokładnej estymacji
- Metody detekcji częstotliwości podstawowej
  - czasowe
  - widmowe
- Realizacja przykładowego algorytmu

Częstotliwość podstawowa

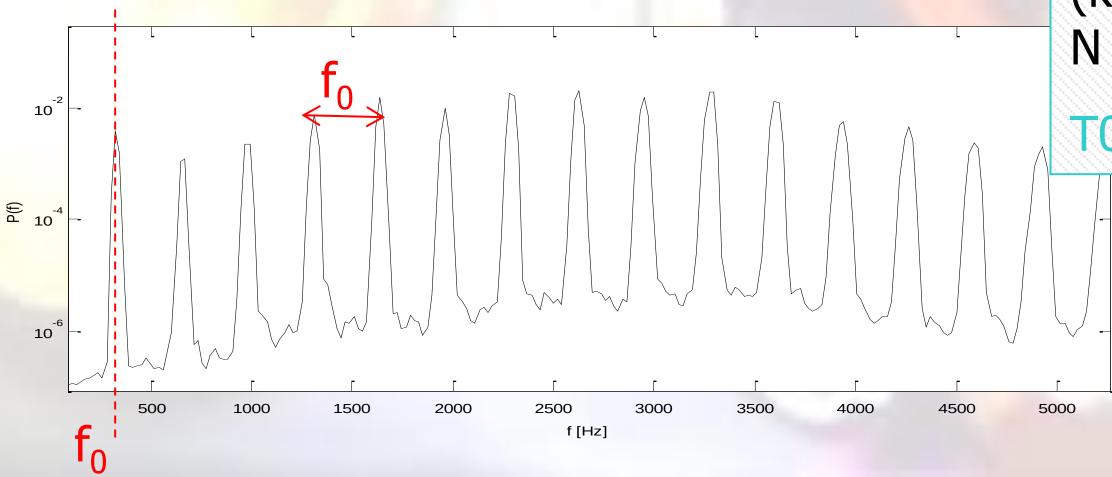
- W instrumencie muzycznym, dla dowolnie dobranej długości struny czy długości kolumny drgającego powietrza istnieje naturalny dźwięk odpowiadający tej długości, złożony z szeregu tonów prostych. Najniższy ton występujący w takim dźwięku nazywany jest **główną składową harmoniczną**, a odpowiadająca mu częstotliwość – **częstotliwością podstawową** lub wysokością dźwięku.
- Amplituda głównej składowej harmonicznej nie musi być największą spośród wszystkich harmonicznych.
- Na barwę dźwięku instrumentu muzycznego decydujący wpływ mają wzajemne relacje między kolejnymi składowymi harmonicznymi.

Częstotliwość podstawowa

- Postać czasowa



- Widmo



f0[Hz]*T0[s]=1

f0[Hz]=f0[k]*fs[Hz]/N

(k – indeks DFT,
N – długość DFT)

T0[s]=T0[smpl]/fs[Hz]

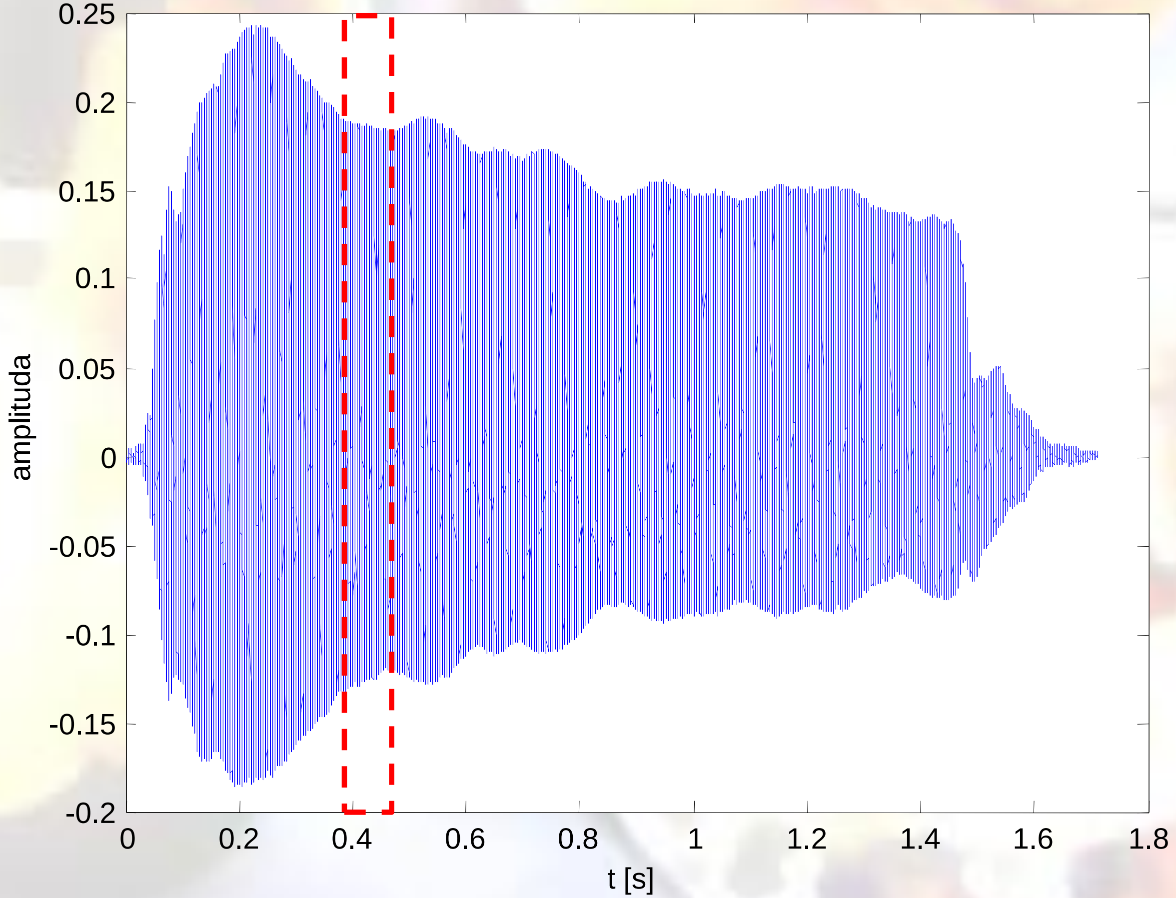
Przykłady zastosowań detekcji częstotliwości podstawowej

- Określanie częstotliwości kolejnych składowych harmonicznych i śledzenie ich zmian czasowych
- Wyznaczanie parametrów widmowych dźwięku
- Klasyfikacja instrumentów
- Automatyczna transkrypcja linii melodycznej do kodu MIDI
- Separacja dźwięków instrumentów muzycznych z nagrań polifonicznych
- Indeksacja i automatyczne wyszukiwanie nagrań

Zasady wyboru ramki sygnału do analizy

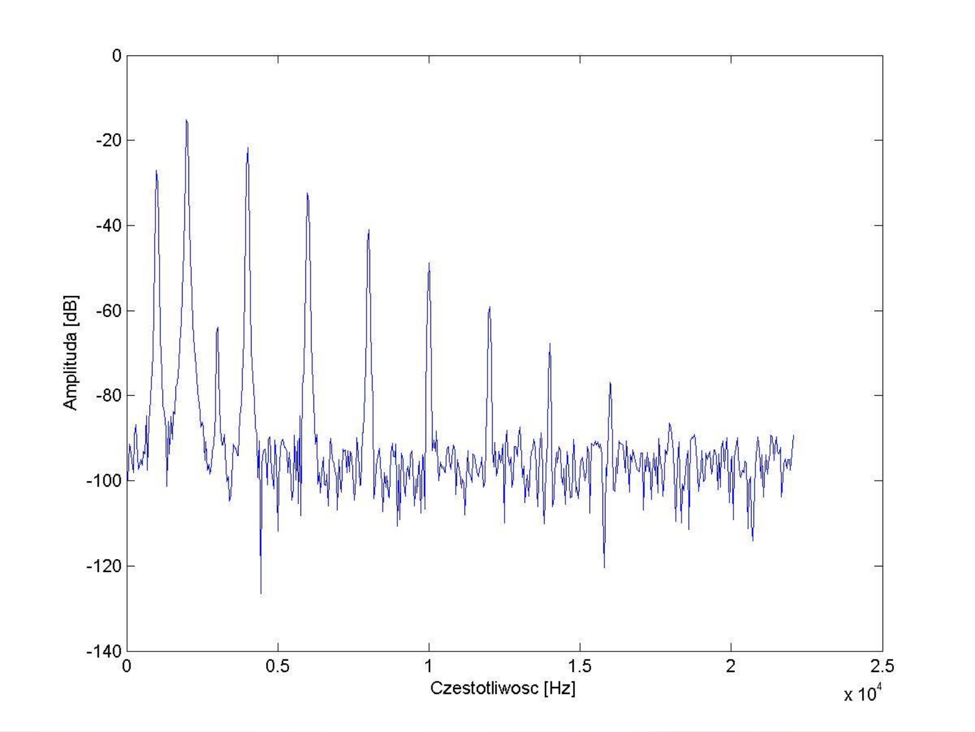
- Długość analizowanej ramki sygnału zależy od wybranej metody detekcji częstotliwości podstawowej, od charakterystyki analizowanego dźwięku oraz od oczekiwanej dokładności wyników
- W ogólności należy wybierać możliwie krótką ramkę sygnału, w której analizowany sygnał jest niezmienny (np. faza „Sustain” w modelu obwiedni dźwięku ADSR przy analizie dźwięków pojedynczych, izolowanych instrumentów muzycznych)
- Dla większości metod, analizowana ramka sygnału powinna zawierać co najmniej kilka okresów sygnału
- W celu poprawy rezultatów można analizować kilka różnych ramek dla tego samego sygnału

Zasady wyboru ramki sygnału do analizy



Błędy oktawowe

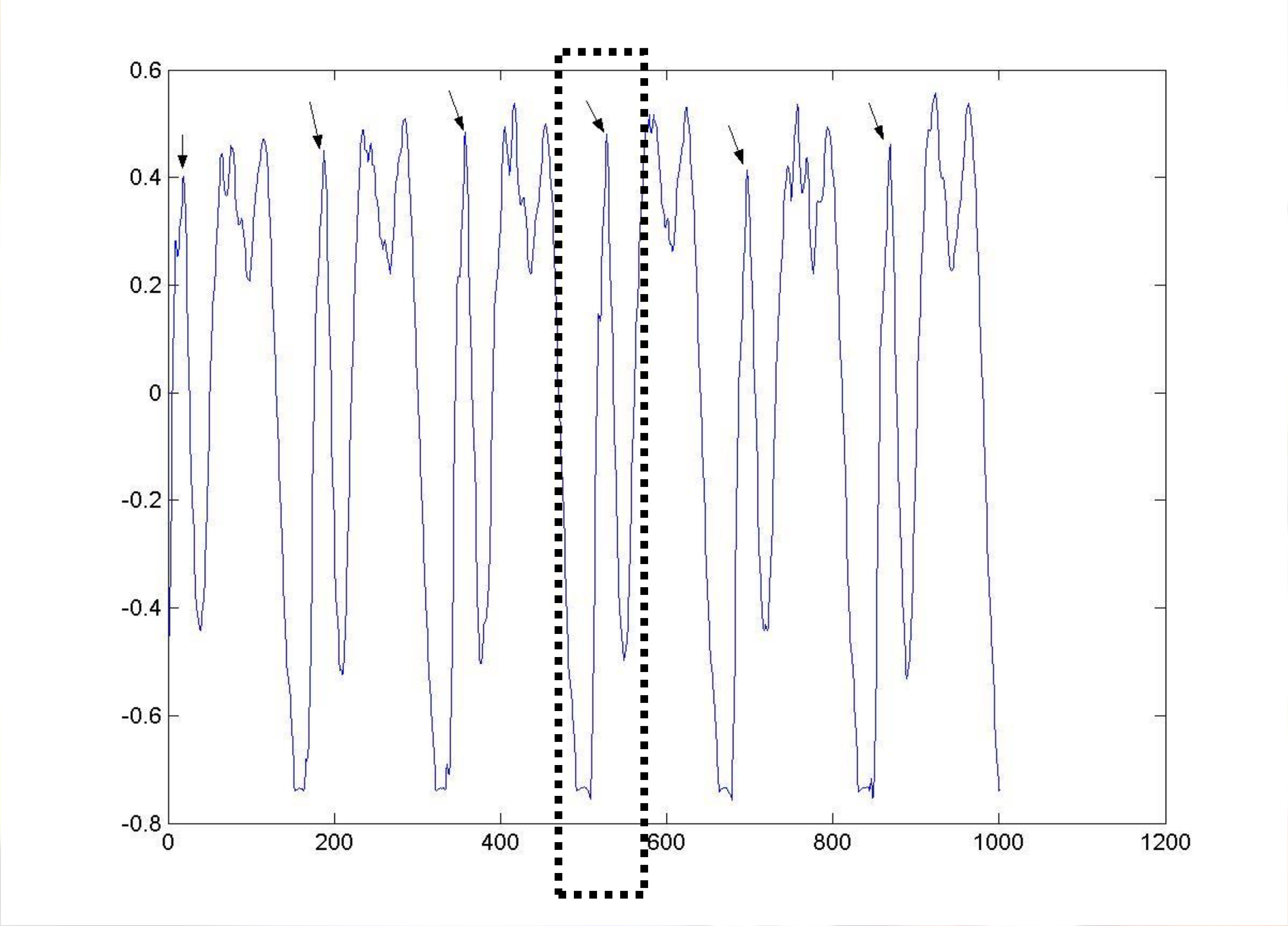
- Błędy oktawowe związane są z trudnością wyznaczenia okresu sygnału (w analizie w dziedzinie czasu), bądź z problemami związanymi z określeniem rzędu składowych harmonicznych wykrytych w widmie.



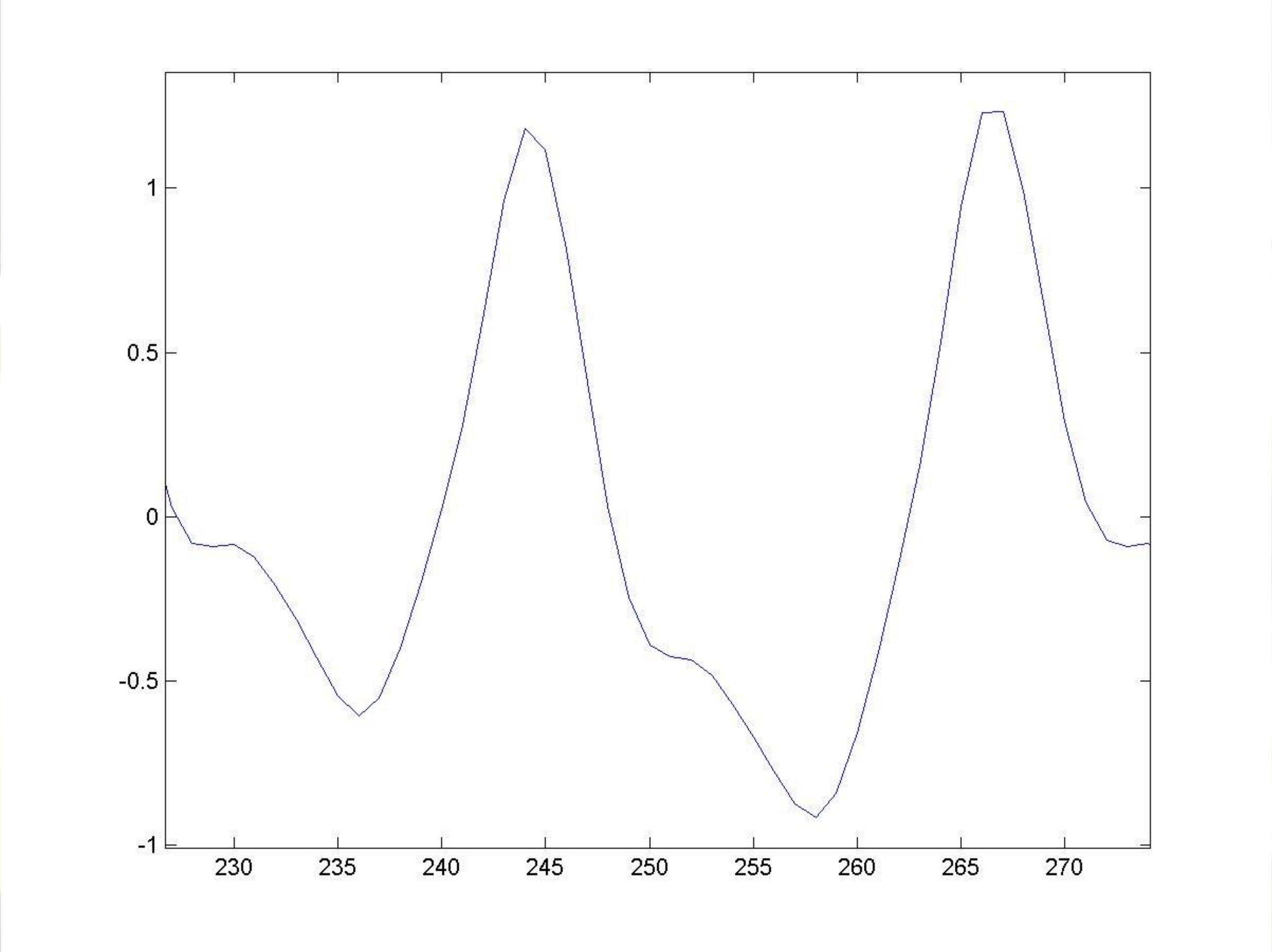
Błędy dokładnej estymacji

- Błędy te wynikają z dyskretnej postaci czasowej sygnałów oraz szumu zakłócającego analizowane sygnały.
- W przypadku reprezentacji czasowych nie zawsze próbka reprezentująca maksimum wypada w rzeczywistym maksimum fali przebiegu.
- W przypadku analizy widmowej piki widma nie zawsze reprezentują częstotliwość. Niedokładne wyznaczenie maksimum składowych harmonicznych wpływa na błąd estymacji.

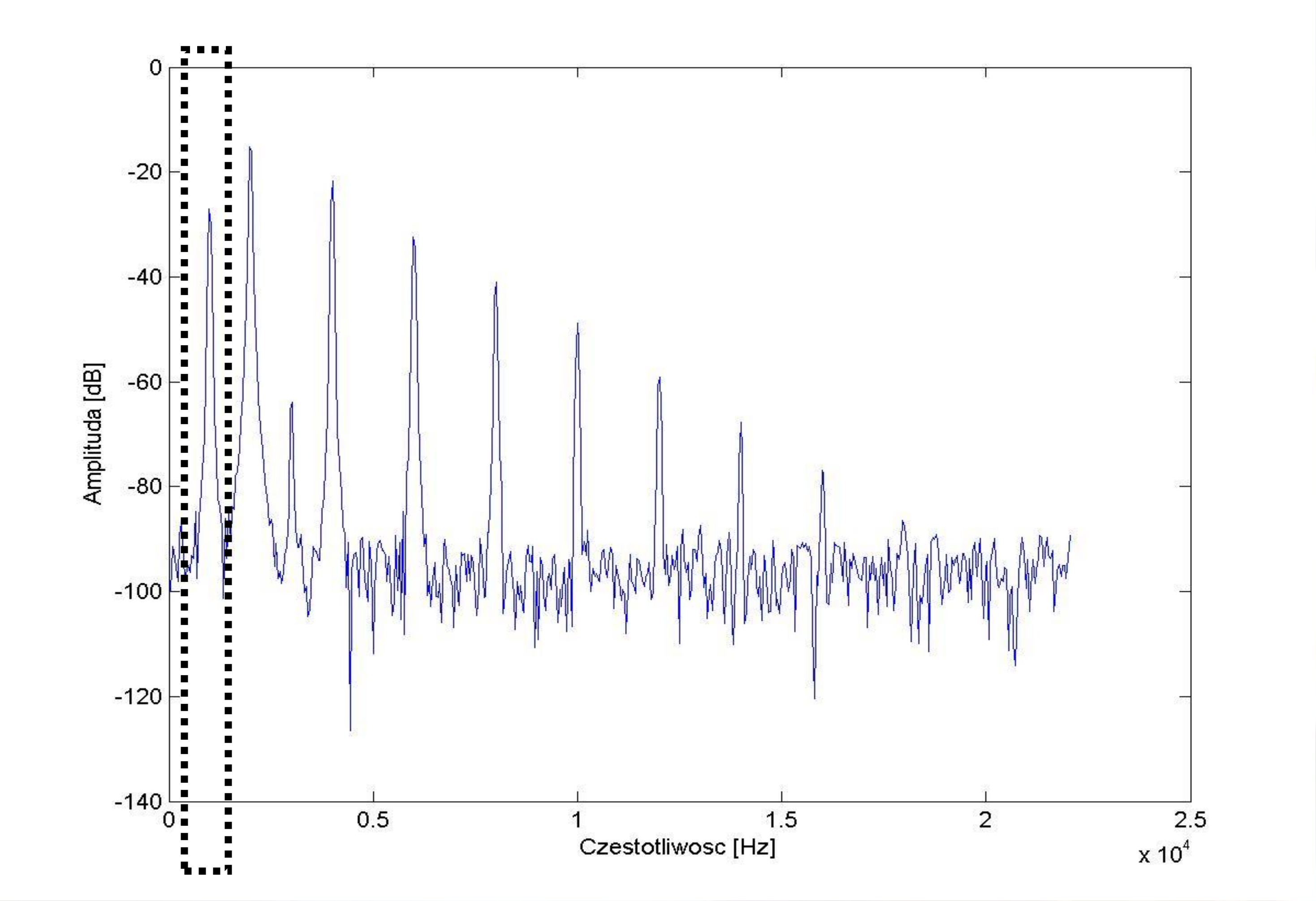
Błędy dokładnej estymacji



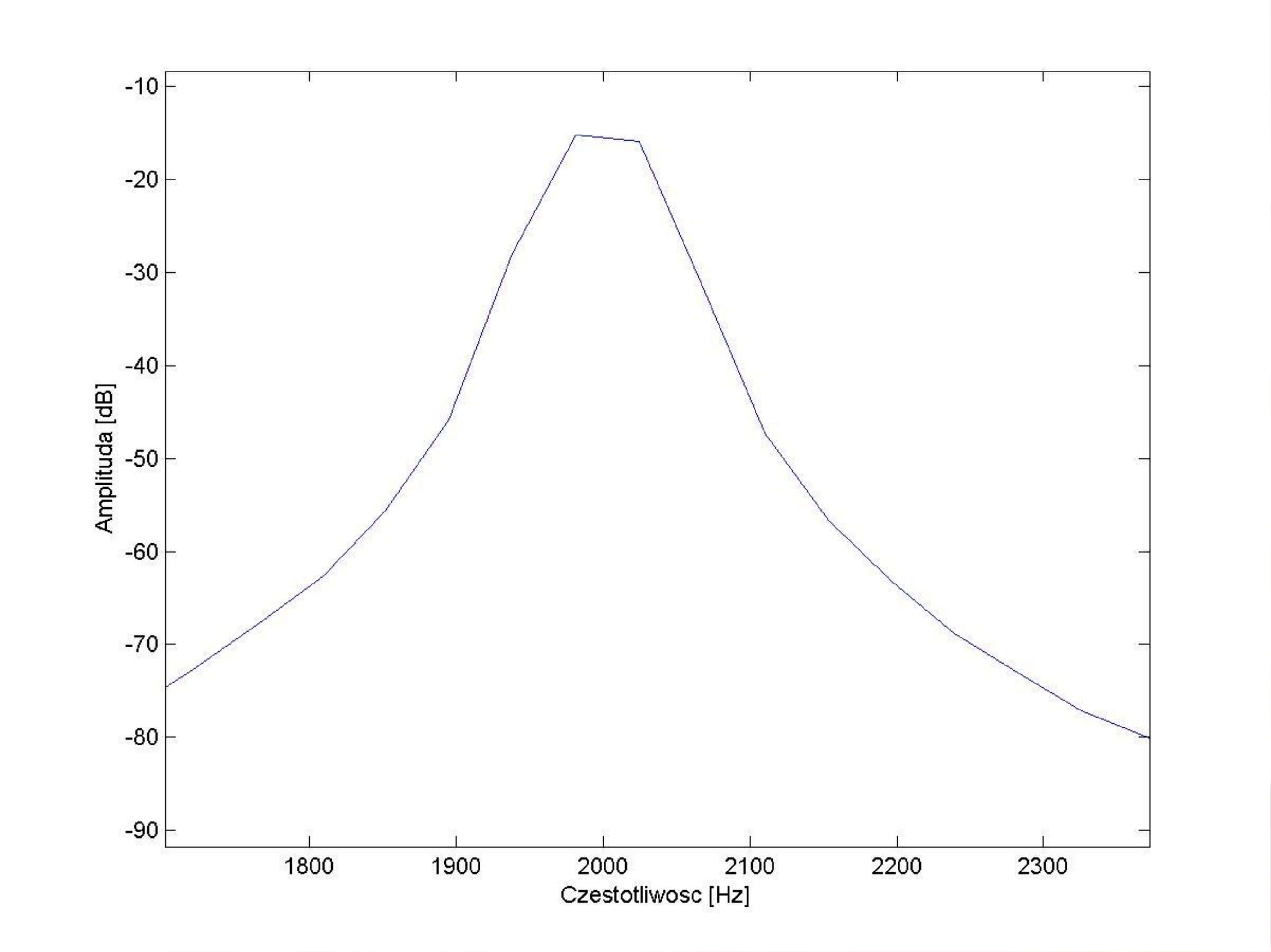
Błędy dokładnej estymacji



Błędy dokładnej estymacji



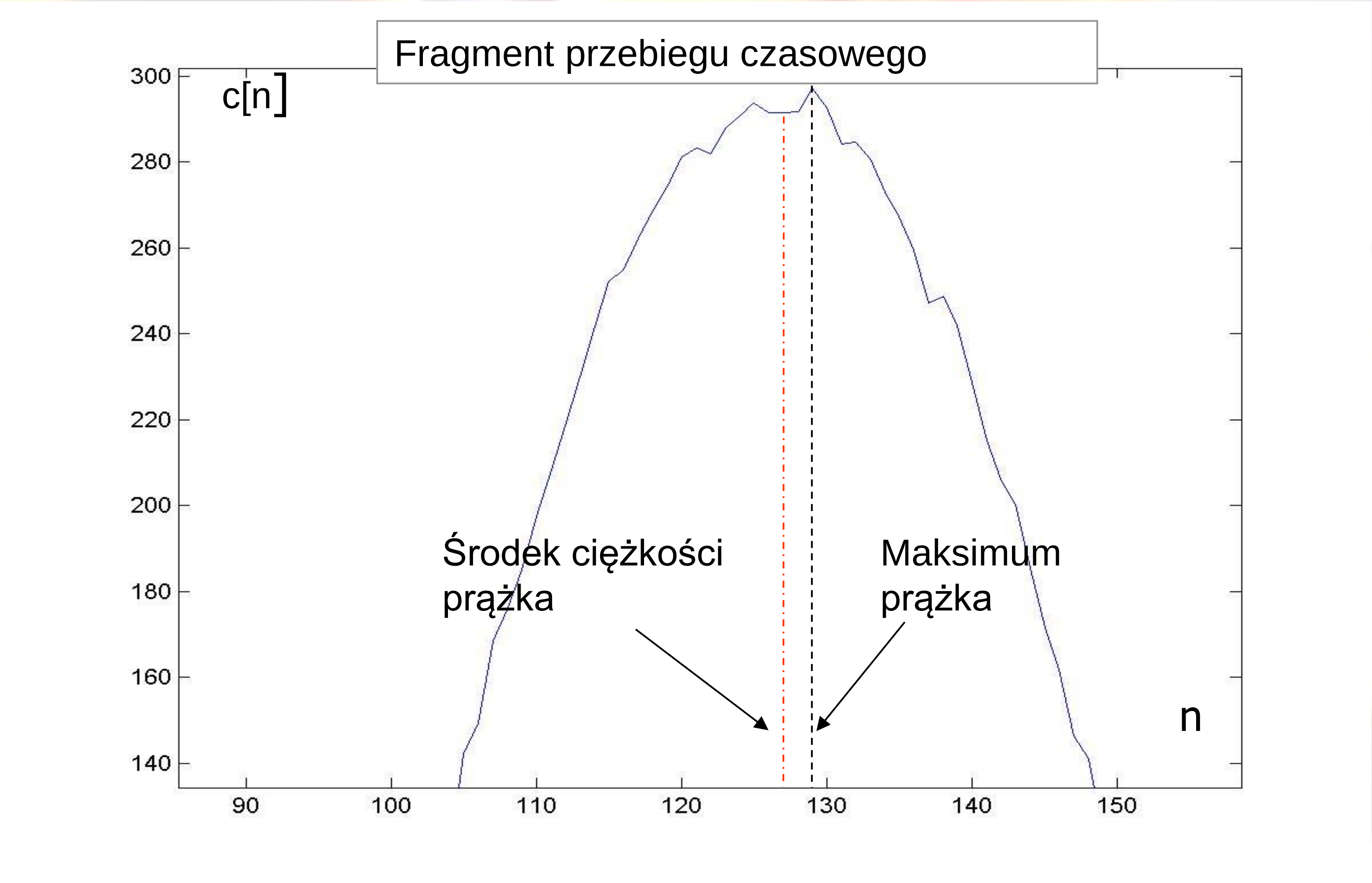
Błędy dokładnej estymacji



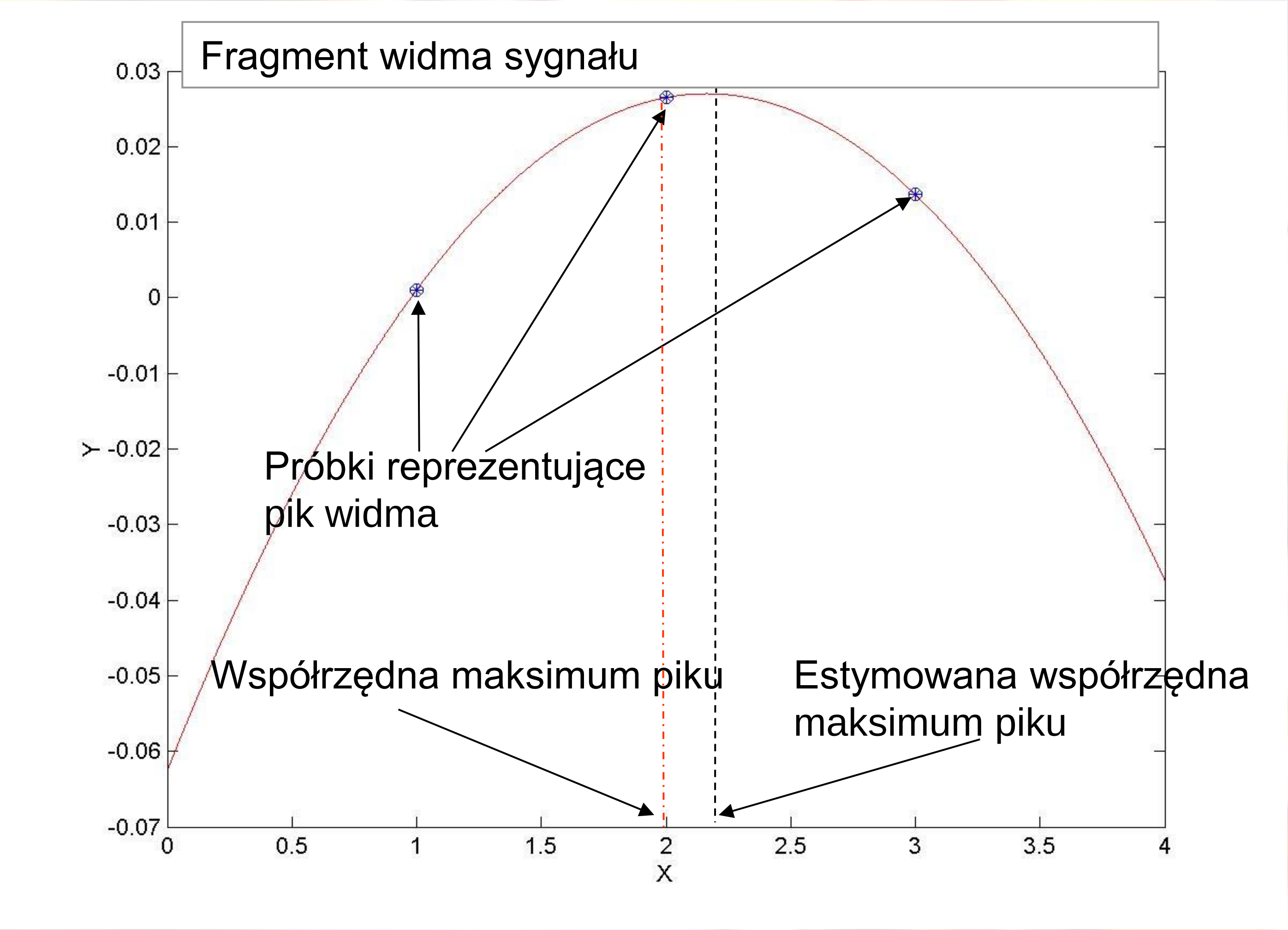
Błędy dokładnej estymacji

- Błędy dokładnej estymacji można zminimalizować przy pomocy technik interpolacyjnych
- Ponieważ w analizie czasowej maksima przebiegów zazwyczaj są reprezentowane przez wiele próbek, skuteczną metodą poprawy estymacji jest obliczanie środka ciężkości analizowanego wycinka sygnału
- W analizie częstotliwościowej piki widma reprezentowane są jedynie przez kilka próbek, przez co skuteczniejsze jest wykorzystanie liniowych (wielomianowych, sklejanych wielomianów), bądź nieliniowych (interpolacja przy pomocy sieci neuronowych) metod interpolacyjnych.

Błędy dokładnej estymacji



Błędy dokładnej estymacji



Podział metod detekcji f_0

- **Metody czasowe**, analizujące bezpośrednio postać czasową sygnału
- **Metody widmowe**, wykorzystujące operacje w dziedzinie częstotliwości

Metody czasowe detekcji f_0

- Najpopularniejsze algorytmy działające w oparciu o reprezentację czasową sygnału to:
  - Analiza sygnału autokorelacji,
  - Analiza sygnału wygenerowanego przy pomocy metody AMDF (ang. Average Magnitude Difference Function)
  - Wykorzystanie powyższych metod dla liniowo i nieliniowo zmodyfikowanego sygnału wejściowego

Metody czasowe detekcji f_0

- **Analiza funkcji autokorelacji**

  Funkcja autokorelacji sygnału dyskretnego:

$$r[n] = \sum_{m} x(m) \cdot x(m+n)$$

  Położenie pierwszego maksimum tej funkcji dla argumentu różnego od zera wyznacza okres sygnału w próbkach. Długość ramki sygnału musi wynosić co najmniej kilka okresów.

  - Bardzo dobra rozdzielczość.
  - Możliwe błędy oktawowe, metoda mało odporna na szum i zakłócenia
  - Trudności w analizie sygnału pozbawionego pierwszej harmonicznej

Metodы czasowe detekcji f_0

- **Analiza funkcji autokorelacji**

  Funkcja autokorelacji sygnału dyskretnego:

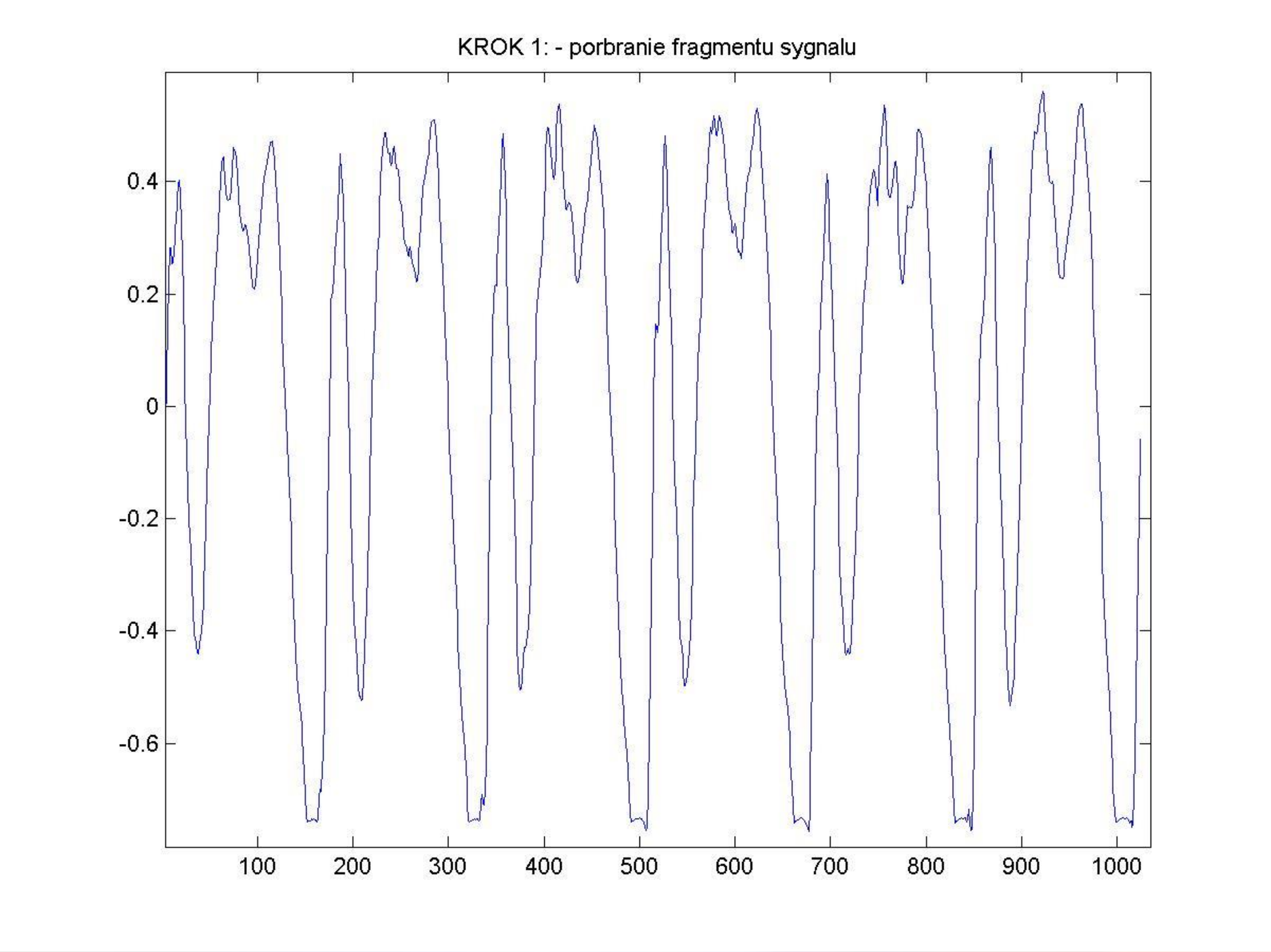
$$r[n] = \sum_m x(m) \cdot x(m+n)$$

Przydatne funkcje **Matlaba**

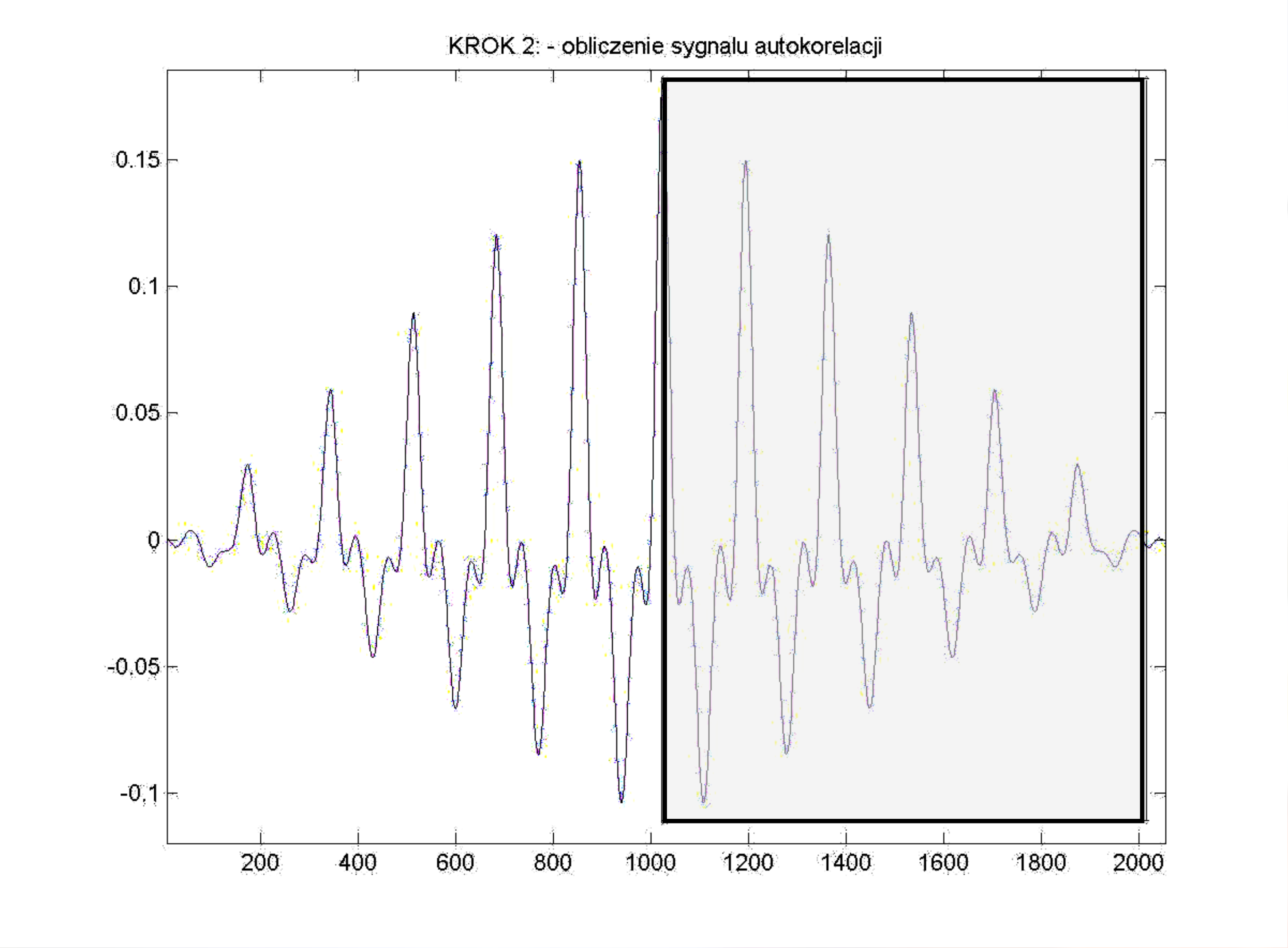
```
y=x(K:L);  %Pobieranie wycinka sygnału x (od próbki K do L)
xs=xcorr(s);  % obliczanie sygnału autokorelacji sygnału s
xs2=xs(length(s):end);  % Pobieranie połowy sygnału autokorelacji
```

- Alternatywny algorytm
  - zwiększyć dwukrotnie długość sygnału *x* poprzez uzupełnienie zerami
  - $r = IFFT\left(FFT(x) \cdot FFT(x)^*\right)$

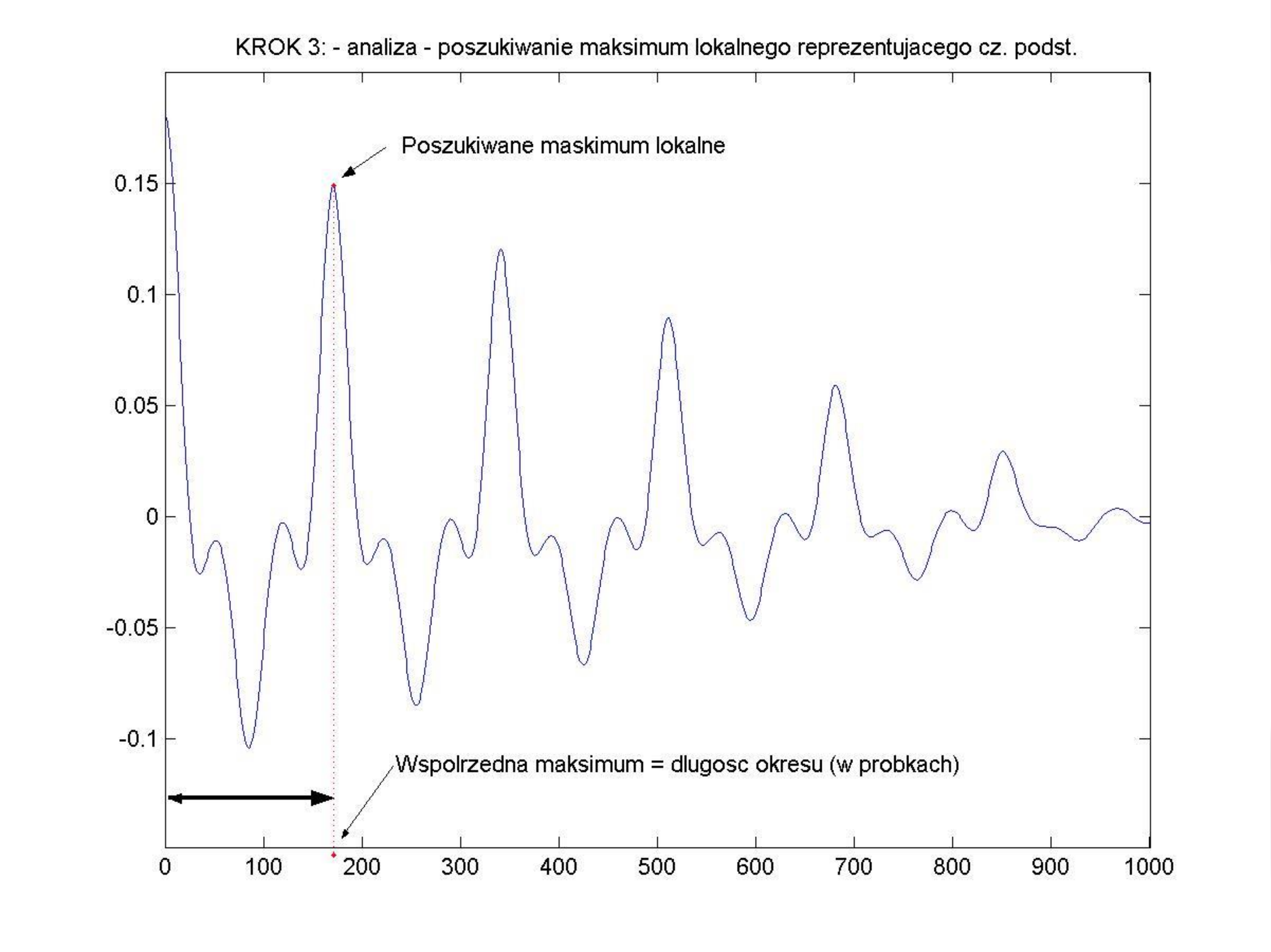
Analiza funkcji autokorelacji



Analiza funkcji autokorelacji



Analiza funkcji autokorelacji



Metody czasowe detekcji f_0

- **Metoda AMDF (Average Magnitude Differential Function)**

  polega na badaniu relacji między sygnałem oryginalnym i opóźnionym

  $$AMDF(n) = \sum_{m=1}^{M} \left| x(m) - x(m+n) \right|^{k}, \quad k = 1$$

  Częstotliwość podstawową sygnału określone jest przez położenie pierwszego minimum lokalnego funkcji AMDF (dla $n \neq 0$). Bardzo mała złożoność obliczeniowa, jednak pojawiają się błędy w przypadku, gdy okres sygnału nie jest całkowitą wielokrotnością okresu próbkowania.

Metody czasowe detekcji f_0

- **Metoda AMDF (Average Magnitude Differential Function)**

  polega na badaniu relacji między sygnałem oryginalnym i opóźnionym

$$AMDF(n) = \sum_{m=1}^{M} \left| x(m) - x(m+n) \right|^k, \quad k = 1$$

Przydatne funkcje **Matlaba**

z=x-y; % Odejmowanie (próbka, po próbce) sygnałów

z=sum(abs(x-y)); % Suma wartości bezwzględnych różnicy sygnałów

z=x.*y; % Mnożenie próbka po próbce sygnałów

z=x*y; % Mnożenie sygnałów (wektorów)

Liniowe i nieliniowe przekształcenia sygnału

Aby zwiększyć energię poszczególnych harmonicznych sygnału trafiającego do detektora częstotliwości podstawowej i tym samym poprawić skuteczność działania algorytmu detekcji (w kontekście minimalizacji błędów oktawowych) często dokonuje się liniowych (filtracja dolno- i górno-przepustowa, splatanie sąsiednich ramek, itp.) bądź nieliniowych (podnoszenie próbek sygnału do potęgi, generowanie przebiegów fazowych) operacji na sygnale

Inne metody czasowe

- **Metody progowe** – polegają na analizie przejść przebiegu czasowego przez wybraną wartość progową. Wyróżnia się:
  - Analizę przejść przez zero ZXABE (ang. zero crossing analysis basic extractor)
  - Analizę przejść przez wartość progową TABE (ang. threshold analysis basic extractor)

  Zaletą metod progowych jest możliwość działania niemal w czasie rzeczywistym (z minimalnym opóźnieniem), ponieważ nie wymagają pełnego zestawu danych. Może to jednak skutkować błędami detekcji.

  Metody posiadają bardzo dobrą rozdzielczość, lecz są mało odporne na addytywny szum zakłócający.

Algorytmy detekcji operujące w dziedzinie widma

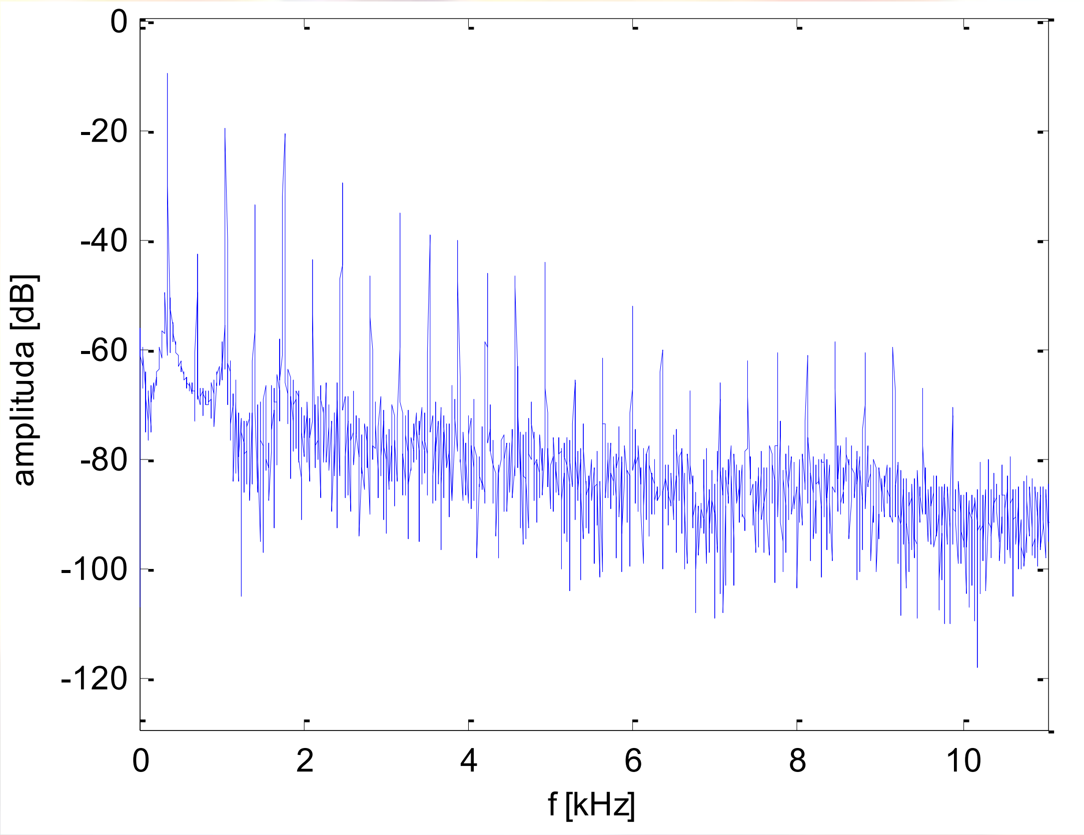
Algorytmy bazujące na reprezentacji widmowej sygnału oraz algorytm pochodne, bazujące na przekształconej reprezentacji widmowej (cepstrum, autokorelacja zlogarytmowanego widma) opierają swoje działanie na detekcji pików reprezentujących składowe harmoniczne sygnału.

Przydatne funkcje **Matlaba**

```
[v, ind]=max(s);   % Wartość maksymalna i jej indeks (sygnału s)
y=filter(ones(1,K)/K, 1,s);  % średnia ruchoma rzędu K sygnału s
y=fft(s);  % obliczanie widma zespolonego
y2=abs(y);  % obliczanie wartości bezwzględnej sygnału zespolonego y
y3=log10(y2);  % obliczanie logarytmu dla wszystkich próbek sygnału y2
plot(20*log10(abs(fft(s))));  % rysowanie widma mocy sygnału s
```

Analiza położenia składowych harmonicznych

Na podstawie odległości pomiędzy składowymi harmonicznymi widma sygnału można określić częstotliwość podstawową sygnału



Metoda cepstralna

- **Metoda cepstralna**

  Obliczana odwrotna transformata Fouriera logarytmu widma amplitudowego analizowanej ramki sygnału, wg wzoru:

$$C_r = \sum_{n=1}^{m} \ln\left(X_n\right) \cdot \cos\left(\frac{r \cdot n\pi}{m}\right)$$

  Częstotliwość podstawowa sygnału w ramce estymowana jest na podstawie położenia maksimum w dziedzinie cepstrum

  Algorytm oparty o analizę cepstralną jest relatywnie niewrażliwy na szum, ale występuje problem pojawiania się błędów oktawowych

Metoda cepstralna

- **Metoda cepstralna**

  Obliczana odwrotna transformata Fouriera logarytmu widma amplitudowego analizowanej ramki sygnału, wg wzoru:

$$C_r = \sum_{n=1}^{m} \ln(X_n) \cdot \cos\left(\frac{r \cdot n\pi}{m}\right)$$

Przydatne funkcje **Matlaba**

z_re=real(z); % Część rzeczywista sygnału zespolonego

z_im=imag(z); % Część urojona sygnału zespolonego

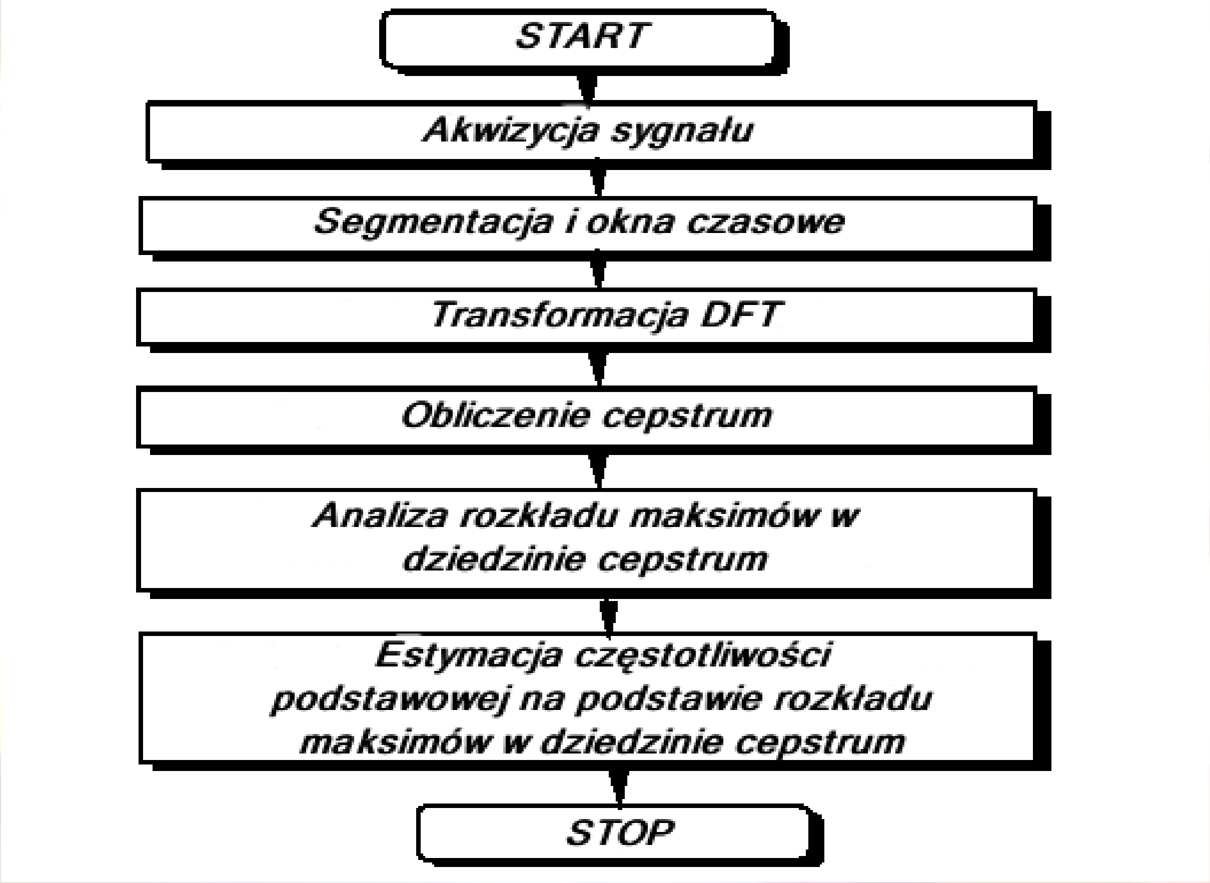
z_an=phase(z); % Faza sygnału zespolonego

x_d=cceps(x); % Obliczanie cepstrum zespolonego sygnału x

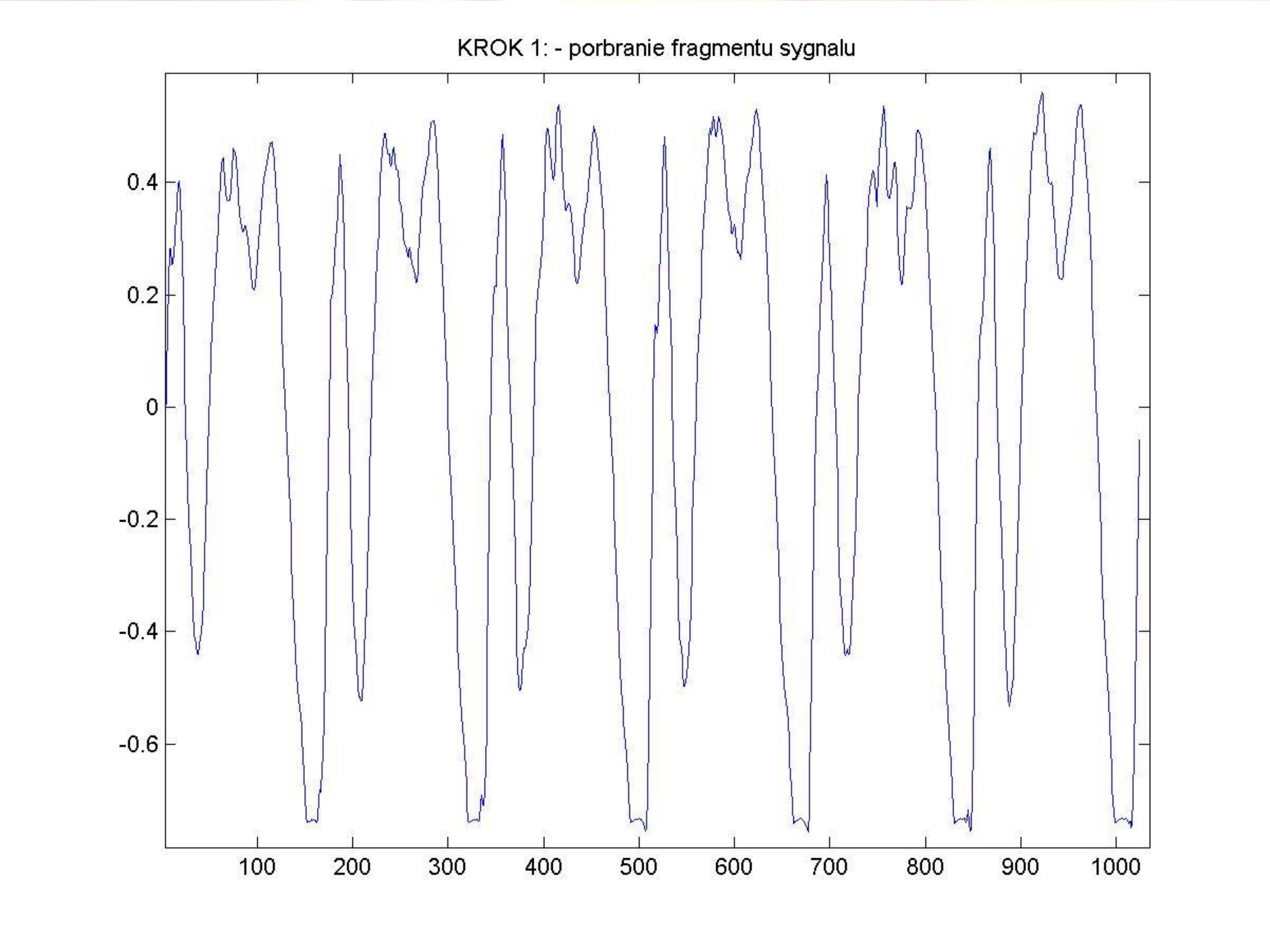
x_d=rceps(x); % Obliczanie cepstrum rzeczywistego sygnału x

Metody widmowe detekcji f_0

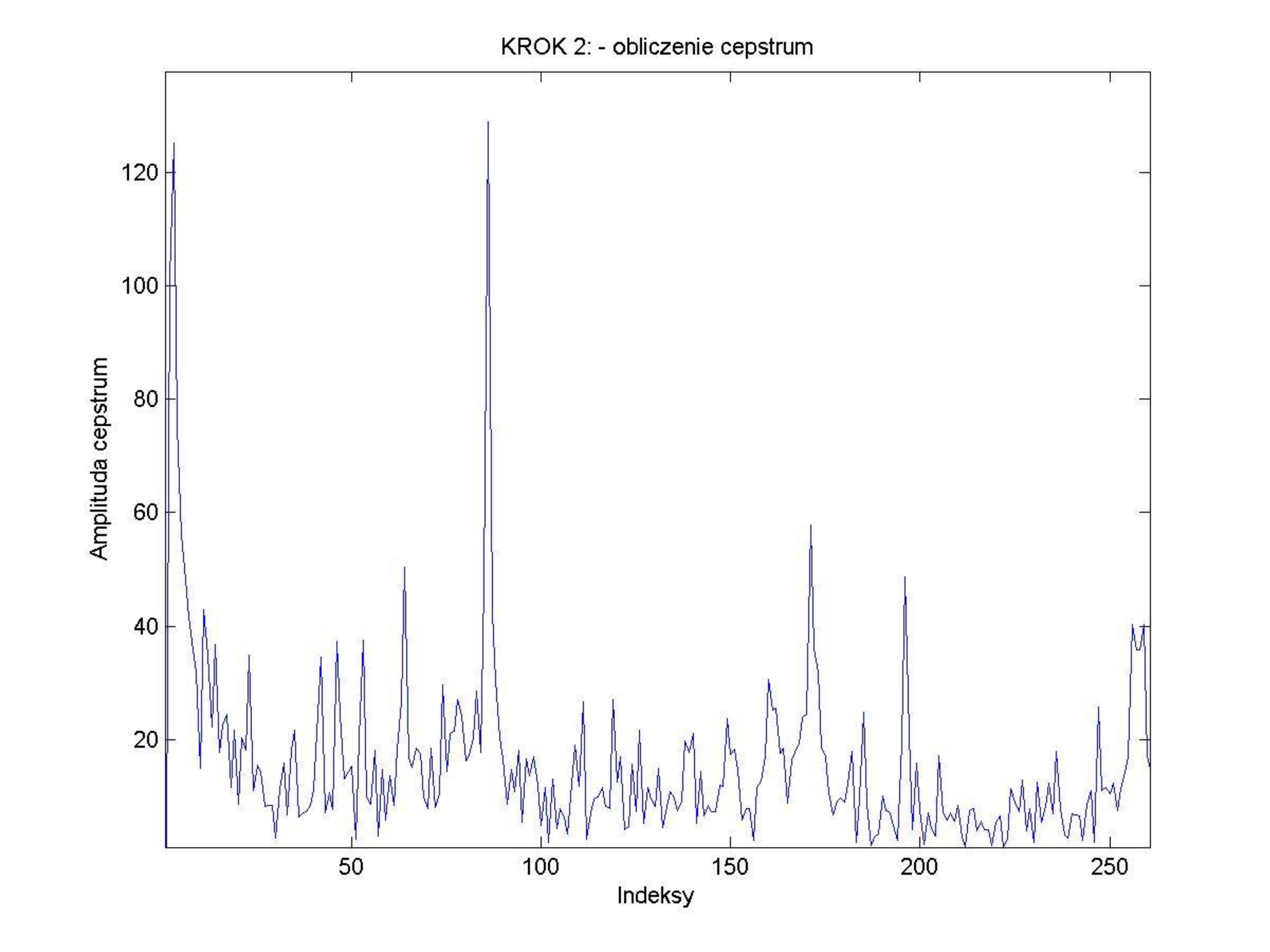
- **Metoda cepstralna**



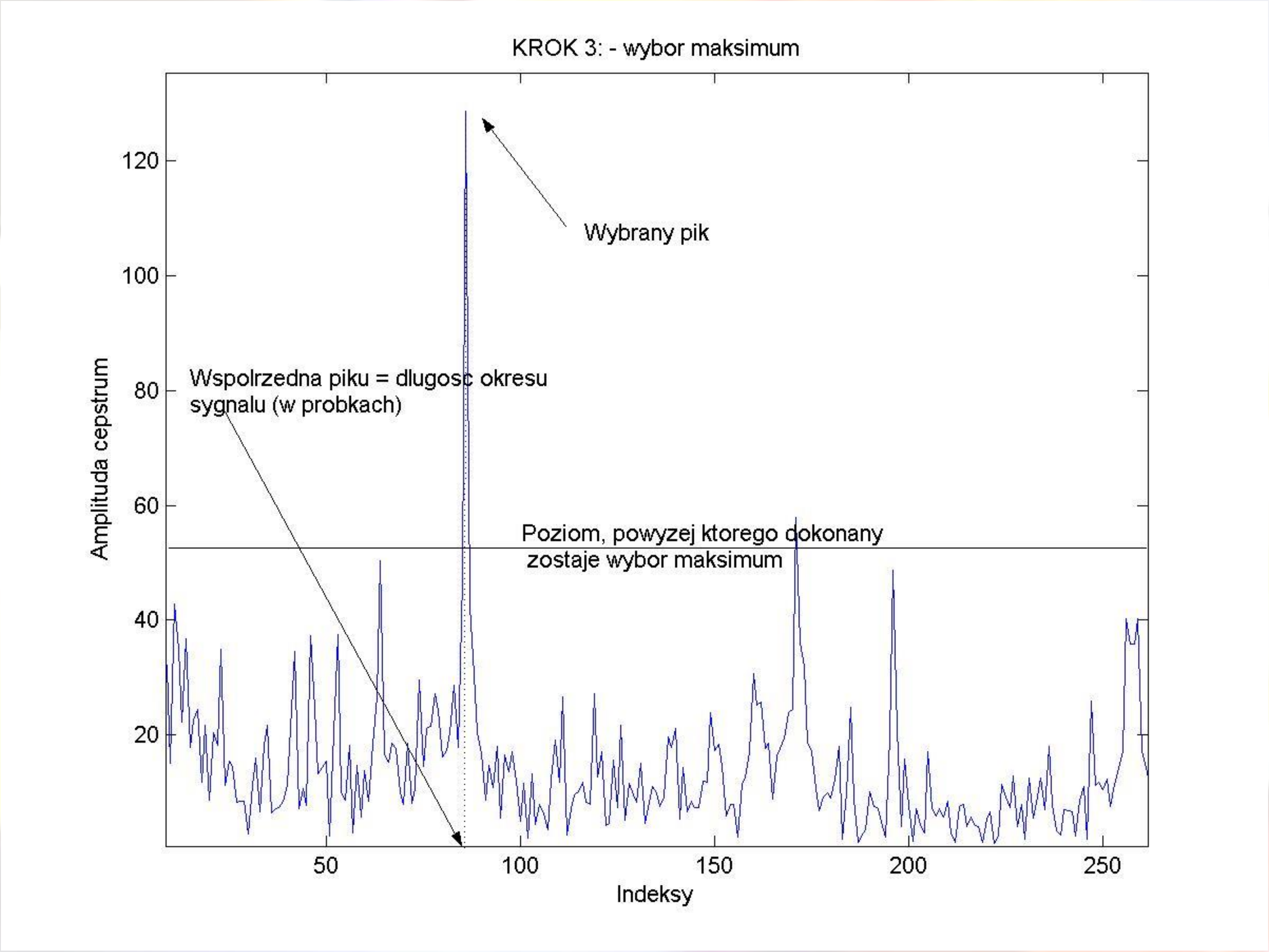
Metoda cepstralna



Metoda cepstralna



Metoda cepstralna



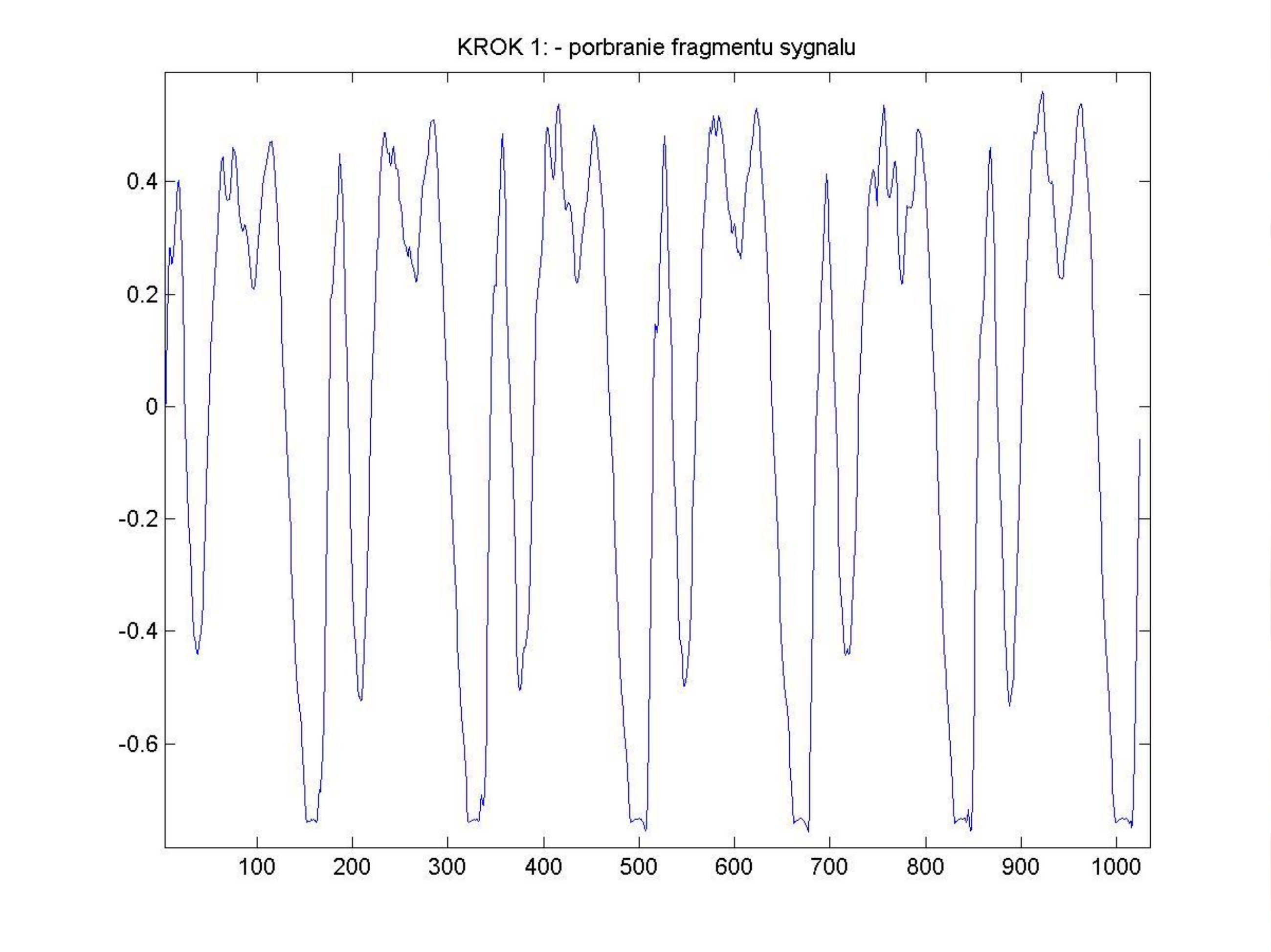
Analiza ACOLS

Metoda ACOLS (ang. Autocorrelation Of Log Spectrum) oparta jest na analizie sygnału autokorelacji obliczonego na podstawie zlogarytmowanego widma sygnału wejściowego, przy czym współrzędna piku reprezentującego częstotliwość podstawową zlokalizowana jest w dziedzinie częstotliwości.

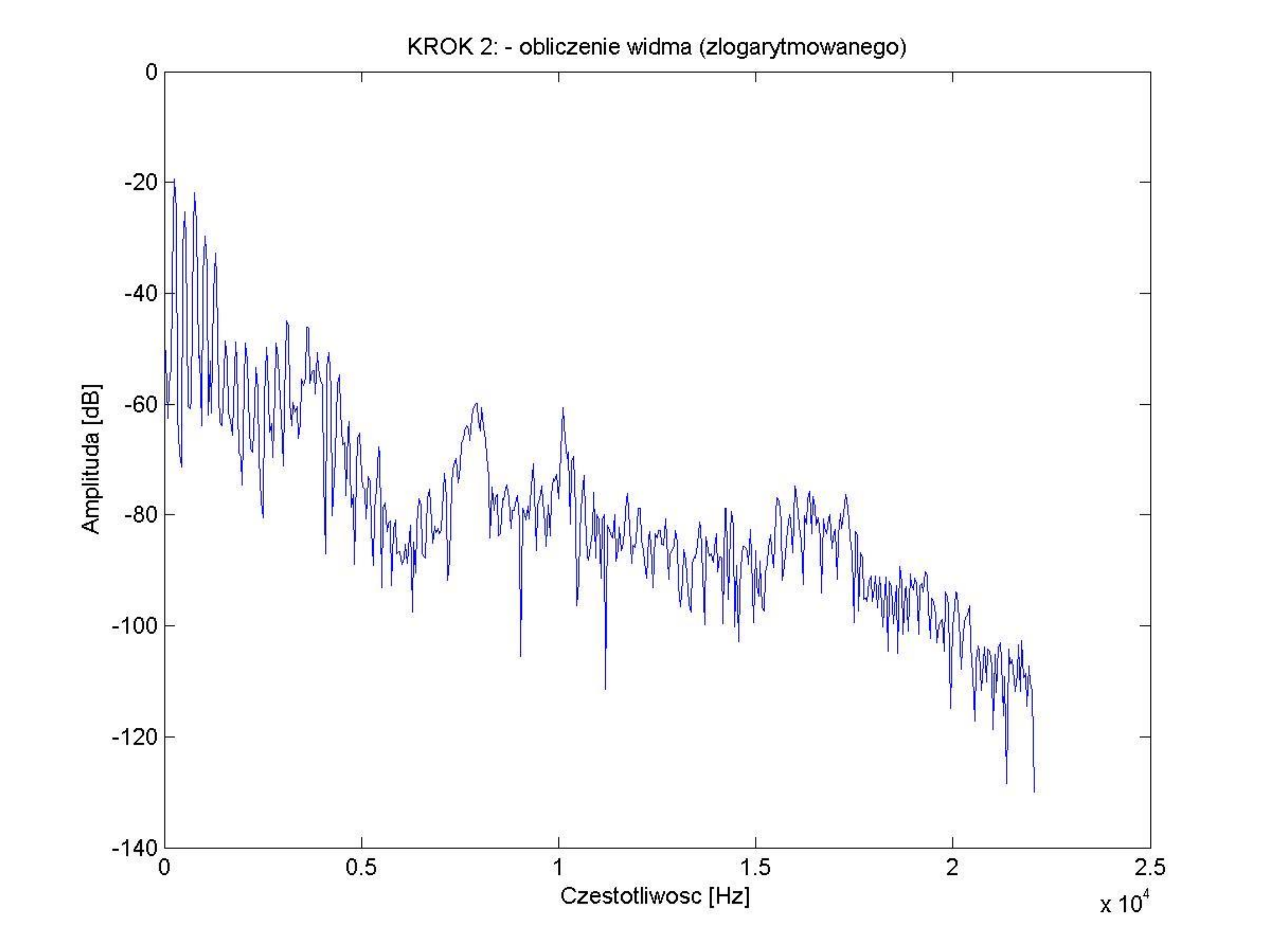
Matlab

```
ACOLS=xcorr(log(abs(fft(s))));
```

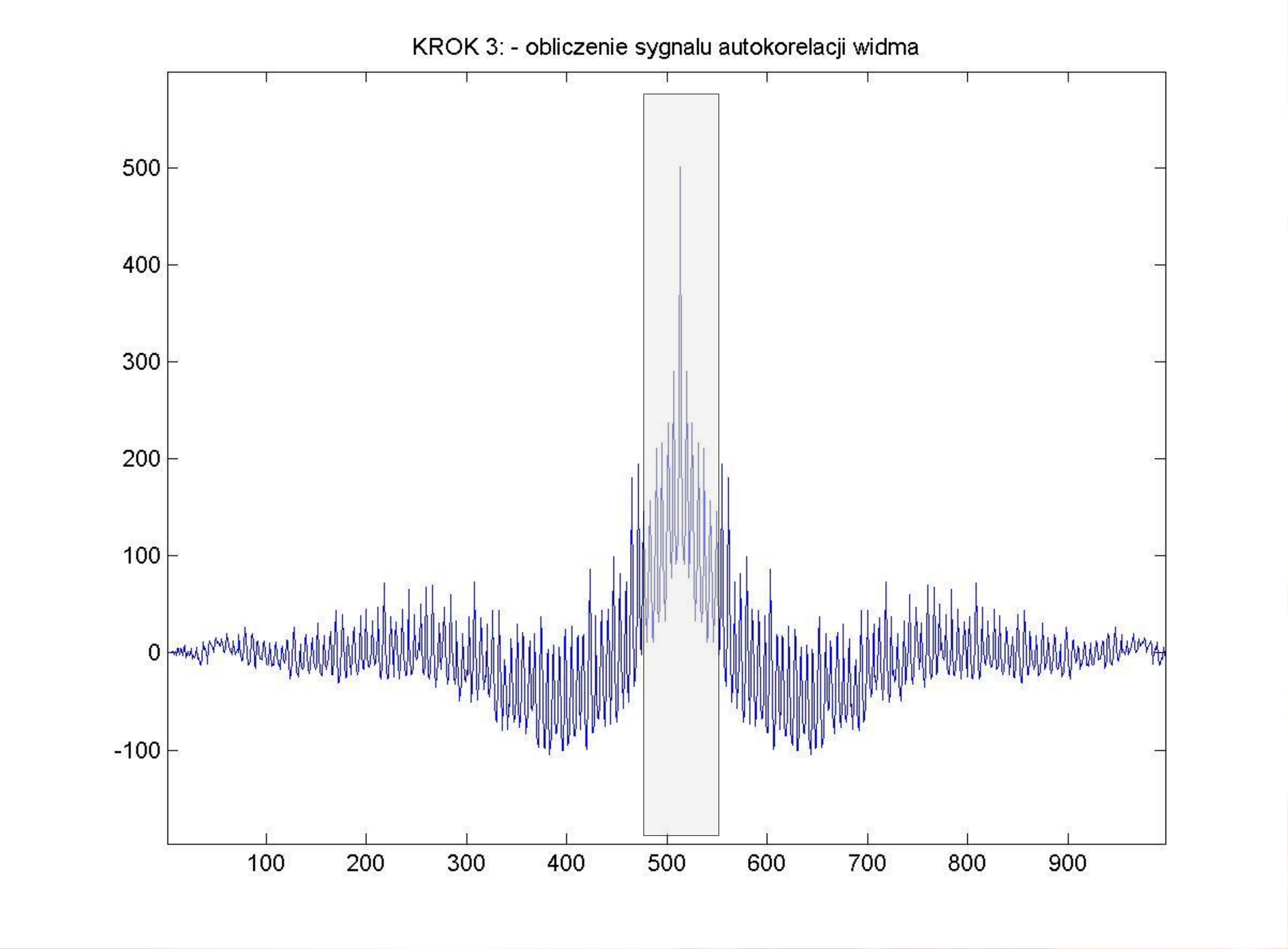
Analiza ACOLS



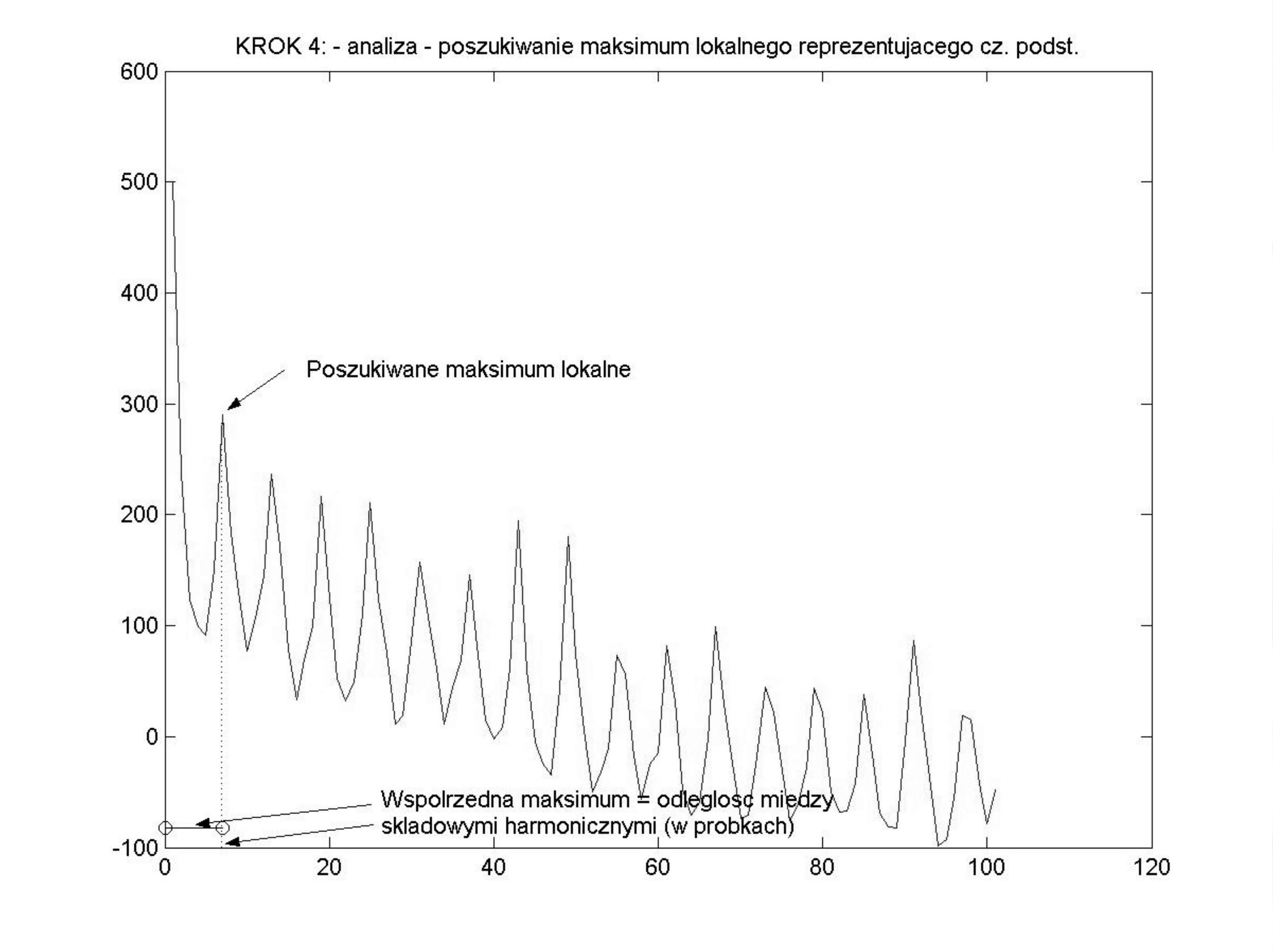
Analiza ACOLS



Analiza ACOLS



Analiza ACOLS



Algorytmy detekcji operujące w przestrzeni czas-częstotliwość

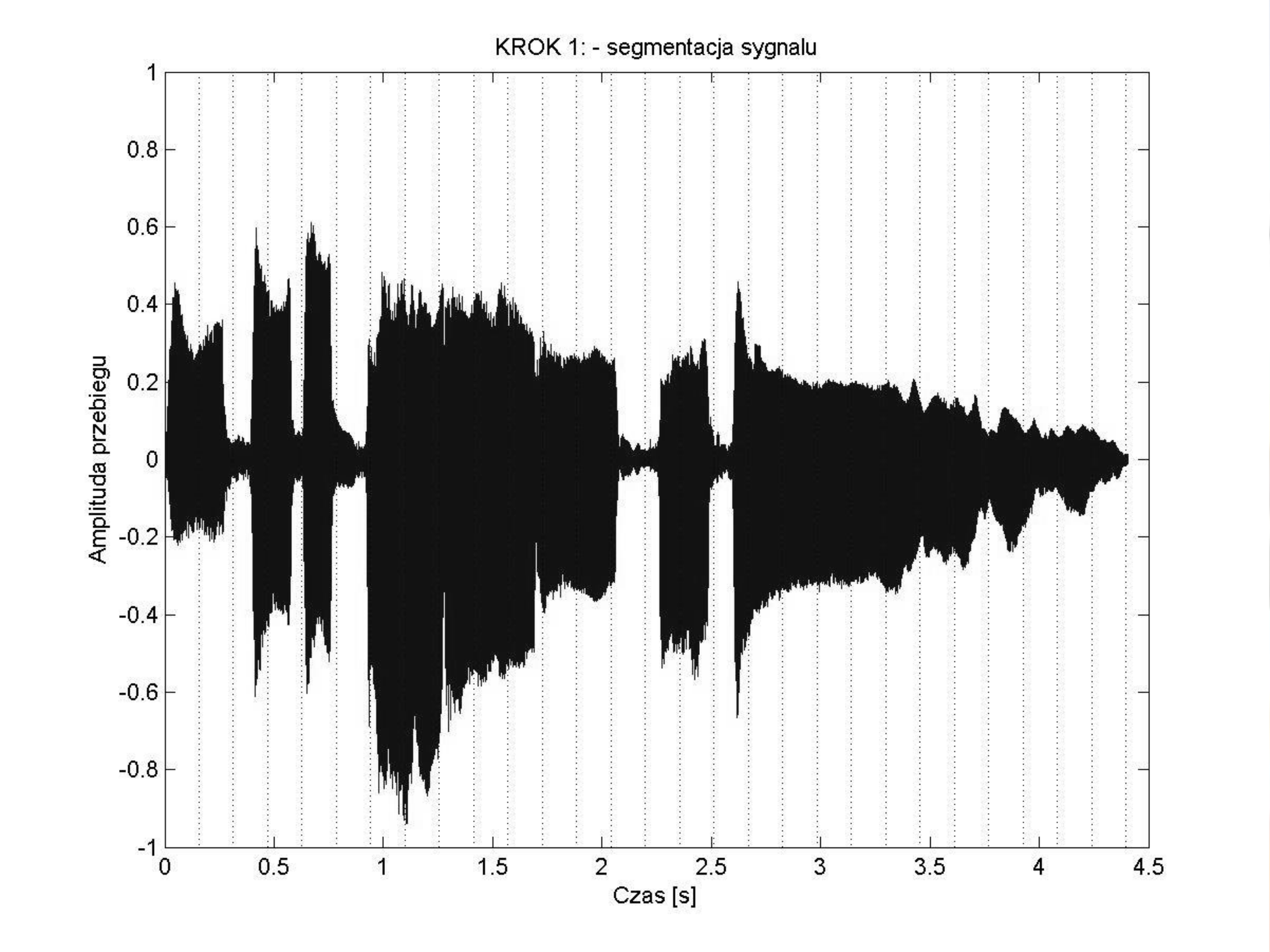
Aby zwiększyć prawdopodobieństwo poprawnej detekcji częstotliwości podstawowej, możliwe jest wykorzystanie informacji zgromadzonych na podstawie analizy przebiegu czasowego oraz reprezentacji widmowej sygnału

Ponadto część algorytmów bada trajektorie wykryte podczas analizy sonograficznej, reprezentujące przebiegi składowych sinusoidalnych w celu ekstrakcji częstotliwości podstawowej

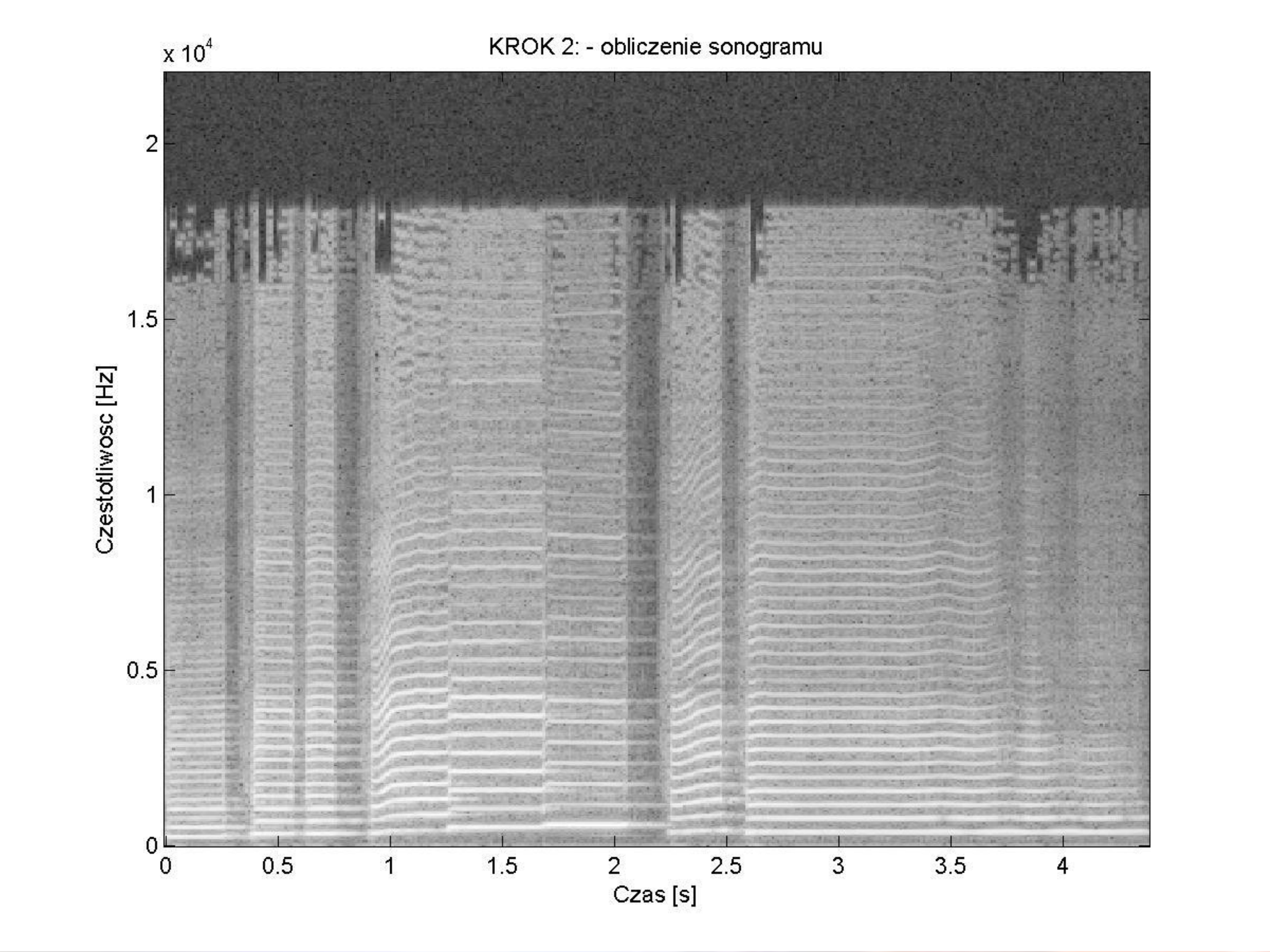
Przydatne funkcje **Matlaba**

S = specgram(s, nFFT, Fs, winType, nOVERLAP); % Obliczanie sonogramu

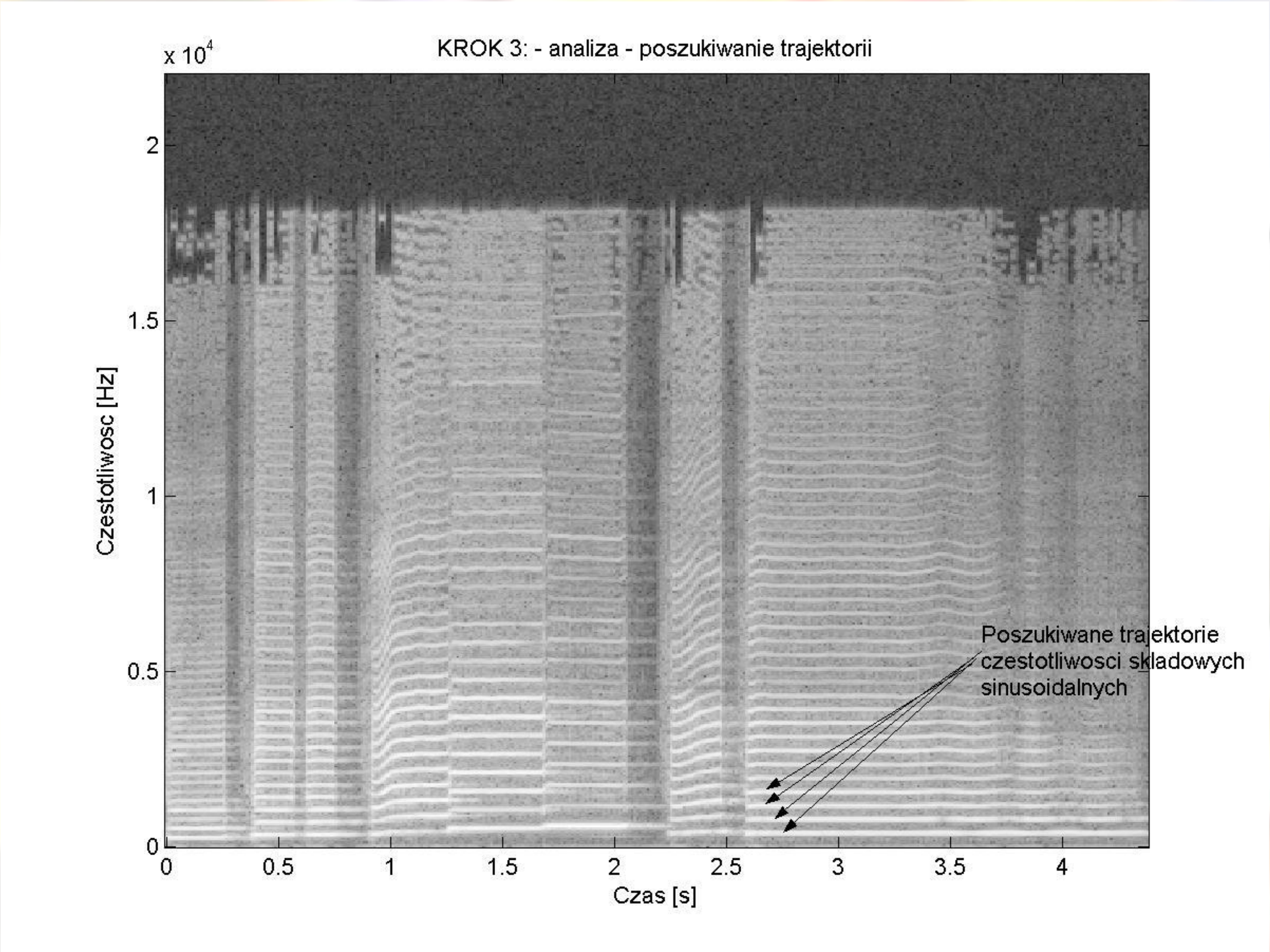
Algorytmy detekcji operujące w przestrzeni czas-częstotliwość



Algorytmy detekcji operujące w przestrzeni czas-częstotliwość



Algorytmy detekcji operujące w przestrzeni czas-częstotliwość



Inne metody widmowe

- **Filtracja grzebieniowa**

  Polega na obliczaniu iloczynów widma sygnału oraz funkcji grzebieniowej o przestrajanej częstotliwości, określającej odległość pomiędzy kolejnymi maksimami lokalnymi funkcji grzebieniowej. Następnie sumuje się wartości prążków po filtracji przez funkcję grzebieniową i przyporządkowuje otrzymane wyniki częstotliwości charakteryzującej funkcję grzebieniową. Położenie maksimum globalnego utworzonej w ten sposób funkcji określa częstotliwość podstawową dźwięku.

Inne metody widmowe

- **Histogram Schroedera**

  Metoda statystyczna polegająca na analizie z osobna częstotliwości każdego prążka.

  Na podstawie rozkładu częstotliwości prążków widma generowany jest histogram częstotliwości. Jeśli wielokrotność częstotliwości analizowanego prążka pokrywa się z częstotliwością innego, to powiększana jest wartość histogramu dla tej właśnie częstotliwości. Częstotliwość podstawowa dźwięku jest wtedy równa częstotliwości, dla której wartość histogramu jest największa.

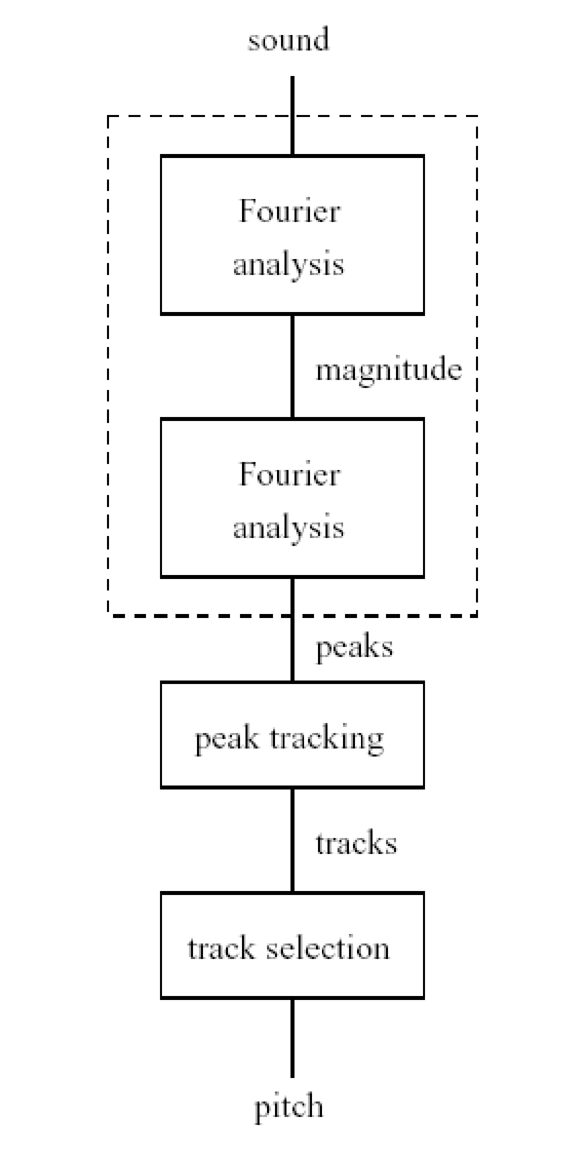
  - duża dokładność
  - odporność na występowanie błędów oktawowych
  - algorytm skuteczny dla sygnałów zaszumionych

Inne metody widmowe

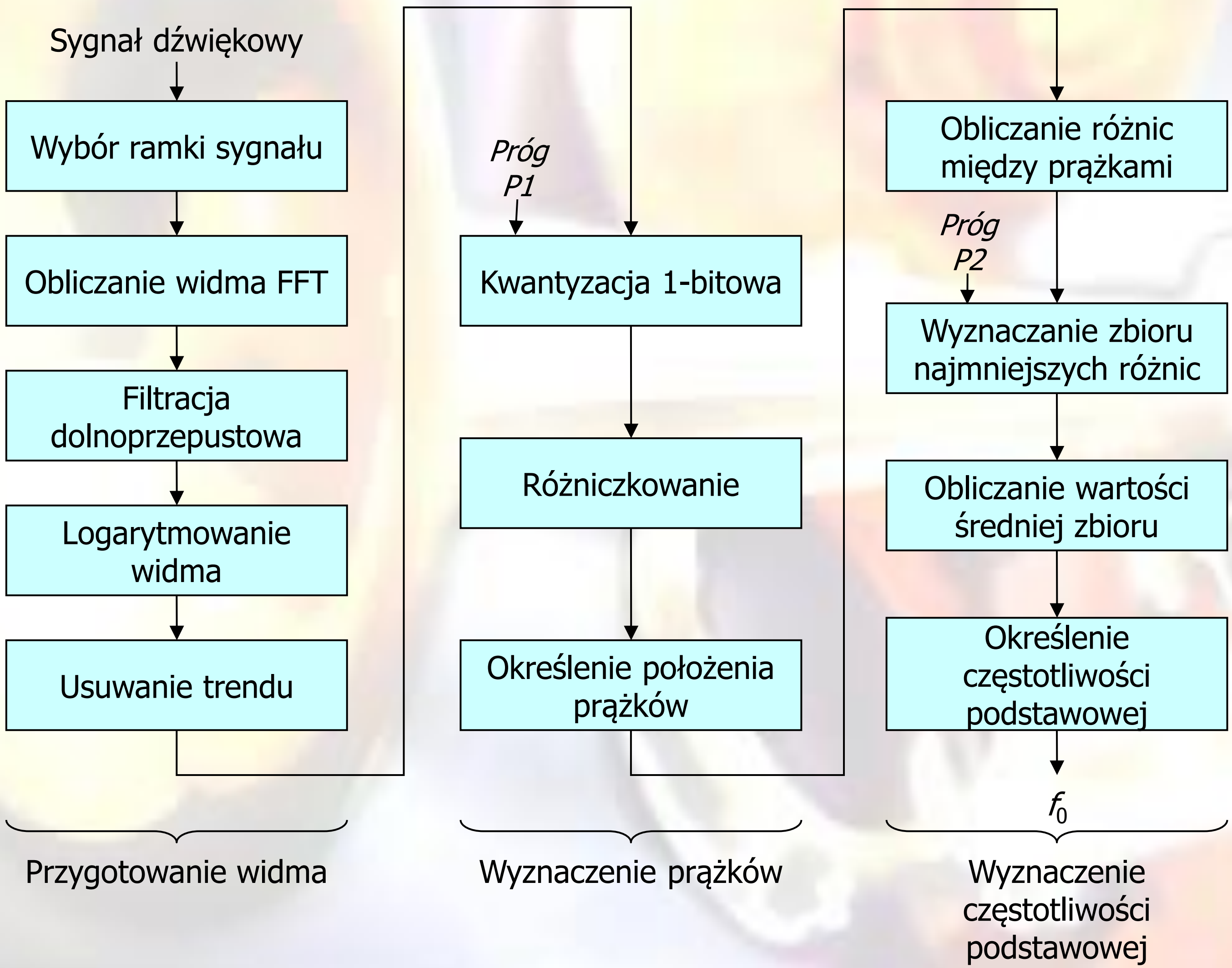
- **Kombinacja transformacji Fouriera**

  Obliczana jest transformata Fouriera widma amplitudowego sygnału.

  Częstotliwość podstawowa odpowiada odległości między prążkami w widmie sygnału, a tym samym największemu maksimum lokalnemu w drugim widmie. Algorytm ten jest skuteczny w przypadku, gdy w widmie sygnału brak jest prążka o częstotliwości odpowiadającej częstotliwości podstawowej.

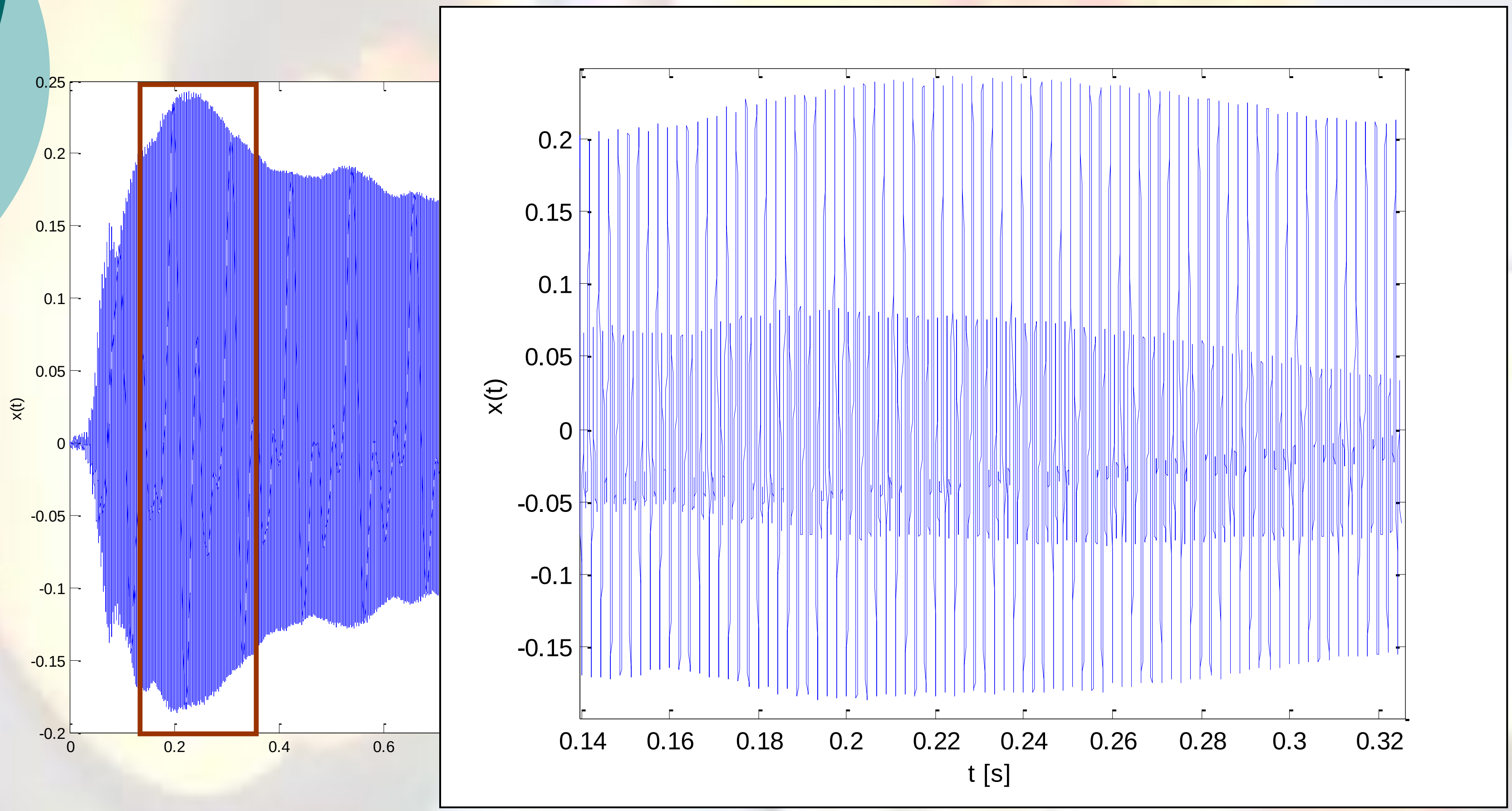


Przykładowy algorytm detekcji f_0



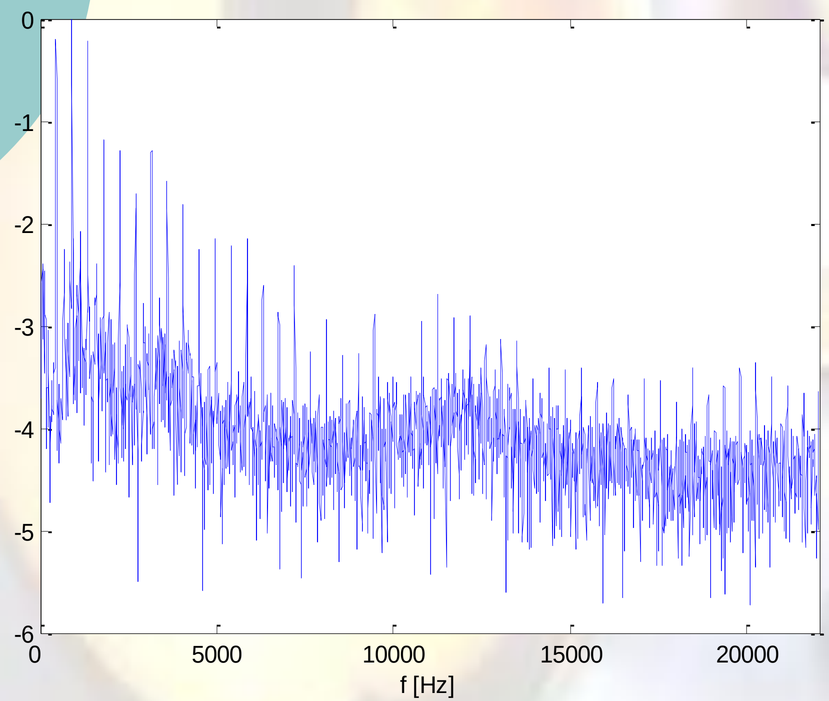
Algorytm detekcji f_0

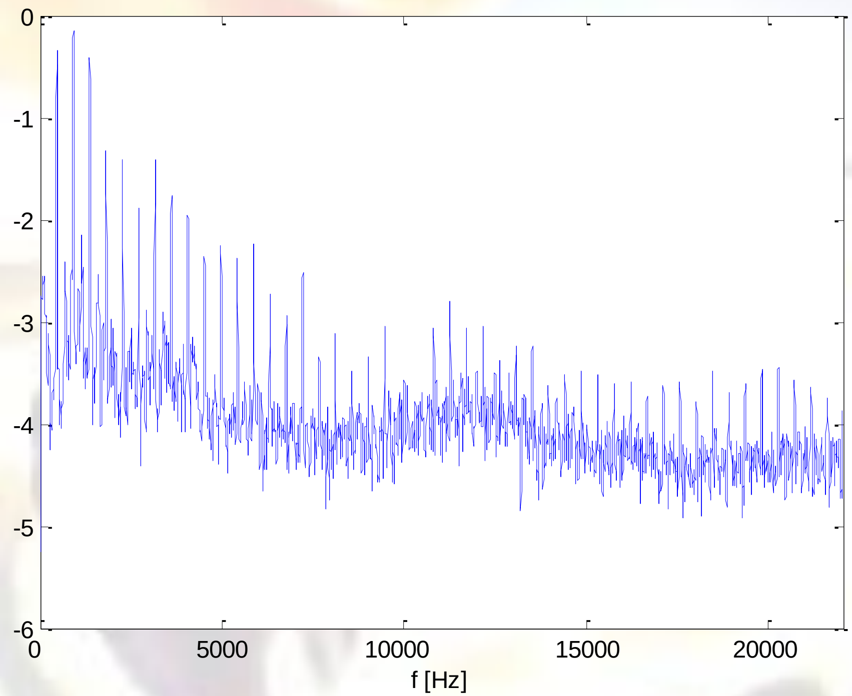
Rożek angielski, dźwięk A4, mezzoforte



Algorytm detekcji f_0

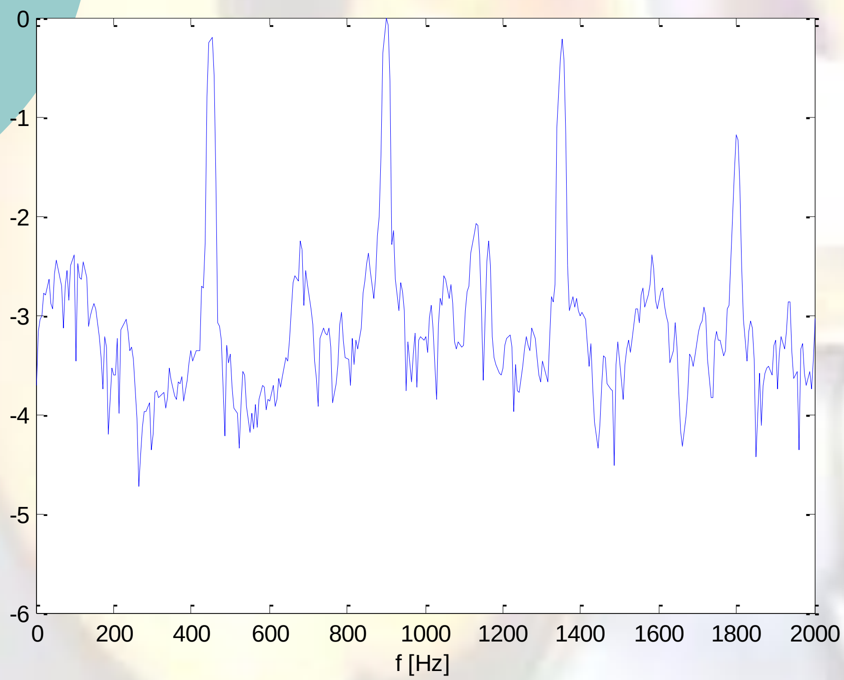
Filtracja dolnoprzepustowa

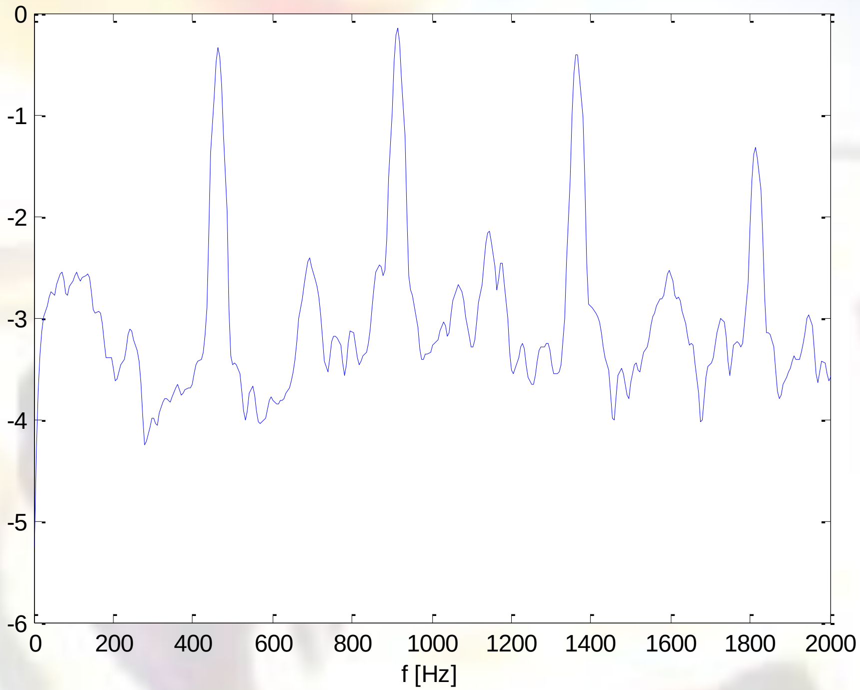




Algorytm detekcji f_0

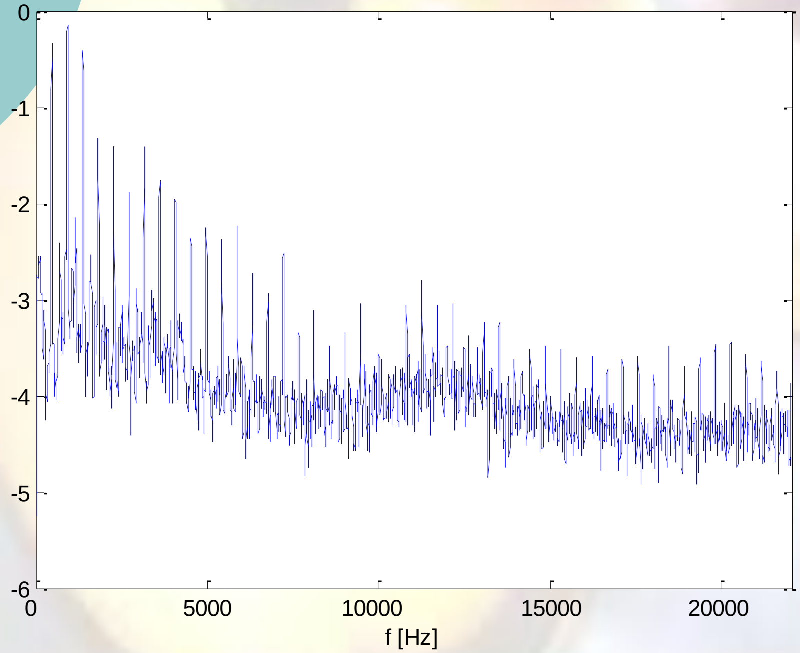
Filtracja dolnoprzepustowa

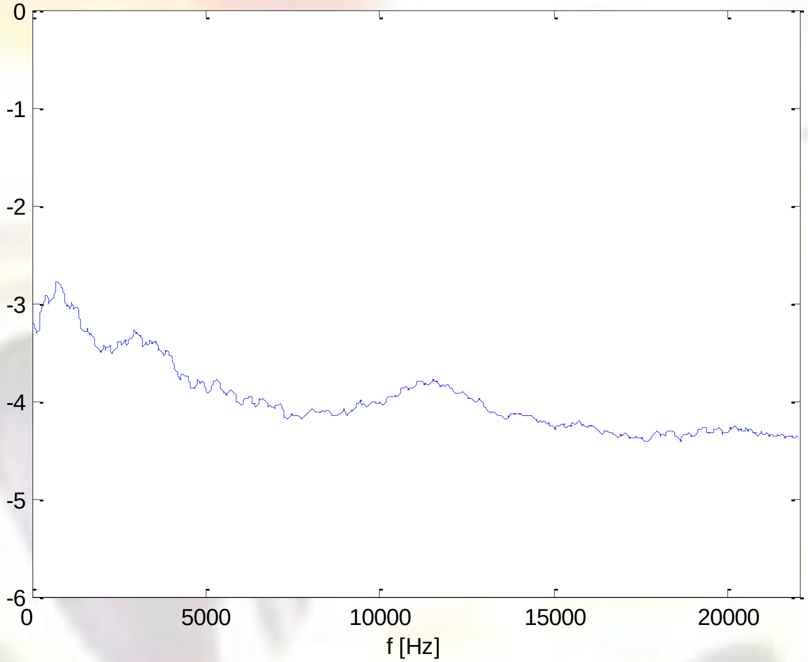




Algorytm detekcji f_0

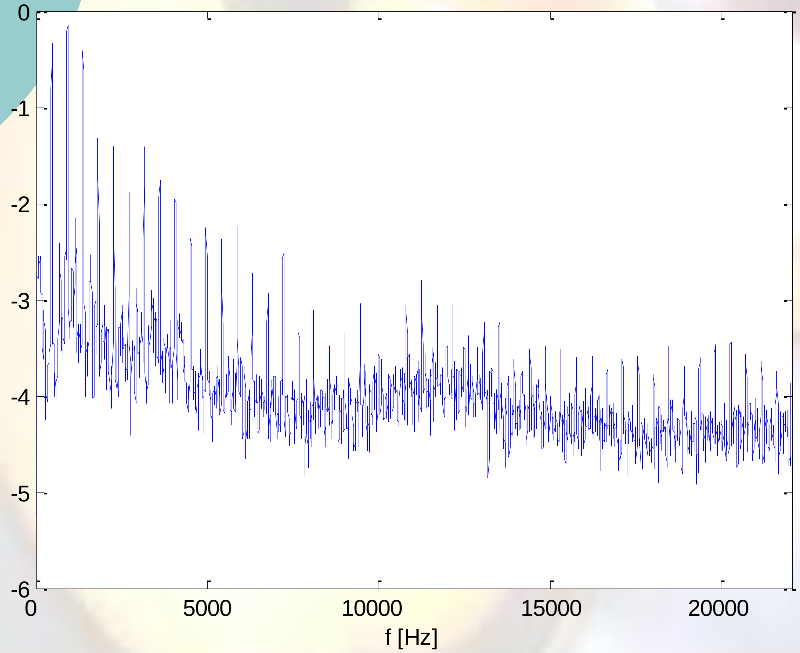
Usuwanie trendu

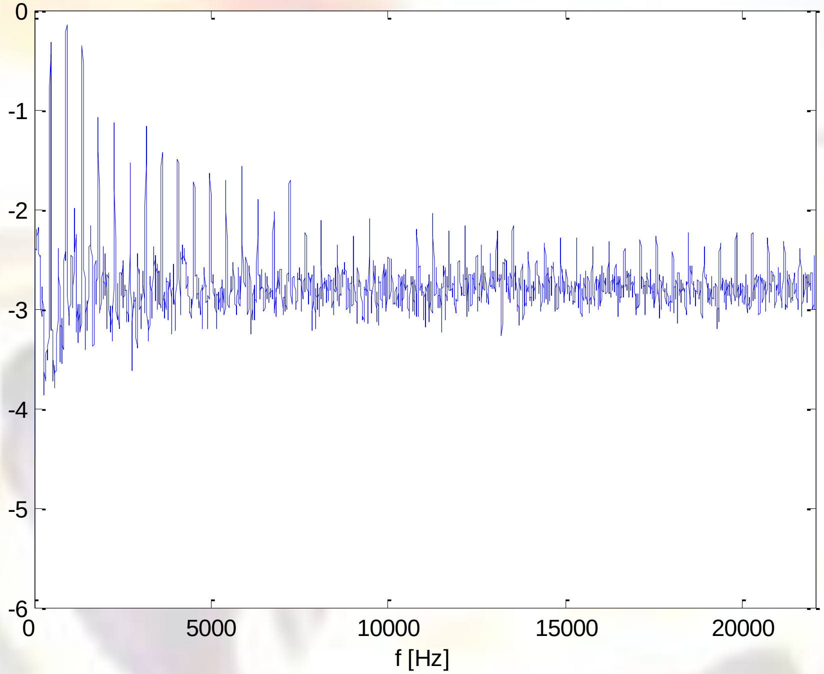




Algorytm detekcji f_0

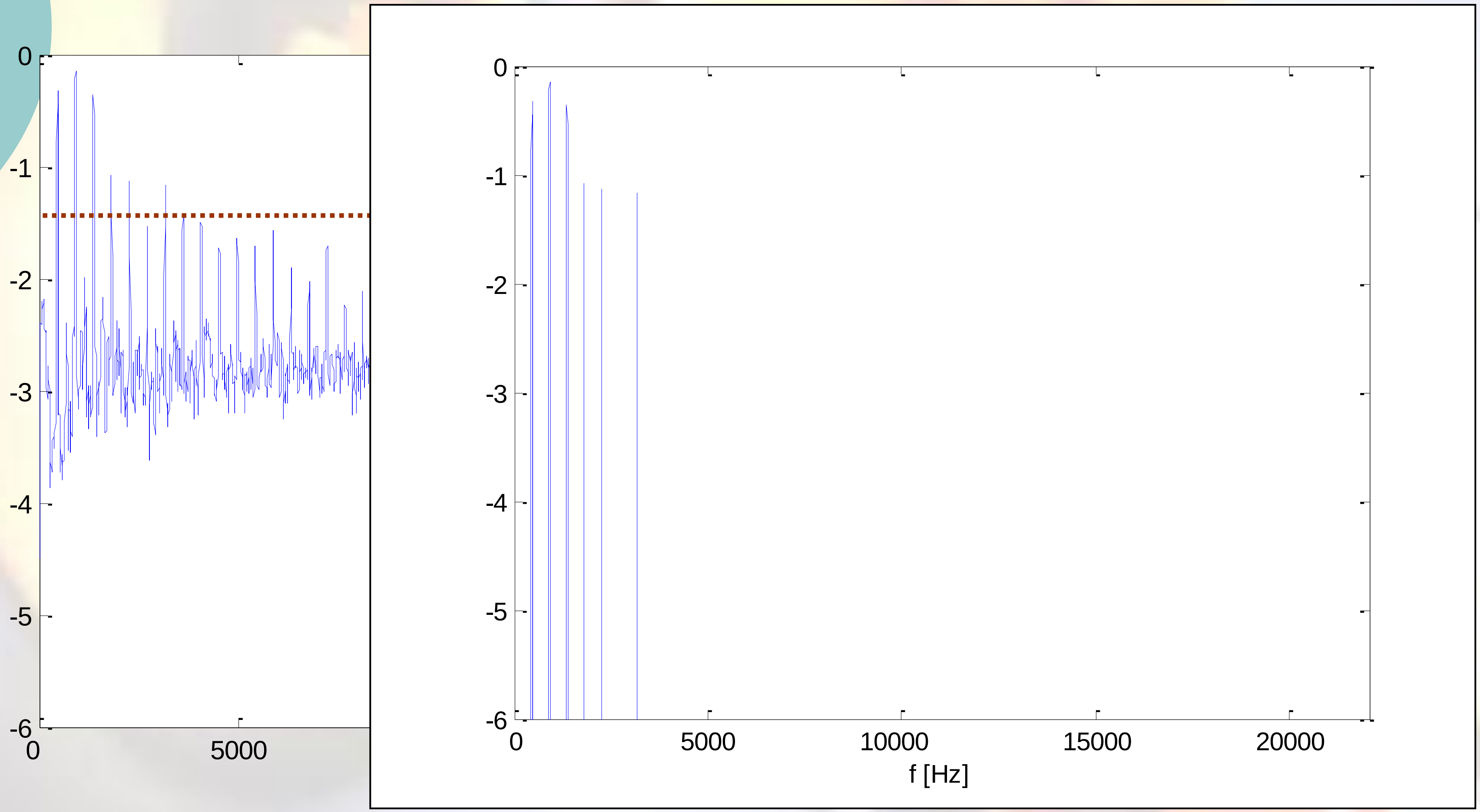
Usuwanie trendu





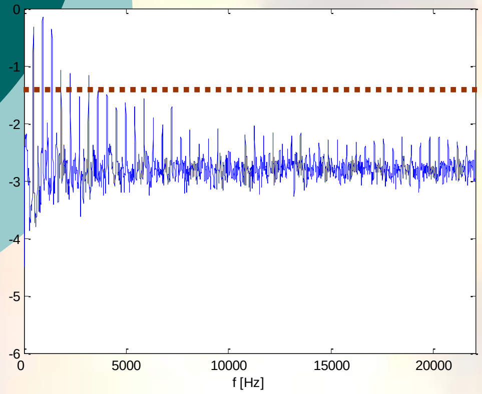
Algorytm detekcji f_0

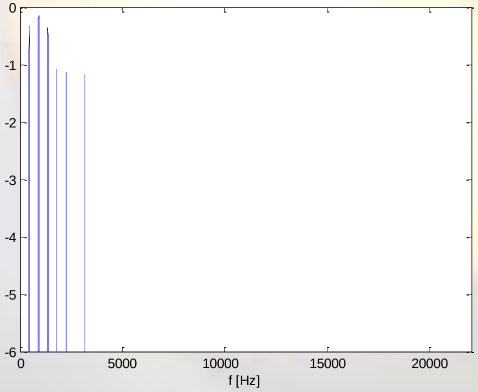
Kwantyzacja 1-bitowa

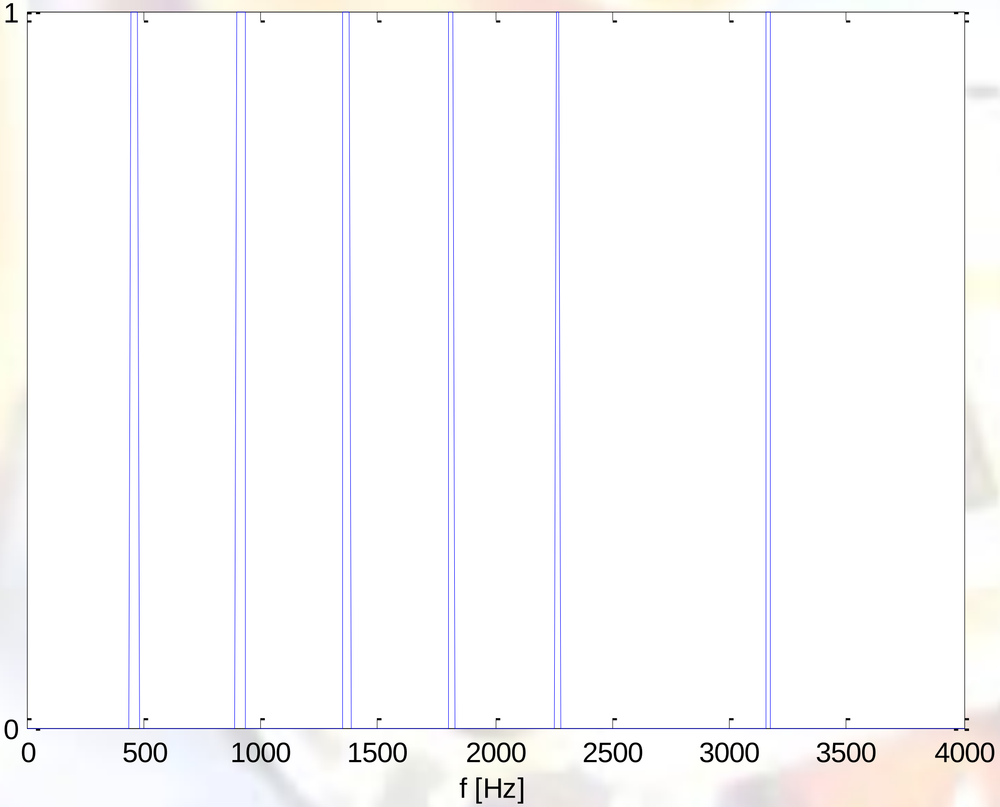


Algorytm detekcji f_0

Kwantyzacja 1-bitowa

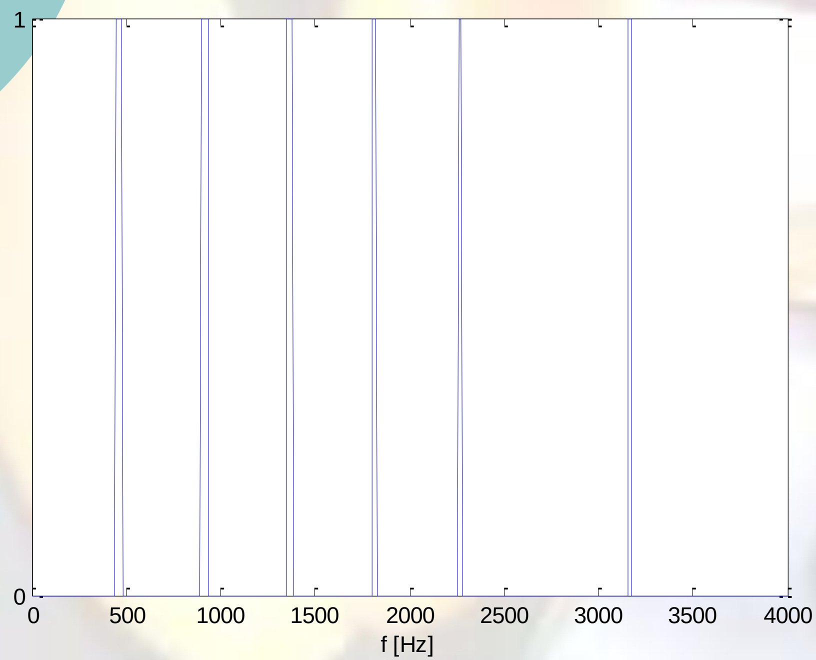


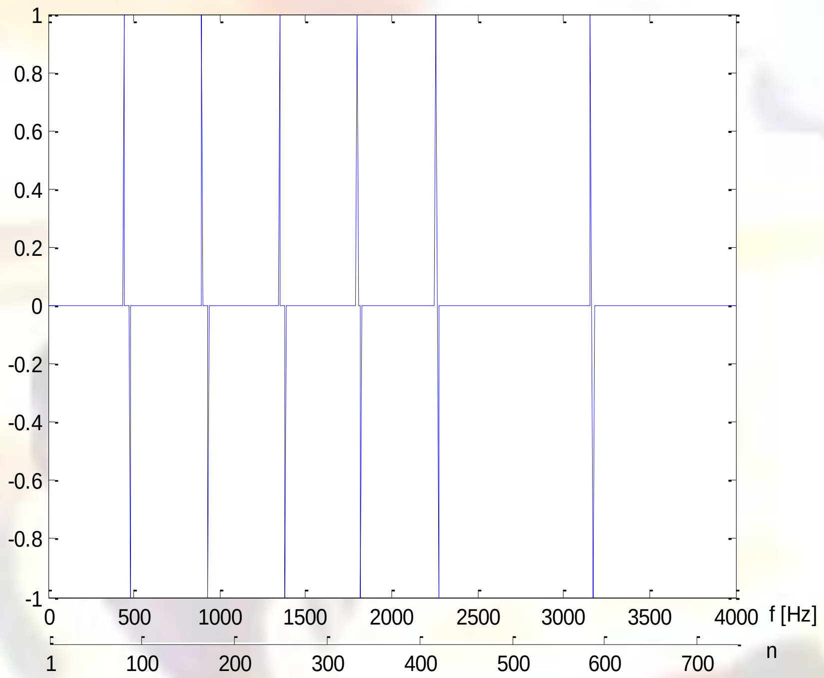




Algorytm detekcji f_0

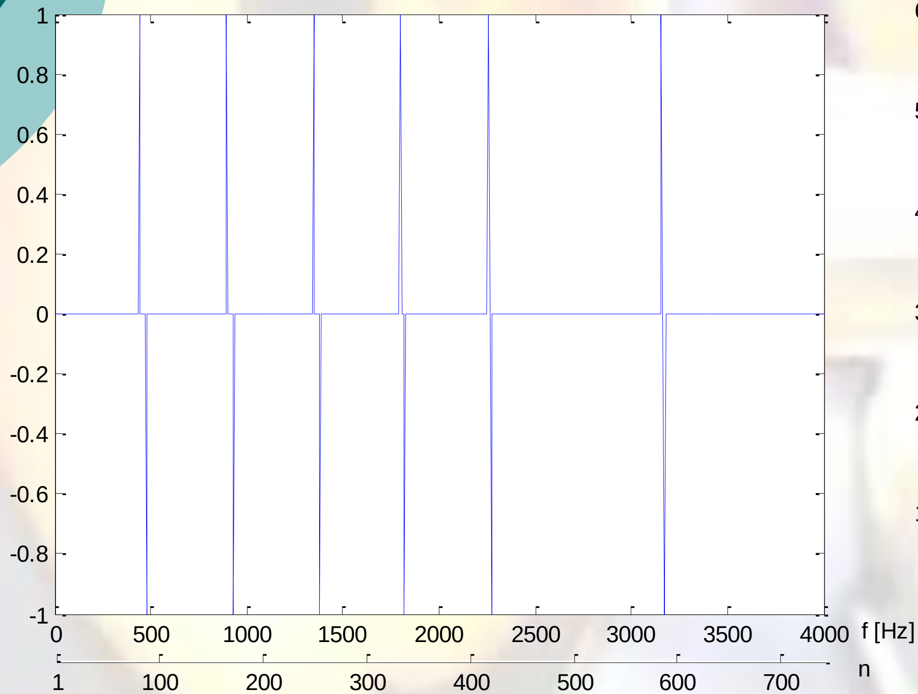
Różniczkowanie

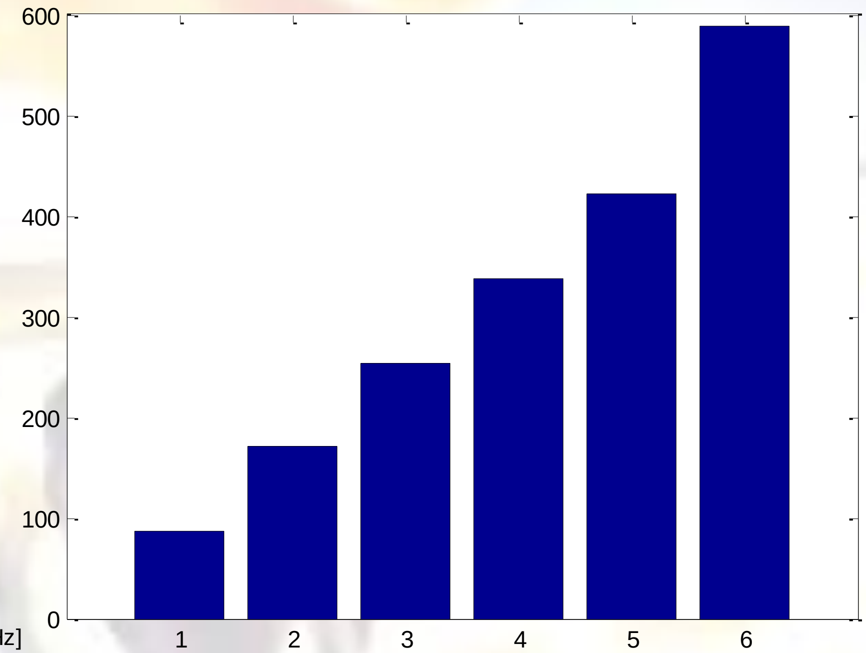




Algorytm detekcji f_0

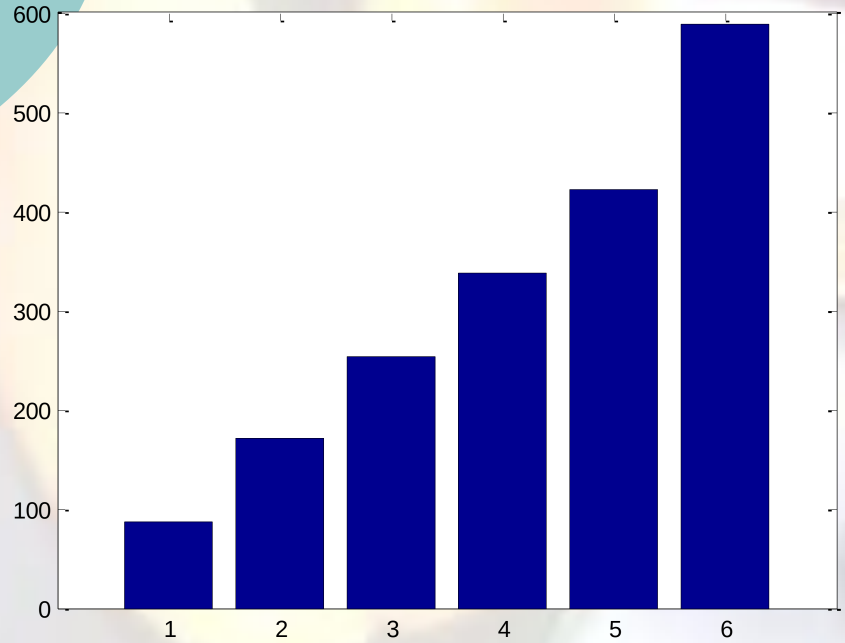
Określenie położenia prążków

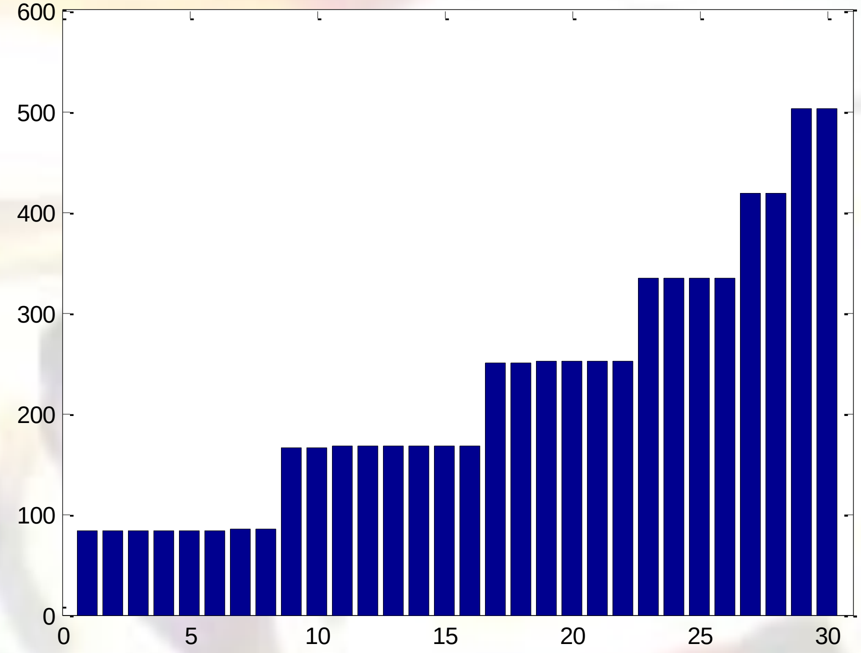




Algorytm detekcji f_0

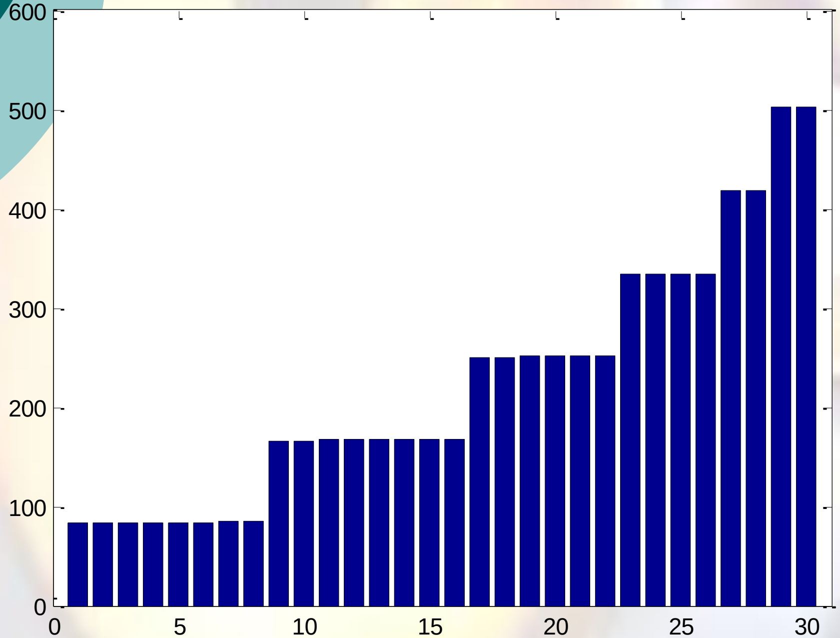
Obliczanie różnic między prążkami

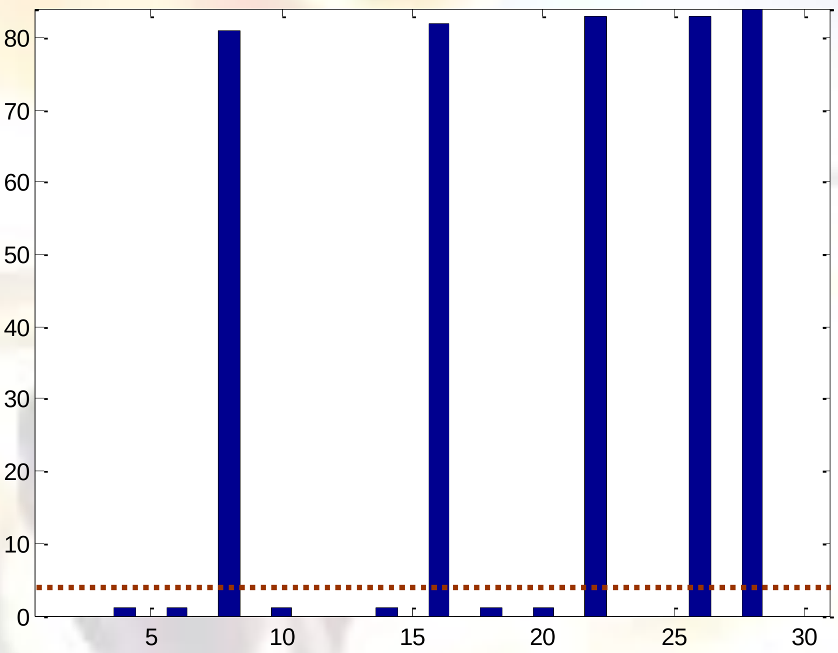




Algorytm detekcji f_0

Określenie częstotliwości podstawowej





f_0 = 450,85Hz A4+42

Podsumowanie

- Detekcja częstotliwości podstawowej jest złożonym zagadnieniem i wykorzystanie odpowiedniego algorytmu uzależnione jest od celu przetwarzania i wymagań stawianych danej metodzie
- Przedstawione algorytmy znacząco różnią się pod względem złożoności obliczeniowej, opóźnienia związanego z rozmiarem przetwarzanych ramek oraz z dokładnością generowanych wyników, która w wielu przypadkach zależna jest od rodzaju analizowanych przebiegów oraz od poziomu szumu w nagraniu