

---

# WYZNACZANIE CZĘSTOTLIWOŚCI PODSTAWOWEJ GŁOSU

---

**Karol Działowski**

Zachodniopomorski Uniwersytet Technologiczny

19 października 2020

## 1 Wstęp

Celem zadania było wyznaczenie częstotliwości podstawowej głosu. Częstotliwość podstawowa głosu (ang. *fundamental frequency*) jest utożsamiana z wrażeniem wysokości dźwięku odbieranym przez nasz mózg. W publikacjach częstotliwość podstawową często oznacza się jako  $F_0$ .

Wyznaczenie częstotliwości podstawowej można przeprowadzić przy zastosowaniu metody autokorelacji. Szacowanie wykonano na dostarczonych próbkach głosu (oznaczanych jako *a\_C3\_ep44.wab* oraz *a\_C4\_ugp44.wav*), dwóch własnych próbkach (*karol\_a\_1.wav* oraz *karol\_a\_2.wav*) a także próbkach syntetycznych *220.wav*, *440.wav*, *440cut.wav*, *774.wav*.

## 2 Wyznaczanie $F_0$ metodą autokorelacji

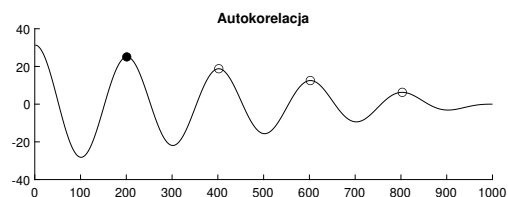
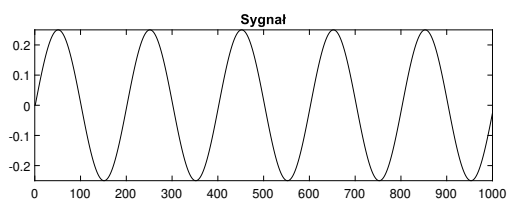
Wyznaczanie częstotliwości podstawowej za pomocą autokorelacji przy użyciu wbudowanych funkcji przedstawiono na listingu 1. Użyto tam funkcji `xcorr` do obliczenia autokorelacji oraz funkcji `findpeaks` do wyznaczenia lokalnych maksimów.

Efekty dla wszystkich wejściowych sygnałów dźwiękowych przedstawiono na wykresie 1.

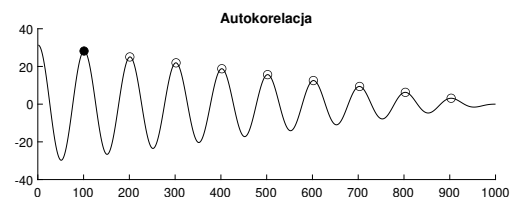
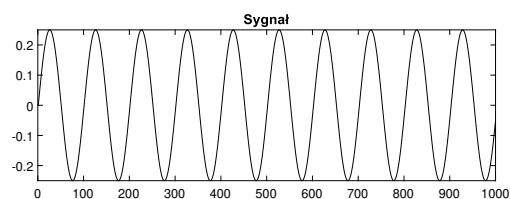
Porównanie wyników przedstawiono w tabeli 1.

```
1 [s,Fs] = audioread(name);
2
3 len = min(1000, length(s));
4 scut = s(1:len);
5 xc = xcorr(scut);
6 xccut=xc(length(scut):end);
7
8 [pks, locs]=findpeaks(xccut);
9 [peak_value, pi] = max(pks);
10 peak_index = locs(pi);
11 f0 = Fs/peak_index
```

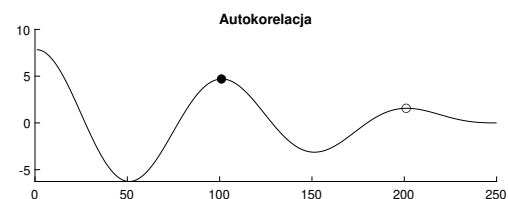
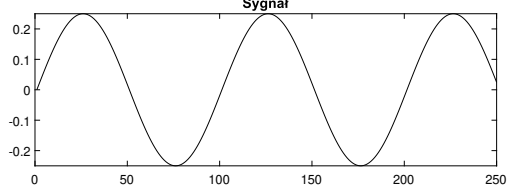
Listing 1: Wyznaczanie  $F_0$ .



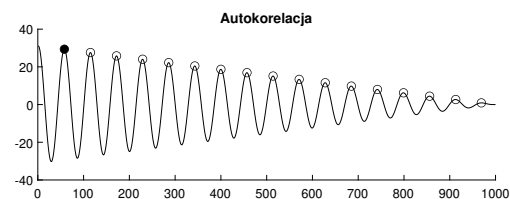
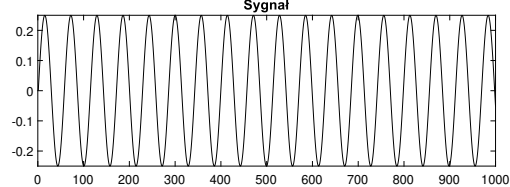
(a) 220 Hz



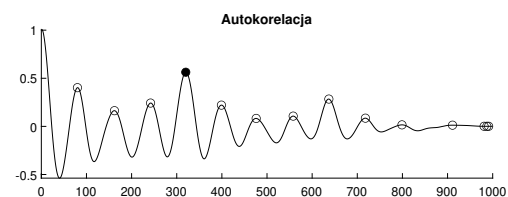
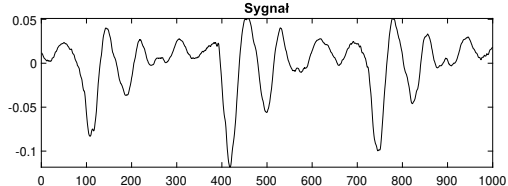
(b) 440 Hz



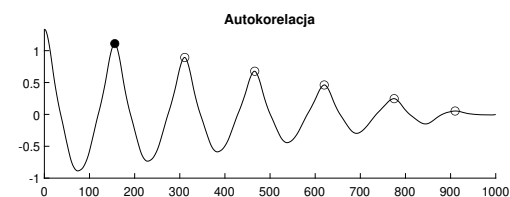
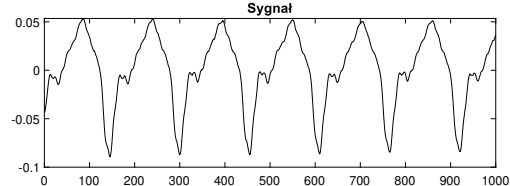
(c) 440 Hz (cut)



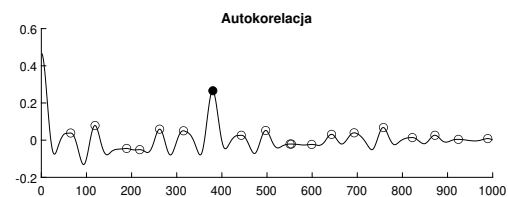
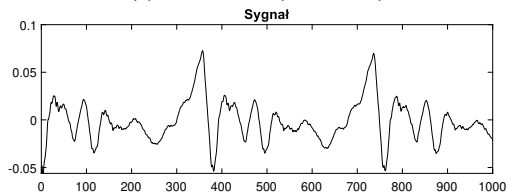
(d) 774 Hz



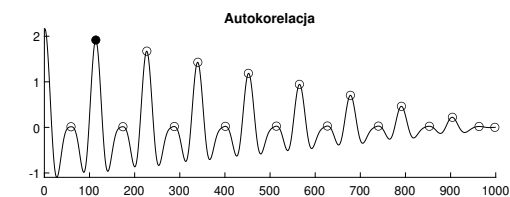
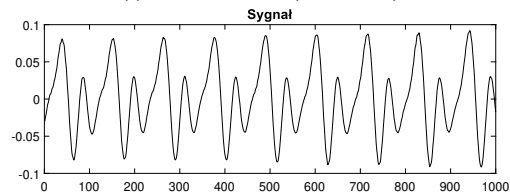
(e) A C3 EP44 (130.81Hz)



(f) A C4 UGP 44 (261.64Hz)



(g) Karol 1 (?Hz)



(h) Karol 2 (?Hz)

Rysunek 1: Sygnały wejściowe. Porównanie wycinka sygnału z autokorelacją wraz z zaznaczonymi maksimumami lokalnymi oraz wybranym punktem maksymalnym.

Nazwa pliku	Wyznaczona częstotliwość $F_0$	Rzeczywista częstotliwość	Błąd
220.wav	219.4030	220	0.5970
440.wav	436.6337	440	3.3663
440cut.wav	436.6337	440	3.3663
774.wav	760.3448	774	13.6552
a_C3_ep44.wav	137.8125	130.81	7.0025
a_C4_ugp44.wav	282.6923	261.64	21.0523
karol_a_1.wav	116.0526		
karol_a_2.wav	386.8421		

Tablica 1: Porównanie wyników otrzymanych w wyniku działania metody bazującej na autokorelacji z wynikami referencyjnymi.

## 2.1 Podsumowanie

Metoda autokorelacji przybliża częstotliwość podstawową w sposób zadowalający w przypadku analizowania wysokości dźwięku śpiewanego przez osobę – błędy wynoszą odpowiednio 7 i 21 Hz.

Zaobserwowano zwiększanie się błędu metody wraz z zwiększaniem częstotliwości.

## 3 Autokorelacja w dziedzinie częstotliwości

Częstotliwość podstawową można też wyznaczyć w dziedzinie częstotliwości za pomocą metody ACOLS. W tym celu należy zrealizować następujące kroki:

1. Pobranie fragmentu nagrania
2. Wyznaczenie widma pobranego fragmentu (fft)
3. Obliczenie wartości bezwzględnej widma
4. Logarytmowanie widma bezwzględnego
5. Autokorelacja
6. Oszacowanie położenia piksu
7. Obliczenie częstotliwości  $F_0$  jako  $f/T$ .

Działanie algorytmu przedstawiono na listingu 2.

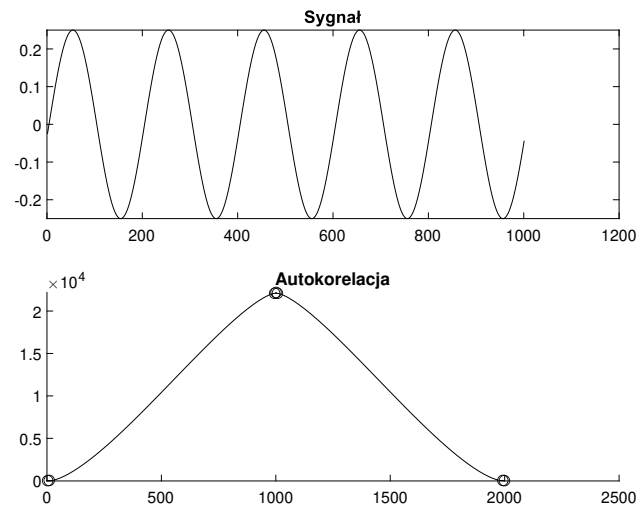
```

1 [s,Fs] = audioread('220.wav');
2
3 ACOLS = xcorr(log(abs(fft(s))));
4
5 [pks, locs]=findpeaks(ACOLS);
6 [peak_value, pi] = max(pks);
7 peak_index = locs(pi);
8 f0 = Fs/peak_index

```

Listing 2: Wyznaczanie  $F_0$  w dziedzinie częstotliwościowej.

Algorytm jednak nie działa prawidłowo. Autokorelacja logarytmu widma bezwzględnego przyjmuje kształt dzwonu, którego maksimum nie przekłada się na częstotliwość fundamentalną. Z tego powodu nie porównano wyników obu metod wyznaczania częstotliwości podstawowej. Działanie metody przedstawiono na rysunku 2.



Rysunek 2: Wyznaczanie  $F_0$  w dziedzinie częstotliwości dla dźwięku 220 Hz.