

Sygnaly 1 – Fizyczne podstawy dźwięku

Projekt „Matematyka dla Ciekawych Świata - edycja XV BIS”
Krzysztof Lasocki

2024-07-22

1 Fizyczne podstawy dźwięku

Dźwięk to zjawisko fizyczne, które odczuwamy słysząc fale dźwiękowe. Jest spowodowany przez fale akustyczne propagujące się (najczęściej) w powietrzu i docierające do naszych uszu.

Fale akustyczne to fale podłużne, czyli takie, w których przesunięcie ośrodka propagacji (np. cząsteczek powietrza) odbywa się w tym samym kierunku, co kierunek rozchodzenia się fali. Drugim rodzajem fal są fale poprzeczne, czyli takie, w których ośrodek porusza się prostopadłe do kierunku rozchodzenia się fali. Przykładem fal poprzecznych są fale na powierzchni wody.

Dźwięk powstaje najczęściej gdy jakiś przedmiot zaczyna wibrować w powietrzu. Może być to np struna wprawiana w wibracje w sposób mechaniczny (np przez młoteczek w fortepianie, albo palce w instrumentach strunowych). Może to też być piszczałka wprawiana w wibracje przez przepływ powietrza, tak jak to ma miejsce w instrumentach dętych. W niektórych przypadkach sam przepływ powietrza może tworzyć fale stojące we wnętrzu, np. flecie albo butelce.

Napięta struna wprawiona w wibracje porusza się, przesuując otaczające ją cząsteczki powietrza. Z powodu chwilowego wzrostu ciśnienia powietrza, sąsiadujące cząsteczki zmieniają położenie, powodując zmianę ciśnienia obok, co porusza kolejne cząsteczki i proces zaczyna się od nowa. Ponieważ cząsteczki gazu mają swoją bezwładność, ich ruch jest opóźniony względem poprzednich. Z tego powodu mija pewien czas aż fala ciśnienia wytworzona w jednym punkcie dotrze do drugiego. Mówimy, że fala propaguje się przez ośrodek z pewną prędkością fazową. Jest to prędkość z którą czoło fali porusza się przez ośrodek, np. powietrze.

2 Wibracje struny

Rozchodzenie się fal opisuje równanie falowe. Jest to równanie różniczkowe drugiego stopnia, którego rozwiązanie zależy od tzw. warunków brzegowych. Nie będziemy wnikać w to jak się je rozwiązuje, natomiast spróbujemy wytłumaczyć jego rozwiązanie dla struny o zamocowanych końcach korzystając z obrazowego opisu zjawiska.

Jak podaliśmy wcześniej, fala akustyczna propaguje się z pewną prędkością v , która jest charakterystyczna dla danego materiału. Aby struna wibrowała stabilnie, długość fali musi być taka, aby propagująca się fala po przebyciu całej długości struny L i odbiciu się od jej obu końców, wróciła do punktu początkowego z taką samą fazą. Inaczej mówiąc, musi być to tzw. **fala stojąca**, czyli taka, której węzły nie zmieniają swojego położenia. Warunkiem koniecznym do tego jest aby w każdym punkcie faza fali była taka sama na początku i na końcu podróży.

Przy odbiciu od zamocowanego końca fala zmienia fazę o 180 stopni („odwraca się”), ale ponieważ następują dwa odbicia, nie wpływa to na nasze rachunki. Łącznie fala przebywa dwie długości struny. Aby faza była identyczna na końcu podróży, fala musi przebyć tę odległość w ciągu jednego okresu T .

Zatem:

$$\frac{2L}{T} = v \quad (1)$$

więc

$$T = \frac{2L}{v} \quad (2)$$

$$f = \frac{1}{T} \quad (3)$$

$$f = \frac{v}{2L} \quad (4)$$

Odwrotność okresu T , czyli częstotliwość f nazywamy częstotliwością podstawową struny. Fala o takiej częstotliwości będzie miała dwa węzły, które znajdują się w umocowanych końcach. Te punkty struny się nie ruszają.

Faza fali zgodzi się również gdy fala przebędzie odległość w czasie będącym wielokrotnością okresu, więc

$$nT = \frac{2L}{v} \quad (5)$$

$$f_n = \frac{nv}{2L} \quad (6)$$

Częstotliwości f_n nazywamy **częstotliwościami harmonicznymi**. Fala o takiej częstotliwości ma więcej węzłów które znajdują się w punktach rozłożonych równo na całej długości struny.

Rzeczywiste drgające struny generują dźwięk będący sumą częstotliwości podstawowej oraz jej harmonicznymi a także zawierające pewne anharmoniczne składowe (np. stuknięcia klawisza w przypadku wysokich dźwięków fortepianu).

W powietrzu, w warunkach normalnych fala akustyczna porusza się z prędkością 343 m/s. Dokładna wartość tej prędkości zależy od wielu czynników takich jak wilgotność, temperatura i ciśnienie powietrza.

Powietrze nie jest jedynym nośnikiem fal akustycznych. Mogą się one propagować w dowolnym innym materiale, np w cieczech i ciałach stałych. W ciałach stałych dźwięk może propagować się również w postaci fal poprzecznych. To zjawisko jest wykorzystywane w kamertonie podczas dotykania go do różnych obiektów aby one również zaczęły emitować dźwięk.

Ciekawostka

Możesz zaobserwować propagację dźwięku w ciałach stałych na różne sposoby, np przykładając ucho do blatu stołu i lekko przesuując po nim palcami lub innym przedmiotem. Będziesz również słyszeć dźwięki, które przepropagowały się z powietrza do blatu.

Zadanie 2.0.1

Oblicz, w jakiej odległości trafił piorun jeżeli od błysku do huków minęło 10 sekund. Załóż, że prędkość dźwięku wynosiła 340 m/s.

Zadanie 2.0.2

Oblicz, jaka jest prędkość dźwięku (to znaczy, z jaką prędkością porusza się fala) w pierwszej strunie gitarowej (E) jeżeli emituje ona częstotliwość podstawową równą $f_E = 329.63\text{Hz}$. Długość strun gitary akustycznej *long scale* wynosi 650mm. Na jaki parametr struny wpływasz, strojąc gitarę?

3 Fale harmoniczne

Najprostszym rodzajem fali jest tzw. fala harmoniczna. Fala harmoniczna jest opisywana funkcją sinusoidalną. Taka fala może rozchodzić się m.in. w napiętej strunie. Fala harmoniczna jest opisywana wzorem:

$$y(t, x) = A \sin(\omega t - kx + \varphi) \quad (7)$$

Argument tej funkcji, zwany fazą, $\omega t - kx + \varphi$, to suma trzech składników:

- ωt to składnik odpowiedzialny za zmianę odchylenia fali w czasie t
- kx to składnik odpowiedzialny za zmianę odchylenia fali w przestrzeni. Dla fali w jednym wymiarze współrzędną oznacza się najczęściej x . Litera k to tzw. liczba falowa, którą zaraz opiszemy dokładniej.
- φ to tzw. faza początkowa, która ustala jakie odchylenie miała fala w punkcie $(t = 0, x = 0)$.

Za natężenie fali odpowiedzialna jest jej **amplituda** A , która mnoży wartość funkcji sinus. Jeżeli rozpatrujemy tylko jedną falę, to możemy pominąć fazę początkową uznając $\varphi = 0$. Wtedy faza to $\omega t - kx$

Przyjrzyjmy się dokładniej wartości fazy. Jest to suma dwóch składników, ωt oraz kx . Zastanówmy się co oznacza ωt . W tym celu ustalmy stałmy $x = 0$. Wtedy, wraz z upływem czasu rośnie wartość ωt , a wraz z nią zmienia się wartość funkcji sinus:

- dla $\omega t = 0$, wartość wychylenia fali wynosi zero ($A \sin(0) = A \cdot 0 = 0$)
- dla $\omega t = \pi/2$, wartość wychylenia fali wynosi A ($A \sin(\pi/2) = A \cdot 1 = A$)
- dla $\omega t = \pi$, wartość wychylenia fali wynosi znów zero ($A \sin(\pi) = A \cdot 0 = 0$)
- dla $\omega t = 3\pi/2$, wartość wychylenia fali wynosi $-A$ ($A \sin(3\pi/2) = A \cdot (-1) = -A$)
- dla $\omega t = 2\pi$, wartość wychylenia fali wynosi znowu zero ($A \sin(2\pi) = A \cdot 0 = 0$)
- dla kolejnych wartości ωt cykl się powtarza, bo sinus to funkcja okresowa z okresem równym 2π .

Jeżeli założymy, że t rośnie z taką szybkością aby okres fali T był równy 1, to wtedy można wyprowadzić wartość ω z warunku

$$\omega T = 2\pi \quad (8)$$

zatem

$$\omega = \frac{2\pi}{T} \quad (9)$$

Jest to wzór na wielkość zwaną **częstotliwością kołową** fali lub po prostu **częstotliwością**¹. Ponieważ wiemy że częstotliwość f oraz okres T to swoje odwrotności, możemy powiązać ω z częstotliwością:

$$\omega = \frac{2\pi}{T} = 2\pi f \quad (10)$$

Zajmijmy się teraz częścią kx . Ustalamy $t = 0$ i podobnie jak dla pierwszego składnika sumy, wnioskujemy, że cały okres sinusa otrzymamy gdy $kx = 2\pi$. Ma ona jakiś związek z przestrzennym aspektem rozchodzenia się fali. Ustalmy znowu, że x ma taką skalę, w której jeden okres fali otrzymujemy gdy $x = 1$. Zatem k musi być równe:

$$k = \frac{2\pi}{x} = \frac{2\pi}{\lambda} \quad (11)$$

Tę specyficzną wartość x nazwiemy długością fali. Oznacza się ją grecką literą λ . Natomiast k to **liczba falowa**. Zauważ, że liczba falowa ma wymiar 1/m.

W ogólnym przypadku rozchodzenia się fali w trzech wymiarach, k staje się wektorem $\vec{k} = [k_x, k_y, k_z]$ i nosi nazwę **wektora falowego**.

Zdefiniowaliśmy prędkość fazową fali jako prędkość rozchodzenia się czoła fali. Doprecyzujemy tę definicję – prędkość fazowa fali to prędkość z jaką rozchodzi się punkt mający stałą fazę, czyli ustaloną wartość argumentu funkcji sinus. W ogólności, dla fal rozchodzących się w trzech wymiarach, mówimy o **powierzchni stałej fazy**.

Zadanie 3.0.1

Gramofon odtwarza muzykę poprzez prowadzenie igły gramofonowej w rowku płyty. Przesunięcia rowka powodują ruch igły który jest zamieniany na sygnał elektryczny przez wkładkę gramofonową. Płyty gramofonowe są nacinane dla kilku prędkościach odtwarzania, m. in. $33\frac{1}{3}$ oraz 45 obrotów na minutę (tzw. “Long play” oraz “Maxi single”).

Z jaką częstotliwością odtworzy się dźwięk A_4 o częstotliwości 440Hz, jeżeli został on nagrany na płycie Long play a odtworzony z prędkością 45 obrotów na minutę?

Wskazówka: Ile razy szybciej będzie wibrować igła?

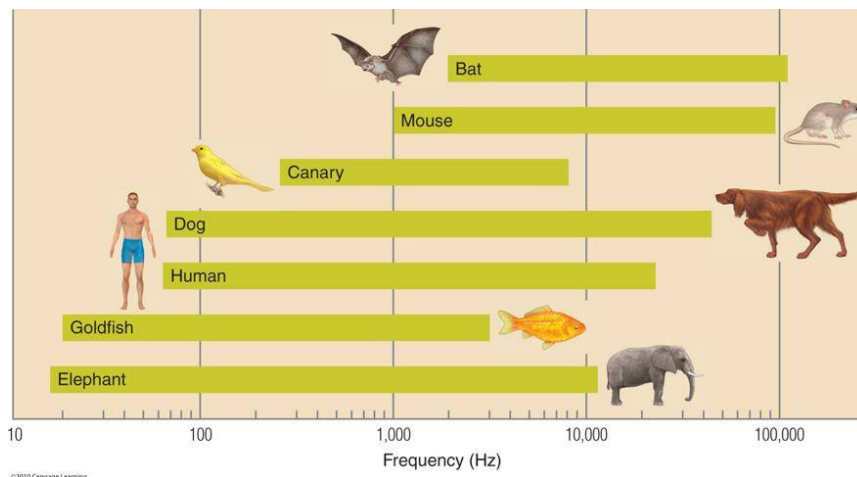
1. Nie należy mylić jej z częstotliwością. W języku polskim częstotliwość i częstotliwość do dwa słowa, jednak w angielskim na częstotliwość kołową mówi się *angular frequency*, a na częstotliwość *frequency*

4 Zakres słyszalnych dźwięków oraz skala muzyczna

4.1 Zakres słyszalności u ludzi i zwierząt

Ludzie słyszą fale akustyczne w zakresie od około 20Hz do 20kHz. Górna granica słyszalności zmniejsza się wraz z wiekiem. Ludzka mowa zawiera się w zakresie od 100 do 5000 Hz.

Zwierzęta mają różne zakresy słyszalnych dźwięków. Wśród zwierząt używających dźwięków do komunikacji, słuch jest najczęściej najbardziej czuły w zakresie który mogą wydawać one same. Najszerszy zakres słyszanych częstotliwości mają koty, delfiny oraz niektóre nietoperze. Najniższe dźwięki (do ok. 7Hz) słyszą niektóre walenie, a najwyższe, delfiny (150kHz) oraz nietoperze (115kHz).



Rysunek 1: Zakres słyszalności zwierząt. Za [4]

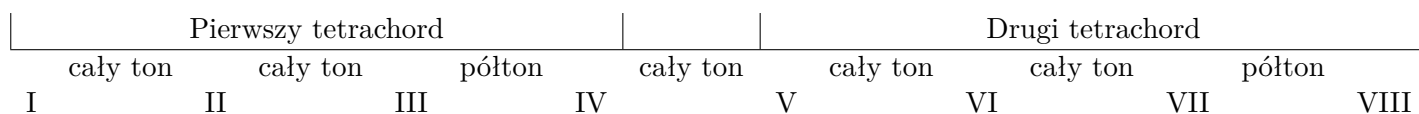
4.2 Skala muzyczna

W tzw. stroju równomiernie temperowanym ustalono częstotliwość dźwięku a^1 . Jest ona zdefiniowana jako 440 Hz. Oktawa jest podzielona na 12 półtonów. Dźwięki oddalone o jedną oktawę mają częstotliwość dwukrotnie wyższą lub niższą, zatem dźwięk a^0 będzie miał 220Hz, a a^2 odpowiada częstotliwość 880Hz.

Strój równomiernie temperowany zakłada, że proporcje częstotliwości kolejnych dźwięków są względem siebie stałe. Ponieważ jedna oktawa to 12 półtonów, to częstotliwości dźwięków oddalonych od siebie o 1 półton są w proporcji $1 : \sqrt[12]{2} \approx 1 : 1.0595$. Odległości między dźwiękami, zwane interwałami, są wielokrotnościami półtonów. Interwał sekundy wielkiej, oddalony o dwa półtony, łączy dźwięki o częstotliwościach w proporcji $1 : (\sqrt[12]{2})^2 \approx 1 : 1.1225$. Tercja wielka to dwa całe tony, czyli 4 półtony, więc proporcja częstotliwości dwóch dźwięków to $1 : (\sqrt[12]{2})^3 \approx 1 : 1.1892$, itd. W końcu oktawa, oddalona o 12 półtonów, daje proporcję $1 : (\sqrt[12]{2})^{12} = 1 : 2$.

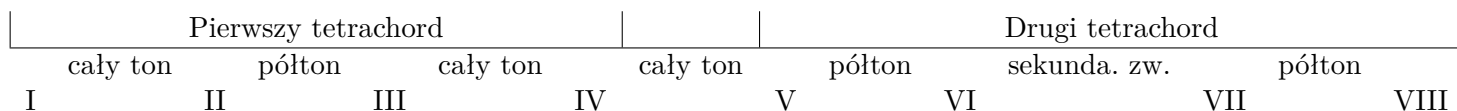
Z tych 12 dźwięków, odpowiadających 12 klawiszom fortepianu w oktawie, w trakcie historii muzyki wyklarowały się różne skale muzyczne. Współcześnie wyróżniamy dwie główne skale składające się z 8 dźwięków, podzielonych na dwa tetrachordy: skalę durową oraz molową.

Skala durowa to skala, w której oba tetrachordy są w proporcji $1 : 1 : 1/2$ tonów, i oddalone od siebie o 1 cały ton:



Rysunek 2: Skala durowa w proporcji $[1 + 1 + 1/2] + 1 + [1 + 1 + 1/2]$.

Skala molowa harmoniczna ma inny układ półtonów – pierwszy tetrachord jest w proporcji 1 : 1/2 : 1, natomiast drugi 1/2 : 3/2 : 1/2:

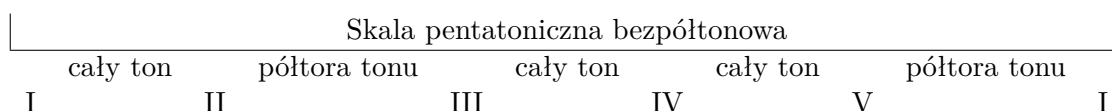


Rysunek 3: Skala molowa naturalna w proporcji $[1 + 1/2 + 1] + 1 + [1/2 + 3/2 + 1/2]$.

Istnieje kilka innych odmian skal molowych, np. melodyczna oraz naturalna. Różnią się one drugim tetrachordem. We wszystkich skalach suma interwałów wynosi 12 półtonów, czyli jedną oktawę.

Zadanie 4.2.1

Znajdź, w jakich proporcjach są do siebie częstotliwości kolejnych dźwięków skali pentatonicznej bezpółtonowej:



Jakie częstotliwości miałyby dźwięki tej skali jeżeli pierwszy byłby o częstotliwości 128Hz?

5 Głośność dźwięku

Częstotliwość dźwięku nie jest jego jedynym parametrem. Dźwięk może być też głośny lub cichy. Fizycznym zjawiskiem odpowiedzialnym za to jest amplituda fali. Większe zmiany ciśnienia powodują głośniejszy dźwięk. Nasze uszy odbierają dźwięk w skali logarytmicznej, oraz mają różną czułość dla różnych częstotliwości.

Do wyrażania wielkości logarytmicznych używa się **decybeli [dB]**, czyli pochodnej jednostek logarytmicznych (Beli). Opisują one proporcję danej wielkości do wielkości odniesienia. Głośność dźwięku używa poziomu odniesienia równego ciśnieniu $20 \mu Pa$. Oznacza to, że dźwięk spowodowany przez falę ciśnienia o amplitudzie $20 \mu Pa$ ma głośność 0 dB SPL (Sound Pressure Level).

Pozostałe wartości są zdefiniowane według wzoru

$$P = 10 \log_{10} \left(\frac{p^2}{p_0^2} \right) \text{ dB} = 20 \log_{10} \left(\frac{p}{p_0} \right) \text{ dB} \quad (12)$$

Liczba 2 przed logarytmem wynika z tego, że natężenie fali to kwadrat jej amplitudy.

Dźwięk, który jest spowodowany falą ciśnienia o natężeniu 10-krotnie większym (100-krotnie większa amplituda) niż ciśnienie odniesienia będzie miał głośność

$$P = 10 \log_{10} \left(\frac{(10p_0)^2}{p_0^2} \right) = 10 \log_{10}(100) = 10 \cdot 2 = 20 \text{ dB SPL} \quad (13)$$

Dźwięk o natężeniu 60 dB SPL (głośność głośnej rozmowy) będzie tworzył fale ciśnienia o natężeniu

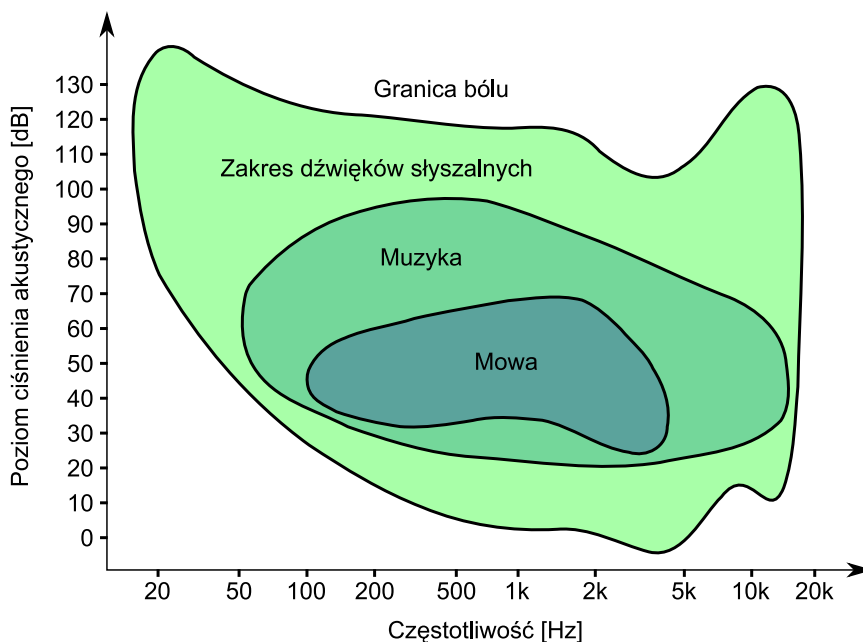
$$60 \text{ dB SPL} = 20 \log_{10} \left(\frac{p}{p_0} \right) \quad (14)$$

$$3 = \log_{10} \left(\frac{p}{p_0} \right) \quad (15)$$

$$10^3 p_0 = p \quad (16)$$

$$p = 20000 \mu Pa = 20 \text{ Pa} \quad (17)$$

Wykres poniżej pokazuje czułość ludzkiego słuchu w zależności od częstotliwości dźwięku. Jak widać, słuch ludzki jest najbardziej czuły w zakresie w którym znajduje się mowa. Najczulszy punkt, znajdujący się w okolicach 3kHz to zakres częstotliwości odpowiadającej płaczu małych dzieci. Górna granica, zwana granicą bólu oznacza natężenia dźwięku tak silne, że wywołują fizyczny ból uszu. Takie natężenia prowadzą do trwałego uszkodzenia słuchu.



Rysunek 4: **Powierzchnia słyszalności**. Na osi pionowej skala dB SPL (domena publiczna) [3]

Zakres słyszalnych głośności dźwięku rozciąga się od 0 do ok. 130dB SPL. Oznacza to, że ludzkie ucho może wykrywać fale dźwiękowe o natężeniach różniących się ponad 10^6 (milion) razy. To ogromny zakres dynamiki!

Zadanie 5.0.1

Jaką zmianę ciśnienia powoduje dźwięk startu odrzutowca słyszanego z odległości 1 m, o głośności 150 dB SPL? Czy takie natężenie dźwięku jest bezpieczne?

6 Wprowadzenie do sygnałów

Na poprzednich zajęciach opisaliśmy fale sinusoidalne oraz jej harmoniczne, a także zdefiniowaliśmy, jakie parametry ma ta fala:

$$y(x, t) = A \sin(\omega t - kx + \varphi) \quad (18)$$

Funkcja $y(t, x)$ opisuje jak zmienia się wartość fali harmoniczej. Załóżmy że jest to fala akustyczna. Wtedy A jest jej amplitudą, czyli mówi, w jakim zakresie fala zmienia lokalne ciśnienie powietrza.

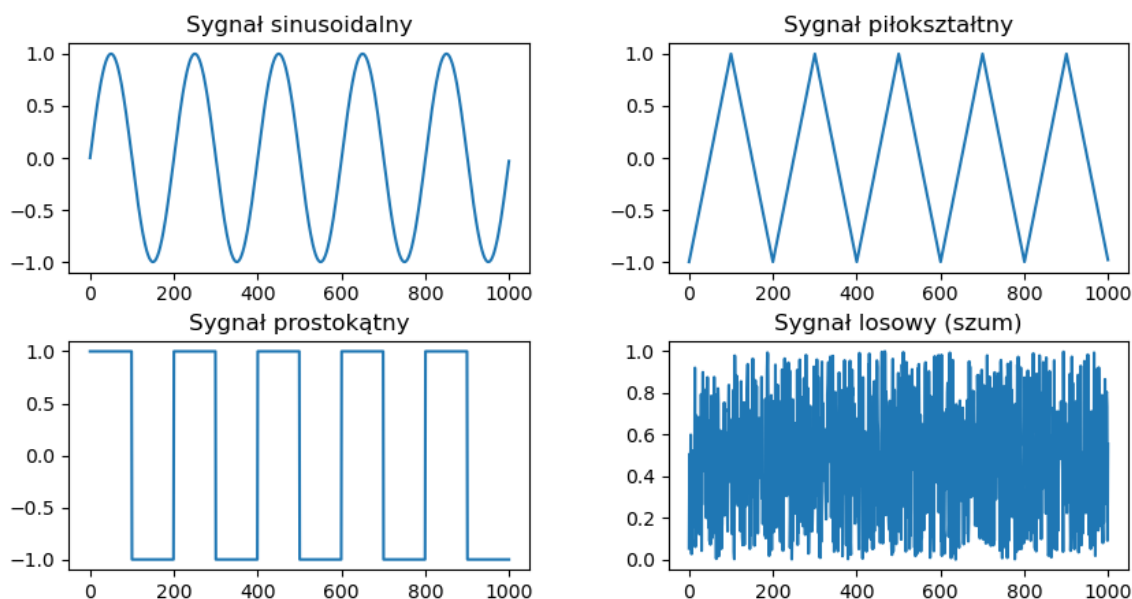
Jeżeli mamy sposób zmierzenia wartości $y(x, t)$ w jakimś punkcie x_0 , np. za pomocą mikrofonu, to otrzymamy funkcję $s(t)$ która jest równa $y(x_0, t)$. x_0 to jakiś wyznaczony punkt, w którym stoi mikrofon. Funkcja $s(t)$, która opisuje jak zmienia się ciśnienie w punkcie x_0 (jaką wartość ma w nim fala akustyczna), jest nazywana **sygnałem**.

W naszym przypadku funkcja $s(t)$ opisana będzie wzorem

$$s(t) = y(x_0, t) = A \sin(\omega t - kx_0 + \varphi) \quad (19)$$

Jeżeli A to jakaś wartość ciśnienia, a s opisuje odchylenie ciśnienia w czasie t , to można powiedzieć, że $s(t)$ to sygnał akustyczny. Podobnie jeśli A byłoby jakąś wartością napięcia, to powiemy, że $s(t)$ opisuje sygnał elektryczny (sygnał napięciowy).

W ogólności dowolną funkcję, która opisuje jak zmienia się jakaś mierzalna wartość w czasie, możemy nazwać **sygnałem**

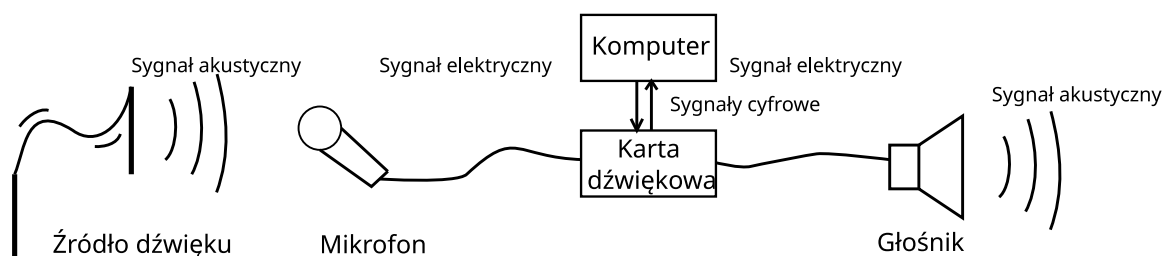


Rysunek 5: **Różne sygnały.**

7 Komputerowe przetwarzanie dźwięku

Korzystając z mikrofonów, toru elektroakustycznego oraz głośników możemy zamieniać sygnały akustyczne na elektryczne oraz odwrotnie. Najprostszym urządzeniem które tego dokonuje, jest telefon analogowy - mikrofon po jednej stronie połączenia zamienia dźwięk (sygnał akustyczny) na sygnał elektryczny, a następnie ten sam sygnał elektryczny jest zamieniany na dźwięk po drugiej stronie.

Możemy również zamieniać sygnał elektryczny (np. napięcie albo natężenie prądu) na wartości liczbowe, oraz odwrotnie - wartości liczbowe na sygnały elektryczne. Służą do tego przetworniki analogowo-cyfrowe (ADC) oraz cyfrowo-analogowe (DAC)². Najprostszym przykładem takiego przetwornika jest karta dźwiękowa komputera, w której przetwornik cyfrowo-analogowy podłączony jest do słuchawek a analogowo-cyfrowy do mikrofonu.



Rysunek 6: **Tor sygnałowy** z wykorzystaniem komputera do przetwarzania dźwięku.

Dzięki karcie dźwiękowej możemy przekonwertować sygnały akustyczne do wartości w pamięci komputera. To pozwala na ich przetwarzanie na bardzo wiele sposobów.

2. ang. *analog to digital converter* oraz *digital to analog converter*

Przykładowo, sygnał z mikrofonu może być zapisany w pamięci komputera, a następnie wysłany do głośnika w późniejszym czasie - w ten sposób za pomocą komputera nagrywamy dźwięk. Można również zsumować kilka sygnałów z mikrofonu (dźwięk fortepianu, śpiew wokalistki), nagranych w różnym czasie - i później dodać je do siebie, aby utworzyć sygnał dźwiękowy, w którym wokalistka śpiewa w akompaniamencie fortepianu.

Sygnał trafiający do głośnika może również zostać wygenerowany sztucznie – na przykład odebrany przez internet (radio internetowe) albo obliczony przez program komputerowy. W ten sposób można syntezować dźwięki nie występujące w naturze.

8 Sygnały dyskretne

W dalszej części zajęć wykorzystamy proste programy w języku Python, aby zgłębiać tajniki syntezy dźwięku. Najpierw jednak przyjrzyjmy się w jaki sposób komputer przechowuje sygnały (takie jak dźwięk) w pamięci.

Aby zreprodukcować sygnał dźwiękowy, powinniśmy za pomocą głośnika wytworzyć taką samą falę jaka padła na membranę mikrofonu, to znaczy, odtworzyć w czasie wartości funkcji 19, która opisuje sygnał dźwiękowy. Problem polega na tym, że aby to zrobić, należy znać wartości $s(t)$ dla każdego t . Zmienna t jest liczbą rzeczywistą, więc nawet w bardzo krótkim okresie czasu, jest nieskończenie wiele tych punktów (mówimy że zbiór \mathbb{R} jest gęsty). Oznacza to że nasz komputer musiałby mieć nieskończenie pojemną pamięć. Takie pamięci z oczywistych przyczyn nie mogą istnieć.

Na szczęście w roku 1949 Harry Nyquist wraz z Claudem Shannonem³sformułowali twierdzenie mówiące, że można odtworzyć dowolny sygnał o ograniczonym zakresie częstotliwości B mając do dyspozycji punkty (próbki) tego sygnału oddalone nie bardziej niż $1/(2B)$. Jest to **Twierdzenie Nyquista-Shannona** zwane również **Twierdzeniem o próbkowaniu**:

Twierdzenie o próbkowaniu

Jeżeli funkcja $x(t)$ zawiera częstotliwości nie większe niż B Hz, to można ją wiernie odtworzyć z szeregu punktów oddalonych od siebie o mniej niż $1/(2B)$ sekund.

Mówiąc inaczej, mając próbki zmierzone co najmniej $2B$ razy na sekundę można odtworzyć sygnał składający się z częstotliwości co najwyżej B Hz.

Twierdzenie o próbkowaniu jest najważniejszym twierdzeniem teorii sygnałów.

O sygnale który jest opisany jako szereg wartości w ustalonych punktach mówimy że jest **dyskretny**, w przeciwieństwie do sygnałów **ciągłych**, które są opisane w każdym punkcie (w danym zakresie czasu).

Sygnały analogowe i cyfrowe

Sygnały analogowe są sygnałami ciągłymi, a cyfrowe dyskretnymi. Sygnały analogowe mają również ograniczone pasmo (nie są w stanie przenieść dowolnie wysokiej częstotliwości), lecz przyczyny tego zjawiska są zupełnie inne.

Sygnały dyskretne w komputerze przechowywane są jako lista (tablica) wartości w pamięci.

Zadanie 8.0.1

Jaka jest najwyższa częstotliwość, którą można odtworzyć z płyty Audio CD? Dźwięk na płycie jest zapisany w formie cyfrowej z częstotliwością próbkowania 44.1kHz. Czy taka częstotliwość próbkowania wystarczy, aby zapisać wszystkie dźwięki słyszane przez ludzi?

3. Oryginalnie to twierdzenie opisał Sir Edmund T. Whittaker w 1915, a w podobnym okresie (1933) zostało ono również sformułowane przez rosyjskiego matematyka V. Kotelnikowa, z tego powodu znane jest również pod ich nazwiskami.

9 Synteza prostych tonów

Do syntezy sygnałów użyjemy języka Python. Będziemy korzystać z gotowych skryptów uruchamianych w środowisku Jupyter.

Zacznijmy od syntezy prostego tonu. Aby to zrobić, wygenerujemy listę kolejnych wartości zmiennej t (czasu), a następnie dla każdej z nich obliczymy $s(t)$ ze wzoru zapisanego w kodzie programu.⁴

Naszą pracę z komputerem musimy zacząć od wczytania kilku bibliotek, które pozwolą nam korzystać z gotowych fragmentów kodu i funkcji:

```
from IPython.lib.display import Audio
import librosa, librosa.display
import matplotlib.pyplot as plt
import numpy as np
from xvbis import *
```

Wygenerujmy prosty ton harmoniczny, obliczając kolejne wartości funkcji \sin . Skorzystamy z funkcji pomocniczej `czas(długość, częstotliwość_próbkowania)` która wygeneruje tablicę kolejnych wartości czasu dla próbek, tak abyśmy uzyskali dźwięk o danym czasie trwania.

```
sr = 16000      # Częstotliwość próbkowania
T = 2          # liczba sekund
# Generujemy wektor wartości czasu jako ciąg kolejnych wartości
t = czas(T, sr)
f = 440        # Częstotliwość w Hz
# Obliczamy s(t) = sin(2*pi*f*t) dla każdego t
ton_sin = np.sin(2*np.pi*f * t)
Audio(ton_sin, rate=sr) # Wynik w postaci pliku dźwiękowego
```

Funkcja `Audio` pozwala nam posłuchać wynikowego sygnału. Musi znajdować się jako ostatnia w komórce.

Możemy wykreślić pierwsze sto próbek za pomocą funkcji `plt.plot`:

```
plt.plot(ton_sin[:100])
```

Początek tablicy próbek w pamięci komputera ma następującą postać

```
print(ton_sin[:100])
```

```
array([0.          , 0.1719291 , 0.33873792, 0.49545867, 0.63742399,
       0.76040597, 0.86074203, 0.93544403, 0.98228725, 0.99987663])
```

Możemy również narysować spektrogram wygenerowanego sygnału:

```
spektrogram(ton_sin)
```

oraz obliczyć jego transformatę Fouriera za pomocą:

```
fourier(ton_sin)
```

4. Zasadniczo przedstawiony program tworzy listę (wektor) wartości t a następnie przekazuje go do funkcji sinus, która zwraca wektor obliczonych wartości. Nie ma tutaj jawnie pętli, natomiast komputer liczy te wartości po kolei.

Aby zsyntezować kilka tonów granych jednocześnie wystarczy zsumować ze sobą kilka sygnałów. Sumę należy podzielić aby nie wyjść poza zakres $(-1; 1)$:

```
sr = 16000
T = 4
t = czas(T, sr)
a4 = np.sin(2*np.pi* 440 * t)
a5 = np.sin(2*np.pi* 880 * t)
oktawa = (a4 + a5)/2
Audio(oktawa, rate=sr)
```

Natomiast aby odegrać dwa sygnały jeden po drugim, musimy złączyć je jeden za drugim (dokonać konkatenacji). Służy to tego funkcja `np.concatenate`:

```
dwatony = np.concatenate((a4, a5))
Audio(dwatony, rate=sr)
```

Zwróć uwagę na podwójne nawiasy w funkcji `np.concatenate`. Są one potrzebne, ponieważ jej argumentem jest krotka⁵, która w Pythonie oznaczana jest przez parę `()`.

Można skonkatelować więcej niż dwa sygnały - wtedy zostaną one zagrane po kolei:

```
T = 1      # sekund
t = czas(T, sr)

# Możesz zdefiniować więcej dźwięków, jeśli znasz ich częstotliwości

c4 = np.sin(2*np.pi*261.62 * t)
e4 = np.sin(2*np.pi*329.62 * t)
g4 = np.sin(2*np.pi*391.99 * t)

# Konkatenacja - sklejanie sygnałów "jeden za drugim"
akord = np.concatenate((c4, e4, g4))
Audio(akord, rate=sr)
```

Zadanie 9.0.1

Stwórz program, który zagra po kolei akord C-Dur (C_4, E_4, G_4) a następnie jego pierwszy przewrót (E_4, G_4, C_5).

Tabela częstotliwości w Hz:

Dźwięk	Częstotliwość	Dźwięk	Częstotliwość
C_4	261.63	G_4	392.00
$C\#_4$	277.18	$G\#_4$	415.30
D_4	293.66	A_4	440.00
$D\#_4$	311.13	$A\#_4$	466.16
E_4	329.63	B_4	493.88
F_4	349.23	C_5	523.25
$F\#_4$	369.99		

5. krotka to struktura danych odpowiadająca matematycznej n -ce (dwójce, trójce, czwórce..). Łączy ona kilka obiektów w grupę.

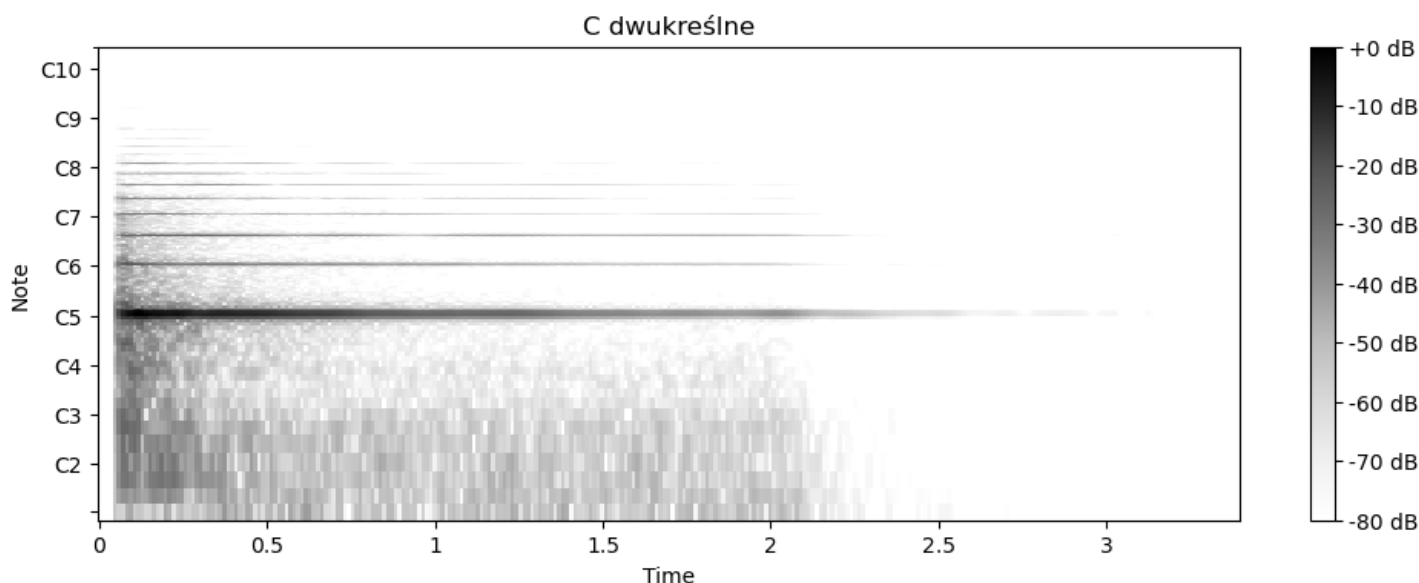
10 Spektrogramy

Jak pokazaliśmy, dźwięk ma różne parametry, które mogą go opisywać. Dźwięk może składać się z pewnej sumy częstotliwości (częstotliwości harmonicznych) oraz mieć różną głośność (natężenie dźwięku). Można zrobić wykres pokazujący jak zmienia się natężenie oraz częstotliwość dźwięku w czasie. Taki wykres nazywamy spektrogramem. Jest to wykres trójwymiarowy - na osi poziomej znajduje się czas, na osi pionowej częstotliwość, a jasność punktu oznacza natężenie dźwięku.

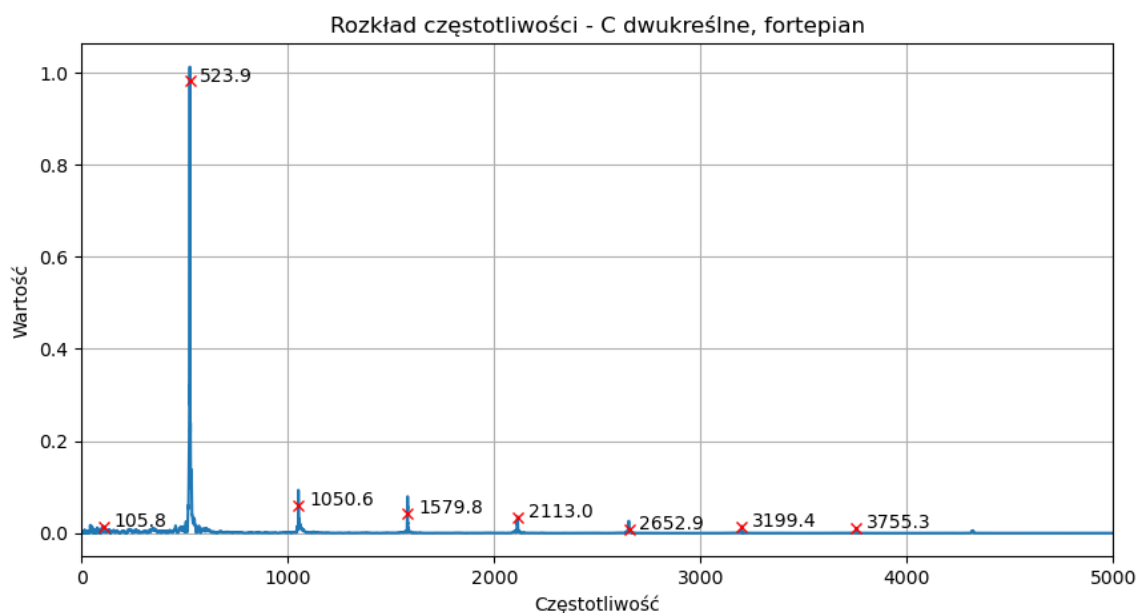
Spektrogramy pozwalają zwizualizować strukturę oraz zawartość nagrania dźwiękowego.

Spektrogram dźwięku, który składa się z pojedynczego sinusoidalnego tonu będzie przedstawiał jedną poziomą linię. Dzieje się tak dlatego, że taki dźwięk ma stałą częstotliwość przez cały czas trwania.

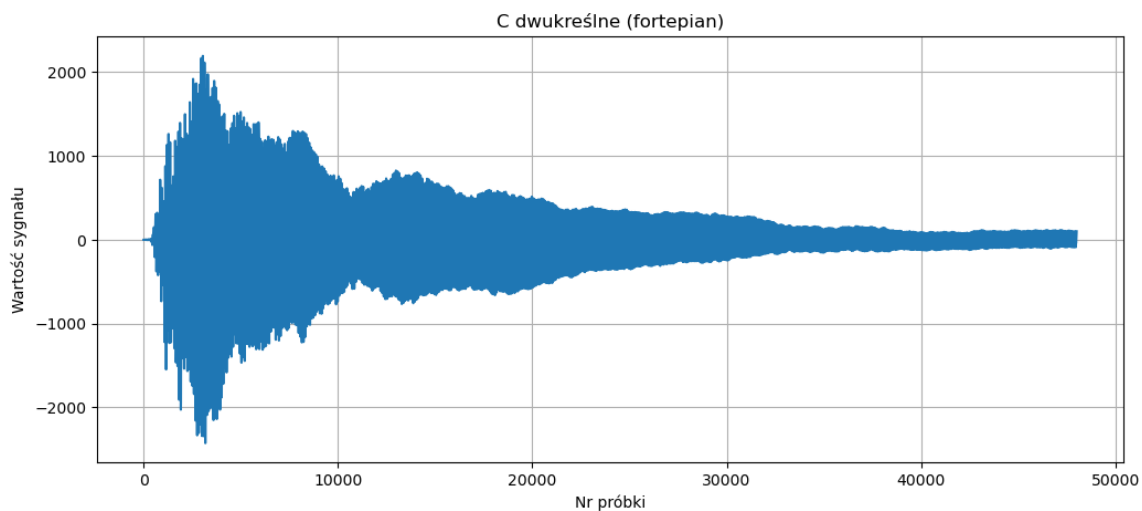
Spektrogram rzeczywistych dźwięków instrumentów będzie zawierać kilka poziomych linii odpowiadających częstotliwości podstawowej oraz jej harmonicznych. Zależnie od barwy tego dźwięku i jego trwania mogą one zanikać szybciej lub wolniej



Rysunek 7: Spektrogram dźwięku c² granego na fortepianie. [6]

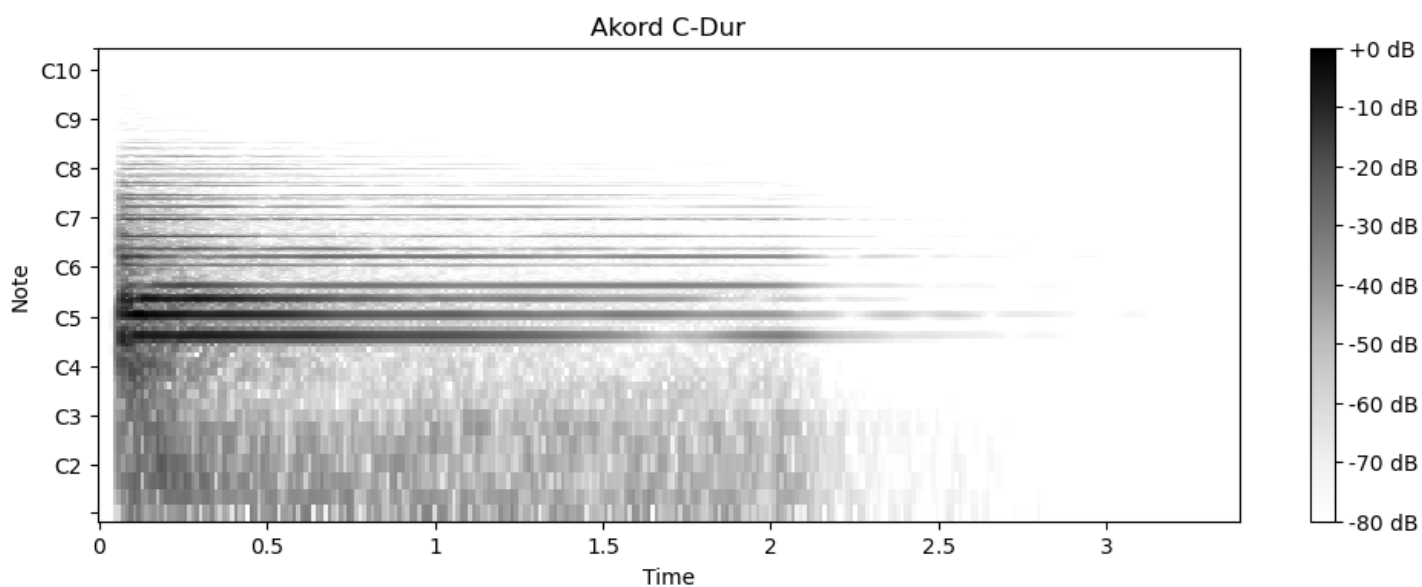


Rysunek 8: Rozkład częstotliwości dźwięku c². Oznaczone są wartości w Hz. [6]



Rysunek 9: Sygnał dźwięku c^2 z poprzedniego spektrogramu. [6]

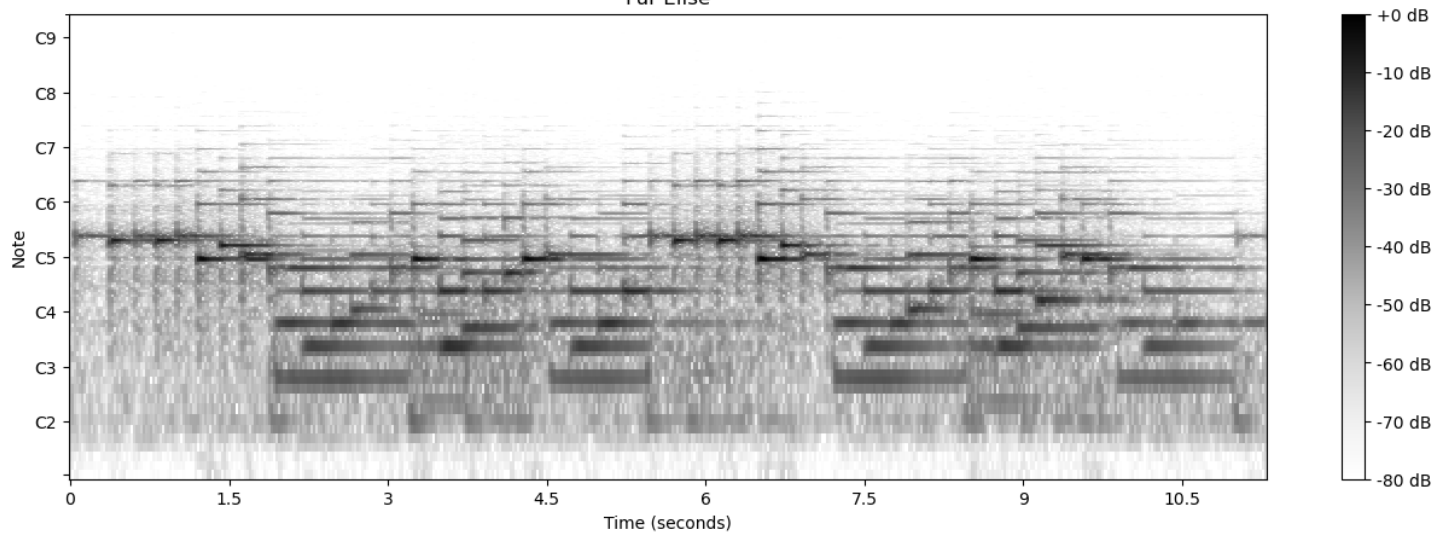
Spektrogram akordu będzie zawierał odpowiednio więcej linii:



Rysunek 10: Spektrogram akordu C-Dur (w przewrocie, G-C-E) granego na fortepianie. [6]

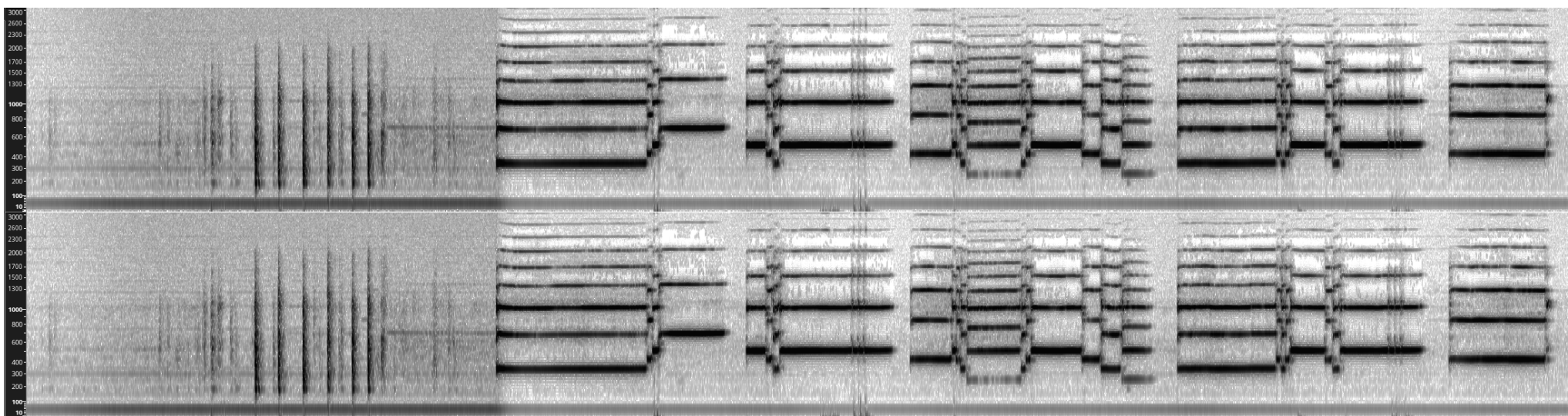
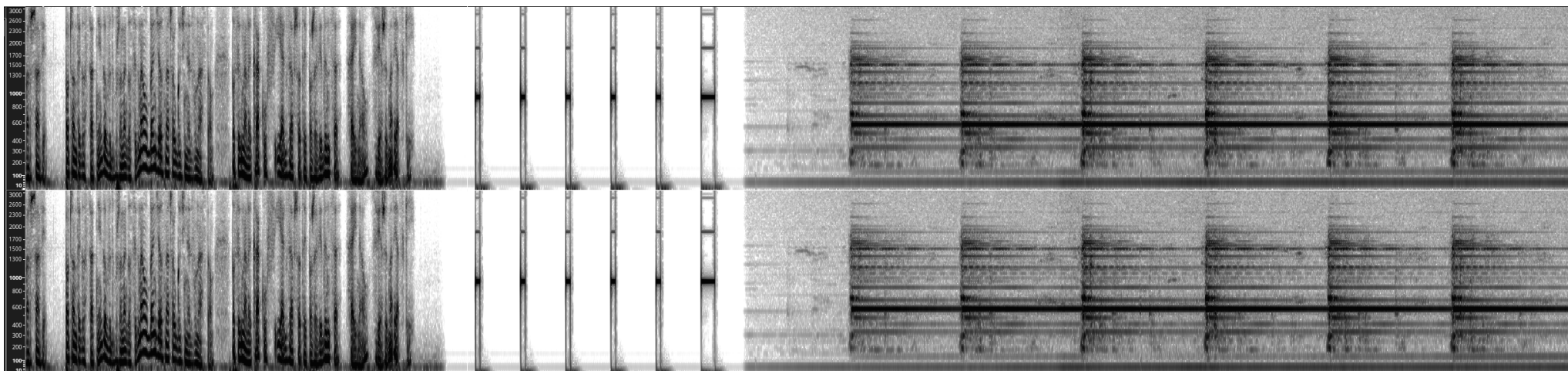
Grając melodię na instrumencie emitujemy z niego kolejno różne wysokości dźwięków. Spektrogram będzie przedstawiał je jako kolejne grupy linii. Tworzy się swojego rodzaju charakterystyczny obraz, z którego można odczytać wysokości kolejnych dźwięków oraz ich barwę a także charakter (staccato, vibrato, legato, itd.)

Für Elise



Rysunek 11: Spektrogram początkowej frazy utworu Ludwika Van Beethovena „Dla Elizy”, granego na fortepianie. [7]

Poniżej znajduje się przykład spektrogramu fragmentu audycji radiowej zawierającej Hejnał Mariacki grany codziennie o godzinie 12. Skala po lewej oznacza częstotliwość w Hz.



Rysunek 12: **Spektrogram transmisji Hejnału**. Pionowa skala to skala mel, która jest wyskalowana tak aby wysokość na skali odpowiadała wysokości dźwięku jaką odbiera ludzka percepcja. Nie jest to skala liniowa względem częstotliwości ani skali harmonicznej. Wartości podane są w Hz.

Zadanie 10.0.1

Przyjrzyj się spektrogramowi hejnału - jak myślisz, dla czego w połowie pierwszego spektrogramu, tuż po sygnale czasu, „tło” robi się szare?

Wskazówka: Co musiał zrobić realizator programu radiowego, aby przenieść słuchaczy z Laboratorium Czasu i Częstotliwości GUM w Warszawie, na Wieżę Mariacką w Krakowie?

11 Zadania dodatkowe

Zadanie 11.0.1

Basia oglądała pokaz sztucznych ogni. Widownia stała na brzegu jeziora, a rakiety były wystrzeliwane z barki stojącej 400m od brzegu. Na jakiej wysokości wybuchła rakietka, jeżeli Basia usłyszała huk po 1.5s od błysku? Załóż, że rakietka leciała pionowo do góry.

Zadanie 11.0.2

Zakładając, że trzecia struna gitarowa emituje dźwięk G_3 o częstotliwości 196Hz, oblicz, w jakim punkcie gitarzysta musi złapać strunę aby skrócić ją tak, by zagrać dźwięk

- C_4
- o 3 półtony wyższy od G_3
- o częstotliwości 220 Hz

Załącz, że długość struny wynosi 650mm

Zadanie 11.0.3

Jaką długość fali będzie miał zapis magnetyczny tonu o częstotliwości 1kHz nagrany na kasecie magnetofonowej? Prędkość przesuwu taśmy wynosi $1 \frac{7}{8}$ cali/sek. $1\text{cal} \approx 25.4\text{mm}$

Zadanie 11.0.4

Skoro już wiesz jak można generować dźwięki oraz akordy o zadanej długości, spróbuj napisać program, który syntezuje wybraną (lub skomponowaną) melodię.

Przydatne wskazówki:

- Możesz skorzystać ze zmiennych aby przechowywać w nich fragmenty melodii, np konkretny akord lub frazę
- Pauza (cisza) to sygnał składający się z samych zer. Możesz go wygenerować za pomocą funkcji np. `zeros(t*sr)`, która wygeneruje Ci t sekund ciszy.
- Częstotliwość dźwięku o oktawę wyżej jest dwukrotnie większa, a o oktawę niżej jest dwukrotnie mniejsza. Możesz użyć operatora mnożenia oraz dzielenia aby Python obliczył ją za Ciebie. (np. wyrażenie `440.00 * 2` w programie da wynik `880.00`)

Tabela częstotliwości w Hz:

Dźwięk	Częstotliwość	Dźwięk	Częstotliwość
C_4	261.63	G_4	392.00
$C\#_4$	277.18	$G\#_4$	415.30
D_4	293.66	A_4	440.00
$D\#_4$	311.13	$A\#_4$	466.16
E_4	329.63	B_4	493.88
F_4	349.23	C_5	523.25
$F\#_4$	369.99		

Zadanie 11.0.5

W jakiej tonacji odsłuchamy Mazurka A-moll F. Chopina, jeżeli przez pomyłkę odtworzymy nagranie z częstotliwością próbkowania 48kHz, gdy oryginalnie było nagrane z częstotliwością próbkowania 44.1kHz?

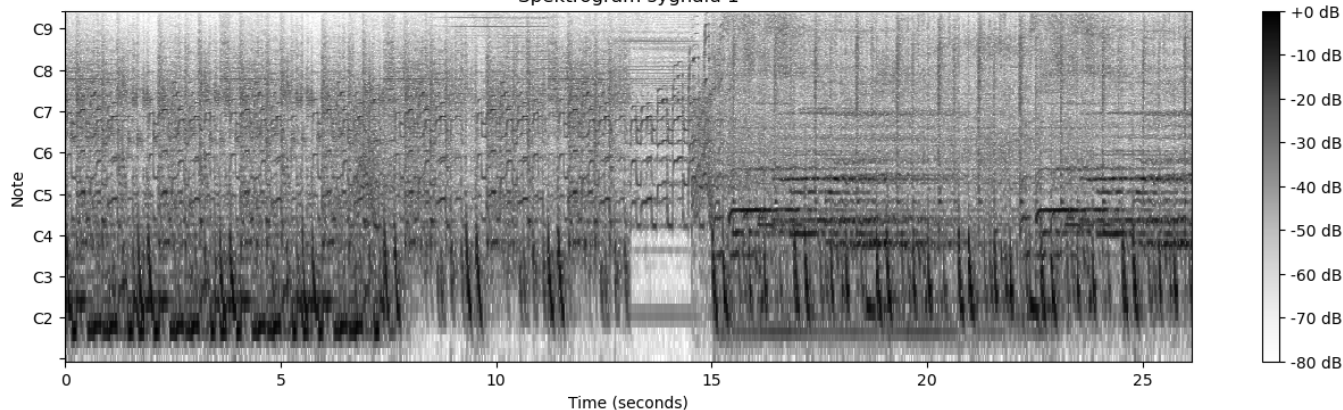
Zadanie 11.0.6

Poniżej przedstawiono trzy spektrogramy około półminutowych fragmentów nagrań dźwiękowych:

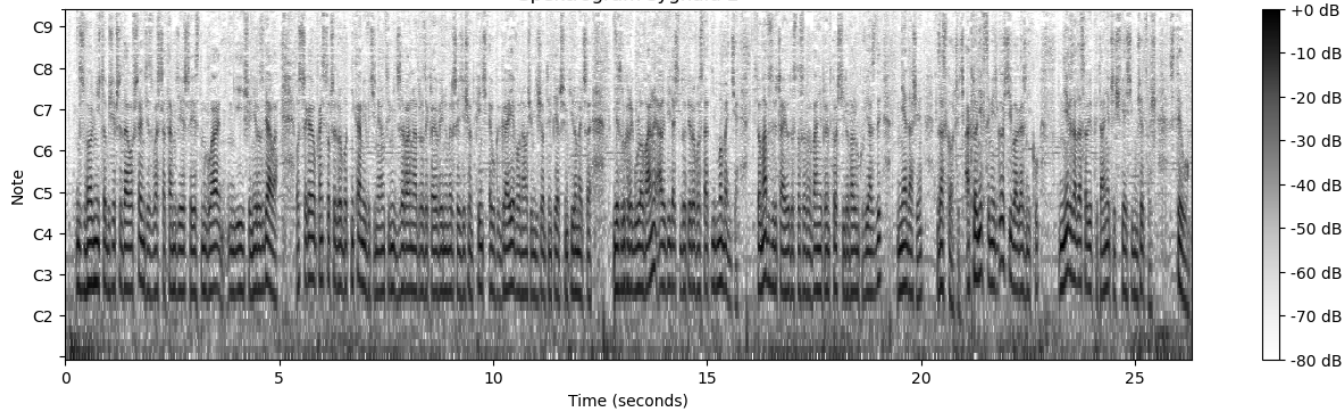
- J.S. Bach - Suita No. 1 w G, gra Misha Maisky, <https://www.youtube.com/watch?v=mGQLXRT13Z0>
- Future Sound of London - Expander, https://www.youtube.com/watch?v=K0vazJ_o6N0
- Fragment audycji radiowej w której spiker czyta wiadomości, <https://www.youtube.com/watch?v=IKWdpGBBtcU>

Spróbuj przyporządkować spektrogram do każdego nagrania.

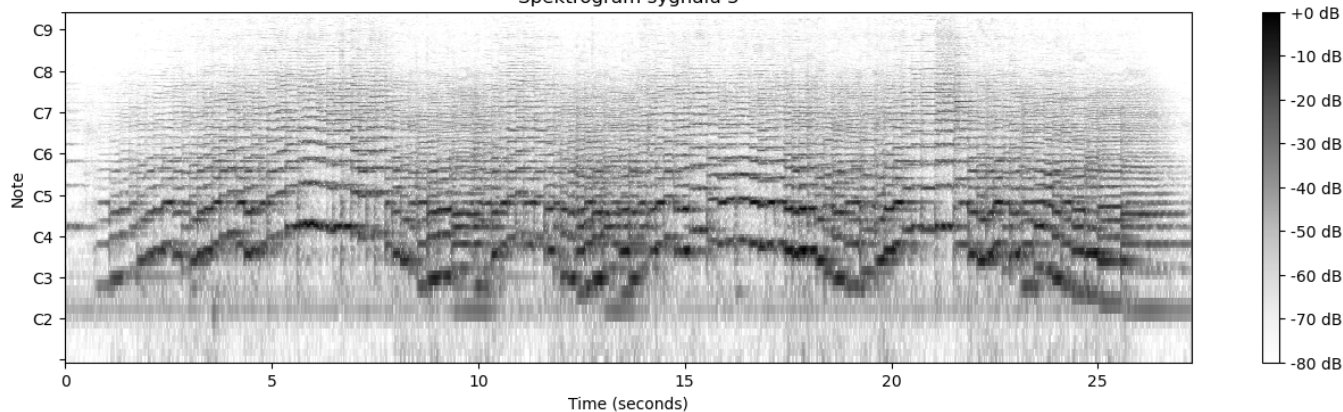
Spektrogram sygnału 1



Spektrogram sygnału 2



Spektrogram sygnału 3



Literatura

- [1] Józef. K. Lasocki - Podstawowe wiadomości z nauki o muzyce, Kraków 2010, PWM SA
- [2] Fritz Winckel - Osobliwości Słyszania Muzycznego, Warszawa 1965, PWN SA
- [3] Powierzchnia Słyszalności, domena publiczna, źródło https://commons.wikimedia.org/wiki/File:Powierzchnia_slyszalnosci.svg
- [4] Źródło: Wstęp do kognitywistyki, B16.1 Czuć i percepcja, Włodzisław Duch, UMK, <http://fizyka.umk.pl/~duch/Wyklady/WDK/16-percepcja.htm>
- [5] Na podstawie <https://www.nti-audio.com/en/support/know-how/fast-fourier-transform-ft>
- [6] Próbki instrumentów muzycznych wykorzystane do wykresów przygotował Daniel Laskowski.
- [7] Na podstawie Mutopia Project - Für Elise, https://www.mutopiaproject.org/ftp/BeethovenLv/Wo059/fur_Elise_Wo059/fur_Elise_Wo059-let.pdf