Instrukcja

Analiza Sygnału
w Dziedzinie
Czasu i Częstości
dr Teodor Buchner

POLITECHNIKA WARSZAWSKA
WYDZIAŁ FIZYKI
PRACOWNIA FIZYKI UKŁADU KRĄŻENIA CZŁOWIEKA
efizyka.if.pw.edu.pl/twiki/bin/view/ACC/

KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

PRZYGOTOWANIE I REALIZACJA SPECJALNOŚCI WSPÓŁFINANSOWANE ZE ŚRODKÓW UNII EUROPEJSKIEJ W RAMACH EUROPEJSKIEGO FUNDUSZU SPOŁECZNEGO

Ćwiczenie 1 *Jaka to melodia?*

# Wprowadzenie

Instrumenty muzyczne dzielimy na dęte strunowe i perkusyjne. W instrumentach strunowych źródłem dźwięku jest struna szarpana (np gitara) lub uderzana (np. fortepian). Teoria opisująca drgania, które zostają wzbudzone w sztywnej strunie jest bardzo złożona, natomiast w pierwszym przybliżeniu można opisać te drgania za pomocą liniowego równania falowego.

$$\frac{\partial^2 y}{\partial t^2} = \nu^2 \frac{\partial^2 y}{\partial x^2} \tag{1}$$

Prędkość rozchodzenia się fali można wyrazić jako:

$$\nu = \sqrt{\frac{T_0}{\rho_0}} \tag{2}$$

gdzie $T_0$ - naprężenie struny w położeniu równowagi, zaś $\rho_0$ - gęstość liniowa (masa jednostki długości) struny. Dla danej struny wartości te są stałe, co oznacza że w pierwszym przybliżeniu prędkość dźwięku w strunie można uznać za stałą, niezależną od wysokości tego dźwięku. W konsekwencji w strunie mogą zostać wzbudzone takie drgania, dla których na obu końcach struny będą znajdować się węzły fali, zaś na długości struny będzie się znajdować nieparzysta wielokrotność połówek długości fali. Długość fali jest związana z jej częstotliwością w sposób następujący:

$$\nu = \lambda f \tag{3}$$

Jeśli spoczynkowa długość struny (tzw menzura) wynosi $L$ (dla swobodnej struny gitary akustycznej menzura wynosi $L = 620\,mm$, dla elektrycznej $L = 624\,mm$), w takim razie możemy spodziewać się wzbudzenia fal o następujących długościach ($k$ oznacza ułamek $\frac{1}{k}$ długości struny odpowiadający połowie długości fali):

$$\lambda_k = \frac{2L}{k} \tag{4}$$

Odpowiada to częstotliwościom:

$$f_k = \frac{\nu}{2L} \cdot k \tag{5}$$

Jeżeli częstość podstawową dla $k = 1$ oznaczymy jako $f_1$, widać że w strunie mogą powstać kolejne wielokrotności tej częstości - tzw *częstotliwości wyższe harmoniczne (overtones)*.

# Elementy akustyki: strój równomierny, temperowany

Podstawowym interwałem stosowanym w akustyce i analizie harmonicznej jest oktawa. Dwa dźwięki różnią się o oktawę jeśli ich częstości mają się do siebie jak 1:2. Jeśli ton podstawowy to $f_1$ to ton o oktawę niższy będzie miał częstotliwość $\frac{f_1}{2}$, zaś ton o oktawę wyższy: $2f_1$. Ton o dwie oktawy wyższy to częstość $4f_1$, zaś o dwie oktawy niższy $\frac{f_1}{4}$ i tak dalej. Oktawa dzielona jest na dwanaście półtonów. Nazwy tych półtonów to:

$$C_m, C\#_m, D_m, D\#_m, E_m, F_m, F\#_m, G_m, G\#_m, A_m, A\#_m, B_m,$$

gdzie $m$-numer oktawy. Częstotliwości półtonów w stosunku do $C_m$ w historii bywały dobierane różnie, natomiast od XIX wieku wprowadzono powszechnie tzw strój równomiernie temperowany (equally tempered) oparty na tonie podstawowym $A_4 = 440Hz$ (warto zwrocić uwagę że wtedy kolejne nuty oznaczone są literami $A, B, C, D, E, F, G$). System równomiernie temperowany jest to strój muzyczny, który polega na podziale oktawy na dwanaście równych półtonów. Stosunek częstotliwości dwóch kolejnych dźwięków w systemie równomiernie temperowanym wynosi $\sqrt[12]{2}$ (Wiki). Za ton podstawowy przyjmuje się ton $A_4$ (dźwięk A w 4-tej oktawie) o częstotliwości $f_A = 440\,Hz$. W muzyce istnieje jeszcze wiele różnych strojów (Wiki). Oktawy są numerowane liczbami naturalnymi od 0. Przykładowo zakres dźwięków w pianinie obejmuje oktawy od 0 ($A_0$) do 7 ($A_7$). Częstotliwość każdego dźwięku $f_p$ można wyrazić w liczbie półtonów $p$, które dzielą go od dźwięku $A_4$:

$$f_p = f_A \cdot 2^{\frac{p}{12}} \tag{6}$$

$$p = 12 \cdot (\ln_2 f - \ln_2 f_A) \tag{7}$$

Widać że skala dźwięków którą człowiek uznaje za równomierną jest w istocie skalą logarytmiczną. Dzieje się tak ponieważ ucho ludzkie nie percepuje bezwzględnej różnicy częstości $\Delta f$ a jedynie względną: $\frac{\Delta f}{f}$, w związku z tym odstępy pomiędzy kolejnymi dźwiękami rosną proporcjonalnie do częstości tych dźwięków. Z relacji (6) można łatwo pokazać iż względna różnica częstotliwości sąsiednich półtonów jest stała, niezależna od ich częstotliwości i równa: $\left(2^{\frac{1}{12}} - 1\right)$.

Ciekawa jest obserwacja położeń kolejnych harmonicznych tonu podstawowego na skali równomiernie temperowanej. Tabela (1) przedstawia rząd harmonicznej oraz interwał tej harmonicznej od tonu podstawowego, wyrażony w liczbie półtonów, zaś wyliczony wg wzoru (7) jako $n_k = 12ln_2k$. Dodatkowo podana jest nazwa interwału używana w muzyce. Tam gdzie interwał ten jest przybliżony, nazwa podana jest w nawiasie. Jak

| Rząd harmonicznej $k$ | Interwał $n_k$ w półtonach | Stosunek częstości | Nazwa interwału |
|---|---|---|---|
| 1 | 0 | 1:1 | pryma |
| 2 | 12 | 2:1 | oktawa |
| 3 | 19.020 | $\sim 3:2$ | (oktawa + kwinta) |
| 4 | 24 | 4:1 | 2 oktawy |
| 5 | 27.863 | $\sim 7:3$ | (2 oktawy + tercja) |
| 6 | 31.020 | $\sim 5:2$ | (2 oktawy + kwinta) |
| 7 | 33.688 | $\sim 14:5$ | (2 oktawy + septyma) |
| 8 | 36 | 8:1 | 3 oktawy |

Tablica 1: Harmoniczne tonu podstawowego na skali równomiernie temperowanej

widać z tabeli (1), wyższe harmoniczne częstości podstawowej przypadają na dźwięki oddalone od tonu podstawowego o jedną lub kilka oktaw ($k = 1, 2, 4, 8$) lub też w przybliżeniu o oktawę lub kilka oktaw i kwintę ($k = 3, 6, 12$). Dzięki temu że odstrojenie częstości jest żadne w przypadku oktaw a niewielkie w przypadku kwint, dźwięki te nazywamy konsonansami doskonałymi. Są one podstawą akordów durowych. Jak widać kolejne grupy harmonicznych ($k = 5, 10, 15$) oraz ($k = 7, 14, 21$) są oddalone od tonu podstawowego odpowiednio o

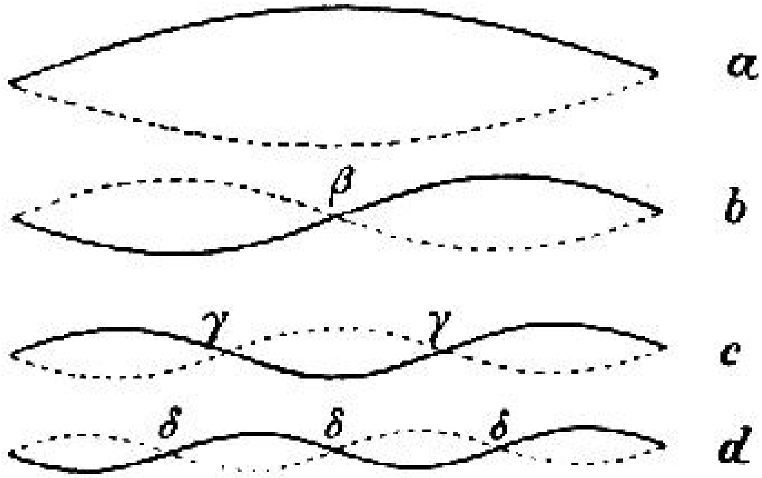



**Fig. 320 Images of the base tones ( *a* ) and the first three overtones ( *b, c, d* ).**

Rysunek 1: Fale stojące dla struny. Wg Rainer Radok, Acoustics

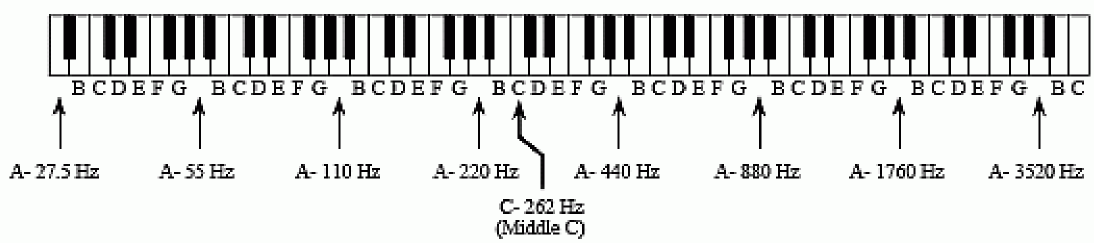
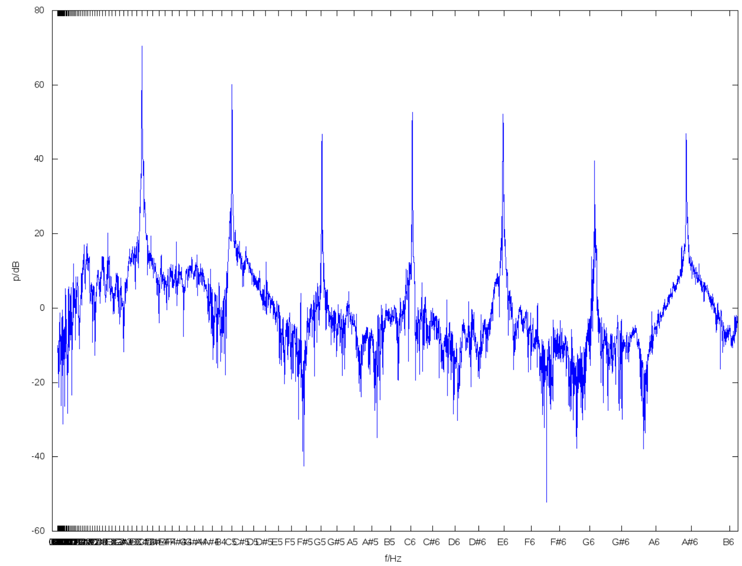



FIGURE 22-4
The Piano keyboard. The keyboard of the piano is a *logarithmic* frequency scale, with the fundamental frequency doubling after every seven white keys. These white keys are the notes: *A, B, C, D, E, F and G*.

Rysunek 2: Podział dźwięków na oktawy. Wg The Scientist and Engineer's Guide to Digital Signal Processing By Steven W. Smith, Ph.D.

`http://www.dspguide.com/ch22/2.htm`

pewną ilość oktaw i tercję bądź septymę. Są to tzw konsonanse niedoskonałe. Już starożytni zauważyli że ładnie współbrzmią (con-sonans - stąd konsonans) dźwięki, których częstotliwości stanowią wielokrotności. Opisuje to zjawisko tzw skala naturalna, w której interwały opisane są przez stosunki częstości postaci $p/q$, gdzie $p$ i $q$ to małe liczby naturalne (Wiki). Niestety skonstruowanie takiej skali, w której wszystkie częstości dźwięków byłyby swoimi wielokrotnościami jest niemożliwe, ponieważ wtedy różnice pomiędzy kolejnymi dźwiękami stają się zbyt duże. Z uwagi na to w wieku XIX zaproponowano skalę równomiernie temperowaną, która wprawdzie wprowadza niewielkie odstrojenia dla kwint i trochę większe dla tercji i septym, za to stosunek częstości sąsiednich dźwięków jest zawsze stały i równy $\sqrt[12]{2}$.




Rysunek 3: Widmo mocy dźwięku $C_4$ pianina.

Przykład widma mocy obejmującego ton podstawowy i sześć pierwszych wyższych harmonicznych przedstawiony jest na rys. 3. Oś pozioma jest w skali liniowej żeby podkreślić fakt iż skala równomiernie temperowana ma charakter logarytmiczny (różnice częstotliwości między kolejnymi dźwiękami zależą od częstotliwości). Tonem podstawowym jest ton $C_4$. Jak widać kolejne harmoniczne przypadają (w przybliżeniu) na dźwięki (w nawiasach podano interwały: $C_5$(oktawa)$, G_5$(kwinta)$, C_6$(oktawa), $E_6$(tercja), $G_6$(kwinta)$, A\#_6$(septyma). W praktyce o barwie dźwięku decyduje pierwszych 10-12 harmonicznych. W przedziałach pomiędzy harmonicznymi widoczne jest widmo o niezerowej mocy. Jego źródła są dwojakie. Pierwszym źródłem są wyższe częstości (nieharmoniczne) - tzw alikwoty. W pierwszym przybliżeniu teorii alikwoty nie występują. To one decydują o barwie dźwięku: dla dwóch różnych instrumentów rozkład mocy alikwot będzie różny.

Wszystkie te dźwięki zagrane jednocześnie z tonem podstawowym będą ładnie współbrzmieć, tworząc konsonans (w przeciwieństwie do dysonansu który ładnie nie brzmi). Przykładowo trójdźwięk toniczny w tonacji $C$ składa się z dźwięków $C$, $E$ i $G$ a więc tonów oddalonych odpowiednio o tercję i kwintę od tonu postawowego. Widać że te interwały nie są przypadkowe a wynikają z lokalizacji częstości harmonicznych drgającej struny. Konsonanse czyli współbrzmienia są podstawą harmonii - nauki traktującej o łączeniu dźwięków.

# Wprowadzenie do Matlaba

Proszę zapoznać się z podstawami Matlaba, np: `http://if.pw.edu.pl/~sobiech/matlab.pdf`.

# Wykonanie ćwiczenia

**Środowisko sugerowane:** *Matlab.*

**Dane:**

- chord.wav - akord w formacie WAV,
- ex1.m - skrypt główny w formacie Matlab,
- tones.m - skrypt pomocniczy w formacie Matlab,

Załaduj plik wav z dysku funkcją wavread lub load (jeśli nie ma wavread). Wykonaj wykres w funkcji czasu. Oceń stacjonarność sygnału (melodia czy akord). Wyznacz dla całego sygnału lub dla odpowiednich części transformatę Fouriera np z oknem Hamminga (lub dowolnym ale świadomie wybranym) a następnie widmo mocy. Wykonaj wykres widma mocy w jednostkach dB/Hz, dla przedziału 16Hz - 4kHz. Następnie zidentyfikuj akord lub melodię nagrane w pliku (= podaj wszystkie dźwięki wchodzące w skład akordu/melodii). Sugerowana metoda postępowania: obetnij wykres widma mocy (ale nie samo widmo) w pionie i w poziomie tak aby zostały tam tylko najwyższe piki w skali podwójnie logarytmicznej. Nastębnie dane zwracane przez funkcję tones wykorzystaj jako źródło położeń ticków oraz etykiet poziomej osi. Sposób użycia podany jest na końcu skryptu `tones.m`. Można także wykorzystać kod funkcji notes i napisać metodę która bierze częstość oraz częstość podstawową i zwraca najbliższy ton ze skali równomiernie temperowanej oraz ewentualnie odstrojenie do niego w jednostkach bezwzględnych (Hz) i względnych (podzielone przez samą częstość albo wyrażone w % interwału). Następnie stwórz metodę, która zidentyfikuje piki a następnie ich centralne częstości zmieni w nazwy półtonów. W przypadku pracy nad melodią trzeba zacząć od metody która zidentyfikuje chwile rozpoczęcia kolejnych dźwięków - dobrze widoczne na wykresie w funkcji czasu.
Procedury do użycia:
`waveread, fft, hamming, conj, log10, plot, loglog, xlabel, ylabel, find, power`.

**Uwagi**:

1. Można korzystać ze źródeł w sieci pod warunkiem podania wszystkich źródeł.
2. Do zamiany obiektu cell na stringa służy instrukcja str = cell2str( cell )
3. Częstości zmierzone mogą różnić się nieco od częstości teoretycznych ponieważ instrument nie był idealnie nastrojony.
4. Uczciwość badacza nakazuje nie odsłuchiwać pliku melodii przed identyfikacją ;-).
5. Każdy dźwięk akordu składa się z tonu podstawowego oraz bogatej kolekcji alikwot (tonów wyższych), których obecność wynika z tego że widmo Fouriera pojedynczego dźwięku nie jest prążkiem.
6. Alikwoty mają częstości wyższe od częstości tonu podstawowego (czytaj: należy zacząć analizę od częstości najniższych).
7. Wśród alikwot pojedycznego dźwięku wyodrębnione są (na tle tła typu 1/f) prążki wyższych harmonicznych tonu podstawowego.
8. Wyższe dźwięki akordu mogą pokrywać się z alikwotami tonów niższych - na tym polega zjawisko współbrzmienia - asonansu (=dlatego jest ładnie).
9. Dźwięki należące do akordu powinny mieć natężenia większe niż natężenia w obrębie alikwot. Ale to nie musi być regułą.

10. Wskazówka. Akord jest w tonacji dur. Można odnaleźć w sieci wszystkie akordy durowe, wyznaczyć ich częstości dla każdej oktawy i porównać ich widmo teoretyczne z widmem otrzymanym. Na przykład można policzyć korelację widma teoretycznego z widmem doświadczalnym. W praktyce widmo teoretyczne lepiej w takim razie opisać jako krzywe Gaussa o pewnej szerokości.

11. Dźwięków w akordzie jest 6, mogą się pojawiać dźwięki wyższe o oktawę lub kilka oktaw od tonu podstawowego lub od dowolnego z tonów trójdźwięku durowego.

**Treść i forma sprawozdania**: forma elektroniczna (Word, TeX, HTML, pdf, ps, ...), sprawozdanie ma zawierać nazwę próbki, widmo mocy, opis metody identyfikacji częstości oraz zidentyfikowany akord: 6 tonów lub melodię zapisaną nazwami półtonów. Do sprawozdania należy dołączyć komplet wykorzystywanych skryptów Matlaba lub zapis sesji jeśli ktoś pracował interaktywnie.