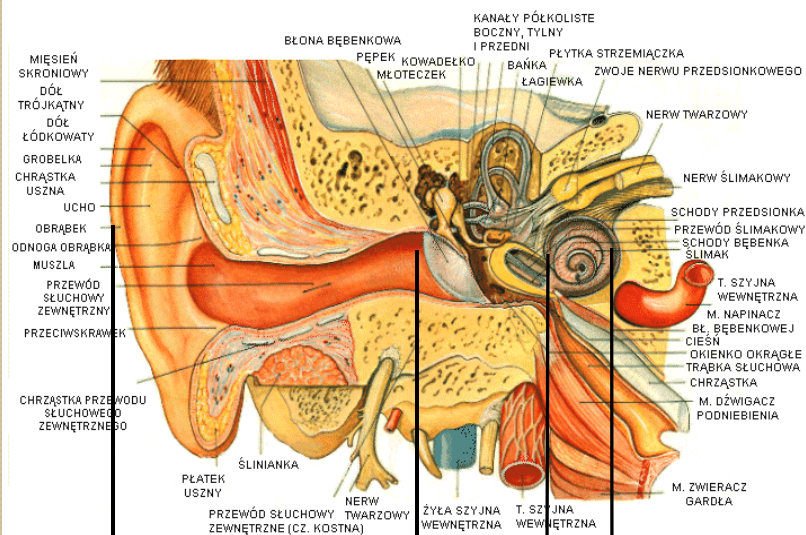


Percepcja dźwięku



Narząd słuchu



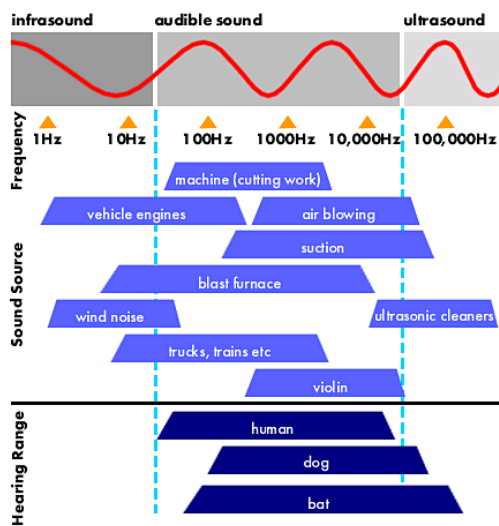
Narząd słuchu

- Ucho zewnętrzne – składa się z małżowiny i kanału usznego, zakończone błoną bębenkową, doprowadza dźwięk do ucha środkowego poprzez drgania błony bębenkowej;
- Ucho środkowe – układ kosteczek słuchowych (młotek, kowadełko, strzemiączko), tworzy transformator między powietrznym środowiskiem ucha zewnętrznego a ciekłym środowiskiem ucha wewnętrznego.
 - Średnie wzmocnienie sygnału wynosi ok. 30 dB (44x).

Narząd słuchu

- Ucho wewnętrzne – zajmuje się „analizą” i „przetworzeniem” dźwięku na szereg bodźców „elektrycznych” i za pomocą nerwów słuchowych przesyła informację dźwiękową do wyższych pięter słuchowych.
 - zasadniczą częścią ucha wewnętrznego jest ślimak, którego istotną częścią jest błona podstawna z narządem Cortiego
 - układ komórek słuchowych na błonie podstawnej i nerwów słuchowych ma układ tonotopowy - działa jak analizator częstotliwościowy

Zakres słyszenia człowieka



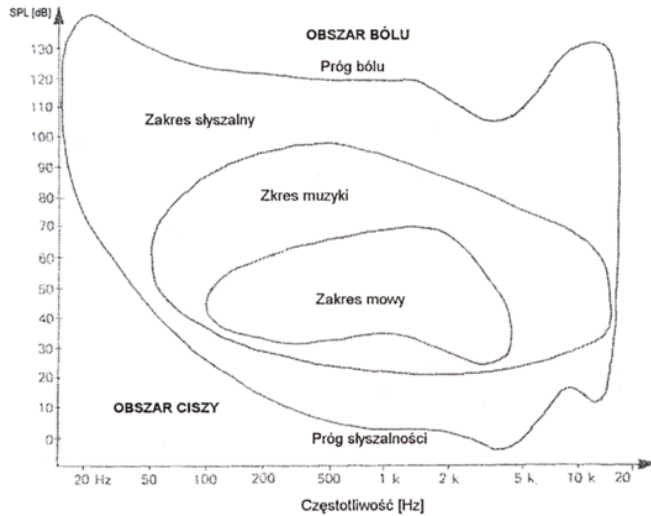
- zakres słyszanych przez człowieka częstotliwości: 16 Hz - 20 kHz
- 10 oktaw

- zakres dynamiki słuchu: 130 dB

Skala muzyczna

- Interwał muzyczny – określony stosunek częstotliwości dwóch dźwięków muzycznych.
- Dźwięk 2x wyższy w muzyce określany jest jako wyższy o oktawę.
- Oktawę tworzy 12 półtonów.
- Różnica częstotliwości sąsiadujących półtonów wynosi: $\sqrt[12]{2}$

Zakres słyszenia człowieka



Dźwięk – zjawisko fizyczne

- Dźwięk jest falą (akustyczną) rozchodzącą się w danym ośrodku sprężystym (ciała stałe, ciecze i gazy);
- Dźwięk jest falą zdolną do wywołania wrażenia słuchowego;

Ton

- Najprostszym rodzajem drgań źródła dźwięku są drgania sinusoidalne;
- Ton prosty to dźwięk składający się z drgań o jednej częstotliwości;

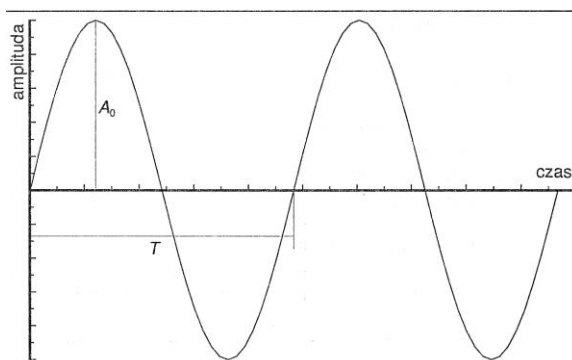
$$A(t) = A_0 \sin(2\pi ft)$$

gdzie:

- A_0 – maksymalne wychylenie (amplituda),
- f – częstotliwość drgań

Ton

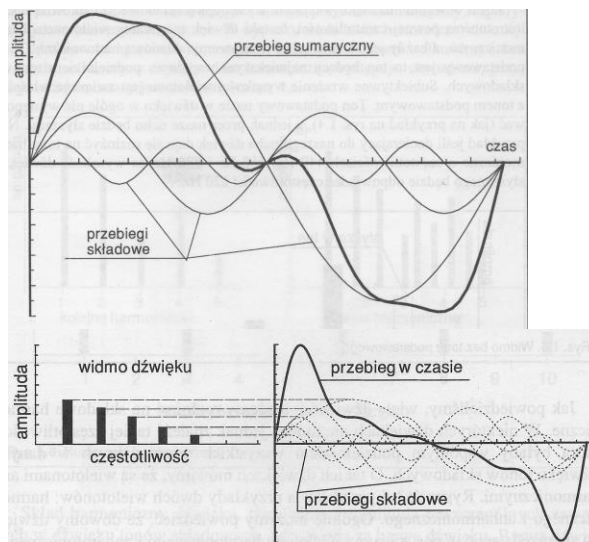
- 1Hz = 1 drganie w ciągu 1 sekundy;
- Okres (T)= czas pełnego jednego drgania;
- $f = 1/T$;



Wieloton

- Na ogół mamy do czynienia z dźwiękami złożonymi, które można przedstawić jako sumę tonów prostych.
- Dokonując analizy dźwięku za pomocą transformacji Fouriera uzyskuje się sumę składowych dźwięków prostych.
- Zbiór składowych sinusoidalnych dźwięku złożonego, określonych przez amplitudę i częstotliwość każdej składowej, nazywa się widmem akustycznym dźwięku;

Wieloton

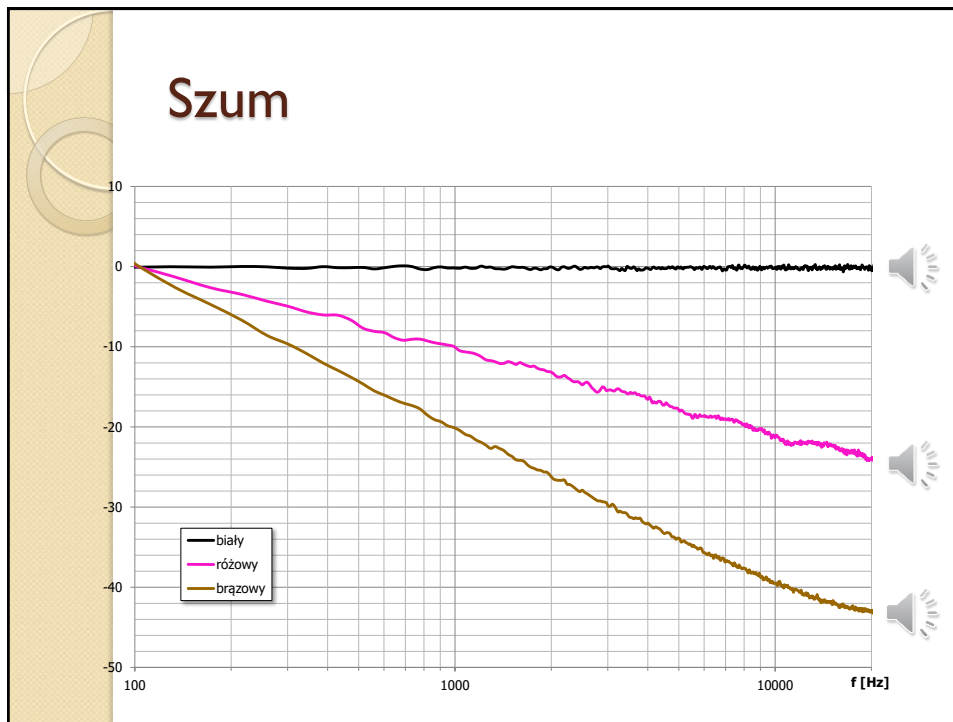


Wieloton

- wieloton harmoniczny
 - w widmie dźwięku częstotliwości składowych są wielokrotnością pewnej częstotliwości, będącej największym wspólnym dzielnikiem tych składowych,
 - kolejne składowe nazywa się harmonicznymi dźwięku;
- Ton podstawowy to ton o częstotliwości będącej największym wspólnym dzielnikiem pozostałych częstotliwości składowych harmonicznymi.
- O wysokości dźwięku decyduje ton podstawowy;

Szum

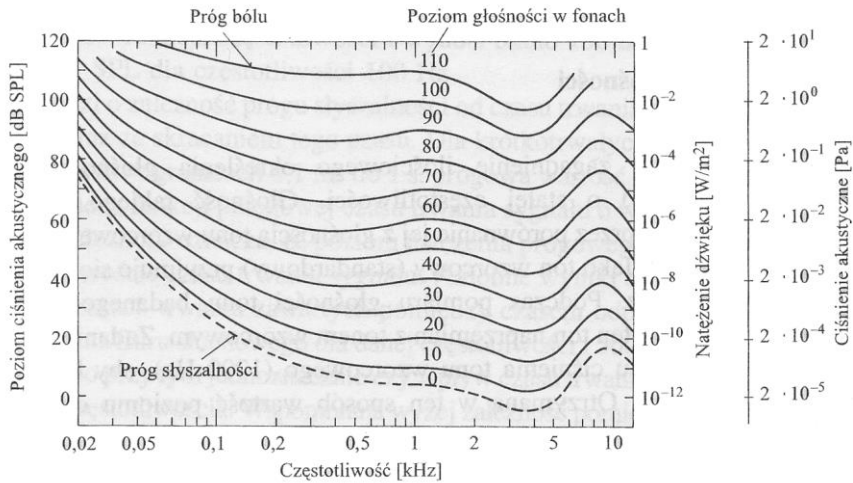
- Szum to dźwięk, którego widmo akustyczne jest ciągłe i ma przypadkowy charakter.
- Szum biały – wszystkie składowe mają taką samą amplitudę;
- Szumy kolorowy (np. brązowy, różowy itp.) – niektóre zakresy częstotliwości mają znacznie większą amplitudę niż pozostałe.



Natężenie i głośność dźwięku

- Natężenie dźwięku to parametr obiektywny, mierzalny, określa moc akustyczną fali przechodzącą przez powierzchnię 1 m^2 , prostopadłą do fali akustycznej. Jednostką natężenia dźwięku jest W/ m^2 (Wat na metr kwadratowy);
- Głośność dźwięku jest subiektywną cechą wrażenia słuchowego, zależną od wielu czynników takich jak: częstotliwość, struktura widmowa, poziom ciśnienia akustycznego, czas trwania bodźca, czułość słuchu, rodzaj i charakter ubytku słuchu.

Krzywe jednakowej głośności



Głośność dźwięku

- Głośność dźwięku zależy od szerokości pasma, a nie zależy od ilości składowych. Im większa szerokość pasma tym wrażenie głośności większe.
- Dla dźwięków krótszych niż 200 ms wrażenie głośności rośnie wraz ze wzrostem czasu trwania dźwięku
- Dla dźwięków o czasie trwania dłuższym niż kilka sekund wrażenie głośności maleje wraz ze wzrostem czasu trwania dźwięku (efekt adaptacji słuchowej).

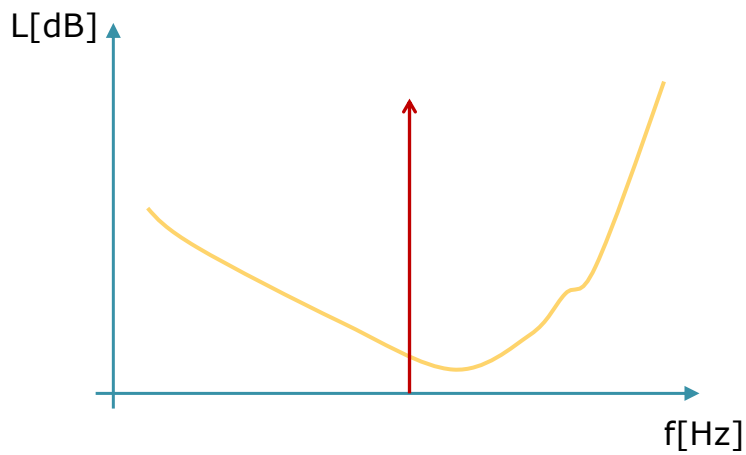
Maskowanie dźwięków

- Maskowanie – proces, w którym próg słyszalności jednego dźwięku (sygnału) podnosi się na skutek obecności innego dźwięku (maskera).
- Maskowanie całkowite występuje gdy głośność jednego dźwięku zanika w obecności innego dźwięku.
- Maskowanie częściowe występuje gdy głośność dźwięku obniża się w obecności innego dźwięku.

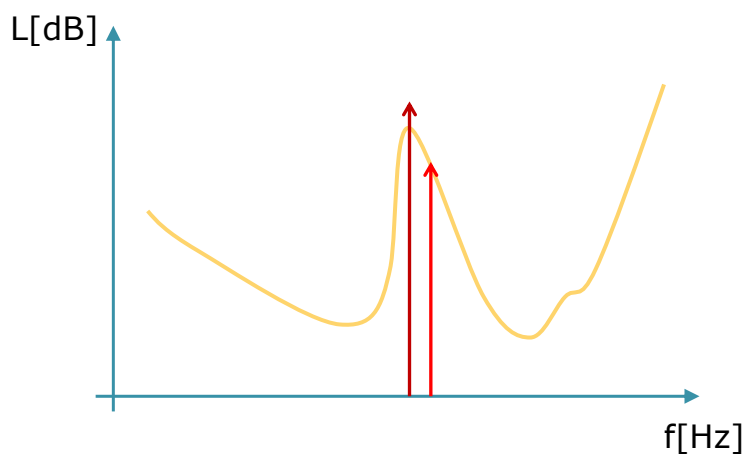
Maskowanie dźwięków

- Maskowanie równoczesne – jeśli sygnał i masker występują w tym samym czasie.
- Maskowanie nierównoczesne – gdy sygnał i masker są względem siebie przesunięte w czasie.
 - premaskowanie (maskowanie wstępne) - wynikające z pewnej bezwładności w zadziałaniu mechanizmu słuchu oraz z faktu, że dźwięki głośniejsze są przetwarzane szybciej niż dźwięki ciche
 - postmaskowanie - spowodowane stosunkowo długim czasem relaksacji neuronów (czas maskowania zależy od natężenia tonu maskującego oraz czasu jego trwania)

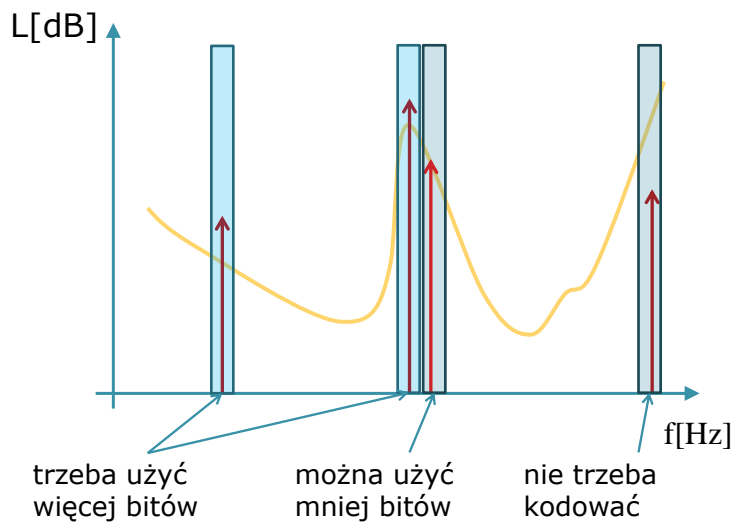
Ilustracja maskowania



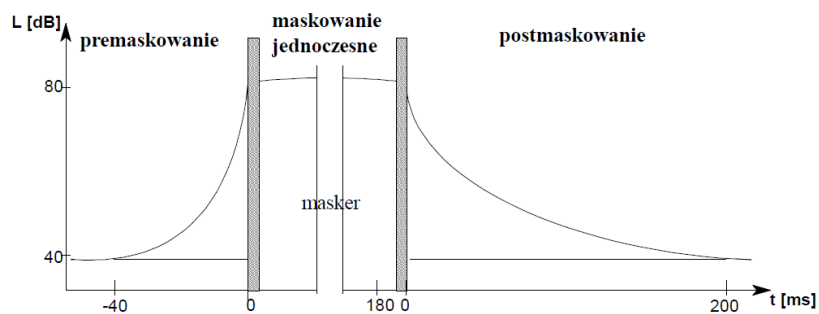
Ilustracja maskowania



Ilustracja maskowania



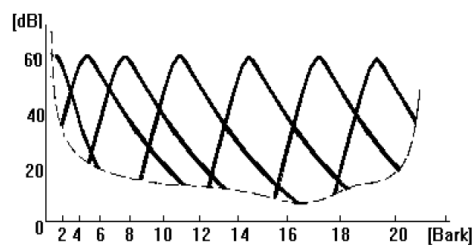
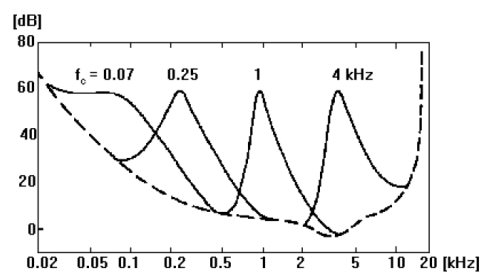
Maskowanie dźwięków



Maskowanie dźwięków

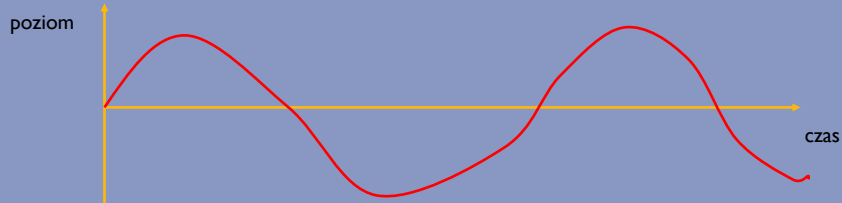
- Pasma krytyczne –maksymalna szerokość pasma szumu, który powoduje podwyższanie progu słyszenia (maskowanie) tonu o częstotliwości równej częstotliwości środkowej pasma krytycznego.
- Maskowanie jest zależne od częstotliwości sygnału i maskera. Jeśli częstotliwość maskera i sygnału należą do tego samego pasma krytycznego to występuje efekt maskowania.

Maskowanie dźwięków

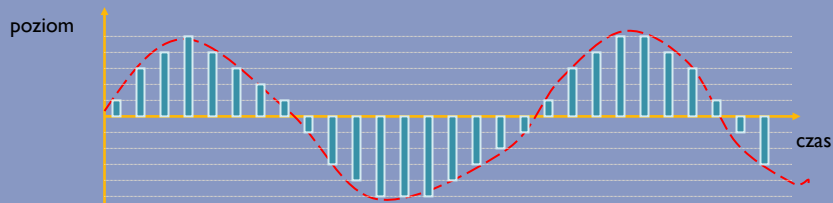


Sygnal foniczny

- analogowy

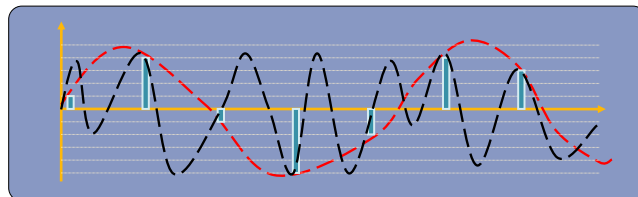


- cyfrowy



Cyfrowy sygnał foniczny

- składa się z tzw. próbek pobieranych z określoną częstotliwością (szybkością) próbkowania
 - im większa częstotliwość próbkowania, tym sygnał cyfrowy lepiej opisuje sygnał analogowy;
 - częstotliwość próbkowania nie może być zbyt mała – bo nie będzie wiadomo jak naprawdę wygląda sygnał – **częstotliwość próbkowania musi być dwa razy większa od maksymalnej częstotliwości sygnału.**



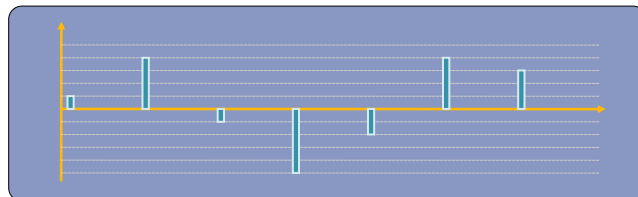
Parametry dźwięku

- częstotliwości próbkowania (w Hz)
 - **8000** – jakość telefoniczna – można zakodować dźwięk do ok. 3,5-4 kHz
 - **22050** – można zakodować dźwięk do ok. 10kHz
 - **32000** – produkcja radiowa, magnetofon DAT, NICAM
 - **44100 – CD-Audio** – można zakodować dźwięk do ok. 20 kHz
 - **48000 – częstotliwość studyjna, DVD, Blu-ray**
 - **96000** – produkcje wysokiej jakości, DVD, Blu-ray
 - **192000** – j.w. – można zakodować dźwięk do ok. 90kHz

Cyfrowy sygnał foniczny

- rozdzielczość bitowa – liczba bitów służąca do opisanie pojedynczej wartości pojedynczej próbki (słupka)
 - im więcej bitów służy do opisu danego dźwięku, tym bardziej dokładnie można opisać dany dźwięk
 - zakres dynamiki konwertera PCM wyraża się wzorem

$$S/N \cong 6n + 1,8 \text{ [dB]}$$



Parametry dźwięku

- rozdzielczości bitowe:
 - 8 bitów - czyli 2^8 możliwych wartości – 256
 - dźwięk zaszumiony, marnej jakości
 - 16 bitów - czyli 2^{16} możliwych wartości – 65.536
 - najbardziej typowa rozdzielczość
 - odstęp sygnał szum rzędu 96dB
 - 24 bity - czyli 2^{24} możliwych wartości – 16.777.216
 - zyskuje na popularności, używana w studiach
 - odstęp sygnał szum rzędu 144dB
 - 32 bity - czyli 2^{32} możliwych wartości – 4.294.967.296
 - używana podczas wewnętrznego przetwarzania i miksowania plików (zapobieganie obcinaniu próbek)

Barwa dźwięku

- Barwa dźwięku jest cechą wrażenia słuchowego zależną od składu widma dźwięku.
- Barwa dźwięku zależy od ilości składowych, ich amplitud i względnych stosunków natężenia.
- Barwa dźwięku pozwala rozróżniać dwa dźwięki o takiej samej wysokości;

Wysokość dźwięku

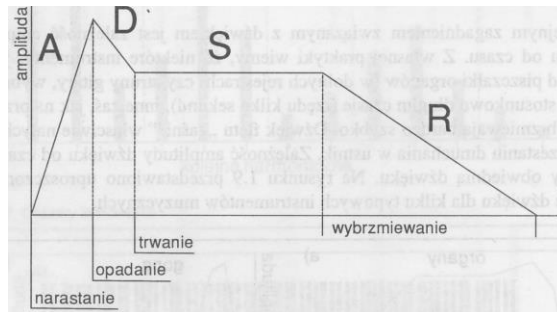
- Wysokość dźwięku to subiektywne wrażenie słuchowe pozwalające szeregować dźwięki od niskich do wysokich.
- Wysokość dźwięku zależy głównie od częstotliwości (ton prosty) lub częstotliwości podstawowej (dźwięki złożone).
- Na ocenę wysokości dźwięku oprócz częstotliwości wpływ ma również: poziom głośności, obwiednia amplitudy, skład widmowy, czas trwania oraz obecność innych dźwięków.

Określanie kierunku dźwięku

- Określenie kierunku dochodzenia dźwięku jest znacznie dokładniejsze w przypadku słyszenia dwuosobnego.
- Człowiek lokalizuje kierunek dźwięku na podstawie analizy różnicy faz dźwięku (czasu dotarcia do ucha lewego i ucha prawego) oraz różnicy natężeń dźwięku (poziom głośności w uchu lewym i w uchu prawym).
- Im wyższa częstotliwość dźwięku tym łatwiej lokalizować kierunek dźwięku.

Obwiednia dźwięku

- Obwiednia dźwięku opisuje zależność amplitudy dźwięku od czasu;
- Zazwyczaj obwiednie dźwięków, zwłaszcza muzycznych, opisuje się za pomocą modelu ADSR (ang. *Attack, Decay, Sustain, Release*);



Obwiednia dźwięku

- Czasy trwania poszczególnych faz są różne i charakterystyczne dla instrumentów muzycznych;
- Nie zawsze muszą występować wszystkie fazy;

