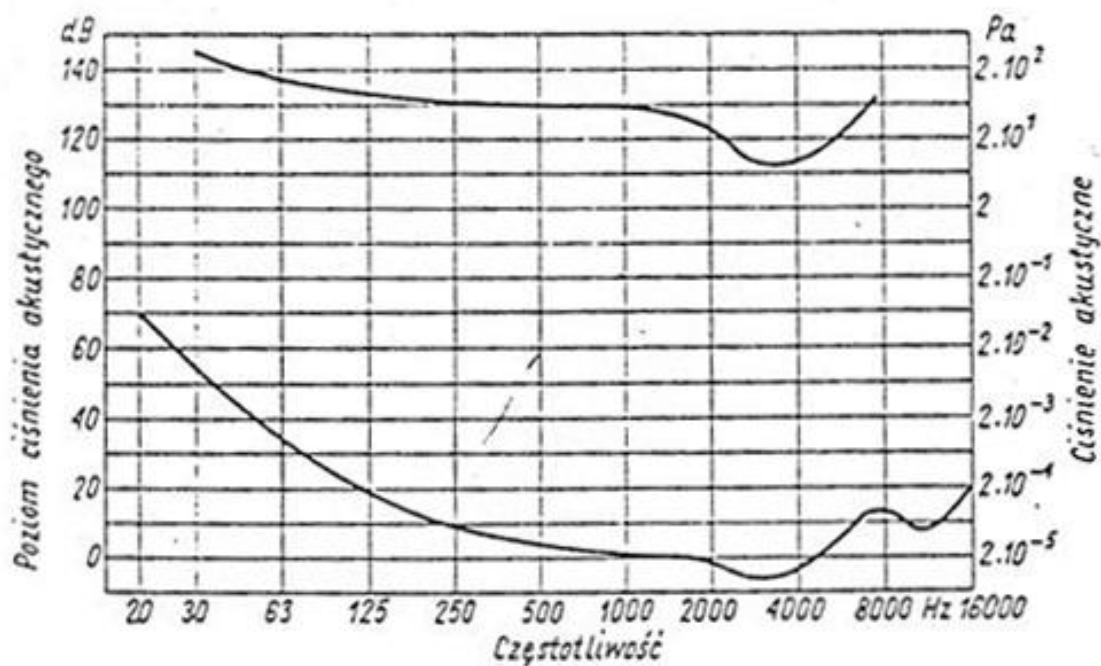
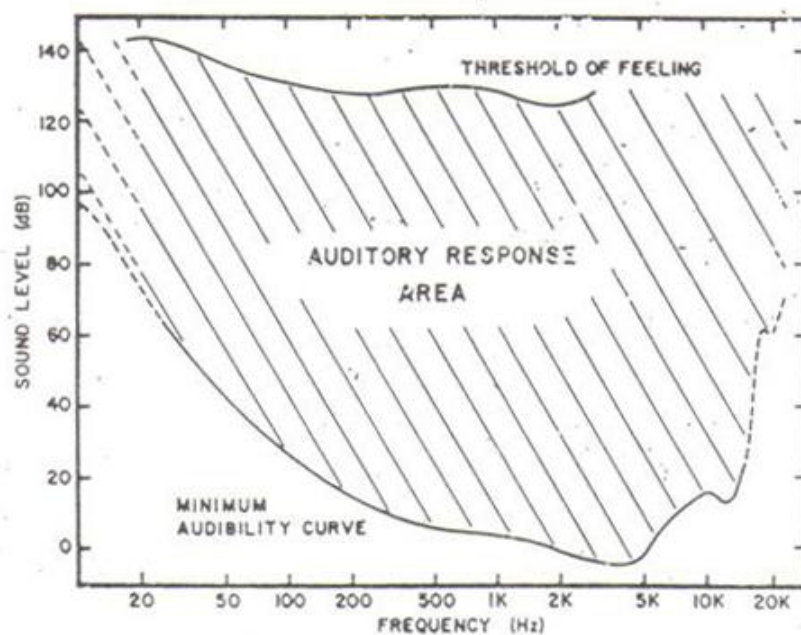


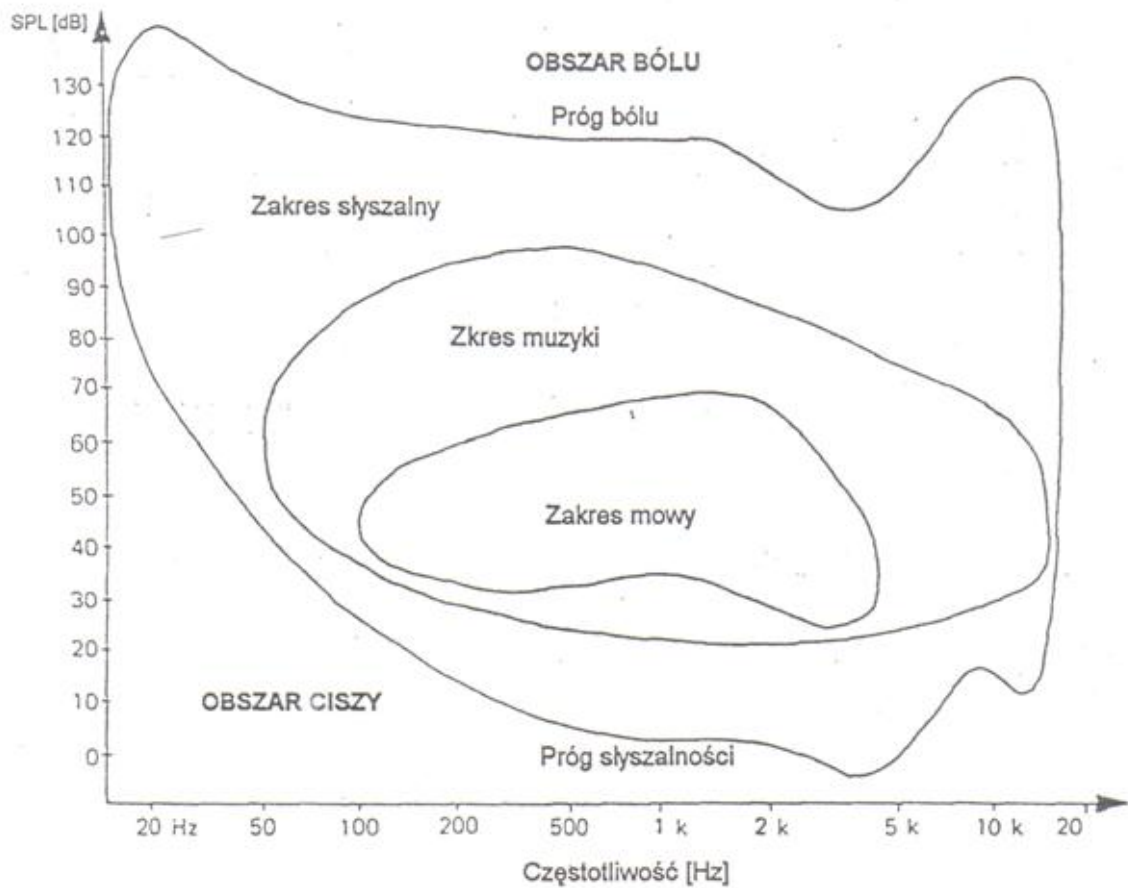
**Przygotowała: prof. Bożena Kostek**



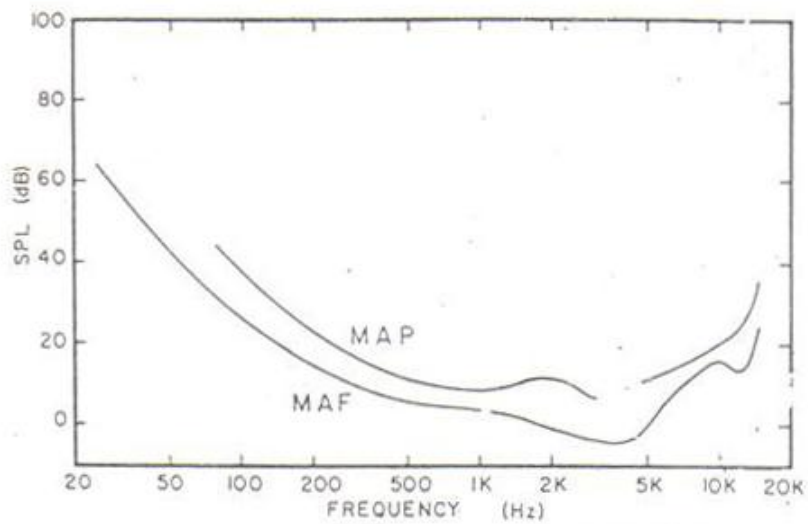
Rys.8 Powierzchnia słyszalności



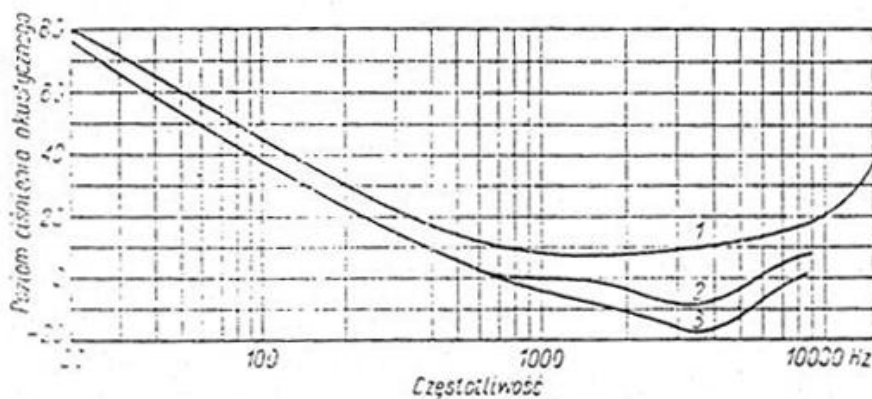
Rys.4 The auditory response area. The minimum audibility curve is based on data of Robinson and Dadson (1956); the low and high frequency portions of the curve (*broken line*) have been extended based on data of Yeowart and Evans (1974) and Corso (1965); respectively. The curve representing the threshold of feeling is based on data of Wegel (1932). Along the frequency axis, the auditory response area is shown to extend from 20 to 20,000 Hz by the *solid cross-hatched lines* although responses to sounds outside these limits are demonstrable (*broken cross-hatched areas*).



Rys.5 Zakresy słyszalności.



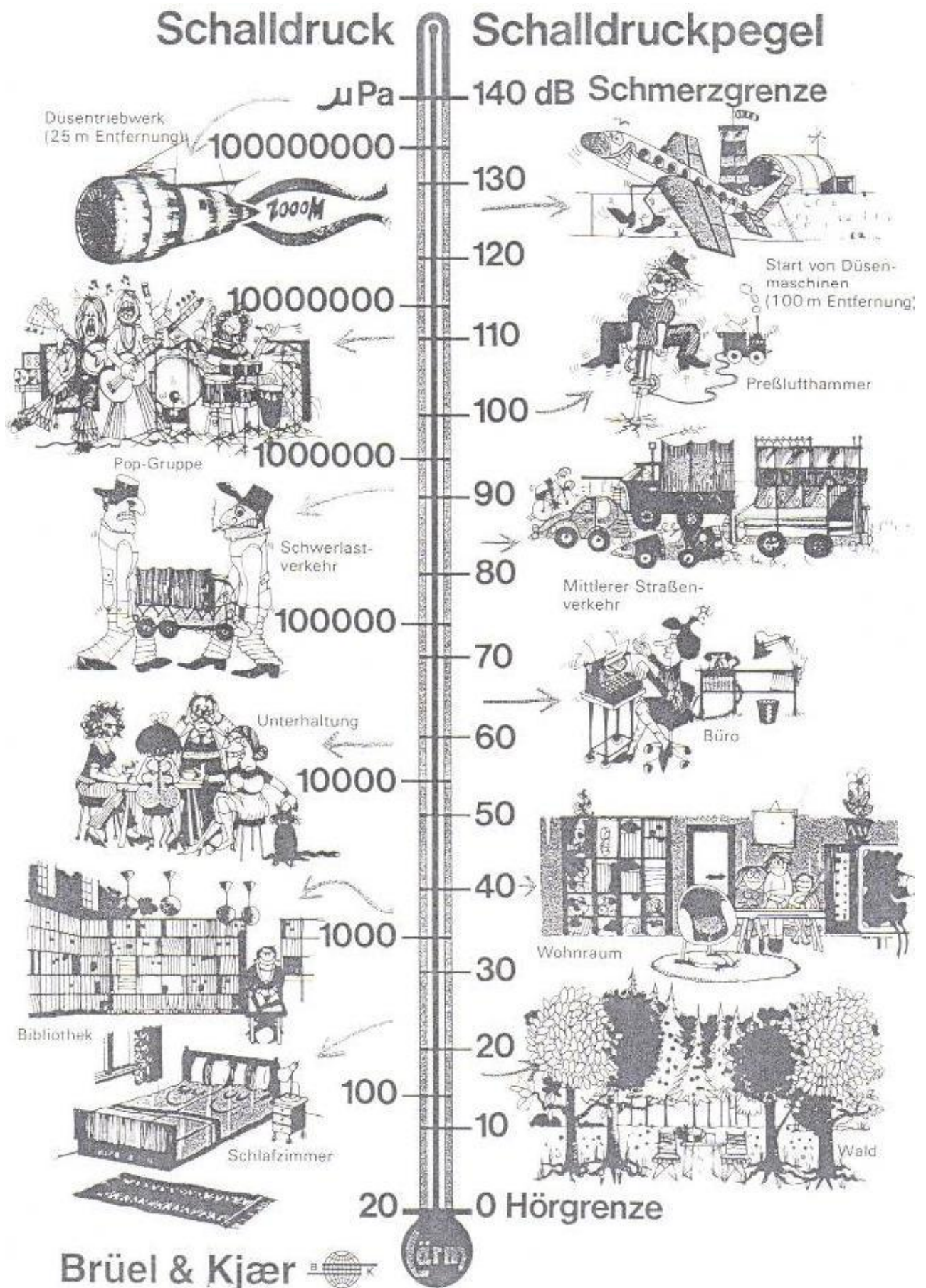
Rys.6 Minimum audible pressure (MAP) and minimum audible field (MAF). The MAP curve is based on Dadson and King's (1952) data wherein sound pressure level was measured under the earphones at the entrance of the ear canal. The MAF curve is the same as that shown in Figure 7.1 based on data from Robinson and Dadson (1956), and represents measurement of binaural free-field (0° azimuth) sensitivity.



Rys.9 Dolna granica słyszalności

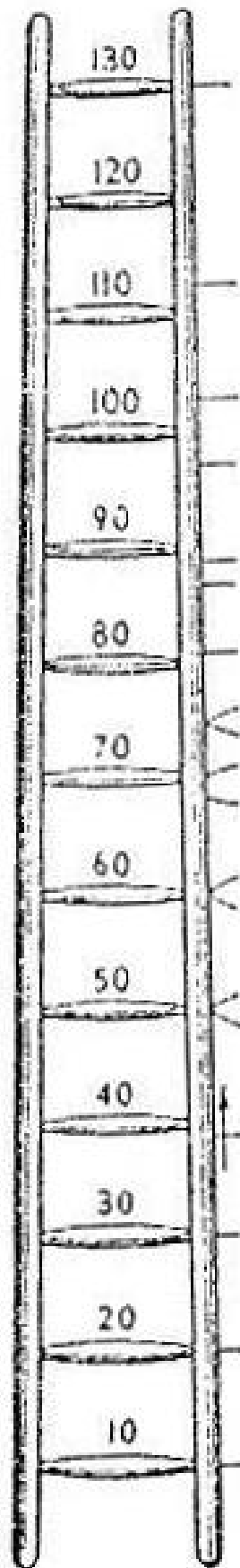
1 — dla słuchania jednousznego; dźwięk jest wytworzony przez słuchawkę i ciśnienie akustyczne jest mierzone u wejścia do przewodu usznego; 2 — dla słuchania dwuusznego; dźwięk jest wytworzony przez źródło punktowe przed słuchaczem i ciśnienie jest mierzone w polu swobodnym; 3 — dla słuchania dwuusznego; dźwięk jest wytworzony przez kilka małych głośników, umieszczonych dowolnie w płaszczyźnie poziomej dokoła słuchacza i ciśnienie akustyczne jest mierzone w polu swobodnym przed umieszczeniem w nim słuchacza

Ze względu na dużą rozpiętość mierzonych wartości ciśnienia (zakres ciśnień akustycznych obejmuje blisko siedem rzędów wartości: od  $2 \times 10^{-5}$  Pa do ponad 10 Pa) wygodniej jest mierzone ciśnienie akustyczne podawać nie w wartościach bezwzględnych, lecz porównywać je do określonej wartości przyjętej za odniesienie i podawać stosunek tych wartości. Jeśli wynik pomiaru jest bardzo duży na ogół podaje się logarytm tego stosunku. W akustyce stosuje się logarytm o podstawie 10 i wynik podaje się w belach lub w ich dziesiątej części zwanej decybelami [dB].



**Brüel & Kjær**  

Rys.7 Przykładowe poziomy dźwięku.



Granica wytrzymałości ucha.

Płyta stalowa, w którą uderza młotem 4 mężczyzn, 2 st. ang.

Sygnal samochodowy głośny, 23 st. ang.

Nitownica, 35 st. ang.

Tramwaj elektryczny na pomoście, 20 st. ang.

Lew ryczący, 18 st. ang.

Bardzo wielkie natężenie ruchu ulicznego, N. York

Samochód ciężar., 30 st. ang.

Człowiek krzyczący, 4 st. ang.

Huk grzmotu.

Znaczne natężenie ruchu ulicznego, Nowy York.

Przecięta ulica handlowa, Chicago.

Znaczne natężenie ruchu Londyn.

Spokojna jazda samochodem, 30 st. ang.

Mowa potoczna, 12 st. ang.

Próg wymiernej reakcji strachu.

Centrum N. Yorku w najspokojniejszej porze nocy.

Spokojna ulica na przedmieściu Londynu.

Szept, 4 st. ang.

Szmer liści przy podmuchach wiatru.

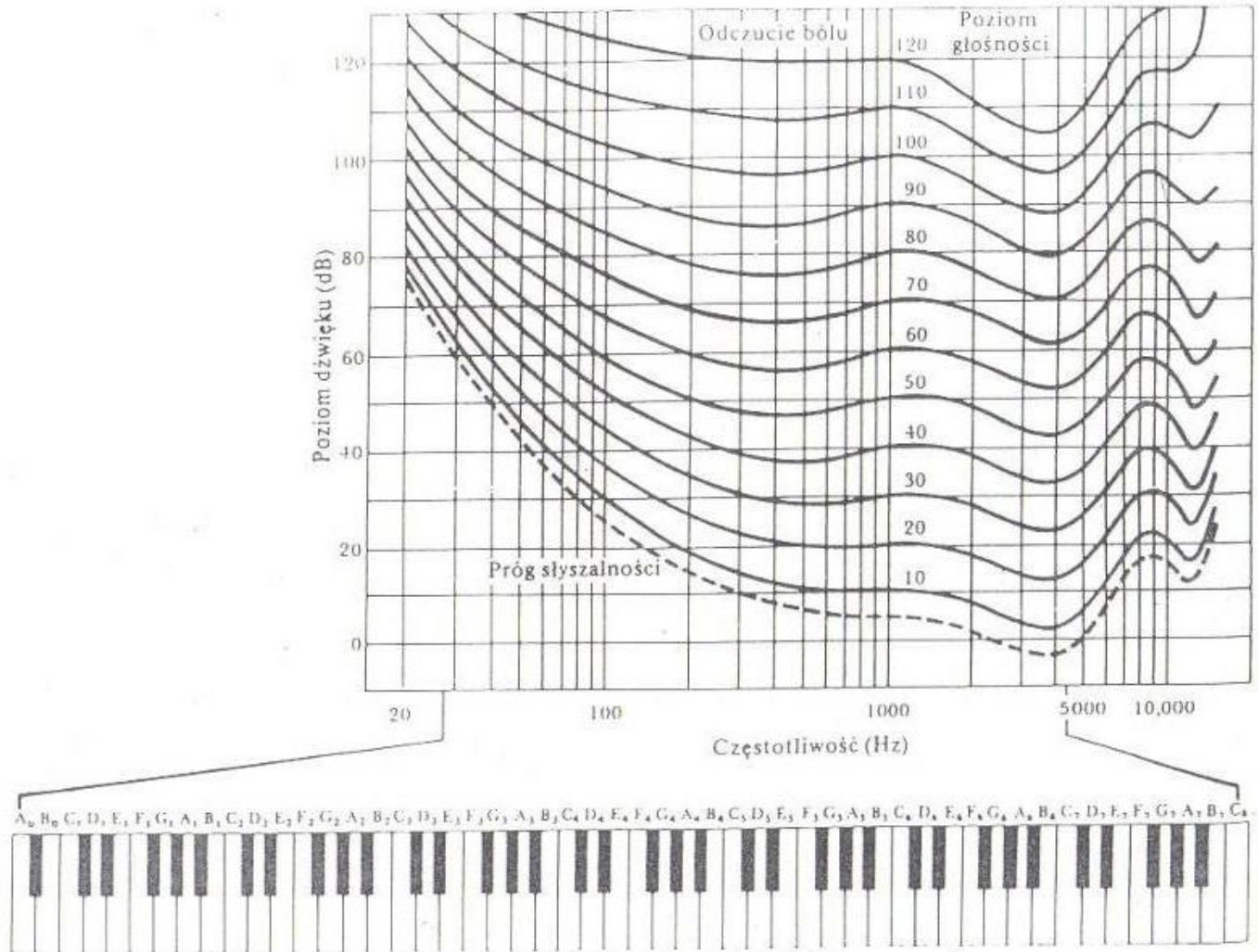
Najslabszy dźwięk możliwy do usłyszenia.

Rys. 3 Drabinka hałasu. Każdy szczebel odpowiada dziesięciokrotnej zmianie natężenia.

$$\Delta W = K \Delta B / B$$

gdzie:  $\Delta W$  - przyrost wrażenia,  $\Delta B$  - przyrost bodźca,  $B$  - bodziec istniejący,  $K$  – stała proporcjonalności

$$W = K \ln B$$



Rys.9 Zależność poziomu dźwięku od częstotliwości

**Poziom głośności** - wielkość będąca porównawczą miarą głośności dźwięku w odniesieniu do głośności dźwięku wzorcowego (wyrażona w fonach). Zestaw krzywych jednakowego poziomu głośności nazywany jest zestawem **krzywych izofonicznych** lub zestawem **izofon**.

**Poziom głośności** dowolnego dźwięku jest wyrażony w fonach, których liczba jest równa poziomowi ciśnienia akustycznego w decybelach, wytwarzanego w punkcie obserwacji przez falę akustyczną biegnącą o częstotliwości 1000 Hz, wywołującą w ściśle określonych warunkach wrażenie takiego samego poziomu głośności jak dźwięk badany. Poziom dźwięku nie jest miarą addytywną ani multiplikatywną.

Pojęcie jednakowego poziomu głośności wynika z porównania poziomu ciśnienia akustycznego danego dźwięku sinusoidalnego z poziomem ciśnienia akustycznego dźwięku sinusoidalnego o częstotliwości 1000Hz.

Poziom głośności wynosi  $n$ -fonów, jeżeli słuchacz o prawidłowym słuchu oceni go jako jednakowo głośny z tonem odniesienia o częstotliwości 1000 Hz, którego poziom wynosi  $n$ -decybeli.

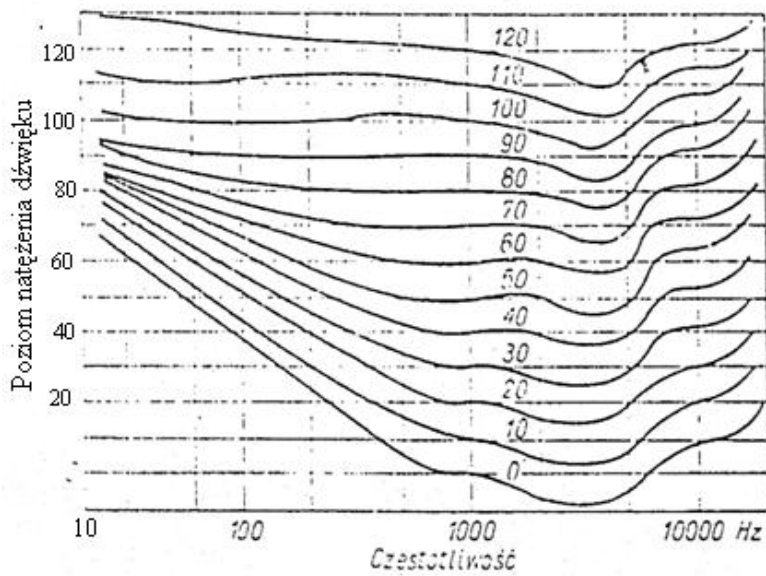
**Prawo Webera-Fechnera** (uzasadnienie stosowania logarytmicznej skali dB)

$$\Delta W = K \Delta B / B$$

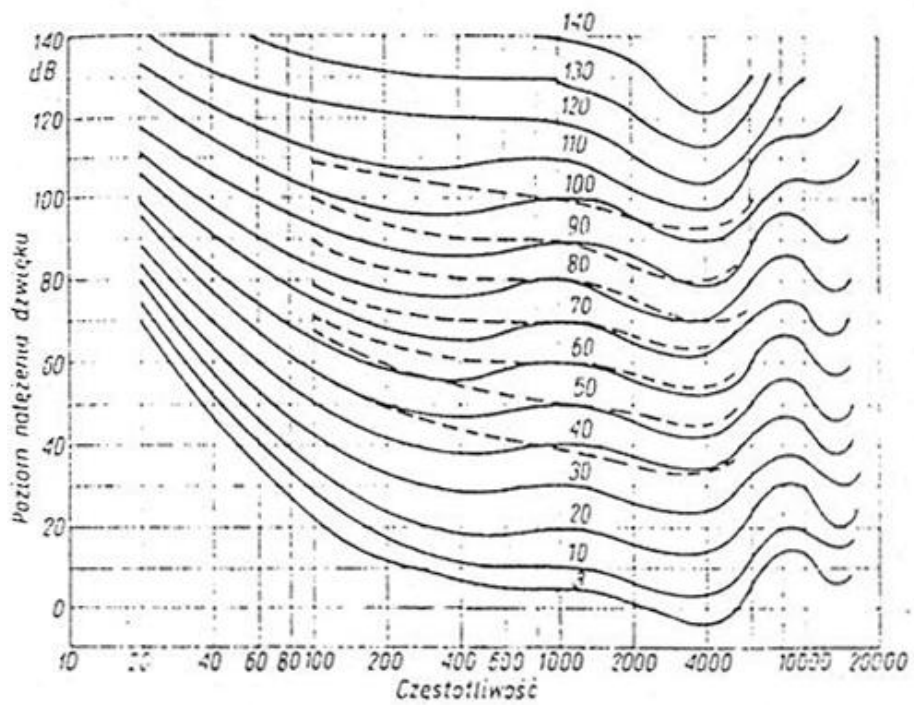
gdzie:  $\Delta W$  - przyrost wrażenia,  $\Delta B$  - przyrost bodźca,  $B$  - bodziec istniejący,  $K$  – stała proporcjonalności

$$W = K \ln B$$

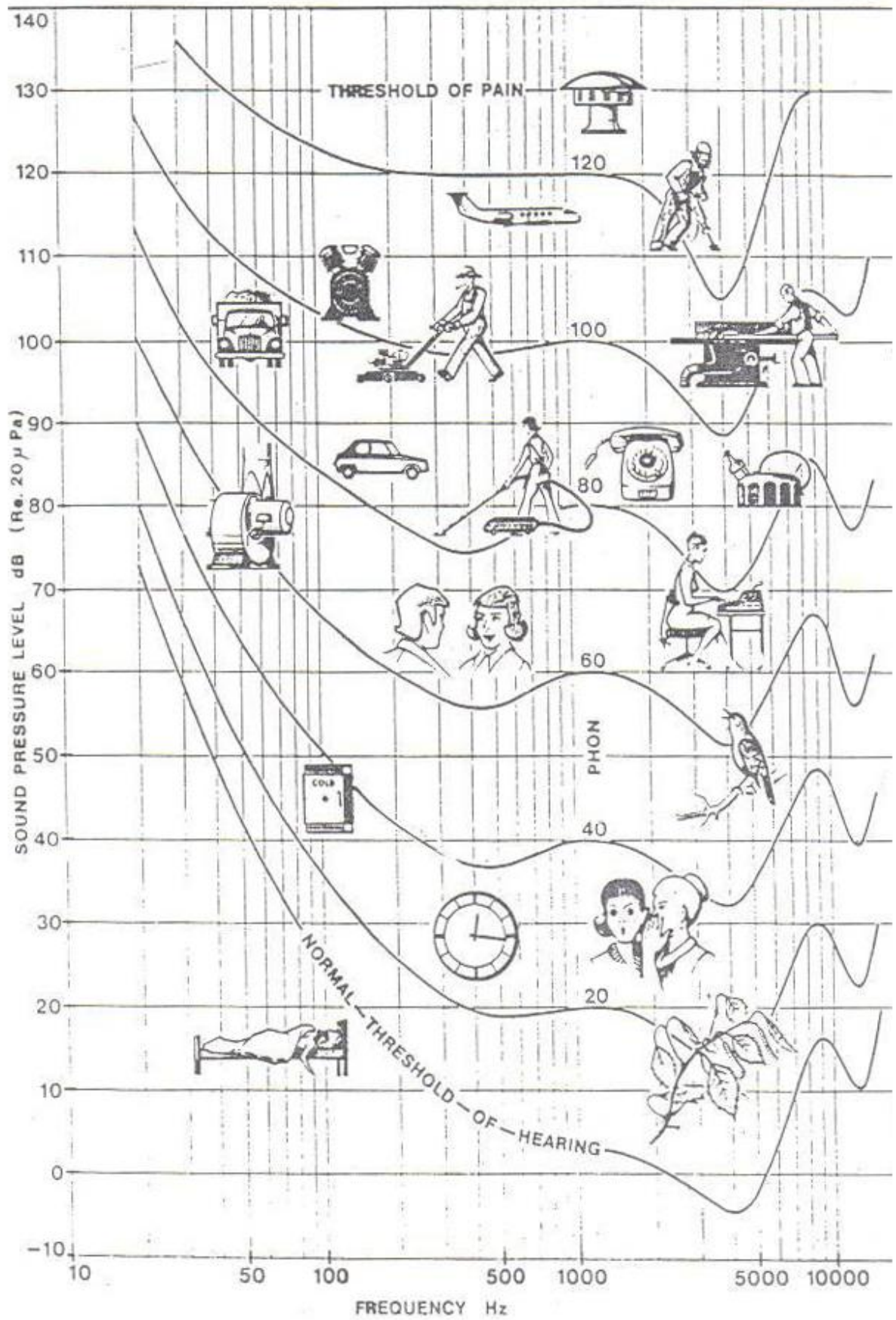




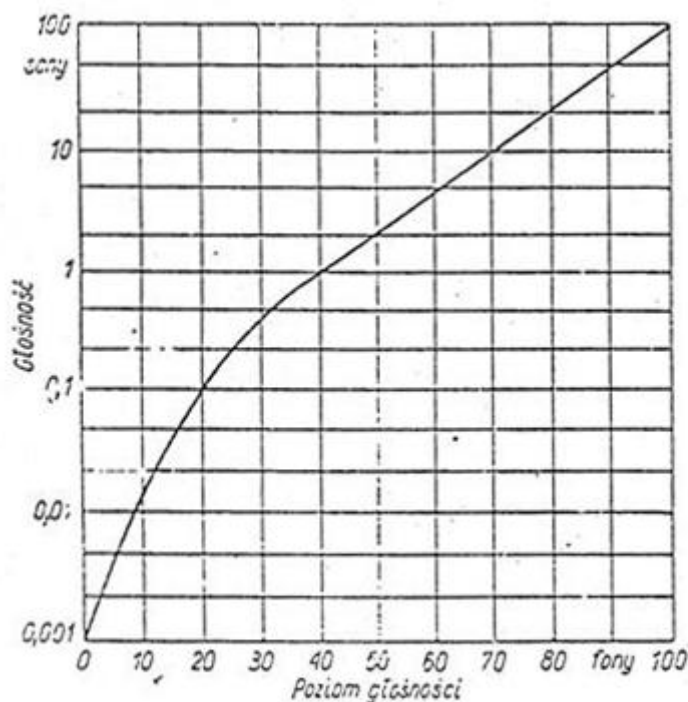
Rys.1 Krzywe jednakowej głośności (izofony) ucha normalnego wg Fletchera i Munsona. Liczby przy krzywych oznaczają poziom głośności w fonach. Poziom natężenia dźwięku 0 dB =  $10^{-12}$  W/m<sup>2</sup> [7]



Rys.2 Krzywe jednakowej głośności ucha normalnego dla tonów wg Robinsona i Dadson (linie ciągłe) i pasm szumów o szerokości 250 melów wg Pollacka (linie przerywane) [19], [17]



Rys.2 Przykładowe poziomy dźwięku



Rys.4 Zależność między głośnością w sonach a poziomem głośności w fonach

**Głośność** jest to **cecha wrażeniowa dźwięku**, pozwalająca na uszeregowanie dźwięków w skali od najcichszych do najgłośniejszych. Skala głośności jest taka sama dla wszystkich częstotliwości.

Dla skali głośności przyjęto, że poziom głośności 40 fonów tonu prostego o częstotliwości = 1000 Hz ma głośność wynoszącą 1 son. Zależność między głośnością  $S$  (w sonach) a poziomem głośności  $L_s$  (w fonach) dla poziomów głośności równych lub większych niż 40 fonów.

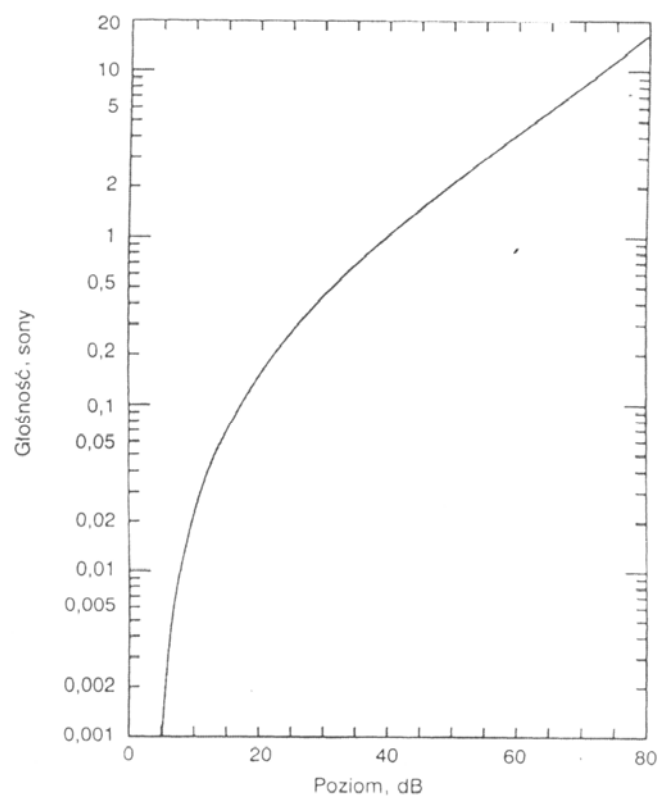
$$S = 2^{\frac{P-40}{10}}$$

$$\log_{10} S = 0,03(P - 40)$$

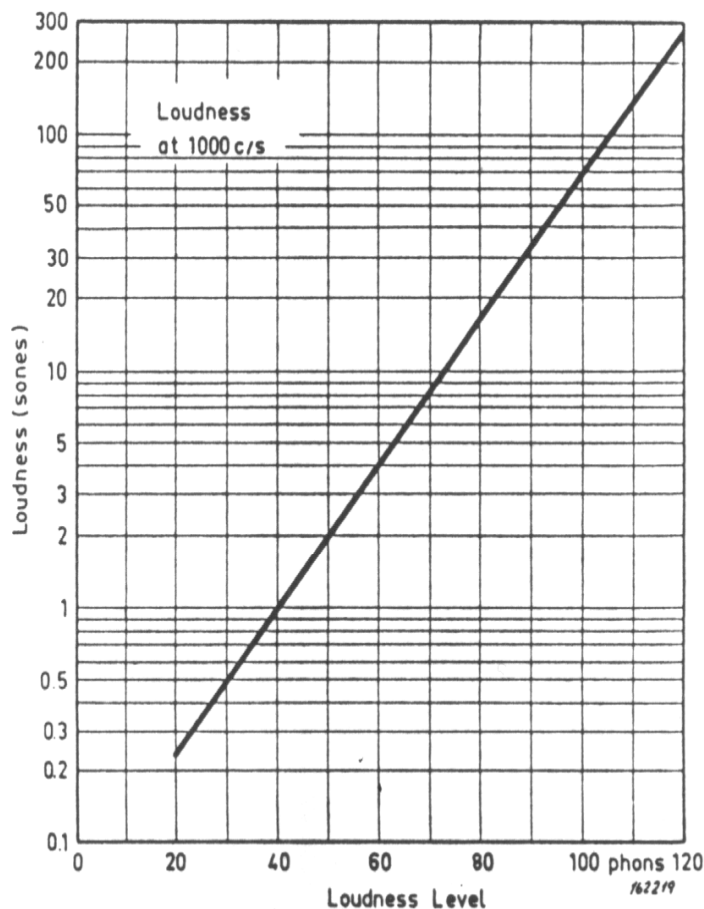
$$P = 40 + \log_{10} S / \log_{10} 2$$

# *S*- poziom głośności w sonach *P* – poziom głośności w fonach

Pomiary Robinsona i innych badaczy doprowadziły do wniosku, że w zakresie od 40 do 120 fonów dźwięk jest odbierany jako dwa razy głośniejszy, gdy jego poziom głośności jest o 10 fonów większy.



Skala głośności w sonach dla tonu o częstotliwości 1000 Hz



The relationship between the loudness in sones and the loudness level in phon. According to the I.S.O. Recommendation ISO/R 131-1959 the relationship may be written as  $S = 2^{(P-40)/10}$  for loudness levels between 20 phon and 120 phon.

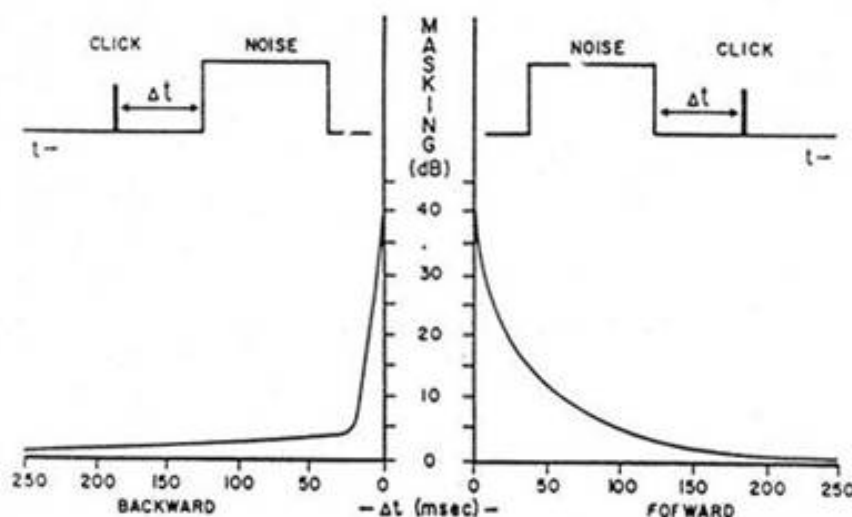
Podwojenie liczby sonów prowadzi do powiększenia o 10 liczby fonów, co daje, że 40 fonów odpowiada 1 sonowi, 50 fonów 2 sonom itd., zaś 120 fonów 256 sonom.

Maskowanie jest to zjawisko (proces) w którym próg słyszalności dźwięku podwyższa się wskutek obecności dźwięku zagłuszającego. Z uwagi na wzajemne relacje słuchanych sygnałów, ich wzajemny stosunek, możemy mówić o zagłuszaniu **całkowitym** (podwyższenie progu) lub **częściowym** (obniżenie głośności). Ze względu na wzajemne relacje czasowe można mówić o:

- o maskowaniu **jednoczesnym** (równoczesnym), gdy zarówno sygnał maskowany, jak i maskujący są długie w czasie,
- o maskowaniu **nierównoczesnym**, wówczas zarówno sygnał, jak i masker są wyraźnie określone w czasie.

Maskowanie nierównoczesne:

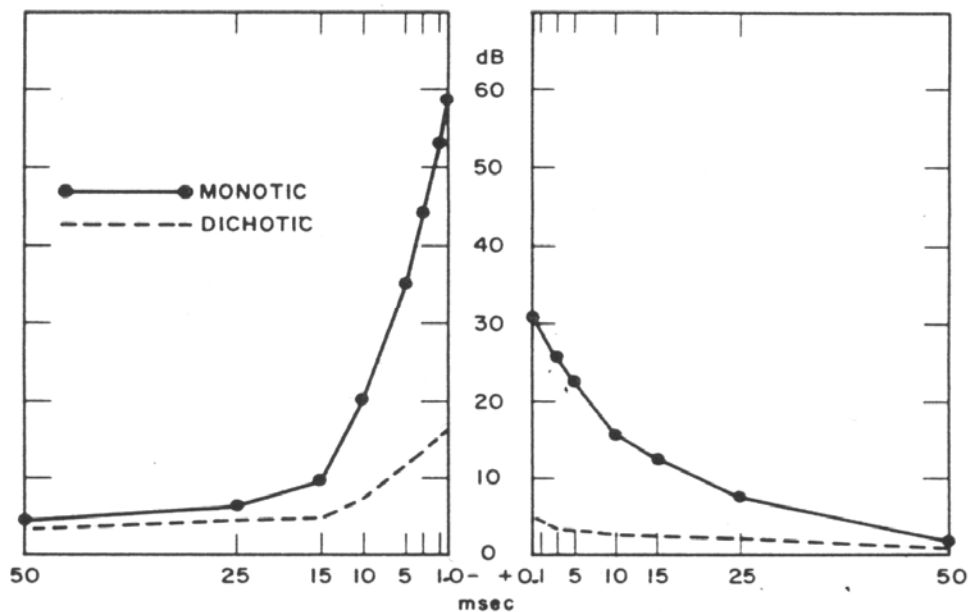
- **Maskowanie przedbodźcowe (wsteczne)** –sygnał (w czasie) występuje przed maskerem
- **Maskowanie jednoczesne (ale bodźcowe)** - wynik maskowania zależy od czasu maskowania (czasu trwania sygnału i maskera)
- **Maskowanie resztkowe, następcze** - masker występuje (w czasie) przed sygnałem



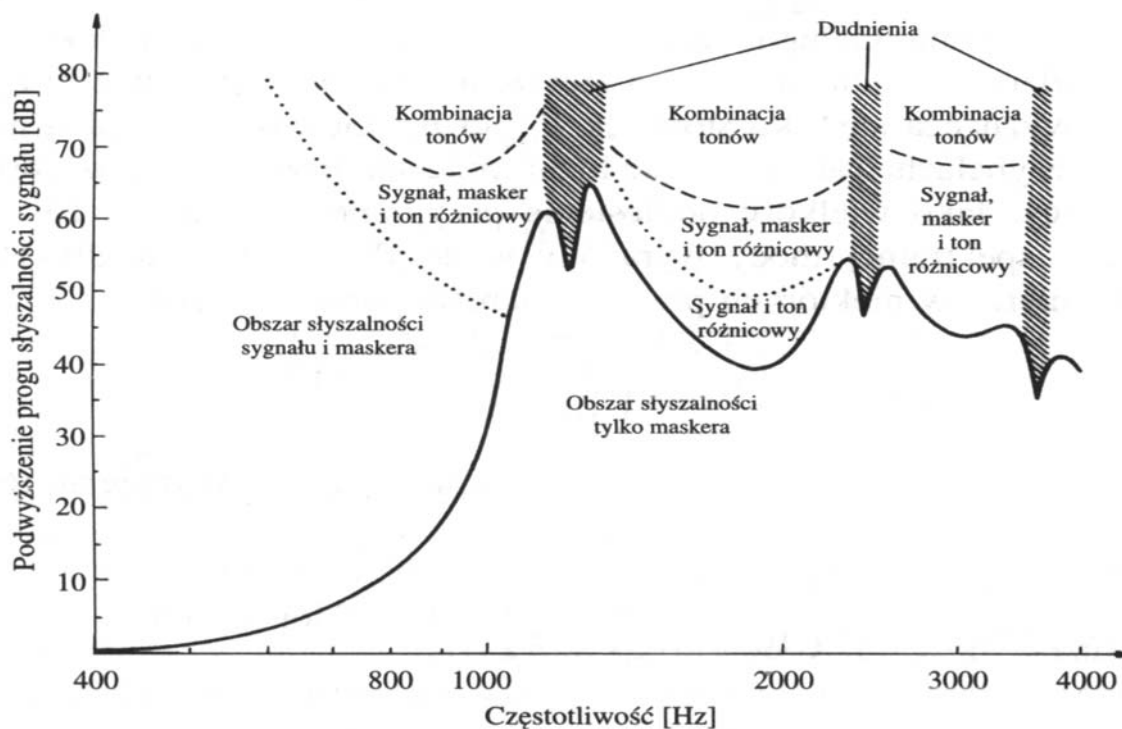
Rys.4 Forward and backward masking effects of a noise masker (white noise, 500 msec duration) on an acoustic click (approximately 0.4 msec duration). The masker and probe stimuli are separated by the time interval  $\Delta t$ . (Based on data of Wilson and Carhart, 1971.)

Z punktu widzenia fizjologicznego można mówić o maskowaniu:

- obwodowym (peryferyjnym; występuje wówczas, gdy sygnał maskujący i maskowany dochodzą do tego samego ucha),
- centralnym (sygnały maskujący i maskowany są podawane oddzielnie do każdego ucha i spotykają się w centralnym układzie nerwowym, a efekty maskowania są od 2. do 3. razy mniejsze niż przy maskowaniu obwodowym).

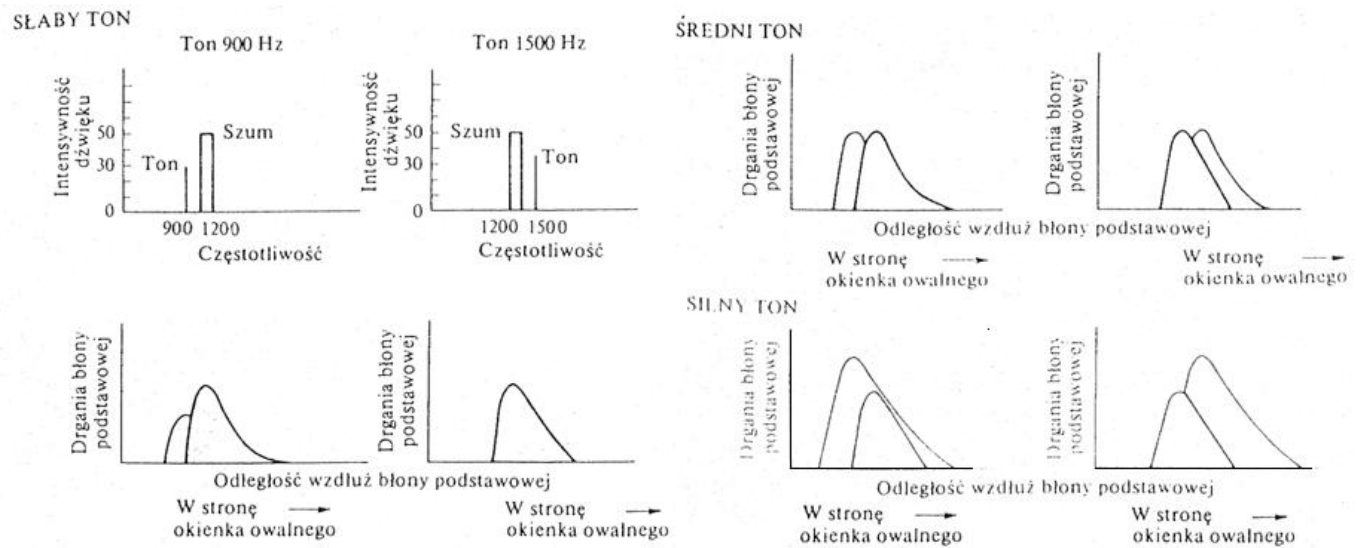


Przykład **maskowania obwodowego** (zagłuszanie tonów prostych przez ton sinusoidalny o zadanym poziomie i częstotliwości).



Dla częstotliwości tonu zagłuszanego równej częstotliwości tonu zagłuszającego i częstotliwości bardzo bliskich zauważa się zjawisko dudnień i nieznaczne zmniejszenie efektu zagłuszania. To samo zauważa się dla częstotliwości bliskich częstotliwościom harmonicznym tonu zagłuszającego, co jest efektem pojawiania się tonów subiektywnych które mogą powodować powstawanie tonów różnicowych.

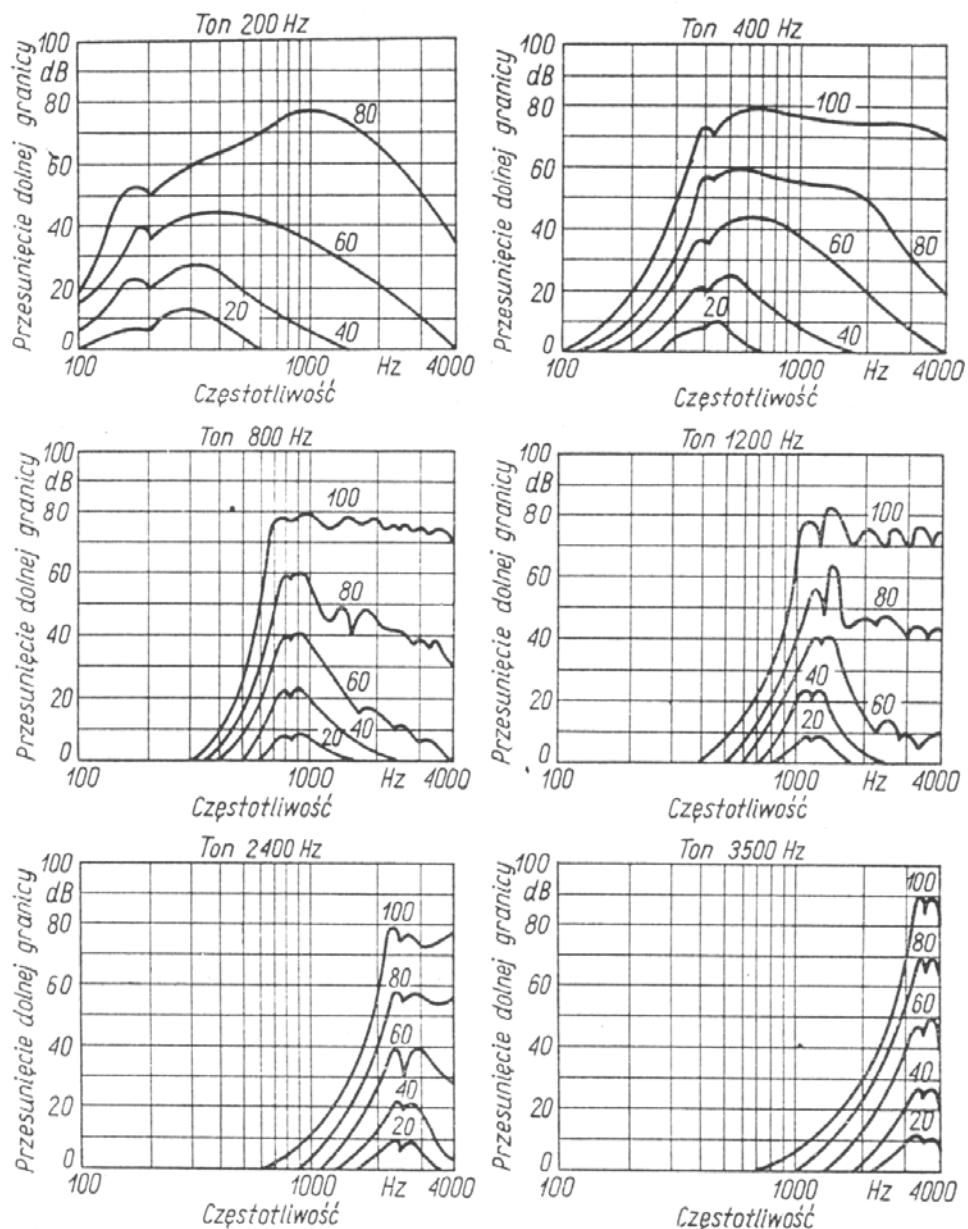
## Przykłady zjawiska maskowania



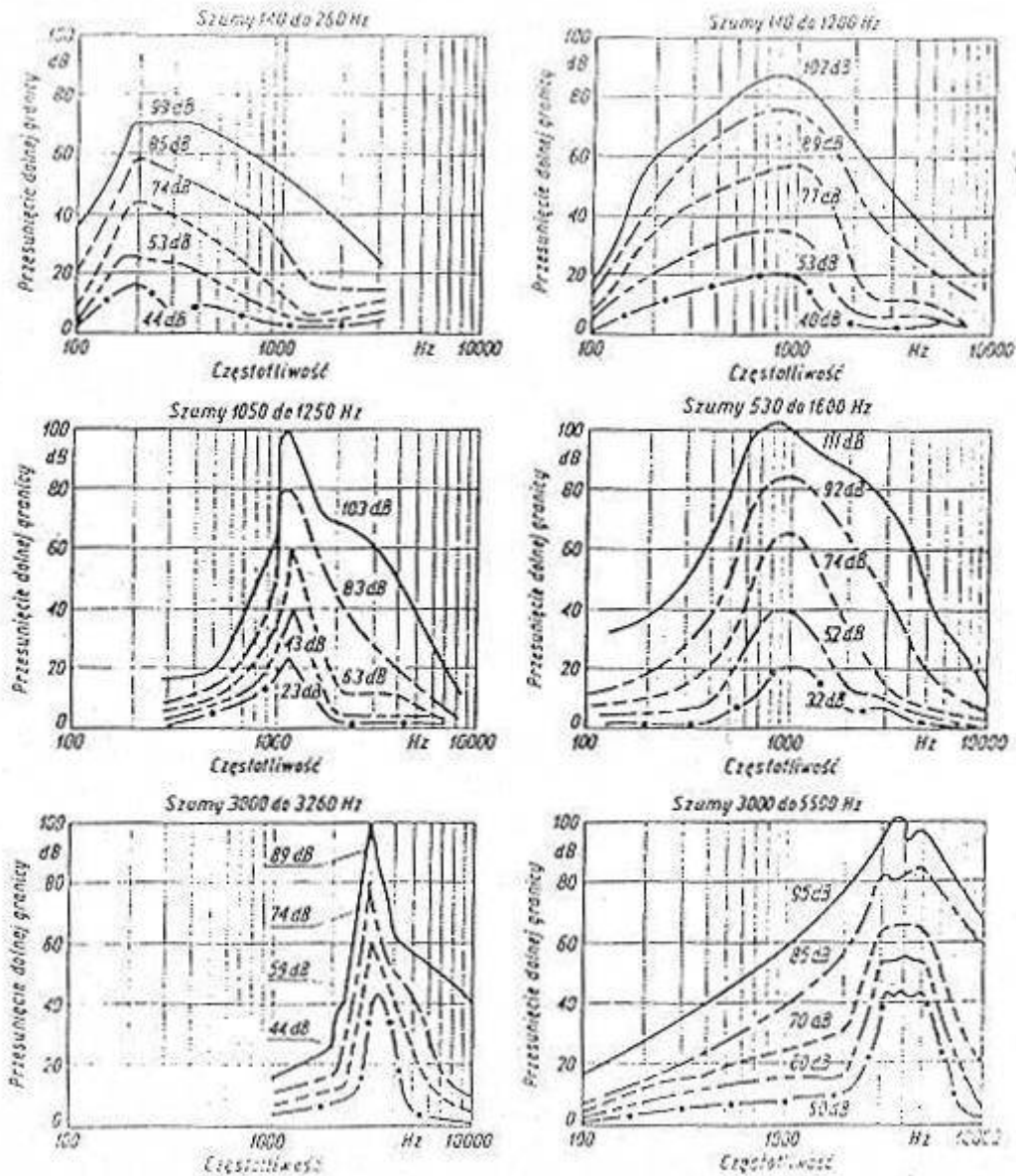
1. Największy efekt zagłuszania bez względu na wartość poziomu tonu zagłuszającego obserwuje się dla częstotliwości zbliżonych do częstotliwości tonu zagłuszającego. Przy czym łatwiej ulegają zagłuszaniu tony o częstotliwościach większych od tonu zagłuszającego, niż tony o częstotliwościach mniejszych
2. Dla częstotliwości tonu zagłuszanego równej częstotliwości tonu zagłuszającego zauważa się zjawisko dudnień. Obserwacja efektów zagłuszania jest wtedy utrudniona, podobnie jak i dla tonów zagłuszanych o częstotliwościach równych harmonicznym tonu zagłuszającego.
3. Dla małych natężeń tonu zagłuszającego, w przypadku odpowiednio dużej odległości w skali częstotliwości, efekt zagłuszania nie występuje.
4. Tony o dostatecznie dużych częstotliwościach i dużych natężeniach zagłuszają wszystkie dźwięki o częstotliwościach większych od częstotliwości tonu zagłuszającego, tony o częstotliwościach małych natomiast, tylko w bezpośrednim swoim sąsiedztwie.

Zależności między przesunięciem dolnej granicy słyszalności wszystkich tonów zagłuszanych tonami o różnych częstotliwościach i poziomach.





Zależności między przesunięciem dolnej granicy słyszalności a częstotliwością tonów zagłuszanych tonami o różnych częstotliwościach i poziomach natężenia dźwięku. Liczby przy krzywych podają poziom natężenia dźwięku zagłuszającego.



Rys. 11.30. Zależność między przesunięciem dolnej granicy słyszalności a częstotliwością tonów, zagłuszanych szumami o wąskich i szerokich pasmach częstotliwości i różnych poziomach natężenia dźwięku

Każdy ton sinusoidalny jest zagłuszany jedynie przez wąskie pasmo szumu położone obok niego, a dalsza część białego szumu nie wpływa na zagłuszanie tego tonu.

Fakt, że jedynie szumy leżące w wąskim paśmie zagłuszają zawarty w nim ton prosty (ton sinusoidalny), a wszystkie szumy poza tym pasmem nie odgrywają żadnej roli prowadzi na do pojęcia **PASM KRYTYCZNYCH**, których szerokość jest zależna od częstotliwości tonu zagłuszanego i zmienia się od 30 Hz przy małych częstotliwościach do kilkuset Hz przy częstotliwościach dużych.

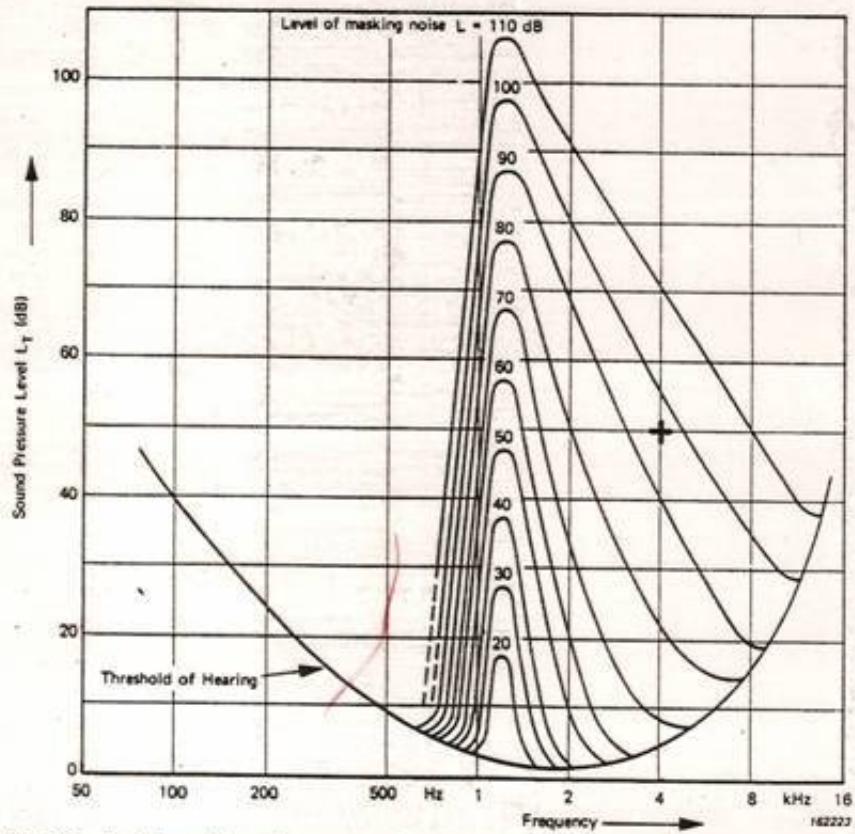
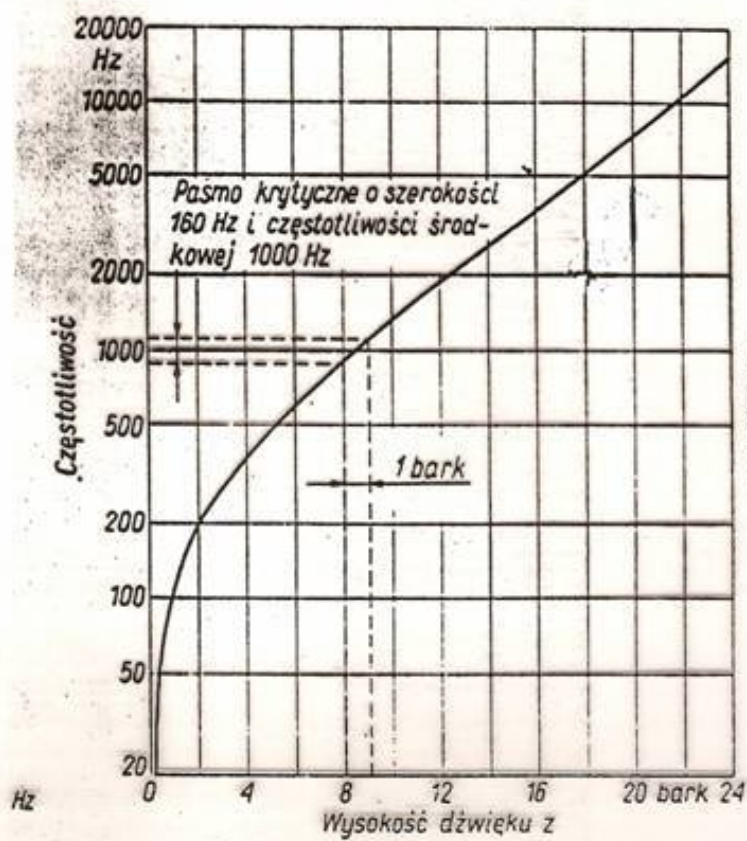


Fig.3.4. Masking effect of a narrow band noise centred at 1200 Hz at various levels (after Zwicker). A 50 dB 4 kHz tone (marked +) can be heard if the masking noise level is 90 dB, but is masked if its level rises to 100 dB

Critical Band (Bark)	1	2	3	4	5	6	7	8
Centre Frequency (Hz)	50	150	250	350	450	570	700	840
Bandwidth $f$ (Hz)	100	100	100	100	110	120	140	150
Critical Band (Bark)	9	10	11	12	13	14	15	16
Centre Frequency (Hz)	1000	1170	1370	1600	1850	2150	2500	2900
Bandwidth $f$ (Hz)	160	190	210	240	280	320	380	450
Critical Band (Bark)	17	18	19	20	21	22	23	24
Centre Frequency (Hz)	3400	4000	4800	5800	7000	8500	10500	13500
Bandwidth $f$ (Hz)	550	700	900	1100	1300	1800	2500	3500

Fig.3.5. Table of critical bands (Frequenz-gruppen)

180122



**Rys. 11.17.** Zależność między wysokością tonu (w barkach) a częstotliwością (w Hz) [24]

## **Pasma krytyczne wg Fletchera**

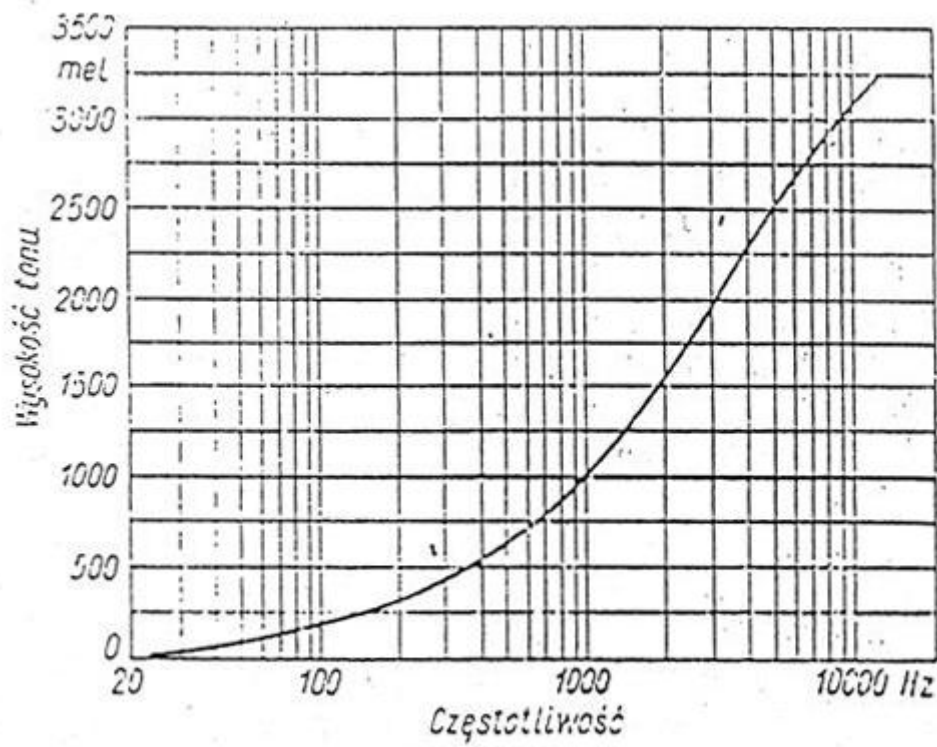
- 1 Dźwięk prosty (ton sinusoidalny) jest głównie maskowany przez częstotliwości zawarte w pewnym paśmie (nazwanym pasmem krytycznym), położonym dookoła pewnej częstotliwości środkowej. Częstotliwości znajdujące się poza tym pasmem (z wyjątkiem dużych poziomów), wpływają na maskowanie tego tonu bardzo mało.
- 2 Dźwięk prosty (ton sinusoidalny) jest wtedy percypowany w białym szumie, gdy jego energia jest co najmniej równa całkowitej energii składowych zawartych wewnątrz pasma krytycznego odpowiadającego częstotliwości tego dźwięku.

Hipoteza ta prowadzi do stwierdzenia:

**- Szerokość pasma krytycznego  $\Delta f_k$  wyrażona w decybelach (jako  $10 \log \Delta f_k$ ) jest równa liczbie decybeli, o jaką trzeba podnieść poziom dźwięku prostego ponad poziom gęstości widmowej szumu maskującego (dB/1Hz), aby był on już słyszalny”.**

**Zwicker** wychodząc z zależności związanych z percepcją poziomu ciśnienia akustycznego zaproponował pasma krytyczne szersze od podanych przez Fletchera i nazwał je „**grupami częstotliwości**”. Według Zwickera całkowite pobudzenie błony podstawnej składa się z sumy pobudzeń częściowych, z których każde odpowiada szerokości jednego pasma krytycznego (grupy częstotliwościowej).

Zakres częstotliwości słyszalnych podzielono na 24 kolejne grupy częstotliwościowe, z których każda ma szerokość jednego barka. Stwierdzono, że grupie częstotliwościowej odpowiada długość błony podstawnej równa około 1,3 mm.



Rys.3 Zależność wysokości tonu od częstotliwości

# **Modelowanie zjawisk percepcyjnych**

- Wyznaczenie absolutnego progu słyszenia
- Modelowanie pasm krytycznych
- Modelowanie maskowania nierównoczesnego
- Modelowanie maskowania równoczesnego
- Aproksymacja wychyleń błony podstawnej
- Pobudzenie błony podstawnej
- Sumowanie pobudzeń
- Aproksymacja sumarycznego wychylenia błony podstawnej
- Globalny próg maskowania

## **Wyznaczenie absolutnego progu słyszenia** (ang. *Absolute Threshold of Hearing*)

Ważną cechą słuchu ludzkiego, charakteryzującego się ogromną dynamiką, jest dolna granica słyszenia, tzw. **absolutny próg słyszenia**.

Dźwięki o głośności usytuowane poniżej tej krzywej są niesłyszalne dla ludzkiego ucha.

Na podstawie danych eksperymentalnych, można znaleźć wyrażenie opisujące krzywą absolutnego progu słyszenia. Najbardziej powszechny jest model zaproponowany przez **Terharda** o funkcji aproksymującej absolutny próg słyszenia:

$$LT_q = 3,64 f^{-0,8} - 6,5 \exp[-0,6 \cdot (f - 3,3)^2] + 10^{-3} f^4$$

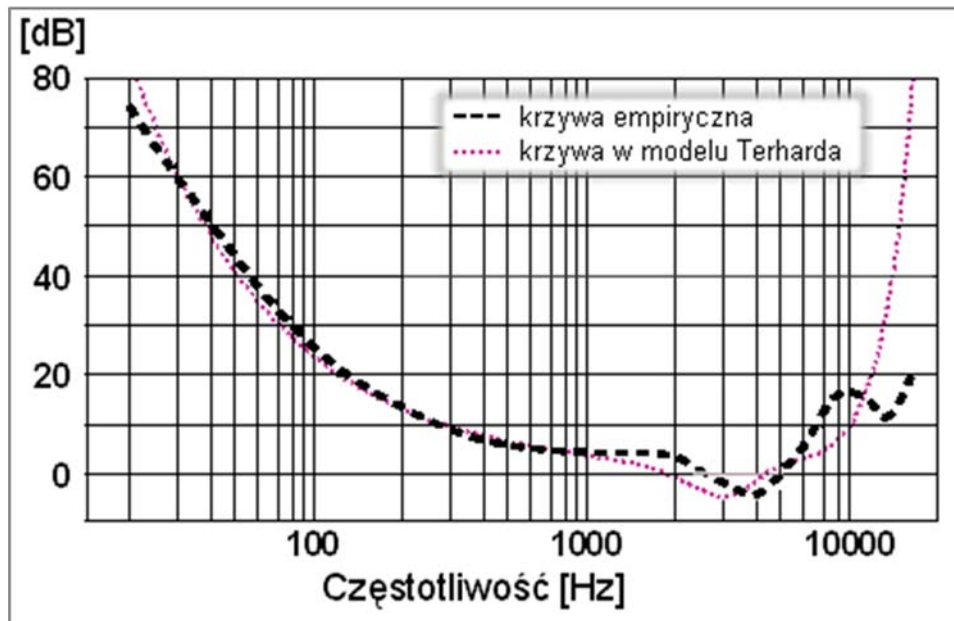
$LT_q$  – poziom progu słyszenia w dB

$f$  – częstotliwość w kHz



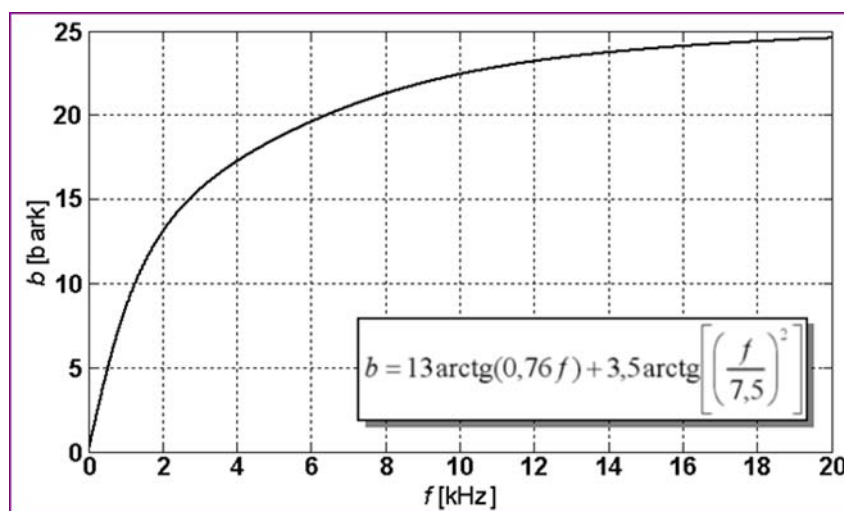
# Wyznaczenie absolutnego progu słyszenia (ang. *Absolute Threshold of Hearing*)

Krzywa absolutnego progu słyszenia w funkcji częstotliwości.



# Modelowanie pasm krytycznych słuchu

**Bark** – perceptualna jednostka dźwięku. Jeden bark odpowiada szerokości pojedynczego pasma krytycznego.



Wykres zależności skali barków od skali Hz według Zwickera

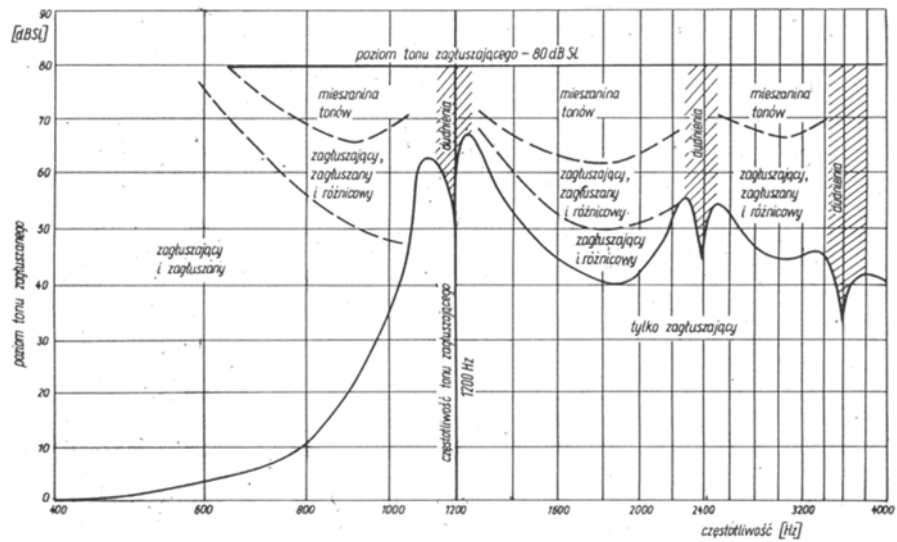
# Maskowanie psychoakustyczne

W algorytmach stratnej kompresji sygnału fonicznego modelowane są zjawiska **maskowania dźwięku**. Zjawiska te są spowodowane wychyleniem błony podstawnej narządu Cortiego pod wpływem bodźców akustycznych.

W związku z niejednorodną podatnością akustyczną błony podstawnej i ograniczoną liczbą komórek nerwowych narządu Cortiego, narząd słuchu zachowuje się jak **równoległy analizator widma o ograniczonej rozdzielczości widmowej i czasowej**.

# Maskowanie równoczesne

Ton o częstotliwości 1200 Hz i poziomie 80 dBSL zagłusza (ze skutkiem widocznym na wykresie) wszystkie tony o częstotliwościach większych, natomiast dla dźwięków o zdecydowanie mniejszych częstotliwościach, efekt zagłuszania prawie nie występuje.



## **Maskowanie równoczesne**

1. Największy efekt zagłuszania bez względu na wartość poziomu tonu zagłuszającego obserwuje się dla częstotliwości zbliżonych do częstotliwości tonu zagłuszającego. Przy czym łatwiej ulegają zagłuszaniu tony o częstotliwościach większych od tonu zagłuszającego, niż tony o częstotliwościach mniejszych

2. Dla częstotliwości tonu zagłuszanego równej częstotliwości tonu zagłuszającego (jak już wspomniano) zauważa się zjawisko dudnień. Obserwacja efektów zagłuszania jest wtedy utrudniona, podobnie jak i dla tonów zagłuszanych o częstotliwościach równych harmonicznym tonu zagłuszającego.

3. Dla małych natężeń tonu zagłuszającego, w przypadku odpowiednio dużej odległości w skali częstotliwości, efekt zagłuszania nie występuje.

## **Maskowanie równoczesne**

4. Tony o dostatecznie dużych częstotliwościach i dużych natężeniach zagłuszają wszystkie dźwięki o częstotliwościach większych od częstotliwości tonu zagłuszającego, tony o częstotliwościach małych natomiast, tylko w bezpośrednim swoim sąsiedztwie.

5. Obok przesunięcia progu słyszalności występują również dodatkowe, towarzyszące zjawiska: pojawianie się tonów różnicowych lub powstawanie mieszaniny tonów i dudnień.

6. W przypadku jednoczesnego istnienia wielu tonów (istnienia mieszaniny tonów) ich wzajemne oddziaływanie (maskowanie) jest związane z poziomem poszczególnych tonów oraz z ewentualną zmianą poziomu ogólnego.

# Podstawy kompresji - podsumowanie

## Maskowanie:

1. maskowanie sąsiednich częstotliwości (maskowanie jednoczesne): ciche dźwięki o częstotliwościach zbliżonych do częstotliwości dźwięku głośnego nie są słyszalne
2. maskowanie dźwięków następujących (maskowanie pobodźcowe): głośny dźwięk potrafi zagłuszyć cichsze dźwięki następujące zaraz po nim
3. maskowanie dźwięków poprzedzających (maskowanie wsteczne): cichy dźwięk poprzedzający w krótkim czasie dźwięk głośny nie jest słyszalny