



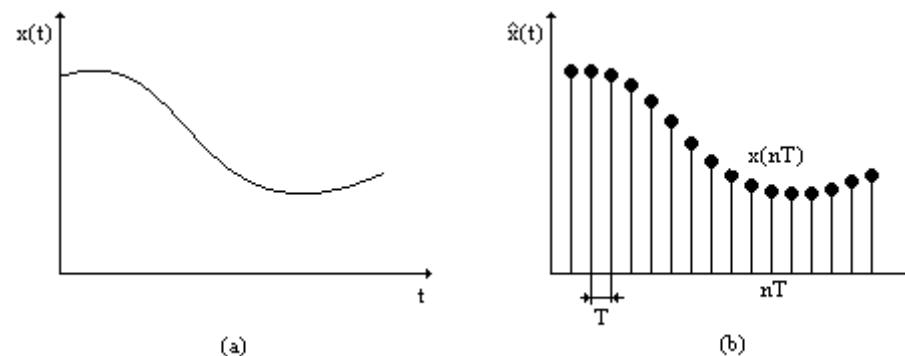
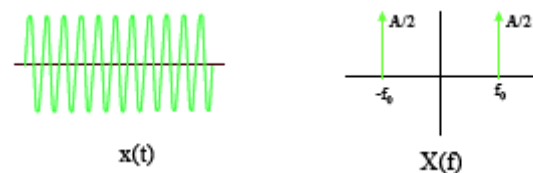
Osnove obrade audio signala: Karakteristike zvuka i osnove psihoakustike

Algoritmi i Arhitecture DSP 2

Osnovne karakteristike zvuka

- Analogni pristup zvuku:
 - promena amplitude u vremenu
 - funkcija frekvencije
- Digitalni zvuk:
 - diskretizacija po vremenu (odabiranje)
 - diskretizacija po amplitudi (kvantovanje)

(T je perioda odabiranja, $F_s = 1/T$ učestalost odabiranja, $x(nT)$ vrednost kontinualnog signala u trenutku odabiranja)



Nikvistova (Nyquist) teorema ili teorema odabiranja: Vremenski kontinualni signal može se potpuno rekonstruisati iz svojih odbiraka ako je, i samo ako je, učestalost odabiranja bar dva puta veća od najviše učestalosti u spektru signala.

Nivo zvučnog pritiska

(eng. Sound pressure level – SPL)

- Nivo zvučnog pritiska: zvuk kao funkcija vazdušnog pritiska u vremenu.
Meri se u paskalima :

$$1Pa = 1 \frac{N}{m^2}$$

- Opseg zvučnog pritiska

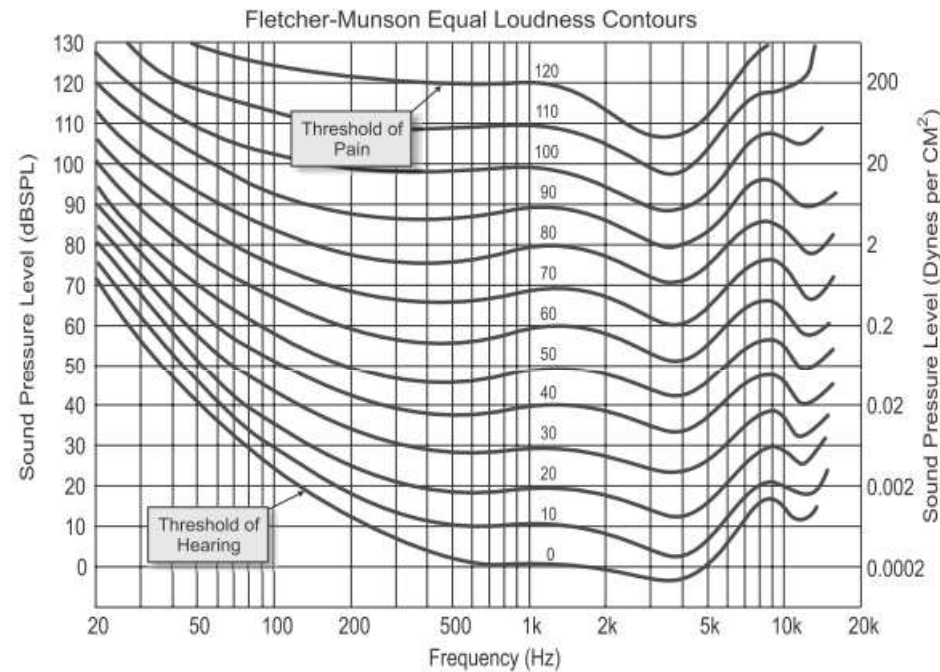
$$SPL = 10 \log_{10} \left(\frac{p}{p_0} \right)^2$$

- p_0 odgovara zvučnom pritisku na pragu čujnosti sa frekvencijom oko 2kHz

Opseg čujnosti

- Ljudi čuju zvuk na frekvenciji od 20 Hz do 20 kHz
- Ljudski govor od 100Hz do 8kHz
- Uobičajeni nivo SPL pri razgovoru je 50-60 dB
- Dinamički opseg muzike od 20 dB do 95 dB
- 100 dB rizik od oštećenja
- 120 dB prag bola

Kriva čujnosti



The graph describes the ear's sensitivity to various frequencies at different sound levels.

Figure 1.6 Fletcher-Munson Equal Loudness Contours

Opseg čujnosti

–logaritamska i linearna skala

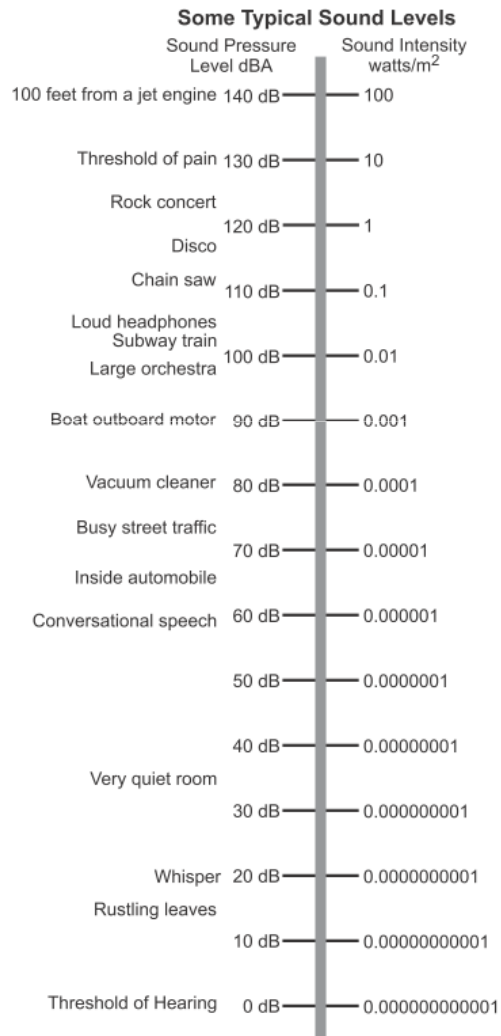
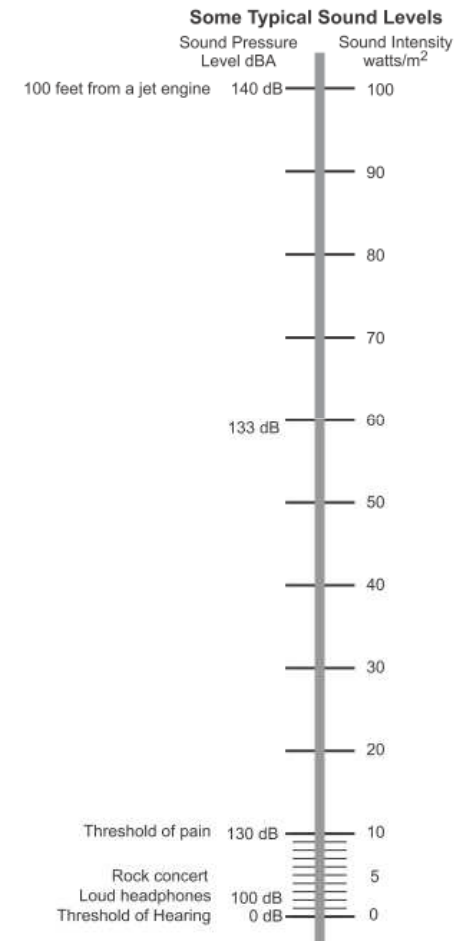


Figure 1.4 Typical Sounds and Their Levels in dB SPL and Intensity in W/m²

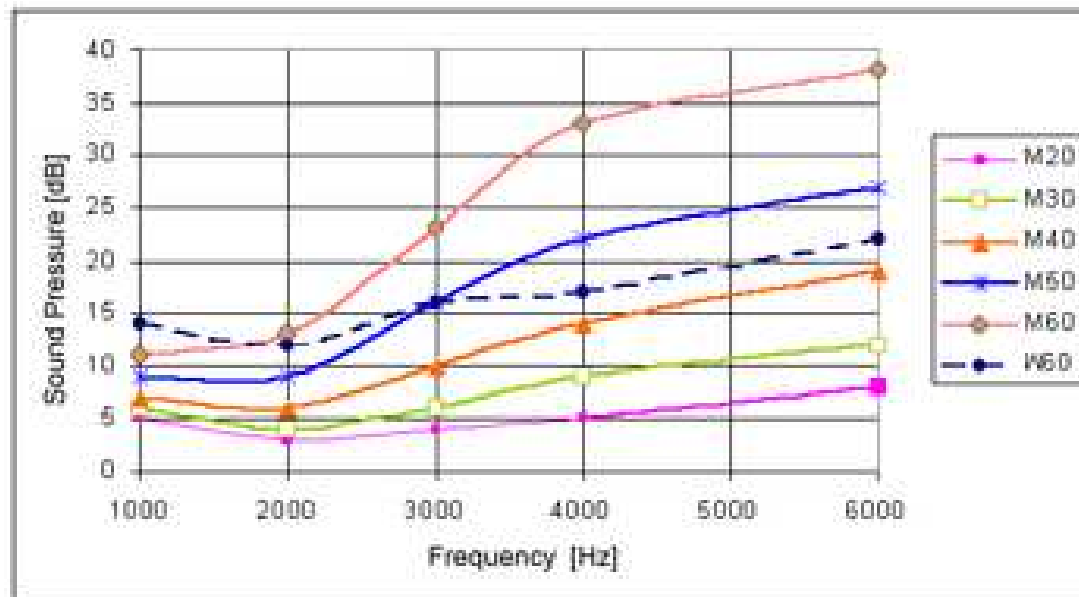


The scale from Figure 1.4 is redrawn as a linear scale showing its absurdity. All of the everyday mid-range intensities are compressed below the lower 10 percent of this scale.

Figure 1.5 The Typical Sound Intensity Levels Graph on a Linear Scale

Apsolutni prag čujnosti - ATH

- Prag čujnosti je najniži zvuk koji može da se čuje na određenoj frekvenciji – apsolutni prag čujnosti (Absolute Threshold of Hearing - ATH)
- Ljudsko uho ne registruje zvuk ispod ATH:
 - Sve frekvencije koje se nalaze ispod praga čujnosti su višak koji nije potrebno kodovati



Apsolutni prag čujnosti – *maksimum dozvoljene energije izobličenja koja se može uneti tokom obrade audio signala*

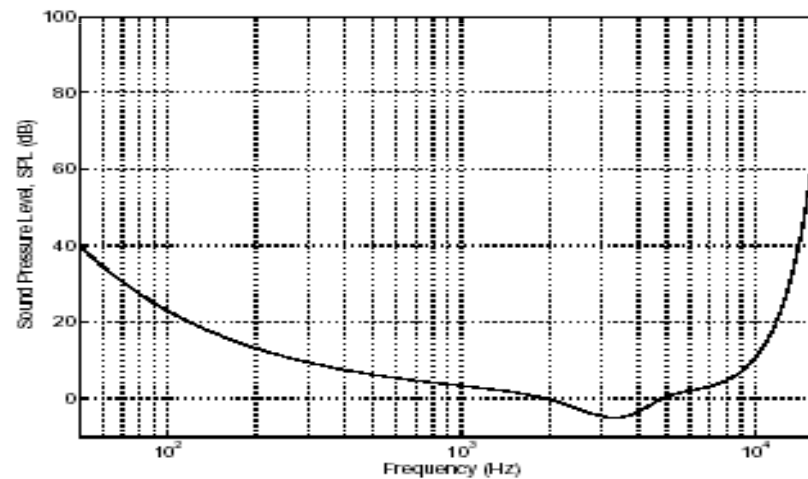
ATH nije savršena mera zbog sledećih razloga:

- aproksimacija krive je dobijena korišćenjem tonala, dok je kvantizaciona greška dobijena u procesu kodovanja značajno složenija u frekventnom domenu
- da bi se signal mogao porediti sa pragom čujnosti, potrebno je znati tačan nivo na kom će se signal reprodukovati (eng. playback level). Ovaj nivo najčešće nije poznat tokom procesa kodovanja

Rešenje drugog problema je da se najniža tačka ATH krive (oko 4 kHz) postavi na zvučni nivo signala sinusa sa amplitudom ± 1 LSB (bit najmanje važnosti).

Aproksimacija apsolutnog praga čujnosti

$$A(f)/_{dB} = 3.64(f/_{Hz})^{-0.8} - 6.5e^{-0.6(f/_{kHz}-3.3)^2} + 10^{-3}(f/_{kHz})^4$$



Kritični opsezi (1)

- Bubna opna vrši frekventno prostornu transformaciju i različiti delovi koštanog dela unutrašnjeg uha (eng. cochlea) 'podešeni' na različite frekventne opsege. Na ovaj način se uho može posmatrati kao neuniformna filter banka sa propusnim opsezima koji se preklapaju, što omogućava definisanje modela kritičnih opsega koji podražavaju ovu karakteristiku uha.
- Estimacija koju su dali **Zwicker i Fastl** (1990), koja opisuje varijaciju širine kritičnog opsega kao funkciju centralne frekvencije maskera :

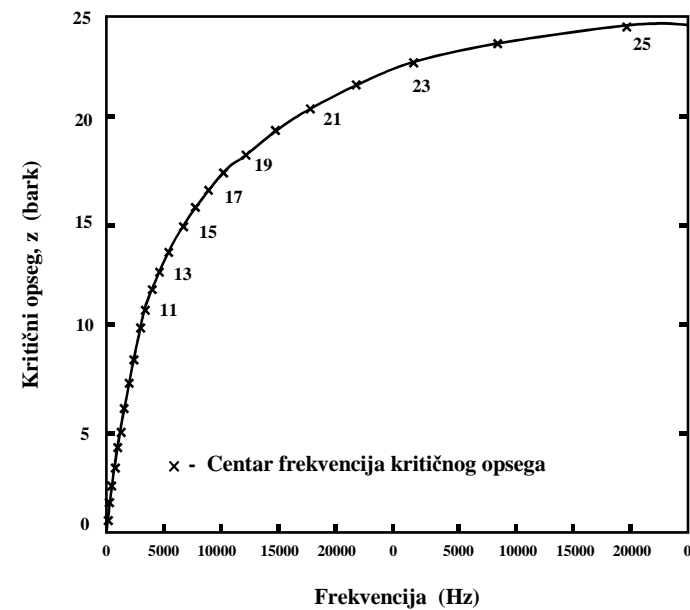
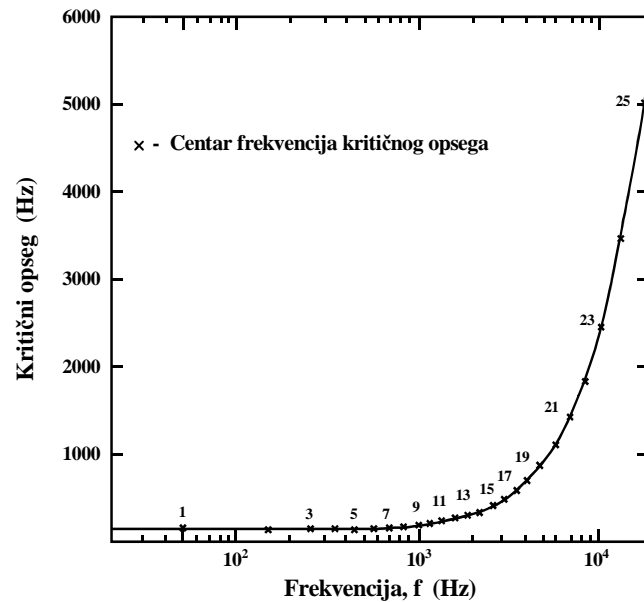
$$\Delta f / \text{Hz} = 25 + 75 \left[1 + 1.4 * \left(f_c / \text{kHz} \right)^2 \right]^{0.69}$$

- Koverzija Hz u Bark skalu, neuniformne u uniformnu skalu:

$$z(f) = 13 * \arctan(0.00076 * f) + 3.5 * \left[\left(f / 7500 \right)^2 \right]$$

Kritični opsezi (2)

- Prikaz kritičnih opsega u Hz i Bark jedinicama:



Idealna filter banka po kritični opsezima:

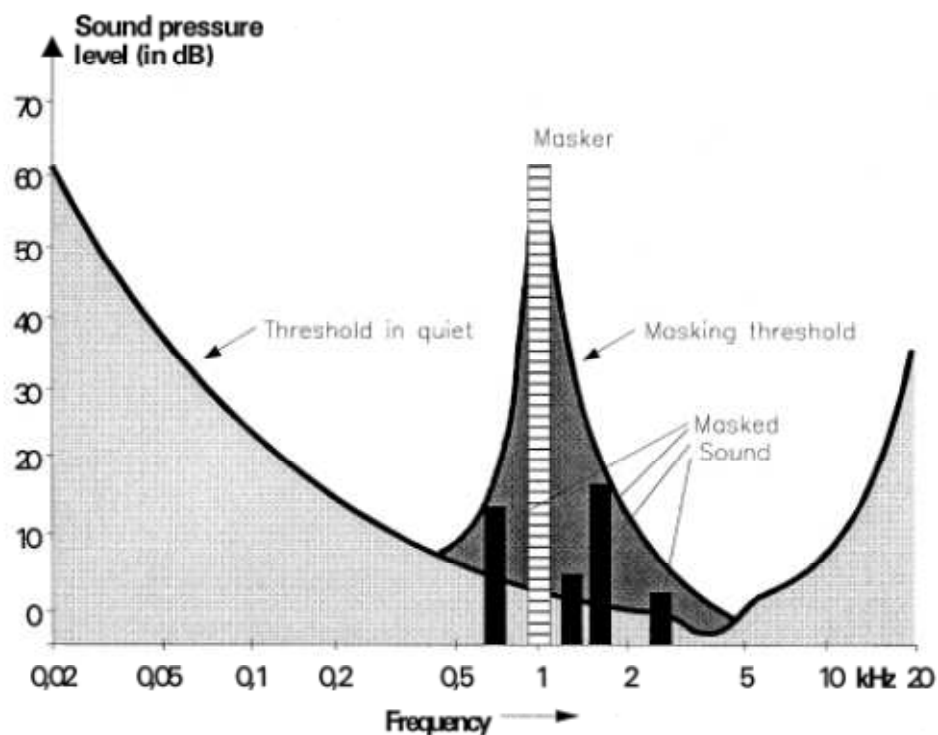
Broj Opsega	Centar Frek. (Hz)	Širina Opsega (Hz)
1	50	-100
2	150	100-200
3	250	200-300
4	350	300-400
5	450	400-510
6	570	510-630
7	700	630-770
8	840	770-920
9	1000	920-1080
11	1370	1080-1480
12	1600	1480-1720
13	1850	1720-2000
14	2150	2000-2320
15	2500	2320-2700
16	2900	2700-3150
17	3400	3150-3700
18	4000	3700-4400
19	4800	4400-5300
20	5800	5300-6400
21	7000	6400-7700
22	8500	7700-9500
23	10.500	9500-12000
24	13.500	12000-15500
25	19.500	15500-

Frekventno maskiranje zvuka (1)

- U svakodnevnom životu često se doživljava maskiranje slabog zvuka jakim zvukom: u trenutku razgovora (maskirani zvuk) prolazi kamion (masker) i više nije moguće čuti razgovor. Sličan primer imamo kada glasan instrument maskira slabiji instrument bliske frekvencije.
- Dominantan vid maskiranja kod stacionarnih, sporo promenjivih signala

Frekventno maskiranje zvuka (2)

- Visokopojasni masker zvučnog nivoa 60dB, 1kHz
- Maskira 4 signala koja su ispod praga maskiranja (iako su iznad ATH)
- Nagib praga maskiranja je strmiji prema nižim frekvencijama, odnosno jače je maskiranje ka višim frekvencijama.

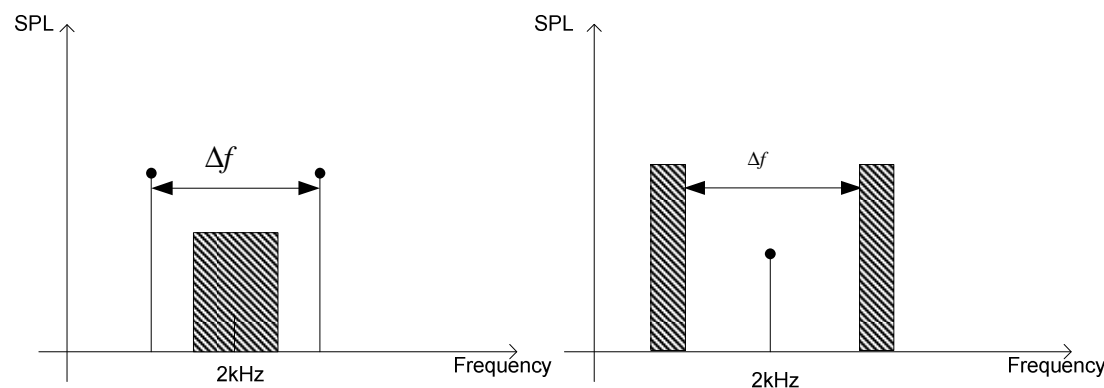


Frekventno maskiranje zvuka (3)

- Postoji uzak pojas maskiranja oko maskera u kome je prag konstantan, ako je širina pojasa manja od kritičnog opsega
- Na slici:
 - Test signal je uskopojasni šum centriran između dva maskera (koji si ton)
 - Test signal je tona između dva maskera (koji su šum)

Ako imamo jak tonal u okviru kritičnog opsega, zbog izraženog efekta maskiranja možemo imati veliku kvantizacionu grešku koja neće biti čujna

Manji razmak daje konstanto maskiranje, a veći lakšu detekciju maskera!



Frekventno maskiranje zvuka (4)

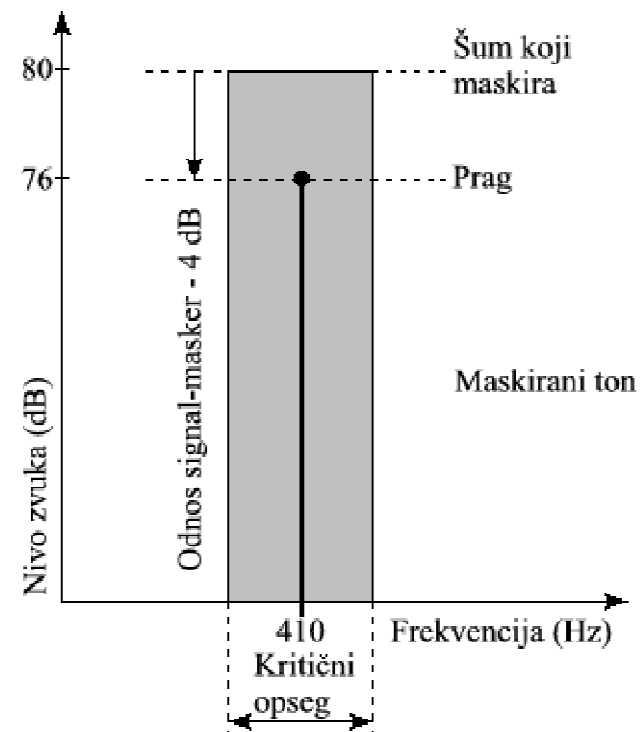
U zavisnosti od vrste jačeg signala, razlikuju se dve vrste maskiranja *unutar* kritičnog opsega (eng. intra-band masking):

- Šum maskira ton (eng. Noise-masking-tone)
- Ton maskira šum (eng. Tone-masking-noise)

Rastojanje između nivoa maskera i praga maskiranja se zove odnos signala i maske (SMR)

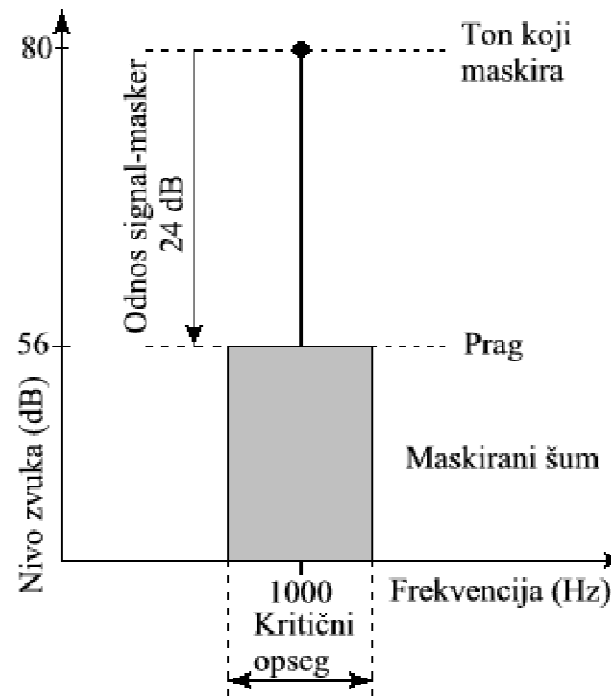
Na slici: šum koji maskira ton.

- SPL šuma 80 dB
- SPL tona 76 dB
- Odnos signala i maske -4dB



Frekventno maskiranje zvuka (5)

- Na slici: ton koji maskira šum (TMN).
 - SPL tona je 80 dB
 - SPL šuma 56 dB
 - SMR iznosi 24 dB

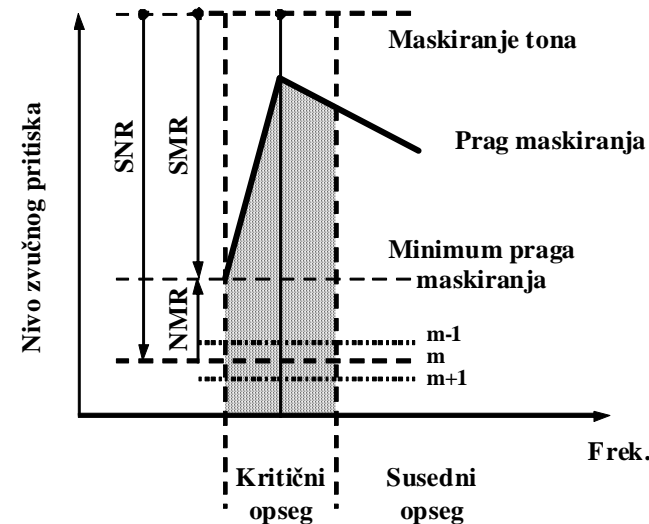


Mnogo veću sposobnost maskiranja imaju šum maskeri!

Prenosna funkcija maskiranja

Maskiranje se ne dešava samo u okviru kritičnih opsega (eng. intra band masking), već se prenosi na susedne opsege. Najjednostavnija aproksimacija prenošenja maskiranja je **trougona prenosna funkcija**, sa nagibima od + 25 dB do -10 dB po Bark-u.

Na slici: u okviru kritičnog opsega signali koji se nalaze ispod praga su maskirani.



- Kod kvantizacije sa m bita kvant. greška se ne čuje ako je SNR veći od SMR
 - Mera čujnog izobličenja je data preko odnosa šuma i maske (**NMR**): dok je negativan greška je nečujna
- $NMR(m) = SMR - SNR(m)$ (in dB)**

Vremensko maskiranje

- Maskiranje može da se dogodi i kada zvukovi nisu istovremeno prisutni: npr. govor posle koga dolazi jak zvuk.
- Fenomen maskiranja se može dogoditi pre i posle vremenskog perioda u kome se nalazi masker. Pre maskiranje je pomalo neočekivano, ono dolazi iz činjenice da nas slušni aparat treba neko vreme da bi povezao informacije za percepciju zvuka, kao i da jači zvuk zahteva duži period integracije od slabijeg zvuka

