

KODOWANIE PERCEPTUALNE

Opracował: Dominik Tyniów, PG pod kier. Prof.. A. Czyżewskiego

Wprowadzenie

Cyfrowy sygnał foniczny w postaci nieskompresowanej charakteryzuje się wysoką przepływnością strumienia bitów.

Przykład:

Sygnał audio o jakości CD:

liczba kanałów: 2

liczba bitów na próbkę: 16 bitów

częstotliwość próbkowania: 44,1 kHz

Przepływność:

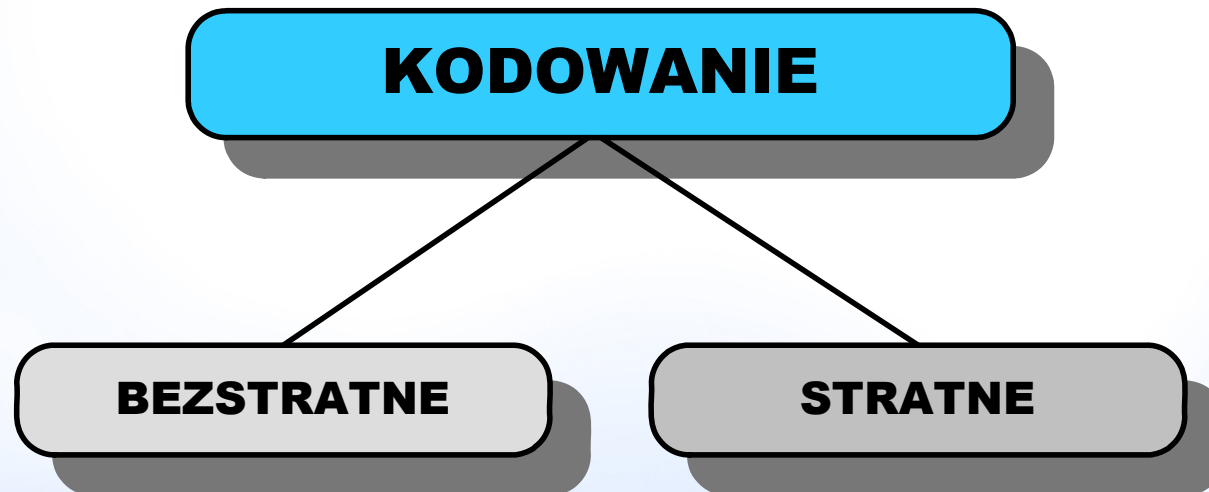
$$2 \times 16 \text{ [bitów]} \times 44100 \text{ [1/s]} = 1411200 \text{ [bitów/s]} \quad (\text{ok. } \mathbf{172 \text{ kB/s}})$$

WNIOSEK: 1 minuta muzyki o jakości płyty CD to ok. 10 MB danych !!!

Stanowi to istotny problem w przypadku archiwizacji oraz przesyłania poprzez sieci komputerowe.

Sposoby ograniczenia przepływności

Zastosowanie odpowiedniego kodowania pozwala na usunięcie redundancji (nadmiarowości) sygnału, a zatem ograniczenie przepływności strumienia bitów.



Kodowanie bezstratne

(ang. *lossless coding*)

Kodowanie bezstratne polega na przyporządkowaniu częściej pojawiającym się wartościom, krótszych słów kodowych.

Zastosowanie bezstratnych metod kodowania, takich jak *kodowanie arytmetyczne*, *kodowanie Huffmana* czy *Lempel-Zip*, bezpośrednio do sygnału audio nie jest zbyt efektywne m.in. ze względu na duży zakres kodowanych wartości.

Cechy bezstratnego kodowania audio:

- zachowana jakość z dokładnością do pojedynczych próbek sygnału
- niewielkie współczynniki kompresji (ok. 2:1)
- kaskadowość – możliwość wielokrotnego kodowania-dekodowania bez utraty jakości

Kodowanie bezstratne

(ang. *lossless coding*)

Zastosowania bezstratnego kodowania:

- archiwizacja
- edycja
- DVD-Audio

Przykłady algorytmów bezstratnej kompresji audio:

- Meridian Lossless Packing (MLP) – stosowany w DVD-Audio
- MPEG-4 Audio Lossless Coding (ALS)
- WavPack
- RKAU
- FLAC
- LPAC



Kodowanie bezstratne

(ang. *lossless coding*)

Tabela 1. Współczynniki kompresji plików audio otrzymane za pomocą standardowego oraz dedykowanego oprogramowania kompresującego.

| Typ | STANDARDOWE | | DEDYKOWANE | | | |
|---------------------------|-------------|--------|---------------|----------------|---------------|--------|
| | WinZip | WinRAR | WavPack | Monkey's Audio | RKAU | FLAC |
| Chór | 1,30:1 | 1,88:1 | 2,36:1 | 2,48:1 | 2,44:1 | 2,29:1 |
| Orkiestra | 1,05:1 | 1,41:1 | 1,77:1 | 1,53:1 | 1,50:1 | 1,46:1 |
| Rock | 1,08:1 | 1,45:1 | 1,67:1 | 1,72:1 | 1,46:1 | 1,43:1 |
| Pop | 1,05:1 | 1,35:1 | 1,47:1 | 1,47:1 | 1,59:1 | 1,52:1 |
| Disco | 1,11:1 | 1,43:1 | 1,57:1 | 1,61:1 | 1,67:1 | 1,59:1 |

Kodowanie stratne

(ang. *lossy coding*)

Kodowanie stratne polega na wyeliminowaniu części sygnału bez wyraźnego pogorszenia jego subiektywnej jakości.

Kodowanie perceptualne jest kodowaniem stratnym i wykorzystuje zjawisko maskowania w pasmach krytycznych słuchu.

Cechy stratnego kodowania audio:

- nieodwracalna utrata informacji
- możliwość osiągnięcia wysokich współczynników kompresji (ok. 12:1 dla jakości porównywalnej z płytą CD)
- utrata jakości sygnału przy kaskadowym łączeniu kodeków

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:

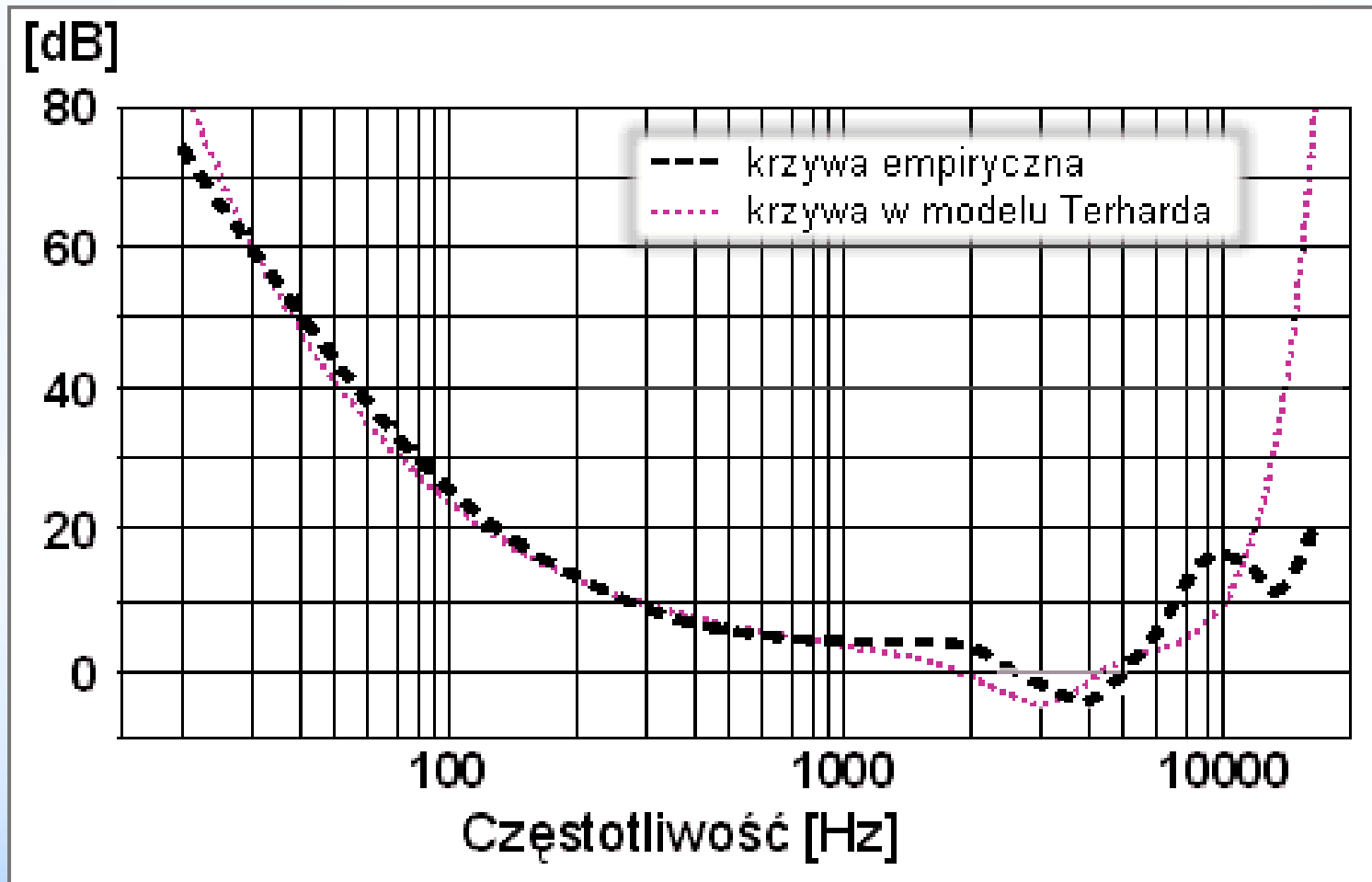
$$LT_q = 3,64 f^{-0,8} - 6,5 \exp[-0,6 \cdot (f - 3,3)^2] + 10^{-3} f^4$$

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

System słuchowy człowieka przetwarza dźwięk w pewnych podpasmach, zwanych *pasdami krytycznymi*.

Definicja pasma krytycznego według *Fletcher*:

pasmo krytyczne jest elementarnym pasmem częstotliwości o szerokości równej Δf , wydzielonym z ciągłego widma mocy szumów i zawierającym w sobie moc akustyczną równą mocy akustycznej tonu prostego o częstotliwości f położonej w środku tego pasma, przy czym rozpatrywany ton prosty ma taką intensywność, że zagłuszany przez nieograniczone widmo szumów ciągłych, znajduje się dokładnie na granicy słyszalności.

Definicja pasma krytycznego według *Zwicker*:

Przy zwiększaniu szerokości pasma szumu, głośność pozostaje na tym samym poziomie, dopóki nie zostanie przekroczona szerokość ***pasma krytycznego***. Wówczas wrażenie głośności ulega zmianie.

Modelowanie pasm krytycznych słuchu

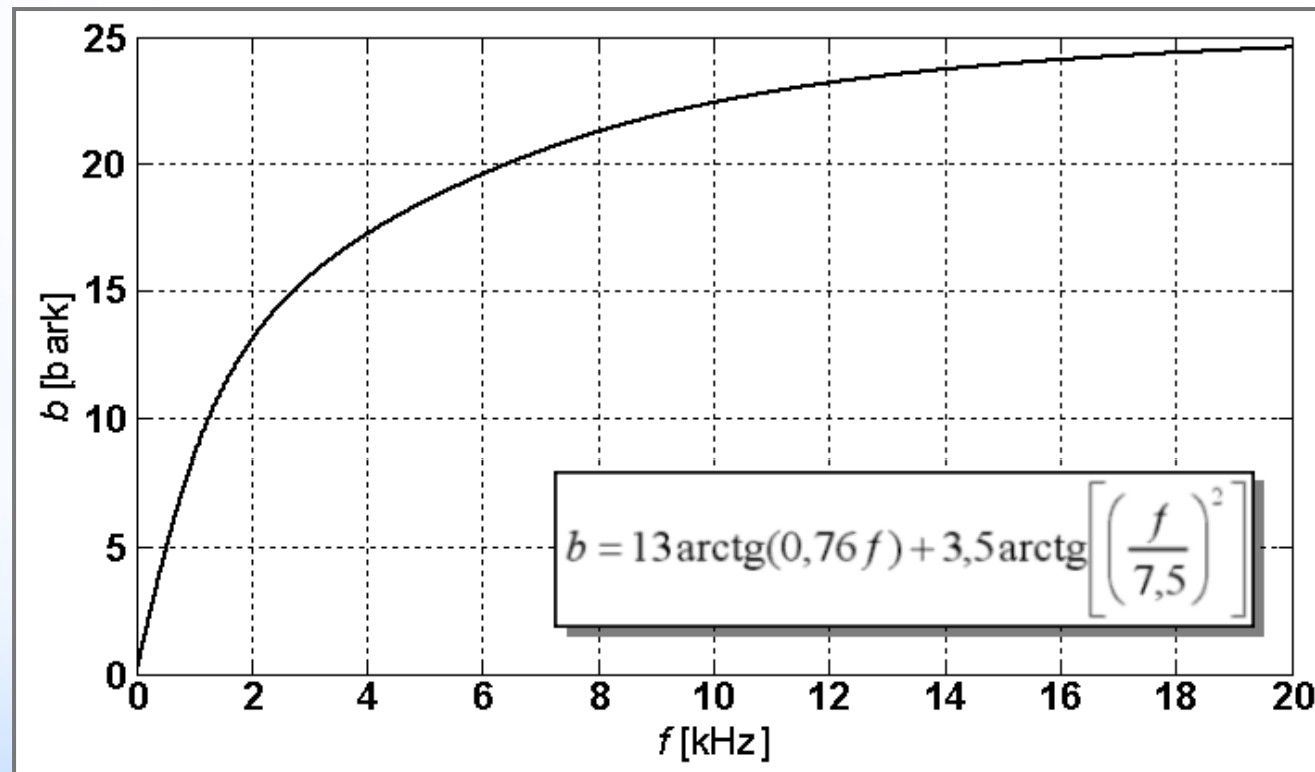
Każdemu pasmu krytycznemu odpowiada pewien odcinek błony podstawnej ślimaka równy ok. 1,3 mm.

System słuchowy może być modelowany jako zestaw filtrów pasmowo-przepustowych, dla których szerokość pasma jest równa szerokości odpowiedniego pasma krytycznego.

Bezwzględne szerokości pasm krytycznych nie są jednakowe. Poniżej częstotliwości 500 Hz szerokość pasma krytycznego jest stała i wynosi ok. 100 Hz, powyżej 500 Hz szerokość każdego następnego pasma krytycznego jest o 20 % większa niż szerokość poprzedniego pasma.

Modelowanie pasm krytycznych słuchu

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



Wykres zależności skali barków od skali Hz według *Zwicker*

Maskowanie psychoakustyczne

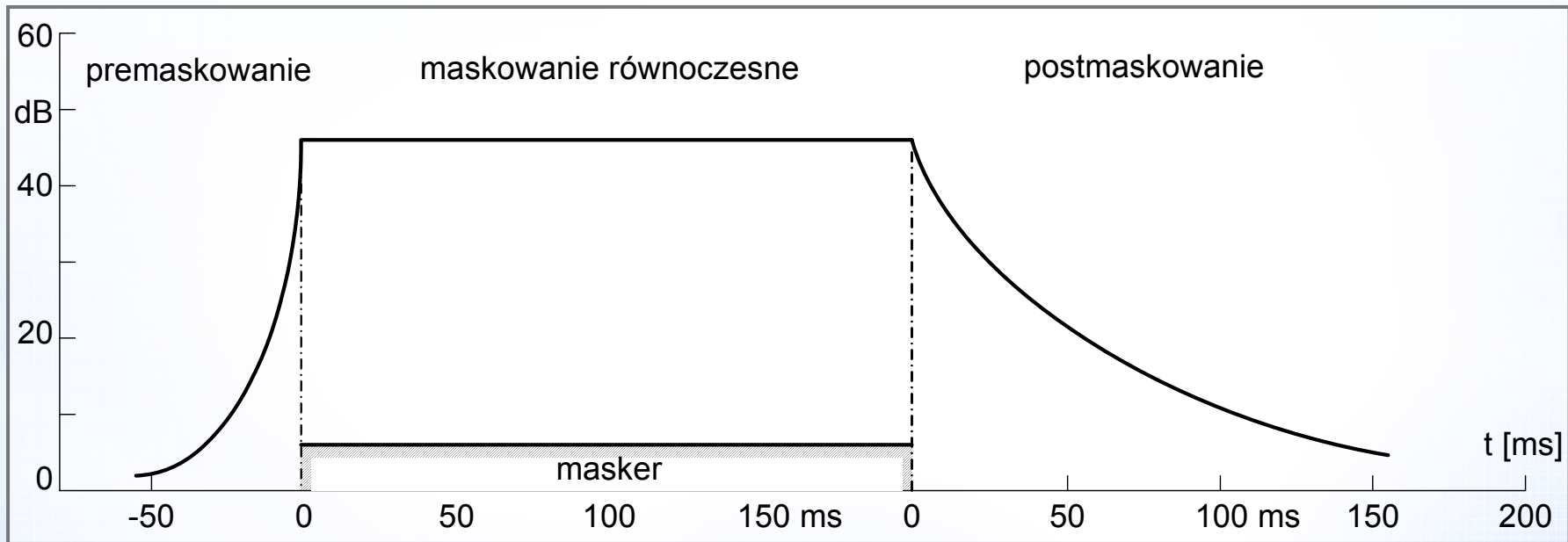
W algorytmach stratnej kompresji sygnału fonicznego modelowane są zjawiska **maskowania dźwięku**. Zjawiska te są spowodowane wychylaniem błony podstawnej narządu Cortiego pod wpływem bodźców akustycznych.

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 nierównoczesne

(ang. *temporal masking*)

Przykład maskowania nierównoczesnego



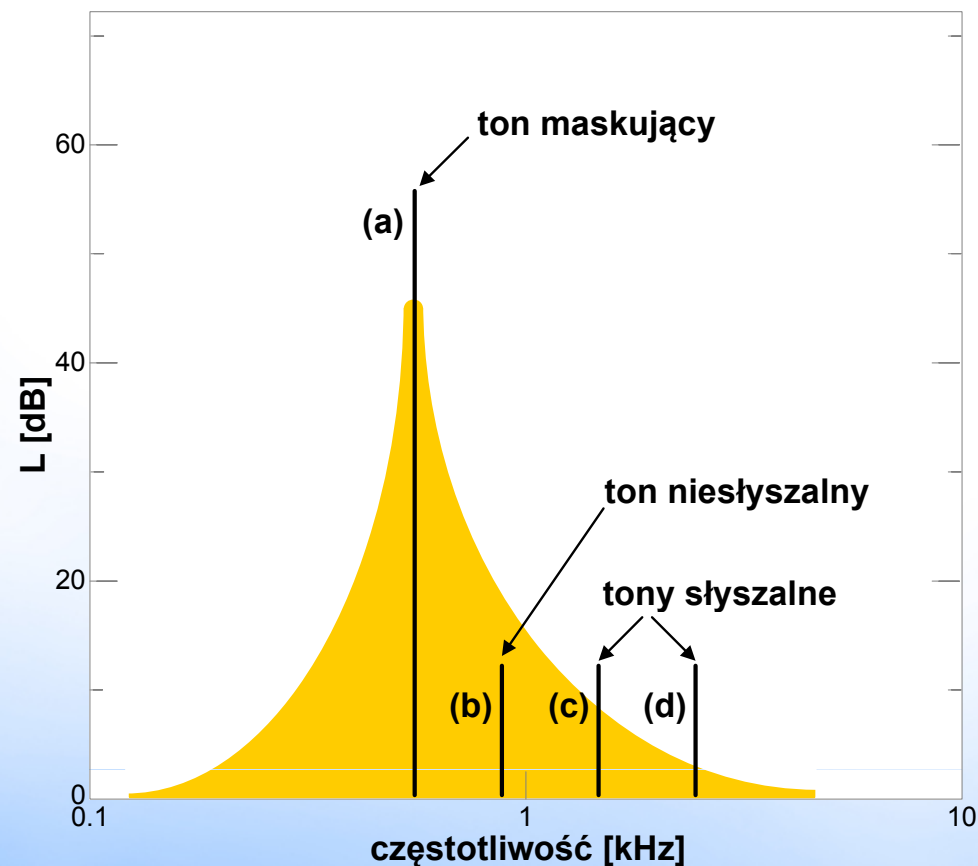
premaskowanie – dźwięk o dużym natężeniu jest w stanie zamaskować tony, które wystąpiły wcześniej. Premaskowanie trwa od 10 do 30 ms.

postmaskowanie – po wystąpieniu głośnego tonu, pozostałe dźwięki mogą nie być słyszane. Postmaskowanie trwa do ok. 200 ms. Czas ten zależy od natężenia oraz czasu trwania tonu maskującego.

Maskowanie równoczesne

(ang. *simultaneous masking*)

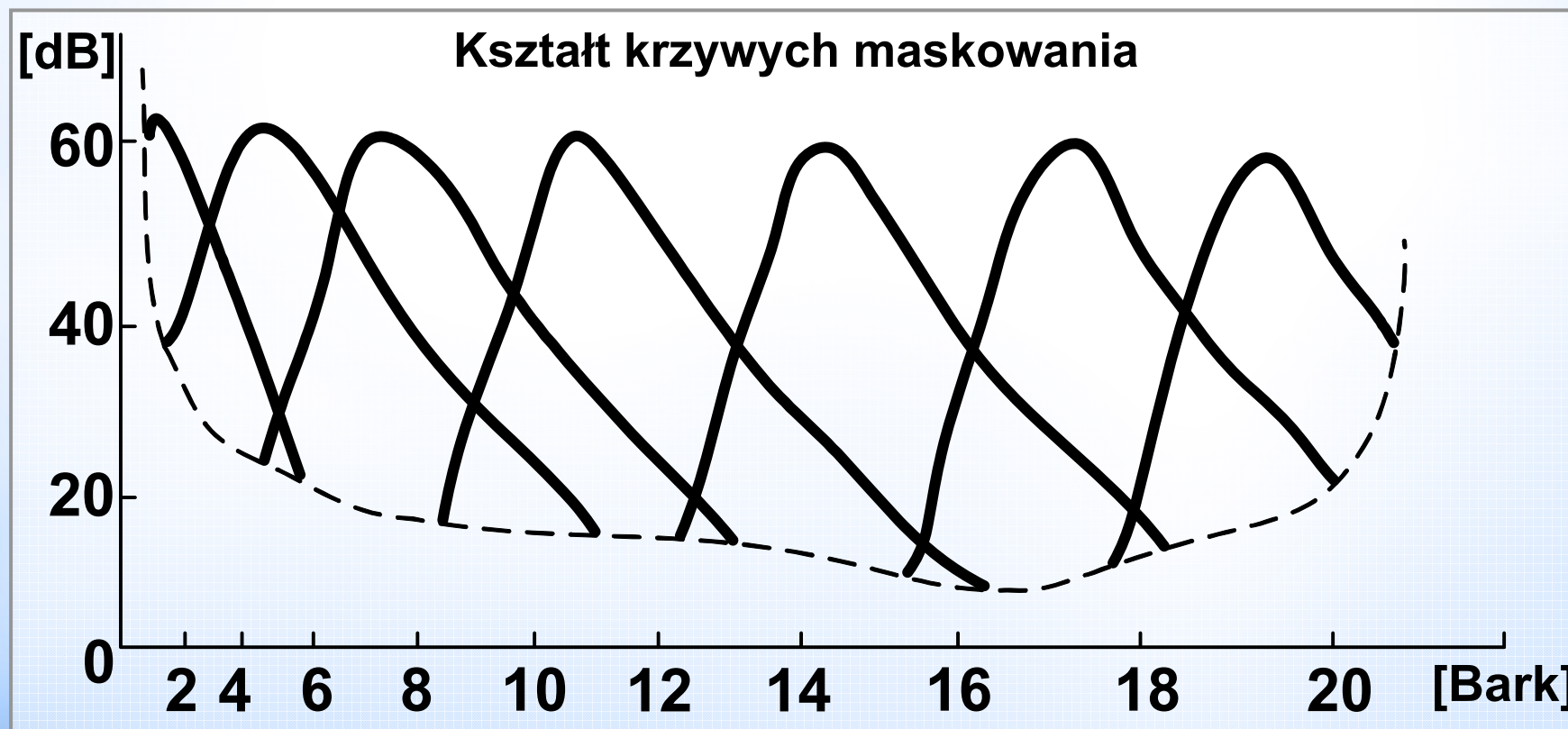
Maskowanie równoczesne charakteryzuje się tym, że pewne tony stają się niesłyszalne w obecności innych – tzw. **maskerów**.



Maskowanie równoczesne

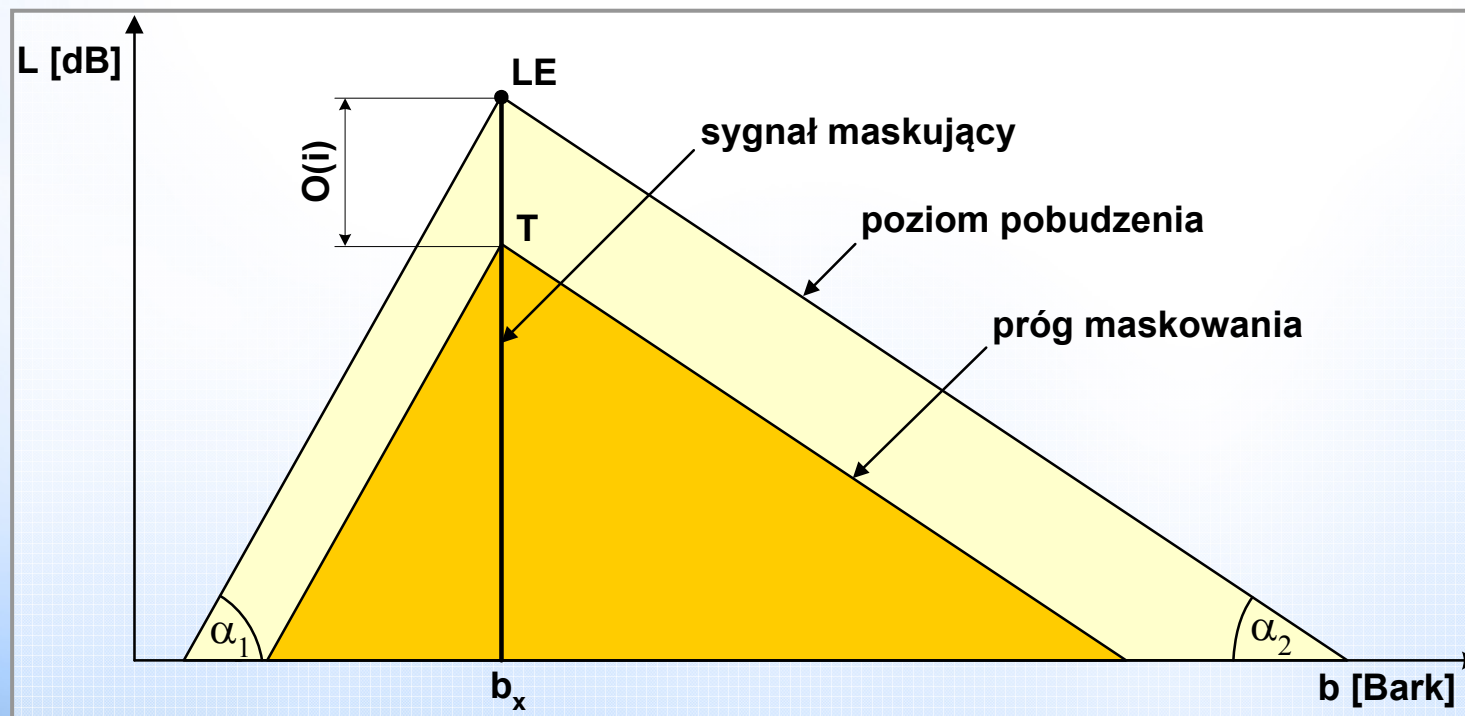
(ang. *simultaneous masking*)

Nachylenie zbocza **krzywych maskowania** po stronie niższych częstotliwości jest praktycznie stałe. Po drugiej stronie – zależy od częstotliwości i głośności poziomu maskera. Im ton maskujący jest głośniejszy, tym zbocze bardziej płaskie i zwiększa się wpływ maskowania na tony o wyższej częstotliwości.



Aproksymacja wychyleń błony podstawnej

W wyniku pobudzenia błony podstawnej zostaje ona wychylona z położenia równowagi. Kształt odkształconej błony, tzw. **poziom pobudzenia**, aproksymuje się przy pomocy dwóch odcinków nachylonych pod kątem α_1 i α_2 względem osi częstotliwości.



Aproksymacja wychyleń błony podstawnej

Nachylenia odcinków aproksymujących wychylenie błony podstawnej przy pobudzeniu LE o częstotliwości f_x można wyrazić przy pomocy zależności:

$$\begin{cases} S_1 = 31 \\ S_2 = 22 + \min(0,23 \cdot f_x^{-1}, 10) - 0,2 \cdot LE \end{cases} \quad \text{albo} \quad \begin{cases} S_1 = 27 \\ S_2 = 24 + 0,23 \cdot f_c^{-1}(i) - 0,2 \cdot LE \end{cases}$$

S_1, S_2 – nachylenia wyrażone w dB/Bark

LE – poziom głośności sygnału maskującego w dB

f_x – częstotliwość w kHz

$f_c(i)$ – częstotliwość środkowa i -tego pasma krytycznego w kHz

Aproksymacja wychyleń błony podstawnej

Próg maskowania wywołany pobudzeniem LE jest aproksymowany przez krzywą T odległą od oszacowanego wychylenia błony podstawnej o wartość $O(i)$:

$$O(i) = \alpha(14,5 + i) + (1 - \alpha)\alpha_v$$

- α – indeks tonalności ($0 \leq \alpha \leq 1$), wskazuje na charakter sygnału pobudzenia. Dla czystego tonu $\alpha = 1$, natomiast dla pobudzenia szumem $\alpha = 0$,
- i – numer pasma krytycznego, w którym nastąpiło pobudzenie
- α_v – indeks maskowania

$$\alpha_v = -2 - 2,05 \cdot \arctg(0,25 f_x) - 0,75 \cdot \arctg\left(\frac{f_x^2}{2,56}\right) \quad \text{albo} \quad \alpha_v = 5,5$$

- f_x – częstotliwość pobudzenia w kHz

Aproksymacja wychyleń błony podstawnej

Indeks tonalności wyznaczany jest na podstawie parametru określającego charakter widmowy sygnału – *SFM* (ang. *Spectral Flatness Measure*). Parametr ten jest zdefiniowany jako stosunek średniej geometrycznej do średniej arytmetycznej widma mocy:

$$SFM = 10 \log_{10} \left(\frac{\left[\prod_{k=1}^{N/2} X_k \right]^{1/N/2}}{\frac{1}{N/2} \sum_{k=1}^{N/2} X_k} \right)$$

$$\alpha = \min\left(\frac{SFM}{SFM_{\max}}, 1\right)$$

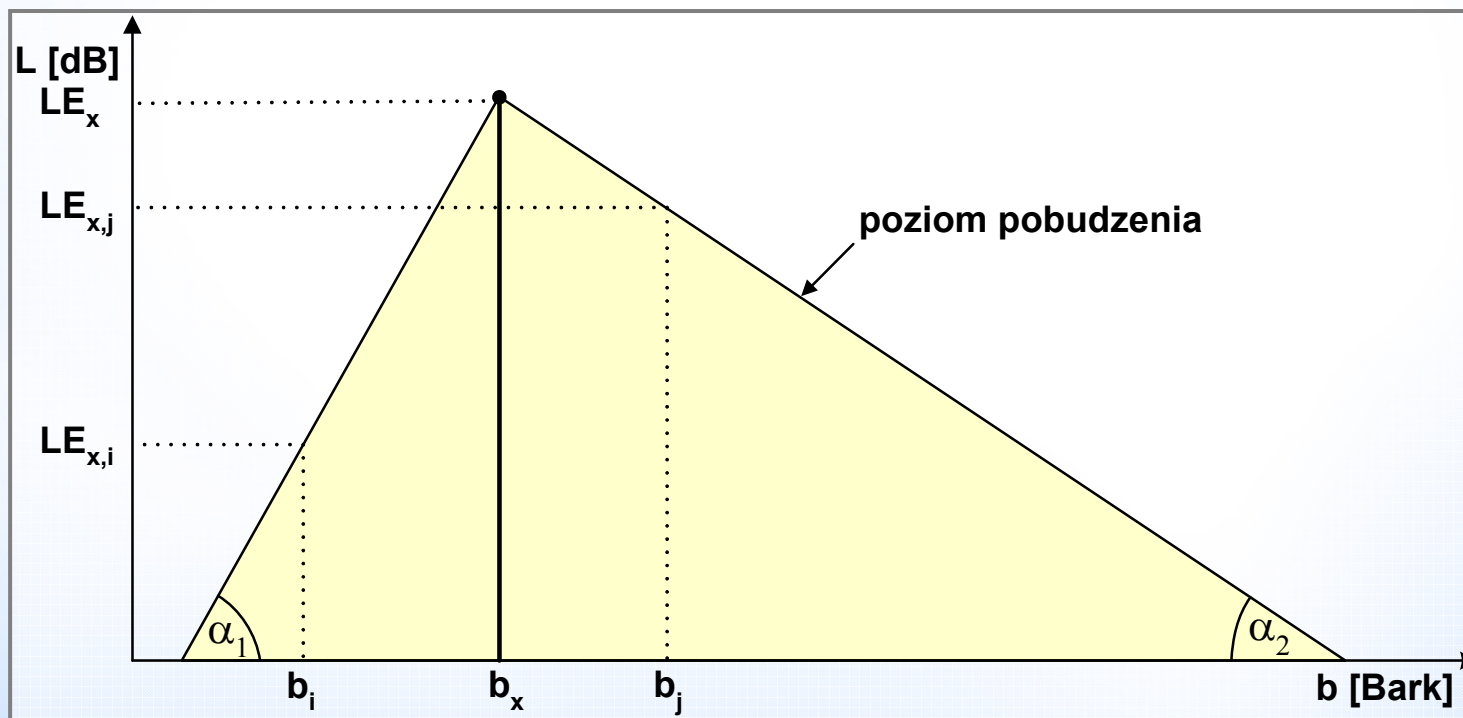
$$SFM_{\max} \stackrel{def}{=} -60 \text{ dB}$$

X_k – widmo gęstości mocy, określane przy pomocy N-punktowej DFT na podstawie widma krótkookresowego

Maskowanie szumem jest efektywniejsze niż maskowanie tonem prostym.

Pobudzenie błony podstawnej

Wpływ pobudzenia na wychylenie błony podstawnej



Pobudzenie błony podstawnej

Wychylenie błony podstawnej w miejscach odpowiadającym częstotliwości b_i ($b_i \leq b_x$) oraz częstotliwości b_j ($b_j > b_x$) przy pobudzeniu sygnałem o częstotliwości b_x można wyrazić za pomocą wzorów:

- w mierze logarytmicznej:

$$\begin{cases} LE_{x,i} = LE_x - S_1(b_x - b_i), & b_i \leq b_x \\ LE_{x,j} = LE_x - S_2(b_j - b_x), & b_j > b_x \end{cases}$$

- w mierze liniowej:

$$\begin{cases} E_{x,i} = E_x \cdot 10^{-\frac{1}{10}S_1\Delta b_{x,i}}, & \Delta b_{x,i} = b_x - b_i \wedge b_i \leq b_x \\ E_{x,j} = E_x \cdot 10^{-\frac{1}{10}S_2\Delta b_{x,j}}, & \Delta b_{x,j} = b_j - b_x \wedge b_j > b_x \end{cases}$$

Sumowanie pobudzeń

Summaryczne pobudzenie E w dowolnym miejscu błony może być modelowane przy pomocy relacji, która jest spełniona zarówno w dziedzinie czasu jak i częstotliwości:

$$E = \left(\sum_{i=1}^n E_i^\alpha \right)^{1/\alpha}, \alpha \leq 2$$

- E_i – liniowe pobudzenie (wychylenie) błony podstawnej w miejscu odpowiadającym częstotliwości b_i
- n – liczba pobudzeń błony podstawnej
- α – współczynnik kompresji, w praktycznej implementacji związanej z tzw. „Modelem Psychoakustycznym 1” (standard MPEG) przyjmuje się, że poziomy pobudzeń pochodzących z różnych miejsc błony podstawnej są addytywne ($\alpha = 2$)

Sumaryczne wychylenie błony podstawnej

Sumaryczne wychylenie błony podstawnej modeluje się przy pomocy splotu, wyrażającego wpływ mocy sygnału $X(i)$ w i -tym paśmie krytycznym na sąsiednie pasma krytyczne:

$$E_G(i) = X(i) * B(i, j) = \sum_{j=1}^{25} X(j)B(i - j)$$

$E_G(i)$ – sumaryczne pobudzenie błony podstawnej w i -tym paśmie krytycznym, wyrażone w skali liniowej,

$B(i, j)$ – funkcja rozkładu energii wzdłuż błony podstawnej

$X(i)$ – moc sygnału w skali liniowej w i -tym paśmie krytycznym

Globalny próg maskowania

Na skutek sumowania się energii poszczególnych pobudzeń błona podstawna ulega wychyleniu, co prowadzi do zamaskowania sygnałów, których energia nie przekracza progu maskowania.

Próg maskowania można określić korzystając z zależności, w oparciu o funkcję wagową $w(i)$:

$$LT(i) = w(i) \cdot LE(i)$$

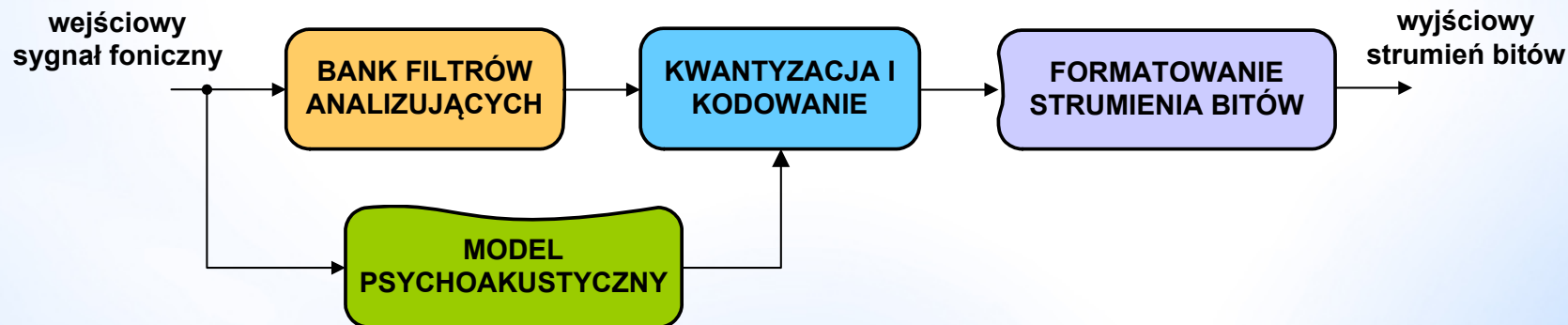
$$10 \log_{10} w(i) \cong -(15,5 + i)$$

$LT(i)$ – próg maskowania dla i -tego pasma krytycznego w dB

$LE(i)$ – pobudzenie w i -tym paśmie krytycznym w dB

Perceptualny koder foniczny

Znakomita większość obecnie stosowanych standardów perceptualnej kompresji dźwięku opiera się na **kwantyzacji poziomego sygnału fonicznego**. W ten sposób uzyskuje się redukcję objętości.



Jeśli produkt uboczny kwantyzacji – **szum kwantyzacji** znajduje się poniżej progu percepcji, to materiał dźwiękowy zachowuje wysoką jakość subiektywną.

Standardy stratnej kompresji dźwięku

Przykłady standardów stratnej kompresji dźwięku:

- **Sony ATRAC** (system MiniDisc)
- **Dolby AC-3** (wielokanałowe systemy kina cyfrowego)
- **PASC** (magnetofon cyfrowy DCC)
- **MUSICAM** (radiofonia cyfrowa DAB)
- **MPEG**
- **AAC** (*Apple*)
- **Ogg Vorbis** (projekt *open-source*)
- **WMA** (*Microsoft*)

Kodek Sony ATRAC

(ang. *Adaptive Transform Acoustic Coding*)



Nośnik MiniDisc:

- Dysk magnetoptyczny o średnicy 64 mm,
- Pojemność dysku ok. 160 MB (74 minuty muzyki),
- Żywotność – min. 30 lat.



Kodek ATRAC:

- Kaskadowe filtry zwierciadlane QMF (ang. *Quadrature Mirror Filters*) dzielą sygnał na trzy podpasma: 0÷5,5 kHz; 5,5÷11 kHz; 11÷22 kHz,
- Zmodyfikowana dyskretna transformacja MDCT (ang. *Modified Discrete Cosine Transform*),

Kodek Sony ATRAC

(ang. *Adaptive Transform Acoustic Coding*)



Kodek ATRAC:

- W zależności od trybu pracy przepływności od 66 kb/s (ATRAC-3 LP-4 – tryb *long-play* o obniżonej jakości) do 266 kb/s (ATRAC SP),
- Możliwość rozwijania systemu bez wprowadzania zmian w dekodерze – algorytm ATRAC nie narzuca konkretnego sposobu **alokacji bitów**. Praca kodera jest całkowicie niezależna od przyjętego modelu psychoakustycznego,
- Zmienna rozdzielczość częstotliwościowa dla różnych pasm częstotliwości,
- Zmienna rozdzielczość czasowa w zależności od parametrów statystycznych sygnału.



więcej: <http://www.minidisc.org/>

Dolby AC-3

- Standard ten opracowano w celu zastosowania w systemie wielokanałowego nagłośnienia kina cyfrowego, następnie wybrano go dla potrzeb kodowania dźwięku w HDTV.
- Wykorzystuje:
 - maskowanie jednoczesne w dziedzinie częstotliwości
 - bank filtrów oparty na transformacji TDAC (ang. *Time Domain Aliasing Cancelation*)
 - zmienną rozdzielczość czasowo-częstotliwościową.
- Koder wykorzystuje dodatkowe 16-bitowe słowo synchronizacji oraz 2 słowa 16-bitowe w celu kontroli błędów (CRC).
- Przepływność binarna kodeka AC-3 jest zawarta między 32 a 640 kbit/s w sumie na wszystkie kanały.
- Typowe całkowite opóźnienie czasowe kodowania ok. 100ms.



więcej: <http://www.dolby.com/tech/ac-3mult.html>

Kodek PASC

(ang. *Precision Adaptive Subband Coding*)



- Standard PASC został opracowany w 1988 roku przez firmę *Philips* wraz z japońskim koncernem *Matsushita* i zastosowany w magnetofonie cyfrowym DCC (ang. *Digital Compact Casette*).
- Analiza sygnału szerokopasmowego odbywa się przy pomocy 32 filtrów FIR, pasmowo-przepustowych o stałej szerokości pasma (750 Hz).
- Przepływność binarna kodeka PASC wynosi 384 kbit/s dla sygnału stereo przy jakości dźwięku płyty CD.
- W 1996 roku firma *Philips* zaprzestała produkcji magnetofonów DCC

System MUSICAM

(ang. *Masking-pattern Universal Subband Integrated Coding And Multiplexing*)



- MUSICAM został opracowany dla potrzeb cyfrowej transmisji radiowej DAB (ang. *Digital Audio Broadcasting*).
- Cyfrowy sygnał foniczny w standardzie EBU/AES (768 kbit/s) jest dzielony na 32 podpasma o szerokości 750 Hz przy pomocy wielofazowego banku filtrów.
- Zastosowano układ realizujący FFT, który oblicza co 24 ms 1024 składowych widma.
- Kodek MUSICAM wykorzystuje maskowanie jednoczesne w dziedzinie częstotliwości i jest kompatybilny z formatami ISO/MPEG (Layer I).
- Przepływność binarna kodeka MUSICAM wynosi 384 kbit/s dla sygnałów stereo przy kompaktowej jakości dźwięku.



więcej: <http://www.worlddab.org/>

Standard MPEG-1

(ang. *Moving Pictures Expert Group*)

- umożliwia kodowanie sygnałów PCM dla częstotliwości próbkowania: 32 kHz, 44,1 kHz oraz 48kHz,
- maskowanie jednoczesne w dziedzinie częstotliwości
- maskowanie w dziedzinie czasu
- zastosowanie banku 32 filtrów wielofazowych wytwarzających podpasma o stałej szerokości 750Hz

Standard MPEG-1

- MPEG umożliwia pracę z zadanyym trybem przepływności:
 - FBR – praca ze stałą przepływnością bitową
 - VBR – praca ze zmienną przepływnością bitową
- Tryby kodowania sygnału:
 - tryb monofoniczny
 - tryb stereofoniczny
 - tryb dwukanałowy (ang. *dual channel*)
 - tryb łączny dwukanałowy (ang. *joint stereo*)

Warstwy modelu MPEG-1

Warstwy modelu MPEG-1

mono
stereo

32 kHz 44,1 kHz 48 kHz



Warstwy modelu MPEG-1

- Warstwa 1 (Layer I):
 - Uproszczona wersja schematu MUSICAM.
 - Przeznaczona głównie do zastosowań takich, jak: cyfrowe nagrywanie w warunkach domowych na kasetach, twarde dyskach i dyskach magnetoptycznych.
 - Warstwa ta używa ramki o czasie trwania 8ms przy częstotliwości próbkowania 48kHz.
 - Przepływności od 32 kbit/s do 448 kbit/s

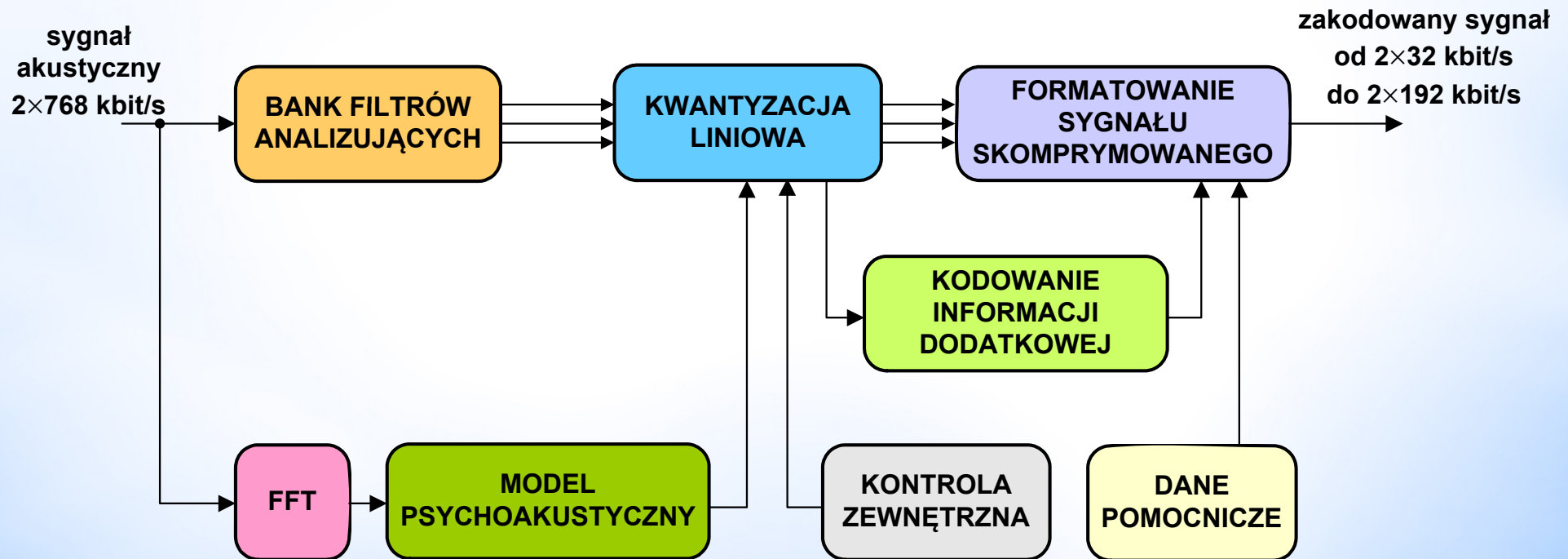
Warstwy modelu MPEG-1

- Warstwa 2 (Layer II):
 - Wprowadza dalsze udoskonalenia w stosunku do warstwy 1.
 - Podstawowa długość ramki to 24ms przy 48kHz.
 - Jest prawie identyczna ze schematem MUSICAM.
 - Ma zastosowanie głównie w profesjonalnej technice dźwiękowej, czyli wykorzystywana jest w rozgłośniach radiowych i telewizyjnych, studiach nagrań oraz w systemach multimedialnych.
 - Przepływności od 32 kbit/s do 384 kbit/s

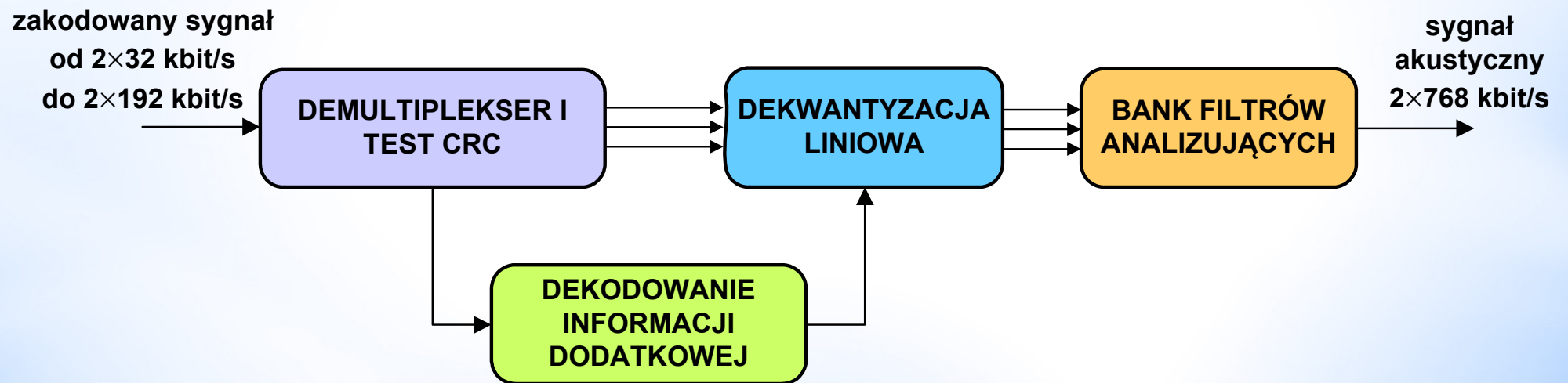
Warstwy modelu MPEG-1

- Warstwa 3 (Layer III):
 - kombinacja efektywnych modułów z kodera ASPEC i MUSICAM.
 - Hybrydowy bank filtrów wprowadza dodatkową rozdzielczość częstotliwościową.
 - Długość ramki jest identyczna z warstwą 2.
 - Używana jest tutaj niejednolita kwantyzacja, adaptacyjna segmentacja i kodowanie entropijne skwantowanych wartości w celu zwiększenia wydajności kodowania.
 - Metoda ta jest najbardziej użyteczna w telekomunikacji, szczególnie w wąskopasmowym ISDN, łączach satelitarnych i wszystkich przypadkach, gdzie wymagana jest wysoka jakość przy niskich przepływnościach.
 - Przepływności od 32 kbit/s do 320 kbit/s

Schemat kodera MPEG



Schemat dekodera MPEG



MPEG Audio - licencje

MPEG Audio nie jest typu „freeware” !!!

KODEKI SOFTWARE'owe:

- Dekodery. Freeware – brak opłat, programy komercyjne – \$0.75 za każdą sprzedaną sztukę oprogramowania albo **\$50,000** jednorazowej opłaty,
- Enkodery. \$2.5 (enc) \$5 (codec) za każdą sztukę albo **\$60,000** jednorazowej opłaty,

KODEKI HARDWARE'owe:

- Dekodery. \$0.75 za sztukę,
- Enkodery. \$2.5 (enc) \$5 (codec) za sztukę,

Standard MPEG-2

- MPEG-2 BC – kompatybilny wstecz
- Obejmuje kompresję sygnału fonicznego dla częstotliwości próbkowania: 16 kHz; 22,05 kHz; 24 kHz – telekonferencje
- Kompresja wielokanałowego sygnału fonicznego typu 3/2+1 lub 5/2+1
- Oferuje obsługę wielokanałowych ścieżek dźwiękowych w wersjach wielojęzycznych.

Standard MPEG-2

- Dodatkowe kanały C (kanał centralny), L_s (lewy surround) i R_s (prawy surround) są przesyłane w pomocniczym polu MPEG-1
- Składa się z dwóch podsystemów:
 - **LSF** (ang. *Lower Sampling Frequency*) dla niższych wartości częstotliwości próbkowania niż w standardzie MPEG-1
 - **Kodowanie wielokanałowe** – do pięciu kanałów fonicznych wraz z opcjonalnym kanałem dla niskiej częstotliwości LFE (ang. *Low Frequency Enhancement*)

Warstwy modelu MPEG-2

Warstwy modelu MPEG-2

Tryb LSF

mono
stereo

16 kHz 22,05 kHz 24 kHz



Tryb wielokanałowy

5 kanałów

32 kHz 44,1 kHz 48 kHz



Subiektywna ocena jakości dźwięku

Tabela 2. Subiektywna ocena jakości dźwięku zakodowanego perceptualnie w standardzie **MPEG-1 Layer III** (44,1 kHz, *joint-stereo*) w zależności od typu muzyki i przepływności bitowej (współczynnika kompresji).

| | PCM | MPEG-1 Layer III | | | |
|-----------|-----------------------------|------------------------------|-----------------------------|----------------------------|------------------------------|
| Chór | <u>1411 kbit/s</u> (1:1) | <u>192 kbit/s</u> (7,3:1) | <u>128 kbit/s</u> (11:1) | <u>64 kbit/s</u> (22:1) | <u>32 kbit/s</u> (44,1:1) |
| Orkiestra | <u>1411 kbit/s</u> (1:1) | <u>192 kbit/s</u> (7,3:1) | <u>128 kbit/s</u> (11:1) | <u>64 kbit/s</u> (22:1) | <u>32 kbit/s</u> (44,1:1) |
| Rock | <u>1411 kbit/s</u> (1:1) | <u>192 kbit/s</u> (7,3:1) | <u>128 kbit/s</u> (11:1) | <u>64 kbit/s</u> (22:1) | <u>32 kbit/s</u> (44,1:1) |
| Pop | <u>1411 kbit/s</u> (1:1) | <u>192 kbit/s</u> (7,3:1) | <u>128 kbit/s</u> (11:1) | <u>64 kbit/s</u> (22:1) | <u>32 kbit/s</u> (44,1:1) |
| Disco | <u>1411 kbit/s</u> (1:1) | <u>192 kbit/s</u> (7,3:1) | <u>128 kbit/s</u> (11:1) | <u>64 kbit/s</u> (22:1) | <u>32 kbit/s</u> (44,1:1) |