

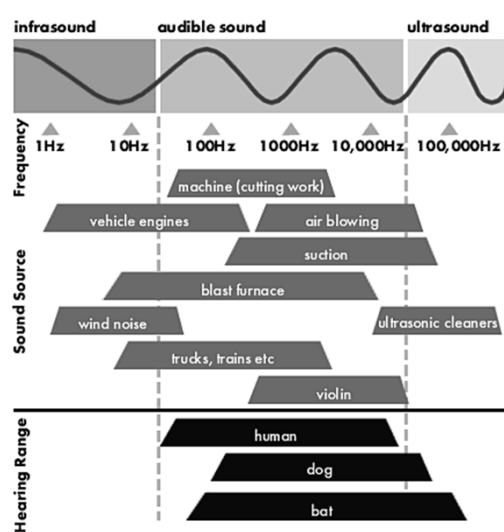
# FORMATY DŹWIĘKU

dr inż. Piotr Ody

2020-03-23

1

## Parametry słuchu



☉ zakres słyszanych przez człowieka częstotliwości: 20 Hz - 20 kHz;

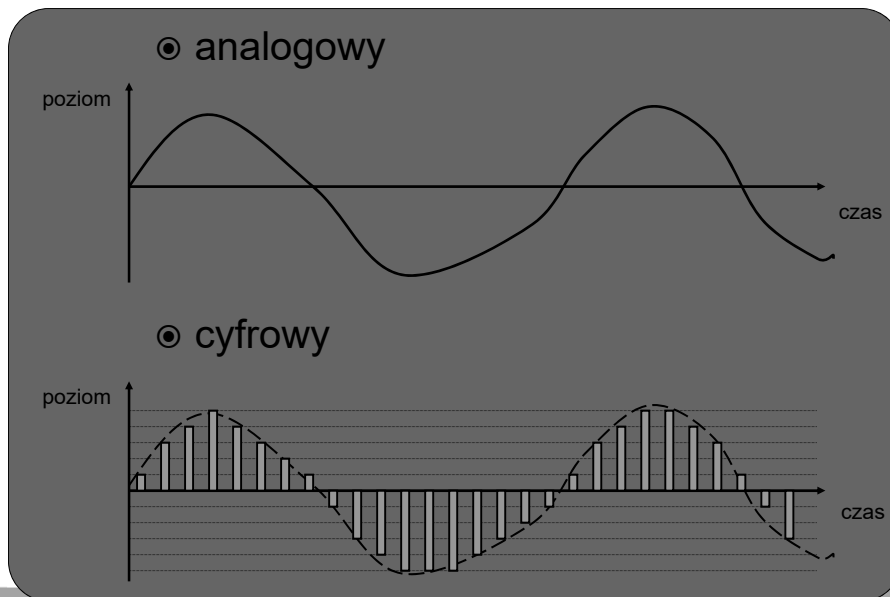
- 10 oktaw
- 20-40-80-160-320-640-1280-2560-5120-10240-20480

☉ zakres dynamiki słuchu: 130 dB

2020-03-23

2

## Sygnal foniczny

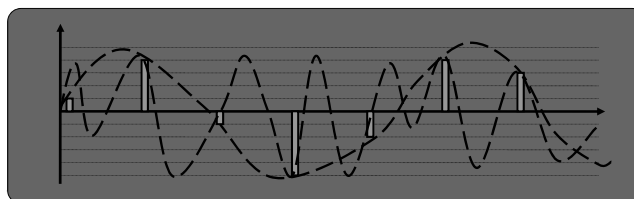


2020-03-23

3

## 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.



2020-03-23

4

## Parametry dźwięku

### ◎ częstotliwości próbkowania (w Hz)

- 8000 – jakość telefoniczna
- 11025 –  $\frac{1}{4}$  częstotliwości 44100
- 16000 – stosowana w standardzie G.722
- 22050 –  $\frac{1}{2}$  częstotliwości 44100
- 32000 – produkcja radiowa, magnetofon DAT
- **44100 – CD-Audio**
- **48000 – częstotliwość studyjna, DVD, Blu-ray**
- 88200 – brak typowych zastosowań
- 96000 – 2x48000, produkcje wysokiej jakości, DVD, Blu-ray
- 192000 – 2x96000, j.w.

2020-03-23

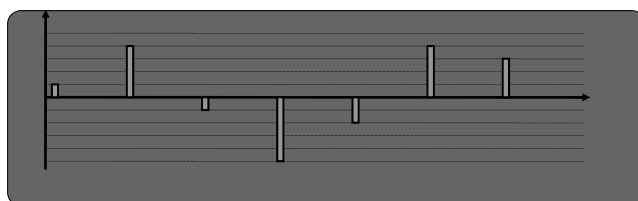
5

## 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

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



2020-03-23

6

## 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)

2020-03-23

7

## Parametry a wielkość pliku

- ◎ 1 minuta nagrania w jakości telefonicznej
  - $60 \text{ [s]} \times 8 \text{ [bit]} \times 8000 \text{ [Sa/s]} \times 1 \text{ [kanał]} = 3,66 \text{ [Mbit]} = 468,75 \text{ [kB]}$
- ◎ 1 minuta nagrania w jakości CD
  - $60 \text{ [s]} \times 16 \text{ [bit]} \times 44100 \text{ [Sa/s]} \times 2 \text{ [kanały]} = 80,75 \text{ [Mbit]} = 10,09 \text{ [MB]}$
- ◎ 1 minuta nagrania w MP3/AAC z jakością zbliżoną do CD
  - $60 \text{ [s]} \times 160 \text{ [kbit/s]} = 9600 \text{ [kbit]} = 1,17 \text{ [MB]}$
- ◎ 1 minuta nagrania 5.1 dla Blu-ray
  - $60 \text{ [s]} \times 24 \text{ [bit]} \times 192000 \text{ [Sa/s]} \times 6 \text{ [kanałów]} = 1582 \text{ [Mbit]} = 197,75 \text{ [MB]}$

2020-03-23

8

## Kompresja

---

- ◎ Metody bezstratne
  - Zakodowany strumień danych po dekompresji jest identyczny z oryginalnymi danymi przed kompresją,
- ◎ Metody stratne
  - W wyniku kompresji część danych (mniej istotnych) jest bezpowrotnie tracona, dane po dekompresji nieznacznie różnią się od oryginalnych danych przed kompresją.

2020-03-23

9

## Kompresja

---

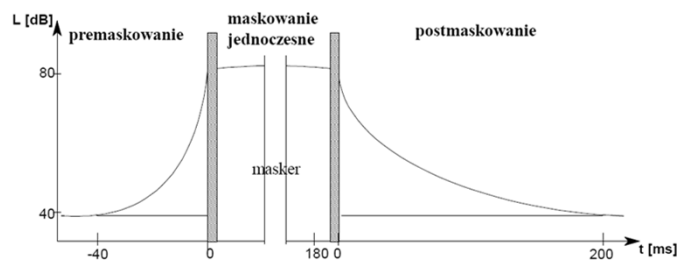
- ◎ Metody bezstratne są mało efektywne
  - typowy stopień kompresji – 10-20%
  - maksymalny stopień kompresji – ok. 40-60%
- ◎ Metody stratne charakteryzują się dużą efektywnością
  - stopień kompresji 90% przy akceptowalnej jakości dźwięku
  - wykorzystują niedoskonałości ludzkiego słuchu (kodowanie perceptualne)

2020-03-23

10

## Kodowanie perceptualne

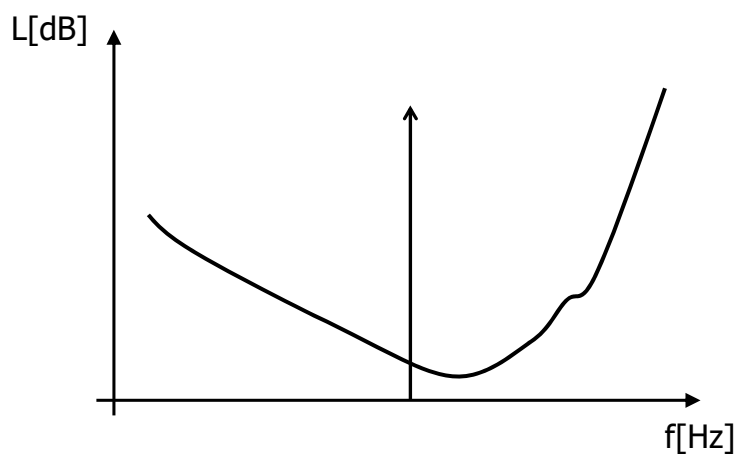
- ◉ wykorzystuje przede wszystkim zjawisko maskowania (jednoczesnego i niejednoczesnego)
  - dźwięki o niższej amplitudzie i zbliżonej częstotliwości są „zagłuszane” przez dźwięki o wyższej amplitudzie



2020-03-23

11

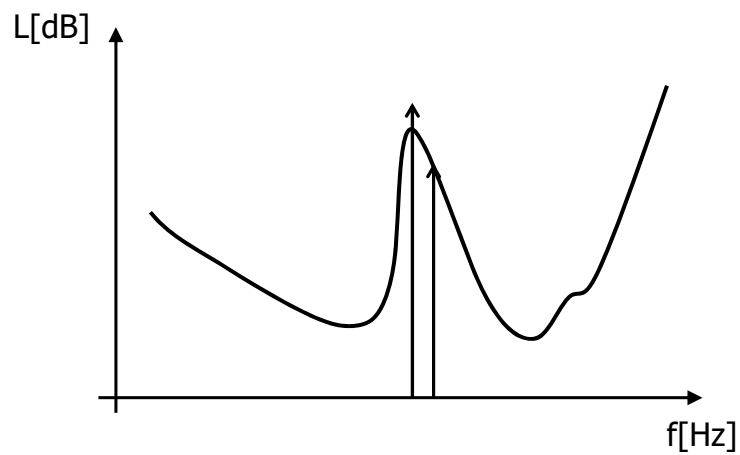
## Ilustracja maskowania



2020-03-23

13

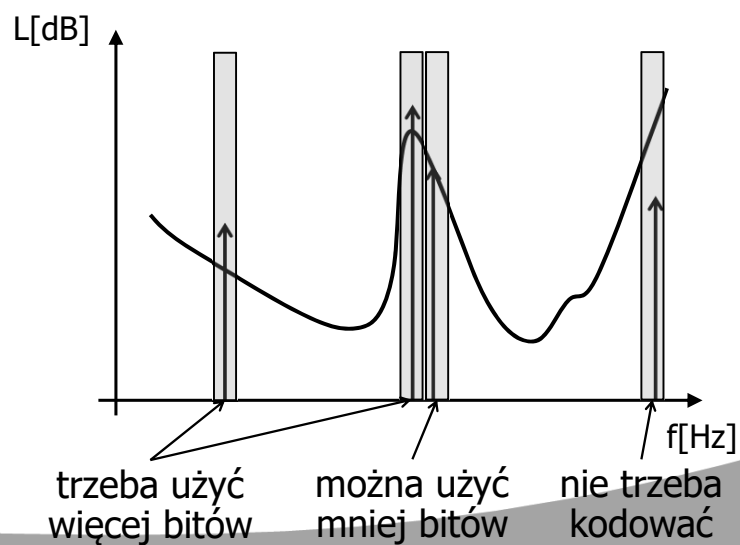
## Ilustracja maskowania



2020-03-23

14

## Ilustracja maskowania

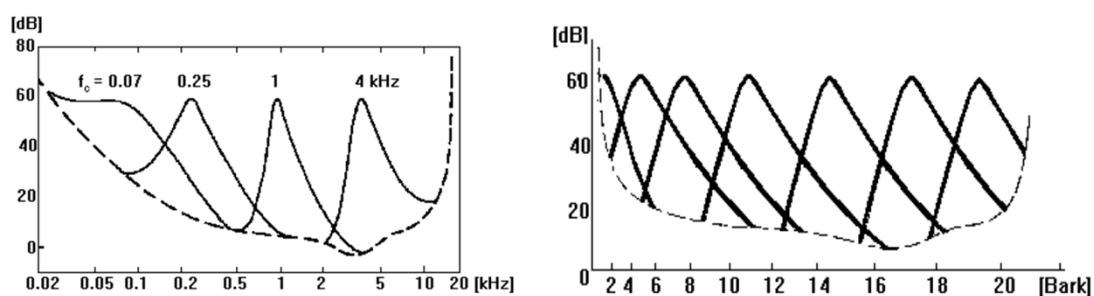


2020-03-23

15

## Maskowanie

- ⦿ maskowanie u każdego człowieka zachodzi nieco inaczej, dlatego kodeki używają uśrednionego modelu psychoakustycznego.
- ⦿ znając składowe dźwięku maskowane w poszczególnych podpasmach, kodek usuwa je z sygnału



2020-03-23

16

## Przykłady formatów

2020-03-23

17



## Formaty dźwięku – WAVE (\*.wav)

- ⦿ jeden z najpopularniejszych formatów w systemie Windows
- ⦿ **typowo** dane zapisane są w formacie PCM
  - możliwe inne formaty danych: ADPCM, u-Law, A-Law, LPC, GSM, CELP, G.721, G.723 a nawet MP3
- ⦿ obsługiwana liczba kanałów: 1, 2, 5.1
- ⦿ problem z plikami większymi od 4GB
  - rozwiązanie: format RF64
- ⦿ zastępowany przez format BWF (Broadcast Wave Format)

2020-03-23

18

## Formaty dźwięku – WAVE (\*.wav)

endian	File offset (bytes)	field name	Field Size (bytes)	
big	0	ChunkID	4	The "RIFF" chunk descriptor
little	4	ChunkSize	4	
big	8	Format	4	
big	12	Subchunk1ID	4	
little	16	Subchunk1 Size	4	The "fmt" sub-chunk
little	20	AudioFormat	2	
little	22	NumChannels	2	
little	24	SampleRate	4	
little	28	ByteRate	4	
little	32	BlockAlign	2	
little	34	BitsPerSample	2	
big	36	Subchunk2ID	4	
little	40	Subchunk2 Size	4	The "data" sub-chunk
little	44	data	Subchunk2Size	

describes the format of the sound information in the data sub-chunk

Indicates the size of the sound information and contains the raw sound data

```
00000000: 52 49 46 46 54 20 E8 12 | 57 41 56 45 66 60 74 20 | RIFFT Ć Ć WAVEfmt
00000010: 10 00 00 00 01 00 02 00 | 44 AC 00 00 10 B1 02 00 | + . 7 D~ +±7
00000020: 04 00 10 00 64 61 74 61 | 30 20 E8 12 00 00 00 00 | J + data0 Ć Ć t/
```

2020-03-23

19

## Formaty dźwięku – MPEG Layer 3 (\*.mp3)

---

- ◎ najpopularniejszy (?) format perceptualnej kompresji stratnej - a zarazem format pliku
- ◎ MPEG-1 Layer 3
  - używa bardziej skomplikowanych modeli psychoakustycznych niż poprzednie warstwy (Layer 2, Layer 1)
  - w efekcie przyjmuje się, że ucho nie dostrzeże różnicy, gdy przepływność na jeden kanał wynosić będzie 96kbit/s
  - obsługiwane przepływności: 32, 40, 48, 56, 64, 80, 96, 112, 128, 144, 160, 192, 224, 256, 320 kbit/s
  - częstotliwości próbkowania: 32, 44.1 i 48 kHz

2020-03-23

23

## Formaty dźwięku – MPEG Layer 3 (\*.mp3)

---

- ◎ MPEG-2 (2,5) Layer 3 (MPEG-2 Backward Compatible)
  - 8, 16, 24, 144 kbit/s
  - częstotliwości próbkowania 8, 11.025, 12, 16, 22.05 i 24 kHz
  - **obsługa dźwięku także w formacie 5.1**

2020-03-23

24

## Formaty dźwięku – MPEG Layer 3 (\*.mp3)

- ⊙ wykorzystywanie podobieństwa kanału lewego i prawego (np. tryb „joint stereo”) w celu poprawy wydajności kompresji
- ⊙ jakość kompresji zależy od implementacji algorytmu
  - dużo formatów pochodnych np. MP3 Pro, MP3 Surround
- ⊙ możliwość zapisu dodatkowych informacji tekstowych (ID3 tags)

2020-03-23

25

## Formaty dźwięku – Windows Media Audio (\*.wma)

- ⊙ format opracowany przez Microsoft
  - dostępne kodeki pozwalające na zapis dźwięku 5.1, kodowanie bezstratne a także kodek dostosowany do mowy
- ⊙ wykorzystanie kontenera ASF (Advanced Systems Format)
  - łatwość tworzenia streamingu
  - możliwość użycia Digital Right Management (DRM)
  - teoretycznie możliwość zawarcia dowolnego kodeka
- ⊙ przepływności od 48kbit/s do 768kbit/s (dla kompresji stratnej)
- ⊙ darmowe narzędzia do tworzenia (Windows Media Encoder, Microsoft Expression Encoder)
- ⊙ porzucony przez Microsoft...

2020-03-23

26

## Formaty dźwięku – Advanced Audio Coding (\*.aac)

- ⊙ następca MP3
- ⊙ standard opisany w MPEG-2 Part 7 i MPEG-4 Part 3
- ⊙ nie jest zachowana kompatybilność w dół z wcześniejszymi wersjami kodeków opartych na standardach MPEG
  - pozwoliło to na osiągnięcie wyższej kompresji
  - wysoka jakość dźwięku 5.1 przy przepływnościach rzędu 320-430kbit/s
- ⊙ teoretycznie pozwala na obsługę do 48 kanałów
  - w tym mono, stereo, 5.1
- ⊙ częstotliwość próbkowania do 96kHz

2020-03-23

27

## Formaty dźwięku – Advanced Audio Coding (\*.aac)

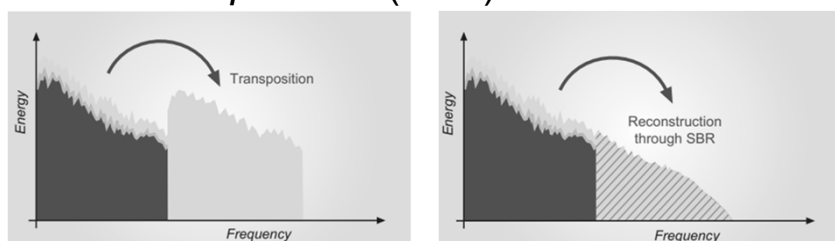
- ⊙ mnóstwo wersji opracowanych pod konkretne zastosowania
  - LC- AAC – Low Complexity AAC
  - HE-AAC – High Efficiency AAC (także jako AAC+)
    - użycie *Spectral Band Replication* i *Parametric Stereo*
  - także wersje bezstratne i dopasowane do mowy
- ⊙ format typowo wykorzystywany w urządzeniach mobilnych
  - ale także w DVB i DAB+ (HE-AAC)

2020-03-23

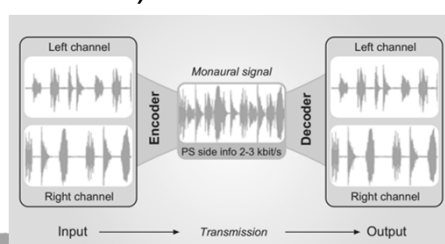
28

## Formaty dźwięku – Advanced Audio Coding (\*.aac)

### ◉ Spectral Band Replication (SBR)



### ◉ Parametric Stereo (PS)



źródło: Metzer Stefan, Moser Gerald (2006)  
MPEG-4 HE-AAC v2 - audio coding for today's  
[digital media world](#)

2020-03-23

29

## Formaty dźwięku – formaty 5.1 / 7.1

### ◉ Dolby Digital (\*.ac3) i DTS – Digital Theatre System (\*.dts)

- dwa konkurujące ze sobą formaty kompresji
- typowo formaty stratne
  - na potrzeby Blu-ray powstały wersje z kodowaniem bezstratnym
- standardowe używane na DVD-Video, Blu-ray i UHD Blu-ray
- typowe przepływności:
  - 448 kbit/s dla DD
  - 768 kbit/s dla DTS
  - należy pamiętać, że oba formaty mogą być również użyte dla dźwięku monofonicznego bądź stereofonicznego



2020-03-23

30

## Formaty dźwięku – kompresja bezstratna

- ⦿ wysoka jakość, ale często konieczne doinstalowanie dodatkowego oprogramowania
- ⦿ Free Lossless Audio Codec (\*.flac)
  - kompresja rzędu 40-50%
  - liczba kanałów: 1 do 8
    - możliwość grupowania kanałów w celu poprawy wydajności kompresji
- ⦿ Monkey's Audio (\*.ape)
  - Open Source
  - przyjmuje się, że stopień kompresji jest wyższy niż dla FLAC-a

2020-03-23

31

## Formaty dźwięku – MIDI (\*.mid)

- ⦿ w zasadzie zapis nutowy utworu
- ⦿ MIDI odtwarza nuty zakodowane w pliku korzystając z dowolnego dostępnego urządzenia dźwiękowego
- ⦿ brzmienie pliku będzie zależało od układu dźwiękowego zainstalowanego u użytkownika
  - synteza FM
  - synteza WaveTable (tablicowa)
  - synteza WaveGuide (falowodowa)

2020-03-23

32

## Przyszłość?

### ◎ MPEG-H -> 3D Audio

- kodowanie nie kanałów, a obiektów, np. dźwięk z trybu, głos komentatora itp.
- miksowanie dźwięku po stronie odbiorcy
- zwiększenie efektywności kompresji powinno umożliwić transmisję 14-18 kanałów przy przepływności rzędu 400 kbit/s
  - algorytmy bazują na AAC, ale nie będzie kompatybilności wstecz

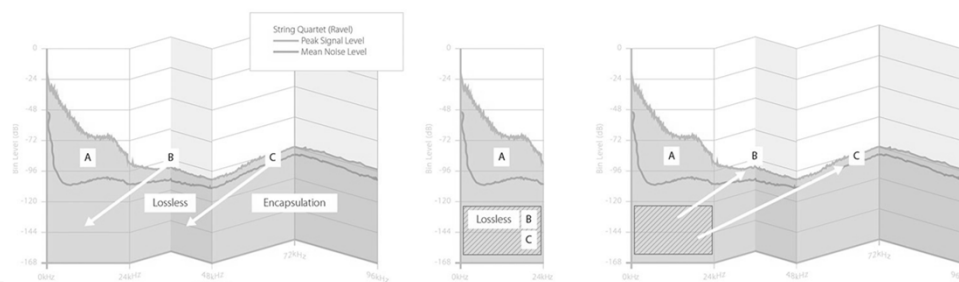
2020-03-23

33

## Przyszłość?

### ◎ MQA (Master Quality Authenticated)

- format kompresji bezstratnej (?)
- polega na upakowaniu składowych wysokoczęstotliwościowych w paśmie do 20kHz
- przeznaczony do streamingu sygnału
- kompatybilny z dotychczasowym sprzętem



2020-03-23

<http://www.digitaltrends.com/home-theater/mqa-best-high-resolution-file-format-htc>

34

## Dla zainteresowanych

---

- © John Watkinson, „The MPEG Handbook”, Focal Press, 2004.
- © <http://www.iis.fraunhofer.de/en/ff/amm.html>

2020-03-23