

SISTEMAS MULTIMÍDIA

ÁUDIO 2

Prof.: Danilo Coimbra
(coimbra.danilo@ufba.br)



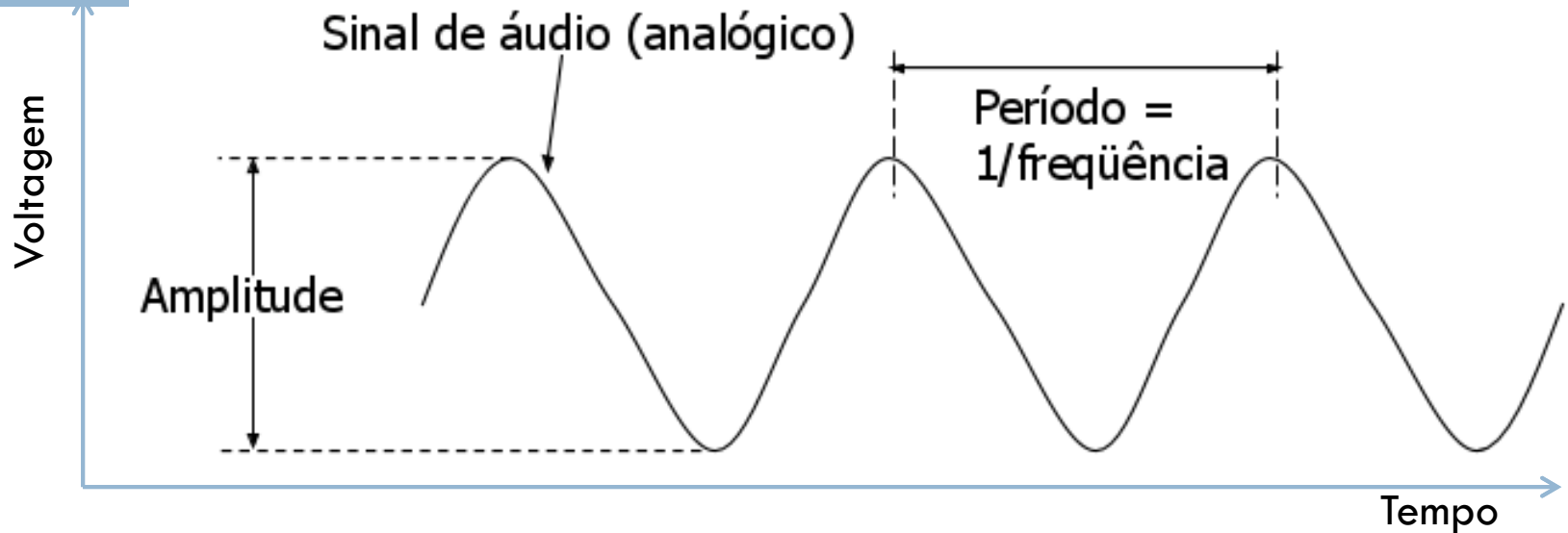
Digitalização

2

- Princípios de digitalização
 - ▣ Para que sistemas computacionais processem e comuniquem sinais de áudio, as ondas mecânicas precisam ser transformadas em elétricas e então em digitais
 - ▣ Duas transformações
- Eletrônica
 - ▣ Conversão de ondas mecânicas em sinais elétricos
- Digital
 - ▣ Conversão de sinais elétricos em bits

Digitalização

3



- **Frequência:** Número de ciclos que as partículas realizam em um segundo (medida em Hertz – Hz)
- **Amplitude:** diferença entre os máximos valores positivos e negativos do sinal de áudio (magnitude)
 - Comum ser expressa observando-se a voltagem (decibéis - dB)

Digitalização

4

- Conversão analógico-digital
 - ▣ Sinal de áudio possui duas dimensões: voltagem e tempo. Elas serão digitalizadas através de três processos:
 - ▣ **Amostragem:** realiza leituras periódicas e instantâneas da voltagem em espaços de tempo uniformes
 - ▣ **Quantização:** converte os valores analógicos amostrados em valores digitais
 - ▣ Codificação

Digitalização - Amostragem

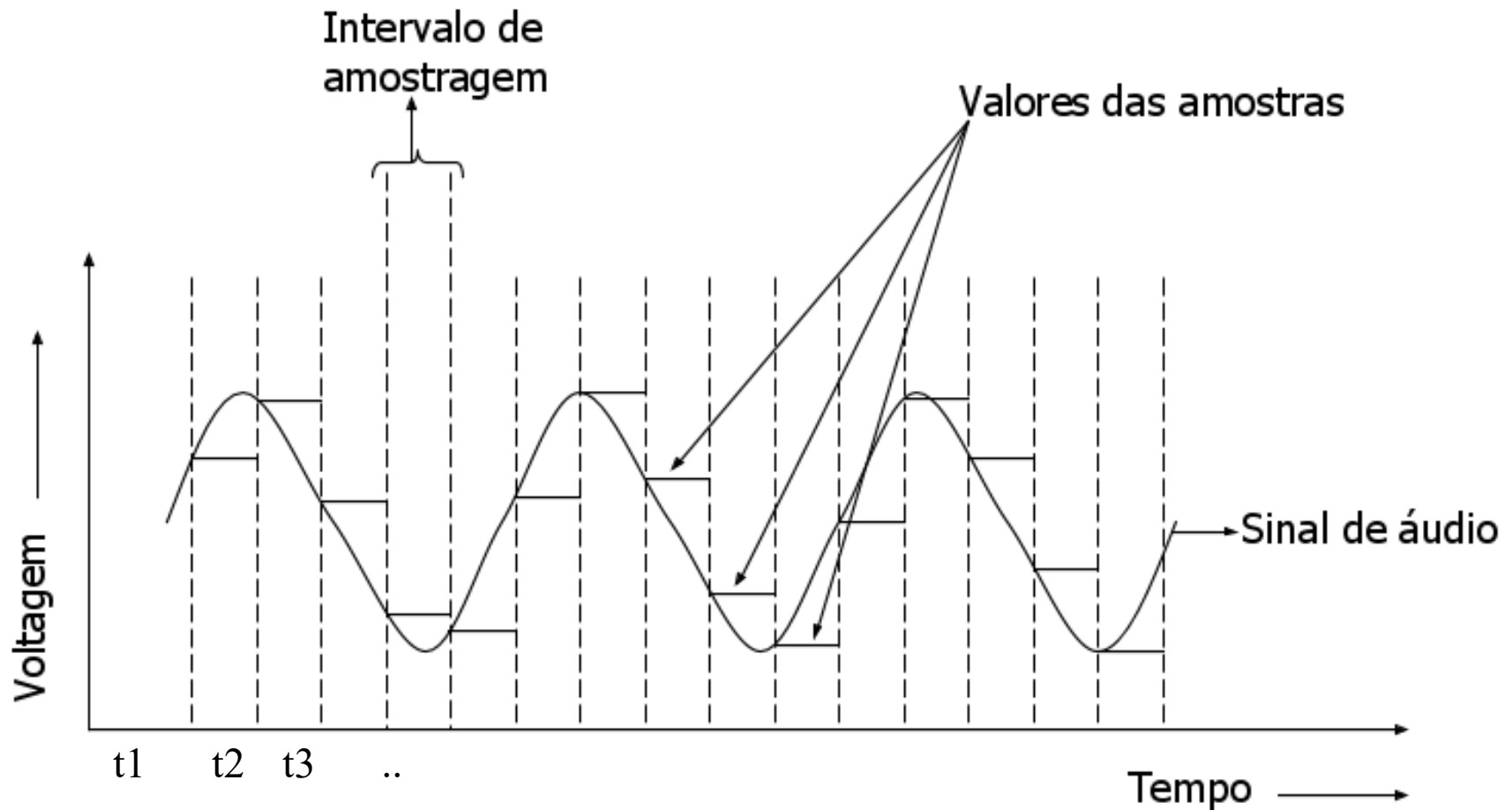
5

□ Amostragem

- O processo de amostragem nada mais é que a obtenção de amostras de um sinal contínuo, em instantes de tempo igualmente espaçados
- Nesta etapa um conjunto de valores analógicos é amostrado em intervalos temporais de periodicidade constante
 - A frequência de relógio é chamada de **taxa de amostragem** ou frequência de amostragem
 - O valor amostrado é mantido constante até o próximo intervalo
 - Dentro de cada intervalo, a amostra tem apenas um valor (discreto no tempo)

Digitalização - Amostragem

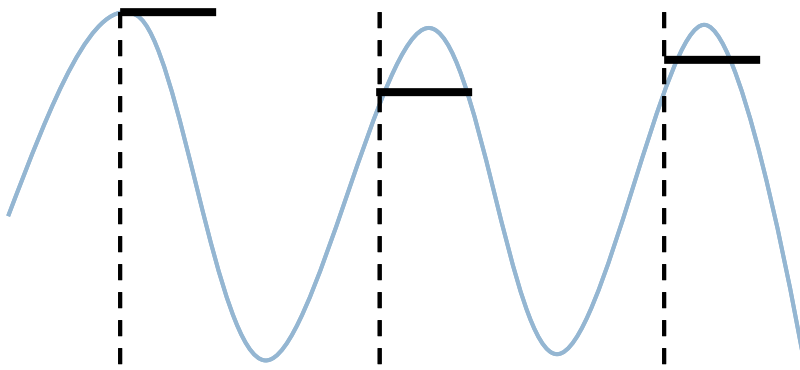
6



Digitalização - Amostragem

7

- O quanto deve ser amostrado?
 - ▣ Reconstruir exatamente o sinal
 - é necessário infinitas amostras!
 - ▣ Se utilizarmos poucas amostras?
 - O sinal ficará distorcido



- Ou seja, não é capturado corretamente o formato real do sinal

Digitalização - Amostragem

8

- ❑ O quanto deve ser amostrado?
- ❑ Uso do teorema de Nyquist
 - “Para obter uma representação precisa de um sinal analógico, sua amplitude deve ser amostrada a uma taxa mínima igual ou superior ao dobro da componente de mais alta frequência presente no sinal”
 - Taxa de Nyquist
- ❑ A partir deste teorema, Sony e Philips estabeleceram uma taxa de amostragem de 44.1 KHz quando conceberam o CD-DA. Por quê?

Digitalização - Amostragem

9

- ❑ O quanto deve ser amostrado?
- ❑ Uso do teorema de Nyquist
 - “Para obter uma representação precisa de um sinal analógico, sua amplitude deve ser amostrada a uma taxa mínima igual ou superior ao dobro da componente de mais alta frequência presente no sinal”
 - Taxa de Nyquist
 - Ex. Se a frequência mais alta do sinal é de 20KHz, para que a reconstrução seja precisa, a amostragem deve ser realizada a 40KHz, ou 40 Ksps.
 - $\text{sps} = \text{samples per second}$

Digitalização - Amostragem

10

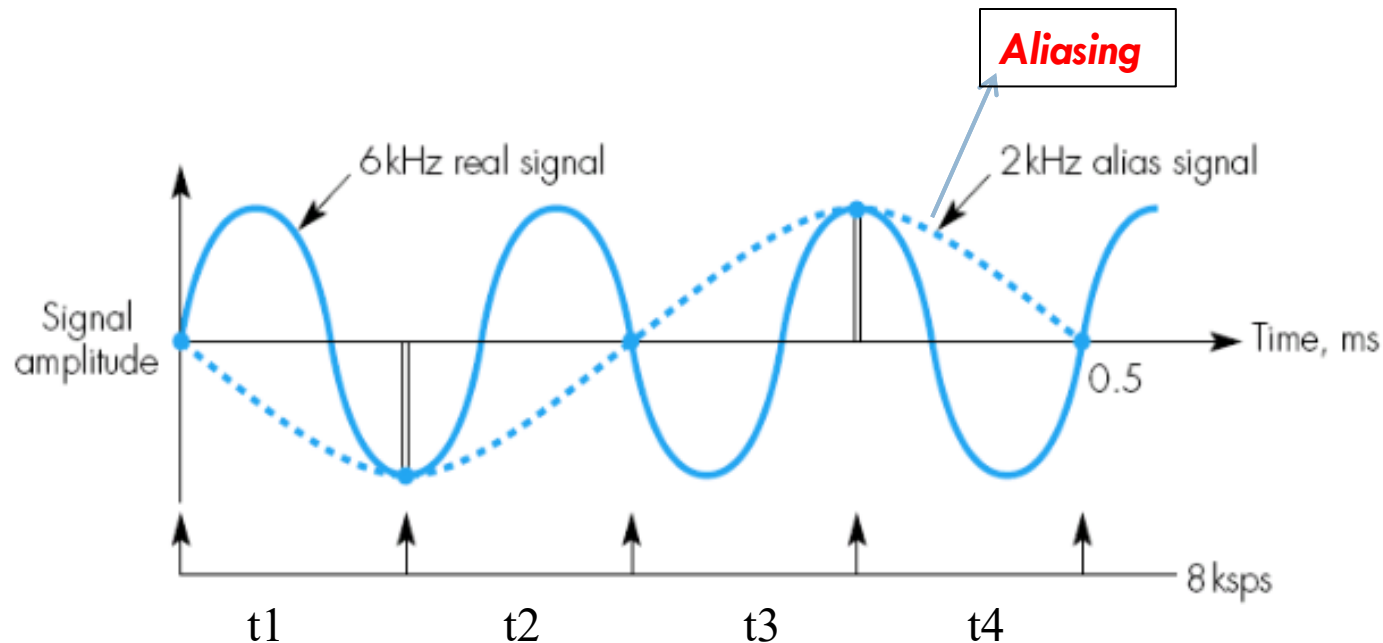
□ E quando não é respeitado o teorema de Nyquist?

□ Ex.:

■ Sinal = 6KHz

■ Amostragem a 8KHz

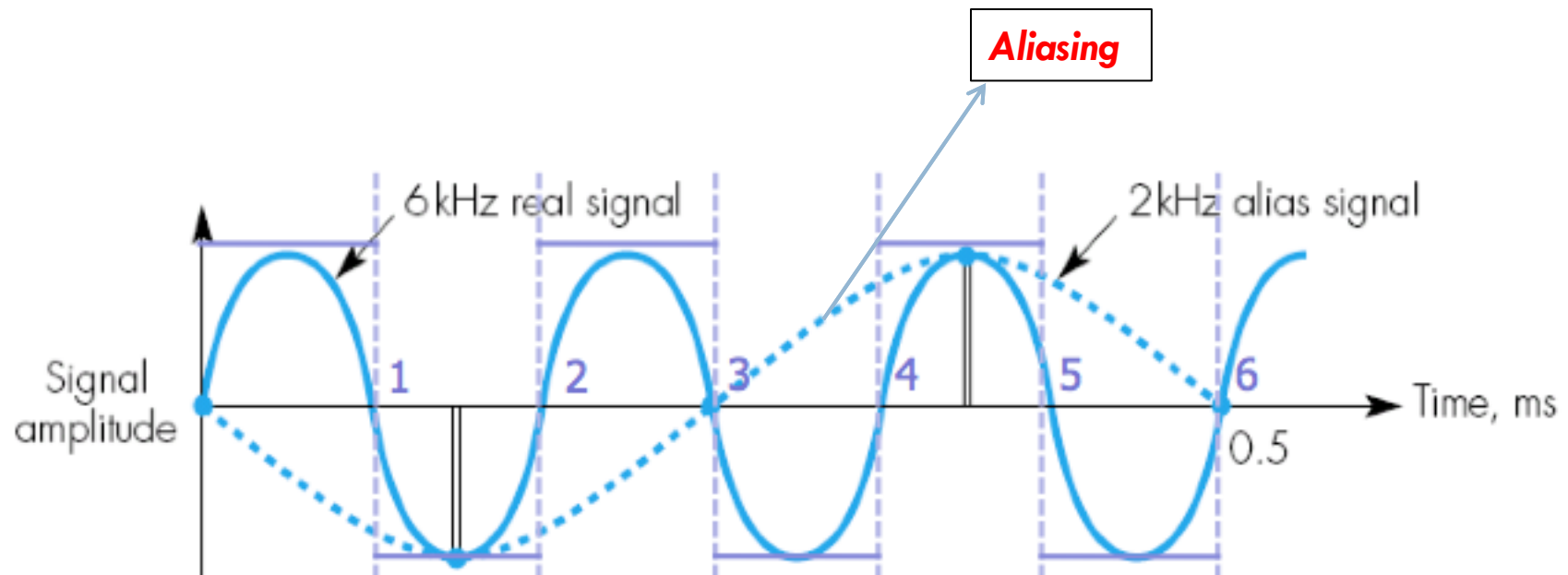
$$\left. \begin{array}{l} f' = |f - \text{amostragem}| \\ f' = |6 - 8| = 2 \text{ KHz} \end{array} \right\}$$



Digitalização - Amostragem

11

- ❑ E quando não é respeitado o teorema de Nyquist?
 - ▣ Teorema de Nyquist = $2 * 6 = 12\text{KHz}$
amostragem (mínimo)



Quantização

12

- Processo pelo qual os **valores analógicos das amostras tomadas da amplitude** do sinal são convertidos em valores digitais

- Para reconstruir exatamente o sinal:
 - ▣ Necessidade de um número infinito de bits

- Usando um número finito de bits:
 - ▣ Representa-se cada amostra através de um número correspondente de níveis discretos

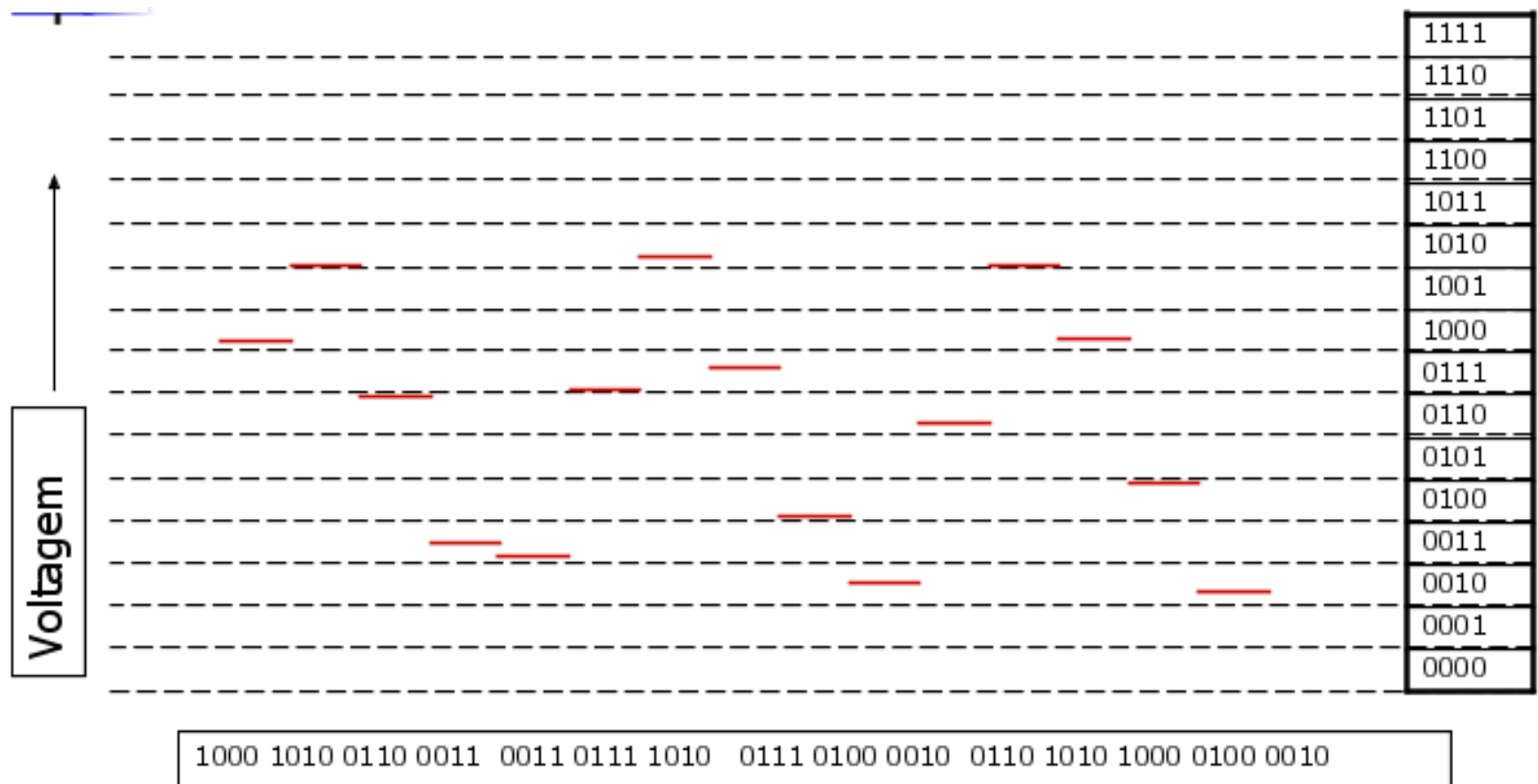
Quantização

13

- Refere-se ao número de bits usados para representar cada amostra
- ▣ Uma amostra representada por apenas um bit poderia receber apenas dois valores: "0" ou "1"
- ▣ Já uma representação com 3 bits poderia receber 8 valores diferentes ($2^3=8$): 000, 001, 010, 100, 110, 101, 011, 111
- ▣ Um CD tem uma resolução de 16 bits o que permite uma resolução binária com 65.534 valores.

Quantização

14



Quantização

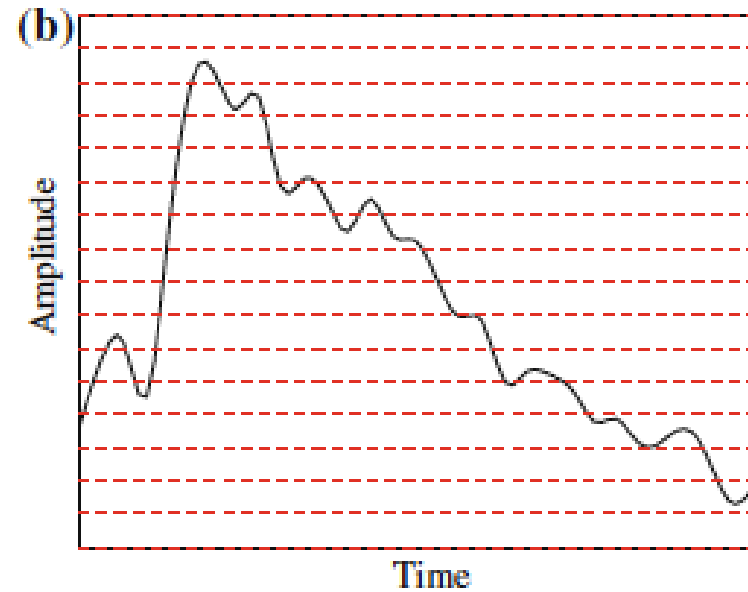
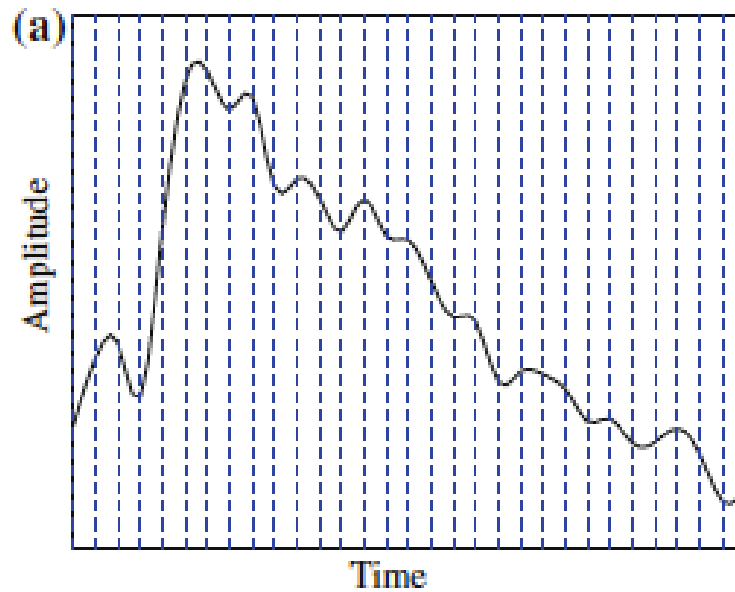
15

□ Exemplo

- ▣ Voz humana varia de 300 a 3.300 Hz
- ▣ 4KHz : frequência da linha telefônica analógica
 - Taxa da amostragem 8KHz (qualidade de telefone)
- ▣ Se considerarmos um CD de 8 bits
 - $8.000 \text{ Hz} * 8 \text{ bits} = 64.000\text{bps}$ ou 64Kbps

Amostragem e Quantização

16

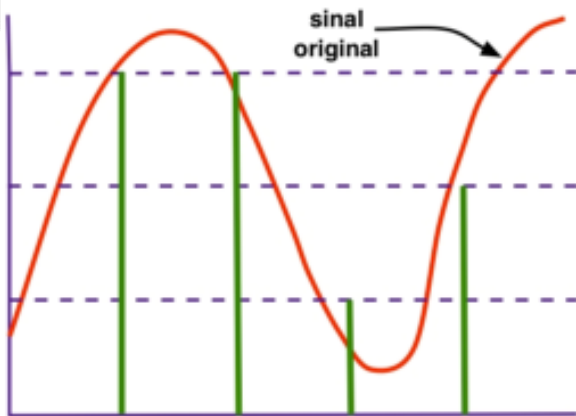


(Liu, 2004)

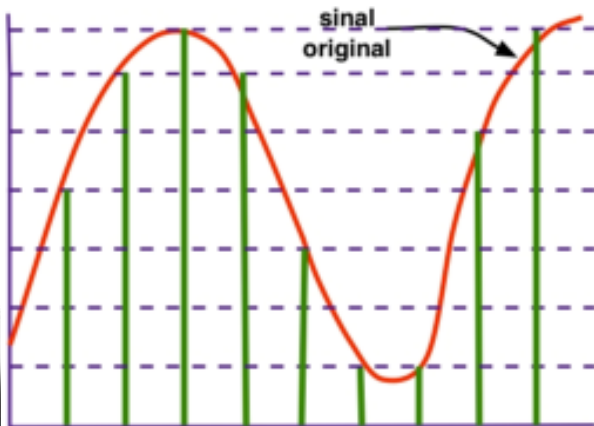
- a) amostragem:
 - ▣ Amostra o sinal analógico em função da dimensão do tempo
- b) quantização:
 - ▣ Amostra o sinal analógico em função da dimensão da amplitude

Amostragem e Quantização

17



Baixa taxa de amostragem e quantização.
Representação muito distorcida



Aumento da taxa de amostragem e quantização.
Representação Mais similar à original

Digitalização

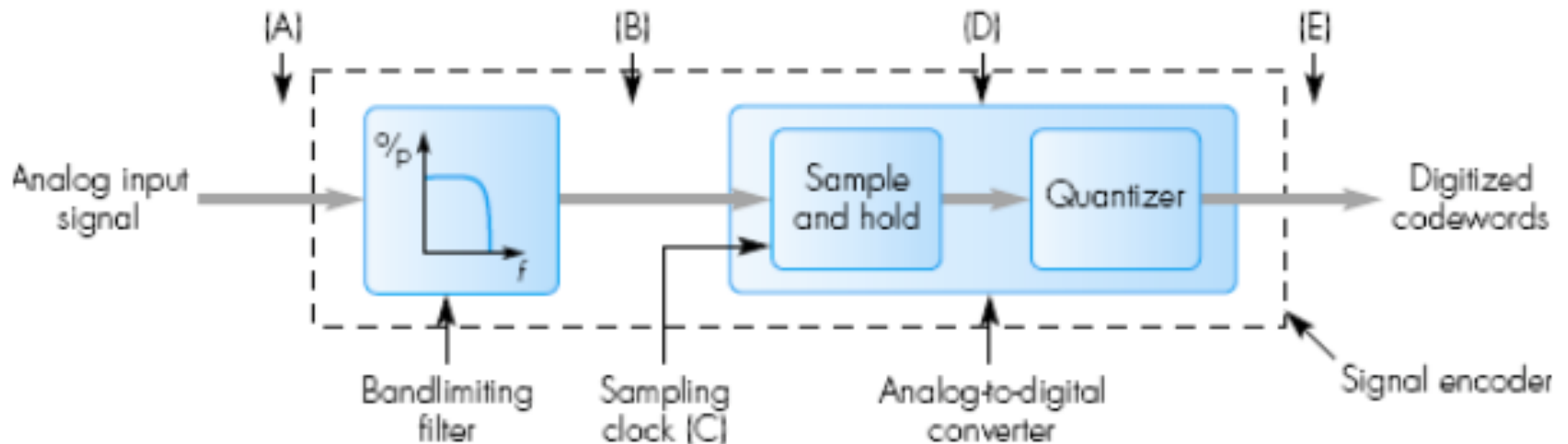
18

- Taxa comuns de amostragem
 - ▣ 8.000Hz, 16.000Hz, 22.050Hz e 44.100Hz (CD)
- Números comuns de bits por amostra
 - ▣ 4, 8, 16 e 24
- Canais de som
 - ▣ 1(mono), 2(estereo), 3, 5, 7
- Qualidade de CD
 - ▣ Amostras a 44.100Hz (44,1 KHz), 16 bits por amostra e 2 canais de som (stereo)

Digitalização

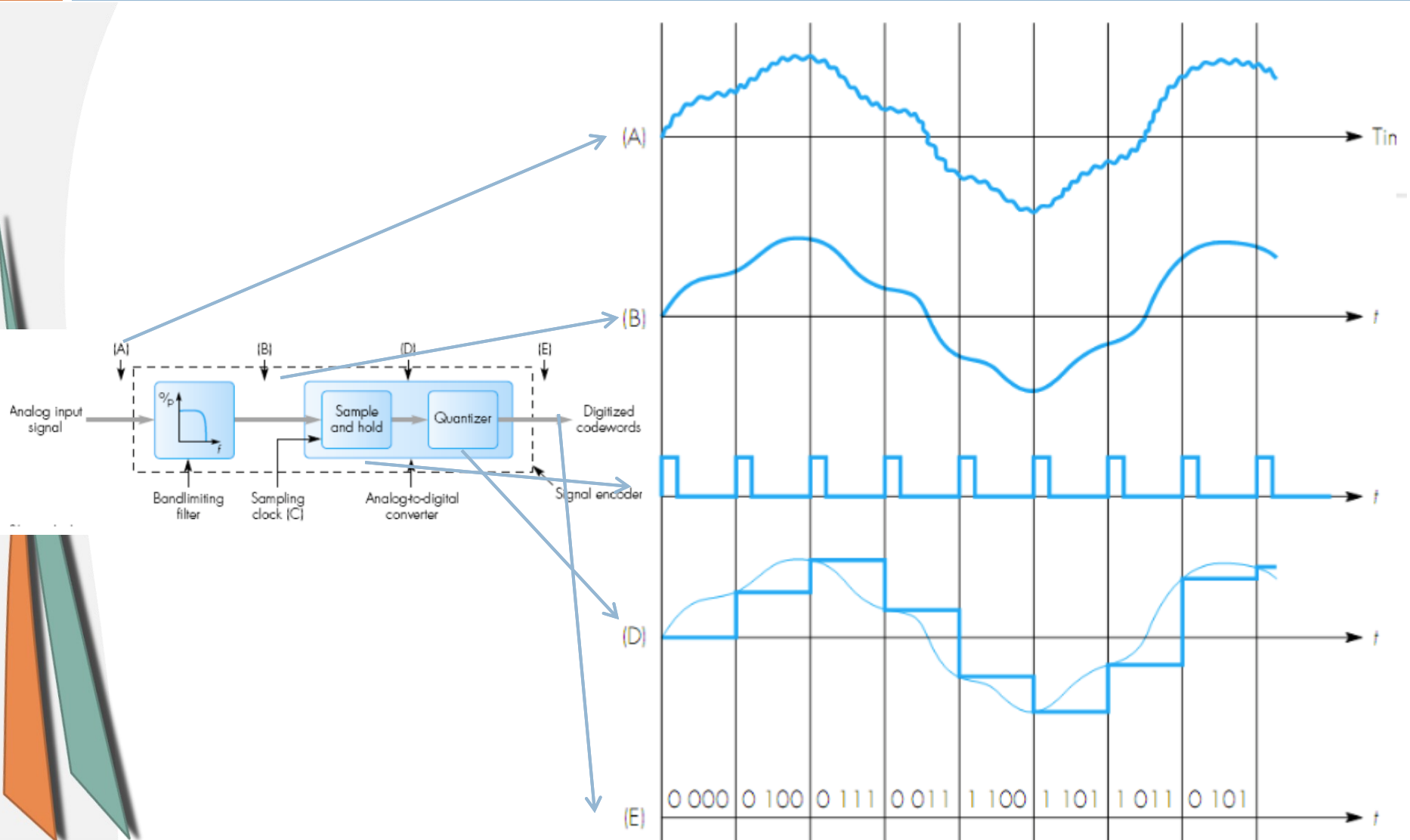
19

- Circuito que realiza amostragem e quantização
 - ▣ Conversor analógico digital (analog to digital converter - ADC)
- Caminho inverso: DAC



Digitalização

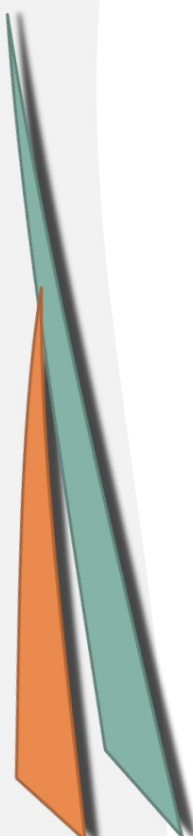
20



Digitalização

21

- Após a captura
 - os dados amostrados e quantizados devem ser “guardados” em algum formato – mídia de representação
 - Exemplo: WAV e MP3



Digitalização

22

- Na prática
- Quantos bytes serão necessários para transmitir 1 segundo de áudio, capturado com qualidade de CD?
 - ▣ Taxa de amostragem: 44100 Hz e 16 bits e som estereo

Digitalização

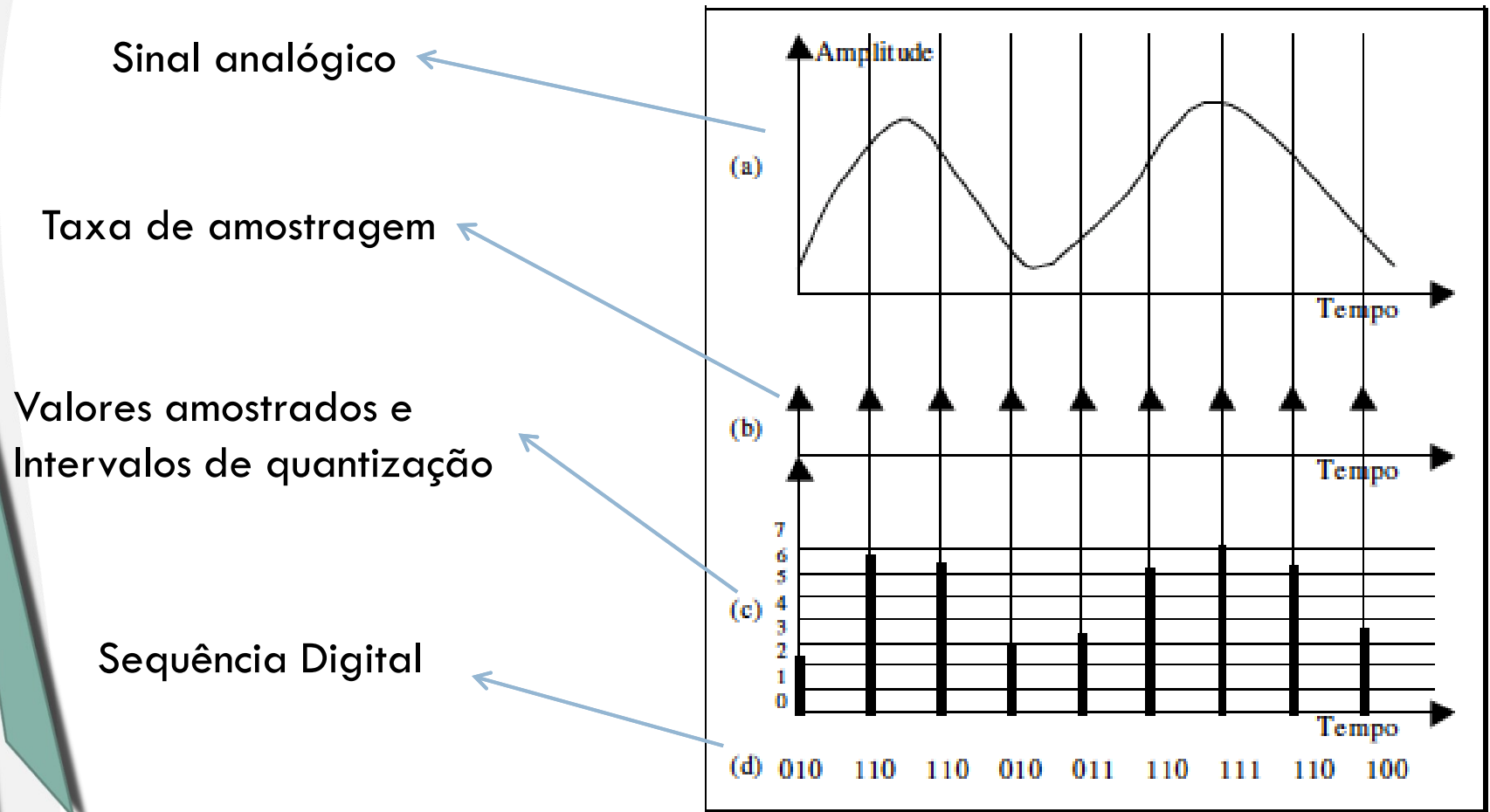
23

- Na prática
- Quantos bytes serão necessários para transmitir 1 segundo de áudio, capturado com qualidade de CD?
 - ▣ Taxa de amostragem: 44100 Hz e 16 bits e som estéreo
 - ▣ $1(\text{segundo}) * 44.100 (\text{taxa de amostragem}) * 2 (\text{bytes por amostra}) * 2 (\text{som estéreo}) = 176.400 \text{ bytes } (*8 \text{ bits})$
- 1,41Mbps! -> largura de banda (*bandwidth*)

Codificação

24

- Consiste em associar um conjunto de dígitos binários (code-word) a cada valor quantificado



Codificações

25

Características de formatos para áudio digital

FORMATO	FREQ. DE AMOSTR. (KHZ)	DIMENSÃO AMOSTRA (BITS)	QUANTIF.	CANAIS	DÉBITO BINÁRIO (KBIT/S)	MÉTODO (CODEC)	QUALIDADE
CD-DA	44,1	16	Linear	2	705/canal	PCM	Alta
DAT	48	16	Linear	2	768/canal	PCM	Alta
G.721	8	8	Linear	1	32/canal	ADPCM	Média
A-LAW/ U-LAW	8	8	Não linear (logarítmica)	1	64/canal	PCM	Telefónica
WAV	44,1	16	Linear	2	705/canal	PCM	Alta
WMA	8 - 96	24	Não uniforme	6 (5.1)	32 - 768	WMA, WMA Pro, WMA Lossless e WMA Voice	Telefónica a alta
MP3	32, 44,1 e 48	16	Não uniforme	2	32 - 320	MPEG-1 Audio Layer III	Média a alta
AAC	8 - 96	24	Não uniforme	48	De 8 até 576/canal	LC-AAC, Main, SSR-AAC	Telefónica a alta
MP4	8 - 96	24	Não uniforme	48	De 8 até 576/canal	LC-AAC, Main, SSR-AAC, HE-AAC, MP1, MP2, MP3	Telefónica a alta
OGG VORBIS	8 - 192	24	Não uniforme	255	45 - 500	Vorbis	Telefónica a alta
REAL AUDIO	8 - 96	24	Não uniforme	Variável	Variável	VSELP, LD-CELP, Dolby AC3, ATRAC3, LC-AAC, HE-AAC.	Telefónica a alta

Débito binário= Taxa de bit (*bit rate*)

(sample rate) * (bit depth) * (number of channels) = kbits per second

Codificações

26

- CD-DA (*Compact Disc-Digital Audio*) e IFF (*Audio Interchange File Format*)
 - ▣ São padrões de codificação sem qualquer compressão
- WAV e AIFF são flexíveis
 - ▣ Suporta qualquer combinação de variáveis da tabela anterior
 - ▣ Também suportam codificações com perdas

Codificação – PCM Linear

27

- Modulação por Código de Pulso (em inglês, *Pulse-code Modulation*)
 - ▣ Método mais conhecido para representar digitalmente amostras de sinais analógicos

- A técnica PCM foi patenteada, em 1939, por Alec. Reeves, quando era engenheiro da ITT¹ na França

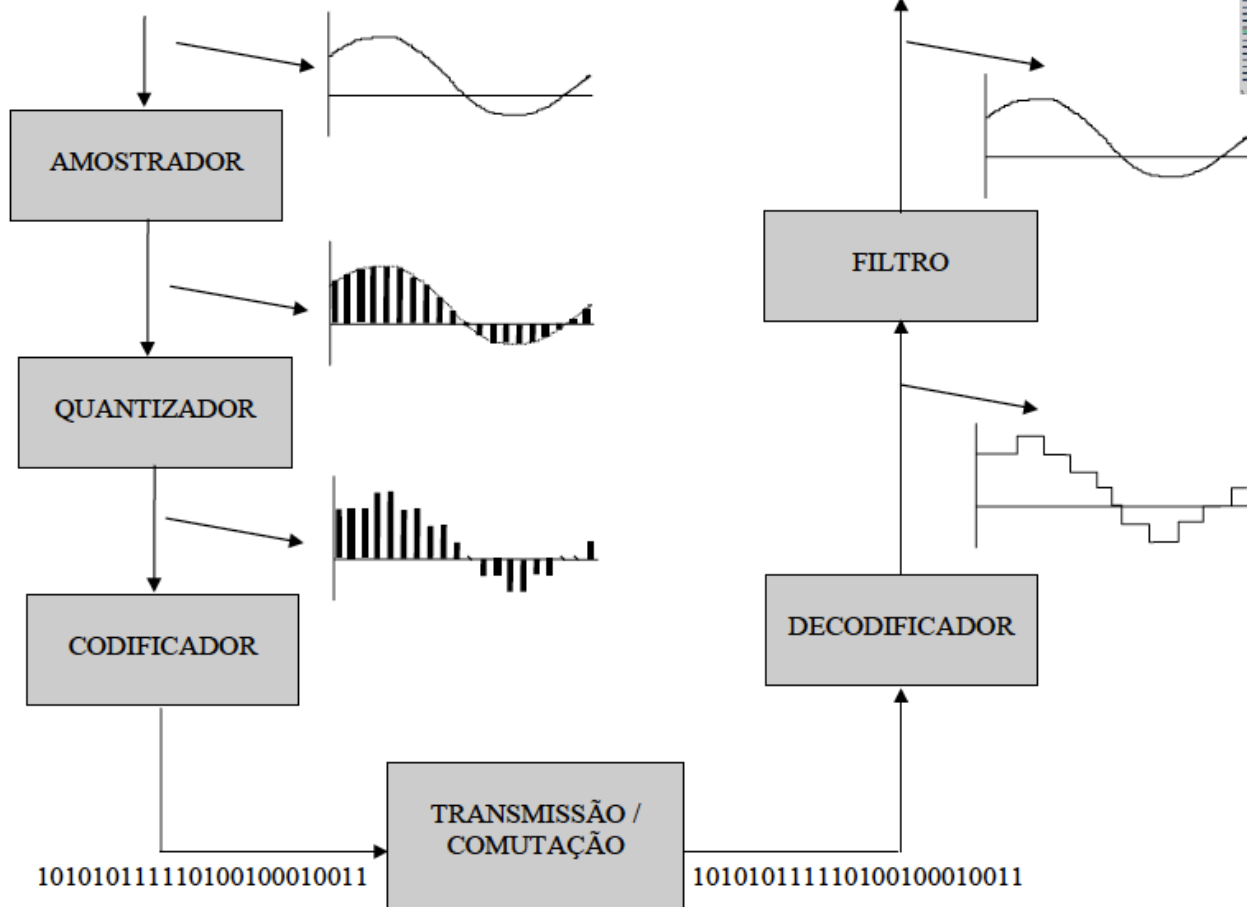
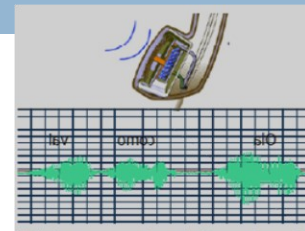
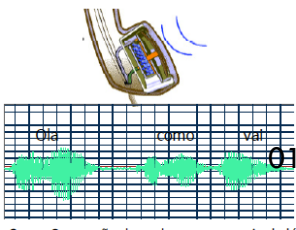
1 - ITT - International Telephone and Telegraph Company

Codificação – PCM Linear

28

- Digitalização por meio de amostragem e quantização
 - ▣ Muito utilizado em telefonia
 - ▣ 8000Hz e 8bits = 64kbps
- Amostragem em intervalos regulares
- Limitações:
 - ▣ Erro de quantização
 - ▣ *Aliasing* caso teorema de Nyquist não seja satisfeito

Codificação – PCM Linear



101010111110100100010011

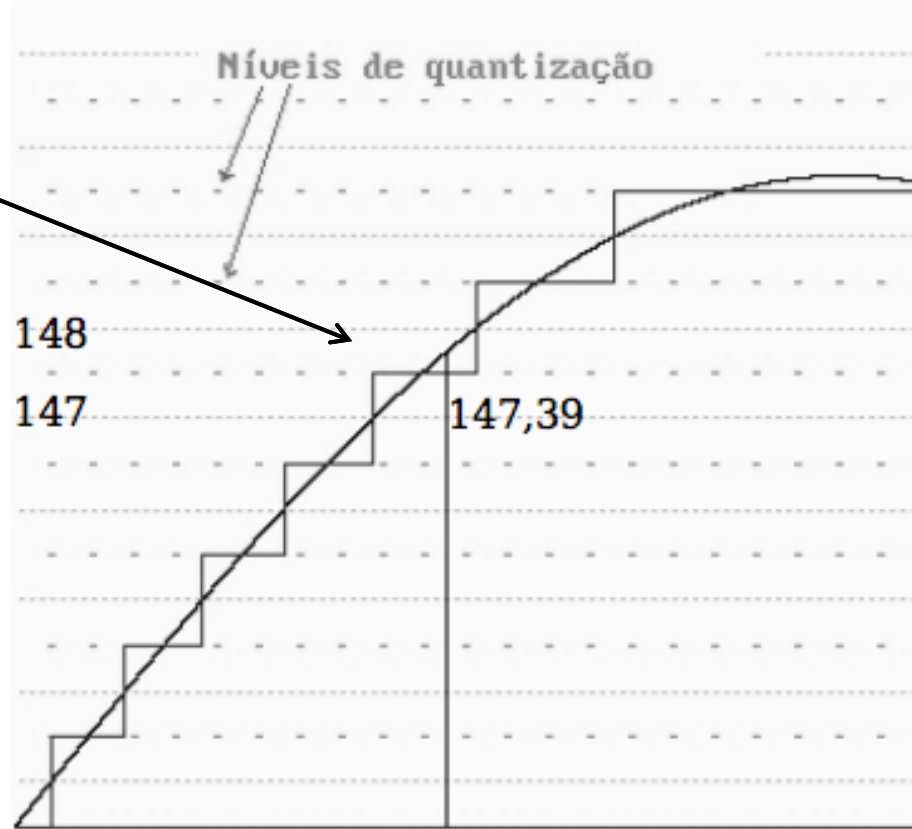
101010111110100100010011

Codificação – PCM Linear

30

Erro de quantização

- Cenário: pulsos elétricos de 0 – 255 volts
 - 8 bits (256 valores)
- Um pulso de 147,39V?
 - Não é possível representar 147,39 com 8 bits



Codificação – PCM Linear

31

Erro de quantização

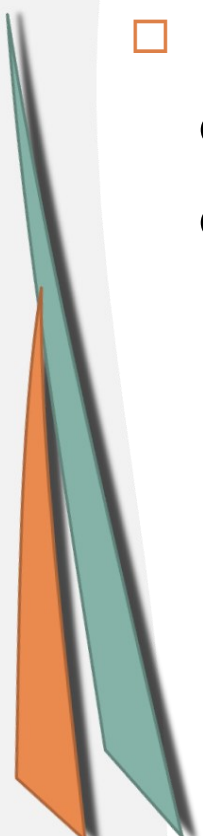
- ▣ Pequenos desvios em relação a amostra original do sinal
- ▣ Teremos um erro de $-0,39V$
 - ▣ se considerarmos nível 147
- ▣ ou $+0,61 V$
 - ▣ Se considerarmos nível 148
- ▣ Esta falta ou excesso no valor do sinal provoca o surgimento de um sinal aleatório
 - ▣ Ruído de quantização

Codificação – PCM Linear

32

Erro de quantização

- Pode ser reduzido se aumentarmos o número de níveis de quantização existentes entre os limites da variação da voltagem



Codificação – *Differential* PCM (DPCM)

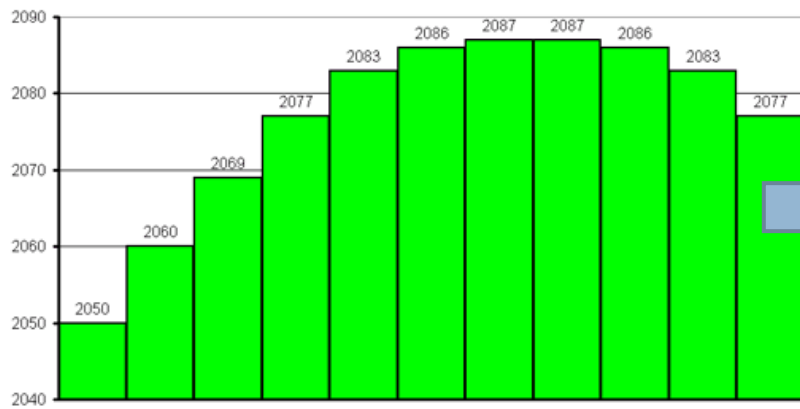
33

- Amostras adjacentes de áudio são parecidas
 - ▣ Diferenças entre os sinais são mínimas
 - ▣ Tira a redundância do sinal do PCM

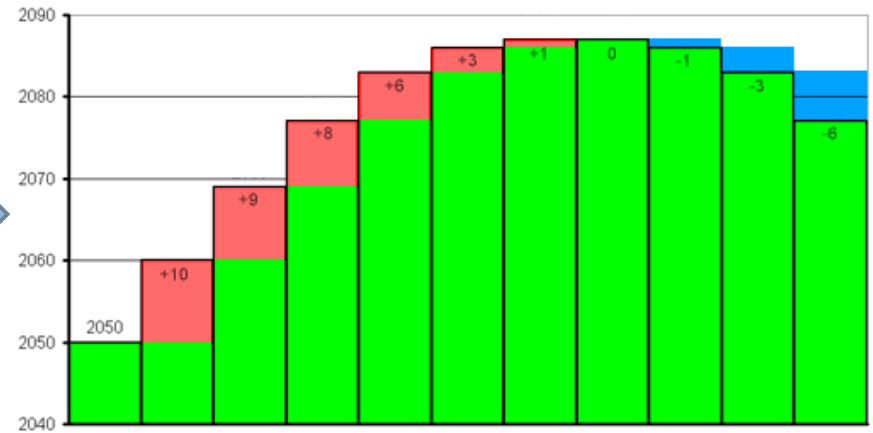
- DPCM faz previsão da amostra seguinte e codifica apenas a diferença e transmite
 - ▣ Analisa a amostra atual e a anterior
 - ▣ Mudanças bruscas entre amostras adjacentes causam distorções

Codificação – *Differential* PCM (DPCM)

34



[2050] [2060] [2069] [2077] [2083] [2086] [2087] [2087] [2086] [2083] [2077]



[2050] [10] [9] [8] [6] [3] [1] [0] [-1] [-3] [-6]

Codificação – *Differential* PCM (DPCM)

35

- Menor quantidade de bits necessários para codificar os valores
 - ▣ Normalmente, salva-se 1 bit para cada amostra
 - ▣ Sinal de voz: redução de 64 kbps (PCM) para 56 kbps (DPCM)

Codificação – ADPCM

36

Adaptative Differential Pulse Code Modulation

- Os níveis de quantização variam com o tempo
 - ▣ De modo a acompanhar a amplitude do sinal
 - ▣ Baseadas nas amostras passadas do sinal

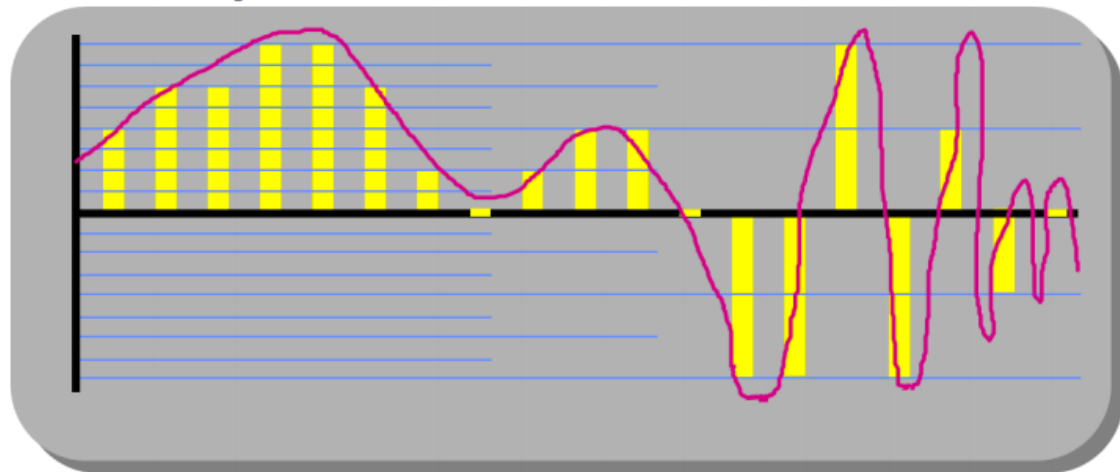
- Objetivo da técnica adaptativa:
 - ▣ Redução na faixa dinâmica do sinal para obter uma redução na taxa final de transmissão

- Sinal de voz apresenta uma correlação entre amostras sucessivas
 - ▣ Amplitude do sinal não varia muito de uma amostra pra outra

Codificação – ADPCM

37

- Ou seja, o modelo ADPCM analisa as diferenças DPCM:
 - ▣ Se a diferença entre sinais é pequena, o ADPCM aumenta o tamanho dos níveis de quantização
 - ▣ Se a diferença é grande, o ADPCM diminui os níveis de quantização
- Adapta os níveis de quantização para o tamanho da diferença dos sinais



Codificação – ADPCM

38

Em suma:

- Utilização de número variável de bits para codificação das diferenças
 - ▣ Menos bits → maiores diferenças
 - ▣ Mais bits → menores diferenças
- O ADPCM diminui a taxa de bits da voz para 32kbps, metade da modulação PCM
 - ▣ Ainda mantendo a mesma qualidade de voz (inteligibilidade)

Codificação – PCM Logarítmico

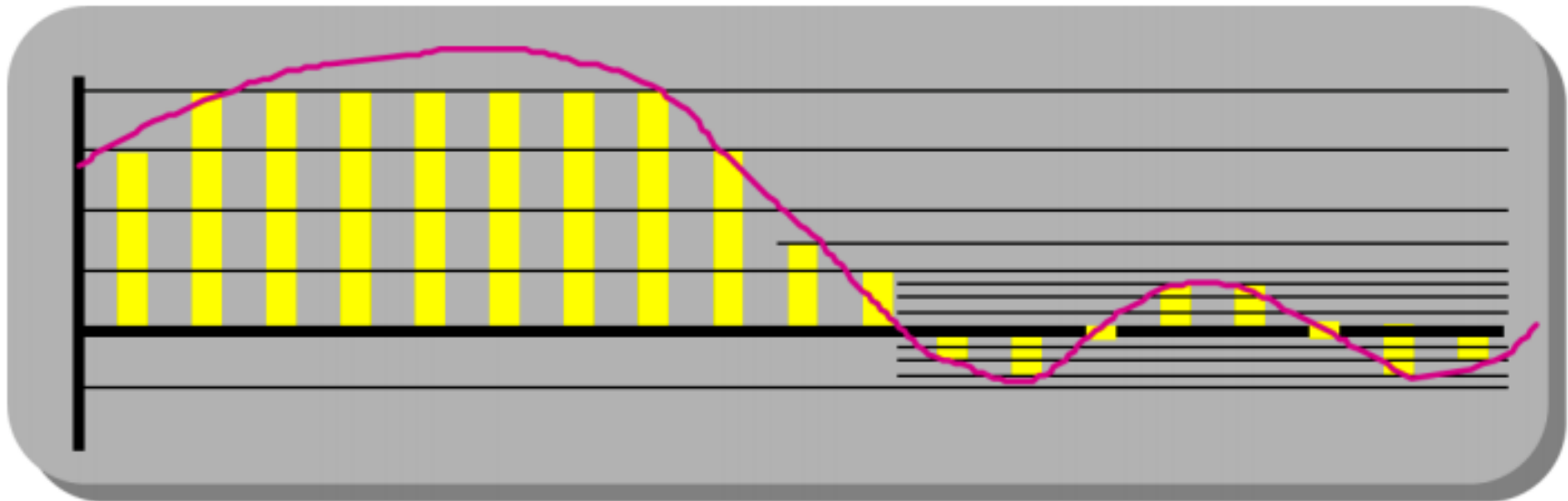
39

- ❑ Amostras de baixa amplitude são amostradas com maior precisão do que as de alta amplitude
 - ❑ Os humanos são menos sensíveis a mudanças de intensidade do som do que em relação a períodos de silêncio
- ❑ Solução: amostrar mais “densamente” as baixas amplitudes e menos “densamente” as amplitudes mais altas (**menos bits no total**)
 - ❑ Comportamento da escala logarítmica
 - ❑ Quantização não linear da amplitude do sinal

Codificação – PCM Logarítmico

40

- ❑ Quantização não linear da amplitude do sinal
 - ❑ O número de bits diminui logaritmicamente com o nível do sinal amostrado
 - ❑ Amostras de baixa amplitude são amostras com mais precisão do que as de alta amplitude



Codificação Perceptual

41

- É uma tecnologia de compressão para sinais de áudio baseada nas imperfeições da audição humana
- Técnica de compressão com perda
 - ▣ Bits decodificados não resultam numa cópia idêntica do som original antes da compressão

Codificação Perceptual

42

- Objetivo:
 - ▣ Ter um fluxo de bits decodificado **que soa exatamente** (ou o máximo possível) **do áudio original** enquanto tenta comprimir o arquivo o máximo possível

- Modelo psico-acústico
 - ▣ Sensibilidade da audição
 - ▣ Mascaramento de frequência
 - ▣ Mascaramento temporal

Codificação Perceptual

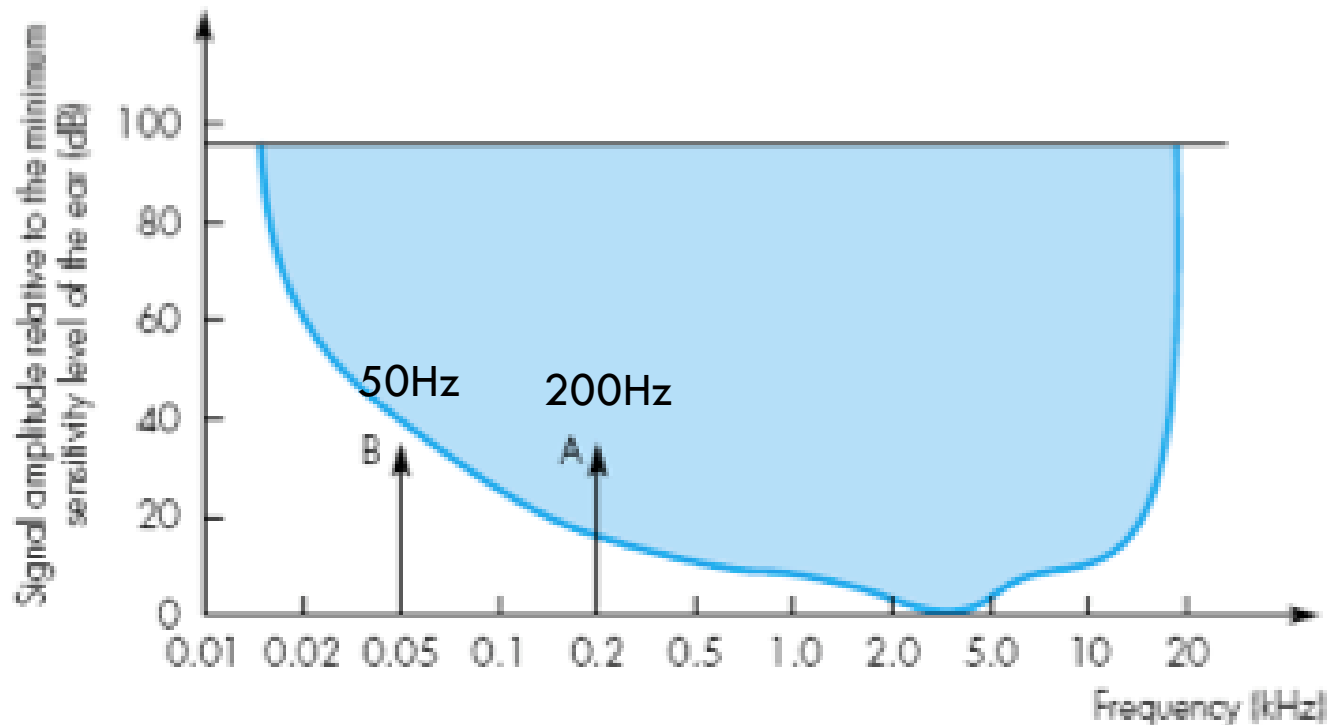
43

- Sensibilidade da audição
 - ▣ A sensibilidade do ouvido humano não é a mesma para todas as frequências
 - ▣ Intervalo entre 20Hz e 20kHz
 - Último valor vai decrescendo a medida que ficamos velhos
 - ▣ Ouvido é mais sensível para frequências entre 500Hz e 5kHz
 - Sensibilidade diminui acima e abaixo desses valores
 - Ou seja, som a 10Hz deve ser mais alto que um som de 1kHz

Codificação Perceptual

44

□ Sensibilidade da audição

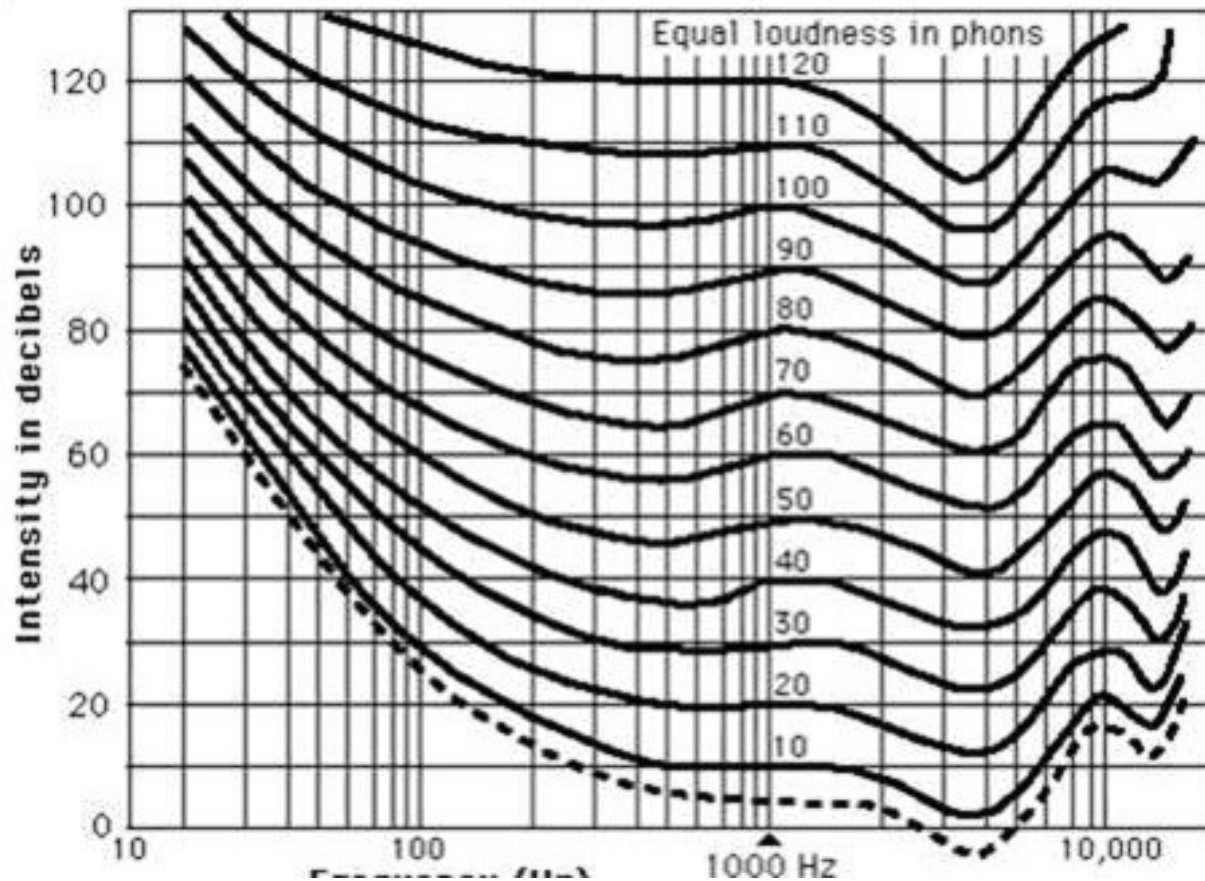


 = Hearing sensitivity of the human ear

Codificação Perceptual

45

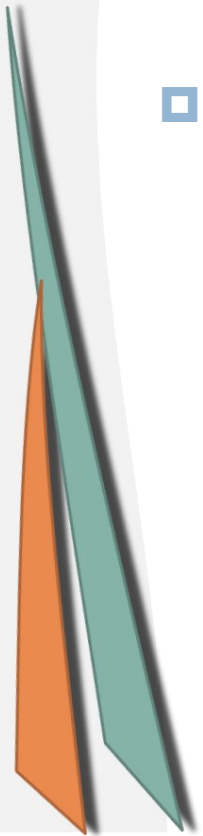
□ Sensibilidade da audição



Codificação Perceptual

46

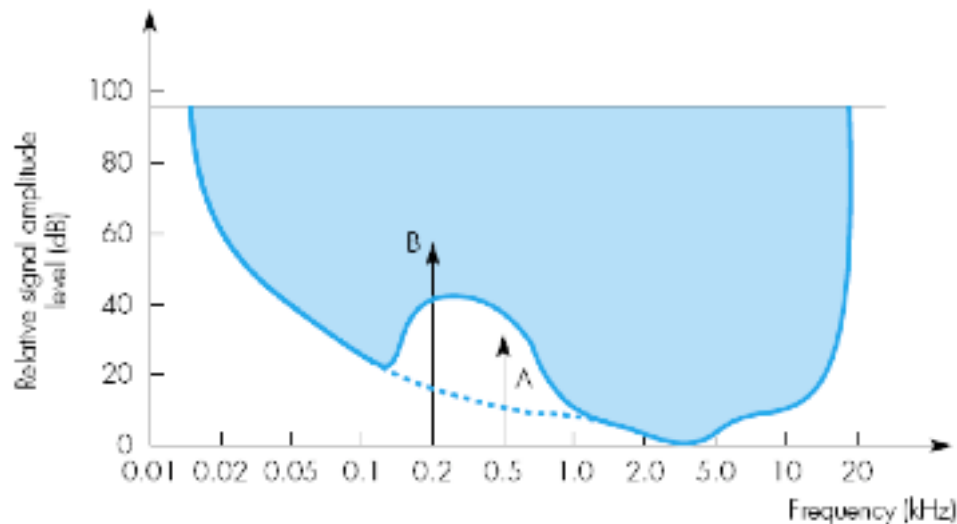
- Mascaramento de frequência
 - ▣ Um som/tom audível pode tornar outro som inaudível
 - ▣ Dois ou mais sons diferentes o estimulam simultaneamente num curto intervalo de tempo e isso consiste no apagamento parcial ou total de algumas componentes do sinal de áudio
 - Um som pode simplesmente, apagar o outro ou então aumentar o seu limiar de audição.



Codificação Perceptual

47

- Mascaramento de frequência
 - Som alto pode tornar sons baixos ináudíveis
 - Resulta no levantamento do limiar da audição



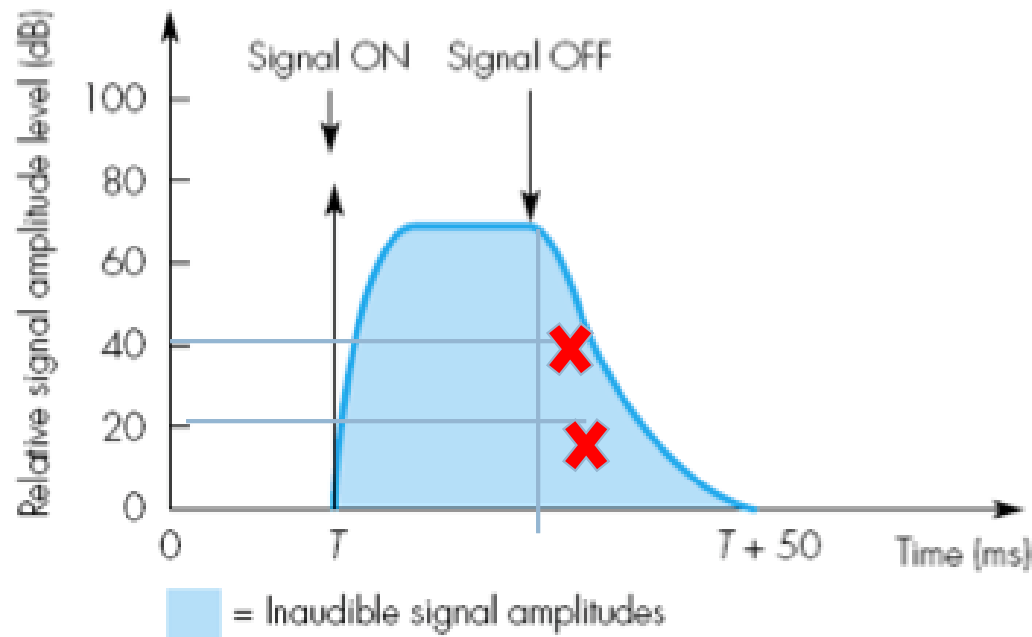
■ = Hearing sensitivity of the human ear

Codificação Perceptual

48

□ Mascaramento temporal

- Após ouvir um som alto, demorará alguns instantes (~ 10 milissegundos) até que o ouvido possa se recuperar e perceber um som mais baixo



Codificação – MP3

49

Definida pela norma *MPEG-1 Audio Layer 3*

- ▣ Utiliza um modelo psico-acústico complexo
- ▣ Elimina as frequências que o ouvido humano não consegue captar
- Estima o que se pode reduzir, **atenuando também a quantidade de informação** do sinal áudio
 - sem que se torne perceptível para o ouvido humano

Codificação – MP3

50

MPEG-1 *Audio Layer 3*

- ▣ Na prática
 - Algumas partes do áudio são mais fáceis de comprimir, nomeadamente os **momentos de silêncio ou música com apenas alguns instrumentos**
 - outras, são mais difíceis de comprimir
- ▣ Compressão com perda
 - As perdas não são perceptíveis
- ▣ Som de alta qualidade e arquivos até 12 vezes menores

Codificação – AAC

51

Advanced Audio Coding

- Melhor qualidade que o MP3 a taxas de bits similares
 - ▣ Foi criado com o objetivo de atingir maior qualidade que o MP3
 - ▣ Utiliza os mesmos padrões base de codificação que o MP3 (baseado em codificação perceptual)
 - mas usa novas ferramentas de codificação de forma a conseguir **taxas de transmissão mais baixas mantendo a qualidade**
- Ganhou popularidade com a empresa Apple, adotando como codificação padrão de seus arquivos de áudio
 - ▣ Apesar de suportarem MP3
 - Nintendo DSi e Wii, PSP (Playstation), Sony Walkman e Android

Codificação – AAC

52

Advanced Audio Coding

- ❑ Parte integrante do MPEG-2 e MPEG-4
- ❑ Melhorias:
 - ❑ Maiores taxas de amostragem (AAC: 8 a 96 kHz; MP3: 16 a 48 kHz)
 - ❑ Até 48 canais (MPEG-1 : 2 canais; MPEG-2: 5.1 canais)
 - ❑ Taxa de bits arbitrária e tamanho variável de amostragem
 - ❑ Extensível para aumentar eficiência da compressão
 - ❑ Melhor e mais simplificado banco de filtros

Codificação – AAC

53

Advanced Audio Coding

- Se comparado ao AAC a taxas superiores a 128kbps, verifica-se que o **MP3 apresenta uma qualidade semelhante ao AAC**
- A maior diferença entre eles ocorre para taxas binárias menores que 128kbps, onde o AAC tem desempenho melhor que o MP3

Codificação – FLAC

54

Free Lossless Audio Codec

- ❑ Codec para compressão de áudio **sem perda**
 - ❑ Formato de código aberto
- ❑ Taxa de compressão de 30 a 50%
 - ❑ ZIP: 10 a 20%
- ❑ Características:
 - ❑ Predição linear para gerar dados residuais
 - ❑ Codificação por carreira para blocos de amostras idênticas (e.g. passagens silenciosas)
 - ❑ Possibilita fazer decodificação rápida e *streaming*

Codificação – Vorbis

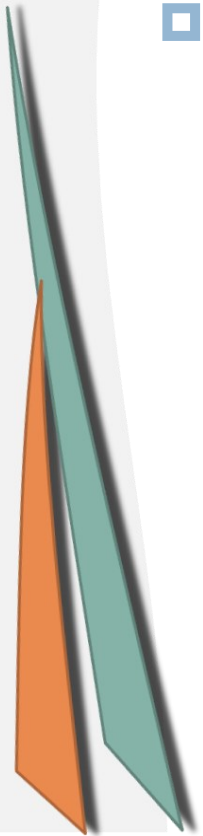
55

- ❑ Codec *lossy* (com perda) em código aberto iniciado em 1993
- ❑ Usado em conjunto com o *container* OGG
- ❑ Taxa de amostragem de 8 kHz até 192 kHz
- ❑ Até 255 canais de áudio
- ❑ Baseado na Transformada Discreta do Cosseno Modificada (MDCT)
 - ▣ Conversão dos dados do domínio temporal para o domínio da frequência
 - ▣ Quantização e codificação por entropia

Codificação – Vorbis

56

- Muito utilizado para serviços de streaming de áudio/música de alta qualidade e definição
 - Spotify



Áudio Binaural

57

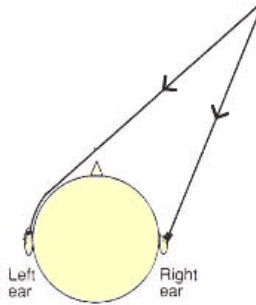
- Técnica que permite uma sensação de imersão muito maior no áudio que está sendo transmitido
 - ▣ Conhecida como áudio 3D
 - ▣ Tecnologia antiga
- Utilizando um molde de cabeça, é colocado dois microfones nas orelhas que transferem o som que nos rodeia
 - ▣ Capta a localização e a distância que está de cada ouvido



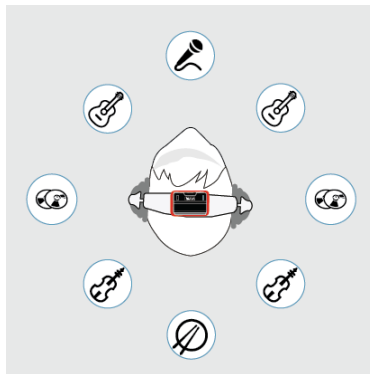
Áudio Binaural

58

- Como temos uma cabeça entre nossos dois ouvidos, conseguimos distinguir de onde vem o som pela microdiferença entre a captação de um ouvido e de outro



- Ao ser transferido para o computador esse áudio é editado para ser ouvido em 360°



Áudio Binaural

59

□ Exemplos

- Créditos do filme Monstros S.A.
- Lou Reed gravou três albuns com essa técnica (fim da década de 70)
 - “Street Hassle”, “Live: Take No Prisoners” e “The Bells”
- Pink Floyd (1983)
 - The Final Cut: explosão 3D na musica “Get Your Filthy Hands Off My Desert”
- Pearl Jam (2000)
 - Música “of the Girl”, trilha do filme “Os fracos não tem vez”.

Referências

60

- ▣ Li, Z.-N.; Drew, M.S. Fundamentals of Multimedia. Pearson/Prentice Hall. 2004. ISBN-10: 0-13-061872-1
- ▣ Luther, A. C. Using Digital Video. AP Professional, 1995.
- ▣ Halsall, F. Multimedia Communications: Applications, Networks, Protocols, and Standards, Addison-Wesley Publishing, 2001. ISBN: 0201398184. Capítulos 2 e 4.
- ▣ Ribeiro N. Multimédia e tecnologias interativas. FCA-Editora Informática; 2012.