



SCC0661 – Multimídia e Hiperemídia

Prof.: Dr. Marcelo Manzato

(mmanzato@icmc.usp.br)

Aula 3 - Áudio

Instituto de Ciências Matemáticas e de Computação - ICMC
Sala 3-160



Sumário

- 1 – Características do Som.
- 2 – Digitalização.
- 3 – Compressão de Áudio.



1. Características do Som

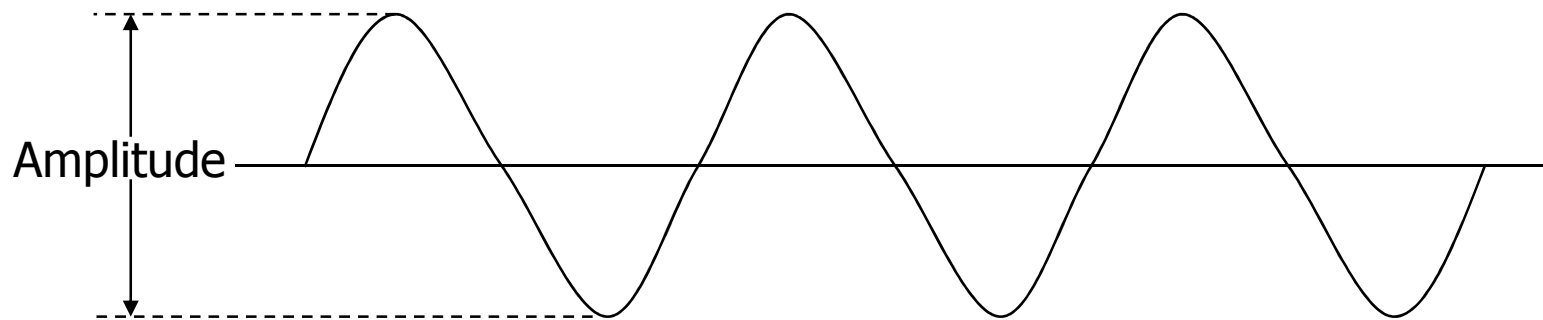


1.1 - O quê é som?

- Som é um fenômeno físico produzido por variações (vibrações) na pressão do ar.
 - Cordas de violino, bater palmas, cordas vocais, ...
- Com as variações
 - as moléculas vizinhas vibram no ar criando um variação de pressão no ar à volta.
 - Essa alteração entre altas pressões e baixas pressões propaga-se no ar, em todas as direções, como uma onda (mecânica).

1.2 - Características físicas do som.

- Som é uma onda mecânica.
 - Possui alguns aspectos, entre eles: amplitude e frequência.



- Amplitude -> Intensidade
 - Está relacionada ao volume do som. Quanto maior a amplitude, mais alto ouvimos o som.

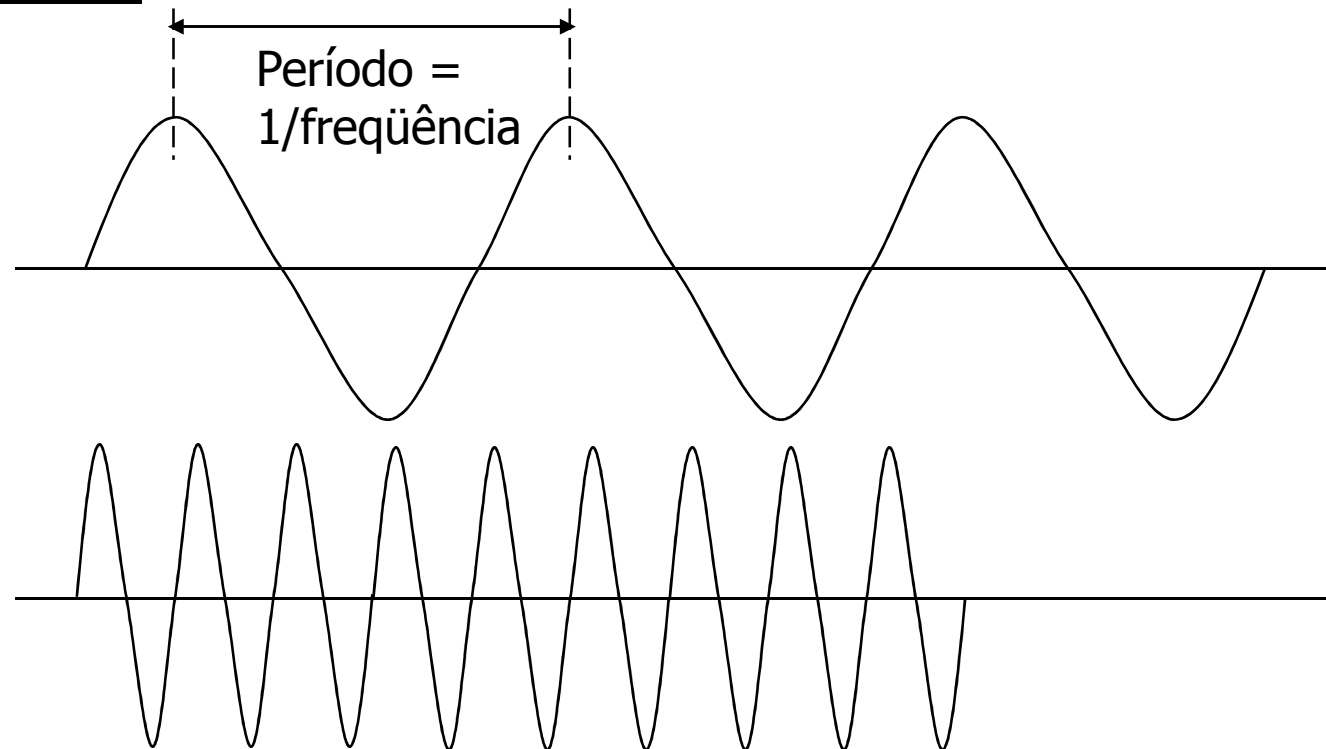
1.2 - Características físicas do som.

- Amplitude -> Intensidade
 - Medida em decibéis (dB).

Intensidade	Exemplos típicos
0dB	Limite da audição
25dB	Estúdio de gravação
50dB	Escritório
70dB	Conversação típica
90dB	<i>Home audio</i>
120dB	Limiar da dor
140dB	Show de rock

1.2 - Características físicas do som.

Frequência



- Frequência determina altura do som (altura \neq volume).
 - Frequências altas = altura maior = sons agudos.
 - Frequências baixas = altura menor = sons graves.

1.2 - Características físicas do som.

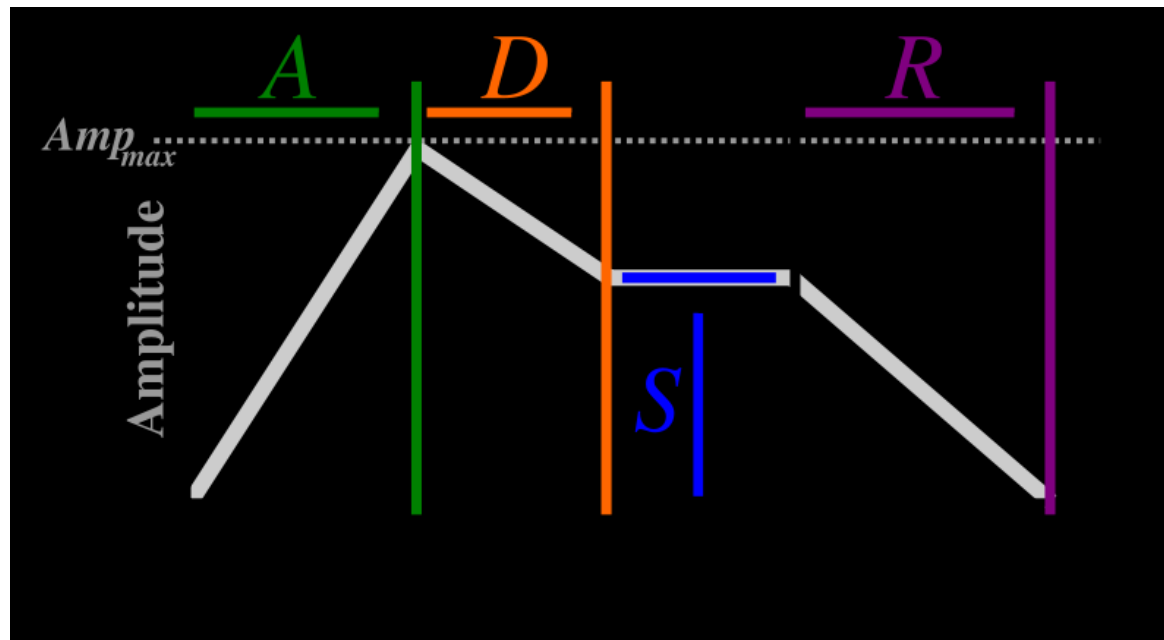


Categoria	Intervalo de Frequência
Infra-som	0 - 20 Hz
Som Audível	20 Hz - 20 KHz
Ultra-som	20 KHz - 1GHz
Hipersom	1 GHz - 10 GHz

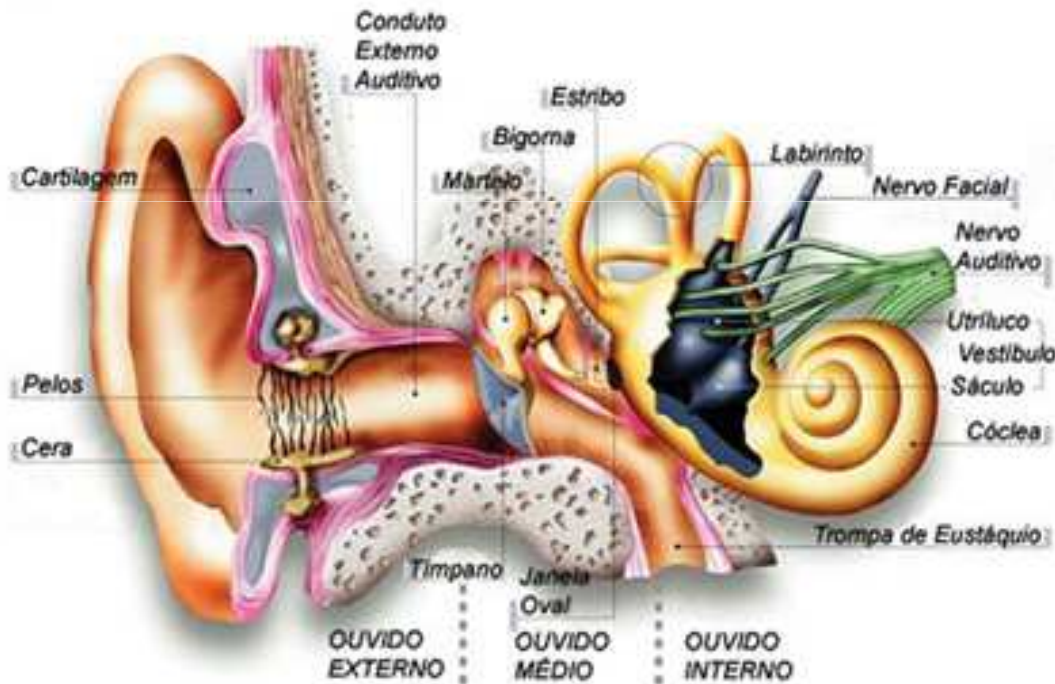
1.2 Características físicas do som

- Envelope

- Ataque (Attack), Decaimento (Decay), Sustentação (sustain) e Libertação (release)



1.3 – Como ouvimos sons?



- As ondas sonoras atingem o tímpano.
- O tímpano faz os ossos do ouvido médio vibrarem.
- Essas vibrações são convertidas em impulsos nervosos que são transmitidos, via o nervo auditivo, para o cérebro.
- Quando esses impulsos chegam ao cérebro, "ouvimos" o som!



1.3 – Como ouvimos sons?

- Assim, o ouvido funciona como um sensor ou transdutor que converte sons em estímulos nervosos que podem ser interpretados pelo cérebro.



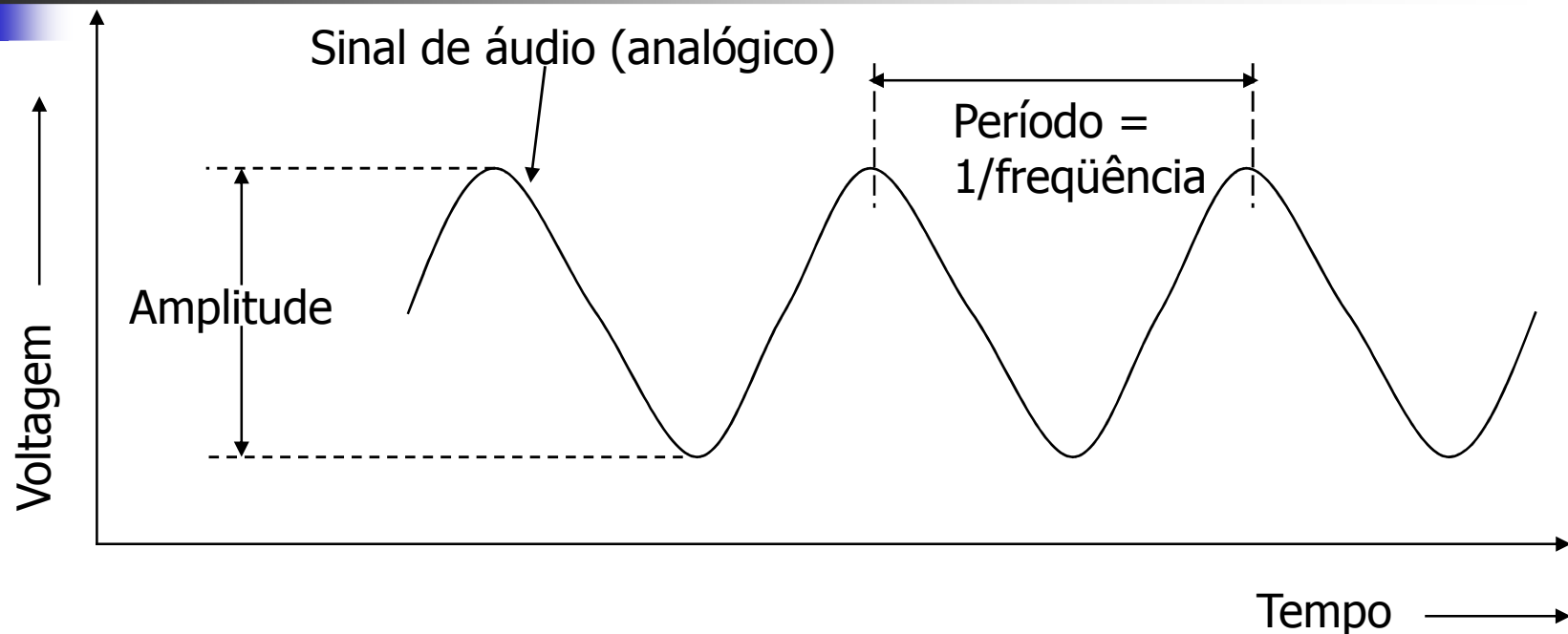
2. Digitalização



2.1 - Princípios de Digitalização.

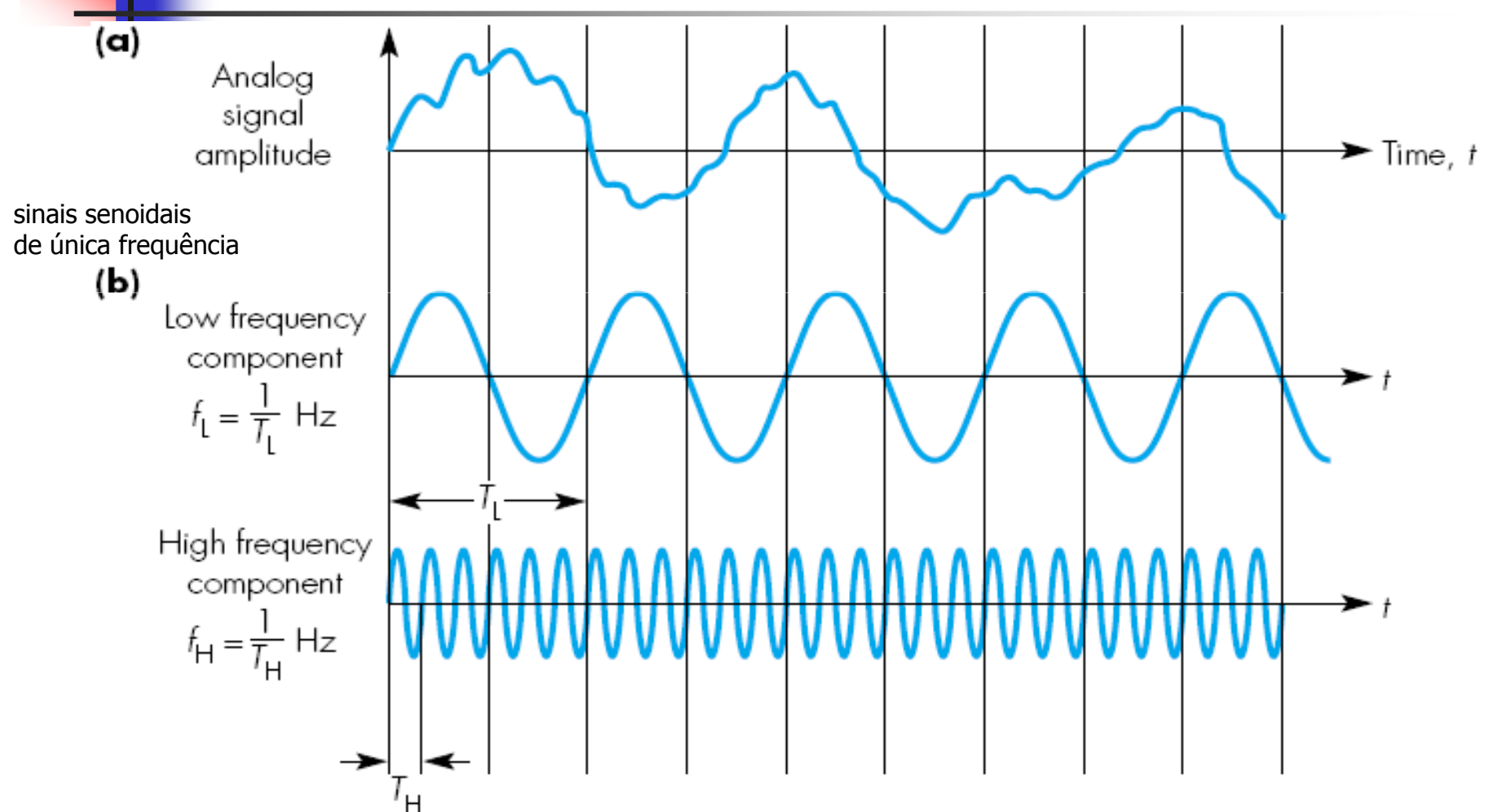
- Para poder ser utilizado em um computador, o som precisa de duas transformações:
 - Eletrônica: conversão de ondas mecânicas em sinais elétricos.
 - Digital: conversão de sinais elétricos em bits.
- Similarmente ao ouvido, o microfone é um transdutor.
 - Converte as variações de pressão do ar em sinais elétricos usáveis pelos equipamentos de áudio.
 - A saída de um microfone é uma voltagem elétrica analógica que varia no tempo do mesmo modo que as ondas mecânicas do som = **Sinal de Audio**

2.1 - Princípios de Digitalização.



- Frequência: taxa com que o sinal varia entre valores positivos e negativos. É medida em Hertz (Hz).
- Amplitude: diferença entre os máximos valores positivos e negativos do sinal de áudio. Pode ser expressa observando-se a voltagem (dependente do sistema). Normalmente expressa em decibéis (dB).

2.1 - Princípios de Digitalização.



sinais senoidais
de única frequência

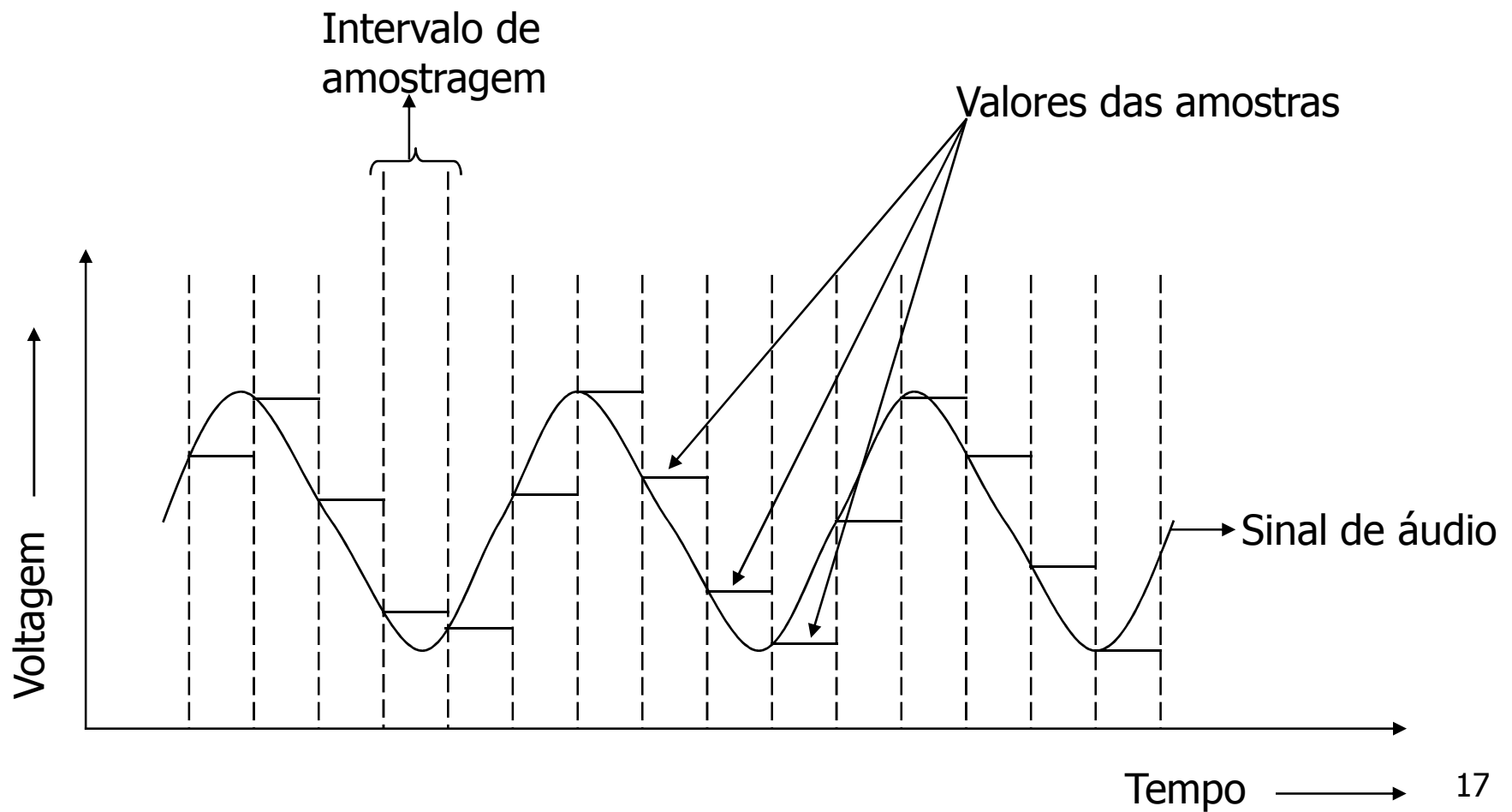
$T_{L/H}$ = time for one cycle = signal period



2.1 - Princípios de Digitalização.

- Conversão analógico-digital.
 - Sinal de áudio possui duas dimensões: voltagem e tempo. As quais serão digitalizadas através de dois processos:
 - **Amostragem**: realiza leituras periódicas e instantâneas da voltagem em espaços de tempo uniformes.
 - **Quantização**: converte os valores analógicos amostrados em valores digitais.
 - Codificador:
 - Filtro digital + ADC (Analog to Digital Converter)

2.2 - Amostragem.





2.2 - Amostragem.

- O quanto deve ser amostrado?
 - Reconstruir exatamente o sinal → infinitas amostras.
 - Poucas amostras → sinal distorcido.

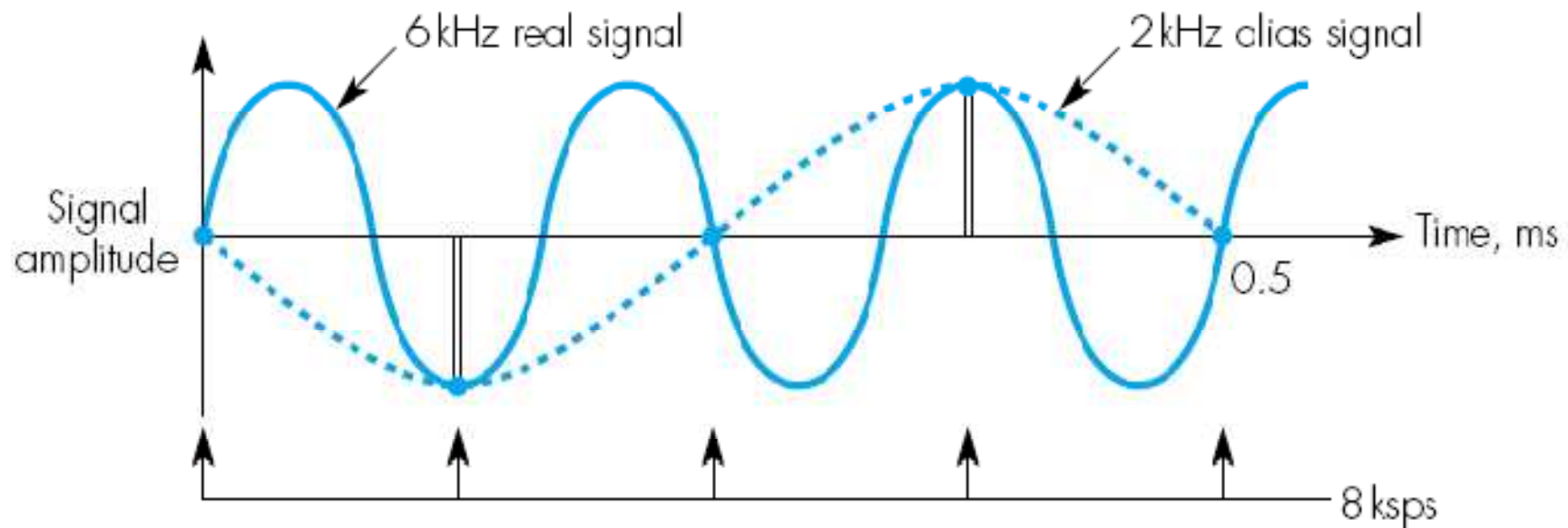


2.2 - Amostragem.

- O quanto deve ser amostrado?
 - **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.
 - sps = samples per second.

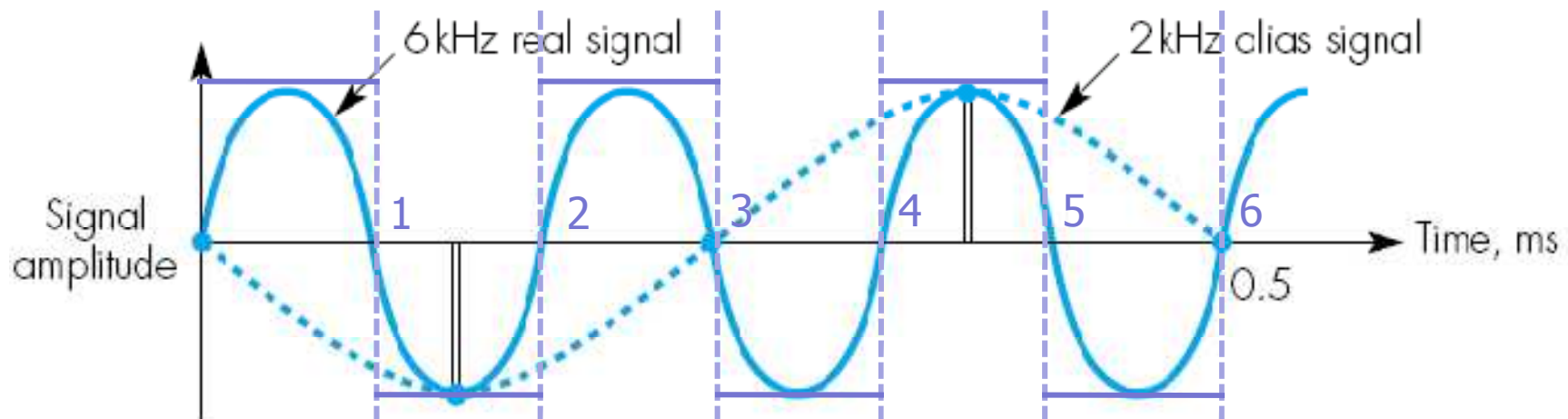
2.2 - Amostragem.

- Aliasing.
 - Sinal = 6 KHz
 - Exemplo: 8 KHz de amostragem



2.2 - Amostragem

- Teorema de Nyquist $\rightarrow 2 \times 6 = 12$ KHz de amostragem (mínimo)

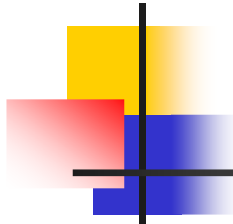




2.2 - Amostragem.

- Filtros anti-aliasing.
 - Removem as componentes acima da taxa de Nyquist (frequências muito altas).

- Em sistemas multimídia:
 - A largura de banda do canal é normalmente menor que a largura de banda do sinal.
 - A taxa de amostragem é determinada pelo largura de banda do canal.
 - A taxa de Nyquist será baseada na frequência mais alta suportada pelo canal.



2.2 - Exercício

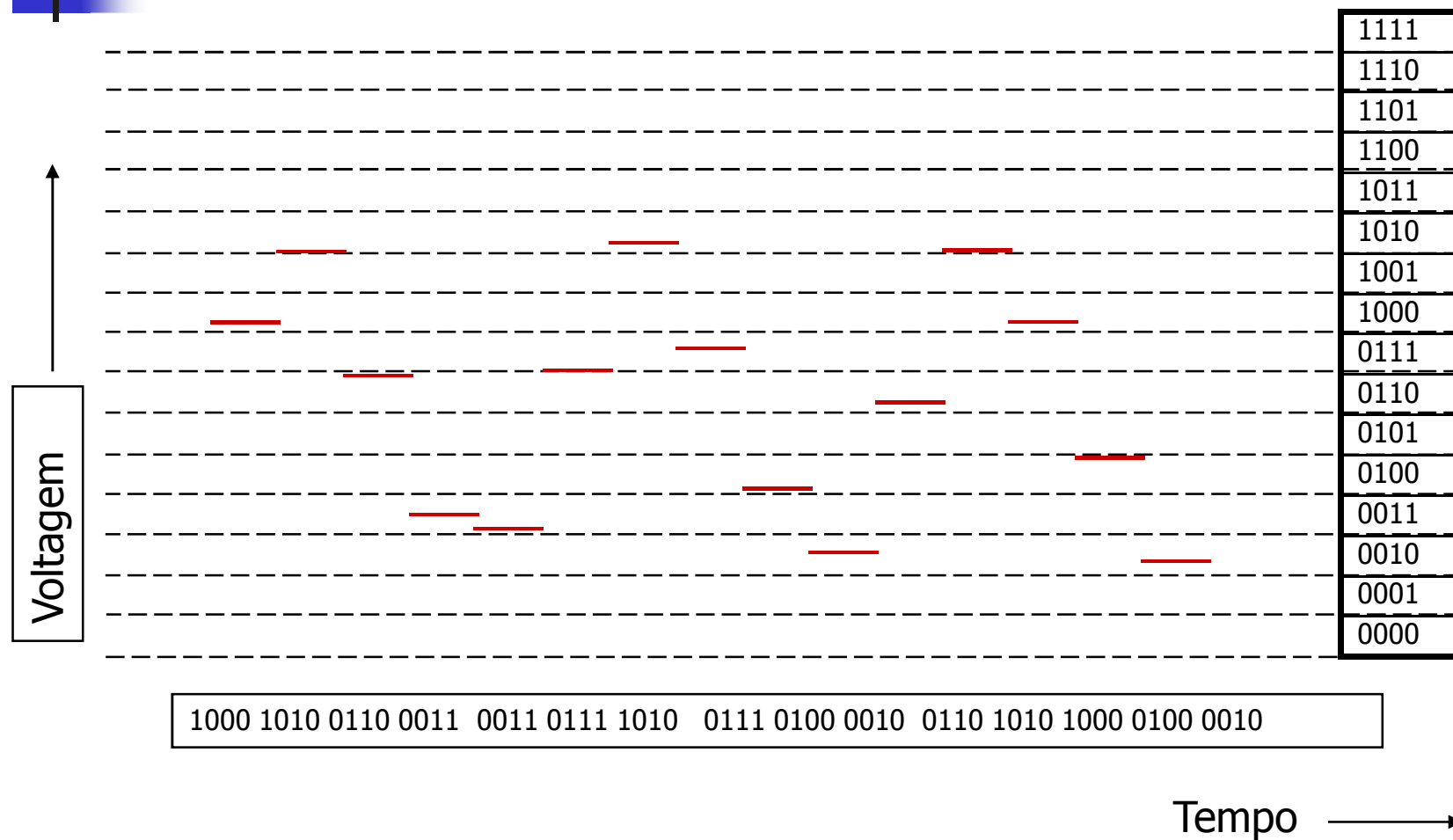
- Determine a taxa de amostragem e a largura de banda de um filtro anti-aliasing necessários para digitalizar um sinal analógico cuja largura de banda varia de 15 Hz até 10 kHz, assumindo que:
 - O sinal digital será armazenado no computador
 - O sinal digital será transmitido em um canal com largura de banda de até 3.4 kHz.



2.3 - Quantização.

- 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.

2.3 - Quantização.





2.3 - Quantização.

- Amostragem e Quantização
 - Número de amostras x número de níveis.
 - Compromisso.
 - Quantização resulta em distorções.



2.4 - Digitalização.

- Taxas comuns de amostragem:
 - 8.000Hz, 11.025Hz, 22.050Hz e 44.100Hz (CD).
- Números comuns de bits por amostra:
 - 4, 8, 16 e 24.
- Canais de som:
 - 1 (mono), 2 (stereo), 3, 5, 7, ...
- Qualidade de CD:
 - Amostras a 44.100Hz (4,1 KHz), 16 bits por amostra e 2 canais de som (stereo).



2.4 - Digitalização.

- Técnica conhecida como **modulação por código de pulso** (linear).
 - *Pulse Code Modulation* – PCM. PCM linear.
- Circuito que realiza amostragem e quantização:
 - Conversor analógico-digital (*analog to digital converter* – ADC).
 - Caminho inverso: DAC. Usado na reprodução de áudio digital.
- PCM é normalmente implementado em hardware.



2.4 - Digitalização.

- Após a captura
 - os dados amostrados e quantizados devem ser “guardados” em algum formato – mídia de representação.
 - WAV e MP3, por exemplo.



2.4 - Digitalização.

- Aspectos quantitativos.
 - Quantos bytes serão necessários para armazenar 1 segundo de áudio, capturado com qualidade de CD?



2.4 - Digitalização.

- Aspectos quantitativos.
 - Quantos bytes serão necessários para armazenar 1 segundo de áudio, capturado com qualidade de CD?
 - $1(\text{segundo}) * 44.100 (\text{taxa de amostragem}) * 2 (16 \text{ bits por amostra}) * 2 (\text{som estéreo}) = 176.400 \text{ bytes.}$
 - Necessidade para transmissão: 1,41Mbps!



Exercício

- Assumindo que a largura de banda de um sinal de fala é de 50 Hz até 10 kHz, e que a largura de banda de um sinal de música é de 15 Hz até 20 kHz, calcule:
 - A taxa de bits que é gerada pelo processo de digitalização em cada caso assumindo a taxa de amostragem Nyquist, e 12 bits por amostra para o sinal de fala e 16 bits por amostra para o sinal de música
 - Calcule a quantidade de memória necessária para armazenar 10 minutos de música em som estéreo

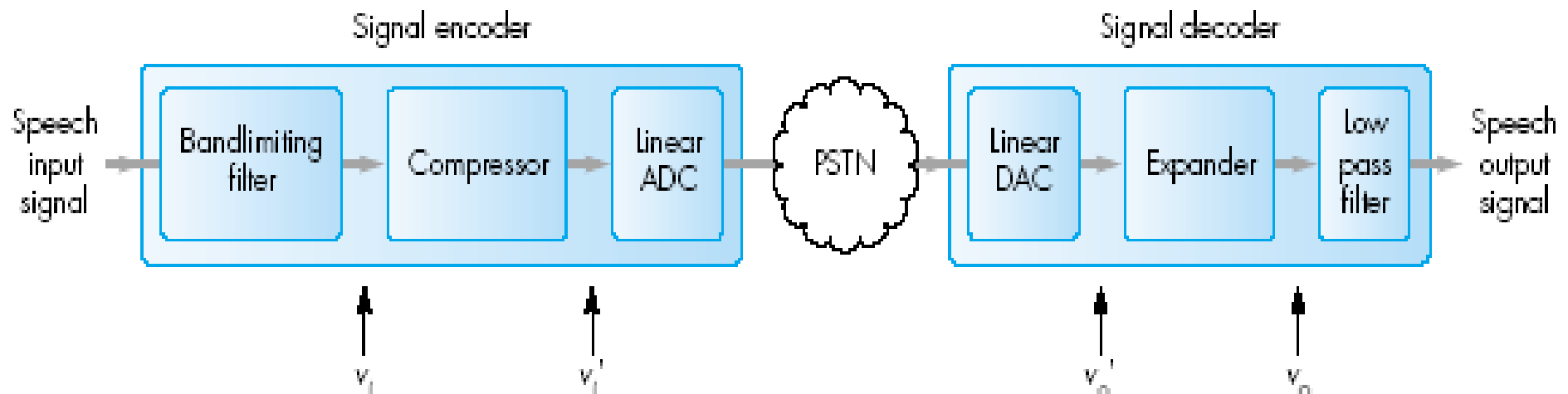


3. Compressão de Áudio

3.1 – PCM Logarítmico

■ Pulse Code Modulation

- Processo de digitalização de áudio para ser usado em redes públicas de telefonia comutada (PSTN)
- Inclui um compressor.
- Recomendação ITU-T G.711.





3.1 – PCM Logarítmico

- Explora quantização não-linear.
 - Amplitudes maiores → maiores intervalos de quantização.
 - Amplitudes menores implicam em maior percepção de ruído de amostragem.
- Desempenho:
 - Utilizando 8 bits equivale à quantização linear com 12 bits.



3.1 – PCM Logarítmico

- PCM A-Law e μ -Law.
 - Usados em telefonia.
 - Largura de banda do sinal de voz: 200Hz a 3.4kHz.
 - Circuito telefônico:
 - 200 Hz até 3.4 kHz (Nyquist: 6.8 kHz)
 - Taxa limite do filtro: 8 kHz devido a imperfeições
 - μ -Law = EUA e Japão, 7 bits.
 - A-Law = Europa e outros, 8bits.

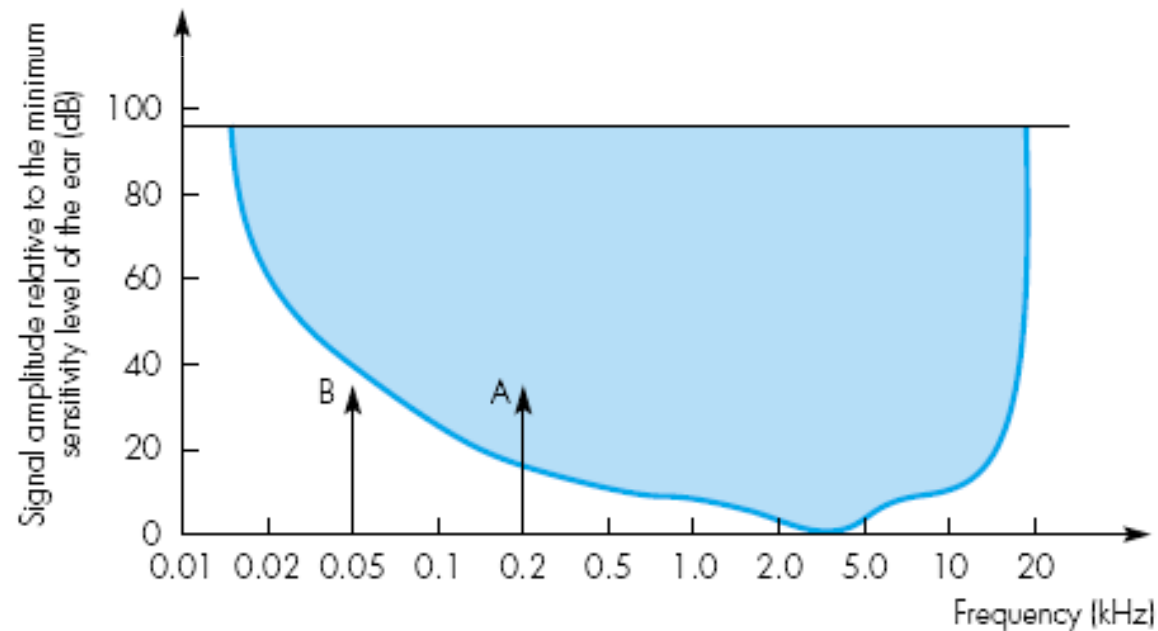


3.2 - ADPCM

- Adaptive Differential PCM.
 - Amostras adjacentes de áudio são parecidas. ADPCM faz previsão da amostra seguinte e codifica apenas a diferença.
 - Mudanças brusca entre amostras adjacentes causam distorções.
 - O algoritmo faz *lookahead* durante a compressão para adaptar a escala de diferenças de acordo com o tamanho da mudança.

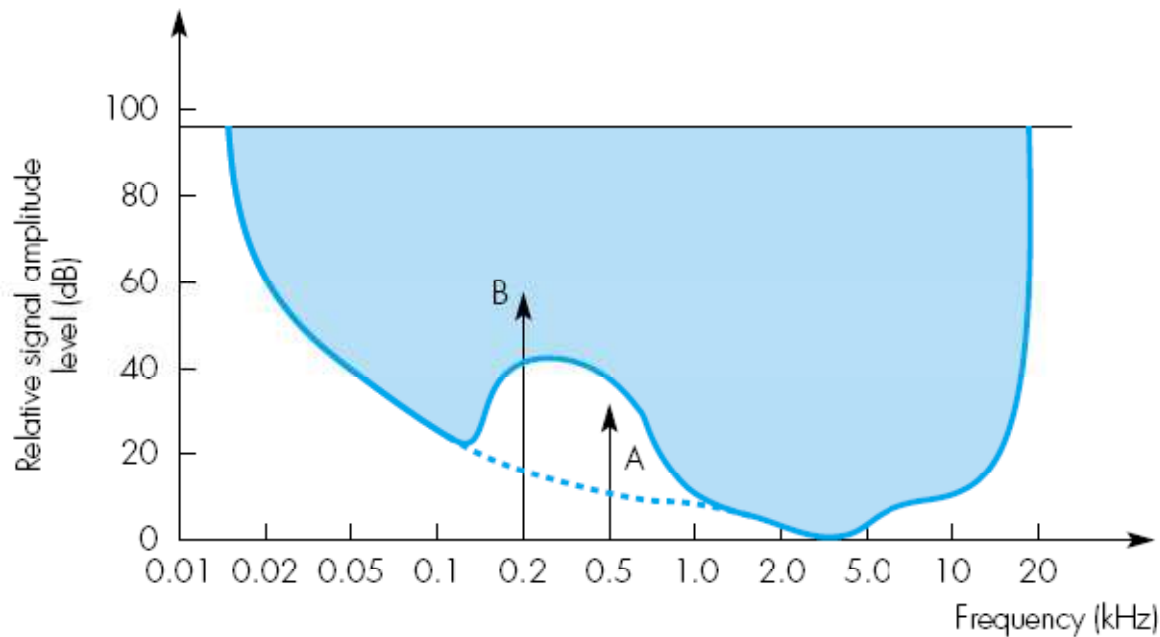
Perceptual Coding – Modelo psico-acústico

- Sensibilidade da audição



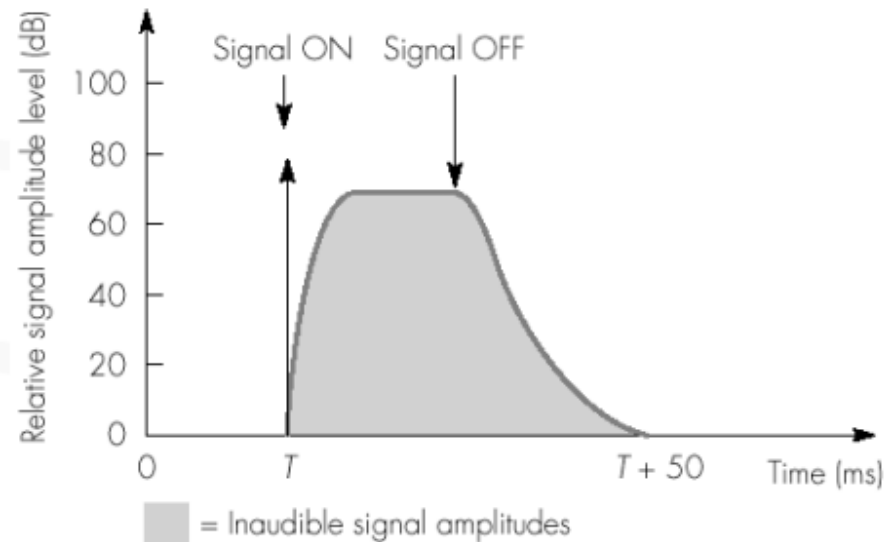
Perceptual Coding – Modelo psico-acústico

- Mascaramento de frequência



Perceptual Coding – Modelo psico-acústico

- Mascaramento temporal
 - Após ouvir um som alto, demorará alguns instantes (~ 50 milisegundos) até que o ouvido possa perceber um som mais baixo

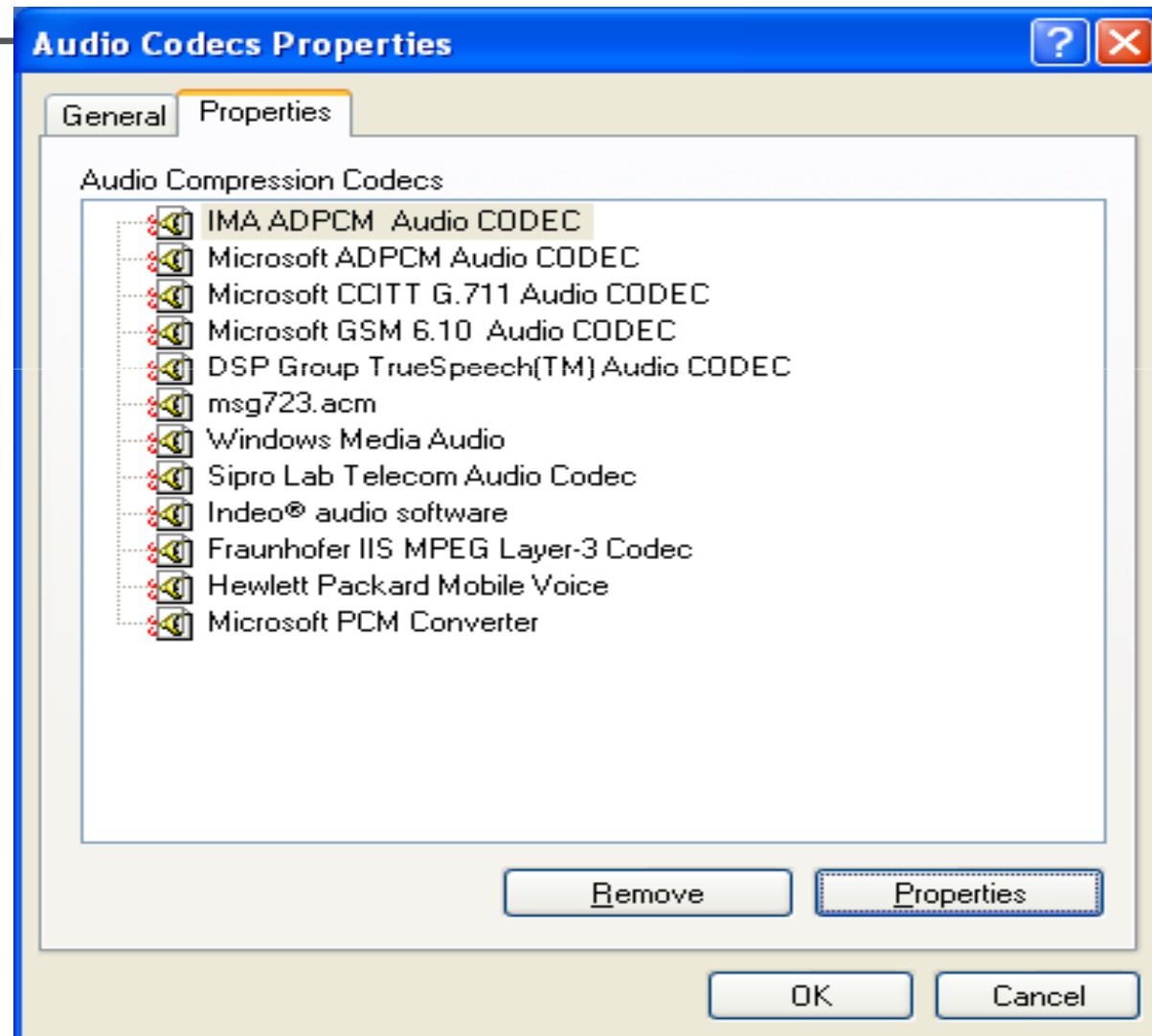




3.3 - MP3

- MPEG-1 Audio Layer 3.
 - Utiliza um modelo psico-acústico complexo.
 - Sensibilidade do ouvido.
 - Mascaramento temporal.
 - Mascaramento de frequência.
 - Elimina as frequências que o ouvido humano não consegue captar.
 - Audição humana: 20Hz a 20KHz; 2 a 4KHz.
 - Compressão com perda. As perdas não são perceptíveis.
 - Som de alta qualidade e a arquivos até 12 vezes menores.

Codecs





3.4 – Áudio sintetizado.

- Áudio pode ser sintetizado.
 - sintetizadores: aparelhos capazes de gerar sons sintéticos a partir de notas musicais.
 - Cada nota é um código.
 - Sintetizadores podem “imitar” o som de diversos instrumentos.
- MIDI – Music Instrument Digital Interface
 - Protocolo para comunicação digital entre instrumentos eletrônicos.
 - Formato das mensagens, conectores, cabos e sinais eletrônicos
 - Arquivos .MID.



Para Saber Mais

- Áudio e digitalização:
 - Luther, A. C. Using Digital Video. AP Professional, 1995.
 - Halsall, F. Multimedia Communications: Applications, Networks, Protocols, and Standards, Addison-Wesley Publishing, 2001. ISBN: 0201398184. Capítulo 2.
- MIDI:
 - <http://www.midi.org/>



Exercícios

- Calcule:
 - A capacidade necessária para armazenar 60 minutos de um áudio com qualidade de CD.
 - Tempo necessário para transmitir 30 segundos do áudio supra-citado usando um canal de transmissão com:
 - 64kbps
 - 1.5 Mbps