



GRVM

SOM

Judith Kelner

jk@cin.ufpe.br

Samuel Macedo

samuel@gprt.ufpe.br

Grupo de Pesquisa em Realidade Virtual Multimídia
Universidade Federal de Pernambuco, Centro de Informática



2011-1º.semestre



Roteiro

- Introdução
- Digitalização
- Processamento
- Compressão e formatos



GRVM

Introdução



2011-1º.semestre



Som

▶ História

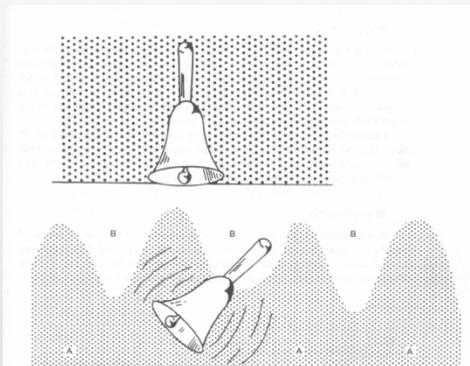
- Pitágoras na Grécia em 500 AC (Monocórdio)

▶ Som

- forma de energia mecânica que se propaga causando compressão e rarefação das moléculas de um meio elástico e inercial (sólido, líquido, gasoso)
- Decaem radialmente com o quadrado da distância da fonte emissora

▶ Quatro elementos

- fonte excitadora (ex. dedos + cordas)
- superfície vibratória (ex. caixa do violão)
- meio de propagação (ex. ar)
- receptor (ex. ouvido)

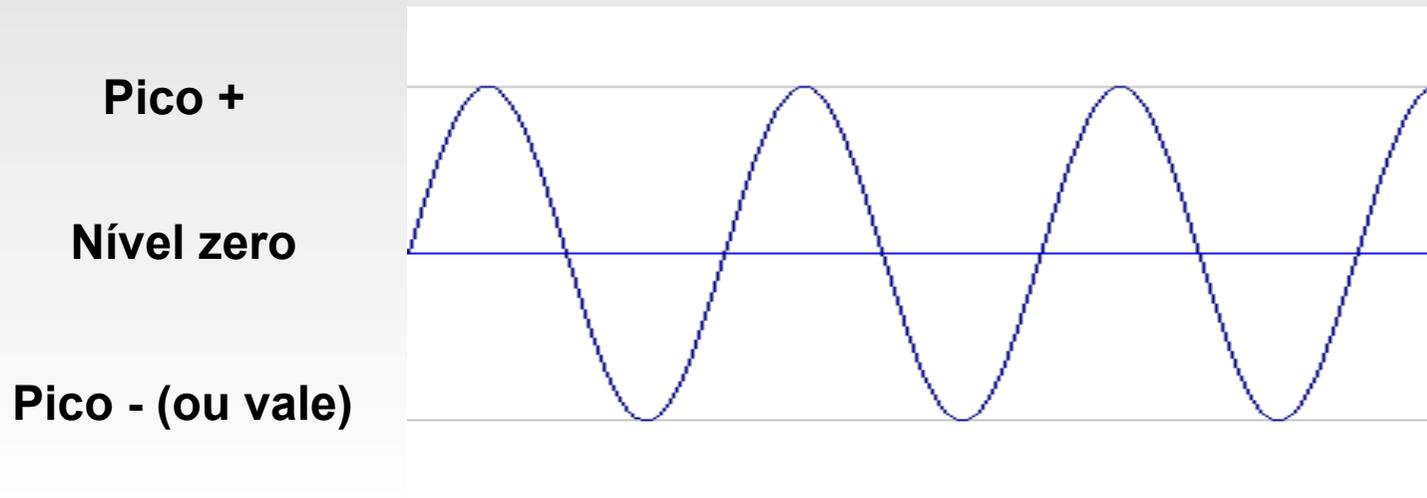


Onda sonora



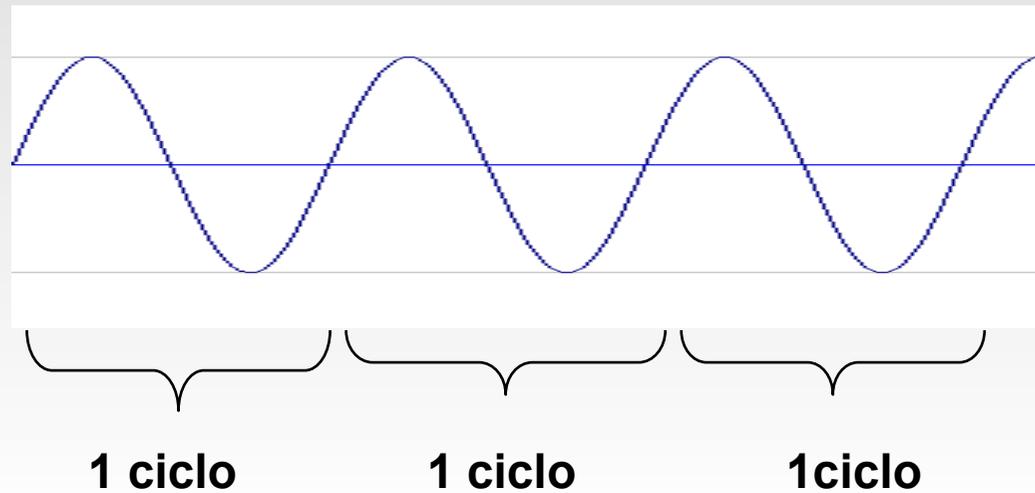
- Características principais
 - amplitude, frequência, comprimento, velocidade, fase, potência, etc.

Amplitude



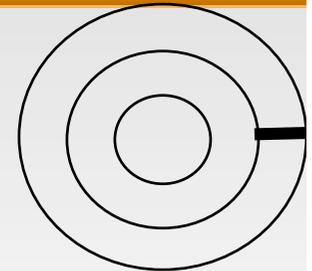
- **Amplitude a**
 - Distância de um ponto da curva ao nível zero
 - Medida instantânea de energia
 - Quanto maior, mais forte o som

Período e Frequência



- **Período T**
 - Tempo (em segundos) de duração de um ciclo
- **Frequência f**
 - Número de ciclos por segundo: Hertz (hz)
 - Inverso do período ($f = 1/T$)
 - Quanto maior a frequência, mais agudo o som
 - Ouve-se de 20 a 20.000 Hz

Comprimento e Velocidade

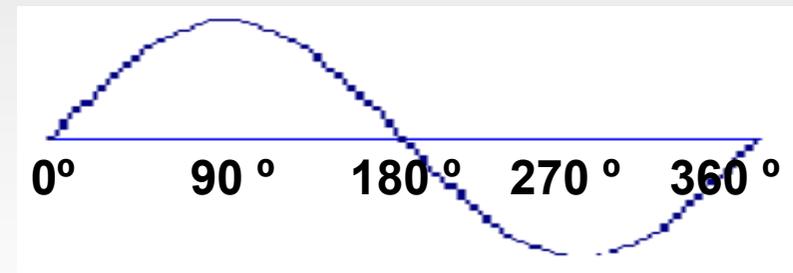
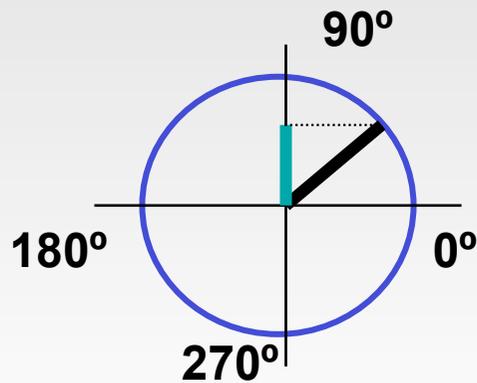


- Comprimento de onda λ
 - Semelhante ao período, só que mede a distância física (milímetros) de um ciclo
 - $\lambda = c/f$
 - onde c é a velocidade do som e f a frequência
 - inversamente proporcional à frequência
 - som agudo => pequeno comprimento
 - som grave => grande comprimento
- Observações
 - graves são dificilmente “localizáveis” em locais “pequenos” (ex. salas) por causa do grande comprimento de onda

Velocidade

- Velocidade de propagação: $c = \lambda f$
 - diretamente proporcional à frequência e ao comprimento de onda
 - depende do meio e da temperatura
 - 344 m/s no ar
 - 1500 m/s na água
 - 5000 m/s no aço
- Efeito Doppler
 - mudança de velocidade causando mudança de frequência
 - ex. ambulância passando

Fase

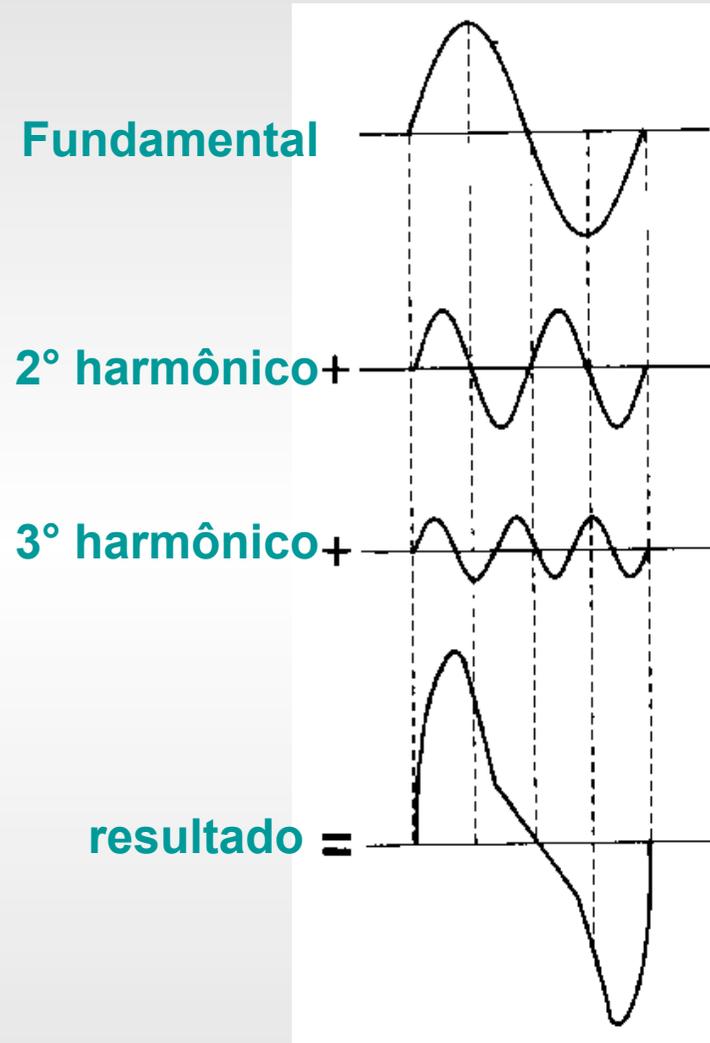


■ Fase

- depende do instante em que a onda começou
- medida em graus, sendo 360° o ciclo completo

Onda complexa

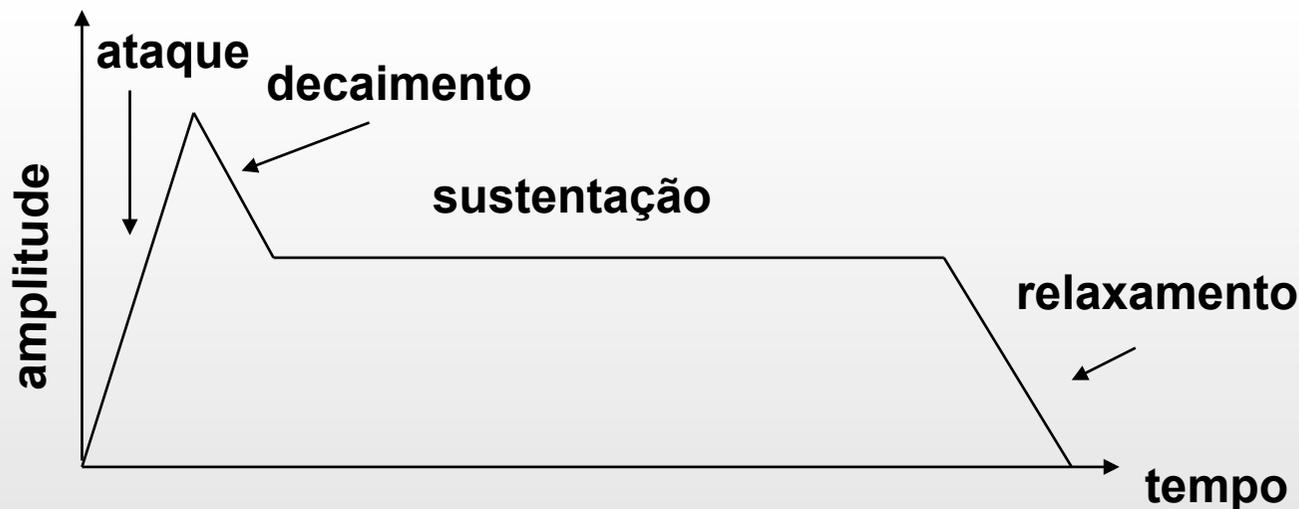
- O conteúdo harmônico
 - É um dos responsáveis pelo timbre de um instrumento
 - É chamado Resposta em Frequência ou Espectro



Envoltória

■ Envoltória:

- Indica como a energia do som se distribui no tempo
- Segundo elemento marcante na definição do timbre. Cada instrumento tem o seu
- O ataque é a característica mais marcante



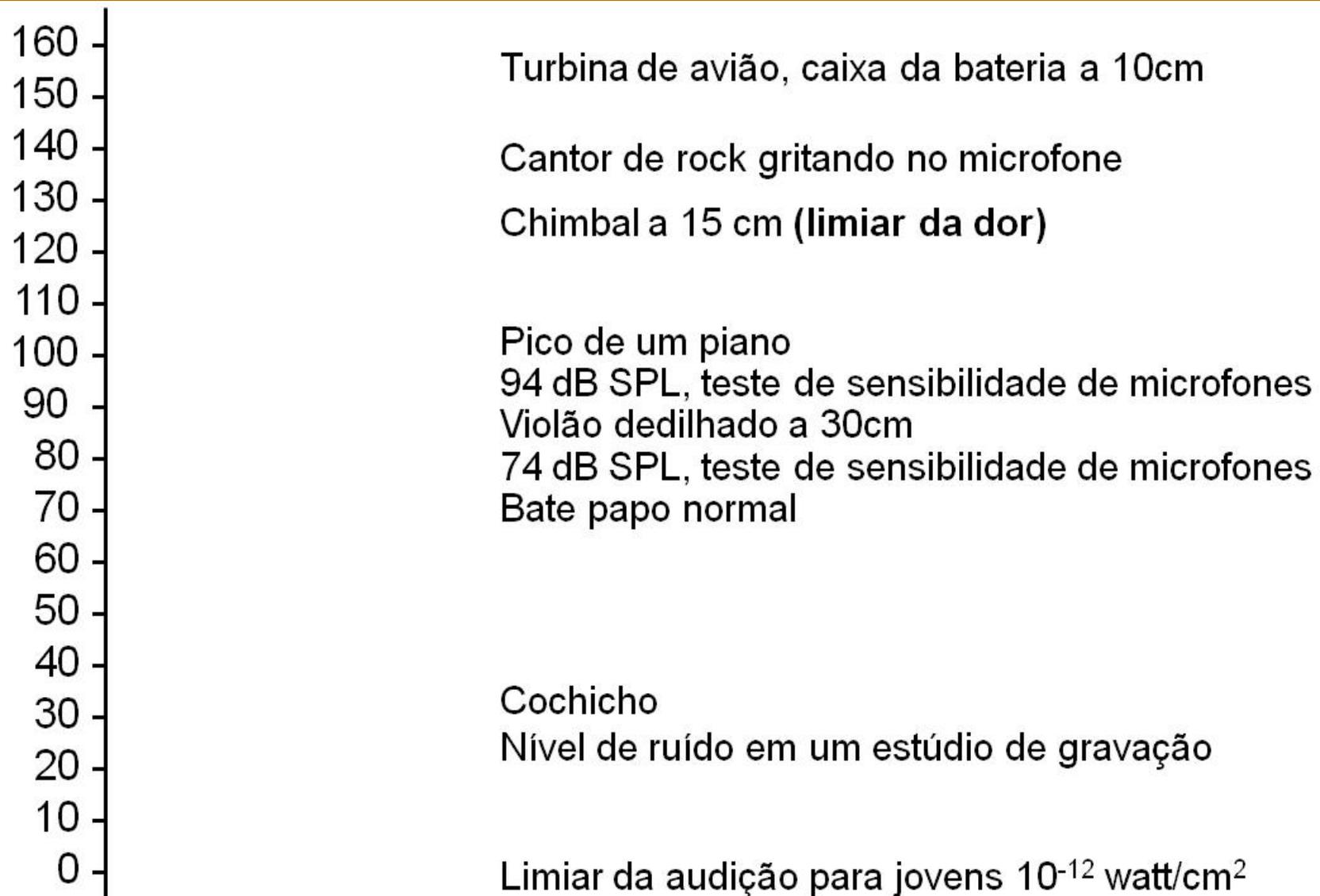
Potência de uma onda sonora

- Volume (nível de áudio): decibel (dB)
 - 1 dB = menor mudança de volume perceptível
 - É uma medida **relativa** entre tensões, correntes, potências ou pressões acústicas

$$\text{dB} = 10 \times \log_{10} (\text{nível/nível de referência})$$

- Existem vários níveis de referência
 - dBm: 1 miliwatt
 - dBu ou dBv: 0.775 volt
 - dBV: 1 volt
 - dB SPL: 10^{-12} watt/cm² (limiar da audição)

Sound Pressure Level (dB-SPL)



Ruído

■ Ruído

- Sinal com espectro de frequência denso e pouco harmônico.

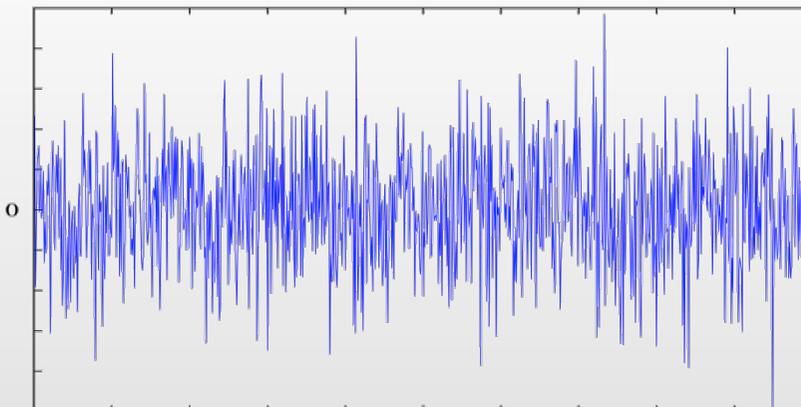
■ Faixas de frequência

■ Ruído rosa:

- mantém a potência (energia) igual entre todas as oitavas sonoras (Aplica-se ao teste de sistemas de áudio)

■ Ruído Branco:

- Aplica-se aos ruídos cuja potência é constante em todas as faixas de frequência, por exemplo, ruído térmico



Relação sinal-ruído

- Relação sinal-ruído (Signal-to-Noise Ratio)
 - deve ser a maior possível
 - Fita cassete NSR = 50 dB
 - CD NSR = 90 dB
- Depende dos meios de armazenamento e/ou captação
- Depende da dinâmica do sinal de áudio
 - dinâmica: variação de volume
 - sinais de baixa potência podem ser mascarados pelo ruído.
 - sinais de alta potência podem sofrer distorções por limitação dos meios de armazenamento e/ou captação



GRVM

Digitalização



2011-1º.semestre



Ondas Sonoras

- Tudo o que chamamos de som são vibrações dos meios físicos
- Quando estas vibrações ocorrem entre 20 e 20 mil vezes por segundo ouvimos um ou mais sons -> 20 Hz até 20 kHz, aproximadamente.



Som Analógico

- Um microfone tem uma membrana que acompanha as vibrações do ar e um circuito que gera uma corrente elétrica
- As vibrações são transformadas em oscilações na voltagem da corrente que ele manda pelo cabo.
- Temos então, o sinal elétrico do áudio, ou o som analógico.



Som Digital

- Converter o som analógico em informações expressas em números.
- Conversor analógico/digital, ou ADC

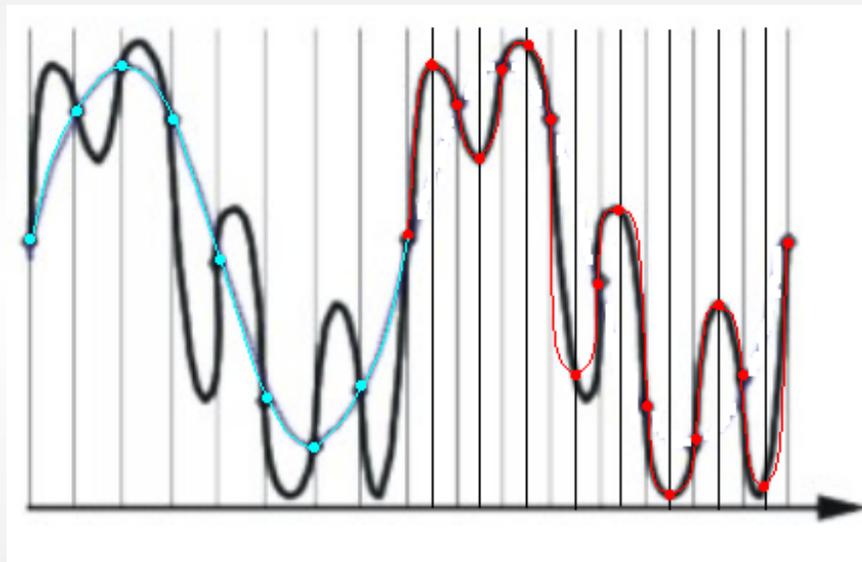


Conversão Analógico/Digital

- **Quantificação** do valor da amplitude do sinal em vários instantes de tempo (Amostragem).
- A quantificação é feita a uma frequência definida como **Taxa de Amostragem** (número de amostras feitas em um segundo).
- Parâmetros no processo de Digitalização:
Quantificação + Taxa de Amostragem

Taxa de Amostragem

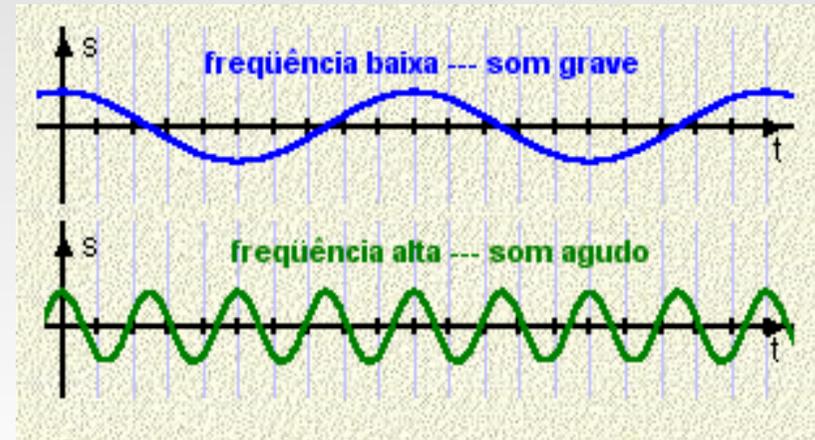
- Número de amostras do sinal analógico que vamos tirar por segundo.



- Uma onda complexa tem frequências perdidas numa baixa taxa de amostragem

Taxa de Amostragem

- Um som grave de 100 Hz, um agudo de 10 kHz e um ADC com taxa de amostragem de 10 kHz.



- apenas uma amostra por ciclo, não permite visualizar (nem ouvir) uma oscilação, o que significa silêncio.

O Teorema de Nyquist - Shannon

Para permitir a audição de uma oscilação numa certa frequência, a amostragem deve ser de, no mínimo, o dobro daquela frequência.

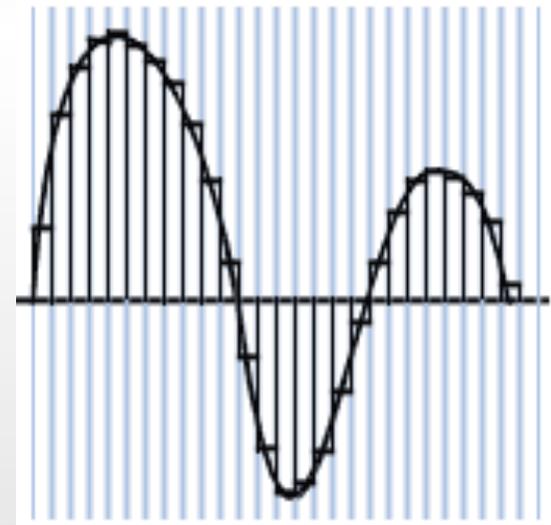
Taxa de Amostragem

- O ouvido humano percebe os sons até cerca de 20 000 Hz.
- É necessário uma frequência de amostragem de pelo menos aproximadamente 40 000 Hz para obter uma qualidade satisfatória

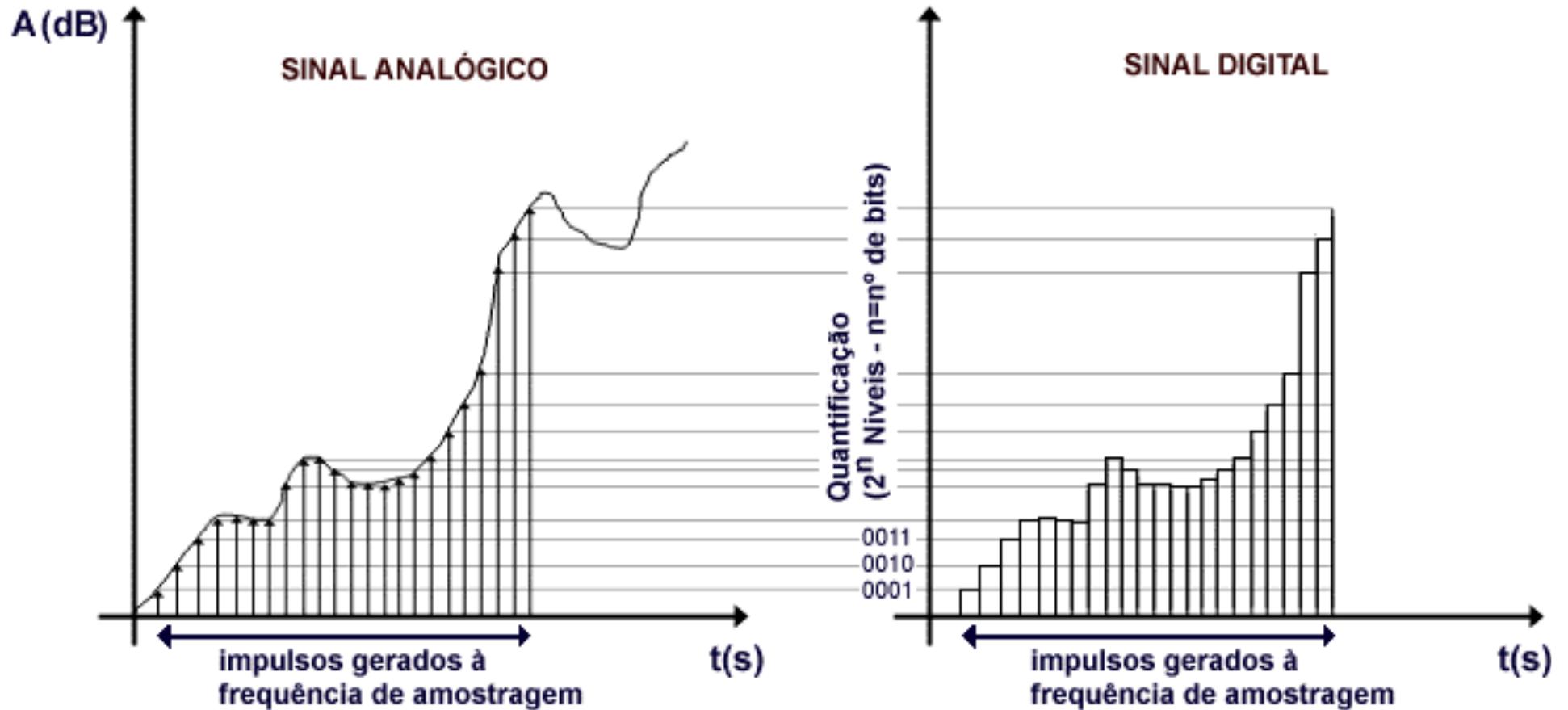
Taxa de Amostragem	Qualidade do Som
44 100 Hz	Qualidade de CD
22 000 Hz	Qualidade de rádio
8 000 Hz	Qualidade de Telefone

Quantificação

- Resolução em termos do valor de Amplitude que o sinal pode ter num determinado instante.
- 8 bits -> 256 valores possíveis
- 16 bits -> 65.536 valores possíveis
- Qualidade do Som X Tamanho

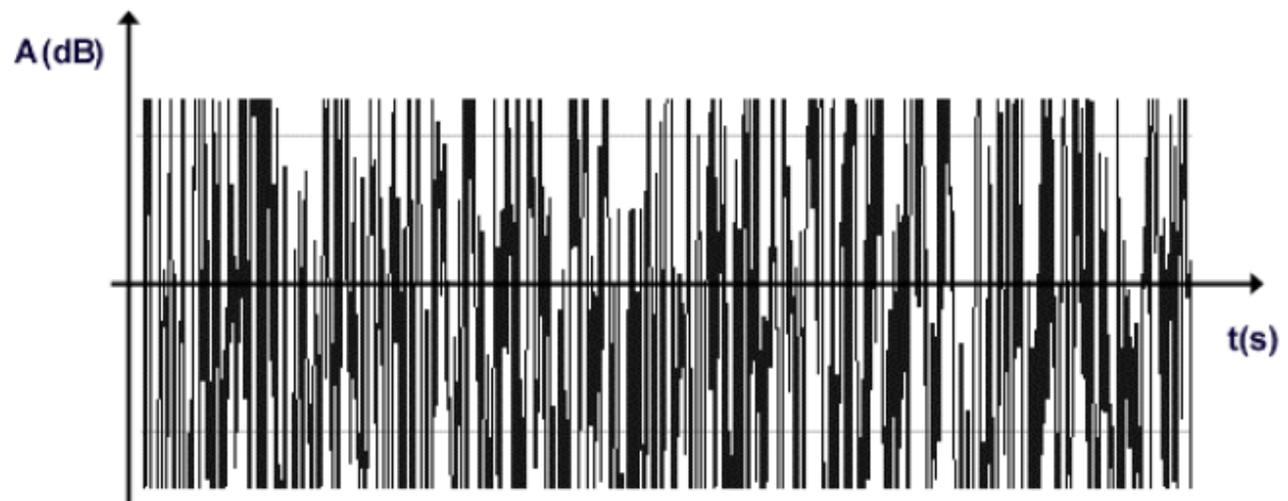
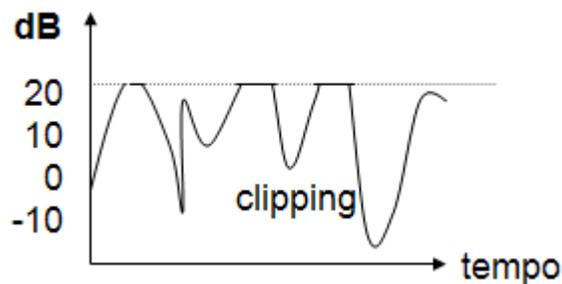


Conversão Analógico/Digital



Clipping

- A amplitude do som excede o valor máximo da gama de quantificação
- O som digitalizado vai tomar o valor máximo quantificável pelo ADC.



Resampling

- Conversão entre formatos de som digital com alteração dos parâmetros de frequência de amostragem e quantificação.

CD (44.1 KHz/16 bits) -> CD-Rom (22.05 KHz/8 bits)

- Retirar amostras ao sinal no caso da frequência de amostragem e arredondar a amplitude das amostras no caso da quantificação.

Resampling

■ Caso inverso

CD-Rom (22.05 KHz/8 bits) -> CD (44.1 KHz/16 bits)

- Aumentar a frequência de amostragem e quantificação
- Utiliza-se técnicas de interpolação, de forma a ser possível obter novas amostras e novas amplitudes que anteriormente não existiam.

Som Estéreo

- Permite reproduzir a diferença de tempo que leva para o som chegar aos ouvidos.
- Dois canais de som (direito - RC e esquerdo - LC) que emitem o som com uma pequena defasagem temporal.
- Ocupa o dobro do espaço em disco que o som digitalizado em mono.

Tamanho dos Arquivos

- A dimensão de um arquivo de áudio será igual a:

Taxa de amostragem \times Número de bits \times segundos \times Número de canais

Tamanho da Amostra	Taxa de Amostragem	Canais	Tamanho do Arquivo
8 bits	11.025 Hz	1 (mono)	11 Kbytes/s
16 bits	44.100 Hz	1 (mono)	88 Kbytes/s
16 bits	44.100 Hz	2 (estéreo)	176 kbytes/s



GRVM

Processamento



2011-1º.semestre



Processamento de Áudio

- Processamento: manipulação feita em um sinal para
 - Melhorar a qualidade do sinal em algum aspecto
 - Adicionar/criar efeitos especiais
 - Extrair informações
- Tecnologia do processamento
 - Digital
 - Analógico

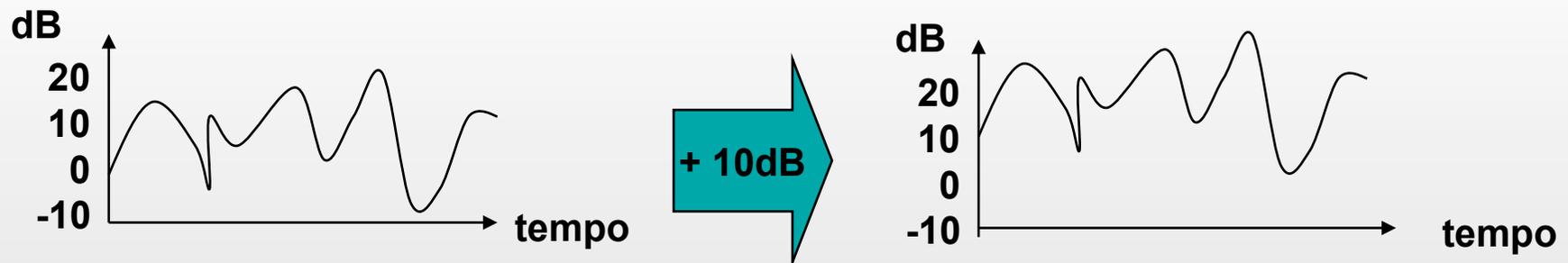
Processadores de Sinais

- Alteração na dinâmica (**amplificação/atenuação**)
 - Compressores, limitadores e noise gates
 - Modelador de envoltória
- Mudança no espectro
 - Filtros e equalizadores
- Efeitos de atraso de tempo
 - Delayers e reverberadores
- Mudança no tempo/altura
 - Flanger, phasing, chorus

Amplificação/Atenuação

■ Amplificação/Atenuação

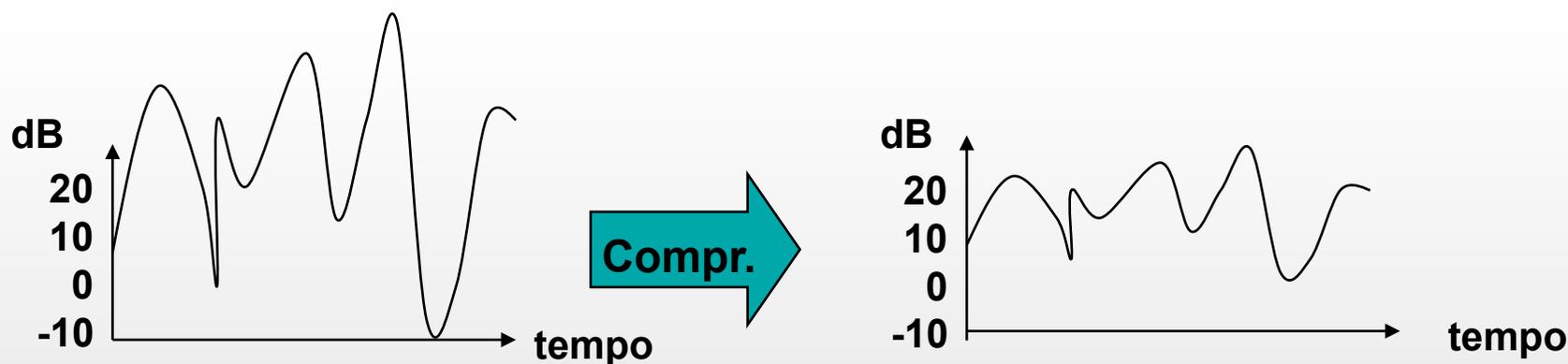
- Operação de aumento/diminuição do volume do sinal realizada pelo amplificador
- Cada ponto é somado/subtraído pelo ganho do amplificador (em dB)



Compressão

■ Compressão

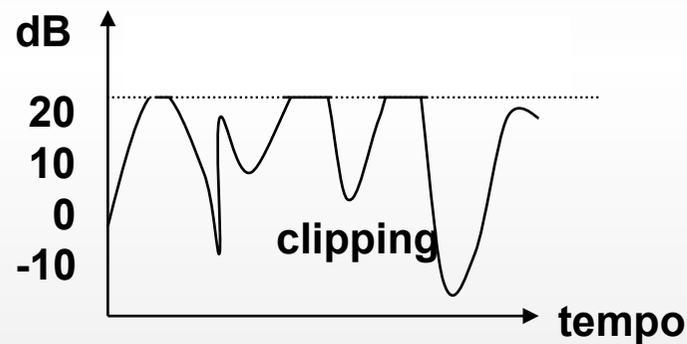
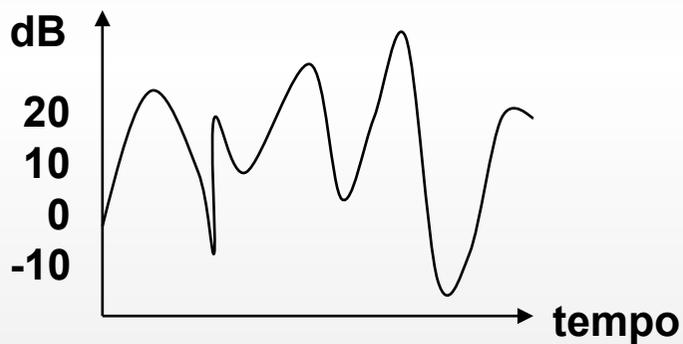
- Amplificação cujo ganho é controlado pelo próprio sinal de entrada: diminui fortes e aumenta fracos, compactando o sinal



Compressão

■ Uso

- Manter nível do sinal mais constante
- Dar mais sustentação ao som
- Minimizar clipping

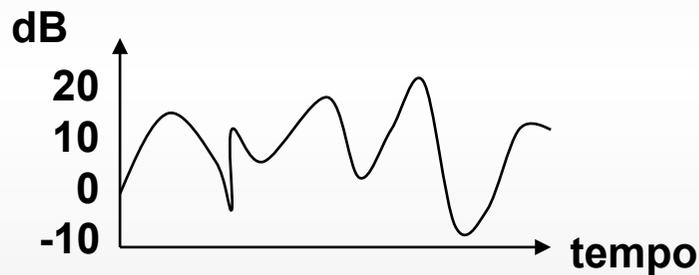


Compressão

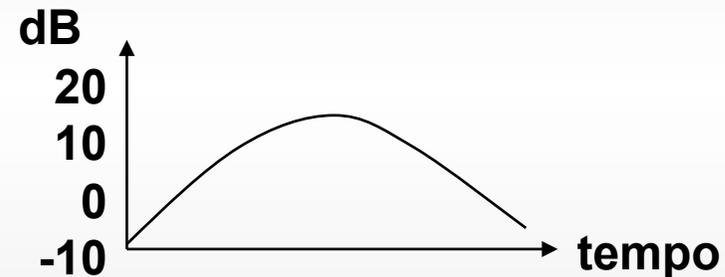
- **Parâmetros principais**
 - taxa de compressão: razão entrada-saída (E/S)
 - tempo de ataque: quão rápida a compressão passa a atuar uma vez ultrapassado um certo limiar
 - tempo de decaimento: quão rápida a compressão cessa depois de que o sinal volta a ficar abaixo de um certo limiar

Modelagem de envoltória

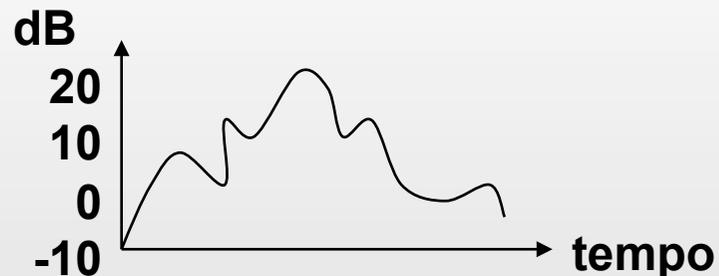
- Modelagem de envoltória
 - Operação mais sofisticada na qual a amplificação é variável no tempo conforme uma envoltória



×

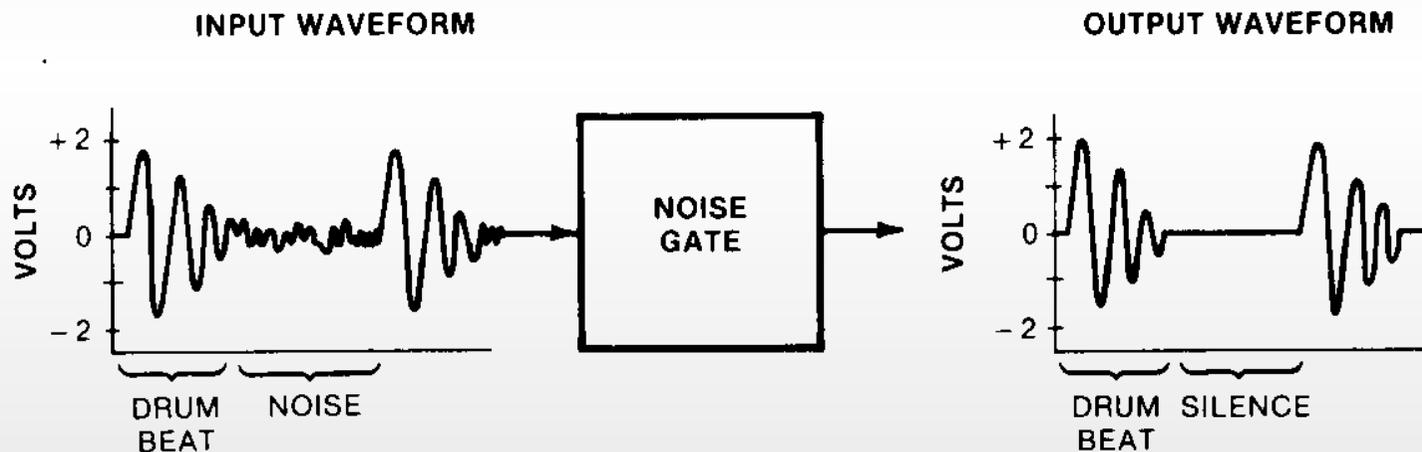


=



Noise gate

- Funciona como uma chave para eliminar ruído
 - Assim que o sinal de áudio cai abaixo do limiar de detecção, elimina o sinal e o ruído que iria ser ouvido



Noise gate

- Parâmetros principais

- Limiar de detecção de sinal (dB), tempo de ataque e de decaimento

- Algoritmo

para toda amostra de entrada i

se $\text{valor}(i) < \text{limiar } L$

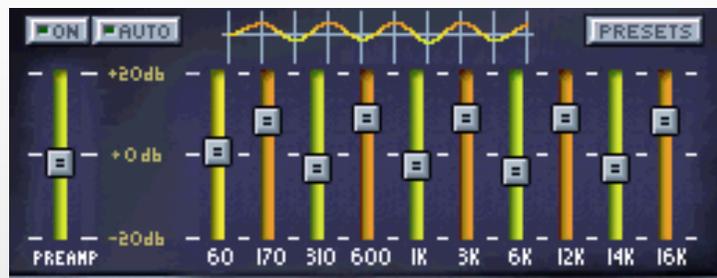
então $o := 0$

senão $o := \text{valor } (i)$

retorne o

Equalização

- Equalizadores
 - Permitem a manipulação das frequências do sinal.
- Tipos de Equalização
 - Controle de Graves e Agudos
 - Gráfica
 - Paramétrica



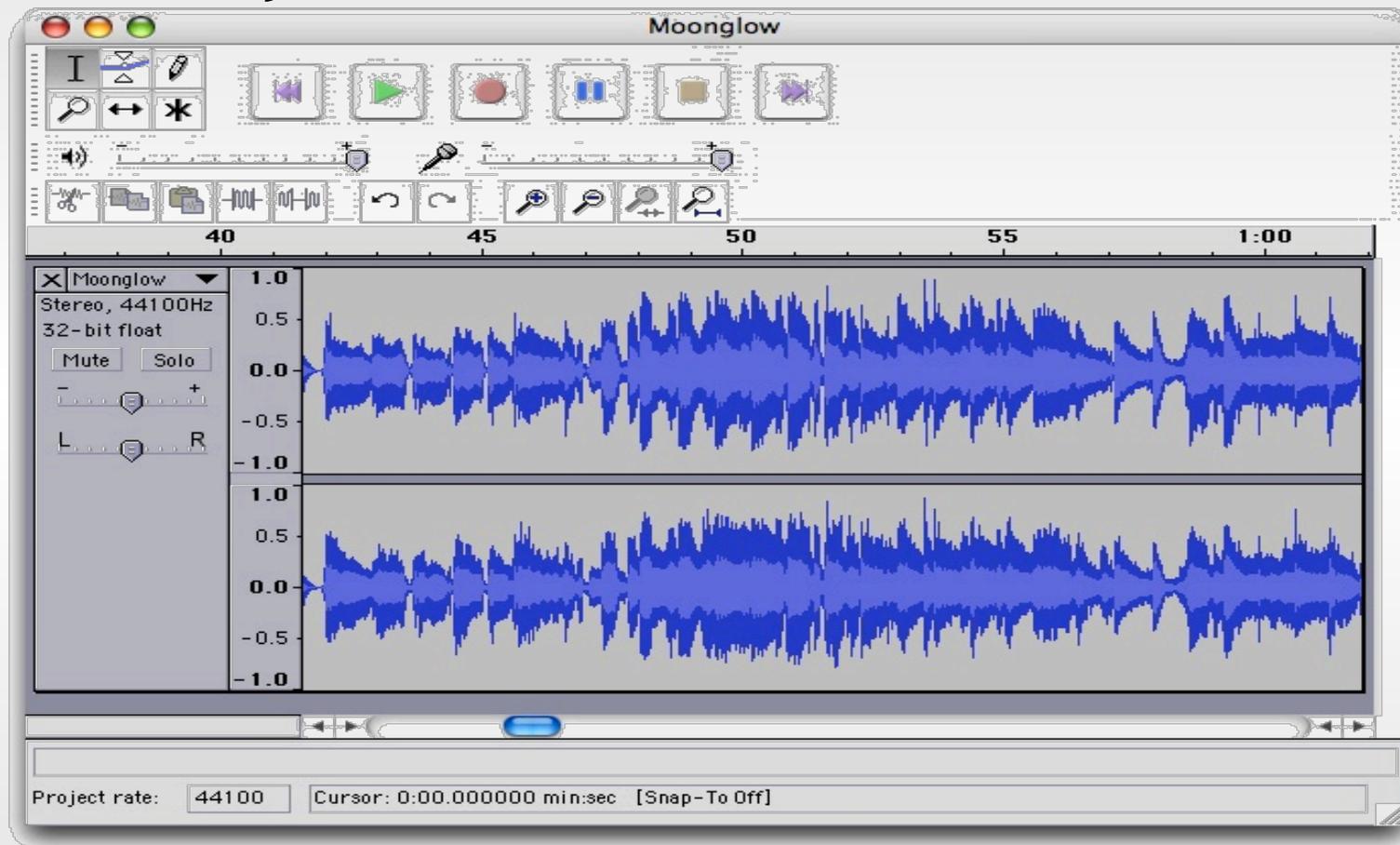
Equalização

■ Aplicação

- Melhorar a qualidade sonora de um instrumento
- Produzir efeitos especiais
- Ressaltar um instrumento mascarado por outros
- Reduzir ruído

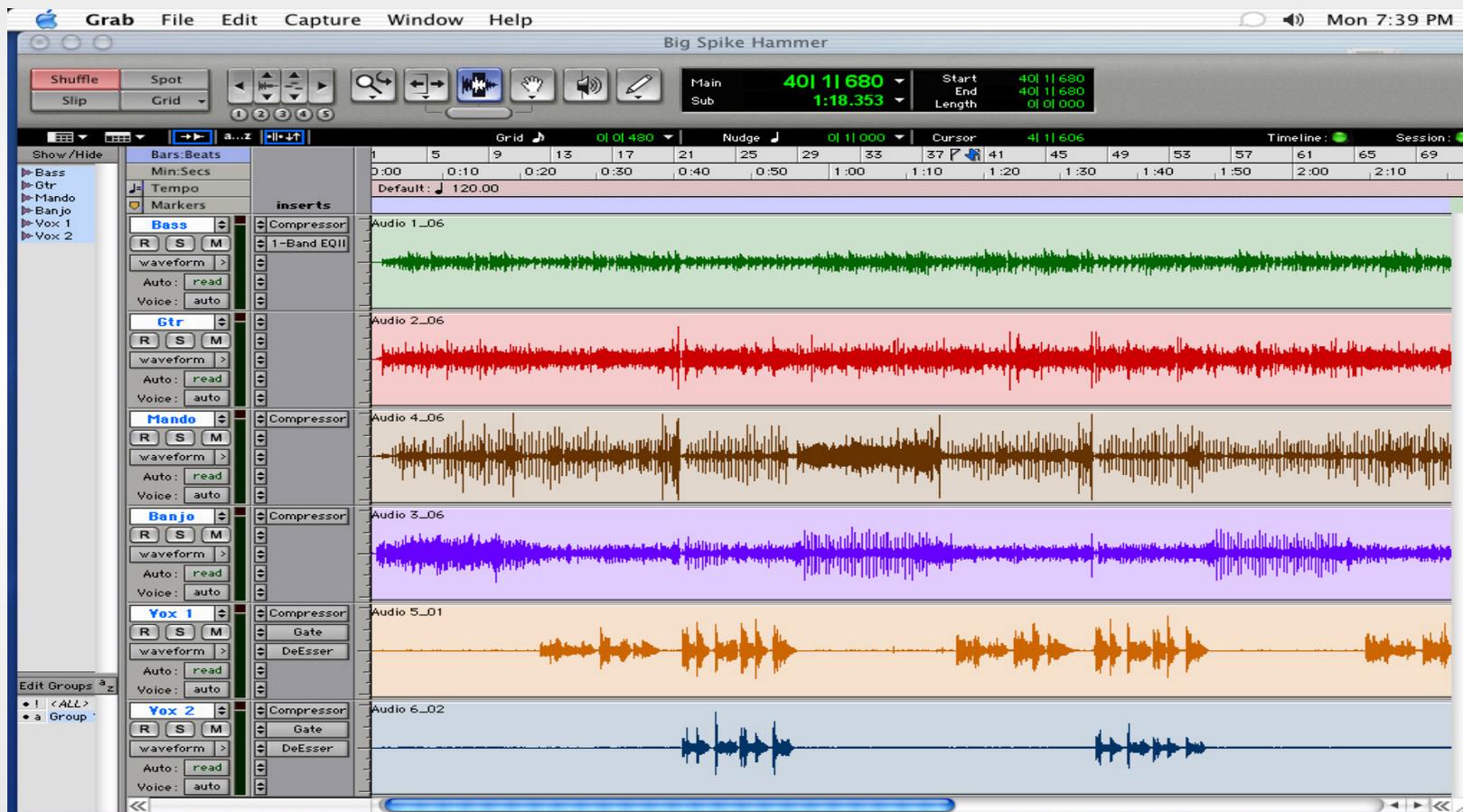
Programas de Edição de Áudio

■ Audacity



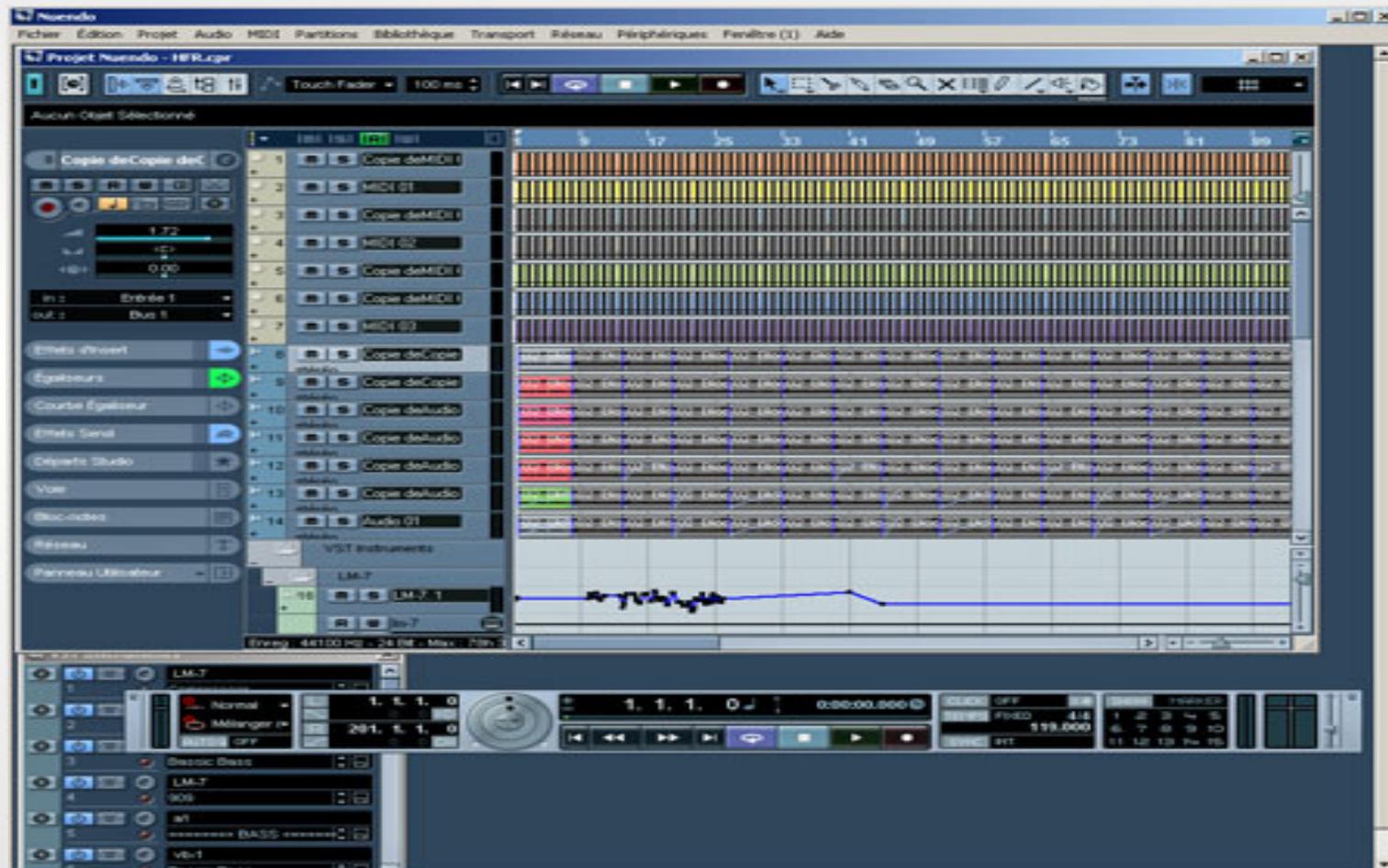
Programas de Edição de Áudio

■ Pro Tools



Programas de Edição de Áudio

■ Nuendo



Pedais e pedaleiras





GRVM

Compressão e formatos de áudio

Compressão e formatos de áudio

- O que é compressão de dados?
 - A ideia de compressão consiste na eliminação de dados redundantes
 - O primeiro passo para isso é a identificação da fonte de redundância
- Do que se trata compressão de áudio?
 - Da mesma forma, trata-se da eliminação de dados redundantes, compondo arquivos de áudio comprimidos

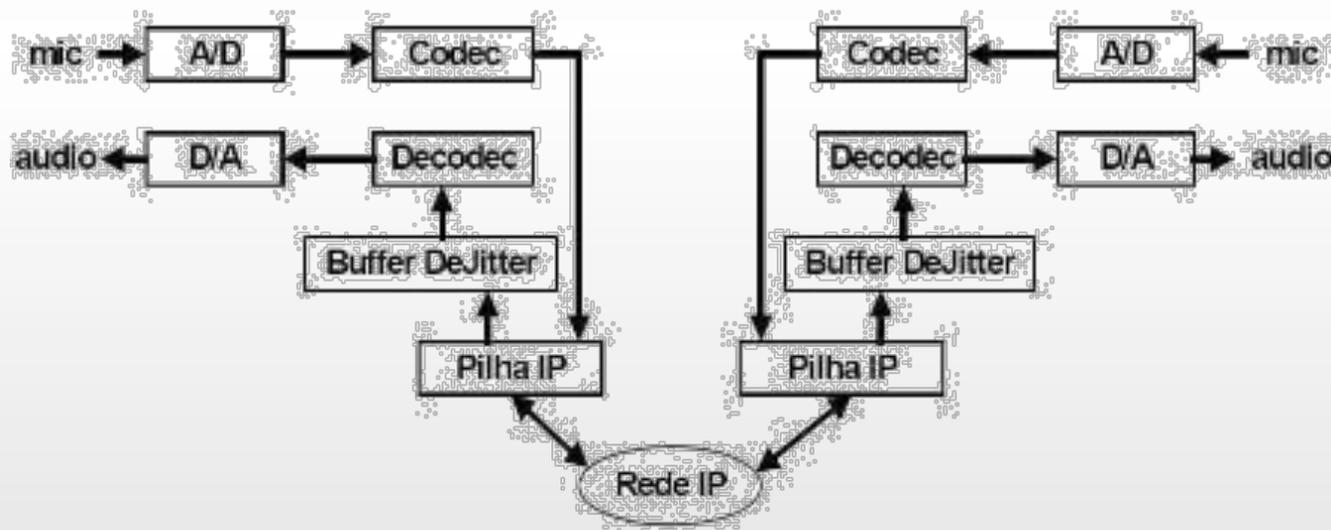


Compressão e formatos de áudio

- Exemplos:
 - Em determinado arquivo de áudio, um período com amostras de som com o mesmo valor, poderia ser representado de maneira breve com a indicação da repetição
 - É possível eliminar informações que são “julgadas ” pelo processo de compressão menos influentes na qualidade do som
- A compressão é composta por duas partes: a codificação e decodificação

Compressão e formatos de áudio

- Os métodos de compressão de som tendem a ser assimétricos
 - O processo de codificação pode ser mais sofisticado, complexo e lento
 - O processo de decodificação precisa ser mais rápido



Compressão e formatos de áudio

- Compressão sem perdas
 - Não há eliminação de informação na mensagem
 - Não tem, de maneira geral, um padrão de alta compressão (cerca de 2:1)
 - Todos os algoritmos necessitam de uma quantidade igual ou maior de bits à da entropia de recurso da informação para a compressão
- Compressão com perdas
 - Sob algum critério, determinadas informações são descartadas
 - Não dispensa a codificação eficiente
 - Grau de compressão x Distorção na mensagem

Compressão e formatos de áudio

■ Técnicas utilizadas na compressão sem perdas

■ Código de Huffman

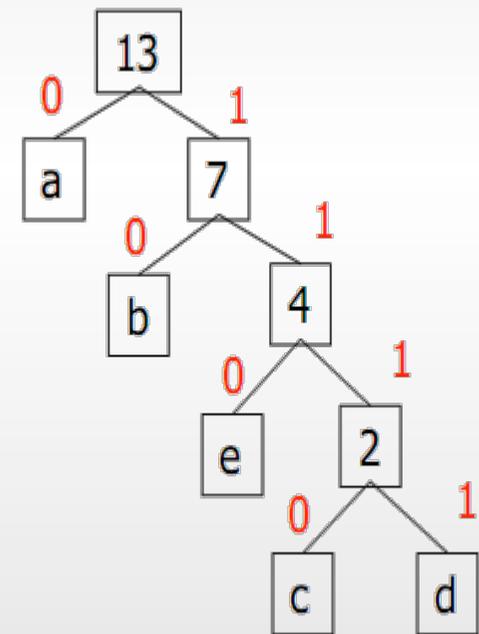
- É uma técnica de compressão de dados estatística
- Reduz o tamanho do código médio a fim de representar determinado alfabeto

■ Decodificação simples

■ Exemplo:

- Considerando a mensagem: aaaaaabbcbdee

símbolo	ocorrência	código
a	6	0
b	3	10
c	1	1110
d	1	1111
e	2	110



Compressão e formatos de áudio

- **Codificação aritmética**
 - Não depende da probabilidade como o código de Huffman
 - Utiliza um intervalo de números reais entre 0 e 1
 - Adiciona um cabeçalho à mensagem
 - Não é muito eficiente para mensagens pequenas
- **Algumas formas de compactação com perdas de áudio são:**
 - G.711 (Pulse Code Modulation)
 - ADPCM (Adaptive Delta Pulse Code Modulation)
 - MPEG3

Compressão e formatos de áudio

■ G.711

- Não é um algoritmo, é um padrão fundamental adequado para comunicação com voz
- Define a modulação por codificação de pulso (PCM - Pulse Code Modulation)
- É a representação digital de um sinal analógico
- Define dois algoritmos principais: μ -law (mais utilizado nos EUA) e o A-law (mais utilizado na Europa)
- Apresenta frequência padrão de 8000 amostras por segundo

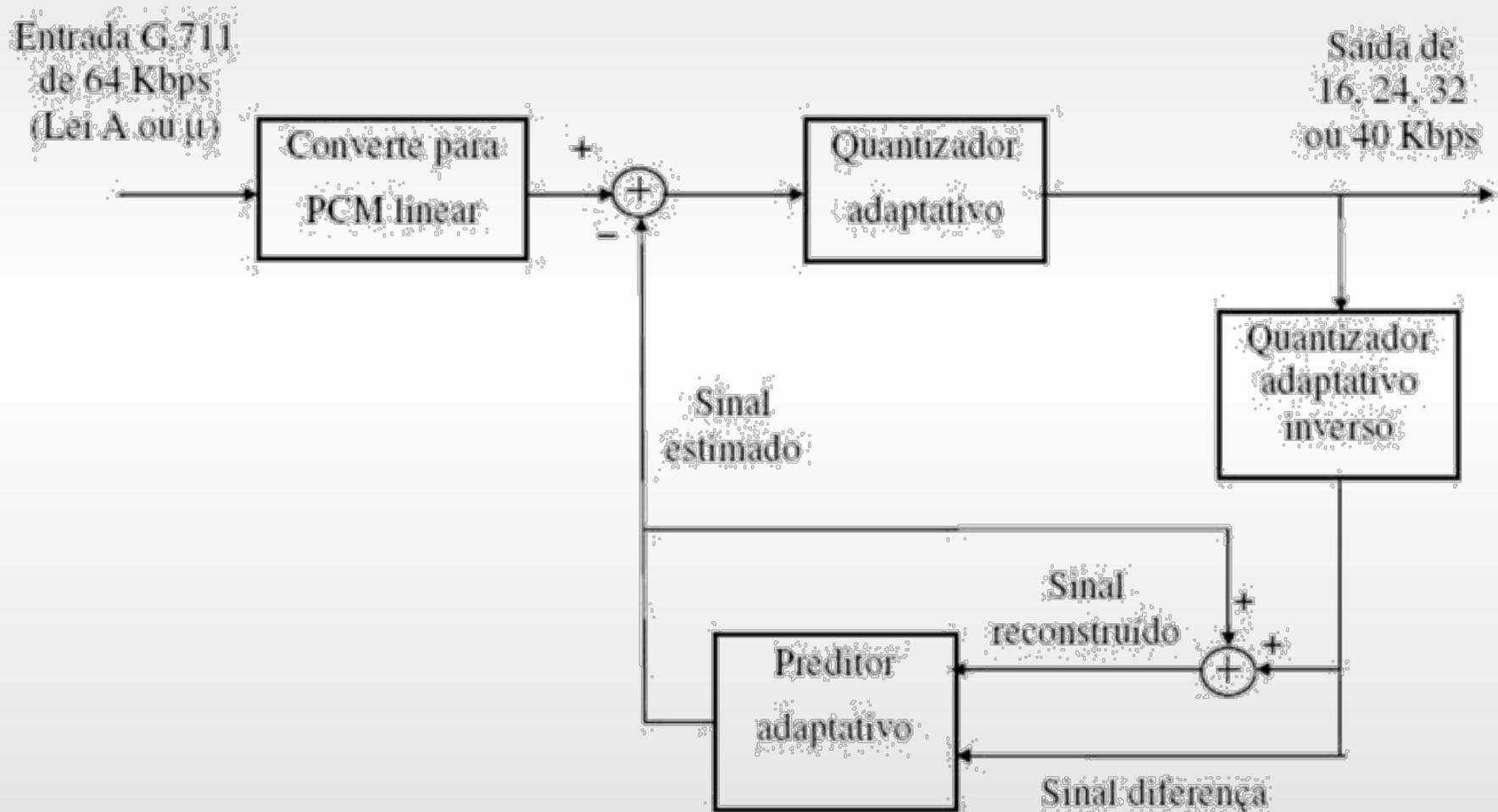
Compressão e formatos de áudio

■ ADPCM

- Comprime arquivos de entrada PCM
- A ideia principal é fazer uma previsão do valor da próxima amostra de áudio, com base nos valores anteriores, e expressar apenas as diferenças identificadas entre o previsto e o real
- Definido pelo padrão G.721 e estendido pelo padrão G.726
- Trabalha em taxas em torno de 16 a 40 Kb/s
- A forma mais comum de saída é 32Kbit/s

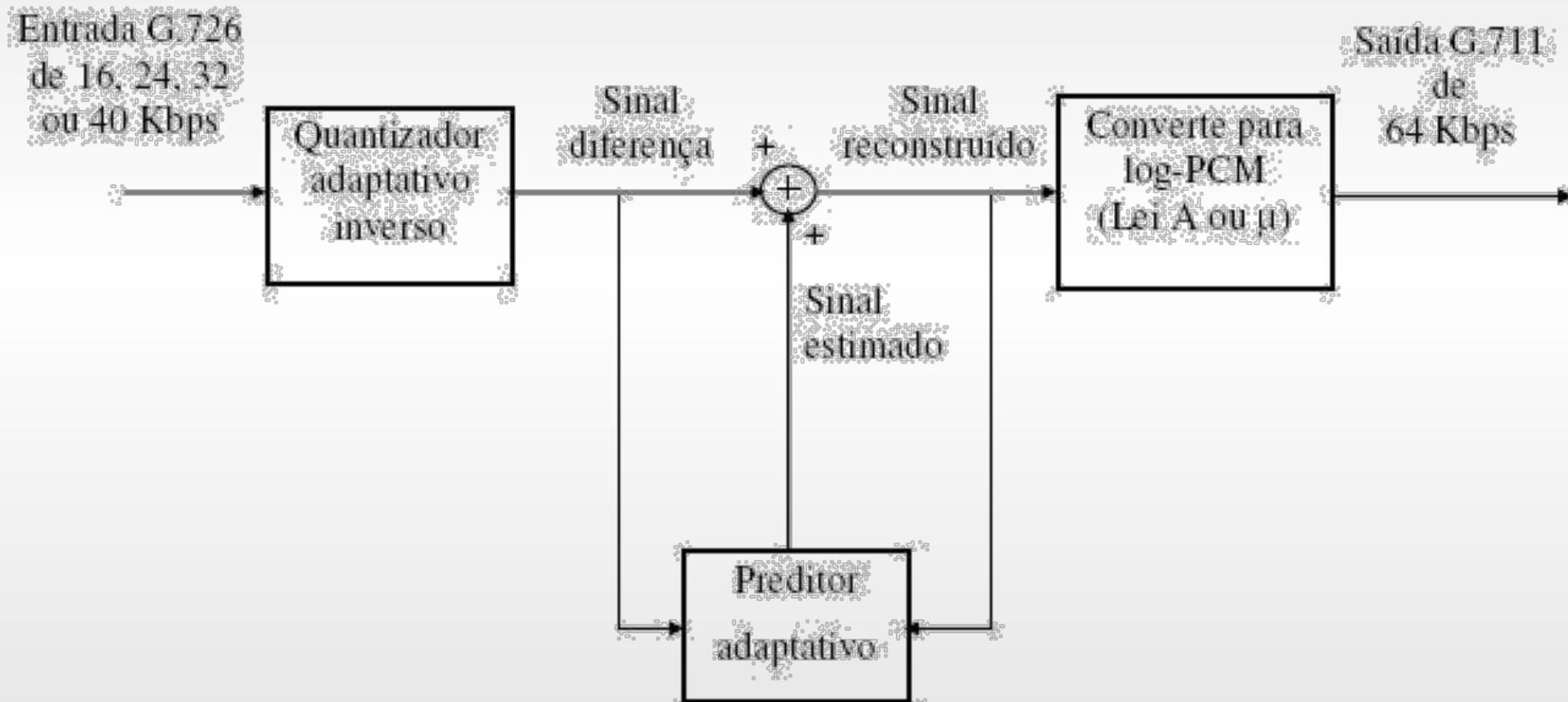
Compressão e formatos de áudio

■ Codificação ADPCM



Compressão e formatos de áudio

■ Decodificação ADPCM



Compressão e formatos de áudio

- O ADPCM possui qualidade semelhante ao G.711 (PCM)
- Quando comparado ao PCM, demonstra um consumo de banda reduzido

Codec	Codificação	Taxa (Kb/s)	Qualidade de voz (MOS)	Delay (ms)	Ano
G.711	PCM	64	4,1	0,75	1972
G.726	ADPCM	40, 32, 24 ou 16	3,85	1	1990

Compressão e formatos de áudio

■ MPEG-3 ou MP3

■ História

- Em 1970, o professor Dieter Seitzer da Universidade Erlangen-Nuremberg na Alemanha iniciou um grupo de pesquisa em codificação de áudio
- Em 1991, surgiu o ASPEC que evoluiu para o codec MP3 - MPEG-1 Layer 3
- Em 1995, o “.mp3” foi escolhido como extensão para arquivos MP3

Compressão e formatos de áudio

■ História

- O primeiro leitor de MP3 de sucesso foi criado em 1997 - AMP MP3 Playback Engine
- Esse leitor foi aprimorado (com uma interface) por dois estudantes, Justin Frankel e Dmitry Boldyrev, e recebeu o nome de Winamp em 1998
- Em 1999 apareceu o Napster, que permitiu a qualquer pessoa encontrar e fazer o download de músicas
- Ainda em 1999, apareceram os primeiros leitores MP3 portáteis

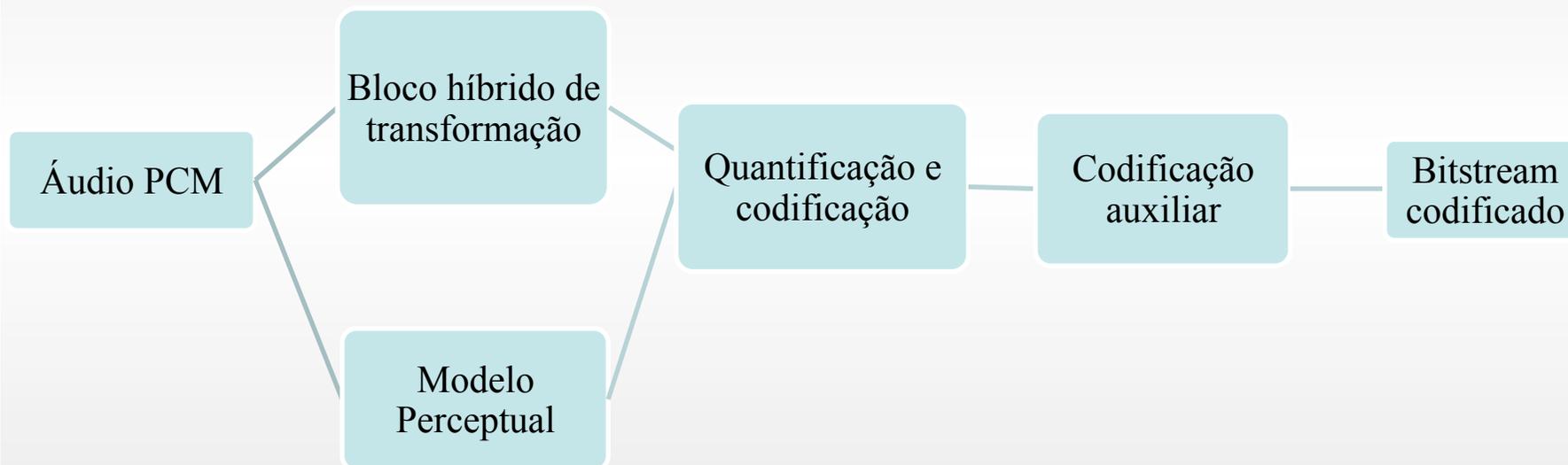
Compressão e formatos de áudio

■ MP3

- A proposta é eliminar frequências sonoras não captadas pelo sistema auditivo humano mas que incham arquivos de som
- Converter um arquivo WAV para MP3, por exemplo, promove uma degradação do som muito pequena não percebida pela maior parte das pessoas

Compressão e formatos de áudio

■ Codificação MP3



Compressão e formatos de áudio

- Bloco híbrido de transformação
 - Composto por duas etapas
 - Filtragem em sub-bandas
 - Transformada discreta modificada do cosseno (MDCT)
 - Filtragem de sub-bandas
 - O sinal original é separado em 32 frequências distintas
 - Essa técnica pretende isolar diferentes de frequência do sinal
 - Transformada discreta modificada do cosseno
 - Essa transformada é aplicada aumentando em 18 vezes a granularidade da divisão do sinal em sub-bandas
 - Sua operação é dependente do modelo psicoacústico (Modelo perceptual)

Compressão e formatos de áudio

■ Modelo Perceptual

■ Composto também por duas etapas

- FFT (Fast Fourier Transform)
- Modelo psicoacústico

■ FFT

- Ocorre em paralelo à filtragem de sub-bandas
- Prepara o sinal para o modelo psicoacústico

■ Modelo psicoacústico

- Determina quais partes do sinal é audível ou não
- Essa informação é utilizada no MDCT, para decidir seu modo de operação, e na Quantificação, para possibilitar a quantificação das linhas de frequência
- Esse modelo detecta os tons dominantes calculando para cada banda crítica um limite de "mascaramento"

Compressão e formatos de áudio

- **Quantificação e codificação**
 - As informações “irrelevantes” detectadas pelo modelo perceptual são descartadas de fato
 - O número de bits para cada banda é determinado
 - A codificação é feita através do Código de Huffman
- **Codificação auxiliar**
 - Todos os parâmetros gerados pelo codificador são utilizados para permitir que o decodificador reproduza o sinal sonoro
 - O sinal comprimido representa o sinal inicial PCM

Compressão e formatos de áudio

■ Decodificação

■ Composto por três etapas principais

■ Decodificação de seqüências de bits

- sincroniza a seqüência de bits codificada
- extrai os coeficientes de freqüência quantificados

■ Dequantificação

- reconstrói os dados a partir dos coeficientes de freqüência gerados pelo bloco da MDCT

■ Mapeamento freqüência - tempo

- constrói o sinal de saída áudio PCM a partir dos coeficientes dequantificados

Referências

- 1) Curtis Roads, The Computer Music Tutorial (Livro-texto), MIT Press. 1996. Cap 1
- 2) Bruce Bartlett, Introduction to Professional Recording Techniques. Howard W. Sams & Co. 1987
- 3) Wilson Guerreiro Pinheiro, Processamento de sinais de áudio. Notas de aula
- 4) Salomon, D., “A Concise Introduction to Data Compression” (2008) - Editora Springer
- 5) [http://wiki.sj.cefetsc.edu.br/wiki/index.php/MP3_\(Artigo_Completo\)](http://wiki.sj.cefetsc.edu.br/wiki/index.php/MP3_(Artigo_Completo))
- 6) Material de computação musical - <http://www.cin.ufpe.br/~musica/2010-1/>
- 7) Barbosa, Álvaro. Edição Digital de Som: Uma abordagem aos fundamentos da escultura sonora orientada para criadores. 1999
- 8) Izecksohn, Sérgio. Som Analógico e Som Digital. Revista Backstage. 2001