

Aula 07

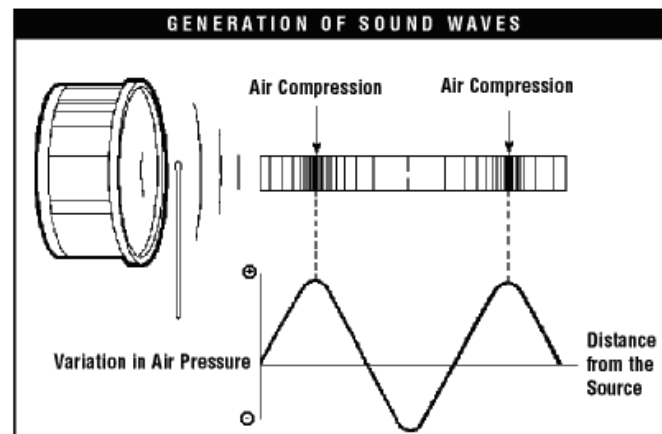
Conceitos Básicos de Áudio Digital

Diogo Pinheiro Fernandes Pedrosa

Universidade Federal do Rio Grande do Norte
Departamento de Ciências Exatas e Naturais
Ciência da Computação

Som – Conceitos Iniciais

- O som consiste em um fenômeno ondulatório → é macroscópico e envolve moléculas de ar sendo comprimidas e expandidas sob a ação de algum equipamento;
- Por exemplo: a vibração de um alto-falante produz uma pressão longitudinal no ar que é percebida como som.

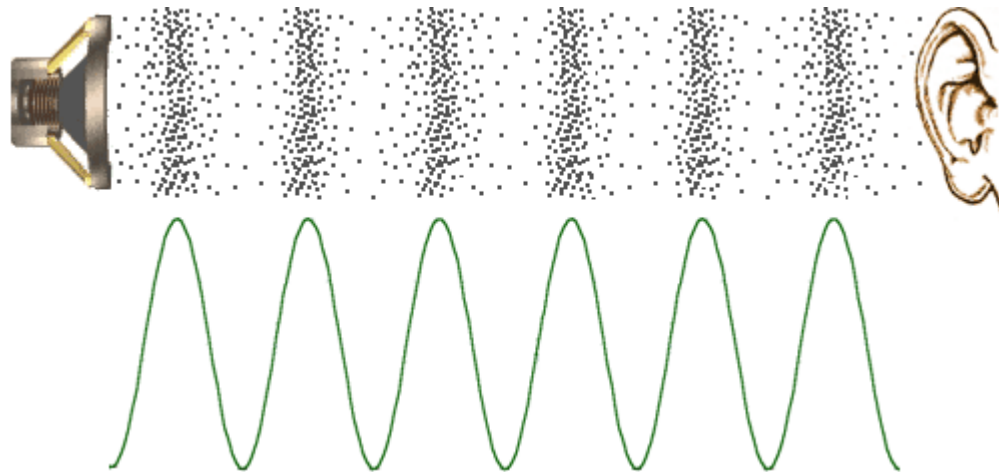


Som – Conceitos Iniciais

- Fato óbvio: sem ar não há som!
- Por ser uma onda de pressão sobre o ar, o som possui as características de **reflexão**, **refração** e **difração** → possibilidade de desenvolvimento de sistemas *surround*.
- O som assume valores contínuos;
 - Para representação digital do som, tem-se que adotar etapas de digitalização;

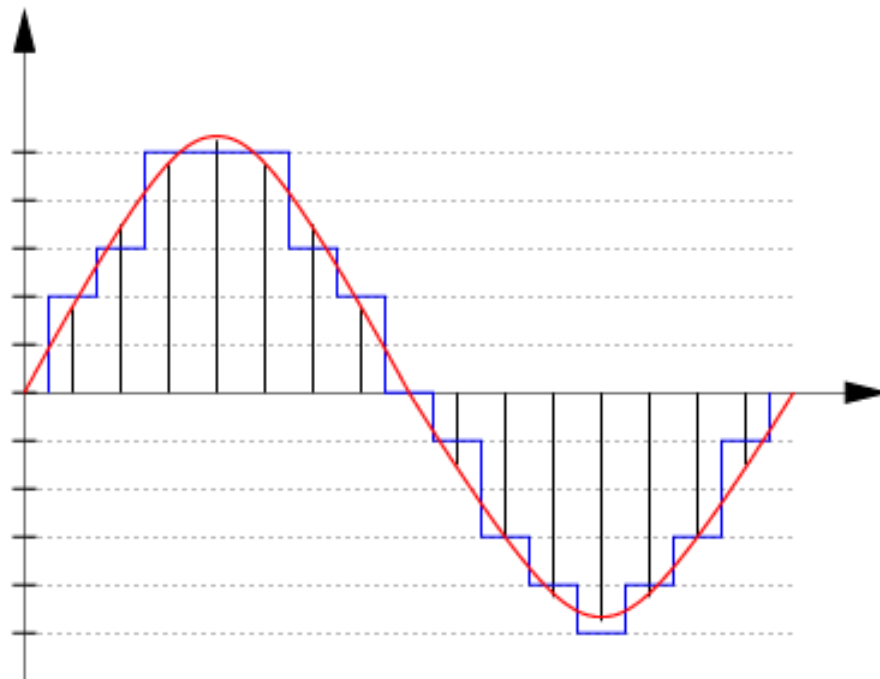
Som – Conceitos Iniciais

- A propagação do som pode ser representado por uma função unidimensional → senóide;
 - *A amplitude da onda varia no tempo;*



Digitalização do Som

- Digitalizar o som tanto na amplitude quanto no tempo;
- Transformar valores contínuos de amplitude em uma sequência de inteiros;



Amostragem →
Medir a quantidade
de interesse em
intervalos
espaçados.

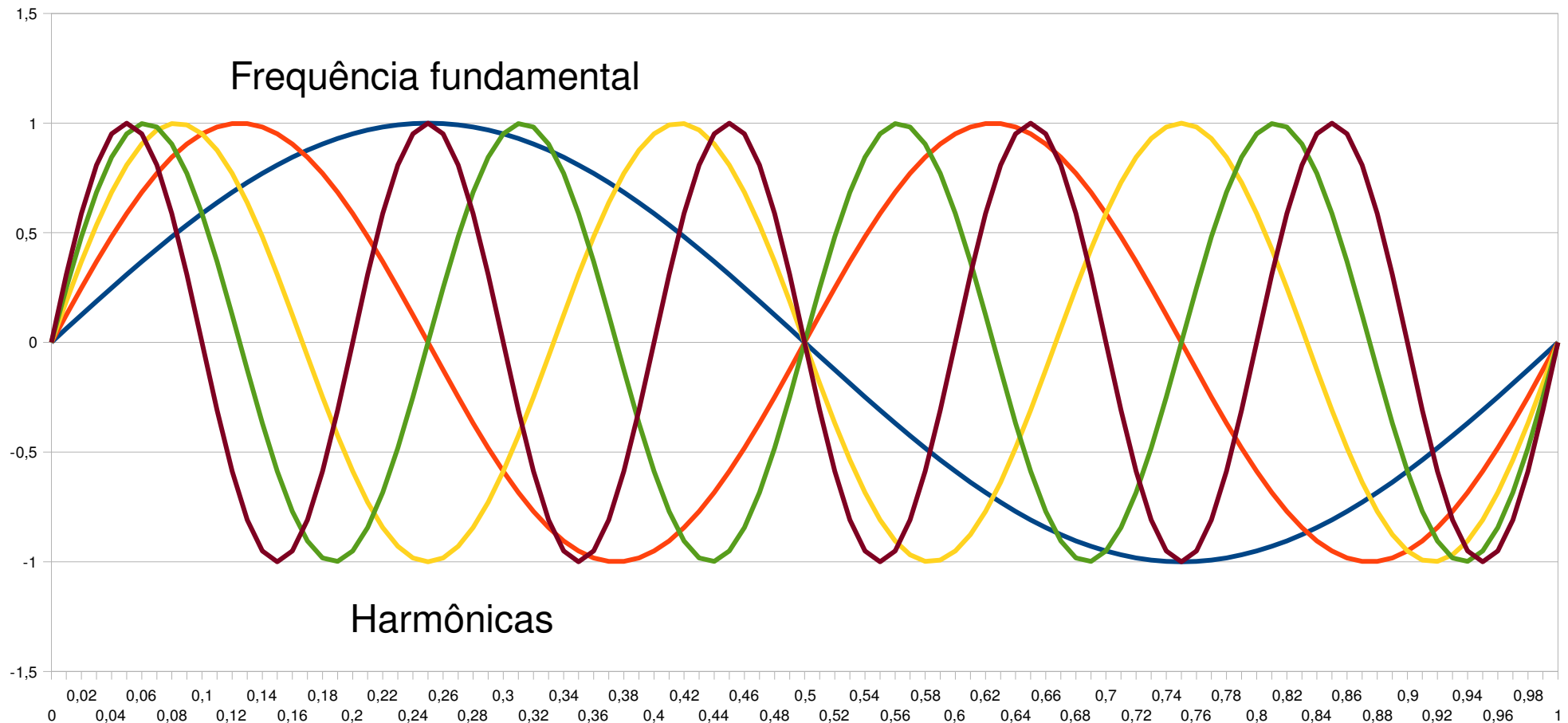
*Frequência de
amostragem* →
taxa com que a
amostragem é
realizada.

Digitalização do Som

- Em áudio, taxas típicas de amostragem variam de 8kHz até 48kHz;
- Amostragem da amplitude da onda → também é chamada de quantização;
 - A quantização pode ser uniforme (com intervalos igualmente espaçados) e não-uniforme;
 - Para quantização uniforme → valores em 8 bits (256 níveis diferentes) ou 16 bits (65.536 níveis diferentes).

Digitalização do Som

- Composição de onda: harmônicas



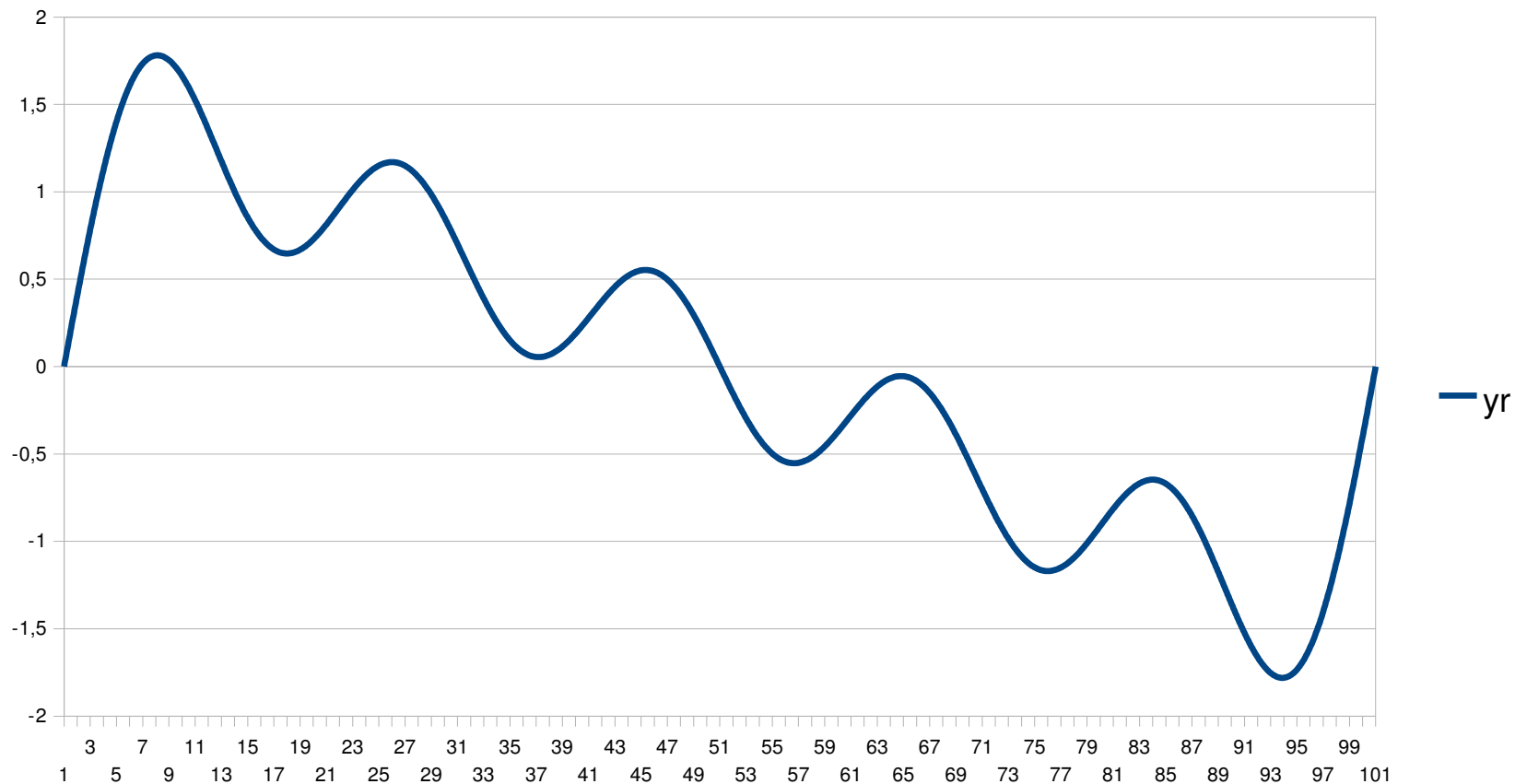
Digitalização do Som

- Composição de onda: combinação linear de senóides (harmônicas)

$$y_r = \text{sen}(f_f) + 0,5 \cdot \text{sen}(2 \cdot f_f) + 0,33 \cdot \text{sen}(3 \cdot f_f) + \\ + 0,25 \cdot \text{sen}(4 \cdot f_f) + 0,5 \cdot \text{sen}(5 \cdot f_f)$$

Digitalização do Som

- Composição de onda:



Digitalização do Som

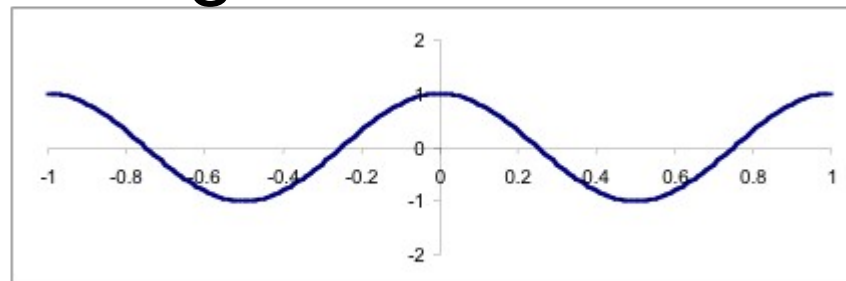
- Teorema de Nyquist:
 - A frequência de amostragem de um sinal contínuo, para que possa ser corretamente reproduzido no seu formato digital, deve ser maior ou igual ao dobro da maior frequência que compõem a onda;

$$f_a \geq 2 \cdot f_{\text{máx.}}$$

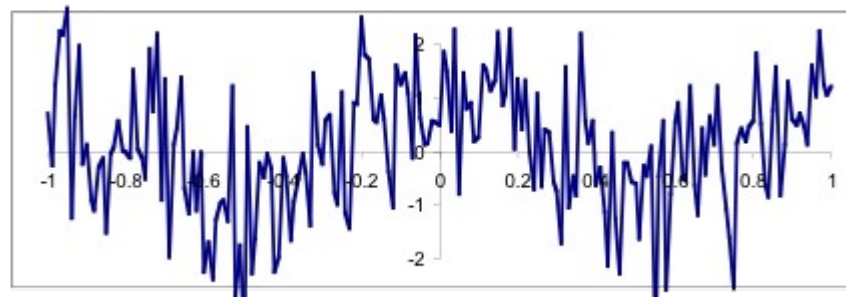
$$\Delta t_a \leq \frac{\Delta t_{\text{máx.}}}{2}$$

Relação Sinal-Ruído

- Sistemas analógicos estão sujeitos a sofrerem interferências de fontes aleatórias;
- Estas interferências fazem com que um *ruído* seja adicionado ao sinal, provocando distorções em sua amostragem.



(a)



(b)

Relação Sinal-Ruído

- Uma medida de qualidade do sinal é chamada de relação sinal-ruído (SNR);
- Essa relação é dada em decibéis (dB):

$$SNR = 10 \cdot \log \left(\frac{V_{\text{sinal}}^2}{V_{\text{ruído}}^2} \right)$$

Relação Sinal – Ruído

- Em som audível, a comparação é feita com um ruído de 1kHz:
 - Conversa normal: 60 dB;
 - Rua movimentada: 70 dB;
 - Limite do desconforto: 120 dB;
 - Limite da dor: 140 dB;
 - Danos ao sistema auditivo: 160 dB.

Relação Sinal/Quantização – Ruído

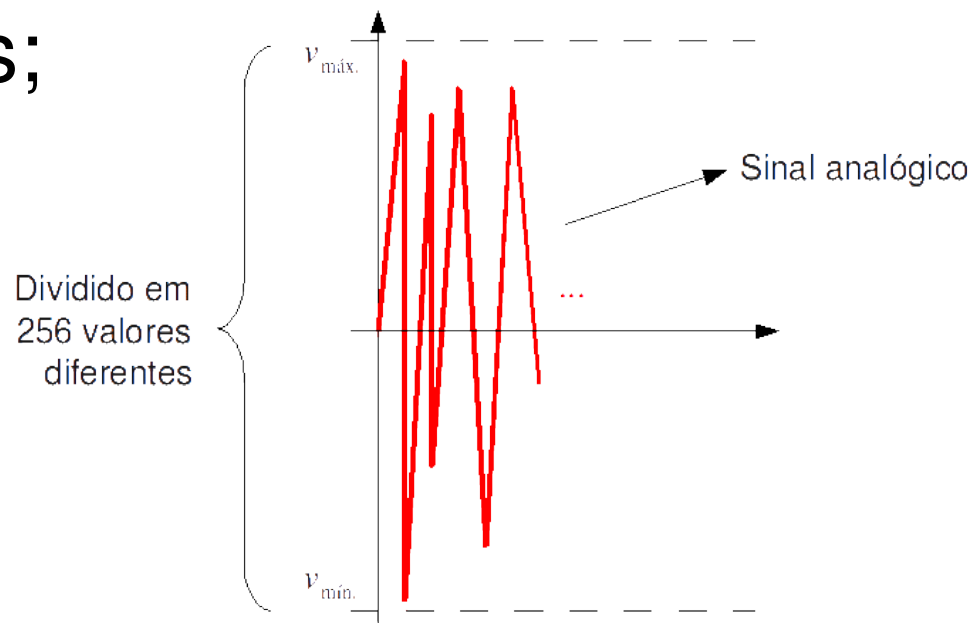
- Nível de qualidade de sinais analógicos → relação sinal – ruído;
- Em sinais digitais, apenas valores quantizados são armazenados:
 - Como medir a qualidade de um sinal digital?
- A precisão de cada amostra é determinada pelo número de bits por amostra → valores típicos: 8 bits ou 16 bits.

Relação Sinal/Quantização – Ruído

- Sinais digitais capturados a partir de um sinal analógico com ruído apresentam erros (o ruído também vai estar presente no valor quantizado);
- Outro tipo de erro em sinais digitais → erro de *quantização*;

Relação Sinal/Quantização – Ruído

- Se há apenas 8 bits para representar as amostras, então todo o intervalo de tensões que descrevem o sinal analógico é representado digitalmente por 256 valores diferentes;



Relação Sinal/Quantização – Ruído

- O erro de quantização é tratado como ruído porque de forma aleatória, de amostra em amostra → erro de arredondamento;
- A qualidade da quantização é caracterizada pela relação sinal/quantização – ruído (SQNR – *Signal-to-Quantization Noise Ratio*);
- Ruído de quantização → diferença entre o valor do sinal analógico e o valor da amostra, em um determinado instante de tempo.

Relação Sinal/Quantização – Ruído

- Cálculo da acurácia da quantização:
 - Considerando N bits por amostra, o intervalo dos sinais digitais é:

$$\left[v_{\text{mín.}}, v_{\text{máx.}} \right] = \left[-v_{\text{máx.}}, v_{\text{máx.}} \right] \Rightarrow \left[-2^{N-1}, 2^{N-1} - 1 \right]$$

Relação Sinal/Quantização – Ruído

- Valor do sinal por amostra:

$$\frac{2 v_{\text{máx.}}}{2^N} = \frac{v_{\text{máx.}}}{2^{N-1}}$$

Relação Sinal/Quantização – Ruído

- O pior caso é:

$$SQNR = 20 \log \frac{V_{\text{sinal}}}{V_{\text{ruído_quant.}}} = 20 \log \frac{2^{N-1}}{1/2} = 20 \cdot N \cdot \log 2 = 6.02 N \text{ (dB)}$$

Quantização Linear e Não – linear

- Quando as amostras obtidas da quantização consistem em vários intervalos com mesma largura, tem-se uma quantização linear;
- Problema:
 - Pequeno número de bits para representação;
- Alternativa:
 - Quantizar o sinal de forma não-linear, atendendo a uma propriedade da lei de Weber.

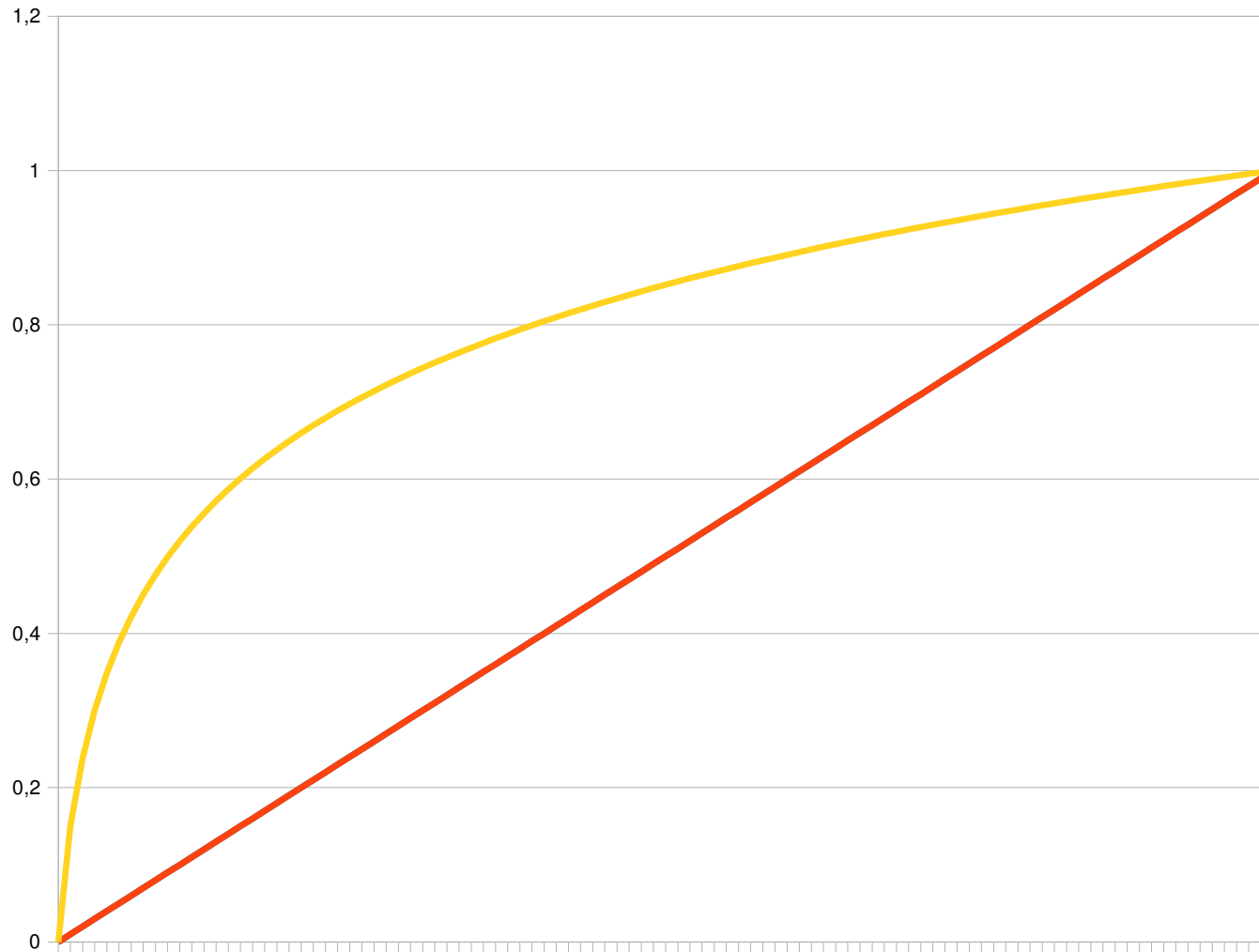
Quantização Linear e Não – linear

- *Lei de Weber*: descreve a relação entre as magnitudes físicas de um estímulo e a intensidade percebida deste estímulo:
 - Diferenças que são igualmente percebidas têm valores proporcionais em relação aos valores absolutos dos estímulos.
 - Exemplo: *se podemos sentir um crescimento no peso de 10 para 11 libras então, se ao invés, iniciarmos com 20 libras, então serão precisos 22 libras para sentirmos uma diferença.*

Quantização Linear e Não – linear

- Como o ouvido humano percebe melhor frequências mais baixas, pode-se usar uma quantização não uniforme para aumentar o número de amostras destas frequências mais baixas e diminuir o número de amostras das frequências mais altas;

Quantização Linear e Não – linear



Quantização Linear e Não – linear

- Quantização μ - law:

$$r = \frac{\text{sgn}(s)}{\ln(1 + \mu)} \cdot \ln \left(1 + \mu \left| \frac{s}{s_p} \right| \right)$$

$$\left| \frac{s}{s_p} \right| \leq 1$$

$\mu \rightarrow$ normalmente é igual a 100 ou a 255;

$s_p \rightarrow$ sinal de pico (maior valor do sinal);

$s \rightarrow$ valor atual do sinal

Quantização Linear e Não – linear

- Quantização *A-law*:

$$r = \begin{cases} \frac{A}{1 + \ln A} \left(\frac{s}{s_p} \right), & \text{se } \left| \frac{s}{s_p} \right| \leq \frac{1}{A} \\ \frac{\text{sgn}(s)}{1 + \ln A} \left[1 + \ln A \left| \frac{s}{s_p} \right| \right], & \text{se } \frac{1}{A} \leq \left| \frac{s}{s_p} \right| \leq 1 \end{cases}$$

$A \rightarrow$ normalmente tem valor 87,6.

Filtragem de Áudio

- Antes da amostragem e conversão analógica-digital (AD), o sinal de áudio é normalmente *filtrado* para remover frequências indesejadas;
- As frequências que são mantidas dependem da aplicação do sinal de áudio;
- O tipo de filtro mais comumente empregado é o filtro passa-banda (ou passa-faixa) → elimina frequências mais altas e mais baixas, deixando intacta apenas uma banda de frequências.

Filtragem de Áudio

- Quando um sinal digital é convertido para uma saída analógica (conversão DA), ele é novamente filtrado para eliminar frequências altas:
 - Um sinal quantizado é basicamente um sinal descontínuo. Teoricamente, entre uma descontinuidade e outra pode haver diversas frequências.
 - Usa-se um filtro passa-baixas → eliminar altas-frequências.

Outras Informações

- Largura de banda (em equipamentos analógicos) → conceitualmente consiste na parte de uma função de transferência de um equipamento que permanece aproximadamente constante, considerando o eixo x como frequências e o eixo y como esta função;
- A largura de banda é expressa em ciclos por segundo (Hz);

Outras Informações

- Para equipamentos digitais, a quantidade de dados que podem ser transmitidos em uma largura de banda fixa é expressa em *bits por segundo*;

Sons Sintéticos

- Necessidade: sons digitais devem ser convertidos para seu formato analógico para que possam ser ouvidos;
- Há duas abordagens:
 - Síntese por modulação da frequência;
 - Síntese por tabela de ondas.

Sons Sintéticos

- Síntese por modulação da frequência:
 - Tipo de síntese presente nas placas de som da *Creative*;
 - Consiste em gerar uma onda sonora a partir da manipulação da frequência de uma onda sonora, chamada de *portadora*, por uma outra onda, chamada de *moduladora*;
 - Portadora e moduladora são ondas descritas por funções senoidais;

Sons Sintéticos

- Síntese por modulação da frequência:
 - Matematicamente, a modificação da portadora é obtida pela inclusão da *função moduladora* no argumento da *função portadora*;

$$x(t) = A(t) \cos(\omega_p \pi \cdot t + I(t) \cos(\omega_m \pi \cdot t + \phi_m) + \phi_p)$$

Função “envelope”:
atribui maior ou menor
intensidade ao som

Frequência das
ondas portadora
e moduladora

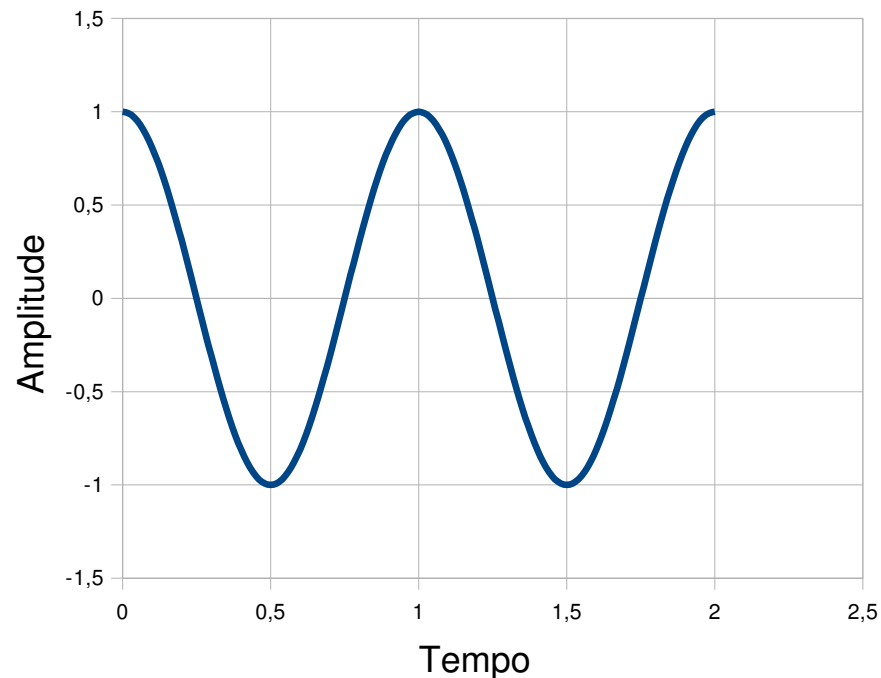
Função que cria
harmônicas

Constantes para
criar deslocamentos
no tempo

Sons Sintéticos

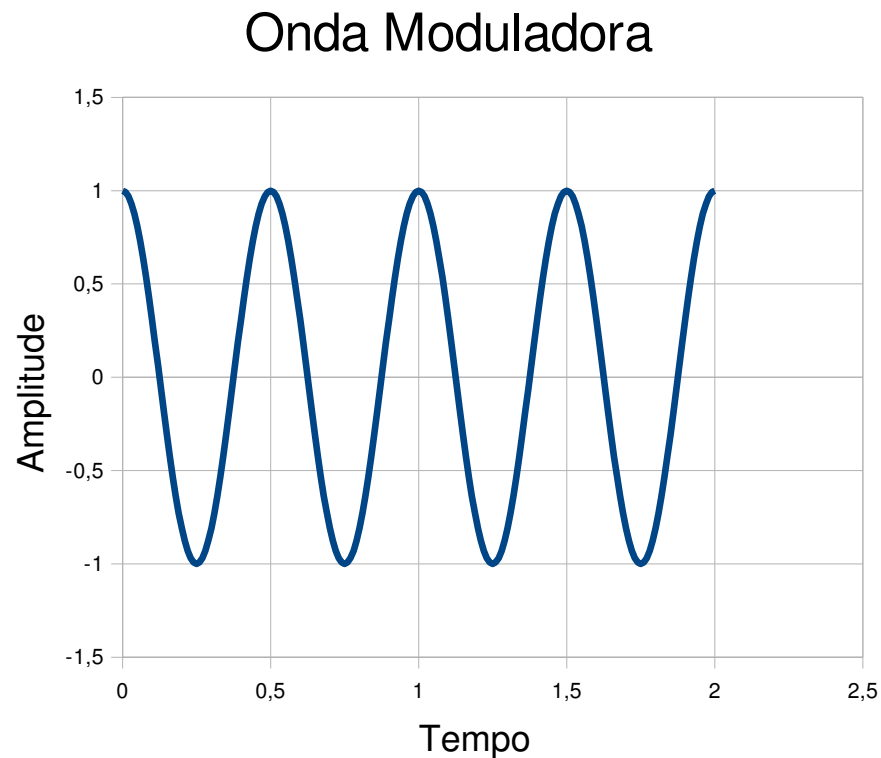
- Exemplo:
 - Onda portadora: $f_p = \cos(2\pi t)$

Onda Portadora



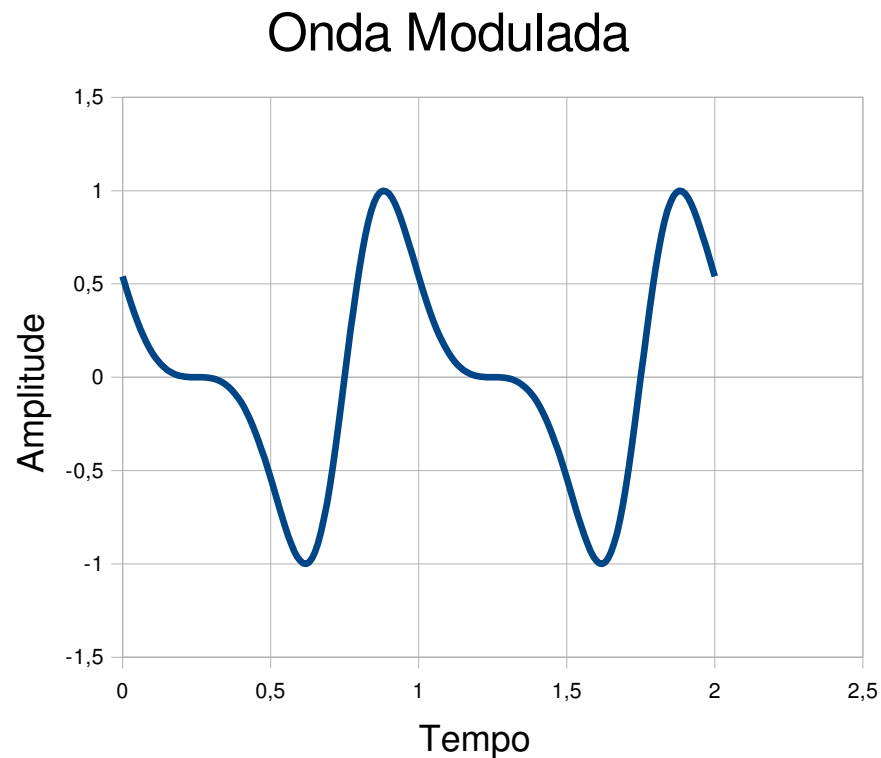
Sons Sintéticos

- Exemplo:
 - Onda moduladora: $f_m = \cos(4\pi t)$



Sons Sintéticos

- Exemplo:
 - Onda modulada: $x(t) = \cos(2\pi t + \cos(4\pi t))$



MIDI

- Acrônimo de *Musical Instrumental Digital Interface*;
- Consiste em um protocolo adotado para permitir a comunicação entre computadores, sintetizadores, teclados e outros equipamentos musicais;
- MIDI caracteriza-se como um *script* que codifica “eventos” (configurações de sons).

Quantização e Transmissão de Áudio

- Para se transmitir um sinal de áudio, ele precisa ser digitalizado;
- Essa digitalização ocorre após uma transformação no sinal → tal fato é conhecido como codificação;
- De uma forma geral:
 - Transformação → μ – law e A – law;
 - Quantização → Modulação por Código de Pulsos.

Quantização e Transmissão de Áudio

- As transformações promovem uma acentuação das amplitudes mais baixas do sinal de áudio:

$$r = \frac{\text{sgn}(s)}{\ln(1 + \mu)} \cdot \ln \left(1 + \mu \left| \frac{s}{s_p} \right| \right)$$

Sinal analógico com baixas amplitudes acentuadas

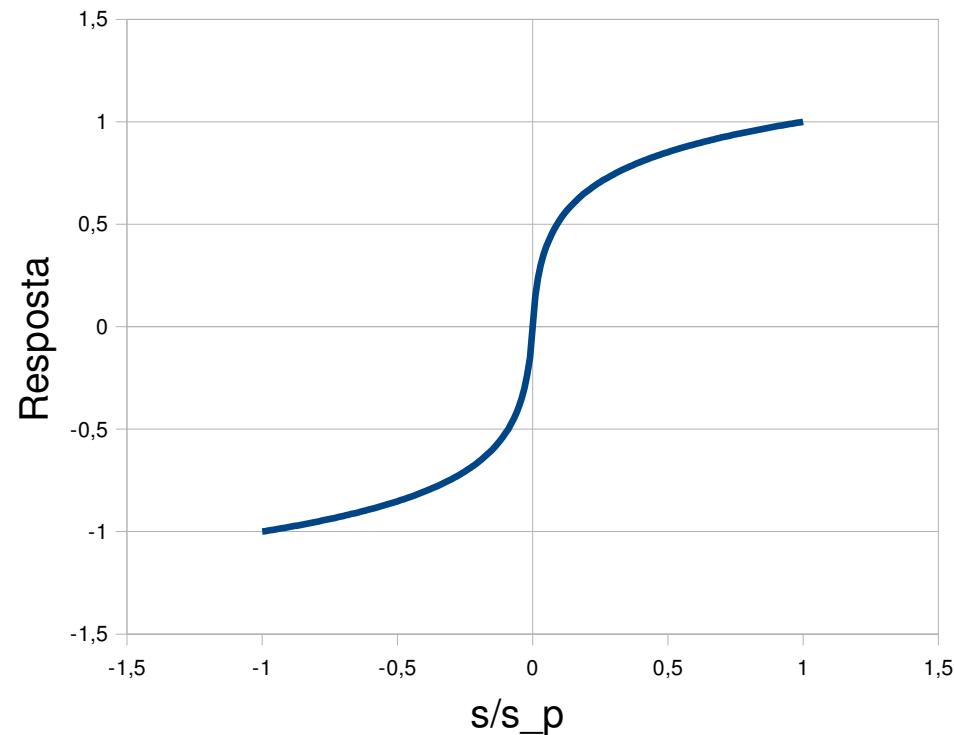
Sinal de pico (valor de maior amplitude)

Sinal lido no instante t

Quantização e Transmissão de Áudio

- As transformações promovem uma acentuação das amplitudes mais baixas do sinal de áudio:

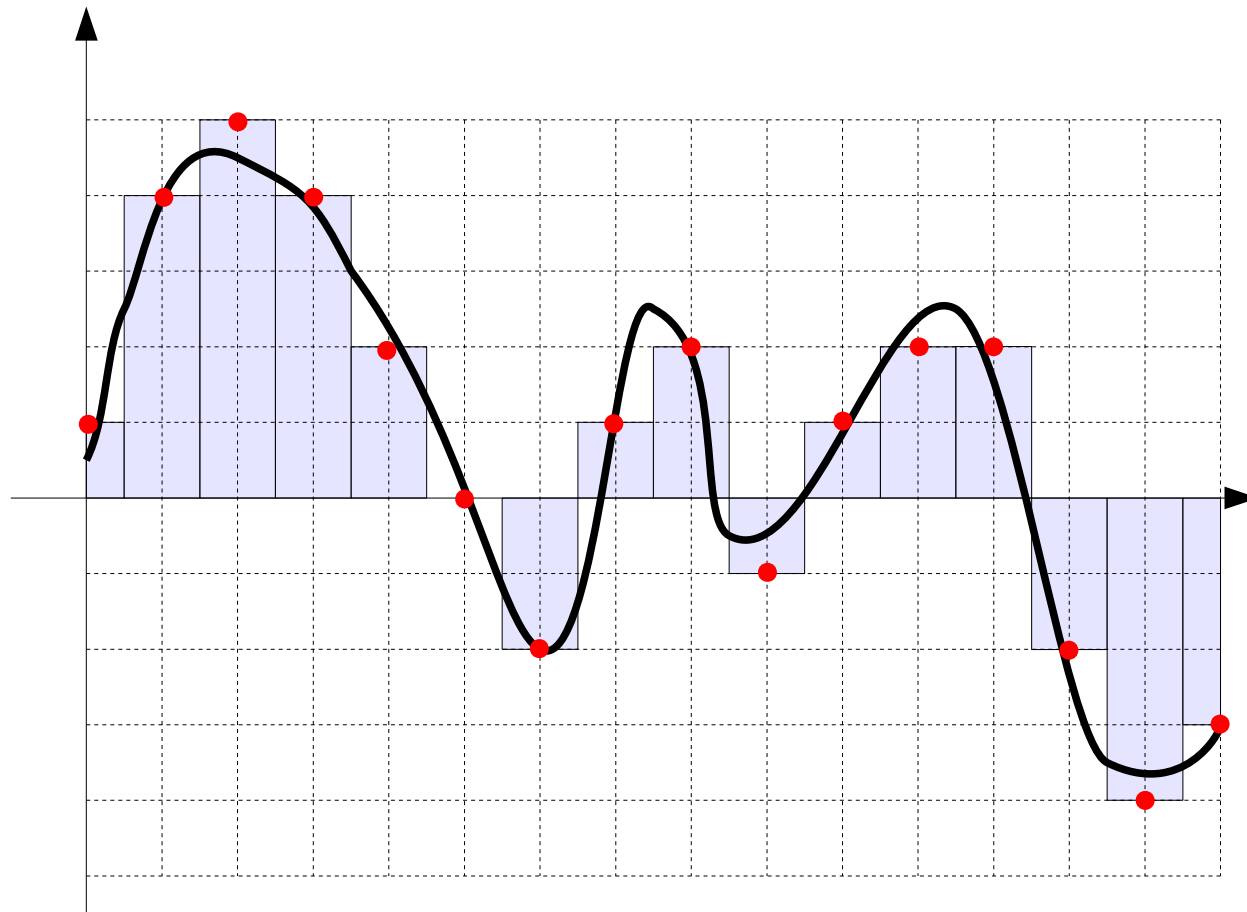
Transformação u-law



Quantização e Transmissão de Áudio

- *A Modulação por Código de Pulso* é a própria quantização → obter valores inteiros a partir de uma amplitude de um sinal analógico, em um instante bem determinado de tempo (a cada intervalo de amostragem);
- Vantagem do PCM → Tanto a amostragem (divisão ao longo do tempo) quanto à quantização (obtenção das amostras) são realizadas de maneira uniforme;

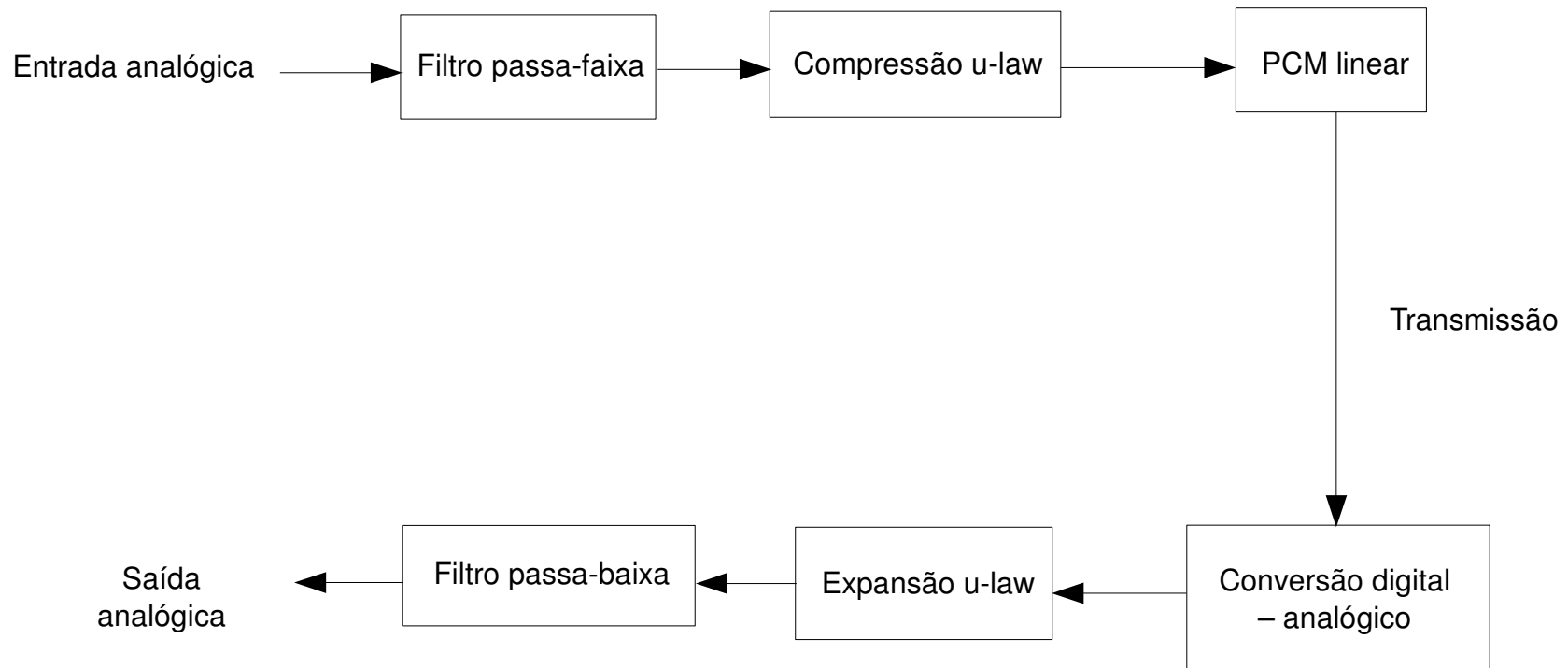
Quantização e Transmissão de Áudio



Quantização e Transmissão de Áudio

- Com a quantização ocorre perda de informação
→ pequeno número de níveis para representar o intervalo de variação da amplitude do sinal;
- Um outro tipo específico de codificação pode ser atribuído à sequência de bits que representa o sinal quantizado;

Quantização e Transmissão de Áudio



Codificação Diferencial de Áudio

- Após quantizado, normalmente o áudio não é armazenado em seu formato PCM mas em um formato que explora as diferenças entre valores de uma amostra em relação a outra;
- Vantagem da codificação diferencial: concentração dos valores quantificados em um intervalo específico, assim pode-se aplicar uma codificação mais curta para os valores.

Codificação Preditiva Sem Perdas

- Codificação Preditiva significa uma forma de transmitir diferenças entre os valores do sinal quantizado;
 - O valor transmitido consiste em:
 - Encontrar uma *previsão* do valor inteiro que representa o áudio no instante de tempo atual;
 - Encontrar o erro existente entre essa previsão e o valor inteiro no instante de tempo atual;
 - Transmitir esse erro.

Codificação Preditiva Sem Perdas

- Considerando f_n , com $n = 1, 2, 3, \dots, N$, uma sequência de valores quantizados, então:

$$\left\{ \begin{array}{l} \hat{f}_n = f_{n-1} \quad \text{ou} \quad \sum_{k=1}^{2 \text{ ou } 4} a_{n-k} f_{n-k} \\ e_n = f_n - \hat{f}_n \end{array} \right.$$

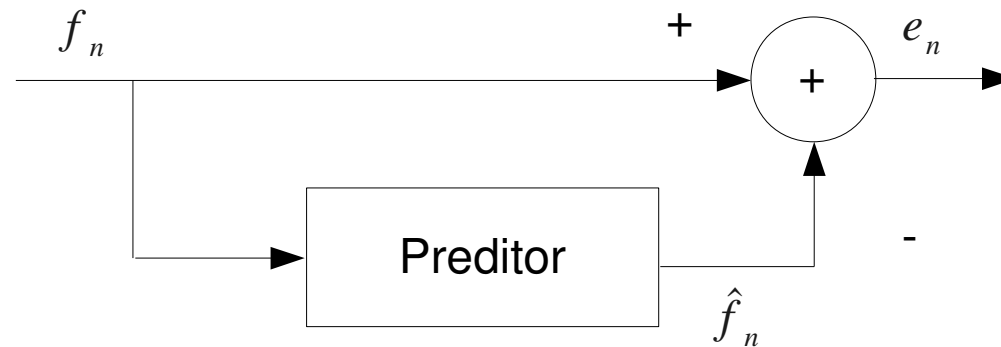
- Exemplo:
$$\left\{ \begin{array}{l} \hat{f}_n = \lfloor \frac{1}{2} (f_{n-1} + f_{n-2}) \rfloor \\ e_n = f_n - \hat{f}_n \end{array} \right.$$

Codificação Preditiva Sem Perdas

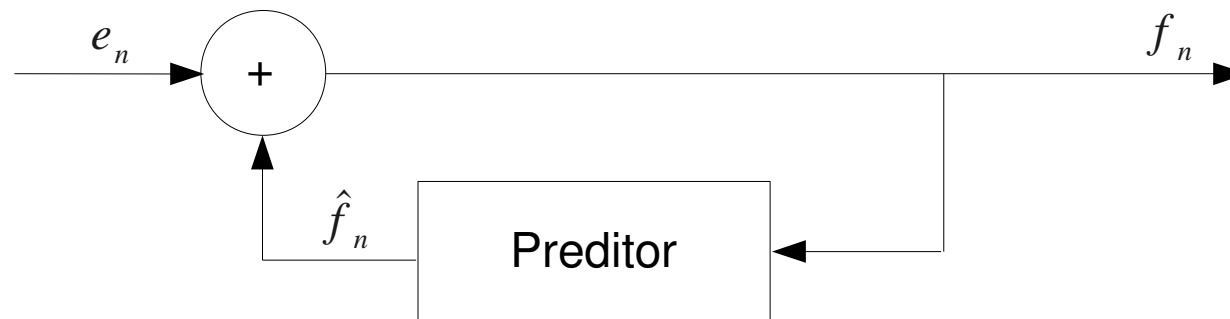
- O termo *sem perdas* vem do fato de que o sinal é totalmente recuperado sem perda alguma;

Codificação Preditiva Sem Perdas

Codificador de Diferenças



Decodificador de Diferenças



Codificação Preditiva Sem Perdas

- Outras codificações preditivas:
 - DPCM – *Differential Pulse Code Modulation*: funciona quase do mesmo modo que a transmissão PCM, contudo incorpora uma etapa adicional de quantização no erro;

Codificação Preditiva Sem Perdas

- Outras codificações preditivas:
 - DM – *Delta Modulation*: versão simplificada do DPCM;

$$\left\{ \begin{array}{l} \hat{f}_n = \tilde{f}_{n-1} \\ e_n = f_n - \hat{f}_n = f_n - \tilde{f}_{n-1} \\ \tilde{e}_n = \begin{cases} +k & \text{se } e_n > 0, \text{ onde } k \text{ é uma constante} \\ -k & \text{caso contrário} \end{cases} \\ \tilde{f}_n = \hat{f}_n + \tilde{e}_n \end{array} \right.$$

Codificação Preditiva Sem Perdas

- Outras codificações preditivas:
 - ADPCM – *Adaptive DPCM*: utiliza algoritmo para encontrar os valores dos coeficientes que definem o valor predito (em uma combinação linear) segundo algum critério de otimização.

$$\hat{f}_n = \sum_{k=1}^4 a_k \tilde{f}_{n-k}$$