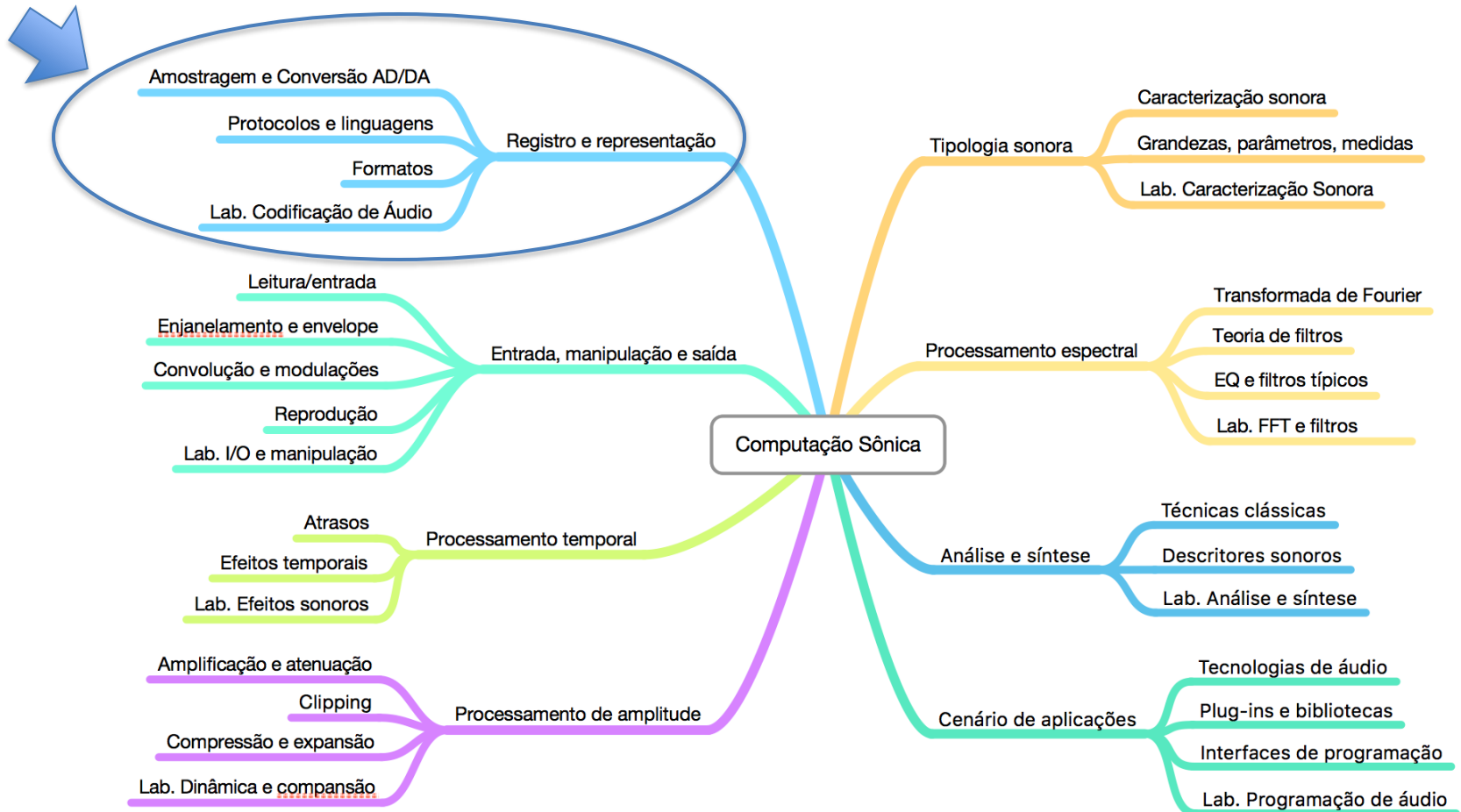


Som analógico e som digital

Prof. Regis Rossi A. Faria



Mapa situacional



Organização do módulo

- Bibliografia recomendada
- Contextualização, teoria e exemplos
- Tópicos fundamentais
 - Ondas acústicas e o som contínuo
 - Som analógico e som digital
 - Teorema da amostragem e a conversão AD/DA
- Tópicos avançados
 - Representação temporal e espectral
 - Formatos de áudio digital
 - Bases e sistemas numéricos de representação
 - Conversores AD/DA
- Revisão e discussões finais

Tópicos abordados no módulo

- Para viabilizar o registro e a posterior reprodução de som dispomos de um legado de tecnologias que trabalham o som na sua forma analógica e na sua forma digital
- Neste módulo abordaremos os princípios básicos do som analógico, do som digital (discreto), os processos de conversão entre analógico e digital e uma visão geral dos formatos de som digital disponíveis

Tópicos abordados no módulo

- Veremos
 - Som analógico: ondas acústicas e a representação contínua do som
 - Som digital: teorema da amostragem e representação do som na forma discreta/digital
 - Conversão analógico-digital (ADC) e digital-analógico (DAC)
 - Sistemas numéricos e formatos de áudio digital (PCM e outros)
 - Razões para se trabalhar com sinais digitais

Bibliografia

- Recomendada
 - Roads, C. (1996). The Computer Music Tutorial. Cambridge, MA: The MIT Press.
 - Pohlmann, K. C. (1995). Principles of digital audio. New York: McGraw-Hill.
 - Russ, M. (1996). Sound synthesis and sampling. Oxford: Focal Press.
- Complementar
 - Chadale, J. (1997) Electric Sound: The Past and Promise of Electronic Music. New Jersey: Prentice Hall.
 - Moscal, T. (1994). Sound Check: The Basics of Sounds and Sound Systems. Milwaukee Hal Leonaard Corporation.
 - Vale, S.d. (1997). Microfone: Tecnologia e Aplicação. Rio de Janeiro: Editora Música & Tecnologia.
- Avançada (referência)
 - Oppenheim, A. V., Schaffer, R. W., Buck, J. R. (1999). Discrete-Time Signal Processing, Prentice Hall, 2a Ed.

Bibliografia online

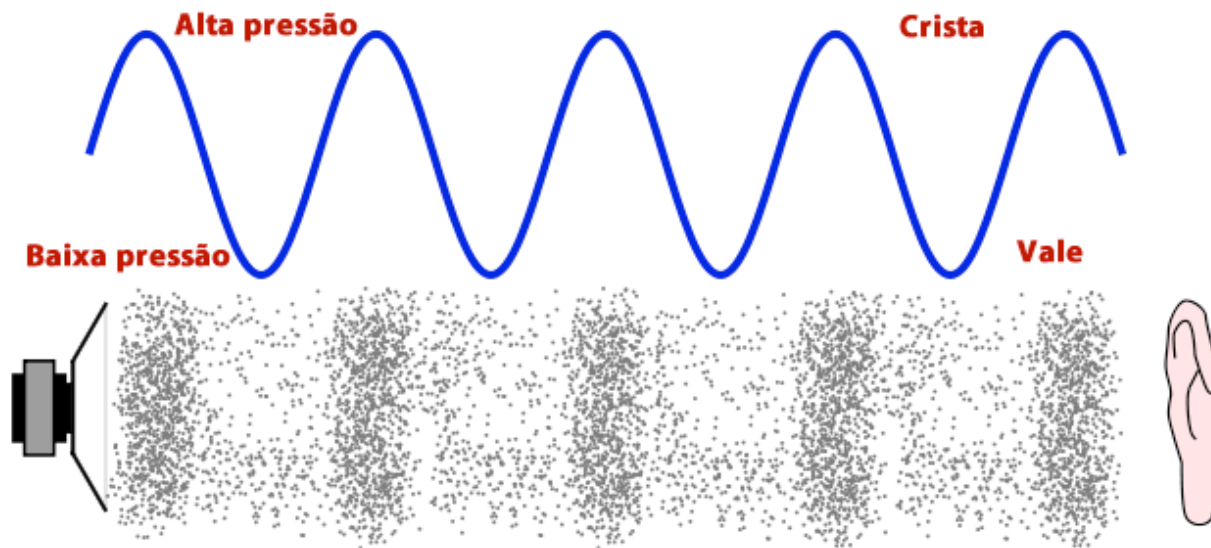
- Julius O. Smith III, Mathematics of the Discrete Fourier Transform (DFT) with Audio Applications
<http://ccrma.stanford.edu/~jos/mdft/>
- Rocchesso, D. (2003). Introduction to Sound Processing. Accesso online: <http://www.scienze.univr.it/~rocchess>. ISBN 88-901126-1-1
- Martin, Geoff (2006). Introduction to Sound Recording. Accesso online: <http://www.tonmeister.ca/main/textbook/>.
- Johannes Kreidler (2009). Programming Electronic Music in Pd. <http://www.pd-tutorial.com/english/ch03s04.html>

Ondas acústicas e o som contínuo

Conceitos da representação do som no formato contínuo

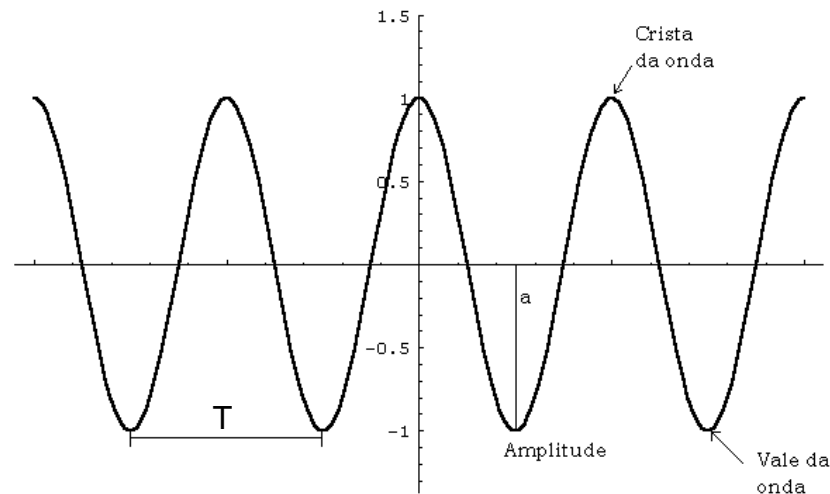
Representação das ondas acústicas

- As ondas acústicas são ondas mecânicas que se propagam no ar
- Tradicionalmente representamos as ondas por meio do registro da variação de alguma grandeza física ao longo do tempo (e do espaço) como por exemplo a *pressão acústica*



Representação das ondas acústicas

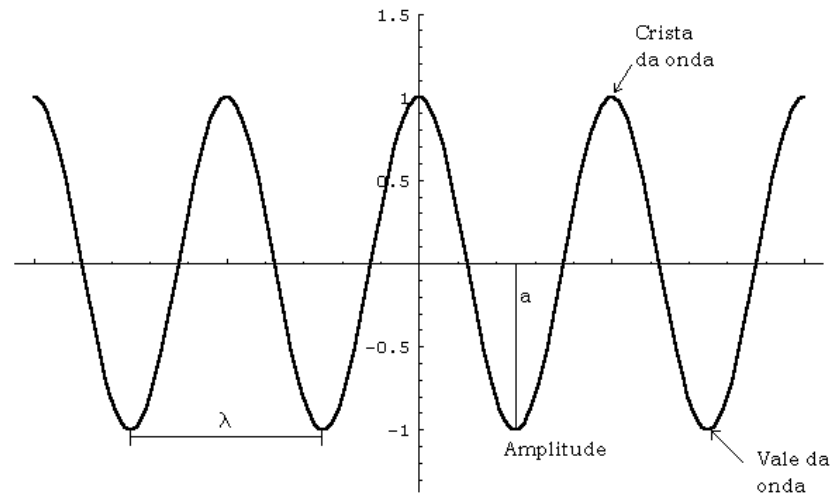
- Graficamente, a sucessão de valores no tempo dá origem a uma **forma de onda contínua**
- Usualmente sendo uma forma oscilante periódica, identificamos vales (pontos de mínimo valor) e cristas (pontos de máximos) da onda, e podemos estimar o período T da onda descrita no tempo, obtido entre pontos de mesma amplitude em ciclos vizinhos



Exemplo de uma forma de onda contínua, no caso: um som de uma só frequência (tom puro, onda senoidal)

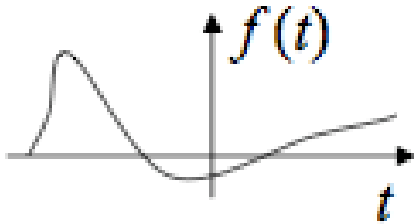
Representação das ondas acústicas

- Sendo a representação gráfica no espaço, podemos estimar o comprimento λ (em metros) da onda descrita no espaço, pela distância entre pontos de mesma amplitude



Representação das ondas acústicas

- Em cada momento do tempo a onda apresentará um valor de amplitude, como numa função matemática:



$f(t)$ = valor da amplitude de f no instante t

- Esta representação equivale a uma tabela (de infinito comprimento) com duas colunas:

t	$f(t)$
0	0
0.1	0.5
0.2	1.2
...	...

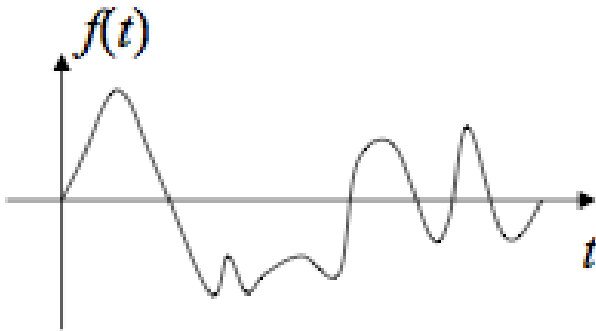
Representação temporal e espectral

Conceitos da representação do som no
domínio do tempo e da frequência

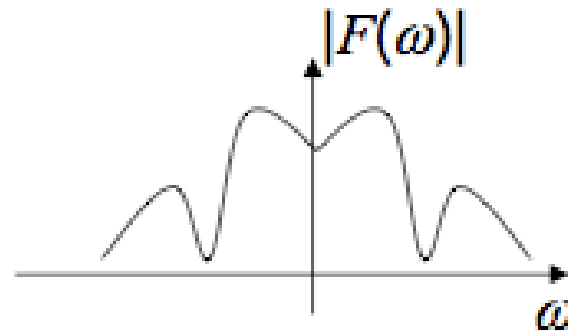
Representação no tempo e frequência

- Os sinais de áudio podem ser representados
 - No *domínio do tempo*: o sinal é uma função no tempo
 - No *domínio da frequência*: o sinal é representado por seu espectro

No tempo



Na frequência



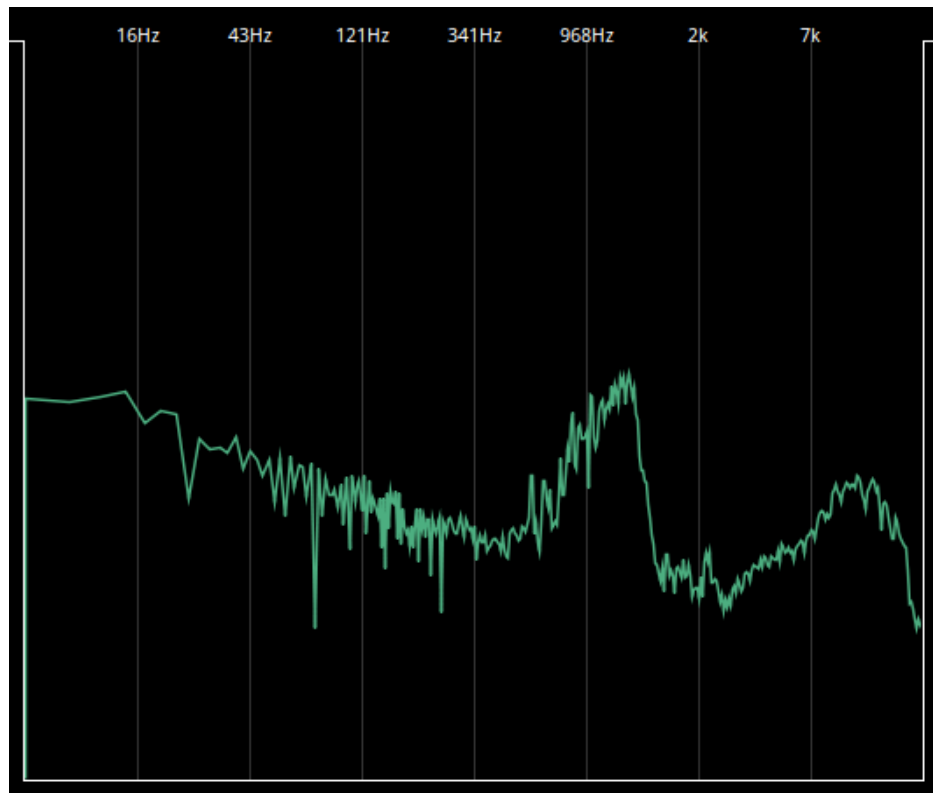
Representação no tempo e frequência

- A representação no domínio da frequência dá origem a *espectros*, *espectrogramas*, gráficos que mostram a intensidade relativa (ou absoluta) de cada frequência contida no sinal, na faixa que está sendo analisada

Espectro do som de gota de água

Note:

- Amplitude relativa das frequências
- Faixa de frequências em análise



Representação no tempo e frequência

- Representações sonoras nos planos *tempo vs. amplitude* e *frequência vs. amplitude* permitem caracterizar o comportamento do sinal no tempo e sua composição de frequências

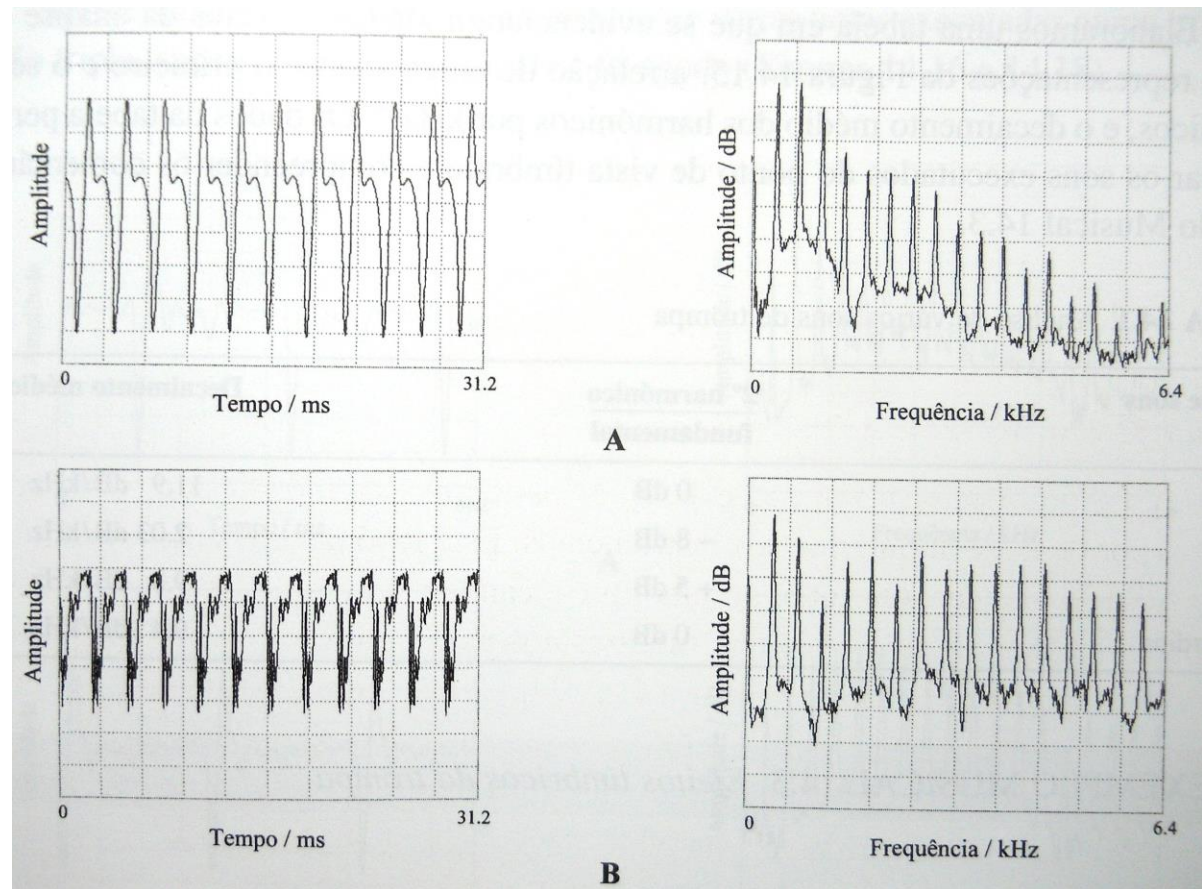


Fig. Sinal de uma nota de trompa emitindo um Fá3 e seu espectro de frequências

Tempo *versus* frequência

- Diversos tipos de processamento sonoro operam sobre o seu espectro, o que exige primeiro que se transforme o sinal do domínio temporal para o domínio da frequência
 - *Vantagem: no domínio da frequência é mais fácil a realização de muitos processamentos sobre o sinal do que no domínio do tempo*
- Para tanto dispomos de ferramentas que permitem converter de uma representação a outra, conhecidas como transformadas, dentre as quais se destaca a *Transformada de Fourier*
- Este assunto é abordado no estudo da representação e da análise espectral

Som analógico

Uma representação contínua da onda
acústica

Som analógico

- Na gravação a *onda acústica* variante no tempo é transformada por um transdutor em uma *onda elétrica* variante no tempo, com sua amplitude medida em Volts (unidade de tensão elétrica)
- A onda elétrica equivale a uma representação análoga à onda acústica (daí o termo analógico)
- Em sua forma elétrica, esta onda pode ser registrada, processada e transmitida por meios de circuitos elétricos convencionais, e posteriormente convertida em onda acústica pelos alto-falantes (transdutores eletro-mecânicos)

Operações e dispositivos

- Operações típicas e dispositivos realizadas com som no formato analógico (sinal contínuo no tempo)

Operação	Dispositivos usados
Gravação	Captação por microfones, transdução em informação elétrica Registro, em meio eletromagnético
Distribuição	Por meio de redes elétricas analógicas Ex: circuitos de uma mesa mixadora, de um pré-amplificador, etc
Processamento	Circuitos analógicos eletro-mecânicos e/ou eletrônicos Ex: Filtros passivos usando resistores, capacitores e indutores (RLC)
Reprodução	Circuitos transdutores eletro-mecânicos

Transformando o sinal analógico

- Em sua representação elétrica, o som pode sofrer várias transformações importantes usando-se os bipolos elementares passivos e ativos:
 - Capacitores (C) (uso em defasadores, filtros)
 - Resistores (R) (uso para atenuação)
 - Indutores (L) (uso em defasadores e filtros)
 - Fontes de tensão (V) e de corrente (I) (geradores de formas de onda)

Vamos vê-los...

Circuitos eletrônicos típicos



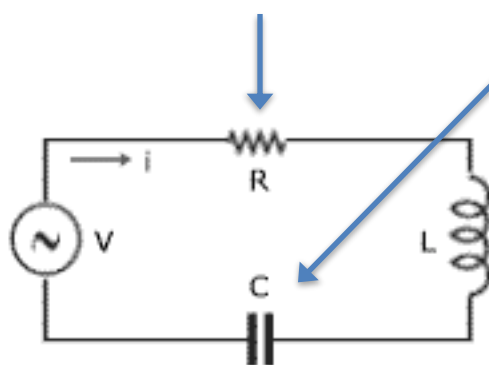
Resistor



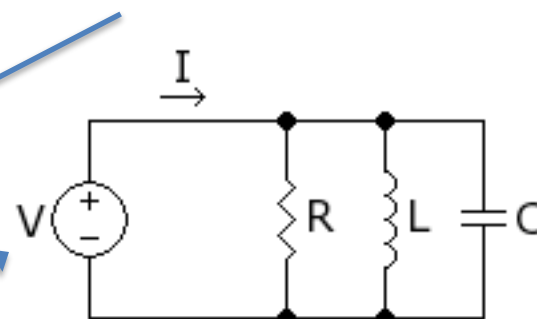
Capacitores



Indutores



**Circuito RLC
série**



Fonte
(de tensão V)

**Circuito RLC
paralelo**

Transformando o sinal analógico

- Redes de circuitos permitem ainda
 - separar sinais,
 - misturar sinais
 - princípio da superposição: comportamento global é igual à soma dos comportamentos individuais
 - linearidade: $R(t) = a.f(t) + b.g(t)$
 - gerar várias saídas de sinais distintas com diferentes características de intensidade, frequências e fases

Som analógico

- Vantagens
 - Simplicidade dos circuitos básicos (RLC) e transdutores
 - Sinal acessível sem necessidade de DAC e ADC
- Desvantagens e limitações
 - Se é desejada precisão alta → custo elevado
 - Sujeito a interferências de todo tipo
 - É impossível realizar certas operações e processamentos no formato analógico (como filtragens com faixa de transição muito estreitas e linhas de atrasos de fase muito longas)

Som digital

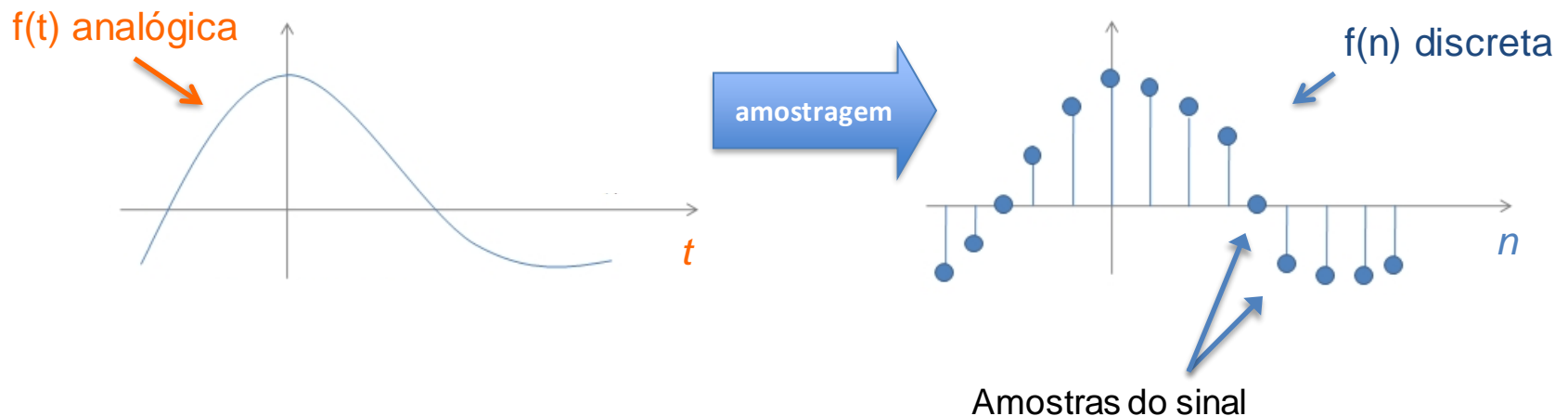
Uma representação discreta da onda
acústica

Som digital

- Para ser processado por computadores e sistemas digitais os sinais sonoros precisam ser reduzidos a sequências de amostras, isto é, precisam ser *discretizados*
- A operação básica que transforma um sinal analógico em um discreto é chamada de amostragem
- Em cada intervalo de tempo T_a uma amostra do sinal é capturada
- A frequência de amostragem será $f_a = 1/T_a$

Digitalização do som

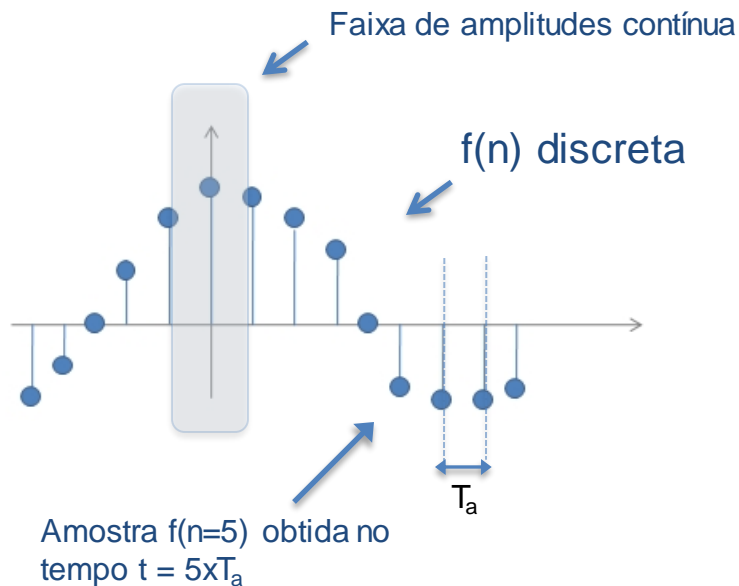
- Formas de onda analógicas podem ser amostradas no tempo gerando sequências numéricas que representam a amplitude instantânea do sinal analógico $f(t)$ em determinados momentos



- Estas sequências são então representadas no tempo discreto indexado por $n \in \mathbb{Z}$ e não mais no tempo contínuo $t \in \mathbb{R}$

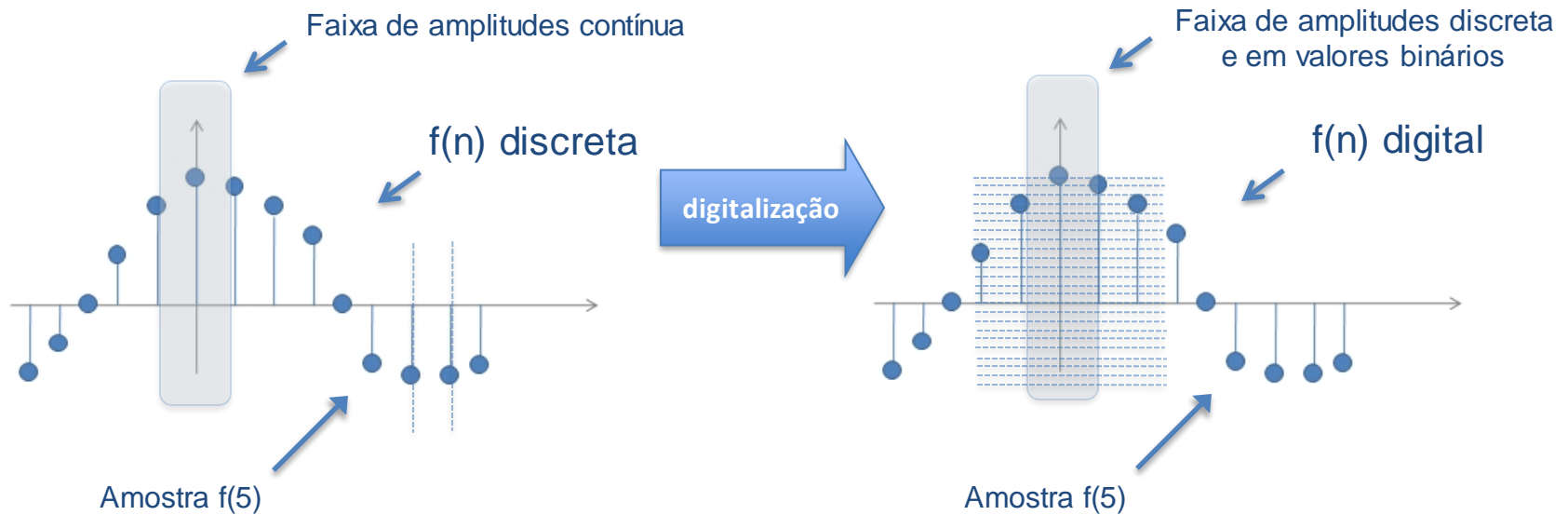
Digitalização do som

- A amostragem é feita periodicamente em tempos múltiplos de um período de amostragem (T_a) e num primeiro momento as amostras possuem suas amplitudes em valores contínuos



Digitalização do som

- A digitalização em si ocorre quando as amplitudes são arredondadas a níveis discretos e finitos, e os números são representados em base binária



Vantagens do som digital

- Razões para se usar som no domínio digital
 - A principal motivação que torna possível trabalhar com o sinal sonoro no domínio discreto é que o sinal analógico original que foi amostrado/discretizado pode ser completamente recuperado por uma operação inversa **caso o Teorema da Amostragem seja obedecido**
 - O sinal sonoro representado numericamente pode ser processado de forma algébrica, por unidades lógicas e aritméticas, permitindo que cálculos de processamento sejam realizados em DSPs e processadores digitais
 - O sinal digital apresenta maior imunidade a ruídos que o analógico

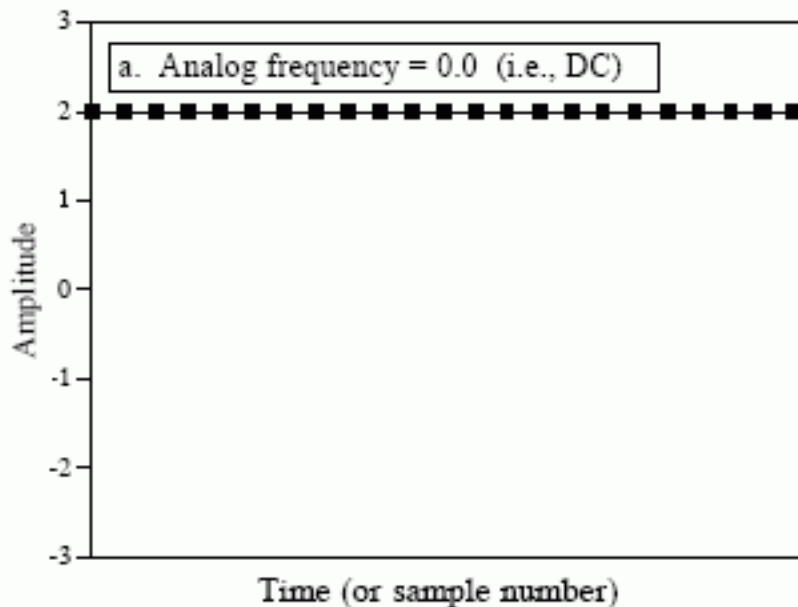
Teorema da amostragem

- Nyquist-Shannon (1928, 1949)
 - “um sinal $f(t)$ cuja maior frequência seja Ω será completamente recuperável a partir de sua versão amostrada $f(n)$ se a frequência de amostragem f_a for igual ou superior a 2Ω ”

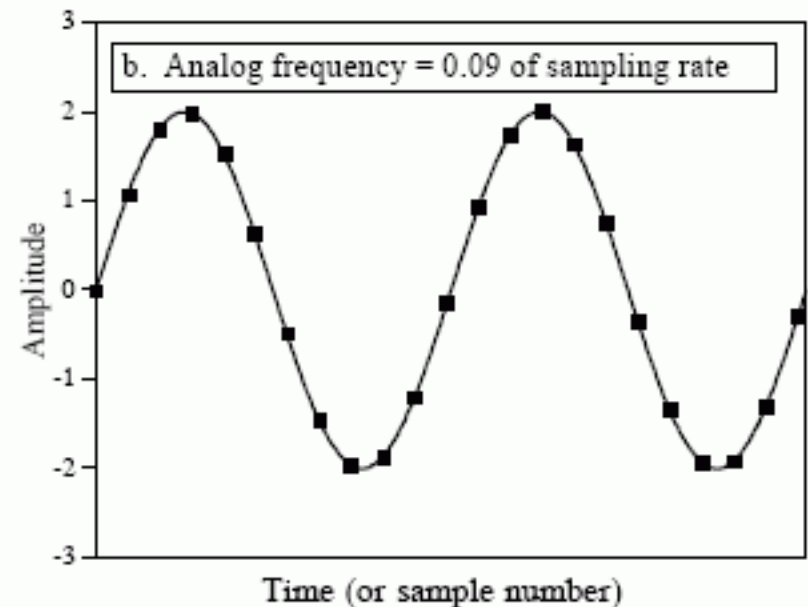
Amostragem com $f_a \geq 2\Omega$

- No exemplo abaixo o sinal original pode ser reconstruído a partir das amostras

$$\Omega=0$$



$$\Omega=0.09f_a \rightarrow f_a \approx 11\Omega$$

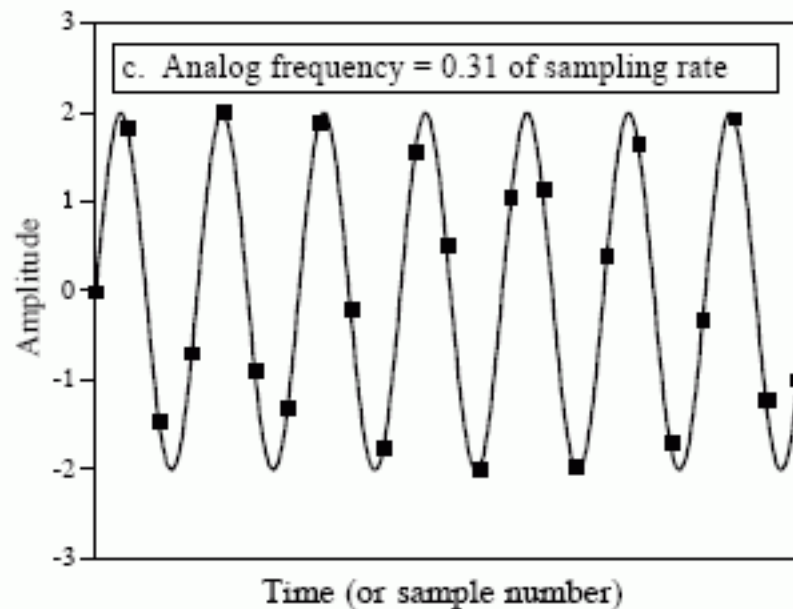


Amostragem com $f_a \geq 2\Omega$

- No exemplo abaixo embora as amostras não pareçam capturar a forma da onda analógica, elas formam um padrão único com o mesmo, e a reconstrução é possível

Na figura: Ω (frequência analógica) é expressa em função de f_a (frequência de amostragem)

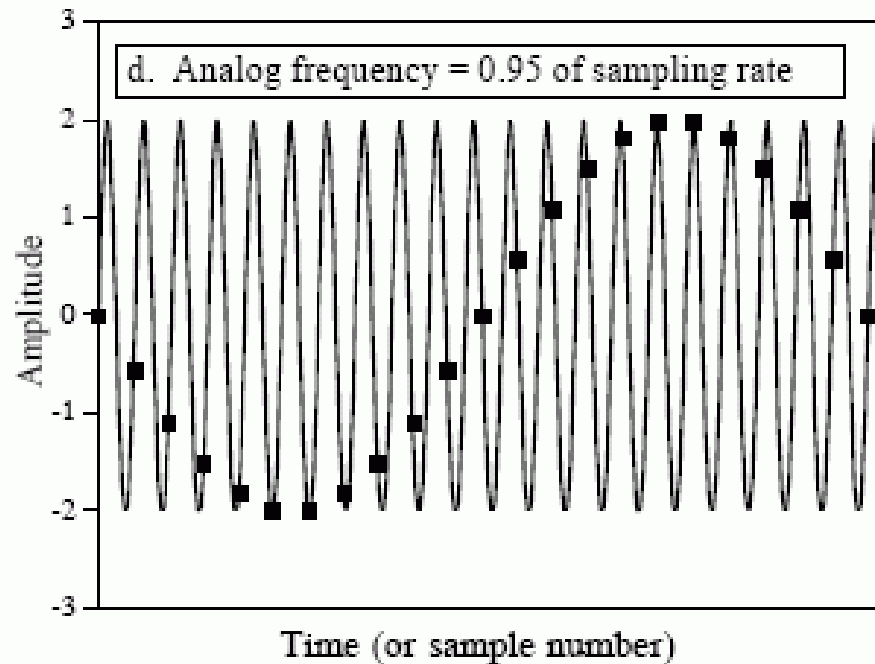
$$\Omega = 0.31f_a \rightarrow f_a \approx 3,22\Omega$$



Amostragem com $f_a < 2\Omega$

- No exemplo abaixo o sinal é amostrado com $f_a = 1,053 \Omega$
- Problema: “**Aliasing**”

$$\Omega = 0.95 f_a \rightarrow f_a \approx 1,053 \Omega$$



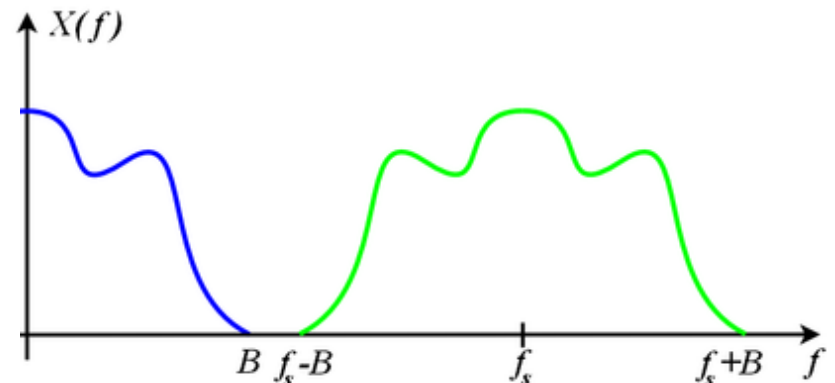
Aliasing

- O aliasing é um efeito produzido pelo aparecimento de frequências devido à sobreposição de uma cópia (ou dobramento) do espectro sonoro sobre ele mesmo
- Consiste em uma ambiguidade na identificação das frequências do som, já que as amostras representam ambigualmente duas ou mais frequências
- O efeito é caracterizado assim como um artefato, uma distorção sonora
- Para ilustrações animadas do efeito veja <https://en.wikipedia.org/wiki/Aliasing>

Espectro do sinal amostrado

- A amostragem provoca o aparecimento de cópias do espectro do sinal original (contínuo) ao longo do eixo das frequências
 - *O espectro $X(f)$ é copiado infinitamente, com cópias centralizadas ao redor de múltiplos da frequência de amostragem f_s*

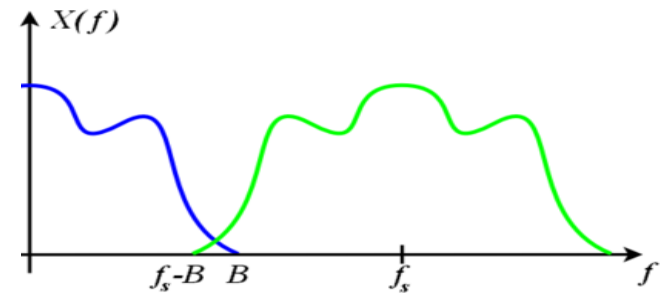
Na figura: $\Omega = B$



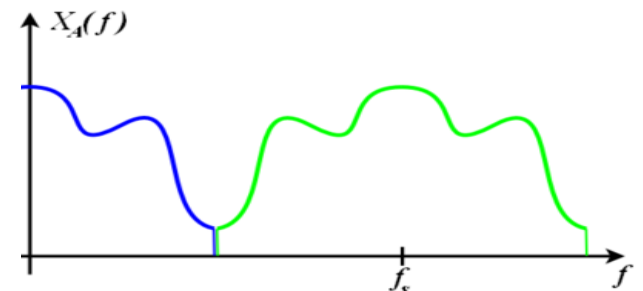
Interpretação espectral do aliasing

- O *aliasing* surge em função da relação não unívoca entre amostra e o sinal amostrado
- Isto ocorre porque as *cópias periódicas* do espectro do sinal amostrado começam a se superpor

Sinal amostrado
insuficientemente

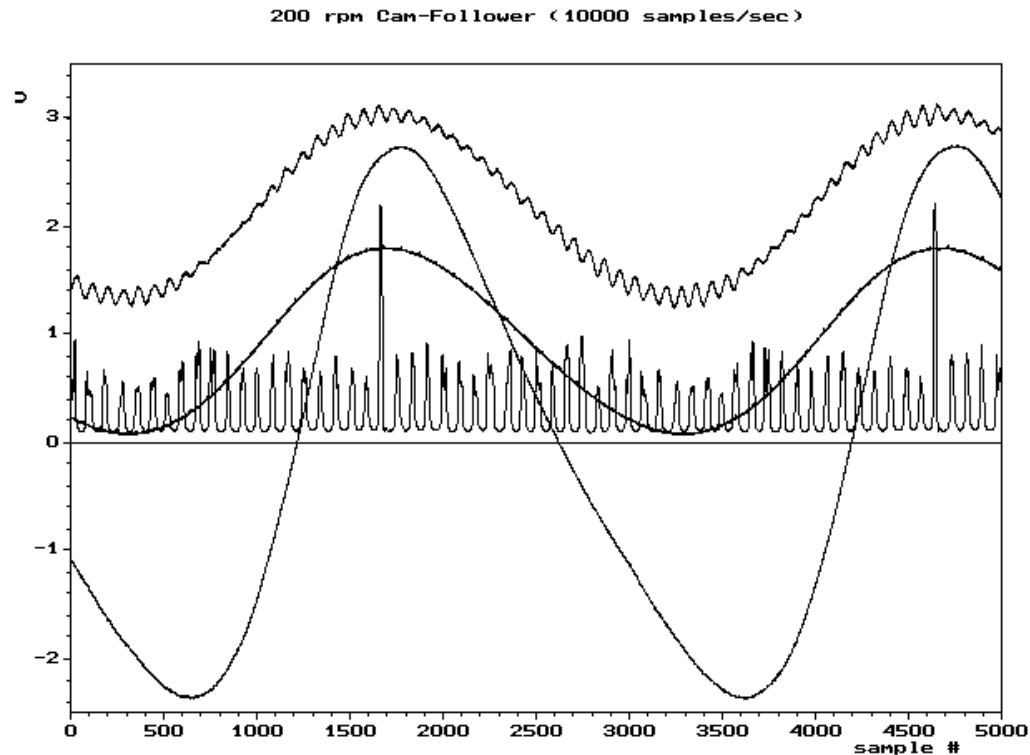


Sinal amostrado
criticamente




Exemplo visual de aliasing

- Veja (abaixo) como os vários sinais, de frequências diferentes, sofrem a influência da variação da taxa de amostragem



Exemplos sonoros de aliasing

1. Tocar arquivo *"sawtooth-aliasingdemo"*

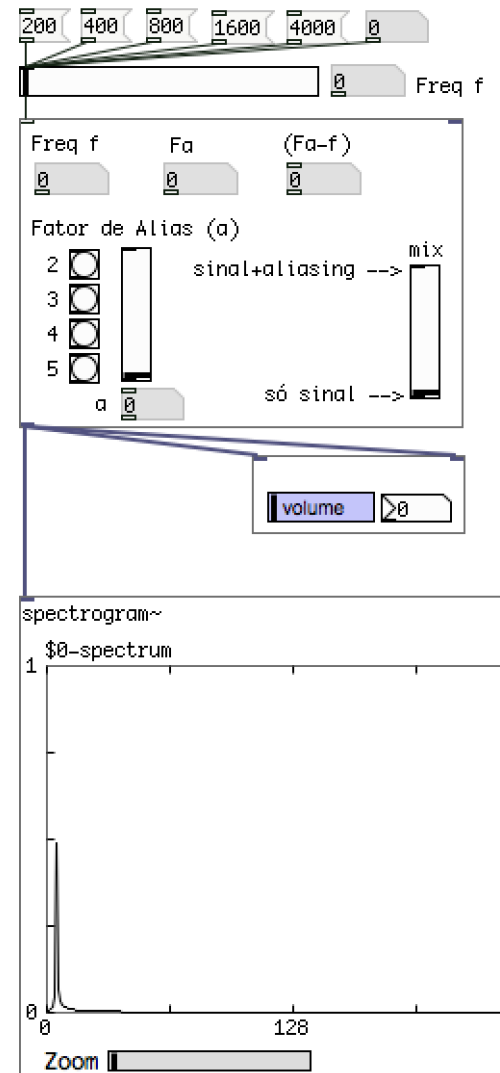
- 3 tons dente-de-serra (441Hz, 882Hz, 1636Hz) tocados em série: primeiro normal, depois com aliasing 

2. Ouça exemplo sonoro interativo de aliasing na página do curso de Sinais e Sistemas da Universidade da Califórnia, Berkeley

- <http://ptolemy.eecs.berkeley.edu/eecs20/week13/aliasing.html>

Demonstração interativa de aliasing

- Abra o *patch* tecmus-aliasing.pd para uma demonstração interativa de distorção devido a aliasing
- O patch permite reproduzir o efeito provocado pelo aliasing em tons puros e em arquivos de som, mudando-se a frequência de amostragem.
- Mostra-se que o *aliasing* distorce o som introduzindo frequências espúrias na parte alta do espectro ao se reduzir a taxa de amostragem



Evitando o aliasing

- O efeito de aparecimento de frequências espúrias devido ao *aliasing* pode ser evitado
 - Aumentando a taxa de amostragem (f_a) p/ no mínimo 2x acima de todas as frequências que estão confundindo suas amostras
 - Aplicando filtros *anti-aliasing*: em teoria “passa-baixas” para restringir a largura de banda do sinal
- Cite vantagens e desvantagens de cada opção

Conversão Analógico-Digital e Digital-Analógico

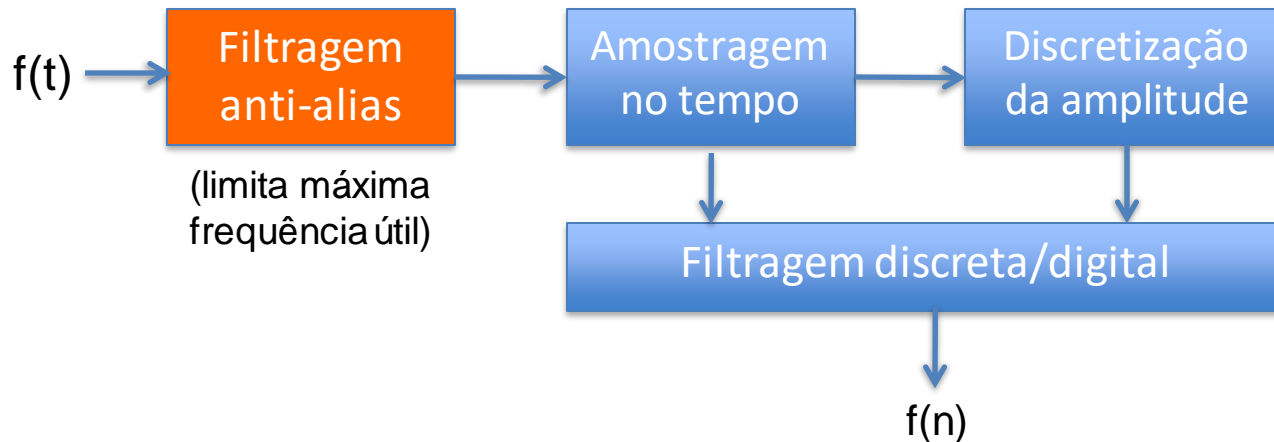
Conversão AD e DA

Motivação

- Há muitas razões para preferir trabalhar com o sinal sonoro no domínio digital, mas o principal argumento que nos permite e justifica trabalhar com sinais amostrados e digitalizados é o fato de que é possível se recuperar o sinal analógico original, que deu origem às amostras, a partir de sua versão amostrada
- O processo que permite a digitalização é a conversão analógico-digital, e o retorno é feito pela conversão digital-analógica
- Vejamos o processo básico de amostragem gerando um sinal discreto e o retorno ao analógico

Visão geral da conversão

- Como discretizar e recuperar o sinal original?



analógico

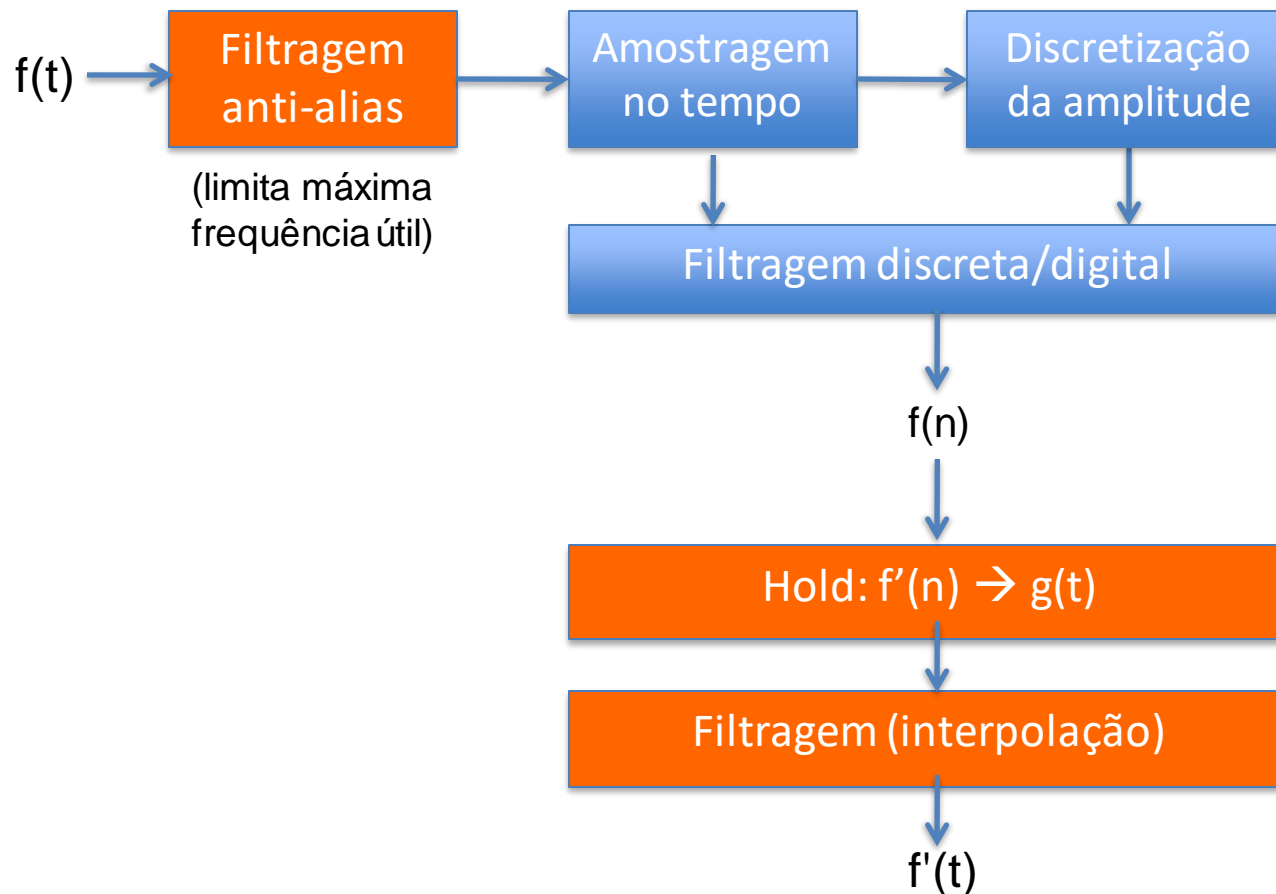
digital

Anti-alias digital:
elimina cópias do
espectro

$f(n)$ é a sequência
discreta do sinal
amostrado

Visão geral da conversão

- Como discretizar e recuperar o sinal original?



analógico

digital

Anti-alias digital:
elimina cópias do
espectro

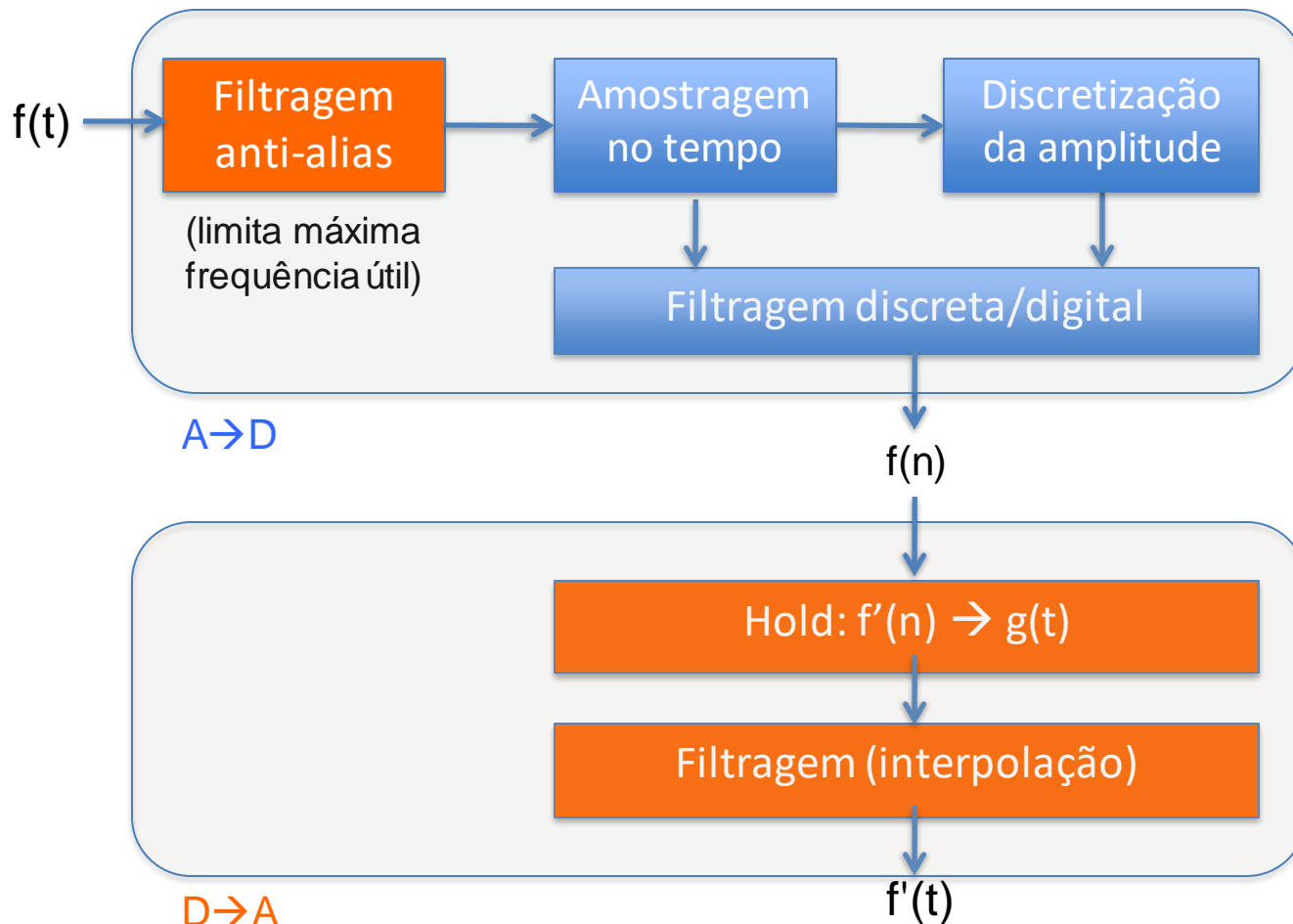
$f(n)$ é a sequência
discreta do sinal
amostrado

$g(t)$ é um sinal
contínuo semelhante
a uma *escadinha*

$f'(t)$ é o sinal contínuo
analógico recuperado

Visão geral da conversão

- Como discretizar e recuperar o sinal original?



analógico

digital

Anti-alias digital:
elimina cópias do
espectro

$f(n)$ é a sequência
discreta do sinal
amostrado

$g(t)$ é um sinal
contínuo semelhante
a uma *escadinha*

$f'(t)$ é o sinal contínuo
analógico recuperado

Roteiro da conversão AD

- Os passos básicos no processo de conversão gerando um sinal digital são
 - (1) Amostragem: discretização do sinal contínuo $f(t)$ produzindo uma sequência discreta $f(n)$ representativa do sinal original, com amostras distando entre si de T_s
 - (2) Este processo faz com que o espectro $F(j\omega)$ de $f(n)$ fique periódico em 2π , isto é, o espectro original é repetido periodicamente ao redor de múltiplos da frequência de amostragem
 - (3) Uma filtragem digital passa-baixa elimina as cópias de $F(j\omega)$ acima e abaixo de $f_s/2$ (π) e deixa o sinal “banda-base” pronto para ser usado em processamentos diversos
 - (4) O sinal $f(n)$ para ser digitalizado precisa ter suas amplitudes arredondadas para valores discretos e finitos padronizados: isto é denominado Quantização
 - (5) Após a quantização $f(n)$ é representado binariamente, o que conclui sua digitalização

Quantização

- Discretização dos valores de amplitude das amostras
 - Após a amostragem, os valores das amplitudes das amostras são ainda representados por valores contínuos
 - Entretanto, para serem trabalhados numericamente faz-se necessário usar um conjunto de valores finito: uma faixa de representação finita
 - A **quantização** discretiza a faixa de amplitudes do sinal (ou faixa dinâmica) a um conjunto de valores finito
 - Após então teremos o sinal discretizado no tempo e na amplitude
 - Ex: sinal amostrado com 44100 amostras/s e quantizado a uma faixa com 65536 níveis de amplitude (caso do CD)

Codificação binária

- O sinal amostrado/discretizado no tempo para ser trabalhado numericamente em DSP e computadores precisa ter os valores de suas amostras codificados em algum sistema numérico binário
- Os valores quantizados das amostras devem então ser representados em base 2 (binária) → nesse ponto teremos o sinal digital representado por sequências de zeros (0) e uns (1)
- Para saber como representar os valores em base 2 estude a seção sobre bases e sistemas numéricos, adiante neste arquivo

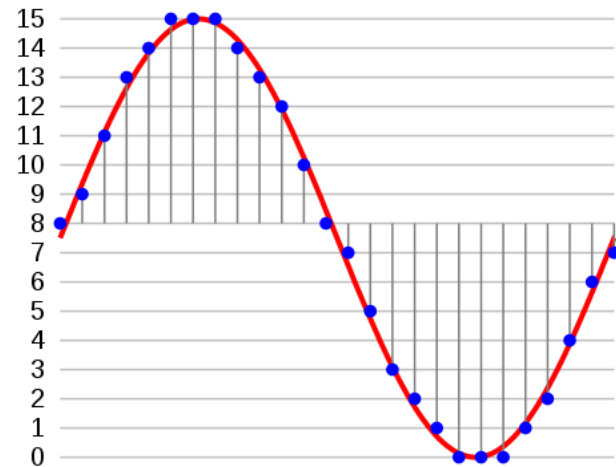
Retorno ao sinal original

- O retorno ao sinal contínuo analógico $f(t)$ a partir de $f(n)$ é resumido pelos passos do roteiro abaixo:
 - (1) Para recuperar o sinal contínuo $f(t)$ primeiro convertemos o sinal $f(n)$ discreto para um sinal contínuo $g(t)$ *segurando-se* o valor entre amostras adjacentes: isto é chamado de “sample & hold”
 - (2) Em seguida suavizamos o sinal $g(t)$, que se parece uma escadinha, realizando uma *filtragem passa-baixa* e gerando o *sinal reconstruído* $f'(t)$ que é uma aproximação (que deve ser fiel) do sinal original $f(t)$

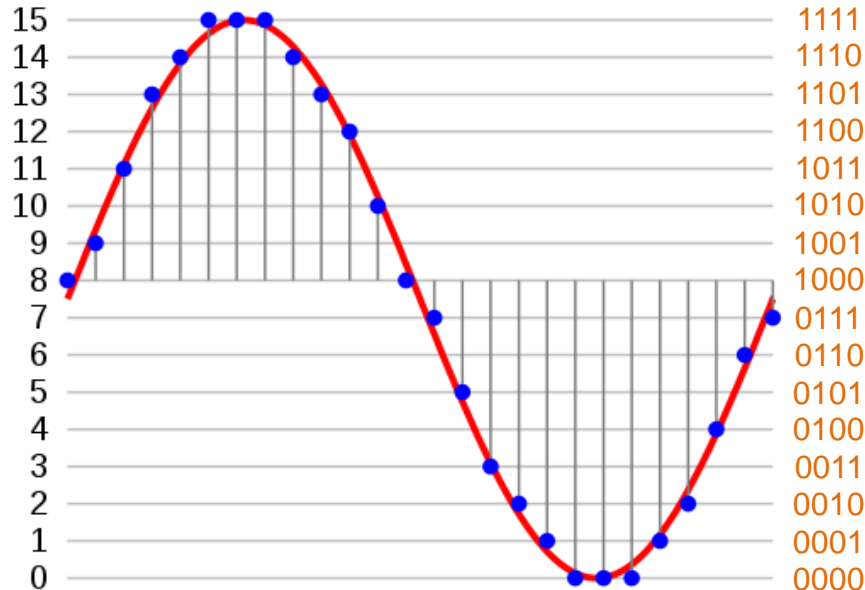
O Conversor PCM

- PCM significa *Modulação por Amplitude de Pulsos* (*Pulse Code Modulation*)
- No PCM um sinal amostrado é representado por uma sequência de pulsos
- Cada pulso corresponde a uma amostra do sinal, que tem um valor de amplitude associado
- A representação da amplitude é mais eficiente se for quantizada, isto é, representada por valores discretos e finitos

Fig. Sinal amostrado e quantizado



O Conversor PCM



- Os valores de amplitude discretos dos pulsos são então codificados usando um **sistema numérico em base binária**, onde são então armazenados na forma de números binários

Amp (pulso n=1) = 1000

Amp (pulso n=2) = 1001

Amp (pulso n=3) = 1011

...

Amp (pulso n= 6) = 1111

Amp (pulso n= 7) = 1111

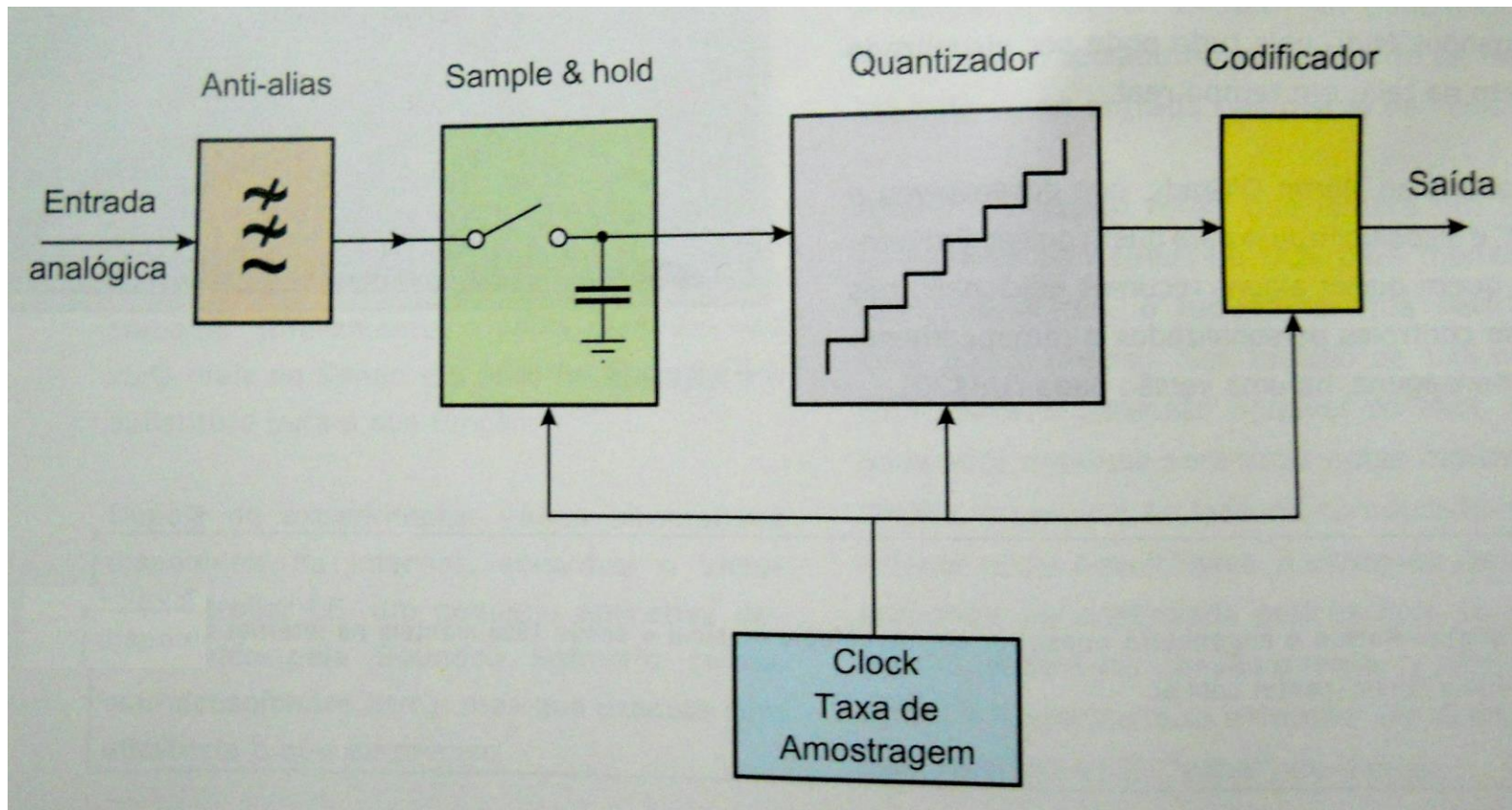
...

O Conversor PCM

- O PCM não é o único método de codificação de áudio digital, mas é *o mais popular*
- Existem muitas variações de conversores PCM, sendo o LPCM (Linear PCM) usado nos arquivos de áudio e caracterizado por ter uma amostragem uniforme
- Vamos ver, todavia, o processo básico de conversão PCM, que envolve 3 operações:
 - Amostragem + Quantização + Codificação

PCM

- Componentes básicos do conversor



Componentes do PCM

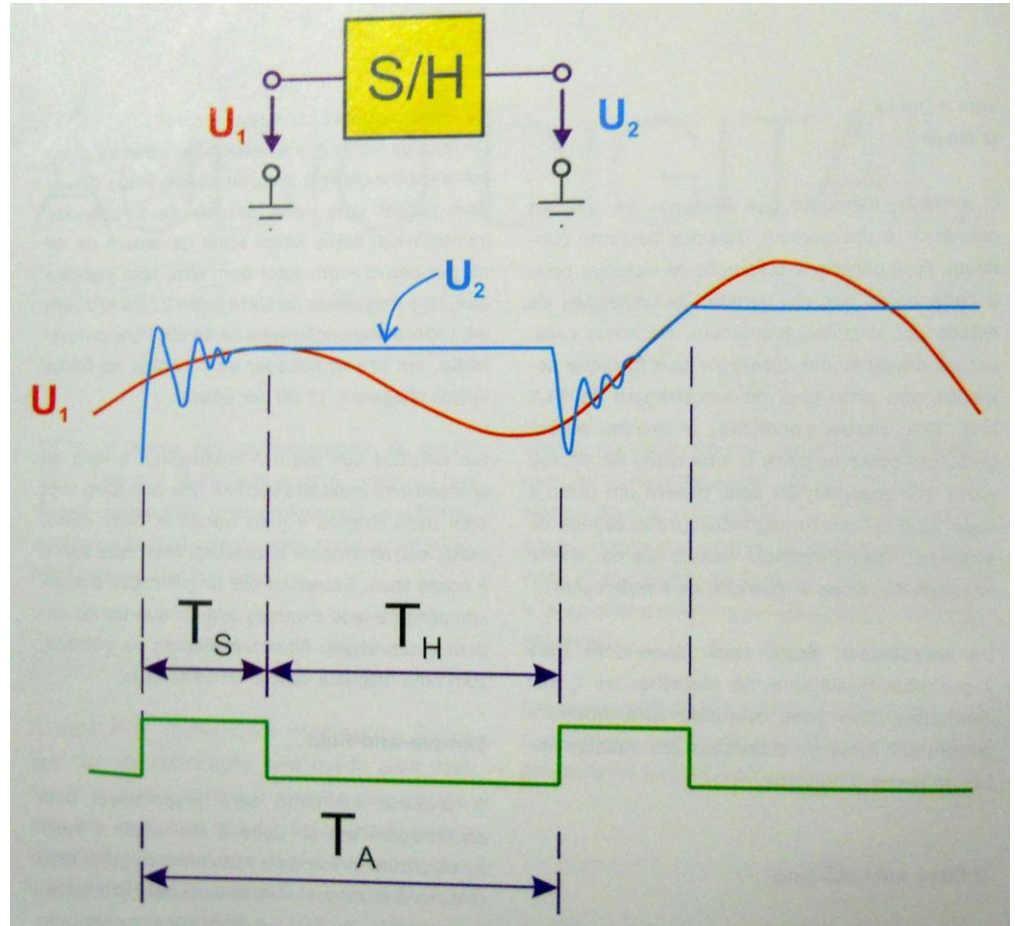
- Filtro anti-alias
 - Elimina as frequências acima do limite de Nyquist (máxima frequência útil no sinal), evitando que sejam convertidas para o domínio amostrado
- *Clock*
 - Base de tempo; coração do sistema
 - dita os tempos em que haverá amostragem

Componentes do PCM

- Amostrador
 - *Sample & Hold* (S/H): coração do amostrador, circuito que faz a leitura do sinal analógico na unidade de tempo e retém por algum tempo o sinal para que seja quantizado
 - $T_{\text{total_amostragem}} = T_{\text{sampling}} + T_{\text{holding}}$

Componentes do PCM

- Amostrador S/H
 - Notar tempos T_s e T_h
 - T_s : Tempo para capturar o sinal e estabilizá-lo
 - T_h : Tempo para convertê-lo



O Conversor PCM

- Quanto maior for a frequência de amostragem f_a maior será o número de amostras retiradas no tempo, e portanto maior será a representação da resolução espectral original
- Uma consequência direta disso é que as frequências mais altas (agudas) serão tão mais preservadas quanto maior for a frequência de amostragem
 - Lembrando que a maior frequência representável será a metade da frequência de amostragem usada, isto é, $\Omega = f_a/2$

Pergunta: qual seriam as frequências mais sujeitas à aliasing: as altas ou baixas frequências de um sinal?

Componentes do PCM

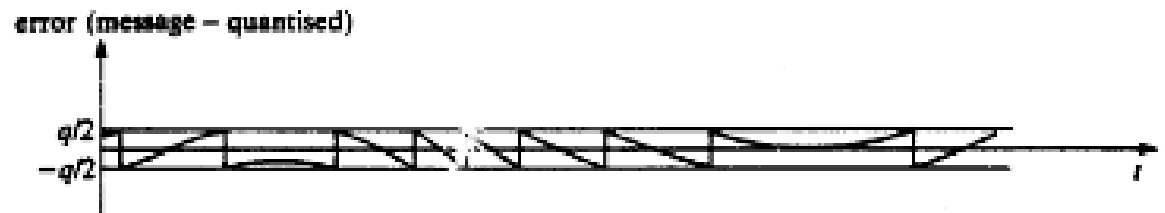
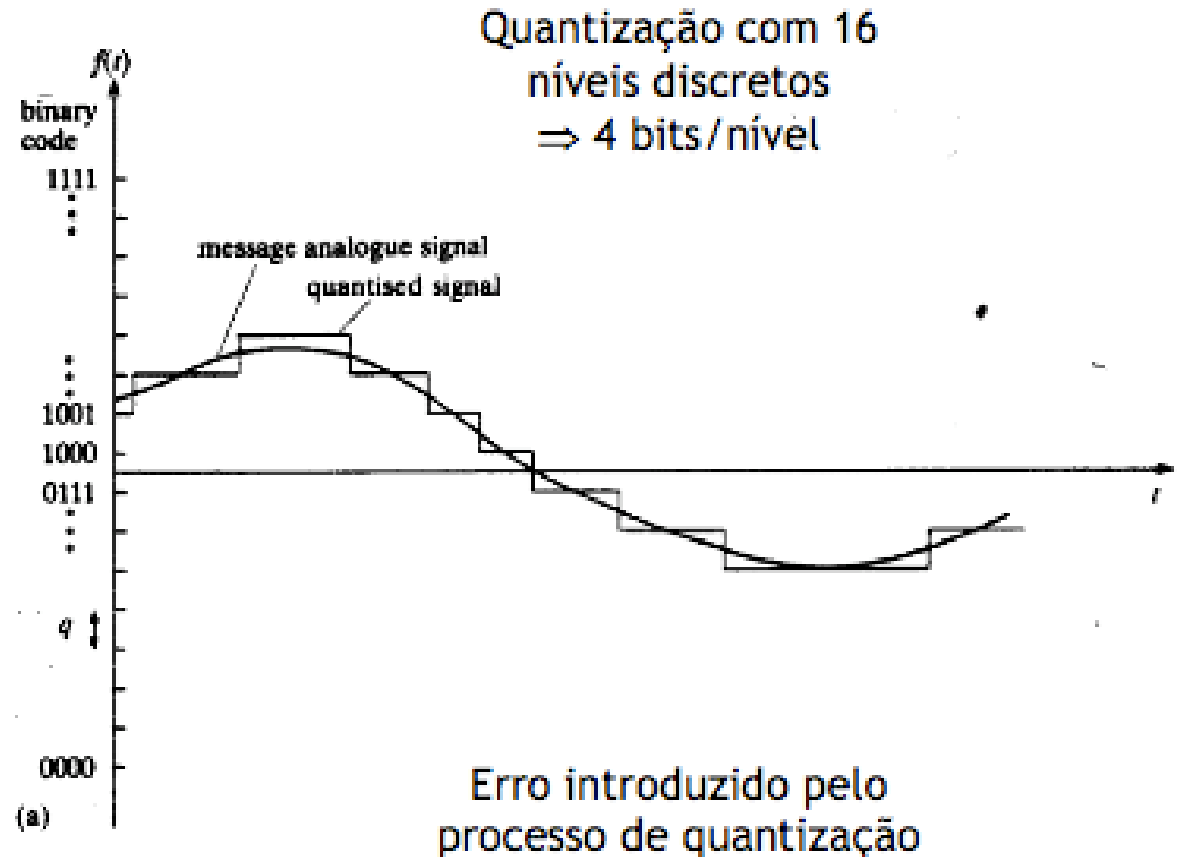
- Quantizador
 - Circuito que relaciona à variável tempo discreta valores também discretos da grandeza medida, isto é, atribui valores de amplitude discretos às amostras
 - A escala contínua de amplitude do sinal é convertida a uma escala discreta, com um número finito de *níveis de quantização*
 - O número de níveis é determinado pelo número de bits usados para representar a escala
 - Ex: 4 bits $\rightarrow 2^4$ níveis = 16 níveis

Componentes do PCM

- Quantizador
 - Na entrada do quantizador, a amplitude da amostra estabilizada (*hold*) em valor contínuo é aproximado para o valor discreto mais próximo de um dos níveis discretos
 - Ex: $3,2 \rightarrow 3$ (*erro*=0,2) | $3,7 \rightarrow 4$ (*erro*: 0,3)
 - O sinal quantizado é caracterizado por uma forma de onda semelhante a uma “escadinha”
 - *Resultado: Medida discreta da amplitude do sinal, com algum erro associado*

Componentes do PCM

- Quantização
 - Ex: c/ 16 níveis
 - Observe:
 - diferença entre valor real do sinal *versus* valor quantizado
 - erro de quantização associado



Componentes do PCM

- Codificador
 - Faz a conversão do valor discreto dos níveis para um sistema de codificação numérico em alguma base
 - Ex: base binária; hexadecimal; etc.
 - A faixa de representação binária de N níveis pode mapeá-los de diferentes modos
 - Ex: com 4 bits, tem-se 16 níveis de quantização, que podem mapear
 - de 0 a 15 ou de 1 a 16 (inteiros positivos) ou
 - de -8 a +7 usando uma representação de números binários positivos e negativos em “complemento de dois” (ver anexos sobre sistemas numéricos binários)

Componentes do PCM

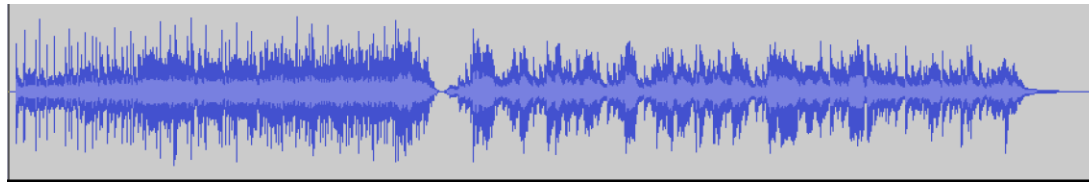
- Codificador
 - Quanto maior o número de bits usados na codificação binária maior será o número de níveis de representação de amplitude, e menor será o erro de aproximação (arredondamento) do valor real para o valor do nível mais próximo
 - Desta forma, maior será a resolução do sinal e maior será a faixa dinâmica representável
 - Ex: cada bit a mais representa aproximadamente mais 6dB de resolução, de forma que 16 bits fornecem aproximadamente 96dB de faixa dinâmica

Exemplos sonoros PCM

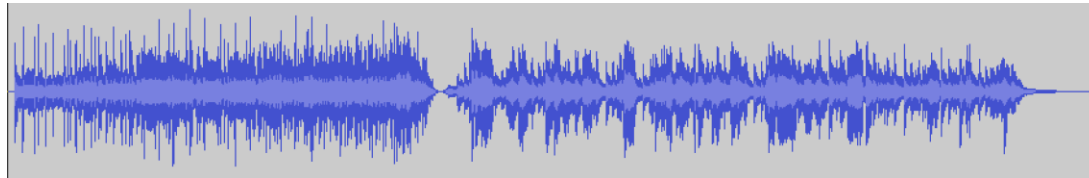
Excertos de Tom Sawyer (Rush)

- Arquivos de som com diferentes resoluções em frequência (largura de banda) mas com mesmo número de níveis de quantização (todos com 16 bits $\rightarrow 2^{16} = 65536$ níveis)

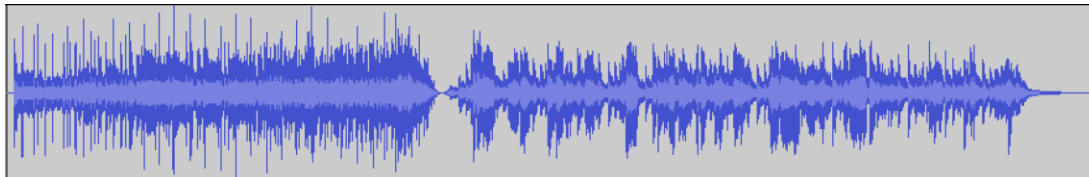
– 8kHz



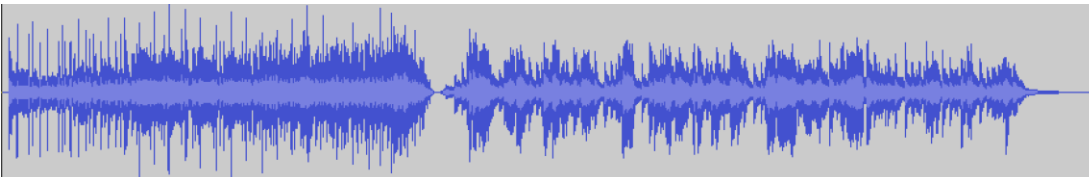
– 11kHz



– 22kHz



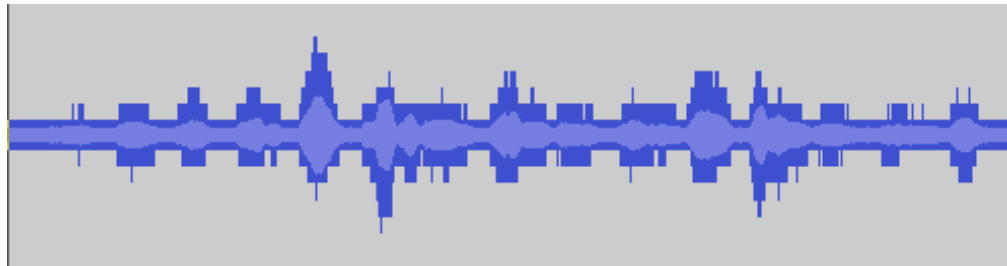
– 44kHz



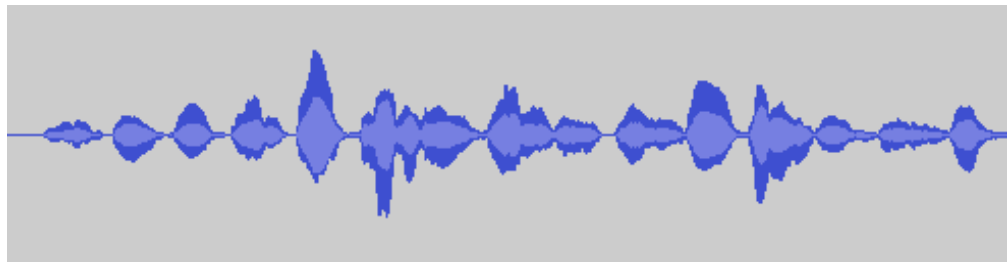
Exemplos sonoros PCM

- Arquivos de som com diferentes números de níveis de quantização mas com mesma resolução em frequência (todos amostrados com 44kHz)

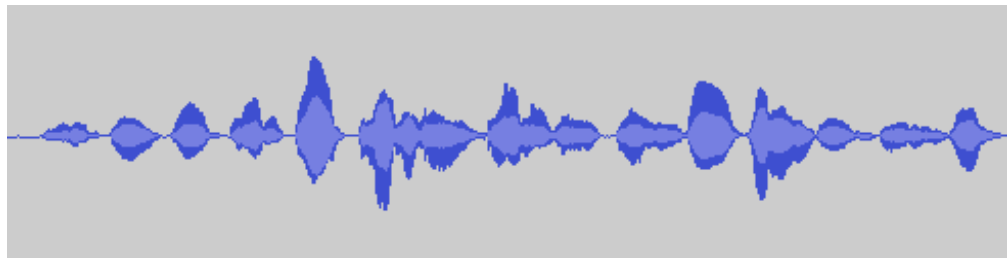
– 44k@4 bits



– 44k@8 bits



– 44k@16 bits



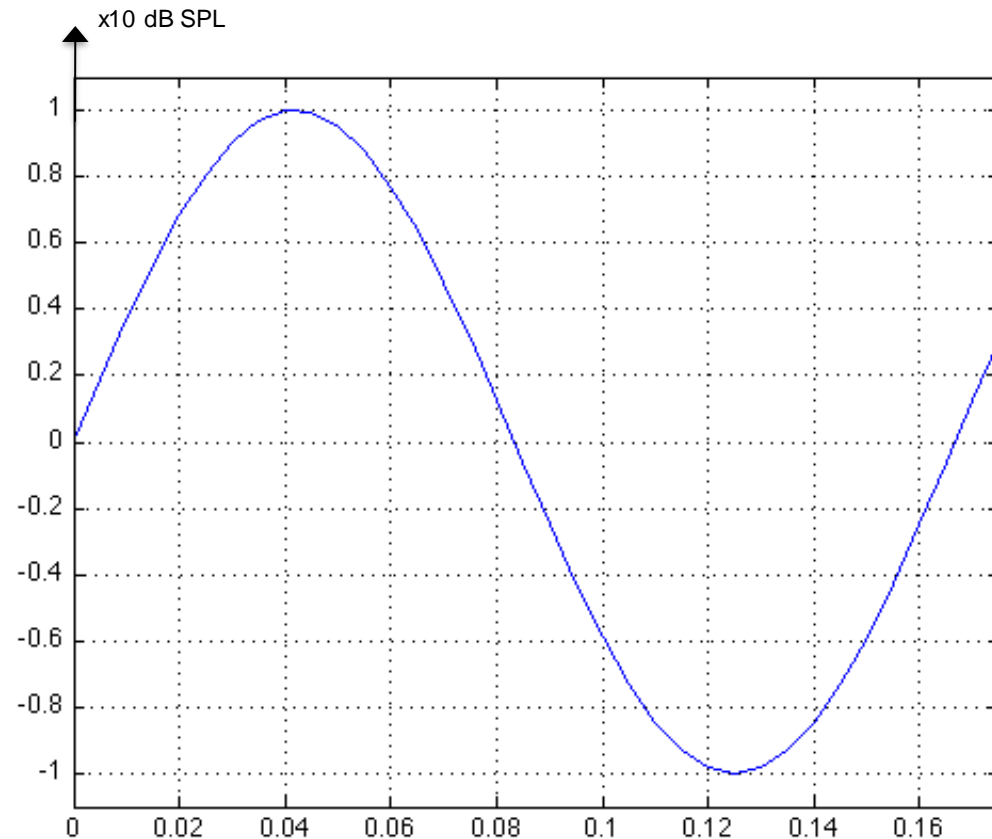
Exemplos sonoros PCM

- Sobre as figuras do slide anterior
 - As figuras mostram arquivos de som com níveis de quantização diferentes: 4 bits, 8 bits, e 16 bits
 - Pode-se observar o efeito do erro de quantização sobre a resolução da representação gráfica do sinal devido a níveis de quantização diferentes
 - Ex: para 4 bits há menos níveis de representação para acomodar o sinal do que para 16 bits, e portanto um erro de quantização maior é mais evidente

Exercício: amostragem e quantização

Roteiro:

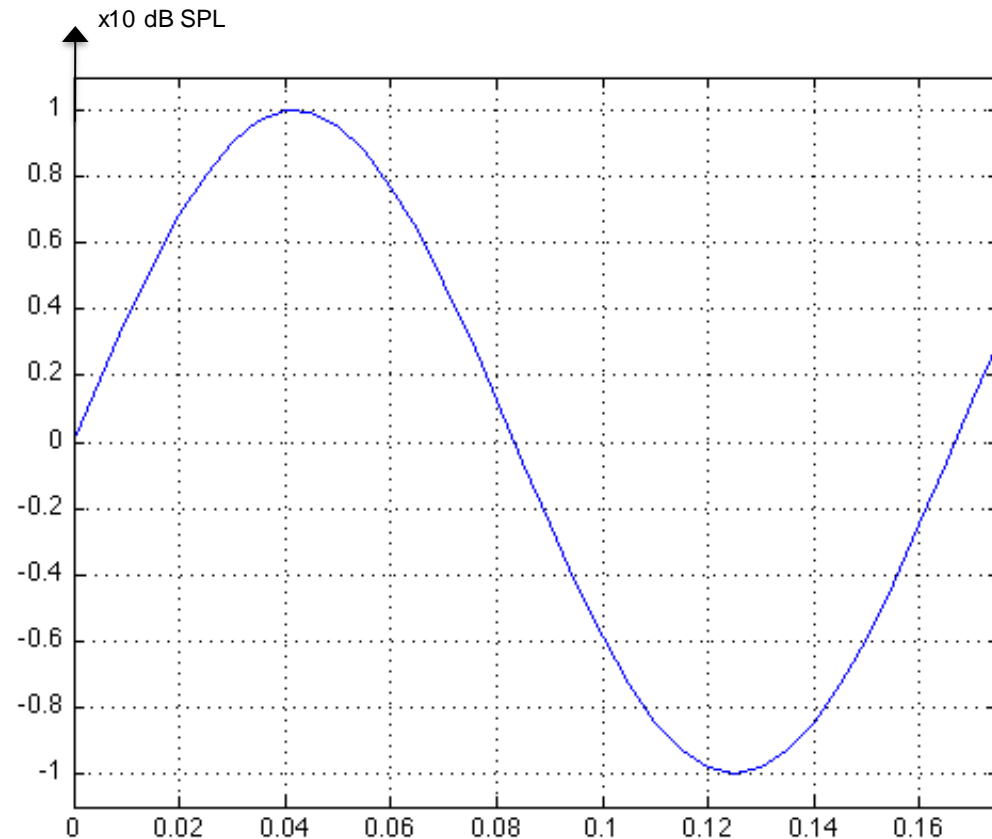
1. Tome o sinal da pressão acústica $p(t)$ registrada de uma fonte sonora, como mostrado no gráfico
2. Amostre-o com uma frequência de 10 amostras por segundo (adote como amostra inicial $t=0$)
3. Informe se haverá aliasing ou não e justifique sua resposta
4. Amostre-o agora com $f_a = 14$ Hz e verifique se haverá aliasing
5. Indique o valor da amplitude das amostras obtidas segundo a escala do eixo vertical. Por conveniência, para lidar com valores de níveis inteiros multiplique os valores de amplitude por x10



Exercício: amostragem e quantização

Roteiro:

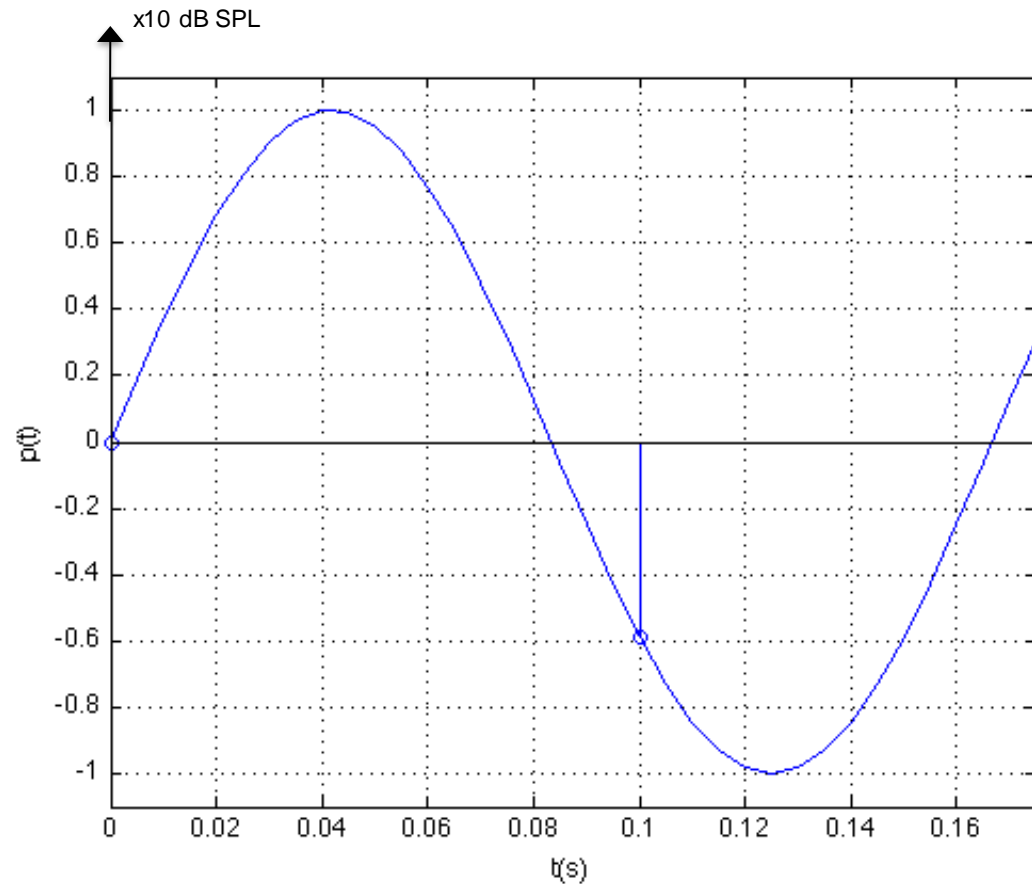
6. Quantize os valores das amostras c/ uma representação binária c/ 4 bits utilizando a codificação em complemento de 2
7. Indique também qual é o erro de aproximação (erro de quantização)
8. Faça comentários em geral sobre o sinal considerado



Solução do Exercício

- $P(t)$ tem período $T \approx 0,167$ s, portanto tem frequência $f = 1/T = 6$ Hz
- $P(t)$ amostrado com $f_a = 10$ Hz (veja amostras em $t=0$ e $t=0,1$ ao lado) terá somente 2 amostras em 1 período, pois o período entre amostras é $T_a = 0,1$ s
- f_a é menor que $2f$ (isto $f_a < 12$ Hz) e desta forma haverá *aliasing*

Obs. O sinal do exercício é um tom puro de 6 Hz (inaudível), que teve uma escala de valores limitada numa faixa de -10 a 10 (unidade arbitrária). Por ser um sinal inaudível, o exemplo só tem caráter didático para exercitar o processo de amostragem e quantização.



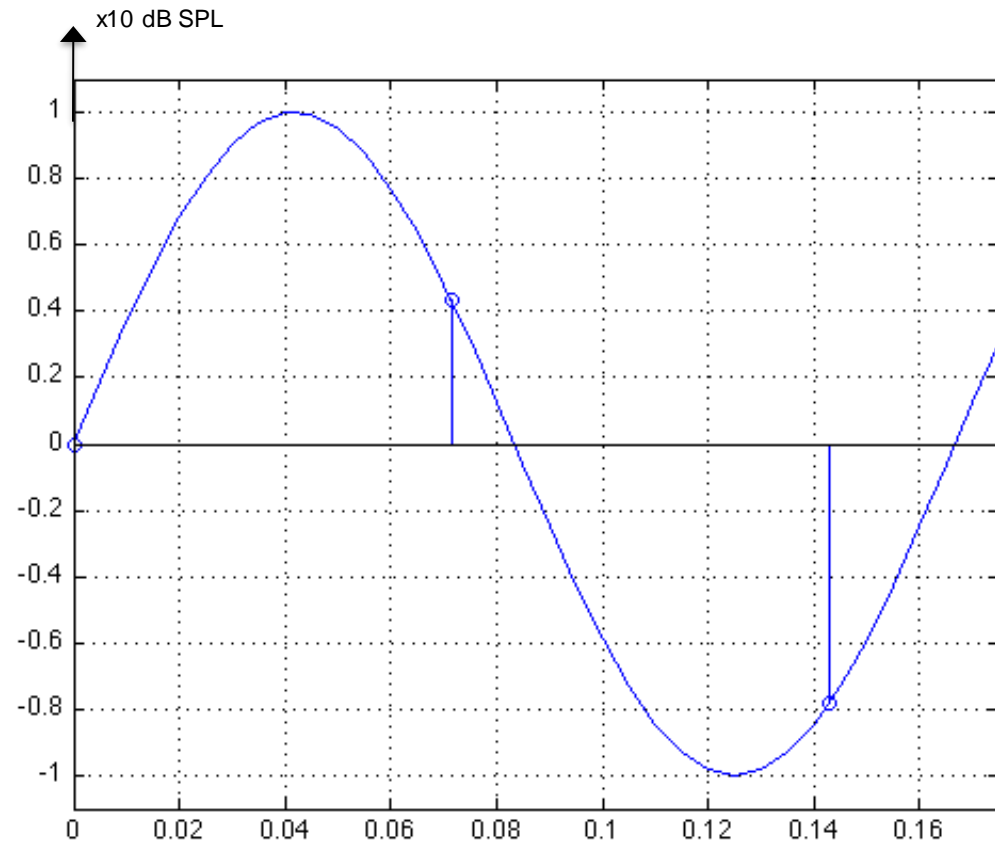
Solução do Exercício

- $P(t)$ amostrado com $f_a = 14$ Hz terá $T_a = 1/14 \approx 0,071$ s entre amostras.
- Como $f_a > 2f$ (isto é $f_a > 12$ Hz) então verificamos que não haverá *aliasing*, havendo amostras suficientes (3 por período) para permitir representar e recuperar o sinal analógico posteriormente
- Indique o valor das amostras obtidas segundo a escala do eixo vertical $\times 10$

$$P_1(t=0) = 0 \text{ dB SPL}$$

$$P_2(t \approx 0,071) \approx 4,5$$

$$P_3(t \approx 0,142) \approx -8,0$$



Solução do Exercício

- Com uma representação binária c/ 4 bits em complemento de dois em ponto fixo (isto é, somente números inteiros) podemos representar valores de -8 a +7, em passos unitários, num total de 16 níveis fixos, o que é insuficiente para representar de -10 a +10. Para tanto usamos então 5 bits, que permitirá representar em complemento de dois de -16 a +15.
- A tabela abaixo quantiza os valores de $p(t)$ para o nível mais próximo dos 32 níveis fixos da representação binária c/ 5 bits em complemento de dois, e indica o erro de aproximação (erro de quantização) máximo
- O nível -8 é obtido tomando +8 = 01000 invertido e somando-se 1 \rightarrow 10111 + 1 = 11000

Amostra	Valor contínuo	Quantizado no nível	Codificado no nível	Erro de quantização máximo
$P_1(t=0)$	= 0	0	00000	0
$P_2(t \approx 0,071)$	$\approx 4,5$	4 ou 5	00100 (ou 00101)	0,5
$P_3(t \approx 0,142)$	$\approx -8,0$	-8	11000	0

Qualidade do PCM

- A qualidade sonora nos conversores PCM é afetada por vários parâmetros, principalmente por:
 - Frequência de amostragem (f_a ou f_s): resolução em frequência (ou espectral) ou largura de banda
 - Número de bits do conversor: erros de quantização e resolução dinâmica
 - Quantização linear/não-linear, uniforme/não uniforme: distorções
 - Não linearidades: distorção harmônica
 - *Jitter*: variações no período de *clock* (base de tempo)

Qualidade do PCM

- A qualidade da conversão é afetada ainda por
 - Linearidade integral (na função de transferência) e linearidade diferencial (uniformidade nos intervalos de quantização) (alto impacto)
 - Monotonicidade (relação de 1 para 1 entre as grandezas relacionadas)
 - Tensão de *offset*: valor constante adicionado ao sinal original (pouco impacto)

Especificações típicas de conversores

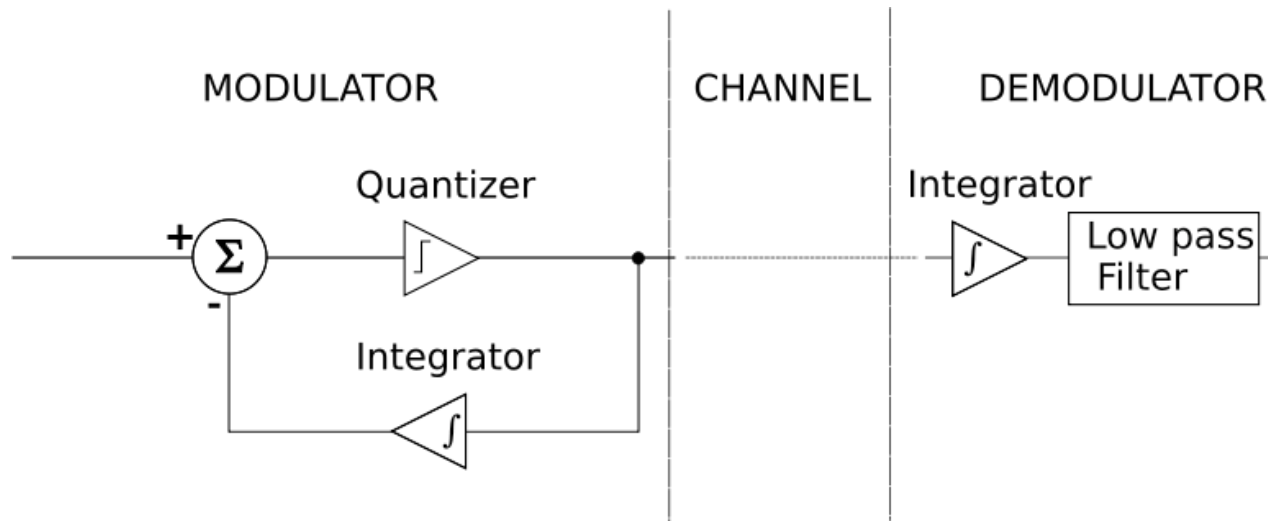
- Exemplo de especificações típicas em *data-sheets* (manuais) de conversores AD e DA PCM
 - Resolução: número de bits, ex: 16 bits
 - SNR (*signal to noise ratio*): ex: 110 dB RMS
 - THD (*total harmonic distortion*): ex: < -104 dB, <0.00032%
 - *Crosstalk* (separação entre canais): >110 dB
 - Internal clock: 800 ps jitter, *random spread spectrum*

Além do PCM

- Além do PCM temos outros métodos de conversão AD/DA, isto é, outras técnicas para amostrar e codificar em bits os sinais de áudio, como a modulação Delta
- A modulação Delta é um método de fazer conversão analógica-digital-analógica baseada numa representação de sinal feita bit a bit
- Uma vantagem deste processo é sua simplicidade, e uma desvantagem o fato de que opera com taxas de amostragem bem mais altas, da ordem de MHz

Modulação Delta

- Ao invés de quantizar o valor de amostras do sinal analógico, o modulador Δ quantiza em cada passo do processo a diferença entre o sinal de entrada e uma média de seus valores anteriores



- A decodificação é ainda mais simples, bastando integrar o *bitstream* recebido para recompor um sinal de banda larga que é posteriormente filtrado para extrair a banda base de áudio.

Formatos de áudio digital

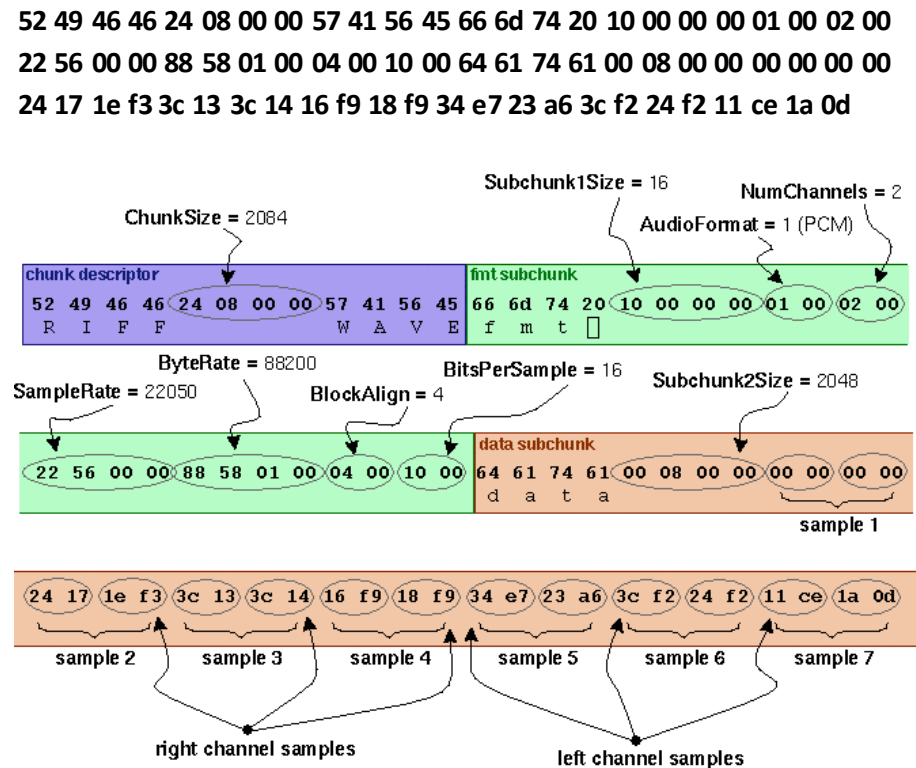
Formatos para codificar e transmitir sinais de
áudio digital

Formatos de áudio digital

- Há uma diversidade de formatos de áudio digital
 - com o sinal representado tanto no domínio do tempo (ex: WAV, AIF, AU, etc.) quanto no da frequência (MPEG, OGG, etc.)
 - com compressão (ex: FLAC, MP3, AAC) ou sem compressão (ex: WAV)
 - para arquivo (ex: WAV, AIF, MP3, MP4, etc.) e para transmissão de áudio (ex: SPDIF, AES3, AES10, AES50, etc.)
- Nos formatos de arquivo sem compressão, como o WAV (Microsoft) e o AIFF (Apple), o áudio em si está representado como um sinal PCM linear (LPCM)

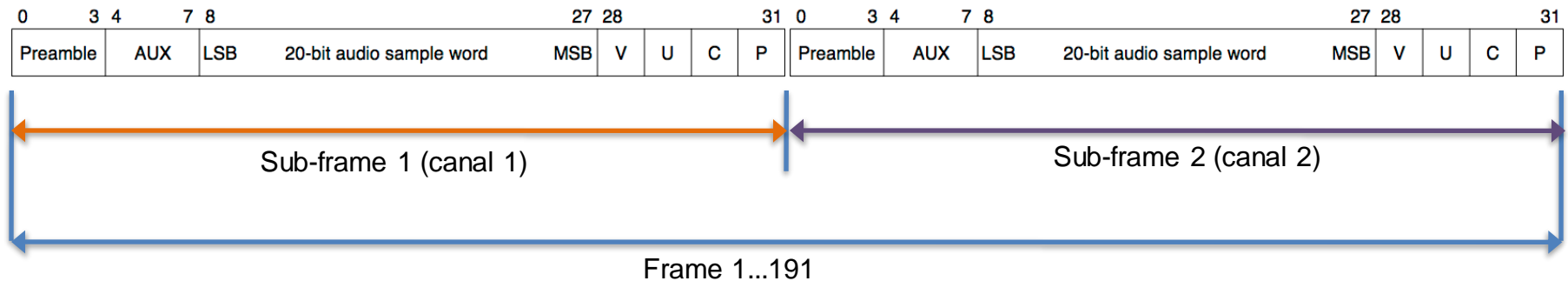
Formatos com áudio PCM

- Nos formatos de arquivo, o *bitstream* informa como os sinais PCM de cada canal são divididos e organizados em grupos (blocos) de bits ao longo do arquivo binário
- Exemplo de formatação no arquivo WAV
 - Descritor de arquivo RIFF
 - Descritor do formato e informações do som (ex: no. de canais, taxa amostragem, bits/amostra)
 - Dados (*payload*)



Formatos com áudio PCM

- A formatação é também importante para a transmissão dos sinais por meio de cabos ou *streamings* digitais, e os formatos mais comuns para transmissão de áudio digital estéreo são o SPDIF e o AES3 (AES/EBU) usados nas interfaces de áudio
- No AES3 os dados de áudio de cada canal são enviados em *sub-frames* de 32 bits juntamente com dados auxiliares do protocolo



- Leia detalhes sobre este formato estereofônico na norma AES3 (1992) bem como sobre sua expansão multicanal AES10 (MADI)
- Referência adicional: <https://www.vbrazilsystems.com/protocolos-digitais-spdifaes-ebu.html>

Formatos comprimidos

- Para se reduzir o tamanho dos arquivos, podem ser usados algoritmos e técnicas de compressão de dados, que reduzem a quantidade de informação a ser armazenada ou transmitida
- Há formatos de áudio com
 - *compressão sem perdas (lossless)* em que toda a informação original pode ser recuperada, e
 - *compressão com perdas*, que utilizam as técnicas de codificação perceptual, e são acompanhados de algum grau de perda de qualidade

Formatos comprimidos

- Formatos usando compressão **sem perdas** incluem:
 - FLAC - *Free Lossless Audio Codec* – um formato de áudio com compressão aritmética dos dados totalmente recuperável
- Formatos usando compressão **com perdas** incluem:
 - [MPEG-1](#) (MP1, MP2, MP3)
 - MPEG-2 e MPEG-4 (AAC - *Advanced Audio Coding*)
 - Ogg Vorbis (livre)
 - Dolby AC-3, DTS e WMA (proprietários)
 - Codecs sobre bluetooth

Formatos comprimidos

- Uma técnica bastante usada para se reduzir a quantidade de informação antes de aplicar uma compressão inclui uma variação do PCM conhecida como **ADPCM (*Adaptive Differential PCM*)**
- Esta é uma técnica que codifica não a amplitude da amostra em si mas sim **a diferença** entre um valor previsto calculado para a amostra e o seu valor real
- As diferenças (sendo menores que o valor integral da amostra) têm menor amplitude e precisam de menos bits para serem armazenadas, o que leva a uma economia de bits na representação

Formatos comprimidos

- Há ainda formatos de representação de áudio PCM não linear, que utilizam curvas de compressão e expansão para melhorar a qualidade sonora quando usamos um número de bits menor
- Para uma mesma qualidade, os arquivos assim terão menor tamanho comparados ao PCM linear
- Estes formatos incluem por exemplo:
 - μ -Law (*mu-law compressed sound format*)
 - A-Law (*a-law compressed sound format*)

Formatos A-Law e μ -Law

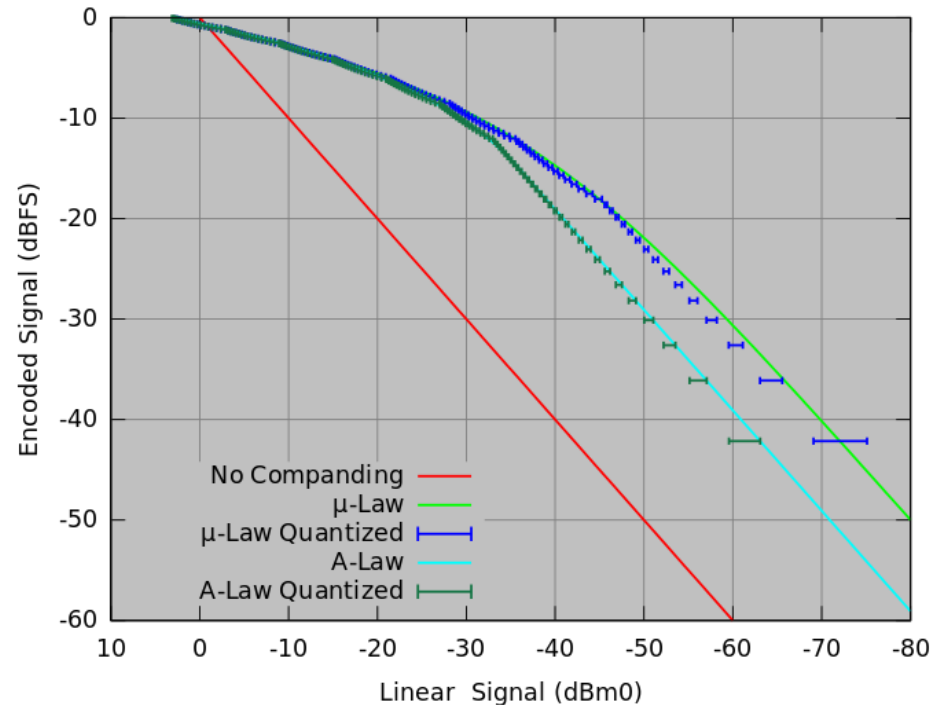
- Ambos os formatos μ -Law e A-Law aproveitam-se do fato da percepção da intensidade pelo ouvido ser **logarítmica** e não linear, e ajustam as amplitudes para uma faixa logarítmica
- Uma forma de fazer isso é **remapear** os valores de amplitude das amostras, originalmente dispostos numa faixa linear, para valores em uma faixa logarítmica, fazendo

$$A_{\text{remapeada}} = \log (A_{\text{original}})$$

- Isso fará que os valores de amplitude mais altas (fortíssimos) sejam comprimidos em uma faixa de representação mais estreita , e os sons menos intensos (pianíssimos) sejam mais bem discriminados
- Em ambos os formatos a quantização usará espaçamentos desiguais entre os níveis, determinados por uma fórmula de compressão logarítmica como esta

Formatos A-Law e μ -Law

- Mapeamento linear (sem transformação): A curva **vermelha** mapeia $A \rightarrow A$, não alterando os valores
 - Ex: $-50\text{dB} \rightarrow -50\text{dB}$
- Mapeamento logarítmico (remapeamento): As outras curvas mapeiam $A \rightarrow A'$, alterando os valores
 - Ex c/ Amplitudes baixas: $(-50\text{dB} \rightarrow -30\text{dB})$ e $(-40\text{dB} \rightarrow -20\text{dB})$, i.e. variação de 10dB mapeada em uma variação de 10dB (preservou mais)
 - Ex c/ Amplitudes altas: $(-30\text{dB} \rightarrow -11\text{dB})$ e $(-20\text{dB} \rightarrow -7\text{dB})$, i.e. variação de 10dB mapeada em uma faixa de 4dB (comprimiu)



- Isto provoca:
 - Redução da faixa dinâmica representável
 - Melhor SNR (relação do nível de sinal sobre o nível de ruído de quantização)

Formatos avançados

- Família MPEG
- Codecs de áudio sobre Bluetooth
- Audio sobre rede (*Audio over IP*)

Formatos avançados

- Família MPEG
 - MPEG-H (3D Audio)
 - MPEG-I (Immersive Audio, em desenvolvimento)
 - Para um cenário geral sobre os padrões MPEG veja <https://mpeg.chiariglione.org/standards/> e <https://www.mpegstandards.org/>

Formatos avançados

- Codecs de áudio sobre Bluetooth (5.0 e 5.1)
 - aptX, aptX HD
 - SBC (sub-band coding)
 - AAC (advanced audio coding)
 - LDAC (sony, <https://www.sony.net/Products/LDAC/>)
 - LHDC
 - LC3
- Tutorial sobre codecs bluetooth:
<https://www.nextpit.com/bluetooth-audio-codecs>

Áudio sobre rede

- Há vários protocolos e tecnologias desenvolvidas por fabricantes diversos para transmitir áudio multicanal sobre redes, usando-se infra-estrutura compatível de rede Ethernet e conectorização RJ-45 de rede comum:
 - Cobranet
 - Dante
 - EtherSound
 - REAC (Roland)
 - NetJACK
- A AES especificou o protocolo AES50 voltado para a transmissão de áudio sobre rede, mas que requer switches próprios, não sendo compatível com equipamentos de rede Ethernet

Áudio sobre rede

- O IEEE também padronizou uma norma para transmissão audiovisual em rede, o *Audio Video Bridging* (AVB, IEEE 802.1 BA)
- Em 2013 a AES publicou uma norma voltada para regular a interoperabilidade entre diferentes protocolos de transmissão de áudio sobre redes IP, o padrão AES67
- Há um fabricante nacional que introduziu um protocolo/equipamentos próprios no mercado: DSPro
- Infelizmente os fabricantes ainda não aderiram aos padrões da AES e IEEE de forma universal compatível
- Tutorial na net sobre Audio sobre IP:
 - Ethernet Audio: Everything You Need To Know About Audio-Over-IP, <https://www.soundonsound.com/techniques/ethernet-audio>

Conversores de formato

- Quase todos os ambientes de áudio digital permitem algum tipo de conversão de formato de áudio
- Conversões podem envolver
 - Mudança de taxa de amostragem e no. de bits (*sample size*)
 - Alteração no tipo de codificação da faixa de valores de amplitude (valores inteiros, inteiros com sinal ou sem sinal, valores fracionários)
 - Normalização e compressão dinâmica (mu-law/A-law), compressão perceptual (mp3, aac, wma, ogg), compressão sem perdas (flac), etc.
 - Alteração no formato de arquivo (dezenas de tipos possíveis)

Conversores de formato

- Alguns conversores de formato são livres, e cobrem um extenso leque de recursos, como por exemplo:
 - Sox (<http://sox.sourceforge.net/>)
 - VLC
- O programa Sox, por exemplo, é operado a partir de comandos de linha, digitados diretamente de um terminal de comandos do Windows ou Linux/Unix, sem interface gráfica
 - Exemplo de comandos de conversão no Sox:

Comando	O que faz
sox piano.wav -b8 piano.wav	Converte o arquivo piano.wav para outro wav com 8 bits de resolução dinâmica
sox piano.wav piano.ogg	Converte o arquivo de wav para ogg (compressão perceptual)
sox piano.wav -signed-integer piano.mp3	Converte piano.wav para piano.mp3 usando faixa de valores inteiros com sinal (+ e -)

Discussões finais

Áudio digital x analógico

- Vantagens?
- Desvantagens?

Vantagens X Desvantagens

- Pontos de discussão comuns:
 - Menos aparelhos
 - Eficácia X eficiência
 - Mais barato (\$)
 - Digital \leftrightarrow software+ hardware e Analógico \leftrightarrow hardware
 - Possível perda de qualidade
 - Fidelidade \leftrightarrow parâmetros como f_s e no. de bits
 - Durabilidade
 - Armazenamento (registro), facilidade de manuseio (peso e dimensões de equipamentos)
 - Chance de corromper o sinal (imunidade a ruído)
 - Registro da faixa dinâmica (compressão ou não)
 - Mais econômico \rightarrow indústria sustentável
 - Facilidade de operação e reconfiguração
 - Chance de fazer processamentos impossíveis em domínio analógico

Aspectos práticos

- Itens que merecem nossa atenção para qualificar aspectos práticos acerca do áudio digital:
 - Amostragem: taxa ou frequência usada
 - Resolução: maior número de bits, maior capacidade de representação de valores de amplitude
 - Faixa de extensão dinâmica: cada bit acrescenta $\sim 6\text{dB}$
 - Relação sinal/ruído (SNR): resolução maior implica em menor SNR
 - Erro de quantização: resolução maior implica em menor erro de quantização
 - Artefatos: amplitudes além da faixa representável provocam *clipping*, que são ouvidos como *clicks* ou distorção
 - Tamanho de arquivos: mais bits e mais amostras implicam em arquivos maiores

Revisão do módulo

- Neste módulo apresentamos as ondas acústicas e suas formas de representação analógica e discreta
- Vimos o teorema da amostragem, o processo de conversão AD e DA, e o formato de representação de áudio digital PCM, usado no CD
- Foram identificados os processos-chaves e limitações relacionados à amostragem e à reconstrução do sinal, como a taxa de amostragem e a quantização da faixa de representação dinâmica
- Finalmente foram vistos outros formatos digitais de áudio e discutidos aspectos práticos relacionados ao áudio digital

Referências adicionais p/ estudo

- Tutorial online sobre áudio digital:
 - http://www.eca.usp.br/prof/iazzetta/tutor/audio/a_digital/a_digital.html

Trabalho Prático

- No ambiente online da disciplina será definido um trabalho prático sobre um dos assuntos vistos nesta aula como
 - Compressão em formatos MPEG
 - Formatos por modulação por pulsos e modulação delta (PCM versus DSD)
 - Formatos de áudio digital sobre redes e para *streaming*
 - Conversão AD/DA

Próximo módulo

- Além da representação do sinal acústico em si, há formatos orientados ou especializados para a representação de certos tipos específicos de informação sonora, como a música
- Próximo módulo: protocolo MIDI e outros formatos para representar música digital
- Tarefa para casa: leitura sobre a especificação MIDI

Bases e sistemas de representação numérica

Representação de números inteiros, fracionários, positivos e negativos em base 2 (binária), 5, 8 (octal), 10 (decimal) e 16 (hexadecimal)

Bases de representação numérica

- Ex: bases de representação

$$- 10 = 1 \times 10^1 + 0 \times 10^0 = 10_{10}$$

$$- 10_{10} = 0 \times 16^1 + 10 \times 16^0 = 0A_{16} = \mathbf{0A_H}$$

$$- 10_{10} = 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 0 \times 2^0 = 1010_2 = \mathbf{1010_2}$$

<i>base 10 (10^n)</i>	<i>base 5 (5^n)</i>	<i>base 2 (2^n)</i>	<i>base 8 (8^n)</i>	<i>base 16 (16^n)</i>
000	0000	0000 0000	000	00
<i>001 (10^0)</i>	<i>0001 (5^0)</i>	<i>0000 0001 (2^0)</i>	<i>001 (8^0)</i>	<i>01 (16^0)</i>
002	0002	<i>0000 0010 (2^1)</i>	002	02
003	0003	0000 0011	003	03
004	0004	<i>0000 0100 (2^2)</i>	004	04
005	<i>0010 (5^1)</i>	0000 0101	005	05
006	0011	0000 0110	006	06
007	0012	0000 0111	007	07
008	0013	<i>0000 1000 (2^3)</i>	<i>010 (8^1)</i>	08
009	0014	0000 1001	011	09
<i>010 (10^1)</i>	0020	0000 1010	012	0A

Sistemas numéricos usados

- Números binários inteiros (positivos e negativos)
 - Complemento de um (bit sinal + valor absoluto)
 - Complemento de dois (bit sinal + valor absoluto)
 - Representação é feita em “ponto fixo”
 - Ponto decimal fixo → escala fixa
 - Representa-se o sinal + valor
- Números binários fracionários (números reais)
 - Representação é feita em “ponto-flutuante”
 - Sinal + expoente + valor → sinal + base + mantissa
 - Ponto decimal “flutua” entre escalas diferentes

Complemento de dois

- O bit mais significativo (à esquerda) é dedicado ao sinal. Com N bits portanto, 1 bit é para o sinal e N-1 bits codificarão 2^{N-1} níveis
 - Ex: 4 bits $\rightarrow 0101 = +101 \rightarrow (+5)$
- Os números negativos são obtidos invertendo-se todos os dígitos (complemento) e somando-se 1
 - Ex: $0101 (+5) \rightarrow 1010 + 1 = 1011 \rightarrow -5$
- Para saber qual número negativo está representado faça a mesma operação
 - Ex: $1010 \rightarrow 0101 + 1 = 0110 \rightarrow |6|$, portanto o número binário 1010 em complemento de dois é -6

Complemento de dois

- A faixa de representação com $N=4$ bits vai de

0111 (+7) que é $+(2^{N-1} - 1) = +(2^3 - 1) = +7$

até

1000 (-8) que $-(2^{N-1}) = -(2^3) = -8$

passando por

0000 (zero) e 1111 (-1)

Representando números inteiros

- Números binários inteiros positivos (sem sinal)
- Inteiros positivos e negativos
 - em complemento de dois ← mais usado
 - em complemento de um

PADRÃO	SEM SINAL	COMPLEMENTO	
		DE DOIS	DE UM
0000 0000	0	+0	+0
0000 0001	1	+1	+1
0000 0010	2	+2	+2
0000 0011	3	+3	+3
.	.	.	.
.	.	.	.
0111 1100	124	+124	+124
0111 1101	125	+125	+125
0111 1110	126	+126	+126
0111 1111	127	+127	+127
1000 0000	128	-128	-127
1000 0001	129	-127	-126
1000 0010	130	-126	-125
1000 0011	131	-125	-124
.	.	.	.
.	.	.	.
1111 1100	252	-4	-3
1111 1101	253	-3	-2
1111 1110	254	-2	-1
1111 1111	255	-1	-0

Representando números reais

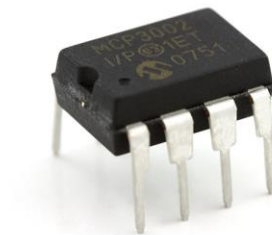
- Representação em ponto flutuante: permite representar números fracionários (em qualquer base) usando-se uma *mantissa + expoente*

Número	Representação em ponto flutuante	mantissa	base	expoente
1532	0.1532×10^4	0.1532	10	4
15.32	0.1532×10^2	0.1532	10	2
0.00255	0.255×10^{-2}	0.255	10	-2
10	0.1×10^2	0.1	10	2
10	0.0010×2^4	0.0010	2	4

Exercício de conversão de base

1. Converta o número decimal 11_{10} para binário
2. Converta os seguintes números binários codificados na forma de *senal + expoente + mantissa* para decimal. Considere que:
 - o sinal toma o bit mais significativo
 - a mantissa usa os 4 bits menos significativos e não está em complemento de dois
 - o expoente está em complemento de dois e usa os bits restantes
 - a) 0110 0011
 - b) 0001 1010

Conversores AD/DA

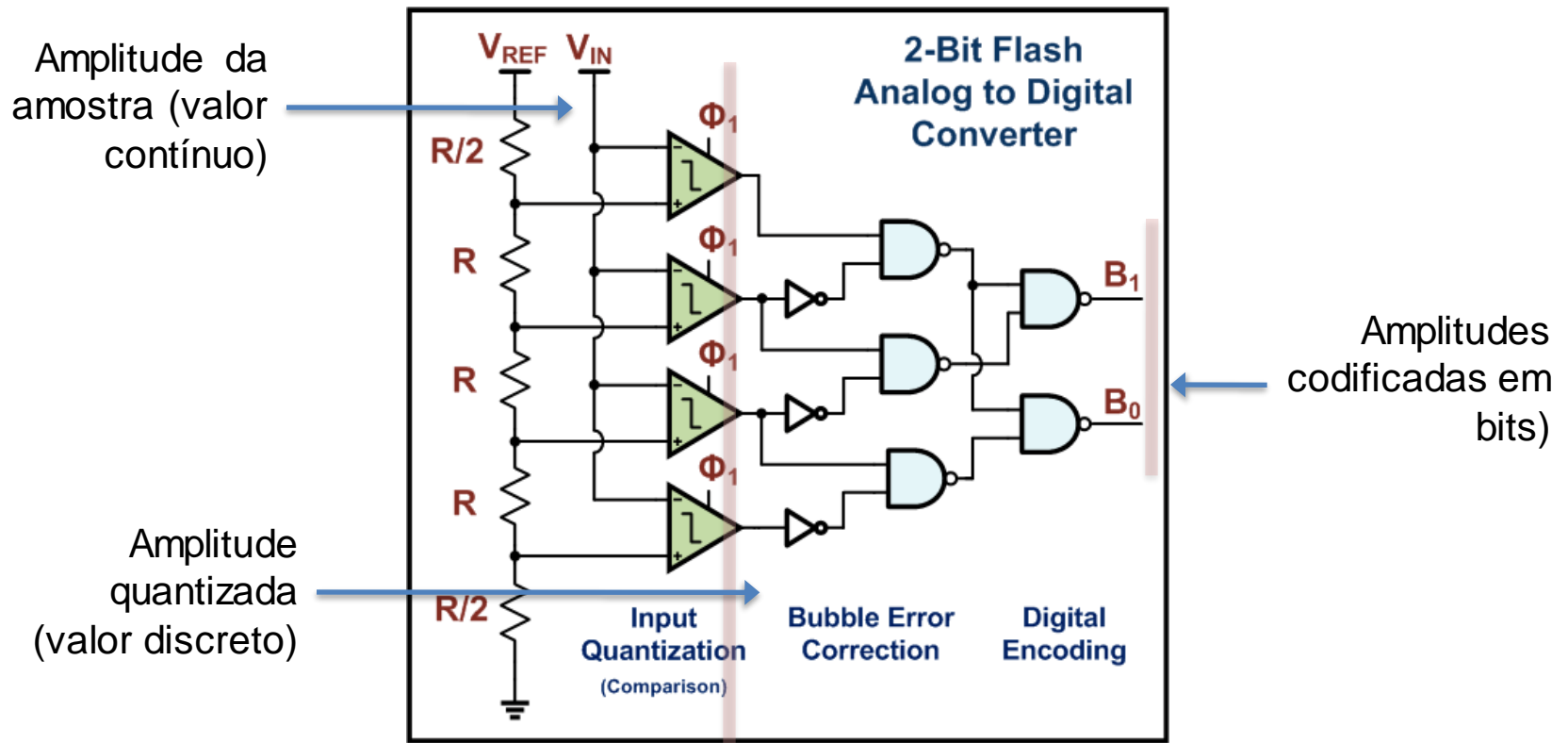


Conversores AD e DA

- Conversores AD e DA são os dispositivos que fazem a amostragem e a reconstrução do sinal original
 - ✓ **Conversores Analógico-Digital ou ADC (analog to digital converter)** usam amostradores “sample & hold” que seguram o valor (amplitude) de uma amostra até a próxima amostra, e convertem este valor para uma representação binária (digital)
 - ✓ **Conversores Digital-Analógico ou DAC (digital to analog converter)** convertem as amostras de valor binário para amostras com valor contínuo, e então submetem o sinal a interpoladores (ex: filtros passa-baixa) que reconstroem um sinal contínuo (analógico)

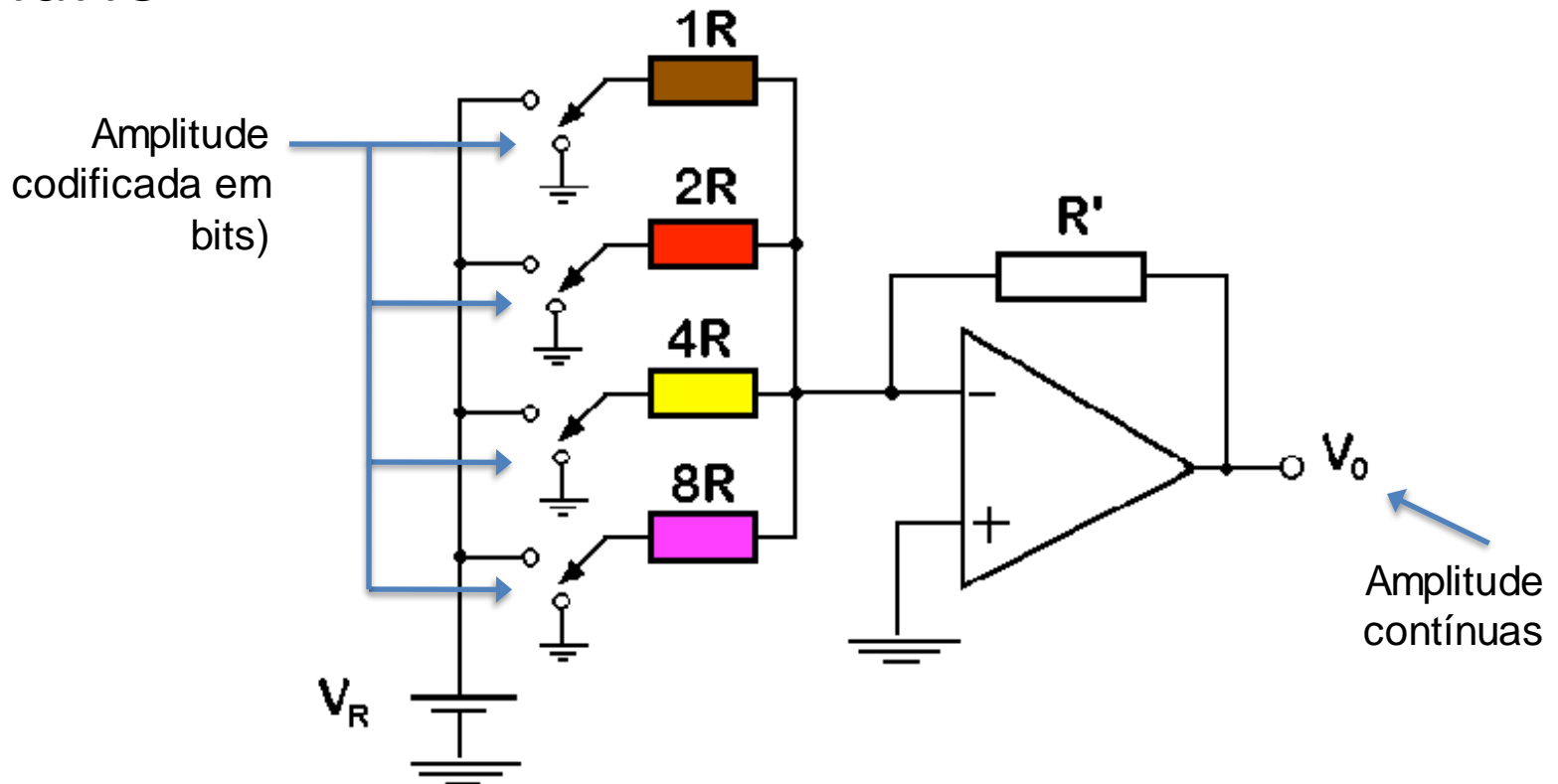
Conversor AD

- Ex: ADC (analog to digital converter) com *flash*
 - $2^n - 1$ comparadores para n bits de resolução



Conversor DA

- Ex: conversor DAC com resistências de peso binário



Conversor DA

- Ex: conversor DAC usando banco de resistências R-2R

