

# Áudio Digital em Gravações Musicais

A popularização das gravações caseiras aconteceu porque os equipamentos analógicos estavam ficando obsoletos e sendo substituídos por novas gerações de interfaces de áudio e equipamentos digitais mais baratos e fáceis de usar que os antigos. E essa tendência continuou desde então.

Hoje em dia, o áudio digital está presente em praticamente todos os estúdios, tanto profissionais quanto amadores. Ainda assim, poucas pessoas compreendem de verdade o que há por trás dele.

Então, para ajudar você nessa questão apresentamos uma introdução aos conceitos básicos relacionados à utilização do Áudio Digital em Gravações Musicais.

São 9 tópicos a serem abordados:

1. A Ascensão da Era Digital
2. Entendendo os Conversores Digitais
3. Taxa de Amostragem
4. Intensidade de Bit
5. Erro de Quantização
6. Dither
7. Latência
8. Relógios Mestre
9. Codificação MP3

## 1. A Ascensão da Era Digital



Embora o áudio digital seja padrão na indústria fonográfica atual, a situação nem sempre foi essa. Originalmente, a informação musical existia apenas na forma de ondas sonoras no ar.

Então, a tecnologia avançou, e as pessoas descobriram novas formas de convertê-la para outros formatos, incluindo:

- notas em uma página
- sinais elétricos em um cabo
- ondas de rádio na atmosfera
- protuberâncias nos discos de vinil

Mas afinal de contas, com a ascensão dos computadores, o áudio digital tornou-se o formato dominante para gravações musicais, porque permitia que as músicas fossem facilmente copiadas e transportadas de forma gratuita.

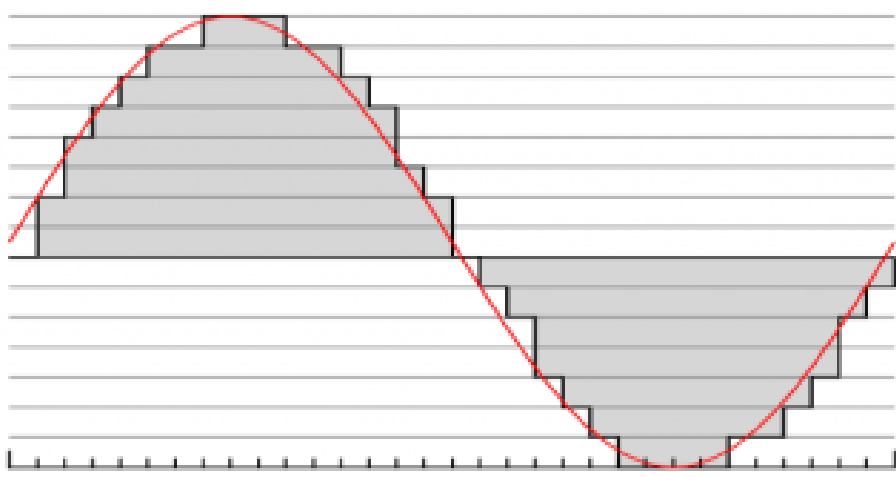
E o dispositivo que tornou tudo isso possível foi o conversor digital.

## 2. Entendendo os Conversores Digitais

Nos estúdios de gravação os conversores digitais existem de 2 formas:

- Como um dispositivo independente em estúdios de ponta, ou
- Como parte da interface de áudio em home studios.

Para converter áudio em código binário, os dispositivos capturam centenas de milhares de imagens (samples) por segundo, para construir um quadro “aproximado” da forma da onda analógica. As imagens dos samples não são exatas, porque entre elas, o conversor precisa basicamente adivinhar o que está acontecendo.



Como você pode ver na figura acima, onde:

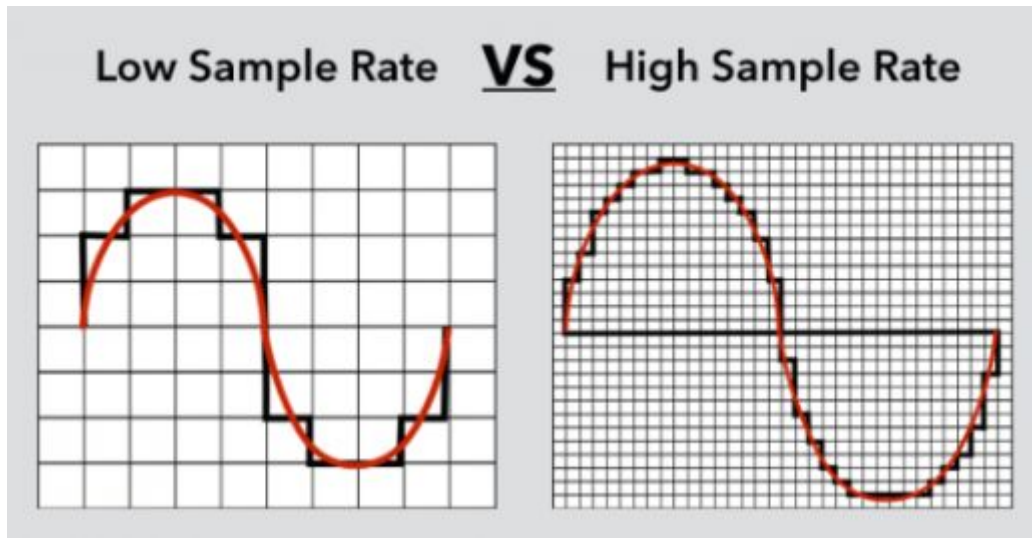
- a curva vermelha é o sinal analógico
- a linha preta é a conversão...

Os resultados não são perfeitos, mas são bons o suficiente para tornar a qualidade do som excelente.

Porém, o quão excelente, exatamente, depende sobretudo de você.

### 3. Taxa de Amostragem

Observe a seguinte imagem:



Como você pode ver, capturando mais amostras (samples) por segundo, as taxas de amostragem mais altas:

- Reúnem mais informações reais,
- Dependem menos de conjecturas,
- Resultam numa imagem muito mais fiel do sinal analógico

E o resultado final é, obviamente, a obtenção de uma melhor qualidade sonora.

Agora, vamos falar de números específicos:

As taxas de amostragem mais comuns no áudio de qualidade profissional:

- 44.1 kHz (CDs de Áudio)
- 48 kHz
- 88.2 kHz
- 96 kHz
- 192 kHz

O valor mínimo de 44.1 kHz existe devido a um princípio matemático conhecido como Teorema da Amostragem de Nyquist-Shannon.

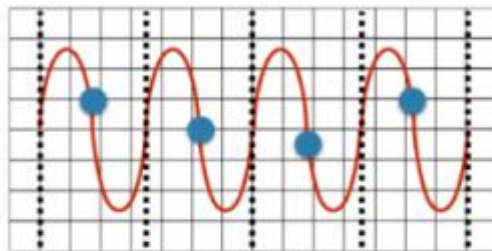
Para gravar sinais digitais com precisão, os conversores precisam capturar o espectro completo da audição humana, entre 20Hz a 20kHz.

De acordo com o Teorema da Amostragem de Nyquist-Shannon a captura de uma frequência específica requer pelo menos 2 samples de cada ciclo para medir tanto os

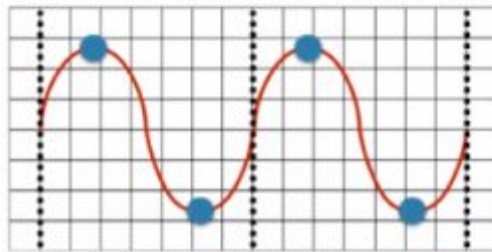
pontos superiores quanto os inferiores da onda. Isso significa que, a gravação das frequências de até 20 kHz requer uma taxa de amostragem de 40 kHz ou mais. Motivo pelo qual, os CDs de áudio estão um pouco acima disso, a 44.1 kHz.

## Nyquist Shannon Sampling Theorem

ONE Digital Sample per Cycle



TWO+ Digital Samples per Cycle



## O Custo das Altas Taxas de Amostragem

Embora as altas taxas de amostragem produzam uma melhor qualidade sonora, os benefícios não são de graça.

Os custos incluem:

- Cargas de processamento maiores
- Quantidade menor de canais
- Arquivos de áudio maiores

Então, uma parte sempre acaba sendo comprometida. Os estúdios profissionais conseguem trabalhar com taxas de amostragem maiores com mais facilidade, porque utilizam equipamentos melhores. No entanto, nos home studios, o melhor é trabalhar a 48 kHz.

## 4. Profundidade de Bit (bit-depth)

Para entender o que é profundidade de bit primeiro precisamos discutir o que são bits. Sendo a abreviação de dígito binário, um bit é uma unidade de código binário que contém o valor 1 ou 0. Quanto mais bits são usados, mais combinações são possíveis. Por exemplo, como você pode ver na figura abaixo, 4 bits resultam num total de 16 combinações.

## 4 Bits = 16 Combinations

0000	1000	1001	1110
0001	0011	0101	1101
0010	0110	1010	1011
0100	1100	0111	1111

Quando usados para codificar informações, cada um desses números é associado a um valor específico. Aumentando a quantidade de bits, o número de valores possíveis cresce exponencialmente.

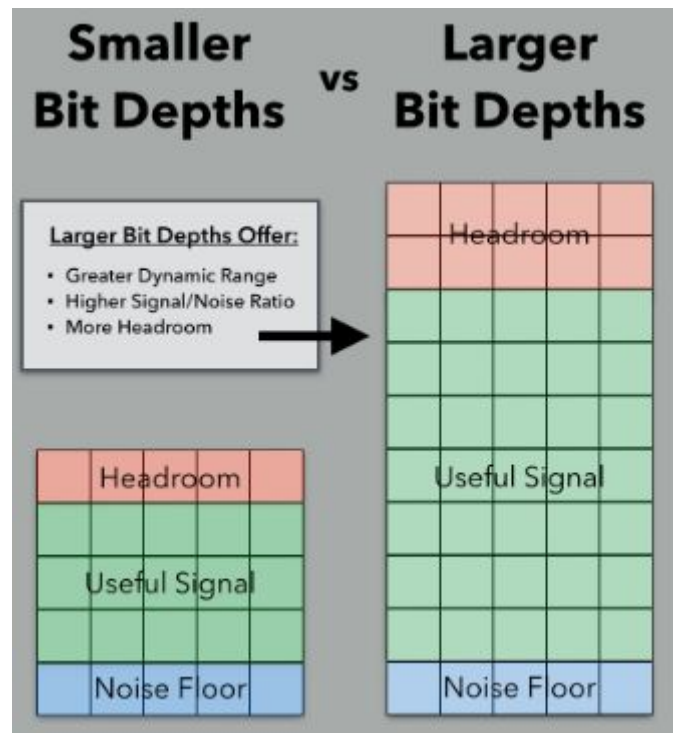
- 4 Bits = 16 valores possíveis
- 8 Bits = 256 valores possíveis
- 16 Bits = 16.536 valores possíveis
- 24 Bits = 16.777.215 de valores possíveis

Com a profundidade do bit no áudio digital, cada valor é atribuído a uma amplitude específica na forma da onda do áudio. Quanto maior a profundidade do bit, maior será a quantidade de incrementos existente entre sons altos e baixos e maior será a faixa dinâmica da gravação. Uma boa regra geral é lembrar que: Para cada “bit” extra, a faixa dinâmica aumenta em 6dB.

Por exemplo:

- 4 Bits = 24 dB
- 8 Bits = 48 dB
- 16 Bits = 96 dB
- 24 Bits = 144 dB

Por fim, isso significa que mais profundidade de bit equivale a menos ruído porque, ao adicionar essa altura extra, o sinal útil (na extremidade alta do espectro) pode ser gravado acima do patamar de ruído (na extremidade suave do espectro).



## 5. Erro de Quantização

Parece impressionante que uma gravação de 24 bits possa produzir quase 17 milhões de possibilidades de valores, certo? Ainda assim, isso é bem menos do que o número infinito de valores possíveis que existem no sinal analógico. Então, no caso de quase todas os samples, o valor real está entre dois valores possíveis. A solução do conversor é simplesmente arredondá-los ou “quantizá-los” para o valor mais próximo. A distorção resultante, conhecida como erro de quantização, acontece em 2 etapas do processo de gravação:

1. no começo, durante a conversão A/D
2. no fim, durante a masterização

Na masterização, a taxa de amostragem/profundidade do bit da faixa final geralmente é reduzida ao ser convertida para um formato digital (CD, mp3 e etc.). Quando isso acontece, algumas informações são deletadas e “requantificadas”, resultando na distorção adicional do som.

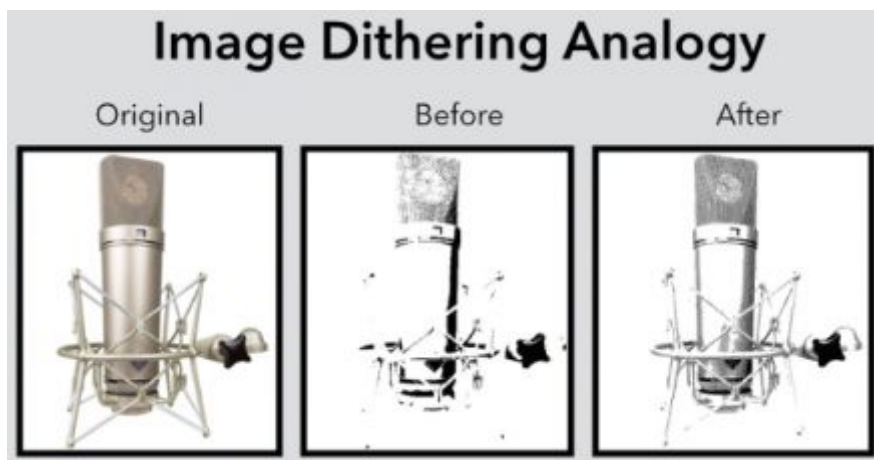
Para lidar com esse problema, há uma solução prática, conhecida como...

## 6. Dither

Ao reduzir um arquivo de 24 bits para 16 bits, o dither é utilizado essencialmente para mascarar uma grande porção da distorção resultante acrescentando um baixo nível de “ruído aleatório” ao sinal de áudio. Já que o conceito é difícil de visualizar com áudio, a analogia mais popular utilizada para explicar esse processo é fazendo dithering com imagens.

Funciona assim:

Quando uma foto colorida é convertida para preto e branco, algumas conjecturas matemáticas são realizadas para determinar se cada pixel colorido deve ser “quantizado” para um pixel preto ou branco. Da mesma forma que acontece na quantização das amostras de áudio digital. Como você pode ver na figura abaixo, a imagem com legenda “before” (ou “antes”) está com uma qualidade bem ruim, correto?



Mas através do processo de dithering uma pequena quantidade de pixels brancos é randomizada nas regiões pretas e uma pequena quantidade de pixels pretos é randomizada nas regiões brancas, adicionando assim, um “ruído aleatório” à imagem, a figura com legenda “after” tem uma qualidade muito superior. Com o processo de dithering de áudio, o conceito é muito parecido.

## 7. Latência

Uma grande falha dos estúdios digitais de hoje em dia é a quantidade de atraso de tempo (latência) que se acumula na cadeia de sinais, especialmente com as DAW's. Com todos os cálculos que acontecem, demora poucos milissegundos ou poucas dúzias de milissegundos para o sinal de áudio sair do sistema.

- Com 0 a 11 ms de delay – é tão curto que as pessoas comuns não percebem nada.
- Com 11 a 22 ms – é possível ouvir um efeito chato de slapback, com o qual você pode demorar para se acostumar.

- Com 22 ms ou + – o delay torna impossível interagir ou cantar em sincronismo com a faixa.

Em uma cadeia de sinais digitais típica, existem 4 recursos que aumentam tempo total de delay:

1. Conversão A/D
2. Buffering do DAW
3. Plugin de Delay
4. Conversão D/A

As conversões A/D e D/A são as que menos causam problemas, aumentando o tempo total de delay em, no máximo, 5 ms. No entanto, o buffer da sua DAW e de certos plugins (incluindo compressores e instrumentos virtuais), podem aumentar o tempo de delay em até 20, 30, 40 ms ou mais.

Para mantê-lo em valores baixos:

1. Desative todos plugins desnecessários enquanto você estiver gravando.
2. Ajuste as configurações de buffer da sua DAW para descobrir qual é o menor tempo que o seu computador suporta sem travar.

Você vai perceber que os tempos de buffer são medidos por samples e não em milissegundos. Para converter de um para outro:

- Divida a quantidade de samples pela taxa de amostragem (em kHz) da sessão para descobrir o tempo de latência em milissegundos.

Por exemplo:  $1024 \text{ samples} \div 44.1 \text{ kHz} = 23 \text{ ms}$

Se você detesta fazer contas, grave os múltiplos aproximados de 44.1:

- 256 amostras = 6 ms
- 512 amostras = 12 ms
- 1024 amostras = 24 ms

Na maioria dos casos, esses passos devem diminuir a latência a um nível controlável, mas se o seu equipamento for muito velho ou de qualidade duvidosa, talvez você não consiga. Neste caso, o último recurso é usar um botão de nome “mix” ou “blend” que muitas interfaces baratas possuem (ambas as palavras significam “mistura” ou “combinação” em inglês), que permite que você combine o playback da sessão com o “sinal ao vivo” que está sendo gravado.

Dividindo o sinal ao vivo do microfone/instrumento e enviando metade dele para o computador, para ser gravado, e a outra metade diretamente para os fones de ouvido, você evita a latência dividindo a cadeia de sinais inteira.



O lado ruim dessa técnica é que você ouve o sinal ao vivo completamente seco, sem nenhum efeito.

Com sorte, uma vez que os computadores estão ficando mais rápidos, logo isso não será mais um problema.

## 8. Master clock



Sempre que dois ou mais dispositivos trocam informações digitais em tempo real os relógios internos deles devem estar sincronizados para que os samples fiquem alinhados impedindo que ocorram aqueles cliques e “pops” chatos no áudio.

Para sincronizá-los, um dispositivo atua como o “mestre”, e os demais, como “escravos”.

Nos home studios simples, o clock da interface de áudio geralmente é o líder por padrão. Nos estúdios

profissionais, que necessitam de conversão digital premium e roteamento de sinal complexo, um dispositivo especial individual, conhecido como relógio master digital (também chamado de word clock) pode ser utilizado. Segundo o que muitos donos afirmam, os benefícios sonoros desses relógios de ponta podem ser muito menos sutis do que você imagina.

## 9. Decodificação Mp3/AAC

No mundo de hoje, arquivos de áudio comprimidos são o padrão do áudio digital, porque devido a limitações de espaço de armazenamento dos iPods e smartphones e de velocidade de conexão para o streaming, todos os arquivos precisam ter o menor tamanho possível.

Utilizando um método de “compressão com perda de dados”, os formatos mp3, AAC e outros similares podem diminuir os arquivos a 1/10 do tamanho original. O processo de codificação funciona usando um princípio da audição humana, conhecido como “mascaramento auditivo” que torna possível deletar muita informação musical, mantendo níveis aceitáveis de qualidade sonora para a maioria dos ouvintes. Os engenheiros de áudio experientes podem até notar algumas diferenças, mas os consumidores que não são da área provavelmente não perceberão.

A quantidade exata de informação que é deletada, depende da taxa de bits do arquivo.

Quando as taxas são mais altas, menos informação é removida e mais detalhes são preservados do arquivo.

Por exemplo, nas faixas mp3:

- 320 kbit/s é a maior taxa de bits possível
- 128 kbit/s é a taxa mínima recomendável
- 256 kbit/s é o “sweet spot” ou ponto ideal que a maioria das pessoas prefere

Para descobrir qual é o formato e a taxa de bits ideal para as suas músicas, sempre verifique as recomendações dadas pelo site ou aplicativo para onde pretende enviá-las (iTunes, YouTube, Soundcloud, etc.).