WHITE PAPER

Faixa de captação do microfone – fatores que influenciam o desempenho de áudio

Setembro 2026

AXIS COMMUNICATIONS

Índice

1	Introdução	3
2	Faixa de gravação física	3
3	Gravações em diferentes ambientes	5
4	Resultados das medições obtidos a partir de gravações em diferentes ambientes	7
1	Apêndice	9

1 Introdução

É raramente simples definir uma faixa máxima definitiva de captação para dispositivos de gravação de áudio, já que o desempenho está ligado ao ambiente acústico. Embora as especificações técnicas forneçam uma linha de base, a faixa prática de gravação depende de uma interação complexa entre fatores ambientais e acústicos que variam de uma instalação para outra.

Este white paper analisa os principais fatores que determinam até que distância um microfone pode captar o som de maneira confiável, com foco na inteligibilidade da fala como referência principal. Esses fatores incluem o nível de ruído de fundo, as características espectrais desse ruído, o tempo de reverberação, a diretividade do microfone e o ruído próprio do microfone e do dispositivo de gravação.

Para estabelecer uma linha de base prática, utilizamos a voz humana como fonte sonora de referência. O ruído é medido com um filtro de ponderação A para obter uma melhor correlação entre o valor da medição e o nível de ruído percebido. A relação sinal-ruído (SNR) da gravação da fala é calculada como a diferença entre o nível da voz e o nível de ruído ponderado A, em decibéis *(consulte o apêndice para obter informações sobre SNR)*. Testes empíricos mostram que a inteligibilidade da fala é aceitável até uma relação sinal-ruído (SNR) de aproximadamente 0 dB (ou seja, quando o ruído e o som têm intensidades iguais), desde que o ruído apresente uma resposta em frequência relativamente plana dentro da faixa de frequências da fala.

Ao compreender essas variáveis, os instaladores e projetistas de sistemas podem tomar decisões mais bem fundamentadas ao selecionar e posicionar dispositivos de áudio em cenários reais.

2 Faixa de gravação física

A capacidade de gravação de um microfone geralmente depende do ruído de fundo do ambiente. Os principais fatores que limitam a faixa de gravação física são:

- **Nível de pressão sonora:** Nem todos os sons percorrem a mesma distância. Por exemplo, um tiro produz um nível de pressão sonora extremamente alto e pode ser ouvido a até um quilômetro de distância. A voz humana, no entanto, é mais baixa e se dissipa mais rapidamente com a distância. Ao estimar uma faixa realista de captação, é preciso primeiro definir o tipo de som que você precisa capturar, para que o nível da fonte determine o limite máximo do que é possível realizar na gravação.
- **Ruído de fundo:** quanto mais barulhento for o ambiente, mais difícil será para um microfone isolar e captar o som desejado. O ruído de fundo varia de um ambiente para outro e compete diretamente com o som alvo. Isso inclui o ruído do vento, do trânsito, de máquinas e da multidão, entre outros. Quanto maior for o ruído de fundo no local onde o microfone está instalado, menor será a faixa de captação efetiva. Isso explica por que o mesmo microfone pode apresentar um desempenho muito diferente em um corredor silencioso em comparação com um cruzamento movimentado.
- **Perfil espectral do ruído:** assim como o volume do ruído é importante, as frequências que ele ocupa também são. A fala humana situa-se principalmente na faixa de frequências médias, aproximadamente entre 300 Hz e 3.400 Hz. Se o ruído de fundo estiver concentrado em frequências muito baixas ou muito altas, é menos provável que interfira na inteligibilidade da fala. No entanto, se o ruído se sobrepor à mesma faixa de frequência do som alvo, torna-se difícil distinguir a fala do ruído, mesmo em níveis moderados de ruído.
- **Tempo de reverberação no local do microfone:** A reverberação é a persistência do som após o cessação da fonte original, causada por reflexões nas paredes, tetos, pisos e outras superfícies. Um tempo de reverberação prolongado pode prejudicar a inteligibilidade da fala, mesmo em situações em que o nível sonoro geral e a relação sinal-ruído pareçam aceitáveis. Em espaços com muita reverberação, como grandes salões vazios ou salas com piso de azulejos, o som da fala chega ao microfone acompanhado de uma quantidade infinita de reflexões da sala, o que torna o som distorcido. Isso dificulta que os ouvintes e os sistemas de gravação distingam claramente as palavras individualmente.
- **Direcionalidade da configuração do microfone:** os microfones variam quanto à sensibilidade aos sons provenientes de diferentes direções. Um microfone omnidirecional capta o som igualmente de todas as direções, o que também significa que capta ruídos de fundo e reverberação de todas as direções. Uma configuração direcional, como vários microfones configurados para a formação de feixe, concentra a sensibilidade em uma direção específica, o que reduz efetivamente a influência do ruído e da reverberação

provenientes de outros ângulos. Isso pode ampliar significativamente a faixa de captação útil em ambientes ruidosos ou com reverberação, mas apenas na direção para a qual o conjunto de microfones está voltado.

- **Ruído próprio de microfones e dispositivos de gravação:** Todo microfone e dispositivo de gravação de áudio gera uma pequena quantidade de ruído elétrico interno, conhecido como ruído próprio. Isso define um limite mínimo absoluto para o que o sistema pode capturar. Por mais silencioso que seja o ambiente, os sons que ficam abaixo do nível de ruído próprio simplesmente não podem ser realizados com clareza. Ao comparar microfones, uma especificação de ruído próprio mais baixa geralmente significa que o microfone consegue captar melhor sons mais baixos, o que pode proporcionar uma faixa de captação maior se o nível de ruído e o tempo de reverberação forem baixos o suficiente.

Para simplificar a análise, podemos partir de um cenário idealizado, no qual não há ruído de fundo, reflexões nem reverberação, e usar a voz humana como nossa fonte sonora de referência. Nessas condições, a questão principal passa a ser: qual relação sinal-ruído (SNR) devemos usar como limite para uma faixa de captação aceitável?

Os testes empíricos nos dão uma resposta clara. Quando aplicamos um filtro de ponderação A ao ruído do sistema e comparamos o nível médio quadrático médio (RMS) do ruído com o nível médio RMS da fala, a inteligibilidade da fala começa a diminuir somente quando a relação sinal-ruído (SNR) cai abaixo de 0 dB. Portanto, utilizamos uma relação sinal-ruído (SNR) de 0 dB como nosso critério limitante para a gravação da voz humana. Isso se aplica desde que o ruído tenha uma resposta em frequência relativamente plana, semelhante à do ruído branco, e esteja dentro da faixa de frequências da fala, o que normalmente é o caso do ruído próprio do microfone.

Por exemplo, um microfone digital pode ter uma sensibilidade de -36 dBFSrms *(consulte o apêndice para obter informações sobre dBFSrms)* quando um som de 1 kHz é reproduzido a 94 dBspl (um sinal sinusoidal digitalizado em escala completa é definido aqui como 0 dBFSrms, de acordo com as folhas de dados do microfone) *(consulte o apêndice para obter informações sobre dBspl)*. Isso se aplica a muitos produtos da Axis com microfones MEMS digitais integrados quando o ganho do microfone está ajustado em 0 dB. Quando você coloca esse microfone em um ambiente muito silencioso, em que o nível de ruído de fundo fica abaixo de 10 dBspl(A) *(consulte o apêndice para obter informações sobre dBspl(A))*, o ruído interno do microfone passa a ser predominante. Neste exemplo, a folha de dados do microfone especifica uma saída digital de -105 dBFSrms(A). A relação sinal-ruído (SNR) do microfone é a diferença entre o nível de referência e o ruído de fundo: –36 – (–105) = 69 dB.

Convertemos esse ruído elétrico de fundo em um nível de ruído acústico equivalente: 94 – 69 = 25 dBspl(A).

Agora, vamos calcular a distância máxima na qual é possível realizar a gravação de uma voz humana com uma relação sinal-ruído (SNR) de 0 dB. De acordo com as normas da ITU, a voz humana possui um nível médio nominal de fala (ASL) de aproximadamente 57 dBspl a 1 m. Em espaço livre, sem objetos que reflitam o som, o nível de pressão sonora diminui com a distância, de acordo com a lei do quadrado inverso:

$$dB = 20\log\frac{reference\ distance}{test\ distance}$$

.

Se a distância de referência for de 1 m, para atingir uma relação sinal-ruído (SNR) de 0 dB neste microfone, o nível da fala deve cair para 25 dBspl(A); uma redução de 32 dB em relação ao nível de referência de 1 m. Reorganizando a fórmula, obtemos:

$$test\ distance = \frac{reference\ distance}{10^{\frac{-(57-noise\ level)}{20}}} = \frac{1}{10^{\frac{-32}{20}}} = 39.8\text{m}$$

.

Isso significa que, em espaço aberto, sem ruído de fundo, sem vento e sem reverberação, esse microfone pode, teoricamente, realizar a gravação de fala inteligível a aproximadamente 40 m. Essa é a faixa máxima de captação teórica.

No entanto, os fatores ambientais têm um impacto maior no nível de ruído do que o ruído de fundo do microfone interno. Um escritório silencioso após o horário de trabalho, com os sistemas de ventilação ainda em

funcionamento, apresenta um nível de ruído de fundo de cerca de 28 a 32 dBspl. Isso por si só já pode reduzir a faixa efetiva de captação em cerca de 50%, diminuindo a faixa máxima para cerca de 20 m. Se fizermos uma gravação ao ar livre em um ambiente urbano, haverá uma combinação de ruídos: ruído do vento, ruído do trânsito, ruídos decorrentes da atividade humana e assim por diante. Cada nível de ruído e cada espectro de ruído terá sua própria faixa de captação correspondente quando se trata de um microfone omnidirecional. Nesse caso, a faixa de captação pode ser de, no máximo, 2 a 4 m, se o ruído do tráfego for intenso.

Uma consideração prática é que as gravações de voz humana feitas no limite de 0 dBSNR ou próximo a ele exigem um ganho digital significativo para se tornarem audíveis, pois o nível da voz gravada é tão baixo que, de outra forma, ficaria abaixo do nosso próprio limite auditivo durante a reprodução. Nos exemplos de gravação a seguir, os arquivos de áudio foram amplificados em +40 dB para compensar o baixo nível da voz.

3 Gravações em diferentes ambientes

Esta seção detalha os locais de gravação, acompanhados de capturas de tela das gravações de vídeo e amostras de voz captadas em várias distâncias.

Exemplo 1: O local fica ao ar livre, com ruído de tráfego proveniente de uma rodovia. O nível de ruído de fundo varia de 62 a 68 dBspl[A] e também é possível ouvir algum ruído do vento.



Figura 3.1 *O local da gravação fica próximo a uma rodovia (veja o ponto vermelho).*



Figura 3.2 *No local de gravação, a distância máxima de gravação é de 4 m de uma câmera corporal da Axis.*

De preferência, use um par de fones de ouvido ao ouvir a gravação.

Exemplo 2: Um corredor silencioso com um nível de ruído de fundo de 32 dBspl[A]. Aqui, o corredor funciona como um longo guia de ondas acústicas. O som se propaga pelo corredor, refletindo nas superfícies refletivas, e chega aos microfones mais alto do que chegaria em um espaço livre e à mesma distância.



Figura 3.3 *A distância máxima de gravação é de 27 m.*

De preferência, use um par de fones de ouvido ao ouvir a gravação.

Exemplo 3: Um amplo saguão com um nível de ruído de fundo de 42 dBspl[A]. Aqui, as paredes possuem materiais de absorção acústica atrás dos painéis difusores, o que faz com que a maioria das reflexões iniciais desapareça. A partir de uma distância maior, o som se propaga pelo canal acústico entre o piso e o teto refletor.

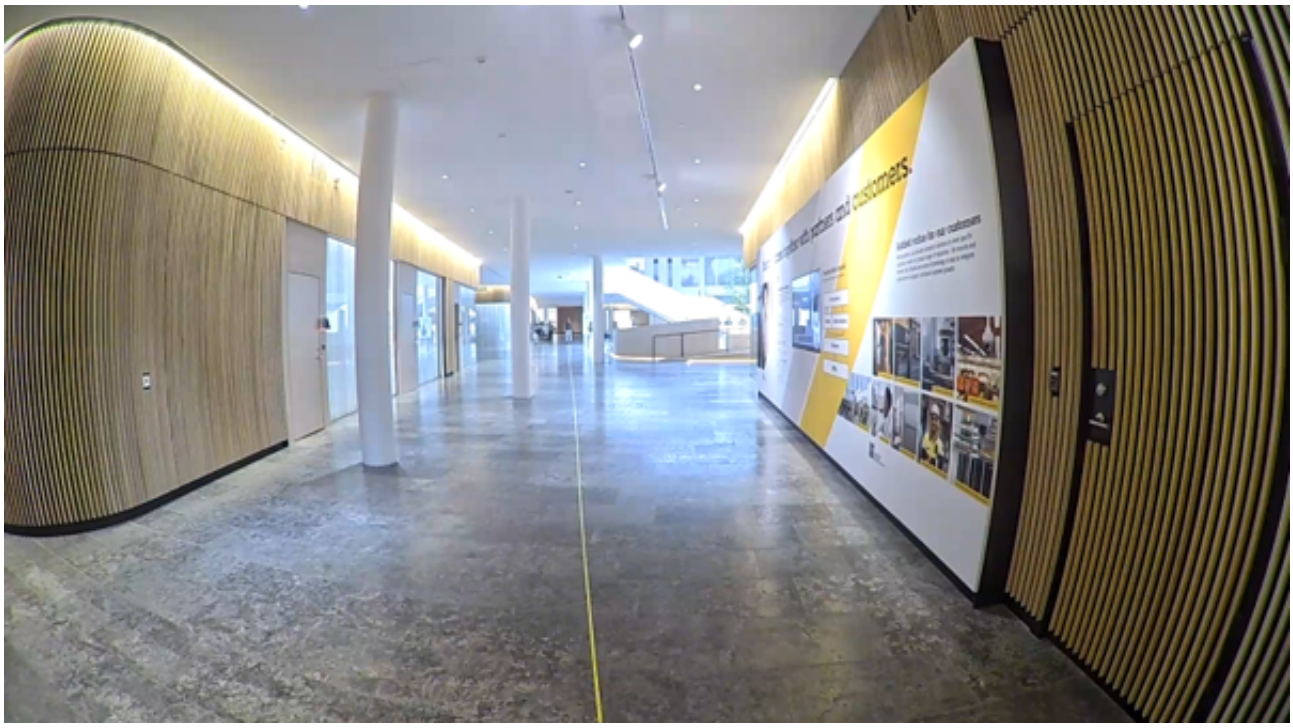


Figura 3.4 *A uma distância máxima de gravação de 40 m, o locutor mal é visível, mas ainda assim pode ser ouvido.*

De preferência, use um par de fones de ouvido ao ouvir a gravação.

4 Resultados das medições obtidos a partir de gravações em diferentes ambientes

O foco da análise está na captação de voz. Sons mais intensos, como tiros, serão ouvidos a uma distância maior. Analisamos e medimos cada frase das gravações de áudio e estimamos a relação sinal-ruído (SNR) para cada frase gravada.

A SNR foi estimada como fala + ruído - ruído de fundo. Uma maneira mais correta seria calcular o nível de fala - o nível de ruído, mas não há como medir o nível de fala sem o ruído. Isso significa que, quando o nível da fala está próximo do nível de ruído, a precisão da SNR diminui.

Localização e ruído de fundo:	No exterior, com ruído do trânsito 61–68 dBspl[A]	Em um corredor silencioso 32 dBspl[A]	Em um grande saguão 42 dBspl[A]	Espaço livre, sem reflexos, silêncio total
Distância (m)	**SNR [dB] (estimado)**	**SNR [dB] (estimado)**	**SNR ([dB] estimado)**	**SNR teórico = ruído do microfone no nível da fala [dB]**
1	2,4	29,4	21,5	32,0
2	1.0	25,6	17,2	26,0
4	0,5	23,4	14,2	20,0
8	0,4	18,9	11,2	13,9

16	Não audível	17,5	6,3	7,9
24	Não audível	14,8	5,4	4,4
32	Não audível	Corredor muito curto	4,6	1,9
40	Não audível	Corredor muito curto	4,7	0,0

Em teoria, o nível de fala deve diminuir em -6 dB quando a distância for duplicada. Isso também significaria que a SNR deveria diminuir proporcionalmente. Nesta tabela, podemos ver que a relação sinal-ruído (SNR) é melhor do que o esperado no corredor e no saguão. A razão é que os reflexos na sala aumentaram os níveis de pressão sonora. Especialmente no corredor, as paredes, o piso e o teto funcionam como um guia de ondas para o som. Isso faz com que o nível sonoro e a SNR sejam cerca de 10 dB melhores do que a SNR teórica no espaço livre, quando testado a uma distância de 24 m, em um corredor.

Sempre que o ruído de fundo excede o ruído próprio acústico equivalente de um único microfone omnidirecional, a faixa de captação depende inteiramente do ruído de fundo do ambiente, e não do microfone.

Apêndice 1 Apêndice

- **Nível de pressão sonora em decibéis (dBspl):** O nível de pressão sonora em decibéis (dB) com referência a 20 µPa. Ao medir os níveis de ruído e de voz, costuma-se utilizar os níveis RMS integrados no tempo. Isso é, às vezes, chamado de nível de potência.

$$RMS = \sqrt{\frac{1}{T2 - T1}\int_{T1}^{T2}[f(t)]^2\,dt}$$

  Os dois tempos de integração mais comuns são o "RMS rápido", com uma janela de tempo de 125 ms, e o "RMS lento", com uma janela de tempo de 1 segundo. Para a análise do ruído de fundo, o nível costuma ser calculado como média ao longo de um período mais longo.

- **Nível de pressão sonora ponderado pelo filtro de ponderação A (dBspl[A]):** às vezes, isso é escrito como dB(A). É o valor ponderado A do nível em dBspl. A ponderação A é uma forma de filtrar o sinal do microfone para obter uma melhor correspondência entre a audição humana e o nível de pressão sonora medido. Isso ocorre porque a audição humana apresenta uma sensibilidade reduzida para frequências baixas abaixo de 500 Hz e frequências altas acima de 10 kHz.

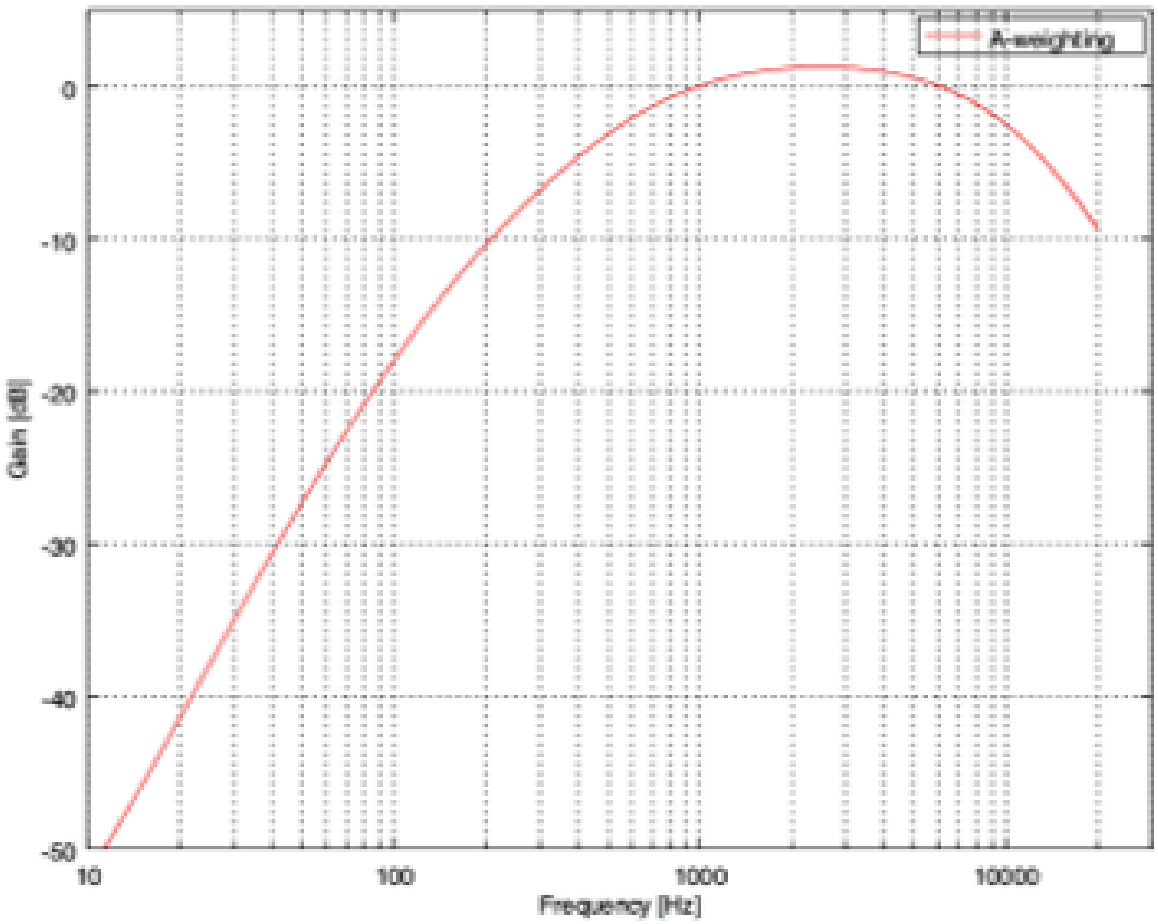


Figura .1 *Filtro de ponderação A.*

- **O pico em escala completa em decibéis (dBFSpeak/dBFSp)** refere-se à amplitude de um sinal de áudio digital em relação à representação digital máxima disponível (escala completa). Uma onda senoidal com amplitude digital máxima, sem distorção por clipping, tem um valor de pico de 0 dBFSp.

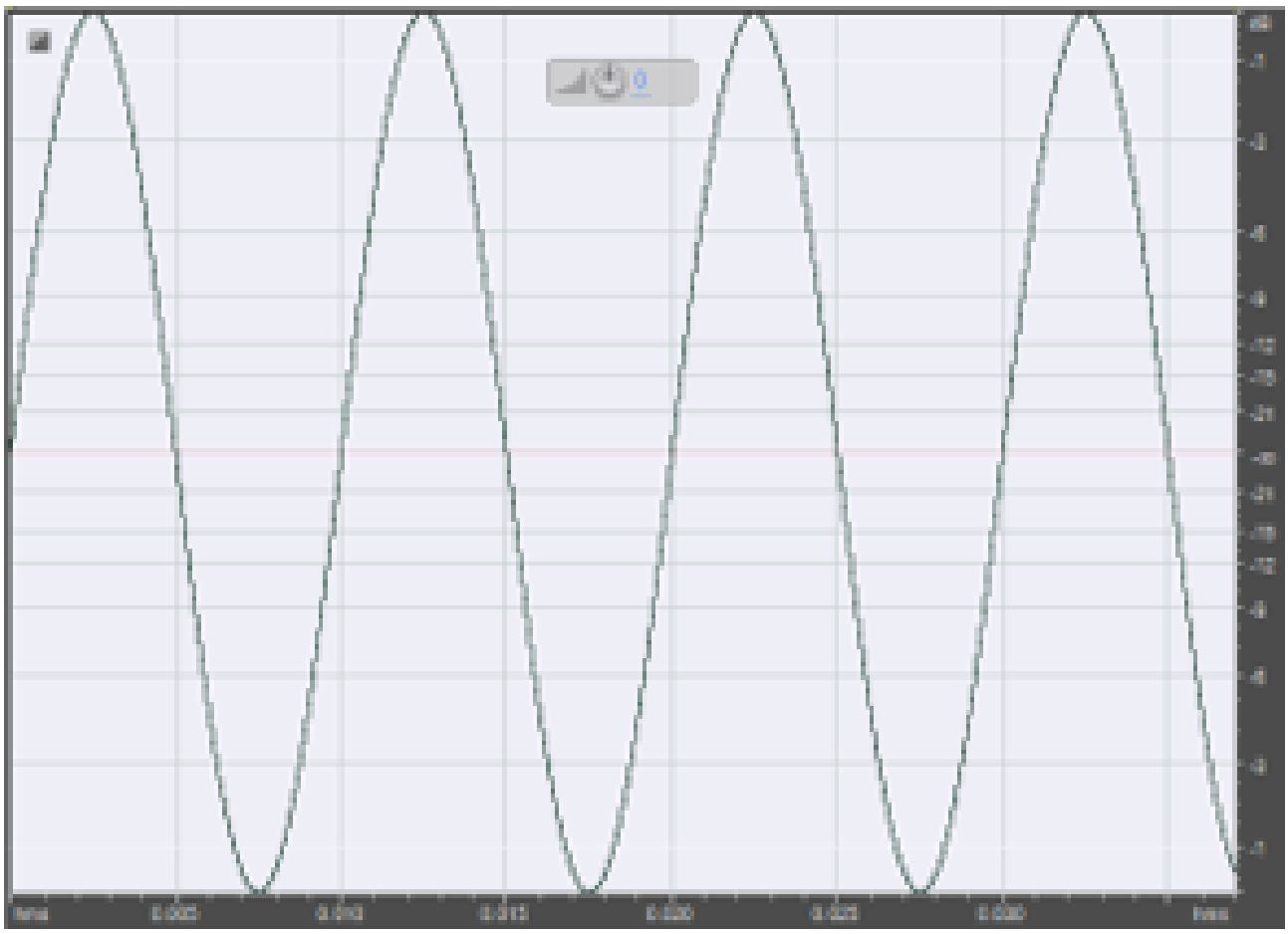


Figura .2 *Um exemplo de sinal com amplitude de 0 dBFSpeak.*

- **dBFSrms** refere-se ao nível de potência de áudio digital de um sinal digital. Um sinal senoidal em escala real com 0 dBFSpeak pode apresentar diferentes níveis RMS, dependendo de como a unidade for definida. Nas folhas de dados de microfones, uma onda senoidal em escala total é definida como 0 dBFSrms; por esse mesmo motivo, utilizamos essa definição neste white paper. De modo geral, a outra definição de um sinal sinusoidal em escala total — -3 dBFSrms — é matematicamente mais correta e mais consistente com a forma como os níveis RMS e os níveis de pico são definidos no domínio analógico. Uma onda quadrada em escala total corresponde a 0 dBFSrms nessa outra definição.
- **A sensibilidade do microfone** é o nível de saída de um microfone quando a entrada acústica é um tom sinusoidal de 1 kHz com 1 Pa, 0 dB a ou 94 dBspl (três formas de expressar o mesmo nível de pressão sonora).
- **Relação Sinal-Ruído (SNR): Quando** os níveis de áudio são expressos em dB, essa relação é calculada como a diferença entre o nível do sinal e o nível de ruído.
- **A relação sinal-ruído (SNR) do microfone** é calculada como a diferença no nível de saída entre um tom de teste de 1 kHz a 94 dBspl e o ruído próprio do microfone, quando este se encontra em silêncio total. Uma relação sinal-ruído (SNR) típica para um microfone MEMS digital é de 69 dBspl[A].
- **A relação sinal-ruído (SNR) da fala** neste white paper, é uma comparação entre a fala humana e o nível de ruído ponderado A quando não há fala. A relação sinal-ruído (SNR) é medida em decibéis. Utilizamos a relação sinal-ruído (SNR) de 0 dB para a fala como definição do SNR mínimo aceitável ao calcular a faixa de captação da fala. Nesse contexto, o ruído pode ser proveniente do ruído ambiental e do ruído elétrico, bem como do ruído próprio do microfone.
- **Nível de ruído do microfone** O nível de ruído digital de um microfone é medido em dBFSrms. Normalmente, os melhores microfones digitais apresentam um ruído de fundo teórico de -105 dBFSrms[A].
- **O ruído acústico equivalente de um microfone** pode ser calculado como o nível de teste de 94 dBspl menos a relação sinal-ruído (SNR) do microfone. Um exemplo comum disso poderia ser 94 – 69 = 25 dBspl[A].

Sobre a Axis Communications

A Axis possibilita um mundo mais inteligente e seguro, aprimorando a segurança, proteção, eficiência operacional e inteligência nos negócios. Como uma empresa de tecnologia em rede e líder do setor, a Axis oferece soluções de videomonitoramento, controle de acesso, intercomunicação e áudio. Essas soluções são aprimoradas por meio de aplicativos de análise inteligentes e apoiadas por treinamentos de alta qualidade.

A Axis conta com cerca de 5.000 funcionários dedicados, em mais de 50 países, e colabora com parceiros de tecnologia e integração de sistemas em todo o mundo para oferecer soluções aos clientes. A Axis foi fundada em 1984 e está sediada em Lund, na Suécia.

© 2026 — Axis Communications AB. AXIS COMMUNICATIONS, AXIS, ARTPEC e VAPIX são marcas comerciais registradas da Axis AB em várias jurisdições. Todas as outras marcas comerciais pertencem aos respectivos proprietários.Reservamo-nos o direito de inserir modificações sem aviso prévio.trademarkdisclaimer