

LIVRE BLANC

# Plage de captation du microphone – facteurs influençant les performances audio

Septembre 2026

# Table des matières

|   |                                                                                        |   |
|---|----------------------------------------------------------------------------------------|---|
| 1 | Introduction                                                                           | 3 |
| 2 | Plage d'enregistrement physique                                                        | 3 |
| 3 | Enregistrements dans différents environnements                                         | 5 |
| 4 | Résultats des mesures issues d'enregistrements réalisés dans différents environnements | 7 |
| 1 | Annexe                                                                                 | 9 |

# 1 Introduction

Il est rarement facile de définir une portée maximale précise pour les dispositifs d'enregistrement audio, car leurs performances dépendent de l'environnement acoustique. Si les spécifications techniques constituent une ligne de base, la plage d'enregistrement réelle dépend d'une interaction complexe entre des facteurs environnementaux et acoustiques qui varient d'une installation à l'autre.

Ce livre blanc examine les principaux facteurs qui déterminent la portée maximale à laquelle un microphone peut capter le son de manière fiable, avec une focalisation sur l'intelligibilité de la parole comme critère de référence principal. Parmi ces facteurs figurent le niveau de bruit de fond, les caractéristiques spectrales de ce bruit, le temps de réverbération, la directivité du microphone, ainsi que le bruit propre du microphone et de l'appareil d'enregistrement.

Pour établir une ligne de base pratique, nous utilisons la voix humaine comme source sonore de référence. Le bruit est mesuré à l'aide d'un filtre de pondération A afin d'obtenir une meilleure corrélation entre la valeur mesurée et le niveau de bruit perçu. Le rapport signal/bruit (SNR) de l'enregistrement de la voix est calculé comme étant la différence, exprimée en décibels, entre le niveau de la voix et le niveau de bruit pondéré A (*voir l'annexe pour plus d'informations sur le rapport signal/bruit (SNR)*). Des tests empiriques montrent que l'intelligibilité de la parole est acceptée jusqu'à un rapport signal/bruit (SNR) d'environ 0 dB (c'est-à-dire lorsque le bruit et le son ont le même niveau), à condition que le bruit présente une réponse en fréquence relativement plate dans la gamme des fréquences de la parole.

En comprenant ces variables, les installateurs et les concepteurs de systèmes peuvent prendre des décisions plus éclairées lors de la sélection et du positionnement des dispositifs audio dans des situations réelles.

## 2 Plage d'enregistrement physique

Les performances d'enregistrement d'un microphone dépendent généralement du bruit de fond ambiant. Les principaux facteurs qui limitent la portée d'enregistrement physique sont les suivants :

- **Niveau de pression sonore** : tous les sons ne se propagent pas sur la même distance. Par exemple, un coup de feu génère un niveau de pression sonore extrêmement élevé et peut être entendu jusqu'à un kilomètre de distance. Une voix humaine est toutefois plus faible et s'estompe plus rapidement avec la distance. Pour estimer une portée de captation réaliste, vous devez d'abord définir le type de son que vous souhaitez capturer, afin que le niveau de la source puisse déterminer la limite supérieure de ce qu'il est possible d'enregistrer.
- **Bruit de fond** : plus l'environnement est bruyant, plus il est difficile pour un microphone d'isoler et de capter le son souhaité. Le bruit de fond varie d'un environnement à l'autre et entre directement en concurrence avec le son recherché. Cela comprend notamment le bruit du vent, de la circulation, des machines et de la foule. Plus le bruit de fond est élevé au lieu où se trouve le microphone, plus la portée de captation effective est réduite. Cela explique pourquoi un même microphone peut fonctionner de manière très différente dans un couloir calme par rapport à un carrefour très fréquenté.
- **Profil spectral du bruit** : tout comme l'intensité sonore du bruit et les fréquences que ce dernier couvre ont leur importance. La parole humaine se situe principalement dans la gamme des fréquences moyennes, soit approximativement entre 300 Hz et 3 400 Hz. Si le bruit de fond se concentre sur des fréquences très basses ou très hautes, il est moins susceptible de nuire à l'intelligibilité de la parole. Cependant, si le bruit se superpose à la même gamme de fréquences que le son cible, il devient difficile de distinguer la parole du bruit, même à des niveaux de bruit modérés.
- **Temps de réverbération au lieu du microphone** : la réverbération est la persistance du son après que la source initiale a cessé d'émettre, due aux réflexions sur les murs, les plafonds, les sols et d'autres surfaces. Un temps de réverbération prolongé peut nuire à l'intelligibilité de la parole, même dans des situations où le niveau sonore global et le rapport signal/bruit (SNR) semblent acceptables. Dans les espaces à forte réverbération, tels que les grands halls vides ou les pièces carrelées, les sons de la parole parviennent au microphone accompagnés d'une quantité infinie de réflexions ambiantes, ce qui rend le son flou. Il est donc difficile, tant pour les auditeurs que pour les systèmes d'enregistrement, de distinguer clairement chaque mot.

- **Directivité de la configuration des microphones** : la sensibilité des microphones aux sons provenant de directions différentes varie d'un modèle à l'autre. Un microphone omnidirectionnel capte le son de manière égale dans toutes les directions, ce qui signifie également qu'il capte les bruits de fond et la réverbération provenant de toutes les directions. Une configuration directionnelle, telle que plusieurs microphones configurés pour la formation de faisceaux, effectue une mise au point sur une direction spécifique, ce qui réduit efficacement l'influence du bruit et de la réverbération provenant d'autres angles. Cela peut étendre considérablement la portée de captation utile dans des environnements bruyants ou réverbérants, mais uniquement dans la direction vers laquelle est orienté le réseau de microphones.
- **Bruit propre des microphones et des dispositifs d'enregistrement audio** : chaque microphone et chaque dispositif d'enregistrement audio génère une faible quantité de bruit électrique interne, appelé « bruit propre ». Cela fixe une limite inférieure absolue à ce que le système est capable de capturer. Même dans un environnement très calme, les sons dont le niveau est inférieur au seuil de bruit propre ne peuvent tout simplement pas être enregistrés proprement. Lorsqu'on compare des microphones, une valeur de bruit propre plus faible indique généralement que le microphone est capable de capter plus efficacement les sons les plus faibles, ce qui peut se traduire par une portée de captation plus étendue si le niveau de bruit et le temps de réverbération sont suffisamment faibles.

Pour simplifier l'analyse, nous pouvons partir d'un scénario idéalisé dans lequel il n'y a ni bruit de fond, ni réflexions, ni réverbération, et prendre la voix humaine comme source sonore de référence. Dans ces conditions, la question essentielle qui se pose est la suivante : quel rapport signal/bruit (SNR) faut-il utiliser comme seuil pour une plage de détection acceptable ?

Les tests empiriques nous apportent une réponse claire. Lorsque l'on applique un filtre de pondération A au bruit du système et que l'on compare le niveau RMS (valeur efficace) moyen du bruit à celui de la parole, l'intelligibilité de la parole ne commence à diminuer que lorsque le rapport signal/bruit (SNR) descend sous 0 dB. Nous utilisons donc un rapport signal/bruit (SNR) de 0 dB comme critère limite pour l'enregistrement d'une voix humaine. Cela est vrai tant que le bruit présente une réponse en fréquence relativement uniforme, similaire à celle du bruit blanc, et qu'il se situe dans la gamme des fréquences vocales, ce qui est généralement le cas pour le bruit propre d'un microphone.

Par exemple, un microphone numérique peut présenter une sensibilité de -36 dBFSrms (*voir l'annexe pour le dBFSrms*) lorsqu'une tonalité de 1 kHz est émise à 94 dBspl (une sinusoïde numérisée à pleine échelle est ici définie comme étant de 0 dBFSrms, conformément aux fiches techniques des microphones) (*voir l'annexe pour le dBspl*). Cela s'applique à de nombreux produits Axis équipés de microphones MEMS numériques intégrés lorsque le gain du microphone est réglé sur 0 dB. Lorsque vous placez ce microphone dans un environnement très calme où le niveau de bruit de fond descend sous 10 dBspl(A) (*voir l'annexe pour le dBspl(A)*), le bruit interne du microphone devient prédominant. Dans cet exemple, la fiche technique du microphone indique une sortie numérique de -105 dBFSrms(A). Le rapport signal/bruit (SNR) du microphone correspond à la différence entre le niveau de référence et le bruit de fond :  $-36 - (-105) = 69$  dB.

Nous convertissons ce niveau de bruit électrique de fond en un niveau de bruit acoustique équivalent :  $94 - 69 = 25$  dBspl(A).

Nous allons maintenant calculer la distance maximale à laquelle il est possible d'effectuer un enregistrement de la voix humaine avec un rapport signal/bruit (SNR) de 0 dB. Selon les normes de l'UIT, la voix humaine présente un niveau moyen nominal de parole (ASL) d'environ 57 dBspl à 1 m. Dans un espace dégagé et en l'absence d'objets réfléchissants, le niveau de pression sonore diminue avec la distance selon la loi du carré inverse :

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

.

Si la distance de référence est de 1 m, pour atteindre un rapport signal/bruit (SNR) de 0 dB au niveau de ce microphone, le niveau de la parole doit descendre à 25 dBspl(A) ; soit une réduction de 32 dB par rapport au niveau de référence de 1 m. En réorganisant la formule, on obtient :

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

Cela signifie que dans un espace dégagé et en l'absence de bruit de fond, de vent et de réverbération, ce microphone peut théoriquement enregistrer une parole intelligible à environ 40 m. Il s'agit de la portée maximale théorique de captation.

Cependant, les facteurs environnementaux ont une incidence plus importante sur le niveau de bruit que le bruit de fond du microphone interne. Un bureau calme après les heures de travail, et avec des systèmes de ventilation qui fonctionnent encore, présente un niveau de bruit de fond compris entre 28 et 32 dBspl environ. Ce seul facteur peut réduire la portée effective de captation d'environ 50 %, ramenant ainsi la portée maximale à environ 20 m. Si l'on effectue un enregistrement en extérieur, dans un milieu urbain, on sera confronté à un mélange de bruits : le bruit du vent, le bruit de la circulation, les bruits liés à l'activité humaine, etc. Avec un microphone omnidirectionnel, chaque niveau de bruit et chaque spectre de bruit ont leur propre plage de captation correspondante. La portée de captation peut alors aller de 2 m à 4 m en cas de bruit de circulation important.

Un aspect pratique à prendre en compte est que les enregistrements de voix humaine réalisés au seuil de 0 dB SNR ou à proximité de celui-ci nécessitent un gain numérique important pour être audibles, car le niveau sonore de la voix enregistrée est si faible qu'il descendrait sinon en dessous de notre propre seuil d'audibilité lors de la relecture. Dans les exemples d'enregistrement qui suivent, les fichiers audio ont été amplifiés de +40 dB pour compenser le niveau distant de la voix.

### 3 Enregistrements dans différents environnements

Cette section présente en détail les lieux d'enregistrement, accompagnés de captures d'écran tirées des enregistrements vidéo et d'extraits audio captés à différentes distances.

**Exemple 1 :** le lieu se trouve à l'extérieur et est exposé au bruit de la circulation provenant d'une autoroute. Le niveau de bruit de fond varie entre 62 et 68 dBspl[A] et un léger bruit de vent est aussi présent.



Figure 3.1 Le lieu de l'enregistrement se trouve à proximité d'une autoroute (voir le point rouge).



Figure 3.2 Sur le lieu de l'enregistrement, la distance maximale d'enregistrement est de 4 m par rapport à une caméra-piéton Axis.

*Il est préférable d'utiliser un casque pour écouter l'enregistrement.*

**Exemple 2 :** un couloir silencieux dont le niveau de bruit de fond est de 32 dBspl[A]. Ici, le couloir agit comme un long guide d'ondes acoustiques. Le son se propage dans le couloir en rebondissant sur les surfaces réfléchissantes et parvient aux microphones avec un volume plus élevé qu'il ne l'aurait fait dans un espace libre, à la même distance.



Figure 3.3 La distance maximale d'enregistrement est de 27 m.

*Il est préférable d'utiliser un casque pour écouter l'enregistrement.*

**Exemple 3 :** un grand hall d'entrée présentant un niveau de bruit de fond de 42 dBspl[A]. Ici, les murs sont équipés de matériaux acoustiques absorbants placés derrière les panneaux de diffusion, ce qui permet d'éliminer la plupart des réflexions précoces. À une plus grande distance, le son se propage dans le canal acoustique formé par le sol et le plafond réfléchissant.

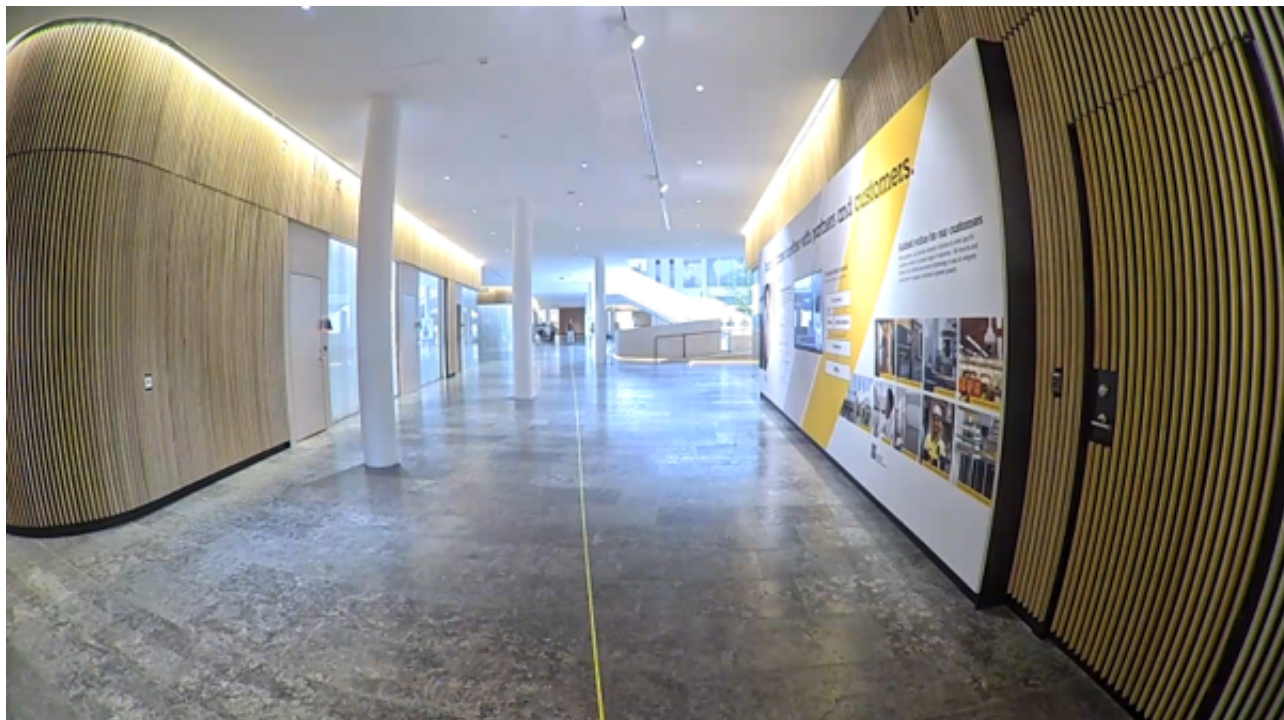


Figure 3.4 À une distance d'enregistrement maximale de 40 m, l'orateur est à peine visible, mais on l'entend tout de même.

*Il est préférable d'utiliser un casque pour écouter l'enregistrement.*

## 4 Résultats des mesures issues d'enregistrements réalisés dans différents environnements

Cette analyse porte sur la captation de la voix. Les sons plus puissants, comme les coups de feu, seront audibles de plus loin. Nous avons analysé et mesuré chaque phrase issue des enregistrements audio, puis estimé le rapport signal/bruit (SNR) pour chacune d'entre elles.

Le rapport signal/bruit (SNR) a été estimé comme suit : parole + bruit - bruit de fond. Une méthode plus appropriée consisterait à calculer le niveau de la parole - le niveau de bruit, mais il n'est pas possible de mesurer le niveau de la parole sans tenir compte du bruit. Cela signifie que lorsque le niveau de la parole est proche du niveau de bruit, la précision du rapport signal/bruit (SNR) diminue.

| Lieu et bruit de fond : | À l'extérieur, avec un bruit de circulation de 61 à 68 dBspl[A] | Dans un couloir silencieux 32 dBspl[A]   | Dans un grand hall d'entrée 42 dBspl[A]  | Espace dégagé, pas de réverbérations, silence total           |
|-------------------------|-----------------------------------------------------------------|------------------------------------------|------------------------------------------|---------------------------------------------------------------|
| Distance (m)            | Rapport signal/bruit (SNR) [dB] (estimé)                        | Rapport signal/bruit (SNR) [dB] (estimé) | Rapport signal/bruit (SNR) ([dB] estimé) | Rapport signal/bruit (SNR) théorique = bruit du microphone au |

|    |             |                    |      | niveau de la parole [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 | Non audible | 17,5               | 6,3  | 7,9                      |
| 24 | Non audible | 14,8               | 5,4  | 4,4                      |
| 32 | Non audible | Couloir trop court | 4,6  | 1,9                      |
| 40 | Non audible | Couloir trop court | 4,7  | 0,0                      |

En théorie, le niveau sonore de la parole devrait diminuer de -6 dB lorsque la distance double. Cela signifierait également que le rapport signal/bruit (SNR) devrait diminuer en conséquence. Dans ce tableau, on constate que le rapport signal/bruit (SNR) est meilleur que prévu dans le couloir et dans le hall d'entrée. Cela s'explique par le fait que les réverbérations dans la pièce ont fait augmenter les niveaux de pression sonore. Dans un couloir en particulier, les murs, le sol et le plafond agissent comme un guide d'ondes pour le son. Cela se traduit par un niveau sonore et un rapport signal/bruit (SNR) supérieurs d'environ 10 dB au rapport signal/bruit théorique dans un espace dégagé, lors d'un test effectué à une distance de 24 m, dans un couloir.

Chaque fois que le bruit de fond dépasse le bruit propre acoustique équivalent d'un microphone omnidirectionnel isolé, la portée de captation dépend entièrement du bruit de fond ambiant plutôt que du microphone lui-même.

## Annexe 1 Annexe

- **Niveau de pression sonore en décibels (dBspl)** : niveau de pression sonore exprimé en décibels (dB), par rapport à une valeur de référence de 20 µPa. Lors de la mesure des niveaux de bruit et de voix, on utilise souvent des niveaux RMS intégrés dans le temps. On parle parfois de « niveau de puissance ».

$$RMS = \sqrt{\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} [f(t)]^2 dt}$$

Les deux durées d'intégration les plus courantes sont le « RMS rapide », avec une fenêtre temporelle de 125 ms, et le « RMS lent », avec une fenêtre temporelle de 1 seconde. Pour l'analyse du bruit de fond, on calcule souvent la moyenne du niveau sur une période plus longue.

- **Niveau de pression sonore pondéré par le filtre de pondération A (dBspl[A])** : cette valeur est parfois exprimée en dB(A). Il s'agit de la valeur pondérée A du niveau en dBspl. La pondération A est une méthode de filtrage du signal du microphone qui permet d'obtenir une meilleure correspondance entre l'audition humaine et le niveau de pression sonore mesuré. En effet, l'audition humaine présente une sensibilité réduite aux basses fréquences inférieures à 500 Hz et aux hautes fréquences supérieures à 10 kHz.

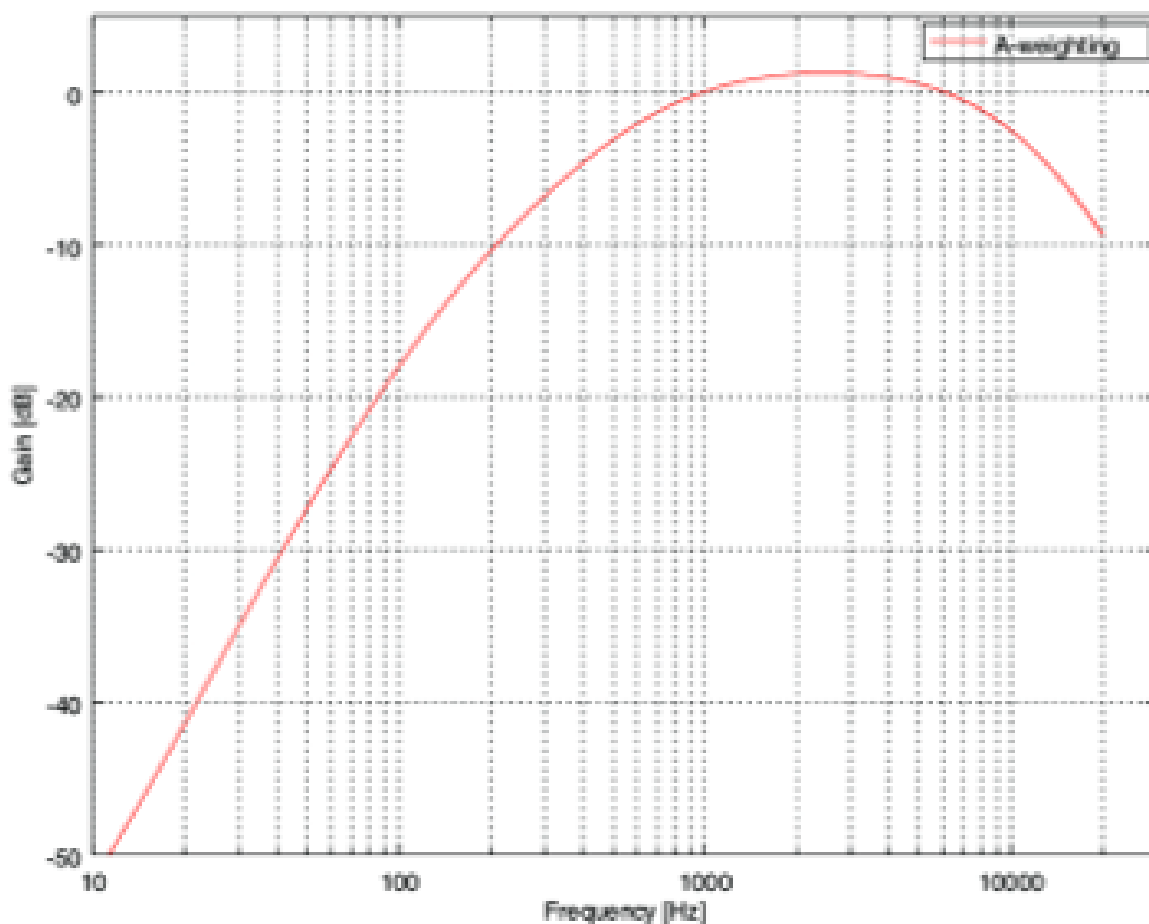


Figure .1 Filtre de pondération A.

- **Le pic en décibels à pleine échelle (dBFSpeak/dBFSp)** désigne l'amplitude d'un signal audio numérique par rapport à la représentation numérique maximale disponible (pleine échelle). Un signal sinusoïdal d'amplitude numérique maximale, sans écrêtage, présente une valeur crête de 0 dBFSp.

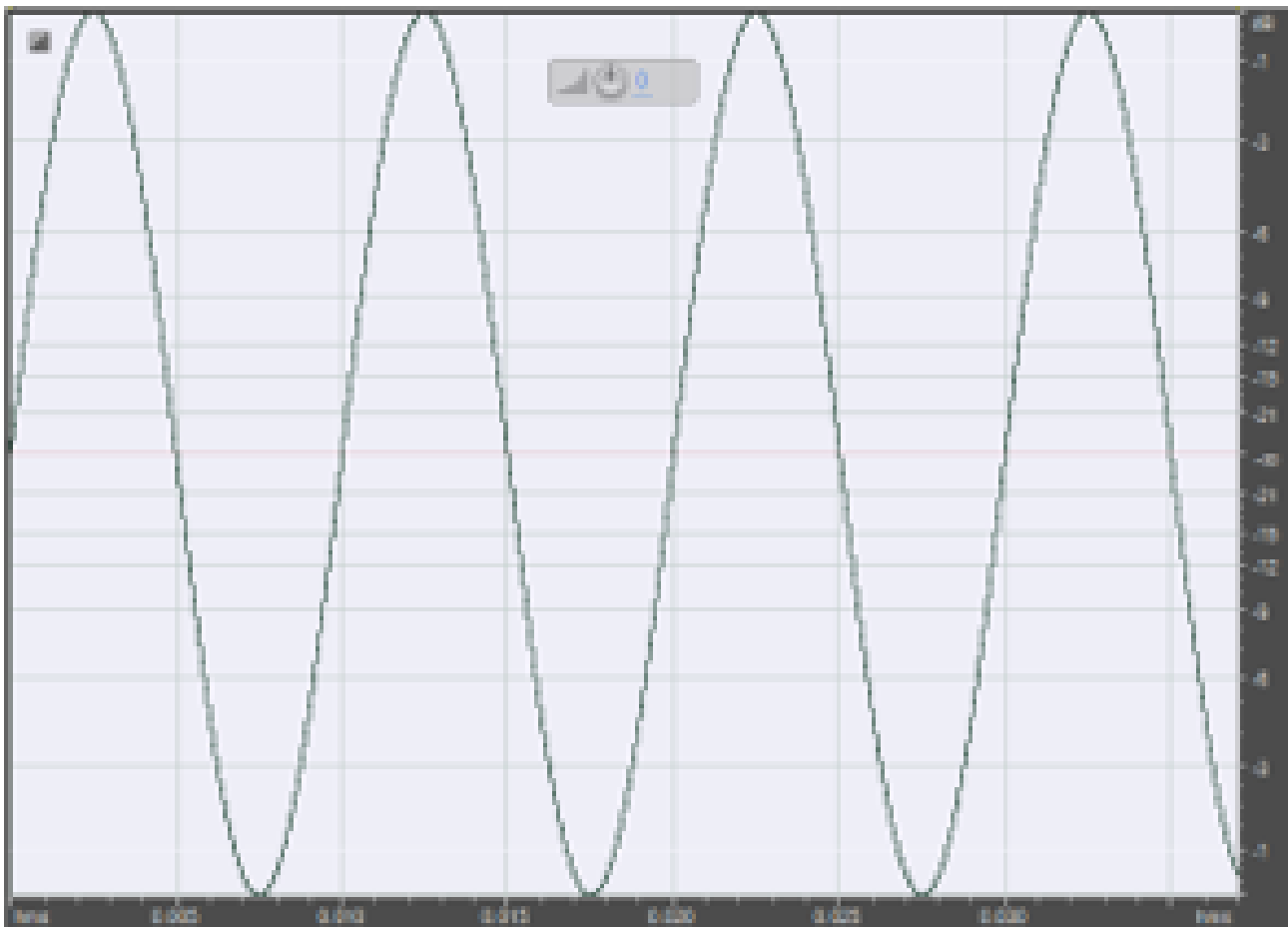


Figure .2 Exemple de signal dont l'amplitude est de 0 dBFSpeak.

- **dBFSrms** désigne un niveau de puissance audio numérique d'un signal numérique. Une onde sinusoïdale à pleine échelle avec 0 dBFSpeak peut présenter différents niveaux efficaces, selon la manière dont l'unité est définie. Dans les fiches techniques des microphones, un signal sinusoïdal à pleine échelle est défini comme étant égal à 0 dBFSrms ; c'est aussi la raison pour laquelle nous utilisons cette définition dans le présent livre blanc. D'une manière générale, l'autre définition d'un signal sinusoïdal à pleine échelle, à savoir -3 dBFSrms, est mathématiquement plus correcte et plus cohérente avec la manière dont les niveaux RMS et les niveaux de crête sont définis dans le domaine analogique. Selon cette autre définition, une onde carrée à pleine échelle correspond à 0 dBFSrms.
- **La sensibilité d'un microphone** correspond au niveau de sortie de celui-ci lorsque le signal d'entrée acoustique est une tonalité sinusoïdale de 1 kHz avec une pression acoustique de 1 Pa, 0 dB a, ou 94 dBspl (trois façons différentes d'exprimer le même niveau de pression sonore).
- **Rapport signal/bruit (SNR)** : lorsque les niveaux audio sont exprimés en dB, ce rapport correspond à la différence entre le niveau du signal et le niveau de bruit.
- **Le rapport signal/bruit (SNR) d'un microphone** est calculé comme étant la différence de niveau de sortie entre une tonalité de test de 1 kHz à 94 dBspl et le bruit propre du microphone, lorsque celui-ci est placé dans un silence absolu. Le rapport signal/bruit (SNR) typique d'un microphone MEMS numérique est de 69 dBspl[A].
- **Dans le présent livre blanc**, le rapport signal/bruit (SNR) de la parole est une comparaison entre la parole humaine et le niveau de bruit pondéré A en l'absence de parole. Le rapport signal/bruit (SNR) est mesuré en décibels. Nous utilisons un rapport signal/bruit (SNR) de 0 dB pour la parole comme seuil minimal de rapport signal/bruit (SNR) acceptable lors du calcul de la portée théorique de captation de la parole. Dans ce contexte, le bruit peut provenir du bruit ambiant et du bruit électrique, ainsi que du bruit propre du microphone.

- **Niveau de bruit d'un microphone** : le niveau de bruit numérique d'un microphone est mesuré en dBFSrms. En règle générale, les meilleurs microphones numériques présentent un bruit de fond théorique de -105 dBFSrms[A].
- **Le bruit équivalent acoustique d'un microphone** peut être calculé comme suit : niveau d'essai de 94 dBspl – rapport signal/bruit (SNR) du microphone. Un exemple courant pourrait être  $94 - 69 = 25$  dBspl[A].

## À propos d'Axis Communications

En améliorant la sûreté, la sécurité, l'efficacité opérationnelle et l'intelligence économique, Axis contribue à un monde plus sûr et plus intelligent. Leader de son secteur dans les technologies sur IP, Axis propose des solutions en vidéosurveillance, contrôle d'accès, visiophonie et systèmes audio. Ces solutions sont enrichies par des applications d'analyse intelligente et soutenues par des formations de haute qualité.

L'entreprise emploie environ 5000 personnes dans plus de 50 pays et collabore avec des partenaires technologiques et intégrateurs de systèmes du monde entier pour fournir des solutions sur mesure à ses clients. Axis a été fondée en 1984, son siège est situé à Lund en Suède.  
aboutaxis\_text2