

## 24/192 Téléchargements de musique

... et pourquoi ils n'ont aucun sens



*Voir également la nouvelle vidéo de Xiph.Org, [Digital Show & Tell](#), pour des démonstrations détaillées de l'échantillonnage numérique en action sur de l'équipement réel!*

Des articles le mois dernier ont révélé que le musicien Neil Young et Steve Jobs d'Apple ont discuté de la possibilité de télécharger de la musique numérique d'une «qualité studio sans compromis». Une grande partie de la presse et des commentaires des utilisateurs étaient particulièrement enthousiastes à l'idée de téléchargements 24 bits 192 kHz non compressés. 24/192 a figuré en bonne place dans mes propres conversations avec le groupe de M. Young il y a plusieurs mois.

Malheureusement, il est inutile de distribuer de la musique au format 24 bits / 192 kHz. Sa fidélité de lecture est légèrement inférieure à 16 / 44,1 ou 16/48, et il occupe 6 fois l'espace.

Il y a quelques problèmes réels avec la qualité audio et «l'expérience» de la musique distribuée numériquement aujourd'hui. 24/192 ne résout aucun d'entre eux. Alors que tout le monde se fixe le 24/192 comme une balle magique, nous n'allons pas voir d'amélioration réelle.

### **Tout d'abord, les mauvaises nouvelles**

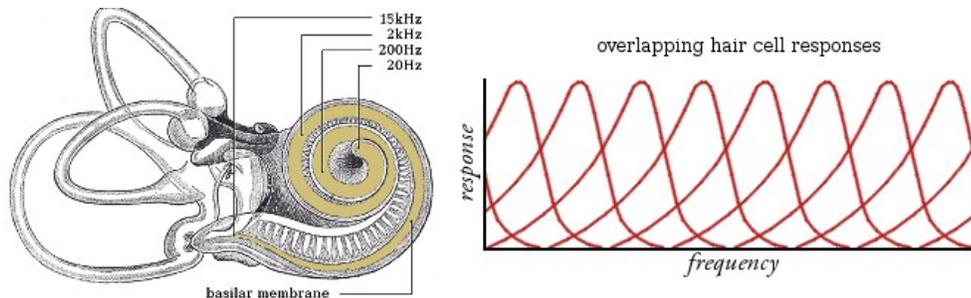
Au cours des dernières semaines, j'ai eu des conversations avec des personnes intelligentes et à l'esprit scientifique qui croient aux téléchargements 24/192 et veulent savoir comment quiconque pourrait éventuellement être en désaccord. Ils ont posé de bonnes questions qui méritent des réponses détaillées.

J'étais également intéressé par ce qui a motivé le plaidoyer audio numérique à haut débit. Les réponses indiquent que peu de gens comprennent la théorie de base du signal ou le [théorème d'échantillonnage](#), ce qui n'est guère surprenant. Des malentendus sur les mathématiques, la technologie et la physiologie ont surgi dans la plupart des conversations, souvent affirmées par des professionnels qui possédaient autrement une expertise audio importante. Certains ont même soutenu que le théorème d'échantillonnage n'explique pas vraiment comment fonctionne l'audio numérique [ [1](#) ].

La désinformation et la superstition ne servent que les charlatans. Alors, couvrons certaines des bases de pourquoi la distribution 24/192 n'a aucun sens avant de suggérer des améliorations qui le font réellement.

## Messieurs, rencontrez vos oreilles

L'oreille entend via les cellules ciliées qui reposent sur la membrane basilaire résonante de la cochlée. Chaque cellule ciliée est effectivement réglée sur une bande de fréquence étroite déterminée par sa position sur la membrane. La sensibilité culmine au milieu de la bande et tombe de chaque côté en forme de cône déséquilibré chevauchant les bandes d'autres cellules ciliées à proximité. Un son est inaudible s'il n'y a pas de cellules ciliées réglées pour l'entendre.



*En haut à gauche: dessin anatomique en coupe d'une cochlée humaine avec la membrane basilaire colorée en beige. La membrane est réglée pour résonner à différentes fréquences sur toute sa longueur, avec des fréquences plus élevées près de la base et des fréquences plus basses à l'apex. Les emplacements approximatifs de plusieurs fréquences sont indiqués.*

*En haut à droite: diagramme schématisé représentant la réponse des cellules ciliées le long de la membrane basilaire sous la forme d'une banque de filtres qui se chevauchent.*

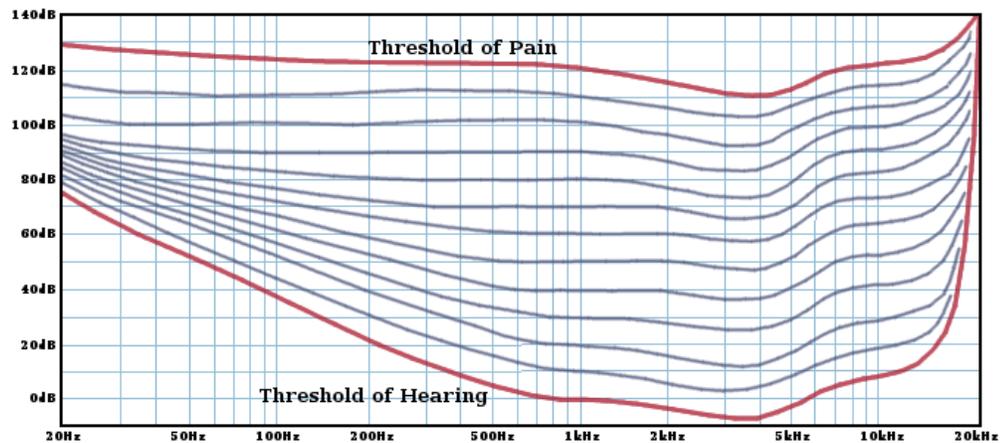
Ceci est similaire à une radio analogique qui capte la fréquence d'une station puissante près de l'endroit où le tuner est réellement réglé. Plus la fréquence de la station est éloignée, plus elle est faible et déformée jusqu'à ce qu'elle disparaisse complètement, quelle que soit sa force. Il existe une limite de fréquence audible supérieure (et inférieure), au-delà de laquelle la sensibilité des dernières cellules ciliées tombe à zéro et l'audition se termine.

## Taux d'échantillonnage et spectre audible

Je suis sûr que vous avez entendu cela de très nombreuses fois: la plage d'audition humaine s'étend de 20 Hz à 20 kHz. Il est important de savoir comment les chercheurs parviennent à ces chiffres spécifiques.

Premièrement, nous mesurons le «seuil absolu d'audition» sur toute la plage audio pour un groupe d'auditeurs. Cela nous donne une courbe représentant le son le plus silencieux que l'oreille humaine puisse percevoir pour une fréquence donnée mesurée dans des circonstances idéales sur des oreilles saines. Un environnement anéchoïque, un équipement de lecture calibré avec précision et une analyse statistique rigoureuse sont la partie la plus facile. Les oreilles et la concentration auditive se fatiguent toutes les deux rapidement, donc les tests doivent être effectués lorsqu'un auditeur est frais. Cela signifie beaucoup de pauses et de pauses. Les tests prennent de plusieurs heures à plusieurs jours selon la méthodologie.

Ensuite, nous collectons des données pour l'extrême opposé, le «seuil de douleur». C'est le point où l'amplitude audio est si élevée que le matériel physique et neuronal de l'oreille est non seulement complètement submergé par l'entrée, mais éprouve une douleur physique. La collecte de ces données est plus délicate. Vous ne voulez pas endommager définitivement l'audition de qui que ce soit au cours du processus.



*Ci-dessus: courbes approximatives de volume égal dérivées de Fletcher et Munson (1933) plus des sources modernes pour des fréquences > 16 kHz. Les courbes de seuil absolu d'audition et de seuil de douleur sont marquées en rouge. Des chercheurs ultérieurs ont affiné ces lectures, aboutissant à l'échelle Phon et aux courbes d'égalité d'intensité standard ISO 226. Les données modernes indiquent que l'oreille est nettement moins sensible aux basses fréquences que les résultats de Fletcher et Munson.*

La limite supérieure de la plage audio humaine est définie comme étant l'endroit où le seuil absolu de la courbe auditive franchit le seuil de la douleur. Pour percevoir même faiblement l'audio à ce point (ou au-delà), il doit simultanément être insupportablement fort.

Aux basses fréquences, la cochlée fonctionne comme une enceinte bass-reflex. L' *hélicotrème* est une ouverture au sommet de la membrane basilaire qui agit comme un port réglé entre 40 Hz et 65 Hz selon les individus. La réponse diminue fortement en dessous de cette fréquence.

Ainsi, 20 Hz - 20 kHz est une plage généreuse. Il couvre à fond le spectre audible, une affirmation soutenue par près d'un siècle de données expérimentales.

## Cadeaux génétiques et oreilles d'or

Sur la base de mes correspondances, de nombreuses personnes croient en des individus dotés de dons auditifs extraordinaires. Ces «oreilles d'or» existent-elles vraiment?

Cela dépend de ce que vous appelez une oreille dorée.

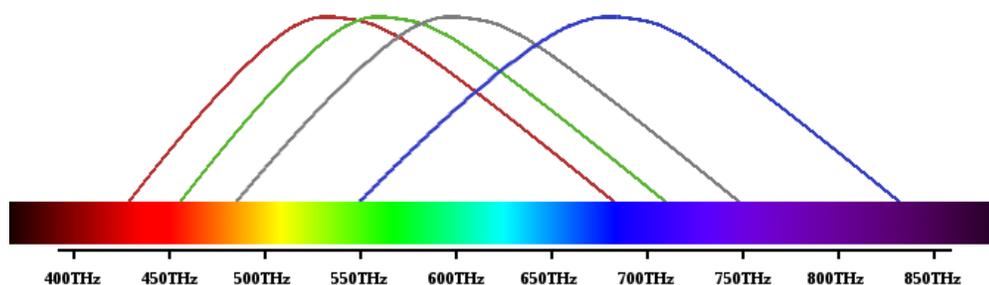
Les oreilles jeunes et saines entendent mieux que les oreilles vieilles ou endommagées. Certaines personnes sont exceptionnellement bien entraînées pour entendre les nuances sonores et musicales que la plupart des gens ignorent même exister. Il fut un temps dans les années 1990 où je pouvais identifier chaque codeur mp3 majeur par son (à l'époque où ils étaient tous assez mauvais), et je pouvais le démontrer de manière fiable dans des tests en double aveugle [ [2](#) ].

Lorsque des oreilles saines se combinent avec des capacités de discrimination hautement qualifiées, j'appellerais cette personne une oreille d'or. Même ainsi, une audition inférieure à la moyenne peut également être formée pour remarquer les détails qui échappent aux auditeurs non formés. Les oreilles dorées sont plus une question d'entraînement que d'entendre au-delà des capacités physiques des mortels moyens.

Les chercheurs en audit auditif aimeraient trouver, tester et documenter des personnes ayant une audition vraiment exceptionnelle, comme une gamme d'audition considérablement étendue. Les gens normaux sont gentils et tout, mais tout le monde veut trouver un phénomène génétique pour un papier vraiment juteux. Nous n'avons pas trouvé de telles personnes au cours des 100 dernières années de tests, donc elles n'existent probablement pas. Désolé. Nous continuerons à chercher.

## Spectrophiles

Vous êtes peut-être sceptique à propos de tout ce que je viens d'écrire; cela va certainement à l'encontre de la plupart des supports marketing. Au lieu de cela, considérons un engouement hypothétique pour la vidéo à large spectre qui ne transporte pas de bagages audiophiles préexistants.



*Ci-dessus: La réponse approximative à l'échelle logarithmique des bâtonnets et des cônes de l'œil humain, superposée au spectre visible. Ces organes sensoriels répondent à la lumière dans des bandes spectrales qui se chevauchent, tout comme les cellules ciliées de l'oreille sont réglées pour répondre à des bandes de fréquences sonores qui se chevauchent.*

L'œil humain voit une gamme limitée de fréquences de lumière, alias le spectre visible. Ceci est directement analogue au spectre audible des ondes sonores. Comme l'oreille, l'œil possède des cellules sensorielles (bâtonnets et cônes) qui détectent la lumière dans des bandes de fréquences différentes mais qui se chevauchent.

Le spectre visible s'étend d'environ 400THz (rouge foncé) à 850THz (violet foncé) [ 3 ]. La perception tombe fortement sur les bords. Au-delà de ces limites approximatives, la puissance lumineuse nécessaire à la moindre perception peut faire friser vos rétines. Ainsi, il s'agit d'une durée généreuse, même pour les personnes jeunes, saines et génétiquement douées, analogue aux limites généreuses du spectre audible.

Dans notre engouement hypothétique pour la vidéo à large spectre, considérons un groupe fervent de spectrophiles qui croient que ces limites ne sont pas assez généreuses. Ils proposent que la vidéo représente non seulement le spectre visible, mais aussi l'infrarouge et l'ultraviolet. Poursuivant la comparaison, il y a une faction encore plus hardcore [et fière d'elle!] Qui insiste sur le fait que cette gamme étendue est encore insuffisante, et que la vidéo semble tellement plus

naturelle lorsqu'elle comprend également les micro-ondes et une partie du spectre des rayons X. Pour un œil d'or, insistent-ils, la différence est le jour et la nuit!

Bien sûr, c'est ridicule.

Personne ne peut voir les rayons X (ou infrarouges, ultraviolets ou micro-ondes). Peu importe combien une personne croit pouvoir le faire. Les rétines n'ont tout simplement pas le matériel sensoriel.

Voici une expérience que tout le monde peut faire: allez chercher votre télécommande infrarouge Apple. La LED émet à 980 nm, soit environ 306 THz, dans le spectre proche infrarouge. Ce n'est pas loin en dehors de la plage visible. Emmenez la télécommande au sous-sol, ou dans la pièce la plus sombre de votre maison, au milieu de la nuit, les lumières éteintes. Laissez vos yeux s'adapter à la noirceur.



*Ci-dessus: télécommande infrarouge Apple photographiée à l'aide d'un appareil photo numérique. Bien que l'émetteur soit assez brillant et que la fréquence émise ne soit pas loin de la partie rouge du spectre visible, il est complètement invisible à l'œil.*

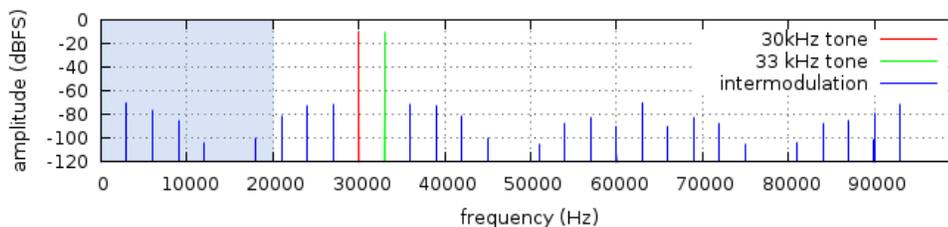
Pouvez-vous voir le flash LED de la télécommande Apple Remote lorsque vous appuyez sur un bouton [ 4 ]? Non? Pas même le moindre montant? Essayez quelques autres télécommandes IR; beaucoup utilisent une longueur d'onde IR un peu plus proche de la bande visible, autour de 310-350THz. Vous ne pourrez pas non plus les voir. Le reste émet juste au bord de la visibilité de 350-380 THz et peut être à peine visible dans une noirceur complète avec des yeux ajustés à l'obscurité [ 5 ]. Tous seraient aveuglément, douloureusement brillants s'ils étaient bien à l'intérieur du spectre visible.

Ces LED proche IR émettent de la limite visible jusqu'à 20% au plus au-delà de la limite de fréquence visible. L'audio 192 kHz s'étend jusqu'à 400% de la limite audible. De peur d'être accusé de comparer des pommes et des oranges, la perception auditive et visuelle chutent de la même manière vers les bords.

### **192 kHz considéré comme dangereux**

Les fichiers musicaux numériques à 192 kHz n'offrent aucun avantage. Ils ne sont pas tout à fait neutres non plus; la fidélité pratique est légèrement pire. Les ultrasons sont un handicap lors de la lecture.

Ni les transducteurs audio ni les amplificateurs de puissance ne sont exempts de distorsion, et la distorsion a tendance à augmenter rapidement aux fréquences les plus basses et les plus hautes. Si le même transducteur reproduit les ultrasons avec le contenu audible, toute non-linéarité déplacera une partie du contenu ultrasonore dans la gamme audible comme une pulvérisation incontrôlée de produits de distorsion d'intermodulation couvrant tout le spectre audible. La non-linéarité dans un amplificateur de puissance produira le même effet. L'effet est très faible, mais des tests d'écoute ont confirmé que les deux effets peuvent être audibles.



*Ci-dessus: Illustration des produits de distorsion résultant de l'intermodulation d'une tonalité de 30 kHz et 33 kHz dans un amplificateur théorique avec une distorsion harmonique totale (THD) non variable d'environ 0,09%. Des produits de distorsion apparaissent dans tout le spectre, y compris à des fréquences inférieures à l'une ou l'autre tonalité.*

*Les ultrasons inaudibles contribuent à la distorsion d'intermodulation dans la gamme audible (zone bleu clair). Les systèmes non conçus pour reproduire les ultrasons ont généralement des niveaux de distorsion bien plus élevés au-dessus de 20 kHz, ce qui contribue encore à l'intermodulation. L'élargissement de la plage de fréquences d'une conception pour tenir compte des ultrasons nécessite des compromis qui réduisent les performances de bruit et de distorsion dans le spectre audible. Dans tous les cas, la reproduction inutile du contenu ultrasonore diminue les performances.*

Il existe plusieurs façons d'éviter la distorsion supplémentaire:

1. Un haut-parleur, un amplificateur et un étage de transition dédiés aux ultrasons pour séparer et reproduire indépendamment les ultrasons que vous n'entendez pas, juste pour qu'ils ne gâchent pas les sons que vous pouvez
2. Amplificateurs et transducteurs conçus pour une reproduction de fréquences plus larges, afin que les ultrasons ne provoquent pas d'intermodulation audible. À coûts et complexité égaux, cette gamme de fréquences supplémentaire doit se faire au prix d'une certaine réduction des performances dans la partie audible du spectre.
3. Haut-parleurs et amplificateurs soigneusement conçus pour ne pas reproduire les ultrasons de toute façon.
4. Ne pas encoder une gamme de fréquences aussi large pour commencer. Vous ne pouvez et n'aurez pas de distorsion d'intermodulation ultrasonique dans la bande audible s'il n'y a pas de contenu ultrasonique.

Ils équivalent tous à la même chose, mais seulement 4) est logique.

Si vous êtes curieux de connaître les performances de votre propre système, les exemples

suivants contiennent une tonalité de 30 kHz et 33 kHz dans un fichier WAV 24/96, une version plus longue dans un fichier FLAC, quelques gazouillis tricolores et un clip de morceau normal décalé jusqu'à 24 kHz afin qu'il soit entièrement dans la plage ultrasonique de 24 kHz à 46 kHz:

- **Tests intermod:**

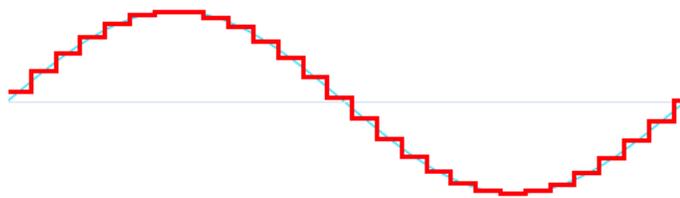
- Tonalité 30 kHz + tonalité 33 kHz (24 bits / 96 kHz) [[WAV 5 secondes](#)] [[FLAC 30 secondes](#)]
- 26kHz - 48kHz tonalités de warbling (24 bits / 96kHz) [[10 secondes WAV](#)]
- 26kHz - 96kHz sons de warbling (24 bits / 192kHz) [[10 secondes WAV](#)]
- Clip de chanson décalé de 24 kHz (WAV 24 bits / 96 kHz) [[WAV 10 secondes](#)] ([version originale du clip ci-dessus](#)) (WAV 16 bits / 44,1 kHz)

En supposant que votre système est réellement capable d'une lecture complète à 96 kHz [ [6](#) ], les fichiers ci-dessus doivent être complètement silencieux, sans bruits, tonalités, sifflets, clics ou autres sons. Si vous entendez quelque chose, votre système présente une non-linéarité provoquant une intermodulation audible des ultrasons. Soyez prudent lorsque vous augmentez le volume; un écrêtage numérique ou analogique, même un écrêtage doux, provoquera soudainement de fortes tonalités d'intermodulation.

En résumé, il n'est pas certain que l'intermodulation par ultrasons soit audible sur un système donné. La distorsion ajoutée pourrait être insignifiante ou elle pourrait être perceptible. Quoi qu'il en soit, le contenu ultrasonique n'est jamais un avantage, et sur de nombreux systèmes, il nuira de manière audible à la fidélité. Sur les systèmes, cela ne fait pas de mal, le coût et la complexité de la manipulation des ultrasons auraient pu être économisés ou dépensés pour améliorer les performances de la gamme audible à la place.

## Échantillonnage des erreurs et des idées fausses

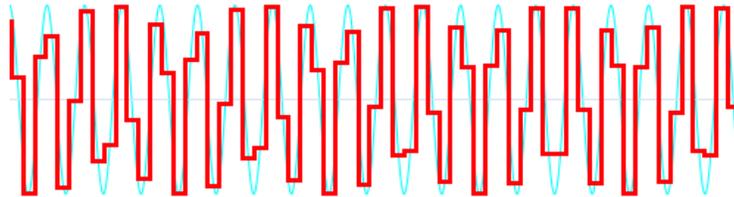
La théorie de l'échantillonnage est souvent peu intuitive sans arrière-plan de traitement du signal. Il n'est pas surprenant que la plupart des gens, même des docteurs brillants dans d'autres domaines, le comprennent régulièrement de manière erronée. Il n'est également pas surprenant que beaucoup de gens ne réalisent même pas qu'ils se trompent.



*Ci-dessus: Les signaux échantillonnés sont souvent représentés comme un escalier brut (rouge) qui semble une mauvaise approximation du signal d'origine. Cependant, la représentation est mathématiquement exacte et le signal retrouve la forme lisse exacte de l'original (bleu) lorsqu'il est reconverti en analogique.*

L'idée fausse la plus courante est que l'échantillonnage est fondamentalement approximatif et avec perte. Un signal échantillonné est souvent décrit comme un fac-similé escalier en escalier aux angles durs de la forme d'onde parfaitement lisse d'origine. Si c'est ainsi que vous envisagez le fonctionnement de l'échantillonnage, vous pouvez penser que plus le taux d'échantillonnage (et plus de bits par échantillon) est rapide, plus la marche d'escalier est fine et plus l'approximation sera proche. Le signal numérique sonnerait de plus en plus près du signal analogique d'origine à mesure que la fréquence d'échantillonnage approche de l'infini.

De même, de nombreuses personnes non DSP examineraient les éléments suivants:



Et dites: "Ugh!" Il peut sembler qu'un signal échantillonné représente mal les formes d'onde analogiques à haute fréquence. Ou, à mesure que la fréquence audio augmente, la qualité échantillonnée diminue et la réponse en fréquence diminue ou devient sensible à la phase d'entrée.

Les regards sont trompeurs. Ces croyances sont incorrectes!

*ajouté le 04/04/2013:*

*Pour faire suite à tout le courrier que j'ai reçu sur les formes d'onde numériques et les marches d'escalier, je démontre le comportement numérique réel sur de l'équipement réel dans notre vidéo [Digital Show & Tell](#), vous n'avez donc pas besoin de me croire sur parole ici!*

Tous les signaux dont le contenu est entièrement inférieur à la fréquence de Nyquist (la moitié du taux d'échantillonnage) sont capturés parfaitement et complètement par échantillonnage; un taux d'échantillonnage infini n'est pas requis. L'échantillonnage n'affecte pas la réponse en fréquence ou la phase. Le signal analogique peut être reconstruit sans perte, en douceur et avec la synchronisation exacte du signal analogique d'origine.

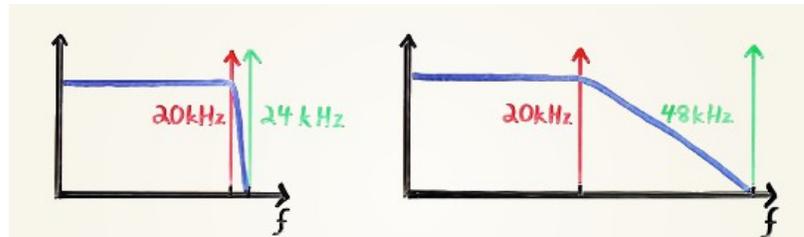
Le calcul est donc idéal, mais qu'en est-il des complications du monde réel? Le plus notoire est l'exigence de limitation de bande. Les signaux dont le contenu dépasse la fréquence de Nyquist doivent être passe-bas avant l'échantillonnage pour éviter une distorsion de repliement; ce passe-bas analogique est le fameux filtre anti-crénelage. L'anticrénelage ne peut pas être idéal dans la pratique, mais les techniques modernes le rapprochent beaucoup. ... et avec cela nous arrivons au suréchantillonnage.

## Suréchantillonnage

Les taux d'échantillonnage supérieurs à 48 kHz ne sont pas pertinents pour les données audio haute fidélité, mais ils sont essentiels en interne à plusieurs techniques audio numériques modernes. *Le suréchantillonnage* est l'exemple le plus pertinent [ [7](#) ].

Le suréchantillonnage est simple et intelligent. Vous vous souviendrez peut-être de mon [A Digital](#)

[Media Primer for Geeks](#) que les taux d'échantillonnage élevés offrent beaucoup plus d'espace entre l'audio à la fréquence la plus élevée qui nous tient à cœur (20 kHz) et la fréquence de Nyquist (la moitié de la fréquence d'échantillonnage). [Cela permet des filtres anti-aliasing analogiques plus simples, plus lisses et plus fiables, et donc une fidélité plus élevée](#). Cet espace supplémentaire entre 20 kHz et la fréquence de Nyquist est essentiellement juste un remplissage spectral pour le filtre analogique.



*Ci-dessus: Diagramme du tableau blanc de [A Digital Media Primer for Geeks](#) illustrant la largeur de bande de transition disponible pour un ADC / DAC 48 kHz (gauche) et un ADC / DAC 96 kHz (droite).*

Ce n'est que la moitié de l'histoire. Parce que les filtres numériques ont peu de limitations pratiques d'un filtre analogique, nous pouvons compléter le processus d'anticrénelage avec une plus grande efficacité et précision numérique. Le signal numérique brut à très haut débit passe à travers un filtre anti-crénelage numérique, qui n'a aucun mal à adapter une bande de transition dans un espace restreint. Après cet anticrénelage numérique supplémentaire, les échantillons de remplissage supplémentaires sont simplement jetés. La lecture suréchantillonnée fonctionne approximativement en sens inverse.

Cela signifie que nous pouvons utiliser un audio à faible débit de 44,1 kHz ou 48 kHz avec tous les avantages de fidélité d'un échantillonnage de 192 kHz ou plus (réponse en fréquence fluide, faible aliasing) et aucun des inconvénients (ultrasons qui provoquent une distorsion d'intermodulation, espace perdu). Presque tous les convertisseurs analogique-numérique (ADC) et convertisseurs numérique-analogique (DAC) d'aujourd'hui suréchantillonnent à des taux très élevés. Peu de gens réalisent que cela se produit parce que c'est complètement automatique et caché.

Les ADC et les DAC n'ont pas toujours suréchantillonné de manière transparente. Il y a trente ans, certaines consoles d'enregistrement enregistrées à des taux d'échantillonnage élevés en utilisant uniquement des filtres analogiques, et la production et le mastering utilisaient simplement ce signal à haut débit. Les étapes d'anti-aliasing et de décimation numérique (rééchantillonnage à un taux inférieur pour les CD ou les DAT) ont eu lieu dans les dernières étapes de la maîtrise. Cela pourrait bien être l'une des premières raisons pour lesquelles 96 kHz et 192 kHz ont été associés à la production musicale professionnelle [ 8 ].

## 16 bits contre 24 bits

OK, donc les fichiers musicaux à 192 kHz n'ont aucun sens. Couvert, fait. Qu'en est-il de l'audio 16 bits contre 24 bits?

Il est vrai que l'audio PCM linéaire 16 bits ne couvre pas tout à fait la plage dynamique théorique de l'oreille humaine dans des conditions idéales. De plus, il y a (et il y aura toujours) des raisons

d'utiliser plus de 16 bits pour l'enregistrement et la production.

Rien de tout cela n'est pertinent pour la lecture; ici, l'audio 24 bits est aussi inutile que l'échantillonnage à 192 kHz. La bonne nouvelle est qu'au moins 24 bits de profondeur ne nuisent pas à la fidélité. Cela n'aide pas et gaspille également de l'espace.

## Revisiter vos oreilles

Nous avons discuté de la plage de fréquences de l'oreille, mais qu'en est-il de la plage dynamique du son le plus doux possible au son le plus fort possible?

Une façon de définir la plage dynamique absolue serait de revoir le seuil absolu de l'audition et le seuil des courbes de douleur. La distance entre le point le plus haut sur le seuil de la courbe de douleur et le point le plus bas sur le seuil absolu de la courbe d'audition est d'environ 140 décibels pour un jeune auditeur en bonne santé. Cela ne durerait pas longtemps cependant; + 130 dB est assez fort pour endommager l'ouïe de façon permanente en quelques secondes à quelques minutes. À titre de référence, un marteau-piqueur à un mètre ne fait que 100-110 dB.

Le seuil absolu d'audition augmente avec l'âge et la perte auditive. Fait intéressant, le seuil de douleur diminue avec l'âge au lieu d'augmenter. Les cellules ciliées de la cochlée elles-mêmes ne possèdent qu'une fraction de la plage de 140 dB de l'oreille; la musculature de l'oreille ajuste continuellement la quantité de son atteignant la cochlée en déplaçant les osselets, tout comme l'iris régule la quantité de lumière pénétrant dans l'œil [ [9](#) ]. Ce mécanisme se raidit avec l'âge, limitant la plage dynamique de l'oreille et réduisant l'efficacité de ses mécanismes de protection [ [10](#) ].

## Bruit environnemental

Peu de gens réalisent à quel point le seuil absolu d'audition est calme.

Le son perceptible le plus silencieux est d'environ -8dB SPL [ [11](#) ]. À l'aide d'une échelle pondérée A, le bourdonnement d'une ampoule à incandescence de 100 watts à un mètre de distance est d'environ 10 dB SPL, soit environ 18 dB de plus. L'ampoule sera beaucoup plus forte sur un gradateur.

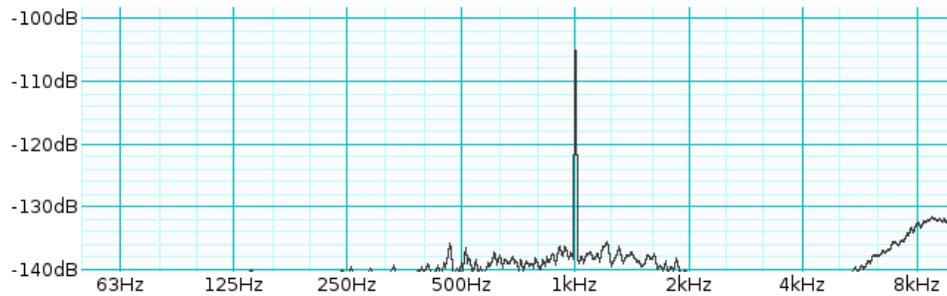
20 dB SPL (ou 28 dB plus fort que le son audible le plus silencieux) est souvent cité pour un studio de diffusion / enregistrement vide ou une salle d'isolation acoustique. C'est la référence pour un environnement exceptionnellement calme, et c'est une des raisons pour lesquelles vous n'avez probablement jamais remarqué entendre une ampoule.

## La plage dynamique de 16 bits

Le PCM linéaire 16 bits a une plage dynamique de 96 dB selon la définition la plus courante, qui calcule la plage dynamique en dB ( $6 * \text{bits}$ ). Beaucoup pensent que l'audio 16 bits ne peut pas représenter des sons arbitraires plus silencieux que -96 dB. Ceci est une erreur.

J'ai lié à deux fichiers audio 16 bits ici; l'un contient une tonalité de 1 kHz à 0 dB (où 0 dB est la tonalité la plus forte possible) et l'autre une tonalité de 1 kHz à -105 dB.

- **Échantillon 1:** [tonalité 1 kHz à 0 dB \(WAV 16 bits / 48 kHz\)](#)
- **Exemple 2:** [tonalité 1 kHz à -105 dB \(WAV 16 bits / 48 kHz\)](#)



*Ci-dessus: Analyse spectrale d'une tonalité de -105 dB codée en PCM 16 bits / 48 kHz. Le PCM 16 bits est clairement plus profond que 96 dB, sinon une tonalité de -105 dB ne pourrait pas être représentée, ni audible.*

Comment est-il possible de coder ce signal, de le coder sans distorsion, et de le coder bien au-dessus du plancher de bruit, lorsque son amplitude de crête est d'un tiers de bit?

Une partie de ce casse-tête est résolue par un tramage approprié, qui rend le bruit de quantification indépendant du signal d'entrée. Par implication, cela signifie que la quantification tramée n'introduit aucune distorsion, juste du bruit non corrélé. Cela implique à son tour que nous pouvons coder des signaux de profondeur arbitraire, même ceux avec des amplitudes de crête beaucoup plus petites qu'un bit [ 12 ]. Cependant, le tramage ne change pas le fait qu'une fois qu'un signal descend sous le plancher de bruit, il devrait effectivement disparaître. Comment la tonalité de -105 dB est-elle toujours clairement audible au-dessus d'un bruit de fond de -96 dB?

La réponse: notre valeur de bruit de fond de -96 dB est effectivement fautive; nous utilisons une définition inappropriée de la plage dynamique.  $(6 * \text{bits})$  dB nous donne le bruit RMS de l'ensemble du signal à large bande, mais chaque cellule ciliée de l'oreille n'est sensible qu'à une fraction étroite de la bande passante totale. Comme chaque cellule ciliée n'entend qu'une fraction de l'énergie totale du plancher de bruit, le plancher de bruit de cette cellule ciliée sera beaucoup plus faible que le chiffre à large bande de -96 dB.

Ainsi, l'audio 16 bits peut aller considérablement plus loin que 96 dB. Avec l'utilisation du tramage en forme, qui déplace l'énergie du bruit de quantification vers des fréquences où il est plus difficile à entendre, la plage dynamique effective de l'audio 16 bits atteint 120 dB en pratique [ 13 ], plus de quinze fois plus profonde que la revendication de 96 dB.

120dB est plus grand que la différence entre un moustique quelque part dans la même pièce et un marteau-piqueur à un pied ... ou la différence entre une pièce "insonorisée" déserte et un son suffisamment fort pour causer des dommages auditifs en quelques secondes.

16 bits est suffisant pour stocker tout ce que nous pouvons entendre, et sera suffisant pour toujours.

## Rapport signal sur bruit

Il convient de mentionner brièvement que le rapport S / N de l'oreille est inférieur à sa plage dynamique absolue. Dans une bande critique donnée, le rapport signal / bruit typique n'est estimé qu'à environ 30 dB. Le rapport signal / bruit relatif n'atteint pas la plage dynamique complète même si l'on considère des bandes largement espacées. Cela garantit que le PCM linéaire 16 bits offre une résolution plus élevée que celle réellement requise.

Il convient également de mentionner que l'augmentation de la profondeur de bits de la représentation audio de 16 à 24 bits n'augmente pas la résolution perceptible ou la « finesse » de l'audio. Il n'augmente que la plage dynamique, la plage entre le son le plus doux possible et le plus fort possible, en abaissant le bruit de fond. Cependant, un bruit de fond de 16 bits est déjà inférieur à ce que nous pouvons entendre.

## Quand le 24 bits est-il important?

Les professionnels utilisent des échantillons 24 bits pour l'enregistrement et la production [ [14](#) ] pour des raisons de marge, de bruit de fond et de commodité.

16 bits suffisent pour couvrir la vraie plage auditive avec de la place à revendre. Il ne couvre pas toute la gamme de signaux possible des équipements audio. La principale raison d'utiliser 24 bits lors de l'enregistrement est d'éviter les erreurs; plutôt que de faire attention à centrer l'enregistrement 16 bits - risquer l'écrêtage si vous devinez trop élevé et ajouter du bruit si vous devinez trop faible - 24 bits permet à un opérateur de définir un niveau approximatif et de ne pas trop s'en inquiéter. Manquer le réglage de gain optimal de quelques bits n'a aucune conséquence et les effets qui compressent dynamiquement la plage enregistrée ont un plancher profond avec lequel travailler.

Un ingénieur a également besoin de plus de 16 bits lors du mixage et du mastering. Les flux de travail modernes peuvent impliquer littéralement des milliers d'effets et d'opérations. Le bruit de quantification et le bruit de fond d'un échantillon de 16 bits peuvent être indétectables pendant la lecture, mais la multiplication de ce bruit par quelques milliers de fois devient finalement perceptible. 24 bits maintient le bruit accumulé à un niveau très bas. Une fois que la musique est prête à être distribuée, il n'y a aucune raison de conserver plus de 16 bits.

## Tests d'écoute

La compréhension est le point de rencontre de la théorie et de la réalité. Une affaire n'est réglée que lorsque les deux sont d'accord.

Les preuves empiriques des tests d'écoute confirment l'affirmation selon laquelle 44,1 kHz / 16 bits offre une lecture de fidélité la plus élevée possible. Il existe de nombreux tests contrôlés confirmant cela, mais je vais brancher un article récent, [Audibility of a CD-Standard A / D / A Loop Inserted into High-Resolution Audio Playback](#) , réalisé par des gens d'ici à la [Boston Audio Society](#) .

Malheureusement, le téléchargement du document complet nécessite une adhésion à l'AES. Cependant, il a été largement discuté dans des articles et sur des forums, avec les auteurs se joignant à. Voici quelques liens:

- [Le nouveau taux d'échantillonnage de l'empereur](#)
- [Fil de discussion du forum Hydrogen Audio](#)
- [Page d'informations supplémentaires à la Boston Audio Society, y compris l'équipement et les listes d'échantillons](#)

Cet article a présenté aux auditeurs un choix entre un contenu DVD-A / SACD à haut débit, choisi par des défenseurs de l'audio haute définition pour montrer la supériorité de la haute définition, et ce même contenu rééchantillonné sur place jusqu'à 16 bits / 44,1 kHz Compact Taux de disque. Les auditeurs ont été mis au défi d'identifier toute différence entre les deux en utilisant une méthodologie ABX. BAS a effectué le test en utilisant un équipement professionnel haut de gamme dans des environnements d'écoute de studio isolés par le bruit avec des auditeurs amateurs et professionnels formés.

Dans 554 essais, les auditeurs ont choisi correctement 49,8% du temps. En d'autres termes, ils devinaient. Pas un seul auditeur tout au long du test n'a pu identifier celui qui était 16 / 44,1 et celui qui était à haut débit [ [15](#) ], et le signal 16 bits n'était même pas tramé!

Une autre étude récente [ [16](#) ] a étudié la possibilité que les ultrasons soient audibles, comme l'ont suggéré des études antérieures. Le test a été conçu pour maximiser la possibilité de détection en plaçant les produits d'intermodulation là où ils seraient le plus audibles. Il a constaté que les sons ultrasoniques n'étaient pas audibles ... mais les produits de distorsion d'intermodulation introduits par les haut-parleurs pourraient l'être.

Cet article a inspiré de nombreuses recherches supplémentaires, dont la plupart avec des résultats mitigés. Une partie de l'ambiguïté s'explique par la constatation que les ultrasons peuvent également induire plus de distorsion d'intermodulation que prévu dans les amplificateurs de puissance. Par exemple, [David Griesinger a reproduit cette expérience](#) [ [17](#) ] et a constaté que sa configuration de haut-parleur n'introduisait pas de distorsion d'intermodulation audible par ultrasons, mais son amplificateur stéréo l'a fait.

## Caveat Lector

Il est important de ne pas sélectionner les articles individuels ou les «commentaires d'experts» hors de leur contexte ou de sources intéressées. Tous les articles ne sont pas entièrement d'accord avec ces résultats (et certains sont en désaccord en grande partie), il est donc facile de trouver des opinions minoritaires qui semblent justifier toutes les conclusions imaginables. *Quoi qu'il en soit, les articles et les liens ci-dessus sont représentatifs du vaste poids et de l'étendue du dossier expérimental.* Aucun article évalué par les pairs qui a résisté à l'épreuve du temps n'est en désaccord avec ces résultats. La controverse n'existe qu'au sein des communautés d'audiophiles consommateurs et passionnés.

Si quoi que ce soit, le nombre de résultats expérimentaux ambigus, non concluants et carrément invalides disponibles via Google souligne à quel point il est difficile de construire un test précis et objectif. Les différences que les chercheurs recherchent sont infimes; ils nécessitent une analyse statistique rigoureuse pour repérer les choix subconscients qui échappent à la conscience des sujets testés. Le fait que nous essayions probablement de «prouver» quelque chose qui n'existe pas rend la tâche encore plus difficile. Prouver une hypothèse nulle revient à prouver le problème

d'arrêt; tu ne peux pas. Vous ne pouvez collecter que des preuves qui donnent un poids écrasant.

Malgré cela, les articles qui confirment l'hypothèse nulle sont des preuves particulièrement solides; confirmer l'inaudibilité est beaucoup plus difficile expérimentalement que de le contester. Des erreurs non découvertes dans les méthodes de test et l'équipement produisent presque toujours des résultats faussement positifs (en introduisant accidentellement des différences audibles) plutôt que de faux négatifs.

Si les chercheurs professionnels ont tellement de mal à tester correctement les différences audibles minuscules et isolées, vous pouvez imaginer à quel point c'est difficile pour les amateurs.

### Comment [par inadvertance] bousiller une comparaison d'écoute

Le commentaire numéro un que j'ai entendu de la part des adeptes de l'audio à très haut débit était [paraphrasant]: "J'ai écouté l'audio à haut débit moi-même et l'amélioration est évidente. Vous me dites sérieusement de ne pas faire confiance à mes propres oreilles?"

Bien sûr, vous pouvez faire confiance à vos oreilles. Ce sont des cerveaux crédules. Je ne veux pas dire cela avec désinvolture; en tant qu'êtres humains, nous sommes tous câblés de cette façon.

### Biais de confirmation, effet placebo et double aveugle

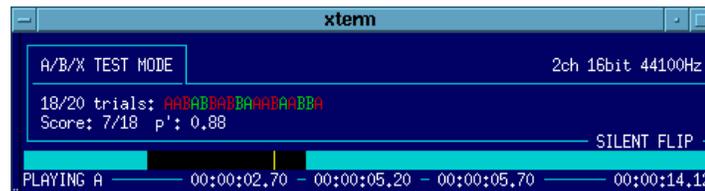
Dans tout test où un auditeur peut distinguer deux choix l'un de l'autre par tout autre moyen que l'écoute, les résultats seront généralement ceux que l'auditeur attendait à l'avance; c'est ce qu'on appelle le [biais de confirmation](#) et il est similaire à l' [effet placebo](#) . Cela signifie que les gens "entendent" les différences en raison de repères et de préférences subconscients qui n'ont rien à voir avec l'audio, comme préférer un amplificateur plus cher (ou plus attrayant) à une option moins chère.

Le cerveau humain est conçu pour remarquer des modèles et des différences, même là où il n'en existe pas. Cette tendance ne peut pas simplement être désactivée lorsqu'une personne est invitée à prendre des décisions objectives; c'est complètement inconscient. Un parti pris ne peut pas non plus être vaincu par un simple scepticisme. *Une expérimentation contrôlée montre que la conscience du biais de confirmation peut augmenter plutôt que diminuer l'effet!* Un test qui n'élimine pas soigneusement le biais de confirmation est sans valeur [ [18](#) ].

Dans les tests en *simple aveugle* , un auditeur ne sait rien à l'avance sur les choix de test et ne reçoit aucune rétroaction au cours du test. Le test en simple aveugle est meilleur que la comparaison occasionnelle, mais il n'élimine pas le [biais de l'expérimentateur](#) . L'administrateur du test peut facilement influencer le test par inadvertance ou transférer son propre biais subconscient à l'auditeur par des signaux involontaires (par exemple, «Êtes-vous sûr que c'est ce que vous entendez?», Langage corporel indiquant un «mauvais» choix, hésitant par inadvertance, etc. ). Il a également été démontré expérimentalement que le biais d'un expérimentateur influence les résultats d'un sujet de test.

Les tests d'écoute en *double aveugle* sont l'étalon-or; dans ces tests, ni l'administrateur du test ni le candidat n'ont aucune connaissance du contenu du test ou des résultats en cours. Les tests ABX exécutés par ordinateur sont l'exemple le plus célèbre, et il existe des outils disponibles

gratuitement pour effectuer des tests ABX sur votre propre ordinateur [ 19 ]. ABX est considéré comme une barre minimale pour qu'un test d'écoute soit significatif; les forums audio réputés tels que [Hydrogen Audio n'autorisent](#) souvent [même pas la discussion des résultats d'écoute à moins qu'ils ne répondent à cette exigence d'objectivité minimale](#) [ 20 ].



*Ci-dessus: Squishyball, un simple outil ABX en ligne de commande, fonctionnant dans un xterm.*

Personnellement, je ne fais aucun test de comparaison de qualité pendant le développement, même occasionnel, sans outil ABX. La science est la science, pas de relâchement.

## Astuces de Loudness

L'oreille humaine peut discriminer consciemment des différences d'amplitude d'environ 1 dB, et les expériences montrent une conscience subconsciente des différences d'amplitude sous 0,2 dB. Les humains considèrent presque universellement un son plus fort pour un meilleur son, et 0,2 dB est suffisant pour établir cette préférence. Toute comparaison qui ne parvient pas à faire correspondre soigneusement l'amplitude aux choix verra le choix le plus fort préféré, même si la différence d'amplitude est trop petite pour être consciente. Les vendeurs stéréo connaissent cette astuce depuis longtemps.

La norme de test professionnelle consiste à faire correspondre les sources à 0,1 dB ou mieux. Cela nécessite souvent l'utilisation d'un oscilloscope ou d'un analyseur de signal. Deviner en tournant les boutons jusqu'à ce que deux sources sonnent de la même manière n'est pas suffisant.

## Coupure

L'écêtage est une autre erreur facile, parfois évidente seulement rétrospectivement. Même quelques échantillons écêtés ou leurs effets secondaires sont faciles à entendre par rapport à un signal non écêté.

Le danger d'écêtage est particulièrement pernicieux dans les tests qui créent, rééchantillonnent ou manipulent autrement des signaux numériques à la volée. Supposons que nous voulons comparer la fidélité de l'échantillonnage à 48 kHz à un échantillon source à 192 kHz. Un moyen typique consiste à sous-échantillonner de 192 kHz à 48 kHz, à le rééchantillonner à 192 kHz, puis à le comparer à l'échantillon original de 192 kHz dans un test ABX [ 21 ]. Cette disposition nous permet d'éliminer toute possibilité de variation d'équipement ou de changement d'échantillon influençant les résultats; nous pouvons utiliser le même DAC pour jouer les deux échantillons et basculer entre sans aucun changement de mode matériel.

Malheureusement, la plupart des échantillons sont maîtrisés pour utiliser la gamme numérique complète. Le rééchantillonnage naïf peut et se coupera souvent de temps en temps. Il est nécessaire de surveiller l'écêtage (et de supprimer l'audio écêté) ou d'éviter l'écêtage par

d'autres moyens tels que l'atténuation.

## Différents médias, différents maîtres

J'ai parcouru quelques articles et articles de blog qui déclarent les vertus du 24 bits ou 96 / 192kHz en comparant un CD à un DVD audio (ou SACD) du " même " enregistrement. Cette comparaison n'est pas valide; les maîtres sont généralement différents.

## Indices par inadvertance

Les signaux audibles par inadvertance sont presque incontournables dans les anciennes configurations de test analogiques et hybrides numériques / analogiques. Les configurations de test purement numériques peuvent éliminer complètement le problème dans certaines formes de test, mais aussi multiplier le potentiel de bogues logiciels complexes. De telles limitations et bogues ont une longue histoire de résultats faussement positifs lors des tests [ [22](#) ].

[The Digital Challenge - More on ABX Testing](#) , raconte une histoire fascinante d'un test d'écoute spécifique réalisé en 1984 pour réfuter les autorités audiophiles de l'époque qui affirmaient que les CD étaient intrinsèquement inférieurs au vinyle. L'article ne s'intéresse pas tellement aux résultats du test (que je suppose que vous pourrez deviner), mais aux processus et au désordre du monde réel impliqués dans la conduite d'un tel test. Par exemple, une erreur de la part des testeurs a révélé par inadvertance qu'un expert audiophile invité n'avait pas fait de choix basés sur la fidélité audio, mais plutôt en écoutant les clics légèrement différents produits par les relais analogiques du commutateur ABX!

Les anecdotes ne remplacent pas les données, mais cette histoire est révélatrice de la facilité avec laquelle les défauts non découverts peuvent biaiser les tests d'écoute. Certaines des croyances audiophiles discutées ici sont également très divertissantes; on espère que certains exemples modernes seront considérés comme aussi idiots dans 20 ans.

## Enfin, la bonne nouvelle

Qu'est-ce qui fonctionne réellement pour améliorer la qualité de l'audio numérique que nous écoutons?

## De meilleurs écouteurs

La solution la plus simple n'est pas numérique. L'amélioration de la fidélité la plus spectaculaire possible pour le coût provient d'une bonne paire d'écouteurs. Sur l'oreille, dans l'oreille, ouverte ou fermée, peu importe. Ils n'ont même pas besoin d'être chers, bien que des écouteurs coûteux puissent en valoir la peine.

Gardez à l'esprit que certains écouteurs sont chers car ils sont bien fabriqués, durables et sonnent bien. D'autres sont chers car ce sont des écouteurs de 20 \$ sous une couche de plusieurs centaines de dollars de style, de marque et de marketing. Je ne ferai pas de recommandations spécifiques ici, mais je dirai que vous ne trouverez probablement pas de bons écouteurs dans un magasin à grande surface, même s'il est spécialisé dans l'électronique ou la musique. Comme dans tous les autres aspects de la hi-fi grand public, faites vos recherches (et mettez en garde

l'émetteur).

## Formats sans perte

Il est assez vrai qu'un fichier Ogg correctement encodé (ou un fichier MP3 ou AAC) ne pourra pas être distingué de l'original à un débit binaire modéré.

Mais qu'en est-il des fichiers mal encodés?

Il y a vingt ans, tous les encodeurs mp3 étaient vraiment mauvais par rapport aux normes d'aujourd'hui. Beaucoup de ces anciens encodeurs défectueux sont toujours utilisés, probablement parce que les licences sont moins chères et la plupart des gens ne peuvent pas faire la différence ou s'en moquent de toute façon. Pourquoi une entreprise dépenserait-elle de l'argent pour réparer ce qu'elle ne sait pas du tout?

Passer à un format plus récent comme [Vorbis](#) ou AAC n'aide pas nécessairement. Par exemple, de nombreuses entreprises et particuliers ont utilisé (et utilisent toujours) [l'encodeur Vorbis intégré de très basse qualité](#) de FFmpeg car il s'agissait de la valeur par défaut dans FFmpeg et ils ne savaient pas à quel point c'était mauvais. AAC a une histoire encore plus longue d'encodeurs largement déployés et de faible qualité; tous les formats avec perte traditionnels le font.

Les formats sans perte comme [FLAC](#) évitent toute possibilité d'endommager la fidélité audio [ [23](#) ] avec un encodeur avec perte de mauvaise qualité, ou même avec un bon encodeur avec perte utilisé de manière incorrecte.

Une deuxième raison de distribuer des formats sans perte est d'éviter la perte générationnelle. Chaque réencodage ou transcodage perd plus de données; même si le premier encodage est transparent, il est très possible que le second ait des artefacts audibles. Cela est important pour quiconque souhaite remixer ou échantillonner à partir de téléchargements. Cela est particulièrement important pour nous, chercheurs en codec; nous avons besoin d'un son clair pour travailler.

## De meilleurs maîtres

Le [test BAS que j'ai lié précédemment](#) mentionne en passant que la version SACD d'un enregistrement *peut* sonner sensiblement mieux que la version CD. Ce n'est pas à cause de l'augmentation de la fréquence d'échantillonnage ou de la profondeur mais parce que le SACD a utilisé un maître de meilleure qualité. Lorsqu'elle est renvoyée sur un CD-R, la version SACD semble toujours aussi bonne que la SACD d'origine et meilleure que la version CD, car l'audio d'origine utilisé pour créer la SACD était meilleur. Une bonne production et un bon mastering contribuent évidemment à la qualité finale de la musique [ [24](#) ].

La récente couverture de «Mastered for iTunes» et d'initiatives similaires d'autres labels de l'industrie est quelque peu encourageante. Ce qui reste à voir, c'est si Apple et les autres «l'obtiennent» ou si ce n'est qu'un crochet pour vendre aux consommateurs une autre copie plus chère de la musique qu'ils possèdent déjà.

## Entourer

Un autre «crochet de vente» possible, que j'accepterais avec enthousiasme, est celui des enregistrements surround. Malheureusement, il y a un péril technique ici.

Le surround discret à l'ancienne avec de nombreux canaux (5.1, 7.1, etc.) est une relique technique remontant aux théâtres des années 1960. Il est inefficace, utilisant plus de canaux que les systèmes concurrents. L'image surround est limitée et a tendance à s'effondrer vers les enceintes les plus proches lorsqu'un auditeur s'assoit ou se déplace hors de position.

Nous pouvons représenter et encoder une localisation excellente et robuste avec des systèmes comme Ambisonics. Les problèmes sont le coût de l'équipement de reproduction et le fait que quelque chose encodé pour un champ sonore naturel sonne mal lorsqu'il est mixé en stéréo et ne peut pas être créé artificiellement de manière convaincante. Il est difficile de simuler des ambisoniques ou un son holographique, un peu comme la façon dont la vidéo 3D semble toujours dégénérer en un gadget criard qui rend de manière fiable 5% de la population malade.

L'audio binaural est également difficile. Vous ne pouvez pas le simuler car cela fonctionne légèrement différemment chez chaque personne. C'est une compétence acquise réglée sur le système d'auto-assemblage des pennes, des canaux auditifs et du traitement neuronal, et elle ne s'assemble jamais exactement de la même manière chez deux individus. Les gens déplacent également inconsciemment la tête pour améliorer la localisation et ne peuvent pas bien localiser à moins qu'ils ne le fassent. C'est quelque chose qui ne peut pas être capturé dans un enregistrement binaural, bien qu'il puisse dans une certaine mesure en surround fixe.

Ce sont des obstacles techniques à peine impossibles. Le surround discret a fait ses preuves sur le marché, et je suis personnellement particulièrement enthousiasmé par les possibilités offertes par Ambisonics.

## Outro

"Je n'ai jamais beaucoup aimé la musique.  
C'est la haute fidélité!"

—Flanders & Swann, [une chanson de reproduction](#)

Le fait est d'apprécier la musique, non? La fidélité de lecture moderne est incompréhensiblement meilleure que les excellents systèmes analogiques disponibles il y a une génération. L'extrême logique est-il plus qu'un simple [problème du premier monde](#) ? Peut-être, mais de mauvais mélanges et codages *ne* me dérange pas; ils me distraient de la musique, et je ne suis probablement pas seul.

Pourquoi repousser le 24/192? Parce que c'est une solution à un problème qui n'existe pas, un modèle d'entreprise basé sur l'ignorance volontaire et l'escroquerie. Plus la pseudoscience est incontrôlée dans le monde en général, plus il est difficile pour la vérité de vaincre la véracité ... même s'il s'agit d'un petit exemple relativement insignifiant.

"Pour moi, il vaut bien mieux saisir l'Univers tel qu'il est réellement que de persister dans l'illusion, aussi satisfaisante et rassurante soit-elle."

-Carl Sagan

## Lectures complémentaires

Les lecteurs m'ont alerté sur une paire d'excellents articles dont je n'étais pas au courant avant de commencer mon propre article. Ils abordent plusieurs des mêmes points que moi plus en détail.

- [Le codage audio numérique de haute qualité](#) par Bob Stuart de Meridian Audio est magnifiquement concis malgré sa plus grande longueur. Nos conclusions diffèrent quelque peu (il considère comme étant donné la nécessité d'une plage de fréquences et d'une profondeur de bits légèrement plus larges sans beaucoup de justification), mais la présentation est claire et facile à suivre. *[Edit: Je ne suis peut-être pas d'accord avec beaucoup d'autres articles de M. Stuart, mais j'aime beaucoup celui-ci.]*
- [La théorie de l'échantillonnage pour l'audio numérique](#) [Lien mis à jour le 2012-10-04] par Dan Lavry de Lavry Engineering est un autre article que plusieurs lecteurs ont souligné. Il développe mes deux pages environ sur l'échantillonnage, le suréchantillonnage et le filtrage en un traitement plus détaillé de 27 pages. Ne vous inquiétez pas, il existe de nombreux graphiques, exemples et références.

Stephane Pigeon de [audiocheck.net](#) écrit pour brancher les tests d'écoute basés sur le navigateur présentés sur son site Web. L'ensemble de tests est encore relativement petit, mais plusieurs étaient directement pertinents dans le contexte de cet article. Ils ont bien fonctionné et j'ai trouvé la qualité assez bonne.

## Notes de bas de page

1. Comme l'a écrit une affiche frustrée:

"[Le théorème d'échantillonnage] n'a pas été inventé pour expliquer le fonctionnement de l'audio numérique, c'est l'inverse. L'audio numérique a été inventé à partir du théorème, si vous ne croyez pas le théorème, vous ne pouvez pas croire non plus à l'audio numérique !! "

<http://www.head-fi.org/t/415361/24bit-vs-16bit-the-myth-exploded>

2. Si ce n'était pas le tour de soirée le plus ennuyeux, c'était assez proche.
3. Il est plus courant de parler de lumière visible comme de longueurs d'onde mesurées en nanomètres ou angströms. J'utilise la fréquence pour être cohérent avec le son. Ils sont équivalents, car la fréquence est juste l'inverse de la longueur d'onde.
4. L'expérience LED ne fonctionne pas avec les LED `` ultraviolettes '', principalement parce qu'elles ne sont pas vraiment ultraviolettes. Ils sont suffisamment violets pour provoquer un peu de fluorescence, mais ils sont toujours bien dans la plage visible. Les vraies LED ultraviolettes coûtent entre 100 \$ et 1 000 \$ chacune et causeraient des dommages aux yeux si elles étaient utilisées pour ce test. Les LED non UV vraiment de qualité grand public émettent également un peu de lumière blanche afin de paraître plus brillantes, de sorte que vous seriez en mesure de les voir même si le pic d'émission était vraiment dans l'ultraviolet.

5. La version originale de cet article indiquait que les LED IR fonctionnent à partir de 300-325THz (environ 920-980nm), des longueurs d'onde invisibles. Un certain nombre de lecteurs ont écrit pour dire qu'ils pouvaient en fait à peine voir les LED dans certaines (ou toutes) de leurs télécommandes. Plusieurs ont eu la gentillesse de me faire savoir de quelles télécommandes il s'agissait, et j'ai pu en tester plusieurs sur un spectromètre. Et voilà, ces télécommandes utilisaient des LED à haute fréquence fonctionnant de 350 à 380 THz (800 à 850 nm), chevauchant simplement le bord extrême de la plage visible.
6. De nombreux systèmes qui ne peuvent pas lire les échantillons à 96 kHz seront sous-échantillonnés en silence à 48 kHz, plutôt que de refuser de lire le fichier. Dans ce cas, les tonalités ne seront pas jouées du tout et la lecture sera silencieuse, quelle que soit la linéarité du système.
7. Le suréchantillonnage n'est pas la seule application pour des taux d'échantillonnage élevés dans le traitement du signal. Il existe quelques avantages théoriques à produire de l'audio à bande limitée à un taux d'échantillonnage élevé en évitant la décimation, même s'il doit être sous-échantillonné pour la distribution. On ne sait pas exactement s'il en est utilisé dans la pratique, car le fonctionnement de la plupart des consoles professionnelles sont des secrets commerciaux.
8. Raisonement historique ou non, il ne fait aucun doute que de nombreux professionnels utilisent aujourd'hui des taux élevés car ils supposent à tort que conserver un contenu au-delà de 20 kHz sonne mieux, tout comme les consommateurs.
9. La sensation de tympan «décroissant» après avoir coupé la musique forte est bien réelle!
10. De jolis diagrammes peuvent être trouvés sur le site HyperPhysics:  
<http://hyperphysics.phy-astr.gsu.edu/hbase/sound/protect.html#c1>
11. 20  $\mu$ Pa est communément défini comme étant 0 dB à des fins de mesure auditive; il est approximativement égal au seuil d'audition à 1 kHz. L'oreille est cependant jusqu'à 8 dB plus sensible entre 2 et 4 kHz.
12. L'article suivant a la meilleure explication du tramage que j'ai rencontré. Bien qu'il s'agisse de tramage d'image, la première moitié couvre la théorie et la pratique du tramage en audio avant d'étendre son utilisation aux images:  
  
Cameron Nicklaus Christou, [Dither optimal et mise en forme du bruit dans le traitement d'image](#)
13. Les ingénieurs DSP peuvent souligner, comme l'a fait l'un de mes propres compatriotes smart-alec, que l'audio 16 bits a une plage dynamique théoriquement infinie pour un son pur si vous êtes autorisé à utiliser une transformée de Fourier infinie pour l'extraire; ce concept est très important pour la radioastronomie.

Bien que l'oreille ne fonctionne pas entièrement contrairement à une transformée de Fourier, sa résolution est relativement limitée. Cela limite la profondeur dynamique pratique maximale des signaux audio 16 bits.

14. La production utilise de plus en plus le flottant 32 bits, à la fois parce qu'il est très pratique sur les processeurs modernes et parce qu'il élimine complètement la possibilité d'écrêtage accidentel à tout moment sans être découvert et ruinant un mix.
15. Plusieurs lecteurs ont voulu savoir comment, si les ultrasons pouvaient provoquer une distorsion d'intermodulation audible, le test de Meyer et Moran 2007 aurait pu produire un résultat nul.

Il devrait être évident que «peut» et «parfois» ne sont pas les mêmes que «volonté» et «toujours». La distorsion d'intermodulation due aux ultrasons est une possibilité, et non une certitude, dans un système donné pour un ensemble donné de matériaux. Le résultat nul de Meyer et Moran indique que la distorsion d'intermodulation était inaudible sur les systèmes utilisés au cours de leurs tests.

Les lecteurs sont invités à [essayer le test de distorsion par intermodulation ultrasonique simple ci-dessus](#) pour une vérification rapide du potentiel d'intermodulation de leur propre équipement.

16. Karou et Shogo, [Détection de seuil pour les tonalités supérieures à 22 kHz](#) (2001). Document de convention 5401 présenté à la 110e convention, 12-15 mai 2001, Amsterdam.

17. Griesinger, [Perception de la distorsion d'intermodulation moyenne et haute fréquence dans les haut-parleurs, et sa relation avec l'audio haute définition](#)

18. Depuis la publication, plusieurs commentateurs m'ont écrit avec des versions similaires de la même anecdote [paraphrasée]: "Une fois, j'ai écouté des écouteurs / amplis / enregistrements qui attendaient le résultat [A] mais j'ai été totalement surpris de trouver [B] à la place! Le biais de confirmation est hooey ! "

J'offre deux réflexions.

Premièrement, le biais de confirmation ne remplace pas tous les résultats corrects par des résultats incorrects. Il biaise les résultats dans une direction incontrôlée d'un montant inconnu. Comment pouvez-vous dire le bien ou le mal *avec certitude* si le test est truqué par votre propre subconscient? Disons que vous vous attendiez à entendre une grande différence, mais que vous avez été choqué d'entendre une petite différence. Et s'il n'y avait en fait aucune différence? Ou, peut-être qu'il y *avait* une différence et, étant conscient d'un biais potentiel, votre scepticisme bien intentionné a été surcompensé? Ou peut-être aviez-vous parfaitement raison? Les tests objectifs, comme ABX, éliminent toute cette incertitude.

Deuxièmement, "Donc, vous pensez que vous n'êtes pas biaisé? Génial! Prouvez-le!" La valeur d'un test objectif réside non seulement dans sa capacité à informer sa propre compréhension, mais aussi à convaincre les autres. Les réclamations nécessitent une preuve. Les réclamations extraordinaires nécessitent une preuve extraordinaire.

19. Les outils les plus simples à utiliser pour les tests ABX sont probablement:

- [Foobar2000](#) avec le [plug-in ABX](#)

- [Squishyball](#) , un outil en ligne de commande Linux que nous utilisons dans Xiph

20. Chez Hydrogen Audio, l'exigence de test objectif est abrégée *TOS8* car il s'agit du huitième élément des conditions d'utilisation.

21. On suppose généralement que le rééchantillonnage endommage irrémédiablement un signal; ce n'est pas le cas. À moins que l'on ne commette une erreur évidente, comme provoquer un écrêtage, le signal sous-échantillonné puis suréchantillonné sera indiscernable de l'original. Il s'agit du test habituel utilisé pour établir que des taux d'échantillonnage plus élevés ne sont pas nécessaires.

22. Ce n'est peut-être pas strictement lié à l'audio, mais ... des neutrinos plus rapides que la lumière, n'importe qui?

23. [Le magazine Wired implique que les formats sans perte comme FLAC ne sont pas toujours complètement sans perte](#) :

"Certains puristes vous diront d'ignorer complètement les FLAC et d'acheter simplement des WAV. [...] En achetant des WAV, vous pouvez éviter la perte de données potentielle encourue lorsque le fichier est compressé en FLAC. Cette perte de données est rare, mais cela arrive . "

C'est faux. Un processus de compression sans perte ne modifie en aucune façon les données d'origine et FLAC ne fait pas exception.

Dans le cas où Wired faisait référence à une corruption matérielle des fichiers de données (défaillance du disque, défaillance de la mémoire, taches solaires), FLAC et WAV seraient tous deux affectés. Un fichier FLAC, cependant, est une somme de contrôle et détecterait la corruption. Le fichier FLAC est également plus petit que le WAV, et donc une corruption aléatoire serait moins probable car il y a moins de données qui pourraient être affectées.

24. La ["guerre du Loudness"](#) est un exemple couramment cité de mauvaises pratiques de mastering dans l'industrie aujourd'hui, même si ce n'est pas le seul. L'intensité est également un phénomène plus ancien que l'article de Wikipédia laisse croire au lecteur; dès les années 1950, les artistes et les producteurs ont poussé pour les enregistrements les plus forts possibles. Les fournisseurs d'équipement ont de plus en plus recherché et commercialisé de nouvelles technologies pour permettre aux maîtres de plus en plus chauds. Un équipement de mastering de vinyle avancé dans les années 1970 et 1980, par exemple, des enveloppes de rainures suivies et imbriquées lorsque cela est possible afin de permettre des amplitudes plus élevées que l'espacement des rainures ne le permettrait normalement.

La technologie numérique d'aujourd'hui a permis d'augmenter le volume à un niveau absurde. Il fournit également une pléthore de plug-ins DAW propriétaires automatiques, très complexes qui sont déployés en masse sans une large compréhension de leur fonctionnement ou de ce qu'ils font vraiment.

—Monty ( [monty@xiph.org](mailto:monty@xiph.org) ) 1er mars 2012

*dernière révision le 25 mars 2012 pour ajouter les améliorations suggérées par les lecteurs.  
Les modifications et corrections apportées après cette date sont marquées en ligne, à l'exception  
des fautes d'orthographe  
repérées le 30 décembre 2012 et le 15 mars 2014, et un «est» supplémentaire supprimé le 1er  
avril 2013]*



Les articles et travaux de démonstration de Monty sont parrainés par Red Hat Emerging Technologies.

(C) Copyright 2012 Red Hat Inc. et Xiph.Org

Un merci spécial à Gregory Maxwell pour ses contributions techniques à cet article



- 
- [Plane  
t.Xiph](#)
- [Dém  
os de  
Mont  
y](#)
- [Intro](#)
- [Physi  
ologi  
e](#)
- [Spect  
re  
audib  
le](#)
- [Oreill  
es  
dorée  
s](#)
- [Spect  
rophil  
es](#)
- [Inter  
modu  
lation](#)
- [Falla  
cies](#)
- [Suré  
chant  
illon  
age](#)
  - [16  
contr  
e 24  
bits](#)
- [Bruit](#)
- [Plage  
dyna  
miqu  
e](#)
- [Signa  
l-bruit](#)
  - [Pour  
quoi  
24  
bits?](#)
- [Tests  
d'éco  
ute](#)
- [Cave  
at](#)