

GOLDMUND

L'audio haute résolution est-il meilleur?

- 12 mars 2019

Les audiophiles, les professionnels de l'audio et les scientifiques ont tendance à ne pas s'entendre sur le point de savoir si l'audio numérique haute résolution sonne mieux que l'audio standard de qualité CD. Cet article examine les problèmes techniques et psychoacoustiques pour trouver une réponse.

Avec l'avènement récent des téléchargements haute résolution, les audiophiles ont adopté avec enthousiasme l'audio numérique haute résolution. Les convertisseurs numérique-analogique (CNA) devraient désormais inclure une capacité de résolution élevée et sont souvent jugés en fonction de la résolution numérique maximale et des formats de fichiers qu'ils prennent en charge. La technologie audio haute résolution est une fonctionnalité standard et attendue dans les équipements audio numériques professionnels depuis plus de dix ans.

Cependant, les chercheurs et les scientifiques de l'audio ont été moins enclins à adopter l'audio haute résolution. Ils signalent des problèmes techniques avec les différents formats. Ils se demandent si les avantages théoriques de l'audio haute résolution peuvent présenter des avantages compte tenu des limites de l'audition humaine et de l'environnement d'écoute typique. Ils soulignent le fait que la différence entre le son haute résolution et le son standard n'est que rarement détectable dans les tests d'écoute aveugle contrôlée.

Quel côté est correct? Ou est la vérité quelque part entre? Ce document examinera les problèmes et aboutira à une conclusion satisfaisante.

Avantages potentiels de l'audio haute résolution

Les partisans de l'audio haute résolution citent deux avantages pour la technologie: une plus grande largeur de bande et une plus grande plage dynamique (ou moins de bruit).

Largeur de bande supérieure : La largeur de bande supérieure de l'audio haute résolution provient de taux d'échantillonnage plus élevés. L'audio numérique à résolution standard à modulation par impulsions et codage (PCM) sur CD a un taux d'échantillonnage de 44,1 kilohertz, ce qui donne une réponse en fréquence d'environ 20 kHz. L'audio PCM haute résolution utilise généralement des fréquences d'échantillonnage de 96 ou 192 kHz, pour une réponse en fréquence légèrement inférieure à 48 et 96 kHz, respectivement. L'audio haute résolution Direct Stream Digital (DSD) dans sa version standard utilise une fréquence d'échantillonnage de 2,8 mégahertz; sa réponse en fréquence est d'environ 100 kHz, bien que ses performances en matière de bruit et de résolution à hautes fréquences soient médiocres.

La fréquence maximale que les humains peuvent entendre est d'environ 20 kHz; ce nombre diminue avec l'âge, surtout chez les hommes. Quel est donc l'avantage de capturer des fréquences plus élevées?

Résolution de synchronisation améliorée: Certains partisans de l'audio haute résolution citent la résolution de synchronisation de l'oreille humaine. Le son arrive presque toujours aux oreilles justifiées et droites à des moments légèrement différents. Les

humains peuvent détecter les différences de synchronisation entre l'oreille droite et l'oreille gauche aussi petites que 15 microsecondes; pour capturer ces informations, une fréquence d'échantillonnage de 66,7 kHz est requise. En omettant les fréquences supérieures à 20 kHz, ces subtiles différences de cadencement sont éliminées, ce qui pourrait, du moins en théorie, nuire à la perception du son.

Requêtes de filtrage réduites: dans l'audio numérique PCM, le signal est filtré pendant l'enregistrement pour éliminer le repliement (interférence avec les signaux de la bande audio par une énergie haute fréquence parasite), et pendant la lecture pour reconstruire une onde sonore lisse à partir de la sortie quantifiée de un convertisseur numérique-analogique. Une pente de filtre très raide (de 36 à 48 dB par octave) est généralement utilisée pour que les fréquences inférieures à la fréquence d'échantillonnage soient affectées le moins possible, tandis que les fréquences supérieures à la fréquence d'échantillonnage sont fortement atténuées. L'inclinaison du filtre introduit un déphasage et des ondulations dans la réponse en fréquence du dispositif, produisant des effets mesurables à des fréquences proches de la fréquence de coupure du filtre.

Avec l'audio PCM haute résolution, étant donné que le taux d'échantillonnage est plus élevé, la fréquence de coupure des filtres anti-aliasing et de reconstruction peut également l'être, et une pente de filtre plus douce telle que 12 dB par octave peut être utilisée. La fréquence d'échantillonnage du DSD standard étant d'environ 2,8 MHz, ses filtres peuvent être réglés sur des fréquences très supérieures à la plage audio. Ainsi, tout effet audible de ces filtres serait fortement réduit ou totalement éliminé.

Plus grande plage dynamique, moins de bruit:La plage dynamique plus grande et le bruit réduit de l'audio haute résolution proviennent de sa résolution d'amplitude accrue. L'audio numérique PCM à résolution standard utilise un échantillonnage sur 16 bits, capable d'enregistrer 65 536 niveaux de signal différents, pour une plage dynamique théorique de 96 dB (6 dB par bit). Cela signifie que le niveau de bruit inhérent est inférieur de -96 dB au niveau sonore le plus élevé que le système puisse enregistrer. Les enregistrements PCM haute résolution utilisant un échantillonnage sur 24 bits peuvent capturer 16 777 216 niveaux de signal différents, pour une plage dynamique de 144 dB. Ainsi, le bruit tombe à -144 dB en dessous du niveau maximum, même si même les meilleurs amplificateurs et convertisseurs numérique-analogique ont un niveau de bruit d'environ -120 dB. (DSD n'offre pas un avantage évident dans ce domaine car sa dynamique varie avec la fréquence,

Les audiophiles ont cité divers avantages de la résolution d'amplitude supplémentaire de l'audio haute résolution, notamment un sens accru des détails et de la dynamique, des arrière-plans plus calmes et un son plus naturel. Les ingénieurs professionnels enregistrent en utilisant des formats haute résolution, en partie parce que cela rend le réglage des niveaux d'enregistrement moins critique; ils peuvent régler le niveau de manière prudente pour éviter de dépasser le niveau de signal maximal (c.-à-d. l'écrêtage) tout en maintenant une plage dynamique étendue. Les professionnels citent également la plus grande précision de l'audio haute résolution pour réduire les artefacts parfois rencontrés dans le post-traitement numérique de l'audio.

Inconvénients potentiels de l'audio haute résolution

Nous avons assisté à de nombreuses discussions dans les publications audio grand public et professionnelles sur les avantages potentiels de l'audio haut de gamme, mais très peu sur les inconvénients techniques potentiels - inconvénients qui sont

probablement plus audibles qu'une augmentation de la plage dynamique ou de la réponse en fréquence.

Augmentation de la distorsion: la principale faille audible de l'audio haute résolution implique une distorsion d'intermodulation (IMD), l'effet créé lorsque deux tonalités audio interfèrent l'une avec l'autre. Avec l'IMD, des sons de somme et de différence sont créés, généralement à des fréquences qui ne sont pas harmoniquement liées aux deux sons d'origine. L'IMD est présent dans tous les équipements audio jusqu'à un certain point, mais des décennies d'évolution de la conception l'ont réduit à l'insignifiance.

Cependant, l'IMD est un problème plus commun aux fréquences ultrasonores. Les équipements qui ne sont pas conçus pour reproduire de telles fréquences - notamment de nombreux amplificateurs et la plupart des haut-parleurs (tweeters) utilisés dans les haut-parleurs actuels - peuvent produire un important IMD s'ils sont forcés de fonctionner à des fréquences qu'ils n'ont pas été conçus pour gérer. Malheureusement, les effets de l'IMD ne se limitent pas aux hautes fréquences.

Par exemple, si un enregistrement haute résolution contient des tonalités à 28 et 30 kHz, un haut-parleur ou un amplificateur sujet aux IMD haute fréquence reproduira (ou tentera de reproduire) non seulement les tonalités à 28 et 30 kHz, mais également la somme et la différence des tons différents. La tonalité de différence - 30 000 moins 28 000 - se produira à 2 kHz, en plein milieu de la gamme de fréquences dans laquelle l'oreille humaine est la plus sensible. Par conséquent, l'hypothèse selon laquelle l'extension de la capacité haute fréquence d'un système audio sera toujours bénéfique est incorrecte.

Un article de 2001 intitulé «Détection de seuil pour des tonalités supérieures à 22 kHz», rédigé par des chercheurs de l'Institut national japonais des sciences et technologies industrielles avancées, a confirmé cette thèse. Les chercheurs ont utilisé des signaux de test combinant une tonalité de 2 kHz jouée avec et sans harmoniques ultrasonores. Lorsqu'un seul haut-parleur était utilisé pour reproduire le son, l'IMD apparaissant dans le système de lecture permettait aux auditeurs de détecter la présence des harmoniques ultrasonores. Lorsqu'un second système haut-parleur / amplificateur était utilisé pour reproduire les harmoniques ultrasonores et que le système haut-parleur / amplificateur d'origine ne reproduisait que la tonalité de 2 kHz, les auditeurs ne pouvaient pas détecter les tonalités ultrasonores. Cela suggère que, même si l'IMD causé par les tonalités ultrasonores était audible, les tonalités ultrasonores ne l'étaient pas, bien qu'elles aient été enregistrées au même niveau que les tonalités à 2 kHz.

Réduction potentielle de la durée de vie des équipements: Les tweeters typiques commencent à atteindre leurs modes de rupture - les fréquences auxquelles leurs composants physiques se comportent de manière non linéaire - à des fréquences comprises entre 25 et 30 kHz. Lorsque les modes de rupture se produisent, le diaphragme du tweeter (dôme) se déforme hors de sa forme d'origine, créant des motifs de vagues dans le diaphragme auparavant lisse. Une distorsion constante du diaphragme en excitant ces modes de rupture peut entraîner une fatigue physique du diaphragme et d'autres composants mécaniques du conducteur, entraînant une distorsion et une défaillance éventuelle du conducteur.

La plupart des circuits d'amplificateurs et de préamplificateurs grand public filtrent les fréquences ultrasonores afin d'éviter les oscillations, un état dans lequel le circuit génère spontanément des tonalités de haute amplitude et de haute fréquence et s'éteint rapidement. Cependant, ce filtrage n'est pas total ou parfait. Comme chaque technicien de l'amplificateur peut probablement le prouver par son expérience personnelle, forcer un circuit audio à reproduire des hautes fréquences à des niveaux élevés provoque

souvent une défaillance des composants électriques du circuit. Ce n'est peut-être pas un problème avec les meilleurs amplificateurs et préamplis haut de gamme, car beaucoup d'entre eux sont spécialement conçus pour gérer les fréquences ultrasonores. Mais cela peut être un problème avec des composants de qualité inférieure connectés à un système haut de gamme, ou avec des systèmes de qualité inférieure installés ailleurs dans une maison.

Est-ce que l'audio haute résolution surpasse vraiment l'audio CD?

Bien que les prétendus avantages techniques de l'audio haute résolution soient souvent vantés dans les documents de marketing et les publications audio, ces avantages sont souvent déjà obtenus dans les systèmes de résolution standard.

Suréchantillonnage pour réduire les exigences de filtrage: Les avantages de l'échantillonnage à des fréquences supérieures à 44,1 kHz peuvent être et sont obtenus en utilisant le suréchantillonnage dans la conversion analogique-numérique et numérique-analogique à définition standard.

Presque toutes les puces de convertisseur analogique-numérique (ADC) et de convertisseur numérique-analogique (DAC) actuellement disponibles suréchantillonnent à des débits élevés, généralement à 96 kHz ou plus. Avec le suréchantillonnage, le filtre anti-repliement dans l'ADC et le filtre de reconstruction dans le CNA peuvent utiliser des fréquences plus élevées et / ou des pentes plus graduelles. Si l'audio est stocké en résolution standard, les échantillons supplémentaires sont simplement ignorés. Cependant, les avantages de la fréquence de filtrage plus élevée et / ou des pentes de filtrage plus graduelles demeurent, tout en éliminant le problème de l'IMD aux fréquences ultrasonores.

Plage dynamique au-delà de 96 dB: S'il est généralement admis que l'enregistrement avec une profondeur de bits supérieure permet de capturer des sons plus calmes, 16 bits suffisent en fait pour capturer tous les sons pouvant être reproduits même avec le meilleur équipement audio.

Le bruit de fond d'un système audio numérique est généralement égal à moins 6 dB par bit. En d'autres termes, un système audio numérique 16 bits a un bruit de fond inférieur de 96 dB au niveau du signal enregistrable maximal. Cependant, la perception commune en audio est que le bruit de fond représente une barrière imperméable sous laquelle rien ne peut être entendu, ce qui est incorrect.

En enregistrement audio numérique PCM, le système ne s'éteint pas simplement lorsque le niveau du signal est inférieur à -96 dB. Dither - une petite quantité de bruit - est ajouté au signal afin que le système n'exclue pas les informations situées en dessous du niveau de signal minimal théorique. En règle générale, le dither est limité aux hautes fréquences où il n'est pas audible. Le résultat peut encore être amélioré par des techniques de mise en forme du bruit qui déplacent le bruit inhérent des systèmes audionumériques vers des fréquences ultrasonores. Il en résulte que le bruit de fond d'un système 16 bits dans la gamme de fréquences audibles peut atteindre, dans la pratique, -110 dB, voire -120 dB.

Résolution dynamique non augmentée: Il est généralement admis que l'augmentation de la résolution en bits de l'audio haute résolution permet une définition plus précise des

niveaux audio, et donc un meilleur rendu de la «microdynamique». Mais là encore, ce n'est pas le cas. La profondeur de bits accrue permet simplement une plus grande plage dynamique: niveaux de signal enregistrables maximum et minimum plus élevés. Par exemple, un échantillon de grosse caisse de +46,324 dB plus fort qu'une note de flûte sera de +46,324 plus fort si vous utilisez un système 16 bits ou 24 bits. Ajouter une plage dynamique au-delà de 16 bits revient à ajouter une plage plus large à un radar utilisé pour mesurer la vitesse des voitures. un radar capable de mesurer des vitesses allant de 0,01 à 100 000 kilomètres par heure n'est pas plus utile dans cette application que celui capable de mesurer des vitesses allant de 1 à 1 000 kilomètres par heure.

Il existe une application dans laquelle une plus grande profondeur de bits peut avoir des avantages: l'enregistrement et l'édition audio professionnels. L'audio subit souvent de nombreuses générations de traitement lors de l'enregistrement, du montage et de la post-production, et chaque étape peut ajouter du bruit. En utilisant un traitement 24 bits, le bruit ajouté est minimisé. Lorsque la production audio est terminée, il peut être converti en 16 bits pour la distribution.

L'utilisation de profondeurs de bits supérieures à 16 bits dans les systèmes audio grand public n'introduit aucun effet néfaste, elle ne fait que gaspiller de l'espace, nécessitant davantage de stockage et des durées de téléchargement plus longues.

Les avantages de l'audio haute résolution sont-ils même détectables?

Même si l'on accepte les prétendus avantages de l'audio haute résolution tout en ignorant les inconvénients potentiels, il est difficile et peut-être même impossible d'affirmer que l'audio haute résolution procure un avantage perceptible même dans les meilleures conditions d'écoute.

Gamme dynamique limitée de l'environnement d'écoute: Les grands systèmes de home cinéma installés par des professionnels sont calibrés à un niveau de pression acoustique de référence (SPL) de 105 dB pour chaque canal principal. Cependant, la plupart des gens considèrent ce niveau trop élevé pour une écoute confortable. (Pour référence, un marteau-piqueur produit environ 100 dB SPL à 1 mètre.) Même la plupart des amateurs de cinéma à domicile écoutent à des niveaux inférieurs de -6 dB environ. Les systèmes audio à deux canaux sont rarement lus à des niveaux supérieurs à 100 dB.

À l'opposé de la plage dynamique, les meilleurs studios d'enregistrement professionnels ont un bruit de fond autour de 30 dB SPL. Pour un home cinéma professionnellement installé, isolé et traité acoustiquement, le bruit de fond peut atteindre environ 40 dB SPL. Pour un salon, le bruit de fond est généralement d'environ 50 dB SPL - ou même plus fort si le salon est ouvert à d'autres parties de la maison.

Ainsi, même en considérant un niveau d'écoute extrêmement élevé de 110 dB, dans une pièce d'écoute isolée acoustiquement avec un bruit de fond de niveau acoustique de 40 dB, la plage dynamique de l'environnement d'écoute n'est que de 70 dB. Cela se situe bien dans la plage dynamique généralement admise de 96 dB de l'audio numérique à résolution standard.

Les fréquences ultrasonores sont indétectables à l'oreille: les recherches n'ont pas permis de démontrer une amélioration audible significative obtenue grâce à des

fréquences d'échantillonnage plus élevées. L'une des rares études ayant montré une capacité statistiquement significative des auditeurs à percevoir les effets d'un taux d'échantillonnage croissant au-delà de 44,1 kHz - «Discrimination du taux d'échantillonnage: 44,1 kHz contre 88,2 kHz», un document de 2014 rédigé par des chercheurs affiliés à l'Université McGill au Canada. - ont conclu que ces différences sont «très subtiles et difficiles à détecter». Il est important de noter que ce test n'a pas évalué la préférence des auditeurs pour le taux d'échantillonnage élevé, mais seulement s'ils pouvaient le distinguer ou non du taux inférieur - ce qui ils ne pouvaient jouer que sur l'une des cinq pièces musicales du test.

La raison pour laquelle la réponse en fréquence étendue des enregistrements haute résolution est si difficile à percevoir est que l'oreille humaine ne possède tout simplement pas les moyens physiques de le détecter. L'oreille détecte les sons à l'aide d'une série de cellules ciliées situées le long de la membrane basilaire. La position de chaque cellule ciliée sur la membrane lui permet de recevoir une gamme de fréquences sonores étroite. Au-dessus de la fréquence d'accord la plus élevée des cellules ciliées, aucun son n'est détecté.

Ainsi, la réponse de l'oreille humaine ne se réduit pas progressivement à l'infini, comme on le suppose généralement. Il décroît à mesure que la fréquence du son approche les 20 kHz, mais si le son est à une fréquence supérieure à celle de la cellule la plus optimisée, aucun son n'est détecté. Cela ressemble beaucoup aux limites de la vue humaine. Nous pouvons détecter une certaine gamme de fréquences lumineuses, mais la lumière infrarouge et ultraviolette est totalement indétectable à travers nos yeux.

Le format DSD offre-t-il des avantages par rapport au PCM?

Certains audiophiles, fabricants d'équipement audio et professionnels de l'enregistrement estiment qu'au-delà de la réponse en fréquence étendue de DSD par rapport au 16 / 44.1 PCM, il offre un son de meilleure qualité et plus naturel que le PCM. Cependant, mis à part les rapports anecdotiques recueillis lors de tests d'écoute pour voyants, peu d'éléments de preuve, voire aucun, ont été présentés à l'appui de cette affirmation.

Un article de 2004 intitulé «DVD-Audio vs. SACD: Discrimination perceptuelle des formats de codage audio numérique» comparait un DSD de 2,8 MHz à un 24 / 176,4 PCM lors d'un test d'écoute aveugle. Ce test a utilisé des enregistrements de performances musicales réalisés dans les deux formats avec des microphones à réponse haute fréquence étendue allant jusqu'à 40 ou 50 kHz. Les deux formats ont été alimentés avec des signaux analogiques directement à partir des microphones, sans mélange. Les tests ont été effectués avec 110 auditeurs. Sur 2 900 comparaisons, il y avait 1 454 choix corrects et 1 446 choix incorrects - à peu près les mêmes résultats que de lancer une pièce de monnaie. Les auteurs ont noté: «Ces personnes, pour la plupart, étaient bien habituées à une écoute critique au niveau professionnel, mais elles ont constaté qu'elles ne pouvaient même pas commencer à reconnaître les différences sonores....»

En outre, bien que les enregistreurs DSD soient maintenant disponibles même à des prix aussi bas que 500 USD, tous les enregistrements DSD, à l'exception des plus simples, doivent être convertis au format PCM pour édition, puis redistribués à DSD pour

distribution. Cette conversion élimine tous les avantages théoriques ou supposés de DSD.

Conclusion

Alors que de nombreuses étiquettes de musique, fabricants d'équipement audio et consommateurs ont vanté les avantages de l'audio haute résolution, il n'existe pour le moment aucune preuve scientifique concluante de son utilité. Il existe toutefois des preuves que l'audio haute résolution peut dans certaines circonstances réduire la fidélité par rapport à l'audio à résolution standard.

L'augmentation de la profondeur de bits de 16 à 24 bits ne nuit pas aux applications grand public, mais elle gaspille de l'espace de stockage et de transmission sans apporter de réel avantage. L'augmentation de la fréquence d'échantillonnage de 44,1 à 96 kHz ou plus n'apporte également aucun avantage réel et peut réellement réduire la fidélité.

Bien sûr, de nombreux amateurs et professionnels de l'audio contestent ces affirmations, mais nous ne connaissons aucune preuve scientifique à l'appui de leurs points de vue.

Rédigé par le laboratoire acoustique de Goldmund en collaboration avec l'expert de l'industrie, M. Brent Butterworth.