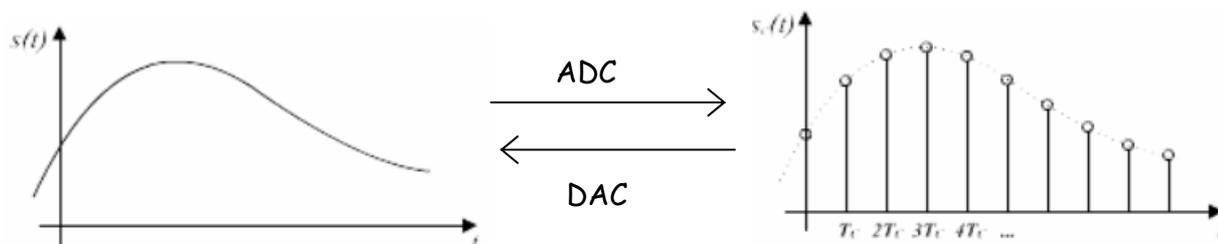


L'audio numérique

I/ Les bases de l'audio numérique

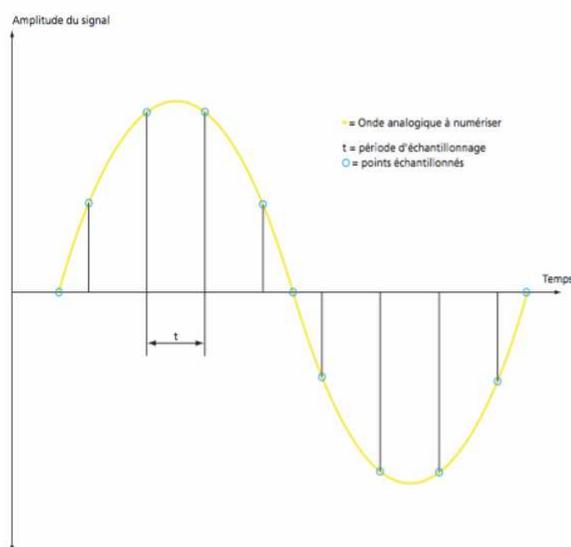
Le son est une vibration de l'air. Un ADC (Analogic to Digital Converter) convertit le signal analogique (la vibration de l'air) en sa représentation numérique (suite de nombres). Le DAC (Digital to Analogic Converter) effectue l'opération inverse.



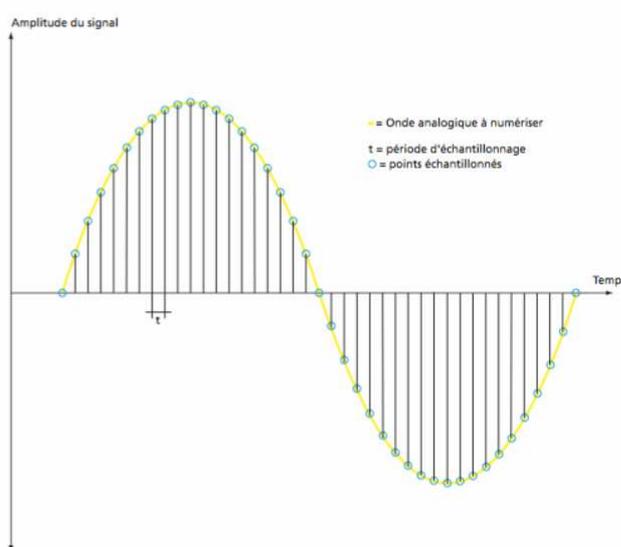
A- La fréquence d'échantillonnage

La **fréquence d'échantillonnage** (sample rate, en kHz) est le nombre d'échantillons (de mesures) par seconde. Cette fréquence traduit l'intervalle de temps entre deux points de la courbe du signal analogique à numériser.

Basse fréquence d'échantillonnage



Haute fréquence d'échantillonnage



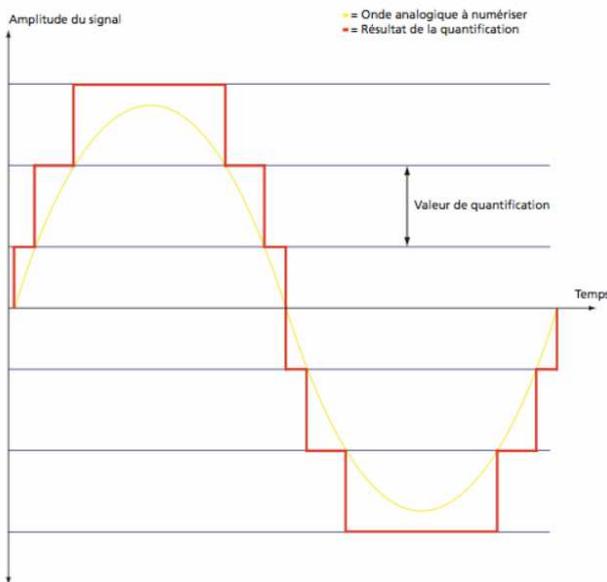
La fréquence d'échantillonnage doit être assez grande pour restituer un son qui semble continu à l'oreille. La majeure partie des humains ne perçoit plus les sons au-delà de 20 000 Hz. Le théorème de Nyquist-Shannon stipule que la fréquence d'échantillonnage doit être égale ou supérieure au double de la fréquence maximale du son à reproduire afin de pouvoir reproduire toutes les fréquences sans perte de données. Il faut donc une fréquence d'échantillonnage d'au moins 40 000 Hz pour obtenir un signal numérique fidèle du son d'origine.

La fréquence d'échantillonnage des CD est de $44\,100\text{ Hz} = 44\,100\text{ échantillons/s} = 44,1\text{ kHz}$

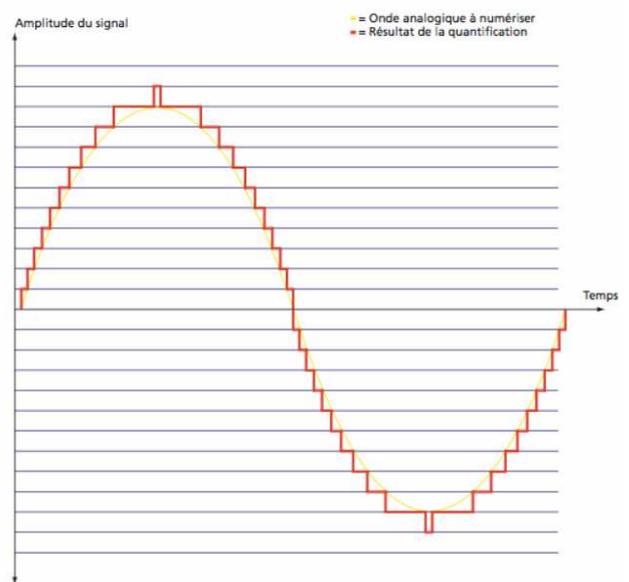
B- L'amplitude d'échantillonnage

L'**amplitude d'échantillonnage** (bit depth= « profondeur de bit », en bit) représente les valeurs possibles de l'amplitude (volume) = graduations de l'axe Y.

Basse quantification



Haute quantification



L'**amplitude d'échantillonnage des CD est de 16 bit** (soit $2^{16} = 65\,536$ valeurs de volume possible).

C- Le débit binaire

Fréquence d'échantillonnage × Amplitude d'échantillonnage = Débit binaire en bit/s

Le **débit binaire** (bitrate, en bit/s, ou kbit/s = kbps) représente la fidélité (qualité) de la représentation numérique du son par rapport au son analogique original.

Le débit binaire des CD est de $44,1\text{ kHz} \times 16\text{ bit} \times 2\text{ canaux (stéréo)} = \boxed{1\,411\text{ kbit/s}} = 176\text{ ko/s}$.

Les preneurs de son (ainsi que leurs ingénieurs) travaillent avec des fichiers à 96 kHz et 24 bits , soit : $96\text{ kHz} \times 24\text{ bit} \times 2\text{ canaux (stéréo)} = \boxed{4\,608\text{ kbit/s}} = 576\text{ ko/s}$.

Ce qui est plus du double de la qualité sonore d'un CD. Ce genre de fichiers, appelés **StudioMaster**, peuvent être achetés sur certains sites web.

Un fichier WAV issu d'un CD et contenant un signal brut PCM aura donc un débit de $1\,411\text{ kbit/s}$. 1h d'audio sur CD pèse donc 620 Mo , ce qui est beaucoup, d'où la nécessité d'une compression (voir ci-dessous).

Remarque : « bit depth » et « bitrate » sont deux expressions qui se ressemblent, d'où la confusion répandue entre l'amplitude d'échantillonnage et le débit binaire. Le bit depth est une amplitude, cad une valeur fixe, alors que le bitrate est une vitesse (débit).

II/ La compression audio

De manière générale, on trouve trois grands groupes de formats audio :

- les formats sans compression : type **.wav** ou **.aiff** = signal brut avec les fréquence et amplitude d'échantillonnages décrits précédemment (PCM = Pulse Code Modulation)
- les formats avec compression non destructive (lossless) : type **.flac** ou **.alac**
- les formats avec compression destructive (lossy) : type **.mp3** **.ogg** **.aac** (= **.m4a**) **.wma** etc...

En supprimant certaines fréquences (**compression destructive = lossy**) on obtient un fichier avec un débit binaire plus faible que 1 411 kbps, et donc un fichier moins lourd. Les **codecs** (codeurs-décodeurs) qui effectuent la compression permettent de diviser le débit jusqu'à 10 fois. Par exemple, un débit courant pour les mp3 est de **128 kbps**.

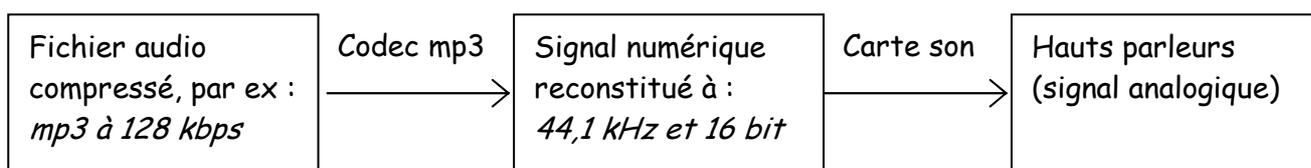
« L'oreille humaine ne peut percevoir que les sons compris entre 20 Hz et 16 kHz et elle est plus sensible entre 2 et 4 kHz. On pourrait donc supprimer tous les sons qui sont au dessous de 100 Hz et au dessus de 16 kHz et favoriser les sons compris entre 2 et 4 kHz.

L'oreille à une autre caractéristique intéressante, lorsqu'une grande énergie est présente sur une fréquence, on observe un pic, elle masque les fréquences voisines, les fréquences qui sont ainsi masquées ne sont pas ou peu perçues par l'oreille humaine. On peut les supprimer sans problème, l'oreille humaine ne percevra aucune différence. »

Il faut aussi ajouter la compression purement informatique (**compression non destructive = lossless**) comme ce que fait Winzip, Winrar, etc... Tout cela est déjà intégré dans les formats de compression comme le mp3, inutile de donc de "zipper" un fichier mp3, la taille du fichier final sera identique à celui d'origine !

Le challenge des différents codecs est d'être le plus efficace possible, cad de minimiser le débit, tout en rendant un signal le plus fidèle possible à l'original. Les codecs ne se valent donc pas tous. Le MP3 est le codec « historique », mais le AAC et le Vorbis sont des codecs plus récents et plus efficaces.

Lors de la lecture d'un fichier compressé, on a donc les étapes suivantes :



Alors que la lecture d'un fichier non compressé donne ceci :



Evidemment, la qualité du signal numérique reconstitué est inférieure à la qualité du signal numérique original, bien que les valeurs « des axes » soient identiques (44,1 kHz et 16 bit).

wikipedia

<http://pascal.ledisque.free.fr/informatique/son/pages/generalites.htm>

<http://thestereobus.com/2008/01/12/sample-rate-and-bitrate-the-guts-of-digital-audio>

<http://www.cuk.ch/articles/5354>