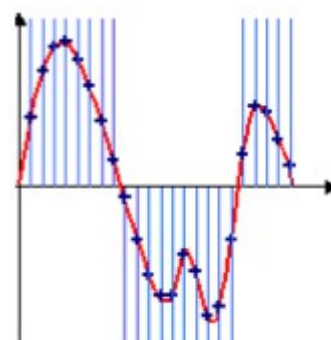


# Le codage des sons

## Table des matières

1. Le monde réel et le monde virtuel.....	2
1.1. Les différents types d'appareil.....	2
1.2. La numérisation.....	2
1.3. Convertisseurs analogique vers numérique.....	3
1.4. Convertisseurs numérique vers analogique.....	3
2. Le son.....	3
2.1. Qu'est - ce que le son ?.....	3
2.2. L'échantillonnage des sons.....	4
2.3. La quantification des sons.....	4
2.4. Mémoire requise pour stocker un son.....	5
3. Les formats audio numériques.....	6
3.1. Le format WAV.....	6
3.2. Le format MPEG, MP3.....	6
3.3. Le format Ogg Vorbis.....	6
3.4. Le format WMA.....	7
3.5. Le format AAC.....	7
3.6. Les Dolby Surround.....	7
3.7. Le DTS.....	7
3.8. Le FLAC.....	7

Pour pouvoir représenter un son sur un ordinateur, il faut arriver à le convertir en valeurs numériques, car celui-ci ne sait travailler que sur ce type de valeurs.



# 1. Le monde réel et le monde virtuel

## 1.1. Les différents types d'appareil

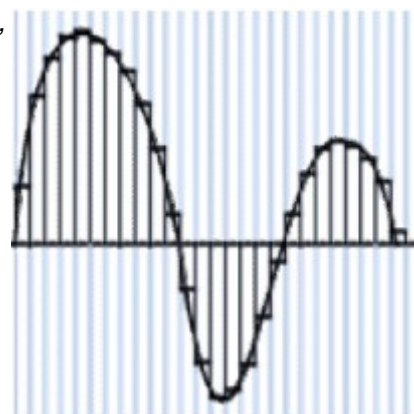
Les phénomènes qui nous entourent sont quasiment tous continus, c'est - à - dire que lorsque ces phénomènes sont quantifiables, ils passent d'une valeur à une autre sans discontinuité.

Ainsi, lorsque l'on désire reproduire les valeurs du phénomène, il s'agit de l'enregistrer sur un support, afin de pouvoir l'interpréter pour reproduire le phénomène original de la façon la plus exacte possible.

Lorsque le support physique peut prendre des valeurs continues, on parle d'enregistrement analogique. Par exemple, une cassette vidéo, une cassette audio ou un disque vinyle sont des supports analogiques.

Par contre, lorsque le signal ne peut prendre que des valeurs bien définies, en nombre limité, on parle alors de signal numérique.

La représentation d'un signal analogique est donc une courbe, tandis qu'un signal numérique pourra être visualisé par un histogramme. De cette façon, il est évident qu'un signal numérique est beaucoup plus facile à reproduire qu'un signal analogique (la copie d'une cassette audio provoque des pertes...).



## 1.2. La numérisation

La transformation d'un signal analogique en signal numérique est appelée numérisation. La numérisation comporte deux activités parallèles : **l'échantillonnage** (en anglais *sampling*) et la **quantification**.

L'échantillonnage consiste à prélever périodiquement des échantillons d'un signal analogique. La quantification consiste à affecter une valeur numérique à chaque échantillon prélevé.

- La qualité du signal numérique dépendra de deux facteurs : la fréquence d'échantillonnage (appelée *taux d'échantillonnage*) : plus celle - ci est grande (c'est - à - dire que les échantillons sont relevés à de petits intervalles de temps) plus le signal numérique sera fidèle à l'original ;
- le nombre de bits sur lequel on code les valeurs (appelé *résolution*) : il s'agit en fait du nombre de valeurs différentes qu'un échantillon peut prendre. Plus celui - ci est grand, meilleure est la qualité.

Ainsi, grâce à la numérisation on peut garantir la qualité d'un signal, ou bien la réduire volontairement pour :

- diminuer le coût de stockage
- diminuer le coût de la numérisation
- diminuer les temps de traitement
- tenir compte du nombre de valeurs nécessaires selon l'application
- tenir compte des limitations matérielles

### 1.3. Convertisseurs analogique vers numérique

Un convertisseur analogique numérique (CAN) est un appareil permettant de transformer en valeurs numériques un phénomène variant dans le temps. Lorsque les valeurs numériques peuvent être stockées sous forme binaire (donc par un ordinateur), on parle de données **multimédia**.

Un ordinateur dit "multimédia" est une machine capable de numériser des documents (papier, audio, vidéo...). Les principaux périphériques comportant des convertisseurs analogique vers numérique sont :

- la souris, l'écran tactile et tout mécanisme de pointage (tablette graphique, trackball, touchpad, ...)
- les scanners
- les cartes de capture sonore (la quasi - totalité des cartes son)
- les cartes d'acquisition vidéo
- les lecteurs (optiques comme le lecteur de CD - ROM, magnétiques comme le disque dur)
- les modems (à la réception)

### 1.4. Convertisseurs numérique vers analogique

Les convertisseurs numérique / analogique permettent de restituer un signal numérique en signal analogique. En effet, si une donnée numérique est plus facile à stocker et à manipuler, il faut tout de même pouvoir l'exploiter. A quoi servirait un son numérique si l'on ne pouvait pas l'entendre...

Ainsi, sur un ordinateur multimédia, on trouve des convertisseurs numérique vers analogique pour la plupart des sorties :

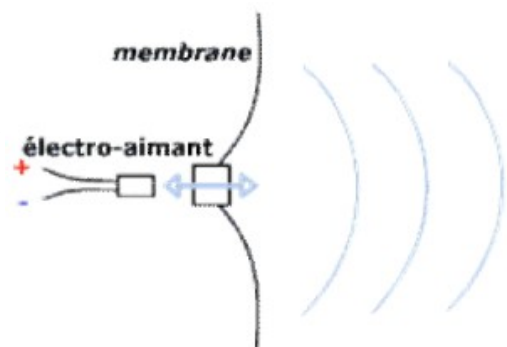
- sorties audio des cartes - sons
- synthétiseur musical
- imprimante
- écran
- modem (à l'émission)
- graveurs CD et DVD

## 2. Le son

### 2.1. Qu'est - ce que le son ?

Le son est une vibration de l'air, c'est - à - dire une suite de surpressions et de dépressions de l'air par rapport à une moyenne, qui est la pression atmosphérique. D'ailleurs, pour s'en convaincre, il suffit de placer un objet bruyant (un réveil par exemple) dans une cloche à vide pour s'apercevoir que l'objet initialement bruyant n'émet plus un seul son dès qu'il n'est plus entouré d'air !

La façon la plus simple de reproduire un son actuellement est de faire vibrer un objet. De cette façon, un violon émet



un son lorsque l'archet fait vibrer ses cordes, un piano émet une note lorsque l'on frappe une touche, car un marteau vient frapper une corde et la fait vibrer.

Pour reproduire des sons, on utilise généralement des haut - parleurs. Il s'agit en fait d'une membrane reliée à un électroaimant, qui, suivant les sollicitations d'un courant électrique va aller en avant et en arrière très rapidement, ce qui provoque une vibration de l'air situé devant lui, c'est - à - dire du son !

La numérisation des sons utilise le codage PCM Pulse Coded Modulation , initialement utilisé pour coder la voix dans le réseau téléphonique.

## 2.2. L'échantillonnage des sons

Pour pouvoir représenter un son sur un ordinateur, il faut arriver à le convertir en valeurs numériques, car celui - ci ne sait travailler que sur ce type de valeurs. Il s'agit donc de relever des petits échantillons de son (ce qui revient à relever des différences de pression) à des intervalles de temps précis. On appelle cette action l'échantillonnage ou la numérisation du son.

L'intervalle de temps entre deux échantillons est appelé taux d'échantillonnage. Etant donné que, pour arriver à restituer un son qui semble continu à l'oreille, il faut des échantillons tous les 10 000 èmes de seconde, il est plus pratique de raisonner sur le nombre d'échantillons par seconde, exprimés en Hertz (Hz). On parlera donc tout aussi bien de taux d'échantillonnage que de fréquence d'échantillonnage.

Il existe un certain nombre de fréquences d'échantillonnage normalisées :

Taux d'échantillonnage	Qualité du son
48000 Hz ou 48 kHz	pour les enregistreurs numériques multipistes professionnels et l'enregistrement grand public (DAT, Mini disc...)
44 100 Hz ou 44,1 kHz	qualité CD
32000 Hz ou 32 kHz	pour la radio FM en numérique
22 000 Hz ou 22 kHz	qualité radio
8 000 Hz ou 8 kHz	qualité téléphone

La valeur du taux d'échantillonnage, pour un CD audio par exemple, n'est pas arbitraire, elle découle en réalité du théorème de Shannon. La fréquence d'échantillonnage doit être suffisamment grande, afin de préserver la forme du signal. Le théorème de Nyquist - Shannon stipule que la fréquence **d'échantillonnage** doit être égale ou supérieure au **double de la fréquence maximale** contenue dans ce signal. Notre oreille perçoit les sons environ jusqu'à 22 000 Hz, il faut donc une fréquence d'échantillonnage au moins de l'ordre de 44 000 Hz pour obtenir une qualité satisfaisante.

## 2.3. La quantification des sons

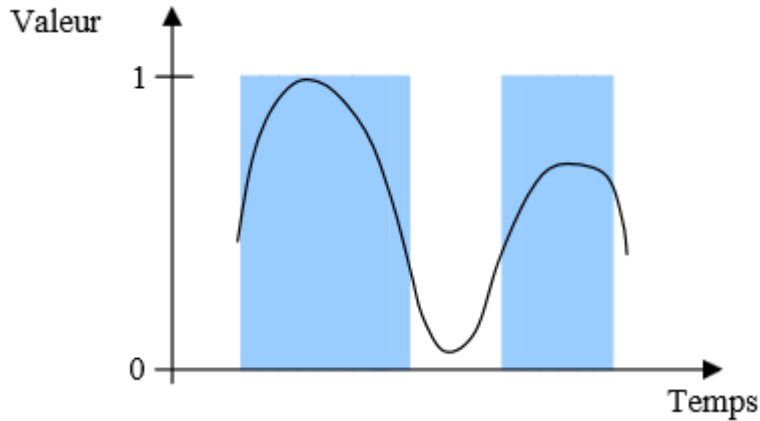
Nous venons de voir que la fréquence d'échantillonnage détermine la qualité du son, mais ce n'est pas tout. La quantification joue un rôle également fondamental.

Dans un ordinateur, toute information est stockée sous forme binaire (des 0 et des 1).

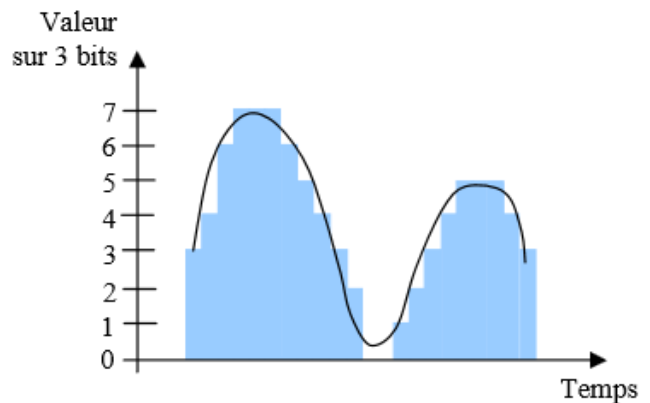
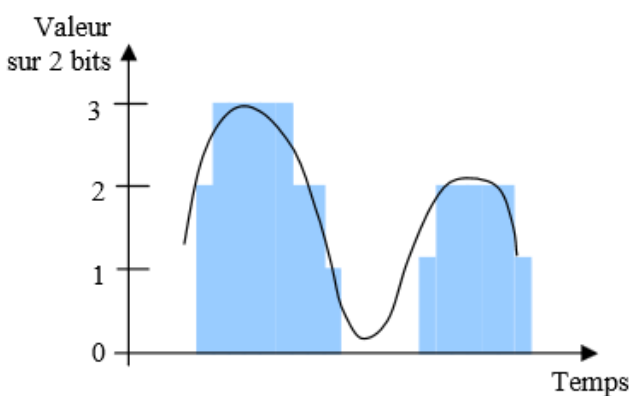
- Si on travaille sur 1 bit, alors on ne peut coder que deux valeurs : 0 et 1.
- Si on travaille sur 2 bits, alors on peut coder quatre valeurs : 0, 1, 2 et 3.
- Si on travaille sur 3 bits, alors on peut coder huit valeurs : 0 à 7.

• ...

À chaque échantillon est associée une valeur qui détermine la valeur de la pression de l'air à ce moment. Si on ne peut coder que 2 valeurs, la courbe va être numérisée ainsi :

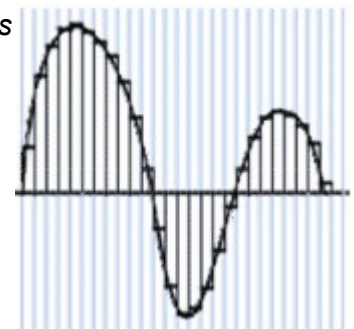


Si on code sur plus de bits, la numérisation s'affine :



Plus la capacité de quantification (le nombre de bits) est grande, plus le son de chaque échantillon sera précis et proche du vrai son.

- Sur 8 bits, on peut coder  $2^8 = 256$  fréquences sonores différentes
- Sur 16 bits, on peut coder  $2^{16} = 65\,536$  fréquences sonores différentes
- Sur 24 bits, on peut coder  $2^{24} = 16\,777\,216$  fréquences sonores différentes
- Sur 32 bits, on peut coder  $2^{32} \sim 4$  milliards de fréquences sonores différentes



Enfin, il ne faut pas oublier qu'un son stéréo est composé de deux sons, l'un dédié au côté gauche, et l'autre au côté droit.

## 2.4. Mémoire requise pour stocker un son

Lorsque l'on travaille avec un ordinateur, les informations en cours de traitement sont stockées dans la mémoire vive. Lorsque vous le souhaitez, vous les enregistrez sur le disque dur, ou sur un autre support, sous la forme d'un fichier.

Il faut donc avoir conscience de la capacité de votre ordinateur, pour comprendre ses limites.

Il est simple de calculer la taille d'une séquence sonore non compressée. En effet, en connaissant le nombre de bits sur lequel est codé un échantillon, on connaît la taille de celui-ci (la taille d'un échantillon est le nombre de bits...). Le nombre d'octets correspondant est facile à calculer, car 8 bits = 1 octet. Ainsi, si le codage de chaque échantillon se fait sur 16 bits, chaque échantillon occupera 2 octets.

Le taux d'échantillonnage permet de connaître le poids de chaque seconde d'enregistrement :

Si on travaille à un taux de 44,1 kHz (sur 16 bits), on aura 44100 échantillons → 2 octets = 88200 octets, soit 88 ko.

Il faut alors connaître la durée de la séquence sonore pour connaître le poids d'une voie : 10 secondes (à 44,1kHz, sur 16 bits) = 10 x 88 ko = 880 ko

Enfin, il faut connaître le nombre de voies. Si le son est stéréo, il faut enregistrer deux voies, soit 1660 ko, pour 10 secondes d'enregistrement audio numérique (à 44,1kHz, sur 16 bits).

La taille en octets d'une séquence sonore est ainsi égale à :

Taux d'échantillonnage x (Nombre de bits ÷ 8) x Nombre de secondes x Nombre de voies

Dans cette qualité d'enregistrement, un cédé audio (700 Mo) peut stocker théoriquement un peu moins de 70 minutes de musique.

### 3. Les formats audio numériques

Les sons numériques peuvent être encodés dans de nombreux formats, certains permettant de les compresser, d'autres non.

#### 3.1. Le format WAV

Un fichier WAV est un conteneur, généralement utilisé pour deux sous-formats :

- Le WAV PCM, format brut sans compression et donc très lourd.
- Le format RIFF - WAV, contenant généralement un fichier mp3 ou WAV, avec des métadonnées, mais très peu utilisé.



#### 3.2. Le format MPEG, MP3

Les formats MPEG sont destinés à l'encodage vidéo.

Il existe plusieurs versions de MPEG, dont le MPEG - 1, MPEG - 2 et le MPEG - 4.

Le MPEG - 1 est subdivisé en plusieurs couches (layers in english).

Chacune des couches permet un travail complémentaire des autres pour encoder des vidéos. La couche 3 permet entre autre l'encodage de la bande son de la vidéo. On parle donc de MPEG - 1 couche 3, le fameux MP3.

Le taux de compression du codec MP3 par rapport au format brut WAV PCM est de l'ordre de 80 % en moyenne sans perte de qualité audible.



#### 3.3. Le format Ogg Vorbis

Ogg Vorbis est un format de compression de données audio développé par la fondation Xiph.org, et largement adopté par la communauté Mac (ordinateurs, iPod, iPhone, ...) et par la communauté

de la musique professionnelle.

Pour préciser, Ogg est un conteneur, format de fichier capable de contenir divers données. Vorbis est un codec audio avec compression destructrice. Il est possible de rencontrer des fichiers Ogg n'utilisant pas le codec Vorbis, mais d'autres comme Flac (audio) ou Theora (Vidéo).

Le taux de compression du codec Vorbis par rapport au format brut WAV PCM est de l'ordre de 80 % en moyenne sans perte de qualité audible.



### 3.4. Le format WMA

Le format WMA (Windows Media Audio) constitue la variante Microsoft de la norme MPEG. La qualité des fichiers WMA étant très semblable à celle des MP3, Microsoft a dû recourir à des astuces plutôt douteuses pour promouvoir son format propriétaire.

En effet, Windows Media Player encode les MP3 avec une qualité inférieure aux fichiers WMA afin de faire croire à l'utilisateur que la qualité est meilleure avec son format.



### 3.5. Le format AAC

AAC (Audio Advanced Coding) est un format de compression de données audio développé par l'institut Fraunhofer en partenariat avec AT&T, Nokia, Sony et Dolby. AAC est un codec audio basé sur la norme Mpeg4, d'où son surnom MP4.

Ce nouveau standard serait vu comme le successeur du célèbre MP3.



### 3.6. Les Dolby Surround

Plusieurs générations de sons Dolby Surround se sont succédées jusqu'à l'instauration du système actuel le Dolby Digital. Les formats Dolby Surround, dont les dernières versions sont le Dolby Digital et le Dolby TrueHD (uniquement sur BluRay Disc), sont des formats audio, utilisés pour les films sur DVD, BluRayDisc et au cinéma, permettant notamment l'utilisation des systèmes 13.1 et un son d'une très grande qualité (24 bits et 96 kHz pour le Dolby TrueHD). Le codec utilisé est nommé AC - 3.



### 3.7. Le DTS

DTS (Digital Theater Sound) est un standard de codage numérique du son créé par les studios Universal, destiné également à la vidéo.



Concurrent direct du Dolby Digital, il compresse moins les données et possède un meilleur encodage que le Dolby Digital, mais moins bon que le Dolby TrueHD.

### 3.8. Le FLAC

Free Lossless Audio Codec (FLAC) est un codec libre de compression audio sans perte.

À l'inverse de codecs tels que MP3 ou Vorbis, il n'enlève aucune information du flux audio. Cette qualité maximale a pour conséquence une quantité d'information plus élevée, qui tout en étant assez variable, se trouve en moyenne être de l'ordre de 50 % de la taille du même fichier non



*compressé au format WAV PCM .*

*Le taux de compression du codec Flac par rapport au format brut WAV PCM est de l'ordre de 50% en moyenne sans perte de qualité audible.*