

SMB Formation - 17 bis, rue Palouzie 93400 Saint-Ouen
Tel : 01 40 11 01 96. Fax : 01 40 11 12 98. Mel. : gblum@smb-formation.com
Georges Blum : N° de Siret : 35212587600024 - Code APE: 9003B - N° de formateur : 11930435493

Traitement numérique du son

1/ - QU'EST-CE QU'UN SON?

1.1 – Définition

D'un point de vue physique, un son est une **énergie** qui se **propage** sous forme de **vibrations** dans un **milieu compressible** (dans l'eau, dans l'air, dans les matériaux solides, mais pas dans le vide!).

Lorsqu'on jette une pierre dans l'eau, on peut facilement observer le phénomène de propagation des ondes à la surface:

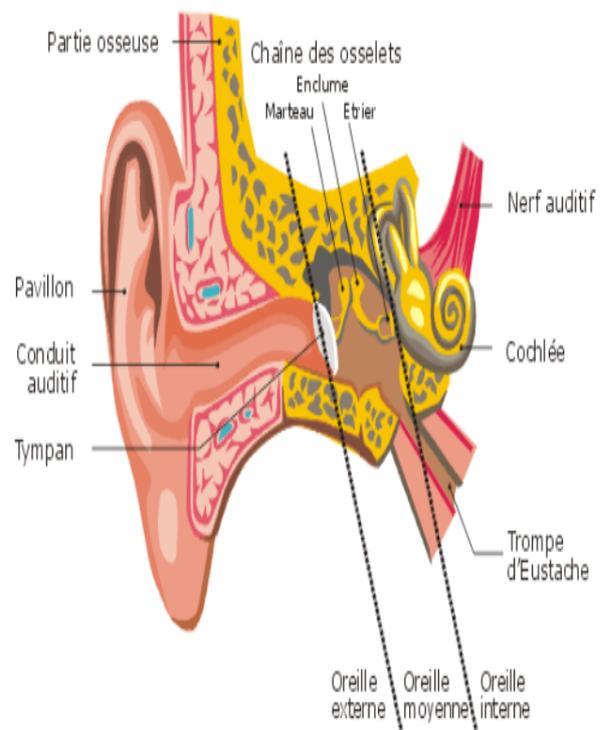
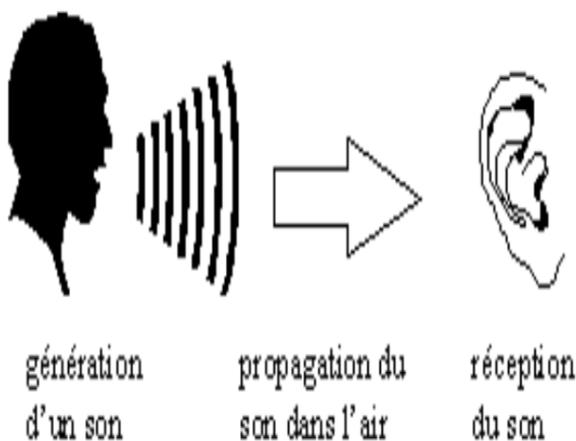


Lors de la diffusion d'un son dans un concert, c'est l'air qui permet sa transmission jusque nos oreilles. De même que l'exemple de l'eau illustré ci-dessus, les molécules d'air transmettent l'énergie et son donc un **support** pour le son.

1/ - QU'EST-CE QU'UN SON?

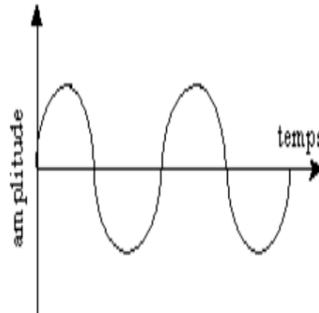
1.2 – Émission, Propagation, Réception

- Pour qu'un son soit émis, une **énergie** doit avant tout mettre en **mouvement un corps** pour produire une vibration. Ainsi, le muscle du larynx, la chute d'un objet sur le sol, ou la tension électrique dans un haut-parleur, provoqueront l'énergie nécessaire pour produire cette vibration.
- Ensuite, pour que ce son puisse se propager, il faut un **milieu élastique favorable à la transmission** de la vibration. En créant des surpressions ou des dépressions, l'air permet la propagation de l'onde. Les matériaux solides ont aussi cette capacité de transmettre le son. Dans le vide par contre, aucun son ne peut se propager, car il n'y a aucun de support.
- Enfin, pour être perçue, il doit y avoir un **récepteur sensible**. Chez l'homme, l'oreille possède une membrane (le tympan) capable de transmettre les informations de vibration en signaux nerveux jusqu'au cerveau, grâce au nerf auditif. De même, le microphone possède également une membrane permettant de transformer les déplacement de l'air en signaux électriques.



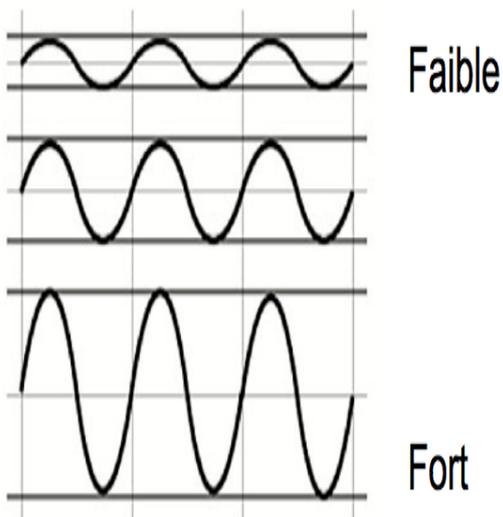
2/ - CARACTERISTIQUE DU SON

Comme tout phénomène vibratoire, le son peut être analysé comme un **signal** qui varie dans le temps. Deux caractéristiques essentielles sont l'**amplitude** et la **fréquence**.



2.1 – L'amplitude

- La première caractéristique d'un son est son **amplitude**. Appelée aussi intensité ou volume sonore, c'est l'expression de la pression de l'air qui se mesure en décibels (dB). 0 dB correspond au minimum que l'oreille humaine puisse percevoir (seuil d'audibilité). Attention, une augmentation de **3db** multiplie la puissance par deux!



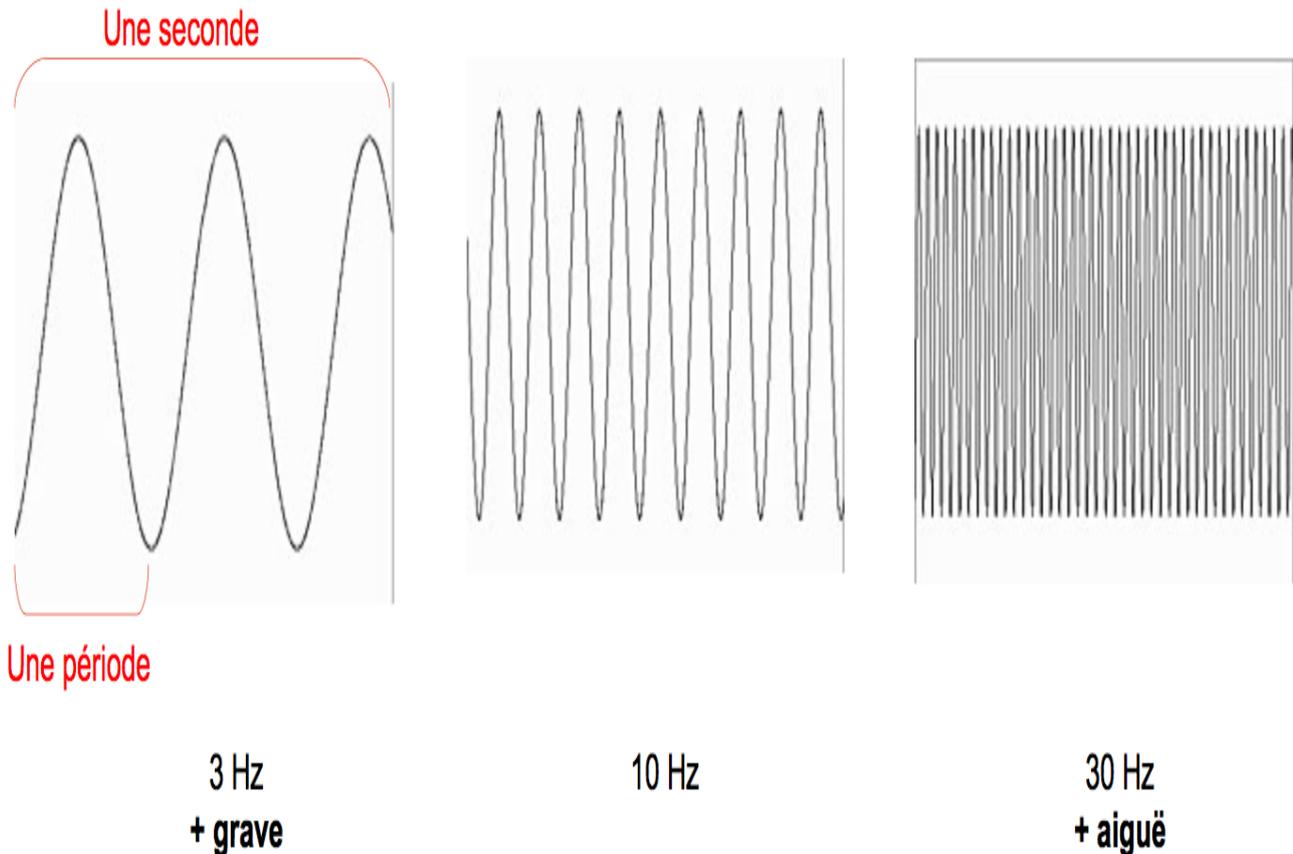
Exemple concret:

- De 0 à 10 dB : Seuil d'audibilité, Désert
- De 30 à 40 dB : forêt
- De 60 à 70 dB : sonnerie de téléphone
- De 80 à 90 dB : tondeuse à gazon, klaxon de voiture
- 120 dB : seuil de la douleur, avion au décollage
- 180 dB : décollage de la fusée Ariane, lancement d'une roquette

- L'évolution de l'amplitude sonore dans le temps s'appelle l'**enveloppe**.

2.2 – Fréquence

- La **fréquence**, exprimée en **Hertz (Hz)**, est le nombre de répétition d'une **période** par seconde. Plus elle est élevée et plus le son paraîtra « aiguë », à l'inverse, il paraîtra « grave ». En musique, la fréquence définit donc la **hauteur** d'un son, soit, la note. (Ex: la note « LA » correspond à 440Hz, soit 440 vibration en une seconde).



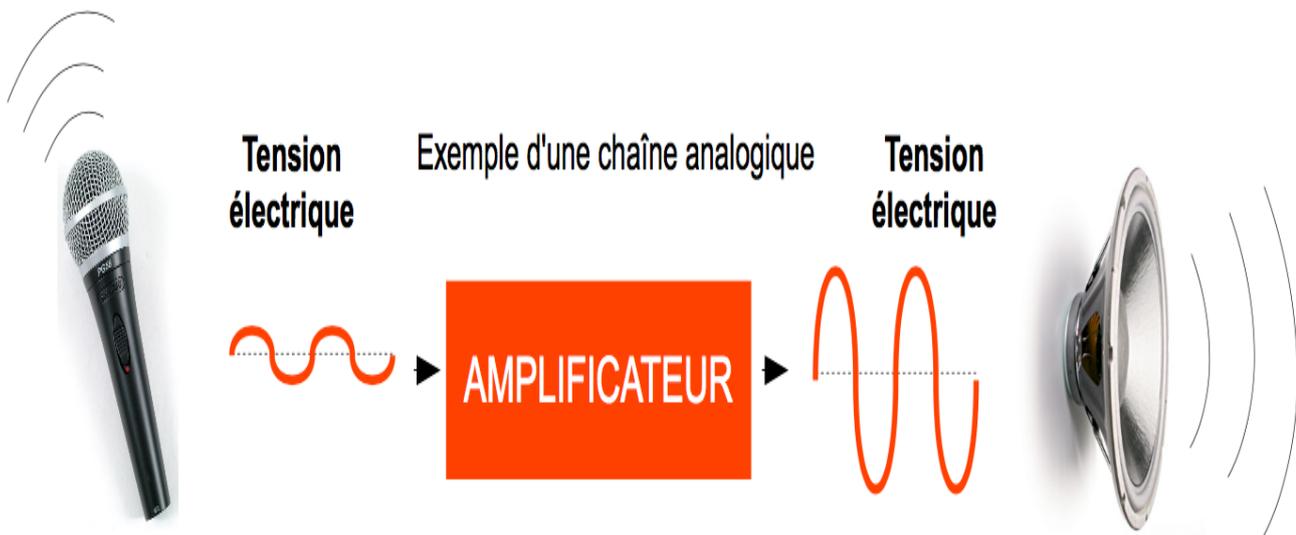
Le spectre de fréquence entendu par l'oreille humaine n'est pas infini, il s'étend environ de 20 Hz à 20 000 Hz (20KHz). En dessous on parle d'**infra-son** et au dessus on parle d'**ultra-son**.

Infra son < 20Hz
Basse 20 - 200Hz
Bas-Médium 200 - 2000Hz
Haut-Médium (ou aigu) 2000 - 12000Hz
aigu (ou sur-aigu) 12000 - 20000Hz
Ultra son >20000Hz

2/ DE L'ANALOGIQUE AU NUMERIQUE

2.1 – Le son analogique : un signal continu

Lorsqu'on capte un son à partir d'un microphone, ce dernier transforme l'énergie mécanique (la pression de l'air exercée sur sa membrane), en une **variation de tension électrique continue**. Ce signal électrique dit « **analogique** » pourra ensuite être amplifié, et envoyé vers un haut-parleur dont la fonction est inverse: transformer à nouveau le signal électrique en une énergie mécanique (on peut observer le déplacement de la membrane d'un haut parleur en marche).



Le son analogique est généralement fixé sur des supports comme les bandes magnétiques, disques vinyles etc. Le problème rencontré par ces supports réside dans l'usure physique des informations au cours de leur utilisation (lecture/écriture). A terme, le signal est affaibli et peut disparaître.

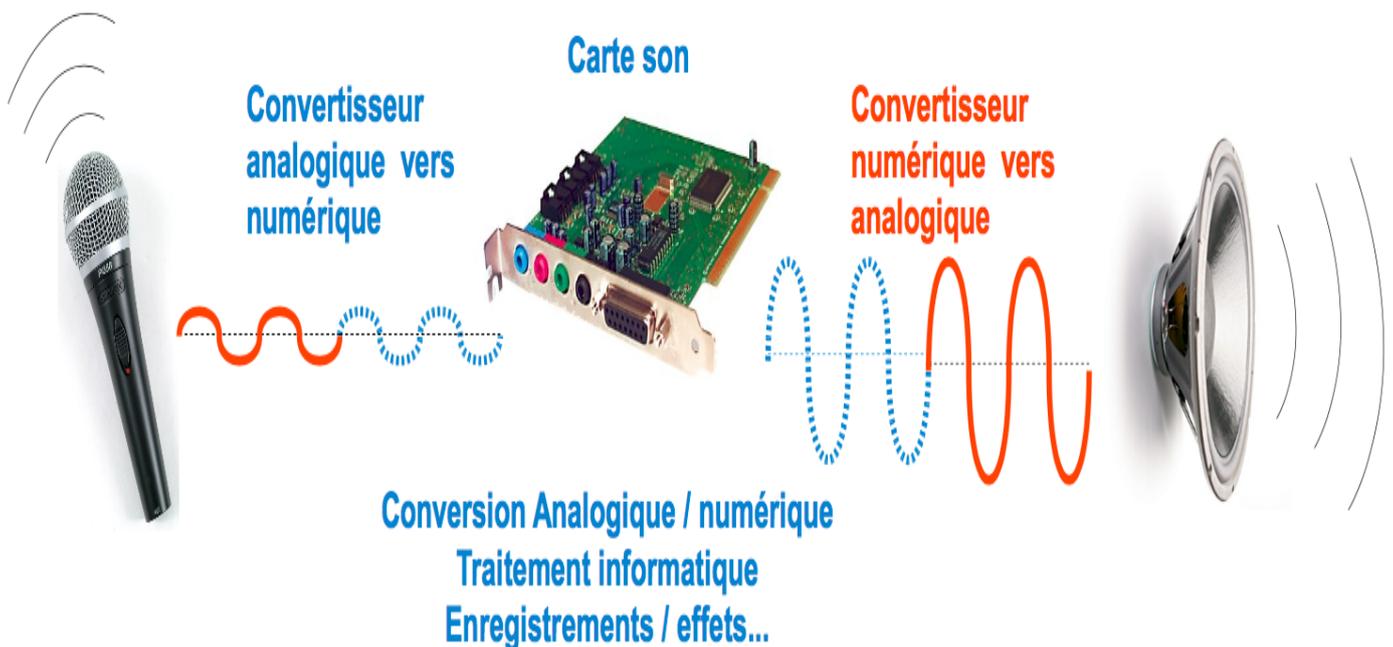


2.2 – Le son en numérique: un signal discontinu

Avec l'informatique, lorsque ce même signal électrique est capturé à partir du micro, il est converti en une **suite de nombre**, on parle alors de **numérisation du signal**. C'est la **carte son** qui s'en charge, elle contient des **entrées** (convertisseurs analogique vers numérique) et des **sorties** (convertisseurs numérique vers analogique).

La première phase appelée numérisation consiste donc à passer d'un signal **continu** (une variation de tension électrique) en une suite de **valeurs mesurées** à intervalles réguliers, donc **discontinu**.

Exemple d'une chaîne numérique



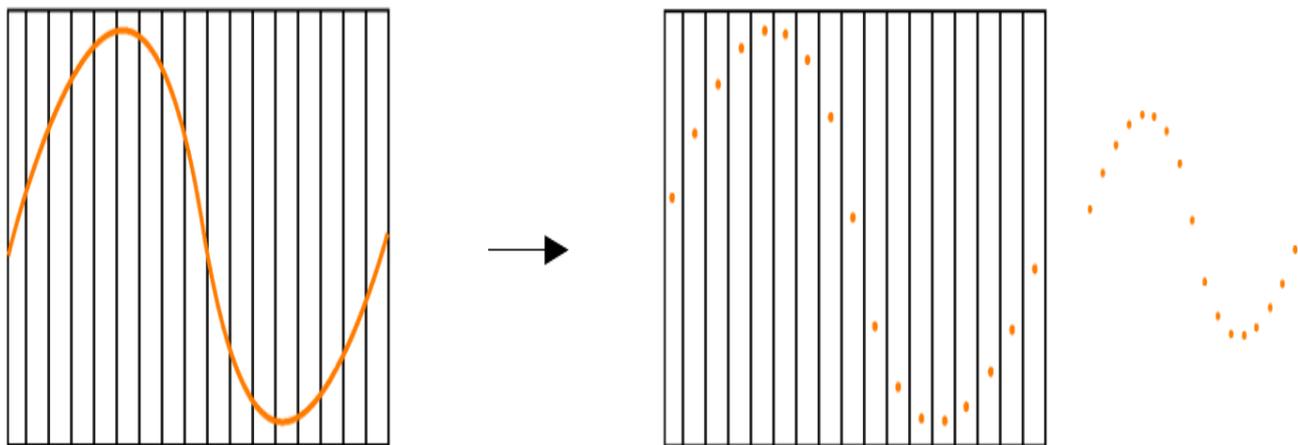
L'avantage du numérique, est la possibilité de lire et de dupliquer autant de fois ce signal sans aucune détérioration, puisqu'il a été réduit en une **suite de nombres** stockée dans un fichier informatique! Cela dit, la compression audio comme le MP3 peut provoquer une perte volontaire du signal afin d'économiser de l'espace de stockage (voir plus bas).

3/ CARACTERISTIQUE DU SON NUMERIQUE

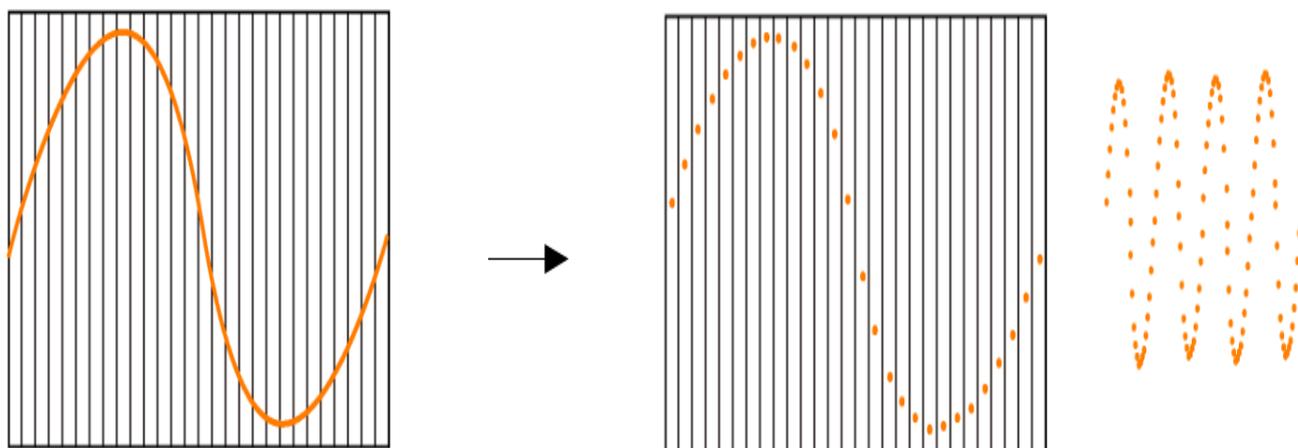
3.1 – Fréquence d'échantillonnage (Hz):

Lorsqu'un son est numérisé, le signal analogique (continu) qui entre dans l'ordinateur est mesuré, un certain nombre de fois par seconde (d'ou la discontinuité). Le son est donc découpé en "tranches", ou échantillons (en anglais « samples »). Le nombre d'échantillons disponibles dans une seconde d'audio s'appel la **fréquence d'échantillonnage exprimée en hertz**.

Pour traduire le plus fidèlement possible le signal analogique de notre micro, il faudra prendre le **plus grand nombre** de mesures possible par seconde. Autrement dit, plus la **fréquence d'échantillonnage** sera élevée, plus la traduction numérique du signal sera proche de l'original analogique.



Une fréquence d'échantillonnage faible peut convenir pour certains sons graves.

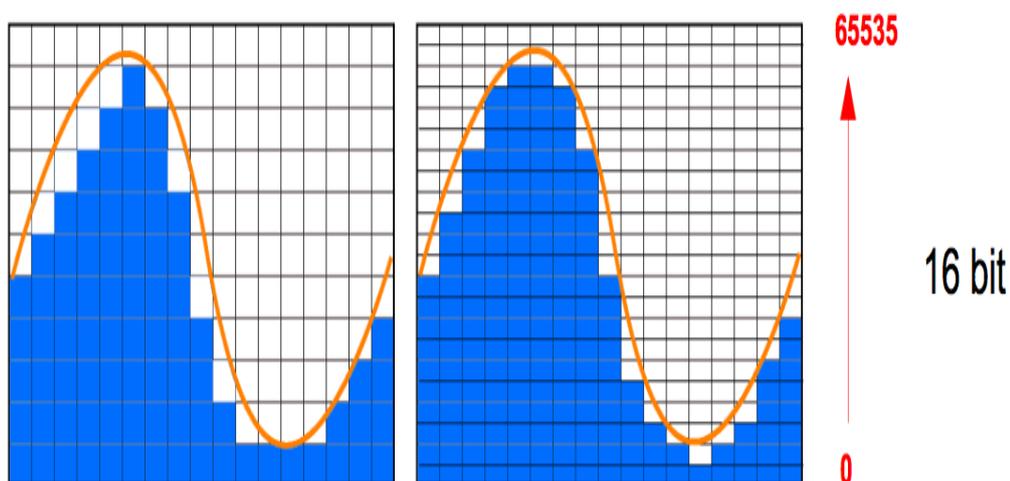


Doubler la fréquence d'échantillonnage améliore la réponse dans les fréquences hautes (les sons aigus).

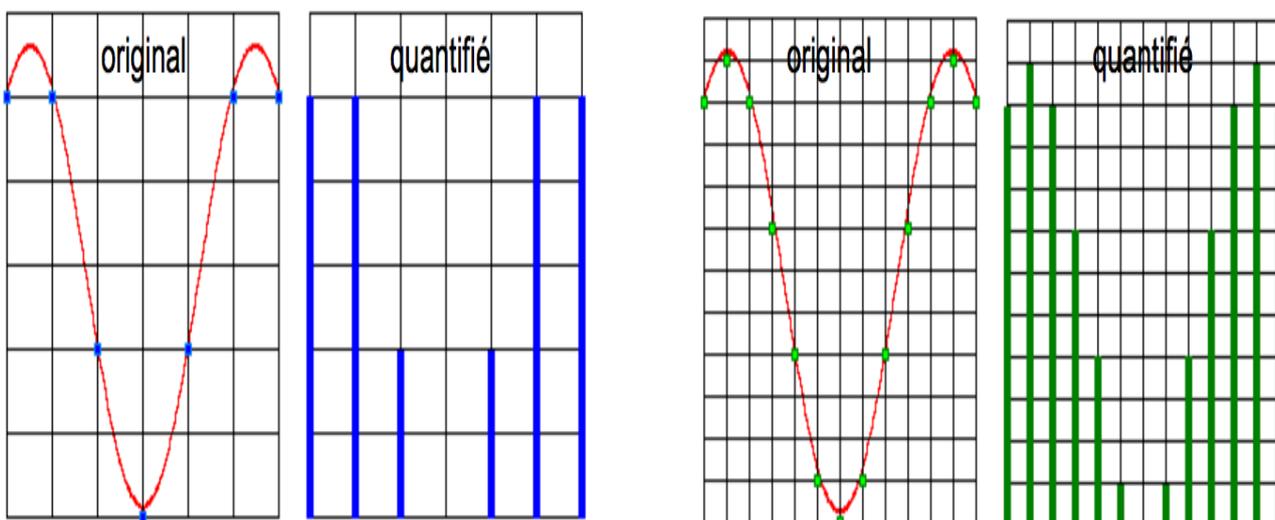
3/ CARACTERISTIQUES DU SON NUMERIQUE

3.2 – Résolution et quantification (bit):

Une autre caractéristique importante est la **résolution** numérique du son, soit le nombre de « niveaux » ou de « paliers » qu'il est possible d'enregistrer pour reproduire l'**amplitude** du signal. Avec une résolution de **16bit**, on dispose de 2^{16} , soit **65535** valeurs possibles pour traduire l'amplitude du son. Ainsi, plus la résolution est élevée, meilleur sera la **dynamique** (l'écart entre le son le plus faible et le plus fort qu'il est possible de reproduire).



La zone bleue montre qu'en doublant la résolution, on est plus proche de la courbe « analogique », soit le signal parfait que l'on souhaite reproduire.



La **quantification** consiste en une deuxième phase où le chiffre de l'amplitude prélevé sera arrondi à l'entier le plus proche.

Quelques exemples de résolutions fréquemment utilisées:

- Son qualité téléphone: 8000 Hz 8bit
- Son qualité radio FM: 22050 Hz 16bit
- Son qualité CD: 44100 Hz 16bit
- Son qualité DVD: 48000 Hz 24bit
- Son audio professionnel: 96000 et 192000 Hz 24 et 32bit

Mémoire requise pour stocker un son:

Il est simple de calculer la taille d'une séquence sonore **non compressée** en connaissant le **nombre d'échantillons par seconde** (fréquence d'échantillonnage), la **résolution (nombre de bits sur lequel est codé un échantillon)**, le **temps de la séquence (en seconde)** et le **nombre de voies utilisées** :

poids (octet) = Fréquence d'échantillonnage (Hz) x Résolution (octet) x Durée (seconde) x Nombre de voies

Exemple: Calcul d'une seconde d'audio qualité CD

Rappel: 1octet = 8bit et 1kilo-octet (ko) = 1024 octet

=> Calculer le poids d'1 minute audio en 44100Hz, 16bit, stéréo. On souhaite une réponse en Mega Octet (Mo).

44100(hz) x 16 (bit) x 60 (sec) x 2 (voies)

On sait que: 1octet = 8 bit, donc 16bit = 2 octets

44100 x 2 x 60 x 2 = 10584000 octet

Conversion de octet vers Kilo-octet (Ko): 10584000 / 1024 = 10335 Ko

Conversion de octet vers Mega-octet (Mo): 10335 / 1024 = 10 Mo

Résumé:

- Un son,

- * Est un phénomène **vibratoire**
- * Se **propage** dans un **milieu** particulier (air, eau, matière...)
- * 3 phases pour exister: émission / transmission / réception

- Ses caractéristiques sont:

- * **L'amplitude**, ou l'intensité du son exprimé en décibel (db)
- * **La fréquence**, répétition d'une période définissant ainsi la hauteur (grave ou aiguë)
- * La nature de son signal: **analogique (variation électrique) ou numérique (codage binaire)**

- Un **son numérique** est traduit par :

- * Sa **fréquence d'échantillonnage**, le nombre de « relevés » effectués chaque seconde.
- * le **nombre de bits** des échantillons, les valeurs numériques disponibles pour traduire l'amplitude du signal (8bit = 2^8 = 256 valeurs, 16bit = 2^{16} 65535 valeurs).
- * le nombre de **voies utilisées** (**Mono** = une voie, **Stéréo** = 2 voies, **Quadriphonie** = 4 voies...)

- Mémoire requise pour stocker un son non compressé en octet:

Fréquence d'échantillonnage(Hz) x Nombre de bits / 8 (octet) x Durée (sec) x Nombre de voies

3/ LES FORMATS DE FICHER SON

3.1 – Principaux formats de fichier non compressés:

Partie en cours de rédaction...

.WAVE : Mis au point par Microsoft et IBM, le format Wave PCM est le format son standard de Windows. Il est limité à un poids de 2Go. Le format "Disque compact" 44.1 kHz, 16 bits et stéréo nous servira de référence pour le calcul du poids et ratio des autres formats.

.AIFF : format de stockage des sons sur les ordinateurs Macintosh d'Apple. C'est l'équivalent du format WAV dans le monde Macintosh. Les résolutions de 8, 16, 20, 24 et 32 bits (à virgule flottante) sont acceptées.

.RAW : Format audio brut

.AU : Le format AU est assez bien répandu grâce à Unix et Linux. La fréquence d'échantillonnage est comprise entre 1 kHz et 200 kHz. Mais les applications de rendu audio ne lisent principalement que trois fréquences d'échantillonnage: 8012.821Hz (codec entré), 22050Hz et 44100Hz.

Les résolutions 8, 16, 20, 24 et 32 bits (flottant) sont acceptées.

...

Pour plus d'infos: http://wiki.univ-paris5.fr/wiki/Format_de_donn%C3%A9es#Les_formats_audio

3.2 – Principaux formats de fichier compressés:

Partie en cours de rédaction...

.MP3 : MPEG-1/2 Audio Layer 3, format de compression très populaire permettant d'occuper quatre à douze fois moins d'espace de données. Pour cela, certaines fréquences inaudibles par l'oreille humaine vont être totalement supprimées dans le fichier. La compression au format MP3 exploite aussi un modèle psycho-acoustique de l'effet dit de « masque » : si deux fréquences d'intensités différentes sont présentes en même temps, l'une peut être moins perçue que l'autre par l'oreille, selon que ces deux fréquences sont proches ou non.

.AAC: Amélioration du format MP3, c'est le format des fichiers audio supportés par Apple au sein de son baladeur numérique iPod et de son logiciel iTunes. L'AAC est un format de compression audio standardisé par l'ISO basé sur les normes du MPEG-4, d'où son nom MP4. Les fréquences d'échantillonnage vont de 8 kHz à 96 kHz (MP3 officiel : 16 à 48 kHz) et il peut gérer jusqu'à 48 canaux.

.WMA : format de compression audio propriétaire développé par Microsoft. Il offre la possibilité de protéger dès l'encodage les fichiers de sortie contre la copie illégale par une technique nommée gestion numérique des droits (ou GND).

.OGG Vorbis (ou OGA): formats et codecs multimédias ouverts, libres et dégagés de tout brevet. Le format de compression audio « Vorbis » est proposé par la fondation Xiph.Org. Moins populaire que le MP3, il lui est pourtant supérieur en terme de qualité à poids égal.

.RA: .ra (real audio), .rv (real video), .rm (real media), .ram (real audio metadata). Famille de codecs audio propriétaires (RealNetworks). Très ancien, il permet de diffuser de la musique sur internet en utilisant la technique du streaming.

3.3 – La compression de donnée:

La compression de donnée consiste à obtenir des fichiers plus légers, afin d'améliorer la vitesse de transfert sur internet ou limité l'espace de stockage utilisé sur un disque dur.

Il existe deux principaux types de compression:

- La compression sans perte: .zip .cab .rar .ace .7z .tar .gzip...

appelée aussi « compactage », cette solution consiste simplement à coder les données binaires de manière plus concise dans un fichier. Elle permet ainsi de retrouver la totalité des informations après une procédure de décompactage.

- La compression avec perte: .mp3 .ogg ...

Concernant essentiellement les fichiers de média (image, son, vidéo), elle consiste en une « réduction » de l'information basée sur notre propre limite de perception. Puisque l'oreille n'est pas sensible à toutes les fréquences audio (seulement entre 20Hz et 20KHz), et que sa sensibilité varie selon certaines d'entre elle, il est possible de réduire la quantité de données de telle sorte que le résultat soit très ressemblant à l'original, voire identique, pour l'oreille humaine.