

Table des matières.

<u>1). Introduction.</u>	2
<u>2). Le son.</u>	
1. Qu'est-ce que le son?	3
2. L'échantillonnage du son.	4
3. Compression du signal.	6
4. La transformée de fourier.	8
5. Représentation du son.	9
6. Mémoire requise pour stocker un son.	10
<u>3). La compression destructive.</u>	11
<u>4). La compression non destructive.</u>	13
<u>5). Le codage de Huffman.</u>	
1. Présentation.	14
2. Création de la table des fréquences d'apparition des fragments.	15
3. Création de l'arbre.	15
4. Résultat.	16
5. Remarque.	17
<u>6). Principaux formats de musique compressée.</u>	18
<u>7). Les enjeux de la compression audio.</u>	23

1). Introduction

Longtemps plébiscité par les professionnels du son, le MP3 arrive maintenant sur l'ordinateur personnel.

Faire des compilations de plus de 150 titres tenant sur un seul Cd ou bien de stocker l'équivalent de 12 albums sur ce même Cd? Et bien c'est possible grâce à la compression audio.

Le MP3 (MPeg Audio Layer 3) est un format de fichier son compressé obtenu par suppression de données.

Ce format est proche du format utilisé par Sony avec ses lecteurs/enregistreurs de MiniDisc. Contrairement au DAT, le MP3, le MiniDisc et le DCC opèrent par suppression de données. Saviez vous que les radios utilisent ce format. Tous les tubes sont stockés sur un serveur. La programmation des morceaux à passer à l'antenne se fait sur ordinateur. Puis, l'ordinateur va chercher les fichiers sons sur le serveur et une carte de décompression (gain de rapidité) les transforme en fichiers sons classiques de qualité CD qui seront transmis sur les ondes.

Le codage Mpeg Layer-3 permet de diminuer d'environ 12 fois la taille d'un fichier audio habituel. Le principal intérêt de ce format est d'atteindre un taux de compression très important sans perte de qualité sonore.

L'avantage est donc le gain de taille. Avec une un ratio de compression de 1:12, les fichiers deviennent plus adaptés à des transferts sur le Net. Prenons un exemple : avec un enregistrement audio de 4 minutes en qualité CD, on atteint facilement les 50 Mo pour le fichier Wav. Après encodage au format MP3, le fichier ne dépassera pas les 4 Mo sans dégradation sonore.

La lecture de tels fichiers se fait via l'intermédiaire de softs dédiés qui permettent le décodage (décompression) de ceux-ci.

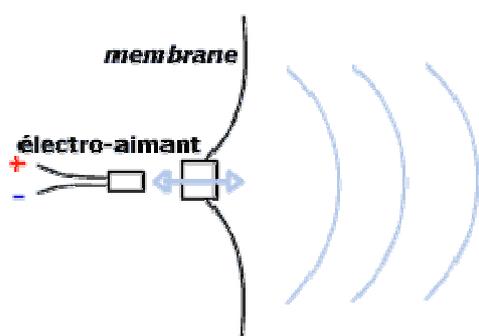
2). Le son

1. Qu'est-ce que le son?

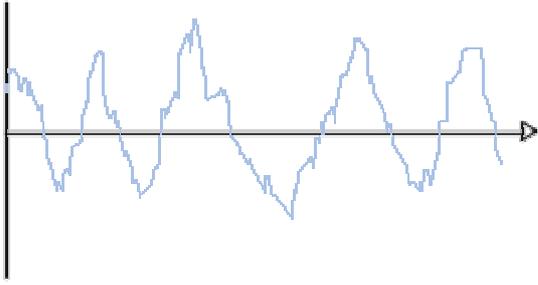
Le son est une vibration de l'air, c'est-à-dire une suite de surpressions et de dépressions de l'air par rapport à une moyenne, qui est la pression atmosphérique. D'ailleurs pour s'en convaincre, il suffit de placer un objet bruyant (un réveil par exemple) dans une cloche à vide pour s'apercevoir que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

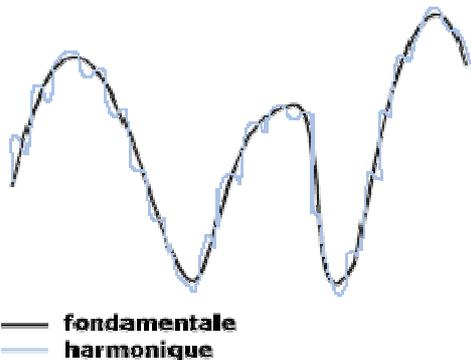
Pour reproduire des sons, on utilise généralement des haut-parleurs. Il s'agit en fait d'une membrane reliée à un électro-aimant, qui, suivant les sollicitations d'un courant électrique va aller en avant et en arrière très rapidement, ce qui provoque une vibration de l'air situé devant lui, c'est-à-dire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



De cette façon on produit des ondes sonores qui peuvent être représentées sur un graphique comme les variations de la pression de l'air (ou bien de l'électricité dans l'électro-aimant) en fonction du temps. On obtient alors une représentation de la forme suivante:



Cette représentation d'un son est appelée spectre de modulation d'amplitude (modulation de l'amplitude d'un son en fonction du temps). Le sonogramme représente par contre la variation des fréquences sonores en fonction du temps. On peut remarquer qu'un sonogramme présente une fréquence fondamentale, à laquelle se superposent des fréquences plus élevées, appelées harmoniques.



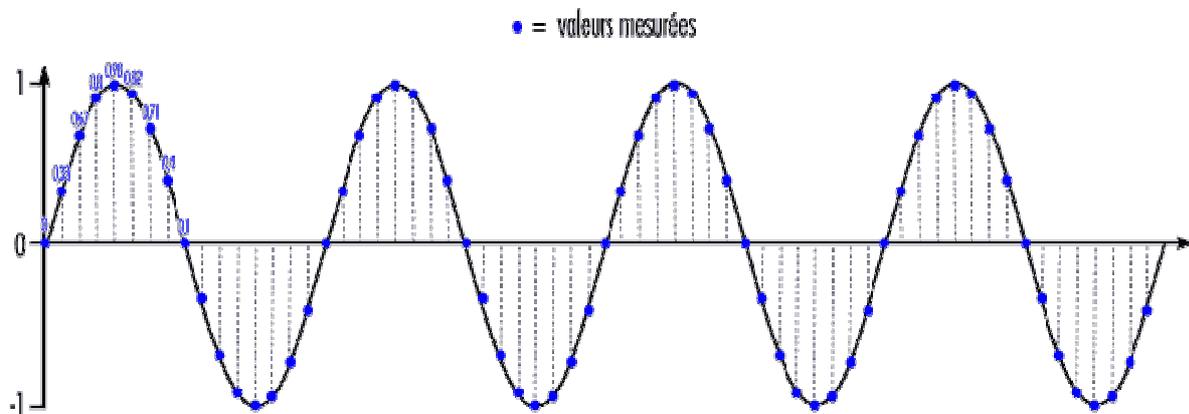
C'est ce qui permet d'arriver à distinguer plusieurs sources sonores: les sons graves auront des fréquences basses, et les sons aigus des fréquences élevées.

2. L'échantillonnage du son.

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.

A chaque échantillon (correspondant à un intervalle de temps) est associé une valeur qui détermine la valeur de la pression de l'air à ce moment, le son n'est

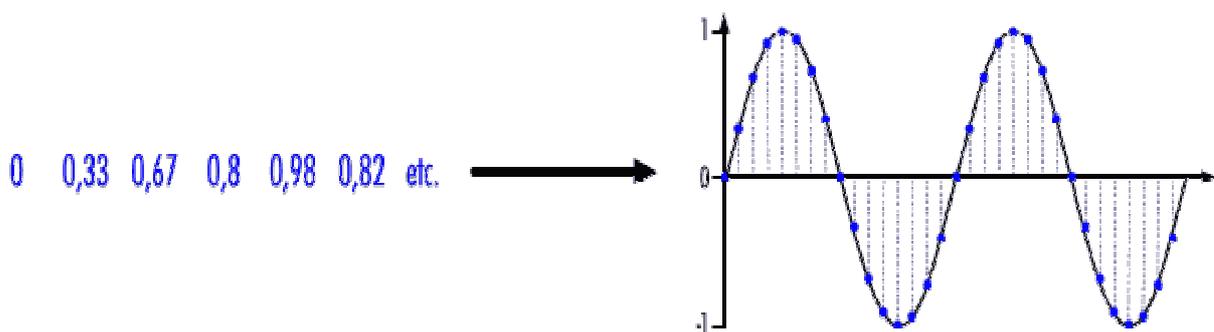
donc plus représenté comme une courbe continue présentant des variations mais comme une suite de valeurs pour chaque intervalle de temps:



Pour numériser (ou "digitaliser") le signal, l'ordinateur en mesure la valeur régulièrement. C'est exactement ce qui se passe quand l'ordinateur enregistre un fichier son : les valeurs mesurées sont enregistrées dans le fichier.

0 0,33 0,67 0,8 0,92 0,82 etc.

Pour rejouer le fichier son, l'ordinateur recrée le signal à partir de chaque valeur enregistrée.



Tous les sons qui nous entourent (musique, bruits...) sont des vibrations que l'on peut échantillonner de cette manière.

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 quelques 100000^{èmes} de seconde, il est plus

pratique de raisonner sur le nombre d'échantillons par seconde, exprimés en Hertz (Hz).

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 20000 Hz, il faut donc une fréquence d'échantillonnage au moins de l'ordre de 40 000 Hz pour obtenir une qualité satisfaisante. Il existe un certain nombre de fréquences d'échantillonnage normalisées :

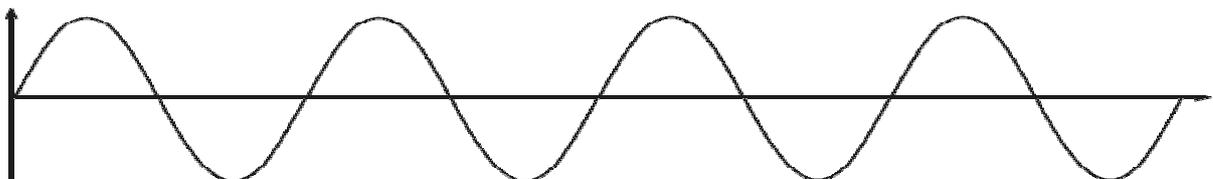
- 32 kHz : pour la radio FM en numérique (bande passante limitée à 15 kHz).
- 44.1 kHz : pour l'audio professionnelle et les compact-disc.
- 48 kHz : pour les enregistreurs numériques multipistes professionnels et l'enregistrement grand public (DAT, Mini-Disc...).

3. Compression du signal.

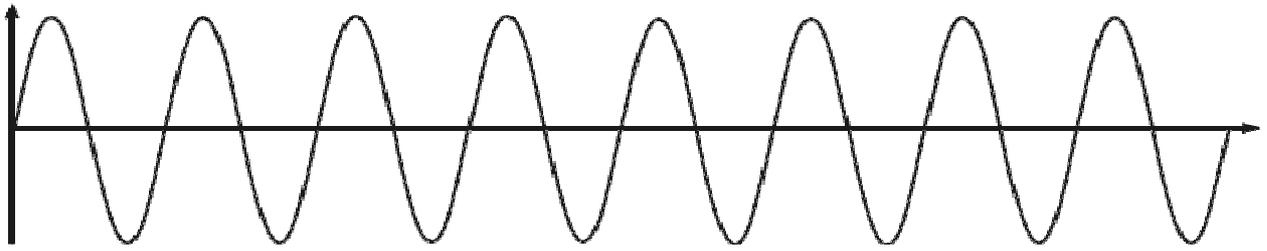
Le problème, c'est qu'il faut enregistrer un très grand nombre de valeurs pour chaque seconde de son. Cela demande beaucoup de ressource système et occupe énormément de place sur le disque dur.

Il a donc fallu trouver des moyens pour gagner de la place. Une des solutions consiste à compresser le signal, c'est à dire d'utiliser moins de données pour mémoriser la même information.

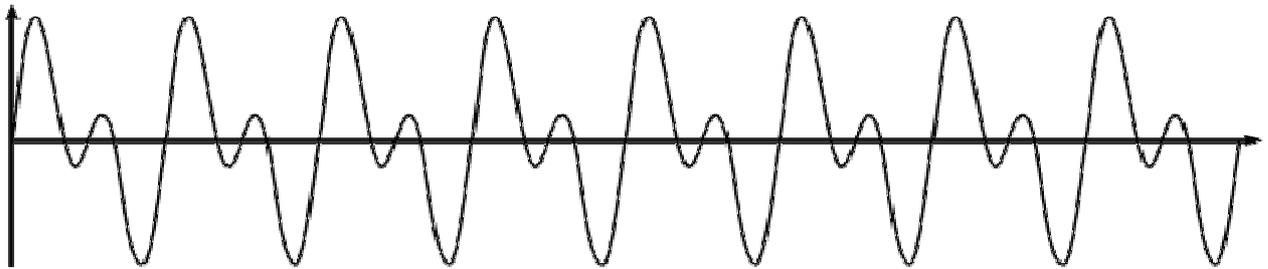
Dans l'exemple suivant, le LA du diapason (celle que vous avez en frappant un diapason ou en décrochant votre téléphone), Le signal se répète de la même manière 440 fois par seconde. On dit qu'il est à 440 Hz.



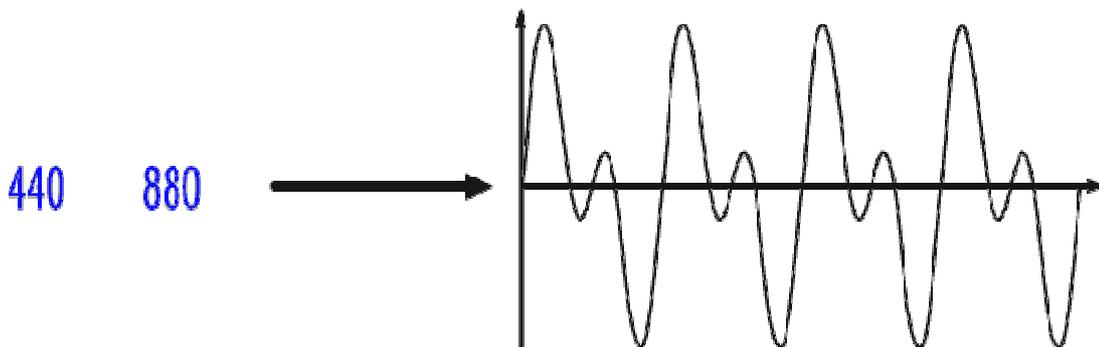
A 880 Hz, la vibration est deux fois plus rapide et le son plus aigu:



On peut mélanger les deux signaux (440 Hz et 880 Hz):



Plutôt que d'enregistrer tous les échantillons (toutes les valeurs) de ce signal, on pourrait n'enregistrer que les valeurs 440 et 880. Ainsi, avec seulement 2 valeurs, on peut reconstituer le signal !



La transformation mathématique qui permet de trouver toutes les fréquences constituant un signal est appelé *transformée de Fourier*.

4. La transformée de fourier.

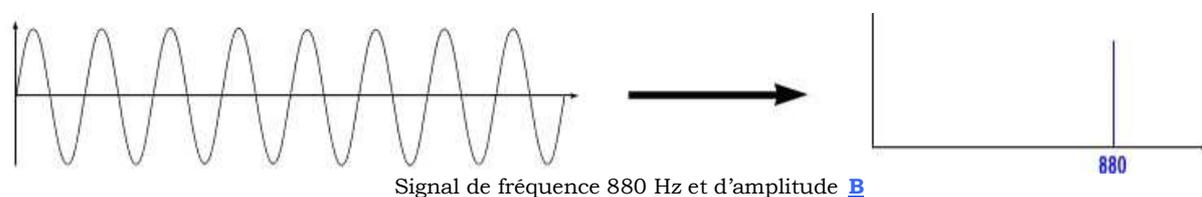
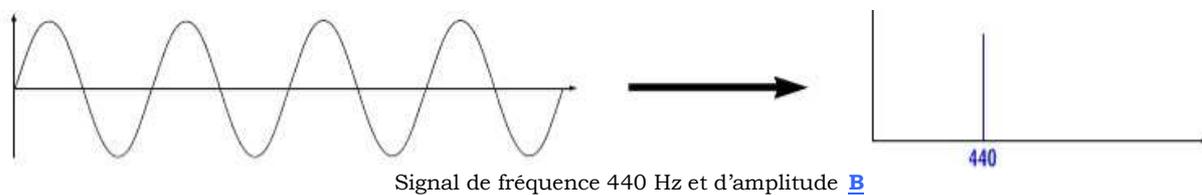
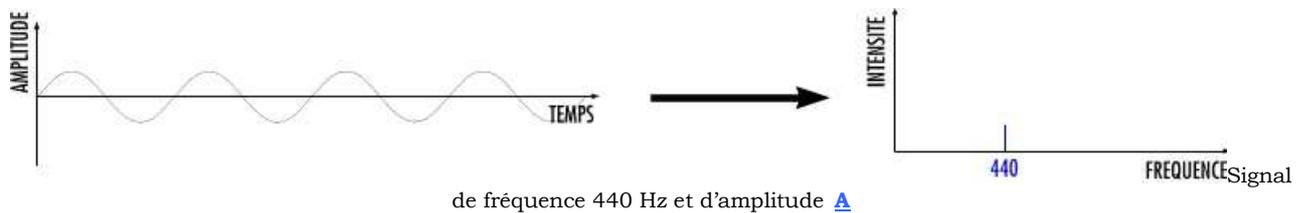
Définition :

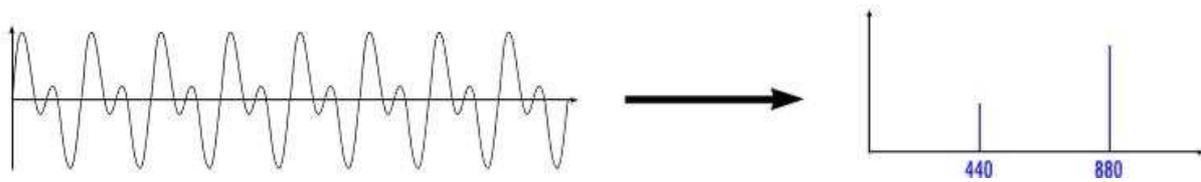
La transformée de Fourier est basée sur la découverte que toute fonction périodique du temps $x(t)$ peut être décomposée en une somme infinie de sinus et cosinus dont les fréquences commencent à zéro et augmentent par multiples entiers d'une fréquence de base $f_0 = 1/T$, où T est la période de $x(t)$.

Cette opération mathématique qui consiste à décomposer une fonction selon ses fréquences. Elle transforme une fonction f qui dépend du temps, en une nouvelle fonction f qui dépend de la fréquence. Cette nouvelle fonction permet de décomposer un signal en différentes fréquences.

Représentation.

Reprenons les exemples précédent avec les fréquences 440 Hz et 880 Hz





Signal de fréquence 440 Hz (d'amplitude **A**) + signal 880 Hz (d'amplitude **B**)

Conclusion.

Une fonction et sa transformée de fourier sont deux aspects de la même information. La fonction temporelle correspondant à un enregistrement de musique montre les variations de la pression de l'air au cours du temps (qui produisent des ondes sonores), mais elle n'indique pas quelles fréquences, quelles notes, composent la musique. La transformée de fourier, en revanche révèle l'information sur les fréquences et cache l'information sur l'évolution temporelle : elle donne, pour notre exemple, les fréquences que la musique comporte, mais il est difficile de savoir quand ces notes sont jouées.

5. Représentation du son.

L'ordinateur travaille avec des bits, il faut donc déterminer le nombre de valeurs que l'échantillon peut prendre, cela revient à fixer le nombre de bits sur lequel on code les valeurs des échantillons.

- Avec un codage sur 8 bits, on a 2^8 possibilités de valeurs, c'est-à-dire 256 valeurs possibles
- Avec un codage sur 16 bits, on a 2^{16} possibilités de valeurs, c'est-à-dire 65536 valeurs possibles

Avec la seconde représentation, on aura bien évidemment une qualité de son bien meilleure, mais aussi un besoin en mémoire beaucoup plus important.

Enfin, la stéréophonie nécessite deux canaux sur lesquels on enregistre individuellement un son qui sera fourni au haut-parleur de gauche, ainsi qu'un son qui sera diffusé sur celui de droite.

Un son est donc représenté (informatiquement) par plusieurs paramètres:

- la fréquence d'échantillonnage.
- le nombre de bits d'un échantillon.
- le nombre de voies (une seule correspond à du mono, deux à de la stéréo, et quatre à de la quadraphonie).

6. Mémoire requise pour stocker un son

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...).

Pour connaître la taille d'une voie, il suffit de connaître le taux d'échantillonnage, qui va nous permettre de savoir le nombre d'échantillons par seconde, donc la taille qu'occupe une seconde de musique. Celle-ci vaut:

Taux d'échantillonnage **X** Nombre de bits.

Ainsi, pour savoir l'espace mémoire que consomme un extrait sonore de plusieurs secondes, il suffit de multiplier la valeur précédente par le nombre de secondes:

Taux d'échantillonnage **X** Nombre de bits **X** Nombre de secondes.

Enfin, la taille finale de l'extrait est à multiplier par le nombre de voies (elle sera alors deux fois plus importante en stéréo qu'en mono...).

La taille en bits d'un extrait sonore est ainsi égale à :

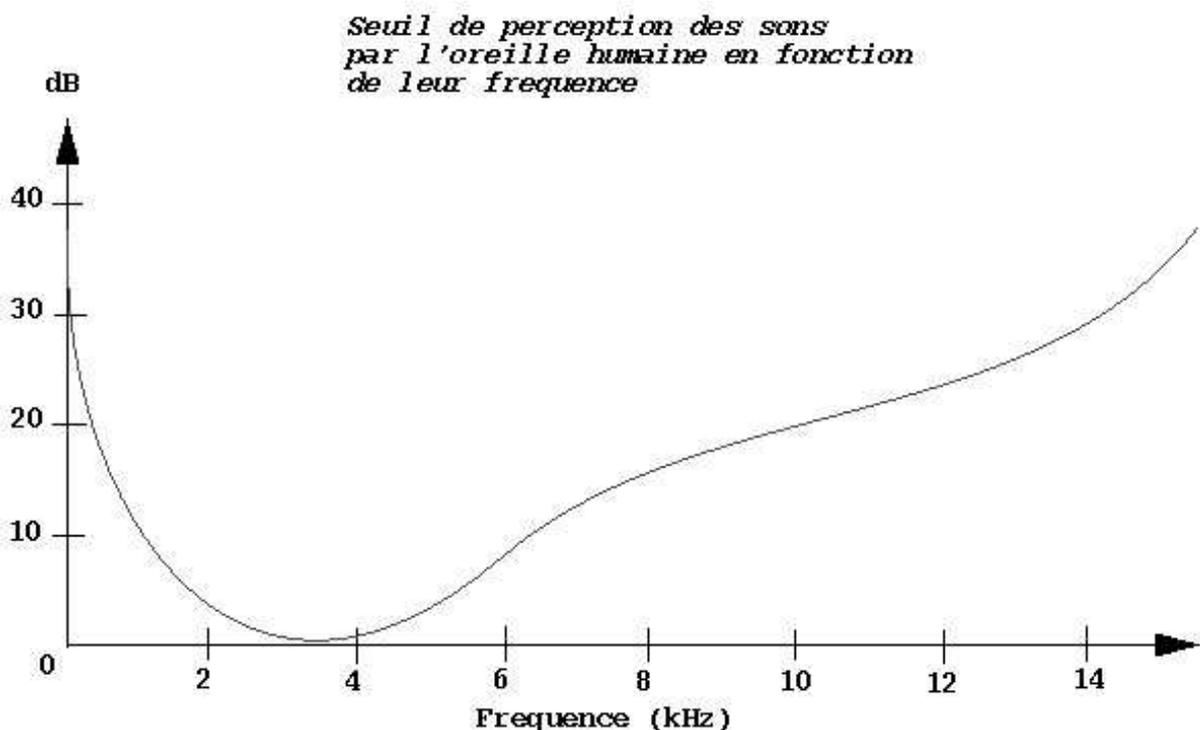
Taux d'échantillonnage X Nombre de bits X Nombre de secondes X Nombre de voies.

3). La compression destructive.

Bien que les compacteurs arrivent déjà à réduire significativement les tailles des fichiers qu'ils compressent, il est encore possible d'améliorer les taux de compression. Cela a d'autant plus d'enjeu qu'une image en haute résolution ou un son dont la définition est de qualité CD prend énormément de place.

L'idée qui préside à cette amélioration vient des limites mêmes de l'homme : un oeil humain n'est pas capable de différencier 2 couleurs très proches parmi les 16 millions affichables par son PC. Pourquoi dans ce cas, ne pas réduire le nombre de couleurs employées par l'image? C'est ce que font les formats de compression JPEG, PING et GIF, et, de ce fait, ils réduisent de manière radicale la taille des images.

Idem pour le son, aucun homme normalement constitué ne peut entendre un son à la fréquence de 20 KHz. Si un morceau contient des fréquences hors de cette gamme, on peut donc purement et simplement les supprimer sans perte de qualité audio puisque l'oreille ne les entend pas. En fait, on entend surtout correctement les fréquences situées dans la gamme 2 kHz 5 kHz. En effet, il faut moins de 5dB pour entendre les fréquences de cette bande alors qu'il faut plus de 20dB pour entendre les fréquences situées en dessous de 100Hz ou au dessus de 10kHz.



Courbe de Fletcher-Munsen

Ces constatations peuvent être exploitées pour réduire la taille des fichiers. On peut par exemple décider que toutes les fréquences au dessus de 15kHz seront supprimées. En supprimant ces informations inaudibles par le commun des mortels, on diminue significativement le nombre de données à prendre en compte et on obtient donc un fichier plus léger.

Poursuivant cette chasse aux informations inutiles, les psycho acousticiens ont établi qu'aucun être humain n'était capable de percevoir la stéréo dans les extrêmes graves. Autant donc passer toutes les fréquences trop basses en mono quitte à rétablir une pseudo stéréo au moment de la lecture du fichier. Et encore quelques octets de gagnés!

Mais le plus futé reste à venir : les mêmes psycho acousticiens ont observé que certaines fréquences en écrasaient d'autres qui, du coup, devenaient superflues.

Pour une nouvelle fois faire simple, imaginez votre groupe de rock : le guitariste avec son Marshal 100 W dont il pousse le volume à 10, le batteur qui enclume comme un forcené et entre les deux, votre petite soeur qui joue du triangle, sans micro pour l'amplifier. Il y a fort à parier que le public n'entende rien de son solo débridé de triangle si les autres ne se décident pas à lui faire un silence royal. Autant, donc, qu'elle ne joue pas du tout.

Fort de ce constat, l'algorithme de compression repère donc les sons "dominants" et retire toutes les données relatives aux sons "dominés". Puisque de toutes façons, on ne les aurait pas entendus, cela ne fait guère de différence au niveau sonore tout en réduisant notablement la taille du fichier.

C'est sur ces principes que reposent de nombreux algorithmes de compression du son tels que le mp3 et ses affiliés, (le WMA, le Real Audio, l'Ogg Vorbis, etc.) : ils allègent le fichier de toutes les informations qu'ils estiment superflues avant de le passer par un algorithme de compression classique.

Parce qu'ils retirent de la matière sonore au fichier traité et dégradent ses qualités d'origine, on dit qu'ils opèrent une compression "destructive". Évidemment, plus on veut un petit fichier, plus la dégradation du son sera grande.

4). La compression non destructive.

Le principe de la compression repose, pour faire simple, sur l'idée qu'un certain nombre de données dans un fichier sont répétitives. Imaginez un fruitier qui fait l'inventaire de ses stocks ; il a deux façons de procéder :

soit il compte fruit par fruit : 1 banane + 1 banane + 1 banane + 1 orange + 1 orange + 1 banane + 1 banane + 1 orange + etc.

Soit il compte par lots : 3 bananes + 2 oranges + 2 bananes + 8 oranges + etc.

La deuxième méthode est nettement moins longue, et prend nettement moins de place. Pour la compression de données, c'est pareil.

Là où le fichier binaire écrit : 111111110000111110000001111

L'algorithme de compression écrit : 8*1,4*0,5*1,6*0,3*1

Le gain de place est évident et peut en outre s'appliquer à tout type de fichiers : les exécutables, les images et bien sûr les sons.

Pour rentrer dans les détails, un son, c'est une fréquence. Une seconde de musique c'est donc une succession de fréquences. Imaginons que dans la suite d'échantillons composant une seconde de musique, nous ayons plusieurs fois de suite la même fréquence, par exemple 10 fois. Si au lieu de stocker ces 10 points, nous en stockions 1 seul et le nombre de fois où il se répète, nous aurions besoin de coder 2 nombres et non 10.

De nombreux compacteurs recourent à l'algorithme de compression de Huffman qui code les données selon leur récurrence statistique. : Plus une séquence de données se répète dans le fichier, plus son code sera court. A l'inverse, les séquences les plus rares sont dotées d'un code plus long.

Cette technique repose sur l'utilisation d'un code de longueur variable et la probabilité d'apparition d'un événement (en l'occurrence ici d'une fréquence). A la décompression, tous les codes sont confrontés à une table de correspondance qui permet de ré attribuer chaque séquence de données au code correspondant. Cette méthode est utilisée dans tous les algorithmes de compression (compression de fichiers texte, compression d'images, compression de sons). C'est la phase finale de la compression. C'est un codage non destructeur.

Evidemment, la démonstration ci-dessus, simplifiée au possible, ne rend pas compte de la complexité des algorithmes de compression mis en œuvre. par des logiciels comme Winzip, WinARJ, WinRAR ou Stuff It. Mais le principe est là, dans les grandes lignes.

5). Le codage de Huffman (1952).

Cet algorithme est couramment utilisé pour la compression, mais son champ d'application ne se restreint pas à ce domaine. Il est aussi utilisé dans la vidéo et dans l'audio. En fait, on lui adjoint la plupart du temps une autre méthode de compression qui, elle, se fait avec perte.

Le principe du codage de Huffman est le suivant : plus un motif est fréquent, plus il sera codé sur un petit nombre de bits. Il offre une méthode pour obtenir la manière de coder les informations à partir de leur probabilité d'apparition.

1. Présentation :

La méthode de compression Huffman consiste à diminuer au maximum le nombre de bits utilisés pour coder un fragment d'information. Cette méthode est une méthode de compression sans perte.

Supposons que notre fichier soit extrêmement simple, et constitué d'une phrase : « Manu fait Huffman ».

Il y a 15 caractères dans ce fichier; chaque caractère étant codé par un octet de 8bits (codage ASCII). Cela signifie donc 15 octets, ou encore 120 bits.

Le fragment d'information sera un caractère. Plus le fragment sera grand, plus les possibilités seront grandes et donc la mise en œuvre complexe à exécuter. L'algorithme de Huffman se base sur la fréquence d'apparition d'un fragment pour le coder : plus un fragment est fréquent, moins on utilisera de bits pour le coder. Dans notre exemple de fichier texte, si on considère que notre fragment est la taille d'un caractère, on peut remarquer qu'ils n'ont pas tous la même fréquence d'apparition : par exemple la lettre 'A' est largement plus fréquente que la lettre 'T' par conséquent la lettre 'A' sera codée sur 2 bits alors que la lettre 'T' en prendra 4.

2. Création de la table des fréquences d'apparition des fragments :

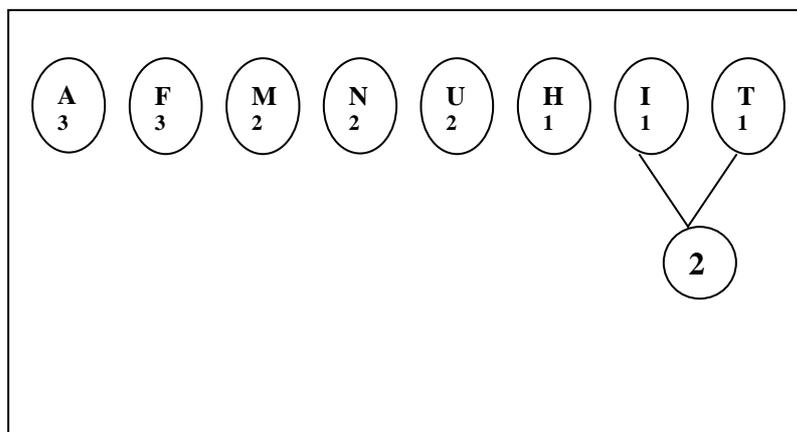
Cette table consiste en un comptage empirique des fragments au sein des données à compresser. Reprenons l'exemple d'un texte : nous allons analyser la phrase : « Manu fait Huffman »

Pour simplifier l'exemple, nous ignorerons la casse et les espaces :

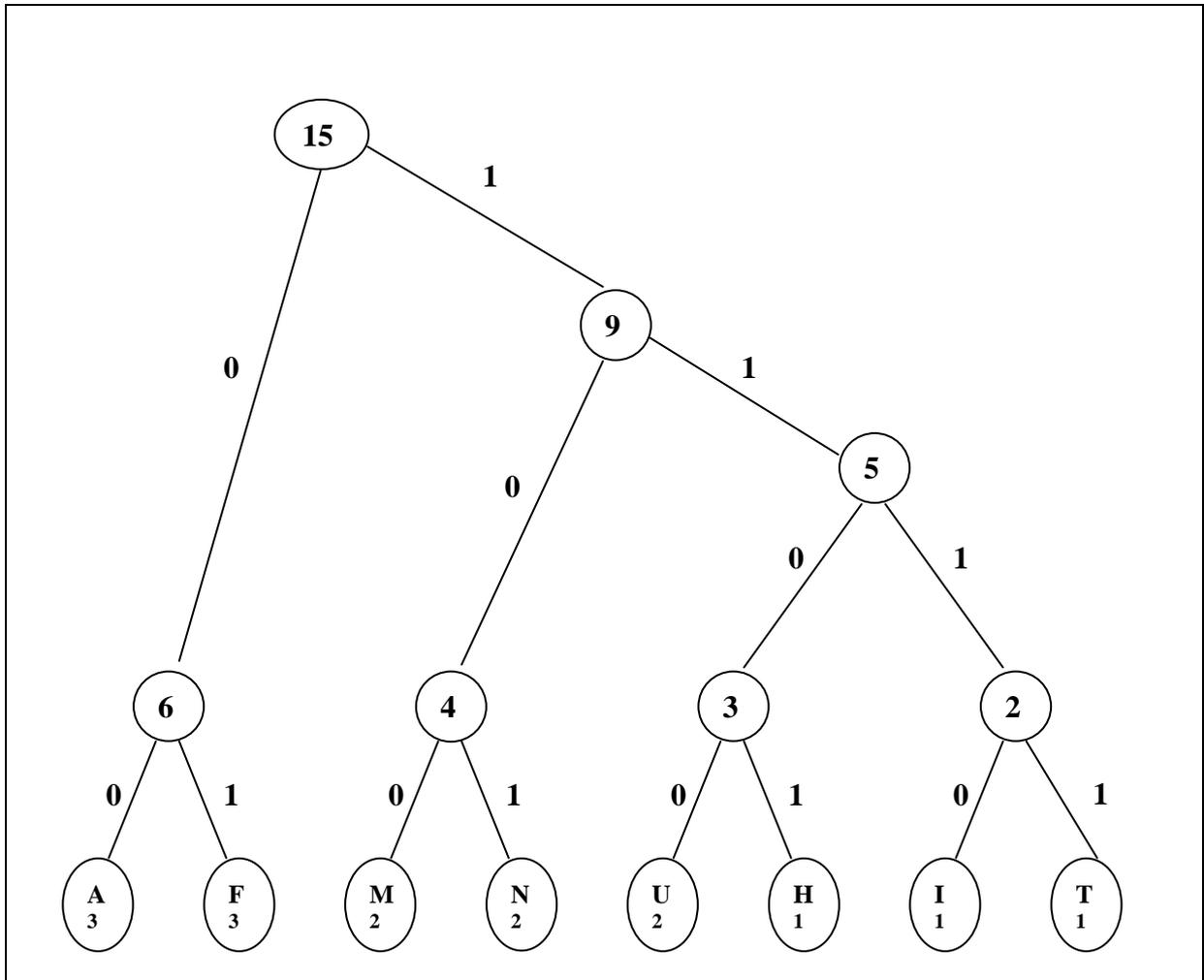
Lettres	Occurrences	Fréquence
A	3	20%
F	3	20%
M	2	13.33%
N	2	13.33%
U	2	13.33%
H	1	6.66%
I	1	6.66%
T	1	6.66%

3. Création de l'arbre.

Nous allons maintenant construire l'arbre à partir de la liste ordonnée de nœuds. La construction est très facile : il suffit de prendre les deux nœuds les moins fréquents (I et T) et de les ajouter comme fils d'un nouveau nœud qui aura pour fréquence la somme des deux, puis continuer en ajoutant toujours les plus petits nœuds.



Il suffit de réitérer cette étape jusqu'à ne plus avoir qu'un seul nœud. Après cela, descendre vers la gauche équivaut à un 0, et descendre vers la droite à un 1.



4. Résultat.

Et voici maintenant, transcrit avec notre nouveau code, la phrase de départ :
10000101110001001110111111011100010110000101

Ce qui nous fait 44 bits, au lieu de 120 au départ! Cela correspond à un taux de compression de 63 %.

Tableau des lettres codées.

Lettres	Occurrences	Code
A	3	00
F	3	01
M	2	100
N	2	101
U	2	1100
H	1	1101
I	1	1110
T	1	1111

Le fait d'avoir généré un code en se servant d'un arbre binaire assure qu'aucun code ne peut être le préfixe d'un autre. Vous pouvez vérifier qu'à l'aide de la table de codage, il n'y a aucune ambiguïté possible pour décoder notre mot compressé! En pratique, la table de codage étant spécifique à chaque fichier, il est indispensable de l'incorporer au fichier compressé, de manière à ce que le décryptage soit possible. Ce qui signifie que la taille du fichier compressé doit être augmentée d'autant. Dans le cas de notre fichier exemple, il faudrait incorporer AU MINIMUM 22 octets de plus pour insérer la table de codage, et le taux de compression n'est plus aussi bon.

Toutefois, pour des fichiers suffisamment larges (à partir de quelques kilooctets) le surplus de taille généré par la table de codage devient négligeable par rapport à l'ensemble du fichier. Concrètement, l'algorithme de Huffman permet d'obtenir des taux de compression typiques compris entre 30% et 60%.

5. Remarque.

On définit le ratio de compression de la manière suivante :

$$P = 1 - \frac{\text{Taille compressée}}{\text{Taille normale}}$$

$$\text{Dans notre cas } P = 1 - \frac{44}{120} = 0.63 \text{ soit } 63 \%$$

6). Principaux formats de musique compressée.

Le plus connu des formats de compression est évidemment ce bon vieux Mp3, mais il est loin d'être le seul au monde. Sentant que la musique numérique compressée avait de grandes chances de bouleverser le marché du disque, de nombreux développeurs se sont lancés dans la conception de formats destinés à offrir le meilleur compromis entre la quantité de données utilisée et la qualité sonore.

Souvent en informatique, le Mp3 doit son statut de standard moins à l'excellence de ses prestations qu'au fait qu'il fut le premier sur le marché. On trouve donc aujourd'hui des codecs (nom donné aux algorithmes de COmpression/DECompression) plus performants et plus riches en termes de fonctionnalités tels le VQF, le MP3 Pro, l'Ogg Vorbis, le Windows Audio Media , le Quick time ou le Real Audio. Par ailleurs, au sein même de la famille MP3, plusieurs algorithmes s'affrontent, émanant chacun de développeurs différents. Ainsi, aux côtés du premier codec mis au point par Fraunhofer, on trouve désormais des concurrents du nom de Lame, Blaze, Gogo ou Radium qui sont tous compatibles entre eux mais qui offrent des différences quant à leur qualité audio ou leur rapidité de traitement... Autant de raisons de se livrer à un petit comparatif. Revue de détail des différents formats.

Le MP3

Star incontestée des codecs, cet algorithme de compression prend naissance dans les laboratoires de Fraunhofer IIS et de l'Université d'Erlangen en 1987. L'ISO en fera un standard dans les années 92-93 (ISO/IEC 11172). Le MP3 est un dérivé de la compression sonore utilisé pour les vidéos au format MPEG-1 (MP3 signifie d'ailleurs MPEG-1 Audio Layer 3).

Il s'agit du format qui offre le panel de paramétrage le plus important. On peut en effet définir un débit ("bitrate" dans le jargon) allant de 32 à 320 kbps. À partir de 128 Kbps, la qualité audio devient suffisante pour encoder des chansons. À 192 Kbps, la qualité est similaire à celle d'un CD audio. À ce taux de compression, une minute de musique équivaut à 1,4 Mo de données. Mais il est également possible d'opter pour un débit variable dit VBR (Variable Bit Rate). Dans ce cas, l'encodeur augmentera le taux de compression sur les passages musicaux moins complexes, ce qui diminuera légèrement la taille du fichier final. Le gain dépend grandement du type de morceau : de 5 à 10% avec un extrait de musique classique ou du jazz, pas grand-chose avec du rock ou du rap. Le gros avantage du MP3 demeure sa compatibilité avec l'ensemble des baladeurs audionumériques, certaines chaînes hi-fi, autoradios et lecteurs de

DVD de salon mais aussi à des applications nécessitant des débits faibles d'où une adhésion très rapide du monde Internet à ce format de compression.

Quasiment universel, il figure sur tous les sites de téléchargement gratuit de type Kazaa ou eDonkey. En revanche, ce format pêche un peu au niveau de la restitution des fréquences aiguës et se montre donc moins adapté à la musique classique, par exemple.

On notera enfin qu'il existe plusieurs algorithmes d'encodage MP3, le "Fraunhofer" étant le plus répandu et le "Lame" le plus efficace. Ses ancêtre peu connu son le mp1 et le mp2. **Le suffixe des fichiers créés est .mp3**

Le MP3 Pro

Avec l'arrivée de formats concurrents et plus efficaces tels le WMA ou l'AAC, les dépositaires du MP3 ont décidé de faire évoluer ce format. D'où l'arrivée du MP3 Pro, sorti à la fin de 2001. Fruit de la collaboration entre Thomson Multimédia et l'Institut Fraunhofer, dans son principe, le MP3 Pro a pour but d'améliorer la principale lacune du MP3 au niveau des fréquences élevées. Ainsi, si le premier supprime purement et simplement les fréquences situées au-delà d'un spectre défini par le taux de compression, le second essaie, autant que faire se peut, de conserver certaines de ces hautes fréquences qui amélioreront la qualité d'écoute. Ce format combine l'algorithme MP3 et un système améliorant la qualité des fichiers compressés appelé SBR pour Spectral Band Replication. Ainsi à un débit de 96 kbps, le MP3 Pro est légèrement plus efficace que le MP3 à 128 kbps, d'où un gain de poids de 25%. Plus le débit sera faible, et plus le gain sera élevé. Et en corollaire, avec un débit supérieur ou égal à 192 kbps, la différence apparaît quasi imperceptible. La qualité CD en MP3 Pro est obtenue à 96 kbps, soit 720 ko pour une minute. Toutefois, et c'est là que le bât blesse et nuit au développement de ce format, il n'existe pas d'encodeur gratuit supportant le MP3 Pro. Signalons que les fichiers MP3 Pro peuvent en théorie être lus par des baladeurs ne supportant que le MP3 mais dans la pratique la qualité sonore obtenue est souvent sujette à caution. **Le suffixe des fichiers créés est .mp3**

Le WMA

Lancé en 1999 par Microsoft, le WMA (*Windows Media Audio*) n'a eu pour but que de contrer l'essor du MP3. Avec des arguments convaincants. Exploitant plus efficacement les caractéristiques de l'oreille humaine au niveau du spectre audible, le WMA parvient à supprimer les fréquences réellement inutiles tout en conservant certaines hautes fréquences qui ont une influence sur la qualité sonore. L'algorithme de compression, relativement efficace, permet de conserver une qualité équivalente à un CD audio avec un débit de 128 kbps, soit 1 Mo pour une minute de chanson.

Totalement gratuit, le WMA a en plus l'avantage de disposer d'un encodeur

intégré au lecteur Windows Media Player. Une façon efficace d'assurer la diffusion de ce support. D'autant que ce format est lié à une gestion pointue des droits d'auteur (DRM ou *Digital Right Management*) qui permet de définir par exemple une durée de vie limitée pour les fichiers ou d'interdire les possibilités de gravure. C'est pourquoi il est utilisé par certains sites de musique en téléchargement. De nombreux baladeurs audionumériques supportent également le WMA et des sharewares réputés tels WinAmp permettent d'encoder dans ce format. **Le suffixe des fichiers créés est .wma**

L'AAC

Contrairement aux idées reçues, l'AAC (*Advanced Audio Coding*) n'est pas un format développé par Apple mais par un consortium au sein duquel se retrouvent l'institut Fraunhofer (le père du MP3), AT&T, Nokia, Sony, ou encore Dolby. Connue également sous le nom de MPEG-2 NBC pour Non Backward Compatible, il est l'un des descendants du MPEG-2 successeur lui-même du MPEG-1. Effectivement, l'AAC est sans aucun doute l'algorithme de compression le plus efficace. Contrairement au MP3 et au WMA, il ne s'appuie pas sur le MPEG-1 mais sur le MPEG-4 (format à l'origine du DivX). Ce choix semble avoir porté ses fruits puisque cela lui permet de disposer d'un meilleur compromis entre le taux de compression et la qualité sonore. Ainsi en AAC, on obtient l'équivalent d'une qualité CD audio avec un débit de seulement 96 kbps. À cela s'ajoute, comme pour le WMA, des propriétés de gestion des droits d'auteur (DRM) et la possibilité de gérer des sons sur 48 canaux différents, ce qui le rend apte à encoder des DVD audio ou vidéo en conservant la spatialisation sonore d'origine (les six canaux du Dolby Digital par exemple), bien que cette possibilité ne soit pour l'heure pas vraiment exploitée, par rapport aux autres formats et compte tenu de sa grande qualité à taux de compression équivalent, nous avons peut-être là le successeur du MP3 mais la durée d'encodage est importante.

Le suffixe des fichiers créés est .aac

L'OGG Vorbis

L'OGG Vorbis est au format de compression audio ce que Linux est aux systèmes d'exploitation. Il s'agit en effet d'un format "ouvert" dont les codes source sont publics et peuvent être adaptés et modifiés par tout un chacun. Il a été mis au point par des programmeurs indépendants sous l'égide de la fondation Xiph.org. Il n'en reste pas moins un format assez efficace. À 128 kbps, on atteint le niveau de qualité d'un CD Audio (1 Mo pour 1 minute). La structure de compression du format OGG est par ailleurs sensiblement différente des MP3, WMA et autres AAC. Il segmente les sources audio en paquets successifs, l'algorithme de compression agissant dans un premier temps sur chaque paquet indépendamment des autres. Cela lui permet de ne pas avoir vraiment de faiblesse sur certaines fréquences et de conserver la même qualité quel que soit le type de musique.

À cela s'ajoutent des fonctions de polyphonie permettant de restituer jusqu'à 255 canaux son. La structure en paquet le rend en plus bien adapté à une utilisation en diffusion continue (*streaming*) sur l'Internet, notamment pour les radios en ligne, autre avantage, ce codec peut être intégré dans les applications sans que le programmeur reverse des royalties aux concepteurs. Ouvert, le OGG Vorbis devrait rapidement évoluer et peut-être se rapprocher de l'AAC en termes de débit. Signalons par ailleurs que la compatibilité descendante est assurée, un fichier OGG Vorbis quelle que soit la version de son encodeur peut être lu par un lecteur plus ancien. De nombreux encodeurs et lecteurs sont compatibles avec l'OGG. Inconvénient, la durée d'encodage est importante. **Le suffixe des fichiers créés est .ogg**

Le VQF

*La norme vqf ou TwinVQ (Transform-domain Weighted Interleave Vector Quantization) est assez récente, elle a été développée par NTT Cyber Space Laboratories et soutenu par Yamaha. On regrettera une durée d'encodage un peu trop longue. Le vqf est réputé pour être meilleur que le mp3, Il fournit une meilleure qualité sonore que le format MP3. En vqf la "qualité CD" est atteinte à 96 Kbits/s ce qui fait que pour une qualité égale à celle du mp3, un fichier vqf prend moins de place, environ 30 % de moins. **Le suffixe des fichiers créés est .vqf,vql ou .vqe.***

Le MPC

Le MPC est un format destructeur de compression audio basé sur le format vidéo Mpeg2. Il est aussi appelé Mp+, c'est pour cela qu'il est également appelé Mpeg Plus ou MP+. C'est un format destructeur qui permet de créer des fichiers audio d'excellente qualité (qualité bien supérieure au MP3 ou Ogg Vorbis). Ses caractéristiques sont une excellente qualité d'écoute, il est très rapide à l'encodage et est gratuit (mais l'encodeur n'est pas Open Source). Ce format n'est intéressant que pour de hauts et très hauts débits (à partir de 180 kbits/s). En raison de l'excellente qualité sonore qu'il permet, la taille des fichiers MPC est supérieure à celle de fichiers encodés en MP3 ou Ogg.

En bitrate de 170kbps environ : très grande transparence (99% des utilisateurs en sont satisfait).

En bitrate de 200kbps environ : Difficile de voir la différence avec l'originale.

EN bitrate de 240kbps environ : Transparence absolue.

Ce format est de plus en plus répandu. Sa qualité d'écoute a su engendrer son succès, et il a su trouver sa place dans les p2p. **Le suffixe des fichiers créés est .mpc**

L'AIF (Audio Interchange File)

Format de fichier audio et ou vidéo contenant du son échantillonné dans un

format cousin d'IFF. D'abord sur l'Amiga, puis chez Apple. Il est l'équivalent du format Wav dans le monde Macintosh. Ce format offre une bonne qualité sonore, il est pris en charge par plusieurs navigateur Internet, et ne requiert pas de plugin. Toute fois la taille importante des fichiers limite sérieusement son utilisation. **Le suffixe des fichiers créés est .aif**

Le Real

Le format Real permet la diffusion de fichiers multimédia sur le net en streaming, c'est à dire sans le chargement préalable du fichier et avec une diffusion en continu. Ce format s'adapte aussi bien à la compression de musiques que d'extraits vidéo. Une méthode bien plus rapide que le téléchargement même si la qualité des musiques subit parfois quelques dégradations. Le lecteur Real player est gratuit et existe sur les plates-formes Windows, MacOS et Unix. . **Le suffixe des fichiers créés est .rmvb** (Real Media Variable Bitrate)

Conclusion.

Tous ces algorithmes de compression fonctionnent sur le même principe : éliminer les fréquences que l'oreille humaine ne perçoit pas du tout, ou très peu, et niveler les fréquences proches les unes des autres à des valeurs identiques de façon à accroître la compression. Ce sont des algorithmes dits "destructifs". Autrement dit, la conversion d'un fichier audio compressé à partir d'un CD audio à nouveau en format CD audio, les données résultantes ne seront pas les mêmes que celles du fichier original. Selon les formats et les besoins, la compression pourra être plus ou moins importante. Elle est définie par ce que l'on appelle le débit, c'est-à-dire le flux d'informations par seconde de musique. Il peut osciller de 32 kbps (kilobits par seconde) à 320 kbps.

7). Les enjeux de la compression audio.

La compression sert, comme nous l'avons vu, à diminuer la taille des fichiers, ce qui permet de stocker plus de données sur un support amovible ou de masse (Disquette, Zip, CD R, DVD ou Disque Dur). En recourant à la compression destructive, il est ainsi possible, par exemple, de faire tenir la discographie complète de Brel sur un simple CD et de régler du même coup bien des problèmes de rangement. Un algorithme non destructif sera quant à lui idéal pour archiver les créations musicales ou pour les porter au mastering...

Mais au-delà de ces considérations basement matérielles, le progrès des algorithmes de compression ouvre aussi le champ à bien des applications. Dans le contexte d'une généralisation des connexions à haut débit (Câble et ADSL), la mise au point du Streaming a notamment permis l'émergence de nouveaux médias : télévision et radio peuvent maintenant passer par le Net.

Si les contenus vidéo souffrent encore d'une bande passante trop étroite pour qu'Internet puisse concurrencer les réseaux hertziens et numériques, l'audio est quant à lui parfaitement au point. La qualité sonore d'une radio Internet est en effet supérieure à celle de la bande FM et la mise en place d'une station est d'une relative facilité technique.

De nombreux groupes ont d'ailleurs entrepris de diffuser certains morceaux en mp3, des plus obscurs aux plus connus (Smashing Pumpkins, Black Crowes, etc.). Cela peut se faire par le biais d'une Web Radio bien sûr, mais cela passe plus souvent par le biais de sites Web dédiés au groupe ou à ce genre d'initiative.

Sur des sites tels que Vitaminic ou mp3.com, des centaines d'artistes présentent ainsi leurs oeuvres qu'on peut écouter en ligne ou télécharger gratuitement. Preuve que ce mode de diffusion est efficace, l'engouement autour de certains titres a amené des maisons de disques à signer des artistes jusqu'ici cantonnés au rang des amateurs : Laurie, les Mules, etc.

Bien décidée à prendre sa part du gâteau et à imposer sa loi face à cette insolente gratuité, l'industrie du disque envisage d'ailleurs d'utiliser la diffusion via Internet comme nouveau support de vente. A termes, l'objectif est de faire disparaître le téléchargement (et peut-être même les disques) au profit d'un système d'abonnements (un forfait mensuel vous donne accès à l'écoute du catalogue de la maison de disque) ou d'un système "pay-per-hear" : Vous payez une somme dérisoire chaque fois que vous voulez écouter un morceau mais devrez repayer autant de fois que vous voudrez l'écouter...

Bien entendu, tout cela est encore en chantier et le disque a encore de beaux jours devant lui. En outre, il n'est pas sûr qu'il soit si facile pour les maisons de disques d'avoir raison de l'esprit matérialiste propre au genre humain : acheter

un disque, posséder l'objet, est aussi un plaisir. J'en veux pour preuve les collectionneurs qui continuent de dépenser des fortunes chez les disquaires alors qu'ils n'auraient pas assez de 3 vies pour écouter les 30000 CD de leur discothèque.

Loin de ce marketing fiction, une autre application du streaming retiendra toute l'attention des musiciens : le Jam via Internet. En effet, à l'image du procédé de Rocket Network implémenté dans Cubase, il est ainsi possible de faire de la musique simultanément avec des gens situés aux quatre coins de la planète. Voilà qui simplifie les échanges et ouvre de nouveaux horizons en termes de rencontres musicales.

Bref, vous le comprenez : Là où les médias généralistes ne voient dans la compression qu'un moyen de "pirater" la musique commerciale, il faut voir une véritable révolution en marche pour les musiciens, le public et l'industrie toute entière. Alors à vos encodeurs!