



S.Ayrinhac (v7. 01/03/2013)

1. Introduction à Audacity

1.1. Audacity est un logiciel libre et gratuit, multiplateformes, facile d'utilisation, qui va nous permettre d'illustrer des notions de traitement du signal appliquées aux sons audibles. Ces notions peuvent être profitables dans d'autres thématiques (traitement des images, géophysique, etc) car le traitement du signal est une discipline transverse. Ce document fait suite à un cours où ont été abordées les notions fondamentales : échantillonnage, théorème de Shannon, quantification, compression.

Audacity a été créé en 1999 par Dominic Mazzoni, et est actuellement développé par une équipe de bénévoles¹.

Le logiciel est disponible à l'adresse <http://audacity.sourceforge.net/>

La version la plus récente d'Audacity est la 2.0. Un manuel et des tutoriaux peuvent être téléchargés à l'adresse :

http://audacity.sourceforge.net/manual-fr-1.2/FR_index.html

L'export de MP3 nécessite l'encodeur LAME. Il faut télécharger et installer un petit add-on à la page : <http://lame1.buanzo.com.ar/> ; puis lancer l'exécutable C:\Program Files\Lame for Audacity\Lame_v3.99.3_for_Windows.exe

1.2. Commandes utiles

- sélectionner une piste à partir du curseur : Maj + avance rapide ;
- sélectionner toutes les pistes : Ctrl + A ;
- copier un bout de piste : Ctrl + C ;
- coller un bout de piste : Ctrl + V ;
- annuler une opération : Ctrl + Z ;
- supprimer une sélection : Suppr
- dézoomer l'échelle verticale : Maj

¹ Voir le livre *Audacity 2, Enregistrez, montez, mixez*, éditions Pearson, Christian Brochec, 2012, ISBN 978-2-7440-9402-6.

2. Plan (correspond à environ 6h de travaux dirigés/travaux pratiques)

1. Introduction à Audacity.

2. Chronogramme.

- 2.1. Lecture, enregistrement et génération d'un son.
- 2.2. Analyse d'un spectrogramme.
- 2.3. Bruit de quantification : quantification uniforme et non-uniforme.

3. Transformée de Fourier (TF)

3.1. Rappels sur la TF : définition mathématique, TF inverse, calcul de la TF d'un créneau, TF de quelques fonctions usuelles.

3.2. Série de Fourier.

3.3. TF discrète (TFD) et TF rapide (FFT).

Visualiser un spectre dans Audacity :

menu Analyse > Tracer le spectre.

3.4. Résolution spectrale et battement.

3.5. Technique du bourrage de zéros (zero-padding).

3.6. Le fenêtrage

3.7. Filtrage et convolution

4. Spectrogramme / sonagramme

4.1. Définition d'un spectrogramme.

Le spectrogramme dans Audacity, et options de visualisation :

menu Edition > Préférences... > Spectrogrammes.

4.2. Repliement du spectre : fréquence instantanée, *chirp* linéaire.

4.3. Le sur-échantillonnage et le sous-échantillonnage (décimation).

4.4. Aller plus loin : transformée en ondelettes (scalogramme).

5. Compression

5.1. Comparaison entre un .wav et un .mp3 du même morceau.

Ce cours propose plusieurs types d'activités (désignés par des couleurs différentes) :

**Cours / rappels
théoriques**

Activité / TP

**Fonctionnement
d'Audacity**

Il s'appuie sur un certain nombre de documents *multimédia* :

- écoute d'un son ◀
- vidéo explicative ▶▶
- liens vers des pages web **www**

2. Chronogramme (forme d'onde, *waveform*)

• 2.1. Importer un son

Activité (2.1.1) Lecture

L'import d'un son se fait via le menu : Fichier > Importer > Audio
 Les durées peuvent être mesurées en cochant en bas la case Durée, et en sélectionnant une unité adaptée (hh:mm:ss + millisecondes par exemple) dans le menu déroulant du chronomètre.
 Le signal sonore peut être finalement stocké sous la forme d'un fichier d'extension .wav.

Activité (2.1.2) Enregistrement

Une fois le microphone branché, vous pouvez commencer l'enregistrement en allant dans le menu en haut > Transport > Enregistrement. Pour manipuler des fichiers de taille raisonnable, veillez à enregistrer le son en mono, 16 bits, 44100 Hz, en ne dépassant pas les 10 s (grâce au panneau qui se trouve à gauche du chronogramme). Le signal sonore peut être finalement stocké sous la forme d'un fichier d'extension .wav. Pour cela, allez dans le menu en haut à gauche > Fichier > Exporter > WAV signé 16 bits PCM

Le VU-mètre indique en vert clair le niveau instantané, en vert foncé le niveau maximum atteint et en rouge une saturation éventuelle (lorsque l'amplitude dépasse 1). Dans ce cas, il faut recommencer l'enregistrement avec un niveau plus faible.

Il est pratique d'enregistrer sur un déclencheur de niveau. Cette possibilité est activée via le menu Transport > Enregistrement automatique. Le réglage du niveau est accessible par Transport > Niveau de l'enregistrement automatique.

Activité (2.1.3) Génération d'un son

Menu Pistes > Ajouter nouvelle > Piste mono

Menu Générer > son...

• 2.2. Analyse d'un chronogramme

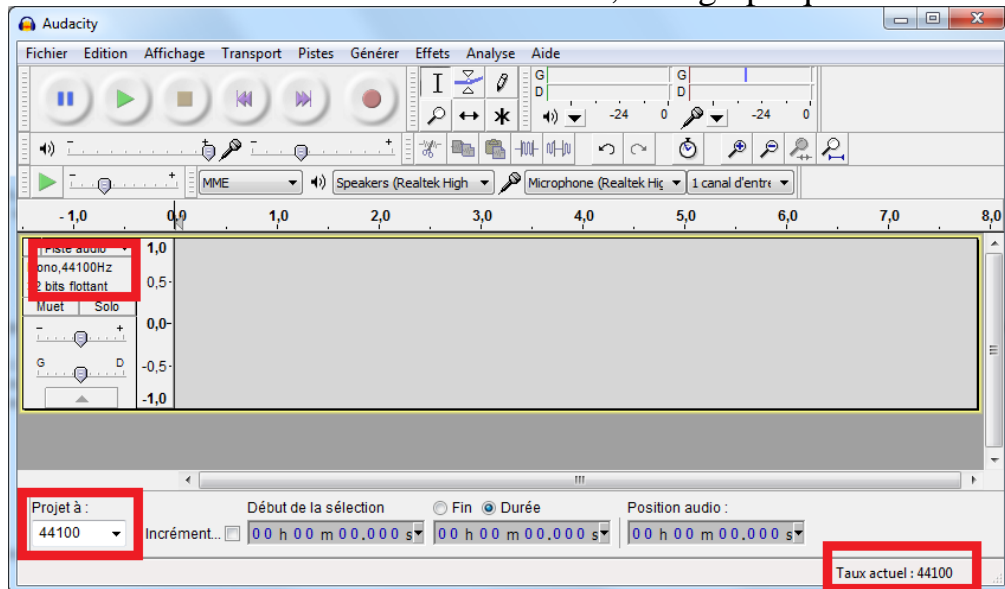
Activité (2.2.1) Chronogramme

Lecture des différents paramètres pertinents :

- durée, nombre d'échantillons : ces grandeurs sont visibles en cochant en bas la case Durée, et en sélectionnant une unité adaptée (hh:mm:ss + millisecondes par exemple) dans le menu déroulant du chronomètre.
- fréquence d'échantillonnage F_e
- nombre de pistes (mono/stéréo)
- quantification (16 bits pour le .wav mais 32 bits pour les fichiers d'Audacity extension .aup).

Note 1 : En général, le signal est un tableau de valeurs et il faut reconstruire le vecteur temps : $t=(0:1:length(y0)-1)/F_e$; (syntaxe Scilab)

Note 2 : La F_e se retrouve 3 fois dans la fenêtre, voir graphique ci-dessous.



La case « Projet à » permet de modifier la F_e de l'ensemble du projet.

Activité (2.2.3) Interférences destructives²

Lorsqu'on écoute en même temps une piste et son inverse, il y a silence. En effet, il y a interférences destructives, car un signal ajouté au même signal en

² Voir l'article de J.Grope, The Physics Teacher, Vol. 49, February 2011

opposition de phase s'annule. Cette idée est à la base du **contrôle actif du bruit**, qui permet une réduction du bruit via l'envoi du même signal mais en opposition de phase.

Activité (2.2.4) Phase

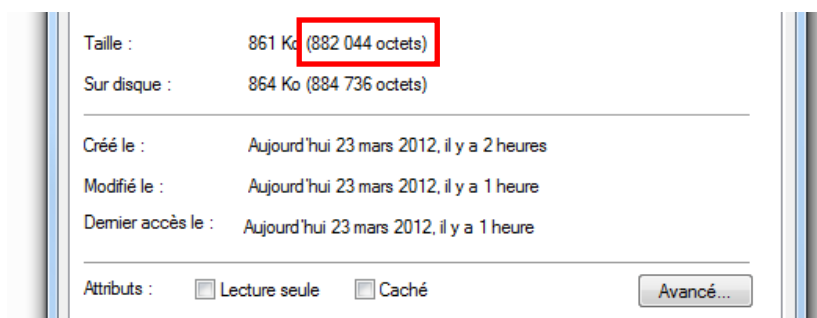
Un signal complexe inversé donne le même son que le signal de départ (Effets > Inverser). L'oreille n'est donc peu sensible à la phase (le fait de commencer par une crête ou un creux). Elle est plus sensible au contenu fréquentiel. Autre exemple avec les deux sons suivants : bien que les chronogrammes soient très différents, les deux sons se révèlent semblables à l'écoute³.

- ❖ http://www.impmc.jussieu.fr/~ayrinhac/sons/somme_de_sinus.wav
- ❖ http://www.impmc.jussieu.fr/~ayrinhac/sons/somme_de_cosinus.wav

Activité (2.2.5) Taille du fichier numérique

La taille d'un fichier audio non compressé est donnée par : $L = FTQP$ où F est la Fe, T le temps d'enregistrement, Q la quantification et P le nombre de pistes (mono, audio, etc). Il faut diviser la taille par 8 pour convertir en octets. Exemple, un CD audio stéréo d'une heure pèse : $44100 \times 16 \times 2 \times 3600 = 635 \text{ Mo}$.

Remarques : les fichiers binaires (les fichiers qui ne contiennent pas du texte) possèdent un en-tête qui décrit le format et donne des informations sur le fichier. Par exemple, un fichier .wav comporte 44 octets d'en-tête. Sous Windows, il y a une différence entre « taille » et « taille sur le disque ». Cette différence provient de la taille des clusters (bloc mémoire ou secteur, en français) occupés par le fichier. Les Ko sont en fait des Kio, c'est-à-dire qu'un Kio=1024 octets et non 1000.



2.3. Bruit de quantification (bdq)

- Rappels : l'amplitude est codée sur b bits, et l'amplitude est découpée en L

³ Voir page 66 du cours de Philippe Guillaume (<http://www-gmm.insa-toulouse.fr/~guillaume/sons-musique.pdf>)

intervalles égaux (quantification uniforme). →FICHE

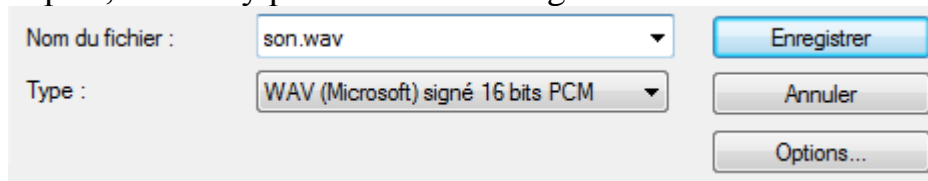
- Quel est le lien mathématique entre L et b ? Réponse : $L=2^b$

Activité (2.3.1) Quantification uniforme et non uniforme

Lors de l'export en .WAV, il est indiqué PCM. Cet acronyme signifie « pulse code modulation ». Ce n'est pas une quantification uniforme. L'amplitude est découpée en L intervalles non égaux. Au contraire, le LPCM pour « linearized pulse code modulation » utilise une quantification uniforme.

Activité (2.3.2) Nombre signé

Lors de l'export, Audacity précise « 16 bits signé » :



En informatique, un entier signé signifie que le premier bit code le signe plus ou moins. Si le codage est effectué sur 16 bits signés, quelle est la plage de valeurs prise ? Réponse : -32768 à 32767 avec un centre à 0 , soit 2^{15} au lieu de 2^{16} car 1 bit est dédié au codage du signe. La valeur 0 apparaît 2 fois.

Activité (2.3.3) Quantification

Écoutons un son avec une quantification qui augmente. Ces sons ont été créés grâce à Scilab (<http://www.scilab.org/>) à partir d'un morceau de Yann Tiersen issu de la bande originale du film Amélie Poulain. Le son original est le suivant :

⚡ <http://www.impmc.jussieu.fr/~ayrinhac/sons/piano-original.wav>

Les sons quantifiés sont les suivants :

⚡ <http://www.impmc.jussieu.fr/~ayrinhac/sons/piano-12bits.wav>

⚡ <http://www.impmc.jussieu.fr/~ayrinhac/sons/piano-10bits.wav>

⚡ <http://www.impmc.jussieu.fr/~ayrinhac/sons/piano-8bits.wav>

⚡ <http://www.impmc.jussieu.fr/~ayrinhac/sons/piano-7bits.wav>

⚡ <http://www.impmc.jussieu.fr/~ayrinhac/sons/piano-6bits.wav>

⚡ <http://www.impmc.jussieu.fr/~ayrinhac/sons/piano-5bits.wav>

⚡ <http://www.impmc.jussieu.fr/~ayrinhac/sons/piano-4bits.wav>

⚡ <http://www.impmc.jussieu.fr/~ayrinhac/sons/piano-3bits.wav>

⚡ <http://www.impmc.jussieu.fr/~ayrinhac/sons/piano-2bits.wav>

Quelle est la quantification maximale ? Réponse : le signal ne contient que 0 ou 1 , soit un codage sur 1 bit (non signé). Écoutons un morceau ainsi quantifié :

⚡ <http://www.impmc.jussieu.fr/~ayrinhac/sons/piano-1bit.wav>

Activité (2.3.4) Quantification des coefficients de Fourier

La quantification employée dans le format .mp3 est une quantification variable

des coefficients de Fourier basée sur la psychoacoustique. Écoutons :

❏ <http://www.impmc.jussieu.fr/~ayrinhac/sons/piano-Fourier2^8+1intervalles.wav>

❏ <http://www.impmc.jussieu.fr/~ayrinhac/sons/piano-Fourier2^4+1intervalles.wav>

Activité (2.3.5) Bruit de quantification

On peut visualiser le bruit de quantification en soustrayant le son quantifié au son originel. Il faut inverser une des deux pistes (Effets > Inverser) et les sommer (Pistes > Mixage et rendu). Sélectionner la totalité d'une piste : touche SHIFT + bouton saut à la fin. Le bruit de quantification a les caractéristiques d'un bruit blanc lorsque la quantification est uniforme (ici ce n'est pas forcément vrai car les .WAV sauvés sont en PCM et non en LPCM).

3. Transformée de Fourier (TF) et analyse spectrale

3.1. Rappels sur la transformée de Fourier

- Définitions mathématiques de la TF et de la TF inverse. →FICHE
- Le spectre est généralement le module de la TF complexe (il y a donc un spectre réel, imaginaire, de phase et d'amplitude ou module). L'énergie est l'amplitude au carré. Le spectre de phase peut être discontinu, c'est-à-dire rester dans l'intervalle $[-\pi, \pi]$ (*wrapped phase*) ou continu (*unwrapped phase*). La phase continue⁴ est la reconstruction de la vraie phase à partir de ces valeurs modulo 2π . Ce phénomène a de nombreuses applications, tel que le radar, l'imagerie à résonance magnétique et l'optique adaptative.
- Lien et différence avec les coefficients de Fourier (notion de série de Fourier): lorsqu'on périodise en temps, on échantillonne en fréquence.
- Exercice : calcul de la TF d'un créneau. →FICHE
 - ▶ <http://www.impmc.jussieu.fr/~ayrinhac/videos/TFporte.mpg>
- Les TFs de quelques fonctions usuelles⁵ :
 - <http://www.impmc.jussieu.fr/~ayrinhac/images/TFpage1.gif>
 - <http://www.impmc.jussieu.fr/~ayrinhac/images/TFpage2.gif>
 - <http://www.impmc.jussieu.fr/~ayrinhac/images/TFpage3.gif>

Activité (3.1.1) Transformée de Fourier d'un sinus cardinal

On peut vérifier que la TF d'un sinus cardinal est une fonction porte.

❏ <http://www.impmc.jussieu.fr/~ayrinhac/sons/sinc.wav>

⁴ Voir <https://pantherfile.uwm.edu/leiyang/public/PhaseUnwrapEncyclopedia.pdf>

⁵ Pages 26-28 du cours de P.Brault (braultp.free.fr/Cours_signal/signal_6juin_sscmp.pdf)

3.2. Rappels sur la série de Fourier

- Définitions mathématiques de la série de Fourier →FICHE

Activité (3.2.1) Motifs périodiques et spectre de raies

En zoomant sur une partie d'un signal, on distingue des motifs périodiques. Ces motifs périodiques sont responsables des pics bien définis que l'on voit dans le spectre. En effet, périodiser en temps signifie discrétiser en fréquence. La fréquence du fondamental est celle de la période la plus grande visible dans le signal. La répartition des harmoniques va varier suivant la forme du signal (série de Fourier) :

(www) <http://www.dspguide.com/ch13/4.htm>

Note : en maintenant Ctrl+V appuyé, on peut copier indéfiniment une période (au préalable il faut avoir copié un morceau de piste).

Activité (3.2.2) Phénomène de Gibbs

Générer un signal carré de 50 Hz puis rééchantillonner à 8000 Hz. On voit apparaître le phénomène de Gibbs (oscillations aux jointures) car la série de Fourier est tronquée.

3.3. TF discrète (TFD) et TF rapide (FFT)

- Liens entre le chronogramme et la TFD →FICHE

► <http://www.impmc.jussieu.fr/~ayrinhac/videos/freqshift.mpg>

- La FFT est une TFD particulière basée (en général) sur l'algorithme de Cooley-Tukey. Gain de temps : la FFT nécessite $N \cdot \log_2(N)$ opérations de calcul contre N^2 pour la TFD. →FICHE

- Il existe une FFT optimisée appelée FFTW

(www) fftw.org

Le spectre dans Audacity

Audacity montre un spectre en énergie⁶. La méthode utilisée est la suivante : le signal est découpé en K intervalles de N points (on applique sur ces intervalles une fenêtre, celle de Hanning par exemple). Les fenêtres ne se recouvrent pas. On

⁶ En fait Audacity fait un « périodogramme moyenné » qui montre « la densité spectrale de puissance » utilisée pour les « signaux à puissance finie » (par exemple les signaux aléatoires). Le périodogramme est l'équivalent du spectrogramme mais il prend en compte l'énergie spectrale au lieu de l'amplitude spectrale.

fait la FFT de chaque intervalle, puis on moyenne les K FFTs, en prenant le module au carré ($|FFT|^2$ ou $\text{Re}(FFT)^2 + \text{Im}(FFT)^2$). Le spectre qui est montré est continu car il y a une interpolation cubique.

- Echelle verticale : énergie, en dB donc logarithmique. Dans cette représentation, le bruit prend autant d'importance que le signal lui-même. Il est impossible de changer dans la version 1.13 du logiciel : une astuce est d'ajouter un bruit blanc (et ainsi réduire l'échelle verticale et zoomer) pour éviter les grandes différences de niveaux. L'échelle verticale est négative (- X dB) : en effet, une amplitude de 1.0 dans le domaine temporel (amplitude max avant saturation) donne une amplitude de 0 dB dans le domaine fréquentiel. Une amplitude de 0 dB correspond donc à un signal non échantillé (sans distorsion), c'est-à-dire au signal de plus forte amplitude possible. C'est une norme inverse des dB audibles où 0 dB correspond au son de plus petite amplitude possible. L'amplitude d'un signal va dépendre aussi de la fenêtre employée : ainsi la fenêtre de Kaiser-Bessel donne la « bonne » amplitude alors que la fenêtre de Hamming donne la « bonne » fréquence.

Les dB ne correspondent pas au niveau sonore réel, cela dépend du niveau sur la piste, mais aussi de l'efficacité des haut-parleurs, et la distance au haut-parleur. Si l'amplificateur est éteint, un silence correspond à $-\infty$. Si l'amplifier est allumé mais le son inaudible, c'est 0 dB.

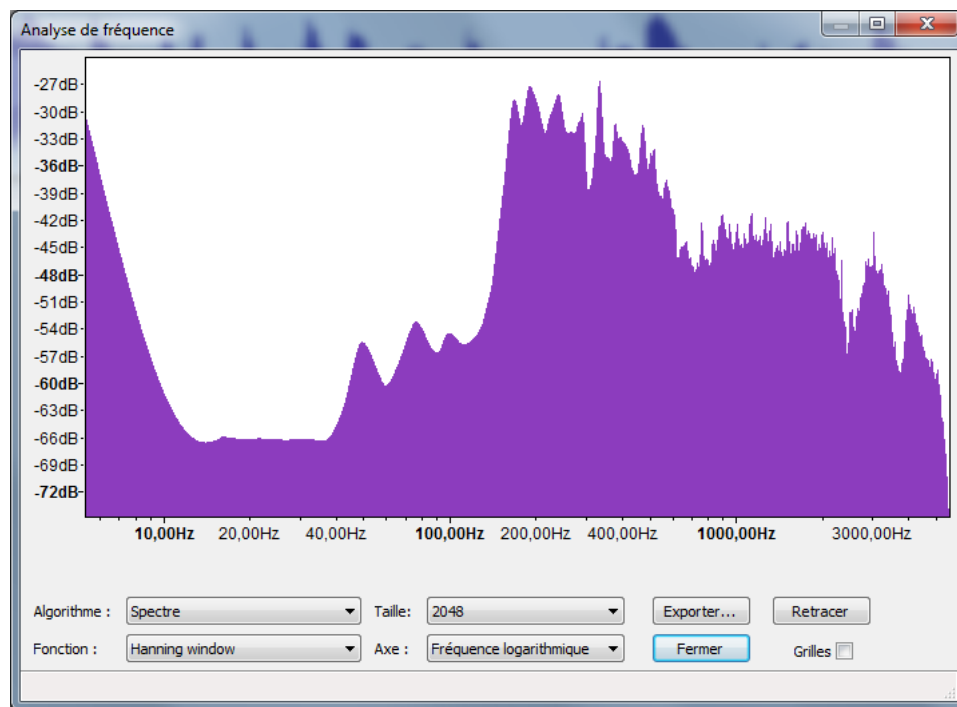
- Echelle horizontale : fréquences en Hz. A fréquence nulle, on trouve la moyenne du signal ! L'échelle peut être logarithmique (auquel cas la valeur 0 est rejetée à moins l'infini) ou linéaire.

Important : la fréquence maximale est $F_e/2$, on visualise $N/2$ points, N est une puissance de 2 car on utilise un algorithme de FFT, la résolution fréquentielle est F_e/N , le paramètre Taille ne peut pas dépasser le nombre de points du signal, le fichier à l'export présente Taille/2 points. Il y a un bug dans l'affichage : le spectre affiché est celui du dernier spectre obtenu même s'il ne peut pas calculer ce spectre.

- Le curseur : La valeur *curseur* indique la fréquence de la croix, tandis que *Crête* donne la fréquence du pic le plus proche. Les notations entre parenthèse traduisent la fréquence absolue en notation MIDI : A#3 signifie « LA dièse de l'octave 3 ».

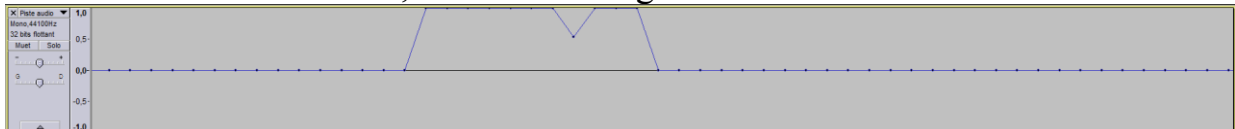
- Durée maximum du signal analysé : la longueur maximum analysable est de 10485760 échantillons. A un taux de 44100 Hz, on trouve 237.77234 secondes. Si on dépasse cette valeur (c'est usuellement le cas dans un morceau de musique courant) un message d'erreur s'affiche : "Too much audio was selected. Only the first %.1f seconds of audio will be analyzed". On peut analyser simultanément plusieurs signaux, mais ils doivent tous avoir le même taux d'échantillonnage.

- L'export : l'export du spectre donne un fichier de 2 colonnes : "Fréquence (Hz), Niveau (dB)". L'auto corrélation standard et les choix suivants donnent un fichier de 3 colonnes : "Décalage (secondes), Fréquence (Hz), Niveau"



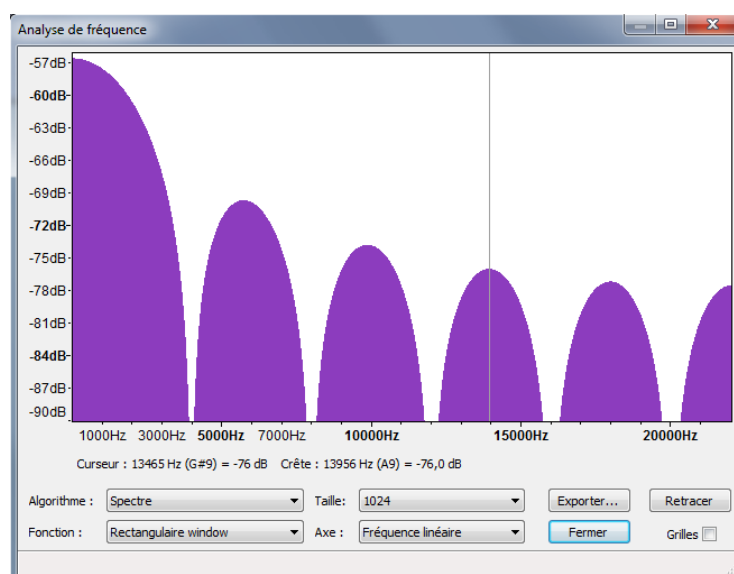
Activité (3.2.1) TF d'un créneau

Grâce à l'outil de retouche, il est facile de générer un créneau :



Comme attendu, le spectre possède un lobe principal centré sur 0 et des lobes secondaires (analyse en rectangulaire) :

En augmentant le créneau, on lobe principal. le lobe les hautes suffit de créneau par



importants fenêtre

la taille du rétrécit le Pour déplacer principal vers fréquences, il moduler le une sinusoïde.

Activité (3.2.2) Valeur moyenne (offset vertical, composante DC)

La valeur à fréquence nulle représente la moyenne du signal. Si celui-ci est centré sur 0, cette valeur est nulle. Dans Audacity, le menu utile est : Effets > normaliser > Supprimer tout décalage DC.

❖ <http://www.impmc.jussieu.fr/~ayrinhac/sons/soundshift.wav>

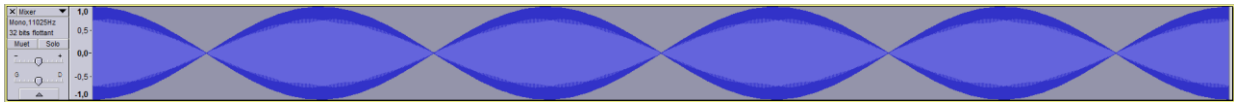
3.4 Résolution spectrale

Activité (3.4.1) Résolution spectrale

Essayons de séparer 2 raies de 200 Hz et de 243 Hz ($F_c=44100\text{Hz}$). On ne distingue les raies qu'au-dessus d'une taille de fenêtre $N=4096$. (Le pouvoir de résolution de Rayleigh est égal à la bande passante à 3 dB). La résolution est F_c/N .

Activité (3.4.2) Battements

Lorsque 2 raies sont proches, on constate l'apparition de battements. →FICHE Le montrer à partir de la relation trigonométrique $\cos(a+b)+\cos(a-b)=2\cos(a)\cos(b)$. C'est un "modulation d'amplitude" (AM) particulière : l'amplitude de la porteuse est modulée pour transmettre de l'information.



Le battement sert à accorder les instruments car un battement est dissonant.

3.5. Technique du zero-padding (ou bourrage de zéros)

On rajoute des zéros supplémentaires à la fin du signal étudié.

- le nombre de points du spectre augmente
- la résolution fréquentielle augmente car $\Delta f = F_c/N$
- Utile pour ajouter des points pour atteindre 2^n valeurs nécessaires à la FFT.
- il n'y a pas d'information supplémentaire par l'ajout de ces zéros.

►► <http://www.impmc.jussieu.fr/~ayrinhac/videos/zero-padding.mpg>

Activité (3.5.1). Zéro-padding dans Audacity

Dans Audacity, le zéro-padding ne peut pas améliorer la résolution spectrale, expliquer pourquoi. Réponse : c'est la technique du périodogramme qui rend inefficace le zéro-padding pour des signaux qui sont plus longs que la plus grande taille de fenêtre.

3.6. Le fenêtrage

→FICHE La fenêtre "naturelle" ou rectangulaire est abrupte. Fenêtre d'apodisation : les fenêtres, leurs TF, lobe principal, lobe secondaire.

► <http://www.impmc.jussieu.fr/~ayrinhac/videos/fenetrage.mpg>

Activité (3.6.1) Nature de la fenêtre d'analyse

Comparaison entre la fenêtre "naturelle" rectangulaire et une fenêtre plus "smooth" de Hanning, 2 sinus de 200 et 243 Hz d'amplitude 0,98 et 0,02. Chronogramme : on n'observe qu'une sinusoïde. Le spectre : on ne sépare les raies qu'avec la fenêtre de Hanning.

Voir ce document qui présente chaque fenêtre et son utilité pratique :

(www) http://www.altracustica.org/docs/fr_analyse_sig_sys.pdf

Activité (3.6.2) Taille de la fenêtre d'analyse

Générer une sinusoïde de fréquence 10 kHz, puis la fenêtrer et observer l'élargissement du lobe principal.

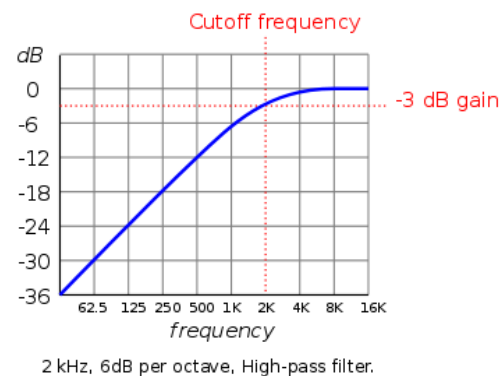
3.7. Filtrage et convolution

- Rappels sur le filtrage : passe-bas, passe-haut, passe-bande, coupe bande (=Notch).
- Relation filtrage-convolution : appliquer un filtre revient à convoluer dans le domaine temporel par un sinus cardinal (sinc). →FICHE

Activité (3.7.1) Filtrage

Test des filtres passe-haut et passe-bas avec la commande "Effets > High Pass Filter". La fréquence de coupure est définie comme la fréquence à -3dB du gain.

Lisse le signal quantifié au max.
Expliquer et Appliquer le filtre Notch sur un créneau temporel. Le filtre Notch élimine une bande de fréquences.



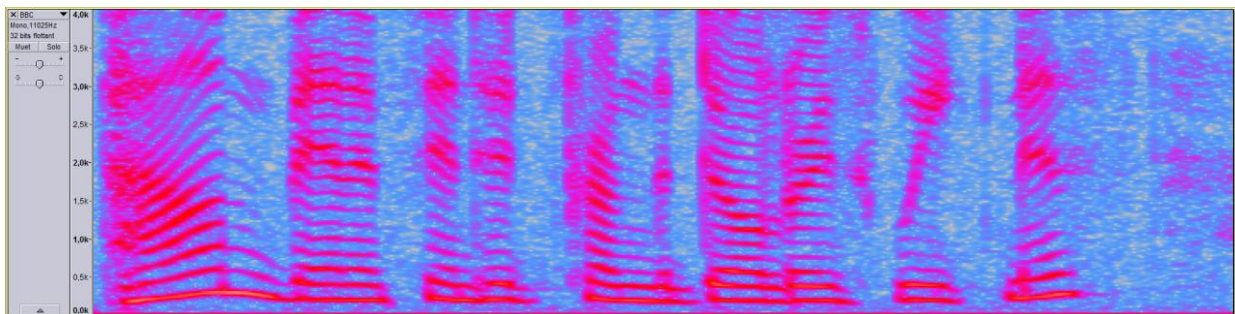
4. Spectrogramme et analyse temps-fréquence

4.1. Nature d'un spectrogramme

TF à fenêtre glissante, les fenêtres peuvent se recouvrir. →FICHE

Le spectrogramme dans Audacity

Un spectrogramme est la visualisation de l'amplitude de Fourier dans le plan temps-fréquence. La figure montrée par Audacity est la projection d'un diagramme 3D amplitude-temps-fréquence en 2D. L'échelle des couleurs indique la valeur de l'amplitude. Plus la fenêtre temporelle est petite et plus la résolution en fréquence est médiocre (analogue au principe d'incertitude de Heisenberg/Gabor en physique $\Delta t \Delta f \geq 1/2$).

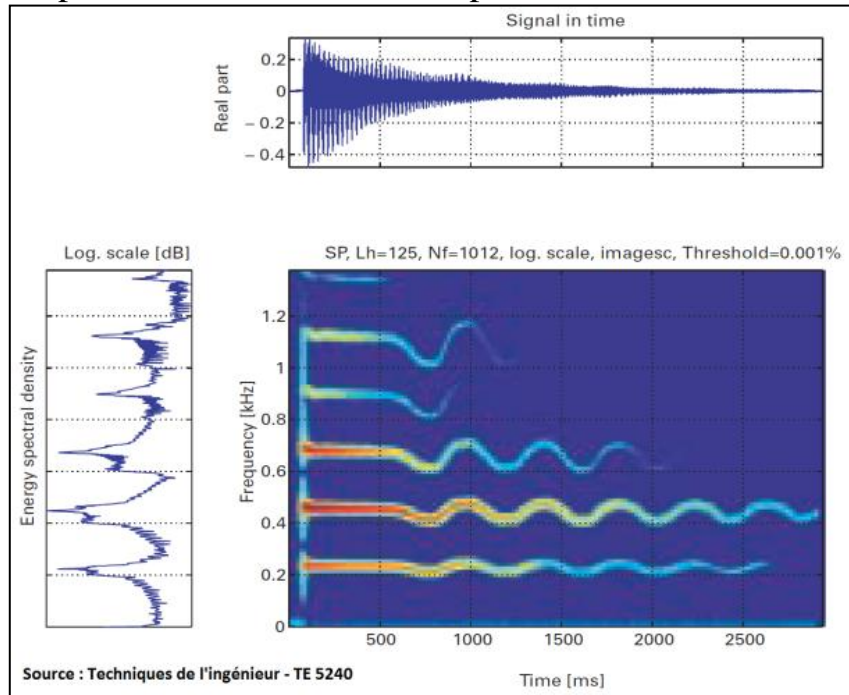


Le spectrogramme est accessible via le panneau de la piste. Les différents paramètres sont accessibles via le menu : Edition > Préférences > Spectrogramme. L'image du spectrogramme peut être sauvée via le menu Help>Screenshots Tools

L'échelle verticale est divisée en $N/2$ bandes où N est la taille de la fenêtre d'analyse. La fréquence max est $F_e/2$. Pour des sons plutôt basses fréquences, sous-échantillonner permet d'améliorer nettement la qualité du spectrogramme.

Il semble y avoir un rapport 10 entre la taille de la fenêtre dans les préférences et la taille de la fenêtre réelle (dépend de F_e , de la fenêtre employée...)

Autre type de représentation courante, en 3 panneaux :



http://www.impmc.jussieu.fr/~ayrinhac/images/Analyse_specgram_vibrato.png

Activité (4.1.1) Spectrogrammes

Le menu spectrogramme dans Audacity, changement des paramètres. Ouverture et visualisation de quelques sons complexes :

chauve-souris (bat.wav) : ◀ <http://tftb.nongnu.org/audio/bat.wav>

un gong (gong.wav) : ◀ <http://tftb.nongnu.org/audio/gong.wav>

baleine (whale.mp3) : ◀

<http://www.impmc.jussieu.fr/~ayrinhac/sons/whales.wav>

dauphin, etc.

Activité (4.1.2) Sonagrammes

Un sonagramme est un spectrogramme de la parole. Les voyelles ('A', 'E') sont reconnaissables grâce aux formants. Les formants sont des bandes d'énergie qui caractérisent les voyelles. Les consonnes ressemblent à une explosion ('P', 'K') ou à un bruit ('S', 'Z'). →FICHE

Voyelle : ◀ http://www.impmc.jussieu.fr/~ayrinhac/sons/voyelles_wiki.wav

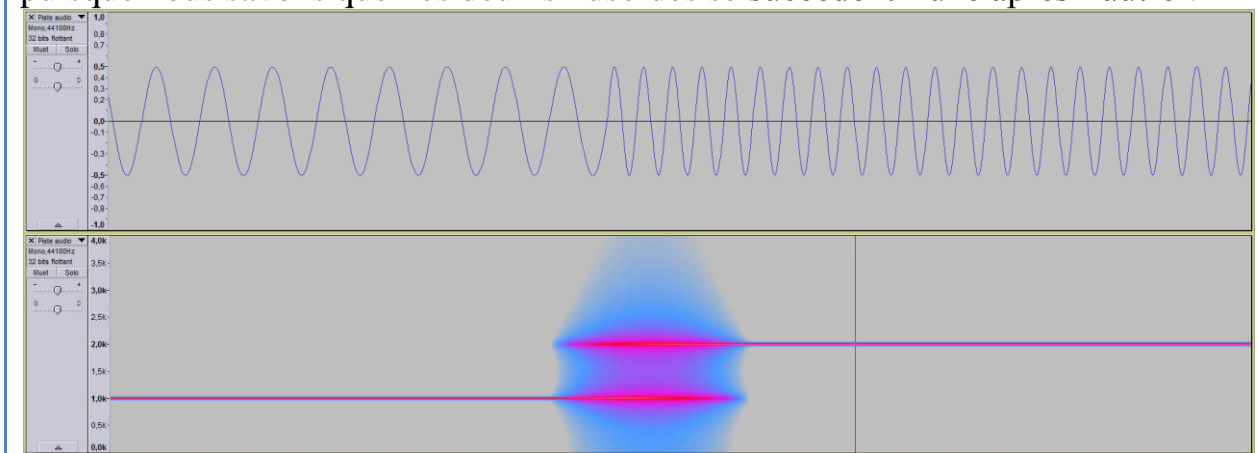
Consonnes : ◀ http://www.impmc.jussieu.fr/~ayrinhac/sons/consonnes_wiki.wav

Activité (4.1.3) Apollo 11

Lorsque Neil Armstrong prononce la phrase « That's one small step for (a) man; one giant leap for mankind », il y a un doute sur la présence de l'article « a ». Visualiser la phrase dans un spectrogramme, et essayer de conclure.

Activité (4.1.4) Localisation 1

créer un son (Fe=8kHz donc fréquence max =4000 Hz) qui comporte deux sinusoïdes collées de fréquence 1000 et 2000 Hz. Utiliser l'outil de glissement temporel pour coller les deux bouts de piste. Dans ce cas on peut faire varier la taille de la fenêtre N à partir de 8. A partir d'un grand N, il devient difficile d'identifier le temps de coupure. Il semble exister un intervalle de temps dans lequel les deux fréquences sont présentes en même temps. C'est **paradoxal**, puisque nous savons que nos deux sinusoïdes se succèdent l'une après l'autre !



Activité (4.1.5) Localisation 2

il suffit de dessiner le spectrogramme d'un son quelconque (musique ou paroles), de prendre une fenêtre de 4096 ou plus, puis de créer un petit silence : celui-ci est invisible. Il devient visible si la fenêtre temporelle est raccourcie.

Activité (4.1.6) Changement d'échelle

si on étale le temps, on concentre les fréquences (exemple avec les disques vinyles lus à la mauvaise vitesse). Pour expérimenter ceci sur Audacity, changer la vitesse de lecture de la piste via Panel>Fréquence d'échantillonnage ou Effet>Changer la vitesse. Un fichier de parole à vitesse très ralentie donne des grognements de fauve. [FICHE COURS] Changer la fréquence d'échantillonnage via le menu de la piste change la vitesse de lecture, et donc la fréquence du son. La commande équivalente est "Effets > Changer la vitesse > slider de pourcentage" Augmenter la fréquence sans augmenter le tempo se fait avec l'outil Effets>Changer la hauteur.

4.2. Repliement du spectre (Activité du théorème de Shannon)

[TD] notion de *fréquence instantanée*

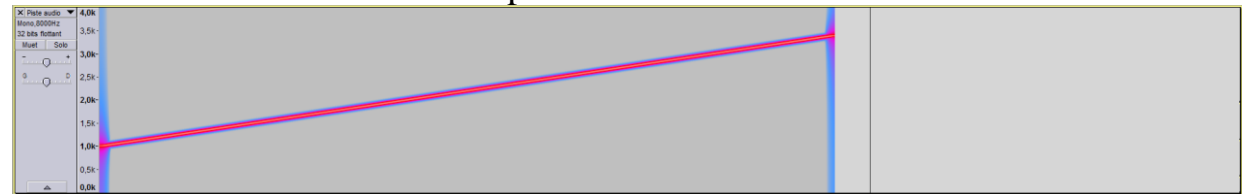
$$f_i = \frac{1}{2\pi} \frac{d\theta}{dt}$$

où θ est l'argument de $y(t)=A\cos[\theta(t)]$. Calculer $\theta(t)$ pour $f_i(t)=F_0+\lambda t$ (variation linéaire de la fréquence). [→FICHE](#)

Activité (4.2.1) Chirp linéaire

Un *chirp* linéaire et la variation linéaire de la hauteur (ou *pitch*) avec le temps. Dans Audacity, il est accessible via le menu Générer > Sifflet > Sinusoïde.

De 1kHz à 4 kHz sur 5 s avec amplitude constante à 1.



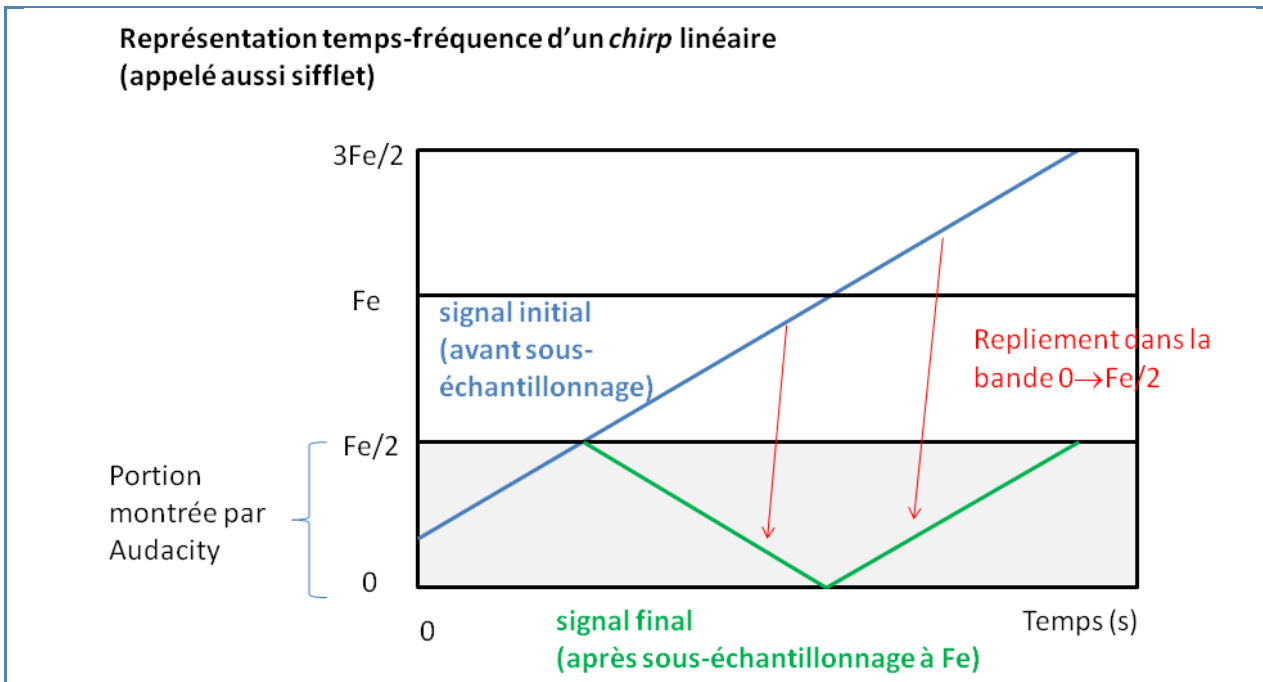
Activité (4.2.2) Repliement

pour créer du repliement, il faut sous-échantillonner à 8000 Hz, puis créer un chirp linéaire jusqu'à 10 kHz via le menu Générer>Sifflet. La hauteur du son monte puis descend !



Note : ce repliement est similaire à la définition de la première zone de Brillouin pour les cristaux en physique.

Le schéma ci-dessous explique cette forme en dents de scie :

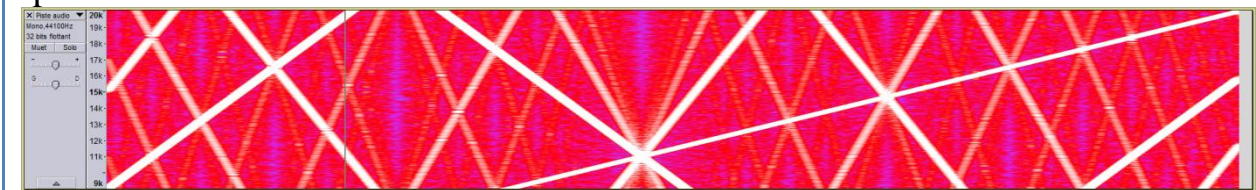


Activité (4.2.3) Filtre anti-repliement

Rééchantillonner avec Audacity : il y a disparition des hautes fréquences. Audacity doit appliquer un filtre passe-bas anti-repliement.

Activité (4.2.4) Sifflet avec forme d'onde carrée

Le spectrogramme d'un sifflet ayant une forme d'onde carrée est très spectaculaire.



4.2. Sous-échantillonnage et sur-échantillonnage

Expliquer le sous-échantillonnage (compression avec pertes des hautes fréquences) et le sur-échantillonnage. [→FICHE](#)

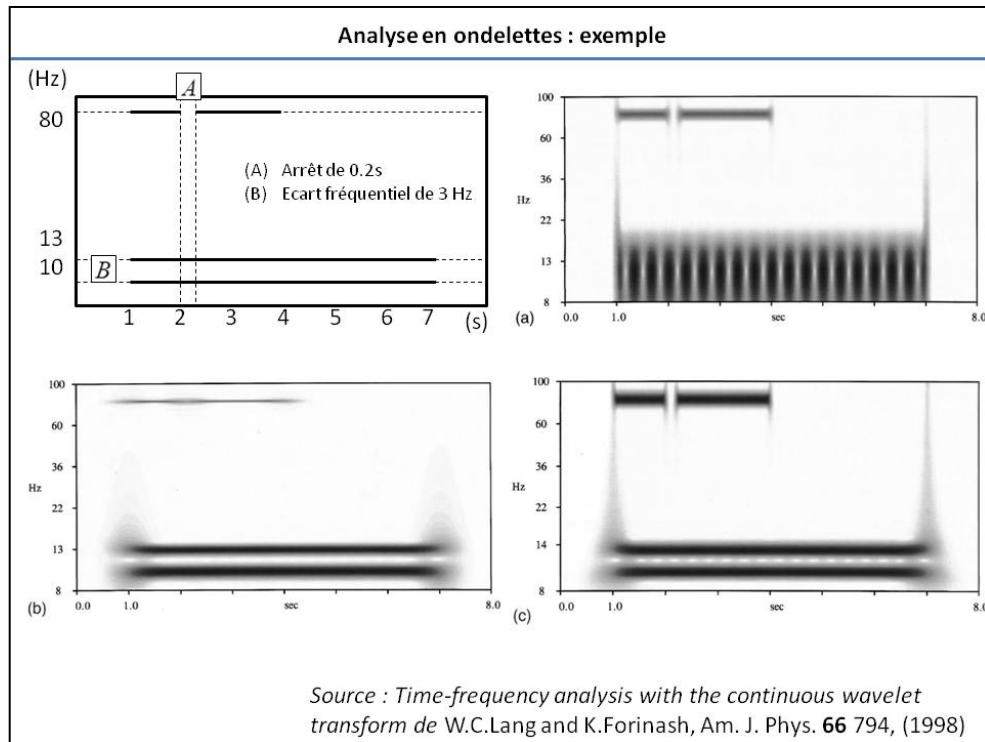
4.4. Ondelettes, scalogramme

La transformée en ondelettes est une analyse temps-échelle où la taille de la fenêtre varie avec la fréquence, et l'ondelette est adaptée au signal à étudier.

[→FICHE](#) Intérêt et applications.

Activité (4.4.1) Analyse en ondelettes

L'analyse par ondelettes permet dans certains cas une meilleure analyse temps-fréquence⁷



http://www.imPMC.jussieu.fr/~ayrinahac/images/Analyse_ondelettes_exemple.png

5. Compression, MP3

5.1. Compression

on a vu des façons simples de compresser :

- quantifier sur moins de bits
- sous-échantillonner par exemple.

Activité (5.1.1) Fichier .mp3

Ouvrir un .wav (générer un bruit blanc et un silence au milieu). L'enregistrer en .mp3, puis l'importer en dessous de la piste originale. Noter la diminution de la taille du fichier. Comparaison des spectrogrammes (fréquence max réglée à la fréquence $F_e/2$). On constate, des « effets de bord » sur le chronogramme mais surtout la disparition des hautes fréquences (au-dessus de 20kHz environ). C'est

⁷ Voir l'article *Time-frequency analysis with the continuous wavelet transform* de W. Christopher Lang and Kyle Forinash, Am. J. Phys. **66**, 794 (1998).

encore plus visible lorsqu'on change le « Frequency Gain » (dB/dec). Prendre une valeur de 10 ou 20. MP3=compression avec pertes (d'informations).

Activité (5.1.2) Fichier .ogg

Contrairement au format MP3, le format OGG Vorbis est libre, il est considéré comme de meilleure qualité pour une compression identique⁸. Noter la diminution de la taille du fichier par rapport au MP3. Ce format est beaucoup utilisé sur Wikipédia.

⁸ Voir livre de Christian Brochec, page 44.

Bibliographie (très sélective)

- (livre) *Audacity 2, Enregistrez, montez, mixez*, éditions Pearson, Christian Brochec, 2012, ISBN 978-2-7440-9402-6.
- (article) *The Hope of Audacity (To teach acoustics)*, Jennifer Groppe, The Physics Teacher **49** 99 (2011)
- (article) *Time-frequency analysis with the continuous wavelet transform* de W. Christopher Lang and Kyle Forinash, Am. J. Phys. **66** 794 (1998).
- (livre) *Mathématiques pour le traitement du signal* M. Bergounioux, DUNOD (2010) ISBN 978-2-10-054781-4
- (livre) *Discrete-Time Signal Processing*, A.V. Oppenheim, R.W.Schafer ISBN 978-0-13-206709-6
- Documents proposés pour l'unité d'ouverture « Musique et sciences » (LP212) de l'université Paris 6 :
<http://www.proba.jussieu.fr/users/lma/MusiqueSciencesL2.html>