



S.Ayrinhac (v13. 14/03/2014)

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 2.0.5. 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 d'enseignements)

Partie A

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 (SF).
- 3.3. TF discrète (TFD) et TF rapide (FFT).
Visualiser un spectre dans Audacity :
menu *Analyse* > *Tracer le spectre*.
- 3.4. Résolution spectrale et phénomène de battement.
- 3.5. Technique du bourrage de zéros (*zero-padding*).
- 3.6. Le fenêtrage

Partie B

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

5. Compression

- 5.1. Comparaison entre un .wav, un .mp3 et un .ogg du même morceau.

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

**Cours / rappels
théoriques**

Activité / TP

**Fonctionnement
d'Audacity**

Cet icône indique que des rappels de cours doivent être faits :

→FICHE

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

- écoute d'un son ◀
- vidéo explicative ▶▶
- document écrit 📄
- 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* ou *Fichier>Ouvrir*.

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.

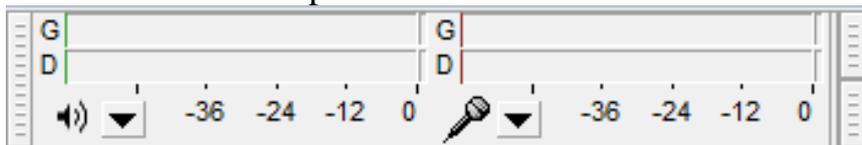
Le signal sonore peut être finalement stocké sous la forme d'un fichier d'extension *.wav*.

Activité (2.1.2) Enregistrement

Une fois le microphone branché, vous pouvez commencer l'enregistrement en allant dans le menu *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 *Fichier > Exporter > WAV signé 16 bits PCM*

Le VU-mètre (voir figure suivante) 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

Audacity permet de générer des sons simples : sinusoïde, créneaux, etc.

Pour générer un son, ajouter une nouvelle piste :

Menu Pistes > Ajouter nouvelle > Piste mono

Puis générer le son lui-même :

Menu Générer > son...

• 2.2. Analyse d'un chronogramme

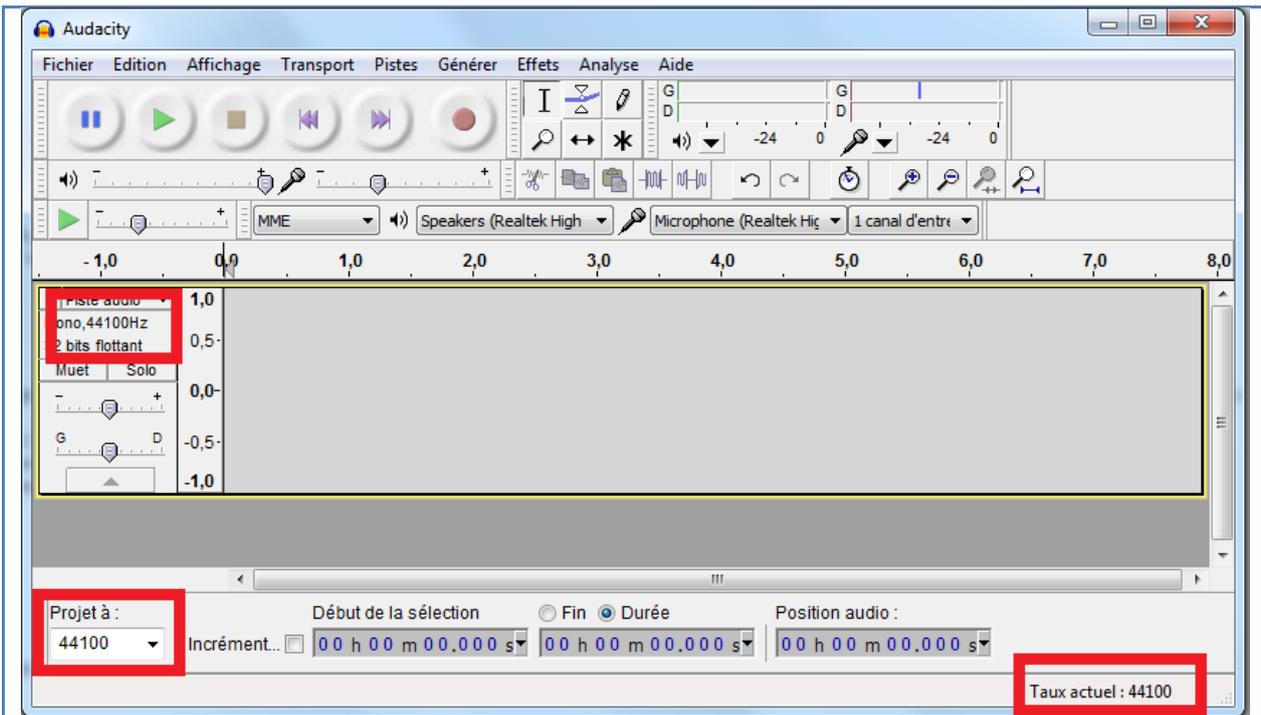
Activité (2.2.1) Chronogramme

Le chronogramme est appelé « forme d'onde » dans Audacity.

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).

Remarque : La F_e est indiquée de plusieurs manières différentes (voir graphique ci-dessous).



La case « Projet à » en bas à gauche permet de modifier la Fe de l'ensemble du projet.

Activité (2.2.3) Interférences destructives²

Question : à partir d'une piste quelconque, la dupliquer, puis l'inverser avec le menu *Effets > Inverse*. Ecouter les deux pistes ensemble. Que constatez-vous ?

Réponse : 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 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 1

Question : inverser un son quelconque avec *Effets > Inverse*. Comparer le son obtenu avec le son d'origine. Que constatez-vous ?

Réponse : un signal complexe inversé donne le même son que le signal de départ. En effet, le cerveau humain est peu sensible à la phase (le fait de commencer par une crête ou un creux) mais il est plus sensible au contenu fréquentiel.

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

Activité (2.2.5) Phase 2

Question : comparer les chronogrammes des deux sons ci-dessous, à la fois visuellement, et à l'écoute. Que constatez-vous ?

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

Réponse : bien que les chronogrammes soient très différents, les deux sons se révèlent semblables à l'écoute³.

Activité (2.2.6) Taille du fichier numérique

La taille L d'un fichier audio non compressé est donnée par la formule suivante :

$$L = F * T * Q * P$$

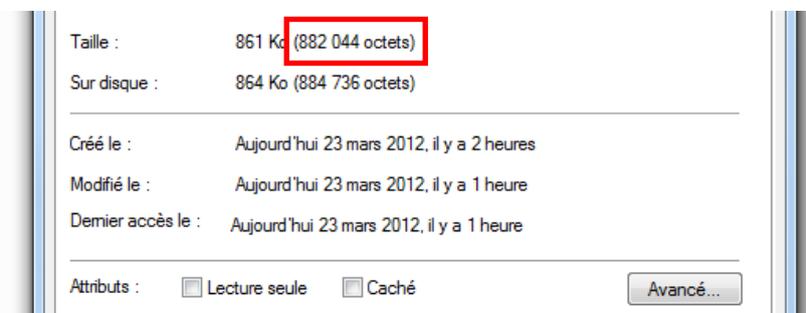
où F est la fréquence d'échantillonnage, T le temps d'enregistrement, Q la quantification et P le nombre de pistes (mono, stéréo, etc). Il faut diviser la taille par 8 pour convertir en octets.

Question : quelle est la taille d'un CD audio standard d'une heure ?

Réponse : $L = 44100 * 16 * 2 * 3600 / 8 = 635 \text{ Mo}$.

Question : prendre un fichier audio quelconque, calculer la taille attendue du fichier, et comparer avec la taille sur le disque. Trouvez-vous des tailles (en octets) identiques ? Pourquoi ?

Réponse : 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.



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

2.3. Bruit de quantification (bdq)

- Rappels : l'amplitude est codée sur b bits, et l'amplitude est découpée en L intervalles égaux (quantification uniforme). →FICHE

Question : quel est le lien mathématique entre L et b ?

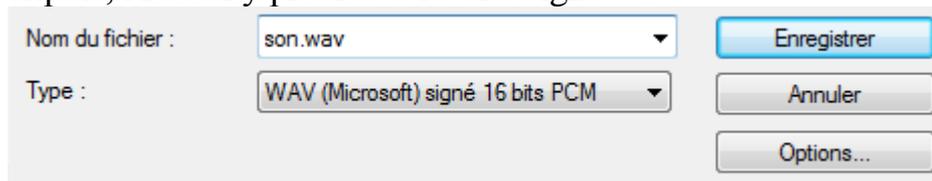
Activité (2.3.1) Quantification uniforme et non uniforme

Question : les fichiers .wav sont-ils codés avec une quantification uniforme ?

Réponse : lors de l'export en .WAV, il est indiqué PCM. Cet acronyme signifie « pulse code modulation ». Ce n'est pas une quantification uniforme, ce qui signifie que l'intervalle de variation de 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.

Question : si le codage est effectué sur 16 bits signés, quelle est la plage de valeurs possibles ?

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 Poulain⁴.

Important : ce n'est pas une « vraie » quantification qui a été effectuée sur les sons ci-dessous. En effet l'amplitude du son est quantifiée, mais le son numérique reste sur 16 bits !

Remarque : il est possible de choisir une « vraie » quantification à l'export via le menu *Fichier > Exporter > Autres formats non compressés > Options > Encodage > Unsigned 8 bit PCM*

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>

⁴ http://fr.wikipedia.org/wiki/Am%C3%A9lie_Poulain

- ◀ <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>

Question : à partir de quelle quantification entendez-vous un bruit parasite apparaître ?

Question : Quelle est la quantification maximale ? Le morceau ainsi quantifié est-il reconnaissable ?

Réponse : le signal ne contient que 0 ou 1, soit un codage sur 1 bit. Écoutons un morceau ainsi quantifié :

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

Activité (2.3.4) Bruit de quantification

On peut essayer de visualiser le bruit de quantification en soustrayant le son quantifié au son original. Pour cela, il faut inverser une des deux pistes (*Effets > Inverser*) et les sommer (*Pistes > Mixage et rendu*).

Question : le bruit obtenu est-il un bruit blanc ? Si non, pourquoi ?

Réponse : en théorie, le bruit de quantification possède les caractéristiques d'un bruit blanc lorsque la quantification est uniforme. En pratique, ce n'est pas le cas, car les .WAV sont sauves en PCM qui est une quantification non uniforme.

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

Exercice : calcul de la TF d'un créneau. →FICHE

Animation : lorsque la largeur de la fonction porte augmente, la largeur du lobe principal du spectre diminue.

- ▶ <http://www.imPMC.jussieu.fr/~ayrinhac/videos/TFporte.mpg>

Voici 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>

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

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.
Voici un sinus cardinal généré avec Scilab :

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

3.2. Rappels sur la série de Fourier

- Définitions mathématiques de la série de Fourier (SF) → [FICHE](#)
 - Lien et différence avec les coefficients de Fourier (notion de série de Fourier).
- Important :** ne pas oublier que lorsqu'on périodise en temps, on discrétise en fréquence (et inversement, lorsqu'on discrétise en temps, on périodise en fréquence).

Voici la décomposition de quelques signaux usuels en séries de Fourier :

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

Activité (3.2.1) Motif périodique et spectre de raies

Grâce à l'outil retouche (icône ci-dessous), générer un motif quelconque.



Exemple de motif :



Dupliquer ce motif pour créer un son. Une méthode possible est la suivante : en maintenant Ctrl+V appuyé, on peut copier indéfiniment une période (au préalable il faut avoir copié un morceau de piste, et placé le curseur en la fin de la piste). Ne pas oublier de supprimer la composante DC si elle existe. Tracer le spectre.

Question : pourquoi a-t-on un spectre de raies ? Identifier le fondamental et les harmoniques.

Réponse : c'est la périodicité du motif qui est responsable des pics bien définis que l'on voit dans le spectre. La fréquence du fondamental est celle de la période la plus grande visible dans le signal (la durée du motif de base). La répartition

des harmoniques va varier suivant la forme du signal (série de Fourier) :
 (www) http://www.dspguide.com/graphics/F_13_10.gif

Activité (3.2.2) Phénomène de Gibbs 1

Générer plusieurs sinusoïdes sur plusieurs pistes différentes avec les fréquences suivantes : 100 Hz, 300 Hz, 500 Hz, 700 Hz, 900 Hz et les amplitudes suivantes : 0.1, 0.033, 0.02, 0.0142, 0.0111.
 Sommer toutes les pistes. Quel est le résultat ?

Activité (3.2.3) Phénomène de Gibbs 2

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.
 Remarque : il est possible aussi d'appliquer un filtrage passe-bas avec un flanc très abrupt, le maximum étant de 48dB par décade (*Effets>Low pass filter>Rolloff 48 dB>Cutoff 8000 Hz*).

3.3. TF discrète (TFD)

Rappels de cours sur la transformée de Fourier discrète (TFD) [→FICHE](#)
http://www.esiee.fr/~bercherj/New/polys/poly_tfd.pdf

Important : quel est le lien entre la TF, la TFD et la SF ?

Voici un tableau de synthèse⁶ :

http://www.impmc.jussieu.fr/~ayrinhac/images/dualite_temps-frequence.png

Cette animation illustre une propriété particulière de la transformée de Fourier discrète : lorsque la fréquence de la sinusoïde est un multiple de l'intervalle spectral, on retrouve le spectre idéal (un Dirac).

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

3.3. Transformée de Fourier rapide (FFT)

[→FICHE](#) La FFT (acronyme de *fast fourier transform*) est une TFD particulière basée (en général) sur l'algorithme de Cooley-Tukey. La FFT apporte un gain de temps : elle nécessite $N \cdot \log_2(N)$ opérations de calcul au lieu

⁶ Ce tableau est issu du site Patrick.furon.free.fr

de N^2 pour la TFD.

Il existe une FFT optimisée appelée FFTW. Pour plus de détails, voir le site suivant :

([www](http://www.fftw.org)) fftw.org

Le spectre dans Audacity

La Transformée de Fourier Discrète (TFD) et la Transformée de Fourier rapide (FFT) sont à la base de l'analyse spectrale en informatique. Audacity utilise la FFT, mais de manière un peu particulière.

La méthode utilisée par Audacity 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 fait la FFT de chaque intervalle, puis on moyenne les K FFTs, en prenant le module au carré ($|FFT|^2$ ou $\text{Re}(FFT)^2 + \text{Im}(FFT)^2$). Au final Audacity montre un spectre en énergie⁷.

- Echelle verticale : elle représente une énergie en dB (c'est donc une échelle logarithmique). Dans cette représentation, le bruit prend autant d'importance que le signal lui-même.

- Il est impossible de changer cette échelle : 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 maximale avant saturation) donne une amplitude de 0 dB dans le domaine fréquentiel. Une amplitude de 0 dB correspond donc à un signal non écrêté (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. Cela explique que l'échelle verticale soit négative. L'amplitude d'un signal va dépendre aussi de la fenêtre employée : ainsi la fenêtre de

⁷ 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.

Kaiser-Bessel donne la « bonne » amplitude alors que la fenêtre de Hamming donne la « bonne » fréquence⁸.

- Les dB ne correspondent pas au niveau sonore réel, cela dépend du niveau de 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.
- **Important** : le spectre obtenu par TFD est forcément discontinu. Or Audacity nous montre une courbe continue. En effet, cette courbe est obtenue par une interpolation cubique (par des polynômes de degré 3) des points issus de la FFT. Lorsqu'on exporte le spectre, on constate que celui-ci est bien constitué de valeurs discrètes.

- Echelle horizontale : fréquences en Hz.

- A fréquence nulle, on trouve la moyenne du signal (appelée aussi « composante DC⁹ »).
- L'échelle peut être logarithmique (auquel cas la valeur 0 est rejetée à moins l'infini) ou linéaire.
- **Important** : la fréquence maximale est $F_e/2$ et on visualise $N/2$ points. En effet, le signal entre $F_e/2$ et F_e est l'image du signal entre 0 et F_e .
- N est une puissance de 2 car on utilise un algorithme de FFT, la résolution fréquentielle est donc 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 pointée par la croix, tandis que *Crête* donne la fréquence du pic le plus proche. Les notations entre parenthèses traduisent la fréquence absolue en notation MIDI : A#3 signifie « LA dièse de l'octave 3 ».

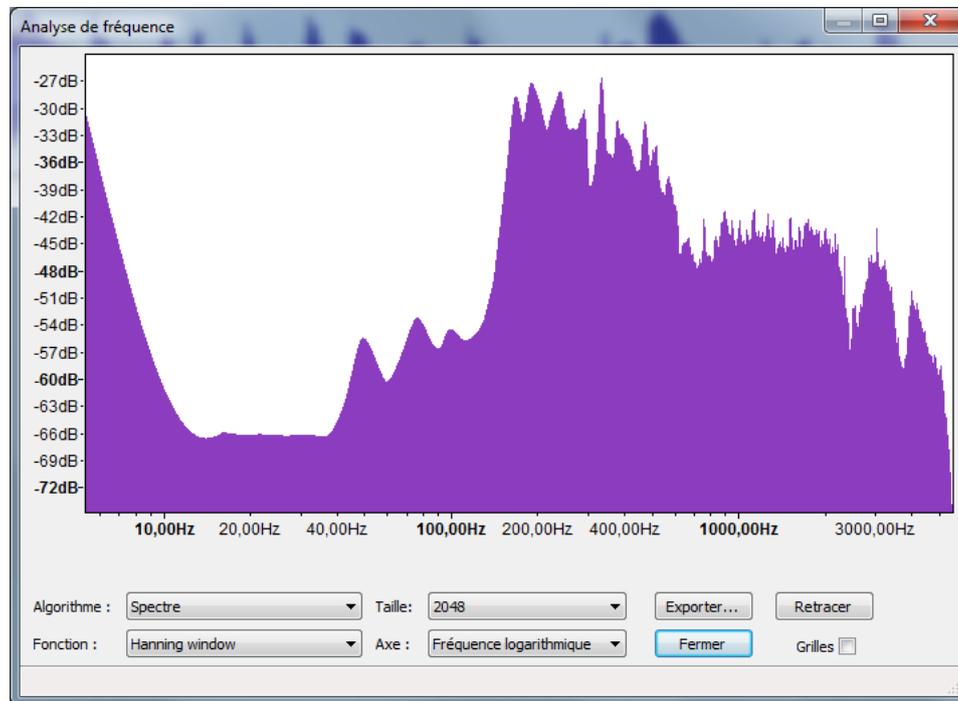
- Durée maximum du signal analysé : la longueur maximum analysable est de 10485760 échantillons. A un taux de 44100 Hz, on trouve 237.77234 secondes. Si on dépasse cette valeur (c'est usuellement le cas dans un morceau de musique courant) un message d'erreur s'affiche : « Too much audio was selected. 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.

⁸Voir le document **Analyse du signal(FFT et Filtrage numérique) & Analyse des systèmes** par J.Dumas et B.Bennevault :

www.altracustica.org/docs/fr_analyse_sig_sys.pdf

⁹ Par analogie avec un signal électrique continu (« DC » signifie *direct current*).

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



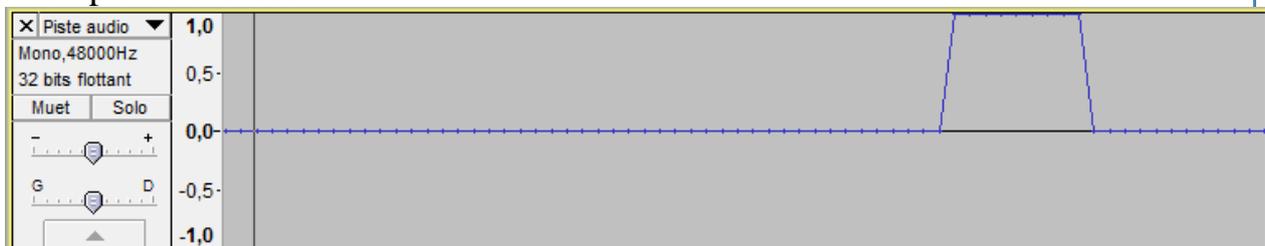
Activité (3.2.1) TF d'un créneau

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

Icône de l'outil de retouche :



Exemple de créneau :

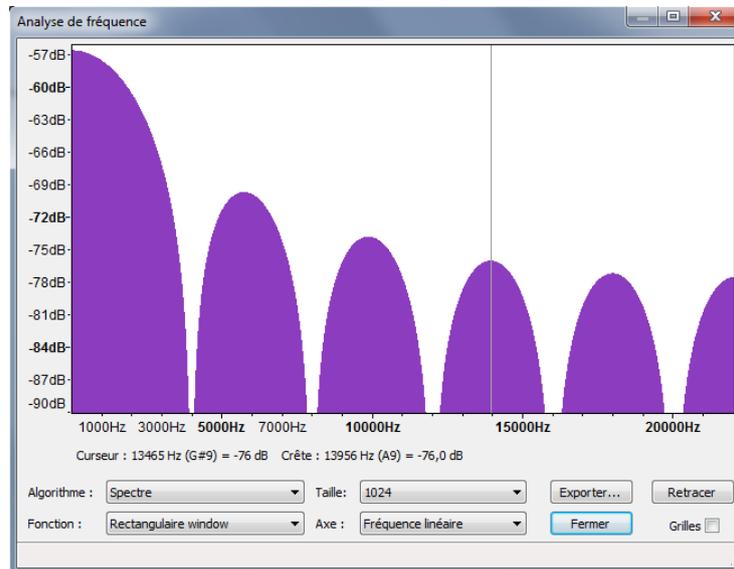


Question : quel est le spectre attendu ? Correspond-t-il au spectre donné par Audacity ? Que se passe-t-il si on augmente la largeur du créneau ?

Réponse : visualisons le spectre : comme attendu, le spectre possède un lobe

principal centré sur 0 et des lobes secondaires importants (analyse en fenêtre rectangulaire).

En augmentant la taille du créneau, on rétrécit le lobe principal. Pour déplacer le lobe principal vers les hautes fréquences, il suffit de moduler le créneau par une sinusoïde.



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

Voici un son qui possède une composante DC :

◀ <http://www.imPMC.jussieu.fr/~ayrinhac/sons/soundshift.wav>

Visualiser le spectre (avec fenêtre de Hanning) puis utiliser l'outil suivant : *Effets* > *normaliser* > *Supprimer tout décalage DC*.

Question : que remarquez-vous ?

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 . Cette résolution donne **l'incertitude grossière** sur la mesure d'une fréquence. Une taille de fenêtre plus grande améliore la précision des mesures de fréquence.

3.4.2. Les battements

Lorsque 2 raies sont proches, on constate l'apparition de battements. →FICHE

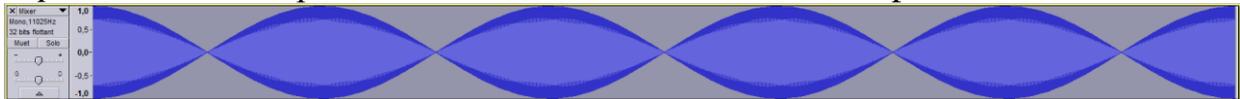
On peut le montrer à partir de la relation trigonométrique $\cos(a+b)+\cos(a-b)=2\cos(a)\cos(b)$. C'est un "modulation d'amplitude" (AM) particulière : l'amplitude de la porteuse est modulée pour transmettre de l'information.

Activité (3.4.2) Battements

Générer 2 pistes de 440 et 442 Hz à 0.5 d'amplitude, séparées. Ecouter.

Question : que remarquez-vous ?

Réponse : voici ce que l'on observe en fusionnant les deux pistes :



Application : 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
- l'intervalle entre 2 points du spectre augmente car $\Delta f = F_e/N$ **mais** la résolution fréquentielle reste la même car la durée du signal reste inchangée.
- 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.
- c'est une technique d'interpolation

Voici une animation qui illustre le zéro-padding (la ligne bleue claire correspond à une TFD avec un pas temporel très faible pour simuler une TF non discrète)

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

Activité (3.5.1). Zéro-padding dans Audacity

Question : comment effectuer le zero-padding simplement sous Audacity ?

Prendre une piste quelconque, effectuer un zéro-padding et comparer les spectres.

La résolution (le plus petit intervalle fréquentiel) est-elle modifiée ?

Réponse : il suffit d'ajouter un silence, la résolution reste identique.

Question : dans Audacity, le zéro-padding ne peut pas améliorer la résolution spectrale, pourquoi ?

Réponse : c'est la technique du périodogramme (utilisée par Audacity pour tracer le spectre) qui rend inefficace le zéro-padding pour des signaux qui sont plus longs que la plus grande taille de fenêtre.

3.6. Le fenêtrage

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

Voici un document qui présente chaque fenêtre et son utilité pratique¹⁰ :

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

Cette animation présente l'effet de la taille de la fenetre « naturelle » sur un signal sinusoidal :

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

Activité (3.6.1) Nature de la fenêtre d'analyse

Observons l'effet du fenêtrage sur le spectre.

Exemple : générer 2 sinus de 200 et 243 Hz d'amplitude 0,98 et 0,01.

Observer le spectre, entre la fenêtre "naturelle" rectangulaire et une fenêtre moins abrupte de Hanning.

Question : avec quelle fenêtre observe-t-on le mieux les deux raies ?

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.

4. Spectrogramme et analyse temps-fréquence

4.1. Nature d'un spectrogramme

Un spectrogramme est une TF à fenêtre glissante (les fenêtres peuvent se recouvrir). →FICHE

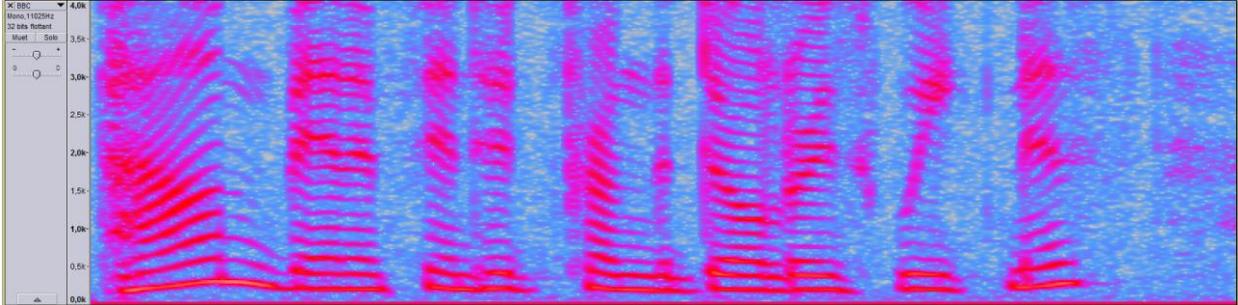
Le spectrogramme dans Audacity

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

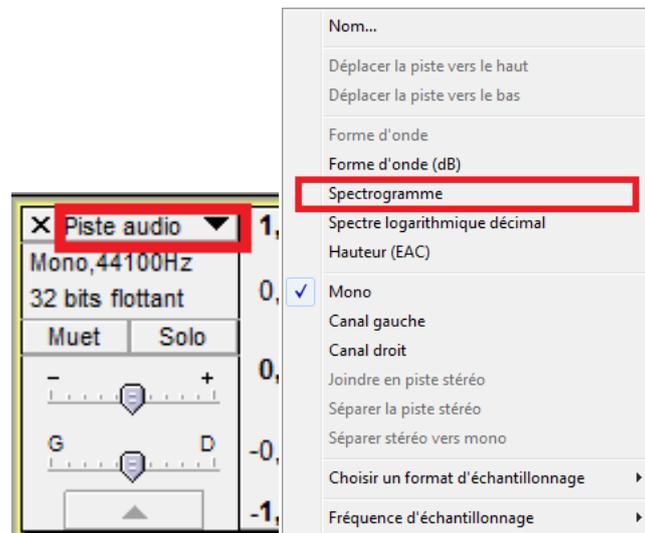
¹⁰ Voir aussi l'article de référence : *On the use of windows for harmonic analysis with the discrete Fourier transform* de Harris F.J. DOI : [10.1109/PROC.1978.10837](https://doi.org/10.1109/PROC.1978.10837)

en fréquence est médiocre. Ce phénomène est analogue au principe d'incertitude de Heisenberg en physique. En traitement du signal, on l'appelle parfois l'inégalité de Heisenberg-Gabor.

Exemple de spectrogramme dans Audacity :



L'affichage du spectrogramme est accessible via le panneau en début de piste (voir les deux figures ci-dessous) :



Important : les différents paramètres sont accessibles via le menu : *Edition* > *Préférences* > *Spectrogramme*.

L'image du spectrogramme peut être sauvée via le menu *Help*>*Screenshots Tools*.

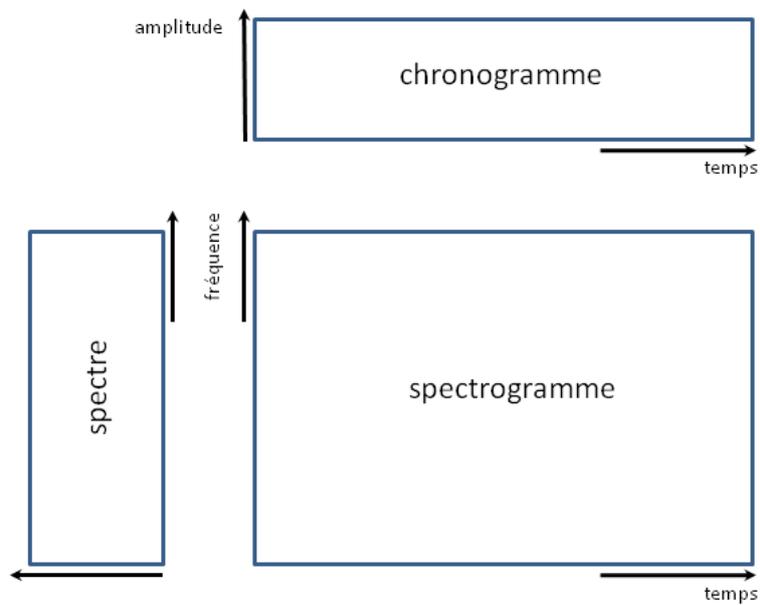
L'échelle verticale est divisée en $N/2$ bandes où N est la taille de la fenêtre d'analyse.

Important : la fréquence max visualisée par le spectrogramme est $F_e/2$ par défaut, de la même manière que le spectre va de 0 à $F_e/2$.

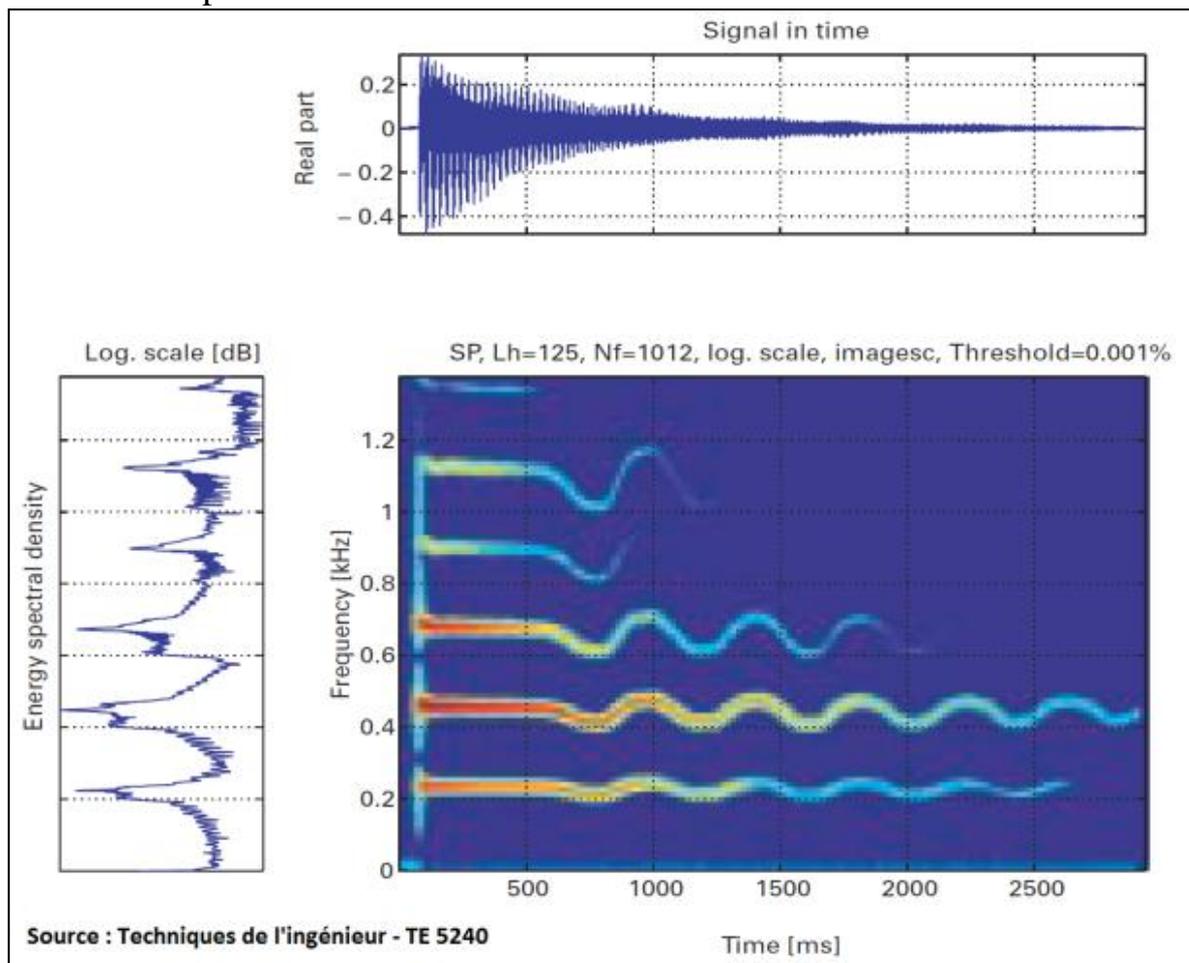
Pour des sons concentrés dans les basses fréquences, sous-échantillonner permet d'améliorer parfois la qualité du spectrogramme.

Il existe un type très courant de représentation courante, en 3 parties, où les

différents axes sont mis en correspondance :



Voici un exemple :



L'image est disponible à cette adresse :

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

Activité (4.1.1) Spectrogrammes

Ici, on vous propose de visualiser le spectrogramme de quelques sons complexes. Réglez au mieux les paramètres pour améliorer la visualisation.

Voici quatre exemples de 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>

vibrato : <http://www.imPMC.jussieu.fr/~ayrinhac/sons/Norma-debut.mp3>

Activité (4.1.2) Spectrogramme et partition

Une partition musicale est la retranscription d'un spectrogramme.

Vous pouvez le vérifier en comparant le morceau « Au clair de la lune » joué au piano, et sa partition :

http://www.imPMC.jussieu.fr/~ayrinhac/sons/au_clair_de_la_lune.wav



Activité (4.1.3) Apollo 11

On s'intéresse ici à une « énigme historique ». Juste après le premier alunissage, le 21 juillet 1969, l'astronaute Neil Armstrong prononça la fameuse 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 » : l'astronaute l'a-t-il prononcé ?

Visualiser la phrase dans un spectrogramme, et essayer de conclure.

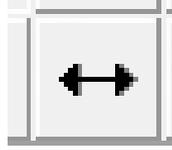
http://www.imPMC.jussieu.fr/~ayrinhac/sons/Neil_Armstrong_One_Small_Step.wav

Le fichier original est librement accessible sur Wikipédia :

http://upload.wikimedia.org/wikipedia/commons/4/48/Frase_de_Neil_Armstrong.ogg

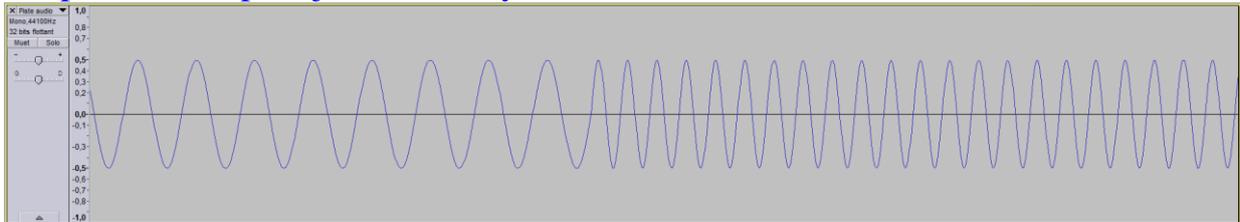
Activité (4.1.4) Localisation 1

Créer un son ($F_e=8\text{kHz}$ 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.



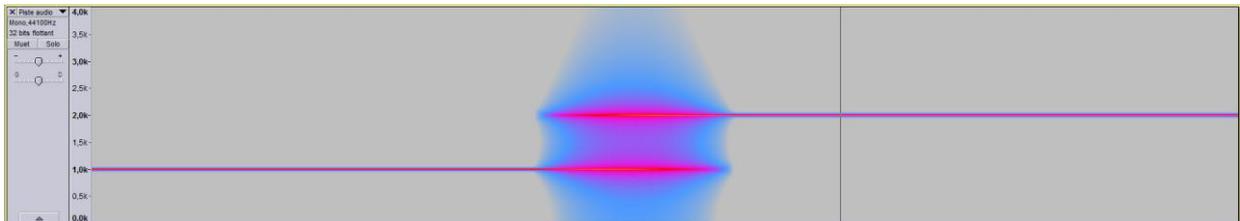
Le fichier est directement disponible ici :

☛ http://www.imPMC.jussieu.fr/~ayrinhac/sons/deux_sins_colles.wav



Visualiser ce son dans un spectrogramme et faire varier la taille de la fenêtre N.

Question : pour quelle taille de fenêtre mesure-t-on le mieux l'instant de raccordement ?



Conclusion : 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

Prendre un son quelconque (musique ou paroles), puis créer un petit silence au milieu de la piste. Dessiner le spectrogramme.

Question : quelle est la taille de fenêtre N qui permet de distinguer ce silence ?

Question : dans le son ci-dessous se cache un silence. Trouver sa position temporelle et sa durée ?

☛ http://www.imPMC.jussieu.fr/~ayrinhac/sons/son_silence_cache.wav

Activité (4.1.6) Changement d'échelle (*time scaling*)

Question : montrer grâce à la transformée de Fourier que si on multiplie le temps par un facteur γ , on divise les fréquences par le même facteur (exemple avec les disques vinyles lus à la mauvaise vitesse).

Pour expérimenter ceci sur Audacity, on peut changer la vitesse de lecture de la piste via *Panel*>*Fréquence d'échantillonnage* ou *Effet*>*Changer la vitesse*. Par

exemple, un fichier audio de parole à vitesse très ralentie donne des grognements de fauve. Augmenter la fréquence sans augmenter le tempo se fait avec l'outil *Effets > Changer la hauteur*.

4.2. Fréquence instantanée.

Lorsque la fréquence est une fonction du temps, on a besoin d'introduire la notion de *fréquence instantanée* F_i :

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

où θ est l'argument du chronogramme $s(t) = A \cdot \cos[\theta(t)]$.

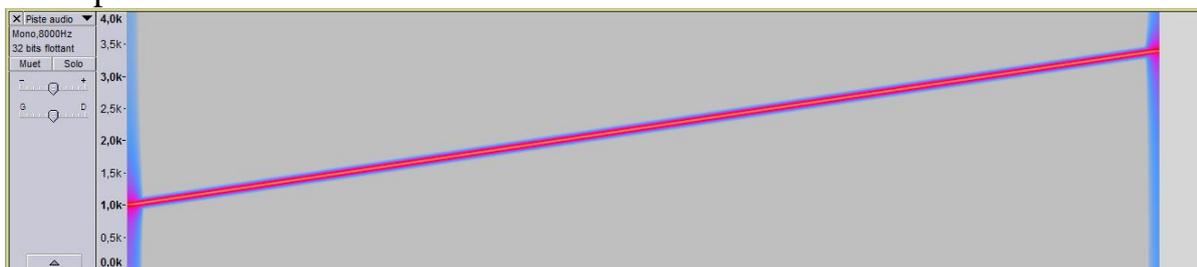
Question : Calculer $\theta(t)$ pour $F_i(t) = F_0 + \lambda t$ (variation linéaire de la fréquence) et en déduire $s(t)$. Ce signal est-il périodique ? (on peut répondre en visualisant le spectre)

Question : que vaut F_i pour un vibrato ? Quelle est la forme de $s(t)$ correspondante ?

Activité (4.2.1) Chirp linéaire (sifflet)

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

Exemple : générer un sifflet de 1kHz à 4 kHz sur 5 s avec amplitude constante. Dessiner le spectrogramme et observer que la fréquence varie linéairement avec le temps.



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

Rappels de cours [→ FICHE](#)

Question : énoncer le théorème de Shannon.

Question : que se passe-t-il si on échantillonne un signal sinusoïdal de 1 kHz à $F_e = 1.6$ kHz ?

Activité (4.2.2) Repliement (*aliasing*)

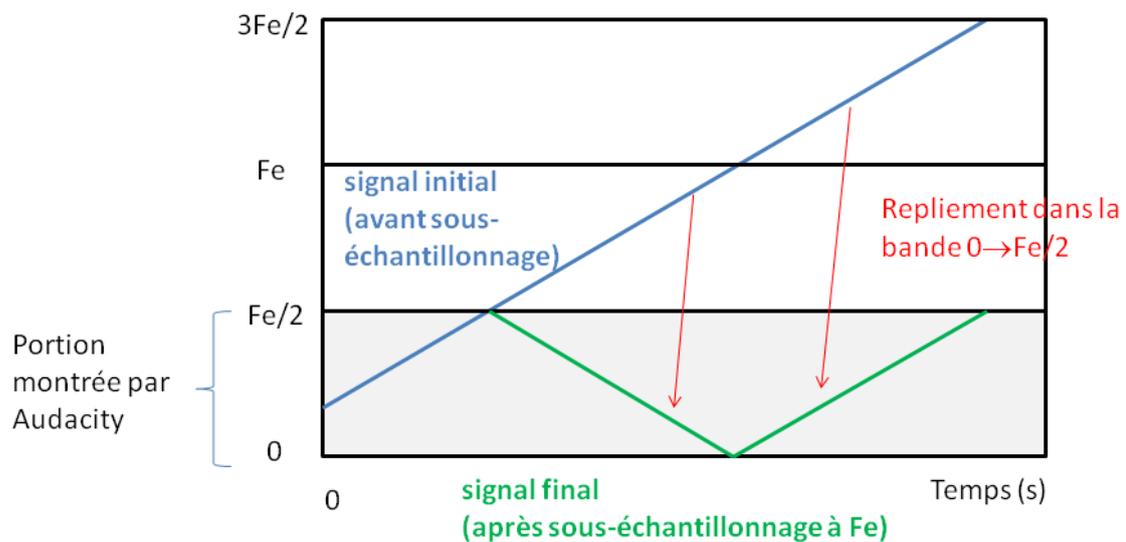
Pour observer du repliement, il faut d'abord sous-échantillonner à 8000 Hz, puis créer un chirp linéaire de 1kHz à 10 kHz via le menu *Générer>Sifflet*.

Question : on observe que la hauteur du son monte puis descend. Pourquoi ?



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

Représentation temps-fréquence d'un *chirp* linéaire (appelé aussi sifflet)



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

Activité (4.2.3) Filtre anti-repliement

Question : on rééchantillonne le sifflet précédent à 100 Hz, on observe la disparition des hautes fréquences. Pourquoi ?

Réponse : Audacity applique un filtre anti-repliement.

Activité (4.2.4) Sifflet avec forme d'onde carrée

Visualiser le spectrogramme d'un sifflet ayant une forme d'onde carrée.



4.3. 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. Intérêt et applications. [→FICHE](#)

Activité (4.3.1) Analyse en ondelettes

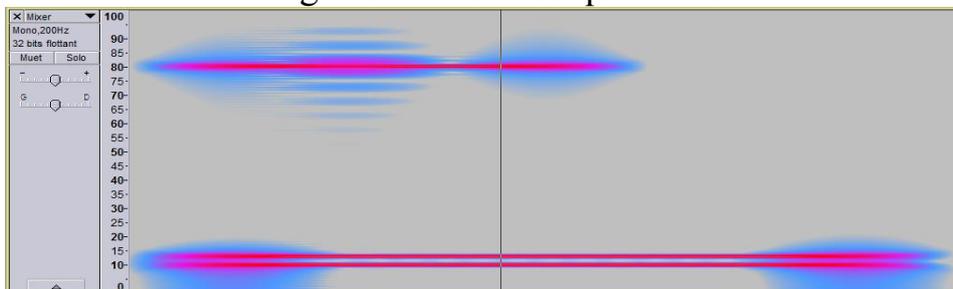
L'analyse par ondelettes permet dans certains cas une meilleure analyse temps-fréquence¹¹.

Exemple avec un son complexe.

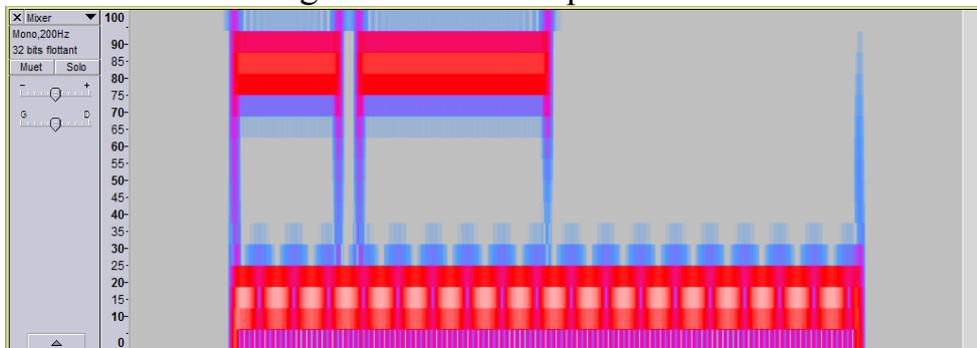
◀ http://www.imPMC.jussieu.fr/~ayrinhac/sons/son_complexe.wav

On rééchantillonne d'abord à 200 Hz .

Définition d'affichage : 512. Voici ce que l'on observe :



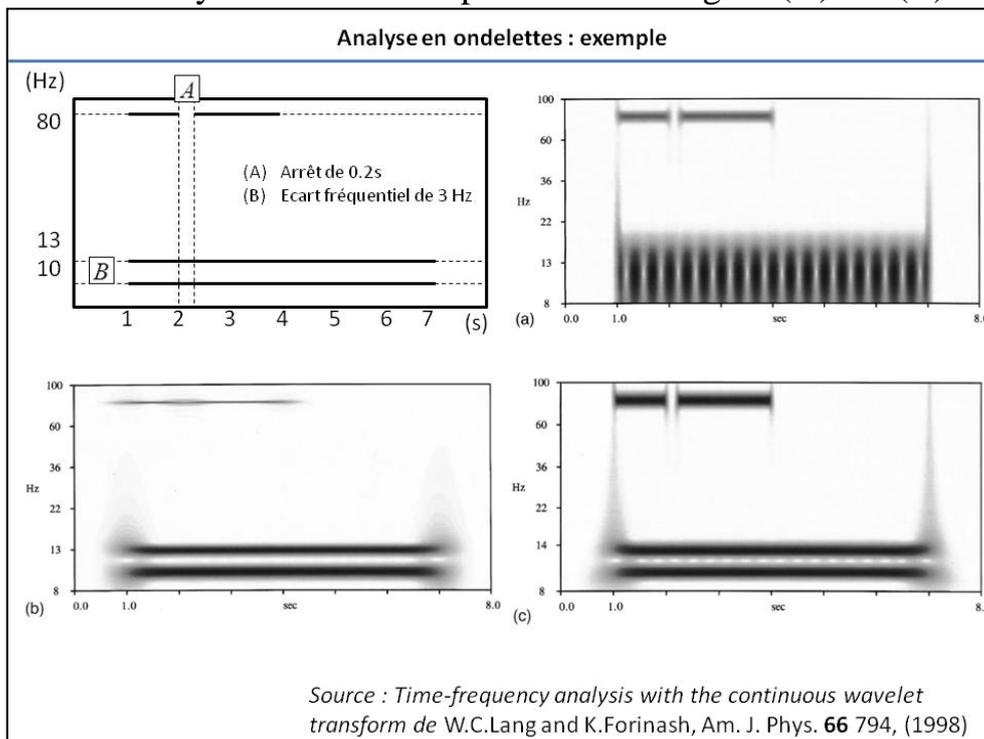
Définition d'affichage : 32. Voici ce que l'on observe :



¹¹ Voir l'article *Time-frequency analysis with the continuous wavelet transform* de W. Christopher Lang and Kyle Forinash, Am. J. Phys. **66**, 794 (1998).

Conclusion : soit on voit le silence sur le signal à 80 Hz, soit on distingue les deux fréquences à 10 Hz et 13 Hz.

Résumé avec la figure ci-dessous : soit un signal complexe qui possède un silence de 0.2s en (A) et deux fréquences très proches en (B). En haut à droite, on a privilégié la résolution temporelle : on distingue (A) mais pas (B). En bas à gauche, on a privilégié la résolution fréquentielle : on voit (B) mais pas (A). En bas à droite une analyse en ondelettes permet de distinguer (A) ET (B).



Cette image est disponible à l'adresse suivante :

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

L'article de W.C.Lang et K.Forinash est disponible ici :

<http://scitation.aip.org/content/aapt/journal/ajp/66/9/10.1119/1.18959>

5. Compression, MP3

5.1. Compression

On a vu des façons simples de compresser :

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

Le format MP3 est une compression avec pertes d'informations.

Activité (5.1.1) Fichier .mp3

Générer un bruit blanc. Enregistrer ce son au format .wav et au format .mp3.

Question : quel est le gain de taille apporté par le format .mp3 ?

Question : comparer les spectrogrammes des deux sons .wav et .mp3. Que constate-t-on ?

Réponse : on constate, des « effets de bord » sur le chronogramme mais surtout la disparition des hautes fréquences (au-dessus de 20kHz environ). C'est encore plus visible lorsqu'on change le « Frequency Gain » (dB/dec). Prendre une valeur de 10 ou 20.

Activité (5.1.2) Fichier .ogg

Contrairement au format MP3, le format OGG Vorbis est libre, et il est considéré comme de meilleure qualité pour une compression identique¹². Ce format est beaucoup utilisé sur Wikipédia.

Question : quel est le gain de taille apporté par le format .ogg ? ce format est-il plus performant que le .mp3 ?

¹² Voir livre de Christian Brochec, page 44.

Bibliographie

1) Sur le traitement du signal :

- (livre) *Mathématiques pour le traitement du signal* M. Bergounioux, DUNOD (2010) ISBN 978-2-10-054781-4
- (livre) *Méthodes et techniques de traitement du signal*, DUNOD (2004) Jacques Max, Jean-Louis Lacoume, EAN13 : 9782100483310
- (doc. en ligne) *Cours de traitement du signal*, Patrice Brault (1999-2000) http://braultp.free.fr/Cours_signal/signal_6juin_sscmp.pdf

2) Sur Audacity :

- (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)

3) Sur les sons musicaux :

- cours de Philippe Guillaume (<http://www-gmm.insa-toulouse.fr/~guillaum/sons-musique.pdf>)
- 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>