De l'acoustique à la musique

Frédéric Faure

Université Grenoble Alpes, France

frederic.faure@univ-grenoble-alpes.fr

pour Licence de Musicologie

(version : 1 décembre 2021)

Table des matières

Dans la version électronique de ce document (format pdf), les couleurs sur le texte sont souvent des liens vers des pages de wikipedia pour avoir plus d'informations ou vers d'autres documents ou vidéos.

Introduction

Ce cours est destiné à des étudiants de musicologie, c'est à dire ayant des **bases de musique** mais on ne suppose **aucune base en science**.

L'objectif du cours est de mettre en valeur les phénomènes acoustiques et physiologiques qui sont présents dans les pratiques musicales.

Il existe une version de ce cours, disponible ici, destinée aux étudiants de musicologie et mathématique/physique qui suit le même plan mais dont l'aspect scientifique est plus approfondi. On y trouvera par exemple la présentation précise de la théorie de Fourier concernant la décomposition d'un son comme somme de sinusoïdes (sons purs), ou différents modèles physiques et mathématiques pour la propagation du son et la génération du son par un instrument de musique.

Chapitre 1 : Le son. Le son correspond aux vibrations de l'air. Le son est le vecteur de l'information musicale. Dans ce chapitre on présente certaines des caractéristiques physiques essentielles du son qui interviennent en musique. On étudiera la propagation des ondes sonores dans l'espace. On étudiera comment un signal sonore (i.e. variations de pression) peut être capté et mesuré en un point donné de l'espace, par un microphone par exemple, pour en faire un signal.

Chapitre 2 : Les signaux sonores. Ce chapitre concerne l'étude des signaux sonores que l'on appelle la **théorie du signal**. On étudiera les signaux qui sont périodique en temps, qui ont de l'importance pour la suite et que l'on appellera **"note musicale** ". Leur importance vient du fait qu'ils sont produits par des phénomènes périodiques comme dans la voix humaine, donc très présents en musique.

Chapitre 3 : perception du son. La perception du son (par les humains) se fait grâce au système auditif qui comporte les oreilles mais aussi des circuits neuronaux spécifiques. L'analyse du son commence par l'oreille. Cette partie est bien étudiée et assez bien comprise : l'onde sonore est transmise dans la cochlée où il y a une membrane et des milliers de cils, chacun étant un résonateur sensible à une étroite plage de hauteur de son (fréquence). Si un cil se met en vibration par résonance, il excite un neurone. L'information est ainsi transmise au cerveau. Ensuite l'analyse est effectuée par le cerveau de façon inconsciente. Cette partie est encore très mal connue, voire **totalement inconnue**. Par des expériences cognitives on peut cependant observer les caractéristiques du son que la conscience perçoit (i.e. le résultat des traitements inconscients). Pour les signaux périodiques, i.e. notes musicales, on a une perception particulière sous forme de **timbre**. Cela est mis en évidence par des expériences d'illusion auditives. De plus pour plusieurs notes musicales de fréquences différentes on ressent comme "consonant" des rapports entres ces fréquences qui sont des petits rationnels et qui correspondent aux intervalles de base de la musique (octaves, quintes, quartes, tierces etc). On parlera aussi de la perception du rythme.

Références et liens conseillées :

- Différents Documents liés au cours.
- Livre (Benson, n.d., p.197), and its web site “Music: a Mathematical Offering”,
- Livre Schnupp *et al.* (2011) “{} Auditory neuroscience: Making sense of sound “ and its web site Auditoryneuroscience web site
- Livre Handbook of Acoustic,(Schroeder *et al.*, 2007).
- Exposé "Voix mathématiques et musique" du 11 septembre 2015 pour la journée de rentrée de l’institut Fourier.
- Page de wikipedia sur l’acoustique musicale

Logiciels gratuits (sur ordinateur) conseillé pour ce cours :

- Audacity pour la gestion de fichiers audio.

Applications Android gratuites conseillées pour ce cours :

- sonomètre : mesure l’intensité d’un son en décibels.
- Lexis Audio Editor : permet de visualiser le signal d’un son enregistré.
- Spectroid : permet de faire une transformée temps-fréquence d’un signal sonore en temps réel.
  - Paramétrage recommandé : Audio/taille du FFT : 1024bins, Affichage : linéaire en fréquence.
- Function generator ou Frequency sound generator

Chapitre 1

Les ondes sonores

1.1 Introduction

Video de cette section.

Le son est le vecteur de l'information musicale. Dans ce chapitre on présente certaines des caractéristiques physiques essentielles du son qui interviennent en musique. On étudiera la propagation des ondes sonores dans l'espace.

Le son est une onde qui se propage dans la matière. Ce peut être dans l'air (gaz), dans l'eau ou dans les solides ou toute autre forme de matière. Le son correspond aux vibrations de cette matière dans un certain régime de fréquences et d'amplitudes. Le son est très utilisé par les organismes vivants pour communiquer, en particulier parmi l'espèce humaine dans l'air, avec la voix et la musique, mais aussi par les poissons et de nombreux animaux dans l'eau, voir video de L. Ballesta.

Dans ce cours on s'intéressera essentiellement aux ondes sonores dans l'air à pression et température ambiante, ce qui est la situation courante en musique.

L'air est un fluide gazeux légèrement visqueux. Le comportement du gaz et parfois simple, laminaire, mais peut être parfois très complexe, on parle alors de turbulence. La propagation du son sous forme ondulatoire fait partie des ces cas simples que l'on va décrire. Par contre la production des tourbillons à l'embouchure d'une flûte par exemple fait partie des cas complexes pas encore bien compris.

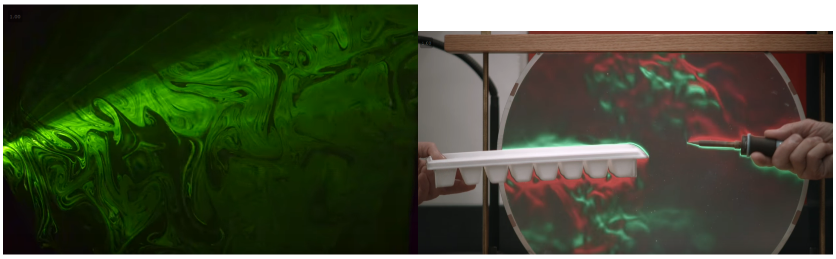


Figure 1.1.1 – (1) Turbulence de l'air éclairé par un faisceau laser. (2) Turbulence observée par l'effet Schlieren.

— Visualisation des mouvements de l'air et des ondes sonores :
  — Turbulence, vortex rings

1.1.1 Le son

Définition 1.1.1. Le **son** est la vibration de la matière dans un régime de petites amplitudes, appelées "**ondes sonores**".

Remarque 1.1.2. Plus précisément, lorsque la matière est un **fluide** (où les molécules sont libres de se déplacer, exemple : air, eau), le son est la **vibration de la pression**. Dans la matière **solide** (où les atomes sont figés les uns par rapport aux autres), les vibrations s'appellent les **"ondes élastiques"** ou "**ondes sismiques**". Par exemple : dans la croûte terrestre, dans une plaque de métal, dans le bois, etc.

Remarque 1.1.3. Avec les capteurs sismiques, loin des routes, mais au centre des continents, les vibrations appelé "bruit sismique" sont dues aux fracas des vagues sur la côtes.

Importance pour les humains : La voix humaine est un son particulièrement important car il permet de communiquer. La voix est apparue entre -2 millions d'années (sons) et-50000 ans (parole). Des recherches récentes font remonter l'apparition de la parole à $-20\,10^6$ ans. La parole a un rôle social pour la communication, pour palabrer, pour l'échange d'informations ...

Plus généralement, l'être humain est réceptif aux sons de la nature (et de la ville) qui l'entoure : vent, rivière, bruit des animaux, bruit du feu, tram, etc. On cherche toujours à comprendre et interpréter les sons que l'on perçoit.

On intéressera donc plus particulièrement aux ondes sonores **dans l'air**.

1.2 L'air

1.2.1 Composition de l'air

L'air est un **gaz** c'est à dire composé de molécules qui bougent **librement** mais s'entrechoquent en échangeant de l'énergie via des changements de vitesse. L'énergie totale est conservée.

Quelques ordres de grandeurs dans l'air avec des conditions normales (température 20C° et pression de une atmosphère)

— L'air est composé essentiellement de molécules di-oxygène (21% du nombre de molécules) et de di–azote (79%).

— La taille d'une molécule est de l'ordre de $R = 10^{-10}$m.

— La distance moyenne entre les molécules de l'air est $d = 35\,R$ (car la densité est $\frac{N}{V} = 2.5\,10^{25}$ molecules/m3)

— La vitesse moyenne d'une molécule de di-oxygène est

$$v = 280\,m/s \tag{1.2.1}$$

(car $v = \left(\frac{RT}{M_{O_2}}\right)^{1/2}$ avec $R = 8.31J/K$ $T = 273K$, $M_{O_2} = 2 \times$ 16g/mole)

— La distance moyenne parcourue entre deux chocs est $l = 3.10^{-7}m = 3000\,R$, appelée libre parcourt moyen.

— Le nombre moyen de collisions pour une molécule est $n_{col.} = \frac{v}{l} = 10^9$ collisions/seconde.

Exercice 1.2.1. Si on zoom par la pensée un cube d'air de $L = 1cm$ de côté, de sorte à transformer la taille d'une molécule $R = 10^{-10}m$ ci-dessus en taille macroscopique : $R' = 1cm$, donner le facteur de zoom qu'il faut faire et les valeurs correspondantes :

Tailles microscopiques :	$R = 10^{-10}$m	$d = 35R$	$L = 1$cm	$l = 3000\,R$
Tailles macroscopique	$R' = 1$cm	$d' = ?$	$L' = ?$	$l' =$
après un zoom d'un facteur $\times$?				

Solution 1.2.2. On obtient un facteur de zoom $\frac{R'}{R} = \frac{10^{-2}}{10^{-10}} = 10^8$ et les valeurs

Tailles microscopiques :	$R = 10^{-10}$m	$d = 35R$	$L = 1$cm	$l = 3000\,R$
Tailles macroscopique	$R' = 1$cm	$d' = 35$cm	$L' = 10^8 cm = 1000$km	$l' = 3000cm = 30$m
après un zoom d'un facteur $\times 10^8$				

Ainsi le cube à la taille $L' = 1000$km de la France !

1.2.2 L'air est un fluide à l'échelle humaine

L'échelle des atomes et molécules s'appelle l'échelle microscopique. L'échelle humaine s'appelle l'**échelle macroscopique**. Il y a une échelle intermédiaire appelée **échelle mésoscopique** à partir de laquelle laquelle l'air est un fluide, c'est à dire que l'on ne perçoit que le comportement collectif des molécule et non plus individuel. A cette échelle la population de molécules a un comportement collectif chaotique appelé **état d'équilibre**.

— Par exemple le fluide peut être au repos (pas de vent) même si les molécules sont en mouvement permanent.
— Pour l'air, l'échelle mésoscopique est précisément à partir de $D = 10^3 d = 3\,10^{-6} m = 3\mu m > l$, de sorte qu'un cube de côté D contient 10^9 molécules qui est grand et D est plus grand que le libre parcourt moyen l. (Dans l'exercice ci-dessus, après le zoom, cela donne $D' = 10^3 d' = 350\text{m}$)
— Par exemple un gaz et un liquide sont des fluides.
— L'état d'équilibre local, résultat de ce comportement collectif est **caractérisé par des grandeurs locales** (bien connues) que sont :
  — la **température** T, qui correspond à la vitesse moyenne des molécules
  — la **pression** P, qui est inversement proportionnelle au volume libre autour de chaque molécule. (d'après la loi des gaz parfaits $PV = nRT$)

Exemple 1.2.3. dans l'atmosphère, la température T et la pression P diminuent avec l'altitude.

1.2.3 La pression de l'air

L'atmosphère a une épaisseur moyenne de 10 km, et la masse de la colonne d'air au dessus d'une surface de la Terre est de

$$1\frac{\mathrm{kg}}{\mathrm{cm}^2} = 100\frac{\mathrm{kg}}{\mathrm{dm}^2} \tag{1.2.2}$$

Cela correspond à une masse d'eau de hauteur 10m.

La densité de l'air au sol terrestre est de $1.2\,\mathrm{kg}/m^3$.

Exemple 1.2.4. "Le verre retourné" : voici une démonstration du poids de l'atmosphère. En principe, cette expérience marche tant que la hauteur du verre est inférieur à 10m.

— La **pression atmosphérique** est la force F par surface exercée par l'air à la surface de la Terre. On calcule sa valeur à partir de (1.2.2) :

$$P = \frac{F}{S} \approx \frac{1\mathrm{kg}\ g}{1\mathrm{cm}^2} = \frac{9.81N}{10^{-4}m^2} \approx 10^5\frac{N}{m^2} = 10^5\mathrm{Pa} = 1\,\mathrm{atmosph.}$$

1.3 Les ondes sonores

1.3.1 La vitesse des ondes sonores

Exemple 1.3.1. Les amplitudes de vibration de pression pour un son de trompette de puissance 90 dB (très fort) est de l'ordre de $\Delta P = 10^{-5}$ atmosph.

Définition 1.3.2. L'**énergie** est une grandeur que possède chaque objet physique (et ondes) et qui s'échange entre objets, mais le total est toujours **conservé**.

Autrement dit, l'énergie prend différents aspects dans la nature mais ni n'apparait, ni ne disparait. Exemples : énergie de vitesse (énergie cinétique), énergie de rayonnement, chaleur (agitation désordonnées des atomes), énergie de masse ($E = mc^2$), etc

Proposition 1.3.3. *Les ondes sonores possèdent une énergie $\mathcal{E}$ proportionelle au carré de leur amplitude A et* ***dans un milieu homogène****, elles se* ***propagent en ligne droite****, à une certaine vitesse c_{son} appelée* ***vitesse du son****. La* ***vitesse du son dans l'air*** *est*

$$c_{\text{son}} = 343\, m/s.$$

Une partie de l'énergie est ***dissipée****, $\frac{\Delta\mathcal{E}}{\mathcal{E}} = 0.5$ par km, convertit sous forme de chaleur.*

Remarque 1.3.4. La formule de la vitesse du son dans l'air est

$$c = \left(\frac{\gamma RT}{M_{air}}\right)^{1/2}$$

avec $\gamma = 7/5$, $R = 8.31$, $T = 293$, $M_{air} = 29$ g/mole.

Remarque 1.3.5. Ainsi le son parcourt $34cm$ en $1ms$ ou $17m$ en 50ms qui est le temps de réaction d'un neurone et donc la limite de la **latence en musique**.

Pour cette raison, il ne faut pas que les musiciens qui jouent ensemble soient trop éloignés les uns des autres, sinon le retard du son est perceptible.

Remarque 1.3.6. La vitesse du son est comparable à la vitesse moyenne d'une molécule (1.2.1). Cela est général aux gaz.

Remarque 1.3.7. Il a été observé récemment (Pour la science juillet 2021) que à la naissance, les chauves souris connaissent précisément la vitesse du son, de façon innée. Des papillons de nuit "arctiinae" captent et émettent (pour brouiller) les sons radars des chauves souris (Ratcliffe and Fullard, 2005).

Exercice 1.3.8. Convertir la valeur de dissipation d'énergie $\frac{\Delta \mathcal{E}}{\mathcal{E}} = 0.5$ par km, en dB par km ?

Solution 1.3.9. $x = 0.5 = \frac{1}{2} = \frac{1}{10^{0.3}} = 10^{-0.3}$ par km, soit $\mathrm{dB}\,(\mathrm{x}) = 10 \times (-0.3) =-$ 3dB par km.

Exemple 1.3.10. Un son de trompette de 90dB (très fort) sur une durée $t = 1s$ correspond à une énergie de $\mathcal{E} = 10^{-3} J/m^2$. Pour comparaison, cette énergie permet de soulever une masse $m = 1g$ (1 pincée de sel) sur une hauteur de $h = 10cm$, d'après la formule $\mathcal{E} = mgh$ avec $g = 9.81m/s^2$.

Exemple 1.3.11. La vitesse du son dans l'eau est

$$c_{\text{son dans l'eau}} = 1.5\,\text{km}/s$$

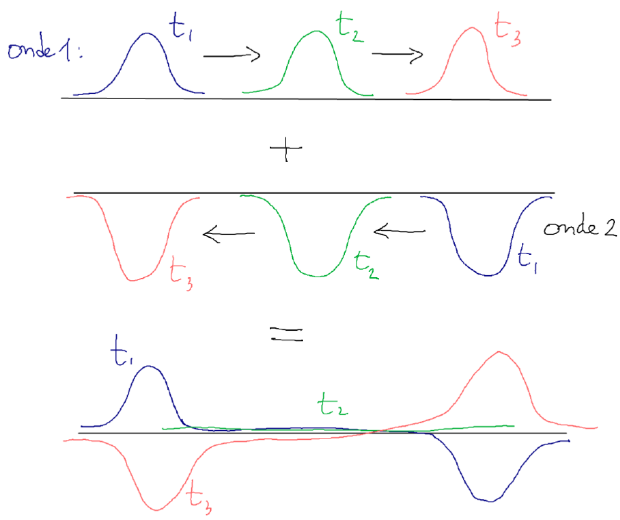
1.3.2 Le principe de superposition

C'est une **propriété des ondes en général**, quelque soit leur nature (et vient du fait qu'elles sont modélisées par une équation linéaire).

Proposition 1.3.12. *Le principe de superposition stipule que pour toute forme d'ondes :*

$$Evolution(onde1 + onde2) = Evolution(onde1) + Evolution(onde2)$$

Exemple 1.3.13. Dans l'exemple suivant, à l'instant intermédiaire t_2, l'onde est nulle :



Cet effet est bien utilisé par les **casques anti-bruits** par exemple.

— **Expérience d'interférences destructives** à 1050Hz ($\lambda = 30$cm) : explication, expérience.

— **Interférences d'ondes sonores (ultra sons)** observées par effet Schlieren et par éclairage stroboscopique :

Les ondes sont émises à gauche et se réfléchissent sur la plaque. On observe des figures d'interférence.

— **Ondes de surface** : interferences

1.3.3 Autres exemples d'ondes dans la nature

1. Le **HOLLA** dans un stade est **une onde qui se déplace**. C'est la "vibration" des humains qui se lèvent et se rassoient **mais ne se déplacent pas**.

2. **Vagues :** ce sont des "ondes de surface". Par exemple à la surface de l'océan, la vibration de hauteur de la mer, peut avoir des vitesses de l'ordre du mètre par seconde. L'amplitude est du centimètre à quelques mètres. L'énergie des ondes se déplacent, mais la matière (eau) ne se déplace pas.

3. **Ondes sonores** : vibration de pression de l'air ou de l'eau.

4. **Ondes sismiques :** vibration du sol

5. **ondes lumineuses :** (ondes électromagnétiques) vibration du champ électrique et magnétique. La vitesse est $c = 3.10^8 m/s$.

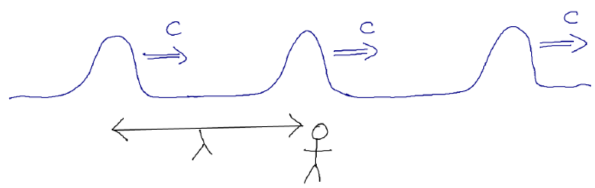
6. **Ondes quantiques :** vibration des champs de matière.

1.3.4 Longueurs d'onde, fréquence, période

Exercice 1.3.14. Si les vagues avancent à la **vitesse** $c = 2m/s$, avec une distance entre les sommets des vagues, appelée **longueur d'onde** $\lambda = 20m$, quelle est la **période** T (i.e. temps entre deux sommets) en un point donné ? quelle est la **fréquence** $f = \frac{1}{T}$ (i.e. nombre de sommets par seconde) ?

Pour cela, utiliser la **formule suivante à connaître** (en pensant à km/h, on retient que la vitesse c est une longueur λ divisée par un temps T) :

$$c = \frac{\lambda}{T}$$



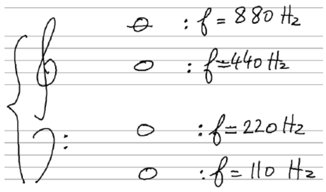
Solution 1.3.15. Donc

$$T = \frac{\lambda}{c} = \frac{20m}{2m/s} = 10s$$

et

$$f = \frac{1}{T} = \frac{1}{10s} = 0.1s^{-1} = 0.1\text{Hz}$$

Exercice 1.3.16. Calculer la longueur d'onde λ pour une onde sonore de fréquence $f = 440\text{Hz}$, (la du diapason) et pour les autres la de la figure ?



Solution 1.3.17. D'après $c = \frac{\lambda}{T} = \lambda f$, on déduit

$$\lambda = \frac{c}{f} = \frac{340m/s}{440s^{-1}} = 0.80m = 80cm$$

Exercice 1.3.18. Les **dauphins** utilisent des sons de fréquences $f \approx 150\,\mathrm{kHz}$ pour localiser les poissons (proies). Quelle est la longueur d'onde λ correspondante ? Même questions pour les **Marsouins** ($f \geq 100$kHz et les **Orques** $f \leq 100$kHz) ?

Solution 1.3.19. Dans l'eau, la vitesse du son est $c = 1500\,\mathrm{m}/s$. Donc pour les dauphins

$$\lambda = \frac{c}{f} = \frac{1500m/s}{150\,000s^{-1}} = 0.01m = 1cm$$

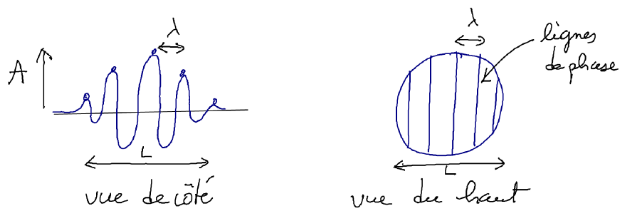
Cette petite longueur d'onde leur permet de localiser les petits poissons par écholocalisation pour se nourrir. Pour les marsouins,

$$\lambda = \frac{c}{f} \leq \frac{1500m/s}{100\,000s^{-1}} = 0.015m = 1.5cm$$

Pour les orques, $\lambda = \frac{c}{f} \geq 1.5cm$.

1.3.5 Décomposition en paquets d'ondes

Définition 1.3.20. Un **paquet d'onde** est une onde particulière de taille donnée L et caractérisée par sa **position** (de son centre), la **longueur d'onde** λ des oscillations spatiales internes et la direction de ces oscillations (lignes ou plans de phase), et son **amplitude** A. En musique, un paquet d'onde correspond à une **note de musique**, qui est un "**son pur**" (voir plus loin).



Exercice 1.3.21. Pour une note de durée $t = \frac{1}{5}s = 0.2s$, calculer la taille L du paquet d'onde produit ?

Solution 1.3.22. $c = \frac{L}{t}$ donc $L = ct = 343m/s \times 0.2s = 70m$.

Proposition 1.3.23. ***Toute onde se décompose*** *en somme (superposition) de* ***paquets d'ondes****, de position, oscillations et amplitudes différentes.*

La proposition précédente est essentielle en science et découle de la théorie de Joseph Fourier.

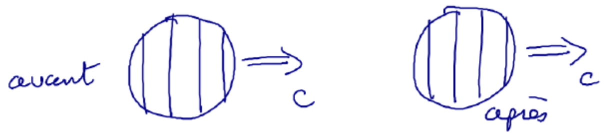
1.3.6 Trajectoire des ondes sonores

1.3.6.1 Dans un milieu homogène

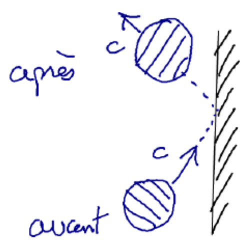
Un **milieu homogène** est un milieu où la **vitesse** c_{son} **du son est constante**, i.e. ne dépend pas du point. Pour cela il faut que la température et la pression soit constantes.

On a déjà donné la proposition suivante :

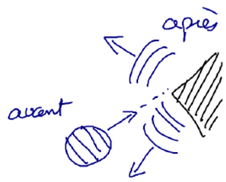
Proposition 1.3.24. *Dans un* ***milieu homogène****, i.e. la vitesse* c_{son} *est constante, alors les paquets d'ondes se propagent en* ***ligne droite*** *perpendiculairement aux lignes (ou plan) de phases. Une faible partie de l'énergie est dissipée pendant la propagation.*



Proposition 1.3.25. *Sur un* ***obstacle plan lisse*** *ou de* ***longueur de variation grande devant la longueur d'onde****, un paquet d'onde* ***se réfléchit*** *et une partie est absorbée par la paroi.*



Proposition 1.3.26. *Sur un* ***obstacle de variation courte devant la longueur d'onde****, un paquet d'onde* ***diffuse****, i.e. produit plusieurs paquets d'ondes avec des amplitudes variables.*

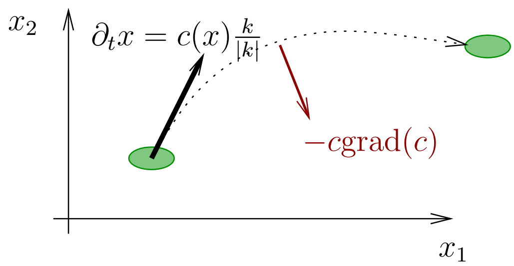


Exemple 1.3.27. Par exemple si l'obstacle est très petit devant la longueur d'onde comme un point, la diffusion est isotrope (on parle de diffusion "s") ou diffusion de Rayleigh. Cela explique la couleur du ciel.

1.3.6.2 Dans un milieu inhomogène

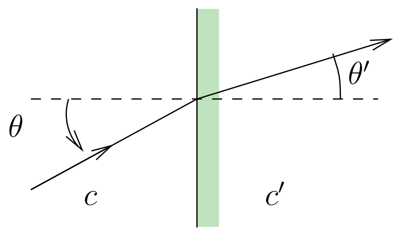
Un **milieu inhomogène** est un milieu où la vitesse c_{son} du son n'est constante, i.e. dépend du point. Ce peut être du à la température ou la pression qui varient.

Proposition 1.3.28. *Dans un **milieu inhomogène**, i.e. si la vitesse $c_{\text{son}}(x)$ dépend du point de façon lente par rapport à la longueur d'onde, alors **la trajectoire d'un paquet d'onde est déviée vers les zones où** $c_{\text{son}}(x)$ **décroit** (i.e. $-\text{grad}c_{\text{son}}$ agit comme une force).*



Cette dernière proposition a de nombreuses conséquences.

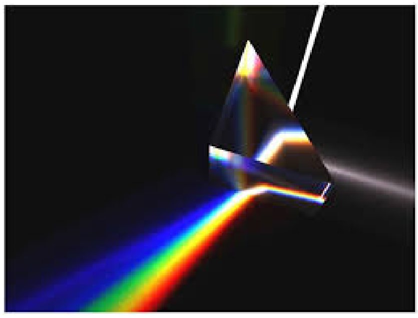
Réfraction entre deux milieux (important pour les ondes lumineuses) Exemple si $c' < c$ (cas du verre par rapport à l'air) alors :

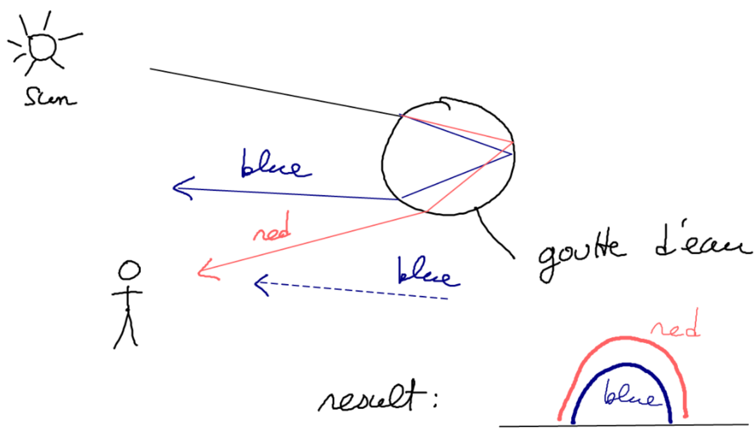


La formule précise appelée Loi de Snell (1621)-Descartes (1637) (trouvée avant par Ibn Sahl (983) à Bagdad) est :

$$\frac{1}{c}\sin\theta = \frac{1}{c'}\sin\theta'$$

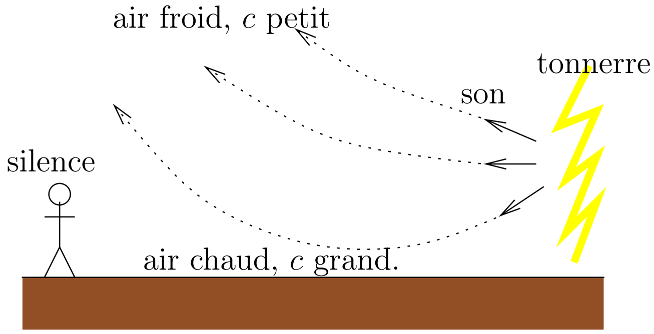
Arc en ciel (important pour les ondes lumineuses) Car dans le verre ou dans l'eau, $c_{\text{bleu}} < c_{\text{rouge}}$, ainsi :



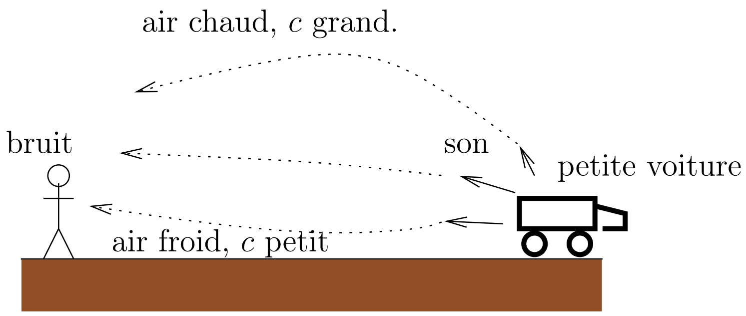


“Mirages acoustique”

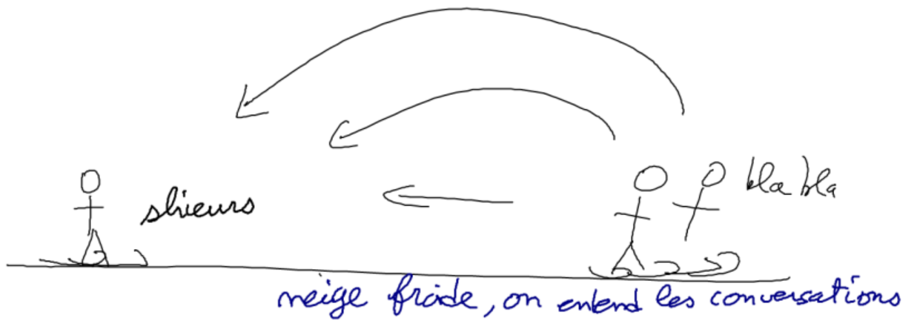
— En été, le soir, l’air est plus chaud au sol que en altitude et donc $c(x)$ décroît aussi avec l’altitude, à pression fixée. Les trajectoires sont donc déviées vers le haut. Il est possible d’observer des orages aux loin, mais de **ne pas entendre de tonnerre**. On parle “d’**éclairs de chaleur**”.



— En hiver, parfois (situation d’inversion de température), l’air est plus froid au sol que en altitude et donc $c(x)$ croit aussi avec l’altitude. Les trajectoires sont donc déviées vers le bas. Il est possible d’**entendre assez fortement une voiture** qui passe pourtant loin dans une campagne calme.



— Sur la neige, entendre des conversations au loin :



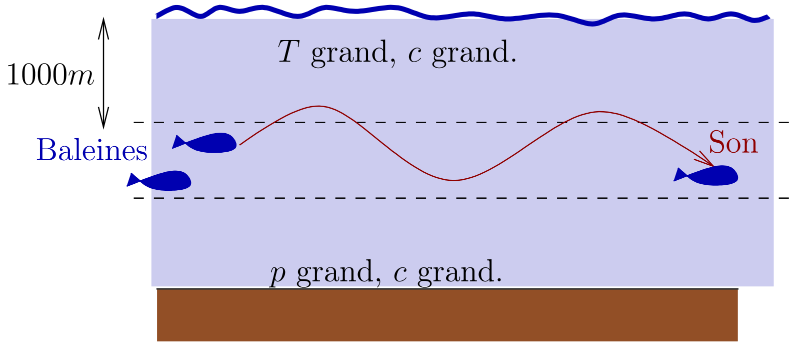
de même à la surface de l'océan froid et calme, dans des barques éloignées, on entend bien les conversations.

— De même dans l'océan si la surface est plus chaude, les sons sous marins n'atteignent pas la surface. Cet effet a été observé sans être compris en 1941, lorsque les américains écoutaient les sous marins depuis la surface et que l'après midi (lorsque la surface est plus chaude) le son des sous marins éloignés ne leur parvenait plus. Ce fut appelé "l'effet de l'après midi", rapidement expliqué par des physiciens.

— Si $c(x)$ varie avec l'altitude mais est minimum dans une strate intermédiaire alors cette strate va agir comme un **canal conducteur** ou "puits de potentiel".

  — L'histoire raconte que les coups de canons pour l'enterrement de la reine victoria le 22/01/1901 n'ont pas été entendus en Angleterre mais en écosse à 400km de distance.

  — **Dans l'océan**, vers une profondeur de 1000m, la valeur de $c(z)$ est minimale, ainsi les rayons sonores sont capturés dans ce canal appelé **canal SOFAR** et découvert en 1943. Les baleines communiquent entre elles à longues distances (1000 km) en se plaçant à cette profondeur. De plus elle envoient des sons qui font échos sur des montagnes sous-marines et leur permette de se localiser dans l'océan. Il semble en effet qu'elles mémorise les positions des ces montagnes dans l'océan. Vidéo, chant des baleines à bosse.



Vidéos

— Cette vidéo est une simulation numérique d'un paquet d'onde dans le stade.

— Retournement temporel d'une onde de surface.

— Ondes (solutions numériques) et représentation dans l'espace des phases. Correspondance classique-ondulatoire. Ondes quantique.

1.3.7 Effet Doppler

On parle d'effet Doppler si la **source sonore est en mouvement**, et/ou le **récepteur est en mouvement par rapport au fluide**. La conséquence est que la **longueur d'onde est différente, ainsi que la fréquence percue.**

Voici des expériences pour expliquer cela.

— Vidéos dans une cuve à ondes de surface.

  — Effet Doppler subsonique

— Effet Doppler supersonique

— Le mur du son

Exemple 1.3.29. Expériences connues : le son d'une sirène d'une voiture en mouvement. Le mur du son créé par un avion supersonique.

1.4 Résonateurs, résonances, ondes stationnaires

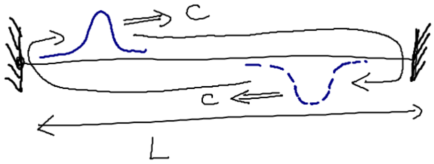
Définition 1.4.1. Un **cavité résonnante** est un volume dans lequel les ondes sont piégées, i.e. se réfléchissent sur les parois avec peu de perte énergétique.

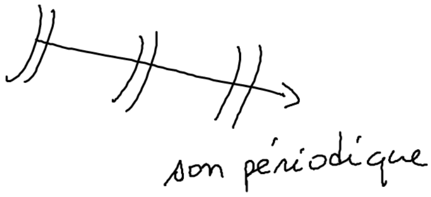
Dans une cavité résonnante, **certaines ondes sont invariantes** (oscillent sur place, voir plus loin ou ici) : leur fréquence f et leur forme dépend de la géométrie de la cavité. Ces ondes sont appelées "**résonances**", "ondes stationnaires", "**modes propres**" de la cavité. Leurs fréquences sont appelées "**fréquences propres**", "**spectre**" de la cavité.

1.4.1 Exemple simple mais important pour la musique : la cavité unidimensionnelle (i.e. selon un axe)

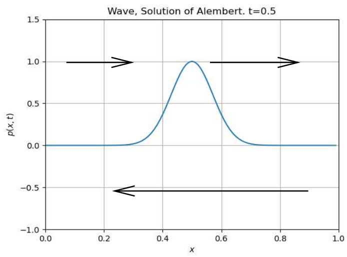
Considérons une cavité unidimensionelle de longueur L. Les ondes sonores se déplacent à $c = 343\text{m/s}$ et se réfléchissent sur les bords. Donc le mouvement est périodique, sur une longueur $2 \times L$, et donc de période T donnée par

$$c = \frac{2L}{T} = 2Lf \Leftrightarrow T = \frac{2L}{c} \Leftrightarrow f = \frac{c}{2L}.$$





Exemple 1.4.2. Voici un exemple d'une onde, vidéo fait par ce programme onde_1dim_segment_superposition.py



Considérons un tuyau d'orgue ou de flute traversière, de longueur $L = 39$ cm (en fait plus court pour des raisons expliquées plus loin). Alors la fréquence propre ou **fréquence de résonance** est :

$$f = \frac{c}{2L} = \frac{343\, m/s}{2 \times 0.39\, m} = 440\,\mathrm{Hz}$$

(c'est un la).

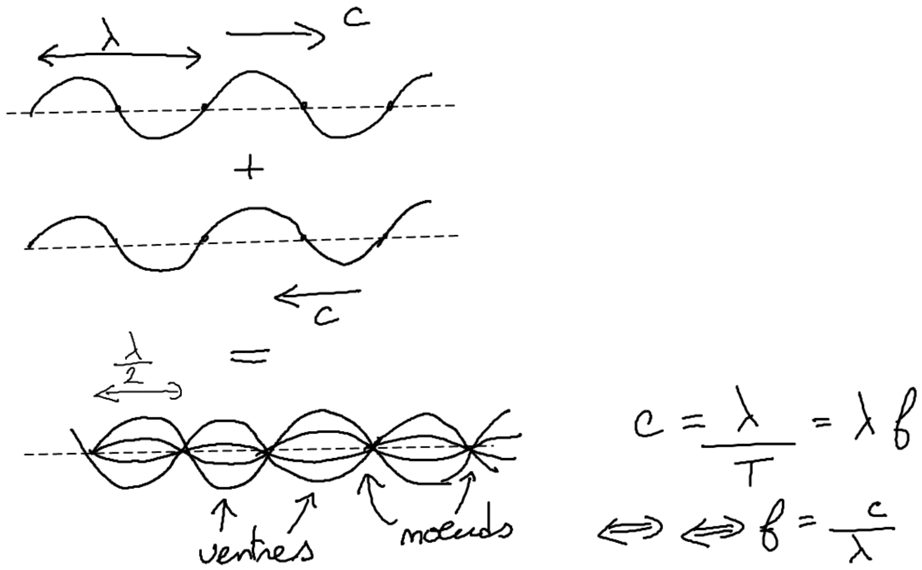
Exercice 1.4.3. Trouver la longueur du tuyau L pour produire un C de fréquence $f_C = 260\mathrm{Hz}$?

Solution 1.4.4. On écrit $f = \frac{c}{2L}$ donc

$$L = \frac{c}{2f} = \frac{343\, m/s}{2 \times 260 s^{-1}} = 0.65\, m = 65\, cm$$

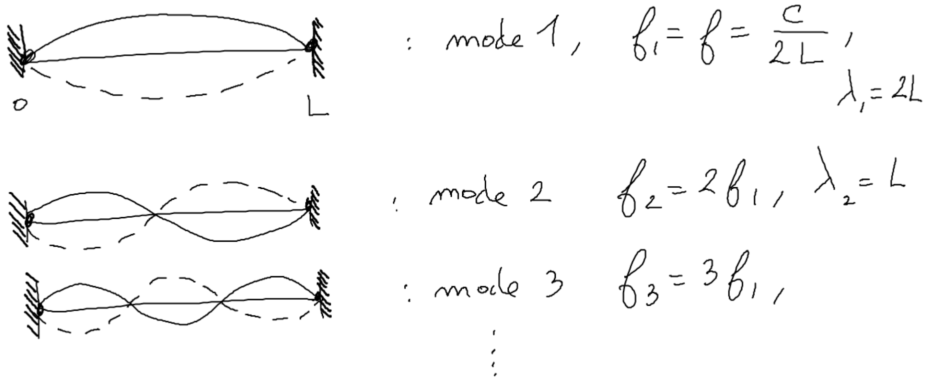
1.4.1.1 Ondes stationnaires (modes propres) de la cavité unidimensionnelle :

— La superposition de deux ondes progressives sinusoidales gauche et droite de longueur d'onde λ donne une onde stationnaire avec des "noeuds" distants de $\lambda/2$, voir cette vidéo :



—

— Par conséquent une **onde stationnaire** d'une cavité unidimensionnelle de longueur L peuvent être pensés comme **une onde progressive qui fait aller et retour** en se superposant et avec la **condition de s'annuler aux extrémités**. Ce sont donc :



avec des fréquences

$$f_1 = f = \frac{c}{2L} \quad \text{: fréquence fondamentale}$$

$$f_2 = 2f \quad \text{: deuxieme harmonique}$$

$$f_3 = 3f \quad \text{: troisieme harmonique}$$

etc

— Le fait que la liste des fréquences propres est précisément les multiples de f, c'est à dire $(f, 2f, 3f, \ldots)$ on dit que le **spectre est harmonique**. Cela découle du fait que l'onde a un mouvement périodique. On verra que les signaux périodiques sont appréciés par notre perception. delà vient le terme "harmonique" comme "harmonieux" et l'utilisation de cordes tendues (cavité 1dim) en musique.

1.4.1.2 Exemple de la corde tendue (guitare, piano, etc)

La corde est un objet solide, donc les ondes d'une corde sont appelées des **ondes élastiques**.

Exercice 1.4.5. On souhaite qu'une corde de guitare (3/4) de longueur $L = 56\text{cm}$, produise un la de fréquence $f = 110\text{Hz}$. Calculer la vitesse v de l'onde élastique sur la corde.

Solution 1.4.6. On a

$$v = \frac{2L}{T} = 2Lf = 2 \times 0.56\text{m} \times 110s^{-1} = 123\,\text{m}/s$$

Exercice 1.4.7. Considérons la corde de mi_4, E_4, $f_{E_4} = 330\text{Hz}$ de la guitare de longueur $L = 56\text{cm}$. A quelle distance L' placer le doigt, pour produire son octave E_5 de fréquence $f_{E_5} = 2f_{E_4} = 660\text{Hz}$? et pour produire la quinte B_4 de fréquence $f_{B_4} = \frac{3}{2} \times f_{E_4} = 495\text{Hz}$?

Solution 1.4.8. D'après la formule ci-dessus

$$f = \frac{v}{2L}$$

on observe que f est inversement proportionnel à L. Ainsi pour doubler la fréquence f il faut diviser par 2 la longueur. Donc pour obtenir $f_{E_5} = 2f_{E_4}$ il faut $L' = \frac{1}{2}L = \frac{1}{2} \times 56 = 28\text{cm}$. Pour obtenir $f_{B_4} = \frac{3}{2}f_{E_4}$ il faut $L' = \frac{2}{3}L = \frac{2}{3} \times 56 = 37,3\text{cm}$.

La **vitesse des ondes** sur une corde tendue de tension T (en N) et masse par unité de longueur μ est (Par Vincinzo Galileo, vers 1570, le père de G. Galilée)

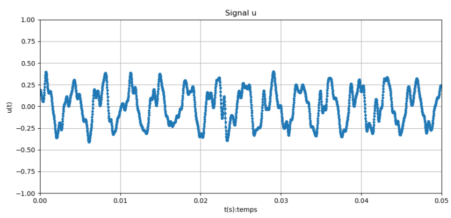
$$v = \sqrt{\frac{T}{\mu}}$$

Cette formule montre ce qui est bien connu des musiciens : la vitesse v (et donc la fréquence f) augmente si la tension de la corde augmente. La vitesse (et donc la fréquence) diminue si la masse de la corde augmente.

1.4.2 Cavité bi-dimensionelle ou autre

Par exemple, une plaque vibrante, ou la table d'harmonie d'un violon ou d'un piano, une pièce, un bloc de matière, etc. Dans ces cas le spectre de fréquences propre est spécifique à la forme de l'objet (ou de la pièce) mais ce ne sont pas les multiples d'une fréquence f donnée, i.e. le spectre n'est pas harmonique, pas appréciable à l'écoute en général (sauf si il s'approche d'un spectre harmonique, voir la conception de cloches ou cymbales, etc).

Exemple 1.4.9. Voici le signal mesuré en un point d'une plaque carré, et le son produit audio_2D.wav avec une fondamentale à $f_0 = 440$Hz, par ce programme sinus_sum_general.py. On le perçoit comme non harmonieux, car le signal est non périodique, le spectre est non harmonique.



1.4.3 Acoustique d'une salle

— Pour une salle de concert, on évite les parois lisses et planes sur lesquelles l'onde se réfléchie parfaitement, ex : mur de salle de bain.

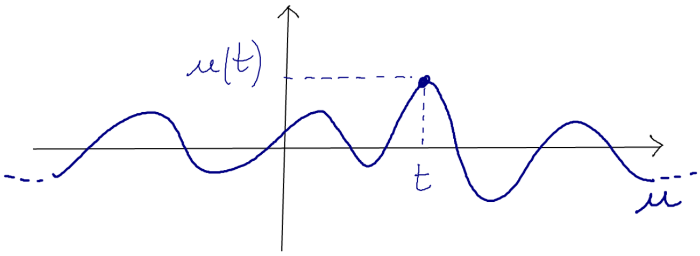
- On évite aussi des parois trop absorbantes de l'énergie de l'onde, qui l'empêche de se réfléchir.
- On préfère des parois qui diffusent les ondes. Pour cela on peut utiliser une structure aléatoire appelée diffuseur de Schroeder (1975), voir ici. Pour quelle raison ? peut être que cette diffusion aléatoire ressemble à la diffusion du son dans la forêt sur les troncs d'arbres ?
- Exemples : diffusion du son
  - dans une église,
  - dans une grotte : voir ce documentaire sur arte qui explique que les points rouge (et mains rouges) marquent les zones de résonances particulières.
- Les algorithmes de "**reverb**" sont modélisés à base de phénomènes de diffusion multiples aléatoires

Chapitre 2

Les signaux

2.1 Signal sonore

Définition 2.1.1. Un **signal sonore** est une suite de valeurs $u(t)$ pour chaque instant t, correspondant par exemple à la **fluctuation de pression** au point d'entrée d'un microphone.



2.1.1 Échantillonnage d'un signal

Éventuellement, lire avant la section A.4 sur le codage des nombres en binaire.

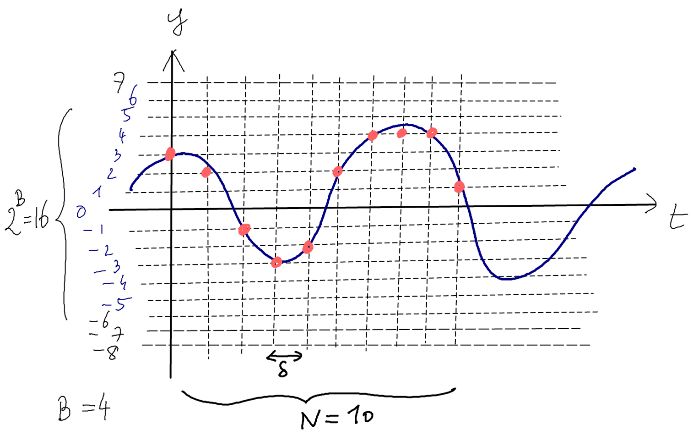
En pratique, le signal sonore est **échantillonné** par une **carte son** (qui permet la conversion analogique/numérique des signaux).

Définition 2.1.2. Un **signal échantillonné** est :

— un nombre fini N de valeurs du signal,

— à des d'instants espacés d'une durée $\delta > 0$ appelé **période d'échantillonnage**, par exemple $\delta = \frac{1}{44100}s$, (ainsi la fréquence d'échantillonnage $f_e = \frac{1}{\delta} = 44100\text{Hz}$ est supérieure à 20000Hz qui est la limite supérieure des fréquences humainement audibles).

— Chaque valeur est codée avec un nombre B de bits. $B = 8, 16, 24, 32, 64$ de sorte qu'une valeur est un entier dans l'intervalle $\left[-2^{B-1}, 2^{B-1}-1\right]$. Ce codage appelé Complément à deux.



Exemple 2.1.3. $N = 10^5$ valeurs, période d'échantillonnage $\delta = \frac{1}{44100}s.$, $B = 4$ donne $2^3 = 8$. $B = 8$ donne $2^7 = 128$, $B = 32$ donne $2^{31} \approx 10^9$.

Exercice 2.1.4. Combien y a t-il d'échantillons N dans un signal de durée $D = 10$ minutes qui est échantillonné avec un pas de temps $\delta = \frac{1}{44100}s.$?

Solution 2.1.5.

$$N = \frac{D}{\delta} = \frac{10\text{mn}}{\frac{1}{44100}\text{s}} = 10 \times 60 \times 44100 = 26.46\,10^6$$

Remarque 2.1.6. Pour les représentations du signal, il peut être utile d'effectuer un changement d'échelle pour avoir des valeurs dans l'intervalle fixe $[-1, 1]$.

Exercice 2.1.7. Pour un signal codé avec $B = 16$ bits, quelles sont les valeurs entières possibles d'un échantillon u_j ? Un signal enregistré sur ordinateur, avec $N = 25\,10^6$ échantillons, codé en 16 bits occupe combien d'octets en mémoire ? (1 octet = 8 bits)

Solution 2.1.8. :

$$u_j \in \left[\!\left[-2^{16-1}, +2^{16-1} - 1 \right]\!\right] = [\![-32768, +32767]\!]$$

16bits$= 2 \times 8$ bits$=$ 2 octets. Donc le signal occupe $N \times 2 = 50\,10^6$ octets$=$ 50 Mo en mémoire.

Exemple 2.1.9. Le format d'enregistrement sur un CD audio est à la fréquence $f = 44100$Hz et avec $B = 16$ bits soit 2 octets par donnée.

2.1.2 Mesure de l'intensité en décibels

Voir la Section A.2. On rappelle que l'**intensité** ou l'énergie d'un signal d'amplitude $|u(t)|$ est donnée par

$$I(t) = (u(t))^2 = |u(t)| \times |u(t)|$$

et que l'on utilise habituellement l'échelle des **décibels** pour l'exprimer :

$$\mathrm{dB}(I) = 10\,\mathrm{Log}_{10}(I)$$

Remarque 2.1.10. On peut écrire $\mathrm{dB}(I)$ directement en fonction de l'amplitude $|u(t)|$:

$$\begin{aligned}\mathrm{dB}(I) &= 10\,\mathrm{Log}_{10}(I)\\ &= 10\,\mathrm{Log}_{10}\left(|u(t)|^2\right) = 20\,\mathrm{Log}_{10}(|u(t)|)\end{aligned}$$

Exercice 2.1.11. Si on multiplie par 2 le signal c'est à dire $u'(t) = 2u(t)$, que devient l'intensité $I(t)$ et que devient la mesure en décibels $\mathrm{dB}(I(t))$?

Solution 2.1.12. On a $I' = (u')^2 = (2u)^2 = 4u^2 = 4I$. Donc

$$\begin{aligned}dB' &= 10\,\mathrm{Log}_{10}(I') = 10\,\mathrm{Log}_{10}(4I) = 10\,\mathrm{Log}_{10}(I) + 10\,\mathrm{Log}_{10}(4)\\ &= dB + 10\,\mathrm{Log}_{10}(4) \approx dB + 10 \times 0.6 \approx dB + 6\end{aligned}$$

Retenir de cet exercice que **un facteur 2 sur l'amplitude est équivalent à une augmentation de 6 dB sur l'intensité** en dB et inversement. Or un facteur 2 correspond à décaler d'un bit l'écriture de l'amplitude en base 2.

Exercice 2.1.13. Si un signal numérique est codé sur $B = 16$ bits, quelle variation en dB (décibels) cela permet t-il entre l'amplitude la plus faible et la plus forte (on parle du **rapport signal/bruit**) ? et pour un signal codé sur $B = 32$ bits ? Comparer aux variations de notre perception (120dB).

Solution 2.1.14.

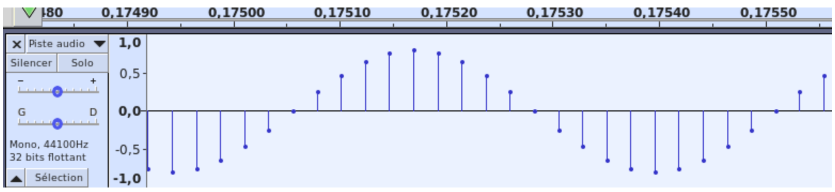
— Si on a un signal codé sur 16bits, le rapport entre l'amplitude minimale 1 et maximale $2^{16-1} = 2^{15}$ est de 2^{15} ce qui correspond à une amplitude en intensité $15 \times 6 = 90$dB (c'est le rapport signal sur bruit du codage). Conséquence pratique : dans ce cas, il est inutile d'utiliser des filtres qui filtrent plus que 90dB.

— Pour 32 bits, cela donne une amplitude $(32 - 1)\,6 = 186$ dB largement supérieure aux variations de notre perception (120dB).

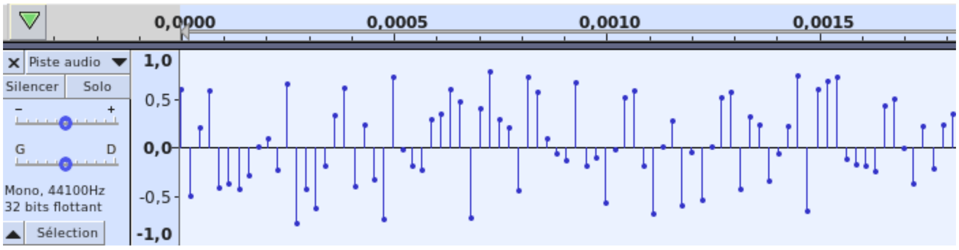
Video des exemples suivants

Exercice 2.1.15. (TP) **"Première utilisation de Audacity"**. Avec le logiciel audacity

1. Créer un signal échantillonné de **forme sinusoidale** en faisant :
   — En bas à gauche, choisir "Taux du projet" : 44100Hz.
   — Menu :Générer/Tonalité/Sinusoide, fréquence 2205Hz, amplitude 0.8, durée 0.5 sec.
   — Observer le signal (zoomer avec la souris). Voir figure ci-dessous.
   — Ecouter.
   — Calculer le nombre d'échantillons par période et vérifier.



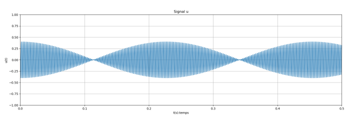
2. Créer un signal où chaque valeur $u_j \in [-1, 1]$ est choisi **au hasard** indépendamment des autres. Ce signal est appelé "**bruit blanc**". Pour cela Menu :Générer/Tonalité/Bruit/Blanc. écouter.



2.1.3 Battements

Exercice 2.1.16. Avec le logiciel audacity, avec la commande "Pistes/Ajouter nouvelle piste mono" et "Générer/tonalité", créer sur deux pistes différentes un signal de fréquence $f = 440$Hz et $f_2 = 444$ Hz. Et faire la somme sur une nouvelle piste à l'aide de la commande "Pistes/Mix/rendu vers nouvelle piste".

On obtient la superposition de deux signaux sinusoidaux de fréquence $f = 440$ Hz et $f_2 = 444$ Hz (on verra que c'est une différence de 1/6 de demi-ton, à la limite du perceptible.



Fichiers battement.wav ou sur canaux séparés : battement_stereo.wav. Le résultat est que l'on entend des **battements** donc la fréquence est

$$f_{\text{batt}} = |f_2 - f_1| = 4\,\text{Hz}$$

(i.e. on entends 4 battements par seconde).

— **Explication simple de cette formule** : si à la date $t = 0$ les deux signaux "sont en phase" c'est à dire que les maxima coïncident alors ensuite ils vont se décaler puis à nouveau se retrouver en phase à la date t lorsque le signal de plus haute fréquence $f_2 = \frac{1}{T_2}$ aura pris une avance d'une période, c'est à dire faisant $N + 1$ oscillations, lorsque le signal de fréquence $f_1 = \frac{1}{T_2}$ fait N oscillations. Cela donne $T_{\text{bat.}} = (N + 1) T_2 = NT_1$, donnant $N = \frac{T_2}{T_1 - T_2}$ et

$$f_{\text{bat}} = \frac{1}{T_{\text{bat}}} = \frac{1}{T_2} - \frac{1}{T_1} = f_2 - f_1.$$

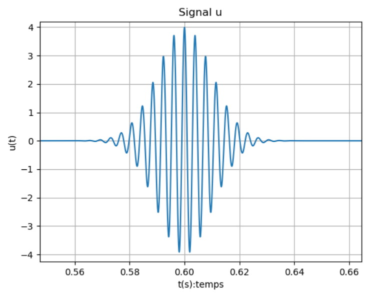
2.2 Sonogramme, transformée par ondelette, transformée de Fourier

2.2.1 Signaux élémentaires : notes de musique, paquets d'ondes Gaussiens (ou ondelettes)

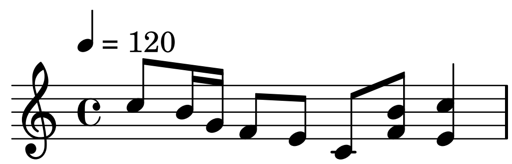
Une fonction parfaitement périodique (comme un sinus) n'est pas réaliste et non plus pas pratique. Nous allons voir qu'il est beaucoup plus intéressant de considérer des fonctions appelées **paquet d'ondes** (ou **ondelette** en théorie du signal) et que d'un point de vue musical on peut appeler **note de musique**, définie ainsi.

Définition 2.2.1. Un **paquet d'onde** est un signal particulier de durée donnée σ et caractérisé par sa **date** t (de son centre), la **période** T des oscillations internes (ou sa fréquence $f = \frac{1}{T}$) et son **amplitude** A. En musique, un paquet d'onde correspond à une **note de musique**, qui est un "**son pur**".

Exemple 2.2.2. Voici un paquet d'onde avec $t = 0.6$, $f = 260$ et $\sigma = 0.01$.



Exemple 2.2.3. Voici une mélodie crée par ordinateur, en superposant des paquets d'ondes (ou notes de musique) de durée $\sigma = 0.02$s reproduisant la mélodie suivante. Visualiser le signal et écouter wave_packet_melody.wav.



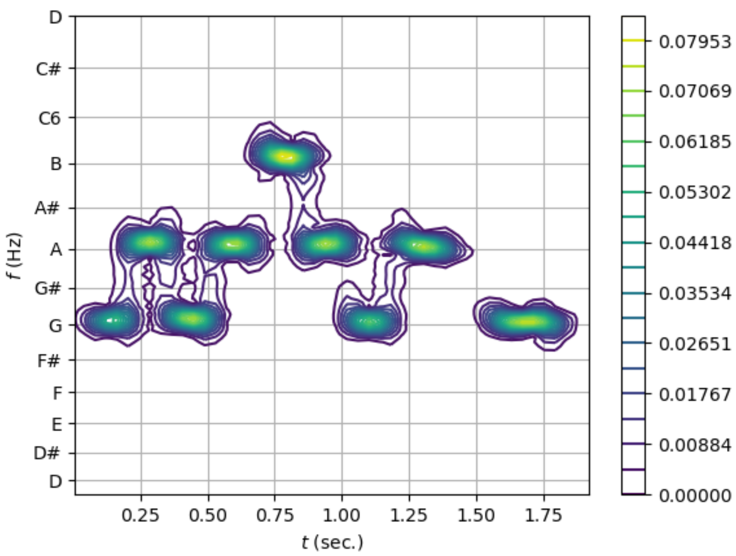
2.2.2 Sonogramme (ou transformée par ondelette)

Proposition 2.2.4. ***Tout signal se décompose*** *en somme (superposition) de* ***paquets d'ondes*** *d'amplitude A pour chaque position temporelle t, et fréquence f.*

La proposition précédente est essentielle en science et découle de la théorie de Joseph Fourier.

Définition 2.2.5. Le **sonogramme d'un signal (ou transformée en paquets d'ondes)** est la représentation de l'intensité $I = |A|^2$ en fonction de t, f dans sa décomposition en paquets d'ondes.

Exemple 2.2.6. A partir du signal sonore flute.wav voici un sonogramme obtenu en fixant la durée $\sigma = 0.04$ sec pour chaque paquet d'onde. Sur la graphe les axes sont t, f où on a utilisé une échelle log pour les fréquences et indicée par le nom de notes. Les couleurs correspondent à une échelle d'intensité. Ainsi on peut "lire la partition".



— Nous recommandons **l'application Spectroid** sous Android, qui permet de faire une transformée temps-fréquence d'un signal sonore en temps réel.

Remarque 2.2.7. Dans la plupart des logiciels qui proposent une transformée de Fourier, il s'agit en fait d'une transformée par paquet d'onde comme définie ci-dessus avec un choix de "fonction fenêtre", window, ou $\sigma = \Delta f$, à préciser. Attention, il est important de **bien choisir la taille** Δf en utilisant le principe d'incertitude ci-dessous.

Exercice 2.2.8. Observer le sonogramme de quelques sons : voix, instrument de musique, etc

2.2.3 Le principe d'incertitude

Remarquer que le nombre N d'oscillations internes dans un paquet d'onde est $N = \frac{\sigma}{T} = \sigma f$. Donc pour un paquet d'onde de durée $\Delta t = \sigma$ fixée, si on change légèrement sa fréquence de Δf, la variation du nombre d'oscillations est $\Delta N = \sigma \Delta f = \Delta t \, \Delta f$ et si $\Delta N < 1$ alors cela ne changera pas son aspect (i.e. nombre d'oscillations internes). Cela correspond à la condition :

$$\Delta N < 1 \Leftrightarrow \quad \Delta f \Delta t < 1.$$

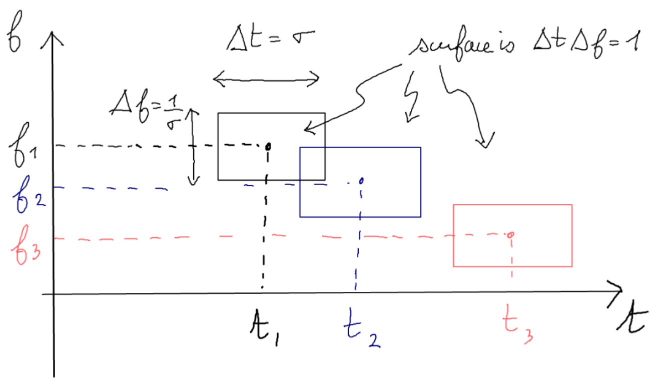
Cela donne le résultat suivant :

Proposition 2.2.9. *La précision temporelle Δt et précision fréquentielle Δf d'un paquet d'onde sont est reliées par*

$$\Delta t \Delta f = 1 \tag{2.2.1}$$

*appelé **principe d'incertitude.***

Comme $\Delta t \Delta f$ est une surface dans le plan temps-fréquence, le principe d'incertitude montre qu'il n'y a pas d'information à une échelle plus fine que la surface 1 :



Remarque 2.2.10. En utilisant une transformée d'un signal en sonogramme et en ne mémorisant que les zones où le signal est suffisement intense, et à raison d'un échantillon par unité de surface $\Delta t \Delta f = 1$ on peut ainsi considérablement réduire la quantité d'information à retenir pour mémoriser un signal. C'est ainsi que fonctionne certains codage audio (et vidéo). Au retour, on reconstruit le signal en superposant de façon adéquate les paquets d'ondes.

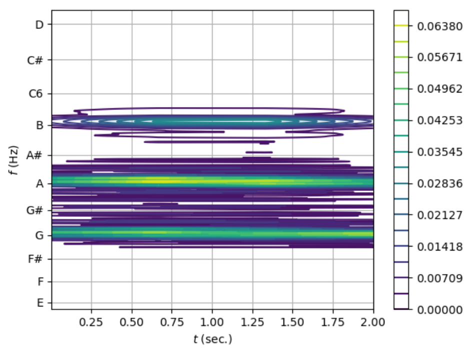
Exemple 2.2.11. Application musicale : supposons une contrebasse qui joue des notes pizzicato autour de la note C_3 dont la fréquence est $f_{C_3} = 65$Hz et la suivante $f_{C\#_3} = 65 \times 1.06 = 69$Hz (voir plus loin). A quel tempo les notes deviennent indiscernables à cause du principe d'incertitude ? (On ignore les harmoniques et considère seulement la composante fondamentale)

Solution 2.2.12. L'écart de fréquence entre ces deux notes voisines est $\Delta f = 69 - 65 = 4$Hz. Pour avoir cette précision, d'après le principe d'incertitude la durée minimale d'une note Δt vérifie $\Delta f \Delta t = 1$. Donc

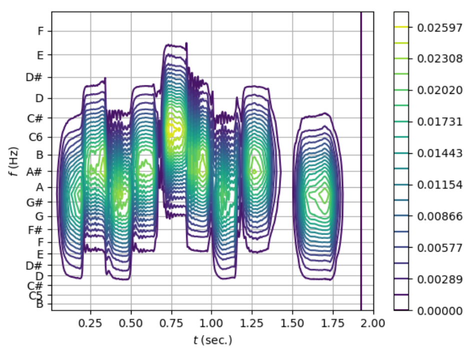
$$\Delta t = \frac{1}{\Delta f} = \frac{1}{4} s.$$

C'est un débit de 4 notes par seconde. A un tempo plus rapide, on ne peut distinguer la fréquence de la note.

Exemple 2.2.13. Dans le sonogramme ci-dessus de la flute, avec $\Delta t = \sigma = 0.5$ on ne distingue plus la résolution temporelle des notes :



avec $\Delta t = \sigma = 0.003$ on ne distingue plus la résolution fréquentielle des notes, car Δf est trop grand :



Liens externes :

— Voir de nombreux exemples de sonogrammes dans le livre Castellengo

(2015).

— spectrogramme en JS

— Video vulgarisée sur le sonagramme

2.3 Signaux périodiques, fréquences, notes musicales et pitch

2.3.1 Signaux périodiques, séries de Fourier

Lecture préalable : On conseille de lire au préalable **la section A.3,** afin de bien comprendre :

— Le mouvement circulaire et la fonction sinus

— La décomposition de Fourier d'un mouvement quelconque en mouvement circulaires (ou en sinus)

— Pour un signal (ou mouvement) périodique de fréquence $f = \frac{1}{T}$, ses fréquences harmoniques $f, 2f, 3f$, etc

Remarque 2.3.1. Les fonctions périodiques sont très particulières d'un point de vue mathématique ou physique. Les signaux périodiques (ou approximativement périodique) sont en effet très rare dans la nature, sauf parmi les animaux, car ils servent à la communication et sont générés par la vibration de "cordes vocales". Ils ont ainsi une grande importance en musique et on les appellera **notes musicales** (bien qu'une "jolie note" ne soit pas un signal parfaitement périodique).

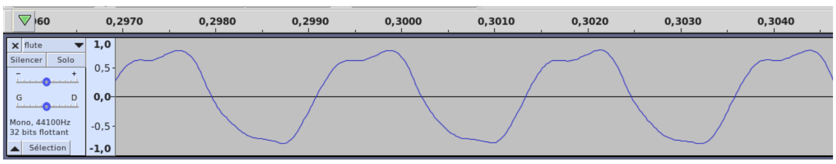
Définition 2.3.2. Un **signal périodique** de **période** T est une fonction $t \to y(t)$ qui vérifie

$$y(t+T) = y(t), \forall t$$

On appelle T la **période fondamentale** si c'est la plus petite valeur de période et $f = 1/T$ est la **fréquence fondamentale**. Une **note musicale** est un signal périodique (ou presque en pratique).

Remarque 2.3.3. Si T est la période fondamentale alors $T, 2T, 3T, \ldots$ sont aussi des périodes car si le signal se répète une fois, alors forcément il se répète ensuite.

Exemple 2.3.4. Avec Audacity, voici le signal d'une note de flûte de période approximative $T = 0.302 - 0.300 = 0.002$sec. donc de fréquence $f = 1/0.002 = 500$Hz :



Exemple 2.3.5. Avec audacity ou l'application Lexis Audio Editor, enregistrer sa voix parlant. Zoomer le signal pour observer que localement, à chaque voyelle, le signal est périodique. Par exemple (voyelle A par un homme) avec Lexis Audio Editor :

Exemple 2.3.6. Sur cet exemple, entre $t = 3.55$ et $t = 3.60$ on observe 6 périodes donc la période T est

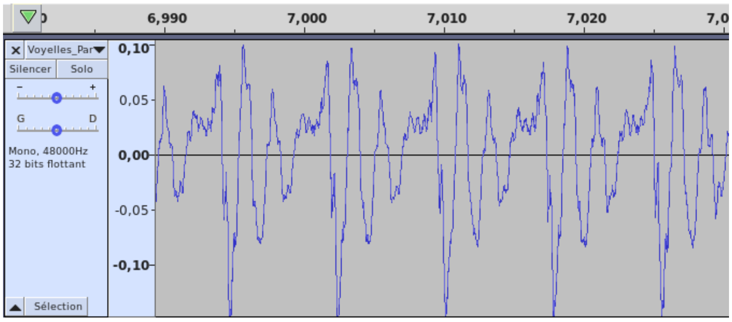
$$T = \frac{3.60 - 3.55}{6} = 0.0083$$

et la fréquence

$$f = \frac{1}{T} = 120\text{Hz}.$$

Exercice 2.3.7. (TP) "**Analyse de la voix**". Avec Audacity ou logiciel équivalent comme Lexis Audio Editor sous android..

1. Enregistrer sa voix prononçant une voyelle, parlée ou chantée. Ecouter. Observer que le signal est (localement) périodique, i.e. c'est une note de musique. Mesurer la période fondamentale T, déduire la fréquence f avec les incertitudes de mesure. Si besoin, voici un fichier déjà enregistré Voyelles_Par_Malik.wav.



2. Selectionner une période (wave-form) et la répéter avec audacity en faisant Ctrl-T (Edition/suppresion spéciale/Rognage audio), puis Effets/Répéter. Quelle différence à l'écoute ?
3. Inverser le sens du signal de départ (Effets/Inverser sens) et écouter. Essayer avec une phrase contenant des consonnes (attaques b,p,t etc) et sans consonnes mais voyelles seules (a,e,i,o,u).
4. Enregistrer sa voix prononçant une voyelle chuchotée. Ecouter. Observer le signal. Quelle différence observez vous avec le cas parlé ?

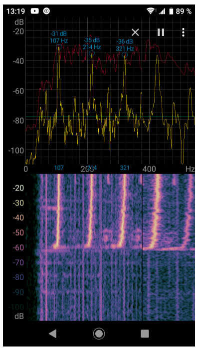
Solution 2.3.8.

1. Partant de la date $t = 7,000$, on mesure $2T \approx 0.016s$ soit $f = \frac{1}{T} =$ 125Hz.
2. On reconnait le son, mais il parait très artificiel. Il manque les variations.
3. Si il y a des consonnes, le son parait étrange. Avec seulement des voyelles, on ne devine pas qu'il a été inversé.
4. Avec le cas parlé (les cordes vocales vibres), le signal est localement périodique. Dans le cas chuchoté, le signal n'est plus localement périodique.

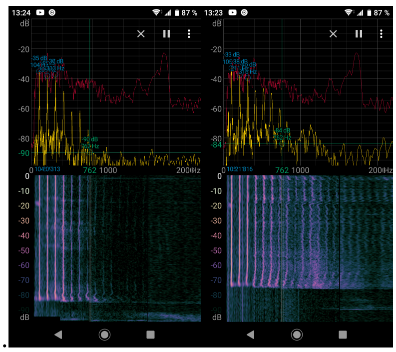
2.3.1.1 Sonogramme d'un signal périodique

Le signal d'un voix réel est (presque) périodique seulement sur des courtes échelles de temps. La fréquence peut varier avec le temps. Pour cela, on a vu l'intérêt d'utiliser la transformée en paquet d'onde ou **sonogramme du signal** pour représenter un signal par une intensité $I(t, f)$ en fonction du temps t et de la fréquence f.

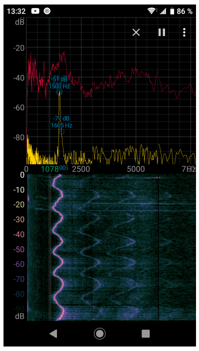
Exemple 2.3.9. Voici le sonogramme d'une voyelle "O" chanté. On observe la fréquence $f = 107\text{Hz}$ et les harmoniques $2f = 214\text{Hz}$, $3f = 321\text{Hz}$ etc



Et voici sur une échelle plus large en fréquences, un "O" et un "A" de même fréquence $f = 104\text{Hz}$. On observe que le "A" a plus d'intensité dans les harmoniques élevées (autour de 1000Hz). C'est ainsi que on distingue les voyelles entre elles.


.

Voici le sonogramme d'un sifflement dont la fréquence varie avec le temps, autour de 1500Hz. On n'observe pas (ou très peu) d'intensité dans les harmoniques supérieures $2f, 3f$ etc. Cela signifie que le sifflement est un signal presque sinusoidal (ou circulaire ou "pur").



2.3.2 Pitch d'un signal périodique

La musique est constituée de notes musicales, c'est à dire de signaux (localement) périodiques. Chaque note a une période T, fréquence $f = \frac{1}{T}$, mais en musique on mesure cela par la "**hauteur de note**" appelée **pitch** et qui correspond à une échelle logarithmique de la fréquence. (De façon analogue à l'intensité I qui est mesurée sur l'échelle logarithmique $\mathrm{dB} = 10\,\mathrm{Log}_{10}\left(\frac{I}{I_0}\right)$ appelée décibels).

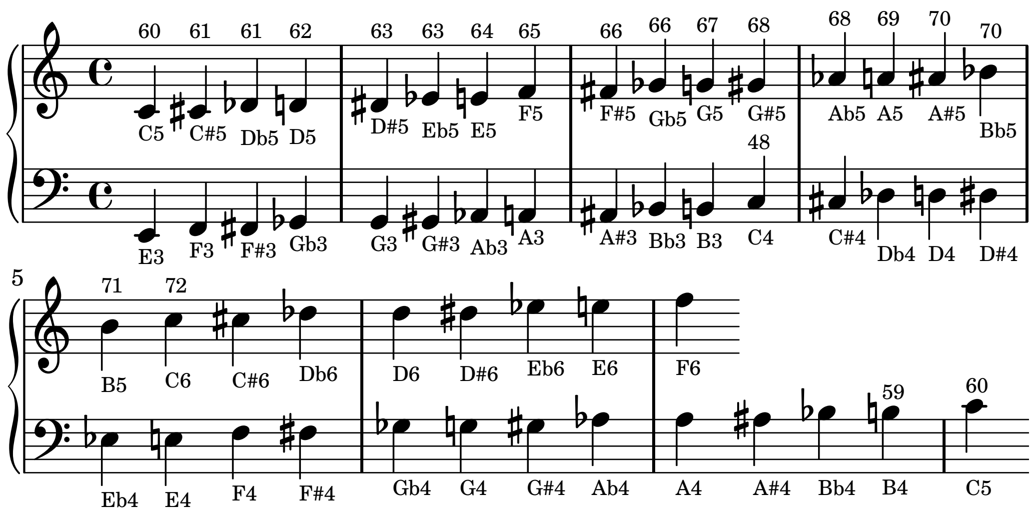
On conseille de lire au préalable **la section B sur la notation musicale.**

2.3.2.1 Les notes sur la portée musicale leur nom et leur pitch (code MIDI) :

Tout d'abord on rappelle le code MIDI de chaque note de la portée du tempérament égal appelé **pitch**. On verra ensuite la correspondance avec la fréquence f de la note.

Définition 2.3.10. La note C_5 (dans l'octave 5) a le pitch $x_{C_5} = 60$ et ensuite on ajoute ± 1 pour chaque demi-ton.

Cela donne la liste suivante des notes et leur pitch :



2.3.2.2 Définition du pitch

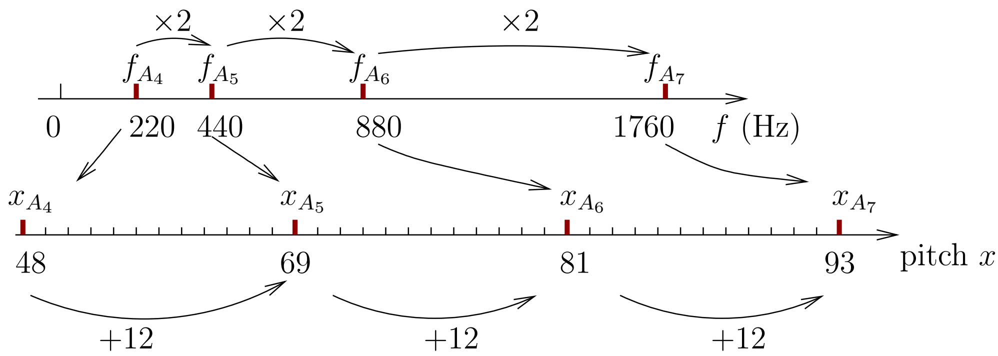
Pour un signal périodique, la définition suivante donne une mesure équivalente des fréquences f mais sur une échelle logarithmique, plus appropriée à la perception humaine et adaptée à la musique.

Définition 2.3.11. Le **pitch MIDI** x est une mesure de la fréquence f sur une échelle logarithmique, i.e. la multiplication correspond à une addition. Par convention,

— la note A_5 de fréquence $f_{A_5} = 440\text{Hz}$ appelé **diapason** a le pitch $x_{A_5} = 69$.

— La multiplication d'une fréquence par $\times 2$ correspond à ajouter $+12$ au pitch.

Comme l'octave est divisée en 12 demi-tons, le pitch est une mesure des fréquences en **demi-tons.**



Remarque 2.3.12. (*) A partir d'une fréquence f donnée, voici une formule qui traduit la définition ci-dessus du pitch

$$x := \left(\frac{12}{\mathrm{Log}_{10}(2)}\right) \mathrm{Log}_{10}\left(\frac{f}{f_{A_5}}\right) + x_{A_5} \quad \in \mathbb{R} \tag{2.3.1}$$

inversement la fréquence f à partir du pitch :

$$f = 10^{\left(\frac{\mathrm{Log}_{10}(2)}{12}\right)\left(x - x_{A_5}\right)} f_{A_5} \tag{2.3.2}$$

ou plus simplement :

$$f = 2^{\frac{1}{12}\left(x - x_{A_5}\right)} f_{A_5} \tag{2.3.3}$$

Exemple 2.3.13. Par exemple

— Un changement de fréquence $f' = 2f$ correspond à un changement de pitch $x' = x + 12$, appelé **intervalle d'octave**

— $x' = x + 1$ est appelé **intervalle de demi-ton**. Cela correspond à $f' = 2^{1/12} f$. Il est utile de se rappeler que

$$2^{\frac{1}{12}} = 1.06 \ldots$$

Ainsi **un demi-ton correspond à une augmentation de** 6% **des fréquences**.

— Un ton est deux demi-tons $x' = x + 2$. Un quart de ton est une moitié de demi-on $x' = x + \frac{1}{2}$ correspondant à $f' = 1.03\, f$.

Exercice 2.3.14. Sachant qu'un demi-ton correspond à une augmentation de 6% de la fréquence, sachant que A_5 a la fréquence $f_{A_5} = 440$, quelle est la fréquence $f_{A\#_5}$? quelle est f_{A_6} (une octave au dessus) et $f_{A\#_6}$?

Solution 2.3.15. $f_{A\#_5} = 440 \times 1.06 = 466\text{Hz}$. $f_{A_6} = 2 \times 440 = 880\text{Hz}$. $f_{A\#_6} = 880 \times 1.06 = 933\text{Hz}$

Exercice 2.3.16. Sachant que l'intervalle le plus petit perceptible est de $\frac{1}{10}$ de demi-ton, cela correspond à quelle pourcentage en fréquence ? Quelle est l'intervalle minimum de fréquence perceptible Δf autour de $f_{A_5} = 440$ Hz ?

Solution 2.3.17. $\frac{1}{10}6\% = 0.6\%$. Ansi $\Delta f = 0.006 \times 440 \approx 3\text{Hz}$.

La définition suivante est parfois utilisée

Définition 2.3.18. Si x est le pitch (unité de demi-ton) on appelle $100x$ les **cents.** Par exemple $x = 60.41$= 61 demi-ton + 41 cents.

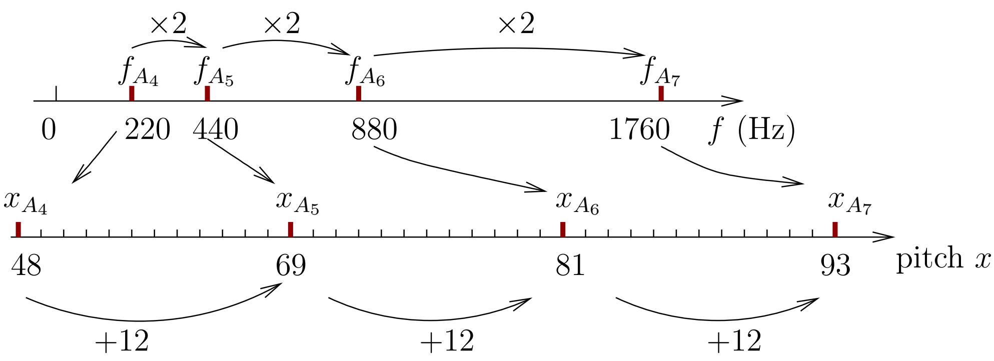
Remarque 2.3.19. On verra que la perception humaine permet de distinguer au mieux un intervalle de 0.1 pitch, c'est à dire un dixième de demi-ton et pas mieux. Cela correspond à un rapport de fréquence de $2^{1/120} = 1.006 = 1 + 0.6\%$.

Exercice 2.3.20. Avec Audacity, créer 10 signaux sinusoidaux entre A_5 de fréquence $f_{A_5} = 440\text{Hz}$ et $A\#_5$. Ecouter et mesurer votre perception la plus fine.

2.3.2.3 Le tempérament égal

Définition 2.3.21. A priori, la définition (2.3.1) précédente toutes les valeurs x du pitch sont possibles. En musique occidentale il est convenu d'utiliser seulement **les valeurs entières du pitch x,** autrement dit la **division de l'octave en 12 parties égales.** Ce choix particulier d'un ensemble discret de fréquences de notes s'appelle le **tempérament égal**.

Sur la figure suivante, le tempérament égal (division de l'octave en 12 parties égales) est représenté par les traits noirs sur l'axe des pitchs x :



Voir l'annexe B pour les différentes conventions et représentations du tempérament égal.

Remarque 2.3.22. Le choix d'utiliser le tempérament égal a été **fait par les luthiers au XIXeme siècle en occident** pour des raisons pratiques. Référence : Benson (n.d., p181,p190,p197). En fait ce choix de se limiter à des valeurs entières de x **limite considérablement le rendu des "couleurs musicales"**, car notre perception est en fait sensible à ± 0.1 pitch (soit 10 cents), voir chapitres suivants et "tempéraments justes".

2.3.3 Comparaison des harmoniques avec le tempérament égal

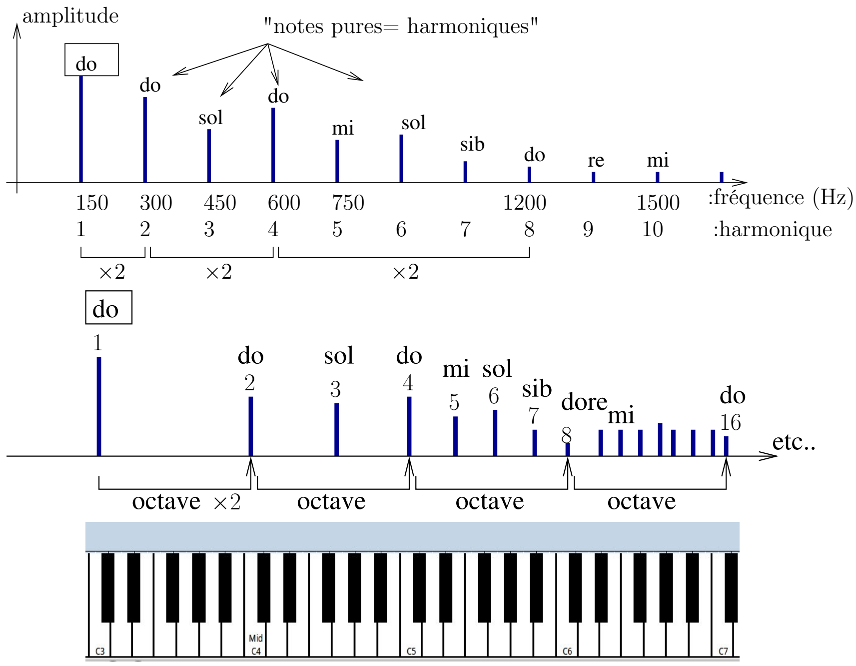
Rappels :

— D'après le Théorème de Fourier A.3.17 que les fréquences d'un signal périodique (ou note de musique) de fréquence f, appelées harmoniques sont $f_a = af$ où $a = 1, 2, 3, \ldots$ est le **numéro de l'harmonique**, c'est à dire :

$$f_1 = f, \quad f_2 = 2f, \quad f_3 = 3f, \text{etc}$$

$f_1 = f$ est appelée la fondamentale.

— L'échelle des pitch (ou hauteur de note en demi-tons) est une échelle logarithmique des fréquences, où une octave $f' = 2 \times f$ est transformée en $x' = x + 12$.



Proposition 2.3.23. *L'intervalle mesuré **en pitch entre la fondamentale et l'harmonique** a **est exactement un entier** (de demi-tons) et correspond donc au tempérament égal, **seulement pour les octaves** : $a = 2, 4, 8, 16$ etc. Pour les autres harmoniques, il y a une correction par rapport au tempérament égal représenté sur la figure suivante.*

Ecoute : Voir cette video qui compare les harmoniques avec le tempérament égal :

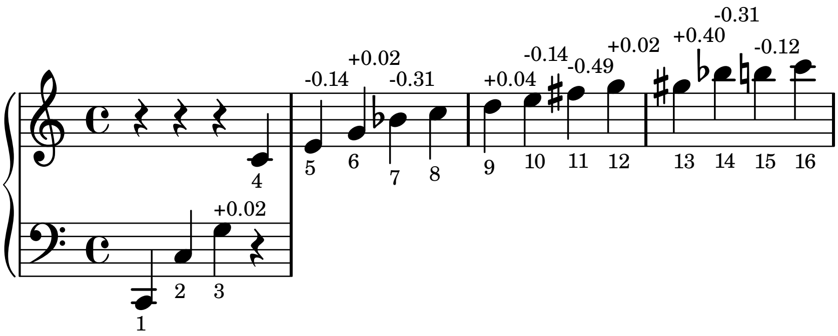


FIGURE 2.3.1 – Harmoniques d'une note fondamentale, ici C_3, représentées sur la portée.

1. Ecouter la difference avec chaque note tempérée.
2. Ecouter les vraies harmoniques et leur superposition : on entend un timbre
3. Ecouter les "fausses harmoniques" (c'est à dire avec le tempérament égal) et leur superposition : on n'entend pas un timbre mais un ensemble complexe de fréquences superposées.

Pour calculer cette correction de la figure 2.3.1, on applique la formule (2.3.1). Par exemple pour l'harmonique $a = 3$, par rapport à l'harmonique $a = 1$, on écrit :

$$x_3 - x_1 = \left(\frac{12}{\mathrm{Log}_{10}(2)}\right)\mathrm{Log}_{10}\left(\frac{3}{1}\right) = 19.02 = \underbrace{12}_{\text{octave}} + \underbrace{7}_{\text{quinte}} + 0.02$$

Exercice 2.3.24. Calculer de même les corrections pour les harmoniques $a = 5, 7, 11, 13$ et vérifier avec la figure ci-dessus.

Remarque 2.3.25. Du fait que $9 = 3 \times 3$, et d'après la formule (A.2.1), on peut écrire :

$$
\begin{aligned}
x_9 - x_1 &= \left(\frac{12}{\mathrm{Log}_{10}(2)}\right)\mathrm{Log}_{10}\left(\frac{9}{1}\right) = \left(\frac{12}{\mathrm{Log}_{10}(2)}\right)\mathrm{Log}_{10}(3\times 3)\\
&= \left(\frac{12}{\mathrm{Log}_{10}(2)}\right)\mathrm{Log}_{10}(3) + \left(\frac{12}{\mathrm{Log}_{10}(2)}\right)\mathrm{Log}_{10}(3) = 2\times 19 + 0.04\\
&= 3\times 12 + 2 + 0.04
\end{aligned}
$$

Exercice 2.3.26. En décomposant les chiffres, calculer de même les corrections pour les harmoniques $a = 6, 7, 10, 12, 14, 15$ et vérifier avec la figure ci-dessus.

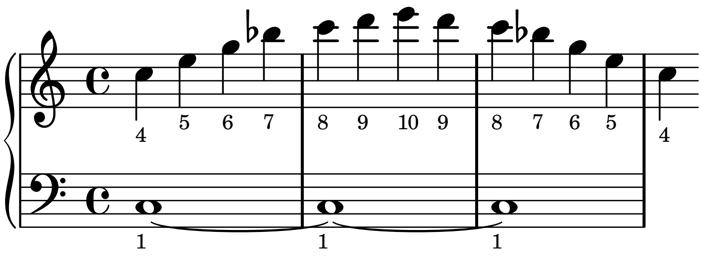
2.3.4 Exemple du chant diphonique

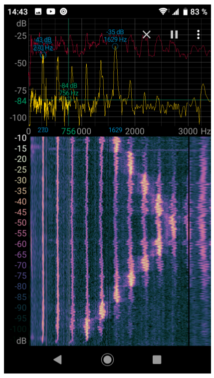
Nous avons dit que la voix humaine est un signal sonore (presque) périodique et contient donc des harmoniques.

Il y a une technique vocale appelée “chant diphonique” qui permet de mettre beaucoup d'intensité dans l'une ou l'autre des harmoniques, Voici cette technique expliquée dans une Conference de Anna-Maria Hefele.

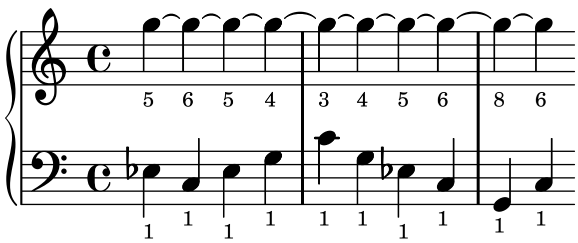
Nous commentons ici quelques exemples précis. Peut écouter le chant diphonique overtone singing - Anna-Maria Hefele sur youtube et on donne la partition montrant la fréquence de la note (en basse) et l'harmonique renforcée sur la portée musicale et avec le numéro n de l'harmonique. Attention : on écrit avec des notes tempérées mais il s'agit d'intervalles justes, comme on verra ensuite, il faudrait ajouter les même corrections que sur la figure 2.3.1.

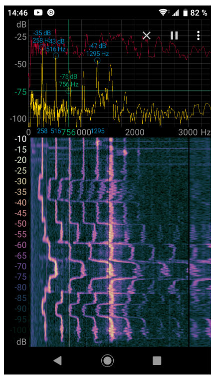
Exemple 2.3.27. Dans ce passage de Anna-Maria Hefele à 29”. Il y a une basse commune f_1.



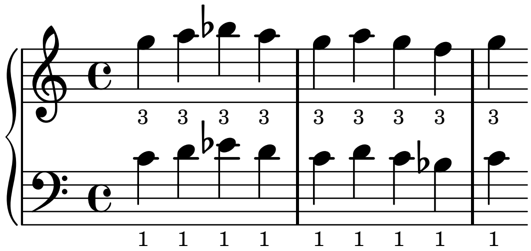


Exemple 2.3.28. Dans ce passage de Anna-Maria Hefele à 1'01". La fréquence f_n de l'harmonique renforcée est fixe, mais le numéro n change, par conséquence la fréquence de la fondamentale f_1 change :

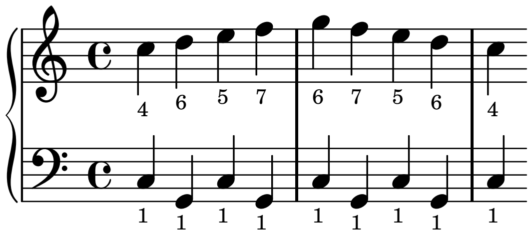




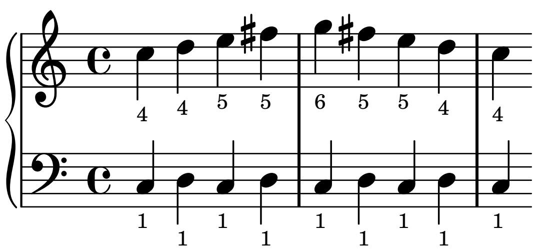
Exemple 2.3.29. Dans ce passage de Anna-Maria Hefele à 2'17". On fixe le numéro $n = 3$ de l'harmonique renforcée et f_1 varie de façon arbitraire, donc f_n varie aussi.



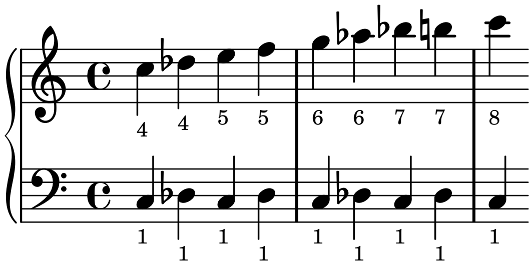
Exemple 2.3.30. Dans ce passage de Anna-Maria Hefele à 2'33'. La note f_1 varie entre deux notes et le numéro n varie de sorte que l'harmonique renforcée f_n fait une mélodie.



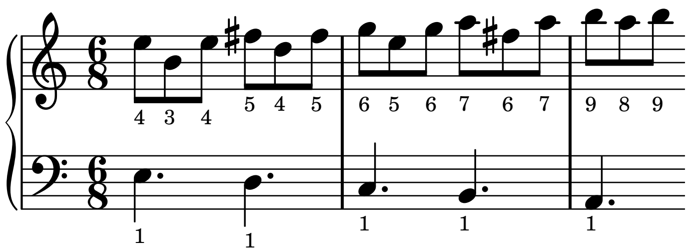
Exemple 2.3.31. Dans ce passage de Anna-Maria Hefele à 2'44'. La note f_1 varie entre deux notes et le numéro n varie de sorte que l'harmonique renforcée f_n fait une mélodie.



Exemple 2.3.32. Dans ce passage de Anna-Maria Hefele à 2'59". La note f_1 varie entre deux notes et le numéro n varie de sorte que l'harmonique renforcée f_n fait une mélodie.



Exemple 2.3.33. Dans ce passage de Anna-Maria Hefele à 3'29". La note f_1 descend et le numéro n varie de sorte que l'harmonique renforcée f_n fait une mélodie en arpèges.



2.3.5 Intervalles justes

2.3.5.1 Définition

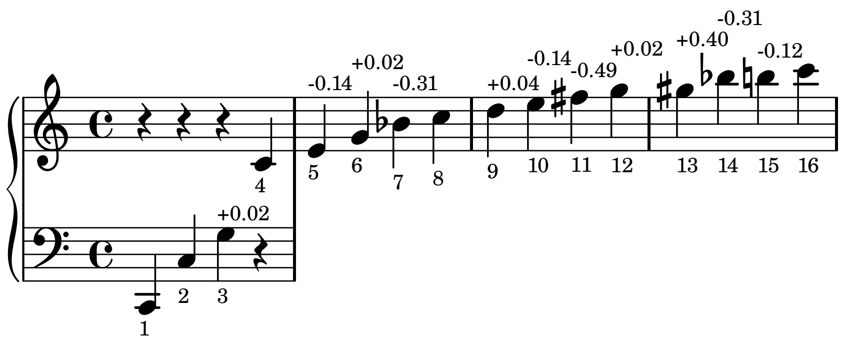
La définition suivante est très importante en musique.

Définition 2.3.34. Un **intervalle juste** est un rapport entre deux fréquences appartenant aux harmoniques $a, b \leq 16$ d'une note donnée.

$$\frac{f_a}{f_b} = \frac{af}{bf} = \frac{a}{b}, \quad a, b \text{ entiers inférieur à } 16.$$

Autrement dit un **intervalle juste** est un intervalle entre deux harmoniques.

On rappelle la suite des harmoniques étudiée ci-dessus avec les corrections de pitch (et on rappelle que on est à peine sensible à ± 0.1 pitch) :



Il est arbitraire de limiter les harmoniques $a, b \leq 16 = 2^4$ ce qui fait 4 octaves. Cela pourrait être plus, et dépend de notre perception (voir chapitre perception).

Exemple 2.3.35. Par exemple, à partir de la figure précédente, on déduit une liste de quelques intervalles justes :

Harmoniques $a \to b$	Nom de l'intervalle	Fraction $\frac{f'}{f}$	Pitch x (en demi-tons)
$1 \to 1$	Unisson	$\frac{1}{1} = 1$	0
$1 \to 2$ ou $2 \to 4$ etc	Octave	$\frac{2}{1} = 2$	12
$2 \to 3$	Quinte	$\frac{3}{2}$	$7 + 0,02..$
$3 \to 4$	Quarte	$\frac{4}{3}$	$5 - 0,02..$
$4 \to 5$	Tierce maj.	$\frac{5}{4}$	$4 - 0,14..$
$5 \to 6$	Tierce min. 1	$\frac{6}{5}$	$3 + 0.02 + 0.14 = 3 + 0.16$
$6 \to 7$	Tierce min. 2	$\frac{7}{6}$	$3 - 0.31 - 0.02 = 3 - 0.33$
$7 \to 8$	Ton 1	$\frac{8}{7}$	$2 + 0.31$
$8 \to 9$	Ton 2	$\frac{9}{8}$	$2 + 0.04$

Exercice 2.3.36. Vérifier les colonnes 3,4 de cette liste et continuer la liste.

Solution 2.3.37. :

Harmoniques $a \to b$	Nom de l'intervalle	Fraction $\frac{f'}{f}$	Pitch x (en demi-tons)
$9 \to 10$	Ton 3	$\frac{10}{9}$	$2 - 0.10$
$10 \to 11$	Ton 4	$\frac{11}{10}$	$2 - 0.35$
$11 \to 12$	Ton 5 ou demi-ton 1	$\frac{12}{11}$	$2 - 0.49 = 1 + 0.51$
$12 \to 13$	demi-ton 2	$\frac{13}{12}$	$1 + 0.38$
$13 \to 14$	demi-ton 3	$\frac{14}{13}$	$1 + 0.29$
$14 \to 15$	demi-ton 4	$\frac{15}{14}$	$1 + 0.19$
$15 \to 18$	demi-ton 5	$\frac{16}{15}$	$1 + 0.12$

2.3.5.2 Quasi justesse de la quinte et quarte tempérée :

On s'aperçoit donc que tous les intervalles justes ne sont pas tempérés à l'exception des **octaves qui sont justes**. Mais de façon remarquable **la**

quarte et la quinte sont proches d'un intervalle juste respectivement 3/2 et 4/3 avec une erreur ± 0.02 pitch qui est imperceptible. Par contre les autres intervalles justes dévient du tempérament égal de façon très perceptible (erreur > 0.1 pitch).

— Ainsi grâce au choix de la division de l'octave en 12 on peut reproduire avec bonne précision les intervalles entre les 4 premières harmoniques (octave, quarte, quinte).
— Le nombre 12 a un autre avantage : il a de nombreux diviseurs : $2, 3, 4, 6$ et ainsi on peut structurer le cercle dodécaphonique en parties égales de multiples façons, comme la journée que l'on a choisit de diviser en 12 heures. Mais cela ne semble pas avoir été un argument au départ en musique. Ça a été utilisé par la suite au XXieme siècle par les modes symétriques de Messiaen ou par la musique sérielle. (Dans la culture, le chiffre 12 est aussi important car il y a 12 mois lunaire dans l'année, cela est seulement du au mouvement relatif de la Terre et de la Lune)

Exercice 2.3.38. Accordage d'un piano en tempérament égal :

— Si f_1,f_2 sont les fréquences de deux notes du piano formant un intervalle de **quinte tempérée**, calculer la fréquence du battement entre l'harmonique 3 de f_1 et l'harmonique 2 de f_2 si $f_1 = 440$Hz. Ce battement est utilisé par les accordeurs de piano, voir cette vidéo sur accordage des quintes.

Solution 2.3.39. Une quinte tempérée est le rapport $f_2 = 2^{\frac{7}{12}} f_1$. La fréquence du battement entre l'harmonique 3 de f_1 et l'harmonique 2 de f_2 est la différence

$$(3f_1 - 2f_2) = f_1 \left(3 - 2 \times 2^{\frac{7}{12}}\right) = f_1 \times 0.0039 = 1.5\text{Hz}$$

Comme $1.5 = \frac{3}{2}$, cela correspond à 3 battements toutes les 2 secondes.

— Si f_1,f_2 sont les fréquences de deux notes du piano formant un intervalle de **tierce tempérée**, calculer la fréquence du battement entre l'harmonique 5 de f_1 et l'harmonique 4 de f_2 si $f_1 = 220$Hz. Ce battement est utilisé par les accordeurs de piano. Video de l'accordage de la tierce (écouter les battements mais attention le discours de la vidéo est plutôt incorrect).

Solution 2.3.40. Une tierce tempérée est le rapport $f_2 = 2^{\frac{4}{12}} f_1$. La fréquence du battement entre l'harmonique 5 de f_1 et l'harmonique 4 de f_2 est la différence

$$(5f_1 - 4f_2) = f_1 \left(5 - 4 \times 2^{\frac{4}{12}}\right) = f_1 \times (-0.039) = -8.7\text{Hz}$$

Cela correspond à 9 battements par seconde.

2.3.5.3 Vidéos sur les intervalles justes

— Jacob Collier explains that Piano is not in tune and shows that major third is out of tune by +0.14 pitch (half tone).

— Comparaison des harmoniques et notes du tempérament égal avec la guitare.

— Harmoniques de la voix

  - Wave form de voyelles chantées par Malik. Harmoniques et notes correspondantes en Video
  - polyphonic overtone singing - Anna-Maria Hefele
  - Conference de Anna-Maria Hefele
- Signaux périodiques ou pas :
  - Lorenz Attractor as audio
  - Bassoon Multiphonics 01
- Video sur l'harmonie des sphères d'après Kepler.
- Video sur comment jouer des intervalles justes sur la guitare

2.4 Le tonnetz et quelques tempéraments justes

Comme l'utilisation des intervalles justes est très importante en musique (musique ancienne et musique du monde) il est utile de représenter les intervalles justes de façon convenable pour ensuite étudier les accords justes, leurs progressions harmoniques et aussi représenter les **"tempéraments justes"** qui sont le choix précis de quelques intervalles justes sur un instrument donné (utilisé en musique ancienne et certaines musiques du monde, autre que "occidentale"). Cette représentation est appelée le réseau **tonnetz** et a été introduite par Euler (en 1739, il a 24 ans), dans son livre « Tentamen novae theoriae musicae ex certissimis harmoniae principiis dilucide expositae ». ("Une nouvelle théorie musicale qui est clairement exposée aux principes d'harmonie les plus fiables")

2.4.1 Décomposition des intervalles justes en intervalles de base

En arithmétique, on peut décomposer chaque entier à partir des nombres premiers : $2, 3, 5, 7, 11, 13, \ldots$. Par exemple $4 = 2\times2$, $6 = 2\times3$, $8 = 2\times2\times2 = 2^3$ etc.

Ces nombres premiers correspondent à des "**intervalles juste de base**" à partir desquels on peut construire tout autre intervalle juste :

Harmoniques $a \to b$	Nom de l'intervalle	Fraction $\frac{f'}{f}$	Pitch x (en demi-tons)
$1 \to 2$	**Octave**	2	12
$1 \to 3$	octave + **Quinte**	3	$12 + 7 + 0,02..$
$1 \to 5$	2 octaves + **Tierce maj.**	5	$2 \times 12 + 4 - 0,14..$
$1 \to 7$	2 octaves + **septieme min.**	7	$2 \times 12 + 10 - 0.33$
$1 \to 11$	3 octaves + **triton bas**	11	$3 \times 12 + 6 - 0.49$
$1 \to 13$	3 octaves + **sixte haute**	13	$3 \times 12 + 9 + 0.40$

Exemple 2.4.1. Par exemple la **tierce mineure** $5 \to 6$, mi $\to$ sol, se décompose de la façon suivante en intervalles de base :

$$\underbrace{(\text{mi} \to \text{sol})}_{\text{tierce min.}} = \underbrace{(\text{mi} \to \text{do})}_{-\text{tierce maj.}} + \underbrace{(\text{do} \to \text{sol})}_{+\text{quinte}}$$

cela correspond au calcul suivant.

$$\frac{6}{5} = \frac{2 \times 3}{5} \equiv \text{oct.+ (oct.+quinte) -(2 oct.+tierce)=quinte - tierce}$$

Ou de façon équivalente et plus directe :

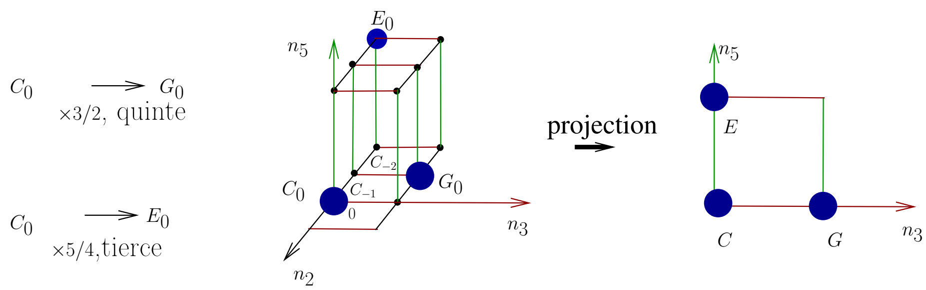
$$\frac{6}{5} = \frac{3}{2} \times \frac{4}{5} = \frac{\frac{3}{2}}{\frac{5}{4}} \equiv \text{quinte - tierce}$$

Lien :

— Vidéo sur les nombres premiers sur Arte, voyage au pays des maths, Conjecture de Riemann.

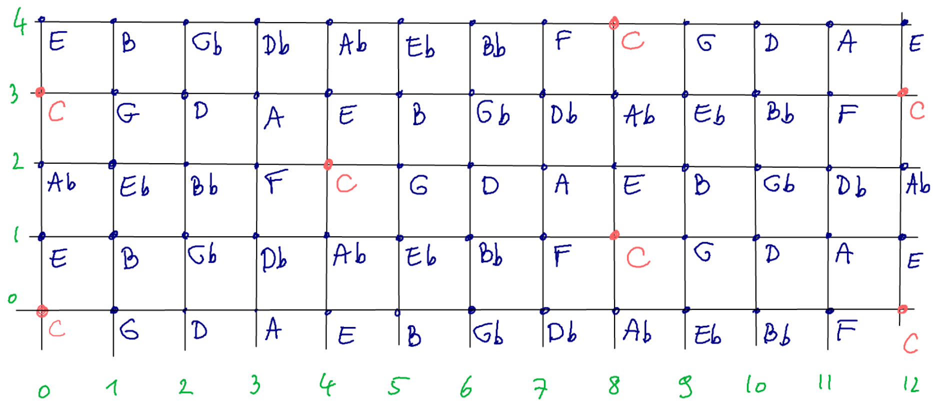
2.4.2 Le tonnetz $2, 3, 5$

Dans le tonnete $2, 3, 5$, par simplicité, on considère seulement les 3 premiers facteurs premiers $2, 3, 5$ et comme changer d'octave n'affecte pas le nom de la note, on omettra de préciser l'octave (i.e. le facteur 2). Les axes n_2, n_3, n_5 correspondent respectivement aux changement de fréquence $\times 2, \times 3, \times 5$ à chaque étape, ce sont dont des axes logarithmiques. Le deuxième schéma est le premier où l'axe n_2 est orthogonal au plan (c'est une projection).

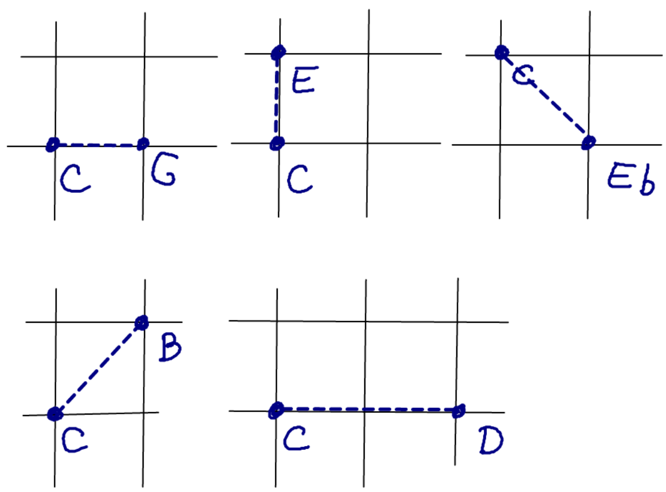


Tonnetz
Tonnetz projetté

Dans la figure suivante, voici la représentation obtenue avec un **axe des quintes justes en horizontal** (facteur 3) et l'**axe des tierces majeures justes en vertical** (facteur 5). Pour aider à la discussion, on a mis des coordonnées en vert. Ce réseau est en fait infini et chaque note représentée a une fréquence différente des autres en principe. On a cependant utilisé les lettres $A, B, \dots G$ pour les notes. Ainsi en rouge, les différentes notes C sont différentes. On précisera ces différences dans la suite.



— Ce qui nous intéressera dans ce réseau, ce sont les **intervalles justes consonants** qui sont des paires de **notes voisines**. Par exemple



— C-G : quinte, facteur $\times 3$,

— C-E : tierce, facteur $\times 5$,

— C-Eb : tierce mineur, facteur $\times \frac{3}{5}$.

— C-D : ton, facteur $\times 3 \times 3 = \times 9$

— Le demi-ton $C - B$, facteur $\times 3 \times 5 = \times 15$ est plutôt dissonant, car ce facteur est trop grand.

Remarques :

- Dans ce réseau tous les intervalles sont justes en principe, chaque note est donc unique, ainsi les notes C en rouge sont en fait toutes différentes, mais proche de l'unisson et sont appelées des **commas**. Par exemple partant du C en $(0,0)$, il y a un autre C en $(4,2)$.
- En terme de pitch (ou demi-ton), on rappelle que d'après le tableau ci-dessus,
  - pour une quinte juste, i.e. déplacement $(1,0)$, il faut ajouter la correction $+0.02$ pitch.
  - pour une tierce majeure juste, i.e. déplacement $(0,1)$, il faut ajouter la correction -0.14 pitch.
- Sur chaque ligne horizontale, appelé **le cycle des quintes**, qui ne se referme pas exactement : il apparaît le **comma pythagoricien** après le déplacement $(12,0)$.
- Sur chaque ligne verticale, appelé **le cycle des tierces majeure**, qui ne se referme pas exactement : il apparait un comma après le déplacement $(0,3)$.
- La **quinte du loup** est le déplacement $(11,0)$. Son pitch est donc $7 + 11 \times 0.02 = 7.22$ pitch. Rappelons que la quinte juste est 7.02 pitch. La quinte du loup est donc perçue comme dissonante.

2.4.2.1 Les commas

La définition suivante n'est pas conventionnelle. Cependant, nous ne connaissons pas de définition de comma en musique juste et celle ci nous semble cohérente avec la suite.

Définition 2.4.2. Un **comma** est un intervalle juste de hauteur plus petit que $1/4$ ton ($= 1/2$ pitch).

Remarque 2.4.3. Il y a donc une infinité de comma. Attention, les comma ne forment pas un sous réseau du tonnetz.

Par exemple :

— Le **comma pythagoricien** est le déplacement $(12, 0)$. C'est donc un intervalle de $12 \times 0.02 = 0.24$ pitch, ($= 1/4$ de demi-ton).

— Le **comma syntonique, (aussi appelé Ptolémé, Dydinis, standard)** est le déplacement $(4, -1)$. C'est donc un intervalle de $4 \times 0.02 - 1 \times (-0.14) = 0.22$ pitch.

— Par exemple le **schisma** est un comma pythagoricien moins un comma syntonique, soit le déplacement

$$(8, 1) = (12, 0) - (4, -1)$$

C'est donc un intervalle de $8 \times 0.02 + 1 \times (-0.14) = 0.02$ pitch, presque imperceptible.

— Le **comma enharmonique** est le déplacement $(0, 3)$.

Exercice 2.4.4. Calculer le pitch du comma $(0, 3)$ et l'exprimer comme combinaison des commas pythagoricien $(12, 0)$ et syntonique $(4, -1)$.

Remarque 2.4.5. le comma syntonique permet de transformer une **tierce pythagoricienne** qui est le déplacement $(4,0)$ en une tierce juste qui est $(0,1)$. En effet

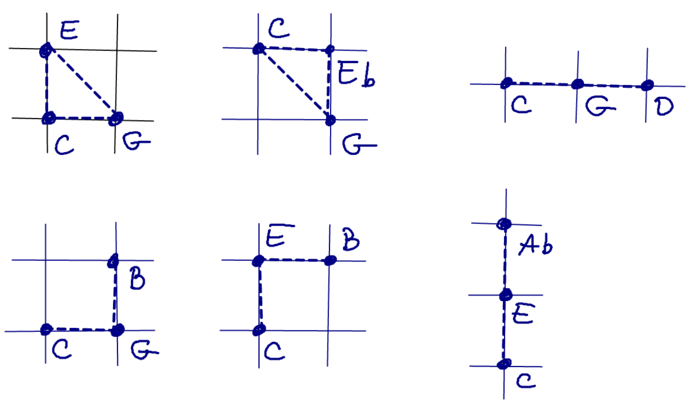
$$(4,0)-(4,-1)=(0,1)\,.$$

Cela est utilisé pour construire le Tempérament mésotonique 1/4 comma.

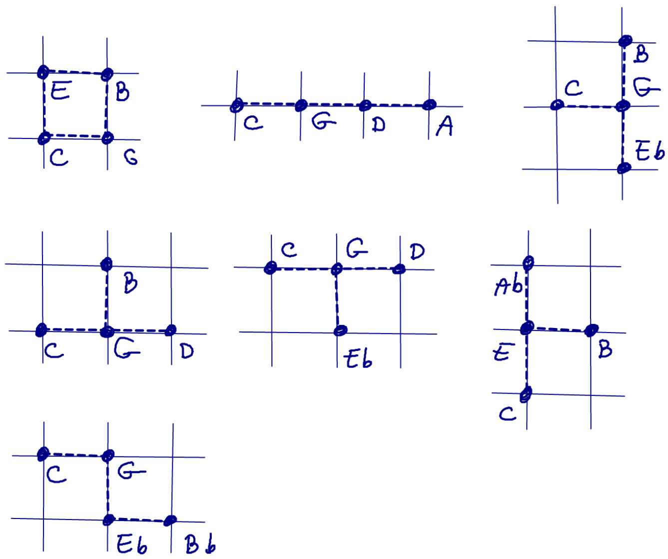
2.4.2.2 Quelques accords justes

Dans le réseau tonnetz 2,3,5, les **accords justes consonants** sont des ensemble de notes formés d'intervalles consonants listés ci-dessus.

Une **triade** est un accord contenant 3 notes. Voici quelques **triades consonantes** remarquables. On remarque la forme géométrique pour chacune d'elle, qui la caractérise.



Voici des accords à 4 notes consonants.



2.4.2.3 Profondeur d'un accord

Chaque accord est caractérisé par sa forme géométrique. Cependant il ressort des propriétés communes. En voici une que l'on appelle la profondeur. Observons :

— Dans la **triade majeure** C,E,G la note C est la fondamentale de toutes ses notes. Donc on perçoit une **sensation de stabilité**. Video

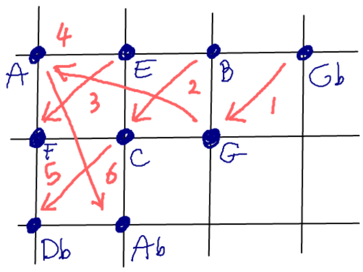
— Dans la **triade mineure** E,G,B aucune des notes présente n'est fondamentale de toutes. La fondamentale serait C, mais elle est absente. On percoit donc une **sensation de manque**. Video

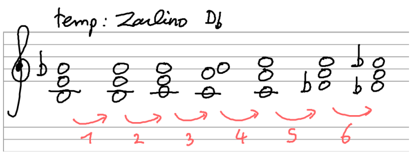
Plus précisément, on peut quantifier cela on mesurant la profondeur en pitch de la fondamentale commune aux notes présentes. (pour une définition précise, voir le cours pour physiciens)

2.4.2.4 Exemple de cadence harmonique en tempérament de Zarlino

Une **cadence harmonique** est un enchaînement d'accords.

Voici un exemple de cadence sur le tempérament de Zarlino, basé sur un mouvement répété de mouvement de demi-ton de déplacement $(-1,-1)$, sauf pour deux mouvements : 3 : ton de déplacement $(-2,1)$ et 5 : demi-ton de déplacement $(1,-2)$.





Exercice 2.4.6. Composer une cadence harmonique sur le tonnetz 2,3,5 avec des accords de 4 ou 5 notes. Pour cela, choisir une liste de déplacements correspond avec une faible variation de pitch que l'on se permettra d'utiliser et partir d'un accord de départ.

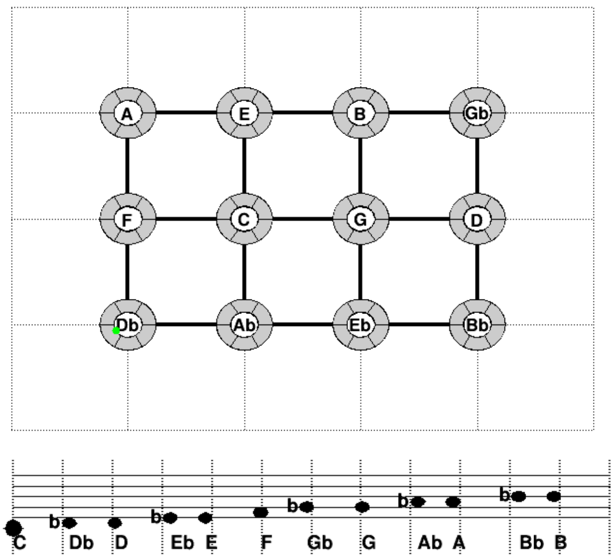
2.4.2.5 Tempéraments justes

Comme le réseau tonnetz contient en principe une infinité de notes, cela pose des difficultés pour jouer ces notes. Sur un instrument traditionnel, en pratique pour jouer avec des intervalles justes, il faut choisir les notes que l'on pourra utiliser. Ce choix s'appelle un tempérament juste.

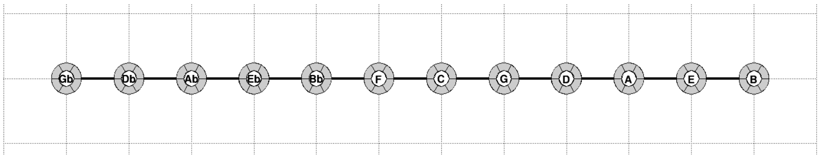
Définition 2.4.7. Un **tempérament juste dans le tonnetz 2,3,5** est **le choix des douze notes chromatiques** $C, Db, D, \dots, Bb, B$ parmi le réseau tonnetz ci-dessus. Ce choix doit être groupé afin de contenir des intervalles et accords justes résonnants que l'on pourra utiliser.

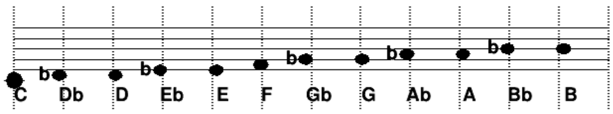
Voici quelques exemple historiques de tempéraments justes pour une certaine tonalité :

— Le **tempérament de Zarlino**, déjà présent dans le tempérament de cloches chinoises de 433 B.C. (ref : article de Cohn 1997, page 63, et Falkenhausen, "suspended music.." p.284) (video) :

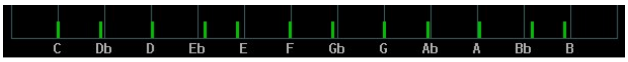


— Le **tempérament de Pythagore**.

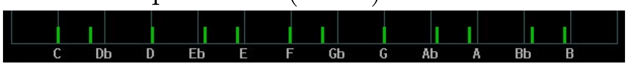




— Le **tempérament de Kepler** dans Harmonices Mundi 1619 (video)



— Le **tempérament de Jean Jacques Rousseau** dans Dictionnaire de la musique" 1768 (video)



— **J.S. Bach**, "Clavier bien tempéré", 1722. (ref : Benson p.188)

The temperament of Vallotti and Young is probably closest to the intentions of J. S. Bach for his Well-Tempered Clavier. According to the re-

— Dans un tempérament juste (non égal), les transpositions donnent des gammes différentes. Voici des conseils de Christian Schubart en 1784, concernant un certain tempérament :

5.13. IRREGULAR TEMPERAMENTS 183

C major	Completely pure. Its character is: innocence, simplicity, naivety, children's talk.
C minor	Declaration of love and at the same time the lament of unhappy love.—All languishing, longing, sighing of the lovesick soul lies in this key.
D♭ major	A leering key, degenerating into grief and rapture. It cannot laugh, but it can smile; it cannot howl, but it can at least grimace its crying.—Consequently only unusual characters and feelings can be brought out in this key.
C♯ minor	Penitential lamentation, intimate conversation with God, the friend and help-meet of life; sighs of disappointed friendship and love lie in its radius.
D major	The key of triumph, of Halelujahs, of war-cries, of victory-rejoicing. Thus, the inviting symphonies, the marches, holiday songs and heaven-rejoicing choruses are set in this key.
D minor	Melancholy womanliness, the spleen and humours brood.
E♭ minor	Feelings of the anxiety of the soul's deepest distress, of brooding despair, of blackest depression, of the most gloomy condition of the soul. Every fear, every hesitation of the shuddering heart, breathes out of horrible E♭ minor. If ghosts could speak, their speech would approximate this key.

Vidéos :

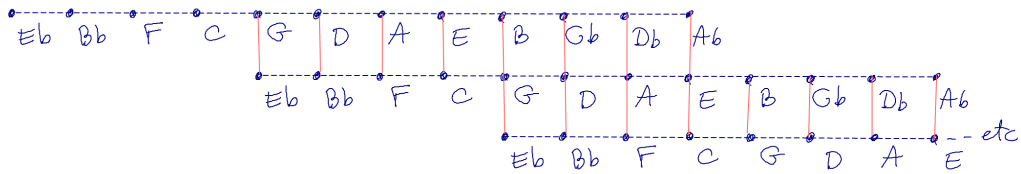
- En tempérament égal :
  - Gymnopédie No. 1 –{} Tonnetz Visualization
  - Tonnetz Analysis of Chopin E-major Prelude
  - Muse - Take A Bow (Tonnetz harmonic analysis)
  - Limitations du tonnetz 2,3,5 : How to create negative harmony using the Tonnetz tone grid

2.4.2.6 Tempérament mésotonique

Le tempérament Tempérament mésotonique 1/4 comma n'est pas un tempérament juste, il n'est donc pas représentable sur le réseau tonnetz. Néanmoins c'est une sorte d'intermédiaire entre le tempérament égal et le tempérament juste de Zarlino ci-dessus. Voici sa définition.

Définition 2.4.8. **"Tempérament mésotonique 1/4 comma"** : Dans la suite du cycle des quintes justes $Eb, Bb, F, C, \ldots, Db, Ab$, on réduit chaque quinte juste par un facteur égal de sorte que la tierce $C - E$ devienne une tierce juste.

D'après la définition, dans la figure ci-dessous toutes les notes de même nom sont égales (i.e. même fréquence ou séparées par des octaves). Horizontalement, les traits pointillés sont des quintes réduites (non justes) de pitch $x = 7 - 0.03 = x_{\text{quinte j.}} - 0.05$ de sorte que verticalement les traits pleins sont des tierces justes de pitch $x_{\text{tierce j.}} = 4 - 0.14$. Ce schéma est périodique (se continue à gauche et à droite).



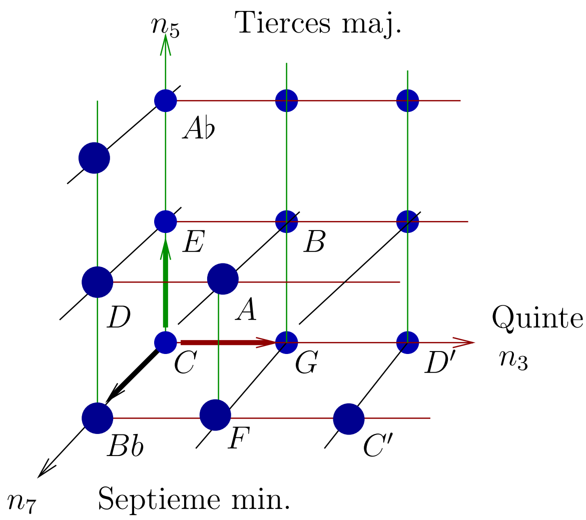
Exercice 2.4.9. Montrer que dans le tempérament mésotonique 1/4 comma défini ci-dessus, les intervalles de quinte qui ont été réduites ont un pitch de quinte juste -0.05 qui est à peine perceptible. Monter que les 8 tierces suivantes sont justes : $Eb - G - B$, $Bb - D - Gb$, $F - A - Db$, $C - E - Ab$.

Solution 2.4.10. Dans la suite des quintes justes entre C et E on a : $C - G - D - A - E$ il y a 4 quintes justes. Cet intervalle $C - E$ s'appelle la **tierce de pythagore** et son pitch est : $x = (7 + 0.02) \times 4 \bmod 12 = 4 + 0.08$. Or la tierce juste est de pitch $x_{tierce} = 4 - 0.14$. Donc dans le tempérament mésotonique, il faut baisser le pitch de $-0.14 - 0.08 = -0.22$ (c'est le comma syntonique de déplacement $(4, -1)$ de pitch 0.22, vu plus haut). On partage cette correction sur les 4 quintes intermédiaires. La correction de chaque quinte est donc $-0.22/4 = -0.05$. Par conséquent après cette correction, dans la suite des quintes corrigées $C-G-D-A-E$ la tierce $C-E$ est juste. Plus généralement dans la suite des quintes corrigées $Eb, Bb, F, C, \ldots, Db, Ab$, les tierces $Eb-G$, $Bb-D$,$F-A$,$C-E$,$G-B$,$D-Gb$,$A-Db$,$E-Ab$ sont justes.

Remarque 2.4.11. D'après wikipédia : c'est le théoricien Pietro Aaron qui aurait imaginé, en 1523 à Venise, le tempérament dit « mésotonique ». Les tempéraments mésotoniques ont été utilisés à la fin de la Renaissance et pendant toute la période baroque. L'intérêt des tempéraments mésotoniques est de pouvoir utiliser des tierces pures, ce qui est possible avec l'exemple ci-dessus dans les huit tonalités majeures Eb, G, Bb, D, F, A, C, E et dans leurs 8 tonalités relatives mineures C, E, G, B, D, Gb, A, Db.

2.4.3 Le tonnetz $2, 3, 5, 7$, tonnetz général et recherche musicale

Si on considère les 4 premiers facteurs premiers $2, 3, 5, 7$ en omettant le facteur 2 de l'octave, on a la représentation suivante des intervalles justes :



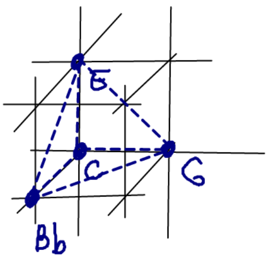
avec l'**axe des quintes** , l'**axe des tierces majeures** et l'**axe des septièmes mineures.**

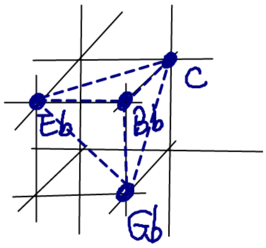
Un intervalle juste est représenté par un déplacement selon les 3 axes. Par exemple

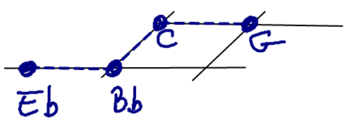
— La **septième** mineure est $(0, 0, 1)$ et correspond au facteur $\frac{7}{4}$.

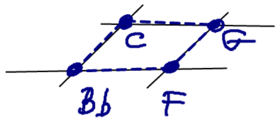
— Le **triton** est $(0, -1, 1)$ et correspond au facteur $\frac{7}{5}$.

Comme avec le tonnetz $2, 3, 5$, on peut énumérer les **accords justes consonants** sur le tonnetz $2, 3, 5, 7$. Voici quelques exemples remarquables :





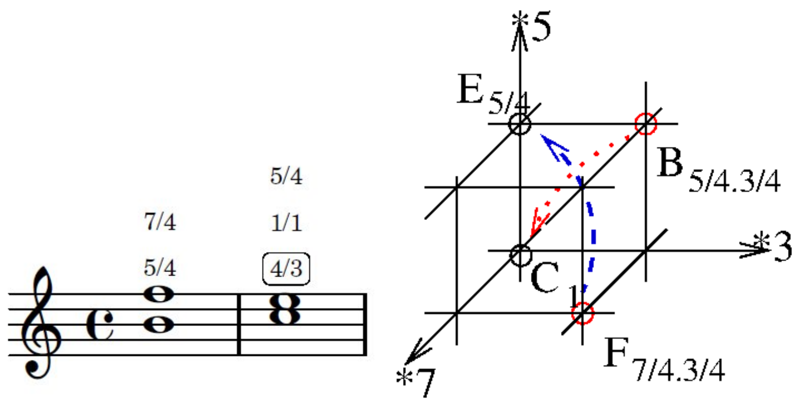




Remarque 2.4.12. Il apparaît des accords bien différents que l'on confondrait en musique à tempérament égal, comme C, Eb, G, Bb qui a une version dans le plan $3, 7$ et une version dans le plan $3, 5$.

2.4.3.1 Résolutions

— **“Résolution du triton $\frac{7}{5}$”** vers la tierce majeure $\frac{5}{4}$: il y a un mouvement d'un demi-ton de $F \to E$ (en tirets bleus) dont la fraction est $\frac{5}{4} \cdot \frac{4}{3} \cdot \frac{4}{7} = \frac{20}{21} \simeq (1.050)^{-1}$. Il y a aussi un mouvement d'un demi-ton $B \to C$ (en pointillés rouges) dont la fraction est $\frac{4}{5} \cdot \frac{4}{3} = \frac{16}{15} \simeq 1.066$. (On rappelle que le demi-ton chromatique est $2^{1/12} \simeq 1.059$). En musique ce dernier mouvement s'appelle “{} résolution de la sensible”. Video, et avec la basse virtuelle : Video.



— **(Cadences II-V-I** Majeure : Video et mineure Video.)

— **Enchaînement d'accords** basée sur trois "petits déplacements" $\frac{20}{21}, \frac{16}{15}, \frac{8}{9}$: Video, et avec la basse virtuelle : Video.

Exercice 2.4.13. Composer une cadence harmonique sur le tonnetz 2,3,5,7 avec des accords de 4 ou 5 notes. Pour cela, choisir une liste de déplacements correspond avec une faible variation de pitch que l'on se permettra d'utiliser et partir d'un accord de départ.

2.5 Les intervalles et accords justes dans les pratiques musicales

La section précédentes nous a donné quelques indications pour analyser les intervalles justes et les accords justes. Considérons maintenant les pratiques musicales. On reviendra plus tard sur les théories musicales. On observe que les intervalles justes sont très présents dans tous les styles de musique, de tout temps et de toute culture. Voici quelques exemples.

— Par ex. l'octave 2/1, la quinte 3/2, la quarte 4/3, sont des intervalles universels dans les cultures humaines. La quarte 4/3 est le premier intervalle de la marseillaise. Son ou Son :

— Dans la fractale de Farey, on a observé des "**gaps**" de résonance, i.e. des intervalles dissonants, tout près des fractions simples, voir remarque **??**. Par exemple, la **ganga** en Dalmatie sont des chants populaires, avec intervalles de notes qui "frottent" (pour le plaisir), ref : carnet de voyage avec Anne-Florence Borneuf. Son

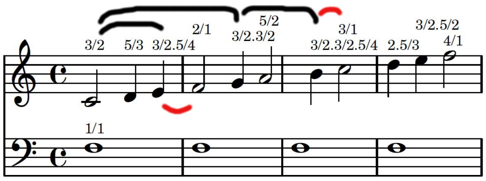
— Organum est un genre musical à **deux voix, du IXème au XIIème siècle en Europe,** constitué d'octaves (2/1), quintes (3/2) et quartes (4/3). Les tierces (6/5) et sixtes (5/3) peuvent être entendues à partir du XIème siècle. ref : Musica Enchiriadis. C'est l'origine du contre-point. Video de Organum duplum : Alleluia, hic Martinus ("organum fleuri").

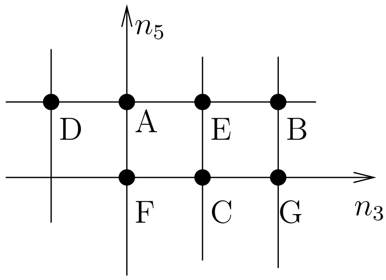
— L' **intervalle 7/5** qui est aussi une fraction proche du triton et remarquable dans la fractale de Farey **??**, est revisité par les bee boppers

(Charlie Parker, Dizzy Gillespie .., 1940'), par ex. dans l'introduction de "salt peanuts" en 1942. Dans cet exemple, il y a un intervalle de quinte diminuée 7/5, puis 14/5 ou 11/4 ?, puis l'octave 2/1. Son

— Les **chants bulgares** (musique traditionnelle actuelle) font entendre des des voix très timbrées (i.e. riches en harmoniques) avec des résonances $\frac{f'}{f} = \frac{a}{b}$ bien marquées qui se concluent souvent par des unissons. On remarque aussi des rythmes en 2+2+3=7. Video.

— En **musique indienne**, on apprécie les résonances entre la **tampura** (riche en harmoniques) et la **flûte**.





Dans cet exemple, on remarque que les rapports rationnels amènent des **relations entres les notes** et pas seulement avec la basse fixe (tampura). Ces relations sont schématisées sur la partition par les traits en noir. On a aussi mis en valeur la proximité de hauteur en rouge. On discutera de cela après.

— Qu'observez vous de particulier concernant le ré (qui est peu joué) ? réponse : c'est le seul dont le dénominateur de la fraction (ici 5/3) n'est pas une puissance de 2. On verra plus loin que si le musicien joue le ré de façon trop insistante, il ressort une couleur « mineur.

— Video

— Des intervalles justes faisant intervenir les entiers 11 et 13 sont présents dans les **chants suisses, et le cor des alpes**. Son (avec la "11#" à 45", et la "13♭" à 1'56"). Remarquer que les musiciens évitent soigneusement de jouer l'harmonique 7 et $14 = 2 \times 7$.

2.5.1 Quelques gammes et modes

2.5.1.1 La gamme pentatonique majeur

Utilisée partout dans le monde.

F	C	G	D	A

— jouée dans l'ordre croissant des fréquences, cela donne : **F**,G,A,C,D,(F) , utilisé par ex. en musique irlandaise.

   — Le mode **C**,D,F,G,A,C est appelé Raga Suddha Saveri en musique indienne, et mode Bac en musique vietnamienne (aussi utilisé en musique chinoise).

   — Le mode **G**,A,C,D,F est le raga Madhyamavati en musique indienne.

   — Le mode **D**,F,G,A,C,D est le raga Udayaravichandrika en musique indienne. (aussi utilisé en musique d'amérique du sud)

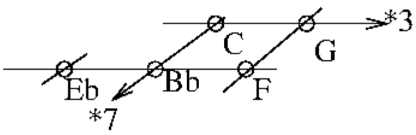
— Video de Bobby McFerrin sur l'universalité de la gamme pentatonique (colloque notes&neurons 2009)

— Gamme pentatonique mineure en namibie, D,F,G,A,C, ref : carnet de voyage avec Emmanuelle Olivier. Son

— Carnet voyages avec Pollo Vallejo, wagogo tanzanie, Rituel d'initiation. à 7'25" : gamme pentatonique sur G,A,B,D,F avec le F qui est en 7/8. Son

— Gamme pentatonique majeure sur le tonnetz : C,E,F,G,Bb Video

2.5.1.2 Gamme pentatonique mineur



— Dans l'ordre : C,Eb,F,G,Bb,(C).

  — elle n'est pas équivalente à la gamme pentatonique majeur Eb,F,G,Bb,C,Eb qui n'utilise que l'axe *3.

— Gamme pentatonique mineure sur le tonnetz : C,Eb,F,G,Bb Video.

2.5.1.3 La Gamme de Zarlino

- Zarlino, XVIeme siècle. Mise en valeur de la quinte juste et tierce juste.

A	E	B	
F	**C**	G	D

- Contient les notes justes des triades de Fa,Do,Sol, cad : (F,A,C),(C,E,G),(G,B,D).
- Cette gamme F,A,C,E,G,B,D est appelée gamme de "**l'intonation juste**".
- Le problème par ex. est que l'accord Dm=(D,F,A) sonne faux.

2.5.1.4 Gamme Pythagoricienne à 7 notes (heptatonique)

- Ce sont 7 notes en rapport de 3/2 (quinte juste).

F	**C**	G	D	A	E	B

- Utilisée au moyen age et renaissance, jusqu'au XVIe siecle, en théorie et en pratique. (rem : Il y a 6 gammes à 7 notes qui contiennent C,G.)
  - Utilisé en musique indienne : Raga Kalyani :

C	G	D	A	E	B	F#

Raga Thodi :

Db	Ab	Eb	Bb	F	**C**	G

2.5.1.5 Mode indien carnatique Shankarabharana

A'	E	B		
F	**C**	G	D	A

Référence : Pham (2003).

Remarques :

— Le problème est que la tierce F-A est fausse.

— A s'appelle “chatusruti dha”. Il est “lumineux, exultant, acidulé. A' s'appelle “trisruti dha” est calme, apaisé, triste.

— Ex cité par F. Pham p.10 Pham (2003) : Si C,D ; A',G ; G,F,E,D,C est joué plusieurs fois, puis C,D ; A,G ; G,F,E,D,C : joué à la fin. Alors le A sonne

2.5.1.6 Raga indiens

— Le Raga carnatique Mohana est

E			
C	G	D	A

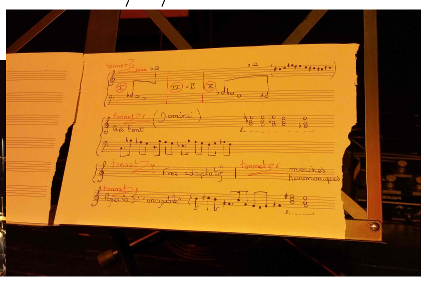
— C'est une gamme pentatonique.

— Le raga carnatique Hindolan (aussi en musique hindoustanie , où il s'appelle Malkauns) ?

Exemple 2.5.1. Exemple d'utilisation du raga Manirangu, is a janya raga, derived from Kharaharapriya which is 22nd on the Melakarta scale : raga Manirangu

2.5.1.7 Recherche avec le tempérament adaptatif

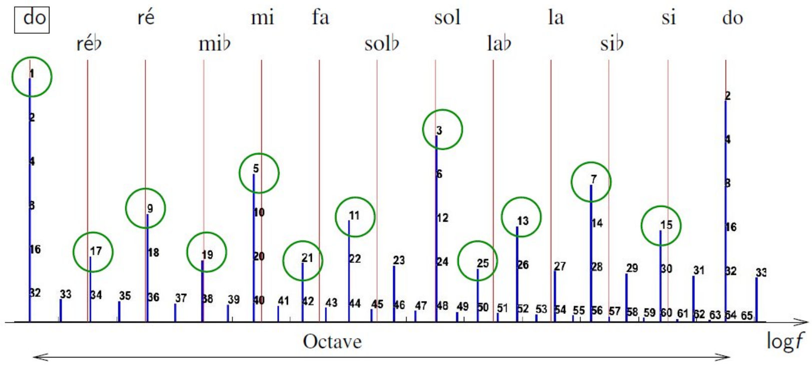
— Avec **Magic Malik en concert le 27/2/2015 :**



— on a fait des morceaux utilisant le **“tempérament des harmoniques”** (cercles verts) :

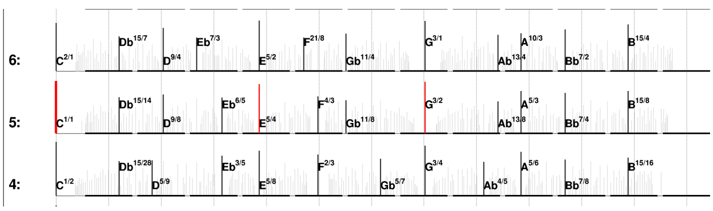
Son à 43”, , Video concert, à 30”.

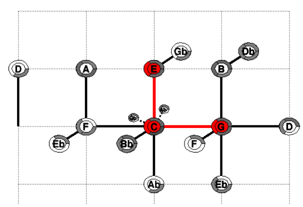
Son : à 7’03” et 11”10”.



— on a fait des morceaux utilisant le **“tempéraments box”** sur le réseau (3,5,7)Video

Définition 2.5.2. Si un accord $A = \{f_1, f_2, \dots f_n\}$ est présent, le **choix d’une prochaine note** de fréquence f_{n+1} dans un intervalle donné de 1 octave/12 ($=\frac{1}{2}$ton) est tel qu’il **minimise la dissonance** du nouvel accord $A' = A \cup \{f_{n+1}\}$.





— Remarques :

- — Invariant par transposition mais **non commutatif** : $C_0 - G_0 - Bb_0 - Eb_0 \neq C_0 - G_0 - Eb_0 - Bb_0$.
- — Décalage possible du « diapason » si on s'éloigne sur le tonnetz.
- — Le résultat dépend très peu du choix de la fonction $D(a/b) = \log(ab)$.

— Exemple du morceau du concert Video du tonnetz avec basse et sifflet. Son, concert à 6'10" et 8'10".

2.6 Traitements particuliers du son musical

2.6.1 Modification d'un son périodique

Comme l'a montré la question 2 de l'exercice 2, un son parfaitement périodique n'est pas très agréable à l'écoute. Il plus agréable de perturber ce son de différentes manières. Par exemple

- — Le **vibrato** est une modulation de la fréquence de la note. Cela se fait au violon, au trombone, à la voix.
- — Le **trémolo** est une modulation de l'amplitude de la note. Ce se fait à la flûte, à la voix.

Cf livre de castellegno page 266.

2.6.2 Détection du pitch d'un signal (presque) périodique

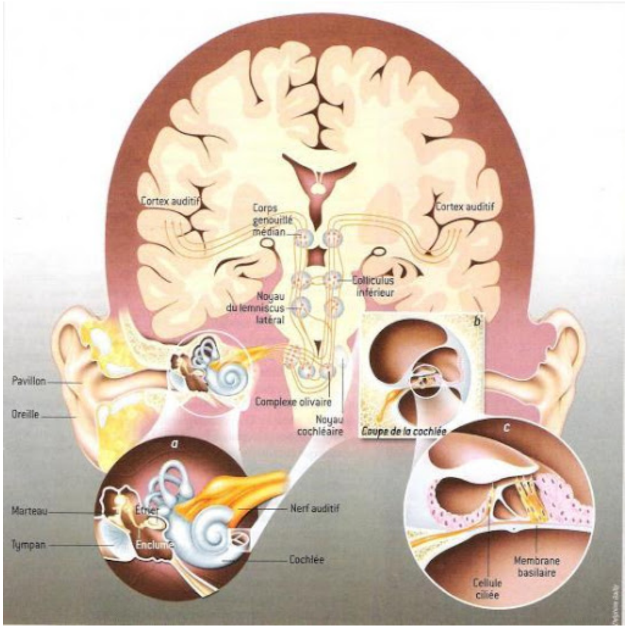
2.6.3 Filtres

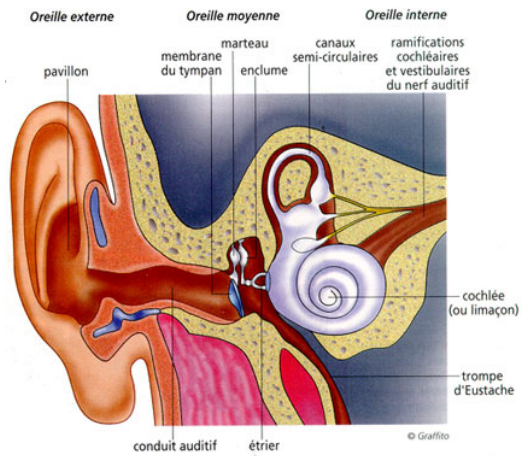
Chapitre 3

Perception du son

Dans ce chapitre on décrit le transfert d'information entre **un signal sonore à l'entrée d'une oreille** (ou de chaque oreille) jusqu'à la **perception consciente d'un son** et ses caractéristiques par l'individu. Tous les mécanismes physiologiques qui participent à ce transfert d'information, s'appelle le **système auditif**. Cette description sera en fait très partielle pour la principale raison que la plupart des mécanismes sont encore mal compris voir totalement inconnus comme ce qu'est la "la conscience" qui est l'étape finale de ce transfert. De plus le signal n'est pas vraiment transféré mais plutôt "analysé","filtré" et "transformé" de façon assez complexe par les étapes situés entre l'oreille et le conscient, appelé "inconscient". Seuls quelques unes des caractéristiques du signal sont extraites et transformées de multiples façon. Cela conduit à des illusions auditives que nous décrirons.

3.1 Description du système auditif





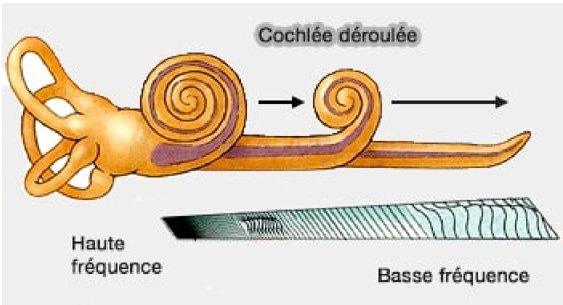
L'appareil auditif

Voici une description rapide des différentes étapes (plus détaillées ensuite) :

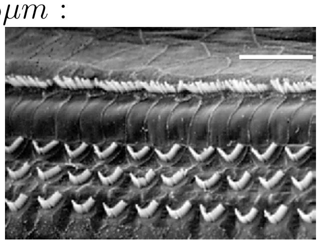
— Un partie de l'onde de pression (le son) qui arrive au voisinage des oreilles est diffusée par **le pavillon** (voir aussi pavillon). Une partie de l'onde entre dans le conduit auditif et atteint la **membrane du tympan**.

— La **membrane du tympan** se met en mouvement et transmet ce mouvement à la cochlée via **trois os** qui sont articulés : marteau, enclume, étrier. Il semblerait que la présence de ces os intermédiaires permette une bonne transmission de l'énergie sonore qui seraient réfléchie sinon. Un autre rôle de ces petits os est qu'ils servent de "fusible" : si le signal est trop fort, les os se bloquent et évitent de transmettre une énergie trop forte aux organes auditifs suivants.

— Dans la **cochlée** il y a un fluide liquide appelé endolymphe qui reçoit les mouvements de vibration de la part de l'enclume, et ces mouvements pénètrent dans la **cochlée** qui est formée en os et sous forme de conduits enroulés en colimaçon. Dans la cochlée, les éléments s'appellent l'organe de Corti : il y a une membrane appelée **membrane basilaire** dont la taille dépliée est 10cm par 1cm , et qui est en contact de tout son long avec des milliers de petits cils, appelées **cellules ciliées**. Les fréquences aiguës sont détectée au début de la cochlée alors que les fréquences graves pénètrent tout le long de la membrane basilaire jusqu'au bout de la cochlée.



— Chaque cil est comme un résonateur. Il peut se mettre en mouvement sous l'influence de la membrane basilaire. Il libère alors des ions qui excitent des terminaisons nerveuses de neurones, des dendrites. Photo de cils, à l'échelle $= 15\mu m$:



— Il y a donc des milliers de dendrites issues de ces cils (i.e. fils conducteurs) qui transmettent des influx nerveux venant spécifiquement de chaque cil. Toutes ces dendrites sont regroupées en parallèle dans le “nerf auditif” qui envoie cette information très riche au cerveau.

— L’information nerveuse est analysée par différentes parties du cerveau. Ne nombreux calculs sont effectués par des algorithmes mal connus ou non connus. Tout cela est inconscient, c’est à dire non perçu par la personne elle même. Certains résultats de tous ces calculs et analyses sont convoyées vers des zones du cerveau qui participent à la “conscience”, c’est à dire aux évènements perçus par la personne.

Voici une description de quelques aspects de cette analyse du son depuis son entrée dans le pavillon de l’oreille jusqu’à ce qu’il atteigne la conscience.

3.1.1 Le pavillon de l’oreille

Nous avons deux oreilles, chacune captant un signal sonore et le pavillon de chaque oreille est très “sculpté”. Cela semble essentiel à la perception tridimensionnelle du son.

— Regarder cette vidéo et les liens donnés qui suggère que la forme complexe du pavillon est utile à la localisation 3 dimensionnelle du son. En particulier regarder cette expérience que vous pouvez faire vous même : elle consiste à coller de la pâte à modeler sur son oreille pour modifier sa géométrie et remarquer, les yeux bandés, que l’on ne sait plus localiser l’origine spatiale des sons. Il paraît que si on garde la nouvelle géométrie quelques semaines, le cerveau se reprogramme, grâce à la synchronisation avec la vue, et la perception spatiale revient (mais re-disparaîtrait si on revenait à la forme initiale de l'oreille !).

— Dans cette vidéo de la BBC , on voit un renard qui se concentre sur la localisation 3 dimensionnelle de sa proie cachée sous une épaisse couche de neige.

Liens :

— Perception 3D : Pourquoi la forme de l'oreille est-elle si compliquée ?

— Localisation 3D d'une proie par un renard.

— Expérience : Virtual Baber shop or allumettes

— Enregistrement binaural sur wikipedia.

— Ref: Labo de Brams

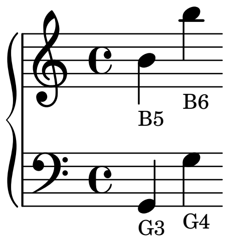
Exercice 3.1.1. (TP) **"Pavillon de l'oreille et localisation du son"**.

Essayer de reproduire les expériences ci-dessus.

3.1.2 Cils

Les cils sont situés dans la cochlée et fonctionnent comme des "résonateurs". Les paramètres importants d'un résonateur sont la fréquence de résonance f_0 et la largeur de résonance Δf_0. On peut convertir f_0 et Δf_0 en unité de pitch, c'est à dire demi-ton, d'après (2.3.1), donnant respectivement $x_0, \Delta x_0$. Voici l'ordre de grandeur pour différents cils, d'après Rossing Fletcher & Rossing (2012), p.74 :

f (Hz)	100	200	500	1000	2000	5000	10000
Δf (Hz)	90	90	110	150	280	700	1200
x_0 (en note)	G_3	G_4	B_5	B_6	B_7	Eb_8	Gb_8
Δx_0 (en $\frac{1}{2}$ ton)	15	7	4	2.5	2	2	3



Il est important de remarquer dans ces données que la précision Δx_0 en unité de pitch (demi-ton) est supérieure à 1 pitch (demi-ton) alors que la perception consciente donne une précision bien meilleure, de l'ordre de 1/10 pitch, voir Section 3.3.3. Cela suggère que le raffinement de la précision se fait dans une étape ultérieure de l'analyse, dans le cerveau.

Liens :

— Video d'un cil qui danse stimulé par un courant électrique.

— Faut-il repenser l'amplificateur cochléaire ?Christine Petit

3.1.3 Physiologie du cerveau

A la date actuelle (2021), on connaît très peu, voire pas du tout, les mécanismes précis d'analyse du son dans le cerveau. Le fonctionnement collectif du cerveau est encore mystérieux [1]. Par exemple on ne sait pas du tout où et comment fonctionne la mémoire, ni même ce qu'est précisément la “conscience” et où et comment fonctionne t-elle. Référence : Dehaene (2014). Voici cependant quelques observations et considérations.

— Noter que le temps de réponse d'un neurone est supérieures à $20ms$ soit une fréquence inférieure à $50Hz$. Ainsi l'analyse du timbre d'un signal audible $u(t)$ sur des fréquences $f \geq 100\text{Hz}$ ne peut pas être une analyse temporelle faite dans le cerveau mais serait plutôt une analyse de la transformée de Fourier $|(\mathcal{F}u)(f)|$? Cela se confirme par le fait que notre perception est insensible aux phases ? Ce temps de réponse détermine la **latence acceptable en musique** : $\Delta t = 20\text{ms}$, i.e. un décalage temporel perceptible. On a déjà vu que sur cette durée Δt le son se propage dans l'air sur une distance $L = c\Delta t = 343 \times 20.10^{-3} m = 7m$. C'est donc une distance au dessus de laquelle la latence peut être perceptible.

— Par imagerie IRM, on observe que les calculs liés à la parole et la résolution temporelle du son sont effectués dans la partie gauche du

1. Remarquer par exemple que par rapport à il y a 10000 ans, les progrès de la connaissance scientifique ont apporté des bouleversement dans nos conditions de vie : transport, logement, nourriture, médecine de la plupart des organes de notre corps. Une exception notable est le cerveau qui reste totalement mystérieux. Une conséquence est que tous les problèmes reliés aux mauvais fonctionnement du cerveau, états d'humeur des personnes, personnalités extravagantes, anxiété, maladie mentales, affectant les relations humaines, tous ces problèmes sont présents de nos jours comme il y a 10000 ans. Quasiment aucun progrès n'a eu lieu sur cet aspect déterminant de la qualité de vie.

cerveau, alors que l'analyse de la résolution fréquentielle, dont la musique sont effectués dans la partie droite, cf Conférence de Christine Petit à 32'.

Liens

- From Tempo to Tears: Early Level Processing of Sound Rhythms in the Mammalian Brain by Anna Montell Magnusson
- From Perception to Pleasure: Music and its Neural Substrates Robert Zatorre
- Cours de Christine Petit au collège de France
- Christine Petit: Perception du rythme et de la métrique

3.2 La voix et les signaux periodiques

L'appareil auditif est très adapté à la perception de la voix humaine. Pour cette raisons, dans cette section nous décrivons la voix humaine, comment elle est générée et perçue.

Chez les humains, la voix est apparue entre -2 millions d'années (sons) et-50000 ans (parole). Des recherches récentes font remonter l'apparition de la parole à -20.10^6 ans. La parole a un rôle social pour la communication, pour palabrer, pour l'échange d'informations ...

On peut imaginer que notre cerveau est programmé (de façon inné et/ou acquise) pour percevoir particulièrement (i.e. analyser) certains sons qui nous sont importants comme :

— Sons de la voix humaine
— Sons d'une source d'eau, de la pluie, de l'orage
— Bruits inquiétants dans le silence : craquements de branches, respirations
— Bruits du cœur de sa maman, de la respiration ...

Références :

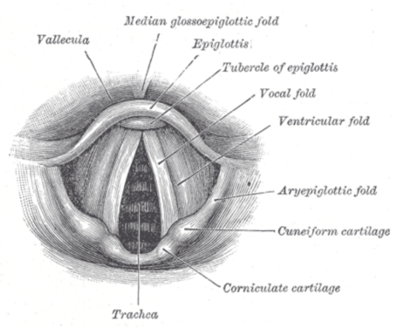
— voir cet article sur l'origine du langage.
— Ce livre de Harari Dauzat & Harari (2015) sur l'histoire et les aptitudes de l'espèce Homo Sapiens.
— Schroeder *et al.* (2007, chap.16).
— Voix du chien Husky, Un phoque qui parle, Vocalisations du signe Marmoset
— Un piano qui parle piano qui parle (2), piano qui parle (3), pinao qui

parle (4)

3.2.1 Observations générales sur la voix

3.2.1.1 Cordes vocales, forme de la cavité buccale

La voix est générée par la mise en vibration des “cordes vocales” qui ne sont pas véritablement des “cordes” mais des muscles sous tension, sous le flux d'air venant des poumons :



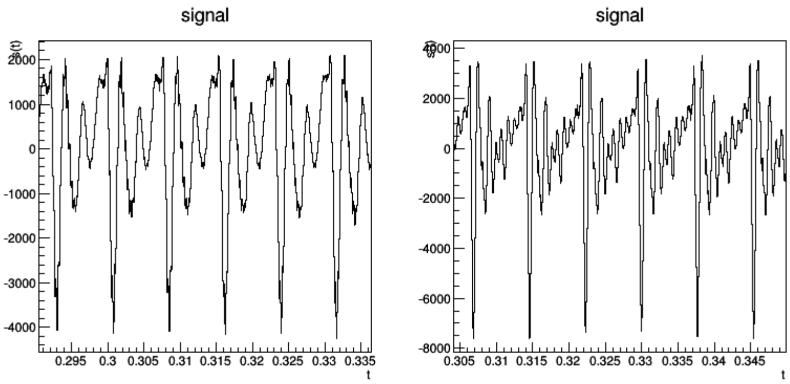
Video des “cordes vocales” en action.

Les cordes vocales produisent un signal sonore (presque) périodique que l'on qualifie de **voyelle**. Voir Section 2.3 sur les signaux périodiques. En fait toute la cavité buccale participe à sculpter ce signal. Cela correspond à des positions précises de la langue, du palais, etc.. Voir cette Video.

Le signal sonore produit et donc la voyelle prononcée dépend de la forme de la cavité buccale.

Exemple 3.2.1. Par exemple, voici des voyelles chantées par Malik Mezzadri sur une fréquence fixe de $f = 130\text{Hz}$ (note C_4). Voici le signal (mesuré avec un micro) pour les voyelles **“ON”** et **“A”** sur quelques périodes $T = 1/f = 7\,ms$.

On observe en effet que le signal est (presque) périodique et dépend de la voyelle :



Remarque 3.2.2. la vitesse du son dans l'air étant $v = 330m/s$ la longueur d'onde est donc $\lambda = v/f = 2,50m$. L'intensité est de l'ordre de 70dB.

3.3 Du signal sonore à la perception consciente

Dans la section 3.1, on a décrit le système auditif. Dans cette Section, on va considérer le système auditif comme une "boite noire", c'est à dire dont on ne connaît ni le contenu ni le fonctionnement. Seulement, on va considérer que le système auditif reçoit en entrée : un signal sonore, et donne en sortie une "perception consciente" c'est à dire des sensations que l'on peut exprimer. On va essayer d'identifier certaines relations entre des entrée et sorties, que l'on appelle "perception".

3.3.1 Perception du temps

On estime que l'on perçoit la position temporelle d'un signal avec une précision supérieure à

$$\Delta t \geq 0.1\text{s} \tag{3.3.1}$$

3.3.2 Perception de l'intensité

En Section **??** on a défini l'intensité d'un signal sonore en décibels (dB). L'intervalle maximal de perception à $f \sim 1000$Hz est

$$I \in [0, 120] \quad \text{dB}.$$

où $I = 0$dB est le seuil minimal de perception, $I = 120$dB est le maximum (ensuite c'est intolérable, voir destructeur) et la précision de perception vers 1000Hz est

$$\Delta I \geq 1 \text{ dB}$$

C'est à dire que l'on ne distingue pas de différence entre deux signaux $u_2 = \lambda u_1$ avec une constante de proportionnalité

$$\frac{1}{10^{1/10}} = 0.8 < \lambda < 10^{1/10} = 1.26.$$

.

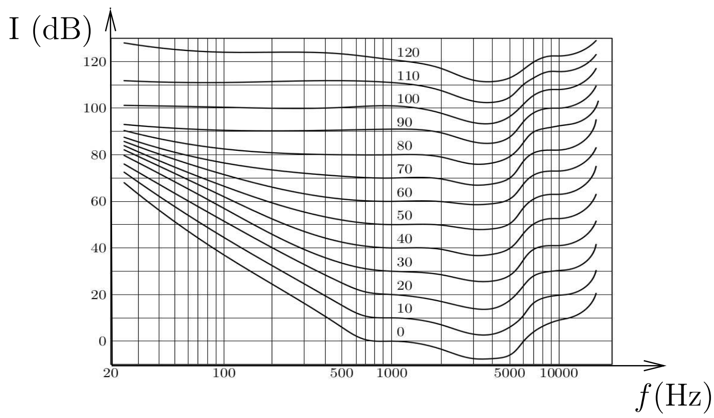


FIGURE 3.3.1 – Ce diagramme appelé courbes isophones de Fletcher-Muson montre des valeurs (f, I) dans le plan fréquence-Intensité qui donne la même perception d'intensité. Par exemple un signal $I = 40\text{dB}$ à $f = 100\text{Hz}$ est à peine audible, comme $I = 0\text{dB}$ à $f = 1000\text{Hz}$.

3.3.2.1 Courbes isophones

— Voir aussi la définition des décibels dBA SPL, voir Benson (n.d., p.10). Les courbes de la figure 3.3.1 sont mesurées par des expériences de psychoacoustiques avec des signaux sinusoidaux : $u(t) = A \sin(\omega t)$, avec $\omega = 2\pi f$

— Observer que sur la figure 3.3.1 l'écart entre les courbes isophones dépend de la fréquence. Une conséquence importante est que en musique, pour un "mixage" de deux (ou plus) instruments jouant à des hauteurs (fréquences) différentes, si l'on souhaite réduire l'ensemble de "10dB" par exemple, il faudra réduire différemment les intensités de chacun selon sa fréquence.

3.3.2.2 Réverbération

Par définition, en acoustique musicale, si un signal s'éteint progressivement, on dit qu'il y a de la **réverbération**. Plus précisément

Définition 3.3.1. Si un signal décroît exponentiellement comme

$$u(t) = u(0)\,10^{-t/\tau}$$

alors on appelle

$$\tau_{\text{reverb}} := 6\tau$$

le **temps de réverbération**. De façon équivalente

$$u(t+\tau_{\text{everb}}) = 10^{-6} u(t)$$

$$I(t+\tau_{\text{everb}}) = I(t) - 60\text{ dB}$$

3.3.2.3 Mesure de l'intensité dans les synthétiseurs

Les synthétiseurs communiquent avec le langage MIDI. L'intensité d'un son est codé par le paramètre appelé "velocité" $v \in [0, 127]$, qui est proportionnel à l'amplitude du signal d'après les spécifications MIDI. Cela est curieux, on penserait plutôt à v fonction affine du log du l'amplitude ?

— Voir ce document.

3.3.2.4 Effet de masquage

Un son grave suffisamment intense masque un son aigu plus faible, c'est à dire que l'on ne perçoit plus le son aigu. Mais l'inverse n'est pas vrai : un son grave faible est toujours perçu même en présence d'un fort son aigu.

Exemple 3.3.2. si il passe un train générant des sons graves forts très de nous, alors on ne s'entend plus parler. Mais si une alarme aiguë sonne, on s'entend encore parler.

L'explication physiologique de ce phénomène serait que le long de la membrane basilaire dans la cochlée, les capteurs des sons aigus sont à l'entrée et ceux des sons graves sont au bout terminal. Le signal d'un son grave passe devant les capteurs des sons aigu pour aller au bout, mais pas l'inverse.

3.3.3 Perception du pitch des notes (fréquences)

On entend seulement les fréquences f comprises dans l'intervalle

$$f \in [20 \text{ Hz}, 20000 \text{ Hz}] .$$

Cela dépend des animaux Benson (n.d., p.9).

Pour des notes pures (i.e. signal périodique sinusoïdal qui est un mode de Fourier) la précision de perception de la fréquence f est en unité de pitch

f (Hz)	250	500	1000	2000	8000
Δx (pitch)	0.17	0.03	0.05	0.04	0.16

Cela est extrêmement précis, largement inférieur à la précision en fréquence des cils, voir Section 3.1.2, par exemple à $f = 1000Hz$:

$$\Delta x = \frac{1}{40} (\Delta x)_{cils}$$

On pense que cette grande précision de perception du pitch est obtenue grâce à des algorithmes dans le cerveau, non encore observés ni compris.

Remarque 3.3.3. La précision de perception du pitch indiquée dans le tableau précédent montre que l'échelle des notes tempérées, Définition **??**, écartées de $\Delta x = 1$ pitch, est très grossière par rapport à notre perception.

Remarque 3.3.4. A cause de cette grande précision en pitch de notre perception, la fabrication d'un instrument de musique demande une grande précision, inférieure à $10^{-2}mm$ Hirschberg *et al.* (1996). (ref : **article 1996, Aeroacoustics of Musical Instruments)**

Liens :

- Christine Petit: Perception de la hauteur sonore, en particulier :
  - à 55'50" : Il existe un "patron dans le cerveau" de detection des harmoniques ?
  - à 57" : detection de son graves grace aux harmoniques.
  - à 1h02' : localisation de la zone « centre du pitch » dans le cerveau.
- Moore: The Perception of Pitch
- Jacob Collier explains that Piano is not in tune and shows that major third is out of tune by $+0.14$ pitch (half tone).

3.3.4 Perception et principe d'incertitude en temps-fréquence

Nous avons mentionné en Section **??**, le principe d'incertitude en temps fréquence, concernant la quantité d'information limitée par la mesure $dt\,df$.

Par exemple, considérons une note C_5 de pitch $x = 60$ avec une précision $\Delta x = 0.5$ (quart de ton) soit $\frac{\Delta f}{f} \underset{(??)}{=} 0.057\,\Delta x$, et de durée $\Delta t = 0.2s$, supérieure au plus petit intervalle de temps perceptible (3.3.1). Cela donne $f = 260$Hz et

$$\Delta t \Delta f = 0.2 \times 260 \times 0.057 \times 0.5 \sim 1.5 \geq 1$$

juste supérieure à 1 donc à la limite du principe d'incertitude. Cela montre que notre perception de l'information dans le plan temps-fréquence est quasi optimale.

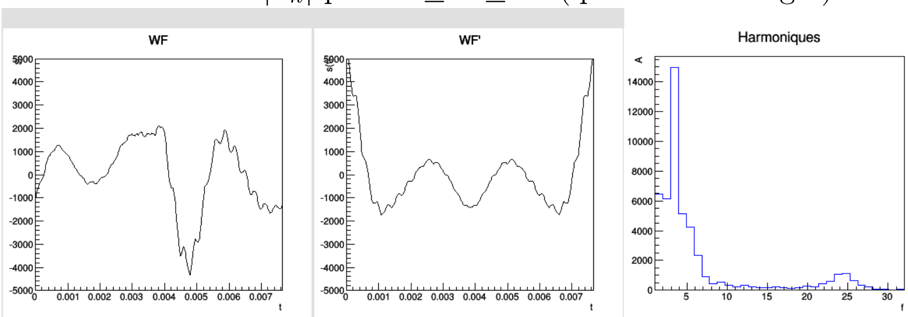
Pour les notes graves, i.e. x petit, la limite $\Delta t \Delta f = 1$ est plus facilement atteinte. Par exemple il est impossible en principe de détecter le pitch précis (au demi-ton) des notes d'une contrebasse qui joue un walking bass rapide (sauf si elle a un timbre riche en harmoniques élevées).

3.3.5 Non perception de la phase

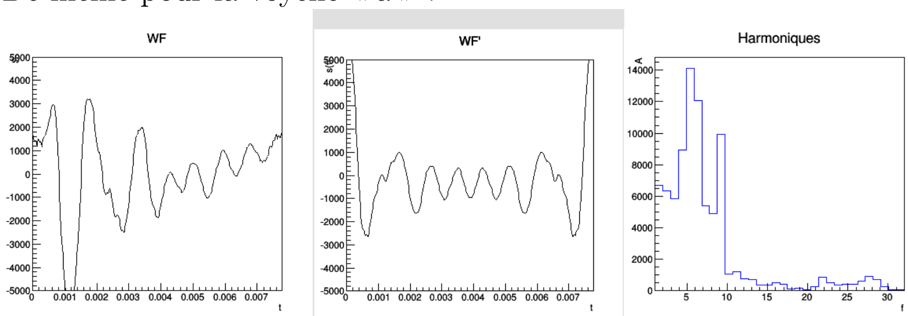
On rappelle d'après le Théorème de Fourier A.3.14 et la définition A.3.3 que dans la décomposition de Fourier d'un signal périodique en signaux sinusoidaux, chaque composante (ou harmonique) de fréquence $f_n = nf$ avec $n = 1, 2, 3, \ldots$ a une phase à l'origine notée ϕ_n (en plus de l'amplitude A_n qui est perceptible via le timbre).

Proposition 3.3.5. ***"Non perception de la phase"***. *Dans les exemples suivants on observe que nous ne percevons pas du tout les phases* $(\phi_n)_n$.

Exemple 3.3.6. Voici un Son $u(t)$. Voici maintenant un **Son** $v(t)$ où toutes les phases sont mises à $\varphi_n = 0, \forall n$. Voici pour la voyelle « on », les formes d'ondes d'origine $u(t)$ et $v(t)$ sur une période, ainsi que les modules des coefficients de Fourier $|\hat{u}_n|$ pour $0 \leq n \leq 32$ (qui sont inchangés) :



De même pour la voyelle « a » :



On remarque que $u(t)$ et $v(t)$ sont très différentes. En particulier $v(t)$ a un pic en $t = 0$ à cause du choix particulier $\varphi_n = 0$.

Exemple 3.3.7. Voici un autre exemple de son artificiel $u(t)$, où chaque harmonique $n = 1, 2 \ldots, 30$ a un module $|\hat{u}_n| = 1/n$ et une **phase** ϕ_n **aléa-**

toire. On représente aussi le signal $u(t)$ sur une période. Remarquer que lorsque les phase sont nulles, le signal $u(t)$ est en « dent de scie ». On le perçoit légèrement, peut être à cause de la “singularité” en zéro qui fausse les rendu.

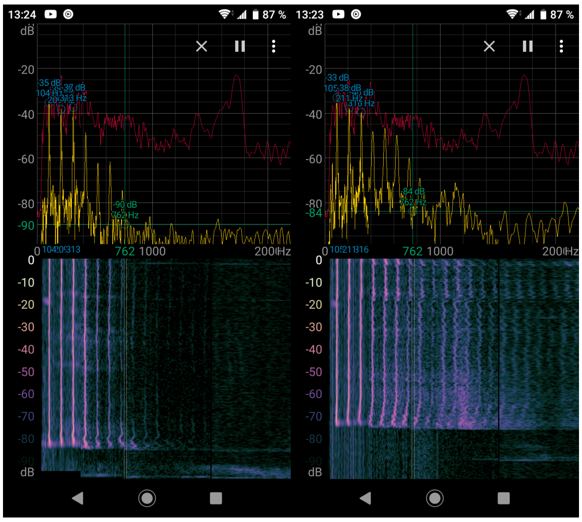
Remarque 3.3.8. Cette non perception des phases ϕ_n suggère que parmi toute l'information du signal, le cerveau perçoit des notes pures quasi-indépendantes et mesure seulement leur fréquence $f_n = \frac{n}{T} = nf$ et intensité A_n.

3.3.6 Perception du timbre

On rappelle d'après le Théorème de Fourier A.3.14 et la définition A.3.3 que dans la décomposition de Fourier d'un signal périodique en signaux sinusoidaux, chaque composante (ou harmonique) de fréquence $f_n = nf$ avec $n = 1, 2, 3, \ldots$ a une phase à l'origine notée ϕ_n et une amplitude A_n. L'ensemble de ces amplitudes $(A_1, A_2, A_3, \ldots)$ s'appelle le **timbre** du signal. (On a vu que les phases $\phi_1, \phi_2, \ldots$ ne sont pas perceptibles).

3.3.6.1 Le timbre et les voyelles. Les formants

Exemple 3.3.9. Et voici sur une échelle plus large en fréquences, un “O” et un “A” de même fréquence $f = 104$Hz. On observe que le “A” a plus d'intensité dans les harmoniques élevées (autour de 1000Hz). C'est ainsi que on distingue les voyelles entre elles.



Remarque 3.3.10. On peut aussi communiquer en parlant sans les cordes vocales, i.e. en “chuchotant”. L’avantage de la voix chantée (i.e. avec les cordes vocales) est peut être qu’elle est mieux perçue parmi d’autres sons ou bruits.

Dans la voix parlée, le timbre est associé à des voyelles, voir l’exemple 3.3.6.

Question 3.3.11. *Comment associer une suite* $(A_n)_{n\geq 1}$ *à une certaine voyelle ?*

Observations : Considérons cet enregistrement d’un signale $u(t)$ de voyelles différentes, i.e. timbres différents mais de même pitch $x = 48$, i.e. $f = 130$Hz, la note est C_4.

— Dans cet exemple, le signal $u(t)$ est filtré, ne gardant que avec les harmoniques 1,2,3 seulement. On constate que **on ne perçoit plus les voyelles**. On perçoit toujours le même pitch x.

— Dans cet exemple, le signal $u(t)$ est filtré, ne gardant que avec les harmoniques $n \geq 5$. On constate que **l’on perçoit les voyelles**. De façon plus surprenante, on perçoit toujours le même pitch x, i.e.

la fréquence f_1 alors qu'elle n'est pas présente dans la transformée de Fourier. La perception des voyelles est surtout liée aux formants, expliqués ci-dessous.

Pour la voix humaine, on observe que l'enveloppe des coefficients de Fourier $(A_n)_n$ possède quelques maxima à des fréquences $F_1, F_2, \ldots$ appelés **formants**. Ces fréquences $F_1, F_2, \ldots$ dépendent de la voyelle prononcée.

D'après Chevalier p.53, la valeur des deux premiers formants F_1, F_2 suffit à déterminer la voyelle prononcée d'après le schéma suivant. Les formants F_3, F_4 apportent des nuances. Certains chanteurs ont un formant supplémentaire spécial entre F_3 et F_4.

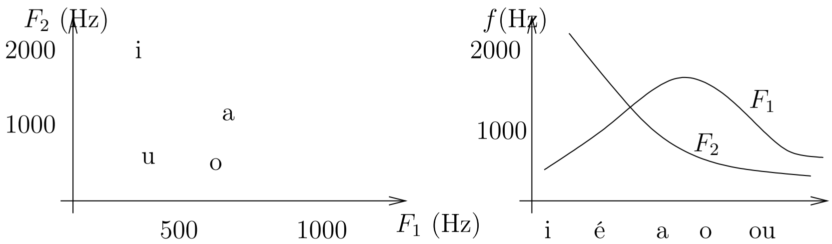


FIGURE 3.3.2 – Valeurs des formants F_1, F_2 en Hz, pour différentes voyelles.

— Exemples sur les formants : voir livre de castellegno page 453 et page 454, page 455

Lien:

— La perception des voyelles n'est pas dans les premières harmoniques: voir Section 2.4 de cet exposé.

— Christine Petit: Perception de la hauteur sonore à 55'50" : Il existe un "patron dans le cerveau" de détection des harmoniques. 57'30" : identifications de neurones (chez le marmoset) qui détectent la hauteur tonale. Identification dans le cerveau du "**centre du pitch**".

3.3.6.2 Paradoxe de la basse virtuelle

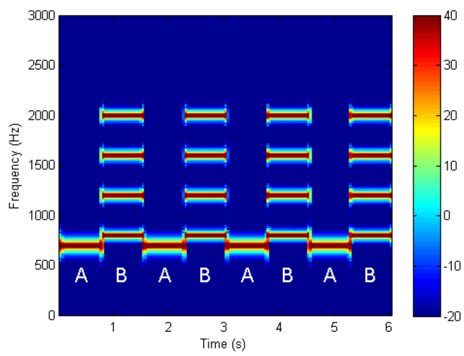
La conclusion importante de observations précédentes est que :

Conclusion 3.3.12. Lorsqu'on entend un signal périodique, on perçoit le timbre

$$(A_n)_{n\geq 1} = (A_1, A_2, A_3, \ldots)$$

et le pitch x même si les premiers coefficients sont nuls, par exemple $A_1 = 0, A_2 = 0, A_3 = 0, A_4 = 0$.

Exemple 3.3.13. Dans cet exemple en entend successivement deux notes A, B. La note A est une note pure de fréquence fondamentale $f_1(A) = 700$Hz. La note B est une note timbrée avec $f_1(B) = 400$Hz, mais pour la note B on a enlevé l'harmonique $f_1(B) = 400$Hz (en posant $A_1 = 0$) de sorte que la fréquence la plus basse émise est $f_2(B) = 800$Hz. A l'écoute on entend cependant $f_1(B) = 400$Hz qui n'est pas présente. (Remarquer que $f_1(A)/f_1(B) = 7/4$ de pitch $x_1(A) - x_2(B) = \frac{12}{\ln 2}\ln\left(\frac{7}{4}\right) = 9.68 \sim 10$ qui est une septième mineure).



Exercice 3.3.14. (TP) « **Illusion de la basse fondamentale manquante** ». Avec audacity Menu/Générer/Tonalité, créer differents signaux sinusoidaux d'amplitude $A = 0.1$ et de fréquences $f_1 = 200$Hz, $f_5 = 1000$Hz, $f_6 = 1200$Hz, $f_7 = 1400$Hz. Remarquer que ce sont des harmoniques de la fréquence fondamentale f_1. Ecouter simultanément les signaux de fréquence f_5, f_6, f_7. Entendez vous la fréquence f_1 ? Cela s'appelle le mystère de la fondamentale manquante actuellement pas très bien compris. On pense que le mécanisme est au niveau du cerveau, et serait utile pour analyser les stimulis venant d'une voix humaine incomplète (dans la nature, la fondamentale peut être brouillée par d'autres bruits).

Remarque : le phénomène du « **troisième son** » ou Combination tone est différent : il fait ressort $f_2 - f_1$ (effets non linéaire dans l'oreille ?). Ecouter par exemple ex. de 3eme son avec $f_1 = 440$ (A_5), $f_2 = 704$ (F_6), on percoit $f_2 - f_1 = 264$ (C_5), mais ici la basse manquante serait $f = f_2/8 = f_1/5 = 88$ (F_4).

Liens:

- voir Section 2.4 de cet exposé.
- https://auditoryneuroscience.com/pitch/missing-fundamentalsMissing funcdamental
- Christine Petit: Perception de la hauteur sonore à 56' : detection de son graves grace aux harmoniques.
- Livre de Castellegno Page 258 : illusion de notes dans une cloche d'église.

3.3.6.3 Fréquences aléatoires :

Voici un exemple artificiel, où des notes pures sont d'amplitude $A_k = 1/k$, de phase nulle, et de fréquence $f_k = k \cdot f_1 \cdot (1 + \epsilon_k)$ avec $f_1 = 130\text{Hz}$ (do) et $\epsilon_k \in [-\epsilon, \epsilon]$ aléatoire. On écoute et représente (f_k, A_k) pour différentes valeurs de ϵ qui mesure la fluctuation des fréquences. Remarquer que l'on perçoit parfois plusieurs notes (par exemple pour $\epsilon = 0.084$). Remarquer que les pics de fréquences f_k, f_{k+1} peuvent se croiser si $\epsilon > 1/k$.

Est-ce en rapport avec l'expérience suivante ?

Y. Menuhin, racontant un étrange concert dans un champ de fleurs à Mykonos : Référence : "La légende du violon", Y Menuhin, Flammarion

> « Je m'arrêtai pour écouter ce bourdonnement entêtant quand, soudain, je m'aperçus qu'il ne s'agissait pas d'un bruit confus et désordonné, mais bien au contraire d'un accord secret de la création : les abeilles émettaient deux notes, et qui plus est ces deux notes formaient une quinte. Je ne sais comment analyser ce prodige de la nature. Toujours est-il que la recherche de la quinte semble constituer une nécessité primitive de l'univers : nous émettons et nous recevons des vibrations. Ces vibrations obéissent à un ordonnancement secret. Une harmonie occulte se met en place. Certains prétendent même que les intervalles des planètes, c'est-à-dire les distances qui les séparent les unes des autres, sont en correspondance intime avec les intervalles de la musique : quintes, tierces, octaves ne seraient que des images musicales des intervalles célestes. »

3.4 Perceptions des intervalles justes et accords justes

3.4.1 Perception des intervalles justes

Comme les rapports de fréquence rationnels $\frac{f'}{f} = \frac{a}{b}$ en (**??**) correspondent aux intervalles juste, voir Section **??**, l'hypothèse (**??**) signifierait que nous percevions particulièrement les intervalles justes.

— **Rappel :** pour une note individuelle, i.e. un signal périodique provenant d'une voix ou d'un instrument, on ne perçoit pas ses harmoniques individuelle, mais on perçoit le timbre $(A_n)_{n\geq 1}$ qui donne une impression collective, interprétée par le cerveau.

— **Question :** que perçoit-on si deux notes (timbrées) de fréquences $f_1 \leq f_2$ sonnent ensemble ?

Voici quelques exemples.

3.4.1.1 Exemples

— Video qui montre une résonance à l'unisson $\frac{f_2}{f_1} = \frac{1}{1}$.

— Video qui montre une résonance à l'octave $\frac{f_2}{f_1} = \frac{2}{1}$.

— Video qui montre une résonance à la quinte $\frac{f_2}{f_1} = \frac{3}{2}$.

— Video qui montre les résonances $\frac{f_2}{f_1} = \frac{a}{b}$ importantes.

— Video qui montre les résonances $\frac{f_2}{f_1} = \frac{a}{b}$ dans l'ordre descendant.

— Video qui montre les résonances $\frac{f_2}{f_1} = \frac{a}{b}$ avec le rajout de la "basse" $f_b = \frac{f_2}{a} = \frac{f_1}{b}$. (la fondamentale manquante en rouge)

— Video qui montre les résonances $\frac{f_2}{f_1} = \frac{a}{b}$ avec la "basse" $f_b = \frac{f_2}{a} = \frac{f_1}{b}$

en rouge et le “sifflet” $f_s = af_1 = bf_2$ en vert.

Résumé 3.4.1. un **intervalle juste** est un ensemble de deux notes dont le rapport des fréquences est $f_2/f_1 = a/b$ avec a, b petits entiers. On ressent cet intervalle comme “résonant” ou “consonant”.

3.4.1.2 Perception des intervalles justes

Sur cette page web.

3.4.2 Perception des accords justes

Voir la Section **??** qui définit le réseau tonnetz, permettant de représenter les accords justes.

Annexe A

Notions de base utiles en mathématiques

Dans ce cours on ne suppose aucune base en science. Cependant on aura besoin de quelques notions élémentaires qui sont expliquées dans cette section.

A.1 Les fractions et équations du premier degré

A.1.1 Additions et soustractions

On désigne par exemple une grandeur numérique recherchée par la lettre x. Par exemple x est le poids d'une pomme, et l'on sait que le poids du saladier est 0.8kg et on pèse au total 1kg pour le saladier qui contient la pomme. On écrit

$$x + 0.8 = 1$$

La solution est

$$x = 1 - 0.8 = 0.2\,\mathrm{kg}$$

Plus généralement, si a, b, c son des chiffres, on a l'équivalence suivante des équations :

$$\begin{aligned} a + b = c \qquad &\Leftrightarrow b = c - a \qquad \Leftrightarrow a = c - b \\ \Leftrightarrow 0 = c - b - a \qquad &\Leftrightarrow 0 = a + b - c \end{aligned}$$

Exercice A.1.1. Vérifiez ces équations avec les valeurs $a = 3$, $b = 2$, $c = 5$.

A.1.2 Multiplication et divisions

On désigne par exemple une grandeur numérique recherchée par la lettre x. Par exemple x est le poids d'une pomme, et l'on sait que 10 pommes pèsent 2kg. On écrit

$$10 \times x = 2 \text{ ou } 10x = 2.$$

La solution est

$$x = \frac{2}{10} = 0.2\,\mathrm{kg}$$

Plus généralement, si a, b, c son des chiffres non nuls, on a l'équivalence suivante des équations :

$$ab = c \Leftrightarrow b = \frac{c}{a} \Leftrightarrow a = \frac{c}{b} \Leftrightarrow 1 = \frac{c}{ab} \Leftrightarrow 1 = \frac{ab}{c}$$

Exercice A.1.2. Vérifiez ces équations avec les valeurs $a = 3$, $b = 2$, $c = 6$.

Notations de puissance :

$$x^2 = x \times x, \qquad x^3 = x \times x \times x, \quad \text{etc}$$

$$x^{-1} = \frac{1}{x}, \quad x^{-2} = \frac{1}{x \times x} = \frac{1}{x^2}, \quad \text{etc}$$

Par commodité, $x^0 = 1$. Quand on écrit x^n, l'entier n s'appelle l'**exposant**.

A.2 Exposants, logarithme et décibels

On présente quels notions importantes pour la suite. Cela peut concerner toute sorte de grandeur physique.

Dans le tableau suivant on a écrit sur la première ligne différentes valeurs d'une grandeur x. Par exemple $x = 100$.

Sur la deuxième ligne, on écrit cette même valeur avec un **exposant**, par exemple $x = 10^2$. La troisième ligne donne la valeur de l'exposant, aussi appelé le **logarithme** de x et noté $\mathrm{Log}_{10}(x)$. Par exemple $\mathrm{Log}_{10}(100) = 2$. La dernière ligne est appelée décibels, c'est seulement l'exposant multiplié par 10, i.e.

$$\mathrm{dB}(x) = 10 \times \mathrm{Log}_{10}(x)$$

Par exemple $\mathrm{dB}(100) = 20$.

Valeurs de x et opérations :	...	$\xrightarrow{\times 10}$	0.1	$\xrightarrow{\times 10}$	1	$\xrightarrow{\times 10}$	10	$\xrightarrow{\times 10}$	100	$\xrightarrow{\times 10}$	100
écriture de x avec l'exposant			$= 10^{-1}$		$= 10^{0}$		$= 10^{1}$		$= 10^{2}$		= 1
$\mathrm{Log}_{10}(x)$: exposant ou logarithme (en base 10)			-1		0		1		2		3
$\mathrm{dB}(x) = 10\,\mathrm{Log}_{10}(x)$:décibel			-10		0		10		20		3

Remarque A.2.1. Sur la ligne de x on passe d'une colonne à la suivante de droite en multipliant par 10. Donc on passe à la colonne de gauche en divisant par 10. Alors dans sur la ligne des logarithme (et des décibels) on passe d'une colonne à la suivante de droite en ajoutant 1 (ou ajoutant 10 pour les décibels). Cette remarque montre que **la fonction logarithme transforme les opérations multiplication en addition**. C'est l'intérêt essentiel de cette fonction. Par exemple $3 = \mathrm{Log}_{10}(100 \times 10) = \mathrm{Log}_{10}(100) + \mathrm{Log}_{10}(10) = 2 + 1$. On retiendra que pour tout nombre positif x, y,

$$\mathrm{Log}_{10}(x \times y) = \mathrm{Log}_{10}(x) + \mathrm{Log}_{10}(y).$$

Remarque A.2.2. Concernant le **cas particulier où on mesure une puissance sonore** p, le décibel est utilisé de sorte à ce que 0 décibels correspond à une puissance de référence, qui est la limite de l'inaudible (i.e. très faible), et correspond à la valeur $p_0 = 10^{-12} W/m^2$. Pour une puissance sonore quelconque p on notera alors $x = \frac{p}{p_0}$ (sans unité) de sorte que $\mathrm{dB}\left(\frac{p_0}{p_0}\right) = \mathrm{dB}(1) = 0$. Ajouter +1dB correspond alors au seuil de perception du changement de puissance. Il apparaît que $\mathrm{dB}(x) = 120\mathrm{dB}$ est un son très (trop) fort au seuil de la douleur, qui est $p = 1W/m^2$.

Remarque A.2.3. Le tableau s'étend à droite et à gauche. Comme valeurs notables,

— il y a $x = 1000000 = 10^6$ appelé million, donnant $\mathrm{Log}_{10}(x) = 6$ et $\mathrm{dB}(x) = 60$.

— il y a $x = 1000\,000\,000 = 10^9$ appelé milliard, donnant $\mathrm{Log}_{10}(x) = 9$ et $\mathrm{dB}(x) = 90$.

— il y a $x = 0.000000001 = 10^{-9}$ appelé nano, donnant $\mathrm{Log}_{10}(x) = -9$ et $\mathrm{dB}(x) = -90$.

— Voici des noms couramment utilisés pour désigner ces exposants :

x :	10^{-9}	10^3	10^6	10^9	10^{12}
nom :	nano	kilo	Mega	Giga	Tera

Par exemple en informatique $2\ 10^9$ octets se dit 2 Giga-octets (on verra une valeur plus précise plus loin).

Exercice A.2.4. Entre la puissance du son le plus fort (seuil de douleur) et le plus faible (perceptible) il y a une différence de 120dB. Quel est le rapport de puissance correspondant ?

Solution A.2.5. Le rapport est $10^{120/10} = 10^{12} =$ 1000 milliard !

Exercice A.2.6. Compléter le tableau suivant

Instrument	trombone		clarinette		voix humaine		chanteur
	fortissimo		fortissimo		parlant		fortissimo
Puissance en W	$5\,W$		0.05W		$5\,10^{-5}W$		0.05W
rapports de puissance		$\xleftarrow{\times ?}$		$\xleftarrow{\times ?}$		$\xrightarrow{\times ?}$	
Puissance en dB	?		?		60		?

Solution A.2.7. Voici la solution :

Instrument	trombone		clarinette		voix humaine		chanteur
	fortissimo		fortissimo				fortissimo
Puissance en W	$5\,W$		0.05W		$5\,10^{-5}W$		0.05W
rapports de puissance		$\xleftarrow{\times 100}$		$\xleftarrow{\times 1000}$		$\xrightarrow{\times 1000}$	
Puissance en dB	110		90		60		90

Remarque A.2.8. La fonction $x \to \mathrm{Log}_{10}(x)$ n'est pas restreinte à $x = 1, 10, 100, \ldots$ mais s'étend à toutes les valeurs intermédiaires. Par exemple $\mathrm{Log}_{10}(2) = 0.301\ldots$ Sur une calculatrice, c'est la touche **Log**. Dans le moteur de recherche google, écrire Log(2) par exemple. On définit toujours les décibels par $\mathrm{dB}(x) = 10\mathrm{Log}_{10}(x)$. Voici une table avec quelques valeurs intermédiaires.

x	1	2	3	5	10
$\mathrm{Log}_{10}(x)$	0	0.30	0.47	0.70	1
$\mathrm{dB}(x)$	0	3	4.7	7	10

Exercice A.2.9.

1. Une clarinette jouant fort émet un son à 90dB. Quel est le niveau sonore de 2 clarinettes (en dB) ?
2. Un trombone jouant fort est à 110dB. Combien faut-il de clarinettes jouant fort pour avoir la même puissance ?

Solution A.2.10.

1. Le niveau sonore de 2 clarinettes est $90 + \mathrm{dB}(2) = 90 + 3 = 93\mathrm{dB}$
2. On cherche le nombre n de clarinettes tel que

$$90 + \mathrm{dB}(n) = 110 \Leftrightarrow \mathrm{dB}(n) = 110 - 90 = 20$$

donc $n = 100$.

Remarque A.2.11. Quelques formules utiles à savoir sur la fonction $x \to \operatorname{Log}(x)$. Pour $x, y > 0$, on a

$$\operatorname{Log}(xy) = \operatorname{Log}(y) + \operatorname{Log}(y) \tag{A.2.1}$$

$$\operatorname{Log}(x^y) = y\operatorname{Log}(x)$$

A.3 Le cercle, sinus et décomposition de Fourier

Bien que peu présent dans la nature, le cercle est un concept fondamental en science.

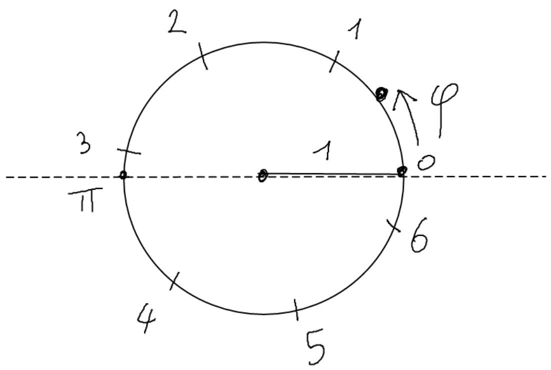
(Dans la nature, on observe rarement des formes quasiment circulaires : silhouette du soleil ou de la pleine lune, arc en ciel, ronds dans l'eau,... ?)

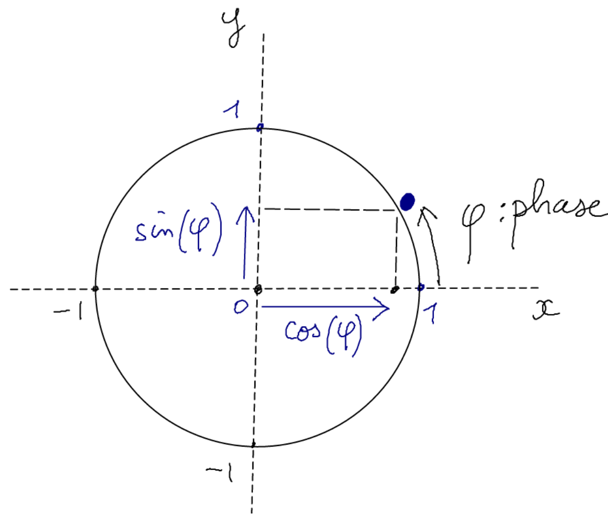
A.3.1 Phase, sinus, cosinus

Définition A.3.1. Sur un cercle, l'angle ou **phase** φ en **radian** est la longueur d'arc mesurée en unité de rayon. En particulier le nombre pi est

$$\pi := \frac{1}{2} \text{ circonférence} \approx 3.14\ldots$$

La projection sur l'axe x s'appelle $\cos\varphi$ (**cosinus** de φ) et la projection sur l'axe y s'appelle $\sin\varphi$ (**sinus** de φ).





Exercice A.3.2. Donner les valeurs de $\cos(0), \sin(0), \cos(\pi/2), \sin(\pi/2), \cos(\pi), \sin(\pi)$.

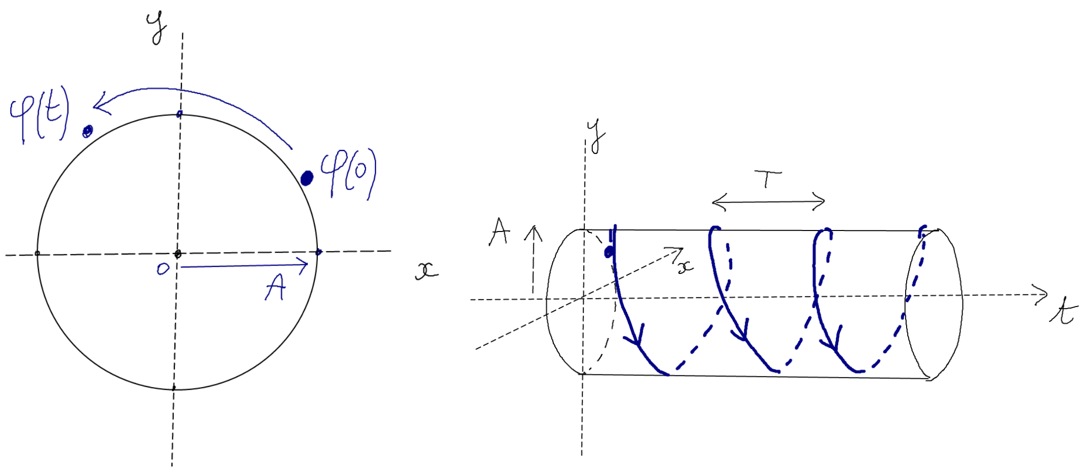
A.3.2 Mouvement circulaire

Définition A.3.3. Un **mouvement circulaire uniforme** est un point qui bouge sur un cercle à vitesse constante. Ce mouvement est caractérisé par :

— sa **période** T qui est le temps en sec. pour faire un tour ou la **fréquence** qui est $f = 1/T$, (nombre de tours par seconde en Hz $= s^{-1}$). La **fréquence est positive** (respect. **négative**) si le mouvement est dans le sens direct, i.e. anti-horaire, (respect. indirect, i.e. sens horaire).

— l'**amplitude** $A > 0$ qui est le rayon du cercle.

— la **phase** φ_0 au départ (à la date $t = 0$).

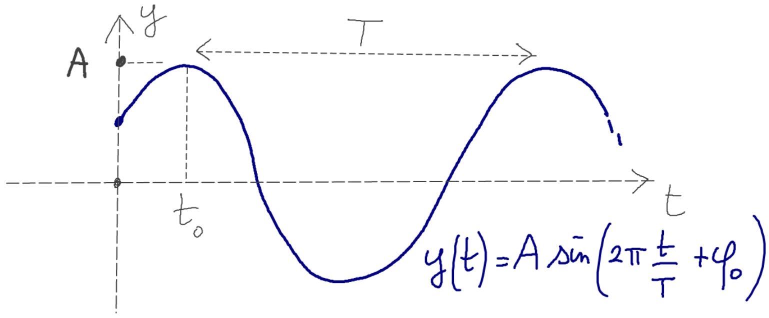


Remarque A.3.4. On note aussi $\omega = 2\pi f$ la vitesse angulaire (en radian par seconde).

Exemple A.3.5. Pour $f = 2$Hz, $T = \frac{1}{f} = 0.5\,s.$, $\omega = 2\pi \times 2$Hz.

Exemple A.3.6. (Le diapason) Pour $f = 440\text{Hz}$, $T = \frac{1}{f} = \frac{1}{440}\,s. = 0.0022\,s. = 2,2\,\text{ms}$. C'est "très rapide".

Proposition A.3.7. *On déduit que la composante du mouvement circulaire ci-dessus sur l'axe y a une* ***forme sinusoidale*** *:*



La formule précise est :

$$y\,(t) = A \sin\left(2\pi \frac{t}{T} + \varphi_0\right) = A \sin\left(2\pi f t + \varphi_0\right)$$

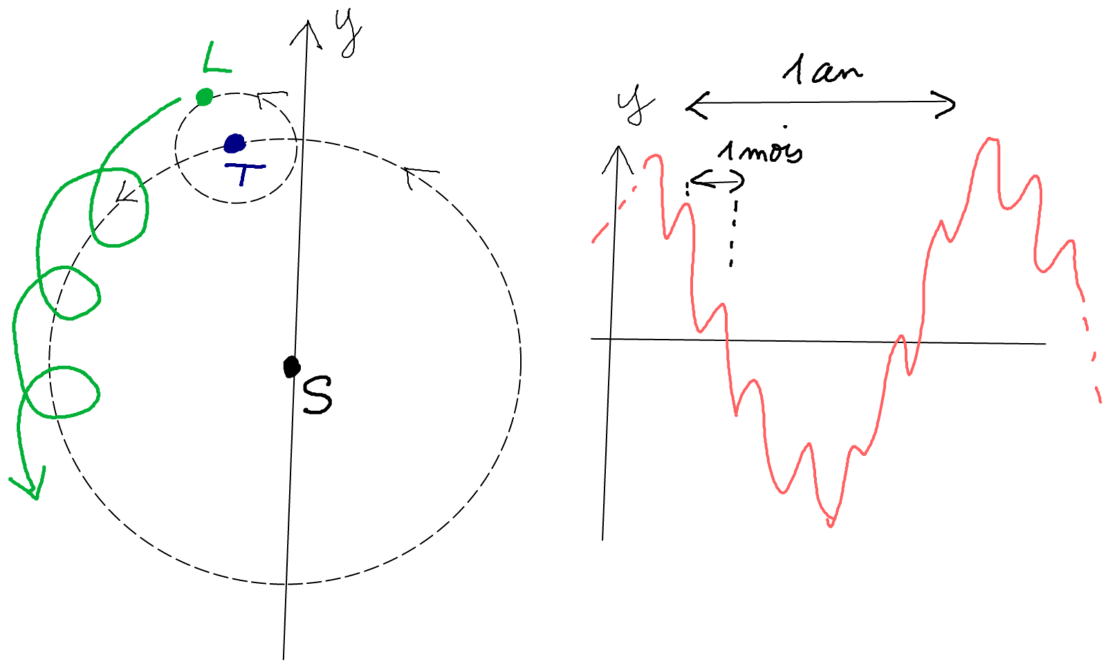
Remarque A.3.8. Pour se convaincre de la forme sinusoidale, observer un **tire-bouchon** de côté.

Remarque A.3.9. La phase φ_0 est reliée à la date t_0 du maximum par $\sin\left(2\pi \frac{t_0}{T} + \varphi_0\right) = \sin \frac{\pi}{2}$ donc $\varphi_0 = -2\pi \frac{t_0}{T} - \frac{\pi}{2}$.

A.3.3 Addition de mouvements circulaires, épicycles et décomposition de Fourier

Exemple A.3.10. la **Terre** a un mouvement circulaire autour du soleil, de période $T_1 = 365.25$ jours et rayon $A = 145\,10^6$km et **la lune** a un mouvement circulaire autour du soleil, de période $T_2 = 27.3$ jours et rayon $A = 0.4\ 10^6$km.

— On déduit que la trajectoire de la lune par rapport au soleil est la somme de ces deux mouvements circulaires et n'est pas périodique (car $\frac{T_1}{T_2} = 13.38\ldots$ n'est pas un entier ni un rationnel). De là vient la difficulté des calendriers basés sur l'orbite terrestre soleil ou sur le mois lunaire (calendrier arabe) (Attention, le dessin n'est pas à l'échelle)



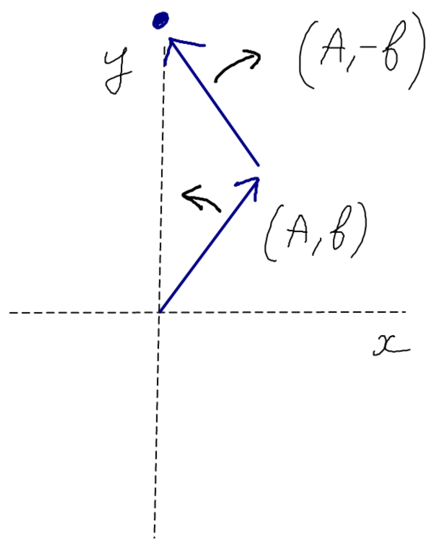
— Voir cette animation de l'orbite lunaire

— De la valeur $\frac{T_1}{T_2} - 1 = 12.38\ldots$ on déduit qu'il y a un peu plus de

12 pleines lunes par an, montrant l'importance du chiffre 12 dans l'histoire ancienne. (l'importance du chiffre 7 vient qu'il y a 7 astres non fixes dans le ciel : lune (lundi), mars (mardi), mercure (mercredi), jupiter (jeudi), venus (vendredi), saturne (samedi), soleil (dimanche, sunday).

Exercice A.3.11. (Difficile) A partir des périodes T_1, T_2 ci-dessus, déduire la période d'un mois synodique (par rapport au soleil) et le nombre de mois lunaires par an.

Exemple A.3.12. La superposition de deux mouvements circulaires de même amplitude A mais fréquences f et $-f$ (égales et opposée), (et phase à l'origine nulle, $\varphi_0 = 0$), est un mouvement selon l'axe y, sinusoïdale de fréquence f et amplitude $2A$.

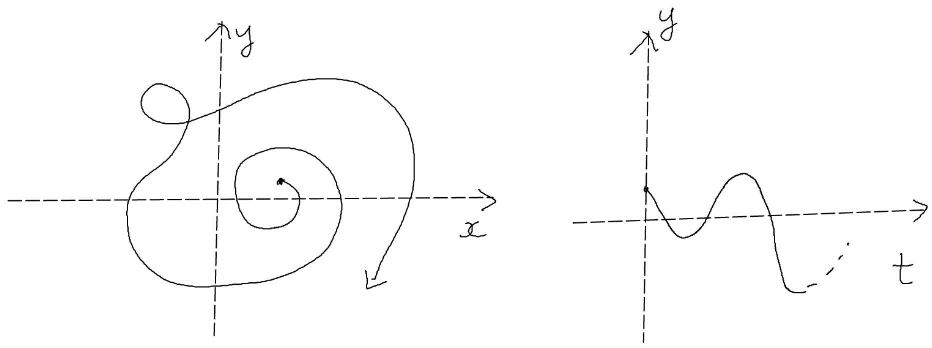


Exercice A.3.13. Faire l' exercice de ce lien.

Dans l'histoire des sciences il y a une longue histoire sur l'étude de l'addition des mouvements circulaires. On peut citer les études des épicyles par les grecs, les Travaux de Al Din Al Touzi au XIXème siècle en perse.

Une étape très importante est le théorème de Fourier obtenu par Joseph Fourier en 1811 à grenoble :

Théorème A.3.14. *(Fourier 1811)* ***Tout mouvement*** $x(t), y(t)$ *dans le plan* ***se décompose de façon unique*** *comme* ***la superposition de mouvements circulaires****, de différentes fréquences* f*, amplitudes* A *et phases à l'origine* φ_0*.*



Voir ces illustrations.

Remarque A.3.15. La **transformée de Fourier** fournit une formule pour obtenir cette décomposition en mouvements circulaires.

A.3.4 Cas particulier d'un mouvement périodique (série de Fourier)

En musique, on s'intéresse particulièrement aux **signaux périodiques**, car donnant un son "harmonieux" à la perception.

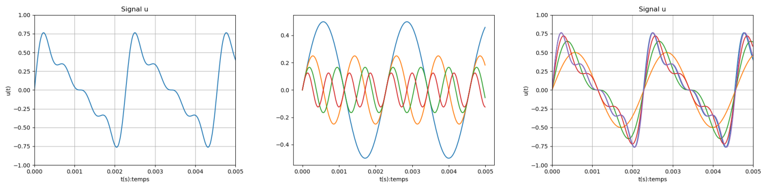
Remarque A.3.16. La justification du terme “harmonique” en musique est que l’on perçoit les signaux périodiques comme “harmonieux” et remonte peut être à l’ouvrage de Kepler Harmonices Mundi “L’Harmonie du monde” 1619 où il relie (sans trop de raison scientifique) musique et orbites planétaires.

Remarquer qu’un mouvement circulaire de période T ou $\frac{1}{2}T$ ou $\frac{1}{3}T$ etc est à fortiori de période T.

Théorème A.3.17. *Inversement d’après le Théorème de Fourier ci-dessus, tout signal périodique de période T se décompose de façon unique comme la superposition de mouvements circulaires de périodes $\frac{1}{n}T$ avec $n = 1, 2, 3, \ldots$ appelé* ***numéro de l’harmonique****. En terme de fréquence, $f = \frac{1}{T}$ est la fréquence fondamentale du signal périodique et* ***la fréquence de l’harmonique*** *$n \geq 1$ est $\frac{n}{T} = nf$, donnant la suite :*

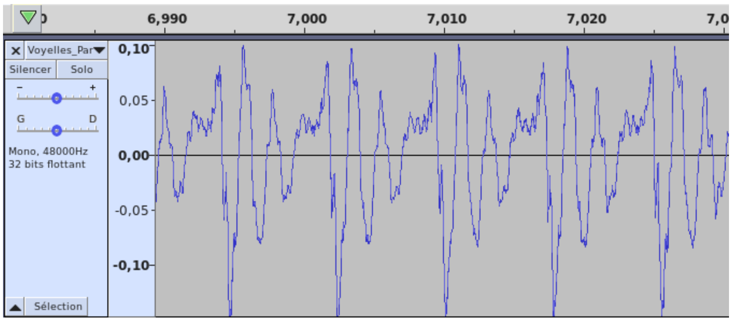
$$f, 2f, 3f, 4f, \ldots$$

Exemple A.3.18. Sur la première figure, voici un signal $y(t)$ obtenu comme somme de quatre fonctions sinus $u_1(t), u_2(t), u_3(t), u_4(t)$ de période respectives $T, \frac{1}{2}T, \frac{1}{3}T, \frac{1}{4}T$ avec $n = 1, 2, 3, 4$ représentées en couleur sur la 2eme figure. La 3eme figure montre les sommes partielles $u_1, u_1 + u_2, u_1 + u_2 + u_3, etc$ pour observer la reconstruction du signal. En rajoutant des harmoniques convenables $n = 5, 6, \ldots$ etc on peut obtenir un signal en “**dent de scie**” à la limite.



Exemple A.3.19. Avec le programme suivant en javascript, vous pouvez changer de façon interactive les amplitudes A_n et phases φ_n et écouter le son résultant. Remarquer que si on change les phases, le signal change fortement, mais à l'écoute on ne perçoit aucun changement.

Exercice A.3.20. Voici un signal Voyelles_Par_Malik.wav qui est le son d'une voyelle "A". On mesure que la période est $T = 0.0625$. Quelles sont les fréquences des harmoniques ?



Solution A.3.21. La fréquence est $f = \frac{1}{T} = \frac{1}{0.0625} = 160\text{Hz}$. Les fréquences des harmoniques sont alors :

$$f = 160Hz, \quad 2f = 320\,\text{Hz}, \quad 3f = 480\text{Hz}, \quad 4f = 640\text{Hz}, \quad \text{etc}$$

A.4 Codage des nombres en base 2 (binaire)

Voir système binaire sur wikipedia.

A.4.1 La base 10

On rappelle qu'en base 10, qui est la base usuelle, on a 10 symboles disponibles : $0, 1, 2, \ldots, 9$ et que les nombres entiers écrits de la façon suivante :

$$0, 1, 2, \ldots 9, 10, 11, 12, \ldots, 19, 20, 21, \ldots$$

Exercice A.4.1. En base 10, combien d'entiers différents peut-on écrire avec $B = 3$ cases $\boxed{}\boxed{}\boxed{}$ et un chiffre dans chaque case ?

Solution A.4.2. On peut écrire les entiers de $0 = 000$ à 999, ce qui fait $1000 = 10^3$ entiers différents. Plus générallement avec B cases, on peut écrire 10^B entiers différents.

A.5 La base 2

De la même façon, en base 2 on a 2 symboles disponibles, par exemple $0, 1$. Les nombres entiers sont écrits de la façon suivante :

$$0, 1, 10, 11, 100, 101, 110, 111, 1000, 1001, \ldots$$

Exercice A.5.1. En base 2, combien d'entiers différents peut-on écrire avec $B = 3$ cases $\boxed{}\boxed{}\boxed{}$ et un chiffre dans chaque case ?

Solution A.5.2. On peut écrire les entiers de $0 = 000$ à 111, ce qui fait $8 = 2^3$ entiers différents. Plus généralement avec B cases, appelé **nombre de bits**, on peut écrire 2^B entiers différents.

Remarque A.5.3. On rappelle qu'un **octet** correspond à 8 bits (i.e. 8 cases).

A.5.1 Le complément à 2

En base 2, avec un certain nombre de cases B fixé (nombre de bits), si on souhaite aussi **écrire des entiers négatifs**, on convient d'utiliser la première case pour le signe : 0 signifie positif et 1 signifie négatif. Par exemple avec $B = 3$ cases : on aura

entier	-4	-3	-2	-1	0	$+1$	$+2$	$+3$
code en base 2	100	101	110	111	000	001	010	011

— Remarquer que pour obtenir la deuxième ligne, on part de 0 et on remplit le tableau de façon croissante et circulaire.

— Plus généralement, avec B cases, on code les entiers de $-\frac{2^B}{2} = -2^{B-1}$ à $\frac{2^B}{2} - 1 = 2^{B-1} - 1$.

— L'avantage de cette convention est que l'opération d'addition est inchangée, par exemple pour calculer $+3 - 2$ on effectue en base 2 : $011 + 110$ ce qui donne 1001, mais en fait 001 car on a que $B = 3$ cases. Le résultat est donc $+1$.

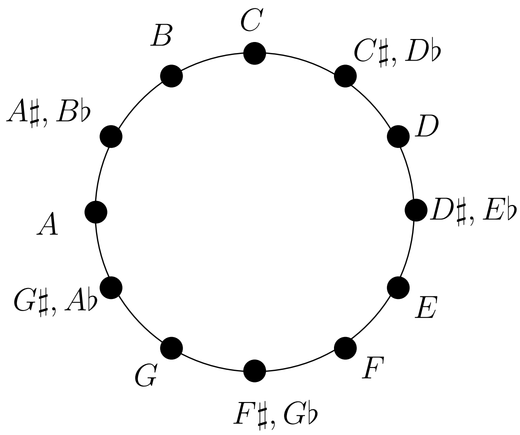
Annexe B

Notations musicales

Comme le tempérament égal est très utilisé, il y a beaucoup de conventions qui lui sont associées et on les rappelle ici.

B.1 Représentation des notes du tempérament égal

Cercle dodécaphonique : Les altérations **dièze** $\sharp$ indique une note au dessus et $\flat$ **bémol** indique une note en dessous. Modulo 12, on peut représenter les notes sur un cercle qui se lit dans le sens horaire.



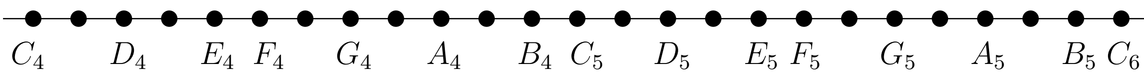
Un tour complet est une **octave**.

Noms francais : En francais on utilise les noms suivants

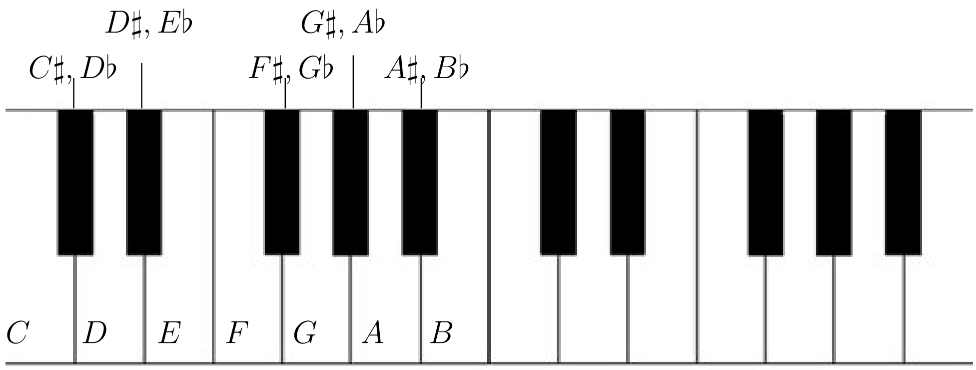
C	D	E	F	G	A	B
Do	Ré	Mi	Fa	Sol	La	Si

Les notes sur la droite $\mathbb{Z}$: Si on déroule le cercle on obtient la droite $\mathbb{Z}$ des pitchs du tempérament égal. On ajoute le numéro de l'octave aux symboles :

...



Les notes sur le clavier du piano :

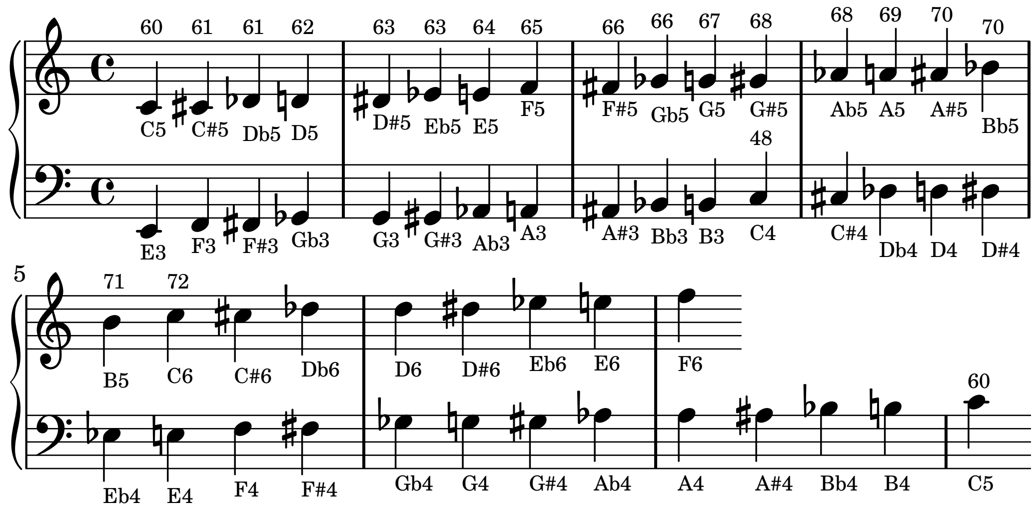


C_5 est la touche « Do » près de la serrure du piano, appelé « do serrure ».

Les notes sur le manche de la guitare :

E_5	F_5	$F\sharp_5, G\flat_5$	G_5	$G\sharp_5, A\flat_5$	A_5	$A\sharp_5, B\flat_5$	B_5	C_6	$C\sharp_6, D\flat_6$	D_6	$D\sharp_6, E\flat_6$	E_6	F_6
B_4	C_5	$C\sharp_5, D\flat_5$	D_5	$D\sharp_5, E\flat_5$	E_5	F_5	$F\sharp_5, G\flat_5$	G_5	$G\sharp_5, A\flat_5$	A_5	$A\sharp_5, B\flat_5$	B_5	C_6
G_4	$G\sharp_4, A\flat_4$	A_4	$A\sharp_4, B\flat_4$	B_4	C_5	$C\sharp_5, D\flat_5$	D_5	$D\sharp_5, E\flat_5$	E_5	F_5	$F\sharp_5, G\flat_5$	G_5	$G\sharp_5, A\flat_5$
D_4	$D\sharp_4, E\flat_4$	E_4	F_4	$F\sharp_4, G\flat_4$	G_4	$G\sharp_4, A\flat_4$	A_4	$A\sharp_4, B\flat_4$	B_4	C_5	$C\sharp_5, D\flat_5$	D_5	$D\sharp_5, E\flat_5$
A_3	$A\sharp_3, B\flat_3$	B_3	C_4	$C\sharp_4, D\flat_4$	D_4	$D\sharp_4, E\flat_4$	E_4	F_4	$F\sharp_4, G\flat_4$	G_4	$G\sharp_4, A\flat_4$	A_4	$A\sharp_4, B\flat_4$
E_3	F_3	$F\sharp_3, G\flat_3$	G_3	$G\sharp_3, A\flat_3$	A_3	$A\sharp_3, B\flat_3$	B_3	C_4	$C\sharp_4, D\flat_4$	D_4	$D\sharp_4, E\flat_4$	E_4	F_4
Fret :	1	2	3	4	5	6	7	8	9	10	11	12	13

Les notes sur la portée musicale leur nom et leur pitch (code MIDI) :



B.2 Les intervalles

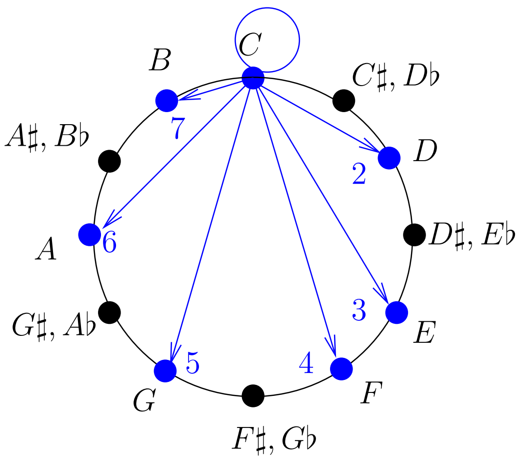
On prend l'exemple d'un intervalle entre C et une autre note supérieure. Les intervalles sont représentés par leur **degré** $1, 2, \ldots 7$ correspondant aux intervalles de la gamme Majeure C, D, E, F, G, A, B.

Représentation des degrés des intervalles sur la droite $\mathbb{Z}$:

Notes :	C		D		E	F		G		A		B	C		D		E	F		G		A
Numero :	1		2		3	4		5		6		7	8		9		10	11		12		13
		2b		3b			5b		5♯		7b			9b		9♯			11♯		13b	

Remarque : dans cet exemple 5# est le $G\#$, c'est la même note (même fréquence) que $6b$ qui est le Ab.

Représentation des degrés de ces intervalles sur le cercle $\mathbb{Z}_{12}$:



Nom des intervalles : Ne pas confondre le degré et le nombre de demi-tons.

Degré	1	2♭	2	3♭
Demi-tons (pitch)	0	1	2	3
Nom	Unisson	Seconde mineure ou 1/2 ton	Seconde ou ton	Tierce mineure

Degré	3	4	5♭	5
Demi-tons (pitch)	4	5	6	7
Nom	Tierce majeure	quarte	quinte diminuée	quinte

Degré	5♯	6	7♭	7
Demi-tons (pitch)	8	9	10	11
Nom	quinte augmentée	Sixte	Septième mineure	Septième Majeure

Degré	8	9♭	9	9♯
Demi-tons (pitch)	12	13	14	15
Nom	Octave	Neuvième mineure	Neuvième Majeure	Neuvième augm

Degré	11	13
Demi-tons (pitch)	16	17
Nom	Onzième	Treizième Majeure

Bibliographie

Benson, DJ. Music : a mathematical offering. *pdf version*.

Castellengo, Michèle. 2015. *Ecoute musicale et acoustique : Avec 420 sons et leurs sonagrammes décryptés.* Eyrolles.

Dauzat, Pierre-Emmanuel, & Harari, Yuval Noah. 2015. *Sapiens : une brève histoire de l'humanité.* Albin Michel.

Dehaene, Stanislas. 2014. *Le Code de la conscience.* Odile Jacob.

Fletcher, Neville H, & Rossing, Thomas D. 2012. *The physics of musical instruments.* Springer Science & Business Media.

Hirschberg, Avraham, Pelorson, Xavier, & Gilbert, JoÃ«l. 1996. Aeroacoustics of Musical Instruments. *Meccanica*, **31**(04), 131–141.

Pham, F. 2003. Gammes naturelles et justesse expressive : l'exemple de la musique indienne. *PrÃ©publ. du laboratoire J.A. DieudonnÃ©*, **50**.

Schnupp, Jan, Nelken, Israel, & King, Andrew. 2011. *Auditory neuroscience : Making sense of sound.* MIT Press, webpage.

Schroeder, Manfred, Rossing, Thomas D, Dunn, F, Hartmann, WM, Campbell, DM, & Fletcher, NH. 2007. Springer handbook of acoustics.