

Chapitre 2

Analyse des signaux sonores

Références :

- Sur analyse des signaux : cours de base : [4 livres de J.O.Smith](#), livre plus avancé : [livre de S. Mallat](#),

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.

2.1 Définitions d'un signal et échantillonnage

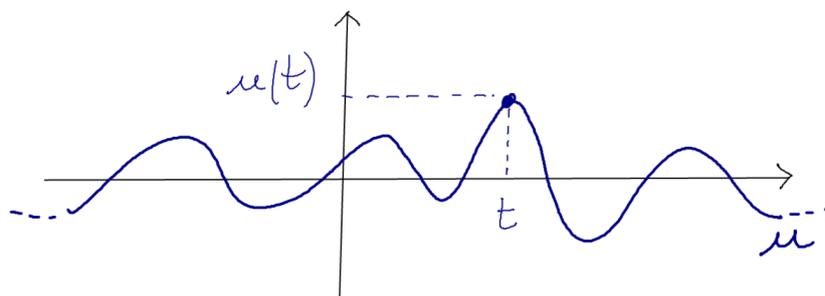
[Video de cette section](#)

2.1.1 Signal sonore

Définition 2.1.1. Un **signal sonore** est une **fonction**

$$u : t \in \mathbb{R} \rightarrow u(t) \in \mathbb{R}$$

où t est la date et $u(t)$ correspond par exemple à la **fluctuation de pression** $u(t) = p(x, t)$ au point d'entrée x fixé d'une oreille ou d'un microphone.



2.1.2 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).

Notation 2.1.2. La notation $\llbracket a, b \rrbracket$ signifie l'ensemble des entiers n dans l'intervalle $a \leq n \leq b$. Par exemple $\llbracket 0, 3 \rrbracket = \{0, 1, 2, 3\}$.

Définition 2.1.3. Un **signal échantillonné** est :

— un nombre fini N de valeurs du signal,

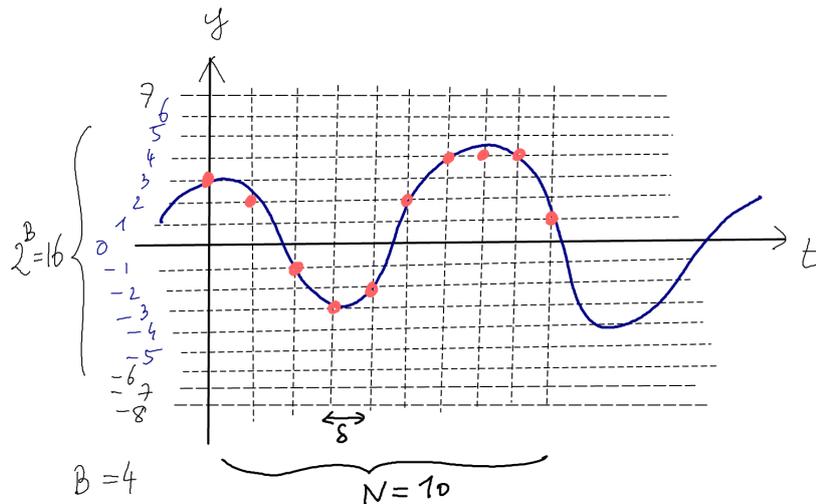
$$u = (u_0, u_1, u_2, \dots, u_{N-1}), \quad u_j \in \llbracket -2^{B-1}, 2^{B-1} - 1 \rrbracket, \quad j \in \llbracket 0, N - 1 \rrbracket,$$

— où chaque valeur est codée avec un nombre B de bits. $B = 8, 16, 24, 32, 64$ de sorte qu'une valeur u_j est un entier dans l'intervalle $\llbracket -2^{B-1}, 2^{B-1} - 1 \rrbracket$. Ce codage appelé **Complément à deux**.

— à des d'instant t_j 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 20000 Hz qui est la limite supérieure des fréquences humainement audibles). Ainsi

$$t_j = j\delta, \quad j \in \llbracket 0, N - 1 \rrbracket,$$

$$u_j \approx u(t_j)$$



Exemple 2.1.4. $N = 10^5$ valeurs, période d'échantillonnage $\delta = \frac{1}{44100}$ s. donnera une durée totale $D = N\delta = \frac{10^5}{44100} = 2.26$ s. La valeur $B = 4$ donne $2^3 = 8$. $B = 8$ donne $2^7 = 128$, $B = 32$ donne $2^{31} \approx 10^9$.

Exercice 2.1.5. Combien y a-t-il d'échantillons N dans un signal de durée $D = 10$ mn qui est échantillonné avec un pas de temps $\delta = \frac{1}{44100}$ s. ?

Solution 2.1.6.

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

Remarque 2.1.7. Pour les représentations du signal, il peut être utile d'effectuer un changement d'échelle pour avoir des valeurs dans l'intervalle fixe $[-1, 1]$. Par exemple avec $B = 8$ bits, on a des échantillons $u_j \in \llbracket -128, +127 \rrbracket$ à valeur entière et si on divise chaque valeur par l'amplitude maximale qui est $128 = 2^7$, on obtient un signal

$$v_j = \frac{1}{2^7} u_j \in [-1, 1],$$

qui varie avec des pas de valeur $\frac{1}{2^7} \approx 0.01$. Dans les exemples ci-dessous, avec les logiciels audacity et python on utilisera ces valeurs v_j redimensionnées.

Exercice 2.1.8. 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 $N = 25 \cdot 10^6$ échantillons, codé en 16 bits occupe combien d'octets en mémoire ? (1 octet = 8 bits)

Solution 2.1.9.

$$u_j \in \llbracket -2^{16-1}, +2^{16-1} - 1 \rrbracket = \llbracket -32768, +32767 \rrbracket$$

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

2.1.2.1 (*) Remarques avancées

Remarque 2.1.10. On peut penser à un signal échantillonné de différentes façons :

1. un **vecteur** à N composantes réelles $u = (u_0, u_1, \dots, u_{N-1})$.
2. une **combinaison linéaire de distributions de Dirac**¹, aux point $t_j = j\delta$:

$$u = \sum_{j=0}^{N-1} u_j \delta_{j\delta}$$

appelé parfois **peigne de Dirac**. (Problème de notations : attention à la confusion possible entre le pas de temps δ et la distribution de Dirac δ)

Remarque 2.1.11. En informatique, les signaux audio échantillonnés (mono ou stéréo) sont stockés en mémoire selon différentes conventions. La plus courante est appelée le **format WAV**, c'est un format où les données (u_0, u_1, \dots) sont écrites à la suite sous formes d'octets. Le format **mp3** est aussi un format très utilisé, mais les données sont écrites après un algorithme de compression.

2.1.3 Mesure de l'intensité en décibels

Pour mesurer l'amplitude d'un signal, on utilise parfois une échelle **logarithmique** qui est plus proche de la perception humaine que l'échelle linéaire. Voici une définition très utilisée appelée les décibels. Il s'agit d'une mesure relative d'intensité. En section 1.3.6 on a déjà présenté cette définition dans un cadre physique, où l'on a considéré l'intensité par rapport à une intensité de référence qui était l'intensité minimale perceptible. Ici, comme on considère un signal numérique $t \rightarrow u(t)$ la valeur de référence sera (arbitrairement) la valeur $u(t) = 1$.

La Section A.2 rappelle la notion de exposants et logarithmes.

1. Par définition, la **distribution de Dirac** δ ou δ_0 appliquée à une fonction u donne sa valeur $u(0)$ en $x = 0$. On écrit $\delta_0(u) = u(0)$. Plus généralement la distribution de Dirac δ_x en x est définie par $\delta_x(u) = u(x)$. Attention, δ_x n'est pas une fonction, mais abusivement on écrit

$$u(x) = \langle \delta_x | u \rangle = \int \delta_x(y) u(y) dy \tag{2.1.1}$$

ou $u(x) = \int \delta(y-x) u(y) dy$. La distribution de Dirac δ_x est aussi appelée **impulsion** en théorie du signal.

Définition 2.1.12. Si $u : t \in \mathbb{R} \rightarrow u(t) \in \mathbb{R}$ est un signal, on définit l'**intensité** en **décibels** (dB) à l'instant t par

$$\text{dB}(t) = 10 \text{Log}_{10}(I(t))$$

où l'**intensité** ou l'énergie d'un signal est donnée par l'amplitude au carré :

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

Remarque 2.1.13. On peut écrire $\text{dB}(I)$ directement en fonction de l'amplitude $|u(t)|$:

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

Le choix du coefficient 10 vient du fait qu'une variation de ± 1 dB correspond à la variation la plus petite décelable par la perception humaine. Par ailleurs $\Delta I = 120\text{dB}$ est à peu près l'intervalle total entre un signal trop faible (non décelable) et trop fort (destructeur de notre appareil auditif).

Exercice 2.1.14. 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 $\text{dB}(I(t))$?

Solution 2.1.15. On a $I' = (u')^2 = (2u)^2 = 4u^2 = 4I$. Donc

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

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 2.1.16. 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).

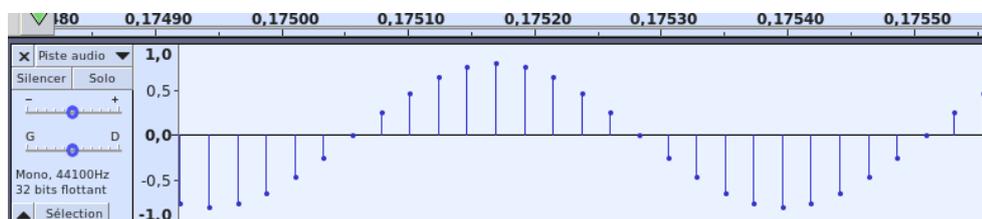
Solution 2.1.17.

- 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 $15 \times 6 = 90\text{dB}$ (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).

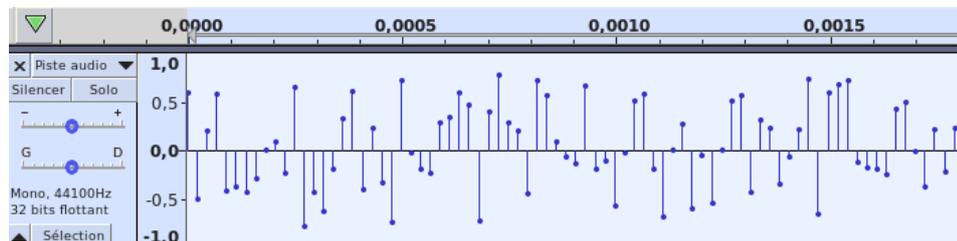
Video des exemples suivants

Exercice 2.1.18. (TP) “**Première utilisation de Audacity**”. Avec le logiciel **audacity**, **A.2**,

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



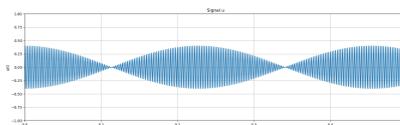
- Créer un signal où chaque valeur $u_j \in [-1, 1]$ est choisi **au hasard** indépendamment des autres. Ce signal est appelé “**bruit blanc**”. Pour cela Menu :Générer/Tonalité/Bruit/Blanc. écouter.



2.1.4 Battements

Exercice 2.1.19. (TP) “**Battements avec Audacity**”.

- 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 sinusoïdaux 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).



2. Vérifier que que l'on entend des **battements** dont la fréquence est

$$f_{\text{batt}} = |f_2 - f_1| = 4 \text{ Hz} \quad (2.1.2)$$

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

3. Donner une explication simple de cette formule ?

Solution 2.1.20.

- 1.
2. Fichiers **battement.wav** ou sur canaux séparés : **battement_stereo.wav**.
3. Explication simple de la formule de battements (2.1.2) (explication plus formelle ci-dessous) : 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 $f_1 = \frac{1}{T_1}$ fait N oscillations. Cela donne $T_{\text{bat.}} = (N + 1)T_2 = NT_1$, donnant $N = \frac{T_2}{T_1 - T_2}$ et

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

Exercice 2.1.21. (TP) “**Battements avec Python**”.

1. Le programme python suivant, **sinus_wav_file.py**, qui génère un signal sinusoïdal de fréquence $f = 440\text{Hz}$ et durée $D = 2\text{sec.}$, défini par :

$$u_j = 0.2 \cos(2\pi f t_j)$$

avec des instants $t_j = j\delta$, $j = 0 \rightarrow N - 1$, avec $N = \lceil \frac{D}{\delta} \rceil$ données et le pas de temps $\delta = 1/44100\text{Hz}$. Ensuite le programme écrit ce signal dans un fichier sinus.wav au format wav. Executer le programme. Ouvrir le fichier sinus.wav avec audacity et écouter.

2. Modifier le programme pour générer le signal suivant

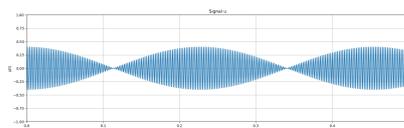
$$u(t) = 0.2 \cos((2\pi) 3ft) + 0.1 \cos((2\pi) 4ft)$$

et l’écouter et visualiser avec audacity. Qu’entendez vous ? le signal est-il périodique ?

3. Modifier le programme pour générer le signal suivant

$$u_1(t) = 0.2 \cos(2\pi ft) + 0.2 \cos((2\pi)(f + \epsilon)t) \quad (2.1.3)$$

avec $\epsilon = 4\text{Hz}$. Expliquer ces phénomènes appelés « **battements** » et calculer leur période, en utilisant une **formule de trigonométrie** que l’on démontrera au préalable à partir de la définition $e^{i\theta} = \cos \theta + i \sin \theta$.



4. Essayer en stéréo : générer les signaux $u_1(t) = 0.2 \cos(2\pi ft)$ et $u_2(t) = 0.2 \cos((2\pi)(f + \epsilon)t)$ sur deux canaux différents. Entendez vous des battements ?

Solution 2.1.22. Programme `battements.py`. Fichiers `battement.wav` et `battement_stereo.wav`.

Explication de la formule de battements (2.1.2) : On a besoin de la formule de trigonométrie

$$\cos a + \cos b = 2 \cos\left(\frac{a+b}{2}\right) \cos\left(\frac{a-b}{2}\right)$$

qui se démontre à partir de la formule de Euler $e^{i\theta} = \cos \theta + i \sin \theta$. On l'applique à (2.1.3) avec $a = 2\pi ft$ et $b = (2\pi)(f + \epsilon)t$, avec $\epsilon = 0.01$, donnant $\frac{a+b}{2} = 2\pi\left(f + \frac{\epsilon}{2}\right)t$, $\frac{b-a}{2} = 2\pi\frac{\epsilon}{2}t$,

$$u_1(t) = 0.4 \underbrace{\cos\left(2\pi\left(f + \frac{\epsilon}{2}\right)t\right)}_{\text{signal}} \underbrace{\cos\left(2\pi\left(\frac{\epsilon}{2}\right)t\right)}_{\text{enveloppe}}$$

La fréquence du signal est donc $\left(f + \frac{\epsilon}{2}\right) = 442\text{Hz} \approx f = 440\text{Hz}$ et la fréquence de l'enveloppe est $\frac{\epsilon}{2} \approx 2.2\text{Hz}$. La fréquence des battements (qui correspondent aux maxima d'amplitudes) est le double, soit $f_{\text{batt}} = \epsilon = 10^{-2}440 = 4.40\text{Hz}$. On entend en effet 4.4 battements par seconde.

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 $t \rightarrow \cos(2\pi ft)$ n'est pas réaliste et pas pratique non plus. Nous allons voir qu'il est beaucoup plus intéressant de considérer des fonctions appelées **paquet d'ondes Gaussien** (ou **ondelette** en théorie du signal) et que l'on appelle **note de musique** d'un point de vue musical, définie ainsi.

Définition 2.2.1. Un **paquet d'onde Gaussien** à l'instant $t_0 \in \mathbb{R}$ et de fréquence $f_0 \in \mathbb{R}$ de durée $\sigma > 0$ est

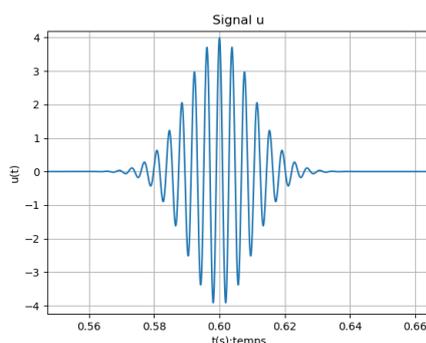
$$\varphi_{t_0, f_0, \sigma}(t) = e^{i2\pi f_0 t} e^{-\frac{1}{2}\left(\frac{t-t_0}{\sigma}\right)^2}. \quad (2.2.1)$$

2.2. SONOGRAMME, TRANSFORMÉE PAR ONDELETTE, TRANSFORMÉE DE FOURIER 67

Remarque 2.2.2. Se rappeler que $\operatorname{Re}(e^{i\theta}) = \cos \theta$ et $\operatorname{Im}(e^{i\theta}) = \sin \theta$ donc la partie réelle du signal (2.2.1) est

$$\operatorname{Re}(\varphi_{t_0, f_0}(t)) = \cos(2\pi f_0 t) e^{-\frac{1}{2}\left(\frac{t-t_0}{\sigma}\right)^2}$$

représenté ci-dessous dans le cas $t_0 = 0.6$, $f_0 = 260$ et $\sigma = 0.01$. La partie imaginaire est semblable.



Le premier facteur $e^{i2\pi f_0 t}$ de (2.2.1) s'appelle la **partie oscillante** et le deuxième facteur $e^{-\frac{1}{2}\left(\frac{t-t_0}{\sigma}\right)^2}$ l'**enveloppe** ou la **fenêtre Gaussienne**. Il peut y avoir d'autres choix intéressants de fenêtres ou **window** en mathématique et théorie du signal. Lire la remarque B.4.1 concernant l'intérêt d'utiliser les modes de Fourier ou ondes planes.

Exercice 2.2.3. (TP) “Une mélodie”. A partir de ce programme python `wave_packet.py`, créer un signal formé de la superposition de paquets d'ondes Gaussiens (ou notes de musique) de durée $\sigma = 0.02$ s reproduisant la mélodie suivante. Visualiser le signal et écouter.



Solution 2.2.4. d'après le tempo, la durée de la noire est $d = \frac{60}{120} = 0.5$ s. Donc si la date de départ est $t_0 > 0$, les dates successives et pitch sont :

Dates après t_0	0	$\frac{1}{2}d$	$\frac{3}{4}d$	d	$d + \frac{1}{2}d$	$2d$	$2d + \frac{1}{2}d$	$2d + \frac{1}{2}d$	$3d$	$3d$
note	C_6	B_5	G_5	F_5	E_5	C_5	F_5	B_5	C_6	E_5
Pitch x	72	71	67	65	64	60	65	71	72	64

Voici le programme `wave_packet_melodie.py`, fichier `wave_packet_melody.wav`.

2.2.1.1 Produit scalaire entre deux paquets d'ondes, principe d'incertitude

Remarque 2.2.5. Avant de lire la suite, vérifier que vous connaissez les **bases d'algèbre linéaire** nécessaires, de la section B.2.

Proposition 2.2.6. *L'angle α entre deux paquets d'ondes Gaussiens $\varphi_{t_1, f_1, \sigma}$ et $\varphi_{t_2, f_2, \sigma}$, définis en (2.2.1), et de même durée $\sigma > 0$ est donné par*

$$\cos \alpha = \frac{|\langle \varphi_{t_1, f_1, \sigma} | \varphi_{t_2, f_2, \sigma} \rangle|}{\|\varphi_{t_1, f_1, \sigma}\| \|\varphi_{t_2, f_2, \sigma}\|} = \exp\left(-\frac{1}{2}d^2\right) \quad (2.2.2)$$

avec la **distance effective** d entre les points (t_1, f_1) et (t_2, f_2) dans le plan temps-fréquence donnée par

$$d^2 = \left(\frac{t_2 - t_1}{\Delta t}\right)^2 + \left(\frac{f_2 - f_1}{\Delta f}\right)^2$$

avec $\Delta t = \sigma$ appelé **incertitude en temps** et $\Delta f = \frac{1}{\sigma}$ appelé **incertitude en fréquence**. Remarquer que d est une distance Euclidienne dans le système d'unité $\Delta t, \Delta f$.

Démonstration. Voir Proposition B.6.2 ou Faure (2010). □

Principe d'incertitude : (2.2.2) signifie d'après (B.2.2) que l'angle α entre les paquets d'ondes φ_{t_1, f_1} et φ_{t_2, f_2} est $\alpha \approx \frac{\pi}{2}$, i.e. les signaux φ_{t_1, f_1} et φ_{t_2, f_2} sont très différents, si et seulement si $d \gg 1$ i.e. si et seulement si

$$|t_2 - t_1| \gg \Delta t \text{ ou } |f_2 - f_1| \gg \Delta f.$$

Remarquer que

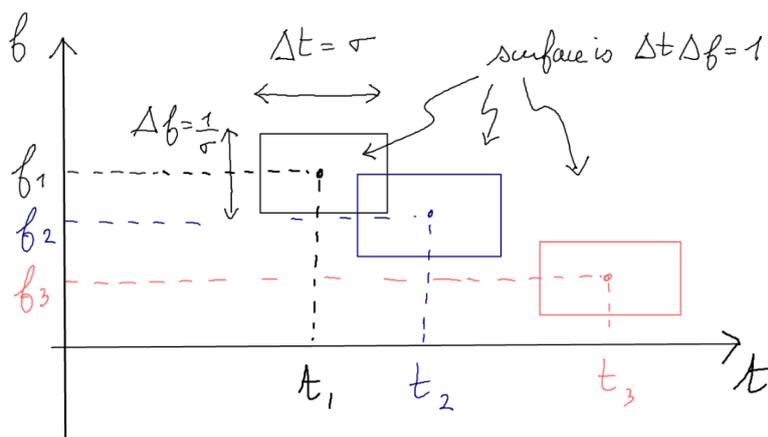
Proposition 2.2.7. *La précision temporelle Δt et précision fréquentielle Δf d'un paquet d'onde sont est reliées par*

$$\Delta t \Delta f = 1 \quad (2.2.3)$$

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 (si on diminue $\sigma = \Delta t$ alors $\Delta f = \frac{1}{\sigma}$ augmente et inversement).

Par exemple, dans le schéma suivant, les signaux $\varphi_{t_1, f_1}, \varphi_{t_2, f_2}$ sont proches car les points (t_1, f_1) et (t_2, f_2) sont à une distance inférieure à 1, alors que φ_{t_3, f_3} est différent (orthogonal) aux autres (distance supérieure à 1).



Remarque 2.2.8. En utilisant une transformée d'un signal en sonogramme et en ne mémorisant que les zones où le signal est suffisamment 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 2.2.9. Application musicale : supposons une contrebasse qui joue des notes pizzicato autour de la note C_3 dont la fréquence est $f_{C_3} = 65\text{Hz}$ et la suivante $f_{C\#_3} = 65 \times 1.06 = 69\text{Hz}$ (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 2.2.10. L'écart de fréquence entre ces deux notes voisines est $\Delta f = 69 - 65 = 4\text{Hz}$. Pour avoir cette précision, d'après le principe d'incertitude la durée minimale d'une note Δt vérifie $\Delta f \Delta t = 1$. Donc

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

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

2.2.2 Sonogramme, transformée de Fourier fenêtrée ou transformée par ondelette

Remarque 2.2.11. Avant de lire la suite, vérifiez que vous connaissez les **bases d'algèbre linéaire** nécessaires, de la section **B.2**.

On utilise les paquets d'ondes (ou notes musicales élémentaires) (2.2.1) et le produit scalaire pour détecter la présence de chaque note $\varphi_{t,f}$ dans un signal u donné.

Définition 2.2.12. Si $u : t \rightarrow u(t)$ est un signal, sa **transformée par paquets d’ondes** (ou **sonogramme**) est la fonction complexe $(Tu)(t, f)$ sur le plan temps-fréquence suivante

$$(Tu)(t, f) := \frac{1}{\|\varphi_{t,f}\|} \langle \varphi_{t,f} | u \rangle \in \mathbb{C} \quad (2.2.4)$$

où

$$\langle \varphi_{t,f} | u \rangle = \int \overline{\varphi_{t,f}(t')} u(t') dt'.$$

Ainsi l’amplitude $|(Tu)(t, f)|$ “mesure la présence” du paquet d’onde élémentaire (ou note) $\varphi_{t,f}$ dans le signal u .

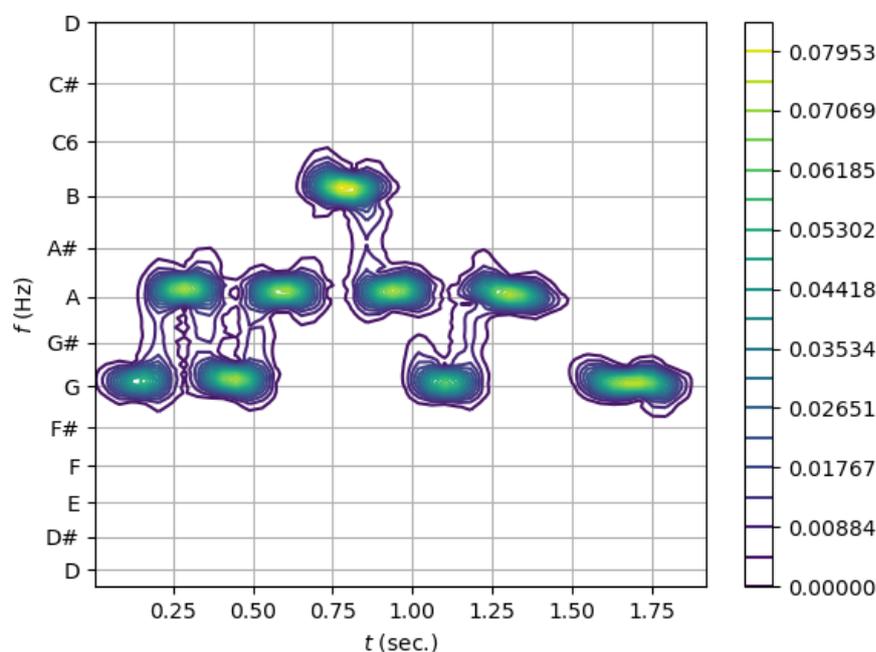
Remarque 2.2.13. T est un opérateur linéaire qui transforme la fonction $u(t)$ à une variable en la fonction $(Tu)(t, f)$ à deux variables. D’après (B.2.5), $(Tu)(t, f)$ s’interprète comme la composante de u sur l’axe du vecteur $\varphi_{t,f}$.

Remarque 2.2.14. Derrière la transformée par paquets d’ondes il y a la transformée de Fourier que l’on détaillera après. La transformée par paquet d’onde est utilisée dans beaucoup de domaines scientifiques sous différents noms : depuis les ondes quantiques, sous le nom de **transformée de Bargmann** Berezin (1950), la géométrie algébrique en mathématiques sous le nom de **plongement de Kodaira**, et plus récemment en théorie du signal avec la **transformée par ondelettes**. Pour améliorer l’analyse, on peut faire varier le paramètres $\sigma(t, f)$ avec le point (t, f) .

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 en (2.2.4) avec un choix de “fonction fenêtre”, **window**, à préciser.

Exercice 2.2.15. (TP) “**Sonogramme d’un signal**”.

1. Ecouter ce signal sonore **flute.wav**. Utiliser ce programme python **sonogram.py** pour effectuer la transformée par paquets d’ondes (ou sonogramme) de ce signal. Choisir une durée $\sigma = 0.04\text{sec}$. Comparer au résultat de l’exercice 2.6.4.



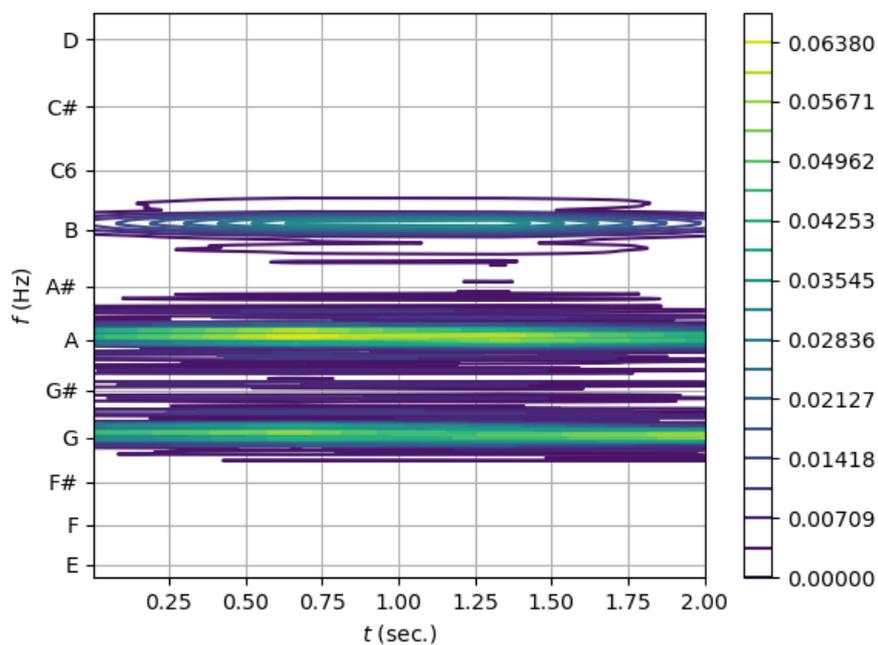
2. Sachant que d'après la proposition 2.2.6, $\Delta t = \sigma$ est l'incertitude en temps et $\Delta f = \frac{1}{\sigma}$ l'incertitude en fréquence du paquet d'onde, quel intervalle acceptable a-t-on pour σ pour bien distinguer les notes ? Changer le paramètre σ dans le programme et observer les résultats si σ est trop grand ou trop petit ?
3. Même exercice avec un fichier de votre création (enregistrement) ou ce fichier [trumpet.wav](#).
4. Même exercice avec la fonction (2.3.3) périodique qui contient 20 harmoniques.

Solution 2.2.16. La durée entre deux notes est ici $\Delta t \approx 0.2s$, il faut donc prendre $\sigma \leq 0.2s$. Par ailleurs on souhaite aussi une précision inférieure à 1/4 ton sur la résolution des fréquence autour de $f = 440Hz$, soit $dx < 0.5$ donc comme

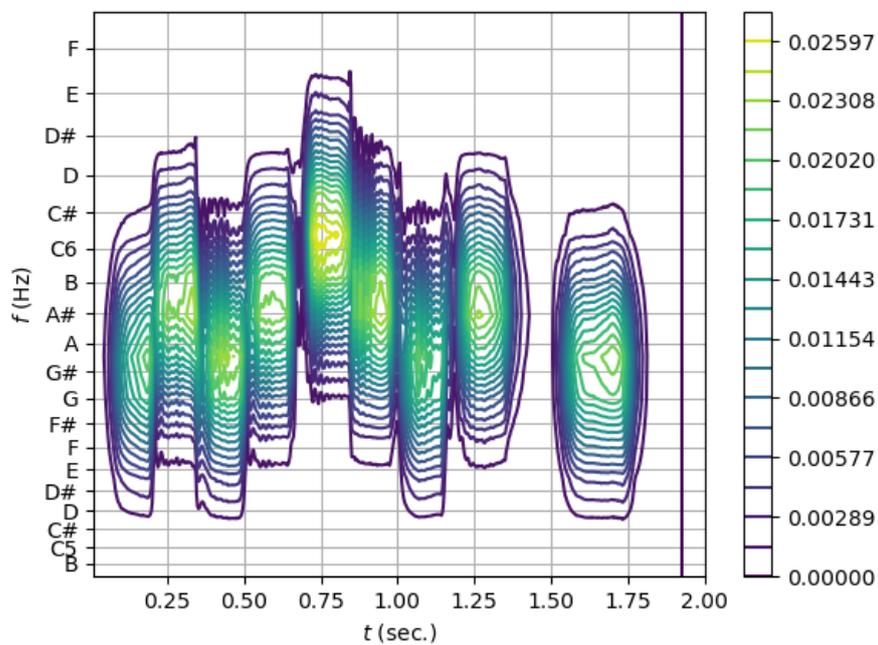
$$dx = \frac{12}{\ln 2} \frac{df}{f} \Leftrightarrow \frac{1}{\sigma} = df = \frac{\ln 2}{12} f dx = 0.06 f dx$$

$$\Leftrightarrow \sigma = \frac{1}{0.06 f dx} > \frac{2}{0.06 \cdot 440} = 0.08$$

, soit $\sigma \geq 0.08s$. L'intervalle suggéré est donc faible. Par exemple, avec $\sigma = 0.5$ on ne distingue plus la résolution temporelle des notes :



Par exemple, avec $\sigma = 0.003$ on ne distingue plus la résolution fréquentielle des notes :



Liens externes :

— Voir de nombreux exemples de sonogrammes dans le livre [Castellengo \(2015\)](#).

- [spectrogramme en JS](#)
- [Video vulgarisée sur le sonogramme](#)

2.2.3 Transformée de Fourier d'un signal

Dans la limite de durée $\sigma \rightarrow \infty$, le paquet d'onde $\varphi_{t,f}$ en (2.2.1) devient tout simplement la fonction suivante appelée **mode de Fourier** ou **onde plane** et ne dépend plus des paramètres t_0, σ mais que de la fréquence f :

$$\varphi_f(t) = e^{i2\pi ft}. \quad (2.2.5)$$

Remarque 2.2.17. La propriété remarquable de cette fonction (en fait la propriété ou définition de la fonction exponentielle elle-même) est qu'elle est vecteur propre de l'opérateur dérivée

$$\frac{d}{dt}\varphi_f = (i2\pi f)\varphi_f$$

Pour cette raison il est souhaitable de décomposer un signal ou une fonction dans la base des modes de Fourier. Mais on perd l'avantage de la localisation temporelle du paquet d'onde (2.2.1). Lire aussi la remarque B.4.1 concernant l'intérêt d'utiliser les modes de Fourier ou ondes planes.

Définition 2.2.18. (Joseph Fourier 1822 à Grenoble) La **transformée de Fourier** d'un signal $u : t \rightarrow u(t)$ est la fonction

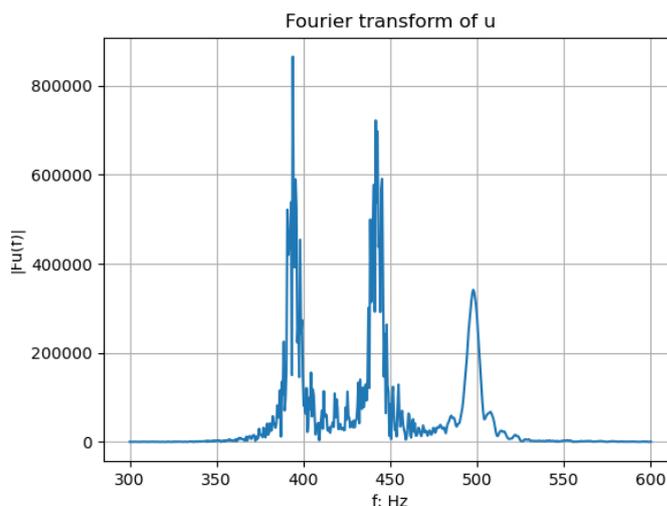
$$f \in \mathbb{R} \rightarrow (\mathcal{F}u)(f) := \langle \varphi_f | u \rangle = \int_{\mathbb{R}} e^{-i2\pi ft} u(t) dt \quad (2.2.6)$$

Remarque 2.2.19. $(\mathcal{F}u)(f) = \langle \varphi_f | u \rangle$, ce produit scalaire mesure la présence du mode de Fourier φ_f dans le signal u .

Lire les sections

- Section B.4.1 qui concerne la **transformée de Fourier** d'une fonction et la **transformée de Fourier inverse**. Les sections suivantes sont des cas particuliers utiles.
- Section B.4.2 concerne la **transformée de Fourier d'un signal échantillonné**,
- Section B.4.3 concerne la **série de Fourier d'un signal périodique**. On a déjà vu en (2.3.1) la décomposition de Fourier d'un signal périodique réel.
- la section B.4.4 concerne la **transformée de Fourier discrète**.

Exercice 2.2.20. (TP) “**Transformée de Fourier discrète d'un signal**”. Voici un programme en python [DFT_of_wavefile.py](#) qui effectue la transformée de Fourier discrète du signal sonore [flute.wav](#) en appliquant directement la formule (B.4.9). On donne aussi le code qui utilise l'algorithme de la transformée de Fourier rapide (le résultat est le même, mais le calcul est plus rapide).



1. Executer ce programme et comparer le résultat avec le sonogramme de l'exercice [2.2.15](#). Commentez l'intérêt du sonogramme ?

Solution 2.2.21. Dans cette figure on distingue trois pic de fréquences, correspondant aux notes G,A,B jouées, mais on ne lit pas par exemple que *G* a été joué quatre fois, ni dans quel ordre. Avec le sonogramme on lit l'information temporelle et l'information fréquentielle simultanément. Avec la transformée de Fourier on ne distingue pas l'information temporelle.

2. Enregistrer sa voix prononçant la voyelle "A" ou un instrument jouant une note tenue et observer la transformée de Fourier (discrète). Observer les harmoniques et les enveloppes du timbre, appelé formants.

2.3 Signaux périodiques, fréquences, notes musicales et pitch

2.3.1 Signaux périodiques, séries de Fourier

[video de cette section](#)

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

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).

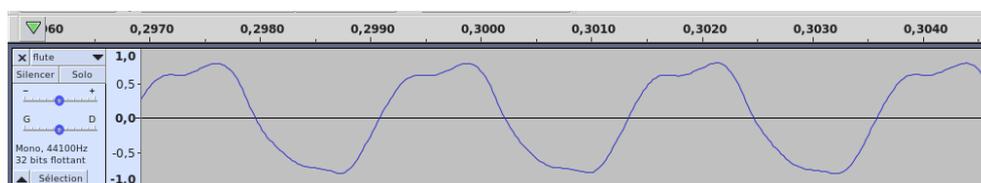
Définition 2.3.1. Un signal périodique de période $T > 0$ est une fonction $t \in \mathbb{R} \rightarrow u(t) \in \mathbb{R}$ qui vérifie

$$u(t + T) = u(t), \forall t \in \mathbb{R}.$$

T est la **période fondamentale**, si c'est la valeur minimale vérifiant cela. On appelle $f = 1/T$ la **fréquence fondamentale** du signal périodique. Une **note musicale** est un signal périodique (ou presque en pratique).

Remarque 2.3.2. Si T est la période fondamentale alors $T, 2T, 3T, \dots$ sont aussi des périodes car si le signal se répète une fois, alors forcément il se répète ensuite.

Exemple 2.3.3. Avec Audacity, voici le signal d'une note de flûte de période approximative $T = 0.302 - 0.300 = 0.002\text{sec}$. donc de fréquence $f = 1/0.002 = 500\text{Hz}$:

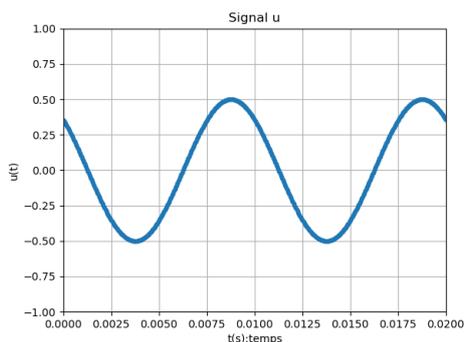
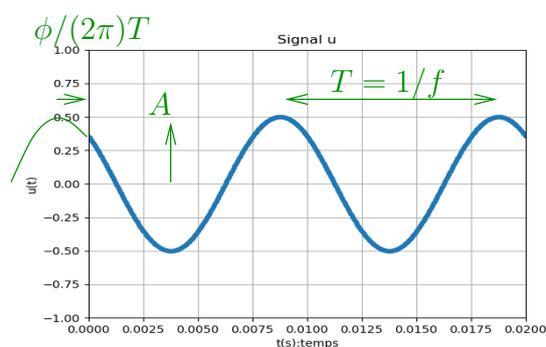


Remarque 2.3.4. On appelle "**wave-form**" ou "forme d'onde" le signal périodique dans l'intervalle d'une période fondamentale : $t \in [0, T] \rightarrow u(t)$. Cette forme peut être quelconque et détermine le signal périodique. La fonction sinus est un exemple particulier de fonction périodique.

Exercice 2.3.5. "Rappels sur la fonction sinus". Tracer l'allure de la fonction suivante

$$u(t) = A \cos(2\pi ft + \varphi)$$

en montrant le rôle des paramètres amplitude A , fréquence f et phase φ . Inversement donner la formule $u(t)$ pour le graphe suivant.

**Solution 2.3.6.**

$$u(t) = A \cos(2\pi ft + \varphi)$$

$$A = 0.5, \quad f = 100, \quad \varphi = \frac{\pi}{4}.$$

Exemple 2.3.7. 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 2.3.8. 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 2.3.9. (*) Si on somme plusieurs fonctions sinus chacune de la forme $A \cos(2\pi ft + \varphi)$, appelé **mode de Fourier**, montrer que la condition nécessaire et suffisante sur ces fonctions pour que le signal résultant soit périodique de période fondamentale T_1 donnée est que

$$f = nf_1$$

avec $n \geq 1$ entier et $f_1 = 1/T_1$ **fréquence fondamentale**.

Solution 2.3.10. Si $v(t) = v(t+T)$ signal périodique donné et on rajoute $u(t) = A \cos(2\pi ft + \varphi)$, cad $w = u + v$, alors à quelle condition w est encore T périodique? Il faut

$$\begin{aligned} w(t+T) = w(t) \forall t, &\Leftrightarrow \cos(2\pi f(t+T) + \varphi) = \cos(2\pi ft + \varphi), \forall t \\ &\Leftrightarrow 2\pi f(t+T) + \varphi = 2\pi ft + \varphi \pmod{2\pi}, \forall t \\ &\Leftrightarrow 2\pi fT = 2\pi n, n \in \mathbb{Z} \Leftrightarrow f = \frac{n}{T_1}, n \in \mathbb{Z} \end{aligned}$$

L'exercice précédent suggère le résultat suivant :

Théorème 2.3.11. “Décomposition de Fourier”.

Un signal périodique réel de période $T = 1/f$ donnée s'écrit de façon unique

$$u(t) = \sum_{n \geq 0} A_n \cos(2\pi(nf)t + \varphi_n) \quad (2.3.1)$$

appelée **décomposition de Fourier** ou **série de Fourier** du signal périodique, avec les amplitudes $A_n \geq 0$, des phases $\varphi_n \in [0, 2\pi]$ qui peuvent être arbitraires mais des fréquences $f_n = nf$ appelées les **harmoniques** du signal.

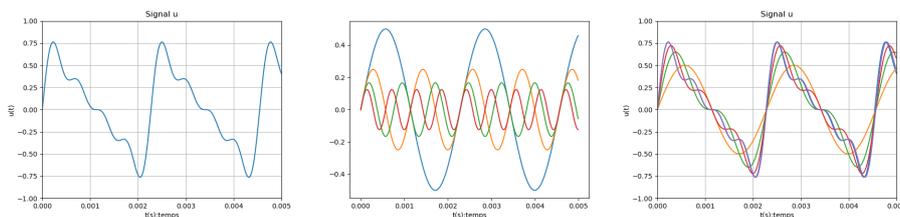
Inversement A_n, φ_n s'obtiennent à partir du signal u par

$$A_n e^{i\varphi_n} = \frac{1}{T} \int_0^T e^{-i2\pi nt/T} u(t) dt, \quad n \in \mathbb{N}. \quad (2.3.2)$$

En particulier $\varphi_0 = 0$ et $A_0 = \frac{1}{T} \int_0^T u(t) dt$ est la moyenne du signal.

Ce théorème est le Théorème B.4.10 en appendice où nous donnons sa preuve.

Exemple 2.3.12. Avec ce programme python `sinus_sum.py` voici un signal $u(t)$ obtenu comme somme de fonctions de Fourier $u_n(t) = \frac{1}{n} \cos(2\pi nft - \frac{\pi}{2}) = \frac{1}{n} \sin(2\pi nft)$ pour $n = 1, 2, 3, 4$ représentées en couleur sur la 2ème figure. La 3ème figure montre les sommes partielles $u_1, u_1 + u_2, u_1 + u_2 + u_3, etc$ pour observer la reconstruction du signal. A la limite, $\sum_{n=1}^{\infty} u_n$ donnerait un signal en “dent de scie”, voir exercice ci-dessous.



Exemple 2.3.13. 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.

Remarque 2.3.14. La justification du terme “harmonique” 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.

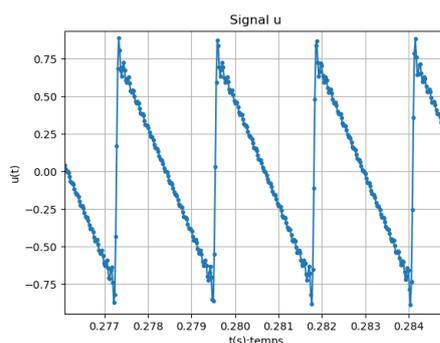
Exercice 2.3.15. (TP) “Série de Fourier”. Avec un programme python

1. Générer le fichier audio.wav du signal suivant sur une durée $D = 2\text{sec}$.

$$u(t) = 0.5 \sum_{n=1}^M \frac{1}{n} \sin(2\pi n f t) \quad (2.3.3)$$

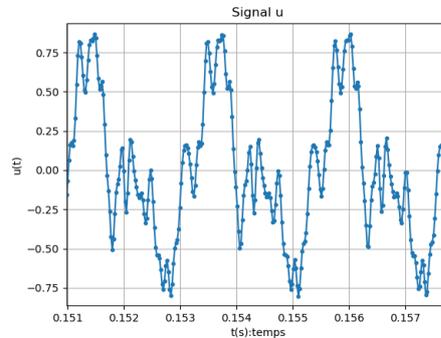
avec $f = 440\text{Hz}$ et $M = 20$ harmoniques. Le signal est-il périodique? L’écouter avec Audacity. Essayer avec d’autres valeurs de M . Que pensez vous du signal si $M \rightarrow \infty$. Le signal sera t-il continu?

Solution 2.3.16. Programme [sinus_sum_wav_file.py](#)



Le signal est périodique de période $T = 1/f = 0.0022\text{sec}$. Si $M \rightarrow \infty$ le signal deviendra discontinu et en forme de “[dent de scie](#)”.

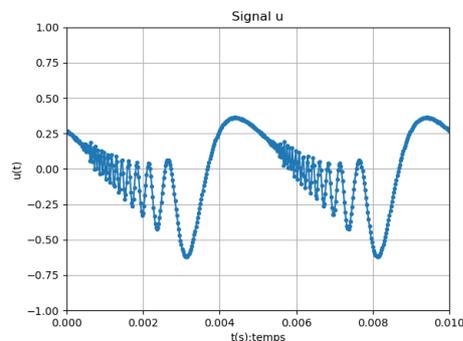
2. Au signal précédent rajouter des phases $\varphi_n \in [0, 2\pi]$ aléatoires (avec loi uniforme). Observer le signal. Ecouter le signal. (On peut montrer que dans la limite $M \rightarrow \infty$, le graphe est continu, c’est un [mouvement Brownien](#) en dim. 1). Essayer les phases $\varphi_n = n \log n$ (série de Hardy-Littlewood, [Katznelson \(2004, p.33\)](#))

Solution 2.3.17. `sinus_sum2.py`

On entend aucune différence avec le signal où $\varphi_n = 0$. Les phases sont inaudibles.

On peut montrer que dans la limite $M \rightarrow \infty$, le signal est (presque sûrement) continu et même dans l'espace $C^{\frac{1}{2}-\epsilon}$, $\forall \epsilon > 0$, cad Hölder d'exposant $\frac{1}{2} - \epsilon$, comme le mouvement Brownien. Mais le signal ne peut pas être $C^{\frac{1}{2}+\epsilon}$, $\forall \epsilon > 0$, [Katznelson \(2004, p.33\)](#).

Série de Hardy-Littlewood avec les phases $\varphi_n = n \log n$ qui est :

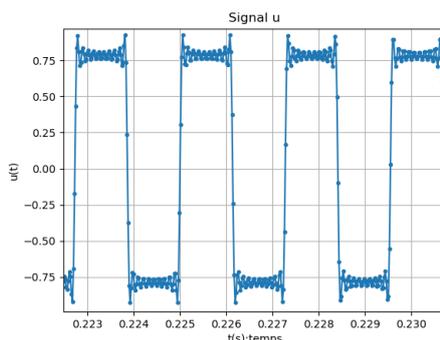


3. Essayer avec le signal

$$u(t) = 0.5 \sum_{n=1}^N \frac{1}{n} (1 - (-1)^n) \sin(2\pi n f t) \quad (2.3.4)$$

Le signal est-il périodique ? que pensez vous du signal si $M \rightarrow \infty$? L'écouter avec Audacity. Ecrire $u(t)$ sous la forme (2.3.1) avec des amplitudes A_n et phases φ_n .

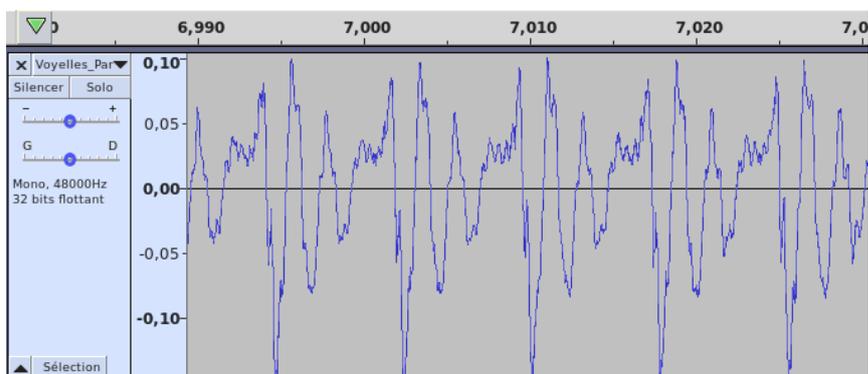
Solution 2.3.18. On obtient



Si $M \rightarrow \infty$ le signal deviendra discontinu et en forme de **signal carré**.

Exercice 2.3.19. (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 . Si besoin, voici un fichier déjà enregistré **Voyelles_Par_Malik.wav**.



2. Sélectionner une période (wave-form) et la répéter avec audacity en faisant Ctrl-T (Edition/suppression 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 2.3.20.

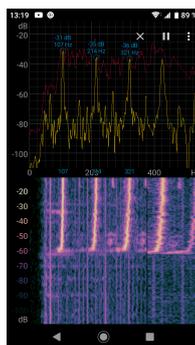
1. Partant de la date $t = 7,000$, on mesure $2T \approx 0.016s$ soit $f = \frac{1}{T} = 125Hz$.

2. On reconnaît le son, mais il paraît très artificiel. Il manque les variations.
3. Si il y a des consonnes, le son paraît étrange. Avec seulement des voyelles, on ne devine pas qu'il a été inversé.
4. Avec le cas parlé (les cordes vocales vibrent), le signal est localement périodique. Dans le cas chuchoté, le signal n'est plus localement périodique.

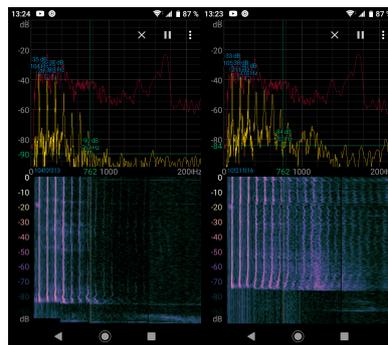
2.3.1.1 Sonogramme d'un signal périodique

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 .

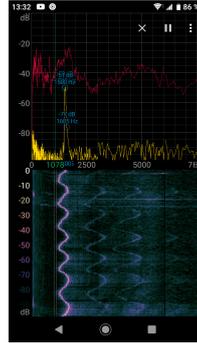
Exemple 2.3.21. Voici le sonogramme d'une voyelle "O" chanté. On observe la fréquence $f = 107\text{Hz}$ et les harmoniques $2f = 214\text{Hz}$, $3f = 321\text{Hz}$ etc



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

[video de cette section](#)

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 $\text{dB} = 10 \text{Log}_{10} \left(\frac{I}{I_0} \right)$ appelée décibels).

On conseille de lire au préalable la section A 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 2.3.22. 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 :

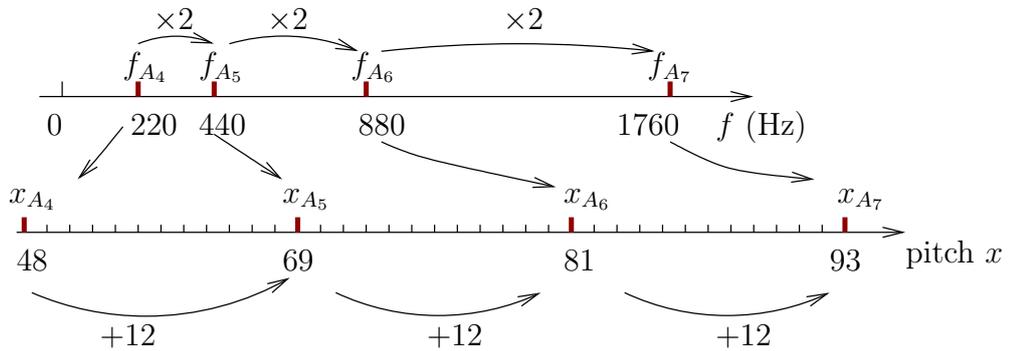
2.3.2.2 Définition du pitch

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 2.3.23. 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 x est une mesure des fréquences f en **demi-tons**.



Remarque 2.3.24. (*) A partir d'une fréquence f donnée, voici une formule qui traduit la définition ci-dessus du pitch

$$x := \left(\frac{12}{\text{Log}_{10}(2)} \right) \text{Log}_{10} \left(\frac{f}{f_{A_5}} \right) + x_{A_5} \in \mathbb{R} \quad (2.3.5)$$

inversement la fréquence f à partir du pitch :

$$f = 10^{\left(\frac{\text{Log}_{10}(2)}{12} \right) (x - x_{A_5})} f_{A_5} \quad (2.3.6)$$

ou plus simplement :

$$f = 2^{\frac{1}{12}(x - x_{A_5})} f_{A_5} \quad (2.3.7)$$

Exemple 2.3.25. 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

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

Ainsi un **demi-ton** correspond à une augmentation de 6% des fréquences. Cela s'obtient aussi si on différencie (2.3.5), on obtient la formule suivante

$$\frac{df}{f} = \frac{\ln 2}{12} dx \sim 0.057 dx \quad (2.3.8)$$

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

Exercice 2.3.26. 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 2.3.27. $f_{A\#_5} = 440 \times 1.06 = 466\text{Hz}$. $f_{A_6} = 2 \times 440 = 880\text{Hz}$. $f_{A\#_6} = 880 \times 1.06 = 933\text{Hz}$

Exercice 2.3.28. (TP) "Mesure du pitch". A partir de la fréquence f mesurée dans l'exercice ??, calculer le pitch x (et son incertitude de mesure) et déduire le nom de la note.

Solution 2.3.29. On avait mesuré $f = 444\text{Hz}$ avec une incertitude $\Delta f = 20\text{Hz}$. On déduit

$$x = \frac{12}{\ln 2} \ln \left(\frac{444}{440} \right) + 69 = 69.15$$

avec une incertitude $\Delta x = \frac{12}{\ln 2} \frac{\Delta f}{f} = \frac{12}{\ln 2} \frac{20}{440} = 0.78$. La note semble donc être A_5 .

Exercice 2.3.30. Si le pitch x est modifié de 3 tons (un "triton"), comment est modifié la fréquence f ?

Solution 2.3.31. Pour un triton = 3 tons = 6 1/2 tons, $x' = x + 6$ donc

$$\frac{f'}{f} = 2^{\frac{6}{12}} = 2^{\frac{1}{2}} = \sqrt{2} = 1.414 \dots$$

Exercice 2.3.32. 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 $f_{A_5} = 440 \text{ Hz}$?

Solution 2.3.33. $\frac{1}{10} 6\% = 0.6\%$. Ainsi $\Delta f = 0.006 \times 440 \approx 3\text{Hz}$.

La définition suivante est parfois utilisée

Définition 2.3.34. 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 2.3.35. 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 2.3.36. (TP) 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

Dans la définition (2.3.5) précédente toutes les valeurs $x \in \mathbb{R}$ sont possibles. En musique occidentale il est convenu d'utiliser seulement les valeurs entières du pitch $x \in \mathbb{Z}$:

Définition 2.3.37. Le **tempérament égal** est le choix de notes (i.e. signaux périodiques) dont le pitch $x \in \mathbb{Z}$ est un **entier**. On décompose alors

$$x = 12 \times o + r$$

avec un entier $o \in \mathbb{Z}$ appelé **octave** et un reste entier $0 \leq r < 12$ appelé degré de la note auquel on associe un **symbole** :

r	0	1	2	3	4	5	6	7	8	9	10	11
symbole	C	C#	D	D#	E	F	F#	G	G#	A	A#	B
nom français	do		ré		mi	fa		sol		la		si

Remarque 2.3.38. De plus on peut ajouter la valeur de l'octave o en indice, par exemples

x	$50 = 12 \times 4 + 2$	$60 = 12 \times 5 + 0$	$67 = 12 \times 5 + 7$	$69 = 12 \times 5 + 9$
symbole	D_4	C_5	G_5	A_5

Remarque 2.3.39. Le choix d'utiliser le tempérament égal a été **fait par les luthiers au XIX^{ème} 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”.

Voir l'annexe [A](#) pour les différentes conventions et représentations du tempérament égal.

Exercice 2.3.40. Calculer les fréquences des notes C_5 et A_4 suivantes ?



Solution 2.3.41. Pour C_5 , $x = 60$ donc $f_{C_5} = 2^{\frac{1}{12}(x-x_{A_5})} f_{A_5} = 2^{\frac{1}{12}(60-69)} 440 = 2^{-9/12} 440 = 262\text{Hz}$. Pour A_4 , $x = x_{A_5} - 12$ donc $f_{C_5} = 2^{\frac{1}{12}(-12)} f_{A_5} = \frac{1}{2} 440 = 220\text{Hz}$

2.3.3 Comparaison des harmoniques avec le tempérament égal

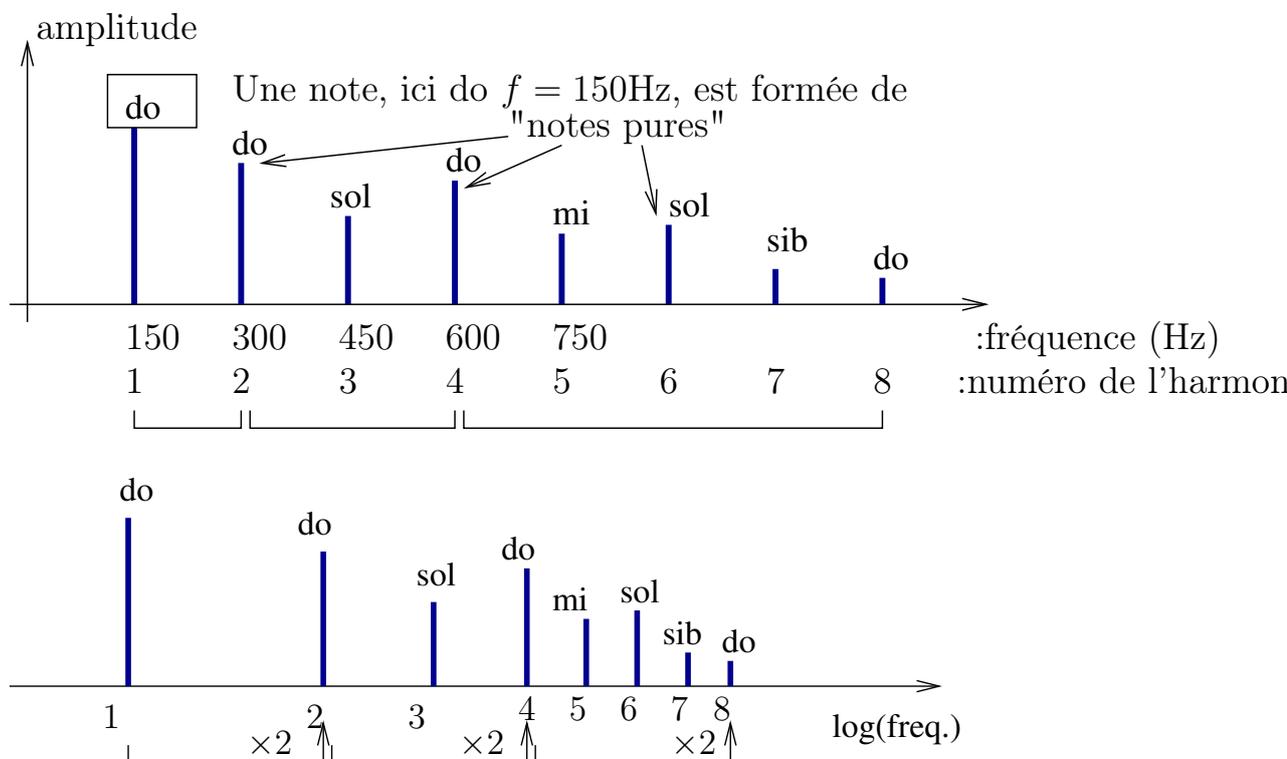
Rappels :

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

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

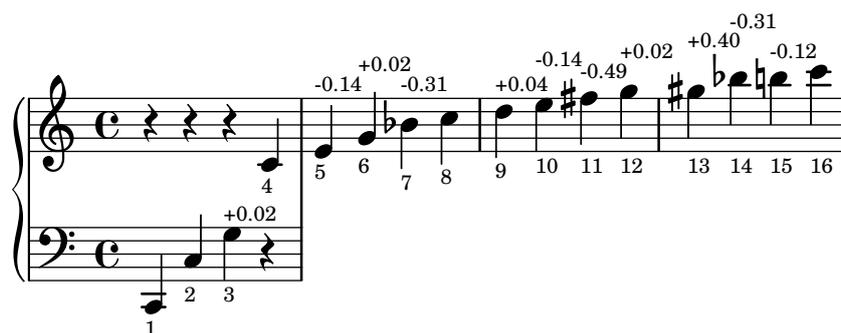
$f_1 = f$ est appelée la fondamentale.

- 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 2.3.42. *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 différence 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 ??, on applique la formule (2.3.5). Par exemple pour l’harmonique $a = 3$, par rapport à l’harmonique $a = 1$, on écrit :

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

Exercice 2.3.43. Calculer de même les corrections pour les harmoniques $a = 5, 7, 11, 13$ et vérifier avec la figure ci-dessus.

Remarque 2.3.44. Du fait que $9 = 3 \times 3$, et d’après la formule (A.2.1), on peut écrire :

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

Exercice 2.3.45. 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.

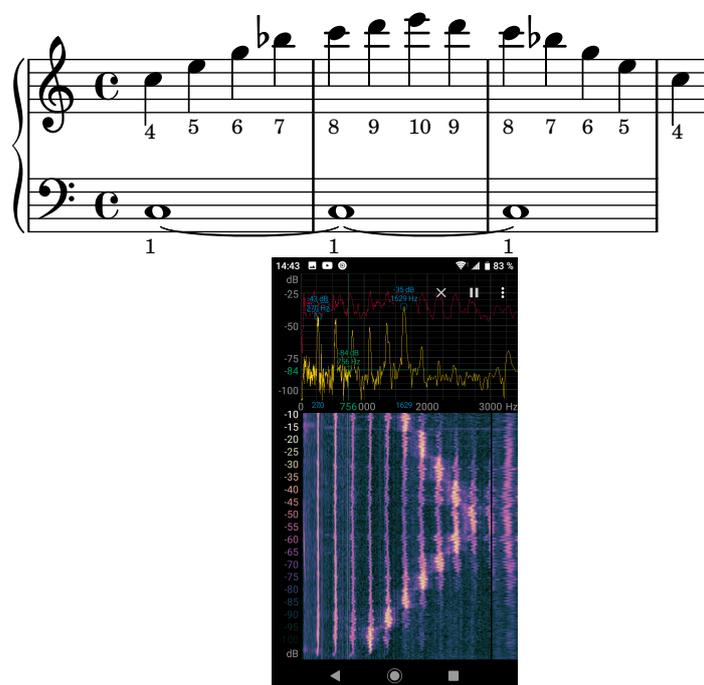
2.3.4 Exemple du chant diphonique

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

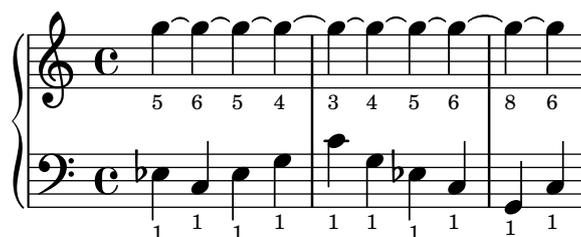
Il y a une technique vocale appelée “**chant diphonique**” qui permet de mettre beaucoup d’intensité dans l’une ou l’autre des harmoniques, Voici cette technique expliquée dans une [Conférence de Anna-Maria Hefe](#).

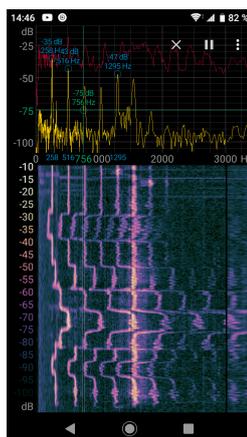
Nous commentons ici quelques exemples précis. Peut écouter le chant **diphonique overtone singing** - [Anna-Maria Hefe](#) sur youtube et on donne la partition montrant la fréquence de la note (en basse) et l’harmonique renforcée sur la portée musicale et avec le numéro n de l’harmonique. Attention : 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 ??.

Exemple 2.3.46. Dans ce passage de [Anna-Maria Hefe](#) à 29”. Il y a une basse commune f_1 .



Exemple 2.3.47. Dans ce passage de [Anna-Maria Hefe](#) à 1’01”. La fréquence f_n de l’harmonique renforcée est fixe, mais le numéro n change, par conséquent la fréquence de la fondamentale f_1 change :





Exemple 2.3.48. 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 2.3.49. 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 2.3.50. 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 2.3.51. 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 2.3.52. 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

[Video de cette section](#)

2.3.5.1 Définition

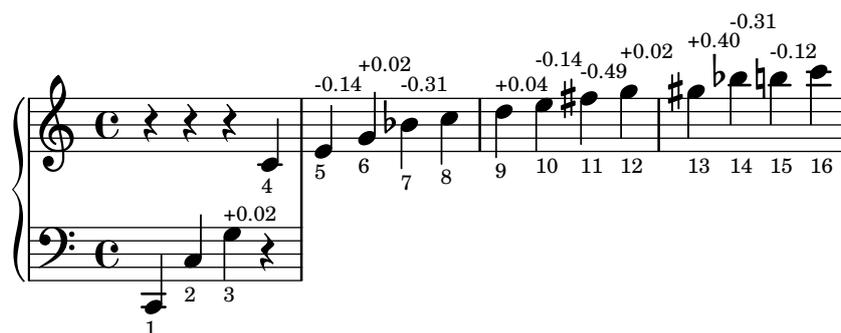
La définition suivante est très importante en musique.

Définition 2.3.53. Un **intervalle juste** est un rapport entre deux fréquences appartenant aux harmoniques $a, b \leq 16$ d'une note donnée.

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

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

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

Exercice 2.3.55. Vérifier les colonnes 3,4 de cette liste et continuer la liste.

Solution 2.3.56. :

Harmoniques $a \rightarrow b$	Nom de l'intervalle	Fraction $\frac{f'}{f}$	Pitch x (en demi-tons)
9 \rightarrow 10	Ton 3	$\frac{10}{9}$	$2 - 0.10$
10 \rightarrow 11	Ton 4	$\frac{11}{10}$	$2 - 0.35$
11 \rightarrow 12	Ton 5 ou demi-ton 1	$\frac{12}{11}$	$2 - 0.49 = 1 + 0.51$
12 \rightarrow 13	demi-ton 2	$\frac{13}{12}$	$1 + 0.38$
13 \rightarrow 14	demi-ton 3	$\frac{14}{13}$	$1 + 0.29$
14 \rightarrow 15	demi-ton 4	$\frac{15}{14}$	$1 + 0.19$
15 \rightarrow 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 XXIème 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 2.3.57. 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 $f_1 = 440\text{Hz}$. Ce battement est utilisé par les accordeurs de piano, voir cette [vidéo sur accordage des quintes](#).

Solution 2.3.58. 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 $f_1 = 220\text{Hz}$. 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 2.3.59. 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).
- Harmoniques de la voix
 - [Wave form de voyelles chantées](#) par Malik. Harmoniques et notes correspondantes en [Video](#)
 - [polyphonic overtone singing - Anna-Maria Hefe](#)

- [Conférence de Anna-Maria Hefe](#)
- 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 **nombre premiers** : 2, 3, 5, 7, 11, 13, ... Par exemple $4 = 2 \times 2$, $6 = 2 \times 3$, $8 = 2 \times 2 \times 2 = 2^3$ etc.

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

Harmoniques $a \rightarrow b$	Nom de l'intervalle	Fraction $\frac{f'}{f}$	Pitch x (en demi-tons)
1 \rightarrow 2	Octave	2	12
1 \rightarrow 3	octave + Quinte	3	12 + 7 + 0,02..
1 \rightarrow 5	2 octaves + Tierce maj.	5	2 \times 12 + 4 - 0,14..
1 \rightarrow 7	2 octaves + septieme min.	7	2 \times 12 + 10 - 0,33
1 \rightarrow 11	3 octaves + triton bas	11	3 \times 12 + 6 - 0,49
1 \rightarrow 13	3 octaves + sixte haute	13	3 \times 12 + 9 + 0,40

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

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

cela correspond au calcul suivant.

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

Ou de façon équivalente et plus directe :

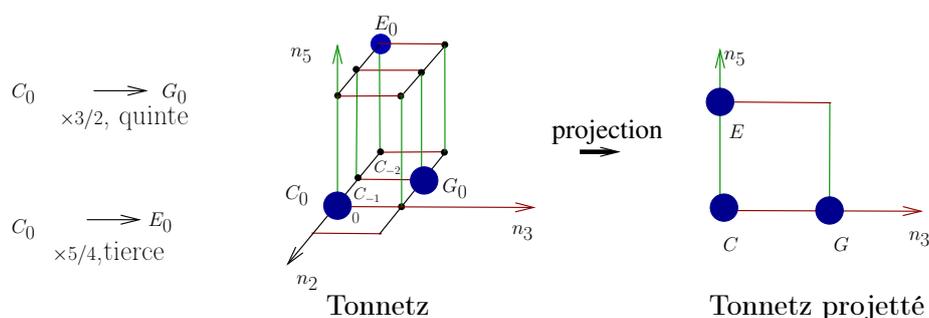
$$\frac{6}{5} = \frac{3}{2} \times \frac{4}{5} = \frac{3}{\frac{5}{4}} \equiv \text{quinte - tierce}$$

Lien :

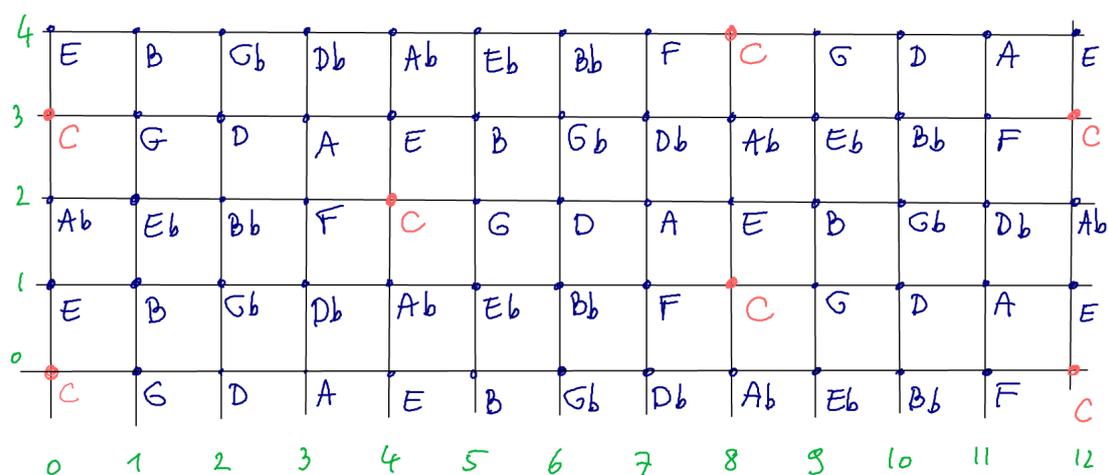
- Vidéo sur les nombres premiers sur [Arte, voyage au pays des maths, Conjecture de Riemann](#).

2.4.2 Le tonnetz 2, 3, 5

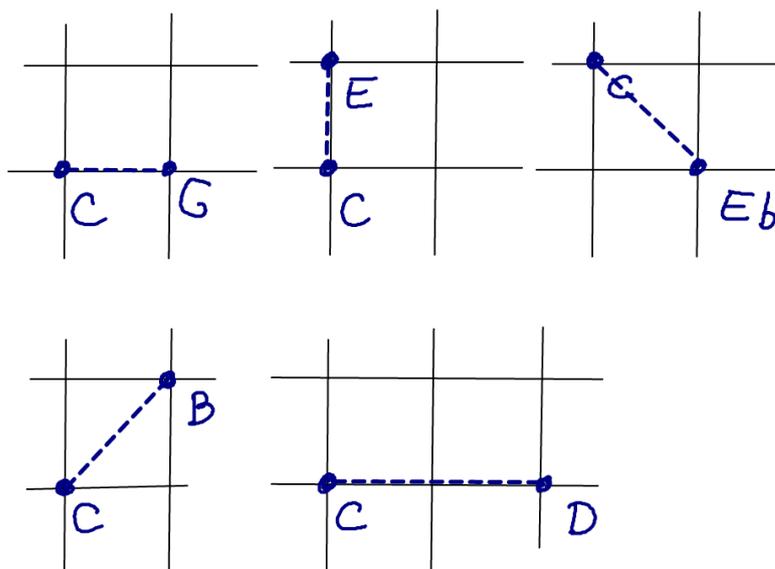
Dans le tonnete 2, 3, 5, par simplicité, on considère seulement les 3 premiers facteurs premiers 2, 3, 5 et comme changer d'octave n'affecte pas le nom de la note, on omettra de préciser l'octave (i.e. le facteur 2). Les axes n_2, n_3, n_5 correspondent respectivement aux changement de fréquence $\times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



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 apparaît 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

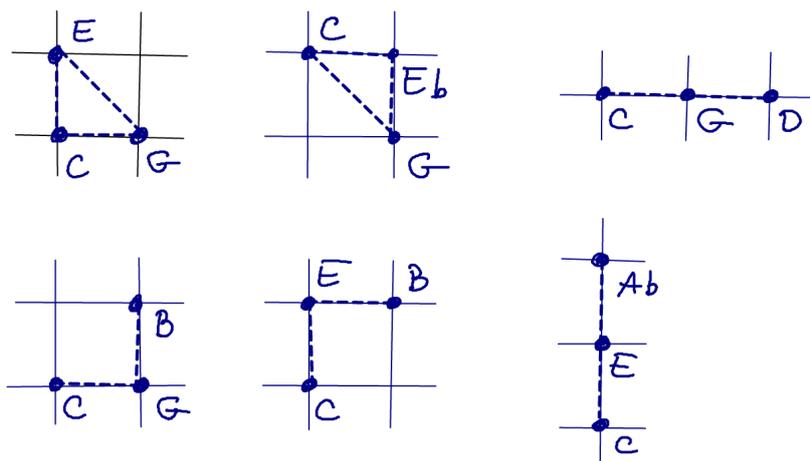
$$(4, 0) - (4, -1) = (0, 1).$$

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

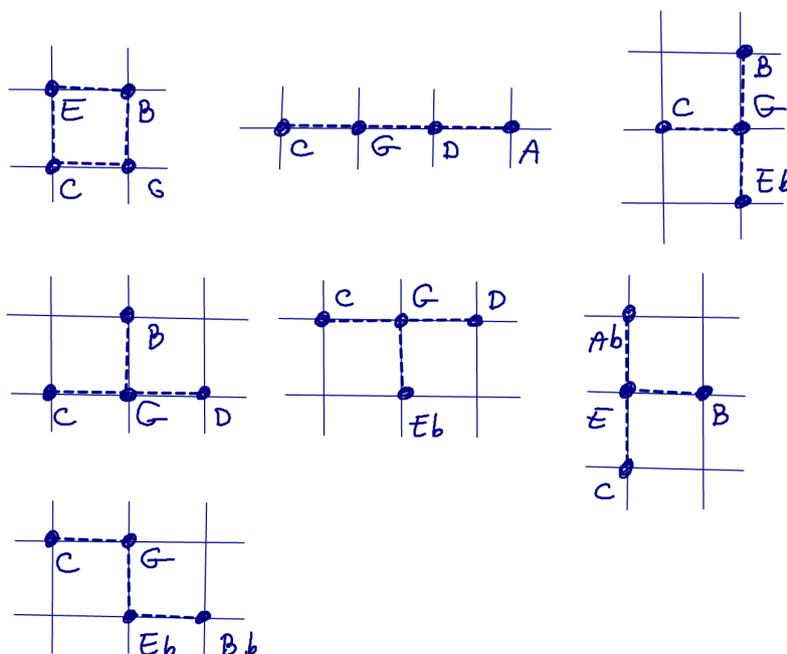
2.4.2.2 Quelques accords justes

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 perçoit donc une **sensation de manque**. [Video](#)

Plus précisément, on peut quantifier cela en 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 $(1, -2)$.

The diagram shows a grid of notes in a Zarlino temperament. The notes are arranged in a 3x4 grid. The top row contains A, E, B, Gb. The middle row contains F, C, G. The bottom row contains Db, Ab. Red arrows indicate the sequence of movements between notes, numbered 1 through 6. Below the grid is a musical staff in G-flat major (one flat) with six chords corresponding to the notes in the grid. Red arrows below the staff indicate the sequence of movements between the chords, numbered 1 through 6.

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

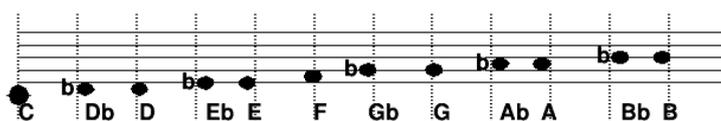
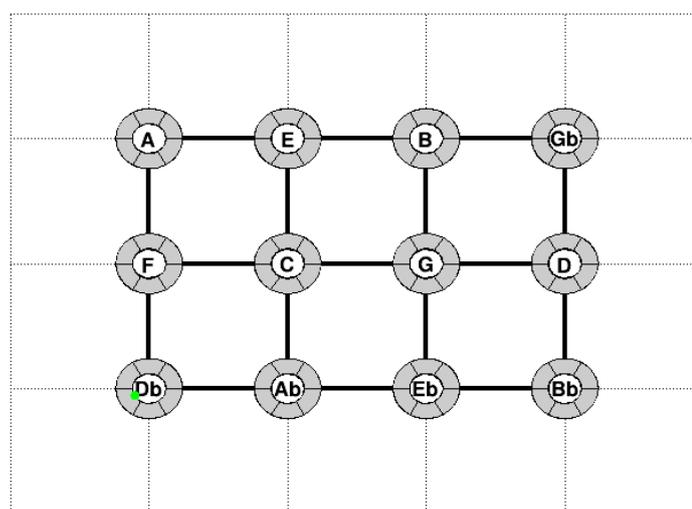
Comme le réseau tonnetz contient en principe une infinité de notes, cela pose des difficultés pour jouer ces notes. Sur un instrument traditionnel, en pratique pour jouer avec

des intervalles justes, il faut choisir les notes que l'on pourra utiliser. Ce choix s'appelle un tempérament juste.

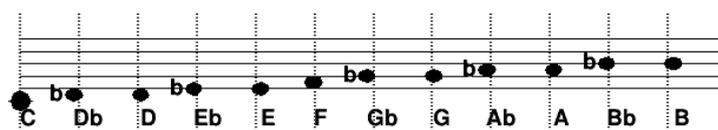
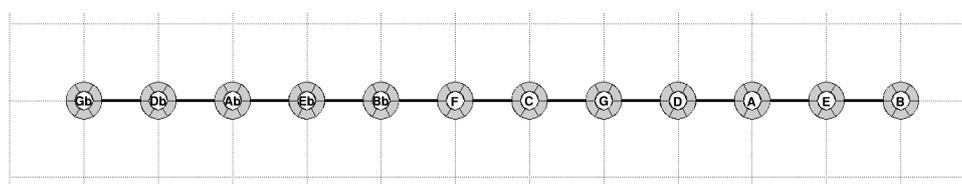
Définition 2.4.7. Un tempérament juste dans le tonnetz 2,3,5 est le choix des douze notes chromatiques C, Db, D, \dots, Bb, B parmi le réseau tonnetz ci-dessus. Ce choix doit être groupé afin de contenir des intervalles et accords justes résonnants que l'on pourra utiliser.

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**.



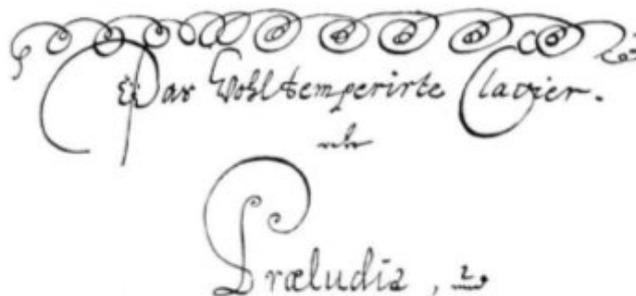
- Le tempérament de **Kepler** dans *Harmonices Mundi* 1619 ([video](#))



- 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 \flat 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 \sharp 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 Hallelujahs, 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 \flat 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 \flat minor. If ghosts could speak, their speech would approximate this key.

Vidéos :

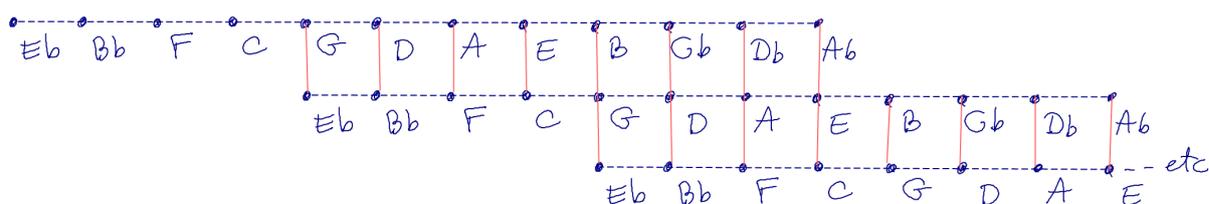
- En tempérament égal :
 - [Gymnopédie No. 1 –{ } Tonnetz Visualization](#)
 - [Tonnetz Analysis of Chopin E-major Prelude](#)
 - [Muse - Take A Bow \(Tonnetz harmonic analysis\)](#)
 - Limitations du tonnetz 2,3,5 : [How to create negative harmony using the Tonnetz tone grid](#)

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 $Eb, Bb, F, C, \dots, Db, Ab$, 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. Montrer que les 8 tierces suivantes sont justes : $Eb - G - B$, $Bb - D - Gb$, $F - A - Db$, $C - E - Ab$.

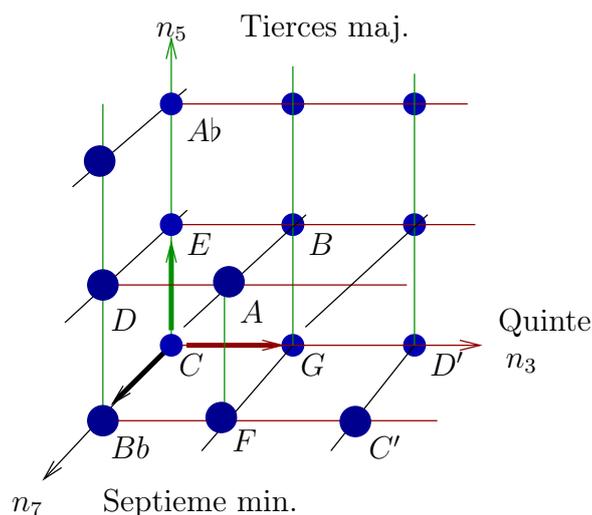
Solution 2.4.10. Dans la suite des quintes justes entre C et E on a : $C - G - D - A - E$ il y a 4 quintes justes. Cet intervalle $C - E$ s'appelle la **tierce de pythagore** et son pitch est : $x = (7 + 0.02) \times 4 \bmod 12 = 4 + 0.08$. Or la tierce juste est de pitch $x_{\text{tierce}} = 4 - 0.14$. 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 $Eb, Bb, F, C, \dots, Db, Ab$, 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 $E_b, G, B_b, D, F, A, C, E$ et dans leurs 8 tonalités relatives mineures $C, E, G, B, D, G_b, A, D_b$.

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 :

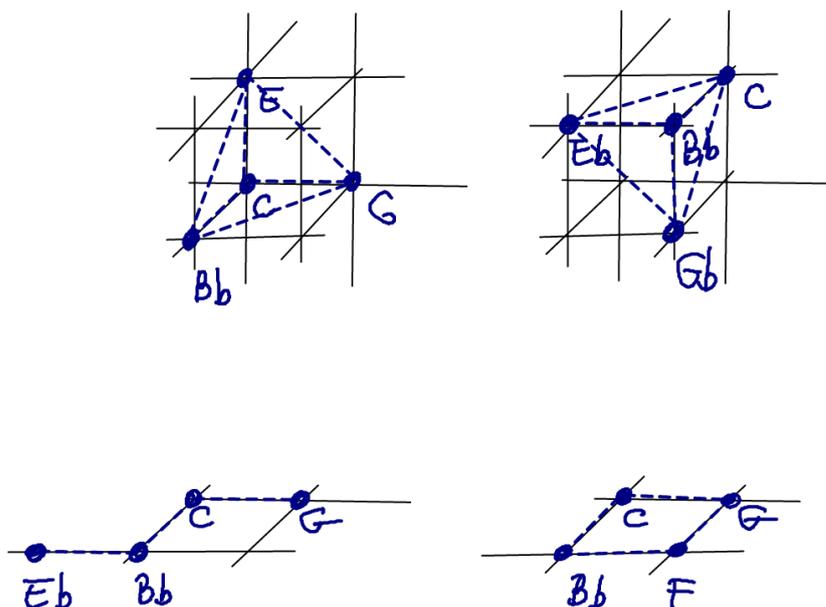


avec l'axe des quintes, l'axe des tierces majeures et l'axe des septièmes mineures.

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 :

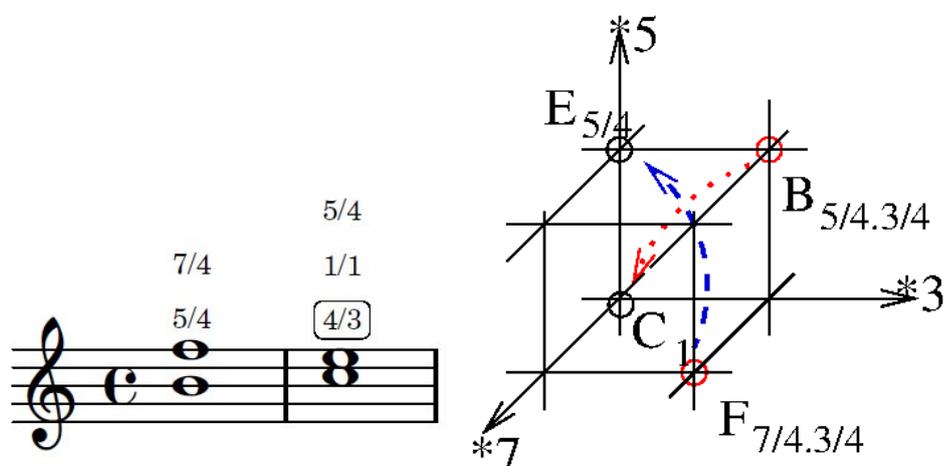


Remarque 2.4.12. Il apparaît des accords bien différents que l'on confondrait en musique à tempérament égal, comme C, Eb, G, Bb qui a une version dans le plan 3, 7 et une version dans le plan 3, 5.

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 $F \rightarrow E$ (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 $B \rightarrow C$ (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 Échantillonnage d'un sinus, effet de repliement (ou effet stroboscopique), Aliasing

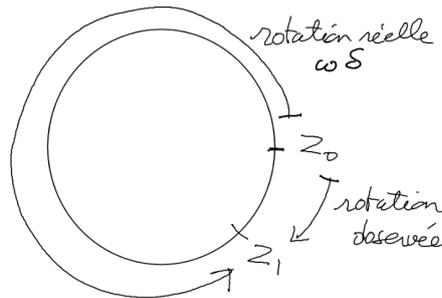
2.5.1 Rappels sur l'effet stroboscopique

L'**effet stroboscopique** correspond au fait qu'un mouvement de rotation $t \in \mathbb{R} \rightarrow z(t) = e^{i2\pi f t} \in \mathbb{C}$ avec une fréquence $f \in \mathbb{R}$, observé à des temps discrets $t_j = j\delta$ avec un pas de temps $\delta > 0$ et $j \in \mathbb{Z}$, donne une suite de valeurs discrètes $z_j = z(t_j) = e^{i2\pi f \delta j}$ qui peut être interprété comme provenant d'une autre fréquence f_k donnée par

$$f_k = f + \frac{1}{\delta}k \quad (2.5.1)$$

avec n'importe quelle valeur entière $k \in \mathbb{Z}$. En effet

$$e^{i2\pi f_k \delta j} = e^{i2\pi (f + \frac{1}{\delta}k) \delta j} = e^{i2\pi f \delta j + 2\pi k j} = e^{i2\pi f \delta j} = z_j.$$



La suite $(f_k)_{k \in \mathbb{Z}}$ forme une **suite périodique de fréquences de période $\frac{1}{\delta}$** . Parmi celles ci, la fréquence perçue est souvent celle de valeur absolue minimale, donc comprise dans l'intervalle

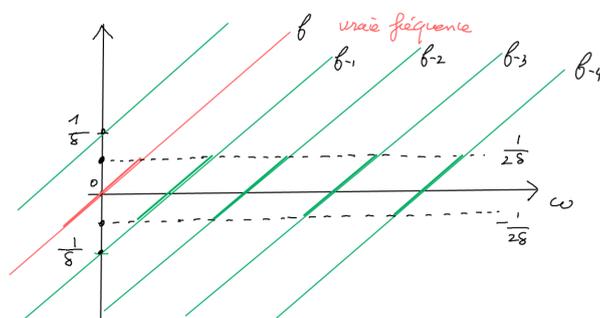
$$f_k \in \left[-\frac{1}{2\delta}, \frac{1}{2\delta} \right].$$

On appelle

$$F_{\text{Nyquist}} = \frac{1}{2\delta}$$

2.5. ÉCHANTILLONNAGE D'UN SINUS, EFFET DE REPLIEMENT (OU EFFET STROBOSCOPIQUE)

la **fréquence de Nyquist**. Supposons que $\delta > 0$ est fixé et ω augmente. Voici le graphe de f_k en fonction de ω en gras, les valeurs f_k qui seront sélectionnées par notre perception, i.e. les plus faibles en valeur absolue.



Exemple 2.5.1. Si on observe une roue qui tourne à la fréquence $f = 49$ tours/sec = 49Hz, éclairée par un néon, $\delta = \frac{1}{50}$ sec alors

$$f_k = f + 50k \text{ tours/sec}$$

et parmi cette liste de fréquences il y a pour $k = -1$, une fréquence faible, $f_{-1} = -1$ tour/sec que notre perception va interpréter comme le vrai mouvement. Il nous semblera donc que la roue tourne en sens contraire, lentement avec -1 tour/sec.

Exercice 2.5.2. (TP*) Faire un programme qui simule cet effet stroboscopique avec une animation. (Utiliser cet **exemple**).

Exercice 2.5.3. Une roue de vélo avec $N = 60$ rayons symétriquement répartis, tourne à une vitesse $f = 5$ tour/sec, éclairée par un néon de fréquence $\delta^{-1} = 50$ Hz. A quelle vitesse f_k nous semble t-il que la roue tourne ?

Solution 2.5.4. La périodicité est maintenant l'écart entre deux rayons soit $\frac{2\pi}{N}$. La condition est donc

$$g_k = g + \frac{1}{\delta N}k = 5 + \frac{50}{60}k$$

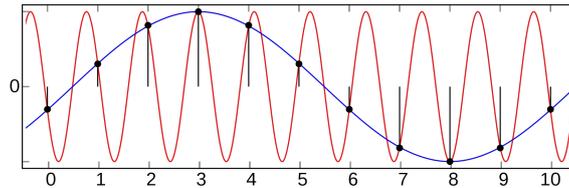
La valeur la plus faible de f_k est obtenue pour $k = -6$ donnant

$$f_{-6} = 5 - 6 \frac{50}{60} = 0$$

La roue paraît donc immobile.

2.5.2 Effet stroboscopique sur une fonction sinus ou cosinus échantillonnée

Observation : From wikipedia : ce graphique montre un signal sinusoïdal de fréquence $f = 0,9$ en rouge, échantillonné avec un pas $\delta = 1,0$. L'échantillonnage pourrait aussi provenir du signal de fréquence $f' = 0,1$ en bleu (ou encore d'autres signaux de fréquences notées $\pm f_k$ dans l'explication ci-dessous). Remarquer ici que $f' = -(f - \frac{1}{\delta})$ noté $-f_{-1}$ ci-dessous.



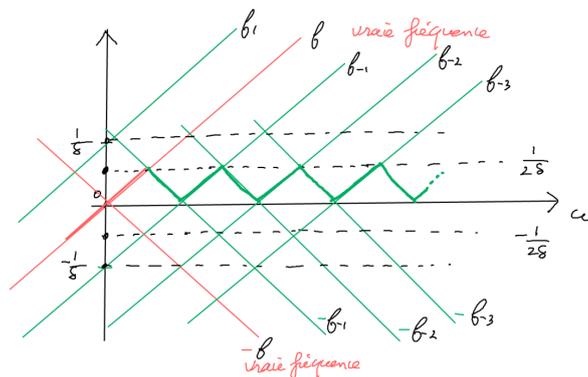
Explication : Le phénomène stroboscopique se produit sur un signal $t \rightarrow \cos(2\pi ft)$ échantillonné avec un pas de temps $\delta > 0$ car

$$\cos(2\pi ft) = \frac{1}{2} (e^{i2\pi ft} + e^{-i2\pi ft})$$

est composé des deux fréquences $f, -f$. Donc d'après (2.5.1), après échantillonnage ce sera équivalent à deux séries de signaux de fréquences

$$\pm f_k = \pm \left(f + \frac{1}{\delta} k \right), \quad k \in \mathbb{Z}, \quad (2.5.2)$$

Voici le graphe de $\pm f_k$ en fonction de f avec δ fixé.



On observe une périodicité et du fait des fréquences négatives $-f_k$, on observe comme un **rebond du graphe des fréquences dans l'intervalle** $[0, \frac{1}{2\delta}]$ appelé **repliement des fréquences** aussi appelé **aliasing**. Rappel : $F_{\text{Nyquist}} = \frac{1}{2\delta}$ s'appelle la **fréquence de Nyquist**. On peut aussi interpréter le résultat comme une symétrie de la fréquence f par

2.5. ÉCHANTILLONNAGE D'UN SINUS, EFFET DE REPLIEMENT (OU EFFET STROBOSCOPIQUE)

rapport à un axe $k\frac{1}{2\delta}$ avec $k \in \mathbb{Z}$ de façon à se ramener dans l'intervalle $[0, \frac{1}{2\delta}]$, mais le plus clair est de retenir que le spectre est composé des fréquences $(-f, f)$ périodisé avec une période $\frac{1}{\delta}$ en fréquence.

Exemple 2.5.5. Voici un signal sinus de fréquence f variable et échantillonné à la fréquence $F = 44100\text{Hz}$. Ecouter [sinus_aliasing.wav](#). Programme python [sinus_aliasing.py](#). Dessin gif animé [sinus_aliasing.gif](#).

2.5.3 Effet stroboscopique (aliasing) sur un signal périodique quelconque

Pour un **signal périodique quelconque** de fréquence fondamentale $f > 0$, on a vu en (2.3.1) qu'il se décompose en somme de cosinus de fréquences $f_n = nf$ avec $n \in \mathbb{N}$ et des amplitudes correspondantes A_n . Il donc réellement composé des fréquences $\pm f_n = \pm nf$. Par conséquent, en appliquant la formule (2.5.2) on obtient le résultat suivant.

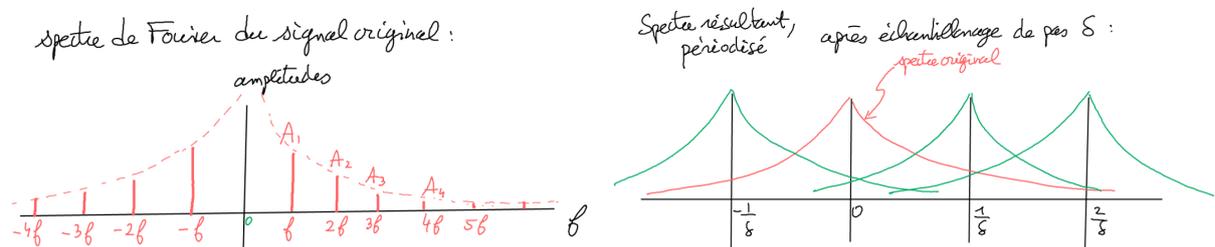
Proposition 2.5.6. *Un signal périodique de fréquence fondamentale $f > 0$ échantillonné avec un pas de temps $\delta > 0$, est composé des fréquences*

$$\pm f_{n,k} = \pm \left(nf + \frac{k}{\delta} \right), \quad \text{avec } n \in \mathbb{N}, k \in \mathbb{Z}.$$

c'est à dire que c'est le spectre original $\pm nf$ périodisé en fréquence avec une période $\frac{1}{\delta}$.

Remarque 2.5.7. On dit aussi qu'il y a un rebond ou repliement (aliasing) de toutes ces fréquences $\pm f_n = nf$ dans l'intervalle $[0, \frac{1}{2\delta}]$.

Voici le schéma que l'on peut faire de ce résultat, correspondant au repliement du spectre des fréquences :



Remarque 2.5.8. si f et $\frac{1}{\delta}$ ne sont pas commensurables, c.a.d. si $f\delta \notin \mathbb{Q}$ alors l'ensemble des fréquences $\pm f_{n,k}$ est dense dans \mathbb{R} . En pratique les amplitudes A_n décroissent de sorte que la mesure est finie.

Conclusion 2.5.9. D’après la figure précédente on observe que les hautes fréquences du signal original donnent composantes à basse fréquence dans le signal périodisé. Ce phénomène est appelé “**Aliasing**” et est **perçu comme des parasites**. Pour l’éviter il est **important de filtrer** le signal original en ne gardant que les fréquences $f \in [-\frac{1}{2\delta}, \frac{1}{2\delta}]$ (appelé filtre basse fréquence à la fréquence de Nyquist $F_{\text{Nyquist}} = \frac{1}{2\delta}$) avant de l’échantillonner.

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

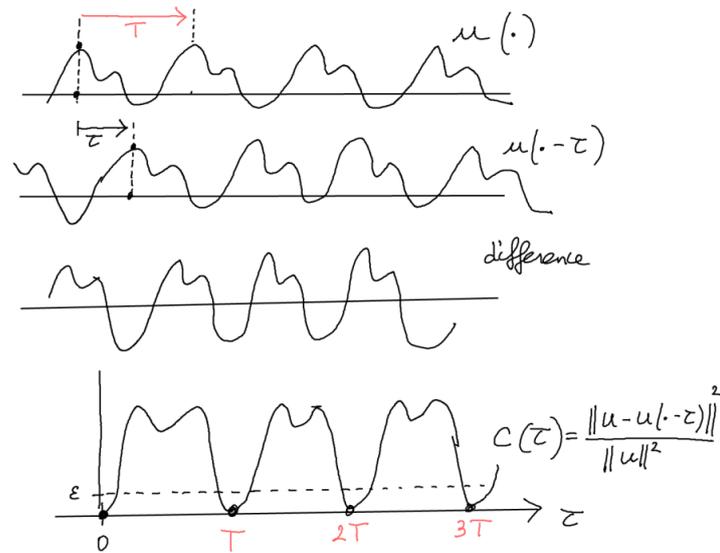
- Le **vibrato** est une modulation de la fréquence. Cela se fait au violon, au trombone, à la voix.
- Le **trémolo** est une modulation de l’amplitude. Ce se fait à la flûte, à la voix.

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

Un **signal monophonique** est un signal qui localement est presque périodique, i.e. correspond localement à une note de musique. Cela provient en général d’une voix humaine ou d’un instrument de musique seul. En musique il peut être intéressant de **connaitre les notes jouées**, pour par exemple écrire la partition correspondante de façon automatique. Dans cette section, on explore cette question.

Question 2.6.1. *à partir d’un signal monophonique (i.e. un instrument seul) donné $t \rightarrow u(t)$ comment détecter automatiquement le pitch (i.e. hauteur de note) au cours du temps $x(t)$?*

Solution 2.6.2. (détails ci-dessous) Nous avons vu que le pitch d’une note seule $x(t)$ est défini par la « période locale » $T(t)$ du signal. Un algorithme efficace consiste, pour un instant t donné, à considérer un court extrait du signal noté $u(\cdot) : t \rightarrow u(t)$ et de le comparer à lui même décalé d’une durée $\tau > 0$. Ce signal décalé est la fonction notée $u(\cdot + \tau) : t \rightarrow u(t + \tau)$. On note $C(\tau) = \frac{\|u(\cdot) - u(\cdot + \tau)\|^2}{\|u(\cdot)\|^2}$ la mesure relative de cette différence où $\|\cdot\|$ est une norme comme la norme L^2 définie ci-dessous. La période $T(t)$ est la valeur de τ qui minimise localement $C(\tau)$.



Remarque 2.6.3. Cet algorithme a une latence (i.e. un retard) inférieur à T_{\min} qui est la période minimale que l'on souhaite détecter. Par exemple si la note la plus grave à détecter est C_3 , de pitch $x_{C_3} = 36$, alors (2.3.7) donne $f_{\min} = 65.4\text{Hz}$, $T_{\max} = 0.015$ sec qui est une latence acceptable. Cet algorithme s'adapte donc très bien à une analyse audio en temps réel.

En peut ensuite faire une transformée de Fourier sur une période T et déduire les coefficients \hat{u}_n , appelé "timbre".

Cet algorithme est implémenté en langage wasm [ici](#), section 5, ce qui permet d'utiliser le programme dans un navigateur web.

2.6.2.1 Détail de l'algorithme et programme

On présente ici les détails de l'algorithme esquissé en section 2.6.2.

Entrée : Soit $t_{\max} > 0$ et une fonction $u : t \in [0, t_{\max}] \rightarrow u(t) \in \mathbb{R}$ non identiquement nulle.

Question : On veut savoir si $u(t)$ se "répète" avec une période $T \in [0, \frac{t_{\max}}{2}]$.

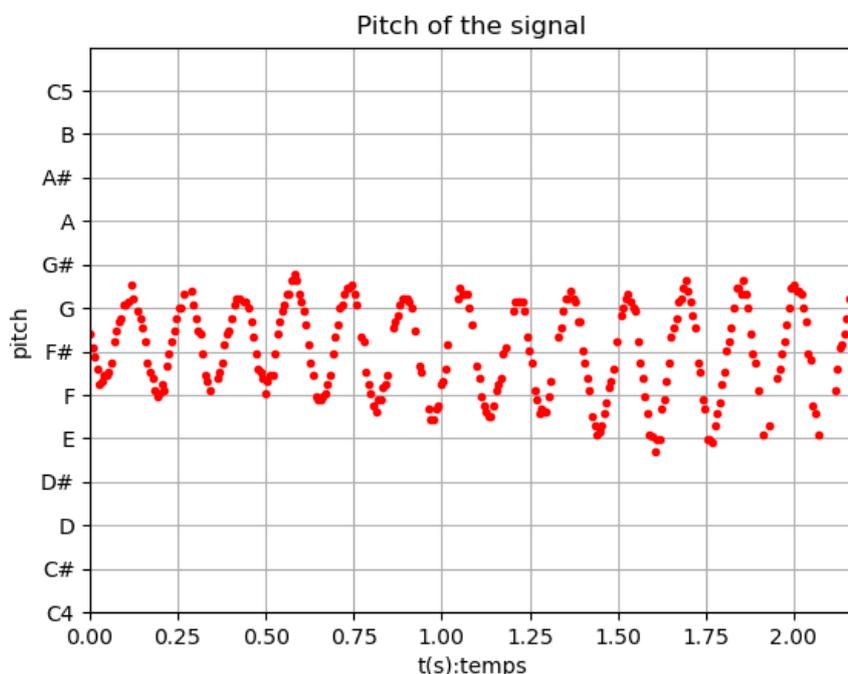
Algorithme : Soit $0 < \tau < \frac{t_{\max}}{2}$. On note T_τ° l'opérateur de translation de τ

$$(T_\tau^\circ u)(t) = u(t + \tau).$$

Soit

$$C(\tau) := \frac{1}{\|u\|_{L^2(0, t_{\max})}^2} \|u - T_\tau^\circ u\|_{L^2(0, \frac{t_{\max}}{2})}^2$$

Solution 2.6.6. Solution Marian Anderson :



On observe une fréquence de 6Hz et une amplitude jusqu'à ± 1 ton.

(*) Améliorer l'algorithme et le programme de détection de pitch, en effectuant une interpolation quadratique pour la recherche du minimum local de $C(\tau)$.

Piste de recherche :

- Comment détecter les pitches d'un signal polyphonique? (mots clefs : multipitch detection)
- Comment écrire la partition automatiquement avec lilypond ou autre? Pour cela il faut détecter les changement de pitch, leur valeur et la date de début et de fin.

2.6.3 Filtres

2.6.3.1 Définition et propriétés d'un filtre ou opérateur

Un **filtre** est une opération T qui transforme un signal $u : t \rightarrow u(t)$ en un autre signal $v : t \rightarrow v(t)$. Mathématiquement c'est donc une application dans l'espace des fonctions. En mathématique on appelle cela un **opérateur**. Dans la suite, on utilisera donc le terme **opérateur** ou **filtre**, de façon équivalente, selon que le contexte est plutôt mathématique ou plutôt analyse du signal. On s'intéresse dans ce chapitre uniquement aux signaux sonores, donc aux fonctions dépendant d'une seule variable $t \in \mathbb{R}$, mais ce qui suit se généralise à des fonctions à plusieurs variables. En analyse des images, on utilise 2 variables x_1, x_2 , en analyse des vidéos, on utilise 3 variables x_1, x_2, t , etc

On va considérer des opérateurs T ayant les propriétés suivantes :

Opérateur linéaire si $\forall u, v$ fonctions et $\forall a, b \in \mathbb{C}$,

$$T(au + bv) = aT(u) + bT(v)$$

Physiquement, cette propriété apparaît de façon approximative dans le régime des faibles amplitudes (d'après le développement de Taylor). On peut alors penser à T comme une matrice et les signaux u, v comme des vecteurs, i.e. appliquer l'algèbre linéaire. Dans le cas d'un signal numérique échantillonné $u = (u_j)_{j=0 \rightarrow N-1}$, on peut écrire l'opérateur T sous la forme matricielle

$$(Tu)_i = \sum_{j=0}^{N-1} T_{i,j} u_j \quad (2.6.1)$$

dans le cas d'un signal discrétisé. Les coefficients $(T_{i,j})_{i,j=0 \rightarrow N-1}$ forment une matrice et caractérisent l'opérateur. Si $\delta_i = \left(0, \dots, 0, \underset{(i)}{1}, 0, \dots\right)$ désigne un vecteur de base, on a

$$T_{i,j} = \langle \delta_i | T \delta_j \rangle.$$

Dans le cas du temps continu $t \in \mathbb{R}$, on écrit

$$(Tu)(t) = \int T(t, t') u(t') dt' \quad (2.6.2)$$

et la fonction $T(t', t)$ s'appelle en mathématique le **noyau de Schwartz** de l'opérateur, et en théorie du signal s'appelle la **fonction de transfert en temps** du filtre. Si δ_t désigne la distribution de Dirac en t , on a²

$$T(t, t') = \langle \delta_t | T \delta_{t'} \rangle.$$

Remarque 2.6.7. D'après (2.1.1) on peut écrire $u(t) = \langle \delta_t | u \rangle$ et donc (2.6.2) s'écrit aussi

$$\langle \delta_t | Tu \rangle = \int \langle \delta_t | T \delta_{t'} \rangle \langle \delta_{t'} | u \rangle dt',$$

qui est analogue à (2.6.1).

Remarque 2.6.8. Avec des signaux d'amplitude trop forte des effets « non linéaires » peuvent apparaître, cad non modélisable par un opérateur linéaire, comme la génération d'harmoniques. Voir [Nonlinear_acoustics](#). En particulier la génération des sons dans les instruments de musique nécessitent des effets non linéaires (ex : instruments à vent, à archet. Les modèles linéaires suffisent pour les percussions et cordes pincées ou frappées), voir [ce cours d'acoustique musicale](#).

2. En effet $\langle \delta_t | T \delta_{t'} \rangle \stackrel{(2.6.2)}{=} \int T(t, t'') \delta_{t'}(t'') dt'' = T(t, t')$.

Opérateur causal pour toute fonction u , et pour tout instant t , la valeur $v(t) = (Tu)(t)$ ne dépend pas des valeurs ultérieures $u(t')$ avec $t' > t$. D'après l'écriture (2.6.1) cela s'exprime par $T_{i,j} = 0$ si $j > i$ et pour (2.6.2) cela s'exprime par $T(t, t') = 0$ si $t' > t$.

Remarque 2.6.9. Si on traite un signal sonore en « temps réel », on ne peut que utiliser des opérateurs causaux (car le signal futur n'est pas connu), mais si on travaille « en différé » c'est à dire avec un signal déjà enregistré, il est possible et parfois utile d'utiliser des opérateurs n'ayant pas la propriété de causalité. (Par exemple le décalage en temps, opérateur de convolution, transformée par ondelettes etc)

Remarque 2.6.10. La causalité en physique est un fait observé, fondamental et presque évident pour le bon sens appelé relation de « cause à effet », appelé en physique aussi **flèche du temps**. Cependant la compréhension de ce phénomène est loin d'être satisfaisante, car les équations de la physique sont symétriques, donc pourquoi on n'aurait pas $T(t, t') = 0$ si $t' < t$ au contraire? En art, cette question est évoquée dans le film récent « TeneT », 2020.

Opérateur stationnaire Un opérateur $T : u \rightarrow v = Tu$ est **stationnaire** si un décalage temporel de u a pour réponse le même décalage temporel de v . Par exemple un filtre d'un ingénieur du son, dont les paramètres sont fixés (ne changent pas au cours du temps), est un « filtre stationnaire ».

En termes mathématiques, si \mathcal{T}_τ est l'opérateur de décalage temporel par τ défini par

$$\mathcal{T}_\tau : u(t) \rightarrow (\mathcal{T}_\tau u)(t) = u(t - \tau)$$

alors l'opérateur T est stationnaire si

$$\forall \tau \in \mathbb{R}, \quad T\mathcal{T}_\tau = \mathcal{T}_\tau T. \quad (2.6.3)$$

Proposition 2.6.11. *si T est un opérateur linéaire stationnaire alors son noyau de Schwartz vérifie*

$$T(t, t') = T(t - t', 0) = \chi(t - t') \quad (2.6.4)$$

*qui est en fait une fonction χ à une seule variable appelée **réponse impulsionnelle du filtre** car c'est l'image de la distribution de Dirac δ_0 par l'opérateur :*

$$\chi(t) := T(t, 0) = (T\delta_0)(t) \quad (2.6.5)$$

Démonstration. on écrit

$$\begin{aligned} (T\mathcal{T}_\tau u)(t) &= \int T(t, t') u(t' - \tau) dt' = \int T(t, t' + \tau) u(t') dt' \\ (\mathcal{T}_\tau T u)(t) &= \int T(t - \tau, t') u(t') dt' \end{aligned}$$

donc (2.6.3) implique $T(t, t' + \tau) = T(t - \tau, t')$, $\forall t, t', \tau$. En particulier, posant $t'' = t' + \tau$ et $\tau = t$ cela donne

$$T(t, t'') = T(t - \tau, t'' - \tau) = T(0, t'' - t), \forall t, t''.$$

Si on applique l'opérateur T à la distribution de Dirac δ_0 , on obtient

$$(T\delta_0)(t) \stackrel{(2.6.2)}{=} \int T(t, t') \delta_0(t') dt' \stackrel{(2.1.1)}{=} T(t, 0).$$

□

Transformée de Fourier appliquée à un opérateur stationnaire Pour une fréquence $f \in \mathbb{R}$ donnée, notons $e^{i2\pi f \cdot} : t \in \mathbb{R} \rightarrow e^{i2\pi f t} \in \mathbb{C}$ la fonction appelée mode de Fourier de fréquence f . Considérons "l'élément de matrice" suivant de l'opérateur dans la base de Fourier :

$$\langle e^{i2\pi f \cdot} | T e^{i2\pi f' \cdot} \rangle = \int \int e^{-i2\pi f t} T(t, t') e^{i2\pi f' t'} dt dt'. \quad (2.6.6)$$

Dans le cas particulier d'un opérateur stationnaire, avec la réponse impulsionnelle $\chi(t) \stackrel{(2.6.5)}{=} T(t, 0)$, on introduit la transformée de Fourier de la réponse impulsionnelle appelée **fonction de transfert** en fréquence :

$$H(f) := (\mathcal{F}\chi)(f) = \int e^{-i2\pi f t} \chi(t) dt \quad (2.6.7)$$

Voici un résultat très important en théorie du signal montrant l'intérêt de la transformée de Fourier

Théorème 2.6.12. *Pour un opérateur stationnaire T de fonction de transfert H , on a*

$$\langle e^{i2\pi f \cdot} | T e^{i2\pi f' \cdot} \rangle = H(f) \delta(f' - f)$$

autrement dit l'opérateur est diagonal dans la base de Fourier.

Démonstration. On observe que

$$\langle e^{i2\pi f \cdot} | T e^{i2\pi f' \cdot} \rangle \stackrel{(2.6.6)}{=} \int \int e^{-i2\pi f t} T(t, t') e^{i2\pi f' t'} dt dt' \stackrel{(2.6.4)}{=} \int \int e^{-i2\pi f t} \chi(t - t') e^{i2\pi f' t'} dt dt'$$

On effectue le changement de variable $t \rightarrow t'' = t - t'$, et on utilisera que $\int e^{i2\pi(f' - f)t'} dt' = \delta(f' - f)$,

$$\begin{aligned} \langle e^{i2\pi f \cdot} | T e^{i2\pi f' \cdot} \rangle &= \int \int e^{-i2\pi f(t'' + t')} \chi(t'') e^{i2\pi f' t'} dt'' dt' = \int \int e^{-i2\pi f t''} \chi(t'') e^{i2\pi(f' - f)t'} dt'' dt' \\ &\stackrel{(2.6.7)}{=} H(f) \delta(f' - f) \end{aligned}$$

□

Remarque 2.6.13. Le Théorème 2.6.12 a été utilisé la première fois par Joseph Fourier pour étudier l'équation de la chaleur, dont le générateur (sur \mathbb{R}^n) est l'opérateur Laplacien Δ qui est un opérateur stationnaire. Il a découvert la transformée de Fourier dans ce but.

Manifestation de la causalité La propriété de causalité de l'opérateur T qui est que $\chi(t) = 0$ si $t < 0$, implique que $f \rightarrow H(f)$ est holomorphe sur $\text{Im}(f) \leq 0$. En effet en posant $f = a + ib$ avec $a, b \in \mathbb{C}$, on a

$$H(f) \stackrel{(2.6.7)}{=} \int_0^{+\infty} e^{-i2\pi ft} \chi(t) dt = \int_0^{+\infty} e^{-i2\pi at} e^{2\pi bt} \chi(t) dt$$

qui est convergent pour tout $b \leq 0$. Voici une caractérisation complète de la propriété de causalité.

Théorème 2.6.14. (*Théorème de Paley-Wiener*). *Considérons un opérateur T stationnaire. T est causal, c'est à dire $\chi(t) = 0$ pour $t < 0$ si et seulement si $H \in H^2(\mathbb{C}_-) = L^2(\mathbb{C}_-) \cap \text{Hol}(\mathbb{C}_-)$ appelé *espace de Hardy* avec $\mathbb{C}_- := \{z \in \mathbb{C}, \text{Im}(z) \leq 0\}$.*

2.6.3.2 Exemple d'un filtre d'ordre 1 à temps continu

Exercice 2.6.15. Considérons le circuit électronique analogique suivant avec une résistance R et un condensateur C :



Le signal en entrée est la tension $u(t)$ et le signal en sortie est la tension $v(t)$. L'opérateur (ou filtre) qui nous intéresse est ici la transformation $T : u \rightarrow v$.

1. Montrer que v est solution de l'équation différentielle ordinaire d'ordre 1 (cad faisant intervenir la dérivée première de v)

$$\frac{dv}{dt} = -\frac{1}{\tau}v(t) + u(t) \quad (2.6.8)$$

avec $\tau = \frac{1}{RC}$.

2. En supposant $u(t') = 0$ et aussi $v(t') = 0$ pour tout $t' < t_0$ (avec t_0 date de départ donnée), montrer que la solution est

$$v(t) = \int_{-\infty}^t e^{(t'-t)/\tau} u(t') dt'. \quad (2.6.9)$$

Remarque 2.6.16. D'après cette expression, remarquer que $v(t)$ s'obtient comme une moyenne de $u(t')$ sur les temps t' antérieurs et un poids exponentiellement décroissant. Ce moyennage efface les fluctuations de u à petites périodes et ne conserve que les fluctuations de grandes périodes $T \gg \tau$ (ou basses fréquences $f = \frac{1}{T} \ll \frac{1}{\tau}$). Ainsi on appelle T un **filtre passe bas** de **fréquence de coupure** $f_c = \frac{1}{2\pi\tau}$.

3. Posant $v = Tu$, T est-il un opérateur linéaire? Quelle est l'expression du noyau de Schwartz (ou fonction de transfert) $T(t, t')$ de T ? T est-il un opérateur causal?
4. Calculer la **réponse impulsionnelle** du filtre $\chi(t)$, c'est à dire la solution v , si en entrée on impose une distribution de Dirac $u = \delta_0$?
5. Calculer la fonction de transfert en fréquence $H(f)$ du filtre et son module $|H(f)|$.

Solution 2.6.17.

1. Si $i(t)$ est le courant dans le circuit, on a $i = C \frac{dv}{dt}$ et $u = Ri + v$ donc

$$\begin{aligned} u &= RC \frac{dv}{dt} + v \\ \Leftrightarrow \frac{dv}{dt} &= -\frac{1}{\tau} v + u \end{aligned}$$

avec $\tau = \frac{1}{RC}$.

2. On peut utiliser la méthode de **de la variation de la constante** ou **principe de Duhamel**. L'équation homogène $\frac{dv}{dt} = -\frac{1}{\tau} v$ a pour solution $v(t) = v(0) e^{-t/\tau}$. On pose alors $v(t) = C(t) e^{-t/\tau}$. Alors

$$\begin{aligned} \frac{dv}{dt} &= -\frac{1}{\tau} v + u \Leftrightarrow \dot{C} e^{-t/\tau} - \frac{1}{\tau} C e^{-t/\tau} = -\frac{1}{\tau} C e^{-t/\tau} + u \\ \Leftrightarrow \dot{C} &= e^{t/\tau} u \end{aligned}$$

donc en supposant $u(t') = 0$ si $t' < t_0$ par exemple,

$$C(t) = \int_{-\infty}^t e^{t'/\tau} u(t') dt'$$

et

$$v(t) = C(t) e^{-t/\tau} = \int_{-\infty}^t e^{(t'-t)/\tau} u(t') dt'.$$

satisfait bien la condition limite $v(t) = 0$ si $t < t_0$.

3. On a bien $T(u_1 + u_2) = Tu_1 + Tu_2$ et $T(\lambda u) = \lambda u$ d'après la linéarité de l'intégrale donc T est bien un opérateur linéaire. D'après la définition (2.6.2), on écrit

$$v(t) = \int_{-\infty}^{+\infty} T(t, t') u(t') dt'$$

avec le noyau de Schwartz

$$T(t, t') = \begin{cases} e^{(t'-t)/\tau} & \text{si } t' \leq t \\ 0 & \text{si } t' > t \end{cases}$$

Donc T est bien un opérateur causal.

4. Si $u = \delta_0$, alors

$$\chi(t) = \int_{-\infty}^{+\infty} T(t, t') \delta_0(t') dt' = T(t, 0) = \begin{cases} e^{-t/\tau} & \text{si } t \geq 0 \\ 0 & \text{si } t < 0 \end{cases} \quad (2.6.10)$$

appelée réponse impulsionnelle du filtre.

5. La fonction de Transfert est la transformée de Fourier de χ :

$$\begin{aligned} H(f) &= \int e^{-i2\pi ft} \chi(t) dt \stackrel{(2.6.10)}{=} \int_0^{+\infty} e^{-i2\pi ft} e^{-t/\tau} dt \\ &= \int_0^{+\infty} e^{(-i2\pi f - \frac{1}{\tau})t} dt = \frac{1}{(-i2\pi f - \frac{1}{\tau})} \left[e^{(-i2\pi f - \frac{1}{\tau})t} \right]_0^{\infty} \\ &= \frac{1}{i2\pi f + \frac{1}{\tau}} \end{aligned}$$

Alors

$$|H(f)|^2 = \frac{1}{\frac{1}{\tau^2} + (2\pi f)^2} = \frac{\tau^2}{1 + \left(2\pi \frac{f}{(1/\tau)}\right)^2}$$

est une **Lorenzienne** centrée en $f = 0$ de largeur en fréquence $f_c = \frac{1}{\tau}$. La fonction

$$f \in \mathbb{C} \rightarrow H(f) = \frac{1}{i2\pi f + \frac{1}{\tau}} = \frac{1}{i2\pi \left(f - \frac{i}{2\pi\tau}\right)}$$

est méromorphe sur \mathbb{C} et admet un pôle en $f = \frac{i}{2\pi\tau}$.

Remarque 2.6.18. En théorie du signal on dit que T est un **filtre linéaire d'ordre 1** car v est solution d'une l'équation différentielle d'ordre 1, (2.6.8).

2.6.3.3 Filtre numérique, temps discret, signaux échantillonnés $(u_j)_j$

Si $v = (v_0, v_1, \dots, v_{N-1}) \in \mathbb{R}^N$ est un signal échantillonné, avec pas de temps $0 < \delta < \tau$, on définit la dérivée discrète par

$$v'_j := \frac{1}{\delta} (v_{j+1} - v_j), \quad j = 0 \dots N - 2$$

Remarque 2.6.19. C'est en fait le taux d'accroissement de la fonction $v(t)$ sur un intervalle de temps δ appelé aussi **différence finie**. On aurait aussi pu définir $v'_j := \frac{1}{\delta} (v_j - v_{j-1})$ ou $v'_j := \frac{1}{2\delta} (v_{j+1} - v_{j-1})$ qui sont des définitions différentes mais deviennent équivalentes dans la limite des « variations lentes » de v .

L'analogue de l'équation différentielle ordinaire (2.6.8) est l'équation suivante appelée équation aux différences finies ;

$$\begin{aligned} \frac{1}{\delta} (v_{j+1} - v_j) &= -\frac{1}{\tau} v_j + u_j, \quad \forall j, \\ \Leftrightarrow v_{j+1} &= v_j + \delta \left(-\frac{1}{\tau} v_j + u_j \right), \quad \forall j \\ &= v_j \left(1 - \frac{\delta}{\tau} \right) + \delta u_j \end{aligned} \tag{2.6.11}$$

Cette dernière formule permet de construire le signal v à partir du signal u par récurrence sur j .

Exercice 2.6.20. (TP) **”Filtre discret d'ordre 1”**. Utilisant la formule (2.6.11), programmer en python une fonction qui en entrée prend le signal $u = (u_0, u_1, \dots, u_{N-1})$ et en sortie renvoie le signal $v = (v_0, v_1, \dots, v_{N-1})$ défini par (2.6.11). Appliquer ce filtre à un signal sonore (pour une fréquence de coupure f_c bien choisie) et écouter.

Exercice 2.6.21. (*) Exprimer la solution de l'équation 2.6.11, c'est à dire exprimer v_j à partir du signal en entrée $(u_j)_j$. Retrouver (2.6.9) dans la limite $\delta \rightarrow 0$. Commenter.

Solution 2.6.22.

1. On a

$$\begin{aligned} v_j &= v_{j-1} \left(1 - \frac{\delta}{\tau} \right) + \delta u_{j-1} \\ v_{j-1} &= v_{j-2} \left(1 - \frac{\delta}{\tau} \right) + \delta u_{j-2} \end{aligned}$$

etc donc

$$\begin{aligned} v_j &= \delta u_{j-1} + \delta \left(1 - \frac{\delta}{\tau}\right) u_{j-2} + \delta \left(1 - \frac{\delta}{\tau}\right)^2 u_{j-3} + \dots \\ &= \delta \sum_{k \geq 0} \left(1 - \frac{\delta}{\tau}\right)^k u_{j-k-1} \end{aligned}$$

qui est une série convergente si $0 < \delta < \tau \Leftrightarrow 0 < \left(1 - \frac{\delta}{\tau}\right) < 1$.

2. Dans la limite $\frac{\delta}{\tau} \ll 1$, vérifions que l'on retrouve l'expression en temps continu (2.6.9). On écrit

$$\begin{aligned} \left(1 - \frac{\delta}{\tau}\right)^k &= e^{k \ln\left(1 - \frac{\delta}{\tau}\right)} \\ \ln\left(1 - \frac{\delta}{\tau}\right) &= -\frac{\delta}{\tau} + O\left(\left(\frac{\delta}{\tau}\right)^2\right) \\ v_j &= \delta \sum_{k \geq 0} \left(1 - \frac{\delta}{\tau}\right)^k u_{j-k-1} = \delta \sum_{k \geq 0} e^{k\left(-\frac{\delta}{\tau} + O\left(\left(\frac{\delta}{\tau}\right)^2\right)\right)} u_{j-k-1} \end{aligned}$$

Posant $t_k = \delta k$, et $v_j = v(t_j)$

$$\begin{aligned} v(t_j) &= \delta \sum_{k \geq 0} e^{\frac{t_k}{\delta}\left(-\frac{\delta}{\tau} + O\left(\left(\frac{\delta}{\tau}\right)^2\right)\right)} u(t_{j-k-1}) \\ &= \delta \sum_{k \geq 0} e^{-t_k/\tau} e^{t_k O\left(\frac{\delta}{\tau^2}\right)} u(t_{j-k-1}) \end{aligned}$$

Or $e^{t_k O\left(\frac{\delta}{\tau^2}\right)} = 1 + t_k O\left(\frac{\delta}{\tau^2}\right)$ et $\delta \sum_{k \geq 0} w(t_k) = \int w(t') dt' (1 + O(\delta))$ donc

$$v(t_j) = \delta \sum_{k \geq 0} e^{-t_k/\tau} \left(1 + t_k O\left(\frac{\delta}{\tau^2}\right)\right) u(t_{j-k-1})$$

Le passage à la limite $\delta \rightarrow 0$ donne

$$v(t) = \int_0^\infty e^{-t'/\tau} u(t - t') dt'$$

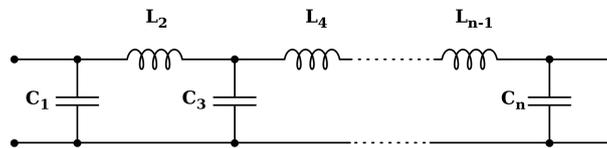
qui est (2.6.9).

Remarque 2.6.23. C'est un cas particulier de ce qui s'appelle **Filtre à réponse impulsionnelle infinie** ou **IIR, Infinite_impulse_response**. En théorie du signal numérique ce sont des filtres très utilisés car ils sont inspirés des phénomènes physique (comme ici un circuit électronique ou filtre analogique) et « sonnent » donc réalistes ou naturels. Le filtre précédent est d'ordre 1, car v_{j+1} dépend de u_j, v_j . Dans un filtre d'ordre n , v_j dépend de v_{j-k}, u_{j-k} avec $k = 1, \dots, n$.

Exercice 2.6.24. (TP) « **perception des voyelles et du langage** » Dans Audacity, charger le fichier **Voyelles_Par_Malik.wav**, l'écouter. Avec Menu/Analyse/Tracer le spectre et choisissant l'option axe fréquence logarithmique, vérifier que la fréquence des voyelles émise est $f_1 = 160\text{Hz}$, soit C_4 . Quelles sont les fréquences de harmoniques f_2, f_3, f_4 ?

1. Avec Menu/Effets/Filtres classiques.. appliquer un **filtre Butterworth passe bas**, d'ordre 6, et de fréquence de coupure $f_c = 500\text{Hz}$. Tracer le spectre et vérifier que les harmoniques hautes sont coupées. Écouter. Reprenez le signal de départ et appliquer le même filtre **passe haut**. Que remarquez vous ?
2. Essayez de même avec votre propre voix enregistrée. Conclusion où se situe l'information perçue sur les voyelles ?

Exercice 2.6.25. (*) Le schéma électronique du **filtre Butterworth** d'ordre n est le suivant.



Ecrire l'équation différentielle ordinaire associée (analogue à (2.6.9)) et écrire l'équation aux différences finies analogue à (2.6.11). Ecrire une fonction en python pour ce filtre.