

MASTERE DE MATHEMATIQUES - M1

UNIVERSITE PARIS VI - PIERRE ET MARIE CURIE

Année Scolaire 2004-2005

*De l'analyse de Fourier traditionnelle aux
ondelettes : une approche par le signal musical*

Sonia Hamnane*

Travail encadré de recherche réalisé sous la direction de Laurent Mazliak

Université Pierre et Marie Curie.

*sonia.hamnane@wanadoo.fr

Je voudrais exprimer mes remerciements à Laurent Mazliak pour avoir accepté d'encadrer mon TER, pour le temps qu'il m'a dédié ainsi que pour ses nombreux conseils et son aide. Mes remerciements vont aussi à Nathalie Delprat pour ses explications précieuses.

Sommaire

1. Introduction.

2. Analyse de Fourier : exemple du signal acoustique, musical.

2.1 Les sons musicaux : généralités relatives au signal sonore, musical.

2.2 Le signal.

2.3 Représentation d'un signal par sa série de Fourier.

2.4 Transformée de Fourier.

2.5 Limites de la transformée de Fourier standard.

3. La transformée de Fourier à fenêtre : exemple du signal acoustique, musical.

3.1 Idée de fenêtre.

3.2 Les formules de Gabor.

3.3 Comparaison Fourier-Gabor et limites de la transformée de Gabor.

4. Les ondelettes

4.1 Idée de base.

4.2 La transformée en ondelettes.

4.3 Exemples.

4.4 Bases orthonormée d'ondelettes et analyse multirésolution.

5. Conclusion.

1 Introduction

La musique définie simplement comme l'art d'organiser les sons, est un objet d'étude pour les sciences, en particulier pour les mathématiques, dans la mesure où celles-ci permettent de comprendre la diversité et la complexité des sons musicaux. Il est alors important de connaître et de comprendre les différents outils mathématiques qui permettent d'organiser, d'analyser, de synthétiser et de modifier les sons musicaux ou tout simplement de les comprendre intimement.

Musique et mathématiques tissent depuis la plus haute antiquité jusqu'à l'époque la plus contemporaine, des liens privilégiés. Mais les relations entre les deux disciplines se modifient considérablement autour des 16ème et 17ème siècles sous l'impulsion de musiciens, théoriciens mais aussi de savants et de philosophes. Ainsi la théorie musicale est repensée radicalement en fonction de nouveaux concepts scientifiques. Penser la musique avec les mathématiques est ainsi une longue histoire où interviennent tour à tour arithmétique, géométrie, algèbre et analyse.

Il s'agira dans notre étude de s'intéresser aux outils mathématiques nécessaires à l'analyse et à la synthèse des sons musicaux. Faisons un bref parcours historique de ces outils depuis le 18ème siècle jusqu'à nos jours, parcours qui nous servira dans le même temps à annoncer la structure de ce travail, puisqu'il respecte un ordre tout autant logique que chronologique. Il faut préciser avant tout que les outils utilisés sont empruntés à la théorie du signal, dans laquelle il s'agit, en particulier, d'extraire d'un signal les informations pertinentes. Les applications industrielles de l'analyse et du traitement des signaux sont nombreuses et variées allant de la géophysique à l'imagerie médicale. Mais il s'agit ici d'aborder le sujet du point de vue de l'analyse sonore.

C'est au 18ème siècle que l'équation des ondes est établie et les travaux de Fourier peu de temps après sont à l'origine de nombreuses et primordiales questions physico-mathématiques. L'analyse de Fourier donne en particulier une impulsion à l'étude des sons musicaux grâce aux outils mathématiques développés. Pendant longtemps les fonctions de base de l'analyse ont été le cosinus, le sinus et l'exponentielle imaginaire. Ainsi les fonctions $\frac{1}{\sqrt{2\pi}}e^{ikx}$ constituent une base orthonormée de l'espace de référence $L^2(0, 2\pi)$ et les séries de Fourier sont les combinaisons linéaires $\sum a_k e^{ikx}$. Mais l'analyse de Fourier présente des inconvénients majeurs qui ne permettent pas une analyse satisfaisante de toute sorte de signaux. En effet, on souhaiterait que

les paramètres réalisent une analyse à la fois en temps et en fréquence, exactement comme à la manière d'une portée musicale où sont indiquées à la fois la fréquence et la durée des notes. C'est pourquoi l'analogie avec la notation musicale est communément utilisée pour introduire l'idée de représentation temps-fréquence, chaque note d'un morceau étant en effet associée à une certaine localisation aussi bien temporelle (son instant d'occurrence et sa durée) que fréquentielle (sa hauteur). On rappellera donc dans un premier temps l'analyse de Fourier en vue d'une application au signal musical et nous verrons dans quelle mesure il est possible de représenter ce signal par sa série de Fourier. Les transformations de Fourier et l'analyse spectrale sont en effet des outils incoutournables du traitement des signaux harmoniques. L'analyse de Fourier est la plus ancienne de ces techniques, elle est performante en particulier grâce à la transformée de Fourier rapide. Les transformations reposent sur un concept de fréquence qui est exclusif de toute évolution temporelle. Ceci ne permettant pas de donner commodément sens à la notion, pourtant intuitive, de fréquence instantanée. Cette difficulté, nous le verrons, est liée de façon fondamentale au fait que les variables de temps et de fréquences sont canoniquement conjuguées et leur représentation associée, sujette à des inégalités de type incertitude.

Par la suite, une approche intuitive est utilisée sur la base de transformée de Fourier glissante ou à fenêtre. Apparaît ainsi la transformée de Gabor, que nous étudierons dans un deuxième temps. Le physicien Denis Gabor l'a proposée en 1940 et on lui doit la formule de reconstruction. Mais cette transformée présente encore des inconvénients et il sera nécessaire d'affiner à nouveau cet outil.

Au début des années 80, de nombreux scientifiques utilisaient déjà les ondelettes comme une alternative à l'analyse de Fourier traditionnelle. En particulier pour le traitement numérique des signaux acoustique (paroles ou musique) mais également pour interpréter les signaux sismiques recueillis au cours de campagnes pétrolières. Les ondelettes n'ont depuis lors cessées de se développer et de trouver de nouveaux champs d'application. C'est ainsi qu'est apparu un parallèle étonnant entre ces méthodes et des techniques développées à des fins totalement différentes. Ces différentes ondelettes ont été regroupées en une thèse cohérente mathématiquement fondée et d'un emploi universel. Les bases orthonormées d'ondelettes que nous étudierons également avec l'analyse multirésolution ne sont apparues que récemment mais des constructions similaires étaient déjà utilisées en mathématiques, physique théorique et en traitement du signal. C'est ce que nous verrons dans un troisième et dernier temps.

2 Analyse de Fourier : exemple du signal acoustique, musical

2.1 Les sons musicaux : généralités relatives au signal sonore, musical

Un son est un ébranlement élastique de l'air, d'un fluide ou d'un solide qui se manifeste par des variations de pressions autour de la pression moyenne du milieu. Lorsque le milieu est homogène, l'onde se propage à vitesse constante appelée célérité. Les ondes sont dites longitudinales.

Un bruit est un ensemble de sons à caractère aléatoire qui ne renseigne l'auditeur que sur la nature de la source qui l'a produit.

Les sons formés par le langage parlé ou la musique forment des éléments d'une phrase verbale ou musicale qui présentent un sens pour l'auditeur. Des informations sont ainsi véhiculées par voie acoustique. Les phonèmes ou les notes constituent l'alphabet de ces langages.

• **Allure temporelle d'une note** : il y a 4 phases successives :

- la phase d'attaque (attack)
- la phase de déclin (decay)
- la phase de maintien (sustain)
- la phase de chute (release)

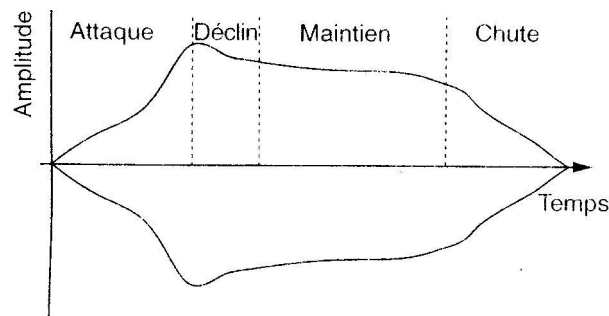


Figure 1 : Evolution temporelle de l'enveloppe d'une note musicale

La forme globale est l'enveloppe de la note.

Dans la phase de maintien (dont la durée est contrôlée par l'instrumentiste) le signal est quasi-périodique. Mais la forme du signal sur une période est une caractéristique propre à chaque instrument. Nous donnons ci-dessous un exemple d'ondes émises par deux instruments.

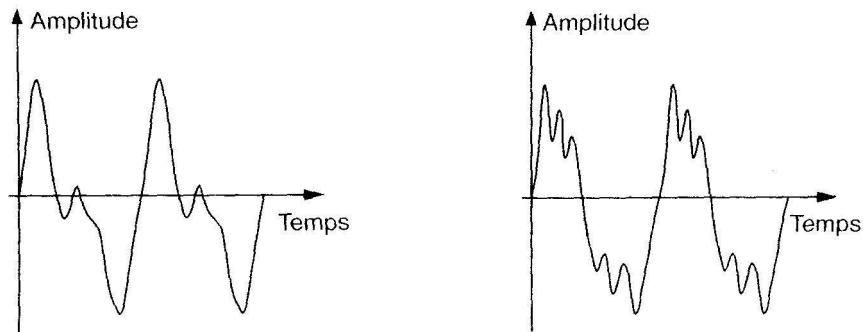


Figure 2 : Exemple d'ondes émises par deux instruments. (Allure temporelle des ondes de pression émises par une clarinette et un piano).

- **Caractéristiques d'un son :**

La hauteur d'un son, qualité qui fait distinguer un son grave d'un son aigu, correspond à la fréquence fondamentale de la forme d'onde dans la phase de maintien. C'est le nombre de vibrations acoustiques produites par l'instrument pendant une seconde. Les sons aigus sont dus au mouvement vibratoire de fréquence élevée et les sons graves au mouvement de fréquence basse. Toutefois l'oreille humaine ne peut percevoir que les sons dont les fréquences sont comprises entre 20 et 20000 Hz pour une ouïe très fine, soit environ une dizaine d'octaves.

L'intensité d'un son est la qualité qui fait distinguer un son fort d'un son faible. Elle est liée à l'amplitude des vibrations sonores. Toutefois, l'oreille ne peut percevoir un son que si son amplitude a une valeur minimale.

Le timbre est la caractéristique qui permet de distinguer deux sons émis par des instruments différents. Il est directement lié à la forme du signal pendant une période. Si le son est musical au sens acoustique du terme, c'est-à-dire créé par un mouvement vibratoire périodique, le son peut être considéré comme la superposition de sons simples harmoniques dont les fréquences sont des multiples entiers de la fréquence d'un son de base, appelé fondamentale. Le timbre d'un son dépend des intensités relatives des différents sons simples harmoniques qui le composent.

- **Analyse et synthèse harmonique :**

On peut analyser mathématiquement le contenu harmonique d'un signal périodique quelconque en utilisant le développement en série de Fourier.

Réciproquement, il est possible de synthétiser des formes d'ondes sonores complexes en ajoutant diverses composantes sinusoïdales en rapport harmonique.

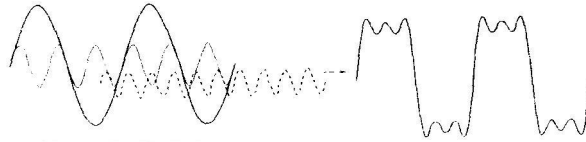


Figure 3 : Exemple de la forme d'onde résultante de 3 sinus

• **Composante fréquentielle des différentes phases**

Durant les phases d'attaque et de chute, l'enveloppe du signal n'est pas constante. Les variations d'amplitude modifient le contenu spectral du son, notamment si ces variations sont brusques. Etudions ce phénomène dans le cas d'un son sinusoïdal pur :

Dans la phase de maintien : la fréquence est fixe et précise.

Phase d'attaque et de chute : si elles sont douces, on perçoit une fréquence unique qui apparaît puis s'évanouit. Si elles sont brutales, il est possible de percevoir un claquement dû à l'élargissement spectral provoqué par la coupure temporelle du signal.

On peut ainsi voir sur la seconde image qu'un signal tronqué (correspondant à une phase de chute brutale) provoque un élargissement spectral.

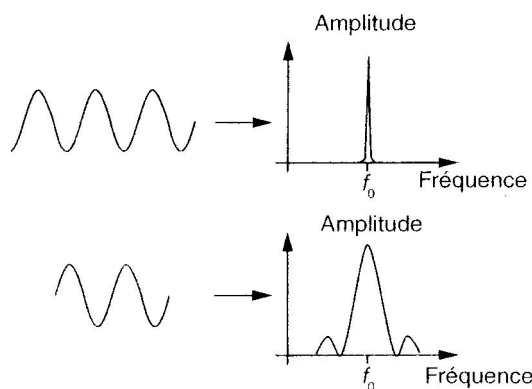


Figure 4 : Elargissement spectral d'un signal tronqué.

2.2 Le signal

La notion de signal est très extensive. Elle ressort de l'observation d'un phénomène, certaines quantités qui dépendent du temps, de l'espace, d'une fréquence ou autre. Ces quantités mesurables sont appelées signaux. On choisit de représenter le signal par la notion de fonction en mathématiques.

On modélisera le signal de façon déterministe, mais on pourrait également le modéliser de façon aléatoire.

Le signal est analogique ($x = x(t)$) si la variable est continue.

Le signal est discret ($x = (x_n)_{n \in \mathbb{N}}$) si la variable est discrète.

Un signal discret est le plus souvent le résultat d'une discrétisation ou d'un échantillonnage d'un signal analogique.

Les valeurs du signal sont considérées comme des valeurs exactes réelles ou complexes.

On appelle quantification l'approximation des valeurs exactes par un nombre fini de données (à stocker sur ordinateur).

Un signal discret et quantifié est appelé signal numérique.

• Exemple d'un signal sinusoïdal pur (ou monochromatique)

Signal de la forme $x(t) = \alpha \cos(\omega t + \varphi)$.

$|\alpha| = \max |x(t)|$ est l'amplitude du signal. ω est la pulsation.

$a = \frac{2\pi}{\omega}$ est la plus petite période. $\lambda = \frac{1}{a}$ est la fréquence. φ est la phase initiale.

Pour des raisons de commodité, on utilise couramment une fonction à valeurs complexe

$$z(t) = \alpha e^{i(\omega t + \varphi)} \quad \text{et} \quad x(t) = \operatorname{Re}(z(t)) = \frac{1}{2}(z(t) + \bar{z}(t)).$$

Remarque 2.1 *Cette écriture n'a pas de sens physique à cause des fréquences négatives.*

Soit $z(t) = ce^{2i\pi\lambda t}$, $\lambda \in \mathbb{R}$ et $c \in \mathbb{C}$. On a adopté la représentation en module et argument dans l'espace des fréquences, avec $c = |c|e^{i\varphi}$.

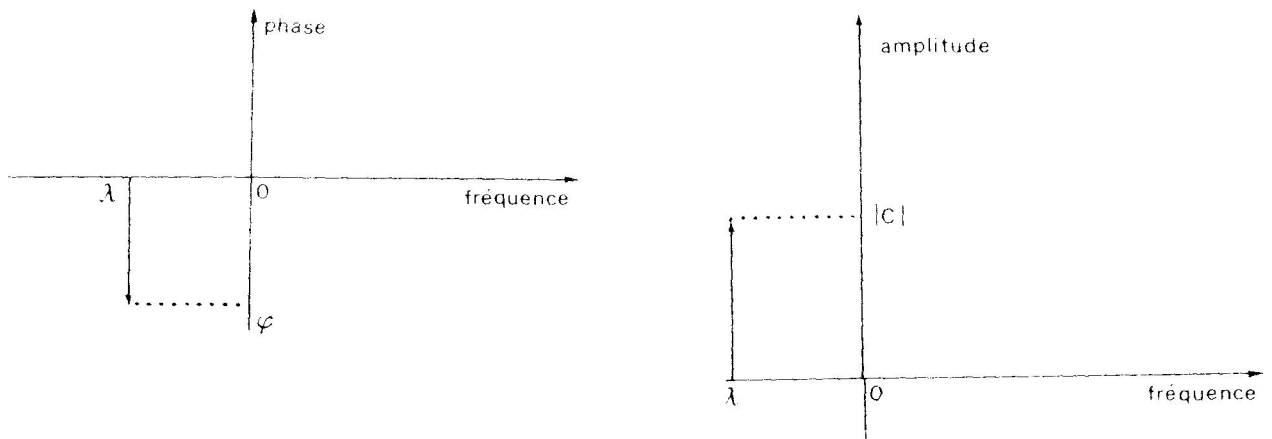


Figure 5 : Représentation en fréquence d'un signal sinusoïdal

Plus généralement,

- **Les signaux trigonométriques**

(les signaux sinusoïdaux entrent dans la structure de tous les signaux périodiques).

On note $e_n(t) = e^{2i\pi n \frac{t}{a}}$. Cette fonction est de période a , il en est donc de même pour tout fonction ρ de la forme

$$\rho(t) = \sum_{n \in I} c_n e^{2i\pi n \frac{t}{a}}$$

où $c_n \in \mathbb{C}$ et où I est un ensemble fini d'entiers fixés. On peut donc supposer que ρ est écrit sous la forme

$$\rho(t) = \sum_{n=-N}^{n=+N} c_n e^{2i\pi n \frac{t}{a}}.$$

On obtient un polynôme trigonométrique de degré inférieur ou égal à N .

On pourra ainsi modéliser les signaux formés d'un nombre fini de signaux monochromatiques.

- **Représentation à l'aide des fonctions trigonométriques**

L'expression $\rho(t) = \sum_{n=-N}^{n=+N} c_n e^{2i\pi n \frac{t}{a}}$ peut être écrite sous forme de combinaison linéaire de sinus et cosinus

$$\rho(t) = \frac{1}{2}a_0 + \sum_{n=-N}^{n=+N} \left(a_n \cos 2\pi n \frac{t}{a} + b_n \sin 2\pi n \frac{t}{a} \right)$$

où $a_n = c_n + c_{-n}$ et $b_n = i(c_n - c_{-n})$.

- **Une base : propriétés d'orthogonalité de la famille des (e_n)**

On note \mathcal{T}_N l'ensemble des polynômes trigonométriques de degré inférieur ou égal à N définis par $p(t) = \sum_{n=-N}^{n=+N} c_n e^{2i\pi n \frac{t}{a}}$ quand les c_n varient. On munit cet espace vectoriel du produit scalaire dans $L_p^2(0, \alpha)$

$$(p, q) = \int_0^\alpha p(t) \bar{q}(t) dt.$$

On a bien sûr

$$(e_n, e_m) = a\delta_{n,m}.$$

Cela exprime le fait que les (e_n) sont orthogonales. Ces fonctions forment donc une famille libre, il est par ailleurs clair qu'elle est génératrice. C'est donc une base de l'espace \mathcal{T}_N qui est par conséquent de dimension $2N + 1$.

- **Expression explicite des coefficients**

Nous avons donc $(e_n, e_m) = 0$ si $n \neq m$ et $\|e_n\|_2 = \sqrt{a}$. Alors $(p, e_n) = ac_n$. Donc (formule de Fourier)

$$c_n = \frac{1}{a} \int_0^a p(t) e^{-2i\pi n \frac{t}{a}} dt,$$

ce qui nous donne une expression explicite des coefficients c_n en fonction de p . On a bien sûr une version réelle de la décomposition précédente avec

$$a_n = \frac{2}{a} \int_0^a p(t) \cos \left(2\pi n \frac{t}{a} \right) dt, \quad b_n = \frac{2}{a} \int_0^a p(t) \sin \left(2\pi n \frac{t}{a} \right) dt.$$

Nous savons à présent résoudre le problème de l'analyse spectrale d'un signal trigonométrique.

Remarque 2.2 L'intégrale $c_n = \frac{1}{a} \int_0^a p(t) e^{-2i\pi n \frac{t}{a}} dt$ peut en fait être prise sur tout intervalle de longueur a . Par exemple $[-\frac{a}{2}, \frac{a}{2}]$. On a alors dans ce cas les propriétés suivantes :

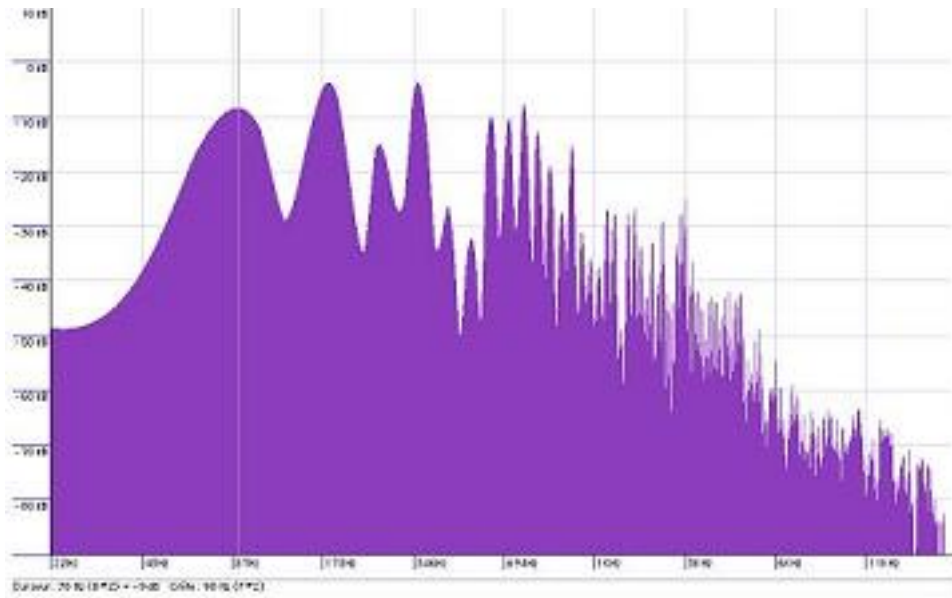
si p est paire alors pour tout n , $c_n = c_{-n}$ et donc pour tout n , $b_n = 0$,
 si p est impaire alors pour tout n , $c_n = -c_{-n}$ et donc pour tout n , $a_n = 0$.

De plus on a $\|p\|_2^2 = \sum_{n=-N}^{n=+N} \sum_{m=-N}^{m=+N} c_n \bar{c}_m (e_n, e_m)$ d'où l'égalité de Parseval pour les polynômes trigonométriques

$$\sum_{n=-N}^{n=+N} |c_n|^2 = \frac{1}{a} \int_0^a |p(t)|^2 dt.$$

2.3 Représentation d'un signal par sa série de Fourier

Figure 6 : Spectre d'instrument : Fa dièse 2 violoncelle



On peut observer la continuité des fréquences. On remarquera un net pic au niveau de la fréquence fondamentale et d'autres aux multiples entiers de cette fréquence.

On a donc une fonction $f : \mathbb{R} \rightarrow \mathbb{C}$ quelconque mais périodique de période a . Peut-on lui associer une décomposition de la forme

$$f(t) = \sum_{n=0}^{+\infty} c_n e^{2i\pi n \frac{t}{a}} ?$$

On peut aborder ce problème d'un point de vue géométrique.

• **Le cadre de l'espace** $L_p^2(0, a)$

Tout d'abord nous ne nous intéressons qu'à des fonctions périodiques par exemple de période a .

D'autre part, nous aurons besoin du fait que $\int_0^a |f(t)|^2 dt$ existe et est finie.

On note

$$L_p^2(0, a) = \{f \in L^2(\mathbb{R} \rightarrow \mathbb{C}) \text{ et périodique de période } a\}.$$

Il s'agit d'un espace vectoriel complexe que l'on munit du produit scalaire

$$(f, g) = \int_0^a f(t)\bar{g}(t)dt$$

et de la norme associée $\|f\|_2 = \sqrt{(f, f)}$.

• **Idée de projection orthogonale sur un sous-espace**

L'égalité $f(t) = \sum_{n=-N}^{+N} c_n e^{2i\pi n \frac{t}{a}}$ ne peut avoir lieu pour toute fonction (le second membre est indéfiniment dérivable) mais nous pouvons tout de même essayer d'approcher f par une somme finie. Le problème devient donc : étant donné N entier naturel fixé, peut on trouver une famille (x_n) telle que

$$\|f - \sum_{n=-N}^{n=+N} x_n e_n\|_2$$

soit minimale ?

Géométriquement il s'agit de trouver un élément f_N du sous-espace \mathcal{T}_N qui soit à une distance minimale de f .

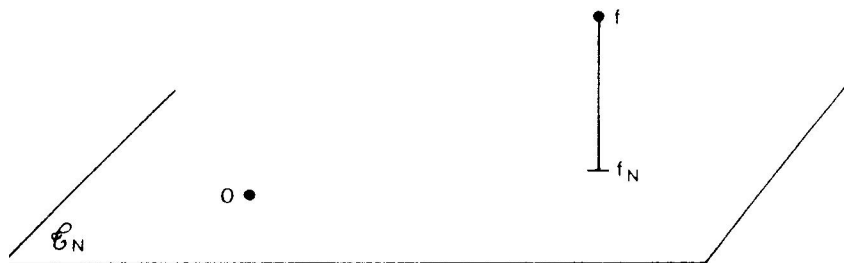


Figure 7 : Projection orthogonale sur un sous espace.

• Résolution du problème d'approximation

On évalue la distance de f à un polynôme trigonométrique quelconque $p = \sum_{n=-N}^{n=+N} x_n e_n$.

On a

$$\|f - p\|_2^2 = \|f\|_2^2 - 2\text{Re}(f, p) + \|p\|_2^2.$$

Or d'après l'égalité de Parseval pour les polynômes trigonométriques, on a

$$\sum_{n=-N}^{n=+N} |c_n|^2 = \frac{1}{a} \int_0^a |p(t)|^2 dt$$

et donc

$$\|p\|_2^2 = a \sum_{n=-N}^{n=+N} |c_n|^2$$

et

$$(f, p) = \sum_{n=-N}^{n=+N} \bar{x}_n (f, e_n).$$

On pose $c_n := c_n(f) = \frac{1}{a} (f, e_n)$. On a alors

$$\|f - p\|_2^2 = \|f\|_2^2 + a \sum_{n=-N}^{n=+N} (|c_n - x_n|^2 - |c_n|^2),$$

le minimum étant donc atteint lorsque $x_n = c_n$ et seulement pour cette valeur. On peut donc en conclure que la meilleure approximation f_N de f existe et est unique ; c'est

$$f_N(t) = \sum_{n=-N}^{n=+N} c_n e_n(t).$$

On peut donc énoncer le

Théorème 2.3 *Il existe un unique polynôme trigonométrique f_N dans \mathcal{T}_N tel que $\|f - f_N\|_2 = \min \{\|f - p\|_2, p \in \mathcal{T}_N\}$, c'est $f_N = \sum_{n=-N}^{n=+N} c_n e_n$ avec*

$$(1) \quad c_n = \frac{1}{a} \int_0^a f(t) e^{-2i\pi n \frac{t}{a}} dt$$

• **Convergence de l'approximation**

On se demande si on a bien $f_N \xrightarrow{N \rightarrow +\infty} f$.

On prend par exemple l'approximation du signal carré ressemblant à l'onde produite par une clarinette.

On a

$$(1) \quad \begin{cases} f(t) = 1 & \text{si } 0 \leq t \leq \pi; \\ f(t) = -1 & \text{si } \pi \leq t \leq 2\pi. \end{cases}$$

Nous obtenons les coefficients de Fourier suivant :

$$a_m = \frac{1}{\pi} \left(\int_0^\pi \cos(m\theta) d\theta - \int_\pi^{2\pi} \cos(m\theta) d\theta \right) = 0$$

car f est impaire,

$$b_m = \frac{1}{\pi} \left(\int_0^\pi \sin(m\theta) d\theta - \int_\pi^{2\pi} \sin(m\theta) d\theta \right) = \begin{cases} \frac{4}{m\pi} & \text{si } m \text{ est impair;} \\ 0 & \text{si } m \text{ est pair.} \end{cases}$$

On a donc pour f le développement en série de Fourier

$$f(\theta) = \frac{4}{\pi} \left(\sin \theta + \frac{1}{3} \sin 3\theta + \frac{1}{5} \sin 5\theta + \dots \right).$$

On a par exemple $f_1(\theta) = \frac{4}{\pi} \sin \theta$, $f_3(\theta) = \frac{4}{\pi} \left(\sin \theta + \frac{1}{3} \sin 3\theta \right)$ et $f_5(\theta) = \frac{4}{\pi} \left(\sin \theta + \frac{1}{3} \sin 3\theta + \frac{1}{5} \sin 5\theta \right)$ qu'on représente ci-après :

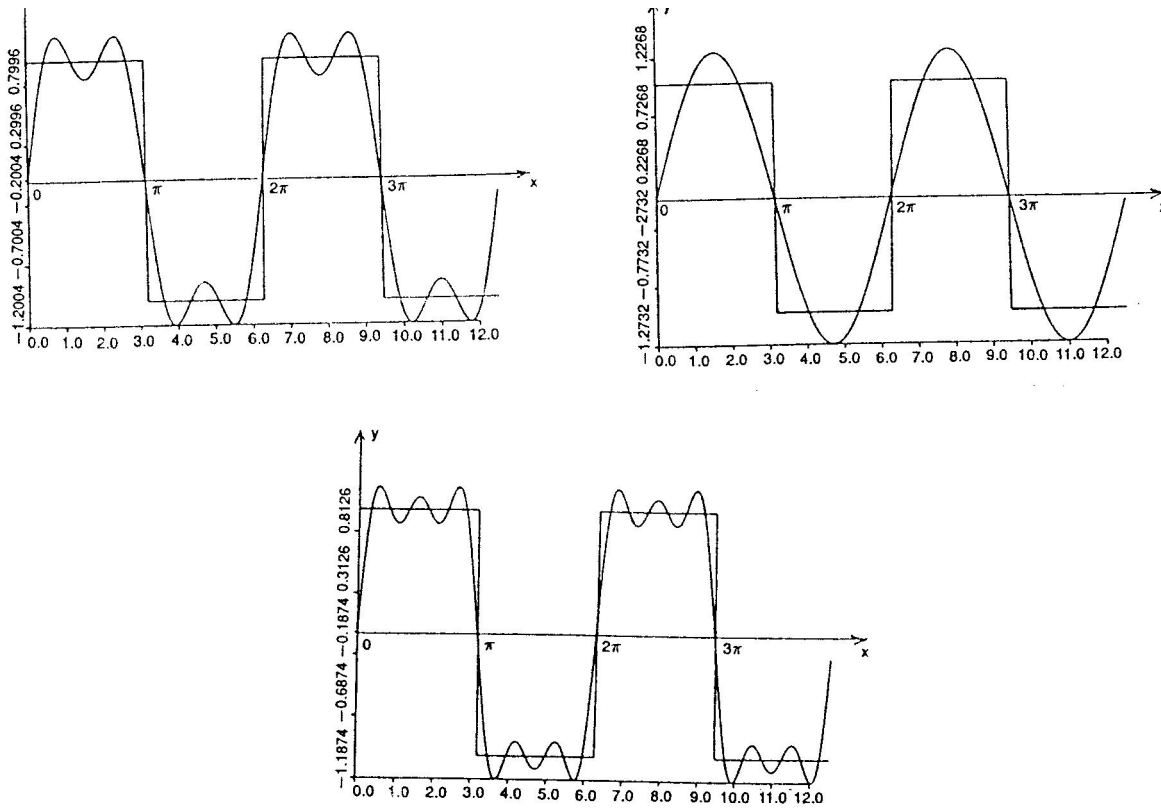


Figure 8

De manière plus générale nous avons le théorème suivant

Théorème 2.4 Si $f \in L^2_p(0, a)$, alors la meilleure approximation f_N de f dans \mathcal{T}_N

$$f_N(t) = \sum_{n=-N}^{+N} c_n e^{2i\pi n \frac{t}{a}} \xrightarrow{N \rightarrow +\infty} f$$

dans $L^2_p(0, a)$ (les c_n étant définis dans la formule (1)). Autrement dit

$$\lim_{N \rightarrow +\infty} \|f - f_N\|_2 = 0.$$

• Notion de spectre

Définition 2.5 Pour un signal f de période a et développé en série de Fourier sous la forme

$$f(t) = \sum_{n \in \mathbb{Z}} c_n e^{2i\pi n \frac{t}{a}},$$

on appelle spectre de f l'ensemble des couples $(\frac{n}{a}, c_n)_{n \in \mathbb{Z}}$.

(cette définition est tirée du livre "Analyse de Fourier et applications" de Gasquet et Witomski, page 49)

Nous avons déjà vu précédemment une représentation en fréquence (en amplitude et en phase) d'un signal sinusoïdal pur. De la même façon, tout signal périodique admet une représentation en fréquence en spectre d'amplitude et en spectre de phase. On les représente grâce à des flèches parallèles à l'axe des ordonnées.

On peut voir un exemple ci-dessous. On y voit que le spectre est constitué de raies espacées de la fréquence $\frac{1}{a}$. Pour $|n| = 1$ les deux raies correspondent à la fréquence fondamentale (la note que l'on entend) les autres raies correspondant aux harmoniques du signal (multiples de la fréquence fondamentale).

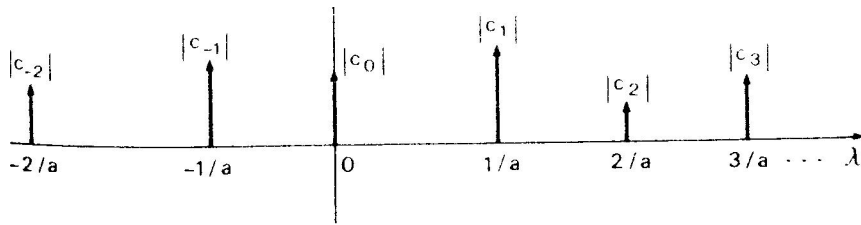


Figure 9 : Spectre d'amplitude.

On représente ici le spectre de phase : ensemble des couples de la forme $(\frac{n}{a}, \theta_n)$ $n \in \mathbb{Z}$ avec $c_n = |c_n|e^{i\theta_n}$ et $\theta_n \in [-\pi, \pi[$.

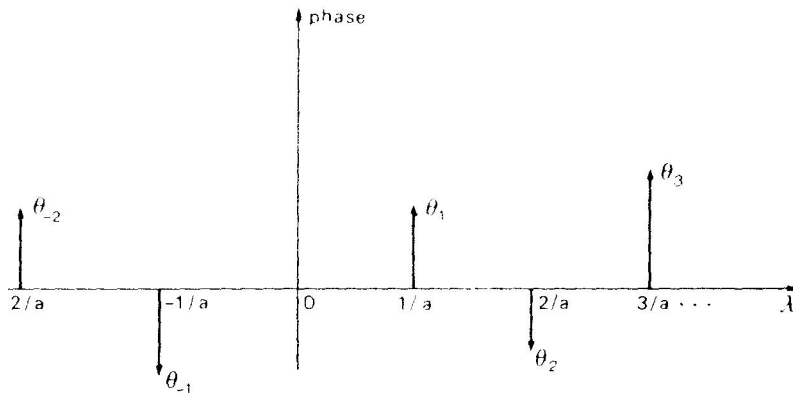


Figure 10 : Spectre de phase.

L'analyse fréquentielle des signaux apporte donc des informations importantes.

Remarque 2.6 *Selon les caractéristiques du signal étudié (c'est-à-dire s'il est continu, discret, périodique ou transitoire) la représentation fréquentielle possède des propriétés particulières. De plus les méthodes utilisées pour calculer ces représentations spectrales seront différentes selon le type de signal (on verra la transformée de Fourier, la transformée de Fourier rapide).*

• Fréquence et gamme

A titre d'exemple on traite ici de la construction des gammes (dans le système occidental) à partir de la notion de fréquence.

Nous percevons un son de fréquence donnée et le son de fréquence double comme étant la même note (l'une étant considérée comme étant plus haute que l'autre). Ainsi on peut donner la

Définition 2.7 *La note associée à une fréquence f est l'ensemble des fréquences de la forme*

$$N = \{2^k f, k \in \mathbb{Z}\}.$$

Ainsi on appelle "la" l'ensemble des fréquences

$$\{\dots, 55, 110, 220, 440, 880, 1760, \dots\}.$$

D'autre part les rapports de fréquences sont perçus comme des intervalles entre les sons correspondants. Par exemple l'intervalle de quinte a pour rapport $\frac{3}{2}$.

Définition 2.8 *Les harmoniques d'une note est l'ensemble des fréquences de la forme*

$$N = \{kf, k \in \mathbb{N}\}$$

Ce sont en ces valeurs qu'on voit apparaître les pics du spectre de la figure 6. On construit ainsi la gamme harmonique :

$$\begin{array}{cccccccc} \text{Do} & \text{Ré} & \text{Mi} & \text{Fa} & \text{Sol} & \text{La} & \text{Si} & \text{Do} \\ f & \frac{9}{8}f & \frac{5}{4}f & \frac{4}{3}f & \frac{3}{2}f & \frac{5}{3}f & \frac{15}{8}f & 2f \end{array}$$

Et la gamme à tempéramment égal. Cette valeur de a permet de rendre les 12 demi-tons de la gamme égaux. On pose $a = 2^{\frac{1}{12}}$:

$$\text{Do} \quad \text{Ré} \quad \text{Mi} \quad \text{Fa} \quad \text{Sol} \quad \text{La} \quad \text{Si} \quad \text{Do}$$

$$f \ a^2 f \ a^4 f \ a^5 f \ a^7 f \ a^9 f \ a^{11} f \ 2f$$

• **Transformée de Fourier des fonctions L^1 et L^2**

On part d'une fonction $f(t)$ définie sur toute la droite réelle, intégrable au sens de Lebesgue sur \mathbb{R} , et on pose

$$\hat{f}(\xi) = \int_{\mathbb{R}} f(t) e^{-it\xi} dt$$

pour $\xi \in \mathbb{R}$. Cette fonction \hat{f} est bien définie sur \mathbb{R} et on l'appelle transformée de Fourier de f . Une de ses propriétés basiques est qu'elle tend vers 0 à l'infini et est continue sur \mathbb{R} .

Une autre propriété utile pour nous est ce que l'on appelle la formule d'inversion de Fourier. Supposons que f et \hat{f} soient toutes les deux dans $L^1(\mathbb{R})$. Alors on a, pour presque tout $x \in \mathbb{R}$

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(t) e^{itx} dt.$$

On peut aussi "prolonger" la transformée de Fourier aux fonctions $L^2(\mathbb{R})$ (ce qui nous donne ce que l'on appelle la transformée de Plancherel). Le point de départ est la densité de $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ dans $L^2(\mathbb{R})$. Si on prend une fonction $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ on voit très facilement que \hat{f} est un élément de $L^2(\mathbb{R})$. On a dans ce cas

$$\int_{\mathbb{R}} |f(t)|^2 dt = \int_{\mathbb{R}} |\hat{f}(t)|^2 dt,$$

ce qui signifie que l'opérateur $\mathcal{F} : f \rightarrow \hat{f}$ de $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ dans $L^2(\mathbb{R})$ est uniformément continue et donc se prolonge de manière unique, grâce à la densité de $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ dans $L^2(\mathbb{R})$.

• **Principe de la transformée de Fourier rapide**

Lorsque l'on considère une fonction continue sur $[0, 2\pi]$, on peut tout d'abord la discrétiser à l'aide de N points. La plupart du temps, cette discrétisation est fait sur un mode dichotomique, c'est-à-dire que N est une puissance de 2. Après avoir fait cette discrétisation, on obtient une nouvelle fonction que l'on peut noter f_N et le calcul de ses N coefficients de Fourier nécessite N^2 opérations élémentaires. La transformée de Fourier rapide est un algorithme permettant de calculer ceux-ci avec une complexité de $2N \log N$. L'idée étant, à chaque calcul d'un nouveau coefficient, de "réutiliser" les calculs faits pour calculer les coefficients précédents.

2.4 Limites de la transformée de Fourier standard

Pour analyser et synthétiser des sons (comme dans les autres applications de l'analyse et du traitement du signal), il s'agit d'extraire du signal, ici signal sonore, des informations pertinentes sous forme de valeurs numériques à la fois caractéristiques du signal en question et pas trop nombreuses.

L'analyse de Fourier est performante pour l'analyse de signaux périodiques et suffisamment réguliers, notamment grâce à la transformée de Fourier rapide. Les signaux numériques caractérisant le signal dans le domaine des fréquences. En effet si x est le temps et $f(x)$ le signal

$$c_n(f) = \frac{1}{T} \int_0^T f(x) e^{-2i\pi \frac{n}{T} x} dx$$

pour un signal de période T ,

$$\hat{f}(\lambda) = \int_{\mathbb{R}} f(x) e^{-2i\pi \lambda x} dx$$

pour f dans $L^1(\mathbb{R})$, donnent le contenu fréquentiel pour la fréquence $\frac{n}{T}$ ou λ .

Cependant si le signal est irrégulier la quantité de paramètres à conserver devient trop importante. Par conséquent, même avec la transformée de Fourier rapide, l'analyse de Fourier présente des inconvénients majeurs pour une analyse satisfaisante du signal.

Le premier et principal inconvénient est que dans le spectre $\hat{f}(\lambda)$ tous les aspects temporels du signal disparaissent, par exemple le début et la fin (attaque et déclin de la note) ou l'instant d'une singularité. En effet, si f n'est pas continue il est quasiment impossible de le détecter à l'aide de \hat{f} . Nous pouvons le voir sur l'exemple suivant où f est un signal créneau : $f = 1_{[-a,a]}$ avec $a > 0$. Sa transformée de Fourier est

$$\hat{f}(\omega) = \frac{\sin(2\pi\omega a)}{\pi\omega}.$$

On représente en dessous la fonction $f = 1_{[-1,1]}$ et sa transformée de Fourier ainsi que la fonction $f = 0.9 1_{[-1,1]} + 0.1 1_{[-2,2]}$ et sa transformée.

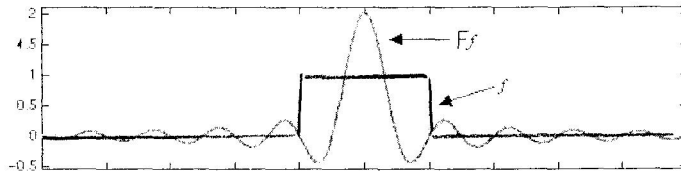


Figure 11a

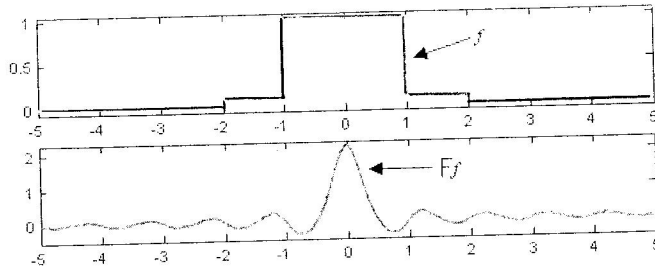


Figure 11b

On remarque que les deux transformées de Fourier sont similaires alors que la seconde fonction possède bien plus de discontinuités que la fonction créneau. On ne peut donc pas repérer les ruptures et les changements de régularité de la fonction f (l'intégration sur \mathbb{R} effectue une moyenne qui masque les discontinuités).

Autre exemple : deux notes de musique jouées l'une après l'autre, ainsi que la transformée de Fourier du signal.

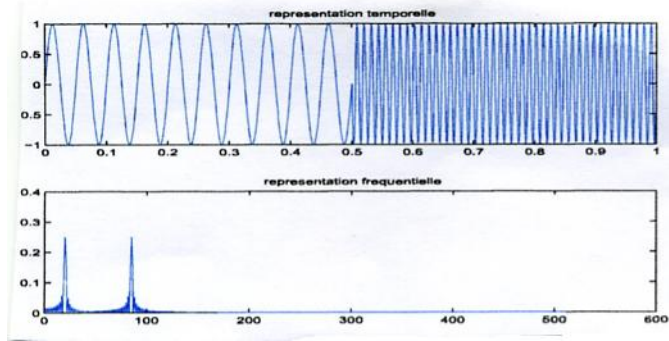


Figure 11c

De la même façon si on écoute un "glissando" joué par un violoncelle par exemple, la fréquence dépend continûment du temps, ce que l'analyse de Fourier ne "voit" pas.

Un autre inconvénient est la non-causalité de la transformée de Fourier \mathcal{F} . Il est clair que le calcul de \hat{f} nécessite la connaissance de f sur \mathbb{R} tout entier. Une analyse en temps réel est donc impossible. On ne peut pas connaître le spectre de \hat{f} d'un signal f si on ignore la fonction.

Par exemple, deux fonctions identiques sur un même intervalle peuvent avoir des transformées de Fourier très différentes.

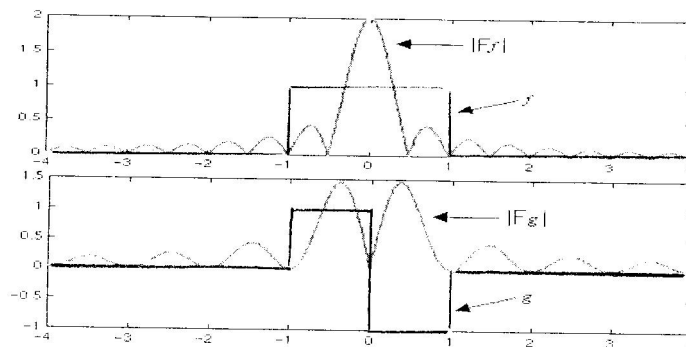


Figure 12

Le troisième inconvénient est dû au principe d'incertitude d'Heisenberg.

Définition 2.9 On appelle *dispersion d'énergie de f en temps*

$$\sigma_f^2 = \int_{\mathbb{R}} x^2 |f(x)|^2 dx$$

où f et xf sont dans $L^2(\mathbb{R})$. On appelle *dispersion d'énergie de f en fréquence*

$$\sigma_{\hat{f}}^2 = \int_{\mathbb{R}} \xi^2 |\hat{f}(\xi)|^2 dx$$

On pose $E_f = \int_{\mathbb{R}} |f(t)|^2 dt$ l'énergie de f ; on introduit aussi

$$(\Delta t)^2 = \frac{(\sigma_f)^2}{E_f},$$

et

$$(\Delta \lambda)^2 = \frac{(\sigma_{\hat{f}})^2}{E_f},$$

Le principe d'incertitude d'Heisenberg précise le lien entre la localisation d'un signal et celle de son spectre, autrement dit entre les dispersions σ_f et $\sigma_{\hat{f}}$, c'est-à-dire :

$$\sigma_f - \sigma_{\hat{f}} \geq \frac{E_f}{4\pi} \implies \Delta t \times \Delta \lambda \geq \frac{1}{4\pi}.$$

Ceci pour une fonction $f : \mathbb{R} \rightarrow \mathbb{C}$ de classe C^1 telle que f , f' et xf sont dans $L^2(\mathbb{R})$.

Ceci se comprend dans le sens où plus le support de f est "petit", plus le support de \hat{f} est "grand" et réciproquement.

Exemple : $\mathcal{F}(\delta) = 1$ et $\mathcal{F}(1) = \delta$.

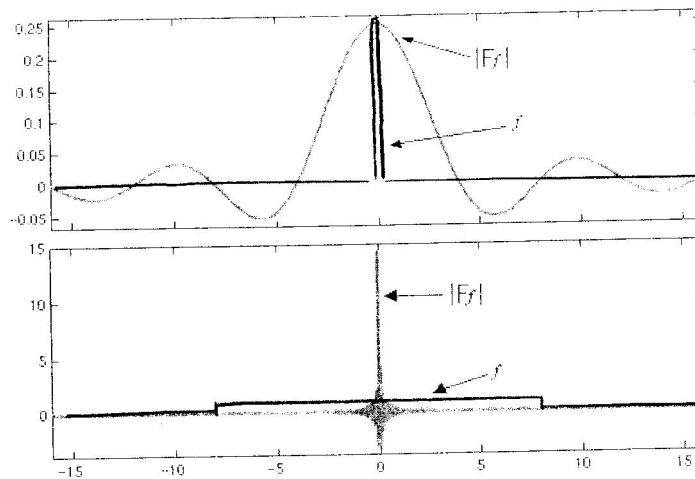


Figure 13 : On représente ici la fonction $f = 1_{[-a,a]}$ où $a = 1/8$ et $a = 8$

Cette propriété implique que la transformée de Fourier inverse peut être numériquement instable puisque les informations utiles pour reconstruire f à partir de \hat{f} par la formule d'inversion peuvent se trouver à très haute fréquence.

En effet, si l'on prend l'exemple d'un enregistrement analogique d'un son régulier sur un disque détérioré en un unique point, la présence du craquement empêche la transformée de Fourier du signal de décroître rapidement ce qui provoquera une reconstruction par la formule d'inversion numériquement instable car en tout point elle est obtenue à partir de l'intégrale d'une fonction très oscillante. Comment localiser cette instabilité numérique ?

Tout ceci montre que la transformée de Fourier est une transformation à caractère global, ce qu'il va falloir corriger.

3 La transformée de Fourier à fenêtre (transformée de Gabor) : exemple d'un signal musical

3.1 Idée de fenêtre

Pour éviter l'inconvénient du caractère global de la transformée de Fourier, une idée naturelle consiste à "tronquer" le signal. On localise ainsi l'analyse en sélectionnant une portion autour d'une position temporelle, puis on fait l'analyse de Fourier de ce signal tronqué. On peut ensuite recommencer pour d'autres positions.

Tout d'abord, il est donc naturel de multiplier le signal $f(t)$ par le créneau $1_{[-a,a]}$ dans le cas où l'on veut "ouvrir une fenêtre" autour de l'origine, ou bien un translaté pour regarder le signal en un autre point.

On regarde donc le signal $g(\lambda) = 1_{[-a,a]}f(\lambda)$ ce qui avec les règles sur les transformées de Fourier revient à effectuer le calcul suivant :

$$\hat{g}(\lambda) = \widehat{1_{[-a,a]}f(\lambda)} = \frac{\sin 2\pi a\lambda}{\pi\lambda} * \hat{f}(\lambda) = (s_a * \hat{f})(\lambda)$$

où $s_a(\lambda) = \frac{\sin 2\pi a\lambda}{\pi\lambda}$ est le sinus cardinal.

La représentation graphique suivante nous permet de comprendre l'insuffisance de l'utilisation du sinus cardinal

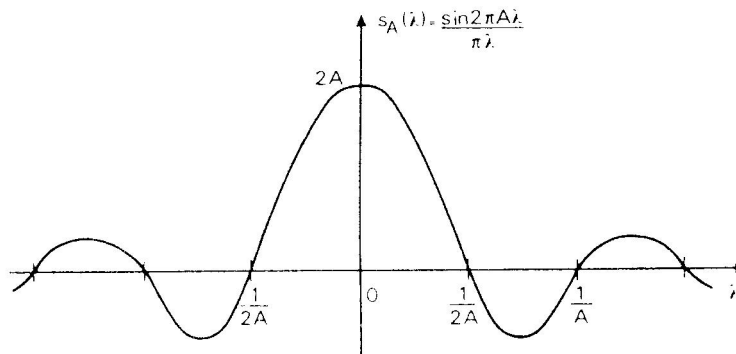


Figure14 : Le sinus cardinal.

En effet, cette fonction s'amortit très lentement et présente d'important lobes près de l'origine.

Remarquons tout de même que plus a est grand, plus l'approximation de \hat{f} par \hat{g} est précise, d'où l'envie d'utiliser des fonctions plus régulières et centrées autour de l'origine.

• Exemples de fenêtre

Fenêtre triangulaire :

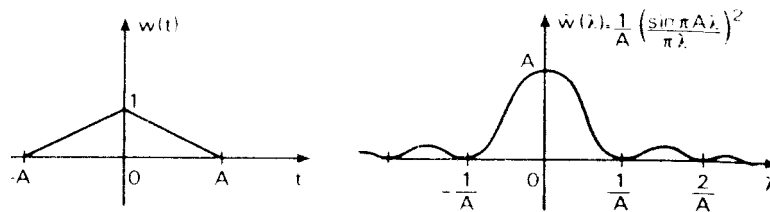


Figure 15 : Fenêtre triangulaire

Fenêtre de Hamming et Hanning :

de la forme $w(t) = [\alpha + (1 - \alpha) \cos 2\pi \frac{t}{a}] r(t)$.

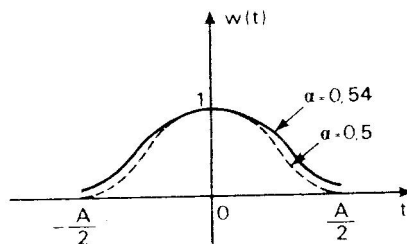


Figure16 Pour $\alpha = 0.54$ on obtient la fenêtre de Hamming et pour $\alpha = 0.5$ on obtient celle de Hanning.

Fenêtre de Gauss :

Fenêtre de Gauss $\omega(t) = A e^{-\alpha t^2}$ ($\alpha, A > 0$)

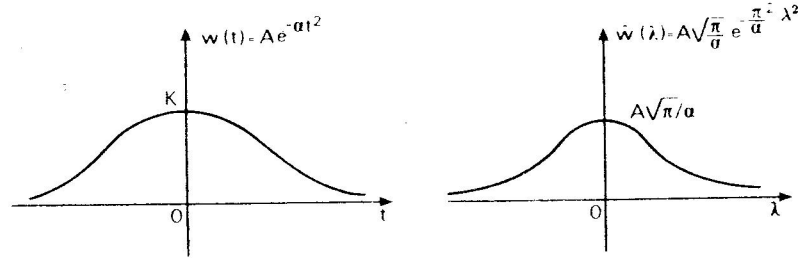


Figure 17

• Définition

Par la suite on est amenés à faire glisser cette fenêtre (centrée en 0 et d'énergie 1) utilisée pour localiser en temps l'analyse.

On obtient alors une famille de coefficients $W_f(\lambda, b)$ remplaçant les valeurs de $\hat{f}(\lambda)$ où λ représente la fréquence, b localise l'analyse et où ω est la fenêtre

$$W_f(\lambda, b) = \int_{-\infty}^{+\infty} f(t)\bar{\omega}(t-b)e^{-2i\pi\lambda t} dt.$$

L'application $f \rightarrow W_f$ s'appelle la transformée de Fourier à fenêtre glissante ou encore la transformée de Gabor. On note alors $\omega_{\lambda,b}(t) = \omega(t-b)e^{-2i\pi\lambda t}$ et on peut l'écrire sous la forme d'un produit scalaire dans L^2

$$W_f(\lambda, b) = (f, \omega_{\lambda,b}).$$

3.2 Les formules de Gabor

Comme pour l'analyse de Fourier, la connaissance des $W_f(\lambda, b)$ pour toutes les valeurs réelles de λ et de b déterminera complètement et de manière unique le signal f .

On a grâce à Gabor la formule inverse donnant f à partir des $W_f(\lambda, b)$ dite formule de reconstruction ou de synthèse et la conservation de l'énergie.

Théorème 3.1 Soit $\omega \in L^1 \cap L^2$ centrée en 0, telle que $|\hat{\omega}|$ soit une fonction paire et $\|\omega\|_2 = 1$. C'est ce qu'on appelle la fenêtre. On a posé $\omega_{\lambda,b}(t) = \omega(t-b)e^{-2i\pi\lambda t}$, λ et b étant des réels, et on considère pour tout signal $f \in L^2$ les

$$W_f(\lambda, b) = \int_{-\infty}^{+\infty} f(t)\bar{\omega}(t-b)e^{-2i\pi\lambda t} dt.$$

On a la conservation de l'énergie

$$\int \int_{\mathbb{R}^2} |W_f(\lambda, b)|^2 d\lambda db = \int_{-\infty}^{+\infty} |f(t)|^2 dt,$$

et la formule de reconstruction

$$f(x) = \int \int_{\mathbb{R}^2} W_f(\lambda, b) \omega_{\lambda, b}(x) d\lambda db,$$

au sens où si $g_A(x) = \int_{b \in \mathbb{R}} \int_{|\lambda| \leq A} W_f(\lambda, b) \omega_{\lambda, b}(x) d\lambda db$ alors $g_A \rightarrow f$ dans L^2 .

3.3 Comparaison Fourier-Gabor et limites de la transformée de Gabor

On peut écrire ces deux transformées :

$$f(x) = \int_{-\infty}^{+\infty} \hat{f}(\xi) e^{2i\pi x \xi} d\xi,$$

et

$$f(x) = \int \int_{\mathbb{R}^2} W_f(\lambda, b) \omega_{\lambda, b}(x) d\lambda db.$$

Nous pouvons voir qu'elles peuvent s'interpréter comme la décomposition du signal f sur une famille de fonctions jouant le rôle d'une base et où les sommes sont remplacées par des intégrales. Dans le cas de Fourier cette famille est évidemment $(e^{2i\pi x \xi})_{\xi \in \mathbb{R}}$ et dans le cas de Gabor $(\omega_{\lambda, b}(x))_{(\lambda, b) \in \mathbb{R}^2}$ où on rappelle que $\omega_{\lambda, b}(t) = \omega(t - b)e^{-2i\pi \lambda t}$.

On appelle les éléments de cette famille les atomes de cette transformée ou bien "ondelettes de Gabor". Graphiquement, on a une fonction sinusoidale fortement amortie pour Gabor, c'est-à-dire une gaussienne modulée en temps.

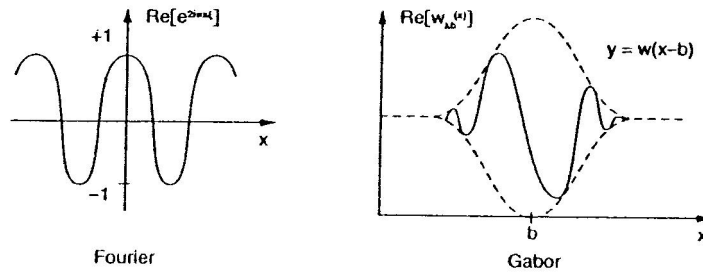


Figure 18 : Fonctions de base dans la décomposition de Fourier et de Gabor.

On obtient de la même façon dans l'espace des fréquences :

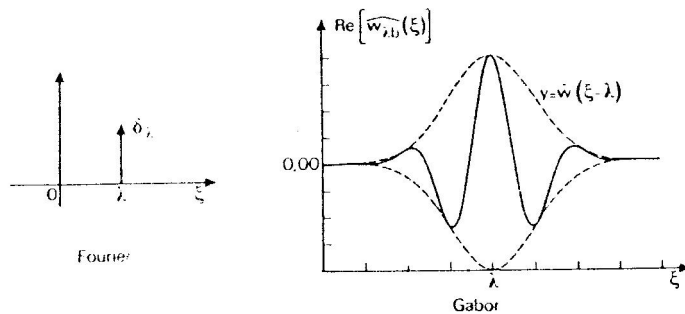


Figure 19

On voit alors que dans la méthode de Fourier les fonctions de base sont totalement concentrées en fréquence (sous forme d'impulsion de Dirac) et totalement réparties sur l'axe du temps car la sinusoïde ne s'amortit pas et s'étend de $-\infty$ à $+\infty$. On voit ainsi que la transformée de Fourier donne le maximum d'informations sur la répartition des fréquences mais perd complètement celles relatives au temps.

La transformée de Gabor nous met en présence d'un objet nouveau appelé Plan temps-fréquence ou espace de phase. Il s'agit ici du plan \mathbb{R}^2 . La transformée de Gabor permet donc de représenter les fonctions de $L^2(\mathbb{R})$ comme des fonctions sur le plan temps-fréquence.

Ceci met clairement en évidence les avantages de la méthode de Gabor sur celle de Fourier.

Prenons l'exemple d'un signal de durée finie. Par la formule de synthèse pour Fourier, il est nécessaire de connaître toutes les valeurs de $\hat{f}(\xi)$ ou du moins avec précision sur une très grande plage car la décroissance à l'infini peut être très lente. Numériquement, les effets de toutes ces sinusoïdes doivent se compenser pour donner la valeur 0 en dehors du support de f ce qui nécessite une précision numérique pour les valeurs de $\hat{f}(\xi)$ beaucoup trop importante.

Dans le cas de la méthode de Gabor ce problème ne se pose pas. En effet, si l'on considère f nulle au voisinage de b_0 par exemple sur l'intervalle $]b_0 - \alpha, b_0 + \alpha[$ assez long et que l'on suppose $\omega(t)$ négligeable pour $|t| \geq 1$, on aura pour b proche de b_0 , $W_f(\lambda, b)$ proche de $\int_{b-1}^{b+1} f(t)\bar{\omega}(t)dt = 0$ et les coefficients $W_f(\lambda, b)$ seront négligeables.

Par contre s'il y a une forte perturbation de f en $t = b_0$, les coefficients seront grands à la fois pour $b = b_0$ mais également pour des valeurs de λ proches de la fréquence locale de f .

- **Inconvénient majeur de Gabor :**

On peut cependant noter un inconvénient majeur de la méthode de Gabor. En effet, la fenêtre est de longueur fixe ce qui est un handicap important lorsqu'on veut traiter des signaux dont les variations peuvent avoir des ordres de grandeur très variables. En particulier, l'attaque d'une note est une phase siège de hautes fréquences. Bien qu'étant très brève, cette phase est caractéristique de l'instrument et de l'interprétation. Elle est donc tout aussi importante que la phase de maintien de la note qui contient en général des fréquences plus basses. Il serait donc intéressant de pouvoir faire varier cette fenêtre non seulement par translation mais également par dilatation et contraction en fonction des caractéristiques du signal et de ce que l'on veut mettre en évidence.

4 Les ondelettes

4.1 Idée de base

a - Intérêt des ondelettes

Nous avons vu précédemment que la transformée de Fourier et la transformée de Fourier à fenêtre présentent un inconvénient majeur. La transformée de Fourier est une transformation globale, la transformée de Gabor est locale mais toutes les deux sont de résolution temporelle fixe.

La transformée en ondelettes va nous permettre de pallier cet inconvénient. Il est en effet nécessaire de disposer d'un outil qui adapte sa résolution à la taille de l'objet ou du détail analysé. Plus généralement, on peut distinguer deux intérêts principaux des ondelettes en tant qu'outil mathématique d'analyse et de représentation des fonctions.

Le premier est, comme dit précédemment, centré autour de l'idée de construire un outil d'analyse local en temps. Nous introduirons alors la transformée continue en ondelettes et nous verrons comment elle résout les problèmes posés par les outils classiques. La transformée en ondelettes est non seulement une analyse locale mais, de plus, sa résolution temporelle est variable. Elle permet donc de décrire le comportement local des signaux à différentes échelles de temps.

Le second est celui du problème de reconstruction et de la recherche de représentations pratiques. En effet, cette transformée nous permet bien de reconstruire le signal de départ et présente des avantages majeurs d'un point de vue numérique.

b - Définition dans un cadre général

La transformation en ondelettes continue consiste à créer, à partir d'une ondelette mère notée ψ (qui ressemble à une petite onde) ou encore appelée ondelette analytique, une famille d'ondelettes

$$\psi_{ab}(t) = \frac{1}{\sqrt{a}} \psi \left(\frac{t-b}{a} \right)$$

où $a \in \mathbb{R}^{+*}$ est l'échelle et $b \in \mathbb{R}$ est la position (a sert à dilater et b à translater). On dit également que tout couple $(a, b) \in \mathbb{R}^{+*} \times \mathbb{R}$ définit un atome de la transformée.

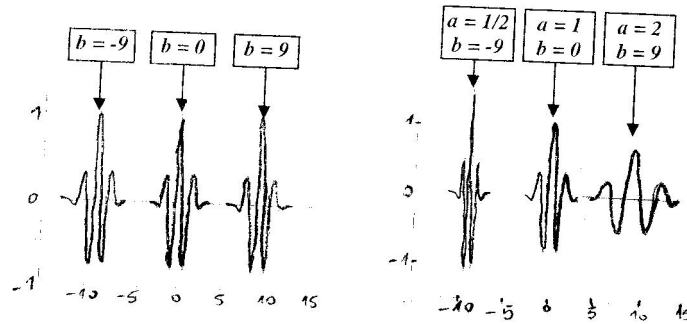


Figure 20 : Exemple de translation et dilatation d'une ondelette.

c - Idée de l'analyse multirésolution et d'analyse temps-échelle

Les ondelettes étant une extension de l'analyse de Fourier, le principe est le même : on reconstruira un signal à partir de sa transformée en ondelettes en "additionnant" des ondelettes de différentes tailles ; de même qu'on reconstruit un signal à partir de sa transformée de Fourier.

Toutefois, le fait de comprimer ou d'étirer les ondelettes pour modifier leur fréquence change quelque chose. Les ondelettes vont s'adapter "automatiquement" aux différentes composantes du signal : elles utilisent une fenêtre étroite pour regarder les composantes transitoires de haute fréquence et une fenêtre large pour regarder les composantes de longue durée, de basse fréquence. C'est cette procédure qu'on appelle multirésolution.

On parle également de microscope mathématique. En effet, en comprimant les ondelettes, on accroît le grossissement du microscope pour révéler des détails de plus en plus fin.

D'autre part, on emploie parfois le terme d'échelle plutôt que celui de résolution, ou encore d'octave.

Le terme résolution évoque le nombre d'échantillons du signal (nombre d'ondelettes utilisées).

Le terme d'échelle désigne la corrélation entre la taille de l'ondelette et la taille des composantes que l'on peut voir.

Le terme d'octave rappelle que doubler la résolution double la fréquence des ondelettes.

On peut alors parler d'analyse temps-échelle pour les ondelettes comme on parlait d'analyse temps-fréquence pour la transformée de Gabor.

d - Caractéristiques de l'ondelette mère et coefficients du signal

Les caractéristiques de l'ondelette mère sont nettement différentes de celles d'une fenêtre. En effet, une fenêtre avait plutôt l'allure d'un créneau, alors que la définition de ψ lui imposera d'être d'intégrale nulle. Par ailleurs, on imposera à ψ une bonne localisation autour de 0 et pour $\hat{\psi}$ une localisation autour d'une certaine fréquence $\omega_0 > 0$. On obtient ainsi des fonctions à décroissance rapide : elles oscillent puis s'amortissent ; ce qui leur donne leur nom.

L'analyse par ondelettes associe à la fonction analysée les coefficients

$$c_f(a, b) = (f, \psi_{ab}) = \int_{-\infty}^{+\infty} f(t) \bar{\psi}_{ab} dt$$

qui comme les coefficients de Gabor décrivent le contenu de la fonction au voisinage de $(b, \frac{\omega_0}{a})$.

e - Un premier exemple : comparaison avec les fonctions de base de Gabor

Les graphiques ci-dessous illustrent les différences de comportement entre les fonctions de base de Gabor $w_{\lambda b}(t)$ à enveloppe rigide et les ondelettes à contraction-dilatation.

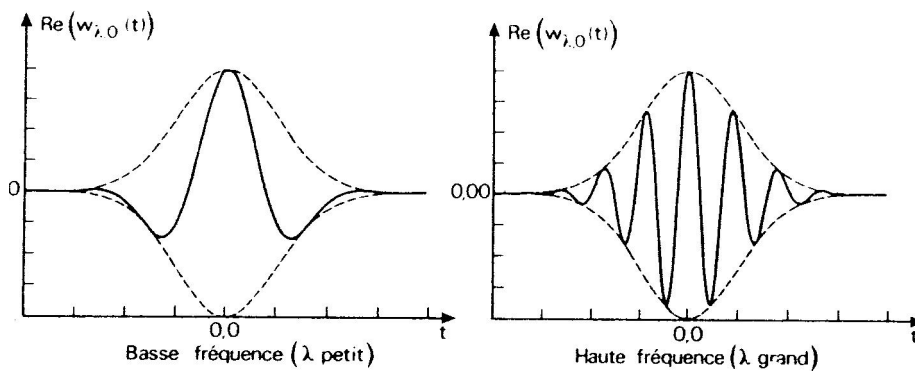


Figure 21 : Fonctions de base de Gabor $w_{\lambda b}(t) = e^{-\frac{1}{2}(t-b)^2} e^{2i\pi\lambda t}$. L'enveloppe est rigide et le nombre d'oscillations augmente avec les fréquences.

Figure 22 : Une ondelette-mère (8ème dérivée gaussienne)

Ondelettes basses et hautes fréquences : elles gardent la même forme et le même nombre d'oscillations. Elles s'étirent (a grand) ou se contractent (a petit). On voit donc que contrairement aux fonctions de base de Gabor, les ondelettes n'ont pas d'enveloppe rigide.

4.2 La transformée en ondelettes

a - Définition d'une ondelette mère et d'une ondelette de base

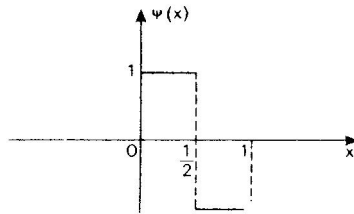
Définition 4.1 Soit $\psi \in L^1 \cap L^2$ telle que

$$(i) \quad K = \int_{-\infty}^{+\infty} \frac{|\hat{\psi}(\lambda)|}{|\lambda|} d\lambda < +\infty$$

et

$$(ii) \quad \|\psi\|_2 = 1.$$

On dit alors que ψ est une ondelette mère.



Remarque 4.2 1- La condition $\|\psi\|_2 = 1$ signifie que ψ a une énergie finie. Elle est de plus normalisée.

2- L'hypothèse (i) implique que $\hat{\psi}(0) = \int_{-\infty}^{+\infty} \psi(t) dt = 0$ car $\hat{\psi}$ est continue. Dans les cas pratiques, cette condition est suffisante pour obtenir les résultats ci-dessus. Si on suppose que ψ et $x\psi$ sont intégrables alors $\hat{\psi} \in C^1$; en utilisant la formule de Taylor en 0 cela montre que $\lambda \rightarrow \frac{|\hat{\psi}(\lambda)|}{|\lambda|}$ est continue en 0, ce qui assure la convergence de l'intégrale sur tout compact de \mathbb{R} . De plus la convergence en l'infini de $\int_{-\infty}^{+\infty} \frac{|\hat{\psi}(\lambda)|}{|\lambda|} d\lambda$ est la conséquence de la formule de Cauchy-Schwarz et du fait que $\psi \in L^2$.

3- Dans le cas de la parole et de l'analyse de son (musical) il est commode d'extraire de l'analyse des informations à propos de la distribution d'énergie et du comportement en phase dans la représentation de la transformée en ondelettes. Ceci peut être fait facilement en utilisant une ondelette à valeur complexe.

Définition 4.3 On appelle ondelette de base

$$\psi_{ab}(t) = \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right)$$

où a et b sont dans \mathbb{R} et a non nul, et on considère pour tout signal $f \in L^2$ les coefficients d'ondelettes

$$c_f(a, b) = \int_{-\infty}^{+\infty} f(t) \bar{\psi}_{ab}(t) dt.$$

Remarque 4.4 1- On peut donner pour les coefficients c_f une définition équivalente

$$c_f(a, b) = \sqrt{a} \int_{-\infty}^{+\infty} \hat{f}(t) \bar{\hat{\psi}}(at) e^{2i\pi tb} dt$$

ou

$$c_f(a, b) = (f * \check{\psi}_a)(b)$$

où $\check{\psi}_a(x) = \frac{1}{\sqrt{a}} \psi\left(-\frac{x}{a}\right)$.

2- Vu du côté temporel (ou spatial) x , $c_f(a, b)$ renseigne sur le signal f autour du point b . Vu du côté fréquentiel ν , $c_f(a, b)$ renseigne sur le signal \hat{f} autour de la fréquence $\frac{1}{a}$.

b - Théorème fondamental

Nous avons le théorème suivant

Théorème 4.5 On a avec les notations précédentes :

a) La conservation de l'énergie

$$\frac{1}{K} \int_{\mathbb{R}^2} |c_f(a, b)|^2 \frac{da db}{2} = \int_{-\infty}^{+\infty} |f(t)|^2 dt.$$

b) La formule de reconstruction

$$f(x) = \frac{1}{K} \int_{\mathbb{R}^2} c_f(a, b) \psi_{ab}(x) \frac{da db}{a^2}$$

au sens où si $f_\varepsilon(x) = \frac{1}{K} \int_{|a| \geq \varepsilon, b \in \mathbb{R}} c_f(a, b) \psi_{ab}(x) \frac{da db}{a^2}$ alors $f_\varepsilon \rightarrow f$ dans L^2 lorsque ε tend vers 0.

4.3 Exemples

a - Ondelette de Morlet

$$\psi(t) = e^{-\frac{t^2}{2}} \cos(5t) \text{ ou } \psi(t) = e^{-\pi t^2} e^{10i\pi x}.$$

On peut d'abord remarquer que celle-ci n'est pas normalisée, mais ceci n'a pas beaucoup d'importance. Par contre l'hypothèse $\int_{-\infty}^{+\infty} \frac{|\hat{\psi}(\lambda)|}{|\lambda|} d\lambda < +\infty$ n'est pas non plus vérifiée puisque $\hat{\psi}(0) = \sqrt{2\pi} e^{-\frac{25}{2}} > 0$. Cependant cette valeur de $\hat{\psi}(0)$ est de l'ordre de 10^{-5} . On peut donc la considérer comme "nulle" dans un calcul numérique effectué sur ordinateur, et on est donc assez proche des conditions d'application du théorème qui reste malgré tout inapplicable (au sens strict du terme, à savoir au sens mathématique).

b - Un exemple simple d'ondelette-mère : l'ondelette de Haar

Elle est définie par

$$(2) \quad \begin{cases} \psi(x) = 1 & \text{si } 0 < x < \frac{1}{2}; \\ \psi(x) = -1 & \text{si } \frac{1}{2} < x < 1; \\ \psi(x) = 0 & \text{sinon .} \end{cases}$$

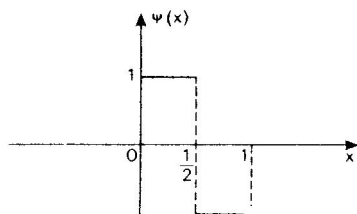


Figure 23 : ondelette de Haar

On a dans ce cas $\hat{\psi}(\xi) = ie^{-i\pi\xi} \frac{1-\cos(\pi\xi)}{\pi\xi}$ et le spectre d'amplitude de l'ondelette de Haar est $|\hat{\psi}(\xi)|$.

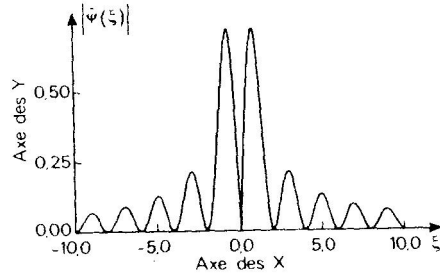


Figure 24 : Spectre d'amplitude de l'ondelette de Haar

On remarque qu'ici la convergence vers 0 à l'infini est très lente, du fait de l'irrégularité de ψ ce qui peut être un fort inconvénient.

c - Les dérivées gaussiennes

Elles sont de la forme $\psi_n(x) = \frac{d^n}{dx^n} e^{-\pi x^2}$, pour $n \geq 1$. On utilise essentiellement la dérivée seconde, qu'on écrit le plus souvent sous la forme :

$$\psi(x) = \frac{2}{\sqrt{3}} \pi^{-\frac{1}{4}} (1 - x^2) e^{-\frac{x^2}{2}}$$

On a alors $\hat{\psi}(\xi) = K \xi^2 e^{-2\pi^2 \xi^2}$ où K est une constante. On peut remarquer que ψ et $\hat{\psi}$ sont à décroissance très rapide et C^∞ , ce sont des éléments de la classe de Schwartz \mathcal{S} .

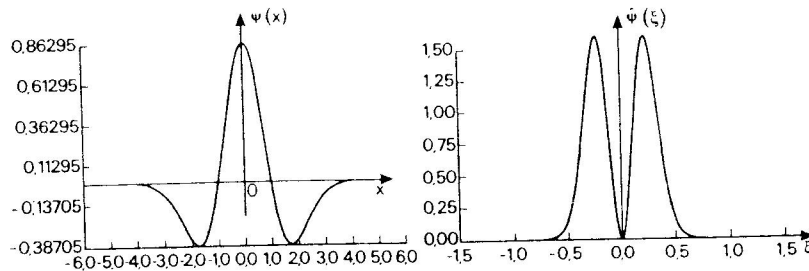


Figure 25 : Deuxième dérivée de gaussienne et son spectre.

En raison de son graphe, on l'appelle souvent le chapeau mexicain.

La huitième dérivée d'une gaussienne et son spectre :

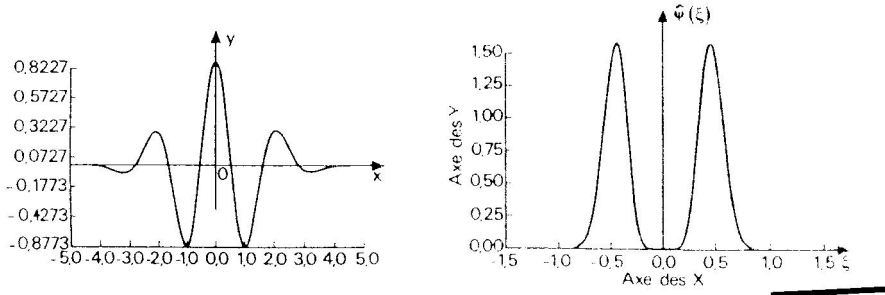


Figure 26 : Huitième dérivée de gaussienne et son spectre.

d - Exemple de la transformée en ondelette d'une fréquence pure

Si $f(x) = \cos(2\pi kx)$ alors

$$c_f(a, b) = \int_{-\infty}^{+\infty} \left(\frac{e^{2i\pi kx} + e^{-2i\pi kx}}{2} \right) \bar{\psi}_{ab}(x) dx = \frac{1}{2} \left[\bar{\psi}_{ab}(k) + \bar{\psi}_{ab}(-k) \right].$$

D'où

$$c_f(a, b) = \frac{\sqrt{a}}{2} \left[\hat{\psi}(ak)e^{2i\pi kb} + \hat{\psi}(-ak)e^{-2i\pi kb} \right].$$

Donc si ψ est complexe analytique $c_f(a, b) = \frac{\sqrt{a}}{2} \hat{\psi}(ak)e^{2i\pi kb}$ et si ψ est réelle $c_f(a, b) = \sqrt{a} \operatorname{Re} \left(\hat{\psi}(ak)e^{-2i\pi kb} \right)$.

4.4 Bases orthonormées d'ondelettes et analyse multirésolution

a - Motivation : le passage à la transformée discrète en ondelettes

Il est facile de constater que la transformée en ondelettes continue est infiniment redondante. Il s'agit donc d'élaborer des méthodes pour extraire rapidement l'information essentielle de ces transformées redondantes, c'est-à-dire de trouver une famille $\{\psi_{j,k}\}$ pour $(j,k) \in I$ d'ondelettes orthogonales sur lesquelles on pourrait décomposer tout signal $f \in L^2$ en série double :

$$f(x) = \sum_{(j,k) \in I} (f, \psi_{j,k}) \psi_{j,k}(x)$$

où I est un ensemble discret de $\mathbb{R}^{*+} \times \mathbb{R}$.

Fixons $a_0 > 1$ et $b_0 > 0$ et prenons $a \in \{a_0^p\}_{p \in \mathbb{Z}}$ et $b \in \{na_0^p b_0\}_{p,n \in \mathbb{Z}}$. Par conséquent, au lieu d'utiliser la famille d'ondelettes ψ_{ab} on se sert pour la transformée discrète de la famille dénombrable d'ondelettes

$$\psi_{p,n} = a_0^{-\frac{p}{2}} \psi(a_0^{-p}t - nb_0)$$

avec $a_0 > 1$, $b_0 > 0$ fixés et p et n dans \mathbb{Z} .

Alors pour $f \in L^2$, on définit la transformée en ondelettes discrète de f par

$$c_f(p,n) = \int_{\mathbb{R}} f(t) \bar{\psi}_{p,n}(t) dt = (f, \psi_{p,n})_{L^2}$$

avec p et n dans \mathbb{Z} .

Le choix usuel pour a_0, b_0 est $a_0 = 2$ et $b_0 = 1$ (théorème d'échantillonnage de Shannon). On a donc

$$\psi_{j,k} = 2^{-\frac{j}{2}} \psi(2^{-j}t - k).$$

Il est naturel de se demander : existe-t-il (et si oui sous quelles conditions) une fonction ψ telle que la famille $\{\psi_{j,k}\}_{(j,k) \in \mathbb{Z}^2}$ soit une base orthonormée de L^2 ?

C'est cette question qui amène à la notion d'analyse multirésolution.

b - Définition et explication

Définition 4.6 Une analyse multirésolution de $L^2(\mathbb{R})$ est une famille $M = \{V_j\}_{j \in \mathbb{Z}}$ de sous-espaces vectoriels emboîtés qui possède les propriétés ci-dessous (que l'on regroupe en trois blocs).

- $\{V_j\}_{j \in \mathbb{Z}}$ est une suite d'espaces d'approximation, c'est-à-dire que
 - (1) V_j est un sous-espace fermé de L^2 ,
 - (2) $V_j \subset V_{j-1}$,
 - (3) $\bigcup_{j \in \mathbb{Z}} V_j = L^2$ et $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$.
- Tous les espaces V_j sont obtenus par la dilatation ou la contraction dyadique des fonctions d'un espace unique (par exemple V_0)
 - (4) $\forall j \in \mathbb{Z}, v(t) \in V_j \iff v(2t) \in V_{j-1}$.
- Une dernière propriété concerne la translation des fonctions. Elle suppose l'existence d'une fonction qui, par translation entière, permet de construire une base de V_0
 - (5) $\exists g \in V_0$ telle que $\{g(t - k)\}_{k \in \mathbb{Z}}$ est une base de Riesz de V_0 .

Définition 4.7 La famille $\{e_k\}_{k \in \mathbb{Z}} \subset L^2$ est une base de Riesz de L^2 si

$$(i) \quad \forall h \in L^2, \exists! \alpha \in l^2(\mathbb{Z}) \text{ tel que } h = \sum_{k \in \mathbb{Z}} \alpha_k e_k.$$

(ii) Il existe $0 < A \leq B < +\infty$ tels que pour tout $h \in L^2$, $A \|\alpha\|_{l^2} \leq \|h\|_2 \leq B \|\alpha\|_{l^2}$ où $\|\alpha\|_{l^2} = \left(\sum_{k \in \mathbb{Z}} |\alpha_k|^2\right)^{\frac{1}{2}}$.

Remarque 4.8 Faisons une remarque à propos de la définition d'une base de Riesz. Une base de Riesz est donc un système générateur, libre et, d'une certaine manière, la propriété (ii) généralise la notion d'orthogonalité. En particulier, dans le cas d'une base orthonormée (base hilbertienne) on a $A = B = 1$ et (ii) est alors simplement l'égalité de Parseval. La donnée de la base de Riesz de L^2 revient à la donnée d'un isomorphisme entre l'espace de fonctions $L^2(\mathbb{R})$ et l'espace des suites $l^2(\mathbb{Z})$.

Remarque 4.9 Maintenant faisons une remarque sur la définition d'analyse multirésolution. Le (1) assure l'existence de la projection orthogonale de f sur chacun des espaces V_j , projection qui approche f . Le (2) traduit l'emboîtement des espaces et l'amélioration de l'approximation lorsque j décroît. Le (3) assure que la suite $\{V_j\}$ converge vers L^2 tout entier et donc que la suite des projections converge vers f . Le (4) caractérise les aspects multirésolution de la suite M et joue un rôle crucial dans la construction des bases d'ondelettes.

c - Interprétation géométrique de la construction

A partir de la famille M , on définit une deuxième famille de sous-espaces notés $\{W_j\}$ où W_j est le supplémentaire orthogonal de V_j dans V_{j-1} , à savoir $V_{j-1} = V_j \oplus W_j$

avec $W_j \perp V_j$. Par opposition aux espaces $\{V_j\}$ qui sont les espaces d'approximation, on dira que les espaces $\{W_j\}$ sont les espaces de détails.

On obtient alors une série de propriétés sur ces espaces $\{W_j\}$ utiles à la compréhension géométrique de la construction :

$$(6) \quad w(t) \in W_j \iff w(2t) \in W_{j-1};$$

$$(7) \quad W_j \perp W_k \text{ pour } j \neq k;$$

$$(8) \quad W_j \perp V_k \text{ pour } j \leq k;$$

$$(9) \quad V_j = V_k \oplus W_k \oplus V_{k-1} \oplus W_{k-1} \oplus \cdots \oplus W_{j+1} \text{ pour } j \leq k;$$

$$(10) \quad V_j = \bigoplus_{i=j+1}^{+\infty} W_i;$$

$$(11) \quad L^2(\mathbb{R}) = V_j \oplus \left(\bigoplus_{i=-\infty}^j W_i \right);$$

$$(12) \quad L^2(\mathbb{R}) = \bigoplus_{i=-\infty}^{+\infty} W_i.$$

Remarque 4.10 (11) indique qu'un élément de L^2 peut s'écrire sous la forme d'une somme orthogonale d'une approximation grossière et d'une infinité de détails plus fins.

(12) exprime le fait que toute fonction de L^2 est une somme infinie de détails orthogonaux. Notons $A^j = P_{V_j} f$ et $D^j = P_{W_j} f$ les projections orthogonales de $f \in L^2$ sur les espaces V_j et W_j . On a alors $A^{j-1} = A^j + D^j$ et $A^j \perp D^j$.

Les espaces $\{V_j\}$ sont les espaces d'approximation au sens suivant : A^j converge vers f dans L^2 lorsque j tend vers $-\infty$. De même les espaces $\{W_j\}$ sont les espaces de détails au sens où, dans L^2 , on a d'une part $D^j \rightarrow 0$ lorsque j tend vers $-\infty$ et d'autre part $f = A^j + \sum_{-\infty}^j D^j$. Autrement dit, pour un niveau d'approximation j fixé, les D^j sont les corrections à ajouter à l'approximation pour retrouver f .

d - Notion de base orthonormées d'ondelettes

On peut à présent énoncer le résultat fondamental associé aux analyses multirésolution. On note

$$f_{j,k} = 2^{-\frac{j}{2}} f(2^{-j}t - k)$$

pour f fonction quelconque.

Théorème 4.11 (*Base orthonormée d'ondelettes*) Soit M une analyse multirésolution de $L^2(\mathbb{R})$. A partir de la fonction g donnée en (5), on peut construire une fonction d'échelle φ puis une ondelette ψ telles que

$$\forall J \in \mathbb{Z}, \left\{ \left\{ \varphi_{j,k} \right\}_{k \in \mathbb{Z}}, \left\{ \psi_{j,k} \right\}_{j,k \in \mathbb{Z}, j \leq J} \right\} \text{ est une base orthonormée de } L^2,$$

et $\left\{ \psi_{j,k} \right\}_{j,k \in \mathbb{Z}}$ est une base orthonormée d'ondelettes de L^2 .

Remarque 4.12 1) Si M est une analyse multirésolution, il existe une infinité de fonctions d'échelle φ et donc une infinité d'ondelettes associées engendrant la même analyse.

2) D'autre part, il existe des bases orthonormées d'ondelettes de L^2 (c'est-à-dire de la forme $\left\{ \psi_{j,k} \right\}_{j,k \in \mathbb{Z}}$) qui ne sont pas associées à une analyse multirésolution. Elles ne sont donc pas associées à une fonction d'échelle. En revanche dès que ψ est suffisamment régulière, il existe une analyse multirésolution sous-jacente.

3) On peut rassembler sous la forme d'un tableau les éléments clés de l'analyse multirésolution.

	Fonctions	Espaces	Bases	$j \nearrow$	$j \searrow$
Approximations	Fonction d'échelle φ	V_j	$\left\{ \varphi_{j,k} \right\}_{k \in \mathbb{Z}}$	Plus grossiers	Plus fins
Détails	Ondelette ψ	W_j	$\left\{ \psi_{j,k} \right\}_{k \in \mathbb{Z}}$		

Figure 27

e - La fonction d'échelle et l'ondelette

On présente ici deux résultats établissant les liens entre les notions d'analyse multirésolution et d'ondelettes orthogonales, et proposant une manière de construire la seconde à partir de la première.

Proposition 4.13 *On considère la fonction d'échelle φ définie à l'aide de sa transformée de Fourier $\hat{\varphi}$ par*

$$(*) \hat{\varphi}(\omega) = \frac{\hat{g}(\omega)}{(\sum_{k \in \mathbb{Z}} |\hat{g}(\omega + k)|^2)^{\frac{1}{2}}}.$$

Alors

$$(1) \quad \varphi \in V_0.$$

$$(2) \quad \{\varphi_{0,k} = \varphi(t - k)\}_{k \in \mathbb{Z}} \text{ est une base orthonormée de } V_0.$$

$$(3) \quad \text{Equation aux deux échelles pour } \varphi : \exists ! a = (a_k)_{k \in \mathbb{Z}}, a \in l^2(\mathbb{Z}), \frac{1}{2}\varphi\left(\frac{t}{2}\right) = \sum_{k \in \mathbb{Z}} a_k \varphi(t - k)$$

dans L^2 .

$$(4) \quad m_0(\omega) = \sum_{k \in \mathbb{Z}} a_k e^{-2i\pi k \omega} \text{ est périodique de période } 1, m_0 \in l^2(0, 1)$$

$$(5) \quad \text{Plus généralement } \forall j \in \mathbb{Z}, \{\varphi_{j,k} = 2^{-\frac{j}{2}}\varphi(2^{-j}t - k)\}_{k \in \mathbb{Z}} \text{ est une base orthonormée de } V_j.$$

• **Remarques et commentaires :**

La relation (*) définit la fonction d'échelle φ à partir de g dans le domaine fréquentiel et conduit à une orthonormalisation dans le domaine temporel.

Les deux premières propriétés montrent qu'il s'agit d'un changement de base dans l'espace V_0 .

On déduit (3) du fait que $\varphi(\frac{t}{2}) \in V_1, V_1 \subset V_0, \{\varphi(t - k)\}_{k \in \mathbb{Z}}$ est une base de V_0 .

• **Construction de l'ondelette**

L'ondelette ψ est définie à l'aide de sa transformée de Fourier $\hat{\psi}$. Soit ρ une fonction périodique de période $\frac{1}{2}, |\rho(\omega)| = 1$ pour presque tout $\omega \in \mathbb{R}$, posons

$$m_1(\omega) = \rho(\omega) e^{-2i\pi\omega} \overline{m_0\left(\omega + \frac{1}{2}\right)},$$

et définissons

$$\hat{\psi}(\omega) = m_1\left(\frac{\omega}{2}\right) \hat{\varphi}\left(\frac{\omega}{2}\right).$$

Alors

$$(1) \quad \psi \in W_0;$$

(2) $\{\psi_{0,k} = \psi(t - k)\}_{k \in \mathbb{Z}}$ est une base orthonormée de W_0 ;

(3) Equation aux deux échelles pour ψ

$$\exists ! b = \{b_k\}_{k \in \mathbb{Z}}, b \in l^2(\mathbb{Z}) \text{ tel que } m_1(\omega) = \sum_{k \in \mathbb{Z}} b_k e^{-2i\pi k \omega}$$

et

$$\frac{1}{2}\psi\left(\frac{t}{2}\right) = \sum_{k \in \mathbb{Z}} b_k \varphi(t - k), \quad \text{dans } L^2.$$

(4) m_1 est périodique de période 1, $m_1 \in L^2(0, 1)$

(5) Plus généralement, pour tout $j \in \mathbb{Z}$, $\{\psi_{j,k} = 2^{-\frac{j}{2}}\psi(2^{-j}t - k)\}_{k \in \mathbb{Z}}$ est une base orthonormée de W_j .

(6) $\{\psi_{j,k}\}_{k \in \mathbb{Z}}$ est une base orthonormée de L^2 .

Faisons quelques remarques et commentaires sur cette construction.

La définition de $\hat{\psi}$ définit l'ondelette à partir de la fonction d'échelle φ dans le domaine fréquentiel au moyen d'un filtre m_1 déduit de m_0 .

Les deux premières propriétés affirment que l'on fabrique ainsi des fonctions de l'espace de détails W_0 et, en considérant toutes les translations entières, $\{\psi(t - k)\}_{k \in \mathbb{Z}}$ est une base orthonormée de W_0 .

La troisième propriété se déduit quant à elle de $\psi(\frac{t}{2}) \in W_1$, et de l'inclusion $W_1 \subset V_0$ et du fait que $\{\psi(t - k)\}_{k \in \mathbb{Z}}$ est une base de V_0 . Elle donne lieu à une seconde équation aux deux échelles définissant l'ondelette. Cette relation est la contrepartie dans le domaine temporel de la relation qui définit m_1 .

5 Conclusion

Nous avons vu en quoi les ondelettes et les gaborettes constituent une alternative à l'analyse de Fourier traditionnelle. En effet nous disposons à présent de méthode de décomposition des signaux en fonctions élémentaires, engendrées par des transformations simples d'une fonction de base. La fonction de base est soit déplacée et modulé (dans le cas des Gaborettes), soit translatée et dilatée (ondelettes). La différence fondamentale entre les deux tient précisément à cette opération de dilatation : les ondelettes s'adaptent d'elles mêmes à la taille des caractéristiques qu'elles recherchent grâce à l'analyse multirésolution qui suggère une interprétation différente de l'analyse par ondelettes, basée sur les idées de lissage, ou d'approximation des fonctions.

Il s'avère ainsi que les différents outils que nous avons étudiés ici sont davantage complémentaires que concurrents, et l'analyse de Fourier classique peut en effet être plus performante sur certains domaines.

Nous verrons lors de la soutenance les différentes applications à l'analyse sonore et en particulier comment les images tirées de cette analyse nous permettent de comprendre intimement le son étudié.

Références

- [1] Daniel Arfib. Des outils mathématiques au service du son. Mathématiques et Musique. Musique et sciences .1997.
- [2] Daniel Arfib, Richard Kronland-Martinet. Transformer le son : Modeler, modéliser. Cahier de l'Ircam numéro 2.1993
- [3] Barbara Burke Hubbard. Ondes et ondelettes, la saga d'un outil mathématique. Pour la science 1995.
- [4] Richard Kronland-Martinet. The wavelet transform for analysis synthesis and processing of speech and music sound. Computer Music Journal 1988.
- [5] Claude Gasquet et Patrick Witomski. Analyse de Fourier et applications. Dunod 1990.
- [6] Stéphane Mallat. Une exploration des signaux en ondelettes. Editions de l'école Polytechnique. 2000.
- [7] Yves Meyer. Ondelettes et opérateurs I . Editions Hermann 1989.
- [8] Curtis Roads, Stephen Travis Pope, Aldo Picciali, Giovanni De Poli. Musical signal processing. Swets ans Zeitlingers Publishers. 1997.
- [9] Bruno Torrèsani. Analyse continue par ondelettes. Interdition CNRS 1995.
- [10] Pierre-Gilles Lemarié-Rieusset, Jean-Pierre Kahane. Srie de Fourier et ondelettes. Cassini 1998.
- [11] Dave Benson. [http ://www.math.uga.edu/ djb/html/math-music.html](http://www.math.uga.edu/~djb/html/math-music.html)