

Analyse harmonique appliquée à la reconnaissance musicale

Andry Rafaralahy

Table des matières

Introduction	3
1 Séries de Fourier	3
1.1 Rappels de théorie de la mesure	3
1.2 Polynômes trigonométriques et séries de Fourier	4
1.3 L'espace $L_P^2(0, a)$	5
1.4 Convergence des séries de Fourier	11
1.4.1 Préliminaires	11
1.4.2 Convergence ponctuelle	15
1.4.3 Convergence uniforme	16
1.4.4 Convergence dans $L_P^2(0, a)$	17
1.5 Représentation spectrale d'une fonction périodique	19
2 Transformée de Fourier discrète	21
2.1 Échantillonnage	21
2.2 Définition de la transformée de Fourier discrète	22
2.3 Quelques propriétés	24
2.4 Transformée de Fourier rapide	25
2.4.1 Principe	25
2.4.2 Coût de l'algorithme	26
2.5 Représentation spectrale d'un signal échantillonné	27
2.5.1 Définitions	27
2.5.2 Exemple	27
3 Application à la reconnaissance musicale	29
3.1 Empreinte acoustique	30
3.2 Spectrogrammes et constellations	30
3.3 Organisation des données	32
3.4 Processus d'identification	33
3.5 Implémentation MATLAB	34

Introduction

L'objectif de ce TIPE est de concevoir un algorithme de reconnaissance musicale. Il doit permettre, en analysant un court échantillon musical (moins de vingt secondes) d'identifier le morceau d'où il est extrait, comme le fait l'application Shazam par exemple. Pour cela nous allons dans un premier temps introduire des outils mathématiques : les séries de Fourier et la transformée de Fourier discrète. Puis, en nous inspirant de ?, nous pourrons construire tel algorithme.

1 Séries de Fourier

Nous voulons décomposer un signal périodique s en une somme (éventuellement infinie) de fonction trigonométriques dont les fréquences soient multiples de la fréquence de s . Cette décomposition, si elle existe, donne ce que l'on appelle la représentation spectrale du signal (voir 1.5). Une telle représentation a de nombreux intérêts pratiques. Nous verrons notamment que dans le cas d'un signal sonore elle peut servir à extraire certaines caractéristiques audibles.

1.1 Rappels de théorie de la mesure

Soit $I \subset \mathbb{R}$ un intervalle.

Définition 1.1.1. On note $\mathcal{M}(I, \mathbb{R})$ l'ensemble des fonctions f de I à valeurs \mathbb{R} mesurables (au sens $(I, \mathcal{B}(I)) - (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mesurables, avec $\mathcal{B}(I)$ et $\mathcal{B}(\mathbb{R})$ les tribus boréliennes sur respectivement I et \mathbb{R} .)

On note $\mathcal{M}(I, \mathbb{R})_+$ les fonctions mesurables positives, c'est-à-dire les fonctions $f \in \mathcal{M}(I, \mathbb{R})$ telles que $f \geq 0$.

Proposition 1.1.1 (Convergence monotone). Soit $(f_n)_{n \in \mathbb{N}}$ une suite d'éléments de $\mathcal{M}(I, \mathbb{R})_+$ telle que pour tout $n \in \mathbb{N}$ et $x \in I$

$$f_{n+1}(x) \geq f_n(x)$$

Alors, en posant pour tout $x \in I$, $f(x) = \lim_{n \rightarrow \infty} f_n(x) \in \overline{\mathbb{R}}_+$, on a $f \in \mathcal{M}(I, \mathbb{R})_+$ et

$$\int_I f_n(t) dt \xrightarrow{n \rightarrow +\infty} \int_I f(t) dt$$

Définition 1.1.2. Soit $p \in [1, +\infty[$. L'ensemble $L^p(I)$ est l'ensemble des fonctions $f: I \rightarrow \mathbb{C}$ vérifiant

$$\int_I |f(t)|^p dt < +\infty$$

Remarque. Si $f, g \in L_P^2(0, a)$, on dira que $f = g$ si elles sont égales partout sauf sur un ensemble de mesure nulle.

Proposition 1.1.2. L'application $\|\cdot\|_p : L^p(I) \rightarrow \mathbb{R}$ définie par

$$\|f\|_p = \left(\int_I |f(t)|^p dt \right)^{1/p}$$

est une norme sur $L^p(I)$, en tenant compte de la remarque 1.1.

Proposition 1.1.3. L'espace $L^p(I)$ muni de la norme $\|\cdot\|_p$ est un espace vectoriel normé.

Proposition 1.1.4 (Convergence dominée dans $L^1(I)$). Soit $(f_n)_{n \in \mathbb{N}}$ une suite d'éléments de $L^1(I)$ telle que

1. $f_n \rightarrow f$ presque partout
2. $\exists F \in L^1(I)$ telle que $|f_n| \leq F$ presque partout pour tout $n \in \mathbb{N}$

Alors $f_n \xrightarrow{n \rightarrow +\infty} f$ dans $L^1(I)$ c'est-à-dire $\|f_n - f\|_1 \rightarrow 0$ quand $n \rightarrow +\infty$. En particulier on a

$$\int_I f_n(t) dt \xrightarrow{n \rightarrow +\infty} \int_I f(t) dt$$

Corollaire 1.1.1 (Convergence dominée dans $L^p(I)$). Soit $(f_n)_{n \in \mathbb{N}}$ une suite d'éléments de $L^p(I)$ telle que

1. $f_n \rightarrow f$ presque partout
2. $\exists F \in L^p(I)$ telle que $|f_n| \leq F$ presque partout pour tout $n \in \mathbb{N}$

Alors $f_n \xrightarrow{n \rightarrow +\infty} f$ dans $L^p(I)$ c'est-à-dire $\|f_n - f\|_p \rightarrow 0$ quand $n \rightarrow +\infty$.

On redémontrera le cas $p = 2$ (voir 1.3.5).

Définition 1.1.3. Soient I un intervalle et $p \in \mathbb{N}^* \cup \{+\infty\}$. On note $\mathcal{C}_c^p(I)$ l'ensemble des fonctions de classe \mathcal{C}^p à support compact.

Proposition 1.1.5. Soit $p \in [1, +\infty]$. L'ensemble $\mathcal{C}_c^\infty(I)$ est dense dans $L^p(I)$.

Corollaire 1.1.2. Soit $p \in [1, +\infty]$. L'ensemble $\mathcal{C}_c^1(I)$ est dense dans $L^p(I)$.

Proposition 1.1.6 (Inégalité de Jensen). Soient J un intervalle, $g: [0, 1] \rightarrow J$ une fonction appartenant à $L^1([0, 1])$ et $\varphi: J \rightarrow \mathbb{R}$ une fonction convexe. Alors

$$\varphi \left(\int_0^1 f(x) dx \right) \leq \int_0^1 \varphi(f(x)) dx$$

1.2 Polynômes trigonométriques et séries de Fourier

Soit $a \in \mathbb{R}_+^*$. Pour tout $n \in \mathbb{Z}$, la fonction $e_n(t) = e^{2i\pi n \frac{t}{a}}$ est périodique de période a . Il en est de même pour toute combinaison linéaire de la forme :

$$p = \sum_{n \in I} c_n e_n$$

où I est un ensemble fini et les $(c_n)_{n \in I}$ sont des coefficients complexes. Dans le cas où :

$$p = \sum_{n=-N}^{+N} c_n e_n$$

on dit que p est un polynôme trigonométrique. Les polynômes trigonométriques sont de carré sommable sur $[0, a]$ c'est-à-dire que :

$$\int_0^a |p(t)|^2 dt < +\infty$$

En effet, pour tout $n \in \mathbb{Z}$ on a :

$$|e_n(t)| = |e^{2i\pi n \frac{t}{a}}| = 1$$

d'où

$$\int_0^a |e_n(t)|^2 dt = a$$

donc les e_n sont de carré sommable. Par linéarité de l'intégrale on en déduit que les polynômes trigonométriques le sont aussi. Introduisons maintenant la notation suivante :

$$L_P^2(0, a) = \{f: \mathbb{R} \rightarrow \mathbb{C} \text{ de période } a \text{ et de carré sommable sur } [0, a]\}$$

Nous venons de montrer que les e_n ainsi que polynômes trigonométriques appartiennent à $L_P^2(0, a)$. En reprenant les notations de la définition 1.1.2 on peut identifier $L_P^2(0, a)$ à $L^2([0, a])$. De ce fait, $L_P^2(0, a)$ hérite des propriétés de L^2 que nous connaissons déjà.

On peut généraliser la notion de polynôme trigonométrique en utilisant \mathbb{Z} comme ensemble d'indices, on obtient des sommes de la forme :

$$\sum_{n=-\infty}^{+\infty} c_n e_n$$

que l'on appelle séries de Fourier. On précisera le sens de cette somme infinie dans les sections suivantes.

1.3 L'espace $L_P^2(0, a)$

L'espace $L_P^2(0, a)$ joue en rôle important dans la théorie des séries de Fourier. Le résultat principal que nous allons démontrer est qu'il existe une norme sur $L_P^2(0, a)$ telle que toute fonction $f \in L_P^2(0, a)$ soit développable en série de Fourier, c'est-à-dire qu'il existe des coefficients $(c_n)_{n \in \mathbb{Z}}$ tels que

$$f = \sum_{n=-\infty}^{+\infty} c_n e_n \tag{1}$$

c'est-à-dire que la suite des sommes partielles $S_N = \sum_{n=-N}^N c_n e_n$ converge vers f pour une certaine norme. Plus précisément, nous allons montrer que :

1. $L_P^2(0, a)$ est un espace de Hilbert pour un certain produit scalaire $\langle \cdot, \cdot \rangle$
2. les $(e_n)_{n \in \mathbb{Z}}$ forment une base hilbertienne de cet espace
3. on peut calculer explicitement les c_n de la formule (1) :

$$\forall n \in \mathbb{Z}, c_n = \langle f, e_n \rangle$$

Proposition 1.3.1. $L_P^2(0, a)$ est un \mathbb{C} -espace vectoriel.

Remarque. Ce résultat fait partie des rappels de la section 1.1.

Démonstration. Il est évident que l'ensemble des fonctions $f: \mathbb{R} \rightarrow \mathbb{C}$ de période a est un \mathbb{C} -espace vectoriel.

Montrons que $L_P^2(0, a)$ en est un sous-espace vectoriel. On vérifie facilement que la fonction nulle appartient bien à $L_P^2(0, a)$. Prenons $f, g \in L_P^2(0, a)$ et $\lambda \in \mathbb{C}$ et montrons que $\lambda f + g \in L_P^2(0, a)$. On a

$$\begin{aligned} \int_0^a |\lambda f(t) + g(t)|^2 dt &= \int_0^a \lambda^2 |f(t)|^2 + |g(t)|^2 + 2|f(t)||g(t)|dt \\ &= \lambda^2 \int_0^a |f(t)|^2 dt + \int_0^a |g(t)|^2 dt + 2 \int_0^a |f(t)||g(t)|dt \end{aligned}$$

Il suffit de montrer que

$$\int_0^a |f(t)||g(t)|dt < +\infty$$

On peut majorer $|f||g|$ comme suit

$$|f||g| \leq \frac{f^2 + g^2}{2}$$

et on en déduit le résultat. \square

Proposition 1.3.2. Si f et g sont deux fonctions de $L_P^2(0, a)$ alors $f\bar{g}$ est intégrable sur $[0, a]$, c'est-à-dire que

$$\int_0^a |f(t)\bar{g}(t)|dt < +\infty$$

Démonstration. En reprenant la majoration de la démonstration précédente

$$|f\bar{g}| = |f||g| \leq \frac{f^2 + g^2}{2}$$

D'où

$$\int_0^a |f(t)\bar{g}(t)|dt < +\infty$$

\square

Proposition 1.3.3. *L'application suivante est un produit scalaire sur $L_P^2(0, a)$:*

$$\begin{aligned}\langle \cdot, \cdot \rangle : L_P^2(0, a) &\rightarrow \mathbb{C} \\ (\langle f, g \rangle) &\mapsto \frac{1}{a} \int_0^a f(t) \overline{g(t)} dt\end{aligned}$$

Démonstration. L'application est bien définie en vertu de la proposition précédente. La linéarité à gauche découle directement de la linéarité de l'intégrale. Montrons que $\langle \cdot, \cdot \rangle$ est hermitienne. Soient $f, g \in L_P^2(0, a)$.

$$\begin{aligned}\overline{\langle f, g \rangle} &= \overline{\frac{1}{a} \int_0^a f(t) \overline{g(t)} dt} \\ &= \frac{1}{a} \int_0^a \overline{f(t) \overline{g(t)}} dt \\ &= \frac{1}{a} \int_0^a \overline{f(t)} g(t) dt \\ &= \langle g, f \rangle\end{aligned}$$

La semi-linéarité à droite peut se déduire de la linéarité à gauche et de l'hermitianité. Reste à montrer que $\langle \cdot, \cdot \rangle$ est définie positive. Pour tout $f \in L_P^2(0, a)$ on a

$$\begin{aligned}\langle f, f \rangle &= \frac{1}{a} \int_0^a f(t) \overline{f(t)} dt \\ &= \frac{1}{a} \int_0^a |f(t)|^2 dt \geq 0\end{aligned}$$

Comme $|f|^2 \geq 0$ alors

$$\frac{1}{a} \int_0^a |f(t)|^2 dt = 0 \implies |f|^2 = 0 \implies f = 0$$

Les deux dernières égalités sont des égalités "presque partout" (voir remarque 1.1).

On a bien montré que $\langle \cdot, \cdot \rangle$ est un produit scalaire sur $L_P^2(0, a)$. \square

On peut définir la notion de série de Fourier associée à une fonction.

Définition 1.3.1. *Soit $f \in L_P^2(0, a)$. On appelle coefficients de Fourier de f les nombres complexes définis par*

$$\forall n \in \mathbb{Z}, c_n(f) = \langle f, e_n \rangle = \frac{1}{a} \int_0^a f(t) \overline{e_n(t)} dt$$

On appelle série de Fourier associée à f la série de Fourier

$$\sum_{n \in \mathbb{Z}} c_n(f) e_n$$

À partir de maintenant $\|\cdot\|_2$ désigne la norme associée à ce produit scalaire.

Remarque. La norme $\|\cdot\|_2$ définie ici coïncide avec celle de 1.1.2 en identifiant $L_P^2(0, a)$ à $L^2([0, a])$.

Pour montrer que $(L_P^2(0, a), \langle \cdot, \cdot \rangle)$ est un espace de Hilbert, il faut montrer qu'il est complet. Nous utiliserons la caractérisation des espaces normés complets suivante :

Proposition 1.3.4. *Un espace normé $(E, \|\cdot\|)$ est complet si et seulement si toute série normalement convergente y est convergente.*

Démonstration. Si E est complet, alors soit $\sum x_n$ une série normalement convergente dans E . On pose, pour tout $n \in \mathbb{N}$, $S_n = \sum_{k=0}^n x_k$ et $T_n = \sum_{k=0}^n \|x_k\|$. T_n converge dans \mathbb{R} donc elle est de Cauchy. Soit $\epsilon > 0$, il existe $N \in \mathbb{N}$ tel que pour tous $m \geq n \geq N$ on ait

$$\|S_m - S_n\| = \left\| \sum_{k=n+1}^m x_k \right\| \leq \sum_{k=n+1}^m \|x_k\| = |T_m - T_n| < \epsilon$$

donc $(S_n)_{n \in \mathbb{N}}$ est de Cauchy dans E . On en déduit que la série $\sum x_n$ converge dans E .

Réiproquement, supposons que dans E toute série normalement convergente soit convergente. Soit $(x_n)_{n \in \mathbb{N}}$ une suite de Cauchy de E . Il existe une sous-suite $(x_{\varphi(n)})$ telle que

$$\|x_{\varphi(n+1)} - x_{\varphi(n)}\| < 2^{-n}, \forall n \in \mathbb{N}$$

En effet, on peut construire une telle sous-suite en prenant

$$\begin{aligned} \varphi(1) &\geq 1 & \text{tel que } \|x_{\varphi(1)} - x_{\varphi(1)}\| &< \frac{1}{2}, & \forall n \geq \varphi(1) \\ \varphi(2) &\geq \varphi(1) & \text{tel que } \|x_{\varphi(2)} - x_{\varphi(1)}\| &< \frac{1}{4}, & \forall n \geq \varphi(2) \\ &\text{etc.} \end{aligned}$$

On a alors que la série de terme général $x_{\varphi(n+1)} - x_{\varphi(n)}$ est absolument convergente, donc convergente par hypothèse. Par conséquent, la suite $(x_{\varphi(n)})$ est convergente et donc la suite de Cauchy (x_n) l'est aussi. \square

Rappelons un théorème important de théorie de la mesure, qui sera d'une grande aide pour démontrer la complétude de $L_P^2(0, a)$.

Proposition 1.3.5 (Convergence dominée dans $L_P^2(0, a)$). *Soit $(f_n)_{n \in \mathbb{N}}$ une suite d'éléments de $L_P^2(0, a)$ telle que*

1. $f_n \rightarrow f$ presque partout
2. $\exists F \in L_P^2(0, a)$ telle que $|f_n| \leq F$ presque partout pour tout $n \in \mathbb{N}$

Alors $f_n \xrightarrow{n \rightarrow +\infty} f$ dans $L_P^2(0, a)$ c'est-à-dire $\|f_n - f\|_2 \rightarrow 0$ quand $n \rightarrow +\infty$.

Démonstration. Pour tout $n \in \mathbb{N}$ posons $h_n = |f_n - f|^2$. On sait que $f_n \rightarrow f$ presque partout, donc $h_n \xrightarrow{n \rightarrow +\infty} 0$ presque partout. Puis on remarque que

$$\begin{aligned} 0 \leq |h_n| &\leq (|f_n| + |f|)^2 \\ &\leq |f_n|^2 + |f|^2 + 2|f_n||f| \\ &\leq 2(|f_n|^2 + |f|^2) \\ &\leq 4F^2 \text{ presque partout} \end{aligned}$$

$4F^2$ est intégrable sur $[0, a]$ puisque F appartient à $L_P^2(0, a)$ donc en appliquant le théorème de convergence dominée pour les fonctions intégrables on obtient

$$\|f_n - f\|_2^2 = \frac{1}{a} \int_0^a |f_n(t) - f(t)|^2 dt = \frac{1}{a} \int_0^a |h_n(t)| dt \xrightarrow{n \rightarrow +\infty} 0$$

□

Proposition 1.3.6. $(L_P^2(0, a), \|\cdot\|_2)$ est un espace complet.

Démonstration. Il suffit de montrer que dans $L_P^2(0, a)$, toute série normalement convergente est convergente. Soit $\sum f_n$ une série normalement convergente de $L_P^2(0, a)$. Notons $A = \sum_{n=0}^{+\infty} \|f_n\|_2$. On pose pour tout $x \in [0, a]$, $g_n(x) = \sum_{k=0}^n |f_k(x)|$. On vérifie que g_n est mesurable positive pour tout $n \in \mathbb{N}$. Soit $f: [0, a] \rightarrow \mathbb{R}$ définie par $F(x) = \lim_{n \rightarrow +\infty} g_n(x)$. On a que la suite (g_n^2) tend en croissant vers F^2 donc par le théorème de convergence monotone

$$\int_0^a g_n(t)^2 dt \xrightarrow{n \rightarrow +\infty} \int_0^a F(t)^2 dt \tag{2}$$

On peut majorer la norme des g_n

$$\begin{aligned} \|g_n\|_2 &= \left\| \sum_{k=0}^n |f_k| \right\|_2 \leq \sum_{k=0}^n \|f_k\|_2 \\ &\leq \sum_{n=0}^{+\infty} \|f_n\|_2 = A < +\infty \end{aligned}$$

On a donc $\|g_n\|_2^2 \leq A^2$ pour tout $n \in \mathbb{N}$ et (2) nous donne

$$\frac{1}{a} \int_0^a |g_n(t)|^2 dt \leq A^2$$

donc en passant à la limite

$$\frac{1}{a} \int_0^a |F(t)|^2 dt \leq A^2 < +\infty \tag{3}$$

Cette dernière inégalité montre que $F < +\infty$ presque partout donc la série $\sum f_n(x)$ est absolument convergente dans \mathbb{R} donc convergente pour presque tout $x \in [0, a]$. On peut définir, pour presque tout x

$$f(x) = \sum_{n=0}^{+\infty} f_n(x)$$

On a $|\sum_{k=0}^n f_k| \leq g_n \leq F$ et $\sum_{k=0}^n f_k \xrightarrow{n \rightarrow +\infty} f$ presque partout donc $|f| \leq F$. L'inégalité (3) montre que f appartient à $L_P^2(0, a)$ donc f appartient aussi à $L_P^2(0, a)$.

Enfin, en appliquant le théorème de convergence dominée dans $L_P^2(0, a)$ à la suite $(\sum_{k=0}^n f_k)_{n \in \mathbb{N}}$, dont on vient de vérifier les hypothèses, on obtient que $\sum_{k=0}^n f_k \xrightarrow{n \rightarrow +\infty} f$ dans $L_P^2(0, a)$. \square

On a ainsi montré que $(L_P^2(0, a), \langle \cdot, \cdot \rangle)$ est un espace de Hilbert.

Proposition 1.3.7. *La famille $(e_n)_{n \in \mathbb{Z}}$ est orthonormée.*

Démonstration. Pour $k, l \in \mathbb{Z}$ on a

$$\begin{aligned} \langle e_k, e_l \rangle &= \frac{1}{a} \int_0^a e^{2i\pi k \frac{t}{a}} e^{-2i\pi l \frac{t}{a}} dt \\ &= \frac{1}{a} \int_0^a e^{2i\pi(k-l) \frac{t}{a}} dt \\ &= \begin{cases} \frac{1}{a} \int_0^a 1 dt & \text{si } k = l, \\ 0 & \text{si } k \neq l \end{cases} \\ &= \delta_{k,l} \end{aligned}$$

\square

Corollaire 1.3.1. *La famille $(e_n)_{n \in \mathbb{Z}}$ est libre.*

Démonstration. Soit $(e_j)_{j \in J}$ une sous-famille finie de $(e_n)_{n \in \mathbb{Z}}$ et $(\lambda_j)_{j \in J}$ tels que

$$\sum_{j \in J} \lambda_j e_j = 0$$

Pour tout $j \in J$ on a

$$\lambda_j = \left\langle \sum_{j' \in J} \lambda_{j'} e_{j'}, e_j \right\rangle = \langle 0, e_j \rangle = 0$$

donc la famille $(e_j)_{j \in J}$ est libre.

On a montré que toute sous-famille finie de $(e_n)_{n \in \mathbb{Z}}$ est libre, donc la famille $(e_n)_{n \in \mathbb{Z}}$ est libre. \square

Proposition 1.3.8 (Inégalité de Bessel). Soit $f \in L_P^2(0, a)$. On a l'inégalité suivante

$$\sum_{n=-\infty}^{+\infty} |c_n(f)| \leq \|f\|_2^2$$

Démonstration. Soit $N \in \mathbb{N}$. D'abord, montrons que $f - \sum_{n=-N}^N c_n(f)e_n$ et $\sum_{n=-N}^N c_n(f)e_n$ sont orthogonaux.

$$\begin{aligned} \left\langle f - \sum_{k=-N}^N c_k(f)e_k, \sum_{n=-N}^N c_n(f)e_n \right\rangle &= \sum_{m=-N}^N \overline{c_m(f)} \left\langle f - \sum_{k=-N}^N c_k(f)e_k, e_m \right\rangle \\ &= \sum_{m=-N}^N \overline{c_m(f)} (\langle f, e_m \rangle - c_m) \text{ par orthonormalité des } (e_n) \\ &= \sum_{m=-N}^N \overline{c_m(f)} (c_m - c_m) \\ &= 0 \end{aligned}$$

Ensuite par le théorème de Pythagore

$$\begin{aligned} \|f\|_2^2 &= \left\| f - \sum_{k=-N}^N c_k(f)e_k + \sum_{k=-N}^N c_k(f)e_k \right\|_2^2 \\ &= \left\| f - \sum_{k=-N}^N c_k(f)e_k \right\|_2^2 + \left\| \sum_{k=-N}^N c_k(f)e_k \right\|_2^2 \\ &= \left\| f - \sum_{k=-N}^N c_k(f)e_k \right\|_2^2 + \sum_{k=-N}^N |c_k(f)|^2 \end{aligned}$$

D'où

$$\sum_{k=-N}^N |c_k(f)|^2 \leq \|f\|_2^2$$

On en déduit

$$\sum_{k=-\infty}^{+\infty} |c_k(f)|^2 \leq \|f\|_2^2$$

en passant à la limite quand $N \rightarrow +\infty$. □

1.4 Convergence des séries de Fourier

1.4.1 Préliminaires

Commençons par démontrer un théorème important de la théorie des séries de Fourier, le lemme de Riemann-Lebesgue. L'une des applications de ce théorème est que la suite des $c_n(f)$ des coefficients de Fourier d'une fonction f tend vers 0 lorsque $|n| \rightarrow +\infty$.

Proposition 1.4.1 (Lemme de Riemann-Lebesgue). *Soient I un intervalle de \mathbb{R} et $f: I \rightarrow \mathbb{C}$ une fonction intégrable sur I . Alors*

$$\lim_{|n| \rightarrow +\infty} \int_I f(t) e_n(t) dt = 0$$

Démonstration. On montre d'abord le résultat pour f appartenant à $\mathcal{C}_c^1(I)$. Remarquons déjà que si on note J le support de f , alors $\int_I f(t) e_n(t) dt = \int_J f(t) e_n(t) dt$. On peut se ramener au cas où $J = [a, b]$. Par intégration par parties, on a, pour tout $n \in \mathbb{N}$

$$\begin{aligned} \int_a^b f(t) e_n(t) dt &= \left[-\frac{ia}{2\pi n} f(t) e_n(t) \right]_b^a + \frac{ia}{2\pi n} \int_a^b f'(t) e_n(t) dt \\ &= \frac{ia}{2\pi n} f(a) e_n(a) - \frac{ia}{2\pi n} f(b) e_n(b) + \frac{ia}{2\pi n} \int_a^b f'(t) e_n(t) dt \end{aligned}$$

En passant au module, et en utilisant le fait que $\forall t \in I, |e_n(t)| = 1$ on trouve

$$\left| \int_a^b f(t) e_n(t) dt \right| \leq \frac{a}{2\pi n} \left(|f(a)| + |f(b)| + \sup_{[a,b]} |f'(t)| \right)$$

Puis comme f' est continue sur $[a, b]$ on a $\sup_{[a,b]} |f'(t)| < +\infty$ donc

$$\left| \int_a^b f(t) e_n(t) dt \right| \xrightarrow{n \rightarrow +\infty} 0$$

Montrons maintenant le cas général. Soit $f \in L^1(I)$. Par densité de $\mathcal{C}_c^1(I)$ dans $L^1(I)$ (voir 1.1.2), il existe une suite $(f_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{C}_c^1(I)$ telle que $\|f_n - f\|_1 \xrightarrow{n \rightarrow +\infty} 0$. Pour $k, n \in \mathbb{N}$ on a

$$\begin{aligned} \left| \int_I f(t) e_n(t) dt \right| &= \left| \int_I f(t) e_n(t) dt - \int_I f_k(t) e_n(t) dt + \int_I f_k(t) e_n(t) dt \right| \\ &\leq \left| \int_I f(t) e_n(t) dt - \int_I f_k(t) e_n(t) dt \right| + \left| \int_I f_k(t) e_n(t) dt \right| \\ &\leq \left| \int_I (f(t) - f_k(t)) e_n(t) dt \right| + \left| \int_I f_k(t) e_n(t) dt \right| \\ &\leq \int_I |(f(t) - f_k(t)) e_n(t)| dt + \left| \int_I f_k(t) e_n(t) dt \right| \\ &\leq \|(f - f_k)e_n\|_1 + \left| \int_I f_k(t) e_n(t) dt \right| \\ &\leq \|f - f_k\|_1 \|e_n\|_1 + \left| \int_I f_k(t) e_n(t) dt \right| \end{aligned} \tag{4}$$

On voit que la quantité $\|e_n\|_1$ est bornée car

$$\|e_n\|_1 = \int_I |e_n(t)| dt \leq \lambda(I) \sup_I |e_n(t)| \leq \lambda(I)$$

avec $\lambda(I)$ la longueur de l'intervalle I.

Soit $\epsilon > 0$. Comme $\|f - f_k\|_1 \xrightarrow{k \rightarrow +\infty} 0$ il existe un entier K tel que pour tout $k \geq K$ on ait

$$\|f - f_k\|_1 < \frac{\epsilon}{2\lambda(I)}$$

Pour tout $k \in \mathbb{N}$, f_k appartient à $\mathcal{C}_c^1(I)$ donc

$$\left| \int_I f_k(t) e_n(t) dt \right| \xrightarrow{n \rightarrow +\infty} 0$$

Par conséquent il existe un entier N tel que

$$\left| \int_I f_k(t) e_n(t) dt \right| < \frac{\epsilon}{2}$$

En reprenant la majoration (4), pour $k \geq K$ et $n \geq N$ on trouve

$$\left| \int_I f(t) e_n(t) dt \right| < \epsilon$$

Donc on a montré que $\int_I f(t) e_n(t) dt \xrightarrow{n \rightarrow +\infty} 0$. On montre le cas $n \rightarrow -\infty$ de manière similaire. \square

On introduit maintenant le noyau de Dirichlet. Cela nous permettra de simplifier l'expression sommes partielles de séries de Fourier (voir par exemple la démonstration de 1.4.6).

Définition 1.4.1 (Noyau de Dirichlet). *Pour $n \in \mathbb{N}$, on appelle noyau de Dirichlet la fonction $D_n : \mathbb{R} \rightarrow \mathbb{C}$ définie par*

$$D_n(t) = \begin{cases} \sum_{k=-n}^n e_k & \text{si } t \in \mathbb{R} \setminus a\mathbb{Z}, \\ 2n+1 & \text{sinon..} \end{cases}$$

Proposition 1.4.2. *Pour tout $n \in \mathbb{N}$, D_n est une fonction paire.*

Démonstration. Pour tout $k \in \llbracket -n, n \rrbracket$ on a $e_k(-t) = e_{-k}(t)$. \square

Proposition 1.4.3. *Pour tout $n \in \mathbb{N}$, D_n est intégrable sur $[0, a]$ et*

$$\int_0^a D_n(t) dt = a$$

Démonstration. L'intégrabilité vient du fait que D_n est égal presque partout à un polynôme trigonométrique. Pour tout $k \in \mathbb{Z} \setminus \{0\}$

$$\int_0^a e_k(t) dt = 0$$

et

$$\int_0^a e_0(t) dt = a$$

d'où le résultat en sommant les intégrales. \square

Proposition 1.4.4. Pour tout $n \in \mathbb{N}$ on a

$$\int_0^{a/2} D_n(t) dt = \frac{a}{2}$$

Démonstration. On a $\int_0^{a/2} e_0(t) dt = \frac{a}{2}$ et pour $k \in \mathbb{Z} \setminus \{0\}$

$$\int_0^{a/2} e_k(t) dt = \frac{ia}{2\pi k} (1 - e^{ik\pi})$$

Donc si k est pair différent de 0 on a $\int_0^{a/2} e_k(t) dt = 0$ et si k est impair $\int_0^{a/2} e_k(t) dt = -\int_0^{a/2} e_{-k}(t) dt$ donc dans la somme les termes impairs d'indices opposés s'annulent. Il ne reste que le terme d'indice 0 d'où

$$\int_0^{a/2} D_n(t) dt = \frac{a}{2}$$

□

Proposition 1.4.5. Pour tout $t \in \mathbb{R} \setminus a\mathbb{Z}$, on a

$$D_n(t) = \frac{\sin\left(\frac{(2n+1)\pi t}{a}\right)}{\sin\left(\frac{\pi t}{a}\right)}$$

Démonstration. Il suffit de calculer

$$\begin{aligned} D_n(t) &= \sum_{k=-n}^n e_k(t) \\ &= \sum_{k=-n}^n e^{2i\pi k \frac{t}{a}} \\ &= e^{-2i\pi n \frac{t}{a}} \sum_{k=0}^{2n} e^{2i\pi k \frac{t}{a}} \\ &= e^{-2i\pi n \frac{t}{a}} \frac{e^{2i\pi \frac{t}{a}(2n+1)} - 1}{e^{2i\pi \frac{t}{a}} - 1} \\ &= e^{-2i\pi n \frac{t}{a}} \frac{e^{i\pi \frac{t}{a}(2n+1)}(e^{i\pi \frac{t}{a}(2n+1)} - e^{-i\pi \frac{t}{a}(2n+1)})}{e^{i\pi \frac{t}{a}}(e^{i\pi \frac{t}{a}} - e^{-i\pi \frac{t}{a}})} \\ &= \frac{\sin\left(\frac{(2n+1)\pi t}{a}\right)}{\sin\left(\frac{\pi t}{a}\right)} \end{aligned}$$

□

1.4.2 Convergence ponctuelle

Proposition 1.4.6 (Théorème de Jordan-Dirichlet). *Soient $f: \mathbb{R} \rightarrow \mathbb{C}$ une fonction a -périodique, continue par morceaux, et $t_0 \in \mathbb{R}$ tels que la fonction*

$$h \rightarrow \frac{f(t_0 + h) - f(t_0 - h) - f(t_0+) - f(t_0-)}{h}$$

soit bornée au voisinage de 0. Alors $\sum_{n \in \mathbb{Z}} c_n(f) e_n(t_0)$ converge et on a

$$\sum_{n=-\infty}^{+\infty} c_n(f) e_n(t_0) = \frac{f(t_0+) + f(t_0-)}{2}$$

Démonstration. On reprend la démonstration de ?, p. 260, en adaptant pour les fonctions a -périodiques.

Quitte à effectuer la translation $t \mapsto t + t_0$, on peut supposer $t_0 = 0$. Pour tout $n \in \mathbb{N}$, on note $s_n = \sum_{k=-n}^n c_k(f)$. Il s'agit de montrer que la suite (u_n) définie par $u_n = s_n - \frac{f(0+) + f(0-)}{2}$ tend vers 0.

On a

$$\begin{aligned} as_n &= \sum_{k=-n}^n \int_{-a/2}^{a/2} f(t) \overline{e_k(t)} dt \\ &= \sum_{k=-n}^n \int_{-a/2}^{a/2} f(t) e_{-k}(t) dt \\ &= \int_{-a/2}^{a/2} f(t) D_n dt \end{aligned}$$

Par ailleurs, D_n est une fonction paire, donc

$$\int_{-a/2}^0 f(t) D_n(t) dt = \int_0^{a/2} f(-t) D_n(t) dt$$

d'où

$$as_n = \int_0^{a/2} (f(t) + f(-t)) D_n dt$$

Remarquons aussi que

$$a \frac{f(0+) + f(0-)}{2} = \int_0^{a/2} (f(0+) + f(0-)) D_n(t) dt \text{ d'après 1.4.4}$$

On en déduit finalement

$$\begin{aligned} as_n &= \int_0^{a/2} (f(t) + f(-t) - f(0+) - f(0-)) D_n(t) dt \\ &= \int_0^{a/2} g(t) \sin\left(\frac{(2n+1)\pi t}{a}\right) dt \end{aligned}$$

où

$$g(t) = \frac{f(t) + f(-t) - f(0+) - f(0-)}{\sin\left(\frac{\pi t}{a}\right)}$$

est continue par morceaux sur $]0, a/2]$ et bornée sur un voisinage de 0 d'après les hypothèses. La fonction g est donc intégrable sur $]0, a/2]$ et le lemme de Riemann-Lebesgue (1.4.1) entraîne $\lim_{n \rightarrow +\infty} au_n = 0$, d'où le résultat. \square

Corollaire 1.4.1. *Si f est a -périodique et de classe C^1 par morceaux, alors pour tout $x \in \mathbb{R}$, la série de Fourier de f converge en ce point x vers $\frac{f(x+) + f(x-)}{2}$. En particulier, si f est continue en x , la série de Fourier de f en x converge vers $f(x)$.*

1.4.3 Convergence uniforme

Proposition 1.4.7. *Soit $f: \mathbb{R} \rightarrow \mathbb{C}$ une fonction a -périodique, continue et C^1 par morceaux. On définit $\varphi: \mathbb{R} \rightarrow \mathbb{C}$ par $\varphi(t) = f'(t)$ si f est dérivable en t et $\varphi(t) = (f'(t+) + f'(t-))/2$ sinon. Les coefficients de Fourier de φ pour tout $n \in \mathbb{Z}$ vérifient*

$$c_n(\varphi) = \frac{2i\pi n}{a} c_n(f)$$

Démonstration. Soit $0 = x_0 < x_1 < \dots < x_p = a$ une subdivision de $[0, a]$ telle que f soit C^1 sur $[x_{k-1}, x_k]$ pour tout $k \in \llbracket 1, p \rrbracket$. En intégrant par parties, on a pour tout $k \in \llbracket 1, p \rrbracket$

$$\int_{x_{k-1}}^{x_k} \varphi(t) \overline{e_n(t)} dt = \left[f(t) e^{-2i\pi k \frac{t}{a}} \right]_{x_{k-1}}^{x_k} + \frac{2i\pi n}{a} \int_{x_{k-1}}^{x_k} f(t) e^{-2i\pi k \frac{t}{a}} dt$$

puis f étant continue, on obtient en sommant cette relation sur k

$$\begin{aligned} \int_0^a \varphi(t) \overline{e_n(t)} dt &= \left[f(t) e^{-2i\pi k \frac{t}{a}} \right]_0^a + \frac{2i\pi n}{a} \int_0^a f(t) e^{-2i\pi k \frac{t}{a}} dt \\ &= \frac{2i\pi n}{a} \int_0^a f(t) \overline{e_n(t)} dt \end{aligned}$$

d'où

$$c_n(\varphi) = \frac{2i\pi n}{a} c_n(f)$$

\square

Proposition 1.4.8. *Si $f: \mathbb{R} \rightarrow \mathbb{C}$ est une fonction a -périodique, continue et C^1 par morceaux, alors la série de Fourier de f converge uniformément vers f sur \mathbb{R} .*

Démonstration. En reprenant les notations de la proposition précédente, on a $c_n(\varphi) = \frac{2i\pi n}{a} c_n(f)$. Ainsi, pour tout $n \in \mathbb{Z}^*$

$$\begin{aligned} |c_n(f)| &= \left| \frac{ac_n(\varphi)}{2i\pi n} \right| \\ &= \left| \frac{a}{2i\pi n} \right| |c_n(\varphi)| \\ &\leq \frac{1}{2} \left(|c_n(\varphi)|^2 + \frac{a^2}{4\pi^2 n^2} \right) \end{aligned}$$

On sait que $\sum_{n \in \mathbb{Z}} |c_n(\varphi)|^2$ converge d'après l'inégalité de Bessel (voir 1.3.8), donc $\sum_{n \in \mathbb{Z}} |c_n(f)|$ converge aussi. Comme

$$\sum_{n \in \mathbb{Z}} \|c_n(f)e_n\|_\infty = \sum_{n \in \mathbb{Z}} |c_n(f)|$$

la série de Fourier de f converge normalement (pour la norme de convergence uniforme) donc elle converge uniformément. Enfin le corollaire 1.4.1 donne qu'elle converge ponctuellement vers f , donc elle converge uniformément vers f . \square

1.4.4 Convergence dans $L_P^2(0, a)$

Proposition 1.4.9. Soit $f \in L_P^2(0, a)$. La série de Fourier associée à f converge vers f dans $L_P^2(0, a)$. Autrement dit

$$\left\| f - \sum_{k=-n}^{k=n} c_k(f) e_k \right\|_2 \xrightarrow{n \rightarrow +\infty} 0$$

Démonstration. Soit $f \in L_P^2(0, a)$. Comme $\mathcal{C}_c^1([0, a])$ est dense dans $L_P^2(0, a)$ (voir 1.1.2), il existe une suite (f_n) d'éléments de $\mathcal{C}_c^1([0, a])$ qui converge vers f selon la norme $\|\cdot\|_2$. Pour $k \in \mathbb{Z}$ et $n \in \mathbb{N}$ on a

$$\begin{aligned} |c_k(f_n) - c_k(f)| &= \frac{1}{a} \left| \int_0^a (f_n(t) - f(t)) \overline{e_k(t)} dt \right| \\ &\leq \frac{1}{a} \int_0^a |f_n(t) - f(t)| |e_k(t)| dt \\ &\leq \frac{1}{a} \|f_n - f\|_1 \|e_k\|_1 \end{aligned} \tag{5}$$

Par changement de variable on obtient

$$\|f_n - f\|_1 = \int_0^a |f_n(t) - f(t)| dt = \frac{1}{a} \int_0^1 \left| f_n \left(\frac{u}{a} \right) - f \left(\frac{u}{a} \right) \right| du$$

et

$$\|f_n - f\|_2^2 = \frac{1}{a} \int_0^a |f_n(t) - f(t)|^2 dt = \frac{1}{a^2} \int_0^1 \left| f_n \left(\frac{u}{a} \right) - f \left(\frac{u}{a} \right) \right|^2 du$$

L'inégalité de Jensen (voir 1.1.6) avec la fonction convexe $x \mapsto x^2$ donne

$$\left(\int_0^1 \left| f_n \left(\frac{u}{a} \right) - f \left(\frac{u}{a} \right) \right| du \right)^2 \leq \int_0^1 \left| f_n \left(\frac{u}{a} \right) - f \left(\frac{u}{a} \right) \right|^2 du$$

Autrement dit on a $\|f_n - f\|_1^2 \leq \|f_n - f\|_2^2$ donc l'inégalité (5) donne que

$$|c_k(f_n) - c_k(f)| \xrightarrow{n \rightarrow +\infty} 0$$

On pose $S_n = \sum_{k=-n}^n c_k(f) e_k$ et $T_{n,m} = \sum_{k=-n}^n c_k(f_m) e_k$. Montrons que $S_n \xrightarrow{n \rightarrow +\infty} f$ dans $L_P^2(0, a)$. Soit $\epsilon > 0$. On a par inégalité triangulaire

$$\|f - S_n\|_2 \leq \|f - f_m\|_2 + \|f_m - T_{n,m}\|_2 + \|T_{n,m} - S_n\|_2 \quad (6)$$

On sait que f_m tend vers f dans $L_P^2(0, a)$ quand $m \rightarrow +\infty$ donc il existe un entier M_1 tel que pour tout $m \geq M_1$ on ait $\|f - f_m\|_2 < \frac{\epsilon}{3}$. On sait que $T_{n,m}$ converge uniformément vers f_m quand $n \rightarrow +\infty$ donc $\|T_{n,m} - f_m\|_2 \xrightarrow{n \rightarrow +\infty} 0$. En effet

$$\begin{aligned} \|T_{n,m} - f_m\|_2 &= \frac{1}{\sqrt{a}} \left(\int_0^a |T_{n,m}(t) - f_m(t)|^2 dt \right)^{\frac{1}{2}} \\ &\leq \frac{1}{\sqrt{a}} \left(a \sup_{[0,a]} |T_{n,m}(t) - f_m(t)|^2 \right)^{\frac{1}{2}} \\ &\leq \|T_{n,m} - f_m\|_\infty \end{aligned}$$

Il existe donc un entier N tel que pour tout $n \geq N$ on ait $\|T_{n,m} - f_m\|_2 < \frac{\epsilon}{3}$. Enfin, comme $|c_k(f_m) - c_k(f)|$ tend vers 0 quand m tend vers l'infini, il existe $M_2(n)$ dépendant de n tel que pour $m \geq M_2(n)$

$$|c_k(f_m) - c_k(f)| < \frac{\epsilon}{3 \sum_{k=-n}^n \|e_k\|_2}$$

ce qui donne

$$\begin{aligned} \|T_{n,m} - S_n\|_2 &= \left\| \sum_{k=-n}^n c_k(f_m) e_k - \sum_{k=-n}^n c_k(f) e_k \right\|_2 \\ &\leq \sum_{k=-n}^n \|c_k(f_m) - c_k(f)\|_2 \|e_k\|_2 \\ &< \frac{\epsilon}{3} \end{aligned}$$

Pour $n \geq N$, en choisissant $m \geq \max(M_1, M_2(N))$, l'inégalité (6) nous donne donc que $\|f - S_n\|_2 < \epsilon$. On a donc bien $\|f - S_n\|_2 \xrightarrow{n \rightarrow +\infty} 0$, ce que l'on voulait montrer. \square

Corollaire 1.4.2. *La famille $(e_n)_{n \in \mathbb{Z}}$ est une base hilbertienne de $L_P^2(0, a)$.*

Démonstration. La famille $(e_n)_{n \in \mathbb{Z}}$ est orthonormée (voir 1.3.7) et elle est totale d'après ce qui précède. \square

1.5 Représentation spectrale d'une fonction périodique

Soit $f: \mathbb{R} \rightarrow \mathbb{C}$ une fonction a -périodique développable en série de Fourier.

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

Les familles de couples $(\frac{n}{a}, |c_n(f)|)_{n \in \mathbb{Z}}$ et $(\frac{n}{a}, \theta_n)_{n \in \mathbb{Z}}$ où $c_n(f) = |c_n| e^{i\theta_n}$, $\theta_n \in [-\pi, \pi[$ forment respectivement le spectre d'amplitude et le spectre de phase de f . Pour ce TIPE, on se focalisera sur les spectres d'amplitude.

Plaçons nous dans le cas où f est à valeurs dans \mathbb{R} et représente un signal sonore périodique. Dans ce cas on a, pour tout $n \in \mathbb{Z}$, $c_n(f) = \overline{c_{-n}(f)}$ donc $|c_n(f)| = |c_{-n}(f)|$ et $\theta_n = -\theta_{-n}$. On peut donc ignorer les indices n négatifs et redéfinir le spectre d'amplitude comme $(\frac{n}{a}, |c_n(f)|)_{n \in \mathbb{N}}$ et le spectre de phase comme $(\frac{n}{a}, \theta_n)_{n \in \mathbb{N}}$. En posant

$$\begin{cases} a_0(f) &= c_0(f) \\ a_n(f) &= c_n(f) + c_{-n}(f) \end{cases} \quad \begin{cases} b_0(f) &= 0 \\ b_n(f) &= i(c_n(f) - c_{-n}(f)) \end{cases}$$

on peut écrire

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

On a que les $a_n(f)$ et $b_n(f)$ sont réels, $|c_n(f)| = \sqrt{a_n(f)^2 + b_n(f)^2}$ et $\theta_n = \text{atan2}(b_n(f), a_n(f))$. Dans cette somme, on appelle le terme d'indices n l'harmonique de rang n . L'harmonique de rang 1 est appelé le fondamental. À l'harmonique de rang n , on associe la grandeur $\frac{n}{a}$ que l'on appelle fréquence de l'harmonique.

Exemple : signal en dent de scie On considère l'exemple de la fonction suivante

$$\begin{cases} s(x) &= \frac{x}{\pi} \quad \text{pour } -\pi < x < \pi \\ s(x + 2k\pi) &= s(x) \quad \text{pour } -\pi < x < \pi \text{ et } k \in \mathbb{Z} \end{cases}$$

Voici son graphe (représentation en temps) :

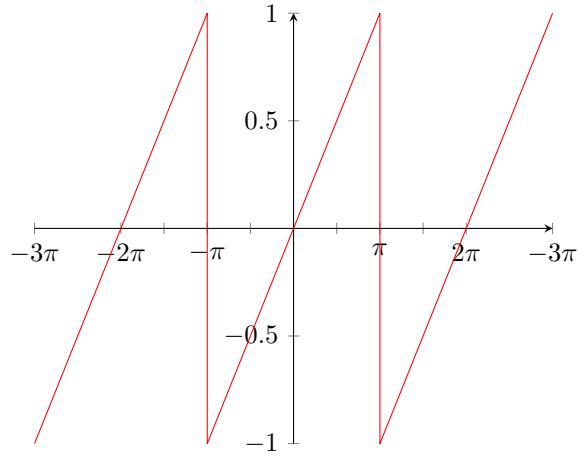


FIGURE 1 – Graphe de s

Voici maintenant son spectre d'amplitude (représentation en fréquence), tronqués aux dix premières valeurs :

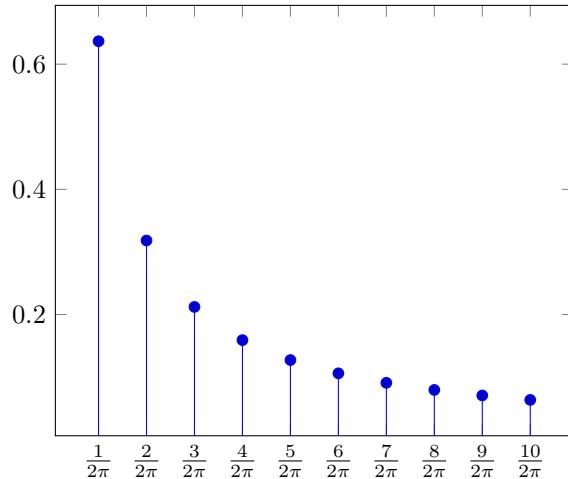


FIGURE 2 – Spectre d'amplitude de s

Application à la musique Un des intérêts d'une telle représentation, c'est qu'elle capture mieux les caractéristiques musicales d'un signal sonore (nous le verrons plus en détails dans les sections suivantes). Cependant, on ne calculera pas de développement en série de Fourier pour l'étude de signaux réels. En effet, ce calcul n'a de sens que pour des signaux parfaitement périodiques. De plus, il existe des moyens plus adapté à une implémentation sur ordinateur pour obtenir la représentation spectrale d'un signal.

2 Transformée de Fourier discrète

On voudrait un algorithme, programmable sur ordinateur, qui permettrait d'obtenir la représentation spectrale d'un signal comme ce qu'on a vu section 1.5. Pour ce faire, il faut déjà numériser le signal, c'est-à-dire en extraire en une suite finie de valeurs. Ensuite, à cette suite finie de valeurs du signal, il nous faut appliquer un équivalent discret du développement en série de Fourier. C'est ce qu'on appelle la transformée de Fourier discrète.

2.1 Échantillonnage

L'échantillonnage est l'un des moyens de numériser un signal. Il consiste à prélever des valeurs du signal à intervalle régulier.

Définition 2.1.1 (Échantillonnage uniforme). *Soient $n \in \mathbb{N}$, $I = [a, b]$ un intervalle et $f: I \rightarrow \mathbb{C}$ une fonction. On définit les nombres complexes y_0, \dots, y_{N-1} comme suit*

$$y_k = f\left(k \frac{(b-a)}{N}\right)$$

Les y_k ainsi définis sont appelés échantillons de f . La grandeur $\frac{N}{b-a}$ est appelée fréquence d'échantillonage.

Théorème 2.1.1 (Nyquist-Shannon). *La représentation discrète d'un signal nécessite une fréquence d'échantillonage supérieure à deux fois la fréquence maximale présente dans ce signal.*

Considérons les fonctions $s(x) = \sin(x)$ et $t(x) = \cos(2x)$. Ces fonctions sont respectivement 2π et π périodiques donc pour échantillonner correctement la deuxième il faut une fréquence d'échantillonage d'au moins $1/\pi$. On remarque par exemple que si l'on échantillonne sur $[-\frac{5\pi}{2}, \frac{7\pi}{2}]$ avec une fréquence d'échantillonage de $\frac{1}{2\pi}$ on obtient les mêmes échantillons pour s et t .

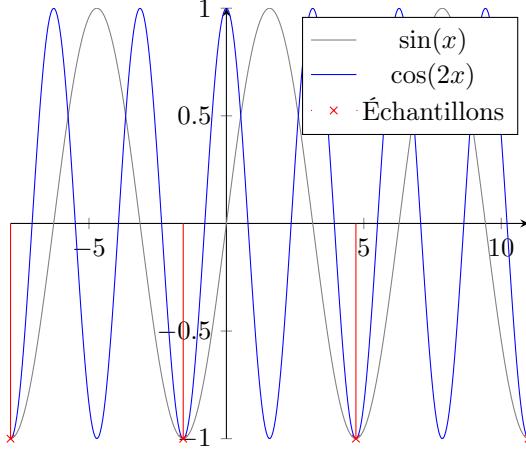


FIGURE 3 – Exemple d'un mauvais échantillonage

2.2 Définition de la transformée de Fourier discrète

Soit f une fonction de \mathbb{R} dans \mathbb{C} (pas forcément périodique). On suppose f échantillonnée par N valeurs sur l'intervalle $[0, a]$, c'est-à-dire que l'on connaît les valeurs y_k telles que

$$y_k = f\left(k \frac{a}{N}\right) \text{ pour } k = 0, 1, 2, \dots, N-1$$

Pour simplifier les écritures, on suppose N pair. On cherche les valeurs c_n telles du polynôme trigonométrique

$$p(t) = \sum_{n=-N/2}^{N/2-1} c_n e_n = \sum_{n=-N/2}^{N/2-1} c_n e^{2i\pi n \frac{t}{a}}$$

qui interpolate f au points $k \frac{a}{N}$, avec $k = 0, 1, 2, \dots, N-1$. Cela revient à résoudre le système linéaire

$$\sum_{n=-N/2}^{N/2-1} c_n \omega^{nk} = y_k \quad k = 0, 1, 2, \dots, N-1 \quad (7)$$

avec $w = e^{\frac{2i\pi}{N}}$. On se ramène à une somme d'indice n dans $\llbracket 0, N-1 \rrbracket$ en effectuant une translation de N sur l'indice de la première moitié de la somme

$$\begin{aligned} \sum_{n=-N/2}^{-1} c_n \omega^{nk} &= \sum_{p=N/2}^{N-1} c_{p-N} \omega^{(p-N)k} \\ &= \sum_{n=N/2}^{-1} c_{p-N} \omega^{nk} \end{aligned} \quad (8)$$

L'égalité (8) s'obtient en remarquant la « N -périodicité» de ω

$$\omega^{(p-N)k} = \omega^{pk-Nk} = e^{\frac{2i\pi pk}{N}} + e^{\frac{2i\pi Nk}{N}} = \omega^{pk}$$

Maintenant, en posant

$$Y_k = \begin{cases} c_n & \text{si } 0 \leq n \leq \frac{N}{2} \\ c_{n-N} & \text{si } \frac{N}{2} \leq n \leq N-1 \end{cases}$$

le système (7) peut s'écrire

$$\sum_{n=0}^{N-1} Y_n \omega^{nk} \quad k = 0, 1, 2, \dots, N-1 \quad (9)$$

Ce système possède une unique solution comme le montre la proposition qui suit.

Proposition 2.2.1. *On pose*

$$y = (y_0 \ y_1 \ \dots \ y_{N-1})^\top \text{ et } Y = (Y_0 \ Y_1 \ \dots \ Y_{N-1})^\top$$

Le système (9) est équivalent au système matriciel suivant

$$\Omega Y = y \quad (10)$$

avec

$$\Omega = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \dots & \omega^{2(N-1)} \\ \vdots & & & & \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \dots & \omega^{(N-1)^2} \end{pmatrix}$$

et la matrice Ω est inversible d'inverse

$$\Omega^{-1} = \frac{1}{N} \bar{\Omega}$$

Démonstration. La relation (10) se prouve facilement en calculant explicitement ΩY . Idem pour l'inversibilité de Ω , il suffit de calculer $\frac{1}{N} \bar{\Omega} \Omega$. \square

On en déduit une formule explicite pour les Y_k

$$Y_n = \frac{1}{N} \sum_{k=0}^{N-1} y_k \omega^{-nk} \quad n = 0, 1, 2, \dots, N-1 \quad (11)$$

Définition 2.2.1. *La formule (11) définit une application $\mathcal{F}_N: \mathbb{C}^N \rightarrow \mathbb{C}^N$:*

$$(y_k) \xrightarrow{\mathcal{F}_N} (Y_n)$$

qu'on appelle transformée de Fourier discrète d'ordre N .

La proposition précédente montre que \mathcal{F}_N est un isomorphisme de \mathbb{C}^N dans lui-même.

Remarque. Une autre approche possible pour définir la TFD est de considérer f comme a -périodique et d'approximer les coefficients de Fourier par la méthode des trapèzes (? , p. 55). On aboutirait aussi, après changement d'indice, à la formule (11).

2.3 Quelques propriétés

On identifie les vecteurs de \mathbb{C}^N à des suites complexes N -périodiques.

Proposition 2.3.1. Soient $(y_k)_{k=0,1,\dots,N-1}$ et $(Y_n)_{n=0,1,\dots,N-1} = \mathcal{F}_N((y_k))$. Alors on a les propriétés suivantes :

- i) $\mathcal{F}_N((y_{-k})) = (Y_{-n})$
- ii) $\mathcal{F}_N((\overline{y_k})) = (\overline{Y_n})$
- iii) $\mathcal{F}_N((\overline{y_{-k}})) = (\overline{Y_n})$

Démonstration. Pour i) :

$$\begin{aligned} Y_{-n} &= \frac{1}{N} \sum_{k=0}^{N-1} y_k \omega^{nk} \\ &= \frac{1}{N} \sum_{k=-N+1}^0 y_{-k} \omega^{-nk} \\ &= \frac{1}{N} \sum_{k=1}^N y_{N-k} \omega^{-nk} \\ &= \frac{1}{N} \sum_{k=0}^{N-1} y_k \omega^{-nk} \quad \text{car } y_{N-N} \omega^{-nN} = y_0 \omega^{-n0} \\ &= \mathcal{F}_N((y_{-k}))_n \end{aligned}$$

Pour ii) :

$$\begin{aligned} \overline{Y_{-n}} &= \overline{\frac{1}{N} \sum_{k=0}^{N-1} y_k \omega^{nk}} \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \overline{y_k} \omega^{-nk} \\ &= \mathcal{F}_N((\overline{y_k}))_n \end{aligned}$$

Enfin (i) et (ii) impliquent (iii). \square

On en déduit une propriété de la TFD de vecteurs réels :

Proposition 2.3.2. Soient $(y_k)_{k=0,1,\dots,N-1}$ un vecteur de \mathbb{R}^N et $(Y_n)_{n=0,1,\dots,N-1} = \mathcal{F}_N((y_k))$. Alors pour tout $n \in \mathbb{Z}$

$$Y_{-n} = \overline{Y_n}$$

Démonstration. Si les y_k sont réels alors pour tout k on a $y_k = \overline{y_{-k}}$ donc d'après les points (i) et (iii) de la proposition précédente on a $Y_{-n} = \overline{Y_n}$. \square

Voyons une conséquence pratique de cette proposition. Soient deux vecteur (x_k) et (y_k) dans \mathbb{R}^N et soient (X_n) et (Y_n) leur TFD respective. Par la proposition on sait que

$$\forall n = 0, 1, \dots, N-1 \quad X_{N-n} = \overline{X_n} \quad \text{et} \quad Y_{N-n} = \overline{Y_n}$$

Soient $z_k = x_k + iy_k$ et (Z_n) la transformée de (z_k) . Par linéarité on a

$$Z_n = X_n + iY_n$$

D'où

$$\begin{cases} X_n &= \frac{1}{2} (Z_n + \overline{Z_{N-n}}) \\ Y_n &= \frac{1}{2i} (Z_n - \overline{Z_{N-n}}) \end{cases} \quad n = 0, 1, \dots, N/2$$

On peut donc calculer (X_n) et (Y_n) en une fois en calculant (Z_n) .

2.4 Transformée de Fourier rapide

2.4.1 Principe

Le calcul d'une TFD (transformée de Fourier discrète) de taille N par la formule (11) nécessite $(N-1)^2$ produits complexes et $N(N-1)$ additions complexes. Dans les applications réelles, N aura souvent une valeur supérieure à 1000. Par exemple dans l'application proposée section 3, on a $N = 4096$. Si on utilisait la formule (11), le calcul d'une seule transformée nécessiterait un nombre de calculs de l'ordre de 30 millions.

Heureusement, il existe des algorithmes de calcul bien plus performants, qu'on appelle les algorithmes de transformée de Fourier rapide. Nous allons aborder ici le plus célèbre, l'algorithme de Cooley-Tuckey. On se place dans le cas où N est une puissance de 2 (appelé cas «radix-2 case»), car il est à la fois facile à démontrer et très utilisé en pratique.

On suppose donc $N = 2^p$. On sépare la somme de la formule (11) en deux parties, l'une avec les termes d'indice pair, l'autre avec les termes d'indice impairs

$$Y_k = \frac{1}{N} \left(\sum_{m=0}^{N/2-1} y_{2m} e^{-\frac{2\pi i}{N} (2m)k} + \sum_{m=0}^{N/2-1} y_{2m+1} e^{-\frac{2\pi i}{N} (2m+1)k} \right)$$

En factorisant la deuxième somme par $e^{-\frac{2\pi i}{N}k}$ et en réarrangeant les facteurs multiplicatifs devant les sommes

$$\begin{aligned} Y_k &= \frac{1}{2} \left(\frac{1}{N/2} \sum_{m=0}^{N/2-1} y_{2m} e^{-\frac{2\pi i}{N/2}mk} + e^{-\frac{2\pi i}{N}k} \frac{1}{N/2} \sum_{m=0}^{N/2-1} y_{2m+1} e^{-\frac{2\pi i}{N/2}mk} \right) \\ &= \frac{1}{2} \left(P_k + e^{-\frac{2\pi i}{N}k} I_k \right) \end{aligned}$$

Puis en remarquant que

$$\begin{cases} P_{k+N/2} &= P_k \\ I_{k+N/2} &= I_k \\ e^{-\frac{2\pi i}{N}(2m)(k+N/2)} &= -e^{-\frac{2\pi i}{N}(2m)k} \end{cases}$$

On trouve

$$Y_{k+N/2} = \frac{1}{2} \left(P_k - e^{-\frac{2\pi i}{N}k} I_k \right)$$

En résumé on a

$$\begin{cases} Y_k &= \frac{1}{2} \left(P_k + e^{-\frac{2\pi i}{N}k} I_k \right) \\ Y_{k+N/2} &= \frac{1}{2} \left(P_k - e^{-\frac{2\pi i}{N}k} I_k \right) \end{cases} \quad k = 0, 1, 2, \dots, N/2-1 \quad (12)$$

Le calcul de la TFD revient donc à calculer les P_k et les I_k pour $k = 0, 1, 2, \dots, N/2-1$. Mais les P_k et les I_k peuvent aussi être obtenus par des TFD

$$\begin{aligned} (y_0, y_2, \dots, y_{N-2}) &\mapsto (P_0, P_1, \dots, P_{N/2-1}) \\ (y_1, y_3, \dots, y_{N-1}) &\mapsto (I_0, I_1, \dots, I_{N/2-1}) \end{aligned}$$

Le même procédé peut donc être appliqué pour calculer les P_k et les I_k . On obtient ainsi un algorithme récursif.

2.4.2 Coût de l'algorithme

Pour $N = 2^p$, on note respectivement M_p et A_p le nombre de multiplications et le nombre d'additions utilisées dans cet algorithme pour le calcul d'une TFD de taille N .

Pour $p = 0$, on calcule la TFD en utilisant directement la définition, ce qui donne $M_0 = 1$ et $A_0 = 0$.

Pour $p \geq 2$, on utilise les formules (12) pour exprimer M_p en fonction de M_{p-1} et A_p en fonction de A_{p-1} . On convient de négliger le coût des divisions par 2 car elles peuvent être remplacées par une division par 2^p à la fin de l'algorithme. Le calcul des P_k et des I_k se fait par deux TFD de taille 2^{p-1} d'où

$$\begin{cases} M_p &= 2M_{p-1} + C_1(p) \\ A_p &= 2A_{p-1} + C_2(p) \end{cases} \quad C_1(p), C_2(p) \geq 0$$

Puis pour chaque $k = 0, 1, \dots, 2^{p-1}$, on doit faire deux additions supplémentaires donc $C_1(p) = 2^p$. Enfin pour $k = 1, 2, \dots, 2^{p-1}$, on doit encore effectuer une multiplication donc $C_1(p) = 2^{p-1} - 1$. On en déduit

$$\begin{cases} M_0 &= 1 \\ M_p &= 2M_{p-1} + 2^{p-1} - 1 \end{cases} \quad \begin{cases} A_0 &= 0 \\ A_p &= 2A_{p-1} + 2^p \end{cases}$$

On peut montrer par récurrence que

$$\begin{cases} M_p &= 2^{p-1}(p-2) + 1 \\ A_p &= 2^p p \end{cases}$$

On trouve donc, en fonction de N un coût de $N/2(\log_2(N)-2)+1$ multiplications et $N \log_2(N)$ additions. Pour $N = 4096$, cela donne environ 70 000 calculs, soit bien moins que les 30 million que l'on aurait eu à faire avec l'algorithme naïf.

2.5 Représentation spectrale d'un signal échantillonné

2.5.1 Définitions

Avec la TFD, on peut donner une représentation spectrale d'un signal échantillonné, d'une manière similaire de ce que l'on a fait section 1.5. Il a cependant une différence majeure avec les séries de Fourier : le signal considéré n'a pas forcément à être périodique, et s'il l'est, sa période n'a pas besoin d'être connue au préalable. En conséquence, les fréquences qu'on représentera sur le spectre ne seront pas forcément des harmoniques.

On définit les spectres d'amplitude et de phase comme suit

Définition 2.5.1. Soient N un entier $(y_k)_{k=0,1,\dots,N-1}$ les échantillons d'un signal échantillonné à une fréquence f , et $(Y_n)_{n=0,1,\dots,N-1}$ leur TFD. On écrit pour tout $n = 0, 1, \dots, N-1$, $Y_n = |Y_n|e^{i\theta_n}$ avec $-\pi \leq \theta_n < \pi$.

On appelle spectre d'amplitude la famille $(n, |Y_n|)_{n=0,1,\dots,N-1}$ et spectre de phase la famille $(n, \theta_n)_{n=0,1,\dots,N-1}$

Définition 2.5.2. Pour $n = 0, 1, \dots, N-1$, on appelle l'intervalle $\left[n\frac{f}{N}, (n+1)\frac{f}{N}\right]$ la n -ième bande de fréquences. En anglais on parle de «frequency bin».

2.5.2 Exemple

En guise d'exemple, comparons les spectres d'amplitude d'une même note jouée au piano et à la guitare. On joue la note sol à 196Hz (Sol2 ou G3). Pour chacune des deux notes on prend 2048 échantillons à une fréquence d'échantillonnage de 44100Hz.

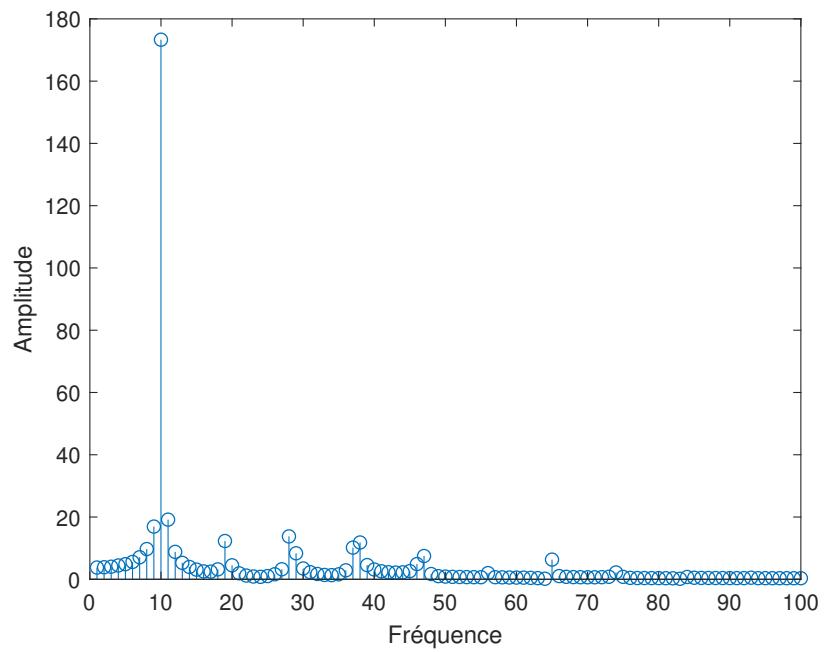


FIGURE 4 – Spectre d'amplitude tronqué de la note sol 196Hz jouée au piano

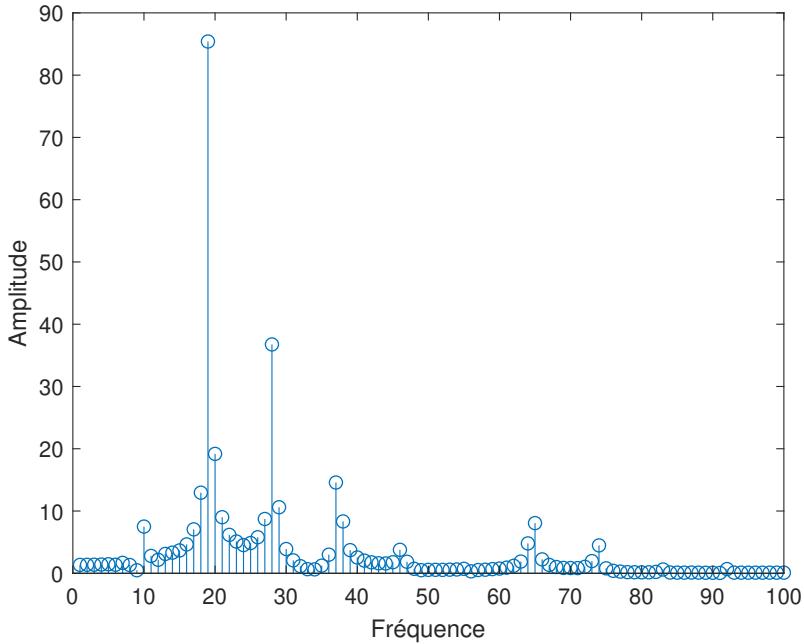


FIGURE 5 – Spectre d'amplitude tronqué de la note sol 196Hz jouée à la guitare

Sur les deux spectres on observe que les pics correspondent bien aux harmoniques. Les bandes de fréquence ont une largeur de $44100/2048 \approx 21\text{Hz}$. Le fondamental correspond donc à la bande numéro 10, le deuxième harmonique à la bande numéro 19, etc.

Pour la note jouée au piano, on remarque que l'harmonique ayant la plus grande amplitude est le fondamental, tandis que pour la guitare c'est le deuxième harmonique. Le fondamental au piano domine complètement les autres harmoniques : il a une amplitude au moins 15 fois plus grande que tous les autres ! On ne retrouve pas ce phénomène à la guitare, où les 3e et 4e harmoniques sont assez présents.

Ces différences dans la décomposition harmonique sont en partie ce qui explique pourquoi une même note jouée sur deux instruments différents ne sonne pas de la même manière.

3 Application à la reconnaissance musicale

Maintenant, on va utiliser la TFD pour résoudre un problème concret : celui de la reconnaissance musicale. Supposons qu'on dispose de fichiers audio qui constituent une base de données. Étant donné un court enregistrement musical

(quelques dizaines de secondes au plus), on veut pouvoir identifier le morceau dont il est extrait. Pour cela on va utiliser l'approche suggérée par ?.

3.1 Empreinte acoustique

Pour chaque fichier audio, on calcule ce que l'on appelle une empreinte acoustique. Cette empreinte acoustique doit résumer le mieux possible les caractéristiques audibles du signal sonore. On utilise le même procédé de calcul d'empreinte pour les fichiers de la base de données et pour les échantillons à identifier. Pour identifier un échantillon, on cherche à faire correspondre son empreinte à l'une de la base de données. Dans l'approche de ?, l'empreinte est composée d'éléments qu'on appelle des hashs. Pour une identification efficace, il suggère que les empreintes acoustiques doivent avoir les qualités suivantes :

- **Spécificité** : deux sons très différentes ne doivent pas avoir la même empreinte.
- **Robustesse** : deux sons perçus comme identique par l'oreille humaine mais différents par leur représentation numérique (à cause de la présence de bruit ou de distorsion du signal par exemple) doivent avoir la même empreinte.
- **Localité temporelle** : chaque hash correspond à un point dans le temps et l'information qu'il porte ne doit concerner qu'une petite partie du signal sonore localisée autour de ce point, de sorte que deux événements distants ne perturbent pas le hash.
- **Invariance par translation** : tous les hashs calculés pour le fichier original doivent être reproductibles indépendamment de leur position temporelle dans le morceau. En d'autres termes, si un hash h correspond à un point t du morceau original, alors l'empreinte d'un extrait allant de t_0 à t_1 avec $t_0 \leq t \leq t_1$, h devrait contenir h . En effet, un échantillon à identifier peut être extrait à partir de n'importe quelle position d'un morceau.

3.2 Spectrogrammes et constellations

La première étape pour générer notre empreinte acoustique est le calcul d'un spectrogramme. Pour cela on utilise ce que l'on appelle la transformée de Fourier discrète à court terme (Short-time Fourier transform en anglais). Cela consiste à découper les données en petits morceaux (qui peuvent se chevaucher) et faire une TFD sur chacun de ces morceaux. Soient (x_k) les échantillons d'un signal audio. On prend de taille $N = 1024$ qui se chevauchent de moitié. On rassemble les résultats dans une matrice $S = (s_{i,j})$. La j -ième colonne de cette matrice correspond à la TFD du j -ième morceau. Concrètement on a

$$(s_{1j}, s_{2j}, \dots, s_{Nj}) = \mathcal{F}_N(x_{jN/2}, x_{jN/2+1}, \dots, x_{jN/2+N})$$

Ensuite on calcule $A = |S| = (|s_{i,j}|)$. Chaque colonne de A est un spectre d'amplitude. On obtient que l'on appelle un spectrogramme d'amplitude.

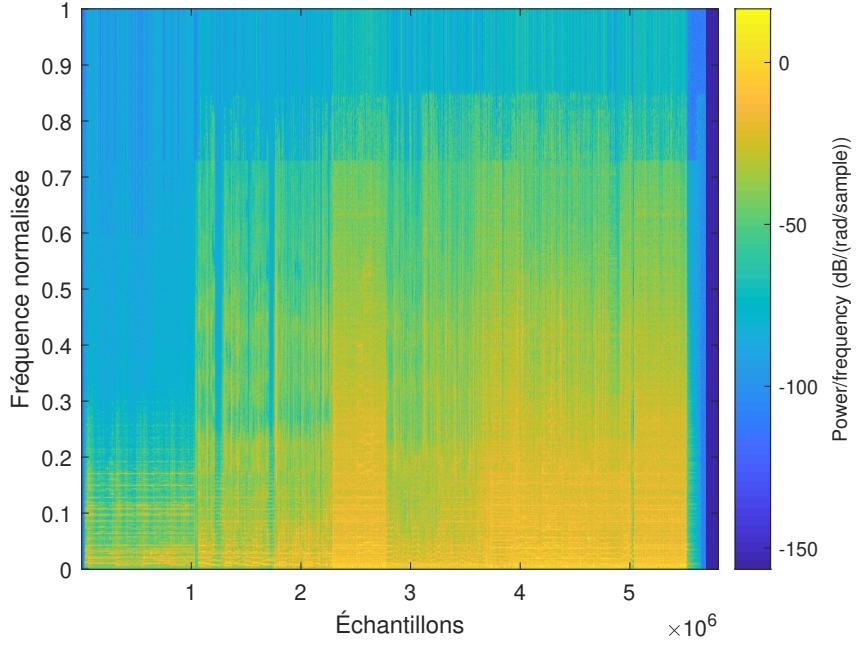


FIGURE 6 – Spectrogramme de « Dream Lantern » des Radwimps

Ensuite, dans ce spectrogramme d'amplitude, on ne garde que les maxima locaux. ? ne définit pas ce que sont les maxima locaux, et à priori il n'y a pas qu'une seule manière de faire¹. On choisit une définition « naturelle » : les maximums locaux sont les (i, j) tels que

$$|s_{i,j}| \geq |s_{k,l}| \quad \forall k, l \quad \max(|i - k|, |j - l|) \leq R$$

R est une constante dont on choisit la valeur qui représente la taille du voisinage que le maximum local doit « dominer ». On construit une matrice $P = (p_{i,j})$ de même taille que A et telle que

$$p_{i,j} = \begin{cases} 1 & \text{si } (i, j) \text{ est un maximum local} \\ 0 & \text{sinon} \end{cases}$$

Ce choix a été fait pour sa simplicité et aussi parce qu'il donne des résultats relativement corrects. Néanmoins, il est possible d'améliorer la précision de l'algorithme en utilisant un procédé de sélection plus sophistiqué.

On appelle la matrice P une constellation. On dit qu'une paire (f, t) est un point de la constellation si et seulement $p_{f,t} = 1$. Ci-dessous un exemple de constellation :

1. Par exemple ? propose une façon qui prend en compte la science psychoacoustique.

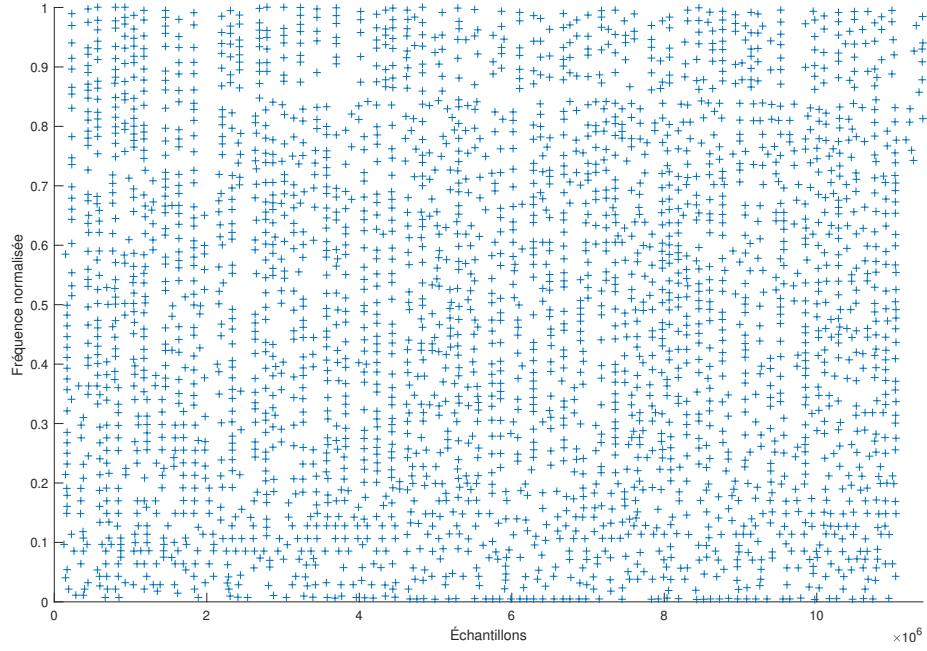


FIGURE 7 – Constellation de « Dream Lantern » des Radwimps

On utilise cette représentation en constellation pour sa robustesse. Le fait de ne garder que des maxima locaux et de se débarrasser ensuite de l'amplitude fait qu'elle est peu impactée par la présence de bruit ou de distorsions du signal.

Avec cela, on peut déjà construire un algorithme d'identification. Pour un extrait donné, on calcule sa constellation puis on la compare avec celle de chaque morceaux de la base en les faisant « glisser » l'une contre l'autre. On garde le morceaux dont la constellation correspond le mieux à celle de l'extrait. Le problème est que bien entendu cet algorithme est trop coûteux en calcul. La base de données peut contenir des millions de morceaux, et pour chacun de ces morceaux il faudrait faire des dizaines de milliers de comparaison de constellations.

3.3 Organisation des données

Les empreintes acoustiques que nous utilisons sont composés d'éléments que l'on appelle des hashs. Les hashs sont générés à partir de paires de points de la constellation. On choisit des points d'ancrages, et à chacun de ce points d'ancrage on associe une zone cible. Puis on associe les points d'ancrage avec chacun des points de la zone cible. Concrètement, si on a un point d'ancrage $a = (f_a, t_a)$, on l'associe avec chacun des points $x = (f_x, t_x)$ de sa zone cible. On obtient des triplets de la forme $(f_a, f_x, t_x - t_a)$ où f_a et f_x sont les numéros de bande de fréquence de a et x et $t_x - t_a$ est la différence de leur position temporelle.

? ne dit pas comment choisir ni les points d'ancrage, ni les zones cibles. Pour notre implémentation, on prend tous les points de la constellation comme points d'ancrage. Puis pour un point d'ancrage $a = (f_a, t_a)$, on prend comme zone cible le plus grand ensemble X_a de points de la constellation tel que :

- $\text{Card}(X_a) \leq F$ avec F une constante choisie.
- Pour tout $x = (f_x, t_x) \in X_a$ on a $t_a \leq t_x$
- Pour tous $x = (f_x, t_x), y = (f_y, t_y), z = (f_z, t_z)$ points de la constellation tels que $x, y \in X_a$ et $t_x \leq t_z \leq t_y$ on a $z \in X_a$

Dans la base de données, à chaque hash on fait correspondre la liste des morceaux qui contiennent ce hash dans leur empreinte.

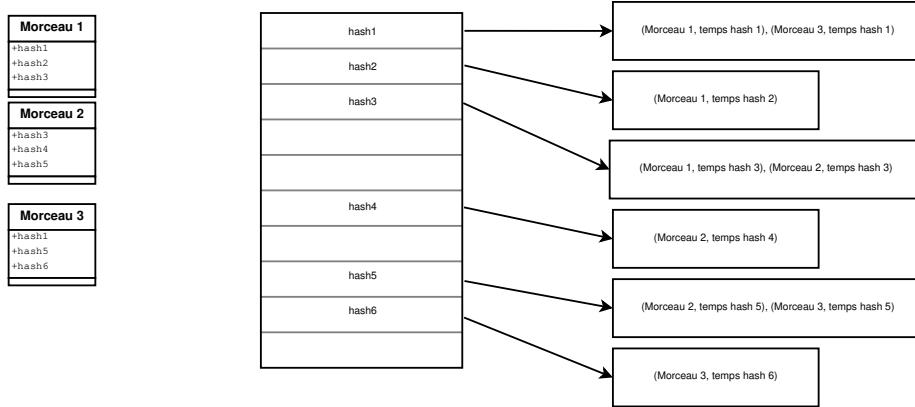


FIGURE 8 – Organisation des données

Pour obtenir une telle structure, on peut par exemple utiliser une table de hachage² (voir ?, p. 238). Dans ce cas, la recherche d'un hash dans la base se fait en temps constant. Cela veut dire que si N désigne le nombre hashs dans la base, la recherche se fait en $O(1)$ opérations. Il existe d'autres alternatives, l'essentiel étant que la recherche d'un hash particulier puisse se faire de manière rapide, et ce quelque soit la taille de la base de données.

3.4 Processus d'identification

Pour identifier un échantillon, on applique ce qui précède pour obtenir un ensemble de couples (hash, temps) (h, t_h) . On recherche ensuite chaque hash dans la base de données. Pour chaque hash on obtient une liste de couples (m, t_m) où m est un morceau hash h et $t_{m,h}$ le temps où le hash apparaît³. Parmi ces

2. Les composants de l'empreinte sont justement appelés hash parce qu'ils sont les clés d'une table de hachage (hash table en anglais)

3. Un hash peut apparaître plusieurs fois dans un morceau, donc $t_{m,h}$ n'est pas bien défini, mais on utilise cette notation pour simplifier.

morceaux, il faut trouver celui qui est le plus cohérent par rapport aux temps associés aux hashs. On prend le morceau le mieux « aligné » temporellement. Pour cela on procède comme suit :

- Pour chaque morceau m , on calcule les différences en temps $\delta_{m,h} = t_h - t_{m,h}$.
- On compte le nombre d'occurrences de chaque $\delta_{m,h}$.
- On choisit le morceau m qui maximise $\max_h(\delta_{m,h})$.

Le morceau choisi doit correspondre à celui de l'extrait, pourvu qu'il soit dans la base de données.

3.5 Implémentation MATLAB

Le programme MATLAB utilise une base données SQLite. Au besoin, on peut facilement adapter le code pour un autre système de base de données relationnelle comme PostgreSQL par exemple.

Deux tables sont utilisées, l'une pour stocker les méta-données⁴ sur les morceaux et l'autre pour les empreintes acoustiques.

```
CREATE TABLE songs (id integer primary key, name varchar not
↪ null)
CREATE TABLE fingerprints (hash varchar not null, song_id integer
↪ not null, song_time integer not null, unique(hash, song_id,
↪ song_time))
```

Deux index sont utilisés pour accélérer les recherches

```
CREATE INDEX idx_fingerprints_hash ON fingerprints(hash)
CREATE INDEX idx_fingerprints_song_id ON fingerprints(song_id)
```

Le premier permet de trouver rapidement un hash dans la base de données, le deuxième sert à recherche des morceaux correspondant à un hash.

La partie du code la plus coûteuse en calcul est la recherche de maxima locaux dans le spectrogramme. Plutôt que de faire une recherche naïve on utilise la fonction `imdilate` de MATLAB.

```
function [P] = peaks(M, radius)
% calcule la matrice des maxima locaux
nhood = ones(radius * 2 + 1, radius * 2 + 1);
nhood(radius+1, radius+1) = 0;
J = imdilate(M, nhood);
P = J >= 0.0005 & M > J; % on prend les max locaux supérieur
↪ à un certain seuil
% le seuil n'est pas obligatoire mais sans lui certain
↪ morceaux
% généreront trop de hashs
end
```

4. Dans notre cas on stocke simplement un identifiant et le nom du fichier, mais on pourrait utiliser cette table pour stocker le nom des artiste, les albums, etc.

La fonction `imdilate` réalise une opération classique en traitement d'images, que l'on appelle « dilatation ». Un avantage à son utilisation est que l'on peut faire effectuer les calculs par une carte graphique. Dans le cas de MATLAB, il faut avoir la Parallel Computing Toolbox et remplacer `imdilate(M, nhood)` par `imdilate(gpuArray(M), nhood)`. Dans d'autres langages, comme le Python par exemple, on pourra utiliser la bibliothèque OpenCV.