
Transformée de Fourier discrète et structures musicales

Pierre BEAUGUITTE

Stage de recherche du Master 2 ATIAM

Effectué de mars à juillet 2011

IRCAM - Équipe Représentations Musicales
Université Pierre et Marie Curie - UPMC

encadré par Carlos AGON (CNRS/IRCAM, Paris),
Emmanuel AMIOT (CPGE, Perpignan),
Moreno ANDREATTA (CNRS/IRCAM, Paris).

Remerciements

Je tiens tout d'abord à remercier mes encadrants Carlos Agon, Moreno Andreatta et Emmanuel Amiot, pour leur disponibilité, leur patience et leur sympathie.

Merci aux collègues du bureau A 107, Louis Bigo, Jean Bresson et Jérémie Garcia, ainsi qu'à l'ensemble de l'équipe Représentations Musicales pour l'accueil chaleureux à l'Ircam.

Je remercie Daniele Ghisi et John Mandereau, mes aînés sur le terrain de la Z-relation, pour m'avoir fait profiter de leur expertise.

Merci enfin à toute l'équipe du Master ATIAM, aux enseignants pour le savoir transmis et aux collègues étudiants pour cette année si enrichissante.

Table des matières

Introduction	1
1 Cadre mathématique, origine de l'approche	3
1.1 Modèle mathématique d'objets musicaux	3
1.2 Les <i>Fourier Properties</i> de David Lewin	4
1.3 Définition et propriétés élémentaires de la transformée de Fourier discrète	5
1.4 Théorème de l'hexacorde	7
1.5 Invariance du module	8
1.5.1 Actions de groupes	8
1.5.2 Z-relation, homométrie	9
1.6 Extensions possibles du modèle	10
1.7 Pistes concernant l'homométrie	12
1.7.1 Espace de définition des phases	12
1.7.2 Quasi-recouvrement du cercle	13
2 Gammes et structures bien réparties	15
2.1 Définition originelle	15
2.2 Typologie des structures bien réparties	17
2.3 Coefficients de Fourier d'une SBR	17
2.4 Caractérisation par la TFD	20
3 Utiliser la TFD pour comparer les rythmes	21
3.1 Définition et propriétés de la distance $d_{\mathcal{F}}$	21
3.2 Évaluation perceptive de $d_{\mathcal{F}}$	22
3.2.1 Protocole d'écoute	22
3.2.2 Analyse des résultats	22
3.3 Perception de la Z-relation rythmique	24
3.3.1 Protocole d'écoute	24
3.3.2 Analyse des résultats	24
3.4 Bilan critique	25
Conclusion	27
Questions ouvertes, perspectives	27

Introduction

La transformée de Fourier discrète permet de mettre en évidence les périodicités d'une fonction. De ce fait elle est l'outil mathématique de référence de la théorie du signal sonore, un son pouvant être décomposé en une combinaison de sinusoïdes. Elle est ainsi massivement utilisée pour l'étude du son, et en particulier de la musique, au niveau de sa manifestation physique. Mais certaines structures musicales abstraites possèdent elles aussi des périodicités : l'octave pour les hauteurs, la mesure pour un rythme joué en boucle. . . Il n'est donc pas surprenant que la transformée de Fourier trouve des applications fructueuses dans l'étude de ces structures au niveau symbolique. On peut citer à ce propos ces phrases de D.T. Vuza, extraite de [27] :

Fourier transforms is of major importance in the mathematical modelling of phenomena involving periodicity, and periodicity is one of the characteristics of major importance in the class of musical phenomena. It is therefore my conviction that in the near future music theory will integrate convolution and Fourier transform as effective investigation tools, music theorists being able to use them in the same way as presently they make use of groups, homomorphisms, group actions, and so forth.

L'idée n'est pas nouvelle, mais est relativement méconnue par rapport aux approches algébriques, moins surprenantes au vu des objets discrets étudiés. Nous allons dans ce rapport donner quelques-uns des résultats remarquables de l'application de la transformée de Fourier discrète aux structures musicales abstraites, sans toujours détailler les démonstrations, qui se trouvent dans les références citées. Nous tenterons de bien montrer la généralité de ces résultats : si dans la littérature les auteurs considèrent souvent une catégorie spécifique d'objets musicaux, nous verrons que les modèles utilisés permettent d'étudier indistinctement différentes dimensions musicales.

Après une première partie donnant des bases théoriques générales, nous nous concentrerons sur l'application de la transformée de Fourier aux gammes bien réparties. Enfin, après cet exposé théorique, nous détaillerons les résultats d'un test d'écoute conçu dans le but de soumettre ces considérations mathématiques à la réalité de la perception musicale.

Chapitre 1

Cadre mathématique, origine de l'approche

1.1 Modèle mathématique d'objets musicaux

Avant de parler de l'opération qu'est la transformée de Fourier discrète et de ses propriétés, il convient de définir les objets mathématiques auxquels nous allons l'appliquer et les objets musicaux qu'ils peuvent modéliser.

On s'intéresse à des sous-ensembles de l'anneau \mathbb{Z}_c , qui désigneront trois sortes d'objets musicaux : des accords, des gammes, ou des rythmes périodiques. Dans les deux premiers cas, les éléments du sous-ensemble désignent des *classes de hauteur* : il s'agit de la hauteur d'une note « modulo l'octave ». Conventionnellement, l'élément 0 désigne la classe *do*, donc indistinctement les do de toutes les octaves. Les entiers suivants désignent, dans l'ordre croissant, les demi-tons successifs (1 pour do#, 2 pour ré, ...). Le système tonal occidental étant le tempérament de référence dans la littérature, la cardinalité c vaudra souvent 12, mais la notion de classe de hauteur se généralise sans peine à la micro-tonalité. Il faut pour cela se ramener aux fréquences des notes, faute de noms : deux hauteurs a et b sont équivalentes si leurs fréquences f_a et f_b sont liées par $f_a = 2^n f_b$, avec $n \in \mathbb{Z}$ (on garde ainsi la notion d'octave, seule la subdivision de celle-ci change).

Dans le cas d'un rythme périodique, les éléments du sous-ensemble sont les instants des attaques du rythme (ou *onset times*), relativement à la période c , correspondant à la mesure. Ainsi \mathbb{Z}_8 peut décrire une métrique à 4 temps divisée en croches binaires, \mathbb{Z}_{12} une métrique à 3 temps divisée en doubles croches binaires ou bien à 4 temps divisée en croches ternaires. . .

On sera amené à utiliser la représentation circulaire, où un objet musical apparaît alors comme une sous-figure du polygone régulier à c côtés, *i.e.* un polygone convexe inscrit dans le cercle et dont les sommets sont des sommets du polygone régulier (voir figure 1.1).

Ces modèles ont leurs limites, et ne prétendent pas pouvoir décrire exhaustivement les gammes, accords et rythmes considérés. Ainsi, il est impossible de tenir compte des

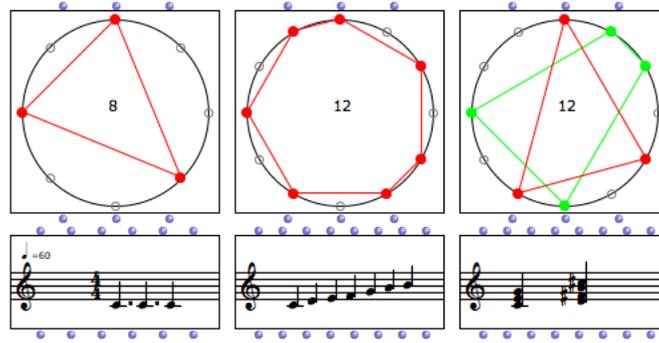


FIGURE 1.1 – Représentations circulaires d’objets musicaux

renversements d’un accord, de la durée ou de l’intensité des éléments constitutifs d’un rythme (contrairement aux *time-spans* de David Lewin, voir [19]). Enfin, ces objets sont uni-dimensionnels, on peut modéliser des hauteurs ou des durées, mais pas les deux en même temps (il faudrait pour cela considérer des produits cartésiens de tels sous-ensembles, une piste que nous n’avons pas suivie dans le cadre de ce stage). Ils fournissent néanmoins un cadre permettant d’étudier un large spectre d’objets musicaux et d’expliciter des résultats intéressants.

1.2 Les *Fourier Properties* de David Lewin

David Lewin, théoricien de la musique américain, a publié en 1959 le premier article faisant usage de la transformée de Fourier discrète à des fins d’analyse musicale autres que l’étude du signal sonore ([17]). L’auteur s’intéresse aux intervalles existant entre les notes de deux accords que l’on modélise par des sous-ensembles A et B de \mathbb{Z}_{12} . Pour $k \in \mathbb{Z}_{12}$, il définit la fonction intervallique comme

$$IFUNC_{A,B}(k) = \text{card} \{ (x, y) \in A \times B \mid x = y + k \}$$

Par exemple pour $A = \{0, 4, 7\}$ (do majeur) et $B = \{2, 5, 9\}$ (ré mineur), on trouve $IFUNC_{A,B}(0) = 0$ car il n’y a aucune note commune entre A et B , $IFUNC_{A,B}(1) = 1$ car on a un unique intervalle de demi-ton entre les notes de A et de B (de mi (4) à fa (5))... La question soulevée par Lewin est alors la suivante : est-on capable, si l’on connaît l’ensemble A et les valeurs de $IFUNC_{A,B}$, de retrouver l’ensemble B ?

Pour y répondre, il remarque que la fonction $IFUNC$ peut s’écrire comme un produit de convolution : $IFUNC_{A,B}(k) = \mathbb{1}_A * \mathbb{1}_B(k)$. En effet

$$\mathbb{1}_A * \mathbb{1}_B(k) = \sum_{x \in \mathbb{Z}_c} \mathbb{1}_A(x) \cdot \mathbb{1}_B(k - x) = \sum_{x \in \mathbb{Z}_c} \mathbb{1}_A(x) \cdot \mathbb{1}_B(x - k) = \sum_{\substack{(x,y) \in A \times B \\ y = x - k}} 1$$

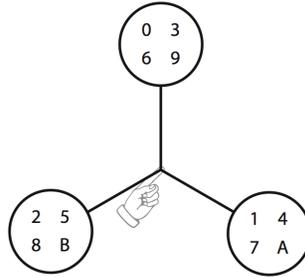


FIGURE 1.2 – La balance de Fourier 4 (source : [23])

La propriété bien connue de la transformée de Fourier de changer un produit de convolution en produit usuel invite naturellement à l’appliquer ici. En définissant \mathcal{F}_P comme la transformée de Fourier de la fonction indicatrice de l’ensemble $P \subset \mathbb{Z}_c$, on a $\mathcal{F}(IFUNC_{A,B}) = \mathcal{F}_A \cdot \mathcal{F}_{-B}$. Puisque les valeurs de la transformée de Fourier d’une fonction caractérisent entièrement cette fonction, il suffit maintenant de connaître les valeurs de \mathcal{F}_{-B} pour répondre à la question. Le calcul de ces valeurs est trivial dès lors que \mathcal{F}_A ne s’annule jamais. Lewin expose donc les cinq formes particulières d’accords pour lesquelles les coefficients de Fourier s’annulent (les *Fourier Properties 1, 2, 3, 4, 6*, le chiffre indiquant quel coefficient s’annule), et qui répondent négativement à la question.

Cet article a trouvé peu d’échos à sa parution, certainement du fait que le *Journal of Music Theory* s’adresse (ou du moins s’adressait) plus à un public de musicologues que de mathématiciens. David Lewin a repris en 2001 ces résultats dans [18], paru dans le même journal, puis Ian Quinn, dans sa thèse [23] a généralisé ces propriétés à d’autres valeurs de c (et donc à la micro-tonalité), et a proposé l’image plus intuitive des *balances de Fourier* (voir figure 1.2) : une fois chaque note de A placée sur le plateau qui lui correspond, si la balance i est équilibrée, alors A possède la *Fourier Property i*.

1.3 Définition et propriétés élémentaires de la transformée de Fourier discrète

Après ce point historique, nous allons maintenant donner la définition et les premières propriétés de la transformée de Fourier discrète.

La transformée de Fourier discrète (TFD) est l’opération définie, pour $f : \mathbb{Z}_c \rightarrow \mathbb{C}$, par :

$$\forall t \in \mathbb{Z}_c, \mathcal{F}(f) : t \mapsto \sum_{k \in \mathbb{Z}_c} f(k) e^{-2i\pi kt/c}$$

Les c valeurs complexes de $\mathcal{F}(f)$ sont les coefficients de Fourier de f . On parle parfois d'espace de Fourier pour désigner l'espace d'arrivée de $\mathcal{F}(f)$.

Elle réalise une bijection entre l'espace des fonctions de \mathbb{Z}_c dans \mathbb{C} et l'espace de Fourier, par l'opération de transformée de Fourier inverse :

$$\forall t \in \mathbb{Z}_c, f(t) = \frac{1}{c} \sum_{k=1}^c \mathcal{F}(f)(k) e^{2i\pi kt/c}$$

Ainsi la donnée des coefficients de Fourier de f est équivalente à la donnée des valeurs de la fonction f elle-même.

La transformée de Fourier discrète est une opération linéaire. Elle transforme un produit de convolution en produit usuel : $\mathcal{F}(f * g) = \mathcal{F}(f) \cdot \mathcal{F}(g)$ (la démonstration de ce résultat est bien connue, voir par exemple [2]).

Dans le cas qui nous intéresse, celui de sous-ensembles du groupe cyclique \mathbb{Z}_c , on définit, à la manière de Lewin, la transformée de Fourier de $A \subset \mathbb{Z}_c$ comme la transformée de Fourier de sa fonction indicatrice :

$$\mathcal{F}_A(t) = \mathcal{F}(\mathbf{1}_A)(t) = \sum_{k \in \mathbb{Z}_c} \mathbf{1}_A(k) e^{-2i\pi kt/c} = \sum_{k \in A} e^{-2i\pi kt/c}$$

On se place en fait dans un cas très particulier de la définition : les fonctions considérées ne sont plus à valeurs dans \mathbb{C} mais dans $\{0, 1\}$. Ceci implique de fortes contraintes quant aux valeurs des coefficients de Fourier, comme nous allons le voir.

D'après la définition, on a directement $\mathcal{F}_A(-t) = \overline{\mathcal{F}_A(t)}$, c'est-à-dire que les coefficients sont réels ou conjugués deux à deux. Le premier coefficient de Fourier $\mathcal{F}_A(0)$ est égal au cardinal de A . Par simple inégalité triangulaire, on a également la relation, pour tout $t \in \mathbb{Z}_c$, $|\mathcal{F}_A(t)| \leq \text{card}(A)$.

Si $A = \mathbb{Z}_c$, on a $\mathcal{F}_{\mathbb{Z}_c} = c \mathbf{1}_{\{0\}}$:

$$\mathcal{F}_A(t) = \sum_{k \in \mathbb{Z}_c} e^{-2i\pi kt/c} = \begin{cases} c & \text{si } t = 0 \\ 0 & \text{sinon} \end{cases} \quad (1.1)$$

Si A est un sous-groupe de \mathbb{Z}_c , c'est-à-dire si A est de la forme $m\{0, \dots, d-1\}$ (que l'on représentera donc par un polygone régulier inscrit dans le cercle) où d est un diviseur de c et $m = c/d$, alors $\mathcal{F}_{m\mathbb{Z}_d} = d \mathbf{1}_{d\mathbb{Z}_m}$:

$$\mathcal{F}_A(t) = \sum_{k \in m\mathbb{Z}_d} e^{-2i\pi kt/c} = \sum_{k \in \mathbb{Z}_d} e^{-2i\pi kt/d} = \begin{cases} d & \text{si } t = nd, n \in \mathbb{N} \\ 0 & \text{sinon} \end{cases}$$

C'est dans ce cas seulement (à translation près) qu'on peut avoir $|\mathcal{F}_A(t)| = \text{card}(A)$ avec $t \neq 0$.

Enfin si A possède une périodicité interne, *i.e.* s'il existe $1 < d < c$ tel que $A + d = A$, alors pour $m = c/d$, si $t \notin m\mathbb{Z}_c$, $\mathcal{F}_A(t) = 0$. En effet, dans ce cas on a pour tout t

$$\mathcal{F}_A(t) = e^{-\frac{2i\pi t}{m}} \cdot \mathcal{F}_A(t)$$

donc si t n'est pas un multiple de m , $\mathcal{F}_A(t)$ est nécessairement nul.

La TFD est enfin un outil de choix pour l'étude des pavages (voir par exemple [14], [13]). On dit qu'une partie A de \mathbb{Z}_c pave \mathbb{Z}_c s'il existe une partie B de \mathbb{Z}_c telle que $\mathbb{Z}_c = A \oplus B$, c'est-à-dire telle que tout entier k s'écrive de manière unique sous la forme $a + b$ (modulo c), avec $(a, b) \in A \times B$. Cette condition est équivalente à $\mathbb{1}_A * \mathbb{1}_B = \mathbb{1}_{\mathbb{Z}_c}$, donc par transformée de Fourier à $\mathcal{F}_A \cdot \mathcal{F}_B = \mathcal{F}_{\mathbb{Z}_c} = c\mathbb{1}_{\{0\}}$. Les ensembles A et B pavent \mathbb{Z}_c si et seulement si

$$\begin{cases} |A| \cdot |B| = c \\ \mathcal{Z}(\mathcal{F}_A) \cup \mathcal{Z}(\mathcal{F}_B) = \mathbb{Z}_c \setminus \{0\} \end{cases}$$

où $\mathcal{Z}(f) = \{x : f(x) = 0\}$ est l'ensemble des zéros de la fonction f .

Par exemple on peut voir ci-dessous un canon rythmique sur \mathbb{Z}_{16} , où $A = \{0, 4, 5, 9\}$, et $B = \{0, 2, 8, 10\}$, représenté à gauche de la figure, et les valeurs de $|\mathcal{F}_A|$ et $|\mathcal{F}_B|$ à droite : on voit qu'à chaque $t \neq 0$, l'un des deux modules s'annule.

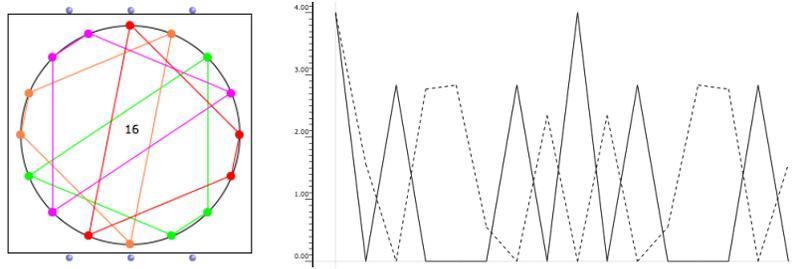


FIGURE 1.3 – Coefficients de Fourier d'un canon rythmique

1.4 Théorème de l'hexacorde

Un résultat bien connu de la *set-theory* est le théorème de l'hexacorde, découvert par Milton Babbitt en 1955 ([5]), puis démontré de différentes façons par Chemillier, Lewin, Wilcox, Mazzola... Ce résultat fait intervenir le *contenu intervallique* d'un accord, qui est simplement la fonction intervallique d'un accord avec lui-même, ce que l'on peut écrire formellement

$$IC_A : k \in \mathbb{Z}_c \mapsto \text{card}\{(x, y) \in A \times A, y - x = k\}$$

Le théorème énonce que tout hexacorde (ensemble de 6 notes) possède le même contenu intervallique que son complémentaire. Emmanuel Amiot a publié dans [1] une preuve très concise de ce résultat utilisant la transformée de Fourier, et que nous allons reprendre ici.

IC_A peut s'écrire comme le produit de convolution $\mathbb{1}_A * \mathbb{1}_{-A}$. Donc, par transformée de Fourier, $\mathcal{F}(IC_A) = \mathcal{F}_A \cdot \mathcal{F}_{-A}$. Par définition, on a immédiatement $\mathcal{F}_{-A} = \overline{\mathcal{F}_A}$, donc

$$\mathcal{F}(IC_A) = |\mathcal{F}_A|^2 \tag{1.2}$$

Soit A un hexacorde et B son complémentaire, alors :

$$\begin{aligned} \mathcal{F}_A(t) + \mathcal{F}_B(t) &= \mathcal{F}_{\mathbb{Z}_{12}}(t) \quad (\text{linéarité de } \mathcal{F}) \\ &= 0 \text{ si } t \neq 0 \text{ (d'après (1.1))} \end{aligned}$$

Donc pour $t \neq 0$, $\mathcal{F}_A(t) = -\mathcal{F}_B(t)$, et $\mathcal{F}_A(0) = \mathcal{F}_B(0) = 6$. Alors $|\mathcal{F}_A| = |\mathcal{F}_B|$, d'où $\mathcal{F}(IC_A) = \mathcal{F}(IC_B)$ et comme la transformée de Fourier caractérise complètement une fonction, $IC_A = IC_B$. \square

Ce résultat s'étend évidemment à toute cardinalité paire n autre que 12, un « hexacorde » désignant alors un ensemble de $n/2$ notes. John Mandereau a étendu dans [20] ce théorème au cas continu, en se plaçant sur un groupe topologique.

1.5 Invariance du module

Le module des coefficients de Fourier possède des propriétés d'invariance par certaines relations mathématiques aptes à modéliser des opérations musicales pertinentes. On va retrouver les opérations de la *set-theory* musicale américaine (voir [11]).

1.5.1 Actions de groupes

Dans notre modèle, la modulation d'un accord ou d'une gamme, qui est une translation dans le domaine des hauteurs, est une translation de l'ensemble A . La transformée de Fourier de l'ensemble translaté $A + p$ est

$$\mathcal{F}_{A+p}(t) = e^{-2i\pi t/c} \mathcal{F}_A(t)$$

L'opération d'inversion ($A \mapsto -A$) laisse elle aussi le module inchangé :

$$\mathcal{F}_{-A}(t) = \overline{\mathcal{F}_A(t)}$$

Le module des coefficients de Fourier est invariant par l'action du groupe diédral D_{2c} . Signalons au passage l'action du groupe affine, utilisé en musique notamment par Estrada : pour d premier avec c ,

$$\mathcal{F}_{dA}(t) = \mathcal{F}_A(dt) \tag{1.3}$$

Contrairement aux opérations du groupe diédral, la multiplication affine permute les coefficients de Fourier, elle ne préserve donc pas la fonction $|\mathcal{F}(t)|$.

Cette préservation des modules par l'action du groupe diédral invite à travailler sur l'ensemble des parties de \mathbb{Z}_c quotienté par D_{2c} , en considérant des classes d'équivalence de sous-ensembles du groupe cyclique. Les actions du groupe diédral correspondent aux opérations T/I (transposition/inversion) de la *set-theory*, les espaces quotients découlant de celles-ci ont été largement étudiés par Quinn, Callender et Tymoczko dans [7]. Mais la relation d'équivalence $A \sim B \Leftrightarrow |\mathcal{F}_A| = |\mathcal{F}_B|$ donne des classes d'équivalence plus grandes, comme nous avons le voir maintenant.

1.5.2 Z-relation, homométrie

La Z-relation a été introduite par Allen Forte dans [11]. On la définit comme suit :

$$A \text{ et } B \text{ sont en Z-relation} \Leftrightarrow IC_A = IC_B$$

D'après l'équation (1.2), cette relation équivaut à $|\mathcal{F}_A| = |\mathcal{F}_B|$. On voit directement que le cas de deux hexacordes complémentaires est un cas particulier de cette relation, et que l'appartenance à la même orbite pour l'action de D_{2c} implique la Z-relation. Dans ce dernier cas, où A et B sont liés par transposition/inversion, on parle de Z-relation *triviale*. La définition originelle de cette relation exclut ce cas trivial, ce qui peut sembler légitime, mais pose problème mathématiquement : la relation définie alors, n'étant pas réflexive, n'est pas une relation d'équivalence. Nous considérons dans la suite la Z-relation au sens large, cas triviaux compris, ce qui nous autorise à écrire $A \sim_Z B$.

La question de savoir, connaissant un ensemble, s'il en existe qui soient en Z-relation non triviale avec lui, est encore ouverte. Le théorème de l'hexacorde permet, pour un cardinal bien précis, et dès lors que les classes d'équivalence sous l'action de D_{2c} de l'ensemble et de son complémentaire sont disjointes, d'expliciter de tels couples. Le plus « simple » apparaît sur \mathbb{Z}_8 , voir figure 1.4.

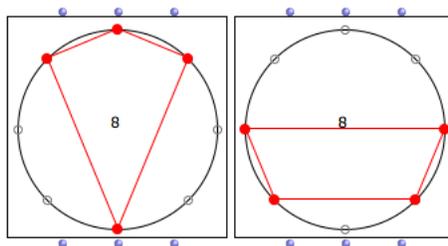


FIGURE 1.4 – Plus petite Z-relation non triviale

Par les propriétés d'invariance du module sous l'action de D_{2c} , les classes d'équivalence de la Z -relation contiennent les orbites de l'action du groupe diédral, elles sont strictement plus grandes s'il existe une relation non triviale. Mandereau, dans [20], a démontré que la Z -relation n'est pas une action de groupe : il n'existe pas de groupe H tel que $\mathcal{P}(\mathbb{Z}_c)/H = \mathcal{P}(\mathbb{Z}_c)/\sim_Z$. De ce fait les grands théorèmes d'algèbre (notamment l'équation de Burnside-Frobenius) ne peuvent malheureusement pas s'appliquer (du moins pas directement) à ce problème.

Le fait que les orbites de $\mathcal{P}(\mathbb{Z}_c)$ sous l'action de D_{2c} soient incluses dans les classes d'équivalence de la Z -relation permet de considérer celle-ci non plus entre des éléments de $\mathcal{P}(\mathbb{Z}_c)$ mais entre des orbites de ceux-ci par translation/inversion. En effet d'après ce qu'on a vu, si $A \sim_Z B$, alors pour tout C dans l'orbite de B sous l'action de D_{2c} , $A \sim_Z C$. D'autre part, puisque pour $t \neq 0$, $\mathcal{F}_{A^c}(t) = -\mathcal{F}_A(t)$ (d'après 1.1), donc $|\mathcal{F}_{A^c}(t)| = |\mathcal{F}_A(t)|$, on a

$$A \sim_Z B \Leftrightarrow A^c \sim_Z B^c$$

De même, d'après (1.3), pour d premier avec c

$$A \sim_Z B \Leftrightarrow dA \sim_Z dB$$

La Z -relation est en fait un cas particulier de l'homométrie, notion originaire du domaine de la cristallographie, où la question est de retrouver la structure d'un cristal à partir de sa figure de diffraction. Musicalement, le but est de reconstruire un accord à partir de son contenu intervallique, ou, de manière équivalente, des modules de ses coefficients de Fourier. Notons que cette question est proche de celle posée par Lewin dans [17]. Cette problématique de la *reconstruction de la phase* est toujours ouverte, et fait l'objet de nombreux travaux de recherche actuels ([16], [21]), où la transformée de Fourier apparaît comme un outil indispensable.

1.6 Extensions possibles du modèle

Il est possible d'envisager des modèles plus riches que le nôtre et qui se prêtent encore bien à l'analyse par la TFD. Tout d'abord, il est possible de passer aisément des sous-ensembles aux distributions sur \mathbb{Z}_c , c'est-à-dire de considérer non plus des fonctions f de \mathbb{Z}_c à valeurs dans $\{0, 1\}$, mais dans $[0, 1]$, voire dans \mathbb{R} . La fonction de départ étant toujours à valeurs réelles, les coefficients restent conjugués deux à deux dans l'espace de Fourier. En effet :

$$\begin{aligned} f \text{ est à valeurs réelles} &\Leftrightarrow f(k) = \overline{f(k)} \text{ pour } k \in \mathbb{Z}_c \\ &\Leftrightarrow \sum_{k \in \mathbb{Z}_c} f(k) e^{-2i\pi kt/c} = \sum_{k \in \mathbb{Z}_c} \overline{f(k)} e^{-2i\pi kt/c} \text{ pour } t \in \mathbb{Z}_c \\ &\Leftrightarrow \sum_{k \in \mathbb{Z}_c} f(k) e^{-2i\pi kt/c} = \overline{\sum_{k \in \mathbb{Z}_c} f(k) e^{2i\pi kt/c}} \\ &\Leftrightarrow \mathcal{F}(f)(t) = \overline{\mathcal{F}(f)(-t)} \end{aligned}$$

Dans ce cas, le premier coefficient $\mathcal{F}(0)$ vaut $\sum_{k \in \mathbb{Z}_c} f(k)$, ce qui étend bien la définition du cardinal d'un ensemble aux distributions.

Ce modèle permet par exemple de tenir compte de nuances au sein d'un rythme ou d'un accord. La présentation de ce travail faite le 18 juin au Palais de la Découverte dans le cadre du festival Agora utilisait ce modèle : il s'agissait d'un dispositif où l'utilisateur module l'intensité de 8 attaques de percussion, jouées en boucle, et peut ainsi créer un rythme puis le modifier en jouant sur les nuances. À l'écran est représenté l'espace de Fourier, où apparaissent les huit coefficients de la distribution f à valeurs dans $[0, 1]$ ainsi définie. Il est ensuite possible de passer au rythme inverse, en prenant les conjugués des coefficients (symétrie par rapport à l'axe réel), ou au « complémentaire » (opération étendue par $x \mapsto 1 - x$), en prenant les opposés des coefficients $\mathcal{F}(t), t \neq 0$ (symétrie par rapport à l'origine), et donnant à $\mathcal{F}(0)$ la valeur $8 - \sum_{k \in \mathbb{Z}_8} f(k)$.

Thomas Noll et Martin Carlé ont proposé l'approche du *Fourier scratching* (voir [4], [22]), impliquant un troisième modèle. Le principe est d'agir directement dans l'espace de Fourier en déplaçant les coefficients librement sur le plan complexe, afin d'obtenir de nouveaux motifs musicaux par transformée de Fourier inverse. La figure 1.5 montre l'interface du dispositif, à gauche est représenté le motif musical, à droite les coefficients de Fourier, tous deux sur une sphère de Riemann.

On a vu que pour qu'une fonction soit réelle, il faut que ses coefficients soient conjugués deux à deux, ce qui impliquerait de déplacer simultanément deux points. Le choix fait ici est, à l'inverse, d'étendre le domaine de la fonction modélisant le motif musical à tout \mathbb{C} . Les valeurs prises par f ne paramètrent donc plus de la même façon le motif musical joué. L'intensité sonore est donnée par le module de f (bien évidemment ajustée par une fonction bornée), et la « couleur sonore » (concrètement des paramètres de synthèse FM) est donnée par la phase.

Ce dernier modèle, bi-dimensionnel, est bien plus sophistiqué que les précédents. S'il permet un rendu sonore musicalement plus intéressant, moins austère, sa complexité rend plus délicate la mise en évidence de propriétés remarquables, et nous n'avons pas cherché à l'exploiter sur le plan théorique. Ce dispositif a d'ailleurs été conçu dans un but de performance musicale et non d'analyse.

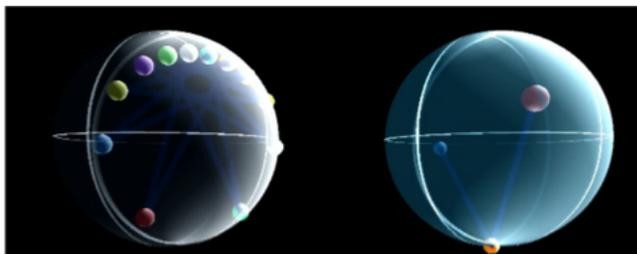


FIGURE 1.5 – Interface du *Fourier scratching* (source : [22])

1.7 Pistes concernant l'homométrie

À défaut d'apporter de réels éléments de réponse au problème de la reconstruction de la phase, nous présentons ici deux pistes novatrices, qui, nous l'espérons, pourront être reprises avec profit à l'avenir.

1.7.1 Espace de définition des phases

La première approche, proposée par Emmanuel Amiot, est topologique. On a constaté que la donnée des modules des coefficients de Fourier caractérisait des classes d'équivalence pour la relation Z , qui sont des réunions d'orbites (éventuellement une seule) pour l'action de D_{2c} sur l'ensemble des parties de \mathbb{Z}_c , elles-mêmes de cardinal variable selon l'existence ou non de symétrie ou période internes. Cela signifie qu'en changeant les phases des coefficients, il est possible de parcourir l'ensemble des éléments de cette classe d'équivalence. La phase étant périodique, c'est-à-dire paramétrable par le cercle \mathbb{S}_1 , l'idée est alors de paramétrer l'ensemble des phases par un point d'un tore.

La phase d'un nombre complexe nul n'étant pas définie, il est nécessaire de faire attention aux coefficients de Fourier nuls (ou encore : aux *FourierProperties*) pour bien définir l'espace auquel on va s'intéresser. Pour un ensemble $A \subset \mathbb{Z}_c$, notons n_A le nombre de coefficients de Fourier nuls : seules $c - n_A$ phases sont donc définies. Si l'on veut rester dans le cadre des distributions réelles sur \mathbb{Z}_c , il est nécessaire d'une part d'avoir $\mathcal{F}(0)$ réel, et d'autre part de conserver la relation $\mathcal{F}(t) = \overline{\mathcal{F}(-t)}$, ce qui va diviser le nombre de degrés de liberté par deux (changer la phase de $\mathcal{F}(t)$ d'un angle θ implique de changer la phase de $\mathcal{F}(-t)$ de $-\theta$). Enfin, dans le cas où c est pair, $\mathcal{F}(c/2)$ doit être réel. Soit $\mathcal{F}_A(c/2)$ est nul, auquel cas la phase, n'étant pas définie, ne pourra pas varier, soit $\mathcal{F}_A(c/2)$ est non nul, et sa phase peut valoir 0 ou π .

Pour résumer : si c est impair, ou bien si c est pair et $\mathcal{F}_A(c/2) = 0$, alors on peut paramétrer l'ensemble des phases des coefficients de Fourier par un point du tore $\mathbb{T}_d = \prod_{k=1\dots d} \mathbb{S}_1$, où $d = \frac{c-n_A-1}{2}$. Si c est pair, et $\mathcal{F}_A(c/2)$ non nul, on doit se placer sur l'espace $\mathbb{T}_d \times \{0, \pi\}$, avec cette fois $d = \frac{c-n_A-2}{2}$, ce qui correspond topologiquement à la réunion de deux tores.

Pour un ensemble A , l'espace ainsi défini contient les points paramétrant chacun des éléments de la classe d'équivalence de A par la relation Z . On peut alors se poser la question de savoir si les différences de cardinal de cette classe (dans le cas où A possède un axe de symétrie, *i.e.* $A = -A + k$, ou une période interne, *i.e.* $A = A + k$, avec $k \neq c$, ce qui correspond aux modes à transpositions limitées d'Olivier Messiaen), et l'existence de relation Z non triviale, correspondent à des cas de figure particuliers de cet espace topologique. De plus cet espace permet de faire varier continûment les phases, et donc de passer d'un sous-ensemble, c'est-à-dire d'une fonction à valeurs dans $\{0, 1\}$, à un autre par l'intermédiaire de distributions à valeurs dans \mathbb{R} .

1.7.2 Quasi-recouvrement du cercle

Cette dernière remarque concerne la structure des *Z-familles*, c'est-à-dire les ensembles d'ensembles en relation Z non triviale (ou encore : les classes d'équivalence de la Z -relation quotientées par le groupe diédral). Cela revient à considérer la relation Z non plus entre des ensembles A et B mais entre leurs orbites par l'action de D_{2c} .

Une structure remarquable de Z -familles est celle des hexacordes complémentaires, pour lesquels, avec les bons représentants des deux ensembles, $A \cup B = \mathbb{Z}_c$. Nous avons observé une structure intéressante, mais non générale, pour d'autres familles.

Sur \mathbb{Z}_{12} , les ensembles $\{0, 1, 3, 7\}$ et $\{0, 1, 7, 9\}$ sont en Z -relation ; on peut choisir des transpositions/inversions particulières A et B de ces deux ensembles telles que $(A \cup B)^c = \{0, 3, 6, 9\} = 3\mathbb{Z}_4$. Pour les ensembles de cardinal 5, on a trois Z -familles, dont deux ont cette même propriété (*i.e.* on peut recouvrir \mathbb{Z}_{12} moins un sous-groupe, dans ce cas $6\mathbb{Z}_2$). Les ensembles de cardinal 5 sur \mathbb{Z}_{15} permettent de même d'avoir $(A \cup B)^c = 3\mathbb{Z}_5$ (voir figure 1.6).

Pour $c = 18$ un phénomène semblable mais plus complexe apparaît : les ensembles $A = \{0, 1, 2, 3, 5, 6, 7, 9, 13\}$ et $B = \{1, 7, 8, 11, 12, 13, 14, 15, 17\}$ sont en Z -relation, et ont la structure remarquable

$$\begin{cases} (A \cap B) = 6\mathbb{Z}_3 + 1 \\ (A \cup B)^c = 6\mathbb{Z}_3 - 2 \end{cases}$$

Si cette observation ne se généralise pas à toutes les Z -familles (sur \mathbb{Z}_{12} il existe une paire d'ensembles de cardinal 5 qui n'a pas cette structure), elle semble être suffisamment fréquente pour mériter d'être étudiée plus en profondeur. Du fait de ses propriétés et valeurs remarquables pour des sous-groupes, la transformée de Fourier peut être un outil adapté pour cela.

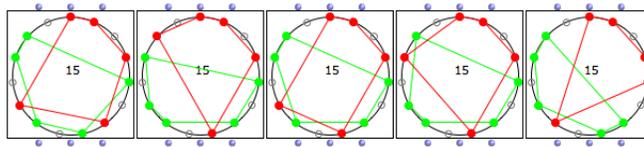


FIGURE 1.6 – Z -familles de cardinal 5 dans \mathbb{Z}_{15}

Chapitre 2

Gammes et structures bien réparties

La notion de gamme bien répartie, ou *GBR*, a été introduite en musique par John Clough et Jack Douthett, théoriciens de la musique appartenant à la tradition diatonique américaine, dans [9]. En plus d'être pertinente sur le plan musical, de part la quantité de structures rythmiques et harmoniques qu'elle permet de caractériser, elle trouve des applications dans d'autres domaines, notamment en physique (voir [10]). Après avoir défini ces objets, nous verrons que les étudier *via* la TFD donne des résultats intéressants. L'essentiel de cette partie est inspiré de [2].

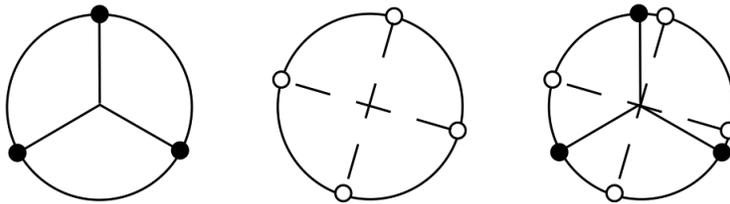
2.1 Définition originelle

Avant d'arriver à un formalisme mathématique clair, nous allons partir de l'approche géométrique, donnant bien l'intuition de ce qu'est une *GBR*. Une gamme bien répartie à d éléments sur \mathbb{Z}_c , notée $GBR_{c,d}$, désigne la répartition de d points parmi c la plus homogène possible. Avec la représentation circulaire, cela revient à trouver la meilleure approximation du polygone régulier à d côtés par une sous-figure du polygone à c côtés. Dans le cas où d divise c , la gamme bien répartie *est* le polygone régulier, ou c'est-à-dire le sous-groupe $m\mathbb{Z}_d$, avec $m = c/d$.

Une première construction possible pour arriver à cette approximation consiste à superposer deux cercles discrétisés l'un en d parties égales, l'autre en $c - d$ parties égales, de façon à ce qu'aucun point n'en recouvre un autre. Ce procédé est illustré figure 2.1 : la répartition des points noirs sur le dernier cercle donne l'ensemble $\{0, 3, 5\} = GBR_{7,3}$.

La rotation appliquée au second cercle peut sembler arbitraire, et c'est bien le cas : un autre choix aurait pu donner un résultat différent pour $GBR_{7,3}$, mais celui-ci aurait été un translaté de $\{0, 3, 5\}$. L'homogénéité de la répartition des points blancs et noirs sur le cercle n'est pas affectée par une rotation de celui-ci, il est donc souhaitable d'arriver à la définition d'une classe d'équivalence par translation plutôt qu'à un unique ensemble.

L'approche géométrique exposée ici, si elle est la plus intuitive, manque d'un formalisme permettant des définitions et des raisonnements mathématiques rigoureux. Il est nécessaire de poser une définition plus formelle des gammes bien réparties.

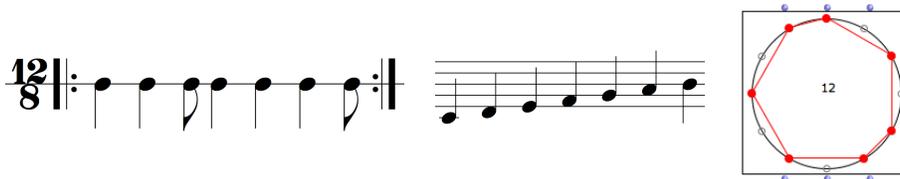

 FIGURE 2.1 – Construction géométrique d'une $GBR_{7,3}$ (source : [10])

La définition proposée dans [9] reprend l'idée de l'approximation d'un polygone régulier avec une approche arithmétique. La gamme bien répartie est définie non comme un ensemble mais comme une suite ordonnée $(a_k)_{k \in \mathbb{Z}_d}$, où pour $\alpha \in \mathbb{Z}_c$,

$$a_k = J_{c,d}^\alpha(k) = \left\lfloor \frac{kc + \alpha}{d} \right\rfloor$$

$\lfloor x \rfloor$ désignant la partie entière de x . Le paramètre α permet de considérer les différentes rotations du cercle. Notons une propriété remarquable (démontrée plus loin) des gammes bien réparties : le complémentaire d'une GBR est encore une GBR . Intuitivement, si l'on revient au modèle purement géométrique exposé figure 2.1, on voit que le complémentaire de la $GBR_{7,3}$ (points noirs) s'obtient en considérant les points blancs, c'est-à-dire en construisant $GBR_{7,4}$ avec la même méthode.

Pour $c = 12, d = 7, \alpha = -1$, les valeurs de a_k décrivent la gamme diatonique, *i.e.* les touches blanches du piano ; pour $d = 5, \alpha = 6$, elles correspondent aux touches noires, c'est-à-dire à la gamme pentatonique, complémentaire de la précédente. De nombreux rythmes courants dans différentes musiques peuvent également être vus comme des gammes bien réparties (voir à ce sujet [25]). La clave *Bembé* par exemple, fréquente en musique africaine ou cubaine, décrit le même polygone que la gamme diatonique : la $GBR_{12,7}$ modélise aussi bien un motif rythmique qu'une échelle harmonique courants dans la musique « traditionnelle ». Pour cette raison, on parlera désormais de *structure* bien répartie, ou *SBR*, afin d'insister sur la généralité de cet objet.


 FIGURE 2.2 – Clave *Bembé*, gamme diatonique, représentation circulaire

2.2 Typologie des structures bien réparties

Nous allons reprendre la classification de [2], inspirée de [23], sans donner les démonstrations de tous les résultats (dont certaines découlent plus simplement de la définition équivalente des *SBR* que nous donnerons à la fin de cette partie). Il existe trois types de structure bien répartie selon la valeur de $m = \text{pgcd}(c, d)$:

- on a une *SBR* de type *I* lorsque $m = 1$. Dans ce cas l'ensemble des valeurs prises par (a_k) est égal, à translation près, à $d^{-1}\{0, \dots, d - 1\}$, où d^{-1} désigne l'inverse de d pour la multiplication sur \mathbb{Z}_c ,
- on a une *SBR* de type II_a dans le cas où $m = d$, c'est-à-dire si d divise c . Alors $SBR_{c,d}$ est la division régulière du cercle en d parties,
- si $1 < m = c - d < d$, on a une structure bien répartie de type II_b . Il s'agit alors du complémentaire d'une structure de type II_a ,
- on a une *SBR* de type *III* dans le cas restant, *i.e.* si $1 < m < d$ et $m \neq c - d$.

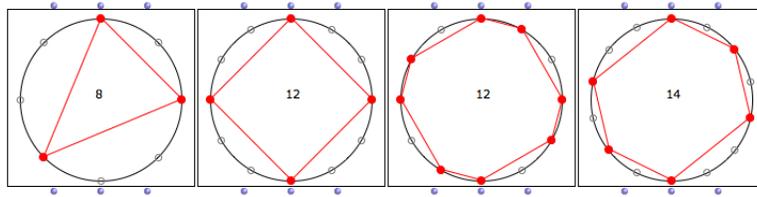


FIGURE 2.3 – Structures bien réparties de type *I*, II_a , II_b et *III*

Les types *II* et *III* regroupent toutes les structures bien réparties possédant une période interne. On peut en fait toujours se ramener à des *SBR* de type *I* en observant que toutes les autres sont constituées de réunions de translatées de $SBR_{c',d'}$, avec $c' = c/m, d' = d/m$ par le sous-groupe $c'\mathbb{Z}_m$. Ainsi, dans les exemples de *SBR* figure 2.3, $SBR_{12,4}$ est bien égal à $SBR_{3,1} \oplus 3\mathbb{Z}_4$, $SBR_{3,1}$ valant simplement $\{0\}$. De même, $SBR_{12,8} = SBR_{3,2} \oplus 3\mathbb{Z}_4$, et $SBR_{14,6} = SBR_{7,3} \oplus 7\mathbb{Z}_2$. On peut donc dire que les *SBR* de type *I* sont *fondamentales*.

Seules les *SBR* de type *I* existent pour toute valeur de c . Celles de type *II* existent dès lors que c est composé, et celles de type *III* si c est composé et $c > 12$ (voir [2]).

2.3 Coefficients de Fourier d'une SBR

Avant de donner une définition équivalente des structures bien réparties utilisant les modules des coefficients de Fourier, nous donnons un résultat inédit, concernant la valeur des coefficients de Fourier d'un tel objet. Pour cela, on ne considère plus une *SBR* comme une suite ordonnée $(a_k)_k$, mais comme l'ensemble $\{a_k\}$ de ses valeurs.

Théorème. Soit d premier avec c , $l = d^{-1}$ l'inverse multiplicatif de d sur \mathbb{Z}_c . Soit A la structure bien répartie à d éléments, et B_l le cluster chromatique à l éléments, i.e. $\{0, \dots, l-1\}$. Pour tout $t \neq 0$,

$$\mathcal{F}_{B_l}(t) = \frac{1}{\overline{\mathcal{F}_A(t)}}$$

Démonstration. A est une structure bien répartie de type I que l'on peut exprimer par $A = -dB_l$. On note B_d le cluster chromatique à d éléments $\{0, \dots, d-1\}$. Considérons alors, dans \mathbb{Z} , l'ensemble $B_d + dB_l$:

$$\begin{aligned} B_d + dB_l &= \{0, \dots, d-1\} \oplus d\{0, \dots, l-1\} \\ &= \{0, \dots, d-1, d, d+1, \dots, d-1 + d(l-1)\} \\ &= \llbracket 0, dl-1 \rrbracket \end{aligned}$$

Notons $a = (dl-1)/c$ (calculé dans \mathbb{Z}). Si on quotiente cet ensemble modulo c , puisque $dl-1 = 0$, on obtient le multi-ensemble $(\{0, \dots, c-1\}, m)$, où la multiplicité $m(t) = a$ pour $t \neq 0$, et $m(0) = a+1$.

On peut écrire le produit de convolution

$$\mathbb{1}_{B_d} * \mathbb{1}_{dB_l}(t) = \sum_{x \in \mathbb{Z}_c} \mathbb{1}_{B_d}(x) \mathbb{1}_{dB_l}(t-x) = m(t)$$

En appliquant la transformée de Fourier à cette expression, on obtient

$$\mathcal{F}_{B_d} \cdot \mathcal{F}_{dB_l} = \mathcal{F}(m)$$

Or, $m = a\mathbb{1}_{\mathbb{Z}_c} + \mathbb{1}_0$, et on sait que pour tout $t \neq 0$, $\mathcal{F}_{\mathbb{Z}_c}(t) = 0$, donc par linéarité de \mathcal{F} , on a, pour $t \neq 0$,

$$\mathcal{F}_{B_d}(t) \cdot \mathcal{F}_{dB_l}(t) = \mathcal{F}_0(t) = 1 \tag{2.1}$$

D'où la relation annoncée dans le théorème : il est possible de passer à l'inverse puisque les deux fonctions \mathcal{F}_{B_d} et \mathcal{F}_{dB_l} sont bien définies sur \mathbb{Z}_c . \square

Géométriquement, cela signifie que les coefficients de Fourier d'une structure bien répartie sont les inverses (au sens de l'inversion par rapport au cercle unité $z \mapsto 1/|z|$), de ceux d'un cluster chromatique.

Corollaire. Les coefficients de Fourier d'un cluster chromatique de cardinal premier avec c sont tous non nuls.

Cette démonstration, utilisant les propriétés remarquables de la transformée de Fourier discrète, est la première que nous ayons proposée pour ce théorème. Emmanuel Amiot a remarqué qu'il était possible de montrer bien plus rapidement la relation (2.1) en revenant à la définition même de \mathcal{F} . Le résultat provient alors simplement des règles élémentaires du calcul complexe.

Pour l'instant seules les structures de type I sont concernées, mais on peut généraliser ce théorème à toutes les autres. Considérons $A = SBR_{c,d}$, avec $m = \text{pgcd}(c, d) \neq 1$. Alors, avec les notations utilisées plus haut, A est c/m -périodique, de la forme $A' \oplus c'\mathbb{Z}_m$, où $A' = SBR_{c',d'}$ est de type I . Notons $B' = \{0, \dots, l-1\}$ le cluster chromatique de cardinal $l = d'^{-1}$ (inverse multiplicatif sur $\mathbb{Z}_{c'}$), et B l'ensemble $B' \oplus c'\mathbb{Z}_m$. On sait alors que pour $t \neq 0$

$$\overline{\mathcal{F}_{A'}(t)} \cdot \mathcal{F}_{B'}(t) = 1$$

Ensuite, on a

$$\begin{aligned} \mathcal{F}_A(mt) &= \sum_{k \in A} e^{\frac{-2i\pi kmt}{c}} \\ &= \sum_{k \in \mathbb{Z}_c} \mathbb{1}_A(k) e^{\frac{-2i\pi kt}{c'}} \\ &= m \sum_{k \in \mathbb{Z}_{c'}} \mathbb{1}_{A'}(k) e^{\frac{-2i\pi kt}{c'}} \\ &= m \cdot \mathcal{F}_{A'}(t) \end{aligned}$$

et on a de même $\mathcal{F}_B(mt) = m\mathcal{F}_{B'}(t)$, d'où

$$\overline{\mathcal{F}_A(mt)} \cdot \mathcal{F}_B(mt) = m^2$$

A et B étant c' -périodiques, on sait que $\mathcal{F}_A(t) = \mathcal{F}_B(t) = 0$ pour tout $t \notin m\mathbb{Z}_{c'}$. On peut donc généraliser le théorème.

Théorème. *Pour $d < c$, soient $m = \text{pgcd}(c, d)$, $c' = c/m$, $d' = d/m$ et $l = d'^{-1}$ (inverse sur l'anneau $\mathbb{Z}_{c'}$). On pose $A = SBR_{c,d}$ et $B = \{0, \dots, l-1\} \oplus \mathbb{Z}_m$, alors*

$$\begin{cases} \overline{\mathcal{F}_A(t)} \cdot \mathcal{F}_B(t) = m^2 & \text{si } t \in m\mathbb{Z}_{c'} \\ \mathcal{F}_A(t) = \mathcal{F}_B(t) = 0 & \text{sinon} \end{cases}$$

Si cette généralisation permet bien de prendre en compte tous les types de SBR , il faut noter que pour les types II_a et II_b , le théorème devient tautologique du fait que $A = B$.

2.4 Caractérisation par la TFD

Il est possible de définir les structures bien réparties grâce aux coefficients de la TFD. Cette approche a été suggérée dans [23], et étudiée en détail dans [2]. L'objet défini n'est plus une suite d'entiers, ni un sous-ensemble de \mathbb{Z}_c , mais une classe d'équivalence de parties de \mathbb{Z}_c pour la relation Z . Par souci de lisibilité, on parlera malgré tout d'*ensemble*, en sous-entendant la classe de celui-ci.

Définition. La structure bien répartie à d éléments dans \mathbb{Z}_c est l'ensemble qui maximise la valeur du d -ième coefficient de Fourier :

$$SBR_{c,d} = \underset{\substack{P \in \mathcal{P}(\mathbb{Z}_c) / \sim_Z \\ |P|=d}}{\operatorname{argmax}} |\mathcal{F}_P(d)|$$

Les classes d'équivalence ainsi définies sont en fait des orbites du groupe des translations \mathbb{Z}_c : il n'existe pas de relation Z non triviale impliquant une SBR , et une SBR possède toujours un axe de symétrie. De plus, le cardinal de cette orbite est strictement inférieur à c dès lors qu'il existe une période interne (types *II* et *III*).

Pour les structures de type *I*, la clé de la démonstration est le fait que $|\mathcal{F}_A(1)|$ est maximal lorsque les points de A sont consécutifs dans \mathbb{Z}_c (voir figure 2.4). Alors, puisqu'une $SBR_{c,d}$ de type *I* s'écrit $A = d\{0, \dots, l-1\}$, et d'après (1.3), la valeur de $|\mathcal{F}_A(d)|$ est bien maximale.

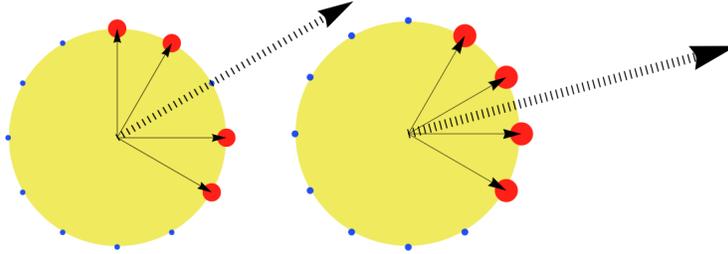


FIGURE 2.4 – Maximisation de $|\mathcal{F}_A(1)|$ (source : [3])

Cette définition se généralise aux autres types de structures bien réparties. Avec cette définition, le fait que le complémentaire d'une structure bien répartie soit encore une structure bien répartie est évident. Signalons l'étude faite par Emmanuel Amiot dans [3] de différents tempéraments, utilisant ces caractérisations par modules maximaux, mais avec un modèle et une définition de \mathcal{F} différents.

Chapitre 3

Utiliser la TFD pour comparer les rythmes

On a vu que la transformée de Fourier, et particulièrement le module des coefficients, constitue un outil puissant pour l'étude théorique d'objets musicaux. On se pose alors la question de savoir si cette opération mathématique peut avoir une dimension perceptive, c'est-à-dire si ses propriétés peuvent correspondre à des phénomènes musicaux audibles ou si elle est cantonnée à l'étude théorique des modèles abstraits d'objets musicaux. Notre étude se concentre sur les rythmes périodiques, objets globalement négligés au profit des ensembles de classes de hauteur, et pour lesquels nous allons définir une distance basée sur les valeurs de la TFD.

3.1 Définition et propriétés de la distance $d_{\mathcal{F}}$

Pour deux rythmes A et B définis sur une période de c pulsations, on pose

$$d_{\mathcal{F}}(A, B) = \sum_{k \in \mathbb{Z}_c} \left| |\mathcal{F}_A(k)| - |\mathcal{F}_B(k)| \right|$$

La symétrie et l'inégalité triangulaire se vérifient trivialement. En revanche pour avoir la propriété de séparation et donc que $d_{\mathcal{F}}$ soit bien une distance au sens mathématique du terme, du fait des propriétés d'invariance du module, il est nécessaire de la considérer sur l'espace des parties de \mathbb{Z}_c quotienté par la relation Z . Là encore nous parlerons d'*ensembles* malgré ce quotient. En particulier les rythmes seront donc considérés à translation près, ce qui semble pertinent sur le plan perceptif : une fois qu'un rythme est joué en boucle, on perçoit le cycle qui s'installe alors, et l'instant de départ par rapport à la mesure devient négligeable (ce fait semblait suffisamment évident sur les exemples sonores pour se dispenser d'un test perceptif spécifique).

Cette distance ne permet de comparer que des rythmes définis sur une même métrique (*i.e.* définis avec la même période c), mais ceux-ci peuvent comporter un nombre de notes différent. Le fait que la TFD mette en évidence les périodicités d'une fonction fait de $d_{\mathcal{F}}$

une distance « sensible » aux périodicités des rythmes. Par exemple, dans \mathbb{Z}_{12} , entre le rythme $\{0, 4, 8\}$ (division en trois temps réguliers) et le rythme $\{0, 1, 4, 5, 8, 9\}$, (les trois temps dédoublés) la périodicité est conservée, et la distance est relativement faible. À l'inverse, entre $\{0, 4, 8\}$ et $\{0, 3, 6, 9\}$ (division en quatre temps) la périodicité change, et la valeur de la distance est alors bien plus grande.

Enfin elle a l'avantage d'être directement calculable, sans qu'il y ait besoin d'exécuter un algorithme comme par exemple pour la distance de Hamming généralisée (voir [6]).

3.2 Évaluation perceptive de $d_{\mathcal{F}}$

La réalisation d'une expérience perceptive est un travail ambitieux et de longue durée, aussi nous avons dû nous limiter à un cadre restreint pour mener à bien une étude cohérente dans le cadre de ce stage : si la distance $d_{\mathcal{F}}$ permet de travailler sur toute cardinalité c , c'est-à-dire sur des rythmes de toutes les métriques, nous avons décidé de ne considérer que le cas $c = 12$, modélisant des structures rythmiques usuelles (4 temps ternaires, trois temps binaires).

3.2.1 Protocole d'écoute

Nous avons choisi aléatoirement 20 rythmes de référence $R_{i,ref}, i = 1, \dots, 20$, puis pour chacun d'eux 4 autres rythmes $R_{i,j}, j = 1, \dots, 4$ en fonction des distances $d_{\mathcal{F}}$ à $R_{i,ref}$: $R_{i,1}$ le plus proche de la référence, $R_{i,4}$ le plus distant, et $R_{i,2}$ et $R_{i,3}$ deux rythmes situés à des distances intermédiaires, choisis aléatoirement. Ces rythmes sont en fait choisis parmi les classes d'équivalence d'ensembles par translation/inversion, donc sur un espace plus grand que le domaine de définition de $d_{\mathcal{F}}$. Ainsi il arrivera que la distance théorique soit nulle, dans le cas d'une Z-relation non triviale, ce qui constitue une légère entorse à la définition formelle de $d_{\mathcal{F}}$.

Nous avons ainsi constitué une base de données de 20 ensembles de 5 rythmes, puis généré les 100 fichiers MIDI correspondants, où le rythme est joué par un son de clave et est répété trois fois. L'auditeur doit alors, pour chacun de ces ensembles, évaluer la ressemblance de chacun des quatre rythmes $R_{i,j}$ avec le rythme de référence $R_{i,ref}$, en déplaçant les curseurs (voir figure 3.1). Pour limiter les effets de contexte, l'ordre des ensembles, et l'attribution des labels A, B, C, D aux quatre rythmes $R_{i,j}$ à comparer, sont aléatoires.

Le test a été effectué par 21 sujets, choisis sans critères particuliers. Le nombre d'écoutes et le temps passé sur chaque ensemble n'étant pas limités, la durée globale de l'expérience dépendait de la stratégie d'écoute de chaque sujet, et a varié de 15 à 45 minutes.

3.2.2 Analyse des résultats

L'utilisateur indique des *distances perceptives* $d_P(R_{i,ref}, R_{i,j})$ pour les 20 ensembles entre les quatre rythmes et la référence, que nous comparons ensuite aux valeurs de la

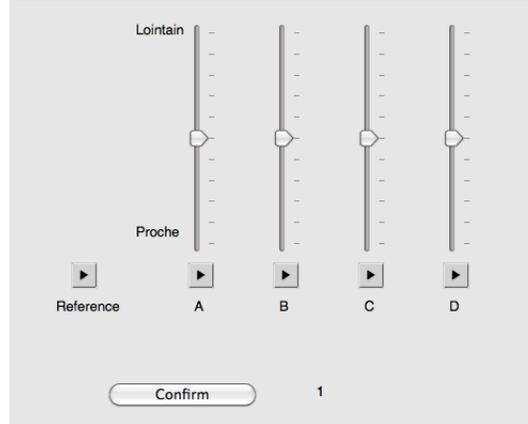
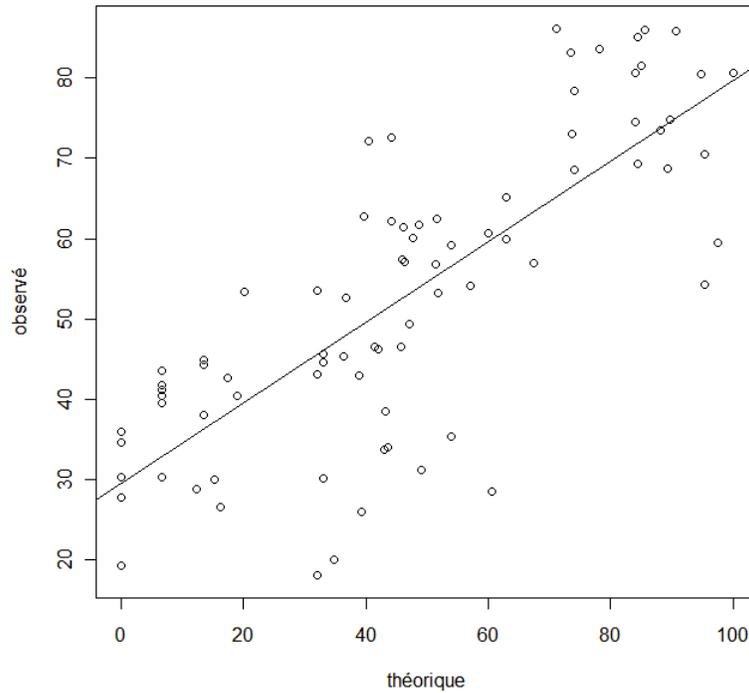


FIGURE 3.1 – Interface du test perceptif

distance théorique $d_{\mathcal{F}}$. Si cette distance est pertinente sur le plan perceptif, c'est-à-dire si une valeur élevée (respectivement faible) de $d_{\mathcal{F}}(X, Y)$ correspond à une sensation de dissemblance (respectivement ressemblance) entre les rythmes X et Y chez l'auditeur, et donc à une valeur élevée (respectivement faible) de d_P , nous devrions arriver à un fort coefficient de corrélation entre les deux grandeurs.

Nous avons dans un premier temps moyenné les valeurs des distances perceptives $d_P(R_{i,ref}, R_{i,j})$ sur l'ensemble des 21 sujets. Une grande variabilité apparaît : l'écart-type moyen sur ces cent variables est supérieur à 20%. Cela étant, en effectuant une régression linéaire sur les données (distances $d_{\mathcal{F}}$ et distances moyennées \tilde{d}_P , normalisées entre 0 et 100 pour permettre l'analyse statistique, voir figure 3.2), on obtient un coefficient de corrélation de 0,779, significatif au seuil de 1 pour 1000.

On constate toutefois sur le graphe que si la tendance globale des données va dans le sens d'une pertinence perceptive de $d_{\mathcal{F}}$, la courbe de régression ne donne pas une relation linéaire : les distances perceptives sont élevées pour des distances théoriques nulles. Ceci peut en fait s'expliquer par le domaine de définition de $d_{\mathcal{F}}$, pour rappel $\mathcal{P}(\mathbb{Z}_c)/\sim_Z$. En effet, s'il est raisonnable de considérer des orbites par translation, comme on l'a déjà expliqué, il est plus discutable de considérer des classes d'équivalence de la Z -relation. Supposer qu'une faible valeur de la distance théorique correspond à une sensation de ressemblance des rythmes revient à supposer que la Z -relation sur les rythmes est perceptible (si $A \sim_Z B$, alors $d_{\mathcal{F}}(A, B) = 0$; A et B sont en fait le même élément du domaine de définition de $d_{\mathcal{F}}$). Nous avons mis en place une seconde expérience perceptive pour tester cette proposition.

FIGURE 3.2 – Relation entre d_P et $d_{\mathcal{F}}$

3.3 Perception de la Z-relation rythmique

3.3.1 Protocole d'écoute

Il existe sur \mathbb{Z}_{12} exactement 23 couples de rythmes (au sens de classes d'équivalence par transposition/inversion) en Z-relation. À chacun de ces couples (A, B) nous avons adjoint un troisième rythme C , de même cardinal que les deux autres. Il était alors demandé à l'auditeur, pour chacun de ces 23 triplets, de « trouver l'intrus », c'est-à-dire d'indiquer quel rythme présente, à l'écoute, le moins de ressemblance avec les autres. Si la relation Z entre les deux premiers rythmes A et B est réellement audible, alors le choix de l'auditeur se portera bien vers le rythme C . Là encore l'ordre des ensembles et l'attribution des labels aux rythmes dans les ensembles sont aléatoires. Le test a été effectué par 15 sujets.

3.3.2 Analyse des résultats

Pour chaque auditeur on a calculé le pourcentage de « bonnes réponses », c'est-à-dire le nombre de cas où l'intrus indiqué était bien le rythme C . En moyennant ce pourcentage sur l'ensemble des sujets, on voit que l'intrus est trouvé dans environ 55% des cas (le hasard valant ici un tiers).

Cette valeur, bien que relativement faible, semble indiquer que la Z-relation est perceptible. Si on regarde plus en détail les 23 paires de rythmes considérées, on constate que 15 d'entre elles sont constituées de rythmes complémentaires (*i.e.* avec les bons représentants des rythmes par transposition/inversion, on a $A \cup B = \mathbb{Z}_{12}$; il s'agit en fait de cas particuliers du théorème de l'hexacorde). En considérant d'un côté ces rythmes complémentaires, et de l'autre les 8 ensembles restants, et en calculant pour chacun les pourcentages de bonnes réponses, on obtient :

- pour les rythmes complémentaires, l'intrus est trouvé dans 66% des cas, donc significativement au dessus du hasard,
- pour les rythmes en Z-relation mais non complémentaires, l'intrus est trouvé dans 33% des cas, valeur du hasard : cela revient à dire que la Z-relation n'est pas perçue dans ce cas.

Ce constat laisse à penser que la relation de complémentarité entre les rythmes (figure d'ailleurs courante en musique, et impliquant la Z-relation dans le cas du théorème de l'hexacorde) est perceptible, mais que la Z-relation dans le cas général ne l'est pas. Il faudrait, pour compléter cette étude, reprendre ce protocole avec des métriques (valeurs de c) différentes et en considérant d'autres ensembles complémentaires que les hexacordes.

3.4 Bilan critique

Les dernières remarques faites sur la perception de la Z-relation tendent à infirmer l'existence d'une dimension perceptive de $d_{\mathcal{F}}$. En effet, puisque deux rythmes en Z-relation ne semblent pas présenter de similarité à l'écoute, le fait qu'elle soit définie sur les classes d'équivalence de cette relation est problématique. Cela explique en particulier les valeurs élevées de d_P à l'origine du graphe 3.2.

Si cette distance semble vouée à rester un outil d'analyse abstrait sans réelle pertinence perceptive, il serait néanmoins intéressant de comparer ses propriétés à celles des nombreuses autres mesures de similarité existantes, à la manière de [24]. Notons par ailleurs que de telles études, visant à confronter des théories musicales abstraites, mathématiques, à la réalité de la perception musicale, sont relativement rares dans la littérature (voir [12], [15]) ; on peut citer à ce sujet l'appel de Castrén [8] qui prône l'« investigation systématique » de telles corrélations.

Enfin, nous ne concluons pas de ces résultats qu'il n'existe aucun lien entre les coefficients de Fourier et la perception des rythmes. Notre définition $d_{\mathcal{F}}$ était motivée par les propriétés d'invariance du module, mais d'autres choix auraient bien sûr été envisageables. À ce sujet, on peut citer l'article [26] de Dmitri Tymoczko, qui met en évidence une forte corrélation entre les valeurs des modules de \mathcal{F} et des distances calculées en terme de *voice leading* (voir figure 3.3).

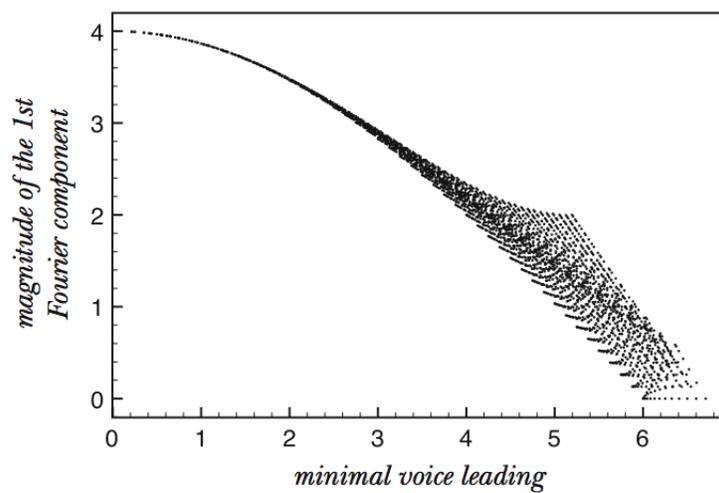


FIGURE 3.3 – Relation entre *voice leading* et module de \mathcal{F}

Conclusion

Le tour d’horizon des principaux résultats impliquant la transformée de Fourier discrète nous a permis d’apprécier la puissance théorique de cet outil pour l’étude de modèles mathématiques d’objets musicaux. En parallèle du travail mathématique, nous avons développé une librairie pour l’environnement OpenMusic, écrite en LISP, permettant d’utiliser les fonctions étudiées dans un environnement de composition assistée par ordinateur, et ainsi d’avoir à disposition à la fois les modèles mathématiques et les objets musicaux eux-mêmes. Si les résultats des tests perceptifs n’ont pas montré de bonnes corrélations entre théorie mathématique et perception, ils ont pu donner quelques éclaircissements, notamment concernant la Z-relation.

Questions ouvertes, perspectives

Les deux pistes concernant l’homométrie évoquées en fin de première partie mériteraient d’être approfondies, l’outil théorique bien adapté au problème que constitue la TFD et la littérature imposante sur le sujet laissent à penser que des résultats pourraient vite être obtenus. Ensuite, définir de nouveaux modèles d’objets musicaux apparaît essentiel, afin de pouvoir travailler d’une part sur des espaces continus, et d’autre part sur plusieurs dimensions.

Enfin, les tests perceptifs effectués mériteraient d’être menés à plus grande échelle, avec plus de sujets, et en prenant en compte un corpus de rythmes plus grand, afin de pouvoir confirmer les hypothèses que nous avons proposées.

Bibliographie

- [1] Emmanuel Amiot. Une preuve élégante du théorème de Babbitt par transformée de Fourier discrète. *Quadrature*, 61, 2006.
- [2] Emmanuel Amiot. Gammes bien réparties et transformée de Fourier discrète. *Mathématiques & Sciences humaines*, 178 :95–118, 2007.
- [3] Emmanuel Amiot. Discrete Fourier transform and good temperament. *Music Theory Online*, 15(2), June 2009.
- [4] Emmanuel Amiot, Thomas Noll, Moreno Andreatta, and Carlos Agon. Fourier oracles for computer-aided improvisation. In *Proceedings ICMC*, 2006.
- [5] Milton Babbitt. Some aspects of twelve-tone composition. *Score*, 12 :53–61, 1955.
- [6] Abraham Bookstein, Vladimir A. Kulyukin, and Timo Raita. Generalized Hamming distance. *Information Retrieval*, 5(4) :353–375, 2002.
- [7] Clifton Callender, Ian Quinn, and Dmitri Tymoczko. Generalized voice-leading spaces. *Science*, 320(5874) :346–348, 2008.
- [8] Marcus Castrén. Pairs of chords as objects illuminating set-class similarity : Some viewpoints and a computer-assisted procedure to create test materials for listener experiments,«. *Electronic Journal of Music Theory and Analysis*, 1(1), 2000.
- [9] John Clough and Jack Douthett. Maximally even sets. *Journal of Music Theory*, 35(1/2) :93–173, 1991.
- [10] Jack Douthett and Richard Krantz. Maximally even sets and configurations : common threads in mathematics, physics, and music. *Journal of Combinatorial Optimization*, 14(4) :385–410, 2007.
- [11] Allen Forte. *The structure of atonal music*. Yale University Press, 1973.
- [12] Ludger Hofmann-Engl. Rhythmic similarity : A theoretical and empirical approach. In *Proceedings of the Seventh International Conference on Music Perception and Cognition*, 2002.
- [13] Mihail N. Kolountzakis. *The study of translational tiling with Fourier Analysis*, pages 131–187. Springer, 2004.
- [14] Mihail N. Kolountzakis and Máté Matolcsi. Algorithms for translational tiling. *Journal of Mathematics and Music*, 3(2) :85–97, 2009.
- [15] Tuire Kuusi. *Set-class and chord : Examining connection between theoretical resemblance and perceived closeness*, volume 12. Citeseer, 2001.

- [16] Guillaume Lachaussée. Théorie des ensembles homométriques. Master's thesis, Ircam/École Polytechnique, 2010.
- [17] David Lewin. Re : Intervallic relations between two collections of notes. *Journal of Music Theory*, 3(2) :298–301, 1959.
- [18] David Lewin. Special cases of the interval function between pitch-class sets x and y . *Journal of Music Theory*, 45(1) :1–29, 2001.
- [19] David Lewin. *Generalized musical intervals and transformations*. Oxford University Press, USA, 2007.
- [20] John Mandereau. Étude des ensembles homométriques et leur application en théorie mathématique de la musique et en composition assistée par ordinateur. Master's thesis, UPMC/Ircam, 2009.
- [21] John Mandereau, Daniele Ghisi, Emmanuel Amiot, Moreno Andreatta, and Carlos Agon. Phase retrieval in musical structures. *Journal of Mathematics and Music*, 2011 (à paraître).
- [22] Andrew J. Milne, Martin Carlé, William A. Sethares, Thomas Noll, and Simon Holland. Scratching the scale labyrinth. *Mathematics and Computation in Music*, pages 180–195, 2011.
- [23] Ian Quinn. *A unified theory of chord quality in equal temperaments*. PhD thesis, University of Rochester, 2004.
- [24] Godfried Toussaint. A comparison of rhythmic similarity measures. In *Proceedings 5th International Conference on Music Information Retrieval*, pages 242–245. Cite-seer, 2004.
- [25] Godfried Toussaint. The geometry of musical rhythm. In *Discrete and Computational Geometry*, pages 198–212. Springer, 2005.
- [26] Dmitri Tymoczko. Set-class similarity, voice leading, and the Fourier transform. *Journal of Music Theory*, 52(2) :251–272, 2008.
- [27] Dan T. Vuza. Supplementary sets and regular complementary unending canons. *Perspectives of New Music*, pages 270–305, 1993.