MMIM Modèles mathématiques
pour l'informatique musicale

Marc Chemillier

Master 2 Atiam (Ircam), 2010-2011

Notions théoriques en combinatoire des mots

- Conjugaison et mots de Lyndon
  - Conjugaison
  - Mots de Lyndon
  - Familles circulaires de factorisation maximale unique
- Combinatoire des mots appliquée à la musique
  - Rythmes asymétriques
  - Chaînes euclidiennes
  - Gammes maximalement réparties
  - Mots de Christoffel et mode lydien
- Construction des rythmes asymétriques

1. Conjugaison et mots de Lyndon

1.1 Conjugaison

On introduit une permutation δ de Σ^* en posant :

$\delta(au) = ua$ pour $u \in \Sigma^*$, $a \in \Sigma$,

$\delta(\varepsilon) = \varepsilon$.

Cela revient à faire passer la première lettre à la fin du mot.

La *classe de conjugaison* de *u* est l'orbite de *u* sous l'action de δ. C'est l'ensemble des permutations circulaires de *u* (ou rotations de *u*).

Exemple : pour le mot *abaa*

$\delta(abaa) = baaa$

$\delta^2(abaa) = aaab$

$\delta^3(abaa) = aaba$

$\delta^4(abaa) = abaa$

Notons que si n est la longueur de u, on a toujours $\delta^n(u) = u$. L'ensemble des puissances δ^k telles que $\delta^k(u) = u$ est un sous-groupe du groupe des permutations de Σ^* engendré par δ et il n'est pas réduit à l'identité. On l'appelle le fixateur de u.

On a un morphisme de l'ensemble des entiers relatifs **Z** (groupe additif) dans le groupe des puissances de δ : $i \rightarrow \delta^i$. Donc l'image réciproque du fixateur de u est un sous-groupe de **Z.** Or tous les sous-groupes de **Z** sont engendrés par un unique élément (monogènes), c'est-à-dire de la forme $d\mathbf{Z}$ avec $d > 0$.

Donc le fixateur de u est engendré par une unique puissance δ^d. De plus, pour tout entier k vérifiant $\delta^k(u) = u$, on a nécessairement k multiple de d.

Rappelons qu'un mot est primitif s'il n'est pas puissance d'un mot plus court.

Proposition. *Un mot primitif u de longueur n a exactement n permutations circulaires distinctes.*

Dem. Supposons que deux permutations circulaires soient égales, c'est-à-dire $\delta^i(u) = \delta^j(u)$ pour $0 \le i < j < n$.

Alors pour $k = j - i < n$, on a $\delta^k(u) = u$. Donc δ^k appartient au fixateur de u. Soit δ^d l'unique générateur du fixateur de u. Alors $d < n$ (car $k < n$). Par ailleurs, n est la longueur de u, donc multiple de d.

On factorise u en mots strictement plus courts de longueur d :

$u = x_1x_2\ldots x_q$

Alors $\delta^d(u) = x_2\ldots x_qx_1 = u = x_1x_2\ldots x_q$, donc $x_1 = x_2\ldots = x_q$ que l'on note x. D'où $u = x^q$, donc u n'est pas primitif.

Corollaire fondamental ([Lothaire 1983, p. 7, proposition 1.3.2]). *Si deux mots u et v commutent $uv = vu$, alors ils sont puissances d'un même mot $u, v \in x^*$.*

Dem. Soient n et $k < n$ les longueurs respectives de uv et de u.

On a $\delta^k(uv) = vu = uv$, donc δ^k appartient au fixateur de uv. Soit δ^d l'unique générateur du fixateur de uv. Alors k et n sont multiples de d, donc $n - k$ est aussi multiple de d.

Comme le mot uv n'a pas n permutations circulaires distinctes, la proposition précédente montre que uv est puissance d'un mot x de longueur d.

Il en résulte $u = x^{k/d}$ et $v = x^{(n-k)/d}$.

1.2 Mots de Lyndon

Un *mot de Lyndon* est un mot

• primitif,

• minimal pour l'ordre alphabétique dans sa classe de conjugaison.

Exemples : *abaa* est-il un mot de Lyndon ?

-> non, le mot de Lyndon associé est *aaab*.

Quel est le mot de Lyndon associé à *xyxxy* ?

-> *xxyxy*

Pour étudier des structures musicales périodiques, les mots de Lyndon fournissent un représentant unique pour chaque classe de conjugaison.

Il existe un algorithme très efficace pour calculer les mots de Lyndon (algorithme de Duval). Cet algorithme repose sur les propriétés suivantes :

Propriété ([Lothaire 1983, p. 66, proposition 5.1.3]). *w mot de Lyndon si et seulement si soit w est une lettre, soit w est un produit de mots de Lyndon plus courts :*

$w = uv$ avec u, v mots de Lyndon tels que $u < v$.

Théorème ([Lothaire 1983, p. 67, théorème 5.1.5]). *Tout mot non vide w admet une factorisation unique en mots de Lyndon $w = u_1u_2\ldots u_n$ tels que $u_1 \geq u_2 \geq \ldots \geq u_n$.*

Exemples :

• toutes les factorisations de *ab* en mots de Lyndon ?

(*a*)(*b*)

(*ab*)

Laquelle est décroissante ?

-> (*ab*) car c'est un mot de Lyndon

• toutes les factorisations de *abaac* en mots de Lyndon ?

(*a*)(*b*)(*a*)(*a*)(*c*)

(*ab*)(*a*)(*a*)(*c*)

(*a*)(*b*)(*a*)(*ac*)

(*ab*)(*a*)(*ac*)

(*a*)(*b*)(*aac*)

(*ab*)(*aac*)

Laquelle est décroissante ?

-> (*ab*)(*aac*)

On note que *abaac* n'est pas un mot de Lyndon. Par contre, *aacab* est un mot de Lyndon et on a bien *aac* < *ab*.

• factorisation décroissante de *aababbaaba* ?

-> (*aababb*)(*aab*)(*a*)

Problème : Pour définir un représentant unique par classe de conjugaison, il peut paraître étrange de devoir passer par une relation d'ordre arbitraire (ordre alphabétique) comme on le fait pour les mots de Lyndon. En fait on montre que toute famille de représentants uniques vérifiant les propriétés de factorisation ci-dessus <u>induit un ordre intrinsèque sur les mots indépendant de l'ordre alphabétique</u>.

1.3 Familles circulaires de factorisation maximale unique (circ-UMFF)

W ensemble de mots de Σ^*

W <u>famille de factorisation</u> (FF) si tout mot $w \in \Sigma^*$ admet une factorisation en mots de W

Une factorisation <u>maximale</u> : tous les facteurs de la décomposition de w sont maximaux, c'est-à-dire qu'il n'existe pas de facteur de w

- qui contienne un élément de sa décomposition, et
- qui soit aussi dans W.

Remarques :

• Si un mot admet une factorisation <u>maximale</u>, celle-ci est <u>unique</u> (s'il y en avait deux, on prendrait le premier facteur qui les distingue et il serait nécessairement plus grand dans l'une que dans l'autre, donc pas maximal)

• Une famille de factorisation FF <u>contient</u> l'alphabet Σ (car les lettres doivent avoir une factorisation)

• Inversement, si W contient l'alphabet, alors W est une famille de factorisation FF

W <u>famille de factorisation maximale unique</u> (UMFF) si toutes les factorisations sont maximales

Lemme *xyz* [Daykin, Daykin & Smyth 2009]. *Une FF notée W est une UMFF si et seulement si* $xy, yz \in W$ *avec* y *non vide impliquent* $xyz \in W$.

Dem. C.N. On montre que si W est une UMFF, W vérifie le lemme. Soient xy et $yz \in W$, on considère la factorisation maximale de $w = xyz$. Son premier facteur w' contient nécessairement xy (sinon il ne serait pas maximal), donc $|w'| \geq |xy|$. De même, son dernier facteur w'' contient yz, donc $|w''| \geq |yz|$. Si ces deux facteurs étaient distincts, on aurait $|w| \geq |w'| + |w''| \geq |x| + 2|y| + |z| > |w|$ car y non vide, ce qui n'est pas possible. Donc le premier facteur est aussi le dernier, donc la factorisation maximale est réduite à un seul facteur égal à w tout entier, donc $w \in W$.

C.S. On montre que si W est une FF qui vérifie le lemme, alors W est une UMFF. Soit $w \in W$, on construit de proche en proche la factorisation maximale de w en prenant le plus long préfixe w_1 de w qui appartient à W, puis le plus long préfixe w_2 du suffixe restant qui appartient à W, etc. Par définition, il n'est pas possible d'augmenter les facteurs obtenus sur la droite en restant dans W, car on les a construits en allant le plus loin possible à droite. On montre qu'il n'est pas possible de les augmenter sur la gauche en restant dans W. Supposons par exemple que yw_iy' soit dans W avec $0 < |y| < |w_{i-1}|$. On peut écrire $w_{i-1} = xy$ avec y non vide. Comme $xy \in W$, $yw_iy' \in W$, le lemme xyz implique $xyw_iy' = w_{i-1}w_iy' \in W$, ce qui est faux car w_{i-1} a été choisi en allant le plus loin à droite possible dans W.

Une <u>famille circulaire de factorisation maximale unique</u> (circ-UMFF) : contient un représentant pour chaque classe de conjugaison de mots primitifs non vides

On définit une relation sur W : $u <_W v$ si et seulement $uv \in W$

Si W est une circ-UMFF :

• on ne peut avoir à la fois $u <_W v$ et $v <_W u$, car uv et vu sont conjugués.

• on montre que l'on a nécessairement l'une des deux relations $u <_W v$ ou $v <_W u$ (Théorème 2.10, [Daykin, Daykin & Smyth 2009]).

• <u>la relation $<_W$ est transitive</u> : Soient $u <_W v$ et $v <_W w$, on montre que l'on ne peut pas avoir $w <_W u$. Si c'était le cas, on aurait uv, vw, $wu \in W$, donc le lemme xyz impliquerait uvw et $wuv \in W$, c'est-à-dire deux permutations distinctes du même mot, ce qui est faux. Donc si l'on n'a pas $w <_W u$, on a nécessairement $u <_W w$, donc la relation $<_W$ est transitive.

La relation $<_W$ est <u>une relation d'ordre strict</u> que l'on peut prolonger en <u>relation d'ordre $\leq_W$ sur W</u>.

Proposition (Lemme 2.12, [Daykin, Daykin & Smyth 2009]). *Un mot $w \in W$ circ-UMFF si et seulement si soit w est une lettre, soit w est un produit de mots plus courts :*

$w = uv$ avec $u, v \in W$ et $u <_W v$.

Théorème (Lemme 2.13, [Daykin, Daykin & Smyth 2009]). *Tout mot non vide w admet une factorisation unique sur W circ-UMFF telle que $w = u_1u_2\ldots u_n$ avec $u_1 \geq_W u_2 \geq_W \ldots \geq_W u_n$.*

On voit qu'une circ-UMFF définit <u>sa propre relation d'ordre</u>, et que les propriétés de décomposition s'expriment avec cette relation d'ordre. Dans le cas où W est définie par les éléments minimaux d'une relation d'ordre arbitraire (mots de Lyndon), celle-ci n'est pas nécessairement identique à l'ordre intrinsèque $<_W$ de W qui intervient dans les propriétés de décomposition.

Exemple : on peut définir l'ensemble W des mots de Lyndon sur <u>l'ordre alphabétique rétrograde $<_\sim$</u>. Si l'on note $x\sim$ le mot rétrograde de x (lu à partir de la fin), on pose $x <_\sim y$ si et seulement $x\sim <_{\text{alph}} y\sim$. On a :

• $abbc <_{alph} abc$, donc en prenant les rétrogrades : $cbba <_{\sim} cba$

• il est facile de voir que *cba, cbba, cbacbba* sont minimaux pour l'ordre $<_{\sim}$ dans leur classe de conjugaison, donc : $cba, cbba, cbacbba \in W$

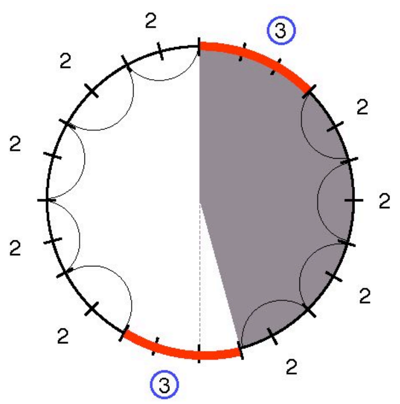
• dès lors, on a $(cba)(cbba) = cbacbba$, donc $cba <_W cbba$

pourtant on a : $cba >_{\sim} cbba$

Dans ce cas, l'ordre $<_{\sim}$ définissant W comme <u>ensemble d'éléments minimaux</u> ne coïncide pas avec l'ordre $<_W$ <u>intrinsèque à sa structure de circ-UMFF</u>.

2. Combinatoire des mots appliquée à la musique

2.1 Rythmes asymétriques vérifiant « l'imparité rythmique »



On considère l'alphabet $\Sigma = \{2, 3\}$ et on étudie les rythmes qui sont des mots de Σ^*. Le <u>poids</u> $h(w)$ d'un mot w est la somme de ses chiffres.

Définition. *Un rythme w vérifie <u>l'imparité rythmique</u> si aucune de ses permutations circulaires ne se factorise en deux mots de même somme, c'est-à-dire de la forme uv avec* $h(u) = h(v)$.

Exemple : $w = 32322$

(32)(322)	$h(32)=5$, $h(322)=7$
(23)(223)	$h(23)=5$, $h(223)=7$
(32)(232)	$h(32)=5$, $h(232)=7$
(223)(23)	$h(223)=7$, $h(23)=5$
(232)(32)	etc.

Il existe un lien très curieux entre l'asymétrie de ces rythmes et l'asymétrie de la gamme diatonique (touches noires du clavier séparées par 2, 3, 2, 2, 3 demi-tons), qui a été mis en évidence pour la première fois, semble-t-il, par Jeff Pressing 1983.

De nombreux travaux ont étudié ces structures asymétriques en introduisant des concepts de combinatoire des mots.

2.2 Chaînes euclidiennes

L'alphabet Σ est l'ensemble des entiers naturels. Comme précédemment, le poids $h(w)$ d'un mot w est la somme de ses chiffres. Notons que les entiers apparaissant dans un mot ne sont pas seulement 2 ou 3 comme pour les rythmes précédents, mais peuvent prendre n'importe quelle valeur.

Définition. *Un mot $w = w_0w_1\ldots w_{n-1}$ est une chaîne euclidienne si w est réduit à une seule lettre (c'est-à-dire un entier), ou si en ajoutant 1 à la première w_0 et en retirant 1 à la dernière w_{n-1}, on obtient un mot $t(w)$ qui est une rotation $\delta^k(w)$ de w.*

Exemple : $w = 22323$,

$t(w) = (2 + 1)232(3 - 1) = 32322 = \delta^2(22323)$

$z = 001010010101$,

$t(z) = (0 + 1)0101001010(1 - 1) = 101010010100 = \delta^7(001010010101)$

Quel rapport entre les mots 22323 et 001010010101 ?

-> Si l'on représente des rythmes en notant 1 pour une frappe et 0 pour un silence, ils apparaissent comme duals l'un de l'autre, c'est-à-dire qu'ils correspondent à deux manières de coder un même rythme (à une rotation près).

Godfried Toussaint s'est intéressé à tous les rythmes *aksak* et autres qui sont des chaînes euclidiennes (2005).

On introduit deux morphismes sur Σ^* définis pour chaque « lettre » de Σ (c'est-à-dire chaque entier x) :

$I(x) = x + 1$,

$X(x) = 01^x$ (c'est-à-dire le mot commençant par 0 suivi de 1 répété x fois).

Exemple : $I(024) = 135$, $X(024) = 001101111$

On montre qu'à une rotation près, il n'existe qu'une seule chaîne euclidienne de longueur n et de poids k donnés notée $E(n, k)$ définie par [Ellis *et al*. 2003] :

$E(n, k) = XE(n - k, k)$ si $n > k$,

$= IE(n, k - n)$ si $n < k$,

$= 1$ si $n = k$,

Exemple :

$E(13, 5) = XE(13 - 5, 5) = XE(8, 5) = XXE(3, 5) = XXIE(3, 2) = XXIXE(1, 2) = XXIXIE(1, 1) = XXIXI(1) = XXIX(2) = XXI(011) = XX(122) = X(01011011) = 0010010100101$.

On obtient un mot de Lyndon w de longueur 13 et de poids 5 tel que :

$t(w) = 1010010100100 = \delta^5(w)$

$E(5, 13) = IE(5, 8) = IIE(5, 3) = IIXE(2, 3) = IIXIE(2, 1) = IIXIXE(1, 1) = IIXIX(1) = IIXI(01) = IIX(12) = II(01011) = I(12122) = 23233$.

Ce mot de longueur 5 et de poids 13 est en quelque sorte le « dual » du précédent.

Il se trouve que tous les rythmes africains vérifiant l'imparité rythmique avec deux groupes de 3 sont des chaînes euclidiennes :

$E(3, 8) = 233$

$E(5, 12) = 22323$

$E(7, 16) = 2223223$

$E(9, 20) = 222232223$

$E(11, 24) = 22222322223$.

On notera qu'ils sont présentés sous forme de mots de Lyndon.

C'est vrai dans le cas général de deux groupes de 3 où l'on a toujours une chaîne euclidienne :

$t(2^n32^{n-1}3) = (2{+}1)2^{n-1}32^{n-1}(3{-}1) = 32^{n-1}32^n = \delta^n(2^n32^{n-1}3)$

Mais ce n'est plus toujours vrai si le nombre de 3 est supérieur à deux.

2.3 Gammes maximalement réparties

Le concept de « gamme maximalement répartie » (*maximally even scale*) a été défini par Clough & Douthett 1991. Il est lié à d'autres concepts comme celui de « gamme bien formée » (*well-formed scale*) de Carey & Clampitt 1989 et conduit à introduire les séries de Fourrier. On en donne ici une définition en termes de combinatoire des mots.

L'alphabet Σ est, comme précédemment, l'ensemble des entiers naturels. On représente une gamme par un mot w donnant les intervalles entre ses degrés consécutifs exprimés en nombre de « demi-tons » du total chromatique (pour la dernière note, on prend l'intervalle avec l'octave supérieur de la première note). On calcule ensuite les « mots cumulés » obtenus en additionnant deux à deux, puis trois à trois, etc. les symboles de w considéré circulairement.

Exemple : Fonction Lisp qui affiche les dix premiers « mots cumulés » pour la gamme diatonique (touches blanches) définie par ses intervalles en nombre de demi-tons

```
(cumulate '(2 2 1 2 2 2 1) 10)
k=1 (2 2 1 2 2 2 1)
k=2 (4 3 3 4 4 3 3)
k=3 (5 5 5 6 5 5 5)
k=4 (7 7 7 7 7 7 6)
k=5 (9 9 8 9 9 8 8)
k=6 (11 10 10 11 10 10 10)
k=7 (12 12 12 12 12 12 12)
k=8 (14 14 13 14 14 14 13)
k=9 (16 15 15 16 16 15 13)
k=10 (17 17 17 18 17 15 13)
```

Remarque :

- On voit facilement que pour $k = 7$ (longueur de w), le mot cumulé ne contient que le nombre 12 = total chromatique.
- Pour $k > 7$, on retrouve les mêmes valeurs que pour $k \leq 7$, mais modulo 12.

Pour simplifier la notation, on convient que les indices sont comptés circulairement, c'est-à-dire qu'un ensemble d'indices de 0 à n est identifié à $\mathbf{Z}/n\mathbf{Z}$ ce qui permet d'écrire x_{i+j} à la place de $x_{i+j \bmod n}$.

Les « mots cumulés » $w(x)$ sont les mots de même longueur d que w définis pour $x \geq 1$ et pour $i \in \mathbf{Z}/d\mathbf{Z}$ par :

$w(x)_i = w_i + \ldots + w_{i+x-1}$.

Définition. *Une gamme est* <u>*maximalement répartie*</u> *si tous les « mots cumulés » w(x) contiennent au plus* <u>*deux entiers différents et consécutifs*</u>.

Remarque :

• Si $d = |w|$ est la longueur de w, il suffit de vérifier la propriété pour les $d - 1$ premiers mots cumulés $w(x)$ avec x de 1 à $d - 1$.

• Que se passe-t-il si le mot w ne contient qu'un seul entier (ex : gamme par tons) ?

-> Il en est de même de tous les mots cumulés $w(x)$.

Les modes correspondant à une <u>division égale de l'octave</u> sont des cas particuliers de gammes maximalement réparties. Par exemple, la « gamme par tons » correspond à w = 222222.

Exemples : La gamme diatonique (touches blanches du piano) est maximalement répartie :

```
(maximally-even '(2 2 1 2 2 2 1))
k=1 (2 2 1 2 2 2 1)
k=2 (4 3 3 4 4 3 3)
k=3 (5 5 5 6 5 5 5)
k=4 (7 7 7 7 7 7 6)
k=5 (9 9 8 9 9 8 8)
k=6 (11 10 10 11 10 10 10)
```

Pour les touches noires du piano, c'est-à-dire la « gamme pentatonique », on a aussi :

```
(maximally-even '(3 2 3 2 2))
k=1 (3 2 3 2 2)
k=2 (5 5 5 4 5)
k=3 (8 7 7 7 7)
k=4 (10 9 10 9 10)
```

Certains rythmes asymétriques d'Afrique centrale apparaissent comme des mots maximalement répartis :

```
(maximally-even '(3 2 2 3 2 2 2))
k=1 (3 2 2 3 2 2 2)
k=2 (5 4 5 5 4 4 5)
k=3 (7 7 7 7 6 7 7)
k=4 (10 9 9 9 9 9 9)
k=5 (12 11 11 12 11 11 12)
k=6 (14 13 14 14 13 14 14)
```

Mais ce n'est pas le cas de l'un des rythmes pygmées vérifiant l'imparité rythmique appelé *mokongo* qui n'est pas maximalement réparti :

```
(maximally-even '(3 3 3 2 3 3 2 3 2))
k=1 (3 3 3 2 3 3 2 3 2)
k=2 (6 6 5 5 6 5 5 5 5)
k=3 (9 8 8 8 8 8 7 8 8)
NOT MAXIMALLY EVEN
```

Proposition ([Demaine *et al.* 2005]). *Toute chaîne euclidienne est maximalement répartie.*

On en déduit, en particulier, que le rythme *mokongo* n'est pas une chaîne euclidienne.

Qu'en est-il de la chaîne euclidienne de même longueur et de même poids ?

-> Le problème est que la longueur 9 et le poids 24 ne sont pas premiers entre eux ($9 = 3 \times 3$ et $24 = 3 \times 8$). Pour ces valeurs, la chaîne euclidienne est :

$E(9, 24) = E(3, 8) = 233$.

D'autres mots rythmiquement impairs ne sont pas maximalement répartis :

```
(maximally-even '(3 3 3 2 3 2 3 3 3 2 2 3 2))
k=1 (3 3 3 2 3 2 3 3 3 2 2 3 2)
k=2 (6 6 5 5 5 5 6 6 5 4 5 5 5)
NOT MAXIMALLY EVEN
```

En revanche, d'autres le sont :

```
(maximally-even '(3 3 2 3 3 2 3 2 3 3 2 3 2))
k=1 (3 3 2 3 3 2 3 2 3 3 2 3 2)
k=2 (6 5 5 6 5 5 5 5 6 5 5 5 5)
k=3 (8 8 8 8 8 7 8 8 8 8 7 8 8)
k=4 (11 11 10 11 10 10 11 10 11 10 10 11 10)
k=5 (14 13 13 13 13 13 13 13 13 13 13 13 13)
k=6 (16 16 15 16 16 15 16 15 16 16 15 16 16)
k=7 (19 18 18 19 18 18 18 18 19 18 18 19 18)
k=8 (21 21 21 21 21 20 21 21 21 21 21 21 21)
k=9 (24 24 23 24 23 23 24 23 24 24 23 24 23)
k=10 (27 26 26 26 26 26 26 26 27 26 26 26 26)
k=11 (29 29 28 29 29 28 29 29 29 29 28 29 29)
k=12 (32 31 31 32 31 31 32 31 32 31 31 32 31)
```

Le dernier mot est la chaîne euclidienne $E(13,34)$, mais le précédent n'en est pas une, car il a même longueur et même poids, or les chaînes euclidiennes sont uniques pour une longueur et un poids donnés.

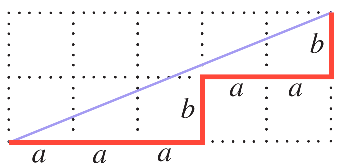
Si l'on récapitule :

- Les chaînes euclidiennes sont **uniques** pour une longueur et un poids donnés,
- Les chaînes euclidiennes sont maximalement réparties,
- Les mots rythmiquement impairs sont uniques pour une longueur donnée avec deux groupes de 3 (type $2^n32^{n-1}3$), mais ce n'est pas vrai sinon,
- Les mots rythmiquement impairs de type $2^n32^{n-1}3$ sont des chaînes euclidiennes
- D'autres mots rythmiquement impairs sont des chaînes euclidiennes, donc sont maximalement répartis,
- Certains mots rythmiquement impairs ne sont ni des chaînes euclidiennes, ni maximalement répartis.

Partie facultative (non traitée en cours) :

2.4 Mots de Christoffel et mode lydien

Un autre concept de combinatoire des mots intervient dans l'étude des gammes (travaux de [Noll 2008]) : c'est celui de mots de Christoffel.



Dans un plan quadrillé, on note a et b les segments unitaires horizontaux et verticaux.

Définition. *Le mot de Christoffel de pente p/q est la suite de a et b définissant le chemin reliant les points (0, 0) et (p, q) le long du quadrillage en restant le plus haut possible sous la droite passant par ces points.*

Considérons le « mode lydien », c'est-à-dire la suite des touches blanches du piano commençant sur *fa* :

fa sol la si do ré mi fa

Thomas Noll 2008 a remarqué qu'en notant a et b les intervalles de ton et demi-ton, on obtient :

$a \quad a \quad a \quad b \quad a \quad a \quad b$

c'est-à-dire exactement le mot de Christoffel de pente 2/5 représenté sur la figure ci-dessus.

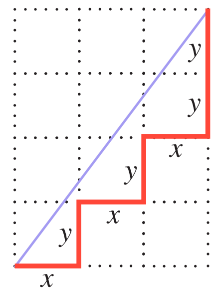
Bien plus, si l'on permute le mode précédent en répartissant les notes selon le cycle des quintes ramenées à l'intérieur de l'octave, on trouve :

fa do sol ré la mi si fa

En notant x et y les intervalles de quinte ascendante et quarte descendante, on a :

$x \quad y \quad x \quad y \quad x \quad y \quad y$

ce qui donne un nouveau mot de Christoffel de pente 4/3 :



Or Thomas Noll a mis en évidence le fait remarquable que ces deux mots de Christoffel sont « duals » l'un de l'autre.

Définition. *Le dual d'un mot de Christoffel w de pente p/q est le mot de Christoffel w^* de pente p^*/q^* où p^* et q^* sont les inverses de p et q dans $\mathbf{Z}/(p+q)\mathbf{Z}$.*

En effet, on a $p = 2$, $q = 5$ et $p + q = 7$:

L'inverse de 2 est 4 dans $\mathbf{Z}/7\mathbf{Z}$, car $2 \times 4 = 8 \approx 1$ mod 7, donc $p^* = 4$.

De meme, l'inverse de 5 est 3, car $5 \times 3 = 15 \approx 1$ mod 7, donc $q^* = 3$.

Les mots de Christoffel ont des propriétés surprenantes liées à la notion de palindrome, c'est-à-dire de mot identique à son image miroir (lu à l'envers) :

Pour le mode lydien $w = aaabaab$, on a une double décomposition en palindromes :

- $w = aub$ où u est un palindrome, soit $w = a(aabaa)b$,
- et $w = a(rabs)b$ où r et s sont deux palindromes, soit $w = a(a)ab(aa)b$

Pour son mot dual $w^* = xyxyxyy$, on a les mêmes décompositions :

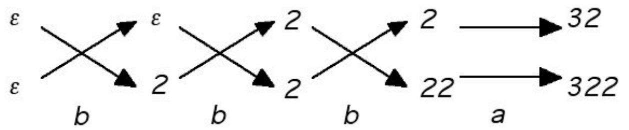
- $w^* = xuy$ où u est un palindrome, soit $w^* = x(yxyxy)y$,
- et $w^* = x(rxys)y$ où r et s sont deux palindromes, soit $w^* = x(yxy)xy()y$.

3. Construction des rythmes asymétriques

3.1 Rythmes vérifiant l'imparité rythmique

Soient deux transformations a et b de $\Sigma^* \times \Sigma^*$ dans $\Sigma^* \times \Sigma^*$ définies par :

$$a(u, v) = (3u, 3v), \qquad b(u, v) = (v, 2u).$$



Le rythme 32322 s'écrit (32)(322) avec (32, 322) = $abbb(\varepsilon, \varepsilon)$.

Théorème ([Chemillier & Truchet 2003]). *Un mot w vérifie l'imparité rythmique si et seulement s'il existe une transformation α obtenue en combinant a et b avec un nombre impair de b telle que $\alpha(\varepsilon, \varepsilon) = (u, v)$ avec $w = uv$ ou vu.*

Soit $D = \{w$ tels que $w = uv$ avec $(u, v) = \alpha(\varepsilon,\varepsilon)$ et α ayant un nombre impair de $b\}$.

Proposition. *Pour w et w' dans D, on a :*

w et w' conjugués si et seulement si α et α' conjugués dans $\{a, b\}^*$.

Pour calculer un représentant par classe de conjugaison de mots w vérifiant l'imparité rythmique, il suffit de calculer les mots de Lyndon pour les mots α sur $\{a, b\}^*$.

On note n_2 et n_3 les nombres de 2 et de 3 dans un rythme vérifiant l'imparité rythmique :

Proposition. *Le nombre de solutions à l'imparité rythmique pour $n_3 = 8$ est la somme des carrés.*

Dem. Les solutions peuvent être associées à des suites α de a et b d'après le théorème précédent, avec quatre a et un nombre impair de b. Soit p la longueur de α. Énumérer les solutions consiste à énumérer les α possibles en plaçant 4 symboles a parmi p dans α, mais en enlevant les permutations circulaires.

Combien y a-t-il de permutations circulaires pour α ?

Supposons α non primitif : $\alpha = \gamma^n$, avec $n \geq 2$. Le nombre de a dans α est n fois le nombre de a dans γ. Mais 4 est divisible seulement par 2 et 4. Dans ces conditions, le nombre de b dans α ne peut pas être impair, ce qui est faux.

D'où α primitif, donc il a exactement p permutations circulaires.

Le nombre de solutions est

$$x = \frac{1}{p} C_p^4 = \frac{(p-3)(p-2)(p-1)}{4!}$$

Comme n_2 est impair, on peut écrire $n_2 = 2k - 1$, donc $p = n_2 + 4 = 2k + 3$, d'où :

$$x = \frac{k(k+1)(2k+1)}{2}$$

Références

• <u>ouvrages fondamentaux</u>

Lothaire M., *Combinatorics on Words*, Encyclopedia of Mathematics, Vol. 17, Addison-Wesley, 1983 (réédité Cambridge University Press, 1997, chap. 1 disponible en ps).

Lothaire M., *Algebraic Combinatorics on Words*, Cambridge University Press, 2002.

Lothaire M., *Applied Combinatorics on Words*, Cambridge University Press, 2005.

Daykin D.E., Daykin J., Smyth W.F., Combinatorics of Unique Maximal Factorization Families (UMFFs), *Fondamenta Informaticae, Special Issue on Stringology*, 97 (3) (2009) 295-309.

Berstel J., Lauve A., Reutenauer C., Saliola F., *Combinatorics on Words : Christoffel Words and Repetitions in Words*, CRM Monograph Series (Centre de Recherches Mathématiques Montréal) , vol. 27, American Mathematical Society, 2009.

• <u>combinatoire des mots appliquée à la musique</u>

Chemillier M., Periodic musical sequences and Lyndon words, *Soft Computing*, special issue on Formal Systems and Music, G. Assayag, V. Cafagna, M. Chemillier (eds.) **8** (9) (2004) 611-616.

Chemillier M., Truchet C., Computation of words satisfying the "rhythmic oddity property" (after Simha Arom's work), *Information Processing Letters* **86** (5) (2003) 255-261.

Chemillier M., Musique et rythme en Afrique centrale, *Pour la science*, dossier n° 47 « Mathématiques exotiques » (2005) 72-76.

Chemillier M., *Les Mathématiques naturelles*, Paris, Odile Jacob, 2007.

Toussaint Godfried T., The Euclidean algorithm generates traditional musical rhythms, *Proc. of BRIDGES: Mathematical Connections in Art, Music and Science*, Banff, Canada, July 31 - August 3 2005.

Ellis John, Frank Ruskey, Joe Sawada, and Jamie Simpson, Euclidean strings, *Theoretical Computer Science*, 301 (2003) 321-340.

Clough John, Jack Douthett, Maximally even sets, *Journal of Music Theory*, 35 (1991) 93-173.

N. Carey, D. Clampitt, Aspects of Well-Formed Scales, *Music Theory Spectrum* 11(2) (1989) 187–206.

Pressing Jeff, Cognitive Isomorphisms Between Pitch and Rhythm in World Musics: West Africa, the Balkans and Western Tonality, *Studies in Music* (Australia) 17 (1983) 38–61.

Noll Thomas, Sturmian sequences and morphisms: a music-theoretical application, *Journée annuelle de la SMF* (Société Math. de France), 2008, p. 79-102

Demaine Erik D., Francisco Gomez-Martin, Henk Meijer, David Rappaport, Perouz Taslakian, Godfried T. Toussaint, Terry Winograd, David R. Wood, The Distance Geometry of Music, *20th Bellairs Winter Workshop on Computational Geometry*, Barbados, January 28-February 4, 2005.