

À PROPOS D'UNE CONJECTURE SUR LES CODES CONJUGUÉS

OMAR KHADIR

RÉSUMÉ. L'objectif de ce travail est d'exposer et de retracer les étapes qui ont mené à la résolution d'une conjecture sur les codes conjugués. Il s'agit de la caractérisation des codes qui commutent, et plus généralement la caractérisation de tous les ensembles ayant cette propriété.

ABSTRACT. The aim of this paper is to set out and retrace the steps that lead to the resolution of an old conjecture on commuting codes. Furthermore, simple proofs are given for some particular cases. This work is related to the characterization of commuting codes, and more generally to the characterization of all sets having this property.

1. Introduction. Les codes sont au centre de la théorie de l'information. Ils ont pris une importance considérable avec les besoins, sans cesse croissants, de communication, de stockage de données, de cryptographie, de compression des images... La question traitée dans cet article concerne une conjecture issue du domaine de l'informatique théorique, publiée en 1989 et résolue en 2002. On définit une loi de composition interne non commutative sur un ensemble particulier A^* , appelé monoïde libre; ensuite on étend cette loi aux sous-ensembles de A^* . On désire caractériser les parties X et Y de A^* qui commutent entre elles.

Dans la section 2, nous allons d'abord préciser les notations et rappeler quelques définitions. Ensuite dans la section 3, nous allons exposer les codes particuliers qui vérifient la conjecture. Enfin, dans la dernière section, après le retour sur la solution globale du problème et la généralisation aux langages quelconques, nous concluons par quelques suggestions sur les perspectives. Le vocabulaire qu'on utilisera ici est essentiellement inspiré des volumes [1] et [8].

2. Notations et définitions.

Définitions 2.1. Soit A un ensemble fixé appelé *alphabet*.

- (i) Une *lettre* est un élément quelconque de A . L'ensemble de toutes les listes à éléments dans A est noté A^* . Formellement,

$$A^* = \bigcup_{n \geq 0} A^n = \{(a_1, a_2, \dots, a_n) \mid n \in \mathbb{N} \text{ et } a_i \in A\} \text{ et } A^0 = \{1\},$$

Reçu le 16 septembre 2003 et, sous forme définitive, le 24 janvier 2005.

où 1 désigne la liste vide.

- (ii) La *concaténation* est la loi de composition interne, notée « $\cdot$ » et définie sur A^* par

$$(a_1, a_2, \dots, a_n) \cdot (b_1, b_2, \dots, b_m) = (a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_m),$$

$$1 \cdot (a_1, a_2, \dots, a_n) = (a_1, a_2, \dots, a_n) \cdot 1 = (a_1, a_2, \dots, a_n).$$

- (iii) Muni de cette loi, l'ensemble A^* admet une structure de *monoïde* dit aussi *monoïde libre*. Pour alléger les notations, on écrira $a_1 a_2 \cdots a_n$ au lieu de $(a_1, a_2, \dots, a_n)$, et l'on dira que le mot $w = a_1 a_2 \cdots a_n$ admet n pour *longueur*, ce que l'on notera par $|w| = n$. La longueur du mot particulier 1 est 0.

À partir de ces définitions, nous avons immédiatement deux **conséquences** :

- (1) Quels que soient les mots u, v, w et t tels que $uv = wt$, nous avons :
- (i) Si $|u| \geq |w|$, alors il existe un mot α tel que $u = w\alpha$ (et donc $\alpha v = t$);
 - (ii) Si $|u| < |w|$, alors il existe un mot α tel que $w = u\alpha$ (et donc $v = \alpha t$).
- (2) Quels que soient les mots u, v, w et t , alors $uv = wt$ et $|u| = |w|$ entraînent $u = w$ (et donc $v = t$).

Définition 2.2. Un *langage* est une partie quelconque de A^* . Un langage X est un *code* si toute relation de la forme

$$x_1 x_2 \cdots x_n = x'_1 x'_2 \cdots x'_m, \quad n, m \geq 1 \quad \text{et} \quad x_i, x'_j \in X,$$

entraîne

$$n = m \quad \text{et} \quad x_i = x'_i \quad \text{pour tout } i \text{ tel que } 1 \leq i \leq n.$$

Autrement dit, si on reçoit un message, son décodage en mots élémentaires ne peut conduire à une confusion. Par exemple, $X = \{ab, ba\}$ est un code. Par contre, $Y = \{a, ba, aba\}$ n'est pas un code car à la réception du message aba il risque d'y avoir une confusion à sa lecture : $aba = (a)(ba)$ ou (aba) .

La proposition suivante est connue depuis les années 50. Elle admet plusieurs démonstrations dont voici une de plus.

Proposition 2.3. Soient u et v deux mots non vides. Les trois énoncés suivants sont équivalents :

- (1) u et v commutent, c'est-à-dire $u \cdot v = v \cdot u$;
- (2) u et v sont puissances d'un même mot w , c'est-à-dire $u = w^n, v = w^m$;
- (3) Il existe $p, q \in \mathbb{N}^*$ tels que $u^p = v^q$.

Démonstration. (1) $\Rightarrow$ (2). Par récurrence sur la longueur du mot produit uv , et en se servant des conséquences précédentes.

(2) $\Rightarrow$ (3). On peut prendre $p = m$ et $q = n$.

(3) $\Rightarrow$ (1). D'une part, u commute avec u^p , donc aussi avec v^q . De plus, v commute avec v^q , donc aussi avec u^p . Puisque u et v^q commutent, nous avons

$$u^{p+1} v^{q+1} = u^p v^q uv.$$

Puisque v et u^p commutent, nous déduisons

$$u^p v^q uv = vu^p v^{q-1} uv.$$

Donc

$$u^{p+1} v^{q+1} = vu^p v^{q-1} uv.$$

D'autre part,

$$u^{p+1} v^{q+1} = uv^{2q+1}.$$

À cause de (2), comme $|uv| = |vu|$, on obtient $uv = vu$. $\square$

On s'est naturellement posé la question de savoir si ces équivalences ne peuvent pas être étendues aux ensembles de mots. Soit X un code; désignons par $\min(X)$ la longueur minimale des éléments de X , et par $X_{\min}$ l'ensemble des mots de X de longueur $\min(X)$.

Ratoandromanana [3] a justifié le résultat suivant et a présenté une démonstration laborieuse dont on retiendra que les exposants p et q ci-dessous peuvent-être choisis de telle sorte que $p = \min(Y)$ et $q = \min(X)$.

Proposition 2.4. *Soient X et Y deux codes. Les deux énoncés suivants sont équivalents :*

- (1) X et Y commutent, c'est-à-dire $X \cdot Y = Y \cdot X$;
- (2) Il existe $p, q \in \mathbb{N}^*$ tels que $X^p = Y^q$.

Dans le même article, on trouve l'énoncé de la conjecture suivante.

Conjecture. *Deux codes commutent si et seulement si ils sont puissances d'un même code.*

3. Codes particuliers vérifiant la conjecture.

Définition 3.1. Un langage X est *préfixe* si pour tout mot $a_1 \cdots a_n$ de X , aucun facteur gauche $a_1 \cdots a_i$ ($1 \leq i < n$) n'est dans X . C'est alors un code.

Exemple. $X = \{ab, aab, b^3\}$ est un code préfixe. Le message

$$m = aabababbbbab \cdots$$

se décode au fur et à mesure qu'on le reçoit :

$$m = (aab)(ab)(ab)(bbb)(ab) \cdots$$

De plus, $Y = \{a, aba\}$ est un code non préfixe car a est un facteur gauche de aba .

Définition 3.2. Un code X est *singulier à gauche* s'il admet au moins un élément x dont aucun facteur gauche n'est dans X et si x n'est pas facteur gauche d'un élément de X . Ainsi les codes préfixes deviennent des codes singuliers particuliers. Un code est *primitif* s'il n'est pas puissance d'un autre code. Un code est *homogène* si tous ses mots ont la même longueur.

D. Perrin [2] a montré, en 1972, que les codes préfixes vérifient la conjecture. Ratoandromanana [3] l'a fait aussi en 1989 pour les codes singuliers.

Lemme 3.3. *Si X, Y sont deux codes homogènes qui commutent, et si X est primitif, alors Y est une puissance de X .*

Démonstration. Les codes homogènes sont un cas particulier des codes préfixes. Donc, d'après le résultat de D. Perrin [2], il existe un code T tel que X et Y s'écrivent sous la forme $X = T^\alpha$ et $Y = T^\beta$. Comme X est primitif, $\alpha = 1$ et par suite, Y est une puissance de X . $\square$

Proposition 3.4. *Si X et Y sont deux codes qui commutent et si $X_{\min}$ est un code primitif, alors Y est une puissance de X .*

Démonstration. Puisque $XY = YX$, on peut vérifier que $X_{\min}Y_{\min} = Y_{\min}X_{\min}$. D'après le lemme précédent, il existe un entier α tel que $Y_{\min} = X_{\min}^\alpha$. D'autre part, $X^{\min(Y)} = Y^{\min(X)}$ et donc $X^{\alpha \min(X)} = Y^{\min(X)}$, ce qui entraîne $Y = X^\alpha$. $\square$

Corollaire 3.5. *Soit n un entier ≥ 1 . Si X et T^n sont deux codes qui commutent et si $T_{\min}$ est primitif, alors X est une puissance de T .*

Démonstration. Comme X et T^n commutent, $X^{\min(T^n)} = T^{n \min(X)}$. Donc $X^{n \min(T)} = (T^{\min(X)})^n$. D'où $X^{\min(T)} = T^{\min(X)}$ et par conséquent, X et T sont des codes qui commutent avec $T_{\min}$ primitif. La proposition précédente permet de conclure. $\square$

Proposition 3.6. *Soient X et Y deux codes tels que la factorisation du cardinal de X vérifie*

$$\text{Card}(X) = p_1^{\alpha_1} \cdots p_r^{\alpha_r} \text{ avec l'un des } \alpha_i = 1.$$

Si X, Y commutent, alors Y est une puissance de X .

C'est le cas par exemple si $\text{Card}(X) = 2, 3, 5, 6, 7, 10, 11, 12, 13, \dots$

Démonstration. Désignons par n et m les entiers respectifs $\min(X)$ et $\min(Y)$. On peut supposer, sans perte de généralité, que $p_1 < \cdots < p_r$ et que $\alpha_1 = 1$. Puisque X et Y sont des codes qui commutent, on a $X^{\min(Y)} = Y^{\min(X)}$; donc Y est de cardinal fini. Par suite, $(\text{Card}(X))^m = (\text{Card}(Y))^n$. Soit

$$\text{Card}(Y) = q_1^{\beta_1} \cdots q_s^{\beta_s} \text{ avec } q_1 < \cdots < q_s$$

la décomposition en facteurs premiers de l'entier $\text{Card}(Y)$. Nous avons alors

$$p_1^{m\alpha_1} \cdots p_r^{m\alpha_r} = q_1^{n\beta_1} \cdots q_s^{n\beta_s}.$$

D'après l'unicité de la décomposition d'un entier en facteurs premiers, on a $p_1 = q_1$ et $m\alpha_1 = n\beta_1$. Comme $\alpha_1 = 1$, nous obtenons $m = n\beta_1$ et la relation $X^{\min(Y)} = Y^{\min(X)}$ entraîne $Y = X^{\beta_1}$. $\square$

Plus généralement, on a :

Corollaire 3.7. *Soit T un code tel que*

$$\text{Card}(T) = p_1^{\alpha_1} \cdots p_r^{\alpha_r}, \text{ avec l'un des } \alpha_i = 1.$$

Pour tout entier naturel non nul n , si T^n commute avec le code X , alors il existe un entier naturel non nul m tel que X soit de la forme T^m .

Démonstration. T^n commute avec X implique que T commute aussi avec X car il existe des entiers p et q tels que $X^p = (T^n)^q$. On applique ensuite directement la proposition précédente. $\square$

En particulier, si l'un des deux codes qui commutent est de la forme $\{u, v\}^n$ où n est un entier non nul et $u, v \in A^* - \{1\}$, alors l'autre code est une puissance de $\{u, v\}$. C'est là un exemple de codes qui ne sont ni codes préfixes ni codes singuliers à gauche.

4. Solution de la conjecture et cas des autres langages. En 2002, Karhumaki *et al.* [6] sont venus à bout de la conjecture en répondant à une question qui peut paraître plus générale : Quels sont les ensembles Y qui commutent avec un code fixé X ?

Théorème 4.1. *Un langage quelconque Y commute avec le code X si et seulement si il existe un code Z , un entier n et un ensemble d'indices I tels que*

$$X = Z^n \text{ et } Y = \bigcup_{i \in I} Z^i.$$

La justification qu'ils ont élaborée consiste à ramener le problème aux séries formelles et plus précisément à un théorème de Cohn. Ce dernier affirme que deux séries formelles commutent si et seulement si elles peuvent être écrites à l'aide d'une seule série formelle commune. Et l'on a, comme simple cas particulier, une réponse affirmative à la conjecture de Ratoandromanana.

Corollaire 4.2. *Deux codes commutent si et seulement si ils sont puissances d'un même code.*

Démonstration. Si X et Y sont des codes qui commutent, d'après le théorème précédent, il existe un code Z , un entier n et un ensemble d'indices I tels que

$$X = Z^n \text{ et } Y = \bigcup_{i \in I} Z^i.$$

Mais si Y est un code, l'ensemble des indices I est obligatoirement un singleton et par suite X et Y sont puissances du même code Z . $\square$

Autre question : Peut-on affirmer que si deux langages finis quelconques X et Y commutent, alors ils sont union de puissances d'un même ensemble? Autrement dit, pour tout $X, Y \subset A^* \setminus \{1\}$, a-t-on

$$XY = YX \implies \exists Z \subset A^* \setminus \{1\}, I, J \subset \mathbb{N} \text{ tels que } X = \bigcup_{i \in I} Z^i \text{ et } Y = \bigcup_{j \in J} Z^j ?$$

Choffrut *et al.* [5] ont résolu le cas où $\text{Card}(X) \leq 2$. Karhumaki *et al.* [6] ont résolu le cas où $\text{Card}(X) = 3$. Les deux donnent des contre-exemples pour $\text{Card}(X) = 4$:

$$X = \{a, ab, ba, b^2\}, \quad Y = X \cup X^2 \cup \{bab, b^3\}$$

où X, Y ne sont pas union de puissances d'un même ensemble, et pourtant ils commutent. En effet,

$$XY = X^2 \cup X^3 \cup X\{bab, b^3\} \text{ et } YX = X^2 \cup X^3 \cup \{bab, b^3\}X.$$

Si on pose $Z = X\{bab, b^3\}$ et $T = \{bab, b^3\}X$, alors on a $Z \subset X^2 \cup X^3 \cup T$ et $T \subset X^2 \cup X^3 \cup Z$ et par conséquent $XY = YX$.

En guise de conclusion, si la conjecture est entièrement résolue, il reste à revoir certaines démonstrations trop compliquées ou trop longues et tenter d'en trouver de plus simples et de plus courtes.

English extended abstract. The aim of this paper is to set out and retrace the steps that lead to the resolution of an old conjecture on commuting codes. Furthermore, simple proofs are given for some particular cases. This work is related to the characterization of commuting codes, and more generally to the characterization of all sets having this property.

First, let us recall some definitions regarding combinatorics on words. Consider a finite set of symbols A that we call *alphabet*. A word u is a finite sequence of elements in A , that is:

$$u = (a_1, \dots, a_n), \quad a_i \in A.$$

We define the operation of *concatenation* as follows:

$$(a_1, \dots, a_n) \cdot (b_1, \dots, b_m) = (a_1, \dots, a_n, b_1, \dots, b_m).$$

We can denote a word u simply by $u = a_1a_2 \cdots a_n$. Moreover we denote u^n the n -th power of u , that is the word obtained by concatenating n times the word u with itself.

A language X is a set of words. A language X is a *code* if every equation $u_1 \cdots u_n = v_1 \cdots v_m$, with $u_i, v_j \in X$, gives $n = m$ and $u_i = v_i$ for all i .

Proposition 2.3. *Let A be a finite alphabet, and u, v two nonempty words. Then the following three statements are equivalent:*

- (1) *u and v are commuting : $uv = vu$.*
- (2) *u and v are powers of the same word w : $u = w^n, v = w^m$, for some $n, m \in \mathbb{N}^*$.*
- (3) *There exist p, q in $\mathbb{N}^*$ such that $u^p = v^q$.*

Ratoandromanana [3] proved that statements (2) and (3) are still true for sets that are codes. Let $XY = \{uv \mid u \in X, v \in Y\}$. Then we have the following result [3].

Proposition 2.4. *The two following statements are equivalent:*

- (1) *The codes X and Y are commuting : $XY = YX$.*
- (2) *There exist p, q in $\mathbb{N}^*$ such that $X^p = Y^q$.*

In the same paper, the author conjectured that two codes X, Y are commuting if and only if they are powers of the same code Z , i.e., $X = Z^n, Y = Z^m$ for some $n, m \in \mathbb{N}^*$.

We have treated some cases and successively showed:

Lemma 2.3. *Let X and Y be two homogeneous codes. If X is primitive and commutes with Y , then Y is a power of X : $Y = X^n$ for some $n \in \mathbb{N}^*$.*

Proposition 3.4. *Let X and Y be two commuting codes. If $X_{\min}$ is a primitive code then Y is a power of X : $Y = X^n$ for some $n \in \mathbb{N}^*$.*

Corollary 3.5. *For all $n \in \mathbb{N}^*$, if X and T^n are two commuting codes and $T_{\min}$ is primitive, then X is a power of T .*

Proposition 3.6. *Let X and Y be two commuting codes such that the factorization of the cardinality of X in prime numbers is*

$$\text{Card}(X) = p_1^{\alpha_1} \cdots p_r^{\alpha_r} \text{ with one of the exponents } \alpha_i \text{ being } 1.$$

If X and Y are commuting, then Y is a power of X .

We have this situation when $\text{Card}(X) = 2, 3, 5, 6, 7, 10, 11, 12, 13, \dots$

Corollary 3.7. *Let T be a code. Suppose that the factorization of the cardinality of T is*

$$\text{Card}(T) = p_1^{\alpha_1} \cdots p_r^{\alpha_r} \text{ with one of the exponents } \alpha_i \text{ being } 1.$$

For all $n \in \mathbb{N}^*$, if T^n commutes with a code X , then X is a power of T .

These were some particular cases. In 2002 Karhumaki et al. [6] resolved the conjecture by answering a question which seems to be more general: *What are the languages Y that are commuting with a fixed code X ?*

Theorem 4.1. *A language Y is commuting with the fixed code X if and only if there exist a code Z , an integer $n \in \mathbb{N}^*$ and a set of indexes I such that*

$$X = Z^n \text{ and } Y = \bigcup_{i \in I} Z^i.$$

Corollary 4.2. *Two codes are commuting if and only if they are powers of the same code.*

Remerciements. Je tiens à remercier le Département de Mathématiques de Mohammedia pour les conditions dans lesquelles ce travail a pu avoir lieu.

BIBLIOGRAPHIE

1. J. Berstel et D. Perrin, *Theory of Codes*, Pure and Applied Mathematics, vol. 117, Academic Press, Orlando, FL., 1985.
2. D. Perrin, *Codes conjugués*, Information and Control **20** (1972), 222-231.
3. B. Ratoandromanana, *Codes et motifs*, Inform. Théor. Appl. **23** (1989), 425-444.
4. J. M. Autebert, L. Boasson et M. Latteux, *Motifs et bases de langages*, RAIRO Inform. Théor. Appl. **23** (1989), 379-393.
5. C. Choffrut, J. Karhumaki et N. Ollinger, *The commutation of finite sets: a challenging problem*, Theoret. Comput. Sci. **273** (2002), 69-79.
6. J. Karhumaki, M. Latteux et I. Petre, *The commutation with codes and ternary sets of words*, Rapport technique no 458, TUCS, Université de Turku, Finlande, 2002.

7. I. Petre, *Commutation problems on sets of words and formal power series*, Thèse de doctorat; Turku, Finlande, Université de Turku, 2002.
8. M. Lothaire, *Combinatorics on words*, Encyclopedia of Mathematics and its Applications, vol. 17, Addison-Wesley, Reading, Mass., 1983.

O. KHADIR

LABORATOIRE D'INFORMATIQUE, MATHÉMATIQUES DISCRÈTES ET ALGÈBRE

FACULTÉ DES SCIENCES ET TECHNIQUES

UNIVERSITÉ HASSAN II-MOHAMMEDIA

B. P. 146, MOHAMMEDIA, MAROC.

COURRIEL : khadir@hotmail.com