# On the $k$-binomial Complexity of Hypercubic Billiard Words


Léo Vivion

Université du Littoral Côte d'Opale, France
leo.vivion@univ-littoral.fr



**Abstract.** We establish that the $k$-binomial complexity of hypercubic billiard words is always equal to their subword complexity.


## 1 Motivations

A hypercubic billiard word in dimension $d$ is an infinite $d$-ary word encoding the faces successively hit by a billiard ball moving in the unit cube of $\mathbb{R}^d$, in which two parallel faces are labeled by the same letter (see Figure 1). In the sequel, the parameter $\theta$ will denote the initial momentum of the ball.

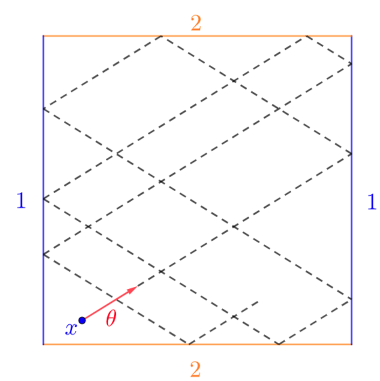


**Fig. 1.** The billiard ball, initially located in $x$ with a momentum $\theta$, generates the infinite word $w = 1211212112...$

Recently, M. Andrieu and the author of the present text computed the abelian complexity of any hypercubic billiard word whose momentum $\theta$ has rationally independent coordinates.

**Proposition 1 (Andrieu, V. [1]).** *Let $d \geq 1$ and $\theta \in \mathbb{R}^d$ with rationally independent coordinates. The abelian complexity of any hypercubic billiard word* *w with momentum $\theta$ is*

$$\rho_w(n) = \sum_{k=0}^{\min(n,d-1)} \binom{d-1}{k}.$$

This expression is surprisingly similar to the well-known subword complexity of hypercubic billiard words, which was obtained with a different method.

**Proposition 2 (Bédaride [3]).** *Let $d \geq 1$ and $\theta \in \mathbb{R}^d$ with rationally independent coordinates. If $d \geq 3$, assume moreover that for any three distinct letters $i, j, k$, the numbers $\theta_i^{-1}$, $\theta_j^{-1}$ and $\theta_k^{-1}$ are rationally independent (♣). Then, the subword complexity of any hypercubic billiard word $w$ with momentum $\theta$ is*

$$p_w(n) = \sum_{k=0}^{\min(n,d-1)} k!\binom{n}{k}\binom{d-1}{k}.$$

*Remark 1.* Note that the condition (♣) is necessary for the last formula to be true, see [2] or [4].

Hence the question: can we explain the similarity between those two expressions? More precisely:

*Question 1.* How factors of hypercubic billiard words are distributed into abelian classes?

Surprisingly, the most natural answer, which consists in believing that

> "The abelian classes of length $n$ can be partitioned into $m+1$ sets $E_0, \dots, E_m$ (where $m := \min(n, d-1)$) such that, for every $k \in \{0, \dots, m\}$, the set $E_k$ contains exactly $\binom{d-1}{k}$ abelian classes, each of them containing exactly $k!\binom{n}{k}$ factors."

is false. Indeed, if such a partition of factors into abelian classes were true, then, for every length $n \in \mathbb{N}$, the set $E_0$ should contain $\binom{d-1}{0} = 1$ abelian class, which, in turn, should contain $0!\binom{n}{0} = 1$ factor. This is in contradiction with the following result.

**Theorem 1.** *For every hypercubic billiard word $w$ whose momentum $\theta$ has rationally independent entries, there exist infinitely many lengths $n \in \mathbb{N}$ for which no abelian class of length $n$ of $w$ contains exactly one factor.*

Then, a legitimate approach towards Question 1 is to compute the $k$-binomial (resp. $k$-abelian) complexities [8,5] of hypercubic billiard words. Indeed, since these complexities form a scale from the abelian complexity to the subword complexity, our idea is to understand how factors are *progressively* partitioned into $k$-binomial (resp. $k$-abelian) classes.

In this extended abstract, we focus on the $k$-binomial complexity of hypercubic billiard words. The computation of their $k$-abelian complexity is the purpose of another ongoing work.

## 2 Definitions and notations

Let $\mathcal{A}$ be a finite set, called *alphabet*. A *finite word* $w$ written over the alphabet $\mathcal{A}$ is an element of $\mathcal{A}^* := \bigcup_{i=0}^{\infty} \mathcal{A}^i$. For $k \in \mathbb{N}_{>0}$, we denote by $\mathcal{A}^{\leq k} := \bigcup_{i=1}^{k} \mathcal{A}^i$ the set of all non-empty finite words of length less than $k$. A (right) *infinite word* $w$ written over the alphabet $\mathcal{A}$ is an element of $\mathcal{A}^{\mathbb{N}}$.

A *factor* $u$ of length $n$ of a word $w$ is a finite word made of $n$ consecutive letters in $w$, while a *scattered factor* of length $n$ of $w$ is a word made of $n$, non-necessarily consecutive, letters of $w$. For instance, if $w = 11212$, then 11 is a factor (and then, a scattered factor) of $w$, 22 is a scattered factor but not a factor of $w$, and 221 is neither a factor nor a scattered factor of $w$.

For a finite word $w \in \mathcal{A}^*$, we denote by $|w|$ its length, by $|w|_a$ the number of occurrences of the letter $a \in \mathcal{A}$ in $w$, by $|w|_u$ the number of occurrences of the finite word $u \in \mathcal{A}^*$ in $w$ as a factor, and by $\binom{w}{u}$ its number of occurrences in $w$ as a scattered factor. For instance, if $w = 11212$, then $|w| = 5$, $|w|_1 = 3$, $|w|_{12} = 2$ and $\binom{w}{12} = 5$.

We say that a word $w$ is *$c$-balanced* ($c \in \mathbb{N}$) when, for every equally long factors $u, v$ of $w$, and every letter $a \in \mathcal{A}$, $\big||u|_a - |v|_a\big| \leq c$. For example, the word $w = 11212$ is 1-balanced while the word $w' = 11122$ is 2-balanced but not 1-balanced.

We say that two finite words $u, v \in \mathcal{A}^*$ are *abelian equivalent* (resp. *$k$-binomially equivalent* (where $k \in \mathbb{N}_{>0}$ is a parameter)) when, for every letter $a \in \mathcal{A}$, $|u|_a = |v|_a$ (resp. when, for every finite word $x \in \mathcal{A}^{\leq k}$, $\binom{u}{x} = \binom{v}{x}$). In that case, we write $u \sim_{\text{ab}} v$ (resp. $u \sim_k v$). For example, if $u = 1212221$ and $v = 2112212$, then $u \sim_{\text{ab}} v$, $u \sim_2 v$ but $u \nsim_3 v$.

These binary relations are equivalence relations. In particular, they partition factors into abelian (resp. $k$-binomial) classes.

The *subword complexity* (resp. *abelian complexity*, resp. *$k$-binomial complexity*) of a word $w$ is the function $p_w : \mathbb{N} \to \mathbb{N}$ (resp. $\rho_w : \mathbb{N} \to \mathbb{N}$, resp. $b_w^k : \mathbb{N} \to \mathbb{N}$) which counts, for every integer $n \in \mathbb{N}$, the number of distinct factors (resp. abelian classes, resp. $k$-binomial classes) of length[1] $n$ of $w$.

## 3 New results

We say that a word $w$ satisfies the property **(P)** if, for every integer $k \geq 2$, its $k$-binomial complexity is equal to its subword complexity. Note that, since for every word $w$ and every $k \in \mathbb{N}_{>0}$ we have $b_w^k \leq b_w^{k+1} \leq p_w$, it is equivalent to only requiring that $b_w^2 = p_w$.

On the one hand, it is known that words satisfying the property **(P)** exist. It is the case, for instance, for Sturmian words [8], and also for the Tribonacci word [6]. On the other hand, there exist words that do not satisfy this property. For

[1] Since two factors are abelian (resp. $k$-binomially) equivalent only if they are of the same length, an abelian (resp. $k$-binomial) class contains only equally long factors. The common length of the factors of a given abelian (resp. $k$-binomial) class is called its length.

example, the Thue-Morse word and, more generally, any aperiodic word obtained as a fixed point of a Parikh constant substitution. Indeed, these words have an unbounded subword complexity, while their $k$-binomial complexity is bounded for every $k \in \mathbb{N}_{>0}$ [8, Theorem 13].

Our main result asserts that hypercubic billiard words satisfy the property **(P)**.

**Theorem 2.** *Let $d \geq 1$. For every integer $k \geq 2$, the $k$-binomial complexity of any hypercubic billiard word in dimension $d$ is equal to its subword complexity.*

*Remark 2.* This theorem is true for *every* hypercubic billiard words, even those for which the subword complexity is not given by Proposition 2.

Our proof strategy also enables us to establish the following result, which is, in the case $d = 2$, a refinement of [8, Theorem 7].

**Theorem 3.** *Let $d \geq 1$. For every integer $k \geq 2$, the $k$-binomial complexity of any $d$-ary, 1-balanced word is equal to its subword complexity.*

*Remark 3.* Since Sturmian words are 1-balanced, since the Tribonacci word is 2 balanced [7], and since any hypercubic billiard word in dimension $d$ is $(d-1)$-balanced [9], one may think that the property **(P)** (or some of its variant allowing that $b_w^k = p_w$ only for sufficiently large values of $k$) could be true for any word satisfying a "good" balancedness property. However, let us recall that the Thue-Morse word is 2-balanced, and yet, does not satisfy the property **(P)**: more precisely, its $k$-binomial complexity is *never* equal to its subword complexity.

## References

1. Andrieu, M., Vivion, L.: Minimal complexities for infinite words written with $d$ letters. In: Combinatorics on Words. Springer International Publishing (2023)
2. Bédaride, N.: Classification of rotations on the torus $\mathbb{T}^2$. Theoretical Computer Science **385**(1), 214–225 (2007)
3. Bédaride, N.: Directional complexity of the hypercubic billiard. Discrete Mathematics **309**(8), 2053–2066 (2009)
4. Borel, J.P.: A geometrical characterization of factors of multidimensional billiard words and some applications. Theoretical Computer Science **380**(3), 286–303 (2007)
5. Karhumäki, J., Saarela, A., Zamboni, L.Q.: On a generalization of abelian equivalence and complexity of infinite words. Journal of Combinatorial Theory, Series A **120**(8), 2189–2206 (2013)
6. Lejeune, M., Rigo, M., Rosenfeld, M.: Templates for the k-binomial complexity of the tribonacci word. Advances in Applied Mathematics **112**, 101947 (2020)
7. Richomme, G., Saari, K., Zamboni, L.Q.: Balance and abelian complexity of the Tribonacci word. Advances in Applied Mathematics **45**(2), 212–231 (2010)
8. Rigo, M., Salimov, P.: Another generalization of abelian equivalence: Binomial complexity of infinite words. Theoretical Computer Science **601**, 47–57 (2015)
9. Vuillon, L.: Balanced words. Bulletin of the Belgian Mathematical Society - Simon Stevin **10**(5), 787–805 (2003)