BULLETIN DE LA S. M. F.

PIERRE ARNOUX
CHRISTIAN MAUDUIT
IEKATA SHIOKAWA
JUN-ICHI TAMURA

Complexity of sequences defined by billiard in the cube

Bulletin de la S. M. F., tome 122, nº 1 (1994), p. 1-12
<http://www.numdam.org/item?id=BSMF_1994__122_1_1_0>

Bull. Soc. math. France,
122, 1994, p. 1–12.

COMPLEXITY OF SEQUENCES DEFINED BY BILLIARD IN THE CUBE

BY

PIERRE ARNOUX, CHRISTIAN MAUDUIT,

IEKATA SHIOKAWA and JUN-ICHI TAMURA (*)

RÉSUMÉ. — Nous démontrons une conjecture de Gérard Rauzy relative à la structure des trajectoires de billard dans un cube. A chaque trajectoire on associe la suite à valeurs dans $\{\mathbf{1}, \mathbf{2}, \mathbf{3}\}$ obtenue en codant par un **1** (resp. **2**, **3**) chaque rebond sur une paroi frontale (resp. latérale, horizontale). Nous montrons que si la direction initiale est totalement irrationnelle, le nombre de sous-mots distincts apparaissant dans cette suite est exactement $n^2 + n + 1$.

ABSTRACT. — We prove a conjecture of Gérard Rauzy related to the structure of billiard trajectories in the cube : let us associate to any such trajectory the sequence with values in $\{\mathbf{1}, \mathbf{2}, \mathbf{3}\}$ given by coding **1** (resp. **2**, **3**) any time the particle rebounds on a frontal (resp. lateral, horizontal) side of the cube. We show that, if the direction is totally irrational, the number of distinct finite words of length n appearing in this sequence is exactly $n^2 + n + 1$.

1. Statement of the result

We consider billiard problems with elastic reflexion on the boundary; the simplest of these problems is the billiard in the square. It is natural to code an orbit for this billiard problem with initial direction (α, β) by the sequence of the sides it meets, coding **1** for vertical sides and **2** for

(*) Texte reçu le 14 janvier 1992.
P. ARNOUX, Laboratoire de Mathématiques Discrètes, UPR 9016, Faculté des Sciences de Luminy, Case 930, 163, avenue de Luminy, 13288 Marseille CEDEX 9 France.
Ch. MAUDUIT, Laboratoire de Logique, Mathématiques Discrètes et Informatique, Université Lyon I, 43 Bd du 11 novembre 1918, 69622 Villeurbanne CEDEX France.
I. SHIOKAWA, Department of Mathematics, Keio University, Hiyoshi, Yokohama, 223 Japan.
J.-I. TAMURA, Faculty of General Education, International Junior College, Ekoda 4-15-1, Nakano-ku, Tokyo, 165 Japan.
AMS classification : 58 F 03, 05 A 15, 05 B 45.

horizontal sides; one can then show, if α and β are rationally independant, that the number of words of length n appearing in this sequence is equal to $n+1$; this is the minimal number of words for a non-periodic sequence, and caracterizes so-called Sturmian sequences (cf. [HM]). This problem is analogous to the study of the intersection of a line of slope β/α in the xy-plane with the net formed by the lines $x=n$ and $y=n$ $(n \in \mathbb{N})$.

An immediate generalization of this problem is the billiard in the 3-dimensional cube. Let us denote by I^3 the unit cube of $\mathbb{R}^3$,

$$I^3 = \{(x_1, x_2, x_3) \mid 0 \le x_i \le 1,\ i = 1, 2, 3\}.$$

The billiard flow in I^3 is the geodesic flow with respect to the natural euclidian metric on the unit tangent bundle, with elastic reflexions on the boundary. This means that, on the face $x_1 = 0$, we identify the two points $(0, x_2, x_3, \alpha_1, \alpha_2, \alpha_3)$ and $(0, x_2, x_3, -\alpha_1, \alpha_2, \alpha_3)$ of the unit tangent bundle $I^3 \times S^3$, and similarly for the other faces. It is clear that the set $I^3 \times (\pm\alpha_1, \pm\alpha_2, \pm\alpha_3)$ is invariant by the flow, and it is classical that, if α_1, α_2 and α_3 are rationally independent (we say that the vector $(\alpha_1, \alpha_2, \alpha_3)$ is *totally irrational*), the restriction of the flow to this set is minimal and uniquely ergodic. In fact, it can then be reduced to a flow on the three-dimensional torus; we will use this fact later.

We can code an orbit by a sequence $(u_n)_{n\in\mathbb{N}} \in \{\mathbf{1}, \mathbf{2}, \mathbf{3}\}^{\mathbb{N}}$, $u_n = \mathbf{i}$ if the n-th face met by the orbit is $x_{\mathbf{i}} = 0$ or $x_{\mathbf{i}} = 1$.

DEFINITION. — Let $u = (u_n)_{n\in\mathbb{N}}$ be a sequence with values in a finite set A. The *language* associated to u is the set $L(u)$ of finite words appearing in $(u_n)_{n\in\mathbb{N}}$, i.e. the set of words $u_i u_{i+1} \dots u_{i+k}$, $(i, k) \in \mathbb{N}^2$. The complexity of the sequence u is the function :

$$p(n) = \#L(u) \cap A^n.$$

In this paper, we prove the following :

THEOREM. — *The complexity of a sequence generated by the cubic billiard with totally irrational initial direction is $p(n) = n^2 + n + 1$.*

This theorem (cf. [R2], [R3]) was first conjectured by Gérard RAUZY in 1981, and proved independently some months ago by the french and the japanese authors of this paper.

To prove the theorem, we will first, in part 2 and 3, reduce the problem to the following : given a certain partition of the torus $\mathbb{T}^2$ in three parallelograms and a translation of $\mathbb{T}^2$, find the number of connected components of the intersection of the partition with its first n images by the translation; we finish the proof by a simple combinatorial argument in part 4. In the last parts, we make a few remarks about possible extensions.

2. Symbolic dynamics associated to the cubic billiard

Along the trajectory, the velocity will be $(\pm\alpha_1, \pm\alpha_2, \pm\alpha_3)$, the signs being determined by the parity of the number of reflexions in the corresponding faces. Since we denote the two parallel faces in each pair by the same symbol, the coding of a trajectory is not changed if we make a symmetry with respect to one of the three planes $x_i = \frac{1}{2},\ i = 1, 2, 3$. So for each point with given direction, we can find by symmetry seven other points with the same coding, and exactly one of these eight points has a direction with all components positive; if we always choose this point for representant, we get the flow with slope $(\alpha_1, \alpha_2, \alpha_3)$ on the torus $\mathbb{T}^3 = \mathbb{R}^3/\mathbb{Z}^3$. This flow has the same coding as the billard, when we code a trajectory by the crossing of the projections of the faces on the torus.

If we go to the universal cover, this gives a very simple definition of the billiard sequences : take a line with direction v, and code the order in which it meets the planes $x_i = n,\ n \in \mathbb{Z}$.

REMARK. — Strictly speaking, we have defined the flow only if the particle meets the interior of a face; we can define the flow when it meets the edge intersection of the faces $x_i = n_i$, $x_j = n_j$, $(n_i, n_j) \in \{0, 1\}^2$ as a succession of reflexions on the two faces; since these reflexions commute, this is well defined. There is no natural way to define a unique coding for this point, so we will accept the two sequences **ij** and **ji** on this edge. This does not cause problems in the totally irrational case, because then the orbit can meet at most once an edge of type **ij**, otherwise it should be contained in a rational plane. In particular, the orbit starting from $(0, 0, 0)$ never meets an edge (In the rational case, things are more complicated; in particular, if the direction is a multiple of an integer vector, each orbit is periodic; if we allow arbitrary coding for meeting of the edges, we could obtain a non-periodic sequence, and in this case, we have to choose a consistent set of rules for the coding). We can extend the flow to the vertices in the same way, accepting the six words **ijk** to code the vertices.

Let us recall that the shift σ on the set of sequences with value in $\{\mathbf{1}, \mathbf{2}, \mathbf{3}\}$ is defined by : $\sigma((u_n)_{n\in\mathbb{N}}) = (v_n)_{n\in\mathbb{N}}$ with $v_n = u_{n+1}$. The coding we have defined is not injective, because we are working with a flow, and not with a map (two points on a segment of orbit that does not meet the faces have same associated sequences); we can restrict to points on the faces, and it is clear that the first return map of the flow on this set is conjugate to the shift on the admissible sequences.

LEMMA. — *Let v be a totally irrational direction, and Ω_v be the set of all admissible sequences for the billiard with initial direction v. The* *set* Ω_v *is a closed subset of* $\{\mathbf{1},\mathbf{2},\mathbf{3}\}^{\mathbb{N}}$, *invariant by* σ, *and the dynamical system* (Ω_v, σ) *is minimal (i.e., every orbit is dense).*

Proof. — Let us consider the map ϕ which associates to an admissible sequence u the point x in the union of faces whose orbit has coding u. This map is well defined, because, by minimality of an irrational flow on the torus, no two points have the same coding; it is continuous, but not injective, and it conjugates the shift to the first return map of the flow.

The inverse map ϕ^{-1}, which associates to a point the set of coding sequences (1,2,4 or 6 depending whether the orbit meets an edge or a vertex) is injective, not continuous (this is impossible for topological reasons), but it has the following property : if the points x_n converge to x, and if we have associated sequences $u(x_n)$ that converge to a sequence u, then u is one of the coding sequences for x. Here, we must take care that each points may have several coding sequences, and also that, if the orbit of x goes through an edge or a vertex, we can find x_n converging to x such that the associated coding sequences do not converge (but all convergent subsequences will go to coding sequences for x).

This implies immediately that Ω_v is invariant by the shift, and closed : if admissible sequences u_n converge to a sequence u, we can suppose, by compacity, that the points $x_n = \phi(u_n)$ converge to x, and by the above property, u is one of the coding sequences associated to x. The minimality of (Ω_v, σ) is an immediate consequence of the minimality of the first return map. □

This implies that, while the sequence depends on the initial position, the set of finite words appearing in this sequence depends only on the initial direction; in particular, all sequences with same initial direction have the same complexity. Instead of counting the number of different words in a given sequence, one can find the complexity by computing the number of different initial segments of length n; more precisely, we can restrict this to trajectories that do not meet an edge before the n-th crossing. Our proof will be based on this fact.

3. Reduction to a translation on $\mathbb{T}^2$

We use the description of admissible sequences as generated by a line in $\mathbb{R}^3$; we need some definitions. We note the canonical basis of $\mathbb{R}^3$ by e_1, e_2, e_3; we define the set of integral points of height n by :

$$\widetilde{P}_n = \{ae_1 + be_2 + ce_3 \mid (a,b,c) \in \mathbb{Z}^3,\ a+b+c = n\}.$$

We call H the diagonal plane $H = \{(x,y,z), x+y+z = 0\}$, we denote by

π the projection of $\mathbb{R}^3$ onto H along the direction $(\alpha_1, \alpha_2, \alpha_3)$, and define $f_i = \pi(e_i)$; remark that $\widetilde{P}_0$ is a lattice in H, generated by $f_i - f_j = e_i - e_j$.

In what follows, $\widetilde{\Sigma}$ (or $\widetilde{P}, \widetilde{F}, \ldots$) will always be some object in $\mathbb{R}^3$, $\bar{\Sigma}$ its projection on H ; this projection will be invariant by the lattice $\widetilde{P}_0$, and Σ will be the quotient by $\widetilde{P}_0$, subset of the torus $H/\widetilde{P}_0$ (on some occasions, Σ will also denote a set of representatives of $\bar{\Sigma}$ modulo $\widetilde{P}_0$, that is a subset of H).

We call $\widetilde{L}_n$ the set of segments joining points of $\widetilde{P}_n$ to points of $\widetilde{P}_{n+1}$ in the three directions :

$$\widetilde{L}_n = \{z + te_i \mid z \in \widetilde{P}_n,\ i = 1,2,3,\ t \in]0,1[\}.$$

We call *face of type* k a face contained in a plane $x_k = 0$, and we note $\widetilde{F}_n^{(k)}$ the set of faces of height n and type k :

$$\widetilde{F}_n^{(k)} = \{z + te_i + se_j \mid z \in \widetilde{P}_n,\ i,j,k \text{ all distinct},\ (s,t) \in]0,1[^2\}.$$

Finally, we call $\widetilde{F}_n$ the set of all faces of height n :

$$\widetilde{F}_n = \bigcup_{k=1,2,3} \widetilde{F}_n^{(k)}$$

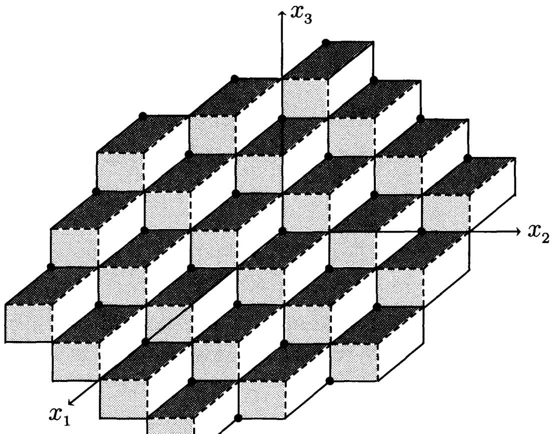

Figure 1. We have represented a part of the stepped surface $\widetilde{\Sigma}_0$; points of $\widetilde{P}_0$ are seen as big dots, edges of $\widetilde{L}_0$ are continuous lines, edges of $\widetilde{L}_1$ are dashed lines, and the three types of faces of $\widetilde{F}_0$ are shaded differently.

The union of $\widetilde{F}_n$ and the adjacent edges, $\widetilde{L}_n$ and $\widetilde{L}_{n+1}$, and vertices, $\widetilde{P}_n$, $\widetilde{P}_{n+1}$ and $\widetilde{P}_{n+2}$ form a cellular complex that we will call the *stepped surface* $\widetilde{\Sigma}_n$ of height n.

It is clear that every orbit crosses each $\widetilde{\Sigma}_n$ exactly once; if it does not go through an edge or a vertex, it will cross $\widetilde{\Sigma}_n$ in $\widetilde{F}_n$, and if we restrict to orbits starting on $\widetilde{\Sigma}_0$ and crossing $\widetilde{F}_n$, the coordinate of rank n of the associated sequence will be the k such that the orbit crosses $\widetilde{F}_n^{(k)}$.

We can parametrize the orbits starting from $\widetilde{\Sigma}_0$ by the plane H. It is clear that the restriction of π to $\widetilde{\Sigma}_n$ is one-to-one, and the projection of the cellular decomposition of $\widetilde{\Sigma}_n$ gives a cellular decomposition $\bar{\Sigma}_n$ of H, invariant by the lattice $\widetilde{P}_0$. Since we have $\widetilde{\Sigma}_n = \widetilde{\Sigma}_0 + ne_1$, the cellular decomposition of H obtained by projection of $\widetilde{\Sigma}_n$ can be deduced by a translation of $n\pi(e_1)$ from $\bar{\Sigma}_0$ (it is also deduced by a translation of $n\pi(e_2)$ or $n\pi(e_3)$, since $\pi(e_i - e_j) = e_i - e_j \in \widetilde{P}_0$).

We now quotient by the lattice $\widetilde{P}_0$, and consider the irrational translation T on $\mathbb{T}^2 = H/\widetilde{P}_0$ of vector $f_1 = f_2 = f_3 \bmod \widetilde{P}_0$, and the cellular complex Σ_0 quotient of $\bar{\Sigma}_0$, which is made of three parallelograms; from the above remarks, the number of different words of length n is equal to the number of sets intersection of a face of this complex with faces of its $(n-1)$ first iterates by the translation. The following lemma allows us to count only the number of connected components of these intersections :

LEMMA. — *Each nonempty intersection of faces of the complexes* $\Sigma_0, T\Sigma_0 = \Sigma_1, \ldots, T^k\Sigma_0 = \Sigma_k$ *is connected.*

Proof. — Consider the three faces $F^{(k)}$ of $\bar{\Sigma}_0$ with a vertex at the origin, defined by :

$$F^{(k)} = \{x_i f_i + x_j f_j \mid 0 < x_i, x_j < 1,\ i \neq k,\ j \neq k\}.$$

They form a fundamental domain for the action of $\widetilde{P}_0$ on H (up to a set of measure 0 : edges and vertices); this domain is a convex hexagon made of three parallelograms.

It is an immediate consequence that, for any two points x, y in the same face of Σ_0, there exists a unique segment joining them and contained in that face : indeed, we can find k and two points $\bar{x}, \bar{y} \in F^{(k)}$, uniquely defined, that project to x, y; the face $F^{(k)}$ being a parallelogram, hence convex, there is a segment joining $\bar{x}$ to $\bar{y}$ and contained $F^{(k)}$, and it projects to a segment joining x to y in a face of Σ_0 : this proves the existence. On the other hand, a segment joining x to y without crossing any ege or vertex can be lifted to a segment completely contained in some $F^{(k)}$ which must be the unique segment from $\bar{x}$ to $\bar{y}$, hence the unicity.

Let us remark here that it is not always true that the intersection of faces of Σ_0 and $\Sigma_k, k > 1$ is connected (we cannot use directly on the torus convexity properties) : two points in this intersection can be joined by a path contained in a face of Σ_0, and also by a path contained in a face of Σ_k, but these path have no reason to be the same. Figure 2 shows an exemple where this intersection is not connected.

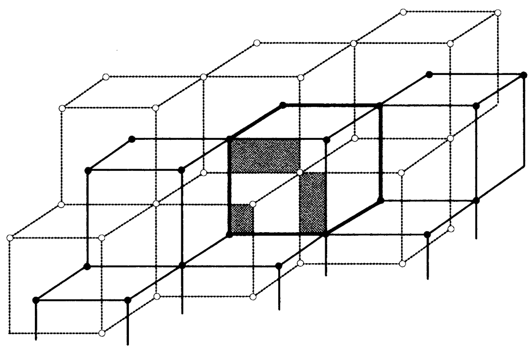

Figure 2. We have represented a complex Σ_0, in full lines, with a fundamental domain in thicker lines ; Σ_0 has three kinds of faces, a big square and two smaller parallelograms. The translate $T^3\Sigma_0$ is represented in dashed lines, and the dashed area shows the intersection of a square face of Σ_0 with the square faces of the translate ; this intersection is not connected.

The situation is different in the case of two consecutive complex Σ_i and Σ_{i+1}. For exemple, there is a convex fundamental domain for the action of $\widetilde{P}_0$ on H which consists of cells of $\bar{\Sigma}_0$, and also of cells of $\bar{\Sigma}_1$: consider the above fondamental domain made of three faces of $\bar{\Sigma}_0$ having the origin as a vertex. There is a unique point in $\widetilde{P}_3$ in this domain ; take the three faces of $\bar{\Sigma}_1$ having this point as a vertex, it is easy to check that this defines the same fundamental domain (cf. figure 3 ; it is the picture one gets when drawing a cube in projection, and this is exactly what we do, since $\widetilde{\Sigma}_0$ and $\widetilde{\Sigma}_1$ are respectively the lower and upper faces of a set of cubes). Now if two points are in the intersection of faces of Σ_0 and Σ_1, we can lift them to this common fundamental domain, and we see that there is a segment joining them in the intersection, since in the plane the intersections of faces are convex sets.

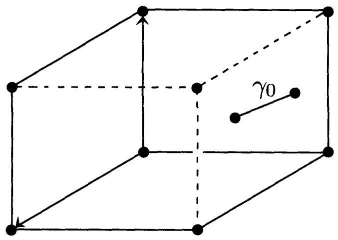

Figure 3

It is an immediate consequence that the intersection of faces of $\Sigma_0, \Sigma_1, \dots \Sigma_k$ is connected, when we take consecutive complexes; for if x, y are two points in the intersection, they are joined by segments $\gamma_0, \gamma_1, \dots, \gamma_k$, with γ_i contained in a face of Σ_i; but we just proved that, in that case, we have $\gamma_i = \gamma_{i+1}$, so that all the segments are equal, and the unique segment obtained is contained in the intersection of the faces. $\square$

4. Proof of the theorem

We can still simplify a little more; call C_n the cellular decomposition of the torus obtained by projection of $\widetilde{L}_n$, $\widetilde{P}_n$ and $\widetilde{P}_{n+1}$ on H and quotient by $\widetilde{P}_0$; this is a very simple cellular complex, with just one face, three edges and two vertices, and it is clear that Σ_n is just the intersection of C_n and C_{n+1}. If we note by D_n the complex given by the intersection of the C_i, $0 \le i \le n$, what we must prove is that the number of faces of D_n is n^2+n+1; we will call A_n this number of faces, E_n the number of edges and V_n the number of vertices. The proof follows easily from two lemmas :

LEMMA 1. — *The edges of C_i and C_j, $i < j$, meet in exactly two points, except if $j = i+1$, in which case they meet in their common extremity P_j.*

Proof. — Since all the partitions C_i are deduced from each other by the translation T, it is enough to consider C_1 and C_j. The simplest way is to consider the lifting of these partitions to the plane; consider the sets $F^{(k)}$ above, which, up to a set of dimension 1, form a fundamental domain for the lattice $\widetilde{P}_0$. The boundary of these sets project to the edges of C_0 and C_1, namely, the segments $\{tf_i \mid 0 < t < 1\}$ project to the edges of C_0, and the segments $\{f_i + tf_j \mid 0 < t < 1, i \neq j\}$ project to the edges of C_1 (cf. figure 4).

This fundamental domain is an hexagon, with three vertices projecting to the point P_1, and the three other vertices to P_2; it is then clear that the edges of C_1 and C_2 meet only in P_2 (cf. figure 4); the complex represented in figure 4 is just the intersection of Σ_0 and Σ_1.

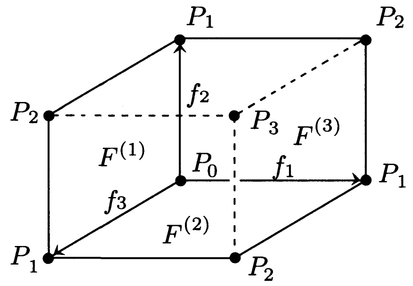

Figure 4

Now, if $i > 2$, it is a consequence of the irrationality that the unique representant of P_i in the fondamental domain is in the interior of one of the $F^{(k)}$. Suppose it is in $F^{(1)}$, then the point $P_i + f_1$ is in the fundamental domain, hence by convexity the corresponding edge of C_i does not intersect C_1, and the points $P_i + f_2$, $P_1 + f_3$ are in adjacent domains, hence each of the corresponding edges meet C_1 in exactly one point (cf. figure 5). $\square$

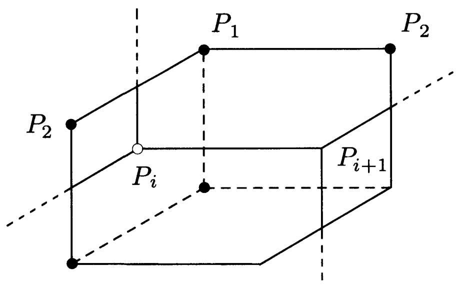

Figure 5

LEMMA 2. — *The number of vertices and edges is given by :*

$$V_n = n^2 + 2, \quad E_n = 2n^2 + n + 3.$$

Proof. — The vertices of D_i are the points P_i, extremities of the edges of C_i, $0 \le i \le n$, and intersection points of edges of C_i and C_j in the interior of the edges. Going from D_n to D_{n+1}, we add the point P_{n+1}, and the intersection points of the edges of C_{n+1} with those of $C_i, 0 \le i \le n$; all these points are distict, by the irrationality condition, so there are two of them for each $i \neq n$ and none for $i = n$, by the preceding lemma, hence :

$$V_{n+1} = V_n + 2n + 1.$$

Since $V_0 = 2$ (it is reduced to points P_0, P_1) we get the formula.

Now the edges consist of the segments of the edges of C_i between two vertices; going from D_n to D_{n+1}, we add three new edges, and we must consider also the intersection with old edges. Each of the $2n$ vertices given by these intersections defines two new edges, hence :

$$E_{n+1} = E_n + 4n + 3$$

and we get the formula, for it is clear that $E_0 = 3$. $\square$

Proof of the theorem. — Since the Euler characteristic of the torus is 0, we have the Euler formula

$$A_n - E_n + V_n = 0$$

from which we get :

$$A_n = n^2 + n + 1. \quad \square$$

5. Higher dimension cases

A natural continuation would be to find the complexity for the billiard in the cube in higher dimensions. We can apply the same method in dimension 4, but counting the vertices and the edges becomes very difficult; in fact, it is not known whether the complexity is independant from the initial direction, and whether it is polynomial. However, if we call $P(n, s)$ the complexity of the irrational $s + 1$-dimensional billiard (assuming that this function is well defined), numerical computations by Jun-ichi Tamura hint to interesting relations between these numbers : it seems that, for $n > s$, $P(n, s) \geq P(s, n)$.

For small n or s, much more is true : if $\inf(n, s) < 3$, then $P(n, s)$ is well defined, and $P(n, s) = P(s, n)$. The number $P(1, s)$ is just the number of symbol used for coding the billiard in dimension $s + 1$, that is the number of pairs of parallel faces for the hypercube in dimension $s + 1$, which is clearly $s + 1$.

$P(2, s)$ is the number of admissible pairs of symbols for the billiard with totally irrational initial direction $(\alpha_1, \ldots, \alpha_{s+1})$. Consider the coding as given by the irrational flow on the torus, by the same analysis as in section 2. There are only $(s+1)^2$ pairs of symbols; all the pairs **ij**, with $\mathbf{i} \neq \mathbf{j}$ are admissible, because segments of trajectories passing near the edge intersection of face number **i** and face number **j** will be coded by **ij** on one side of the edge, and by **ji** on the other. But, by minimality of the irrational flow, any trajectory will pass in a given neighborhood of such a segment. The only pairs left are of the type **ii**. If we suppose that $\alpha_1 > \alpha_i$, for all $i > 1$, then it is easy to prove that the pair **11** is admissible, and not the other ones : this amounts to prove that there are times t, t' such that $t\alpha_1 = k \in \mathbb{N}$, $t'\alpha_1 = k+1$, and $\tau\alpha_j \notin \mathbb{N}$ for all $j > 1$ and all $\tau \in]t, t'[$ (it suffices to take $t = 0$), while it is impossible to have the same property for α_j. Hence there are exactly s^2+s+1 pairs of admissible symbols, and we see that $P(2,s) = P(s,2)$.

We can also remark that $P(n,s)$ is bounded by $(s+1)^n$; if, for fixed n, $P(n,s)$ is a polynomial in s, the degree of this polynomial is at most n. In fact, for $n = 0, 1, 2$, it is a unitary polynomial with integral coefficients.

If we make the hypotheses that $P(n,s)$ is defined for all integers n and s, symmetric $(P(n,s) = P(s,n))$, and that, for fixed n, it is a unitary polynomial in s of degree n, it is easy to see that it is uniquely defined by this conditions : if we know $P(n,s)$ for $n > N$, then, by symmetry, we know N value of $P(N,s)$, and we can determine the coefficients of this polynomial. For example $P(3,s)$ is equal to $s^3+2s+1 = s(s-1)(s-2)+3s(s-1)+3s+1$, and if we note $x^{\underline{n}} = x(x-1)\dots(x-n+1)$, it is easily proved that $P(n,s) = \sum_{i=0}^{n} \binom{n}{i} s^{\underline{i}}$, giving the hypothetical general formula :

$$P(n,s) = \sum_{i=0}^{\inf(n,s)} \frac{n\,!\,s\,!}{(n-i)\,!\,i\,!\,(s-i)\,!}.$$

This formula is in quite good agreement with the numerical data found up to now.

6. Additional remarks

It is possible to find the complexity also in the case where the vector is not totally irrational; if the vector is completely rational, then the sequence is periodic, and the complexity is bounded by a function of the direction.

An open problem is to characterize exactly the set of sequences that can be obtained from cubic billiards; in the case of the square, these are exactly the sturmian sequences. In particular, it should be interesting to know the recurrence function of these sequences, defined as the function which, to any integer n, associates the smallest integer $R(n)$ such that every word of length $R(n)$ contains all words of length n. This function is well defined, because the system is minimal, so every word appearing in the sequence appears an infinite number of times, with bounded gaps. In the case of sturmian sequences, the recurrence function is related to the development in continued fraction of the frequency of one of the symbols (cf. [HM]). Finding the recurrence function for cubic billard could lead to a 2-dimensional continued fraction expansion, and could also give an algorithm for the multiplication of ordinary continued fractions.

The complexity depends, not only on the system, but also on the coding chosen for this system; in some cases, it is possible to chose a coding for a translation of the torus by sequences of the much lower complexity $2n+1$ (cf. [AR], [R1]). It should be interesting to know the minimal complexity for an injective coding of a translation of the torus or a cubic billiards. We can remark that this minimal complexity can only be an asymptotic notion : chosing a partition in two atoms with an atom much bigger than the other, we can, for any given integer N, find an injective coding with complexity $p(n) = n + 1$ for all $n < N$, which is the minimum possible complexity for a nonperiodic sequence.

BIBLIOGRAPHY

[AR] Arnoux (P.) et Rauzy (G.). — *Représentation géométrique de suites de complexité* $2n+1$, Bull. Soc. Math. France, t. **119**, 1991, p. 199–215.

[HM] Hedlund (G.A.) and Morse (M.). — *Symbolic dynamics II. Sturmian trajectories*, Amer. J. Math., t. **62**, 1940, p. 1–42.

[R1] Rauzy (G.). — *Nombres algébriques et substitutions*, Bull. Soc. Math. France, t. **110**, 1982, p. 147–178.

[R2] Rauzy (G.). — *Suites à termes dans un alphabet fini*, Séminaire de Théorie des Nombres de Bordeaux, t. **25**, 1982–1983, p. 1–16.

[R3] Rauzy (G.). — Mots infinis en arithmétique, in *Automata on infinite words*, Lecture Notes in Computer Sciences **192**, 1985, p. 165–171.