Commun. Math. Phys. 174, 43 – 56 (1995)

Communications in
Mathematical
Physics

Complexity of Trajectories in Rectangular Billiards

Yu. Baryshnikov★

Department of Mathematics, University of Osnabrück, D-49069 Osnabrück, Germany
E-mail: yuliy@aphrodite.mathematik.uni-osnabrueck.de

Received: 20 May 1994/in revised form: 30 January 1995

Abstract: To a trajectory of the billiard in a cube we assign its symbolic trajectory – the sequence of numbers of coordinate planes, to which the faces met by the trajectory are parallel. The complexity of the trajectory is the number of different words of length n occurring in it. We prove that for generic trajectories the complexity is well defined and calculate it, confirming the conjecture of Arnoux, Mauduit, Shiokawa and Tamura [AMST].

0. Introduction

Consider a rectangular billiard in $\mathbb{R}^{s+1}$, that is the dynamical system defined by the free motion of the point between collisions with the boundary of the billiard domain and elastic reflections at the collision instants, with the billiard domain being a $(s+1)$-dimensional cube with the faces parallel to coordinate planes.

This dynamical system is equivalent to the trivial system with constant velocities on a torus and is studied in much detail (see [T] for a survey). However, there are questions still attracting a lot of attention in the literature, such as the question of the coding of trajectories by listing its consecutive collisions with the boundary.

Specifically, to a trajectory one associates an infinite word in alphabet $\mathscr{A} = \{\mathbf{0}, \dots, \mathbf{s}\}$ as follows: each time the trajectory meets a face of the cube parallel to the j^{th} coordinate plane, one writes down $\mathbf{j}$. The resulting infinite word will be called a *symbolic trajectory*. In exceptional cases the trajectory meets more than one face simultaneously, but such cases are not generic and will not be considered.

The resulting symbolic trajectories arise in numerous problems related to number theory, quasicrystals, computer graphics, etc. These trajectories were abundantly studied in the two-dimensional case (where they bear also such names as Sturmian trajectories or Beatty or Wythoff sequences); a sample bibliography can be found in [B, LP, S]. Although multidimensional generalizations are investigated much less, one can find quite a lot of results on those in the papers mentioned.

★ The author was supported by DFG.

The *complexity* of a trajectory is defined as the number of different words of length n occurring in the associated symbolic trajectory considered as a function of n. The problem of the determination of the complexity for rectangular billiards was apparently first studied by M. Morse and G.A. Hedlund in [MH], where it was completely solved for two-dimensional billiards. They have shown that the complexity is independent of the trajectory (provided that the coordinate projection of velocity are rationally incommensurable) and is equal to $n + 1$.

Of course, the most striking fact here is the independence of the complexity of a particular trajectory. This independence persists in the three dimensional case, which was considered by Arnoux, Mauduit, Shiokawa and Tamura [AMST]. They have shown that the complexity of symbolic trajectories is $n^2 + n + 1$, as was conjectured by Rauzy in [R]. The authors made their own conjecture concerning the complexity of the symbolic trajectories in the multidimensional case, based on some quite mysterious assumption of symmetry in s and n. In fact, their formula follows immediately from the independence of the trajectory result (see part 5 of the present paper).

Here we prove their conjecture, giving the general formula for the complexity of symbolic trajectories associated with rectangular billiards in arbitrary dimension.

The method of the solution is as follows. The set of all subwords of length n of a symbolic trajectory we call the n-thesaurus. It is obvious, that if the *velocities* are rationally independent, then the thesaurus does not depend on the initial point of the trajectory (since the system is minimal), that allows us to speak about the thesaurus of the (generic) velocities vector.

First, we write down explicitly the condition for a word of length n to belong to the n-thesaurus of a velocity in terms of consistency of a certain system of linear equations and inequalities. Further we investigate the change of the n-thesaurus when the vector of velocities varies. These changes happen when the inverse velocities are rationally dependent only, that is when a *resonance* occurs. If the resonance is simple, that is there is at most one (up to multiples) vanishing integer combination of inverse velocities, then we show that there is a one-to-one correspondence between words leaving the thesaurus and the words coming into the thesaurus when the resonant value is traversed. That proves that the complexity of the billiard is a well defined function of n only (that is it takes the same value for all nonresonant velocities). To finish, we calculate the n-thesaurus for a special velocity vector, which yields the main result of the paper:

Theorem. *The size of the n-thesaurus of the generic velocities vector (that is such that both its component are independent over* $\mathbb{Q}$ *and their inverses are) is*

$$\sum_{k=0}^{\min(s,n)} k! \binom{s}{k} \binom{n}{k} .$$

1. Basic Constructions

Let $B \subset \mathbb{R}^{s+1}$ be a rectangular area bounded by the hyperplanes $\{x_i = 0\}$, $\{x_i = l_i\}$, $i = 1, \ldots, s$. Without loss of generality, we will take all $l_i = 1$ to restrict ourselves to the unit cube case. The movement of the particle in B is defined as follows: it moves freely with velocity $v = (v_0, \ldots, v_s)$ until it reaches the boundary where it reflects elastically (that means that if the collision point belongs to the face parallel to the i^{th} coordinate plane, then $v_i \mapsto -v_i$). The usual procedure of the first 2^{s+1}-fold covering of B by the torus and then of the covering of the torus by $\mathbb{R}^{s+1}$ leads to the description of the motion of the point as the projection of the free motion in $\mathbb{R}^{s+1}$ with velocity v; the collision instants correspond to the instants of the intersection by the lifted trajectory of the hyperplanes $x_j = n$, $n \in \mathbb{Z}$. Excluding lower dimensional cases we can assume that all $v_j \neq 0$; without loss of generality, we can even take all $v_i > 0$. For any j the instants of intersections of trajectories with the hyperplanes $x_j = n$ form, obviously, an arithmetic progression with the difference $a_j = (v_j)^{-1}$. Therefore the symbolic trajectory, corresponding to a billiard trajectory with the given velocity v can be described as follows: we mark points in $\mathbb{R}$ belonging to the j^{th} progression by $\mathbf{j}$ and then read all the marks in their natural order. We assume that no point is marked simultaneously by more than one number; this is true for almost all trajectories with given velocity vector (trajectories with the given velocity are parameterized by their starting point modulo a shift along the trajectory, which gives the s-dimensional torus as the space of trajectories; trajectories for which the corresponding arithmetic progression have common points form a countable union of $(s-1)$-dimensional tori). Trajectories for which no point is marked by more than one letter (or, equivalently, which never hits $(s-1)$-dimensional faces of the cube) will be called generic.

It is more convenient to work with the vector $a = (a_0, \ldots, a_s)$ of *inverse velocities* or differences of the arithmetic progressions in question. We will say that a word q in the alphabet $\mathscr{A} = \{\mathbf{0}, \ldots, \mathbf{s}\}$ of length n is a-admissible, if there exists a generic trajectory with velocities inverse to a, such that q is a subword of the length n in its symbolic trajectory (it follows immediately that if the differences for a_j's are $\mathbb{Q}$-independent, then q is a subword of the symbolic trajectory for any generic trajectory as the system is minimal). The union of all a-admissible words of length n is called the n-thesaurus for a and is denoted as $\mathscr{T}(a)$.

We will represent the presence of q in $\mathscr{T}(a)$ as some condition on a polyhedron depending on the word and velocities. Introduce the following $(3(s+1)+n)$-dimensional space W with coordinates

$$x_0^-, \ldots, x_s^-; \quad x_1, \ldots, x_n; \quad x_0^+, \ldots, x_s^+;$$
$$a_0, \ldots, a_s . \tag{1.1}$$

The meaning of the coordinates x is the following. To an n-subword of the symbolic trajectory n consequent instances correspond when the particle hits the boundary. The numbers $x_1, \ldots, x_n$ represent just these instants. The number x_j^- (x_j^+, respectively) represents the last moment before x_1 (the first after x_n) when the particle hits a face parallel to the j^{th} coordinate plane.

We will often consider W as the direct sum of its $2(s+1)+n$-dimensional x-part W_x and $(s+1)$-dimensional a-part W_a; the projection of W on its a part will be denoted as p_a.

The conditions of precedence mentioned above are encoded by the following inequalities defining a polyhedral cone $C \subset W$:

$$\max(x_0^-, \ldots, x_s^-) \leqq x_1 \leqq x_2 \leqq \cdots \leqq x_n \leqq \min(x_0^+, \ldots, x_s^+) . \tag{1.2}$$

Now for any word q of length n in the alphabet $\mathscr{A}$ we define the linear space $W(q) \in W$ as follows: let $I_j = \{i_1^j, \ldots, i_{|I_j|}^j\} \subset \{1, \ldots, n\}$ be the subsequence of indices in $\{1, \ldots, n\}$ for which $q_i = j$. Then the linear subspace $W(q)$ is defined by the conditions that the sequences

$$x_j^-, x_{i_1^j}, \dots, x_{i_{|I_j|}^j}, x_j^+$$

form arithmetical progressions with the differences a_j for $j = 0, \dots, s$.

A simple count shows that the dimension of $W(q)$ is $(2s+2)$ independently of q: one can vary the first terms of the arithmetical progressions and their differences arbitrarily.

Further, we define the convex polyhedral cone

$$P(q) = C \cap W(q)\,, \tag{1.3}$$

and its intersection with the fibers of the projection p_a of W on its a-part

$$P(q,a) = P(q) \cap p_a^{-1}(a)\,. \tag{1.4}$$

These polyhedra play in the sequel quite a fundamental role because of the following.

Lemma 1.5. *The word q belongs to the thesaurus $\mathcal{T}(a)$ exactly when the following equivalent conditions hold*:

1. *$P(q)$ has the maximal dimension $2(s+1)$ and a point $(x,a) \in W$ is interior in $P(q)$*;
2. *$P(q,a)$ has the maximal dimension $(s+1)$ in the fiber $p_a^{-1}(a)$.*

Proof. Let the word q be a part of the symbolic trajectory associated to a generic trajectory with the vector of inverse velocities a. Let $\dots t_{-1}, t_0, t_1, \dots$ be the instances of collisions. Then, assuming that the word starts, say, at the first term, one can by setting $x_i = t_i$, $i = 1, \dots, n$ and attaching to x_j^- (x_j^+) the last moment of the appearance of **j** before t_1 (the first moment of occurrence of **j** after t_n) get an interior vector of $P(q)$ as all the inequalities defining $P(q)$ are in fact strict at the point. That means that a small vicinity of the point (x,a) (with $x = (x_0^-, \dots, x_s^-, x_1, \dots, x_n, x_0^+, \dots, x_s^+)$) in $W(q)$ belongs to $P(q)$, and the projection of the vicinity to W_a is open there, that proves 1.

Assertion 2 follows immediately from 1.

Assume 2. It implies that there exists a point $(x,a) \in P(q,a)$ such that all the inequalities defining $P(q)$ are strictly satisfied. Having such a vector x one easily constructs a piece of trajectory with differences a which has the word q as a part of its symbolic trajectory. Extending the trajectory to both sides (which can be done unambiguously) and disturbing a little the arithmetical progressions to avoid multiple points in their union – that always can be done as the fact that the considered point is interior in $P(q,a)$ means that they are different and small distortion do not change their order – gives the desired generic trajectory.

The following statement generalizes Corollary 4 of [LP]:

Corollary 1.6. *If a word q belongs to the thesaurus $\mathcal{T}(a)$, then the reversed word q^{in} also belongs to it.*

Proof. The mapping

$$x_j^- \mapsto -x_j^+, \quad x_i \mapsto -x_{n-i}, \quad x_j^+ \mapsto -x_j^-$$

takes a point in $P(q,a)$ into a point in $P(q^{\text{in}},a)$.

2. Changes of Thesaurus, Flows in Graphs and Linear Programming

Now we are going to study the changes of the thesaurus when a varies in some way. Our goal will be to establish the constancy of the thesaurus size. To prove it we join two arbitrary vectors of generic inverse velocities a_1, a_2 by the segment I and consider its preimage under p_a in W. The intersection of the preimage with any of the cones $P(q)$ is a convex polyhedron $P_I(q)$ again, and Lemma 1.5 implies that the changes of the thesaurus occur exactly in those points of I where some of the polyhedra $P(q,a)$ lose their full dimension. Such points we will call *critical.*

The totality of all polyhedra $P(q)$ (where q are words of length n) is finite, and each of these polyhedra has a finite number of faces, so the set of the critical points in I is finite. Choose one such point a_* and two neighboring points a_i, a_o, between which no further critical point besides a_* occurs. That implies that the thesaurus *before* a_* is that at a_i and the thesaurus *after* a_* is that in a_o. Choose a linear function l on the W_a such that $l(a_i) < l(a_o)$, and lift it to the whole W.

Lemma 2.1. *The word q belongs to $\mathscr{T}(a_i) - \mathscr{T}(a_o)$ if and only if $P_I(q)$ has maximal dimension $s+2$ and l reaches its maximal value over $P_I(q)$ on $P(q,a_*)$ – a face of $P(q)_I$.*

Proof. The fact that $q \in \mathscr{T}(a_i)$ means by Lemma 1.5 that $P(q,a)$ has the dimension $s+1$ for all $a \in I$ between a_i and a_*, thus the dimension of $P_I(q)$ is $s+2$. If for certain a between a_* and a_o the polyhedron $P(q,a)$ were nonempty, then the polyhedron $P_I(q)$ would contain the cone spanned by a point in $P(q,a)$ and $P(q,a_i)$ and for any a' between a_* and a the fiber $P(q,a')$ would have full dimension $s+1$, so that the word q would belong (by Lemma 1.5) to $\mathscr{T}(a')$ and also to $\mathscr{T}(a_o)$. Therefore all the sections $P(q,a)$ are empty for a after a_* and the maximum of l is attained on $P(q,a_*)$.

Inversely, if the maximum of l is attained on $P(q,a_*)$, then all the polyhedra $P(q,a)$ for a after a_* are empty. Further, if the dimension of $P_I(q)$ is $s+2$, then some of the fibers $P(q,a)$ have the dimension $s+1$ and thus all of them before a_*, as $P(q,a_*)$ is nonempty.

To investigate implications of the criticality of a point we introduce a graph and a flow on it. The vertices of the graph correspond to x-coordinates in W and the edges to constraints defining the polyhedra $P(q)$.

Let q be a word in $\mathscr{A}$ of length n. The graph $\Gamma(q)$ has $2(s+1)+n$ vertices $v_o^-, \dots, v_s^-$; $v_1, \dots, v_n$; $v_o^+, \dots, v_s^+$. The (oriented) edges of $\Gamma(q)$ are of two types: first, independent of q, are following: connecting each $v_j^-, j = 0, \dots, s$ to v_1; v_1 to v_2, v_2 to $v_3, \dots, v_{n-1}$ to v_n and v_n to each of $v_j^+, j = 0, \dots, s$. This q-independent part is a tree and will be denoted as Γ; the edges of this tree will be referred to as m-edges. The edges of second type are q specific and connect v_j^- to $v_{i_1^j}$; this latter one to $v_{i_2^j}$ and so on until $v_{i^j_{|I_j|}}$. This last vertex is connected to v_j^+. (Recall that $I_j = (i_1, \dots, i_{|I_j|})$ is the subset of $\{1, \dots, n\}$ consisting of indices i such that $q_i = j$; if the subset is empty v_j^- is connected directly to v_j^+.) We will call them l-edges and will mark them by the corresponding letters of $\mathscr{A}$.

The edges of the graph $\Gamma(q)$ correspond to constraints defining the polyhedron $P(q)$; m-edges corresponding to inequalities defining the cone C and l-edges to equalities defining $W(q)$.

Recall, that a *closed flow* in an oriented graph is a function on its edges such that the Kirchhoff's current law is satisfied: for any vertex the sum of all flows (values of the function) over the *in*-edges is equal to that over all *out*-edges.

Recall also some basic facts from the linear programming theory which will be used in the proof of the following proposition (see, e.g. [FF]):

Fact. *A. Let P be a polyhedron in* $\mathbb{R}^S$ *defined by a system of linear equations and inequalities*

$$\psi_\alpha \geqq 0; \qquad \phi_\beta = 0\,;$$

$\alpha = 1,\ldots,A$; $\beta = 1,\ldots,$ B; *ϕ's and ψ's – linear inhomogeneous functions. Assume that the maximum of a linear function l is attained on a face F of the polyhedron, and y is a relatively interior point of the face. Then there exists a linear combination (whose coefficients are called Lagrange multipliers) of the linear functions constraining the polyhedron (i.e. ψ's and ϕ's), with nonnegative coefficients of ψ's, whose sum with l is a constant. Moreover, the linear combination can be chosen in such a way, that the coefficient for ψ_α is positive exactly when $\psi_\alpha(y) = 0$ (LP duality).*

B. Inversely, if a linear combination of the constraining function with nonnegative coefficients for the linear functions entering inequalities defining P plus l is constant, and all the constraining functions with nonzero coefficients vanish at $y \in P$, then l achieves its maximum over P at y.

Proposition 2.2. *Let a_* be critical and the word q either vanishes from the thesaurus or appears there when a varies through I. Then there exists a non-zero closed flow on $\Gamma(q)$, and if x is a relatively interior point of $P(q,a_*)$, then m-edges in the support of the flow correspond exactly to those inequalities which becomes equalities on x and the flow through any of these m-edges is positive.*

Proof. Choose Lagrange multipliers

$$\begin{aligned} m_j^- &\geqq 0 \quad \text{for} \quad x_1 - x_j^- \geqq 0\,, \\ m_i &\geqq 0 \quad \text{for} \quad x_{i+1} - x_i \geqq 0\,, \\ m_j^+ &\geqq 0 \quad \text{for} \quad x_j^+ - x_n \geqq 0 \end{aligned} \tag{2.3}$$

for inequalities defining cone C.

The constraints, defining the linear subspace $W(q)$ are either

$$x_i - x_j^- - a_j = 0, \quad \text{or} \quad x_i - x_{i'} - a_j = 0, \quad \text{or} \quad x_j^+ - x_i - a_j = 0\,,$$

and we attach Lagrange multipliers $l_j^-, l_{i,i'}$ and l_j^+ to them respectively. Thus to each constraint and, consequently, to each edge of $\Gamma(q)$ a Lagrange multiplier is associated.

We will need one more set of multipliers for the constraints forcing a to belong to the line through a_i, a_o. (Notice that there is no need to introduce multipliers for $a_j \geqq 0$ as the inverse velocities are positive by assumption.)

So, according to the Fact of the linear programming theory (as stated above), if a_* is critical, and (x,a_*) is an interior point of $P(q,a_*)$, we can choose a set of coefficients (Lagrange multipliers), such that the resulting linear combination plus l is constant.

Consider now these multipliers as defining a flow Φ on $\Gamma(q)$. Indeed, to each of them corresponds a unique edge in the graph, so that the x-part of the corresponding function is the difference of the x's at the ends of the edge. The equality of the linear combination to $-l$ plus a constant means exactly that the defined flow is closed: the coefficient for any of the x's is the algebraic sum of flows to the corresponding vertex, and l does not depend on the x's (in other words, the Lagrange multipliers define a closed 1-cocycle on $\Gamma(q)$).

The statement about m-edges in the support of the flow is tantamount to the LP duality.

A cycle in a graph is called simple if it passes through any edge at most once. A simple (non-oriented) cycle in the oriented graph $\Gamma(q)$ is said to be *subordinated to the flow* Φ, if the only m-edges it contains are from the support of Φ and it traverses those m-edges in accordance with their orientation. To any simple cycle a closed flow corresponds; it takes a constant (positive) value on all edges of the cycle. Such closed flows will also be called simple and subordinated if the underlying cycle is.

Simple subordinated cycles are important as they generate bounded resonances at a_*. A *resonance* is a vanishing linear integer combination of a_j's. A resonance will be called bounded if the sum of absolute values of its coefficients does not exceed the number of edges in $\Gamma(q)$.

Lemma 2.4. *Let a_* be critical; x be an interior point in $P(q, a_*)$ and Φ be a closed flow defined by the Lagrange multipliers at x. Then to any simple cycle subordinated to Φ a bounded resonance at a_* corresponds.*

Proof. Any edge of the cycle corresponds either to equality $x_i - x_{i'} - a_j = 0$, or to equality of one of the following formats:

$$x_j^- - x_1 = 0 \quad \text{or} \quad x_{i+1} - x_i = 0 \quad \text{or} \quad x_j^+ - x_n = 0$$

(the m-edges are in the cycle if the corresponding Lagrange multiplier is positive only, which makes them equalities). Summing up all of them we obtain a bounded resonance at a_*.

Lemma 2.5. *For any closed flow Φ subordinated cycles exist. Moreover, any closed flow can be decomposed into a positive linear combination of simple subordinated flows.*

Proof. Reverse the orientations of all l-edges with negative flow, so that the flow through any edge is nonnegative. A subordinated cycle can be then found by the following algorithm: pick any edge in the support of the flow and go along the arrow. In the reached vertex choose a new adjoining edge along which the movement according to its orientation is possible – such an edge always exists since the flow is closed. Iterating we will reach a vertex already seen at a stage, thus getting a subordinate cycle. To prove the decomposition part of the lemma, define the closed flow on this chosen cycle by assigning to each edge in it the minimal flow of the edges gone through. Subtracting the resulting closed flow from the initial one we will obtain a flow with a smaller support. Iterating finishes the proof.

3. Simple Resonances

We say that the critical point a_* is *simple* if all bounded resonances at the point are integer multiples of a single resonance $n \cdot a_* = 0$. The condition of simplicity strongly restricts the structure of possible simple cycles subordinated to the flows associated with the critical point a_*. In fact one can prove that the support of such a flow is necessarily a simple cycle or contains l-edges of only two types. We will prove here only a weaker result needed in what follows.

Proposition 3.1. *Let $q \in \mathcal{T}(a_i) - \mathcal{T}(a_o)$; $a_* \in I$ be a simple critical point on the segment I; x be an interior point of $P(q, a_*)$ and Φ be the associated closed flow. Then at most one of the edges $x_j^- \to x_1$ and at most one of the edges $x_n \to x_j^+$ belong to the support of Φ.*

Proof. Consider a cycle subordinated to the flow Φ. For any mark $\mathbf{j}$ we define a $\mathbf{j}$-segment in the cycle as a sequence of l-edges with this mark bounded by edges of other (necessarily not l-) types.

First, we will prove, that for any simple cycle subordinated to the associated flow and for any mark $\mathbf{j}$ there is at most one $\mathbf{j}$-segment in the cycle. Indeed, let s_1, s_2 be two $\mathbf{j}$-segments separated by some pieces of the cycle c_1, c_2, so that the whole cycle has the form

$$-s_1 - c_1 - s_2 - c_2 - s_1 - \ .$$

We can now form two new cycles joining the ends of c_1 and c_2 by $\mathbf{j}$-segments (Fig. 1). These two cycles are clearly simple and subordinate as the m-edges remain unchanged.

An easy check shows that the closed flow whose support is the initial simple cycle is now the sum of thus constructed closed flows. Decomposing if necessary these flows further we arrive at a stage of the situation when each of the new cycles has at most one $\mathbf{j}$ segment. If the initial cycle had more than one $\mathbf{j}$ segment, then among these cycles exist both cycles going through $\mathbf{j}$ edges in positive and negative directions.

Each of the cycles generates a nontrivial (as $n_j \neq 0$) resonance, a multiple of n by assumption. The fact that both positive and negative multiples of n occur means that both l and $-l$ can be represented as linear combinations of the constraining functions with nonnegative m's (as the constructed flows are subordinated), and thus both l and $-l$ achieve their maxima on P_q at the point (x, a_*) (part B of the linear programming Fact). It follows that $P_l(q) = P(q, a_*)$ in contradiction with the assumption, that $P_l(q)$ has dimension $s + 2$.

Thus each of the simple subordinated cycles has at most one $\mathbf{j}$ segment, and the direction of traversing them is the same for all cycles.

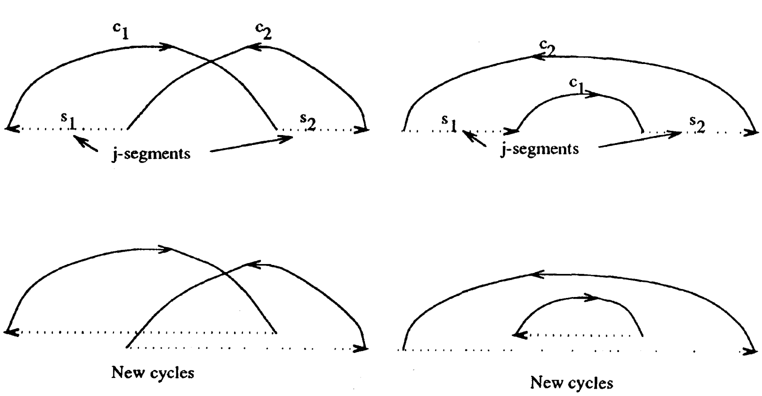

Fig. 1. Simplification of simple subordinated cycles.

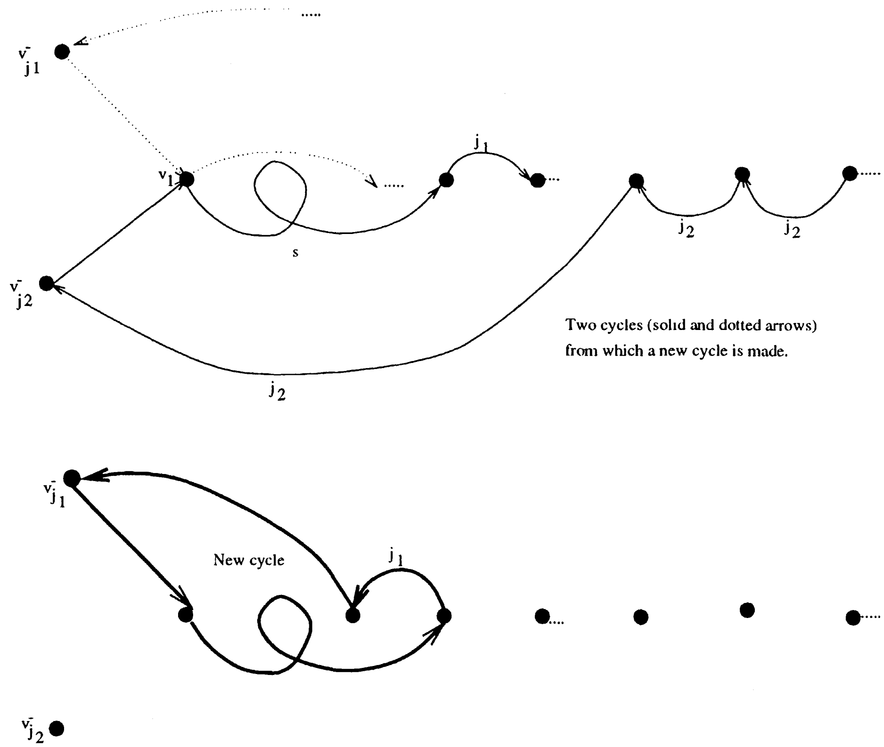

Fig. 2.

Assume now that the support of the associated flow contains two (m-)edges $v_{j_1}^- \to v_1$ and $v_{j_2}^- \to v_1$. Let γ_1, γ_2 be two simple subordinated cycles containing these edges correspondingly. Let s be the segment of the cycle γ_2 starting in v_1 and ending at the beginning of the first $\mathbf{j}_1$ edge. Then cutting short along $\mathbf{j}_1$-edges to $v_{j_1}^-$ and from there to v_1 we obtain a new subordinate cycle (Fig. 2). This cycle contains no $\mathbf{j}_2$ edges as they obviously cannot belong to s, but contain more $\mathbf{j}_1$ edges than γ_2. Therefore the resonance defined by the cycle cannot be a multiple of n – a contradiction. Similar reasoning gives x_j^+ part.

Corollary 3.2. *For any interior point x of $P(q, a_*)$ at most one of inequalities $x_j^- \leqq x_1 (x_n \leqq x_j^+)$ becomes equality.*

Proof. It follows immediately from Propositions 2.2 and 3.1.

4. In-Out Correspondence

Now we will prove the crucial result: if the critical point of the segment I is simple (that is satisfies at most one resonance modulo natural multiples) then the number of words disappearing from the thesaurus equals the number of words appearing there. To do it we construct the *in-out correspondence* as follows.

Let x be a vector in W_x, and q a word of length n. We can consider these data as defining a marking of $\mathbb{R}$.

A *marking of real line* by elements of $\mathscr{A}$ is a mapping from $\mathbb{R}$ to the set of subsets of $\mathscr{A}$. We will consider only finite markings, that is such that the number of points mapped into a nonempty subset of $\mathscr{A}$ is finite. The size of the subset to which a point is sent is called the multiplicity of the point; the sum of all multiplicities is the multiplicity of the marking.

The marking defined by $x \in W_x$ and word q is constructed as follows. We assume that points $x_j^{\pm}$ are marked by $\mathbf{j}$, and x_i is marked by q_i. If a point $\xi \in \mathbb{R}$ is marked by several letters (that is some of the coordinates $x_j^{\pm}$ or x_i coincide), then we assemble all these letters together and send the point ξ into the corresponding subset. All points which are not equal to any of coordinates $x_j^{\pm}, x_i$ are sent to an empty subset of $\mathscr{A}$. The multiplicity of this marking is $2(s+1)+n$.

We will call a marking *good in the middle* if there is a segment $[\xi_-, \xi_+] \subset \mathbb{R}$, such that

1. The sum of multiplicities of the points in each of the half lines $(-\infty, \xi_-)$ and (ξ_+, ∞) is $(s+1)$, and
2. The marking of $[\xi_-, \xi_+]$ is simple, that is each point in the segment has multiplicity at most one.

A good in the middle marking defines a word of length n: one just reads these middle n letters in their natural order. This word we denote by $Q(x;q)$; it coincides with q if x is an interior point of C.

When a line in W is given (with some parameter l), the coordinates x_j^-, x_i, x_j^+ become a linear function of l and a movement along the line can be considered as an evolution of these points steadily moving in $\mathbb{R}$. We will be using this convenient terminology throughout this section.

Let a_* be our simple critical point, the word q belongs to $\mathscr{T}(a_i) - \mathscr{T}(a_o)$ and $P(q, a_*)$ be the face of the polyhedron $P_l(q)$, where l attains its maximum. Let x_* be a relatively interior point of the polyhedron. Choose a line L through (x_*, a_*) which projects onto I under p_a and such that one of the halflines on which (x_*, a_*) divides it belongs to $P_l(q)$. We will call this halfline *the a_i-side*, and the other one – *the a_o-side*. It is clear that one can choose the resonance at a_* as a parameter on L, so that $l(a_*) = 0$ and l is nonpositive on $P_l(q)$.

Lemma 4.1. *Points of the line L close enough to (x_*, a_*) define a good in the middle marking of* $\mathbb{R}$.

Proof. The statement is trivial for points on the a_i-side of the line. If that is not the case for the a_o-side, then a pair of the marked points coincide identically along the line L. The pair cannot include any of x_i points, as each of them is separated from the rest on the a_i side. Equally, it cannot be a pair $x_{j_1}^-, x_{j_2}^+$. So we assume that $x_{j_1}^- = x_{j_2}^+$ identically on the line. To affect the middle part of the marking on the a_0-side, the pair should move through x_1 when $a = a_*$. That yields $x_{j_1}^- = x_{j_2}^- = x_1$ at x – a relatively interior point of $P(q, a)$, and thus

$$x_1 - x_{j_1}^- = 0; \quad x_1 - x_{j_2}^- = 0$$

at $x_* \in P(q, a_*)$. This contradicts Corollary 3.2.

The x^+ case is similar.

Thus we have a good in the middle marking defined by (x_*, a_*) on the a_o-side of the line and form the word $q' = Q(q,x)$. This correspondence $q \mapsto q'$ is called an *in-out* one.

To use it we have to prove first that it is unambiguous.

Proposition 4.2. *The in-out correspondence is defined unambiguously, that is, does not depend on the choice of the interior point $x_* \in P(q,a_*)$ and of line through x_*.*

Proof. Let L, L' be two lines through a relatively interior point x of $P(q,a)$ which p_a maps onto I, such that the a_i-sides of both of them belong to $P_I(q)$. Assume that the markings on their a_o-sides define *different* words q', q''. Let V be the two-dimensional plane spanned by L, L'. This plane is fibered by the level lines of p_a and both the lines L, L' are transversal to the fibration. These lines cut segments in each fiber of p_a. The difference of the words defined by the middle parts of the corresponding markings implies that inside the segment a couple of these middle points coincide. The finiteness of the number of such couples yields existence of a line L'' in V between L and L' which is projected onto I and along which a pair of x-coordinates coincide identically. The pair cannot contain any of the x_i's, as it would contradict the assumption that the a_i-sides of both L, L' are in $P_I(q)$, where these points are distinct from the other ones. Similarly, the pair cannot be $x^-_{j_1}, x^+_{j_2}$. So, the pair is of the type $x^-_{j_1}, x^-_{j_2}$ or $x^+_{j_1}, x^+_{j_2}$. The fact that the points in the pair belong to the middle n points on the a_o-side of L'' implies that this coinciding pair passes through the point x_1 or x_n respectively in the course of the movement along the line L'' from a_i to the a_o-side. That would mean that at x_* $x_1 - x^-_{j_1} = 0$; $x_1 - x^-_{j_2} = 0$, or $x_n - x^+_{j_1} = 0$; $x_n - x^+_{j_2} = 0$, in contradiction with Corollary 3.2.

Let now x, x' be two different relatively interior points in $P(q,a)$. Choose two lines $L \ni x$, $L' \ni x'$ with their a_i parts in $P_I(q)$ lying in a two-dimensional plane V and projecting onto I. The assumption that the words defined by their a_o parts are different means again that between L and L' in V there is a line along which some pair of x coordinates coincide. Reasonings as above prove it impossible.

Now we can give a combinatorial description of in-out correspondence. Take an interior point x of $P(q,a_*)$. If $x^-_j = x_1$ at x, then we add the letter $\mathbf{j}$ at the beginning of q, if $x^+_j = x_n$, then the letter $\mathbf{j}$ is appended to q. These operations are unambiguous by Corollary 3.2. *A block* in the thus extended word is a maximal subword of consequent letters such that the corresponding coordinates coincide at x. For example, if $q = \mathbf{001201}$ and $x^-_1 = x_1 < x_2 = x_3 < x_4 < x_5 = x_6 < \ldots$, then the extended word is $(\mathbf{10})(\mathbf{01})(\mathbf{2})(\mathbf{01})$ and blocks are bracketed.

Proposition 4.3. *The word q' is obtained from q by inversing all blocks and erasing letters at the beginning/end of the word if they were appended there.*

Proof is obvious. A line through x with the a_i-side in $P_I(q)$ defines a family of markings with marked points moving steadily in $\mathbb{R}$. If at the critical moment some points are distinct, then their order in $\mathbb{R}$ remains the same some time around the instant. If some number of points crash together, then their order after the collision reverses.

In the example above, the word q' will be then given by reversing blocks: $(\mathbf{01})(\mathbf{10})(\mathbf{2})(\mathbf{10})$ and by erasing the first letter: $q' = \mathbf{110210}$. (Notice, that this transformation corresponds to the resonance $a_0 = a_1$.)

One further thing we need to show is that if q leaves the thesaurus during the movement from a_i to a_o, then q' appears there.

Proposition 4.4. *If q, q' are in in-out correspondence, and $q \in \mathscr{T}(a_i) - \mathscr{T}(a_o)$, then $q' \in \mathscr{T}(a_o) - \mathscr{T}(a_i)$.*

Proof. By the construction of q', the polyhedron $P(q')$ has full dimension since the fact that the marking defined by a point (x', a') on the a_o-side of the line L passing through (x_*, a_*) is good in the middle means that all inequalities constraining $P(q')$ are strict at the point corresponding to the marking. So we know that $q' \in \mathscr{T}(a_o)$ and we need to prove only that it does not belong to $\mathscr{T}(a_i)$.

Let γ be a simple cycle in the graph $\Gamma(q)$ subordinated to the flow Φ defined at $x \in P(q, a_*)$. The constraints corresponding to the edges in γ are all equalities at x and the sum of the corresponding linear functions is a negative multiple of l on $P_l(q)$.

One can consider this cycle also as a cyclical chain of letters of the extended word used in the combinatorial description of in-out correspondence. It is obvious that the m-edges join the letters within a block and that l-edges join (equal) letters in different blocks. The transformation that reverses blocks sends this cyclical chain into a cyclical chain with the same properties in the extended word q'. This cyclical chain defines in its turn a simple cycle in the graph $\Gamma(q')$: l-edges continue to connect the letters as in each block all letters are different (otherwise the corresponding points were unable to collide) and an l_j-edge continues to connect consequent letters $\mathbf{j}$ even if they are swapped inside their blocks by transformation; the m-edges in the chain connect neighboring letters in blocks and they remain such. We will denote this new cycle in $\Gamma(q')$ as γ'.

The point $x \in P(q, a_*)$ defines a trajectory (degenerate) which, by construction of q', can be disturbed slightly to produce a trajectory with q' in the thesaurus. Let us consider the corresponding point $x' \in P(q', a_*)$. At this point all the inequalities corresponding to the m-edges entering the cycle γ' become equalities, as they join points inside blocks which have equal ξ-coordinates. Reverse the orientation in γ' to restore the original directions in m edges. Then the sum of all the linear functions in constraints corresponding to edges in γ' is a *positive* multiple of l, because the collections of l-edges in γ and γ' coincide, but the edges have reverted orientations. It follows (by Fact B of the linear programming theory) that the functional l reaches its *minimum* at x' and thus $\mathscr{T}(a_i)$ is empty.

The last step is to prove the reflectivity of the in-out correspondence.

Proposition 4.5. *If $q \mapsto q'$, then $q' \mapsto q$.*

Proof. Let $x \in P(q, a_*)$ be an interior point. It defines a trajectory which can be slightly disturbed to produce a generic trajectory with q' in its thesaurus. That means that we can associate to this trajectory a point on the boundary of $P_l(q')$; in $P(q', a_*)$ to be specific. Altogether this gives a mapping from the interior of $P(q, a_*)$ into $P(q', a_*)$. It is immediate that the mapping is an affine embedding. If the dimensions of these polyhedra $P(q, a_*)$ coincide, then an interior point $x \in P(q, a_*)$ corresponds to an interior point $x' \in P(q, a_*)$; the line $L \ni x$ used to define q' corresponds to a line $L' \ni x'$ with the a_o-side in $P_l(q')$ and with the good in the middle marking on its a_i-side defining the word q. This, clearly, implies the reflectivity.

It is enough to prove that the dimension of $P(q', a_*)$ is not greater than that of $P(q, a_*)$, as the embedding of the latter polyhedron into the former one yields that the dimensions are equal.

The dimension of polyhedron $P(q,a_*)$ is less than the generic one $(s+1)$ because some of the inequalities constraining the polyhedron become identical equalities and thus shifts of some of the arithmetical progressions forming the trajectory become dependent. These equalities correspond to m-edges entering the closed flow Φ constructed in Proposition 2.2. Each of these m-edges can be extended to a simple cycle γ in $\Gamma(q)$. Reasoning as in the proof of Proposition 4.4, one can construct a simple cycle γ' in $\Gamma(q')$. Under this operation m-edges go to m-edges and the letters which mark the vertices they connect remain the same. As the sum of constraints corresponding to the edges of γ' vanishes to a_* (as a multiple of the resonance), all the inequalities among the constraints are in fact equalities. It follows that the dependence of shifts of arithmetical progressions present on $P(q,a_*)$ persists on $P(q',a_*)$ and the claim follows.

Corollary 4.6. *The size of the thesaurus is constant in two endpoints a_ι, a_o of a segment containing a single simple resonance point a_*.*

5. Proof of the Main Theorem

Theorem. *The size of the thesaurus for a generic velocity vector is given by the formula*

$$\sum_{k=0}^{\min n,s} k! \binom{s}{k} \binom{n}{k} .$$

Proof. Let a_1, a_2 be inverse velocities. Choose a piecewise linear path joining them in the space of inverse velocities, such that endpoints of each segment of the path were linearly independent over $\mathbb{Q}$ and each point of the path belonged to at most one bounded resonance hyperplane $(n, \cdot\,) = 0$. Such a path can be chosen as the number of bounded resonances is finite and the points where multiple resonances occur have codimension 2. Using Corollary 4.6 we see that the size of the thesaurus is constant in the endpoints of the segments forming the path and thus the sizes of thesaurus in a_1 and a_2 are equal. To find it we choose a special vector of inverse velocities. Namely, let $a_0 = 1$ and $a_1, \ldots, a_s$ be arbitrary numbers larger than n, such that the whole tuple $a_0, \ldots, a_s$ is $\mathbb{Q}$-independent. Then the thesaurus can be easily described: a word q belongs to the thesaurus exactly when it contains at most one letter $\mathbf{j}$ with $j \geqq 1$. The number of such words can be calculated immediately. Indeed, each of the words is specified by the number k of letters $\mathbf{j}$ with $j \geqq 1$; by the positions in the word of letters of these letters $\left(\binom{n}{k}\text{ possibilities}\right)$; by the set of letters $\mathbf{j}$ with $j \geqq 1$ used $\left(\binom{s}{k}\text{ possibilities}\right)$ and by one of $k!$ variants of their allocation there. Summing it all up one gets the stated answer.

6. Concluding Remarks

6.1. As a corollary of the presented result we get the remarkable symmetry noted in [AMST]: the size of the n-thesaurus in $\mathbb{R}^{s+1}$ is a function symmetric in n and s. Actually, as it was shown in [AMST] this property almost characterizes this function and thus it would be very interesting to have a construction directly proving this symmetry. I believe that the size of the n-thesaurus can be described in terms of the facet combinatorics of an appropriately chosen polyhedron and that $n \leftrightarrow s$ symmetry would follow from Dehn–Sommerville relations. Such an approach would give a much clearer insight into the combinatorics of symbolic trajectories, but so far I do not know how to realize it.

6.2. Lemma 1.5 provides in principle an algorithmic method to define whether a word is a piece of symbolic trajectory of a billiard. One has to check the compatibility of a system of linear inequalities, which can be effectively done.

6.3. Using the geometric approach of [AMST] one can generate a subdivision of the s-dimensional torus parameterizing the billiard trajectories with given velocity into the convex polyhedra corresponding to different words of the thesaurus of size n. The volumes of these polyhedra closely relate to the reccurrence function of symbolic trajectories, that is the size of the word in which all words of the n-thesaurus appear. The moving of the inverse velocity through a resonance results in the degeneration of some of these polyhedra, and thus enables one to describe the asymptotics of the minimal of their volumes. That gives an approach to the investigation of the reccurrence function. I hope to return to the question in a separate paper.

References

[AMST] Arnoux, P., Mauduit, C., Shiokawa, I., Tamura, J.-I.: Complexity of sequences defined by billiard in the cube. Bull. Soc. Math. France **122**, 1–12 (1994)

[B] Bruckstein, A. M.: Self-similarity properties of digitized straight lines. In: Vision geometry, Proc. AMS Spec. Sess., 851st Meet., Hoboken/NJ (USA) 1989, Contemp. Math. **119**, 1–20 (1991)

[FF] Ford, L.R., Fulkerson, D.R.: Flows in Networks. Princeton, NJ: Princeton Univ. Press, 1962

[LP] Lunnon, W.F., Pleasants, P.A.B.: Characterization of two-distance sequences. J. Aust. Math. Soc., Ser. A **53**, No. 2, 198–218 (1992)

[MH] Morse, M., Hedlund, G.A.: Symbolic dynamics II. Sturmian trajectories. Am. J. Math. **62**, 1–42 (1940)

[R] Rauzy, G.: Mots infinis et arithmetique. In: Automata on Infinite Words, Lect. Notes in Comp. Sciences **192**, Berlin, Heidelberg, New York: Springer (1985) pp. 165–171

[S] Stolarsky, K.B.: Beatty sequences, continuous fractions and certain shift operators. Can. Math. Bull. **19**, 473–482 (1976)

[T] Tabachnikov, S.: Billiards. Preprint (1994)

Communicated by Ya.G. Sinai