

Consider a billiard in a polygon $Q \subset \mathbf{R}^2$ having all angles commensurate with π . For the majority of initial directions, density of every infinite semitrajectory in configuration space is proved. Also proved is the typicality of polygons for which some billiard trajectory is dense in phase space.

Introduction

A billiard in a polygon Q in the Euclidean plane \mathbf{R}^2 is a dynamical system produced by the frictionless motion of a point-sphere inside Q with elastic reflection from the boundary ∂Q of the polygon; the velocity of the sphere may be taken to be unity. This motion is unbounded in time, so long as the sphere does not run into a vertex of the polygon. In the contrary case, the motion is defined in time (in the positive or negative sense) up to collision with a vertex. The first of these cases will be termed the general case; the second will be termed the exceptional case.

The phase space $M = M(Q)$ of this dynamical system is obtained from the direct product $Q \times S^1$, where S^1 is the circle of unit velocities, by identifying pairs of the forms (q, v) , (q, v') for $q \in \partial Q$, v and $v' \in S^1$ and $v - v' = 2(n, v)n$; here, $n = n(q)$ is the unit exterior normal to ∂Q at the point q . Consider the case that the velocity vector makes an angle φ with some chosen direction e . Then, for each point q in a side AB of the polygon the law for the above-described identification takes the form

$$(q, \varphi) \sim (q, \varphi'), \quad \varphi' = 2\alpha - \varphi, \tag{1}$$

where α is the angle between AB and the direction e .

We denote by $\{T_t\} = \{T_t^Q\}$ the phase flow of the system. The transformations T_t are defined for all t only on the subset $M' \subset M$ of the elements $(q, v) \in M$ whose carriers q never meet vertices of the polygon Q . For the remaining elements of M the transformations T_t are defined not for all t but just up to collision of the carrier with a vertex. A volume element $dm = dq \cdot d\varphi$ in $Q \times S^1$ induces on M a finite measure m invariant under the flow $\{T_t\}$; note that $m(M \setminus M') = 0$.

We denote by $Q(\varphi)$ the polygon $Q \times \{v(\varphi)\} \subset Q \times S^1$. Trajectories of the flow $\{T_t\}$ which come from points $(q, \varphi) \in Q(\varphi)$ are line segments in $Q(\varphi)$ which form the angle φ with a chosen direction e . After reflection from a side AB these trajectories "cut over" into the polygon $Q(\varphi')$, where φ' is defined by formula (1). If one views the polygons $Q(\varphi)$ and $Q(\varphi')$ as lying in a plane, with $Q(\varphi')$ obtained from $Q(\varphi)$ by reflection with respect to the side AB (we write $Q(\varphi') = S_{AB}Q(\varphi)$), then the pieces of trajectories which lie in $Q(\varphi) \cup Q(\varphi')$ comprise not a broken line but rather a straight-line segment. The identification described above amounts to gluing $Q(\varphi)$ to $Q(\varphi')$ along side AB and defines a natural smooth structure of class C^∞ on the set $\text{Int}(Q(\varphi) \cup Q(\varphi'))$. By successive reflections with respect to sides of the polygons $Q(\varphi)$, $Q(\varphi')$, ... one may "straighten out" any finite part of a trajectory of the flow $\{T_t\}$, i.e., one may correspond to this part a segment in the plane.

Finally, we may note that the direct product topology in $Q \times S^1$ induces a topology in M , and the function $T_t(x)$ of the variables t and $x = (q, v) \in M$ is continuous in t and x at each point of its region of definition.

M. V. Lomonosov Moscow State University; Central Economicomathematical Institute, Academy of Sciences of the USSR. Translated from *Matematicheskie Zametki*, Vol. 18, No. 2, pp. 291-300, August, 1975. Original article submitted September 26, 1974.

©1976 Plenum Publishing Corporation, 227 West 17th Street, New York, N.Y. 10011. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission of the publisher. A copy of this article is available from the publisher for \$15.00.

§ 1. Invariant Manifolds of a Billiard

In this section and the next we shall assume that all angles of the polygon Q are commensurable with π . We shall fix a direction e for reckoning angles, by taking e to lie along one of the sides of the polygon. We may then write down angles to e formed by the remaining sides in the form $\alpha_r = \pi m_r / 2n_r$, $r = 1, 2, \dots$. Let $N = N(Q)$ be the least common denominator of the fractions m_r/n_r , $r = 1, 2, \dots$; then from (1) it follows that the function

$$F(x) = F(q, v) = F(\varphi) = |\varphi| \pmod{\pi/N} \quad (2)$$

is well-defined on M and is invariant under the flow $\{T_t\}$. Thus, for each number c , $0 \leq c \leq \pi/(2N)$, there is a $\{T_t\}$ -invariant subset $M_c = M_c(Q) = \{x : F(x) = c\}$. Suppose $c \neq 0, \pi/(2N)$. Then, as is evident from (2), the set M_c may be obtained as the union of $4N$ nonintersecting polygons $Q(\varphi_s^\pm)$, where

$$\varphi_s^\pm = \pm c + s\pi/N, \quad s = 0, 1, \dots, 2N - 1, \quad (3)$$

by pairwise identification of their sides in accord with formula (1).

PROPOSITION 1. For $c \neq 0, \pi/(2N)$ the set M_c with the topology induced from M is homeomorphic with a two-dimensional manifold. In fact, one may introduce on M_c a smooth structure of class C^∞ which coincides with the structure described in the introduction except at the vertices of the finite set (3) of polygons. The genus p of the manifold $M_c = M_c(Q)$ is determined by the shape of the polygon Q (i.e., p does not depend on c).

Proof. First of all, it is necessary to verify the existence, at those points of M_c which are identified with vertices of the polygons $Q(\varphi_s^\pm)$, of neighborhoods which are homeomorphic with the disk and which possess a smooth structure which agrees with the smooth structure at the remaining points. Let Q_0 be one of the polygons and let B, A, D be three consecutive vertices of Q_0 . Let $\alpha = \pi m/n$ be the size of the angle BAD where m and n are relatively prime. We shall describe a neighborhood of that point of M_c which is obtained from A ; for this purpose we shall use the viewpoint whereby, with the aid of reflections, polygons are glued together (i.e., their sides are identified). The polygon Q_1 from the set (3) is joined to Q_0 along the side AB ; this polygon Q_1 may be viewed as the image of Q_0 under the reflection S_{AB} with respect to the side $AB: Q_1 = S_{AB}Q_0$. The polygon $Q_2 = S_{AC}Q_1$ is joined to Q_1 along the "free" side AC of Q_1 ; etc. It is easily seen that the polygon Q_{2n-1} , which joins to Q_{2n-2} along some other side emanating from the vertex A , must also be joined to the polygon Q_0 along the side AD . Thus, a neighborhood U_A of the point A in M_c is glued from $2n$ sectors having angle $\pi m/n$ and constitutes an m -sheeted surface with branch-point A . If we regard the k -th sector as lying in the complex plane

$$R_k = \{z \in \mathbb{C} : |z| < \varepsilon, [(k-1)m/n] \cdot \pi \leq \arg z \leq (km/n) \cdot \pi\},$$

we can define a homeomorphism of U_A onto the circle by using the transformation $z \mapsto w = \sqrt[n]{z}$. Thus, on U_A there is defined a chart which smoothly matches the charts on the sets $\text{Int } Q(\omega)$ and $\text{Int } (Q(\omega) \cup Q(\omega'))$, so that on M_c there is defined a smooth C^∞ manifold structure. From formula (1) and the choice of the number $N = N(Q)$ it follows that the manifold M_c is connected and that the gluing of the polygons of the set (3) does not depend on the value of the constant c ; consequently, the genus p of the manifold $M_c = M_c(Q)$ does not depend on c : $p(M_c(Q)) = p(Q)$. Proposition 1 is proved.

Let us now describe the topology of M_c when $c = 0$ or $c = \pi/(2N)$. The invariant subset M_0 is glued out of $2N$ polygons $Q(\omega_k)$, where $\omega_k = k\pi/N$, $0 \leq k < 2N$. From formula (1) it follows that if a side AB of the polygon $Q(\omega_k)$ forms an angle $\alpha = \omega_k$ with the chosen direction, then its points are not identified with any other points: $\varphi' = 2\alpha - \varphi = 2\omega_k - \varphi_k = \varphi_k = \varphi$. Such sides of the polygons $Q(\omega_k)$ will be called "boundaries." Arguments in the proof of Proposition 1 show that M_0 is a manifold with boundary, with the boundary of M_0 consisting of the points of the "boundary" sides. (Suppose a vertex A of an angle of measure $\pi m/n$, where m and n are relatively prime, belongs to one of the "boundary" sides. Then, when n is odd a neighborhood of A is glued from n sectors with angles $\pi m/n$, and when n is even a neighborhood of A is glued from $n/2$ such sectors. In both cases the neighborhood is homeomorphic to a semidisk.) Analogous considerations show that the invariant subset $M_{\pi/(2N)}$ is likewise homeomorphic to a two-dimensional manifold with boundary. These two manifolds can be either orientable or nonorientable. (For example, when Q is an equilateral triangle, both manifolds are Möbius bands.)

Since in every phase space the sets M_0 and $M_{\pi/(2N)}$ are of measure zero, we shall henceforth consider only the closed manifolds M_c described in Proposition 1, where $0 < c < \pi/(2N)$.

Definition 1. Points of the manifold M_c which are obtained by identifying vertices of angles of measure $\pi m/n$, where $m > 1$, will be called branch points of order m .

Let us now describe the flow $\{T_t^c\} = \{T_t^c|_{M_c}\}$ induced on the manifold M_c . It is clear that in the neighborhood of each point of M_c the flow $\{T_t^c\}$ either is a local flow of parallel displacements or extends by continuity to such a flow (in neighborhoods of vertices of angles π/n). If $A \in M_c$ is a branch point of order m , then on each sheet of an m -sheeted z -neighborhood of the point A the flow $\{T_t^c\}$ likewise is a flow of displacements (in one and the same direction); moreover, on each sheet precisely one trajectory ends at A and precisely one trajectory begins at A : altogether, m "entering" and m "exiting" singular trajectories. The vector field V^c of the flow $\{T_t^c\}$ has in the coordinates z the form $V^c(z) = e^{i\varphi}$, where $\varphi = \varphi(c)$ is a constant; in the coordinates $w = \sqrt[m]{z}$ this field is given by the formula $V^c(w) = w^{1-m} e^{i\varphi}/m$ (here $z \neq 0$ and $w \neq 0$).

The flow $\{T_t^c\}$ preserves the finite measure m_c which is induced on M_c and which coincides with Lebesgue measure on the polygons $Q(\varphi_s^\pm)$. In a neighborhood of a branch point of order m the measure m_c likewise coincides with Lebesgue measure in local coordinates z , while in local coordinates w , m_c is given by the density $m|w|^{m-1}$.

PROPOSITION 2. On the manifold M_c there is a smooth transformation of time which vanishes only at branch points and which transforms the flow $\{T_t^c\}$ into a smooth flow $\{\tilde{T}_t^c\}$ having the following properties:

(K0) The set $\Omega(\{\tilde{T}_t^c\})$ of nondiffuse points of the flow $\{\tilde{T}_t^c\}$ coincides with M_c .

(K1) The fixed points of the flow $\{\tilde{T}_t^c\}$ are precisely the branch points of the manifold M_c .

(K2) A branch point of order m is a multisaddle point of index $1-m$ for the flow $\{\tilde{T}_t^c\}$; in other words, the phase portrait of the flow in a neighborhood of this point consists of $2m$ hyperbolic sectors separated by m entering and m exiting separatrices.

Proof. Multiply the vector field V^c of the flow $\{T_t^c\}$ by a smooth function f , where f is defined in a small neighborhood of each branch point by the formula $f(w) = |w|^m$, m being the order of the branch point, and is distinct from 0 at all other points. We obtain a vector field \bar{V}^c which satisfies a Lipschitz condition. The flow $\{\bar{T}_t^c\}$ produced by this field preserves the measure \bar{m}_c , where $d\bar{m}_c(x) = dm_c(x)/f(x)$. In neighborhoods of branch points the measure \bar{m}_c is given by the density $m|w|^{-1}$; consequently, $\bar{m}_c(M_c) < \infty$ and therefore $\Omega(\{\bar{T}_t^c\}) = M_c$. If we multiply the vector field \bar{V}^c by a smooth function which vanishes only at branch points and decreases sufficiently rapidly in neighborhoods of these points, we obtain a smooth field \tilde{V}^c which generates a smooth flow $\{\tilde{T}_t^c\}$. Clearly, $\Omega(\{\tilde{T}_t^c\}) = \Omega(\{\bar{T}_t^c\})$ and therefore $\{\tilde{T}_t^c\}$ possesses property (K0). Assertions (K1) and (K2) follow directly from the properties of the time transformation and from the description of the structure of the flow $\{T_t^c\}$ given before the statement of the present proposition.

If all angles of Q have the form π/n then the manifold M_c has no branch points and $\{T_t^c\}$ is an everywhere-defined smooth flow without singular points. Hence, M_c is a torus and $p(M_c) = p(Q) = 1$. This case is realized only for billiards in rectangles and in triangles of three types: with angles $(\pi/3, \pi/3, \pi/3)$, $(\pi/2, \pi/4, \pi/4)$ and $(\pi/2, \pi/3, \pi/6)$. As is well known, the topological structure of a smooth flow on the torus which has no fixed points and which preserves Lebesgue measure is completely determined by the rotation number of the flow. It is not difficult to see that in the situation being considered here the rotation number of the flow $\{T_t^c\}$ is a nonconstant analytic function of c . Consequently, for all values of c , except a countable set of values corresponding to rational rotation numbers, the flow $\{T_t^c\}$ is topologically transitive on the torus M_c and is ergodic – indeed, strongly ergodic – with respect to the measure m_c . In the next section we will show that this picture is partially preserved even in the general case, when $p(Q) > 1$ and M_c is not a torus (see Propositions 3 and 4). We make two remarks.

1. From the point of view of analytic mechanics a billiard is a Hamiltonian system with two degrees of freedom. In the case of a polygon whose angles are commensurable with π , this system has two independent first integrals: the energy integral $H = |v|^2/2$ and the above-defined integral F . The functions H and F occur in involution "almost everywhere," and a billiard may be regarded as "almost" integrable in the sense of a Liouville system. The violation of commutativity of H and F at points with carriers which meet vertices of the polygon entails that the phase space of the system decomposes not only into tori but also into invariant knots (i.e., two-dimensional manifolds of genus $p > 1$).

2. It is well known that a system of two solid spheres of masses m_1 and m_2 in an interval (with elastic reflections from the ends of the interval) is equivalent to a billiard in a right triangle with acute angle $\alpha = \arctan \sqrt{m_1/m_2}$ ([1], Lecture 10). If the angle α is commensurate with π , then the system of two spheres has two first integrals and reduces to flows on invariant knots in phase space (the invariant surfaces are tori only when $m_1 = m_2$, $m_2 = 3m_1$, or $m_1 = 3m_2$).

§ 2. Topological Transitivity in Invariant Manifolds

Definition 2. A flow $\{S_t\}$ on a two-dimensional manifold M^2 will be called quasiminimal if the following conditions are satisfied:

- (KM1) the number of singular points of the flow is finite;
- (KM2) each singular point is a multisaddle point with a finite number of entering and exiting separatrices;
- (KM3) each positive or negative semitrajectory which is not a separatrix is dense in M^2 .

A. G. Maier [2] dealt with the study of smooth flows on orientable surfaces and having a finite number of equilibrium conditions for boundary cycles and separatrices. We shall need a corollary of the basic result ([2], Theorem IX) (a corollary used by one of the authors in [3]).

COROLLARY TO MAIER'S THEOREM. Suppose that the flow $\{S_t\}$ of class C^1 on the closed orientable two-dimensional manifold M^2 satisfies the conditions (KM1) and (KM2), and in addition the conditions (M0) $\Omega(\{S_t\}) = M^2$:

- (M1) $\{S_t\}$ has no periodic trajectories;
- (M2) no separatrix of the flow S_t goes from one singular point to another.

Then, the flow $\{S_t\}$ possesses the property (KM3) (and, therefore, is quasiminimal).*

PROPOSITION 3. For all values $c \in (0, \pi/N)$ except a countable set of values, the flow $\{\tilde{T}_t^c\}$ is quasiminimal (and therefore the flow $\{T_t^c\}$ is topologically transitive).

Proof. By virtue of Proposition 2 the flow $\{\tilde{T}_t^c\}$ satisfies conditions (M0), (KM1), and (KM2) of the corollary stated above. Therefore it is sufficient to verify that for a billiard in a polygon Q , periodic trajectories and "trajectories" which go between vertices (corresponding to separatrices which join multi-saddles) can go along only a countable number of directions. Let us apply to trajectories of these two types the "straightening-out" procedure described in the introduction. We find that each of these trajectories has on the plane \mathbf{R}^2 a direction of the form AZ , where A is some vertex of the polygon $Q \subset \mathbf{R}^2$ and Z is the image of one of the vertices of Q under a composition of reflections of the plane \mathbf{R}^2 with respect to the polygon Q . The set of such points Z in the plane is countable, and Proposition 3 is proved.

Definition 3 (see [3]). Suppose that the flow $\{S_t\}$ of class C^1 on the two-dimensional manifold M^2 has a finite number of fixed points. A Borel measure μ on M^2 is called a nontrivial invariant measure of the flow $\{S_t\}$ if it is invariant, the measure of each trajectory of the flow is zero, and for every open neighborhood U of the set of fixed points $\mu(M^2 \setminus U) < \infty$.

The following assertion comes essentially from Theorem 1 of [3], although formally that theorem concerns flow with nondegenerate saddles.

PROPOSITION 4. If the flow $\{\tilde{T}_t^c\}$ is quasiminimal, then this flow (and therefore also the flow $\{T_t^c\}$) has at most p distinct (up to multipliers) ergodic nontrivial invariant measures, where p is the genus of the manifold $M_c(Q)$.

§ 3. Topological Transitivity in Phase Space

We now deduce some consequences for billiards in arbitrary polygons.

Denote by \mathcal{P}_n the set of all orientable nondegenerate n -gons with indexed vertices, considered as distinct up to similarity. An element $Q \in \mathcal{P}_n$ shall be identified with the ordered set $\{a, b, q_1, \dots, q_{n-2}\}$ of the vertices of an n -gon in the plane \mathbf{R}^2 having two points a and b fixed. The embedding $\mathcal{P}_n \rightarrow \mathbf{R}^{2(n-2)}$ induces a topology in \mathcal{P}_n ; it is not difficult to construct a metric for this topology in which the space \mathcal{P}_n is complete.

* It is possible to prove that for manifolds of genus $p > 1$ condition (M1) follows from the remaining conditions of this assertion.

Let $\mathcal{P}_n^0 \subset \mathcal{P}_n$ be the set of those n -gons whose angles are commensurate with π , and let $\mathcal{P}_n^N = \{Q \in \mathcal{P}_n^0: N(Q) \geq N\}$ (the number $N(Q)$ for $Q \in \mathcal{P}_n^0$ is defined at the beginning of § 1). It is clear that for each N the set \mathcal{P}_n^N is everywhere dense in \mathcal{P}_n .

PROPOSITION 5. Let \mathcal{R} be a closed subset of \mathcal{P}_n such that for each N the set $\mathcal{P}_n^N \cap \mathcal{R}$ is everywhere dense in \mathcal{R} . Then \mathcal{R} includes an everywhere dense subset \mathcal{R}_* of type G_δ such that for each n -gon $Q \in \mathcal{R}_*$ the flow $\{T_t^Q\}$ in $M(Q)$ is topologically transitive.

Proof. We introduce a metric in the subset $M^0(Q) = \text{Int } Q \times S^1 \subset M(Q)$ by the formula

$$d_Q((q, v), (q', v')) = \max\{|q - q'|, d_1(v, v')\},$$

where d_1 is the natural metric in $S^1 = \mathbf{R}/\mathbf{Z}$. In distinction to $M(Q)$, all spaces $M^0(Q)$ are pairwise homeomorphic. Let us fix some "standard" n -gon $Q_0 \in \mathcal{P}_n$ and then define a homeomorphism $h_Q: M^0(Q_0) \rightarrow M^0(Q)$ by the formula $h_Q(q, v) = (f_Q(q), v)$, where $q' = f_Q(q) = f(q, Q) = f(q; q_1, \dots, q_{n-2}) \in \mathbf{R}^2$ depends continuously on q, q_1, \dots, q_{n-2} ; here, for each $Q \in \mathcal{P}_n, f_Q: Q_0 \rightarrow Q$ is a homeomorphism and f_{Q_0} is the identity transformation. Finally, let us choose in $M^0(Q_0)$ a countable basis of open sets consisting of open balls B_k of radius r_k in the metric $d_{Q_0}, k = 1, 2, \dots$.

Let \mathcal{R}_{kl} be the set of those $Q \in \mathcal{R}$ for which there is a trajectory of the flow $\{T_t^Q\}$ which intersects both $h_Q B_k$ and $h_Q B_l$, even if it may not be defined for all t . We set $\mathcal{R}_* = \bigcap_{k,l=1}^{\infty} \mathcal{R}_{kl}$ and we shall show that the set \mathcal{R}_* possess all the properties mentioned in the statement of the proposition.

From the continuity of the function f and from the continuous dependence of each finite piece of the configuration of a trajectory of the flow $\{T_t^Q\}$, i.e., of a broken line $q(t; q_0, v_0; q_1, \dots, q_{n-2}) = q(t; x_0; Q)$ in the plane \mathbf{R}^2 , on the initial condition $x_0 = (q_0, v_0)$ and on the vertices q_1, \dots, q_{n-2} of an n -gon $Q \in \mathcal{R}$, it follows that all of the sets \mathcal{R}_{kl} are open. Moreover, if $N > \min\{r_k^{-1}, r_l^{-1}\}$ and $Q \in \mathcal{P}_n^N \cap \mathcal{R}$, then from the definition of h_Q , and B_k and formulas (3) it follows that for every c the manifold $M_c(Q')$ intersects both $h_Q B_k$ and $h_Q B_l$, and by virtue of Proposition 3 $Q' \in \mathcal{R}_{kl}$; by the condition that the sets $\mathcal{P}_n^N \cap \mathcal{R}$ are everywhere dense in \mathcal{R} , the sets \mathcal{R}_{kl} are likewise everywhere dense in \mathcal{R} . Thus, \mathcal{R}_* is an everywhere dense subset of \mathcal{R} and of type G_δ (we recall that \mathcal{P}_n is complete, while \mathcal{R} is closed).

Since for every Q the sets $h_Q B_k, k = 1, 2, \dots$ form a basis of open sets in $M^0(Q)$ (and in $M(Q)$), it follows from the definition of \mathcal{R}_{kl} and \mathcal{R}_* that if $Q \in \mathcal{R}_*$, then for every two open sets $U, V \subset M(Q)$ there is a trajectory of the flow $\{T_t^Q\}$ which intersects both U and V . Since, by a time transformation which nowhere vanishes on $M^0(Q)$, the flow $\{T_t^Q\}$ may be transformed into a continuous flow, the topological transitivity of $\{T_t^Q\}$ follows. Proposition 5 is proved.

COROLLARY. A system of two solid spheres in an interval (with elastic reflections from the ends of the interval) is topologically transitive on each surface of constant energy for values of the ratio m_1/m_2 belonging to an everywhere dense subset of the half-line of type G_δ .

For the proof it is sufficient to apply Proposition 5 to the set $\mathcal{R} \subset \mathcal{P}_3$ of right triangles and to take into account Remark 2 of § 1.

Propositions 3 and 5 proved above characterize the topological picture of the behavior of billiard trajectories in polygons; the corresponding metric picture remains very unclear. The following two questions strike us as being interesting.

1. In a polygon whose angles are commensurate with π , does ergodicity follow from topological transitivity of the billiard flow on an invariant manifold?
2. Are there polygons for which the billiard flow is ergodic in all of phase space? How large is the set of these polygons (from the point of view of density, category, measure, etc.) in \mathcal{P}_n ?

LITERATURE CITED

1. Ya. G. Sinai, Introduction to Ergodic Theory [in Russian], Erevan University Press, Erevan (1973).
2. A. G. Maier, "Trajectories on orientable surfaces," Matem. Sbornik, 12 (54), No. 1, 71-84 (1943).
3. A. B. Katok, "Invariant measures of flows on orientable surfaces," Dokl. Akad. Nauk SSSR, 211, No. 4, 775-778 (1973).