Записки научных
семинаров ПОМИ
Том 300, 2003 г.

Y. Baryshnikov and V. Zharnitsky

BILLIARDS AND NONHOLONOMIC DISTRIBUTIONS

Abstract. In this note, billiards with full families of periodic orbits are considered. It is shown that construction of a convex billiard with a "rational" caustic (i.e., carrying only periodic orbits) can be reformulated as a problem of finding a closed curve tangent to a $(N-1)$-dimensional distribution on a $(2N-1)$-dimensional manifold. The properties of this distribution are described as well as some important consequences for the billiards with rational caustics. A very particular application of our construction states that an ellipse can be infinitesimally perturbed so that any chosen rational elliptic caustic will persist.

1. Introduction

A classical 2-dimensional billiard is the mechanical dynamical system describing successive bounces of an elastic ball from the boundary of a plane domain. When the ball hits the boundary, the normal component of the velocity reverses and the tangential one is conserved; between bounces no forces act on the ball. One typically abstracts this (continuous time) dynamical system to its discrete time version, recording only the boundary points and the incidence angles at which the ball hits the boundary. This latter dynamical system is called *Classical (Birkhoff) billiard* if the boundary is convex and smooth. The billiard boundary curvature must be at least of bounded variation so the billiard flow is well defined [9]. Birkhoff billiards serve as archetypal Hamiltonian systems, with the bonus that there is an intricate interplay between the geometry of the billiard table and the properties of the system.

Periodic points feature prominently in studies of billiard dynamics, playing significant role both in ergodic and integrable or nearly integrable systems. In this note, we are interested in billiard tables having *many* periodic points, in fact 1- or 2-parametric families of them. Motivations for this interest are manifold: integrable billiard systems should have 1-parametric families of periodic points; spectral properties of the Laplace operator in a billiard domain depend on whether the periodic points of the billiard dynamical system have Lebesgue measure zero; there are also some conjectured applications in photonic physics...

The construction of billiard tables with full families of periodic orbits (that is, families parameterized by the circle, such that each point of the orbit traverses the billiard table boundary as parameter varies) has been done in the past. Thus, billiards with full families of 2-period orbits correspond to curves of constant width), see, e.g., [5] for some examples and more references. It is considerably harder to find examples of billiards with full families of 3- or higher-periodic orbits. To the best of our knowledge there are only two references [11], [12, unpublished]. In [11], the differential relations are derived which must be satisfied by billiard boundaries with full families of 3-period orbits and a particular solution is found. These differential relations are however rather complicated, so that this result does not reveal the structure of the set of billiards with full families.

It turns out, that one can describe billiard domains having one-parameter family of periodic orbits as closed curves tangent to certain nonholonomic distribution (i.e., nonintegrable subbundle of the tangent bundle) on a manifold. This approach has been used implicitly in the studies of curves of constant width (which bound exactly the billiard domains with 1-parametric families of 2-periodic trajectories) [8]. Here we generalize this approach and develop the necessary machinery.

Roughly, we consider the set of N-tuples of points (ordered cyclically) on the plane. If these points form an N-periodic trajectory of the billiard system, the tangent lines to the billiard domain boundary should be orthogonal to the bisector of the angle formed by consecutive segments of the periodic trajectory. Thus, infinitesimal shifts of N-periodic orbits form an N-dimensional distribution, the *Birkhoff distribution*. This distribution can be shown to be nonholonomic in the manifold of polygons of given circumference with the growth vector $(N, 2N-1)$. A parametric family of N-periodic orbits would then be immersed into the space of N-tuples of points on plane as an *integral manifold* for the Birkhoff distribution. This simple observation motivated the present study.

2. Billiards with rational caustics carrying 3-period orbits

In this section, we give an outline of our construction for special case of 3-period orbit families. General constructions are given in the following sections.

For any triangular periodic orbit in a Birkhoff billiard, the bisector at each vertex is orthogonal to the tangent line to the billiard boundary. Hence, a 1-parametric family can be obtained by taking the starting triangle and letting its vertices move orthogonally to the bisectors (which change of course with the triangle) with the some velocities $u_1(t)$, $u_2(t)$, $u_3(t) > 0$ (where $u_i > 0$ corresponds to the clockwise direction).

Using this observation, we can now consider an associated nonholonomic system of three points on the plane with nonholonomic constraints: they are to move only orthogonally to the bisectors: $b_i \perp L_i$.

Each triangle $z = (z_1, z_2, z_3)$ can be considered as a point in $\mathbb{R}^6$ where nonholonomic constraints correspond to a distribution $\mathcal{H} \subset \mathbb{R}^6$. Obviously, this distribution is not completely nonholonomic as the perimeter of the triangle does not change when vertices move orthogonally to the bisectors, i.e.,

$$P(z_1, z_2, z_3) = |z_1 - z_2| + |z_2 - z_3| + |z_3 - z_1| = P_0.$$

It turns out that there are no other integrable relations and, thus, this distribution is of type $(3, 5)$ on M_1 (the manifold of triangles with $P = 1$).

One can choose coordinates on $M(1)$ to be the angles of bisectors with a fixed line and cartesian coordinates of the point of intersection of bisectors $(\alpha_1, \alpha_2, \alpha_3, x_c, y_c)$.

In these coordinates, the nonholonomic constraints take the form

$$dx_c = \sum_{i=1}^{3} A_i d\alpha_i$$
$$dy_c = \sum_{i=1}^{3} B_i d\alpha_i,$$

where A_i, B_i depend only on α_j.

3. Notation and Main results

Fix an integer $N \geq 2$. Let $\mathbf{N} = \mathbb{Z}/N\mathbb{Z}$ be the cyclic group of order N. Let $V = \mathbb{R}^2$ be the Euclidean plane, and $S = S^*V$ the spherization of its cotangent bundle: in other words a point in S is the pair (z, e), where z is a point in V and e is a unit norm covector at based at z. Denote by $V_N = V^{\mathbf{N}}$, $S_N = S^{\mathbf{N}}$ the set of V- or S-valued functions on $\mathbf{N}$, and by $p_N : S_N \to V_N$ the corresponding projection (forgetting covectors). We will write a generic element of V_N as $\mathbf{z} = (z_0, \ldots, z_{N-1})$, and similarly $\mathbf{s} = (s_0, \ldots, s_{N-1})$, $s_i = (z_i, e_i)$ for a generic element of S_N.

Recall that a *distribution* or a *differential system* [4] is a (constant rank) subbundle) of the tangent bundle. Spherization of the cotangent bundle of V carries a natural distribution H of rank 2 (*the contact structure*): it is defined by the condition that e vanishes on the velocity of the base point z.

The manifold S_N carries therefore the distribution H_N of rank $2N$ which is just the product of the distributions in the factors of S_N. Notice that H contains $\ker(p_1)_*$ (one can rotate e arbitrarily insofar z is constant) and consequently H_N contains the kernel of $(p_N)_*$

Let U be an open subset of V_N and $l : U \to S_N$ be a section of the fibration p_N over U. We denote the image of U under l as $L_l := l(U)$, or simply L when no confusion threatens. Obviously, the intersection of the subbundles TL and $\ker(p_N)_*$ over L is trivial, and the dimension count implies

Lemma 1. *The intersection of subbundles over* $\ker(p_N)_* \cap TL$ *defines a subbundle on* L *of constant rank* N.

We denote this distribution on L, as well as its preimage under the diffeomorphism $l : U \to L_l$ by $\mathcal{D}_l$.

To smooth enough function $f : V_N \to \mathbb{R}$ one can associate a section of p_N (over an open subset of V_N) by setting

$$e_i = \frac{d_i f}{|d_i f|},$$

where $d_i f$ is just the differential of f considered as a function of ith variable z_i. Then on $U(f) = \{\mathbf{z} : d_i f \neq 0, i \in \mathbf{N}\}$ these unit norm covectors form a section which we will denote by $l(f) : U(f) \to S_N$.

This section gives rise to a distribution $\mathcal{D}_{l(f)}$ on $U(f)$.

Obviously, the distribution $\mathcal{D}_{l(f)}$ is tangent to the level sets $M(a) = f^{-1}(a)$ of function f. Therefore it cannot be completely nonholonomic in $U(f)$, but at most in the manifolds $M(a)$. This is true indeed for generic f, and, what is more important to us, this is also true for special functions f relevant to the billiard problem:

Proposition 1. *Let* ϕ *be a smooth function with positive derivative on the positive half-axis. Then the distribution* $\mathcal{D}_{l(f)}$ *corresponding to function*

$$f : \mathbf{z} \mapsto \sum_{i \in \mathbf{N}} \phi(|z_i - z_{i+1}|) \tag{1}$$

is completely nonholonomic with growth vector $(N, 2N-1)$.

In other words, already the first Lie brackets of the vector fields tangent to $\mathcal{D}_{l(f)}$ generate the full tangent bundle of the level manifolds $M(a)$.

Remark 1. Function $\phi(s) = s$, which corresponds to extremization total length, is central in our approach to the billiard problem. However, other functions are of interest as well: the function $\phi(s) = s^2$ describes trajectories extremization total energy (one might think of N points connected cyclically by identical elastic strings) and the function $\phi(s) = \sqrt{s}$ corresponds to polygons of extremal (algebraic) area.

The open subset $U(f)$ over which the section $l(f)$ is defined unambiguously is given for functions (1) by

$$U(f) = \{\phi_i e_i - \phi_{i+1} e_{i+1} \neq 0, i \in \mathbf{N}\},$$

where

$$e_i := \frac{z_{i+1} - z_i}{|z_{i+1} - z_i|}$$

is the unit vector pointing along ith leg of trajectory $\mathbf{z}$, $l_i = |z_{i+1} - z_i|$ is its length and

$$\phi_i = d\phi/ds|_{s=l_i}.$$

For $\mathbf{z} \in U(f)$ one can choose as the basis for $\mathcal{D}_{l(f)}$ the tangent vectors $L_i, i \in \mathbf{N}$ defined by

$$L_i z_j = \delta_{ij} J(\phi_i e_i - \phi_{i+1} e_{i+1}), \tag{2}$$

where J is the rotation by 90 degrees.

We concentrate however from now on the billiard problem proper.

Definition 1. *Distribution $\mathcal{D}_{l(f)}$ for $f = \sum_{i \in \mathbf{N}} |z_i - z_{i+1}|$ is called the* `Birkhoff distribution`.

4. Billiards tables

Let B be a billiard table, that is a disk in V with smooth (or piece-wise smooth) boundary which we will assume is parameterized by $\gamma : \mathbb{T}^1 \to V$ (say, counterclockwise by the length) so that

$$\partial B = \gamma(\mathbb{T}^1).$$

Taking into account the outer normal $e(t)$ at the boundary point $\gamma(t)$ defines the map

$$S\gamma : \mathbb{T}^1 \to S, S\gamma(t) = (\gamma(t), e(t)).$$

Exponentiating by $\mathbf{N}$ we obtain a mapping

$$S\gamma_N : \mathbb{T}^{\mathbf{N}} \to S_N$$

whose image $\Gamma_N \subset S_N$ is an embedded N-torus.

The definition implies

Lemma 2. *The torus Γ_N is an integral submanifold of the distribution H_N.*

Another immediate consequence of the definition is

Lemma 3. *A point $\mathbf{s} \in \Gamma_N \cap L_{l(f)}$ if and only if $\mathbf{z} = p_N(\mathbf{s})$ is a critical point of f restricted to Γ_N. In particular, for $f = \sum_{i \in \mathbf{N}} |z_i - z_{i+1}|$, the sequence $\mathbf{z} = (z_0, \ldots, z_{N-1}) \in p_N(\Gamma_N \cap L_{l(f)})$ is (V-component of) a closed trajectory of the billiard mapping corresponding to the table B.*

5. Caustics

Let B be a billiard table. A 1-parametric family of N-periodic billiard trajectories in B can be lifted, by constructions of previous section, to a curve in $\Gamma_N \cap L_{l(f)}$. This curve is clearly tangent to $\mathcal{D}_{l(f)}$.

We say that a billiard table B is N-*Lazutkin* if there exists a family of N-periodic trajectories traversing the boundary of B in counterclockwise direction, or, equivalently, if there exists an N-*periodic caustic* entirely within B (recall that a caustic is a unicursal piece-wise smooth curve such that the tangency to this curve of one leg of the billiard trajectory implies the tangency of the next, and thus all, legs; the caustic is N-periodic if all such trajectories are N-periodic).

We denote by $\sigma_N : \mathbf{N} \to \mathbf{N}$ shift by $+1$, and also corresponding mappings of V_N, S_N (i.e., $\sigma_N : (z_0, z_1, \ldots) \mapsto (z_1, z_2, \ldots, z_0)$ etc).

Proposition 2. *The billiard table is N-Lazutkin, if and only if there exists a mapping $\lambda : I = [0,1] \to S_N, \lambda(t) = (s_0(t), \ldots, s_{N-1}(t))$ such that*

- *$\sigma\lambda(0) = \lambda(1)$;*
- *$dp_N(\lambda(t))/dt \neq 0$ is a linear combination of vector fields L_i defined in (2) with positive coefficients;*
- *$\sigma(t) \in \Gamma_N \cap L_{l(f)}$ for all $t \in I$.*

Conversely, traces $\{z_i(t)\}, t \in I, i \in \mathbf{N}$ of any curve $\lambda : I = [0,1] \to S_N$ satisfying these three conditions define an N-Lazutkin billiard table

This reduces the study of billiard tables with closed families of periodic trajectories to (much more accessible) study of curves tangent to the Birkhoff distribution.

Denote the space of horizontal curves λ satisfying the conditions of Proposition 2 with endpoints $\mathbf{z}, \sigma\mathbf{z}$ as $\mathcal{L}_N(\mathbf{z})$, and $\mathcal{L}_N = \cup_{\mathbf{z}} \mathcal{L}_N(\mathbf{z})$.

6. N-Lazutkin billiards are deformable

Consider an N-Lazutkin billiard. How freely can one deform the billiard table *preserving N-Lazutkin property*? More precisely, consider the Banach manifold $\mathcal{C}$ of smooth (say, C^∞) curves $\gamma : \mathbb{T} \to V$, and there the subset of curves having the N-Lazutkin property. What is the local structure of this subset? By Proposition 2, the set $\mathcal{L}_N$ parameterizes N-Lazutkin billiards. The structure of $\mathcal{L}_N$ is described by the following theorem:

Theorem 1. *$\mathcal{L}_N$ is a smooth Banach manifold modelled on subspaces of codimension $N-1$ in the Banach manifold $V_N \times C^\infty(I, \mathbb{R}^{N-1})$.*

Sketch of the proof. Let $\lambda \in \mathcal{L}_N(\mathbf{z})$, and $\gamma(\mathbb{T}^1)$ the corresponding billiard table boundary. Extend the natural parameter t on $\gamma(\mathbb{T}^1)$ to a function on its (collar) neighborhood $\gamma \subset \mathcal{N} \subset V$.

The fibers of the fibration $T : \mathcal{N}^{\mathbf{N}} \to \mathbb{T}^{\mathbf{N}}$ given by $T : \mathbf{z} \mapsto (t(z_0), \ldots, t(z_{N-1})$ are transversal to distribution $\mathcal{D}$. Hence any horizontal curve can be determined by its starting point $\mathbf{z}$ and its T-projection $\theta = T\lambda$ (we will call this horizontal curve *the lift of* θ). Consider a deformation $(\tilde{\theta}, \tilde{\mathbf{z}} \in T^{-1}(\theta(0)))$ of the data $(\theta = T\lambda, \mathbf{z})$, $\lambda \in \mathcal{L}(\mathbf{z})$ and the corresponding lift λ'. The curve $\tilde{\lambda}$ will again determine a point in $\mathcal{L}_N$ if

1. if $\tilde{\theta}(1) = \sigma\tilde{\theta}(0)$;
2. the velocity of the lift of $\tilde{\theta}$ remains in the open cone spanned by L_i's;
3. $\tilde{\lambda}(1) = \sigma\tilde{\lambda}(0)$ (remark that these two points belong to the same fiber $T^{-1}(\tilde{\theta}(1))$ as σ commutes with T).

The first condition is open and thus satisfied for deformations close to λ. Second condition fixes the endpoints of the curve $\tilde{\theta}$. The space of such curves in the base $\mathbb{T}^{\mathbf{N}}$ can be identified (locally) with with $C^\infty(I, \mathbb{R}^{N-1})$. Let $E : C^\infty(I, \mathbb{R}^{N-1}) \to T^{-1}(T\sigma\tilde{\mathbf{z}})$ be the endpoint mapping. Third condition implies that $\mathcal{L}(\mathbf{z}')$ is locally parameterized by $E^{-1}(\sigma\mathbf{z})$, and to finish the proof it remains to show that this mapping has full rank $(N-1)$ (the dimension of the fiber). The mapping sending the curve on the base to the endpoint may indeed be degenerate in our situation. However, this can happen only if the velocity of the lift is everywhere on $[0,1]$ belongs to the kernel of the curvature, a $TS_N/\mathcal{D}$-valued 2-form on $\mathcal{D}$ [6]. A computation shows that this is impossible for Birkhoff distributions (indeed, for all distributions (1)), as long as the velocity is a linear combination of L_i's with coefficients of the same sign.

A natural (and so far only?) examples of N-Lazutkin billiards were ellipses. Deformability of N-Lazutkin billiards implies, inter alia, that the family of such billiards is infinite-dimensional.

7. On the set of 3-periodic points

We conclude this note with a short proof that the set of 3-periodic points of the billiard mapping has there there is no open 2-dimensional disk consisting of 3-periodic points of Birkhoff billiard mappings. This result is well-known (at least 3 proofs have been published so far, starting with [13]). Yet we would like to emphasize how naturally nonholonomic distributions enter the picture.

Theorem 2. *There exists no open 2-dimensional disk consisting of 3-periodic trajectories of Birkhoff billiard mapping.*

Proof. Indeed, existence of such a ball would imply by Proposition 2 that there exists a smooth 2-dimensional disk C embedded into $\Gamma_N \cap L_l$. By Lemmas 3 and 2, this disk is an integral manifold for $\mathcal{D}_l$. It remains to show that $\mathcal{D}_l$ admits no 2-dimensional integral manifolds for $N = 3$. Indeed, the growth vector in this case is $(3, 5)$. It is well known (see [6]), that all such distributions can be extended to distribution of the growth $(2, 3, 5)$. Denote corresponding subbundles as $\mathcal{D}^2 \subset \mathcal{D}^3 = \mathcal{D}_l$. Further, one can compute easily that a 2-manifold integral with respect to $\mathcal{D}$ should be also integrable with respect to $\mathcal{D}^2$. This, however, is in contradiction with the growth vector, as commutators of the vector fields tangent to $\mathcal{D}^2$ vanish on C and cannot generate $\mathcal{D}^3$ there.

References

1. V.I. Arnold, *Geometrical methods in the theory of ordinary differential equations*, Springer (1983).
2. G. D. Birkhoff, *On the periodic motions of dynamical systems*, Acta Math. **50** (1927), 359–379.
3. G. D. Birkhoff, in: *Dynamical Systems, Dynamical systems*, AMS Colloquium Publications, Vol. IX, AMS, Providence, R.I. (1966).
4. P. A. Griffiths, *Exterior differential systems and the calculus of variations*, Progress in Mathematics **25**, Birkhäuser, Boston, Mass. (1983).
5. A. Katok, Gutkin, *Caustics for inner and outer billiards*, Comm. Math. Phys. **173**, No. 1 (1995), 101–133.
6. R. Montgomery, *A tour of subriemannian geometries, their geodesics and applications* Mathematical Surveys and Monographs **91**, AMS, Providence, RI (2002).
7. A. M. Vershik, V. Y. Gershkovich, *Nonholonomic dynamical systems. Geometry of distributions and variational problems*, in: *Current Problems in Mathematics. Fundamental Directions*, Vol. 16, Moscow (1987), pp. 5–85.

8. M. Ghandehari, *An optimal control formulation of the Blashke–Lebesgue theorem*, J. Math. Anal. Appl. **200**, No. 2 (1996), 322–331.
9. V. Zharnitsky, *Invariant tori in the systems of billiard type*, Commun. Math. Phys. **211** (2000), 289–302.
10. B. Halpern, *Strange billiard tables*, Trans. Amer. Math. Soc. **232** (1977), 297–305.
11. Nobuhiro Innami, *Convex curves whose points are vertices of billiard triangles*, Kodai Math. J. **11** (1988), 17–24.
12. C. Berge, *Unpublished result*, communicated to the authors by S. Tabachnikov.
13. M. Rychlik, *Periodic points of the billiard ball map in a convex domain*, J. Diff. Geom. **30**, No. 1 (1989), 191–205.

Mathematical Sciences Research
Bell Laboratories, Lucent Technologies
Murray Hill, NJ 07974

Поступило 30 ноября 2002 г.