TWO APPLICATIONS OF JACOBI FIELDS TO THE BILLIARD BALL PROBLEM.

Maciej P. Wojtkowski

September 13, 93

Abstract.
We present new proofs of two results on the billiard ball problem by Rychlik [R] and Bialy [B].

§0. INTRODUCTION.

We will give new proofs of two results on the billiard ball problem by Rychlik [R] and Bialy [B]. The original proofs were based on variational considerations. In our approach the variational context is absent, the dynamical system takes the center stage. We hope the simplifications provided by our method will make possible some progress on the conjectures for which these results lend partial support.

§1.THE DYNAMICAL SYSTEM.

Let us consider a convex domain Q in the plane. The billiard ball system is the flow Φ^t on $Q \times \mathbb{S}^1$ defined by the free motion of a point particle in Q, with elastic reflections at the boundary ∂Q (the angle of reflection equal to the angle of incidence). The circle $\mathbb{S}^1$ represents unit velocities. Strictly speaking, we need to identify the velocities at the boundary according to the collision law. The flow Φ^t preserves the Liouville measure ν equal to the product of the Lebesgue measures in Q and $\mathbb{S}^1$.

Birkhoff [Bir] thought that this dynamical system is a very good model for Hamiltonian dynamics. In the last thirty years his belief proved to be strikingly accurate. We understand as much about the low dimensional Hamiltonian dynamics, as we know about the billiard system.

The flow Φ^t has a natural section map $T : \mathcal{M} \to \mathcal{M}$, where $\mathcal{M} = \partial Q \times [0, \pi]$. The map T describes the dynamics "from collision to collision". The space $\mathcal{M}$ is the set of unit tangent vectors attached at the boundary and pointing inwards. It can be coordinatized by (s, φ), where s is the arclength parameter taken modulo p,

---

We would like to thank Gil Bor, Jian Cheng, Victor Donnay and Lenny Friedlander for helpful and enlightening discussions. The suggestions of the referee are also gratefully acknowledged.

the perimeter of ∂Q, and $\varphi \in [0, \pi]$ is the angle that the unit velocity vector makes with the tangent vector to the boundary.

The map T, the billiard ball map, preserves the measure $\mu = \sin\varphi ds d\varphi$. This measure is obtained from ν by the contraction (internal product) with the velocity vector field of the flow Φ^t. It follows that

$$\int_{\mathcal{M}} \tau(x) d\mu(x) = \nu(Q \times \mathbb{S}^1) = Area(Q) 2\pi, \tag{1}$$

where $\tau(x)$ for $x \in \mathcal{M}$ is the time to the next collision. This formula is also well known in integral geometry (cf. [S], formula (3.6)).

§2. JACOBI FIELDS.

Let us consider a family of billiard orbits (or straight lines parametrized by the arc length)

$$p(\epsilon) = p(\epsilon, t) = \gamma(\epsilon) + v(\epsilon)t, \quad |\epsilon| < \epsilon_0, \ -\infty < t < +\infty.$$

The Jacobi field $J(t)$ is defined as

$$J(t) = \frac{\partial p}{\partial \epsilon}|_{\epsilon=0}. \tag{2}$$

The Jacobi field is the infinitesimal description of the family of billiard orbits (lines) around the orbit $\epsilon = 0$

$$J(t) = \gamma'(0) + v'(0)t.$$

The Jacobi fields satisfy the following differential equation (the curvature of the plane is equal to zero)

$$\frac{d^2}{dt^2} J = 0.$$

At any point in the phase space the Jacobi fields form the tangent subspace. They can be identified with pairs of vectors $(J(0), J'(0)) = (\gamma'(0), v'(0))$. Note that the prime denotes differentiation with respect to t or ϵ, whichever applies. It is natural to consider only the Jacobi fields for which $J(0) = \gamma'(0)$ is orthogonal to the velocity vector $v(0)$. Note that $J' = v'(0)$ is automatically orthogonal to $v(0)$. We will call such Jacobi fields transversal. One can also think about the (two dimensional) space of transversal Jacobi fields as the factor space by the (one dimensional) subspace of Jacobi fields corresponding to variations of the initial point along the orbit (time variations).

Transversal Jacobi fields can be considered as tangent vectors to the section $\mathcal{M}$. The connection is given by the formulas

$$\begin{aligned} J &= \sin\varphi ds, \\ J' &= -\, k ds - d\varphi, \end{aligned} \tag{3}$$

where $(J,\, J') \in \mathbb{R}^2$ are now linear coordinates after the appropriate choice of a unit vector orthogonal to $v(0)$ was made, and $k = k(s)$ is the curvature of the boundary ∂Q.

The evolution of $(J,\, J')$ between collisions is given by the matrix

$$(4) \qquad \begin{pmatrix} 1 & \tau \\ 0 & 1 \end{pmatrix}.$$

The Jacobi fields can be naturally extended beyond reflections in the boundary: with every family of lines including a segment of a billiard orbit we can associate the family of reflected lines which include the segment of the reflected orbit. The reflected Jacobi field is now calculated by (2). A transversal Jacobi field $(J,\, J')$ is transformed at a reflection by the linear map

$$(5) \qquad \begin{pmatrix} -1 & 0 \\ \frac{2}{d} & -1 \end{pmatrix}, \quad \text{where} \quad d = \frac{\sin\varphi}{k(s)}.$$

§3.THE THEOREM OF BIALY.

In the recent paper [B] Bialy proved the following theorem

Theorem 1([B])). *If the phase space $\mathcal{M}$ of the billiard ball map is foliated by continuous closed invariant curves not null-homotopic in $\mathcal{M}$, then Q is a disk.*

It is well known ([Bir])that the billiard ball map in an ellipse is integrable. No other examples of convex domains with integrable billiard ball maps are known. It makes the following conjecture, attributed to Birkhoff, plausible.

Conjecture. *If a neighborhood of the boundary of $\mathcal{M}$ is foliated by continuous closed invariant curves not null-homotopic in $\mathcal{M}$, then Q is an ellipse.*

We will first present the proof of a weaker result then Bialy's Theorem, to make the line of thought perfectly clear. We will show then that the same argument does apply in the general case.

Definition. *A smooth closed curve contained in the domain Q is called a smooth caustic, if every billiard orbit with one segment tangent to the curve has all its segments tangent to the curve.*

The interior of a disk with the center removed is foliated by smooth caustics, the concentric circles.

Theorem 2. *If the domain Q is foliated by smooth caustics in such a way that almost every orbit is tangent to a caustic, then Q is a disk.*

Proof. Let us fix $x \in M$ and the corresponding billiard orbit tangent to a smooth caustic. We denote by $l(x)$ the distance from the boundary to the nearest point of

Figure 1 Billiard table and caustic.

tangency in the past with the smooth caustic. The nearest point of tangency in the future is then equal to $\tau(x) - l(Tx)$ (see Fig.1).

The transversal Jacobi field defined by the nearby orbits tangent to the same smooth caustic vanishes at the points of tangency. We get from (3) and (4)

$$\begin{pmatrix} 1 & \tau(x) - l(Tx) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ \frac{2}{d(x)} & -1 \end{pmatrix} \begin{pmatrix} 1 & l(x) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ * \end{pmatrix},$$

where the $*$ indicates that the respective component is not restricted. We obtain immediately

$$d(x) = \frac{2l(x)(\tau(x) - l(Tx))}{l(x) + (\tau(x) - l(Tx))}. \tag{6}$$

The last formula is the classical "mirror equation" of the geometric optics.

Since the harmonic mean of two positive numbers does not exceed the arithmetic mean we conclude that

$$2d(x) \le \tau(x) + l(x) - l(Tx). \tag{7}$$

Let us stress that to get (7) from (6), it is necessary to have

$$l(x) > 0 \quad \text{and} \quad l(Tx) < \tau(x),$$

(note that $d(x) > 0$). We will now obtain the desired conclusion by integrating (7) and comparing with the Isoperimetric Inequality, just as it was done in [B]. Integrating the right hand side of the inequality we obtain from (1)

$$\int_{\mathcal{M}} \tau(x) + l(x) - l(Tx) d\mu(x) = \int_{\mathcal{M}} \tau(x) d\mu(x) = 2\pi Area(Q).$$

We have used the fact that $l(x)$ is bounded and defined almost everywhere in $\mathcal{M}$. The integration of the left hand side yields

$$\int_{\mathcal{M}} 2d(x) d\mu(x) = \int_0^p \int_0^\pi \frac{2\sin^2\varphi}{k(s)} ds d\varphi = \pi \int_0^p \frac{1}{k(s)} ds.$$

By the Buniakovski - Cauchy - Schwartz inequality

$$\int_0^p \frac{1}{k(s)} ds \int_0^p k(s) ds \geq p^2,$$

with the equality only for the constant curvature, i.e. for the disk. Since

$$\int_0^p k(s) ds = 2\pi \tag{8}$$

we conclude that

$$2\pi Area(Q) \geq \frac{p^2}{2},$$

which violates the Isoperimetric Inequality except when it is the equality, i.e., when Q is a disk. $\square$

Let us consider a measurable one dimensional subbundle L of the tangent bundle of $\mathcal{M}$, i.e., for almost every $x \in \mathcal{M}$ we have a one dimensional subspace (a line) $L(x) \subset \mathcal{T}_x\mathcal{M}$ which depends measurably on x. We assume that for almost every $x \in \mathcal{M}$ the subspace $L(x)$ is not vertical, i.e., it is not tangent to the curve $\{s = const\}$. We choose an orientation of nonvertical lines by the condition $ds > 0$. Let us call such an oriented subbundle a monotone subbundle of the tangent bundle of $\mathcal{M}$. We call a monotone subbundle invariant ,if it is preserved, including the orientation, under the action of the derivative of T.

Let us recall that vectors tangent to $\mathcal{M}$ can be naturally identified with transversal Jacobi fields. If we have an invariant monotone subbundle L, then Jacobi fields which correspond to nonzero tangent vectors in the subbundle must vanish inside the billiard table, almost everywhere in $\mathcal{M}$. Indeed let $(J(x), J'(x))$, $x \in \mathcal{M}$ be a measurable family of nonzero Jacobi fields spanning the oriented subbundle $L(x)$, $x \in \mathcal{M}$. We have $J(x) > 0$ almost everywhere in $\mathcal{M}$ and there is a positive measurable function $\alpha : \mathcal{M} \to \mathbb{R}^+$ such that

$$\begin{pmatrix} -1 & 0 \\ \frac{2}{d(Tx)} & -1 \end{pmatrix} \begin{pmatrix} 1 & \tau(x) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} J(x) \\ J'(x) \end{pmatrix} = \alpha(x) \begin{pmatrix} J(Tx) \\ J'(Tx) \end{pmatrix}.$$

It follows that $J(x)+\tau(x)J'(x)=-\alpha(x)J(Tx)$ is almost everywhere negative and hence almost everywhere the Jacobi field vanishes inside the billiard table, i.e., there is a measurable function $f(x)$, $0<f(x)<\tau(x)$ such that

$$J(x)+f(x)J'(x)=0.$$

Let us now consider the measurable function

$$l(x)=\tau(T^{-1}x)-f(T^{-1}x).$$

We have $0<l(x)<\tau(T^{-1}x)$ almost everywhere. Moreover, as a consequence of the invariance of the subbundle, the function $l(x)$ satisfies (6). Hence we can repeat the proof of Theorem 2. We obtain the following

Theorem 3. *If the billiard ball map T has an invariant monotone subbundle of the tangent bundle of $\mathcal{M}$, then the billiard table Q must be a disk.*

Let us show how Bialy's Theorem 1 follows from Theorem 3. The mapping $T:\mathcal{M}\to\mathcal{M}$ is a measure preserving twist map. By the Birkhoff Theorem [H], any closed invariant curve not null-homotopic in $\mathcal{M}$ is the graph of a Lipschitz function $\varphi=\varphi(s)$. Although a general Lipschitz function may be nondifferentiable at some points, it must have finite upper and lower derivatives

$$\frac{d^+\varphi}{ds}=\limsup_{s_1,s_2\to s}\frac{\varphi(s_2)-\varphi(s_1)}{s_2-s_1},\quad \frac{d^-\varphi}{ds}=\liminf_{s_1,s_2\to s}\frac{\varphi(s_2)-\varphi(s_1)}{s_2-s_1}.$$

The upper and lower derivatives define respectively the upper and lower tangent lines to the graph. The bundle of upper (or lower) tangent lines to the invariant curves is defined everywhere and is clearly measurable and monotone (never vertical). The Jacobi field which spans this oriented tangent line can be chosen as (cf. (3)) $J=\sin\varphi(s)$, $J'=-k(s)-\frac{d^+\varphi}{ds}$. In turn, the positive function $l(x)$ ("the distance to the caustic") is given by the formula

$$l(s,\varphi)=\frac{\sin\varphi}{k(s)-\frac{d^{\pm}\varphi}{ds}}.$$

It follows from the invariance of the curves that this monotone subbundle is invariant. Indeed, the positive orientation of the curves is preserved under T.

To summarize, the tangents to the invariant curves of the family, such as described in the Bialy's Theorem 1, give us an invariant monotone subbundle. So Theorem 3 implies Theorem 1.

Remark.

Theorem 3 has an interesting connection to billiard tables with hyperbolic billiard ball map. The stable and unstable subbundles are invariant but by Theorem 3 they cannot be invariant as monotone subbundles. Indeed, in the case of Bunimovich's stadium and its generalizations by Donnay, [D], the positive orientation of the (un)stable subbundle is reversed at some of the segments of the orbit between the two flat pieces of the boundary (connecting the strictly convex pieces).

In the case of billiard tables bounded by closed convex scattering curves, introduced in [W], the situation is different. These are closed curves given by the natural equation $k = k(s)$, where $k(s) > 0$ is a smooth function in the interval $[0, p]$ such that

$$\frac{d^2}{ds^2}\frac{1}{k(s)} \leq 0$$

with strict inequality holding almost everywhere in s. Hence the closed curve is smooth except at one point. The examples include perturbations of the circle and the cardioid. For such billiard tables we have hyperbolicity in all of the phase space $\mathcal{M}$ and the positive orientation of the (un)stable subbundle is preserved (cf.[W]), so we do have two invariant monotone subbundles. The apparent contradiction with Theorem 3 is resolved by the failure of the formula (8). Because of the singularity, we have $\int_0^p k(s)ds > 2\pi$.

When we perturb the circle to a closed convex scattering curve, the monotone invariant subbundle of the billiard ball map in the disk splits into two subbundles (the stable and unstable ones). The proof of Theorem 2 shows that it is only the singularity of the boundary that provides room for such a splitting.

§4. THE THEOREM OF RYCHLIK.

The problem of the measure of periodic orbits for the billiard ball map arises in the spectral geometry [M-M]. A version of Kupka-Smale Theorem for billiard ball maps in convex domains was proven by Lazutkin [L]. It follows that for generic convex domains the periodic orbits of bounded period are isolated, and hence there are only countably many periodic orbits.

It is fairly clear that for convex domains with real analytic boundary the measure of periodic orbits must be zero. At the same time it seems plausible that by modifying the boundary in the C^∞ category one could produce a whole neighborhood filled with periodic orbits. Rychlik proved in [R] that it is not possible for orbits of period 3.

Theorem [R]**.** *The set of periodic orbits of period* 3 *of the billiard ball map is nowhere dense.*

Rychlik shows that actually the set must have measure zero. His proof is based on the study of the length functional, and it required computer assistance. Recently Stojanov [St] simplified the calculations, so that there is no need for the use of symbolic computation systems. We propose a more dynamical proof.

Proof. Let $\{x_0, x_1, x_2\}$ be a periodic orbit of period 3 with a neighborhood filled by periodic orbits of period 3.

We have that T^3 and its derivative DT^3 is equal to identity in the neighborhood of x_0. The last property can be formulated in the language of Jacobi fields. In

Figure 2 Periodic orbit with period 3.

particular we get from (4) and (5), see Fig. 2,

$$\begin{pmatrix} 1 & \tau(x_0) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ \frac{2}{d(x_0)} & -1 \end{pmatrix} \begin{pmatrix} 1 & \tau(x_2) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} =$$
$$\begin{pmatrix} -1 & 0 \\ -\frac{2}{d(x_1)} & -1 \end{pmatrix} \begin{pmatrix} 1 & -\tau(x_1) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ -\frac{2}{d(x_2)} & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The equality of the first components of the two vectors yields

$$\tau(x_0) + \tau(x_2) - \tau(x_1) = \frac{2\tau(x_0)\tau(x_2)}{d(x_0)}. \tag{8}$$

The total length of the periodic orbits in our neighborhood must be constant. Let us denote it by L,

$$L = \tau(x_0) + \tau(x_1) + \tau(x_2).$$

¿From the Cosine Formula we get

$$\tau(x_0) + \tau(x_2) - \tau(x_1) = \frac{4\tau(x_0)\tau(x_2)\sin^2\varphi_0}{L} \tag{9}$$

where $x_0 = (s_0, \varphi_0)$. Combining (8) and (9) we obtain

$$k(s_0) = \frac{2\sin^3\varphi_0}{L}. \tag{10}$$

Such a relation has to hold also for all nearby orbits. In particular for all nearby orbits starting at the same point of the boundary with different velocity vectors, i.e., at $x(\varphi) = (s_0, \varphi)$ with φ close to φ_0. But this is impossible because the function on the right hand side of (10) is not constant in any interval. $\square$

In the same fashion as it was done in [R], one can now argue that the set of periodic orbits of period 3 must have measure zero. Indeed, let us assume that the set of periodic points of period 3 has positive measure. If x_0 is the Lebesgue density point of the set of periodic points with period 3, then although T^3 may be different from identity in the neighborhood of x_0, DT^3 must be equal to identity at x_0. Hence the formulas (8) and (10) must hold for such a point. We obtain that for any Lebesgue density point $x = (s, \varphi)$

$$L(s, \varphi)k(s) - 2\sin^3\varphi = 0, \tag{11}$$

where $L(s, \varphi)$ is defined as the perimeter of the triangle with the vertices at the foot points of $T^{-1}x, x$ and Tx. Every periodic point with period 3 is the critical point of the function $L(s, \varphi)$ defined in this way.

It follows from (11) that at every Lebesgue density point of the set of periodic points of period 3 we must have also

$$\frac{\partial}{\partial\varphi}\left(L(s, \varphi)k(s) - 2\sin^3\varphi\right) = 0. \tag{12}$$

In view of the criticality of L at the periodic points, the partial derivative (12) is zero only if φ assumes values $0, \frac{\pi}{2}, \pi$. But this is impossible for an orbit of period 3. Hence the set of density points is empty, contrary to our assumption.

References

[B] M. Bialy, *Convex Billiards and a Theorem by E. Hopf*, preprint (1992).

[Bir] G.D. Birkhoff, *Dynamical Systems. Colloq. Pub., Vol IX*, Amer. Math. Soc., Providence, RI, 1927.

[D] V. Donnay, *Using integrability to produce chaos: billiards with positive entropy*, Comm.Math.Phys. **141** (1991), 225 - 257.

[H] M.R. Herman, *Sur les courbes invariantes par les difféomorphismes de l'anneau*, Astérisque **103 -104** (1983).

[L] V.F. Lazutkin, *The Convex Billiard and Eigenfunctions of the Laplace Operator*, Izd. Leningrad. Univ.,, St. Petersburg, 1981.

[M-M] S. Marvizi, R. Melrose, *Spectral Invariants of Convex Planar Regions*, J.Diff.Geom. **17** (1982), 475-502.

[R] M. R. Rychlik, *Periodic Points of the Billiard Ball Map in a Convex Domain*, J.Diff.Geom. **30** (1989), 191–205.

[S] L.A. Santalo, *Integral Geometry and Geometric Probability*, Addison-Wesley, Reading, MA, 1976.

[St] L. Stojanov, *Note on the Periodic Points of the Billiards*, J.Diff.Geom. **34** (1991), 835 - 837.

[W] M.P. Wojtkowski, *Principles for the design of billiards with nonvanishing Lyapunov exponents*, Comm. Math. Phys. **105** (1986), 391 – 414.