Ergod. Th. & Dynam. Sys., (2024), **44**, 1741–1756

doi:10.1017/etds.2023.70

Effective rigidity away from the boundary for centrally symmetric billiards

MISHA BIALY

School of Mathematical Sciences,
Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University,
Tel Aviv, Israel
(e-mail: bialy@tauex.tau.ac.il)

(*Received* 4 *October* 2022 *and accepted in revised form* 28 *August* 2023)

Abstract. In this paper, we study centrally symmetric Birkhoff billiard tables. We introduce a closed invariant set $\mathcal{M}_{\mathcal{B}}$ consisting of locally maximizing orbits of the billiard map lying inside the region $\mathcal{B}$ bounded by two invariant curves of 4-periodic orbits. We give an effective bound from above on the measure of this invariant set in terms of the isoperimetric defect of the curve. The equality case occurs if and only if the curve is a circle.

Key words: minimal action, integrable billiards, rigidity
2020 Mathematics Subject Classification: 37E40 (Primary); 37J35, 37C83 (Secondary)

1. *Introduction*

In this paper, we study Birkhoff billiards for centrally symmetric C^2-smooth strictly convex curves in the plane. We introduce the set $\mathcal{M}_{\mathcal{B}}$ lying in the region $\mathcal{B}$ between two invariant curves $\alpha, \bar{\alpha}$ in the phase space (see Figure 1). The set $\mathcal{M}_{\mathcal{B}}$, by definition, consists of those orbits such that any finite sub-segment is locally maximizing, for the length functional $\mathcal{L}$ associated to the billiard table. We assume that $\alpha, \bar{\alpha}$ consist of 4-periodic orbits of rotation numbers $1/4$ and $3/4$, respectively. It then follows that the set $\mathcal{M}_{\mathcal{B}}$ is a closed set which is invariant under the billiard map T. Our goal in this paper is to get an upper bound on the measure of the set $\mathcal{M}_{\mathcal{B}}$ which is sharp, that is, the case when $\mathcal{M}_{\mathcal{B}}$ occupies the whole of $\mathcal{B}$ occurs if and only if the billiard table is circular. Thus, we show that the measure of the complement set $\Delta_{\mathcal{B}} := \mathcal{B} \setminus \mathcal{M}_{\mathcal{B}}$ can be estimated from below in terms of the isoperimetric defect of the billiard domain.

These bounds are of obvious importance for classical dynamics (and probably also for quantum properties), because all 'rotational' invariant curves, as well as Aubry–Mather sets, are filled by orbits which are locally length maximizing (we refer to the monographs [**3**, **17**–**19**] for background material).

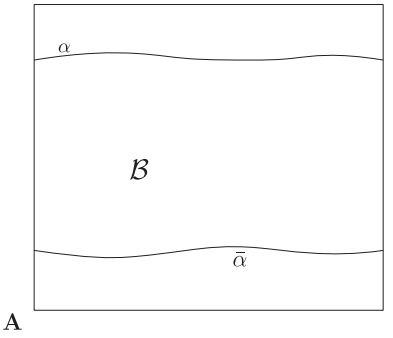

FIGURE 1. The region $\mathcal{B}$.

Estimates of this type were obtained previously in [6, 11], as an effective version of the so called the E. Hopf rigidity phenomenon for billiards.

The estimate presented here is related to the recent progress in the Birkhoff conjecture [9] for centrally symmetric billiard tables. Similarly to [9], we consider here the class of C^2-billiard tables having invariant curve consisting of 4-periodic orbits and use its properties. We refer here to papers [1, 2, 13–16] for other powerful recent approaches. However, the main novelty of the present paper is that the region $\mathcal{B}$ lies away from the boundary of the phase cylinder.

It is an open question how to remove the restriction of central symmetry of the billiard table. It is also interesting if effective bounds can be found for a region between two arbitrary invariant curves in the phase space.

We now turn to the needed background and the formulation of the main result. Let γ be a C^2-smooth simple closed convex curve of positive curvature in $\mathbb{R}^2$. We fix the counterclockwise orientation on γ. We shall use the arclength parametrization s as well as the parametrization by the angle ψ formed by the outer unit normal n to γ with a fixed direction. These two parametrizations are related by $d\psi = k(s)\,ds$, where $k(s)$ is the curvature at the point $\gamma(s)$.

The natural phase space of the Birkhoff billiard inside γ is the space $\mathbf{A}$ of all oriented lines that intersect γ. This space is topologically a cylinder and we shall refer to it as the phase cylinder of T. The billiard map T acts on $\mathbf{A}$ by the reflection law in γ. The phase cylinder carries a natural symplectic structure that can be described as follows.

Each oriented line is identified with the pair $(\cos\delta, s)$, $\delta \in (0, \pi)$, where $\gamma(s)$ is the incoming point and δ is the angle between the line and the tangent $\gamma'(s)$. In these coordinates, the symplectic form is $d\lambda$, where $\lambda = \cos\delta\, ds$ and $\cos\delta$ plays the role of momentum variable. We shall denote by μ the corresponding invariant measure on the phase space $\mathbf{A}$. The billiard map T is a symplectic map and the chord length $L(s, s_1) = |\gamma(s) - \gamma(s_1)|$ is a generating function of T (see Figure 2). Namely,

$$T^*\lambda - \lambda = \cos\delta_1\, ds_1 - \cos\delta\, ds = dL.$$

Moreover, one can check that T satisfies the twist condition:

$$L_{12}(s, s_1) > 0, \tag{1}$$

https://doi.org/10.1017/etds.2023.70 Published online by Cambridge University Press

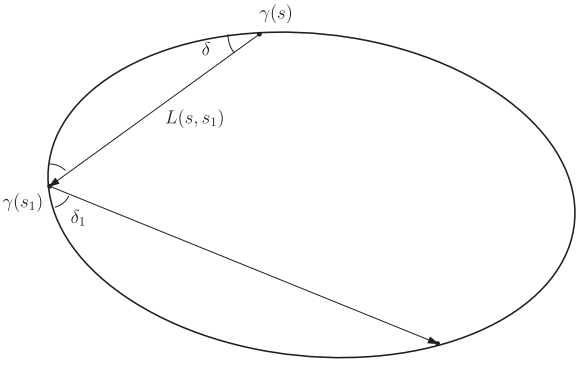

FIGURE 2. Generating function L corresponding to the 1-form λ.

meaning that T is a *negative* twist symplectic map (here and below, we use subindex 1 and/or 2 for the partial derivative with respect to the first or the second argument, respectively).

Remark. Traditionally, the generating function is the negative of ours, that is, the negative chord length. However, we prefer, for convenience, sign + for the generating function and hence the twist condition in equation (1) for the billiard map. Consequently, we deal with maximizing (and not minimizing) orbits.

For the generating function L, we can naturally define the variational principle as follows. For the configuration sequence $\{s_n\}$, we associate the formal sum

$$\mathcal{L}\{s_n\} = \sum_n L(s_n, s_{n+1}).$$

Configurations $\{s_n\}$, corresponding to billiard trajectories, are critical points of the functional $\mathcal{L}$.

We shall consider *locally maximizing* configurations, that is, those configurations which give *local maximum* for the functional between any two end-points. We shall call such configurations *m-configurations*, and the corresponding orbits on the phase cylinder $\mathbf{A}$, *m-orbits*. We denote by $\mathcal{M} \subset \mathbf{A}$ the set swept by all m-orbits corresponding to the variational principle for the generating function L. We shall also use the following notation:

$$\mathcal{M}_\mathcal{B} := \mathcal{M} \cap \mathcal{B}, \quad \Delta_\mathcal{B} := \mathcal{B} \setminus \mathcal{M}_\mathcal{B}.$$

Let $\gamma \subset \mathbb{R}^2$ be a C^2-smooth, centrally symmetric, convex closed curve of positive curvature. We shall assume that the billiard map corresponding to γ has a rotational (that is, winding once around the cylinder and simple) invariant curve $\alpha \subset \mathbf{A}$ consisting of 4-periodic orbits. We shall denote by $\bar{\alpha}$ the corresponding invariant curve of rotation number $\frac{3}{4}$. This curve consists of the same billiard trajectories but with the reversed orientation of the lines. Our main result is the following.

THEOREM 1.1. *Suppose that the billiard ball map T of γ has a continuous rotational invariant curve $\alpha \subset \mathbf{A}$ of rotation number $1/4$, consisting of 4-periodic orbits. Let $\bar{\alpha}$ be*

https://doi.org/10.1017/etds.2023.70 Published online by Cambridge University Press

the corresponding invariant curve of rotation number $\frac{3}{4}$. *Let* $\mathcal{B} \subset \mathbf{A}$ *be the domain between the curves* α *and* $\bar{\alpha}$ *(see Figure 1). Then the following estimate holds:*

$$\frac{3\beta}{16}(P^2 - 4\pi A) \leq \mu(\Delta_{\mathcal{B}}), \tag{2}$$

where P, A *denote the perimeter and the area of* γ, *and* $\beta > 0$ *is the minimal curvature of* γ.

Sharp estimates for $\mathcal{M}$ were obtained first in [6] and then in [11] as a quantitative version of the so called E.Hopf rigidity phenomenon for billiards discovered in [4] and then [5, 20]. In [11], the region between the invariant curve α and the boundary of the phase cylinder was considered, while in the present paper, the region $\mathcal{B}$ lies away from the boundary. The significance of the invariant curve of 4-$periodic$ orbits was first understood in [9], and we shall remember the properties of this curve here and use them below.

Here are some useful corollaries of Theorem 1.1.

COROLLARY 1.2. *Set* $\mathcal{M}_{\mathcal{B}}$ *of locally maximizing orbits occupies the whole region* $\mathcal{B}$ *if and only if* γ *is a circle.*

In fact, one can reformulate Corollary 1.2 in a dynamical way.

COROLLARY 1.3. *Suppose that the restriction of billiard map* T *to* $\mathcal{B}$ *has an invariant measurable field of non-vertical oriented lines, with the orientation chosen on the lines coherently by the condition* $ds > 0$. *Then* γ *is a circle.*

This is especially useful in establishing the following geometric fact.

COROLLARY 1.4. *If* γ *is not a circle, then there always exist a point* $x \in \mathcal{B}$ *and a vertical tangent vector* $v \in T_x\mathcal{B}$ *such that for some positive integer n, the vector* $DT^n(v)$ *is again vertical (this exactly means that the points x and* $T^n x$ *are conjugate).*

Corollary 1.3 follows immediately from Theorem 1.1 applying the criterion of local maximality in terms of Jacobi fields [11, Theorem 1.1].

To deduce Corollary 1.4, one can argue analogously to [4]. More precisely, suppose, by contradiction, that for any vertical vector $v \in T_x\mathcal{B}$ and any positive integer n, the vector $DT^n(v)$ is not vertical. This implies that any finite segment of a billiard trajectory $\{\gamma(s_n), n \in [M, N]\}$ has a non-degenerate matrix of second variation $\delta^2\mathcal{L}_{MN}$. Then, by a continuity argument, all the matrices $\delta^2\mathcal{L}_{MN}$ must be negative definite (because this holds true for orbits lying on the rotational invariant curve α). Hence, all billiard configurations, corresponding to the orbits lying in $\mathcal{B}$, are locally maximizing. Therefore, Theorem 1.1 applies and the curve γ is a circle, which is a contradiction.

2. *Important tools*

2.1. *Non-standard generating function.* Another way to get the same symplectic form is to fix an origin in $\mathbb{R}^2$ (we shall fix the origin at the center of γ) and to introduce the coordinates (p, φ) on the space of all oriented lines, so that φ is the angle between the right unit normal to the line and the horizontal, and p is the signed distance to

https://doi.org/10.1017/etds.2023.70 Published online by Cambridge University Press

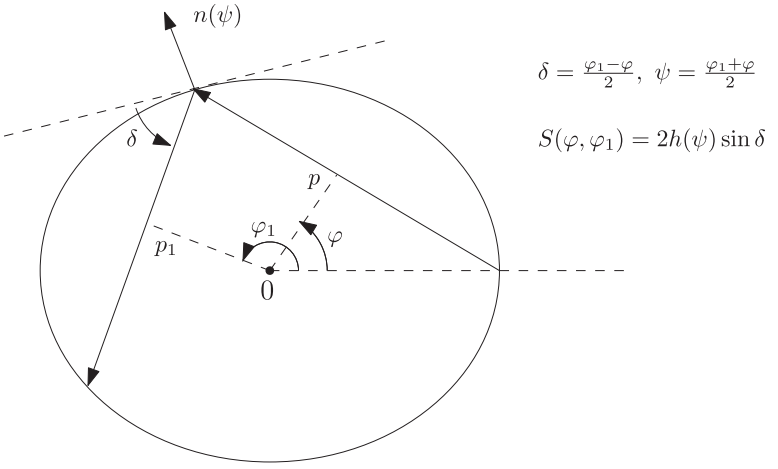

FIGURE 3. Generating function S corresponding to the 1-form β.

the line (see Figure 3). In this way, the space of oriented lines is identified with T^*S^1. Moreover, the standard symplectic form $d\beta$ with $\beta = pd\varphi$ coincides with the symplectic form described before. In this description, p plays the role of momentum variable.

For the second choice of the coordinates (p, φ), the generating function was found first in [8] for the two-dimensional case and then in [7] for higher dimensions (see [10] for further applications). This function S is determined by the formulas:

$$T^*\beta - \beta = p_1\, d\varphi_1 - p\, d\varphi = dS, \quad S(\varphi, \varphi_1) = 2h(\psi)\sin\delta,$$

where

$$\psi := \frac{\varphi_1 + \varphi}{2}, \quad \delta := \frac{\varphi_1 - \varphi}{2}.$$

Here and throughout this paper, we denote by h the support function of γ with respect to 0:

$$h(\psi) := \max_{\gamma} \langle \gamma, n_\psi \rangle,$$

where n_ψ is the unit outer normal to γ in the direction ψ. The fact that S is the generating function for T means that the line with coordinates (p, φ) is mapped into the line (p_1, φ_1) (see Figure 3) if and only if

$$\begin{aligned} p &= -S_1(\varphi, \varphi_1) = h(\psi)\cos\delta - h'(\psi)\sin\delta, \\ p_1 &= S_2(\varphi, \varphi_1) = h(\psi)\cos\delta + h'(\psi)\sin\delta. \end{aligned} \tag{3}$$

It follows from the direct computation (see below Proposition 2.4) that the map T satisfies the twist condition with respect to the symplectic coordinates (p, φ) meaning that the cross-derivative satisfies $S_{12} = \frac{1}{2}\rho(\psi)\sin\delta > 0$, where $\rho(\psi) = h''(\psi) + h(\psi) > 0$ is the radius of curvature.

https://doi.org/10.1017/etds.2023.70 Published online by Cambridge University Press

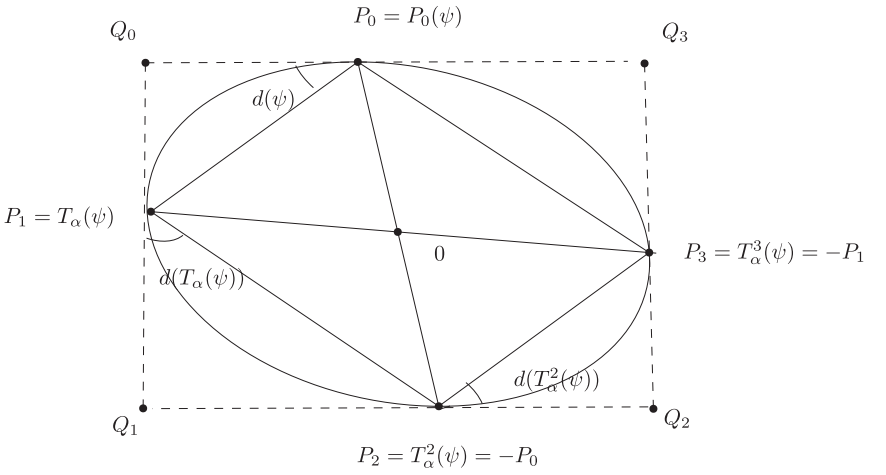

FIGURE 4. Rectangle $Q_0Q_1Q_2Q_3$ corresponding to the 4-periodic orbit forming parallelogram $P_0P_1P_2P_3$.

2.2. *Two variational principles.* One can associate variational principle $\mathcal{S}$ also for the function S:

$$\mathcal{S}\{\varphi_n\} = \sum_n S(\varphi_n, \varphi_{n+1}).$$

In [11], we gave a criterion for an orbit to be locally maximizing. It then follows from this criterion that the set $\mathcal{M}$ does not depend on which generating function L or S is used for the map T. We shall use the function S to prove Theorem 1.1.

Remark. It appears that vertical vector in the statement of the Corollary 1.4 can be understood with respect to each of the vertical foliations $\{s = \text{const}\}$ or $\{\varphi = \text{const}\}$. This follows from the proof of Corollary 1.4 and the fact, proven in [11], that the classes of locally maximizing orbits corresponding to the generating functions L, S coincide.

In particular, the existence of conjugate points with respect to the vertical foliation $\{\varphi = \text{const}\}$ implies that one can find a beam of parallel lines such that after n reflections, the beam becomes parallel (infinitesimally) again.

2.3. *Properties of the invariant curve of* 4*-periodic orbits.* If the billiard curve γ is an ellipse, then there exists a rotational invariant curve α consisting of 4-periodic orbits. The corresponding quadrilaterals inscribed in γ are called Poncelet 4-gons. It is well known (see [12] for several proofs) that all Poncelet 4-gons for an ellipse are parallelograms. This fact can be generalized from the case of an ellipse to any centrally symmetric billiard table. We now turn to state the results from [9] and refer to [9] for the proofs. The next theorem is illustrated in Figure 4.

THEOREM 2.1. *Let γ be a centrally symmetric billiard table. Assume that billiard ball map $T : \mathbf{A} \to \mathbf{A}$ has a continuous rotational invariant curve $\alpha = \{\delta = d(\psi)\}$ of rotation number $\frac{1}{4}$ consisting of* 4*-periodic orbits of T. Then the following properties hold.*

https://doi.org/10.1017/etds.2023.70 Published online by Cambridge University Press

(A) *Function $d(\psi)$ is π-periodic and the billiard quadrilaterals corresponding to the traces of the orbits contained in the invariant curve α are parallelograms.*
(B) *The tangent lines to γ at the vertices of the parallelogram form a rectangle.*
(C) $0 < d(\psi) < \pi/2, \quad d(\psi + \pi/2) = \pi/2 - d(\psi)$.
(D) *The functions d and h satisfy the identities*

$$\tan d(\psi) = \frac{h(\psi)}{h(\psi + \pi/2)} = -\frac{h'(\psi + \pi/2)}{h'(\psi)},$$

and

$$h^2(\psi) + h^2\left(\psi + \frac{\pi}{2}\right) = R^2 = \text{const}.$$

Remark. It follows from Theorem 2.1 item (D) that the orthoptic curve associated with γ is a circle of radius R (like in the case of an ellipse). Here the orthoptic curve of γ, by definition, is the locus of points Q, such that the two tangents to γ passing through Q form a right angle.

COROLLARY 2.2. *Let γ be a convex centrally symmetric billiard table. Let $\alpha = \{\delta = d(\psi)\} \subset \mathbf{A}$ be an invariant curve consisting of 4-periodic orbits. It then follows from Theorem 2.1 item (D) that*

$$h(\psi) = R \sin d(\psi), \quad h\left(\psi + \frac{\pi}{2}\right) = R \cos d(\psi),$$

for a positive constant R.

COROLLARY 2.3. *The explicit formulas of item (D) show that the invariant curve α is necessarily C^2-smooth, since the support function h is C^2-smooth by assumption.*

2.4. *Function ω and an inequality.* It turns out that one can introduce a measurable bounded function ω on the set $\mathcal{M}$ satisfying the inequality:

$$\omega(p_1, \varphi_1) - \omega(p, \varphi) \geq S_{11}(\varphi, \varphi_1) + S_{22}(\varphi, \varphi_1) + 2S_{12}(\varphi, \varphi_1). \tag{4}$$

The construction of this function (see [4]) was inspired by the celebrated E.Hopf theorem on tori with no conjugate points. Let us sketch this construction. Let $\{(p_n, \varphi_n)\}$ be a locally maximizing orbit of the point $z = (p_0, \varphi_0)$. It then follows that there exists an invariant vector field $\{(\delta p_n, \delta\varphi_n)\}$ along the orbit $\{(p_n, \varphi_n)\}$ such that the corresponding field $\delta\varphi_n$ is a Jacobi field along the billiard configuration $\{\varphi_n\}$ (normalized by $\delta\varphi_0 = 1$) and is strictly positive. Remember, a *Jacobi field* along a configuration $\{\varphi_n\}$ is a sequence $\{\delta\varphi_n\}$ satisfying the *discrete Jacobi equation*:

$$b_{n-1}\delta\varphi_{n-1} + a_n\delta\varphi_n + b_n\delta\varphi_{n+1} = 0, \tag{5}$$

where, as before,

$$a_n = S_{22}(\varphi_{n-1}, \varphi_n) + S_{11}(\varphi_n, \varphi_{n+1}), \; b_n = S_{12}(\varphi_n, \varphi_{n+1}).$$

https://doi.org/10.1017/etds.2023.70 Published online by Cambridge University Press

Then the invariance of the field $\{(\delta p_n, \delta\varphi_n)\}$ along the orbit implies (by differentiating the formula $p_n = -S_1(\varphi_n, \varphi_{n+1})$):

$$\delta p_n = -S_{11}(\varphi_n, \varphi_{n+1})\delta\varphi_n - S_{12}(\varphi_n, \varphi_{n+1})\delta\varphi_{n+1},$$

or equivalently, due to the Jacobi equation:

$$\delta p_n = S_{22}(\varphi_{n-1}, \varphi_n)\delta\varphi_n + S_{12}(\varphi_{n-1}, \varphi_n)\delta\varphi_{n-1}.$$

Then one defines $\omega(p_n, \varphi_n) := \delta p_n/\delta\varphi_n$. One can prove that ω is a measurable function and satisfies the relations:

$$\begin{cases} \omega(T(p, \varphi)) = S_{22}(\varphi, \varphi_1) + S_{12}(\varphi, \varphi_1)\delta\varphi_1(\varphi, p)^{-1}, \\ \omega(p, \varphi) = -S_{11}(\varphi, \varphi_1) - S_{12}(\varphi, \varphi_1)\delta\varphi_1(\varphi, p). \end{cases} \tag{6}$$

Subtracting the second equation from the first one and using $S_{12} > 0$, $\delta\varphi_1 > 0$, we get the inequality in equation (4).

Also notice that from equation (6), we have the inequality

$$S_{22}(\varphi_{-1}, \varphi) < \omega(p, \varphi) < -S_{11}(\varphi, \varphi_1),$$

since in equation (6), $S_{12}, \delta\varphi_1, \delta\varphi_{-1}$ are positive. Using Proposition (2.4), it then follows that function ω is bounded on $\mathcal{M}_{\mathcal{B}}$:

$$|\omega| < \max_{\mathcal{B}}\{|S_{11}|, |S_{22}|\} < K(\gamma),$$

where $K(\gamma)$ depends only on γ (for example, one can set $K(\gamma) = \max_{\gamma}\{\rho + h + |h'|\}$, using the formulas of Proposition (2.4)).

2.5. *Derivatives of generating function S.* The derivatives of the generating function S can be immediately computed.

PROPOSITION 2.4. *The second partial derivatives of S are*

$$\begin{aligned} S_{11}(\varphi, \varphi_1) &= \tfrac{1}{2}(h''(\psi) - h(\psi))\sin\delta - h'(\psi)\cos\delta; \\ S_{22}(\varphi, \varphi_1) &= \tfrac{1}{2}(h''(\psi) - h(\psi))\sin\delta + h'(\psi)\cos\delta; \\ S_{12}(\varphi, \varphi_1) &= \tfrac{1}{2}(h''(\psi) + h(\psi))\sin\delta, \end{aligned}$$

where $\psi := (\varphi_1 + \varphi)/2, \quad \delta := (\varphi_1 - \varphi)/2.$

3. *Proof of Theorem 1.1*

In the following, we shall work with the coordinates (p, φ) and the function ω constructed above for the generating function S. We start the proof of Theorem 1.1 integrating equation (4) over $\mathcal{M}_{\mathcal{B}}$ with respect to the invariant measure $d\mu = dpd\varphi$.

To perform the integration, we compute the invariant measure as follows.

The symplectic form $dp \wedge d\varphi$ can be written using generating function in equation (3):

$$dp \wedge d\varphi = -d(S_1(\varphi, \varphi_1)) \wedge d\varphi = S_{12}d\varphi \wedge d\varphi_1.$$

https://doi.org/10.1017/etds.2023.70 Published online by Cambridge University Press

Since T is symplectic, the measure

$$d\mu = dpd\varphi = S_{12}d\varphi d\varphi_1$$

is invariant. Using the explicit formula for the second derivative (Proposition 2.4), we compute

$$d\mu = S_{12}d\varphi d\varphi_1 = \left(\tfrac{1}{2}\rho(\psi)\sin\delta\right)d\varphi d\varphi_1 = \rho(\psi)\sin\delta d\psi d\delta,$$

where (see Figure 3 showing all the notation)

$$\rho(\psi) = h''(\psi) + h(\psi)$$

is the radius of curvature of γ, and

$$\psi := \frac{\varphi_1 + \varphi}{2}, \quad \delta := \frac{\varphi_1 - \varphi}{2}.$$

Hence, integrating the inequality in equation (4) with respect to the invariant measure $d\mu$, we obtain

$$0 \geq \int_{\mathcal{M}_\mathcal{B}} [S_{11}(\varphi, \varphi_1) + 2S_{12}(\varphi, \varphi_1) + S_{22}(\varphi, \varphi_1)]d\mu.$$

Moreover, we get from Proposition 2.4, after obvious simplifications,

$$S_{11}(\varphi, \varphi_1) + 2S_{12}(\varphi, \varphi_1) + S_{22}(\varphi, \varphi_1) = 2h''(\psi)\sin\delta.$$

Thus, equation (4) yields the inequality:

$$0 \geq \int_{\mathcal{M}_\mathcal{B}} [(h''(\psi)\sin\delta]d\mu. \tag{7}$$

Since $\mathcal{M}_\mathcal{B} = \mathcal{B} \setminus \Delta_\mathcal{B}$, we get

$$\int_\mathcal{B} [(h''(\psi)\sin\delta]\, d\mu \leq \int_{\Delta_\mathcal{B}} [(h''(\psi)\sin\delta]\, d\mu. \tag{8}$$

Let us denote

$$I := \int_\mathcal{B} [(h''(\psi)\sin\delta]\, d\mu. \tag{9}$$

We shall give an upper bound for the right-hand side of equation (8), and a lower bound on the left-hand side I, and together we get the required bound. For the right-hand side of equation (8), write

$$\begin{aligned}\int\limits_{\Delta_\mathcal{B}} h''(\psi)\sin\delta d\mu &\leq \left|\int\limits_{\Delta_\mathcal{B}} h''(\psi)\sin\delta d\mu\right| \\ &\leq \int\limits_{\Delta_\mathcal{B}} |h''(\psi)\sin\delta| d\mu \leq \mu(\Delta_\mathcal{B}) \max_{\Delta_\mathcal{B}} |h''|.\end{aligned} \tag{10}$$

Since $h(\psi) + h''(\psi) = \rho(\psi)$, where $\rho(\psi)$ is the radius of curvature, then

$$|h''| \leq \rho + \max h.$$

https://doi.org/10.1017/etds.2023.70 Published online by Cambridge University Press

Since γ is centrally symmetric, we have $\max h \leq D/2$, where D is the diameter. Also, the maximal radius of curvature of γ is $1/\beta$, where β is the minimal curvature of γ. This gives us the estimate

$$I \leq \int\limits_{\Delta_{\mathcal{B}}} h''(\psi) \sin\delta \, d\mu \leq \left(\frac{D}{2} + \frac{1}{\beta}\right)\mu(\Delta_{\mathcal{B}}) \leq \frac{2}{\beta}\mu(\Delta_{\mathcal{B}}), \tag{11}$$

where we used Blaschke's rolling disk theorem, stating that γ is contained inside a disk with radius equal to the maximal radius of curvature of γ, and this means that $D \leq 2/\beta$.

We now turn to estimate I from below. Namely, we shall prove in the next section the following.

THEOREM 3.1. *Integral I can be estimated from below:*

$$I \geq \tfrac{3}{8}(P^2 - 4\pi A).$$

Proof of Theorem 1.1 follows immediately from equation (11) and Theorem 3.1.

4. *Proof of Theorem 3.1*

Substituting into the integral I the explicit expression $d\mu = \rho(\psi)\sin\delta \, d\psi \, d\delta$ and integrating first with respect to δ, we get from equation (9):

$$I = \int_0^{2\pi} d\psi \left[h''(h + h'') \int_{d(\psi)}^{\pi - d(\psi)} d\delta \, \sin^2\delta \right]. \tag{12}$$

Here we used the fact that in the coordinates (ψ, δ), the domain of integration takes the form

$$\mathcal{B} = \{(\psi, \delta) : \psi \in [0, 2\pi], \ \delta \in [d(\psi), \pi - d(\psi)]\}.$$

Here and below, $d(\psi)$ is the function described in §2.3.

Integrating in equation (12) with respect to δ, we obtain

$$I = \int_0^{2\pi} \left[h''(\psi)(h''(\psi) + h(\psi))\right]\left(\frac{\pi}{2} - d(\psi) + \frac{1}{2}\sin 2d(\psi)\right) d\psi. \tag{13}$$

Now we substitute into equation (13) the expressions for h, h', h'' via $d(\psi)$ using Corollary 2.2 of Theorem 2.1:

$$\begin{cases} h = R\sin d, \\ h' = R\cos d \, d', \\ h'' = R\cos d \, d'' - R\sin d \, (d')^2, \\ d\left(\psi + \dfrac{\pi}{2}\right) = \dfrac{\pi}{2} - d(\psi). \end{cases} \tag{14}$$

In what follows, we usually omit the arguments for the functions h, d and their derivatives.

https://doi.org/10.1017/etds.2023.70 Published online by Cambridge University Press

Thus, we get from equation (11) the following equality on the function d:

$$\begin{aligned} I &= R^2 \int_0^{2\pi} (\sin d - \sin d\, d'^2 + \cos d\, d'')(-\sin d\, d'^2 + \cos d\, d'') \\ &\quad \times \left(\frac{\pi}{2} - d + \frac{1}{2}\sin 2d\right) d\psi \\ &= R^2 \int_0^{2\pi} U\, d\psi, \end{aligned} \tag{15}$$

where we introduced U by the formula

$$U := (\sin d - \sin d\, d'^2 + \cos d\, d'')(-\sin d\, d'^2 + \cos d\, d'') \left(\frac{\pi}{2} - d + \frac{1}{2}\sin 2d\right).$$

The assumption of central symmetry implies that $h(\psi)$, $d(\psi)$ are π-periodic. Hence,

$$\int_0^{2\pi} U(\psi)\, d\psi = 2\int_0^{\pi} U(\psi)\, d\psi.$$

We shall prove now the following estimate.

THEOREM 4.1.

$$\int_0^{\pi} U(\psi)\, d\psi \geq \frac{3}{16R^2}(P^2 - 4\pi A).$$

Proof. The idea of the proof is to proceed in three steps: 'symmetrization', integration by parts, and Wirtinger inequality. Doing this, we pass to a new integrand, $\tilde{U}$, satisfying the inequality $\tilde{U} \geq \text{const}\, h'^2$. Moreover, integrating this inequality, we will be able to estimate the integral of $\tilde{U}$ from below by isoperimetric defect.

We write

$$U = U_1 + U_2 + U_3 + U_4 + U_5,$$

where

$$\begin{aligned} U_1 &= d''^2 \cos^2 d\left(\frac{\pi}{2} - d + \frac{1}{2}\sin 2d\right), \\ U_2 &= -2d''d'^2 \sin d \cos d\left(\frac{\pi}{2} - d + \frac{1}{2}\sin 2d\right), \\ U_3 &= d'' \sin d \cos d\left(\frac{\pi}{2} - d + \frac{1}{2}\sin 2d\right), \\ U_4 &= d'^4 \sin^2 d\left(\frac{\pi}{2} - d + \frac{1}{2}\sin 2d\right), \\ U_5 &= -d'^2 \sin^2 d\left(\frac{\pi}{2} - d + \frac{1}{2}\sin 2d\right). \end{aligned}$$

Step 1. Symmetrization. We perform the change of the integration variable by the rule $\psi \to \psi + \pi/2$. By equation (14), which is the consequence of Theorem 2.1 and

https://doi.org/10.1017/etds.2023.70 Published online by Cambridge University Press

Corollary 2.2, this intertwines $\sin(d)$ with $\cos(d)$ and changes the sign of d''. Denote the changed integrand by $\hat{U}_j$.

Also denote the 'symmetrized' integrand by

$$V_j := U_j + \hat{U}_j.$$

Then we have

$$\int_0^\pi U_j(\psi)\, d\psi = \int_0^\pi \hat{U}_j(\psi)\, d\psi = \frac{1}{2}\int_0^\pi V_j(\psi)\, d\psi,$$

where V_j can be written as

$$\begin{aligned}
V_1 &= d''^2\left(\frac{\pi}{4} + \left(\frac{\pi}{4} - d\right)\cos 2d + \frac{1}{2}\sin 2d\right),\\
V_2 &= d''d'^2 \sin 2d\left(2d - \frac{\pi}{2}\right),\\
V_3 &= d'' \sin 2d\left(\frac{\pi}{4} - d\right),\\
V_4 &= d'^4\left(\frac{\pi}{4} + \left(d - \frac{\pi}{4}\right)\cos 2d + \frac{1}{2}\sin 2d\right),\\
V_5 &= -d'^2\left(\frac{\pi}{4} + \left(d - \frac{\pi}{4}\right)\cos 2d + \frac{1}{2}\sin 2d\right).
\end{aligned}$$

Step 2. Integration by parts. Terms V_2 and V_3 contain d'' in the first power. Therefore, we apply integration by parts for V_2, V_3 to get rid of the second derivative d''. Notice that thanks to the π-periodicity of the integrands, the off-integration terms vanish. Thus, we get new integrands W_i, $i = 1, \ldots, 5$, where

$$\begin{aligned}
W_1 &= V_1 = d''^2\left(\frac{\pi}{4} + \left(\frac{\pi}{4} - d\right)\cos 2d + \frac{1}{2}\sin 2d\right),\\
W_2 &= d'^4\left(-\frac{4}{3}\cos 2d\left(d - \frac{\pi}{4}\right) - \frac{2}{3}\sin 2d\right),\\
W_3 &= d'^2\left(2\cos 2d\left(d - \frac{\pi}{4}\right) + \sin 2d\right),\\
W_4 &= V_4 = d'^4\left(\frac{\pi}{4} + \left(d - \frac{\pi}{4}\right)\cos 2d + \frac{1}{2}\sin 2d\right),\\
W_5 &= V_5 = -d'^2\left(\frac{\pi}{4} + \left(d - \frac{\pi}{4}\right)\cos 2d + \frac{1}{2}\sin 2d\right).
\end{aligned}$$

Thus, we get for the integral of U:

$$\int_0^\pi U\, d\psi = \int_0^\pi \sum_{i=1}^5 U_i\, d\psi = \frac{1}{2}\int_0^\pi \sum_{i=1}^5 V_i\, d\psi = \frac{1}{2}\int_0^\pi \sum_{i=1}^5 W_i\, d\psi. \qquad (16)$$

https://doi.org/10.1017/etds.2023.70 Published online by Cambridge University Press

Summing $W_2 + W_4$ and $W_3 + W_5$, we rewrite using only three summands:

$$X_1 := W_1 = d'^2\left(\frac{\pi}{4} + \left(\frac{\pi}{4} - d\right)\cos 2d + \frac{1}{2}\sin 2d\right),$$
$$X_2 := W_2 + W_4 = d'^4\left(\frac{\pi}{4} - \frac{1}{3}\left(d - \frac{\pi}{4}\right)\cos 2d - \frac{1}{6}\sin 2d\right),$$
$$X_3 := W_3 + W_5 = d'^2\left(-\frac{\pi}{4} + \left(d - \frac{\pi}{4}\right)\cos 2d + \frac{1}{2}\sin 2d\right).$$

Thus, we have

$$\int_0^\pi U\, d\psi = \frac{1}{2}\int_0^\pi (X_1 + X_2 + X_3)\, d\psi. \tag{17}$$

Step 3. Use of the Wirtinger inequality. Let us introduce the function of d which is the multiplier in X_1:

$$f(d) := \frac{\pi}{4} + \left(\frac{\pi}{4} - d\right)\cos 2d + \frac{1}{2}\sin 2d.$$

This function is strictly positive since d varies in $(0, \pi/2)$. In fact, one can say more precisely

$$f \in \left[\frac{1}{2} + \frac{\pi}{4}, \frac{\pi}{2}\right).$$

Also we can write

$$X_2 = d'^4 f_2, \quad f_2 := \frac{\pi}{4} - \frac{1}{3}\left(d - \frac{\pi}{4}\right)\cos 2d - \frac{1}{6}\sin 2d,$$

and one can see that f_2 is positive as well.

Similarly for X_3, we have

$$X_3 = d'^2 f_3, \quad f_3 := -\frac{\pi}{4} + \left(d - \frac{\pi}{4}\right)\cos 2d + \frac{1}{2}\sin 2d = (\sin 2d - f).$$

However, the function f_3 is not necessarily positive. To bypass this difficulty, we shall use the Wirtinger inequality, which we apply to the function

$$Y := d'\sqrt{f}.$$

Notice that Y is π-periodic and has zero average, since it can be written as a complete derivative. Hence,

$$\int_0^\pi (Y'^2 - 4Y^2)\, d\psi \geq 0.$$

We have the following expressions:

$$Y' = \sqrt{f}d'' + \frac{f'}{2\sqrt{f}}d'^2 \Rightarrow Y'^2 = f\, d''^2 + f'\, d''d'^2 + \frac{f'^2}{4f}d'^4.$$

https://doi.org/10.1017/etds.2023.70 Published online by Cambridge University Press

Therefore,

$$\int_0^\pi (Y'^2 - 4Y^2)\, d\psi = \int_0^\pi \left(f\, d''^2 + f'\, d'' d'^2 + \frac{f'^2}{4f} d'^4 - 4f d'^2 \right) d\psi$$
$$= \int_0^\pi \left(f\, d''^2 - \frac{f''}{3} d'^4 + \frac{f'^2}{4f} d'^4 - 4f d'^2 \right) d\psi = \int_0^\pi g\, d\psi \geq 0,$$

where $g := (f\, d''^2 - f'' 3d'^4 + (f'^2/4f)d'^4 - 4f d'^2)$ and we performed integration by parts again.

Thus, finally we can write

$$X_1 + X_2 + X_3 = g + d'^4 \left(f_2 + \frac{f''}{3} - \frac{f'^2}{4f} \right) + (f_3 + 4f)$$
$$= g + d'^4 \left(f_2 + \frac{f''}{3} - \frac{f'^2}{4f} \right) + d'^2 (\sin 2d + 3f).$$

The following claim is crucial.

LEMMA 4.2. *Both expressions* $(\sin 2d + 3f)$ *and* $(f_2 + f''/3 - f'^2/4f)$ *of the last formula are strictly positive.*

Proof. (1) Since $f \in [\frac{1}{2} + \pi/4, \pi/2)$, then $(3f + \sin 2d) \geq \frac{3}{2} + 3\pi/4$. Analyzing the behavior of the function f, one can claim more:

$$(3f + \sin 2d) \geq 3f(0) = \frac{3\pi}{2}. \tag{18}$$

(2) For the expression $(f_2 + f''/3 - f'^2/4f)$, we need to compute

$$f' = -2 \left(\frac{\pi}{4} - d \right) \sin 2d,$$

$$f'' = -4 \left(\frac{\pi}{4} - d \right) \cos 2d + 2 \sin 2d.$$

We substitute f_2 and the second derivative of f:

$$\left(f_2 + \frac{f''}{3} - \frac{f'^2}{4f} \right) = \frac{f''}{3} - \frac{f'^2}{4f} + \frac{\pi}{4} - \frac{1}{3} \left(d - \frac{\pi}{4} \right) \cos 2d - \frac{1}{6} \sin 2d$$
$$= -\frac{f'^2}{4f} + \frac{1}{3} \left[-4 \left(\frac{\pi}{4} - d \right) \cos 2d + 2 \sin 2d \right] + \frac{\pi}{4} - \frac{1}{3} \left(d - \frac{\pi}{4} \right) \cos 2d - \frac{1}{6} \sin 2d$$
$$= -\frac{f'^2}{4f} + \frac{\pi}{4} + \left(d - \frac{\pi}{4} \right) \cos 2d + \frac{1}{2} \sin 2d.$$

Thus, we need to check the sign of the expression:

$$-\left(\frac{\pi}{4} - d \right)^2 \sin^2 2d + \left(\frac{\pi}{4} + \left(\frac{\pi}{4} - d \right) \cos 2d + \frac{1}{2} \sin 2d \right)$$
$$\times \left(\frac{\pi}{4} + \left(d - \frac{\pi}{4} \right) \cos 2d + \frac{1}{2} \sin 2d \right)$$

https://doi.org/10.1017/etds.2023.70 Published online by Cambridge University Press

$$= -\left(d - \frac{\pi}{4}\right)^2 + \frac{\pi^2}{4^2} + \frac{1}{4}\sin^2 2d + \frac{\pi}{4}\sin 2d$$
$$= -\left(d - \frac{\pi}{4}\right)^2 + \left(\frac{\pi}{4} + \frac{1}{2}\sin 2d\right)^2.$$

Notice that since $d \in (0, \pi/2)$, then $|d - \pi/4| < \pi/4$ and hence the last expression is strictly positive. This completes the proof of Lemma 4.2. □

We are now in position to finish the proof of Theorem 4.1. Using Lemma 4.2, we can deduce from equation (17) with the help of equation (18),

$$2\int_0^\pi U\, d\psi \geq \int_0^\pi d'^2(\sin 2d + 3f)\, d\psi \geq \int_0^\pi \frac{3\pi}{2} d'^2\, d\psi$$
$$\geq \int_0^\pi \frac{3\pi}{2}\cos^2 d\, d'^2\, d\psi = \frac{3\pi}{2R^2}\int_0^\pi h'^2\, d\psi, \tag{19}$$

where we used $h' = R\cos d\, d'$ of equation (14) in the last equality.

Now consider the isoperimetric defect $P^2 - 4\pi A$ for the curve γ. We have the classical formulas:

$$P = \int_0^{2\pi} h\, d\psi, \quad A = \frac{1}{2}\int_0^{2\pi}(h^2 - h'^2)\, d\psi.$$

By Cauchy–Schwartz inequality, we have

$$P^2 \leq 2\pi\int_0^{2\pi} h^2\, d\psi = 2\pi\left(2A + \int_0^{2\pi} h'^2\, d\psi\right).$$

Hence, using equation (19), we get

$$P^2 - 4\pi A \leq 2\pi\int_0^{2\pi} h'^2\, d\psi = 4\pi\int_0^\pi h'^2\, d\psi \leq \frac{16}{3}R^2\int_0^\pi U\, d\psi.$$

This completes the proof of Theorem 4.1. □

5. *Discussion*

It is very natural to ask if one can reconstruct elliptic billiards by sharp inequalities containing the measures $\mu(\Delta_{\mathcal{B}})$, $\mu(\mathcal{M}_{\mathcal{B}})$ (similarly to Theorem 1.1).

It would be very interesting to extend the ideas used in this paper to other Hamiltonian systems such as twist symplectic maps, as well as to continuous time systems.

An important goal in the study of Birkhoff billiards, as well as of general twist maps, in particular of standard-like maps, is to understand the dynamical behavior between two invariant curves. Our result can be considered as a step in this direction. It is not clear, however, how to approach this goal for arbitrary invariant curves and also how to remove the central-symmetry assumption.

Acknowledgements. MB was partially supported by ISF grant 580/20 and DFG grant MA-2565/7-1 within the Middle East Collaboration Program.

References

[1] M. Arnold and M. Bialy. Nonsmooth convex caustics for Birkhoff billiards. *Pacific J. Math.* **295**(2) (2018), 257–269.

[2] A. Avila, V. Kaloshin and J. De Simoi. An integrable deformation of an ellipse of small eccentricity is an ellipse. *Ann. of Math. (2)* **184** (2016), 527–558.

[3] V. Bangert. Mather set for twist maps and Geodesics on Tori. *Dynamics Reported (Dynamics Reported. New Series, 1)*. Eds. U. Kirchgraber and H.-O. Walther. Wiley, Chichester, 1988, pp. 1–56.

[4] M. Bialy. Convex billiards and a theorem by E. Hopf. *Math. Z.* **214**(1) (1993), 147–154.

[5] M. Bialy. Hopf rigidity for convex billiards on the hemisphere and hyperbolic plane. *Discrete Contin. Dyn. Syst.* **33**(9) (2013), 3903–3913.

[6] M. Bialy. Effective bounds in E. Hopf rigidity for billiards and geodesic flows. *Comment. Math. Helv.* **90**(1) (2015), 139–153.

[7] M. Bialy. Gutkin billiard tables in higher dimensions and rigidity. *Nonlinearity* **31** (2018), 2281–2293.

[8] M. Bialy and A. E. Mironov. Angular billiard and algebraic Birkhoff conjecture. *Adv. Math.* **313** (2017), 102–126.

[9] M. Bialy and A. E. Mironov. The Birkhoff–Poritsky conjecture for centrally-symmetric billiard tables. *Ann. of Math. (2)* **196**(1) (2022), 389–413.

[10] M. Bialy and S. Tabachnikov. Dan Reznik identities and more. *Eur. J. Math.* **8**(4) (2022), 1341–1354.

[11] M. Bialy and D. Tsodikovich. Locally maximizing orbits for the non-standard generating function of convex billiards and applications. *Nonlinearity* **36**(3) (2023), 2001–2019.

[12] A. Connes and D. Zagier. A property of parallelograms inscribed in ellipses. *Amer. Math. Monthly* **114** (2007), 909–914.

[13] A. Glutsyuk. On polynomially integrable Birkhoff billiards on surfaces of constant curvature. *J. Eur. Math. Soc. (JEMS)* **23**(3) (2021), 995–1049.

[14] G. Huang, V. Kaloshin and A. Sorrentino. Nearly circular domains which are integrable close to the boundary are ellipses. *Geom. Funct. Anal.* **28**(2) (2018), 334–392.

[15] V. Kaloshin and A. Sorrentino. On the local Birkhoff conjecture for convex billiards. *Ann. of Math. (2)* **188** (2018), 315–380.

[16] V. Kaloshin and A. Sorrentino. On the integrability of Birkhoff billiards. *Philos. Trans. Roy. Soc. A* **376** (2018), 20170419; doi:10.1098/rsta.2017.0419.

[17] V. V. Kozlov and D. V. Treshchëv. *Billiards. A Genetic Introduction to the Dynamics of Systems with Impacts (Translations of Mathematical Monographs, 89)*. American Mathematical Society, Providence, RI, 1991.

[18] K. F. Siburg. *The Principle of Least Action in Geometry and Dynamics (Lecture Notes in Mathematics, 1844)*. Springer-Verlag, Berlin, 2004.

[19] S. Tabachnikov. *Geometry and Billiards*. American Mathematical Society, Providence, RI, 2005.

[20] M. P. Wojtkowski. Two applications of Jacobi fields to the billiard ball problem. *J. Differential Geom.* **40**(1) (1994), 155–164.