Shortest billiard trajectories [*]

Dániel Bezdek   Károly Bezdek [†]

October 23, 2008

Abstract

In this paper we prove that any convex body of the d-dimensional Euclidean space $(d \geq 2)$ possesses at least one shortest generalized billiard trajectory moreover, any of its shortest generalized billiard trajectories is of period at most $d + 1$. Actually, in the Euclidean plane we improve this theorem as follows. A disk-polygon with parameter $r > 0$ is simply the intersection of finitely many (closed) circular disks of radii r, called generating disks, having some interior point in common in the Euclidean plane. Also, we say that a disk-polygon with parameter $r > 0$ is a fat disk-polygon if the pairwise distances between the centers of its generating disks are at most r. We prove that any of the shortest generalized billiard trajectories of an arbitrary fat disk-polygon is a 2-periodic one. Also, we give a proof of the analogue result for ε-rounded disk-polygons obtained from fat disk-polygons by rounding them off using circular disks of radii $\varepsilon > 0$. Our theorems give partial answers to the very recent question raised by S. Zelditch on characterizing convex bodies whose shortest periodic billiard trajectories are of period 2.

1   Introduction

Mathematical billiards have been studied intensively for quite some time and have a fascinating history. (See for example the book [7] of Tabachnikov for

an elegant introduction to billiards.) The main motivation for writing this paper was to get an understanding of the geometry of the shortest billiard trajectories in convex bodies. The more exact details of the results proved in this paper are as follows.

For our purposes it seems natural to define billiard trajectories in the following way. This introduces a larger class of polygonal pathes for billiard trajectories than the traditional definition widely used in the literature (see [7]). Let $\mathbf{C}$ be an arbitrary convex body in the d-dimensional Euclidean space $\mathbb{E}^d, d \geq 2$ that is a compact convex set with non-empty interior in $\mathbb{E}^d$. Then we say that the closed polygonal path $\mathbf{P}$ (possibly with self-intersections) is a *generalized billiard trajectory* of $\mathbf{C}$ if all the vertices of $\mathbf{P}$ lie on the boundary of $\mathbf{C}$ and if each of the inner angle bisectors of $\mathbf{P}$ between two consecutive sides of $\mathbf{P}$ is perpendicular to a supporting hyperplane of $\mathbf{C}$ passing through the corresponding vertex of $\mathbf{P}$ (Figure 1). If $\mathbf{P}$ has n sides, then we say that $\mathbf{P}$ is an *n-periodic generalized billiard trajectory* in $\mathbf{C}$. Note that our definition of generalized billiard trajectories coincides with the traditional definition of billiard trajectories whenever the billiard table has no corner points. It seems that the paper [2] was among the first suggesting a detailed study of generalized billiard trajectories in convex domains. For some analogue higher dimensional investigations we refer the interested reader to the recent paper [6]. Generalized billiard trajectories have the following fundamental property.

Theorem 1.1 *Let $\mathbf{C}$ be a convex body in $\mathbb{E}^d, d \geq 2$. Then $\mathbf{C}$ possesses at least one shortest generalized billiard trajectory moreover, any of the shortest generalized billiard trajectories in $\mathbf{C}$ is of period at most $d + 1$.*

For the sake of completeness we mention that according to the main result of [1] any d-dimensional billiard table with a smooth boundary, but not necessarily convex, has a k-periodic billiard trajectory with $k \leq d+1$, which is a closely related result. Actually, in $\mathbb{E}^2$ we can improve Theorem 1.1 as follows. In order to state that theorem in a possibly short form let us introduce the following terminology. A *disk-polygon* is simply the intersection of finitely many congruent closed circular disks having some interior point in common in $\mathbb{E}^2$. Disk-polygons form a large class of convex domains (that is of compact convex sets with non-empty interior) in $\mathbb{E}^2$ including for example the well-known family of *Reuleaux polygons*. (See [4] for a substantial list of references and results on Reuleaux polygons and related topics.) Disk-polygons
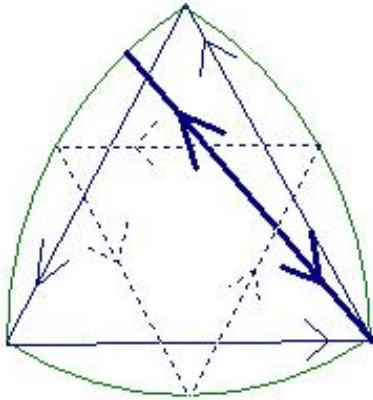
Figure 1: The Reuleaux triangle and some of its generalized billiard trajectories.

and their higher dimensional analogues which are called ball-polyhedra have many interesting geometric properties and so, it is somewhat surprising that the systematic study of this class of convex sets which is in a way a "curved analogue" of the class of convex polygons (resp., convex polytopes), has been proposed only very recently in [3]. In this paper we assume that whenever we take a disk-polygon, the disks generating it (simply called *generating disks*) are all needed that is each of them contributes to the boundary of the disk-polygon through a circular arc. Finally, let $\mathbf{D}$ be a disk-polygon in $\mathbb{E}^2$ having the property that the pairwise distances between the centers of its generating disks of radii r are at most r. In short, we say that $\mathbf{D}$ is a *fat disk-polygon* with parameter $r > 0$. (We note that for the purpose of this introduction it would be sufficient to assume that $r = 1$. However, as in the following proofs (fat) disk-polygons are showing up with different generating radii r, therefore in order to accomodate them properly, we had to insist on not to fix the value of the parameter r.)

Theorem 1.2 *Let* $\mathbf{D}$ *be a fat disk-polygon in* $\mathbb{E}^2$. *Then any of the shortest generalized billiard trajectories in* $\mathbf{D}$ *is a 2-periodic one.*

In the recent paper [6], the following fundamental question is studied that was raised by Zelditch in [8] motivated by applications to inverse spectral

3

problems: In which convex bodies are the shortest periodic billiard trajectories of period 2? It is proved in [6] that any convex body whose inscribed ball touches the boundary of the given convex body at two diametrically opposite points has that property. Theorem 1.2 shows that the family of fat disk-polygons possesses that property as well.

According to Birkhoff's well-known theorem ([7]) if $\mathbf{B}$ is a strictly convex billiard table with smooth boundary (that is if the boundary of $\mathbf{B}$ is a simple, closed, smooth and strictly convex curve) in $\mathbb{E}^2$, then for every positive integer $N > 1$ there exist (at least two) N-periodic billiard trajectories in $\mathbf{B}$. (In fact, here the rotation number of the billiard trajectory in question can be pre-assigned as well. Also, it is well-known that neither the convexity nor smoothness can be removed from the assumptions in order to have the same conclusion. Last but not least, for a higher dimensional analogue of the Birkhoff's theorem we refer the interested reader to [5].) Billiard tables suitable for Birkhoff's theorem can be easily constructed from disk-polygons as follows. Take a disk-polygon $\mathbf{D}$ in $\mathbb{E}^2$. Then choose a positive ε not larger than the inradius of $\mathbf{D}$ (which is the radius of the largest circular disk contained in $\mathbf{D}$) and take the union of all circular disks of radius ε that lie in $\mathbf{D}$. The set obtained in this way we call the ε-rounded disk-polygon of $\mathbf{D}$ and denote it by $\mathbf{D}(\varepsilon)$. The following theorem shows also the complexity of the problem of Zelditch [8] on characterizing convex domains whose shortest periodic billiard trajectories are of period 2.

Theorem 1.3 *Let $\mathbf{D}$ be a fat disk-polygon in $\mathbb{E}^2$. Then any of the shortest (generalized) billiard trajectories in the ε-rounded disk-polygon $\mathbf{D}(\varepsilon)$ is a 2-periodic one for all $\varepsilon > 0$ being sufficiently small.*

Actually, we believe that the following even stronger statement holds.

Conjecture 1.4 *Let $\mathbf{D}$ be a fat disk-polygon in $\mathbb{E}^2$. Then any of the shortest (generalized) billiard trajectories in the ε-rounded disk-polygon $\mathbf{D}(\varepsilon)$ is a 2-periodic one for all ε being at most as large as the inradius of $\mathbf{D}$.*

2   Proof of Theorem 1.1

We start with the following rather natural statement that can be proved easily with the help of Helly's theorem [4].

Lemma 2.1 *Let $\mathbf{F}$ be a finite set of at least $d+1$ points and $\mathbf{C}$ be a convex set in $\mathbb{E}^d, d \geq 2$. Then $\mathbf{C}$ has a translate that covers $\mathbf{F}$ if and only if every $d+1$ points of $\mathbf{F}$ can be covered by a translate of $\mathbf{C}$.*

Proof: For each point $p \in \mathbf{F}$ let $\mathbf{C}_p$ denote the set of all translation vectors in $\mathbb{E}^d$ with which one can translate $\mathbf{C}$ such that it contains p i.e. let $\mathbf{C}_p := \{t \in \mathbb{E}^d \mid p \in t + \mathbf{C}\}$. Now, it is easy to see that $\mathbf{C}_p$ is a convex set of $\mathbb{E}^d$ for all $p \in \mathbf{F}$ moreover, $\mathbf{F} \subset t + \mathbf{C}$ if and only if $t \in \cap_{p \in \mathbf{F}} \mathbf{C}_p$. Thus, Helly's theorem [4] applied to the convex sets $\{\mathbf{C}_p \mid p \in \mathbf{F}\}$ implies that $\mathbf{F} \subset t + \mathbf{C}$ if and only if $\mathbf{C}_{p_1} \cap \mathbf{C}_{p_2} \cap \cdots \cap \mathbf{C}_{p_{d+1}} \neq \emptyset$ holds for any $p_1, p_2, \ldots, p_{d+1} \in \mathbf{F}$, i.e. if and only if any $p_1, p_2, \ldots, p_{d+1} \in \mathbf{F}$ can be covered by a translate of $\mathbf{C}$, finishing the proof of Lemma 2.1. $\qquad\square$

Also the following statement is going to play a central role in our investigations. This is a generalization of the analogue 2-dimensional statement proved in [2].

Lemma 2.2 *Let $\mathbf{F} = \{f_1, f_2, \ldots, f_n\}$ be a finite set of points and $\mathbf{C}$ be a convex body in $\mathbb{E}^d, d \geq 2$. Then $\mathbf{F}$ cannot be translated into the interior of $\mathbf{C}$ if and only if the following two conditions hold: There are closed supporting halfspaces $H_{i_1}^+, H_{i_2}^+, \ldots, H_{i_s}^+$ of $\mathbf{C}$ assigned to some points of $\mathbf{F}$ say, to $f_{i_1}, f_{i_2}, \ldots, f_{i_s}$ with $1 \leq i_1 < i_2 < \cdots < i_s \leq n$ and a translation vector $t \in \mathbb{E}^d$ such that*

(i) the translated point $t + f_{i_j}$ belongs to the closed halfspace $H_{i_j}^-$ for all $1 \leq j \leq s$, where the interior of $H_{i_j}^-$ is disjoint from the interior of $H_{i_j}^+$ and its boundary hyperplane is identical to the boundary hyperplane of $H_{i_j}^+$ (which is in fact, a supporting hyperplane of $\mathbf{C}$);

(ii) the intersection $\cap_{j=1}^s H_{i_j}^+$ is nearly bounded meaning that it lies between two parallel hyperplanes of $\mathbb{E}^d$.

Proof: First, we assume that there are closed supporting halfspaces $H_{i_1}^+$, $H_{i_2}^+, \ldots, H_{i_s}^+$ of $\mathbf{C}$ assigned to some points of $\mathbf{F}$ say, to $f_{i_1}, f_{i_2}, \ldots, f_{i_s}$ with $1 \leq i_1 < i_2 < \cdots < i_s \leq n$ and a translation vector $t \in \mathbb{E}^d$ satisfying (i) as well as (ii). Based on this our goal is to show that $\mathbf{F}$ cannot be translated into the interior of $\mathbf{C}$ or equivalently that $\mathbf{F}$ cannot be covered by a translate of the interior $\mathrm{int}\mathbf{C}$ of $\mathbf{C}$. We prove this in an indirect way i.e. we assume that $\mathbf{F}$ can be covered by a translate of $\mathrm{int}\mathbf{C}$ and look for a contradiction. Indeed, if $\mathbf{F}$ can be covered by a translate of $\mathrm{int}\mathbf{C}$, then $t + \mathbf{F}$

can be covered by a translate of int$\mathbf{C}$ i.e. there is a translation vector $t^* \in \mathbb{E}^d$ such that $t + \mathbf{F} \subset t^* + \text{int}\mathbf{C}$. In particular, if $\mathbf{F}^* := \{f_{i_1}, f_{i_2}, \ldots, f_{i_s}\}$, then $t + \mathbf{F}^* \subset t^* + \text{int}\mathbf{C}$. Clearly, this implies that $\cap_{j=1}^s H_{i_j}^+ \subset \text{int}\left(\cap_{j=1}^s t^* + H_{i_j}^+\right)$, a contradiction to (ii).

Second, we assume that $\mathbf{F}$ cannot be translated into the interior of $\mathbf{C}$ and look for closed supporting halfspaces $H_{i_1}^+, H_{i_2}^+, \ldots, H_{i_s}^+$ of $\mathbf{C}$ assigned to some points of $\mathbf{F}$ say, to $f_{i_1}, f_{i_2}, \ldots, f_{i_s}$ with $1 \leq i_1 < i_2 < \cdots < i_s \leq n$ and a translation vector $t \in \mathbb{E}^d$ satisfying (i) as well as (ii). In order to simplify matters let us start to investigate the case when $\mathbf{C}$ is a smooth convex body in $\mathbb{E}^d$ i.e. through each boundary point of $\mathbf{C}$ there exists precisely one supporting hyperplane of $\mathbf{C}$. (Also, without loss of generality we can assume that the origin o of $\mathbb{E}^d$ is an interior point of $\mathbf{C}$.) As $\mathbf{F}$ cannot be translated into int$\mathbf{C}$ therefore Lemma 2.1 implies that there are $m \leq d+1$ points of $\mathbf{F}$ say, $\mathbf{F}_m := \{f_{j_1}, f_{j_2}, \ldots, f_{j_m}\}$ with $1 \leq j_1 < j_2 < \cdots < j_m \leq n$ such that $\mathbf{F}_m$ cannot be translated into int$\mathbf{C}$. Now, let $\lambda_0 := \inf\{\lambda > 0 \mid \lambda\mathbf{F}_m$ cannot be translated into int$\mathbf{C}\}$. Clearly, $\lambda_0 \leq 1$ and $\lambda_0\mathbf{F}_m$ cannot be translated into int$\mathbf{C}$ moreover, as $\lambda_0 = \sup\{\delta > 0 \mid \delta\mathbf{F}_m$ can be translated into $\mathbf{C}\}$, therefore there exists a translation vector $t \in \mathbb{E}^d$ such that $t + \lambda_0\mathbf{F}_m \subset \mathbf{C}$. Let $t + \lambda_0 f_{i_1}, t + \lambda_0 f_{i_2}, \ldots, t + \lambda_0 f_{i_s}$ with $1 \leq i_1 < i_2 < \cdots < i_s \leq n, 2 \leq s \leq m \leq d+1$ denote the points of $t + \lambda_0\mathbf{F}_m$ that are boundary points of $\mathbf{C}$ and let $H_{i_1}^+, H_{i_2}^+, \ldots, H_{i_s}^+$ be the corresponding closed supporting halfspaces of $\mathbf{C}$. We claim that $\mathbf{H}^+ := \cap_{k=1}^s H_{i_k}^+$ is nearly bounded. Indeed, if $\mathbf{H}^+$ were not nearly bounded, then there would be a translation vector $t' \in \mathbb{E}^d$ with $\mathbf{H}^+ \subset t' + \text{int}\mathbf{H}^+$. As $\mathbf{C}$ is a smooth convex body therefore this would imply the existence of a sufficiently small $\mu > 0$ with the property that $\{t + \lambda_0 f_{i_1}, t + \lambda_0 f_{i_2}, \ldots, t + \lambda_0 f_{i_s}\} \subset \mu t' + \text{int}\mathbf{C}$, a contradiction. Thus, as $o \in \text{int}\mathbf{C}$ therefore the points $f_{i_1}, f_{i_2}, \ldots, f_{i_s}$ and the closed supporting halfspaces $H_{i_1}^+, H_{i_2}^+, \ldots, H_{i_s}^+$ and the translation vector $t \in \mathbb{E}^d$ satisfy (i) as well as (ii). We are left with the case when $\mathbf{C}$ is not necessarily a smooth convex body in $\mathbb{E}^d$. In this case let $\mathbf{C}_N, N = 1, 2, \ldots$ be a sequence of smooth convex bodies lying in int$\mathbf{C}$ with $\lim_{N \to +\infty} \mathbf{C}_N = \mathbf{C}$. As $\mathbf{F}$ cannot be translated into the interior of $\mathbf{C}_N$ for all $N = 1, 2, \ldots$ therefore applying the method described above to each $\mathbf{C}_N$ and taking proper subsequences if necessary we end up with some points of $\mathbf{F}$ say, $f_{i_1}, f_{i_2}, \ldots, f_{i_s}$ with $1 \leq i_1 < i_2 < \cdots < i_s \leq n$ and with s convergent sequences of closed supporting halfspaces $H_{N,i_1}^+, H_{N,i_2}^+, \ldots, H_{N,i_s}^+$ of $\mathbf{C}_N$ and a convergent sequence of translation vectors t_N that satisfy (i) and (ii) for each N. By taking the limits $H_{i_1}^+ := \lim_{N \to +\infty} H_{N,i_1}^+, H_{i_2}^+ := \lim_{N \to +\infty} H_{N,i_2}^+, \ldots, H_{i_s}^+ := \lim_{N \to +\infty} H_{N,i_s}^+$

and $t := \lim_{N \to +\infty} t_N$ we get the desired nearly bounded family of closed supporting halfspaces of $\mathbf{C}$ and the translation vector $t \in \mathbb{E}^d$ satisfying (i) as well as (ii). This completes the proof of Lemma 2.2. $\qquad\square$

Lemma 2.3 *Let $\mathbf{C}$ be a convex body in $\mathbb{E}^d, d \geq 2$. If $\mathbf{P}$ is a generalized billiard trajectory in $\mathbf{C}$, then $\mathbf{P}$ cannot be translated into the interior of $\mathbf{C}$.*

Proof: Let $p_1, p_2, \ldots, p_n$ be the vertices of $\mathbf{P}$ and let $v_1, v_2, \ldots, v_n$ be the points of the unit sphere $\mathbb{S}^{d-1}$ centered at the origin o in $\mathbb{E}^d$ whose position vectors are parallel to the inner angle bisectors (halflines) of $\mathbf{P}$ at the vertices $p_1, p_2, \ldots, p_n$ of $\mathbf{P}$. Moreover, let $H_1^+, H_2^+, \ldots, H_n^+$ denote the closed supporting halfspaces of $\mathbf{C}$ whose boundary hyperplanes are perpendicular to the inner angle bisectors of $\mathbf{P}$ at the vertices $p_1, p_2, \ldots, p_n$. Based on Lemma 2.2 in order to prove that $\mathbf{P}$ cannot be translated into the interior of $\mathbf{C}$ it is sufficient to show that $\cap_{i=1}^n H_i^+$ is nearly bounded or equivalently that $o \in \mathrm{conv}(\{v_1, v_2, \ldots, v_n\})$, where $\mathrm{conv}(.)$ denotes the convex hull of the corresponding set in $\mathbb{E}^d$. It is easy to see that $o \in \mathrm{conv}(\{v_1, v_2, \ldots, v_n\})$ if and only if for any hyperplane H of $\mathbb{E}^d$ passing through o and for any of the two closed halfspaces bounded by H say, for H^+, we have that $H^+ \cap \mathrm{conv}(\{v_1, v_2, \ldots, v_n\}) \neq \emptyset$. Indeed, for a given H^+ let $t \in \mathbb{E}^d$ be chosen so that $t + H^+$ is a supporting halfspace of $\mathrm{conv}(\{p_1, p_2, \ldots, p_n\})$. Clearly, at least one vertex of $\mathbf{P}$ say, p_{i_0} must belong to the boundary of $t + H^+$ and therefore $v_{i_0} \in H^+ \cap \mathrm{conv}(\{v_1, v_2, \ldots, v_n\})$, finishing the proof of Lemma 2.3. $\qquad\square$

For the purpose of the following statement it seems natural to introduce *generalized $(d+1)$-gons* in $\mathbb{E}^d$ as closed polygonal path (possibly with self-intersections) having at most $d + 1$ sides.

Lemma 2.4 *Let $\mathbf{C}$ be a convex body in $\mathbb{E}^d, d \geq 2$ and let $\mathcal{F}_{d+1}(\mathbf{C})$ denote the family of all generalized $(d+1)$-gons of $\mathbb{E}^d$ that cannot be translated into the interior of $\mathbf{C}$. Then $\mathcal{F}_{d+1}(\mathbf{C})$ possesses a minimal length member moreover, the shortest perimeter members of $\mathcal{F}_{d+1}(\mathbf{C})$ are identical (up to translations) with the shortest generalized billiard trajectories of $\mathbf{C}$.*

Proof: If $\mathbf{P}$ is an arbitrary generalized billiard trajectory of the convex body $\mathbf{C}$ in $\mathbb{E}^d$ with vertices $p_1, p_2, \ldots, p_n$, then according to Lemma 2.3 $\mathbf{P}$ cannot be translated into the interior of $\mathbf{C}$. Thus, by Lemma 2.1 $\mathbf{P}$ possesses at most $d + 1$ vertices say, $p_{i_1}, p_{i_2}, \ldots, p_{i_{d+1}}$ with $1 \leq i_1 \leq i_2 \leq \cdots \leq i_{d+1} \leq$

7

n such that $p_{i_1}, p_{i_2}, \ldots, p_{i_{d+1}}$ cannot be translated into the interior of $\mathbf{D}$. This implies that by connecting the consecutive points of $p_{i_1}, p_{i_2}, \ldots, p_{i_{d+1}}$ by line segments according to their cyclic ordering the generalized $(d+1)$-gon $\mathbf{P}_{d+1}$ obtained, has length $l(\mathbf{P}_{d+1})$ at most as large as the length $l(\mathbf{P})$ of $\mathbf{P}$ moreover, $\mathbf{P}_{d+1}$ cannot be covered by a translate of int$\mathbf{C}$ i.e. $\mathbf{P}_{d+1} \in \mathcal{F}_{d+1}(\mathbf{C})$. Now, by looking at only those members of $\mathcal{F}_{d+1}(\mathbf{C})$ that lie in a d-dimensional ball of sufficiently large radius in $\mathbb{E}^d$ we get via a standard compactness argument and Lemma 2.2 that $\mathcal{F}_{d+1}(\mathbf{C})$ possesses a member of minimal length say, $\Delta_{d+1}(\mathbf{C})$. As the inequalities $l(\Delta_{d+1}(\mathbf{C})) \leq l(\mathbf{P}_{d+1}) \leq l(\mathbf{P})$ hold for any generalized billiard trajectory $\mathbf{P}$ of $\mathbf{C}$ therefore in order to finish our proof it is sufficient to show that $\Delta_{d+1}(\mathbf{C})$ is a generalized billiard trajectory of $\mathbf{C}$. Indeed, as $\Delta_{d+1}(\mathbf{C}) \in \mathcal{F}_{d+1}(\mathbf{C})$ therefore $\Delta_{d+1}(\mathbf{C})$ cannot be translated into int$\mathbf{C}$. Thus, the minimality of $\Delta_{d+1}(\mathbf{C})$ and Lemma 2.2 imply that if $q_1, q_2, \ldots, q_m$ denote the vertices of $\Delta_{d+1}(\mathbf{C})$ with $m \leq d+1$, then there are closed supporting halfspaces $H_1^+, H_2^+, \ldots, H_m^+$ of $\mathbf{C}$ whose boundary hyperplanes $H_1, H_2, \ldots, H_m$ pass through the points $q_1, q_2, \ldots, q_m$ (each being a boundary point of $\mathbf{C}$) and have the property that $\cap_{i=1}^m H_i^+$ is nearly bounded in $\mathbb{E}^d$. If the inner angle bisector at a vertex of $\Delta_{d+1}(\mathbf{C})$ say, at q_i were not perpendicular to H_i, then it is easy to see via Lemma 2.2 that one could slightly move q_i along H_i to a new position q_i' (typically an exterior point of $\mathbf{C}$ on H_i) such that the new generalized $(d+1)$-gon $\Delta_{d+1}'(\mathbf{C}) \in \mathcal{F}_{d+1}(\mathbf{C})$ would have a shorter length, a contradiction. This completes the proof of Lemma 2.4. □

Finally, notice that Theorem 1.1 follows from Lemma 2.4 in a straightforward way.

3 Proof of Theorem 1.2

Let $\mathbf{D}$ be a fat disk-polygon in $\mathbb{E}^2$ and assume that the conclusion of Theorem 1.2 does not hold for $\mathbf{D}$. Based on Theorem 1.1 this means that there is a shortest generalized billiard trajectory $\mathbf{P}$ in $\mathbf{D}$ of period 3. Thus, $\mathbf{P}$ is a generalized billiard triangle of $\mathbf{D}$ whose perimeter is the smallest among the perimeters of all generalized billiard trajectories of $\mathbf{D}$. Our goal is to show that this assumption leads to a contradiction, namely we will prove that $\mathrm{per}(\mathbf{P}) > 2 \cdot \mathrm{width}(\mathbf{D})$, where $\mathrm{width}(.,.)$ stands for the *minimal width* of the corresponding convex set denoting the smallest distance between two

parrallel supporting lines of the given convex set.(This is indeed a contradiction because the double covered line segment represented by a chord of $\mathbf{D}$ of length equal to the minimal width of $\mathbf{D}$ is in fact, a generalized billiard trajectory of period 2 in $\mathbf{D}$.)

Sublemma 3.1 $\mathbf{P}$ *is a billiard triangle in* $\mathbf{D}$ *meaning that it is a generalized billiard triangle of* $\mathbf{D}$ *with the additional property that each vertex of* $\mathbf{P}$ *is a smooth boundary point of* $\mathbf{D}$.

Proof: Lemma 2.4 and Lemma 2.2 imply the existence of three closed supporting halfplanes of $\mathbf{D}$ whose boundary lines pass through the corresponding vertices of $\mathbf{P}$ and whose intersection is a triangle say, $\boldsymbol{\Delta}_1^+$. Now, imagine that a vertex say, p of $\mathbf{P}$ is not a smooth boundary point of $\mathbf{D}$. This means that we can slightly perturb the boundary line L_1 of the closed supporting halfplane of $\mathbf{D}$ assigned to p such that we create a new closed supporting halfplane for $\mathbf{D}$ with boundary line L_2 and obtain a new triangle $\boldsymbol{\Delta}_2^+$ into which $\mathbf{P}$ is inscribed. Clearly, we can arrange all this so that the inner angle bisector of $\mathbf{P}$ at p is not perpendicular to L_2. As a result we can slightly move p along L_2 to a new position p', such that the new triangle $\mathbf{P}'$ obtained from $\mathbf{P}$ by replacing p with p' possesses a strictly smaller perimeter than $\mathbf{P}$ moreover, using Lemma 2.2 it is easy to see that $\mathbf{P}'$ cannot be translated into the interior of $\mathbf{D}$, and therefore via Lemma 2.4 it leads to a contradiction. This finishes the proof of Sublemma 3.1. $\qquad\square$

Let the vertices of $\mathbf{P}$ be denoted by p_1, p_2 and p_3. Based on Sublemma 3.1 let $\mathbf{D}_i$ be the uniquely determined generating disk of $\mathbf{D}$ with center c_i and radius r whose boundary circle S_i passes through p_i , $1 \leq i \leq 3$. Clearly $\mathbf{P}$ is a billiard triangle of the disk-triangle disk$\boldsymbol{\Delta}_r(\mathbf{P}) = \cap_{i=1}^3 \mathbf{D}_i$. Hence, if c denotes the center of the incircle of $\mathbf{P}$, then the points p_i, c and c_i are all lying on a line for each $1 \leq i \leq 3$. Also, as $\mathbf{D} \subset \text{disk}\boldsymbol{\Delta}_r(\mathbf{P})$ therefore

$$(1) \qquad 2 \cdot \text{width}(\text{disk}\boldsymbol{\Delta}_r(\mathbf{P})) \geq 2 \cdot \text{width}(\mathbf{D}).$$

Moreover, as $c_1, c_2, c_3 \in \mathbf{D}$ therefore

$$(2) \qquad c_1, c_2, c_3 \in \text{disk}\boldsymbol{\Delta}_r(\mathbf{P}).$$

As a next step we continuously increase the radius r to a maximum value r' by moving each c_i along the halfline starting at p_i and passing through c,
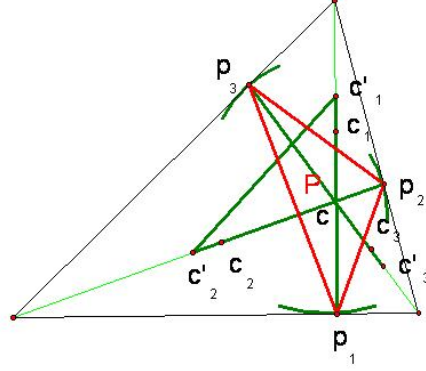
Figure 2: Constructing $\text{disk}\mathbf{\Delta}_{r'}(\mathbf{P})$ from $\text{disk}\mathbf{\Delta}_{r}(\mathbf{P})$.

to a new position c_i' with $\text{dist}\,(p_i, c_i') = r'$, $1 \le i \le 3$ and having the property that

$$(2') \qquad c_1', c_2', c_3' \in \text{disk}\mathbf{\Delta}_{r'}(\mathbf{P}),$$

where $\text{dist}\,(.,.)$ stands for Euclidean distance between points and $\text{disk}\mathbf{\Delta}_{r'}(\mathbf{P})$ denotes the intersection of the closed circular disks $\mathbf{D}_i'$ of radius r' centered at the points c_i', $1 \le i \le 3$ (Figure 2). It is clear that due to the maximality of r' with respect to $(2')$ the diameter of the set $\{c_1', c_2', c_3'\}$ has to be equal to r' with say, $\text{dist}\,(c_1', c_2') = r'$. This means that if S_i' denotes the boundary circle of $\mathbf{D}_i'$, $1 \le i \le 3$, then

$$(3) \qquad c_1' \in S_2' \cap \text{bd}(\text{disk}\mathbf{\Delta}_{r'}(\mathbf{P})), c_2' \in S_1' \cap \text{bd}(\text{disk}\mathbf{\Delta}_{r'}(\mathbf{P})), c_3' \in \text{disk}\mathbf{\Delta}_{r'}(\mathbf{P}),$$

where $\text{bd}(.)$ stands for the boundary of the corresponding set. Also, as $\text{disk}\mathbf{\Delta}_{r}(\mathbf{P}) \subset \text{disk}\mathbf{\Delta}_{r'}(\mathbf{P})$ therefore

$$(4) \qquad 2r' = 2 \cdot \text{width}(\text{disk}\mathbf{\Delta}_{r'}(\mathbf{P})) \ge 2 \cdot \text{width}(\text{disk}\mathbf{\Delta}_{r}(\mathbf{P})).$$

Last but not least $\mathbf{P}$ is a billiard triangle in the disk-triangle $\text{disk}\mathbf{\Delta}_{r'}(\mathbf{P})$ as well with

$$(5) \qquad p_i \in S_i', 1 \le i \le 3.$$

Now, let H^* denote the closed halfplane bounded by the line $c_1' c_2'$ that contains c_3' and let c_3^* be the corner point of the 2-sided disk-polygon $\mathbf{D}_1' \cap \mathbf{D}_2'$ that lies in H^* (Figure 3).
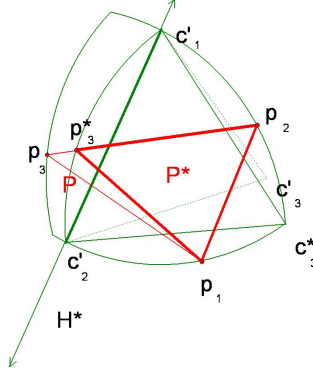
Figure 3: Constructing $\mathbf{P}^*$ from $\mathbf{P}$.

If $\mathbf{D}_3^*$ denotes the closed circular disk of radius r' centered at c_3^*, then $\mathbf{Re} = \mathbf{D}_1' \cap \mathbf{D}_2' \cap \mathbf{D}_3^*$ is a Reuleaux triangle of constant width r' with

$$(6) \qquad \mathbf{Re} \subset \mathrm{disk}\mathbf{\Delta}_{r'}(\mathbf{P}).$$

Now, (5) via (3) holds in the following stronger form

$$(5') \qquad p_1 \in S_1' \cap \mathrm{bd}(\mathbf{Re}), p_2 \in S_2' \cap \mathrm{bd}(\mathbf{Re}).$$

If S_3^* denotes the boundary circle of $\mathbf{D}_3^*$, then $(5), (5')$ and (6) imply the existence of a point p_3^* with

$$(7) \qquad p_3^* \in \mathbf{P} \cap S_3^* \cap \mathrm{bd}(\mathbf{Re}).$$

If $\mathbf{P}^*$ denotes the triangle determined by the points p_1, p_2 and p_3^*, then Lemma 2.2, $(5')$ and (7) imply in a straighforward way that $\mathbf{P}^*$ cannot be translated into the interior of $\mathbf{Re}$. Moreover, as $\mathbf{P}^* \subset \mathbf{P}$ therefore

$$(8) \qquad \mathrm{per}(\mathbf{P}) \geq \mathrm{per}(\mathbf{P}^*).$$

Sublemma 3.2 $\mathrm{per}(\mathbf{P}^*) > 2r'$.

Proof: Recall the statement proved in [2] that the shortest generalized billiard trajectories of $\mathbf{Re}$ among all generalized billiard trajectories of period at most 3 are the double covered chords of $\mathbf{Re}$ having length equal to $2r'$, where r' stands for the constant width of $\mathbf{Re}$. Thus, Lemma 2.4 implies in a straightforward way that indeed $\mathrm{per}(\mathbf{P}^*) > 2r'$. □

Hence, (1), (4), (8) and Sublemma 3.2 imply that $\mathrm{per}(\mathbf{P}) > 2 \cdot \mathrm{width}(\mathbf{D})$, a contradiction. This finishes the proof of Theorem 1.2.

4 Proof of Theorem 1.3

Assume that $\mathbf{D}$ is a fat disk-polygon moreover, assume that the claim of Theorem 1.3 does not hold for $\mathbf{D}$. In what follows, using Theorem 1.2, we show that these assumptions lead to a contradiction therefore providing a proof for Theorem 1.3.

The above assumptions imply that there exists a decreasing sequence $\epsilon_n > 0$ with $\lim_{n\to\infty} \epsilon_n = 0$ such that for each ϵ_n, the ϵ_n-rounded disk-polygon $\mathbf{D}(\epsilon_n)$ has a shortest (generalized) billiard trajectory $\mathbf{P}(\epsilon_n)$ of period at least 3. Now, Lemma 2.4 implies that $\mathbf{P}(\epsilon_n)$ is in fact, a (generalized) billiard triangle in $\mathbf{D}(\epsilon_n)$ for all n. According to the well-known Blaschke selection theorem (applied to the container $\mathbf{D}$ with $\mathbf{D}(\epsilon_n) \subset \mathbf{D}$ and to the sequence $\mathbf{P}(\epsilon_n)$) one can find a subsequence of $\mathbf{P}(\epsilon_n)$ whose (generalized) billiard triangles converge to a triangle or a double covered line segment labeled by $\mathbf{P}$. For the sake of simplicity we assume that the subsequence in question is identical to the sequence $\mathbf{P}(\epsilon_n)$ and so, we assume that the sequence $\mathbf{P}(\epsilon_n)$ converges to $\mathbf{P}$. It is easy to see that $\mathbf{P}$ is a generalized billiard trajectory of $\mathbf{D}$ which is either a triangle or a double covered line segment. Thus, we have the following two cases:

Case (A): $\mathbf{P}$ is a generalized billiard triangle of $\mathbf{D}$;

Case (B): $\mathbf{P}$ is a double covered line segment of $\mathbf{D}$ with the property that the two lines passing through the end points of $\mathbf{P}$ and being perpendicular to $\mathbf{P}$ are supporting lines of $\mathbf{D}$.

Case (A): On the one hand, clearly $\mathrm{per}(\mathbf{P}(\epsilon_n)) \leq 2 \cdot \mathrm{width}(\mathbf{D}(\epsilon_n))$ for all n and therefore

$$(9) \qquad \mathrm{per}(\mathbf{P}) \leq 2 \cdot \mathrm{width}(\mathbf{D}).$$

On the other hand, Theorem 1.2 and (9) imply that $\mathbf{P}$ must be one of the shortest generalized billiard trajectories in $\mathbf{D}$ of period 2, a contradiction. This finishes the proof of Theorem 1.3 in Case (A).

Case (B): Actually, our goal is to show that this case cannot occur. Let the vertices of $\mathbf{P}(\epsilon_n)$ be denoted by a_n, b_n and c_n with $a_n, b_n, c_n \in \mathrm{bd}(\mathbf{D}(\epsilon_n))$. As $\mathbf{P}(\epsilon_n)$ converges to the double covered line segment $\mathbf{P}$ represented by a chord of $\mathbf{D}$ therefore one can assume that for all sufficiently large n two vertices of $\mathbf{P}(\epsilon_n)$ say, a_n and b_n belong to $\mathrm{bd}(\mathbf{D}(\epsilon_n))$ as well as to the boundary of a closed circular disk of radius ϵ_n centered at some interior point of $\mathbf{D}$ say, o_n. Based on this, as $\mathbf{P}(\epsilon_n)$ is a (generalized) billiard triangle in $\mathbf{D}(\epsilon_n)$ therefore $\mathbf{P}(\epsilon_n)$ must be an isosceles triangle say, with $\mathrm{dist}(a_n, c_n) = \mathrm{dist}(b_n, c_n)$

and with o_n being the intersection of the inner angle bisectors of $\mathbf{P}(\epsilon_n)$ for all sufficiently large n. If $2\alpha_n$ denotes the angle of the triangle $\mathbf{P}(\epsilon_n)$ at the vertices a_n and b_n, then an easy computation shows that

$$(10) \quad \text{dist}(a_n, b_n) = 2\epsilon_n \cos \alpha_n \text{ and } \text{dist}(a_n, c_n) = \text{dist}(b_n, c_n) = \frac{\epsilon_n \cos \alpha_n}{\cos 2\alpha_n}.$$

Also, it is easy to see that

$$(11) \quad \text{dist}(c_n, o_n) = \epsilon_n \sqrt{1 + \frac{\cos^2 \alpha_n}{\cos^2 2\alpha_n} - 2\frac{\cos^2 \alpha_n}{\cos 2\alpha_n}}.$$

Furthermore, if d_n denotes the midpoint of the circular arc of radius ϵ_n connecting a_n to b_n and belonging to $\text{bd}(\mathbf{D}(\epsilon_n))$, then the lines passing through c_n and d_n and being perpendicular to the line $c_n d_n$ are supporting lines for $\mathbf{D}(\epsilon_n)$. Thus, the double covered line segment with end points c_n and d_n represents a (generalized) billiard trajectory of period 2 in $\mathbf{D}(\epsilon_n)$ and so,

$$(12) \quad \text{dist}(c_n, d_n) \geq \frac{1}{2}\text{per}(\mathbf{P}(\epsilon_n)).$$

Using (10) and (11) the inequality (12) is equivalent to the following inequality that should hold for all sufficiently large n:

$$(13) \quad 0 \geq \left(\cos \alpha_n + \frac{\cos \alpha_n}{\cos 2\alpha_n} \right) - \left(1 + \sqrt{1 + \frac{\cos^2 \alpha_n}{\cos^2 2\alpha_n} - 2\frac{\cos^2 \alpha_n}{\cos 2\alpha_n}} \right).$$

Clearly, as n tends to infinity, ϵ_n approaches 0, meanwhile the line segment with end points c_n and o_n converges to the chord $\mathbf{P}$ of $\mathbf{D}$. This forces α_n to converge to $\frac{\pi}{4}$. Using Maple one can conveniently check that the function $g(x) = \left(\cos x + \frac{\cos x}{\cos 2x} \right) - \left(1 + \sqrt{1 + \frac{\cos^2 x}{\cos^2 2x} - 2\frac{\cos^2 x}{\cos 2x}} \right)$ defined over the open interval $(0.7, \frac{\pi}{4})$ is a decreasing positive function. In fact, it turns out that $g(x) > 0.4$ for all $0.7 < x < \frac{\pi}{4}$. (Of course, this can be worked out "by hand" as well, using the proper analytic tools from calculus and thus, avoiding completely the use of Maple.) This contradicts (13) finishing the proof of Theorem 1.3 in Case (B).

References

[1] V. Benci and F. Giannoni, *Periodic bounce trajectories with a low number of bounce points*, Ann. Inst. H. Poincar Anal. Non Linaire 6(1) (1989), 73-93.

[2] K. Bezdek and R. Connelly, *Covering curves by translates of a convex set*, Amer. Math. Monthly 96(9) (1989), 789-806.

[3] K. Bezdek, L. Lángi, M. Naszódi and P. Papez, *Ball-polyhedra*, Discrete Comput. Geom. 38 (2007), 201-230.

[4] T. Bonnesen and W. Fenchel, *Theorie der konvexen Körper*, Springer Verlag, Berlin, 1974.

[5] M. Farber and S. Tabachnikov, *Topology of cyclic configuration spaces and periodic trajectories of multi-dimensional billiards*, Topology 41(3) (2002), 553-589.

[6] M. Ghomi, *Shortest periodic billiard trajectories in convex bodies*, Geom. Funct. Anal. 14 (2004), 295-302.

[7] S. Tabachnikov, *Geometry and billiards*, Amer. Math. Soc., 2005.

[8] S. Zelditch, *Spectral determination of analytic bi-axisymmetric plane domains*, Geom. Funct. Anal. 10(3) (2000), 628-677.

Dániel Bezdek, Haskayne School of Business, 2500 University drive N.W., University of Calgary, Calgary, Alberta, Canada, T2N 1N4.
e-mail: daniel.bezdek@gmail.com

Károly Bezdek, Department of Mathematics and Statistics, 2500 University drive N.W., University of Calgary, Calgary, Alberta, Canada, T2N 1N4.
e-mail: bezdek@math.ucalgary.ca